

Search for the Higgs Boson Produced in Association with Top Quarks and Decaying into Bottom Quarks with the ATLAS Detector

Shunsuke Honda



February 2018

**Search for the Higgs Boson Produced in
Association with Top Quarks and Decaying into
Bottom Quarks with the ATLAS Detector**

Shunsuke Honda

Doctoral Program in Physics

**Submitted to the Graduate School of
Pure and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Science**

**at the
University of Tsukuba**

Abstract

The Higgs boson coupling to the top quarks (top Yukawa-coupling) is one of the most important quantities to be measured experimentally. The top quark has by far the heaviest mass among the Standard Model (SM) particles and with a value approximately $m_{\text{top}} \sim v/\sqrt{2}$ ($v = 246$ GeV). The corresponding top Yukawa-coupling is almost unity, which may imply a special relationship between the Higgs boson and the top quark. Higgs boson production in association with a pair of top quarks, $t\bar{t}H$, offers a unique production mode allowing a direct measurement of the top Yukawa-coupling. Although the production cross-section is only $\sim 1\%$ of the total Higgs boson production, the $t\bar{t}H$ process offers a distinct signature due to the numerous final state objects that depend on the Higgs boson and $t\bar{t}$ decays. Using the largest branching ratio for the Higgs boson, $H \rightarrow b\bar{b}$ of 58%, the search for $t\bar{t}H$ production was performed in this thesis. In order to trigger signal events with significant rejections of QCD and multi-jet backgrounds, the $t\bar{t}H(H \rightarrow b\bar{b})$ channel with at least one lepton emitted from $t\bar{t}$ decay was chosen. The difficulties for this channel include a low efficiency in reconstructing all final-state objects, combinatorial ambiguity from multiple b -jets in the final state, and large backgrounds from production of $t\bar{t}$ plus additional heavy flavor jets such as b -jets and c -jets. The $t\bar{t}$ plus jets background is difficult to model and gives essentially the same kinematics as the $t\bar{t}H$ process.

The analysis used 36.07 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV, collected with the ATLAS detector at the Large Hadron Collider, corresponding to the full pp data sets from 2015-2016 runs. The developed analysis is characterized by the following points, (1) categorize events according to the jet and the b -tagged jets with the multiplicity and b -tagging working points to maximize the signal sensitivity; (2) utilize two $t\bar{t}H$ reconstruction algorithms, one using both Higgs and $t\bar{t}$ kinematics, while the other using only $t\bar{t}$ kinematics to provide non-biased Higgs kinematics; (3) utilize two discriminant variables for separating the $t\bar{t}H$ signal and the $t\bar{t} + b\bar{b}$ background using a likelihood and a matrix-element method; (4) adopt a multi-variate analysis to obtain the best sensitivity constructed by likelihood and matrix-element discriminants; (5) perform a global fit which simultaneously determines the contributions from the $t\bar{t}H$ signal and the major backgrounds.

The $t\bar{t}H$ signal strength, defined as the ratio of the measured $t\bar{t}H$ cross section to the Standard Model expectation is measured to be $\mu = 0.84^{+0.64}_{-0.61}$ assuming a Higgs boson mass of 125 GeV. In the result combining all $t\bar{t}H$ analysis channels, the $t\bar{t}H$ signal strength is measured to be $1.17^{+0.34}_{-0.32}$. This corresponds to an observed significance of 4.2σ , while 3.7σ was expected from a simulation predicted by the SM. The observed $t\bar{t}H$ production and decay rates are also interpreted in a coupling framework to evaluate the top Yukawa-coupling, which are found to be compatible with the SM prediction.

Acknowledgements

I would like to thank all the people who supported this analysis and my activity in the doctoral program. I would like to express the most sincere gratitude to my supervisor Kazuhiko Hara for his greatful advice and encouragements since I started to belong to the experimental high energy physics laboratory in the University of Tsukuba. He always gave clear and pertinent comments and suggestions for not only my studies but everything during my research life. He also often gave me jokes. He is certainly a great physicist.

I also really appreciate to Yuji Enari, Koji Nakamura, and Yasuyuki Horii who started to analyze $t\bar{t}H(H \rightarrow b\bar{b})$ with me and gave a lot of great advice through my study. They discussed many topics with me, and offered the farsighted strategies and suggestions to improve our analysis. I am really encouraged in every meeting by them, and admire their impressive ideas and passions in discussions. Without their helps, I could not contribute so much to the ATLAS analysis group.

I would like to thank ATLAS $t\bar{t}H(H \rightarrow b\bar{b})$ and HTop working group leaders: Michele Pianisimo, Jahred Adelman, Elizaveta Shabalina, Marine Kuna who organized our analysis group, gave intuitive comments and suggestions to me. I learned a lot of things from discussions with them. Georges Aad and Snezana Nektarijevic also gave a lot of suggestions and comments to my study and answered my naive askings. Of course, I am also thankful to all people in the ATLAS $t\bar{t}H(H \rightarrow b\bar{b})$ working group members.

I participated in the operation team of the ATLAS semiconductor tracker (SCT) during my stay at CERN. I am grateful to all people in the group, and especially would like to thank Dave Robinson and Hidetoshi Otono with their lots of discussions and supports to work on the SCT operation. I had a lot of invaluable experiences in the SCT group. The summary talk of the SCT in the "ATLAS week" was one of the best experiences for me.

I also really appreciate to all staffs in the University of Tsukuba ATLAS group. Fumihiko Ukegawa organized our Tsukuba ATLAS group and gave lots of comments to understand my study with deep knowledges. Koji Sato also gave meaningful and incisive comments. I often found some points to be left out, and investigated with his comments and advises. Hideki Okawa organized my qualification task. He also often made opportunities to think deeply why its analysis method was adopted. I also would like to thank all staffs in my laboratory. Especially, Yuji Takeuchi discussed various topics with me about the particle interactions and properties, which gave me lots of deep understandings and attractions to the particle physics. Shinhong Kim often gave me meaningful advice with his deep insight in coffee breaks.

I also want to thank the ATLAS collaboration.

Last but not the least, I would like to thank my parents, sister and brother with their great emotional supports for my life.

Contents

1	Standard Model and Higgs Boson	1
1.1	The Standard Model	1
1.2	Brout-Englert-Higgs Mechanism	6
1.2.1	Higgs boson and vector bosons	6
1.2.2	Yukawa Couplings	9
1.3	Standard Model Particles	9
1.4	Higgs Boson	11
1.5	Top Quark	13
1.6	Higgs Boson and Top Quark	18
1.7	Top-Yukawa-Coupling and Beyond the SM	20
1.8	$t\bar{t}H$ Production Analysis	25
1.8.1	$t\bar{t}H(H \rightarrow b\bar{b})$ Analysis	25
2	LHC and ATLAS Detector	29
2.1	Large Hadron Collider	29
2.2	ATLAS Detector	31
2.3	The ATLAS Coordinate System	32
2.4	Magnet	33
2.5	Inner Detector	34
2.5.1	PIXEL	37
2.5.2	SCT	37
2.5.3	TRT	38
2.6	Calorimeter	40
2.6.1	EM Calorimeter	43
2.6.2	HAD Calorimeter	43
2.6.3	Forward Calorimeter	44
2.7	Muon Spectrometer	45
2.7.1	Precision Chambers	45
2.7.2	Trigger Chambers	46
2.8	Trigger and DAQ System	47
2.9	Data Recording and Quality	49
2.10	Luminosity Acquired in Run 2	50
3	Object Reconstruction and Event Selection in ATLAS	53
3.1	PreSelection	53
3.1.1	Primary Vertex	53
3.1.2	Event Cleaning	53
3.2	Trigger	53
3.3	Offline Leptons	55
3.4	Electron Reconstruction	56
3.5	Muon Reconstruction	57

3.6	Lepton Isolation	59
3.7	Jet	60
3.7.1	Jet Reconstruction	60
3.7.2	Jet Calibration	61
3.7.3	Systematic Uncertainties Related to Jet Energy Scale Calibrations	66
3.7.4	Jet Selection	67
3.8	Missing Transverse Energy	69
3.9	Overlap Removal	70
3.10	Flavor Tagging	71
3.10.1	Impact Parameter Based Algorithms	71
3.10.2	Secondary Vertex Finder	73
3.10.3	Multi-vertex Fitter	74
3.10.4	Discriminant Score	75
3.10.5	Flavor Tagging Calibration	76
3.11	$t\bar{t}H(H \rightarrow b\bar{b})$ Objects	81
3.11.1	Event Selections	81
3.11.2	Truth-Matching	81
4	Run 1 Analysis	85
5	Analysis Overview	89
6	Monte-Carlo Simulation	91
6.1	Event Generator	91
6.2	Parton Distribution Function	92
6.3	Parton Shower and Hadronization	93
6.4	MC Samples for $t\bar{t}H(H \rightarrow b\bar{b})$ Analysis	93
6.4.1	$t\bar{t}H$	94
6.4.2	WtH and tH	94
6.4.3	$t\bar{t}V$	95
6.4.4	$t\bar{t}+\text{jets}$ and $t\bar{t}+\geq 1b$ Reweighting	95
6.4.5	Wt and single-top	98
6.4.6	Other background	98
6.4.7	Fake-Leptons	99
7	Data Driven Background Estimate	101
7.1	Fake-Leptons	101
7.2	$V+\text{jets}$	102
8	Region Definition	107
8.1	Impact of Pseudo-Continuous b -tagging Systematics	109
8.2	Simple Usage of Pseudo-Continuous b -tagging	110
8.3	Features with Pseudo-Continuous b -tagging	112
8.4	S/B Based Categorization	113
8.5	Background Based Categorization	116
8.6	Adopted Categorization	116
8.7	Boosted Region	121
8.8	Minor Modification Considerations	122
8.8.1	Loose UPR	122
8.8.2	More Controls for $t\bar{t}+\geq 1c$ Process	123

8.9	Dilepton Regions	124
8.10	Comparison of Data and Simulation in Defined Regions	128
8.11	H_T^{had} and H_T^{all} impacts on the fit	136
8.12	Event Yields for Defined Regions	140
9	MultiVariate Discriminants for $t\bar{t}H$ Signal Enhancement	143
9.1	RecoBDT	143
9.1.1	Single-lepton Channel	143
9.1.2	Dilepton Channel	166
9.2	Matrix Element Discriminant	167
9.3	Likelihood Discriminant	169
10	Final Discriminant with Multivariate Analysis	175
10.1	Single-lepton Channel	175
10.2	Single-lepton Boosted Channel	187
10.3	Dilepton Channel	187
11	Systematic Uncertainties	197
11.1	Luminosity and Pile-up	197
11.2	Lepton and E_T^{miss}	197
11.3	Jet	198
11.4	Flavor tagging	198
11.5	Cross Sections	200
11.6	$t\bar{t}H$ Modeling	202
11.7	$t\bar{t}$ Modeling	202
11.8	Other Background Modeling	206
11.9	Pruning and Categorization of Systematics	206
12	Simultaneous Global Fit and Results	209
12.1	Fit Overview	209
12.2	Fit Result	210
12.3	Fit Validations	227
12.3.1	Noticeable Distributions	227
12.3.2	Higgs Mass Distributions	234
12.3.3	Fitting with Alternative Discriminants	238
12.4	Event Yields after Fitting	239
13	Combined Result with Other $t\bar{t}H$ Channels	243
13.1	Combined Results	243
13.2	Coupling Measurement	245
14	Conclusions	249
A	MVA Algorithms for Machine Learning	251
A.1	Boosted-Decision-Tree	251
A.2	Neural-Network	252
A.3	MVA Performance	254
A.4	Comparison between BDT and NN	254

B	Tips on Plots	257
B.1	Sample Composition	257
B.2	Ratio Plot	258
B.3	χ^2 Probability	258
B.4	Automatic Binning	258
C	Tips on Systematic Uncertainty	261
C.1	Systematic Variation	261
C.2	Smoothing	261
D	Detailed Fit Information	263
D.1	Full Systematic Variations	263

List of Figures

1.1	Standard model particles	2
1.2	Feynman diagrams with virtual loops	5
1.3	Measurements of the running coupling strengths	6
1.4	Higgs potential	7
1.5	Summary of several Standard Model total and fiducial production cross section measurements	10
1.6	W and Z boson decay channels	10
1.7	Four main Feynman diagrams for Higgs boson production at the LHC	11
1.8	Summary of Higgs boson mass measurements from the individual and combined analyses, compared to the combined Run 1 measurement by the CMS and ATLAS	13
1.9	The fiducial and total cross sections of Higgs boson production measured in the 4ℓ final state	14
1.10	Differential fiducial cross sections obtained in $H \rightarrow ZZ^*$ channel	14
1.11	Cross section measurements for ggF , VBF , VH , and $t\bar{t}H$ processes	15
1.12	Four main Feynman diagrams for top quark production at the LHC	16
1.13	$t\bar{t}$ decay channels	16
1.14	Inclusive and differential cross section measurements of the top-quark pair production	17
1.15	Various top mass measurements from the CMS and ATLAS experiments	17
1.16	68% and 95% CL contours of the m_W and m_{top} indirect determination compared with the direct measurements	18
1.17	Effects on the Higgs self-coupling and potential by changing the top-Yukawa coupling value	19
1.18	Energy scale where the Higgs self-coupling becoming negative depending on the top quark Yukawa	20
1.19	Phase diagram of vacuum stability in the (m_H, m_{top}) plane	21
1.20	Phase space for VLQ contributions with respect to modifications in the signal strengths and in Yukawa couplings	23
1.21	95% CL exclusion of the vector-like T quark mass	24
1.22	Analysis channels in the search for the Higgs boson produced via $t\bar{t}H$	25
1.23	Feynman diagrams of $t\bar{t}H(H \rightarrow b\bar{b})$	26
2.1	Schematic view of the CERN accelerator complex	29
2.2	Illustration of the ATLAS detector	31
2.3	Illustration of the detection of stable particles are detected in the ATLAS detector	32
2.4	ATLAS coordination	33
2.5	Illustration of the magnet and predicted magnetic field integral as a function of $ \eta $	34
2.6	Illustration of the inner detector in the barrel and end-cap regions	35
2.7	Illustration of an electron trajectory traversing PIXEL and SCT	36
2.8	Intrinsic impact parameter resolutions of tracks	37

2.9	Schematics of PIXEL and its module	38
2.10	Schematics of Planar sensor and 3D sensor for PIXEL	38
2.11	Schematics of SCT module for barrel and end-cap regions	39
2.12	Transverse view of a quarter section of the barrel TRT straw layers	39
2.13	Illustration of the calorimeter system	40
2.14	Illustration of the EM calorimeter system	43
2.15	Illustration of the HAD calorimeter system	44
2.16	Illustration of the muon spectrometer	45
2.17	Schematics of MDT station	46
2.18	Schematics of CSC	46
2.19	Cross-section of TGC modules	47
2.20	The ATLAS TDAQ system in Run 2	48
2.21	Luminosity weighted relative fraction of good quality data delivery by the various components of the ATLAS detector/trigger subsystems in 2016	51
2.22	Integrated luminosity	51
2.23	Instantaneous luminosity	51
2.24	Luminosity-weighted distribution of the mean number of interactions per crossing	52
3.1	Electron trigger efficiency	54
3.2	Muon trigger efficiency	55
3.3	Electron ID performance	56
3.4	Electron reconstruction and ID efficiency measurements	58
3.5	Muon reconstruction and ID efficiency measurements	59
3.6	Typical reconstructed jets	61
3.7	Jet origin correction	62
3.8	Pile-up energy density distribution	63
3.9	Pile-up correction to the jet p_T	64
3.10	Average energy response and η_{det} difference between the reconstructed and truth jets	64
3.11	Relative response of jets	66
3.12	Fractional jet energy scale systematic uncertainty components	67
3.13	Distributions of the corrJVF and R_{p_T}	69
3.14	Efficiencies of the JVT	69
3.15	Comparisons of JVT between the data and the simulation	70
3.16	Illustration of b -tagging	71
3.17	Significance of the transverse signed track impact parameter, and LLR for the IP2D algorithm	73
3.18	Discriminants into the MVA for b -tagging used in the secondary vertex finder algorithm	75
3.19	JetFitter vertex reconstruction efficiency as a function of jet p_T	75
3.20	Discriminants in the MVA for b -tagging used in JetFitter	76
3.21	c -jet and light-jet rejection performance versus b -jet efficiency at various MV2 b -tagging settings	76
3.22	Distributions of the b -tagging score	77
3.23	b -tagging efficiency and ratio	78
3.24	Ratio of c -jet mis-tagging efficiency	79
3.25	MV2c10 and MV2c10Flip distributions	80
3.26	Ratio of the light-jet mis-tagging efficiency	80
3.27	p_T v.s. ΔR for two quarks from W_{had} decay and two b -quarks from Higgs boson decay	82

4.1	Background compositions in regions in the Run 1	86
4.2	Summary of signal strength measurements of the $t\bar{t}H$ production in Run 1	86
4.3	Background compositions in regions in Run 1	87
6.1	Sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator and parton shower	92
6.2	NNLO PDFs evaluated at $\mu^2 = 10 \text{ GeV}^2$ and $\mu^2 = 10^4 \text{ GeV}^2$	93
6.3	Feynman diagrams for $t\bar{t}H$ production	95
6.4	Feynman diagrams for WtH and tH productions	95
6.5	Feynman diagrams for $t\bar{t}Z$ and $t\bar{t}W$ productions	95
6.6	Feynman diagrams for $t\bar{t}$ +jets	96
6.7	Fractions of the $t\bar{t} + b$, $t\bar{t} + b\bar{b}$, $t\bar{t} + B$ and $t\bar{t} + \geq 3b$ sub-categories on the truth-level	97
6.8	Feynman diagrams for single-top production	98
6.9	Feynman diagrams for minor background sources	99
7.1	Efficiencies for fake-lepton estimation in the electron channel and 2015 runs	103
7.2	Efficiencies for fake-lepton estimation in the electron channel and 2016 runs	104
7.3	Efficiencies for fake-lepton estimation in the muon channel and 2015 runs	104
7.4	Efficiencies for fake-lepton estimation in the muon channel and 2016 runs	105
8.1	Background fractions in the nine analysis regions defined simply with single b -tagging WP at 70%	108
8.2	Distributions of MV2c10 b -tagging scores and b -tagging tag weight bins for b -, c - or light-jets	109
8.3	Highly constrained systematic uncertainties in various b -tagging systematic options	110
8.4	Background fractions in the twelve analysis regions defined with two b -tagging WPs at 60% and 77%	111
8.5	Number of $t\bar{t}H(H \rightarrow b\bar{b})$ events expected by the SM in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n\text{Jets} \geq 6$	112
8.6	S/B ratio in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n\text{Jets} \geq 6$	113
8.7	Background composition in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n\text{Jets} \geq 6$	113
8.8	Ratio of $t\bar{t}H(H \rightarrow b\bar{b})$ to the $t\bar{t} + \geq 2b$ background process in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n\text{Jets} \geq 6$	113
8.9	Region Definition with two b -tagging WPs at 60% and 77% for $n\text{Jets} \geq 6$	114
8.10	Region categorization based on S/B	114
8.11	Background fractions in the nine analysis regions for ≥ 6 jets defined by S/B ordering	115
8.12	Region categorization based on BKG composition	116
8.13	Background fractions in the nine analysis regions for ≥ 6 jets defined by background components	117
8.14	Background fractions in the eleven analysis regions for the single-lepton channel	118
8.15	Region categorization for ≥ 6 jets events in the single-lepton channel	119
8.16	Region categorization for 5 jets events in the single-lepton channel	120
8.17	Truth Higgs p_T and truth hadTop p_T	121
8.18	Background fractions in the subdivided $\text{CR}(t\bar{t} + \geq 1c)$ regions for the single-lepton channel	123
8.19	Number of $t\bar{t}H(H \rightarrow b\bar{b})$ in ≥ 4 jets categorized in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$	124
8.20	S/B ratio in ≥ 4 jets categorized in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$	124
8.21	Background fractions in the seven analysis regions for the dilepton channel	125
8.22	Region categorization for ≥ 4 jets events in the dilepton channel	126

8.23	Region categorization for 3 jets events in the dilepton channel	127
8.24	Number of jets in the single-lepton and dilepton channels	128
8.25	Tag weight bin numbering to convert into the one dimensional distribution. Column bins are merged to avoid poor statistics.	129
8.26	Number of b -tagged jets at each WP in the single-lepton channel	130
8.27	Number of b -tagged jets at each WP in the dilepton channel	131
8.28	H_T^{had} distributions in ≥ 6 j regions for the single-lepton channel	132
8.29	H_T^{had} distributions in 5 j regions and boosted region for the single-lepton channel	133
8.30	H_T^{all} distributions in ≥ 4 j regions for the dilepton channel	134
8.31	H_T^{all} distributions in 3 j regions for the dilepton channel	135
8.32	Highly constrained or pulled systematic variations with the H_T^{had} distribution in $t\bar{t}$ +light control regions	137
8.33	H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with nominal setup	138
8.34	H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with onebin setup	138
8.35	H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with optimum setup	138
8.36	Highly constrained or pulled systematic variations	139
9.1	RecoBDT scores in $\geq 6\text{jSR}(t\bar{t}H)$	147
9.2	RecoBDT scores in $\geq 6\text{jSR}(t\bar{t}+\geq 2b, \text{Hi})$	147
9.3	RecoBDT scores in $\geq 6\text{jSR}(t\bar{t}+\geq 2b, \text{Lo})$	148
9.4	RecoBDT scores in $5\text{jSR}(t\bar{t}H)$	148
9.5	RecoBDT scores in $5\text{jSR}(t\bar{t}+\geq 2b, \text{Hi})$	149
9.6	Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT w/o H in ≥ 6 jets	150
9.7	Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT with H in ≥ 6 jets	151
9.8	Masses of two-object systems constructed by the RecoBDT w/o H in ≥ 6 jets	152
9.9	Masses of two-object systems constructed by the RecoBDT with H in ≥ 6 jets	153
9.10	ΔR variables constructed by the RecoBDT w/o H in ≥ 6 jets (1)	154
9.11	ΔR variables constructed by the RecoBDT w/o H in ≥ 6 jets (2)	155
9.12	ΔR variables constructed by the RecoBDT with H in ≥ 6 jets (1)	156
9.13	ΔR variables constructed by the RecoBDT with H in ≥ 6 jets (2)	157
9.14	Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT w/o H in 5 jets	158
9.15	Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT with H in 5 jets	159
9.16	Masses of two-object systems constructed by the RecoBDT w/o H in 5 jets	160
9.17	Masses of two-object systems constructed by the RecoBDT with H in 5 jets	161
9.18	ΔR variables constructed by the RecoBDT w/o H in 5 jets (1)	162
9.19	ΔR variables constructed by the RecoBDT w/o H in 5 jets (2)	163
9.20	ΔR variables constructed by the RecoBDT with H in 5 jets (1)	164
9.21	ΔR variables constructed by the RecoBDT with H in 5 jets (2)	165
9.22	MEM _{D1} distribution in $\geq 6\text{jSR}(t\bar{t}H)$	168
9.23	Kinematic variables used in $\mathcal{P}_S^{\text{kin}}$ and $\mathcal{P}_B^{\text{kin}}$ showing noticeable differences among $t\bar{t}H$, $t\bar{t} + b\bar{b}$, and $t\bar{t} + bj$ in ≥ 6 jets and ≥ 4 b -tagged jets.	171
9.24	Kinematic variables used in $\mathcal{P}_S^{\text{kin}}$ and $\mathcal{P}_B^{\text{kin}}$ showing noticeable differences in ≥ 6 jets and ≥ 4 b -tagged jets.	172
9.25	LHD distributions in $\geq 6\text{jSR}(t\bar{t}H)$	172

9.26	LHD distributions in ≥ 6 jets SR($t\bar{t}+\geq 2b$,Hi) and SR($t\bar{t}+\geq 2b$,Lo).	173
9.27	LHD distributions in 5 jets SR($t\bar{t}H$) and SR($t\bar{t}+\geq 2b$).	173
10.1	Linear correlations among input variables in ClassBDT	179
10.2	Input variables (event kinematics, LHD, and MEM) in ClassBDT for $\geq 6j$ SR($t\bar{t}H$)	180
10.3	Input variables (RecoBDT with H) in ClassBDT for $\geq 6j$ SR($t\bar{t}H$)	181
10.4	Input variables (RecoBDT w/o H) in ClassBDT for $\geq 6j$ SR($t\bar{t}H$)	181
10.5	Input variables (event kinematics and LHD) in ClassBDT for ≥ 6 jets signal regions	182
10.6	Input variables (RecoBDT with H) in ClassBDT for ≥ 6 jets signal regions	183
10.7	Input variables (RecoBDT w/o H) in ClassBDT for ≥ 6 jets signal regions	183
10.8	Input variables (b -tagging variables) in ClassBDT for ≥ 6 jets signal regions	184
10.9	Input variables (event kinematics and LHD) in ClassBDT for 5 jets signal regions	185
10.10	Input variables (RecoBDT with H) in ClassBDT for 5 jets signal regions	186
10.11	Input variables (RecoBDT w/o H) in ClassBDT for 5 jets signal regions	186
10.12	Input variables (b -tagging variables) in ClassBDT for 5 jets signal regions	186
10.13	input variables in ClassBDT for boosted regions	188
10.14	Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}H$) region	190
10.15	Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}H$) region	191
10.16	Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}+\geq 2b$) region	192
10.17	Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}+\geq 2b$) region	193
10.18	Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}+1b$) region	194
10.19	Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4j$ SR($t\bar{t}+1b$) region	195
11.1	Systematic uncertainties of eigen-variations in the jet energy scale	198
11.2	Systematic uncertainties of calibration parameters in the jet energy scale and energy resolution	199
11.3	Systematic uncertainties on b -jet tagging efficiency	199
11.4	Systematic uncertainties on c -jet mis-tagging efficiency	200
11.5	Systematic uncertainties on light-jet mis-tagging efficiency	200
11.6	PS modeling variations for $t\bar{t}H$ events (HERWIG++ vs. PYTHIA8) in $\geq 6j$ SR($t\bar{t}H$).	202
11.7	H_T^{had} variations by $t\bar{t}$ modeling (POWHEG+PYTHIA 8 vs. SHERPA5F) in $\geq 6j$ CR($t\bar{t}+\geq 1c$).	204
11.8	H_T^{had} variations by $t\bar{t}$ PS modeling (PYTHIA 8 vs. HERWIG7) in $\geq 6j$ CR($t\bar{t}+\geq 1c$).	204
11.9	H_T^{had} variations by $t\bar{t}$ ISR / FSR modeling in $\geq 6j$ CR($t\bar{t}+\geq 1c$).	204
11.10	H_T^{had} variations by $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ flavor scheme modeling in $\geq 6j$ CR($t\bar{t}+\geq 1c$).	204
11.11	ClassBDT variations by $t\bar{t}$ modeling in $\geq 6j$ SR($t\bar{t}H$).	205
11.12	Variations by $t\bar{t}Z$ modeling (SHERPA vs. MG5_aMC@NLO+PYTHIA8)	206
12.1	Fraction of background components and sensitivities in analysis regions	211
12.2	Fit pulls and constraints	212
12.3	Linear correlation matrix between two nuisance parameters	213
12.4	Summary of the observed and expected best-fit values of $\mu_{t\bar{t}H}$ in individual channels and their combination	215
12.5	Ranking of the nuisance parameters included in the fit according to the impact on the measured signal strength	216
12.6	Analysis regions in the single-lepton channel	217
12.7	Analysis regions in the dilepton channel	218
12.8	H_T^{had} distributions in $5j$ CR($t\bar{t}+\geq 1c$)	219
12.9	ClassBDT distributions in $5j$ SR($t\bar{t}+\geq 2b$)	219
12.10	H_T^{had} distributions in $5j$ SR($t\bar{t}H$)	220

12.11 H_T^{had} distributions in the boosted region	220
12.12 H_T^{had} distributions in $\geq 6j_{\text{CR}}(t\bar{t}+\geq 1c)$	221
12.13 ClassBDT distributions in $\geq 6j_{\text{SR}}(t\bar{t}+\geq 2b, \text{Lo})$	221
12.14 ClassBDT distributions in $\geq 6j_{\text{SR}}(t\bar{t}+\geq 2b, \text{Hi})$	222
12.15 ClassBDT distributions in $\geq 6j_{\text{SR}}(t\bar{t}H)$	222
12.16 ClassBDT distributions in $\geq 4j_{\text{SR}}(t\bar{t}+1b)$	223
12.17 ClassBDT distributions in $\geq 4j_{\text{SR}}(t\bar{t}+\geq 2b)$	223
12.18 ClassBDT distributions in $\geq 4j_{\text{SR}}(t\bar{t}H)$	224
12.19 All analysis region distributions for both lepton channels sorted by $\log(S/B)$ after fitting	225
12.20 Number of jets in the single-lepton and dilepton channels	227
12.21 Number of b -tagged jets at each WP in the single-lepton channel	228
12.22 Number of b -tagged jets at each WP in the dilepton channel	229
12.23 H_T^{had} distributions in $\geq 6j$ regions for the single-lepton channel after fitting	230
12.24 H_T^{had} distributions in $5j$ regions and boosted region for the single-lepton channel after fitting	231
12.25 H_T^{all} distributions in $\geq 4j$ regions for the dilepton channel after fitting	232
12.26 H_T^{all} distributions in $3j$ regions for the dilepton channel after fitting	233
12.27 m_{bb}^{Higgs} (Reco BDT w/o H) distributions in $\geq 6j_{\text{SR}}(t\bar{t}H)$	234
12.30 m_{bb}^{Higgs} (Reco BDT with H) distributions in $5j_{\text{SR}}(t\bar{t}H)$	235
12.31 m_{bb}^{Higgs} (Reco BDT w/o H) distributions in $\geq 4j_{\text{SR}}(t\bar{t}H)$	236
12.32 m_{bb}^{Higgs} (Reco BDT with H) distributions in $\geq 4j_{\text{SR}}(t\bar{t}H)$	236
12.33 $m_{\text{rcjet}}^{\text{Higgs}}$ distributions in the boosted SR	237
12.34 Summary of the observed best-fit values of $\mu_{t\bar{t}H}$ comparing the different sets of distributions in the signal regions	238
13.1 Summary of $t\bar{t}H$ measurements of μ from various Higgs decay channels and the combined result	244
13.2 Profiled likelihood ratio as a function of κ_{top} using $t\bar{t}H$ ($H \rightarrow \gamma\gamma$) channel	245
13.3 Profiled likelihood ratio as a function of κ_{top}	246
13.4 Allowed regions at 68% and 95% CL in the κ_F - κ_V plane obtained from the combination of all $t\bar{t}H$ channels analyses	247
A.1 Schematics of decision tree algorithm	252
A.2 Schematics of BDT algorithm	253
A.3 Schematics of NN algorithm	253
A.4 Output scores of BDT and NN algorithms	254
A.5 ROC curves of BDT and NN algorithms	255
B.1 BDT scores before and after automatic binning	257
C.1 Systematic shapes before and after smoothing	262
D.1 Fit pulls and constraints for all systematic variations	264
D.2 Fit pulls and constraints for all systematic variations	265

List of Tables

1.1	Lagrangian and equation of motion for a free particle with a spin number	1
1.2	Production cross sections for the Higgs boson	11
1.3	Decay channels of the Higgs boson	12
2.1	Machine parameters of the LHC	30
2.2	Typical resolutions and detector parameters of the inner detector sub-systems	36
2.3	$\delta\eta$ - $\delta\phi$ segments and η coverages of the calorimeter sub-systems	41
3.1	Triggers used for analysis	54
3.2	Definitions of electron discriminating variables	57
3.3	Muon ID efficiency	59
3.4	Input variables for the b -tagging MVA	72
3.5	b -tagging working points for MV2c10	77
3.6	Truth parton acceptance efficiency	82
3.7	Truth-level boosted fraction	83
3.8	Truth matching efficiency	83
4.1	S/B and $S/\sqrt{B}$ obtained in Run 1 regions	85
6.1	Nominal samples used in the analysis	94
6.2	Summary of the settings used for the simulation of the $t\bar{t}$ samples	98
8.1	S/B and $S/\sqrt{B}$ in the nine analysis regions defined simply with single b -tagging WP at 70%	108
8.2	Impacts on the sensitivities with various b -tagging systematic options	110
8.3	S/B and $S/\sqrt{B}$ in the twelve analysis regions defined simply with single b -tagging WP at 70%	111
8.4	Impacts on the sensitivities with removing regions	112
8.5	S/B and $S/\sqrt{B}$ in the nine analysis regions defined by S/B ordering	115
8.6	Uncertainty on $\mu_{t\bar{t}H}$, binned expected significance Z_{exp} , and uncertainties on $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalizations with merging regions defined by S/B ordering.	115
8.7	S/B and $S/\sqrt{B}$ in the four analysis regions defined by background components	117
8.8	S/B and $S/\sqrt{B}$ in the eleven analysis regions for the single-lepton channel	118
8.9	Statistics in signal regions for the single-lepton channel	122
8.10	S/B and $S/\sqrt{B}$ in the seven analysis regions for the dilepton channel	125
8.11	Yields of the analysis regions in ≥ 6 jets for the single-lepton channel for 36.1 fb^{-1}	140
8.12	Yields of the analysis regions in 5 jets for the single-lepton channel for 36.1 fb^{-1}	141
8.13	Yields of the analysis regions for the dilepton channel for 36.1 fb^{-1}	142
9.1	Input variables to RecoBDT in the single-lepton channel	145
9.2	Reconstruction efficiency	146
9.3	Input variables to RecoBDT in the dilepton channel	166

9.4	Efficiency of fully truth-matched jets in $t\bar{t}H$ events in different selections	170
9.5	Efficiency of fully truth-matched jets in $t\bar{t} + b\bar{b}$ events in different selections	170
10.1	Input variables used to the ClassBDTs in the single-lepton signal regions	177
10.2	Ranking of the input variables used to the ClassBDTs in the single-lepton signal regions	178
10.3	Input variables to the ClassBDTs in the boosted single-lepton signal region.	187
10.4	Variables used to the ClassBDTs in the dilepton signal regions.	189
11.1	List of cross section normalization uncertainties	201
11.2	Summary of the systematic uncertainty sources in $t\bar{t}$ +jets modeling	203
11.3	List of systematic uncertainties considered in the analysis	207
12.1	Fitted normalization factors for $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$	214
12.2	Summary of the contributions to the signal strength uncertainty	214
12.3	Fitted normalization factors for $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ with the alternative sets of distributions in the signal regions	238
12.4	Yields of the analysis regions in ≥ 6 jets for the single-lepton channel for 36.1 fb^{-1} after fitting	239
12.5	Yields of the analysis regions in 5 jets for the single-lepton channel for 36.1 fb^{-1} after fitting	240
12.6	Yields of the analysis regions for the dilepton channel for 36.1 fb^{-1} after fitting	241
13.1	Summary of the uncertainties affecting the combined $\mu_{t\bar{t}H}$ value	244

Chapter 1

Standard Model and Higgs Boson

1.1 The Standard Model

The Standard Model (SM) is a renormalizable quantum field theory describing the interactions between elementary particles through fundamental forces also carried by the elementary particles. There are seventeen particles that are categorized into three groups based on their spin numbers and their different roles in forming the universe:

- twelve fermions described by spin-1/2 fields, which form the matter complying with the Pauli exclusion principle
- four vector bosons (gauge bosons) described by spin-1 fields, which mediate interactions between particles
- one scalar-boson (Higgs boson) described by a spin-0 field, which gives masses to gauge bosons and fermions.

All particles have their associated anti-particles with same masses but opposite quantum numbers. In general, the field (φ) associated with spin number can be described by a Lagrangian satisfying the Euler-Lagrange equation:

$$\frac{\delta}{\delta\varphi}S = -\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) + \frac{\partial \mathcal{L}}{\partial \varphi} = 0. \quad (1.1)$$

The equation of motion is derived from the Lagrangian as described by this equation, and is summarized in Table 1.1.

TABLE 1.1: Lagrangian and equation of motion for a free particle with a spin number. γ_μ is the gamma matrix. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

spin	field	Lagrangian	motion equation
0	ϕ (scalar): Klein-Gordon field	$\frac{1}{2} ((\partial_\mu \phi) (\partial^\mu \phi) - m^2 \phi^2)$	$(\square + m^2) \phi = 0$
1/2	ψ (spinor): Dirac field	$\bar{\psi} (i\gamma_\mu \partial^\mu - m) \psi$	$(i\gamma_\mu \partial^\mu - m) \psi = 0$
1	A (vector): Proca field (Maxwell field if $m=0$)	$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu - j_\mu A^\mu$	$\square A^\mu - \partial^\mu (\partial_\nu A^\nu) + m^2 A^\mu = j^\mu$

Three fundamental forces in the SM are described as gauge theory where interactions mediated by the gauge bosons: photon (γ) for the electromagnetic interaction, Z^0 and $W^\pm$ bosons for the

weak interaction, and gluon (g) for the strong interaction. Fermions can be subdivided in two categories (quarks and leptons) based on which interactions they experienced. All of them undergo weak interactions, while only quarks experience strong interactions. Quarks have non-integer charges $2/3$ (up-type quarks) or $-1/3$ (down-type quarks), while leptons have integer charges of 0 (neutrinos) or -1 (charged leptons). Pairs (weak-doublets) are constructed using a neutrino and charged-lepton, and using a up-type and down-type quarks. Quarks and leptons are categorized in three generations characterized by their masses. A summary of the particles in the SM is shown in Figure 1.1.

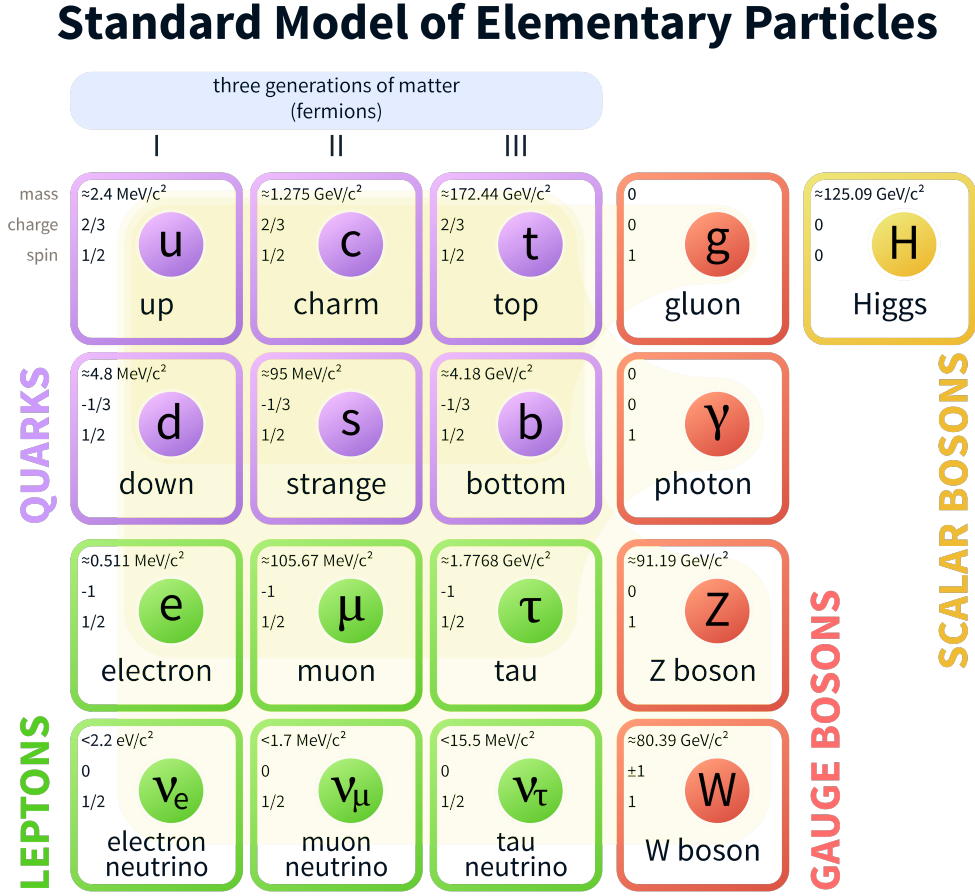


FIGURE 1.1: Standard model particles.

In gauge theory, the Lagrangian should be invariant under the local gauge transformations (U) for the fermion spinor(ψ):

$$\psi \rightarrow \psi' = U\psi. \quad (1.2)$$

The electroweak and strong interactions are based on different symmetry groups that require invariance under local gauge transformations:

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (1.3)$$

Although gravity is also expected to be a fundamental interaction in the universe, its quantum theory has not yet been established for conserving the Lagrangian under a local gauge transformation, and it is not included in the SM.

$SU(3)_C$ is a gauge group for strong interactions mediated by gluons, and their dynamics are described by quantum chromodynamics (QCD). In QCD, quarks are described as the color triplet eigenstates (denoted by color names, R , G and B), while gluons are color octets carrying a pair of color-anticolor. They interact through exchanging their colors at a quark-gluon or gluon-gluon vertex. The colored particles are unstable, and free quarks and gluons are connected to each other as a color singlet stable particle (non-colored state, e.g. $1/\sqrt{3}(R\bar{R} + B\bar{B} + G\bar{G})$, $1/\sqrt{6}(RGB - RBG + GBR - GRB + BRG - BGR)$). This composite particle including two quarks (mesons) and three quarks (baryons) is called hadron. Therefore, the internal quark properties are experimentally measured in the corresponding hadron observables.

$SU(2)_L \times U(1)_Y$ is the unified gauge group for the electromagnetic and weak (EW) interactions mediated by photons, $W^\pm$, and Z^0 bosons. The fermion field ψ can be expressed by the left-handed and right-handed components as

$$\psi_R = \frac{1}{2}(1 \mp \gamma^5)\psi. \quad (1.4)$$

The $SU(2)_L$ weak interaction only acts on the $V-A$ form corresponding to left-handed fermions. Therefore, the EW eigenstate of the fermion ψ is composed of a left-handed weak-isospin doublet ($\psi_L = Q_L$ for quarks, L_L for leptons) denoted as $T = 1/2$, $T_3 = \pm 1/2$ and two right-handed weak-isospin singlets ($\psi_R \in \{\mathcal{U}_R, \mathcal{D}_R\}$ for quarks, $\{\nu_R, \ell_R\}$ for leptons) denoted as $T = 0$, $T_3 = 0$. Here, T refers to the weak-isospin number and T_3 refers the third component of the weak-isospin. One doublet and two singlets are generally defined as

$$Q_L = \begin{pmatrix} \mathcal{U} \\ \mathcal{D} \end{pmatrix}_L = \left\{ \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\}, \quad \mathcal{U}_R = \{u_R, c_R, t_R\}, \quad \mathcal{D}_R = \{d_R, s_R, b_R\}, \quad (1.5)$$

for quark fields, and

$$L_L = \begin{pmatrix} \nu \\ \ell \end{pmatrix}_L = \left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \right\}, \quad \ell_R = \{e_R, \mu_R, \tau_R\}, \quad \nu_R = \{\nu_{e,R}, \nu_{\mu,R}, \nu_{\tau,R}\} \quad (1.6)$$

for lepton fields.

The hypercharge (Y) acts on the charged particle in the group $U(1)_Y$,

$$Q = T_3 + \frac{Y}{2}. \quad (1.7)$$

The SM Lagrangian describes all interactions among gauge fields, fermion fields, and Higgs fields as:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}}. \quad (1.8)$$

The $\mathcal{L}_{\text{Gauge}}$ is the kinematic and self-interacting terms of gauge fields. It is constructed by the $U(1)_Y$ gauge field B_μ , the three $SU(2)_L$ gauge fields W_μ^i and the eight $SU(3)_C$ gauge fields G_μ^a ,

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4}G_{\mu\nu}^a G_{\mu\nu}^a - \frac{1}{4}W_{\mu\nu}^i W_{\mu\nu}^i - \frac{1}{4}B_{\mu\nu} B_{\mu\nu} \quad (1.9)$$

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_{b,\mu} G_{c,\nu} \quad a, b, c \in (0, 1, \dots, 7) \quad (1.10)$$

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g_w \varepsilon^{ijk} W_{j,\mu} W_{k,\nu} \quad i, j, k \in (1, 2, 3) \quad (1.11)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.12)$$

$$(1.13)$$

where $G_{\mu\nu}^a$, $W_{\mu\nu}^i$ and $B_{\mu\nu}$ are gauge field strength tensors, and g_s and g_w are coupling strengths for the strong and weak interactions ($SU(3)_C$ and $SU(2)_L$), respectively. The electroweak bosons are observed not in eigenstates of their interactions (W_μ^i and B_μ), but rather in their linear mixed fields $W_\mu^\pm$, Z_μ^0 and A_μ having mass eigenstates, written as

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \quad (1.14)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2), \quad (1.15)$$

where θ_W is the weak-mixing angle (Weinberg angle) satisfying

$$\cos \theta_W = \frac{g_w}{\sqrt{g_w^2 + g_w'^2}}, \quad \sin \theta_W = \frac{g_w'}{\sqrt{g_w^2 + g_w'^2}}. \quad (1.16)$$

The coupling $g_w(g_w')$ is referred to the coupling strength through the weak isospin (weak hypercharge) quantum number.

The $\mathcal{L}_{\text{fermion}}$ describes fermion kinematic terms and interactions with gauge bosons.

$$\mathcal{L}_{\text{fermion}} = \bar{\psi}_L i \gamma^\mu D_\mu \psi_L + \bar{\psi}_R i \gamma^\mu D_\mu \psi_R, \quad (1.17)$$

where D_μ is a covariant derivative promoting the gauge symmetry from global to local and replacing simple ∂_μ :

$$D_\mu = \partial_\mu - i g_s \hat{T}_{s,a} G_\mu^a - i g_w \hat{T}_{w,i} W_\mu^i - i g_w' \frac{\hat{Y}}{2} B_\mu. \quad (1.18)$$

$\hat{T}_{s,a}$, $\hat{T}_{w,i}$ and $\hat{Y}/2$ are interaction generators and are described as:

$$\hat{T}_{s,a} = \begin{cases} \lambda_a/2 & \text{for color triplet} \\ 0 & \text{for color singlet} \end{cases} \quad (1.19)$$

$$\hat{T}_{w,i} = \begin{cases} \tau_i/2 & \text{for weak isospin doublet} \\ 0 & \text{for weak isospin singlet} \end{cases} \quad (1.20)$$

$$\hat{Y}/2 = Y/2 \quad (1.21)$$

where $\lambda_a/2$ ($\tau_i/2$) is a set of linearly independent traceless 3×3 (2×2) matrices.

The coupling strength has a dependence on the energy scale (Q) of the interaction. A gauge boson field can create pairs of virtual particles for a very short time followed by annihilation back to the original gauge boson. This process is illustrated using a loop diagram, as shown in Figure 1.2a. The strong and the weak interactions have also self-interaction terms ($g_s f^{abc} G_{b,\mu} G_{c,\nu}$ and $g_w \epsilon^{ijk} W_{j,\mu} W_{k,\nu}$) in Equations 1.10 and 1.11. Feynman diagrams are also shown in Figure 1.2b. In quantum electrodynamics (QED), the effective coupling g runs with the scale of momentum transfer (corresponding to the energy scale Q) due to virtual fermion loops phenomenon, which is called vacuum polarization. The running amplitude is given by

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 - \frac{\alpha(\mu^2)}{3\pi} z_f \log\left(\frac{Q^2}{\mu^2}\right)} \quad (1.22)$$

which is renormalized at an energy scale μ as a cut-off for larger momentum contributions in the

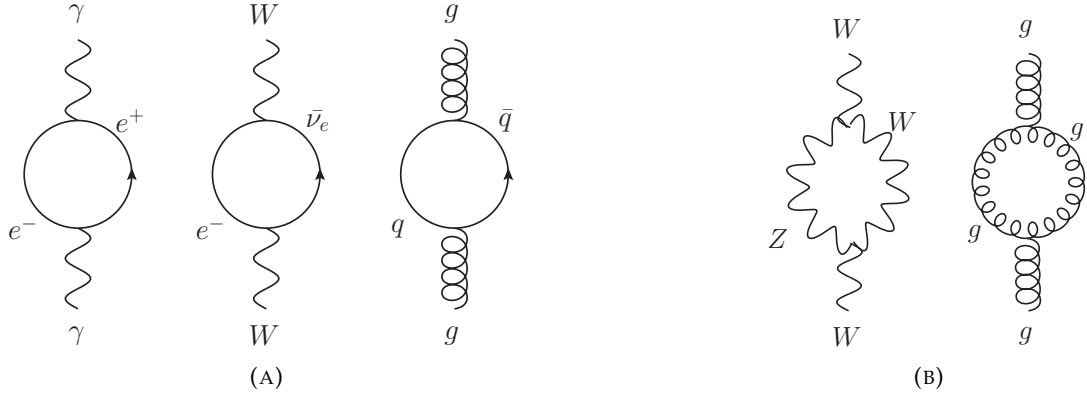


FIGURE 1.2: Feynman diagrams with virtual fermion loops (loops by themselves) contributing to the coupling strengths in the left (right) plot

loop integration for preventing the divergence of the coupling. The z_f is a sum of squared active fermion charges. From this equation, the QED coupling increases with energy scale, as shown in Figure 1.3a. This is experimentally observed by the LEP experiment.

On the other hand, the weak and the strong couplings decrease at larger energy scale due to their self-interaction terms. The running of the strong interaction in approximation with one-loop corrections is described as,

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \log\left(\frac{Q^2}{\Lambda^2}\right)}, \quad (1.23)$$

where Λ is the upper limit of the energy scale in the loop integration, and the β -function can be written using the number of fermions considered in the loops (n_f) as

$$\beta(\alpha_s) = -\frac{\beta_0}{(4\pi)^2} \alpha_s^3, \quad \beta_0 = \frac{33 - 2 \cdot n_f}{3}. \quad (1.24)$$

α_s decreases with increasing Q^2 , a feature known as asymptotic freedom, which is crucial to conduct an experiment using a hadron-hadron collider at high energies. When two hadrons collide at high energies, such as $\sim O(\text{TeV})$, the partons in the hadrons interact as two free particles. On the other hand, at low energies, the coupling strength reaches a divergence point called color confinement. Therefore, QCD calculations require two cut-off scales, such as in the MC parton generation with hadronization, to describe physics processes and decay products. The running feature of the QCD coupling strength is shown in Figure 1.3b. This figure compiles the DØ, H1, ZEUS and ATLAS experiments at different energy scales.

The dynamics of the weak interaction are similar to the dynamics of the strong interaction, and its coupling decreases with energy scale.

Both Lagrangians of gauge bosons and fermions conserve gauge symmetry under a local gauge transformation. However, if the mass terms are introduced by hand, their gauge invariances are not conserved. For instance, the mass terms of weak vector bosons are described as:

$$m_W^2 W_\mu^\dagger W^\mu + \frac{1}{2} m_Z^2 Z_\mu Z^\mu \quad (1.25)$$

As for the experimental results, the weak vector bosons have non-zero masses. To give masses to the weak vector bosons while keeping the conservation under a local gauge transformation, a

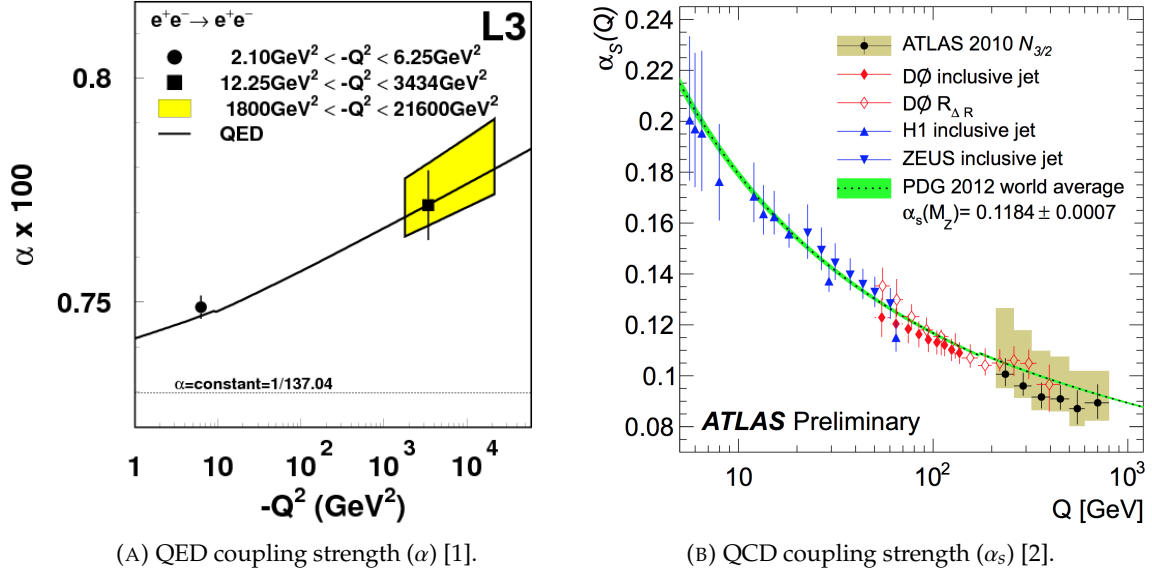


FIGURE 1.3: Measurements of the running coupling strengths. (Left) The QED coupling strength by LEP-L3 measured at three points with different energy scales as shown by the points and the yellow band. (Right) The QCD coupling strength measured by four experiments (DØ, H1, ZEUS, and ATLAS experiments). The PDG 2012 world average was obtained by combining results from DØ, H1, and ZEUS experiments. The data points from the ATLAS experiment are on the extrapolated curve within their uncertainties.

spontaneous symmetry breaking (SSB) is introduced, where $SU(2)_L \times U(1)_Y$ breaks down into a $U(1)_{EM}$ symmetry group. Such a mechanism (e.g. the Brout-Englert-Higgs mechanism) can give masses to fermions and vector bosons through their interactions with the new scalar particle, the Higgs boson.

1.2 Brout-Englert-Higgs Mechanism

1.2.1 Higgs boson and vector bosons

The Higgs field is introduced as a weak isospin doublet of complex scalar fields:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (1.26)$$

where $\phi^+(\phi^0)$ has a weak isospin number 1/2 with its third component +1/2(−1/2) and an electric charge +1(0), therefore its hyper-charge is +1(+1).

The Lagrangian of the complex scalar field is described as:

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \quad (1.27)$$

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (1.28)$$

where D_μ is the same as Equation 1.18, but this complex scalar field is color-singlet. The renormalizable Higgs potential $V(\phi)$ is described by two parameters, μ^2 and λ . If $\lambda < 0$, the potential

minimum becomes $-\infty$ as $\phi \rightarrow \pm\infty$. Hence, this case leads to an unphysical and non-stable vacuum of the universe. In the case of $\lambda > 0$, two options can be considered for the Higgs potential as shown in Figure 1.4. If $\mu^2 > 0$, the potential minimum is at $|\phi| = 0$, and the vacuum expectation value (VEV) $\langle 0|\phi|0\rangle = 0$ with conserved symmetry around $|\phi| = 0$.

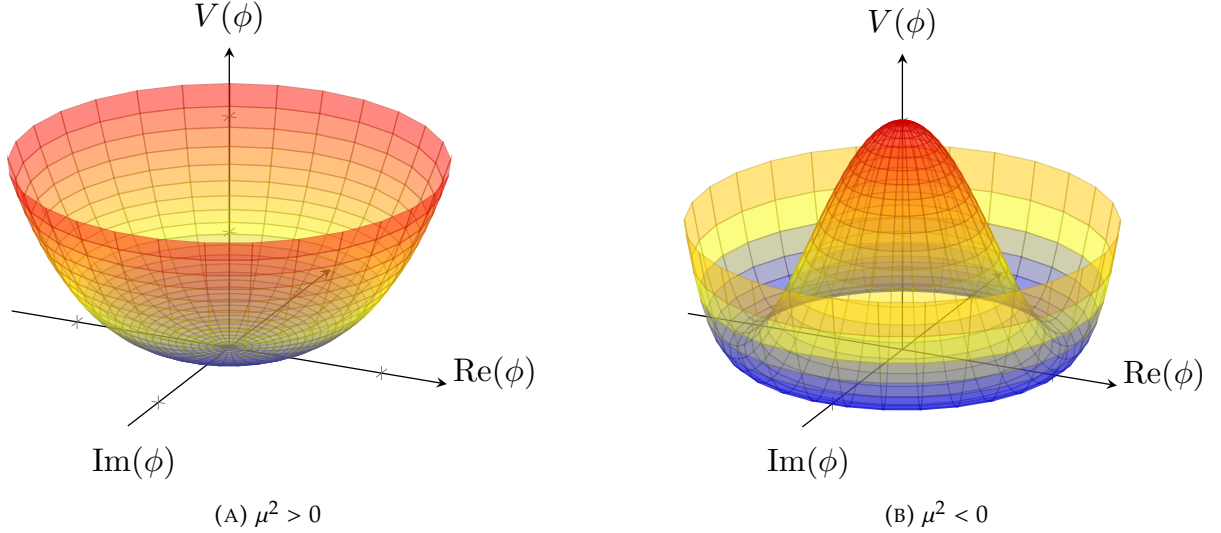


FIGURE 1.4: Higgs potential. The potential is drawn for parameters (A) $\lambda > 0$ and $\mu^2 > 0$, and (B) $\lambda > 0$ and $\mu^2 < 0$.

For $\mu^2 < 0$, the potential is at a minimum

$$\langle \phi^\dagger \phi \rangle = -\frac{\mu^2}{2\lambda} \quad (1.29)$$

where there are no unique minima point in the Higgs field. However, the VEV at the minimum potential is $\langle 0|\phi|0\rangle = v/\sqrt{2}$, where $v = \sqrt{-\mu^2/\lambda} \in \mathbb{R}$. Once a particular ground state is chosen, the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken to $U(1)_{EM}$ as:

$$\phi|_{\text{ground state}} = \phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.30)$$

For an infinitesimal phase shift by the generators of $SU(2)_L$ or $U(1)_Y$, its symmetry can be confirmed as follows:

$$\phi_0 \rightarrow e^{\frac{i}{2}\tau_j\theta_j}\phi_0 = (1 + i\tau_j\theta_j)\phi_0 \quad \text{for } SU(2)_L, \quad (1.31)$$

$$\phi_0 \rightarrow e^{\frac{i}{2}Y\theta_0}\phi_0 = (1 + iY\theta_0)\phi_0. \quad \text{for } U(1)_Y. \quad (1.32)$$

Under the independent transformations of the $SU(2)_L$ and $U(1)_Y$ generators, Equations 1.31 and 1.32, ϕ_0 will be different from the original ϕ_0 . However, if the transformation operator is $\tau_3/2 + Y/2$ (electric charge generator), the Higgs field ϕ_0 is conserved. This shows the $U(1)_{EM}$ symmetry of the Higgs field after the spontaneous breaking. At the minimum of the Higgs potential, the

Lagrangian of the Higgs term is described as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi_0)^\dagger (D^\mu \phi_0) - \frac{1}{2}\mu^2 v^2 - \frac{1}{4}\lambda v^4 \quad (1.33)$$

$$\begin{aligned} (D_\mu \phi_0)^\dagger (D^\mu \phi_0) &= \frac{1}{8}v^2 \left[g_w^2 \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) \right. \\ &\quad \left. + \frac{1}{8}v^2 \left[\left(g_w(W_\mu^3) - g'_w(B_\mu) \right)^2 + 0 \cdot \left(g'_w(W_\mu^3) + g_w(B_\mu) \right)^2 \right] \right] \end{aligned} \quad (1.34)$$

$$= 2 \cdot \frac{1}{2} \left(\frac{1}{2}v g_w \right)^2 W_\mu^\dagger W^\mu + \frac{1}{2} \left(\frac{1}{2}v \sqrt{g_w^2 + g_w'^2} \right)^2 Z_\mu^2 + 0 \cdot A_\mu^2 \quad (1.35)$$

From Equation 1.35 and the mass term of the Proca Lagrangian (see Table 1.1), the masses of gauge bosons are naturally derived with spontaneous breaking:

$$m_{W^+} = m_{W^-} = \frac{1}{2}v g_w \quad (1.36)$$

$$m_Z = \frac{1}{2}v \sqrt{g_w^2 + g_w'^2} \quad (1.37)$$

$$m_\gamma = 0 \quad (1.38)$$

The Higgs field can be expanded around this minimum as:

$$\phi(x) = \frac{1}{\sqrt{2}} e^{\frac{i}{2}(\tau_j \theta_j(x) - Y \theta_0(x))} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.39)$$

where $\theta_j(x)$ and $\theta_0(x)$ are absorbed by the $SU(2)_L$ and $U(1)_Y$ gauge fields, respectively, and $h(x)$ is the real scalar field with one degree of freedom, i.e. the Higgs boson. With this choice of vacuum and gauge, the Lagrangian of the Higgs term is described as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - \frac{1}{2}\mu^2(v + h)^2 - \frac{1}{4}\lambda(v + h)^4 \quad (1.40)$$

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2}(\partial_\mu h)^2 + \frac{1}{4}g_w(v + h)^2 \left(W_\mu^\dagger W^\mu + \frac{1}{2} \frac{g_w^2 + g_w'^2}{g_w^2} Z^2 \right) \quad (1.41)$$

$$\frac{1}{2}\mu^2(v + h)^2 + \frac{1}{4}\lambda(v + h)^4 = \frac{1}{2}\mu^2 v^2 + \frac{1}{4}\lambda v^4 + \left(\frac{1}{2}\mu^2 + \frac{3}{2}v^2\lambda \right) h^2 + \lambda v h^3 + \frac{1}{4}\lambda h^4 \quad (1.42)$$

where the phase transformation does not change the Lagrangian. Only the real scalar field acts on the Lagrangian representing the Higgs interactions with SM particles. The $(D_\mu \phi)^\dagger (D^\mu \phi)$ describes the interaction terms between the Higgs boson and the vector bosons, the kinematic terms of the Higgs field, and the mass terms of the vector bosons, as shown in Equation 1.35. The remaining potential terms are the Higgs boson mass term (h^2) and self-coupling terms (h^3, h^4). From the scalar Lagrangian form shown in Table 1.1, the mass of the Higgs boson is:

$$m_H = \sqrt{2\lambda v^2}. \quad (1.43)$$

Taking the coefficients of h^3 and h^4 , λ indicates the Higgs self-coupling strength.

1.2.2 Yukawa Couplings

The Higgs field can also yield masses and interaction terms with fermions through their Yukawa couplings. The Lagrangian is:

$$\mathcal{L}_{\text{Yukawa}} = -(\bar{L}_L^\ell \phi) Y^\ell \ell_R - (\bar{L}_L^\nu \tilde{\phi}) Y^\nu \nu_R - (\bar{U}_L^u \tilde{\phi}) Y^u u_R - (\bar{D}_L^d \phi) Y^d d_R + h.c. \quad (1.44)$$

$$= -\frac{y_\ell v}{\sqrt{2}} \bar{\ell} \ell - \frac{y_\nu v}{\sqrt{2}} \bar{\nu} \nu - \frac{y_u v}{\sqrt{2}} \bar{u} u - \frac{y_d v}{\sqrt{2}} \bar{d} d - \left(\frac{y_\ell}{\sqrt{2}} \bar{\ell} \ell + \frac{y_\nu}{\sqrt{2}} \bar{\nu} \nu + \frac{y_u}{\sqrt{2}} \bar{u} u + \frac{y_d}{\sqrt{2}} \bar{d} d \right) h \quad (1.45)$$

where $\ell = (e, \mu, \tau)$, $\nu = (\nu_e, \nu_\mu, \nu_\tau)$, $u = (u, c, t)$, and $d = (d, s, b)$. All of these are in terms of mass eigenstates. The operator Y^f denotes the Yukawa coupling operator of the fermion and y_f is its coupling. The first four terms in Equation 1.47 describe the masses of fermions. Due to the Higgs field form in Equation 1.30, it can interact only with down-type fields. To restore the interaction with up-type fields, the charge conjugate ($i\tau^2$) is adopted to the Higgs field towards up-type fields. τ^2 describes one of the phases along the absorbed gauge fields, and it has no effects on the physics description of the Higgs. Furthermore, the charge conjugated form is orthogonal to the form in Equation 1.30:

$$\tilde{\phi} = i\tau^2 \phi = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ v+h \end{pmatrix} = \begin{pmatrix} v+h \\ 0 \end{pmatrix}. \quad (1.46)$$

From the mass term of the fermion Lagrangian in Table 1.1, the mass of the fermion is defined as:

$$m_f = \frac{v}{\sqrt{2}} y_f. \quad (1.47)$$

While the gauge boson masses can be determined from the known values of the weak interaction couplings, the fermion masses are completely free parameters and their Yukawa couplings are not predicted by the SM.

1.3 Standard Model Particles

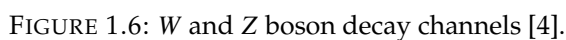
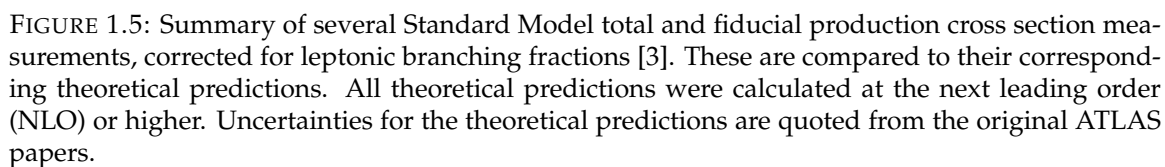
After all the SM particles were discovered, a wide range of measurements for SM properties has been performed at the Large Hadron Collider (LHC) by the CMS and ATLAS experiments. These experiments investigated the nature of gauge bosons and top quarks at the highest energy scales in the world ($O(\text{TeV} \sim 10 \text{ TeV})$). A summary of the SM production cross section measurements is shown in Figure 1.5, where all measurements agree with the SM predictions.

In the high energy experiments, electron, muon, neutrino, and photon are considered stable due to their long lifetime. Quarks and gluon other than top quark develop bremsstrahlung cascades of narrowly collimated gluons and quark-antiquark pairs via the strong interaction, forming the "jet".

For the W and Z boson, the decay channels are shown in Figure 1.6.

The τ lepton decays into $\bar{\nu}_\tau + e\bar{\nu}_e$, $\bar{\nu}_\tau + \mu\bar{\nu}_\mu$, or $\bar{\nu}_\tau + \bar{u}d$. The two decay channels including charged leptons in the final state are called leptonic decays, while one decay channel including quarks is called hadronic decay. The hadronic decay of the τ lepton generates a jet by pions originating from $\bar{u}d$.

The Higgs boson and top quark decays are described in following sections.



1.4 Higgs Boson

The Higgs boson was discovered in 2012 by the CMS and ATLAS experiments at the LHC. The main Higgs boson productions in the LHC experiments are shown in Figure 1.7, and cross sections are summarized in Table 1.2.

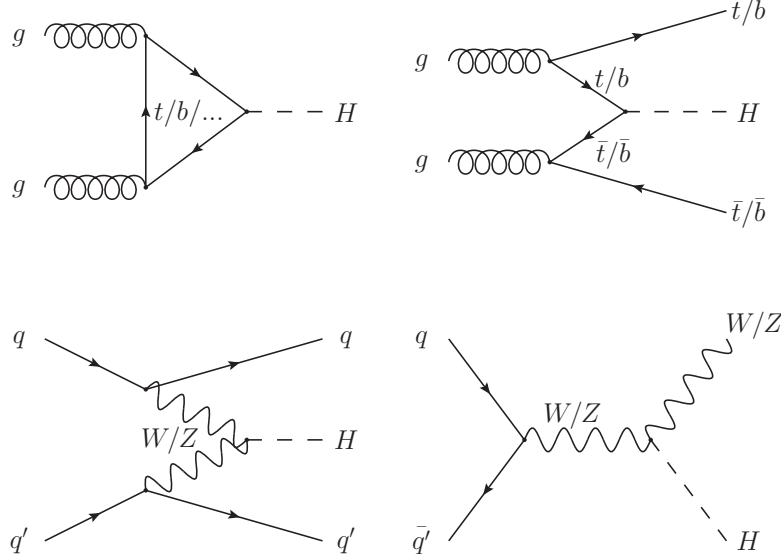


FIGURE 1.7: Four main Feynman diagrams for Higgs boson production at the LHC. Top-left shows gluon-gluon fusion (ggF), top-right is associated with a top/bottom-quark pair ($t\bar{t}H/b\bar{b}H$), bottom-left shows vector boson fusion (VBF), and bottom-right is associated with a vector boson (VH).

TABLE 1.2: Production cross sections for the Higgs boson assuming $m_H = 125.0$ GeV [5].

process	ggF	VBF	WH	ZH	$t\bar{t}H$	$b\bar{b}H$
cross section [pb]	44.14	3.782	1.373	0.8839	5.071	0.4880
unc.: $\pm$ QCD scale	+7.6% -8.1%	+0.4% -0.3%	+0.5% -0.7%	+3.8% -3.1%	+5.8% -9.2%	+20.2% -23.9%
unc.: $\pm$ (PDF+ α_S)	$\pm 3.1\%$	$\pm 2.1\%$	$\pm 1.9\%$	$\pm 1.6\%$	$\pm 3.6\%$	

The production cross sections are determined by a composition of the initial state partons from protons and their interaction matrix elements. The coupling strengths between the Higgs boson and SM particles are described by $h(x)$ terms in $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$. Gluon-gluon fusion (ggF) mediated by fermion loops shows the largest production cross section: the massless gluon does not couple directly to the Higgs boson. The larger contribution in the loop arises from the fermion with the larger coupling strength to the Higgs boson, corresponding to the fermion mass. Therefore, the top quark and bottom quark are the main contributors. The second and third largest production cross sections of the Higgs boson are vector boson fusion (VBF) and association with a vector boson (VH), respectively. They have a relatively large direct coupling between the Higgs boson and the vector boson. The associated productions with a quark pair are relatively small. So far, only $t\bar{t}H$ and $b\bar{b}H$ are accessible as an association with a top-quark and bottom-quark pair, respectively. They have a direct coupling between the Higgs boson and the quark. Because the top quark is heavier than the Higgs boson, its direct coupling can be observed only through the $t\bar{t}H$ production.

The decay channels of the Higgs boson are summarized in Table 1.3. The largest branching ra-

TABLE 1.3: Decay channels of the Higgs boson for $m_H = 125.0$ GeV [5].

process	$H \rightarrow b\bar{b}$	$H \rightarrow \tau\tau$	$H \rightarrow \mu\mu$	$H \rightarrow c\bar{c}$
Branching ratio	0.5824	6.272×10^{-2}	2.176×10^{-2}	2.891×10^{-2}
unc.: $\pm$ Theory	$\pm 0.65\%$	$+1.17\%$ -1.16%	$\pm 1.23\%$	$\pm 1.2\%$
unc.: $\pm$ mass parameter	$+0.72\%$ -0.74%	$+0.98\%$ -0.99%	$+0.97\%$ -0.99%	$+5.26\%$ -0.98%
unc.: $\pm \alpha_S$ parameter	$+0.78\%$ -0.80%	$\pm 0.62\%$	$+0.59\%$ -0.64%	$\pm 1.25\%$

process	$H \rightarrow gg$	$H \rightarrow \gamma\gamma$	$H \rightarrow Z\gamma$	$H \rightarrow WW^*$	$H \rightarrow ZZ^*$
Branching ratio	8.187×10^{-2}	2.270×10^{-3}	1.533×10^{-3}	0.2137	2.619×10^{-2}
unc.: $\pm$ Theory	$+3.40\%$ -3.41%	$+1.73\%$ -1.72%	$\pm 5.71\%$	$\pm 0.99\%$	$\pm 0.99\%$
unc.: $\pm$ mass parameter	$+1.12\%$ -1.13%	$+0.93\%$ -0.99%	$+0.98\%$ -1.01%	$+0.99\%$ -0.98%	$+0.99\%$ -0.98%
unc.: $\pm \alpha_S$ parameter	$+3.69\%$ -3.61%	$+0.61\%$ -0.62%	$+0.58\%$ -0.65%	$+0.66\%$ -0.63%	$+0.66\%$ -0.63%

tio is $H \rightarrow b\bar{b}$, where the bottom quark has the largest mass among particles with $m < m_H/2$. The analysis of this decay channel suffers from large background sources and a lack of well-designed triggers to obtain a signal with high efficiency while simultaneously rejecting backgrounds. The third largest is $H \rightarrow gg$, which connects the Higgs boson to gluons through heavy quark loops, similar to ggF production. This is more difficult because gluons produced by Higgs boson decay cannot be tagged unlike b -quarks, and backgrounds from the pile-up and other physics processes are enormous. The $H \rightarrow \tau\tau$ channel requires high performance τ identifications, and it is important to understand the missing E_T to reconstruct the Higgs mass. Afterwards, decay channels into vector bosons contain high momentum leptons or photons in the final states. These are significant even with their small decay ratios (e.g. $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$ and $H \rightarrow Z\gamma$). The mass of the Higgs boson can be also reconstructed in these channels. The $H \rightarrow \gamma\gamma$ process occurs via loops of charged fermions and a W boson. Because the coupling strength is largest between the Higgs boson and a W boson, a W boson is the largest contributor in this loop, followed by top and bottom quark loops. The $H \rightarrow \mu\mu$ channel is a very clean channel, but its decay ratio is too small to observe with available collision data.

A broad range of Higgs property measurements has been performed using proton-proton collision data produced by the LHC. Due to the clean signature and high S/B ratio, $H \rightarrow ZZ^*$ and $\gamma\gamma$ channels provide good signal sensitivities for the measurements, despite their small branching ratios. The largest background is non-resonant processes with continuous mass spectra. The ability to fully reconstruct the Higgs mass with excellent resolutions has been critical in establishing Higgs signals. Using two channels, the mass, cross sections, and kinematics of the production processes were measured.

The most precise Higgs boson mass was measured to be 125.09 ± 0.24 GeV by combining results from the CMS and ATLAS experiments in Run 1 [6]. The ATLAS experiment also measured the Higgs mass 124.98 ± 0.28 GeV with Run 2 data, as summarized in Figure 1.8. Its uncertainty was reduced by 32% from the ATLAS Run 1 measurement of 125.36 ± 0.41 GeV [7]. The main sources of systematic uncertainties come from experimental effects such as LAr calorimeter calibrations and material energy loss estimates in the $H \rightarrow \gamma\gamma$ channel.

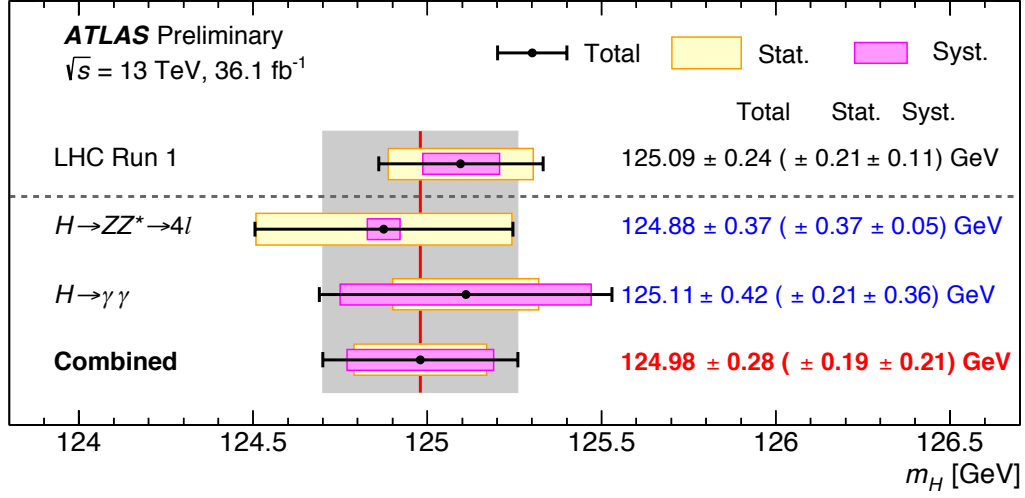


FIGURE 1.8: Summary of Higgs boson mass measurements from the individual and combined analyses, compared to the combined Run 1 measurement by the CMS and ATLAS [8].

Inclusive and differential fiducial cross sections of Higgs production were measured in $H \rightarrow ZZ^*$ and $\gamma\gamma$ channels [9, 10]. All of them indicate good compatibility with the standard model predictions, although they have still large statistical uncertainties with the pp collision data of 36.1 fb^{-1} at center-of-mass energy 13 TeV in Run 2. Figure 1.9 shows inclusive cross sections in the $H \rightarrow ZZ^*$ channel. Figure 1.10a shows the differential cross sections in terms of $p_{T,4\ell}$, which corresponds to $p_{T,\text{Higgs}}$. Thus, the production kinematics of the Higgs boson can be probed. This is the most precise measurement for $p_{T,\text{Higgs}}$. Differential distributions of the leading jet p_T were also measured in Figure 1.10b, which can probe QCD activity. The data tend to be larger than predictions in higher p_T bins in differential distributions. These deviations should be studied further with more accurate QCD calculations and more data to reduce statistical uncertainties. The same tendencies are seen also in the $H \rightarrow \gamma\gamma$ [10].

Combined measurements of Higgs boson production cross sections are presented in [11]. The result is shown in Figure 1.11 for the individual production processes of ggF, VBF, VH , and $t\bar{t}H$. No significant deviations from predictions are observed, although VBF is larger than the value predicted by the SM hypothesis by about 2 standard deviations.

1.5 Top Quark

The top quark has the largest mass among the SM particles ($m_{\text{top}} \sim 172.5 \text{ GeV}$), discovered in 1995 by the DØ and CDF experiments at the TEVATRON $p\bar{p}$ collisions ($\sqrt{s} = 1.8, 1.92 \text{ TeV}$). After that, the unique features and properties of top quark have been measured precisely by the CMS and ATLAS experiments using pp collisions at the LHC ($\sqrt{s} = 7, 8 \text{ and } 13 \text{ TeV}$). The top quark is suggested to play a special role in some beyond SM theories [12, 13, 14, 15] because of its large mass compared to other fermions. Furthermore, the top quark has a special feature in the mass value itself. The top quark mass value is approximately $m_{\text{top}} \sim v/\sqrt{2}$ ($v = 246 \text{ GeV}$), which corresponds to a top Yukawa-coupling of almost 1. This coupling value may imply a special relationship between the Higgs boson and the top quark. Even if the top quark has no connection with such a new theory, its measurement is very interesting because the top quark decays before it hadronizes due to its large mass and corresponding short lifetime. Its feature allows the bare

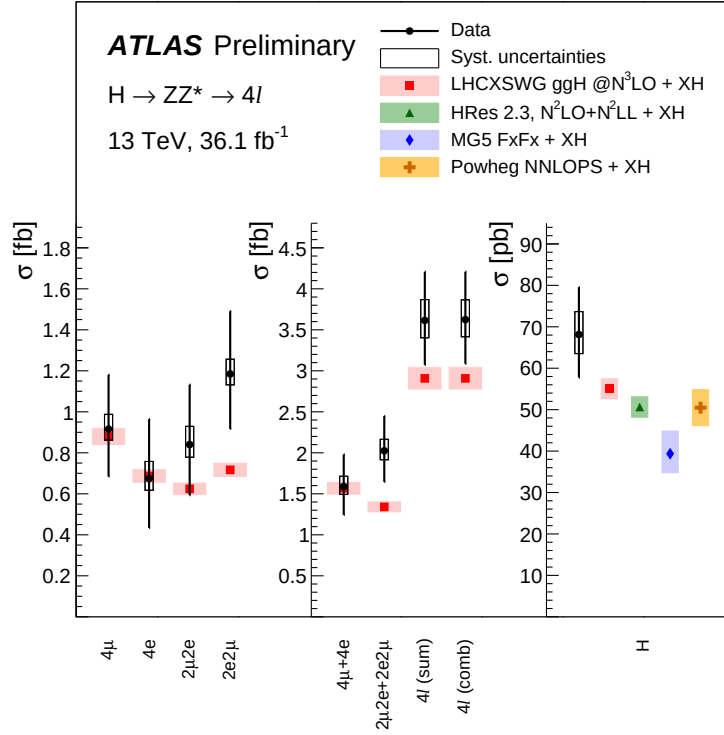


FIGURE 1.9: The fiducial cross sections (left two panels) and total cross section (right panel) of Higgs boson production measured in the 4ℓ final state [9]. The inclusive fiducial cross section is measured as the sum of all channels, as well as by combining the per-channel measurements assuming SM branching ratios.

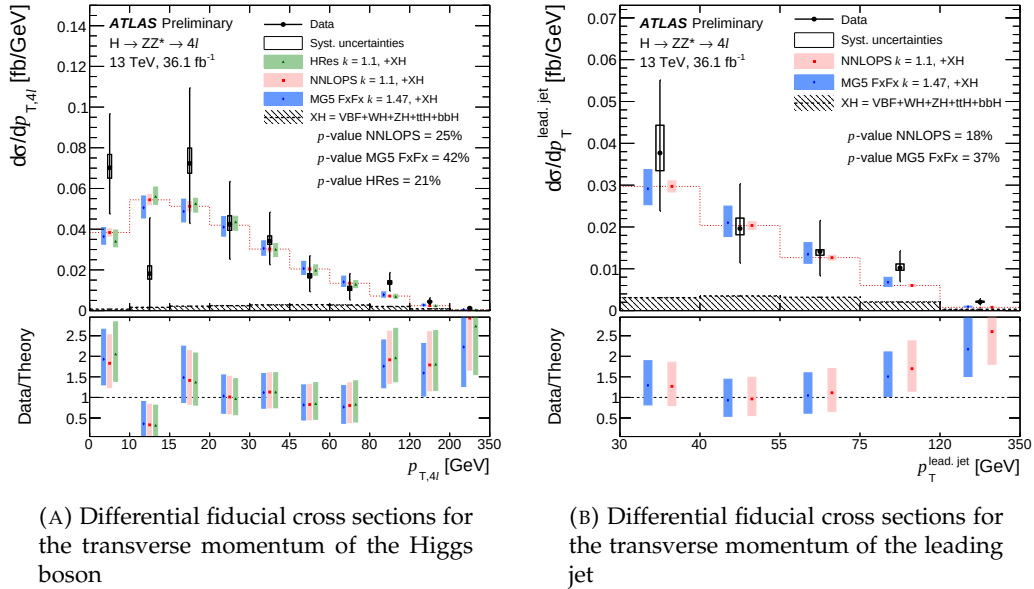


FIGURE 1.10: Differential fiducial cross sections obtained in $H \rightarrow ZZ^*$ channel [9].

quark properties to be investigated such that its spin information is directly transferred to the

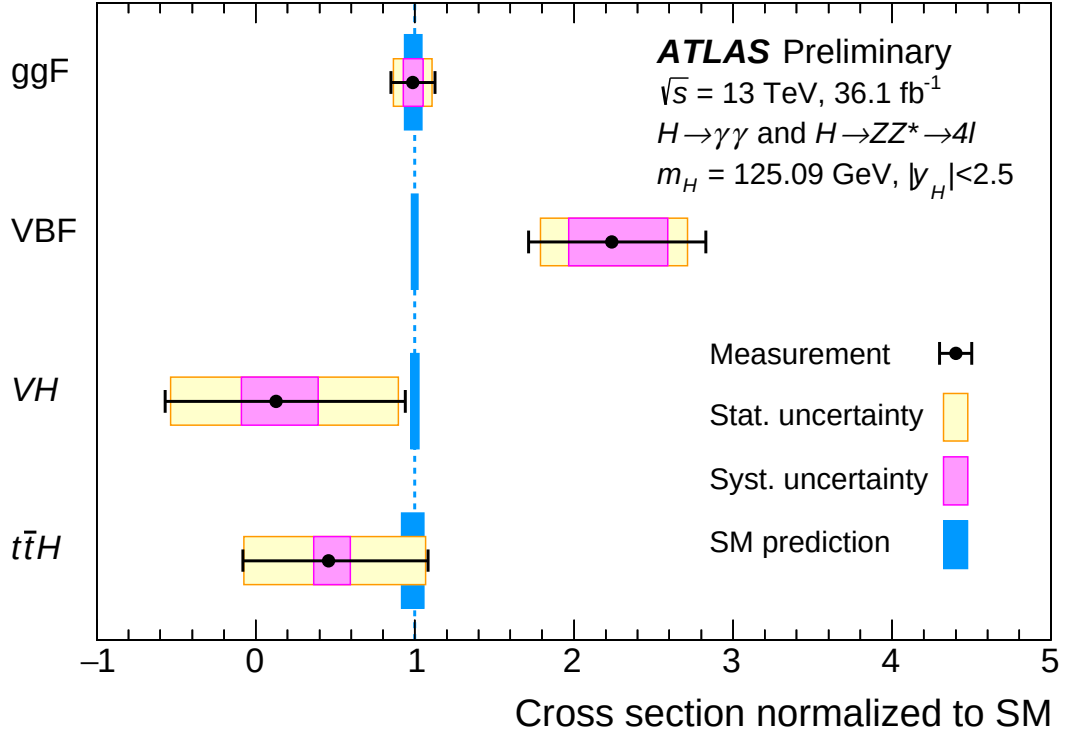


FIGURE 1.11: Cross sections for ggF, VBF, VH, and $t\bar{t}H$ processes measured with the assumption of predicted branching ratios [11]. The values are normalized with respect to the SM predictions.

decay products. In particular, top quark measurements are important for inspecting QCD, and thereby constrain QCD parameters [16, 17].

The top-quark is produced through the strong or EW interactions. By the strong interaction, a top-quark pair ($t\bar{t}$) is generated by gluon splitting, which has the largest cross section in top quark production (see Figure 1.5). A single top quark is produced in association with a W boson or bottom quark by the EW interaction. Main Feynman diagrams are shown in Figure 1.12.

The top quark decays via weak interactions accounting for the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Each of the matrix elements describes the coupling strength between up- and down-type quarks through the W bosonic weak interaction [4]:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix}. \quad (1.48)$$

Therefore, the top quark decays exclusively as $t \rightarrow Wb$ (Wtb vertex). The $t \rightarrow Wd$ and Ws channels are highly suppressed by $|V_{td}|$ and $|V_{ts}|$ terms corresponding to vertices of Wtd and Wts , respectively. Other decay modes into Zq or γq are called flavor-changing-neutral-current (FCNC), and they are also highly suppressed. The FCNC process is forbidden in the tree level, but it can occur at the one-loop level. However the GIM mechanism strongly suppresses loop contributions: $\text{BR}(t \rightarrow Zq) \sim \mathcal{O}(10^{-14})$. The experimentally observed (expected) upper limits at 95% confidence level are set on the $\text{BR}(t \rightarrow Zu)$ to 1.7×10^{-4} (2.4×10^{-4}), and on the $\text{BR}(t \rightarrow Zc)$ to 2.3×10^{-4} (3.2×10^{-4}) [18].

The $t\bar{t}$ analyses are characterized by the number of leptons emitted from two W bosons, as

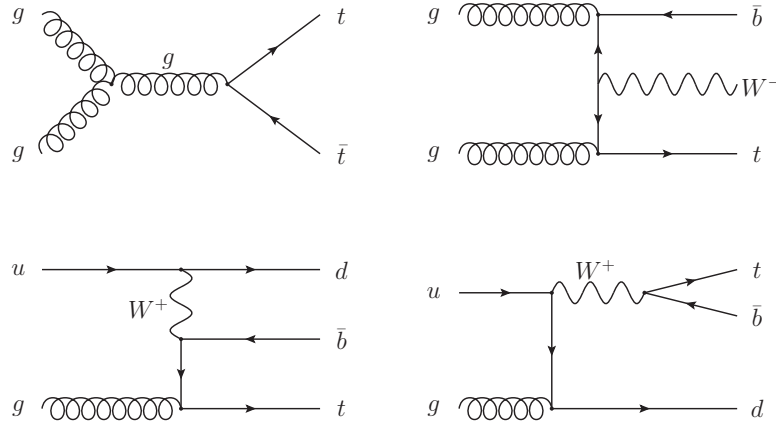


FIGURE 1.12: Four main Feynman diagrams for top quark production at the LHC. Top-left shows top quark pair production via the strong interaction. The other three show productions via the EW interaction. Top-right shows a single top associated with a W boson (Wt). Bottom-left and bottom-right are single top t and s channel productions, respectively.

shown in Figure 1.13. In the all-hadron channel, both W bosons decay hadronically, and no leptons

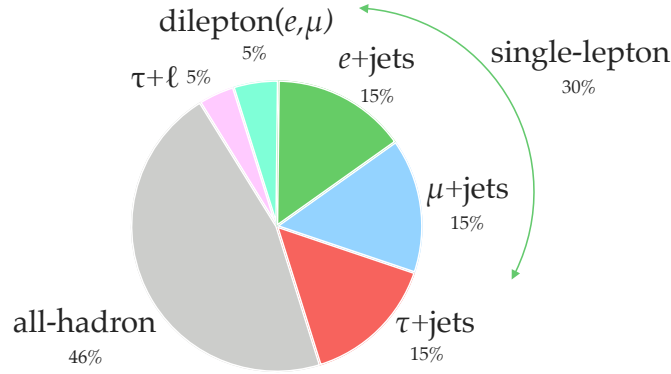


FIGURE 1.13: $t\bar{t}$ decay channels [4].

are included. Although the all-hadron channel has a largest fraction of the $t\bar{t}$ decays, the large QCD background worsens the sensitivity. The dilepton channel has almost 5% fraction of the $t\bar{t}$ decays but small background sources. This channel realizes the high sensitivity if a large amount of data is available. The single-lepton channel is a golden channel of the $t\bar{t}$ analysis, where the fraction of the $t\bar{t}$ decays is relatively large ($\sim 30\%$) and QCD background is suppressed by the existence of one lepton. The τ lepton is not stable unlike the light leptons (electron and muon), therefore, $t\bar{t}$ events including τ leptons are treated separately from events including only light leptons.

With large $t\bar{t}$ event statistics, precise cross section measurements have been performed both inclusively and differentially at the LHC as shown in Figure 1.14. Measurements of $t\bar{t}$ production cross section as functions of top and $t\bar{t}$ kinematic observables are especially important for verification of the QCD calculation in the SM framework. Furthermore, differential measurements are expected to be sensitive to new phenomena, although all the present results are consistent to the SM predictions.

The top quark mass is also important to constrain the vacuum stability combined with the mass of the Higgs boson as discussed later. The latest results of the CMS and ATLAS experiments are shown in Figure 1.15. The uncertainty on the top quark mass is about ± 0.5 GeV.

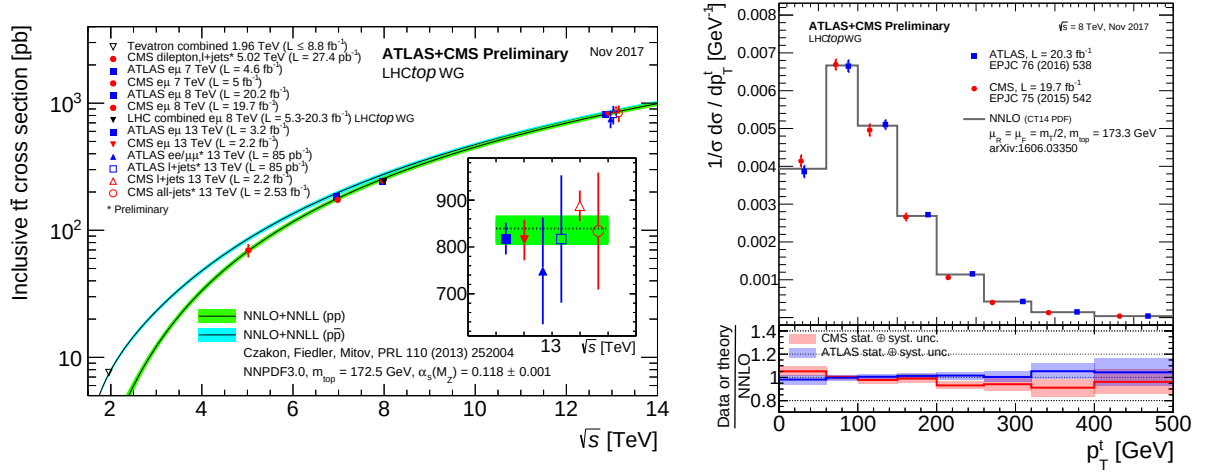


FIGURE 1.14: The top-quark pair production cross section [19]. The left plot shows inclusive cross sections at various energies at the LHC and TEVATRON colliders. The right plot shows the differential cross section as a function of the top quark's transverse momentum. All the results are consistent with SM predictions.

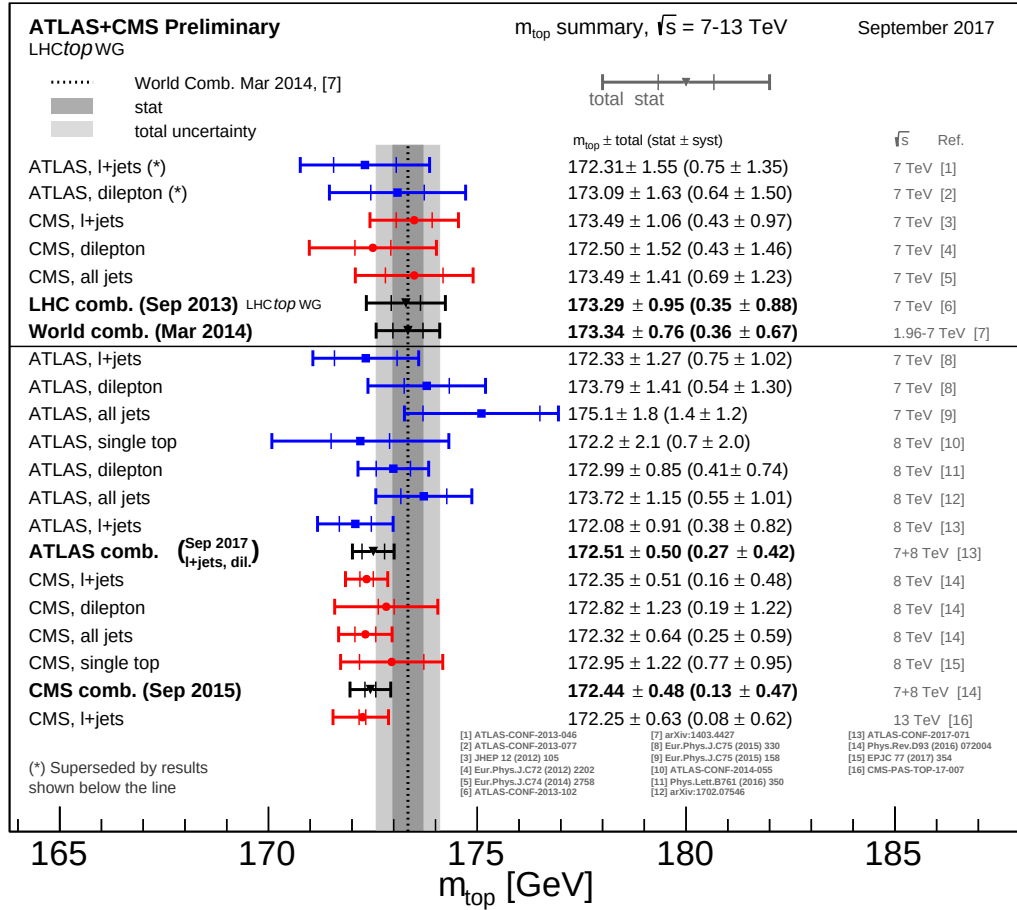


FIGURE 1.15: Various top mass measurements from the CMS and ATLAS experiments at $\sqrt{s} = 7, 8, 13$ TeV [19].

An important consistency test of the SM [20] was performed by simultaneously determining m_{top} and m_W from other various SM parameters, such as m_Z and its width, and g_s . A scan of the confidence level profile of m_W versus m_{top} is shown in Figure 1.16 with or without including the results of the m_H measurement (blue or grey). Both contours agree with the direct measurements (green bands or on the ellipse for two degrees of freedom).

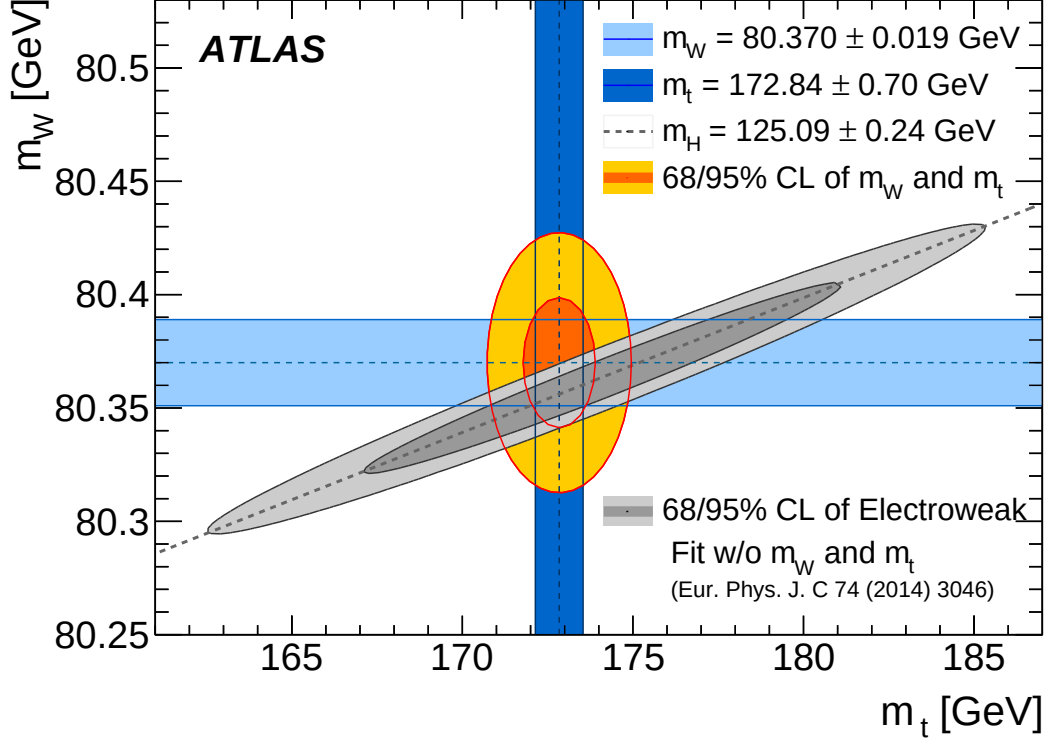


FIGURE 1.16: The 68% and 95% CL contours (red and yellow) of the m_W and m_{top} indirect determination from the global electroweak fit [20] are compared with the 68% CL contours of the ATLAS measurements of the top quark and W boson masses (blue) [21]. The determination from the electroweak fit (grey) uses as input the LHC measurement of the Higgs boson mass, $m_H = 125.09 \pm 0.24$ GeV [6].

1.6 Higgs Boson and Top Quark

By the success of the SM up to $\mathcal{O}(\text{TeV})$, we have not obtained any ideas for the energy scale of new physics, although the SM is known not to be complete to explain phenomena such as the dark matter and baryon asymmetry in the universe. In this section, the Higgs self-coupling is discussed with the renormalization group (RG) evolution to look for the energy scale where the SM becomes theoretically in contradiction with some observations under the assumption of absence of beyond the SM (BSM) signals [22].

The Higgs self-coupling strength (λ) is defined using Equations 1.43 and 1.36 as,

$$\lambda = \frac{m_H^2}{2v^2} = \frac{m_H^2 g_w^2}{8m_W^2} = \frac{m_H^2 G_F}{\sqrt{2}}, \quad (1.49)$$

where $G_F = \sqrt{2}g_w^2/8m_W^2$ is the fermi-constant precisely determined from the measurement of the muon life-time as $1.1663788(7) \times 10^5 \text{ GeV}^{-2}$ (0.6 ppm).

The RG evolution of the Higgs self-coupling at one loop is described as [22],

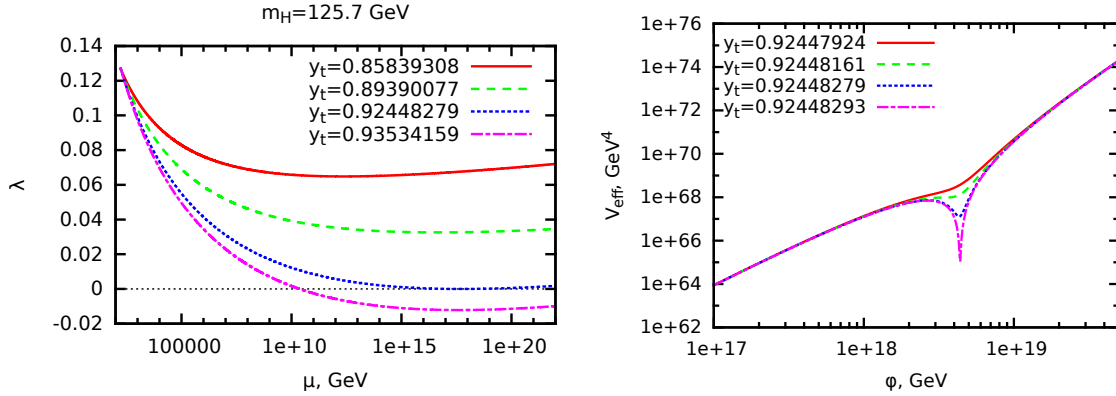
$$16\pi^2 \frac{d\lambda}{d \ln \mu} = 24\lambda^2 + 12\lambda y_t^2 - 9\lambda(g_w^2 + \frac{1}{3}g_w'^2) - 6y_t^4 + \frac{9}{8}g_w^4 + \frac{3}{8}g_w'^4 + \frac{3}{4}g_w^2 g_w'^2, \quad (1.50)$$

where y_t denotes the top Yukawa-coupling at the normalization point $\mu = 173.5 \text{ GeV}$ ($\sim m_{\text{top}}$). The evolution depends on the positive contributions from the weak bosons and the negative contributions from the top quark, where the contribution y_t^4 is largest. The SM parameters other than y_t are fixed to the experimental values and no BSM signals are assumed until the Planck scale, $\mathcal{O}(10^{19} \text{ GeV})$. As shown in Figure 1.17a, the Higgs self-coupling can become negative and the vacuum becomes unstable at some energy scale depending on the y_t value. The largest y_t value with $\lambda > 0$ up to the Planck scale is defined as y_t' . Taking the running feature with the top Yukawa, the Higgs potential $V(\phi)$ may show another local minimum ($\phi = \phi'_0$) with the consideration of the loop contribution as:

$$V(\phi) = \mu^2 \phi^2 + \lambda \phi^4 + \delta V_{\text{top}}(\phi) \quad (1.51)$$

$$\delta V_{\text{top}}(\phi) \propto -y_t^4 \phi^2 \log\left(\frac{\phi^2}{\mu^2}\right), \quad (1.52)$$

where $\delta V_{\text{top}}(\phi)$ is the loop contribution propagated by the top quark. Illustration of $V(\phi)$ is shown for several y_t settings in Figure 1.17b. When the current vacuum (ϕ_0) is degenerated with the new



(A) Running of the Higgs self-coupling λ for the Higgs mass $m_H = 125.7 \text{ GeV}$ and several values of the top-Yukawa y_t ($\mu = 173.2 \text{ GeV}$). Calculations consider λ evolution at four loops.

(B) Higgs potentials with various top-Yukawa y_t ($\mu = 173.2 \text{ GeV}$). A very small change in the top-Yukawa coupling converts the monotonic behavior of the Higgs potential to that with an extra minimum at the large energy scale.

FIGURE 1.17: Effects on (A) the Higgs self-coupling λ and (B) the Higgs potential by changing the top-Yukawa coupling value [22].

minimum (ϕ'_0) at a certain energy scale, $V(\phi = \phi'_0) = V(\phi = \phi_0)$, the state is called "critical" as $V(\phi = \phi_0) = 0$ in the present configuration. For the top Yukawa at the critical state (y_t^{crit}), four scenarios are possible for the vacuum stability:

- $y_t < y_t'$: the current vacuum at $\phi = \phi_0$ is the unique minimum in the Higgs potential,

- $y'_t < y_t < y_t^{\text{crit}}$: the current vacuum has the deeper potential than the new minimum, and the current vacuum is at the global minimum,
- $y_t^{\text{crit}} < y_t < y_t^{\text{crit}} + \delta y_t$: the new vacuum is deeper than the present vacuum. As the lifetime of the current vacuum is longer than the age of the universe, the current vacuum is meta-stable,
- $y_t^{\text{crit}} + \delta y_t < y_t$: the lifetime of the current vacuum is shorter than the age of the universe, and the current vacuum is unstable,

where numerically $y'_t = y_t^{\text{crit}} - 1.2 \times 10^{-6}$ and $\delta y_t = 0.04$, highly dependent on the loop corrections of the Higgs potential. When y_t is smaller than y_t^{crit} , the current vacuum is truly stable and there are no hints of BSM from this discussion. However, otherwise, the current vacuum is meta-stable or unstable dropping to the true minimum. The BSM is expected to appear to make the Higgs potential stable and prevent the negative Higgs self-coupling at the energy scale around or below μ_{new} . Figure 1.18 shows the dependence of the scale μ_{new} on the observed top-Yukawa value. Figure 1.19 shows the vacuum structure with relation to the Higgs mass and the top mass converted from the top-Yukawa with the assumption of the SM relationship. The top Yukawa coupling is highly important in discussion of the vacuum stability, and its observation may indicate the BSM energy scale.

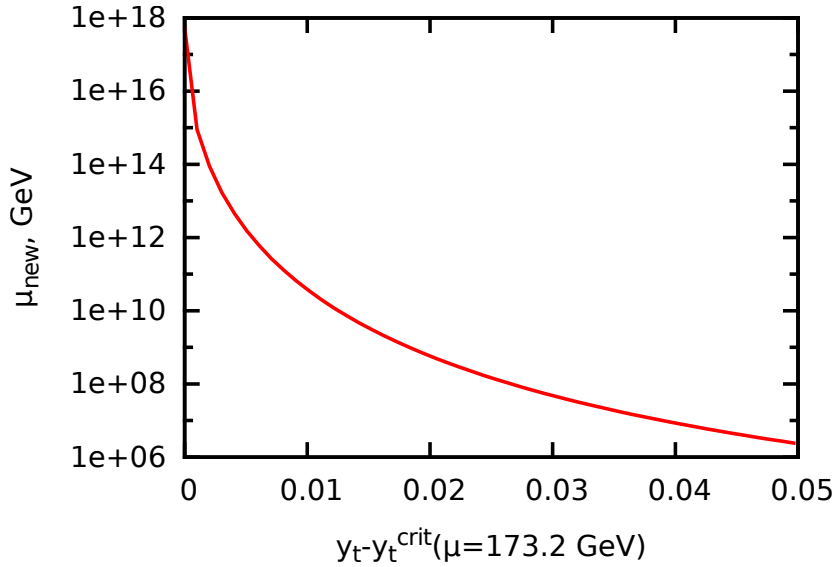


FIGURE 1.18: Energy scale where the Higgs self-coupling becoming negative (possibly requiring new physics at lower energies) depending on the top quark Yukawa [22].

1.7 Top-Yukawa-Coupling and Beyond the SM

If there is an enhancement in the top-Yukawa over the SM prediction, adding a small set of vector-like partners of the top and bottom quarks with masses at $O(\text{TeV})$ would give a good explanation [24]. The vector-like quarks (VLQ) that are partners of the heavy top and bottom quarks are predicted in many extensions of the SM. Their left- and right-handed fields transform in the same representation of $SU(2)$. A simple scenario [24] contains the usual SM gauge groups and matter fields, requiring addition of a vector-like quark $SU(2)$ doublet ($Q_{L,R}$) and two vector-like quarks

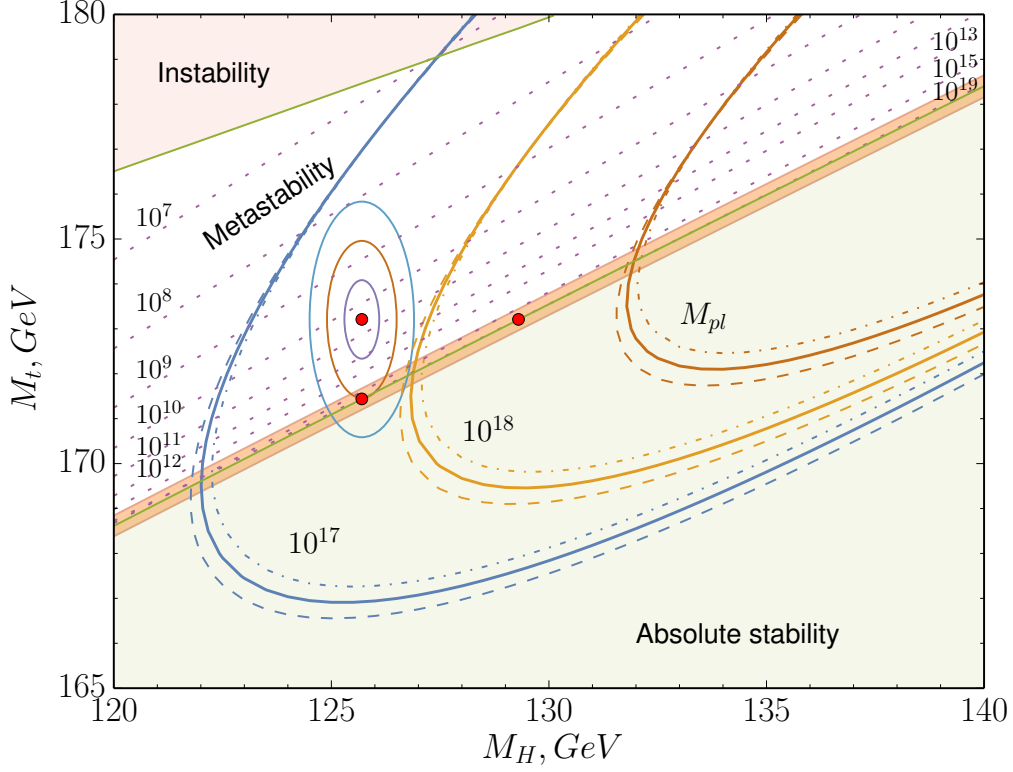


FIGURE 1.19: Phase diagram of vacuum stability (light-green shaded area), metastability (white area), and instability (pink shaded area) in the (m_H, m_{top}) plane [23]. Purple dotted curves shows the scale μ_{new} . Contours of $\beta_\lambda(\mu_{\text{new}}) = 0$ are denoted by solid parabola-like lines with uncertainties due to 1σ error (dashed and dot-dashed lines). The observed (m_H, m_{top}) (upper left red bullet) with the uncertain 1σ (purple ellipse), 2σ (brown ellipse), and 3σ (blue ellipse) are marked for reference. Critical state is shown by the solid green line with uncertainty due to 1σ error (orange shaded band), and the critical points with observed m_H (lower red bullet) and m_{top} (right red bullet) are also shown.

$SU(2)$ singlets ($T_{L,R}$ and $B_{L,R}$), one with up-type weak charge and another with down-type weak charge:

$$Q_{L,R} = \begin{pmatrix} T \\ B \end{pmatrix}_{L,R}, \quad T_{L,R}, \quad B_{L,R}, \quad (1.53)$$

They are assumed to obtain their masses not through the Higgs coupling. These VLQs can couple with SM quarks through Yukawa interactions, and mixing with other quarks can modify the Yukawa coupling strength:

$$(\bar{\mathcal{U}}_L^t \tilde{\phi}) Y^t t_R \rightarrow (\bar{U}_L^i \tilde{\phi}) Y^t_{ij} K_R^j, \quad (1.54)$$

where U_L and K_R are sets of left- and right-handed fields described as,

$$U_L = (\mathcal{U}_L^t, Q_L, T_L), \quad (1.55)$$

$$K_R = (t_R, Q_R, T_R). \quad (1.56)$$

The Yukawa operator using the above fields as the bases is described as the matrix as,

$$\mathbf{Y}^t = \frac{1}{\sqrt{2}} \begin{pmatrix} vy_t & 0 & vy_{qT} \\ vy_{Q_t} & 0 & vy_1^t \\ 0 & vy_2^t & 0 \end{pmatrix}. \quad (1.57)$$

The Higgs Yukawa couplings ($\mathbf{Y}^t$) are obtained after transforming the Yukawa matrices into the physical basis. Then, the relationship of Equation 1.47 between the top quark mass and the top Yukawa-coupling is not valid anymore. The mixing term with the VLQs results in the modification of the Yukawa coupling as,

$$y_t^{\text{phys}} = y_t + \delta y_t. \quad (1.58)$$

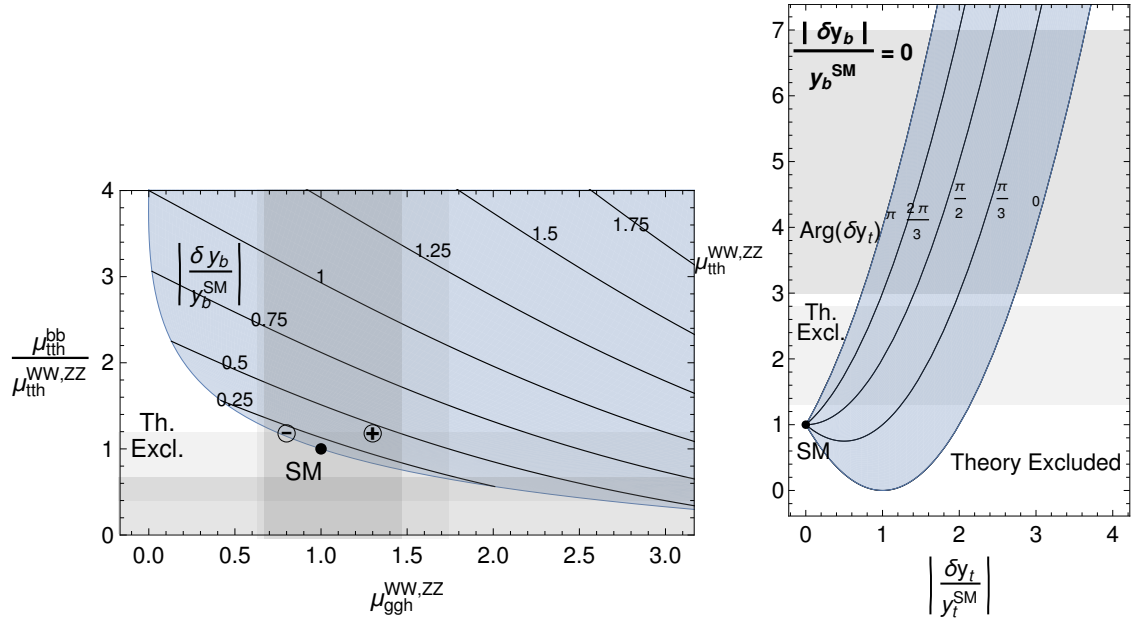
The bottom Yukawa operator ($\mathbf{Y}^b$) and coupling (y_b^{phys}) can be defined similarly as for the top.

The Higgs production cross section in ggF has been observed to be in good agreement with the SM predictions with the smallest uncertainty among four production modes (Figure 1.11). Therefore, the top-Yukawa modification should be assumed not to enhance the ggF production. To suppress the loop contributions from the top and bottom quarks, one constraint called "Brane Higgs Limit" is imposed as,

$$\det \mathbf{Y}^t = \det \mathbf{Y}^b = 0. \quad (1.59)$$

The most interesting feature of this model is allowing different Yukawa deviations depending on the Higgs decay channels. The available phase is shown in Figure 1.20. This model has a simple theoretical realization, is highly predictable, and can be tested (or ruled out) by precise measurements of the Higgs signal strength. Especially, the $t\bar{t}H$ production (top Yukawa) measurements are a strong discriminant to test the VLQ models.

Direct searches for the VLQs have been performed by the CMS and ATLAS experiments. As shown in Figure 1.21, no excesses are observed.



(A) Contours of the bottom quark Yukawa correction with respect to the signal strength of ggF with $H \rightarrow VV$ and the ratio of the $t\tau H$ cross sections with $H \rightarrow b\bar{b}$ and $H \rightarrow VV$. Each contour is traced by varying the phase of δy_b .

(B) Higgs signal strength $t\tau H$ with $H \rightarrow VV$ with respect to the top quark Yukawa coupling correction. The contours trace points with constant value of the phase of δy_t . We consider a parameter space point where the bottom Yukawa coupling is SM-like ($\delta y_b = 0$).

FIGURE 1.20: Phase space for VLQ contributions with respect to modifications in the signal strengths and in Yukawa couplings [24]. The horizontal and vertical grey bands represent the experimentally allowed bounds ($\pm 1\sigma$) set by the LHC Run 1 (darker) and the preliminary data from LHC early Run 2 (lighter). The "Th. Excl." region comprises all points excluded by the Brane Higgs Limit constraint.

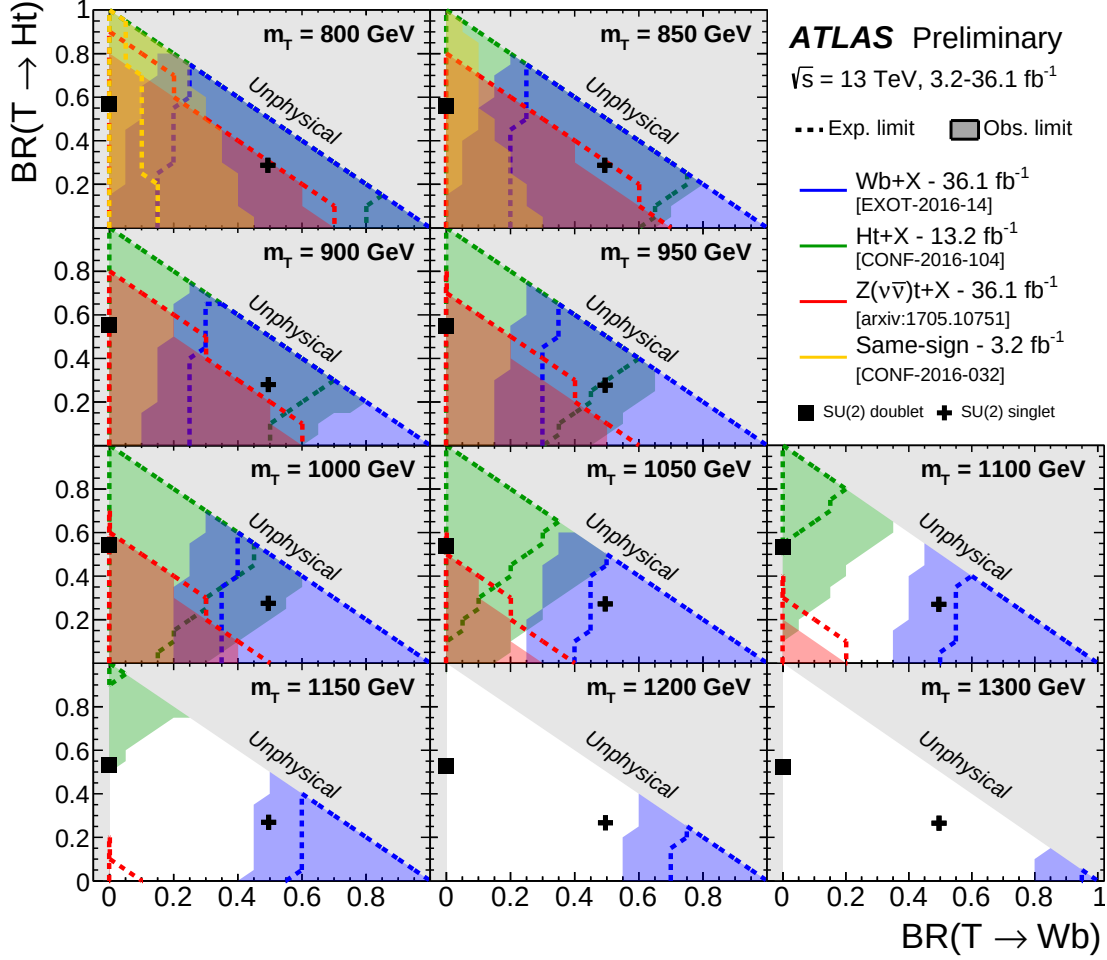


FIGURE 1.21: Observed (filled area) and expected (dashed line) 95% CL exclusion of the vector-like T quark mass in the plane of $\text{BR}(T \rightarrow Ht)$ versus $\text{BR}(T \rightarrow Wb)$ [25]. Results of four search channels are plotted: $Wb + X$ (blue), $Ht + X$ (green), $Z(\nu\bar{\nu})t + X$ (red) and the same-sign leptons (yellow). The grey area corresponds to the unphysical region where the sum of branching ratios exceeds unity. The default branching ratio values for the weak-isospin singlet and doublet cases are shown as cross and square symbols, respectively.

1.8 $t\bar{t}H$ Production Analysis

The $t\bar{t}H$ production has not been discovered. Although the $t\bar{t}H$ production cross section is around 1% of the total Higgs boson production, the $t\bar{t}H$ process offers a distinct signature due to the numerous final state objects that depend on the Higgs boson and $t\bar{t}$ decay. The search for the Higgs boson produced via $t\bar{t}H$ is performed in several analyses depending on the Higgs decay:

- $H \rightarrow b\bar{b}$ [26] (discussed in this thesis)
- $H \rightarrow \gamma\gamma$ [27]
- $H \rightarrow ZZ^* \rightarrow 4\ell$ [28]
- $H \rightarrow WW^*, H \rightarrow \tau\tau$ ($H \rightarrow$ multileptons) [29]

Each analysis is sub-divided by the number of leptons and hadronically decaying τ s in the final state. The analysis categorization is summarized in Figure 1.22.

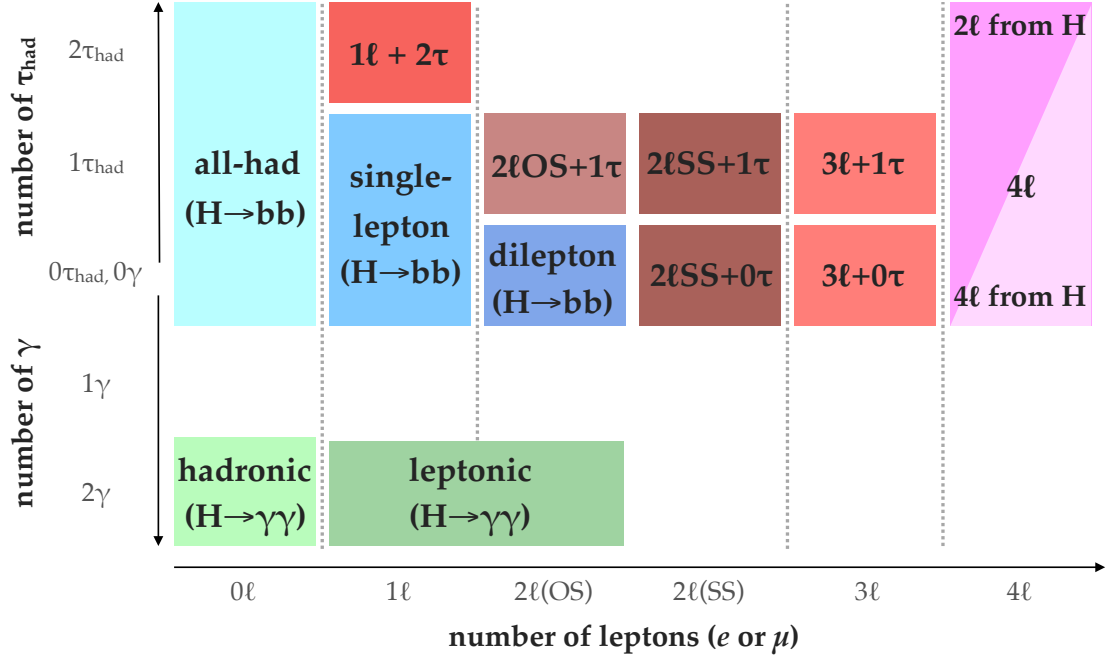


FIGURE 1.22: Analysis channels in the search for the Higgs boson produced via $t\bar{t}H$. Channels are separated by the number of leptons, hadronically decaying τ s, and photons. The channels with green, red, and blue squares are combined into the $H \rightarrow \gamma\gamma$, $H \rightarrow b\bar{b}$, and $H \rightarrow$ multileptons analysis, respectively.

1.8.1 $t\bar{t}H(H \rightarrow b\bar{b})$ Analysis

Using the largest branching ratio for the Higgs boson, $H \rightarrow b\bar{b}$ of 58%, the search for this production was performed in this thesis. In order to trigger on signal events with significant rejections of QCD and multi-jet backgrounds, the $t\bar{t}H(H \rightarrow b\bar{b})$ channel with at least one lepton emitted in $t\bar{t}$ (single-lepton and dilepton channels) were chosen. With the large $H \rightarrow b\bar{b}$ branching ratio and

detection of associated $t\bar{t}$ in leptonic decay channels, we expect large statistics and significant rejection of dominating QCD backgrounds.

In the single-lepton analysis channel, one of the W bosons from $t\bar{t}$ decays into $\ell\nu$ (ℓ is electron or muon), and the other decays into quarks:

$$t\bar{t}H \rightarrow bW^+\bar{b}W^-b\bar{b} \rightarrow b\nu\ell^+ \bar{b}q\bar{q}' b\bar{b} \text{ or } bq\bar{q}' \bar{b}\nu\ell^- b\bar{b}.$$

The single-lepton channel has one lepton (electron or muon) and 6 jets in the final state, where four of jets are b -jets. In the dilepton channel, two top quarks in the $t\bar{t}H$ system decay into b -quark and W boson, and both W bosons decay into $\ell\nu$. The final state objects of the dilepton $t\bar{t}H$ are 2 leptons with opposite charges (e^+e^- , $\mu^+\mu^-$, or $e^{+/-}\mu^{-/+}$) and 4 b -jets.

$$t\bar{t}H \rightarrow bW^+\bar{b}W^-b\bar{b} \rightarrow b\nu\ell^+ \bar{b}\nu\ell^- b\bar{b}$$

Both Feynman diagrams are shown in Figure 1.23.

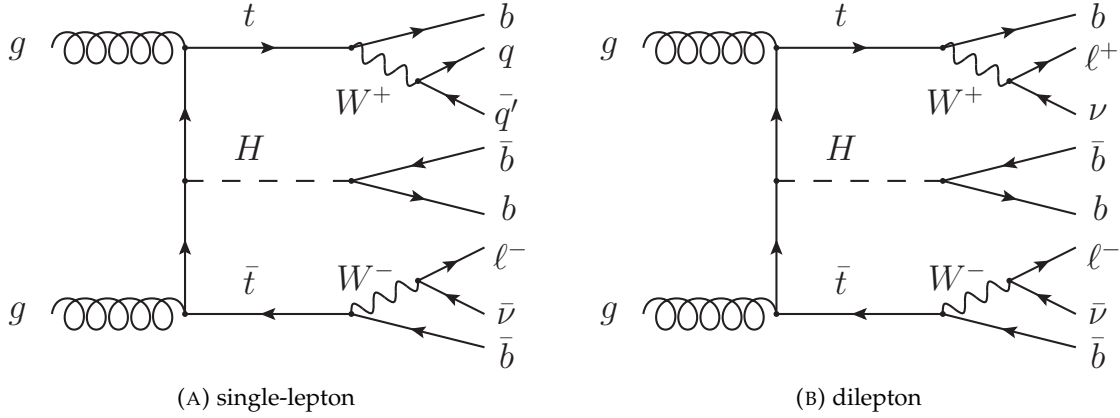


FIGURE 1.23: Feynman diagrams of $t\bar{t}H(H \rightarrow b\bar{b})$.

The partons and reconstructed objects in the $t\bar{t}H$ single-lepton channel are specified as:

- $b1_{\text{Higgs}}$: leading p_T b -jet (parton) of the decay product from Higgs boson
- $b2_{\text{Higgs}}$: sub-leading p_T b -jet (parton) of the decay product from Higgs boson
- hadTop : top-quark decaying to b and a hadronically decaying W boson
- lepTop : top-quark decaying to b and a leptonically decaying W boson
- b_{hadTop} : b -jet (parton) of the decay product from hadTop
- W_{had} : hadronically decaying W boson from hadTop decay
- $q1_{\text{hadW}}$: leading p_T light-jet (parton) of the decay product from W_{had}
- $q2_{\text{hadW}}$: subleading p_T light-jet (parton) of the decay product from W_{had}
- b_{lepTop} : b -jet (parton) of the decay product from lepTop
- W_{lep} : leptonically decaying W boson from lepTop decays
- $b1_{t\bar{t}}$: leading p_T b -jet (parton) among the decay products from $t\bar{t}$

- $b_{2t\bar{t}}$: subleading p_T b -jet (parton) among the decay products from $t\bar{t}$

The partons and reconstructed objects in the dilepton channel are also specified as:

- $b_{1\text{Higgs}}$: leading p_T b -jet (parton) of the decay product from the Higgs boson
- $b_{2\text{Higgs}}$: subleading p_T b -jet (parton) of the decay product from the Higgs boson
- top: top-quark decaying to b and a leptonically decaying W^+ boson
- anti-top: anti-top-quark decaying to b and a leptonically decaying W^- boson
- b_{top} : b -jet (parton) of the decay product from top decay
- W_{top}^+ : leptonically decaying W^+ boson from top decay
- $b_{\text{anti-top}}$: b -jet (parton) of the decay product from anti-top decay
- $W_{\text{anti-top}}^-$: leptonically decaying W^- boson from anti-top decay

Chapter 2

LHC and ATLAS Detector

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) [30] is a circular proton-proton collider with a circumference of 27 km located at the European Council for Nuclear Research (CERN). A schematic view is illustrated in Figure 2.1.

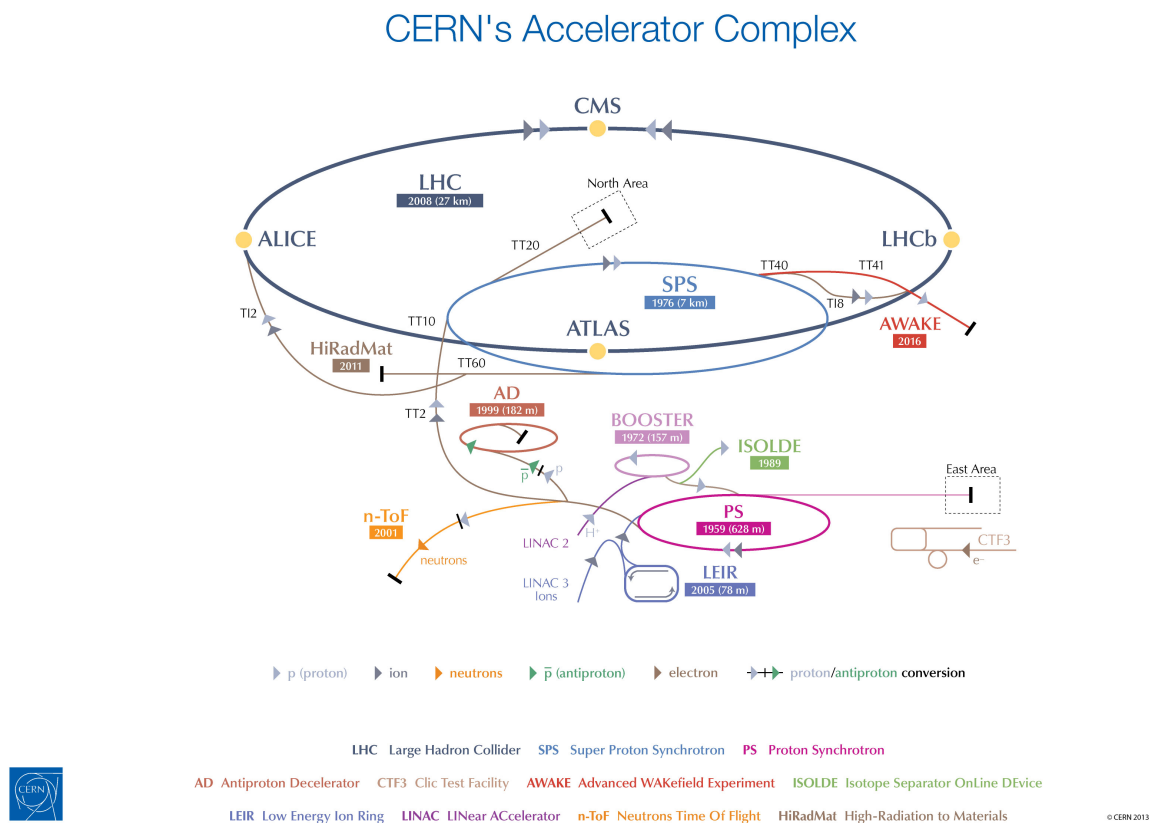


FIGURE 2.1: Schematic view of the CERN accelerator complex [31]. The four main LHC experiments are shown at the interaction points.

The LHC beams are gradually increased by subsequent acceleration steps up to the highest energy of 6.5 TeV (current beam energy). Protons are produced from the ionized hydrogen gas. The generated protons move through the LINAC2 and are accelerated up to 50 MeV. They are injected to the Proton Synchrotron Booster (PSB) 157 m in circumference, which increases their

energy up to 1.4 GeV. Afterwards, protons are injected to the Proton Synchrotron (PS) ring with a circumference of 628 m and are accelerated to 25 GeV. The PS squeezes protons into tight bunches that are the base bunch structure of the LHC. The last chain of proton acceleration occurs in the 7-km-circumference Super Proton Synchrotron (SPS), where the protons are brought to an energy of 450 GeV. These protons are injected into the LHC through two 2.5-km-long transfer lines. The main acceleration up to a maximum of 7 TeV can be achieved with eight resonant radio frequency (RF) cavities per beam. The RF electric field oscillates at 400 MHz and increases the beam energy by ~ 0.5 MeV per turn. The field intensity used to accelerate the 7 TeV beam is ~ 5.5 MV/m. In addition to the RF cavities, a total of 1232 dipole magnets have been assembled in the LHC to maintain the energy at 7 TeV. Each magnet produces a bending field of up to 8.33 T using superconducting coils made of NbTi. Temperature is maintained at 1.9 K in helium super-fluid. The LHC was operated at 3.5 TeV beam energy in 2011, 4 TeV in 2012 (Run 1), and 6.5 TeV since 2015 (Run 2). The machine parameters are shown in Table 2.1.

TABLE 2.1: Machine parameters of the LHC. Values taken from [32].

parameter	2016	2015	2012
$m_{\text{proton}}\gamma$: beam energy [TeV]	6.5	6.5	4
$1/f_{\text{rev}}$: bunch spacing [ns]	25	25	50
β^* at the interaction point [cm]	40	80	60
θ_c : crossing angle [μrad]	140	145	145
ε : normalized emittance at the start filling [μm]	2.0	3.5	2.2
n : max. bunch population [$10^{11}p/\text{bunch}$]	1.15	1.15	1.6
max. number of bunches per injected train	96	144	144
N_b : max. number of bunches	2220	2244	1374
$\mathcal{L}(t)$: instantaneous luminosity [$10^{34}\text{cm}^{-2}\text{s}^{-1}$]	1.4	0.5	> 0.7

The LHC beam intensity is described by its instantaneous luminosity, $\mathcal{L}(t)$, which is equal to the frequency of particles encountering each other per unit of area and time. If a physics process with cross section σ occurs with $\mathcal{L}(t)$ luminosity, the interaction rate is $\sigma \times \mathcal{L}(t)$. The total number of interactions (N) is described as:

$$N = \int_{t_1}^{t_2} [\sigma \times \mathcal{L}(t)] dt = \sigma \times \mathcal{L}_{\text{int}}. \quad (2.1)$$

If two beams are assumed to have the same parameters and approximately Gaussian lateral distributions, the instantaneous luminosity of the beam can be represented as:

$$\mathcal{L}(t) = N_b \times \mathcal{F} \frac{n^2 f_{\text{rev}} \gamma}{4\pi \varepsilon \beta^*} \quad (2.2)$$

where N_b is the number of bunches, n is the number of protons per bunch, f_{rev} is the revolution frequency of the beam, γ is the Lorentz factor, β^* is the beta-function at the interaction point, ε is the normalized beam emittance, and $\mathcal{F}$ is a geometric luminosity reduction factor that takes into

account the beam crossing angle at the interaction point. The $\mathcal{F}$ is generally described as:

$$\mathcal{F} = 1/\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma_{xy}^*}\right)^2} \quad (2.3)$$

where θ_c is the full crossing angle of the two beams at the interaction point, σ_z is the bunch length, and σ_{xy}^* is the transverse beam size.

2.2 ATLAS Detector

A Toroidal LHC Apparatus (ATLAS) detector [33] is built at one of the collision points in the LHC. The ATLAS detector (Figure 2.2) is designed as a general-purpose detector and used for measurements of various physics processes. The detector has well balanced detector performance and is used to identify and measure the generated objects, such as electrons, muons, photons, and hadrons. A schematic illustration of particle identification is shown in Figure 2.3.

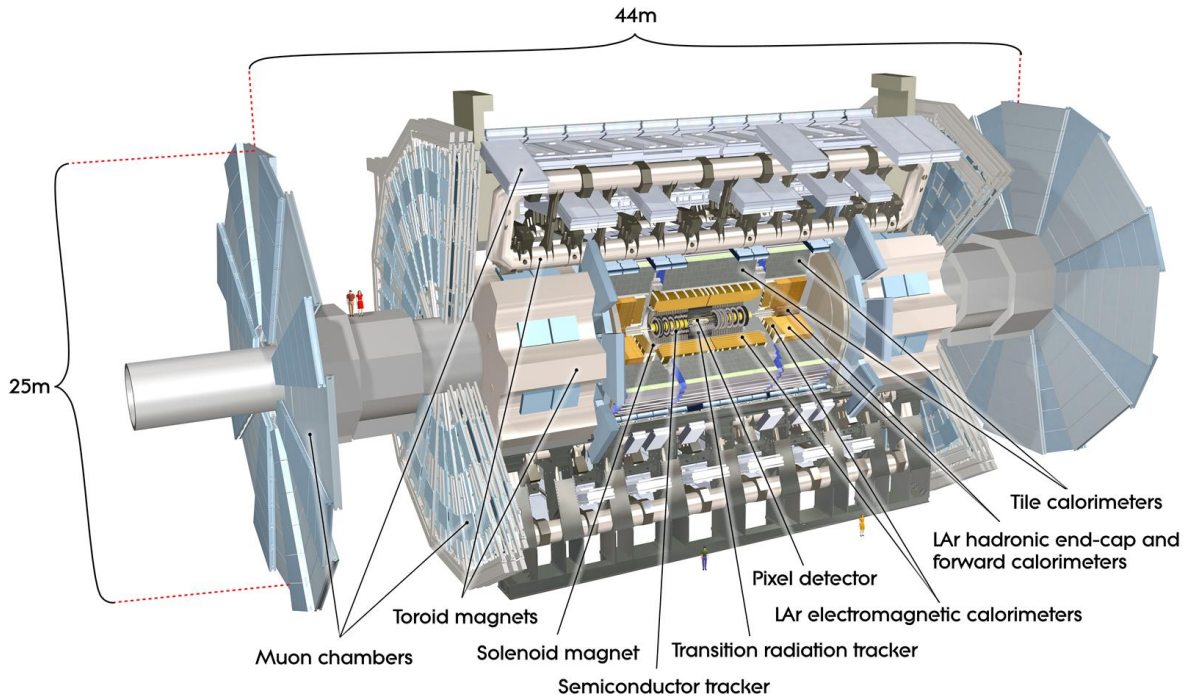


FIGURE 2.2: Illustration of the ATLAS detector [33].

The ATLAS detector is primarily composed of three components. The most inner part is the inner detector, which measures momenta of charged particles in the solenoid magnet. The middle part is a calorimeter which measures particle energies by stopping particles inside the calorimeter. The electromagnetic (EM) calorimeter measures particle energies via electromagnetic interactions, while the hadronic (HAD) calorimeter measures via both electromagnetic and strong interactions. The outer part is a muon spectrometer, which measures momenta of charged particles penetrating through the calorimeter.

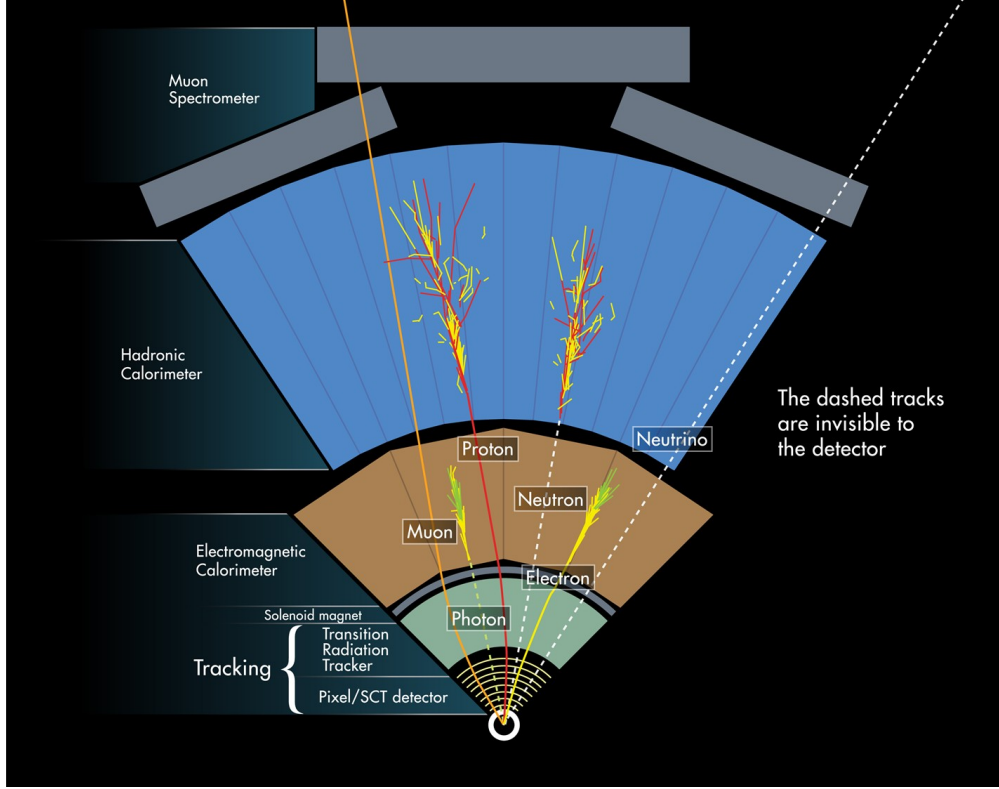


FIGURE 2.3: Illustration of the detection of stable particles are detected in the ATLAS detector [34]. Neutrinos escape from the ATLAS detector, leaving missing energy.

2.3 The ATLAS Coordinate System

The ATLAS detector employs the right-handed coordinate system with the x -axis pointing to the LHC center, the vertical y -axis pointing up and the z -axis along the beam line. The origin is set at the center of the ATLAS detector. The ATLAS coordinate system is shown in Figure 2.4. The transverse plane is defined as the x - y plane, and a cylindrical coordinate system (r, ϕ, z) is adopted where r is the radial distance on the transverse plane and ϕ denotes the azimuth angle around the beam line. The concentric and cylindrical sub-detectors in the central part are called "barrel", and the disk-shaped sub-detectors at the sides of the barrel are called "end-cap". The end-cap part in the $z > 0$ and $z < 0$ denote A-side and C-side, respectively.

In hadron collider experiments, it is useful to define the pseudo-rapidity η instead of the polar angle θ :

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (2.4)$$

η becomes zero if a particle moves in the transverse plane ($\theta = \pi/2$), while η becomes $\pm\infty$ if a particle moves along the z -axis ($\theta = 0, \pi$).

Using this coordinate system, a four-vector is presented as (p_T, η, ϕ, E) , where the transverse momentum p_T is the momentum of the particle projected on the transverse plane, and E denotes the energy of the particle. The p_T can be described with the three-dimensional momentum p ,

$$p_T = p \sin \theta. \quad (2.5)$$

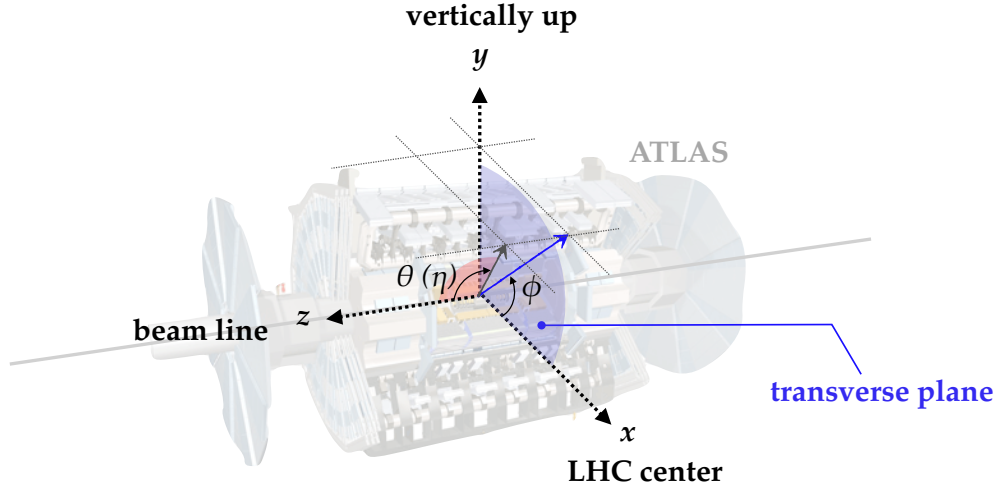


FIGURE 2.4: ATLAS coordinate system

Transverse energy $E_T = E \sin \theta$ is used in addition to p_T . The rapidity of the particle y is defined using the longitudinal momentum component along the z -axis p_z ,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (2.6)$$

The distance between two objects is often described as the angular distance using η or y :

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \quad (2.7)$$

$$\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}. \quad (2.8)$$

The interaction points where two protons collide in the ATLAS detector are called vertices. The closest approach of the object's trajectory to the vertex is called the impact parameter (IP), which is denoted as d_0 on the transverse plane, and z_0 along the z -axis.

2.4 Magnet

The magnet system is composed of four large superconducting magnet sub-systems cooled to 4.5 K by liquid helium, and is designed to provide magnetic fields that bend the trajectory of the charged particle. It consists of one central solenoid (CS) and three open-air toroids, as shown in Figure 2.5. The combination usage has the advantage of extending the coverage of $|\eta|$ coverage up to 3. Furthermore, the open-air design of the toroids reduces the impact of multiple scattering on momentum resolution.

The CS with a diameter of 2.5 m covers the inner detector volume. It provides a magnetic field to the inner detector parallel to the beam axis. The magnetic field is uniform with 2 T field strength. The field strength decreases slightly along the beam direction due to the finite length of the CS. The toroidal systems generate the magnetic field in the ϕ direction for the muon spectrometer. There are two 5-m long cylindrical end-cap toroid (ECT) systems, each with an external diameter of 10.7 m at the extremities of the detector, and one 25.3-m long cylindrical barrel toroid (BT) system with a diameter of 20 m that is centrally located around the calorimeters. Each toroid system is composed of eight rectangular coils arranged in the ϕ around the beam axis. The generated

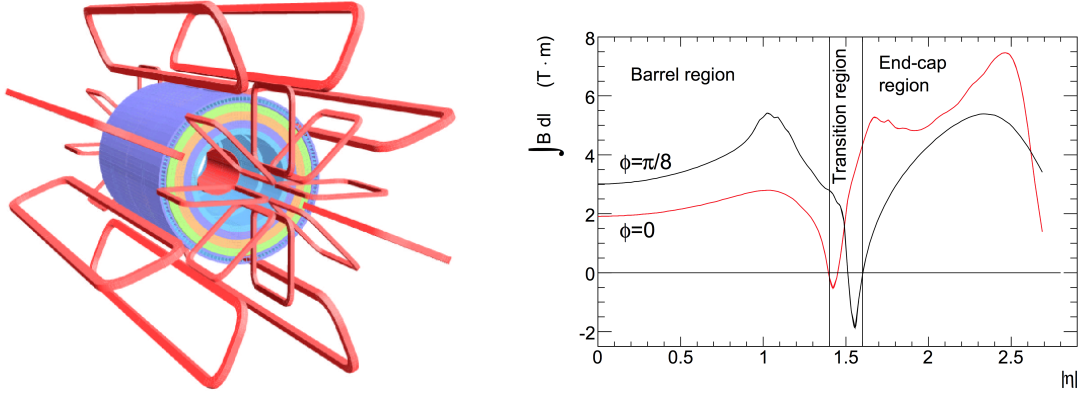


FIGURE 2.5: (Left) Illustration of the magnet [35]. Outer red rings show toroids and the inner red cylinder shows the CS. (Right) Predicted field as a function of $|\eta|$ integrated from the innermost to the outermost MDT layers in one toroid octant, for infinite momentum muons. The curves correspond to $\phi = 0$ (red) and $\phi = \pi/8$ (black).

magnetic fields provide 1.5 to 5.5 Tm of bending power in $0 < |\eta| < 1.4$, and ~ 1 to 7.5 Tm in $1.6 < |\eta| < 2.7$. In the transition region of $1.4 < |\eta| < 1.6$, the BT extends over the ECTs.

2.5 Inner Detector

From the most inner part, the pixel detector (PIXEL), semiconductor tracker (SCT), and transition radiation tracker (TRT) are installed in the inner detector covering $|\eta| < 2.5$ as shown in Figure 2.6. In the barrel, layers of all sub-detectors are composed of concentric cylinders and in the end-cap region are composed of disks or wheels arranged orthogonally to the beam direction. Whole length of the inner detector is 7 m, with its external radius 1.15 m, and fully contained in the CS magnetic field. The PIXEL and SCT are silicon detectors to achieve fast response and high spatial resolution, while TRT is a gaseous detector to cover the large volume. Performance of the components is shown in Table 2.2.

The inner detector measures the trajectory of the charged particle in the magnetic field by detecting the position at each detector layer traversed by the particle. The particle track is reconstructed by connecting the hit points. By the direction of bending of the trajectory, the sign of the particle charge is identified. The transverse momentum (p_T) is calculated by the sagitta ($S[\text{m}]$) and chord ($L[\text{m}]$) of the trajectory on the transverse plane:

$$p_T \sim \frac{0.3BL^2}{8S}, \quad (2.9)$$

where $B[\text{T}]$ is the magnetic field strength. Figure 2.7 shows a trajectory traversing PIXEL and SCT layers, and its sagitta and chord. The momentum resolution of the track neglecting the multiple scattering effect is obtained as:

$$\frac{\sigma(p_T)}{p_T} = \frac{40\sigma(x)p_T}{BL^2} \sqrt{\frac{5}{N+4}} \quad \text{for } N \geq 10, \quad (2.10)$$

where $\sigma(x)$ is the intrinsic position resolution of individual detector layer and N is the number of hits associated to the track. Because a particle with higher p_T generates the smaller sagitta of the trajectory, the resolution of the p_T of the charged particle is worse at higher p_T . The material can

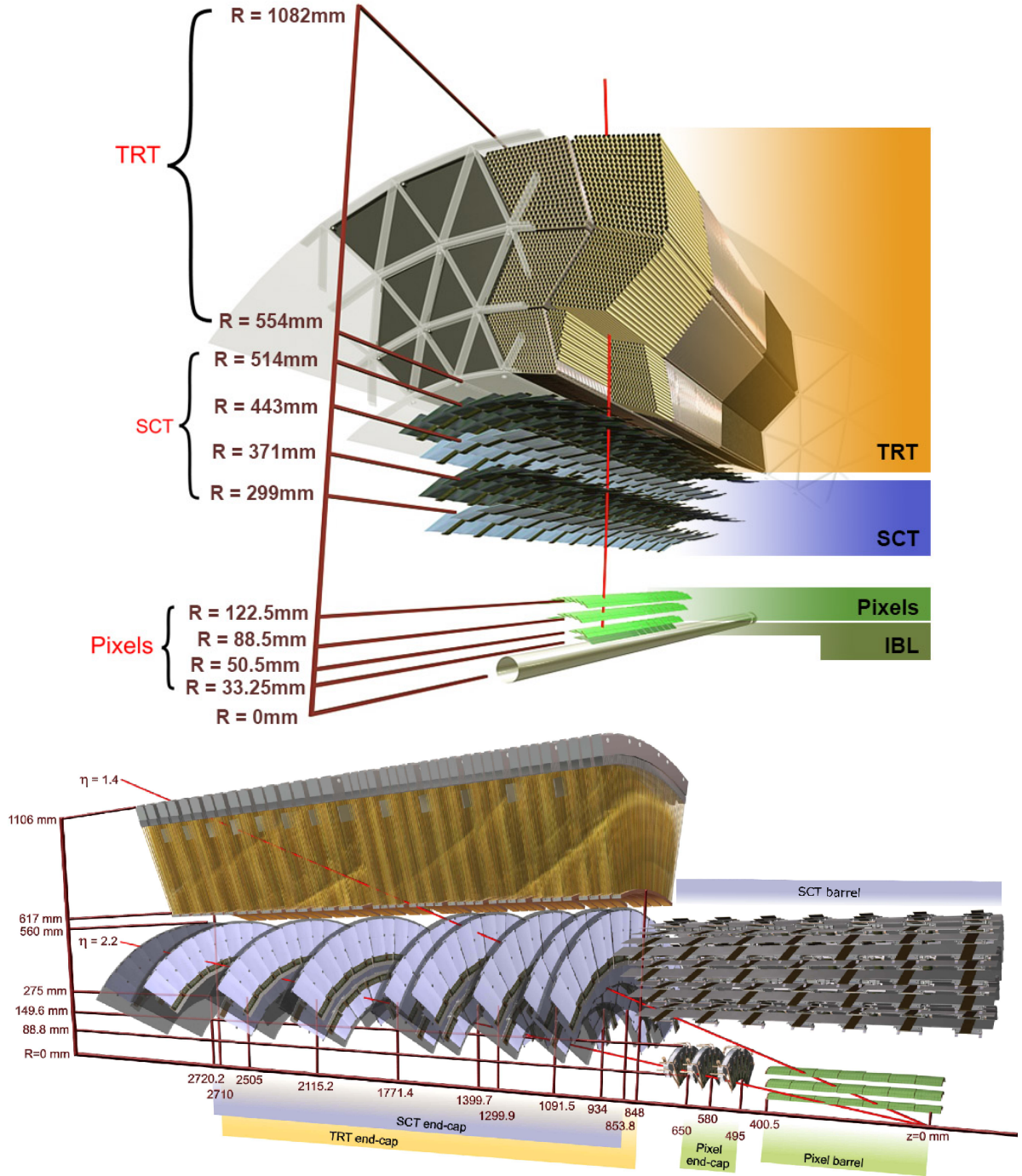


FIGURE 2.6: Illustration of the inner detector in (upper) the barrel and (lower) the end-cap region. [36, 37]

degrade the momentum resolution due to multiple-scattering for lower p_T . The overall thickness of the inner detector is about 0.4 radiation lengths ($0.4X_0$) at $\eta = 0$, and increases up to $1.5X_0$ in the forward region due to the presence of services (e.g. cables for the electronics and the cooling system). Combining the effect of the multiple-scattering in the inner detector, the momentum

TABLE 2.2: Typical resolutions and detector parameters of the inner detector sub-systems [38, 39]. The Insertable B-layer (IBL) is newly installed in 2014.

	Radial extension, R [mm]	Length, z [mm]	# layers	Intrinsic resolution [μm]
beam pipe	$25 < R < 30$	–	–	–
PIXEL				
IBL	$R = 33.25$	$0 < z < 310$	1	$r\text{-}\phi \times z = 8.5 \times 47$
PIXEL barrel	$50.5 < R < 122.5$	$0 < z < 400.5$	3	$r\text{-}\phi \times z = 12 \times 66$
PIXEL end-cap	$88.8 < R < 149.6$	$495 < z < 650$	2×3	$r\text{-}\phi \times r = 12 \times 77$
SCT				
SCT barrel	$299 < R < 514$	$0 < z < 749$	4	$r\text{-}\phi \times z = 16 \times 580$
SCT end-cap	$275 < R < 560$	$839 < z < 2735$	2×9	$r\text{-}\phi \times z = 16 \times 580$
TRT				
TRT barrel	$563 < R < 1082$	$0 < z < 712$	73	$r\text{-}\phi = 170$
TRT end-cap	$644 < R < 1004$	$848 < z < 2710$	160	$r\text{-}\phi = 170$

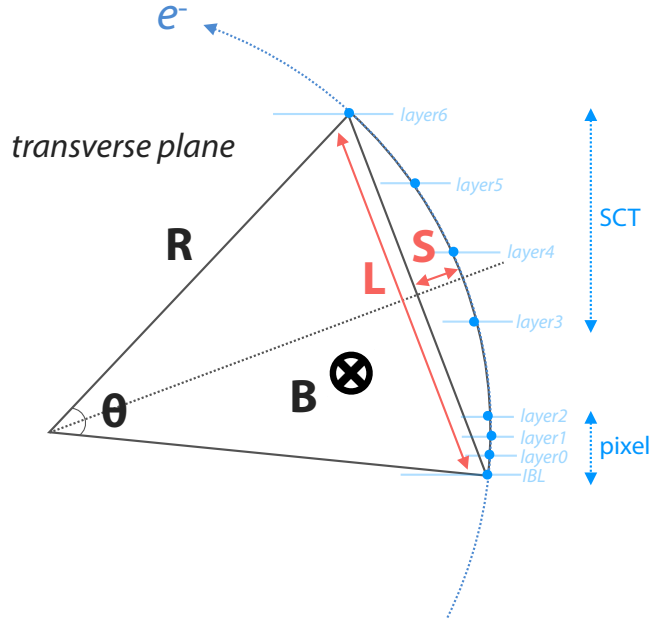


FIGURE 2.7: Illustration of an electron trajectory traversing PIXEL (IBL and layer0,1,2) and SCT (layer3, 4, 5, 6). S and L are the sagitta and chord of the trajectory, and $B[T]$ is the magnetic field.

resolution was measured using cosmic-ray muons in 2011 [40] as:

$$\frac{\sigma(p_T)}{p_T} = (0.016 \pm 0.0001) \oplus [(5.3 \pm 0.2) \times 10^{-5} \times p_T/[\text{GeV}]], \quad (2.11)$$

where the first term denotes the multiple-scattering effect and the second term corresponds to Equation 2.10. The resolution of the impact parameter is also highly affected by the intrinsic resolution of the inner detector. The measured resolutions of d_0 and z_0 are shown in Figure 2.8. The good performance of the tracking and impact parameters is crucial to the performance of such

as the b -tagging as discussed in Chapter 3.

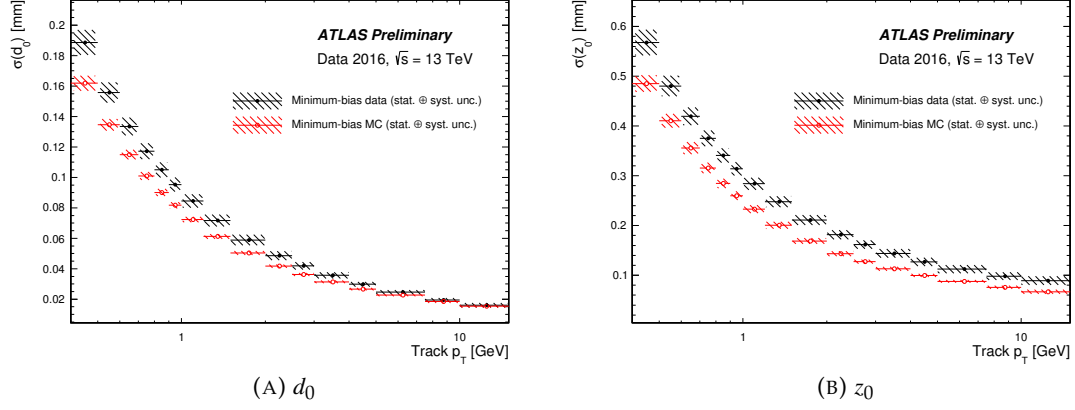


FIGURE 2.8: Intrinsic impact parameter resolutions of tracks, $\sigma(d_0)$ and $\sigma(z_0)$ measured in minimum-bias simulation and in 2016 data [41]. They were extracted in fine bins of p_T and η of the tracks using iterative Gaussian fits, then averaged over η . Systematic uncertainties were taken into account: primary vertex resolution uncertainty, unfolding procedure, non-Gaussian tails of the resolution fit, data-taking period dependence assessed by comparing events collected in 2015 and in 2016 runs.

2.5.1 PIXEL

The PIXEL is the innermost detector of the inner detector (Figure 2.9). Therefore, it is designed to provide the most precise hit positions of charged particles in a very dense track environment close to the interaction point. The inner most layer (IBL) is exposed to 2.5 MGy of Total Ionization Dose (TID) and $5 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ of Non Ionizing Energy Loss (NIEL) over the detector lifetime. Two sensor technologies, planar and 3D sensors, were adopted for 75% and 25% of the IBL sensors in central and high η regions, respectively. The schematics are shown in Figure 2.10. The planar sensor is also adopted in outer layers, which has been developed since the beginning of the PIXEL system design. It is fabricated in n-type Si bulk wafer with surface electrodes by n^+ -implant. The thickness of the sensor is $250 \mu\text{m}$. On the other hand, the 3D sensor is fabricated in p-type Si bulk wafer with pillar electrodes penetrating the bulk by n^+ - and p^+ -implants. It has a lower depletion voltage and faster charge collection than the planar, although it has a higher input capacitance resulting in more noise, and most notably higher costs for the production. The thickness of the sensor is $230 \mu\text{m}$. The pixel pitches of the IBL are $50 \times 250 \mu\text{m}^2$ and other pixels are $50 \times 400 \mu\text{m}^2$. The pixel matrix in each module is 80×328 for the IBL and 144×328 for the other layers. In total, almost 2500 modules were assembled corresponding to 92 million readout channels. The PIXEL is operated at a low temperature ($\sim -10^\circ\text{C}$) to keep good performance and minimize the radiation damage. The PIXEL has IBL and three pixel-layers in the barrel, and three pixel-disks in each end-cap. The inner most pixel-layer in the barrel (layer-0 in Figure 2.7) is called b-layer.

2.5.2 SCT

The SCT consists of four double-sided micro-strip silicon detector layers in the barrel and nine disks in each end-cap region providing at least eight hit points. The design of the barrel and end-cap modules are different as shown in Figure 2.11. The module consists of two $6.38 \times 6.4 \text{ mm}^2$

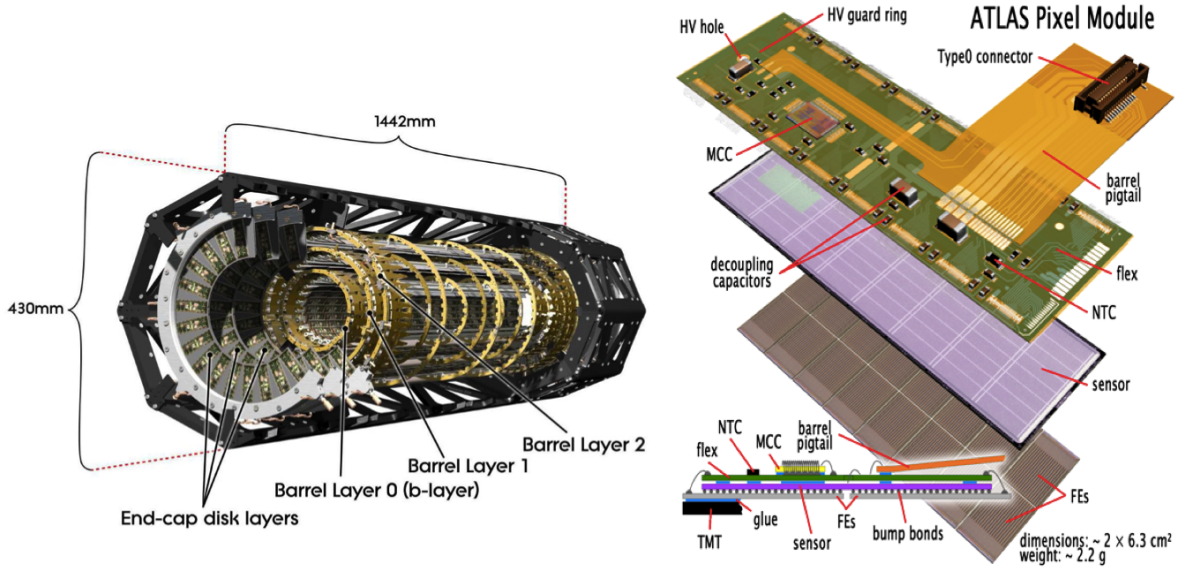


FIGURE 2.9: Schematics of (left) PIXEL and (right) its module [42]

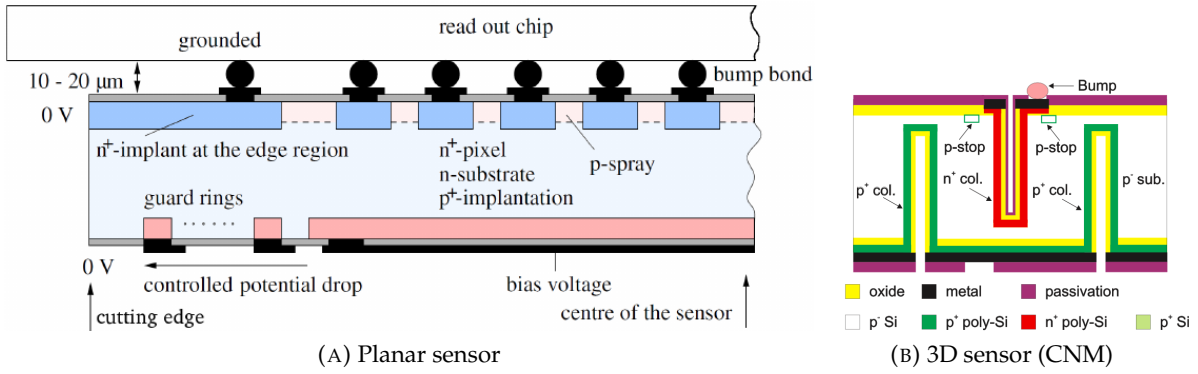


FIGURE 2.10: Schematics of (A) Planar sensor [43] and (B) 3D sensor designed by CNM [44].

p-on-n single-sided sensors each with 768 active readout strips of 12 cm length with $80 \mu\text{m}$ pitches. The thickness of the sensor is $285 \mu\text{m}$. Two sensors are glued back-to-back with a stereo angle of 40 mrad to obtain two-dimensional hit position. The total number of the modules is 4088 for a total of 6.2 million readout channels.

2.5.3 TRT

The TRT is composed of 370,000 cylindrical drift straw tubes. It improves the momentum measurement and provides information of electron separated from other charged particles by transition radiation, as the transition radiation intensity is proportional to the Lorentz boost of the particle, $\gamma = E/m$. In the barrel region (73 layers), there are 52,544 4-mm diameter straw tubes of $60\text{-}\mu\text{m}$ thick carbon fiber film with Kapton reinforcement acting as the cathode. In the center of the tube, a $30\text{-}\mu\text{m}$ diameter gold-plated tungsten wire is strung as the anode. The straws are filled with a gas mixture of 70% Xenon (for good X-ray absorption), 27% CO_2 (for increasing the electron drift velocity and photon-quenching) and 3% O_2 (for enhancing the operation stability). The applied voltage 1.5 kV to the cathode corresponds the maximum drift time of 45 ns. The signal is extracted through both ends of the straw to identify the position of the traversed particle.

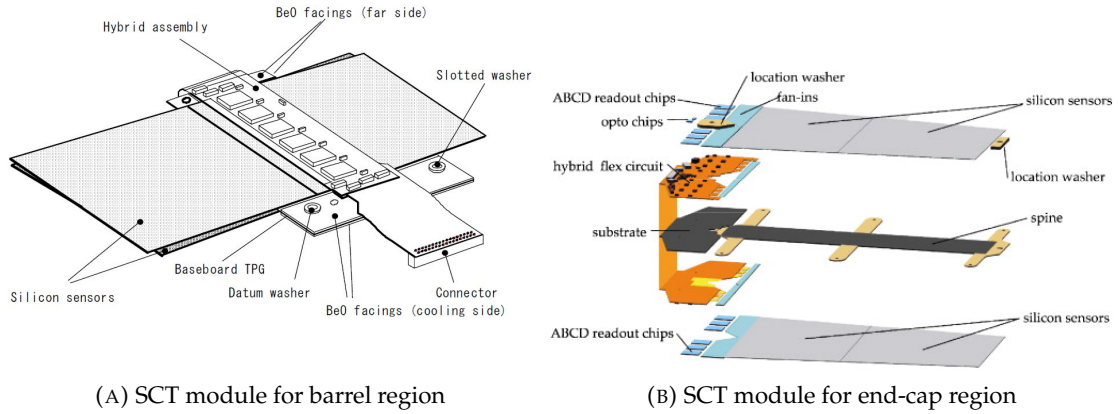


FIGURE 2.11: Schematics of SCT module for (A) barrel and (B) end-cap regions [45]. Each Si sensor is a single-sided micro strip sensor. Glued two sensors with a stereo angle provide two-dimensional hit positions from one module.

The schematics is shown in Figure 2.12. In the end-cap region, 122,880 straw tubes (160 layers) are arranged perpendicularly to the beam axis.

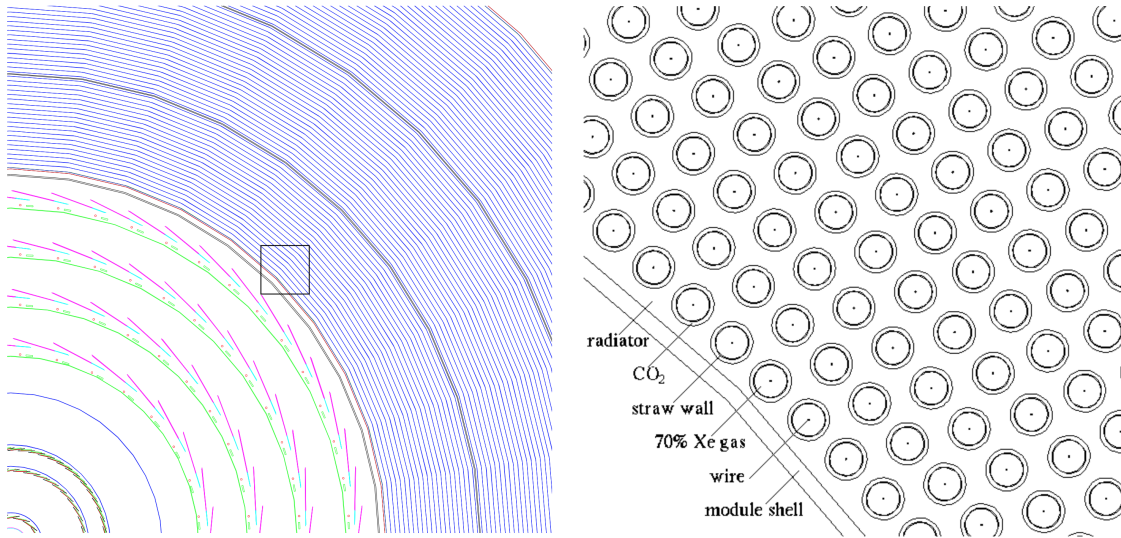


FIGURE 2.12: (Left) Transverse view of a quarter section of the barrel TRT straw layers shown by approximated arcs of circles (Right) detailed view of straws in TRT corresponds to box drawn in the left figure [45].

The spaces between the straws are filled with polymer fibers (barrel) and foils (end-caps) to create transition radiation, which is emitted by the charged particles and depends on γ of the particle as they traverse a material boundary. The emitted X-rays are absorbed by the Xe gas, resulting into additional collected energy. The front end system employs two thresholds: low threshold about 300 eV for minimum ionizing particle tracking and high threshold about 6 or 7 keV for electron identification.

The intrinsic single-point resolution of the TRT is limited compared to the PIXEL and SCT, but the large number of hits per track, typically more than 36 hits (22 hits in the transition region:

$0.8 < |\eta| < 1.0$) for a charged particle with $p_T > 0.5$ GeV and $|\eta| < 2.0$, improve the resolution of track momentum.

2.6 Calorimeter

The calorimeter system (Figure 2.13) surrounds the inner detector and consists of an inner EM and outer HAD calorimeter sub-systems. They stop various particles including neutral particles and measure the energies in the wide range of pseudo-rapidity up to 4.9. A fast electronics system employed to the calorimeter provides the first level (L1) trigger. The ATLAS calorimeter is a sampling type which alternates the absorber plate initiating particle showers and the active layer to perform energy measurements. The calorimeters are composed of different absorber and active layer combinations with different granularity depending on the η range. The $\delta\eta$ - $\delta\phi$ segments and η coverages of the components are shown in Table 2.3.

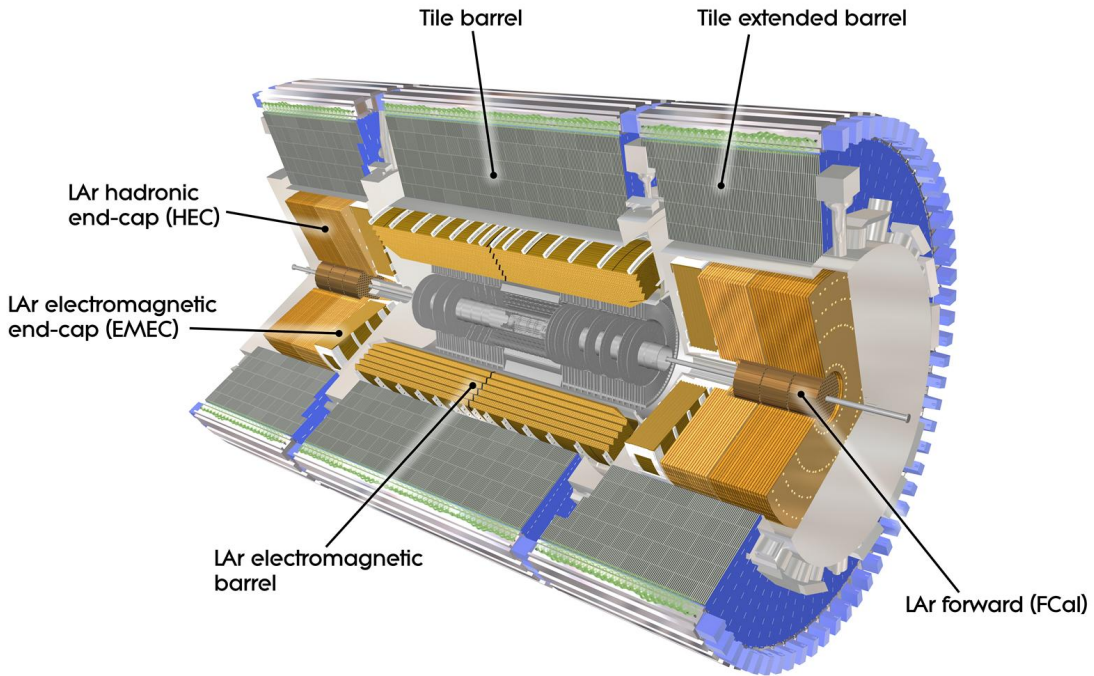


FIGURE 2.13: Illustration of the calorimeter system [46].

Particles passing through calorimeters interact with materials electromagnetically and/or hadronically until its energy becomes lower than the critical energy (E_c) which is the minimum energy to radiate photons or continue nuclear interactions. The particles with $E > E_c$ produce multiple particles and they radiate photons or continue nuclear interactions again. As the result of the chain of interactions, a "shower" is formed in the calorimeter. The shower formed through electro-magnetic (hadronic) interaction is called EM-shower (hadronic-shower). Generated charged particles ionize the active material in the calorimeter. The energy is calculated by measuring the ionization or equivalently the number of particles in the shower, which is proportional to the energy of incident

TABLE 2.3: $\delta\eta$ - $\delta\phi$ segments and η coverages of the calorimeter sub-systems. "PS" denotes the pre-sampler layer.

	$ \eta $ coverage	$\delta\eta \times \delta\phi$	absorber	active material
EM-calorimeter (LAr barrel)				
PS	< 1.52	0.025×0.1	Pb	LAr
layer-1(inner)	< 1.4	$0.025/8 \times 0.025$		
layer-2(middle)	$1.4 \text{ to } 1.475$	0.025×0.025		
	< 1.4	0.025×0.025		
layer-3(outer)	$1.4 \text{ to } 1.475$	0.075×0.025		
	< 1.35	0.05×0.025		
EM-calorimeter (LAr end-cap)				
PS	$1.5 \text{ to } 1.8$	0.025×0.1	Pb	LAr
layer-1(inner)	$1.375 \text{ to } 1.425$	0.05×0.1		
	$1.425 \text{ to } 1.5$	0.025×0.1		
	$1.5 \text{ to } 1.8$	$0.025/8 \times 0.1$		
	$1.8 \text{ to } 2.0$	$0.025/6 \times 0.1$		
	$2.0 \text{ to } 2.4$	$0.025/4 \times 0.1$		
	$2.4 \text{ to } 2.5$	0.025×0.1		
layer-2(middle)	$2.5 \text{ to } 3.2$	0.1×0.1		
	$1.375 \text{ to } 1.425$	0.05×0.025		
	$1.425 \text{ to } 2.5$	0.025×0.1		
layer-3(outer)	$2.5 \text{ to } 3.2$	0.1×0.1		
	$1.5 \text{ to } 2.5$	0.05×0.025		
HAD-calorimeter (Tile barrel)				
layer-1,2	< 1.0	0.1×0.1	Fe	scintillating tile
layer-3	< 1.0	0.2×0.1		
HAD-calorimeter (Tile extended-barrel)				
layer-1,2	$0.8 \text{ to } 1.7$	0.1×0.1	Fe	scintillating tile
layer-3	$0.8 \text{ to } 1.7$	0.2×0.1		
HAD-calorimeter (LAr end-cap)				
HEC-inner	$1.5 \text{ to } 3.2$	0.1×0.1	Cu	LAr
HEC-outer	$1.5 \text{ to } 3.2$	0.2×0.2		
EM/HAD-calorimeter (LAr forward)				
FCal1(inner)	$3.15 \text{ to } 4.83$		Cu	LAr
FCal2(middle)	$3.24 \text{ to } 4.81$		W	LAr
FCal3(outer)	$3.32 \text{ to } 4.75$			

particle. Therefore, the calorimeter energy resolution is described as:

$$\frac{\sigma(E)}{E} = \frac{A}{\sqrt{E/[\text{GeV}]}} \oplus B, \quad (2.12)$$

where A and B denote the constant values depending on the absorber and active materials, and the sampling fraction in the calorimeter. The term of $A/\sqrt{E}$ is called the "stochastic" term describing the shower intrinsic fluctuations. The second constant term, B includes the uncertainties related to response non-uniformities, calibration, effect of dead materials and energy losses before the calorimeter. Required performances are $A \leq 10\%$ and $B \leq 1\%$ for the EM calorimeter, $A \leq 50\%$ and $B \leq 3\%$ for the HAD calorimeter in $|\eta| < 3$. For the HAD calorimeter in $|\eta| > 3$, the resolution of the transverse energy ($\sigma(E_T)/E_T$) is considered with $A \leq 100\%$ and $B \leq 10\%$.

The incoming electron and photon mainly undergo bremsstrahlung and pair-creation, respectively, and generate the EM-shower in the calorimeter. The radiation length X_0 [g/cm²] and Moliere radius R_M [g/cm²] are parameters to describe the characteristic length and lateral size of EM-shower. The radiation length is defined using the energy loss equation by the bremsstrahlung as:

$$-\frac{dE}{dx} = \frac{E}{X_0} \quad \text{for } X_0 \text{ [g/cm}^2\text{]} = \frac{A}{4\alpha N_A Z^2 r_e^2 \ln(183/Z^{1/3})}, \quad (2.13)$$

where

- A : mass number of the material
- Z : atomic number of the material
- N_A : Avogadro's number
- r_e : classical electron radius
- α : fine structure constant.

Using $t_{\max}$ defined as the longitudinal position where maximum energy is deposited in the shower, the shower longitudinal length of the cylinder containing 95% of the shower energy is approximately described as:

$$L_{95\%} [X_0] = t_{\max} + 0.08Z + 9.6, \quad (2.14)$$

where $t_{\max} = \ln(E/E_c) - 1.0$ for electron-induced shower and $t_{\max} = \ln(E/E_c) - 0.5$ for photon-induced shower. The Moliere radius is defined as a radius of the cylinder containing 90% of the shower energy, which is approximately given by:

$$R_M \text{ [g/cm}^2\text{]} = \frac{21}{E_c/[\text{MeV}]} X_0. \quad (2.15)$$

The critical energy where dE/dx by ionization and bremsstrahlung are equal, depends on the absorber material, 0.56 cm for Pb.

The hadronic-shower has complicate structure because of the strong interactions. In the hadronic-shower description, the interaction length (λ) is used:

$$\lambda \text{ [g/cm}^2\text{]} = \frac{A}{N_A \sigma_{\text{inel},pA}} \sim \frac{A^{1/3}}{N_A \sigma_{\text{inel},pp}}, \quad (2.16)$$

where $\sigma_{\text{inel},pA}$ is the cross section of the inelastic collision of a proton and nucleon, and approximately equal to $\sigma_{\text{inel},pp} A^{2/3}$. Approximately, the ratio of the interaction length to the radiation length is proportional to $A^{4/3}$ ($\lambda/X_0 \sim 10$ in case of Fe). Using the interaction length, $t_{\max}$ and $L_{95\%}$ are approximately given as:

$$t_{\max}/\lambda = 0.2 \ln(E/[\text{GeV}]) + 0.7 \quad (2.17)$$

$$L_{95\%}/\lambda = t_{\max} + 2.5 (E/[\text{GeV}])^{0.3} \quad (2.18)$$

2.6.1 EM Calorimeter

The EM calorimeter is composed of LAr active medium equipped with Kapton electrodes, and lead absorber plates with accordion shape which provides complete ϕ symmetry without azimuth cracks. The EM calorimeter is separated into the barrel region ($|\eta| < 1.475$) and the end-cap region ($1.375 < |\eta| < 3.2$). The barrel calorimeter consists of two identical cylinders separated for positive and negative η regions with a 6-mm gap between them at $z = 0$. Each end-cap calorimeter is subdivided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$ and an inner wheel covering the region $2.5 < |\eta| < 3.2$. Figure 2.14 shows the detail geometry of the EM calorimeter in the barrel region. The wave of the accordion electrode structure in the barrel (end-cap) region grows along ϕ -direction (z -direction) and the LAr gap is constant (larger with the radius). The thickness of EM calorimeter is designed to have a radiation length of more than $24X_0$ ($26X_0$) in the barrel (end-cap) region to stop the electrons and photons with the energies up to $O(\text{TeV})$. The calorimeter has three longitudinal segments for a precision measurement. The finer granularity of the layer-1 in the central is required to be sensitive to the position of the EM-shower to be matched to the track information available in the inner detector. The layer-2 absorbs the most of the energy of the shower with its large radiation length, and the layer-3 covering up to $|\eta| < 2.5$ collects the remaining tail of the shower. The pre-sampler (PS) is also installed in front of the layer-1, which is composed of a single active LAr layer and estimates the energy loss in upstream of the calorimeter. The total number of channels for the entire EM calorimeter is about 190,000.

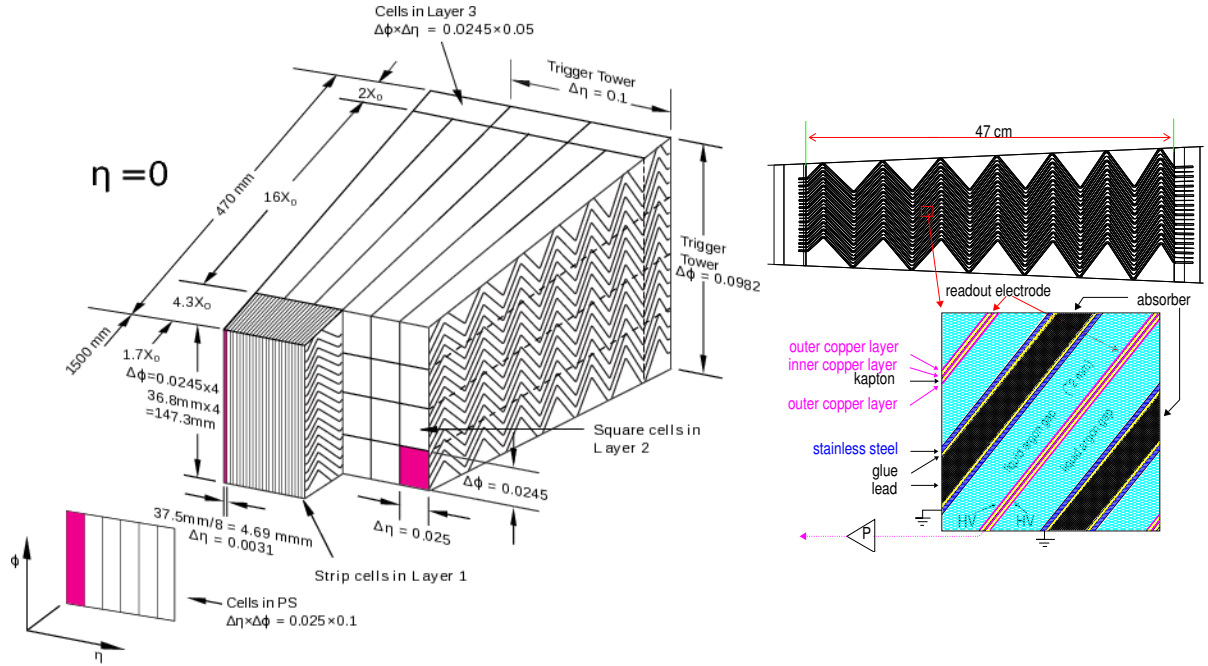


FIGURE 2.14: Illustration of the EM calorimeter system (barrel region) [45, 47].

2.6.2 HAD Calorimeter

The HAD calorimeter is separated in two types depending on the materials and η coverage, the tile calorimeter and LAr end-cap calorimeter. The tile calorimeter shown in Figure 2.15 is composed of steel absorber plates and polystyrene scintillating tiles, subdivided by the η coverage into barrel ($|\eta| < 1.0$) and extended-barrel ($0.8 < |\eta| < 1.7$) regions. Each tile calorimeter has three

longitudinal segments with interaction lengths 1.5λ , 4.1λ , and 1.8λ from the inner segment. The total number of readout channels is approximately 10,000.

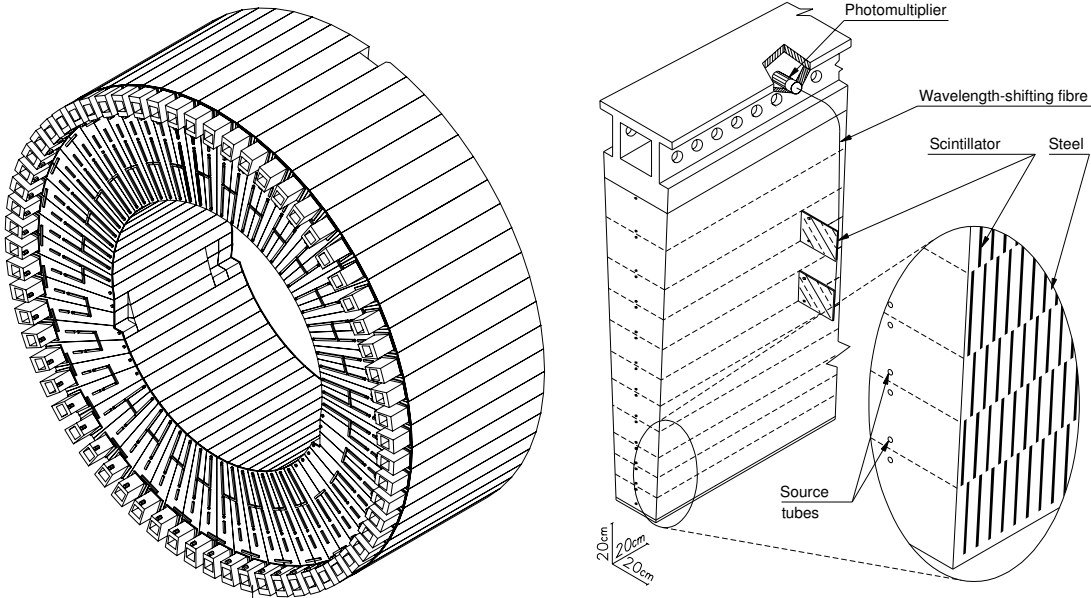


FIGURE 2.15: Illustration of the HAD calorimeter system [45]. The module is constructed from steel plates and scintillator tiles. The scintillator tiles are placed radially for easier coupling of wavelength shifting fibers to the scintillating fiber. Wavelength shifting fibers are coupled to two sides of the scintillating tiles connected separately to two photomultiplier tubes placed back at the module.

The LAr Hadronic End-Cap calorimeter (HEC) is composed of copper plate absorbers and LAr active material, similar to the EM calorimeter because LAr is radiation-hard. The HEC consists of two wheels, the inner-wheel and outer-wheel. Each wheel contains two longitudinal sections and 32 identical wedge-shaped modules. The module in the inner-wheels (outer-wheel) has 24 (16) plates of 25-mm (50-mm) thick copper and a 12.5-mm (25-mm) thick front plate with active LAr layer of 8.5 mm. Seven stainless-steel tie-rods penetrate the modules to keep the rigidity of the module. The number of readout channels is 5,632.

2.6.3 Forward Calorimeter

The LAr Forward Calorimeter (FCal) is composed of three 45-cm thick modules named FCal1, 2, and 3. FCal1 consists of copper absorber for the EM shower optimizing the energy resolution and heat removal. FCal2 and 3 are made of tungsten absorber to shorten the hadronic shower size. All of three modules use the LAr active material. The FCal covers the very high η range, $3.1 < |\eta| < 4.9$ with an interaction length of $\sim 10\lambda$.

2.7 Muon Spectrometer

Muon spectrometer is a set of detectors for detecting charged particles which penetrate through the calorimeter, most likely muons. The muon spectrometer is designed to identify the muon and measure its momentum. Therefore, the muon spectrometer has the multiple layers and each layer measures the position of the traversing muon. Same as the inner detector, the track is reconstructed by hit positions in the muon spectrometer, and the muon momentum is calculated using the sagitta of the trajectory. The muon spectrometer also provides the trigger signal to take the data including high p_T muon candidate. The muon spectrometer consists of four gaseous sub-detectors covering the huge area outside of the calorimeter: the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC), and the Thin Gap Chambers (TGC). The MDT and CSC are wire drift detectors for precision measurements of the muon momentum in $|\eta| < 2.7$, while the RPC and TGC are installed for fast response dedicated for the trigger in $|\eta| < 2.4$. As the precision tracking system measures the muon trajectory within $O(\mu s)$, the trigger chambers provide fast response identifying the beam crossing (< 25 ns). Precision chambers have a transverse momentum resolution of approximately 10% for 1 TeV muons corresponding to a position resolution of $50 \mu m$ for sagitta of $500 \mu m$.

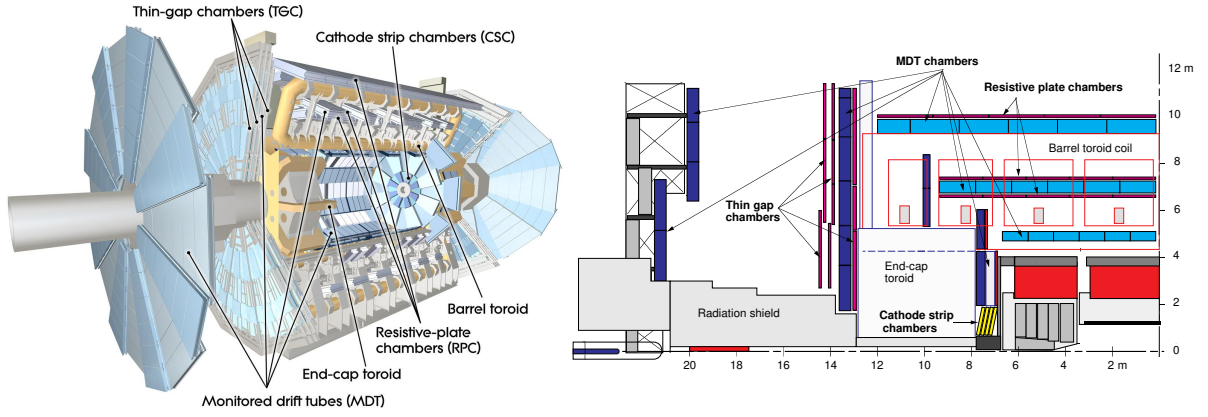


FIGURE 2.16: Illustration of the muon spectrometer. [48]

2.7.1 Precision Chambers

MDTs are proportional chambers based on the drift tube technology. The MDT uses an aluminum tube of 30-mm diameter with a central W-Re wire of $50\text{-}\mu m$ diameter. The tubes are filled with a gas mixture of 91% Ar, 4% N_2 , and 5% CH_4 . The maximum drift time is 700 ns with a small Lorentz angle. The MDT chambers are arranged in concentric cylinder shapes in the central and the perpendicular wheels in the end-cap. The barrel cylindrical MDT consists of three stations to measure the trajectory sagitta by the toroid magnet field. To achieve the precise position resolution, MDT chamber stations are constructed from four tube-layers for the inner layer, and three tube-layers each for the middle and outer layer. The schematics for the station in the inner layer is shown in Figure 2.17. Their tubes are placed transverse to the beam axis. Each cylindrical MDT station provides a position resolution of $40 \mu m$, and the combined resolution for three station is $30 \mu m$. The end-cap MDT wheels are installed perpendicular to the beam axis in similar stations to the barrel as shown in Figure 2.17. Due to their reliability, mechanical robustness and simpler operation, MDT chambers are employed to cover the large η range ($|\eta| < 2.7$ for the middle and outer layers, and $|\eta| < 2$ for the inner layer).

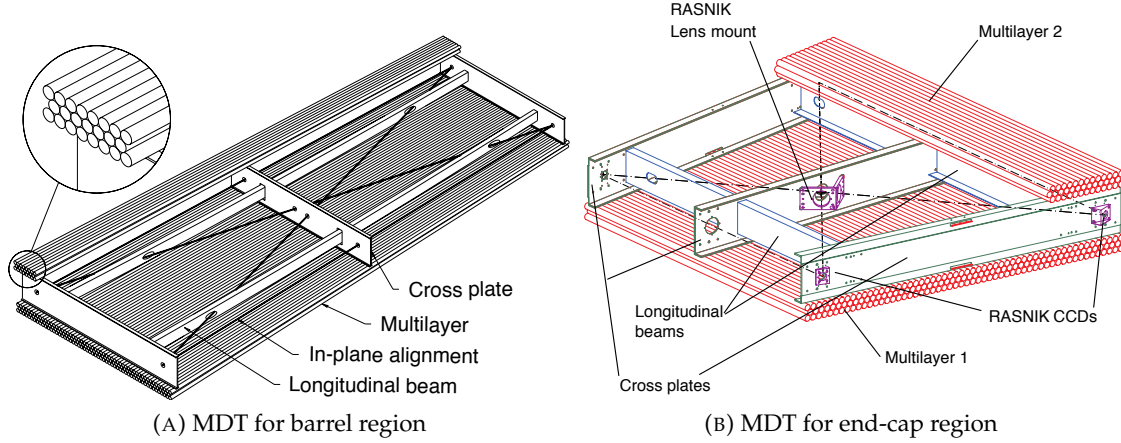


FIGURE 2.17: Schematics of MDT station for (A) barrel (B) end-cap region [35].

The CSC is installed in $2 < |\eta| < 2.7$ where the hit rate reaches $\sim 1 \text{ kHz/cm}^2$ exceeding the 150 Hz/cm^2 limit sustainable by the MDT. The CSC is the multi-wire proportional chamber (MWPC) with W-Re anode wires ($30 \mu\text{m}$ diameter and 2.5 mm pitch). The signal is read out by two cathode-strip layers made of copper-clad laminated by poly urethane foam perpendicular to each other. The finer readout is achieved at readout pitch of $5.3\text{-}5.6 \text{ mm}$ in the bending direction, as shown in Figure 2.18. The gap between the anode wire and cathode strip is designed to be equal to the pitch of the anode wires. The chamber is filled with a gas mixture of 30% Ar, 50% CO_2 , and 20% CF_4 . The CSCs are wheel-shaped chambers arranged in 4 layers in each station. The precise position determination is achieved by measuring the charge induced on the segmented cathode by the avalanche on the anode wire, and by interpolating the charges between neighbor strips. The r.m.s. resolutions of $\leq 60 \mu\text{m}$ have been measured in several prototypes.

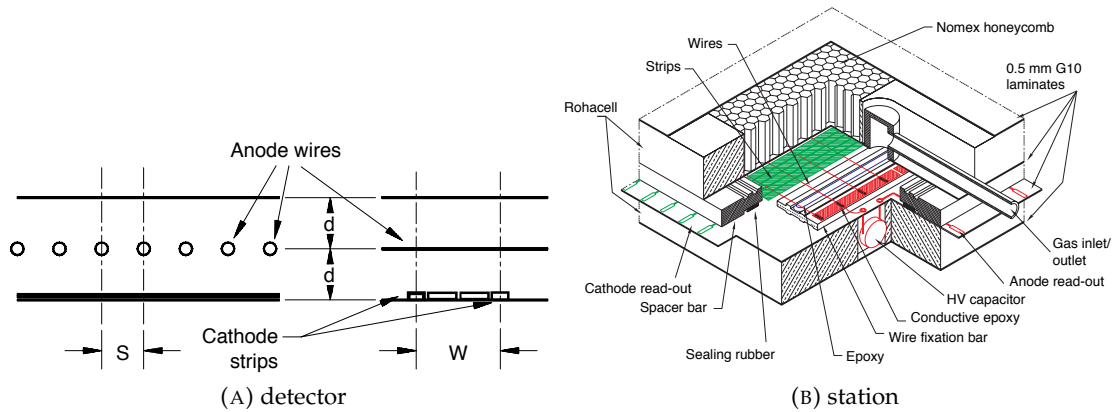


FIGURE 2.18: Schematics of CSC (A) detector and (B) station [35].

2.7.2 Trigger Chambers

The muon trigger system employs the RPC and TGC for the barrel and the end-cap regions, respectively.

The RPC is a gaseous detector providing a typical space-time resolution of $1\text{ cm} \times 1\text{ ns}$. The basic RPC unit is a narrow gas gap (2 mm) formed by two parallel resistive bakelite plates, separated by insulating spacers. The primary ionization electrons are amplified in avalanche mode gases with a mixture of 97% $\text{C}_2\text{H}_2\text{F}_4$ and 3% C_4H_{10} operated at a voltage of 9.8 kV. The signal is read out via capacitive coupling by orthogonal series of metal strips on both sides of the chamber (30.0-39.5 mm pitches): the one is parallel to the MDT wires providing the bending view as the trigger detector, the other is orthogonal to the MDT wires required in the offline pattern recognition. The RPC has three stations each of which is composed of two detector layers corresponding to four orthogonal sets of readout strips.

The TGC (Figure 2.19) is designed similarly to the MWPC, but the anode-cathode spacing (1.4 mm) is thinner than the anode wire pitch (1.8 mm). The anode wires (50 μm diameter) are arranged parallel to the MDT wires providing the trigger information with cathode strips arranged orthogonal to the anode wires. The orthogonality provides the position measurement of the traversing muon. The TGC is filled with a gas mixture of 55% CO_2 and 45% of $n\text{-C}_5\text{H}_{12}$. The cathode plane consists of 1.6-mm thick G-10 plates on which the graphite cathodes are deposited. The TGC modules are constructed from two or three series of TGCs named "doublet" or "triplet". Two of the back side of the cathode plates per doublet or triplet modules are patterned with copper strips (1.46-4.91 cm width). The TGC modules are located at three positions in the end-cap: one triplet module layer before the second MDT station, two doublet module layers behind the second MDT station. With the high granularity of the TGC, the anode and cathode signals are transferred

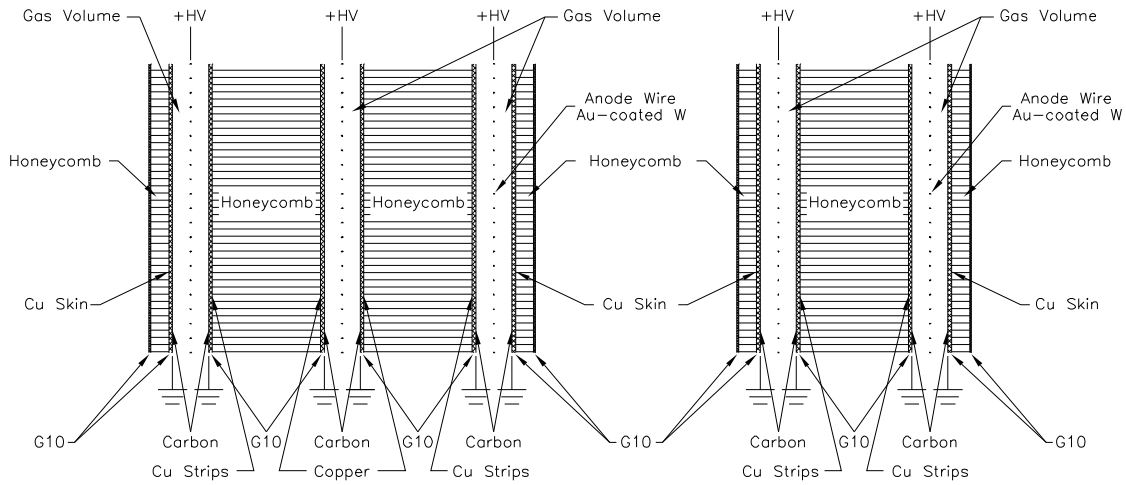


FIGURE 2.19: Cross-section of TGC triplet and doublet modules [35].

by grouping 4-20 wires to realize the desired granularity as a function of η corresponding to the anode readout pitch 7.2-36.0 mm. The signal is digitized at more than 99% efficiency with 25 ns gate window.

2.8 Trigger and DAQ System

The ATLAS detector does not record every collision produced by the LHC, since the expected amount of data would reach 100 PB/s.

The ATLAS trigger and data acquisition (TDAQ) system selects and records the events with interesting physics characteristics. The multi-level TDAQ system is shown schematically in Figure 2.20, implemented to handle events produced in high interaction-rate environment. The trigger system consists of a hardware level 1 (L1) trigger and a software-based high-level trigger (HLT). The event rate is reduced from the bunch-crossing rate of 40 MHz (event rate ~ 1 GHz) to 100 kHz at L1 and to an average 1 kHz recording rate at the HLT. At L1, fast custom-made electronics finds regions of interest (RoI) within $2.5 \mu\text{s}$ using the following systems:

- L1Calo: based on the calorimeter tower granularity in the RoI. Local E_T maxima are searched in 2×2 towers for triggering the electrons, photons, and taus, and 4×4 to 8×8 towers summed over EM and HAD calorimeters for the jets.
- L1Muon: identifying the muon by the muon trigger chambers RPCs (TGCs) in the barrel (end-caps).

At the HLT, fast algorithms access data from an RoI, and then full-event algorithms similar to the offline are performed within a processing time of 0.2 s on average.

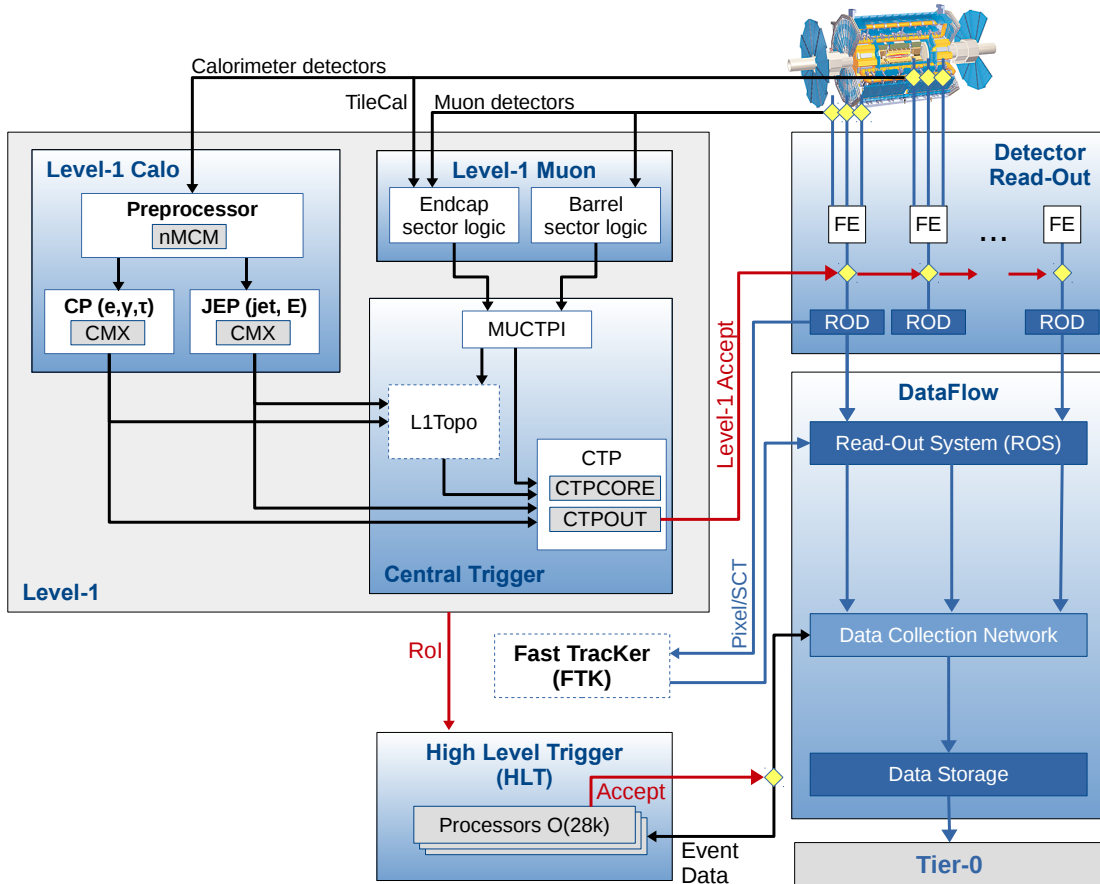


FIGURE 2.20: The ATLAS TDAQ system in Run 2 with an emphasis on the relevant triggering components [49].

2.9 Data Recording and Quality

The data passing the HLT trigger requirement is processed and recorded through the readout-drivers (RODs) which are VME modules devoted to data processing, configuration and control. One ROD publishes one series of bit-collection. The event-by-event data is to be recorded to one format gathering all ROD data from the whole ATLAS detectors. The event data is built by the aggregation of the readout-buffers (ROBs), where each ROB fragment saves corresponding one ROD data bit-collection.

The gathered data is grouped in the luminosity-block typically corresponding to the one to a few minute duration depending on the operational condition. The luminosity-blocks in one whole filling of proton beams are grouped as a run.

After recording one run data, each luminosity-block is qualified using the full offline reconstruction information. The qualifications are performed based on sub-detector information and combined object based information. If the intolerable issues are found in the luminosity-block, the luminosity-block is not used in the physics analysis by putting the defect tag (intolerable-defect). If the issue is tolerable but should be notified, the luminosity-block is also marked as the tolerable defect tag.

Among the ATLAS sub-detectors, the SCT has been operated in highly stable conditions. As an example, the SCT data quality procedure is described in this section. Checking the SCT data quality is performed using byte-stream errors, efficiencies, noise, and number of tracks. The SCT is the strip detector having almost 6 million strips. Those strips are grouped for management of the SCT condition. 128 strips are processed in one front-end chip, and the data from six chips equipped in one side of the module is called link. One module has two links, and up to 48 modules are managed by one ROD. The SCT has in total 128 RODs for the 4088 modules.

For checking the data quality in a manageable procedure, we want to remove events with the known common problems. Because the SCT is one of the most stable sub-detector systems and it is significantly rare for all SCT modules to fall into the problematic condition at the same time, the events with most of SCT modules disabled or judged problematic can be treated as the events having the problem in the whole ATLAS detector. Therefore, "SCTFlag" is put on the event-by-event if there are ROBFragment errors from > 1000 ($> 12\%$) links and/or Level1ID errors from > 500 ($> 6\%$) links in the SCT. The ROBFragment error indicates a broken status in the ROB header due to such as no RODs found in the event. The Level1ID error shows the synchronization problem between the level1 ID counters in the link and in the central Timing, Trigger and Control (TTC) system. The SCTFlagged events are removed from the data quality checking in other detectors as well as from the physics analysis described in Chapter 3.

In the SCT data quality, the number of links publishing byte-stream errors averaged over the events in the luminosity-block is given for checking the critical issues. If more than two disabled-layers are found due to the byte-stream errors in same η - ϕ spaces in the luminosity-block, track quality is reduced significantly from the nominal condition in that region. If such area is more than 5% of the inner-detector coverage in η - ϕ space, intolerable defect tag is put on the luminosity-block. Even for coverage loss less than 5%, tolerable defect tag is put with following procedures. To put tolerable defects efficiently, the byte-stream errors are categorized by the level where the error is recognized: ROD-level and link-level. The link-level errors can happen in an event-by-event basis or a few events. When there are > 80 links with link-level errors, the tolerable defect tag is put on the luminosity-block. When the ROD-level error is published, the same error bit is given to all links in the ROD, and ROD recovery procedure takes place lasting for about a few seconds. If the ROD error happens in the end of the luminosity-block and recovered in the beginning of next luminosity-block, the number of links with ROD-level error in each of the two luminosity-blocks

may become small by averaging. To put the tolerable defect tag in such luminosity-blocks, the threshold of the ROD-level error links is set at 1 for putting the tolerable defect tag.

Finally, the overall SCT conditions through the run are checked. If each of the following criteria is satisfied, the tolerable defect tag is set on the run:

- number of disabled links is more than 80
- number of links with any byte-stream errors is more than 80
- number of links with the noise occupancy $> 0.15\%$ is more than 80
- number of links with the efficiency $< 0.95\%$ is more than 80
- SCT total hit efficiency $< 98\%$ or SCT total hit efficiency for first BCID events $< 99\%$ in any of barrel, end-cap A, and end-cap C sides
- number of tracks is less than 200 in any of barrel, end-cap A, and end-cap C sides

We also put the defects depending on technical conditions of SCT in particular luminosity-blocks:

- (tolerable) high-voltage is not set to the nominal value (150 V during operation or 50 V at standby)
- (tolerable) SCT charge threshold is not set at 1 fC
- (tolerable) during the SCT timing scan
- (tolerable) SCT system are excluded from the cooling loop
- (intolerable) during the SCT HV ramp up to 150 V. The data taking is required to start after the high-voltage is set for the PIXEL but not required for the SCT in order not to lose events.
- (intolerable) during the SCT re-configuration process
- (intolerable) during the ATLAS excluding the SCT system

Figure 2.21 summarizes the data quality efficiency in each sub-component of the ATLAS. In the 2016 run, the SCT operation was very stable even under the high instantaneous luminosity operations.

2.10 Luminosity Acquired in Run 2

The integrated (instantaneous) luminosity collected by the ATLAS detector in Run 2 (2015 and 2016) is plotted in Figure 2.22 (2.23). The ATLAS data-taking efficiency was above 90% in both years. In total $4.2 + 38.5 \text{ fb}^{-1}$ was delivered by the LHC machine and $3.9 + 35.6 \text{ fb}^{-1}$ was collected by the ATLAS detector. The maximum instantaneous luminosity exceeded $13 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ in Run 2.

The mean number of interactions per crossing ($\langle\mu\rangle$ value) is shown in Figure 2.24. The $\langle\mu\rangle$ value is the mean of the Poisson distribution of the number of interactions per crossing calculated from the instantaneous per-bunch luminosity as:

$$\mu = \mathcal{L}_{\text{bunch}} \times \sigma_{\text{inel}} / f_r, \quad (2.19)$$

where $\mathcal{L}_{\text{bunch}}$ is the instantaneous luminosity per bunch, σ_{inel} is the inelastic pp cross section of 80 mb for 13 TeV collisions, and f_r is the LHC revolution frequency. The luminosity shown is preliminary 13 TeV value derived using calibration released in February 2017, based on van-der-Meer beam-profile scans performed in 2016.

ATLAS pp 25ns run: April-October 2016											
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets		Trigger
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid	L1
98.9	99.9	99.7	99.3	98.9	99.8	99.8	99.9	99.9	99.1	97.2	98.3
Good for physics: 93-95% (33.3-33.9 fb⁻¹)											
Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collisions with 25ns bunch spacing at $\sqrt{s}=13$ TeV between April-October 2016, corresponding to an integrated luminosity of 35.9 fb ⁻¹ . The toroid magnet was off for some runs, leading to a loss of 0.7 fb ⁻¹ . Analyses that don't require the toroid magnet can use that data.											

FIGURE 2.21: Luminosity weighted relative fraction of good quality data delivery by the various components of the ATLAS detector/trigger subsystems during LHC fills with stable beams in pp collisions at 13 TeV, and after switching the tracking detectors on [50]. Runs were taken between 28th April and 26th October 2016, corresponding to a recorded integrated luminosity of 35.9 fb⁻¹.

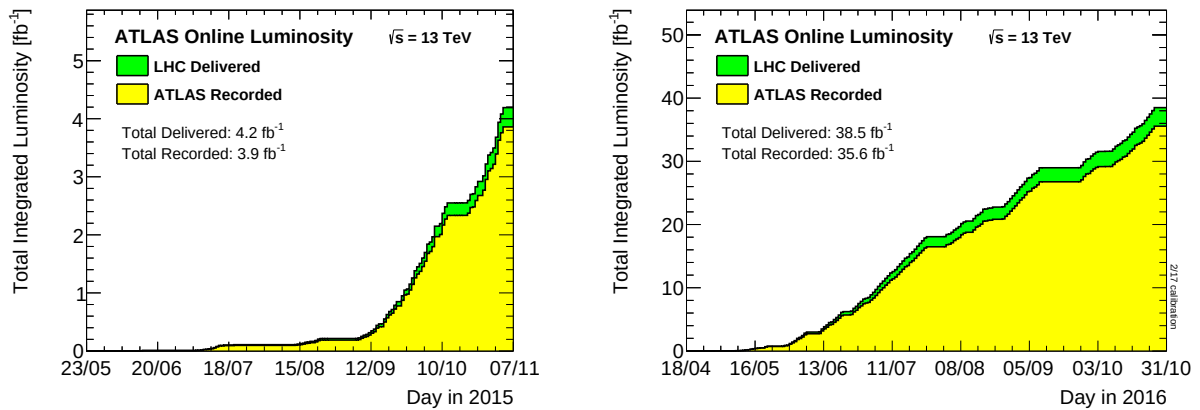


FIGURE 2.22: The integrated luminosity in 2015 (left) and 2016 (right). The luminosity delivered by the LHC machine is shown in green with the luminosity collected by the ATLAS detector reported in yellow [51].

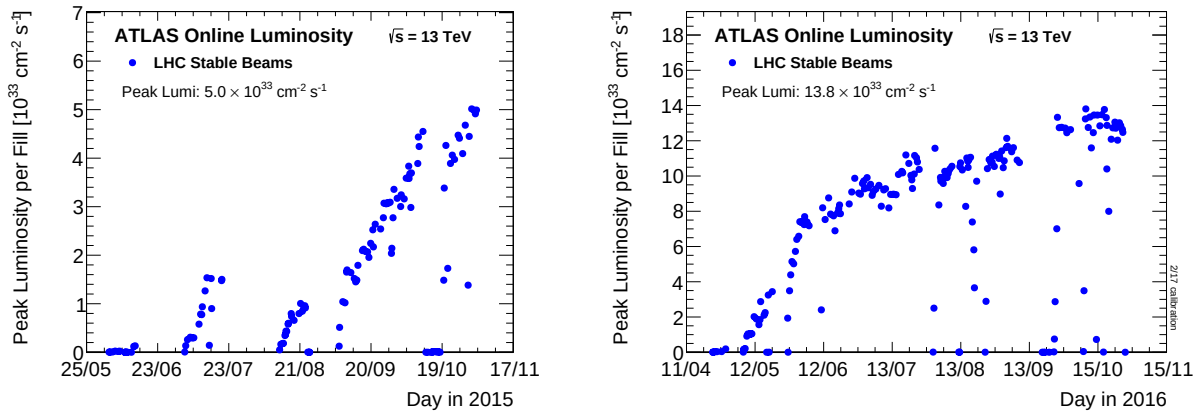


FIGURE 2.23: The instantaneous luminosity in 2015 (left) and 2016 (right) [51].

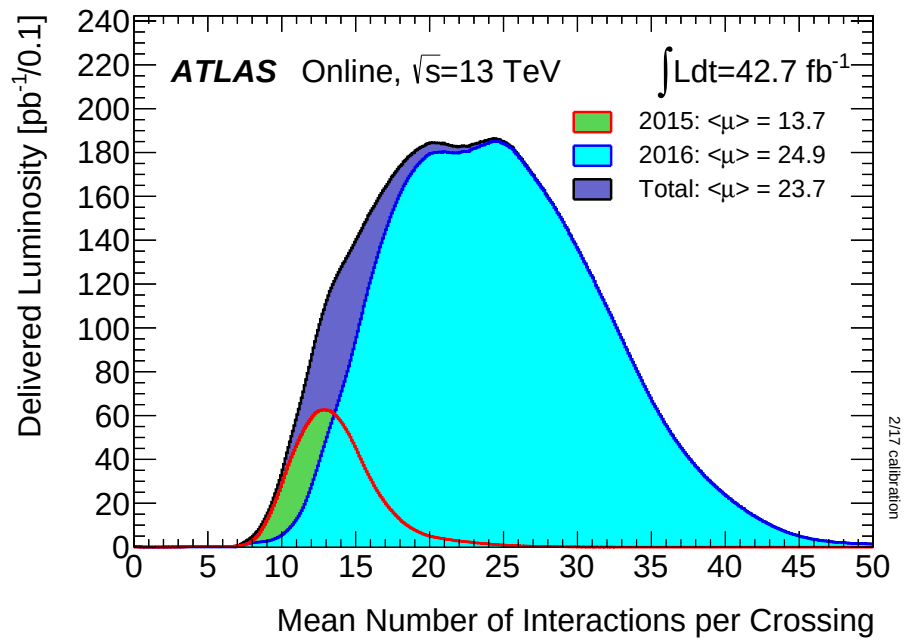


FIGURE 2.24: Luminosity-weighted distribution of the mean number of interactions per crossing ($\langle\mu\rangle$) at 13 TeV pp collision data in 2015 and 2016 runs [51].

Chapter 3

Object Reconstruction and Event Selection in ATLAS

3.1 PreSelection

3.1.1 Primary Vertex

At the LHC, multiple interactions occur per beam crossing (typically 20 – 40) with high intensities. The interaction points referred to as primary vertices (PVs) were reconstructed by associating at least two tracks with $p_T > 400$ MeV. They were used to identify the vertex of the hard scattering process, a candidate of interesting physics objects. Among the reconstructed PV candidates, the PV with the highest sum of the squared transverse momenta of the accompanying tracks was regarded to be associated with hard scattering. The rest of PVs were considered as pileup interactions (pile-up PVs). Objects associated to the hard scattering (pile-up) PV were referred as hard scattering (pile-up) objects.

If the vertices were reconstructed apart from the beam collision points, they were defined as secondary vertices. They were used to identify long lived heavy-flavor hadron decays, in order to identify jets originating from the b -quark and c -quark as described later.

3.1.2 Event Cleaning

The events arising from noise bursts and/or data corruption in the calorimeter were vetoed. The noise burst is the coherent noise from a lot of cells, which is localized and has very short time less than $1 \mu s$ depending on the instantaneous luminosity.

Events with the "SCTFlag" discussed in Chapter 2 are also removed from the analysis, which indicates problems in whole ATLAS detector.

3.2 Trigger

This analysis for $t\bar{t}H(H \rightarrow b\bar{b})$ used data acquired by single-lepton trigger chains, which allowed the lowest transverse momentum leptons without pre-scaling in the trigger.

As for the L1 triggers, events were primarily triggered by L1EM20VH and L1MU15 in the 2015 runs, and L1EM22VH and L1MU20 in the 2016 runs. Here, L1EM20VH and L1EM22VHI required events to have at least one cluster reconstructed in the EM calorimeter with $E_T > 20$ GeV and $E_T > 22$ GeV, respectively. The threshold varied depending on η to account for energy losses (denoted as "V") and required the ratio of E_T in the hadronic to that in the EM calorimeters to be less than a certain value (denoted as "H"). L1EM22VHI required that the isolation in the EM cluster (denoted as "I"). L1MU15 and L1MU20 were required to have at least one muon reconstructed in the muon spectrometer with $p_T > 15$ GeV and $p_T > 20$ GeV, respectively.

Afterwards, HLT chains were applied to keep only interesting events for physics analysis. The triggers are shown in Table 3.1. Multiple triggers were adopted to maximize the efficiency in a wide range of lepton p_T values, where the trigger with a higher p_T threshold had looser isolation requirement.

TABLE 3.1: Triggers used for analysis. See electron and muon sections for definition of ID criteria and track-to-vertex-association (TTVA) cuts, and lepton isolation section for definitions of isolation variables. In the 2015 runs, HLTs with the lowest p_T thresholds required corresponding L1 triggers explicitly.

year	lepton	ID	p_T [GeV]	isolation cut	TTVA	L1 requirements
2015	e	mediumLH	24	–	✓	L1EM20VH
	e	mediumLH	60	–	✓	–
	e	looseLH	120	–	✓	–
	μ	medium	20	$p_T^{\text{cone20}}/p_T < 0.12$	✓	L1MU15
	μ	medium	50	–	✓	–
2016	e	tightLH	26	$E_T^{\text{varcone20}}/E_T < 0.1$	–	–
	e	mediumLH	60	–	–	–
	e	looseLH	140	–	–	–
	μ	medium	26	$p_T^{\text{varcone30}}/p_T < 0.07$	✓	–
	μ	medium	50	–	✓	–

The L1 triggers and HLTs were different between the 2015 and 2016 runs because more pile-up events were observed in 2016 than in 2015. Therefore, tighter trigger requirements were imposed in 2016 to maintain the overall data acquisition rate. The lepton p_T threshold increased to 26 GeV in 2016 in both electron and muon HLT triggers. The trigger efficiencies for electrons and muons in the 2016 runs in each lepton channel are shown in Figures 3.1 and 3.2.

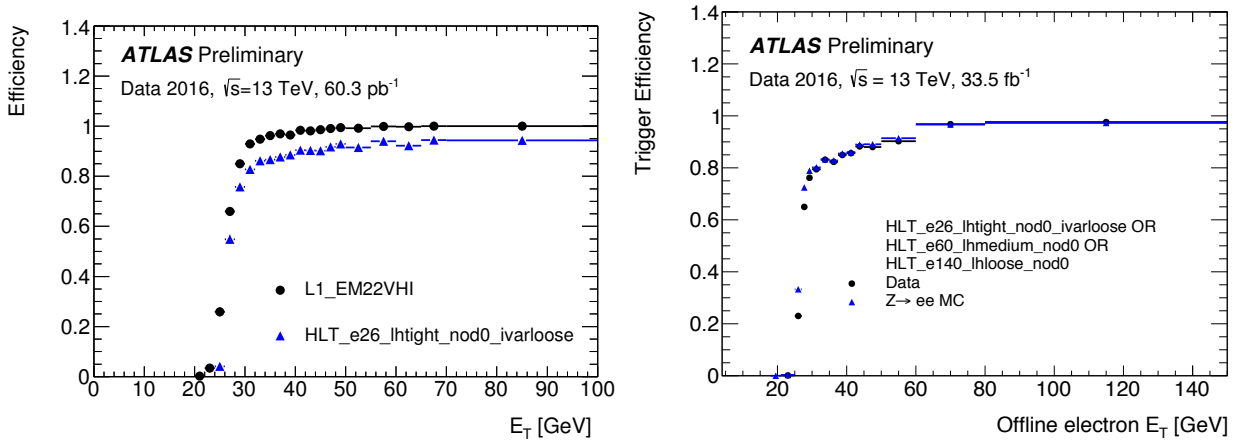


FIGURE 3.1: Left (Right) plot shows electron L1 trigger (HLT) efficiency in the 2016 runs [52]. All considered HLTs are combined as a logical sum.

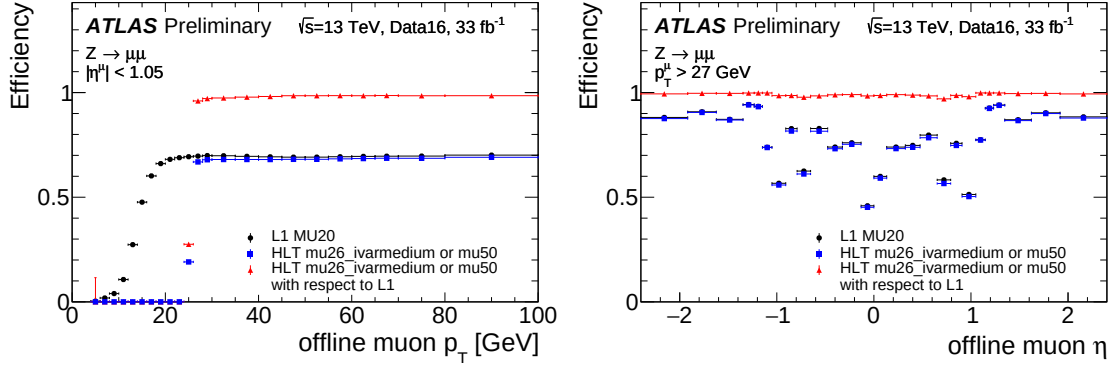


FIGURE 3.2: Left (Right) plot shows muon trigger efficiencies with relation to p_T (η) in the 2016 runs [53]. All considered HLTs are combined as a logical sum.

3.3 Offline Leptons

Leptons were reconstructed and identified efficiently using their corresponding detector information, and specific selections were applied to veto fake leptons. The selection criteria for the final offline leptons are summarized. Each lepton should pass the identification (ID) criterion, track-to-vertex-association (TTVA) cuts to ensure the track to originate from the hard scattering PV, isolation to reduce fake sources, p_T and η cuts for the trigger and detector coverage, and trigger matching with the reconstructed object.

The offline electrons were required to pass the following:

- tightLH ID criteria
- TTVA cuts
- isolation cut
- $|\eta| < 2.47$, but excluding the crack region ($1.37 < |\eta| < 1.52$) of the EM calorimeter coverage.
- leading lepton $p_T > 27$ GeV as required from the lowest p_T trigger threshold used in 2016, and the second leading lepton $p_T > 10$ GeV to veto fake components.
- leading p_T lepton is required to match the trigger object with $\Delta R < 0.15$ (see trigger section)

The offline muons used in this analysis are required to pass:

- medium combined ID criteria
- TTVA cuts
- isolation cut
- $|\eta| < 2.5$ as limited by the coverage of the inner detector
- leading lepton $p_T > 27$ GeV as required from the lowest p_T trigger threshold used in 2016, and the second leading lepton $p_T > 10$ GeV to veto fake components
- leading p_T lepton is required to match the trigger object with $\Delta R < 0.15$ (see trigger section)

3.4 Electron Reconstruction

Electrons were reconstructed using the measured energy deposited in the EM calorimeter and the tracks reconstructed in the inner detector. At first, seed EM clusters were built as electron candidates with a cluster window defined as 3×5 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ at the middle layer of the EM calorimeter. The measured energy clusters were constructed by merging nearby energy deposits. This cluster reconstruction had 95% efficiency at $p_T = 7$ GeV, and more than 99% for $E_T > 15$ GeV. As an electron can leave a track in the inner detectors, track-to-cluster matching was also performed to identify a cluster as an electron candidate.

Electron candidates were further evaluated with the following criteria to veto fake electrons and electrons from pile-up PVs.

- $p_T \geq 7$ GeV
- $|\eta| < 2.47$
- TTVA cuts: $|d_0^{BL}/\sigma_{d_0^{BL}}| < 5$ and $|(z_0^{BL} - z_0^{VTX}) \sin \theta| < 0.5$ mm

Here, η was calculated for the cluster position at the 2nd layer of the EM calorimeter. The d_0^{BL} (z_0^{BL}) is the transverse (longitudinal) impact parameter with respect to the beam line.

The electron ID algorithm was a likelihood-based (LH) method using probability density functions (PDFs) of variables as listed in Table 3.2 [54]. The discriminant, $d_{\mathcal{L}}$ was calculated as

$$d_{\mathcal{L}} = \frac{\mathcal{L}_s}{\mathcal{L}_s + \mathcal{L}_b}, \quad \mathcal{L}_{s(b)} = \prod_{\text{variable}; i} P_{s(b),i} \quad (3.1)$$

Three operating points were provided (tightLH, mediumLH, and looseLH) for the $d_{\mathcal{L}}$ selection thresholds, depending on p_T and η . The tightLH had the highest electron purity, which was adopted in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis. Some performance plots are shown in Figure 3.3.

The electron reconstruction and identification efficiency were measured using the 2016 data (see Figure 3.4) for electron p_T and η . The difference between the data and simulations was defined as a scale factor, including the uncertainties shown as error-bars in Figure 3.4.

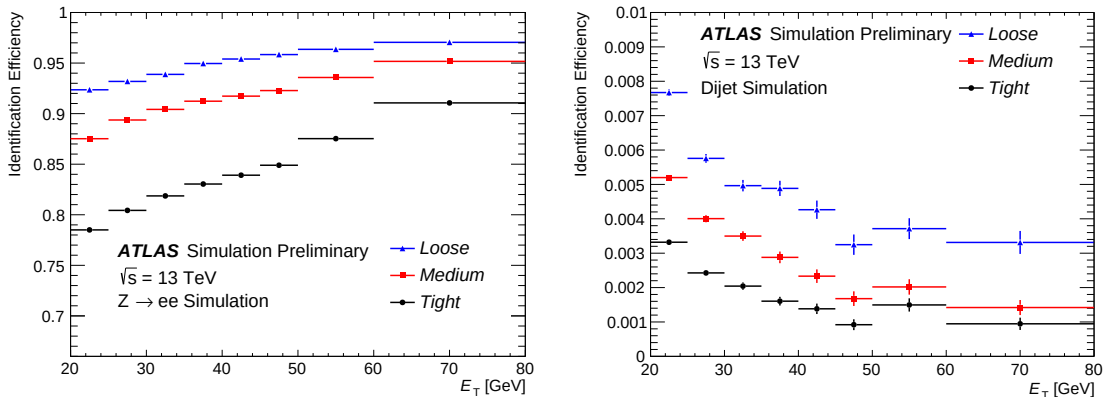


FIGURE 3.3: Electron ID performance plots as a function of the electron E_T [54]. The left plot shows the ID efficiency of real electrons estimated in a simulated Z +jets ($Z \rightarrow e^+e^-$) sample. The right plot shows the fake efficiency identifying jets as electrons evaluated in a dijets simulation sample.

TABLE 3.2: Definitions of electron discriminating variables [54].

Type	Description
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used in $ \eta < 0.8$ or $ \eta > 1.37$)
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used in $0.8 < \eta < 1.37$)
Back layer of EM calorimeter	Ratio of the energy in the back layer to the total energy in the EM calorimeter. This variable is used only below 100 GeV because of physical back layer leakage at higher energies
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - (\sum E_i \eta_i)^2/(\sum E_i)^2}$, where E_i is the energy of cell i at pseudo-rapidity η_i . The sum is calculated within a window of 3×5 cells
	Ratio of the energy difference between the largest and second largest energy deposits in the 3×5 cells
	Ratio of the energy in the strip layer to the total energy in the EM calorimeter
Track conditions	Number of hits in the innermost pixel layer to discriminate against photon conversions
	Number of hits in the pixel detector
	Number of total hits in the pixel and SCT detectors
	Transverse impact parameter with respect to the beam-line
	Significance of transverse impact parameter defined as the ratio of d_0 and its uncertainty
TRT	Momentum lost in the track between the perigee and the last measurement point divided by the original momentum
	Likelihood probability based on transition radiation in the TRT
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track
	$\Delta\phi$ between the cluster position in the middle layer and the track extrapolated from the perigee
	$\Delta\phi_2$, similar to $\Delta\phi$ but the track momentum is re-scaled to the cluster energy before extrapolating the track from the perigee to the middle layer of the calorimeter
	Ratio of the cluster energy to the track momentum

3.5 Muon Reconstruction

The muon ID criteria also had three working points of identification quality (tight, medium, and loose) in the trade-off between muon purity and efficiency. In this analysis, the medium muon was chosen, which had the smallest systematic uncertainties in muon reconstruction. The medium muon was also required to be reconstructed in both the muon spectrometer (MS) and the inner detector (ID), denoted as a combined muon. Combined muons were reconstructed with high quality identification, although the available $|\eta|$ range was limited to $|\eta| < 2.5$ due to the inner detector coverage.

The inner detector and the muon spectrometer first reconstructed the track segment separately. The tracks reconstructed in the muon spectrometers were defined as muon candidates, and a global fit was performed to match these muon candidates to the tracks in the inner detector.

Muon candidates were reconstructed in the muon spectrometers to pass the following criteria to reduce fake muons:

- $p_T > 4$ GeV

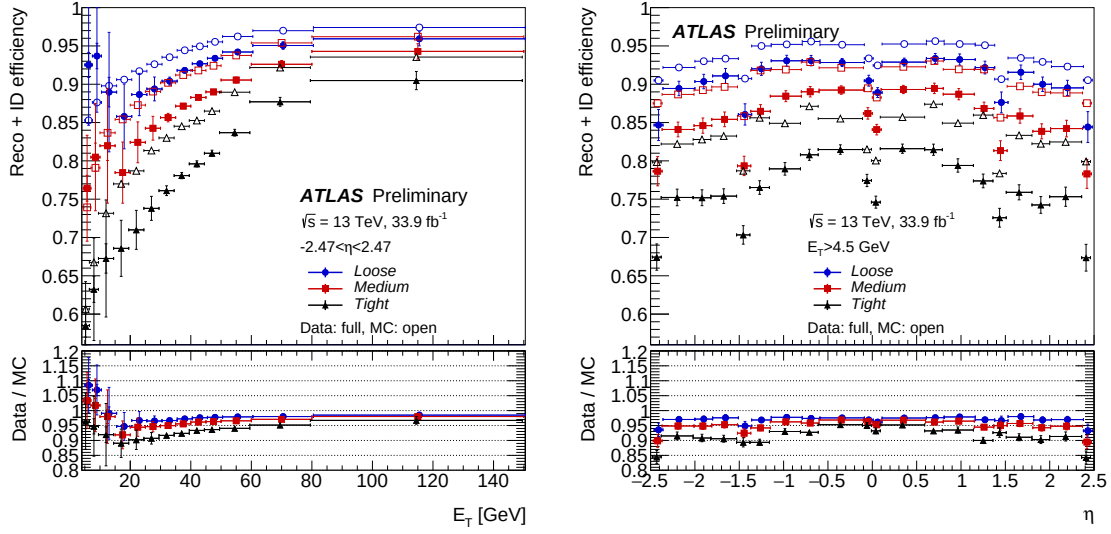


FIGURE 3.4: Electron reconstruction and ID efficiency measurements [55]. For the sample with dominant $Z+\text{jets}(Z \rightarrow e^+e^-)$ events, the efficiencies were measured for the three operating points. The left(right) plot shows efficiencies with relation to p_T (η). The results of two tag-and-probe methods are combined in these plots. One method requires the mass window of the Z boson, and the other only requires isolation around the probe electrons. Full procedures are discussed in [54]. The data efficiency is slightly lower than for the MC due to the fact that the MC does not properly represent the 2016 TRT conditions and known mis-modeling of calorimeter shower shapes in the GEANT4 detector simulation.

- $|\eta| < 2.5$
- $q/p \text{ significance} = ((q/p)_{\text{MS}} - (q/p)_{\text{ID}}) / \sqrt{\sigma_{(q/p)_{\text{MS}}}^2 + \sigma_{(q/p)_{\text{ID}}}^2} < 7$ for the charge q and the muon momentum p measured in the two segments MS and ID.
- $|\eta| > 0.1$, $n_{\text{precisionLayers}} > 1$. Otherwise, $n_{\text{precisionLayers}} == 1$ and $n_{\text{precisionHoleLayers}} < 2$.
- $|d_0^{\text{BL}} / \sigma_{d_0^{\text{BL}}}| < 3$ for the transverse IP significance
- $|(z_0^{\text{BL}} - z_0^{\text{VTX}}) \sin \theta| < 0.5 \text{ mm}$ for the longitudinal IP agreement

Then, the inner detector track segment was required to match muon candidates according to the following criteria:

- Number of pixel hits + number of dead pixels along the track ≥ 1
- Number of SCT hits + number of dead SCT strips along the track ≥ 5
- Number of pixel holes + number of SCT holes along the track ≤ 2
- In $0.1 < |\eta| < 1.9$ (= TRT acceptance), require $n > 5$ and $n_{\text{TRTOutliers}} < 0.9n$ where n_{TRTHits} ($n_{\text{TRTOutliers}}$) denotes the number of TRT hits (outliers) along the track, $n = n_{\text{TRTHits}} + n_{\text{TRTOutliers}}$.

The identification performance was estimated from the $t\bar{t}$ simulation sample, where prompt muons from W decay were used to evaluate the efficiency while hadron jets were used to examine the muon mis-identification fraction. The results are shown in Table 3.3.

TABLE 3.3: Muon ID efficiencies estimated using the $t\bar{t}$ MC sample [56]. The prompt muons from W decay are treated as genuine muons, while "muons" from hadrons are treated as fake muons to be rejected. An investigation was performed with muons for $20 < p_T < 100$ GeV and $|\eta| < 2.5$. The results for the three considered criteria are shown for comparison. Detailed procedures are described in [56].

Criteria	$\varepsilon_\mu[\%]$	$\varepsilon_{\text{Hadron}}[\%]$
loose	98.1	0.76
medium	96.1	0.17
tight	91.8	0.11

The combined muon reconstruction and identification efficiency was measured using the 2016 data. The results are shown in Figure 3.5 in relation to the muon p_T and η . The difference between data and simulation was taken into account as a scale factor, which was included in the systematic uncertainty.

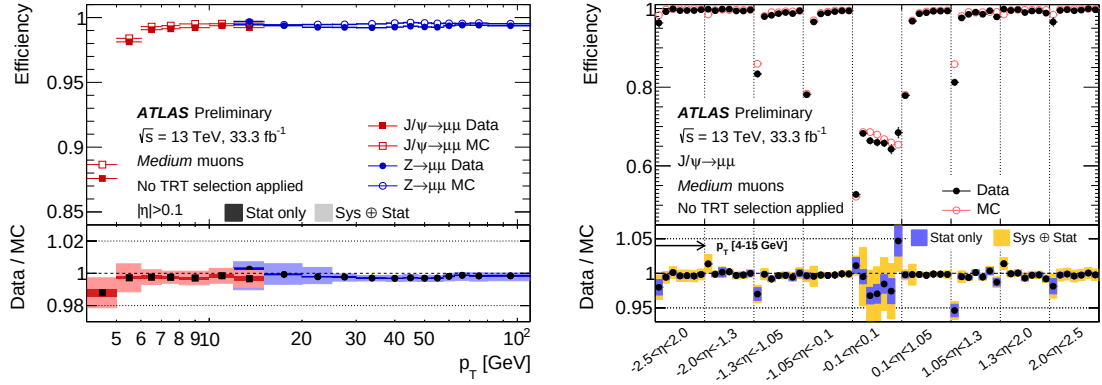


FIGURE 3.5: Muon reconstruction and ID efficiency measurements [57]. The Z +jets ($Z \rightarrow \mu^+\mu^-$) events, and J/ψ +jets ($J/\psi \rightarrow \mu^+\mu^-$) events selected for high and low p_T regions are used to measure efficiencies of medium combined muon selection. The left (right) plot shows efficiencies as a function of p_T (η) for $0.1 < |\eta| < 2.5$ (for $4 < p_T < 15$ GeV). Muon candidates with $4 < p_T < 15$ GeV are used for the muon η efficiency.

3.6 Lepton Isolation

Objects faking as leptons are:

- jets and photons that are accidentally identified as an electron accidentally
- electrons generated from photon conversion
- muons generated as leptonic decay products from heavy flavor hadrons, such as B hadrons

To reduce these contributions, lepton isolation was applied. There are two types of isolation schemes based on the calorimeter and tracker information.

- Calorimeter-based discriminating variable, $E_T^{\text{cone}0.2}/E_T$:
 $E_T^{\text{cone}0.2}$ is the sum of transverse energies for topological clusters within a cone of $\Delta R = 0.2$ around the lepton candidate, not including the lepton E_T of itself.

- Track-based discriminating variable for the electron (muon), $p_T^{\text{cone0.2}}/E_T$ ($p_T^{\text{cone0.3}}/p_T$):
 $p_T^{\text{cone0.2}}$ is the sum of transverse momenta of all tracks within a cone of $\Delta R = \min(0.2, 10 \text{ GeV}/E_T)$ around the electron, and $p_T^{\text{cone0.3}}$ is sum of transverse momenta of all tracks within a cone of $\Delta R = \min(0.3, 10 \text{ GeV}/p_T)$ around the muon candidate, not including the lepton E_T or p_T of itself. Required tracks have $p_T > 1 \text{ GeV}$, $|(z_0^{\text{BL}} - z_0^{\text{VTX}}) \sin \theta| < 3 \text{ mm}$, and hits on the pixel or SCT detectors.

The isolation efficiency ε_{iso} is given as a function of E_T (p_T) for electrons (muons). The operating point of the efficiency was $0.1143\% \times E_T(p_T) + 92.14\%$ in both schema. The performance of the isolation estimated using the Z+jets ($Z \rightarrow e^+e^-(\mu^+\mu^-)$) simulation sample resulted in typical efficiencies of 90/99% for $E_T(p_T) = 25/60 \text{ GeV}$.

3.7 Jet

3.7.1 Jet Reconstruction

Jets were reconstructed from three-dimensional topological energy clusters (topo-clusters)[58]. The topo-clusters were built from energy deposits in calorimeter cells according to the following procedures, which suppress contributions from pile-up of soft clusters and electronics noise in cells:

- Find the cell with $S = |E|/\sigma_{\text{noise in the cell}} > 4$ defined as "seed" of the cluster. The seed clusters are ordered in decreasing S .
- If cells with $S > 2$, defined as "neighbors", adjoin to the seed or its neighbors, they are merged with the cluster.
- Repeat until there are no neighbors.
- Perimeter cells with $S > 0$ are also merged with the cluster.
- Search for local maxima with $E > 500 \text{ MeV}$ each having at least four neighbors, but no neighbors with larger energy.
- Re-cluster around the local maxima with the same algorithm above, but with no thresholds and no merging applied. Cells at the cluster border are shared with weights depending on energies and distances from their local maxima.

All clusters were initially calibrated to the electromagnetic (EM) scale, namely the cluster energy is correct for electromagnetic showers.

After clustering, jets were reconstructed by the anti- k_t algorithm with a radius parameter of $R = 0.4$ [59]. In the anti- k_t algorithm, two distance parameters between clusters were calculated as:

$$d_{ij} = \min(k_{T,i}^{2p}, k_{T,j}^{2p}) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R} \quad (3.2)$$

$$d_{iB} = k_{T,i}^{2p} \quad (3.3)$$

where i and j are cluster indices, $k_{T,i}$, y_i and ϕ_i are the transverse momentum, rapidity, and azimuth of the i -th cluster, R is the radius parameter, p is the algorithm parameter which equals to -1 for the anti- k_t . If $d_{ij} < d_{iB}$, the i -th and j -th clusters are merged. Otherwise, the i -th cluster is defined as a jet and removed from the cluster list. The anti- k_t algorithm favors merging of soft

clusters around the cluster with high transverse momentum. Typical jets reconstructed in an event are shown in Figure 3.6.

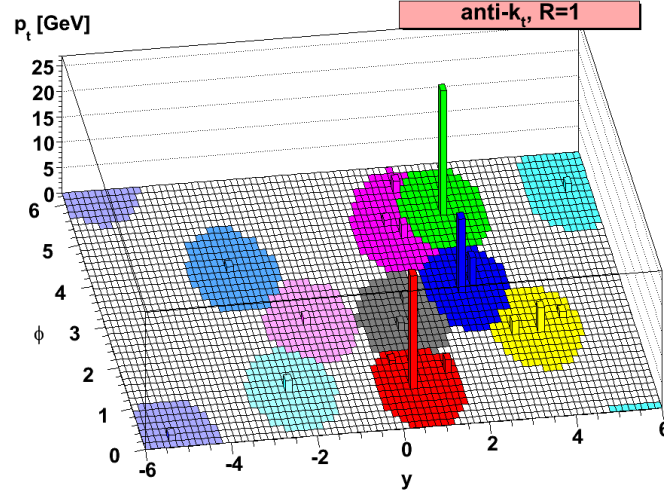


FIGURE 3.6: Typical jets reconstructed using anti- k_t algorithm with $R = 1.0$ [59]. Partons were generated with HERWIG. The soft components are associated to the energetic partons as illustrated by differently colored regions.

The jets reconstructed in the calorimeters do not have good resolution about their vertices. The tracks reconstructed in the inner detector would be associated to the jet for precise measurement of its vertex. In some simulation samples, it is often useful to assign the origin of partons to the reconstructed jets, such as in evaluations of the flavor tagging and of the correct assignments of jets to $t\bar{t}H$ objects. The ghost-association technique [60] is able to properly account for association of the tracks and parton origin. In the ghost-association technique, tracks or partons are added to the list of clusters in Equation 3.3 with their energies set to $\mathcal{O}(10^{-100} \text{ GeV})$. After clustering jets with the modified list, the tracks or truth partons are clustered together in the jet, giving proof that they are treated to be associated to a jet. This approach is properly adopted also for jets with irregular lateral distributions.

3.7.2 Jet Calibration

Jet calibration [61] was performed in several steps to achieve accurate jet description using both MC-base and in-situ calibration.

Origin Correction

The jet direction was corrected to indicate hard scattering PV instead of the origin of the ATLAS coordinate without changing the jet energy. The performance was evaluated in Run 1 as shown in Figure 3.7.

Pile-up Corrections[63]

Pile-up corrections are applied to hard scattering jets in order to correct their energy by subtracting the contamination from pile-up PV activities according to the following two schema.

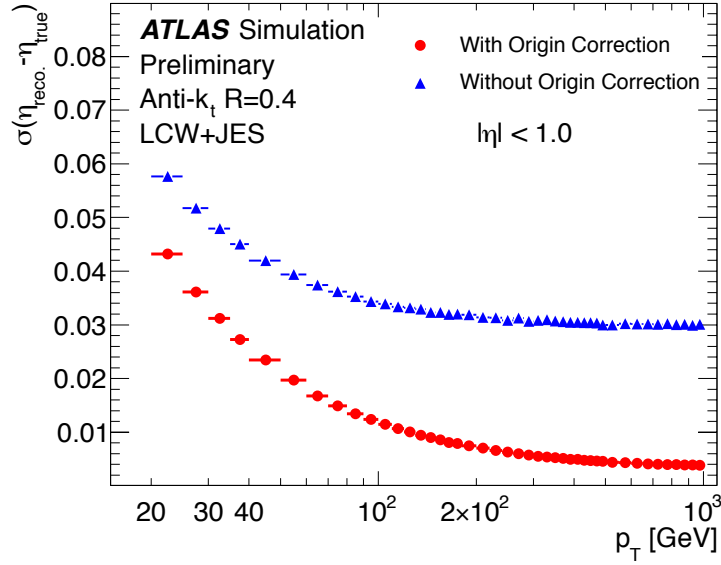


FIGURE 3.7: Pseudo-rapidity uncertainty with and without origin correction [62]. Although clusters in this plot are calibrated to the local cluster weighting (LCW) scale instead of the EM scale, their performance is unchanged.

The first scheme is a correction by the area-based subtraction according to the jet area and average energy density in $\eta \times \phi$ plane. This is used to subtract pile-up activities that uniformly overlap in the jet. The jet area A is defined by an active-area algorithm, where ghost particles of $p_T = \mathcal{O}(10^{-100} \text{ GeV})$ are uniformly added to the event. After the jet clustering, the number of ghosts clustered into each jet defines the jet area. The pile-up energy density in the range of $|\eta| < 2.0$ is calculated using the k_t algorithm, which is the same procedure as the anti- k_t but with $p = +1$ in Equation 3.3. The k_t algorithm is more sensitive to soft clusters. Then the pile-up energy density ρ is defined as the median of the jet energy densities (p_T/A) in each event. The median provides a good description of the pile-up energy density, because more pile-up jets are reconstructed in each event with smaller p_T than hard scattering jets. Typical distributions of the pile-up energy density are shown in Figure 3.8. The ρ topology uncertainty is assigned by the uncertainty in the underlying event contribution to ρ as estimated using several distinct MC generators and final-state topologies.

The other pile-up correction called the residual correction was applied after the area-based correction. The correction is subdivided into the in-time and out-of-time corrections. The former removes energy contributions from additional interactions in the same bunch crossing, which depends on the number of primary vertices N_{PV} . The latter removes contributions from multiple interactions before and after a few bunches, which depends on the average number of interactions per bunch crossing $\langle \mu \rangle$. To evaluate these contributions, the difference between reconstructed p_T^{reco} and p_T^{truth} which does not include any pile-up activities was studied using dijet simulation sample. The p_T^{truth} is defined as the transverse momentum of the "truth" jet which is reconstructed by the anti- k_t algorithm but does not use clustered cells on the calorimeter. Truth jets were required to match to reconstructed jets with $\Delta R < 0.3$. The difference, $\Delta p_T = p_T^{\text{reco}} - p_T^{\text{truth}}$, was fitted to a linear function of N_{PV} or $\langle \mu \rangle$ separately in bins of p_T^{truth} (20 to 200 GeV) and η^{truth} (0 to 4.5), where their fitted coefficients were denoted as $\alpha(p_T)$ and $\beta(p_T)$ for the N_{PV} and $\langle \mu \rangle$ dependences, respectively. Furthermore, due to the logarithmic dependence of $\alpha(p_T)$ and $\beta(p_T)$ on p_T , these coefficients are determined by additional fits in the full p_T^{reco} range separately in four η bins ($|\eta| =$

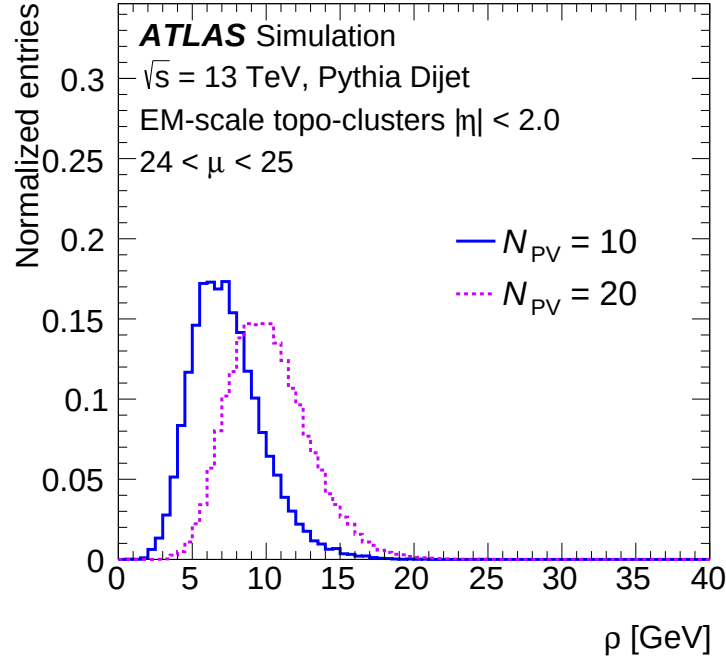


FIGURE 3.8: Pile-up energy density distribution, ρ for the average number of interactions per bunch crossing, $\langle\mu\rangle = 24$ to 25 [61]. The distributions are shown for 10 and 20 reconstructed PVs.

up to 1.2, 1.2 to 2.2, 2.2 to 2.8, 2.8 to 4.5). Their fitted coefficients are denoted as k_α and k_β as given in Equation 3.4, where the dependence on N_{PV} is expressed by normalizing the coefficient at $p_T = 25$ GeV bin,

$$\alpha(p_T) = \frac{\partial \Delta p_T}{\partial N_{PV}} = \left. \frac{\partial \Delta p_T}{\partial N_{PV}} \right|_{p_T=25 \text{ GeV}} + k_\alpha \times \log \left(\frac{p_T}{25 \text{ GeV}} \right). \quad (3.4)$$

The $\langle\mu\rangle$ dependence coefficient $\beta(p_T)$ was defined similarly.

To summarize the two pile-up correction schema, the corrected jet p_T is described as

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \rho \times A - \alpha(p_T^{\text{reco}} - \rho \times A) \times (N_{PV} - 1) - \beta(p_T^{\text{reco}} - \rho \times A) \times \langle\mu\rangle \quad (3.5)$$

where ρ is the pile-up energy density defined event-by-event, A is the jet area, α (β) is the in-time (out-of-time) coefficient as given in Equation 3.4 in the $|\eta|$ regions. The performance of the pile-up correction is shown in Figure 3.9.

Uncertainties in the modeling of N_{PV} and $\langle\mu\rangle$ were derived from the difference between the MC and data. The p_T -dependence of the linear coefficients also provided uncertainties from the fits in the full range of p_T^{reco} .

MC EtaJES Calibration

This calibration corrects for the jet energy dependence on η . The differences between reconstructed and truth jets are mainly from lower energy responses due to gaps in the calorimeter granularity and transitions between different calorimeter sub-components. For this calibration, the deficit in energy response was constructed from the mean of a Gaussian fit to the $E_{\text{reco}}/E_{\text{truth}}$ for each jet in bins of E_{truth} and η_{det} using the dijet MC sample, where η_{det} denotes the pseudo-rapidity of

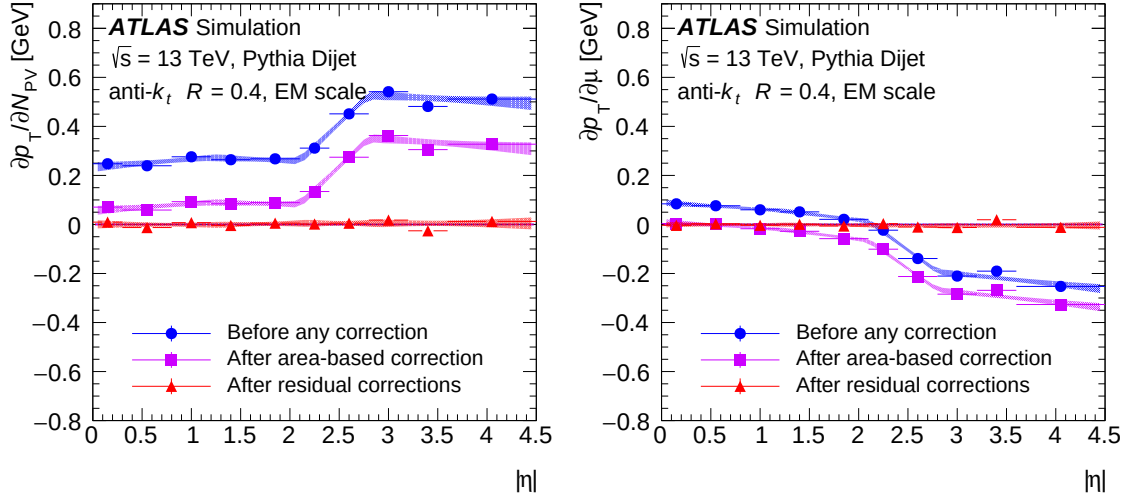


FIGURE 3.9: Pile-up correction to the jet p_T [61]. Left (Right) plot shows the in-time (out-of-time) pile-up dependence as a function of $|\eta|$ for $p_T^{\text{truth}} = 25$ GeV. Without any corrections, jets depend on the pile-up in terms of jet p_T ($\partial p_T / \partial N_{\text{PU}} \neq 0$ or $\partial p_T / \partial \mu \neq 0$), and on jet $|\eta|$ (non flat distributions). With both corrections applied, the pile-up dependences nearly vanishes.

the jet with the origin taken at the center of the ATLAS detector. The average energy response as a function of η_{det} is shown in Figure 3.10. After the correction, the η_{det} gaps between reconstructed and truth jets were observed in the barrel-endcap ($|\eta_{\text{det}}| \sim 1.4$) and in the endcap-forward ($|\eta_{\text{det}}| \sim 3.1$) transition regions, as shown in Figure 3.10. The η_{det} differences between the truth and reconstructed jets were also corrected.

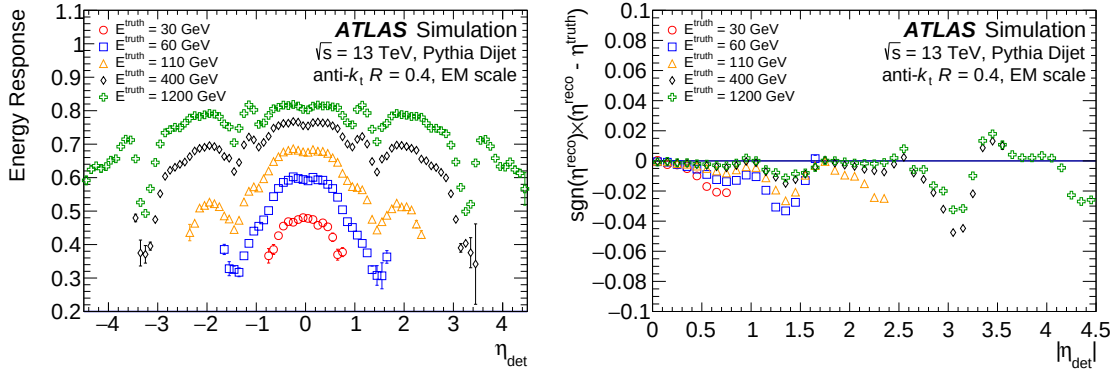


FIGURE 3.10: (Left) Average energy response as a function of $|\eta_{\text{det}}|$ for $p_T^{\text{truth}} = 30, 60, 110, 400$, and 1200 GeV. (Right) Inverse of the average response is applied as a correction factors. There remain several spikes in $|\eta_{\text{det}}| \sim 1.4$ and 3.1. These are also corrected as a function of η_{det} . [61]

Global Sequential Calibration

Global sequential calibration (GSC) calibrates the jet flavor composition and the shower shape, which are affected by the initial jet flavor (heavy quark, light quark, or gluon). Five variables were used sequentially to correct the jet responses as functions of p_T^{truth} and $|\eta_{\text{det}}|$ conserving the overall jet energy scale fixed at the EM+JES scale. The five variables are:

- f_{Tile0} : fraction of jet energy measured in the first layer of the hadronic calorimeter for $|\eta_{\text{det}}| < 1.7$.
- f_{LAR3} : fraction of jet energy measured in the third layer of the EM calorimeter for $|\eta_{\text{det}}| < 3.5$.
- n_{trk} : number of tracks with $p_T > 1$ GeV ghost-associated with the jet for $|\eta_{\text{det}}| < 2.5$.
- $\mathcal{W}_{\text{trk}}$: average p_T -weighted transverse distance in the η - ϕ plane between the jet axis and all tracks of $p_T > 1$ GeV ghost-associated to the jet for $|\eta_{\text{det}}| < 2.5$.
- n_{segments} : number of muon segment tracks ghost-associated with the jet for $|\eta_{\text{det}}| < 2.7$. These tracks are expected to come from very high-energy jets that are not fully contained in the calorimeter, i.e. punch-through jets.

The dependence of the jet response on each observable is reduced to less than 2% after the full GSC method is applied.

In-situ Energy Calibration

This calibrates differences of the energy and η response between data and MC simulation using the jet η and p_T against response to well-known physics objects as a reference.

Corrections to jet p_T values were performed with γ +jets, Z+jets and dijets events to compensate the residual p_T difference. At first, in-situ calibrations with γ +jets and Z+jets were performed. The leading jet p_T value (p_T^{probe}) is expected to be balanced to the well-measured electro-weak object (p_T^{ref}), such as γ and Z ($Z \rightarrow e^+e^-$ or $\mu^+\mu^-$). The individual sample had a statistical limit covering the jet p_T region corresponding to $20 < p_T < 500$ GeV for Z+jets and $36 < p_T < 950$ GeV for γ +jets.

In the γ +jets, events were required such that $\Delta\phi(\text{jet}, \gamma) > 2.8$ to be back-to-back, and that the second leading jet $p_T < \max(15 \text{ GeV}, 0.1 \times p_T^{\text{ref}})$ in order to reduce the multi-jet contamination. Afterwards, the direct balance (DB) technique was utilized in the following calculations:

$$R = p_T^{\text{probe}} / p_T^{\text{ref}} \quad \text{with} \quad p_T^{\text{ref}} = p_T^\gamma \cos(\phi^\gamma - \phi^{\text{probe}}) \quad (3.6)$$

where ϕ^{probe} is the jet azimuth to be calibrated, and $R \sim 1$ means the jet energy is well calibrated. The $\Delta\phi$ correction to p_T^γ was applied to reduce the effects of the un-balance due to multiple parton radiations, and the energy corrected p_T^γ was used as the reference.

Z+jets utilizes the missing- E_T projection fraction (MPF) technique rather than the DB technique, which has less sensitivity to the QCD radiation. Therefore, Z+jets events required less stringent selections compared to γ +jets events, such that $\Delta\phi(\text{jet}, Z) > 2.9$ and the second leading jet $p_T < \max(12 \text{ GeV}, 0.3 \times p_T^{\text{ref}})$. The relative jet response is defined by the average of the MPF response as

$$R = \left\langle 1 + \frac{\hat{n}_{\text{ref}} \cdot \vec{E}_T^{\text{miss}}}{p_T^{\text{ref}}} \right\rangle \quad (3.7)$$

where $\hat{n}_{\text{ref}}$ means the direction of the reference object Z , $\vec{E}_T^{\text{miss}}$ is defined in Section 3.8. As the MPF is defined independent to the probe jet objects, the MPF is also less sensitive to the jet reconstruction scheme than the DB technique. However, ϕ -symmetric pile-up and underlying-event activity could not be corrected well with the MPF.

With these two in-situ calibrations, the jet energy was well calibrated, especially in the low p_T region. Another in-situ calibration using dijet events can extend the p_T calibration up to 2 TeV. In dijet events, two objects are expected to be balanced between the largest p_T jet (p_T^{probe}) and the

collection of other well-calibrated low p_T jets (p_T^{ref}). The jet response is defined as $R = p_T^{\text{probe}}/p_T^{\text{ref}}$. The second largest p_T value was required to be less than 950 GeV to ensure a fully calibrated jet could be used as the reference.

All correction factors regarding residual p_T differences between data and MC simulations are shown in Figure 3.11.

A correction using jet η , η inter-calibration correction, was also performed with a dijet sample for the forward jets ($0.8 < |\eta| < 4.5$) with reference to the well-calibrated central jets ($|\eta| < 0.8$). Data and MC simulations of the η inter-calibration correction are plotted in Figure 3.11. The differences between MC simulations are generally within 1%, which was assigned as the systematic uncertainty.

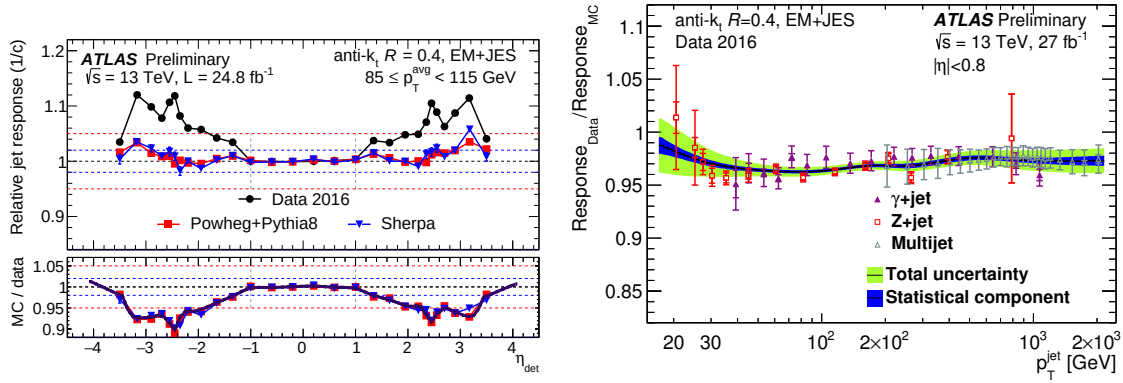


FIGURE 3.11: Relative response of jets in the data and simulation samples from in-situ calibration. The left plot shows the jet response calculated in the data and the dijets simulation sample used in the η inter-calibration correction. The difference between the data and POWHEG+PYTHIA simulated points are assigned as correction factors. The difference between two simulation samples are assigned as the systematic uncertainty. The right plot shows the ratio of the jet response between data and simulation samples. This calibration is performed using three simulation samples (γ +jets, Z+jets, and dijets), and the combined correction is shown in the black curve with blue/green uncertainty bands. [64]

3.7.3 Systematic Uncertainties Related to Jet Energy Scale Calibrations

The all calibrations described in the previous subsections provide a set of 80 jet energy scale (JES) systematic uncertainty terms:

- 23 sources from the in-situ calibration using Z+jets, including 13 statistical uncertainties in 13 jet p_T bins.
- 23 sources from the in-situ calibration using γ +jets, including 15 statistical uncertainties in 15 jet p_T bins.
- 21 sources from the in-situ calibration using dijets, including 16 statistical uncertainties in the 16 leading jet p_T bins.
- 3 sources from the η inter-calibration.
- 3 sources from the jet flavor.
- Uncertainty in the GSC punch-through correction.
- 4 sources from the pile-up correction.

- Non-closure uncertainty from difference in the absolute JES calibration using a fast detector response simulation [65] instead of a full detector simulation.
- High- p_T jet ($p_T > 2$ TeV) uncertainty from single-particle response and test-beam measurements.

The systematic uncertainties in the three in-situ calibrations were derived from the object reconstructions, selection variations as well as modeling differences, where no discernible benefits were found keeping full correlations among all systematic variations. To reduce the number of uncertainties to handle, global reduction [66] was performed through factorization of the matrix constructed by a total of 67 residual p_T uncertainties and their correlation factors. The seven uncertainties with the largest magnitudes were kept, and the rest uncertainties were merged into a single uncertainty, thus reducing to a set of eight completely uncorrelated uncertainties for residual p_T of the JES.

Three η inter-calibration uncertainties were derived from the differences between the MC simulation generators, the statistical uncertainty on each bin for the correction, and the non-closure uncertainty from difference using the measurement in $2.0 < |\eta| < 2.6$.

The three flavor response uncertainties were derived by comparing the average jet response for each jet flavor using two generators. On the other hand, the flavor composition uncertainty depends on the physics process, and it is difficult to find a common value for different physics analyses. Therefore, the value was derived from a MC simulation by assuming a 50% quark and 50% gluon composition, but with a conservative 100% uncertainty. The composition uncertainties between b and light-flavor jets were also derived in a similar manner as the gluon uncertainty.

Four sources from the pile-up correction come from the uncertainties in N_{PV} and $\langle\mu\rangle$ modeling, uncertainty in the residual p_T dependence, and the ρ topology uncertainty.

In total, the jet calibration procedures retain $8 + 13 = 21$ uncertainties, which are summarized in Figure 3.12.

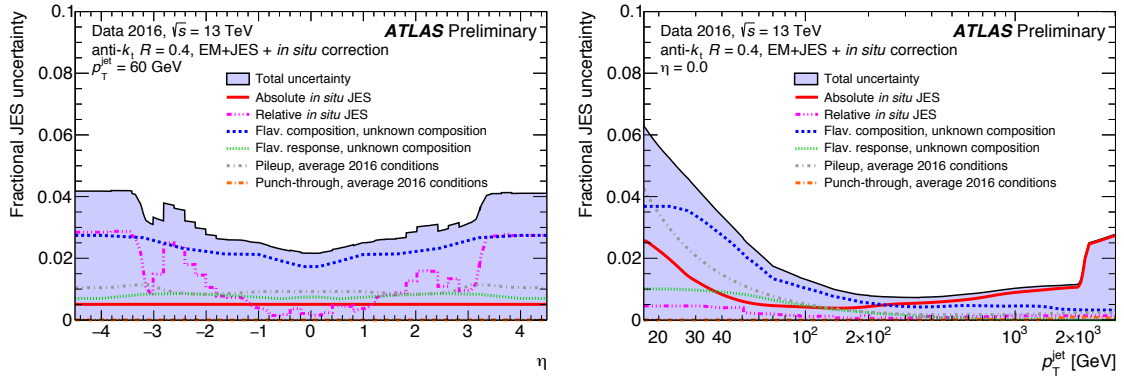


FIGURE 3.12: Fractional jet energy scale systematic uncertainty components. The red line shows the reduced residual p_T uncertainties composed of eight separate uncertainties. The pink dotted line shows the η -intercalibration uncertainty. The flavor composition and response uncertainties assume a constant quark and gluon composition in the simulation. [64]

3.7.4 Jet Selection

Reconstructed and calibrated jets were required for further criteria, which were used to reject jets from pile-up interactions:

- $p_T > 25$ GeV

- $|\eta| < 2.5$
- jet-vertex-tagger (JVT) cut
- if "bad" or "ugly" jets are in the event, the whole event is removed

The p_T selection reduces pileup jets significantly and large JES systematic uncertainties in the low p_T region (see Figure 3.12).

The JVT constructed from two variables, corrJVF and R_{p_T} , is a discriminant variable used to reject pileup jets [67]. This discriminant was used in the $20 < p_T < 60$ GeV range, where pile-up jets tend to be reconstructed as frequently as jets originating from the hardest primary vertex (hard scattered jets).

The jet-vertex-fraction (JVF) is defined as:

$$\text{JVF} = \frac{\sum_i p_T^{\text{track}_i}(\text{PV}_0)}{\sum_j \sum_i p_T^{\text{track}_i}(\text{PV}_j)} \quad (3.8)$$

where $p_T^{\text{track}_i}(\text{PV}_j)$ denotes the transverse momentum of the i -th track associated to the jet under evaluation and to the j -th primary vertex. PV_0 is the hardest vertex among the multiple primary vertices reconstructed in the same bunch crossing. JVF provides p_T fraction of tracks which are associated to the jet and originate from the hard scattering PV.

corrJVF is the corrected JVF used to reduce the dependence on the number of primary vertices. With the number of primary vertices, the sum of the track p_T from pile-up interactions increases, and as a result, the JVF decreases in high pile-up environments. To apply the JVF in various pile-up environments, corrJVF is defined as:

$$\text{corrJVF} = \frac{\sum_i p_T^{\text{track}_i}(\text{PV}_0)}{\sum_i p_T^{\text{track}_i}(\text{PV}_0) + \frac{\sum_{j>0} \sum_i p_T^{\text{track}_i}(\text{PV}_j)}{0.01 n_{\text{pileup track}}}} \quad (3.9)$$

where $n_{\text{pileup track}}$ is the number of tracks associated to the jet originating from any vertex other than hardest primary vertex. The coefficient 0.01 in $n_{\text{pileup track}}$, is the slope of the sum of the track p_T from pile-up interactions.

The R_{p_T} is defined as:

$$R_{p_T} = \frac{\sum_i p_T^{\text{track}_i}(\text{PV}_0)}{p_T^{\text{jet}}}, \quad (3.10)$$

which becomes zero for pile-up jets if all tracks from PV_0 are associated with jets from PV_0 . Similar to corrJVF , R_{p_T} does not directly depend on the number of primary vertices.

The JVT was performed using these two variables, thus constructing a single two-dimensional likelihood. The efficiency and fake rate described in this section were evaluated by a dijet simulation sample at $\sqrt{s} = 8$ TeV as shown in Figure 3.14. The JVT efficiency curve has the highest performance among all variables. JVT efficiencies with relation to the number of primary vertices for JVT cut at 0.6 or 0.85 are shown in Figure 3.14.

In the plot, the JVF curves are also shown for comparison. Both corrJVF and R_{p_T} are designed such that they do not depend on the number of primary vertices. Therefore, the JVT is almost flat against N_{PV} . For suppressing pile-up jets, the JVT cut was applied as $\text{JVT} > 0.59$ for jets with $p_T < 60$ GeV. The overall efficiency of the hard scattered jet evaluated in the 2015-2016 data is 92% after the JVT cut as shown in Figure 3.15. The evaluation was performed by selecting Z +jets ($Z \rightarrow \mu^+ \mu^-$) events with at least one jet balanced to the Z boson.

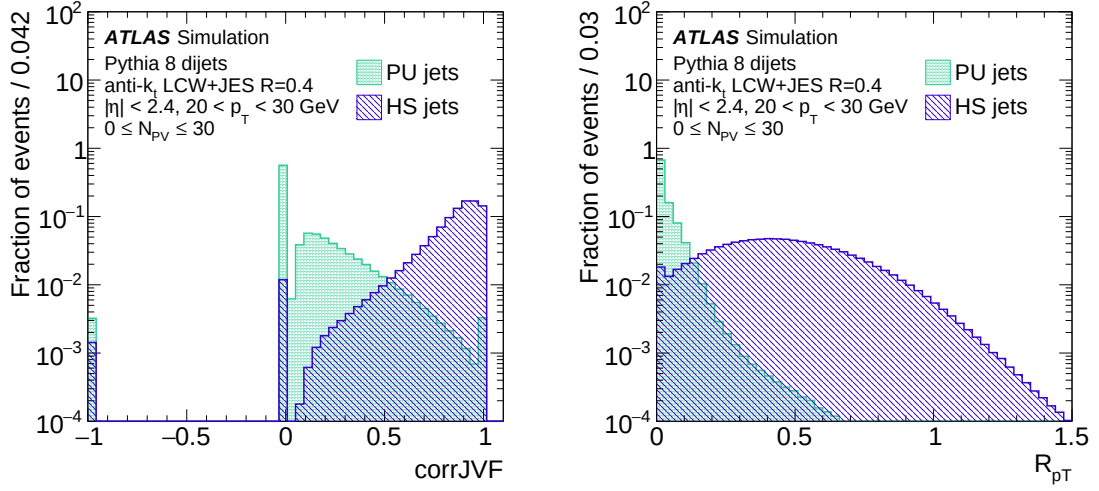


FIGURE 3.13: Distributions of the corrJVF (left) and R_{pT} (right). Each plot shows distributions of the hard scattered (HS) and pile-up (PU) jets. The content at corrJVF = -1 denotes that the jet has no associated tracks. These plots are evaluated with jets calibrated to the LCW scale at $\sqrt{s} = 8$ TeV, but the performance is same as the EM scale at $\sqrt{s} = 13$ TeV.[67]

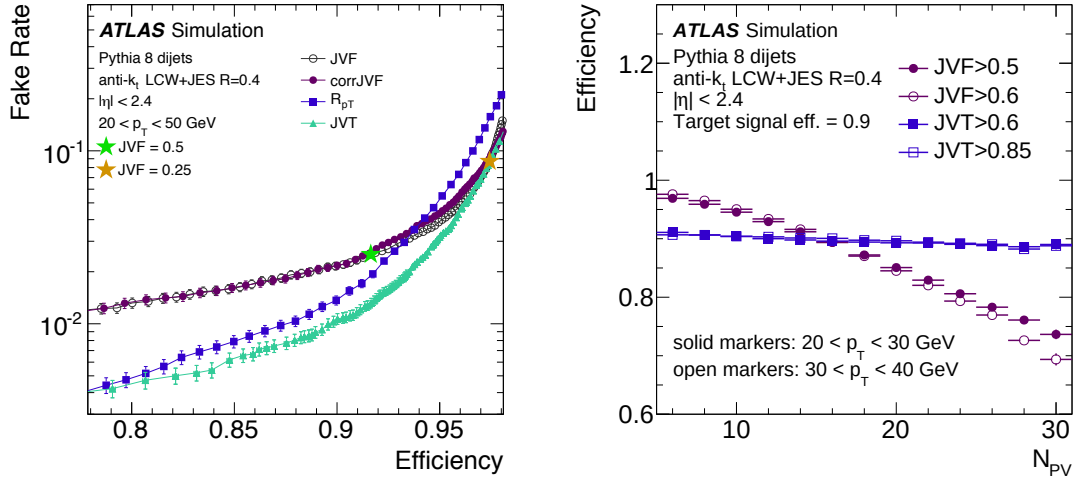


FIGURE 3.14: (Left) Efficiency curves for picking hard-scatter jets versus fake rate of picking pile-up jets using JVF, corrJVF, R_{pT} , and JVT. The JVF working points with cut values 0.25 and 0.5 are indicated with gold and green stars. (Right) The primary vertex dependence of the hard scattered jet efficiency for $20 < p_T < 30$ GeV (solid markers) and $30 < p_T < 40$ GeV (open markers) ranges for fixed cuts of JVT (blue squares) and JVF (violet circles) such that the inclusive efficiency is 90%. These plots are evaluated with jets calibrated to the LCW scale at $\sqrt{s} = 8$ TeV, but the tendency is same as the EM scale at $\sqrt{s} = 13$ TeV.[67]

3.8 Missing Transverse Energy

At the hadron collider, some fraction of the energy of the incoming particles escapes down the beam-pipe. Consequently, net momentum conservation is meaningful only in the transverse plane against the beam axis, and the missing transverse momentum is defined as the opposite direction

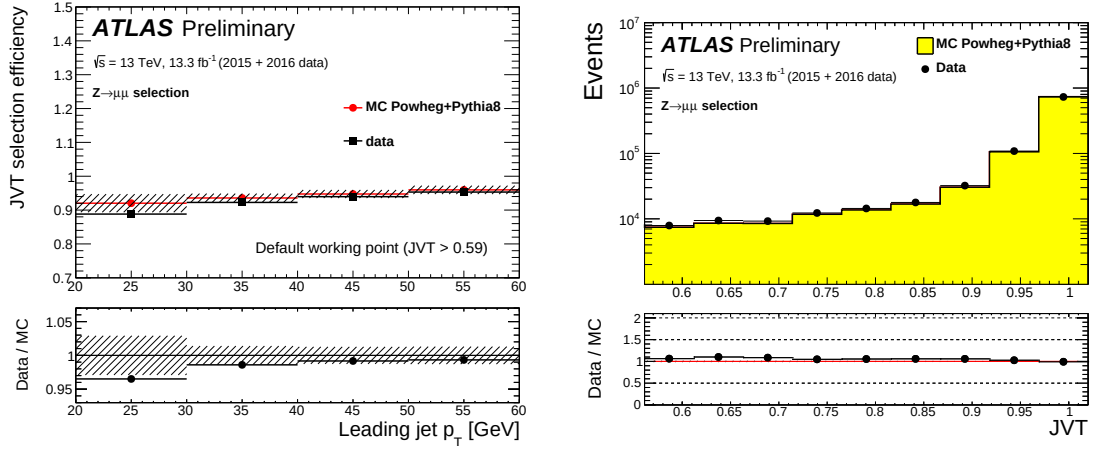


FIGURE 3.15: Comparisons between the data and the simulation for the JVT[68]. (Left) JVT efficiency for the hard scattered jets in POWHEG+PYTHIA6 Z+jets simulation sample and in 2015+2016 data, at a JVT > 0.59 cut and with Z ($Z \rightarrow \mu^+\mu^-$) selection. The evaluated jet is chosen to be balanced against Z boson. The difference between the MC and data is used as a scale factor. (Right) The JVT distribution from both MC and data, where the simulation has good agreement with the data.

of the vector sum of the transverse momenta for all objects, including soft tracks. Here, the soft tracks term is estimated from low p_T tracks associated to the primary vertex, but not associated to any reconstructed objects.

$$\vec{E}_T^{\text{miss}} = - \left(\sum_{obj} \vec{p}_T + \sum_{soft} \vec{p}_T \right) \quad (3.11)$$

The missing transverse energy (E_T^{miss}) is the magnitude of this vector.

3.9 Overlap Removal

After all objects were defined, an overlap removal procedure was applied to avoid double counting the same object of multiple types (electron, muon and jet candidates). If the object is assigned to multiple object candidates, the overlap removal is applied in the following order:

- If an electron candidate is found within $\Delta R < 0.2$ for any jet candidates, remove the jet object with smallest ΔR from the list of jet candidates. This avoids double counting of the electron as a jet at the same time.
- If an electron candidate is found within $\Delta R < 0.4$ for any jet candidates, remove the electron object. This avoids counting an electron emitted in the jet hadronization.
- If a muon candidate is found within $\Delta R < 0.4$ for any jet candidates and has ≥ 3 associated tracks, remove the muon object. This also avoids counting a muon emitted in the jet hadronization, especially from heavy hadron decay.
- If a muon candidate is found within $\Delta R < 0.4$ for any jet candidates but has less than 3 associated tracks, remove the jet candidate. This object is expected to be a high energy muon that can deposit significant energy in the calorimeter.

3.10 Flavor Tagging

Reconstructed jets originating from b -quark (b -jets) can be identified by finding a hadron including b -quarks which travels a few milli-meters before decaying. This algorithm is called " b -tagging", and this is one of the most important discriminants used in $t\bar{t}H(H \rightarrow b\bar{b})$ analysis. An illustration of b -tagging is shown in Figure 3.16. Hadron which includes b -quarks (c -quarks), is called B -hadron (C -hadron).

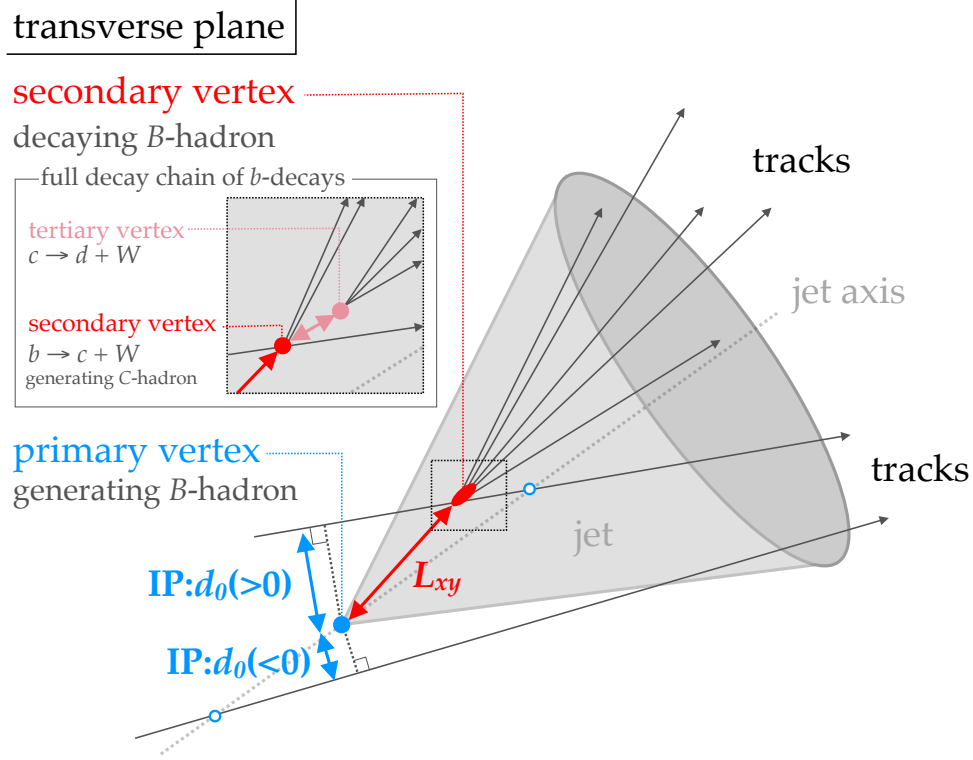


FIGURE 3.16: Illustration of b -tagging. A B -hadron travels L_{xy} in the transverse plane before decaying. Then the B -hadron decays into multiple hadrons, thus constructing the secondary vertex. Each track has an impact parameter (IP) defined by its three-dimensional or two-dimensional distance from the primary vertex. The sign of the track IP is positive (negative) if the track intersects the jet axis in the same (opposite) hemisphere of the jet direction.

A multivariate analysis (MVA) is used for b -tagging. The input variables used in b -tagging are the kinematics of the jet, and the variables related to three algorithms. The list of the input variables is shown in Table 3.4. Detailed explanations of the three algorithms are described in the following sub-sections.

3.10.1 Impact Parameter Based Algorithms

The impact parameter based algorithm calculates the sum of the log-likelihood of IP for all tracks associated to the jet. The likelihood describes the PDF of the IP significance under the assumptions of three flavors (b -jet, c -jet, and light-jet).

The tracks used in this algorithm are:

- $p_T^{\text{track}} > 1 \text{ GeV}$
- $|d_0| < 1 \text{ mm}$ and $|z_0 \sin \theta| < 1.5 \text{ mm}$

TABLE 3.4: Input variables for the b -tagging MVA constructed from the jet kinematics, and variables related to three algorithms. LLR is a log-likelihood ratio.

category	description
kinematics	transverse momentum of the jet
	pseudo rapidity of the jet
IP2D	LLR of d_0 between the b -jet and light-jet assumptions
	LLR of d_0 between the b -jet and c -jet assumptions
	LLR of d_0 between the c -jet and light-jet assumptions
IP3D	LLR of three dimensional IP between the b -jet and light-jet assumptions
	LLR of three dimensional IP between the b -jet and c -jet assumptions
	LLR of three dimensional IP between the c -jet and light-jet assumptions
SV1	Number of tracks from the secondary vertex
	Number of vertex candidates with two tracks
	Invariant mass of tracks from the secondary vertex, assuming pion masses
	Fraction of the charged jet energy in the secondary vertices
	Transverse distance between the primary and secondary vertices
	Distance between the primary and secondary vertices
	Significance of the average distance between the primary and secondary vertices
	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
JetFitter	Number of two-track vertex candidates (prior to the decay chain fit)
	Number of tracks from displaced vertices with ≥ 2 tracks
	Number of displaced vertices with more than one track
	Number of displaced vertices with one track
	Invariant mass of tracks from displaced vertices assuming pion masses
	Fraction of the charged jet energy in the secondary vertices
	ΔR between the jet axis and the vertical sum of the momenta for all tracks attached to the displaced vertices
	Significance of the average distance between the primary and displaced vertices

- ≥ 1 hit in the pixel detector allowing at most 1 hole
- ≥ 7 hits in the pixel and SCT detectors with at most 2 holes
- rejected if the vertex constructed by pairs of tracks originates from a light-hadron (K_s^0 or Λ^0) decay, photon conversion or interactions with material.

where "hole" is defined as a hit associated to the track traversing through the disabled modules. Signed impact parameters d_0 or z_0 are defined for each track with respect to the primary vertex, as shown in Figure 3.16. The significances d_0/σ_{d_0} or $z_0 \sin \theta / \sigma_{z_0 \sin \theta}$ are used for the likelihood. The IP2D algorithm adopts the significance of d_0 into the PDF for each flavored jet (see Figure 3.17). On the other hand, the two significances are used in a two-dimensional template PDF in the IP3D algorithm accounting for their correlation.

All PDFs are evaluated by using the $t\bar{t}$ simulation sample. However, if the IBL and the b-layer pixel has no hits when a hit is expected, the PDFs are evaluated by the $t\bar{t}$ and $Z' \rightarrow t\bar{t}$ simulation samples ($m_{Z'} = 4$ TeV). The large amount of high p_T b -jets from $Z' \rightarrow t\bar{t}$ allows the construction of PDFs without statistical limitation, which is difficult with the $t\bar{t}$ sample only.

As the final discriminant to separate b -jets from light-jets, a log-likelihood ratio (LLR) is evaluated as $\text{LLR} = \sum_{i=1}^N \log \left(\frac{p_b}{p_u} \right)$, where the b -jet (light-jet) likelihood is denoted as $p_b(p_u)$. This LLR distribution is shown in Figure 3.17. Similarly, other flavor hypothesis combinations of the flavor hypotheses are calculated and used as the input variables into the MVA.

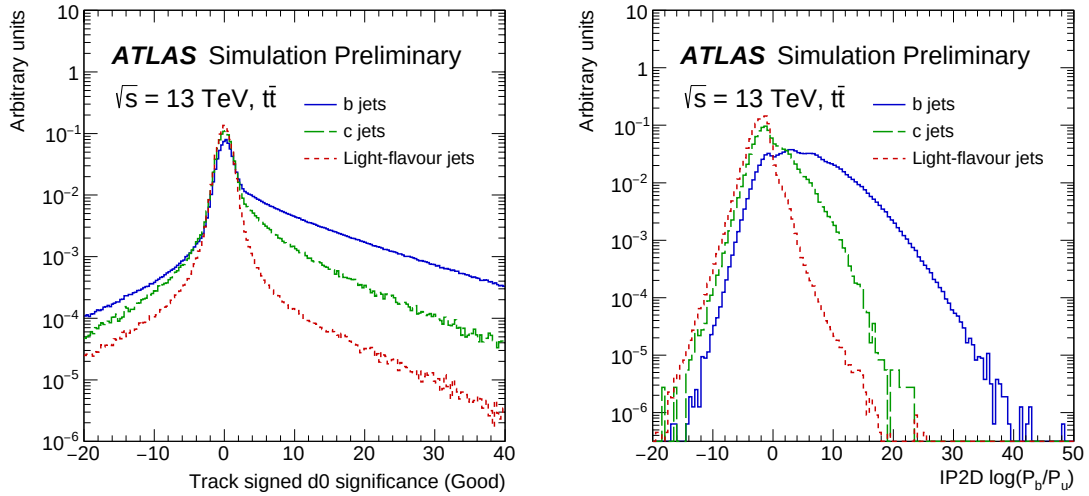


FIGURE 3.17: Significance of the transverse signed track impact parameter, and LLR for the IP2D algorithm [69]. For b -jets (solid blue), c -jets (dashed green) and light-jets (dotted red), the LLR was evaluated using the $t\bar{t}$ simulation sample with the b -jet and light-jet hypotheses.

3.10.2 Secondary Vertex Finder

The secondary vertex finder (SV) explicitly reconstructs the displaced secondary vertex within the jet. In the SV algorithm, all pairs of tracks within a jet are considered as two-track vertices. If the two-track vertex is likely to originate from long-lived light-hadron decays, photon conversions or material interactions, it is discarded. Afterwards, all remaining vertices and their associated tracks are re-fit to one secondary vertex with any outliers removed.

Tracks included in this algorithm are required to pass the following selections:

- $p_T^{\text{track}} > 0.4 \text{ GeV}$
- $|d_0| < 3.5 \text{ mm}$
- reject tracks with small d_0 and large z_0 significances to reduce fake vertices ($d_0/\sigma_{d_0} < 2$ and $z_0/\sigma_{z_0} > 6$)
- ≥ 1 hit in the pixel detector
- ≥ 7 hits in the pixel detector and SCT, or ≥ 8 hits in $|\eta| > 1.5$, since hadronic interactions could occur more in the high η than in the central region due to the increasing amount of detector materials. This makes track reconstruction difficult, and results in worse track quality in this region if the same requirements used in the central region are applied.
- at most 1 hit shared with other tracks
- the χ^2 statistic for the track quality is less than 3
- leading 25 tracks are allowed at most in each jet. High energy jets (e.g. $p_T > 300 \text{ GeV}$) tend to have many tracks from fragmentation, and the vertices constructed by these tracks can be mis-tagged as the secondary vertex of the b -jet.

b -jet is tagged by looking for a two-track vertex that is displaced from the primary vertex, where vertices originating from fake sources should be suppressed. Therefore, several selections on the constructed two-track vertices are considered:

- significance of the distance between the two-track vertex and the primary vertex > 2
- sum of impact parameter significances of the tracks associated to the two-track vertex > 2
- χ^2 of the fitted tracks < 4.5
- mass of two-track system $< 6 \text{ GeV}$
- tracks emitted from K_s^0 and Λ^0 decays can be rejected by using the mass of the constructed vertex. The invariant mass spectra can identify the peaks due to K_s^0 and Λ^0 decays. Charged tracks coming from such vertices are rejected

Some distributions of the input variables to the MVA after combining associated all qualified tracks into one secondary vertex are shown in Figure 3.18.

3.10.3 Multi-vertex Fitter

The multi-vertex fitter algorithm (JetFitter) focuses on full reconstruction of the decay chain for the B -hadron. The decay positions of $PV \rightarrow b \rightarrow c$ are reconstructed through full reconstruction of the decay chain. JetFitter finds a common line along which all three decay vertices lie and provides positions of the B - and C -hadron decays, even when only one track can be attached to one of these decay processes. The efficiencies of vertex reconstructions with at least one or two tracks are shown in Figure 3.19. The reconstruction efficiency becomes significantly higher by including one-track vertices, although mis-identified efficiencies for light-jets also increases due to accidental coincidence of large numbers of tracks in high p_T jet fragmentations. Typical distributions used in the b -tagging MVA are shown in Figure 3.20.

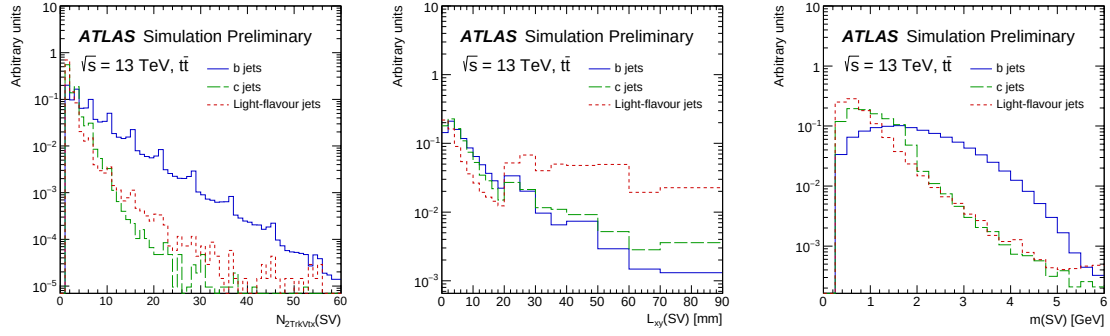


FIGURE 3.18: Discriminants into the MVA for b -tagging used in secondary vertex finder algorithm for b -jets (solid blue), c -jets (dashed green) and light-jets (dotted red), as evaluated using the $t\bar{t}$ simulation sample [69]. The number of two-track vertices reconstructed within the jet (left), the transverse decay length (center), and the invariant mass (right). The rate of light-flavour jets stays flat at large transverse decay length due to material interactions.

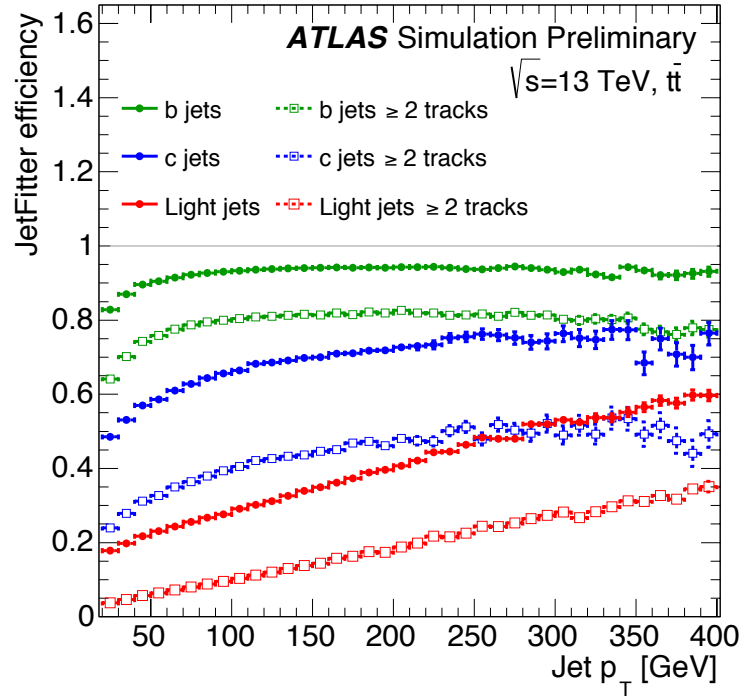


FIGURE 3.19: JetFitter vertex reconstruction efficiency as a function of jet p_T [70].

3.10.4 Discriminant Score

To maximize b -tagging performance, the three algorithms are combined to one MVA output distribution, MV2, using the boosted-decision-tree (BDT) algorithm. BDT training was performed with b -jets as the signal and non b -jets (mixture of c -jets and light-jets) as the background to maximize the b -jets efficiency and c -jets and light-jets rejections. The mixture ratio affects the rejection performance for c -jets and light-jets. Three mixture ratios were considered in the 2015 and 2016 data, which are denoted by the c -jet fraction in the background sample: MV2c00, MV2c10, and MV2c20, with 0%, 7%, and 15% c -jet fractions, respectively. The efficiency of b -jets with relation to the rejection factor is shown in Figure 3.21. In this analysis, importance is placed on the separation

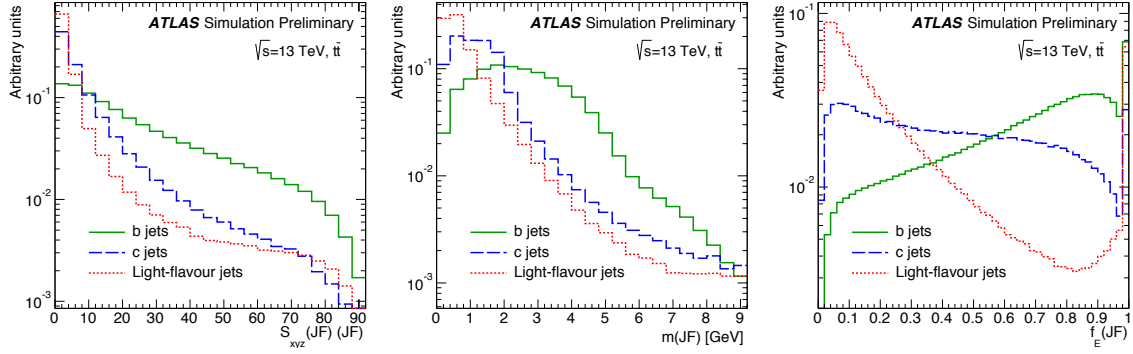


FIGURE 3.20: Discriminants in the MVA for b -tagging used in JetFitter for b -jets (solid green), c -jets (dashed blue) and light-jets (dotted red) evaluated using the $t\bar{t}$ simulation sample [70]. The average flight length significance of the reconstructed vertices (left), the invariant mass of tracks fitted to one or more displaced vertices (center), and the energy fraction defined as the energy of the tracks associated to the displaced vertex relative to all tracks reconstructed within the jet (right).

between b -jets and c -jets, rather than between b -jets and light-jets. MV2c10 and MV2c20 have the highest c -jet rejection performance, and MV2c10 has a higher rejection performance for light-jets than MV2c20. Therefore, the analysis adopted MV2c10 as the b -tagging discriminant. The BDT output distributions are shown in Figure 3.22.

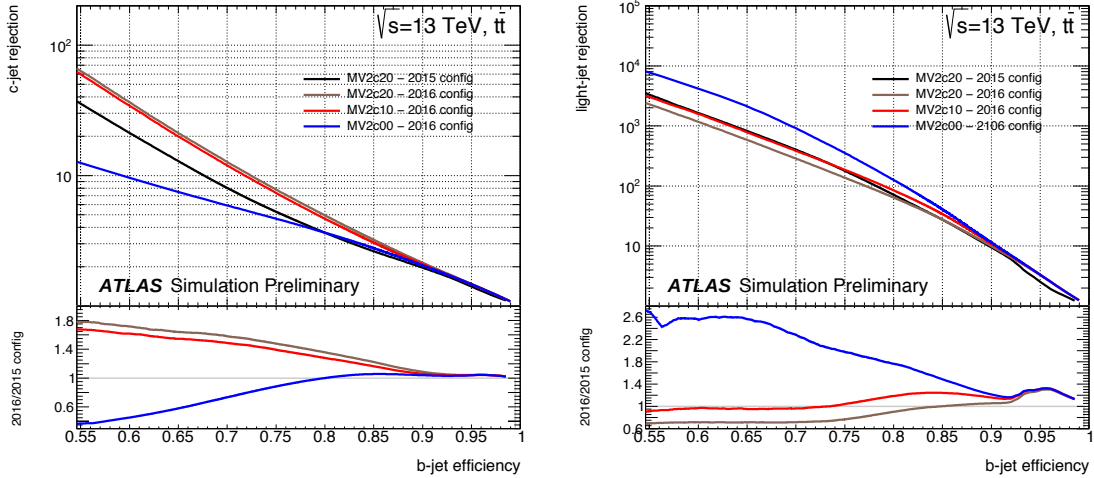


FIGURE 3.21: c -jet and light-jet rejection performance versus b -jet efficiency at various MV2 b -tagging settings with different c -jet fractions in the BDT training data (MV2c00, MV2c10, and MV2c20) [69]. These are evaluated using $t\bar{t}$ simulated events. Results in the previous configuration used in 2015 are also shown. From the 2015 configuration, the IP2(3)D and SV algorithms in 2016 have shown several improvements by optimizing the track and vertex requirements.

Four working points (WPs) are implemented for MV2c10, and their performances are shown in Table 3.5.

3.10.5 Flavor Tagging Calibration

The b -tagging calibrations were performed by measuring the tagging efficiency rate of b -jets, c -jets, and light-jets respectively. The efficiency difference between the data and simulation was treated as a correction factor and is defined as $\varepsilon_{\text{data}}/\varepsilon_{\text{MC}}$, where $\varepsilon_{\text{data}}$ (ε_{MC}) denotes the efficiency measured

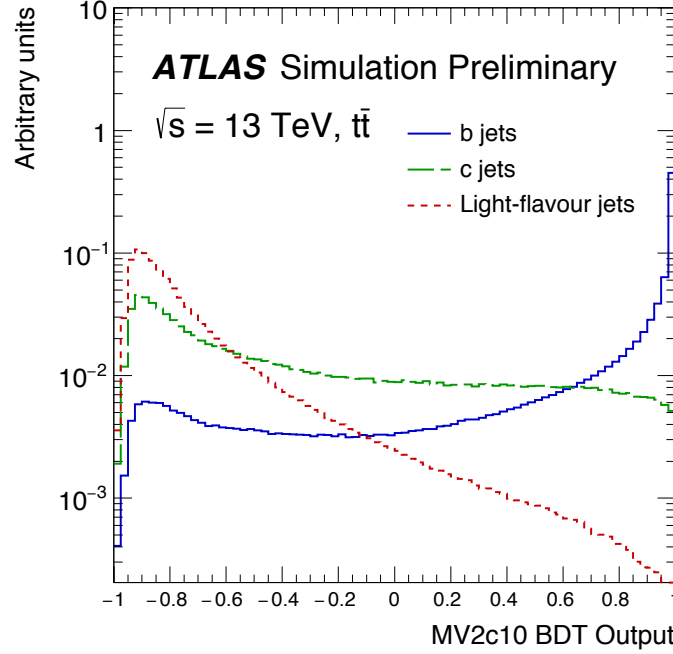


FIGURE 3.22: Distributions of the b -tagging score for b -jets (solid green), c -jets (dashed blue) and light-jets (dotted red) as evaluated using $t\bar{t}$ simulation sample [69].

TABLE 3.5: b -tagging working points for MV2c10 [69]. All values are evaluated for jets with $p_T > 20$ GeV, $|\eta| < 2.5$, and JVT > 0.59 if $p_T < 60$ GeV in the $t\bar{t}$ simulation sample. Each WP has a fixed threshold for the MV2c10 b -tagging discriminant in the full range of jet p_T and η as shown in the "MV2c10" column.

name	b -jet efficiency	purity	c -jet rejection	τ -jet rejection	light-jet rejection	MV2c10
60%	60.03 %	99.00 %	34.54	183.98	1538.78	0.935
70%	69.97 %	97.46 %	12.17	54.72	381.32	0.824
77%	76.97 %	95.17 %	6.21	22.04	134.34	0.646
85%	84.95 %	89.66 %	3.10	8.17	33.53	0.176

in the data (simulation). Different methods are employed to measure the efficiency in each flavor to maximize the calibration performance and to reduce their uncertainties. It is better to take a phase space near that is similar to the phase space used in this analysis, but the two phase space should not overlap. The calibration methods are summarized below, where the events have no overlaps with the events studied in this analysis:

- $t\bar{t}$ PDF likelihood method for the calibration for b -jet efficiency [71]
- use W boson decays in $t\bar{t}$ events [72] or $W+c$ events [73] for the calibration for c -jet rejection
- negative tag method for the calibration for light-jet rejection [74]

The b -jet tagging efficiency was measured in dileptonically decaying $t\bar{t}$ events requiring exactly two leptons with opposite charges and 2 or 3 jets. To maximize the performance, four categories were defined, depending on the number of jets (2 or 3) and flavors of the two final state leptons ($e\mu$ or $ee, \mu\mu$). In the categories with the same lepton flavors, $E_T^{\text{miss}} > 60$ GeV and $m_{\ell\ell} > 50$ GeV

were required to reduce the fake leptons and low-mass resonances decaying into ee or $\mu\mu$. A mass region in $80 < m_{\ell\ell} < 100$ GeV dominated by Z bosons was excluded from the b -tagging efficiency estimation. A BDT was trained using the $t\bar{t}$ sample to select $\bar{b}$ events from $t\bar{t}$ decays as signal, and reject all other events as background. Input variables were based on the kinematics and topologies of the largest p_T dijet system. After excluding events with the BDT low score, a MV2c10 PDF per b -jet p_T bin, $\text{PDF}_b(w|p_T)$ for b -jet was parametrized to maximize the likelihood in each of the four categories. For example, the likelihood in 2-jet categories is defined as

$$\begin{aligned} \mathcal{L}(p_{T1}, p_{T2}, w_1, w_2) = & (f_{bb}\text{PDF}_{bb}(p_{T1}, p_{T2})\text{PDF}_b(w_1|p_{T1})\text{PDF}_b(w_2|p_{T2}) \\ & + f_{bl}\text{PDF}_{bl}(p_{T1}, p_{T2})\text{PDF}_b(w_1|p_{T1})\text{PDF}_l(w_2|p_{T2}) \\ & + f_{ll}\text{PDF}_{ll}(p_{T1}, p_{T2})\text{PDF}_l(w_1|p_{T1})\text{PDF}_l(w_2|p_{T2}) \\ & + 1 \Leftrightarrow 2) / 2. \end{aligned} \quad (3.12)$$

w_i and p_{Ti} are MV2c10 score and p_T of the i -th jet in each event, respectively. f_{bl} is the fraction of events having one b -jet and one light-jet. $\text{PDF}_{bl}(p_{T1}, p_{T2})$ is the two-dimensional PDF for $[p_{T1}, p_{T2}]$ for the flavor combination [b -jet, light-jet]. Fractions and PDFs having other flavor combinations (bb and ll) were defined similarly. $\text{PDF}_l(w, p_T)$ was evaluated using the $t\bar{t}$ simulation sample. The $\text{PDF}_{bl}(w, p_T)$ is the two-dimensional PDF of the probability for two jets to have p_{T1} and p_{T2} , which was also estimated from the simulation. If a single working point is used, $\text{PDF}_b(w, p_T)$ in each jet p_T bin is a two-binned histogram. On the other hand, if all four working points are used, $\text{PDF}_b(w, p_T)$ is a five-binned histogram. Evaluated $\text{PDF}_b(w, p_T)$'s in four categories were combined into one PDF, and the difference between the data and simulation was assigned as a scale factor. An example performance for b -jet tagging efficiency of 77% WP is shown in Figure 3.23. Calibration uncertainties are typically a few % for jet p_T values up to 300 GeV.

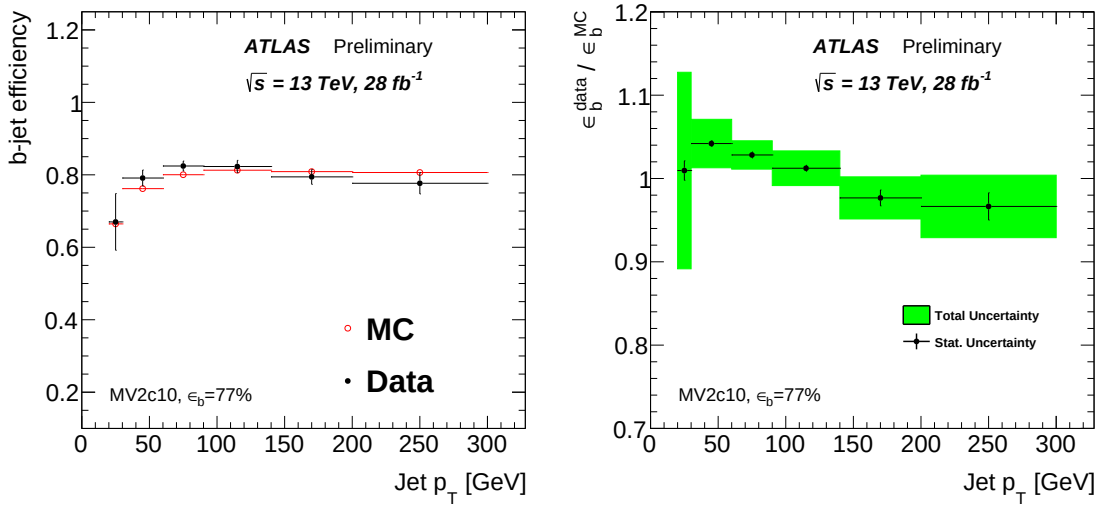


FIGURE 3.23: (Left) b -tagging efficiency and (right) the ratio of the data to simulation at the 77% WP as a function of jet p_T . These results were obtained using the $t\bar{t}$ PDF method [75].

The c -jet calibration was performed with single-leptonic $t\bar{t}$ or $W+c$ events. Single-leptonic $t\bar{t}$ events were required to contain one lepton and $E_T^{\text{miss}} > 20$ GeV to reduce fake lepton contributions. Exactly 4 jets, including at least 2 b -tagged jets at 70% WP, were also required to reject non $t\bar{t}$ events. Kinematic likelihood fitter (KLFitter) [76] was used to determine the jet assignment to the partons in $t\bar{t}$ system. KLFitter was performed assuming masses and widths of W boson and top quark. After excluding events with the low KLFitter score, the highly enriched $t\bar{t}$ events ($\sim 96\%$)

are retained. The mis-tagging efficiency of c -jets was measured using two jets assigned to decay products from W , where at most one jet was allowed to be b -tagged. The log likelihood function was defined by the number of observed and expected events divided into two dimensional bins depending on leading and sub-leading p_T values of the two jets. The c -tagging efficiency among the b -tagged jets of the W decays was parametrized in the likelihood. Flavor compositions and number of expected pre-tagged events were derived from the simulation. Tagging efficiencies for b -jets and light-jets after the calibrations described in this section were used. The log-likelihood function simultaneously corrects the normalizations of the $t\bar{t}$ sample bin-by-bin (in total six correction factors). The efficiencies evaluated in every bin were compared to the simulated values and their differences were treated as scale factors. In this sample, the fraction of W decays including at least one c -jet is $\sim 40\%$. Therefore, the c -jet efficiency depends on the calibration of tagging efficiencies for b -jets and light-jets. The obtained c -jet calibration uncertainties are less than typically 20% for jet p_T values up to 140 GeV. The c -jet mis-tagging efficiency at the 77% WP is shown in Figure 3.24.

The $t\bar{t}$ based calibration has a phase space near $t\bar{t}H(H \rightarrow b\bar{b})$. The c -jet efficiency measurement was also performed using $W+c$ events. This calibration was used instead of the $t\bar{t}$ based calibration when we used the Run 1 based signal and control regions that overlap with the region used in $t\bar{t}$ based calibration. The c -jet mis-tagging efficiency in $W+c$ events at the 77% WP is shown in Figure 3.24. The efficiency uncertainties are not so different, but the evaluated phase space is apart from the signal region for $t\bar{t}H(H \rightarrow b\bar{b})$.

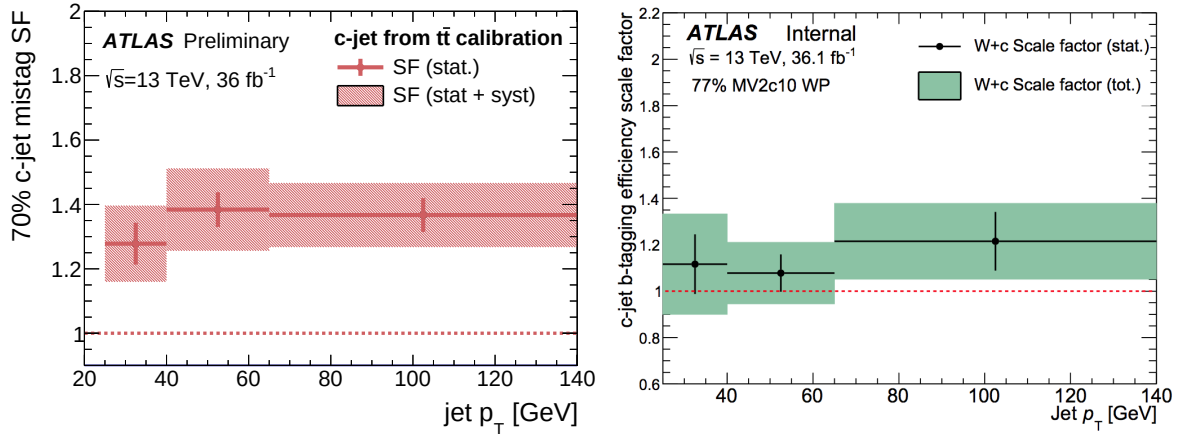


FIGURE 3.24: Ratio of c -jet mis-tagging efficiency from the data and the simulation at the 77% WP as a function of jet p_T . These values were obtained using $t\bar{t}$ events, where one of the W boson decays into c and s quarks (left), or using $W+c$ events (right).

Due to the large light-jet rejections, it is difficult to measure the light-jet mis-tagging efficiency. The light-jet mis-tagging is caused by the un-completeness of the detector response and performance. To calibrate the light-jet efficiency, another dedicated MV2c10 algorithm (MV2c10Flip) was developed, which has a similar efficiency for light-jets but much lower efficiencies for b -jets and c -jets than MV2c10. The intrinsic discriminant for b -tagging is the impact parameter, as shown in Figure 3.17. The positive tails in these distributions were used for b -tagging, while the light-jets shapes are almost symmetric between the positive and negative impact parameters. Therefore, MV2c10Flip used the same input variables as the three basic algorithms, but the tracks and vertices were modified by flipping the sign of the impact parameters and all related variables, or by using only negatively signed tracks. The performance of MV2c10Flip was confirmed using a multi-jet simulation sample, as shown in Figure 3.25. The light-jet events have similar shapes

($\lesssim 50\%$ difference) in both MV2c10 and MV2c10Flip below ~ 0.9 . Therefore, this calibration is more reliable with loose working points.

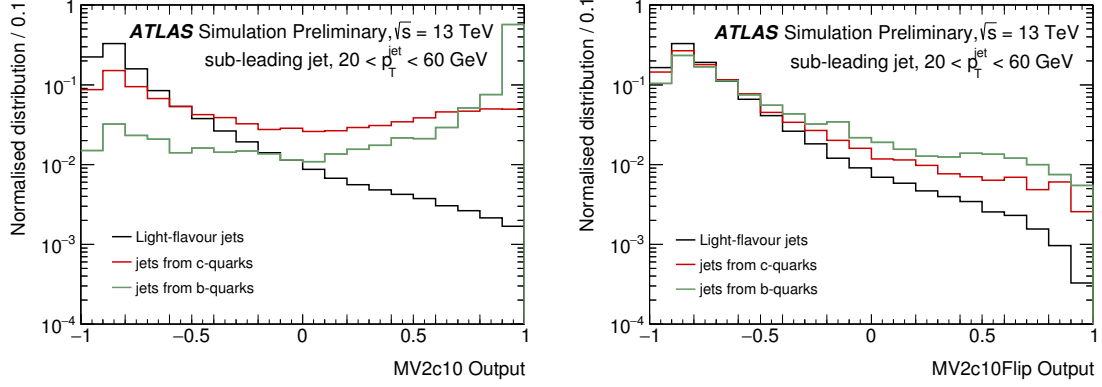


FIGURE 3.25: MV2c10 (left) and MV2c10Flip (right) distributions for the three flavored jets in multi-jet events [77]. The second highest p_T jet is $20 < p_T < 60$ GeV to avoid trigger biases.

The efficiency measurement was performed by mathematical calculation of the light-jet efficiency ε_l using the tagging efficiency in the MV2c10Flip, $\varepsilon_{\text{flip}}^{\text{data}}$:

$$\varepsilon_{\text{flip}}^{\text{data}} = f_{\text{HF}} \varepsilon_{\text{HF, flip}} + (1 - f_{\text{HF}}) \varepsilon_{l, \text{flip}} \quad (3.13)$$

$$\varepsilon_l = \frac{\varepsilon_l^{\text{MC}}}{\varepsilon_{l, \text{flip}}^{\text{MC}}} \cdot \varepsilon_{l, \text{flip}} = \frac{\varepsilon_l^{\text{MC}}}{\varepsilon_{l, \text{flip}}^{\text{MC}}} \cdot \frac{1}{f_{\text{HF}} \left(\frac{\varepsilon_{\text{HF, flip}}^{\text{MC}}}{\varepsilon_{l, \text{flip}}^{\text{MC}}} \right) + (1 - f_{\text{HF}})} \cdot \varepsilon_{\text{flip}}^{\text{data}}. \quad (3.14)$$

The heavy flavor fraction (f_{HF}) in the events, the ratio of light-jet efficiencies on MV2c10 and MV2c10Flip ($\varepsilon_l^{\text{MC}}/\varepsilon_{l, \text{flip}}^{\text{MC}}$), and the ratio of efficiencies on MV2c10Flip for heavy flavor jets and light-jets ($\varepsilon_{\text{HF, flip}}^{\text{MC}}/\varepsilon_{l, \text{flip}}^{\text{MC}}$) were taken from the simulation.

The light-jet efficiency was measured in two dimensional bins of jet p_T and η , as shown in Figure 3.26 for the light-jet mis-tagging efficiency at the 77% WP.

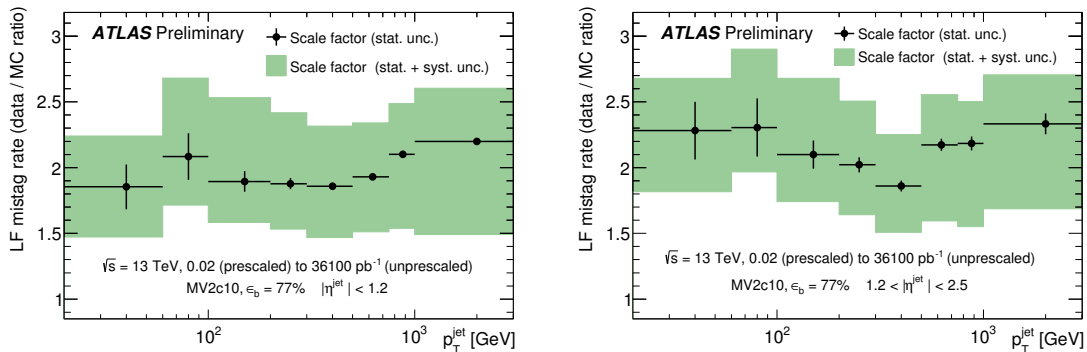


FIGURE 3.26: Ratio of light-jet mis-tagging efficiency from the data and simulation at the 77% WP as a function of jet p_T using the negative tag method [77]. The left plot is evaluated when $|\eta|$ for the jet less than 1.2, while the right plot is evaluated when $1.2 < |\eta| < 2.5$. These are applied as scale factors in the jet p_T and η bins.

Uncertainties for the scale factors in b -tagging calibrations were derived by varying each uncertainty source by $\pm 1\sigma$. These uncertainties were fed into an eigen-variation model, similar to the one performed for the JES uncertainties, in order to reduce the number of systematic uncertainty parameters to handle. Some of the largest fractions were kept as separate uncertainties, while all other uncertainty sources were merged into a single variation. When a single working point is used, a total of 6, 4, and 16 uncorrelated uncertainties are considered for b -jets, c -jets and light-jets, respectively. For multiple working points, as used in the pseudo-continuous b -tagging, these numbers are multiplied by five tag weight bins, resulting in 30, 20, and 80 for b -jets, c -jets and light-jets, respectively. These systematic uncertainties were taken as uncorrelated among three flavors.

3.11 $t\bar{t}H(H \rightarrow b\bar{b})$ Objects

3.11.1 Event Selections

In the single-lepton channel, no kinematic selections regarding leptons and E_T^{miss} were applied in order to maintain high signal efficiency. In the dilepton channel, dilepton mass was required to be above 15 GeV and outside of the Z boson mass window of $83 \sim 99$ GeV in order to remove events including Z bosons decaying into $\ell^+\ell^-$.

In the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, fake-lepton and W +jets events have large contributions, which should be efficiently removed via b -tagging requirements. Therefore, it is better to require a minimum pre-selection by b -tagging and reject only events in a region far from the signal enriched phase space. At least two b -tagged jets at 85% WP are required. More precise discussions on the number of jets (nJets) and the number of b -tagged jets (nBTags) are given in Chapter 8.

There are overlap events in the single-lepton and dilepton channels. For example, if the events include two leptons, where one has the p_T value more than 27 GeV, and the other has the p_T value in $10 < p_T < 27$ GeV, they pass the criterion in both channels. These overlapped events were removed from the single-lepton channel.

3.11.2 Truth-Matching

In the $t\bar{t}H$ analysis, one or two leptons in the single-lepton or dilepton channel, respectively, were required to pass the tight ID and isolation requirements, and one of them was tagged by the trigger. Therefore, the reconstructed leptons can be matched to $t\bar{t}H$ leptons with least ambiguity.

On the other hand, the six or four jets in the single-lepton or dilepton channel, respectively, are difficult to reconstruct all of them as separate jets. Parton quarks outside of the detector η range or with p_T values lower than 25 GeV would be not reconstructed correctly as jets. Kinematic acceptance efficiencies for quarks at the truth-level are summarized in Table 3.6. Almost 50% of $q2_{\text{had}W}$ are missed, mainly due to their low p_T values. The $b2_{\text{Higgs}}$ has 72% efficiency, which is smallest among the b -partons. The efficiency that all quarks are inside the acceptance range is 56-59%.

If two quarks are emitted close to each other, they may be merged into one jet when reconstructed. Because the jet radius parameter is set to 0.4, the partons should be reconstructed as individual jets if ΔR is not less than 0.4. The ΔR of two particles becomes small if the parent object has a high p_T value, which is called "boosted". Relations between parent p_T and ΔR between its decay objects are shown in Figure 3.27 for the W_{had} and the Higgs boson in $\text{nJets} \geq 6$ with $\text{nBTags}_{85\%} \geq 4$. In this figure, the boosted fraction is 6.8%, where two decay objects from either of the W_{had} or the Higgs boson are close ($\Delta R < 0.4$). Because of the multiple quarks in the event, two quarks can have $\Delta R < 0.4$ accidentally. This probability is around 7% per quark, and in total 20%

TABLE 3.6: Truth parton acceptance efficiency with kinematic selections at the reconstructed-level from single-lepton $t\bar{t}H$ events.

partons	$b1_{\text{Higgs}}$	$b2_{\text{Higgs}}$	$b1_{t\bar{t}}$	$b2_{t\bar{t}}$	$q1_{\text{hadW}}$	$q2_{\text{hadW}}$	$4b$	$4b+q1_{\text{hadW}}$	all
nJets ≥ 6 with nBTags _{85%} ≥ 4 in the reconstructed-level									
$p_T > 25$ GeV	1.00	0.93	1.00	0.94	0.99	0.71	0.87	0.87	0.61
$ \eta < 2.4$	0.99	0.97	0.99	0.97	0.98	0.94	0.93	0.91	0.86
$\rightarrow \text{Both}$	0.99	0.92	0.99	0.92	0.98	0.69	0.83	0.82	0.56
nJets ≥ 6 with nBTags _{77%} ≥ 4 in the reconstructed-level									
$p_T > 25$ GeV	1.00	0.95	1.00	0.95	1.00	0.70	0.91	0.90	0.63
$ \eta < 2.4$	0.99	0.98	0.99	0.98	0.99	0.94	0.94	0.93	0.88
$\rightarrow \text{Both}$	0.99	0.94	0.99	0.94	0.98	0.68	0.87	0.86	0.59
nJets ≥ 6 with nBTags _{70%} ≥ 4 in the reconstructed-level									
$p_T > 25$ GeV	1.00	0.96	1.00	0.96	0.99	0.69	0.93	0.92	0.64
$ \eta < 2.4$	0.99	0.98	0.99	0.98	0.99	0.94	0.95	0.94	0.88
$\rightarrow \text{Both}$	0.99	0.95	0.99	0.95	0.98	0.67	0.89	0.88	0.59
nJets ≥ 6 with nBTags _{60%} ≥ 4 in the reconstructed-level									
$p_T > 25$ GeV	1.00	0.97	1.00	0.98	0.99	0.68	0.95	0.94	0.64
$ \eta < 2.4$	0.99	0.99	1.00	0.98	0.98	0.92	0.96	0.95	0.88
$\rightarrow \text{Both}$	0.99	0.96	1.00	0.96	0.98	0.66	0.92	0.90	0.59

of events have at least one pair of quarks with $\Delta R < 0.4$. To summarize, the total fraction of the "good" events is 46%, where all truth quarks have $p_T > 25$ GeV, $|\eta| < 2.4$, and are separated from each other by $\Delta R \geq 0.4$. In "good" events, all jets are to be reconstructed and correctly assigned to partons. The results for other b -tagging WP cases are shown in Table 3.7. Noticeable differences were not observed.

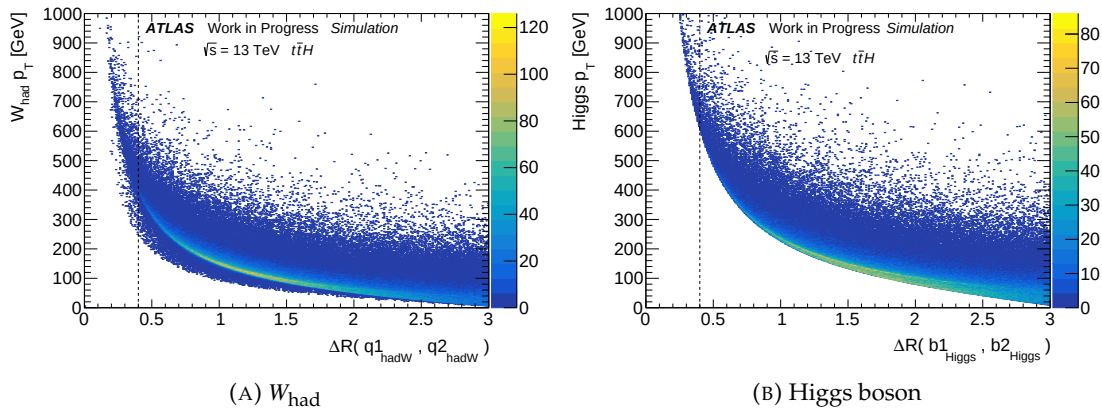


FIGURE 3.27: p_T v.s. ΔR for (A) two quarks from W_{had} decay and (B) two b -quarks from Higgs boson decay. Fractions of $\Delta R < 0.4$ are 1.2% and 5.8% for W_{had} and Higgs boson with corresponding p_T values of approximately 350 GeV and 600 GeV, respectively. The fraction of either decay objects with $\Delta R < 0.4$ is 6.8%.

TABLE 3.7: Truth-level boosted fraction in the single-lepton $t\bar{t}H$ events. The first and second columns show the fraction of two decay partons from the W_{had} and Higgs boson with $\Delta R < 0.4$ individually. The third column shows the fraction of two decay partons with $\Delta R < 0.4$ arising either from the W_{had} or Higgs boson. The forth column shows the total fraction that all partons have $\Delta R > 0.4$ against any other partons. The last column shows the fraction of "good" events, where all partons have $p_T > 25$ GeV, $|\eta| < 2.4$, and are separated from each other by $\Delta R > 0.4$.

boosted fraction			all isolated partons	"good" events
Higgs boson	W_{had}	either of two		
nJets ≥ 6 with nBTags _{85%} ≥ 4 in the reconstructed-level				
0.00038	0.0049	0.0052	0.83	0.46
nJets ≥ 6 with nBTags _{77%} ≥ 4 in the reconstructed-level				
0	0.0052	0.0052	0.85	0.49
nJets ≥ 6 with nBTags _{70%} ≥ 4 in the reconstructed-level				
0	0.0050	0.0050	0.85	0.50
nJets ≥ 6 with nBTags _{60%} ≥ 4 in the reconstructed-level				
0.00017	0.0050	0.0051	0.86	0.51

Truth-matching was performed by requiring $\Delta R < 0.35$ between the reconstructed jet and truth parton. The results are summarized in Table 3.8. The matching efficiency losses are explained by changes in object directions due to scattering in the detectors or by high momentum parton radiation. These losses are significant, especially for low momentum objects.

TABLE 3.8: Truth matching efficiency. "Inclusive" means all $t\bar{t}H$ events used to calculate truth matching efficiencies, while "good events" means using only events where all quarks have $p_T > 25$ GeV, $|\eta| < 2.4$, and are separated from each other by $\Delta R > 0.4$.

partons	$b1_{\text{Higgs}}$	$b2_{\text{Higgs}}$	$b1_{t\bar{t}}$	$b2_{t\bar{t}}$	$q1_{\text{hadW}}$	$q2_{\text{hadW}}$	$4b$	$4b+q1_{\text{hadW}}$	all
nJets ≥ 6 with nBTags _{85%} ≥ 4									
inclusive	0.98	0.85	0.98	0.87	0.88	0.57	0.71	0.62	0.34
good events	0.99	0.91	0.98	0.93	0.94	0.77	0.82	0.77	0.59
nJets ≥ 6 with nBTags _{77%} ≥ 4									
inclusive	0.99	0.91	0.99	0.92	0.88	0.56	0.82	0.71	0.39
good events	1.00	0.94	0.99	0.96	0.94	0.77	0.89	0.83	0.64
nJets ≥ 6 with nBTags _{70%} ≥ 4									
inclusive	0.99	0.94	0.99	0.94	0.88	0.55	0.87	0.76	0.42
good events	1.00	0.96	1.00	0.97	0.94	0.76	0.93	0.87	0.67
nJets ≥ 6 with nBTags _{60%} ≥ 4									
inclusive	1.00	0.95	1.00	0.96	0.88	0.53	0.91	0.80	0.43
good events	1.00	0.97	1.00	0.98	0.94	0.75	0.95	0.89	0.68

The fraction of fully truth-matched $t\bar{t}H$ events are from 34% to 43% in ≥ 6 jets events. Therefore, the analysis should take the events with some missing partons/jets into consideration when optimizing the discriminants to separate $t\bar{t}H$ and background sources.

Chapter 4

Run 1 Analysis

Previous studies of the search for $t\bar{t}H(H \rightarrow b\bar{b})$ have been reported in [78]. The analysis used 20.3 fb^{-1} of pp collision data at $\sqrt{s} = 8 \text{ TeV}$ taken by the ATLAS detector in Run 1 of the LHC from 2010 to 2012. Signal and control regions were defined using the number of leptons, jets and b -tagged jets at 70% working point of the Run 1 b -tagging algorithm [79] as summarized in Table 4.1 and Figure 4.1.

TABLE 4.1: S/B and $S/\sqrt{B}$ obtained in Run 1 regions shown separately for the single-lepton and dilepton analysis channels [78]. The number of jets considered in the single-lepton (dilepton) channel is 4, 5 and ≥ 6 (2, 3, ≥ 4), while the number of b -tagged jets is 2, 3 and ≥ 4 at 70% working point in both channels. S/B and $S/\sqrt{B}$ are calculated with simulation samples with the signal predicted by the SM. The "distribution" column shows the variable used in the fit to determine the $t\bar{t}H(H \rightarrow b\bar{b})$ signal strength. Signal regions utilize multi-variate score distributions using the Neural Network (NN) to separate signal and background events, while control regions utilize the scalar sum of p_T of jets (H_T^{had}) in the single-lepton channel, and the scalar sum of p_T of jets, leptons and E_T^{miss} (H_T^{all}) in the dilepton channel to control the systematic uncertainties coming from jet performances and $t\bar{t}$ +jets modelings.

channel	region	nJets	nBTags _{70%}	S/B	$S/\sqrt{B}$	distribution
single-lepton	(4j, 2b)	4	2	3.2×10^{-4}	0.099	H_T^{had}
	(4j, 3b)	4	3	1.6×10^{-3}	0.14	H_T^{had}
	(4j, 4b)	4	4	1.4×10^{-2}	0.17	H_T^{had}
	(5j, 2b)	5	2	8.4×10^{-4}	0.19	H_T^{had}
	(5j, 3b)	5	3	4.0×10^{-3}	0.30	H_T^{had}
	(5j, $\geq 4b$)	5	≥ 4	2.5×10^{-2}	0.40	NN
	($\geq 6j$, 2b)	≥ 6	2	2.4×10^{-3}	0.39	H_T^{had}
	($\geq 6j$, 3b)	≥ 6	3	9.7×10^{-3}	0.63	NN
	($\geq 6j$, $\geq 4b$)	≥ 6	≥ 4	4.0×10^{-2}	0.82	NN
dilepton	(2j, 2b)	2	2	1.0×10^{-4}	0.013	H_T^{all}
	(3j, 2b)	3	2	5.6×10^{-4}	0.055	H_T^{all}
	(3j, 3b)	3	3	7.1×10^{-3}	0.13	H_T^{all}
	($\geq 4j$, 2b)	≥ 4	2	2.7×10^{-3}	0.20	H_T^{all}
	($\geq 4j$, 3b)	≥ 4	3	1.8×10^{-2}	0.40	NN
	($\geq 4j$, $\geq 4b$)	≥ 4	≥ 4	6.7×10^{-2}	0.43	NN

A hypothesis test for the presence of a signal predicted by the SM was performed on a profile likelihood, which takes into account the systematic uncertainties as nuisance parameters to be fitted to the data using the distributions in the fifteen analysis regions listed in Table 4.1. This

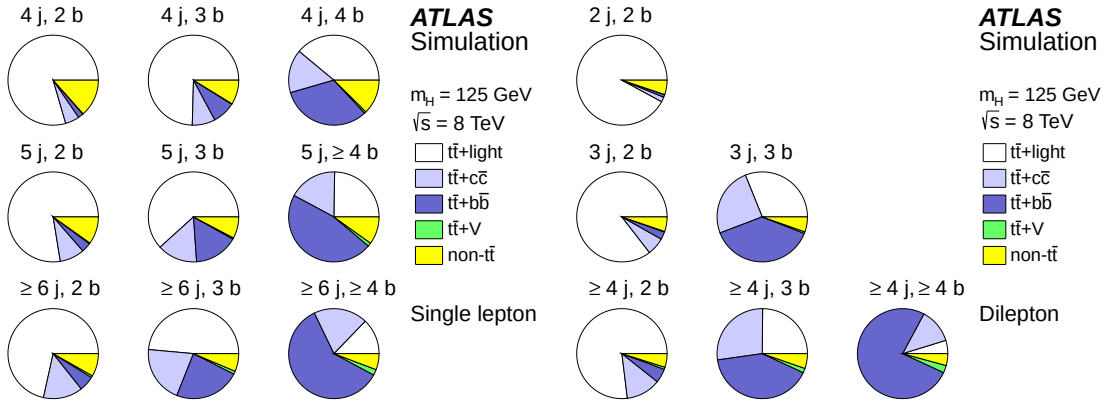
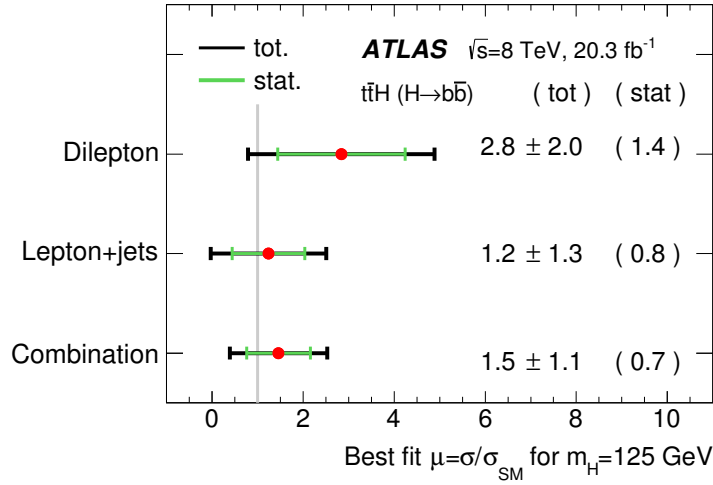


FIGURE 4.1: Background compositions in regions used in Run 1 [78].

procedure allowed to reduce the impact of systematic uncertainties on the sensitivity by taking advantage of high-statistics background-dominated control distributions included in the likelihood fit. Fits were performed under the signal-plus-background hypothesis, and results were represented as the ratio of cross sections, signal strength $\mu_{t\bar{t}H} = \sigma_{t\bar{t}H}^{\text{obs,exp}} / \sigma_{t\bar{t}H}^{\text{SM}}$ as shown in Figure 4.2. The Run 1 result indicated $\sim 60\%$ of the uncertainty on $\mu_{t\bar{t}H}$ is from the systematic uncertainty. Figure 4.3 shows the fifteen largest systematic uncertainties affecting the signal strength where uncertainties in $t\bar{t}$ modelings especially for $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ are ranked predominantly.

FIGURE 4.2: Summary of signal strength measurements of the $t\bar{t}H$ production in Run 1 shown for individual channels and in combination [78].

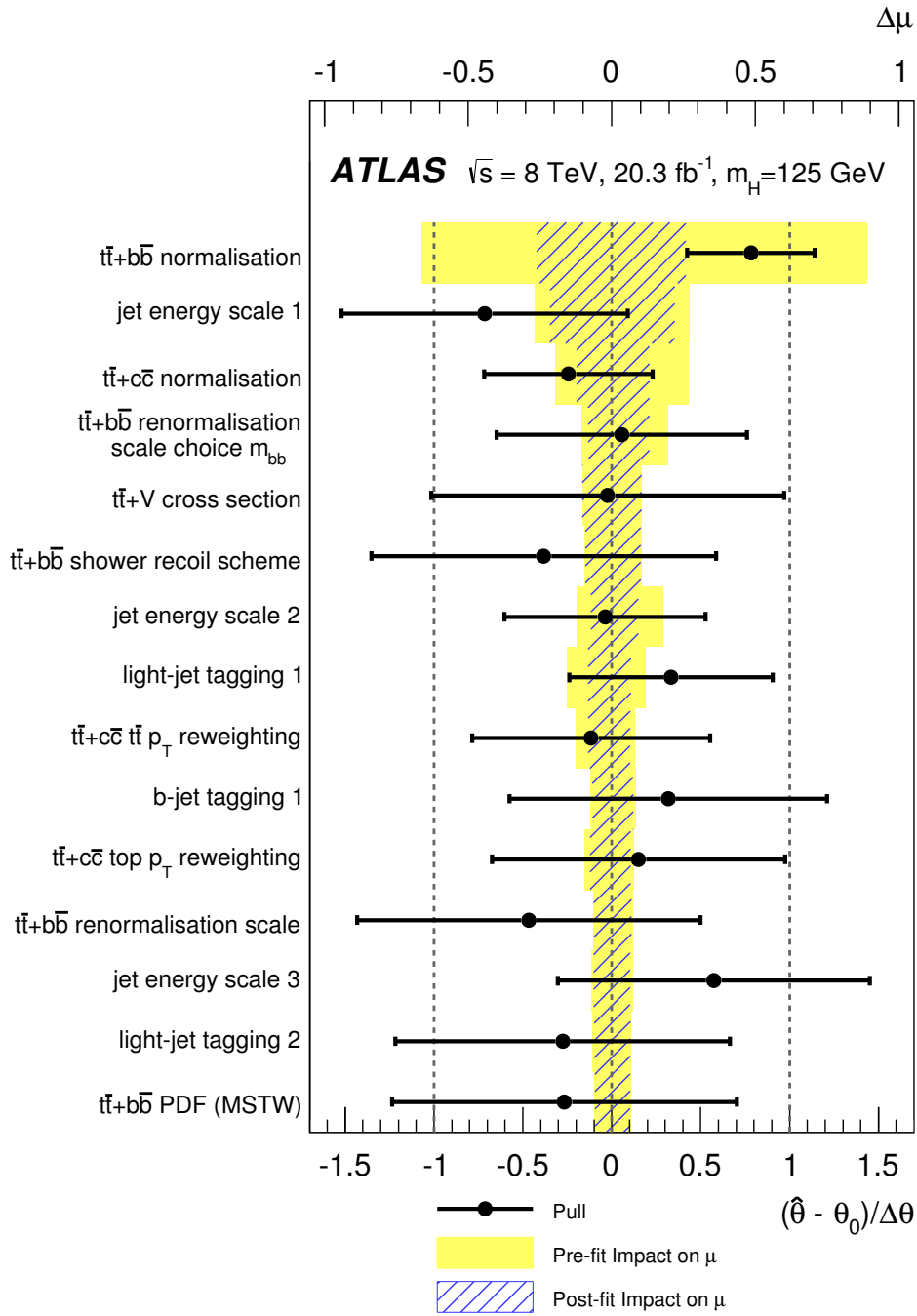


FIGURE 4.3: Ranking of the nuisance parameters used in the fit according to the impact on the signal strength ($\mu_{t\bar{t}H}$) measurement uncertainty in Run 1 [78]. Top fifteen parameters are shown. Yellow band corresponds to the impact before the global fit ("pre-fit") and dashed blue is the impact after the global fit ("post-fit") on $\mu_{t\bar{t}H}$ (top coordinate). The "pulls" (bottom coordinate) are shifts of the nuisance parameters after the fit divided by the uncertainty.

Chapter 5

Analysis Overview

The $t\bar{t}H$ production cross section is $\sim 1\%$ of the total Higgs boson production cross section. This becomes 3.8 times larger at $\sqrt{s} = 13$ TeV in Run 2 than $\sqrt{s} = 8$ TeV in Run 1, while the production cross sections for other background sources such as $t\bar{t}$ and $t\bar{t}Z$ is 3.3 and 3.5 times larger at $\sqrt{s} = 13$ TeV than $\sqrt{s} = 8$ TeV, respectively.

As discussed in the previous chapter, the $t\bar{t}$ +jets background is dominant in this analysis, and their modeling uncertainties limit the $t\bar{t}H(H \rightarrow b\bar{b})$ sensitivity. The $t\bar{t}$ +jets modeling is difficult to be derived using data due to their extremely similar kinematics to the $t\bar{t}H$ process. For the further sensitivity, a signal region is required where reducible $t\bar{t}$ +jets events are excluded as much as possible. In the signal regions, uncorrelated various distributions with more separating powers from the irreducible background source $t\bar{t} + b\bar{b}$ process are required. They form the one best discriminant by the multi-variate analysis. In addition, good controls for the $t\bar{t}$ +jets are also required. Each control region is required to be as much dominant as possible by $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$, or $t\bar{t}$ +light process. The accurate $t\bar{t}$ MC sample is also required. Considering these points, the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis is performed by following procedures:

- Accurate MC samples with enough statistics are prepared for various physics processes. The QCD fake lepton events are estimated using data. (Chapters 6 and 7)
- Signal and control regions are defined by categorizing events according to their jet and the b -tagged jets with various multiplicities and working points. (Chapter 8)
- Utilizing two reconstruction algorithms; one is for full reconstruction of the $t\bar{t}H$ system using Higgs boson kinematics as well as $t\bar{t}$ kinematics, while the other uses only $t\bar{t}$ kinematics and provides non-biased Higgs kinematics. (Chapter 9)
- Adopting two strong discriminant variables for separating the $t\bar{t}H$ signal and $t\bar{t} + b\bar{b}$ background using a likelihood and matrix-element methods. (Chapter 9)
- Performing multi-variate analysis to obtain the best sensitivity from two reconstruction information and adopting new discriminants as well as event kinematics. (Chapter 10)
- A global fit which simultaneously determines the contributions from the $t\bar{t}H$ signal and the major backgrounds. (Chapter 12)

Chapter 6

Monte-Carlo Simulation

All considerable physics processes are simulated by a Monte-Carlo (MC) method with matrix-element generators and parton shower calculations using the full ATLAS detector simulation [80] based on GEANT4 [81].

In the LHC pp collisions at energy scale $O(\text{TeV})$, a good physics simulation requires a good description of the QCD processes as well as the proton structure functions which describe the momentum composition of three valence quarks as well as sea quarks and gluons surrounding them.

Due to the asymptotic freedom feature of QCD, partons can be regarded as free particles if they have large momentum and correspondingly in short time scale. QCD hard scattering interactions and their behaviors are described using perturbation. The generator models all possible Feynman diagrams for certain physics interaction using initial parton information up to certain limited orders. The MC simulates randomly initial/final state particles according to their probabilities. The hadronization evolution of colored partons into stable hadrons occurring at low energy scale $O(\text{MeV})$ is also needed to be considered where non-perturbative QCD can handle the phenomena. Afterwards, the MC also simulates the interactions of generated particles with the full ATLAS detector material converting them into detector responses in order to process the simulation data same as for the real data. The set of particles before the detector simulations is called truth-partons, which includes the exact identities and decay chains of elementary-particles and hadrons not biased by the detector limitation.

6.1 Event Generator

The event simulation begins with the collision with large momentum transfer between two partons coming from protons. The partons behave as asymptotically free and are described by a perturbative QCD. The cross section for general physics process $pp \rightarrow X$ is defined according to the factorization theory [83] as

$$\sigma_{pp \rightarrow X} = \sigma_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(x_a p_a, x_b p_b, \mu_R^2, \mu_F^2), \quad (6.1)$$

where a and b are possible partons to be summed over for $ab \rightarrow X$ processes, and a parton distribution function (PDF), $f_i(x_i, \mu_F^2)$ represents the probability of parton i having momentum fraction (x_i) inside the proton when protons are probed at a scale μ_F . The factorization scale μ_F is an estimate separating adoption of the perturbative and non-perturbative QCD. At the energy scale μ_F , the perturbative calculation fails to be converged due to the phenomena such as UV divergence. The cross section calculation depends on the choice of the renormalization scale μ_R of QCD, at which α_s is evaluated. μ_F and μ_R are usually set to the quantities characterizing the modeled process such as the mass of the heaviest particle and the sum of the momenta of generated particles.

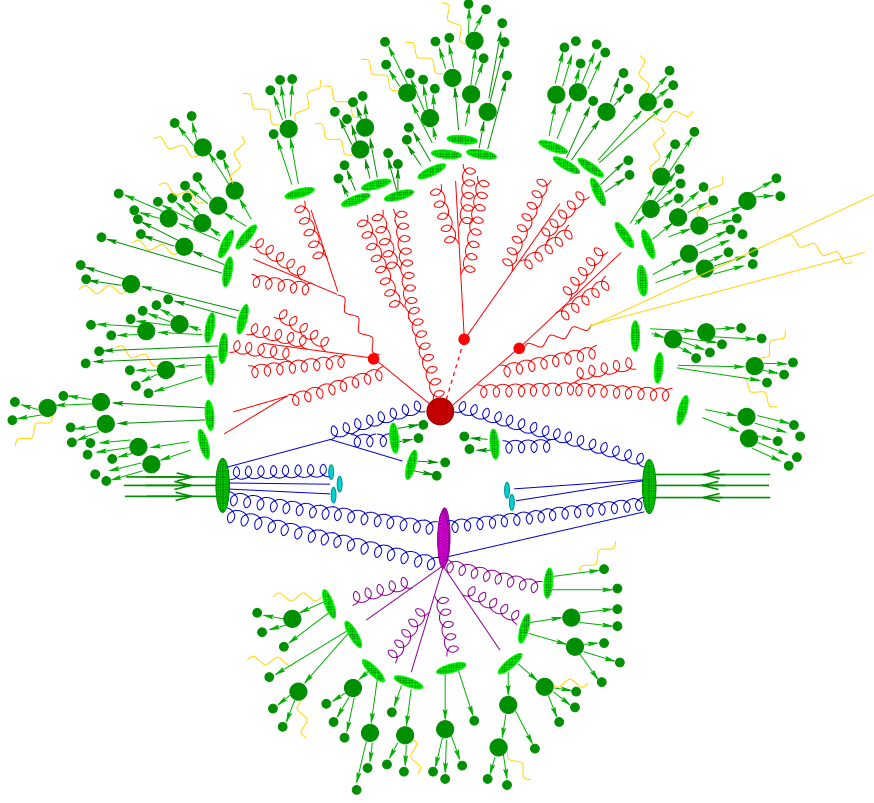


FIGURE 6.1: Sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator and parton shower [82]. The red blob in the center represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation.

The $\hat{\sigma}_{ab \rightarrow X}$ can be expressed in perturbation theory as:

$$\hat{\sigma}_{ab \rightarrow X} = \sum_{k=0}^{\infty} \int d\Phi_{X+k} \left| \sum_{l=0}^{\infty} \mathcal{M}_{X+k}^{(l)} \right|^2 \quad (6.2)$$

where $\mathcal{M}_{X+k}^{(l)}$ is the matrix element of $ab \rightarrow X$ with k additional parton emissions and with l additional loops. The phase space with k additional parton emissions is represented as Φ_{X+k} . The fixed order to calculate the cross section is defined by $k + l$:

- $k = 0, l = 0$: leading order (LO), which is the production with minimum vertices
- $k = n, l = 0$: LO for X plus n partons production
- $k + n, n$: (next-) n LO (N^n LO), which is the production with N^{n-1} LO of X plus one parton, with N^{n-2} LO for X plus two partons, ... , and with LO for X plus n partons

6.2 Parton Distribution Function

The PDFs are measured experimentally at specific energy scales. Then the PDFs are extrapolated to the energy scale in concern using Altarelli-Parisi (AP) equations. There are several sets of PDFs,

depending on which experimental data are used and the Q^2 -evolution scheme in the AP equations.

PDF sets CTEQ-10 (CT10) [84], NNPDF 3.0 [85] are mainly used in the $t\bar{t}H$ analysis. Typical PDF distributions are shown in Figure 6.2 as used in NNPDF 3.1.

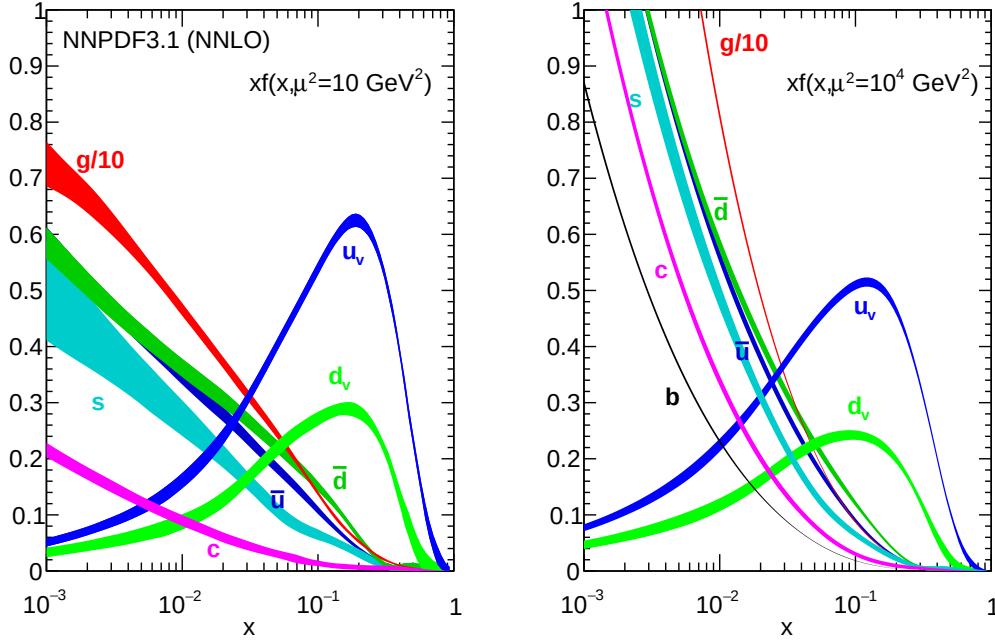


FIGURE 6.2: The NNPDF3.1 NNLO PDFs, evaluated at $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10^4 \text{ GeV}^2$ (right) [86].

6.3 Parton Shower and Hadronization

Partons involved in a hard-scatter process normally have $Q > 1 \text{ GeV}$ corresponding to $\alpha_S < 1$. Quarks and gluons involved in the interaction radiate off gluons carrying energies and colors. Emitted gluons then split into further gluons or quark-antiquark pairs, which is called parton shower (PS). This process continues until all generated quarks break up into hadrons. Emissions produced by PS in collinear divergent or soft (IR divergent) regions are treated by an all-order resummation using an approximation scheme with a leading-log accuracy. Final-state radiation (FSR), a gluon radiated off from a final-state parton, is generated through the PS. For the initial-state radiation (ISR), the momenta of the partons undergoing the hard process need to be precisely adjusted. The momentum of the incoming partons is simulated first. After that, the momentum and angle of the ISR are simulated by backwards evolution. Due to the color connections, the direction of the ISR tends to be aligned to that of incoming parton.

6.4 MC Samples for $t\bar{t}H(H \rightarrow b\bar{b})$ Analysis

All considerable physics processes contributing in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis are simulated by the MC method, as summarized in Table 6.1.

TABLE 6.1: Samples used in the analysis. POWHEG-BOX, MG5_AMC and SHERPA calculate the matrix element at NLO accuracy in QCD, while MADGRAPH calculates at LO accuracy in QCD. The PDFs with "4F" refer to four flavor (u, d, c and s quarks) PDF. The "Tune" column refers to the underlying-event (UE) tuned parameters of the PS.

Process	Generator	PDF _{Gen}	PS	PDF _{ps}	Tune	Normalization
$t\bar{t}H$	MG5_AMC	NNPDF3.0NLO	PYTHIA8	NNPDF2.3LO	A14	NNLO
WtH	MG5_AMC	CT10	HERWIG++	CTEQ6L1	UEEE5	NLO
tH	MADGRAPH	CT10/4F	PYTHIA8	CTEQ6L1	A14	NLO
$t\bar{t}V$	MG5_AMC	NNPDF3.0NLO	PYTHIA8	NNPDF2.3LO	A14	NLO
$t\bar{t}$	POWHEG-BOX v2	NNPDF3.0NLO	PYTHIA8	NNPDF2.3LO	A14	NNLO+NNLL
Wt	POWHEG-BOX v1	CT10	PYTHIA6	CTEQ6L1	P2012	aNNLO
single-top(t -ch)	POWHEG-BOX v1	CT10/4F	PYTHIA6	CTEQ6L1	P2012	aNNLO
single-top(s -ch)	POWHEG-BOX v1	CT10	PYTHIA6	CTEQ6L1	P2012	aNNLO
WtZ	MG5_AMC	NNPDF3.0NLO	PYTHIA8	NNPDF2.3LO	A14	NLO
tZ	MADGRAPH	CTEQ6L1/4F	PYTHIA6	CTEQ6L1	P2012	NLO
$t\bar{t}WW$	MADGRAPH	NNPDF2.3LO/4F	PYTHIA8	NNPDF2.3LO	A14	NLO
$t\bar{t}t\bar{t}$	MADGRAPH	NNPDF2.3LO	PYTHIA8	NNPDF2.3LO	A14	NLO
W +jets	SHERPA	CT10	SHERPA	CTEQ6L1	SHERPA	NNLO
Z +jets	SHERPA	CT10	SHERPA	CTEQ6L1	SHERPA	NNLO
VV	SHERPA	CT10	SHERPA	CTEQ6L1	SHERPA	NLO

For modeling the uncertainties on $t\bar{t}H$, $t\bar{t}$, $t\bar{t}V$, Wt and single-top (t -channel), MC samples with different generators, PS models and radiation parameters were prepared. The differences are implemented as systematic uncertainties in Chapter 11. To realize enough statistics but rapid calculations, a fast simulation of the detector response [65] (AFII) is adopted instead of the full ATLAS detector simulation (FS) in the systematic samples other than $t\bar{t}H$.

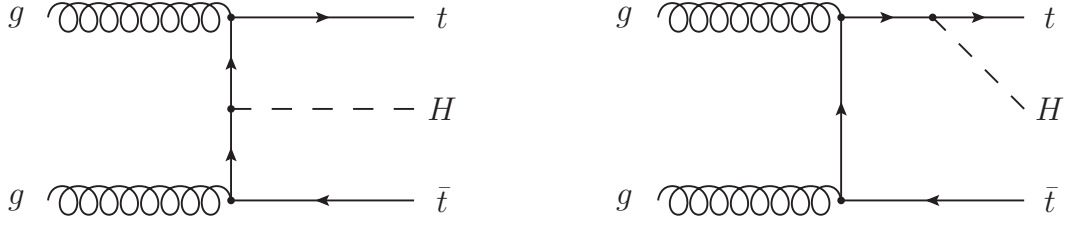
Minor background sources such as single-top(s -channel), WtZ , tZ , $t\bar{t}WW$ and $t\bar{t}t\bar{t}$ are merged together as the "OtherTop". The background sources without top quarks (VV and V +jets) are also merged together.

6.4.1 $t\bar{t}H$

Signal $t\bar{t}H$ processes (Figure 6.3) were produced using MG5_AMC [87] with NLO matrix elements and NNPDF3.0NLO [85] PDF set which is interfaced to PYTHIA8 [88] with A14 NNPDF23LO UE tune [89]. Factorization (μ_F) and renormalization (μ_R) scales for MG5_aMC@NLO generator were set to $\mu_F = \mu_R = H_T/2$, where H_T is the transverse mass of the event defined as the sum of $\sqrt{p_T^2 + m^2}$ over all particles. The Higgs-boson mass is set to 125 GeV where its decay rate is calculated. Alternative PS was examined to estimate the systematic variation using HERWIG++ [90] instead of PYTHIA8.

6.4.2 WtH and tH

These samples have not been discovered and should be negligible in the SM prediction in this analysis, but are included as the background sources. Higgs boson produced in association with single-top and W boson was produced with MG5_AMC interfaced to HERWIG++ with CTEQ6L1 PDF

FIGURE 6.3: Feynman diagrams for $t\bar{t}H$ production.

set. Higgs boson produced in association with single-top and additional partons was produced with MADGRAPH (LO generator) interfaced to PYTHIA8 with CT10 [91, 92] PDF set. The Feynman diagrams are shown in Figure 6.4.

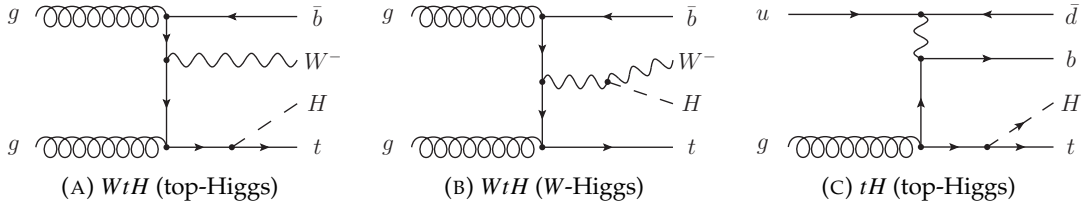
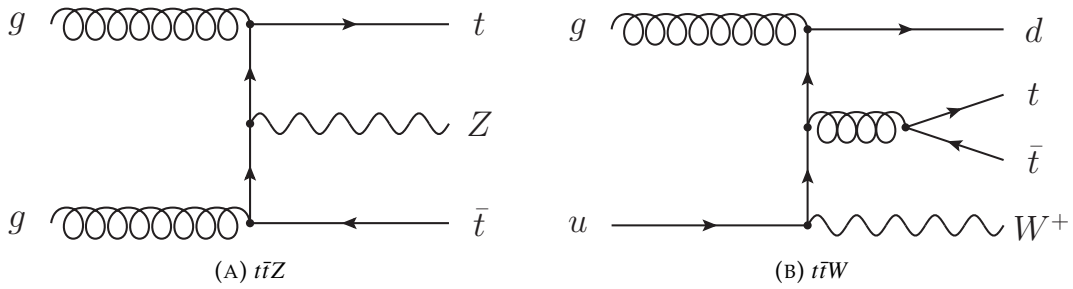


FIGURE 6.4: Feynman diagrams for WtH and tH productions. The WtH production has two kinds of couplings with the Higgs boson: top-Higgs in (A) and W-Higgs in (B). The tH Feynman diagram shown in (C) includes top-Higgs coupling only.

6.4.3 $t\bar{t}V$

$t\bar{t}V$ ($t\bar{t}W$ and $t\bar{t}Z$ shown in Figure 6.5) used MG5_aMC@NLO with NLO matrix elements and NNPDF3.0NLO PDF set interfaced to PYTHIA8 with A14 NNPDF23LO UE tune, same as for $t\bar{t}H$ production. The systematic variation was evaluated by alternative samples generated with SHERPA as the different choices of the generator and PS.

FIGURE 6.5: Feynman diagrams for $t\bar{t}Z$ and $t\bar{t}W$ productions.

6.4.4 $t\bar{t}+\text{jets}$ and $t\bar{t}+\geq 1b$ Reweighting

The $t\bar{t}$ process was generated using POWHEG-BOX [93, 94, 95, 96] v2 with NLO matrix elements and NNPDF3.0NLO PDF. The h_{dump} parameter [96] which regulates high- p_T emissions against $t\bar{t}$,

was set to 1.5 times the top quark mass at $m_{\text{top}} = 172.5$ GeV. PYTHIA8 with A14 NNPDF23LO UE tune was interfaced to the generator.

The yield is normalized with a QCD NNLO prediction by top++2.0 [97] including NNLL resummation for soft gluon terms [98, 98, 99, 100, 101], resulting in a cross section of 832^{+46}_{-51} pb.

The generated NLO $t\bar{t}$ sample is well modeled such as in the top-quark p_T distribution, showing less than a few percent difference between NLO and NNLO predictions.

To model $t\bar{t}$ +jets events more precisely, they were categorized using flavors of additional truth jets not from $t\bar{t}$ decay, requiring $p_T > 15$ GeV and $|\eta| < 2.5$. Those jets were matched to B -/ C -hadrons with $p_T > 5$ GeV within $\Delta R < 0.3$. Using this information, the events were categorized with following selections:

- $t\bar{t} + b\bar{b}$: two truth jets each matched to a B -hadron
- $t\bar{t} + b$: only one truth jet matched to a B -hadron (the other is missed)
- $t\bar{t} + \geq 3b$: at least three of truth jets each matched to a B -hadron
- $t\bar{t} + B$: at least one truth jet, each matched to more than one B -hadron
- $t\bar{t} + c\bar{c}$: at least two truth jets each matched to a C -hadron
- $t\bar{t} + c$: only one truth jet matched to a C -hadron
- $t\bar{t} + C$: at least one truth jet each matched to more than one C -hadrons
- $t\bar{t}$ +light : other events

The samples including $t\bar{t} + b\bar{b}$, $t\bar{t} + \geq 3b$, $t\bar{t} + b$ and $t\bar{t} + B$ are inclusively referred as $t\bar{t} + \geq 1b$ as well as those for $t\bar{t} + c\bar{c}$, $t\bar{t} + C$ and $t\bar{t} + c$ referred as $t\bar{t} + \geq 1c$. Furthermore, the events having $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ are inclusively referred as $t\bar{t}$ +HF. Feynman diagrams for $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t}$ +light are shown in Figure 6.6.

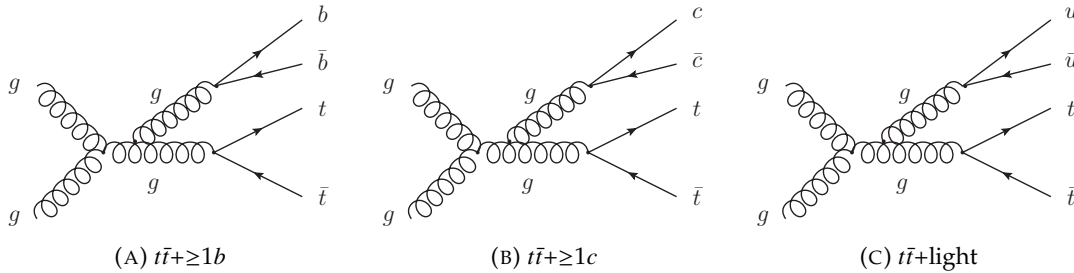


FIGURE 6.6: Feynman diagrams for $t\bar{t}$ +jets. $t\bar{t}$ +jets are categorized by the additional flavor emitted from the gluon, $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t}$ +light.

For improvement of $t\bar{t} + \geq 1b$ modeling, which is the largest background source in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, $t\bar{t} + \geq 1b$ events were reweighted to full $t\bar{t} + b\bar{b}$ NLO production using SHERPA+OPENLOOPS (SHERPAOL) [82, 102, 103] with four flavor-scheme (4F) for the CT10 PDF set, while the nominal POWHEG+PYTHIA8 predicts additional $b\bar{b}$ pairs at LO via the PS. SHERPAOL is a fully automated generator for tree and one loop amplitudes of $2 \rightarrow 2$ to 5 SM QCD and electro-weak interactions at NLO based on SHERPA2.2.0. The renormalization scale μ_R was set to the CMMSPS value $\mu_{\text{CMMSPS}} = \prod_{i=t\bar{t}b\bar{b}} E_{T,i}^{1/4}$ and the factorization scale was set to $H_T/2 = \frac{1}{2} \sum_{i=t\bar{t}b\bar{b}} E_{T,i}$. The resummation scale μ_Q is set to $H_T/2$. The NLO QCD reduces the predicted scale uncertainties to 20 ~ 30%

from 70 ~ 80% at LO. In addition, $t\bar{t} + b\bar{b}$ produced with second additional $b\bar{b}$ pair at LO is also predicted better than with POWHEG+PYTHIA8; significant differences are observed in the cross section value including such processes. SHERPAOL using 4F PDF does not include b -quarks in the PDF, but describes b -quarks as massive particles of 4.75 GeV. This can describe the kinematics of b -quarks more precisely including collinear gluon splittings into a $b\bar{b}$ pair, for example, $b\bar{b}$ distribution with very small ΔR such as in $t\bar{t} + B$ events and very soft gluon splittings contributing in $t\bar{t} + b$ (+ one missing b -jet) events. Although 4F PDF cannot describe initial state radiations $g \rightarrow b\bar{b}$, such cross sections are sub-dominant in $t\bar{t} + b\bar{b}$ processes [103].

The $t\bar{t} + \geq 1b$ reweighting is performed by scaling the relative contributions of the different sub-categories of $t\bar{t} + \geq 1b$ ($t\bar{t} + b\bar{b}$, $t\bar{t} + B$, $t\bar{t} + b$ and $t\bar{t} + \geq 3b$) obtained from POWHEG+PYTHIA8 to SHERPAOL (4F), as shown in Figure 6.7.

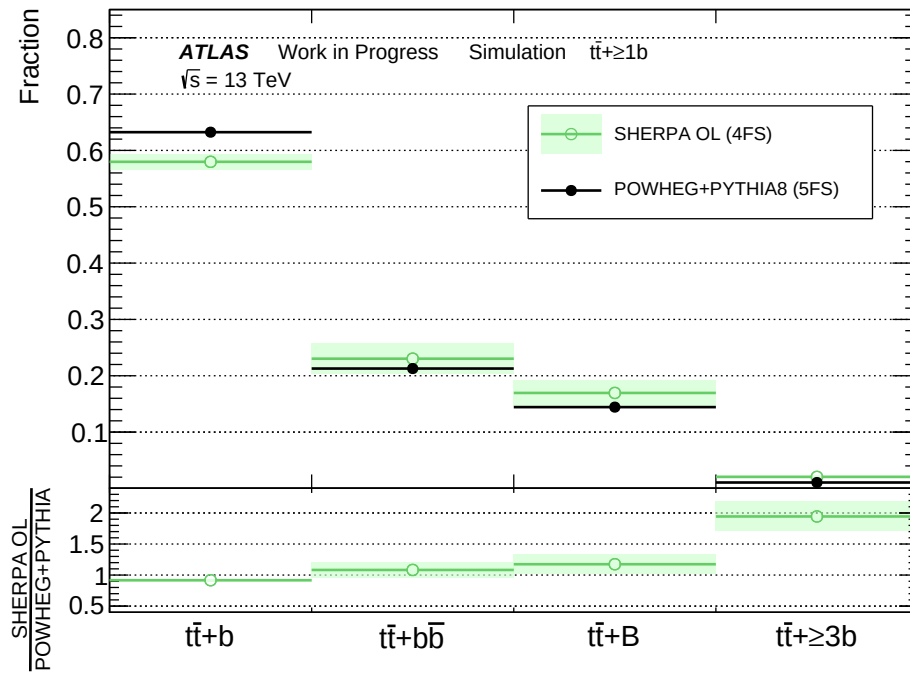


FIGURE 6.7: The fractions of the $t\bar{t} + b$, $t\bar{t} + b\bar{b}$, $t\bar{t} + B$ and $t\bar{t} + \geq 3b$ sub-categories on the truth-level. The $t\bar{t} + b\bar{b}$ inclusive generation by POWHEG+PYTHIA8 (5F) is compared to the dedicated generation of $t\bar{t} + b\bar{b}$ by SHERPAOL (4F). The uncertainties of the SHERPAOL are described in Chapter 11 (green shaded area). The fractions of the nominal POWHEG+PYTHIA8 are to be reweighted to SHERPAOL.

Alternative samples with varying the MC parameters were also prepared for the study of systematic variations as summarized in Table 6.2. Two samples denoted as radHi and radLo were produced with the same MC setups as the nominal but with varied radiation parameters to model the initial and final state radiations. Setting the renormalization and factorization scales by a factor of 1/2 and twice the h_{damp} parameter was to increase the radiation process (radHi), while the variation set them with opposite directions was to suppress radiation (radLo). Another sample was generated to model the PS using HERWIG7 instead of PYTHIA 8. SHERPA+OPENLOOPS (5F) was produced as a variation of the generator and the PS. Its radiation parametrization was also different from the nominal. SHERPA+OPENLOOPS (4F) was also produced to cover the difference between 5F and 4F in the PDF. The reweighting to the SHERPA+OPENLOOPS 2.1 4F was performed in the truth-level, and differences not fully reweighted were treated as the systematic variation.

TABLE 6.2: Summary of the settings used for the simulation of the $t\bar{t}$ samples. For the renormalization and factorization scales, $m_{T,t} = \sqrt{m_t^2 + p_{T,t}^2}$ ($m_{T,\bar{t}} = \sqrt{m_{\bar{t}}^2 + p_{T,\bar{t}}^2}$) which is the transverse mass of the top (anti-top) quark, for $p_{T,t}$ ($p_{T,\bar{t}}$) p_T of the top (anti-top) quark in the $t\bar{t}$ center-of-mass reference frame. For the SHERPAOL $t\bar{t} + b\bar{b}$ sample in the last column, $\mu_{\text{CMMPs}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$ and $H_T/2 = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$.

Generator PS	POWHEG PYTHIA 8	POWHEG PYTHIA 8	POWHEG PYTHIA 8	POWHEG HERWIG7	SHERPA 2.2.1	SHERPA 2.1
Ren. scale	$m_{T,t}$	$\frac{1}{2}m_{T,t}$	$2m_{T,t}$	$m_{T,t}$	$\sqrt{\frac{m_{T,t}^2 + m_{T,\bar{t}}^2}{2}}$	μ_{CMMPs}
Fact. scale	$m_{T,t}$	$\frac{1}{2}m_{T,t}$	$2m_{T,t}$	$m_{T,t}$	$\sqrt{\frac{m_{T,t}^2 + m_{T,\bar{t}}^2}{2}}$	$H_T/2$
h_{damp}	$1.5m_t$	$3m_t$	$1.5m_t$	$1.5m_t$	–	–
Tune	A14	A14 Var3c up	A14 Var3c down	H7-UE-MMHT	Author's tune	Author's tune
PDF Flavor	5	5	5	5	5	4
variation	nominal	radHi	radLo	PS	full modeling	flavor scheme

6.4.5 Wt and single-top

Wt and single-top (t- and s-channel) samples, shown in Figure 6.8, were generated through POWHEG-BOX v1 with NLO matrix elements using CT10 5F PDF set for Wt and single-top (s-channel), and CT10 4F PDF set for single-top (t-channel) interfaced to PYTHIA6 with Perugia 2012 (P2012) UE tune [104]. Wt events have an overlap with $t\bar{t}$ events. To subtract the overlap, the diagram removal scheme [105] was taken as the nominal.

Alternative samples of Wt were generated using HERWIG++ instead of PYTHIA6 as the PS modeling variation as well as different radiation parameters to model the initial and final state radiations denoted as radHi and radLo. Another overlap removal method using the diagram subtraction scheme was also prepared as a systematic sample for Wt .

Alternative sample was prepared also for single-top t-channel for variation of the PS modeling as well as the radiation parameters.

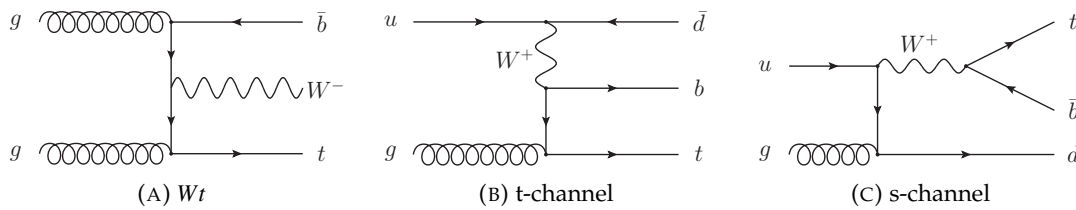


FIGURE 6.8: Feynman diagrams for single-top production.

6.4.6 Other background

Other background productions have very small contributions in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis. The Feynman diagrams are shown in Figure 6.9.

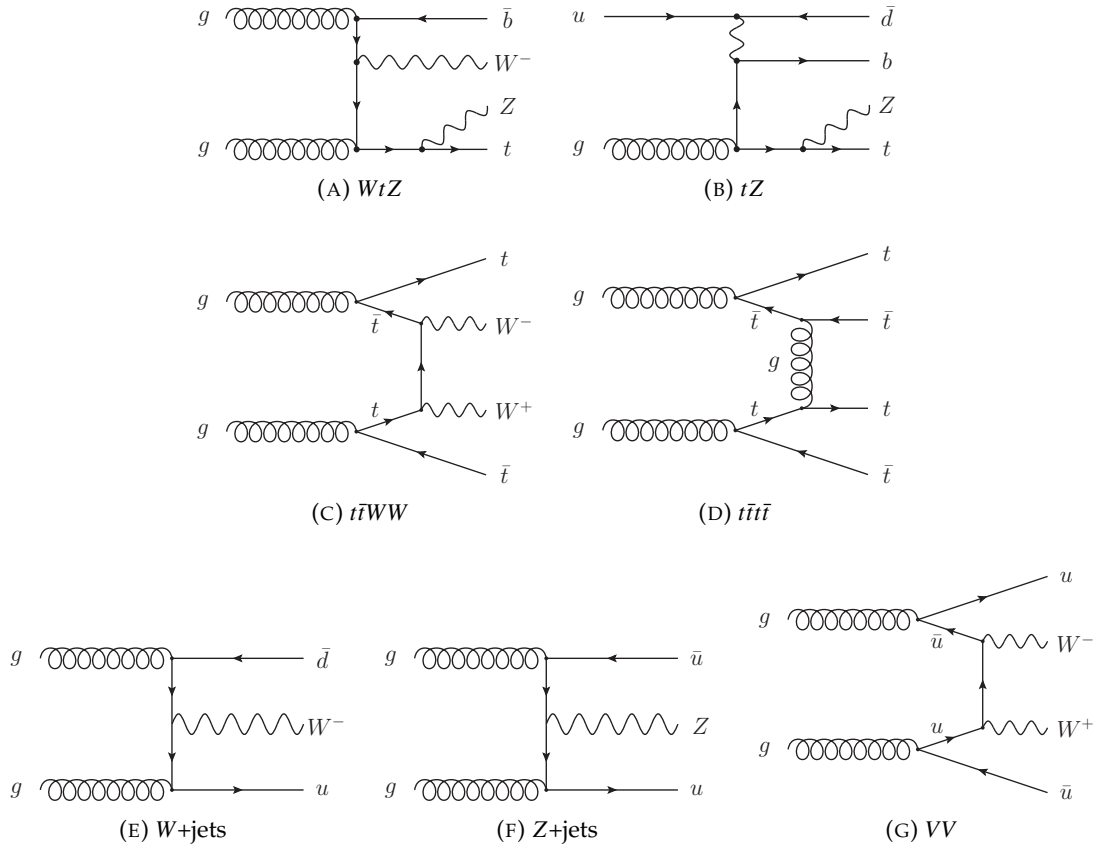


FIGURE 6.9: Feynman diagrams for minor background sources.

6.4.7 Fake-Leptons

Fake-lepton sources were estimated in different methods in both channels. All MC events with jet passing the final lepton offline selections were rejected.

In the single-lepton channel, fake-lepton sources were estimated using the data as described in Chapter 7. The dilepton channel regarded the rejected MC events as a fake-lepton source. If either one of two leptons is non-prompt lepton, the event was also rejected from the nominal background source and treated as another fake-lepton source.

Chapter 7

Data Driven Background Estimate

7.1 Fake-Leptons

Fake-lepton background where fake-leptons are from non-prompt leptons and jet misidentified as leptons, is one of the reducible background sources.

The main fake-lepton source is the multi-jet production with the fake lepton from misidentified jet, photon conversion, decay product of heavy flavor quark such as $b \rightarrow c + \ell + \bar{\nu}_\ell$ which are accidentally associated to the primary vertex. This source has a non-prompt lepton relatively near jets, with small E_T^{miss} and m_T^W distribution with no W mass peak. However, in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, E_T^{miss} and m_T^W cuts were not applied in event selections not to lose signal events, and this increased the contribution from the fake-lepton especially in the control regions. The multi-jet production is not modeled well in the simulation, and significantly low selection efficiency makes it difficult to simulate this production with enough statistics. Therefore, data driven techniques were adopted to estimate this production with the Matrix method.

To take a fake-lepton enriched dataset, looser lepton selections were applied:

- Electron identification: MediumLH (TightLH for nominal)
- Muon identification: Medium (same as the nominal)
- Isolation cut: No requirements

The Matrix method calculates the number of fake-leptons in the nominal selections using the efficiency for loose and nominal datasets for the fake-lepton (denoted as "fake") and other background sources with a real prompt-lepton (denoted as "real"). Number of events after nominal or loose selection can be described as $N_{\text{nominal/loose}} = N_{\text{nominal/loose}}^{\text{fake}} + N_{\text{nominal/loose}}^{\text{real}}$, where the number to be estimated is $N_{\text{nominal}}^{\text{fake}}$. Using the real and fake efficiencies, f_{real} and f_{fake} , N_{nominal} is described as

$$N_{\text{nominal}} = N_{\text{nominal}}^{\text{fake}} + N_{\text{nominal}}^{\text{real}} \quad (7.1)$$

$$= f_{\text{fake}} N_{\text{loose}}^{\text{fake}} + f_{\text{real}} N_{\text{loose}}^{\text{real}} \quad (7.2)$$

Then,

$$N_{\text{nominal}}^{\text{fake}} = f_{\text{fake}} N_{\text{loose}}^{\text{fake}} = \frac{f_{\text{fake}}}{f_{\text{real}} - f_{\text{fake}}} (f_{\text{real}} N_{\text{loose}} - N_{\text{nominal}}) \quad (7.3)$$

This can be obtained counting the events passing loose criterion with a weight of

$$w = \frac{f_{\text{fake}}}{f_{\text{real}} - f_{\text{fake}}} (f_{\text{real}} - \theta) \quad (7.4)$$

where θ is 1 if the event passes the nominal criterion otherwise it becomes 0.

For a measurement of the real efficiency (f_{real}), a tag-and-probe method was utilized for the $Z \rightarrow \ell\ell$ events. Events having two leptons with same flavors, opposite signs and $60 < m_{\ell\ell} < 120$ GeV were required to pass the loose selection. Events were also required to have at least two jets to keep statistics. Afterwards, one of the two leptons was required to be the lepton passing the nominal lepton criterion, denoted as tagged lepton. The efficiency was evaluated to see whether the other lepton (probe lepton) passes the nominal criterion or not, $f_{\text{real}} = (\text{number of probe leptons passing nominal criterion})/(\text{number of probe leptons})$.

The fake efficiency is measured as

$$f_{\text{fake}} = \frac{N_{\text{nominal}}^{\text{data}} - N_{\text{nominal}}^{\text{MC}}}{N_{\text{loose}}^{\text{data}} - N_{\text{loose}}^{\text{MC}}} \quad (7.5)$$

where the differences between the data and MC are the expected number of events of fake-lepton sources. Used MC samples are $t\bar{t}$, single-top, $W(Z)$ +jets, VV with an assumption that the simulation samples correctly describe the data to an acceptable precision. To obtain fake-lepton enriched events, following selections are applied for the fake efficiency measurements:

- $E_{\text{T}}^{\text{miss}} < 20$ GeV for the electron channel
- $|d_0| > 5$ mm for the muon channel

Required number of jets is also at least two to keep sufficient statistics.

The magnitudes of the fake-lepton contribution are estimated for leptons (e or μ) and the data taking years (2015 or 2016), individually. The efficiencies were measured as functions of lepton p_{T} , η , largest jet p_{T} and minimum ΔR between lepton and jet in addition to inclusive efficiencies. All results are shown in Figures 7.1 to 7.4.

Obtained different efficiency values were combined into one fake or real efficiency as:

$$f_{\text{real/fake}} = f_{\text{real/fake}}^{\text{av}} \cdot \prod_v \frac{f_{\text{real/fake}}^v}{f_{\text{real/fake}}^{\text{av}}} \quad (7.6)$$

where $f_{\text{real/fake}}^{\text{av}}$ is the inclusive efficiency and $f_{\text{real/fake}}^v$ is the efficiency for the variable v .

7.2 V+jets

V +jets productions are not well described by SHERPA generator especially when they are generated with additional heavy-flavor jets. Therefore, V +jets data driven estimate was performed.

Scale factors for Z +jets were investigated in the dilepton channel with ≥ 2 jets and without the dilepton mass selection. Z +jets MC sample was categorized by additional two jet flavors:

- $Z_{4\text{HF}}$: Z boson produced with at least four heavy-flavor jets
- $Z_{3\text{HF}}$: Z boson produced with three heavy-flavor jets
- $Z_{2\text{HF}}$: Z boson produced with two heavy-flavor jets
- $Z_{\text{HF,light}}$: Z boson produced with one heavy-flavor jets
- Z_{light} : Z boson produced with no heavy-flavor jets

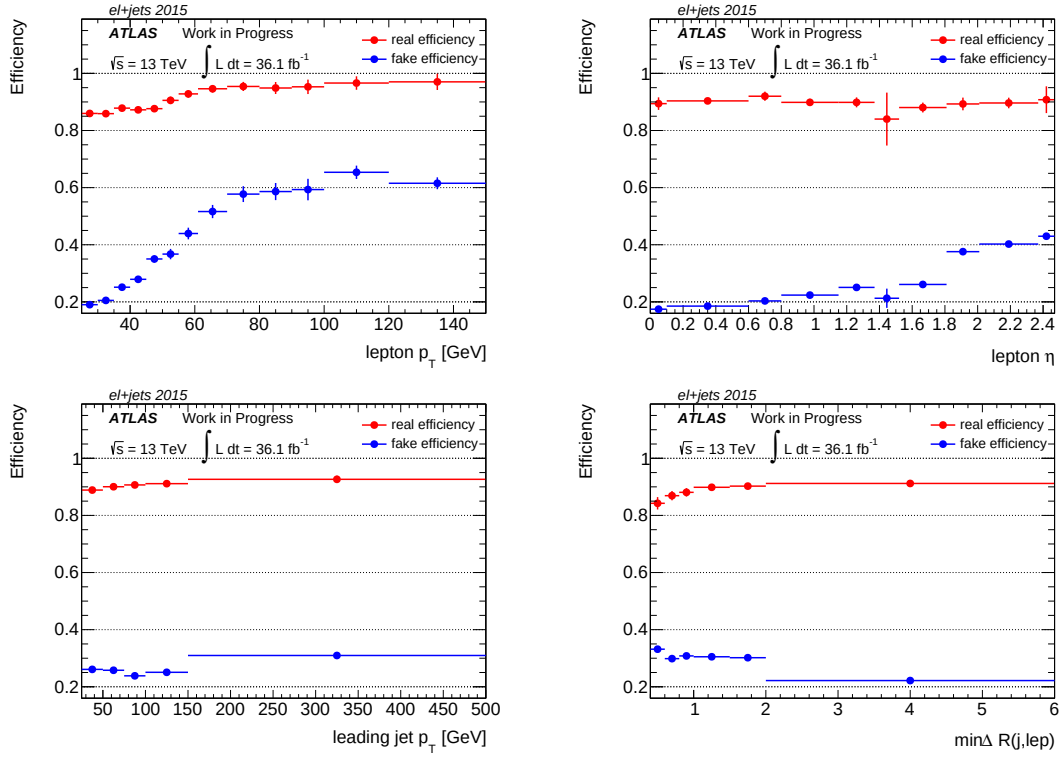


FIGURE 7.1: Efficiencies for fake-lepton estimation in the electron channel and 2015 runs shown as a function of lepton p_T , η , largest jet p_T and minimum ΔR between the lepton and jet.

Scale factors were determined in the five categories individually. Events were divided into five by the number of b -tagged jets at 77% WP, $n\text{BTags}_{77\%} = 0, 1, 2, 3, \geq 4$. To determine the scale factors, a fit was done where all Z +jets scale factors are free parameters, as well as the normalization of all other MC samples grouped together which is decorrelated among fit categories. The dilepton mass distributions were used in the fit to simultaneously allow the Z mass peak to drive the Z +jets scale factors and the side-bands to drive the normalizations of the other samples in five categories. The normalization of other samples remains close to 1 for $n\text{BTags}_{77\%} \leq 2$ categories and increases for categories with more b -tagged jets, which come from the mis-modeling of the $t\bar{t}$ +HF. As the result, one scale factor applied to the $Z_{\geq 1\text{HF}}$ samples was $1.3 \pm 0.4 (\pm 35\%)$. All Z +jets scale factors are compatible with this estimate.

It is difficult to take W +jets dominant region, therefore, larger normalization uncertainties were assigned instead of deriving scale factors.

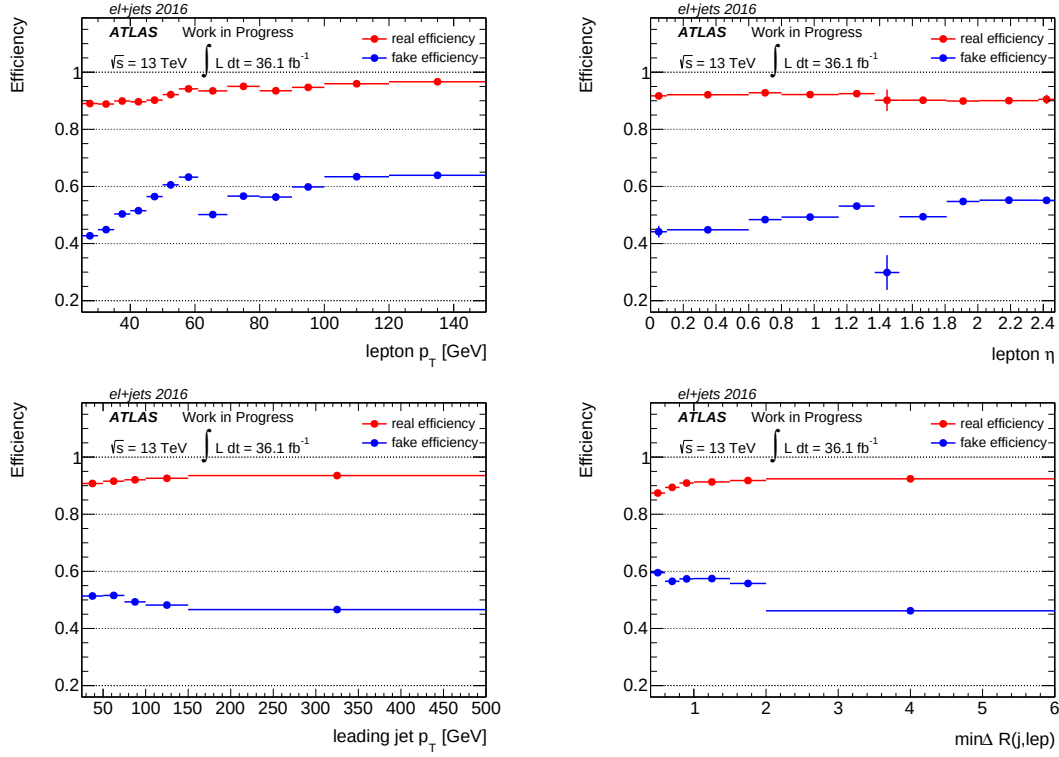


FIGURE 7.2: Efficiencies for fake-lepton estimation in the electron channel and 2016 shown as a function of lepton p_T , η , largest jet p_T and minimum ΔR between the lepton and jet.

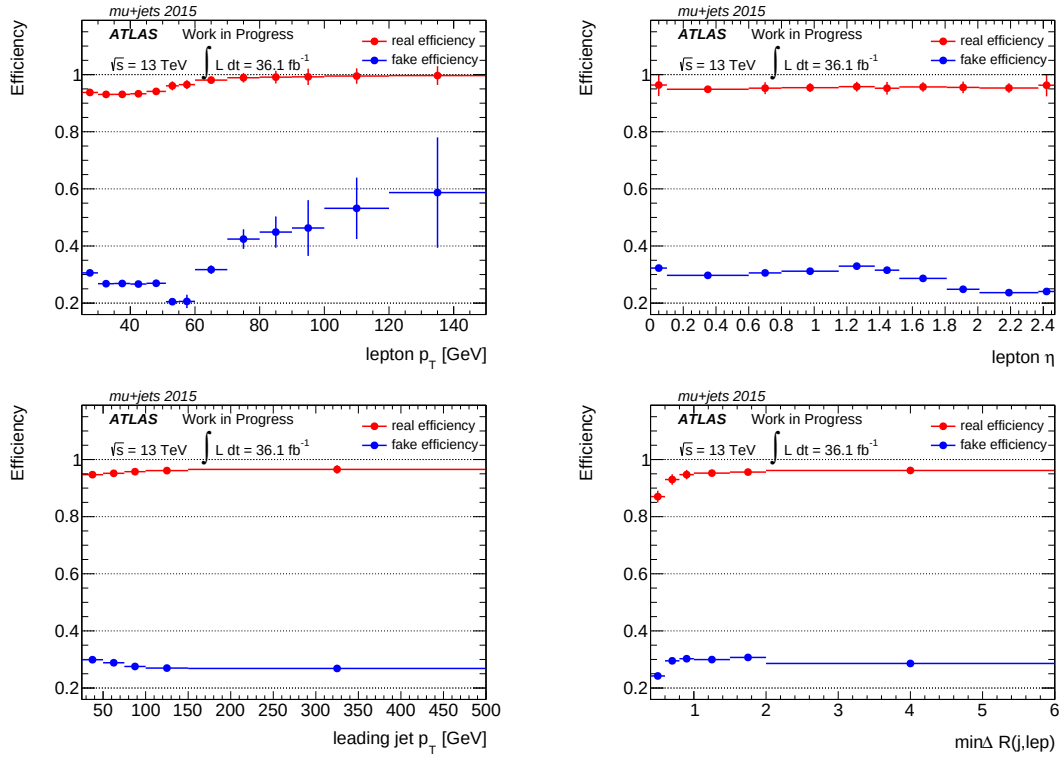


FIGURE 7.3: Efficiencies for fake-lepton estimation in the muon channel and 2015 runs shown as a function of lepton p_T , η , largest jet p_T and minimum ΔR between the lepton and jet.

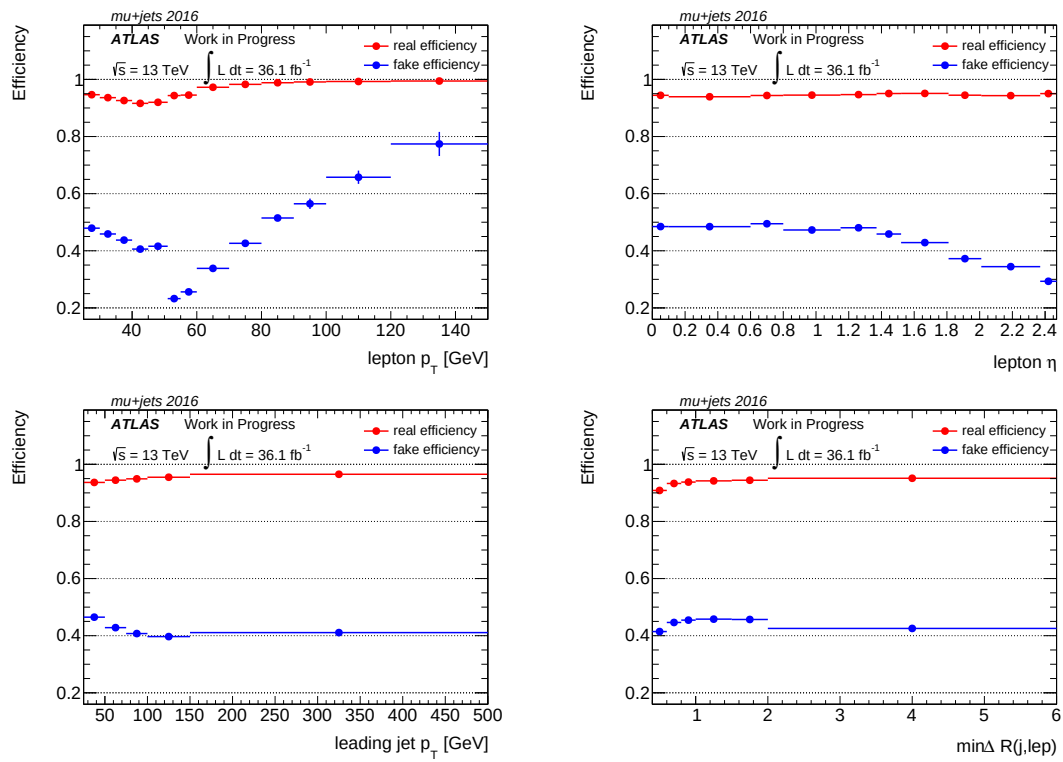


FIGURE 7.4: Efficiencies for fake-lepton estimation in the muon channel and 2016 runs shown as a function of lepton p_T , η , largest jet p_T and minimum ΔR between the lepton and jet.

Chapter 8

Region Definition

We define the signal regions where the $t\bar{t}H$ events are enriched, and control regions where backgrounds are evaluated by comparing with MC simulations. The signal and control regions are categorized based on the number of jets (nJets) and number of b -jets. The $t\bar{t}H$ contains six jets including four b -jets in the single-lepton channel if all partons are correctly and separately reconstructed. However in the case where some jets are missed due to p_T smaller than 25 GeV and/or $|\eta|$ higher than 2.5, full reconstruction of the $t\bar{t}H$ system cannot be performed and discriminant variables should be optimized also for the events including some missing objects. For example, $q2_{\text{had}W}$ is often missed due to low p_T as shown in Table 3.6. Therefore, region definitions were performed in different nJets regions: 4, 5, and ≥ 6 jets for the single-lepton channel, 3 and ≥ 4 jets in the dilepton channel respectively. They are denoted as 4 j, 5 j, and ≥ 6 j regions for the single-lepton channel, 3 j and ≥ 4 j regions for the dilepton channel.

The main background sources in the $t\bar{t}H$ analysis are from $t\bar{t}$ +jets processes. They are separated into three components depending on the flavor of additional jets, $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$, and $t\bar{t}$ +light. In categorizing with b -tagging information, the flavors of $t\bar{t}$ +jets distribute with different fractions in the signal and control regions affecting the constraints on their systematic uncertainties. There is little separation between $t\bar{t}H$ signal and $t\bar{t}+\geq 1b$ because they have the same final state particles. Therefore, the region definition is crucial to control the reducible background sources such as $t\bar{t}+\geq 1c$ and $t\bar{t}$ +light to reject as much as possible while keeping high signal efficiency in the signal regions. The region definition was studied for the events for the single-lepton channel with ≥ 6 jets. Afterwards, the best definition was also adopted similarly in other regions such as the 5j region in the single-lepton channel, and the $\geq 4j$ and 3j regions in the dilepton channel. In this section, the number of $t\bar{t}H$ events is assumed to be the SM expectation value and $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalizations are also given by the MC. Possible deviations have little effects in the region definition.

To calculate the expected sensitivity of the $t\bar{t}H$ signal, the global fits were performed in all options described in Chapter 12. Three parameters (the signal strength described as the ratio of the number of $t\bar{t}H$ signal events to the SM cross section expectation $\mu_{t\bar{t}H}$, and normalization corrections for $t\bar{t}+\geq 1b$ events $k_{t\bar{t}+\geq 1b}$ and $t\bar{t}+\geq 1c$ events $k_{t\bar{t}+\geq 1c}$) were fitted with all distributions fixed to the SM prediction. The mean values of the three parameters and all systematic nuisance parameters were fixed to 1 in the expected fits. Then, uncertainties on these parameters were compared as the sensitivities. Distributions of defined regions used in the fits were determined as followings:

- Normalization (one bin distribution) if $t\bar{t}$ +light is more than 60%. The regions dominated by $t\bar{t}$ +light process has phase spaces far away from the signal regions, and tends to have a large number of events. To avoid too tight constraints from these regions, any binned distributions are not adopted.

- ClassBDT distribution if the S/B is larger than 1.5%. The region is sensitive to determine the number of signal events, and treated as the signal region. The detailed explanation of the ClassBDT is given in Chapter 10. If the pseudo-continuous b -tagging is used for systematic variations, the ClassBDT for inclusive ≥ 6 jets is utilized as the distribution. On the other hand, if the systematic variations for a single WP is used, the other ClassBDT for inclusive ≥ 6 jets without any tag weight bins and LHD is trained for the distribution.
- Otherwise, H_T^{had} distribution can constrain systematic uncertainties related to jet calibrations.

As a reference, a simple definition was adopted at first by utilizing a single b -tagging WP. (Available WPs are shown in Table 3.5.) In the Run 1 analysis, the region definition was performed with a single b -tagging WP corresponding to a 70% tagging efficiency for b -jets with a light-jet mistagging rate of $\leq 1\%$ [78], where the b -tagged jets were defined as the jet having larger b -tagging score than the WP threshold. Nine analysis regions were categorized using the number of b -tagged jets $n\text{Tags}_{70\%} = (2, 3 \text{ or } \geq 4)$ in each of $n\text{Jets} = (4, 5 \text{ or } \geq 6)$ regions. The simple definition here also adopted the 70% WP which performs well balancing the b -jet purity and non b -jets rejection. The fitted sensitivities are:

$$\mu_{t\bar{t}H} = 1.00^{+1.11}_{-1.14}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.21}_{-0.19}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.48}_{-0.41}. \quad (8.1)$$

This sensitivity improvement from the Run 1 analysis shown in Figure 4.2 is the result of statistic gain in this analysis because of the increased center of mass energy.

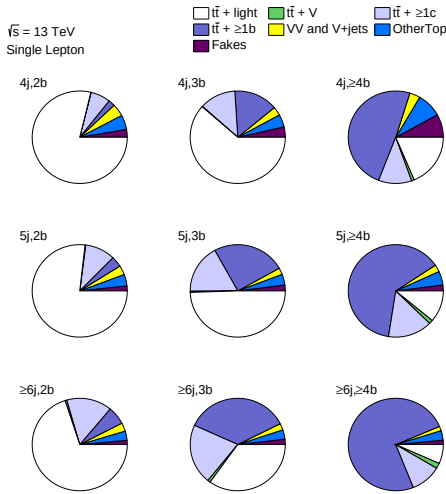


FIGURE 8.1: Background fractions in the nine analysis regions defined simply with single b -tagging WP at 70%.

TABLE 8.1: S/B and $S/\sqrt{B}$ in the nine analysis regions defined simply with single b -tagging WP at 70%. The "distribution" column shows the variable used in the fit. The ClassBDT for regions is trained without any tag weight bins.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
4j,2b	3.0×10^{-4}	0.22	Normalization	1
4j,3b	2.2×10^{-3}	0.38	Normalization	1
4j,≥4b	2.0×10^{-2}	0.41	H_T^{had}	3
5j,2b	7.6×10^{-4}	0.44	Normalization	1
5j,3b	5.2×10^{-3}	0.83	H_T^{had}	8
5j,≥4b	3.1×10^{-2}	1.0	ClassBDT	8
≥6j,2b	2.3×10^{-3}	1.1	Normalization	1
≥6j,3b	1.2×10^{-2}	2.0	ClassBDT	8
≥6j,≥4b	4.1×10^{-2}	2.3	ClassBDT	8

As shown in Figure 8.1, the fractions of $t\bar{t}+\geq 1c$ events in the most sensitive signal region ($\geq 6j, \geq 4b$) are large, and $t\bar{t}H$ signal events are not extracted reliably due to large uncertainties in both $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ modelings. For a better control of reducible background sources and optimizing further the signal regions with a help of various S/B levels, multiple b -tagging working points were adopted. All four WPs were implemented where the b -tagging score was binned into five according to the WP thresholds. This is called pseudo-continuous b -tagging method.

Fully continuous b -tagging score and binned score distributions are shown in Figure 8.2. In the five binned distribution, each bin called "tag weight bin" contains the b -tagging scores exclusively. For example, the tag weight bin "70" contains jets having b -tagging score exceeding the 70% WP threshold but below the 60% WP threshold. The four largest tag weight bins are chosen per event for the region definition. Ideally we retain four b -jets in $t\bar{t}H$ and $t\bar{t} + b\bar{b}$ events, three b -jets and one non b -jet in $t\bar{t} + b$ and $t\bar{t} + B$ events, and two b -jets and two non b -jets in $t\bar{t} + c\bar{c}$ and $t\bar{t}$ +light events. The b -tagged jets are used to classify separate categories ($B_{j1}, B_{j2}, B_{j3}, B_{j4}$), where B_{ji} corresponds to the tag weight bin of the i -th jet in the order of the b -tagging score.

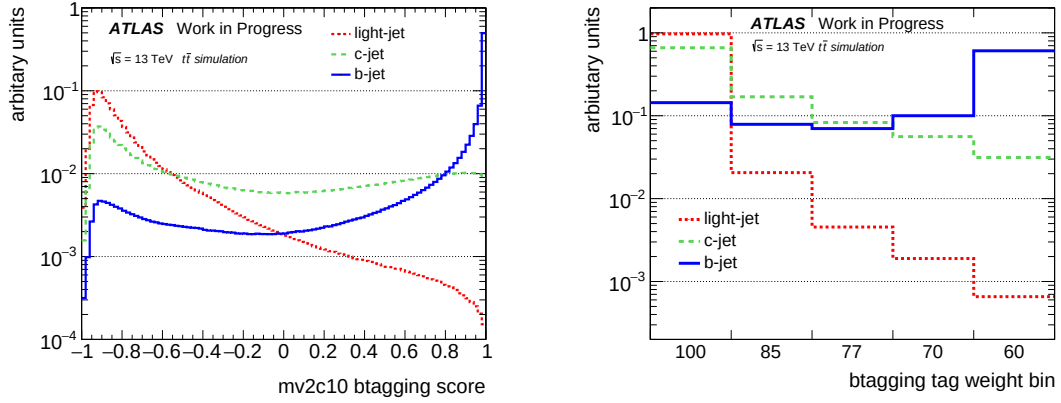


FIGURE 8.2: Distributions of (left) MV2c10 b -tagging scores and (right) b -tagging tag weight bins for b -, c - or light-jets. The areas are normalized to 1.

8.1 Impact of Pseudo-Continuous b -tagging Systematics

Before using the multiple WPs, the systematic uncertainty effects are checked in this section. To simultaneously calibrate all five bins of the b -tagging score, the five times larger number of systematic variations is assigned when the multiple WPs are used in the analysis. These b -tagging systematic variations are typically larger than the single-WP systematic variations. Therefore, if we do not get a significant gain over the systematic variations, no improvements cannot be expected in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis by utilizing pseudo-continuous b -tagging.

The same calibration sources for b -jet and light-jet are used in the pseudo-continuous and single-WP b -tagging. The b -tagging with a single WP can choose two calibration sources for the c -jet mis-tagging, single-leptonic $t\bar{t}$ or $W+c$ events, while the pseudo-continuous b -tagging has only single-leptonic $t\bar{t}$ source. For the single-leptonic $t\bar{t}$ event set, 4 j regions are used for the calibration, and the corresponding events must be removed from the analysis regions.

The systematic impacts on the sensitivities with various options using b -tagging are summarized in Table 8.2. Comparing the uncertainties of the $t\bar{t} + \geq 1c$ normalization, the $t\bar{t}$ calibration has a better description of the phase space in the signal enriched regions, resulting a smaller uncertainty on the $t\bar{t} + \geq 1c$ normalization. This results the better sensitivity on the $t\bar{t}H$ signal. This tendency is also seen in the systematic constraints by the fit, as shown in Figure 8.3. The c -jet mis-tagging constraints are relaxed, and their mean values are closed to the original values provided by the $t\bar{t}$ calibration. Pseudo-continuous b -tagging calibration has also a better sensitivity than the single WP calibration, because of the precise description of b -jets in the pseudo-continuous b -tagging with fine binned calibration. The constraints of b -tagging systematic uncertainties are relaxed by using the pseudo-continuous b -tagging calibration. Although the number of systematic variations

increases by using the pseudo-continuous b -tagging, most of them are less than 1% in all bins in the analysis regions.

The usage of the pseudo-continuous b -tagging itself provides the better signal sensitivity, and the impacts of their systematic uncertainties are negligibly small.

TABLE 8.2: Impacts on the sensitivities with various b -tagging systematic options. The full regions means the nine regions shown in Figure 8.1. If 4 j regions are removed, the number of regions is reduced to six.

categorization	systematics	c -jet calib.	$\pm\sigma(\mu_{t\bar{t}H})$	$\pm\sigma(k_{t\bar{t}+\geq 1b})$	$\pm\sigma(k_{t\bar{t}+\geq 1c})$
full regions in ≥ 4 jets	70% WP	$W+c$	+1.11 -1.14	+0.21 -0.19	+0.48 -0.41
full regions in ≥ 4 jets	70% WP	$t\bar{t}$	+1.06 -1.09	+0.21 -0.19	+0.38 -0.35
full regions in ≥ 4 jets	PCB	$t\bar{t}$	+0.99 -0.99	+0.21 -0.19	+0.35 -0.33
removed 4 j regions	70% WP	$t\bar{t}$	+1.09 -1.13	+0.24 -0.20	+0.40 -0.36
removed 4 j regions	PCB	$t\bar{t}$	+1.05 -1.06	+0.22 -0.19	+0.37 -0.34

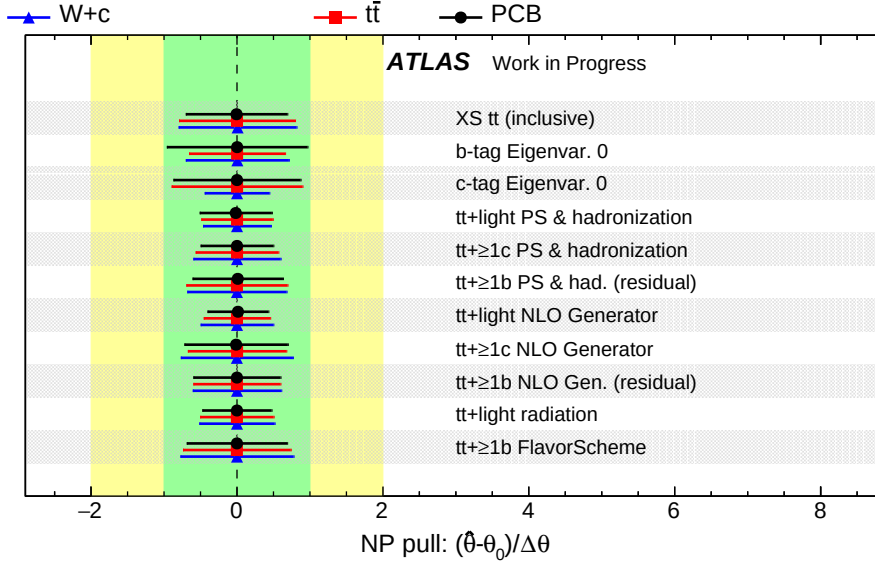


FIGURE 8.3: Highly constrained systematic uncertainties in various b -tagging systematic options. Only systematic uncertainty sources with $> 20\%$ constraints from the original uncertainty band are shown in the plot. Black and red points denote the fit with $W+c$ and $t\bar{t}$ c -jet mis-tagging calibration for a single-WP of 70% b -tagging. Blue points show the fit with $t\bar{t}$ c -jet mis-tagging calibration for the pseudo-continuous b -tagging. The analysis regions are common among the three options, where full nine regions are used in ≥ 4 jets.

8.2 Simple Usage of Pseudo-Continuous b -tagging

As the baseline of the pseudo-continuous b -tagging usage, the simple region definition using two WPs (77% and 60%) was implemented for 5 and ≥ 6 jets. At first, analysis regions were defined

using number of b -tagged jets at loose WP (77%) cumulatively. After that, each region is subdivided into two regions, Hi and Lo. The Hi regions require the same numbers of b -tagged jets at 60% WP instead of 77% WP, while remained events are categorized into Lo regions, as shown in Figure 8.4. The fitted sensitivities are

$$\mu_{t\bar{t}H} = 1.00^{+0.73}_{-0.73}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.12}_{-0.11}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.31}_{-0.28}. \quad (8.2)$$

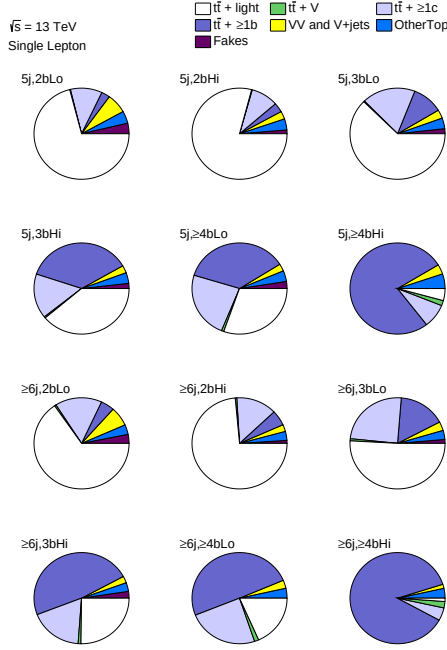


FIGURE 8.4: Background fractions in the twelve analysis regions defined with two b -tagging WPs at 60% and 77%.

TABLE 8.3: S/B and $S/\sqrt{B}$ in the nine analysis regions defined simply with single b -tagging WP at 70%. The "distribution" column shows the variable used in the fit.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
5j,2b Lo	6.4×10^{-4}	0.28	Normalization	1
5j,2b Hi	6.7×10^{-4}	0.31	Normalization	1
5j,3b Lo	2.1×10^{-3}	0.46	H_T^{had}	6
5j,3b Hi	7.4×10^{-3}	0.72	H_T^{had}	6
5j,\ge 4b Lo	1.2×10^{-2}	0.68	ClassBDT	8
5j,\ge 4b Hi	4.5×10^{-2}	0.86	ClassBDT	8
\ge 6j,2b Lo	1.9×10^{-3}	0.69	Normalization	1
\ge 6j,2b Hi	2.0×10^{-3}	0.75	Normalization	1
\ge 6j,3b Lo	5.3×10^{-3}	1.2	H_T^{had}	8
\ge 6j,3b Hi	1.4×10^{-2}	1.5	ClassBDT	8
\ge 6j,\ge 4b Lo	2.2×10^{-2}	1.8	ClassBDT	8
\ge 6j,\ge 4b Hi	5.4×10^{-2}	1.9	ClassBDT	8

With this procedure, the total number of regions increased twice from the single-WP option, part of which have almost no constraints on both of the $t\bar{t}H$ signal strength and $t\bar{t}+\text{HF}$ normalizations. Possible sensitivity degradation by removing some regions is summarized in Table 8.4. If the regions have non-negligible contributions from non- $t\bar{t}$ background sources, and have phase spaces far from the signal enriched regions, they should be removed to avoid unexpected constraints and sensitivity degradations by them. The $\geq 5j,2b$ Lo and $\geq 5j,2b$ Hi regions have almost no impacts on the sensitivities and similar phase spaces. However, these regions have a good control for determining the number of the inclusive $t\bar{t}$ events, and removing both of them from the analysis regions worsen the sensitivities. Therefore, as the pre-selection for the region definition, $\text{nBTags}_{77\%} \geq 3$ or $\text{nBTags}_{60\%} \geq 2$ is required to reject $5j,2b$ Lo and $\geq 6j,2b$ Lo regions but to keep $5j,2b$ Hi and $\geq 6j,2b$ Hi regions.

For an optimization of the region definition with pseudo-continuous b -tagging, the fit using only $\geq 6j$ regions was done, resulting sensitivities as:

$$\mu_{t\bar{t}H} = 1.00^{+0.95}_{-0.95}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.16}_{-0.15}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.38}_{-0.31}. \quad (8.3)$$

TABLE 8.4: Impacts on the sensitivities with removing regions. The baseline is twelve analysis regions defined with two b -tagging WP at 60% and 77% as shown in Figure 8.4.

removed regions	$\pm\sigma(\mu_{t\bar{t}H})$	$\pm\sigma(k_{t\bar{t}+ \geq 1b})$	$\pm\sigma(k_{t\bar{t}+ \geq 1c})$
–	+0.73 –0.73	+0.12 –0.11	+0.31 –0.28
5j,2b Lo and $\geq 6j,2b$ Lo	+0.73 –0.73	+0.12 –0.11	+0.31 –0.29
5j,2b Hi and $\geq 6j,2b$ Hi	+0.73 –0.73	+0.12 –0.11	+0.31 –0.29
5j,3b Lo and $\geq 6j,3b$ Lo	+0.74 –0.74	+0.13 –0.12	+0.39 –0.35
5j,3b Hi and $\geq 6j,3b$ Hi	+0.81 –0.81	+0.18 –0.15	+0.33 –0.31
5j, $\geq 4b$ Lo and $\geq 6j,\geq 4b$ Lo	+0.86 –0.86	+0.13 –0.12	+0.35 –0.32
5j, $\geq 4b$ Hi and $\geq 6j,\geq 4b$ Hi	+0.99 –0.99	+0.14 –0.13	+0.32 –0.30
5j,2b Hi/Lo and $\geq 6j,2b$ Hi/Lo	+0.74 –0.74	+0.13 –0.12	+0.33 –0.30

8.3 Features with Pseudo-Continuous b -tagging

To see the sensitivity of pseudo-continuous b -tagging, the number of $t\bar{t}H$ signals, S/B and background compositions are plotted in Figures 8.5 to 8.7 for all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$. As expected, the combinations with tighter WPs have larger number of the $t\bar{t}H$ signal events and higher S/B . The background sources are almost from $t\bar{t}$ +jets processes in all bins. There are larger fractions from $t\bar{t}+\geq 2b$ (corresponding to $t\bar{t}+b\bar{b}$ and $t\bar{t}+\geq 3b$) in bins with tighter WP combinations, and larger fractions from $t\bar{t}+1b$ (corresponding to $t\bar{t}+b$ and $t\bar{t}+B$) in bins with tighter WP combinations under $B_{j4} = 100$. The loose combination bins have large contributions from $t\bar{t}+\geq 1c$. If at least one of WPs is 100, the fraction of $t\bar{t}$ +light increases. The region definition does not provide a good separation power between $t\bar{t}H$ and $t\bar{t}+\geq 2b$, as shown in Figure 8.8.

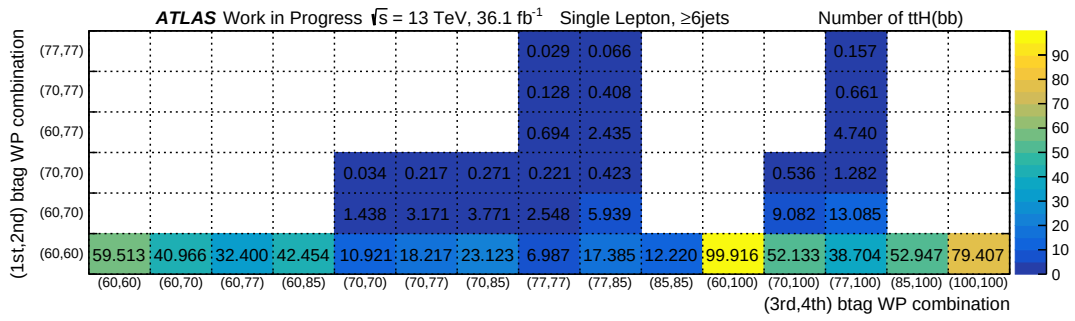
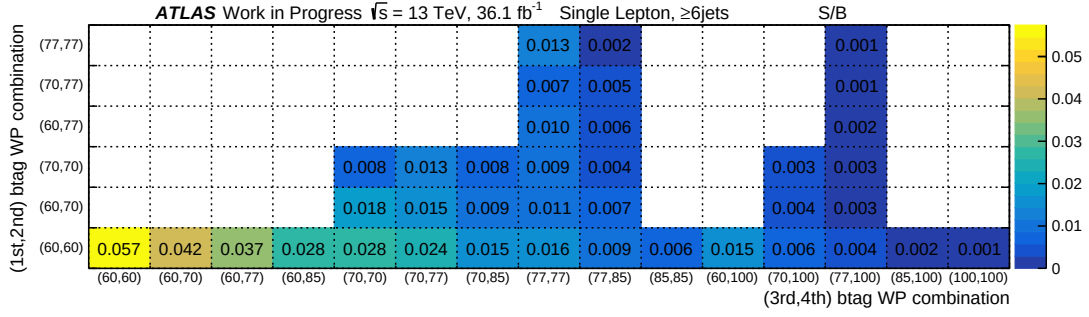
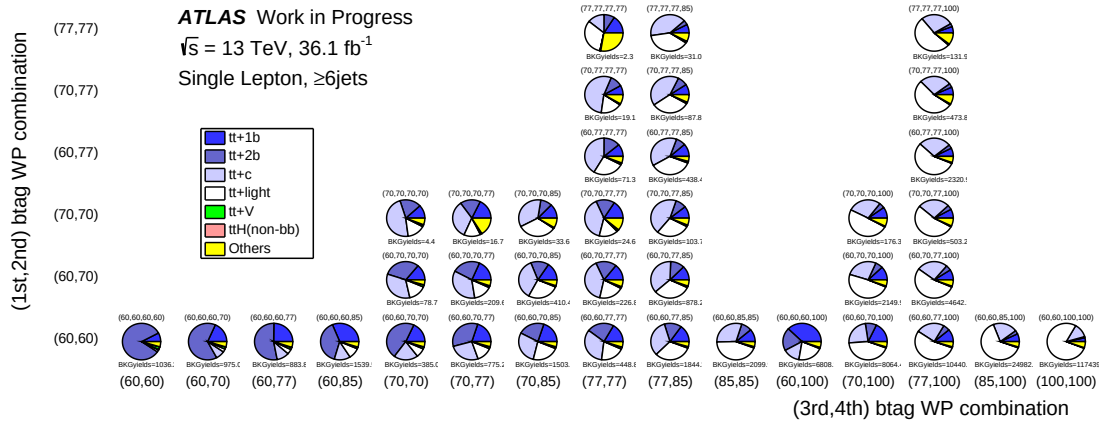
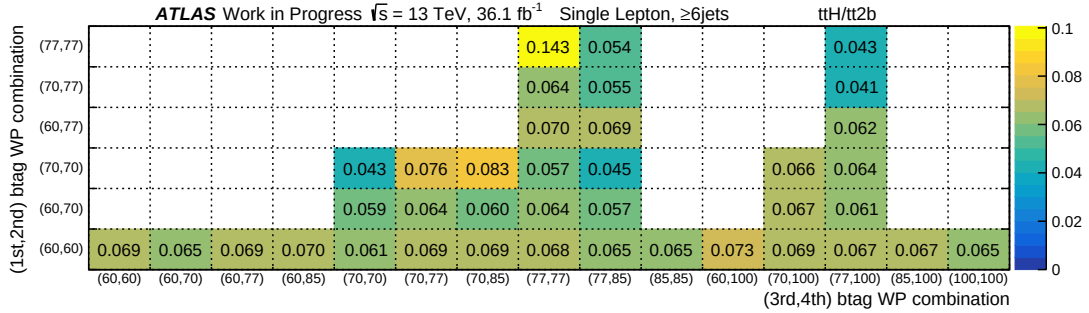


FIGURE 8.5: Number of $t\bar{t}H(H \rightarrow b\bar{b})$ events expected by the SM in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n_{\text{jets}} \geq 6$. The x -axis shows (B_{j3}, B_{j4}) combinations, and the y -axis shows (B_{j1}, B_{j2}) combinations.

Some of bins in Figures 8.5 to 8.7 are insignificant due to very small statistics, similar S/B values or background compositions, therefore, they are merged for simplicity. The region definition in the combination map for the simple option using 60% and 77% WPs is shown in Figure 8.9.

The details of the procedure are given in the following sections.

FIGURE 8.6: S/B ratio in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n_{\text{Jets}} \geq 6$.FIGURE 8.7: Background composition in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n_{\text{Jets}} \geq 6$. $t\bar{t}H \rightarrow b\bar{b}$ background includes $t\bar{t} + b\bar{b}$ and $t\bar{t} + \geq 3b$, and $t\bar{t} + 1b$ includes $t\bar{t} + b$ and $t\bar{t} + B$.FIGURE 8.8: Ratio of $t\bar{t}H(H \rightarrow b\bar{b})$ to $t\bar{t} + \geq 2b$ background process in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$ for $n_{\text{Jets}} \geq 6$.

8.4 S/B Based Categorization

The $t\bar{t}H$ signal separation is optimized adopting the selections best optimized in each of the various S/B range regions. This procedure can define the signal regions using the ClassBDT distributions. Therefore, merging procedure based on S/B was as following:

- 1) choose the bin with maximum S/B
- 2) add other bins in the order of S/B until the recalculated S/B is reduced by 10% from the original S/B

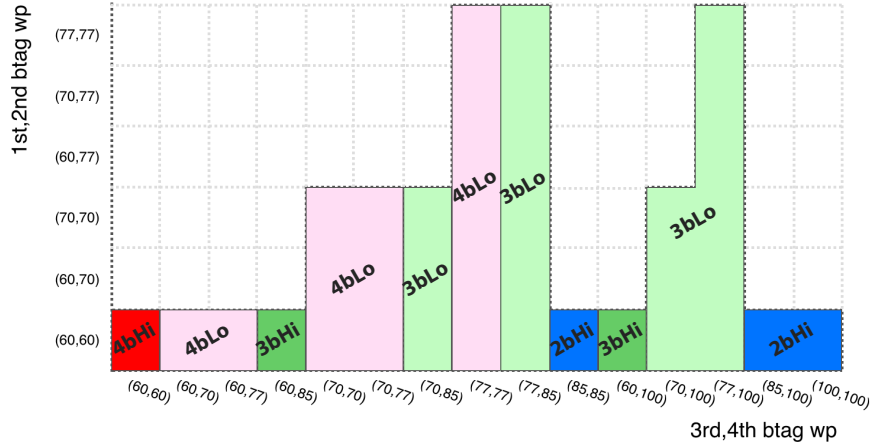


FIGURE 8.9: Region Definition with two b -tagging WPs at 60% and 77% for $n_{\text{Jets}} \geq 6$. The bins with the same numbering are regarded as the same region.

3) repeat until all regions are examined

The tag-weight-bin combination is categorized as shown in Figure 8.10. In total nine regions are defined by this merging. The obtained region performance is shown in Figures 8.11, and Table 8.5. To keep the loss of significance as small as possible, the ClassBDT distribution was utilized in the six regions with $t\bar{t} + \text{light} < 60\%$.

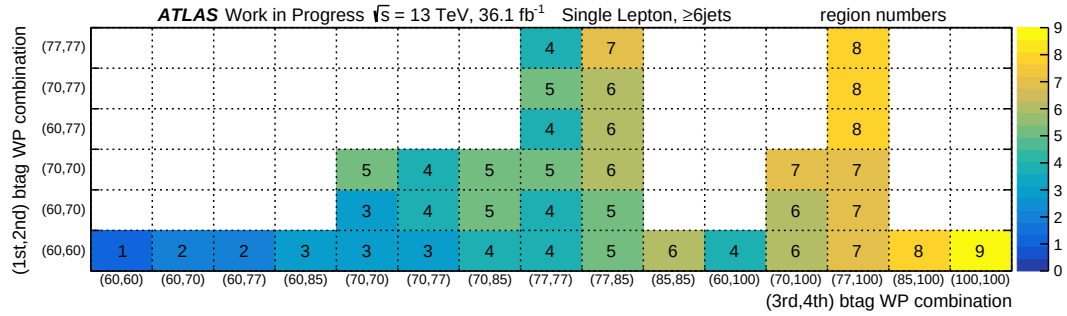


FIGURE 8.10: Region categorization based on S/B . The bins with the same numbering are regarded as the same region.

For the defined six regions using ClassBDT, the optimum number of signal regions was determined by merging these regions one by one from lower S/B regions. The impact on the sensitivities was also checked if three regions using the normalization are merged. To determine the optimum number of regions, the uncertainty on the $\mu_{t\bar{t}H}$ and binned expected significance Z_{exp} are used as the discriminants. The binned expected significance is defined as:

$$Z_{\text{exp}} = \sqrt{2 \sum_{i=\text{bin}} \left((S_i + B_i) \ln \left(1 + \frac{S_i}{B_i} \right) - S_i \right)}, \quad (8.4)$$

where S_i and B_i are the numbers of signal and background events, respectively. The Z_{exp} is reduced to $\sqrt{\sum_{i=\text{bin}} (S_i / \sqrt{B_i})^2}$ if S_i and B_i have large statistics. The result is shown in Table 8.6. Both Z_{exp} and uncertainty on the $\mu_{t\bar{t}H}$ become much worse if $\geq 6j, \text{SR1}$ is merged to other regions. This

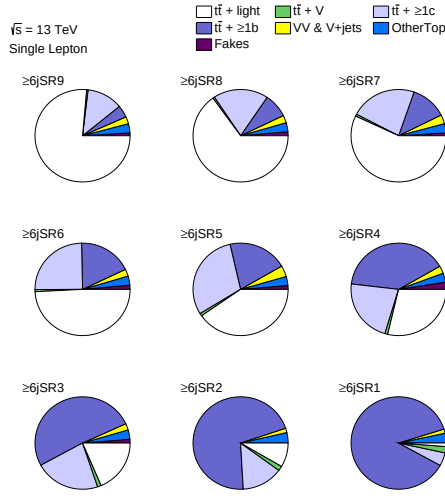


FIGURE 8.11: Background fractions in the nine analysis regions for ≥ 6 jets defined by S/B ordering.

TABLE 8.5: S/B and $S/\sqrt{B}$ in the nine analysis regions defined by S/B ordering. The "distribution" column shows the variable used in the fit.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
$\geq 6j, SR9$	1.6×10^{-3}	0.52	Normalization	1
$\geq 6j, SR8$	2.9×10^{-3}	0.55	Normalization	1
$\geq 6j, SR7$	3.9×10^{-3}	0.56	Normalization	1
$\geq 6j, SR6$	5.8×10^{-3}	0.77	ClassBDT	8
$\geq 6j, SR5$	7.5×10^{-3}	0.57	ClassBDT	8
$\geq 6j, SR4$	1.2×10^{-2}	1.3	ClassBDT	8
$\geq 6j, SR3$	2.1×10^{-2}	1.3	ClassBDT	8
$\geq 6j, SR2$	3.3×10^{-2}	1.6	ClassBDT	8
$\geq 6j, SR1$	5.4×10^{-2}	1.8	ClassBDT	8

region requires the tightest selection as $(B_{j1}, B_{j2}, B_{j3}, B_{j4}) = (60, 60, 60, 60)$, which is the same requirement to $\geq 6j, \geq 4b$ Hi region. This region is called "ultra-pure-region" (UPR). The second and third largest S/B regions ($\geq 6j, SR2$ and $\geq 6j, SR3$) also have significant losses in both Z_{exp} and uncertainty on $\mu_{t\bar{t}H}$ if these regions are merged. From $\geq 6j, SR4$ to $\geq 6j, SR6$, the sensitivities are not different if they are merged together. Therefore, these regions should be used for controls of the background processes. The $\geq 6j, SR1, SR2$, and $SR3$ are better to be kept as separated signal regions and use ClassBDTs as the distributions for a global fit. One merged region of $\geq 6j, SR4, 5, 6$ is good to be used as a control region. Other three regions with $t\bar{t} + \text{light} > 60\%$ also have little impacts on the sensitivities, and they are reasonable to be treated as one merged region.

TABLE 8.6: Uncertainty on $\mu_{t\bar{t}H}$ and binned expected significance Z_{exp} with merging regions defined by S/B ordering. Regions using ClassBDT and normalization are merged separately. Merged regions use same bins and variables before merged.

merged regions	# of regions	# of bins	unc. on $\pm\sigma(\mu_{t\bar{t}H})$	Z_{exp}	$\pm\sigma(k_{t\bar{t}+\geq 1b})$	$\pm\sigma(k_{t\bar{t}+\geq 1c})$
(No merging)	9	51	+0.94 -0.91	4.44	+0.17 -0.15	+0.40 -0.33
$\geq 6j, SR7, 8, 9$	7	49	+0.96 -0.95	4.43	+0.19 -0.17	+0.40 -0.34
$\geq 6j, SR5, 6$ and $\geq 6j, SR7, 8, 9$	6	41	+0.94 -0.94	4.43	+0.25 -0.20	+0.46 -0.39
$\geq 6j, SR4, 5, 6$ and $\geq 6j, SR7, 8, 9$	5	33	+0.98 -1.0	4.42	+0.25 -0.20	+0.49 -0.41
$\geq 6j, SR3, 4, 5, 6$ and $\geq 6j, SR7, 8, 9$	4	25	+1.0 -1.1	4.39	+0.26 -0.20	+0.59 -0.45
$\geq 6j, SR2, 3, 4, 5, 6$ and $\geq 6j, SR7, 8, 9$	3	17	+1.2 -1.3	4.31	+0.24 -0.19	+0.54 -0.48
$\geq 6j, SR1, 2, 3, 4, 5, 6$ and $\geq 6j, SR7, 8, 9$	2	9	+3.2 -2.9	4.13	+0.53 -0.56	+0.88 -0.78

8.5 Background Based Categorization

To control the background components, here we take control regions where certain background component is dominant. The $t\bar{t} + b\bar{b}$ dominant region should have large $t\bar{t}H$ signal events because of the flat ratio of the $t\bar{t}H$ signal and $t\bar{t} + b\bar{b}$ background (Figure 8.8). Therefore, the merging procedure takes into account of significant background compositions of $t\bar{t}$ +jets events in the order:

- merge bins where $t\bar{t} + \geq 2b > 30\%$.
- merge bins where $t\bar{t} + 1b > 30\%$, but not $t\bar{t} + \geq 2b > 30\%$.
- merge bins where $t\bar{t} + \geq 1c > 30\%$, but neither $t\bar{t} + \geq 2b > 30\%$, nor $t\bar{t} + 1b > 30\%$.
- merge other bins. This is a $t\bar{t}$ +light dominant region

The tag-weight-bin combination is categorized as shown in Figure 8.12. In total four regions are defined by this merging. The obtained region performance is shown in Figures 8.13, and Table 8.7.

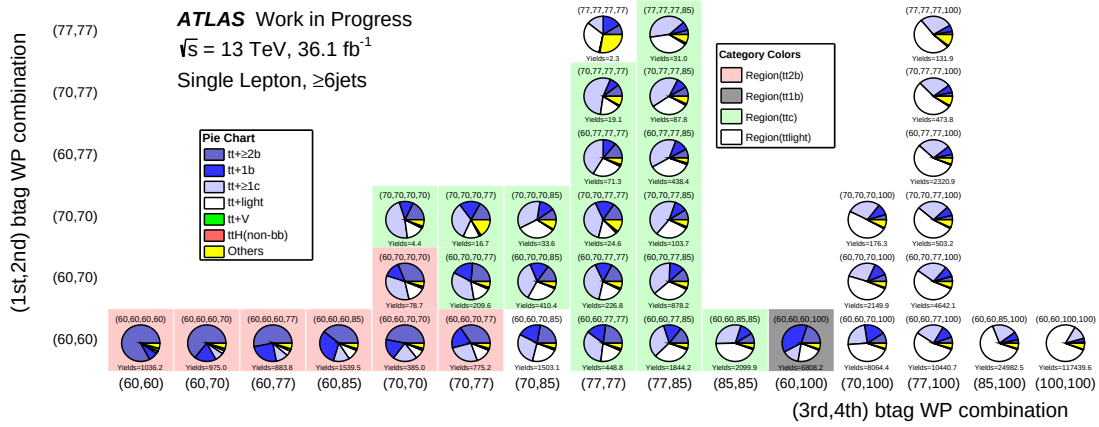


FIGURE 8.12: Region categorization based on BKG composition. The bins with same color are grouped into a region, resulting in four regions in total.

The sensitivities are obtained as:

$$\mu_{t\bar{t}H} = 1.00^{+1.2}_{-1.2}, \quad k_{t\bar{t} + \geq 1b} = 1.00^{+0.20}_{-0.18}, \quad k_{t\bar{t} + \geq 1c} = 1.00^{+0.55}_{-0.45}. \quad (8.5)$$

The $t\bar{t} + \geq 1c$ dominant region almost corresponds to the merged regions SR4, SR5, and SR6 in the S/B based definition. The $t\bar{t}$ +light dominant region also corresponds to the merged regions SR7, SR8, and SR9 in the S/B based definition. The $t\bar{t} + \geq 1b$ normalization is determined better than the S/B based definition because of the separated $t\bar{t} + 1b$ dominant region. The sensitive regions are merged into one $t\bar{t} + \geq 2b$ dominant region, and the sensitivity is not so high. Especially the UPR is not isolated as separate region any more.

8.6 Adopted Categorization

The two methods to define the analysis regions have similar results. However, the S/B based strategy has a reasonable explanation to have three signal regions, and the background based strategy shows that the $t\bar{t} + 1b$ dominant region reduces the $t\bar{t} + \geq 1b$ normalization uncertainty. To maximize the sensitivity in this analysis, the S/B and background based categorizations are

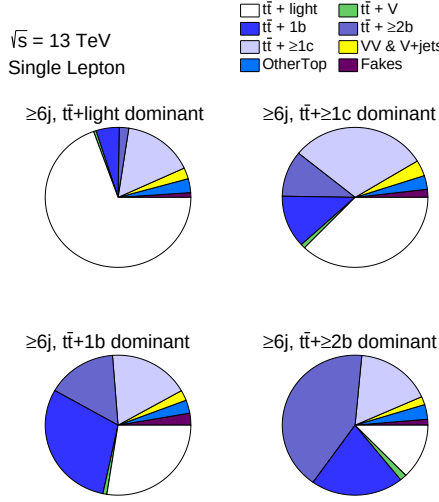


TABLE 8.7: S/B and $S/\sqrt{B}$ in the four analysis regions defined by background components. The "distribution" column shows the variable used in the fit.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
$\geq 6j, t\bar{t}+\text{light}$	2.5×10^{-3}	1.1	Normalization	1
$\geq 6j, t\bar{t}+\geq 1c$	8.7×10^{-3}	0.93	H_T^{had}	8
$\geq 6j, t\bar{t}+1b$	1.2×10^{-2}	1.12	H_T^{had}	8
$\geq 6j, t\bar{t}+\geq 2b$	2.9×10^{-2}	2.6	ClassBDT	8

FIGURE 8.13: Background fractions in the nine analysis regions for ≥ 6 jets defined by background components.

combined. At first, the background based categorization was performed to take control regions (CRs) to suppress the modeling uncertainties of the $t\bar{t}$ processes. Afterwards, defined $t\bar{t}+\geq 2b$ region was subdivided to increase the sensitivity by the categorization based on S/B , where three signal regions were constructed. The ratio of $t\bar{t}H$ and $t\bar{t}+\geq 2b$ events is almost constant in the background based $t\bar{t}+\geq 2b$ region, therefore, the fraction of $t\bar{t}+\geq 2b$ can be used instead of S/B for the categorization. Region definition in the ≥ 6 jets is summarized as:

- $\text{SR}(t\bar{t}H)$: merge bins where $t\bar{t}+\geq 2b$ of at least 60% (UPR).
- $\text{SR}(t\bar{t}+\geq 2b, \text{Hi})$: merge remaining bins where $t\bar{t}+\geq 2b$ of at least 45%.
- $\text{SR}(t\bar{t}+\geq 2b, \text{Lo})$: merge remaining bins where $t\bar{t}+\geq 2b$ of at least 30%.
- $\text{CR}(t\bar{t}+1b)$: merge all remaining bins with a $t\bar{t}+1b$ content of at least 30%.
- $\text{CR}(t\bar{t}+\geq 1c)$: merge all remaining bins with a $t\bar{t}+\geq 1c$ content of at least 30%.
- $\text{CR}(t\bar{t}+\text{light})$: Merge all remaining bins.

The same procedure was taken in the 5 jets:

- $\text{SR}(t\bar{t}H)$: merge bins where $t\bar{t}+\geq 2b$ of at least 60% (UPR).
- $\text{SR}(t\bar{t}+\geq 2b)$: merge remaining bins where $t\bar{t}+\geq 2b$ of at least 20%.
- $\text{CR}(t\bar{t}+1b)$: merge all remaining bins with a $t\bar{t}+1b$ content of at least 20%.
- $\text{CR}(t\bar{t}+\geq 1c)$: merge all remaining bins with a $t\bar{t}+\geq 1c$ content of at least 20%.
- $\text{CR}(t\bar{t}+\text{light})$: Merge all remaining bins.

The numbers of events in 5 j regions are smaller than ≥ 6 j regions, therefore, two signal regions are defined to avoid small statistics in 5jSR($t\bar{t}+\geq 2b$) region. The obtained region performance is shown in Figures 8.14, and Table 8.8. The merged tag weight bins in each region in 5 and ≥ 6 jets are shown in Figures 8.15 and 8.16. The SR($t\bar{t}H$) has higher S/B without much less $t\bar{t}+\geq 1c$ contribution compared with the single WP categorization. Therefore, the signal sensitivity depends much less on $t\bar{t}+\geq 1c$ modeling. The sensitivities with the ≥ 6 j regions are :

$$\mu_{t\bar{t}H} = 1.00^{+0.82}_{-0.78}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.18}_{-0.16}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.46}_{-0.38}. \quad (8.6)$$

Their uncertainties are less than both S/B and background based strategies. We can determine the signal strength simultaneously with the $t\bar{t}+HF$ normalizations.

The total sensitivity using all 5 jets and ≥ 6 j regions are obtained as:

$$\mu_{t\bar{t}H} = 1.00^{+0.62}_{-0.59}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.12}_{-0.12}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.34}_{-0.29}. \quad (8.7)$$

The uncertainties are smaller with this region definition compared to the simple region definition using two WPs, as well as the region definition using a single WP.

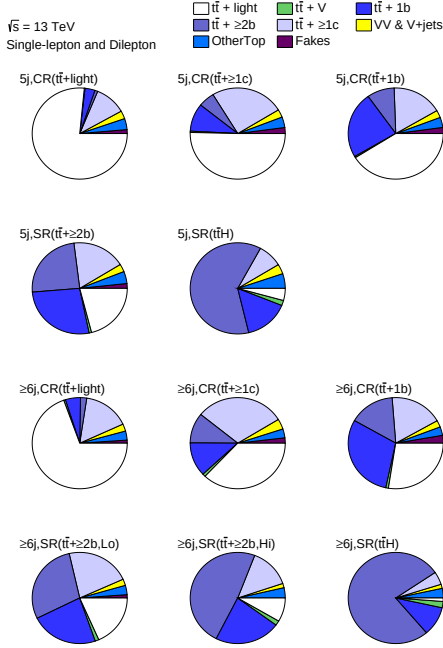


FIGURE 8.14: Background fractions in the eleven analysis regions for the single-lepton channel.

TABLE 8.8: S/B and $S/\sqrt{B}$ in the eleven analysis regions for the single-lepton channel. The "distribution" column shows the variable used in the fit.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
5jCR($t\bar{t}+\text{light}$)	8.8×10^{-4}	0.44	Normalization	1
5jCR($t\bar{t}+\geq 1c$)	3.7×10^{-3}	0.26	H_T^{had}	6
5jCR($t\bar{t}+1b$)	7.2×10^{-3}	0.70	H_T^{had}	6
5jSR($t\bar{t}+\geq 2b$)	1.8×10^{-2}	0.86	ClassBDT	8
5jSR($t\bar{t}H$)	4.8×10^{-2}	0.87	ClassBDT	8
$\geq 6j$ CR($t\bar{t}+\text{light}$)	2.5×10^{-3}	1.1	Normalization	1
$\geq 6j$ CR($t\bar{t}+\geq 1c$)	9.1×10^{-3}	0.96	H_T^{had}	8
$\geq 6j$ CR($t\bar{t}+1b$)	1.4×10^{-2}	1.2	H_T^{had}	8
$\geq 6j$ SR($t\bar{t}+\geq 2b, \text{Lo}$)	2.3×10^{-2}	1.4	ClassBDT	8
$\geq 6j$ SR($t\bar{t}+\geq 2b, \text{Hi}$)	3.6×10^{-2}	1.7	ClassBDT	8
$\geq 6j$ SR($t\bar{t}H$)	5.4×10^{-2}	1.9	ClassBDT	8

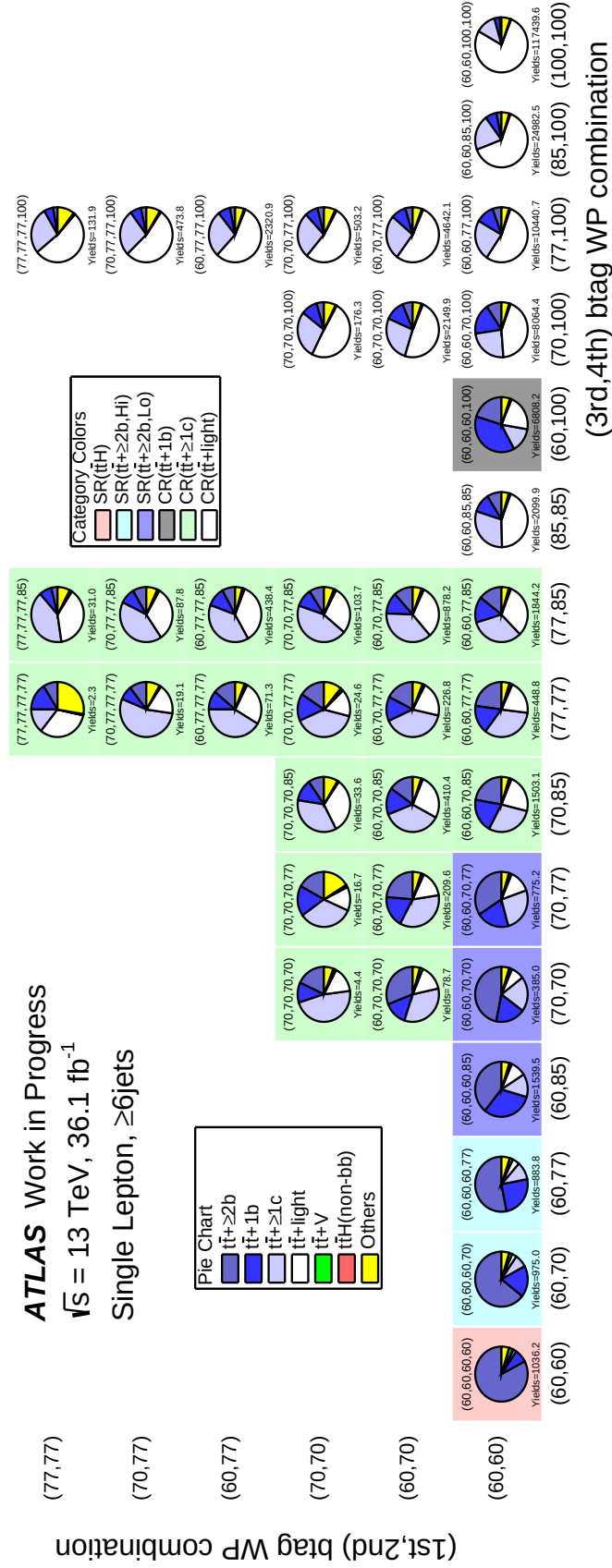


FIGURE 8.15: Region categorization for ≥ 6 jets events in the single-lepton channel. The three signal regions and three control regions are defined.

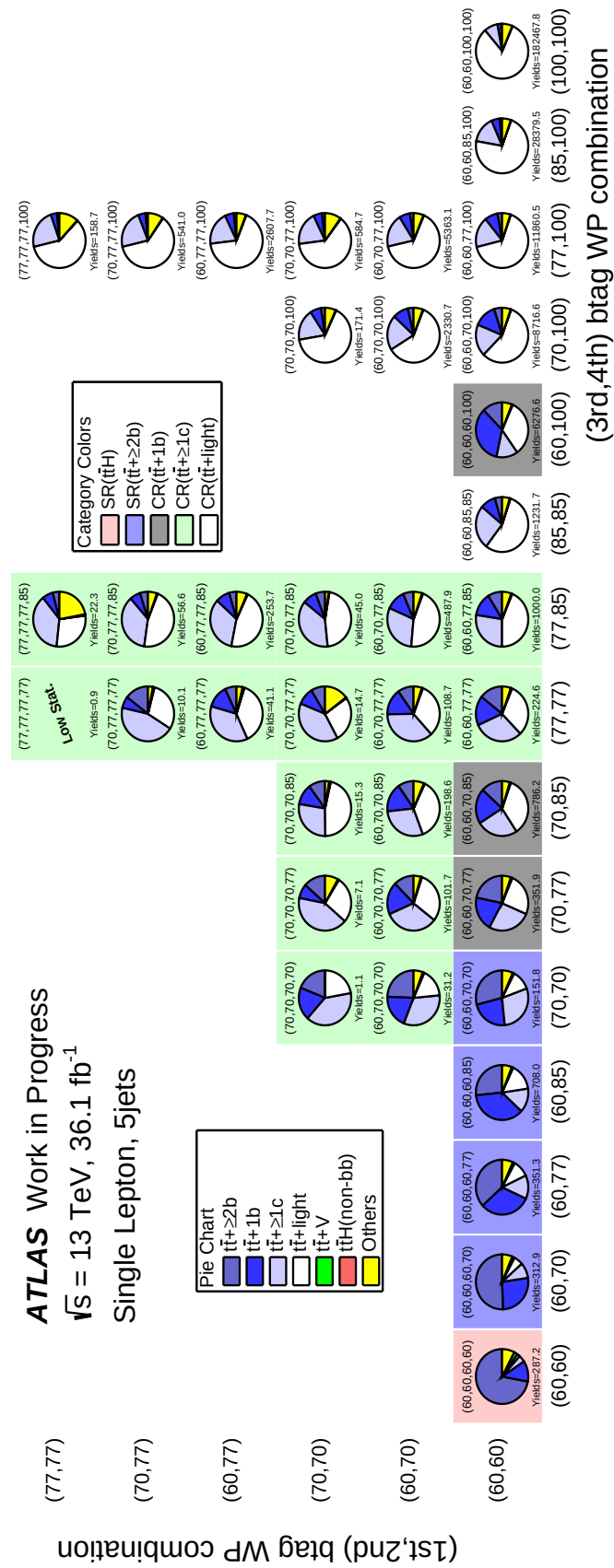


FIGURE 8.16: Region categorization for 5 jets events in the single-lepton channel. The two signal regions and three control regions are defined.

8.7 Boosted Region

In the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, identification of $H \rightarrow b\bar{b}$ is difficult to obtain without biases on the Higgs mass distribution (discussed in Chapter 9). An additional analysis channel was adopted targeting events where the Higgs-boson and hadTop have high momenta such that their decay products are close by. This boosted channel was implemented in the single lepton channel. The event selection is based on large- R jets formed by reclustering the standard $R = 0.4$ jets into $R = 1.0$ jets, which is denoted as rcjet. The boosted channel events were built from events with one lepton and ≥ 5 jets including $\text{nBTags}_{85\%} \geq 4$ as followings:

- require rcjets with $200 < p_T < 1500$ GeV, $m_{\text{rcjet}} > 50$ GeV, $|\eta| < 2$, and at least two sub-constituents $R = 0.4$ jets
- require at least one rcjet with $p_T > 200$ GeV and including ≥ 2 b -tagged jets at 85% WP
- choose one Higgs-tagged rcjet with the largest sum of b -tagging tag weight bins of sub-constituents $R = 0.4$ jets
- require at least one rcjet with $p_T > 250$ GeV and including ≥ 1 b -tagged jets and ≥ 1 non- b -tagged jets at 85% WP
- choose one top-tagged rcjet with the largest rcjet mass
- require at least one b -tagged jets at 85% WP out of Higgs- and top-tagged rcjets, which corresponds to the b -jet from lepTop

These events were removed from the single lepton channel regions discussed above, and kept as one signal region called "boosted SR". The fraction of events removed by this overlap removal procedure is 4% from $t\bar{t}H$ events and 3% from $t\bar{t}$ events in the $\geq 6\text{jSR}(t\bar{t}H)$. The truth Higgs p_T and truth hadTop p_T in events before and after boosted selections are shown in Figure 8.17.

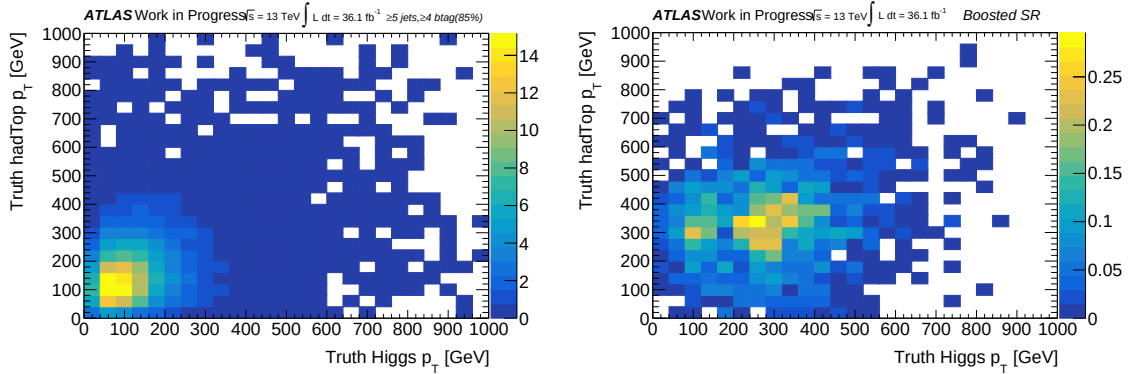


FIGURE 8.17: Truth Higgs p_T and truth hadTop p_T . (Left) ≥ 5 jets including $\text{nBTags}_{85\%} \geq 4$ events are used. (Right) The events with boosted selections are used.

In the boosted region, high p_T Higgs and hadTop are chosen. The Higgs-tagged rcjet in the boosted SR contains two b -tagged jets truth-matched to the Higgs decay products in 49% of the $t\bar{t}H$ events. In the events where two Higgs decay products are truth-matched, the probability of the Higgs-tagged rcjet containing two truth-matched b -tagged jets is 55%. The probability of the Higgs reconstruction in the boosted SR is higher than resolved regions (described in Chapter 9). The S/B and $S/\sqrt{B}$ in the boosted region are 0.025 and 0.65, respectively. The background composition is not dominated by $t\bar{t}+\geq 1b$ unlike signal regions defined in the previous section, therefore the boosted topology is expected a good sensitivity in this analysis.

8.8 Minor Modification Considerations

8.8.1 Loose UPR

The adopted UPR events are after highly tight selection, and the MC sample has small statistics accordingly. A loose UPR is considered, which includes $(B_{j1}, B_{j2}, B_{j3}, B_{j4}) = (60, 60, 60, 70)$ bin in addition to $(60, 60, 60, 60)$. If the two sensitivities are not so different, the loose UPR can be adopted the $SR(t\bar{t}H)$ to avoid small statistics in the most sensitive signal region. Therefore, the signal regions are modified as an alternative strategy defined as:

- Loose $\geq 6jSR(t\bar{t}H)$: $(60, 60, 60, 60)$ and $(60, 60, 60, 70)$ in ≥ 6 jets
- Loose $\geq 6jSR(t\bar{t}+\geq 2b, Hi)$: $(60, 60, 60, 77)$ and $(60, 60, 60, 85)$ in ≥ 6 jets
- Loose $\geq 6jSR(t\bar{t}+\geq 2b, Lo)$: $(60, 60, 70, 70)$ and $(60, 60, 70, 77)$ in ≥ 6 jets
- Loose $5jSR(t\bar{t}H)$: $(60, 60, 60, 60)$ and $(60, 60, 60, 70)$ in 5 jets
- Loose $5jSR(t\bar{t}+\geq 2b)$: $(60, 60, 60, 77)$, $(60, 60, 60, 85)$, and $(60, 60, 70, 70)$ in 5 jets

Their statistic uncertainty, S/B , and $S/\sqrt{B}$ comparisons are shown in Table 8.9. The loose signal regions have small statistical simulation uncertainties especially in the SHERPA $t\bar{t}+\geq 1b$ systematic sample.

TABLE 8.9: Statistics in signal regions for the single-lepton channel. The column named "tag weight bins" shows the combinations of the tag weight bins for each signal region. The nominal (POWHEG+PYTHIA8) and SHERPA $t\bar{t}+\geq 1b$ events are shown for the bench marks of the simulation statistics. The uncertainties show statistical uncertainties on the simulation samples, and values with brackets show their relative uncertainties. All events are used in eight binned Class-BDT distributions, and the statistical uncertainty in each bin is larger than the whole normalization uncertainty.

regions	# nominal $t\bar{t}+\geq 1b$	# SHERPA $t\bar{t}+\geq 1b$	S/B	$S/\sqrt{B}$
nominal signal regions				
$\geq 6jSR(t\bar{t}H)$	1010 ± 13 (1.3%)	870 ± 26 (3.1%)	5.4×10^{-2}	1.8
$\geq 6jSR(t\bar{t}+\geq 2b, Hi)$	1750 ± 18 (1.0%)	1590 ± 30 (1.9%)	3.3×10^{-2}	1.6
$\geq 6jSR(t\bar{t}+\geq 2b, Lo)$	2100 ± 19 (0.90%)	1920 ± 40 (2.1%)	2.1×10^{-2}	1.3
$5jSR(t\bar{t}H)$	273 ± 6 (2.2%)	266 ± 12 (4.5%)	4.5×10^{-2}	0.85
$5jSR(t\bar{t}+\geq 2b)$	1230 ± 14 (1.1%)	1164 ± 26 (2.2%)	1.7×10^{-2}	0.82
alternative loose signal regions				
loose $\geq 6jSR(t\bar{t}H)$	1930 ± 20 (0.98%)	1680 ± 30 (1.8%)	4.5×10^{-2}	2.2
loose $\geq 6jSR(t\bar{t}+\geq 2b, Hi)$	1130 ± 10 (1.2%)	1050 ± 30 (2.9%)	2.8×10^{-2}	1.2
loose $\geq 6jSR(t\bar{t}+\geq 2b, Lo)$	1800 ± 20 (1.1%)	1660 ± 40 (2.4%)	2.1×10^{-2}	1.2
loose $5jSR(t\bar{t}H)$	561 ± 9 (1.6%)	525 ± 17 (3.3%)	3.6×10^{-2}	1.0
loose $5jSR(t\bar{t}+\geq 2b)$	939 ± 13 (1.4%)	905 ± 23 (2.5%)	1.5×10^{-2}	0.63

The sensitivities with the loose UPRs are obtained as:

$$\mu_{t\bar{t}H} = 1.00^{+0.73}_{-0.73}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.13}_{-0.12}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.35}_{-0.30}. \quad (8.8)$$

The uncertainties are much worse if the loose UPR is taken as the $SR(t\bar{t}H)$, and the loose signal regions are not adopted in this analysis.

8.8.2 More Controls for $t\bar{t}+\geq 1c$ Process

To take better control of the $t\bar{t}+\geq 1c$ process, $CR(t\bar{t}+\geq 1c)$ is subdivided into two control regions as following requirements:

- $CR(t\bar{t}+\geq 1c)$ (vs. $t\bar{t}+\geq 1b$): require $t\bar{t}+\text{light} < 20\%$.
- $CR(t\bar{t}+\geq 1c)$ (vs. $t\bar{t}+\text{light}$): remaining events

Their background compositions are shown in Figure 8.18. The sensitivities are obtained as:

$$\mu_{t\bar{t}H} = 1.00^{+0.60}_{-0.56}, \quad k_{t\bar{t}+\geq 1b} = 1.00^{+0.11}_{-0.11}, \quad k_{t\bar{t}+\geq 1c} = 1.00^{+0.31}_{-0.27}. \quad (8.9)$$

The sensitivity is slightly better than the nominal region definitions, however this option is not adopted for this analysis to avoid having small statistics regions.

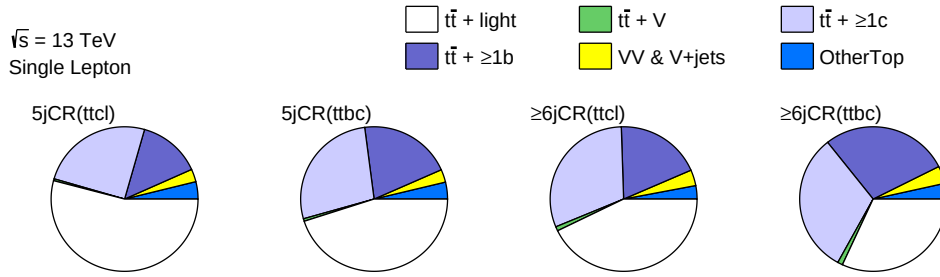


FIGURE 8.18: Background fractions in the subdivided $CR(t\bar{t}+\geq 1c)$ regions for the single-lepton channel.

8.9 Dilepton Regions

Same procedure was conducted also in the dilepton channel. To reduce the non $t\bar{t}$ processes, the pre-selection was also applied in the dilepton channel as $n\text{BTags}_{77\%} \geq 2$. Number of $t\bar{t}H(H \rightarrow b\bar{b})$ and S/B ratio in the ≥ 4 j region in all tag-weight-bin combinations are shown in Figures 8.19 and 8.20.

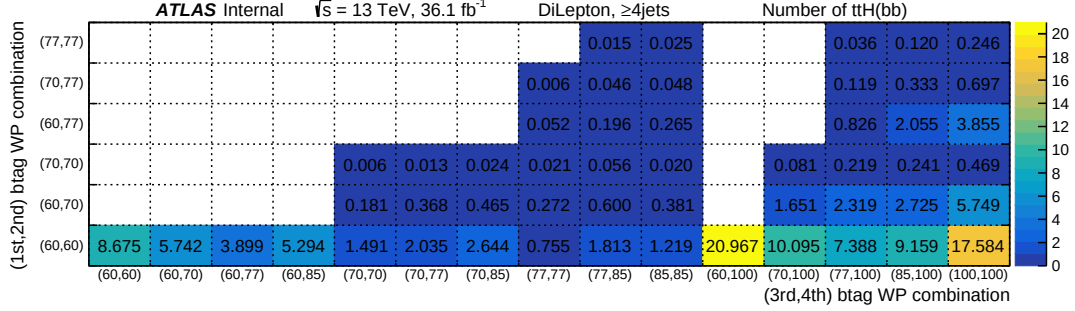


FIGURE 8.19: Number of $t\bar{t}H(H \rightarrow b\bar{b})$ in ≥ 4 jets categorized in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$.

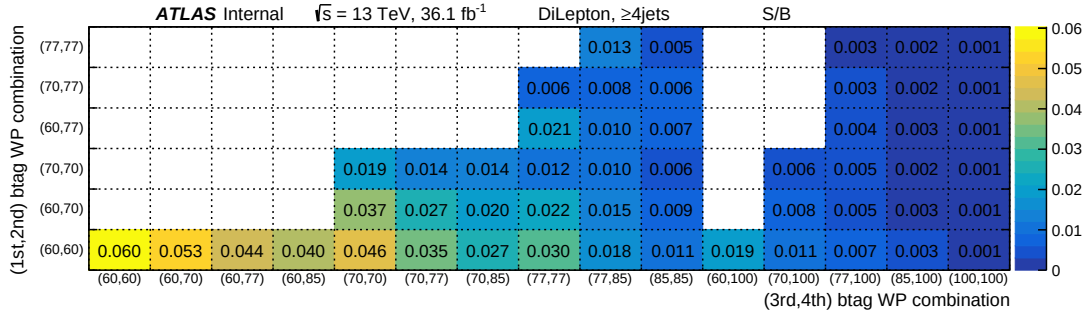


FIGURE 8.20: S/B ratio in ≥ 4 jets categorized in all combinations of $(B_{j1}, B_{j2}, B_{j3}, B_{j4})$.

Dilepton events are also categorized based on the background composition at first. Afterwards, $t\bar{t} + \geq 2b$ region was subdivided to maximize the sensitivity with S/B variations but not to lose statistics in each region. The final state of the dilepton channel has two leptons, therefore, control regions utilized H_T^{all} distributions to control lepton kinematics as well as jets, instead of H_T^{had} . The H_T^{all} is calculated by the scalar p_T sum over all jets and leptons in each event. For the dilepton channel with ≥ 4 jets:

- $\text{SR}(t\bar{t}H)$: Merge all bins with a $t\bar{t} + \geq 2b$ content of at least 70% (loose UPR).
- $\text{SR}(t\bar{t}+1b)$: Merge all remaining bins with a $t\bar{t} + 1b$ content of at least 50%.
- $\text{SR}(t\bar{t}+\geq 2b)$: Merge all remaining bins with a $t\bar{t} + \geq 2b$ content of at least 20%.
- $\text{CR}(t\bar{t}+\geq 1c)$: Merge all remaining bins with a $t\bar{t} + \geq 1c$ content of at least 30%.
- $\text{CR}(t\bar{t}+\text{light})$: Merge all remaining bins.

For the dilepton channel with 3 jets:

- $\text{CR}(t\bar{t}+\geq 1b)$: Merge all bins with a $t\bar{t} + \geq 1b$ content of at least 30%.

- CR($t\bar{t}$ +light): Merge all remaining bins.

Defined region categorizations in 3 and ≥ 4 j regions are shown in Figures 8.22 and 8.23.

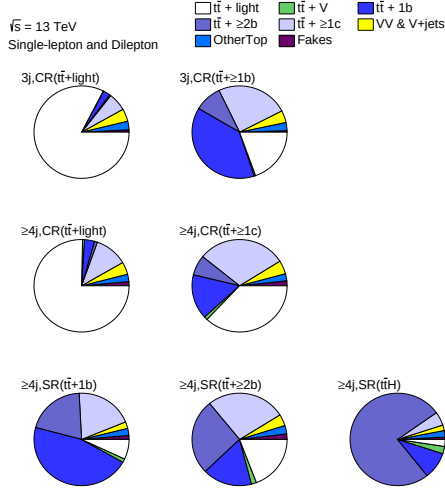


FIGURE 8.21: Background fractions in the seven analysis regions for the dilepton channel.

TABLE 8.10: S/B and $S/\sqrt{B}$ in the eleven analysis regions for the single-lepton channel. The "distribution" column shows the variable used in the fit.

region	S/B	$S/\sqrt{B}$	distribution	# of bins
3jCR($t\bar{t}$ +light)	4.3×10^{-4}	0.12	Normalization	1
3jCR($t\bar{t}+\geq 1b$)	5.8×10^{-3}	0.23	H_T^{all}	2
≥ 4 jCR($t\bar{t}$ +light)	2.1×10^{-3}	0.50	Normalization	1
≥ 4 jCR($t\bar{t}+\geq 1c$)	7.8×10^{-3}	0.53	H_T^{all}	5
≥ 4 jSR($t\bar{t}+1b$)	1.8×10^{-2}	0.64	ClassBDT	8
≥ 4 jSR($t\bar{t}+\geq 2b$)	2.3×10^{-2}	0.83	ClassBDT	8
≥ 4 jSR($t\bar{t}H$)	6.0×10^{-2}	0.97	ClassBDT	8

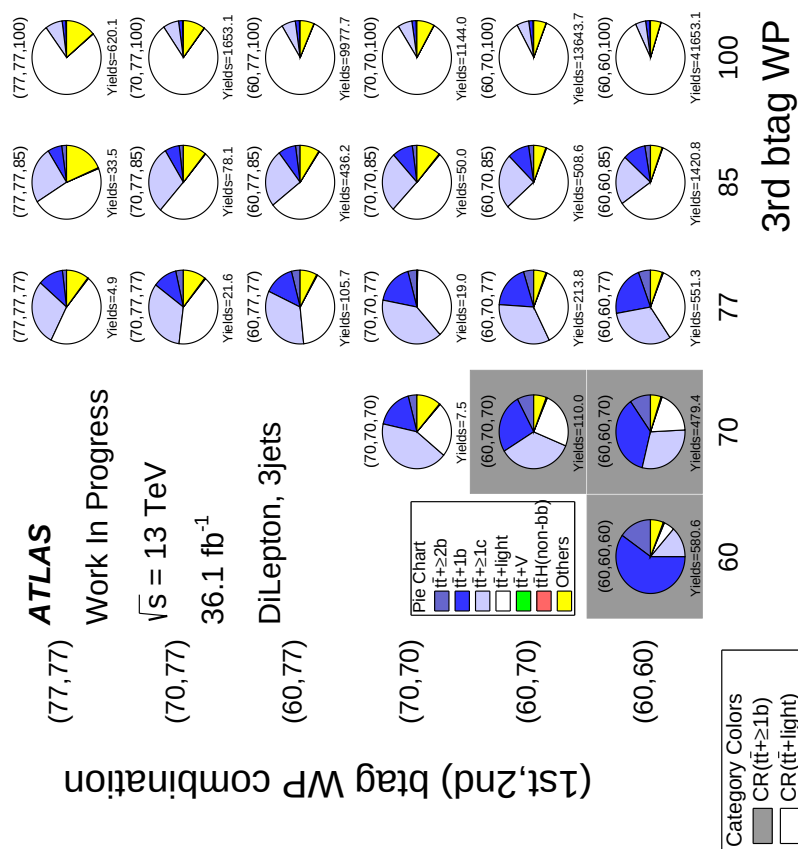


FIGURE 8.23: Region categorization for 3 jets events in the dilepton channel. The two control regions are defined.

8.10 Comparison of Data and Simulation in Defined Regions

For the validation of the prediction, the data and simulation consistencies with various distributions were checked in defined regions.

At first, the numbers of jets was checked inclusively in each lepton channel as shown in Figure 8.24. The numbers of b -tagged jets at four WPs were also checked inclusively in all analysis regions for each lepton channel. The each tag-weight-bin combination for the pseudo-continuous b -tagging was also checked. The two-dimensional tag-weight-bin map is converted to one-dimensional distribution, where the bin numbering is shown in Figure 8.25. The two-dimensional bins with small statistics were merged in the conversion to avoid the large statistical uncertainties in any bins in one-dimensional distribution. The b -tagging distributions were shown in Figures 8.26 and 8.27 for the single-lepton and dilepton channels, respectively. In these plots, events passing the pre-selection ($\text{nBTags}_{77\%} \geq 3$ or $\text{nBTags}_{60\%} \geq 2$) or the boosted-selection ($\text{nBTags}_{85\%} \geq 4$) are included for the single-lepton channel, while events passing the pre-selection ($\text{nBTags}_{77\%} \geq 2$) are included for the dilepton channel.

As the object momentum validation, the H_T^{had} and H_T^{all} distributions are shown in each region for the single-lepton and dilepton channels, respectively. (Figures 8.28 to 8.31) Control regions dominated by the $t\bar{t} + \geq 1c$ or $t\bar{t} + 1b$ processes utilize their distributions as the global fit for the determination of the signal cross section.

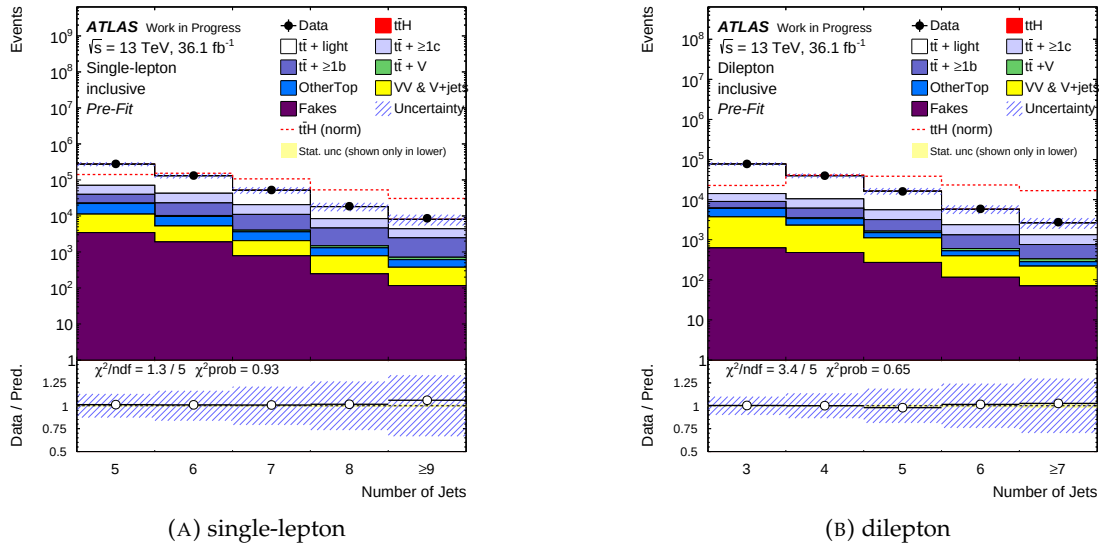


FIGURE 8.24: Number of jets (A) in the single-lepton channel and (B) in the dilepton channel.

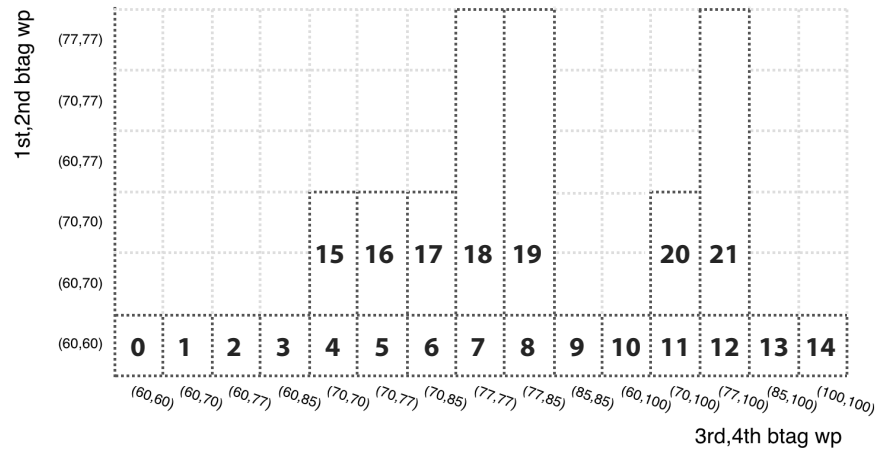
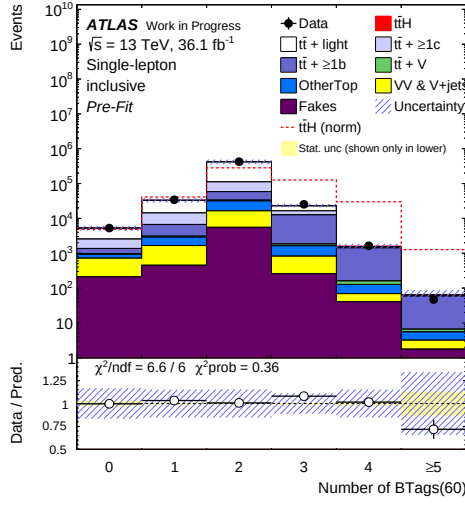
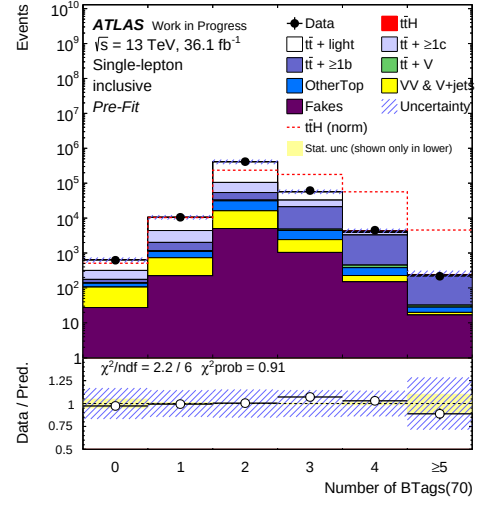
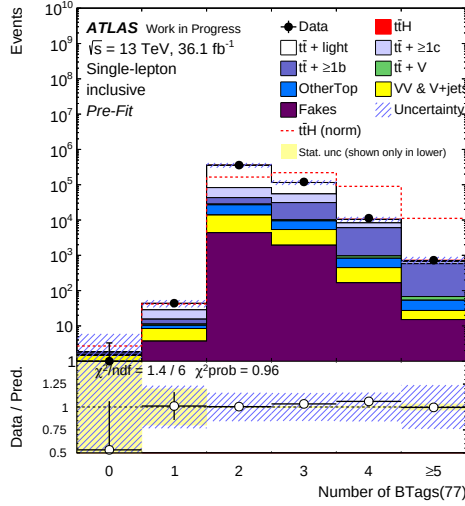
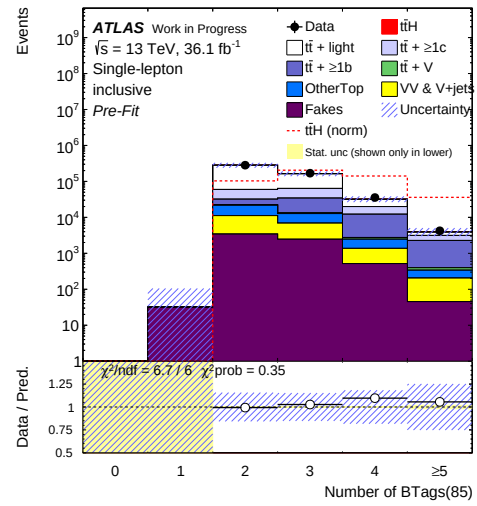
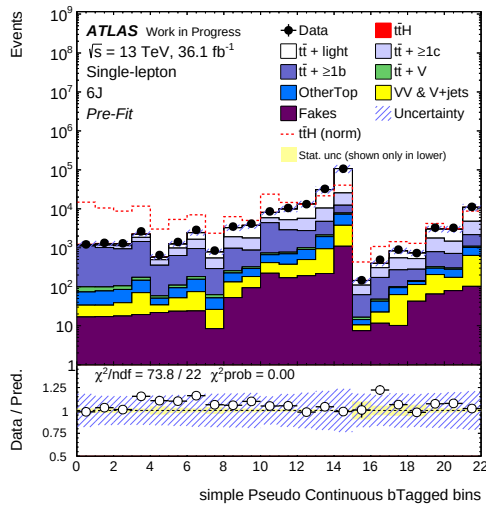
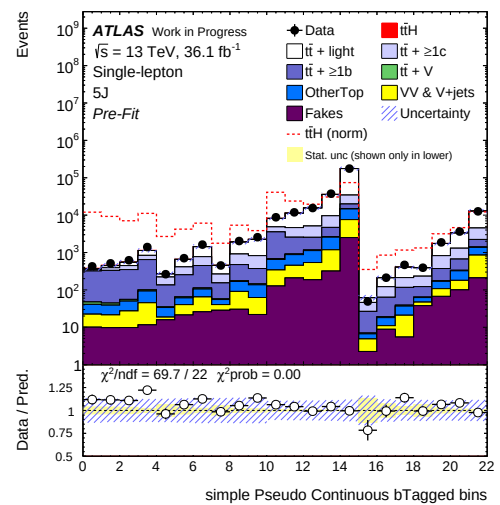
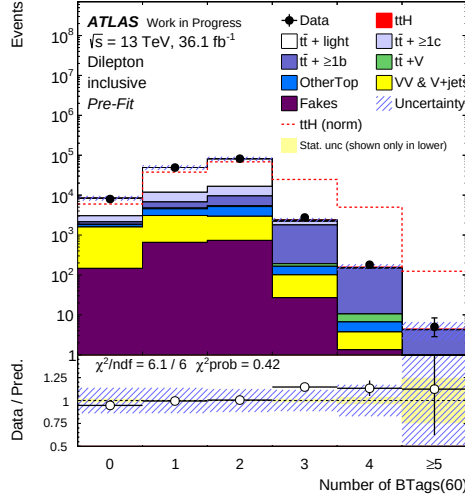
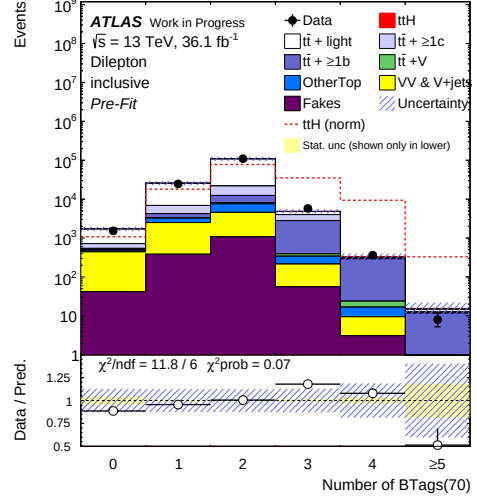
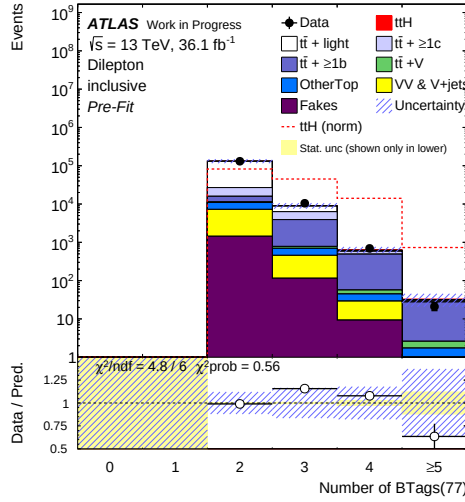
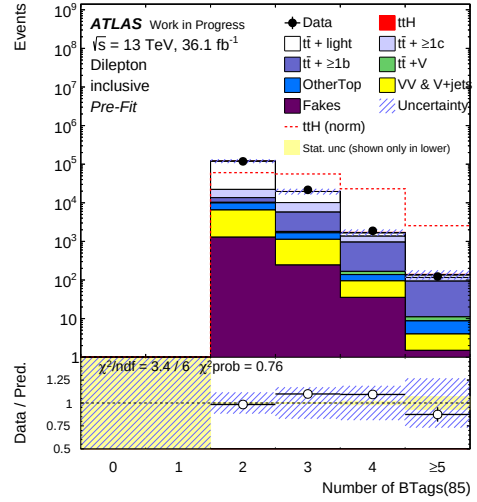
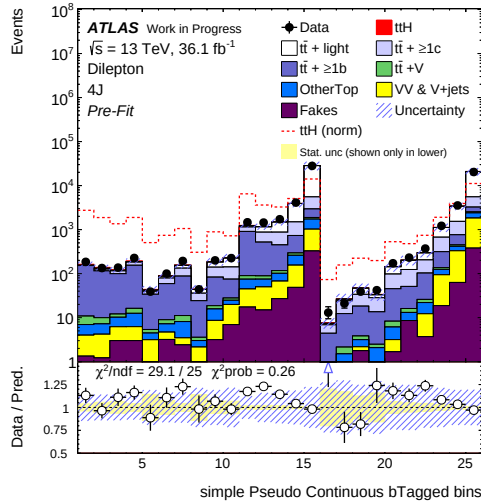
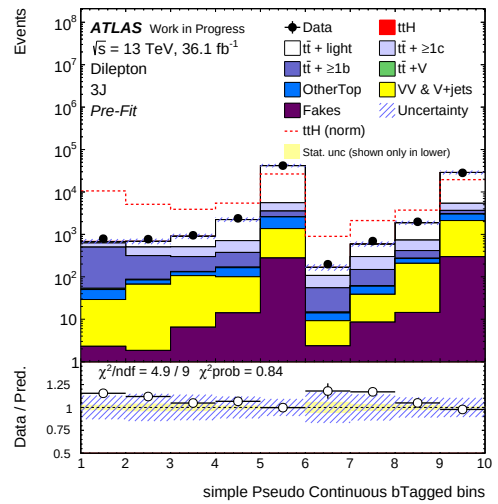


FIGURE 8.25: Tag weight bin numbering to convert into the one dimensional distribution. Column bins are merged to avoid poor statistics.

(A) # b -tagged jets at 60% WP(B) # b -tagged jets at 70% WP(C) # b -tagged jets at 77% WP(D) # b -tagged jets at 85% WP(E) tag weight bins in the ≥ 6 j region

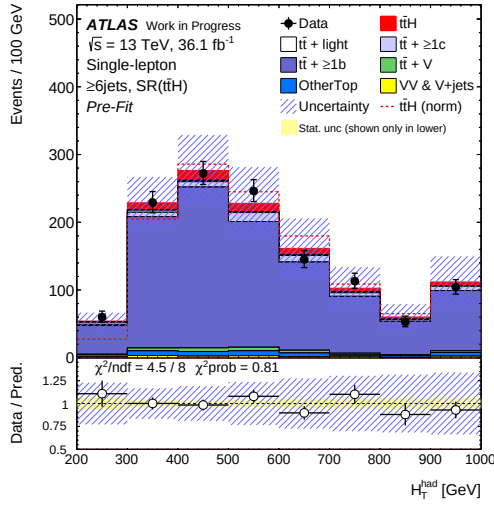
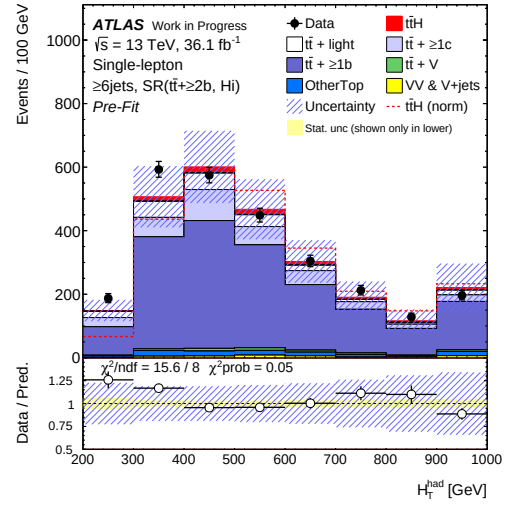
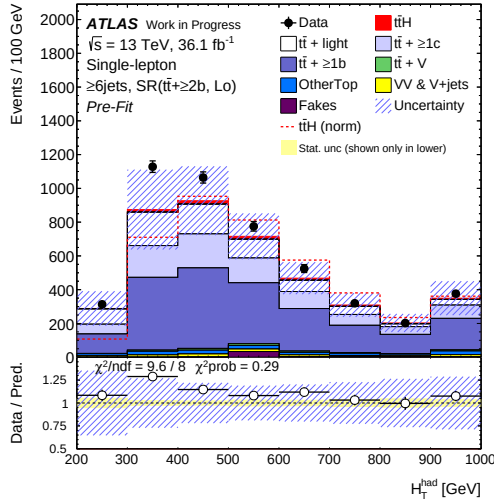
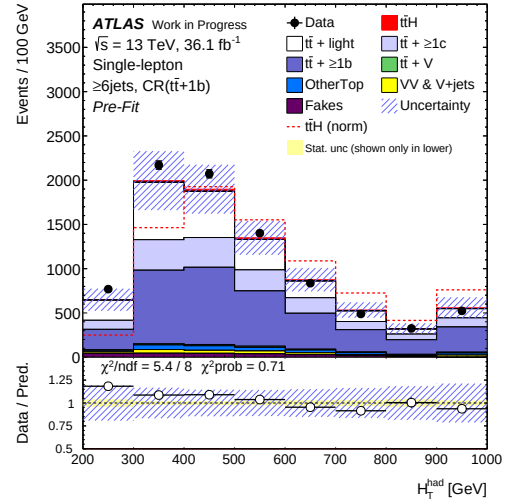
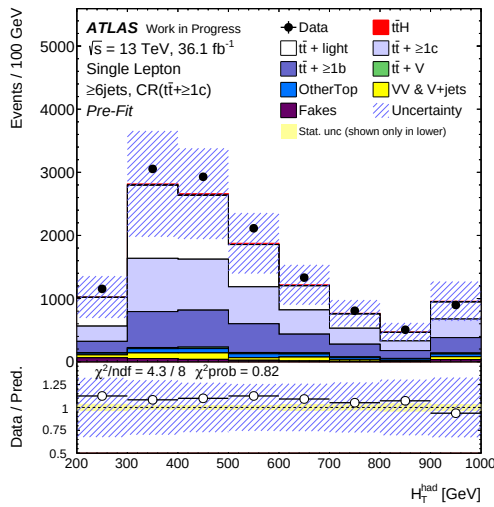
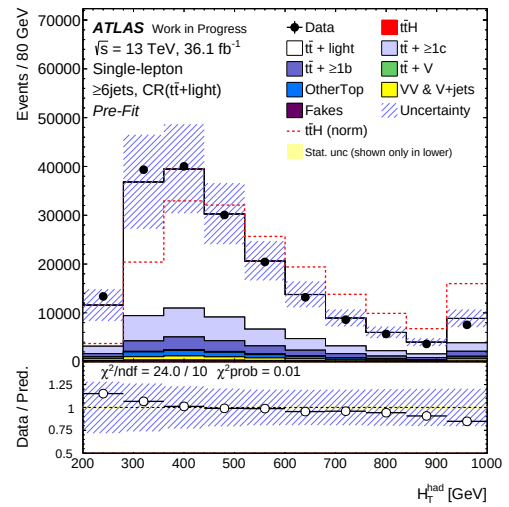
(F) tag weight bins in the 5 j region

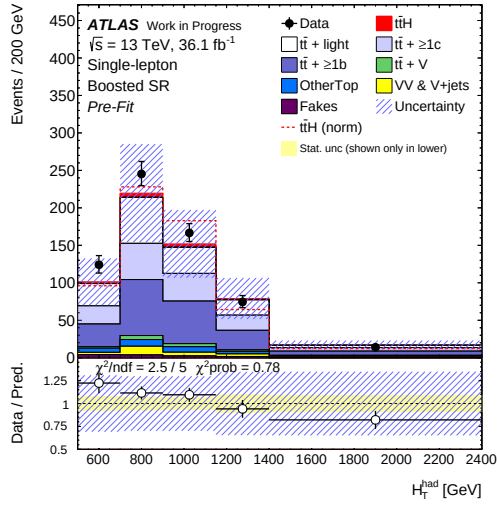
FIGURE 8.26: Number of b -tagged jets at each WP in the single-lepton channel. The numbering for the distributions of the tag weight bin in (E) and (F) is shown in Figure 8.25.

(A) # b -tagged jets at 60% WP(B) # b -tagged jets at 70% WP(C) # b -tagged jets at 77% WP(D) # b -tagged jets at 85% WP(E) tag weight bins in the ≥ 4 j region

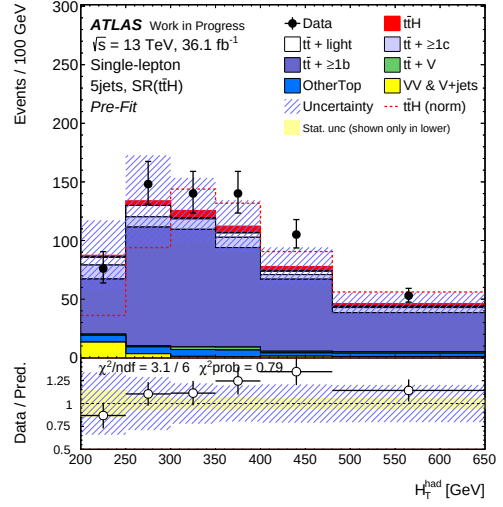
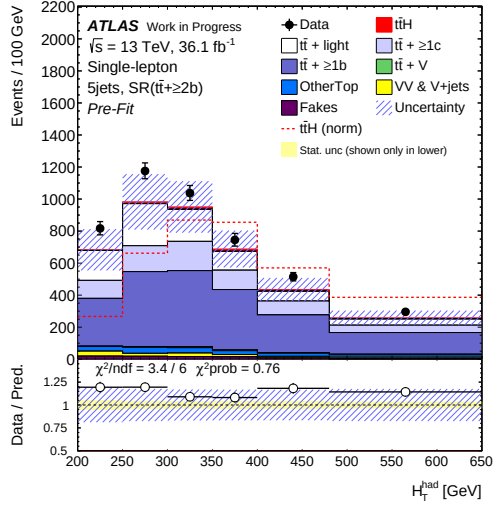
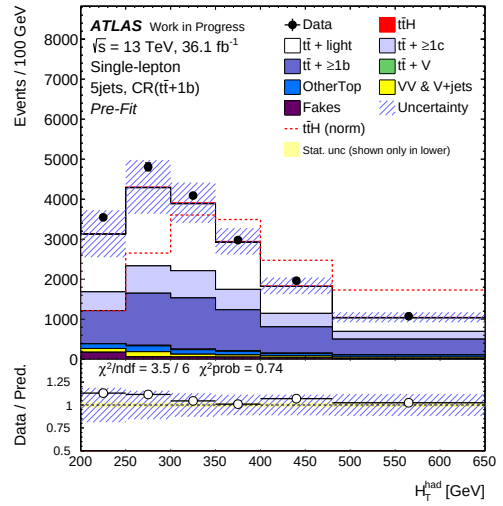
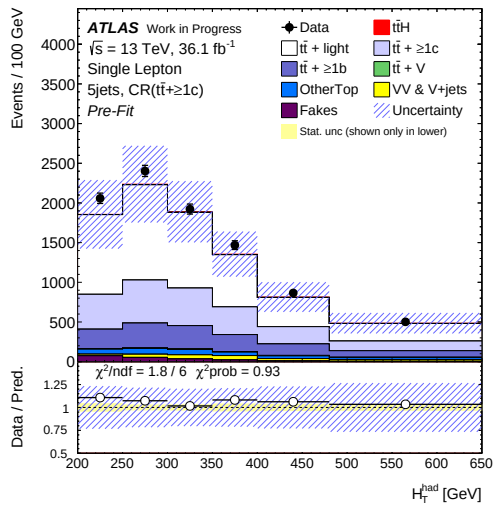
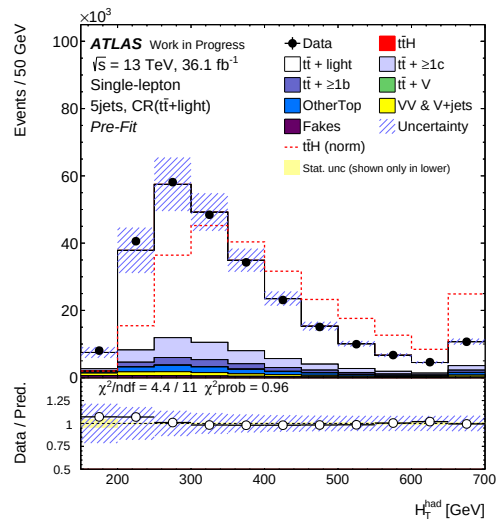
(F) tag weight bins in the 3 j region

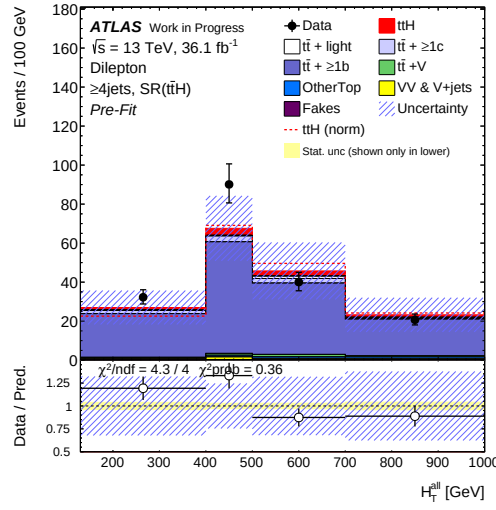
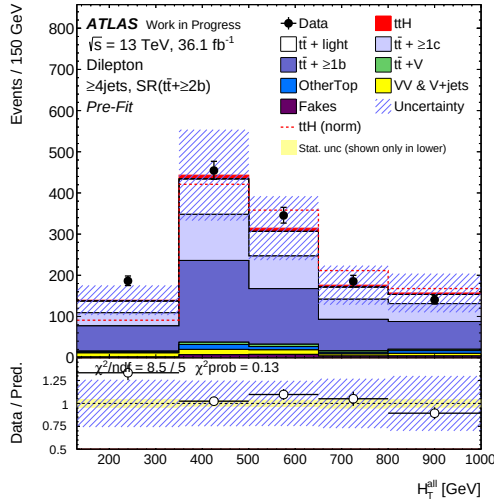
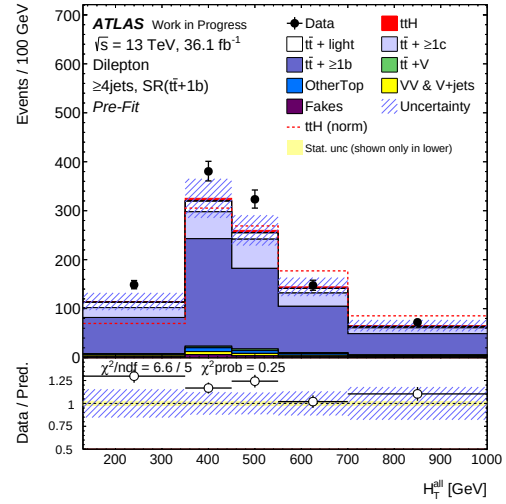
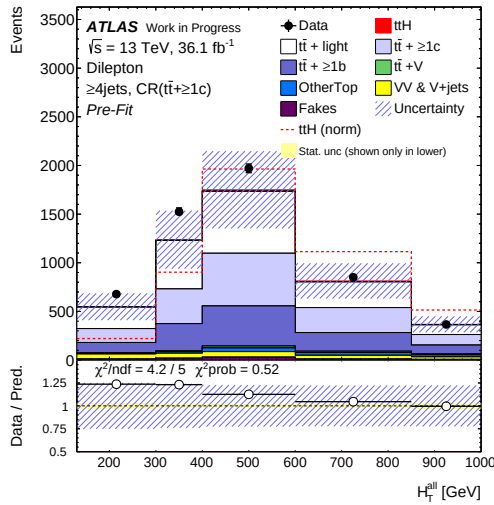
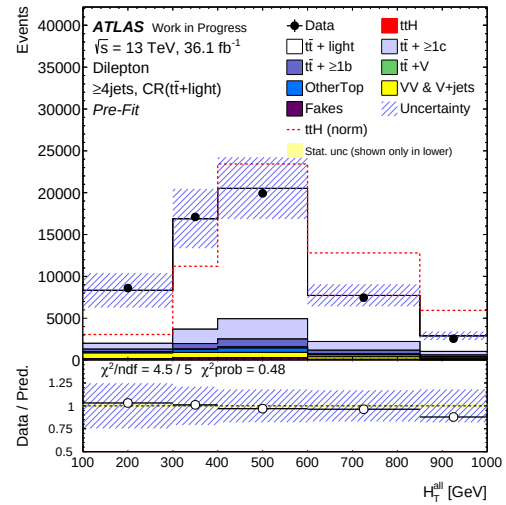
FIGURE 8.27: Number of b -tagged jets at each WP in the dilepton channel. The numbering for the distributions of the tag weight bin in (E) and (F) is shown in Figure 8.25.

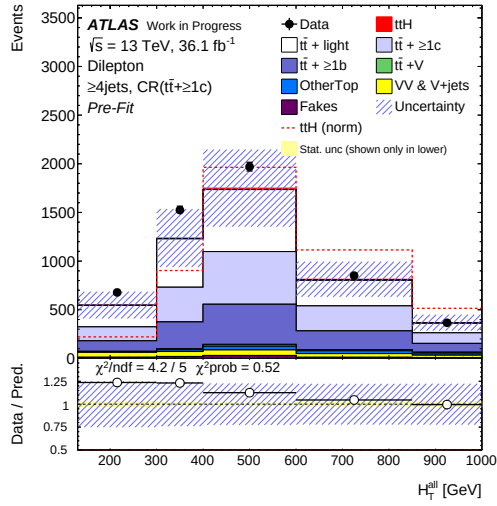
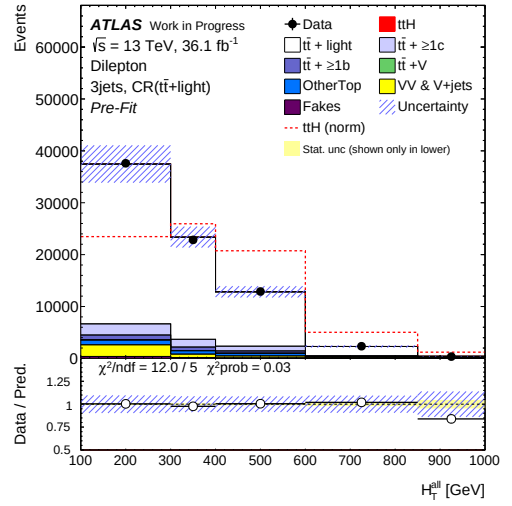
(A) $\geq 6 \text{ jets SR}(t\bar{t}H)$ (B) $\geq 6 \text{ jets SR}(t\bar{t} + \geq 2b, \text{Hi})$ (C) $\geq 6 \text{ jets SR}(t\bar{t} + \geq 2b, \text{Lo})$ (D) $\geq 6 \text{ jets CR}(t\bar{t} + 1b)$ (E) $\geq 6 \text{ jets CR}(t\bar{t} + \geq 1c)$ (F) $\geq 6 \text{ jets CR}(t\bar{t} + \text{light})$ FIGURE 8.28: H_T^{had} distributions in $\geq 6 \text{ j}$ regions for the single-lepton channel



(A) Boosted SR

(B) 5jSR($t\bar{t}H$)(C) 5jSR($t\bar{t} + \geq 2b$)(D) 5jSR($t\bar{t} + 1b$)(E) 5jCR($t\bar{t} + \geq 1c$)(F) 5jCR($t\bar{t} + \text{light}$)FIGURE 8.29: H_T^{had} distributions in 5 j regions and boosted region for the single-lepton channel

(A) $\geq 4\text{jSR}(t\bar{t}H)$ (B) $\geq 4\text{jSR}(t\bar{t} + \geq 2b)$ (C) $\geq 4\text{jSR}(t\bar{t} + 1b)$ (D) $\geq 4\text{jCR}(t\bar{t} + \geq 1c)$ (E) $\geq 4\text{jCR}(t\bar{t} + \text{light})$ FIGURE 8.30: H_T^{all} distributions in $\geq 4 \text{ j}$ regions for the dilepton channel

(A) $3j\text{CR}(t\bar{t} + \geq 1c)$ (B) $3j\text{CR}(t\bar{t} + \text{light})$ FIGURE 8.31: H_T^{all} distributions in 3 j regions for the dilepton channel

8.11 H_T^{had} and H_T^{all} impacts on the fit

Only the normalization was adjusted in the $t\bar{t}$ +light control regions. The differences of the fit results were discussed in this section when the H_T^{had} or H_T^{all} distributions are examined also in $t\bar{t}$ +light control regions ($\geq 6j\text{CR}(t\bar{t}+\text{light})$, $5j\text{CR}(t\bar{t}+\text{light})$, $\geq 4j\text{CR}(t\bar{t}+\geq 1c)$, and $3j\text{CR}(t\bar{t}+\text{light})$). The b -only hypothesis fits were used for their studies, where the signal events are fixed to zero in all region. The free-floating normalization factors ($k_{t\bar{t}+\geq 1b}$ and $k_{t\bar{t}+\geq 1c}$) and systematic pulls and constraints were checked without looking at the signal behaviors to avoid a biased conclusion. To suppress the signal contributions, histogram bins with $S/B > 5\%$ were blinded, and not used in the fits. Any H_T^{had} and H_T^{all} histogram bins in the control regions were not blinded in this requirement. Noticeable differences in the systematic pulls and constraints are shown in Figure 8.32. More information in $t\bar{t}$ +light control regions provide the higher constraints on uncertainties for $t\bar{t}$ +light modeling. The large differences in the systematic pulls show that they have phase spaces away from our analysis signal regions. The large statistics can move pulls to realize the better data-prediction agreement in $t\bar{t}$ +light control regions rather than in signal regions. Furthermore, the disagreement between the data and prediction can be seen in $t\bar{t}$ +light control regions (especially in $\geq 6j\text{CR}(t\bar{t}+\text{light})$), where the slope appears in the lower ratio plot in each distribution. This results the larger systematic pulls.

In Figures 8.28 to 8.30, the slopes are also seen in the ratio plots in $t\bar{t}+\geq 1c$ and $t\bar{t} + 1b$ control regions. If they exhibit large systematic pulls, the signal may be suffered by them. Therefore, the b -only data fits were tested. Three test options are considered as following:

- nominal setup: H_T^{had} utilized in $\text{CR}(t\bar{t}+1b)$ and $\text{CR}(t\bar{t}+\geq 1c)$ for the single-lepton channel, and H_T^{all} utilized in $\text{CR}(t\bar{t}+\geq 1c)$ for the dilepton channel, as shown in previous sections.
- onebin setup: normalizations utilized in all control regions for both lepton channels.
- optimum setup: H_T^{had} utilized in $\text{CR}(t\bar{t}+\geq 1c)$ for the single-lepton channel, and normalizations utilized in other regions. The H_T^{had} distributions in single-lepton $\text{CR}(t\bar{t}+\geq 1c)$ have smallest discrepancies between data and prediction. This setup is expected to have jet systematic constraints with a least impact by the simulation mis-modeling in H_T^{had} and H_T^{all} .

After each fitting, the H_T^{had} and H_T^{all} distributions are compensated by the systematic pulls if the distributions are used in the fit. The H_T^{had} distributions in $\geq 6j\text{CR}(t\bar{t}+1b)$ and $\geq 6j\text{CR}(t\bar{t}+\geq 1c)$, and H_T^{all} distributions in $\geq 4j\text{CR}(t\bar{t}+\geq 1c)$ after the b -only hypothesis fitting are shown in Figures 8.33 to 8.35. With three setups, the expected sensitivities were also checked and resulted:

$$\mu_{t\bar{t}H} = 1.00_{-0.52}^{+0.56}, \quad k_{t\bar{t}+\geq 1b} = 1.00_{-0.08}^{+0.08}, \quad k_{t\bar{t}+\geq 1c} = 1.00_{-0.19}^{+0.19} \quad \text{with the nominal setup.} \quad (8.10)$$

$$\mu_{t\bar{t}H} = 1.00_{-0.60}^{+0.64}, \quad k_{t\bar{t}+\geq 1b} = 1.00_{-0.09}^{+0.10}, \quad k_{t\bar{t}+\geq 1c} = 1.00_{-0.23}^{+0.23} \quad \text{with the onebin setup.} \quad (8.11)$$

$$\mu_{t\bar{t}H} = 1.00_{-0.57}^{+0.61}, \quad k_{t\bar{t}+\geq 1b} = 1.00_{-0.08}^{+0.09}, \quad k_{t\bar{t}+\geq 1c} = 1.00_{-0.20}^{+0.20} \quad \text{with the optimum setup.} \quad (8.12)$$

Dropping the shape information in H_T^{had} and H_T^{all} worsen the sensitivities. The noticeable differences of the systematic pulls among three setups are summarized in Figure 8.36. The largest difference is the c -tagging EV1 which is back towards the nominal value with onebin setup. This pull is from the shape compensation of H_T^{had} in $\text{CR}(t\bar{t}+1b)$ region in the single-lepton channel. Between the onebin and optimum setups, large differences are not seen in the systematic pulls. In

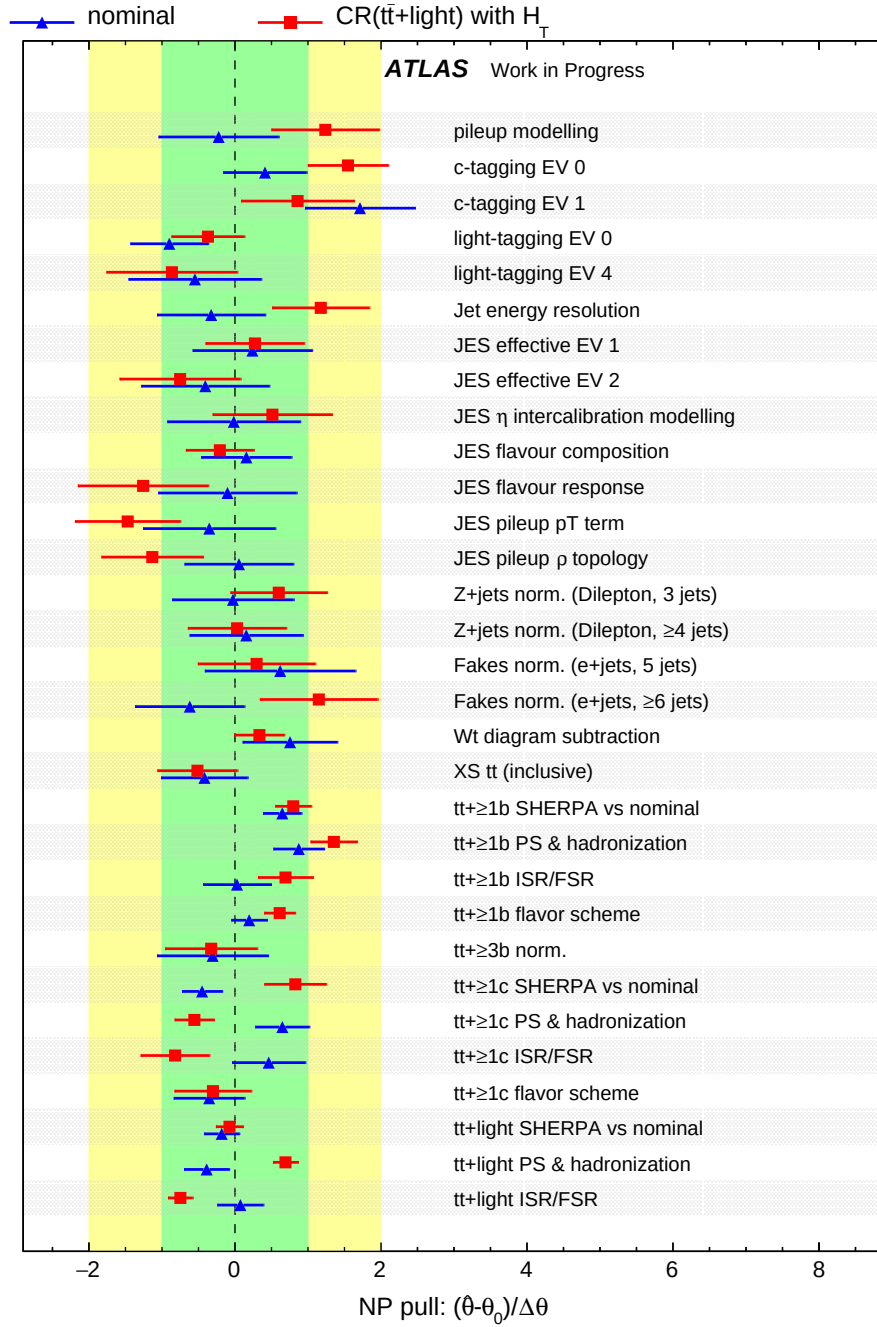


FIGURE 8.32: Highly constrained or pulled systematic variations with the H_T^{had} distribution in $t\bar{t}$ +light control regions. Only systematic uncertainty sources with $> 0.5\sigma$ pulls or $> 20\%$ constraints from the original uncertainty band are shown in the plot. Blue lines show the nominal setups as shown in Tables 8.8 and 8.10. Red lines show the alternative option with H_T^{had} or H_T^{all} distributions in $t\bar{t}$ +light control regions.

this analysis, the optimum option was adopted with least affecting the H_T^{had} and H_T^{all} discrepancies, therefore, all control regions in the dilepton channel and two CR($t\bar{t}+1b$) in the single-lepton channel utilize the normalization. The two CR($t\bar{t}+\geq 1c$) regions in the single-lepton channel are expected to have a significant constraint to the jet systematic uncertainties.

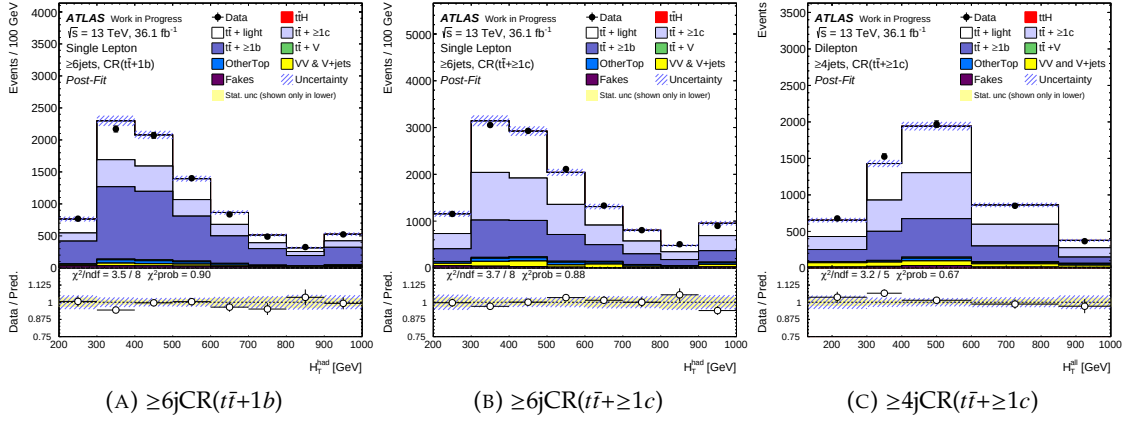


FIGURE 8.33: H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with nominal setup. In the fit, H_T^{had} was used in $\text{CR}(t\bar{t}+1b)$ and $\text{CR}(t\bar{t}+\geq 1c)$ for the single-lepton channel, and H_T^{all} was used in $\text{CR}(t\bar{t}+\geq 1c)$ for the dilepton channel.

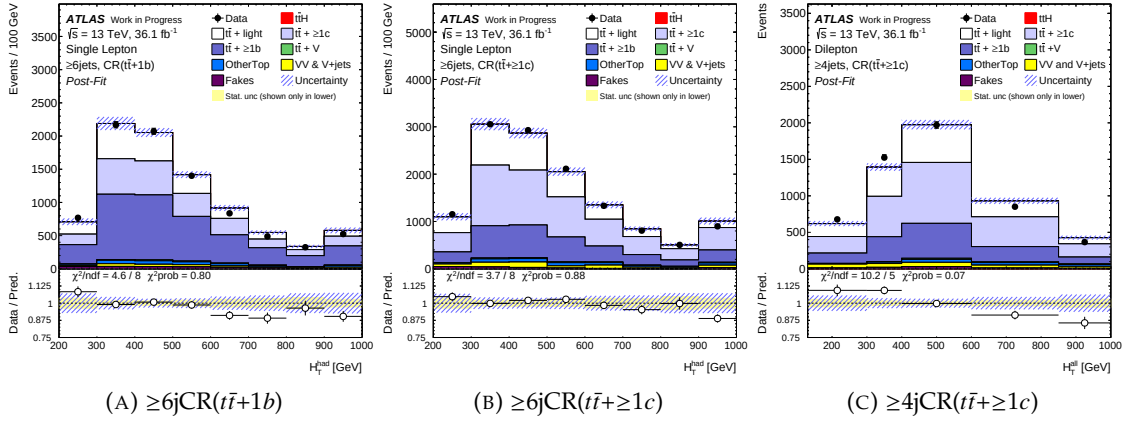


FIGURE 8.34: H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with onebin setup. In the fit, normalization was adjusted in all control regions.

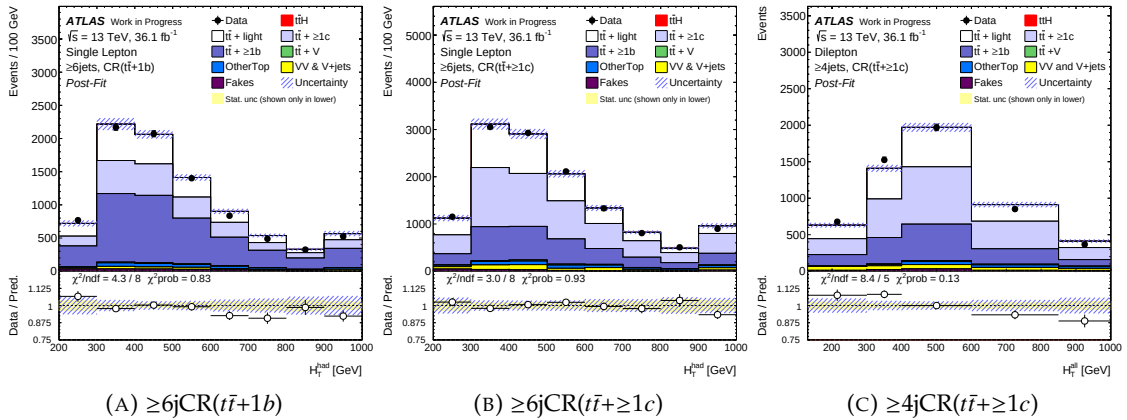


FIGURE 8.35: H_T^{had} and H_T^{all} distributions after b only hypothesis fitting with optimum setup. In the fit, H_T^{had} was used in single-lepton $\text{CR}(t\bar{t}+\geq 1c)$, and normalization was adjusted in other regions.

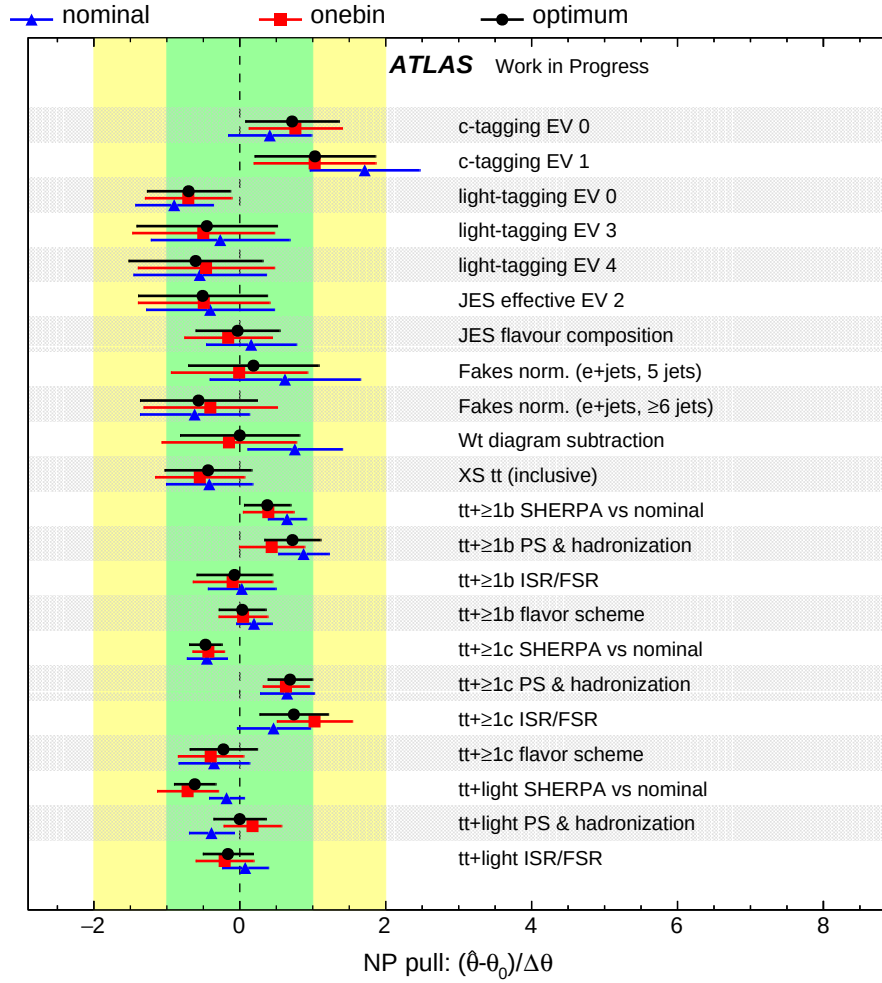


FIGURE 8.36: Highly constrained or pulled systematic variations. Only systematic uncertainty sources with $> 0.5\sigma$ pulls or $> 20\%$ constraints from the original uncertainty band are shown.

8.12 Event Yields for Defined Regions

Event yields in all analysis regions are summarized in Tables 12.4 to 12.6. All expected events were predicted by the Standard Model MC.

TABLE 8.11: Yields of the analysis regions in ≥ 6 jets for the single-lepton channel for 36.1 fb^{-1} . The " $t\bar{t}V$ " includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " has WtZ and tH processes.

	≥ 6 jets, CR($t\bar{t}$ +light)	≥ 6 jets, CR($t\bar{t}$ + $\geq 1c$)	≥ 6 jets, CR($t\bar{t}$ + $1b$)
$t\bar{t}$ + light	$125\,000 \pm 34\,600$	4330 ± 2010	2220 ± 521
$t\bar{t}$ + $\geq 1c$	$28\,400 \pm 7260$	3560 ± 1350	1460 ± 333
$t\bar{t}$ + $\geq 1b$	$13\,100 \pm 1840$	2660 ± 543	3670 ± 503
$t\bar{t}V$	1010 ± 120	118 ± 20.9	70.5 ± 8.54
single-top	5740 ± 1840	355 ± 138	234 ± 81.0
$t\bar{t}WW$	15.0 ± 1.93	1.74 ± 0.343	0.572 ± 0.128
$t\bar{t}t\bar{t}$	18.5 ± 9.29	8.22 ± 4.15	5.30 ± 2.67
non top	6590 ± 2000	649 ± 266	418 ± 127
tH	23.1 ± 2.73	4.78 ± 0.685	5.33 ± 0.641
$t\bar{t}H$	450 ± 49.5	102 ± 13.5	100 ± 12.5
Total	$181\,000 \pm 39\,500$	$11\,800 \pm 3210$	8180 ± 1140
Data	181706	12778	8576

	≥ 6 jets, SR($t\bar{t}$ + $\geq 2b$, Lo)	≥ 6 jets, SR($t\bar{t}$ + $\geq 2b$, Hi)	≥ 6 jets, SR($t\bar{t}H$)
$t\bar{t}$ + light	745 ± 371	209 ± 209	14.1 ± 10.2
$t\bar{t}$ + $\geq 1c$	884 ± 350	346 ± 104	53.3 ± 32.9
$t\bar{t}$ + $\geq 1b$	2100 ± 422	1750 ± 370	1010 ± 240
$t\bar{t}V$	51.2 ± 7.56	40.8 ± 5.85	25.8 ± 3.77
single-top	124 ± 56.5	80.3 ± 46.6	33.2 ± 17.3
$t\bar{t}WW$	0.477 ± 0.117	0.253 ± 0.0854	0.0704 ± 0.0408
$t\bar{t}t\bar{t}$	6.66 ± 3.36	6.33 ± 3.19	4.13 ± 2.08
non top	139 ± 49.1	37.6 ± 15.7	16.9 ± 6.12
tH	3.64 ± 0.501	3.52 ± 0.485	2.25 ± 0.298
$t\bar{t}H$	84.7 ± 10.7	81.2 ± 10.5	62.5 ± 11.1
Total	4140 ± 859	2550 ± 512	1220 ± 255
Data	4698	2641	1222

TABLE 8.12: Yields of the analysis regions in 5 jets for the single-lepton channel for 36.1 fb^{-1} . The " $t\bar{t}V$ " includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " has WtZ and tH processes.

	5 jets, CR($t\bar{t}$ +light)	5 jets, CR($t\bar{t}+\geq 1c$)	5 jets, CR($t\bar{t}+1b$)
$t\bar{t}$ + light	$197\,000 \pm 26\,000$	2580 ± 721	4250 ± 921
$t\bar{t} + \geq 1c$	$27\,500 \pm 4360$	1280 ± 503	1770 ± 273
$t\bar{t} + \geq 1b$	$11\,300 \pm 1150$	791 ± 128	3400 ± 443
$t\bar{t}V$	589 ± 57.1	23.2 ± 4.20	48.1 ± 5.85
single-top	$10\,100 \pm 2310$	192 ± 53.7	366 ± 97.8
$t\bar{t}WW$	2.91 ± 0.457	0.199 ± 0.0956	0.148 ± 0.0640
$t\bar{t}t\bar{t}$	0.855 ± 0.461	0.0568 ± 0.0378	0.263 ± 0.141
non top	$10\,900 \pm 3220$	252 ± 88.7	495 ± 157
tH	19.3 ± 1.62	1.39 ± 0.225	6.29 ± 0.576
$t\bar{t}H$	224 ± 22.9	18.6 ± 2.56	68.1 ± 7.63
Total	$258\,000 \pm 29\,600$	5140 ± 1110	$10\,400 \pm 1290$
Data	259320	5465	11095
	5 jets, SR($t\bar{t}+\geq 2b$)	5 jets, SR($t\bar{t}H$)	Boosted SR
$t\bar{t}$ + light	502 ± 209	15.2 ± 32.6	177 ± 123
$t\bar{t} + \geq 1c$	436 ± 92.7	29.6 ± 16.7	168 ± 70.1
$t\bar{t} + \geq 1b$	1230 ± 204	273 ± 53.2	236 ± 89.2
$t\bar{t}V$	19.9 ± 2.92	6.42 ± 1.33	16.1 ± 2.98
single-top	99.9 ± 43.8	17.2 ± 5.95	27.1 ± 19.0
$t\bar{t}WW$	0.	0.	0.519 ± 0.135
$t\bar{t}t\bar{t}$	0.205 ± 0.114	0.0360 ± 0.0201	4.07 ± 2.06
non top	112 ± 40.3	12.4 ± 8.24	47.0 ± 17.0
tH	3.26 ± 0.376	1.41 ± 0.183	1.97 ± 0.283
$t\bar{t}H$	40.1 ± 5.18	16.0 ± 2.16	16.9 ± 1.99
Total	2440 ± 393	371 ± 69.0	695 ± 200
Data	2798	426	740

TABLE 8.13: Yields of the analysis regions for the dilepton channel for 36.1 fb^{-1} . The " $t\bar{t}V$ " includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " has WtZ and tH processes.

	3 jets, CR($t\bar{t}$ +light)	3 jets, CR($t\bar{t}+\geq 1b$)	≥ 4 jets, CR($t\bar{t}$ +light)	≥ 4 jets, CR($t\bar{t}+\geq 1c$)
$t\bar{t}$ + light	$63\,100 \pm 5680$	292 ± 108	$42\,500 \pm 9760$	1730 ± 728
$t\bar{t} + \geq 1c$	4770 ± 2130	364 ± 155	6310 ± 2810	1410 ± 587
$t\bar{t} + \geq 1b$	2130 ± 233	714 ± 143	2510 ± 286	1080 ± 121
$t\bar{t}V$	113 ± 30.6	6.76 ± 26.9	346 ± 183	52.0 ± 42.1
single-top	2280 ± 692	42.1 ± 12.8	1430 ± 459	110 ± 45.8
$t\bar{t}WW$	0.486 ± 0.0875	$0.008\,78 \pm 0.008\,78$	6.47 ± 0.266	1.16 ± 0.138
$t\bar{t}t\bar{t}$	0.0582 ± 0.0152	$0.0122 \pm 0.005\,35$	3.82 ± 0.0878	2.20 ± 0.0710
non top	3980 ± 1250	67.3 ± 25.2	3250 ± 951	301 ± 97.3
tH	2.66 ± 0.277	0.546 ± 0.113	6.05 ± 0.571	1.78 ± 0.209
$t\bar{t}H$	32.2 ± 3.85	8.66 ± 1.08	114 ± 11.3	35.3 ± 3.66
Total	$76\,400 \pm 6750$	1500 ± 260	$56\,400 \pm 10\,800$	4720 ± 1070
Data	76025	1744	55627	5389
	≥ 4 jets, SR($t\bar{t}+1b$)	≥ 4 jets, SR($t\bar{t}+\geq 2b$)	≥ 4 jets, SR($t\bar{t}H$)	
$t\bar{t}$ + light	83.4 ± 41.0	248 ± 115	6.41 ± 9.95	
$t\bar{t} + \geq 1c$	235 ± 61.1	345 ± 205	12.6 ± 9.32	
$t\bar{t} + \geq 1b$	819 ± 88.8	590 ± 97.9	247 ± 60.9	
$t\bar{t}V$	15.1 ± 35.3	21.7 ± 38.1	6.62 ± 56.0	
single-top	27.7 ± 12.3	31.8 ± 18.9	3.82 ± 2.77	
$t\bar{t}WW$	0.0901 ± 0.0275	0.449 ± 0.0866	0.0121 ± 0.0171	
$t\bar{t}t\bar{t}$	1.29 ± 0.0524	3.92 ± 0.115	1.71 ± 0.0652	
non top	45.3 ± 11.7	77.7 ± 27.0	7.57 ± 2.21	
tH	0.751 ± 0.114	1.14 ± 0.215	0.441 ± 0.0704	
$t\bar{t}H$	21.9 ± 2.54	29.1 ± 4.31	15.6 ± 2.52	
Total	1250 ± 143	1350 ± 324	302 ± 85.7	
Data	1467	1444	319	

Chapter 9

MultiVariate Discriminants for $t\bar{t}H$ Signal Enhancement

The most significant uncertainty on the sensitivity of $t\bar{t}H(H \rightarrow b\bar{b})$ is the modeling of $t\bar{t} + \geq 1b$, which is a non-reducible background source due to the same final state as $t\bar{t}H(H \rightarrow b\bar{b})$. To separate $t\bar{t}H(H \rightarrow b\bar{b})$ from $t\bar{t} + \geq 1b$ in the signal regions, a full reconstruction of $t\bar{t}H$ system in both single-lepton and dilepton channels was conducted. In the single-lepton channel, additional discriminants were also utilized for better signal separation from $t\bar{t}$ +jets background events. They were implemented separately but combined into one BDT score with event kinematic variables as described in the final analysis in Chapter 10.

9.1 RecoBDT

RecoBDT is a multivariate analysis using BDT algorithm to reconstruct a $t\bar{t}H$ -system from reconstructed objects by solving the assignment of reconstructed jets to the partons. For the RecoBDT training, combinations with all truth partons correctly assigned to reconstructed jets were treated as the signal, while other combinations were treated as the background using $t\bar{t}H(H \rightarrow b\bar{b})$ simulation sample. If all $t\bar{t}H$ objects are used for input variables into RecoBDT, background events are also biased to the $t\bar{t}H$ signal. For instance, the mass of the reconstructed Higgs-boson tends to show a peak at m_{Higgs} even for the background events. On the other hand, if only $t\bar{t}$ objects are used for input variables with assigning remaining jets to the Higgs-boson, the reconstruction efficiency is not as high as RecoBDT with Higgs information, but non-biased Higgs variables can be obtained. Therefore, to maximize the performance of RecoBDT, two sets of input variables are prepared. One uses as much information as possible to obtain the correct assignment of all jets (RecoBDT with H), while the other uses only $t\bar{t}$ information not to bias the Higgs-boson properties (RecoBDT w/o H). Each RecoBDT provides the best jet permutation maximizing the output score of the RecoBDT. They were trained in five inclusive signal regions with following requirements:

- single-lepton channel in ≥ 6 jets
- single-lepton channel in 5 jets with the assumption of one jet missed from the W_{had} decay
- dilepton channel in ≥ 4 jets with ≥ 4 b -jets passing 85% working point threshold
- dilepton channel in ≥ 4 jets with 3 b -jets passing 85% working point threshold

9.1.1 Single-lepton Channel

As discussed in Chapter 3, $q2_{\text{had}W}$ tends to be missed in the reconstructed level. Therefore, RecoBDT was trained with truth-matched events where all four truth b -partons and at least one

light-quark parton from hadronic W boson decay were matched within $\Delta R < 0.35$ to the reconstructed jets in both ≥ 6 j and 5 j regions. RecoBDT is expected to solve jets assignment other than $q2_{\text{had}W}$ correctly even when $q2_{\text{had}W}$ assignment is wrong. In the 5 j region, RecoBDT assumes $q2_{\text{had}W}$ is always missed, and W_{had} is constructed only by $q1_{\text{had}W}$.

In addition to the assignment of jets, RecoBDT in the single-lepton channel also chose the best solution of the neutrino p_z . For the reconstruction of leptonic W boson, E_T^{miss} was used as the transverse momentum of the neutrino and neutrino p_z was calculated with E_T^{miss} and lepton four-momentum constraining the invariant mass of them to the W boson mass. This can be solved with the quadratic equation as,

$$p_{z,\nu} = \frac{1}{2} \frac{p_{z,\ell}\beta \pm \sqrt{\Delta}}{E_\ell^2 - p_{z,\ell}^2} \quad (9.1)$$

where

$$\beta = m_W^2 - m_\ell^2 + 2p_{x,\ell}p_{x,\nu} + 2p_{y,\ell}p_{y,\nu} \quad (9.2)$$

$$\Delta = E_\ell^2 \left(\beta^2 + (2p_{z,\ell}p_{T,\nu})^2 - (2E_\ell p_{T,\nu})^2 \right). \quad (9.3)$$

If there are two real solutions for p_z , both of them are used to calculate with the permutation of jets, and the higher RecoBDT scored solution was taken as the correct one. If there are two imaginary solutions, the real part is taken as the solution.

Furthermore, in order to profit from b -tagging information, four largest b -tagging scored jets were assigned only to b -jet candidates in the permutation, and only remaining jets were assigned to light-jet candidates. Therefore, the number of permutations considered in each event is 24, 36 or 6 for nJets = 6, 7, or 5, respectively.

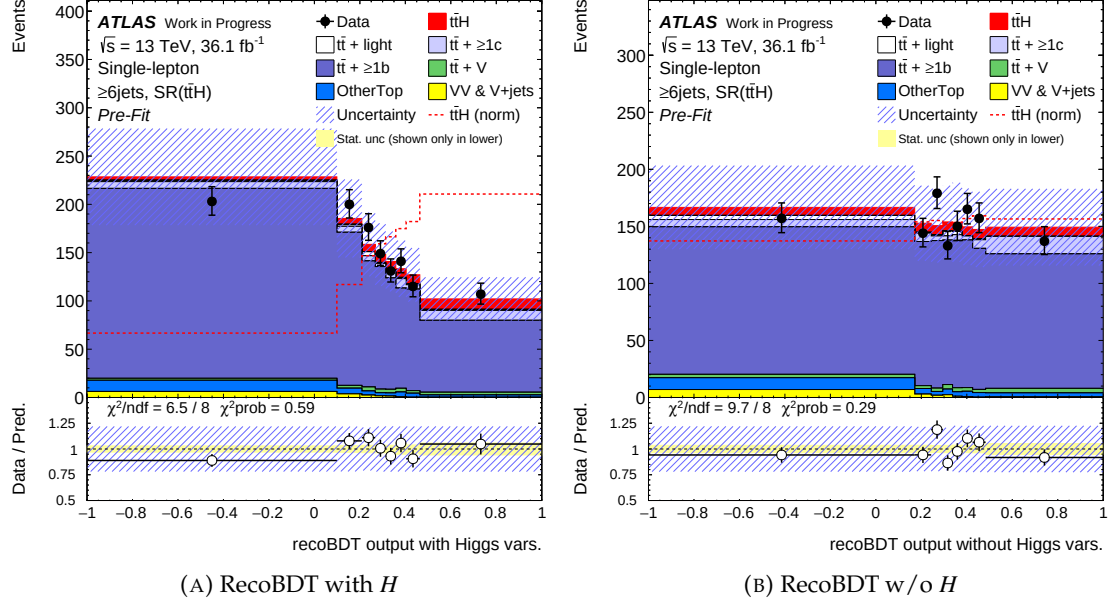
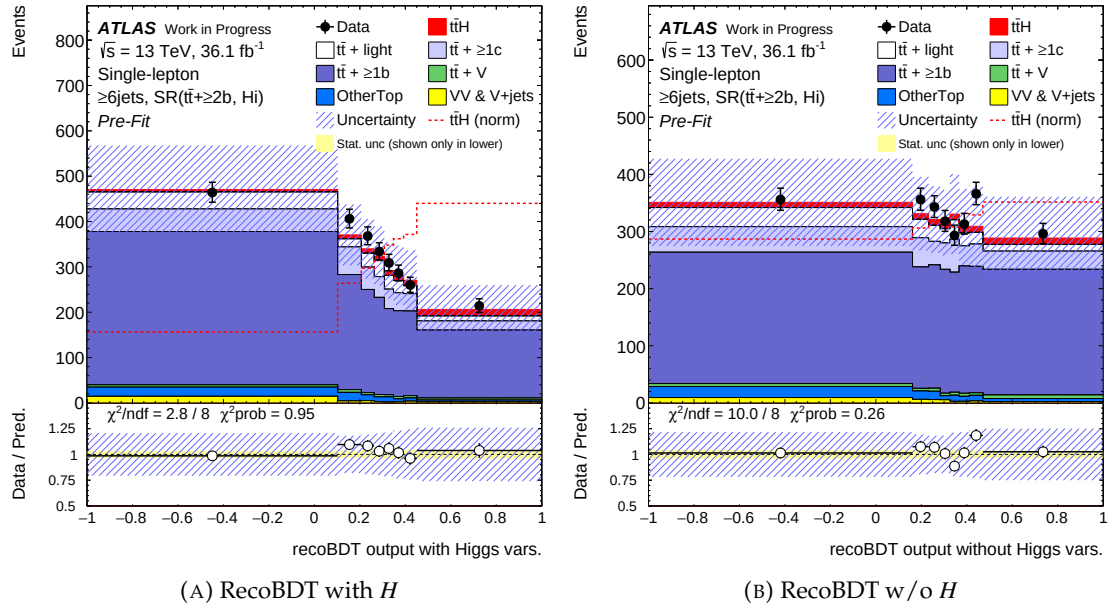
Input variables for RecoBDTs are summarized in Table 9.1, and their output scores in each signal region are shown in Figures 9.1 to 9.5. By the Higgs constraint, the output score of the RecoBDT with H has a separation power between the $t\bar{t}H$ signal and the $t\bar{t}$ +jets background. The obtained best permutations were checked whether they are correctly assigned. The reconstruction efficiencies are summarized in Table 9.2. In the most sensitive signal region, $\geq 6\text{jSR}(t\bar{t}H)$, the $b\bar{b}$ pair from the Higgs boson decay is correctly assigned as 42% and 28% by RecoBDT with and w/o H , respectively. Using the best jet permutation, the distributions used as the input variables are also shown in Figures 9.6 to 9.16. The invariant masses of the lepTop, W_{had} , and hadTop have peaks at their expected masses in ≥ 6 jets. In the 5 j regions, the W_{had} is constructed by one light jet assigned as $q1_{\text{had}W}$ where $q2_{\text{had}W}$ is assumed to be missed. Therefore, the invariant masses of the W_{had} and hadTop in 5 jets are not peaked at ideal W and top mass. The mass distribution of the $b\bar{b}$ pair from the Higgs boson decay is biased to the Higgs mass by the RecoBDT with H in all processes, as expected. On the other hand, the mass distribution of the $b\bar{b}$ pair from the Higgs boson decay is not biased by the RecoBDT w/o H , and has a good separation between the $t\bar{t}H$ signal and other background processes.

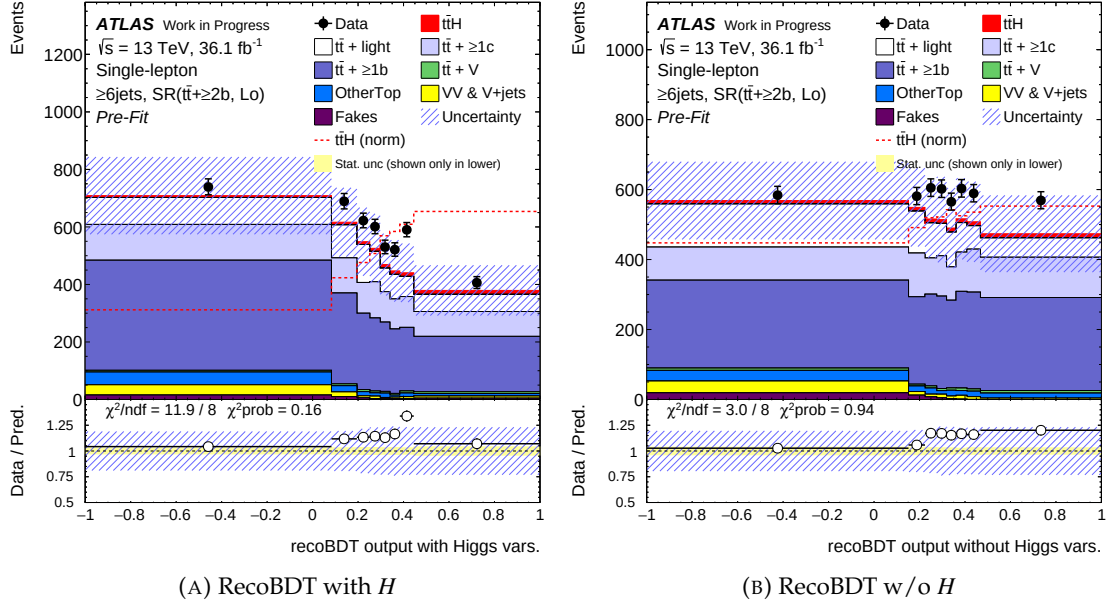
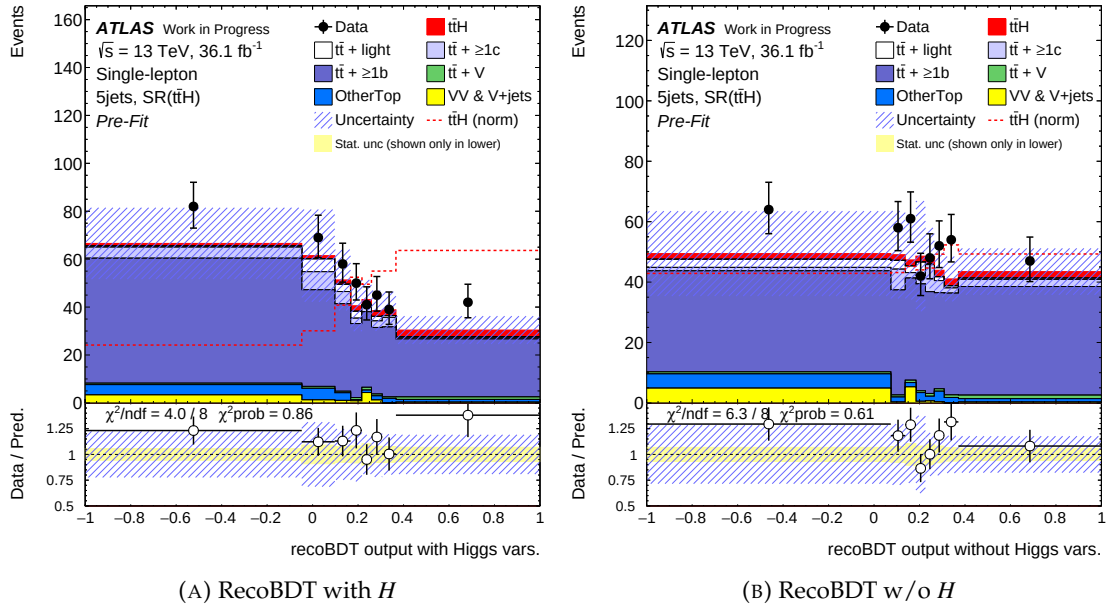
TABLE 9.1: Input variables to RecoBDT to assign reconstructed objects to $t\bar{t}H$ partons in the single lepton channel. Variables with $\checkmark$ are used in both RecoBDTs with/without Higgs information. The variables with $\bullet$ ($\circ$) mean that they are used only in RecoBDT with(without) Higgs information.

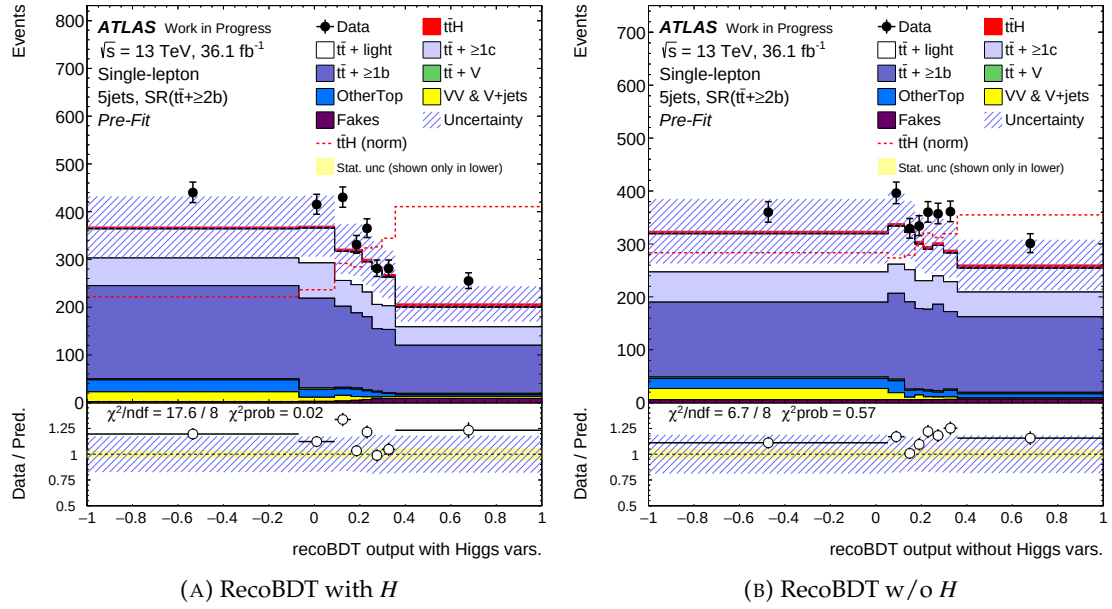
Variable	$\geq 6\text{jets}$	5jets
Topological information from $t\bar{t}$ decay		
lepTop mass	$\checkmark$	$\checkmark$
hadTop mass	$\checkmark$	$\checkmark$
W_{had} mass	$\checkmark$	—
mass of W_{had} and b_{hadTop} system	$\checkmark$	$\checkmark$
mass of W_{lep} and b_{lepTop} system	$\checkmark$	$\checkmark$
$\Delta R(W_{\text{had}}, b_{\text{lepTop}})$	$\checkmark$	$\checkmark$
$\Delta R(W_{\text{had}}, b_{\text{hadTop}})$	$\checkmark$	$\checkmark$
$\Delta R(\ell, b_{\text{lepTop}})$	$\checkmark$	$\checkmark$
$\Delta R(\ell, b_{\text{hadTop}})$	$\checkmark$	$\checkmark$
$\Delta R(b_{\text{lepTop}}, b_{\text{hadTop}})$	$\checkmark$	$\checkmark$
$\Delta R(q_{1\text{hadW}}, q_{2\text{hadW}})$	$\checkmark$	—
$\Delta R(b_{\text{hadTop}}, q_{1\text{hadW}})$	$\checkmark$	—
$\Delta R(b_{\text{hadTop}}, q_{2\text{hadW}})$	$\checkmark$	—
min. $\Delta R(b_{\text{hadTop}}, q_{i\text{hadW}})$	$\checkmark$	—
$\Delta R(\ell, b_{\text{lepTop}}) - \text{min. } \Delta R(b_{\text{hadTop}}, q_{i\text{hadW}})$	$\checkmark$	$\checkmark$
Topological information from Higgs decay		
Higgs boson mass	$\bullet$	$\bullet$
Mass of Higgs and q_1	$\bullet$	$\bullet$
$\Delta R(b_{1\text{Higgs}}, b_{2\text{Higgs}})$	$\bullet$	$\bullet$
$\Delta R(b_{1\text{Higgs}}, \ell)$	$\bullet$	$\bullet$
$\Delta R(b_{1\text{Higgs}}, b_{\text{lepTop}})$	—	$\bullet$
$\Delta R(b_{1\text{Higgs}}, b_{\text{hadTop}})$	—	$\bullet$

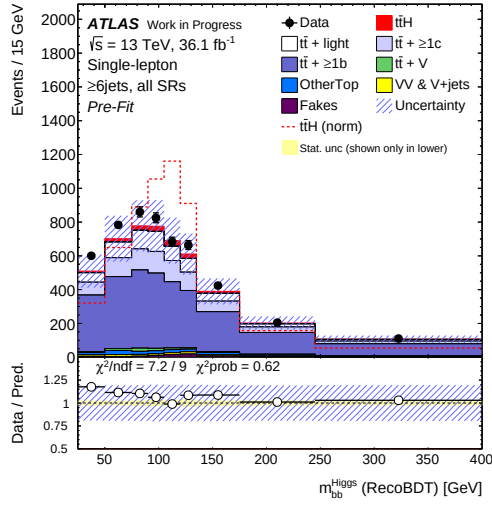
TABLE 9.2: Reconstruction efficiency. The probabilities to match the jet assignment correctly by truth matching and by the RecoBDT. The row "all" means all events are used to calculate the probabilities, while the row "truth matched" means only fully truth-matched events are used where all partons are matched with $\Delta R < 0.35$ to individual jet (4b and two light quarks in ≥ 6 jets, 4b and one of q_{hadW} in 5 jets).

partons	$b1_{\text{Higgs}}$	$b2_{\text{Higgs}}$	$b1_{t\bar{t}}$	$b2_{t\bar{t}}$	$q1_{\text{hadW}}$	$q2_{\text{hadW}}$	Higgs	4b+ q_{hadW}	all
RecoBDT with H : nJets ≥ 6 with nBTags _{85%} ≥ 4									
all	0.48	0.45	0.56	0.44	0.41	0.27	0.33	0.14	0.074
truth matched	0.55	0.55	0.64	0.56	0.55	0.55	0.43	0.25	0.22
RecoBDT with H : ≥ 6 jets signal regions									
all	0.51	0.49	0.60	0.48	0.44	0.28	0.37	0.17	0.09
truth matched	0.57	0.57	0.67	0.58	0.57	0.57	0.45	0.28	0.24
RecoBDT with H : ≥ 6 SR($t\bar{t}H$)									
all	0.56	0.55	0.65	0.55	0.49	0.30	0.42	0.23	0.12
truth matched	0.61	0.61	0.70	0.63	0.61	0.60	0.50	0.33	0.28
RecoBDT w/o H : nJets ≥ 6 with nBTags _{85%} ≥ 4									
all	0.42	0.38	0.56	0.4	0.42	0.27	0.21	0.1	0.055
truth matched	0.46	0.47	0.63	0.48	0.56	0.55	0.29	0.18	0.16
RecoBDT w/o H : ≥ 6 jets signal regions									
all	0.43	0.41	0.59	0.43	0.45	0.29	0.24	0.12	0.067
truth matched	0.48	0.48	0.65	0.50	0.59	0.58	0.30	0.20	0.18
RecoBDT w/o H : ≥ 6 SR($t\bar{t}H$)									
all	0.45	0.46	0.62	0.47	0.50	0.30	0.28	0.16	0.087
truth matched	0.51	0.51	0.68	0.53	0.62	0.61	0.34	0.23	0.20
RecoBDT with H : nJets= 5 with nBTags _{85%} ≥ 4									
all	0.46	0.43	0.54	0.41		0.62	0.30	0.17	
truth matched	0.55	0.58	0.66	0.57		0.87	0.44	0.37	
RecoBDT with H : 5jSR($t\bar{t}H$)									
all	0.55	0.57	0.64	0.55		0.73	0.42	0.30	
truth matched	0.60	0.63	0.71	0.63		0.95	0.50	0.44	
RecoBDT w/o H : nJets=5 with nBTags _{85%} ≥ 4									
all	0.38	0.36	0.53	0.36		0.62	0.18	0.11	
truth matched	0.44	0.51	0.61	0.46		0.87	0.29	0.25	
RecoBDT w/o H : 5jSR($t\bar{t}H$)									
all	0.44	0.47	0.59	0.46		0.73	0.26	0.20	
truth matched	0.48	0.54	0.65	0.52		0.95	0.32	0.29	

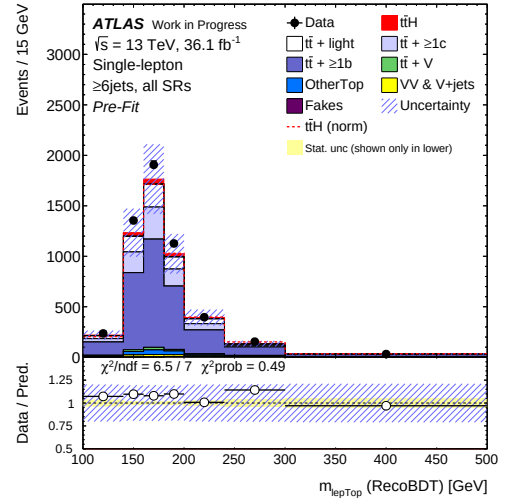
FIGURE 9.1: RecoBDT scores in $\geq 6jSR(tH)$ FIGURE 9.2: RecoBDT scores in $\geq 6jSR(tH+\geq 2b, Hi)$

FIGURE 9.3: RecoBDT scores in $\geq 6jSR(t\bar{t}+\geq 2b, Lo)$ FIGURE 9.4: RecoBDT scores in $5jSR(t\bar{t}H)$

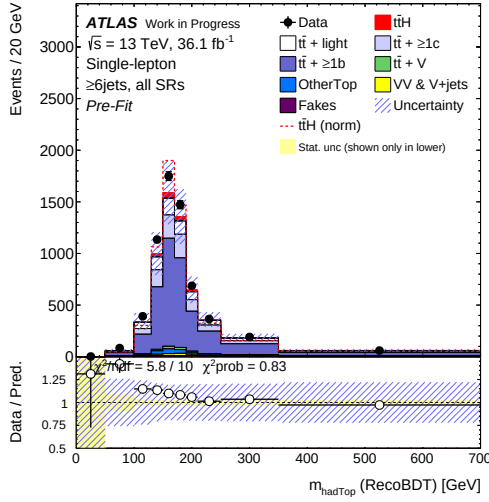
FIGURE 9.5: RecoBDT scores in 5jSR($t\bar{t} + \geq 2b$, Hi)



(A) Higgs boson mass



(B) lepTop mass



(C) hadTop mass

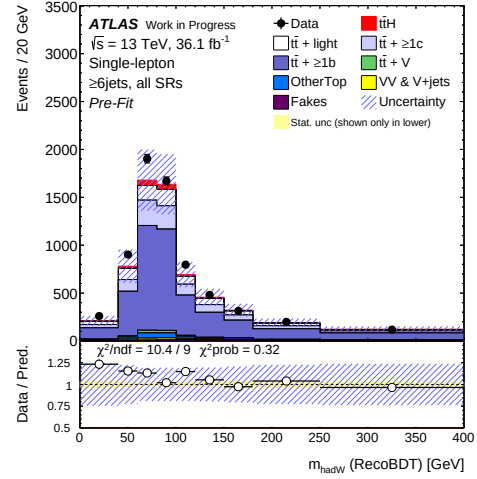
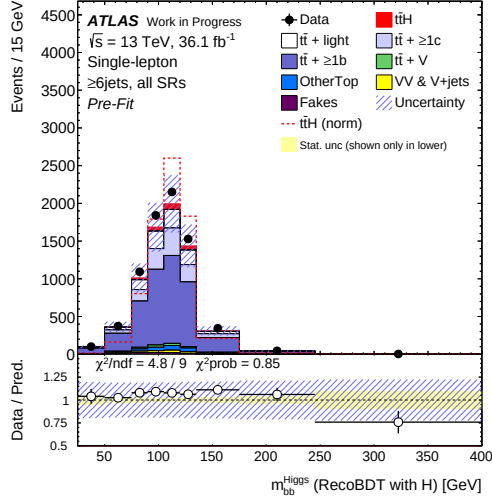
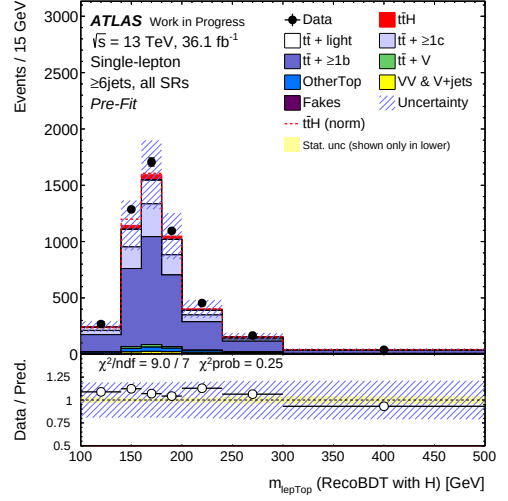
(D) W_{had} mass

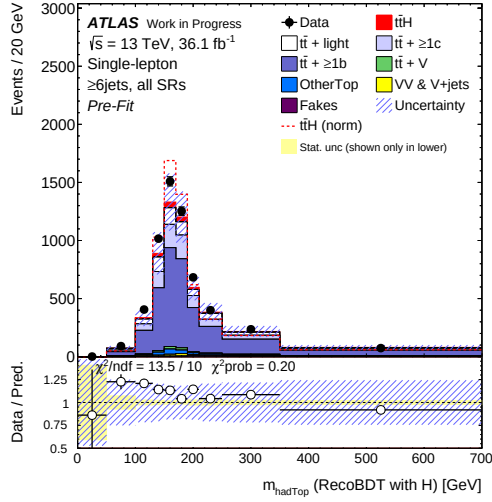
FIGURE 9.6: Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT w/o H in ≥ 6 jets



(A) Higgs boson mass



(B) lepTop mass



(C) hadTop mass

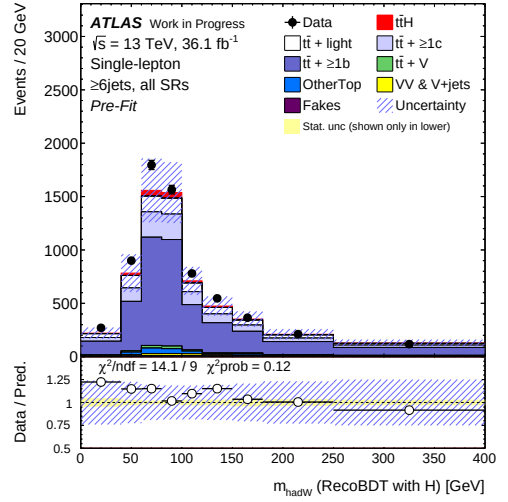
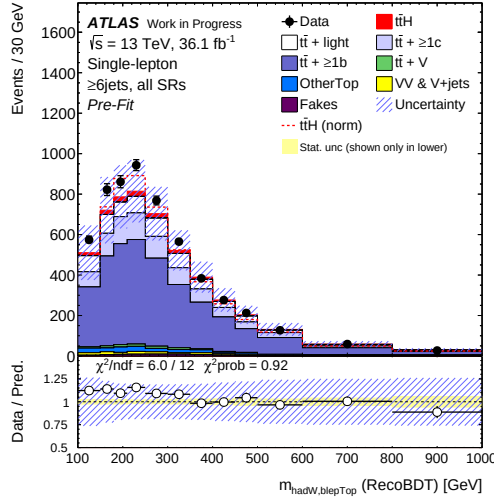
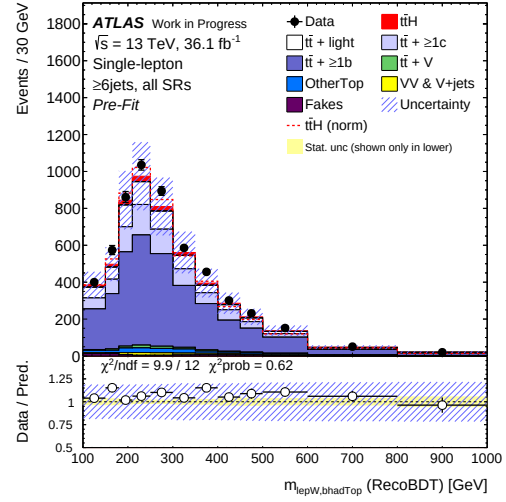
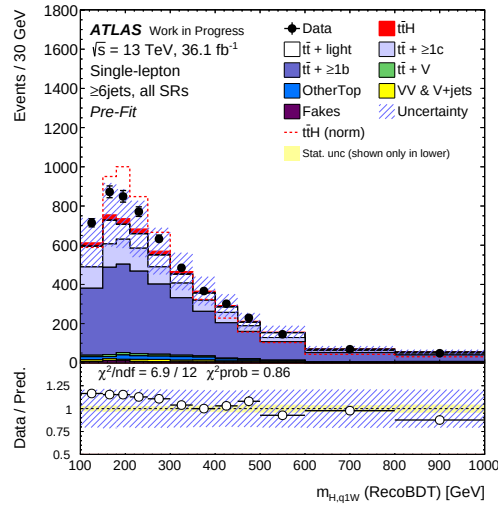
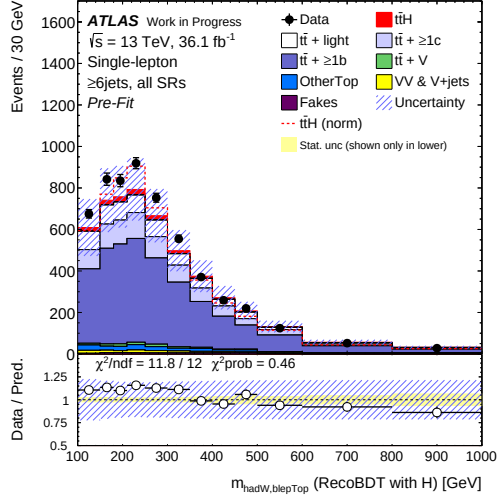
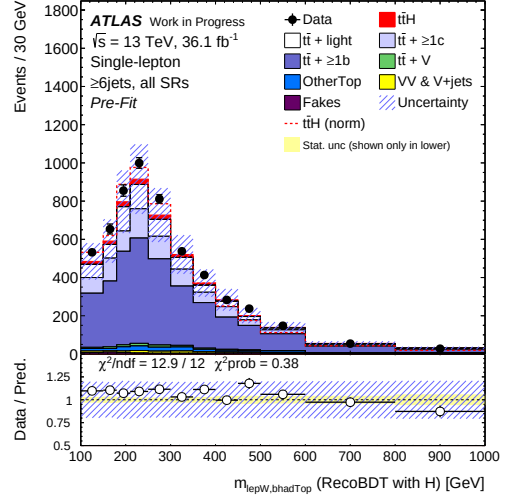
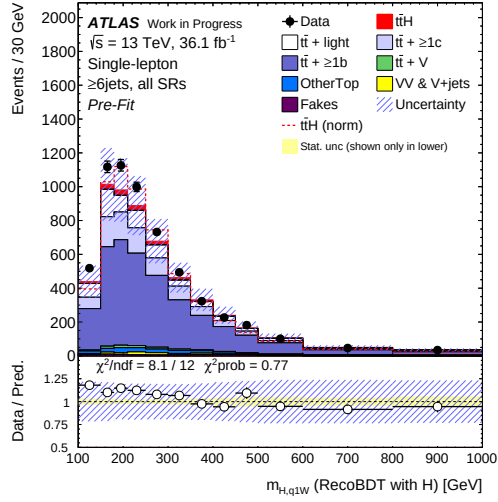
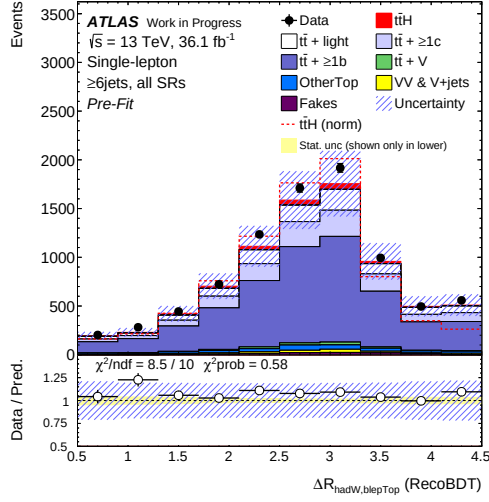
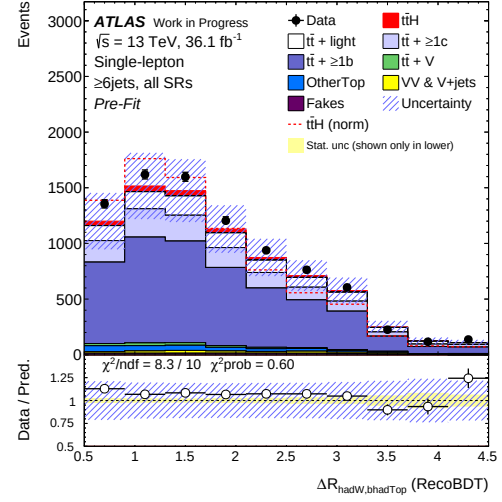
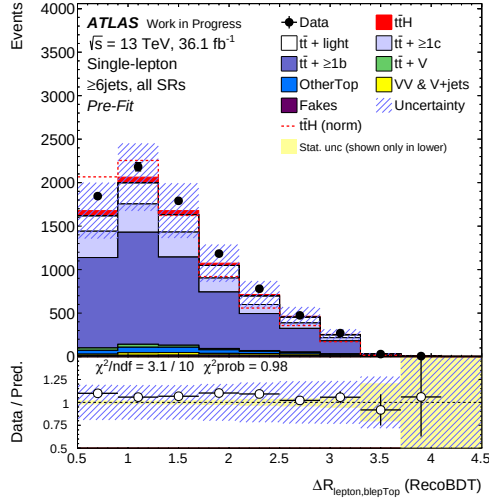
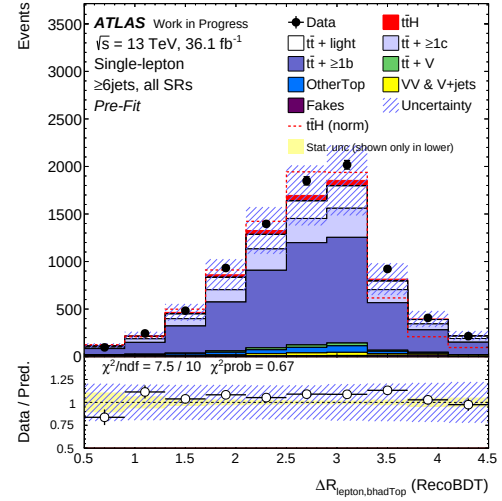
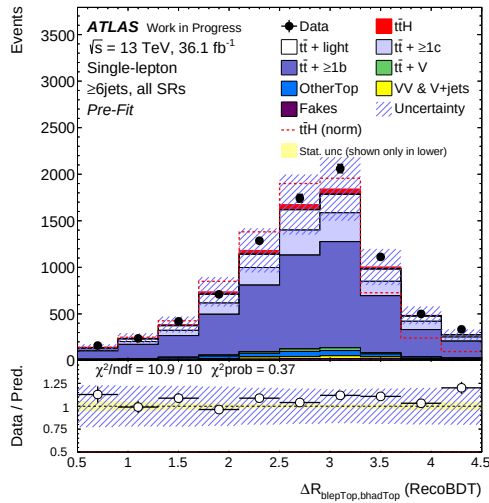
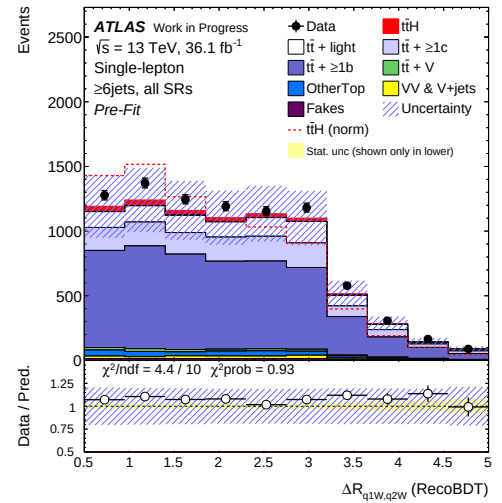
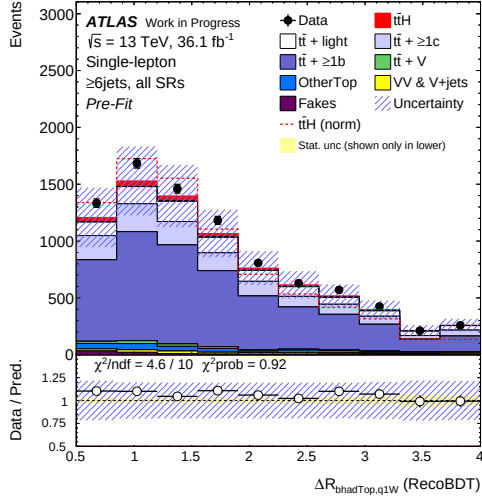
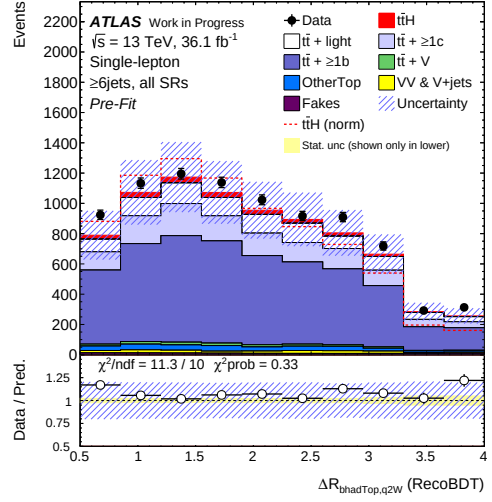
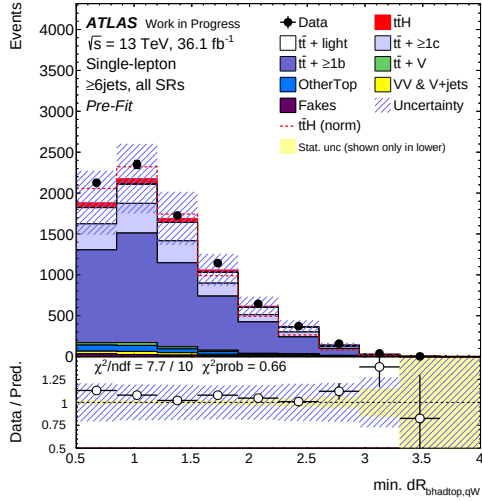
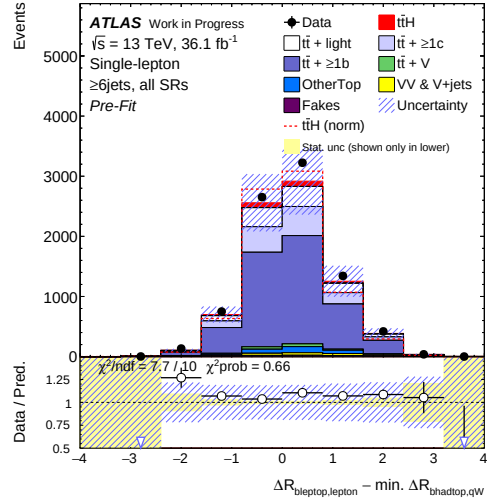
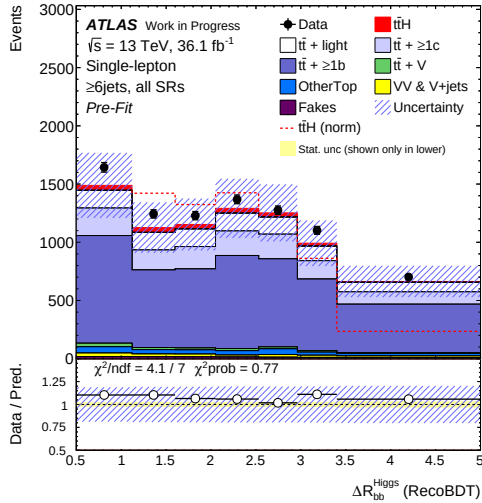
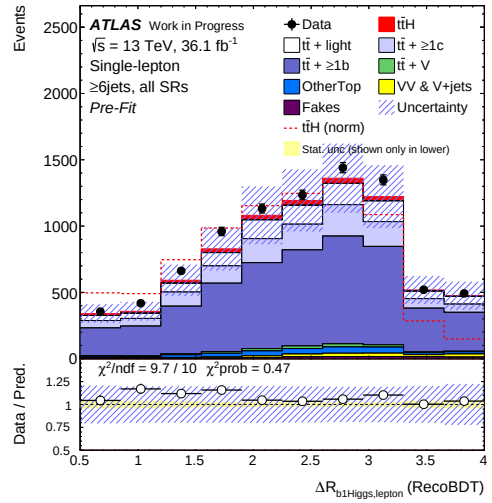
(D) W_{had} mass

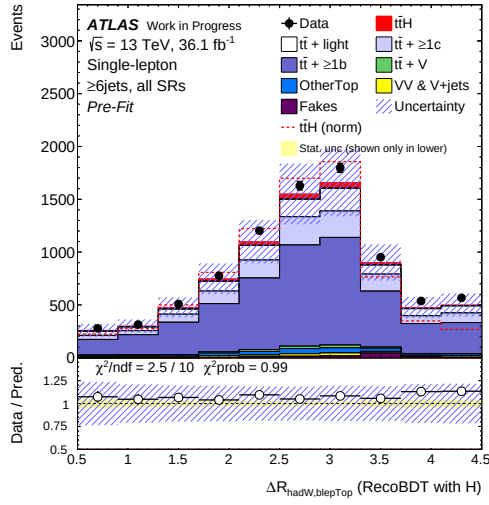
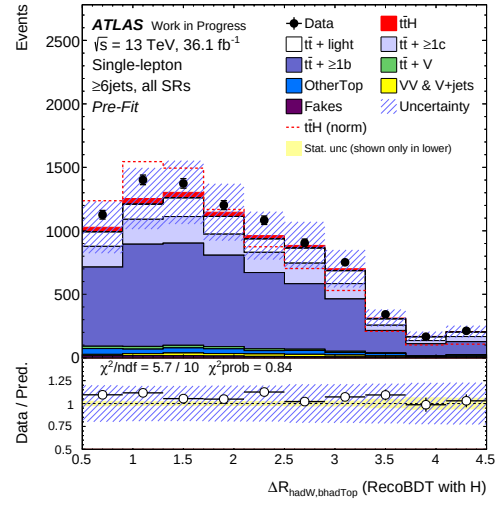
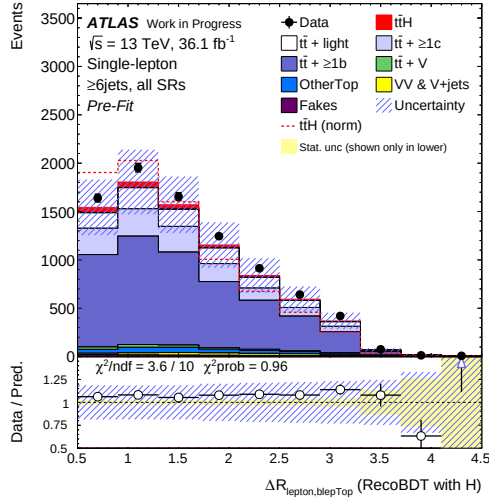
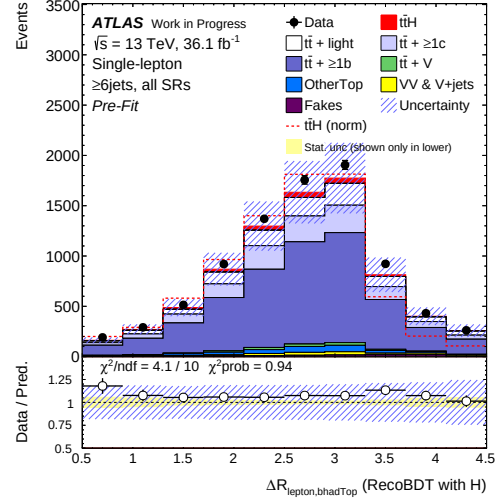
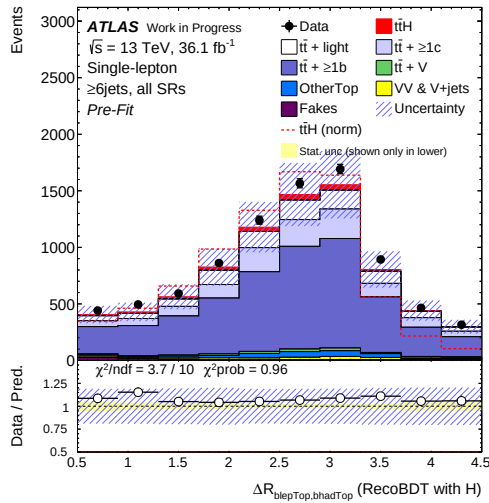
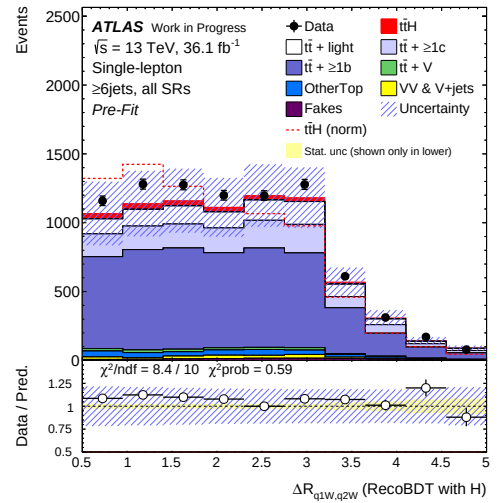
FIGURE 9.7: Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT with H in ≥ 6 jets

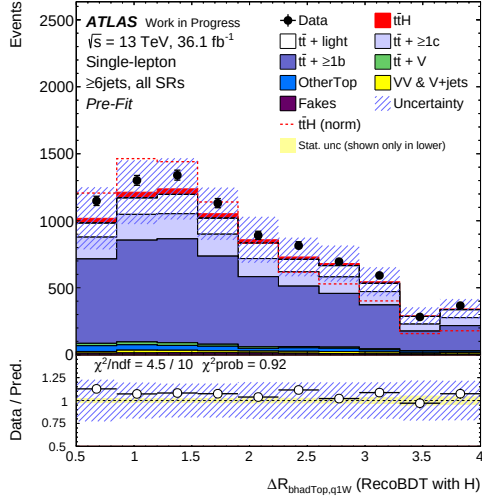
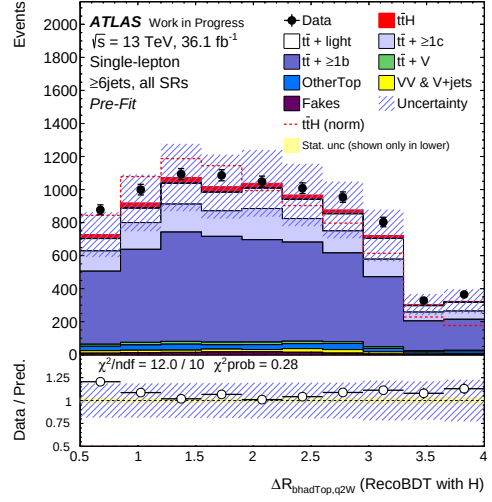
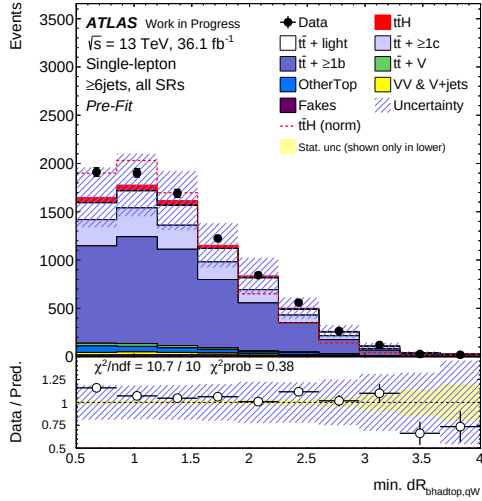
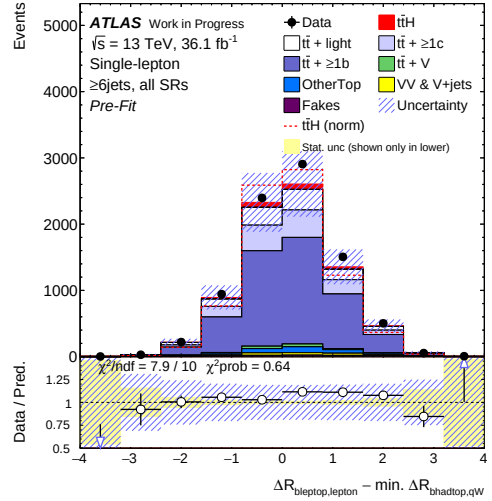
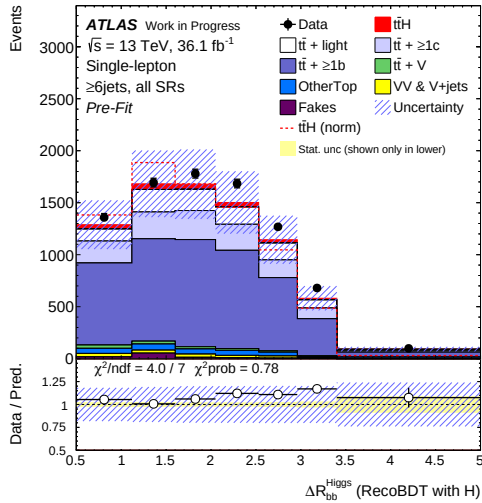
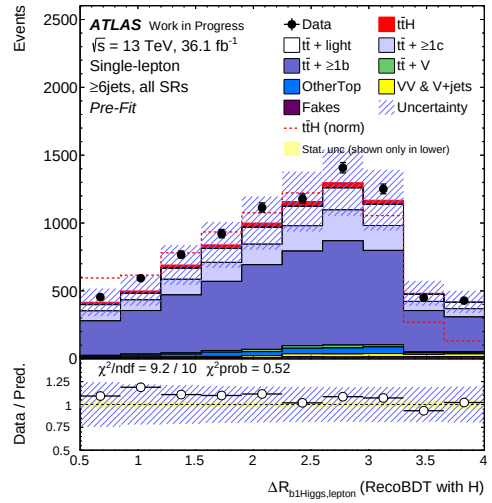
(A) mass of W_{had} and b_{hadTop} system(B) mass of W_{lep} and b_{lepTop} system(C) mass of Higgs and $q1_{\text{had}W}$ systemFIGURE 9.8: Masses of two-object systems constructed by the RecoBDT w/o H in ≥ 6 jets

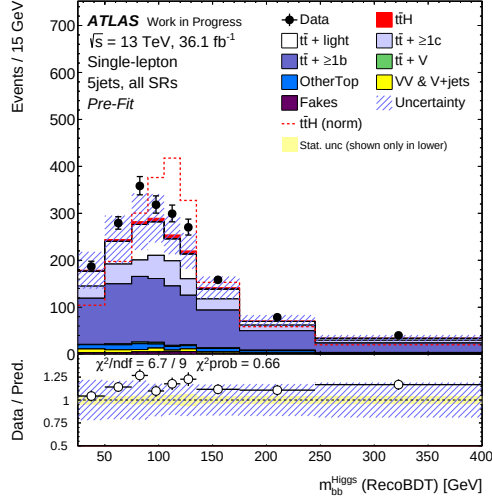
(A) mass of W_{had} and b_{hadTop} system(B) mass of W_{lep} and b_{lepTop} system(C) mass of Higgs and $q1_{\text{had}W}$ systemFIGURE 9.9: Masses of two-object systems constructed by the RecoBDT with H in ≥ 6 jets

(A) $\Delta R(W_{\text{had}}, b_{\text{lepTop}})$ (B) $\Delta R(W_{\text{had}}, b_{\text{hadTop}})$ (C) $\Delta R(\ell, b_{\text{lepTop}})$ (D) $\Delta R(\ell, b_{\text{hadTop}})$ (E) $\Delta R(b_{\text{lepTop}}, b_{\text{hadTop}})$ (F) $\Delta R(q_{1\text{hadW}}, q_{2\text{hadW}})$ FIGURE 9.10: ΔR variables constructed by the RecoBDT w/o H in ≥ 6 jets (1)

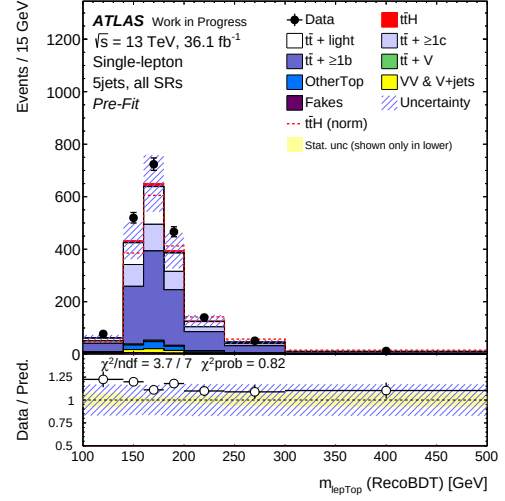
(A) $\Delta R(b_{\text{hadTop}}, q_{1\text{hadW}})$ (B) $\Delta R(b_{\text{hadTop}}, q_{2\text{hadW}})$ (C) min. $\Delta R(b_{\text{hadTop}}, q_{1\text{hadW}})$ (D) $\Delta R(\ell, b_{\text{lepTop}}) - \min. \Delta R(b_{\text{hadTop}}, q_{1\text{hadW}})$ (E) $\Delta R(b_{1\text{Higgs}}, b_{2\text{Higgs}})$ (F) $\Delta R(b_{1\text{Higgs}}, \ell)$ FIGURE 9.11: ΔR variables constructed by the RecoBDT w/o H in ≥ 6 jets (2)

(A) $\Delta R(W_{\text{had}}, b_{\text{lepTop}})$ (B) $\Delta R(W_{\text{had}}, b_{\text{hadTop}})$ (C) $\Delta R(\ell, b_{\text{lepTop}})$ (D) $\Delta R(\ell, b_{\text{hadTop}})$ (E) $\Delta R(b_{\text{lepTop}}, b_{\text{hadTop}})$ (F) $\Delta R(q_{1\text{hadW}}, q_{2\text{hadW}})$ FIGURE 9.12: ΔR variables constructed by the RecoBDT with H in ≥ 6 jets (1)

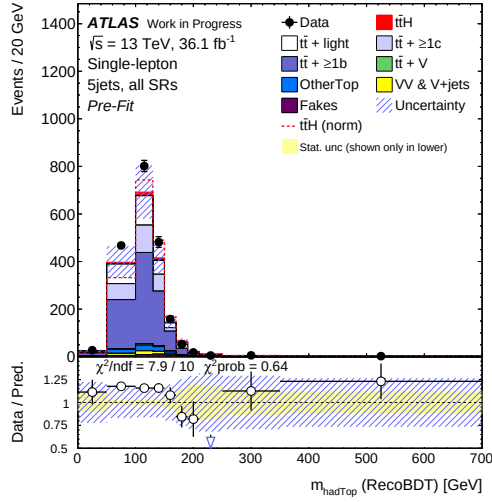
(A) $\Delta R(b_{\text{hadTop}}, q1_{\text{hadW}})$ (B) $\Delta R(b_{\text{hadTop}}, q2_{\text{hadW}})$ (C) $\min. \Delta R(b_{\text{hadTop}}, q1_{\text{hadW}})$ (D) $\Delta R(\ell, b_{\text{lepTop}}) - \min. \Delta R(b_{\text{hadTop}}, q1_{\text{hadW}})$ (E) $\Delta R(b1_{\text{Higgs}}, b2_{\text{Higgs}})$ (F) $\Delta R(b1_{\text{Higgs}}, \ell)$ FIGURE 9.13: ΔR variables constructed by the RecoBDT with H in ≥ 6 jets (2)



(A) Higgs boson mass



(B) lepTop mass



(C) hadTop mass

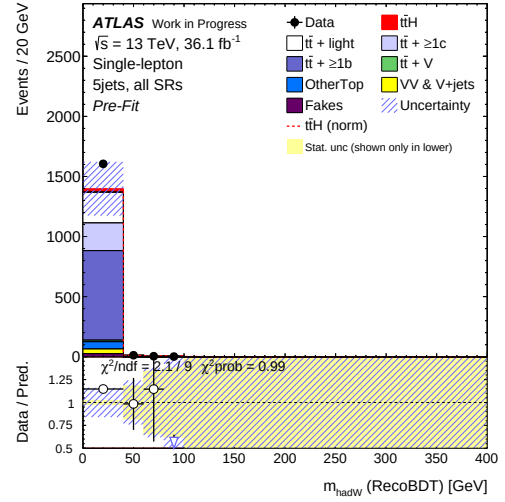
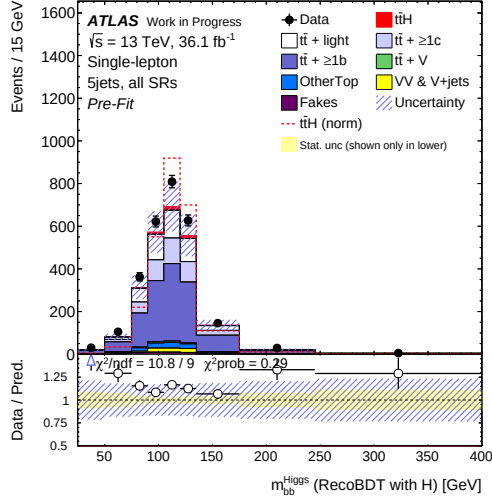
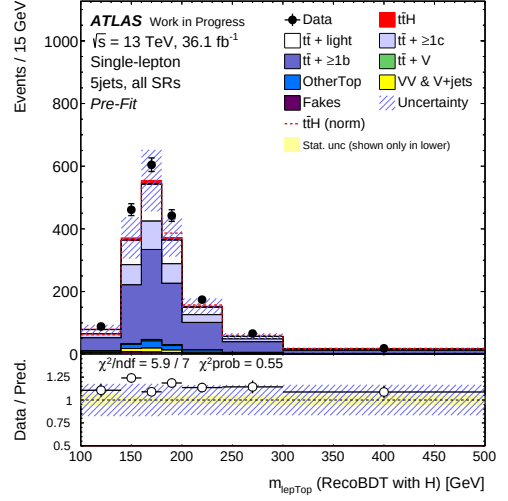
(D) W_{had} mass

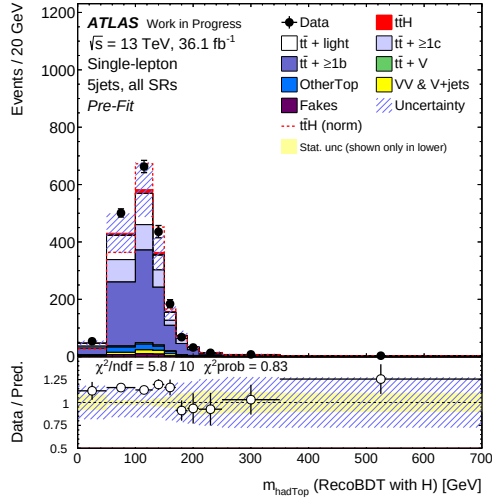
FIGURE 9.14: Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT w/o H in 5 jets



(A) Higgs boson mass



(B) lepTop mass



(C) hadTop mass

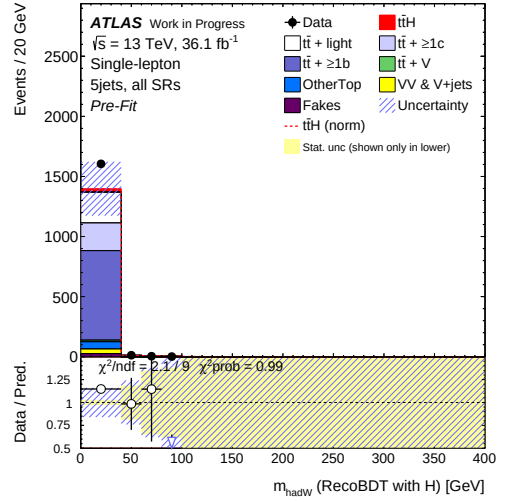
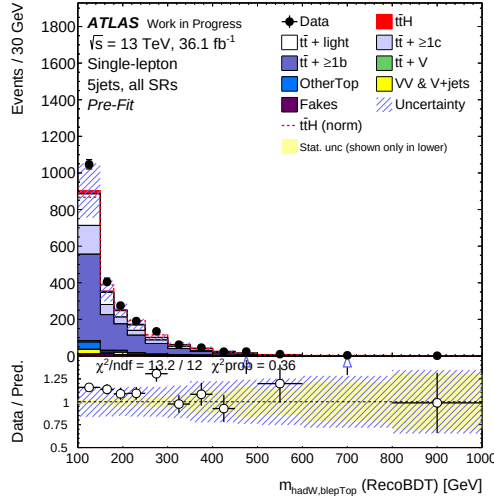
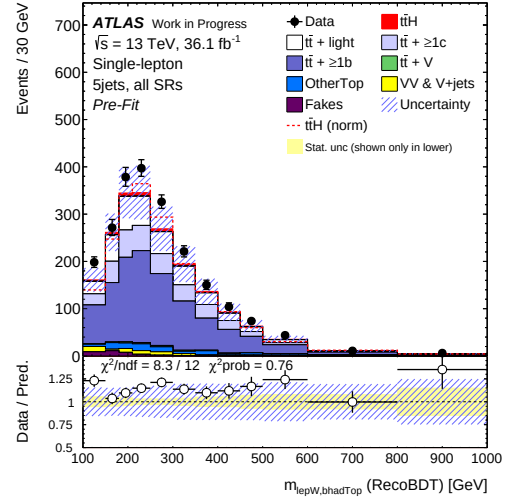
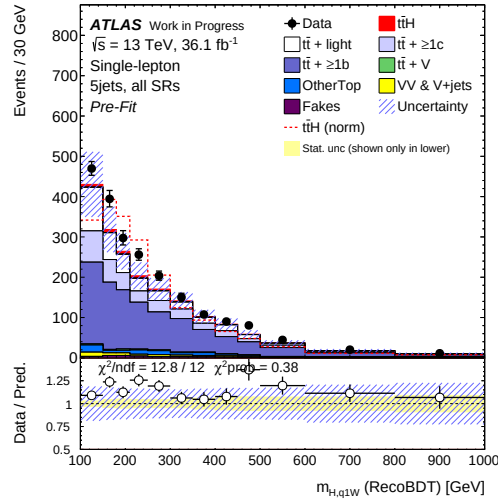
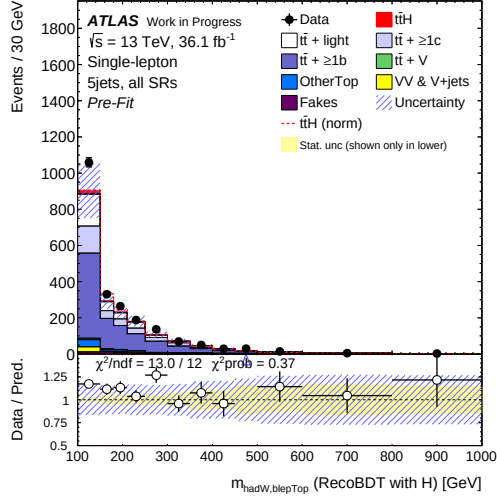
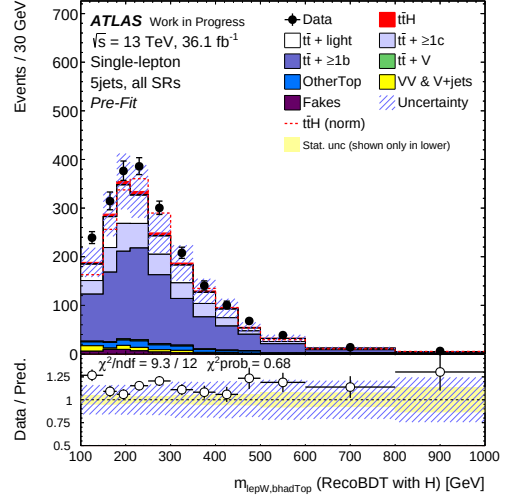
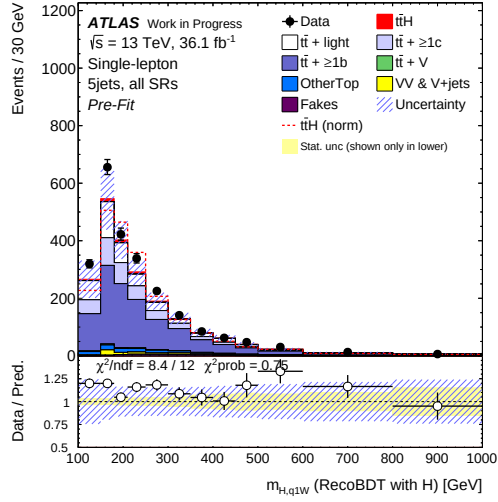
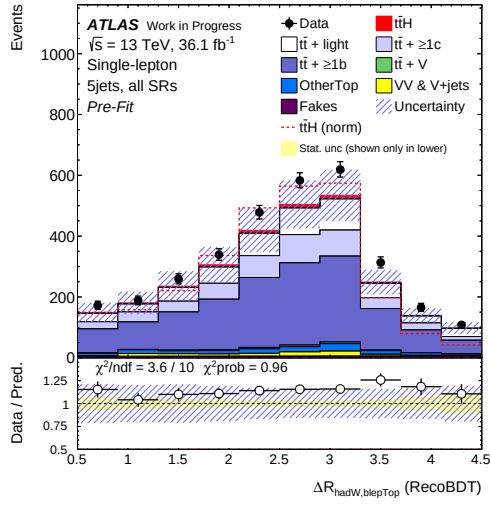
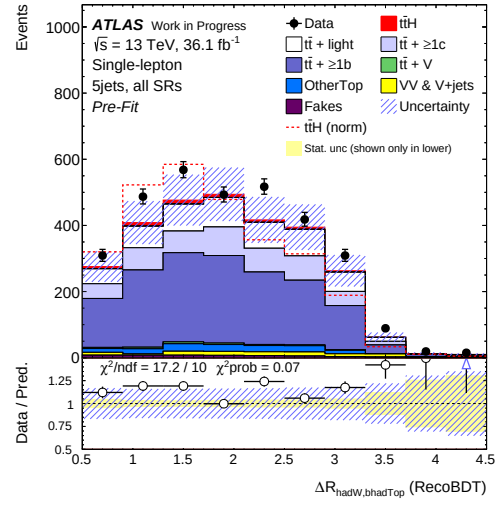
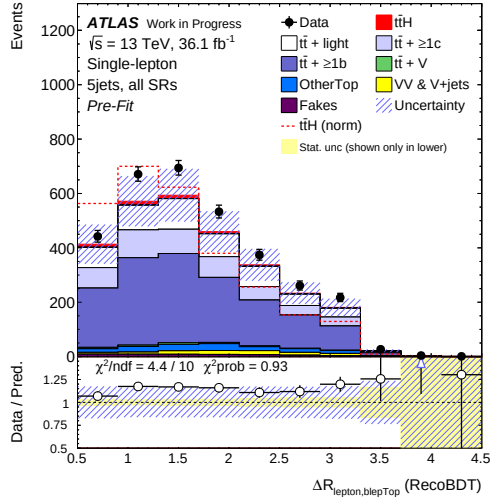
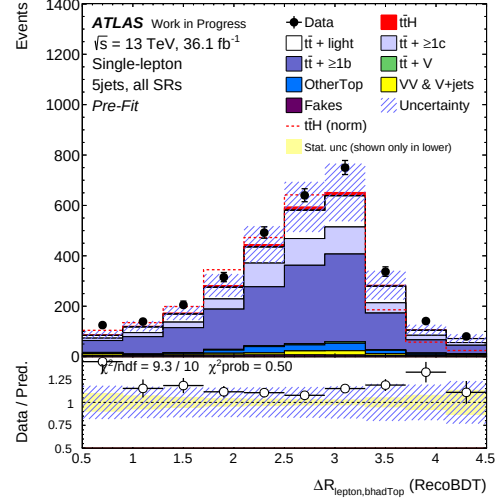
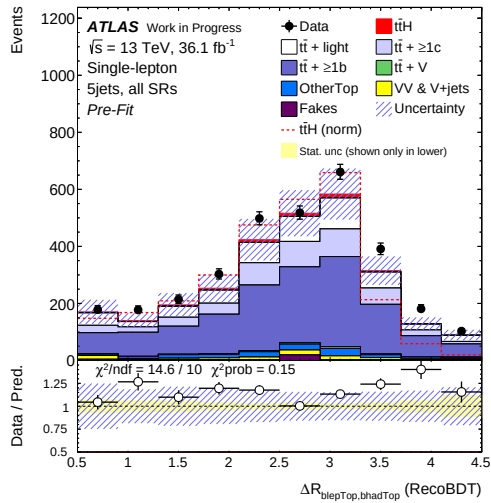
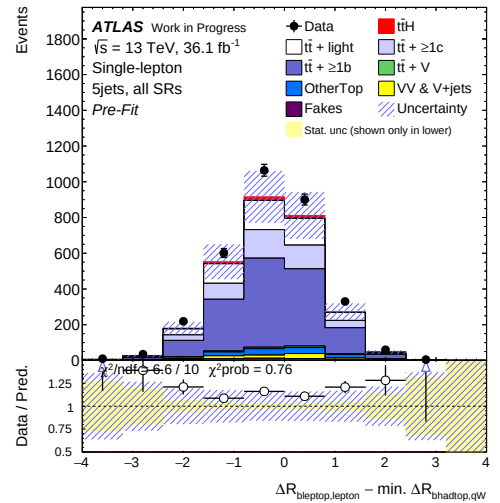
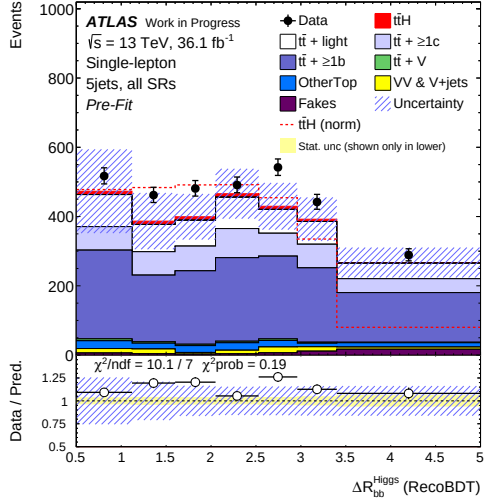
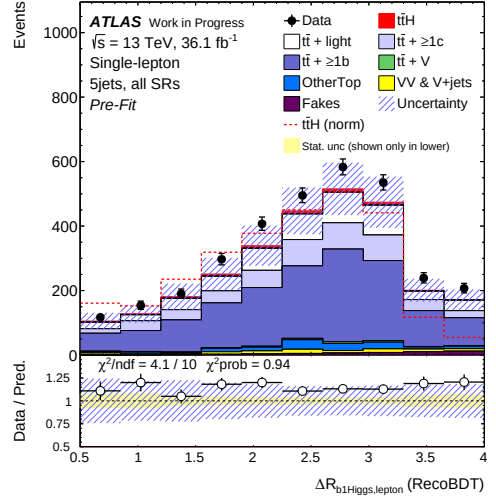
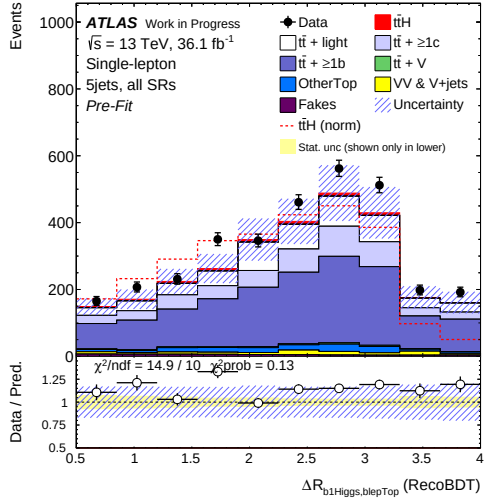
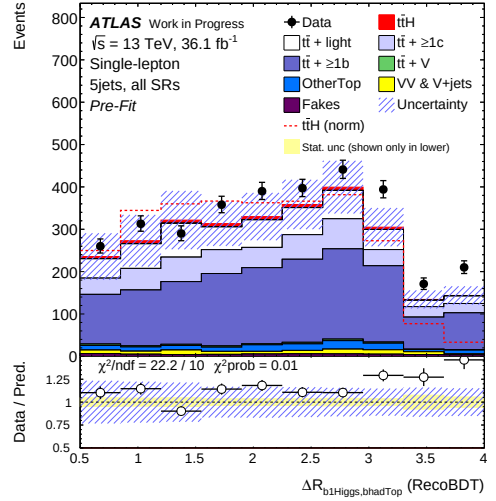
(D) W_{had} mass

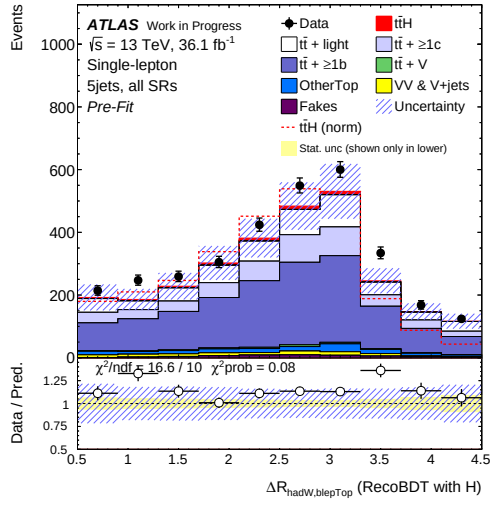
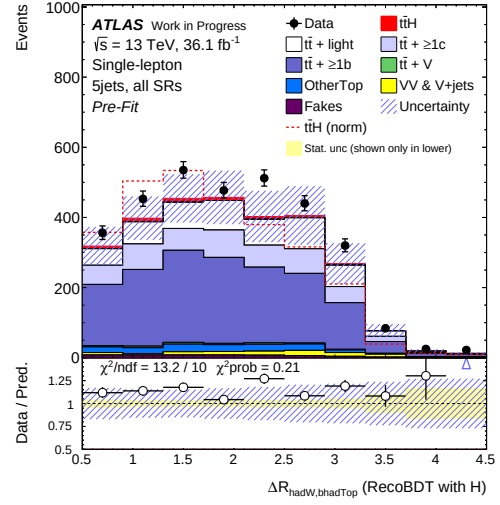
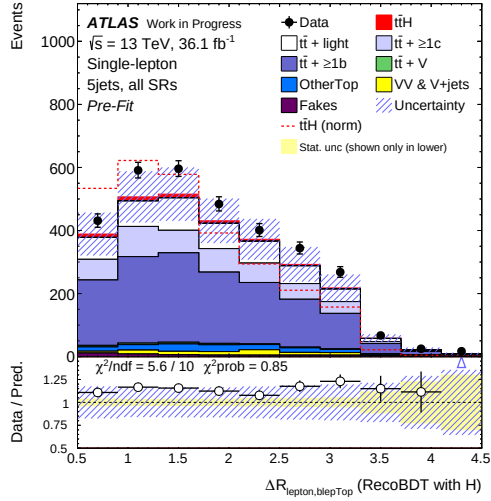
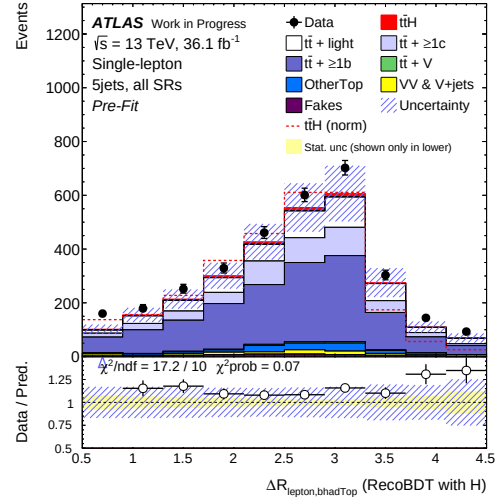
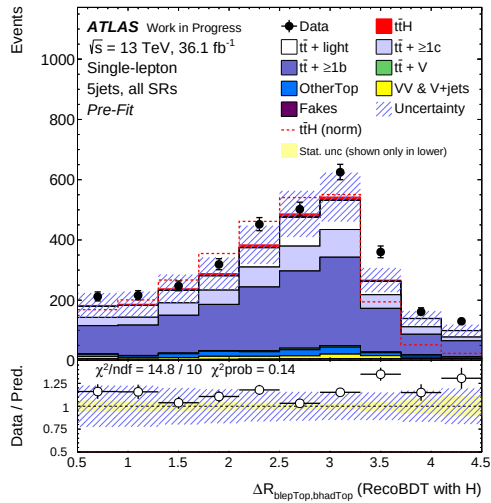
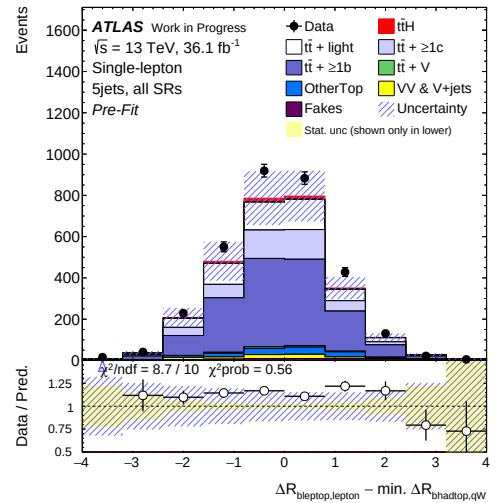
FIGURE 9.15: Invariant masses of lepTop, hadTop, W_{had} , and Higgs, which were constructed by the RecoBDT with H in 5 jets

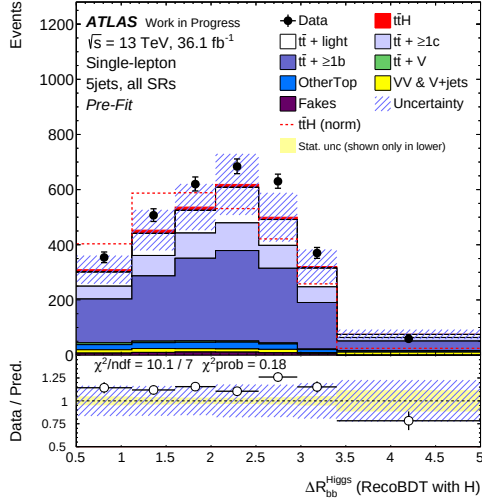
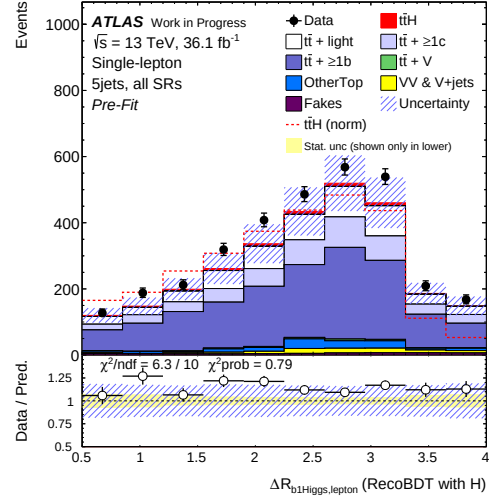
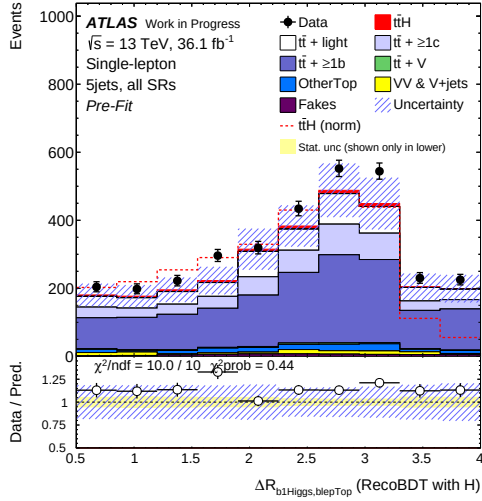
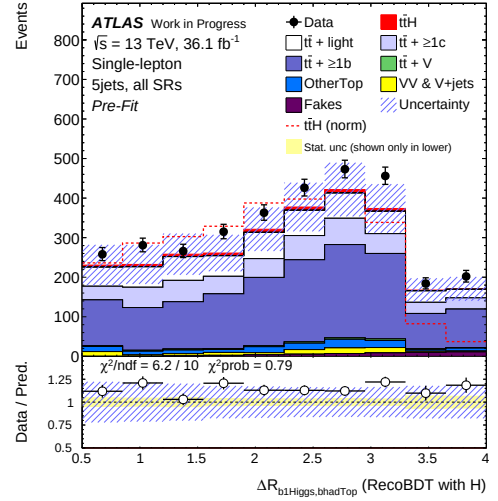
(A) mass of W_{had} and b_{hadTop} system(B) mass of W_{lep} and b_{lepTop} system(C) mass of Higgs and $q1_{\text{hadW}}$ systemFIGURE 9.16: Masses of two-object systems constructed by the RecoBDT w/o H in 5 jets

(A) mass of W_{had} and b_{hadTop} system(B) mass of W_{lep} and b_{lepTop} system(C) mass of Higgs and $q1_{\text{had}W}$ systemFIGURE 9.17: Masses of two-object systems constructed by the RecoBDT with H in 5 jets

(A) $\Delta R(W_{\text{had}}, b_{\text{lepTop}})$ (B) $\Delta R(W_{\text{had}}, b_{\text{hadTop}})$ (C) $\Delta R(\ell, b_{\text{lepTop}})$ (D) $\Delta R(\ell, b_{\text{hadTop}})$ (E) $\Delta R(b_{\text{lepTop}}, b_{\text{hadTop}})$ (F) $\Delta R(\ell, b_{\text{lepTop}}) - \min. \Delta R(b_{\text{hadTop}}, q_{i\text{had}W})$ FIGURE 9.18: ΔR variables constructed by the RecoBDT w/o H in 5 jets (1)

(A) $\Delta R(b_1Higgs, b_2Higgs)$ (B) $\Delta R(b_1Higgs, \ell)$ (C) $\Delta R(b_1Higgs, b_{lepTop})$ (D) $\Delta R(b_1Higgs, b_{hadTop})$ FIGURE 9.19: ΔR variables constructed by the RecoBDT w/o H in 5 jets (2)

(A) $\Delta R(W_{\text{had}}, b_{\text{lepTop}})$ (B) $\Delta R(W_{\text{had}}, b_{\text{hadTop}})$ (C) $\Delta R(\ell, b_{\text{lepTop}})$ (D) $\Delta R(\ell, b_{\text{hadTop}})$ (E) $\Delta R(b_{\text{lepTop}}, b_{\text{hadTop}})$ (F) $\Delta R(\ell, b_{\text{lepTop}})$
 $\min. \Delta R(b_{\text{hadTop}}, q_{i\text{hadW}})$ FIGURE 9.20: ΔR variables constructed by the RecoBDT with H in 5 jets (1)

(A) $\Delta R(b_{1\text{Higgs}}, b_{2\text{Higgs}})$ (B) $\Delta R(b_{1\text{Higgs}}, \ell)$ (C) $\Delta R(b_{1\text{Higgs}}, b_{1\text{lepTop}})$ (D) $\Delta R(b_{1\text{Higgs}}, b_{1\text{hadTop}})$ FIGURE 9.21: ΔR variables constructed by the RecoBDT with H in 5 jets (2)

9.1.2 Dilepton Channel

RecoBDT in the dilepton channel used truth-matched events where all four truth b -partons are matched with $\Delta R < 0.4$ to the individual reconstructed jet. For RecoBDT w/o H , exact truth-matching between the two b -jets from the Higgs boson decay is not required: if two b -jets from the Higgs-boson are swapped each other, those events were also included as the signal events.

In the reconstruction of the top/anti-top, $W_{\text{top/anti-top}}$ were constructed only by the positive/negative charged lepton (E_T^{miss} was not used), and four momenta of top/anti-top were calculated using these $W_{\text{top/anti-top}}$ (which equals to positive/negative charged lepton) and b -jets.

TABLE 9.3: Input variables to RecoBDT to assign reconstructed objects to $t\bar{t}H$ partons in the dilepton channel. Variables with $\checkmark$ are used in both RecoBDTs with/without Higgs information. The variables with $\bullet(\circ)$ mean that they are used only in RecoBDT with(without) Higgs information.

Variable	≥ 4 b -tags	3 b -tags
Topological information from $t\bar{t}$ decay		
$\Delta m(\text{top}, \text{anti-top})$	$\checkmark$	$\checkmark$
top mass	$\checkmark$	$\checkmark$
$\Delta R(\ell^+, b_{\text{top}})$	$\checkmark$	$\checkmark$
anti-top mass	$\checkmark$	$\checkmark$
$\Delta R(\ell^-, b_{\text{anti-top}})$	$\checkmark$	$\checkmark$
$\Delta\phi(\text{top}, \text{anti-top})$	$\circ$	$\checkmark$
$\Delta R(b_{\text{top}}, b_{\text{anti-top}})$	$\bullet$	—
p_T of b_{top}	$\circ$	$\circ$
p_T of $b_{\text{anti-top}}$	$\circ$	$\circ$
$\Delta(\Delta R(\ell^+, b_{\text{top}}), \Delta R(\ell^-, b_{\text{anti-top}}))$	$\circ$	$\circ$
min. $\Delta\eta(b_{(\text{anti-})\text{top}}, \ell^\pm)$	$\circ$	$\circ$
Topological information from Higgs decay		
Higgs boson mass	$\bullet$	$\bullet$
$\Delta\phi(\text{Higgs}, t\bar{t})$	—	$\bullet$
$\Delta R(\text{Higgs}, t\bar{t})$	$\bullet$	—
$\Delta R(b_{1\text{Higgs}}, b_{2\text{Higgs}})$	$\bullet$	$\bullet$
p_T of $b_{2\text{Higgs}}$	—	$\bullet$
min. $\Delta R(b_{i\text{Higgs}}, \ell^\pm)$	—	$\bullet$
max. $\Delta R(\text{Higgs}, b)$	$\bullet$	—

9.2 Matrix Element Discriminant

The Matrix Element Method (MEM) calculates the likelihood under an assumption that each event originates from a specific production, which is the most powerful discriminant when all final state objects are correctly detected and used as input information. Therefore, the MEM is only utilized in $\geq 6jSR(t\bar{t}H)$ in the single-lepton channel where this region has the largest truth-matching fraction and the highest sensitivity among all regions, thus, this region is expected to have a significant benefit using the MEM discriminant. In this analysis, MEM provides signal and background likelihoods ($\mathcal{L}_s$ and $\mathcal{L}_b$), where signal likelihood assumes that an event is produced according to the $t\bar{t}H(H \rightarrow b\bar{b})$ Feynman diagrams, while background likelihood assumes that an event is produced according to the $t\bar{t} + b\bar{b}$ Feynman diagrams. As a discriminant to separate them, MEM_{D1} is defined as

$$MEM_{D1} = \log(\mathcal{L}_s) - \log(\mathcal{L}_b). \quad (9.4)$$

Each likelihood for the physics process i is defined with the matrix element $\mathcal{M}_i(\mathbf{Y})$ calculated by MG5_AMC with CT10 PDF,

$$\mathcal{L}_i = \sum \int \frac{f_1(x_1, Q^2) f_2(x_2, Q^2)}{|\vec{q}_1||\vec{q}_2|} |\mathcal{M}_i(\mathbf{Y})|^2 T(\mathbf{X}; \mathbf{Y}) d\Phi(\mathbf{Y}). \quad (9.5)$$

Because the matrix element is calculated using four momenta of all final state partons as inputs, all reconstructed objects were transformed to truth partons with transfer functions, $T(\mathbf{X}; \mathbf{Y})$'s, which are PDFs of parton level parameters $\mathbf{Y}$'s of a given parton to be reconstructed as $\mathbf{X}$. Integrating the PDFs over $\mathbf{Y}$'s phase space $\Phi(\mathbf{Y})$ gives the measured $\mathbf{X}$ through $T(\mathbf{X}; \mathbf{Y})$. In the MEM calculation, the jet assignment is required. To reduce the number of combinations, four largest b -tagging scored jets were retained for the assignment to four b -partons, and a pair of remaining jets with a mass m_{jj} closest to $m_W = 80.4$ GeV were referred to the light-quark partons emitted from W_{had} . Therefore, there are $12(= 4!/2)$ permutations per event to be summed up in the likelihood.

The matrix element also requires information about the initial state particles. A parton distribution function $f_i(x_i, Q^2)$ of i -th proton was prepared with momentum $\vec{q}_i$ to carry energy fraction x_i of the proton in a collision energy scale $Q = (\sum_i E_i)^2 - (\sum_i p_{z,i})^2$ (i runs over all eight final state partons). The $t\bar{t}H(H \rightarrow b\bar{b})$ and $t\bar{t} + b\bar{b}$ productions from $q\bar{q}$ have significantly small fractions and were not used in the MEM calculation.

The most affecting parameters in the integration are jet energies, as they have large uncertainties and reconstructed parameters could differ from the parton levels. So in the integration of Equation 9.5, six following parameters were converted from reconstructed to truth parton level by integration:

- energy of b_{lepTop}
- energy of b_{hadTop}
- energy of $b1_{\text{Higgs}}$ ($b1_{\text{gluon}}$ if $t\bar{t} + b\bar{b}$)
- energy of $b2_{\text{Higgs}}$ ($b2_{\text{gluon}}$ if $t\bar{t} + b\bar{b}$)
- energy of $q1_{\text{hadW}}$
- energy of $q2_{\text{hadW}}$

The integration ranges of jet energy were determined from transfer functions. Other parameters (jet directions, lepton and neutrino four momenta) were assumed to be same, namely use

δ -function in integration. The neutrino p_z is analytically solved by assuming the W boson mass, and all possible solutions in the range of $-1 \text{ TeV} < p_z < 1 \text{ TeV}$ were summed up.

Their transfer functions were determined by fitting inclusive $t\bar{t}$ MC distributions ($n_{\text{jets}} \geq 4$ and $n_{\text{BTags}_{85\%}} \geq 2$). Different PDFs and parametrizations were adopted in light-jets and b -jets with the energy at reconstructed level (E_r) and at truth parton level (E_t):

$$T^{\text{light}}(E_r; E_t) = \frac{1}{\sqrt{2\pi}(\sigma_1 + A\sigma_2)} \left(\exp \left[\frac{-(E_t - E_r - \mu_1)^2}{2\sigma_1^2} \right] + A \cdot \exp \left[\frac{-(E_t - E_r - \mu_2)^2}{2\sigma_2^2} \right] \right) \quad (9.6)$$

$$T^b(E_r; E_t) = N \cdot \exp \left[\frac{-(E_t - E_r - \mu)^2}{2\sigma^2} \right] \quad \text{for } \frac{E_t - E_r - \mu}{\sigma} < \alpha \quad (9.7)$$

$$T^b(E_r; E_t) = N \cdot \left(\frac{n}{|\alpha|} \right)^n \exp \left[-\frac{|\alpha|^2}{2} \right] \left(\frac{n}{|\alpha|} - |\alpha| + \frac{E_t - E_r - \mu}{\sigma} \right)^{-n} \quad \text{for } \frac{E_t - E_r - \mu}{\sigma} \geq \alpha \quad (9.8)$$

where the double-Gaussian function was parametrized by $A, \mu_1, \sigma_1, \mu_2, \sigma_2$ for light-jet PDF and the crystal ball function was parametrized by μ, σ, α and n for b -jet PDF. Crystal ball function can describe the difference between the truth- and reconstructed-level which comes from the missing neutrino contributions in the semi-leptonic b -quark decay ($b \rightarrow c\mu\bar{\nu}_\mu$).

The obtained MEM_{D1} is shown in Figure 9.22.

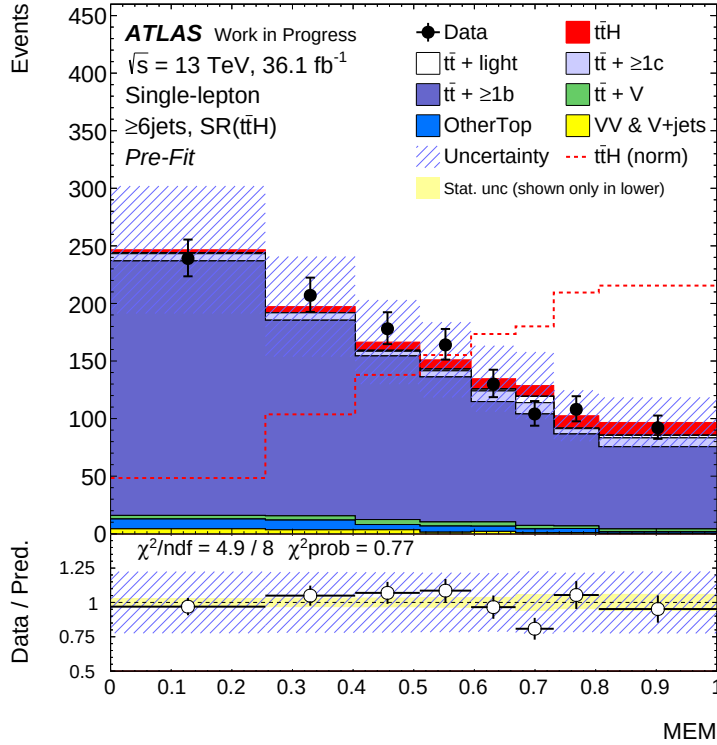


FIGURE 9.22: MEM_{D1} distribution in $\geq 6j\text{SR}(t\bar{t}H)$. The distribution was converted to the range of $[0, 1]$ by a sigmoid function: $1/(1 + \exp[-\text{MEM}_{\text{D1}} - 4])$.

9.3 Likelihood Discriminant

The likelihood discriminant (LHD) is to separate the $t\bar{t}H$ signal from $t\bar{t}+b\bar{b}$ background sources in the wide phase space, and was implemented in the single-lepton channel. Similar to the MEM, likelihoods ($\mathcal{P}_S$ and $\mathcal{P}_B$ for the signal and background) were calculated by assuming the $t\bar{t}H(H \rightarrow b\bar{b})$ and $t\bar{t}+b\bar{b}$ processes, respectively. The LHD additionally took into consideration the probability of missing jets from the $t\bar{t}H$ or $t\bar{t}+b\bar{b}$ system in the reconstructed-level. Therefore, it was utilized in all signal regions for both ≥ 6 jets and 5 jets events. Using two likelihoods, the LHD is calculated as:

$$\text{LHD} = \frac{\mathcal{P}_S}{\mathcal{P}_S + \mathcal{P}_B}. \quad (9.9)$$

Each likelihood was constructed by summation of combinations of jet assignments to the $t\bar{t}H$ or $t\bar{t}$ +jets system weighted by flavor-tagging information, where at most leading eight jets were used in order of p_T (the maximum number of permutations is $20,160 (= 8C_6 \cdot 6!)$). In each jet assignment, the weight of flavor-tagging information as the form of the likelihood, $\mathcal{P}_{\text{b-tag}}$ was defined by production of b -, c -, or light-jet tagging likelihoods of each jet assigned to b -, c -, or light-parton. The flavored jet PDFs are shown in Figure 8.2. Various likelihoods were also considered in the jet assignment by using PDFs of invariant masses and their differences for the purpose of reducing the correlations in the jet assignments for the Higgs boson, top quarks and W bosons. Likelihoods for the angles between two of $t\bar{t}H$ or $t\bar{t}$ +jets objects were also used to obtain larger significance for the separation between the signal and background sources. They were combined into one kinematic likelihood score as $\mathcal{P}_K^{\text{kin}}$ with $K = (S, B)$:

$$\begin{aligned} \mathcal{P}_K^{\text{kin}} = & \mathcal{P}_K(M_{j_1 j_2}) \mathcal{P}_K(M_{\text{lepTop}}) \mathcal{P}_K(M_{\text{hadTop}} - M_{\text{hadW}}) \mathcal{P}_K(M_{\text{hadW}}) \\ & \cdot \mathcal{P}_K(M_{t\bar{t}} - M_{\text{hadTop}} - M_{\text{lepTop}}) \mathcal{P}_K(M_{t\bar{t}+jj} - M_{t\bar{t}} - M_{j_1 j_2}) \\ & \cdot \mathcal{P}_K(\cos(\theta_{j_1(j_1 j_2 RF)} - \theta_{j_1 j_2})) \mathcal{P}_K(\cos(\theta_{jj(t\bar{t}+jj RF)} - \theta_{t\bar{t}+jj})) \end{aligned} \quad (9.10)$$

where j_1 and j_2 denote additional two jets besides $t\bar{t}$ ($b1_{\text{Higgs}}$ and $b2_{\text{Higgs}}$ for the $t\bar{t}H$ signal, $b1_{\text{gluon}}$ and $b2_{\text{gluon}}$ for the $t\bar{t}+b\bar{b}$ background). Several terms took PDFs of invariant masses which were subtracted with their sub-component masses to reduce correlations between them. The $\cos(\theta_{j_1(j_1 j_2 RF)} - \theta_{j_1 j_2})$ is cosine angle between j_1 in the $j_1 j_2$ rest-frame and the initial direction of $j_1 j_2$, and the $\cos(\theta_{jj(t\bar{t}+jj RF)} - \theta_{t\bar{t}+jj})$ is cosine angle between $j_1 j_2$ in the $t\bar{t}+jj$ rest-frame and the initial direction of the $t\bar{t}+jj$. In 5 j regions, W_{had} was constructed only by $q1_{\text{hadW}}$. The total likelihood ($\mathcal{P}_K$) is defined by $\mathcal{P}_{\text{b-tag}}$ and $\mathcal{P}_K^{\text{kin}}$ as:

$$\mathcal{P}_K = \frac{\sum (\mathcal{P}_{\text{b-tag}} \cdot \mathcal{P}_K^{\text{kin}})}{\sum \mathcal{P}_{\text{b-tag}}} \quad (9.11)$$

To cover the wide range of the phase space, the probabilities of missing jets were estimated with the $t\bar{t}H$ signal sample separately in $(n\text{Jets}, n\text{BTags}_{70\%}) = (\geq 6, \geq 4), (\geq 6, 3), (5, \geq 4)$ shown in Table 9.4. Three regions have different fractions for missing jets, however, $q2_{\text{hadW}}$ has the largest probability to miss in the reconstructed-level. Two likelihood were built and combined under different hypotheses such that all jets are fully matched to truth partons or that one of q_{hadW} jets is missed. In the case of missing q_{hadW} , W_{had} was reconstructed by remaining q_{hadW} and added a new jet replacing the missing q_{hadW} .

One of the additional b -jets also has the large fraction to be missed in the $t\bar{t}+b\bar{b}$ background

TABLE 9.4: Efficiency of fully truth-matched jets in $t\bar{t}H$ events in different selections. Selection column means (nJets,nBTags_{70%}). If at least one parton is un-matched, the event is not treated as fully matched. These fractions are listed in "missing jet". Missing fractions per origin are shown for the cases that only one jet is missing.

selection	fully matched	missing jets	missing parton for one jet missing events					
			$b1_{\text{Higgs}}$	$b2_{\text{Higgs}}$	b_{lepTop}	b_{hadTop}	$q1_{\text{hadW}}$	$q2_{\text{hadW}}$
($\geq 6, \geq 4$)	0.40	0.60	0.01	0.07	0.03	0.04	0.08	0.77
($\geq 6, = 3$)	0.24	0.76	0.01	0.04	0.02	0.02	0.09	0.82
($= 5, \geq 4$)	–	1.0	0.02	0.27	0.12	0.14	0.04	0.41

source (Table 9.5). Therefore, the probability of missing one of additional b -jets was also considered similarly to q_{hadW} jets. Two background hypotheses with $t\bar{t} + b\bar{b}$ and $t\bar{t} + bj$ (one of additional b -jets missed) were combined into one weighted by their probabilities. In the case of missing one of the additional b -jets, M_{j1j2} was constructed by the remaining b_{Higgs} with adding a new jet replacing the missing b_{Higgs} .

TABLE 9.5: Efficiency of fully truth-matched jets in $t\bar{t} + b\bar{b}$ events in different selections. Column "selection" refers to (nJets,nBTags_{70%}). Columns named $t\bar{t} + bc$ and $t\bar{t} + bl$ show the gluon fractions for the added jet in $t\bar{t} + bj$ events.

selection	$t\bar{t} + b\bar{b}$	$t\bar{t} + bj$	$t\bar{t} + bc$	$t\bar{t} + bl$
($\geq 6, \geq 4$), fully matched	0.82	0.18	0.14	0.86
($\geq 6, \geq 4$), missing one q_{hadW}	0.84	0.16	0.16	0.84
($\geq 6, = 3$), fully matched	0.42	0.58	0.07	0.93
($\geq 6, = 3$), missing one q_{hadW}	0.46	0.54	0.08	0.92
($= 5, \geq 4$), missing one q_{hadW}	0.71	0.29	0.17	0.83

In addition, for the calculation of M_{lepTop} , neutrino p_z was determined analytically using the constraint of W boson mass. The probability of obtaining $p_{\nu z,1}$ ($p_{\nu z,2}$) closer to the truth neutrino was estimated from the $t\bar{t}H$ simulation to be 65% (35%), if two real solutions ($p_{\nu z,1}$ and $p_{\nu z,2}$ for $|p_{\nu z,1}| < |p_{\nu z,2}|$) are obtained. In the LHD, two solutions were summed over weighted by this probability ($\mathcal{P}_K^{\text{kin}} = 0.65 \times \mathcal{P}_K^{\text{kin}}(p_{\nu z,1}) + 0.35 \times \mathcal{P}_K^{\text{kin}}(p_{\nu z,2})$ if two real solutions were obtained for $p_{\nu z}$).

In the calculation of $\mathcal{P}_{\text{b-tag}}$, one of q_{hadW} can be a c -jet or light-jet as a decay product of $W \rightarrow cs$ or ud , thus, both possibilities were considered with fractions estimated by each sample ($t\bar{t}H$ for the signal and $t\bar{t} + b\bar{b}$ for the background). In the case of missing jets, added new jet has the same situation. Therefore, the b -tagging likelihood was described in the fully matched hypothesis in the $\geq 6j$ regions:

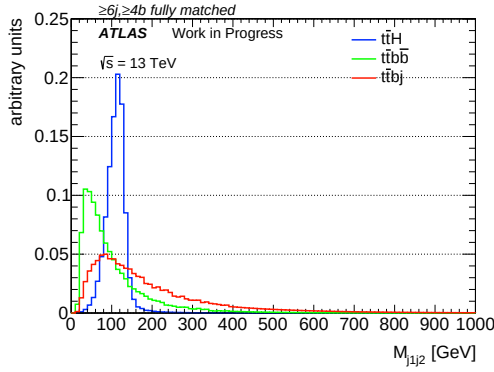
$$\begin{aligned} \mathcal{P}_{\text{b-tag}} = & \mathcal{P}^b(b_{\text{lepTop}})\mathcal{P}^b(b_{\text{hadTop}})\mathcal{P}^b(j1)\mathcal{P}^b(j2) \\ & \cdot [f^c\mathcal{P}^c(q1_{\text{hadW}}) + f^l\mathcal{P}^l(q1_{\text{hadW}})]\mathcal{P}^l(q2_{\text{hadW}}) \end{aligned} \quad (9.12)$$

where $\mathcal{P}^b$, $\mathcal{P}^c$, and $\mathcal{P}^l$ are PDFs of b -tagging tag weight bins for b -, c -, and light-jets (Figure 8.2). In the term for q_{hadW} probability, both c - and light-jet PDFs were considered with their fractions (f^c and f^l). In the case of the missing one of additional b -jets for the $t\bar{t} + b\bar{b}$ background hypothesis, both c - and light-jet PDFs were considered also for the new added jet replacing the missing one with the fractions shown in Table 9.5. For the new added jet in the case of missing one of q_{hadW} , only light-jet was considered.

All PDFs were prepared per region and hypothesis. Considerable number of PDFs for one variable was in total at most 45 regions:

$$\begin{pmatrix} t\bar{t}H \\ t\bar{t} + b\bar{b} \\ t\bar{t} + bj \end{pmatrix} \times \begin{pmatrix} \geq 6j, \geq 4b \\ \geq 6j, = 3b \\ = 5j, \geq 4b \end{pmatrix} \times \begin{pmatrix} \text{missing } q_{\text{hadW}} \text{ jets} \\ \text{fully matched } (\geq 6j) \end{pmatrix} \times \begin{pmatrix} p_{vz,1} & \text{for 2 solutions} \\ p_{vz,2} & \text{for 1 solution} \end{pmatrix}. \quad (9.13)$$

Variables showing noticeable differences are shown in Figures 9.23 to 9.24 under different hypotheses.



(A) additional dijet mass

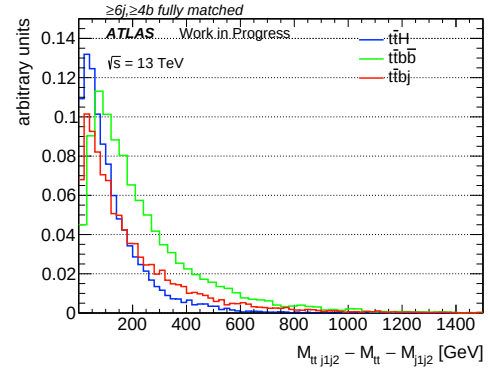
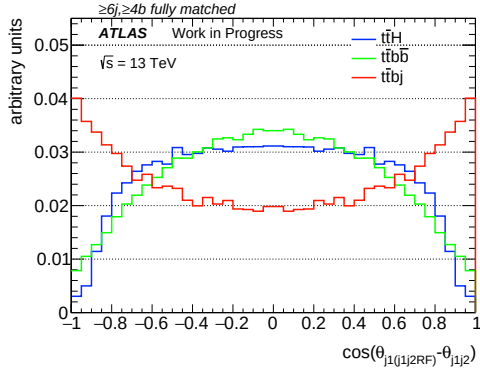
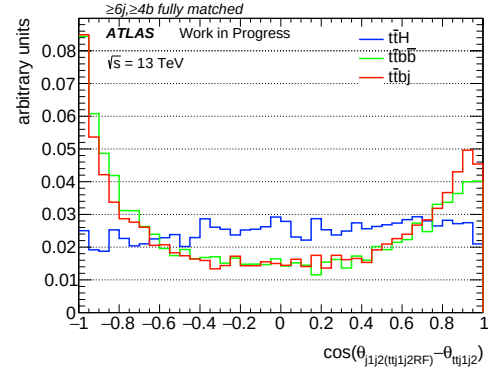
(B) $t\bar{t}+jj$ mass subtracted with $t\bar{t}$ mass and additional dijet mass(C) cosine angle between j_1 in the j_1j_2 ($t\bar{t} + jj$) rest-frame and the initial direction of j_1j_2 (D) cosine angle between j_1j_2 in the $t\bar{t} + jj$ rest-frame and the initial direction of $t\bar{t} + jj$

FIGURE 9.23: Kinematic variables used in $\mathcal{P}_S^{\text{kin}}$ and $\mathcal{P}_B^{\text{kin}}$ showing noticeable differences among $t\bar{t}H$, $t\bar{t} + b\bar{b}$, and $t\bar{t} + bj$ in ≥ 6 jets and ≥ 4 b -tagged jets.

Obtained LHD distributions in each signal region are shown in Figures 9.25 to 9.27.

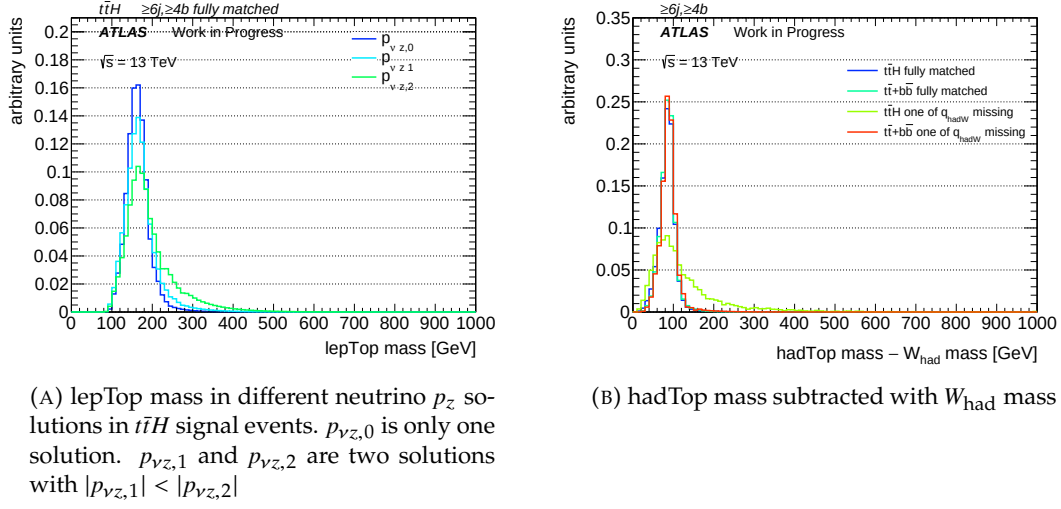


FIGURE 9.24: Kinematic variables used in $\mathcal{P}_S^{\text{kin}}$ and $\mathcal{P}_B^{\text{kin}}$ showing noticeable differences in ≥ 6 jets and ≥ 4 b -tagged jets.

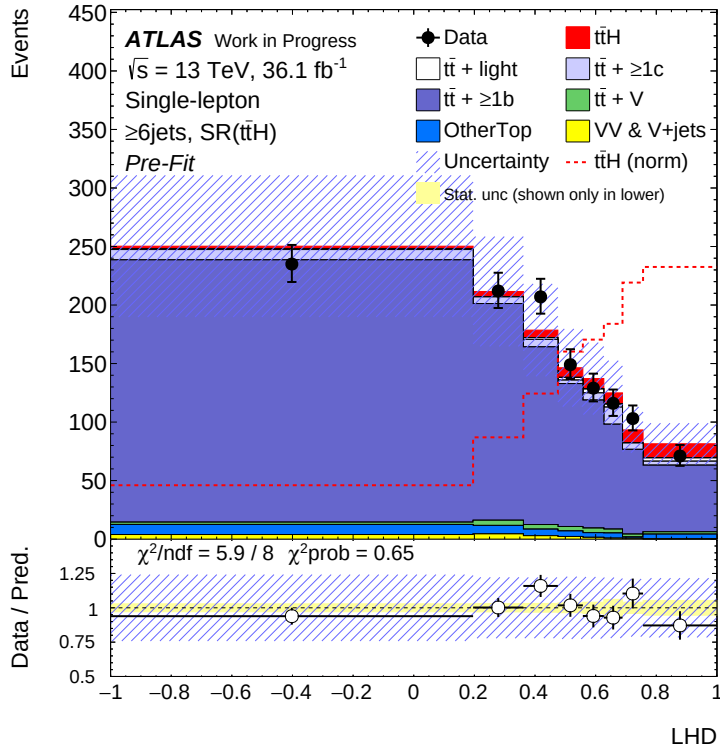
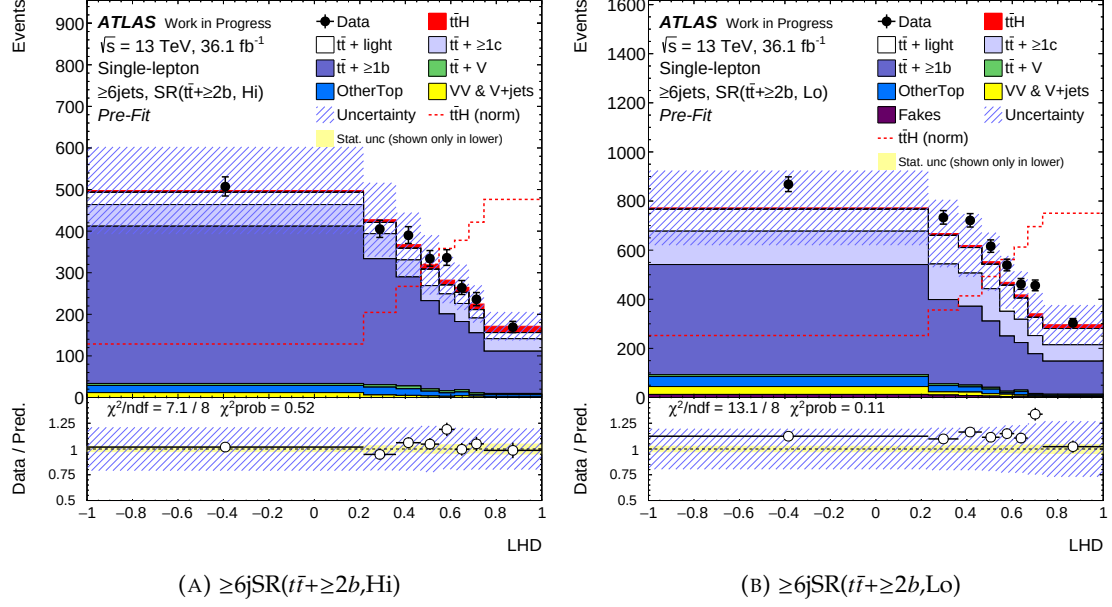
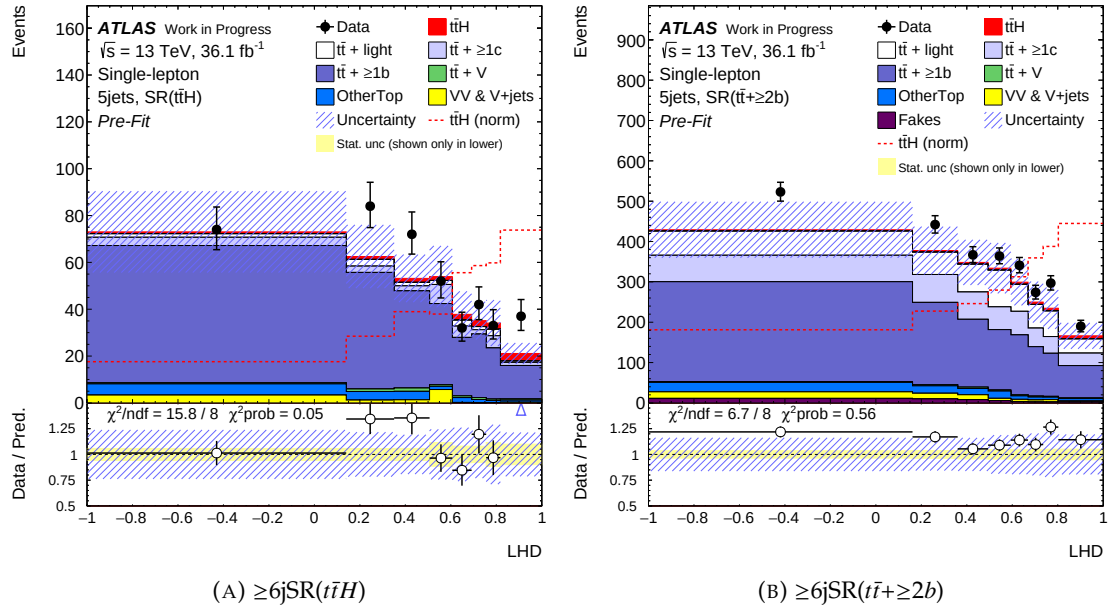


FIGURE 9.25: LHD distributions in $\geq 6j\text{SR}(t\bar{t}H)$.

FIGURE 9.26: LHD distributions in ≥ 6 jets $\text{SR}(t\bar{t} + \geq 2b, \text{Hi})$ and $\text{SR}(t\bar{t} + \geq 2b, \text{Lo})$.FIGURE 9.27: LHD distributions in 5 jets $\text{SR}(t\bar{t}H)$ and $\text{SR}(t\bar{t} + \geq 2b)$.

Chapter 10

Final Discriminant with Multivariate Analysis

To maximize the sensitivity of $t\bar{t}H(H \rightarrow b\bar{b})$ separation from the $t\bar{t}+\geq 1b$ background source, a multivariate analysis using a BDT algorithm, ClassBDT, was adopted as the final discriminant and to determine the signal strength. The ClassBDT was trained separately in ≥ 6 jets and 5 jets in the single-lepton channel and in each of the three signal regions in the dilepton channel using $t\bar{t}H(H \rightarrow b\bar{b})$ events as the signal and inclusive $t\bar{t}$ +jets ($t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$ +light) events as the background.

10.1 Single-lepton Channel

The events for training the ClassBDTs in ≥ 6 jets and 5 jets required at least 4 b -tagged jets at 85% WP to keep statistics. In addition, a dedicated training was performed with the MEM_{D1} in the most sensitive signal region, $\geq 6\text{jSR}(t\bar{t}H)$.

For the input variables, simple kinematics were taken such as a pair of jets (including b -tagged jets), event shape variables, b -tagging scores, RecoBDT information and dedicated discriminants (LHD and MEM_{D1}). All input variables are summarized in Table 10.1, and their linear correlations and rankings of their contributions to the ClassBDT scores are shown in Figure 10.1 and 10.2, respectively. The data and MC comparisons are shown in Figures 10.5 to 10.12. The b -tagged jets are defined as the four jets with largest b -tagging tag weight bins. (The jets in a same tag weight bin are sorted by larger p_T .)

Two event shape variables were taken as input variables in the ClassBDT. The first one, Aplanarity, is calculated from the momentum tensor [106] defined as,

$$M_{xyz} = \sum_i \begin{pmatrix} p_{x,i}^2 & p_{x,i}p_{y,i} & p_{x,i}p_{z,i} \\ p_{y,i}p_{x,i} & p_{y,i}^2 & p_{y,i}p_{z,i} \\ p_{z,i}p_{x,i} & p_{z,i}p_{y,i} & p_{z,i}^2 \end{pmatrix} \quad (10.1)$$

where $p_{k,i}$ is the momentum of the i -th object along the k -axis. M_{xyz} is summed over all objects in consideration. Eigenvalues, $\lambda_0, \lambda_1, \lambda_2$ ($\lambda_0 > \lambda_1 > \lambda_2 > 0$) are used to construct event shape variables such as Aplanarity = $\frac{3}{2}\lambda_2$. The second one, H_l , is defined using Fox-Wolfram moment [107],

$$H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{(\sum_k |\vec{p}_k|)^2} P_l(\cos \Omega_{ij}) \quad (10.2)$$

where $P_l(\cos \Omega_{ij})$ is the Legendre Polynomial which can be described as $P_0(x) = 1$ for $l = 0$ and $P_1(x) = x$ for $l = 1$, and H_l is summed up all i, j combinations of objects in consideration. For

example, H_1 is described as

$$H_1 = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j| \cos \Omega_{ij}}{(\sum_k |\vec{p}_k|)^2} = \frac{\sum_{i,j} (\vec{p}_i \cdot \vec{p}_j)}{(\sum_k |\vec{p}_k|)^2} = \frac{(\sum_i \vec{p}_i)^2}{(\sum_k |\vec{p}_k|)^2}. \quad (10.3)$$

H_1 is the ratio of the vector sum and the scalar sum of the momenta, and becomes zero if the system at rest due to $\sum_i \vec{p}_i = 0$, while H_1 becomes one if all considered objects move along the same direction. In this ClassBDTs, $\text{Aplanarity}_{\text{jet}}$ was summed up over all jets and H_1^{all} was summed up for all jets and leptons.

From RecoBDT information, the best jet assignments with the largest score and the resulting kinematic variables were provided into the input variables for ClassBDTs. RecoBDT w/o H does not bias the Higgs variables on both signal and background, therefore, Higgs related variables with this jet assignment have large separating powers. On the other hand, RecoBDT with H has a large separating power in its score itself.

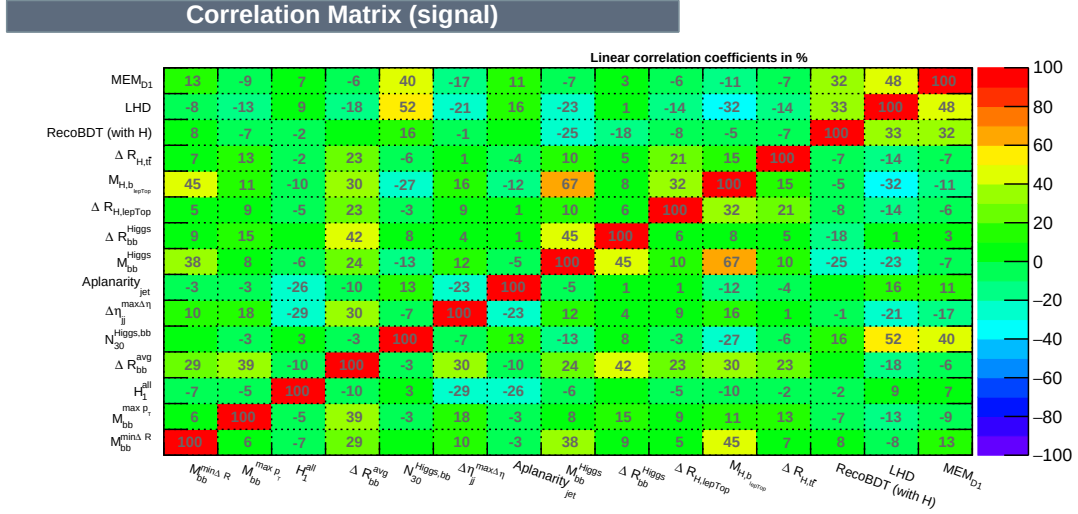
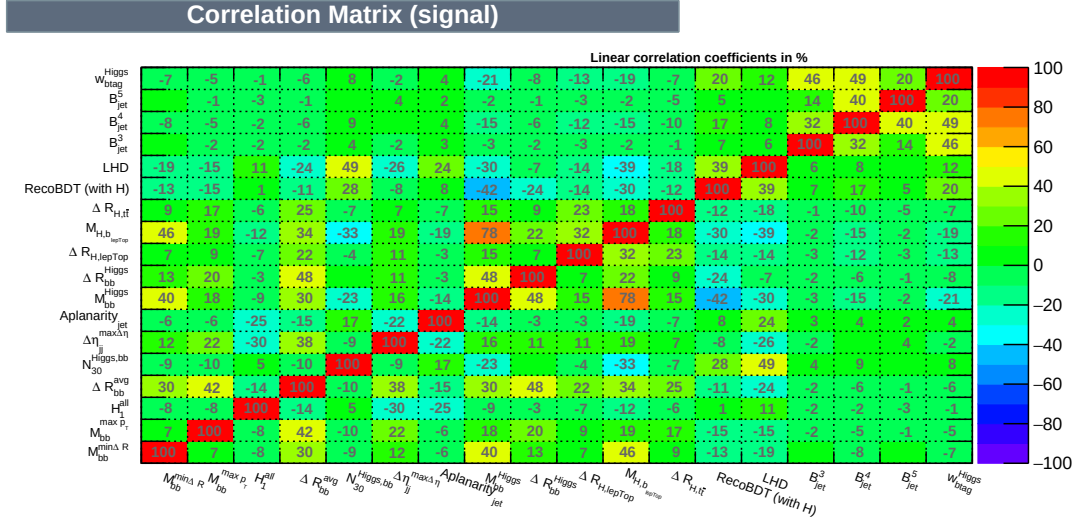
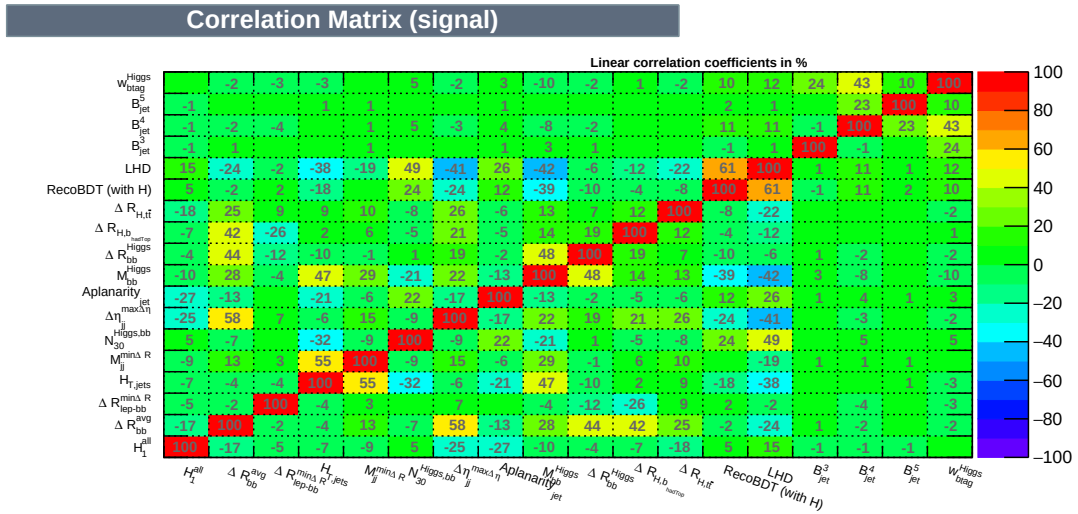
The variables using b -tagging tag weight bins have still powers for rejecting $t\bar{t} + \geq 1c$, $t\bar{t} + \text{light}$ and other non- $t\bar{t}$ productions in most signal regions. To use them in the BDT input variables, $B_{\text{jet}} = \{60, 70, 77, 85, 100\}$ was converted to the continuous integer score $B_{\text{jet}} = \{5, 4, 3, 2, 1\}$. These were not used in the dedicated training in $\geq 6\text{jSR}(t\bar{t}H)$ because all $B_{\text{jet}}^{1,2,3,4}$ are 5 and B_{jet}^5 does not have a separation power sufficient to be added.

TABLE 10.1: Input variables used to the ClassBDTs in the single-lepton signal regions. The MEM_{D1} variable was only used in $\geq 6\text{jSR}(t\bar{t}H)$, where variables based on the b -tagging discriminant were not used.

Variable	Description	$\geq 6\text{ jets}$	$\geq 6\text{jSR}(t\bar{t}H)$	5 jets
General kinematic variables				
$\Delta R_{bb}^{\text{avg}}$	Average ΔR for all b -tagged jet pairs	✓	✓	✓
$\Delta R_{bb}^{\text{max } p_T}$	ΔR between the two b -tagged jets with the largest p_{Tbb}	✓	✓	–
$\Delta \eta_{jj}^{\text{max } \Delta \eta}$	Maximum $\Delta \eta$ between any two jets	✓	✓	✓
$m_{bb}^{\text{min } \Delta R}$	Mass of two b -tagged jets system with the smallest ΔR_{bb}	✓	✓	–
$m_{jj}^{\text{min } \Delta R}$	Mass of two jets system with the smallest ΔR_{jj}	–	–	✓
$N_{30}^{\text{Higgs}, bb}$	Number of b -tagged jet pairs with m_{bb} within the Higgs mass $\pm 30\text{ GeV}$	✓	✓	✓
H_T^{had}	Scalar sum of jet p_T	–	–	✓
$\Delta R_{\text{lep}-bb}^{\text{min } \Delta R}$	Smallest ΔR between the lepton and the two b -tagged jets system	–	–	✓
Event shape variables				
$\text{Aplanarity}_{\text{jet}}$	1.5 times the 2nd eigenvalue of the momentum tensor using all jets	✓	✓	✓
H_1^{all}	Second Fox-Wolfram moment computed using all jets and the lepton	✓	✓	✓
Likelihood and matrix element method calculations				
LHD	Likelihood discriminant	✓	✓	✓
MEM_{D1}	Matrix element discriminant	–	✓	–
Variables from RecoBDT with H				
RecoBDT score	Output score of the RecoBDT with H	✓	✓	✓
$\Delta R_{H, t\bar{t}}$	ΔR between Higgs and $t\bar{t}$ system	✓	✓	✓
$\Delta R_{H, b_{\text{hadTop}}}$	ΔR between Higgs and b -jet from hadronic top	–	–	✓
Variables from RecoBDT w/o H				
m_{bb}^{Higgs}	Higgs mass	✓	✓	✓
$m_{H, b_{\text{lepTop}}}$	Mass of Higgs and b -jet from leptonic top	✓	✓	–
$\Delta R_{bb}^{\text{Higgs}}$	ΔR between b -jets from the Higgs	✓	✓	✓
$\Delta R_{H, \text{lepTop}}$	ΔR between Higgs and leptonic top candidate	✓	✓	–
$w_{b\text{-tag}}^{\text{Higgs}(bb)}$	Sum of tag weight bins of jets from Higgs in best assignment provided by RecoBDT w/o H	✓	✓	✓
Variables from b -tagging				
B_{jet}^3	3rd largest tag weight bin	✓	–	✓
B_{jet}^4	4th largest tag weight bin	✓	–	✓
B_{jet}^5	5th largest tag weight bin	✓	–	✓

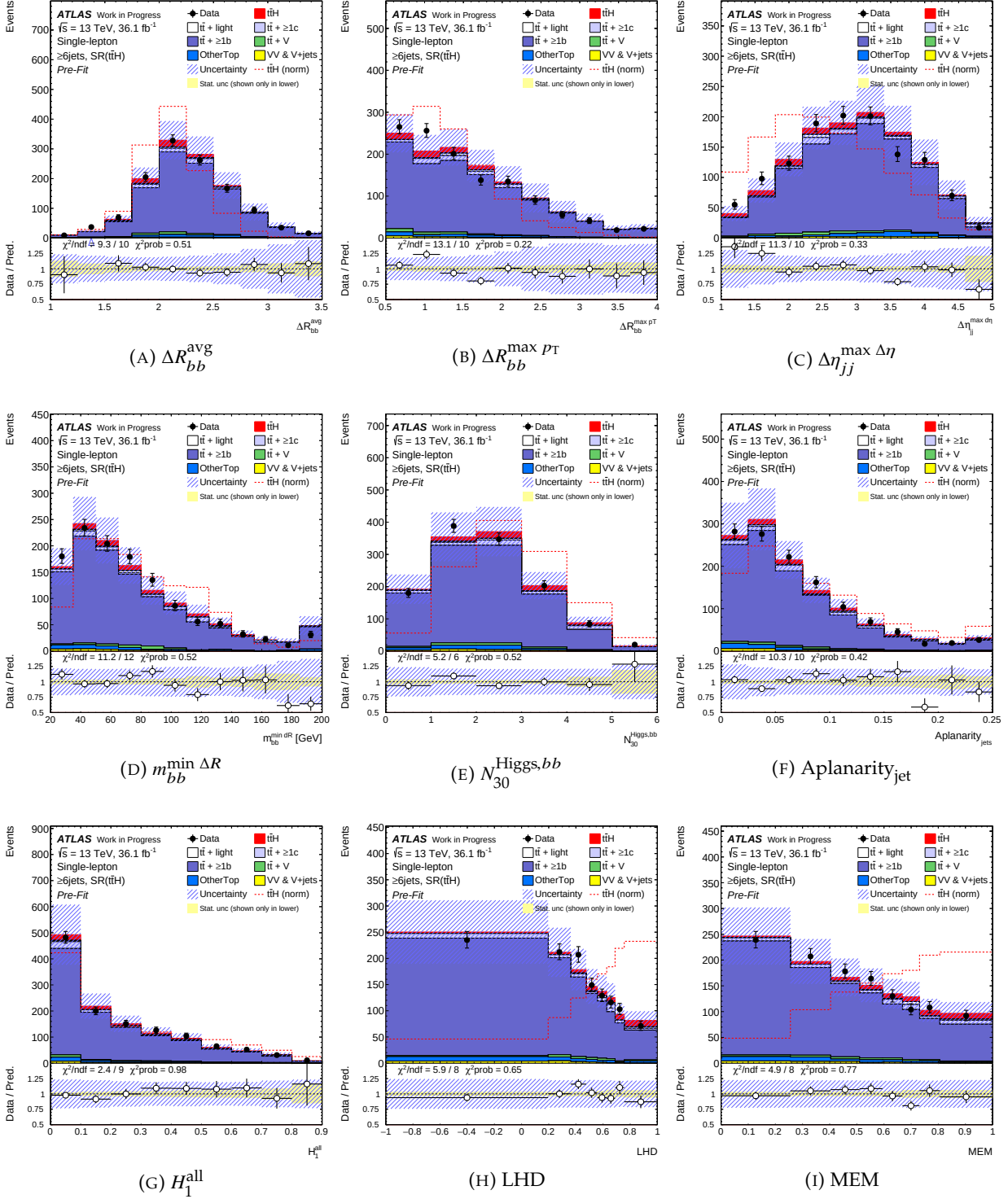
TABLE 10.2: Ranking of the input variables used to the ClassBDTs in the single-lepton signal regions. The variables are sorted according to the BDT score gained by adding the variable in the ClassBDT.

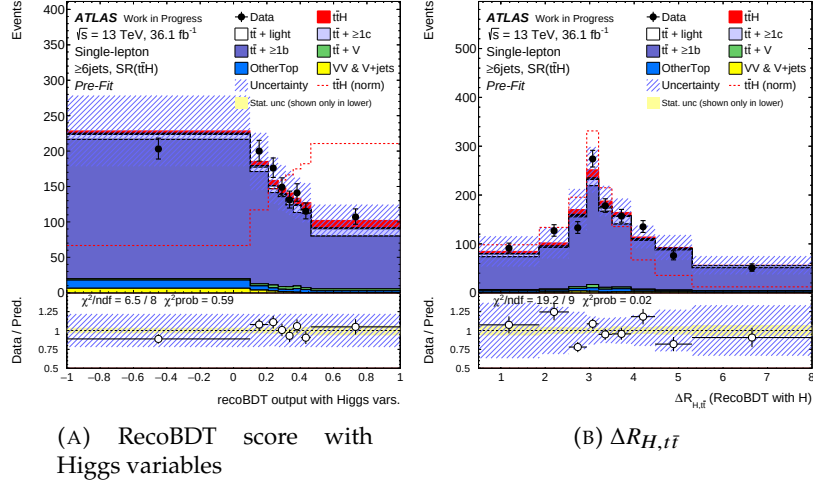
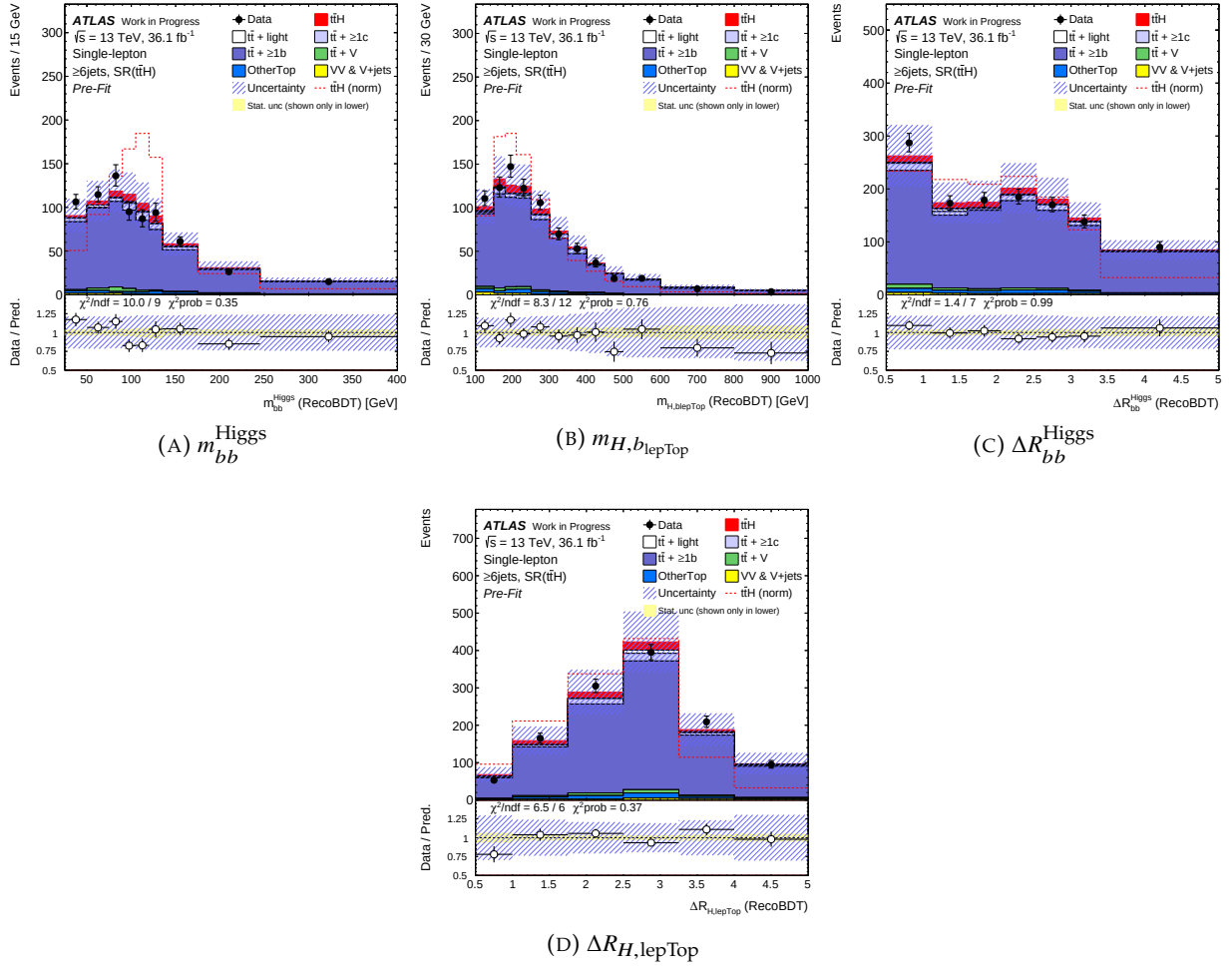
Ranking	≥ 6 jets	≥ 6 SR($t\bar{t}H$)	5 jets
1	B_{jet}^4	LHD	B_{jet}^4
2	LHD	MEM_{D1}	$\Delta\eta_{jj}^{\max \Delta\eta}$
3	B_{jet}^3	RecoBDT score	LHD
4	$\Delta R_{bb}^{\text{avg}}$	$\Delta R_{bb}^{\max p_T}$	H_T^{had}
5	RecoBDT score	$m_{bb}^{\min \Delta R}$	$\Delta R_{bb}^{\text{avg}}$
6	$w_{b\text{-tag}}^{\text{Higgs}(bb)}$	m_{bb}^{Higgs}	B_{jet}^3
7	$m_{bb}^{\min \Delta R}$	$\Delta R_{bb}^{\text{Higgs}}$	$w_{b\text{-tag}}^{\text{Higgs}(bb)}$
8	$\Delta\eta_{jj}^{\max \Delta\eta}$	$\Delta R_{bb}^{\text{avg}}$	$m_{jj}^{\min \Delta R}$
9	$\Delta R_{H,\text{lepTop}}$	$\Delta\eta_{jj}^{\max \Delta\eta}$	RecoBDT score
10	$\Delta R_{bb}^{\text{Higgs}}$	$m_{H,b_{\text{lepTop}}}$	$\Delta R_{H,b_{\text{hadTop}}}$
11	$\Delta R_{bb}^{\max p_T}$	H_1^{all}	$\Delta R_{bb}^{\text{Higgs}}$
12	Aplanarity _{jet}	$\Delta R_{H,t\bar{t}}$	$\Delta R_{\text{lep-}bb}^{\min \Delta R}$
13	$m_{H,b_{\text{lepTop}}}$	Aplanarity _{jet}	H_1^{all}
14	m_{bb}^{Higgs}	$\Delta R_{H,\text{lepTop}}$	$\Delta R_{H,t\bar{t}}$
15	$\Delta R_{H,t\bar{t}}$	$N_{30}^{\text{Higgs},bb}$	m_{bb}^{Higgs}
16	B_{jet}^5	—	Aplanarity _{jet}
17	H_1^{all}	—	$N_{30}^{\text{Higgs},bb}$
18	$N_{30}^{\text{Higgs},bb}$	—	B_{jet}^5

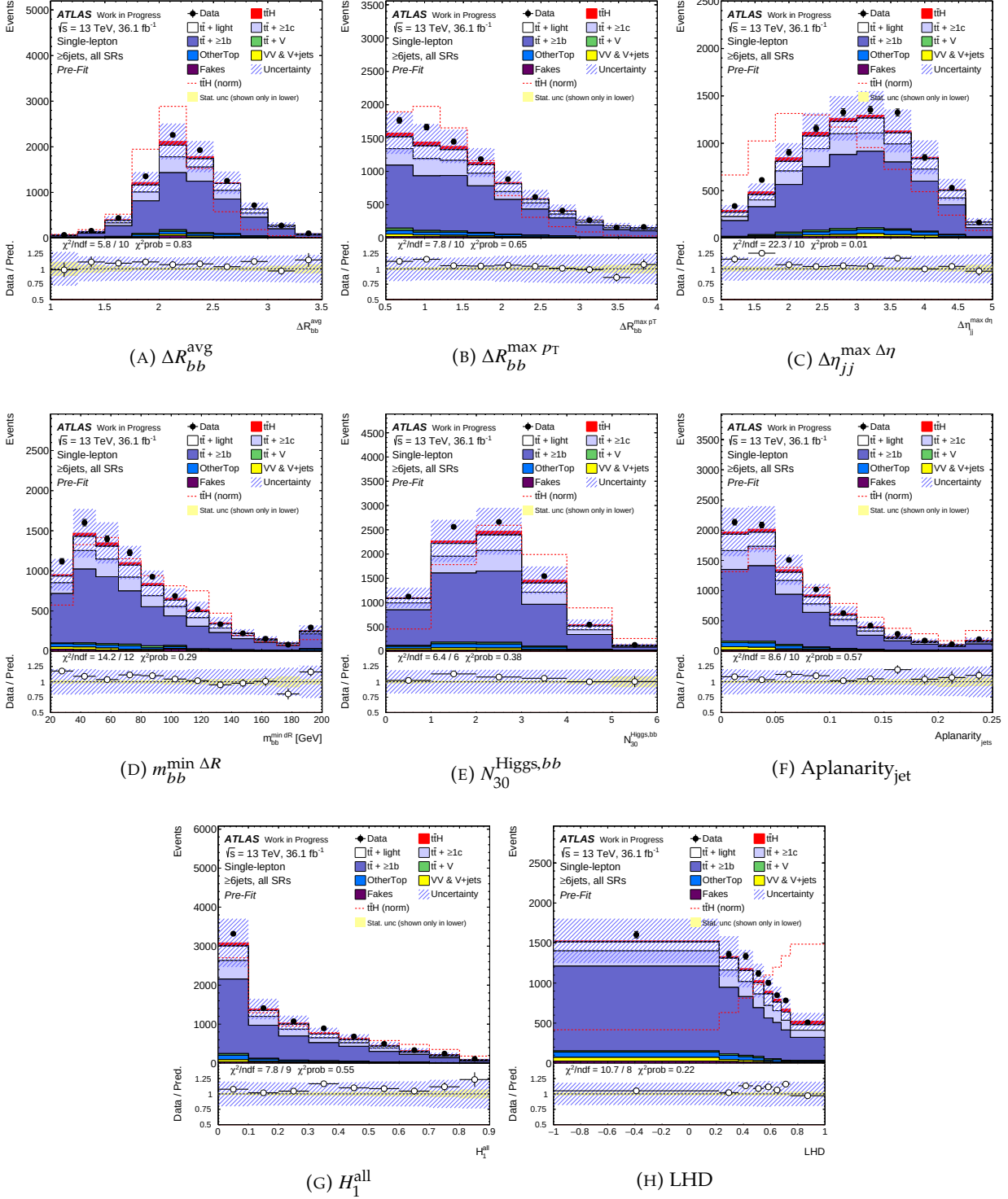
(A) $\geq 6jSR(t\bar{t}H)$ (B) ≥ 6 jets signal regions

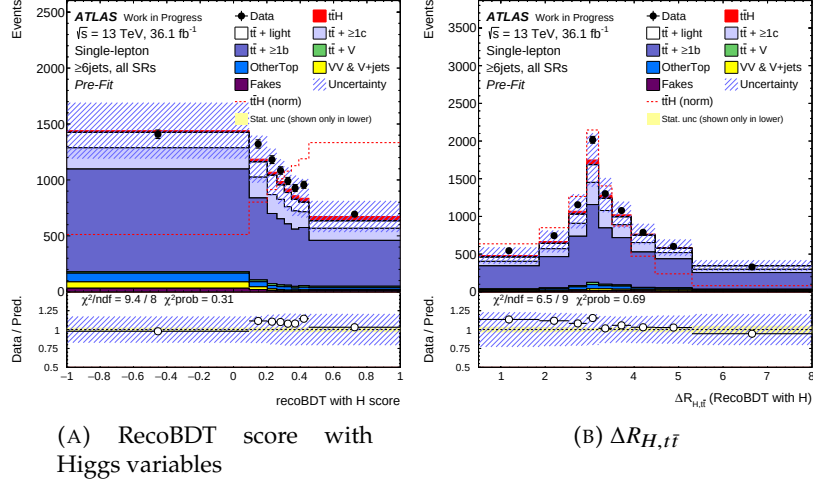
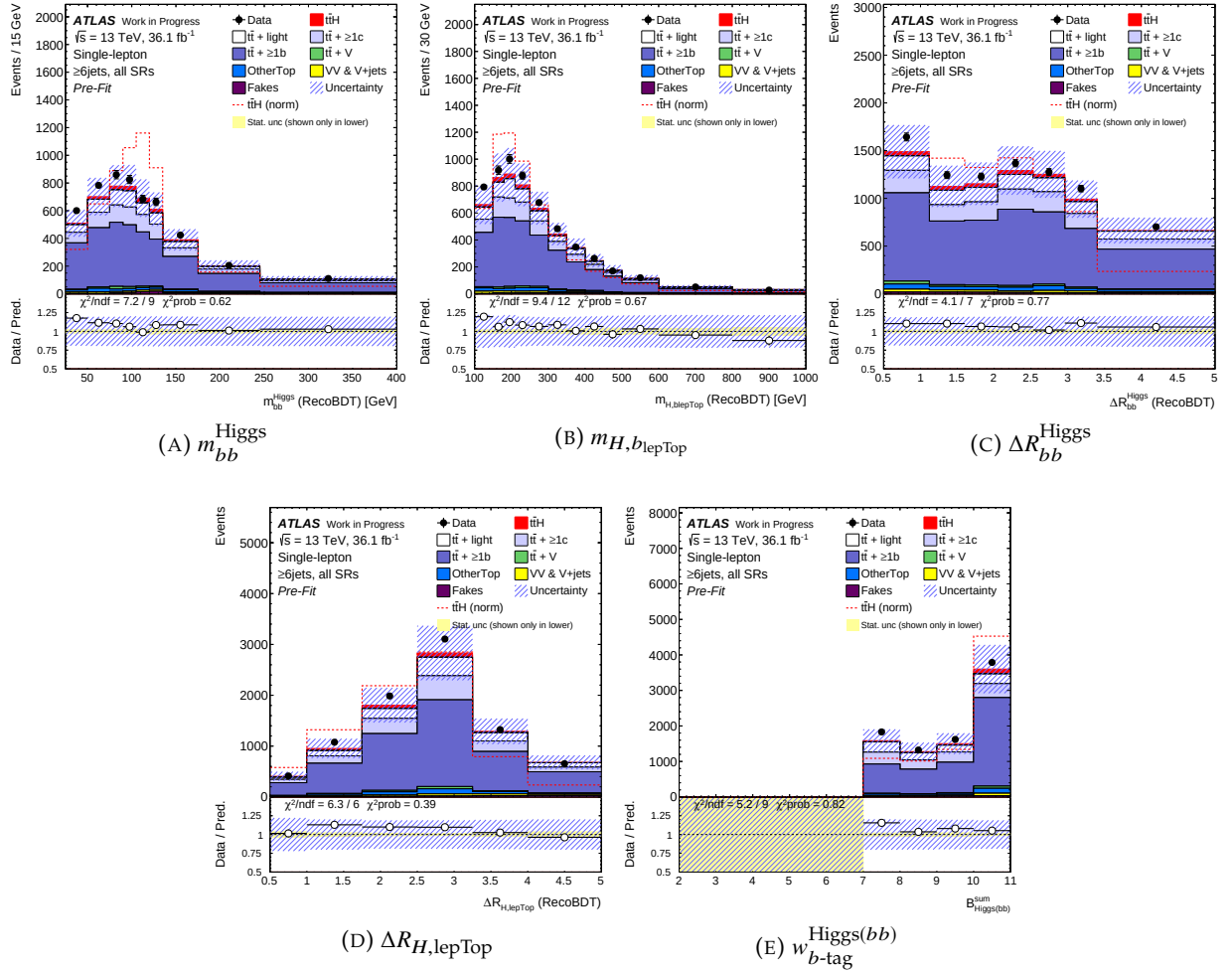
(C) 5 jets signal regions

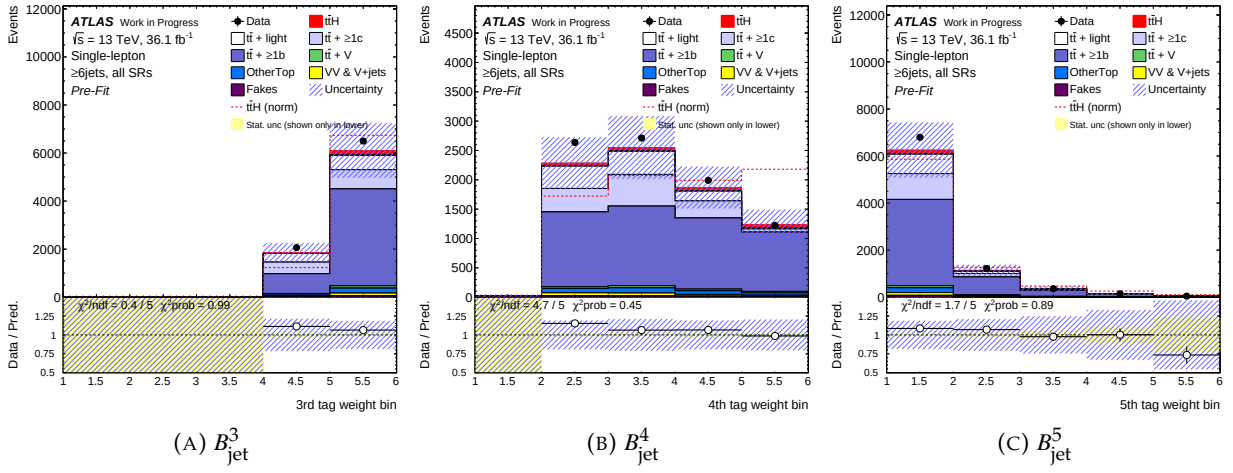
FIGURE 10.1: Linear correlations among input variables in ClassBDT for (A) $\geq 6jSR(t\bar{t}H)$, (B) ≥ 6 jets signal regions, and (c) 5 jets signal regions. Correlations were calculated with the signal events.

FIGURE 10.2: Input variables (event kinematics, LHD, and MEM) in ClassBDT for $\geq 6j\text{SR}(t\bar{t}H)$

FIGURE 10.3: Input variables (RecoBDT with H) in ClassBDT for $\geq 6jSR(t\bar{t}H)$ FIGURE 10.4: Input variables (RecoBDT w/o H) in ClassBDT for $\geq 6jSR(t\bar{t}H)$

FIGURE 10.5: Input variables (event kinematics and LHD) in ClassBDT for ≥ 6 jets signal regions

FIGURE 10.6: Input variables (RecoBDT with H) in ClassBDT for ≥ 6 jets signal regionsFIGURE 10.7: Input variables (RecoBDT w/o H) in ClassBDT for ≥ 6 jets signal regions

FIGURE 10.8: Input variables (b -tagging variables) in ClassBDT for ≥ 6 jets signal regions

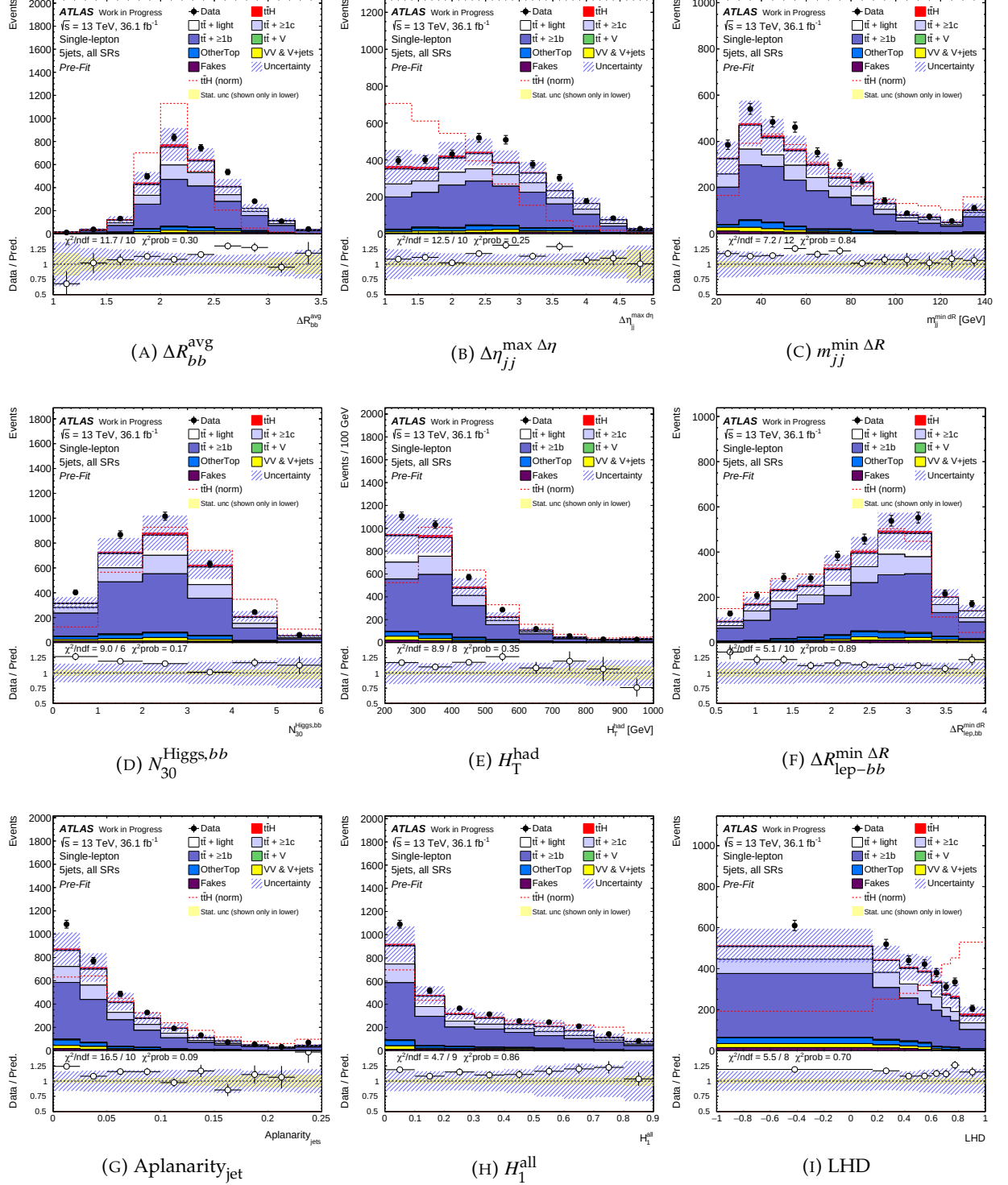


FIGURE 10.9: Input variables (event kinematics and LHD) in ClassBDT for 5 jets signal regions

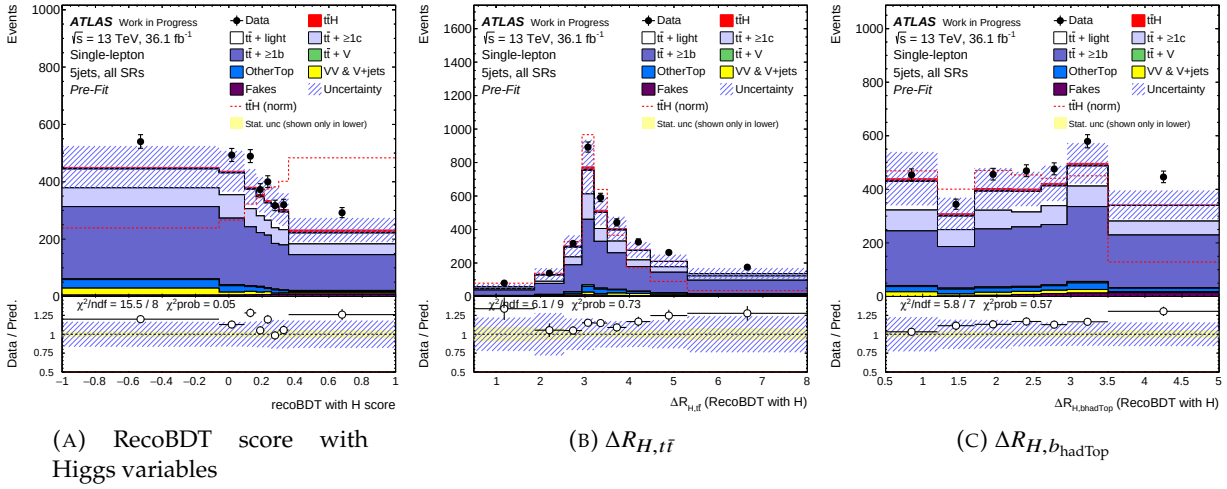
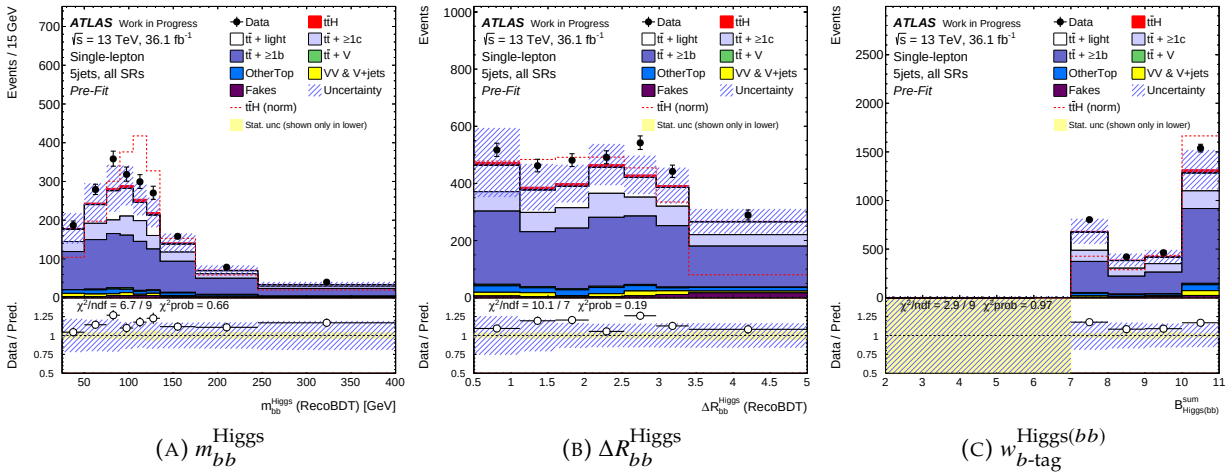
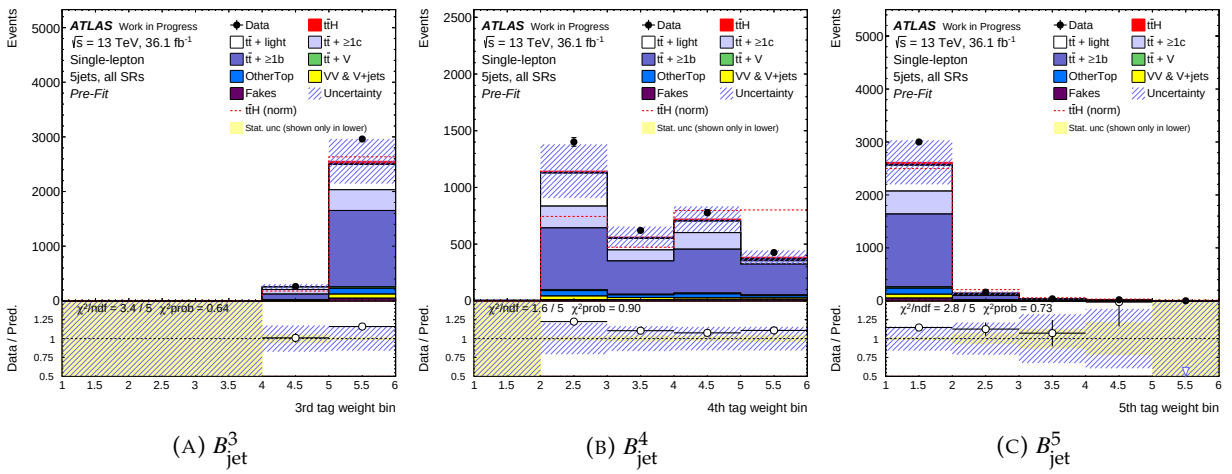
FIGURE 10.10: Input variables (RecoBDT with H) in ClassBDT for 5 jets signal regionsFIGURE 10.11: Input variables (RecoBDT w/o H) in ClassBDT for 5 jets signal regions

FIGURE 10.12: Input variables (b-tagging variables) in ClassBDT for 5 jets signal regions

10.2 Single-lepton Boosted Channel

For the boosted channel, one dedicated ClassBDT was built using reclustering information summarized in Table 10.3. The comparisons of data-MC for input variables are in Figure 10.13.

TABLE 10.3: Input variables to the ClassBDTs in the boosted single-lepton signal region.

Variable	Description
Reclustering variables	
$m_{\text{rcjet}}^{\text{Higgs}}$	Higgs-tagged rcjet mass
$d12^{\text{top}}$	First splitting scale for top-tagged rcjet
$\Delta R_{H,b}$	ΔR between Higgs-tagged rcjet and b -jet out of Higgs- or top-tagged rcjets
$\Delta R_{T,b}$	ΔR between top-tagged rcjet and b -jet out of Higgs- or top-tagged rcjets
$\Delta R_{H,T}$	ΔR between Higgs-tagged rcjet and top-tagged rcjet
$\Delta R_{H,\ell}$	ΔR between Higgs-tagged rcjet and lepton
Variables from b -tagging	
$w_{b\text{-tag}}^{\text{sum}}$	Sum of b -tagging tag weight bins of all jets
$w_{b\text{-tag}}^b / w_{b\text{-tag}}^{\text{sum}}$	Ratio of sum of b -tagging tag weight bins of b -jets out of Higgs- or top-tagged rcjet to all jets

10.3 Dilepton Channel

In the dilepton channel, dedicated trainings in three signal regions, $\geq 4\text{jSR}(t\bar{t}H)$, $\geq 4\text{jSR}(t\bar{t}+\geq 2b)$ and $\geq 4\text{jSR}(t\bar{t}+1b)$, were applied to maximize their sensitivities. Same as for the single-lepton channel, various input variables were used such as simple kinematics built by a pair of any objects, event shape variable, b -tagging score variable, RecoBDT information and reconstructed kinematics of $t\bar{t}H$ system using assignments from RecoBDT. All input variables are shown in Table 10.4.

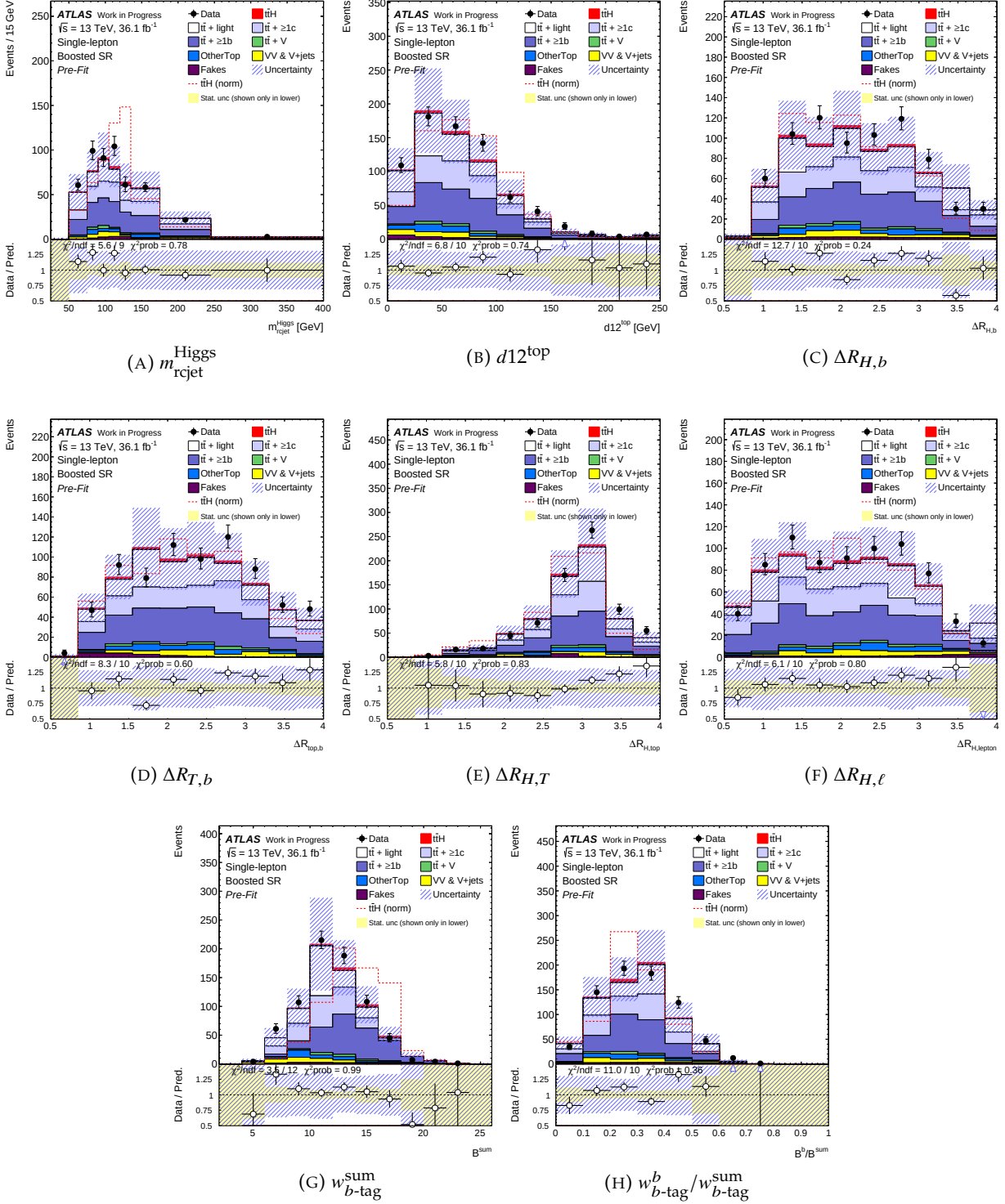
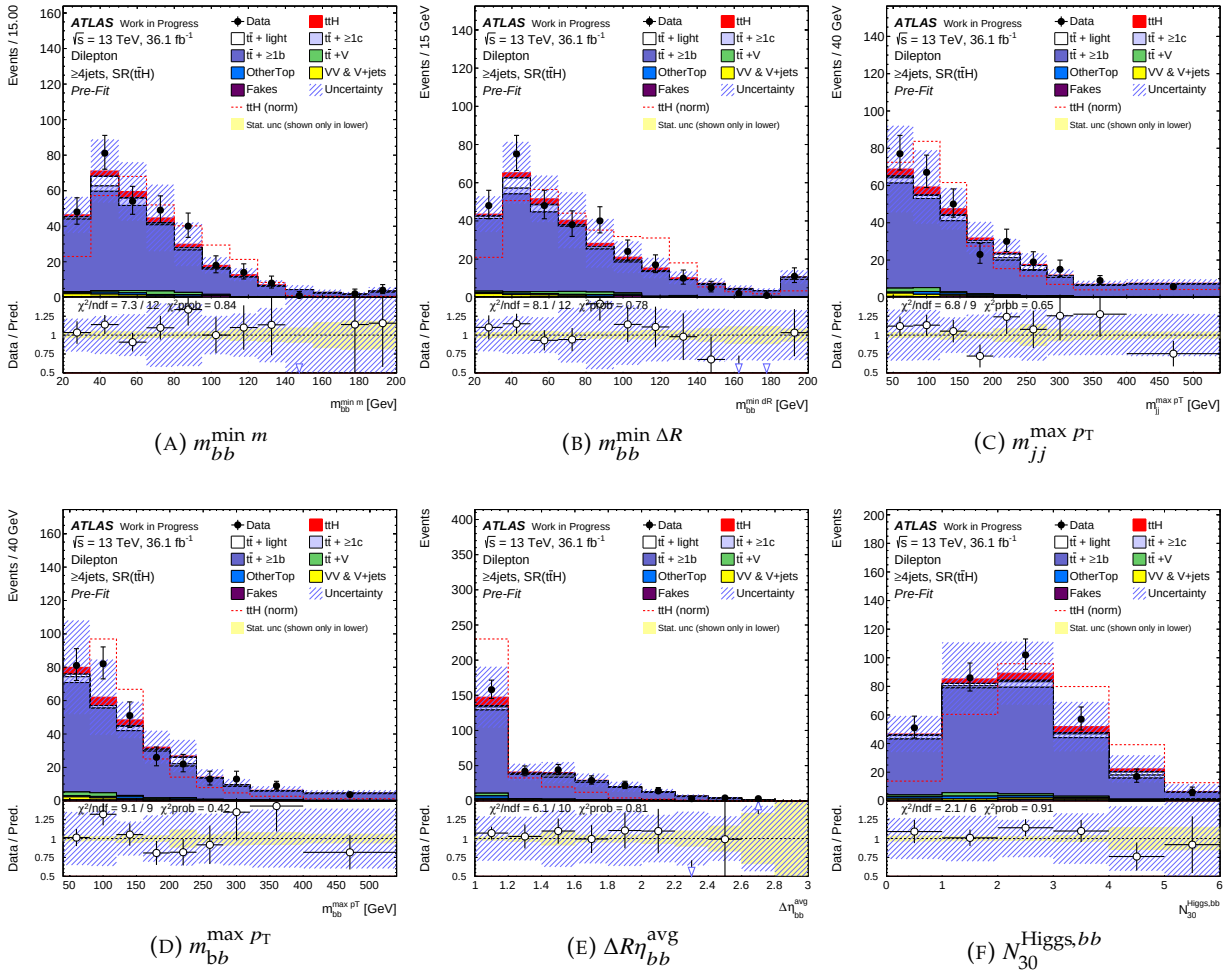


FIGURE 10.13: input variables in ClassBDT for boosted regions

TABLE 10.4: Variables used to the ClassBDTs in the dilepton signal regions.

Variable	Definition	SR($t\bar{t}H$)	SR($t\bar{t}+\geq 2b$)	SR($t\bar{t}+1b$)
General kinematic variables				
$m_{bb}^{\min m}$	Minimum invariant mass of a b -tagged jet pair	✓	✓	–
$m_{bb}^{\max m}$	Maximum invariant mass of a b -tagged jet pair	–	–	✓
$m_{bb}^{\min \Delta R}$	Invariant mass of the b -tagged jet pair with minimum ΔR	✓	–	✓
$m_{jj}^{\max p_T}$	Invariant mass of the jet pair with maximum p_T	✓	–	–
$m_{bb}^{\max p_T}$	Invariant mass of the b -tagged jet pair with maximum p_T	✓	–	✓
$\Delta R\eta_{bb}^{\text{avg}}$	Average $\Delta\eta$ for all b -tagged jet pairs	✓	✓	✓
$\Delta\eta_{lj}^{\max \Delta\eta}$	Maximum $\Delta\eta$ between a jet and a lepton	–	✓	✓
$\Delta R_{bb}^{\max p_T}$	ΔR between two b -tagged jets with maximum p_T	–	✓	✓
$N_{30}^{\text{Higgs},bb}$	Number of b -tagged jet pairs with m_{bb} within the Higgs mass ± 30 GeV	✓	✓	–
$n_{\text{jets}}^{p_T > 40}$	Number of jets with $p_T > 40$ GeV	–	✓	✓
Aplanarity $_b$	1.5 times the 2nd eigenvalue of the momentum tensor using b -tagged jets [106]	–	✓	–
H_T^{all}	Scalar sum of p_T of all jets and leptons	–	–	✓
Variables from RecoBDT with H				
RecoBDT score	Output score of the RecoBDT with H	✓	✓	–
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs and $t\bar{t}$ system	✓	–	–
Variables from RecoBDT w/o H				
RecoBDT score	Output score of the RecoBDT w/o H	✓	✓	✓
m_{bb}^{Higgs}	Higgs mass	✓	–	✓
$\Delta R_{H,\ell}^{\min}$	Minimum ΔR between Higgs and a lepton	✓	✓	✓
$\Delta R_{H,b}^{\min}$	Minimum ΔR between Higgs and one of b -jets from $t\bar{t}$	✓	✓	–
$\Delta R_{H,b}^{\max}$	Maximum ΔR between Higgs and one of b -jets from $t\bar{t}$	–	✓	–
$\Delta R_{bb}^{\text{Higgs}}$	ΔR between b -jets from the Higgs	–	✓	–
Variables from b -tagging				
$w_{b\text{-tag}}^{\text{Higgs}(bb)}$	Sum of b -tagging tag weight bins of jets from Higgs best assigned by the RecoBDT w/o H	–	✓	–

FIGURE 10.14: Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j\text{SR}(t\bar{t}H)$ region

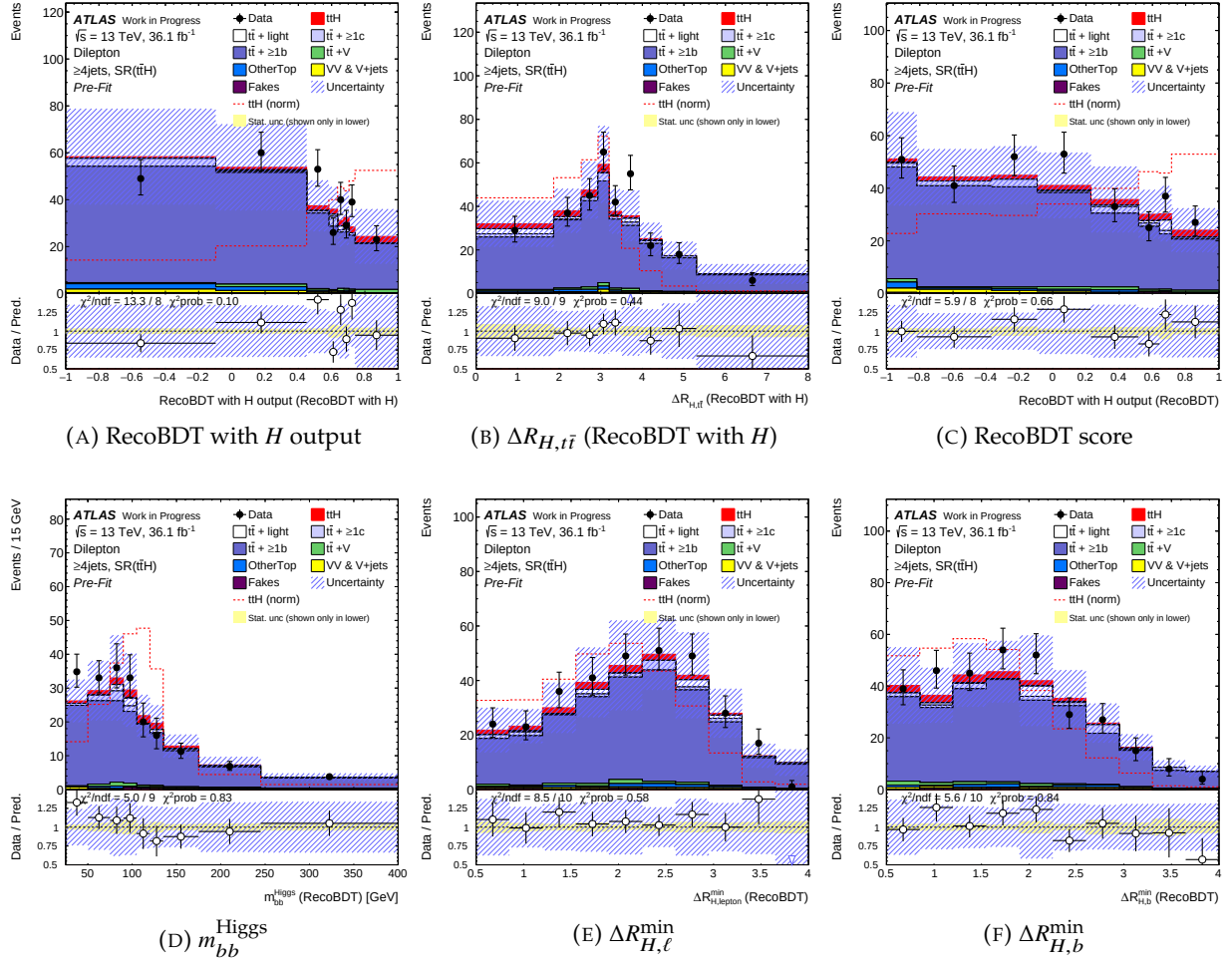
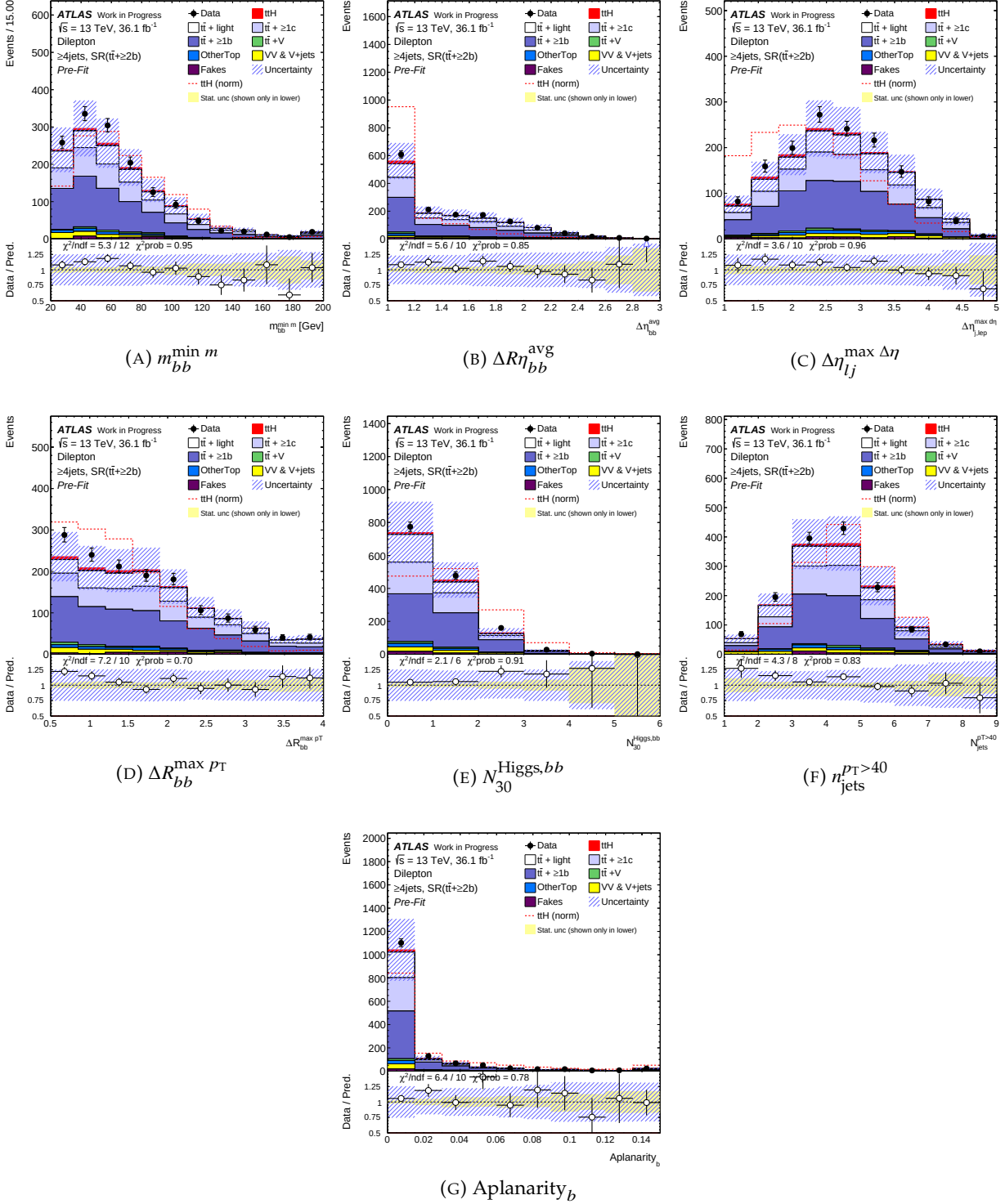


FIGURE 10.15: Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4jSR(t\bar{t}H)$ region. (A) and (B) are constructed by the RecoBDT with H , and remaining variables, (C) to (F) are constructed by the RecoBDT w/o H .

FIGURE 10.16: Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j\text{SR}(t\bar{t}+\geq 2b)$ region

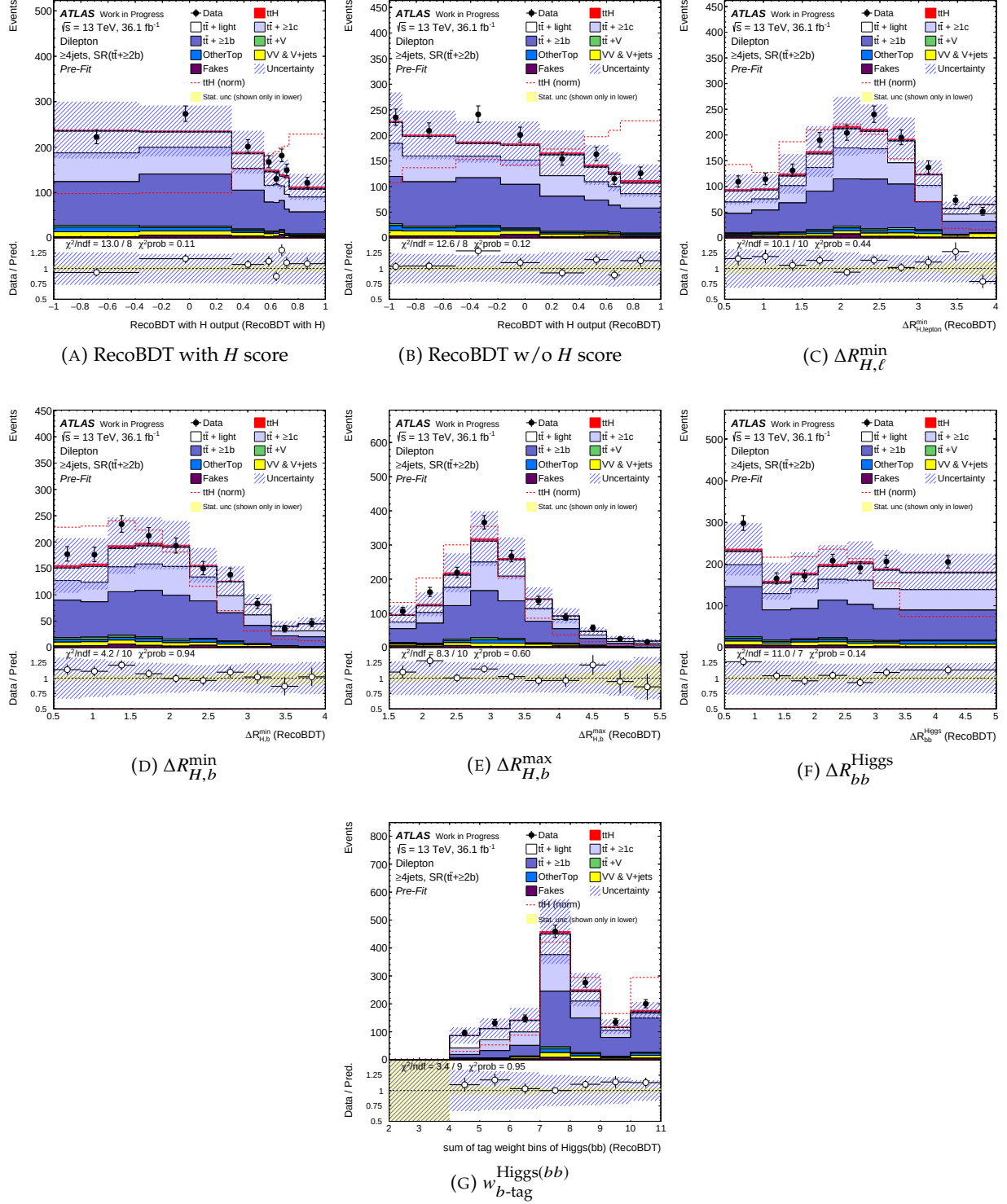
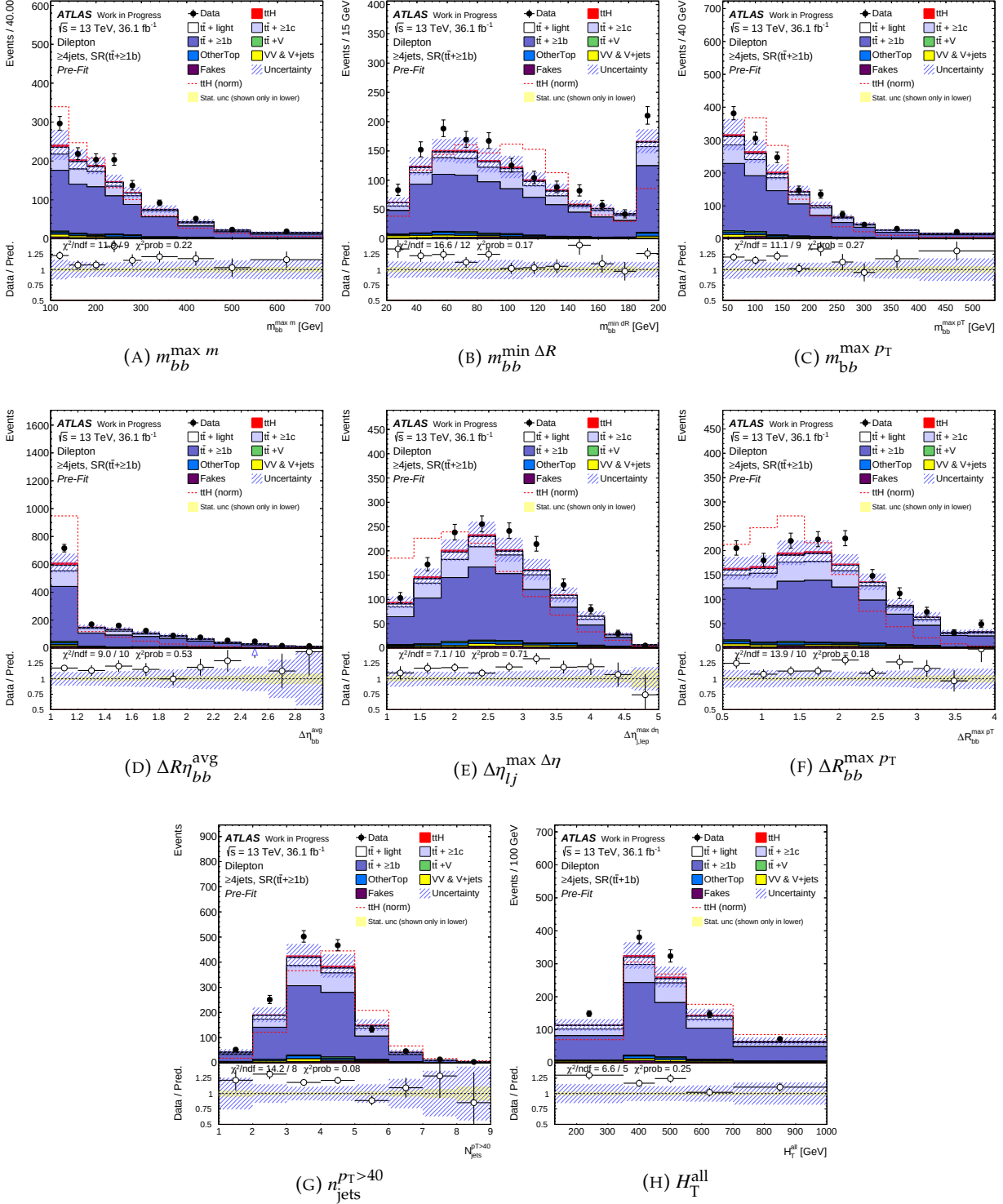


FIGURE 10.17: Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4jSR(t\bar{t} \geq 2b)$ region. (A) is constructed by the RecoBDT with H , and (B) to (F) are constructed by the RecoBDT w/o H .

FIGURE 10.18: Input variables (event kinematics) in ClassBDT for dilepton $\geq 4j\text{SR}(t\bar{t}+1b)$ region

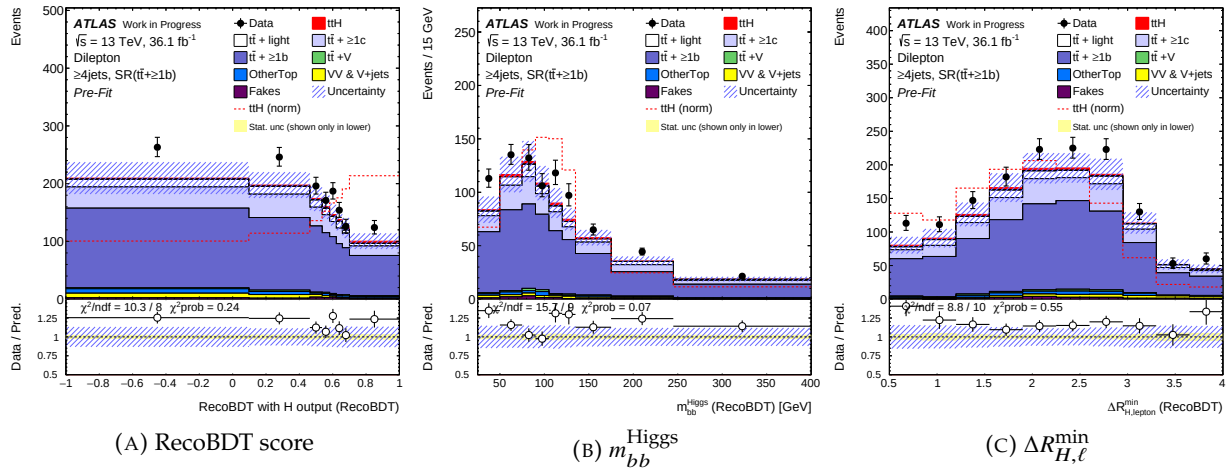


FIGURE 10.19: Input variables (RecoBDT) in ClassBDT for dilepton $\geq 4jSR(t\bar{t}+1b)$ region. All variables are constructed by the RecoBDT w/o H .

Chapter 11

Systematic Uncertainties

In the hadron collider experiment, various systematic uncertainties arise from various components.

We present typical uncertainties are shown in this chapter to explain the possible systematic variation. The H_T^{had} distributions of $t\bar{t}$ simulation in $\geq 6j\text{CR}(t\bar{t}+\geq 1c)$ show variations by the detector performance and object calibrations (Figures 11.1 to 11.5). To see the variation sensitivity to the signal strength, the ClassBDT distributions in $\geq 6j\text{SR}(t\bar{t}H)$ are also shown for the $t\bar{t}+\geq 1b$ and other large impact variations (Figures 11.6 to 11.12). In each plot, the simulation statistic uncertainties are also shown.

11.1 Luminosity and Pile-up

The collected luminosity is 36.1 fb^{-1} with relative uncertainty of $\pm 2.1\%$. The ATLAS luminosity scale was calibrated using data from dedicated x - y beam-separation scans (van der Meer scans) performed in August 2015 and May 2016 [108]. The largest uncertainty contribution comes from the transformation of the absolute luminosity scale from the low-rate vdM-scan regime to the high-luminosity conditions of physics operation.

The pileup reweighting uncertainty covers the uncertainty in the ratio of the predicted and measured inelastic cross sections in the fiducial volume [109].

11.2 Lepton and E_T^{miss}

Uncertainties in the lepton efficiency measurement (for the trigger, reconstruction, ID and isolation) were evaluated using the tag-and-probe method in $Z \rightarrow \ell\ell$ (ℓ is electron or muon) dominant event sample. For the muon efficiencies, all uncertainties were uncorrelated between statistic and systematic uncertainties. Any of the uncertainties were smaller than 0.5% for electrons with $p_T > 30 \text{ GeV}$ and for muons in whole p_T range. The energy scale and its resolution uncertainties were defined as $\pm 1\sigma$ variations of the lepton momentum propagating to distributions of masses of $Z \rightarrow \ell\ell$ and $J/\psi \rightarrow \ell\ell$ as well as E/p of $W \rightarrow \ell\nu$. For the muon, the sagitta measurement also provided systematic uncertainties.

The missing E_T (E_T^{miss}) uncertainties are mostly affected by uncertainties associated to leptons and jet energy scales and resolutions which are already considered. However, soft term contributions should be taken into account in addition. Soft term energy resolution uncertainties are calculated in two axis components, parallel and perpendicular, to minimize their uncertainties. The perpendicular axis is defined as the unit vector of sum of the unit momentum vector of all reconstructed objects. The parallel axis is defined as the vertical vector to the perpendicular vector.

Kinematics of the lepton and E_T^{miss} were not used in any selections in this analysis, and their systematic uncertainties are not significant in any distributions in the fit for determining the signal

strength. Therefore, they have almost no impacts on $t\bar{t}H(H \rightarrow b\bar{b})$ analysis. Most of them are less than 1% variations on any bins.

11.3 Jet

Uncertainties related to jets in principle come from the efficiency of jet reconstruction and its calibration (JES and jet energy resolution) as well as JVT variable discussed in Chapter 3, where reducing the JES residual p_T correction uncertainties into eight effective eigen-variations were described. Thirteen uncertainties were derived from variations of calibration parameters for reconstructed jets. Although these uncertainties are not large, ranging from 1% to 5.5% per jet depending on its p_T , the effects become larger in the high jet multiplicity environment such as in the $t\bar{t}H(H \rightarrow b\bar{b})$ final state. The uncertainties in the jet energy resolution and in the JVT efficiency were also considered. Because the jet energy resolution uncertainties in two regions, 5jCR($t\bar{t}+\geq 1c$) and 3jCR($t\bar{t}$ +light), have different behaviors from those in other regions in the fit, the jet energy resolution uncertainty was divided into two uncorrelated components, one is for 5jCR($t\bar{t}+\geq 1c$) and 3jCR($t\bar{t}$ +light), the other is for other regions. The systematic variations which have large impacts on the signal strength or large correlations, pulls, and constraints are shown in Figures 11.1 to 11.2.

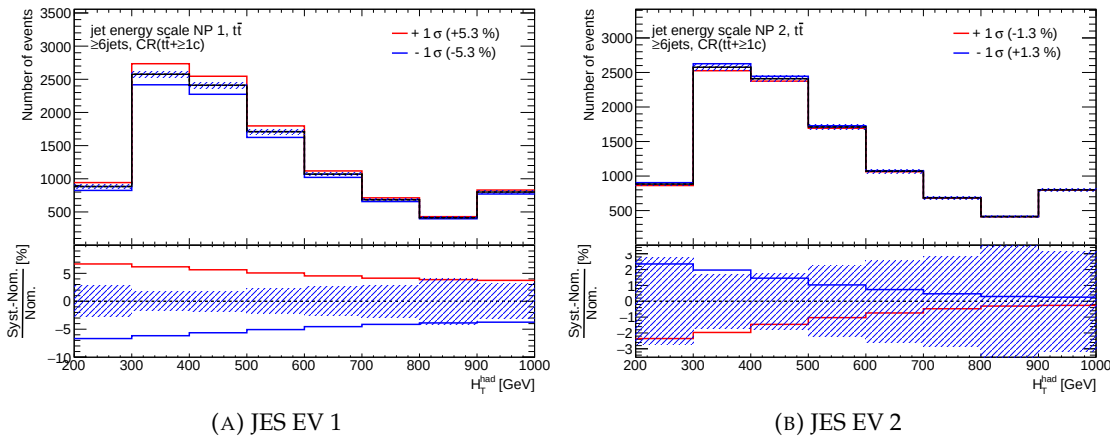


FIGURE 11.1: Systematic uncertainties of eigen-variations in the jet energy scale

11.4 Flavor tagging

Flavor tagging correction scale factors for b -, c - and light-jets were considered separately in dedicated calibrations. For b -jet and c -jet, scale factors were derived as a function of p_T , while the light-jet scale factors were derived as functions of two dimensional bins of p_T and η . Each of flavor jet uncertainties was provided by effective eigen-variation, and totally 30, 20 and 80 uncertainties were provided for b -jets, c -jets, and light-jets, respectively. The uncertainty associated to the b -jets ranges from 2 to 10% depending on the p_T and tag-weight. The size of variations on c mis-tagging is between 5 to 20%, and that on light-flavor mis-tagging 10 to 50%. Large amount of uncertainties are pruned out due to very small contributions less than 1% variations in whole regions, therefore, only large systematic variations are shown in Figures 11.3 to 11.5.

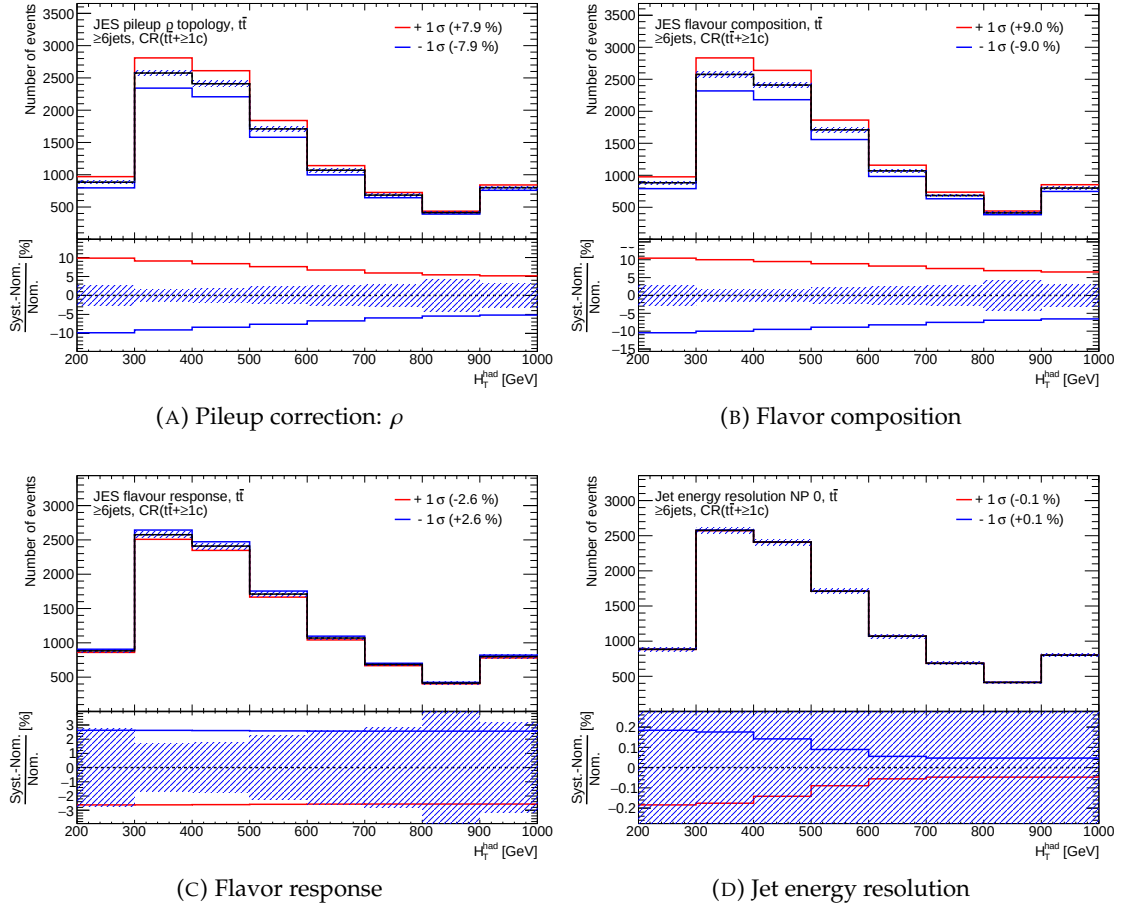
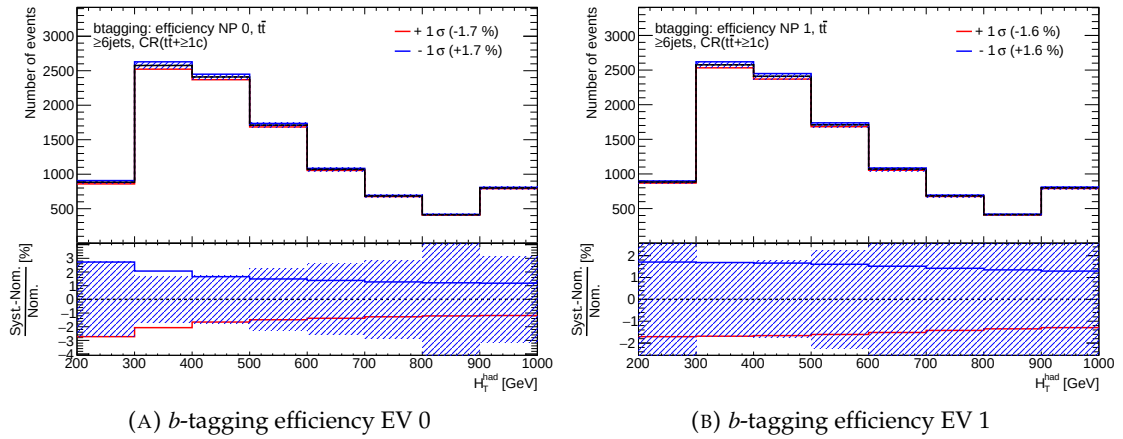


FIGURE 11.2: Systematic uncertainties of calibration parameters in the jet energy scale and energy resolution

FIGURE 11.3: Systematic uncertainties on b -jet tagging efficiency.

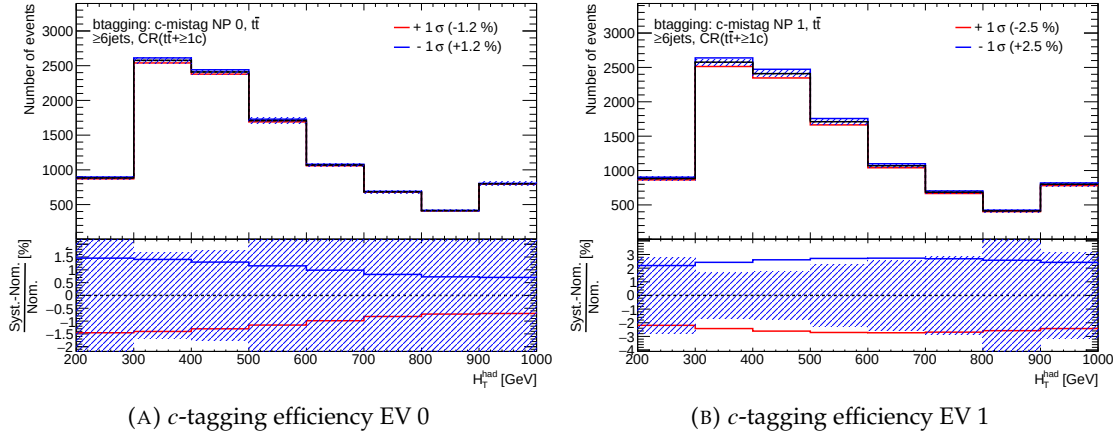
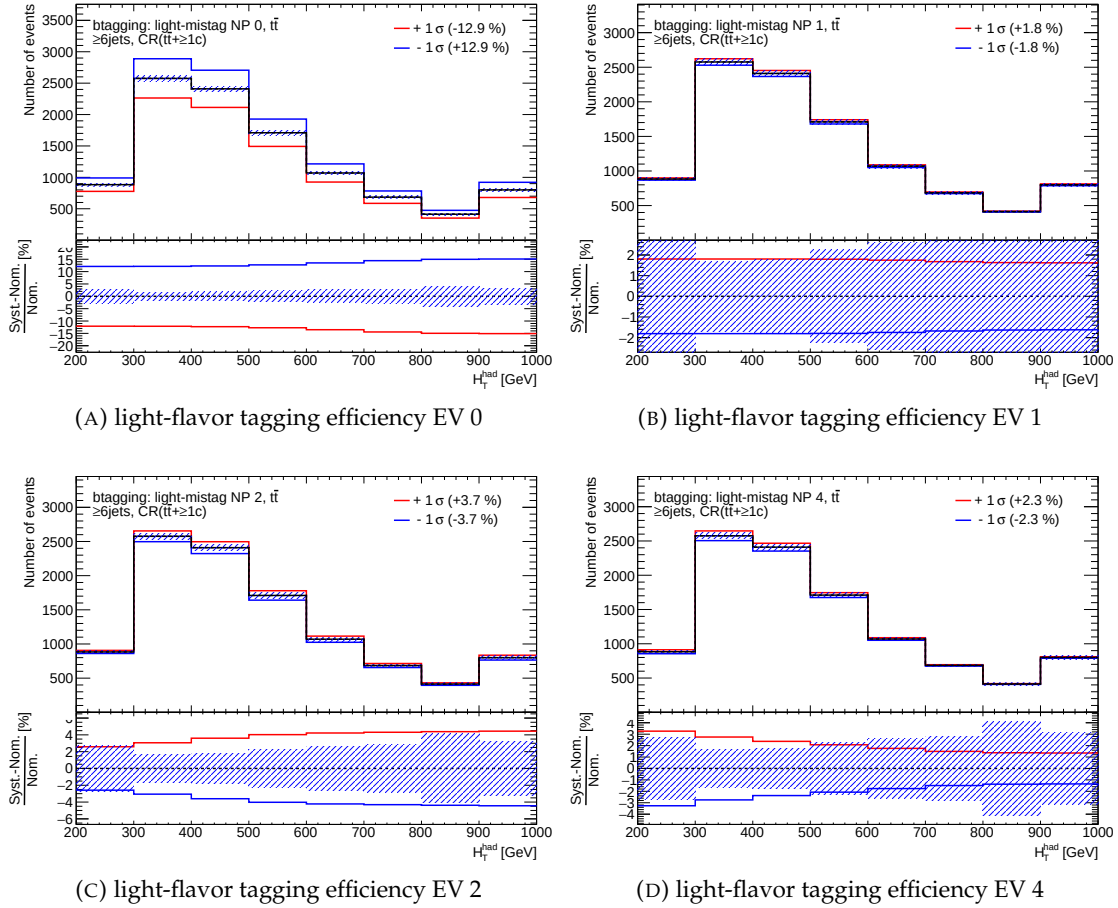
FIGURE 11.4: Systematic uncertainties on c -jet mis-tagging efficiency.

FIGURE 11.5: Systematic uncertainties on light-jet mis-tagging efficiency.

11.5 Cross Sections

All MC samples used the theoretical cross sections to normalized the event samples to the data with 36.1 fb^{-1} . The uncertainties of the normalization were provided from the QCD scale variation

and from the PDF and α_S variations. For insignificant backgrounds, the overall normalization uncertainties were given conservatively covering the detector acceptance uncertainties as well as QCD, PDF and α_S variations. The cross section and normalization uncertainties used in this analysis are summarized in Table 11.1.

Considering the uncertainty in $t\bar{t}$, two normalization factors to correct the cross sections of $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ were assigned in addition to the $t\bar{t}$ inclusive cross section uncertainty because of poor $t\bar{t}$ +HF modeling.

An uncertainty of 40% was assumed for the W +jets cross section, with an additional 30% normalization uncertainty used for W plus heavy-flavor jets, uncorrelated with 2 and ≥ 3 heavy-flavor jets in the single-lepton channel. These uncertainties were based on variations of the factorization and renormalization parameters in the SHERPA simulation. An uncertainty of 35% was applied to the Z +jets normalization, uncorrelated with the 3 j and ≥ 4 j regions to account for the variations in the SHERPA simulation as well as the uncertainty from data driven estimate for the heavy flavor component in the dilepton channel. A 50% normalization uncertainty on the VV was used, which includes uncertainties on the inclusive cross section and additional jet production.

A 50% uncertainty was assigned to the data driven estimated non-prompt leptons in the single lepton channel, uncorrelated into six uncertainties by 5j, $\geq 6j$, and boosted categories and by lepton flavors. In the dilepton channel, the non-prompt lepton background was assigned an overall 25% uncertainty.

TABLE 11.1: List of cross section normalization uncertainties. Sample name with 1ℓ and 2ℓ denote the sample used in the single-lepton and dilepton channels, respectively.

Sample	QCD scale	PDF and α_S	normalization
$t\bar{t}H$	+5.8%/-9.2%	$\pm 2.6\%$	cross section is free parameter as $\mu_{t\bar{t}H} = \sigma_{t\bar{t}H} / \sigma_{t\bar{t}H}^{\text{SM}}$
WtH	+6.5%/-6.7%	+6.3%	—
tH	+6.5%/-14.9%	+3.7%	—
$t\bar{t}Z$	+9.6%/-11.3%	+4.0%	—
$t\bar{t}W$	+12.9%/-11.5%	+3.4%	—
$t\bar{t}$	$\pm 6\%$		additional normalizations to $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$
Wt	+5%/-4%		—
single-top(t-ch)	+5%/-4%		—
single-top(s-ch)	+5%/-4%		—
WtZ	$\pm 50\%$		—
tZ	+7.7%/-7.9%	$\pm 0.9\%$	—
$t\bar{t}WW$	+10.9%/-11.8%	$\pm 2.1\%$	—
$t\bar{t}t\bar{t}$	$\pm 50\%$		—
W +jets(1ℓ)	$\pm 40\%$		$\pm 40\%$ on 2-HF and ≥ 3 -HF categories
Z +jets(1ℓ)	$\pm 35\%$		—
Z +jets(2ℓ)	—		$\pm 35\%$ on 3-jet and ≥ 4 -jet categories
VV	$\pm 50\%$		—
fake-leptons(1ℓ)	—		$\pm 50\%$ on $e+5j$ $\mu+5j$ $e+\geq 6j$ $\mu+\geq 6j$ e +boosted, μ +boosted
fake-leptons(2ℓ)	$\pm 25\%$		—

Furthermore, other prediction uncertainties are on the branching-ratios of the Higgs-boson decays, 2.2% for $H \rightarrow b\bar{b}$ [110].

11.6 $t\bar{t}H$ Modeling

As the $t\bar{t}H$ signal modeling besides the normalization uncertainties, the PS modeling variation was assigned by comparing MG5_aMC@NLO+HERWIG++ to the nominal MG5_aMC@NLO+PYTHIA8. Typical uncertainty shapes of this modeling are shown in Figure 11.6 in the $\geq 6jSR(t\bar{t}H)$.

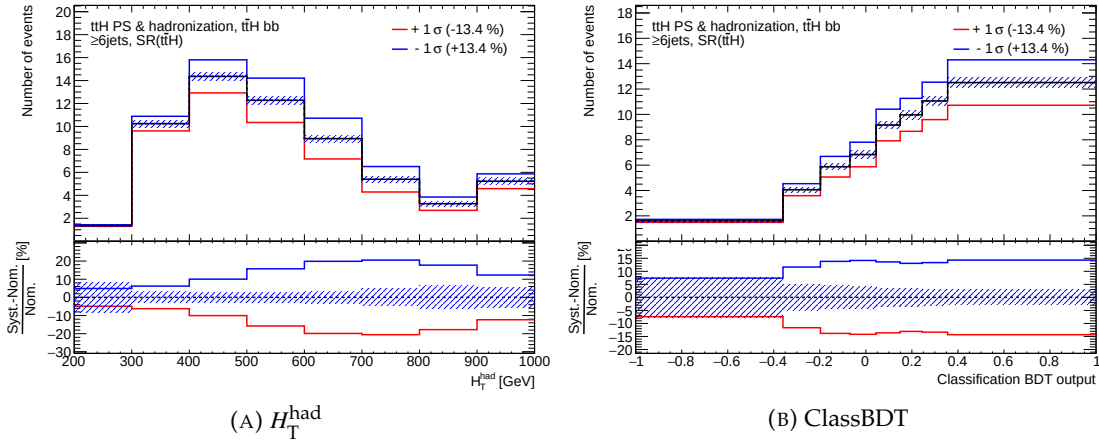


FIGURE 11.6: PS modeling variations for $t\bar{t}H$ events (HERWIG++ vs. PYTHIA8) in $\geq 6jSR(t\bar{t}H)$.

11.7 $t\bar{t}$ Modeling

In addition to cross section uncertainties, various $t\bar{t}$ modeling uncertainties were taken into account considering many aspects of modeling tunes and alternative samples with optional parameters as described in Chapter 6. Those systematic uncertainty sources are shown in Table 11.2. At first, the $t\bar{t}$ events have one systematic variation uncorrelated with flavor components ($t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$, and $t\bar{t}+\text{light}$) by comparing alternative MC sample (SHERPA5F) to the nominal sample POWHEG+PYTHIA8 as to examine the inclusive variation of generators, PS and ISR/FSR tunes. Typical variations in $\geq 6jCR(t\bar{t}+\geq 1c)$ are shown in Figure 11.7. In addition, two kinds of dedicated comparisons were also assigned as variations of the PS comparing POWHEG+HERWIG7 to the nominal, and variations of the ISR/FSR tunes using radHi and radLo samples, which are shown in Figures 11.8 and 11.9. Alternative samples were generated by the AFII detector simulation instead of the full detector simulation (FS). Therefore the relative differences between nominal sample but produced with AFII and the alternative sample were applied to the nominal (FS) as 1σ systematic variations.

The $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ samples have further systematic sources with respect to the flavor-scheme in the PDF. Two alternative samples were generated for $t\bar{t} + c\bar{c}$ and $t\bar{t} + b\bar{b}$ productions at NLO. A dedicated $t\bar{t}+\geq 1c$ sample was produced by MG5_aMC+HERWIG++ with three flavor-scheme (3F) PDF, while dedicated $t\bar{t}+\geq 1b$ sample was produced by SHERPAOL with four flavor-scheme (4F) PDF. The $t\bar{t}+\geq 1b$ variation was assigned by comparing SHERPAOL 4F and nominal POWHEG+PYTHIA8, as shown in Figure 11.10. Furthermore, $t\bar{t} + b\bar{b}$ dedicated NLO calculation by SHERPAOL 4F cannot treat events with $b\bar{b}$ from the multi-parton interactions (MPI) which

were generated through the PS. To cover this uncertainty, a variation of $\pm 50\%$ was assigned to the $t\bar{t} + \geq 1b$ (b -quarks from MPI) normalization. The $t\bar{t} + \geq 3b$ was assigned with an additional variation of $\pm 50\%$ because its reweighting factor ~ 2 is large requiring conservative treatment. For $t\bar{t} + \geq 1c$ events, variations were assigned as reweighting factors from the POWHEG+PYTHIA 8 to MG5_AMC+ HERWIG++3F for $t\bar{t} + \geq 1c$ subcomponents defined as following:

- $t\bar{t} + c\bar{c}$ reweighted using top p_T , ttbar p_T , ΔR between additional $c\bar{c}$ and p_T of the additional $c\bar{c}$ system sequentially.
- $t\bar{t} + c$ reweighted using top p_T , ttbar p_T and additional c -jet p_T sequentially.
- $t\bar{t} + C$ reweighted using top p_T , ttbar p_T and additional c -jet p_T sequentially.

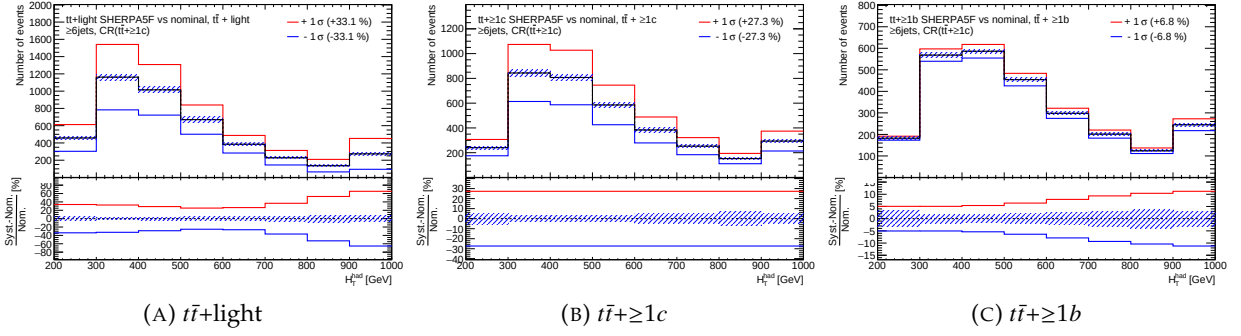
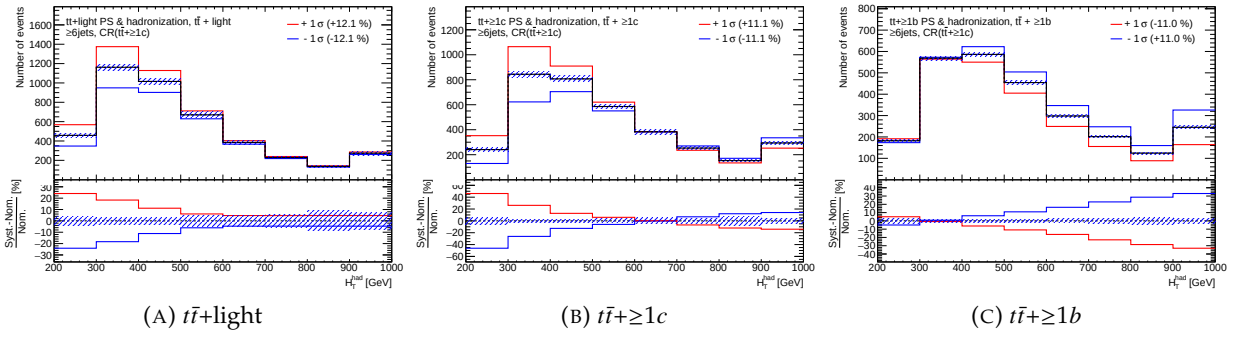
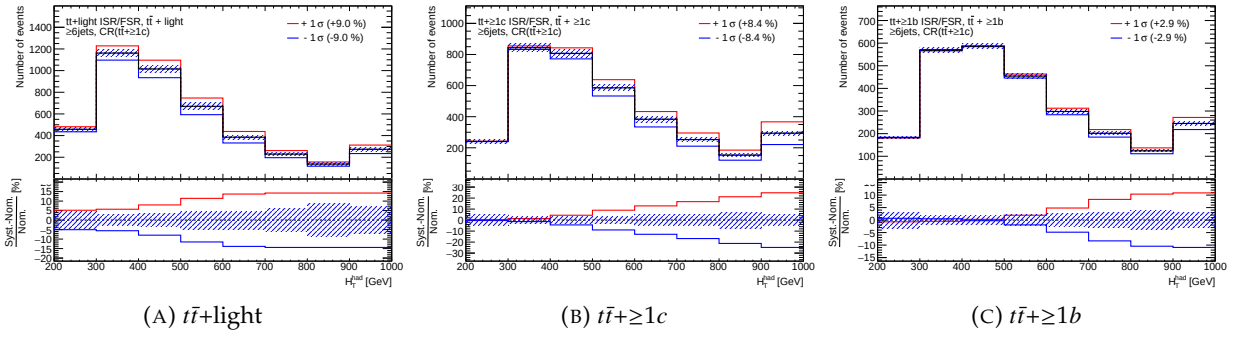
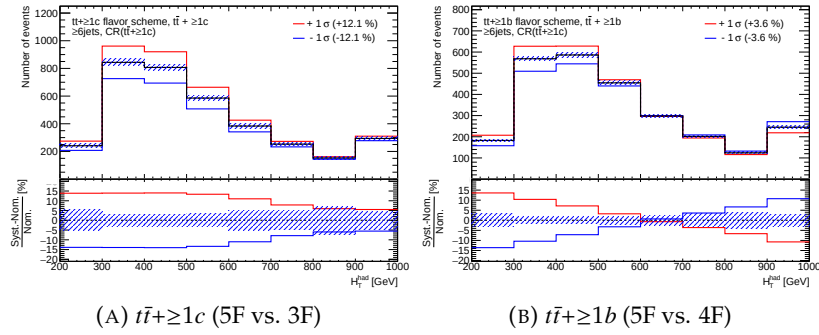
This is also shown in Figure 11.10.

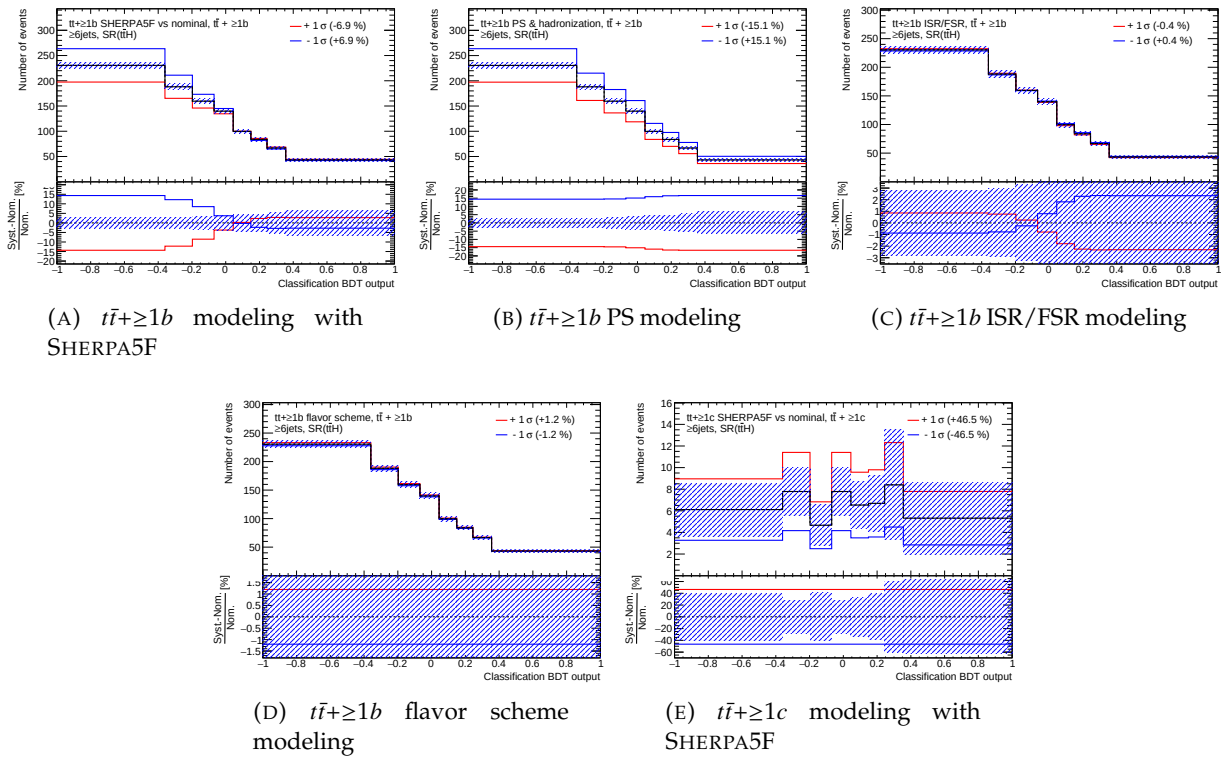
In addition to the $t\bar{t} + \geq 1b$ inclusive variation, sub-components contributing to $t\bar{t} + \geq 1b$ were also varied with scale and PDF parameters used in SHERPAOL 4F. The systematic variations were assigned by reweighting POWHEG+PYTHIA 8 to seven varied setups of SHERPAOL instead of nominal SHERPAOL.

TABLE 11.2: Summary of the systematic uncertainty sources in $t\bar{t}$ +jets modeling

Systematic variation	Description	applied categories
$t\bar{t}$ variations (uncorrelated in $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t}$ +light)		
SHERPA vs. nominal	Choice of generator, PS and ISR/FSR tunes	All
PS & hadronization	POWHEG+HERWIG7 vs. POWHEG+PYTHIA 8 (nominal)	All
ISR / FSR	radHi(radLo) vs. nominal for up(down) variation	All
$t\bar{t} + \geq 1c$ flavor scheme	$t\bar{t} + c\bar{c}$ dedicated production by MG5_AMC+ HERWIG++(3F) vs. inclusive nominal production (5F)	$t\bar{t} + \geq 1c$
$t\bar{t} + \geq 1b$ flavor scheme	$t\bar{t} + b\bar{b}$ dedicated production by SHERPAOL (4F) vs. inclusive nominal production (5F)	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ subcomponent variations by tuning parameters of SHERPA4F for reweighting		
$t\bar{t} + \geq 1b$ renormalization scale	Up or down by a factor of two	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ resummation scale	Vary μ_Q from $H_T/2$ to μ_{CMMPs}	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ global scales	Set μ_Q , μ_R , and μ_F to μ_{CMMPs}	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ PDF (MSTW)	MSTW vs. CT10(nominal)	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ PDF (NNPDF)	NNPDF vs. CT10(nominal)	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ UE	Alternative set of tunes for the underlying events	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ subcomponent normalization variations		
$t\bar{t} + \geq 1b$ MPI	Up or down by 50%	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 3b$ normalization	Up or down by 50%	$t\bar{t} + \geq 1b$

These variations are expected to have large impacts on the sensitivity of the signal strength. The variations in the ClassBDT distributions in $\geq 6j\text{SR}(t\bar{t}H)$ are shown in Figure 11.11.

FIGURE 11.7: H_T^{had} variations by $t\bar{t}$ modeling (POWHEG+PYTHIA 8 vs. SHERPA5F) in $\geq 6j\text{CR}(t\bar{t} \geq 1c)$.FIGURE 11.8: H_T^{had} variations by $t\bar{t}$ PS modeling (PYTHIA 8 vs. HERWIG7) in $\geq 6j\text{CR}(t\bar{t} \geq 1c)$.FIGURE 11.9: H_T^{had} variations by $t\bar{t}$ ISR / FSR modeling in $\geq 6j\text{CR}(t\bar{t} \geq 1c)$.FIGURE 11.10: H_T^{had} variations by $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ flavor scheme modeling in $\geq 6j\text{CR}(t\bar{t} \geq 1c)$.

FIGURE 11.11: ClassBDT variations by $t\bar{t}$ modeling in $\geq 6jSR(t\bar{t}H)$.

11.8 Other Background Modeling

For $t\bar{t}V$ modeling, the modeling difference between SHERPA and MG5_aMC@NLO+PYTHIA8 was assigned as the systematic uncertainty. The $t\bar{t}Z$ variations are shown in Figure 11.12.

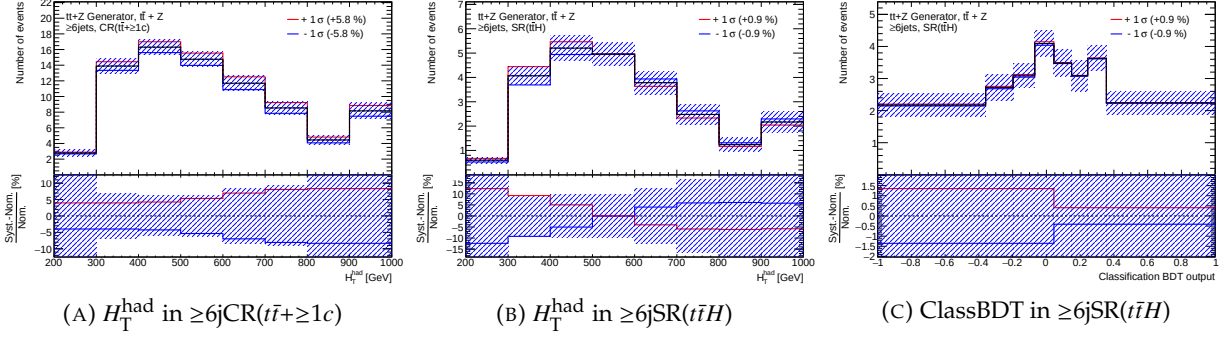


FIGURE 11.12: Variations by $t\bar{t}Z$ modeling (SHERPA vs. MG5_aMC@NLO+PYTHIA8)

In Wt and single-top t-channel, the systematic uncertainty was assigned with the PS and hadronization variation comparing POWHEG+PYTHIA6 (nominal) and POWHEG+HERWIG++ (alternative sample), and the ISR/FSR variation comparing different radiation tune parameters (radHi and radLo). In addition, the Wt sample has another systematic uncertainty estimated comparing different overlap subtraction schema against $t\bar{t}$ events.

The alternative samples for $t\bar{t}V$, Wt and single-top t-channel use AFII except for the study of the overlap subtraction variation in Wt . Therefore, same as for $t\bar{t}$ sample, the nominal $t\bar{t}V$, Wt and single-top t-channel samples using AFII instead of FS were compared to the alternative samples, and their relative differences were applied to the FS nominal samples.

11.9 Pruning and Categorization of Systematics

Among the various systematic uncertainties in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, some of them have only small variations and almost no impacts on the result. If a certain uncertainty has less than 1% effect to an analysis region, its normalization variation is removed from the global fit. If the variations associated to a certain systematic uncertainty are less than 1% in any histogram bins of an analysis region, its shape variation is also removed from the fit.

The uncertainties discussed in the above sections are summarized in Table 11.3 in terms of the number of uncorrelated parameters, their types (normalization and/or shape variations) and total provided numbers in all analysis regions.

TABLE 11.3: List of systematic uncertainties considered in the analysis. An "N" ("S") means that the uncertainty is taken as a normalization factor (a shape variation) for all processes and channels. If both S and N are assigned in the uncertainty, the normalization and shape are fully correlated. The "Components" column shows the number of uncorrelated variations. The column denoted as "before" shows the number of parameters defined as the number of components multiplied by number of analysis regions using the NPs. Many parameters were removed by pruning in some regions, and the number of remaining parameters are shown in the column denoted as "after". The uncertainties in the single-lepton channel are shown in this table.

Systematic uncertainty	Type		Components	before		after	
	S	N		S	N	S	N
Luminosity	–	✓	1	–	168	–	168
Reconstructed Objects							
Electron trigger+reco+ID+isolation	✓	✓	4	672	672	2	0
Electron energy scale+resolution	✓	✓	2	336	336	1	6
Muon trigger+reco+ID+isolation	✓	✓	10	1680	1680	5	0
Muon momentum scale+resolution+sagitta	✓	✓	5	840	840	3	5
Tau detector, in-situ and model	✓	✓	3	504	504	0	0
Pileup modelling	✓	✓	1	168	168	20	95
Jet vertex tagger	✓	✓	1	168	168	0	33
Jet energy scale	✓	✓	20	3360	3360	208	945
Jet energy resolution	✓	✓	2	168	168	32	120
Missing transverse energy scale+resolution	✓	✓	3	504	504	16	18
b -tagging efficiency	✓	✓	30	5040	5040	42	487
c -mistagging efficiency	✓	✓	20	3360	3360	27	286
Light-mistagging efficiency	✓	✓	60	10080	10080	71	374
Mistag extrapolation $c \rightarrow \tau$	✓	✓	1	168	168	0	0
Background and Signal Modeling							
$t\bar{t}$ inclusive cross section	–	✓	1	–	36	–	36
$t\bar{t}+\geq 1c$: normalization (free floating)	–	✓	1	–	12	–	–
$t\bar{t}+\geq 1b$: normalization (free floating)	–	✓	1	–	12	–	–
$t\bar{t}+\geq 1b$: $t\bar{t}+\geq 3b$ normalization	✓	✓	1	12	12	6	11
$t\bar{t}+\geq 1b$: reweighting to SHERPAOL 4F	✓	✓	9	96	96	1	30
$t\bar{t}+\geq 1c$: flavor scheme modeling	✓	✓	1	12	12	5	12
$t\bar{t}+\geq 1b$: flavor scheme modeling	✓	✓	1	12	12	6	12
$t\bar{t}$ modeling: ISR/FSR	✓	✓	3	36	36	14	30
$t\bar{t}$ modeling: SHERPA vs. nominal	✓	✓	3	36	36	15	34
$t\bar{t}$ modeling: PS & hadronization	✓	✓	3	36	36	15	35
non top simulation cross section/normalization	–	✓	5	–	60	–	57
other top production cross sections	✓	✓	11	132	132	25	80
Wt and single-top (t-chan.) cross section	–	✓	3	12	36	4	32
Single-top modeling	✓	✓	5	60	60	5	53
Fakes normalization	–	✓	6	–	18	–	17
$t\bar{t}V$ cross section	–	✓	4	–	48	–	48
$t\bar{t}V$ modeling	✓	✓	2	24	24	7	21
$t\bar{t}H$ cross section	–	✓	2	–	72	–	72
$t\bar{t}H$ branching ratios	–	✓	3	–	36	–	36
$t\bar{t}H$ modeling	✓	✓	1	36	36	9	31

Chapter 12

Simultaneous Global Fit and Results

12.1 Fit Overview

A simultaneous global fit was performed to determine the number of $t\bar{t}H$ signal events with all systematic uncertainties included where two normalization factors for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ mis-modelings were set free-floating. The signal was determined as the ratio to the SM cross section expectation ($\sigma_{t\bar{t}H}^{\text{SM}}$), referred as the signal strength $\mu_{t\bar{t}H} = \sigma_{t\bar{t}H} / \sigma_{t\bar{t}H}^{\text{SM}}$.

The three parameters $\mathbf{x} = (\mu_{t\bar{t}H}, k_{t\bar{t}+\geq 1b}, k_{t\bar{t}+\geq 1c})$ were determined simultaneously by maximizing the likelihood of the agreement between the data and the simulation defined as:

$$\mathcal{L} = \prod_{i=0}^{\text{regions}} \left[\prod_{k=0}^{\text{bins}} \mathcal{P}_{\text{POIS}} \left(N_{ik}^{\text{obs}} \middle| N_{ik}^{\text{exp}}(\mathbf{x}, \gamma_{ik}, \boldsymbol{\alpha}) \right) \cdot \mathcal{P}_{\text{LN}} \left(1 \middle| \gamma_{ik}, \sigma_{\text{MCstat}}^{\text{rel}} \right) \right] \cdot \prod_{j=0}^{\text{syst}} \mathcal{P}_{\text{GAUS}}(0 \middle| \alpha_j, 1) \quad (12.1)$$

$$N_{ik}^{\text{exp}} = \left(\mu_{t\bar{t}H} \cdot N_{ik,t\bar{t}H}^{\text{exp}}(\boldsymbol{\alpha}) + k_{t\bar{t}+\geq 1b} \cdot N_{ik,t\bar{t}+\geq 1b}^{\text{exp}}(\boldsymbol{\alpha}) + k_{t\bar{t}+\geq 1c} \cdot N_{ik,t\bar{t}+\geq 1c}^{\text{exp}}(\boldsymbol{\alpha}) + N_{ik,\text{BKG}}^{\text{exp}}(\boldsymbol{\alpha}) \right) \cdot \gamma_{ik} \quad (12.2)$$

where the likelihood is expressed by Poisson probabilities ($\mathcal{P}_{\text{POIS}}$) describing the agreement of the data and simulated expectation in the k -th bin of the distribution in the i -th analysis region. The three free floating parameters, $\mathbf{x}$ are set to maximize the likelihood. To take into account of the fluctuation of the total expected events due to the MC statistical uncertainty, one nuisance parameter (NP) γ_{ik} was implemented as a scale factor corresponding to the relative MC statistical uncertainty ($\sigma_{\text{MCstat}}^{\text{rel}}$) for i -th bin in k -th region. The likelihood was maximized to change γ_{ik} with the penalty term which decreases the likelihood score when the NP deviates from the nominal value ($\gamma_{0,ik} = 1$). The penalty term is described as a log-normal probability using $\gamma_{ik} > 0$,

$$\mathcal{P}_{\text{LN}} \left(1 \middle| \gamma_{ik}, \sigma_{\text{MCstat}}^{\text{rel}} \right) = \frac{1}{\sqrt{2\pi} \ln \sigma_{\text{MCstat}}^{\text{rel}}} \exp \left[-\frac{(\ln \gamma_{ik})^2}{2(\ln \sigma_{\text{MCstat}}^{\text{rel}})^2} \right]. \quad (12.3)$$

Furthermore, fluctuations of all systematic uncertainties were also considered, as Gaussian probabilities using NPs θ_j for j -th systematic uncertainty ($\Delta\theta_j$). The systematic pulls were defined by $\alpha_j = (\theta_j - \theta_{0,j})/\Delta\theta_j$ where $\theta_{0,j}$ is the nominal value of the NP. The α_j should be zero if the NP stays at the nominal, and ± 1 if the NP deviates by $\pm 1\sigma$ of its systematic uncertainty. Therefore, each Gaussian probability is described as:

$$\mathcal{P}_{\text{GAUS}}(0 \middle| \alpha_j, 1) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(1 - \alpha_j)^2}{2} \right]. \quad (12.4)$$

For the instrumental uncertainties, all systematic pulls should be typically within $\pm 1\sigma$ because these uncertainties and nominal values were measured and optimized to be commonly adoptable for various physics analyses. On the other hand, for the modeling uncertainties, all systematic

pulls should be also inside the ± 1 band, but some could be constrained if their uncertainties are assigned conservatively.

Afterwards, the test statistic q_μ was built to set a limit for a given signal strength μ_L by the profiled likelihood ratio:

$$q_\mu = -2 \ln \Lambda = -2 \ln \left(\frac{\mathcal{L}(\text{data}|\mu_L, \hat{\hat{\theta}})}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})} \right) \quad \text{for } 0 \leq \hat{\mu} \leq \mu \quad (12.5)$$

where $\hat{\mu}$ and $\hat{\theta}$ are parameters that maximize the likelihood, and $\hat{\hat{\theta}}$ are nuisance parameters that maximize the likelihood with $\mu = \mu_L$. The probability (p -value) to have the observed data is defined under the two hypotheses:

- b only hypothesis, where no signal process is assumed ($\mu = 0$).
- $s + b$ hypothesis, where the signal strength is equal to μ_L .

Their p -values are described as:

$$p_0 = \int_{-\infty}^{q_0^{\text{obs}}} f(q_\mu|0) dq_\mu \quad (12.6)$$

$$p_{\mu_L} = \int_{q_{\mu_L}^{\text{obs}}}^{\infty} f(q_\mu|\mu_L) dq_\mu \quad (12.7)$$

where $f(q_\mu|\mu')$ denotes the PDF of the test statistics under the $\mu = \mu'$ hypothesis. When $p_{\mu_L}/(1 - p_0) \leq \alpha$, the hypothesis $\mu = \mu_L$ is excluded at $(1 - \alpha)$ confidence level (CL). To quote the 95% CL upper limit on the signal strength, μ_L is profiled until α is 0.05.

12.2 Fit Result

The analysis regions implemented in the fit are in total twelve for the single-lepton channel and seven for the dilepton channel. All signal regions, four (two) for the single-lepton (dilepton) channel, utilized ClassBDT score as the fit input variables, while two control regions, $\geq 6j\text{CR}(t\bar{t} \geq 1c)$, $5j\text{CR}(t\bar{t} \geq 1c)$ in the single-lepton channel, utilized the sum of jet p_T (H_T^{had}) as discussed in Chapter 8. The only normalizations, the numbers of events, were used in remaining regions in both channels. The background fractions, S/B and $S/\sqrt{B}$ values are summarized in Figure 12.1.

The following four fit configurations were considered:

- fit in the single-lepton channel
- fit in the dilepton channel
- two- μ combined fit: using both lepton channels with correlated NPs and normalizations ($k_{t\bar{t} \geq 1b}$ and $k_{t\bar{t} \geq 1c}$), and decorrelated $\mu_{t\bar{t}H}$
- combined fit: using both lepton channels with correlated NPs, normalizations ($k_{t\bar{t} \geq 1b}$ and $k_{t\bar{t} \geq 1c}$), and $\mu_{t\bar{t}H}$

The uncertainty sources resulted in noticeable fit pulls ($\hat{\alpha}_j$) and constraints $((\Delta\theta_j - \Delta\hat{\theta}_j)/\Delta\theta_j)$, where $\hat{\theta}_j$ is the j -th systematic uncertainty after the fit) are shown for different fit setups in Figure 12.2, where $\hat{\alpha}_j$ is the deviation from the original input normalized by the uncertainty originally

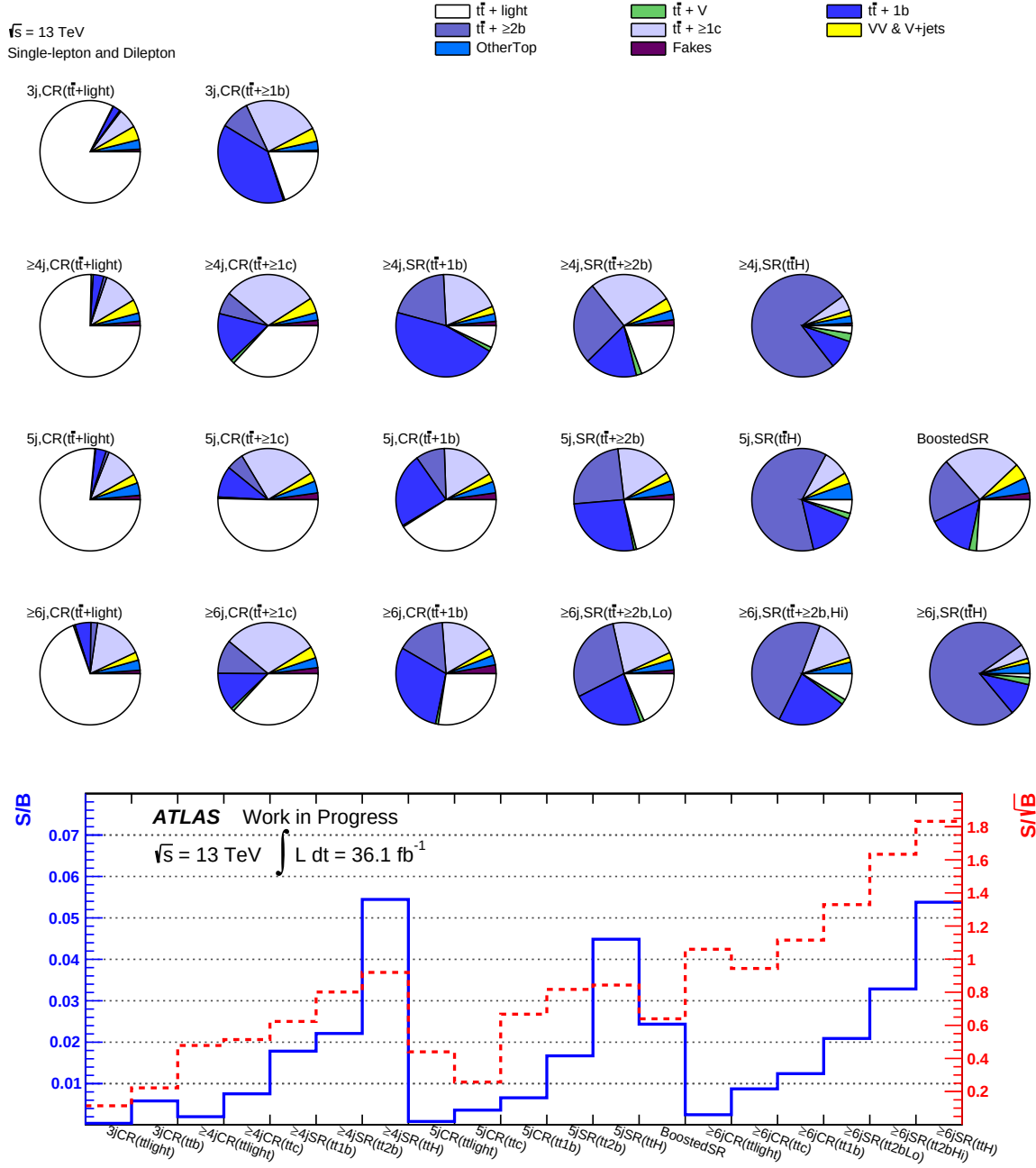


FIGURE 12.1: (Top) Fraction of background components in analysis regions. The first two rows show seven dilepton regions, and the next two rows show twelve single-lepton regions including one boosted signal region. In total nine signal regions (SRs) are defined, and remaining ten regions are treated as control regions (CRs). The region is named with the dominant physics process among $t\bar{t}H$ and $t\bar{t}$ +jets. (Bottom) S/B and $S/\sqrt{B}$ sensitivities in analysis regions. The blue line shows S/B in the left axis, and the red line shows $S/\sqrt{B}$ in the right axis.

evaluated and $\Delta\hat{\theta}_j/\Delta\theta_j$ is the ratio of the uncertainty between before and after the fit. Comparing the two- μ combined fit and the combined fit, all systematic uncertainties and normalization factors were not different each other (less than 5% differences), therefore, only combined fit results are shown in Figure 12.2 and Table 12.1. Pulls and constraints for all systematic variations

are shown in Appendix D. In all fit setups, $t\bar{t}$ modeling systematic uncertainties were constrained through fitting. All pulls are inside $\pm 1\sigma$ of original uncertainties (green band in Figure 12.2). The pulls agree with those derived from all fits. The c -mistagging NPs and $t\bar{t}+\geq 1c$ systematic variations have slightly larger pulls in the combined fit than individual fits because of lack of the $t\bar{t}+\geq 1c$ constraint in the individual fit.

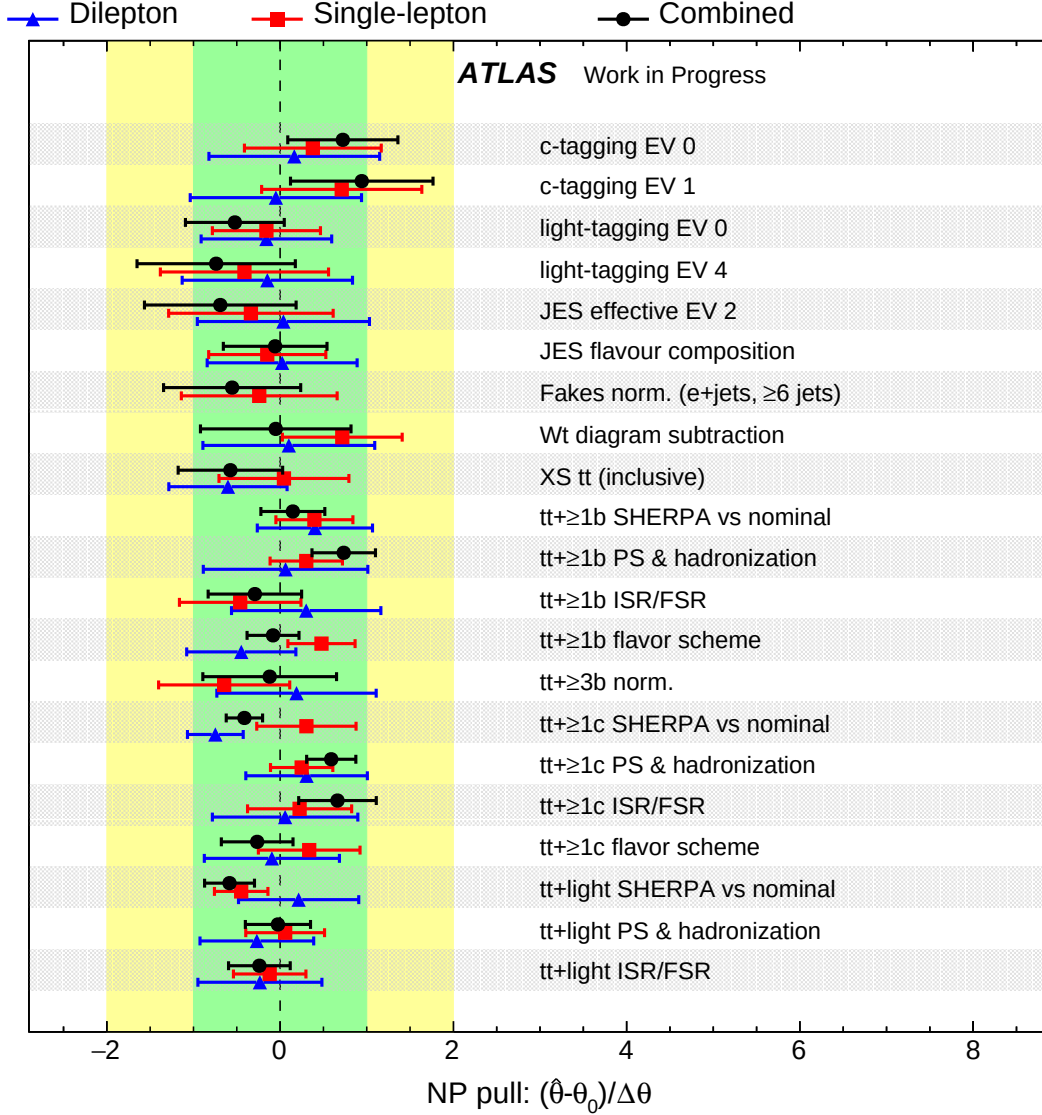


FIGURE 12.2: Fit pulls and constraints. Only systematic uncertainty sources with $> 0.5\sigma$ pulls or $> 20\%$ constraints from the original uncertainty band are shown in the plot. Black, red and blue points denote the fit with dilepton, single-lepton, and combined channel, respectively.

The linear correlations among the nuisance parameters are shown in Figure 12.3. The nuisance parameters having more than 40% correlations to other parameters are shown in this figure. Almost all parameters were from the $t\bar{t}$ +jets modeling uncertainties, and some of them had large correlations with the signal strength, $\mu_{t\bar{t}H}$.

Obtained normalization factors, $k_{t\bar{t}+\geq 1b}$ and $k_{t\bar{t}+\geq 1c}$ in all fits are summarized in Table 12.1. The expected values are obtained by fits with fixing the signal cross section at the SM prediction, the normalization factors at one, and all systematic variations at nominal values.

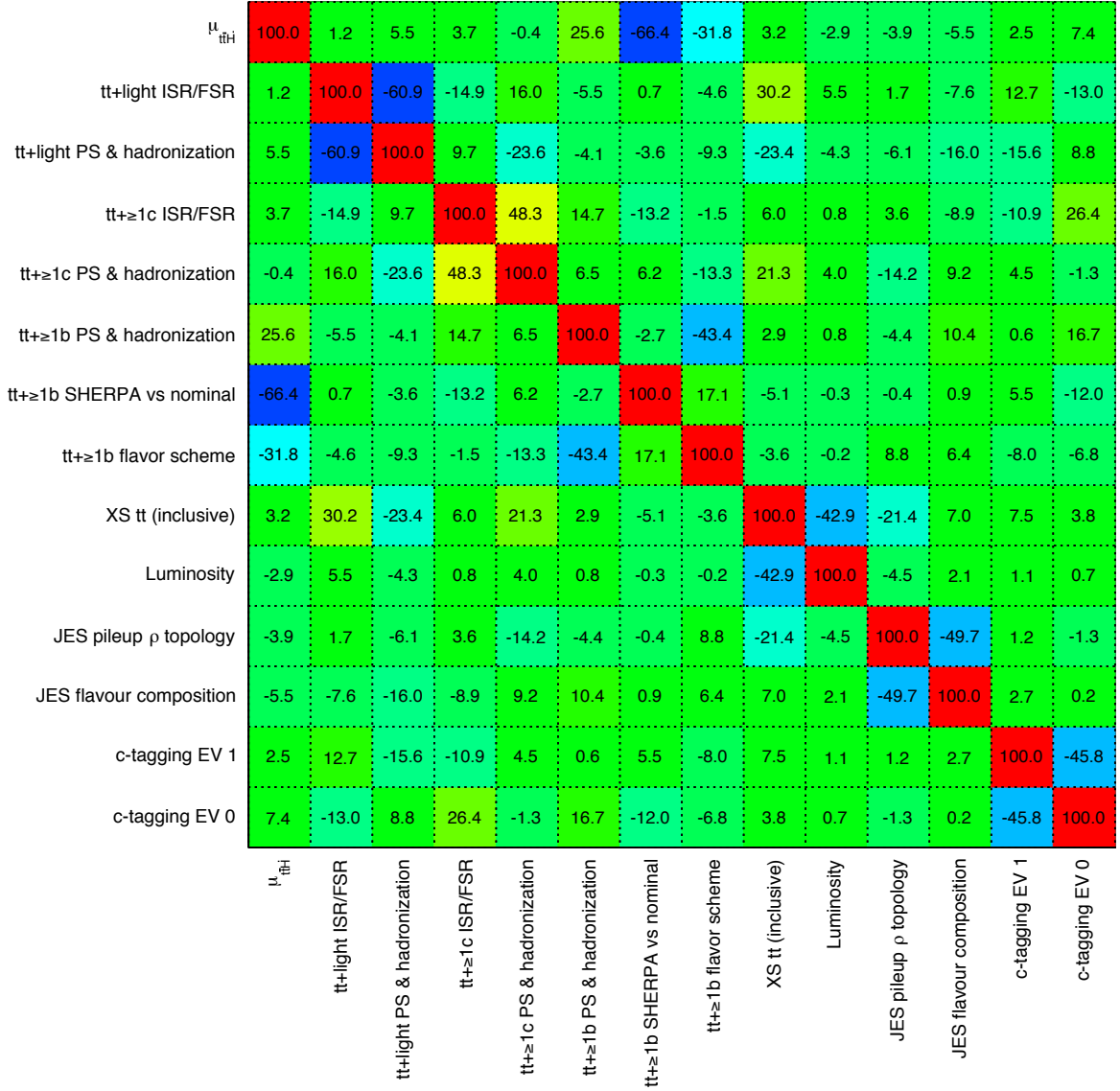


FIGURE 12.3: Linear correlation matrix between two nuisance parameters. Each bin has a correlation value [%].

The $\mu_{t\bar{t}H}$ values obtained by all fits are shown in Figure 12.4. The fitted signal strength $\mu_{t\bar{t}H}$ obtained from the combined fit is:

$$0.85^{+0.64}_{-0.61} = 0.85^{+0.29}_{-0.29}(\text{stat})^{+0.57}_{-0.53}(\text{syst}). \quad (12.8)$$

No significant excess is observed in data compared to the no $t\bar{t}H$ signal hypothesis ($\mu_{t\bar{t}H} = 0$). The observed $t\bar{t}H$ signal significance corresponds 1.4 standard deviations (1.4σ) with an expectation of 1.6σ . A signal strength larger than 2.0 is excluded at the 95% confidence level.

The contributions to the uncertainties were grouped into fourteen sources, and summarized in Table 12.2. Figure 12.5 shows the top ten systematic uncertainties which have largest impacts on the $\mu_{t\bar{t}H}$ after the fit. Comparing this with the result in Run 1, $t\bar{t}+\geq 1c$ modeling uncertainties have less impacts on the $\mu_{t\bar{t}H}$ in this result because of optimized selection of the analysis regions. However, $t\bar{t}+\geq 1b$ modeling systematic uncertainty is still largest in this analysis although many

TABLE 12.1: Fitted normalization factors for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$.

	dilepton	single-lepton	two- μ combined	combined
Expected values				
$k_{t\bar{t}+\geq 1b}$	$1.00^{+0.13}_{-0.12}$	$1.00^{+0.16}_{-0.16}$	$1.00^{+0.09}_{-0.09}$	$1.00^{+0.09}_{-0.08}$
$k_{t\bar{t}+\geq 1c}$	$1.00^{+0.37}_{-0.32}$	$1.00^{+0.52}_{-0.42}$	$1.00^{+0.20}_{-0.20}$	$1.00^{+0.20}_{-0.20}$
Observed values				
$k_{t\bar{t}+\geq 1b}$	$1.25^{+0.19}_{-0.19}$	$1.19^{+0.14}_{-0.13}$	$1.27^{+0.11}_{-0.11}$	$1.23^{+0.10}_{-0.10}$
$k_{t\bar{t}+\geq 1c}$	$1.80^{+0.65}_{-0.52}$	$1.27^{+0.41}_{-0.34}$	$1.59^{+0.23}_{-0.22}$	$1.61^{+0.23}_{-0.22}$

variables were implemented to separate signal events from $t\bar{t}+\geq 1b$ background events.

TABLE 12.2: Summary of the contributions to the signal strength uncertainty. The contribution sources were categorized by their origins. The "Background statistics" is the statistical uncertainties on the simulated samples used as the background physics processes. This also includes the statistical uncertainties on the data-driven fake lepton process. The $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalization is determined by the data, therefore, their uncertainties are categorized in the statistical uncertainties, which will reduce by the upcoming increased data. "Intrinsic statistical unc." equals to the data statistical uncertainty. The squared sum of the $\Delta\mu$ of all categories is not same as the total uncertainty because of correlations among categories.

Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	+0.46	-0.45
Background statistics	+0.35	-0.31
Flavor tagging	+0.16	-0.16
Jet energy scale and resolution	+0.12	-0.12
$t\bar{t}H$ modeling	+0.24	-0.06
$t\bar{t}+\geq 1c$ modeling	+0.09	-0.11
Pileup modeling	+0.01	-0.04
Other background modeling	+0.05	-0.05
$t\bar{t}+\text{light}$ modeling	+0.07	-0.05
Luminosity	+0.01	-0.02
Lepton and E_T^{miss} related unc.	+0.03	-0.04
Total systematic uncertainty	+ 0.57	-0.53
$t\bar{t}+\geq 1b$ normalization	+0.08	-0.08
$t\bar{t}+\geq 1c$ normalization	+0.03	-0.05
Intrinsic statistical unc.	+0.20	-0.20
Total statistical uncertainty	+0.29	-0.29
Total uncertainty	+0.64	-0.61

Distributions and number of events before and after the combined fit (denoted as pre-fit and post-fit, respectively) are shown in Figures 12.6 to 12.18.

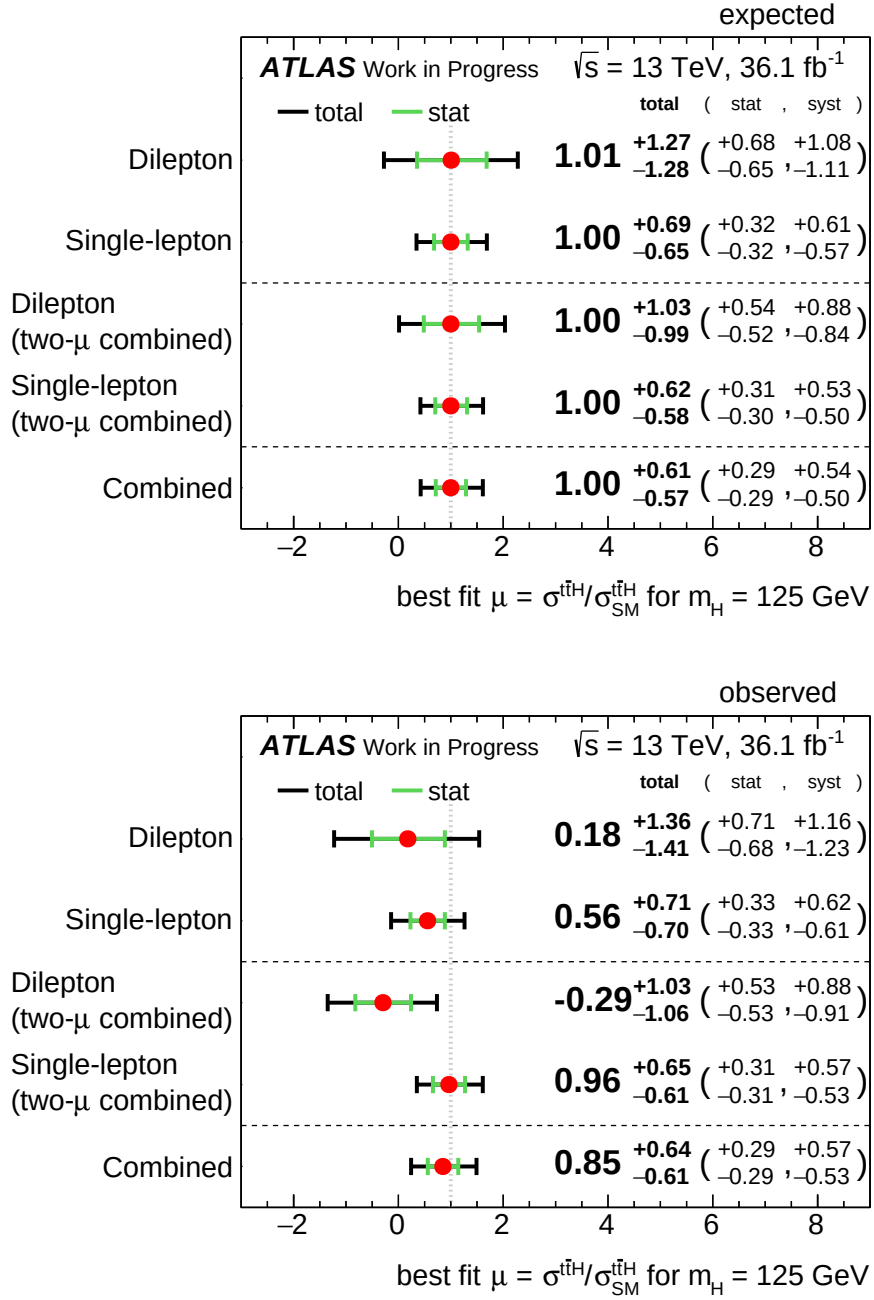


FIGURE 12.4: Summary of the (top) expected and (bottom) observed best-fit values of $\mu_{t\bar{t}H}$. In each figure, top two rows show the results in individual channels in individual fits, the next two rows show the results in individual channels in the two- μ combined fit, the last row shows the combined result.

Figure 12.19 concludes the present search, showing all histogram bins in all analysis region for both lepton channels sorted by $\log(S/B)$ after fitting.

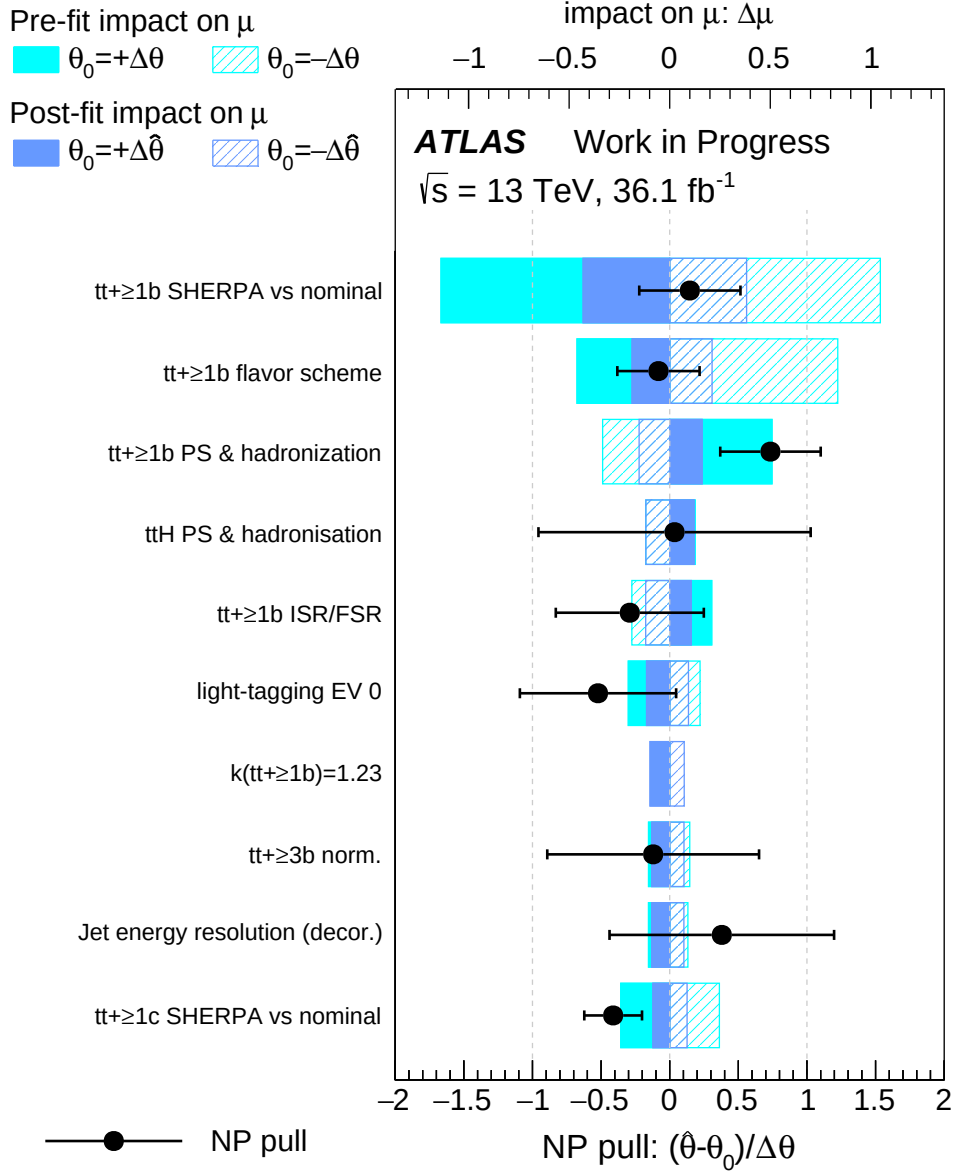


FIGURE 12.5: Ranking of the nuisance parameters included in the fit according to the impact on the measured signal strength. The top 10 parameters are shown. The cyan rectangles correspond to the pre-fit and the blue ones to the post-fit impact on $\mu_{t\bar{t}H}$, both referring to the upper scale. The impact of each NP, $\Delta\mu_{t\bar{t}H}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered NP to its best-fit value ($\hat{\theta}$) shifted its pre-fit and post-fit uncertainties by $\pm\Delta\theta$. The black points show the pulls of the nuisance parameters with respect to their nominal values (θ_0). These pulls and their relative post-fit errors ($\hat{\theta}$) refer to the lower scale. The pull of $k_{t\bar{t}+\geq 1b}$ is not plotted.

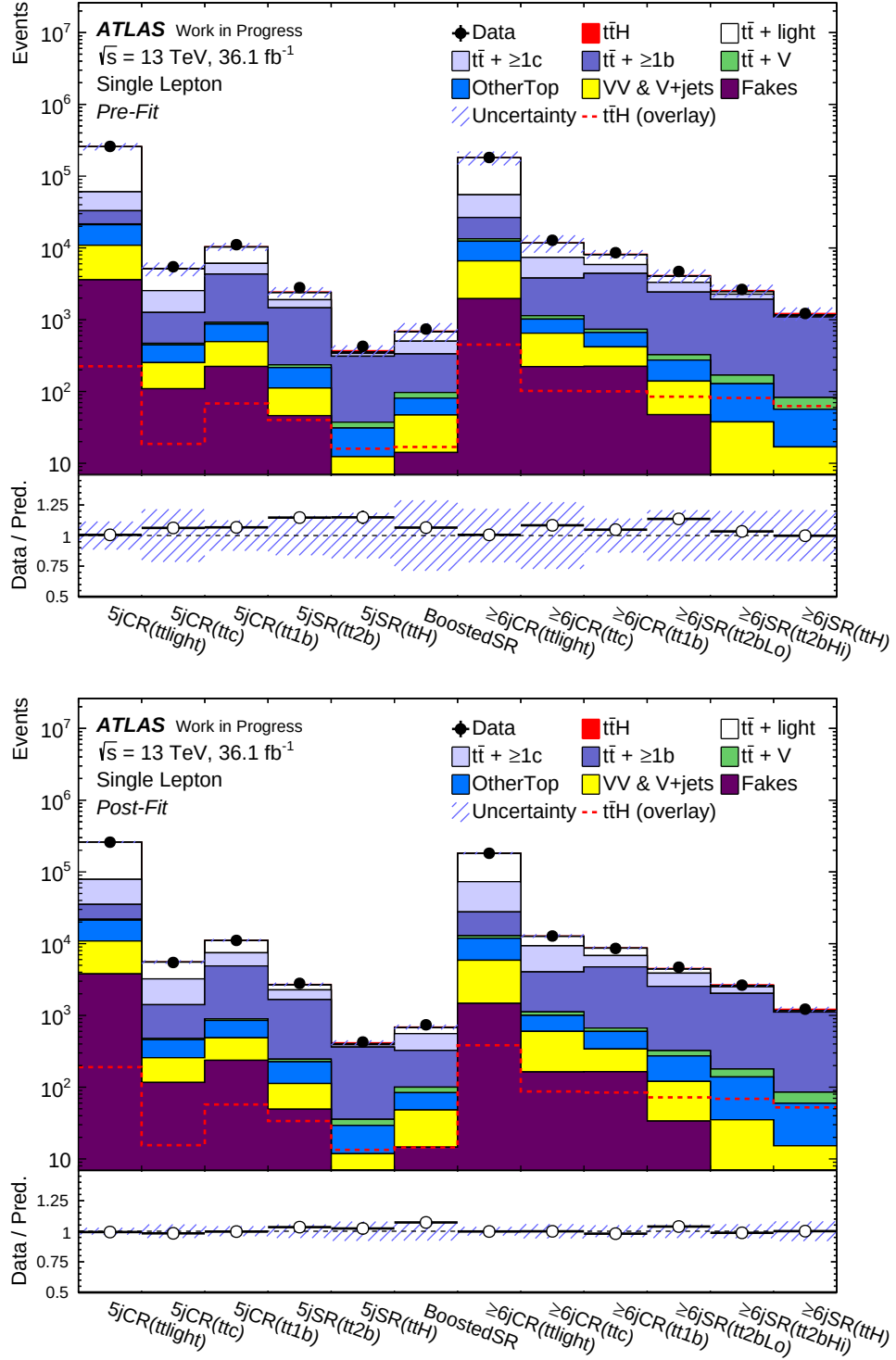


FIGURE 12.6: Analysis regions in the single-lepton channel. Only number of events are used in the fit in $\geq 6jCR(t\bar{t}+1b)$, $5jCR(t\bar{t}+1b)$, $\geq 6jCR(t\bar{t}+light)$ and $5jCR(t\bar{t}+light)$, while in $\geq 6jCR(t\bar{t}+\geq 1c)$, $5jCR(t\bar{t}+\geq 1c)$ and signal regions, distributions are used to constrain the systematic uncertainties. In the pre-fit plot, the number of $t\bar{t}H$ signal is assumed to the SM prediction. In the ratio plots, the data divided by the predictions including the signal events are shown. Shaded area denotes the total uncertainties including all systematic uncertainties and MC statistic uncertainties.

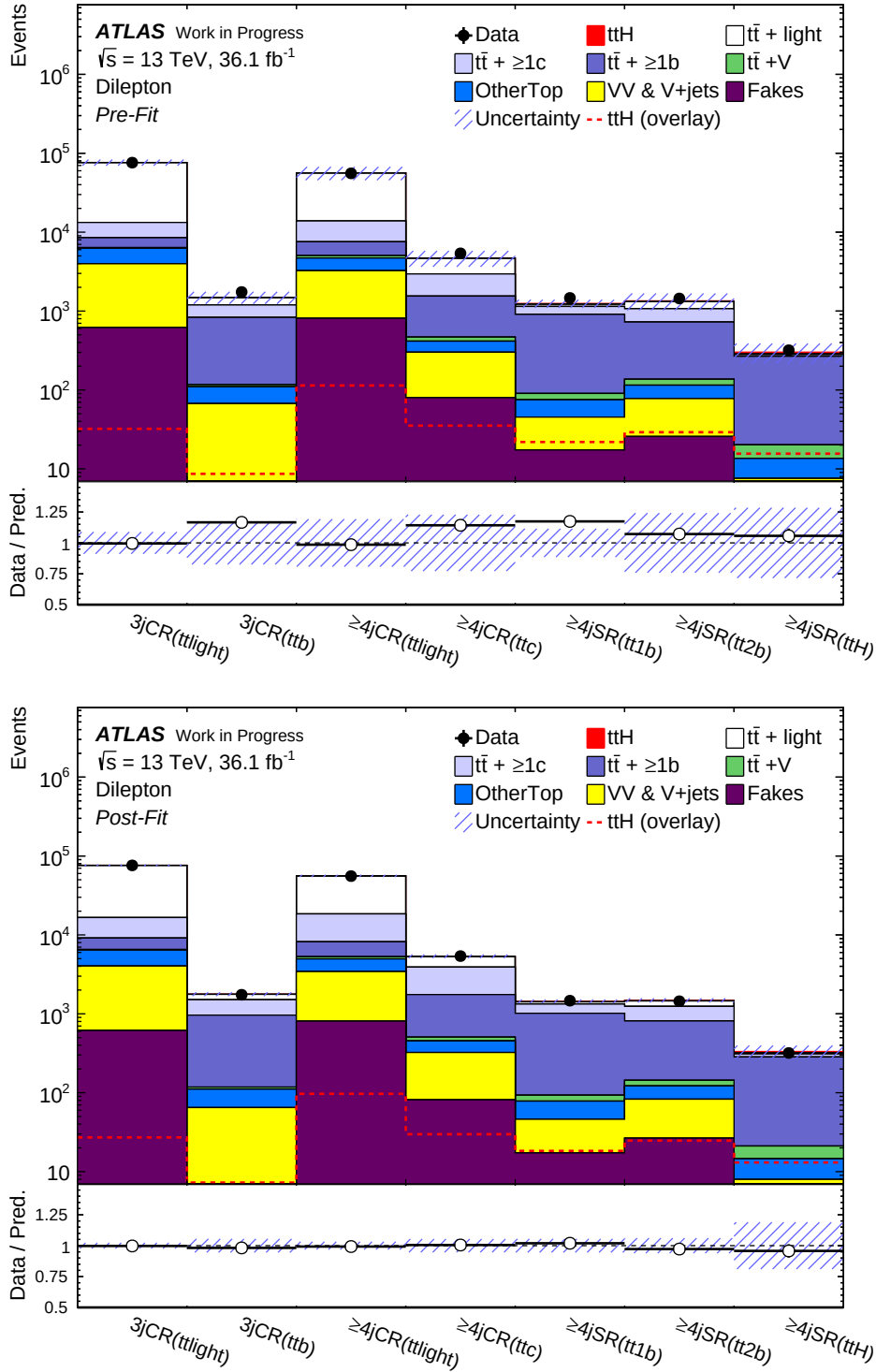
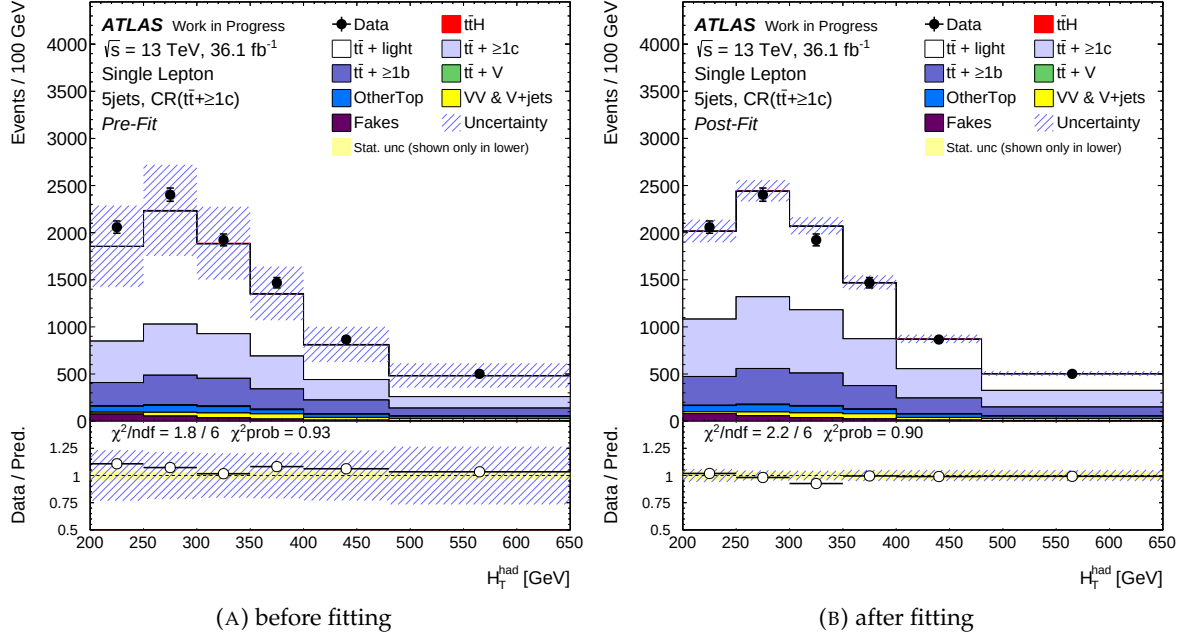
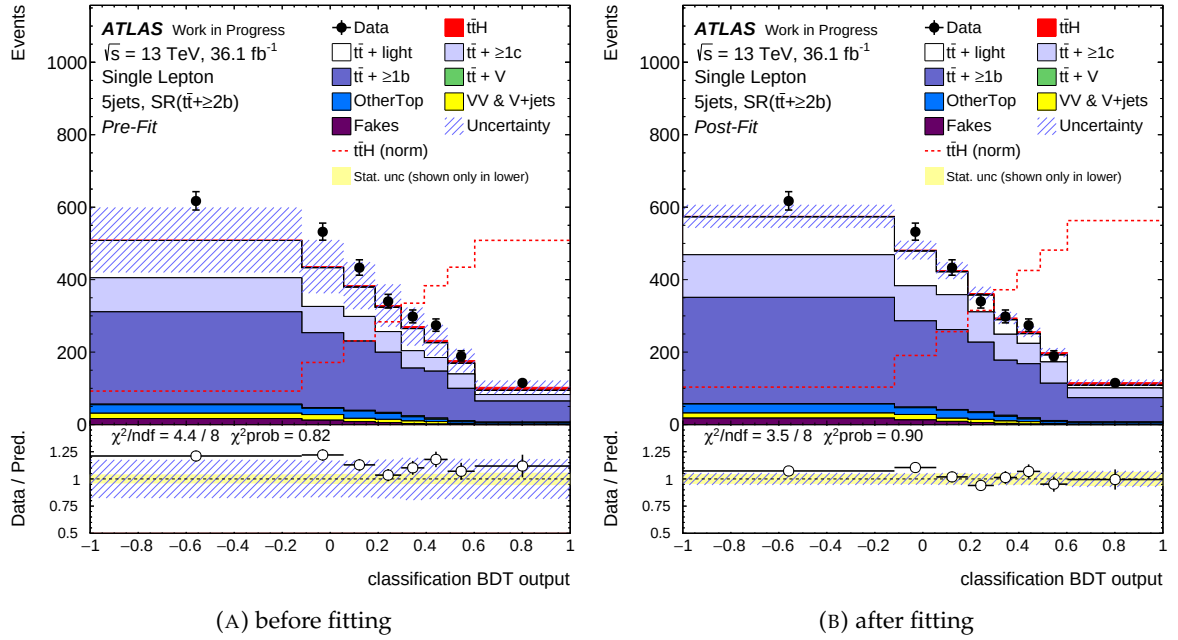
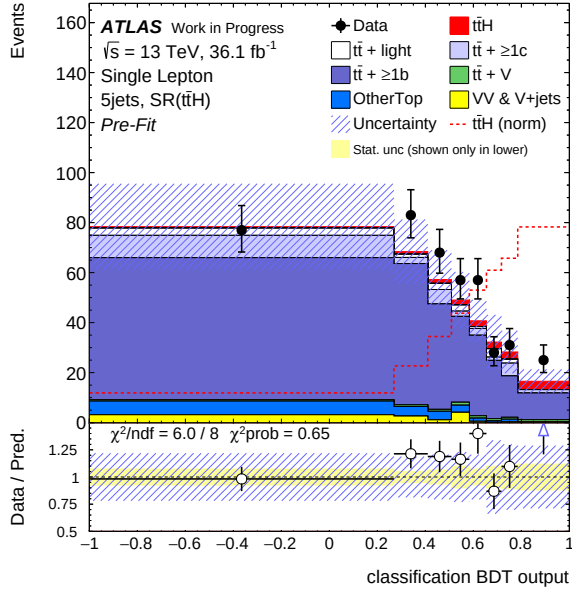
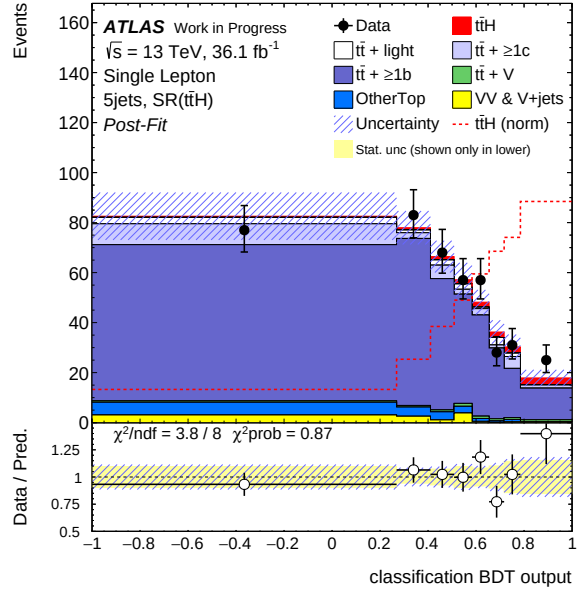


FIGURE 12.7: Analysis regions in the dilepton channel. Only number of events are used in the fit in $\geq 4j\text{CR}(t\bar{t}+\geq 1c)$, $\geq 4j\text{CR}(t\bar{t}+\text{light})$, $3j\text{CR}(t\bar{t}+\geq 1b)$ and $3j\text{CR}(t\bar{t}+\text{light})$, while in signal regions, distributions are used to constrain the systematic uncertainties. In the pre-fit plot, the number of $t\bar{t}H$ signal is assumed to the SM prediction. In the ratio plots, the data divided by the predictions including the signal events are shown. Shaded area denotes the total uncertainties including all systematic uncertainties and MC statistic uncertainties.

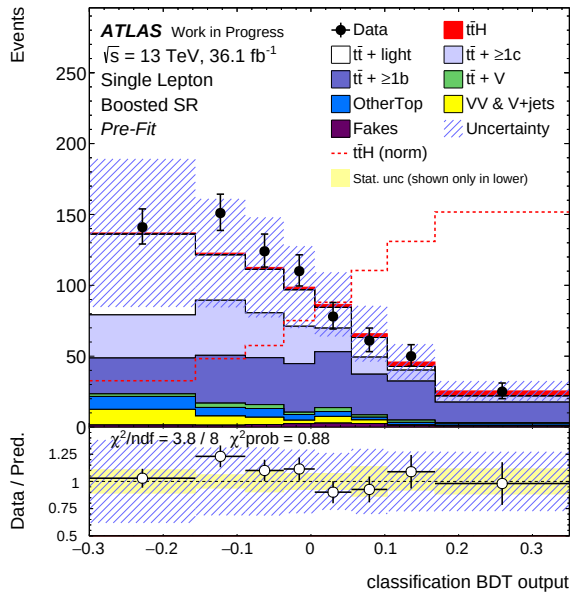
FIGURE 12.8: H_T^{had} distributions in 5jCR($t\bar{t} \geq 1c$)FIGURE 12.9: ClassBDT distributions in 5jSR($t\bar{t} \geq 2b$)



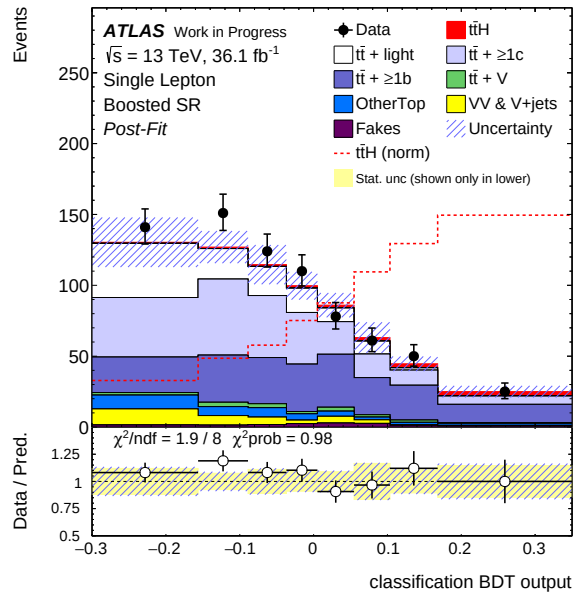
(A) before fitting



(B) after fitting

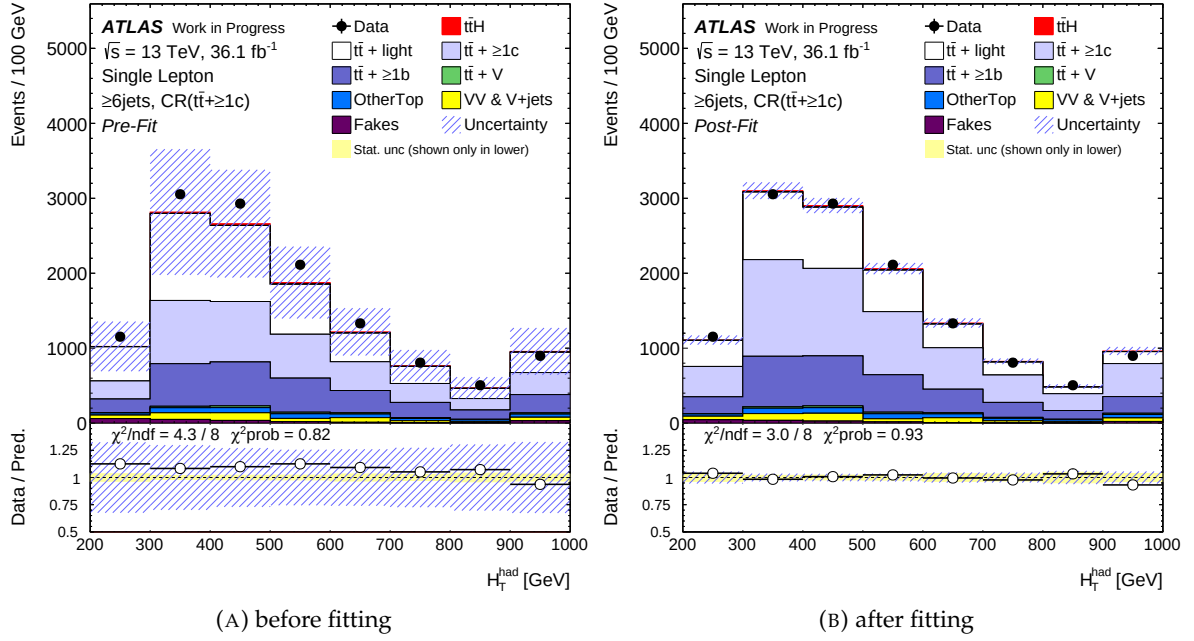
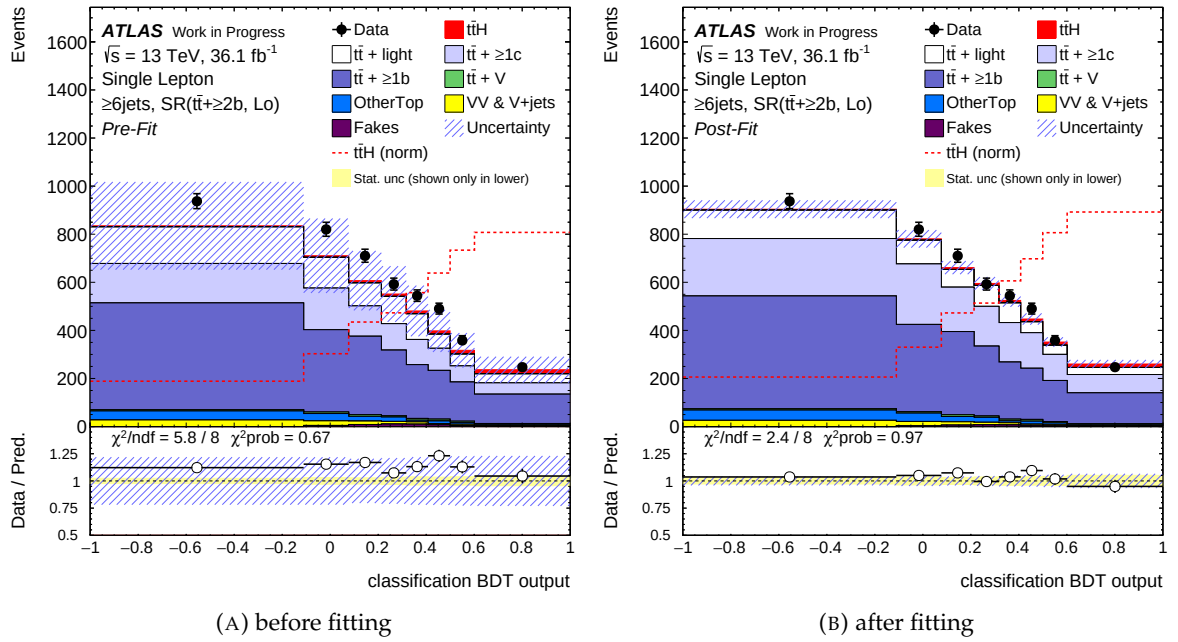
FIGURE 12.10: H_T^{had} distributions in 5jSR($t\bar{t}H$)

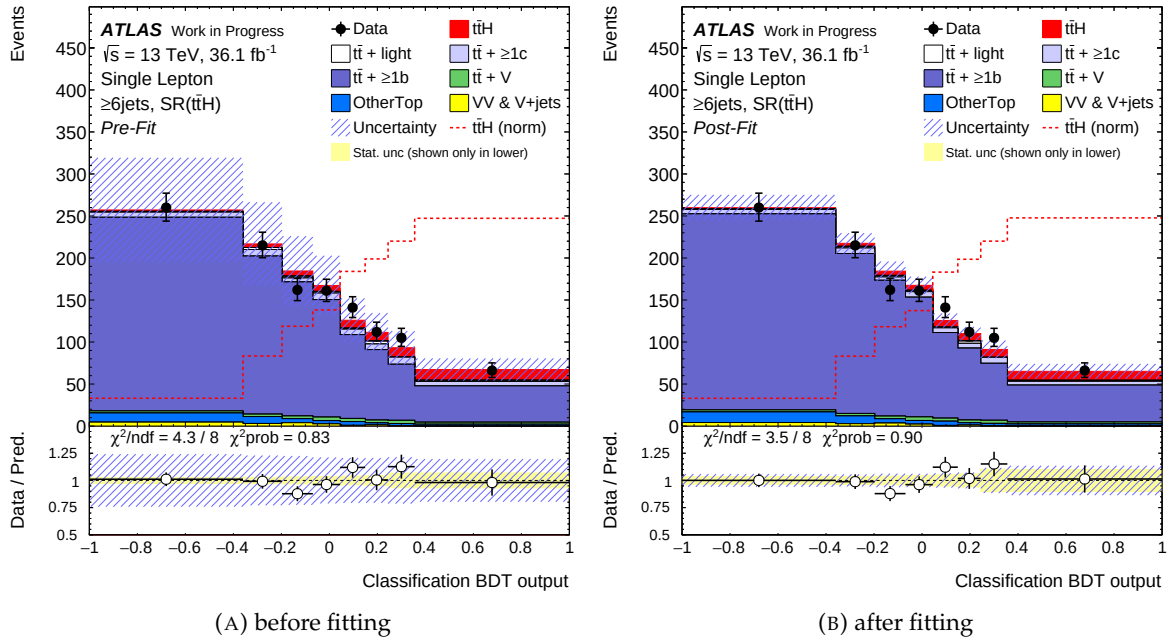
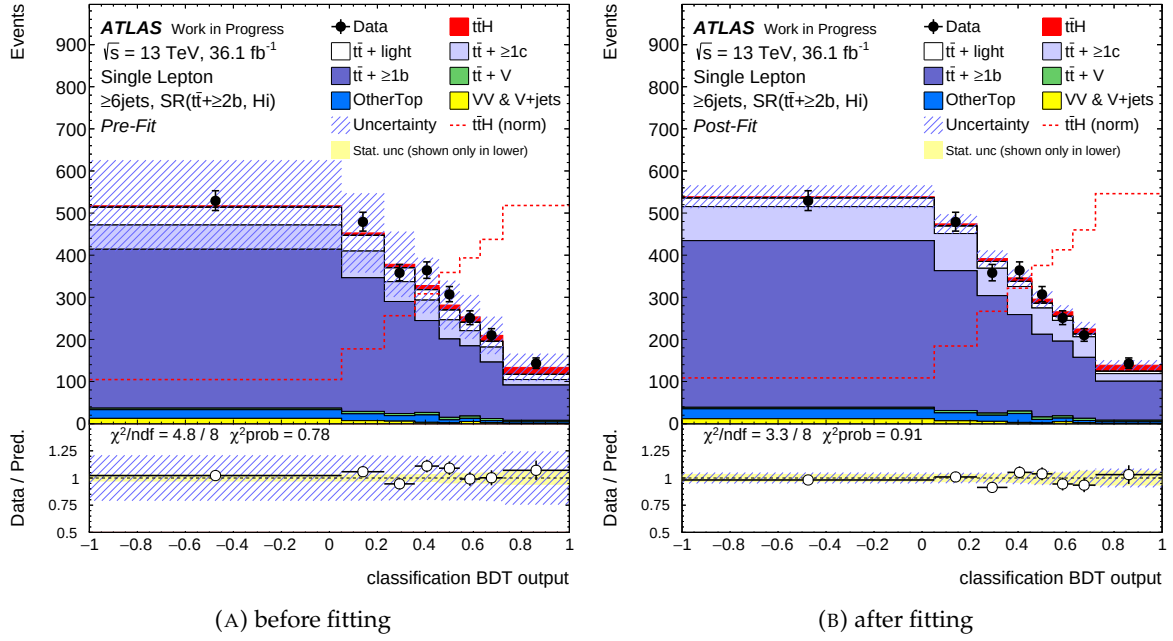
(A) before fitting

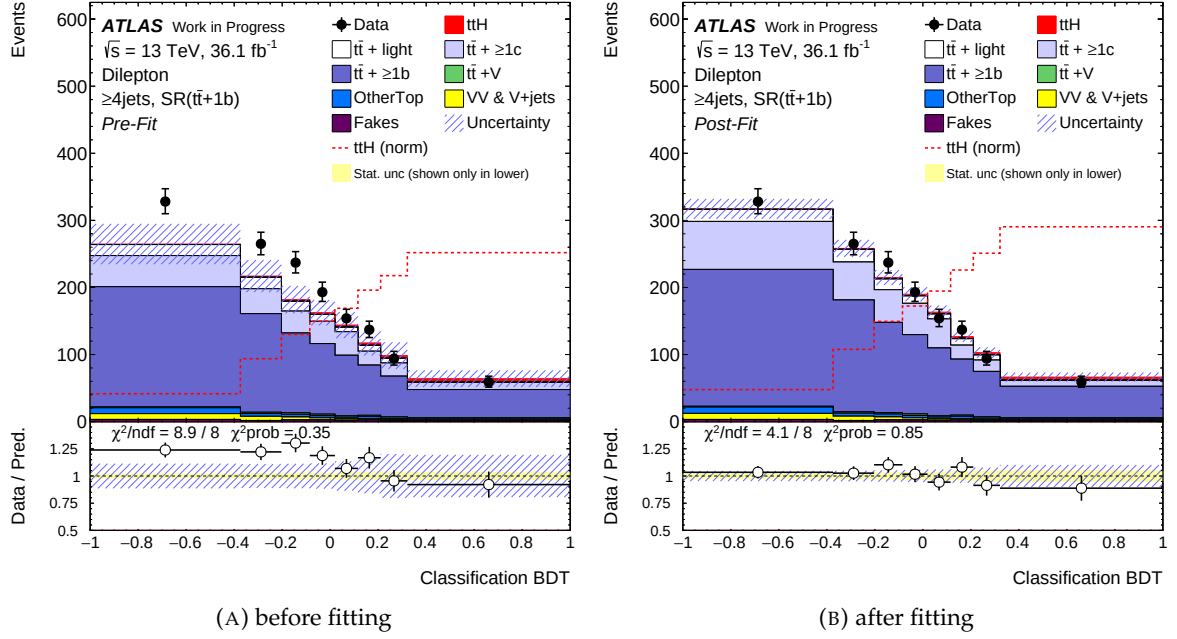
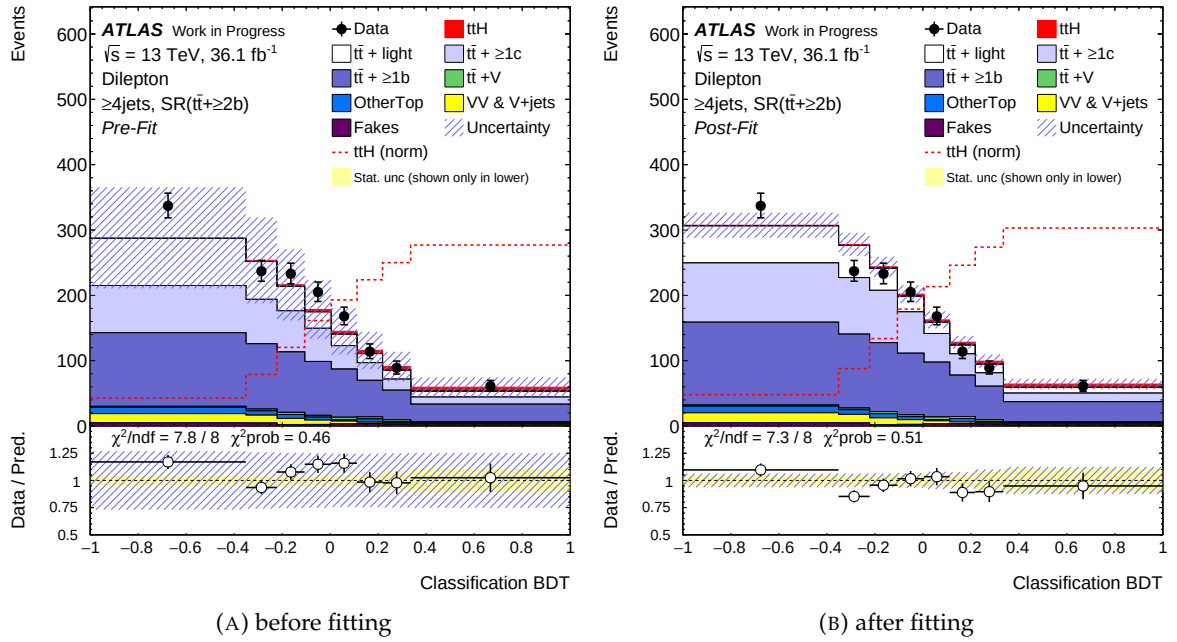


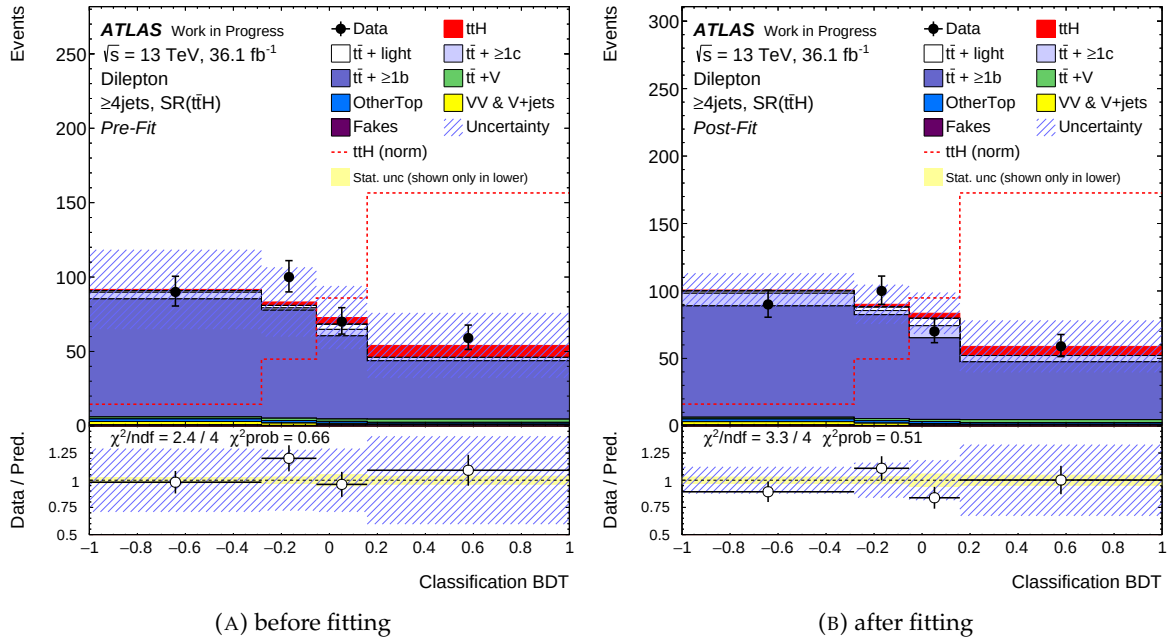
(B) after fitting

FIGURE 12.11: H_T^{had} distributions in the boosted region

FIGURE 12.12: H_T^{had} distributions in $\geq 6\text{jCR}(t\bar{t}+\geq 1c)$ FIGURE 12.13: ClassBDT distributions in $\geq 6\text{jSR}(t\bar{t}+\geq 2b, \text{Lo})$



FIGURE 12.16: ClassBDT distributions in $\geq 4jSR(t\bar{t}+1b)$ FIGURE 12.17: ClassBDT distributions in $\geq 4jSR(t\bar{t}+\geq 2b)$

FIGURE 12.18: ClassBDT distributions in $\geq 4jSR(t\bar{t}H)$

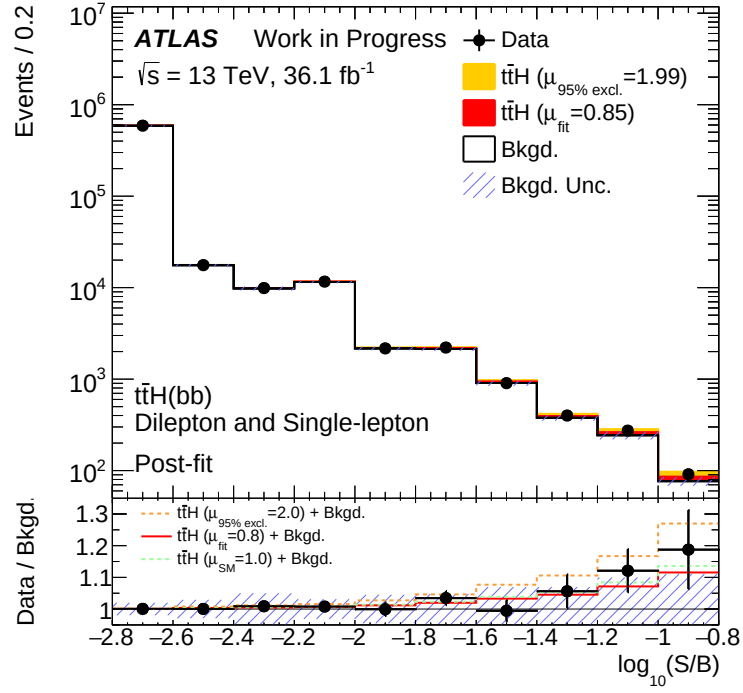


FIGURE 12.19: All analysis region distributions for both lepton channels sorted by $\log(S/B)$ after fitting. The background events are merged in one color (white). The several benchmarks of the signal strength are shown in red and yellow color rectangles stacked on the background events, where the red color shows the signal events at the best fit value, and yellow color shows the signal events at the value excluded at the 95% CL. The lower plot shows the relative pulls of the data to the predicted background events. The data pulls are also compared to the several benchmarks. Red and yellow denote the same color definition in the upper plot. Green line shows the signal events predicted by the SM. Shaded region shows the total systematic uncertainty other than the signal modeling. The first bin includes the underflow.

12.3 Fit Validations

12.3.1 Noticeable Distributions

To validate the fit, number of jets and b -tagged jets, H_T^{had} and H_T^{all} distributions were checked after fitting in Figures 12.20 to 12.26. The regions utilizing normalizations have slopes in the ratio between data and prediction, as discussed in Chapter 8, which were not compensated. The mis-modeling did not affect significantly to the global fit and to the result.

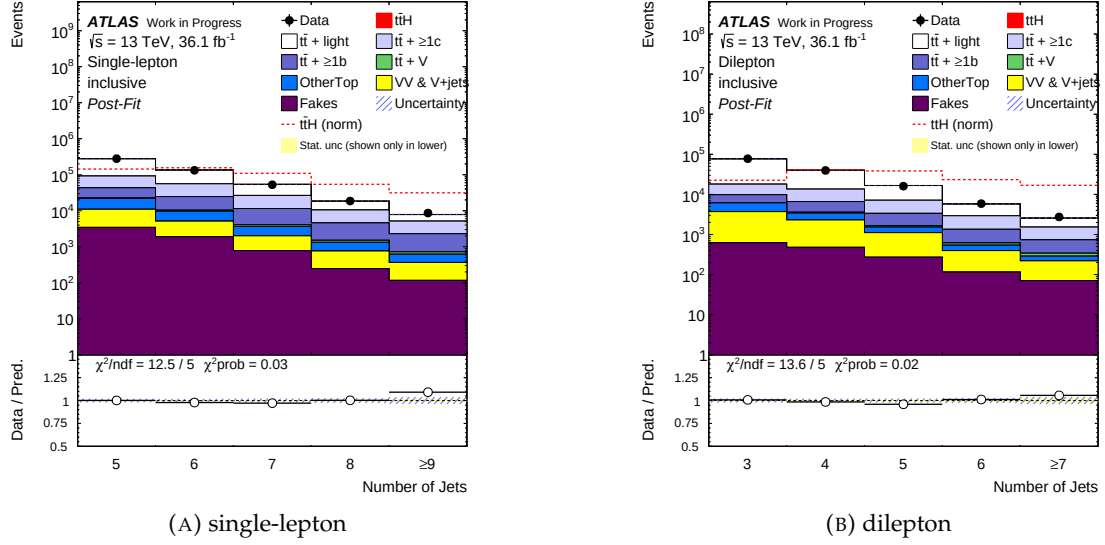
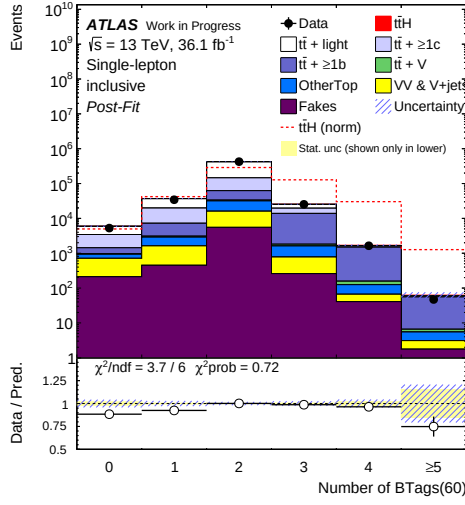
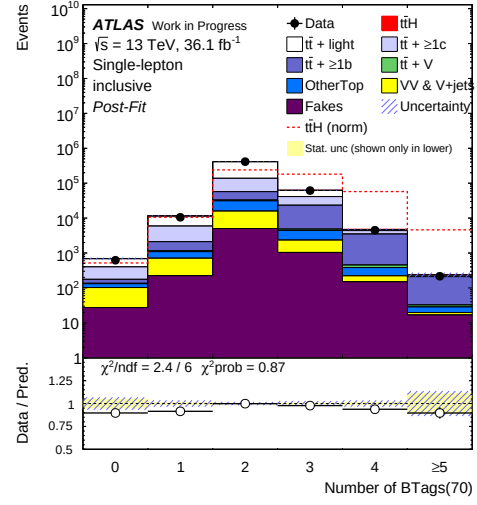
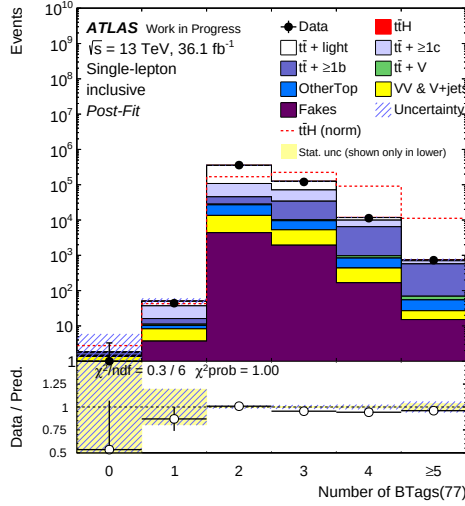
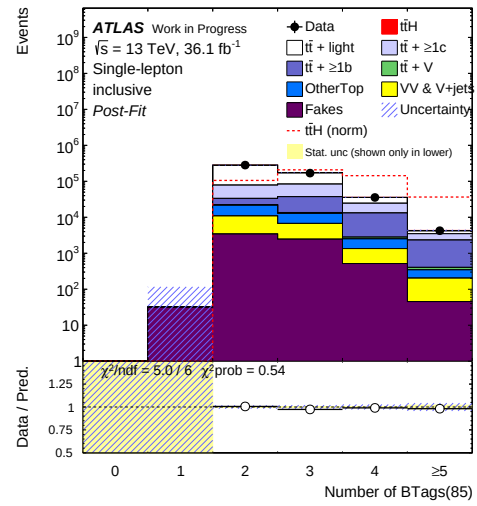
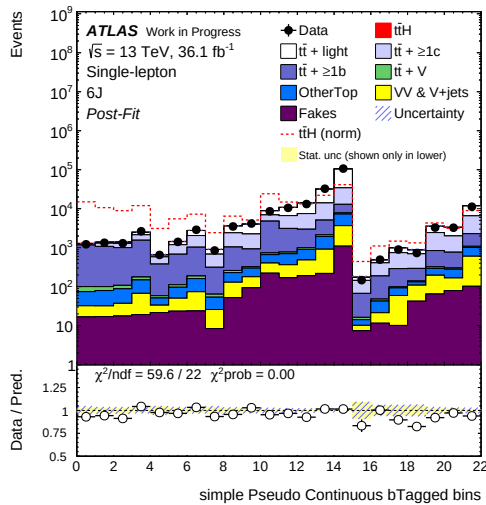
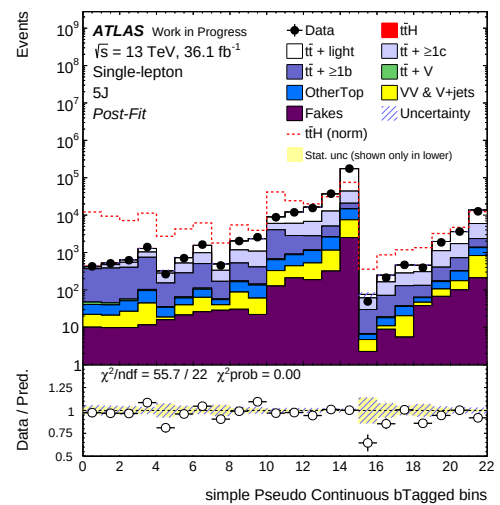
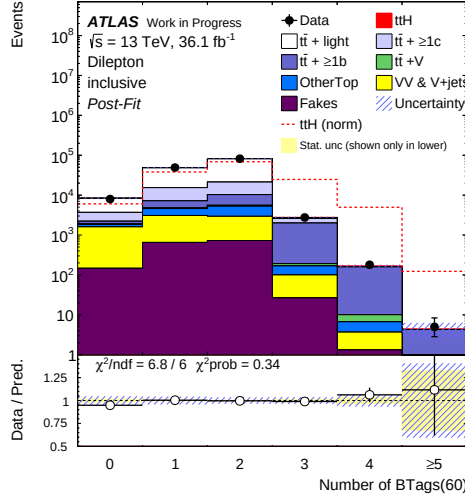
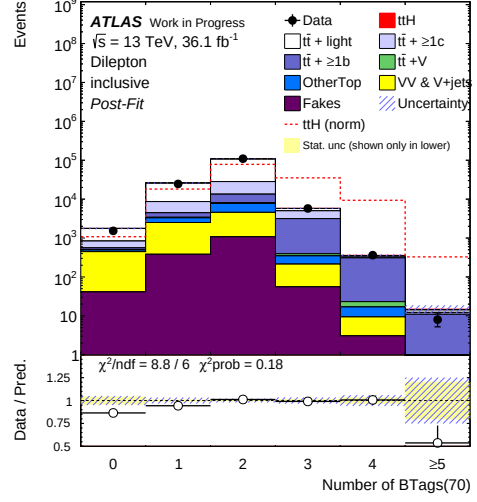
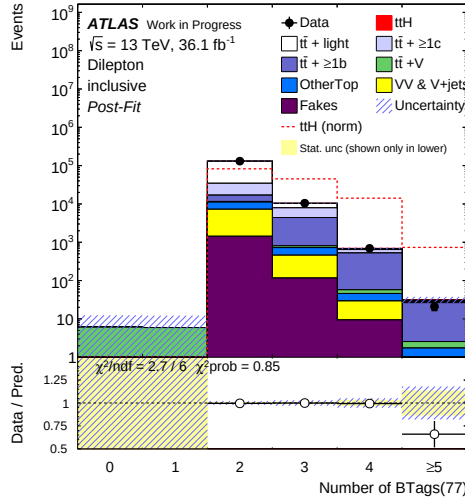
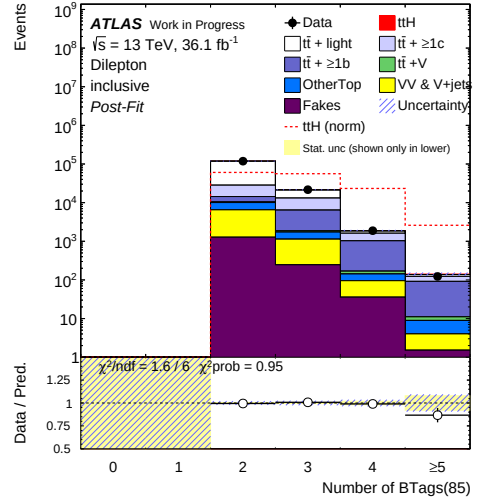
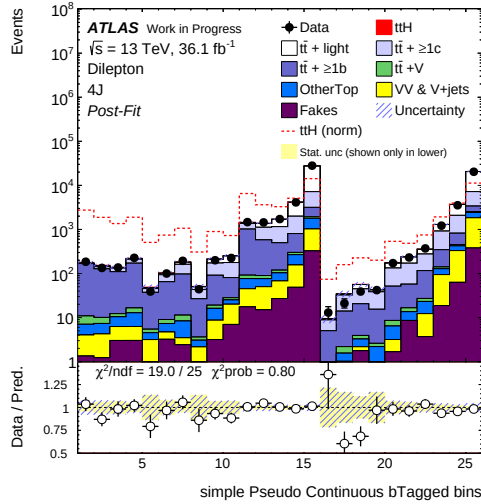
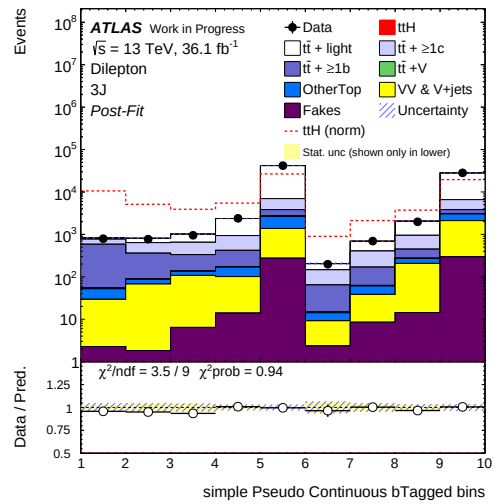


FIGURE 12.20: Number of jets (A) in the single-lepton channel and (B) in the dilepton channel after fitting.

(A) # b -tagged jets at 60% WP(B) # b -tagged jets at 70% WP(C) # b -tagged jets at 77% WP(D) # b -tagged jets at 85% WP(E) tag weight bins in the ≥ 6 j region

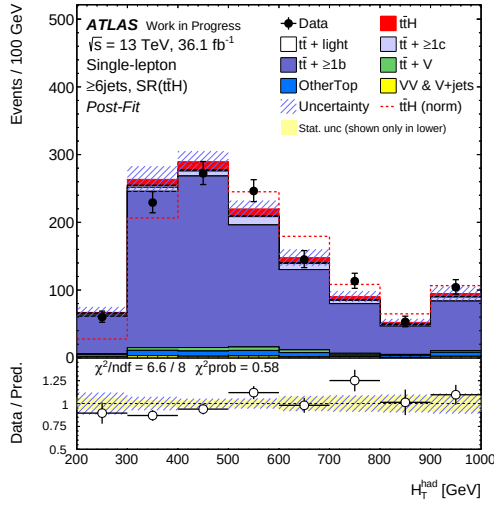
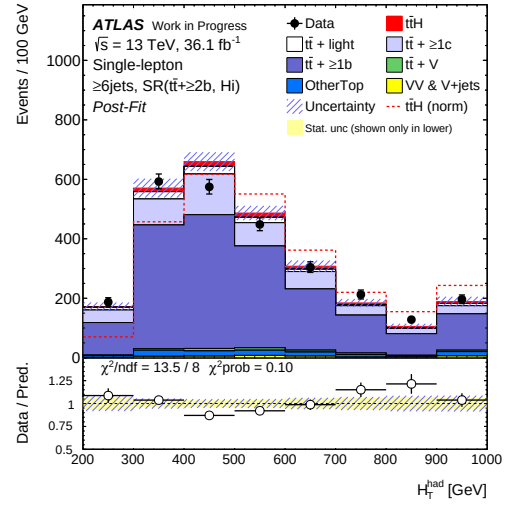
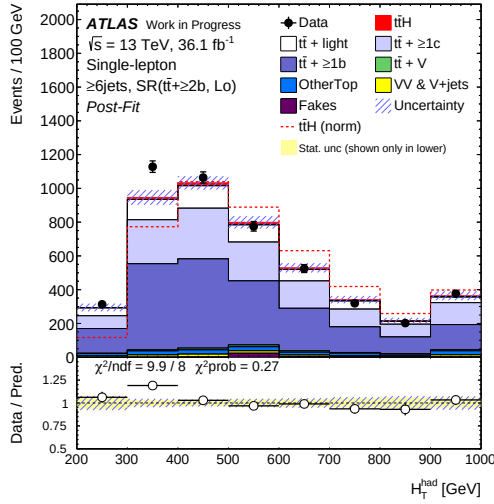
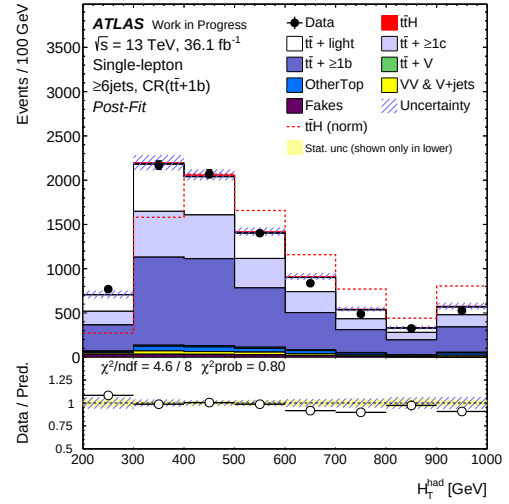
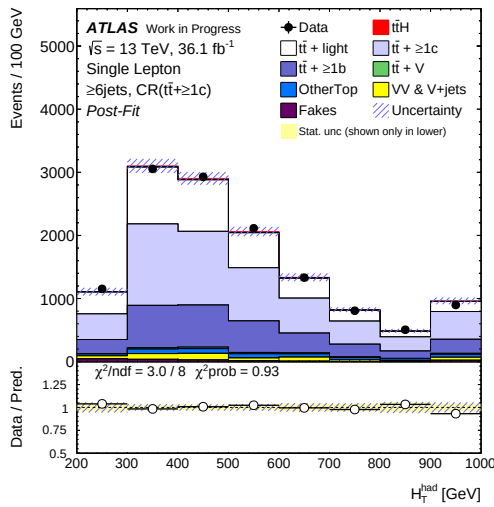
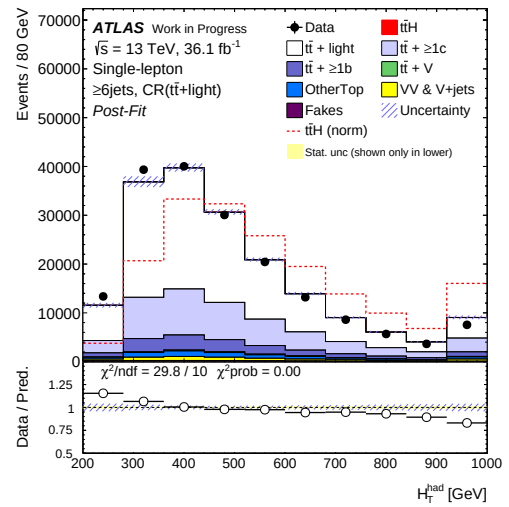
(F) tag weight bins in the 5 j region

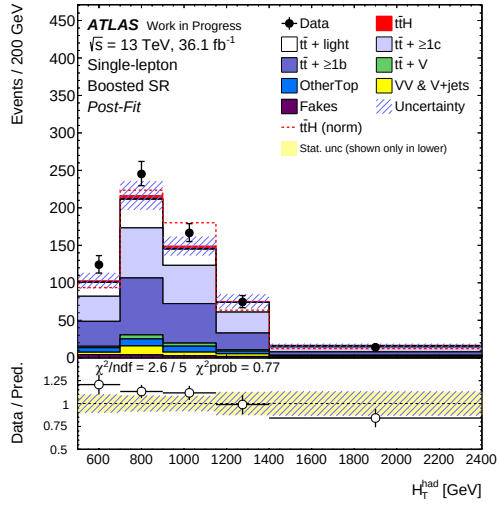
FIGURE 12.21: Number of b -tagged jets at each WP in the single-lepton channel after fitting. The numbering for the distributions of the tag weight bin in (E) and (F) is shown in Figure 8.25.

(A) # b -tagged jets at 60% WP(B) # b -tagged jets at 70% WP(C) # b -tagged jets at 77% WP(D) # b -tagged jets at 85% WP(E) tag weight bins in the ≥ 4 j region

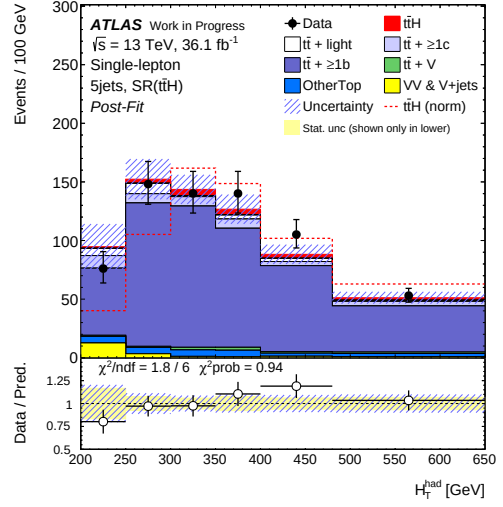
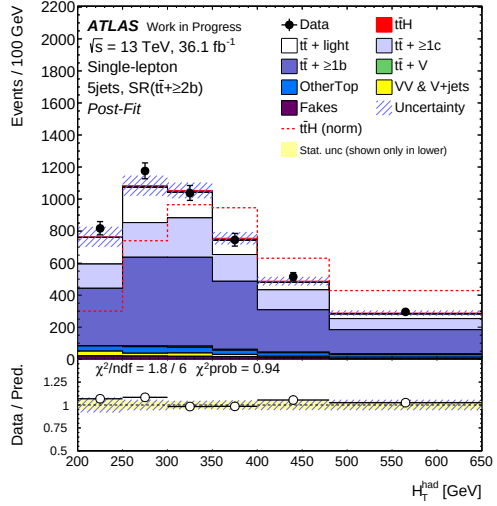
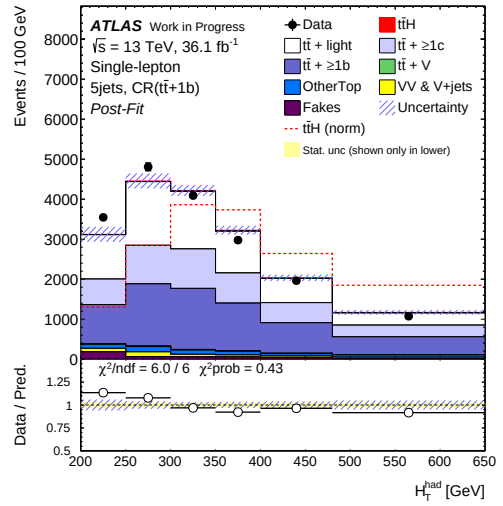
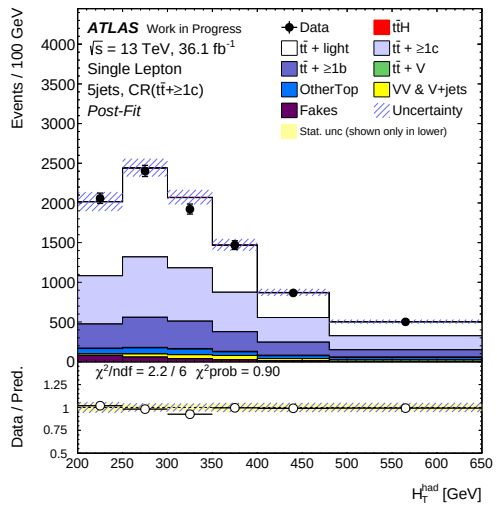
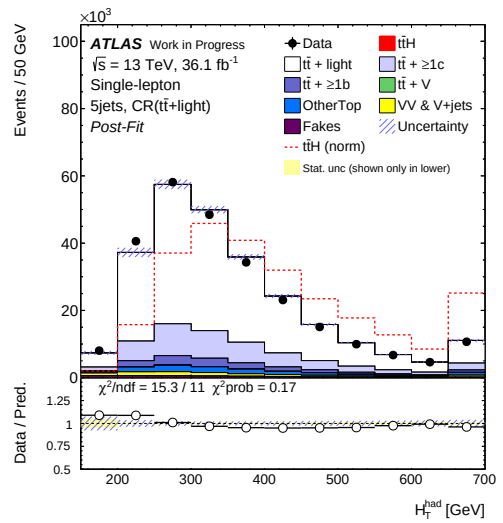
(F) tag weight bins in the 3 j region

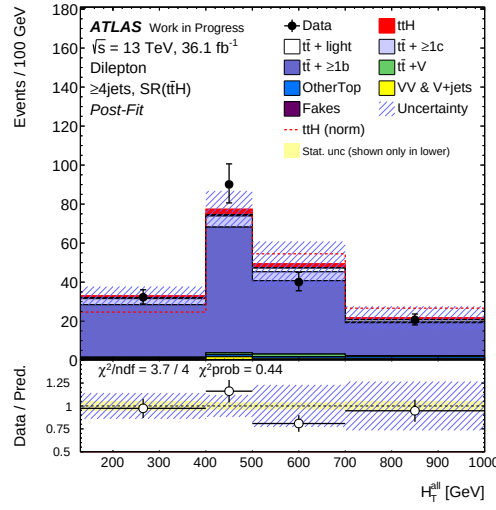
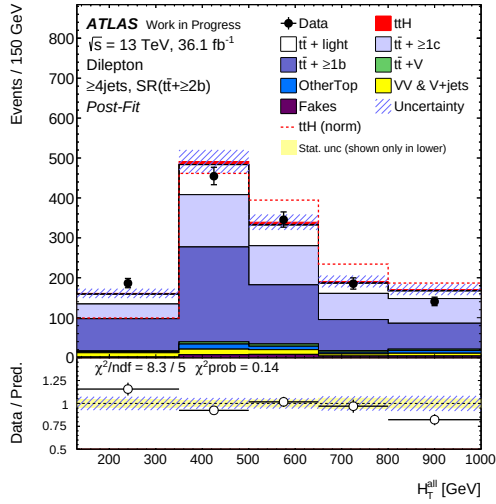
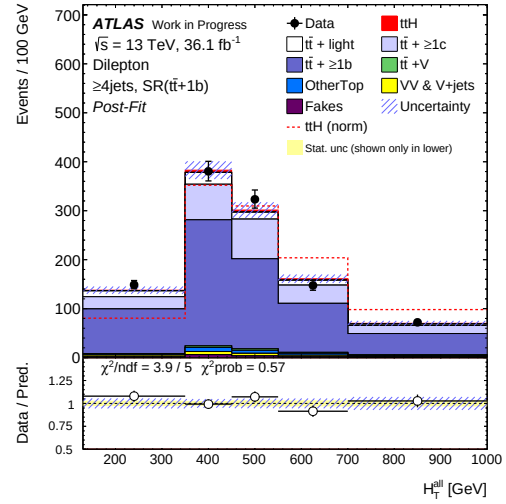
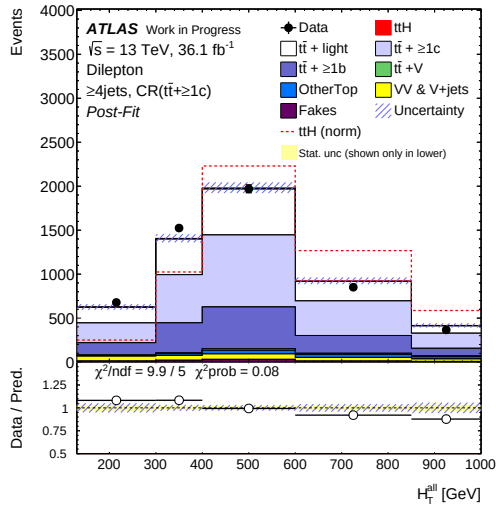
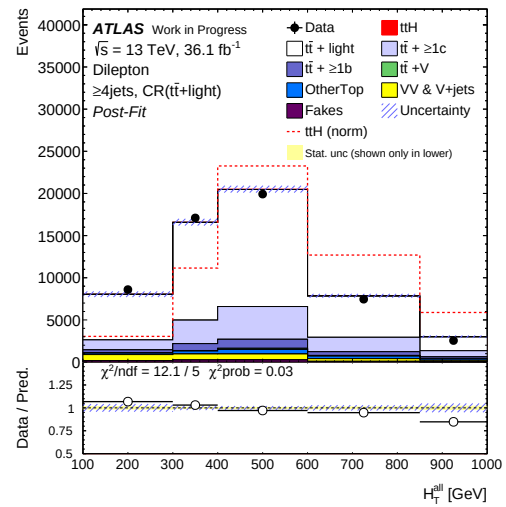
FIGURE 12.22: Number of b -tagged jets at each WP in the dilepton channel after fitting. The numbering for the distributions of the tag weight bin in (E) and (F) is shown in Figure 8.25.

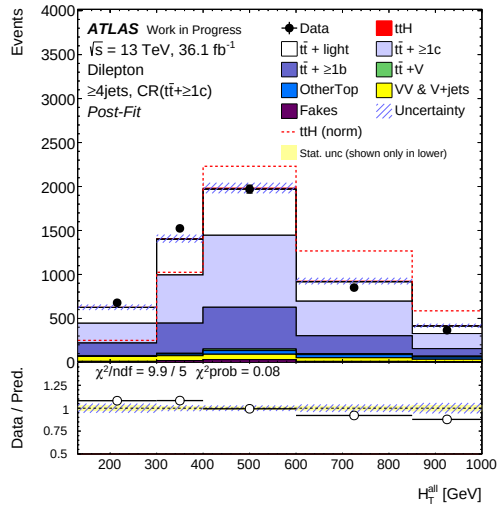
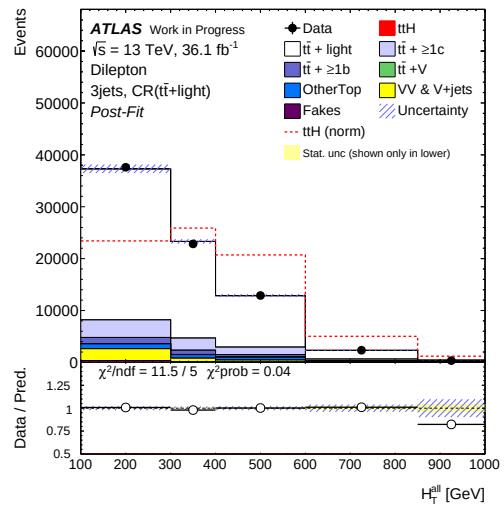
(A) $\geq 6\text{jetsSR}(t\bar{t}H)$ (B) $\geq 6\text{jetsSR}(t\bar{t} + \geq 2b, \text{Hi})$ (C) $\geq 6\text{jetsSR}(t\bar{t} + \geq 2b, \text{Lo})$ (D) $\geq 6\text{jetsCR}(t\bar{t} + 1b)$ (E) $\geq 6\text{jetsCR}(t\bar{t} + \geq 1c)$ (F) $\geq 6\text{jetsCR}(t\bar{t} + \text{light})$ FIGURE 12.23: H_T^{had} distributions in $\geq 6 \text{ j}$ regions for the single-lepton channel after fitting



(A) Boosted SR

(B) 5jSR($t\bar{t}H$)(C) 5jSR($t\bar{t}+\geq 2b$)(D) 5jSR($t\bar{t}+1b$)(E) 5jCR($t\bar{t}+\geq 1c$)(F) 5jCR($t\bar{t}+\text{light}$)FIGURE 12.24: H_T^{had} distributions in 5 j regions and boosted region for the single-lepton channel after fitting

(A) $\geq 4 \text{ jets SR}(t\bar{t}H)$ (B) $\geq 4 \text{ jets SR}(t\bar{t} + \geq 2b)$ (C) $\geq 4 \text{ jets SR}(t\bar{t} + 1b)$ (D) $\geq 4 \text{ jets CR}(t\bar{t} + \geq 1c)$ (E) $\geq 4 \text{ jets CR}(t\bar{t} + \text{light})$ FIGURE 12.25: H_T^{all} distributions in $\geq 4 \text{ j}$ regions for the dilepton channel after fitting

(A) $3j\text{CR}(t\bar{t} + \geq 1c)$ (B) $3j\text{CR}(t\bar{t} + \text{light})$ FIGURE 12.26: H_T^{all} distributions in 3 j regions for the dilepton channel after fitting

12.3.2 Higgs Mass Distributions

m_{bb}^{Higgs} distributions in the signal regions are shown in Figures 12.27 to 12.33. All data distributions are consistent to the predictions with and without $t\bar{t}H$ signal events.

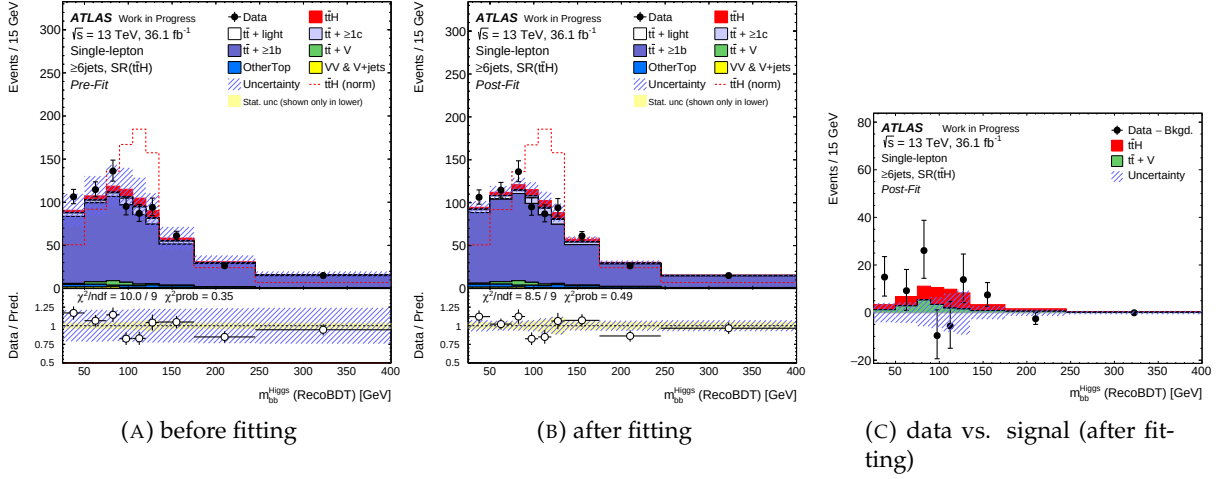


FIGURE 12.27: m_{bb}^{Higgs} (Reco BDT w/o H) distributions in $\geq 6j\text{SR}(t\bar{t}H)$. (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

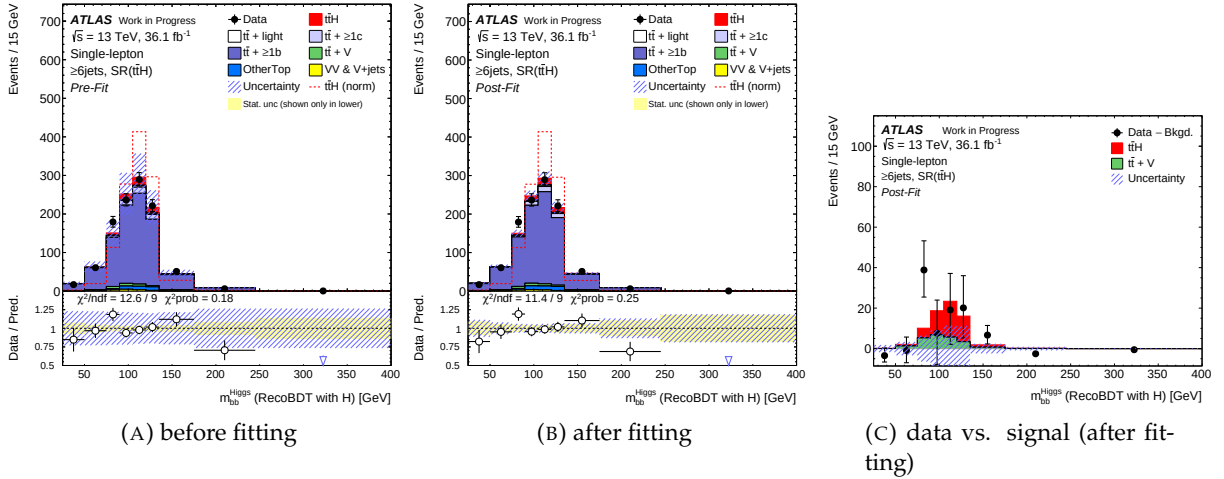


FIGURE 12.28: m_{bb}^{Higgs} (Reco BDT with H) distributions in $\geq 6j\text{SR}(t\bar{t}H)$. (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

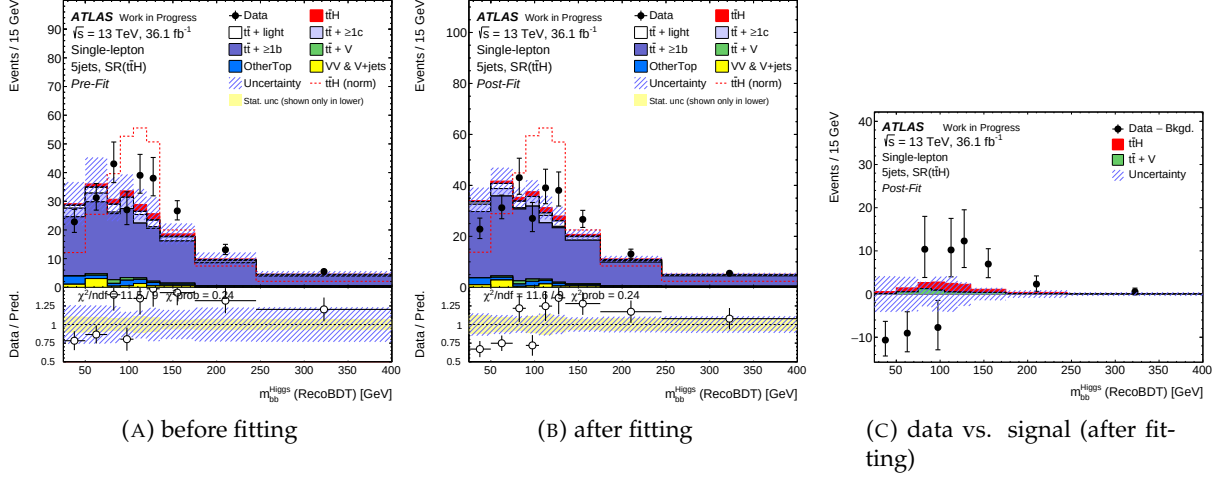


FIGURE 12.29: m_{bb}^{Higgs} (Reco BDT w/o H) distributions in 5jSR($t\bar{t}H$). (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

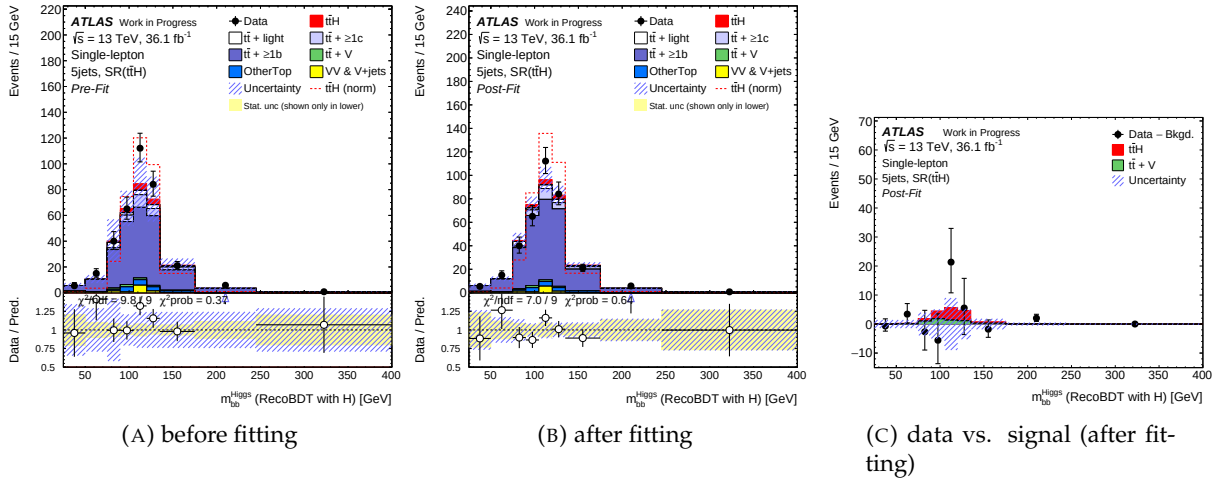


FIGURE 12.30: m_{bb}^{Higgs} (Reco BDT with H) distributions in 5jSR($t\bar{t}H$). (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

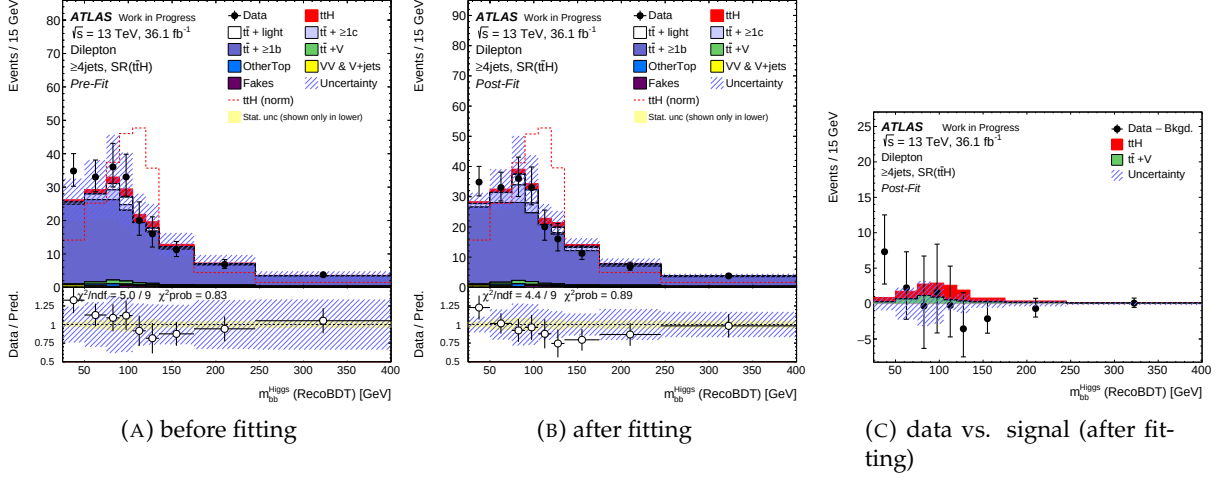


FIGURE 12.31: m_{bb}^{Higgs} (Reco BDT w/o H) distributions in $\geq 4j\text{SR}(t\bar{t}H)$. (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

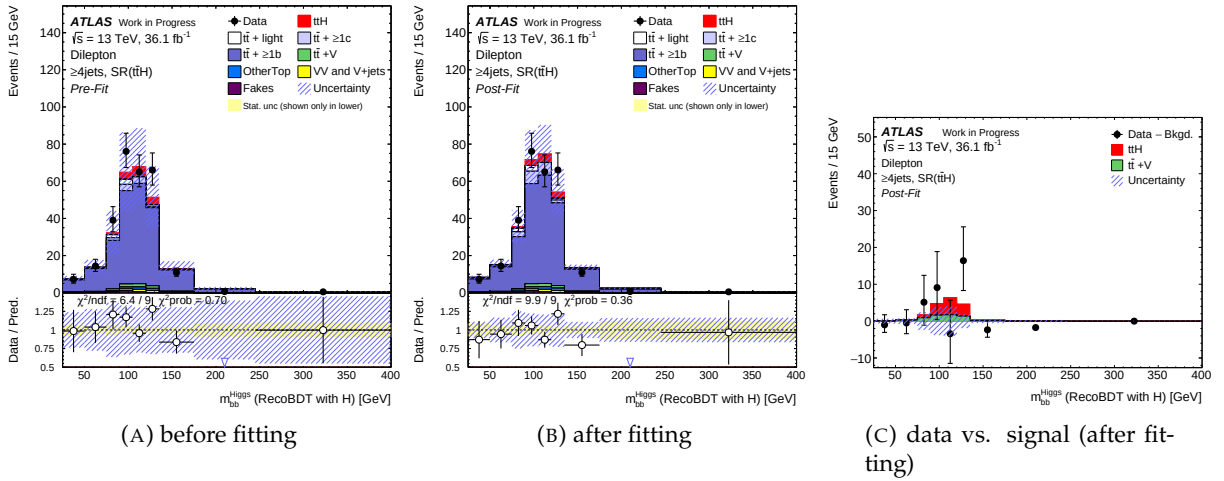


FIGURE 12.32: m_{bb}^{Higgs} (Reco BDT with H) distributions in $\geq 4j\text{SR}(t\bar{t}H)$. (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

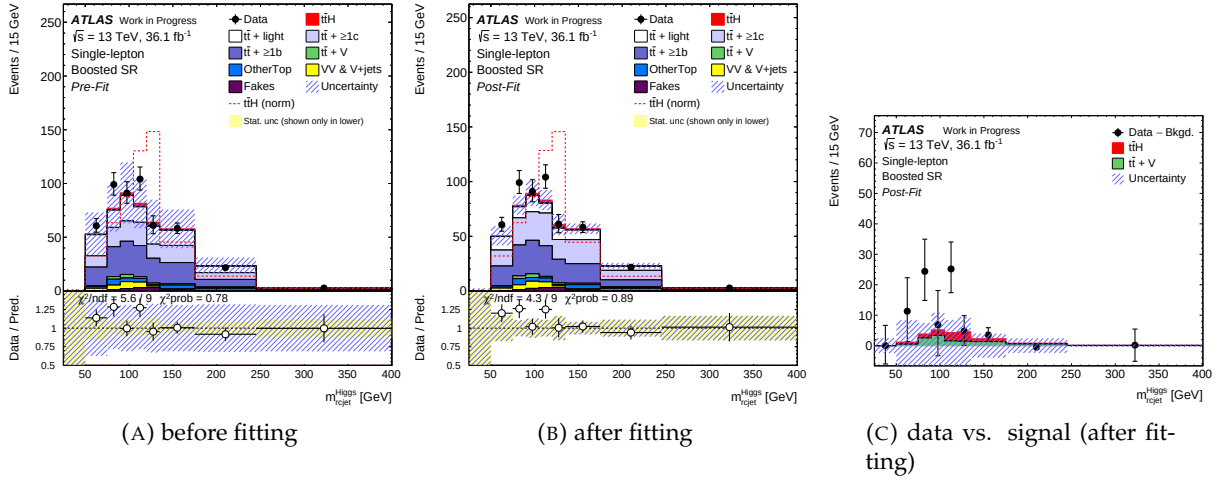


FIGURE 12.33: m_{rcjet}^{Higgs} distributions in the boosted SR. (A) Before fitting, the number of $t\bar{t}H$ signal is assumed to the SM prediction. (B) After the fitting, the determined signal strength (0.83), pulls and constrained systematic uncertainties were applied. The data and MC are in good agreement. (C) Comparison of the $t\bar{t}H$ plus $t\bar{t}V$ expectation and data where other background events are subtracted.

12.3.3 Fitting with Alternative Discriminants

As the fitting validation, alternative sets of fitting were checked using output score of the RecoBDT with H and LHD. In the control regions, the same input distributions were used, and the distributions in the signal regions were replaced as following:

- RecoBDT fitting: Output scores of the RecoBDT with H were utilized in all signal regions for both lepton channels.
- LHD fitting: LHD in all signal regions were utilized in the single-lepton channel, and the nominal ClassBDTs were utilized in signal regions in the dilepton channel.

The results are compared in Figure 12.34 and Table 12.3. The observed best-fit values of $\mu_{t\bar{t}H}$ is fluctuated by the different discriminant in the signal regions, however the three results are consistent to each other in $\pm 1\sigma$.

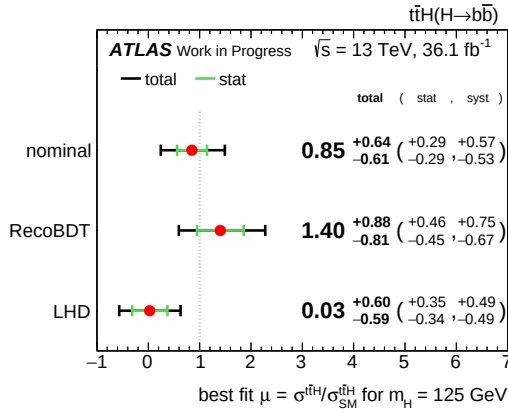


TABLE 12.3: Fitted normalization factors for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ with the alternative sets of distributions in the signal regions.

	nominal	RecoBDT	LHD
$k_{t\bar{t}+\geq 1b}$	$1.23^{+0.10}_{-0.10}$	$1.29^{+0.10}_{-0.10}$	$1.34^{+0.10}_{-0.10}$
$k_{t\bar{t}+\geq 1c}$	$1.61^{+0.23}_{-0.22}$	$1.48^{+0.24}_{-0.23}$	$1.42^{+0.22}_{-0.21}$

FIGURE 12.34: Summary of the observed best-fit values of $\mu_{t\bar{t}H}$ comparing the different sets of distributions in the signal regions.

12.4 Event Yields after Fitting

TABLE 12.4: Yields of the analysis regions in ≥ 6 jets for the single-lepton channel for 36.1 fb^{-1} after fitting. The " $t\bar{t}V$ " row includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" row includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " row has WtZ and tH processes.

	$\geq 6 \text{ jets, CR}(t\bar{t}+\text{light})$	$\geq 6 \text{ jets, CR}(t\bar{t}+\geq 1c)$	$\geq 6 \text{ jets, CR}(t\bar{t}+1b)$
$t\bar{t} + \text{light}$	$109\,000 \pm 4490$	3380 ± 427	1840 ± 165
$t\bar{t} + \geq 1c$	$45\,400 \pm 4930$	5290 ± 672	2080 ± 294
$t\bar{t} + \geq 1b$	$14\,800 \pm 1370$	2920 ± 274	4070 ± 313
$t\bar{t}V$	1010 ± 94.0	119 ± 14.0	68.1 ± 7.21
single-top	5930 ± 1510	387 ± 115	247 ± 72.2
$t\bar{t}WW$	15.1 ± 1.84	1.75 ± 0.278	0.524 ± 0.120
$t\bar{t}t\bar{t}$	18.5 ± 9.23	8.28 ± 4.13	5.27 ± 2.63
non top	5870 ± 1370	599 ± 166	340 ± 82.4
tH	23.2 ± 1.99	4.89 ± 0.525	5.28 ± 0.481
$t\bar{t}H$	384 ± 278	87.1 ± 63.4	84.3 ± 61.2
Total	$182\,000 \pm 3190$	$12\,800 \pm 355$	8750 ± 223
Data	181706	12778	8576

	$\geq 6 \text{ jets, SR}(t\bar{t}+\geq 2b, \text{Lo})$	$\geq 6 \text{ jets, SR}(t\bar{t}+\geq 2b, \text{Hi})$	$\geq 6 \text{ jets, SR}(t\bar{t}H)$
$t\bar{t} + \text{light}$	575 ± 93.9	101 ± 33.4	11.7 ± 5.67
$t\bar{t} + \geq 1c$	1340 ± 183	479 ± 99.9	44.9 ± 20.0
$t\bar{t} + \geq 1b$	2220 ± 163	1850 ± 134	1030 ± 62.4
$t\bar{t}V$	50.7 ± 5.94	40.4 ± 4.89	25.5 ± 3.33
single-top	142 ± 51.5	93.9 ± 45.2	38.3 ± 16.8
$t\bar{t}WW$	0.450 ± 0.107	0.241 ± 0.0793	0.0610 ± 0.0389
$t\bar{t}t\bar{t}$	6.73 ± 3.36	6.27 ± 3.13	4.05 ± 2.03
non top	120 ± 35.6	35.1 ± 11.6	15.2 ± 4.76
tH	3.69 ± 0.399	3.57 ± 0.393	2.22 ± 0.256
$t\bar{t}H$	72.1 ± 52.1	68.8 ± 49.7	52.3 ± 37.9
Total	4520 ± 131	2680 ± 96.6	1220 ± 50.6
Data	4698	2641	1222

TABLE 12.5: Yields of the analysis regions in 5 jets for the single-lepton channel for 36.1 fb^{-1} after fitting. The " $t\bar{t}V$ " row includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" row includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " row has WtZ and tH processes.

	5 jets, CR($t\bar{t}$ +light)	5 jets, CR($t\bar{t}$ + $\geq 1c$)	5 jets, CR($t\bar{t}$ + $1b$)
$t\bar{t}$ + light	$182\,000 \pm 5570$	2320 ± 217	3610 ± 239
$t\bar{t}$ + $\geq 1c$	$43\,400 \pm 5350$	1820 ± 249	2570 ± 379
$t\bar{t}$ + $\geq 1b$	$13\,300 \pm 1200$	929 ± 88.9	3990 ± 317
$t\bar{t}V$	585 ± 54.7	22.1 ± 2.97	46.7 ± 5.38
single-top	$10\,400 \pm 1820$	200 ± 40.5	357 ± 79.2
$t\bar{t}WW$	2.87 ± 0.391	0.178 ± 0.0791	0.151 ± 0.0604
$t\bar{t}t\bar{t}$	0.844 ± 0.427	0.0592 ± 0.0350	0.256 ± 0.131
non top	$10\,900 \pm 2620$	257 ± 74.2	487 ± 134
tH	19.3 ± 1.51	1.44 ± 0.202	6.27 ± 0.545
$t\bar{t}H$	190 ± 138	15.5 ± 11.4	57.3 ± 41.7
Total	$261\,000 \pm 4970$	5560 ± 188	$11\,100 \pm 273$
Data	259320	5465	11095

	5 jets, SR($t\bar{t}$ + $\geq 2b$)	5 jets, SR($t\bar{t}H$)	Boosted SR
$t\bar{t}$ + light	406 ± 68.4	13.4 ± 9.88	119 ± 33.8
$t\bar{t}$ + $\geq 1c$	608 ± 102	27.7 ± 14.4	232 ± 38.7
$t\bar{t}$ + $\geq 1b$	1410 ± 106	327 ± 25.6	225 ± 32.3
$t\bar{t}V$	19.5 ± 2.43	6.34 ± 1.18	16.2 ± 2.43
single-top	110 ± 38.6	16.1 ± 5.29	29.3 ± 18.7
$t\bar{t}WW$	0.	0.	0.526 ± 0.114
$t\bar{t}t\bar{t}$	0.209 ± 0.113	0.0342 ± 0.0189	4.13 ± 2.07
non top	112 ± 34.6	11.8 ± 6.86	48.1 ± 12.6
tH	3.24 ± 0.341	1.40 ± 0.174	2.00 ± 0.249
$t\bar{t}H$	33.8 ± 24.5	13.4 ± 9.75	14.5 ± 10.5
Total	2710 ± 84.6	417 ± 24.5	691 ± 39.8
Data	2798	426	740

TABLE 12.6: Yields of the analysis regions for the dilepton channel for 36.1 fb^{-1} after fitting. The " $t\bar{t}V$ " row includes $t\bar{t}W$ and $t\bar{t}Z$. The Wt , single-top (s,t-channel), WtZ , and tZ events are shown in the "single-top" row. The "non top" row includes W +jets, Z +jets, VV , and fake-leptons processes. The " tH " row has WtZ and tH processes.

	3 jets, CR($t\bar{t}$ +light)	3 jets, CR($t\bar{t}+\geq 1b$)	≥ 4 jets, CR($t\bar{t}$ +light)	≥ 4 jets, CR($t\bar{t}+\geq 1c$)
$t\bar{t}$ + light	$59\,600 \pm 1720$	260 ± 44.9	$37\,500 \pm 1360$	1420 ± 174
$t\bar{t} + \geq 1c$	7470 ± 1010	547 ± 85.7	$10\,300 \pm 1270$	2160 ± 283
$t\bar{t} + \geq 1b$	2600 ± 236	845 ± 75.7	2890 ± 295	1250 ± 105
$t\bar{t}V$	112 ± 29.1	6.79 ± 27.3	331 ± 170	52.0 ± 42.1
single-top	2380 ± 575	45.7 ± 11.4	1520 ± 399	126 ± 46.8
$t\bar{t}WW$	0.489 ± 0.0811	$0.008\,78 \pm 0.008\,78$	6.47 ± 0.266	1.16 ± 0.135
$t\bar{t}t\bar{t}$	0.0588 ± 0.0126	$0.0122 \pm 0.005\,33$	3.82 ± 0.0875	2.20 ± 0.0708
non top	4030 ± 1020	64.9 ± 19.9	3440 ± 819	324 ± 82.0
tH	2.64 ± 0.245	0.537 ± 0.0979	6.06 ± 0.525	1.78 ± 0.199
$t\bar{t}H$	27.1 ± 19.9	7.27 ± 5.37	96.9 ± 70.0	29.8 ± 21.7
Total	$76\,200 \pm 1360$	1780 ± 67.9	$56\,000 \pm 1130$	5360 ± 157
Data	76025	1744	55627	5389

	≥ 4 jets, SR($t\bar{t}+1b$)	≥ 4 jets, SR($t\bar{t}+\geq 2b$)	≥ 4 jets, SR($t\bar{t}H$)	
$t\bar{t}$ + light	91.5 ± 28.3	217 ± 43.0	10.0 ± 8.56	
$t\bar{t} + \geq 1c$	313 ± 52.2	431 ± 87.6	26.1 ± 7.80	
$t\bar{t} + \geq 1b$	923 ± 72.7	669 ± 58.0	263 ± 20.1	
$t\bar{t}V$	14.8 ± 34.9	21.7 ± 38.3	6.56 ± 56.2	
single-top	30.2 ± 12.0	33.9 ± 18.3	4.55 ± 2.84	
$t\bar{t}WW$	0.0902 ± 0.0275	0.447 ± 0.0832	0.0121 ± 0.0171	
$t\bar{t}t\bar{t}$	1.30 ± 0.0507	3.91 ± 0.107	1.71 ± 0.0649	
non top	46.1 ± 10.4	82.9 ± 22.3	7.98 ± 2.07	
tH	0.741 ± 0.111	1.18 ± 0.184	0.444 ± 0.0686	
$t\bar{t}H$	18.3 ± 13.3	24.7 ± 17.8	13.1 ± 9.46	
Total	1440 ± 63.0	1490 ± 70.9	333 ± 59.0	
Data	1467	1444	319	

Chapter 13

Combined Result with Other $t\bar{t}H$ Channels

13.1 Combined Results

The $t\bar{t}H$ production analysis has been performed by ATLAS in the following four channels:

- $H \rightarrow b\bar{b}$ [26] (discussed in this thesis)
- $H \rightarrow \gamma\gamma$ [27]
- $H \rightarrow ZZ^* \rightarrow 4\ell$ [28]
- $H \rightarrow WW^*, H \rightarrow \tau\tau$ (multileptons) [29]

All analyses use the same data (36.1 fb^{-1}) and the same MC generators for $t\bar{t}H$ production, and use the nominal Higgs boson decay branching fractions from Table 1.3 assuming $m_H = 125 \text{ GeV}$. For the extraction of the $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$, the associated single top Higgs boson production processes $tHjb$ and WtH were considered as backgrounds and fixed to their SM expectations with theoretical uncertainties. While $H \rightarrow \text{multileptons}$ and $H \rightarrow b\bar{b}$ channels can neglect contaminations from other Higgs productions in their searches, $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ utilize a global fit with all Higgs-production signal strengths. In combining the results, only the $t\bar{t}H$ enhanced regions were considered and other productions were fixed to the SM expectation values. Contaminations by other Higgs productions are 4-21% in the $t\bar{t}H$ enriched region, 21-64% in the $tHjb/WtH$ regions for $H \rightarrow \gamma\gamma$, 23% in the $t\bar{t}H$ enriched region for $H \rightarrow ZZ^* \rightarrow 4\ell$.

The individual best fit results and their combined result are shown in Figure 13.1. As no events were observed in the $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis, a 68% confidence-level (CL) upper limit on $\mu_{t\bar{t}H}$ was reported. The combined $t\bar{t}H$ signal strength was:

$$\mu_{t\bar{t}H} = 1.17 \pm 0.19 \text{ (stat)}^{+0.27}_{-0.23} \text{ (syst)},$$

where no-signal assumption $\mu_{t\bar{t}H}$ was excluded at 4.2σ with an expectation of 3.8σ in the case of the SM signal assumption, first evidence for $t\bar{t}H$ production. The corresponding cross section value of $t\bar{t}H$ production is $590^{+160}_{-150} \text{ fb}$ as compared to the SM prediction $507^{+35}_{-50} \text{ fb}$. The impact of various uncertainties on the combination is shown in Table 13.1. The contribution from $t\bar{t}H$ modeling has the largest impact in the combination result. The $t\bar{t}$ +jets modeling has the second largest impact. More accurate modeling of $t\bar{t}H$ and $t\bar{t}$ are required in future for more precise top-Yukawa measurement.

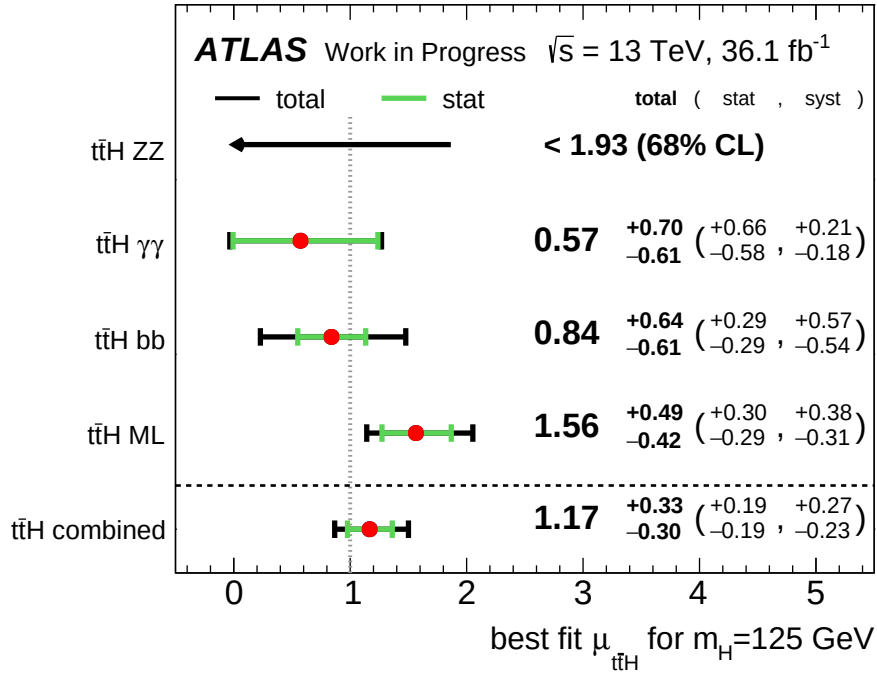


FIGURE 13.1: Summary of $t\bar{t}H$ measurements of μ from various Higgs decay channels and the combined result. "ML" refers to the multilepton channel. The best-fit values of μ for the individual analyses are extracted independently. The systematic uncertainty nuisance parameters are correlated for the combination. As no events are observed in the $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis, a 68% CL upper limit is shown.

TABLE 13.1: Summary of the uncertainties affecting the combined $\mu_{t\bar{t}H}$ value.

Uncertainty source	$\Delta\mu_{t\bar{t}H}$	
$t\bar{t}H$ modeling	+0.15	-0.09
$t\bar{t}$ modeling	+0.12	-0.12
Fake lepton estimates	+0.08	-0.08
$t\bar{t}V$ modeling	+0.07	-0.07
Background statistics	+0.07	-0.07
Jet energy scale and resolution	+0.06	-0.06
Other background modeling	+0.06	-0.06
Flavor tagging	+0.05	-0.05
Luminosity	+0.04	-0.04
Other experimental unc.	+0.04	-0.03
Other Higgs prod. modeling	+0.01	-0.01
Total systematic uncertainty	+0.27	-0.23
Statistical uncertainty	+0.19	-0.19
Total uncertainty	+0.33	-0.30

13.2 Coupling Measurement

The $t\bar{t}H$ analyses including various Higgs decay channels are sensitive to the Yukawa couplings of the top quark, bottom quark, and τ -lepton, and the gauge couplings (W and Z). The top-Yukawa measurement only via the $t\bar{t}H$ production has no constraints to the sign of the top-Yukawa with relation to the gauge couplings. However, interferences in the $H \rightarrow \gamma\gamma$ loop and in the $tHjb$ and WtH productions increase the sensitivity to the strength measurements as well as to the signs of interference between fermion and boson couplings. In this section, following observables were used:

- $t\bar{t}H$ production cross section
- Higgs decay branching ratios ($H \rightarrow b\bar{b}$, $\tau\tau$, WW^* , and ZZ^*)
- tH ($tHjb$ and WtH) production cross section (sensitive to W - H and t - H interference)
- W - t interference in the $H \rightarrow \gamma\gamma$ loop (assuming only SM contributions in the loop)

The coupling measurement was performed using the κ -parametrization, in which the Higgs boson coupling to particle i is parametrized as $\kappa_i = \kappa_i^{\text{obs}} / \kappa_i^{\text{SM}}$. At first, for the simplicity, only top-Yukawa κ_{top} was set to be changed by the global fit, and other κ modifications were fixed to SM predictions ($\kappa_{b,\tau,W,Z} = 1$). The impact from the interferences on the top-Yukawa measurement was checked using $t\bar{t}H$ ($H \rightarrow \gamma\gamma$) channel, as shown in Figure 13.2. The top-Yukawa κ_{top} was measured as

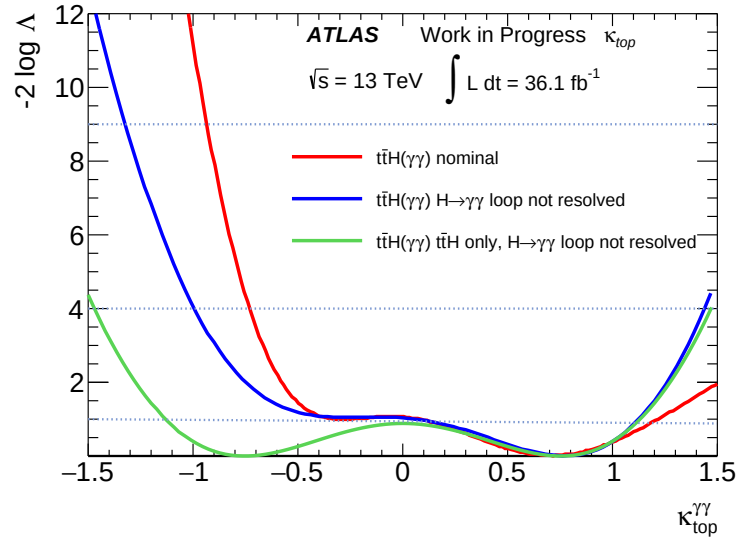


FIGURE 13.2: Profiled likelihood ratio (q_μ) as a function of κ_{top} using $t\bar{t}H$ ($H \rightarrow \gamma\gamma$) channel. The each profiled likelihood ratio curve is subtracted by its minimum value. Three options were tested. The red line used both interferences in $tHjb$ and WtH production cross sections and $H \rightarrow \gamma\gamma$ loop. The blue line used the interference in $tHjb$ and WtH production cross sections, and the $H \rightarrow \gamma\gamma$ branching ratio was fixed to SM prediction. The green line used no interferences, and has no constraints on the sign of κ_{top} . The gray dashed lines show exclusion limits for 1σ , 2σ , and 3σ . Using interferences, minus sign of κ_{top} is excluded at almost 1σ .

shown in Figure 13.3. The direct top-Yukawa measurement was achieved by not resolving the $H \rightarrow \gamma\gamma$ loop contribution, where the plus sign of κ_{top} is preferred. The κ_{top} value at the minimum profiled likelihood ratio is 1.08 ± 0.15 . The κ_{top} value at the local minima profiled likelihood ratio

in the minus sign of κ_{top} is -0.87 with $-2\ln\Lambda = 2.56$. This corresponds 1.6σ deviation from the minus signed κ_{top} . Using the $H \rightarrow \gamma\gamma$ loop contribution, further precise measurement can be achieved.

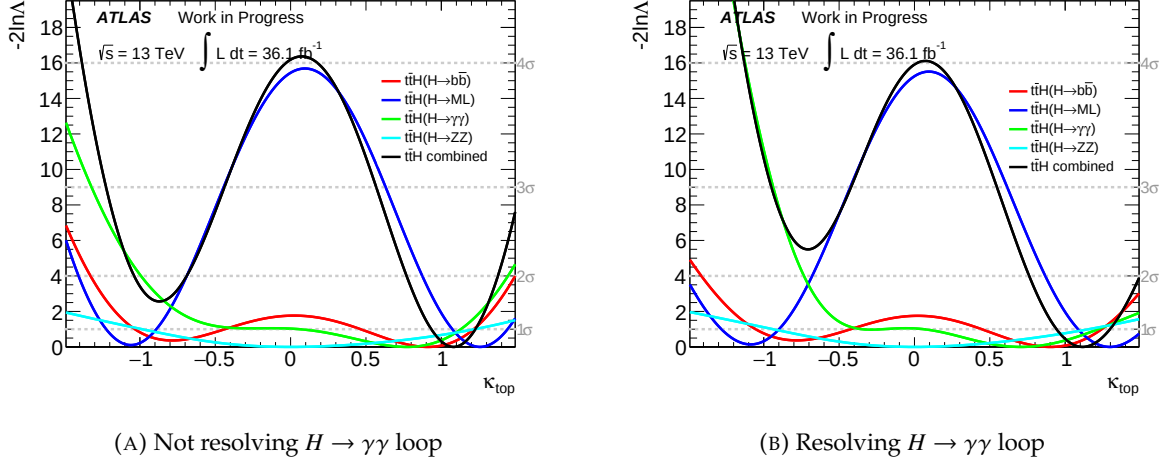


FIGURE 13.3: Profiled likelihood ratio (q_μ) as a function of κ_{top} using all $t\bar{t}H$ analysis channels. The $H \rightarrow \gamma\gamma$ loop was not resolved in the left plot, and the direct top-Yukawa measurement was performed. In the left plot, the $H \rightarrow \gamma\gamma$ loop was resolved to κ_{top} , κ_b , and κ_W by assuming no BSM contributions in the loop. The each profiled likelihood ratio curve is subtracted by its minimum value. The individual curves for different analysis channels and combined curve are shown. The gray dashed lines show exclusion limits for 1σ , 2σ , 3σ , and 4σ .

For a further limit to the Yukawa-coupling, a two-parameter fit was made assuming that all the boson couplings scale with the same coupling modifier κ_V , while all the fermion couplings scale with the same coupling modifier κ_F . This constrained parameterization is motivated by the intrinsic difference between the Higgs boson couplings to vector bosons, which originate from the EW symmetry breaking, and the Yukawa couplings to the fermions.

The results are shown in Figure 13.4, and are in good agreement with $\kappa_F = \kappa_V = 1$. The best fit values are $\kappa_F = 1.1$ and $\kappa_V = 1.0$. The κ_F value at the local minima of the profiled likelihood ratio in the minus signed κ_F is -0.71 , where the profiled likelihood ratio is 4.0 (2σ deviation from the minus signed κ_{top} hypothesis). The possibility of $\kappa_F < 0$ is excluded at more than 95% CL.

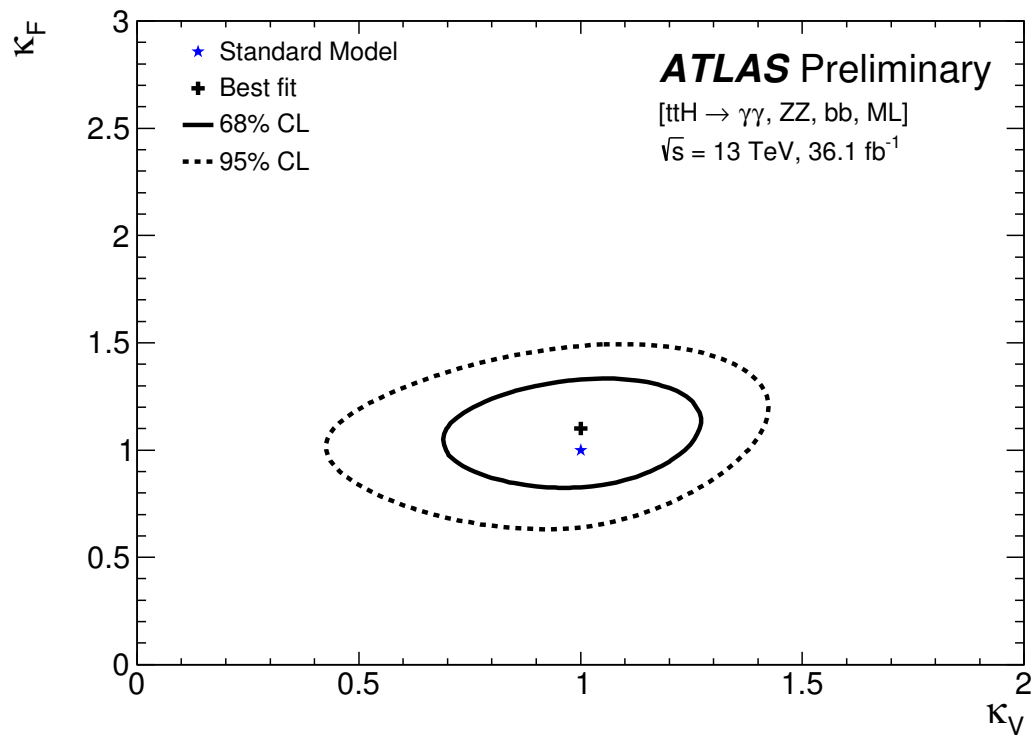


FIGURE 13.4: Allowed regions at 68% and 95% CL in the κ_F - κ_V plane obtained from the combination of all $t\bar{t}H$ channels analyses [29]. The Higgs boson is assumed not to couple to any particles beyond the Standard Model, and the $H \rightarrow \gamma\gamma$ and $H \rightarrow gg$ couplings are expressed in terms of κ_F and κ_V .

Chapter 14

Conclusions

A search for the Standard Model Higgs boson produced in association with top quarks ($t\bar{t}H$ production) was presented. The 36.1 fb^{-1} data of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ was collected with the ATLAS detector at the Large Hadron Collider in 2015 and 2016. The search focused on the Higgs boson decaying to $b\bar{b}$ and the top-quarks pair decaying to final states containing one or two leptons. With the large $H \rightarrow b\bar{b}$ branching ratio and detection of associated $t\bar{t}$ in leptonic decay channels, we expect large statistics and significant rejection of dominating QCD backgrounds. This analysis was limited by the modeling uncertainty of $t\bar{t}$ plus additional jets productions. To maximize the sensitivity and control $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ productions, the analysis regions were optimized using b -tagging with four working points. Multivariate techniques were then utilized to discriminate between the $t\bar{t}H$ signal events and background events dominated by $t\bar{t} + b\bar{b}$ production. The number of signal events determined as the ratio of the SM expectation, signal strength = $\mu_{t\bar{t}H} = \sigma_{t\bar{t}H} / \sigma_{t\bar{t}H}^{\text{SM}}$, was measured to be $0.84^{+0.64}_{-0.61}$ corresponding to the observed $t\bar{t}H$ signal significance of 1.4σ with an expectation of 1.6σ . The measurement uncertainty is dominated by systematic uncertainties mainly in the theoretical knowledge of the $t\bar{t} + \geq 1b$ production process. An improved modeling of this background will be important for future efforts to determine the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) process more precisely.

Although no significant excess was observed in data compared to the no $t\bar{t}H$ hypothesis in the $H \rightarrow b\bar{b}$ analysis channel alone, with the combination of available channels of $t\bar{t}H$ searches, the no-signal assumption $\mu_{t\bar{t}H}$ was excluded at 4.2σ with an expectation of 3.8σ in the case of the SM signal assumption. This provides the first evidence for $t\bar{t}H$ production. The corresponding cross section value of the $t\bar{t}H$ production is $590^{+160}_{-150} \text{ fb}$ as compared to the SM prediction of $507^{+35}_{-50} \text{ fb}$.

Appendix A

MVA Algorithms for Machine Learning

The multi-variate analysis (MVA) combines a lot of weak classifiers (called "input variables"), and forms one powerful discriminant to classify a set of events into two categories (the signal and background categories). The MVA using a machine learning requires "training" events to know the two category tendencies from input variables. Each training event has true information which category it should be classified in. To avoid the bias, the set of training events has to be different from the set of events to be analyzed. If the training sample has not enough statistics towards the number of input variables, the training will learn the statistical fluctuations in the training sample, and not fully optimized resulting a worse MVA performance than the expected, which is called "over-training". In this thesis, two sets of training are prepared according to the sequential event number added to each event:

- odd-numbered analysis uses the MVA discriminant trained by even-numbered events.
- even-numbered analysis uses the MVA discriminant trained by odd-numbered events.

Training events are completely orthogonal to the events to be analyzed in this configuration. In addition, additional set of events is not required, although over-training may be seen due to the half statistics of the set of events for each training sample.

The MVA techniques are not unique in the high energy physics, but widely used in various areas using the statistic analysis. Therefore, a lot of algorithms for how to combine information of the input variables have been developed for various purposes. In this section, two algorithms used in this thesis are explained, BDT and NN.

All MVAs described in this thesis are implemented in the toolkit for MVA (TMVA) package [111].

A.1 Boosted-Decision-Tree

A decision tree (DT) is an extension of the cut selections with multiple variables. The DT has a consecutive set of variable selections, and each selection has one threshold to separate two categories. Each selection depends on the former classification, and the final discriminant classifies each event into two categories after a given maximum number of selections, called "depth". The DT chooses variables for criterion by maximizing the separation gain computed as Gini-index $\sim p(1-p)$, where p is the signal purity of the signal category. The advantages of the DT are easy to understand and interpretable by human, and fast training. However, a single tree does not usually have enough power for a good separation in the final discriminant. The schematics of the DT algorithm is shown in Figure A.1.

The boosted decision tree (BDT) algorithm is a set of DT iterations to enhance the separation power. The AdBoost method is adopted in this thesis. If there are mis-classified events in the

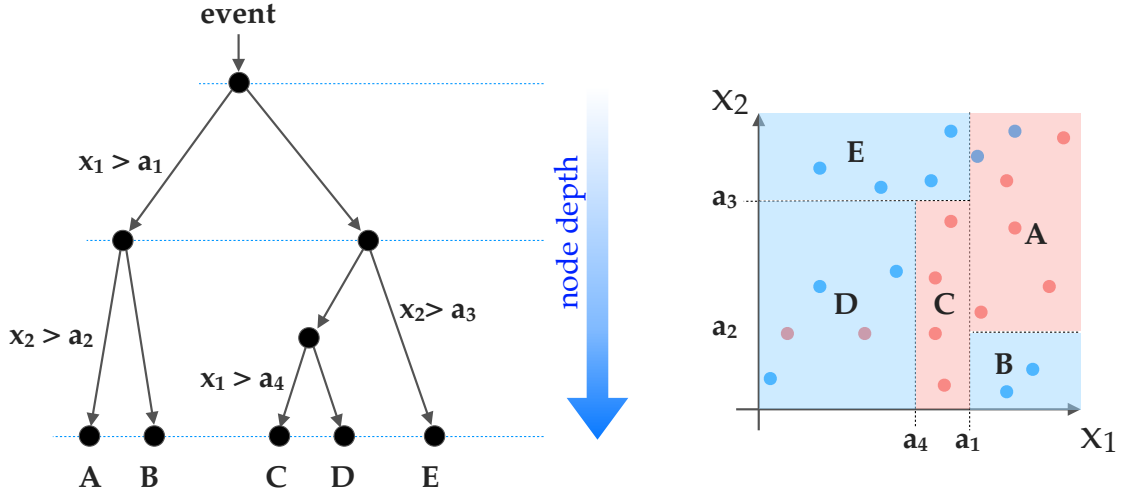


FIGURE A.1: (Left) Schematics of decision tree algorithm. Two input variables (x_1 and x_2) are used for explanation. In each node, the threshold for the variable selection is optimized as a_i ($i = 1, 2, 3, 4$). The final discriminant classify events into five regions (A, B, C, D, E). Regions named as "A" and "C" are signal categories, and other regions are background categories. (Right) The DT with two input variables can be illustrated as two dimensional map. Red (Blue) points show the signal (background) category events. A few points are mis-identified by the DT, such as blue points in the A region and red points in the D region.

first DT algorithm, the weights are applied to these events by $(1 - f_{\text{err}})/f_{\text{err}}$, where f_{err} is mis-classified fraction in the training events. In the next iteration, these weights are used to calculate the Gini-index in each selection optimization (weighted Gini-index = $\sum w \times p(1 - p)$). The iteration is continued until the mis-identification rate is reduced to the given value. Final discriminant is the weighted sum of all classifiers as:

$$\text{BDT output}(x) = \frac{1}{N_{\text{DT}}} \sum_i^{N_{\text{DT}}} \log \left(\frac{1 - f_{\text{err}}^i}{f_{\text{err}}^i} \right) C^i(x), \quad (\text{A.1})$$

where x is the category index, N_{DT} is the number of DTs corresponding to the number of iterations, and $C^i(x)$ is the classifier in the i -th iteration ($C^i(x) = 1$ assigns the event to the signal category and $C^i(x) = -1$ assigns to the background category). The BDT is optimized to reduce the mis-classified rate, resulting a better performance than a single DT. The schematics of the BDT procedure is shown in Figure A.2.

The BDT is widely used in the ATLAS experiment. The flavor-tagging discriminant (MV2c10) adopts the BDT to combine three methods to identify b -jets with the rejection of c -jets and light-jets, as discussed in Chapter 3. The $t\bar{t}H(H \rightarrow b\bar{b})$ analysis utilizes the BDT method in the reconstruction of $t\bar{t}H$ system, and in the final discriminant to separate $t\bar{t}H$ signal from the $t\bar{t}$ -jets background processes, as discussed in Chapters 9 and 10.

A.2 Neural-Network

The neural network (NN) is powerful and flexible algorithm originated from modeling the brain function, which is a collection of interconnected neurons. At first, each input variable is normalized to range from -1 to 1, and assigned to a neuron. These neurons compose the input layer. All

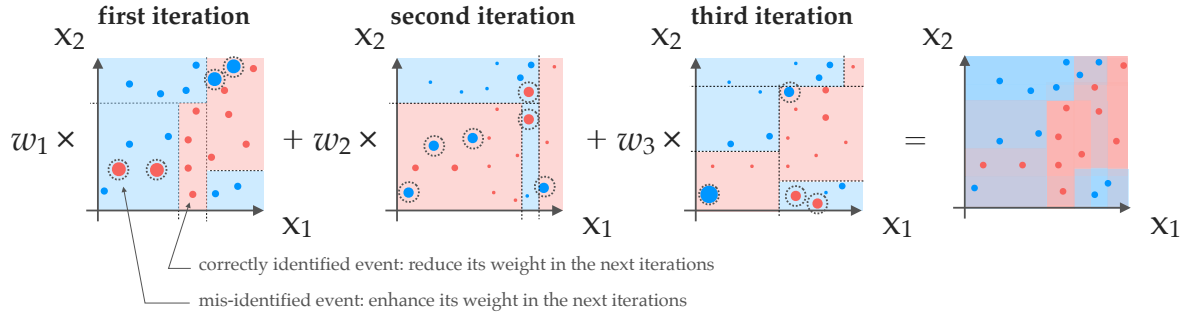


FIGURE A.2: Schematics of BDT algorithm. Same as in Figure A.1, the BDT with two input variables (x_1 and x_2) is considered. Each two dimensional map shows each DT algorithm result. Red (Blue) points show the signal (background) category events, and their sizes correspond to event weights. Large (small) weights are applied to the mis-classified (correctly assigned) events in the iteration. After the iteration, the final discriminant is weighted sum of the DT classifiers resulting complicated borders for the categorization on the map.

neurons in the input layer are combined with a series of weights in a neuron. Taking various sets of weights, various neurons are also formed. They are put in together as the hidden layer, which is intermediate state for the final discriminant of NN and is not seen by users. In the combining the input neurons, offset term is also summed together. After that, all neurons in the hidden layer are combined with weighted summation into one neuron. This neuron variable is the final discriminant of the NN. To significantly separate between the signal and background, the activated function is applied for each neurons in the hidden layer before the summation. The schematics of the NN procedure is shown in Figure A.3.

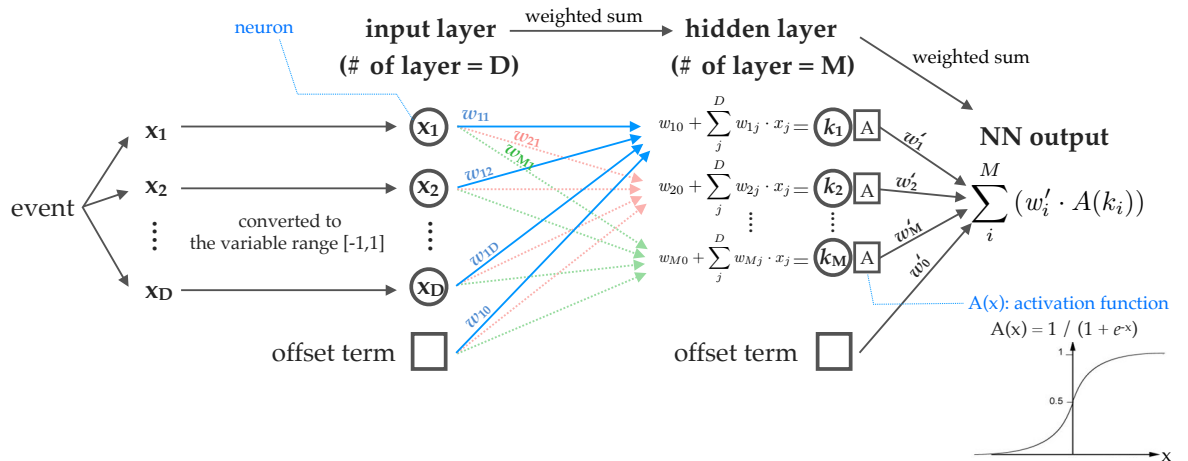


FIGURE A.3: Schematics of NN algorithm. The number of input variables is denoted by D , and the number of neurons in the hidden layer is denoted by M . x_i is the i -th input variable. w_{ji} is a weight value for the i -th input variable to calculate j -th neuron variable (k_j) in the hidden layer. w'_i is a weight value for k_j to calculate the NN output. The activated function is applied to k_j before the summation for the NN output. As the activated function, the sigmoid function is shown in this figure.

The training optimizes a series of weights used in the NN to realize the best separation of two categories. The correlations between the output score and the input variables can be nonlinear when neurons have nonlinear responses to their input through the activated function.

The NN is used in the $t\bar{t}H(H \rightarrow b\bar{b})$ Run 1 analysis. The concept in the NN is complicated than the BDT, but the performance is not so different as verified by inputting the same variables to both algorithms (see Appendix A.4).

A.3 MVA Performance

To see the performance of the MVA, the receiver operating characteristic (ROC) curve is useful. The ROC curve shows the background event rejection over the signal event efficiency in the analysis events. Generally, the better classifier can be identified by the larger area under curve (AUC). And overtraining can be seen by the smoothness of the ROC curve. Input variable optimization is performed by maximizing the AUC of the ROC curve.

A.4 Comparison between BDT and NN

For a comparison of the MVA performance between the BDT and NN algorithms, one NN was trained with the same input variables and same training samples to the ClassBDT for the inclusive $6j$ region of the present analysis as shown in Table 10.1. The output score of both algorithms are shown in Figure A.4, and the ROC curves are shown in Figure A.5.

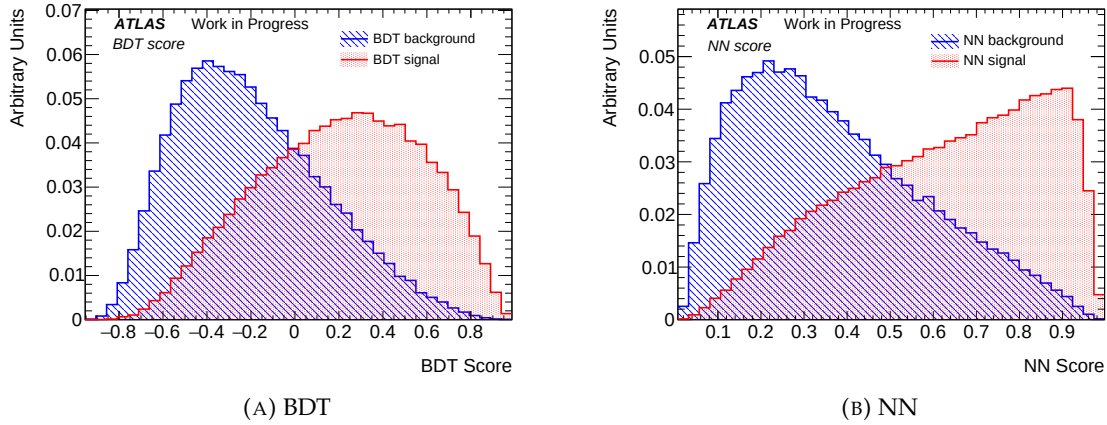


FIGURE A.4: Output scores of BDT and NN algorithms. The input variables are listed in the " ≥ 6 jets" column in Table 10.1. The $t\bar{t}H$ and $t\bar{t}$ +jets simulations are used as the signal and background samples, respectively. Red lines show the signal scores, and blue lines show the background scores in both figures.

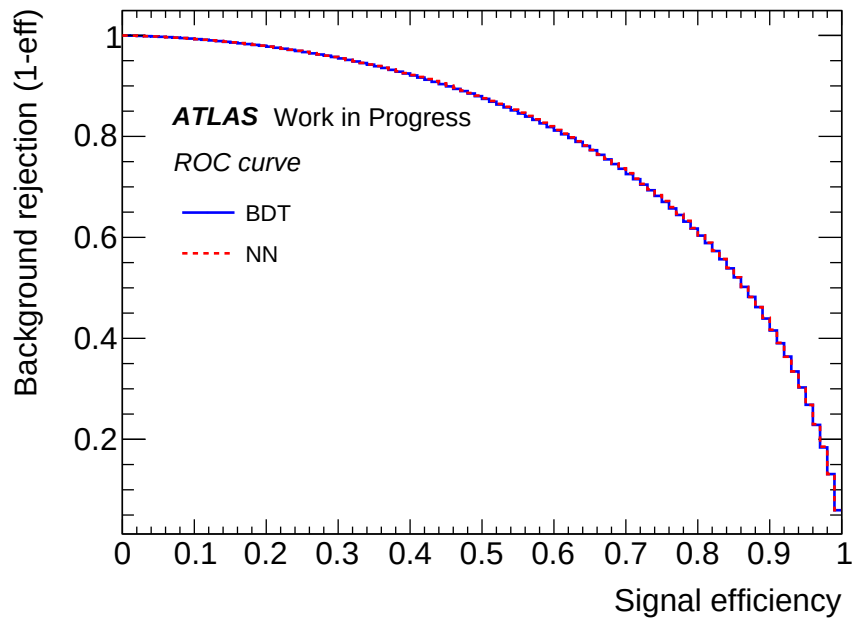


FIGURE A.5: ROC curves of BDT and NN algorithms. Two algorithms have almost same curves. The differences are less than 0.5% in all bins. The AUCs of the ROC curves are 0.78 in both algorithms.

Appendix B

Tips on Plots

Explanations for plots comparing the data and prediction are given in this section. The example plot is shown in Figure B.1.

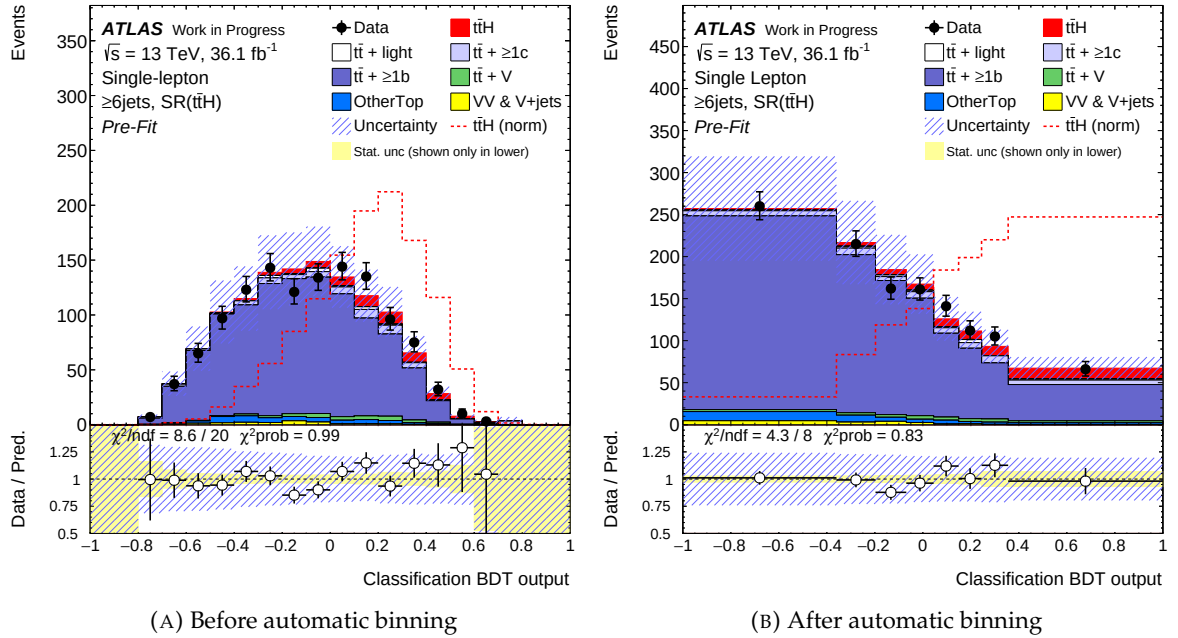


FIGURE B.1: BDT scores (A) before and (B) after automatic binning in $\geq 6j\text{SR}(t\bar{t}H)$ in the single-lepton channel. Before automatic binning, BDT distribution was divided into 20 bins with same bin widths. In the upper plots, the number of events in each bin is shown. Each color shows different prediction sample by the SM MC simulation. Total uncertainty including both systematic and MC statistical uncertainties is shown with the shaded band in each bins. Red dotted line shows the signal event shape normalized to the total background prediction. Black points are data shown with its statistical uncertainties. The 36.1 fb^{-1} data was collected with pp -collisions at $\sqrt{s} = 13 \text{ TeV}$. Both plots used original uncertainties before the global fitting for $t\bar{t}H$ signal determination (denoted as "Pre-Fit"). In the lower plots, the ratios of the data and prediction are shown with white points. The yellow band shows the MC statistic uncertainties. To qualify the consistency between the data and prediction, the χ^2 value and its probability are shown.

B.1 Sample Composition

In the data-prediction comparison plots, the minor samples are merged together. The samples including top quarks (WtH , tH , Wt , single-top(t/s -ch), WtZ , tZ , $t\bar{t}WW$, and $t\bar{t}t\bar{t}$) are shown together

as "OtherTop". The MC samples not including top quarks (VV , W +jets, and Z +jets) are shown together as " VV & V +jets".

B.2 Ratio Plot

In the lower part of the data-prediction comparison plot, the ratio of the data and prediction is shown. The prediction includes all background MC sources, fake leptons estimated by the data, and the signal events predicted by the SM.

B.3 χ^2 Probability

To qualify the consistency between the data and prediction, the χ^2 value and its probability are shown in each data and prediction plot. The χ^2 is defined as:

$$\chi^2 = \sum_i^{N_{\text{bins}}} \sum_j^{N_{\text{bins}}} \frac{(\Delta_i \times \Delta_j)}{\sigma_{ij}^2}, \quad (\text{B.1})$$

where i and j are bin indices. $\sigma_{ij}(i \neq j)$ is total systematic correlations with the j -th bin, assigned to the i -th bin. σ_{ii} is total uncertainty on the i -th bin, which includes all systematic uncertainties, data and MC statistical uncertainties.

To compare the χ^2 value to the other plots with different number of bins, the number of bins (ndf) is shown in each plot. Using two quantities, χ^2 and ndf, the null hypothesis probability is calculated through the cumulative density function of the χ^2 function. The calculated probability describes the probability that the plot have consistency between data and prediction less than the observed χ^2 score.

B.4 Automatic Binning

In the signal sensitive distributions which have the signal rich bins on either side and the background rich bins on opposite side, automatic binning procedure is performed to maximize the sensitivity of the signal. The procedure is given two parameters, the parameters weighting the signal and background events, z_S and z_B , respectively. Input histograms are characterized by very finer bins at first, such as 10000. The binning score is defined as:

$$Z = \frac{z_S \cdot n_S}{N_S} + \frac{z_B \cdot n_B}{N_B}, \quad (\text{B.2})$$

where the total number of events of the signal and background is denoted as N_S and N_B , respectively. n_S and n_B denote the number of the signal and background in the merged bin. The merging procedure starts from the signal enriched end bin. The score Z is calculated with merging every bin. If $Z > 1$, merging of the bin stops and the merged bin is kept as a new bin of the plot. From the next bin, the new merging procedure starts again until all bins are merged.

In this thesis, both two parameters (z_S and z_B) are set to 4 for signal regions in the single-lepton channel. They are set to 2, 3, and 4 for $\text{SR}(t\bar{t}H)$, $\text{SR}(t\bar{t}+\geq 2b)$, and $\text{SR}(t\bar{t}+1b)$ in ≥ 4 jets in the dilepton channel not to have any bins with large statistical uncertainties. In Figure B.1, the ClassBDT distribution in the $\geq 6j\text{SR}(t\bar{t}H)$ (see 10) is shown before and after automatic binning.

Binning in kinematic plots is optimized not to have any bins with large MC and data statistical uncertainties.

Appendix C

Tips on Systematic Uncertainty

C.1 Systematic Variation

The systematic variations in two directions are defined as $\pm 1\sigma$ symmetric variations about the nominal value. If the variation is not symmetric, the variation is modified to be symmetric:

$$\text{modified } \pm 1\sigma \text{ variation} = \pm [(+1\sigma \text{ variation}) + (-1\sigma \text{ variation})] / 2. \quad (\text{C.1})$$

On the other hand, for the systematic variation defined only in one direction ($+1\sigma$), the other direction is defined as the reversed $+1\sigma$ variation with relation to the nominal.

C.2 Smoothing

To assign the systematic uncertainty shapes as shown in Figures 11.1 to 11.12, the uncertainty shapes were smoothed [112] to flatten any bumps created due to poor statistics in the systematic sources. The modification in each bin is written as:

$$N'_i = \frac{1}{4}N_{i-1} + \frac{1}{2}N_i + \frac{1}{4}N_{i+1}, \quad (\text{C.2})$$

where N_i and N'_i is the number of events in the i -th bin before and after smoothing. This smoothing works well to remove unphysical spikes. However, if the histogram has real peaks, valleys, and any regions with large second derivatives, smoothing works to vanish them. To avoid these over-smoothing and keep the smoothness in the histogram, the second order smoothing was further applied. The final smoothed number of events in the i -th bin, N''_i is described as:

$$N''_i = N'_i + \left(\frac{1}{4}(N'_{i-1} - N_{i-1}) + \frac{1}{2}(N'_i - N_i) + \frac{1}{4}(N'_{i+1} - N_{i+1}) \right). \quad (\text{C.3})$$

Several systematic shapes before and after smoothing are shown in Figure C.1. To obtain the enough smoothed shape for each systematic variation, the smoothing procedure in all bins by Equation C.3 was applied 40 times.

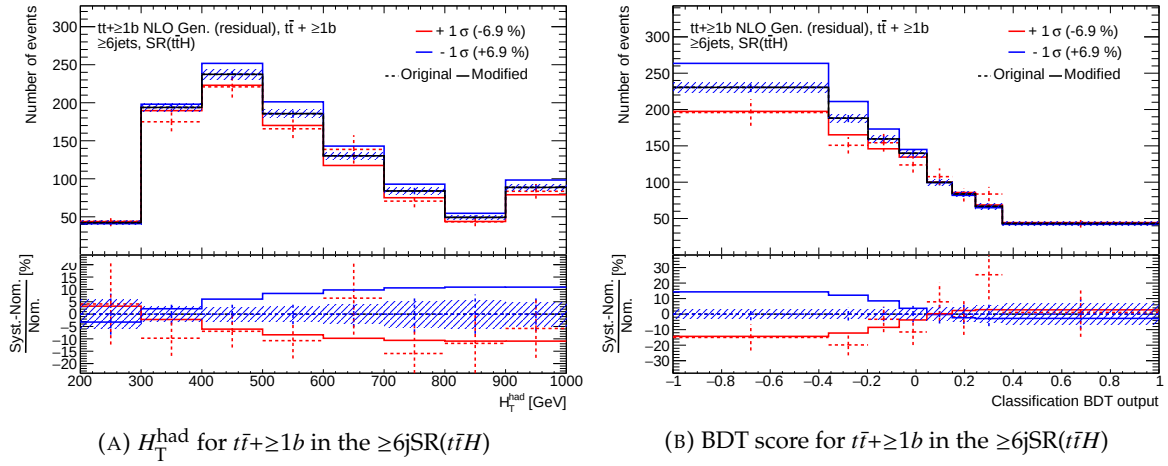


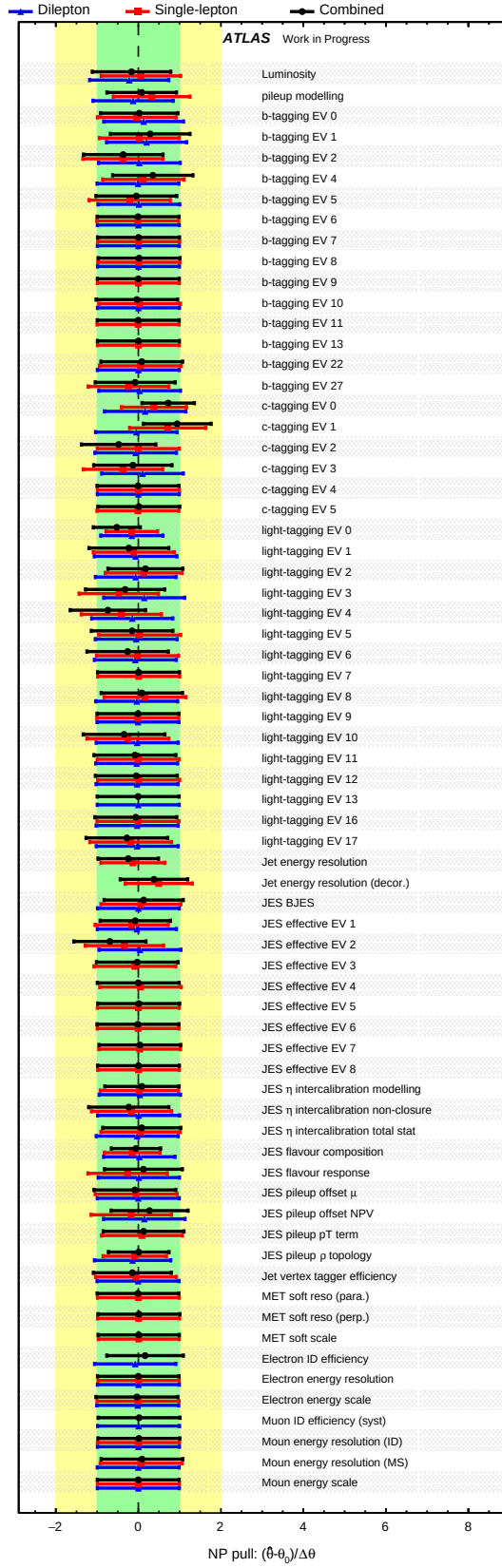
FIGURE C.1: Systematic shapes before and after smoothing. The systematic variation of $t\bar{t} + \geq 1b$ generator comparison (POWHEG+PYTHIA 8 vs. SHERPA5F). Dotted lines show the original variation before smoothing, and solid lines show the smoothed lines.

Appendix D

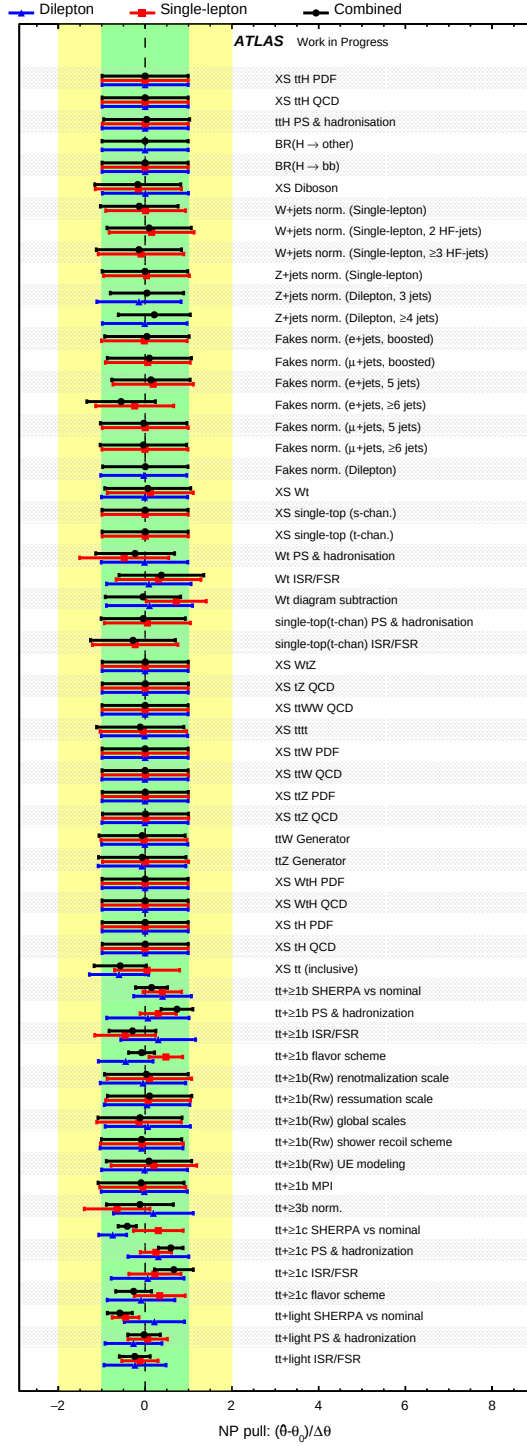
Detailed Fit Information

D.1 Full Systematic Variations

All of the systematic pulls and constraints are shown in Figure D.1 for the comparison between the individual fits and combined fit, and in Figure D.2 for the comparison between the two- μ combined and full combined fits.

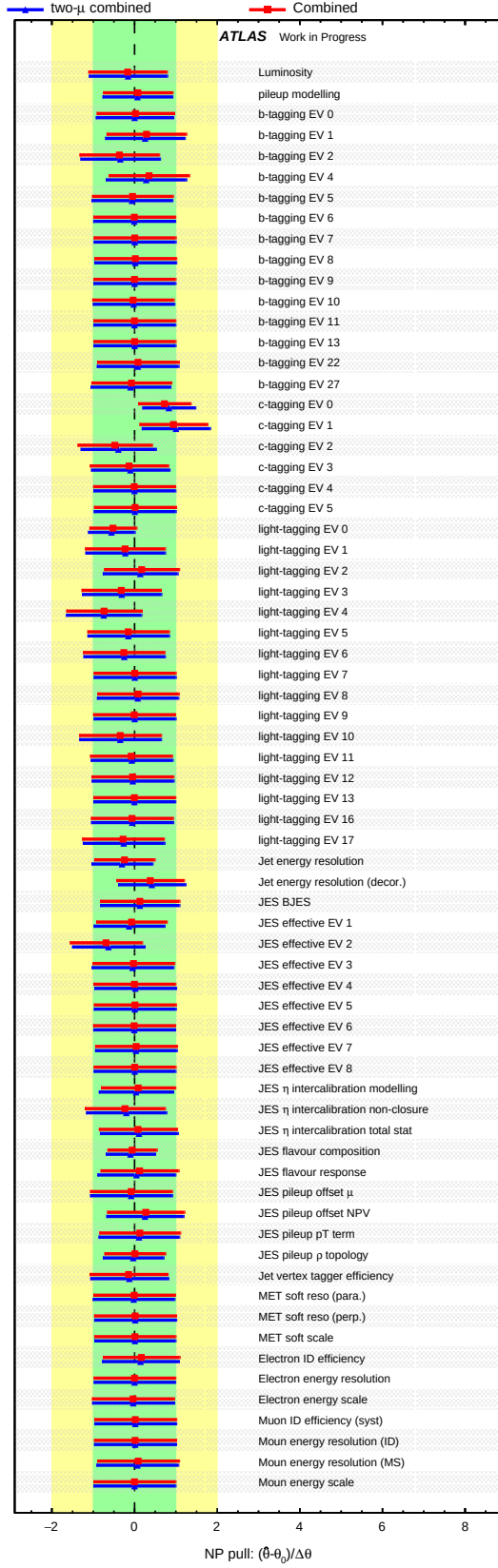


(A) detector systematics

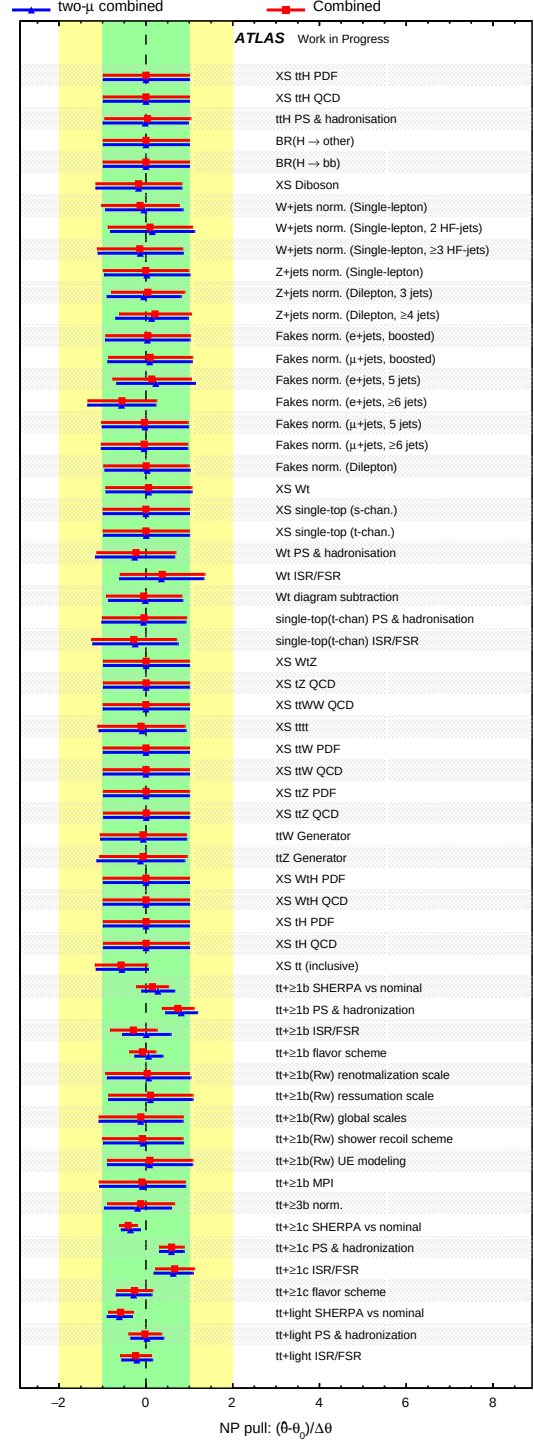


(B) modeling systematics

FIGURE D.1: Fit pulls and constraints for all systematic variations. Blue, red points denote the individual fits with dilepton and single-lepton, respectively. Black points show the combined fit.



(A) detector systematics



(B) modeling systematics

FIGURE D.2: Fit pulls and constraints for all systematic variations. Blue, red points denote the two- μ combined and full combined fits, respectively.

Bibliography

- [1] P. Achard et al., “Measurement of the running of the electromagnetic coupling at large momentum-transfer at LEP”, Phys. Lett. B623 (2005) 26.
- [2] The ATLAS collaboration, “Measurement of multi-jet cross-section ratios and determination of the strong coupling constant in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector.”, ATLAS-CONF-2013-041 (2013).
- [3] ATLAS Standard Model physics group, “Summary plots from the ATLAS Standard Model physics group”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/>.
- [4] C. Patrignani et al., “Review of Particle Physics”, Chin. Phys. C40 (2016) 100001.
- [5] Ansgar Denner et al., “Standard Model input parameters for Higgs physics”, LHCHXSWG-INT-2015-006 (2015).
- [6] The ATLAS and CMS Collaborations, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”, Phys. Rev. Lett. 114 (2015) 191803.
- [7] The ATLAS Collaboration, “Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb⁻¹ of pp collision data”, Phys. Rev. D90 (2014) 052004.
- [8] The ATLAS Collaboration, “Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s}=13$ TeV pp collisions using the ATLAS detector”, ATLAS-CONF-2017-046 (2017).
- [9] The ATLAS Collaboration, “Measurement of inclusive and differential cross sections in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel at 13 TeV with the ATLAS detector”, ATLAS-CONF-2017-032 (2017).
- [10] The ATLAS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel with 36.1 fb⁻¹ pp collision data at the center-of-mass energy of 13 TeV with the ATLAS detector”, ATLAS-CONF-2017-045 (2017).
- [11] The ATLAS Collaboration, “Combined measurements of Higgs boson production and decay in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels using $\sqrt{s} = 13$ TeV pp collision data collected with the ATLAS experiment”, ATLAS-CONF-2017-047 (2017).
- [12] Christopher T. Hill, “Topcolor assisted technicolor”, Phys. Lett. B345 (1995) 483.
- [13] Robert M. Harris, Christopher T. Hill, and Stephen J. Parke, “Cross-section for topcolor Z-prime(t) decaying to t anti-t: Version 2.6”, FERMILAB-FN-0687, CDF-ANAL-EXOTIC-PUBLIC-5064 (1999).
- [14] Robert M. Harris and Supriya Jain, “Cross Sections for Leptophobic Topcolor Z' Decaying to Top-Antitop”, Eur. Phys. J. C72 (2012) 2072.
- [15] Ben Lillie, Lisa Randall, and Lian-Tao Wang, “The Bulk RS KK-gluon at the LHC”, JHEP 09 (2007) 074.

- [16] Serguei Chatrchyan et al., “Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Lett. B* 728 (2014) 496.
- [17] The ATLAS Collaboration, “Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”, *Eur. Phys. J. C* 76 (2016) 538.
- [18] The ATLAS Collaboration, “Search for flavour-changing neutral current top quark decays $t \rightarrow qZ$ in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector”, ATLAS-CONF-2017-070 (2017).
- [19] LHC Top Physics Working Group, “LHCTopWG Summary Plots”, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>.
- [20] M. Baak et al., “The global electroweak fit at NNLO and prospects for the LHC and ILC”, *Eur. Phys. J. C* 74 (2014) 3046.
- [21] The ATLAS Collaboration, “Measurement of the W -boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, CERN-EP-2016-305 (2017).
- [22] Fedor Bezrukov and Mikhail Shaposhnikov, “Why should we care about the top quark Yukawa coupling?”, *J. Exp. Theor. Phys.* 120 (2015) 335.
- [23] A. V. Bednyakov et al., “Stability of the Electroweak Vacuum: Gauge Independence and Advanced Precision”, *Phys. Rev. Lett.* 115 (2015) 201802.
- [24] Gilles Couture et al., “Top and bottom partners, Higgs boson on the brane, and the $t\bar{t}h$ signal”, *Phys. Rev. D* 95 (2017) 095038.
- [25] The ATLAS Exotic physics group, “Summary plots from the ATLAS Exotic physics group”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/>.
- [26] ATLAS Collaboration, “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, CERN-EP-2017-291 (2017).
- [27] ATLAS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} pp collision data at the center-of-mass energy of 13 TeV with the ATLAS detector”, CERN-EP-2017-288 (2017).
- [28] ATLAS Collaboration, “Measurement of the Higgs boson coupling properties in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector”, CERN-EP-2017-206 (2017).
- [29] ATLAS Collaboration, “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector”, ATLAS-CONF-2017-077 (2017).
- [30] Lyndon Evans and Philip Bryant, “LHC Machine”, *JINST* 3 (2008) S08001.
- [31] Fabienne Marcastel, “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN”, OPEN-PHO-CHART-2013-001 (2013), General Photo.
- [32] Elias Métral, “LHC: status, prospects and future challenges”, *PoS LHCP2016* (2016) 002.
- [33] The ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* 3 (2008) S08003.
- [34] Joao Pequeno and Paul Schaffner, “An computer generated image representing how ATLAS detects particles”, CERN-EX-1301009 (2013).

- [35] The ATLAS Collaboration, “ATLAS muon spectrometer: Technical design report”, CERN-LHCC-97-22, ATLAS-TDR-10 (1997).
- [36] Karolos Potamianos, “The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider”, PoS EPS-HEP2015 (2015) 261.
- [37] Manuel Kayl, “Tracking Performance of the ATLAS Inner Detector and Observation of Known Hadrons”, Hadron collider physics. Proceedings, 22nd Conference, HCP 2010, Toronto, Canada, August 23-27, 2010 (2010).
- [38] The ATLAS Collaboration, “ATLAS inner detector: Technical Design Report, 1”, CERN-LHCC-97-016, ATLAS-TDR-4, Technical Design Report ATLAS ().
- [39] The ATLAS Collaboration, “ATLAS Insertable B-Layer Technical Design Report”, ATLAS-TDR-19 (2010).
- [40] The ATLAS Collaboration, “Studies of the performance of the ATLAS detector using cosmic-ray muons”, Eur. Phys. J. C71 (2011) 1593.
- [41] The ATLAS Tracking Performance Group, “Impact Parameter Resolution Using 2016 MB Data”, (2016), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-018/>.
- [42] The ATLAS Collaboration, “ATLAS pixel detector electronics and sensors”, JINST 3 (2008) P07007.
- [43] S. Altenheiner et al., “Planar slim-edge pixel sensors for the ATLAS upgrades”, JINST 7 (2012) C02051.
- [44] J Albert et al., “Prototype ATLAS IBL Modules using the FE-I4A Front-End Readout Chip”, JINST 7 (2012) P11010.
- [45] The ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, JINST 3 (2008) S08003.
- [46] Joao Pequeno, “Computer Generated image of the ATLAS calorimeter”, CERN-GE-0803015 (2008).
- [47] The ATLAS Collaboration, “Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons”, Eur. Phys. J. C70 (2010) 755.
- [48] Joao Pequeno, “Computer generated image of the ATLAS Muons subsystem”, CERN-GE-0803017 (2008).
- [49] The ATLAS Collaboration, “Performance of the ATLAS Trigger System in 2015”, Eur. Phys. J. C77 (2017) 317.
- [50] ATLAS Collaboration, “Data Quality Information for Data”, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010/>.
- [51] ATLAS Collaboration, “LuminosityPublicResultsRun2”, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2/>.
- [52] The ATLAS collaboration, “Electron and photon trigger efficiencies using the full 2016 ATLAS data”, (2017), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Electron_and_photon_trigger_AN2.
- [53] The ATLAS collaboration, “Muon trigger performances in 2016”, (2017), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonTriggerPublicResults#Muon_trigger_performances_in_AN1.

- [54] The ATLAS collaboration, “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data”, ATLAS-CONF-2016-024 (2016).
- [55] The ATLAS collaboration, “Electron reconstruction and identification efficiency measurements using an integrated luminosity of 33.9 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp collision data recorded in 2016.”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2017-003>.
- [56] The ATLAS collaboration, “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}=13 \text{ TeV}$ ”, Eur. Phys. J. C 76 (2016) 292. 45.
- [57] The ATLAS collaboration, “ATLAS Muon Combined Performance with the full 2016 dataset”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-001>.
- [58] W. Lampl et al., “Calorimeter Clustering Algorithms: Description and Performance”, ATL-LARG-PUB-2008-002 (2008).
- [59] M Cacciari, G Salam, and G Soyez, “The anti- k_t jet clustering algorithm”, JHEP 04 (2008) 063.
- [60] The ATLAS collaboration, “Pile-up subtraction and suppression for jets in ATLAS”, ATLAS-CONF-2013-083 (2013).
- [61] The ATLAS collaboration, “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”, (2017), submitted to Phys. Rev. D.,
- [62] The ATLAS collaboration, “Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS”, ATLAS-CONF-2015-037 (2015).
- [63] The ATLAS collaboration, “Pile-up subtraction and suppression for jets in ATLAS”, ATLAS-CONF-2013-083 (2013).
- [64] The ATLAS collaboration, “JES Public Plots for Moriond 2017”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2017-003/>.
- [65] W Lukas, “Fast Simulation for ATLAS: Atlfast-II and ISF”, ATL-SOFT-PROC-2012-065 (2012).
- [66] The ATLAS Collaboration, “Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector”, Eur. Phys. J. C 75 (2015) 17.
- [67] The ATLAS collaboration, “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ using the ATLAS detector”, Eur. Phys. J. C 76 (2015) 581. 39.
- [68] The ATLAS collaboration, “JVT Public Plots for ICHEP 2016”, (2017), URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-011>.
- [69] The ATLAS collaboration, “Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run”, ATL-PHYS-PUB-2016-012 (2016).
- [70] The ATLAS collaboration, “Expected performance of the ATLAS b -tagging algorithms in Run-2”, ATL-PHYS-PUB-2015-022 (2015).
- [71] The ATLAS collaboration, “Calibration of b -tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment”, ATLAS-CONF-2014-004 (2014).
- [72] The ATLAS collaboration, “Measurement of c -jet tagging efficiency in $t\bar{t}$ events using a likelihood approach”, ATLAS-COM-CONF-2017-083 (2017).

- [73] The ATLAS collaboration, “Calibration of the b -tagging efficiency for c jets with the ATLAS detector using events with a W boson produced in association with a single c quark”, ATLAS-CONF-2013-109 (2013).
- [74] The ATLAS collaboration, “Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data”, ATLAS-CONF-2014-046 (2014).
- [75] The ATLAS collaboration, “ b -tagging calibration plots using dileptonic $t\bar{t}$ events produced in pp collisions at $\sqrt{s} = 13$ TeV and a combinatorial likelihood approach”, (2017), URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2016-003/>.
- [76] Johannes Erdmann et al., “A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework”, Nucl. Instrum. Meth. A 748 (2014) 18.
- [77] The ATLAS collaboration, “ b -tagging calibration plots in light-flavour jets from 2015+2016 data using the negative-tag method”, (2017), URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-002/>.
- [78] The ATLAS collaboration, “Search for the Standard Model Higgs boson produced in association with top quarks and decaying to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC”, ATLAS-CONF-2014-011 (2014).
- [79] The ATLAS Collaboration, “Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data”, ATLAS-CONF-2014-046 (2014).
- [80] The ATLAS Collaboration, “The ATLAS Simulation Infrastructure”, Eur. Phys. J. C 70 (2010) 823.
- [81] GEANT4 Collaboration, S. Agostinelli et al., “GEANT4: A Simulation toolkit”, Nucl. Instrum. Meth. A 506 (2003) 250.
- [82] T. Gleisberg et al., “Event generation with SHERPA 1.1”, JHEP 02 (2009) 007.
- [83] John C. Collins, Davison E. Soper, and George F. Sterman, “Factorization of Hard Processes in QCD”, Adv. Ser. Direct. High Energy Phys. 5 (1989) 1.
- [84] Michiel Botje et al., “The PDF4LHC Working Group Interim Recommendations”, (2011), arXiv: 1101.0538 [hep-ph].
- [85] Richard D. Ball et al., “Parton distributions for the LHC Run II”, JHEP 04 (2015) 040.
- [86] Richard D. Ball et al., “Parton distributions from high-precision collider data”, Eur. Phys. J. C 77 (2017) 663.
- [87] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, JHEP 07 (2014) 079.
- [88] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, “A Brief Introduction to PYTHIA 8.1”, Comput. Phys. Commun. 178 (2008) 852.
- [89] The ATLAS Collaboration, “ATLAS Pythia 8 tunes to 7 TeV data”, ATL-PHYS-PUB-2014-021 (2014).
- [90] M. Bahr et al., “Herwig++ Physics and Manual”, Eur. Phys. J. C 58 (2008) 639.
- [91] Emanuele Re, “Single-top Wt -channel production matched with parton showers using the POWHEG method”, Eur. Phys. J. C 71 (2011) 1547.
- [92] Simone Alioli et al., “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, JHEP 0909 (2009) 111.
- [93] Paolo Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, JHEP 0411 (2004) 040.

- [94] Stefano Frixione, Paolo Nason, and Carlo Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, JHEP 0711 (2007) 070.
- [95] Simone Alioli et al., “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, JHEP 1006 (2010) 043.
- [96] John M. Campbell et al., “Top-pair production and decay at NLO matched with parton showers”, JHEP 04 (2015) 114.
- [97] Michal Czakon and Alexander Mitov, “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”, Comput. Phys. Commun. 185 (2014) 2930.
- [98] Matteo Cacciari et al., “Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation”, Phys. Lett. B 710 (2012) 612.
- [99] Michal Czakon and Alexander Mitov, “NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels”, JHEP 1212 (2012) 054.
- [100] Michal Czakon and Alexander Mitov, “NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction”, JHEP 1301 (2013) 080.
- [101] Michal Czakon, Paul Fiedler, and Alexander Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$ ”, Phys. Rev. Lett. 110 (2013) 252004.
- [102] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini, “Scattering Amplitudes with Open Loops”, Phys. Rev. Lett. 108 (2012) 111601.
- [103] Tomas Jezo et al., “New NLOPS predictions for $t\bar{t} + b$ -jet production at the LHC ”, (2018), arXiv: 1802.00426 [hep-ph].
- [104] Peter Zeiler Skands, “Tuning Monte Carlo Generators: The Perugia Tunes”, Phys. Rev. D 82 (2010) 074018.
- [105] Stefano Frixione et al., “Single-top hadroproduction in association with a W boson”, JHEP 0807 (2008) 029.
- [106] V. Barger, J. Ohnemus, and R. Phillips, “Event shape criteria for single lepton top signals”, Phys. Rev. D 48 (1993) 3953.
- [107] Catherine Bernaciak et al., “Fox-Wolfram Moments in Higgs Physics”, Phys. Rev. D 87 (2013) 073014.
- [108] Morad Aaboud et al., “Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC”, Eur. Phys. J. C 76 (2016) 653.
- [109] The ATLAS Collaboration, “Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC”, Phys. Rev. Lett. 117 (2016) 182002.
- [110] D. de Florian et al., “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”, FERMILAB-FN-1025-T, CERN-2017-002-M (2016).
- [111] Andreas Hocker et al., “TMVA - Toolkit for Multivariate Data Analysis”, PoS ACAT (2007) 040.
- [112] CERN, “3rd CERN School of Computing ”, (1974), URL: <https://cds.cern.ch/record/186223>.