

Search for the Associated Production of Higgs boson with top anti-top pairs

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2013

James Paul Agnew

School of Physics and Astronomy

Contents

Abstract	17
Declaration	18
Copyright	19
Acknowledgements	20
1 Introduction	21
2 Theory	22
2.1 Standard Model	22
2.1.1 Forces in the Standard Model	22
2.1.2 Leptons and Quarks	27
2.1.3 The Top Quark	29
3 The Tevatron and the D0 Detector	33
3.1 The Tevatron	33
3.2 The D0 Detector	35
3.2.1 Tracking	36
3.2.2 Calorimeters	38
3.2.3 Muon Systems	40
3.2.4 Luminosity Detectors	41

3.2.5	Triggering	43
4	Identification of Particles in the Detector	46
4.1	Tracks in the detector and vertex identification	47
4.1.1	Tracks	47
4.1.2	Vertices	48
4.2	Electron and Muon Identification	49
4.2.1	Electrons	49
4.2.2	Muons	52
4.3	Missing Transverse Energy	54
4.4	Jet Reconstruction	54
4.4.1	Jets	54
4.4.2	b-jet Identification	55
4.5	Secondary Vertex Mass	56
4.6	Jet Lifetime Probability Tagger	57
4.7	The Multivariate Tagging Method	58
5	Data and Simulated Samples	60
5.1	Data	60
5.2	Background Processes	61
5.3	Signal Processes	63
6	Data-driven Background Determination	65
6.1	V+jets	65
6.2	Multijet	66
7	Event Selection	69
7.1	Validation Plots	70
7.2	B-tagging Selection	79

8	Search for $t\bar{t}H$ production	82
8.0.1	Separation of signal from background	82
8.0.2	Event yields and Limit setting	84
8.0.3	Systematic uncertainties	86
8.0.4	Results	86
9	Search Using a Multivariate Approach	100
9.1	Multivariate Analysis Method	100
9.2	Discriminating Variables	101
10	Summary	113
	Appendices	115
A	Validation Plots	115
B	Limit Plots for Individual Lepton Channels	141
C	MVA Output Distributions	143
	Bibliography	154

List of Tables

1	EMID operating point values for real electrons and fake electrons in the CC and EC	53
2	Breakdown of integrated luminosity of RunII data after quality cuts	61
3	Cross section times branching ratios for $t\bar{t}H$ with $H \rightarrow b\bar{b}$ in the range $90 < m_H < 155$ GeV	63
4	b -tagging category names and operating points. Note that for the 2 tag categories the mean BL for the two jets is taken.	81
5	Summary of expected and observed yields in the various channels from the 4 jet 1 tight b -tag bin (4j1tt) to the 4 jet ≥ 3 b -tag bin (4j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.	84
6	Summary of expected and observed yields in the various channels from the 5 jet 1 tight b -tag bin (5j1tt) to the 5 jet ≥ 3 b -tag bin (5j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.	96

7	Summary of expected and observed yields in the various channels from the ≥ 6 jet 1 tight b -tag bin (6j1tt) to the ≥ 6 jet ≥ 3 b -tag bin (6j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.	97
8	Summary of H_T -independent systematic uncertainties used as input for the limit derivation.	98
9	Expected and observed ratios of excluded $t\bar{t}H$ cross section times $H \rightarrow b\bar{b}$ branching fraction over SM expectation for different values of the Higgs mass.	99
10	Expected and observed ratios of excluded $t\bar{t}H$ cross section times $H \rightarrow b\bar{b}$ branching fraction over SM expectation given to 3 s. f. for different values of the Higgs mass.	112

List of Figures

2.1	Branching ratios of the Higgs for a range of Higgs masses [8].	27
2.2	The Higgs decay to two photons through (a), (b) a W boson pair and (c) a top quark pair [9].	28
2.3	Branching fractions of the possible decays for two W bosons.	31
2.4	Feynman diagrams of the production mechanisms for $t\bar{t}$ production at the Tevatron [18].	32
3.1	The Tevatron accelerator chain [19]	34
3.2	A cutaway view of the D0 detector [20]	35
3.3	A side view of the D0 tracking system [20].	37
3.4	A 3/4 view of the D0 SMT [20]	38
3.5	A cutaway view of the D0 calorimeter system [20].	40
3.6	The Muon System [22].	41
3.7	Location of the luminosity monitors on the z -axis [22].	42
3.8	The D0 trigger and DAQ system [22].	43
4.1	BDT output for (a) CC electrons in low luminosity events, (b) CC electrons in high luminosity events, (c) EC electrons in low luminosity events and (d) EC electrons in high luminosity events. In each figure, the solid red distribution is from real electrons and blue from fake electrons, with the barred points from a testing sample [31].	52
4.2	A jet showering in the detector [35]	55

4.3	MVA output [38]	56
4.4	The M_{svt} distributions for b -, c - and light-jets [38]	57
4.5	The JLIP probability [38]	58
5.1	The total integrated luminosity delivered by the Tevatron accelerator (green) and recorded by the D0 detector (blue) before quality cuts. . .	61
6.1	Binned distribution of m_T^W for determination of scale factors for multijet (green) and V +jets (yellow) when fitted to background subtracted data for events with 4 jets and 1 tight electron	66
6.2	Scale factors derived from the fitting process for events with at least 4 jets and 1 tight electron. The scale factor derived for 5 jets is also used in the at least 6 jet case.	67
7.1	Legend for validation plots.	72
7.2	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 4 jets and 1 electron.	73
7.3	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 5 jets and 1 electron.	74
7.4	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with at least 6 jets and 1 electron.	75

7.5	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 4 jets and 1 muon.	76
7.6	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 5 jets and 1 muon.	77
7.7	Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with at least 6 jets and 1 muon.	78
7.8	Average b -ID discriminant output of two tagged jets. This plot is taken from a search for associated production of Higgs with a vector boson [54], and as such the signal shown in this plot reflects that search rather than the one documented in this thesis.	80
8.1	Normalized distribution of H_T (applying a $p_t^{jet} > 15$ GeV cut) for $t\bar{t}H$ where the Higgs mass was set to 115 GeV (red), $t\bar{t}$ (blue), W +jets (green) and multijet (brown) production. H_T is defined as the scalar sum of the transverse momenta of the 4 leading jets.	88
8.2	Normalized distribution of the number of good jets (a) and the number of good jets which have a NN loose b -tag (b) for $t\bar{t}H$ where the Higgs mass was set to 115 GeV (red), $t\bar{t}$ (blue), W +jets (green) and multijet (brown) production.	88
8.3	Legend for input plots, with signal scaling factor 1000.	89

8.4	<i>H_T</i> distributions corresponding to the <i>e</i> +jets data set of 9.7 fb ⁻¹ requiring 1 loose <i>b</i> -tag (upper left), 2 loose <i>b</i> -tags (upper right), 2 medium <i>b</i> -tags (middle left), 2 tight <i>b</i> -tags (middle right) and ≥ 3 <i>b</i> -tags (lower) for events with 4 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	90
8.5	<i>H_T</i> distributions corresponding to the <i>e</i> +jets data set of 9.7 fb ⁻¹ requiring 1 loose <i>b</i> -tag (upper left), 2 loose <i>b</i> -tags (upper right), 2 medium <i>b</i> -tags (middle left), 2 tight <i>b</i> -tags (middle right) and ≥ 3 <i>b</i> -tags (lower) for events with 5 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	91
8.6	<i>H_T</i> distributions corresponding to the <i>e</i> +jets data set of 9.7 fb ⁻¹ requiring 1 loose <i>b</i> -tag (upper left), 2 loose <i>b</i> -tags (upper right), 2 medium <i>b</i> -tags (middle left), 2 tight <i>b</i> -tags (middle right) and ≥ 3 <i>b</i> -tags (lower) for events with ≥ 6 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	92
8.7	<i>H_T</i> distributions corresponding to the <i>μ</i> +jets data set of 9.7 fb ⁻¹ requiring 1 loose <i>b</i> -tag (upper left), 2 loose <i>b</i> -tags (upper right), 2 medium <i>b</i> -tags (middle left), 2 tight <i>b</i> -tags (middle right) and ≥ 3 <i>b</i> -tags (lower) for events with 4 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	93
8.8	<i>H_T</i> distributions corresponding to the <i>μ</i> +jets data set of 9.7 fb ⁻¹ requiring 1 loose <i>b</i> -tag (upper left), 2 loose <i>b</i> -tags (upper right), 2 medium <i>b</i> -tags (middle left), 2 tight <i>b</i> -tags (middle right) and ≥ 3 <i>b</i> -tags (lower) for events with 5 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	94

8.9	H_T distributions corresponding to the μ +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with ≥ 6 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).	95
8.10	The 95% CL upper limit on the $\sigma_{\bar{t}tH}$ cross section times branching ratio over the SM expectation in NLO QCD as a function of the Higgs mass. The 1 and 2 σ uncertainty bands for the expected limit are indicated by the green and green+yellow bands.	98
9.1	BDT distributions corresponding to the $e + 4$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	105
9.2	BDT distributions corresponding to the $e + 5$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	106
9.3	BDT distributions corresponding to the $e + \geq 6$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	107
9.4	BDT distributions corresponding to the $\mu + 4$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	108
9.5	BDT distributions corresponding to the $\mu + 5$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	109
9.6	BDT distributions corresponding to the $\mu + \geq 6$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).	110

9.7	Limits obtained for the process $\sigma(t\bar{t}H) \times BR(H \rightarrow b\bar{b})$ at D0 in the range of Higgs masses $90 < m_H < 155$ GeV	111
A.1	Legend for validation plots.	116
A.2	Validation plots for the jet with highest transverse momentum p_T for events with 1 electron, exactly 4 jets and A.2(a) 1 tight tag, A.2(b) 2 loose tags, A.2(c) 2 medium tags, A.2(d) 2 tight tags and A.2(e) 3 tags.	117
A.3	Validation plots for the jet with second highest transverse momentum p_T for events with 1 electron, exactly 4 jets and A.3(a) 1 tight tag, A.3(b) 2 loose tags, A.3(c) 2 medium tags, A.3(d) 2 tight tags and A.3(e) 3 tags.	118
A.4	Validation plots for the transverse momentum p_T of the lepton for events with 1 electron, exactly 4 jets and A.4(a) 1 tight tag, A.4(b) 2 loose tags, A.4(c) 2 medium tags, A.4(d) 2 tight tags and A.4(e) 3 tags.	119
A.5	Validation plots for the transverse momentum p_T of the neutrino for events with 1 electron, exactly 4 jets and A.5(a) 1 tight tag, A.5(b) 2 loose tags, A.5(c) 2 medium tags, A.5(d) 2 tight tags and A.5(e) 3 tags.	120
A.6	Validation plots for the jet with highest transverse momentum p_T for events with 1 electron, exactly 5 jets and A.6(a) 1 tight tag, A.6(b) 2 loose tags, A.6(c) 2 medium tags, A.6(d) 2 tight tags and A.6(e) 3 tags.	121
A.7	Validation plots for the jet with second highest transverse momentum p_T for events with 1 electron, exactly 5 jets and A.7(a) 1 tight tag, A.7(b) 2 loose tags, A.7(c) 2 medium tags, A.7(d) 2 tight tags and A.7(e) 3 tags.	122

A.8	Validation plots for the transverse momentum p_T of the lepton for events with 1 electron, exactly 5 jets and A.8(a) 1 tight tag, A.8(b) 2 loose tags, A.8(c) 2 medium tags, A.8(d) 2 tight tags and A.8(e) 3 tags.	123
A.9	Validation plots for the transverse momentum p_T of the neutrino for events with 1 electron, exactly 5 jets and A.9(a) 1 tight tag, A.9(b) 2 loose tags, A.9(c) 2 medium tags, A.9(d) 2 tight tags and A.9(e) 3 tags.	124
A.10	Validation plots for the jet with highest transverse momentum p_T for events with 1 electron, at least 6 jets and A.10(a) 1 tight tag, A.10(b) 2 loose tags, A.10(c) 2 medium tags, A.10(d) 2 tight tags and A.10(e) 3 tags.	125
A.11	Validation plots for the jet with second highest transverse momentum p_T for events with 1 electron, at least 6 jets and A.11(a) 1 tight tag, A.11(b) 2 loose tags, A.11(c) 2 medium tags, A.11(d) 2 tight tags and A.11(e) 3 tags.	126
A.12	Validation plots for the transverse momentum p_T of the lepton for events with 1 electron, at least 6 jets and A.12(a) 1 tight tag, A.12(b) 2 loose tags, A.12(c) 2 medium tags, A.12(d) 2 tight tags and A.12(e) 3 tags.	127
A.13	Validation plots for the transverse momentum p_T of the neutrino for events with 1 electron, at least 6 jets and A.13(a) 1 tight tag, A.13(b) 2 loose tags, A.13(c) 2 medium tags, A.13(d) 2 tight tags and A.13(e) 3 tags.	128

A.14 Validation plots for the jet with highest transverse momentum p_T for events with 1 muon, exactly 4 jets and A.14(a) 1 tight tag, A.14(b) 2 loose tags, A.14(c) 2 medium tags, A.14(d) 2 tight tags and A.14(e) 3 tags.	129
A.15 Validation plots for the jet with second highest transverse momentum p_T for events with 1 muon, exactly 4 jets and A.15(a) 1 tight tag, A.15(b) 2 loose tags, A.15(c) 2 medium tags, A.15(d) 2 tight tags and A.15(e) 3 tags.	130
A.16 Validation plots for the transverse momentum p_T of the lepton for events with 1 muon, exactly 4 jets and A.16(a) 1 tight tag, A.16(b) 2 loose tags, A.16(c) 2 medium tags, A.16(d) 2 tight tags and A.16(e) 3 tags.	131
A.17 Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, exactly 4 jets and A.17(a) 1 tight tag, A.17(b) 2 loose tags, A.17(c) 2 medium tags, A.17(d) 2 tight tags and A.17(e) 3 tags.	132
A.18 Validation plots for the jet with highest transverse momentum p_T for events with 1 muon, exactly 5 jets and A.18(a) 1 tight tag, A.18(b) 2 loose tags, A.18(c) 2 medium tags, A.18(d) 2 tight tags and A.18(e) 3 tags.	133
A.19 Validation plots for the jet with second highest transverse momentum p_T for events with 1 muon, exactly 5 jets and A.19(a) 1 tight tag, A.19(b) 2 loose tags, A.19(c) 2 medium tags, A.19(d) 2 tight tags and A.19(e) 3 tags.	134

A.20	Validation plots for the transverse momentum p_T of the lepton for events with 1 muon, exactly 5 jets and A.20(a) 1 tight tag, A.20(b) 2 loose tags, A.20(c) 2 medium tags, A.20(d) 2 tight tags and A.20(e) 3 tags.	135
A.21	Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, exactly 5 jets and A.21(a) 1 tight tag, A.21(b) 2 loose tags, A.21(c) 2 medium tags, A.21(d) 2 tight tags and A.21(e) 3 tags.	136
A.22	Validation plots for the jet with highest transverse momentum p_T for events with 1 muon, at least 6 jets and A.22(a) 1 tight tag, A.22(b) 2 loose tags, A.22(c) 2 medium tags, A.22(d) 2 tight tags and A.22(e) 3 tags.	137
A.23	Validation plots for the jet with second highest transverse momentum p_T for events with 1 muon, at least 6 jets and A.23(a) 1 tight tag, A.23(b) 2 loose tags, A.23(c) 2 medium tags, A.23(d) 2 tight tags and A.23(e) 3 tags.	138
A.24	Validation plots for the transverse momentum p_T of the lepton for events with 1 muon, at least 6 jets and A.24(a) 1 tight tag, A.24(b) 2 loose tags, A.24(c) 2 medium tags, A.24(d) 2 tight tags and A.24(e) 3 tags.	139
A.25	Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, at least 6 jets and A.25(a) 1 tight tag, A.25(b) 2 loose tags, A.25(c) 2 medium tags, A.25(d) 2 tight tags and A.25(e) 3 tags.	140
B.1	The 95% CL upper limit on the $\sigma_{t\bar{t}H}$ cross section times branching ratio over the SM expectation in NLO QCD as a function of the Higgs mass in (a) electron and (b) muon channels.	142

C.1	Variables used for BDT training for 4 jet, 1 tight tag category. Signal distributions are shown in blue and background distributions in red. . .	144
C.2	Variables used for BDT training for 4 jet, 2 loose tag category. Signal distributions are shown in blue and background distributions in red. . .	145
C.3	Variables used for BDT training for 4 jet, 2 medium tag category. Signal distributions are shown in blue and background distributions in red. . .	146
C.4	Variables used for BDT training for 4 jet, 2 tight tag category. Signal distributions are shown in blue and background distributions in red. . .	147
C.5	Variables used for BDT training for 4 jet, 3 tag category. Signal distributions are shown in blue and background distributions in red.	148
C.6	Variables used for BDT training for 5 jet, 1 tight tag category. Signal distributions are shown in blue and background distributions in red. . .	149
C.7	Variables used for BDT training for 5 jet, 2 loose tag category. Signal distributions are shown in blue and background distributions in red. . .	150
C.8	Variables used for BDT training for 5 jet, 2 medium tag category. Signal distributions are shown in blue and background distributions in red. . .	151
C.9	Variables used for BDT training for 5 jet, 2 tight tag category. Signal distributions are shown in blue and background distributions in red. . .	152
C.10	Variables used for BDT training for 5 jet, 3 tag category. Signal distributions are shown in blue and background distributions in red.	153

Abstract

This thesis describes the search for the associated production of a Higgs boson with top anti-top pairs in proton anti-proton collisions at a centre-of-mass energy of 1.96 TeV. The data used were collected by the D0 detector during the RunII data taking period which ran from 2001 to 2011 at the Tevatron Collider located at the Fermi National Accelerator Laboratory (or Fermilab).

Distributions of the H_T variable separated into 3 jet, 5 b -tagging and 2 lepton categories were used as inputs to a modified frequentist limit setting procedure, which was in turn used to search for the $t\bar{t}H$ process with $H \rightarrow b\bar{b}$. Since this process was not observed, limits were set on the cross section times branching ratio $\sigma(t\bar{t}H) \times BR(H \rightarrow b\bar{b})$, with an expected (observed) limit of 24.7 (74.3) measured for a Higgs mass value of 125 GeV.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

James Paul Agnew

School of Physics and Astronomy

University of Manchester

Oxford Road

Manchester

M13 9PL

September 2013

Copyright

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in The University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the School of Physics and Astronomy.

Acknowledgements

It has been a privilege to work under the supervision of Professors Terry Wyatt and Christian Schwanenberger, and I thank them both for their help and guidance.

I thank Yvonne Peters for her expertise in the early stages of my postgraduate studies, and Bjoern Penning for guiding me while at Fermilab and for the enormous effort he contributed in the final stages. They were both more patient with me than I really deserved.

I also thank the members of the WH group, particularly Lidija, Mike and Savanna, for helping me to navigate the WH framework.

Thank you to Professor Fred Loebinger for accepting me on to the course in the first place, and for keeping us all in good spirits.

Thank you to all my Fermifriends for keeping me thoroughly entertained while I spent time at Fermilab and long after, and to my Manchester based friends for football, Korean food and fun.

Thank you to my family for their enduring support and effort to understand what it is I've been doing.

Finally, thank you to my wife Imogen for your support, love and sacrifice. Without you I would never have gotten this far.

Chapter 1

Introduction

In this thesis I report on a search for the Standard Model process $t\bar{t}H$ in a final state consisting of one lepton, at least 4 jets and missing transverse energy, performed as part of a postgraduate degree at the University of Manchester.

I will begin by briefly summarizing the theoretical background for the Standard Model, paying particular attention to areas directly related to the process in question and the relevant final state. This will be followed by an overview of the steps taken to accelerate protons and anti-protons such that they collide at centre of mass energy of $\sqrt{s} = 1.96$ TeV, and of the D0 detector, at which the data used in the analysis was recorded. A discussion of the object identification will follow from this, before moving on to talk about methods of data simulation and correction factors which had to be applied. Once all of these important considerations have been accounted for, the main topics of this thesis will then follow as I describe the experimental methods and results for the search for $t\bar{t}H$.

Chapter 2

Theory

2.1 Standard Model

The Standard Model (SM) is a gauge-invariant quantum field theory that describes spin-1/2 particles, fermions, and their interactions through forces, which are mediated by particles of integer spin, or bosons. Its structure arises from the symmetry group $SU(3) \times SU(2) \times U(1)$.

The Standard Model has so far provided an accurate description of all known particles and their interactions over many orders of magnitude of energies.

2.1.1 Forces in the Standard Model

There are four fundamental forces in nature:

- The strong force, which is described by the $SU(3)$ symmetry group
- The weak force, which is described by the $SU(2)$ symmetry group
- The electromagnetic force, which is described by the $U(1)$ symmetry group

- The gravitational force. This force is many orders of magnitude weaker than the other three fundamental forces and cannot yet be incorporated into the SM. Hence it is neglected for the remainder of this thesis.

The Strong Force

The mediator of the strong force is the gluon, which is massless and carries a colour charge. There are 8 gluons, identical except for their colour charge. As the gluons themselves carry colour charge and only interact with particles that carry colour charge, they are able to interact with themselves. In terms of their symmetry group, this behaviour is non-abelian.

The strong force coupling parameter, α_S , decreases with energy, with its behaviour described by the equation:

$$\frac{\partial \alpha_S^2(Q^2)}{\partial \ln(Q^2)} = - \left(11 - \frac{2n_f}{3} \right) \frac{\alpha_S}{4\pi} \quad (2.1)$$

where Q^2 is the energy scale of the interaction and n_f is the number of families of fermions. In the SM, $n_f = 6$.

The Electromagnetic Force

The mediator of the electromagnetic (EM) force is the photon, γ , which is massless and couples to charged particles, though is itself uncharged. Because it is massless, the range of the EM force is infinite.

The Weak Force

The mediators of the weak force are the W^+ , W^- and Z^0 bosons. Unlike the photon and gluons, these are massive particles, with their masses determined experimentally to be 80.385 ± 0.015 GeV [1] for the W^\pm bosons and 91.1876 ± 0.0021 GeV [1] for the

Z^0 . The high masses of these bosons mean that they can only act over short distances. The bosons of the weak force couple to particles which have weak isospin, though the Z^0 can couple with particles with weak isospin or electric charge.

Electroweak Symmetry Breaking and the Higgs Field

The coupling strength of the EM and weak forces both increase with energy. The Glashow-Salam-Weinberg model [2, 3, 4] showed that at high enough energies, the EM and weak forces combine to yield the electroweak (EW) force, which is thus described by a $SU(2) \times U(1)$ group. From the gauge invariance requirement, four gauge bosons will result. The $SU(2)$ component yields three of these: W^1 , W^2 and W^3 , and since $SU(2)$ is non-abelian, this means that these three bosons are able to interact with each other. The remaining gauge boson is the B and arises from the $U(1)$ group. The photon and weak bosons may then be described as a mixture of the $SU(2) \times U(1)$ gauge bosons as follows:

$$W^\pm = \frac{1}{\sqrt{2}} (W^1 \mp W^2) \quad (2.2)$$

and

$$\begin{pmatrix} Z \\ \gamma \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix} \quad (2.3)$$

where θ_W is known as the weak mixing angle and derives from the coupling constants of the $SU(2)$ (g_W) and $U(1)$ (g'_W) groups such that:

$$\sin \theta_W = \frac{g'_W}{\sqrt{g_W^2 + g'^2_W}} \quad (2.4)$$

With their respective forces combined, the discrepancy between the masses of the

weak bosons and that of the photon must be explained. In 1964, a mechanism for this process was proposed by 3 groups of researchers: Brout and Englert [5], Guralnik, Hagen and Kibble [6] and Higgs [7]. Higgs proposed a massive scalar boson identified with the field, which we call the Higgs boson, and hence the Higgs field, earning him and Englert the 2013 Nobel Prize in Physics.

The Higgs mechanism allowed for the symmetry of the massless bosons to be broken spontaneously through interaction with a complex scalar field ϕ :

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.5)$$

The field ϕ has potential energy V :

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.6)$$

with $\mu^2 < 0$ and $\lambda > 0$. Setting $\phi^+ = 0$ at the minimum of the potential V_0 (with ϕ_0 the field at the minimum potential) and allowing ϕ^0 a non-vanishing component:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.7)$$

then the potential has minima at $\phi_0 = \pm v$. At some point in the early Universe then, the Higgs field acquired a vacuum expectation value (or VEV) at v , meaning that it now has an inherent value at every point in space:

$$v = \frac{|\mu|}{\sqrt{\lambda}} \quad (2.8)$$

v has a value of approximately 246 GeV [1]. Taking perturbations h about v ,

$$\phi = \phi_0 + h = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \quad (2.9)$$

Using this expression for ϕ in the potential then gives

$$V = V_0 + \lambda v^2 h^2 \quad (2.10)$$

The Higgs boson corresponds to these perturbations in ϕ , with a mass term $M_H = \sqrt{2\lambda}v$. Masses of the EW bosons are then given by the following expressions:

$$M_{W^\pm} = \frac{vg}{2} \quad (2.11)$$

$$M_Z = \frac{v\sqrt{g^2 + g'^2}}{2} \quad (2.12)$$

$$M_\gamma = 0 \quad (2.13)$$

where g and g' are the coupling constants of the $U(1)$ and $SU(2)$ groups.

The Higgs boson decays into pairs of massive particles, provided that this is kinematically viable. Figure 2.1 shows the branching ratio of a Higgs over a range of mass values. Note that a Higgs decay to two photons is included, even though there is no direct coupling between a Higgs and a massless particle - this decay occurs indirectly through either a boson or fermion loop (see Figure 2.2 for possible decay scenarios).

In 2012, collider experiments at CERN confirmed the existence of a Higgs-like boson and in 2013 identified it as a Higgs boson [10, 11] with a mass of approximately 126 GeV, a huge success for the Standard Model. In 2012, at the Fermilab Tevatron, the CDF and D0 experiments saw first evidence for the fermionic decay of such a particle, specifically to $b\bar{b}$ [12].

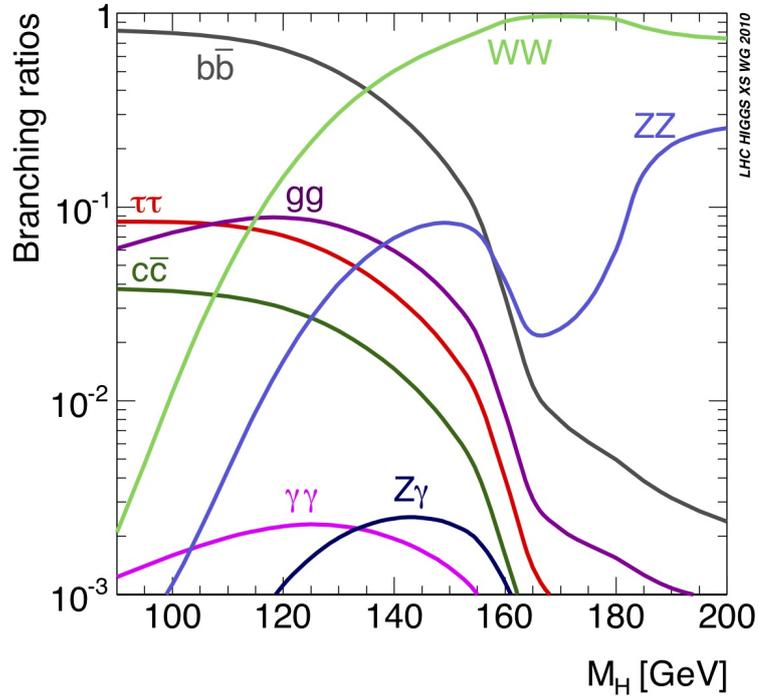


Figure 2.1: Branching ratios of the Higgs for a range of Higgs masses [8].

2.1.2 Leptons and Quarks

There are 3 flavours of leptons in the Standard Model: electron (e), muon (μ) and tau (τ). For each flavour there is a charged and neutral lepton. The charged leptons are named as their respective flavours and carry an electric charge of -1, while the neutral leptons are called neutrinos. The 3 flavours can be written as 3 families of weak isospin doublets.

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

For each lepton there exists an anti-lepton, which has the same mass but opposite charge. As well as electric charge, the leptons also carry weak isospin.

There are 6 flavours of quark in the SM. These are: up, down, charm, strange, top and bottom and, as is the case for leptons, for each there exists an anti-quark which has the same mass but opposite electric charge. The charges carried by quarks are as follows:

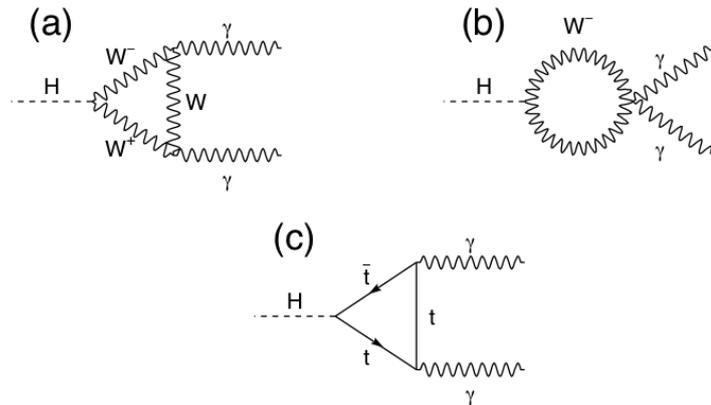


Figure 2.2: The Higgs decay to two photons through (a), (b) a W boson pair and (c) a top quark pair [9].

- Colour charge (red, green or blue)
- EM charge
- Weak isospin.

Their colour charge allows them to interact with the strong force. At low energies, this dominates their interactions. At such energies, they exist only as colourless objects (that is, they must either consist of a combination of all 3 colours or anti-colour charges, or of a colour-anti-colour pair) in bound states known as hadrons. Hadrons consisting of 3 quarks are known as baryons, with the proton and neutron both being examples of these. Hadrons consisting of 2 quarks are known as mesons, with examples including the π^0 and the B^0 . At much higher energies, the quarks approach what is known as asymptotic freedom as the strong force is weaker, so they are almost able to move freely, without as much interaction with other quarks in the bound state.

As with leptons, the quarks can be grouped into 3 families:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}.$$

The mass eigenstates of the down-type quarks are not the same as the flavour eigenstates. Hence, there is mixing between the quark flavours. This is described by the

Cabibbo-Kobayashi-Maskawa (CKM) matrix [13, 14]:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

The element of the CKM matrix most relevant to this analysis is V_{tb} . It has been measured experimentally to take a value of $|V_{tb}| = 0.999146^{+0.000021}_{-0.000046}$ [1]. For the purpose of this thesis, we assume it takes a value of 1. This means that all top quarks would decay into bottom quarks via the weak interaction, so releasing a W boson as well.

2.1.3 The Top Quark

The top quark is markedly different from the other quarks, primarily in that its mass is so much larger than the other quark masses. Experimentally it has been measured as 173.20 ± 0.87 GeV at the Tevatron [15]. This is much heavier than its closest neighbour the bottom quark, which has a mass of 4.18 ± 0.03 GeV [1].

The top quark is also unique in that it decays before hadronizing. All other quarks will fragment into a spray of hadrons, which are called jets. Since the top quark decays to Wb before it hadronizes, it allows the study of a naked quark without the measurement being affected by soft colour effects, which will arise from interactions between quarks in a hadron both before and after the decay of the quark in question.

As noted above, the b quark will hadronize immediately, while the W decays into quarks or leptons. Since we assume that the top can only decay into b quarks, it is the W decay that characterizes the top decay process. If both W 's in our $t\bar{t}$ events decay leptonically (that is, into a charged lepton plus its corresponding neutrino), we describe the $t\bar{t}$ decay as ‘dileptonic’. If both W 's decay hadronically, we describe the event as ‘all hadronic’ (or ‘all jets’). We mainly consider events that fall between these two categories; where one W decays leptonically and the other hadronically (which

we refer to as ‘lepton plus jets’ events), though dilepton decays are also considered for the SM $t\bar{t}$ process.

There are advantages to choosing such events for our search. While the all hadronic decay channel represents the greatest fraction of decay possibilities, as shown in Figure 2.3, it is difficult to identify all of the jets in the event correctly. This would be even more of a problem when looking for an additional pair of b jets coming from a Higgs decay.

In contrast, the dileptonic channel has a clean signal of two high-momentum charged leptons as well as missing energy and two high-momentum jets (from the hadronized b quarks). However, less than 10% of events decay in this way. Furthermore, the fact that the missing energy in the event must be attributed to two particles can lead to reconstruction difficulties.

In this thesis the lepton+jets decays of $t\bar{t}$ pairs produced in association with a Higgs boson was investigated. With a low mass Higgs, this is the only direct way to measure the Yukawa coupling of the top quark, which is expected to be ~ 1 . The large Yukawa coupling may indicate that the top quark plays an important role in electroweak symmetry breaking, such as in certain ‘Little Higgs’ models which include the existence of a heavy $SU(2)$ singlet quark, T [16], in which the Higgs arises as a pseudo-Goldstone boson. The existence of a fourth generation vector-like quark t' , resulting from the decay of a vector colour octet, may also affect the top quark Yukawa coupling, as the Higgs could also interact with such a particle, thus reducing the coupling to the known quarks [17].

Top Quark Production

Top quarks can either be produced via the strong or the electroweak interaction.

At the Tevatron the $t\bar{t}$ pair production via strong interaction dominates, while the production of a single top quark via the electroweak interaction occurs less frequently.

Top Pair Branching Fractions

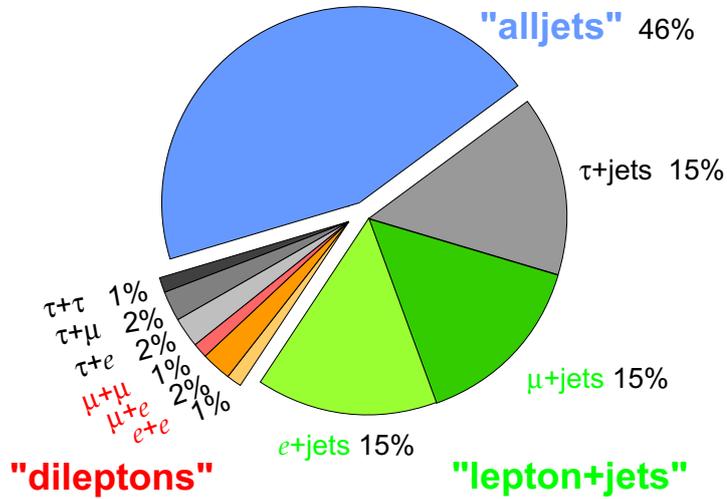


Figure 2.3: Branching fractions of the possible decays for two W bosons.

Pair production of top quarks results either from $q\bar{q}$ annihilation, as in Fig 2.4(a), which contributes around 85% of the $t\bar{t}$ production at the Tevatron, or by gluon-gluon fusion, as in Fig 2.4(b), which contributes the remaining 15% of the production cross section. In contrast, At the LHC, the gluon-gluon fusion dominates the $t\bar{t}$ production cross section with a contribution of 90%.

Single top quark production can occur either via the s-channel process $q\bar{q} \rightarrow W + t\bar{b}$, or by the t-channel process $gq \rightarrow qt\bar{b}$, and generally produce final states with less jets.

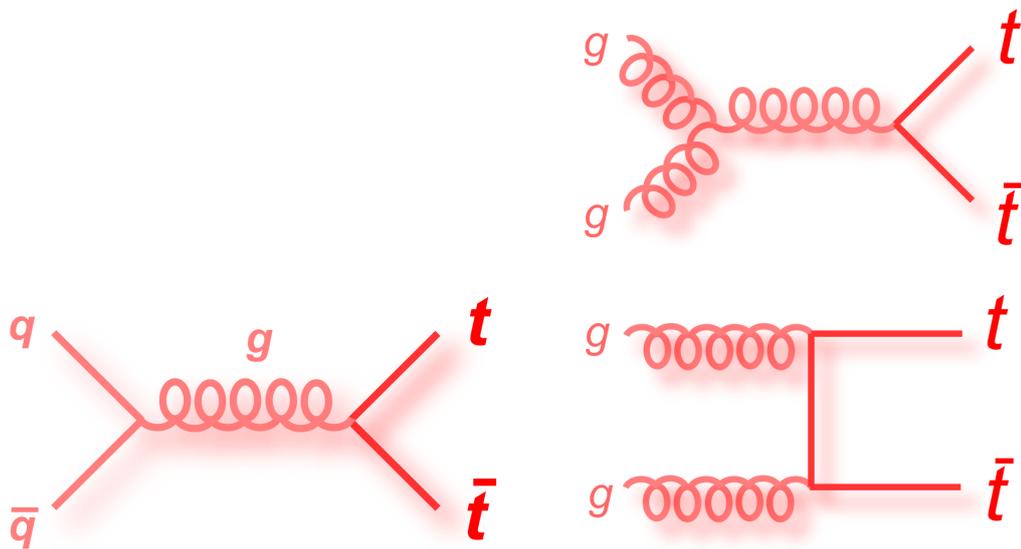


Figure 2.4: Feynman diagrams of the production mechanisms for $t\bar{t}$ production at the Tevatron [18].

Chapter 3

The Tevatron and the D0 Detector

3.1 The Tevatron

The Tevatron is a proton-antiproton collider that operated from 1983 - 2011. Measuring almost 4 miles (6.4 km) in circumference, it is located at the Fermi National Accelerator Laboratory (FNAL, or Fermilab) near Batavia, Illinois. It played host to two major particle physics experiments: CDF and D0, both of which have active collaborations still.

In order to collide protons and antiprotons, they must first be accelerated through a chain of processes. To make the proton beam, hydrogen atoms are ionised and then accelerated to an energy of 750 keV in the Cockcroft-Walton. A linear accelerator (or 'linac') then accelerates the ions to an energy of 400 MeV, after which any remaining electrons in the beam are removed when the beam passes through a carbon foil. The Booster accelerates the proton beam to 8 GeV, and from here the beam goes into the Main Injector, where it is accelerated to 120 GeV.

In the Main Injector, the proton beam is directed at a nickel target to produce antiprotons (as well as other particles, which are separated using a magnetic charge-mass spectrometer). Their momenta are partly homogenised in the Debuncher, before being moved to the Recycler, where their momenta are further homogenised by the process

FERMILAB'S ACCELERATOR CHAIN

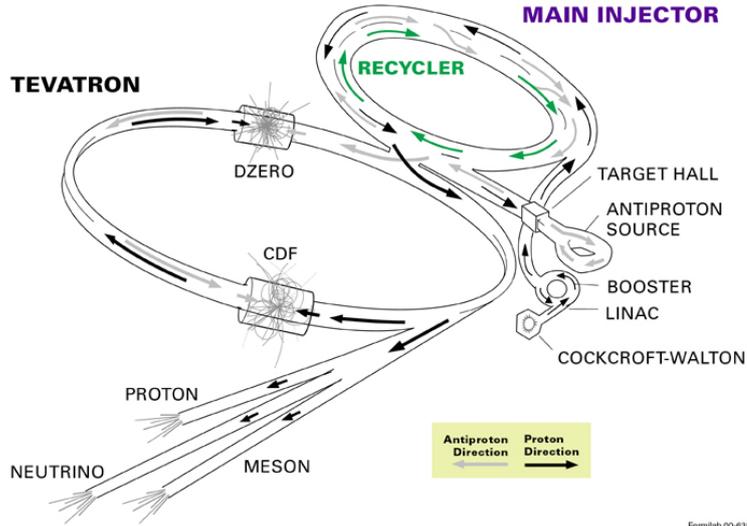


Figure 3.1: The Tevatron accelerator chain [19]

of stochastic cooling. The protons, in the meantime, are stored in the Accumulator. Once enough protons and antiprotons have been produced, they are transferred back in to the Main Injector to be accelerated to an energy of 150 GeV, before finally being transferred to the Tevatron ring itself. Here, each beam is accelerated to energies of 980 GeV in opposite directions around the ring, providing a centre of mass collision energy of $\sqrt{s} = 1.96$ TeV, giving rise to the Tevatron's name.

The beams each consist of 36 'bunches', with 3 groups of 12 bunches each separated by $7 \mu\text{s}$, and with $0.396 \mu\text{s}$ separating each bunch within a group. In total there are around 10^{11} protons and 10^{10} antiprotons in each bunch in their respective beams, allowing for a high rate of collisions - essential for probing rare physics interactions.

There are several collision points around the ring, with the two general-purpose detectors CDF and D0 each placed around one of them. The ring itself consists of around 1000 superconducting dipole magnets, each cooled with liquid helium to a temperature of 4.2 K, providing a 4.2 T magnetic field.

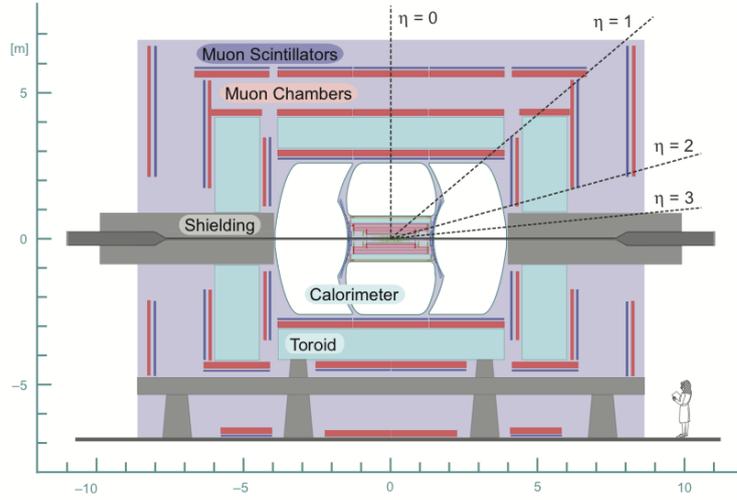


Figure 3.2: A cutaway view of the D0 detector [20]

3.2 The D0 Detector

D0 is an all-purpose detector, with a focus on ability to accurately measure events involving high mass particles and with high transverse momentum, p_T . The high mass particles tend to decay close to the interaction point, and the detector is designed such that their decay products deposit their energy in different parts of the detector, depending on the products themselves. To this end, D0 consists of a central tracking system, EM and hadronic calorimeters and a muon detector. If there is good understanding of the D0 detector then we can infer which particles were produced in the interaction by reconstructing the decay products.

D0 uses a right-handed coordinate system. In Cartesian coordinates, the positive z -axis points along the proton beam, with the y -axis pointing vertically upwards and the x -axis points into the Tevatron ring. Because the detector is cylindrically symmetrical about the beamline, cylindrical coordinates are used in the analyses, with:

$$r = \sqrt{x^2 + y^2} \quad (3.1)$$

$$\phi = \arctan \frac{x}{y} \quad (3.2)$$

$$\eta = -\ln \left[\tan \frac{\theta}{2} \right] \quad (3.3)$$

where θ is the azimuthal angle with respect to the z axis. η is called the pseudorapidity, and for massless particles it is equal to the rapidity:

$$y = \frac{1}{2} \ln \left[\frac{E + p_z}{E - p_z} \right] \quad (3.4)$$

3.2.1 Tracking

The tracking system at the centre of D0 is important for precisely recording the interaction point, as well as any displaced vertices, which are key hallmarks of b jets - key objects in this analysis. It is also important for measuring the momenta of particle tracks and has a momentum resolution of

$$\delta p_T/p_T = 0.02 \oplus 0.002 p_T \quad (3.5)$$

where the first term on the right-hand side of the equation represents a multiple scattering term.

The tracking system consists of a Silicon Microstrip Tracker (SMT), located at the collision point in the very centre of the detector, and Central Fiber Tracker (CFT) which is positioned around the SMT. Both are housed within a superconducting solenoid magnet with a field of 2T, with a cryostat to provide the necessary cooling. In addition, there are preshower detectors located outside the solenoid. These detectors increase the efficiency of distinguishing between electrons and photons (discussed further in chapter 4). There are two sets of detectors; the central preshower detector (CPS) is located between the toroid and the calorimeter, and covers the region $\eta < 1.3$, while the forward preshower detectors (FPS) are located in front of the endcap calorimeter

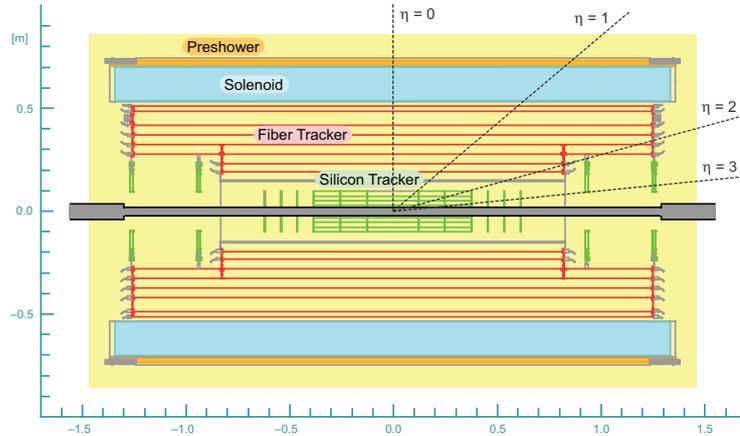


Figure 3.3: A side view of the D0 tracking system [20].

and cover $1.5 < \eta < 2.5$. Figure 3.3 shows a side view of the tracking system as a whole.

Silicon Microstrip Tracker

The SMT consists of 6 'barrels' and 16 discs (see Figure 3.4). The barrels provide tracking information in the $r - \phi$ plane and are particularly useful for 'central' vertices (vertices that lie in an η range of $|\eta| < 1$), while the discs provide information for the $r - z$ plane and are used for identifying 'forward' vertices (vertices whose corresponding tracks lie in an η range of $|\eta| > 1$). Taken together, they allow tracks to be reconstructed in 3 dimensions.

Each barrel has 4 layers of silicon readouts, arranged in modules called ladders. The layers are further divided into two sub-layers. The first and second layers each consist of 12 ladders, while the third and fourth layer each have 24.

The four central barrels use double-sided double-metal (DSDM) detectors in layers 1 and 3, while layers 2 and 4 in use a single-sided (SS) detector in all barrels. The outermost barrels use double-sided (DS) detectors. The DSDM detectors consist of axial and stereo strips with a stereo angle of 2° , while the SS detectors only consist of axial

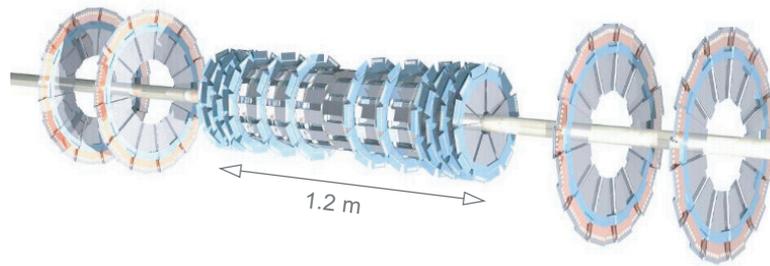


Figure 3.4: A 3/4 view of the D0 SMT [20]

strips.

An additional layer, layer 0, was added to the SMT in 2006, providing improved vertex resolution for the RunIIb data taking period. It is the closest layer to the interaction point, with two sensors at a distance of 16.1 and 17.6 mm from the beam axis ??, and is located inside layer 1.

Central Fiber Tracker

The CFT has 16 doublet layers of fibers, arranged in 8 cylinders each containing 2 doublet layers. In each cylinder, one layer is aligned along the beam axis while the second layers are offset by an angle of $\pm 3^\circ$, where the sign alternates between the doublet layers in the cylinders. These doublet layers contain scintillating fibres which are connected by waveguides made from clear fibre to visible light photon counters (VLPC). These VLPC are located below the detector, which means that the waveguides connected to scintillators at $\phi = \pi/2$ are necessarily longer than those connected at $\phi = 3\pi/2$, giving rise to a ϕ -dependency in the tracking efficiency which must be corrected for.

3.2.2 Calorimeters

The calorimeter is designed to measure the energy of electrons, photons and jets. This information is further used to determine if there is any missing transverse energy in the

event, which points to the presence of an object which does not interact with matter much, such as a neutrino, as these will carry away a portion of the energy of the event without depositing any of it within the detector.

The calorimeter consists of three sections: the central calorimeter (CC) and the two endcap calorimeters (EC). The CC covers the range $|\eta| < 1.1$ while the EC sections each cover the range $1.4 < |\eta| < 4$. The gap between these two η regions is partly covered by the Inter Cryostat Detectors (ICD) in the range $1.1 < |\eta| < 1.3$, which are so-called because they lie outside of the cryostats used to maintain a low temperature for the liquid argon contained within the cells of the CC and EC.

The three main calorimeter sections are segmented radially in four electromagnetic (EM) layers, which are then surrounded by three fine and one coarse hadronic layers, where fine and coarse refer to the relative size of the cells in the layers. The cells are grouped into “towers” roughly along lines of constant η , and in all but the third layer of the EM calorimeter their transverse extent is comparable to the transverse size of the hadronic and electromagnetic showers at $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The third EM calorimeter layer has cells with half the size in η and ϕ as it is in this layer where EM showers most often reach their maximum extent, thus allowing for more precise measurements of the energy deposits.

Absorber plates made of depleted uranium with a thickness of 3 (4) mm in the CC (EC) regions of the EM calorimeter compel the particles passing through them to impart their energy as they are dense, presenting abundant material for the objects to interact with. Assuming that all of their energy is deposited in the calorimeter, then a measurement can be made on the total energy of the object through the reconstruction techniques described in chapter 4.

The coarse hadronic modules contain thicker absorber plates - 46.5 mm but made from copper in the CC and steel in the EC. A potential difference of 2kV is applied to the

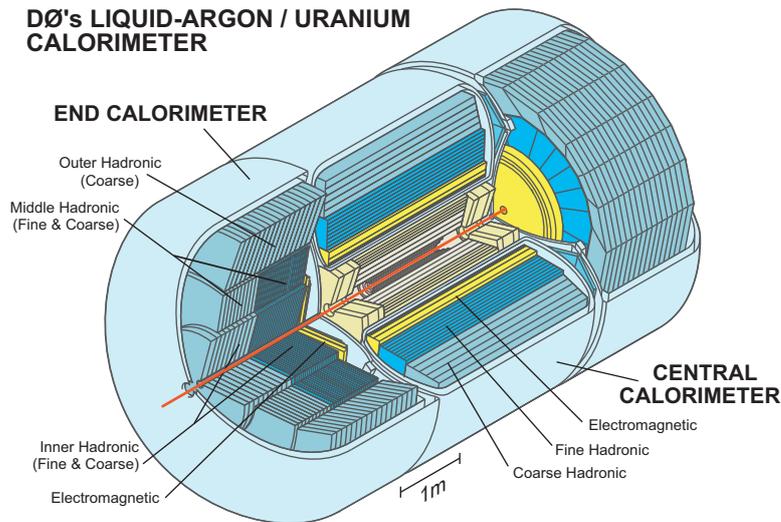


Figure 3.5: A cutaway view of the D0 calorimeter system [20].

small (2.3 mm) gap between the absorber plates. This creates a potential difference, allowing for the collection of charge generated in the liquid argon, which is proportional to the energy deposited in the calorimeter.

3.2.3 Muon Systems

The muon system is situated furthest from the interaction point as muons as they are minimally interacting particles, or MIPS, and as such leave little or no trace in the calorimeter.

The system is composed of drift tubes and scintillators as well as a toroidal magnetic which allows a measurement of the tracks' momenta to be performed which is independent of the measurement performed for tracks in the central detector region which utilises the 2T magnetic field, which ultimately helps to reduce uncertainties on the momentum measurement for muons.

There are three layers in the muon system, labelled A (located before the toroidal magnet in the outward radial direction from the interaction point), B and C (both located after the toroidal magnet), which consist of layers of drift tubes and scintillators. The

sections of the muon system located in the forward regions ($1 \leq |\eta| \leq 2$) use ‘mini’ drift tubes (MDT), while those located in the central region ($|\eta| < 1$) have proportional drift tubes. Scintillator counters in the system are used in the triggering process (discussed briefly in section 3.2.5, whereas the modules with wire chambers are used for measurement as well as triggering.

Figure 3.6 shows the muon system in an exploded view.

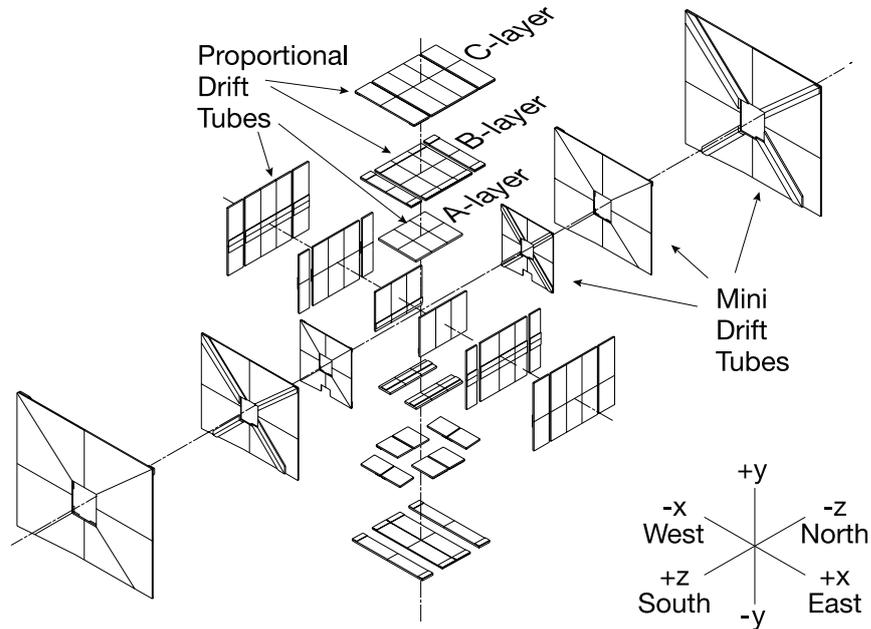


Figure 3.6: The Muon System [22].

3.2.4 Luminosity Detectors

Measuring the luminosity is an important component of physics analyses, so D0 uses luminosity monitors (LM), which are located at $z = \pm 140$ cm (see figure 3.7). These detect inelastic $p\bar{p}$ collisions and consist of 24 15 cm long scintillation counters, with each of these counters occupying the range $2.7 < |\eta| < 4.4$. A timing resolution of 0.3 ns allows the LM to distinguish between particles which come from the interaction point or close to it and those which come from the beam halo (a ring of particles around the beam which result from the beam dynamics due to the Tevatron’s bending

magnets).

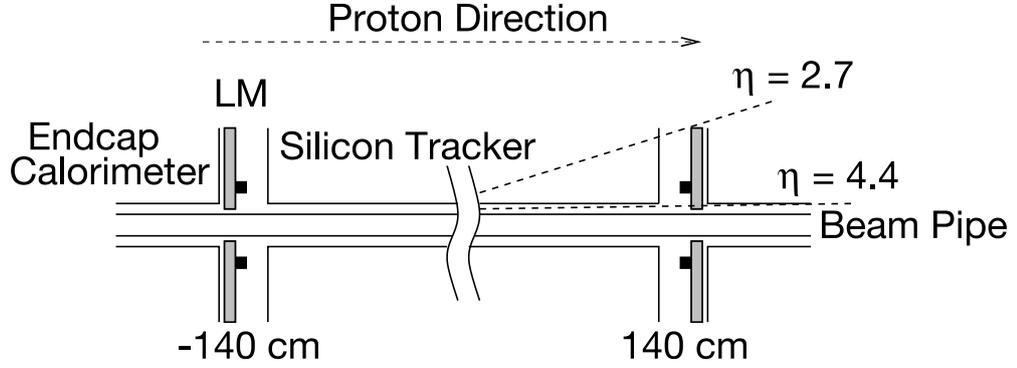


Figure 3.7: Location of the luminosity monitors on the z -axis [22].

The count rate measurement is used to determine the instantaneous luminosity [23]. Between 2006 and 2011, instantaneous luminosity varied from $(5-420) \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ [23]. Instantaneous luminosity \mathcal{L} is given by the equation:

$$\mathcal{L} = \frac{1}{\sigma_{eff}} \frac{dN}{dt} \quad (3.6)$$

where σ_{eff} is the effective inelastic cross section measured by the LM and N the number of interactions. The effective cross section is derived from the inelastic cross section $\sigma_{inelastic}(1.96 \text{ TeV}) = 60.7 \pm 2.4 \text{ mb}$, after taking into account acceptance effects and the efficiency of the LM.

The integrated luminosity is calculated in so-called luminosity blocks. Each luminosity block is assigned a luminosity block number (LBN) which increments at the end of every beam store, or after 60 seconds. The LBN then dictates the fundamental unit of time for the luminosity measurement, which is chosen such that for each LBN the integrated luminosity is approximately constant.

3.2.5 Triggering

There are almost 2 million collisions per second at each beam crossing point of the Tevatron. These collisions mostly consist of soft scattering processes and noise, while only a small fraction of events involve a sufficiently large momentum transfer necessary to produce particles that D0 analyses aim to study. Hence, some way of filtering out the less interesting events is required, as there are limits on the rate and amount of information pertaining to an event which can be recorded. At D0, this rate is limited to 100 Hz. As such, a three-step process [22] is employed to make sure that the events which are most likely to contain interesting physics are recorded. Figure 3.8 provides an overview of the D0 trigger and data acquisition (DAQ) system while the following sections detail each of the three steps in the triggering process.

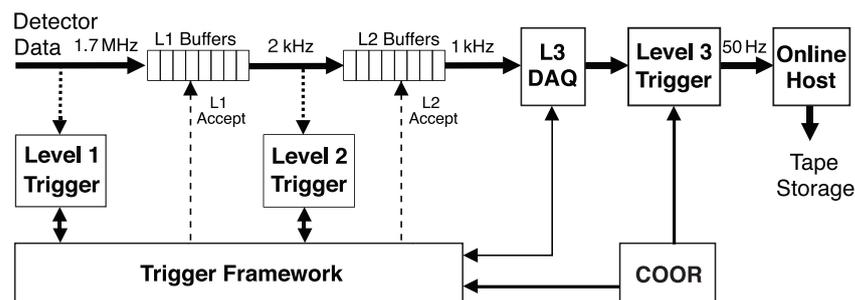


Figure 3.8: The D0 trigger and DAQ system [22].

The L1 Trigger

The L1 trigger is based on specialized hardware and consists of the following components:

- The Forward Proton Trigger (L1FPD). This selects events by triggering on protons or antiprotons scattered at small angles, which are indicative of diffractive events with little momentum transfer.

- The Central Track Trigger (L1CTT) and Muon System Trigger(L1Muon), which identify tracks which pass momentum thresholds, and partly work in conjunction with each other, though the Muon System Trigger is only used to identify muons with momenta which pass the threshold.
- Calorimeter Trigger (L1Cal), which triggers on objects in the calorimeter which exceed threshold values of the transverse energy.

The L1 trigger rate is limited by the maximum readout rates of the subsystems involved in the event, as well as a need to minimize the dead-time associated with the readout.

The L2 Trigger

The L2 trigger consists of two stages. In the preprocessor stage, information from the subsystems is gathered from the L1 system and the data is analyzed in order to reconstruct objects (see Chapter 4 for further details about this process). This stage consists of the Level 2 versions of the triggers listed above for L1, with the exception that the L2 version of the Central Track Trigger consists of separate triggers for the SMT and the CTT, as well as utilising a preshower detector trigger (L2PS). In the global stage (L2Global), the data from across the subsystems is combined to form physics objects. The decision as to whether an event is then kept or not is made within $100\mu s$, which reduces the data acquisition rate to ~ 1 kHz.

The L3 Trigger

The final L3 trigger reduces the amount of events deemed to be uninteresting further still, without entailing a volume of output which would be too great to store permanently. The L3 trigger uses ‘farm’ of computing nodes to perform a rudimentary reconstruction of events.

Every event which passes the L1 and L2 trigger stages gets fully read-out before being

sent to the L3 farm. A single farm node processes all data blocks associated with a single data event.

Chapter 4

Identification of Particles in the Detector

When particles interact at or near the collision point, some of the products of these interactions leave behind signatures in the detector thanks to the detector's design. Using these signatures, physical objects can be inferred using algorithms which reconstruct the objects.

This chapter will describe those signatures which are relevant to this analysis, namely:

- Tracks in the detector and vertex identification
- Electron and muon identification
- Quantification of missing transverse energy
- Jet reconstruction and b -jet identification

The other major identification and reconstruction processes are of tau leptons and photons. Tau reconstruction is a complex process due to the multiple types of decay available to tau leptons, but is not relevant to this thesis owing to the requirement of

electrons or muons as the only leptons in the final state (see chapter Event Selection for further information). We also do not require photons in the final state, but as photon identification is part of the electron ID process it will be mentioned in the appropriate section.

4.1 Tracks in the detector and vertex identification

4.1.1 Tracks

Tracks are left in the detector when particles with electric charge pass through the SMT and CFT and leave 'hits'; that is, interact with the detector material. Clusters of hits are reconstructed to form tracks using a combination of three algorithms.

The Histogram Track Finder (HTF) [24] algorithm and Alternative Algorithm (AA) [25] are pattern-finding algorithms which group clusters to identify track candidates in the detector. The HTF populates histograms based on track parameter values (the curvature of the track and its azimuthal angle at the distance of closest approach to the centre of the detector), with a peak in the histogram indicating a track candidate. The AA forms track candidates by taking combinations of three hits in the barrel or discs and extrapolating this proto-track into the next layer of the SMT or CFT. New hits are incorporated into the proto-track if a fit of the hits satisfies a likelihood constraint based on the curvature of the proto-track and the angle between the hits. If there are multiple hits within the layer that satisfy the constraints, additional track candidates are formed and the extrapolation continues from each hit. If a layer has no hits which satisfy the track constraints, this is recorded as a 'miss'. The HTF is most useful for high- p_T tracks, while the AA is best for low- p_T tracks, so using both is optimal for a wide range of analyses, particularly ones in which a range of objects are expected [26]. The Kalman track fit algorithm [27] iterates track candidates from the HTF and AA

and propagates them through the D0 tracking system using the D0 interacting propagator [28]. From the fitting process the final track attributes are assigned, along with their associated uncertainties, with some of the candidates being deemed unsuitable. The magnetic field from the 2T solenoid causes the tracks to curve, which allows the momenta and charge of the particles to be determined; another key tool in the reconstruction process.

4.1.2 Vertices

Once tracks have been reconstructed, it is then possible to look for vertices in the event. Primary vertices (PV) are points from which a number of tracks originate and indicate that a hard interaction has taken place there. A secondary vertex (SV) can indicate that an object such as a B meson was produced at the PV and decayed in flight. The tracks which then trace back to the SV are most likely to be its decay products.

Accurate and precise identification of PV is essential for separating the objects which result from the hard interaction from the underlying event (such as QCD processes and scattering) as well as from SVs. Identifying the objects which are originated from the hard interaction is essential for accurately quantifying the missing transverse energy in the event (see Section 4.3 for further details), which is required for this analysis (see Chapter 7 for further details).

Primary vertices are reconstructed using the Adaptive Primary Vertex (APV) algorithm [29] as an iterative Kalman fitter (the same as mentioned previously). Tracks with $p_T > 0.5$ GeV and two or more hits in the SMT are fitted simultaneously and the tracks which contribute the highest χ^2 value are iteratively removed until the fit achieves a χ^2/NDF (where NDF is the number of degrees of freedom in the fit) value of less than 10. The solution will not be unique as many interactions can occur in a bunch crossing, so the PV are ordered in terms of the p_T of the tracks associated

with them and the one associated with the highest p_T tracks is taken to be the PV as it is the least likely to have come from soft scattering in the underlying event.

4.2 Electron and Muon Identification

4.2.1 Electrons

Electrons are characterised by having tracks in the SMT and CFT (as they are charged particles) and by depositing energy in the calorimeter. As such, electrons are reconstructed by matching tracks to clusters of energy deposits.

For identifying clusters, an algorithm [30] is used to identify so-called ‘seed’ cells which have energy above a certain threshold deposited in them. The threshold is required to prevent signal noise from the detector (from radioactive material present in the detector and from electronics) contributing to the clusters. Cells adjacent to seed cells with energy deposits below this threshold may still be considered part of the cluster, however, thanks to the algorithm. This helps to ensure that electrons can be accurately reconstructed.

From the collections of seed cells (and the cells which immediately surround them), ‘towers’ are identified in $\eta - \phi$ space, which themselves are added together if they lie within a cone of $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$ of the tower with the highest energy. These towers form the clusters which must then be matched to a track if it is to be counted as an electron. If no track can be matched, the cluster is counted as a photon.

Photons may fake an electron signal by producing an electron-positron pair in the tracking system, while jets may do so by having a charged track with a π^0 decaying to two photons in the EM calorimeter. Hence, a number of variables are used to help distinguish true electrons from the various sources of fake ones. The CC and EC regions are treated separately for this, with the intervening region $1.1 < \|\eta_{det}\| < 1.5$ excluded. Key variables are listed below:

- The cluster must be matched spatially to a track. This is based on a χ^2 fit defined as:

$$\chi_{Spatial}^2 = \left(\frac{\Delta\phi}{\sigma_\phi} \right)^2 + \left(\frac{\Delta z}{\sigma_z} \right)^2 \quad (4.1)$$

where $\Delta\phi$ and Δz are the spatial coordinates centred on the cluster in the third EM layer.

- E_T/p_T , as for high p_T electrons, their total energy will be dominated by their momenta, hence E_T/p_T should be close to 1
- Seven or eight variables (depending on whether the cluster is in the CC or the EC) relating to the shape of the cluster ‘shower’ are combined to form a likelihood fit called the ‘H-matrix’. The χ^2 of the fit should be lower for electrons than for other objects.
- Since electrons should deposit most of their energy in the EM calorimeter, f_{EM} should be large (that is, close to 1), where

$$f_{EM} = \frac{E_{EM}}{E_{Tot}} \quad (4.2)$$

and E_{EM} is the amount of energy of the cluster which is deposited in the EM calorimeter while E_{Tot} is the energy of the cluster as a whole. In contrast, jets would have a lower value of f_{EM} as most of their energy will be deposited in the hadronic calorimeter layers.

- z , the distance from the track to the PV
- The number of tracks in a cone of $R = 0.05$ around the matched track, with vertices within 2 cm of the candidate track’s vertex (as the track should be relatively isolated just as the cluster is)
- Σp_T for tracks within a cone of $R = 0.4$ around the matched track, but not

including the matched track itself.

- Electrons should be relatively isolated from other activity in the calorimeter, so an isolation fraction is defined as:

$$f_{Iso} = \frac{E_{Tot} - E_{EM}}{E_{EM}} \quad (4.3)$$

where E_{Tot} is confined to a cone of $R = 0.4$ and E_{EM} to a cone of $R = 0.2$ when considering track isolation.

These variables are ultimately combined, along with others [31] in a Boosted Decision Tree (BDT), a TMVA package from the ROOT [32] framework used for increasing signal-to-background ratio (see section 9 for more details) in regions of the final BDT distribution, allowing for cuts to be placed which reject events which do not meet the requirements. The output for the BDT can be seen in Figures 4.1(a) - 4.1(d), with a samples of real electrons shown in red and fake electrons in blue.

The BDT is used to positively identify electrons, and the operating points are identified by the suffix `_eff`. It is also used to identify fake electrons, and for this the suffix `_fake` is used. Furthermore, for both cases we wish to identify ‘loose’ and ‘tight’ electrons for the purposes of estimating the number of QCD events (see section 7 for further details). As such, we use the operating point `emvPoint05_eff` and `emvPoint1_fake` for loose electrons and `emvPoint1_eff` and `emvPoint2_fake` for tight electrons. The analysis in this thesis uses the cuts on values of the BDT which are given in Table 1, with these values needing to be exceeded to pass the cut.

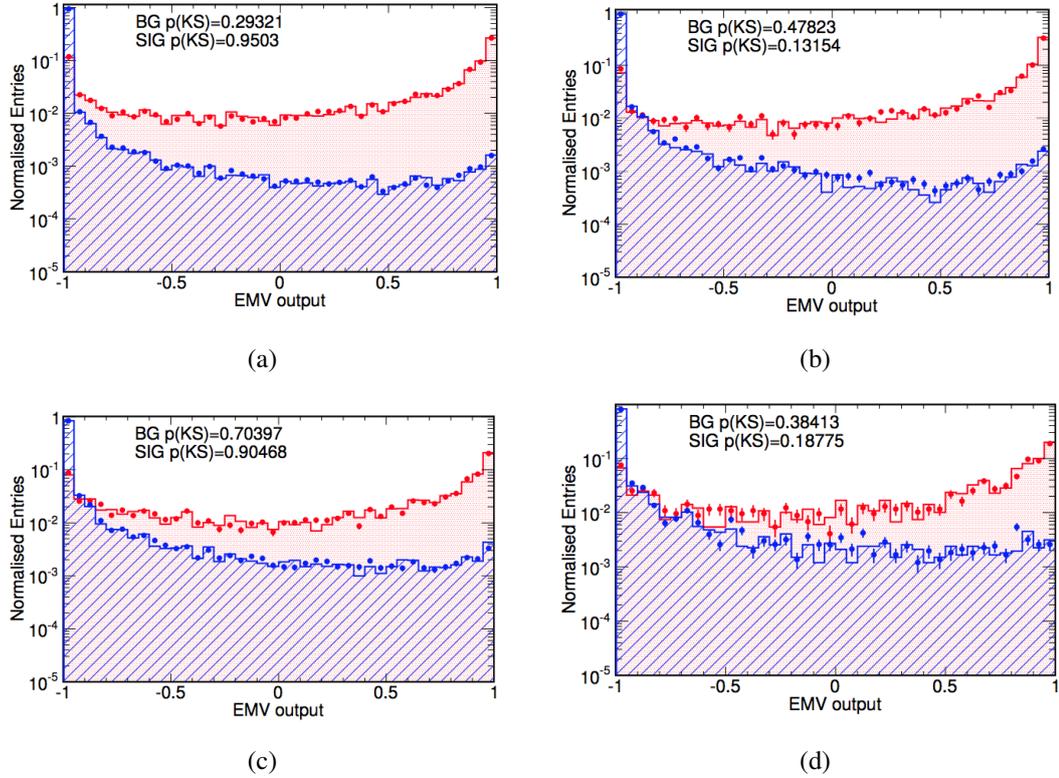


Figure 4.1: BDT output for (a) CC electrons in low luminosity events, (b) CC electrons in high luminosity events, (c) EC electrons in low luminosity events and (d) EC electrons in high luminosity events. In each figure, the solid red distribution is from real electrons and blue from fake electrons, with the barred points from a testing sample [31].

4.2.2 Muons

Muons are identified by making hits in the muon system (as other particles should very rarely leave the calorimeter without having deposited all of their energy) and by identifying associated tracks in the central tracking regions. Tracks in the muon system are reconstructed from hits in the scintillators and layer wires of each of the three layers separately, which are then matched to form track candidates. If these can be matched with a track in the central tracking system, it is called a central track matched muon. Information about the number of hits, the number of layers in the muon system which contain hits and the track matched in the central tracker is used to assign the reconstructed muon into one of several categories which determine its quality [33]. As with

Operating point		
	Value	
Name	CC	EC
emvPoint05_eff	-0.96	-0.86
emvPoint1_eff	-0.22	-0.74
emvPoint1_fake	-0.86	-0.86
emvPoint2_fake	-0.32	-0.02

Table 1: EMID operating point values for real electrons and fake electrons in the CC and EC

the electron identification process, loose and tight muons are identified for estimating the multijet background later on.

All muons coming from the hard scatter should have a high p_T , so a criteria of $p_T > 12$ GeV is required. In addition to these criteria, a loose muon satisfies the following:

- Tracks must be isolated, as non-isolated tracks are more likely to come from the decay of heavy bound hadrons. Hence, $\mathcal{I}_{trk} < 2.5$ GeV, where $\mathcal{I}_{trk} = \sum_{r_i > 0.1}^{r_i < 0.4} p_T^i$ and r_i gives the distance from the track in $\eta \times \phi$.
- Similarly, the muon must be isolated in the calorimeter; $\mathcal{I}_{cal} < 2.5$ GeV, where $\mathcal{I}_{cal} = \sum_{r_i > 0.1}^{r_i < 0.4} E_T^i$
- There is at least one hit in the SMT.
- The χ^2 fit of the track is less than 4.
- Distance of closest approach of the track to the collision point in the transverse plane is less than 2cm.

Tight muons are then identified using the following criteria:

- $\text{Halo}(0.1, 0.4)/p_T(\mu) < 0.4$, where $\text{Halo}(0.1, 0.4)$ is the scalar sum of transverse calorimeter energy clusters (excluding the coarse hadronic calorimeter as its towers encounter significantly greater noise) in a hollow cone around the muon within a spatial window of $0.1 < \Delta R < 0.4$.

- $\text{TrkCone}(0.5)/pT(\mu) < 0.12$, where $\text{TrkCone}(0.5)$ is the scalar sum of the transverse momentum of all tracks within a cone of radius $\Delta R = 0.5$ around the muon and the track matched to the muon is excluded from the summation.

Scale factors are applied to the simulation based on these identification, track matching and isolation criteria.

4.3 Missing Transverse Energy

When neutrinos are produced in an event, they do not leave a signature in the detector as they interact so weakly with matter. As such, their presence in an event must be inferred from the imbalance in the sum of transverse energy of the objects in the data [36]. Since the partons involved in the hard scattering process have no (or negligible) initial energy in the transverse plane, their collision products cannot have any either.

The transverse energy deposited in the calorimeter cells is added vectorially (except for contributions from the coarse hadronic calorimeter as this can suffer from large amounts of noise). Corrections for muon E_T (which do not deposit their energy in the calorimeter) and detector effects are applied, leaving a value that is taken to be the ‘missing’ transverse energy, \cancel{E}_T .

4.4 Jet Reconstruction

4.4.1 Jets

Hadronization of quarks produced in the hard interaction of an event, as well as soft gluon radiation, lead to the formation of ‘jets’ in the detector. These are characterised by large amounts of energy deposited in the hadronic calorimeter. This is shown schematically in Fig. 4.2. Jets with higher energy will deposit their energy in more

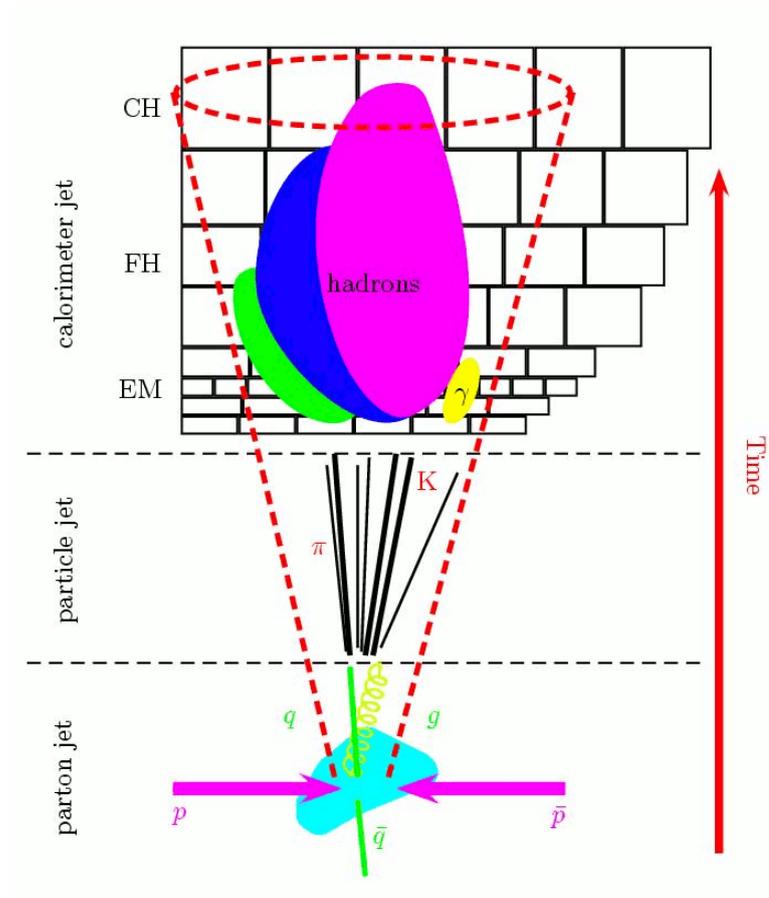


Figure 4.2: A jet showering in the detector [35]

collimated showers than jets of lower energy.

At D0, jets are reconstructed using the jet cone algorithm with $\Delta R = 0.5$ [34].

4.4.2 b-jet Identification

B-jet identification, or ‘*b*-tagging’, identifies those jets that are most likely to have come from the hadronization of *b* quarks. In order to do this, a number of variables are identified which exhibit differences between jets which are *b*-like and those which are not, such as the distance between the primary and secondary vertices which, as mentioned earlier, can indicate that a hadron with a long life time was produced in the

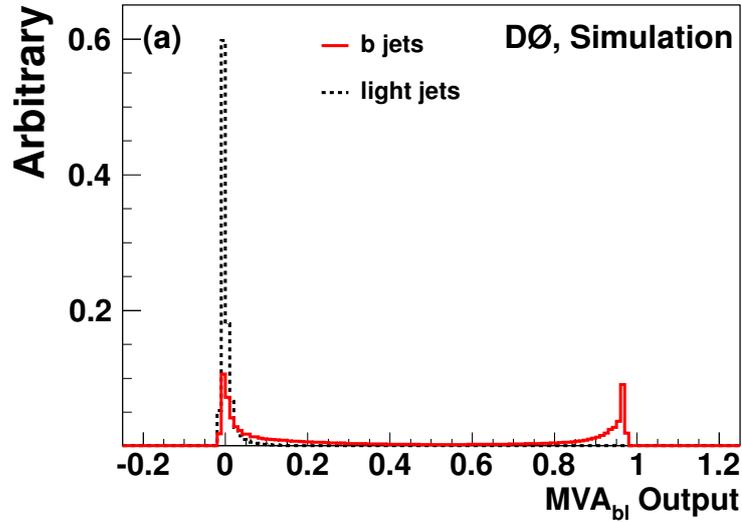


Figure 4.3: MVA output [38]

event and subsequently decayed - a characteristic property of a B meson. A multivariate analysis (MVA) [37] is then used to exploit such differences in order to construct a continuous distribution in which those jets which are b -like are close to the upper limit of the distribution at 1, while the light jets are close to the lower limit at 0. Figure 4.3 shows the output of the MVA for b -, c - and light-jets. Some of the variables used as inputs to this MVA are discussed below.

4.5 Secondary Vertex Mass

The secondary vertex mass, M_{svt} , is the invariant mass constructed from the momenta of all tracks associated with the most significant secondary vertex. B -jets tend to have a higher M_{svt} value. Figure 4.4 shows the M_{svt} distributions for b -, c - and light-jets.

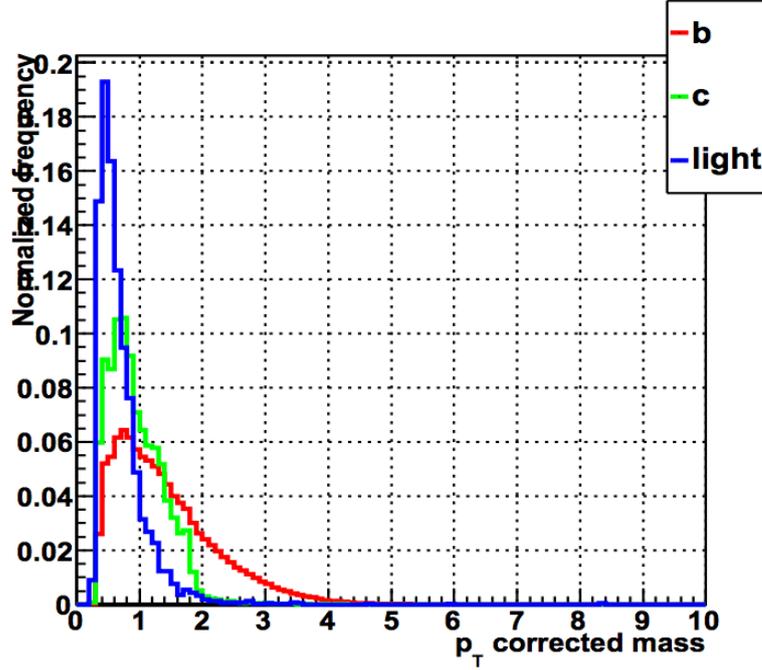


Figure 4.4: The M_{svt} distributions for b -, c - and light-jets [38]

4.6 Jet Lifetime Probability Tagger

The jet lifetime probability tagger [39] (JLIP) uses the impact parameter (IP) of tracks associated with a calorimeter jet in order to calculate the probability that the jet originated at the PV. IP are considered to be positive or negative depending on whether the extrapolation of the tracks which they are associated with cross the axis of the jet they have been matched to before (positive) the PV or after (negative). Any tracks with a positive IP are then assigned a probability, P , which quantifies the likelihood of that track having originated at the PV. Negative IP values were not assigned a probability at the time that the analysis documented in this thesis was undertaken due to modelling difficulties. These were included in a later tagging algorithm, which achieved a 5% gain in b -jet identification efficiency, and may be a source of improvement in future studies. A JLIP value is calculated for each jet using the following equation:

$$JLIP = \prod_{i=0}^{N=Tracks} P^i \quad (4.4)$$

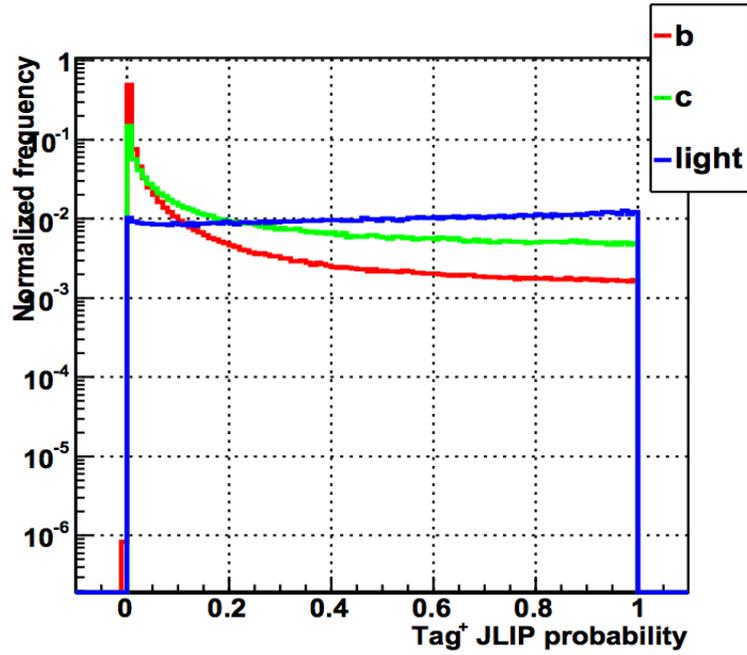


Figure 4.5: The JLIP probability [38]

Figure 4.5 shows the $-\ln(\text{JLIP})$ distributions for b -, c - and light-jets.

4.7 The Multivariate Tagging Method

In addition to the M_{svt} and JLIP variables discussed above, the MVA makes use of a further 7 input variables, which are as follows:

- The reduced JLIP (rJLIP). This is the value that the JLIP would take if the track which is least likely to have originated from the PV is removed.
- The number of tracks that are used in the SV reconstruction.
- The number of SV.
- The value of χ^2/NDF in the SV fit
- The decay length significance in the transverse plane with respect to the PV, S_{xy} .

- ΔR in (η, ϕ) space between the jet axis and the vector constructed from the difference in PV and SV locations.
- A composite variable based on the number of tracks with an IP significance greater than some optimized value. The optimization (see [37] for further details on this variable; CSIP \mathcal{N}_{CSIP}).

Chapter 5

Data and Simulated Samples

5.1 Data

The data used in the analyses in this thesis were taken during the RunII data-taking period at D0. The RunII data is split into two epochs. The first is RunIIa, which was recorded up until the 2006 shutdown, while the second, RunIIb, was taken from 2006 until the final Tevatron run in 2011. RunIIb is further broken down into four run periods: RunIIb1, RunIIb2, RunIIb3 and RunIIb4. Due to differences in reconstruction algorithms, tracking efficiencies and other small changes across the run periods, different simulated samples were used for each of RunIIa, RunIIb1, RunIIb2 and RunIIb3+4. During the course of RunII, D0 recorded over 10fb^{-1} of data. After imposing quality cuts based on the performance of D0's subdetectors, the total integrated luminosity used for these analyses was 9.7fb^{-1} . The breakdown of this integrated luminosity across the run periods is given in Table 2, while the the development of the integrated luminosity over time is shown in Fig. 5.1.

Run Period	Integrated Luminosity (fb^{-1})
RunIIa	1.08
RunIIb1	1.22
RunIIb2	3.04
RunIIb3	1.99
RunIIb4	2.40

Table 2: Breakdown of integrated luminosity of RunII data after quality cuts

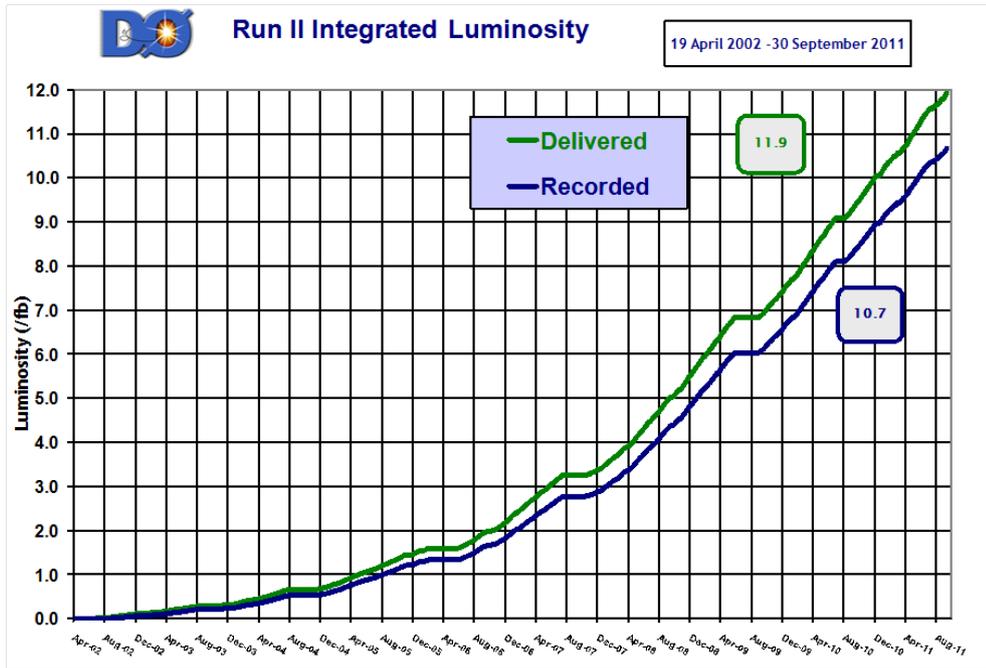


Figure 5.1: The total integrated luminosity delivered by the Tevatron accelerator (green) and recorded by the D0 detector (blue) before quality cuts.

5.2 Background Processes

In this analysis, most of the background processes were generated using the Monte Carlo (MC) simulated event generator method, with the exception of the multijet background (see section 6.2).

The $t\bar{t}$, $W + \text{jets}$ and $Z + \text{jets}$ samples were generated using the ALPGEN v2.3 [41] generator, which calculates the matrix element level of the processes. They are then interfaced with PYTHIA v6.409 [42] to model the subsequent parton showering and hadronization, using the leading-order parton distribution functions (PDFs) from

CTEQ6L1 [43, 44]. The cross section of the $t\bar{t}$ process is calculated to Next-to-Next-to-Leading Order (NNLO) level in α_S , while the $W + \text{jets}$ and $Z + \text{jets}$ cross sections are scaled to Next-to-Leading Order (NLO) [45] levels, using so-called k -factors, which are the ratio of the NLO and LO calculations.

Separate samples containing the $t\bar{t}g$ process (where a soft gluon is emitted in the $t\bar{t}$ process) with $g \rightarrow b\bar{b}$ were generated, again using ALPGEN with PYTHIA for showering. This represents a subset of the samples generated to simulate the $t\bar{t}$ process, but by processing them separately it allowed for a larger sample to be used for multivariate training (see Chapter 9). To avoid double counting such events, a filter was applied to the $t\bar{t}$ samples which removed any events which satisfied the same kinematic criteria with which the $t\bar{t}g$ samples were produced.

Additionally, samples were generated to simulate the $t\bar{t}Z$ process, where $Z \rightarrow b\bar{b}$. This only has a very small cross section at the centre-of-mass energy of the Tevatron [46], but had to be taken into account for the multivariate training as two extra b jets whose energies sum to ~ 100 GeV produced in association with a $t\bar{t}$ pair may look like the $t\bar{t}H$ signal in its kinematics. This process was simulated using MADGRAPH 5 [47] for the matrix elements and again interfaced with PYTHIA for the showering. This process was not a subset of the $t\bar{t}$ sample, and as such did not require a filter in the same manner as the $t\bar{t}g$ samples.

For the $W + \text{jets}$ and $Z + \text{jets}$ processes, samples containing only so-called light jets (that is, jets which arose from the hadronization of u , d or s quarks) and samples containing two heavy flavour jets (jets arising from the hadronization of c or b quarks) are generated separately. A skimming process was then employed to remove any events which could be double-counted across samples, so $W + b\bar{b}$ and $W + c\bar{c}$ events are removed from the $W + jj$ (where each j can be a jet from a u , d or s hadron) samples. Similarly, $W + c\bar{c}$ events are removed from the $W + b\bar{b}$ samples, while the $W + c\bar{c}$ are not skimmed. Scaling factors are applied to the $W + \text{jets}$ and $Z + \text{jets}$ processes,

m_H (GeV)	$\sigma(t\bar{t}H) \times BR(H \rightarrow b\bar{b})$ (fb)
90	8.27
95	7.25
100	6.33
105	5.47
110	4.66
115	3.88
120	3.15
125	2.47
130	1.86
135	1.34
140	0.92
145	0.60
150	0.36
155	0.19

Table 3: Cross section times branching ratios for $t\bar{t}H$ with $H \rightarrow b\bar{b}$ in the range $90 < m_H < 155$ GeV

which are discussed in 6.1.

Diboson events, which are those events with a pair of W bosons, a pair of Z bosons or one each of the two electroweak bosons (WZ), were simulated only with PYTHIA, while single-top background samples were generated using COMPEP [52]. The cross sections of these samples were also calculated at the NLO level.

5.3 Signal Processes

As with the simulated background processes, the simulated $t\bar{t}H$ signal samples were produced using the MC method. ALPGEN was used for the hard scattering process and PYTHIA for the showering. They were produced for a range of Higgs masses, from 90 to 155 GeV in 5 GeV intervals and have been normalised to the NLO QCD cross sections [48, 49].

The cross section times branching ratios for $t\bar{t}H$ with $H \rightarrow b\bar{b}$ are shown in Table 3.

All of the simulated samples, including those for the background, were processed through the D0 detector simulation `d0gstar` [50] (which is based on the detector material simulation package `GEANT` [51]), the D0 detector electronics simulation `d0sim`, and the reconstruction software `d0reco`. Due to several major upgrades, such as the installation of Layer 0 between RunIIa and RunIIb, different sets of simulated samples are used in this analysis, which also account for changes in the different algorithms used and for different calibrations.

Chapter 6

Data-driven Background

Determination

6.1 V+jets

Scaling factors are derived for the $W + \text{jets}$ and $Z + \text{jets}$ processes before the b -tagging stage of selection (see Chapter 7 and Section 7.2), further to the k -factors mentioned in the previous section, to account for the large uncertainties associated with their cross sections. All other backgrounds (apart from multijet, which is treated in the same way as the $W + \text{jets}$ and $Z + \text{jets}$; see section 6.2) are subtracted from the data in binned distributions of the reconstructed transverse W boson mass, m_T^W , and the $W + \text{jets}$ and $Z + \text{jets}$ are fitted as one sample to the remainder of the distribution using the ROOT [32] TFractionFitter, respecting their predicted relative proportions. Electron and muon samples are treated separately, and the fit is performed separately on each jet multiplicity bin as well, with the exception of the (at least) 6 jet bin which suffers from insufficient statistics to achieve a stable fit. Hence, the scaling factors derived from the 5 jet case are used instead. Separate scale factors are also derived for loose and tight lepton identification categories as well.

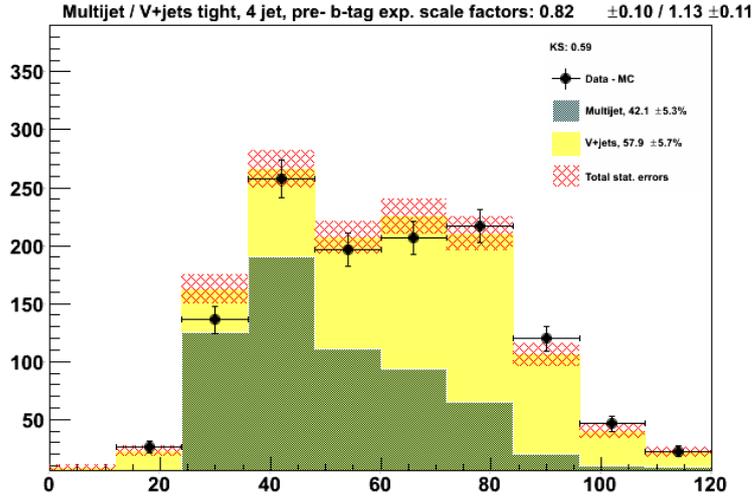


Figure 6.1: Binned distribution of m_T^W for determination of scale factors for multijet (green) and V +jets (yellow) when fitted to background subtracted data for events with 4 jets and 1 tight electron

Figure 6.1 shows the binned distributions for the 4 jet case in the electron channel for RunIIb3 + 4, while Figure 6.2 shows the derived scaling factors for the 4, 5 and (at least) 6 jet multiplicities (where the 6 jet matches the 5 jet scaling factor by design) for this run epoch and lepton. Both figures are for the tight lepton ID case.

6.2 Multijet

Multijet samples are derived from data events using a modified version of the Matrix Method. In the usual Matrix Method, the number of data events which contain a lepton which pass the loose identification criteria, N_L , and the number which pass the tight criteria, N_T , are given by the following equations:

$$\begin{aligned}
 N_L &= N_l + N_{mj}, \\
 N_T &= \epsilon_l N_l + \epsilon_{mj} N_{mj} = N'_l + N'_{mj}
 \end{aligned}
 \tag{6.1}$$

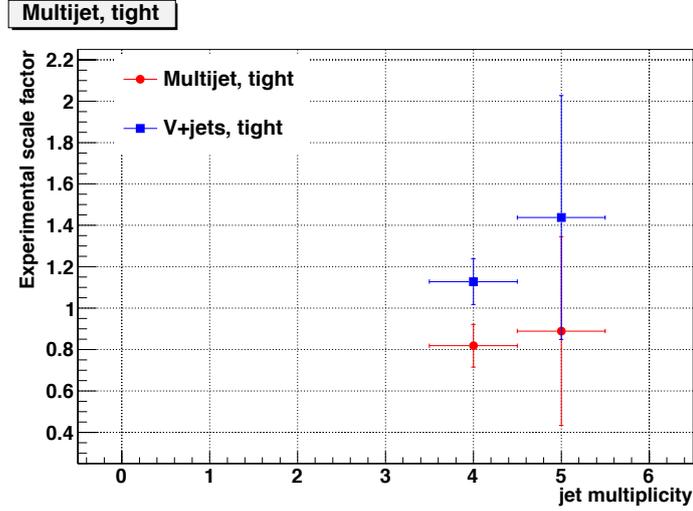


Figure 6.2: Scale factors derived from the fitting process for events with at least 4 jets and 1 tight electron. The scale factor derived for 5 jets is also used in the at least 6 jet case.

where N_l is the number of events containing a lepton which passes the loose selection criteria, N_{mj} is the number of events containing no leptons but with a multijet event being misidentified and passing the loose selection criteria, ϵ_l the efficiency of a lepton which passes the loose selection to then pass the tight selection criteria, and ϵ_{mj} the same but for a misidentified multijet event.

Rearranging this to find N'_{mj} , the number of multijet events expected in the tight lepton samples, gives:

$$N'_{mj} = \frac{\epsilon_{mj}}{\epsilon_l - \epsilon_{mj}} (\epsilon_l N_L - N_T) \quad (6.2)$$

In the usual matrix method, weights which are parameterized by event kinematics k_i are derived from this equation thus:

$$w_i = \frac{\epsilon_{mj}(k_i)}{\epsilon_l(k_i) - \epsilon_{mj}(k_i)} (\epsilon_l(k_i) - \Theta_i^T) \quad (6.3)$$

where $\Theta_i^T = 1$ if the event satisfies the tight lepton criteria and 0 otherwise. These weights are then applied to each event such that:

$$N'_{mj} = \sum_{i=1}^{N_L} w_i \quad (6.4)$$

However, in accordance with recommendations from studies performed [53], a slightly different method is employed so that the estimate of the multijet yield is not so closely correlated with data events, and so that events containing a lepton passing the tight selection do not contribute a negative weight due to the Θ_i^T term. Hence, only loose events which do not satisfy the tight selection are used, and designated ‘loose-not-tight’. In this case,

$$w_i = \frac{\epsilon_{mj}}{1 - \epsilon_{mj}} \quad (6.5)$$

and

$$N'_{mj} = \sum_{i=1}^{N_{L-n-T}} w_i \quad (6.6)$$

where N_{L-n-T} is the number of these loose-not-tight events. However, this sample may include real leptons, and as such a degree of double-counting may occur. To avoid this, the W or Z + jets events (collectively V + jets) are reweighted by w_l , where

$$w_l = 1 - \frac{\epsilon_{mj}(1 - \epsilon_l)}{(1 - \epsilon_{mj})\epsilon_l} \quad (6.7)$$

Further scaling factors are applied to the multijet yields, calculated in the same way as those which are applied to the V + jets samples, to improve the accuracy of the yield estimate. Again, these scaling factors are calculated only for the cases where there are exactly 4 or 5 jets in the event, with the 5 jet scaling factors applied to the (at least) 6 jet yields.

Chapter 7

Event Selection

Selection criteria are applied to the dataset and simulated processes to increase the signal-to-background ratio. That is, to cut away those events in which are unlikely to allow the signal process to be seen, either because it has not occurred (for instance if there are no leptons identified in the event) or because events containing other processes which occur at a much higher rate than the signal would outnumber events which are signal-like and thus worsening the statistical likelihood that a signal-like process has been observed.

A single lepton with high transverse momentum is required, as well as large missing transverse energy in the event as these criteria characterise a leptonically decaying W boson. Hence, a p_T cut of 15 GeV is required for the lepton, which helps to reduce the multijet background. The leptons must also be within the acceptance of the appropriate detector, which for the electron means the calorimeter, so either $|\eta| < 1.1$ for the CC or $1.5 < |\eta| < 2.5$ for the EC is required. As the muon system only extends to a value of 2 in η , the muon must also be within this range. \cancel{E}_T is required to be greater than 15 GeV if the lepton is identified as an electron and 20 GeV in the case of a muon, in part due to the signature of a neutrino from a leptonically decaying W , but also to ensure that there is no overlap with the selection criteria used for the data-driven multijet estimation. The object reconstructed from the lepton and \cancel{E}_T seen in the detector is

expected to be the W boson, and a transverse mass cut is placed on this object at the same time as the \cancel{E}_T itself to further suppress multijet background.

High p_T jets are also required in the event, due to the hadronically decaying W boson and the two b -jets from the $t\bar{t}$ decay, as well as the two b -jets we expect from the decay of the Higgs boson. Two high p_T jets are required in the trigger, and a further two are required in the selection, with further categories defined by the number of jets in the event. A p_T cut of 20 GeV is imposed on the jets, with a detector acceptance of $\eta < 2.5$.

The selection criteria are summarised below:

- Exactly one lepton (e or μ) with $p_T \geq 15$ GeV within $|\eta| < 1.1$ or $1.5 < |\eta| < 2.5$ in the case of electron or $|\eta| < 2$ for muon
- $\cancel{E}_T > 15$ (20) GeV in the case of the electron (muon) channel
- Exactly 4, 5 or ≥ 6 jets with $p_T > 20$ GeV within $\eta < 2.5$. For RunIIb epochs, these must also be vertex confirmed.
- Triangular cut $W_{M_T} > 40 - 0.5 \times \cancel{E}_T$ GeV

7.1 Validation Plots

Included here are plots of a number of kinematic variables which should be well-understood for the type of analysis performed for this thesis. Plots for electrons and muons are shown separately and for the following variables:

- Leading jet p_T
- Second-leading jet p_T
- Lepton p_T

- Missing transverse energy

Plots are also separated according to the jet categories of exactly 4, 5 and at least 6 jets, before tagging criteria (see Section 7.2) are imposed ('pre-tag'). In each plot, the signal is scaled by a factor as indicated in the legend and overlaid (solid red histogram). Figure 7.1 shows a larger version of the legend with a signal scaling factor of 200.

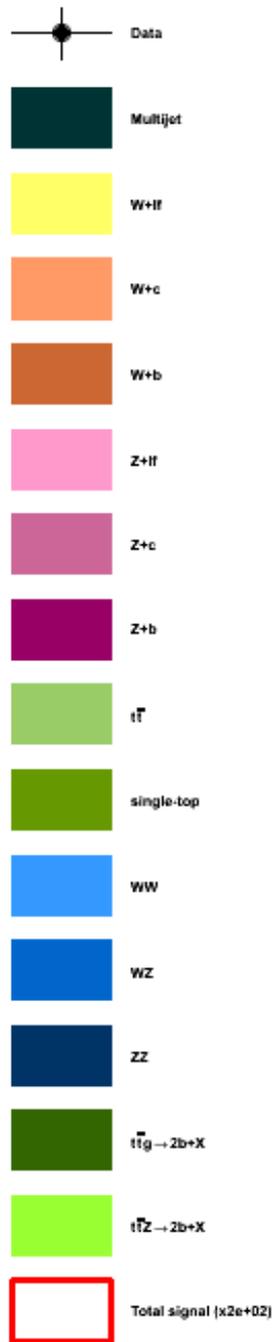


Figure 7.1: Legend for validation plots.

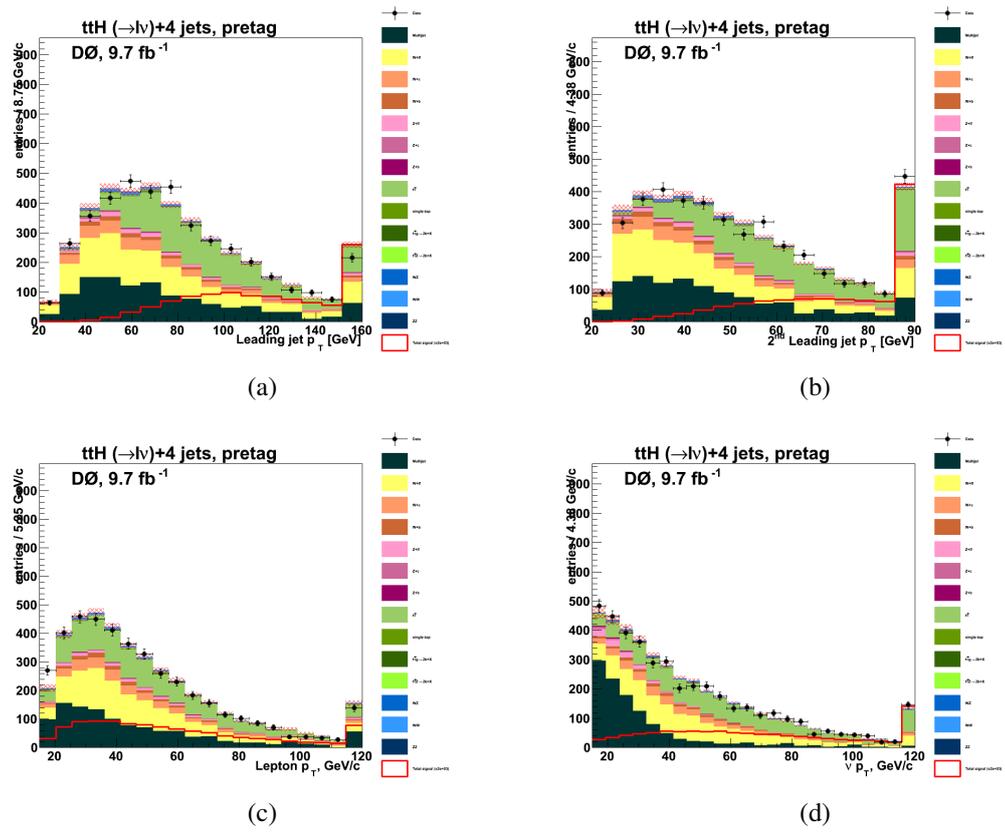


Figure 7.2: Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 4 jets and 1 electron.

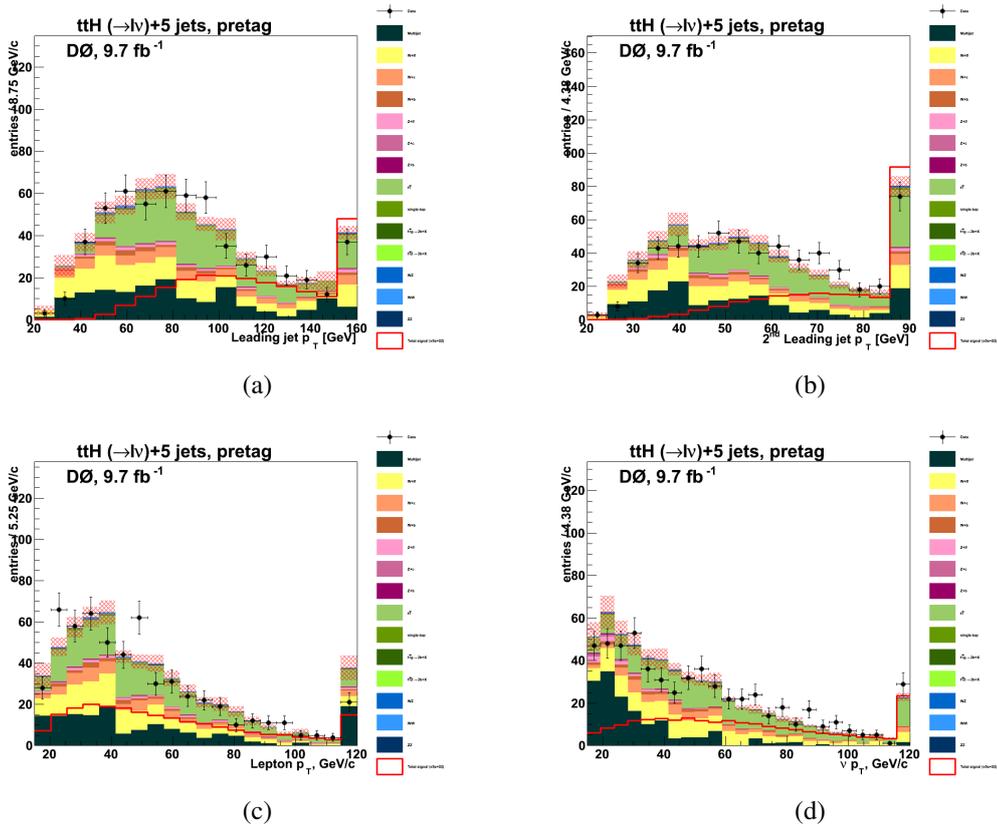


Figure 7.3: Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 5 jets and 1 electron.

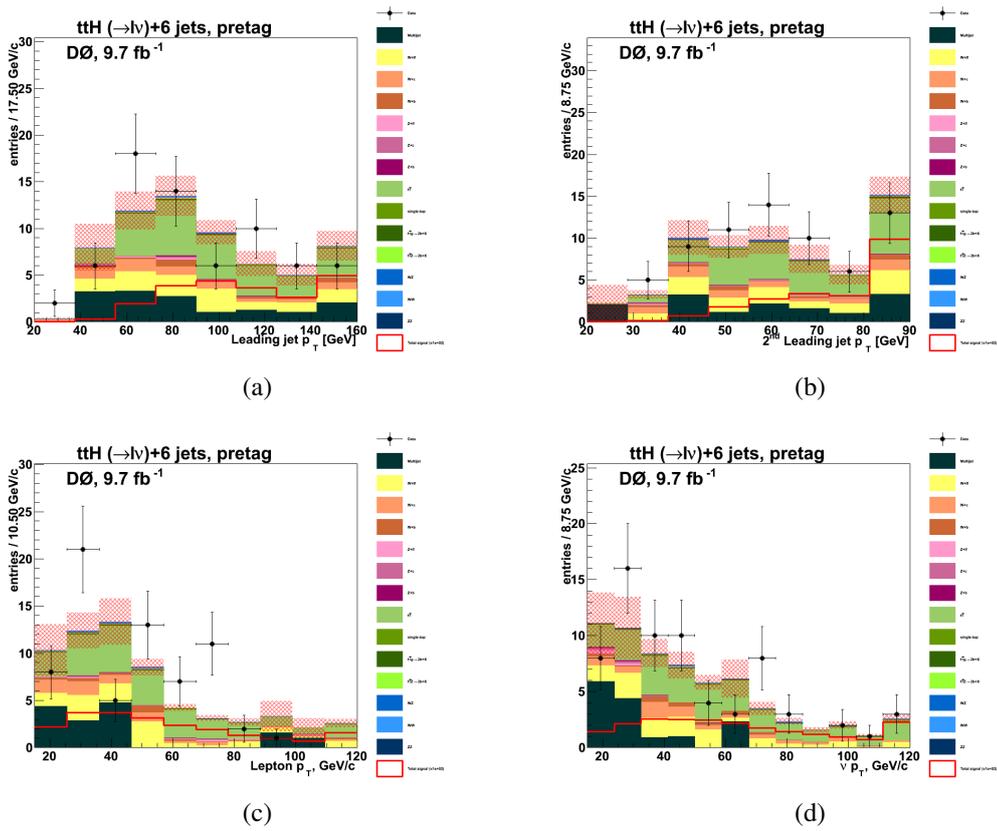


Figure 7.4: Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with at least 6 jets and 1 electron.

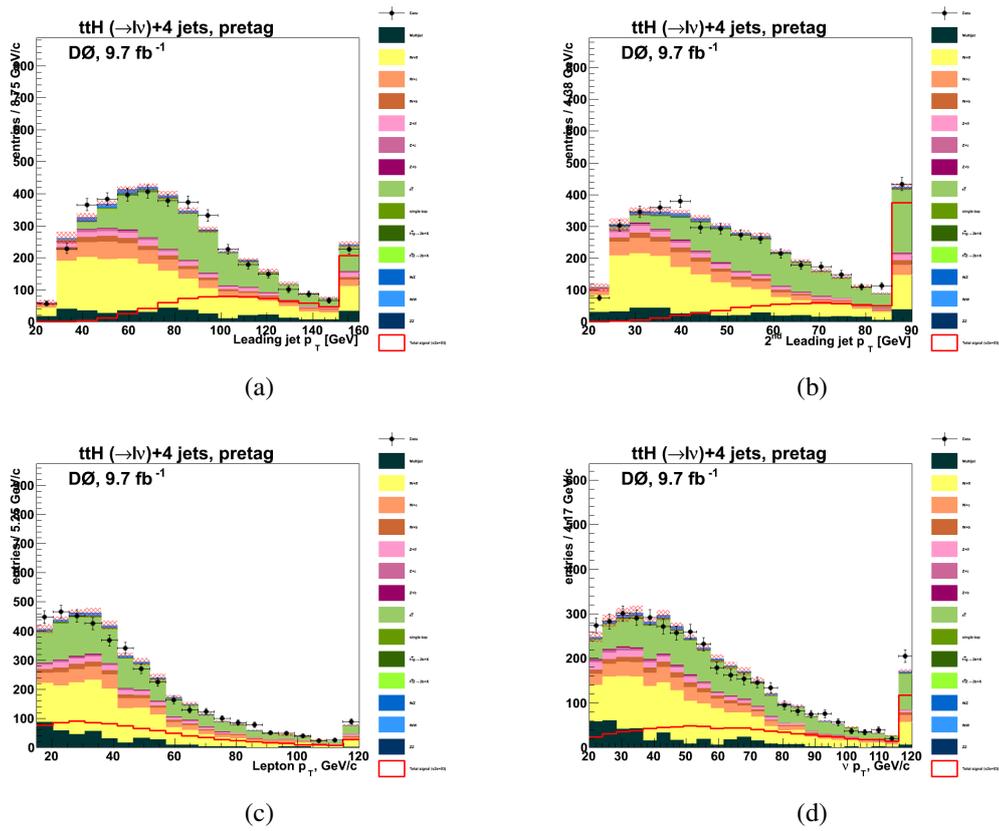


Figure 7.5: Pre-tag validation plots for (a) transverse momentum of the leading jet, (b) transverse momentum of the second-leading jet, (c) transverse momentum of the lepton and (d) missing transverse energy, for events with exactly 4 jets and 1 muon.

7.2 B-tagging Selection

We employ a direct tagging method for identifying jets coming from decaying b -hadrons in signal, data and background samples. The exception to this is in multijet where we use a tag rate function (TRF) method. This is due to the low statistics associated with requiring 5 or more jets, preventing our fitting procedure from stabilising.

Further to the selection criteria outlined already, identification of b -jets is used to further refine the selection. We define orthogonal tagging samples, with zero, 1, 2 or at least 3 jets passing at least the most loose b -tagging operating point, which is designated L6, for which the MVA output discussed in Section 4.4.2 has a value greater than 0.02. The MVA output will hereafter be referred to as BL (or BL_j for the BL value of a single jet j). The 1- and 2-tag categories are further divided into two and three subcategories respectively to give seven in total, which are summarised in Table 4 with their respective tagging criteria. In the case of 2 tagged jets, the mean BL is taken as this allows the definition of the 2 medium tag category with some signal-to-background improvement over imposing cuts on the BL of individual jets, as shown in Figure 7.8 [54]. Applying cuts to the BL of individual jets results in a flat distribution for both signal and background in the region $0.35 \leq BL < 0.55$, and so by taking the mean an additional tagging category with reasonably high signal-to-background ratio can be used in the analysis.

If an event satisfies multiple tagging criteria, the event is placed in the sample containing the most tags. That is, 2 loose is preferred over 1 tight and 3 tags preferred over any of the 2 tag categories. In order to subdivide the samples in this way, it is essential that the MVA b -tagger output is well understood, which is achieved by introducing a continuous tagging approach using updated TRFs from the b-ID group [55]. Even though seven categories are defined, only 5 are used in the analysis (indicated by a (*) symbol in Table 4). These 5 categories were chosen so as to ensure orthogonality

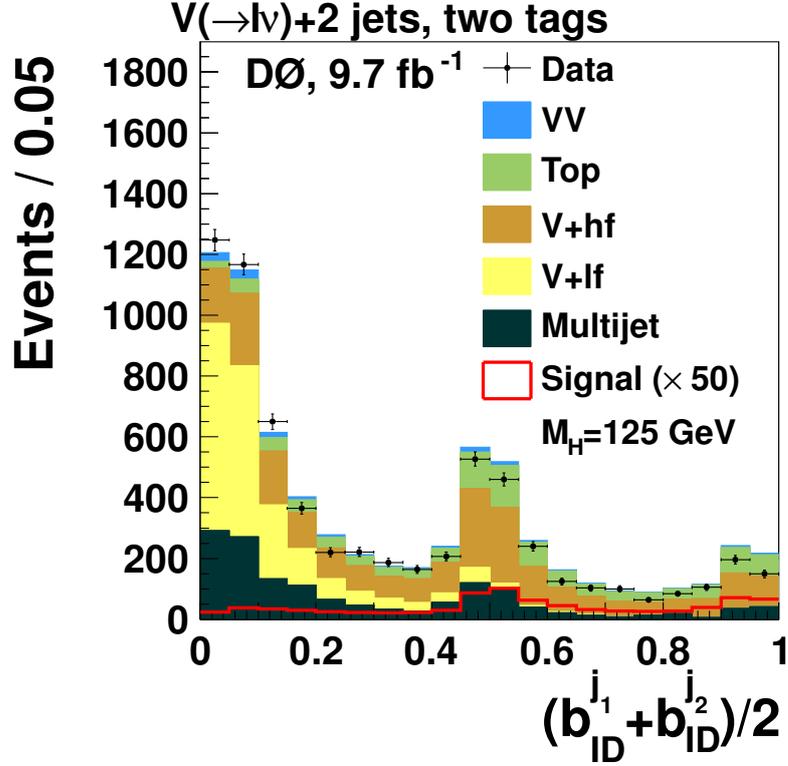


Figure 7.8: Average b -ID discriminant output of two tagged jets. This plot is taken from a search for associated production of Higgs with a vector boson [54], and as such the signal shown in this plot reflects that search rather than the one documented in this thesis.

between this analysis and the D0 searches for associated production of Higgs with a vector boson [54] (when jet categories are also taken into account). The 0 tag and 1 loose tag categories are included as a check, performed to ensure that unitarity is preserved after applying b -tagging;

$$N_{pretag} = N_{0-tag} + N_{1loose} + N_{1tight} + N_{2loose} + N_{2medium} + N_{2tight} + N_{3-tag}$$

This is found to be the case to within 0.5% in MC, within the uncertainty of the b -tagging scale factors. If no scaling factors were applied, this sum would balance exactly.

Further validation plots are shown here for the leading jet p_T and lepton p_T variables in each of the tagging and jet categories. Validation plots for the remaining variables from section 7.1 can be found in Appendix A

Tagging Category Name	<i>b</i> -ID discriminant cut
0 tag	All jets having $BL_j < 0.02$
1 loose tag	Exactly one jet with $0.02 < BL_j < 0.15$
1 tight tag (*)	Exactly one jet with $BL_j > 0.15$
2 loose tag (*)	Exactly two jets, with $0.02 \leq (BL_{j1} + BL_{j2})/2 < 0.35$
2 medium tag (*)	Exactly two jets, with $0.35 \leq (BL_{j1} + BL_{j2})/2 < 0.55$
2 tight tag (*)	Exactly two jets, with $(BL_{j1} + BL_{j2})/2 > 0.55$
3 tag (*)	3 or more jets with $BL_j > 0.02$

Table 4: *b*-tagging category names and operating points. Note that for the 2 tag categories the mean *BL* for the two jets is taken.

It is shown that the background MC processes describe the distribution of the data well enough to proceed with the analysis based on these selection criteria.

Chapter 8

Search for $t\bar{t}H$ production

Since the Higgs boson is now known to be less than the mass of the top quark, the Higgs boson cannot decay into a $t\bar{t}$ pair. Hence, we must look to the $t\bar{t}H$ process to infer the Yukawa coupling, which is expected to be ~ 1 as it is the heaviest (known) quark, so is expected to have the strongest coupling with the Higgs.

For this analysis, we measure the cross section times branching ratio of the $t\bar{t}H$ process in a lepton plus jets final state, assuming that the Higgs decays to a $b\bar{b}$ pair 100% of the time. The sample is split into events containing 4, 5 and ≥ 6 jets, and the case of the lepton being identified as an electron or muon are treated separately.

8.0.1 Separation of signal from background

Studies were performed to compare the event kinematics of W +jets, multijet and $t\bar{t}$ background events with the $t\bar{t}H$ signal in order to find variables with discrimination power between signal and background. For this purpose, samples were generated containing events with at least 4 jets. Figs. 8.1 and 8.2 show distributions of H_T (the sum of the transverse momentum of the event), the number of jets and the number of b -tagged jets for W +jets, multijet, $t\bar{t}$ and $t\bar{t}H$ simulated events. These variables give the

greatest separation power between signal and background - a better separation power than the invariant dijet masses, which has been shown to be the most discriminating variable in the VH analyses [12]. We define subsamples with 4, 5 or ≥ 5 jets, with 1 tight b -tag, 2 loose b -tags, 2 medium b -tags, 2 tight b -tags or ≥ 3 b -tags, for e +jets and μ +jets final states, and utilise the H_T distribution for each of those 30 channels. Although the $t\bar{t}H$ contribution is small for events with 1 or 2 b -tags these bins help to constrain the $t\bar{t}$ background and thus improve the sensitivity by $\approx 15\%$.

Figures 8.4-8.9 show the H_T distributions for all channels with 4, 5 or ≥ 6 jets and 1 tight, 2 loose, 2 medium, 2 tight or ≥ 3 b -tags for the full data set of 9.7 fb^{-1} . The data are compared to the different sources of background. The contribution of the $t\bar{t}H$ signal for a Higgs boson mass of 125 GeV is multiplied by a factor as indicated in the legend and overlaid (solid red histogram). A larger version of the legend is shown in Figure 8.3 for the 4 jet, 1 tight tag, electron case, with a signal scaling factor of 1000. In each plot, the $t\bar{t}$ cross section is normalized to 7.46 pb corresponding to a top quark mass of 172.5 GeV. The $t\bar{t}H$ signal is for a Higgs mass of 125 GeV and $\sigma(t\bar{t}H) \times B(H \rightarrow b\bar{b})$ to 4.28 fb.

8.0.2 Event yields and Limit setting

The expected and observed numbers of events in the different channels are summarized in Tabs. 5-7. The yields are shown for a SM Higgs boson of mass 125 GeV. The $t\bar{t}$ contribution is calculated for a theoretical $t\bar{t}$ cross section of $\sigma_{t\bar{t}} = 7.46$ pb [56] for a top quark mass of 172.5 GeV which is compatible with the world average value [57].

	e+jets				
	4j1tt	4j2lt	4j2mt	4j2tt	4j3t
Signal	0.072	0.049	0.063	0.084	0.17
$t\bar{t}$ +single top	290	133	169	219	167
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.68	0.45	0.52	0.69	1.0
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.02
WW	8.0	5.0	1.0	0.47	0.55
VZ	2.1	0.91	0.61	0.53	0.41
V+HF jets	94	46	22	13	11
V+LF jets	55	52	3.2	0.19	5.5
multijets	106	77	27	20	22
sum Bkg	556 ± 24	314 ± 18	223 ± 15	253 ± 16	207 ± 14
Observed	592	347	231	216	202

	μ +jets				
	4j1tt	4j2lt	4j2mt	4j2tt	4j3t
Signal	0.060	0.041	0.050	0.070	0.14
$t\bar{t}$ +single top	290	133	170	228	164
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.63	0.37	0.46	0.64	0.97
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.02
WW	8.2	4.5	1.7	0.37	0.79
VZ	2.1	1.4	0.60	0.54	0.40
V+HF jets	122	57	27	15	18
V+LF jets	71	66	2.8	0.39	6.0
multijets	59	28	17	10	10
sum Bkg	554 ± 24	291 ± 17	219 ± 15	256 ± 16	200 ± 14
Observed	571	315	229	255	231

Table 5: Summary of expected and observed yields in the various channels from the 4 jet 1 tight b -tag bin (4j1tt) to the 4 jet ≥ 3 b -tag bin (4j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.

In all channels the number of candidate events is consistent with the background

expectation within the statistical and systematical uncertainties. This is interesting because for the first time the ≥ 6 -jet channels and ≥ 3 b -tag channels are analysed. Since there is no evidence for SM $t\bar{t}H$ production with the current statistics, we set 95% C.L. limits on the production cross section times branching ratio $\sigma(t\bar{t}H) \times B(H \rightarrow b\bar{b})$. As input for the limit calculation we use the H_T distributions using the binning as displayed in Figs. 8.4-8.9.

To set limits on the SM Higgs boson production cross section, a modified frequentist approach [58] is used, where the signal confidence level CL_s , defined as the ratio of the confidence level for the signal-plus-background hypothesis to the background-only hypothesis ($CL_s = CL_{s+b}/CL_b$), is calculated by integration of the distributions of a test statistic over the outcomes of pseudo-experiments, generated according to Poisson statistics, for the signal+background and background-only hypotheses. The test statistic is calculated as a joint log-likelihood ratio (LLR) obtained by summing LLR values over the bins of the H_T distributions. Systematic uncertainties are incorporated via Gaussian smearing of the Poisson probability distributions for signal and backgrounds within the pseudo-experiments. All correlations between signal and backgrounds are maintained. To reduce the impact of systematic uncertainties on the sensitivity of the analysis, the individual signal and background contributions are fitted to the data (and pseudo-data). This is done for both the signal-plus-background and the background-only hypotheses independently by maximizing a profile likelihood function for each hypothesis [59]. The profile likelihood is constructed via a joint Poisson probability over the number of bins in the calculation and is a function of the nuisance parameters in the system and their uncertainties, which are given an additional Gaussian constraint associated with their prior predictions. Apart from systematics we use the SM $t\bar{t}$ cross section as a nuisance parameter taking the uncertainty as a Gaussian prior. The maximization of the likelihood function is performed over the nuisance parameters.

8.0.3 Systematic uncertainties

The main uncertainties only change event yields, rather than the H_T distribution shapes. These uncertainties are in the jet energy scale, jet energy resolution, lepton identification, luminosity, b -tagging and W , $\sigma_{t\bar{t}}$ and $t\bar{t}b\bar{b}$ background models. Another uncertainty on the event preselection is caused by the primary vertex selection and data quality requirements. All of these are summarized in Table 8.

The uncertainties on the jet energy scale and b -tag probabilities for light, c , and b -quark jets were treated initially as shape dependent uncertainties. To do so, we vary these functions, determined from data, by plus or minus one standard deviation from their central values to find the modifications to the shape of the H_T distributions. However, it was seen that these shape modifications were driven by the low systematics in the event, and thus were included as flat systematics instead.

8.0.4 Results

Fig. 8.10 shows the ratio of the $\sigma_{t\bar{t}H}$ cross section times branching ratio limit over the SM NLO prediction (left plot). The observed limit is shown in red with the expected limit, defined as the median of the limits obtained in background-only pseudo experiments, as the black dashed line. For a 125 GeV Higgs mass, the expected and observed limits on the $t\bar{t}H$ cross section times branching fraction for $H \rightarrow b\bar{b}$ are 24.7 and 74.3 times larger than the SM value, respectively. Table 9 gives the numerical values of the expected and observed limits for different Higgs masses. Compared to the preliminary result [60], the limits have improved as expected from the increase of integrated luminosity from 2.1 fb^{-1} to 9.7 fb^{-1} .

An excess is seen across the mass range, which has been investigated. It is robust, and is deemed to be a feature in data from the muon channel, as can be seen from Figures B.1(a) and B.1(b) in the Appendix section. It is also seen in similar analyses

at D0.

Searches for SM $t\bar{t}H$ production have also been performed by the CMS collaboration at the LHC [61] and by the CDF collaboration at the Tevatron [62]. Both experiments are able to set lower limits on $t\bar{t}H$ production than could be achieved in this analysis. At a Higgs mass of 125 GeV, CMS obtained an observed (expected) limit of 5.8 (5.2) times the Standard Model expectation using 5.0 fb^{-1} (5.1 fb^{-1}) of data at a collision energy of 7 TeV (8 TeV), while the CDF observed (expected) limit was 20.5 (12.6) using 9.45 fb^{-1} of data collected at the Tevatron. Both analyses used a form of Neural Network (NN), a type of MVA. While direct comparisons with the CMS experiment's results are difficult due to the higher collision energy (which entails a greatly enhanced $t\bar{t}H$ production cross section of approximately 85 fb at 7 TeV [63], though with $t\bar{t}$ background also greatly enhanced), a more straightforward comparison can be made with the limits obtained by the CDF collaboration. Both their expected and observed limits were approximately a factor of 2 lower than that which was achieved in this analysis, although the reason(s) for this are not clear. The use of two separate tagging algorithms offers a higher tagging efficiency than one alone [62], which may enhance signal acceptance in categories with higher tagging multiplicity, thus improving the signal-to-background ratio. The use of an MVA may also help to improve sensitivity. An MVA method was used for this analysis and is documented in Chapter 9.

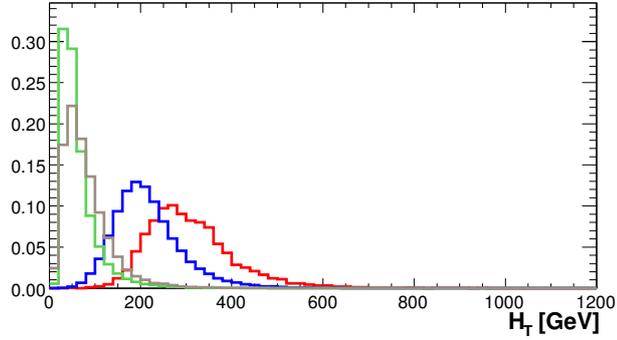
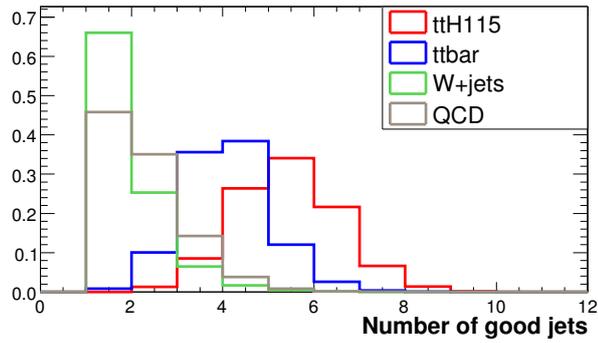
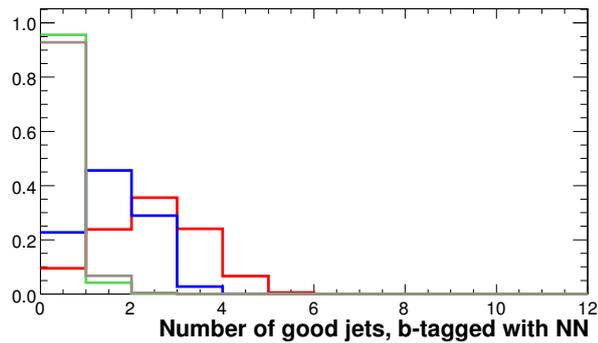


Figure 8.1: Normalized distribution of H_T (applying a $p_t^{jet} > 15$ GeV cut) for $t\bar{t}H$ where the Higgs mass was set to 115 GeV (red), $t\bar{t}$ (blue), W +jets (green) and multijet (brown) production. H_T is defined as the scalar sum of the transverse momenta of the 4 leading jets.



(a)



(b)

Figure 8.2: Normalized distribution of the number of good jets (a) and the number of good jets which have a NN loose b -tag (b) for $t\bar{t}H$ where the Higgs mass was set to 115 GeV (red), $t\bar{t}$ (blue), W +jets (green) and multijet (brown) production.

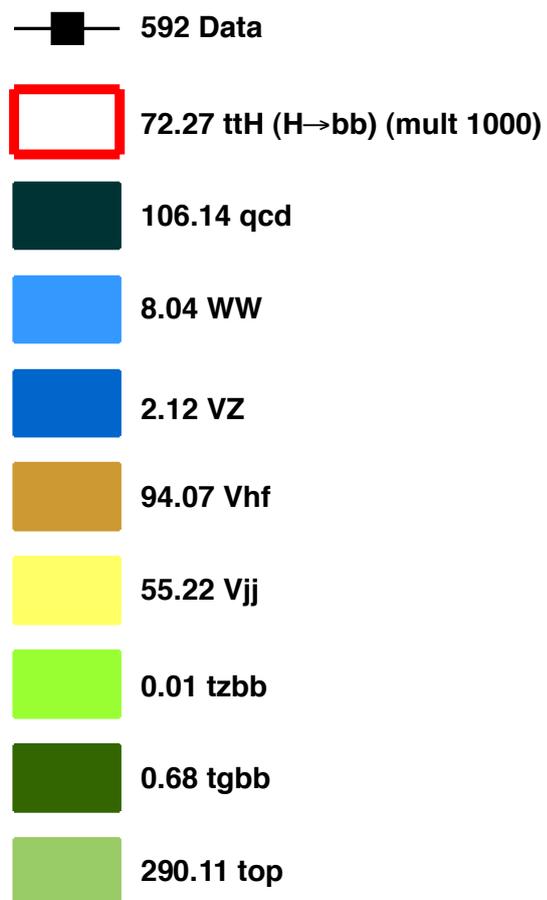


Figure 8.3: Legend for input plots, with signal scaling factor 1000.

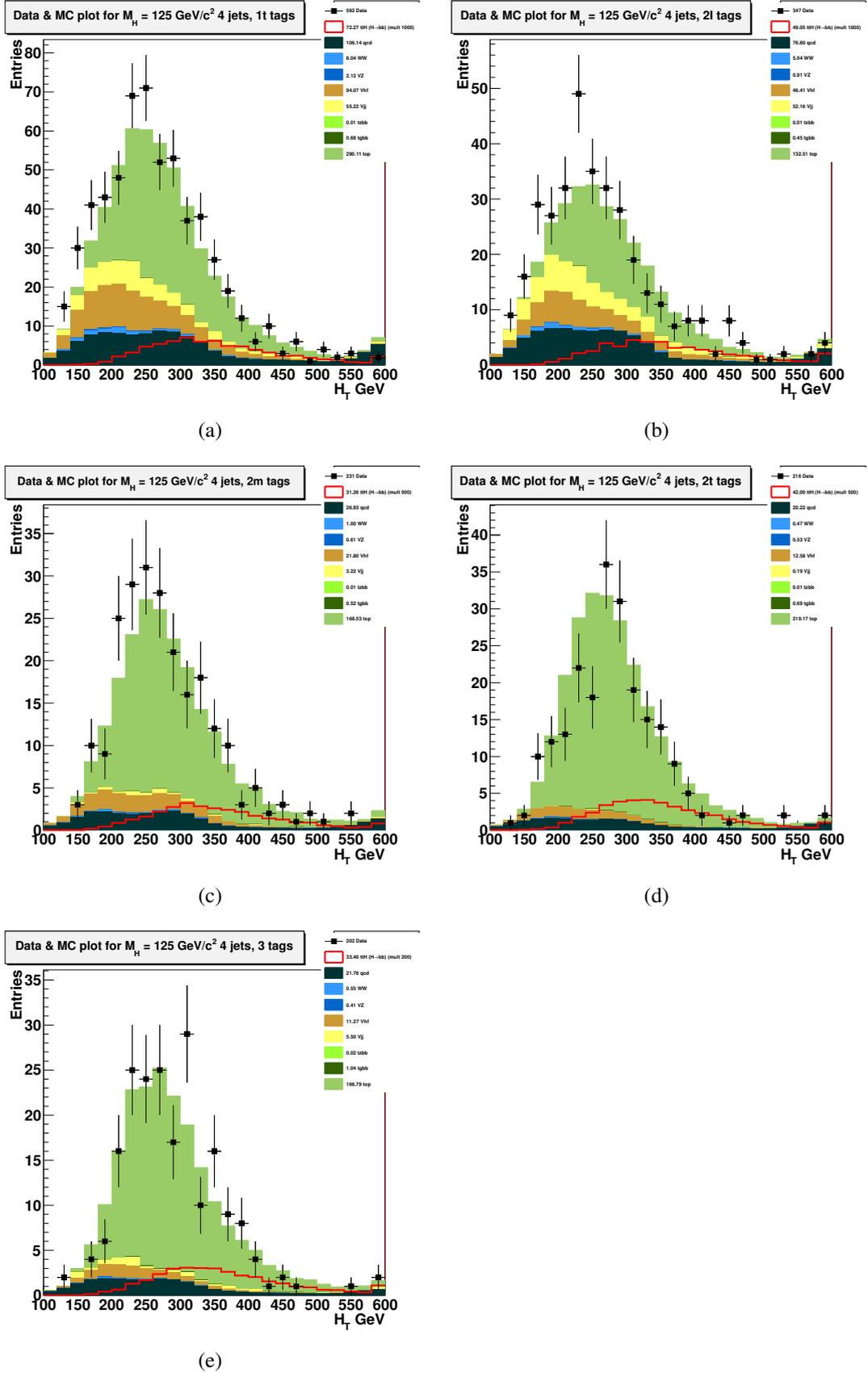


Figure 8.4: H_T distributions corresponding to the e +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with 4 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

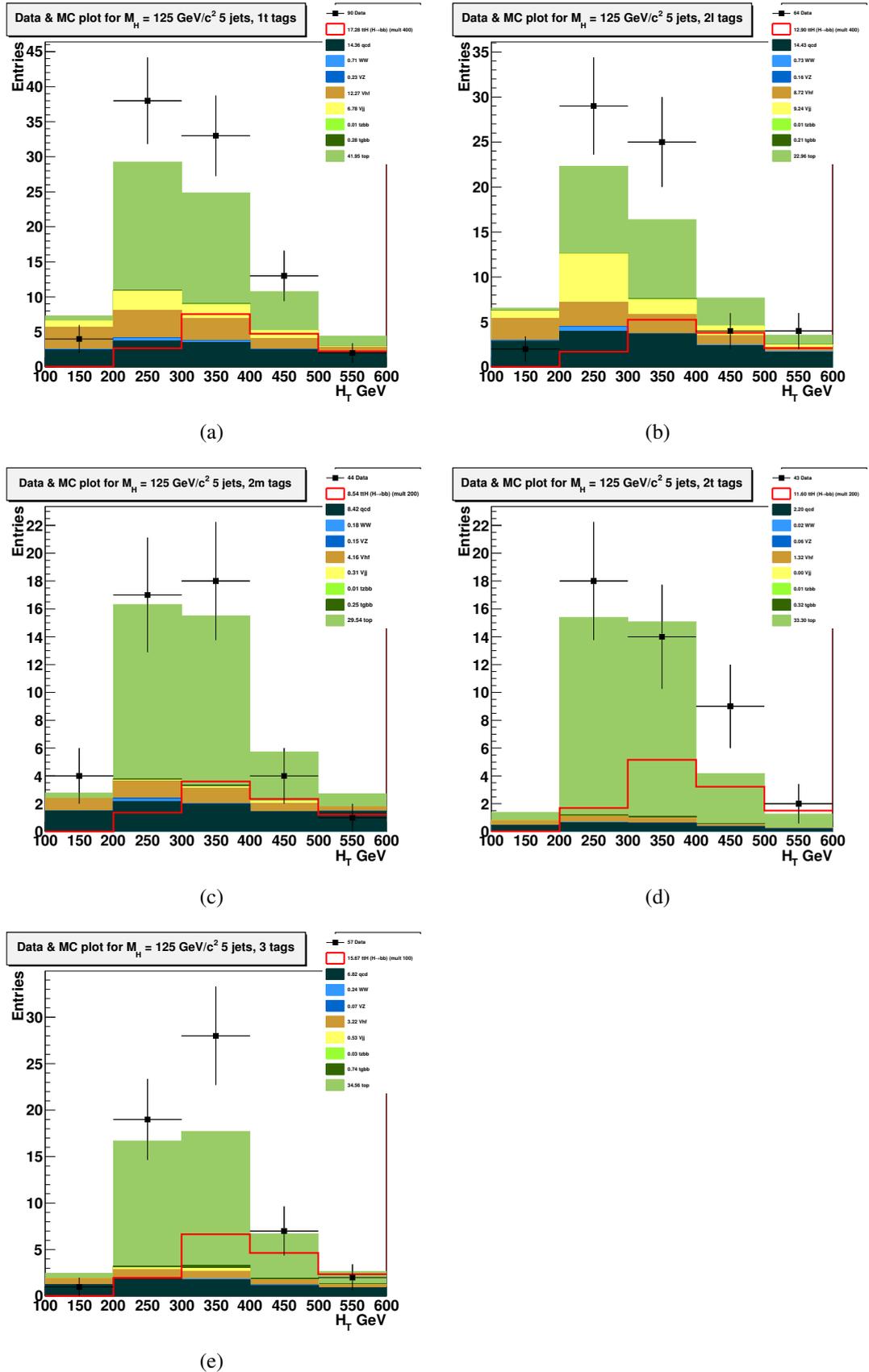


Figure 8.5: H_T distributions corresponding to the e +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with 5 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

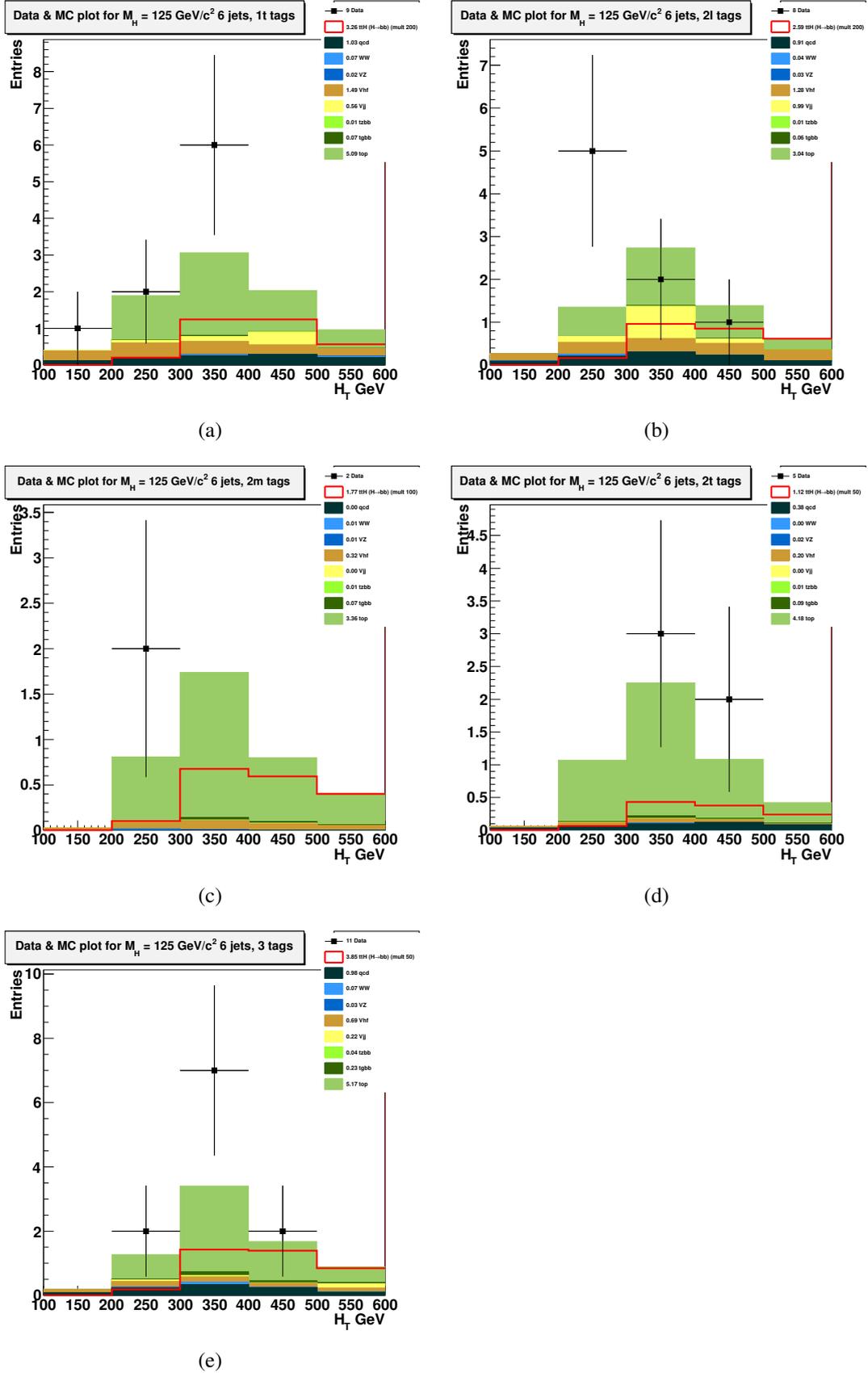


Figure 8.6: H_T distributions corresponding to the e +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with ≥ 6 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

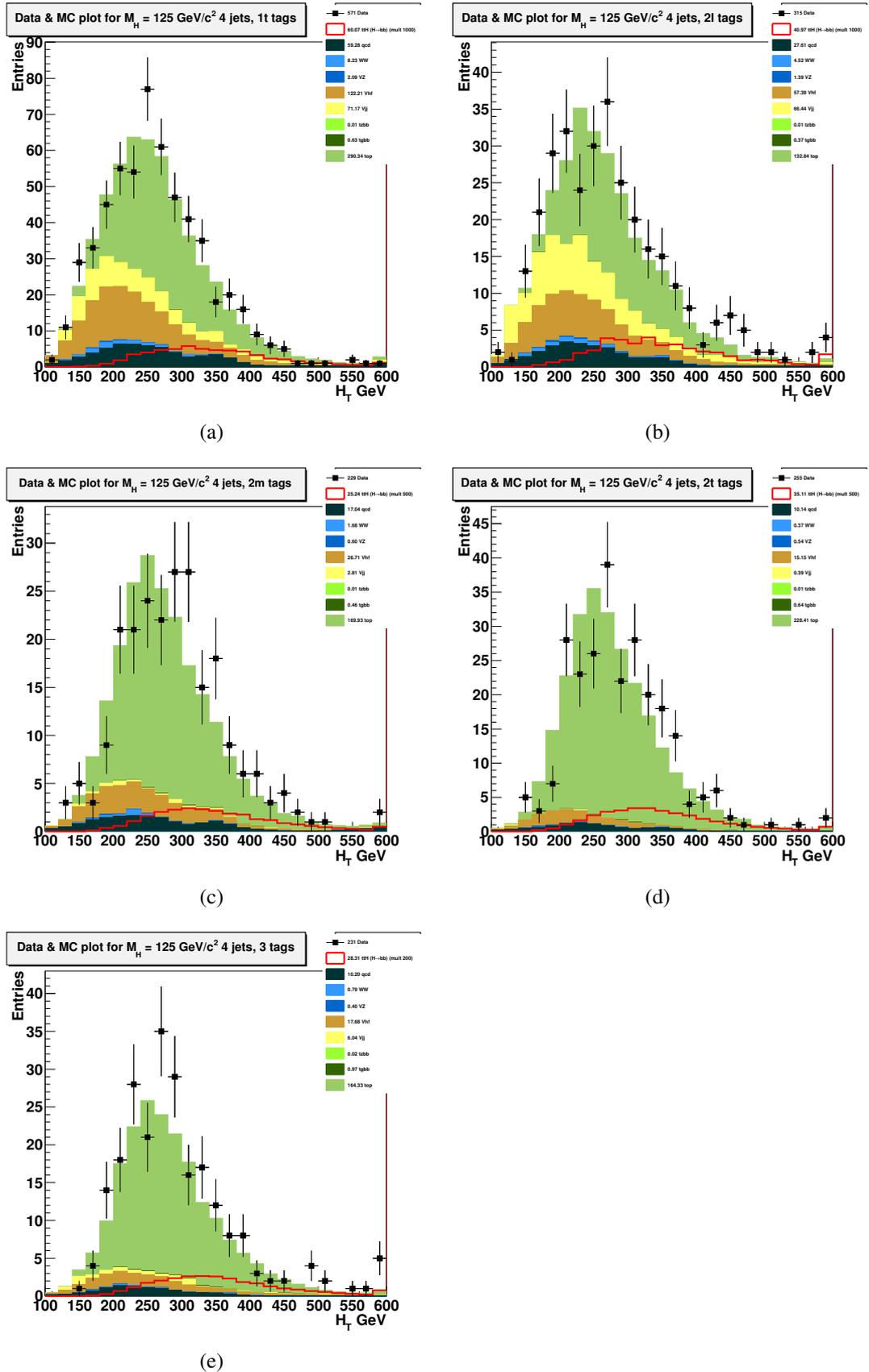


Figure 8.7: H_T distributions corresponding to the μ +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with 4 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

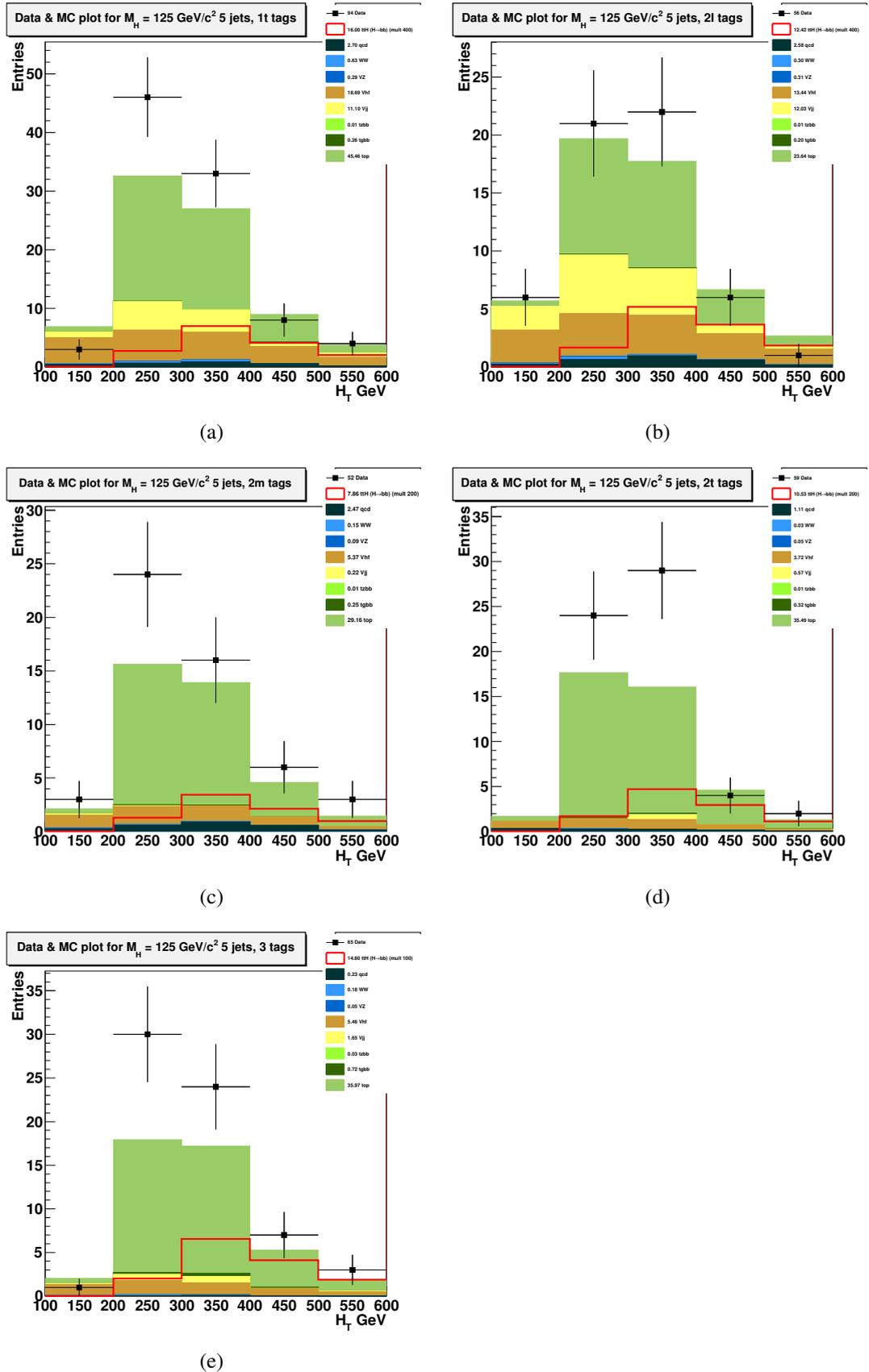


Figure 8.8: H_T distributions corresponding to the μ +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with 5 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

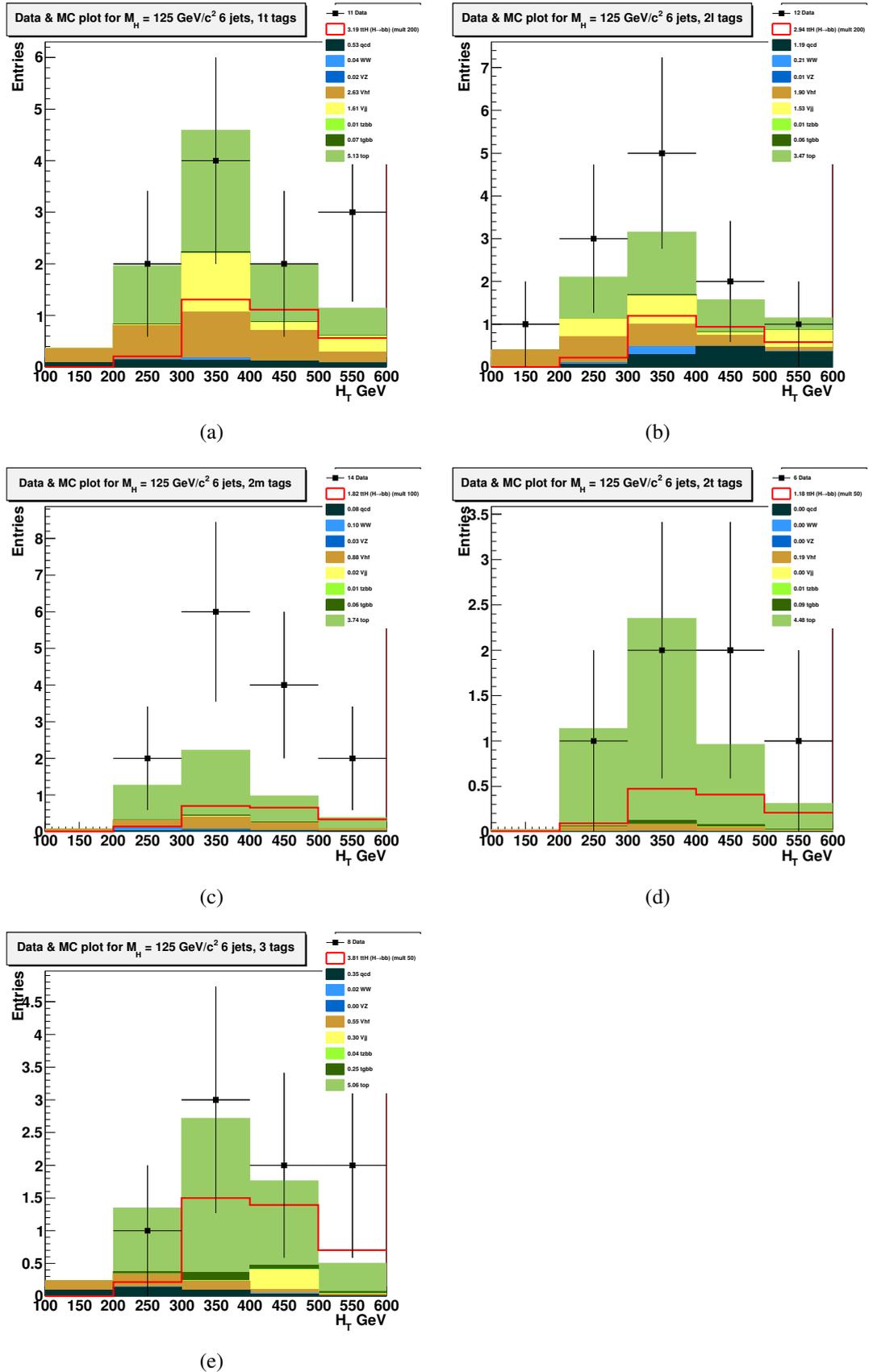


Figure 8.9: H_T distributions corresponding to the μ +jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (upper left), 2 loose b -tags (upper right), 2 medium b -tags (middle left), 2 tight b -tags (middle right) and ≥ 3 b -tags (lower) for events with ≥ 6 jets. The signal is enhanced by a factor as indicated in the legend and overlaid (solid red histogram).

	e+jets				
	5j1tt	5j2lt	5j2mt	5j2tt	5j3t
Signal	0.043	0.032	0.043	0.058	0.16
$t\bar{t}$ +single top	41	23	30	33	35
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.28	0.21	0.25	0.32	0.74
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.03
WW	0.71	0.73	0.18	0.02	0.24
VZ	0.23	0.16	0.15	0.06	0.07
V+HF jets	12	8.7	4.2	1.3	3.2
V+LF jets	6.8	9.2	0.31	< 0.01	0.53
multijets	14	14	8.4	2.2	6.8
sum Bkg	77 ± 9	56 ± 8	43 ± 7	37 ± 6	46 ± 7
Observed	90	64	44	43	57

	μ +jets				
	5j1tt	5j2lt	5j2mt	5j2tt	5j3t
Signal	0.040	0.031	0.039	0.053	0.15
$t\bar{t}$ +single top	45	24	29	35	36
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.26	0.20	0.25	0.32	0.72
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.03
WW	0.63	0.30	0.15	0.03	0.18
VZ	0.29	0.31	0.09	0.05	0.05
V+HF jets	19	13	5.4	3.7	5.5
V+LF jets	11	12	0.22	0.57	1.7
multijets	2.7	2.6	2.5	1.1	0.23
sum Bkg	79 ± 9	53 ± 7	38 ± 6	41 ± 6	44 ± 7
Observed	94	56	52	59	65

Table 6: Summary of expected and observed yields in the various channels from the 5 jet 1 tight b -tag bin (5j1tt) to the 5 jet ≥ 3 b -tag bin (5j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.

	e+jets				
	6j1tt	6j2lt	6j2mt	6j2tt	6j3t
Signal	0.016	0.013	0.018	0.022	0.077
$t\bar{t}$ +single top	5.1	3.0	3.4	4.2	5.2
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.07	0.06	0.07	0.09	0.23
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.04
WW	0.07	0.04	0.01	< 0.01	0.07
VZ	0.02	0.03	0.01	0.02	0.03
V+HF jets	1.5	1.3	0.32	0.20	0.69
V+LF jets	0.56	0.99	< 0.01	< 0.01	0.22
multijets	1.0	0.91	< 0.01	0.38	0.98
sum Bkg	8.3 ± 2.9	6.4 ± 2.5	3.8 ± 1.9	4.9 ± 2.2	7.4 ± 2.7
Observed	9	8	2	5	11

	μ +jets				
	6j1tt	6j2lt	6j2mt	6j2tt	6j3t
Signal	0.016	0.015	0.018	0.024	0.076
$t\bar{t}$ +single top	5.1	3.5	3.7	4.5	5.1
$t\bar{t}g \rightarrow t\bar{t}b\bar{b}$	0.07	0.06	0.06	0.09	0.25
$t\bar{t}Z \rightarrow t\bar{t}b\bar{b}$	0.01	0.01	0.01	0.01	0.04
WW	0.04	0.21	0.10	< 0.01	0.02
VZ	0.02	0.01	0.03	< 0.01	< 0.01
V+HF jets	2.6	1.90	0.88	0.19	0.55
V+LF jets	1.61	1.5	0.02	< 0.01	0.30
multijets	0.53	1.2	0.08	< 0.01	0.35
sum Bkg	10 ± 3	8.4 ± 2.9	4.9 ± 2.2	4.8 ± 2.2	6.6 ± 2.6
Observed	11	12	14	6	8

Table 7: Summary of expected and observed yields in the various channels from the ≥ 6 jet 1 tight b -tag bin (6j1tt) to the ≥ 6 jet ≥ 3 b -tag bin (6j3t). The background is given for $\sigma_{t\bar{t}} = 7.46$ pb. The expectations are shown for a Higgs mass of 125 GeV. The uncertainties are statistical and Gaussian only.

Source	Value
Event preselection	1.2%
Muon identification	4%
Electron identification	3%
Luminosity	6.1%
PDF model	2%
Jet energy scale	6%
Jet energy resolution	5%
Heavy background models	20%
Light flavour background models	6%
Diboson models	6%
Multijet background model	50%
Uncertainty on $\sigma_{t\bar{t}}$	7%
Uncertainty on $t\bar{t}b\bar{b}$	50%
Uncertainty on $t\bar{t}Z$	25%

Table 8: Summary of H_T -independent systematic uncertainties used as input for the limit derivation.

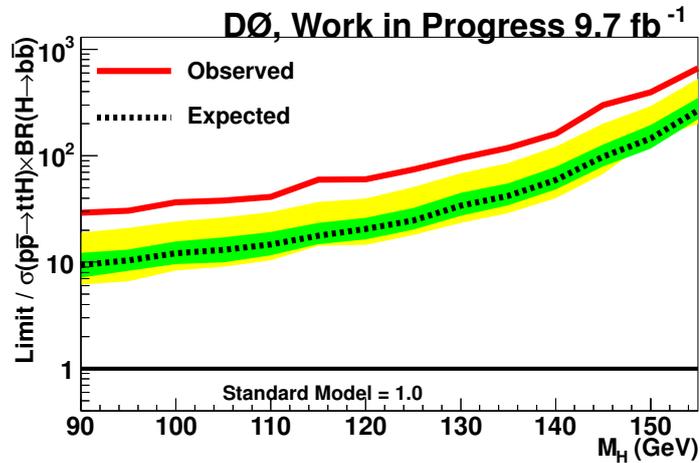


Figure 8.10: The 95% CL upper limit on the $\sigma_{t\bar{t}H}$ cross section times branching ratio over the SM expectation in NLO QCD as a function of the Higgs mass. The 1 and 2 σ uncertainty bands for the expected limit are indicated by the green and green+yellow bands.

Higgs mass (GeV)	Expected	Observed
90	9.39	29.1
95	10.4	30.5
100	12.2	36.5
105	13.1	38.0
110	14.7	41.0
115	17.8	59.4
120	20.5	60.0
125	24.7	74.3
130	34.0	95.1
135	42.0	118
140	59.3	160
145	98.2	300
150	146	394
155	265	664

Table 9: Expected and observed ratios of excluded $t\bar{t}H$ cross section times $H \rightarrow b\bar{b}$ branching fraction over SM expectation for different values of the Higgs mass.

Chapter 9

Search Using a Multivariate Approach

9.1 Multivariate Analysis Method

In the course of the analysis, a study was undertaken to assess the performance of multivariate analysis (MVA) techniques, as these have shown promising results in other analyses at D0, such as single top measurements [64] and Higgs searches [?], as well as the $t\bar{t}H$ searches at CMS [61] and CDF [62].

For this technique, the sample was split into 3 subsamples: training, testing and evaluating. This is to ensure that the final distributions used as inputs into our limit setting procedure are not biased by the training or testing of the MVAs.

Training was only used on the 4 and 5 jet categories, as the low yields in the 6 jet samples led to large statistical fluctuations in the output distributions and became wildly overtrained. Hence, the BDT weights derived from the 5 jet category were applied to events in the 6 jet category. Because the derived weights are necessarily applied to different events in the 5 and 6 jet categories, different distributions are still acquired, so there is no concern about having identical input distributions in the limit setting procedure and thus introducing any level of degeneracy.

A Boosted Decision Tree (BDT) method was used as this method is more reliable than some other TMVA methods when training on samples which are low in statistics.

The decision trees were trained on samples of $t\bar{t}$, $t\bar{t}g$ and $t\bar{t}Z$, with the $t\bar{t}H$ process taken as the signal sample for training. The $t\bar{t}$ background, along with its subset the $t\bar{t}g$ background, represent the majority of the data events which could be mistaken for signal events due to the much larger production cross section of the $t\bar{t}$ process when compared with $t\bar{t}H$ as well as the inherent difficulty in identifying all jets associated with the hard scattering. Distributions of variables which can be used to discriminate between the signal and background processes therefore rely mainly on the presence of two additional energetic jets.

The $t\bar{t}Z$ process, while boasting only a very low production cross section, is more signal-like in its kinematic distributions than either the $t\bar{t}$ or $t\bar{t}g$. Hence, any variables which demonstrate some discriminating power between $t\bar{t}Z$ and $t\bar{t}H$ are also of value.

9.2 Discriminating Variables

The distributions that are used as input for the training of the BDTs are described in this section. Not all distributions were included for all channels, in some cases due to poor description of the data, or due to lack of discriminating power compared with other variables.

In addition to the variables already found in the existing framework, the fitting algorithm HITFIT [65] as used by the Top Group at D0 was incorporated into the analysis in order to employ variables that exhibit differences between the signal and $t\bar{t}$ background which arise from certain properties of the top quark. HITFIT is a kinematic fitter used for determining the most likely jet pairings of the $t\bar{t}$ pair with regard to the hadronically decaying top quark. Due to the 2 extra b jets in the signal, this presents HITFIT with more permutations than in the $t\bar{t}$ case, and so the fitter will not always match the correct jets, leading to differences between the distributions in signal

and background.

The list below gives a brief description of each variable along with the variable's name within the framework for comparison with the input distributions for both signal and background (in each jet, tag and lepton category) shown in C. The variables are described here in order of the ranking of their separation power for the 5 jet, 3 tag (electron) case. Since not all variables were used in this category, the remainder are listed alphabetically by their framework nomenclature.

1. Mean transverse energy of jets in the event; `jet_et_mean`
2. Sum of transverse energy of jets; `topo_jets_sumet`
3. Sum of visible transverse momentum p_T in the event; `topo_vis_sumpt`
4. Transverse energy of the third leading jet; `jet_et[2]`
5. Maximum transverse energy of a single jet; `jet_et_max`
6. Transverse mass of the HITFIT-determined hadronically decaying top quark; `topo_hf_mt`
7. Transverse energy of the second leading jet; `jet_et[1]`
8. Value of the χ^2 fit from HITFIT; `topo_hf_chisq`
9. Invariant mass of lepton and visible part of neutrino, $(p_x^{miss}, p_y^{miss}, 0, MET)$ and two leading jets; `lnujj_zeronupz_m`
10. Mass of the HITFIT-determined $t\bar{t}$ pair; `topo_hf_m_tt`
11. Maximum transverse energy of a single tagged jet; `jet_et_max_tag`
12. Sum of transverse momentum of the visible lepton and missing transverse energy; `lnu_sumpt`

13. Ratio of energy of pair of jets with the highest invariant mass which is closest to the W mass; `lnujjjj-jp1-jj-eratio`
14. Ratio of energy of pair of jets with the lowest invariant mass which is closest to the W mass; `lnujjjj-jp2-jj-eratio`
15. Pseudorapidity angle η of the HITFIT-determined $t\bar{t}$ pair; `topo_hf_eta_tt`
16. Missing transverse energy; `nu_pt`
17. Transverse momentum of the HITFIT-determined hadronically decaying top quark; `topo_hf_pt_t2`
18. Angle ΔR between the visible lepton and missing transverse energy; `lnu_angle`
19. Lepton σ , where σ is defined as $(p_T(l) * \Delta R(W_{lep}, l) + MET * \Delta R(W_{lep}, MET)) / (p_T(l) + MET)$; `lep_sigma`
20. Pseudorapidity angle η of missing transverse energy; `nu_eta`
21. Angle between the visible lepton and missing transverse energy; `lnu_dphi`
22. $\cos \theta$ between lepton and missing transverse energy in centre of mass frame; `lnu_cm_cstheta`
23. Maximum angle between the lepton and leading jet, $\Delta\eta(l,j)$; `jetlep_detamax`
24. Pseudorapidity angle η for the 4th highest p_T jet; `jet_eta[3]`
25. Difference in pseudorapidity angle of the third leading jet and missing transverse energy; `jetnu_deta[2]`
26. Energy of the lepton; `lep_e`
27. Difference in pseudorapidity angle between the visible lepton and missing transverse energy; `lnu_deta`

28. Difference in pseudorapidity angle between lepton and leptonically decaying W ;
`lnu_lep_delta`
29. Transverse momentum of the leptonically decaying W with respect to the thrust;
`lnu_recoilpt`
30. Transverse mass of lepton, missing transverse energy \cancel{E}_T and leading 4 jets;
`lnujjjj_mt`

Using the weights which are obtained from the MVA process, they are then applied to data, background and signal events which were not used in the training or testing phases to produce characteristic distributions, seen in Figs C.1 - C.10 in the Appendix.

These distributions are then rebinned using an automated method which aims to make each bin which contains data events to also contain signal events so that the limit setting procedure (which is described in more detail in section 8) does not try to fit a signal of 0 to a non-zero data yield, as this would require an undefined scaling factor. The binned distributions are shown in Figs 9.1 - 9.6. In each distribution, the $t\bar{t}$ cross section is normalized to 7.46 pb, corresponding to a top quark mass of 172.5 GeV. The $t\bar{t}H$ signal is for a Higgs mass of 125 GeV and $\sigma(t\bar{t}H) \times B(H \rightarrow b\bar{b})$ to 4.28 fb. The signal is enhanced by a factor as indicated in each legend and overlaid (solid red histogram). Yields of other processes are included in the legend and are unscaled.

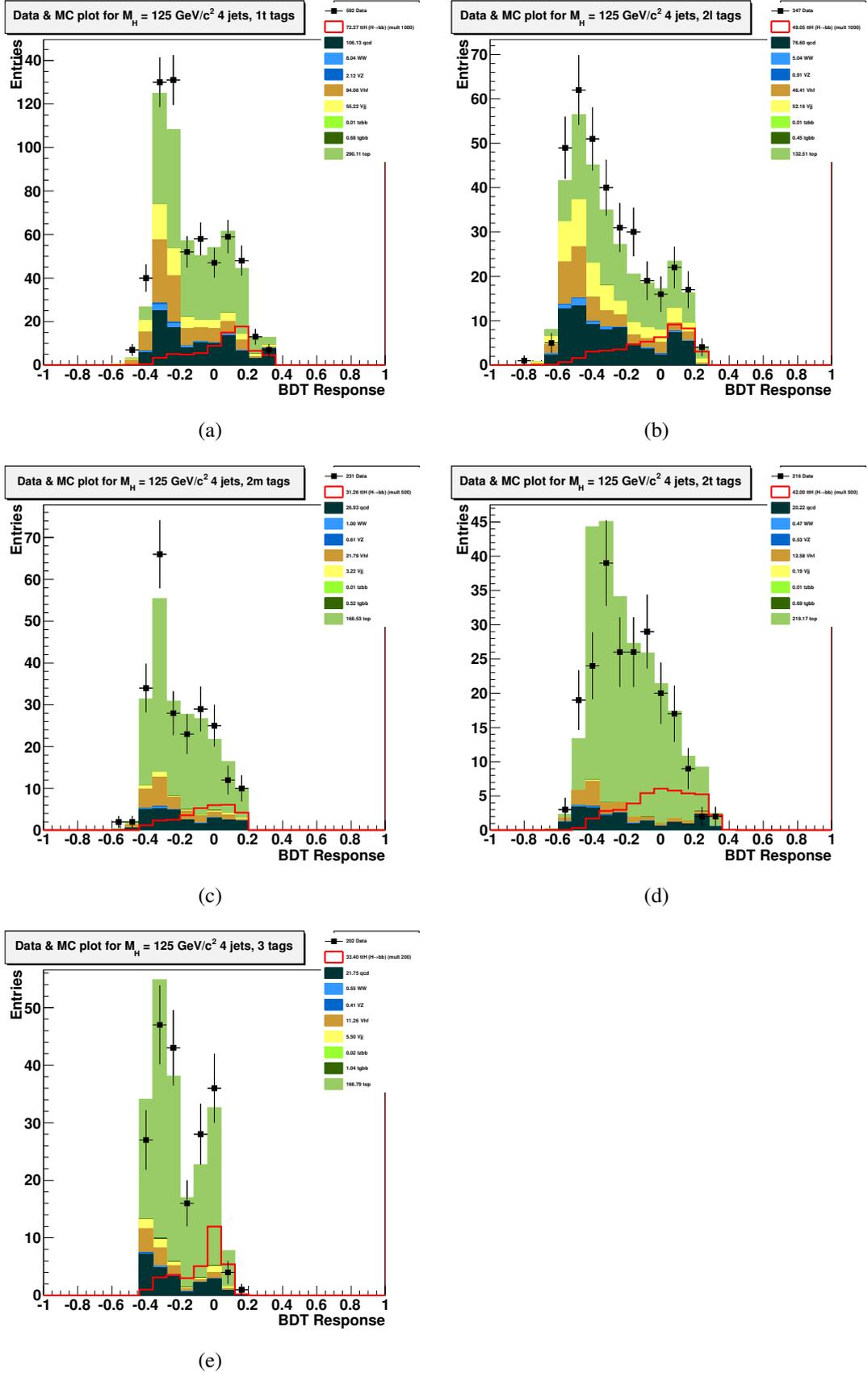
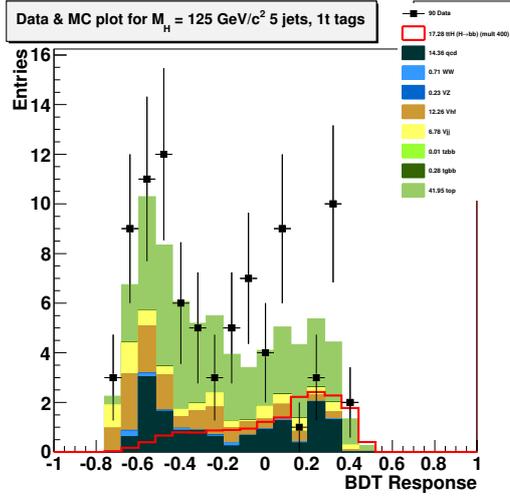
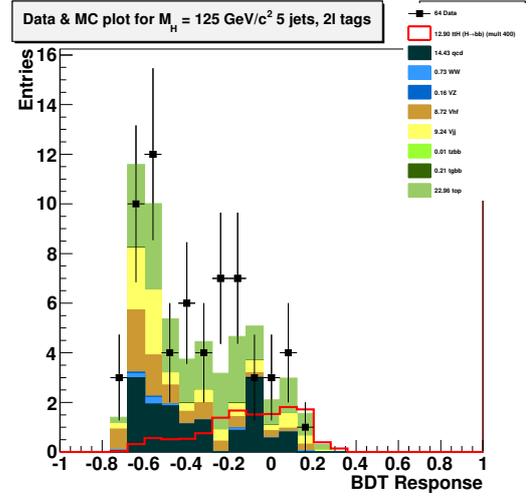


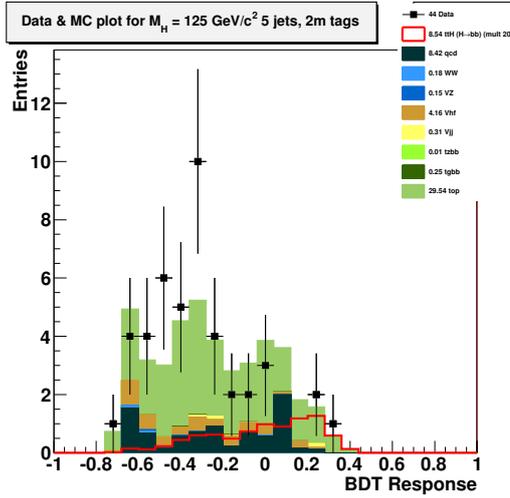
Figure 9.1: BDT distributions corresponding to the $e + 4$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).



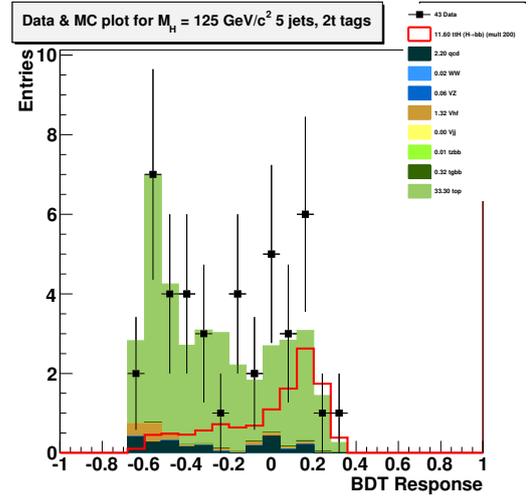
(a)



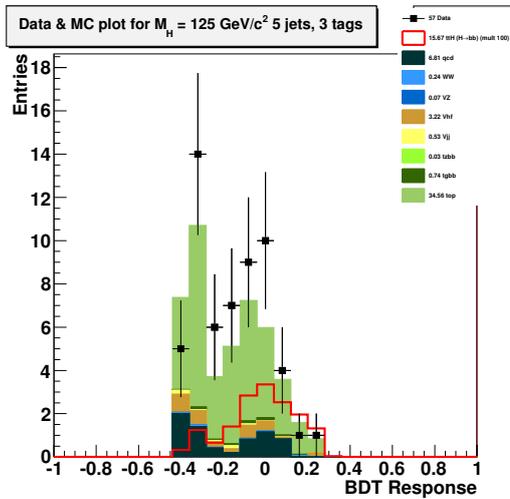
(b)



(c)



(d)



(e)

Figure 9.2: BDT distributions corresponding to the $e + 5$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).

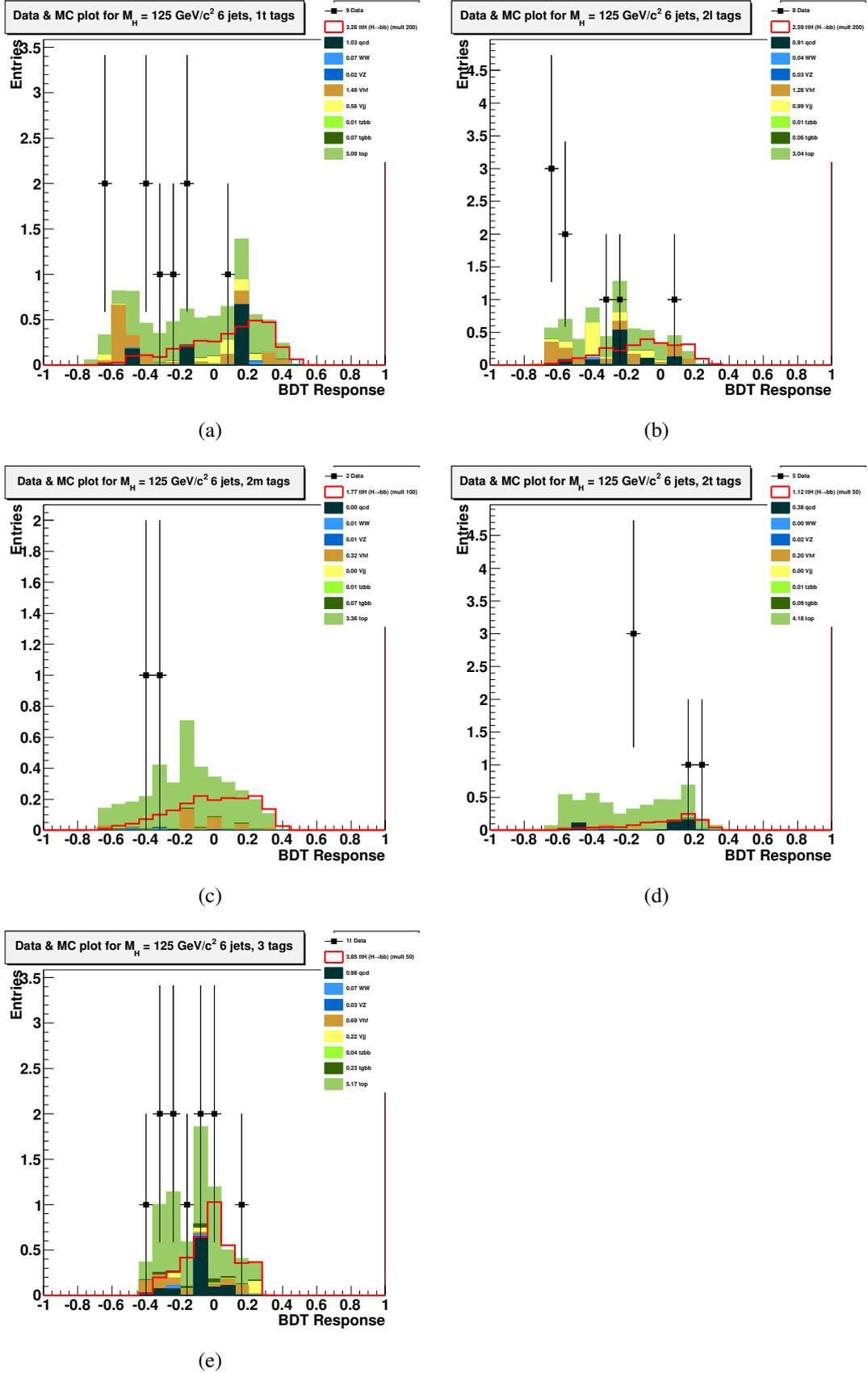
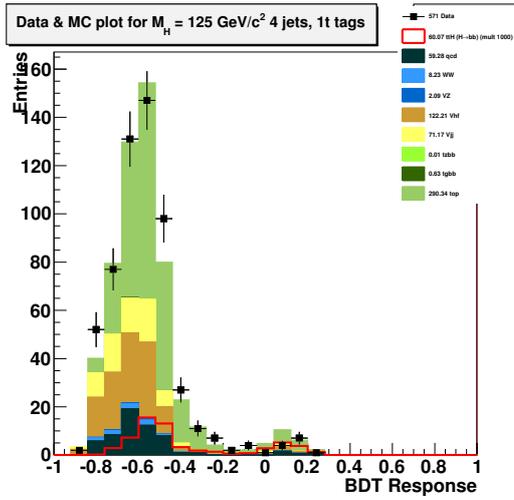
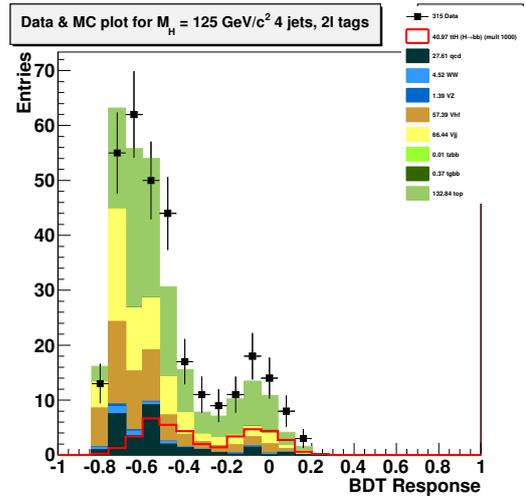


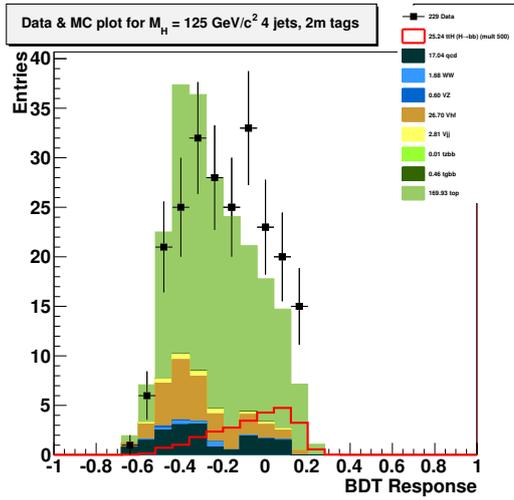
Figure 9.3: BDT distributions corresponding to the $e^+ \geq 6$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).



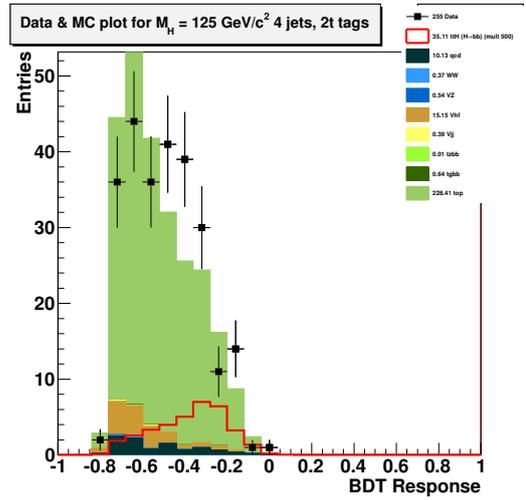
(a)



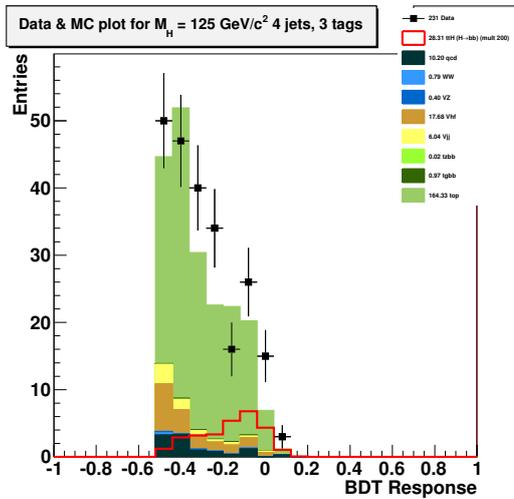
(b)



(c)

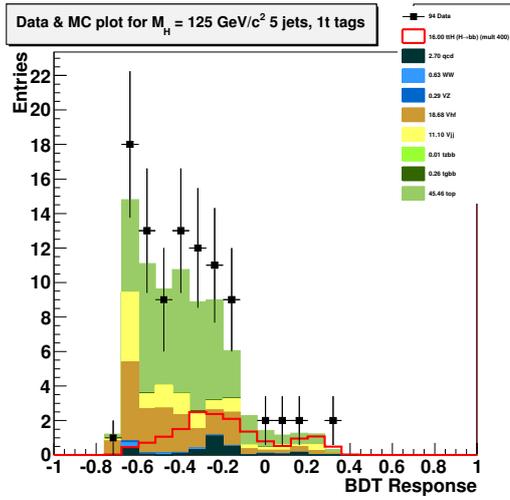


(d)

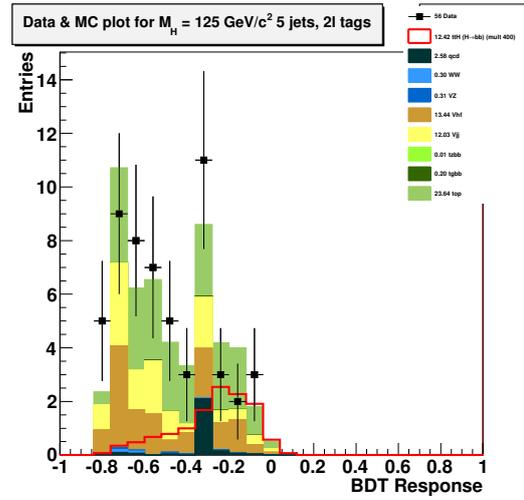


(e)

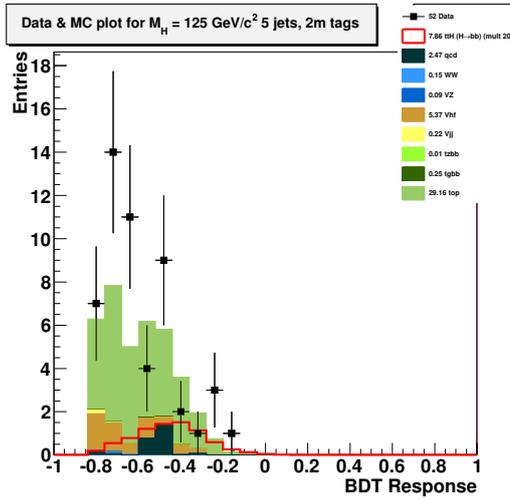
Figure 9.4: BDT distributions corresponding to the $\mu + 4$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).



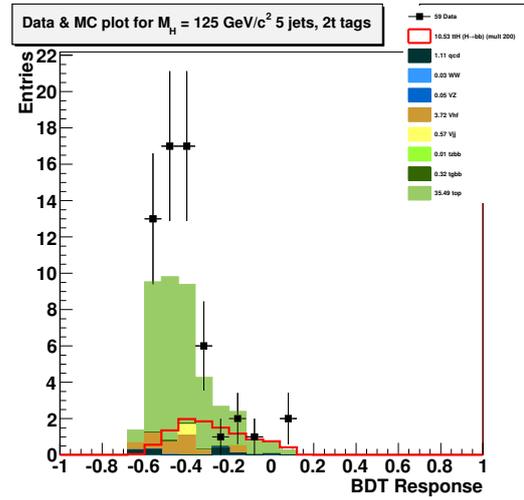
(a)



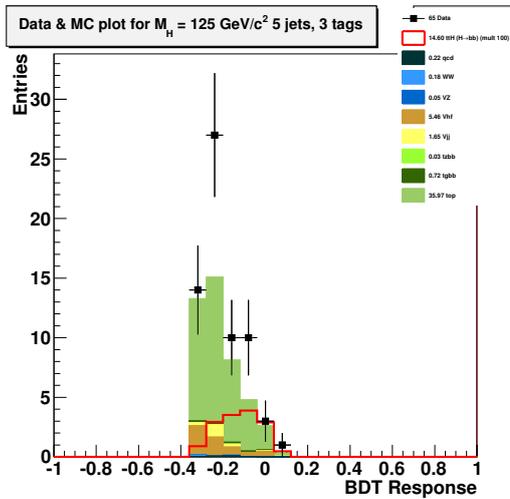
(b)



(c)



(d)



(e)

Figure 9.5: BDT distributions corresponding to the $\mu + 5$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).

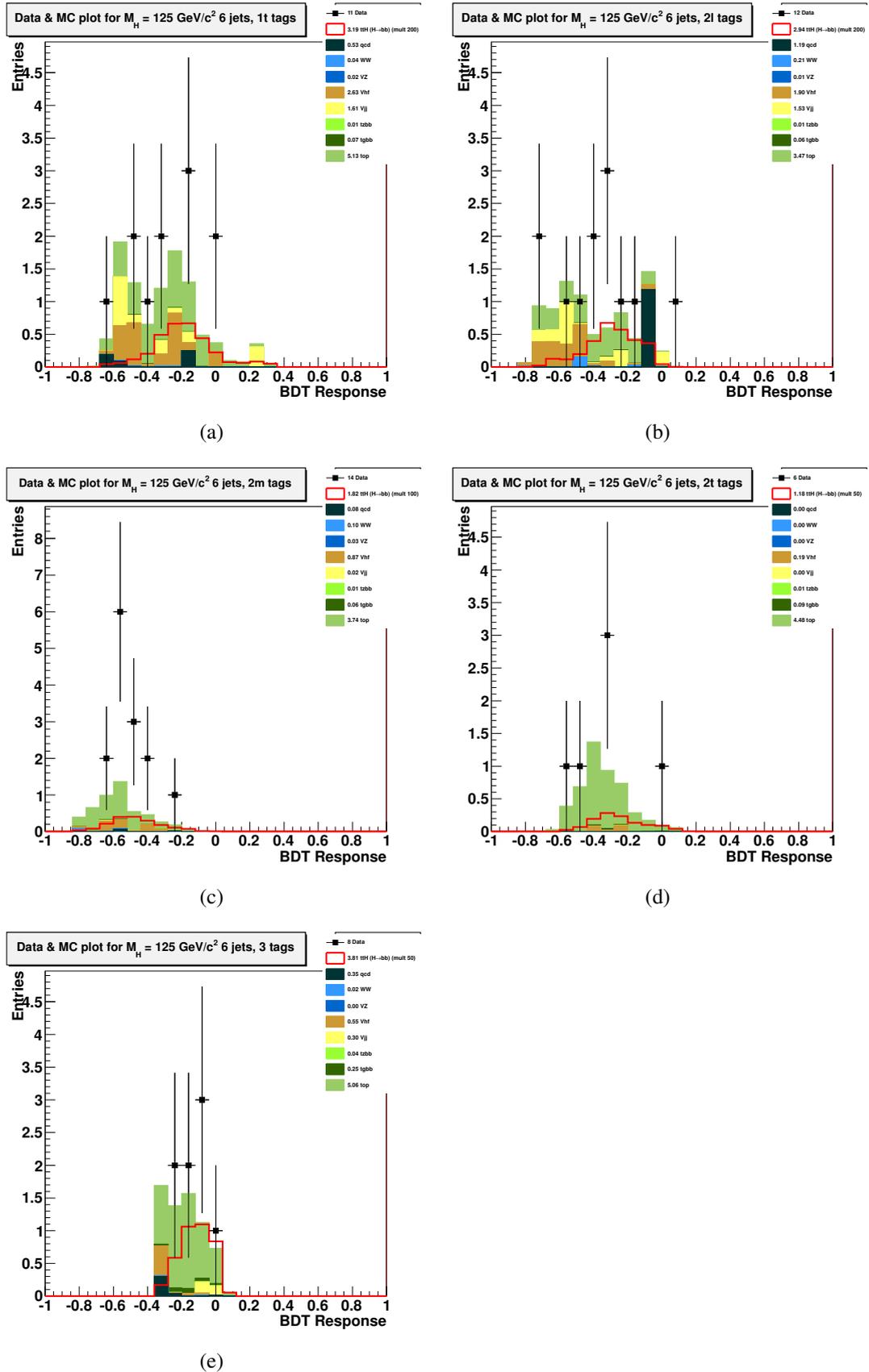


Figure 9.6: BDT distributions corresponding to the $\mu^+ \geq 6$ jets data set of 9.7 fb^{-1} requiring 1 loose b -tag (a), 2 loose b -tags (b), 2 medium b -tags (c), 2 tight b -tags (d) and ≥ 3 b -tags (e).

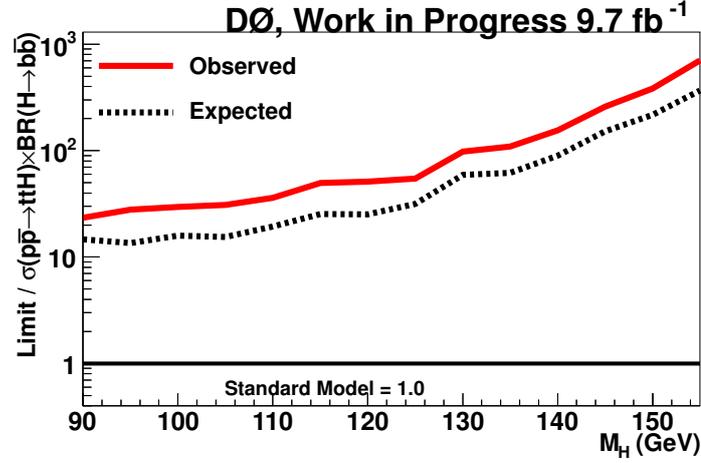


Figure 9.7: Limits obtained for the process $\sigma(t\bar{t}H) \times BR(H \rightarrow b\bar{b})$ at D0 in the range of Higgs masses $90 < m_H < 155$ GeV

Using the standard D0 limit setting procedure [59], limits relative to the SM predicted cross section are derived. These limits are scale factors that quantify how much the yield derived from the signal cross section times branching ratio would need to be multiplied by to account for discrepancies between the data and the background in these input distributions. Figure 9.7 shows the scale factors which were derived using this method (without systematic uncertainty bands), and they are summarised in table 10.

For a Higgs mass of 125 GeV, the limits are 31.6 (54.6) expected (observed) times the SM $t\bar{t}H \rightarrow t\bar{t}b\bar{b}$.

Since these nominal limits do not show improvement over those obtained using H_T alone as a discriminating variable, further analysis using the inputs derived from the BDT method given above was not pursued, as using a single, well-understood variable yields the same or similar results while being less prone to systematic uncertainties.

It is not clear why the BDT method does not show improvement as was expected.

Higgs mass (GeV)	Expected	Observed
90	14.7	23.3
95	13.6	27.8
100	15.9	29.5
105	15.5	30.9
110	19.3	36.1
115	25.4	49.6
120	25.3	51.1
125	31.6	54.6
130	59.2	97.6
135	61.6	109
140	89.8	156
145	152	259
150	217	385
155	367	705

Table 10: Expected and observed ratios of excluded $t\bar{t}H$ cross section times $H \rightarrow b\bar{b}$ branching fraction over SM expectation given to 3 s. f. for different values of the Higgs mass.

One reason may be due to the inability to train on the 6 jet samples, particularly the higher tagging multiplicities, as this is where the signal-to-background ratio is highest. Further work would be required to establish stable BDT training parameters for these jet and tagging categories. The rebinning method may also be partially responsible, as this tends to reduce the spread of both signal and background MVA distributions by construction, in order to avoid errors arising from a division by 0. While this is clearly necessary, there may be scope for optimisation within the method employed such that a greater spread (and thus more bins in which a countable difference between signal and background can be utilised) of signal and background can be achieved.

Chapter 10

Summary

In this thesis a search for the production of the SM Higgs boson in association with top and antitop quarks in a data set of 9.7 fb^{-1} is performed. Kinematical information is analysed in different bins of jets multiplicity and b -tagged jets multiplicity. A combination of many different distributions being able to separate between signal and the main backgrounds is performed using multi-variate analysis techniques. A BDT output distribution is found to be able to separate signal and background well. However, using the distribution of the scalar sum of all final state objects (H_T) gives a similar sensitivity and is therefore preferred.

The channels with 4, 5 or ≥ 6 jets and 1 tight, 2 loose, 2 medium, 2 tight and ≥ 3 b -tags are investigated separately. In all channels within the uncertainties agreement between the observed and expected number of events is found; no hint of associated Higgs production is found. Upper limits on $t\bar{t}H$ production are derived. For a Higgs mass of 125 GeV the expected limit for $\sigma(t\bar{t}H) \times B(H \rightarrow b\bar{b})$ is 27 times larger than the SM calculation. The observed limit is a factor of 48 larger than the SM prediction. These results improve on previous limits derived by the D0 collaboration.

Further work could be done using this analysis by searching for a signal which has the same final state, but resulting from a different initial state such as a massive color-octet vector boson decaying to a top quark and a heavy t' quark, which would

then decay into a Higgs boson and a top quark.

Appendix A

Validation Plots

Below are presented the validation plots from section 7.1 separated into the five tagging categories. For each plot, the signal is scaled by a factor as indicated in the legend and overlaid (solid red histogram).

A legend indicating the data, background and signal samples (signal sample with scaling factor of 200, as in the case of 5 jets and 2 tight tags).

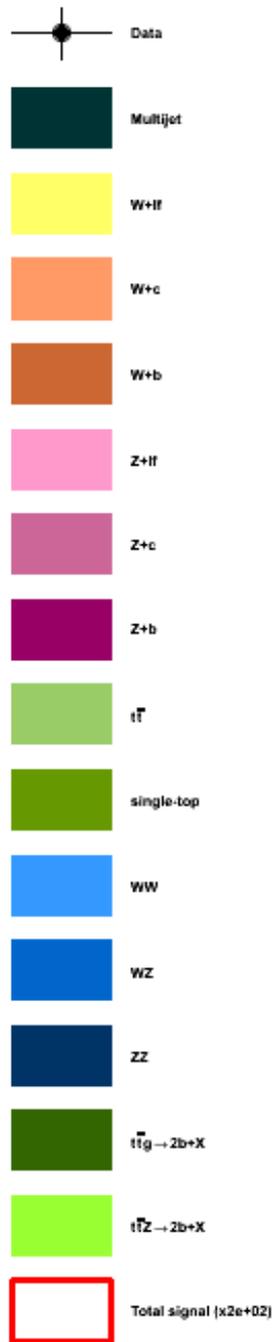


Figure A.1: Legend for validation plots.

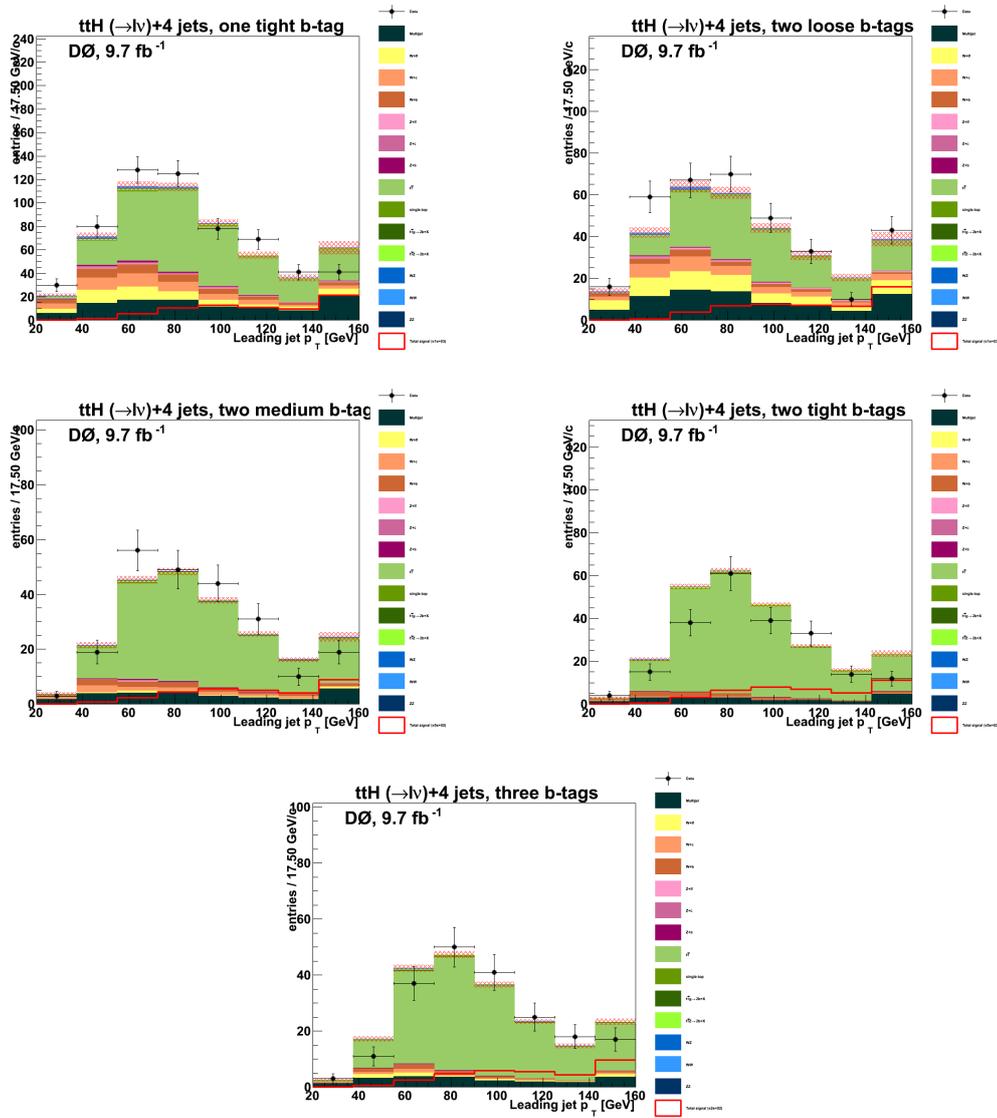


Figure A.2: Validation plots for the jet with highest transverse momentum p_T for events with 1 electron, exactly 4 jets and A.2(a) 1 tight tag, A.2(b) 2 loose tags, A.2(c) 2 medium tags, A.2(d) 2 tight tags and A.2(e) 3 tags.

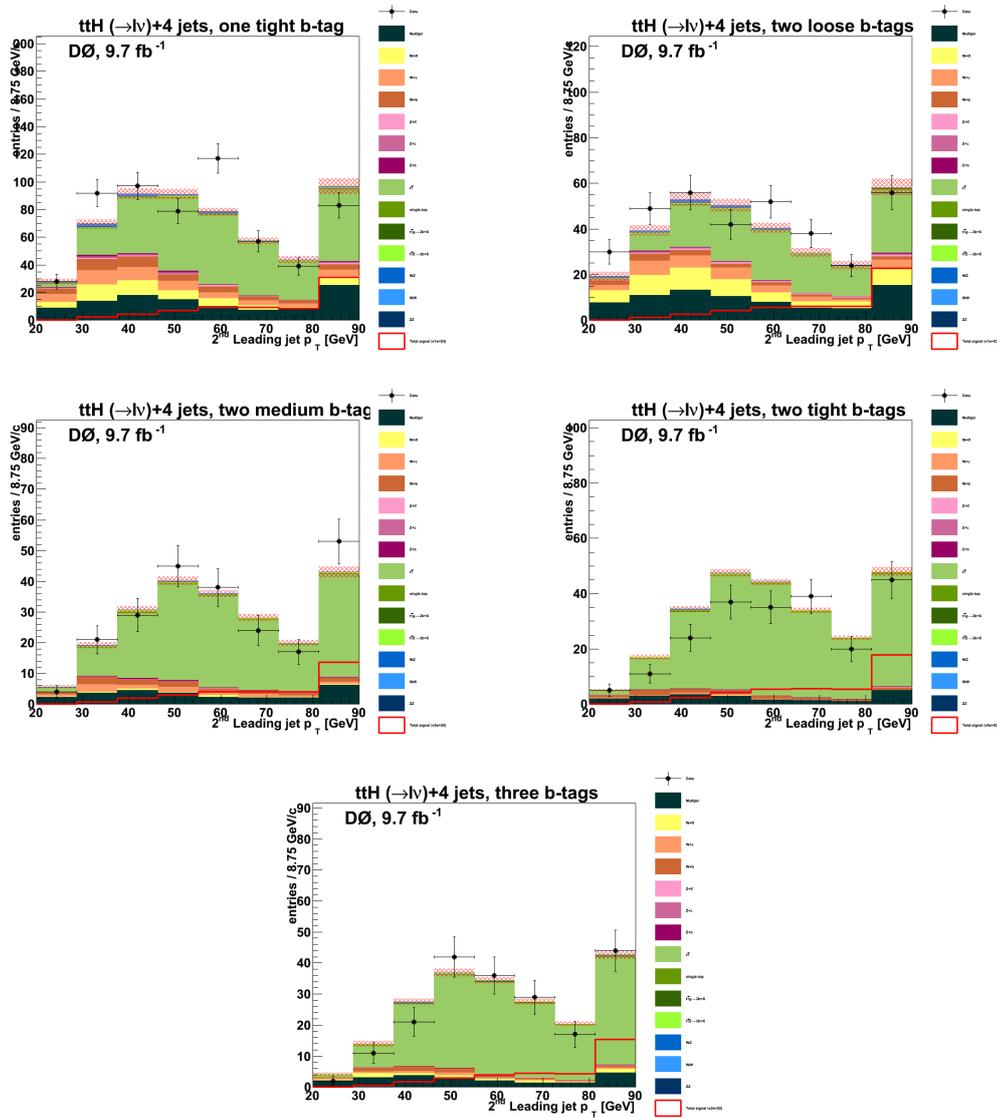


Figure A.3: Validation plots for the jet with second highest transverse momentum p_T for events with 1 electron, exactly 4 jets and A.3(a) 1 tight tag, A.3(b) 2 loose tags, A.3(c) 2 medium tags, A.3(d) 2 tight tags and A.3(e) 3 tags.

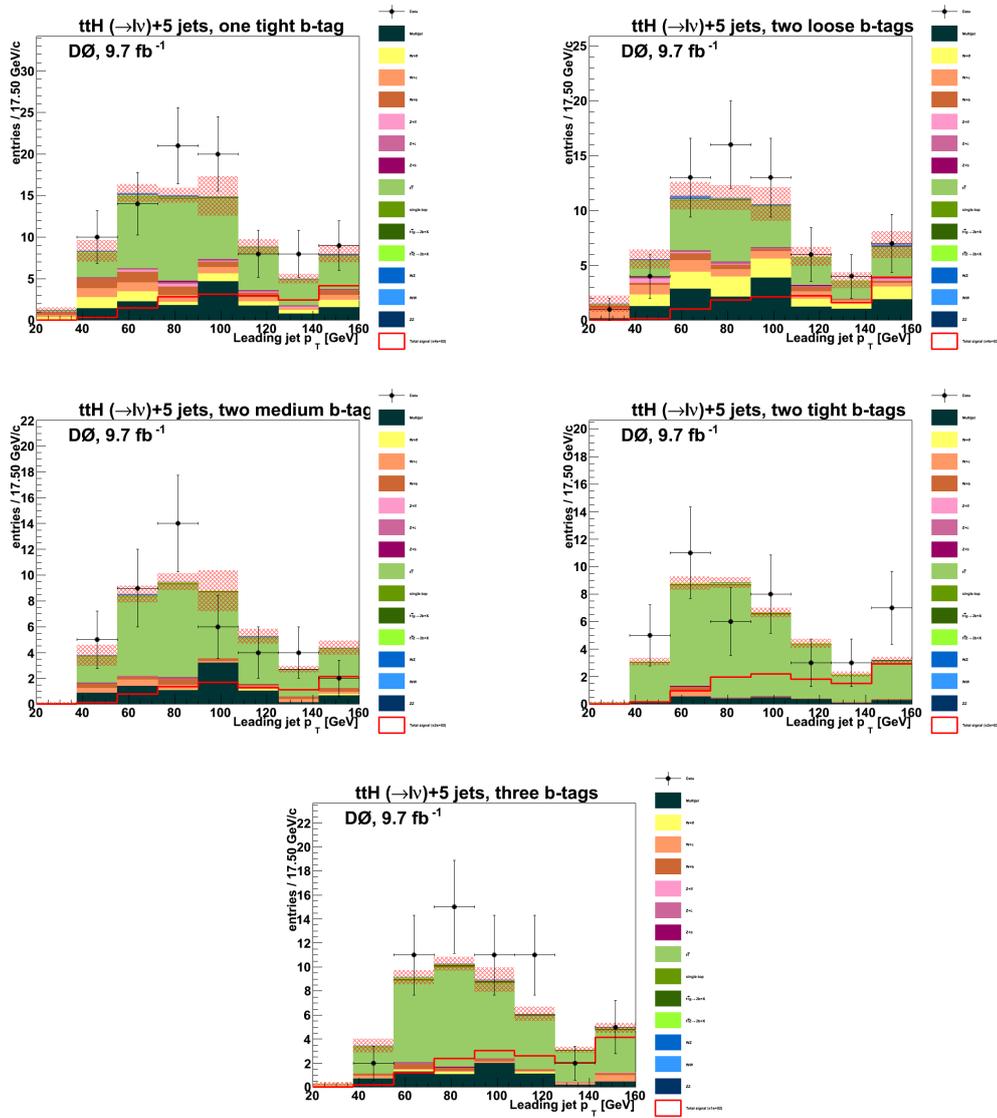


Figure A.6: Validation plots for the jet with highest transverse momentum p_T for events with 1 electron, exactly 5 jets and A.6(a) 1 tight tag, A.6(b) 2 loose tags, A.6(c) 2 medium tags, A.6(d) 2 tight tags and A.6(e) 3 tags.

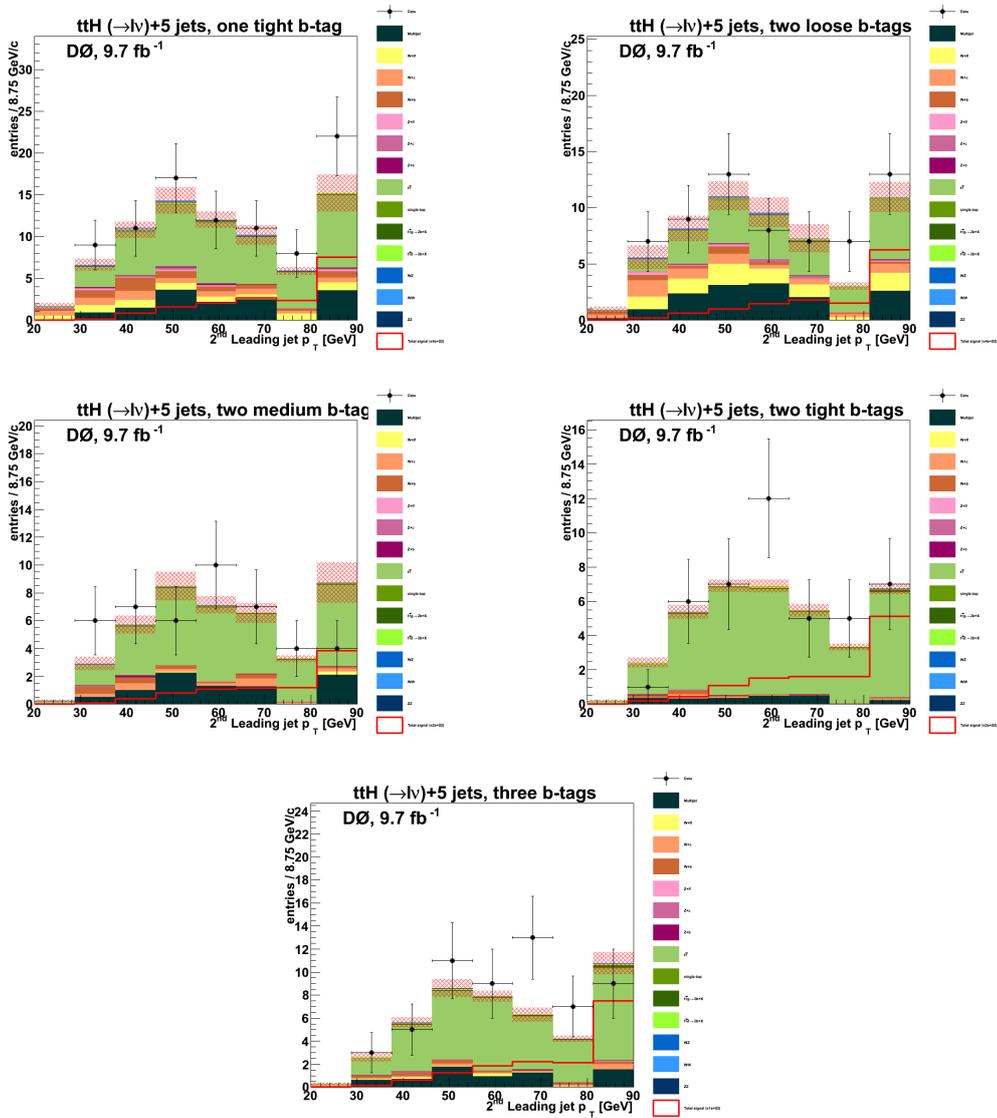


Figure A.7: Validation plots for the jet with second highest transverse momentum p_T for events with 1 electron, exactly 5 jets and A.7(a) 1 tight tag, A.7(b) 2 loose tags, A.7(c) 2 medium tags, A.7(d) 2 tight tags and A.7(e) 3 tags.

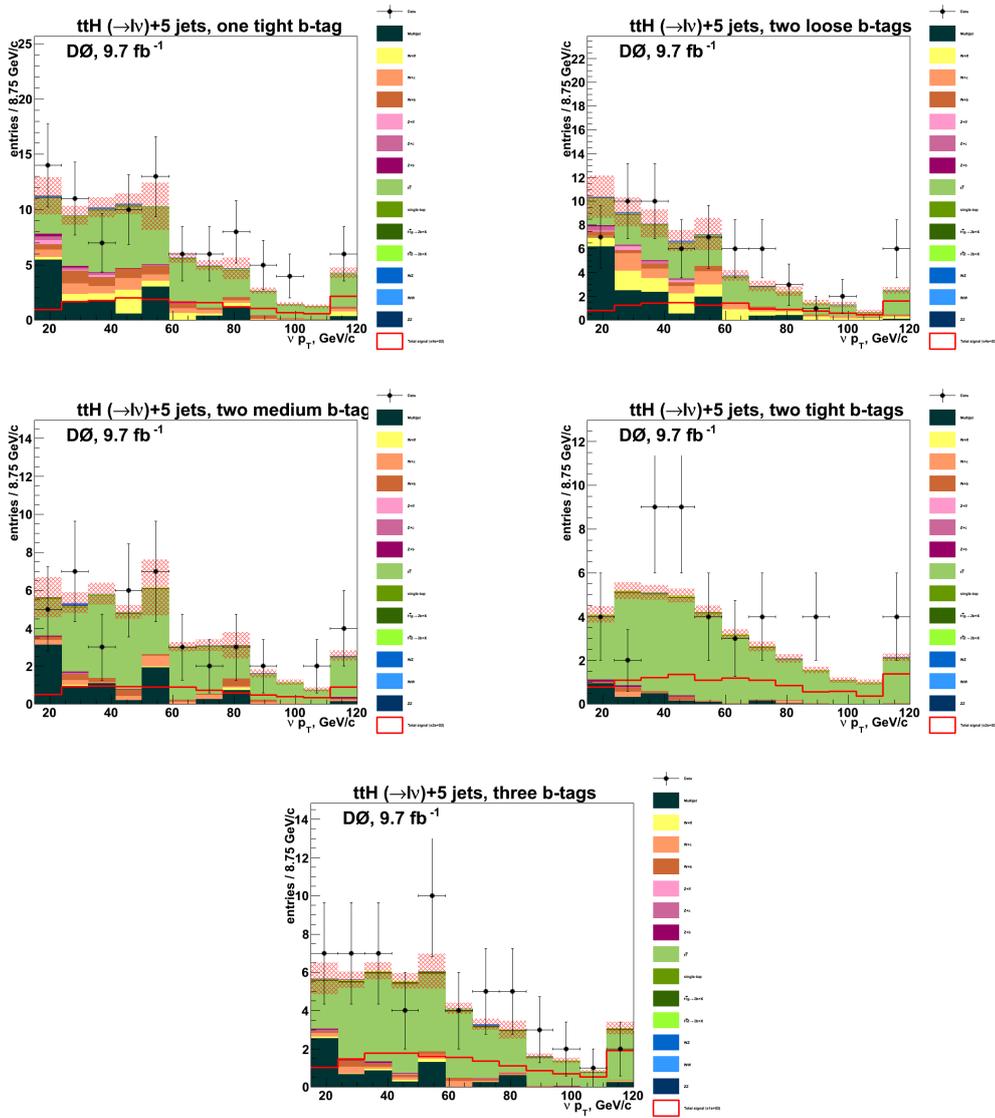


Figure A.9: Validation plots for the transverse momentum p_T of the neutrino for events with 1 electron, exactly 5 jets and A.9(a) 1 tight tag, A.9(b) 2 loose tags, A.9(c) 2 medium tags, A.9(d) 2 tight tags and A.9(e) 3 tags.

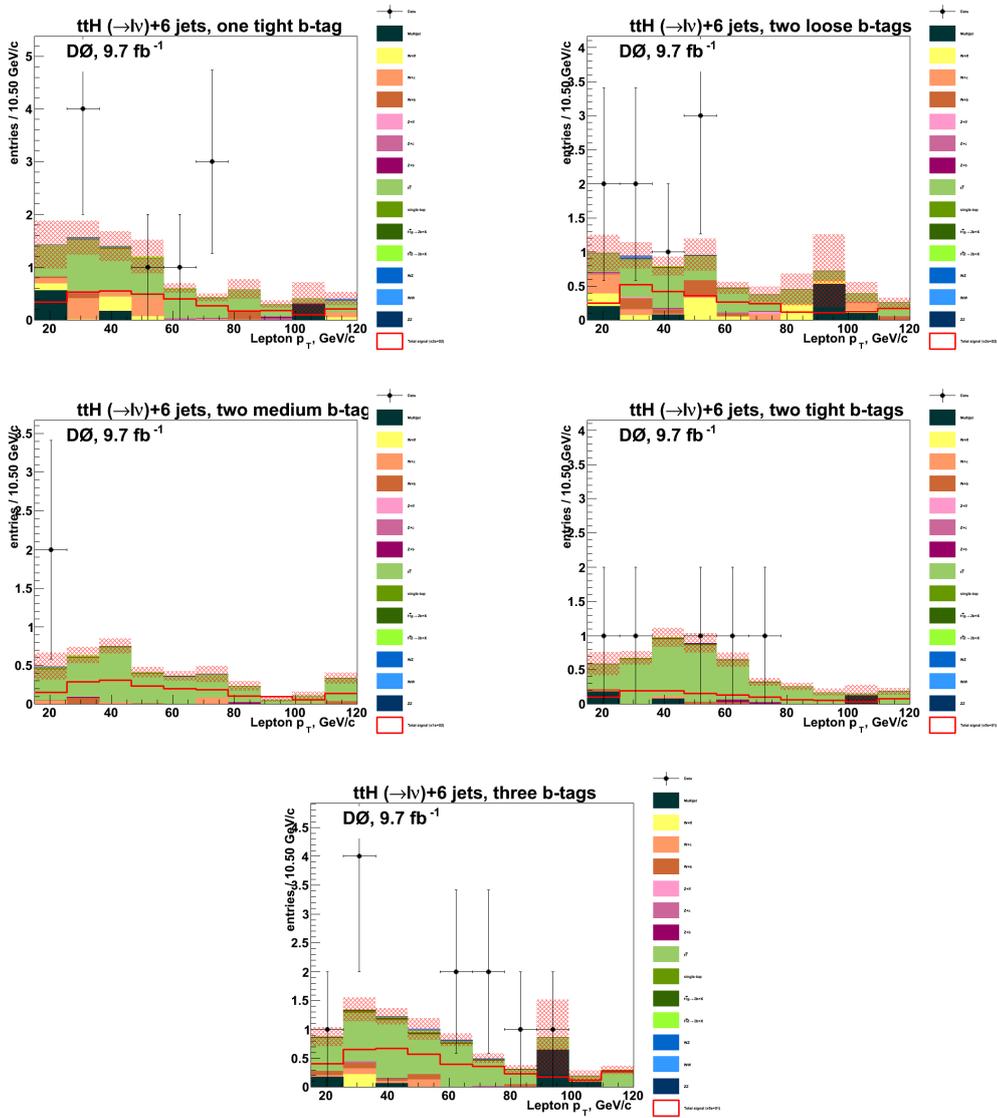


Figure A.12: Validation plots for the transverse momentum p_T of the lepton for events with 1 electron, at least 6 jets and A.12(a) 1 tight tag, A.12(b) 2 loose tags, A.12(c) 2 medium tags, A.12(d) 2 tight tags and A.12(e) 3 tags.

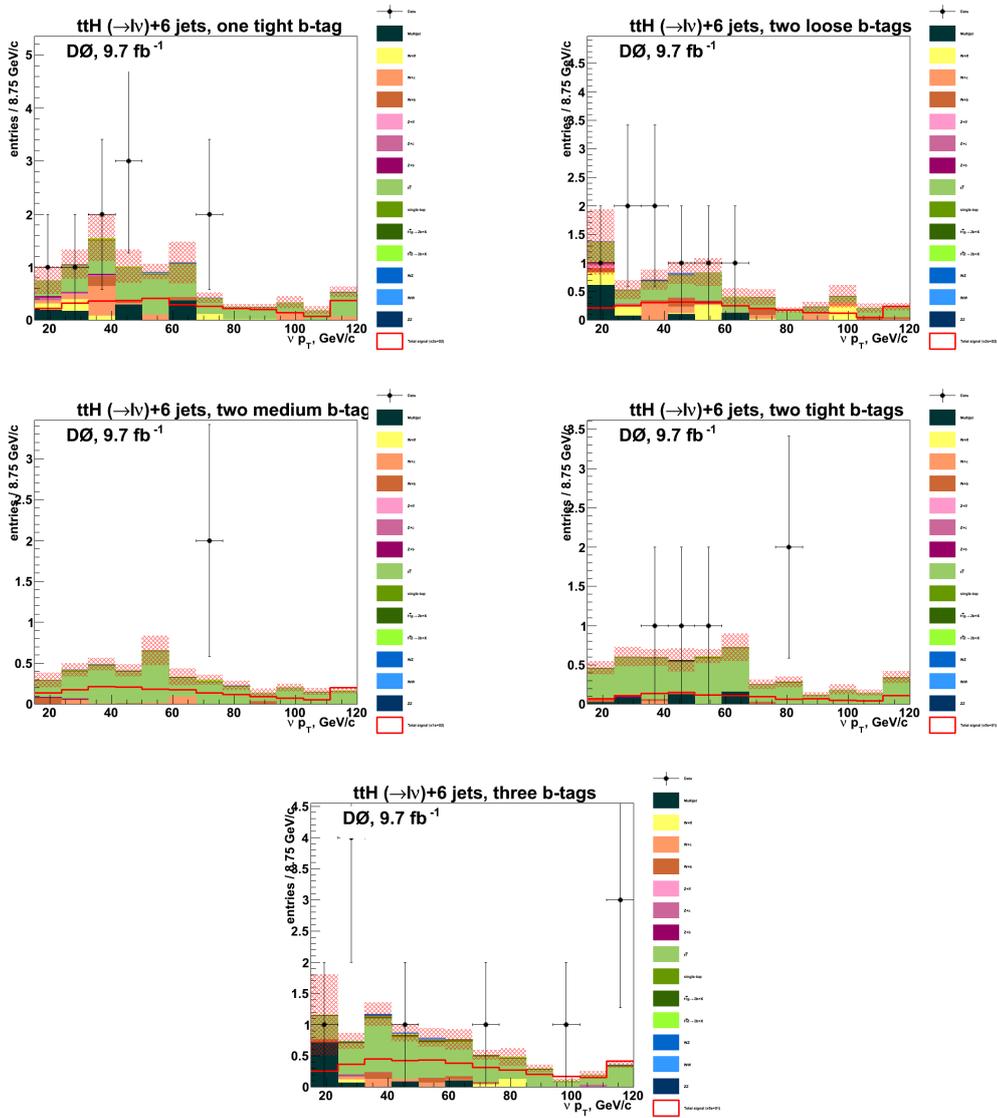


Figure A.13: Validation plots for the transverse momentum p_T of the neutrino for events with 1 electron, at least 6 jets and A.13(a) 1 tight tag, A.13(b) 2 loose tags, A.13(c) 2 medium tags, A.13(d) 2 tight tags and A.13(e) 3 tags.

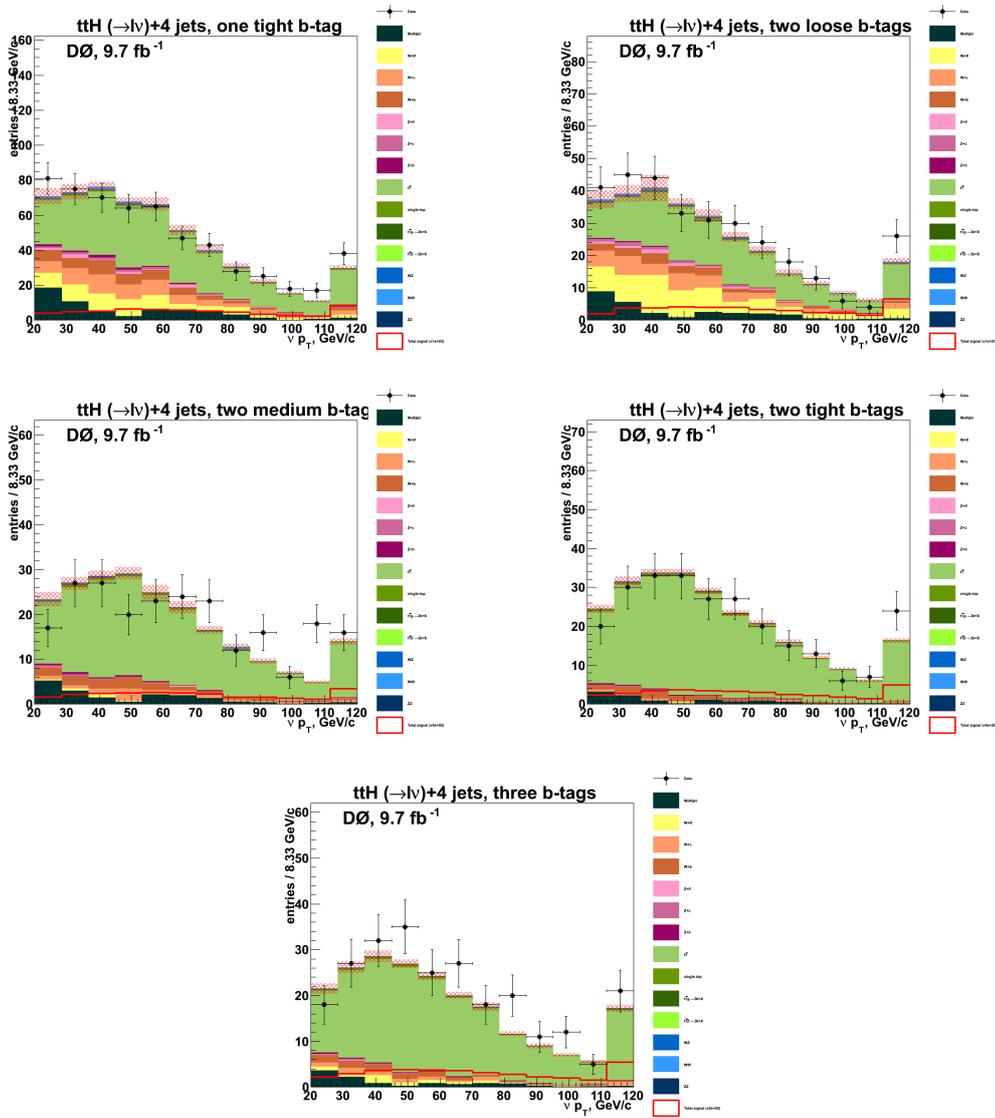


Figure A.17: Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, exactly 4 jets and A.17(a) 1 tight tag, A.17(b) 2 loose tags, A.17(c) 2 medium tags, A.17(d) 2 tight tags and A.17(e) 3 tags.

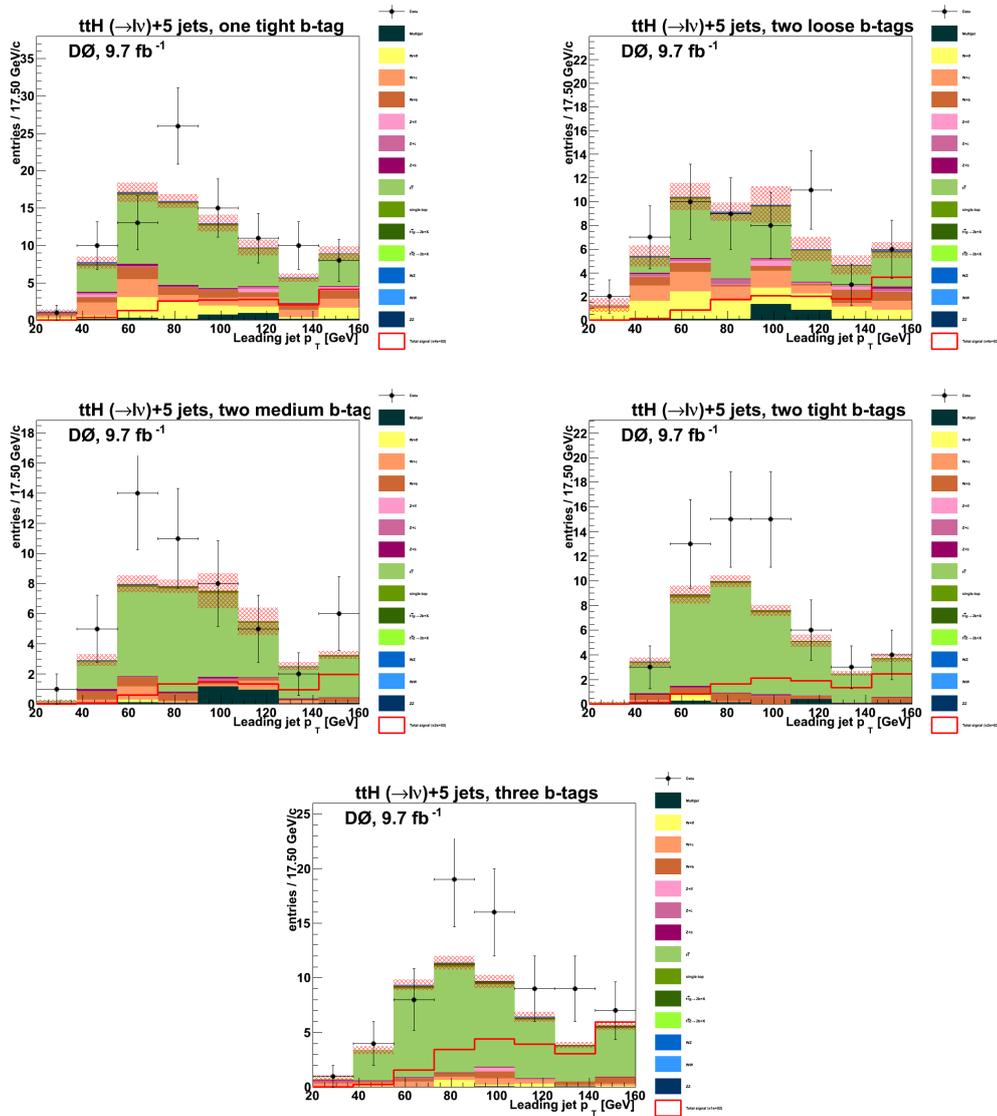


Figure A.18: Validation plots for the jet with highest transverse momentum p_T for events with 1 muon, exactly 5 jets and A.18(a) 1 tight tag, A.18(b) 2 loose tags, A.18(c) 2 medium tags, A.18(d) 2 tight tags and A.18(e) 3 tags.

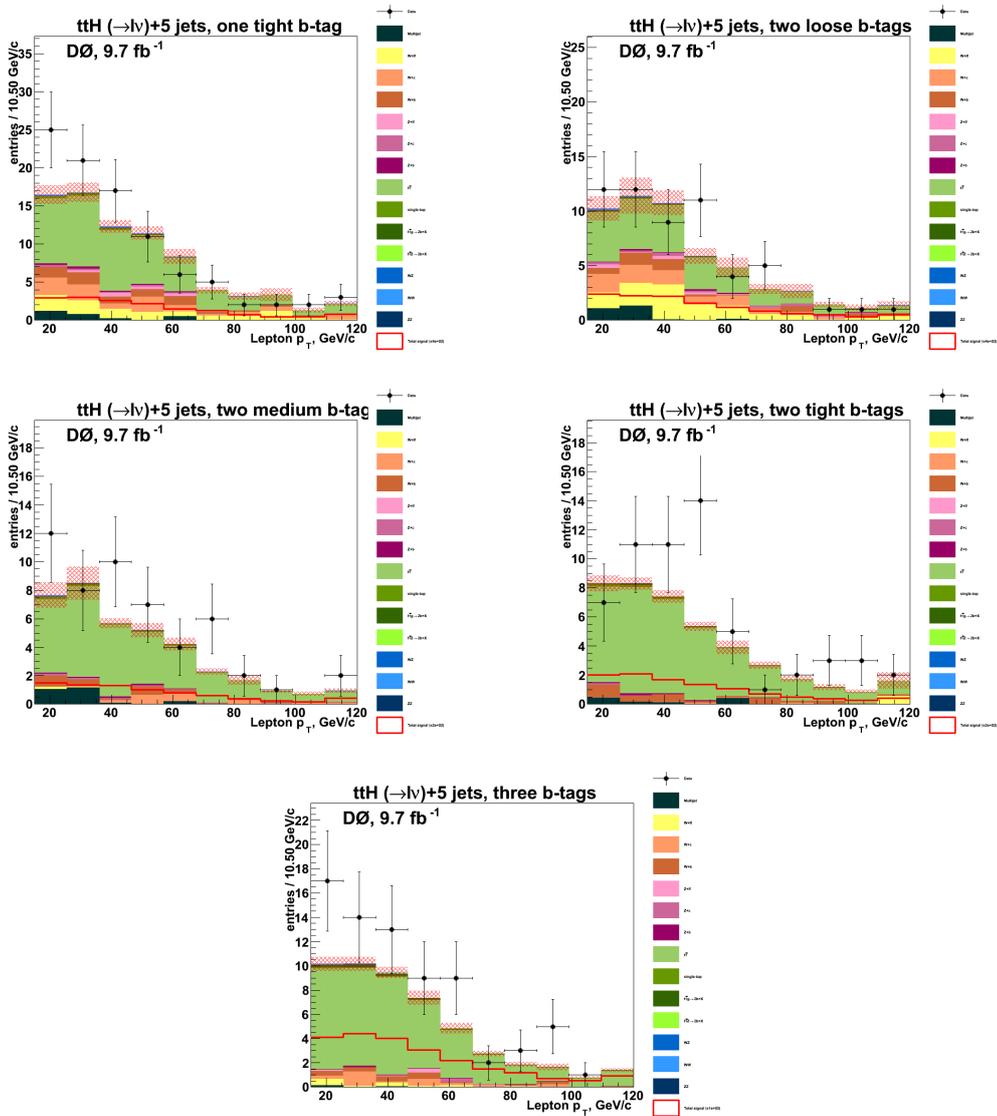


Figure A.20: Validation plots for the transverse momentum p_T of the lepton for events with 1 muon, exactly 5 jets and A.20(a) 1 tight tag, A.20(b) 2 loose tags, A.20(c) 2 medium tags, A.20(d) 2 tight tags and A.20(e) 3 tags.

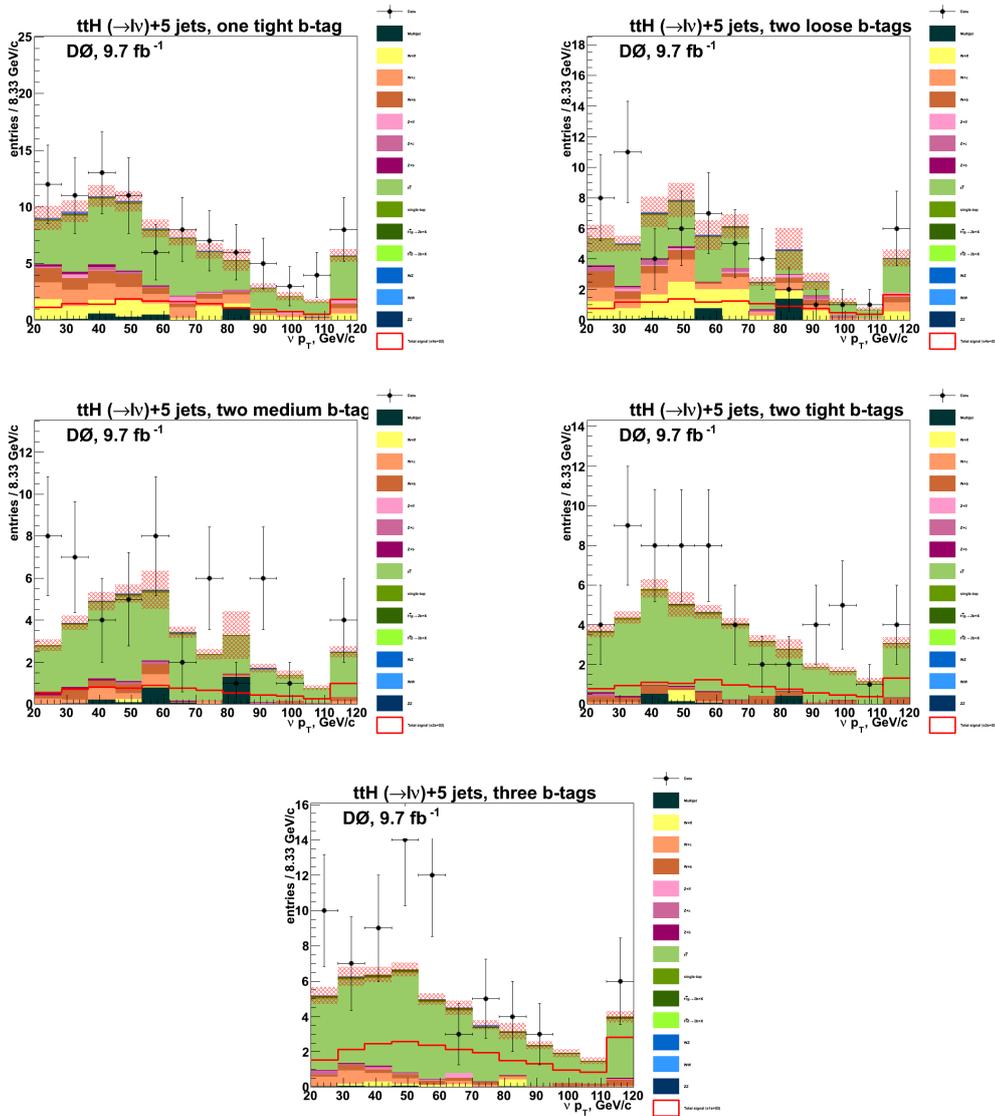


Figure A.21: Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, exactly 5 jets and A.21(a) 1 tight tag, A.21(b) 2 loose tags, A.21(c) 2 medium tags, A.21(d) 2 tight tags and A.21(e) 3 tags.

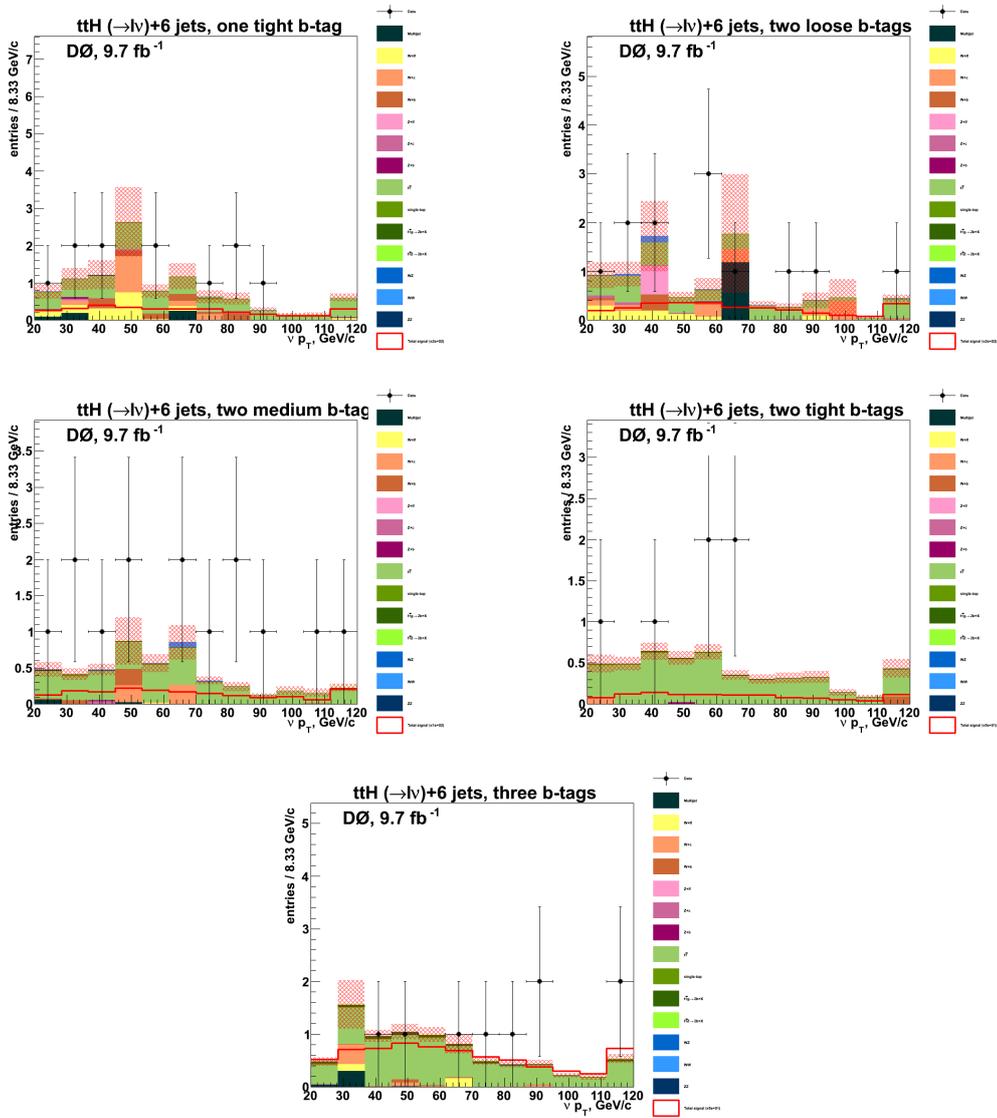
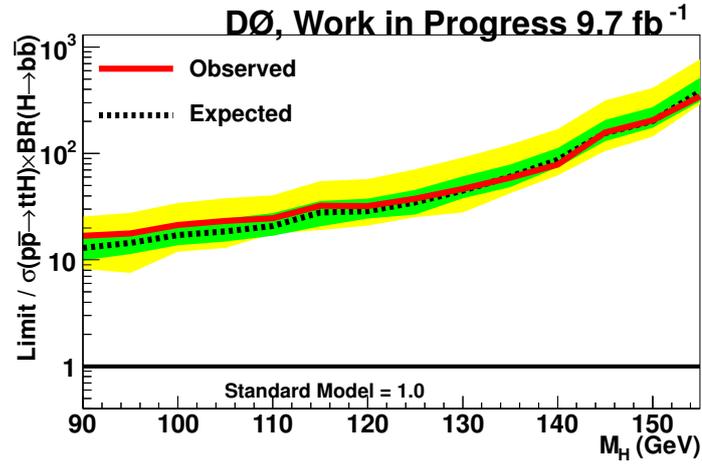


Figure A.25: Validation plots for the transverse momentum p_T of the neutrino for events with 1 muon, at least 6 jets and A.25(a) 1 tight tag, A.25(b) 2 loose tags, A.25(c) 2 medium tags, A.25(d) 2 tight tags and A.25(e) 3 tags.

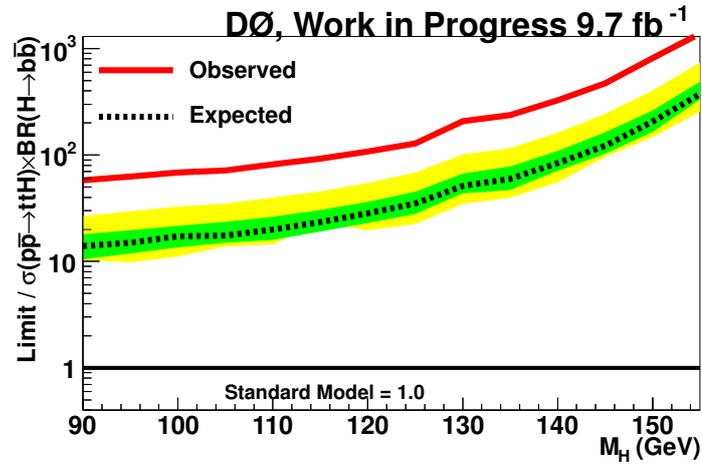
Appendix B

Limit Plots for Individual Lepton Channels

Below are presented separate plots for the electron and muon channels which show the ratio of the $\sigma_{t\bar{t}H}$ cross section times branching ratio limit over the SM NLO prediction with uncertainties. The observed limit is shown in red with the expected limit, defined as the median of the limits obtained in background-only pseudo experiments, as the black dashed line. The 1 and 2 σ uncertainty bands for the expected limit are indicated by the green and green+yellow bands.



(a)



(b)

Figure B.1: The 95% CL upper limit on the $\sigma_{t\bar{t}H}$ cross section times branching ratio over the SM expectation in NLO QCD as a function of the Higgs mass in (a) electron and (b) muon channels.

Appendix C

MVA Output Distributions

Below are presented plots depicting the distributions of the discriminating variables from section 9.2 used in the MVA analysis for both signal (blue) and background (red). The distributions are split into the jet, tagging and lepton categories.

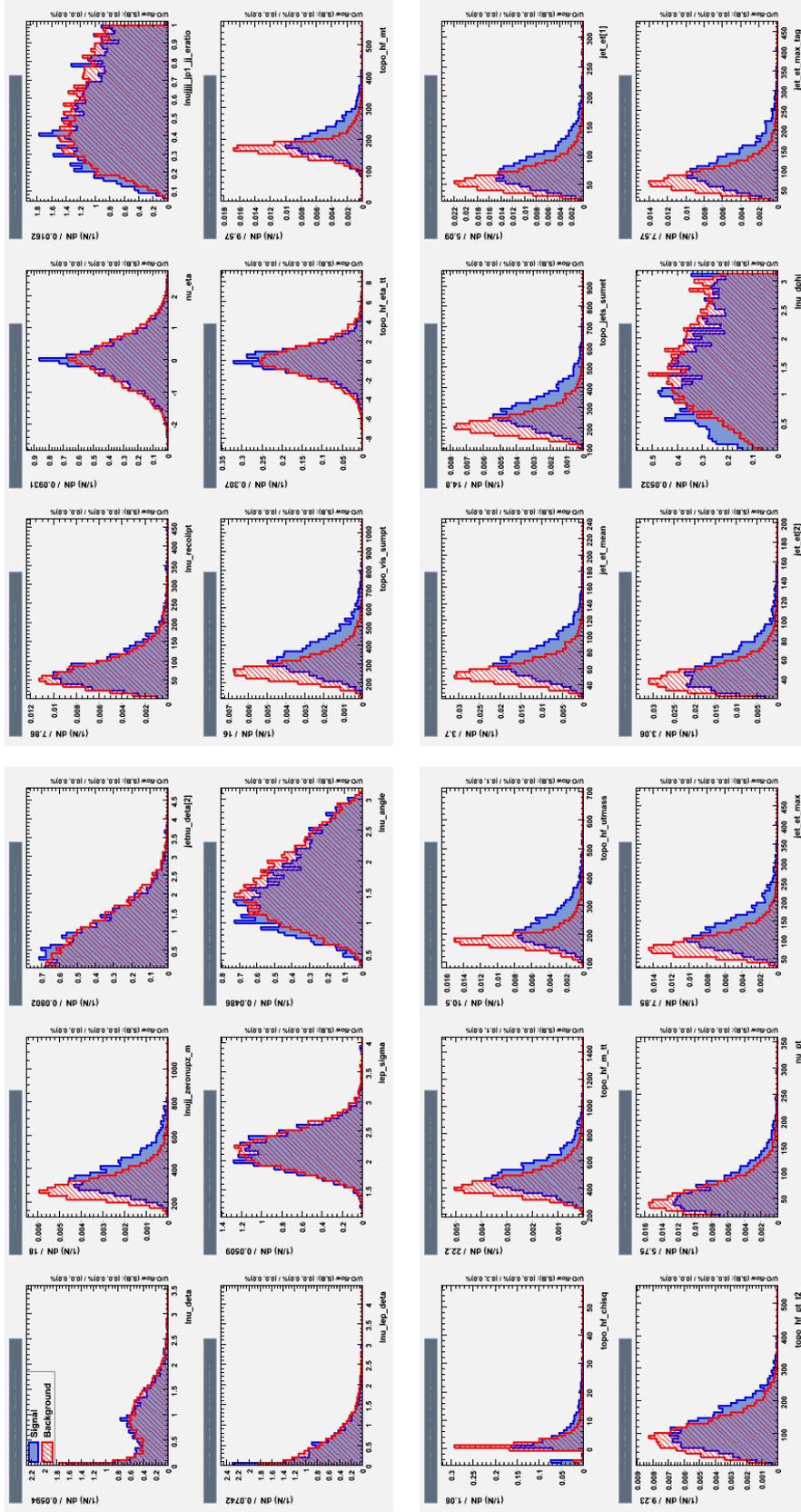


Figure C.1: Variables used for BDT training for 4 jet, 1 tight tag category. Signal distributions are shown in blue and background distributions in red.

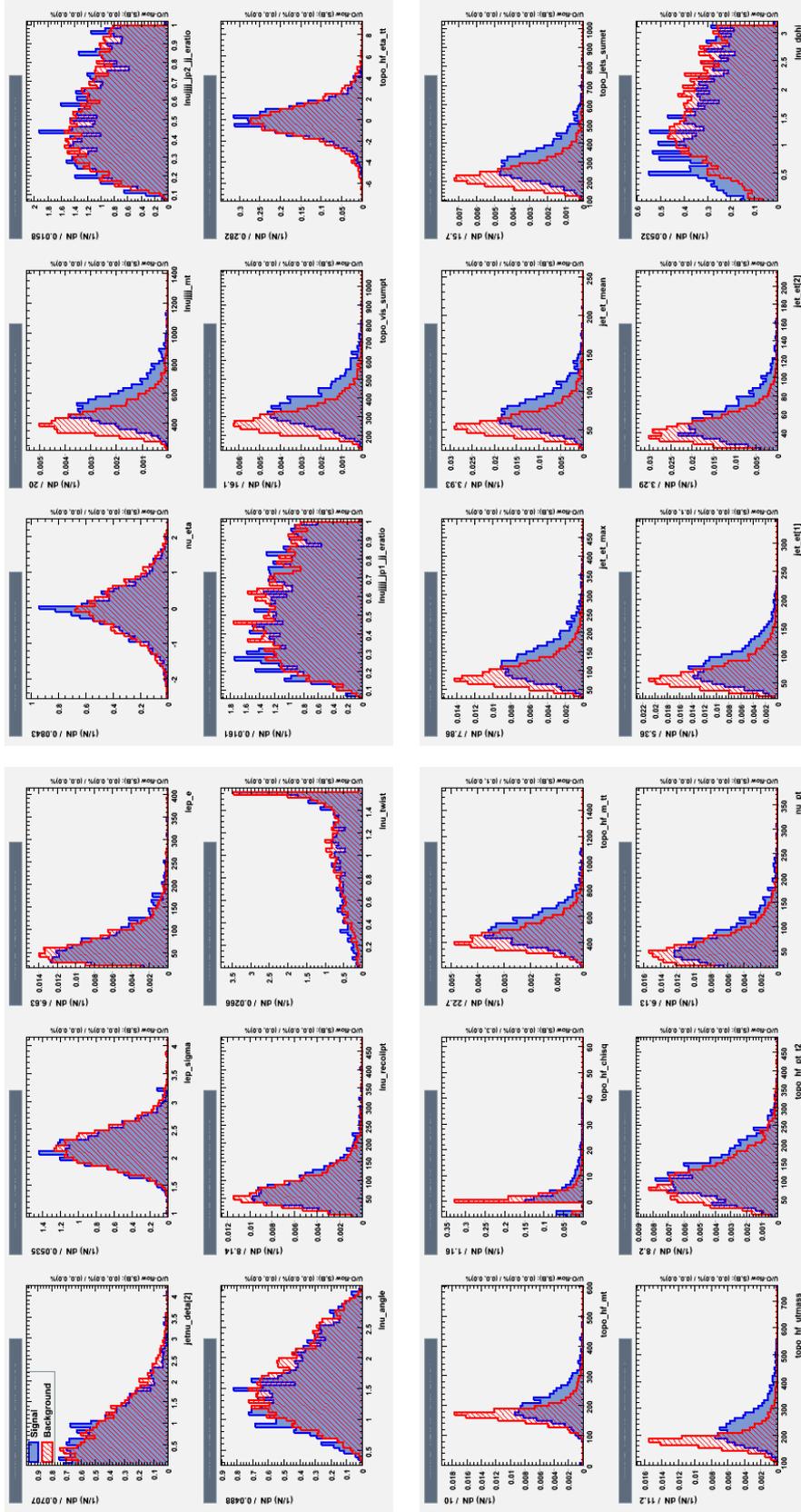


Figure C.2: Variables used for BDT training for 4 jet, 2 loose tag category. Signal distributions are shown in blue and background distributions in red.

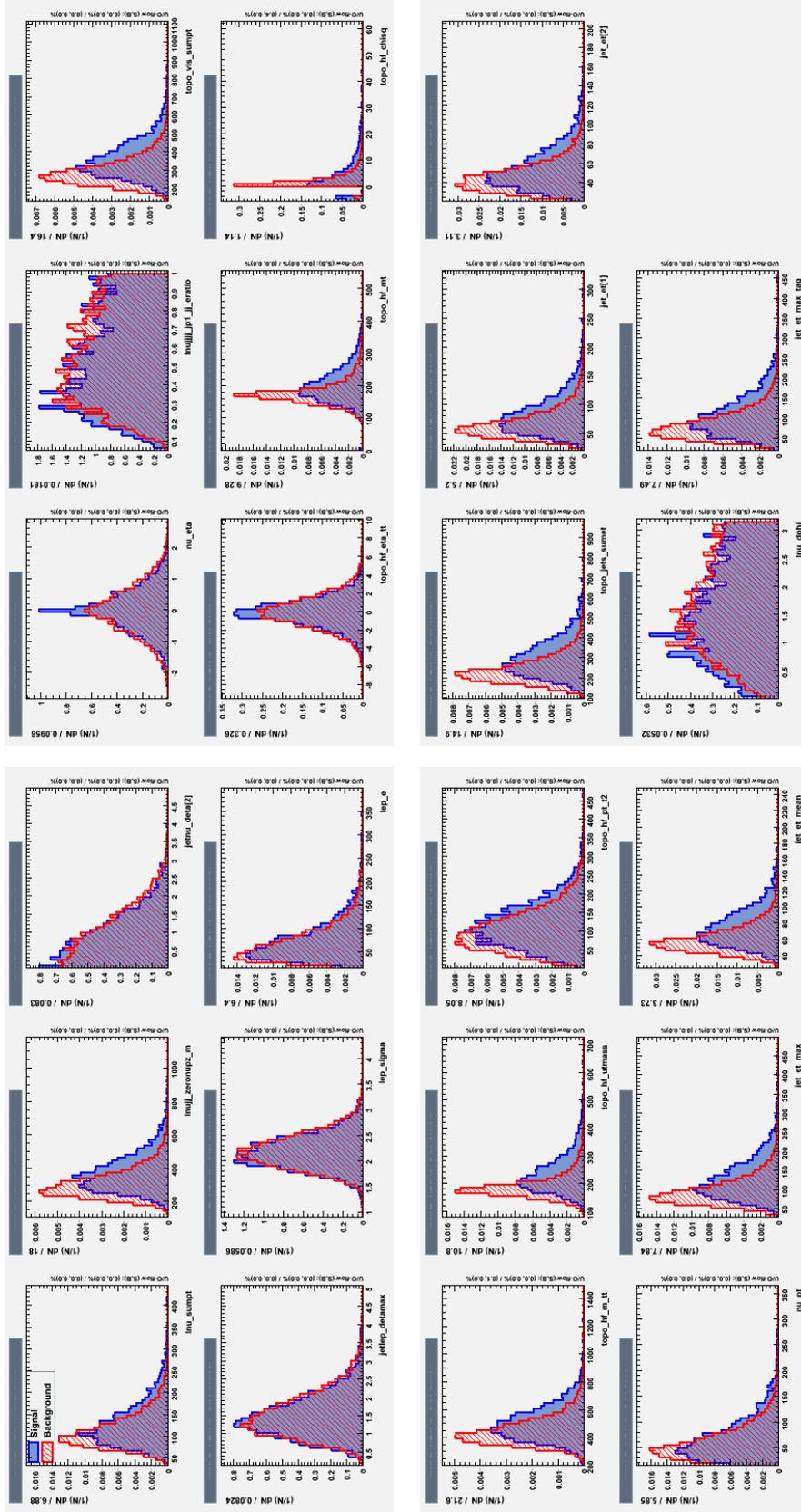


Figure C.3: Variables used for BDT training for 4 jet, 2 medium tag category. Signal distributions are shown in blue and background distributions in red.

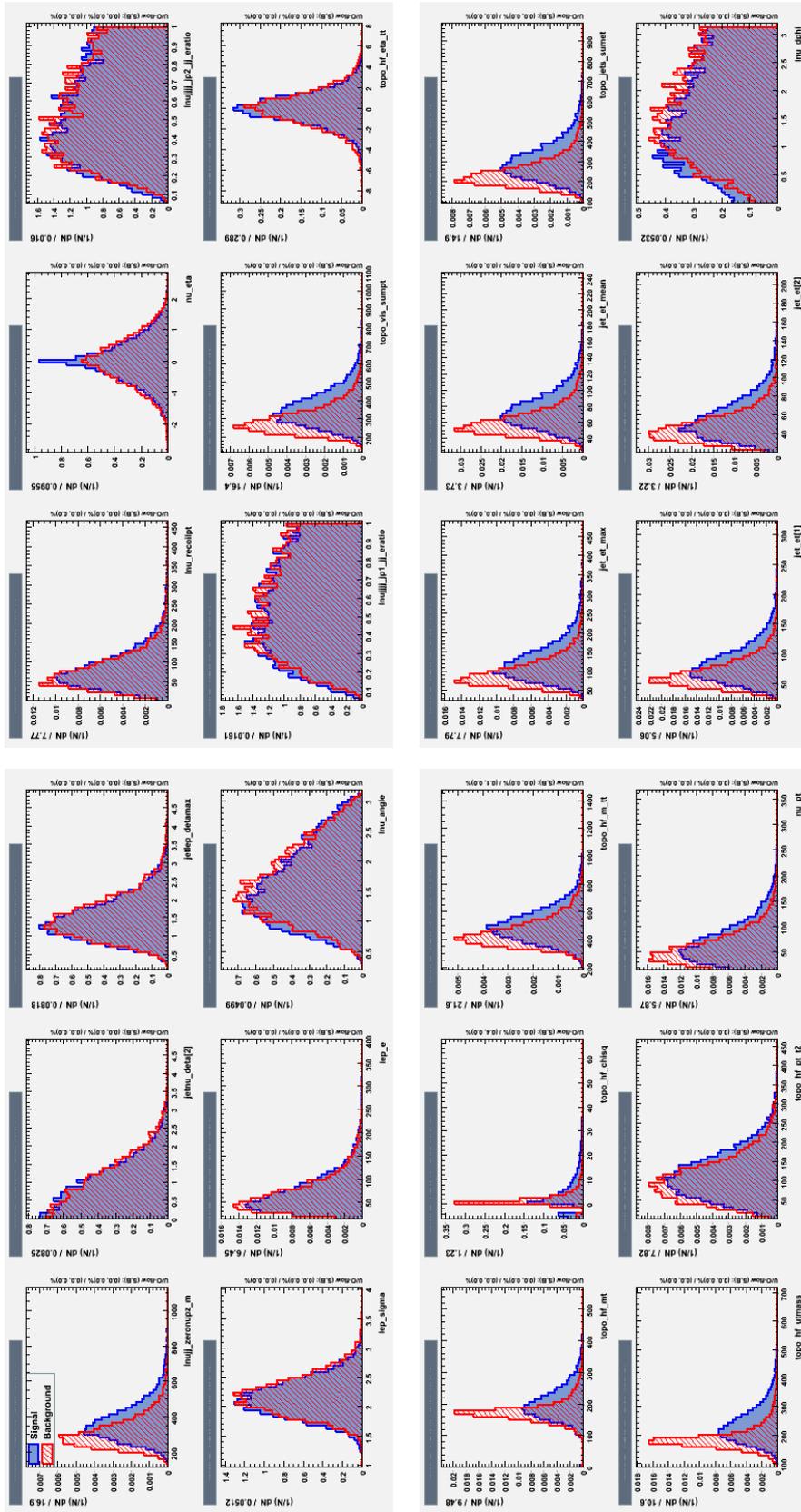


Figure C.5: Variables used for BDT training for 4 jet, 3 tag category. Signal distributions are shown in blue and background distributions in red.

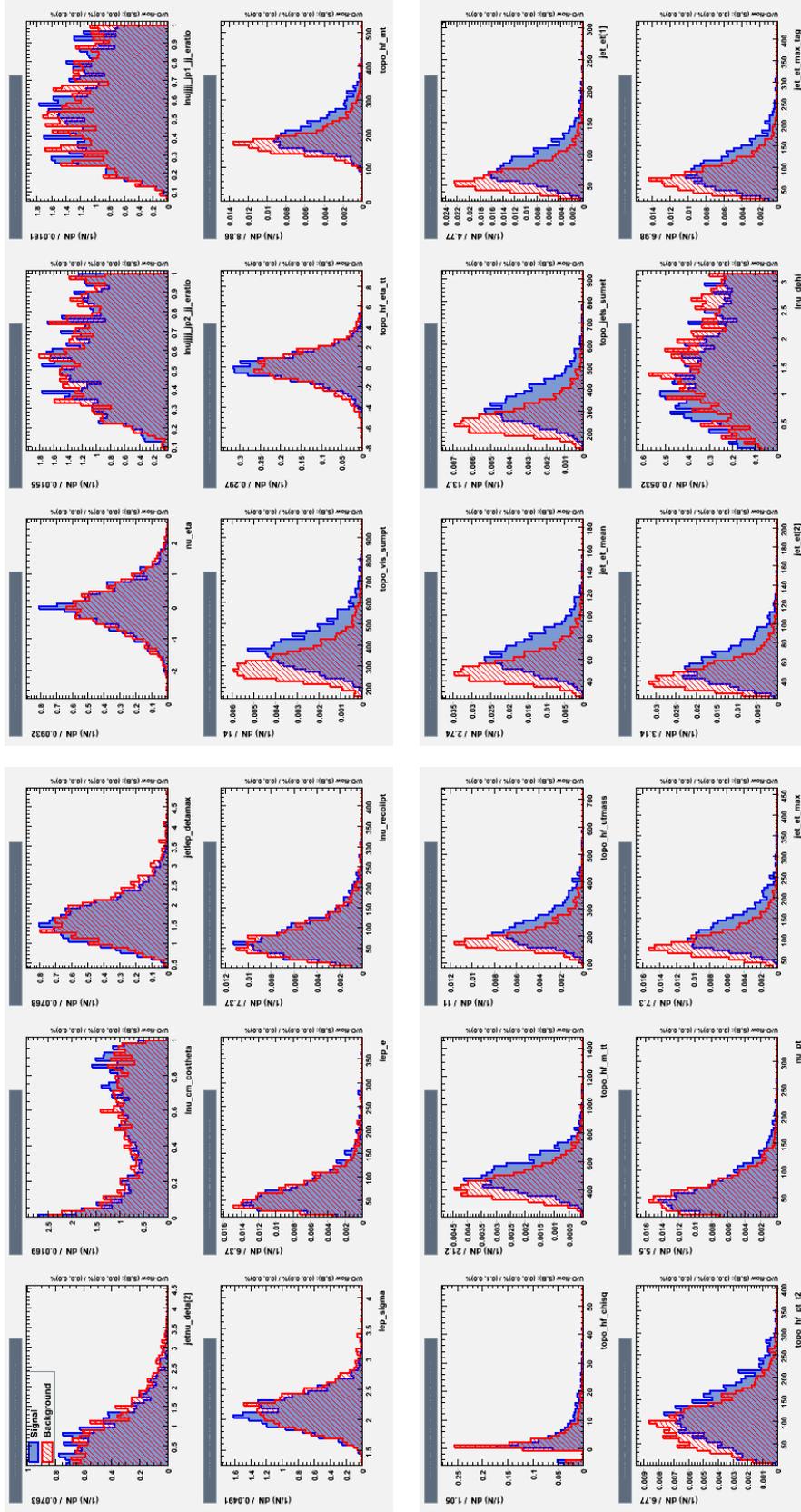


Figure C.6: Variables used for BDT training for 5 jet, 1 tight tag category. Signal distributions are shown in blue and background distributions in red.

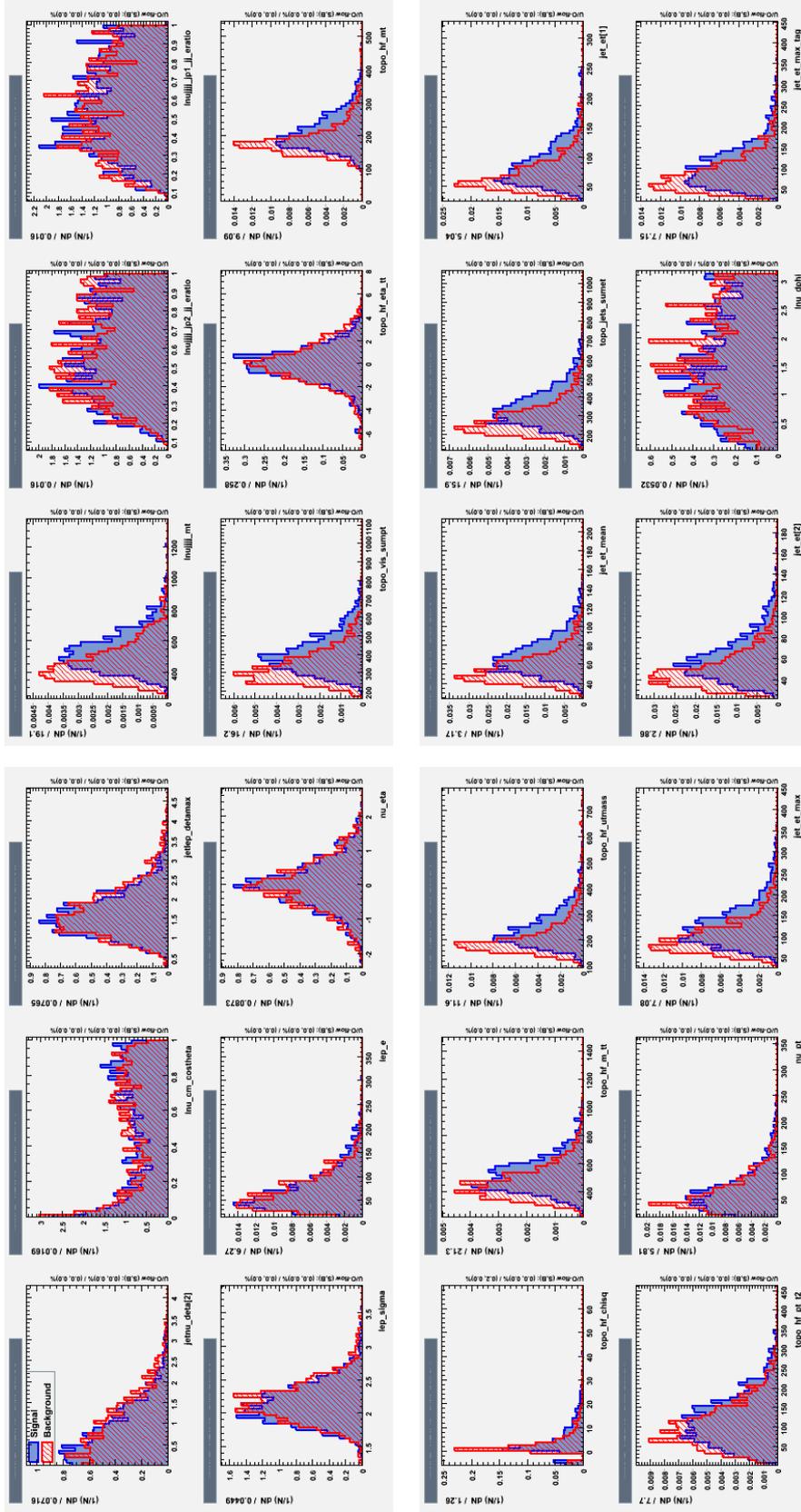


Figure C.7: Variables used for BDT training for 5 jet, 2 loose tag category. Signal distributions are shown in blue and background distributions in red.

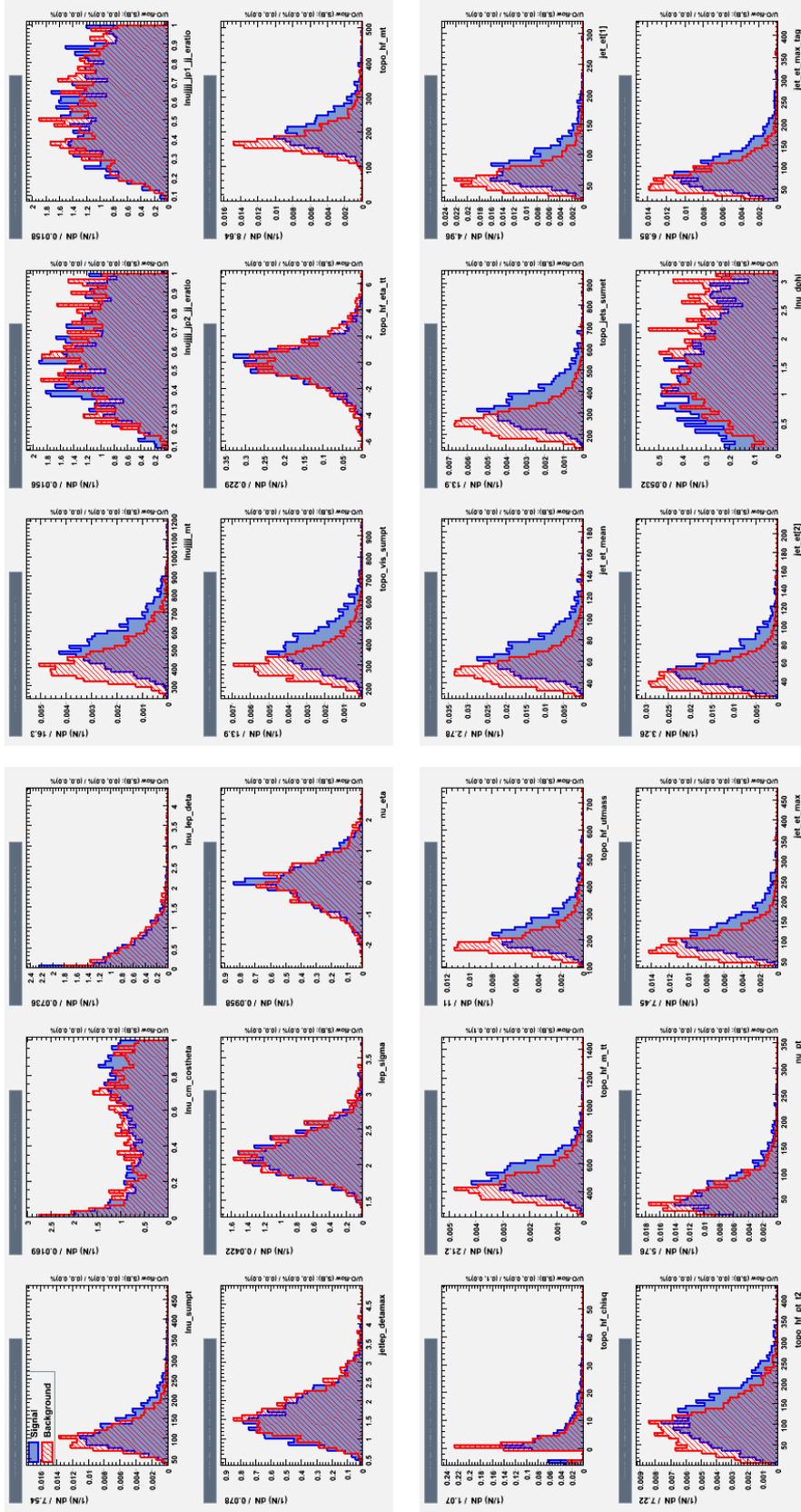


Figure C.8: Variables used for BDT training for 5 jet, 2 medium tag category. Signal distributions are shown in blue and background distributions in red.

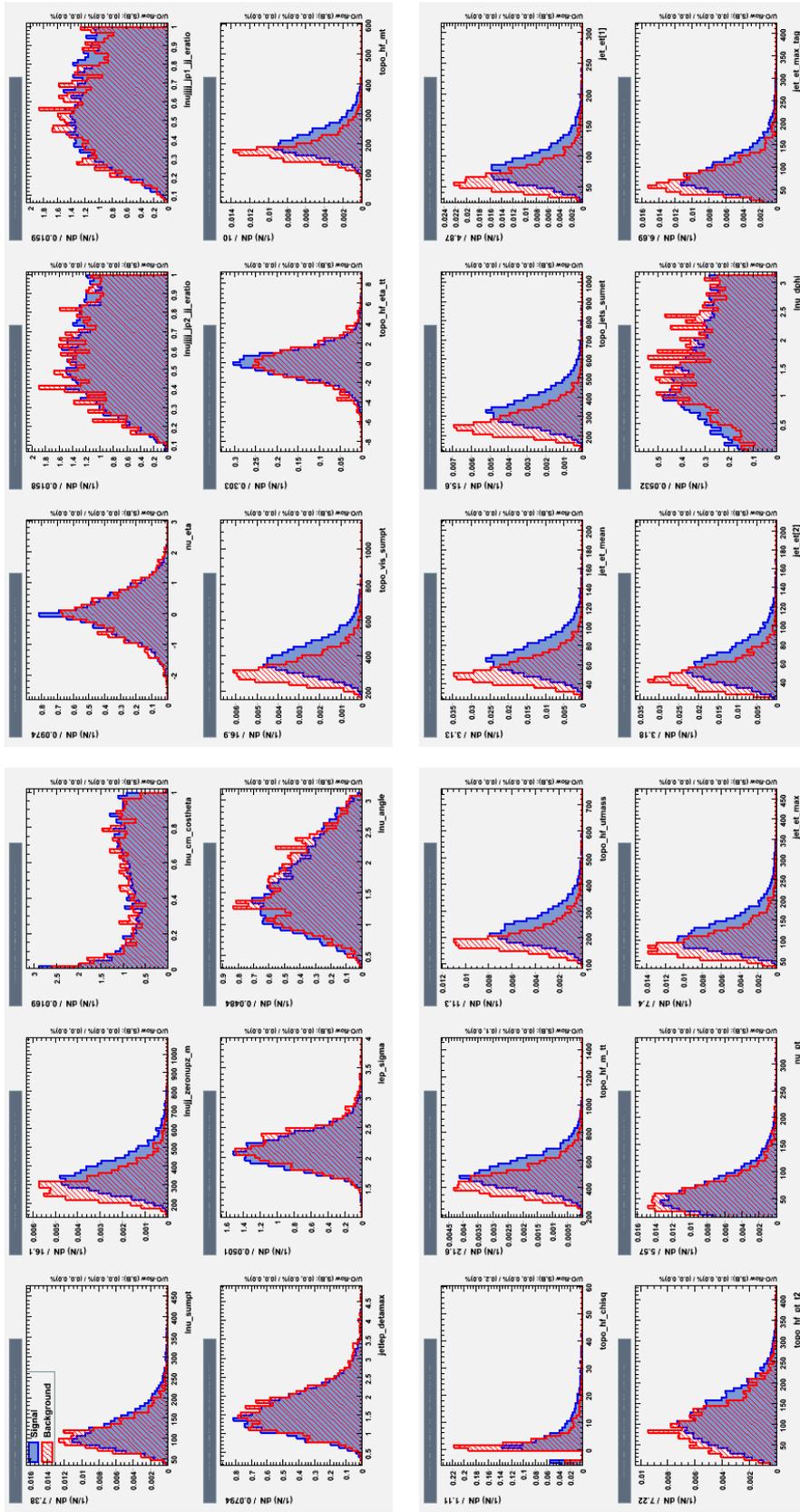


Figure C.10: Variables used for BDT training for 5 jet, 3 tag category. Signal distributions are shown in blue and background distributions in red.

Bibliography

- [1] J. Beringer et al. (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).
- [2] S. L. Glashow, “Partial Symmetries of Weak Interactions”, Nucl. Phys. **22**, 579-588 (1961)
- [3] J. Goldstone, A. Salam, S. Weinberg, “Broken Symmetries”, Phys. Rev. **127**, 965-970 (1962)
- [4] S. Weinberg, “A Model of Leptons”, Phys. Rev. Lett. **19**, 1264-1266 (1967)
- [5] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, Physical Review Letters **13** (9), 321-323 (1964)
- [6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, Physical Review Letters **13** (20) 585-587 (1964)
- [7] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, Physical Review Letters **13** (16), 508509 (1964)
- [8] University of Edinburgh School of Physics and Astronomy, http://www.ph.ed.ac.uk/sites/default/files/images/higgs/YRHS_BR_fig1.jpg
- [9] http://www.physikblog.eu/wp-content/uploads/2011/04/higgs-gammagamma_egede.png
- [10] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012)

- [11] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012)
- [12] T. Aaltonen, V. M. Abazov *et al.* [CDF and D0 Collaborations] “Higgs Boson Studies at the Tevatron”, Phys. Rev. D **88** 052014 (2013)
- [13] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963)
- [14] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [15] M. Muether *et al.* [Tevatron Electroweak Working Group and CDF and D0 Collaborations], “Combination of CDF and D0 results on the mass of the top quark using up to 8.7 fb^{-1} at the Tevatron”, arXiv:1305.3929 [hep-ex] (2013).
- [16] M. Perelstein, M. E. Peskin and A. Pierce, “Top quarks and electroweak symmetry breaking in little Higgs models,” Phys. Rev. D **69** (2004) 075002 [hep-ph/0310039].
- [17] B. A. Dobrescu, K. Kong and R. Mahbubani, “Prospects for top-prime quark discovery at the Tevatron,” JHEP **0906** (2009) 001 [arXiv:0902.0792 [hep-ph]].
- [18] Ann Heinson,
http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html
- [19] http://www.fnal.gov/pub/today/archive/archive_2008/today08-06-18.html
- [20] Ann Heinson,
http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_dzero_detector.html
- [21] R. Angstadt *et al.*, “The Layer 0 Inner Silicon Detector of the D0 Experiment”, Nucl. Instrum. Methods in Phys. Res. Sect. A, **622** (298):298-310 (2010)
- [22] V.M. Abazov *et al* [D0 Collaboration], “The Upgraded D0 detector”, Nucl. Instrum. Meth., **A565**:463-537, (2006)

- [23] B. C. K. Casey, M. Corcoran, K. DeVaughan, Y. Enari, E. Gallas, I. Katsanos, J. Linnemann and J. Orduna *et al.*, “The D0 Run IIB Luminosity Measurement”, Nucl. Instrum. Meth. A **698**, 208 (2013)
- [24] A. Khanov, “HTF: histogramming method for finding tracks, the algorithm description”, **D0 Note 3778** (2000)
- [25] G. Borissov, “AA track finding algorithm”, Presentation given at D0 plenary meeting on 28th February 2003
- [26] Diehl et al, “Tracking Algorithm Recommendation Committee (TARC) Report II” (2003)
- [27] H. Greenlee, “The D0 Kalman Track Fit”, **D0 note 4304** (2004)
- [28] H. Greenlee, “The D0 Interacting Propagator” **D0 Note 4293** (2003)
- [29] A. Schwartzman and C. Tully. “Primary vertex reconstruction by means of adaptive vertex fitting”, **D0 Note 4918** (2005)
- [30] G. Bernardi, E. Busato, and J. Vlimant, “Improvements from the T42 algorithm on calorimeter objects reconstruction”, **D0 Note 4335** (2004)
- [31] T. Head, “Top quark spin correlations and leptonic forward-backward asymmetries at D0”, PhD thesis, University of Manchester (2012)
- [32] ROOT <http://root.cern.ch>.
- [33] O. Brandt et al., “Muon Identification Certification for the Summer 2009 Extended Dataset (Run IIB-1 and -2)”, **D0 Note 6025** (2010)
- [34] A. Harel, “Jet ID Optimization”, **D0 Note 4919** (2006)
- [35] http://www-d0.fnal.gov/phys_id/jets/jetid.html

- [36] U. Bassler and G. Bernardi, “Towards a Coherent Treatment of Calorimetric Energies: Missing Transverse Energy, Jets, E.M. Objects and the T42 Algorithm”, **D0 Note 4124** (2003)
- [37] V. M. Abazov *et al* [D0 Collaboration], “b-Jet Identification in the D0 Experiment” Nucl. Instrum. Meth. A **620** 490 (2010)
- [38] S. Greder and I. Ripp-Baudot “B-tagging with combined multivariate techniques”, **D0 Note 5936** (2011)
- [39] D. Bloch et al, “Performance of the JLIP b-tagger in p14”, **D0 Note 4348** (2004)
- [40] http://d0server1.fnal.gov/Projects/Operations/D0RunII_DataTaking_files/image005.png
- [41] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003)
- [42] T. Sjostrand, L. Lonnblad, S. Mrenna, and P. Z. Skands, (2003)
- [43] H. L. Lai et al., Phys. Rev. **D55**, 1280 (1997)
- [44] J. Pumplin et al., JHEP **07**, 012 (2002)
- [45] The Tevatron New Phenomena and Higgs Working Group, CDF Note 10474 (2012)
- [46] J. Goldstein et al., “Prospects of Observing $t\bar{t}H$ at RunII: A discovery mode for the Higgs boson?”, International Journal of Modern Physics A, Vol. 16, Suppl. 1A, 308-310 (2001)
- [47] F. Maltoni and T. Stelzer, “MadEvent: Automatic event generation with MadGraph,” JHEP **0302** 027 (2003)

- [48] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira and P. M. Zerwas, “NLO QCD corrections to t anti- t H production in hadron collisions,” Nucl. Phys. B **653**, 151 (2003) [hep-ph/0211352].
- [49] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackerth, “Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections,” Phys. Rev. D **68**, 034022 (2003) [hep-ph/0305087].
- [50] Y. Fisyak and J. Womersley, “D0gstar D0 GEANT Simulation of the Total Apparatus Response”, **D0 note 3191** (1997)
- [51] R. Brun and F. Carminati, “GEANT Detector Description and Simulation Tool”, CERN Program Library Long Writeup W5013, unpublished (1993)
- [52] E. Boos et al., “CompHEP”, Nucl. Instrum. Meth. **A534**, 250 (2004)
- [53] C. Gerber et al., “Studies on the instrumental background in electron+jets+met final state selections using p17 Monte Carlo”, **D0 Note 5955** (2009).
- [54] V. M. Abazov *et al* [D0 Collaboration], “Search for the Standard Model Higgs boson in $l\nu$ +jets final states in 9.7 fb^{-1} of $p\bar{p}$ collisions with the D0 detector”, Phys. Rev. D, **88** 052008 (2013)
- [55] T. Scanlon https://plone4.fnal.gov/P1/D0Wiki/object-id/bid/TRFS_Moriond2012 (2011)
- [56] S. Moch and P. Uwer, “Theoretical status and prospects for top-quark pair production at hadron colliders”, Phys. Rev. D, **78** (2008)
- [57] T. Aaltonen *et al.* [CDF and D0 Collaborations], “Combination of the top-quark mass measurements from the Tevatron collider,” Phys. Rev. D **86** 092003 (2012)
- [58] A. L. Read, “Modified frequentist analysis of search results (the CL_s method)”, CERN-OPEN-2000-205 (2000)

- [59] W. Fisher, FERMILAB-TM-2386-E
- [60] V. M. Abazov *et al* [D0 Collaboration], “Search for the Standard Model Higgs boson in the $t\bar{t}H \rightarrow t\bar{t}b\bar{b}$ channel”, **D0 Note 5739-CONF** (2009)
- [61] S. Chatrchyan *et al.* [CMS Collaboration], “Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC,” **JHEP 1305** (2013) 145 [arXiv:1303.0763 [hep-ex]].
- [62] T. Aaltonen *et al.* [CDF Collaboration], “Search for the standard model Higgs boson produced in association with top quarks using the full CDF data set,” *Phys. Rev. Lett.* **109** (2012) 181802 [arXiv:1208.2662 [hep-ex]].
- [63] S. Dittmaier *et al.* [LHC Higgs Cross Section Working Group Collaboration], “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,” arXiv:1101.0593 [hep-ph].
- [64] V. M. Abazov *et al* [D0 Collaboration], “Evidence for s-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”, Accepted to *Phys. Lett. B* (2013)
- [65] S. Snyder, “Measurement of the Top Quark Mass at D0”, Ph.D. thesis, Stony Brook University (1995)