

MEASUREMENT AND BEYOND THE STANDARD MODEL
INTERPRETATIONS OF THE $t\bar{t}t\bar{t}$ PRODUCTION CROSS
SECTION FROM 140fb^{-1} OF pp COLLISION DATA AT
 $\sqrt{s} = 13\text{TeV}$ IN ATLAS

by

Rachel Lindley

Copyright © Rachel Lindley 2023

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHYSICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2023



THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: *Rachel Lindley*, titled: *Measurement and Beyond the Standard Model interpretations of the $4t$ top production cross section from 140fb^{-1} of pp collision data at $\sqrt{s}=13\text{TeV}$ in ATLAS*

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Kenneth Johns

Kenneth Johns (Jul 20, 2023 21:39 PDT)

Date: Jul 20, 2023

Kenneth Johns

Elliott Cheu

Elliott Cheu

Date: Jul 21, 2023

John Rutherford

John Rutherford

Date: Jul 21, 2023

Shufang Su

Shufang Su (Jul 22, 2023 08:40 GMT+8)

Date: Jul 22, 2023

Erich W. Varnes

Erich W. Varnes (Jul 21, 2023 17:56 PDT)

Date: Jul 21, 2023

Erich Varnes

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Kenneth Johns

Kenneth Johns (Jul 20, 2023 21:39 PDT)

Date: Jul 20, 2023

Kenneth Johns

Dissertation Committee Chair

Physics

ACKNOWLEDGEMENTS

I would like to thank my graduate advisor, Ken Johns, for his support and guidance. The other faculty members of the experimental high-energy analysis group, Elliott Cheu and Erich Varnes, have also encouraged me and taught me about thinking critically in a particle physics context. Thanks to John Rutherford and Shufang Su for their time and consideration as part of my thesis committee. I would also like to thank the post-docs I have worked with: Simon Berlendis, Matt LeBlanc, Chris Delitzsch, and Jad Sardain. In particular, Simon provided me with an introduction to analysis and graciously answered many novice questions. I am grateful for the camaraderie of my fellow grad students: Billie Braimah and Hao Zhou, for being friendly office-mates and conference buddies; Michael Schott, Nick Lopez, and Zhaoyuan “Maxwell” Cui for boba-tea, conversations and laughter. Thanks to Michelle Solis and Casey Frantz for the friendships that somehow formed from the afternoons we spent scowling at oscilloscopes.

I have also received incredible support from outside of ATLAS. Thanks to Ian Marsh and Maia Wilenius (and Pico, of course), my Tucson family. Thanks to Will Lockhart and Heather Charles for supporting me at my defense even though it was early in the morning, and Brianna Billingsly who was also there, and who gave me cute math stickers once. Thanks to Kelly Kurtz and Leslie Ralph for helping me face myself, a task arguably more challenging than understanding the Higgs Mechanism. Thanks to everyone who immediately stamped “Doctor” on everything addressed to me after they learned I had defended my thesis. Thank you to my Fairbanks family— especially my sister, Emily, who has always believed in me no matter what.

I gratefully acknowledge the United States Department of Energy (DOE) for funding the work presented in this thesis under DOE grant number DE-SC0009913. Many thanks to the Fan family for sponsoring the Fanfare Travel Award, which allowed me to travel to CERN.

DEDICATION

For my family.

TABLE OF CONTENTS

LIST OF FIGURES	9
LIST OF TABLES	15
ABSTRACT	17
CHAPTER 1 Introduction	18
CHAPTER 2 Theoretical Background	21
2.1 The Standard Model	21
2.1.1 Overview	21
2.1.2 Gauge Invariance and Field Theories	21
2.1.3 Electroweak Theory	24
2.1.4 The Higgs Mechanism and Yukawa Coupling	25
2.1.5 QCD and the Strong Force	29
2.2 Top Quark Physics	30
2.3 Beyond the Standard Model	33
2.3.1 2HDM	35
2.3.2 Higgs Oblique Models	37
CHAPTER 3 Experimental Apparatus	41
3.1 Introduction	41
3.2 LHC	42
3.3 Experiments	42
3.4 The ATLAS Detector	44
3.4.1 The Geometry of the Detector	44
3.4.2 ATLAS Inner Detector (ID)	45
3.4.3 ATLAS Calorimeters	46
3.4.4 Muon Spectrometers	49
3.5 ATLAS Upgrades for the HL-LHC	52
3.5.1 sTGC	52
3.5.2 MicroMegas Detectors	52

TABLE OF CONTENTS – *Continued*

CHAPTER 4	Data and Monte Carlo Samples	55
4.1	Introduction	55
4.2	Data	55
4.2.1	Triggers	57
4.2.2	Data Processing	60
4.3	Monte Carlo	61
4.3.1	Matrix Element Calculation and PDFs	61
4.3.2	Parton showering	62
4.3.3	Samples and Alternative Samples	63
CHAPTER 5	Data Reconstruction	70
5.1	Introduction	70
5.2	Electrons	70
5.2.1	Reconstruction	70
5.2.2	Identification	72
5.2.3	Isolation	74
5.2.4	Other Criteria	75
5.3	Muons	75
5.3.1	Reconstruction	75
5.3.2	Identification	77
5.3.3	Isolation	77
5.3.4	Other Criteria	77
5.4	Small Radius Jets	78
5.4.1	Particle Flow Reconstruction	78
5.4.2	Jet Clustering	80
5.4.3	Jet Calibration	80
5.4.4	Other Jet Selections	81
5.5	b-Jets	83
5.6	Reclustered Jets	84
5.7	Overlap Removal	85
5.8	Missing Transverse Energy	87
CHAPTER 6	Systematic Uncertainties	89
6.1	Introduction	89
6.1.1	HBSM Analysis: Theoretical Uncertainties	89
6.1.2	Higgs Oblique Analysis: Theoretical Uncertainties	93

TABLE OF CONTENTS – *Continued*

6.2	Experimental Uncertainties	95
CHAPTER 7	Data-Driven Reweighting	100
7.1	HBSM Analysis	100
7.1.1	Introduction	100
7.1.2	HF Rescaling	100
7.1.3	NN Based Kinematic Reweighting	101
7.2	Higgs Oblique Analysis	103
7.2.1	QmisID	103
7.2.2	$t\bar{t}W$ Data-Driven Reweighting	107
CHAPTER 8	Signal Discrimination and MVA	111
8.1	Introduction	111
8.2	Region Definition	111
8.2.1	HBSM Analysis	111
8.2.2	Higgs Oblique Analysis	113
8.3	Multivariate Analysis	115
8.4	Machine Learning Algorithms	117
8.4.1	Graph Neural Network	117
8.4.2	Decision Trees	119
8.4.3	HBSM Analysis	121
8.4.4	Higgs Oblique Analysis	123
8.4.5	Quark-Gluon Tagging in HBSM	126
CHAPTER 9	TRExFitter and Profile Likelihood Fits	134
9.1	Introduction	134
9.2	Profile Likelihood Fitting	134
9.3	Setting Limits	139
9.4	TRExFitter	139
9.5	Preliminary Profile Likelihood Fit Results	141
9.5.1	HBSM Analysis	141
9.5.2	Higgs Oblique Analysis	142
CHAPTER 10	Results	157
10.1	HBSM Analysis	157
10.2	Higgs Oblique Analysis	158

TABLE OF CONTENTS – *Continued*

CHAPTER 11 Conclusion	172
Bibliography	175

LIST OF FIGURES

2.1	Branching ratios for top quark decay.	31
2.2	Comparison of measurements of the $t\bar{t}$ cross section to theoretical predictions, as a function of center-of-mass energy $\sqrt{s}$ [19].	32
2.3	Feynman diagrams associated with single top quark production[23].	32
2.4	Feynman diagrams for $t\bar{t}t\bar{t}$ production with QCD and EW interactions[24].	33
2.5	Branching ratios for $t\bar{t}t\bar{t}$ decay.	34
2.6	$t\bar{t}t\bar{t}$ production in association with a CP-even Heavy Higgs, H , or pseudoscalar Heavy Higgs, A	38
2.7	Cross sections of $t\bar{t}t\bar{t}$ production associated with Heavy Higgs H or pseudoscalar Heavy Higgs A . Different lines correspond to different values of the parameter $\tan\beta$	39
2.8	Higgs propagator affected by a virtual BSM particle. The addition to the self-interaction energy is expressed as $\Delta\mathcal{O}(p^2)$, where p is the momentum.	39
2.9	Impact of the $\hat{H}$ parameter on the $t\bar{t}t\bar{t}$ cross section. The cross section is calculated for 15 $\hat{H}$ values, using MadGraph5_aMC@NLO, and compared to the “Theory curve” given in Equation 2.37. Two cases are considered: one at $\sqrt{s} = 14\text{TeV}$ without Z/γ interactions, and one at $\sqrt{s} = 13\text{TeV}$ with Z/γ interactions.	40
3.1	Layout of the LHC complex, showing the various accelerators, the LHC itself, and the positions of the different experiments[35].	43
3.2	One of the modules of the barrel EM calorimeter, displaying the accordion geometry found in the EM barrel and endcap calorimeters [41]. The different dimensions and layers of the calorimeter are also labeled.	49
3.3	Layout of the muon spectrometer system on the ATLAS detector, showing the various types of muon detectors [43].	50
3.4	Schematic of the Micromegas design, displaying the path of an ionizing particle within the detector [46].	53

LIST OF FIGURES – *Continued*

6.1	Distribution of GNN score in the $\geq 10j4b$ region. The distribution is shown for MC samples with a 5FS $t\bar{t} + \geq 1b$ sample, and for MC samples with a 4FS $t\bar{t} + \geq 1b$ sample (shown as “data” points). The left plot shows the distributions without a flavor scheme uncertainty, and the right plot shows the same distribution with the flavor scheme uncertainty added. The discrepancy in the highest GNN score bins is still not completely covered by the extra uncertainty, but shows improvement.	92
6.2	Comparison between 5FS and 4FS expressed as a fit of 5FS MC samples to 4FS “data.” The large value of extracted $\mu_{t\bar{t}t\bar{t}}$ indicates that the discrepancy between 4FS and 5FS is being interpreted by the fit as a large $t\bar{t}t\bar{t}$ -like signal. Exact agreement between the two flavor schemes would result in $\mu_{t\bar{t}t\bar{t}} = 1$. It can be seen in the top plot (without FS uncertainties) that the value of $\mu_{t\bar{t}t\bar{t}}$ is not 1 even within uncertainties. The bottom plot incorporates FS uncertainties, so that the total uncertainty now encompasses the value of $\mu_{t\bar{t}t\bar{t}} = 1$	93
7.1	The extracted values of the three normalization factors for the flavor of the $t\bar{t}$ associated jet(s).	101
7.2	MC yields for each rescaling region, compared to data. The left plot shows the comparison before rescaling, and the right shows the same comparison after rescaling is applied.	102
7.3	Training datasets for the reweighting NN. The two classes the NN differentiates are MC and data.	104
7.4	The H_T distribution for 1L sample (left) and 2LOS sample (right). The data distribution is compared to the MC $t\bar{t} + jets$ distribution before and after kinematic reweighting.	104
7.5	The N_{jets} distribution for 1L sample before reweighting (left) and after (right).	105
7.6	The N_{jets} distribution for the 1b $t\bar{t}W$ reweighting regions. The data distribution is compared to the MC $t\bar{t}W + jets$ distribution after reweighting.	109
7.7	The N_{jets} distribution for the four $t\bar{t}W$ DD reweighting regions. The data distribution is compared to the MC $t\bar{t}W + jets$ distribution after reweighting.	110

LIST OF FIGURES – *Continued*

8.1	Visualization of region definitions based on jet and b-tagged jet multiplicity [116]. The 1L regions are shown on the left and the 2LOS shown on the right.	112
8.2	Distributions for two correlated variables, x_1 (left) and x_2 (right). The top row shows the distributions for Class A and Class B, and the bottom row shows the combined distribution. When each variable is considered individually, Class A and Class B are not easily separable.	116
8.3	Left: Three-dimensional plot of x_1 and x_2 , showing the distribution for Class A in gray and Class B in green. Right: Class A and Class B have narrow and separable distributions for x_1+x_2	116
8.4	Flow of information in a GNN [127]. The first step processes messages between each pair of elements, including relational information, while the second step aggregates that information. The third step uses a neural network to update the elements' values, and the final step calculates a single output value for each pp event.	119
8.5	BDT distributions for a testing/training sample for both signal $t\bar{t}\bar{t}\bar{t}$ and background $t\bar{t}+\text{jets}$. The distribution is shown for two different m_H mass points in the $1L \geq 9j \geq 3b$ region.	124
8.6	BDT distributions for all testing/training datasets for both signal $t\bar{t}\bar{t}\bar{t}$ and background $t\bar{t} + \text{jets}$. The distribution is shown for two different m_H mass points in the $1L \geq 9j \geq 3b$ region.	124
8.7	Comparison of GNN and BDT performance when classifying $t\bar{t}$ and $t\bar{t}\bar{t}\bar{t}$ events. The ROC curve for the GNN is shown in the left plot, and the ROC curve for the BDT is shown on the right. Note that curves are different because the axes display different efficiencies. For both, however, an ideal classifier will maximize the AUC.	126
8.8	A direct comparison of the sensitivity between the two MVA methods. The orange dotted line shows the mean sensitivity as a function of mass using the BDT. The dotted black line shows the mean sensitivity when using the GNN. The bands indicate the 1 and 2σ confidence intervals. There are also two solid lines showing curves for theoretical cross sections of the Heavy Higgs-associated $t\bar{t}\bar{t}\bar{t}$ process.	127
8.9	GNN distributions for each testing dataset i , compared to GNN distributions in datasets not equal to i . The $t\bar{t}\bar{t}\bar{t}$ LO sample is shown, which was used for the training of the GNN. This was done to avoid the negative weights present in the NLO sample.	128

LIST OF FIGURES – *Continued*

8.10	ROC curves for each testing dataset i as well as the combined result. The dashed line represents the efficiency of a random classifier.	129
8.11	Composition of jets for events with ≥ 3 b-jets, with the focus on percentage of b-jets as different cuts are applied.	130
8.12	Composition of jets for events with ≥ 3 b-jets, with the focus on percentage of gluon jets as different cuts are applied.	131
8.13	ROC curves for the BDT classifier, separating $t\bar{t}$ events from $t\bar{t}t\bar{t}$ events associated with Heavy Higgs of $m_H = 400\text{GeV}$. The AUC improves by about 10% when the q/g tagging variable is added.	132
8.14	AUC values for the GNN classifier, as a function of Heavy Higgs mass associated with the $t\bar{t}t\bar{t}$ signal. The improvement when adding the q/g tagging variable is small but consistent across mass points.	133
9.1	Visualization of p-value and significance as the difference between the mean of the hypothesis distribution and the measured value.	135
9.2	Pre-fit distributions for 1L channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot. Some plots, particularly those for high b-jet multiplicity, have a single bin with large uncertainty.	143
9.3	Pre-fit distributions for 2LOS channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.	144
9.4	Post-fit distributions for 1L channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.	145
9.5	Post-fit distributions for 2LOS channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.	146

LIST OF FIGURES – *Continued*

9.6	Nuisance parameter constraints after fitting to Asimov data.	147
9.7	Ranking of systematic uncertainties for 1LOS fit to Asimov data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit). . . .	148
9.8	Pre-fit distributions for 1L channel, associated with different signal and background samples, compared with blinded data yields.	149
9.9	Pre-fit distributions for 2LOS channel, associated with different signal and background samples, compared with blinded data yields.	150
9.10	Post-fit distributions for 1L channel, associated with different signal and background samples, compared with blinded data yields.	151
9.11	Post-fit distributions for 2LOS channel, associated with different signal and background samples, compared with blinded data yields. . . .	152
9.12	Nuisance parameter pulls and constraints after fitting to blinded data.	153
9.13	Quadratic fit to $t\bar{t}t\bar{t}$ yields as a function of $\hat{H}$ in each of the 17 bins of GNN in the SR.	154
9.14	Log likelihood plot showing the likelihood at each value of $\hat{H}$ relative to the likelihood maximum, which in this case is set up to be at $\hat{H} = 0$.	155
9.15	Post-fit distributions for SSML CR's and SR, associated with different signal and background samples.	156
10.1	Limits of HBSM analysis sensitivity as a function of Heavy Higgs mass, m_H . This limit plot is obtained from a fit to Asimov data. The left plot shows the comparison between using both 1L and 2LOS channels to using 1L alone. The right plot shows the comparison between using both 1L and 2LOS channels to using 2LOS alone. In both cases, the combined channels have greatest sensitivity.	157
10.2	Post-fit distribution of the SSML signal region. There is a clear discrepancy between the data yield and the background-only yield in the bins with highest GNN score.	159
10.3	Negative log likelihood as a function of $t\bar{t}t\bar{t}$ (x-axis) and $t\bar{t}t$ (y-axis). The darkest color indicates the smallest negative log likelihood value, and thus the best fit.	161

LIST OF FIGURES – *Continued*

10.4	Log likelihood scan associated with the SM $t\bar{t}t\bar{t}$ measurement. The observed scan has a minimum log likelihood at $\sigma_{t\bar{t}t\bar{t}} = 22.5\text{fb}$, where the background hypothesis is 6.1 standard deviations away.	162
10.5	Pre-fit distributions for SSML CR's and SR, associated with different signal and background samples.	163
10.6	Post-fit distributions for SSML CR's and SR, associated with different signal and background samples.	164
10.7	Normalization factors for background yields for SSML fit to unblinded data.	165
10.8	Correlation matrix for SSML fit to unblinded data.	166
10.9	Nuisance parameter pulls and constraints for SSML fit to unblinded data. The plot has been divided into six parts to enlarge text.	167
10.10	Ranking of systematic uncertainties for SSML fit to unblinded data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit).	168
10.11	Log likelihood scans as a function of $\bar{H}$ value. The dotted line represents the SM hypothesis of $\hat{H} = 0$ and the solid line represents the measured log likelihood scan.	169
10.12	Correlation matrix for Higgs Oblique fit to unblinded data.	170
10.13	Ranking of systematic uncertainties for Higgs Oblique fit to unblinded data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit).	171

LIST OF TABLES

2.1	List of isospin doublets, organized by generation	26
4.1	Single lepton triggers used for the HBSM $t\bar{t}t\bar{t}$ analysis.	57
4.2	Single lepton and dilepton triggers used for the Higgs Oblique $t\bar{t}t\bar{t}$ analysis.	58
4.3	Summary of MC samples for HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4). Table 1 of 2.	64
4.4	Summary of MC samples for HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4). Table 2 of 2.	64
4.5	Summary of alternative MC samples for the HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4).	65
4.6	Summary of nominal MC samples for the SSML Higgs Oblique analysis. Table 1 of 2.	67
4.7	Summary of nominal MC samples for the SSML Higgs Oblique analysis. Table 2 of 2.	68
4.8	Summary of alternative MC samples for the SSML Higgs Oblique analysis.	69
5.1	List of input variables for electron identification likelihood evaluation.	73
5.2	List of requirements for reconstructed tight electrons for HBSM $t\bar{t}t\bar{t}$ analysis.	75
5.3	List of requirements for reconstructed tight electrons for Higgs Oblique $t\bar{t}t\bar{t}$ analysis.	76
5.4	List of requirements for reconstructed tight muons in HBSM $t\bar{t}t\bar{t}$ analysis.	78
5.5	List of requirements for reconstructed tight muons in Higgs Oblique $t\bar{t}t\bar{t}$ analysis.	78

LIST OF TABLES – *Continued*

5.6	List of requirements for reconstructed small-radius jets in HBSM $t\bar{t}\bar{t}$ analysis.	82
5.7	List of requirements for reconstructed small-radius jets in Higgs Oblique $t\bar{t}\bar{t}$ analysis.	83
5.8	Overlap removal procedure.	85
5.9	List of requirements for reconstructed loose electrons in HBSM $t\bar{t}\bar{t}$ analysis.	86
5.10	List of requirements for reconstructed loose electrons in Higgs Oblique $t\bar{t}\bar{t}$ analysis.	86
5.11	List of requirements for reconstructed loose muons in HBSM $t\bar{t}\bar{t}$ analysis.	86
5.12	List of requirements for reconstructed loose muons in Higgs Oblique $t\bar{t}\bar{t}$ analysis.	87
6.1	List of uncertainties for reconstructed objects	99
7.1	Range of Z mass window and sidebands, in GeV.	106
7.2	The normalization factors derived from a fit to the CR1b(+) and the CR1b(-) regions.	109
8.1	Event selection requirements for the 3bL, 3bH, and 3bV regions. The b-jet tagging efficiencies are 60%, 70%, and 85%. Note that 3bL only requires ≤ 2 b-jets in an event for the tightest working point efficiency, but three for the 70% efficiency. This means that an event with no b-jets at the 60% efficiency but three b-jets at the 70% efficiency would be a 3bL event.	112
8.2	Descriptions of output values for the DFCAA tool.	113
8.3	Selections for each region of the SSML channels.	114
8.4	Input variables for HBSM BDT.	122
8.5	Input variables for HBSM GNN.	123
8.6	Input variables for Higgs Oblique GNN.	125

ABSTRACT

This dissertation presents the measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 13\text{TeV}$ at the ATLAS detector at CERN, with an integrated luminosity of 140fb^{-1} . Two independent analyses are conducted: one using the single-lepton and opposite-charge dilepton final states (1LOS), and the other using the same-charge dilepton and multilepton final states (SSML). The 1LOS analysis measures the $t\bar{t}t\bar{t}$ cross section in the context of a Two Higgs Doublet Model, with $t\bar{t}t\bar{t}$ cross section enhanced by contributions from $t\bar{t}H$ production, where H is a Heavy Higgs boson. The SSML analysis measures the $t\bar{t}t\bar{t}$ cross section and extracts limits on the Beyond the Standard Model (BSM) Higgs Oblique parameter, $\hat{H}$. Both analyses use signal regions with high particle jet and b-tagged jet multiplicity, as well as multivariate classifiers to isolate the $t\bar{t}t\bar{t}$ signal. Statistical analysis is performed using a profile likelihood method to obtain the cross section and the uncertainty on it, as well as the signal significance.

The SSML analysis measures $22.5_{-4.3}^{+4.7}(\text{stat})_{-3.4}^{+4.6}(\text{syst}) = 22.5_{-5.5}^{+6.6}\text{fb}$ for the $t\bar{t}t\bar{t}$ cross section, with an observed significance of 6.1 standard deviations above the background-only hypothesis, and a significance of 4.3 assuming a Standard Model cross section. This constitutes the first discovery of $t\bar{t}t\bar{t}$ production. The extracted $\hat{H}$ parameter is $\hat{H} = 0.09$ with the upper end of the 95% CL region at $\hat{H} = 0.2$, which is also the upper limit of the model's validity. The 1LOS analysis has not yet been unblinded, but fits to Asimov data place upper limits at the 95% CL on the $t\bar{t}t\bar{t}$ cross section in association with a Heavy Higgs of mass m_H . The limits range from $40 - 50\text{fb}$ for $m_H = 400\text{GeV}$ to $10 - 20\text{fb}$ for $m_H = 1000\text{GeV}$.

CHAPTER 1

Introduction

Particle physics seeks to understand the composition of the world, to identify the properties and interactions of the fundamental and indivisible constituents of matter called particles. The Standard Model of particle physics represents the current paradigm through which this field is understood. In the Standard Model, fermions interact with each other via the exchange of force-carrying gauge bosons, and gain mass through their interactions with the Higgs boson. This theory had perhaps its most dramatic triumph in 2012 with the discovery of the Higgs boson by the ATLAS and CMS experiments at the CERN Large Hadron Collider (LHC) [1, 2]. The Higgs boson was the last undiscovered particle predicted by the Standard Model, and its discovery represented not only the theoretical accuracy of the Standard Model's predictions, but also a century of technological development in designing particle colliders and detectors.

Today, the field of particle physics is no less challenging. While the Standard Model continues to dominate the understanding of small-scale particle interactions, the large-scale observations in cosmology demonstrate that the particles of the Standard Model comprise only about 1/6 of the total matter in the universe [3]. Moreover, the Standard Model cannot explain certain mysteries such as baryogenesis or the particular mass of the Higgs boson, discussed more in Chapter 2. The current consensus is that the Standard Model is incomplete, an effective field theory of a larger theory. It is the objective of particle physicists to rigorously test the predictions of the Standard Model to discover clues regarding this larger theory as well as search for particles that exist outside of the Standard Model.

The focus of this thesis is the measurement of the cross section for $t\bar{t}t\bar{t}$ production in pp collisions at $\sqrt{s} = 13\text{TeV}$. Equivalently this is an analysis on four top quark (or four top) production. This process is predicted to occur in the Standard Model with a cross section of $11.97^{+18\%}_{-21\%}\text{fb}$ [4]. For comparison, the production cross section for the Higgs boson at the same center-of-mass energy is $56 \pm 4\text{pb}$, more than 4,000 times larger [5].

The $t\bar{t}t\bar{t}$ process is a particularly good process to study in the context of certain Beyond the Standard Model (BSM) theories. Many BSM models affect Higgs boson interactions, and, as the most massive Standard Model particle, the top quark will be sensitive to these effects. The very small Standard Model $t\bar{t}t\bar{t}$ cross section means that even small BSM effects will be more apparent.

Evidence for the $t\bar{t}t\bar{t}$ process was previously measured by ATLAS and CMS in 2020. These results include the measured cross section, the observed significance for the measured cross section, and the expected significance assuming a Standard Model cross section. The ATLAS measurement was $24 \pm 5(\text{stat}) \text{ }^{+5}_{-4}(\text{syst}) = 24^{+7}_{-6}\text{fb}$ with an observed (expected) significance of 4.7 (2.6) standard deviations above the background-only hypothesis [6]. The CMS result was $17 \pm 4(\text{stat}) \pm 3(\text{syst})\text{fb}$ with an observed (expected) significance of 4.0 (3.2) standard deviations [7]. For both ATLAS and CMS searches, the measured cross section is larger than the Standard Model value. Although the discrepancy is not statistically significant, it adds motivation to conduct $t\bar{t}t\bar{t}$ cross section measurements using BSM interpretations. This thesis will cover two such interpretations, a two Higgs doublet model (2HDM, also referred to as HBSM) and a Higgs Oblique model [8, 9]. Chapter 2 will describe these models in greater detail.

Searches are conducted in several final state decay channels, grouped into two sets: one lepton/opposite sign (1LOS) and same-sign/multilepton (SSML). The 1LOS analysis searches for $t\bar{t}t\bar{t}$ production with decays into states with a single lepton or two oppositely-charged leptons. These channels are used in the HBSM analysis. The SSML analysis searches for $t\bar{t}t\bar{t}$ production with decays into states with two same-charge leptons or more than two leptons. These channels are used in the Higgs Oblique analysis. Although distinct, both the 1LOS and SSML channels are useful for the search for $t\bar{t}t\bar{t}$ production, and the decision to use one particular set of channels for each analysis is somewhat arbitrary. Indeed, the SSML channels have been used in a parallel HBSM analysis [10] outside the scope of this thesis.

For data collection, experimental physicists have a formidable array of tools to work with. The Large Hadron Collider is the world’s largest particle collider, accelerating protons to within $(1 - 1 \times 10^{-8})c$. The resulting proton-proton collisions have an energy of 13TeV and are able to produce every Standard Model particle. A subset of the collisions are recorded by particle detectors such as the ATLAS detector, a multi-purpose detector with layered tracking systems, calorimeters, and muon spectrometers. The LHC and the ATLAS detector are described in more detail in Chapter 3. The collected data, as well as simulated collision events, are described

in Chapter 4. The data and simulated samples are reconstructed into physics objects using a number of selections and algorithms, described in Chapter 5. There are also systematic uncertainties associated with both data and simulated samples, and these are described in Chapter 6. Some simulated samples are reweighted to better match data. This process is described in Chapter 7.

The two analyses exploit different advantages and confront different challenges, discussed in more depth in Chapter 8. The search itself is performed using machine learning to isolate the $t\bar{t}t\bar{t}$ signal from background processes, and then profile likelihood fitting to determine if the signal excess above the background hypothesis is statistically significant. Chapter 9 describes the statistical reasoning behind profile likelihood fitting as well as the tool used to implement it, TRExFitter. Lastly, the results are given in Chapter 10 and conclusions are presented in Chapter 11.

CHAPTER 2

Theoretical Background

2.1 The Standard Model

2.1.1 Overview

Particle physics seeks to discover and understand elementary particles and their interactions. By elementary it is meant that the particles cannot be further divided into smaller components. By the 1970's, the Standard Model of particle physics was fully developed. Although it has limitations, this theory provides accurate means of calculating quantities such as the probability of interaction, or cross section, describes the ways in which particles can be produced and decay, and provides a mechanism through which the particles acquire a mass.

The Standard Model contains 17 elementary particles and anti-particles: six quarks, three charged leptons, three neutral leptons (neutrinos), four gauge bosons, and one scalar boson. These particles interact via three forces: the strong force, the weak force, and the electromagnetic force. Although the fundamental particles have a mass, gravity is not currently incorporated into the Standard Model. Given that the other three forces are much stronger, it is neglected in this discussion.

The quarks interact via all three forces, the charged leptons interact via the weak and electromagnetic force, and the neutrinos interact only via the weak force. The bosons are interpreted as “force-carrying” particles, where the strong force is conveyed via gluon exchange, the electromagnetic through photon exchange, and the weak through the exchange of W and Z bosons. The last particle, the Higgs boson, represents the excitation of the Higgs field. The quarks, leptons, and other bosons interact with the Higgs field, and it is this mechanism that generates particle mass.

2.1.2 Gauge Invariance and Field Theories

The Standard Model is a field theory, meaning that the equations of motion describe continuous fields rather than discrete particles. There are a few reasons to use fields rather than particles: to maintain consistency in a relativistic framework, and

because the uncertainty principle implies that, for an arbitrarily small time scale, an arbitrary number of particles can be spontaneously created [11]. Particles are then interpreted as excitations of a field.

The equations of motion must also follow certain requirements. They must be Lorentz-invariant in order to satisfy the principle that the laws of physics are the same in every relativistic frame. Maxwell's equations follow from this invariance condition:

$$\partial^\mu F_{\mu\nu} = 0 \quad (2.1)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor. The Lagrangian for the free electromagnetic field is written as:

$$\mathcal{L}_{EM} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.2)$$

The Lagrangian for a field $\psi(x)$ corresponding to a fermion of mass m is written as:

$$\mathcal{L}_F = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) \quad (2.3)$$

Within electromagnetic interactions, the charge and current are conserved. Noether's Theorem connects all conservation laws with a symmetry in the equations of motion that describe them. In the case of electromagnetism, the symmetry is in the invariance of the electromagnetic field under a transformation of the electromagnetic potential, A_μ . The potential can vary by some scalar function f such that:

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}f \quad (2.4)$$

and the field will remain the same. The particular choice of $\vec{A}$ is called a gauge, and the field invariance is therefore a gauge invariance. If f is not dependent on spatial coordinates $\vec{x}$, then the gauge invariance is *globally* invariant.

The global gauge invariance of the potential can now be written as an invariance of the field under a U(1) transformation:

$$\psi(x) \rightarrow e^{i\alpha}\psi(x) \quad (2.5)$$

The invariance can be seen in the fermion Lagrangian because

$$\partial_\mu \psi'(x) = \partial_\mu e^{i\alpha}\psi(x) = e^{i\alpha}\partial_\mu \psi(x) \quad (2.6)$$

and so

$$\mathcal{L}'_F = \bar{\psi}'(x)(i\gamma^\mu\partial_\mu - m)\psi'(x) = e^{-i\alpha}\bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)e^{i\alpha}\psi(x) = \mathcal{L}_F \quad (2.7)$$

However, the invariance of the fermion Lagrangian does not hold for a *local* transformation, i.e. if α varies as a function of x . In the case of a local transformation, the differentiation of the field leads to

$$\partial_\mu \psi'(x) = \partial_\mu e^{i\alpha(\vec{x})}\psi = e^{i\alpha}\partial_\mu \psi(x) + \mathbf{i}e^{i\alpha(\vec{x})}\psi\partial_\mu(\alpha(\vec{x})) \quad (2.8)$$

where the term in bold will not cancel out. In order to create a Lagrangian with local gauge invariance, the differentiation term must be redefined to include a gauge field, A_μ :

$$D_\mu \equiv \partial_\mu - ieA_\mu \quad (2.9)$$

where the Lagrangian is invariant under the transformation

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e}\partial_\mu(\alpha(\vec{x})) \quad (2.10)$$

This essentially couples the fermion Lagrangian to the Lagrangian for the electromagnetic field. The new gauge field becomes the photon, and the combined Lagrangian becomes

$$\mathcal{L}_{EM+F} = \bar{\psi}(x)(i\gamma^\mu(\partial_\mu - ieA_\mu) - m)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.11)$$

In Equation 2.11 can be seen the kinetic term describing motion of the fermion; the interaction term, or the coupling of the fermion to a photon; and lastly the mass term. All terms are operating on the fermion field. This conceptualization of the electromagnetic interaction is known as Quantum Electrodynamics (QED).

2.1.3 Electroweak Theory

The same principle of creating a Lagrangian with Lorentz invariance and local gauge invariance is also used in describing the weak force. In fact, the weak force is combined with the electromagnetic force, and the combined electroweak force observes $SU(2)_L \times U(1)_Y$ invariance. The Y refers to hypercharge and the L refers to left-handedness, both discussed below. The $SU(2)$ generators are $I_1, I_2, I_3 = \frac{1}{2}\sigma_1, \frac{1}{2}\sigma_2, \frac{1}{2}\sigma_3$, where σ_i represent the Pauli matrices. The local gauge transformation by an amount α is:

$$\psi(x) \rightarrow \psi'(x) = e^{-i(\vec{\alpha} \cdot \vec{I})}\psi(x) \quad (2.12)$$

The $U(1)_Y$ transformation is similar to that for QED, but the generator is now the ‘hypercharge’ quantity, Y , instead of electric charge, Q . The hypercharge is expressed as $Y = 2(Q - I_3)$, and thus is connected to both the electric force and weak force generators.

The $SU(2)_L \times U(1)_Y$ local gauge invariance requirement leads to four massless gauge fields: A^1, A^2, A^3 , and B^0 , similar to the field A_μ that was introduced in the QED theory above. These fields manifest as the gauge bosons for the electroweak force: W^+ , W^- , Z , and γ .

$$\begin{aligned} W^\pm &= \frac{1}{\sqrt{2}}(A^1 \mp iA^2) \\ Z^0 &= \cos(\theta_w)A^3 - \sin(\theta_w)B^0 \\ \gamma &= \sin(\theta_w)A^3 + \cos(\theta_w)B^0 \end{aligned} \quad (2.13)$$

The angle θ_w is the weak mixing angle. Without the Higgs mechanism, these particles would remain massless to ensure local gauge invariance, contradicting the short-range nature of the weak force. However, the gauge fields couple to the Higgs field, which has a potential with nonzero minima. The Higgs mechanism breaks the symmetry of the electromagnetic and weak forces, gives mass to the $W^\pm$ and the Z bosons, and leaves the photon massless. More details are given below in Section 2.1.4.

Several terms in the SM Lagrangian dictate how fermions interact via the weak force. The fermion fields can be written as:

$$\psi_R = \frac{1}{2}(1 + \gamma^5)\psi, \psi_L = \frac{1}{2}(1 - \gamma^5)\psi \quad (2.14)$$

where the subscripts R and L denote handedness. Left-handed fermions travel in a direction anti-parallel to their spin; right-handed fermions travel in a direction parallel to their spin. Handedness is a relative quantity, as the momentum of a fermion relative to one observer might be different from the momentum relative to another observer. Handedness is also called ‘helicity.’ This is distinct from chirality, which is the actual quantity governing weak interactions, but helicity presents a way to conceptually understand chirality, and the two quantities are approximately identical at relativistic energies.

The weak force isospin current interacts only with left-handed fermions (or right-handed anti-fermions). This means that left-handed fermions form doublets with the other fermion of the same generation. For example, left-handed up-quarks and down-quarks form a doublet (see Table 2.1). There are three generations, and so three doublets for quarks and three doublets for leptons. Right-handed fermions do not interact via the weak force, and form singlets. As neutrinos only interact via the weak force, there are therefore no right-handed neutrinos.

2.1.4 The Higgs Mechanism and Yukawa Coupling

Without the Higgs field, the Standard Model presents a picture of massless gauge bosons and fermions. This Lagrangian satisfies local gauge invariance, but requires that all particles remain massless or else violate that invariance. The Higgs mechanism combines local gauge invariance with spontaneous symmetry breaking. In this process, some gauge bosons acquire mass and others remain massless. The concept of spontaneous symmetry breaking and the Higgs mechanism was introduced by Brout,

Isospin Doublets			
Generation	1	2	3
Quarks	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$
Leptons	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$

Table 2.1: List of isospin doublets, organized by generation

Englert, Guralnik, Hagen, Higgs, and Kibble via several papers in 1964 [12], [13], [14].

The principle of spontaneous symmetry breaking is to consider a potential V that varies as a function of the scalar field ϕ such that:

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - V(\phi) \quad (2.15)$$

and

$$V(\phi) = -\mu^2\phi\bar{\phi} + \frac{\lambda}{2}(\phi\bar{\phi})^2 \quad (2.16)$$

This potential has two nonzero minima. These are the “vacuum expectation values” or vev [11]. The field $\phi(x)$ can be chosen to exist near one vev or another, but once the choice is made, the Lagrangian is no longer symmetric for the transformation $\phi \rightarrow -\phi$. Because it is the lowest energy configuration for ϕ , this potential generates *spontaneously* broken symmetry.

When this case is generalized to a continuous rotational symmetry in N dimensions for N scalar fields, the result is the generation of a set of $N - 1$ massless fields, called “Goldstone bosons.” In fact, Goldstone’s theorem states that a massless particle is created for each spontaneously broken continuous symmetry.

The masses arise when the local gauge symmetry and spontaneous symmetry breaking are considered together. Local gauge symmetry demands the use of a covariant derivative. For $SU(2) \times U(1)$ local symmetry, this derivative is:

$$D_\mu = \partial_\mu - igA_\mu^a I^a - i\frac{1}{2}g'B_\mu \quad (2.17)$$

The $SU(2)$ gauge bosons are A_μ^a with a coupling constant of g , and the $U(1)$ gauge boson is B^μ with a coupling constant of g' . The Higgs field is a complex doublet scalar field with $SU(2) \times U(1)$ symmetry, invariant under the transformation

$$\phi \rightarrow \phi' = e^{\vec{\alpha} \cdot \vec{I}} e^{iY} \phi \quad (2.18)$$

Here, the generators are the isospin and hypercharge discussed in Section 2.1.3. Consider a potential such that the field ϕ has a vev of:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.19)$$

When this field interacts with the gauge bosons in the covariant derivative, the Lagrangian becomes

$$\mathcal{L} = \bar{\phi} D^\mu D_\mu \phi \quad (2.20)$$

or, focusing on the gauge boson terms $\Delta\mathcal{L} = \mathcal{L} - \partial_\mu$ terms:

$$\Delta\mathcal{L} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} (gA_\mu^a I^a + \frac{1}{2}g'B^\mu) (gA^{b\mu} I^b + \frac{1}{2}g'B^\mu) \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.21)$$

which becomes

$$\Delta\mathcal{L} = \frac{1}{2} \frac{v^2}{4} [g^2(A_\mu^1)^2 + g^2(A_\mu^2)^2 + (-gA_\mu^3 + g'B_\mu)^2] \quad (2.22)$$

In order to get the form $\Delta\mathcal{L} = -m^2|\phi^2|$, the masses and fields must be:

$$\begin{aligned}
W^\pm &= \frac{1}{\sqrt{2}}(A^1 \mp iA^2), m_W = g\frac{v}{2} \\
Z^0 &= \frac{g}{\sqrt{g^2 + g'^2}}A^3 - \frac{g'}{\sqrt{g^2 + g'^2}}B^0, m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \\
\gamma &= \frac{g'}{\sqrt{g^2 + g'^2}}A^3 + \frac{g}{\sqrt{g^2 + g'^2}}B^0, m_\gamma = 0
\end{aligned} \tag{2.23}$$

The weak mixing angle can be defined as:

$$\begin{aligned}
\cos \theta_W &= \frac{g}{\sqrt{g^2 + g'^2}} \\
\sin \theta_W &= \frac{g'}{\sqrt{g^2 + g'^2}}
\end{aligned} \tag{2.24}$$

The real gauge bosons and their masses are now identified.

The fermions have left- and right-handed components with different gauge quantum numbers, and, without interaction with the Higgs boson, cannot have mass terms that preserve gauge invariance. For fermions, the interaction with the scalar Higgs field is a type of Yukawa interaction, and can be described by the Lagrangian of a fermion field (such as that of an electron, where the left-handed doublet is the electron and its neutrino) coupling to the scalar Higgs field ϕ . This is written as a function of the left-handed doublet, L_L , the Higgs field, the right-handed electron e_R , and the Hermitian conjugate:

$$\mathcal{L}_{Yukawa} = -g_e(\bar{L}_L\phi e_R + \bar{e}_R\phi^\dagger L_L) \tag{2.25}$$

Here, g_e is the strength of the interaction. When the symmetry of the Higgs field is broken, the electron gains a mass proportional to the interaction strength and the vacuum expectation value of the Higgs potential, v .

$$\begin{aligned}
\mathcal{L}_{Yukawa} &= -g_e v(\bar{e}_L e_R + \bar{e}_R e_L) - g_e(\bar{e}_L e_R + \bar{e}_R e_L)H \\
&= -m_e(\bar{e}e) - g_e(\bar{e}e)H
\end{aligned}
\tag{2.26}$$

The first term describes the mass of the fermion, where $m_f = g_e v$, and the second term is the interaction with the Higgs particle H . The Yukawa interaction provides a mechanism for fermions to acquire mass, but does not predict the values of the masses.

2.1.5 QCD and the Strong Force

The strong force is described by the field theory of quantum chromodynamics (QCD), so named because the relevant conserved quantity is ‘color’ charge with three possible colors: red, blue, and green. The strong force interacts only between colored quarks and gluons, the massless gluon being the force-carrying boson. The gluon has a nonzero color charge, which causes phenomena such as asymptotic freedom and hadronization.

QCD arises from $SU(3)_c$ local gauge invariance, wherein the equations of motion remain unchanged under a transformation:

$$\psi(x) \rightarrow e^{-i(\vec{\alpha} \cdot \vec{T})} \psi(x) \tag{2.27}$$

Just as the transformation in $SU(2)$ space affects the isospin of weakly interacting particles, the transformation in $SU(3)$ space affects the color charge. The generators T_a are related to the 3×3 Gell-Mann matrices: $T_a = \frac{1}{2}\lambda_a$, which are analogous to the 2×2 Pauli matrices of $SU(2)$. Eight gauge bosons, or gluons, arise from the generators. As with the other field theories, the QCD Lagrangian contains terms for the kinetic energy and mass of the fermion field, as well as for the free gluon field:

$$\mathcal{L}_{QCD} = \bar{\psi}(x)(i\gamma^\mu(D_\mu - m\partial))\psi(x) - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \tag{2.28}$$

The strong force binds quarks and gluons together in a particular way such that, at small distances, the partons experience virtually no force. This phenomenon

is called “asymptotic freedom.” For all forces, the coupling strength changes as a function of energy (or, equivalently, interaction distance)—this is called the “running” of the coupling. For the electromagnetic force, for example, the coupling tends towards infinity at high energies. The charge at short distances is sufficient to polarize the vacuum. Any electron-positron pairs tend to align themselves with the field in the vicinity of the point charge, and their intervention effectively screens the point charge, causing its coupling value to drop to a smaller, constant value except at very short distances. For the strong force, a different phenomenon occurs. The coupling tends towards zero, and the vacuum screening is performed by gluons as well as quarks. What results is asymptotic freedom.

If the distance between two partons is increased, they begin to experience a binding force. This force does not decrease with distance and, at some point, the energy in the color field between the two partons is sufficient to spontaneously create two new partons. This process is called “hadronization.” The pairs and trios of bound quarks and gluons are known as hadrons: baryons contain three valence quarks, and mesons contain a quark and anti-quark pair. Unlike partons, hadrons can exist as free particles because they have net zero color charge.

Because the strong force is so strong for large distances (and, equivalently, low energies), predictions for low-energy QCD interactions cannot be solved analytically. Instead, the strong force is modeled phenomenologically using models such as the Lund string model[15] or lattice QCD [16]. Bound partons are modeled by empirical parton distribution functions (PDF’s), discussed more in Chapter 4.

2.2 Top Quark Physics

With a mass of $172.69 \pm 0.30\text{GeV}$, the top quark is the most massive of the quarks and also of the SM particles [5]. It is an up-type quark with a fractional electric charge of $+\frac{2}{3}$. With a lifetime on the order of 10^{-25} seconds, the top quark decays before it hadronizes.

The top quark decays via the electroweak force, with a branching ratio of nearly 100%, into a W^+ boson and b-quark. From there, the b-quark decays into a collimated shower of particles (a jet), and the W^+ boson can decay to either a pair of hadrons or a pair of leptons, as seen in Figure 2.1.

The most common mode of production for top quarks in pp collisions is the production of a $t\bar{t}$ pair via the strong force. Including up to NNLO QCD contri-

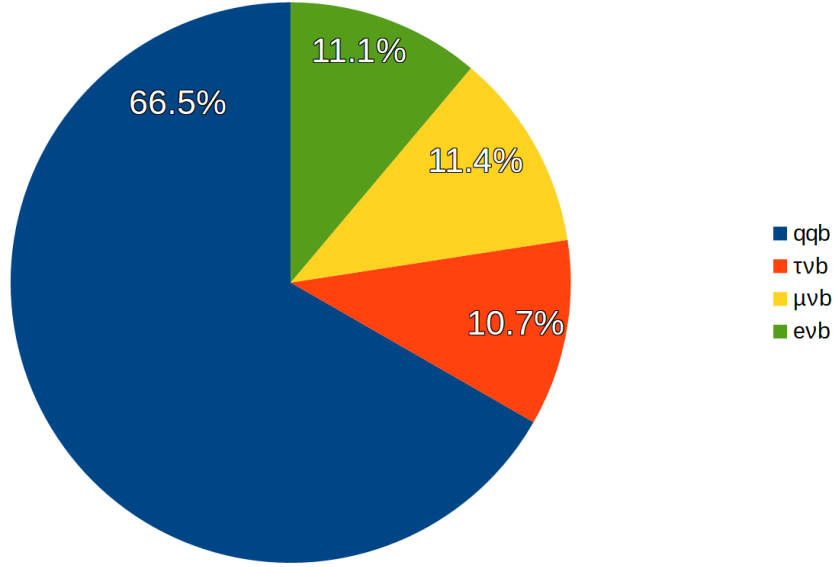


Figure 2.1: Branching ratios for top quark decay.

butions, the production cross section is calculated to be $831.8^{+19.8+35.1}_{-29.2-35.1}$ pb for a pp collision at energy $\sqrt{s} = 13\text{TeV}$ [5, 17]. At this high energy, 90% of production occurs via gluon-gluon fusion, with the remaining production coming from quark collisions. A recent measurement of the $t\bar{t}$ cross section by ATLAS agrees within uncertainties, with a measurement of $\sigma_{t\bar{t}} = 830 \pm 0.4$ (stat.) ± 36 (syst.) ± 14 (lumi.)pb [18]. Generally, measurements from colliders have agreed with the theory predictions for the $t\bar{t}$ cross section, as shown in Figure 2.2.

Single top quarks can be created via the weak force, involving the interactions of a W boson and a b-quark as seen in Figure 2.3. At 13TeV, the production cross section for single top quarks is lower than for $t\bar{t}$: the cross section for the t-channel is 215pb, approximately 10pb for the s-channel, and 72pb for the Wt channel[20–22].

The focus of this analysis is on the production of $t\bar{t}t\bar{t}$, which has an even smaller cross section. In the SM, the $t\bar{t}t\bar{t}$ process occurs through QCD interactions, as seen in Figure 2.4, or alternatively in association with the production of an off-shell electroweak Z or H boson.

The current SM estimate of the cross section is calculated by incorporating QCD interactions and EW interactions up to the next-to-leading order (NLO). Generally the EW interactions, with a smaller coupling constant, have a smaller impact

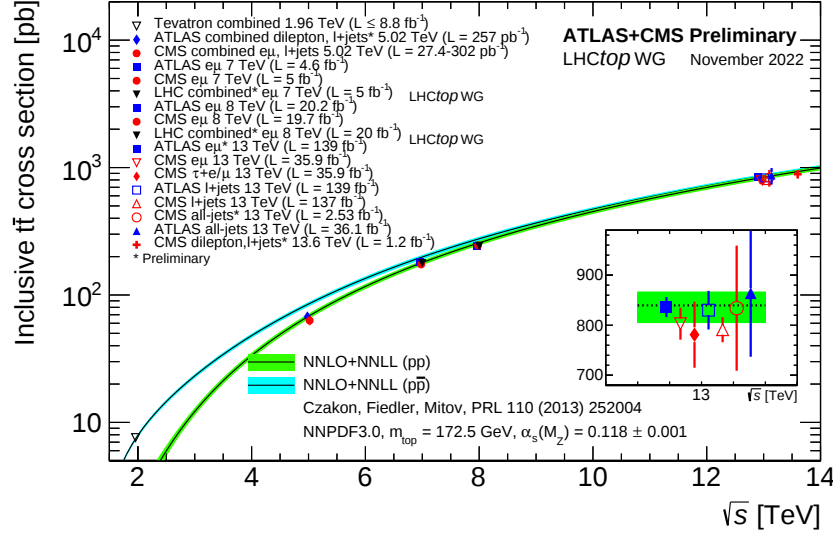


Figure 2.2: Comparison of measurements of the $t\bar{t}$ cross section to theoretical predictions, as a function of center-of-mass energy $\sqrt{s}$ [19].

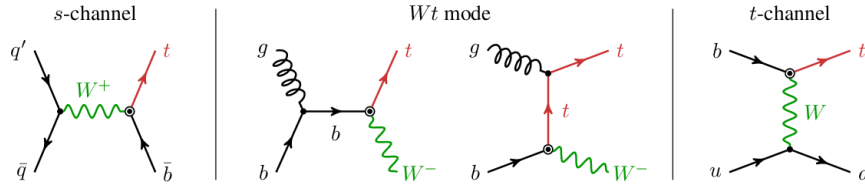


Figure 2.3: Feynman diagrams associated with single top quark production[23].

on the cross section than the QCD interactions. For SM $t\bar{t}t\bar{t}$ including both QCD and EW NLO corrections, the cross section is calculated at $\sqrt{s} = 13\text{TeV}$ to be $11.97^{+18\%}_{-21\%}\text{fb}$ [4].

The $t\bar{t}t\bar{t}$ process has several possible final states depending on the decay of the four W bosons, with the different final state branching ratios shown in Figure 2.5. The Higgs Oblique analysis uses the same-sign dilepton and multilepton decay channels (SSML) and the 2HDM analysis uses the single lepton and opposite-sign dilepton channels (1LOS). The two sets of decay channels, SSML and 1LOS, have different characteristics and challenges to overcome. The SSML channels have a much smaller branching fraction, and therefore lower $t\bar{t}t\bar{t}$ yield, but also have much lower levels of background contamination. The main background for the SSML channels

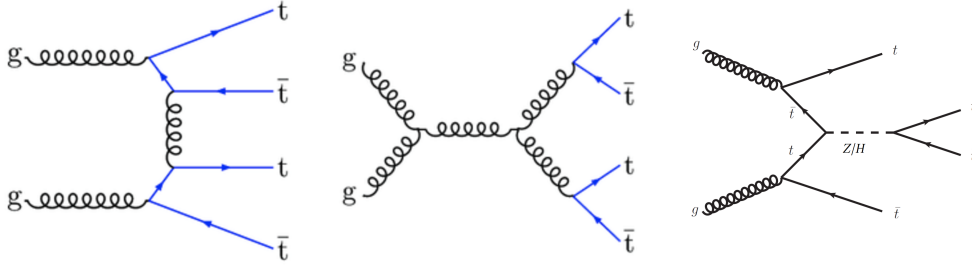


Figure 2.4: Feynman diagrams for $t\bar{t}t\bar{t}$ production with QCD and EW interactions[24].

comes from $t\bar{t}$ production with an associated Higgs, Z , or W . The 1LOS channels constitute a majority of the $t\bar{t}t\bar{t}$ decays, but have a large irreducible background of $t\bar{t}$ + jets events.

ATLAS made the first measurement of evidence of $t\bar{t}t\bar{t}$ production in 2020, with a cross section measurement of $\sigma_{t\bar{t}t\bar{t}} = 24_{-6}^{+7}\text{fb}$ [24]. The significance of the measurement was observed to be 4.3 standard deviations above the background prediction. The *expected* significance of the measurement refers to the sensitivity assuming a signal strength equal to the Standard Model prediction. The expected significance for the 2020 ATLAS measurement was 2.4 standard deviations above the background prediction. This measurement was obtained using the SSML channels. Later, the cross section in the 1LOS channel was measured and combined with the SSML result to give an observed (expected) significance of 4.7 (2.6) standard deviations [6].

Recently, a re-optimization of the SSML analysis produced the first discovery measurement of $\sigma_{t\bar{t}t\bar{t}} = 22.5_{-5.5}^{+6.6}\text{fb}$ with an observed (expected) significance of 6.1 (4.3) standard deviations [25]. The re-optimized version of this analysis is described throughout this thesis, with an additional focus on the Beyond the Standard Model (BSM) interpretations described below.

2.3 Beyond the Standard Model

The combination of electroweak theory and QCD into the Standard Model of particle physics has been highly successful, as the Standard Model predicts the behavior of small-scale particle interactions with extreme accuracy. Measurements of the electron and positron magnetic moments agree with the SM to more than ten decimal places [26]. However, there are phenomena that the Standard Model does not explain, and so new Beyond the Standard Model (BSM) theories are sought to address these

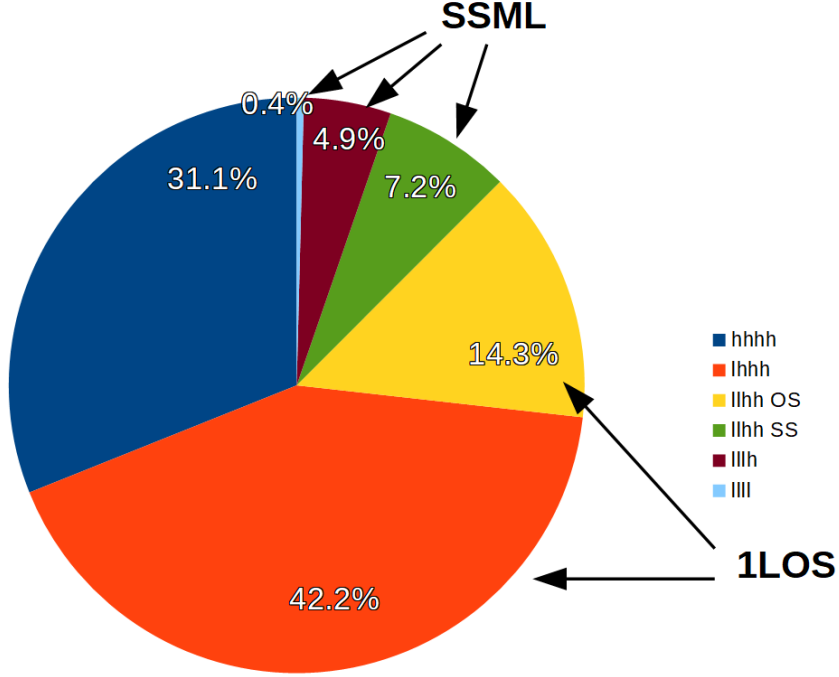


Figure 2.5: Branching ratios for $t\bar{t}t\bar{t}$ decay.

issues.

Hierarchy Problem: The Standard Model cannot explain why the mass of the Higgs boson is roughly 125GeV. There are correction terms to the mass that are much larger (on the scale of a unification energy) that must very precisely cancel in order to give the measured value of the Higgs mass.

Another way to envision this is to consider the interaction of the Higgs particle with virtual particles spontaneously created in the vacuum[27]. The interactions affect the Higgs mass in a way proportional to the energy scale Λ of the vacuum. The leading contributions to the mass correction include terms dependent on self-interactions of strength λ , interactions with gauge bosons of strength g and g' , and top quark Yukawa interactions of strength y_t [28]:

$$\delta m_H^2 = \frac{\Lambda^2}{32\pi^2} [6\lambda + \frac{1}{4}(9g^2 + 3g'^2) - y_t^2] \quad (2.29)$$

However, if the energy scale becomes very large, as it must to unify gravity

with the other forces, then the contributions to the Higgs mass become enormous. In order for m_H to be its relatively small measured value of 125GeV, there must be other correction terms from new physics that cancel out the large value of δm_H^2 . Without new physics, the cancellation can only be accomplished by balancing out contributions in a very precise way, and the amount of precision required is considered unnatural. The problematic relation of the small Higgs mass to the much larger unification energy scale is known as the ‘hierarchy problem.’

Dark matter: Astronomical observations of galaxies reveal that their rotational velocity is much higher than would be expected from the amount of visible matter present[29]. Imaging using gravitational lensing provides further evidence of clusters of invisible matter[3]. In order to explain these phenomena, there must be additional gravity provided by as-yet unseen matter, called ‘dark’ matter. Neutrinos are a possible dark matter candidate within the Standard Model, but their small mass is inconsistent with the behavior of observed dark matter. There is currently no experimental proof regarding the nature of dark matter. Many BSM theories, such as supersymmetry and extended Higgs models, propose additional particles that could be candidates for dark matter[8, 30].

Baryogenesis: Astronomical data, such as surveys of cosmic rays, indicate that the observable universe is composed of mostly matter rather than anti-matter. Indeed, if there was anti-matter present, it would quickly annihilate with the surrounding matter[31]. The Standard Model does not offer an explanation for why there is such an imbalance. There are CP-violating weak interactions that favor the production of particles over their anti-particles, but the strength of the CP violation does not produce a sufficient imbalance. BSM theories offer more CP-violating interactions that can account for the discrepancy.

There are many BSM theories that have been developed to address these issues. The work presented here involves two such BSM interpretations: a two Higgs doublet model (2HDM), and a BSM addition to the self-energy, or oblique, correction to the Higgs propagator.

2.3.1 2HDM

For the 2HDM interpretation, the BSM extension is applied by adding a second Higgs doublet to the existing SM scalar doublet [8]. In the SM case, the spontaneous symmetry breaking of the Higgs potential leads to one massive, scalar particle. For the 2HD case, eight fields are created, which leads to five particles: the SM Higgs,

one other neutral scalar H , two charged scalars $H^\pm$, and one pseudoscalar particle A .

This simple extension to the Standard Model is popular for several reasons. The existence of a second neutral scalar particle in the 2HDM aligns with many supersymmetric models, which require two Higgs doublets. Supersymmetry directly addresses the hierarchy problem by assigning supersymmetric ‘partners’ to fermions and bosons [30]. The 2HDM model can also directly address the baryogenesis issue by providing new channels for CP violation to occur.

The SM Higgs potential for a single scalar field ϕ is written as

$$V = m_1^2 \bar{\phi}\phi + \lambda(\bar{\phi}\phi)^2 \quad (2.30)$$

For two scalar fields, the potential becomes:

$$\begin{aligned} V = m_{11}^2 \bar{\phi}_1 \phi_1 + m_{22}^2 \bar{\phi}_2 \phi_2 - m_{12}^2 (\bar{\phi}_1 \phi_2 + \bar{\phi}_2 \phi_1) + \frac{\lambda_1}{2} (\bar{\phi}_1 \phi_1)^2 + \frac{\lambda_2}{2} (\bar{\phi}_2 \phi_2)^2 \\ + \lambda_3 \bar{\phi}_1 \phi_1 \bar{\phi}_2 \phi_2 + \lambda_4 \bar{\phi}_1 \phi_2 \bar{\phi}_2 \phi_1 + \frac{\lambda_5}{2} [(\bar{\phi}_1 \phi_2)^2 + (\bar{\phi}_2 \phi_1)^2] \end{aligned} \quad (2.31)$$

One defining parameter of the 2HDM is the ratio of the vacuum expectation values for the two Higgs doublets. This is defined such that, for vacuum expectation values:

$$\begin{aligned} \langle \Phi_1 \rangle_0 &= \begin{pmatrix} 0 \\ v_1/\sqrt{2} \end{pmatrix} \\ \langle \Phi_2 \rangle_0 &= \begin{pmatrix} 0 \\ v_2/\sqrt{2} \end{pmatrix} \end{aligned} \quad (2.32)$$

the parameter β is:

$$\tan \beta \equiv \frac{v_2}{v_1} \quad (2.33)$$

Also relevant is the parameter α , which is the rotation that diagonalizes the mass squared matrix:

$$\begin{pmatrix} m_{12}^2 \frac{v_2}{v_1} + \lambda_1 v_1^2 & -m_{12}^2 + \lambda_{345} v_1 v_2 \\ -m_{12}^2 + \lambda_{345} v_1 v_2 & m_{12}^2 \frac{v_1}{v_2} + \lambda_2 v_2^2 \end{pmatrix} \quad (2.34)$$

where $\lambda_{345} = \lambda_3 + \lambda_4 + \lambda_5$. The values of α and β determine the interaction strength of the Higgs field with other particles, in essence giving them their particular mass. Constraining β by using the measured $t\bar{t}t\bar{t}$ cross section is a goal of the 2HDM $t\bar{t}t\bar{t}$ analysis.

There are several versions of 2HDM models. This $t\bar{t}t\bar{t}$ search is based on the Type II model, where the $Q = \frac{2}{3}$ right-handed quarks couple to one Higgs doublet and the $Q = -\frac{1}{3}$ right-handed quarks couple to the other.

The model adds Higgs particles that can couple to top quarks, providing additional production modes for $t\bar{t}t\bar{t}$ as shown in Figure 2.6. This enhances the $t\bar{t}t\bar{t}$ cross section by an amount determined by the parameter β . The cross sections for different Higgs masses are obtained from the LHC Higgs Cross Section Working Group, and are shown in Figure 2.7[32]. The cross sections are given for the combination of $t\bar{t}t\bar{t}$ production in association with H and in association with A .

2.3.2 Higgs Oblique Models

BSM models generally assume the existence of particles and interactions that have not yet been experimentally detected. An Effective Field Theory (EFT) treats the Standard Model as a low-energy approximation of a higher-order theory, where BSM physics can be detected in the context of corrections to SM propagators.

Some interpretations, such as the 2HDM model discussed above, present a means of direct detection. The Higgs Oblique BSM interpretation, in contrast, seeks to establish the presence of BSM particles through their impact on the interactions of SM particles [9]. In particular, an ‘‘oblique’’ parameter affects interactions indirectly through self-energy corrections to electroweak propagators. This can be visualized in the Feynman diagram shown in Figure 2.8.

The propagator for the Higgs boson in the SM is given as:

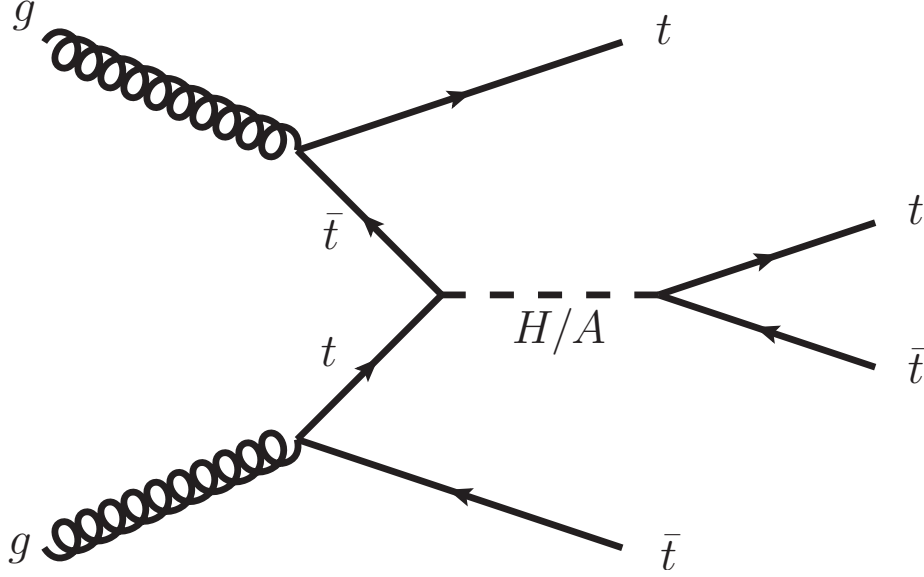


Figure 2.6: $t\bar{t}t\bar{t}$ production in association with a CP-even Heavy Higgs, H , or pseudoscalar Heavy Higgs, A .

$$\Delta_h(p^2) = \frac{1}{p^2 - m_h^2} \quad (2.35)$$

where m_h is the SM Higgs boson mass. The strength of the BSM self-energy term is parameterized by the Higgs Oblique parameter, $\hat{H}$, and affects the propagator:

$$\Delta_h(p^2) = \frac{1}{p^2 - m_h^2} - \frac{\hat{H}}{m_h^2} \quad (2.36)$$

which alters the cross section of every interaction involving the Higgs. However, many common processes are not ideal probes of $\hat{H}$ because there are cancellations in $pp \rightarrow h \rightarrow VV$ processes that make the $\hat{H}$ impact on couplings to gauge bosons disappear. The top quark has the largest mass of any SM particle, and therefore the strongest coupling to the Higgs, so any nonzero $\hat{H}$ value will have a greater impact on the cross sections of interactions with top quarks. This impact will also be

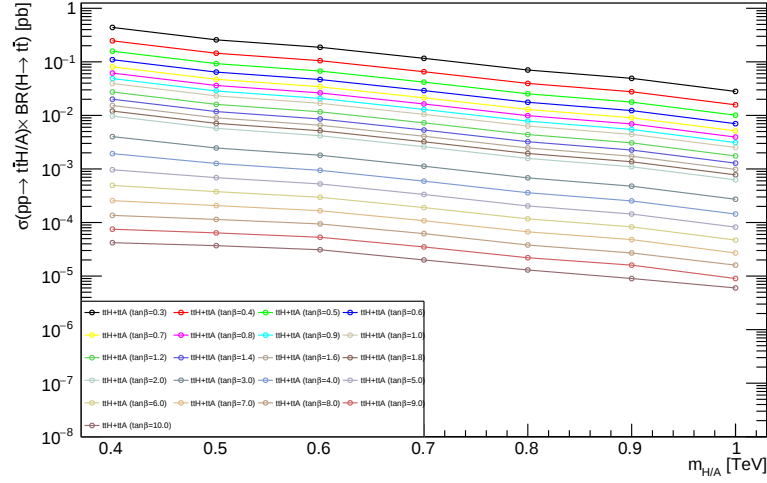


Figure 2.7: Cross sections of $t\bar{t}t\bar{t}$ production associated with Heavy Higgs H or pseudoscalar Heavy Higgs A . Different lines correspond to different values of the parameter $\tan\beta$.

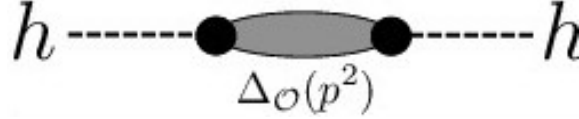


Figure 2.8: Higgs propagator affected by a virtual BSM particle. The addition to the self-interaction energy is expressed as $\Delta\mathcal{O}(p^2)$, where p is the momentum.

more apparent for processes with small SM cross sections. Therefore, the $t\bar{t}t\bar{t}$ process associated with Higgs production is a particularly good probe of the $\hat{H}$ value.

When the Higgs Oblique parameter is applied to the $t\bar{t}t\bar{t}$ production cross section, it is predicted at $\sqrt{s} = 14\text{TeV}$ collision energy to alter the cross section in a quadratic fashion [9]:

$$\delta\sigma_{t\bar{t}t\bar{t}} \equiv \frac{\sigma_{\hat{H}} - \sigma_{SM}}{\sigma_{SM}} \approx 0.03 \left(\frac{\hat{H}}{0.04} \right) + 0.15 \left(\frac{\hat{H}}{0.04} \right)^2 \quad (2.37)$$

This function assumes small $\hat{H}$ values. In fact, above $\hat{H} = 0.2$, the theory

no longer satisfies unitarity, and thus values above 0.2 are forbidden. Equation 2.37 does not include interactions with other electroweak particles, specifically the Z and photon. Figure 2.9 shows that calculations made using MadGraph5_aMC@NLO can reproduce Equation 2.37 assuming a center-of-mass energy of $\sqrt{s} = 14\text{TeV}$. In order to better replicate LHC conditions, the impact is also shown for $\sqrt{s} = 13\text{TeV}$, and including interactions with Z and γ .

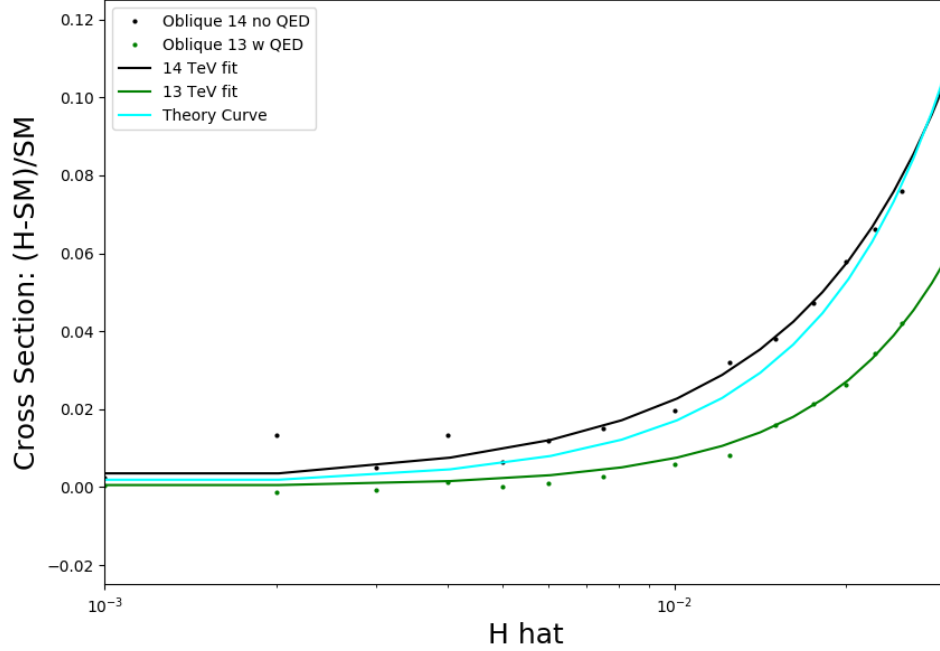


Figure 2.9: Impact of the $\hat{H}$ parameter on the $t\bar{t}t\bar{t}$ cross section. The cross section is calculated for 15 $\hat{H}$ values, using MadGraph5_aMC@NLO, and compared to the “Theory curve” given in Equation 2.37. Two cases are considered: one at $\sqrt{s} = 14\text{TeV}$ without Z/γ interactions, and one at $\sqrt{s} = 13\text{TeV}$ with Z/γ interactions.

CHAPTER 3

Experimental Apparatus

3.1 Introduction

The general trend of particle physics experiments has been towards observation of particle interactions at higher energy, and thus smaller distance scales. This realm is the least understood in part because it is so difficult to achieve. Many modern high-energy investigations rely upon specialized particle colliders; the largest of these, the Large Hadron Collider (LHC), is the most energetic machine in the world.

The LHC became operational in 2008, 79 years after the invention of the first circular accelerator. The first circular accelerator, a cyclotron developed by Ernest Lawrence, accelerated protons to 80keV, and was small enough to fit in one hand. In contrast, the LHC is designed to accelerate protons to a maximum 7TeV, and its 27km circumference lies 100m below the ground at the border of France and Switzerland[33]. The LHC was constructed by the European Organization for Nuclear Research, historically abbreviated as CERN. Originally a collaboration between the countries of Western Europe, CERN now hosts tens of thousands of scientists from more than 70 countries.

The operation of the LHC comprises many parts. Proton beams must be produced, accelerated, and then focused at specific collision sites, known as Interaction Points (IP's). These IP's are dictated by the placement of particle detectors, which record aspects of the collision such as particle tracks and energy depositions. The LHC has four main detectors. The Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) are the two general purpose detectors. ALICE and LHCb are designed for more specific studies. In addition to these main experiments, there are three smaller detectors: TOTEM, LHCf, and MoEDAL. This section will describe both the LHC itself and its experiments, with further details provided for the ATLAS detector, which was the detector used for this research.

3.2 LHC

The LHC produces proton-proton (pp) collisions at energies in the TeV range. For the first data-taking period (Run I), pp collisions occurred at center-of-mass energies of $\sqrt{s} = 7\text{TeV}$ and $\sqrt{s} = 8\text{TeV}$. For Run II, the energy was increased to $\sqrt{s} = 13\text{TeV}$. Regardless of final energy, the protons must undergo a series of sequential accelerations, which make use of the facilities of previous CERN experiments.

First, protons must be produced. This is accomplished by using an electric field to strip electrons from a hydrogen sample. These protons undergo a series of accelerations. A linear accelerator (the Linac2) accelerates protons to 50MeV. There are three additional circular accelerators: the Proton Synchrotron Booster, Proton Synchrotron, and Super Proton Synchrotron, which further accelerate the protons to energies of 1.5GeV, 25GeV, and 450GeV respectively [34]. Lastly, these pre-accelerated protons are injected into the LHC itself, reaching a final individual beam energy of 6.5TeV. To achieve this high energy, the LHC makes use of a 400MHz Radio Frequency (RF) for acceleration and an array of 1232 superconducting dipole magnets at 8T to produce the circular path of the beam. Additionally, magnets with higher-order fields (quadrupole, sextupole, and octopole) direct and focus the beam at the aforementioned IP's where the detector experiments are located.

3.3 Experiments

The diversity of experiments at the LHC reflects the diversity of questions that still exist in the field of particle physics. The following are the specialized detectors at the LHC, constructed for a few specific purposes:

ALICE, located at IP 2, records the interactions of lead nuclei collisions ($Pb - Pb$) and collisions of other heavy nuclei. The aim of these ion collisions is to create quark-gluon plasma and observe its properties, in particular the behavior of free partons prior to color confinement [36].

LHCb, located at IP 8, exploits the large energy and luminosity of the LHC as a means of manufacturing b-quarks. Experiments at LHCb seek to compare properties of b-hadrons with predictions from the Standard Model as well as to address possible causes of the matter-antimatter imbalance in the universe [37].

TOTEM is located in the region forward of the CMS detector, and provides information regarding the small-angle collisions of the proton beams: cross-sections

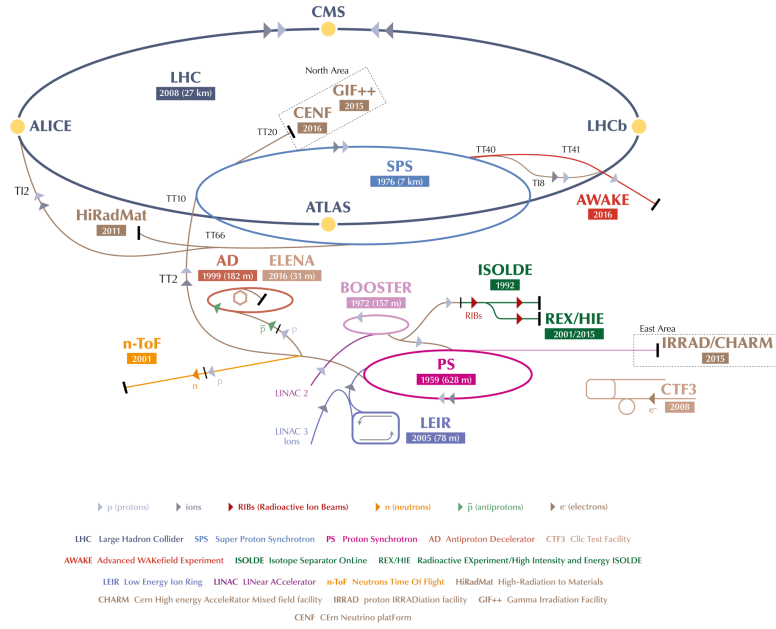


Figure 3.1: Layout of the LHC complex, showing the various accelerators, the LHC itself, and the positions of the different experiments[35].

of hard scattering as well as elastic scattering and diffractive processes [38].

MoEDAL shares the LHCb interaction point. Its detectors are sensitive to highly ionizing particles. The experiment was designed to search for magnetic monopoles [39].

The two general-purpose detectors, CMS and ATLAS, consist of a series of different particle detectors layered in a cylindrical design. The inner layers of the detector measure the momentum of charged particles using a solenoidal magnetic field. The middle layers consist of calorimeters that measure energy deposits of all particles except neutrinos and muons. The outer layers of the detector measure the momentum of muons using toroidal magnetic fields. The details of this layout are discussed in the next section, in the context of the ATLAS detector.

CMS, the Compact Muon Solenoid detector, is located at IP 5 in the LHC [40]. It shares many similarities to ATLAS in both its design and physics goals,

though the magnet configuration is different: CMS has a single internal solenoid with a 4T field, where ATLAS has a 2T internal solenoid, and additional 0.5-1T toroid magnets placed just inside the muon spectrometers.

3.4 The ATLAS Detector

3.4.1 The Geometry of the Detector

To properly describe the geometry of the detector and particle interactions, standardized coordinates are used. The coordinate system is nearly the same as the cylindrical system, with r as a radial component and ϕ as the angular component. Instead of using θ as the angular coordinate along the beam axis, pseudorapidity η is used. This is a coordinate specialized for use in high-energy collisions. The difference between two η values is approximately invariant under a Lorentz boost along the beam axis, and is defined as

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

where θ is the polar angle with respect to the beam axis. In terms of momentum $\vec{p}$ and longitudinal momentum p_L , the pseudorapidity is

$$\eta = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_L}{|\vec{p}| - p_L} \right) \quad (3.2)$$

The rapidity y is the true Lorentz-invariant quantity, defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) \quad (3.3)$$

At very high energies such that $|\vec{p}| \approx E$, then $\eta \approx y$ and the pseudorapidity becomes approximately Lorentz invariant. When a particle's trajectory is nearly parallel to the beam axis, θ approaches zero and η approaches infinity.

To further describe particle properties, variables often make use of the “transverse” component. For example, transverse momentum, p_T , represents the particle momentum in directions normal to the beam axis (i.e. a combination of radial and

angular coordinates). This is useful for calculations because the initial transverse momentum of any collision is very close to zero, and hard scatter physics will appear in the transverse plane. In contrast, longitudinal momentum of the partons will always depend on parton distribution functions.

3.4.2 ATLAS Inner Detector (ID)

The ID consists of three main components: the Silicon Pixel Tracker, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT) [41]. Detectors in this region are designed to operate with high temporal and spatial resolution while also withstanding the extreme radiation from their location closest to the IP.

The Silicon Pixel Tracker is the innermost detector, with the innermost layer located at a radius of 5cm from the IP. The Pixel Tracker consists of three layers of silicon detectors which yield three points of trajectory for charged particles as they curve in the presence of the inner magnetic field. Each barrel pixel is $50\mu\text{m}$ in $R-\phi$ (corresponding to the arc length) $\times 400\mu\text{m}$ in z , and each endcap pixel is $50\mu\text{m}$ in $R-\phi \times 400\mu\text{m}$ in R , providing a resolution of $10\mu\text{m} \times 115\mu\text{m}$ for the region $|\eta| < 2.5$. The endcap detector contains 13 million pixels, and the barrel contains 67 million pixels, collectively covering an area of 1.7m^2 .

The SCT is located radially outside of the Pixel Tracker, at a distance of 30-56cm from the IP. Its objective is to provide similar tracking information at a larger radius where individual pixels become intractable. The SCT is constructed of 4088 silicon-strip detectors forming four barrel layers and two endcap layers, also covering up to $|\eta| < 2.5$. There are 2112 barrel sensors with $80\mu\text{m}$ pitch, and 1976 endcap sensors with a rotational pitch of 40 mrad.

The TRT uses proportional drift tubes to provide $R-\phi$ tracking information at a radius of 56.3-106.6cm from the IP. The tubes are arranged in three layers in the barrel, parallel to the beam axis and forming 32 separate sections. In the endcap, they are oriented radially. The tubes provide a $R-\phi$ precision of $130\mu\text{m}$. There are 351,000 channels covering the region up to $|\eta| = 2.0$. The TRT is a “transition radiation” detector designed to generate photons from particles as they pass through different materials. As light particles will radiate more, the TRT provides differentiation between electrons and hadrons.

Also contained within the ID region is the Beam Conditions Monitor (BCM), which is used not for data-taking but for quality control of the beam. Sixteen pCVD diamond sensors are located in the forward and backward regions of the detector,

1.84m along the beam axis from the IP and 5cm away in the radial direction. The rapid, 1ns rise time of the pCVD sensors ensures the BCM has capability to discriminate a signal bunch-crossing (occurring every 25ns) from background collisions. This information is used for both triggering and for the beam abort signal.

The detectors in the ID rely on an external magnetic field to bend particle trajectories according to their charge and transverse momentum. This magnetic field is produced by the Central Solenoid. A 2T field is produced by a superconducting NbTi coil of wire, cooled to 4.5K and wrapped 1173 times around the ID in such a way that the magnetic field runs parallel to the beam axis. Due to its proximity to the IP, particles may still interact with and be influenced by the material in the solenoid. To alleviate that issue, the solenoid is designed to be as thin as possible, supported by a structure made of specialized, high strength aluminum.

3.4.3 ATLAS Calorimeters

The role of the calorimeters in the ATLAS detector is to initiate hadronic and electromagnetic showers and to measure their energy and position. There are two types of calorimeters: one designed to accommodate electromagnetic (EM) showers, and the other dedicated to hadronic showers. Given the shorter range of electromagnetic showers, the EM calorimeters are located closer to the IP. The hadronic particles typically pass through all or much of the material of the EM calorimeters and enact their own larger showers within the material of the hadronic calorimeters. These two types of calorimeter are located in both the barrel and endcap regions of ATLAS.

Showers originating from electrons and photons proceed via the electromagnetic interaction. These showers are generally more compact and less complex than showers initiated by hadrons. At high energies, the primary interactions are photons interacting with material to produce an electron/positron pair (“pair production”) and leptons emitting photons via *bremsstrahlung*. Materials have a “critical energy,” E_c , at which the rates of energy loss from lepton *bremsstrahlung* and ionization are equal[5]. Showers build until reaching E_c , and shower depth is dependent on both E_c and the “radiation length” X_0 of the material. The radiation length is the distance a particle can travel before losing $1/e$ of its energy. The shower depth is expressed as:

$$X = X_0 \frac{\ln(E_0/E_c)}{\ln(2)} \quad (3.4)$$

This results in EM showers that are compact in longitudinal and transverse size.

Showers originating from hadrons tend to be much broader in transverse and longitudinal size and involve more complex QCD and nuclear interactions. These showers are initiated by a collision between a particle and an atomic nucleus, which is subsequently broken apart and/or excited. The “nuclear interaction length” is a material property defined by the distance a particle can travel before undergoing an inelastic nuclear collision. The interaction length increases as a material’s atomic number increases, and decreases with material density. Hadronic showers also contain varying fractions of electromagnetic showers, primarily originating from the decay of neutral pions into photon pairs.

The calorimeter system is designed to cover the largest solid angle possible, increasing the accuracy of missing E_T estimates. There are six calorimeters: the Forward Calorimeters (both EM and hadronic), the Endcap Calorimeters (EM and hadronic), the EM Barrel Calorimeter (EMB), and the Hadronic Tile Calorimeter, which is also located in the barrel region. Though the absorber material is different between EM and hadronic calorimeters, all four non-barrel calorimeters make use of Liquid Argon as a detection material, exploiting its radiation tolerance and linear, time-stable response.

While tracking detectors experience reduced resolution as the particle momentum increases, calorimeters experience the opposite. In the barrel and endcap calorimeters, electrons and pions have fractional energy resolution of approximately 1-3% for beams 200-10GeV respectively. In the forward calorimeters, electrons have a resolution of about 5-10% for beams 200-10GeV, and pions have a resolution of about 10-25%.

Electromagnetic Barrel Calorimeter: The EM barrel calorimeter covers the $|\eta| < 1.475$ region, arranged in an accordion geometry as seen in Figure 3.2, with the LAr detectors interspersed with lead absorber plating. The barrel calorimeter consists of two half-barrels that meet at the $z=0$ point with a gap of 4mm. The EM barrel calorimeter is designed to accommodate at least 22 radiation lengths: 22-30 for $|\eta| = 0-0.8$, and 24-33 for $|\eta| = 0.8-1.3$.

Hadronic Tile Calorimeter: The tile calorimeter covers the $|\eta| < 1.0$ region with its central barrel, and two extended barrels on each end cover the region $0.8 < |\eta| < 1.7$. The absorbing material is steel and the detectors are scintillator tiles. There are three layers extending radially outward for each barrel. For the

central barrel, these layers are 1.5, 4.1, and 1.8 interaction lengths thick. For the extended barrel, the layers are 1.5, 2.6, and 3.3 interaction lengths thick.

EM Endcap (EMEC): The endcap calorimeters are placed at a smaller rapidity than the forward calorimeters. The EMEC covers the $1.374 < |\eta| < 3.2$ region. The lead-LAr detectors are arranged in the same accordion-structure of lead and LAr that comprises the barrel calorimeter.

Hadronic Endcap (HEC): The HEC covers the pseudorapidity region of $1.5 < |\eta| < 3.2$, and consists of 8.5mm gaps of LAr layered with copper absorber. At each endcap there are two independent wheels consisting of 32 modules.

Forward Calorimeters: The tendency of increased resolution for increased particle momentum allows for the placement of calorimeters in the far forward region of $3.1 < \eta < 4.9$, where tracking detectors would be of little use. The forward calorimeters begin at a distance of 5m along the beam axis from the IP, and form three layers: FCal1, FCal2, and FCal3. FCal1 uses a copper absorber, which acts as both an efficient conductor of heat and an absorber of EM showers. FCal2 and FCal3 use high-Z tungsten as their absorber to instigate hadronic showers. Lastly, a brass plug is used to absorb any remaining radiation, which reduces radiation damage to machinery in the forward region, prevents formation of radioactive material, and reduces background in the endcap muon detectors.

Layered between the absorbers are hexagonal arrays for the LAr calorimeter readout. The calorimeters are electrode tubes consisting of a grounded, cylindrical exterior, a gap filled with LAr (the dimension varies between FCals, but is 0.27mm for FCal1), and a positively charged interior rod. The small gap spacing ensures that the positive Argon ions can quickly neutralize on the electrode's exterior, giving the calorimeter a drift time of 61ns in FCal1.

The University of Arizona played a leading role in the design and construction of the forward calorimeters, having first developed the proof of concept and prototype in 1993 [42].

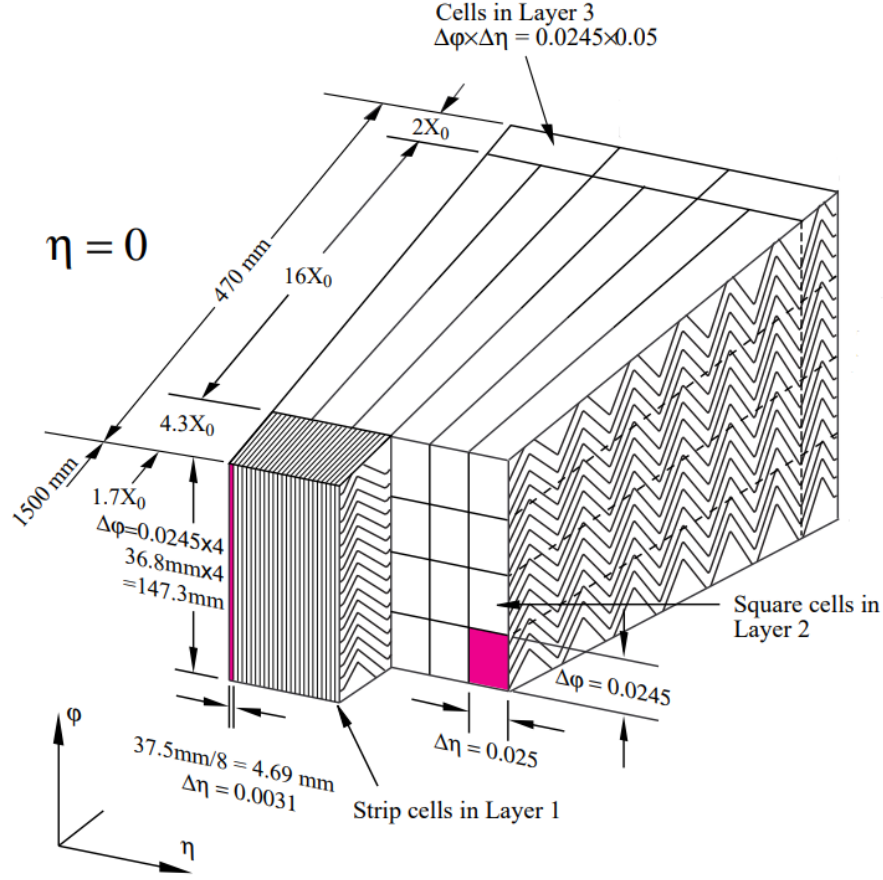


Figure 3.2: One of the modules of the barrel EM calorimeter, displaying the accordion geometry found in the EM barrel and endcap calorimeters [41]. The different dimensions and layers of the calorimeter are also labeled.

3.4.4 Muon Spectrometers

The muon detectors are located on the outermost part of the ATLAS detector, where muons can be detected with low contamination from other particles. The system is designed to track the deflection of muons as they move through the toroidal magnetic field. There is a layer of muon detectors between the hadronic calorimeters and the toroid magnets, a layer within the toroid magnets, and a layer outside of the toroid magnets. In terms of rapidity, there are three different sections: the barrel covering $|\eta| < 1.4$ measuring muons deflected by the barrel toroids, a transition region $1.4 < |\eta| < 1.6$ where muons may be deflected by either the barrel or endcap magnet,

and the endcap covering $1.6 < |\eta| < 2.7$ measuring muons deflected by the endcap toroids. The muon spectrometer layout is shown in Figure 3.3.

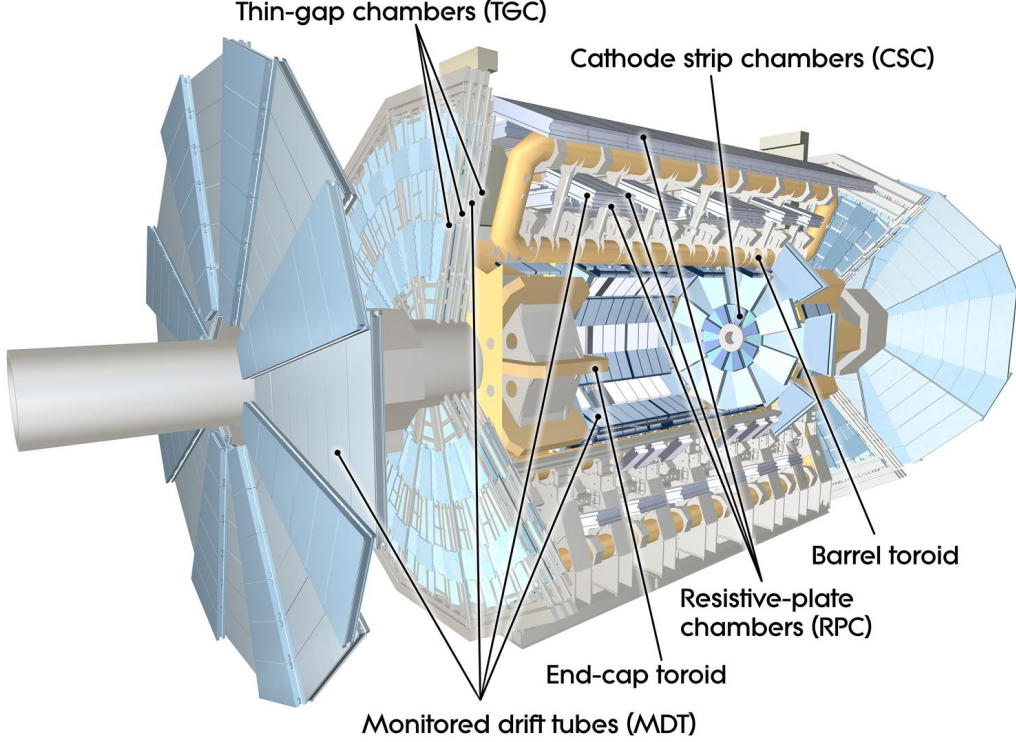


Figure 3.3: Layout of the muon spectrometer system on the ATLAS detector, showing the various types of muon detectors [43].

There are four types of muon detector: Monitored Drift Tubes (MDT's), Cathode Strip Chambers (CSC's), Resistive Plate Chambers (RPC's), and Thin Gap Chambers (TGC's). MDT's and CSC's are used for muon tracking, RPC's are used for trigger purposes, and TGC's are used for both tracking and trigger.

The MDT's are tubes of Argon (93%) and CO_2 (7%) gas with a $50\mu\text{m}$ anode wire in the center held at a 3080V. When a muon passes through, it ionizes the gas, and the electrons are drawn to the anode, generating a signal. The use of many independent tubes offers both structural stability and robustness against individual failure. The MDT's are organized in rectangular chambers in the barrel, and trapezoidal chambers in the endcap, achieving resolution of $80\mu\text{m}$ per tube and $35\mu\text{m}$ per chamber. However, due to the high maximum drift time of 700ns, the MDT's cannot

tolerate high signal rate. In most of the endcap region they cover up to $|\eta| < 2.7$, but in the innermost Small Wheel, they only cover up to $|\eta| < 2$.

The CSC's have a much faster drift time of 40ns, and are employed in the inner (Small) wheel of the endcap, covering regions $2 < |\eta| < 2.7$. They are composed of many anode wires, held at 1900V, contained within a rectangular gap between two cathodes comprised of parallel strips. The gap is filled with a similar gas mixture of 80% Argon and 20% CO₂. The distance between each anode is 2.5mm, and the distance between an anode and a cathode plate is also 2.5mm. There are large and small CSC chambers, with 250 wires in the small chamber and 402 wires in the large chamber. In addition to the faster response time, the CSC's have a better ability to resolve multiple tracks, a necessary feature in the forward region. The CSC resolution is 60 μ m in the bending (ϕ) plane and 5mm in the non-bending plane.

The RPC's are used for the barrel region of the trigger system. Their design is comprised of parallel resistive plates with an electric field of 4.9kV/mm between them. The ionization of the gas within the gap generates an electron avalanche, which induces a signal in strips located outside of the resistive plates. The RPC's are arranged in rectangular chambers in the barrel, covering the rapidity region $|\eta| \leq 1.05$. They have a spatial resolution of 23-35mm in η and ϕ .

The TGC's have a similar design to the CSC's, with a series of anode wires within a gas chamber. The spacing is smaller—1.8mm between anodes, and 1.4mm between an anode and an inner graphite layer—and the gas is a mixture of 55% CO₂ and 45% n-pentane rather than CO₂/Argon. The TGC's usually receive a signal from an ionizing muon within 25ns, and four layers of detectors allow the TGC's to handle the high signal rate in the region $1.05 \leq |\eta| \leq 2.4$. The TGC's have an azimuthal resolution of 2-3mrad.

The toroid magnets are a part of the muon spectrometer system, bending the paths of muons between the layers of muon detectors. There are two endcap toroids with a field of about 1T and one barrel toroid with 0.5T, each constructed of eight coils arranged in a radially symmetric fashion. The barrel toroid offers 1.5-5.5T $\times$ m of bending power in the barrel region ($|\eta| < 1.4$), and the endcap toroids have a bending power of 1-7.5T $\times$ m in the region $1.6 < |\eta| < 2.7$. There is less bending power in the transition region between $1.4 < |\eta| < 1.6$.

3.5 ATLAS Upgrades for the HL-LHC

After the end of Run II in 2018, the ATLAS detector underwent the Phase-I upgrade, where the primary objective was to accommodate the increase in pp collisions from the High Luminosity LHC. The HL-LHC is expected to operate with an instantaneous luminosity of $5\text{--}7 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, as compared to the peak luminosity in 2018 of $2.1 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ [44],[45].

One of the Phase-I upgrades was the replacement of the muon spectrometer's Small Wheels with the New Small Wheels. The NSW have faster responsiveness to provide more robust triggering in the forward region, owing to track stub formation in the NSW. This allows the trigger to maintain the same muon p_T thresholds that were used in Run II without significantly increasing the trigger rate. In previous runs, the endcap muon triggers only used the middle layer of the endcap muon detectors. In Run II, approximately 90% of the muon triggers were fake, arising from background hits in the muon chambers. Adding the NSW to the Level-1 trigger system will drastically reduce the contribution from these false signals.

The New Small Wheels are a combination of MicroMegas (MM) detectors and small Thin-Gap Chambers (sTGC's) designed to cover the same region as the Small Wheels: $1.3 < |\eta| < 2.7$. They consist of sixteen sectors in two multilayers, where each multilayer is composed of four layers of sTGC detectors and four layers of MM detectors.

3.5.1 sTGC

Much like its predecessor, the sTGC is designed as a chamber with anode wires spaced 1.8mm apart, positioned 1.4mm away from cathode plates. While the TGC has a strip pitch between 10.8mm-55.8mm, the sTGC has much finer granularity, and the strips are 3.2mm apart. This refines the resolution of the detector from 2-3mrad for the TGC to 1mrad for the sTGC. The sTGC are the primary trigger detector for the NSW, but can also be used in muon reconstruction.

3.5.2 MicroMegas Detectors

With a maximum response time of 100ns and a $100\mu\text{m}$ tracking resolution, the MM detectors are capable of handling the high numbers of incident particles present in the HL-LHC. The MM layout is shown in Figure 3.4. There is a chamber containing 93% Ar, 5% CO_2 gas, and 2% C_4H_{10} to prevent sparking. A grounded mesh is suspended 5mm from the -300V cathode plate and $128\mu\text{m}$ from the 550V anode plate. The

original designs of the MicroMegas were susceptible to sparking, and to prevent this, resistive strips are placed above the readout strips.

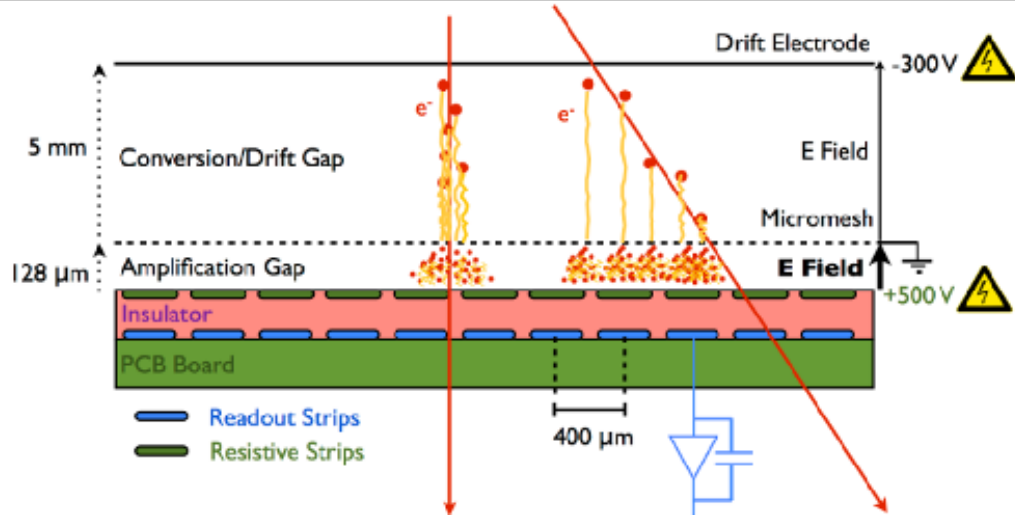


Figure 3.4: Schematic of the MicroMegas design, displaying the path of an ionizing particle within the detector [46].

There are two regions defined by the mesh: the drift region above the mesh, with an electric field of 600V/cm ; and the amplification region below the mesh, with a much greater electric field of 40kV/cm . A muon passing into the chamber will ionize the gas, and individual electrons will be drawn towards the mesh in the drift region. Once they reach the mesh, the amplification region will generate an avalanche of electrons that comprises a very fast pulse, within a fraction of a nanosecond. The positive ions created in the amplification region will return to the mesh within 100ns .

The readout technology includes printed circuit boards (PCB's) containing 1024 readout strips each, which have a pitch of approximately 0.4mm . There are roughly 2 million readout channels for the entire MM system. Each PCB contains eight VMM's, a type of Application Specific Integrated Circuit (ASIC), and each VMM processes 64 channels. The VMM converts the analog pulse signal into a digital signal with charge amplification, discrimination, timing information, and pulse amplitude information [47]. The events stored by the VMM are then passed to the Read-Out Controller (ROC), which formats the information for the backend processing. The MM are the primary detector for muon reconstruction, but can also provide information for triggering.

A research team at the University of Arizona was responsible for the development and testing of the PCB's, which are called MicroMegas FrontEnd 8 (MMFE8) in reference to the eight VMM's per board. Prior to the main production of the 4,096 MMFE8 boards installed on the NSW, the prototype MMFE8 designs underwent considerable development and testing, including reducing noise level and assessing operation in a magnetic field. A set of tests was created to certify the functionality of each part of the MMFE8, and was implemented when preparing the full production of MMFE8's for installation on the NSW. The tests included visual inspection, measurements of voltages and currents, data collection using both internal and external pulsers, and tests of the trigger output. Of the 4742 boards produced, there were 449 failures, approximately 100 of which came from failures of individual VMM's [48].

CHAPTER 4

Data and Monte Carlo Samples

4.1 Introduction

This chapter describes the data and Monte Carlo (MC) samples used in this analysis. MC samples are used to simulate the known SM processes and predicted BSM processes that constitute the signal and background for the analysis. The analysis sensitivity is evaluated and optimized using these MC samples as described in Chapter 9. For the measurement of the $t\bar{t}t\bar{t}$ cross section, data are collected from the ATLAS particle detector. The data undergo several steps of processing, reconstruction, and selection in order to be useful for physics analyses, as described in Chapter 5. This chapter focuses on the collection process itself: the integrated luminosity, the performance of the detector, and the identification of the high quality data suitable for the analysis.

In order to avoid bias within the analysis, data are “blinded” in the phase space most likely to contain events of interest. The relevant tools and selections are instead developed on simulated pp collision events, which are produced via an MC simulation. These MC samples provide initial estimates of signal and background behavior, assuming an input model and cross-checks performed in kinematically similar regions of phase space. MC samples also have “truth” information regarding the identity of all simulated particles in an event. The samples are employed to optimize the analysis strategy. Once the analysis process is finalized, the data are unblinded.

The reliance on MC requires that the physics processes in pp collisions, the LHC pileup environment, and particle detection in ATLAS all be modelled with exquisite accuracy. This chapter describes the data collected and the MC samples used.

4.2 Data

The analysis uses the data collected from the full Run II campaign at ATLAS: $140.1 \pm 1.2\text{fb}^{-1}$ collected between 2015-2018 at $\sqrt{s} = 13\text{TeV}$ [49]. The integrated

luminosity of the pp beams is measured primarily using the LUCID2 Cherenkov detector, exploiting its luminosity sensitivity [50, 51]. The beam conditions monitor (BCM) diamond detectors are also used, as are the SCT and pixel detectors, the endcap calorimeters, and the TileCal E-cells. The E-cells are segments of the TileCal, located in the extended calorimeter, that are placed in the forwardmost region of the TileCal.

The detectors are calibrated during specialized LHC runs at low luminosity, wherein the beams are systematically displaced using the van der Meer (vdM) method [52]. This calibration is then applied to the measurements of high-luminosity runs. The calibrations were performed yearly and the differences in measurements between luminosity detectors allow for estimation of uncertainty [51].

The uncertainty on the luminosity is approximately 0.83% of the total. Many uncertainties arise from the vdM scans, which are dependent on measurements of beam size, bunch population, and the algorithms used for luminosity calculation. In particular, the vdM technique assumes that beam sizes in the x and y directions are completely separable. The reality is that correlations between x and y produce ‘non-factorization’ effects[51]. The largest source of uncertainty in the luminosity calculation comes from extrapolating the luminosity values from the low-luminosity vdM regime to the high-luminosity regime, where physics measurements actually take place. This uncertainty is known as ‘calibration transfer’ and is 1.3-1.6% of the total luminosity value for a given year. The peak instantaneous luminosity in 2015 was $5 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$, and steadily increased over the campaign to a peak of $21 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$ in 2018[45]. The average number of pp collisions per bunch crossing, also called pileup collisions, was 33.7 for the full Run II data-taking.

The pp bunch collision rate is 40MHz. This rate is reduced to 100kHz by the Level 1 trigger filtering, which uses hardware to select only certain collisions of interest. This 100kHz rate is further reduced by the software-based High Level Trigger, to a rate of about 1kHz [49]. These events are stored and monitored for quality. Major issues arising in the detector—such as electrical surges, magnet failures, and readout software errors—can lead to data being rejected for use in physics analyses. In general, the efficiency of ATLAS data collection was very high, increasing from 88.8% in 2015 to 97.5% in 2018, and with an average efficiency of 95.6% [49].

The data are organized based on LHC intervals called ‘fills’ and ATLAS intervals called ‘runs.’ Each run has a unique number associated to it, and is divided further into luminosity blocks (LBs). LBs are defined by approximately constant

conditions for the detector, trigger, data quality, and instantaneous luminosity. LBs generally last 60 seconds or until conditions are observed to change.

4.2.1 Triggers

Data are collected based on selections made by the L1 and associated HLT triggers. For the HBSM $t\bar{t}t\bar{t}$ analysis, the data are selected using single electron and single muon triggers, shown in Table 4.1.

For this analysis, the naming convention identifies the HLT p_T threshold first (“eX” or “muX”, where X is the p_T threshold in GeV), then the isolation/likelihood requirements, if any. Isolation is based on the presence of other objects in a variable-sized cone, ΔR^{var} , that decreases in size as a function of object p_T . The L1 trigger may also be identified in the HLT name, but this is not the case for every trigger. There are a few other notation conventions:

A **nod0** HLT trigger does not include the transverse impact parameter d_0 in the likelihood algorithm. The **ivar** tag indicates a track isolation requiring $p_T^{iso}(\Delta R^{var} < 0.2)/p_T < 0.10$. **V** indicates an L1 η dependence on the E_T threshold. **H** indicates an L1 hadronic activity veto. **I** indicates that an L1 EM isolation is applied.

Trigger List		
Data-taking Year	HLT Trigger Name	L1 Trigger
Single Electron Triggers		
2015	e24_lhmedium_L1EM20VH	L1EM20VH
2016, 2017, 2018	e60_lhmedium, e120_lhloose	L1EM20VH
	e24_lhmedium_L1EM20VH	L1EM22VHI
	e60_lhmedium, e120_lhloose	L1EM22VHI
	e26_lhtight_nod0_ivarloose	L1EM22VHI
	60_lhmedium_nod0, e140_lhloose_nod0	L1EM22VHI
Single Muon Triggers		
2015	mu20_iloose_L1MU15	L1MU15
2016, 2017, 2018	mu50	L1MU20
	mu26_ivarmedium, mu50	L1MU20

Table 4.1: Single lepton triggers used for the HBSM $t\bar{t}t\bar{t}$ analysis.

The Higgs Oblique analysis uses a similar collection of single lepton triggers, with the addition of di-lepton triggers shown in Table 4.2[53]. The naming conven-

tions are the same, with the addition of “muXnoL1,” as in mu22_mu8noL1. This means that the selection is made by applying the single muon trigger mu22, and then locates a subleading muon with $p_T > 8\text{GeV}$.

Trigger List		
Data-taking Year	HLT Trigger Name	L1 Trigger
Single Electron Triggers		
2015	e24_lhmedium_L1EM20VH	EM20VH
2016, 2017, 2018	e60_lhmedium, e120_lhloose	EM20VH
	e26_lhtight_nod0_ivarloose	EM22VHI
	e60_lhmedium_nod0, e140_lhloose_nod0	EM22VHI
Single Muon Triggers		
2015	mu20_iloose_L1MU15	MU15
2016, 2017, 2018	mu50	MU20
	mu26_ivarmedium, mu50	MU20
Di-electron Triggers		
2015	2e12_lhloose_L12EM10VH	2EM10VH
2016	2e17_lhvloose_nod0	2EM15VH
2017, 2018	2e24_lhvloose_nod0	2EM20VH
Di-muon Triggers		
2015	2mu10	2MU10
2016, 2017, 2018	mu18_mu8noL1	MU15
	2mu14	2MU10
	mu22_mu8noL1	MU20
Electron-muon Triggers		
2015	e17_lhloose_mu14	EM15VH_MU10
2016, 2017, 2018	e7_lhmedium_mu24	MU20
	e17_lhloose_nod0_mu14	EM15VH_MU10
	e7_lhmedium_nod0_mu24	MU20

Table 4.2: Single lepton and dilepton triggers used for the Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

The electron triggers are a combination of selections performed by the L1 trigger and the HLT [54]. The L1 trigger first identifies regions of interest (RoI's) within the EM calorimeter, based on the maximum E_T within a given region of the central calorimeter ($|\eta| < 2.5$). For $E_T < 50\text{GeV}$, the electron candidate must also pass an E_T threshold (20GeV for e24_lhmedium_L1EM20VH), and the E_T in the corresponding area of the hadronic calorimeter must be sufficiently low. An

isolation requirement rejects electron candidates with high- E_T cells nearby in the EM calorimeter.

Events passing the L1 trigger are forwarded to the HLT, which first uses a fast reconstruction to confirm electron candidates, and later performs a precision reconstruction on the remaining events. The HLT fast reconstruction uses two different algorithms: a cut-based algorithm which is applied to electron candidates with $E_T < 15\text{GeV}$, and a neural-network “Ringer” algorithm which is applied to electrons with $E_T \geq 15\text{GeV}$ [54]. The Ringer algorithm identifies electrons based on their tendency to produce conical showers, which means that energy deposited in concentric rings within the EM and hadronic calorimeters acts as a useful discriminant.

The precision reconstruction involves matching tracks within the RoI to energy deposits in the second layer of the EM calorimeter, which must be suitably close ($\Delta\eta \leq 0.05$, $\Delta\phi \leq 0.05\text{rad}$). Electron selection is applied using the likelihood classification described in Chapter 5.

The muon triggers similarly use the L1 trigger and the associated HLT triggers[55]. The L1 trigger identifies muon candidates and their associated RoIs using a combination of tracking and p_T measurements. The muon trajectory is affected by the toroid magnets, and deviation from a straight line (infinite p_T) is converted to a p_T value. Tracking of the muon candidate is nontrivial due to the far radial distance and structure of the MS. Numerous gaps and necessary dead spaces for detector servicing place limitations on where muons can be measured, and showers produced at detector edges constitute a significant background of hits. Tracking is addressed separately in the barrel ($|\eta| < 1.05$) and the endcap ($1.05 < |\eta| < 2.4$) regions.

In the barrel, there are four layers of RPC detectors. The single muon triggers use a high- p_T threshold, which requires a coincidence in the inner three chambers. The outermost layer, also known as the “feet” chambers, is a newer addition to the RPC array and is used in coincidence with Layer 3 to detect muon candidates that pass through gaps in the inner two chambers.

In the endcap, triggering is based on coincidences between the TGC layers. A muon candidate must have a coincidence either between the Endcap Inner (EI) TGC and the middle layer of the Big Wheel TGCs, or between the Forward Inner (FI) TGC and the middle layer of the Big Wheel TGCs. There are also coincidence requirements between the hadronic TileCal and the TGCs in the lower eta ($1.05 < |\eta| < 1.3$) region.

The fast algorithm for the muon HLT, also known as the stand-alone (SA) algorithm, uses the MDT hits in the RoI to improve the track reconstruction provided by the L1 trigger. For the single muon trigger, the precision algorithm uses track matching in the ID and the MS to select muon candidates.

4.2.2 Data Processing

Data passing the HLT are stored at the Tier-0 computing facility at CERN as RAW data [56]. The data undergo several processing stages to be used for analysis: quality inspection; reconstruction and calibration into Analysis Object Data (AODs); derivation (described below) and size reduction into derived AODs (DAODs); and finally further variable definition and organization on an analysis level via a framework, which produces the Root-based ntuples [57].

If data-taking conditions do not pass inspection from the ATLAS data quality group, the LB is rejected. A bad LB generally occurs during a large-scale failure of a detector, and accounts for $<1\%$ of the data collected from each individual system, or $<5\%$ of the total ATLAS data [56]. The good LB's are indicated in an XML file called a Good Run List, or GRL. The GRL recommendations for Run II are provided for each of the four data-taking years [58], and the GRL's themselves include the good and bad LB's for each data-taking run [59–62].

The reconstruction process involves taking the RAW data as input and creating physics objects. Localized energy depositions in the calorimeters are reconstructed as jets, tracks from the ID and hits in the muon spectrometer are reconstructed as muons, and ID tracks and energy depositions in the EM calorimeter are reconstructed as electrons or photons. Then, calibration is performed to correct the absolute energy and energy resolution of objects such as jets and electrons [63, 64]. This is discussed more in Chapter 5. This analysis uses the ATLAS software release 21 for reconstruction.

The current state of the collisions, such as the location of the beam spot and the noise, is recorded in a separate data stream and applied to the physics data stream to increase reconstruction accuracy. For example, a noisy channel appearing in a certain LB will be masked, and its information will be disregarded during further analysis. The LB is the smallest data-taking window that can be rejected [56].

The data files with physics objects, AODs, are very large and contain a variety of events. A derivation step is performed to reduce the file size and to select only events relevant to a particular set of analyses. For this analysis, the AnalysisTop

software is used for derivation, producing a DAOD [65]. The DAOD used for both the Higgs Oblique and the HBSM analysis is TOPQ1. An analysis-specific framework is then used to further refine the variables kept for analysis, as well as to add options such as multivariate output variables. Both data and MC ROOT files are identical in structure, and are analyzed in a similar manner.

4.3 Monte Carlo

MC samples are used both to simulate the signal of interest as well as model many of the physics backgrounds. MC samples allow for both the refinement of an analysis as well as an expected result that can be compared to data.

Modeling a pp collision involves Standard Model predictions and simulation of particle interactions with the detector. The latter is well-modelled using the GEANT4 toolkit [66]. The former is challenging because there are both perturbative and non-perturbative regimes in QCD [67]. Normally there is a hard parton-parton scatter followed by parton radiation showering followed by hadronization followed by ATLAS detector simulation.

4.3.1 Matrix Element Calculation and PDFs

The matrix element (ME) calculation takes as an input the center-of-mass energy, a parton distribution function, and the type of partons involved in the hard scatter interaction. Different models yield more accurate results for different types of interactions, so separate samples are produced for each interaction type. For the $t\bar{t}t\bar{t}$ signal, the matrix element calculation is performed at NLO. A k-factor, or the ratio of the NLO cross section to the LO cross section, is not needed as the calculated cross section is already at NLO. The $t\bar{t}t\bar{t}$ leading order cross section is $7.59^{+64\%}_{-36\%}\text{fb}$, which when NLO corrections are applied becomes $11.97^{+18\%}_{-21\%}$. Even higher order corrections scale approximately with the QCD and EW coupling constants, 0.1 and 0.01 respectively [68]. This means that corrections of higher order than NLO are relatively small, and the cross section calculation at NLO is reasonably accurate.

The hard process generators used in this analysis are MadGraph_aMC@NLO [69], PowHegBox v2 (NLO) [70], and Sherpa 2.2.[71]. For each sample, one nominal generator is used, and alternative samples may also be produced with the other generators. These alternative samples are used to quantify the generator uncertainty.

The cross section for a hard process explicitly depends on the initial momen-

tum of the interacting partons. The probability that an interaction will occur for a specific type of parton containing some fraction x of the proton’s longitudinal momentum is known as a parton distribution function (PDF). PDF’s are non-perturbative and determined empirically by fits to experimental measurements [72]. The PDF sets used for this analysis are NNPDF3.1NLO, NNPDF2.3NLO, NNPDF3.0NLO, NNPDF3.0nlof4 [73], and MMHT2014LO [74]. The NNPDF sets are created using neural-network optimization of parameters, fitting to a broad range of hard scattering processes. Examples include cross sections collected from deep inelastic scattering experiments, vector boson production data, and Drell-Yan production data [73]. The MMHT2014LO PDF set performs a similar fit, but parameterization is based on a different fitting formula using Chebyshev polynomials [74].

4.3.2 Parton showering

The output of the generator calculation is information about the final state partons of the produced hard scatter. This is then passed to a parton showering simulator. The parton showering simulator connects the high-energy, perturbative hard scattering regime to the low-energy, non-perturbative hadronization regime [75]. The probability of a decay is analytically solvable, but has divergences in edge cases: for example, the decay of a particle into two exactly collinear particles. These divergences are resolved by placing thresholds on energy and opening angle resolution. Practically speaking, this corresponds to the fact that the detector is not infinitely sensitive and can only measure particle energies and particle separations above certain values. The MC generator constructs a branching probability at each iteration. Initial state radiation is modelled in the same way, and summed with final state radiation to yield a realistic estimate of the parton showering.

Hadronization refers to the confinement of low-energy partons into hadrons. This process is non-perturbative and cannot be calculated from QCD first principles. Instead, models such as the Lund string model track the input partons and hadronize them according to low energy data [15]. The probability of a particular hadronization final state is a function of properties such as the distance between partons.

In addition to parton shower modelling, the underlying event, initial state radiation, and final state radiation must be modelled. The free parameters in underlying event modelling are optimized to agree well with data, and a particular set of parameters is known as a “tune” [76]. The Pythia8 showering package is used with the A14 tune for this analysis [77]. Sherpa is also used for parton showering, with dedicated tuning developed by Sherpa authors [71].

4.3.3 Samples and Alternative Samples

For the HBSM analysis (1LOS channels), Tables 4.3 and 4.4 show the MC signal and background samples used. Two separate $t\bar{t}t\bar{t}$ samples are produced. One is the SM $t\bar{t}t\bar{t}$ sample, normalized to the SM predicted NLO cross section of 12fb. The other is the BSM $t\bar{t}t\bar{t}$, where $t\bar{t}$ is produced in association with a Heavy Higgs (see Figure 2.7). This process is also normalized to 12fb, following the excess observed in the previous $t\bar{t}t\bar{t}$ measurement [24]. The main background, $t\bar{t} + jets$, includes $t\bar{t} + \geq 1c$, $t\bar{t} + \geq 1b$, and $t\bar{t} + light$. After it is produced, the $t\bar{t} + jets$ sample is divided into its three constituents as described in Section 7.1.2.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
312440-46	BSM $t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
500326	SM $t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
410470 410472	$t\bar{t} + jets$	PowhegBox v2	Pythia 8.230	GEANT4
411068-69	$t\bar{t}b\bar{b}$	PowhegResBox	Pythia 8.230	GEANT4
410658-59 410644-45	singletop	PowhegBox v2	Pythia 8.230	GEANT4
410646-47	tW	PowhegBox v2	Pythia 8.230	GEANT4
700168 700205	$t\bar{t} + W$	Sherpa 2.2.1	Sherpa 2.2.1	GEANT4
346344-45	$t\bar{t} + H$	PowhegBox v2	Pythia 8.230	GEANT4
413023	$t\bar{t} + Z$	Sherpa 2.2.1	Sherpa 2.2.1	GEANT4
410408	$t + W + Z$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
410560	$t + Z$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
410081	$t\bar{t} + WW$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
304014	$t\bar{t}t$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
364170-83 364156-59 364184-97 364114-27 364100-13 364128-41 364204-09 364198-203 364210-15	$V + jets$	Sherpa 2.2.1	Sherpa 2.2.1	GEANT4

Table 4.3: Summary of MC samples for HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4). Table 1 of 2.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
500460-63 410081	$t\bar{t}+VV$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
410389	$t\bar{t} + \gamma$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
366140-54	$V + \gamma$	Sherpa 2.2.4	Sherpa 2.2.4	GEANT4
342284-85	$V + H$		Pythia 8.230	GEANT4
364250 364253-55 364283-85 364287-90 345705-06 345723 363355-60 363489 363494	VV	Sherpa 2.2.2 Sherpa 2.2.1	Sherpa 2.2.2 Sherpa 2.2.1	GEANT4
364242-49	VVV	Sherpa 2.2.2	Sherpa 2.2.2	GEANT4
410156-57 410218-20 410276-78	$t\bar{t}+X$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4

Table 4.4: Summary of MC samples for HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4). Table 2 of 2.

In order to quantify the systematic uncertainty associated with the simulation process, the nominal samples are compared to alternative samples, which are produced using a different generator or parton shower choice. The difference in yield is considered as the uncertainty on the generator/parton shower machinery. The alternative samples are shown in Table 4.5.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
700046 700075 700355 412043 412044 412115	SM $t\bar{t}t\bar{t}$	Sherpa 2.2.10 Sherpa 2.2.11 MadGraph5_aMC@NLO MadGraph5_aMC@NLO MadGraph5_aMC@NLO	Sherpa 2.2.10 Sherpa 2.2.11 Pythia 8.230 Herwig 7 Pythia 8.230	GEANT4
410464-65 412066-71 407348-50 411233-34 410557-58 411082-87 407354-56 410250-52	$t\bar{t}$ +jets	MadGraph5_aMC@NLO PowhegBox v2 Sherpa 2.2.10	Pythia 8.230 Herwig 7 Sherpa 2.2.10	GEANT4
411032-35 412004 412005	singletop	PowhegBox v2 MadGraph5_aMC@NLO	Herwig 7 Pythia 8.230	GEANT4
410654-55 411036-37 412002	tW	PowhegBox v2 PowhegBox v2 MadGraph5_aMC@NLO	Pythia 8.230 Herwig 7 Pythia 8.230	GEANT4
410155 412123 413008 700000	$t\bar{t}+W$	MadGraph5_aMC@NLO Sherpa 2.2.1 Sherpa 2.2.8	Pythia 8.230 Sherpa 2.2.1 Sherpa 2.2.8	GEANT4
345940-42	$t\bar{t}+H$	MadGraph5_aMC@NLO	Pythia 8.230	GEANT4
410142	$t\bar{t}+Z$	Sherpa 2.2.10	Sherpa 2.2.10	GEANT4
361100-08 346413	V +jets	PowhegBox v2	Pythia 8.230	GEANT4

Table 4.5: Summary of alternative MC samples for the HBSM analysis. All events were generated with PDF set NNPDF3.0NLO, except SM $t\bar{t}t\bar{t}$ (NNPDF3.1NLO), $t\bar{t}b\bar{b}$ (NNPDF3.0nlof4), and the t-channel for singletop (NNPDF3.0nlof4).

The samples used in the Higgs Oblique analysis (SSML channels) are shown in Tables 4.6, 4.7, and 4.8. For the Higgs Oblique analysis, there are five $t\bar{t}t\bar{t}$ samples produced: one corresponding to SM $t\bar{t}t\bar{t}$ production, and four samples at different values of $\hat{H}$. Some of the samples used the AltFastII simulation for detector response

[78]. This simulation performs a faster, less thorough, approximation of the energy response and distributions in the calorimeters, which is the most time-consuming part of the GEANT4 simulation.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
510087	$t\bar{t}\bar{t}\bar{t} \hat{H} = 0.05$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
510088	$t\bar{t}\bar{t}\bar{t} \hat{H} = 0.08$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
510089	$t\bar{t}\bar{t}\bar{t} \hat{H} = 0.12$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
510090	$t\bar{t}\bar{t}\bar{t} \hat{H} = 0.17$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
514721	$t\bar{t}\bar{t}\bar{t} \hat{H} = 0$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
412043	$t\bar{t}\bar{t}\bar{t}$	MadGraph5_aMC@NLO	Pythia 8.230	AltFast II
304014	3top	MadGraph5_aMC@NLO	Pythia 8	
410081	$t\bar{t}WW$	MadGraph5_aMC@NLO	Pythia 8	
505074-76 506085-87 700309 410156-57 410218-20 410276-78	$t\bar{t}ll$	MadGraph5_aMC@NLO Sherpa 2.2.11	Pythia 8 Sherpa 2.2.11	
500460-63	$t\bar{t}VV$	MadGraph5_aMC@NLO	Pythia 8	
410658-59 410646-47 410644-45	singletop	PowhegBox v2	Pythia 8	
410560 410408	tZ tWZ	MadGraph5_aMC@NLO MadGraph5_aMC@NLO	Pythia 8 Pythia 8	
410470 410472 407342-44	$t\bar{t}$	PowhegBox v2	Pythia 8	GEANT4
410389	$t\bar{t}\gamma$	MadGraph5_aMC@NLO	Pythia 8	
346344-45	$t\bar{t}H$	PowhegBox v2	Pythia 8	GEANT4
700205 700168	$t\bar{t}W$	Sherpa 2.2.10	Sherpa 2.2.10	GEANT4
366140-54	$V\gamma$	Sherpa 2.2.4	Sherpa 2.2.4	
342284-85	VH	Pythia 8	Pythia 8	
410156-57 410218-20 410276-78	$t\bar{t}Z$	MadGraph5_aMC@NLO	Pythia 8	GEANT4

Table 4.6: Summary of nominal MC samples for the SSML Higgs Oblique analysis.
Table 1 of 2.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
364170-83 364156-59 364184-97 364114-27 364100-13 364128-41 364204-09 364198-203 364210-15	V +jets	Sherpa 2.2.1	Sherpa 2.2.1	GEANT4
364250 364253-55 364283-85 364287-90 345705-06 345723 363355-60 363489 363494	VV	Sherpa 2.2.2 Sherpa 2.2.1	Sherpa 2.2.2 Sherpa 2.2.1	GEANT4
364242-49	VVV	Sherpa 2.2.2	Sherpa 2.2.2	

Table 4.7: Summary of nominal MC samples for the SSML Higgs Oblique analysis.
Table 2 of 2.

Sample Production				
Sample DSID	Process	Generator	Shower	Detector Sim.
700075 700355 412044	$t\bar{t}t\bar{t}$	Sherpa 2.2.10 Sherpa 2.2.11 MadGraph5_aMC@NLO	Sherpa 2.2.10 Sherpa 2.2.11 Herwig 7	
411032-33 411036-37 411034-35 412004 412002 412005	singletop	PowhegBox v2 MadGraph5_aMC@NLO	Herwig 7 Pythia 8	
345940-41	$t\bar{t}H$	MadGraph5_aMC@NLO	Pythia 8	AltFastII
600793-96 600787-90 410155 412123 700000	$t\bar{t}W$	PowhegBox v2 PowhegBox v2 MadGraph5_aMC@NLO Sherpa 2.2.8	Pythia 8 Herwig 7 Pythia 8 Sherpa 2.2.8	GEANT4
361100-08 700320-34 700338-49	V +jets	PowhegBox v2 Sherpa 2.2.11	Pythia 8 Sherpa 2.2.11	

Table 4.8: Summary of alternative MC samples for the SSML Higgs Oblique analysis.

CHAPTER 5

Data Reconstruction

5.1 Introduction

Reconstruction involves translating the digitized signals from the ATLAS detector into physics objects that can be used for analysis. Object reconstruction is challenged by contributions from pileup, noise, and fluctuations in the physical processes that give rise to the signal being reconstructed. The goal of reconstruction is to provide well-measured, well-defined objects with relatively high efficiency.

5.2 Electrons

Creating electron objects proceeds through several steps[79]. First, electron candidates are found in the reconstruction process. Identification and isolation criteria are subsequently applied. For this analysis, additional selections are used to ensure the quality of reconstructed electrons is high.

5.2.1 Reconstruction

As charged particles, electrons produce curved tracks within the inner detector (ID) and localized energy deposits within the EM calorimeter. The first part of this process involves the reconstruction of these tracks and energy deposits.

A standard track-pattern reconstruction is used for all parts of the ID[80]. Hits in the silicon detectors create seeds for the track, which are represented in the reconstruction algorithm as points in 3-dimensional space. Starting from the IP, hits are added iteratively moving outwards using a combinatorial Kalman filter[81].

Meanwhile, clusters are reconstructed in the EM calorimeter using a topological cluster algorithm[82]. The energy of each calorimeter cell is assessed relative to the expected background noise. If the cell energy is above a certain threshold, t_{seed} , then the cell acts as a seed for a proto-cluster. Every neighboring cell is assessed; if its energy is above a threshold $t_{neighbor}$, then it is added to the proto-cluster, and its neighbors are considered in the next iteration. If a neighbor cell is below $t_{neighbor}$

but above a threshold t_{cell} , then it is added to the proto-cluster but its neighbor cells are not considered in the next iteration. This continues iteratively until there are no more cells satisfying the threshold requirements.

There are also conditions governing the boundaries of the proto-clusters. If a cell satisfying $t_{neighbor}$ is directly adjacent to two seeds, then the proto-clusters are merged. If a proto-cluster has two or more maxima with energy $> 500\text{MeV}$ and at least four lower energy neighbors, then the proto-cluster is split.

The expected background energy is calculated as the RMS of the expected electronic noise, which depends on gain and other detector settings at the time of data acquisition. The background energy also incorporates the expected noise from pileup, calculated from the average instantaneous luminosity for Run II. For electrons, $t_{seed} = \frac{E_{signal}}{E_{background}} \geq 4$, $t_{neighbor} \geq 2$, and there is no threshold requirement for t_{cell} . This is known as the “4-2-0” clustering strategy.

To accommodate larger electron showers, superclusters are also constructed from the EM calorimeter deposits. The inputs for a supercluster are the EM topo-clusters. Starting with the topo-cluster with the highest E_T , each topo-cluster is assessed, and becomes a supercluster seed if $E_T > 1\text{GeV}$ and it is matched to a track with ≥ 4 hits in the silicon detectors. For each supercluster seed, additional topo-clusters within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are added as satellites. If a topo-cluster within $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ shares a best-matched track with the supercluster seed, it is also added to the supercluster. These satellites represent the secondary EM showers created by the electron.

Once tracks and clusters, and superclusters have been reconstructed, they form ROI’s that can be matched. Matching is based on the closest point between the track and the location of the cluster in the second layer of the EM calorimeter, and requires that $|\Delta\eta| < 0.05$ and $-0.10 < q(\phi_{track} - \phi_{clus}) < 0.05$, where q is the reconstructed track charge.

Once the electron candidates are reconstructed, their energy must be calibrated. This ensures that the signal from the calorimeter, which is a proportional analog pulse, is recovered with an accurate value for the energy [64]. The calibration first estimates the energy of calorimeter deposits using a multivariate regression algorithm, then corrects for: different energy scales in different calorimeter layers, effects of non-uniformities such as those at detector boundaries, and for overall energy scale. For the overall energy scale, data selected to represent the $Z \rightarrow e^+e^-$

process are compared to simulated samples of the same process. The selection is of events centered around the Z mass of 91GeV.

For all topological clusters, including those used for reconstructing jets as described in Section 5.4, local cell weighting (LCW) is applied. This is intended to correct for calorimeter non-compensation, or discrepancies in EM and hadronic energy deposition [83]. For each topo-cluster, the probability P_{clus}^{EM} is calculated. This represents the probability that the topo-cluster is electromagnetic in origin, where $0 \leq P_{clus}^{EM} \leq 1$. The probability is calculated using the average cell signal density and the depth at which the topo-cluster originated. Cells within the topo-cluster are weighted according to their probability:

$$w_{cell}^{cal} = P_{clus}^{EM} w_{cell}^{EM} + (1 - P_{clus}^{EM}) w_{cell}^{had} \quad (5.1)$$

The weight is determined by the ratio of the energy deposition relative to the electromagnetic energy deposition. The weights and the probabilities are determined using MC samples of neutral pions (for electromagnetic topo-clusters) and charged pions (for hadronic topo-clusters).

5.2.2 Identification

Once electrons are reconstructed, the identification process filters electron candidates for quality [79]. Identification is performed by a likelihood (LH) discriminant, which takes as input variables associated with the development of the EM shower in the calorimeter, the reconstruction of the track in the ID, and the quality of the track/cluster match. The input variables are listed below in Table 5.1. The likelihood for signal (background) is given as the product of the probability distribution functions P for each variable x , in each bin i :

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^N P_{S(B),i}(x_i) \quad (5.2)$$

Electron Likelihood Inputs		
Category	Description	Name
Hadronic leakage	Ratio of first layer hadronic cal. E_T to E_T of EM cluster	R_{had_1}
	Ratio of entire hadronic cal. E_T to E_T of EM cluster	R_{had}
EM third layer	Ratio of third layer EM cal. energy to total EM cal. energy	f_3
EM second layer	Ratio of the sum of energies for a window of $3 \times 7 \eta \times \phi$ cells to the sum of energies for a window of 7×7 cells	R_η
	Lateral shower width as a function of where and how much energy is deposited. For a window of $3 \times 5 \eta \times \phi$, width is expressed as $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$	w_{η_2}
	Ratio of the sum of energies for a window of $3 \times 3 \eta \times \phi$ cells to the sum of energies for a window of 3×7 cells	R_ϕ
EM first layer	Total lateral shower width, calculated using energy and cell distance from the highest energy cell, i_{max} . Index i runs over a window of $\Delta\eta \approx 0.0625$. Expressed as $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$	w_{stot}
	Ratio of energy difference between highest cell energy and second-highest cell energy to the energy sum of the two maxima, $\frac{(E_1 - E_2)}{(E_1 + E_2)}$	E_{ratio}
Track conditions	Number of hits in innermost pixel layer	$n_{innermost}$
	Number of hits in the pixel detector	n_{Pixel}
	Total number of hits in pixel and SCT detectors	n_{Si}
	Transverse impact parameter	d_0
	Significance of d_0 , where $\sigma(d_0)$ is the d_0 uncertainty	$ d_0/\sigma(d_0) $
	Momentum lost by track between the perigee and the last measurement point, divided by the momentum measured at perigee	$\Delta p/p$
	LH probability based on transition radiation in the TRT	eProbabilityHT
Track-cluster matching	$\Delta\eta$ between cluster position in the first layer of the EM cal. and the track	$\Delta\eta_1$
	$\Delta\phi$ between cluster position in the second layer of the EM cal. and the charge q times the momentum-rescaled track, extrapolated from the perigee	$\Delta\phi_{res}$
	Ratio of cluster energy to measured track momentum	E/p

Table 5.1: List of input variables for electron identification likelihood evaluation.

The discriminating variable is d_L , which is the natural logarithm of the ratio of the electron signal likelihood to the non-electron background likelihood:

$$d_L = \ln \frac{L_S}{L_B} \quad (5.3)$$

The values of d_L are set in nine bins of an $|\eta|$ distribution by 12 bins of an E_T distribution in order to meet a fixed electron selection efficiency, averaging 93%, 88%, and 80% for the Loose, Medium, and Tight Working Points (WP's) respectively. For this analysis, electrons are required to pass the 'tightLH' selection. Electrons passing the Loose criteria are used for overlap removal.

5.2.3 Isolation

There are requirements for both calorimeter cluster isolation and for track isolation. Isolation is expressed as the amount of energy present within a cone of radius ΔR that does not come from the reconstructed object. For cluster isolation, the total energy around the cluster $E_{T,raw}^{isol}$ is calculated by summing the transverse energy of all positive energy topo-clusters whose center is located within ΔR from the center of the electron cluster. Then, the energy of the electron cluster is subtracted. This is done by subtracting the energy from the cells in a 5 cell $\times$ 7 cell, or $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$, rectangle around the center of the electron cluster. There is extra "leakage" energy, which is calculated using MC samples without pileup, and this is also subtracted. There is also a correction for energy deposited by pileup and underlying events. These are all subtracted from $E_{T,raw}^{isol}$ to determine how much energy from other physics objects is present within the ΔR cone.

Track isolation is performed in a similar way to energy cluster isolation, wherein the p_T of tracks within a cone is summed, excluding the p_T from the actual electron track of interest. However, the cone size is variable to accommodate the fact that, in high-momentum decays, the resulting electrons may have accompanying decay products nearby. Thus for larger p_T , the cone size shrinks.

The electrons used in this analysis are required to pass the fixed-cut (FC) Tight criteria. These isolation requirements set limits on the value of E_T^{isol} evaluated for radius $\Delta R = 0.20$ (topoetcone20) and p_T^{isol} evaluated for radius $\Delta R = 0.30$ (ptvarcone30) [84]. There are also additional requirements on the track-to-vertex association (TTVA) for the track. The FCTight requirements are $E_T^{cone20}/p_T < 0.06$ for the calorimeter isolation and $ptvarcone30_TightTTVA_pt1000/p_T < 0.06$ for the

track isolation, where p_T refers to the momentum of the track or cluster.

5.2.4 Other Criteria

Electrons must pass a p_T threshold of 28GeV and $|\eta|$ region cuts excluding the LAr crack region: $|\eta| < 1.37$ and $1.57 < |\eta| < 2.47$. Candidates from bad calorimeter clusters are removed, as well as candidates associated with certain known bugs, according to the recommendations of EGamma group[85]. To ensure the quality of electron tracks, the TTVA tool is employed, using transverse and longitudinal impact parameters. These ensure that the vertex from which the electron originated is sufficiently close to the IP. These are listed in Table 5.2. For dilepton events involving electrons, further filtering is performed to remove events containing electrons with misidentified charge. This is done using the Electron Charge ID Selector Tool (ECIDS)[86]. The ECIDS tool uses a Boosted Decision Tree (described in Chapter 8) with electron track variables as input.

Tight Electron Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 1.37$ or $1.52 < \eta < 2.47$
Transverse momentum	$p_T > 28\text{GeV}$
Track to vertex association	$ d_0^{BL}significance < 5$ and $ \Delta z_0^{BL}sin\theta < 0.5\text{mm}$
Identification	TightLH
Isolation	FCTight
Charge ID	pass ECIDS (for $e^\pm e^\mp$ and $e^\pm \mu^\mp$ only)
Other object quality	not from bad calorimeter cluster

Table 5.2: List of requirements for reconstructed tight electrons for HBSM $t\bar{t}t\bar{t}$ analysis.

5.3 Muons

5.3.1 Reconstruction

Muon reconstruction mostly incorporates tracks reconstructed in the ID and the MS[87]. In some cases, such as a muon with large energy loss, information from the calorimeters may also be considered. ID track reconstruction for muons is similar to that for electrons. MS track reconstruction begins with the individual hits in the MS detectors, which are used to construct short, straight lines via a Hough Transform[88]. The straight lines are assembled into a larger scale curved track

Tight Electron Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 1.37$ or $1.52 < \eta < 2.47$
Transverse momentum	$p_T > 15\text{GeV}$
Track to vertex association	$ d_0^{BL} \text{significance} < 5$ and $ \Delta z_0^{BL} \sin\theta < 0.5\text{mm}$
Identification	TightLH
Isolation	PLImprovedTight
Charge ID	pass ECIDS (for $e^\pm e^\mp$ and $e^\pm \mu^\mp$ only)

Table 5.3: List of requirements for reconstructed tight electrons for Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

using the IP location and a parabolic approximation of the muon’s curved path in the magnetic field. The curve occurs in a 2-dimensional plane, with some additional 3rd dimensional coordinates that are added using trigger detector measurements. This creates 3-dimensional track candidates. These candidates are fit to a curve using a global χ^2 that incorporates MS tracks, ID tracks, muon interactions with other detector materials, and detector misalignments. This is an iterative process that refines which MS hits are used to reconstruct the tracks, and which selects higher quality tracks when two tracks are in close proximity.

There are five methods of reconstructing muon candidates: combined (CB), inside-out combined (IO), muon-spectrometer extrapolated (ME), segment-tagged (ST), and calorimeter-tagged (CT). CB candidates are reconstructed by matching MS tracks to ID tracks. IO candidates are created in a reverse process, matching ID tracks to MS tracks. ME candidates are formed when no ID track match can be found for an MS track. In this case, the MS track is extrapolated by itself to the IP. The ME method is useful for high $|\eta|$ (up to $|\eta| = 2.7$). Muon ST candidates are formed when there is tight matching between ID and MS tracks. In this case, track information is taken from only the ID track. The CT method reconstructs muon candidates that deposit their energy in the calorimeter, and uses ID and calorimeter data. Using these methods, the track is extrapolated back to the IP, including estimated energy loss in the calorimeters, and the p_T at the IP is estimated.

CB muons are used to correct the momentum of the muons [89]. The correction factors are derived using a fit to the invariant mass of $J/\Psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ samples to data, and are a function of both the momentum smearing (the limits on p_T resolution) and momentum scaling. Some muon momentum corrections also

account for a sagitta bias, a momentum bias in μ^+ vs. μ^- tracks. However, this analysis does not incorporate this bias correction.

5.3.2 Identification

This analysis uses muons satisfying a Medium identification requirement, which is designed to accept high quality prompt muons and reject backgrounds such as misidentified light hadrons. The Medium reconstruction places restrictions on four different parameters. The first is $\frac{q}{p}$ significance. The charge (q)/momentum (p) ratio recovered from the ID is compared to the ratio recovered from the MS:

$$\frac{q}{p} \text{significance} = \frac{|q/p_{ID} - q/p_{MS}|}{\sqrt{\sigma^2(q/p_{ID}) + \sigma^2(q/p_{MS})}} \quad (5.4)$$

where $\sigma(q/p_{ID})$ is the uncertainty associated with the ID ratio, and $\sigma(q/p_{MS})$ is the uncertainty for the MS ratio. The second parameter is the number of precision stations, which are defined as MS stations where the muon has at least three hits in either the MDT or CSC. Similarly, a precision hole station is defined as a MS station where the muon has less than three hits but is missing three or more expected hits based on the muon trajectory. The last parameter used is the $|\eta|$ of the track.

Medium identification muons are required to be either CB or IO muons. The charge momentum ratio must be $\frac{q}{p} > 7$. For the region $0.1 < |\eta| < 2.5$, muon candidates are required to have at least two precision stations. For the region $|\eta| < 0.1$, muon candidates must have at least one precision station and no more than one precision hole station.

5.3.3 Isolation

Muon isolation is performed in a similar way to electron isolation, with selections on the ID track isolation[90]. In particular, the isolation follows the requirement $\text{ptvarcone30_TightTTVA_pt1000}/p_T < 0.06$, referring to a p_T isolation cut calculated with radius $\Delta R = 0.3$.

5.3.4 Other Criteria

To select high-quality muon objects, muons must pass a p_T threshold (28GeV for the HBSM analysis, 15GeV for the Higgs Oblique analysis) and an $|\eta|$ selection of $|\eta| < 2.5$. The track-to-vertex tool places cuts on the transverse and longitudinal impact

parameters, ensuring that the vertex from which the muon originated is sufficiently close to the IP. These requirements are listed in Table 5.4 for the HBSM analysis and Table 5.5 for the Higgs Oblique analysis. Besides the p_T threshold, the two analyses have slightly different isolation requirements for muons.

Tight Muon Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 2.5$
Momentum calibration	Sagitta bias correction not used
Transverse momentum	$p_T > 28\text{GeV}$
Track to vertex association	$ d_0^{BL}significance < 3$ and $ \Delta z_0^{BL}sin\theta < 0.5\text{mm}$
Identification	Medium
Isolation	FCTightTrackOnly

Table 5.4: List of requirements for reconstructed tight muons in HBSM $t\bar{t}t\bar{t}$ analysis.

Tight Muon Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 2.5$
Momentum calibration	Sagitta bias correction not used
Transverse momentum	$p_T > 15\text{GeV}$
Track to vertex association	$ d_0^{BL}significance < 3$ and $ \Delta z_0^{BL}sin\theta < 0.5\text{mm}$
Identification	Medium
Isolation	PLImprovedTight

Table 5.5: List of requirements for reconstructed tight muons in Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

5.4 Small Radius Jets

5.4.1 Particle Flow Reconstruction

Particle flow (PFlow) reconstruction incorporates both calorimeter and track information and is intended to be an improvement over reconstruction using only the EMTopo clusters [91]. While EMTopo jets rely on calorimeter data to reconstruct jet energy, PFlow jets make use of track momentum measurements to increase the accuracy of jet energy measurements from charged particles. The energy of charged particles is subtracted from the calorimeter, leaving only the energy of neutral par-

ticles. The total PFlow jet energy is the sum of charged-particle track energy and neutral-particle calorimeter deposits.

The topo-clusters are constructed from calorimeter cells according to the usual method [83]. Topo-clusters are matched with ID tracks according to their $\Delta\eta$ and $\Delta\phi$. The expected average energy deposition, $\langle E_{dep} \rangle$, from a particle with track momentum p^{trk} is calculated as

$$\langle E_{dep} \rangle = p^{trk} \langle E_{ref}^{clus} / p_{ref}^{trk} \rangle \quad (5.5)$$

The ratio $\langle E_{ref}^{clus} / p_{ref}^{trk} \rangle$ is the mean response. The response is calculated using single-pion samples without pileup [92]. The topo-cluster energy E_{ref}^{clus} is summed in a $\Delta R = 0.4$ cone around the track position, which has been extrapolated to the second layer of the EM calorimeter.

Once calculated, the energy $\langle E_{dep} \rangle$ is subtracted from calorimeter cells, starting with the region called the layer of highest energy density (LHED). The LHED is determined as the layer of the calorimeter with the (1) largest rate of increase in energy density as a function of layer depth and (2) closest proximity to the extrapolated position of the track that was matched to the energy deposit. Rings are created around the extrapolated track position of the LHED, based on the shower profile for a pion sample. Subtraction takes place cell-by-cell in these rings. This process removes the expected calorimeter energy depositions coming from charged-particle tracks. If the remaining energy is less than the uncertainty $1.5\sigma E_{dep}$, it is assumed to be from fluctuations in deposition from the charged-particle track, and the energy is removed. If the remaining energy is greater than $1.5\sigma E_{dep}$, it is assumed to be deposited by other (neutral) particles and the energy is kept.

For well-isolated and well-matched ID tracks and topo-clusters from charged particles, all of the energy E_{dep} is subtracted from the topo-cluster, and the jet momentum information is determined solely by the ID track. For topo-clusters without matches, with neutral-particle contributions, or for ID tracks that form multiple topo-clusters, jet energy may be determined both by ID tracks and partially subtracted topo-cluster energy, or entirely by topo-cluster energy.

In particular, high- p_T tracks are often near other high- p_T tracks in the core of jets, which makes associating a single track and single topo-cluster difficult. For $p_T^{trk} > 100\text{GeV}$, no track information is used and the PFlow jet is identical to the

EMTopo jet. For $p_T^{trk} < 100\text{GeV}$, the track information is not used if the energy in the cluster, compared to the average energy deposition of a pion, $\langle E_{dep} \rangle$, is greater than a preset value scaled by the track momentum:

$$\frac{E^{clus} - \langle E_{dep} \rangle}{\sigma(E_{dep})} > 33.2 \times \log_{10}(40\text{GeV}/p_T^{trk}) \quad (5.6)$$

where $\sigma(E_{dep})$ is the standard deviation of the pion deposition energy.

The PFlow reconstruction process ends with combining any remaining energy in the topo-clusters with the track momentum.

5.4.2 Jet Clustering

The anti- k_T algorithm takes the PFlow jet constituents (ID tracks and modified-energy topo-clusters) as inputs and constructs conical jets of a specified radius, R [93]. Common choices of R include $R=0.4$ for smaller jets and $R=1$ for larger jets. The essence of the algorithm is to add constituents i and j based on their p_T (k_{ti} and k_{tj} , respectively) and distance Δ_{ij} , such that they represent a minimum parameterized by d_{ij} :

$$d_{ij} = \min \left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2} \right) \frac{\Delta_{ij}^2}{R^2} \quad (5.7)$$

This condition prioritizes clustering between the hard jet “seed” and the closest soft particles surrounding it. If there are two or more jet seeds within radius R , then two jets will be formed, and their relative hardness will determine how the boundary between them is defined. If the p_T between seed 1 and seed 2 is equal, $k_{t1} = k_{t2}$, then they will be separated by an equidistant line. When they are not equal, the higher p_T jet seed will claim more area, up to the extreme case $k_{t1} \gg k_{t2}$, where the jet formed by seed 1 will be a complete cone and jet 2 will not occupy any calorimeter cells or tracks in the overlapping region.

5.4.3 Jet Calibration

Several corrections are applied to the jets reconstructed from data[91]. First, extra energy from pileup collisions is removed from the jets. This process is determined by jet area and p_T density in the jet, and then further corrected using comparisons to

MC. Next, three calibrations are performed. The first is a correction to the jet energy and direction. This accounts for energy loss in non-detecting material, as well as resolution limitations and transitions in detector geometry that might bias the jet $|\eta|$. The second calibration is performed to correct for the fact that jets originating from different particles have different deposition patterns in the calorimeter, and therefore the reconstructed jet energy has a dependence on jet flavor prior to calibration. The third calibration is applied to data only. Since the data are calibrated to MC samples, the in situ calibration applies a correction, c , to account for differences in data and MC:

$$c = \frac{R_{insitu}^{data}}{R_{insitu}^{MC}} \quad (5.8)$$

where R_{insitu}^{data} and R_{insitu}^{MC} are the jet responses measured in data and MC respectively. The in situ calibration uses specific reference events, such that the response is defined as the average ratio of jet p_T to reference object p_T .

5.4.4 Other Jet Selections

There is a selection requiring jet $p_T > 25\text{GeV}$ and $|\eta| < 2.5$. Events containing one or more jets of quality LooseBad are also removed by jet cleaning. Reconstructed jets at this stage may still be affected by pileup, particularly the presence of jets produced in simultaneous collisions. A tool called the Jet Vertex Tool (JVT) is used to identify and remove pileup contributions to jets [94, 95]. This tool is an improvement upon the Jet Vertex Fraction (JVF) used previously to identify pileup jets. The JVF uses the sum of momentum over all tracks in the jet that originated at the primary vertex, $\sum p_T^{hard}$. This is expressed as a fraction of total track momentum in the jet, including tracks originating elsewhere, p_T^{PU} :

$$JVF = \frac{\sum p_T^{hard}}{\sum p_T^{hard} + \sum p_T^{PU}} \quad (5.9)$$

It is apparent that the total jet p_T increases as the amount of pileup, n_{trk}^{PU} , increases. Thus the efficiency of a JVF cut is dependent on the amount of pileup. The JVT is an improvement because it remains stable over different amounts of pileup. The JVT is a multivariate combination of two track-based variables: $corrJVF$ and R_{pT} . The first variable, $corrJVF$, is similar to JVF but the p_T of the non-hard-

scatter constituent tracks is scaled by the total amount of pileup, n_{trk}^{PU} :

$$corrJVF = \frac{\sum p_T^{hard}}{\sum p_T^{hard} + \frac{\sum p_T^{PU}}{kn_{trk}^{PU}}} \quad (5.10)$$

where k is set to 0.01 and is roughly the scaling factor of $\langle p_T^{PU} \rangle / n_{trk}^{PU}$. The second variable, R_{pT} , is fraction of momentum coming from all hard tracks in the jet compared to the total calibrated momentum of the jet:

$$R_{pT} = \frac{\sum p_T^{hard}}{p_T^{jet}} \quad (5.11)$$

Because pileup jets have very few constituent tracks from the hard-scatter vertex, most pileup jets will have an R_{pT} value close to zero. The distribution of R_{pT} for pileup and hard-scatter jets is different, and offers additional discrimination power. The JVT combines these two variables in a 2-dimensional plane and uses a nearest-neighbor algorithm to calculate a likelihood and to identify pileup jets [96]. This analysis uses a Tight WP, defined for EMPFlow jets as $JVT > 0.5$, with an average efficiency of 96%[97]. This is applied to all jets with $20 < p_T < 60\text{GeV}$ and $|\eta| < 2.4$.

Jet Selection	
Feature	Requirements
Jet constituents	EMPFlow
Jet clustering algorithm	anti- k_t
Pseudorapidity range	$ \eta < 2.5$
Transverse momentum	$p_T > 25\text{GeV}$
JVT	JVT > 0.5
Jet quality	events with LooseBad jets are entirely discarded

Table 5.6: List of requirements for reconstructed small-radius jets in HBSM $t\bar{t}t\bar{t}$ analysis.

Jet Selection	
Feature	Requirements
Jet constituents	EMPFLOW
Jet clustering algorithm	anti- k_t
Pseudorapidity range	$ \eta < 2.5$
Transverse momentum	$p_T > 20\text{GeV}$
JVT	JVT > 0.5

Table 5.7: List of requirements for reconstructed small-radius jets in Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

5.5 b-Jets

Because the top quark decays almost exclusively into b-quarks, the presence of at least four b-quarks in an event is a useful indicator of $t\bar{t}t\bar{t}$ production. This analysis uses the Flavor Tagging recommendations and identifies jets originating from b-quarks via the DL1r tagger[98].

Generally, b-tagging algorithms exploit several features of b-hadrons: their long lifetime, their mass, and their tendency to decay into muons. B-hadrons have an average lifetime of roughly 1.5ps, which, depending on the energy, means that they will decay at a point several millimeters from the primary vertex where they originated[99]. B-jets will have a secondary vertex that is displaced from the primary vertex. The tracks created by b-hadron decay will therefore have large impact parameter values with respect to the primary vertex.

Low-level taggers directly use the track properties to assign probabilities of jet flavors based on the impact parameter and reconstructed secondary vertex. These taggers in turn act as inputs for the high-level DL1 Tagger, which employs an Artificial Deep Neural Network to output probabilities for each jet to originate from a light, c, or b-hadron[100]. In addition to using the low-level taggers, the DL1r variant of the DL1 Tagger incorporates track correlation information from a Recurrent Neural Network (RNN) [101, 102]. The track correlation results from multiple tracks being produced by a displaced vertex. If one track is found with a displaced vertex, another is more likely to be found. From the three probability outputs p_b , p_c , p_{light} , and from the adjustable parameter f_c , the DL1r assigns each jet a single value:

$$DL1r = \frac{p_b}{f_c p_c + (1 - f_c) p_{light}} \quad (5.12)$$

The f_c parameter is used to prioritize c -jet rejection versus *light*-jet rejection. For DL1r, the fraction f_c is set to 0.018 [101].

The DL1r tagger has four calibrated working points, which are defined by their b-tagging efficiency in $t\bar{t}$ MC samples: 60%, 70%, 77%, and 85%. The HBSM analysis uses all four WP's for various region definitions. In addition, jets are assigned a value based on how many of the four WP's they pass, where 1 corresponds to passing none, 2 corresponds to passing the 85% WP, continuing up to 5 which corresponds to passing the 60% WP. This process is referred to as pseudo-continuous b-tagging. The Higgs Oblique analysis defines a jet as b-tagged if it passes a DL1r cut at the 77% efficiency WP.

5.6 Reclustered Jets

For a given momentum, the decay products of heavy particles are generally expected to have higher angular separation than those of light particles. Thus, jets from lower-momentum top quark decays are expected to produce a larger radius jet. In order to account for jets with larger structure than $R=0.4$, this analysis makes use of reclustered jets[103].

Reclustered jets are formed using the small-r jets as input. Small-r jets must be calibrated when they are reconstructed, which means that changing the radius of a small-r jet is a nontrivial. Reclustering is useful because the calibration of the small-r jets used as inputs is applied automatically to the reclustered jet. Therefore, the radius of the reclustered jet can be easily adjusted to accommodate properties of the analysis signal, such as the larger radius of top decay jets.

Small-r jets must be above a certain energy threshold in order to be calibrated. Therefore, reclustered jets come with the additional advantage of some initial trimming, which removes contributions from pileup and the underlying event. Reclustered jets are further trimmed by removing all small-r constituents that contribute $< 5\%$ to the total p_T of the reclustered jet [104].

Besides the inputs, the reclustering process is otherwise identical to the jet reconstruction process. For this analysis, reclustered jets are constructed using the anti- k_t algorithm, with $R=1.0$. The input for reclustering is $R=0.4$ jets constructed

using the anti- k_t algorithm. Reclustered jets are required to have $p_T > 200\text{GeV}$ and $|\eta| < 2.0$ for this analysis [105].

5.7 Overlap Removal

Overlap removal refers to the procedure that ensures that a single detector signal does not contribute to multiple reconstructed objects. The objects are prioritized according to Table 5.8 and a track or calorimeter deposit contributing to multiple objects is instead assigned only to the object with the highest priority.

Overlap removal takes place after object reconstruction, and uses as input the reconstructed electrons, muons, and jets passing a loose selection criteria. The loose criteria are less strenuous than those used to define the objects used in the analysis, and are detailed in Tables 5.9 and 5.11.

Overlap removal starts with the objects at the top of the list in Table 5.8 and progresses downwards. The first objects to be compared are electron candidates that share a track. The electron candidate with the lower reconstructed p_T is removed from the list of electron candidates. Calorimeter muons are removed if they share a track with an electron. Conversely, electrons sharing an ID track with MS muons are removed. If a jet is within $\Delta R < 0.2$ of an electron, the jet is removed. If the jet is within $0.2 < \Delta R < 0.4$ of an electron, the electron is removed. The last stages of the overlap removal procedure compare muons and jets. If a jet with fewer than three tracks is ghost-matched (considering only track p_T direction in η and ϕ , not magnitude) to a muon ID track, the jet is removed. Lastly, if a muon with momentum $p_{T\mu}$ is a distance $\Delta R < 0.4 + 10\text{GeV}/p_{T\mu}$ from a jet, the muon is removed.

Overlap Removal Procedure		
Reject	Against	Criteria
Electron	Electron	shared track, $p_{T,1} < p_{T,2}$
Muon	Electron	is Calo-Muon and shared ID track
Electron	Muon	shared ID track
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < 0.4$
Jet	Muon	NumTrack < 3 AND (ghost-associated OR $\Delta R < 0.2$)
Muon	Jet	$\Delta < \min(0.4, 0.04 + 10\text{GeV}/p_T(\mu))$

Table 5.8: Overlap removal procedure.

Loose Electron Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 1.37$ or $1.52 < \eta < 2.47$
Transverse momentum	$p_T > 10\text{GeV}$
Track to vertex association	$ d_0^{BL} significance < 5$ and $ \Delta z_0^{BL} \sin\theta < 0.5\text{mm}$
Identification	MediumLH
Isolation	-
Charge ID	pass ECIDS (for $e^\pm e^\mp$ and $e^\pm \mu^\mp$ only)
Other object quality	not from bad calorimeter cluster

Table 5.9: List of requirements for reconstructed loose electrons in HBSM $t\bar{t}t\bar{t}$ analysis.

Loose Electron Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 1.37$ or $1.52 < \eta < 2.47$
Transverse momentum	$p_T > 15\text{GeV}$
Track to vertex association	$ d_0^{BL} significance < 5$ and $ \Delta z_0^{BL} \sin\theta < 0.5\text{mm}$
Identification	MediumLH
Isolation	Loose_varRad
Charge ID	pass ECIDS (for $e^\pm e^\mp$ and $e^\pm \mu^\mp$ only)

Table 5.10: List of requirements for reconstructed loose electrons in Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

Loose Muon Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 2.5$
Momentum calibration	Sagitta correction not used
Transverse momentum	$p_T > 10\text{GeV}$
Track to vertex association	$ d_0^{BL} significance < 3$ and $ \Delta z_0^{BL} \sin\theta < 0.5\text{mm}$
Reconstruction	Medium
Isolation	-

Table 5.11: List of requirements for reconstructed loose muons in HBSM $t\bar{t}t\bar{t}$ analysis.

Loose Muon Selection	
Feature	Requirements
Pseudorapidity range	$ \eta < 2.5$
Momentum calibration	Sagitta correction not used
Transverse momentum	$p_T > 15\text{GeV}$
Track to vertex association	$ d_0^{BL} \text{significance} < 3$ and $ \Delta z_0^{BL} \sin\theta < 0.5\text{mm}$
Reconstruction	Medium
Isolation	PflowLoose.FixRad

Table 5.12: List of requirements for reconstructed loose muons in Higgs Oblique $t\bar{t}t\bar{t}$ analysis.

5.8 Missing Transverse Energy

For collision events, there will be particles produced, particularly neutrinos, that are not detected by the ATLAS detector. These particles carry away energy that cannot be reconstructed. Therefore, the initial transverse momentum of zero TeV will result in a nonzero final transverse momentum. This is expressed as missing transverse energy, or E_{miss}^T , and can be used to quantify neutrino energy [106].

E_{miss}^T is calculated as the negative of the total E^T measured in the detector. This total is separated into x and y components of the transverse plane, and consists of contributions from reconstructed e , μ , photon, hadronic τ , and jet objects (hard terms), as well as charged-particle tracks associated with the primary vertex but not matched to any reconstructed objects (soft terms).

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,soft} \quad (5.13)$$

The hard terms come from fully reconstructed and calibrated objects. Because each type of object undergoes an independent reconstruction process, a detector signal may be included in both an electron and a jet. This results in double-counting of terms in the E_T^{miss} calculation. To prevent this, a signal ambiguity resolution is implemented. This procedure prioritizes objects with shared signals in the following order: electrons, photons, hadronic τ -jets, and jets. Muons are not included in ambiguity resolution because their reconstruction process, in the ID and MS, generally results in little overlap with E_T measurements of other objects.

The soft terms are calculated from the p_T of charged-particle tracks in the ID. These tracks are associated with the hard-scatter vertex, but not with any of the reconstructed hard objects. Tracks from jets rejected by the ambiguity resolution are also included. Because the tracks are required to be associated with the hard-scatter vertex, track-only $E_T^{miss,soft}$ is mostly unaffected by pileup.

CHAPTER 6

Systematic Uncertainties

6.1 Introduction

For any measurement, there is an associated uncertainty. The results from this analysis are comprised of many measurements: signals from the ATLAS detector reconstructed into physics objects, interpretations of physics objects (such as b-tagging), and values calculated from theoretical predictions. The purpose of this chapter is to list the sources of systematic uncertainty that have been incorporated into the measurement of the $t\bar{t}t\bar{t}$ cross section.

The uncertainties are generally divided into two types: those coming from the MC modeling of samples, and data uncertainties from the ATLAS detector.

6.1.1 HBSM Analysis: Theoretical Uncertainties

Theoretical uncertainties contribute to the inaccuracies in the predicted yields for each type of background interaction. Since the dominant background is $t\bar{t} + \text{jets}$, the uncertainties for this sample are given particular attention. Below are listed the systematic uncertainties included for each sample. Generally, systematic uncertainties are applied as a normalization factor to the cross section only. However, generator and parton shower choice uncertainties are applied to both cross section and distribution shape.

$t\bar{t} + \text{jets}$: The $t\bar{t} + \text{jets}$ sample is first divided into $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$, and $t\bar{t} + \text{light}$ samples, and then corrected using the two-step reweighting process described in Chapter 7. Every uncertainty correspondingly undergoes the same categorization and reweighting. The $t\bar{t} + \geq 1b$ sample is then further divided into four categories:

1. $t\bar{t}b$: $t\bar{t} + \text{jets}$ events with one extra jet matched to a b-hadron
2. $t\bar{t}B$: $t\bar{t} + \text{jets}$ events with one extra jet matched to at least two b-hadrons
3. $t\bar{t}bb$: $t\bar{t} + \text{jets}$ events with two extra jets matched to two b-hadrons

4. $t\bar{t}bbb$: all $t\bar{t}+ \geq 1b$ events that do not enter in the previous categories

The nominal hard-scatter sample is produced using an NLO calculation, neglecting contributions from higher-order corrections. To estimate the uncertainty created by the missing higher-order corrections, the factorization scale is varied by a factor of 2 and $1/2$, and the renormalization scale is independently varied by a factor of 2 and $1/2$. This is done for each of the three categories of $t\bar{t} + \text{jets}$, producing six nuisance parameters (NPs).

There is also uncertainty caused by the initial state radiation (ISR) and final state radiation (FSR). This radiation is the emission of a gluon or photon immediately before or after the pp collision, which affects the incoming and outgoing momentum of the hard process partons. The ISR and FSR are accounted for in parton shower modelling. The ISR uncertainty is estimated by varying the input of α_S according to the recommendations for the A14 tune, giving $\alpha_S = 0.127^{+0.013}_{-0.012}$ [76]. The uncertainty in the quantity of FSR is estimated by varying the factorization in the A14 tune by a factor of $1/2$ and 2.

Different matrix element calculators and parton shower (PS) models will give somewhat different results. Therefore, there is uncertainty associated with the choice of generator and parton shower simulator. To estimate these uncertainties, the nominal choices are compared to alternatives. For the ME generator, PowHegBox v2 is compared to MadGraph5_aMC@NLO. Differences are expressed as 12 NP's: shape and acceptance variations for $t\bar{t} + \text{light}$, $t\bar{t}+ \geq 1c$, $t\bar{t}b$, $t\bar{t}B$, $t\bar{t}bb$, and $t\bar{t}bbb$. For PS modeling, Pythia 8.230 is compared to Herwig 7, both using PowHegBox v2 as the ME generator. This also produces a similar set of 12 NP's.

Another theoretical uncertainty for the $t\bar{t}$ samples arises from the normalization factors obtained from HF re-scaling. In this case, the normalization factor is given a conservative flat uncertainty of 50%, applied to the factor for $t\bar{t}+ \geq 1c$ and all four sub-categories of $t\bar{t}+ \geq 1b$.

Lastly, there is an uncertainty associated with the flavor scheme used to produce the $t\bar{t} + \text{jets}$ samples. The 'flavor scheme' refers to whether or not the b-quark is included in the list of possible partons for the matrix element modeling of a pp collision [107]. The four-flavor scheme (4FS) does not include b-quark partons, only the four lightest quarks. The five-flavor scheme (5FS) does include b-quarks as partons.

The HBSM analysis uses 5FS for MC sample production. However, it is

unclear which of the two flavor schemes provides a better approximation of the hard process, and indeed the answer may change depending on the type of event being modeled. Moreover, events generated with the 4FS have different kinematic distributions than events generated with 5FS. The greatest discrepancy is in events with a large number of b-quarks, in particular $t\bar{t} + \geq 1b$ events. This can be seen in Figure 6.1, where the MC samples are all background events generated with 5FS, and the “data points” are MC samples with $t\bar{t} + \geq 1b$ produced with 4FS. The discrepancy increases for higher jet and b-jet multiplicity, as well as at higher GNN scores.

To account for the discrepancy between 4FS and 5FS, the HBSM analysis incorporates an uncertainty associated with the modeling of $t\bar{t} + \geq 1b$ events. The extra uncertainty helps to account for the difference in the two flavor scheme distributions. Figure 6.2 expresses the discrepancy as an extracted μ value, where $\mu = 1$ would indicate that the two flavor schemes matched exactly. Though the μ value is not equal to one, when the flavor scheme uncertainty is added, $\mu = 1$ is within uncertainties of the fitted value.

$t\bar{t}t\bar{t}$: There are uncertainties associated with the MC modeling of the signal samples. QCD corrections are applied by varying the renormalization and factorization scales by a factor of 2 and 1/2 in the nominal ME generator, MadGraph5_aMC@NLO. The PDF uncertainties were found to be negligible in previous studies, and so are accounted for by a 1% uncertainty[6]. A cross section uncertainty of +7.8% and -13.3% is applied to account for the uncertainty in the SM prediction. These values come from a recent calculation, which predicts a cross section of $13.37^{+1.04}_{-1.78}\text{fb}$ [108]. Note that this is a different prediction than is used for the Higgs Oblique analysis, which was conducted earlier and used an older cross section prediction.

$t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$: The QCD corrections are applied by varying the renormalization and factorization scales by a factor of 2 and 1/2 in the nominal ME generator: MadGraph5_aMC@NLO for $t\bar{t}W$ and $t\bar{t}Z$, and PowHegBox v2 for $t\bar{t}H$.

Generator choice uncertainties are estimated by comparing the nominal generator with an alternative. For $t\bar{t}W$ and $t\bar{t}Z$, the nominal MadGraph5_aMC@NLO is compared to Sherpa 2.2.10 and Sherpa 2.2.1, respectively. For $t\bar{t}H$, the nominal PowHegBox v2 is compared to MadGraph5_aMC@NLO.

There is uncertainty associated with modeling the samples in the high jet multiplicity regions. This uncertainty is quantified based on the levels of mismod-

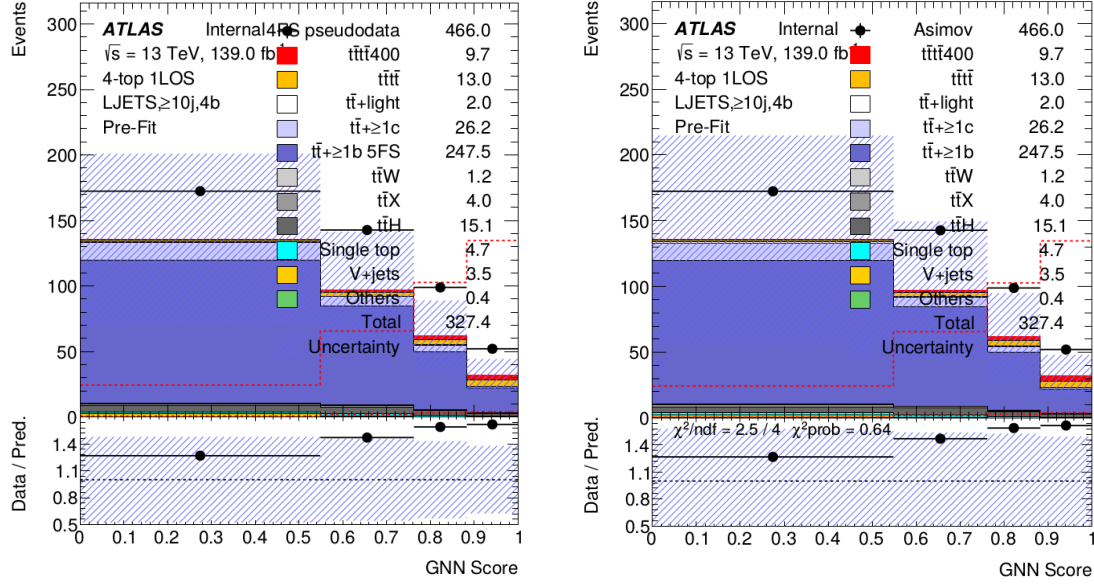


Figure 6.1: Distribution of GNN score in the $\geq 10j4b$ region. The distribution is shown for MC samples with a 5FS $t\bar{t}+\geq 1b$ sample, and for MC samples with a 4FS $t\bar{t}+\geq 1b$ sample (shown as “data” points). The left plot shows the distributions without a flavor scheme uncertainty, and the right plot shows the same distribution with the flavor scheme uncertainty added. The discrepancy in the highest GNN score bins is still not completely covered by the extra uncertainty, but shows improvement.

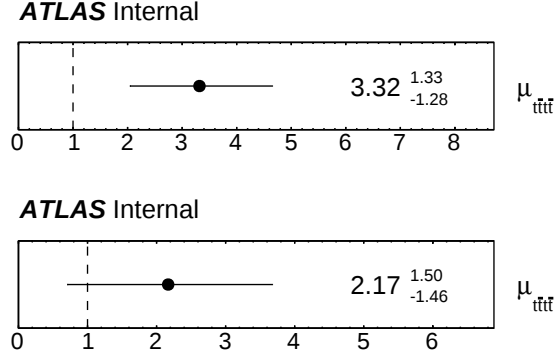


Figure 6.2: Comparison between 5FS and 4FS expressed as a fit of 5FS MC samples to 4FS “data.” The large value of extracted $\mu_{t\bar{t}t}$ indicates that the discrepancy between 4FS and 5FS is being interpreted by the fit as a large $t\bar{t}t$ -like signal. Exact agreement between the two flavor schemes would result in $\mu_{t\bar{t}t} = 1$. It can be seen in the top plot (without FS uncertainties) that the value of $\mu_{t\bar{t}t}$ is not 1 even within uncertainties. The bottom plot incorporates FS uncertainties, so that the total uncertainty now encompasses the value of $\mu_{t\bar{t}t} = 1$.

elling of $t\bar{t} + jets$ for N_{jets} distributions in the $2b$ region. A 10%, 20%, and 30% uncertainty is applied respectively to events with 9(7), 10(8), and 11(9) jets in the 1L(OS) channel.

There is a 10% uncertainty on the theoretical cross section value for $t\bar{t}H$, and 12% for $t\bar{t}Z$ [109]. These uncertainties are applied in addition to those listed above.

$t\bar{t}t$: A theoretical cross section uncertainty of 35% is applied to the $t\bar{t}t$ cross section. Note that this background sample is not yet included in the results shown in Chapters 9 and 10.

Other processes: This includes VV , VH , and VVV , as well as $t\bar{t}WW$, $t\bar{t}ZZ$, $t\bar{t}WZ$, $t\bar{t}HH$, and $t\bar{t}WH$. For VV , a 40% uncertainty is applied to the cross section[110]. For all others, a conservative 50% uncertainty is applied to the cross section.

6.1.2 Higgs Oblique Analysis: Theoretical Uncertainties

$t\bar{t}t\bar{t}$: Modelling uncertainties are derived through comparison to the alternative samples. QCD corrections are applied by varying the renormalization and factorization

scales by a factor of 2 and 1/2 in the nominal ME generator, MadGraph5_aMC@NLO. The PDF uncertainties were found to be negligible in previous studies, and so are accounted for by a 1% uncertainty[6]. A 20% uncertainty is applied to the cross section to account for the uncertainty in the SM prediction of $11.97^{+18\%}_{-21\%}\text{fb}$ [4].

$t\bar{t}W$: Modelling uncertainties are derived through comparison to the alternative samples. For $t\bar{t}W$, the alternative samples are reweighted using the same data-driven formula discussed in Chapter 7. QCD corrections are applied by varying the renormalization and factorization scales by a factor of 2 and 1/2 in the nominal ME generator. No PDF or cross section uncertainties are applied to $t\bar{t}W$ samples because the N_{jets} normalization factors are a part of the fit. There is an additional 50% uncertainty applied to events with exactly one additional truth b-jet not from top decay. There is also a 50% uncertainty added to events with at least two additional truth b-jets not from top decay.

$t\bar{t}H$: Modelling uncertainties are derived through comparison to the alternative samples. QCD corrections are applied by varying the renormalization and factorization scales by a factor of 2 and 1/2 in the nominal ME generator. PDF uncertainty is 1%. The uncertainty on the theoretical cross section value for $t\bar{t}H$ is 10%[109]. There is an additional 50% uncertainty applied to events with exactly one additional truth b-jet not from top decay. There is also a 50% uncertainty added to events with at least two additional truth b-jets not from top decay.

$t\bar{t}Z$: Modelling uncertainties are derived through comparison to the alternative samples. QCD corrections are applied by varying the renormalization and factorization scales by a factor of 2 and 1/2 in the nominal ME generator. PDF uncertainty is given as 1%. The uncertainty on the theoretical cross section value for $t\bar{t}Z$ is 12%[109]. There is an additional 50% uncertainty applied to events with exactly one additional truth b-jet not from top decay. There is also a 50% uncertainty added to events with at least two additional truth b-jets not from top decay.

tZ and tWH : Both the tZ and tWH samples have a 30% cross section uncertainty applied to them [111, 112].

$t\bar{t}t$: There is a 30% cumulative uncertainty that accounts for the theoretical cross section uncertainty, QCD uncertainty, PDF uncertainty, and uncertainty associated with choice of flavor scheme. There is an additional 50% uncertainty applied to events with at least four truth b-jets.

VV : For the VV samples, the cross section uncertainty varies by jet mul-

tiplicity and is set as 20%, 50%, and 60% for events with ≤ 3 , 4, and ≥ 5 jets respectively [110].

$t\bar{t}WW$: A conservative cross section uncertainty of 50% is applied to this rare process.

Other: For VH , VVV , $t\bar{t}ZZ$, $t\bar{t}WZ$, $t\bar{t}HH$, $t\bar{t}WH$, there is a conservative total cross section uncertainty of 50%.

QmisID There are three uncertainty values associated with the estimation of QmisID rates, a process described in Chapter 7. The first is statistical uncertainty. There is also uncertainty on the choice of the Z -mass window. The value of this uncertainty is obtained by taking the largest variation in the QmisID rate after changing the window from 3σ to 5σ in the Breit-Wigner fit. The third uncertainty comes from non-closure, which is obtained by comparing the QmisID results to Sherpa 2.2.1 $Z \rightarrow e^+e^-$ truth samples.

6.2 Experimental Uncertainties

Experimental uncertainties cover the systematic errors that may occur at any stage in data-taking, from measurement of luminosity to the reconstruction of physics objects. These uncertainties are the same for both analyses.

Electrons: Events selected using the single-electron trigger have an associated uncertainty due to the trigger efficiency[113]. Electrons are reconstructed by matching tracks in the ID with deposits in the EM calorimeter[79]. Further filters are applied to ensure purity of the reconstructed electrons, including an identification step based on a likelihood discriminant and an evaluation of electron isolation. Electrons are modelled in MC using either a full simulation or AltFastII (AFII)[78].

There are differences in efficiency between data and MC for trigger, reconstruction, identification, and isolation. Scale factors are derived using the tag-and-probe method applied to MC and data[79]. There are four scale factors, one each for trigger, reconstruction, identification, and isolation, and there are four associated NP's for these scale factors.

In addition to the four scale factor uncertainties, there is one NP for the AFII energy scale, one NP for the full simulation energy scale, one NP for electron energy resolution, and one NP associated with the efficiency of the ECIDS tool used to remove events with incorrect electron charge[86]. These are included to account

for differences between data and MC.

Electron uncertainties are summarized in Table 6.1.

Muons: For events containing muons, there are 15 total NP's. There are NP's associated with scale factors applied to trigger, identification, and isolation efficiencies. Like electron uncertainties, these are obtained by comparison of data and MC[87]. Unlike electrons, muons are reconstructed using information from the ID and the MS, where tracks reconstructed from both detectors are matched to each other to yield a candidate muon. This track-to-vertex (TTVA) efficiency has one NP for statistical uncertainty and one NP for systematic uncertainty.

In addition, muons also have one NP for track smearing in the ID and one NP for track smearing in the MS. There is one NP associated with the charge independent momentum scale, and two NP's for the charge dependent momentum scale.

Muon uncertainties are summarized in Table 6.1.

Jets: This analysis uses particle-flow jets reconstructed using the anti- k_t algorithm. The small-R jets are reconstructed using radius parameter $R=0.4$. Pileup reduction is performed using the Jet Vertex Tagger[95]. There are uncertainties associated with the JVT, the jet energy scale (JES), the jet energy resolution (JER), reclustering jets, and heavy-flavor tagging.

The JVT operates by applying selections to jet p_T and $|\eta|$. There are scale factors applied to MC to correct for discrepancies in JVT efficiency between MC and data. The JVT uncertainty is the statistical uncertainty combined with a 30% uncertainty on pileup estimation after pileup suppression has been applied[97].

JES is a correction applied to reconstructed jets to more closely match the energy of truth jets[63]. The JetETMiss group at ATLAS provides a recommended set of reduced uncertainties to be used for the JES and JER[114]. Reduced uncertainties are obtained by combining the least significant NP's in a way that preserves their correlation [115]. For JES, there are 15 effective NP's: two detector-related, four modeling-related, three mixed modeling and detector-related, and six statistical uncertainties[116]. There is one NP to account for the difference in response between gluon and quark jets (jet flavor composition)[63]. There is one NP to account for the difference in gluon jet response between sample generators (jet flavor response). There are four NP's associated with pileup subtraction from jets, labeled OffsetMu, OffsetNPV, PtTerm, and RhoTopology. For jet punch-through, one of two NP's is

applied: MC16 if the sample is produced with full simulation, and AFII if produced with AltFastII. There is one NP for non-closure between AFII and full simulation, one NP for single particle response at high p_T , and one NP for b-jet response. For JER, there are seven effective NP's, as well as two NP's for data/MC agreement.

The reclustered (RC) jets are constructed by inputting small-R jets into the anti- k_t algorithm with radius $R=1.0$ [117]. Because RC jets are composed of small-R jets, the uncertainties associated with small-R jets are simply propagated to the RC jets.

Heavy-flavor jets include those originating from a b-quark or c-quark decay, and heavy-flavor tagging involves the use of the DL1r tagger[102]. The tagging process has uncertainties that arise from limited statistics, from calibration of the detector and from theoretical modeling of the jets. These uncertainties are consolidated into a reduced set, including 44 NP's for b-tagging, and 19 NP's each for c-tagging and light-tagging[116].

MET: The missing transverse energy is reconstructed as an object[106]. There is a hard component comprised of every successfully reconstructed and identified object in an event. Detector activity that cannot be matched to an object is used to calculate the soft component of E_T^{miss} . Because the hard component is composed of other reconstructed objects, the uncertainty of the hard component is simply propagated to E_T^{miss} . It is only for the soft component that separate NP's are required. There are three NP's: one based on the projection of p_T^{soft} parallel to p_T^{hard} ; one based on the resolution of the parallel component $p_{\parallel}^{soft}$, calculated as the RMS of $p_{\parallel}^{soft}$; and one for the resolution of the perpendicular component $p_{\perp}^{soft}$, calculated as the RMS of $p_{\perp}^{soft}$.

Luminosity: Luminosity is determined largely by the LUCID2 detectors, as detailed in Chapter 4. Uncertainties arise from this procedure, such as background subtraction and scan-to-scan reproducibility, and also from the fact that the calibration is performed at low luminosities and must be extrapolated to the high luminosity at which physics data-taking is performed. This extrapolation process, called calibration transfer, is the largest source of uncertainty for the luminosity. The calibration transfer uncertainty is estimated by comparing low-luminosity fills with a pair of consecutive high-energy fills, with both measurements taken from the TileCal's forwardmost segments, the E-cells. The total luminosity uncertainty is calculated as 0.83% and is applied as a flat uncertainty to all events [51].

Pileup Reweighting: There is a discrepancy between pileup distributions in data and MC, and the MC samples are reweighted to better match data. In the process, the data estimate of $\langle\mu\rangle$, or number of inelastic collisions per event, is rescaled. The rescaling is a ratio of 1.00/1.03, reducing the pileup estimate. This rescaling has an associated uncertainty, which is calculated by increasing the data estimate (“up” variation) and decreasing (“down” variation). The up variation is calculated by applying a factor of 1.00/0.99, and the down variation is calculated by applying a factor of 1.00/1.06[116].

A full list of systematics is shown in Table 6.1.

Systematics	
Systematic Uncertainty	Description
ATLAS_Lumi	Uncertainty of Run II int. luminosity
ATLAS_pileup_reweighting	Correction to MC pileup dist.
Jets	
ATLAS_JVT	Jet vertex tagging uncertainty
ATLAS_JER_EffectiveNP	Effective NP’s for JER
ATLAS_JES_BJES	JES uncertainty b-jet response
ATLAS_JES_SinglePart	JES uncertainty single particle resp.
ATLAS_JES_PU_Rho	JES uncertainty pileup subtraction
ATLAS_JES_PU_PtTerm	
ATLAS_JES_PU_OffsetNPV	
ATLAS_JES_PU_OffsetMu	
ATLAS_JES_Flavor_Resp	JES uncertainty flavor (response)
ATLAS_JES_Flavor_Comp	JES uncertainty flavor (composition)
ATLAS_JES_EtaInter_NonClosure_highE	JES eta-intercalibration
ATLAS_JES_EtaInter_NonClosure_negEta	
ATLAS_JES_EtaInter_NonClosure_posEta	
ATLAS_JES_EtaInter_Stat	
ATLAS_JES_EtaInter_Model	
ATLAS_JES_EffectiveNP	Effective NP’s for jet energy scale
ATLAS_JES_PunchThrough	Uncertainty for punch-through jets
ATLAS_FTAG	Uncertainties for DL1r tagging
Electrons	
ATLAS_EG_RESOLUTION	Electron energy resolution
ATLAS_EG_SCALE	Electron energy scale (Full Sim.)
ATLAS_EG_SCALE_AF2	Electron energy scale (AFII)
ATLAS_EL_ID	Electron identification uncertainty
ATLAS_EL_Isol	Electron isolation uncertainty

ATLAS_EL_Reco ATLAS_EL_Trigger	Electron reconstruction uncertainty Single electron trigger uncertainty
Muons	
ATLAS_MUON_ID	Muon ID track smearing uncertainty
ATLAS_MUON_MS	Muon MS track smearing uncertainty
ATLAS_MUON_SAGITTA_RESBIAS	Charge dependent momentum scale
ATLAS_MUON_SAGITTA_RHO	
ATLAS_MUON_SCALE	Charge independent momentum scale
ATLAS_MU_ID_SYST	Muon identification uncertainty
ATLAS_MU_ID_SYST_LOWPT	
ATLAS_MU_ID_STAT	
ATLAS_MU_ID_STAT_LOWPT	
ATLAS_MU_Isol_SYST	Muon isolation uncertainty
ATLAS_MU_Isol_STAT	
ATLAS_MU_Trigger_SYST	Single muon trigger uncertainty
ATLAS_MU_Trigger_STAT	
ATLAS_MU_TTVA_SYST	Track-to-track vertex uncertainty
ATLAS_MU_TTVA_STAT	
E_T^{miss}	
ATLAS_MET_SoftScale	E_T^{miss} soft component uncertainty
ATLAS_MET_SoftResPerp	
ATLAS_MET_SoftResPara	

Table 6.1: List of uncertainties for reconstructed objects

CHAPTER 7

Data-Driven Reweighting

7.1 HBSM Analysis

7.1.1 Introduction

In the single lepton and opposite sign dilepton final states, the background for $t\bar{t}t\bar{t}$ production is dominated by the production of $t\bar{t}$ in association with additional jets. In the phase space of interest, at high jet and b-jet multiplicities, the MC modeling of $t\bar{t}$ events becomes inaccurate. The contribution of $t\bar{t} + \text{heavy flavor (HF) jets}$ is known to be underestimated [118]. The distributions of kinematic variables are also affected by the mismodeling.

To address the inaccuracies, the MC events are reweighted via a two-step process. In the first step, the flavor contributions of the additional jets in $t\bar{t} + \text{jets}$ is adjusted to match data in regions with negligible signal contribution. In the second step, the newly scaled events are used as inputs for an NN-based reweighting of the kinematic distributions.

7.1.2 HF Rescaling

The $t\bar{t} + \text{jets}$ background is comprised of different types of events that can be categorized by the flavor of quark(s) initiating the additional jet(s), where HF jets are those originating from a c-quark or b-quark. The $t\bar{t} + \text{jets}$ events are separated into three groups: $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$, and $t\bar{t} + \text{light}$. The separation is performed using the truth-level variables `HF_Classification` and `jet_truthflav` to identify truth b-hadrons and c-hadrons that did not come from the decay of a top quark. Particle jets not originating from the $t\bar{t}$ interaction are matched to truth hadrons, requiring a distance of $\Delta R < 0.4$. An event is classified as $t\bar{t} + \geq 1b$ if at least one particle jet not in the $t\bar{t}$ system is matched to a truth b-hadron, and that b-hadron has $p_T > 5\text{GeV}$. An event is classified as $t\bar{t} + \geq 1c$ if it is not identified as a $t\bar{t} + \geq 1b$ and at least one particle jet is matched to a c-hadron with $p_T > 5\text{GeV}$. This means that events with b-jets that decay to c-jets are still classified as $t\bar{t} + \geq 1b$. All remaining $t\bar{t}$ events are classified as $t\bar{t} + \text{light}$.

Normalization factors for the three categories are defined: μ_{TTL} , μ_{TTC} , and μ_{TTB} . The values of these parameters are determined by a likelihood fit to data in regions where $N_{jets} \geq 6j$ (2LOS) and $N_{jets} \geq 8j$ (1L), separated by b-jet multiplicity into 2b, 3bL, 3bH, and $\geq 4b$ regions (3bL and 3bH are defined in Table 8.1). These regions have $\leq 2.6\%$ $t\bar{t}t\bar{t}$ signal contamination, assuming a 12fb SM $t\bar{t}t\bar{t}$ cross section and a 12fb BSM $t\bar{t}t\bar{t}$ cross section. This is considered suitable for background-only modeling [116]. The normalization factors resulting from the fit are shown in Figure 7.1. These factors are then applied to the $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$, and $t\bar{t} + light$ samples for all regions. The data/MC comparison before and after HF rescaling is shown in Figure 7.2.

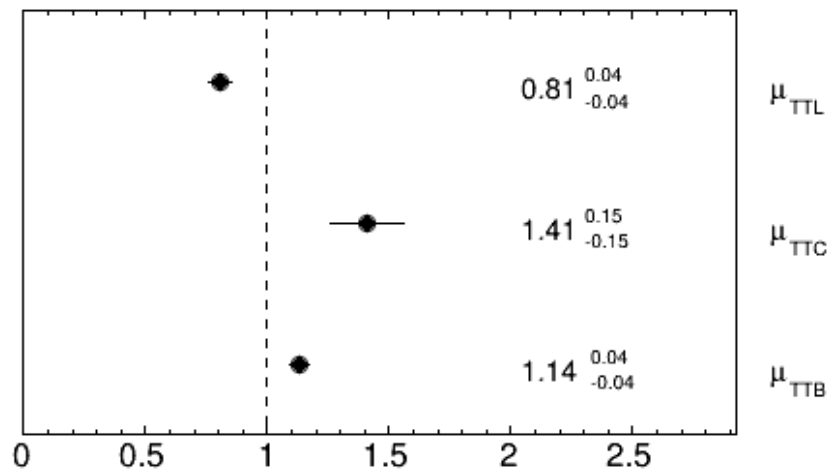


Figure 7.1: The extracted values of the three normalization factors for the flavor of the $t\bar{t}$ associated jet(s).

7.1.3 NN Based Kinematic Reweighting

The HF rescaling ensures that the appropriate flavor fractions of $t\bar{t} + jets$ are used, but does not correct for shape discrepancies between data and MC in kinematic distributions. This discrepancy is most apparent at high jet p_T , high jet activity (sum of jet p_T), and high jet multiplicity [119]. After HF rescaling, the MC events undergo a further reweighting derived from a data-driven NN method.

The method involves an NN classifier discriminating between two classes of events, denoted data and MC, which each consist of N variables forming a vector $\mathbf{x}$

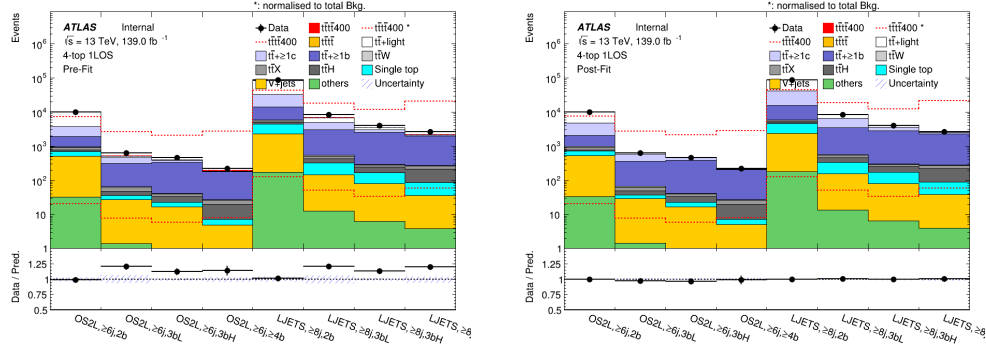


Figure 7.2: MC yields for each rescaling region, compared to data. The left plot shows the comparison before rescaling, and the right shows the same comparison after rescaling is applied.

[120]. The fraction of events used in training that are data is α_{data} and the fraction of events that are MC is α_{MC} , such that $\alpha_{data} + \alpha_{MC} = 1$. Whether an event is coming from data or MC is of course known in this analysis, but the intent of this particular classifier is to determine the ratio of the probability distributions, $P_{data}(x)$ and $P_{MC}(x)$. The classifier will have some output, $o(\mathbf{x})$, based on $P_{data}(x)$ and $P_{MC}(x)$. The classifier is trained on both data and MC. After training and minimizing, the output $o_{min}(\mathbf{x})$ is expressed as the probability of an event being classified as ‘data’ given input variables $\mathbf{x}$:

$$o_{min}(\mathbf{x}) = P(data|\mathbf{x}) = \frac{\alpha_{data}P_{data}(\mathbf{x})}{\alpha_{data}P_{data}(\mathbf{x}) + \alpha_{MC}P_{MC}(\mathbf{x})} \quad (7.1)$$

This expression can be rearranged to become a function of the ratio of probability distributions:

$$\frac{o_{min}(\mathbf{x})}{1 - o_{min}(\mathbf{x})} = \frac{\alpha_{data}P_{data}(\mathbf{x})}{\alpha_{MC}P_{MC}(\mathbf{x})} = w(\mathbf{x}) \quad (7.2)$$

The $w(\mathbf{x})$ term represents the reweighting factor that is then used to scale only MC events. This method does not require binning or prior knowledge of individual probability distributions $P_{sig}(\mathbf{x})$ and $P_{bkg}(\mathbf{x})$. Instead, the focus is on extracting the ratio of the two probability distributions. After the NN is trained, each MC event is individually input into the trained NN, the output $o_{min}(\mathbf{x})$ is calculated, and the

MC event is scaled by $w(\mathbf{x})$.

For this analysis, the kinematic reweighting is performed by training the NN in the 7j (5j) and 2b region for 1L (2LOS). To mimic a $t\bar{t}$ + jets only dataset, the data include non- $t\bar{t}$ MC events with negative weighting, effectively subtracting the non- $t\bar{t}$ contribution from data, as seen in Figure 7.3. Modeling is different between the nominal and alternative $t\bar{t}$ + jets samples, so the kinematic corrections must be performed independently for each nominal and alternative sample. A hyperparameter scan was performed to determine the optimal choices for settings such as loss function [116]. For both 1L and 2LOS samples, an exponential loss function is used. This loss function compares the NN output y for an event i to its expected value, $\hat{y}_i$:

$$\mathcal{L} = \sum_{i=1}^{i=N} e^{-y_i \hat{y}_i} \quad (7.3)$$

where $N = 1000$ is the number of events in the training sample. The exponential function is particularly useful for accounting for the lower statistics of the 2LOS samples. The variables $\mathbf{x}$ used for the reweighting are the jet p_T , number of jets, number of reclustered jets (described in Chapter 5), missing E_T , and lepton jet p_T .

The correction to the H_T distribution is shown for both the 1L and 2LOS channels in Figure 7.4. The ratio of data $t\bar{t}$ yield to nominal Powheg Pythia MC $t\bar{t}$ samples without kinematic reweighting is shown in the black points. Without reweighting, there is as much as a 50% underestimation in MC yield at low H_T , and an overestimation at high H_T . The ratio of data to reweighted MC is shown in the red points, demonstrating better agreement. As a second example, the distribution for N_{jets} is shown in Figure 7.5. Again, the distribution shows better data/MC agreement after reweighting is applied.

7.2 Higgs Oblique Analysis

7.2.1 QmisID

The QmisID rate is corrected by fitting in a region dominated by the $Z \rightarrow e^+e^-$ process. It is assumed that events within the Z -mass window will consist of $Z \rightarrow e^+e^-$ and a uniform background, which can be estimated using events in the side bands around the window. In the Z -mass window region, all same-sign (SS) electron pairs

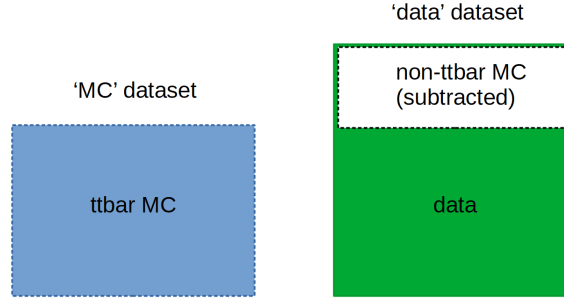


Figure 7.3: Training datasets for the reweighting NN. The two classes the NN differentiates are MC and data.

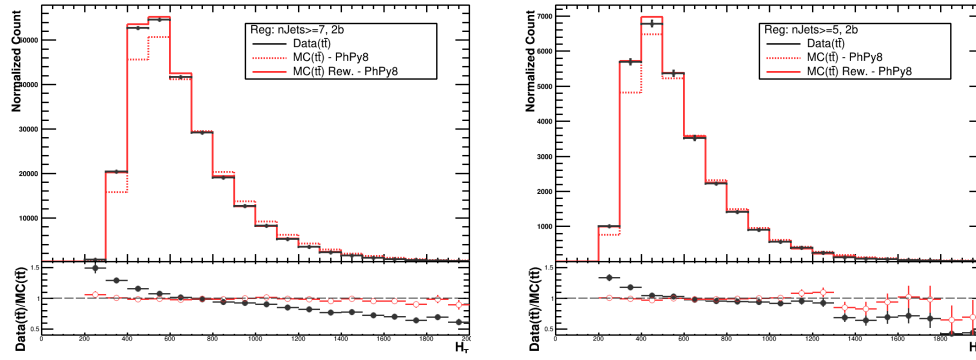


Figure 7.4: The H_T distribution for 1L sample (left) and 2LOS sample (right). The data distribution is compared to the MC $t\bar{t} + jets$ distribution before and after kinematic reweighting.

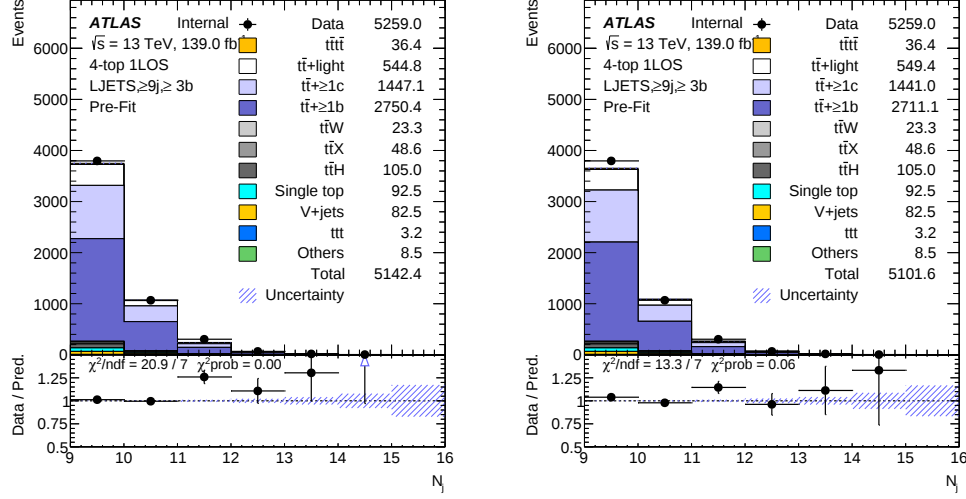


Figure 7.5: The N_{jets} distribution for 1L sample before reweighting (left) and after (right).

are assumed to include one electron with misidentified charge. The misidentification can originate from high electron p_T resulting in a straight track in the ID. Misidentification can also come from one of the electrons radiating a photon, which subsequently produces another e^+e^- pair. This process of photon conversion is also discussed in Section 8.2.2.

For each SS electron pair, the number of events with the leading electron in the i^{th} 2D bin of p_T/η and the sub-leading electron in the j^{th} bin, N_{ij}^{SS} , is expressed as:

$$N_{ij}^{SS} = N_{ij}^{OS+SS}(\epsilon_i(1 - \epsilon_j) + \epsilon_j(1 - \epsilon_i)) \quad (7.4)$$

Here, the QmisID rates for the i^{th} and j^{th} bin are ϵ_i and ϵ_j . The number of SS events is assumed to follow a Poisson distribution centered around the expected value, $\bar{N}_{ij}^{SS}$.

$$f(N_{ij}^{SS} | \bar{N}_{ij}^{SS}(\epsilon_i, \epsilon_j)) = \frac{(\bar{N}_{ij}^{SS})^{N_{ij}^{SS}} e^{-\bar{N}_{ij}^{SS}}}{N_{ij}^{SS}!} \quad (7.5)$$

The Poisson likelihood function is then constructed, following the same logic discussed in Chapter 9.

$$L(\epsilon|N^{SS}) = \prod_{i,j} f(N_{ij}^{SS}|\bar{N}^{SS}(\epsilon_i, \epsilon_j)) \quad (7.6)$$

This is fit to data in the $Z \rightarrow e^+e^-$ region, and the likelihood is minimized to extract the values of ϵ in 2D bins of jet p_T and $|\eta|$. There are four bins in $|\eta|$ and three bins in p_T for the CRttW⁺, CRttW⁻, and material conversion regions. Due to low statistics, only one bin is used for the low m_{γ^*} region.

Events are selected around the Z mass. The boundaries for the Z mass window and sidebands are shown in Table 7.1.

	Z Window	Lower Side Band	Upper Side Band
SS	[67.07, 112.03]	[44.59, 67.07]	[112.03, 134.51]
OS	[71.96, 109.32]	[53.28, 71.96]	[109.32, 128.00]

Table 7.1: Range of Z mass window and sidebands, in GeV.

To find the mean and variance of the Z mass, the yield distribution is fit to a Breit-Wigner function. The mean mass value for SS was measured as 89.55GeV and for OS was measured as 90.64GeV. Since many SS events are due to photon conversion, there is a shift in the mean mass for SS events from the energy loss. The Z -mass window is defined as the region within 4σ of the SS mean, and the side bands are in the area between twice the Z -mass window and the boundaries of the window. The yield in the side bands is averaged and subtracted from the Z peak to get the total number of $Z \rightarrow e^+e^-$ events.

The $Z \rightarrow e^+e^-$ MC samples are then fit to the data in 2D bins of p_T and $|\eta|$. Fitting MC samples to data is described in Chapter 9. Each bin must have a converging fit and an uncertainty of less than 30%, otherwise the binning is re-optimized. A binning setup was found that satisfied both of these requirements.

Once the values for ϵ_i and ϵ_j are calculated, they are applied as weights to the OS pair events in the region of $N_{jets} \geq 6$, $N_{b-jets} \geq 2$, and $H_T > 500\text{GeV}$. The weight is expressed as:

$$w = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} \quad (7.7)$$

7.2.2 $t\bar{t}W$ Data-Driven Reweighting

In the previous search for the SM production of $t\bar{t}t\bar{t}$ in the SSML channels, the modelling uncertainties on the $t\bar{t}W$ + jets background were among the most impactful in affecting the analysis sensitivity [6]. To reduce this uncertainty, a data-driven reweighting is applied to the N_{jets} distribution of $t\bar{t}W$ samples.

The N_{jets} distribution is assumed to follow scaling for QCD jets [121]. This scaling is a parameterized function, which is expressed as the ratio of the number of events with $n+1$ jets to the number of events with n jets. There are two components to the parameterized function: an exponential decay and a Poisson distribution.

The exponential function assumes that the yield decreases exponentially as a function of N_{jets} :

$$\frac{\text{Yield}_{n+1}}{\text{Yield}_n} = e^{-b} = a_0 \quad (7.8)$$

or, rearranging:

$$\text{Yield}_{n+1} = \text{Yield}_n \times a_0 \quad (7.9)$$

This can be extrapolated to the case where there are $n + m = J$ total jets:

$$\text{Yield}_J = \text{Yield}_n \times a_0^m \quad (7.10)$$

There are four hard process jets in a $t\bar{t}W$ dilepton event. Considering this process, where $n = 4$, the $t\bar{t}W$ yield for total jets $J = m + 4$ is:

$$\text{Yield}_J = \text{Yield}_{4j} \times a_0^{(J-4)} \quad (7.11)$$

The Poisson distribution follows the form, for n total jets with expectation value $\bar{n}$:

$$P_n = \frac{\bar{n}^n e^{-\bar{n}}}{n!} \quad (7.12)$$

and

$$P_{n+1} = \frac{\bar{n}^{n+1} e^{-\bar{n}}}{(n+1)!} \quad (7.13)$$

The ratio, then, is

$$P_{(n+1)/n} = \frac{\bar{n}}{(n+1)} = \frac{a_1}{1 + (J-4)} \quad (7.14)$$

Scaling both functions by the yield at $N_{jets} = 4$, the combined function is a product of all successive bins j from $N_{jets} = 5$ to $N_{jets} = J$:

$$\begin{aligned} \text{Yield}_J &= \text{Yield}_{4j} \prod_{j=4}^{j=J-1} \left[a_0 + \frac{a_1}{1 + (j-4)} \right] = \\ \text{Yield}_{t\bar{t}W^+, 4j} \prod_{j=4}^{j=J-1} \left[a_0 + \frac{a_1}{1 + (j-4)} \right] &+ \text{Yield}_{t\bar{t}W^-, 4j} \prod_{j=4}^{j=J-1} \left[a_0 + \frac{a_1}{1 + (j-4)} \right] \end{aligned} \quad (7.15)$$

The $t\bar{t}W^+$ and $t\bar{t}W^-$ events are treated separately and fitted in separate regions to account for the large difference in cross section between $t\bar{t}W^+$ and $t\bar{t}W^-$. This difference stems from the fact that $t\bar{t}W$ production comes from quark-quark interactions in the pp collisions, and the PDF's of the protons favor $t\bar{t}W^+$ production.

The $t\bar{t}W$ samples are fit to data in four regions defined (as shown in Table 8.3) to enhance $t\bar{t}W$ events above other background events. There are four parameters: $NF_{t\bar{t}W^+, 4j}$, $NF_{t\bar{t}W^-, 4j}$, a_0 , and a_1 . The NF's are fit in the four regions, and the correction factor is applied to all $t\bar{t}W$ sample events in the analysis. Table 7.2 shows the NF's for a fit to the CR1b(+) and CR1b(-) regions. The NF's associated with the yield at $N_{jets} = 4$ are large, indicating an underestimation of $t\bar{t}W$ events by the MC modeling prior to fitting. Figure 7.6 shows the CR1b(+) and CR1b(-) regions after the fit. Generally, the MC and data are in good agreement.

The post-fit N_{jets} distributions for each of the four regions is shown in Figure 7.7. There is good MC/data agreement in these regions. The NF's associated with the final fit to unblinded data are in Figure 10.7 in Chapter 10.

Parameter	Fit value
a1	$0.82^{+0.22}_{-0.21}$
a0	$0.26^{+0.09}_{-0.09}$
$N_{t\bar{t}W}$ 4j plus	$36.60^{+5.28}_{-4.75}$
$N_{t\bar{t}W}$ 4j minus	$20.66^{+3.47}_{-3.06}$

Table 7.2: The normalization factors derived from a fit to the CR1b(+) and the CR1b(-) regions.

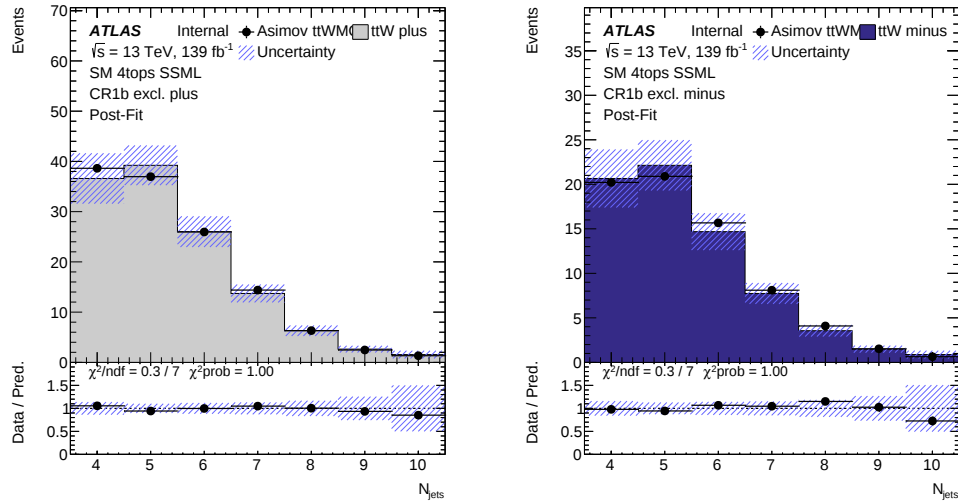


Figure 7.6: The N_{jets} distribution for the 1b $t\bar{t}W$ reweighting regions. The data distribution is compared to the MC $t\bar{t}W$ + jets distribution after reweighting.

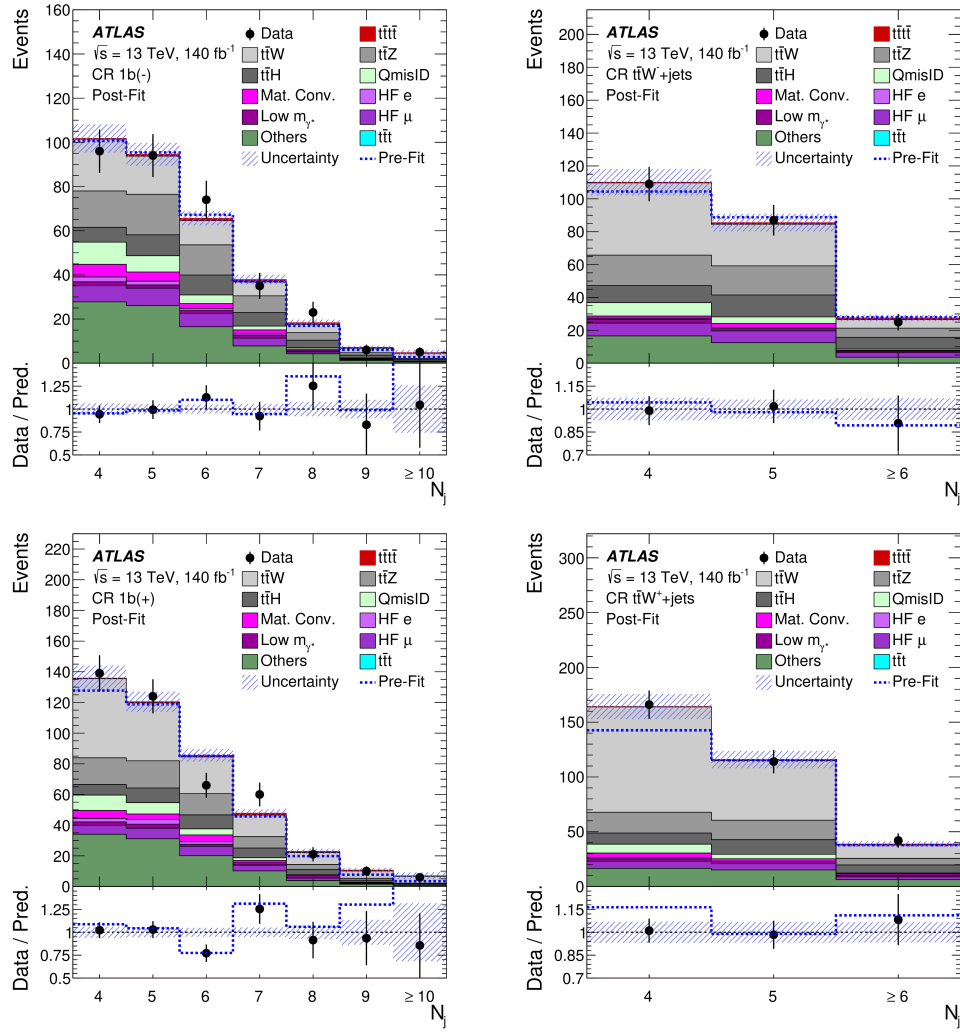


Figure 7.7: The N_{jets} distribution for the four $t\bar{t}W$ DD reweighting regions. The data distribution is compared to the MC $t\bar{t}W$ + jets distribution after reweighting.

CHAPTER 8

Signal Discrimination and MVA

8.1 Introduction

The observation of $t\bar{t}t\bar{t}$ production is challenging both due to the low Standard Model cross section, as well as the large and irreducible number of background events in the 1L and 2LOS decay channels. In particular, there are many $t\bar{t}$ events, produced with additional jets, that closely resemble the signal. The analysis strategy centers on choosing a region of phase space such that the $t\bar{t}t\bar{t}$ contribution in that region is significant.

The first step is to define signal regions with anticipated higher numbers of signal events, while excluding as much of the background as possible. One of the clearest differences between $t\bar{t}t\bar{t}$ and $t\bar{t}$ is that the former is expected to have, on average, twice as many jets and b-tagged jets as the latter. Therefore, the two major variables that define the regions are the jet and b-jet multiplicity. Within each region, the analysis further isolates the $t\bar{t}t\bar{t}$ events using a machine learning (ML) classifier. The output of the ML classifier ranks events by their probability of being $t\bar{t}t\bar{t}$, with high output values indicating a more “ $t\bar{t}t\bar{t}$ -like” event. The result within a region is a distribution with the highest value bins having most of the signal, and little of the background.

8.2 Region Definition

8.2.1 HBSM Analysis

The regions are divided by both number of jets per event and number of jets passing a b-jet selection [116]. For the 1L channel, the number of jets is 8, 9, or ≥ 10 . The number of b-tagged jets is 3, 4, or ≥ 5 . There is also a distinction made between three b-jets passing a higher-efficiency loose cut (3bL), and three b-jets passing a lower-efficiency tight cut (3bH). The efficiency definitions for 3b regions are shown in Table 8.1, including the 3b validation region. There are six background regions: 8j3bL, 8j3bH, 8j4b, 8j ≥ 5 b, 9j3bL, and 9j3bH. There are also six signal regions: 9j4b,

$9j \geq 5b, \geq 10j3bL, \geq 10j3bH, \geq 10j4b, \geq 10j \geq 5b$.

Name	$N_b^{60\%}$	$N_b^{70\%}$	$N_b^{85\%}$
3bL	≤ 2	$=3$	-
3bH	$=3$	$=3$	≥ 4
3bV	$=3$	$=3$	$=3$

Table 8.1: Event selection requirements for the 3bL, 3bH, and 3bV regions. The b-jet tagging efficiencies are 60%, 70%, and 85%. Note that 3bL only requires ≤ 2 b-jets in an event for the tightest working point efficiency, but three for the 70% efficiency. This means that an event with no b-jets at the 60% efficiency but three b-jets at the 70% efficiency would be a 3bL event.

For the 2LOS channel, the divisions are similar. The $t\bar{t}t\bar{t}$ event decays to the same number of b-quarks, resulting in the same number of expected b-jets. However, there is one less hadronic decay of the W , and so correspondingly fewer expected total jets. There are five background regions: 6j3bL, 6j3bH, 7j3bL, 7j3bH. There are four signal regions: $7j \geq 4b, \geq 8j3bL, \geq 8j3bH, \geq 8j \geq 4b$.

The 1L and 2LOS regions of phase space are shown in Figure 8.1. The signal regions are the regions with highest jet and b-jet multiplicity. There are also control regions, composed of mostly background events, which are used in the profile likelihood fit to constrain uncertainties (see Chapter 9). The validation regions are not used in the fit, but their distributions are plotted to ensure that the fit results have accurately modeled the background. There are also 2b regions that are employed for corrections to background samples, in particular for the MC $t\bar{t}$ samples as described in Chapter 7.

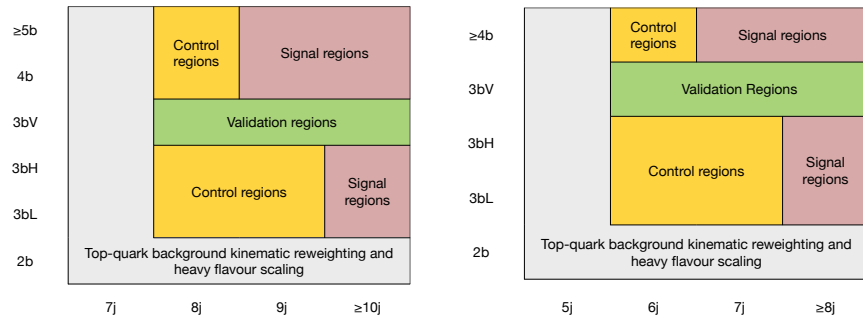


Figure 8.1: Visualization of region definitions based on jet and b-tagged jet multiplicity [116]. The 1L regions are shown on the left and the 2LOS shown on the right.

8.2.2 Higgs Oblique Analysis

For the SSML analysis, the control regions are defined with the intent of isolating a single type of background process. A fit to all of the control regions thus ensures that every type of major background process is well-modeled in the high jet and b-jet multiplicity signal region.

There are four regions designed to account for background processes involving the production of extra leptons. These leptons come from the creation of an e^+e^- pair from a virtual photon γ^* , from material conversion of a real photon interacting with detector material, and the creation of an electron or muon from a heavy flavor (HF) decay.

The γ^* internal conversion is usually caused by radiation of a photon from a W boson, with the W subsequently decaying into a lepton/neutrino. The virtual photon then creates an e^+e^- pair. Because the photon is off-shell, it is generally short-lived, and the resulting pair of leptons is generally low mass. The internal conversion happens at a smaller radius from the IP than material conversion, and can be identified by finding a pair of opposite-sign tracks within $\Delta R < 0.3$ and reconstructing the mass of the track pair.

Material conversion also creates a pair of tracks displaced from the primary vertex, with a larger electron decay radius but also low mass. Material conversion, internal conversion, and other fake lepton events are categorized using this information by the DFCommonAddAmbiguity (DFCAA) tool. The output of the DFCAA is shown in Table 8.2.

DFCAA Output	Description
-1	No 2nd track found
0	2nd track found but no conversion found
1	γ^* Internal conversion candidate
2	Material conversion candidate

Table 8.2: Descriptions of output values for the DFCAA tool.

The CR low m_{γ^*} region and CR material conversion region therefore have the same selections except for the DFCAA, which requires the material conversion region to have DFCAA = 2 for at least one lepton and DFCAA = 1 and $\neq 2$ for the low m_{γ^*} region.

Region	Channel	N_j	N_b	Other selection	Fitted var.
CR Low m_{γ^*}	SSee SSem	$4 \leq N_j \leq 6$	≥ 1	DFCAA(l_0 or l_1)=1 DFCAA(l_0 and l_1)!=2	1 bin
CR Mat. Conv.	SSee SSem	$4 \leq N_j \leq 6$	≥ 1	DFCAA(l_0 or l_1)=2	1 bin
CR HF μ	emm mmm		=1	$100 < H_T < 300\text{GeV}$ $E_T^{miss} > 50\text{GeV}$ DFCAA(l_0 and l_1 and l_2) < 0 abs(charge) = 1	p_T^{lep2}
CR HF e	eee eem		=1	$100 < H_T < 275\text{GeV}$ $E_T^{miss} > 35\text{MeV}$ DFCAA(l_0 and l_1) < 0 abs(charge) = 1	p_T^{lep2}
CR $t\bar{t}W^+$	SSem SSmm	≥ 4	≥ 2	$M_{ee} < 0$ or $M_{ee} > 0.1$, $ \eta(e) < 1.5$ for $N_b=2$, $H_T < 500\text{GeV}$ or $N_j < 6$ for $N_b \geq 3$, $H_T <$ DFCAA(l_0 and l_1) < 0 $\sum(charge) > 0$	N_{jets}
CR $t\bar{t}W^-$	SSem SSmm	≥ 4	≥ 2	$M_{ee} < 0$ or $M_{ee} > 0.1$, $ \eta(e) < 1.5$ for $N_b=2$, $H_T < 500\text{GeV}$ or $N_j < 6$ for $N_b \geq 3$, $H_T <$ DFCAA(l_0 and l_1) < 0 $\sum(charge) < 0$	N_{jets}
CR1b(+)	2LSS + 3L	≥ 4	=1	(SSee SSem) DFCAA(l_0 and l_1) ≤ 0 $H_T > 500\text{GeV}$ $\sum(charge) > 0$	N_{jets}
CR1b(-)	2LSS + 3L	≥ 4	=1	(SSee SSem) DFCAA(l_0 and l_1) ≤ 0 $H_T > 500\text{GeV}$ $\sum(charge) < 0$	N_{jets}
CR Low GNN	2LSS + 3L	≥ 6	≥ 2	$H_T > 500\text{GeV}$, GNN < 0.2	GNN
SR	2LSS + 3L	≥ 6	≥ 2	$H_T > 500\text{GeV}$	GNN

Table 8.3: Selections for each region of the SSML channels.

The HF decay regions are defined to have DFCAA < 0 for the two leading

leptons to exclude all possible conversion leptons. There are also cuts on H_T and missing E_T , shown below in the regions definition table, Table 8.3.

The CR low m_{γ^*} region, CR material conversion, and HF decay regions are used in a template fit to extract normalization factors (NF's) for the rate of fake and non-prompt leptons. A fit to data in these four regions allows for the extraction of NF's for HF decay to electron (NF_{HF_e}), HF decay to muon (NF_{HF_μ}), material conversion ($NF_{Mat.Conv.}$), and internal conversion ($NF_{Lowm_{\gamma^*}}$).

There are four regions designed to provide better modeling for the $t\bar{t}W^\pm$ processes, which constitute a major background in the SSML channel. There is a region enhancing the $t\bar{t}W^+$ process, one enhancing the $t\bar{t}W^-$ process, a low N-bjets region with positively charged leptons, and a low N-bjets region with negatively charged leptons. The regions definitions are shown in Table 8.3. The NF's for the $t\bar{t}W^\pm$ processes are estimated using a data-driven method described in Chapter 7.

There is a single signal region in the SSML analysis, defined by a high jet and b-jet multiplicity, and by large jet activity.

8.3 Multivariate Analysis

The basic idea behind multivariate analysis is that, instead of using a single variable, using the information from several variables at once can yield more powerful results [122]. An example, shown in Figures 8.2 and 8.3, is to consider the separation of two categories, A and B , each of which has events with the same associated variables, in this case two variables $\mathbf{x} = (x_1, x_2)$. In this example, the distributions of x_1 for A and B have large overlap, and so do the distributions for x_2 . However, the distributions for $x_1 + x_2$ are easily separable for the two classes. This is a case where a simple cut on either variable will not discriminate A from B , but a multivariate cut will.

This example can be expanded to events with N variables $\mathbf{x} = (x_1, x_2, \dots, x_N)$. Multivariate analysis then takes place within N -dimensional space. The objective is to construct some function of $\mathbf{x}$ and coefficients ω such that events in Class A are easily distinguishable from events in Class B.

Constructing such a function is usually impractical to do analytically. Instead, machine learning is employed to approximate the function. Specific algorithms, such as decision trees (DT's) and neural networks (NN's), have different approaches, the details of which will be discussed below. This analysis uses supervised machine learning, in which an algorithm is given both the input and the desired

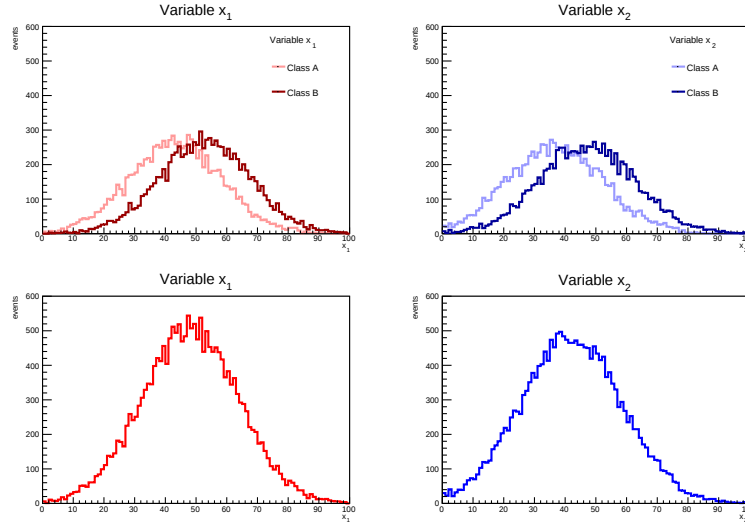


Figure 8.2: Distributions for two correlated variables, x_1 (left) and x_2 (right). The top row shows the distributions for Class A and Class B, and the bottom row shows the combined distribution. When each variable is considered individually, Class A and Class B are not easily separable.

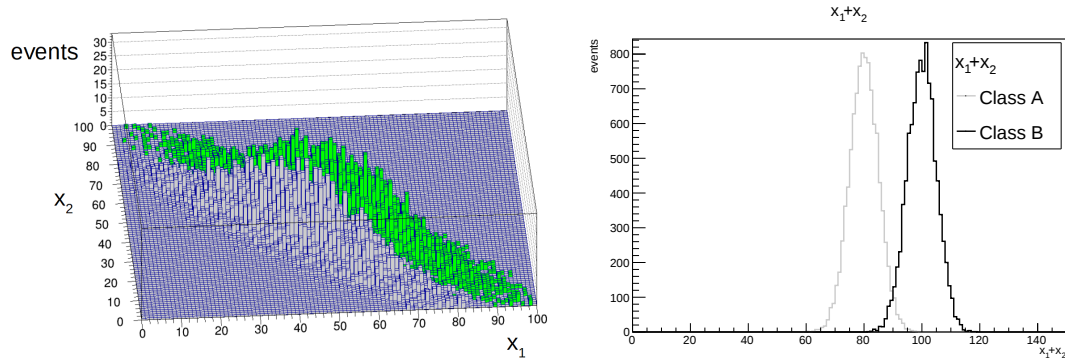


Figure 8.3: Left: Three-dimensional plot of x_1 and x_2 , showing the distribution for Class A in gray and Class B in green. Right: Class A and Class B have narrow and separable distributions for x_1+x_2 .

output, and is “trained” by learning to replicate the output. The trained algorithm can then be used on datasets for which the output is unknown. Machine learning, applied to the task of multivariate analysis, is a powerful tool that can often yield higher significance and greater separation power than simple cuts on single variables. However, some of the decision-making process is not transparent, and the function created by the algorithm is derived empirically.

8.4 Machine Learning Algorithms

As stated previously, the algorithms are solving a specific problem: for a given input set of variables or features $\mathbf{x}$, determine some $f(\omega, \mathbf{x})$ output that will provide information about how to categorize the event. For the case of this analysis, $f(\omega, \mathbf{x})$ is discrete and the available values represent the probability that an event is a $t\bar{t}t\bar{t}$ event. Each “step” of algorithm training involves recursively calculating an answer, comparing this answer with the target answer via a loss function, and adjusting the function parameters to reduce the loss in the following step. The process repeats until specific stopping conditions are reached, such as a decision tree’s Gini Index (defined below).

The quality of MVA algorithms is assessed by creating a receiver operating characteristic (ROC) curve, which for a binary classifier plots the rate of true positive results against the false positives [123]. The area under the ROC curve (AUC) value quantifies the efficacy of the algorithm. A larger AUC indicates a higher rate of true positives for a lower rate of false negatives, and therefore a better classifier.

The Boosted Decision Tree (BDT) was the primary classifier used for the SM measurement that resulted in evidence for $t\bar{t}t\bar{t}$ production [6]. In this analysis it is used to validate the results produced using a Graph Neural Network (GNN). The choice to use GNN results is based on the improved sensitivity when employing the GNN. This is shown in Figure 8.7, where the GNN classifier ROC curve has a higher AUC, and in Figure 8.8 where analysis with the GNN achieves a higher significance than analysis using the BDT.

8.4.1 Graph Neural Network

A neural network is comprised of neurons in the sense that information is aggregated in nodes, which are connected to each other and which can be “activated” or “suppressed” by the information they receive from other nodes [124]. In a feedforward neural network, there are several layers of nodes. Each node in a layer receives in-

formation from all of the nodes in the previous layer. This information is processed by a parameterizable function, and the output is usually a value within a continuous function ranging from $[0,1]$ [125]. This output then becomes the input for the next layer, until all layers have completed their processing.

This feedforward neural network can be improved in several ways. Firstly, a node could, in addition to considering all previous nodes, use the information on its own state during the previous iteration. This allows for short-term memory within the network and is especially useful for processing sequential information. The second improvement is to more efficiently organize the input information. This is a defining feature of a Graph Neural Network, or GNN: the variables of input events are either nodes corresponding to object data, such as jet p_T , or else edges that quantify the relationships between two nodes. For example, two jet p_T nodes would be connected by an edge that represents the ΔR between the jets.

Figure 8.4 demonstrates the steps of an iteration within the message-passing GNN algorithm:

1. There is a pair of input nodes, h_i and h_j , and they are related by an edge variable e_{ij} . Their values at time t are input into a neural network M_t , which uses a rectified linear unit (ReLU) function as the activation function [126]. A message m_{ij}^t is formed as the output.
2. All messages created between node i and the other nodes (m_{ij}^t , m_{ik}^t , m_{im}^t) are aggregated via a sum or mean to form m_i^t , the message for i at time t .
3. The node variables are then updated. The current value for node i , h_i^t , is combined with the message m_i^t via another neural network, H_t . Note that this step involves the node self-referencing its current value. After the information is combined in H_t , the value of the node is updated for the next iteration at time $t + 1$, becoming h_i^{t+1} .
4. There are three message-passing layers which update the node values. Once node updating is completed at time T , all the finalized node values (h_i^T , h_j^T , h_k^T) are used as input into a function R , which outputs a single classification value y . In the case of this analysis, y is a value between $[0,1]$ that represents how likely the event is to be a $t\bar{t}t\bar{t}$ event.

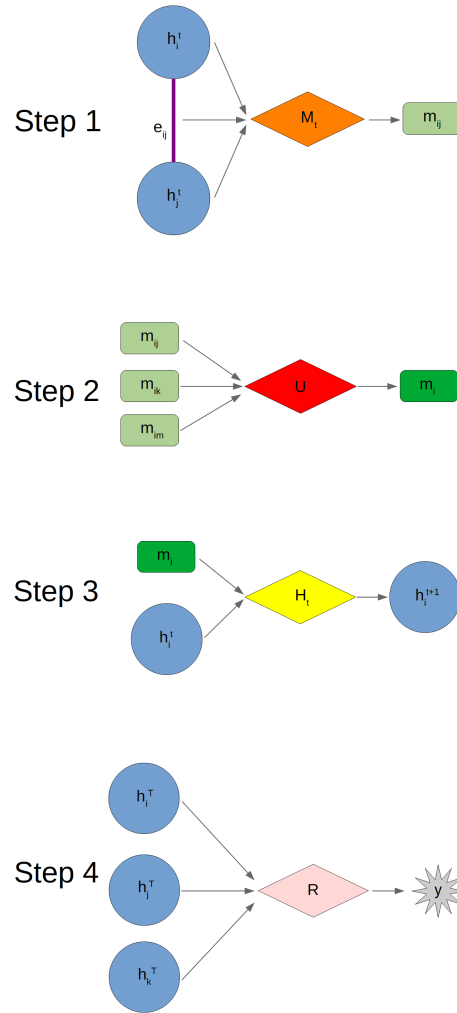


Figure 8.4: Flow of information in a GNN [127]. The first step processes messages between each pair of elements, including relational information, while the second step aggregates that information. The third step uses a neural network to update the elements' values, and the final step calculates a single output value for each pp event.

8.4.2 Decision Trees

Each step of a decision tree involves applying a cut at a single value for a single variable. The algorithm begins with a full dataset, consisting of both signal (s) and background (b) events described by a set of N variables $\mathbf{x} = x_1, x_2, \dots, x_N$. The

decision tree operates as follows [128]:

1. Evaluate the separation power when splitting the dataset by some value of x_1 .
2. Evaluate the separation power when splitting the dataset by some value of $x_2 \dots x_N$.
3. If a stopping criteria is satisfied, then do not split the dataset and exit the algorithm.
4. If no stopping criteria is satisfied, split the dataset according to the value of x_i that yields the greatest separation power.
5. Iterate on each of the new daughter datasets.
6. The algorithm is finished when all datasets have satisfied a stopping criteria.

A singular decision tree has the advantage of transparency and stability against missing or irrelevant variables, but its separation power is lower than other methods and can be sensitive to minor changes in the training dataset. To address these disadvantages, multiple decision trees are used in a method called “boosting.” A Boosted Decision Tree (BDT) averages its results, giving greater weight to trees with a lower error. Trees can also be trained multiple times and resampling random events. This is called “bagging” and, because typical events are more likely to be selected than outlier events, bagging reduces the impact of statistical fluctuations on the BDT output.

This analysis uses BDTG, a BDT available through TMVA [96]. The BDTG evaluates the separation power at 30 different values for each variable. A purity p is defined as the ratio of signal events to total events. In particular, the separation power is expressed as a Gini Index:

$$p \times (1 - p) \tag{8.1}$$

The stop criterion for this BDT is dictated by the size of the dataset. When a splitting reduces the dataset in that category below 100 events, the BDT algorithm terminates on that splitting.

The boosting algorithm, GradientBoost, combines a number M of the decision trees via a weighted sum. Each individual decision tree has an output $f(x; a_m)$,

where x is the set of input variables for an event and a_m is an adjustable parameter. These trees are weighted by another adjustable parameter, β_m , and combined to form an overall output, $F(x; P)$:

$$F(\mathbf{x}; P) = \sum_{m=0}^{M-1} \beta_m f(x; a_m); P \in \{\beta_m; a_m\}_0^M \quad (8.2)$$

The boosted output, $F(x; P)$ is then optimized by adjusting the parameters P to minimize the loss function. The loss function, in this case, is a binomial log-likelihood function:

$$L(F, y) = \ln(1 + e^{-2F(\mathbf{x}; P)y}) \quad (8.3)$$

where y is the true output value. This analysis makes use of 600 decision trees to boost the classifier. These trees each sample 60% of the available training dataset in order to bag the classifier.

8.4.3 HBSM Analysis

The HBSM GNN is constructed using PyTorch Geometric[129], inspired by the graph_nets framework[130]. A rectified linear unit (ReLU) function is used as the activation function for the inner NN's, and a sigmoid activation function is used for the output of the final layer. A binary cross-entropy function is used as the loss function.

The BDT framework is BDTG[96], and the loss function is a binomial log-likelihood, defined in Section 8.4.2.

Both the BDT and the GNN are trained to separate $t\bar{t}t\bar{t}$ from its dominant background, $t\bar{t} + jets$. No other background samples are used. For each, two MVA's are developed: one for the 1L channel trained on $n_{jets} \geq 9$, $n_{b-jets} \geq 3$, and one for the 2LOS channel trained on $n_{jets} \geq 7$, $n_{b-jets} \geq 3$. Each MVA has hyperparameters, such as the number of trees used in the BDT. These are optimized based on the ROC curve as well as the stability of the MVA. The input variables for the HBSM BDT are shown in Table 8.4 and the input variables for the HBSM GNN are shown in Table 8.5.

Variable	Description
H_T	sum of transverse momentum for all leptons and jets
$\sum_i p_{Ti} / \sum_i E_i$	centrality
N_{jets}	jet multiplicity
p_{T0}	leading jet transverse momentum
ΔR_{jj}^{avg}	average angular separation of all jet pairs
$M_{jjj}^{min, \Delta R}$	invariant mass of jet triplet with minimum angular separation
$M_{jjj}^{\Delta R < 3}$	average invariant mass of all jet triplets with $\Delta R < 3$
$\sum_{i=0}^3 p_{Ti} / \sum_{i \geq 4} p_{Ti}$	ratio of p_T sum of first four jets to p_T sum of remaining jets
ΔR_{bl}^{min}	minimum ΔR among pairs of b-tagged jet and lepton
ΔR_{bb}^{min}	minimum angular separation among all pairs of b-tagged jets
M_{bbb}^{avg}	average invariant mass of all b-tagged jet triplets
M_{bb}^{min}	minimum invariant mass of b-tagged jet pair
$\sum_{i=0}^{i < 4} pcb_i$	sum of pcb score of the four jets with highest DL1r score
N_{RC100}	number of RC-jets with a mass higher than 100GeV
$\sum d_{12}$	sum of the first k_t splitting scale d_{12} of all RC-jets
$\sum d_{23}$	sum of the second k_t splitting scale d_{23} of all RC-jets
E_T^{miss}	missing transverse energy
$(m_T(l, E_T^{miss}))$	W reconstructed transverse mass (1L only)
m_{ll}	invariant mass of two leptons (2LOS only)

Table 8.4: Input variables for HBSM BDT.

An MVA's process can become too closely adapted to its training dataset. This makes its classification abilities good for its training dataset, but when applied to any other dataset, the classifier's separation power is worse. This phenomenon is known as "overtraining." Generally, a comparison between a testing and training dataset will show the effects of overtraining. As seen in the ROC curves of Figures 8.7 and 8.8, the AUC difference is very small, indicating that neither the BDT nor the GNN are highly affected by overtraining.

To further ensure that overtraining is not adversely impacting the analysis, a two-fold validation is used. The dataset is split into two equal parts, designated "odd" and "even." The MVA is first trained to minimize the loss function: once using the odd samples, and once using the even samples. The two MVAs' performances are optimized separately. If the odd samples were used to minimize the loss function, the even samples are used in the performance step, and vice-versa. This ensures that no MVA is used on the dataset upon which it was trained. This is the case for both the BDT and GNN.

Variable	Description
Nodes	
p_T	object transverse momentum
η	object pseudo-rapidity (0 for E_T^{miss} node)
E	object energy
pcb	pseudo-continuous DL1r b-tagging score (0 for non-jets)
-	object type (-2 for muons, -1 for electrons, 0 for E_T^{miss} , 1 for jets)
Edges	
$\Delta\phi$	angular separation between pairs of objects
$\Delta\eta$	
ΔR	
Globals	
H_T	sum of transverse momentum for all leptons and jets
$\sum_{i=0}^3 p_{Ti} / \sum_{i \geq 4} p_{Ti}$	ratio of p_T sum of first four jets to p_T sum of remaining jets
$M_{jjj}^{\Delta R < 3}$	average invariant mass of all jet triplets with $\Delta R < 3$
$(m_T(l, E_T^{miss}))$	W reconstructed transverse mass (1L only)
m_{ll}	invariant mass of two leptons (2LOS only)
$\sum_i p_{Ti} / \sum_i E_i$	centrality
N_{jets}	jet multiplicity
N_{RC100}	number of RC-jets with mass greater than 100GeV
M_{bbb}^{avg}	average invariant mass of all b-tagged jet triplets
ΔR_{bl}^{min}	minimum ΔR among pairs of b-tagged jet and lepton
ΔR_{bb}^{min}	minimum angular separation among all pairs of b-tagged jets
ΔR_{jj}^{avg}	average angular separation of all jet pairs
$\sum_{i=0}^{i < 6} pcb_i$	sum of pcb score of the six jets with highest DL1r score
$\sum d_{12}$	sum of the first k_t splitting scale d_{12} of all RC-jets
$\sum d_{23}$	sum of the second k_t splitting scale d_{23} of all RC-jets

Table 8.5: Input variables for HBSM GNN.

The BDT distributions of signal and background for both a sample testing and training set are shown in Figure 8.5. The GNN distributions for all testing and training sets are shown in Figure 8.6.

8.4.4 Higgs Oblique Analysis

The Higgs Oblique analysis uses a GNN as the signal/background classifier, as the GNN was shown to improve analysis sensitivity over the BDT [25]. The analysis uses the graph_nets framework and open-source code [130]. The GNN uses a log loss

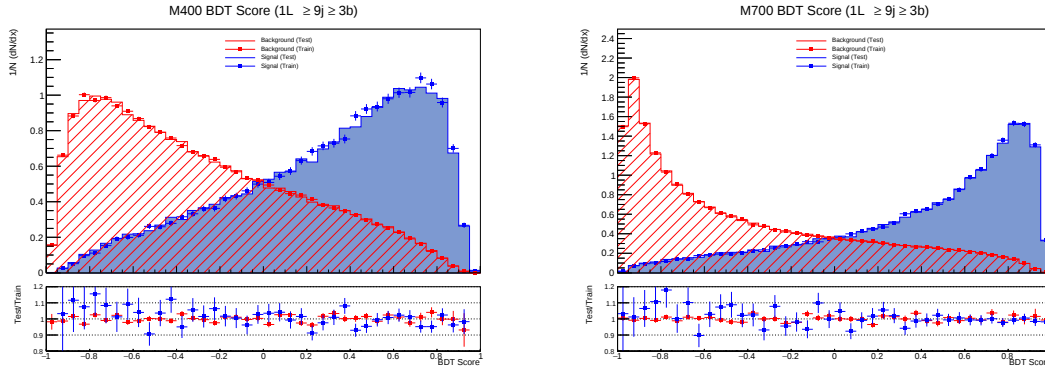


Figure 8.5: BDT distributions for a testing/training sample for both signal $t\bar{t}t\bar{t}$ and background $t\bar{t}+jets$. The distribution is shown for two different m_H mass points in the $1L \geq 9j \geq 3b$ region.

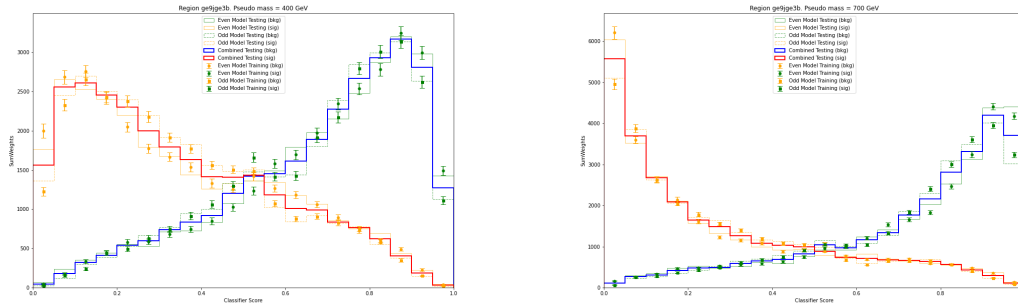


Figure 8.6: BDT distributions for all testing/training datasets for both signal $t\bar{t}t\bar{t}$ and background $t\bar{t}+jets$. The distribution is shown for two different m_H mass points in the $1L \geq 9j \geq 3b$ region.

Variable	Description
Nodes	
p_T	object transverse momentum
pcb	pseudo-continuous DL1r b-tagging score (0 for non-jets)
-	lepton charge
-	object type (-2 for muons, -1 for electrons, 0 for E_T^{miss} , 1 for jets)
Edges	
$\Delta\phi$	angular separation between pairs of objects
$\Delta\eta$	
ΔR	
Globals	
N_{jets}	jet multiplicity

Table 8.6: Input variables for Higgs Oblique GNN.

function and a sigmoid activation function.

The HBSM analysis uses the 1LOS decay channels and has one dominant background. In contrast, the Higgs Oblique analysis uses the SSML decay channels, and so the $t\bar{t}t\bar{t}$ signal must be isolated from several important backgrounds. Therefore, the dataset for the background is comprised of all MC pre-fit backgrounds, including a QmisID $t\bar{t}$ background. The signal is the LO SM $t\bar{t}t\bar{t}$ sample.

Weights for $t\bar{t}W$ sample were optimized to maximize the AUC of the ROC curve. Better discrimination from the $t\bar{t}W$ background was found to increase the effectiveness of the GNN as a whole. Weights between 1-9 were tested on the $t\bar{t}W$ sample, with a value of 3 giving an optimum result. All other samples have a weight of 1. The input variables for the Higgs Oblique GNN are shown in Table 8.6.

To alleviate overtraining effects, the k -folding method is used, with the number of folds being $k = 6$. The entire dataset is divided into six independent subsets based on eventNumber. For each subset i , the GNN is trained on all subsets except i , and then analysis is performed on i . The number $k = 6$ was determined based on overall performance and AUC.

The GNN distributions of background and $t\bar{t}t\bar{t}$ NLO are shown in Figure 8.9. The ROC curve is shown in Figure 8.10.

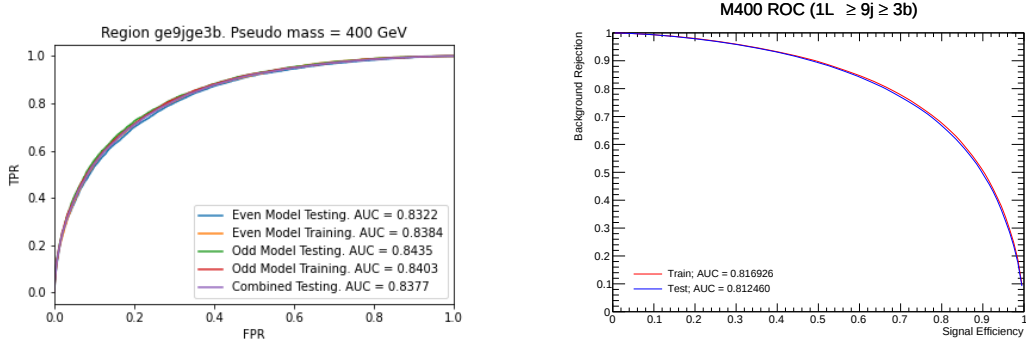


Figure 8.7: Comparison of GNN and BDT performance when classifying $t\bar{t}$ and $t\bar{t}t\bar{t}$ events. The ROC curve for the GNN is shown in the left plot, and the ROC curve for the BDT is shown on the right. Note that curves are different because the axes display different efficiencies. For both, however, an ideal classifier will maximize the AUC.

8.4.5 Quark-Gluon Tagging in HBSM

For the 1LOS channel, the high jet multiplicity of the signal regions means that any background events must have several extra jets. For example, to qualify for the 9j regions, $t\bar{t}$ events must have, in addition to the two b-jets produced by the top decays, seven extra jets. One of the W bosons must decay leptonically to be part of the 1L channel. If the other W boson decays hadronically, creating two jets, then there are five jets that are not produced as part of the hard $t\bar{t}$ process. These extra jets are typically softer and come from a light quark or gluon decay.

A quark/gluon tagger was introduced to exploit this feature of background event jets and to enhance the existing $t\bar{t}t\bar{t}$ vs. $t\bar{t}$ discriminant. The tagger is itself a BDT classifier designed to use jet information to distinguish quarks from gluons. The input variables are n_{track} , jet p_T , $|\eta|$, and two variables expressing jet width and momentum correlation. The jet width w_{track} is defined by the sum of angular separation of track i from the jet, weighted by track p_T . The two point momentum correlation, C_β , is calculated using the angular separation of track i from track j , where β is a parameter set to $\beta = 0.2$

$$w_{track} = \frac{\sum_i p_{T,i} \Delta R_i}{\sum_i p_{T,i}} \quad (8.4)$$

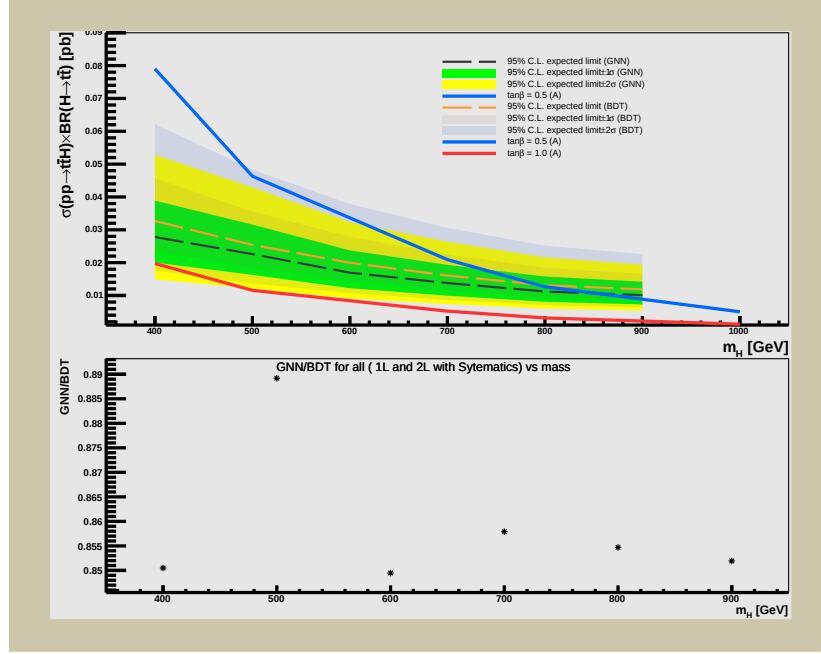


Figure 8.8: A direct comparison of the sensitivity between the two MVA methods. The orange dotted line shows the mean sensitivity as a function of mass using the BDT. The dotted black line shows the mean sensitivity when using the GNN. The bands indicate the 1 and 2σ confidence intervals. There are also two solid lines showing curves for theoretical cross sections of the Heavy Higgs-associated $t\bar{t}t\bar{t}$ process.

$$C_{\beta=0.2} = \frac{\sum_{ij} p_{T,i} (\Delta R_{ij})^\beta}{(\sum_i p_{T,i})^2} \quad (8.5)$$

The quark/gluon (q/g) tagger has three working point (WP) efficiencies for 50%, 80%, and 90%. There are three separate cut values for quark-tagging and for gluon-tagging. In each instance, the q/g tagger was added as an input to the discriminant MVA, expressed as the number of jets in an event passing a particular tagging cut. The MVA was trained on $t\bar{t}$ and $t\bar{t}t\bar{t}$ samples in the $\geq 9j$ and $\geq 3b$ regions for 1L and $\geq 7j$ and $\geq 3b$ regions for 2LOS. The impact of various configurations of the q/g tagger was tested by comparing the resulting ROC curve with the nominal ROC curve without any q/g tagger variable.

The greatest improvement was found to be from an 80% gluon tag cut on jets

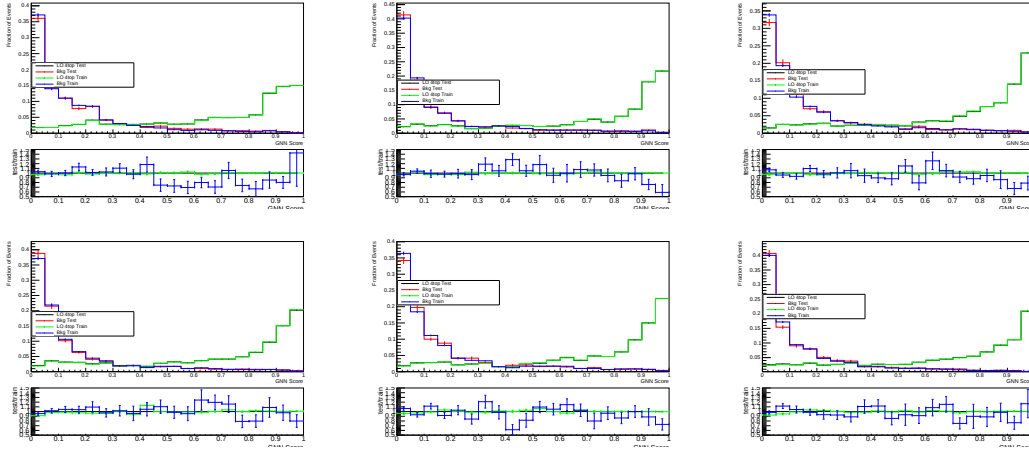


Figure 8.9: GNN distributions for each testing dataset i , compared to GNN distributions in datasets not equal to i . The $t\bar{t}t\bar{t}$ LO sample is shown, which was used for the training of the GNN. This was done to avoid the negative weights present in the NLO sample.

failing a DL1r b-tag cut at 85% WP efficiency. The addition of the b-jet exclusion enhanced the gluon tagging because the q/g tagger is better at distinguishing between light/c-jets and gluon jets, but does not perform as well distinguishing between b-jets and gluon jets. Thus the tagger is only applied to jets that are not identified as b-jets by the b-jet tagger.

This can be seen in Figure 8.11, which shows the percentage of jets in the $1L \geq 3b$ region passing certain cuts. 26% of the total $1L \geq 3b$ jets are truth b-jets, and 17% of the total $1L \geq 3b$ jets pass the 80% WP gluon tag. In other words, 66% of true b-jets are tagged as gluons. When the 85% WP b-exclusion is applied, less than 7% of truth b-jets pass the cut. In other words, more than 93% of true b-jets are excluded by the DL1r cut. Of these, 71% pass the gluon tag. This illustrates the poor b-jet/gluon-jet discrimination power of the q/g tagger, but also shows how much the addition of the b-tagger can aid the q/g tagger.

Figure 8.12 shows that, of the total jets in events with ≥ 3 b-jets, about 45% of jets from $t\bar{t}$ are truth gluon-jets. Then, 77% of truth gluon-jets are correctly tagged by the q/g tagger. Independently, 82% of truth gluon-jets are excluded by the DL1r b-exclusion cut. The last point on the plot shows that 63% of truth gluon-jets in the $t\bar{t}$ sample are both b-excluded and gluon-tagged. This figure also illustrates

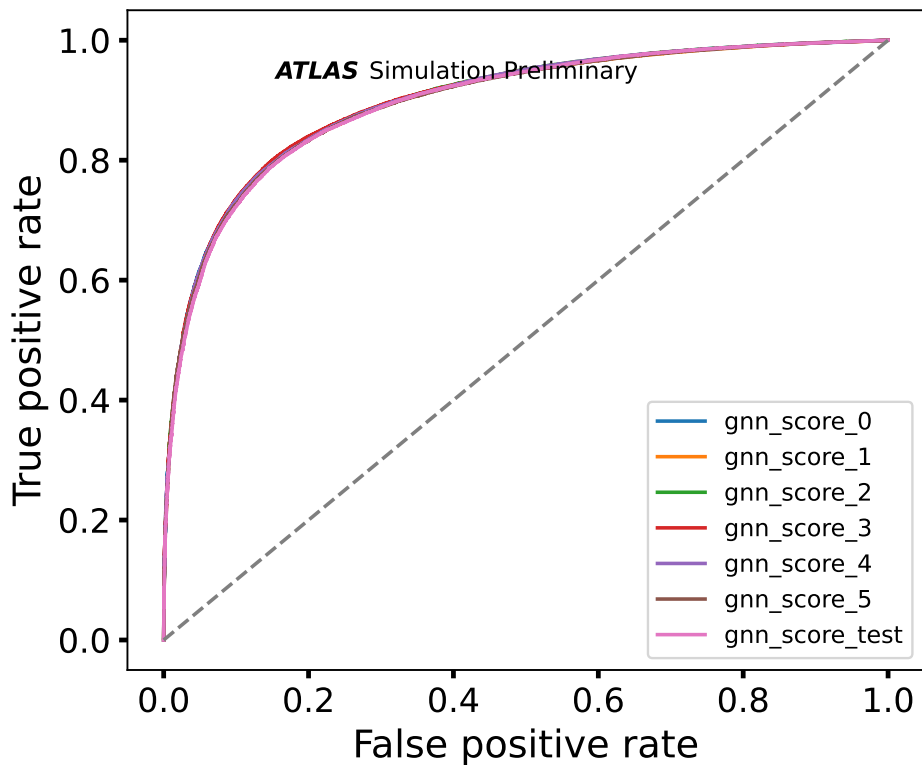


Figure 8.10: ROC curves for each testing dataset i as well as the combined result. The dashed line represents the efficiency of a random classifier.

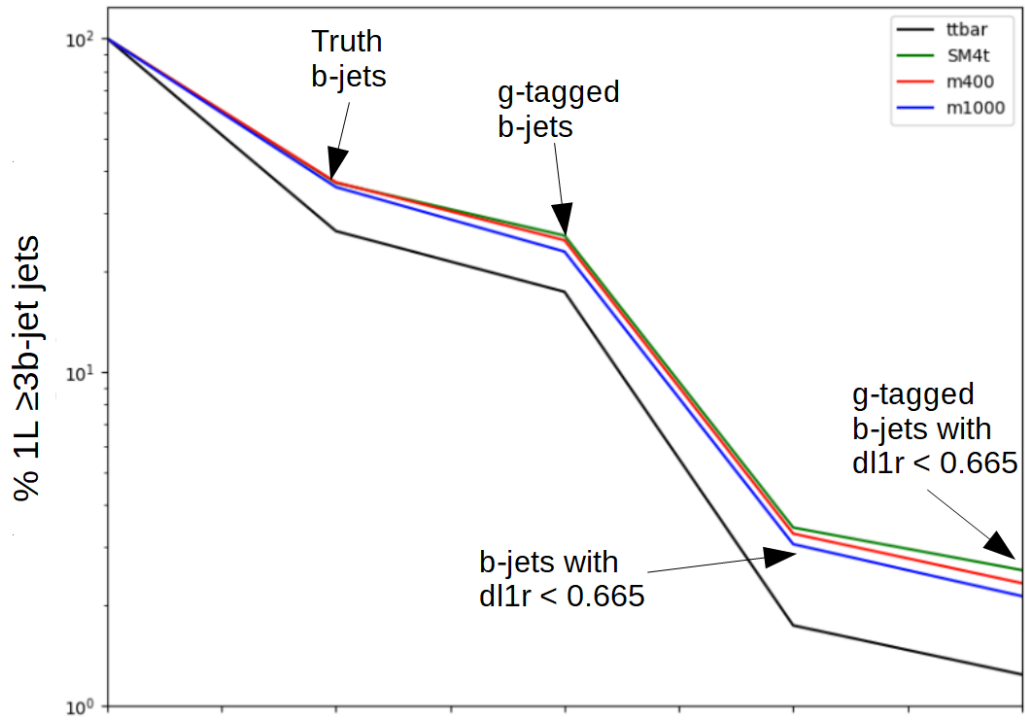


Figure 8.11: Composition of jets for events with ≥ 3 b-jets, with the focus on percentage of b-jets as different cuts are applied.

the large difference in the number of gluon-jets between the $t\bar{t}$ sample and any of the $t\bar{t}t\bar{t}$ samples.

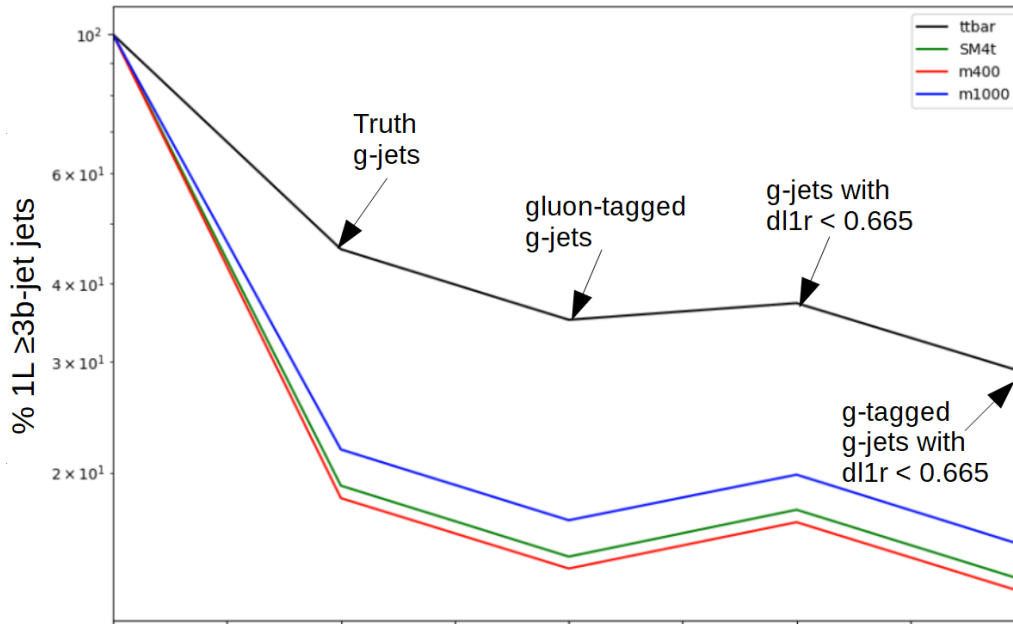


Figure 8.12: Composition of jets for events with ≥ 3 b-jets, with the focus on percentage of gluon jets as different cuts are applied.

Figure 8.13 shows a comparison of the ROC curves for the BDT discriminant with and without the b-excluded gluon tagging. The AUC improvement for $t\bar{t}$ vs. $t\bar{t}t\bar{t}$ associated with a heavy Higgs of mass $m_H = 400\text{GeV}$ is roughly 10%. Figure 8.14 shows the GNN AUC as a function of m_H , demonstrating noticeable improvement when the q/g tagging variable is added. The performance of both MVA's is enhanced by the introduction of the q/g tagger as a b-excluded gluon-tagger.

The q/g tagger was not incorporated into the final HBSM analysis strategy because the fully developed q/g tagger was not available on the timeline of the analysis. However, future analyses involving signal regions with high jet multiplicity might benefit from the addition of a b-excluded gluon tagging option.

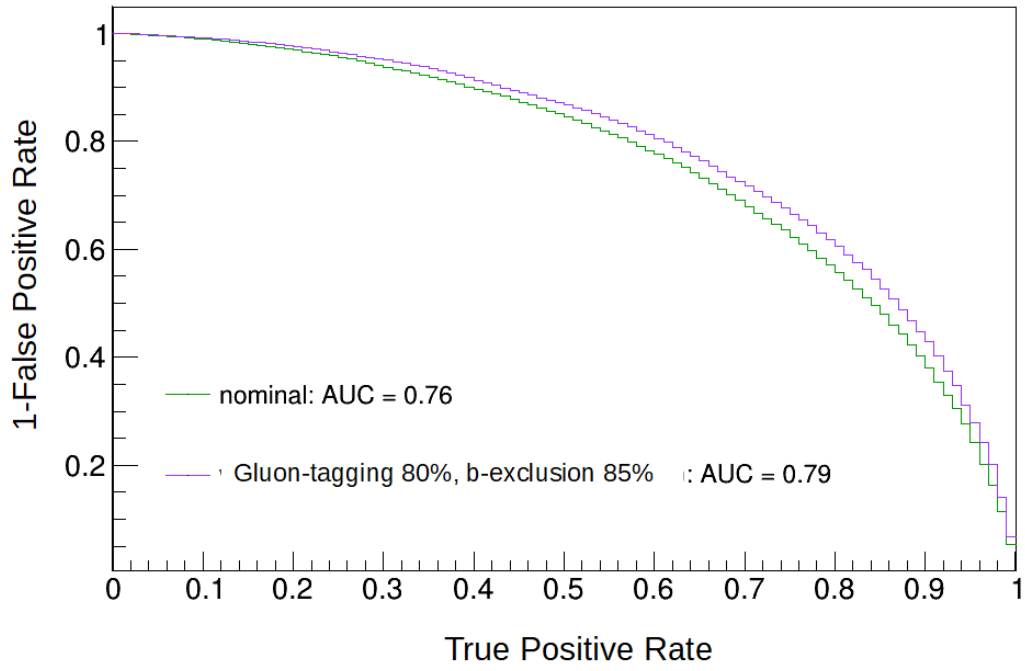


Figure 8.13: ROC curves for the BDT classifier, separating $t\bar{t}$ events from $t\bar{t}t\bar{t}$ events associated with Heavy Higgs of $m_H = 400\text{GeV}$. The AUC improves by about 10% when the q/g tagging variable is added.

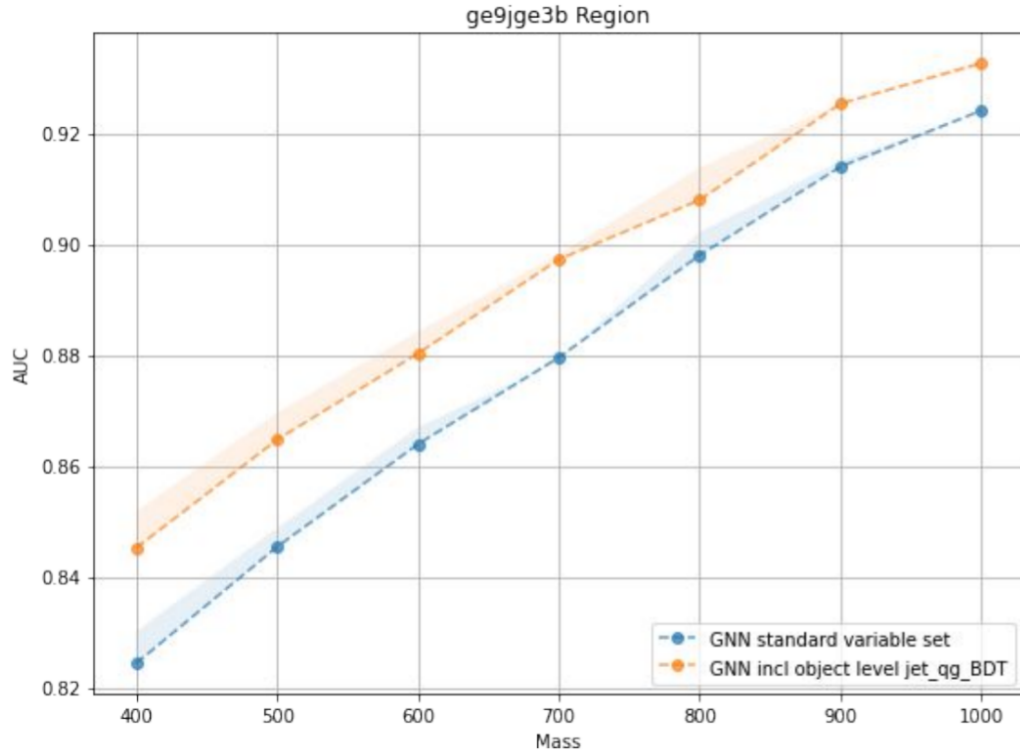


Figure 8.14: AUC values for the GNN classifier, as a function of Heavy Higgs mass associated with the $t\bar{t}t\bar{t}$ signal. The improvement when adding the q/g tagging variable is small but consistent across mass points.

CHAPTER 9

TRExFitter and Profile Likelihood Fits

9.1 Introduction

For new particle searches or rare processes, there is a convention for reporting results that acknowledges the probabilistic nature of the processes being studied. The convention is not to declare whether or not a process exists, but rather to quantify the level of confidence that it does. For example, the data are analyzed as a binned distribution of a chosen variable. There is a prediction of what this distribution would look like with only known background contributions (the null hypothesis) as well as a prediction of how the added signal process would affect the distribution. The objective is to determine how well the null hypothesis and the signal hypothesis match the data.

Profile likelihood fitting is used to accomplish this objective. The profile likelihood fitting process performs several tasks. First, it expresses the compatibility of a hypothesis to data as a single, meaningful value. Secondly, it incorporates uncertainties in the data-taking process and analysis. Lastly, it can output the most likely value for a signal cross section. This is expressed as a “strength” or ratio of the observed cross section to the predicted cross section, where the predicted cross section is the SM cross section. A strength of one corresponds to agreement with the prediction. Alternatively, profile likelihood fitting can place limits on the upper value of the signal strength. The program called TRExFitter is used to perform likelihood fitting for this analysis.

9.2 Profile Likelihood Fitting

Profile likelihood fitting is a series of steps by which a distribution is transformed into a probability function depending on signal strength and systematics, then into an optimized ratio, then into a test statistic depending only on signal strength. For this analysis, the input distribution is the output of a multivariate analysis (MVA) classifier, wherein an event scoring a high value is judged as more likely to be a $t\bar{t}\bar{t}$ event (see Chapter 8). The systematic uncertainties are those involved in the data-

taking process at the LHC, in the MC modelling of the background and signal events, and within the theory. The outputs of the profile likelihood fitting are threefold: adjustments to the uncertainty values, or “pulls,” reductions of the possible range of uncertainty values, or “constraints,” and the value of the $t\bar{t}t\bar{t}$ cross section that best matches the input distribution relative to the SM cross section, or “signal strength”.

There is also convention regarding the results. In particular, the test statistic that is produced by the fitting is turned into a p-value and its equivalence, Z . Conceptually, the p-value is the probability that random fluctuations in the null hypothesis could reproduce the data. A low p-value indicates that such a fluctuation is unlikely and, by extension, there is likely to be signal present. The significance is obtained by considering a normalized Gaussian distribution. The value of the significance is the number of standard deviations required to create an area under the curve equal to the p-value, as shown in Figure 9.1. If the p-value is small, then a larger number of standard deviations from the norm are required, resulting in a high significance. The standard for the field of high-energy physics is to claim ‘evidence’ for a process at a significance of 3σ , or a p-value of 0.0013, and to claim ‘discovery’ at a significance of 5σ and a p-value of 2.87×10^{-7} [131].

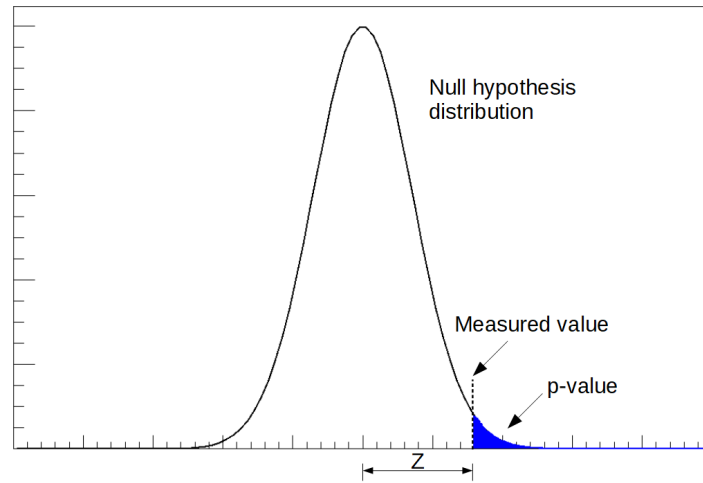


Figure 9.1: Visualization of p-value and significance as the difference between the mean of the hypothesis distribution and the measured value.

The first step of profile likelihood fitting is to transform an input distribution into a likelihood function. Consider a distribution that has N bins corresponding to different values of a variable x . There are background contributions to each bin and also signal contributions, the number of which depend on signal strength μ . There

are also a set of systematic uncertainties $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_b$ corresponding to the signal and background. The signal and background yields (s_i and b_i , respectively) in the i th bin are described by probability density functions $f(x; \theta)$ such that:

$$s_i = s_{tot} \int_{bin\ i} f_s(x; \boldsymbol{\theta}_s) dx, \quad b_i = b_{tot} \int_{bin\ i} f_b(x; \boldsymbol{\theta}_b) dx \quad (9.1)$$

The values s_{tot} and b_{tot} are the total mean number of signal and background events. The expectation value E for observed events n_i is:

$$E[n_i] = \mu s_i + b_i \quad (9.2)$$

In addition to the phase space where signal contribution is expected, there are also validation regions of the phase space. These background-dominated regions are included to better characterize the values of θ . For a histogram with M bins, the expectation value E for observed events m_i is dependent on some calculable background prediction u_i :

$$E[m_i] = u_i(\boldsymbol{\theta}) \quad (9.3)$$

Conceptually, the likelihood function is a means of comparing the observed and expected yields for each bin. In particular, a Poisson probability is used, of the form:

$$P = \frac{\lambda^k e^{-\lambda}}{k!} \quad (9.4)$$

where λ corresponds to the number of expected events in a bin and k is the number of observed events. For the signal region histograms, this Poisson probability is written as a product of N bins:

$$L_{sig}(\mu, \boldsymbol{\theta}) = \prod_{j=1}^N \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \quad (9.5)$$

Here, n_j is the number of observed events in the j th bin. For the background histograms (in validation and control regions), the same probability is expressed for M bins, where m_k is the number of observed events in the k th bin and u_k , the number of expected events, is also implicitly dependent on the uncertainties $\boldsymbol{\theta}$:

$$L_{bkg}(\mu, \boldsymbol{\theta}) = \prod_{k=1}^M \frac{u_k^{m_k}}{m_k!} e^{-u_k} \quad (9.6)$$

The likelihood function combines the Poisson probabilities associated with all bins for both the background and signal histograms:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^N \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^M \frac{u_k^{m_k}}{m_k!} e^{-u_k} \quad (9.7)$$

A large likelihood value indicates a high probability that the nuisance parameters $\boldsymbol{\theta}$ and signal strength μ will match the observed yields. Both $\boldsymbol{\theta}$ and μ can be adjusted, and in particular can be optimized to achieve the highest possible value of L . This is called the “unconditional likelihood maximum,” expressed as $L(\hat{\mu}, \hat{\boldsymbol{\theta}})$. The “conditional likelihood maximum” is the maximum value of L for a fixed value of μ , and is expressed as $L(\mu, \hat{\boldsymbol{\theta}})$.

The conditional maximum will always be less than or equal to the unconditional maximum. Therefore, a ratio can be defined that is dependent only on signal strength μ :

$$\lambda(\mu) = \frac{L(\mu, \hat{\boldsymbol{\theta}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \quad (9.8)$$

The unconditional likelihood maximum represents the case where μ and $\boldsymbol{\theta}$ have been adjusted to match the observed data as closely as possible. The conditional likelihood maximum represents the closest match to data when μ is at a fixed value. Therefore, the value of $\lambda(\mu)$ scales from 0 to 1, where 1 indicates that the given value of μ is the same as the unconditional value, i.e. the value that most closely matches what is observed in the original input data distributions. Maximizing the ratio in Equation 9.8 leads to an output ideal μ and the nuisance parameter set $\boldsymbol{\theta}$. This is

referred to as a “profile likelihood fit” and the resulting μ and θ are assumed to best represent the data.

This ratio can be expressed as an equivalent value called the test statistic:

$$t_\mu = -2\ln\lambda(\mu) \quad (9.9)$$

There is a probability density function (PDF) associated with t_μ . This is an independent input and, for this analysis, is assumed to be a noncentral chi-square distribution [131]. The value of t_μ obtained by the likelihood fitting is denoted as $t_{\mu,obs}$ and the p-value of the fit is thus the area under the t_μ pdf beginning at $t_{\mu,obs}$:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu) dt_\mu \quad (9.10)$$

This can then be transformed into a significance:

$$Z = \Phi^{-1}(1 - p) \quad (9.11)$$

The Φ^{-1} represents the quantile, or the number of standard deviations required to get probability $1 - p$. This is the same definition as was discussed in the introduction.

High compatibility between data and the null hypothesis will result in a large profile likelihood ratio, which in turn translates to a small test statistic, a large p-value, and a low significance. The inverse is also true. Comparing the input data to a null hypothesis means calculating Z using a test statistic pdf assuming a background-only distribution. Using $t_{\mu,obs}$ associated with the maximum likelihood ratio yields an “observed” p-value, and using t_μ associated with $\mu = 1$ yields an “expected” p-value.

In the case of this analysis, the data are compared to a background-only hypothesis that includes all Standard Model processes except for the predicted $t\bar{t}t\bar{t}$ production.

9.3 Setting Limits

Often in an analysis, it is meaningful to constrain the parameter of interest. Even if there is not sufficient evidence for a process, the analysis can still draw conclusions about certain relevant quantities. For example, this analysis seeks to demonstrate the existence of BSM $t\bar{t}t\bar{t}$ production, extracting a value for either Heavy Higgs mass or the Higgs Oblique parameter. However, the analysis can also constrain the limits for these parameters regardless of the significance.

Instead of finding a significance based on the extracted signal strength, the reverse is performed. A desired significance (or equivalent p-value) is used as the input. Using similar formulae as before, a signal strength can be calculated. In particular, assuming a signal that only enhances the cross section, a slightly different test statistic is defined:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \geq \mu \\ 0 & \hat{\mu} < \mu \end{cases} \quad (9.12)$$

Where, as in Equation 9.8, $\hat{\mu}$ represents the value of μ corresponding to the unconditional likelihood maximum. And the formula for the p-value is largely the same as before:

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu \quad (9.13)$$

Conceptually, a limit represents how much signal strength is needed to achieve a certain significance. The analysis strategy plays an important role in such a limit. Consider two cases: an analysis that uses regions of phase space where there is large background and small fractions of signal yield, and an analysis where at least a few regions have large fractions of signal yield. The former is going to require higher signal strength than the latter in order to achieve the same p-value. This is why finding limits is also useful for assessing the sensitivity of an analysis strategy.

9.4 TRexFitter

For this analysis, profile likelihood fitting is performed using a package called TRexFitter. TRexFitter uses a RooFit framework, which allows the various distributions,

probabilities, and functions to be expressed in ROOT-compatible C++ language [132]. TRExFitter reads in the distributions associated with background and signal, as well as any nuisance parameters. A separate tool, HistFactory, is employed to create the likelihood functions [133]. The likelihood functions are stored in an XML workspace, and, during the fitting process of TRExFitter, the functions are adjusted to maximize the likelihood ratio in Equation 9.8. In addition to the statistical analysis, TRExFitter also has plotting capabilities to visualize both yields and nuisance parameters.

The input histograms for TRExFitter include both ATLAS data and Monte Carlo simulated pp collision events. For the background regions, the likelihood fit is performed for distributions of the sum of transverse jet momentum, or HT. For the signal regions, the likelihood fit is performed for distributions of a GNN output designed to identify $t\bar{t}t\bar{t}$ events (see Chapter 8). The systematics are either explicitly included from independent calculations, such as the luminosity, or else included as their own input distributions. For a given source of systematic error, the input histograms are the same as the nominal histograms but with the yields adjusted by $\pm 1\sigma$ of the systematic error in question.

The profile likelihood fit is first performed using an Asimov dataset as an input. The Asimov dataset is a special dataset such that, when the likelihood fit is performed, it returns the same values for the error parameters θ that were used as input. In other words, the signal and background MC are put into the fit with their nuisance parameters at their nominal values, and the fit assumes that these values are already optimal. The fit does not change the mean values, but it can still constrain the range of each nuisance parameter value, reducing it from its original range of 1σ . The new range is obtained from the range of $\hat{\theta}$ values obtained in the likelihood ratio (Equation 9.8).

The role of the Asimov fit is to test the sensitivity of the analysis without relying on data, which must be blinded. The Asimov fit can also be used to assess the quality of the fit, which not only includes the sensitivity, but also the stability against fluctuations in data, reasonable values for the nuisance parameters, and no problematic nuisance parameter constraints. The definition of a “problematic” constraint is somewhat arbitrary, but generally it is considered a problematic constraint if A) the constraint is on a small, less impactful nuisance parameter, B) the nuisance parameter is *perfectly* constrained (with a range of zero), or C) if there is a large difference in constraint on the same nuisance parameter between two iterations of a fit.

A blinded fit is a fit performed using data, but which has been blinded in regions of phase space anticipated to have the highest signal to background ratio. For this analysis, bins for BDT/GNN scores above a certain value are blinded. The cut on BDT/GNN score is applied to each region separately such that the blinded bins have an MC-modeled signal percentage $> 5\%$.

A realistic Asimov fit is a two-part process. First, a fit to blinded data is performed to extract the constrained/pulled values of the systematics and of the background yields. The results from the blinded fit are used to create “realistic” Asimov data. The output of the realistic Asimov fit gives a more accurate assessment of the analysis sensitivity before unblinding. The realistic Asimov fit is used in the Higgs Oblique analysis, shown below in Section 9.5.2.

9.5 Preliminary Profile Likelihood Fit Results

9.5.1 HBSM Analysis

The Asimov fit is performed using the TRExFitter tag 00-04-16 [116]. The six 1L control regions are fit to the jet activity (H_T) distribution, and the six 1L signal regions are fit to the MVA distribution. There are two MVA distributions used: a BDT and a GNN, discussed in Chapter 8. For each MVA type, two different MVA’s are used: one optimized for the 1L channel, and one optimized for the 2LOS channel. Therefore there are two separate fits, one for each MVA type, and four trained MVA’s used for each mass point. The $t\bar{t}H \rightarrow t\bar{t}t\bar{t}$ process is investigated separately for seven different mass values of H , from 400GeV to 1000GeV in 100GeV increments.

The Asimov fit is performed over all control and signal regions, and with all available samples and systematics. The dominant $t\bar{t}$ background is reweighted using first a data-driven heavy-flavor (HF) scaling, and then reweighting is applied to kinetic variable distributions using a neural network method (See Chapter 7).

The pre-fit yield plots for signal and background distributions are shown in Figures 9.2 and 9.3, and the post-fit yield plots are shown in Figure 9.4 and 9.5. Plots using GNN distributions for a representative mass point of $m_H = 500\text{GeV}$ are shown. As expected, the post-fit distributions have reduced uncertainty from the constraints placed on nuisance parameters during the fit. The post-fit NP constraints are shown in Figure 9.6. For the Asimov fit, the NP values are assumed to be correct, and the values themselves are not adjusted (pulled) by the fit; only their range is constrained. The largest constraints occur for the $t\bar{t}$ modeling systematics and the

cross section systematics. The modeling systematics also have the highest impact on the calculated value of the signal strength, μ , as indicated in the blue bars of Figure 9.7. This large impact is expected, and as such the constraints on the modeling systematics are not considered excessive. However, the constraint on the $t\bar{t} + \geq 1c$ cross section is larger than expected and is currently under investigation.

A preliminary fit to data is performed using the same TRExFitter tag. The setup is the same as for the Asimov fit, except that instead of assuming that the yields and NP values are already correct, they are adjusted to maximize the likelihood ratio with respect to data. In the preliminary data fit, regions expected to have high signal yield, with S/R ratio $> 5\%$, are blinded. This fit is performed assuming that there is no $t\bar{t}t\bar{t}$ contribution, i.e. a background-only fit. The blinded fit provides information on the modeling agreement in background-dominated regions of phase space.

The pre-fit yield plots for signal and background distributions are shown in Figure 9.8 and 9.9, and the post-fit yield plots are shown in Figure 9.10 and 9.11. Plots using the GNN distributions for the representative mass point of $m_H = 700\text{GeV}$ are shown. The post-fit distributions have reduced uncertainty due to data/MC matching in the fit being used to constrain the nuisance parameters (NPs). The post-fit NP constraints are shown in Figure 9.12. For the data fit, the NP values can be both adjusted and constrained by the fit. There are no anomalous large pulls, which would indicate issues with the fit.

9.5.2 Higgs Oblique Analysis

The Higgs Oblique analysis seeks to extract, from the signal strength of $t\bar{t}t\bar{t}$, a value and associated uncertainty for the Higgs Oblique parameter, $\hat{H}$. To accomplish this, an EFT adaptation of the TRExFitter tool is used. Instead of having one signal (the SM $t\bar{t}t\bar{t}$), there are five samples representing the $t\bar{t}t\bar{t}$ signal at different values of $\hat{H}$: $\hat{H} = 0.05, 0.08, 0.12$, and 0.17 . The EFTRExFitter tool fits the $t\bar{t}t\bar{t}$ yield as a function of $\hat{H}^2$, where the quadratic dependence comes from Equation 2.37. The EFTRExFitter tool was used to produce the quadratic fits shown in Figure 9.13. In this case, the fit extracts the coefficients for this quadratic function.

Once the $t\bar{t}t\bar{t}$ yield has been described as a function of $\hat{H}$, the fit to data commences in the same way that it does for the HBSM analysis: there is a profile likelihood fit that constrains systematic uncertainties and estimates a value of $t\bar{t}t\bar{t}$ that maximizes the likelihood. However, instead of extracting a $t\bar{t}t\bar{t}$ signal strength in terms of the SM cross section, the signal strength is expressed as the value of $\hat{H}$

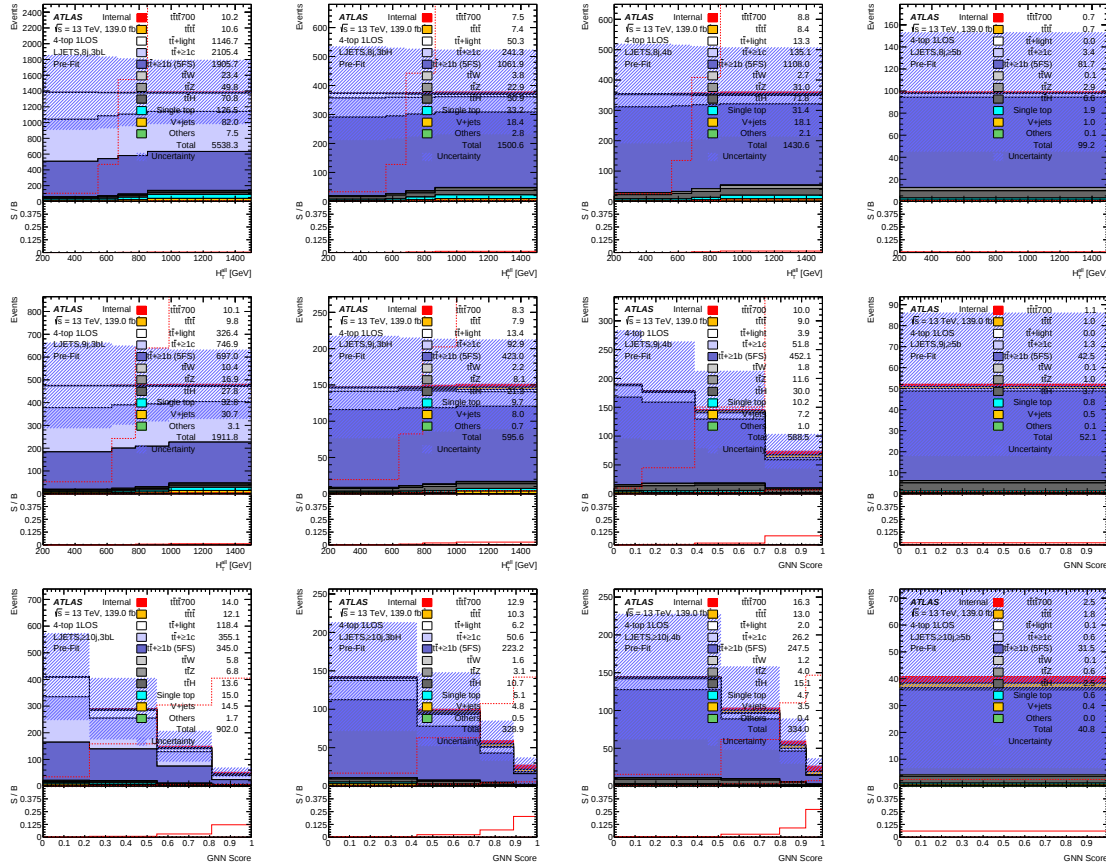


Figure 9.2: Pre-fit distributions for 1L channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot. Some plots, particularly those for high b-jet multiplicity, have a single bin with large uncertainty.

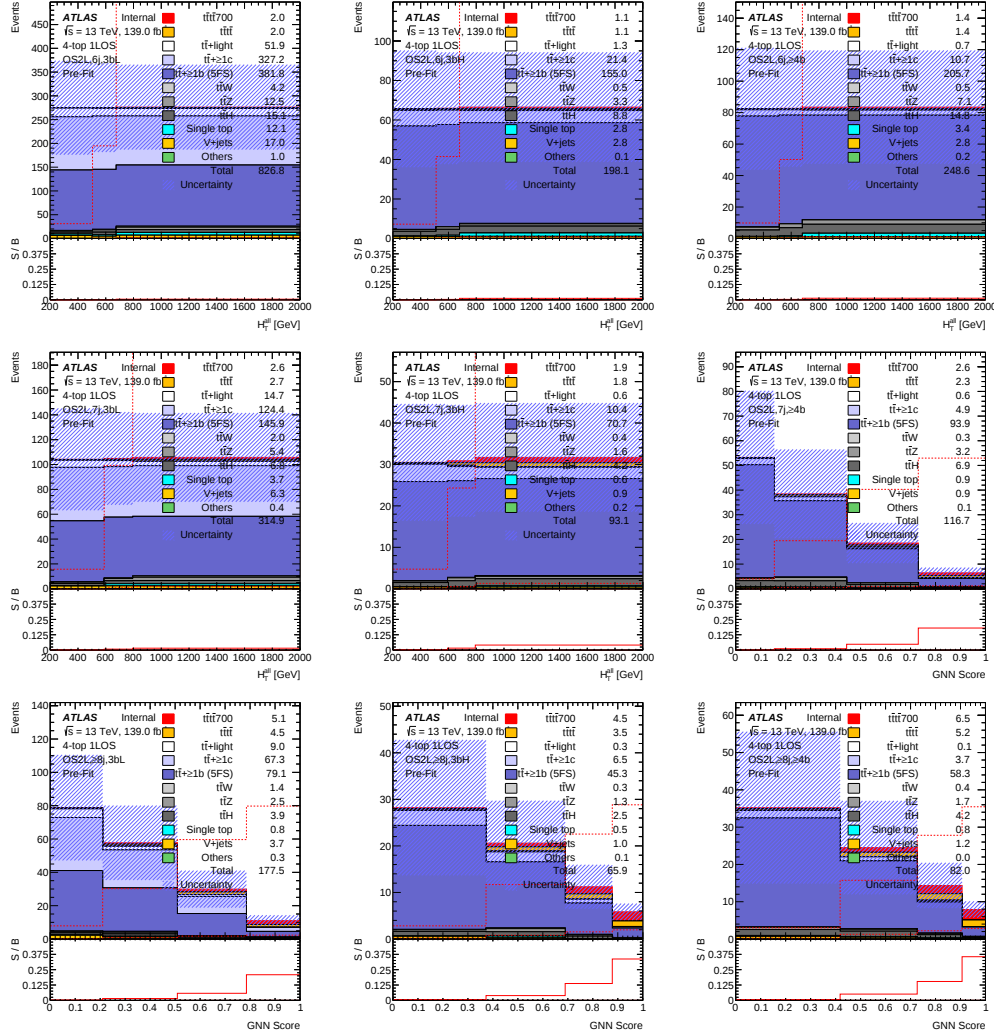


Figure 9.3: Pre-fit distributions for 2LOS channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.

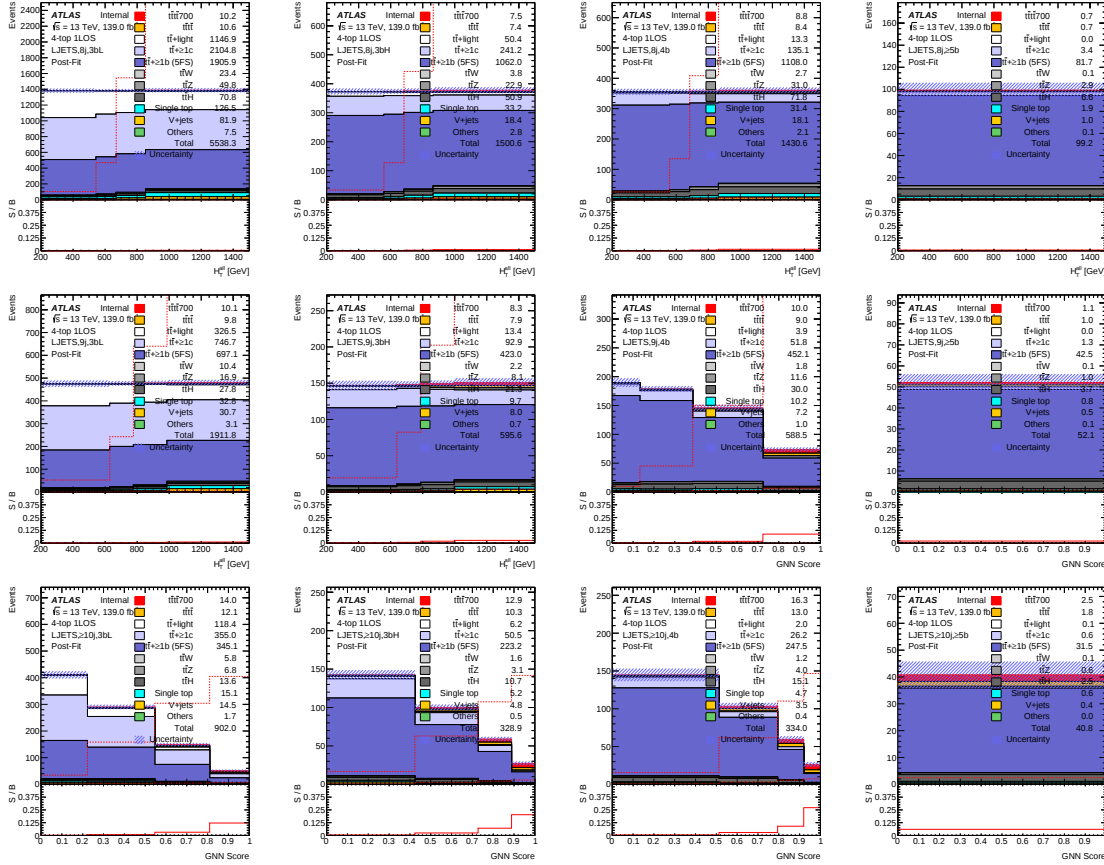


Figure 9.4: Post-fit distributions for 1L channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.

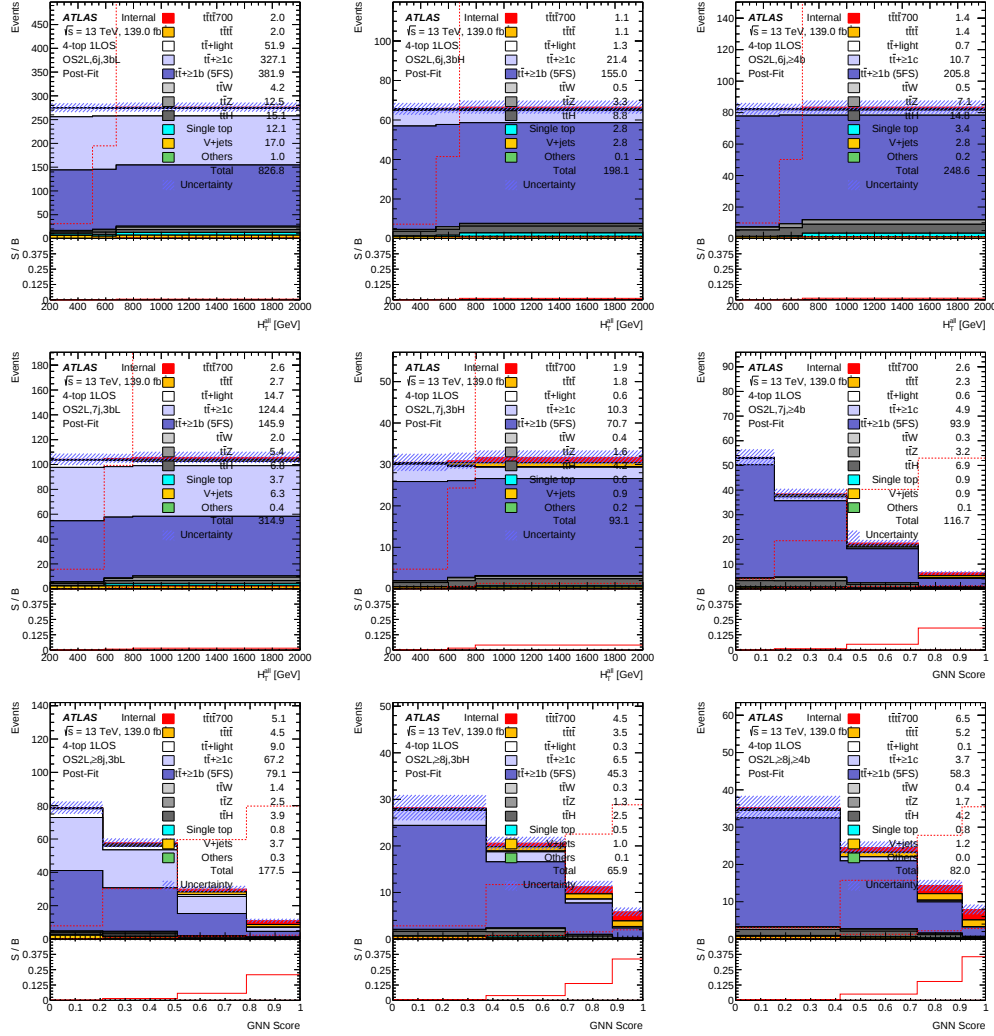


Figure 9.5: Post-fit distributions for 2LOS channel, associated with different signal and background samples. The event yield is on the y-axis. For the control regions, the H_T value is on the x-axis, and for the signal regions, the GNN score is on the x-axis. The signal-to-background ratio is shown in the bottom of the plot.

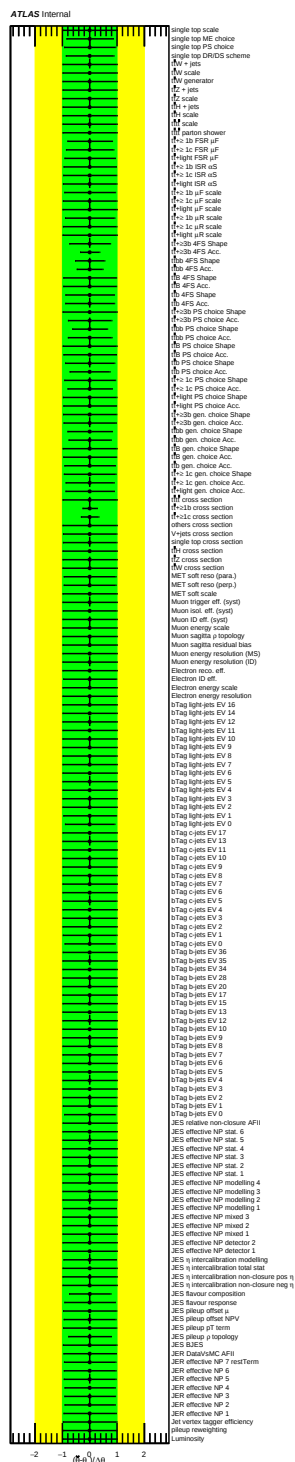


Figure 9.6: Nuisance parameter constraints after fitting to Asimov data.

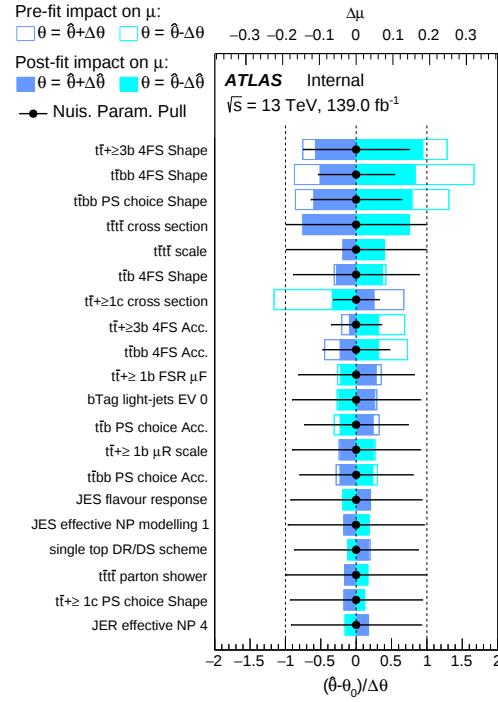


Figure 9.7: Ranking of systematic uncertainties for 1LOS fit to Asimov data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit).

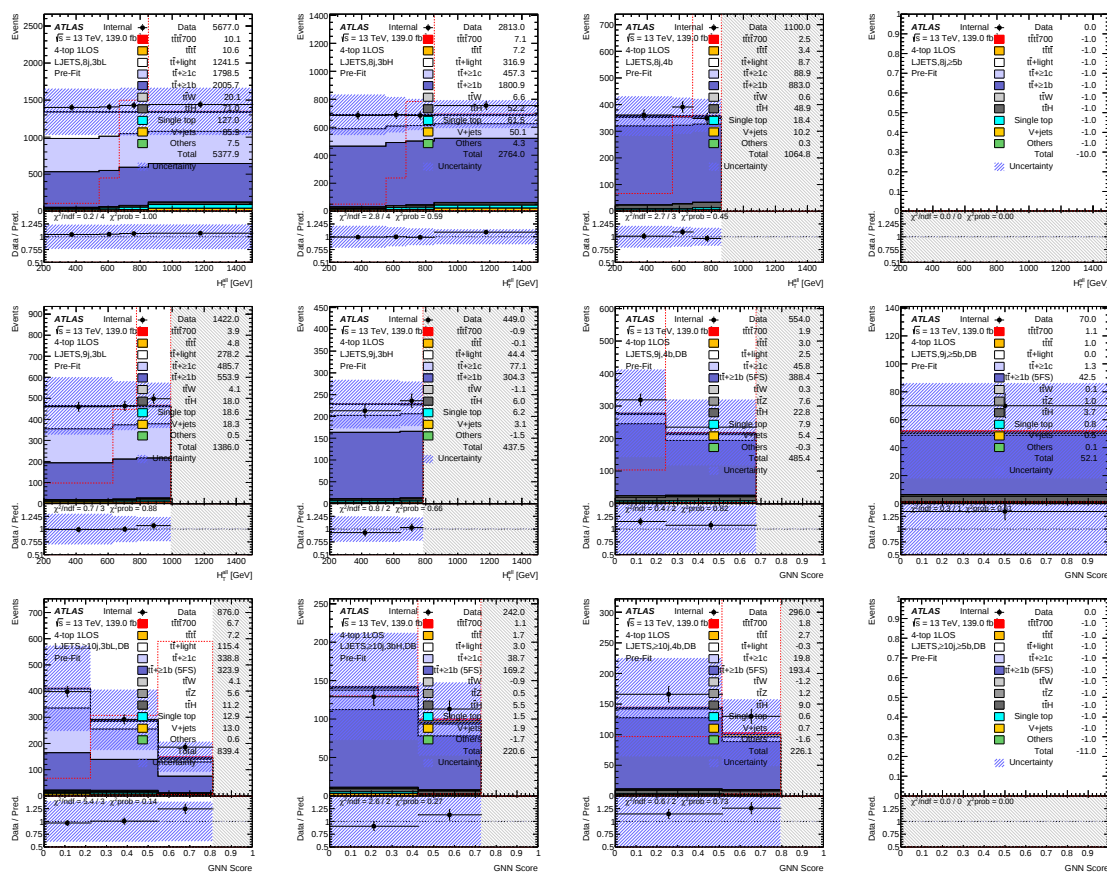


Figure 9.8: Pre-fit distributions for 1L channel, associated with different signal and background samples, compared with blinded data yields.

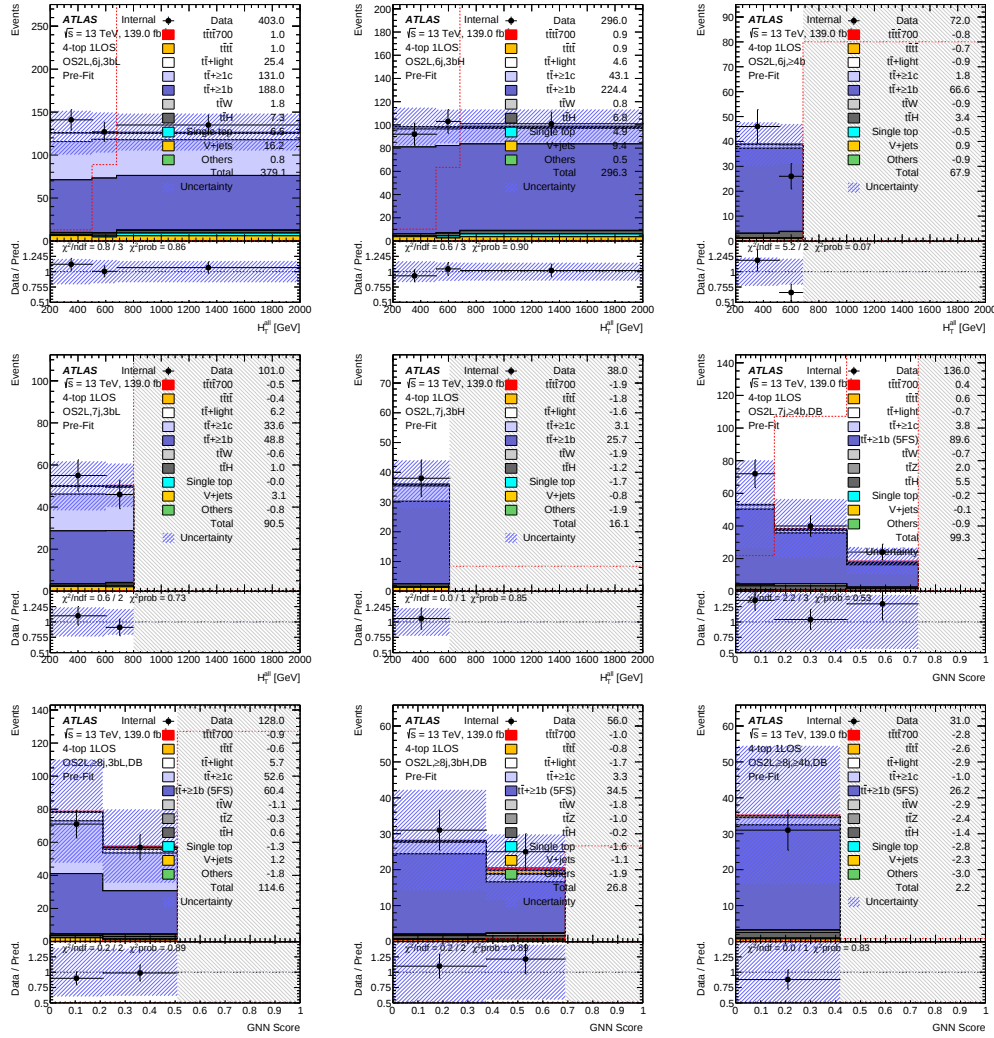


Figure 9.9: Pre-fit distributions for 2LOS channel, associated with different signal and background samples, compared with blinded data yields.

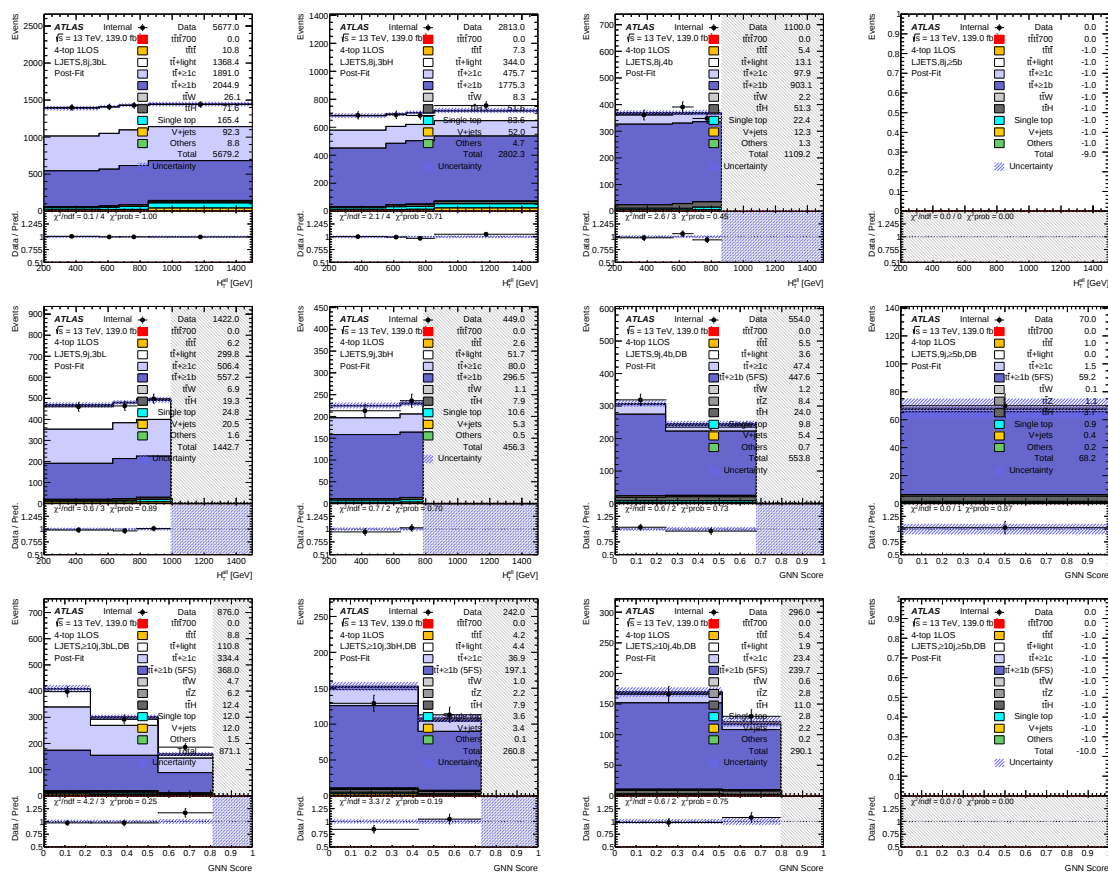


Figure 9.10: Post-fit distributions for 1L channel, associated with different signal and background samples, compared with blinded data yields.

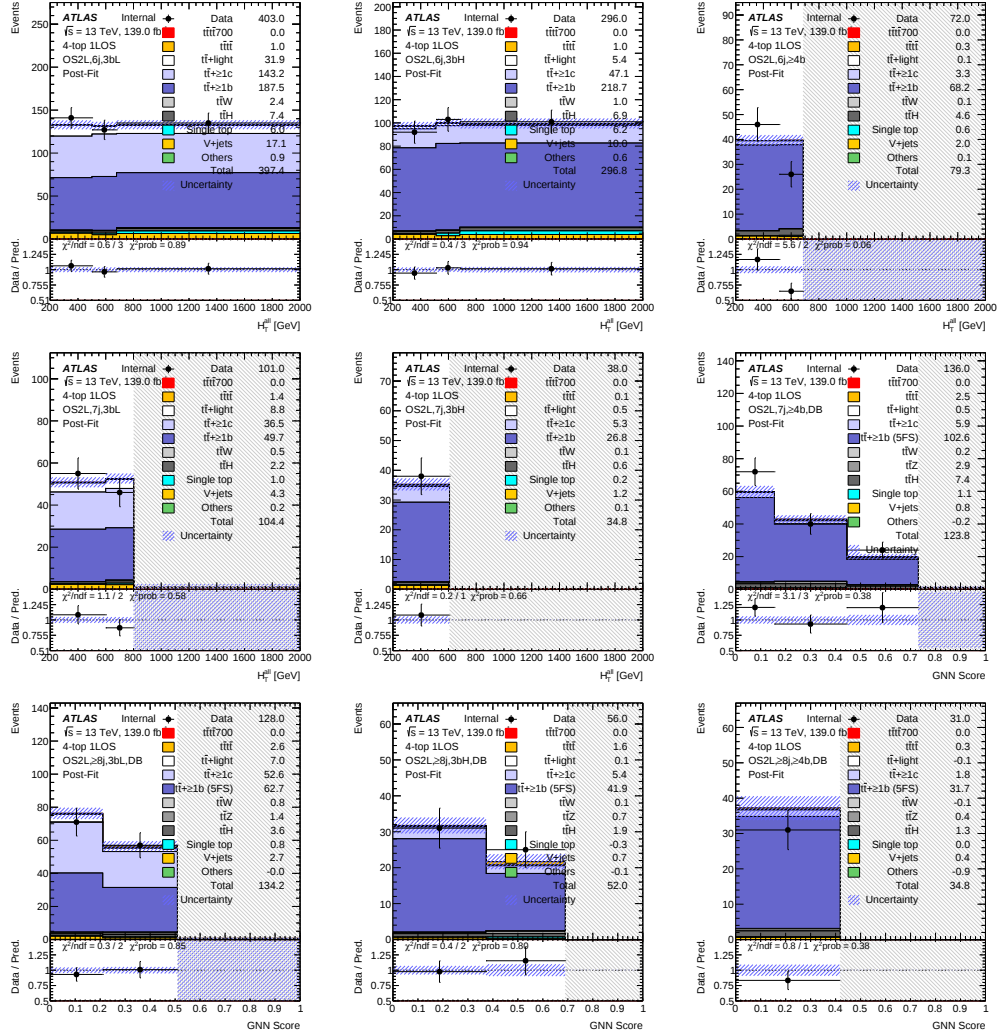


Figure 9.11: Post-fit distributions for 2LOS channel, associated with different signal and background samples, compared with blinded data yields.

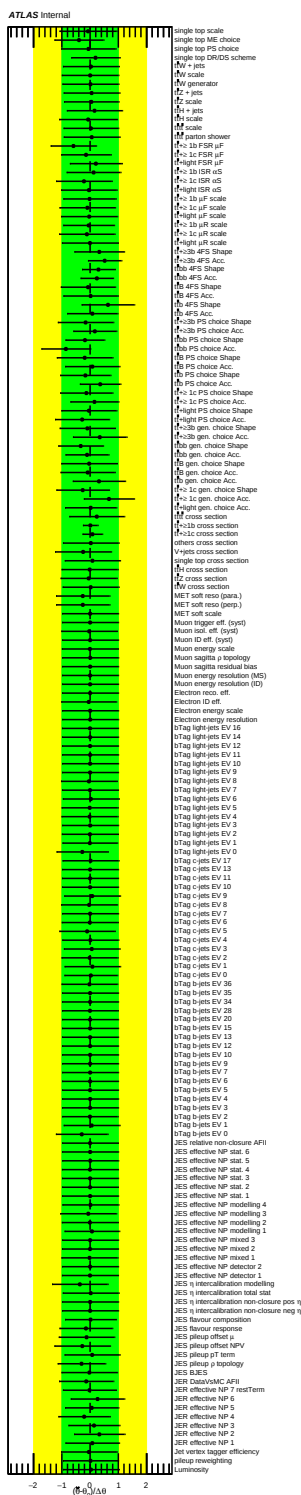


Figure 9.12: Nuisance parameter pulls and constraints after fitting to blinded data.

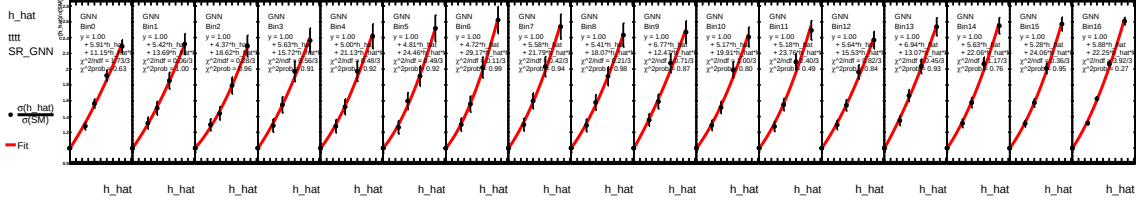


Figure 9.13: Quadratic fit to $t\bar{t}t\bar{t}$ yields as a function of $\hat{H}$ in each of the 17 bins of GNN in the SR.

that minimizes the negative log likelihood ratio. In this case, the ratio, defined in 9.8, is between the unconditional likelihood maximum and the conditional maximum assuming a value of $\hat{H}$. In this case, μ is replaced with the Higgs Oblique parameter.

Also taken into consideration is the treatment of the $t\bar{t}H$ background. Because the $t\bar{t}H$ process also involves a Higgs boson, it is impacted by the additional term in the Higgs propagator. Three versions of $t\bar{t}H$ background were investigated: fixed, where the $t\bar{t}H$ cross section is held at its SM value; fitted, where the $t\bar{t}H$ cross section is given a linear normalization factor; and parameterized, where the $t\bar{t}H$ cross section is a function of $\hat{H}$:

$$\mu_{t\bar{t}H} = (1 - \hat{H})^2 \quad (9.14)$$

and $\mu_{t\bar{t}H}$ is the $t\bar{t}H$ cross section normalization factor. The parameterized treatment of $t\bar{t}H$ has the most physical motivation, and is the treatment used in the final fit discussed in Chapter 10.

The profile likelihood ratio is expressed in the analysis as $-2\Delta\ln(L)$. A value of zero indicates that the chosen value of $\hat{H}$ maximizes the likelihood, and is the value of $\hat{H}$ that best matches the data. Because the SM predicts a value of $\hat{H} = 0$, this is the expected value. Tests of the fit using $\hat{H} = 0$ MC samples as pseudodata demonstrate that the fit is able to correctly extract the expected value, with a 95% CL of roughly 0.12. This can be seen in Figure 9.14, which displays the $-2\Delta\ln(L)$ for the three different treatments of the $t\bar{t}H$ background. At large values of $\hat{H}$, the EFT expansion of the Higgs propagator becomes non-unitary, which would cause unphysical probabilities for scattering mediated by a Higgs boson. The non-unitary regime is shown in Figure 9.14 as a shaded region for $\hat{H} > 0.2$.

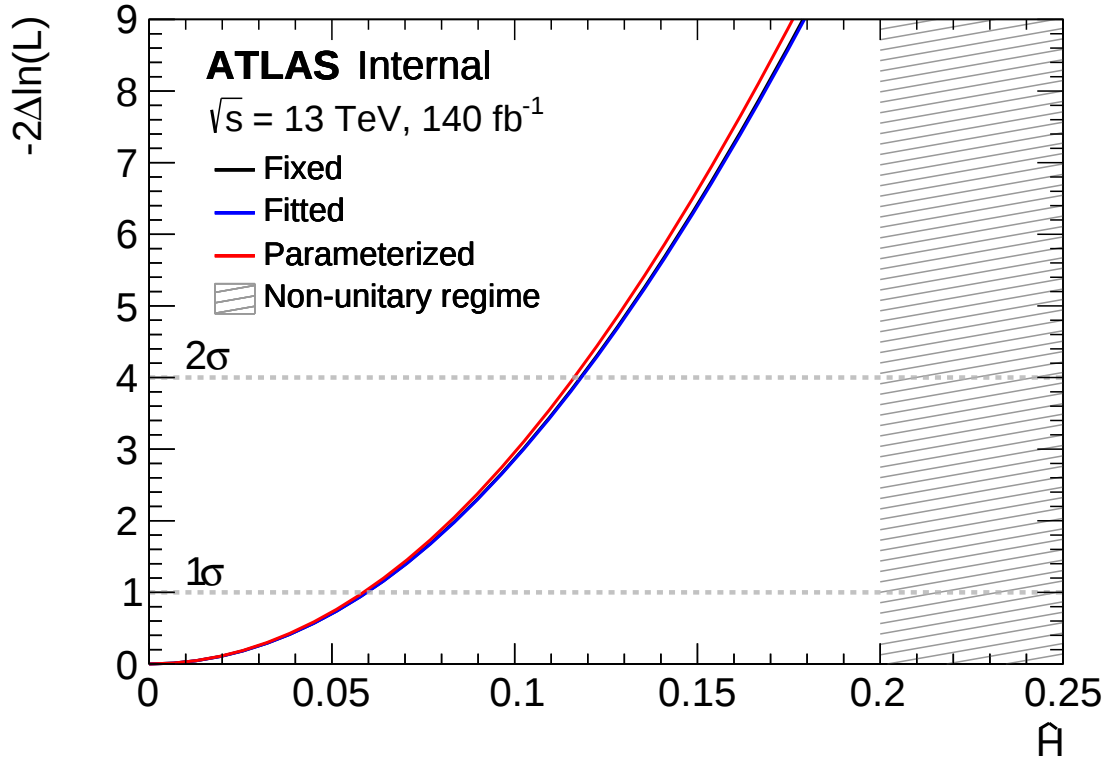


Figure 9.14: Log likelihood plot showing the likelihood at each value of $\hat{H}$ relative to the likelihood maximum, which in this case is set up to be at $\hat{H} = 0$.

Figure 9.15 shows the pre-fit results in the control regions and signal region, whose definitions are discussed in Chapter 8.

Figure 9.15: Post-fit distributions for SSML CR's and SR, associated with different signal and background samples.

CHAPTER 10

Results

10.1 HBSM Analysis

The unblinded results of the HBSM analysis are not currently available. However, there are plots available for the expected limits on m_H , based on the Asimov fit described in Section 9.5.1. Figure 10.1 shows how the individual channels compare to the combined channel limits. The 2LOS channel is seen to be somewhat less sensitive than the 1L channel, but still contributes to the sensitivity of the combined limit. Based on this plot, the analysis is expected to exclude the $pp \rightarrow t\bar{t}H \rightarrow t\bar{t}t\bar{t}$ process associated with Heavy Higgs H of $m_H = 1000\text{GeV}$ with a cross section of $10 - 20\text{fb}$ at the 95% confidence level. For lower m_H , the analysis is less sensitive, excluding the process above a cross section of roughly $40 - 50\text{fb}$.

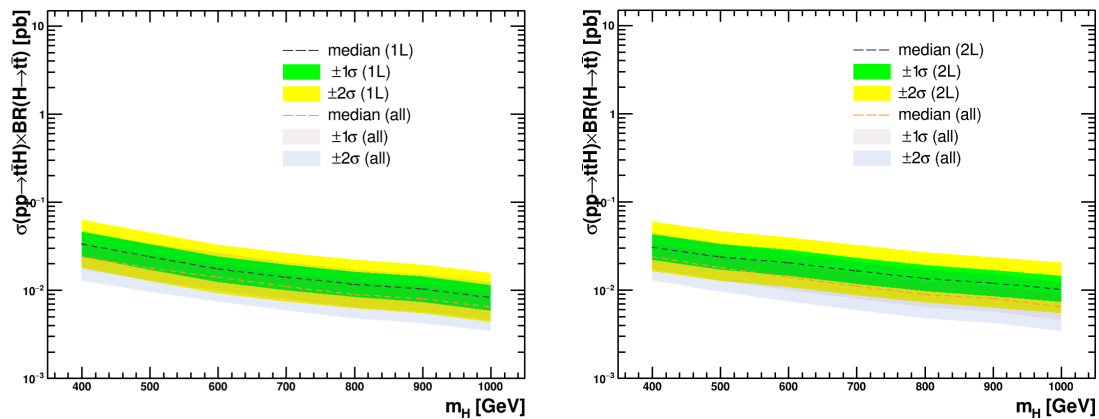


Figure 10.1: Limits of HBSM analysis sensitivity as a function of Heavy Higgs mass, m_H . This limit plot is obtained from a fit to Asimov data. The left plot shows the comparison between using both 1L and 2LOS channels to using 1L alone. The right plot shows the comparison between using both 1L and 2LOS channels to using 2LOS alone. In both cases, the combined channels have greatest sensitivity.

These are only preliminary results fit to Asimov data that assume a BSM cross section of 12fb , based off the previous $t\bar{t}t\bar{t}$ cross section measurement of 24fb

[24] minus the SM $t\bar{t}t\bar{t}$ cross section estimate of 12fb (see Section 4.3.3). To complete the analysis, the ATLAS data will need to be unblinded. The analysis is currently being finalized. There are a few MC samples and uncertainties that need to be added, including the $t\bar{t}t$ sample and its corresponding cross section uncertainty, the $t\bar{t}t\bar{t}$ cross section and generator uncertainties, the $t\bar{t}H$ generator uncertainty, uncertainties associated with $t\bar{t}$ reweighting, and extra uncertainty on flavor-tagging for high- p_T jets.

The uncertainties with the largest impact on the BSM $t\bar{t}t\bar{t}$ cross section are the $t\bar{t} + jets$ modeling uncertainties. The large impact is a combination of the fact that $t\bar{t} + jets$ is poorly modeled at such a high jet multiplicity as well as the fact that $t\bar{t} + jets$ constitutes the majority of the background. Future analyses will need to address this source of systematic error, perhaps with improved modeling and data-driven reweighting.

10.2 Higgs Oblique Analysis

The results of the profile-likelihood fit give a $t\bar{t}t\bar{t}$ signal strength of:

$$\mu = 1.9 \pm 0.4 \text{ (stat.)}_{-0.4}^{+0.7} \text{ (syst.)} = 1.9_{-0.5}^{+0.8}.$$

This corresponds to a measured $t\bar{t}t\bar{t}$ cross section of:

$$\sigma_{t\bar{t}t\bar{t}} = 22.5_{-4.3}^{+4.7} \text{ (stat.)}_{-3.4}^{+4.6} \text{ (syst.)} = 22.5_{-5.5}^{+6.6} \text{fb}.$$

The measured $t\bar{t}t\bar{t}$ signal is within 1.8 standard deviations of the SM prediction for the $t\bar{t}t\bar{t}$ cross section of $11.97_{-21\%}^{+18\%} \text{fb}$ [4].

The pre-fit and post-fit distributions for all control regions and the signal region are shown in Figures 10.5 and 10.6. The post-fit distributions show good agreement between MC and data. The post-fit distribution for the signal region is also displayed individually in Figure 10.2. There are several normalization factors associated with the yields of background samples. These values are shown in Figure 10.7.

The correlation matrix, shown in Figure 10.8, shows reasonable correlations between the systematic uncertainties. The largest correlations are between the signal strength $\mu_{t\bar{t}t\bar{t}}$ and theoretical SM $t\bar{t}t\bar{t}$ cross section, between $\mu_{t\bar{t}t\bar{t}}$ and the $t\bar{t}t\bar{t}$ generator choice, between the $t\bar{t}W$ modeling parameters a_0 and a_1 , and between the normalization factors for $t\bar{t}W^+$ and $t\bar{t}W^-$ for $N_{jets} = 4$. Generally, it makes sense

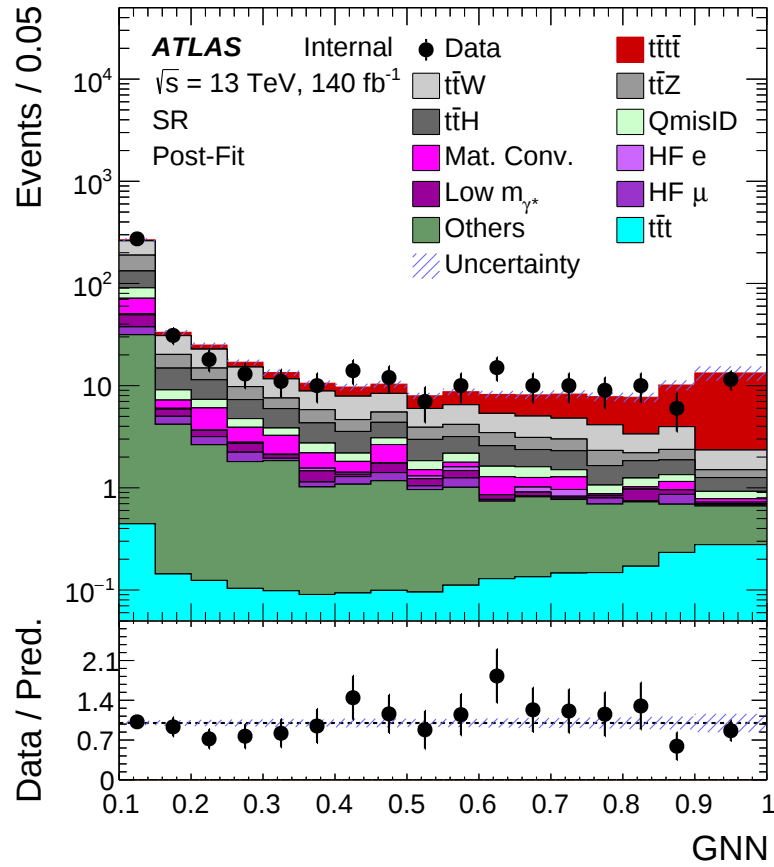


Figure 10.2: Post-fit distribution of the SSML signal region. There is a clear discrepancy between the data yield and the background-only yield in the bins with highest GNN score.

for these parameters to be correlated.

The constraints and pulls on the uncertainties are shown in Figure 10.9 and do not show any large pulls/constraints indicating issues with the fit.

The ranking plot shown in Figure 10.10 lists the total uncertainty, and then the systematic uncertainties that most impact the measurement. The bottom axis measures how much the nuisance parameter (NP) has been pulled from its original value θ_0 to its post-fit value $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The NP pulls are represented by black dots, with the post-fit uncertainty $\Delta\hat{\theta}$ as the error bars. Many of the normalization factors are poorly estimated in pre-fit, so NPs such as $NF_{t\bar{t}W^+,4j}$ have a large value $(\hat{\theta} - \theta_0)/\Delta\theta > 1$. NPs such as a_0 with small error bars are highly constrained by the fit.

The upper axis corresponds to the impact on the value of the signal strength, $\Delta\mu$. The blue bars express the impact of each NP, where the extent of the pre-fit impact, $\hat{\theta} \pm \Delta\theta$ is shown in the unfilled bars and the extent of the post-fit impact is shown in the filled bars. The NPs are ranked according to their impact on μ , with the largest $\Delta\mu$ indicating the NP whose variations most affect μ .

The MC modeling of the $t\bar{t}t\bar{t}$ contributes the most uncertainty, with the generator choice being the most impactful. The extraction of the a_0 parameter in the $t\bar{t}W$ data-driven reweighting is also impactful. The jet flavor tagging parameters Light0 and B0 are the most impactful experimental uncertainties. The C0 parameter is also in the ranking, somewhat lower.

The treatment of the three top ($t\bar{t}t/t\bar{t}\bar{t}$) background is considered. The $t\bar{t}t$ process has many kinematic similarities to the $t\bar{t}t\bar{t}$ signal and is difficult to distinguish when using the GNN classifier. Figure 10.3 shows a 2-dimensional log likelihood plot where the $t\bar{t}t$ and $t\bar{t}t\bar{t}$ cross sections are both floating fit parameters. It is not apparent why the best fit occurs at such a large value of $\sigma_{t\bar{t}t}$ and a small value of $\sigma_{t\bar{t}t\bar{t}}$, but the best fit is still within 2.1 standard deviations of the SM prediction. The anti-correlation between $\sigma_{t\bar{t}t}$ and $\sigma_{t\bar{t}t\bar{t}}$ is clear. The SM cross section for $t\bar{t}t$ is calculated at NLO in QCD using MadGraph to be 1.67 fb. For standalone measurement of the $t\bar{t}t\bar{t}$ cross section, the $t\bar{t}t$ background cross section is fixed to this SM value.

As seen in Figure 10.4, the measured $t\bar{t}t\bar{t}$ signal has an observed significance of 6.1 standard deviations above the background-only hypothesis. The expected sensitivity of the analysis is 4.3 standard deviations above the background-only hypothesis. The observed significance above five standard deviations represents the

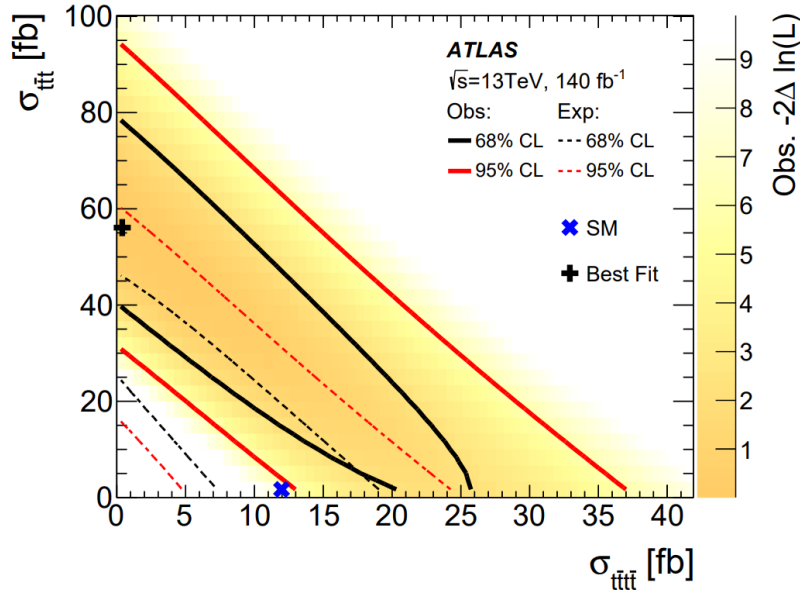


Figure 10.3: Negative log likelihood as a function of $t\bar{t}t\bar{t}$ (x-axis) and $t\bar{t}t$ (y-axis). The darkest color indicates the smallest negative log likelihood value, and thus the best fit.

first discovery of the $t\bar{t}t\bar{t}$ process. Concurrent with the ATLAS discovery, CMS announced a measurement of the $t\bar{t}t\bar{t}$ cross section with a significance of 5.5 standard deviations [134].

A fit is performed to extract the Higgs Oblique parameter $\hat{H}$ from the unblinded data. The $t\bar{t}H$ background is parameterized as a function of $\hat{H}$, and the $\hat{H}$ value is obtained as the minimum of a log likelihood scan as described in Section 9.5.2. The likelihood scan is shown in Figure 10.11, with the solid line representing the observed scan and the dotted line representing the expected scan for the case of $\hat{H} = 0$. The observed value is $\hat{H} = 0.09$. The observed (expected) 95% CL is $\hat{H} = 0.2$ (0.12). This limit is at the boundaries of the model validity, which for $\hat{H} > 0.2$ violates unitarity [9]. A recent measurement from CMS placed the 95% CL upper limit at $\hat{H} = 0.12$ [135].

The correlation matrix in Figure 10.12 shows large negative correlation between the $\hat{H}$ value and the SM $t\bar{t}t\bar{t}$ cross section. This is expected, given that if the SM $t\bar{t}t\bar{t}$ cross section is larger, the Higgs Oblique contribution to the measured $t\bar{t}t\bar{t}$ cross section must be correspondingly smaller. Figure 10.13 shows the uncertainties

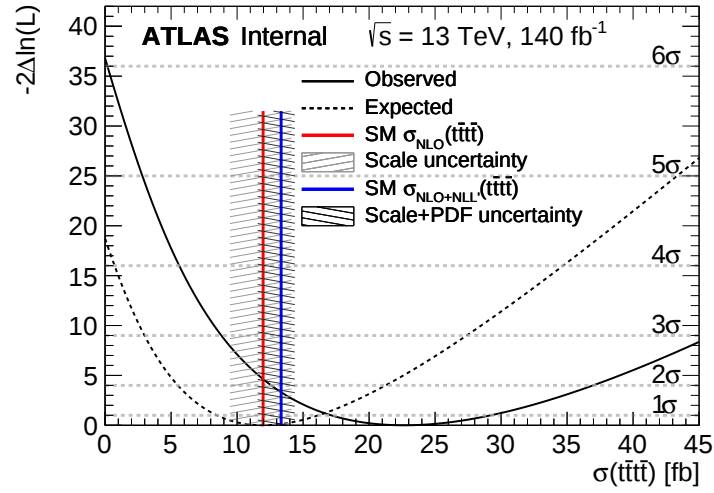


Figure 10.4: Log likelihood scan associated with the SM $t\bar{t}t\bar{t}$ measurement. The observed scan has a minimum log likelihood at $\sigma_{t\bar{t}t\bar{t}} = 22.5\text{fb}$, where the background hypothesis is 6.1 standard deviations away.

with the largest impact on the fit. The $t\bar{t}W$ background fitting parameters a_0 and a_1 contribute the most uncertainty to the fit.

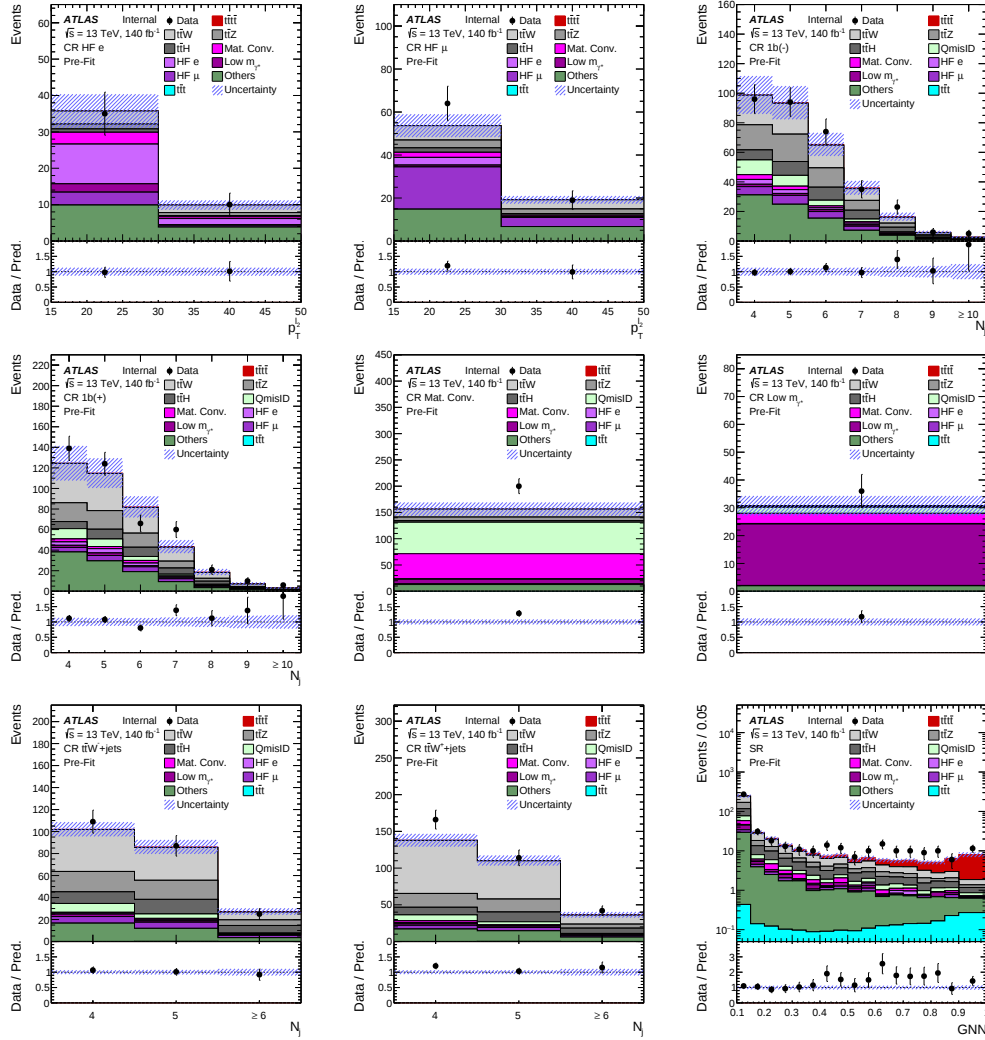


Figure 10.5: Pre-fit distributions for SSML CR's and SR, associated with different signal and background samples.

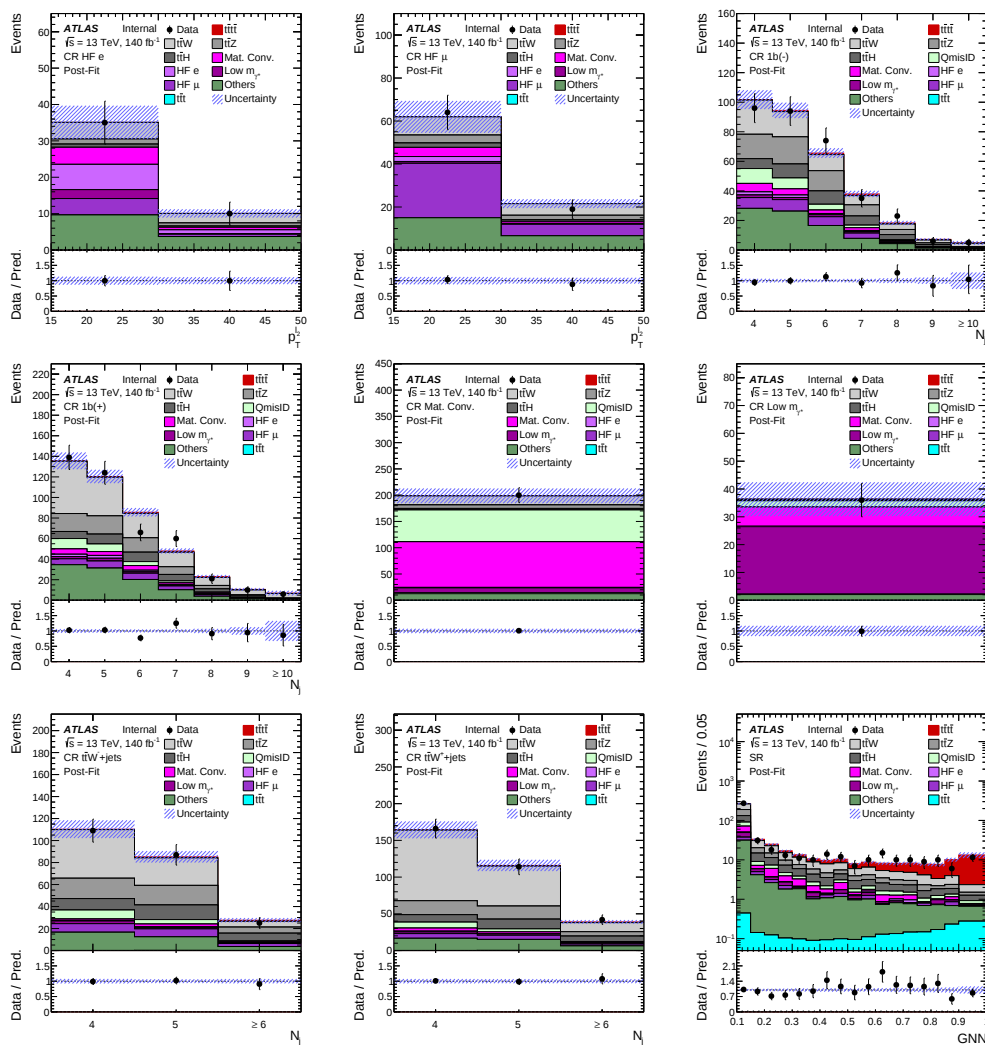


Figure 10.6: Post-fit distributions for SSML CR's and SR, associated with different signal and background samples.

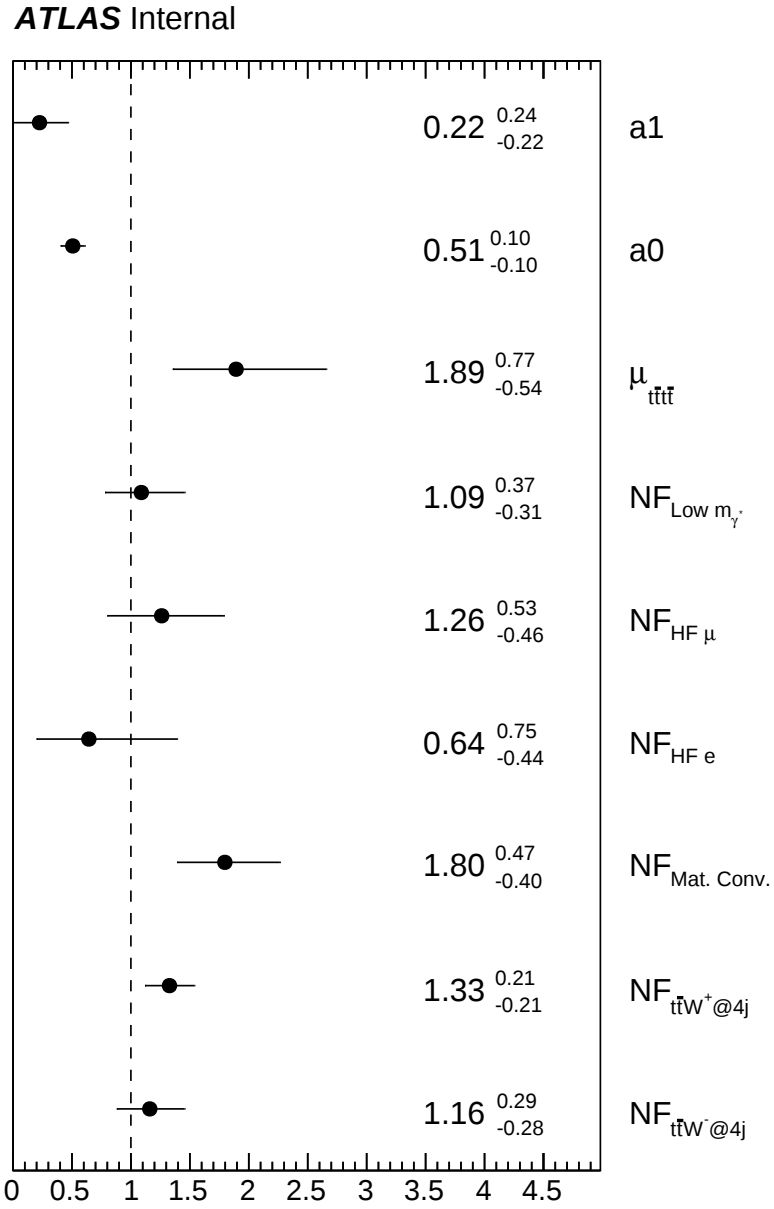


Figure 10.7: Normalization factors for background yields for SSML fit to unblinded data.

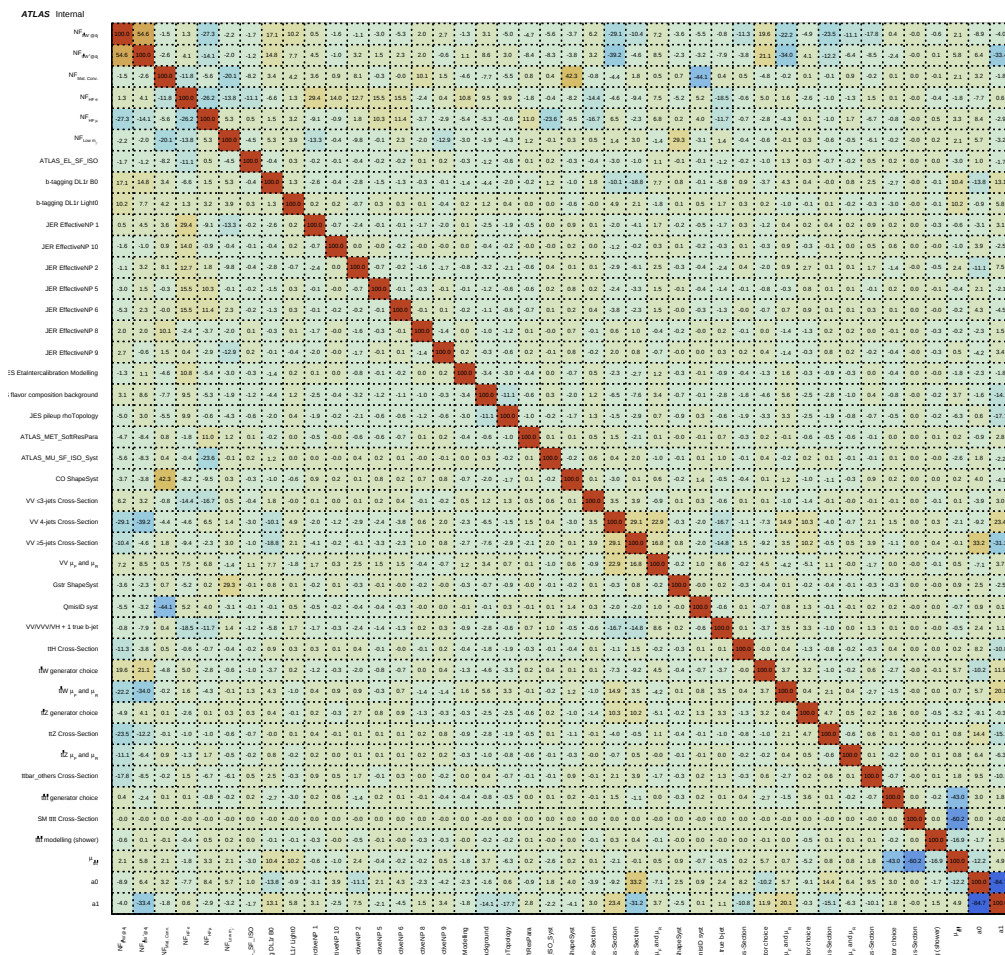


Figure 10.8: Correlation matrix for SSML fit to unblinded data.

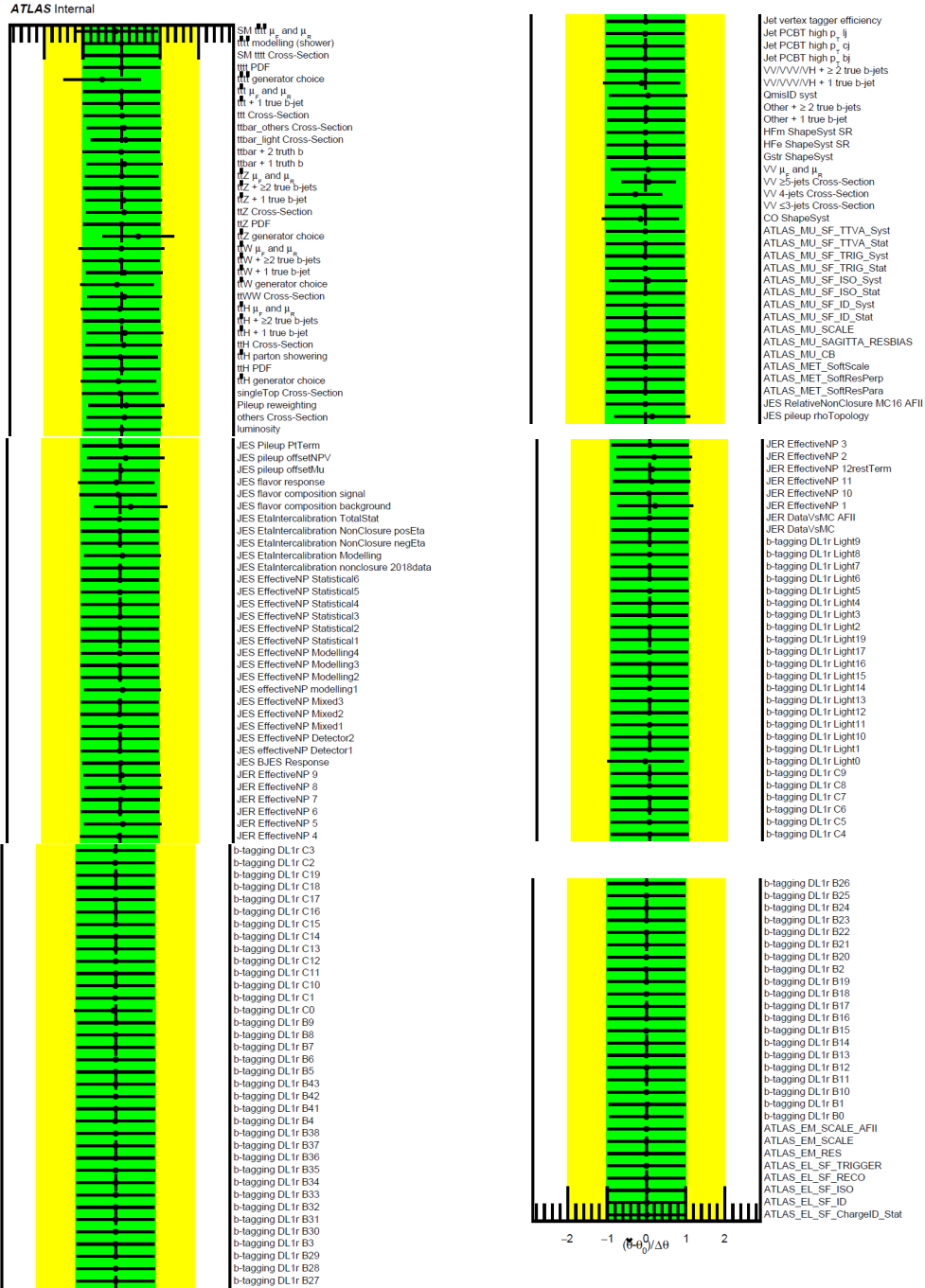


Figure 10.9: Nuisance parameter pulls and constraints for SSML fit to unblinded data. The plot has been divided into six parts to enlarge text.

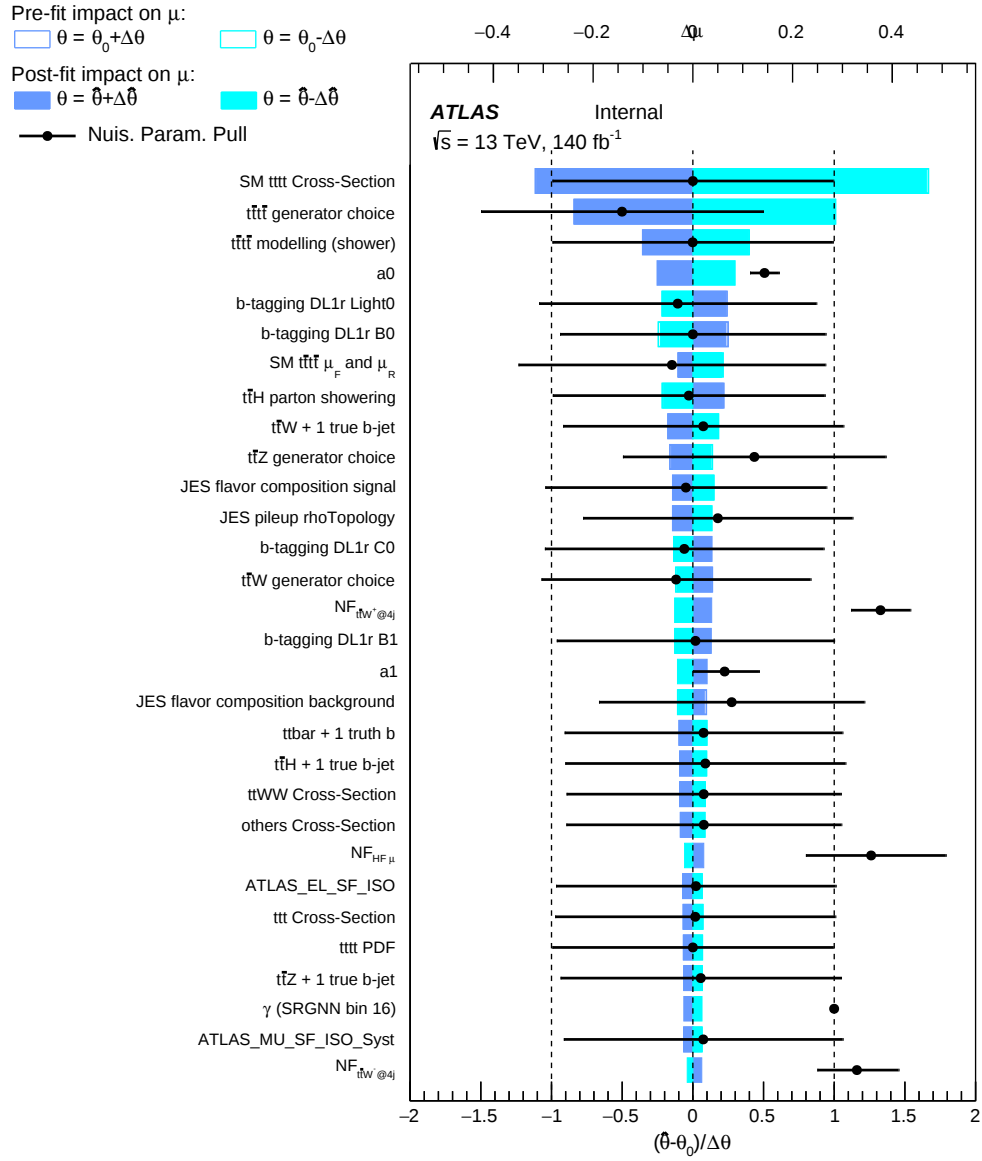


Figure 10.10: Ranking of systematic uncertainties for SSML fit to unblinded data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit).

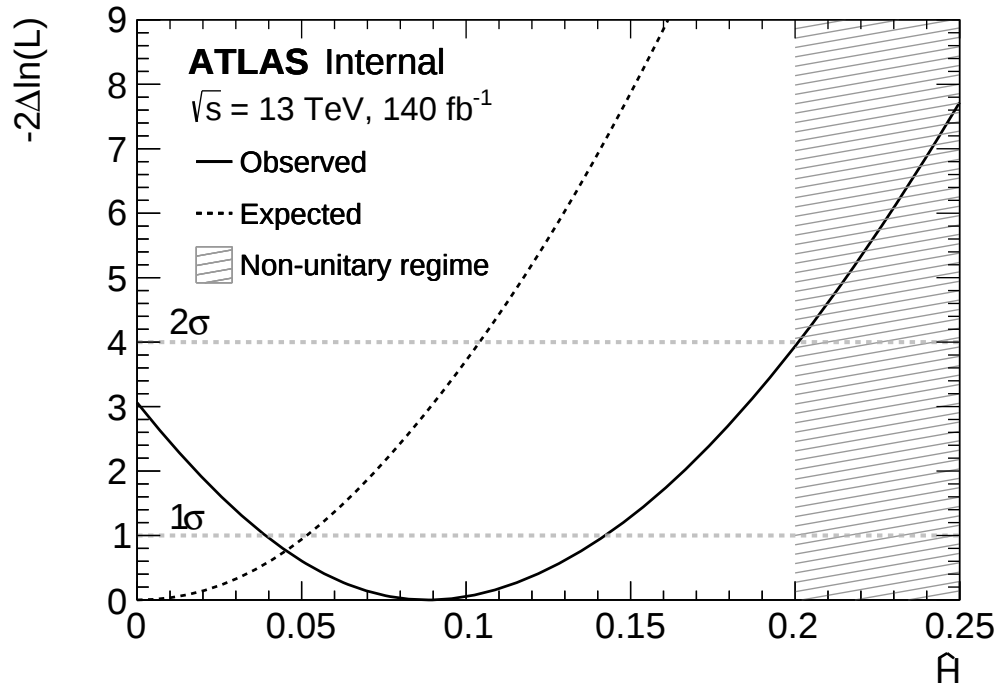


Figure 10.11: Log likelihood scans as a function of $\hat{H}$ value. The dotted line represents the SM hypothesis of $\hat{H} = 0$ and the solid line represents the measured log likelihood scan.

Figure 10.12: Correlation matrix for Higgs Oblique fit to unblinded data.

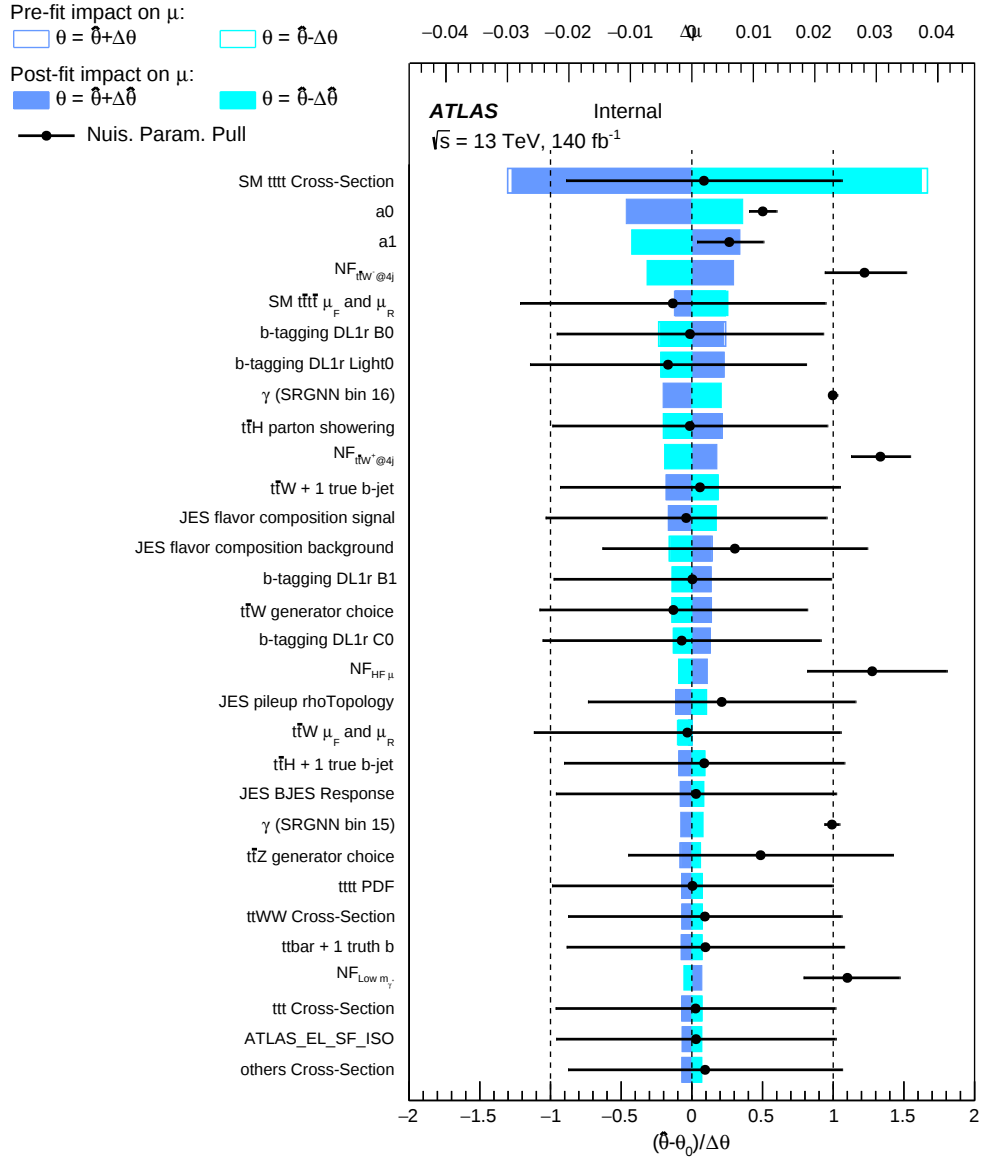


Figure 10.13: Ranking of systematic uncertainties for Higgs Oblique fit to unblinded data. The bottom axis and black dots refer to the difference between the nuisance parameter pre-fit value θ_0 and the post-fit unconditional likelihood maximum $\hat{\theta}$, scaled by the pre-fit uncertainty $\Delta\theta$. The top axis and blue bars refer to the nuisance parameter's impact on the value of μ . The edge of each bar represents the value of $\Delta\mu$ when the parameter is varied by its uncertainty, $\Delta\theta$ (for pre-fit) or $\Delta\hat{\theta}$ (for post-fit).

CHAPTER 11

Conclusion

Particle physics studies the properties of the fundamental constituents of matter. The current view is summarized in the Standard Model, which uses field theory to describe the interactions of fermions via the exchange of force-mediating bosons. The photon mediates the electromagnetic force between all particles with electromagnetic charge, the W and Z bosons mediate the weak force between left-handed isospin doublets, and the gluon mediates the strong force, which affects particles with color charge. Interactions with the Higgs boson give fermions and the W and Z boson their masses. After the discovery of the Higgs boson in 2012, all Standard Model particles have been discovered.

However, the Standard Model is known to present an incomplete picture of the world. There are outstanding questions regarding the nature of dark matter observed in cosmological studies, baryogenesis, and the small mass of the Higgs. The Standard model does not provide an explanation for why particle masses are at their particular value, how neutrinos have mass, or how gravity fits into the fundamental forces of nature. There are many theoretical extensions grouped as Beyond the Standard Model that can address these issues, and experiments can verify which of these are viable and which must be rejected.

The ATLAS detector at the CERN Large Hadron Collider is one such experiment. This multi-purpose detector measures particle tracks and energy depositions that result from proton-proton collisions at $\sqrt{s} = 13\text{TeV}$. The Run II campaign of ATLAS lasted from 2015-2018 and produced $140 \pm 1.2\text{fb}^{-1}$ of data [49].

The $t\bar{t}t\bar{t}$ process is a rare process with a SM predicted cross section at NLO of $11.97^{+18\%}_{-21\%}\text{fb}$ at collision energies of $\sqrt{s} = 13\text{TeV}$ [4]. The small cross section and the fact that the top quark has a large coupling to the Higgs boson means that $t\bar{t}t\bar{t}$ production is particularly sensitive to BSM phenomena affecting Higgs interactions. This thesis presents two measurements of the $t\bar{t}t\bar{t}$ cross section: one in the context of a two Higgs Doublet Model (HBSM), and the other in the context of a Higgs Oblique parameter with BSM contributions. The HBSM analysis uses the single

lepton and opposite-sign dilepton decay channels, while the Higgs Oblique analysis uses the same-sign dilepton and multi-lepton decay channels.

Each top quark in the $t\bar{t}t\bar{t}$ process decays into a b-quark and a W boson. Thus $t\bar{t}t\bar{t}$ events are characterized by high jet and b-jet multiplicity. Both analyses leverage this fact to define regions that isolate signal from various SM backgrounds. There are also control regions to help constrain the modeling of important backgrounds, and validation regions that serve to further check background modeling constraints. Some of the background processes have known issues with MC modeling, and so both analyses make use of data-driven reweighting to correct these samples.

To further isolate signal events from background, a multivariate classifier is used. The HBSM analysis uses a graph neural network (GNN) to discriminate $t\bar{t}t\bar{t}$ from its primary background, $t\bar{t} + jets$. A boosted decision tree (BDT) is also used to cross-check the results from the GNN. The Higgs Oblique analysis uses a GNN to separate $t\bar{t}t\bar{t}$ events from all background events.

Profile likelihood fitting is performed to assess the statistical significance of the $t\bar{t}t\bar{t}$ signal. This fitting is performed using TRExFitter, a statistical analysis package with a RooFit framework. The results are expressed as a $t\bar{t}t\bar{t}$ cross section measurement with signal strength μ , where $\mu = 1$ corresponds to the SM cross section. For the HBSM analysis, limits are placed on the sensitivity to Heavy Higgs detection as a function of the Heavy Higgs mass. For the Higgs Oblique analysis, an $\hat{H}$ value is extracted and a 95% CL upper limit is placed.

Although there are no current unblinded results for the HBSM analysis, the limits obtained from the Asimov fit show the sensitivity of the analysis assuming a BSM $t\bar{t}t\bar{t}$ cross section of 12fb. For the highest value of the Heavy Higgs mass, $m_H = 1000\text{GeV}$, the analysis is expected to exclude a BSM $t\bar{t}t\bar{t}$ cross section of 10 – 20fb, which is close to the expected value of $\sigma_{BSMt\bar{t}t\bar{t}} = 12\text{fb}$.

The Higgs Oblique analysis measures a $t\bar{t}t\bar{t}$ cross section of $\sigma_{t\bar{t}t\bar{t}} = 22.5^{+4.7}_{-4.3}$ (stat.) $^{+4.6}_{-3.4}$ (syst.) = $22.5^{+6.6}_{-5.5}\text{fb}$. This corresponds to a signal strength of $\mu = 1.9 \pm 0.4$ (stat.) $^{+0.7}_{-0.4}$ (syst.) = $1.9^{+0.8}_{-0.5}$. The observed (expected) significance is 6.1 (4.3) standard deviations above the background-only hypothesis. The extracted $\hat{H}$ value is $\hat{H} = 0.09$, and the upper limit at a 95% CL is $\hat{H} = 0.2$, at the boundary of the Higgs Oblique model's validity.

The discovery of the $t\bar{t}t\bar{t}$ process at a significance greater than 5σ is an exciting development in the study of this rare process. The measured cross section

is within 1.8 standard deviations of the expected SM value, and more data is needed to confirm whether the large measured cross section is truly in agreement with the SM or not. The $t\bar{t}t\bar{t}$ process remains an intriguing probe for many BSM studies.

Future studies can improve upon both analyses. The HBSM analysis is limited by $t\bar{t} + jets$ modeling uncertainties, while the Higgs Oblique analysis is affected by modeling uncertainties for $t\bar{t}t\bar{t}$ and $t\bar{t}W$ samples. The accuracy of these MC modeling tools may increase with greater computational sophistication as well as better measurements of SM processes. The statistical uncertainty of the $t\bar{t}t\bar{t}$ cross section measurement will decrease with more data. To that end, the ATLAS detector is already collecting data for its Run III campaign, which will run from 2022-2025 at an unprecedented collision energy of $\sqrt{s} = 13.6\text{TeV}$. Run III plans to deliver 450fb^{-1} of collision data [136]. The following runs will be part of the High Luminosity LHC program, which will run at $5 - 7.5\times$ the LHC nominal luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Runs IV, V, and beyond are anticipated to deliver as much as 4000fb^{-1} of collision data.

This thesis describes the first firm establishment of the SM $t\bar{t}t\bar{t}$ process. While the existence of this process is not surprising, the measured cross section is roughly twice as large as the SM prediction, an exciting result. There are several BSM interpretations for which the $t\bar{t}t\bar{t}$ process serves as a useful probe, including the two interpretations detailed in this thesis. Hopefully Runs III and IV will reveal whether or not there is new physics involved in the $t\bar{t}t\bar{t}$ process.

Bibliography

- ¹ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Physics Letters B* **716**, 1–29 (2012).
- ²CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Physics Letters B* **716**, 30–61 (2012).
- ³R. Massey, T. Kitching, and J. Richard, “The dark matter of gravitational lensing”, *Reports on Progress in Physics* **73**, 086901 (2010).
- ⁴R. Frederix, D. Pagani, and M. Zaro, “Large NLO corrections in $t\bar{t}W^\pm$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions”, *Journal of High Energy Physics* **2018**, 10.1007/jhep02(2018)031 (2018).
- ⁵R. L. Workman et al. (Particle Data Group), “Review of Particle Physics”, *PTEP* **2022**, 083C01 (2022).
- ⁶ATLAS Collaboration, “Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Journal of High Energy Physics* **11** (2021).
- ⁷CMS Collaboration, *Evidence for four-top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, 2023.
- ⁸G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, and J. P. Silva, “Theory and phenomenology of two-Higgs-doublet models”, *Physics Reports* **516**, Theory and phenomenology of two-Higgs-doublet models, 1–102 (2012).
- ⁹C. Englert, G. F. Giudice, A. Greljo, and M. McCullough, “The $\hat{H}$ -parameter: an oblique Higgs view”, *Journal of High Energy Physics* **2019**, 10.1007/jhep09(2019)041 (2019).
- ¹⁰D. K. Paredes Hernandez, N. Orlando, M.-J. Tsai, Z. Zheng, S. Huang, Y. Tu, S. Berlendis, C. J. Birch-Sykes, K. Johns, A. Juste Rozas, R. E. Lindley, R. Peters, J. Qian, A. Sonay, K. C. Tam, C. Chauhan, and P. Berta, *Search for $t\bar{t}H \rightarrow t\bar{t}t\bar{t}$ production in the multilepton final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep. (CERN, Geneva, 2022).
- ¹¹M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Advanced book classics, Avalon Publishing, 1995).

- ¹²P. Higgs, “Broken symmetries and the masses of gauge bosons”, *Physical Review Letters* **13** (1964).
- ¹³F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector bosons”, *Physical Review Letters* **13** (1964).
- ¹⁴T. K. G.S Guralnik C.R. Hagen, “Global conservation laws and massless particles”, *Physical Review Letters* **13** (1964).
- ¹⁵B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton fragmentation and string dynamics”, *Physics Reports* **97**, 31–145 (1983).
- ¹⁶R. Gupta, *Introduction to lattice qcd*, 1998.
- ¹⁷M. Czakon, P. Fiedler, and A. Mitov, “The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_s^4)$ ”, *Physical Review Letters* **110**, 10.1103/physrevlett.110.252004 (2013).
- ¹⁸ATLAS Collaboration, “Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment”, *Physics Letters B* **810**, 135797 (2020).
- ¹⁹”*LHCTopWG Summary Plots*”.
- ²⁰E. L. Berger, J. Gao, and H. X. Zhu, “Differential distributions for t-channel single top-quark production and decay at next-to-next-to-leading order in QCD”, *Journal of High Energy Physics* **2017**, 10.1007/jhep11(2017)158 (2017).
- ²¹N. Kidonakis, “Next-to-next-to-leading logarithm resummation for s-channel single top quark production”, *Physical Review D* **81**, 10.1103/physrevd.81.054028 (2010).
- ²²N. Kidonakis, “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- ”, *Physical Review D* **82**, 10.1103/physrevd.82.054018 (2010).
- ²³S. Mergelmeyer, “P o s (i c h e p 2 0 1 6) 6 6 4 single top-quark production cross section using the atlas detector at the lhc”, in (2016).
- ²⁴ATLAS Collaboration, “Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *European Journal of Physics C* **80** (2020).
- ²⁵ATLAS Collaboration, *Observation of four-top-quark production in the multilepton final state with the ATLAS detector*, 2023.
- ²⁶G. Gabrielse, S. E. Fayer, T. G. Myers, and X. Fan, *Towards an Improved Test of the Standard Model’s Most Precise Prediction*, 2019.

- ²⁷G. F. Giudice, “Naturally Speaking: The Naturalness Criterion and Physics at the LHC”, in *Perspectives on LHC physics* (WORLD SCIENTIFIC, Jan. 2008), pp. 155–178.
- ²⁸C. Csáki and P. Tanedo, *Beyond the Standard Model*, en, 2015.
- ²⁹L. Bergström, “Dark matter evidence, particle physics candidates and detection methods”, *Annalen der Physik* **524**, 479–496 (2012).
- ³⁰H. Nilles, “Supersymmetry, supergravity and particle physics”, *Physics Reports* **110**, 1–162 (1984).
- ³¹A. Collaboration, “Antiproton Flux, Antiproton-to-Proton Flux Ratio, and Properties of Elementary Particle Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station”, *Phys. Rev. Lett.* **117**, 091103 (2016).
- ³²”*2HDM Benchmark Recommendations*”.
- ³³L. Evans and P. Bryant, “LHC machine”, *Journal of Instrumentation* **3**, S08001 (2008).
- ³⁴K. Schindl, “The Injector Chain for the LHC; rev. version”, (1999).
- ³⁵E. Mobs, “The CERN accelerator complex. Complexe des accélérateurs du CERN”, General Photo (2016).
- ³⁶ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation* **3**, S08002 (2008).
- ³⁷LHCb Collaboration, “The LHCb Detector at the LHC”, *Journal of Instrumentation* **3**, S08005 (2008).
- ³⁸TOTEM Collaboration, “The TOTEM detector at LHC”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **617**, 11th Pisa Meeting on Advanced Detectors, 62–66 (2010).
- ³⁹MoEDAL Collaboration, “The physics programme of the MoEDAL experiment at the LHC”, *International Journal of Modern Physics A* **29**, 1430050 (2014).
- ⁴⁰CMS Collaboration, “The CMS experiment at the CERN LHC”, *Journal of Instrumentation* **3**, S08004 (2008).
- ⁴¹ATLAS Collaboration (ATLAS), “The ATLAS experiment at the CERN Large Hadron Collider”, *JINST* **3**, Also published by CERN Geneva in 2010, S08003 (2008).

- ⁴² "The Forward Calorimeter Prototype 1993".
- ⁴³ J. Pequeno, "Computer generated image of the ATLAS Muons subsystem", 2008.
- ⁴⁴ T. Kawamoto, S. Vlachos, L. Pontecorvo, J. Dubbert, G. Mikenberg, P. Iengo, C. Dallapiccola, C. Amelung, L. Levinson, R. Richter, and D. Lellouch, *New Small Wheel Technical Design Report*, tech. rep., ATLAS New Small Wheel Technical Design Report (2013).
- ⁴⁵ "Public ATLAS Luminosity Results for Run-2 of the LHC", <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- ⁴⁶ J. Allard, N. Andari, M. Anfreville, D. Attié, E. Aubernon, S. Aune, H. Bachacou, F. Balli, F. Bauer, J. Beltramelli, J. Bennet, T. Benoit, H. Bervas, T. Bey, S. Bouaziz, M. Boyer, T. Challey, G. Cara, T. Chaleil, T. Chevalérias, X. Coppollani, J. Costa, G. Decock, F. Deliot, D. Denysiuk, D. Desforge, G. Disset, G. Durand, R. Durand, J. Elman, E. F. Ribas, M. Fontaine, A. Formica, J. Galán, W. Gamache, A. Giganon, J. Giraud, P. Giraud, G. Glonti, C. Goblin, P. Graffin, J. Guillard, S. Hassani, S. Hervé, S. Javello, F. Jeanneau, D. Jourde, S. Jurie, M. Kebbiri, T. Kawamoto, C. Lampoudis, J. Laporte, D. Leboeuf, M. Lefevre, M. Lohan, C. Loiseau, P. Magnier, I. Mandjavidze, J. Manjarrés, P. Mas, M. Mur, R. Nikolaidou, A. Peyaud, D. Pierrepont, Y. Piret, P. Ponsot, G. Prono, M. Riallot, F. Rossi, P. Schune, T. Vacher, M. Vandenbroucke, A. Vigier, C. Vuillemin, M. Usseglio, and Z. Wu, "The large inner Micromegas modules for the Atlas Muon Spectrometer Upgrade: construction, quality control and characterization", *Nucl. Instrum. Methods Phys. Res., A* **1026**, To be submitted to NIMA, 166143 (2022).
- ⁴⁷ G. Iakovidis, V. Polychronakos, and G. De Geronimo, *VMM - An ASIC for micropattern detectors*, tech. rep. (CERN, Geneva, 2015).
- ⁴⁸ G. Iakovidis, *MMFE8 production and refurbishment status*, 2021.
- ⁴⁹ ATLAS Collaboration, "ATLAS data quality operations and performance for 2015-2018 data-taking", *JINST* **15**, P04003 (2020).
- ⁵⁰ e. a. G. Avoni, "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS", *JINST* **13**, P07017 (2018).
- ⁵¹ ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, 2022.
- ⁵² S. van der Meer, "Calibration of the effective beam height in the ISR", CERN-ISR-PO-68-31 (1968).
- ⁵³ "TriggerPhysicsMenu", ATLAS twiki.

- ⁵⁴ATLAS Collaboration, “Performance of electron and photon triggers in ATLAS during LHC Run 2”, *European Physical Journal C* **80**,47 (2020).
- ⁵⁵ATLAS Collaboration, “Performance of the ATLAS muon triggers in Run 2”, *JINST* **15** P09015 (2020).
- ⁵⁶S. Wang, “ATLAS Computing and Data Preparation”, in ATLAS Induction Day and Software Tutorial, January 20-24 2020 (2020).
- ⁵⁷N. Krumnack, “XAOD Analysis Introduction”, in ATLAS Induction Day and Software Tutorial, January 20-24, 2020 (2020).
- ⁵⁸“GoodRunListsForAnalysisRun2”, ATLAS Twiki.
- ⁵⁹“PHYS_StandardGRL_All_Good_25ns”, ATLAS DQM website.
- ⁶⁰“PHYS_StandardGRL_All_Good_25ns”, ATLAS DQM website.
- ⁶¹“PHYS_StandardGRL_All_Good_25ns_TriggerNo17e33prim”, ATLAS DQM website.
- ⁶²“PHYS_StandardGRL_All_Good_25ns_TriggerNo17e33prim”, ATLAS DQM website.
- ⁶³*Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-015> (CERN, Geneva, 2015).
- ⁶⁴ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data”, *Journal of Instrumentation* **14**, P03017 (2019).
- ⁶⁵“TopxAODStartGuideR21”, ATLAS Twiki.
- ⁶⁶S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytrcek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell’Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin,

- M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O’Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. Wellisch, T. Wenaus, D. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche, “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**, 250–303 (2003).
- ⁶⁷P. Skands, “Introduction to QCD”, in *Searching for new physics at small and large scales* (Sept. 2013).
- ⁶⁸M. Z. R. Frederix D. Pagani, “Large NLO corrections in $t\bar{t}W^\pm$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions”, *JHEP* **02** (2018) 031 (2018).
- ⁶⁹J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *Journal of High Energy Physics* **2014**, 10.1007/jhep07(2014)079 (2014).
- ⁷⁰S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX”, *Journal of High Energy Physics* **2010**, 10.1007/jhep06(2010)043 (2010).
- ⁷¹E. Bothmann, G. S. Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz, S. Schumann, and F. Siegert, “Event generation with sherpa 2.2”, *SciPost Physics* **7**, 10.21468/scipostphys.7.3.034 (2019).
- ⁷²M. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2014).
- ⁷³R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. D. Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, “Parton distributions for the LHC run II”, *Journal of High Energy Physics* **2015**, 10.1007/jhep04(2015)040 (2015).

- ⁷⁴L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, “Parton distributions in the LHC era: MMHT 2014 PDFs”, *The European Physical Journal C* **75**, 10.1140/epjc/s10052-015-3397-6 (2015).
- ⁷⁵A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Höche, H. Hoeth, F. Krauss, L. Lönnblad, E. Nurse, P. Richardson, S. Schumann, M. H. Seymour, T. Sjöstrand, P. Skands, and B. Webber, “General-purpose event generators for lhc physics”, *Physics Reports* **504**, 145–233 (2011).
- ⁷⁶*ATLAS Pythia 8 tunes to 7 TeV data*, tech. rep. (CERN, Geneva, 2014).
- ⁷⁷T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to PYTHIA 8.1”, *Computer Physics Communications* **178**, 852–867 (2008).
- ⁷⁸C. ATLAS, M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi, J. Virzi, I. Vivarelli, M. Werner, S. Yamamoto, and T. Yamanaka, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, tech. rep. (CERN, Geneva, 2010).
- ⁷⁹ATLAS Collaboration, “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”, *Journal of Instrumentation* **14**, P12006–P12006 (2019).
- ⁸⁰T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-SOFT-PUB-2007-007> (CERN, Geneva, 2007).
- ⁸¹R. Frühwirth, “Application of kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262**, 444–450 (1987).
- ⁸²W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-LARG-PUB-2008-002> (CERN, Geneva, 2008).
- ⁸³ATLAS Collaboration, “Topological cell clustering in the ATLAS calorimeters and its performance in LHC run 1”, *The European Physical Journal C* **77**, 10.1140/epjc/s10052-017-5004-5 (2017).
- ⁸⁴“IsolationSelectionTool”, ATLAS Twiki.

- ⁸⁵“Electron and Photon Selection and Identification for Run2 and Run3”, ATLAS Twiki.
- ⁸⁶“Electron Charge ID Selector Tool”, ATLAS Twiki.
- ⁸⁷ATLAS Collaboration, “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”, European Journal of Physics C **81** (2021).
- ⁸⁸J. Illingworth and J. Kittler, “A survey of the hough transform”, Computer Vision, Graphics, and Image Processing **44**, 87–116 (1988).
- ⁸⁹ATLAS Collaboration, ““Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV””, The European Physical Journal C **76** (2016).
- ⁹⁰“IsolationSelectionTool”, ATLAS Twiki.
- ⁹¹ATLAS Collaboration, “Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, European Journal of Physics C **81** (2021).
- ⁹²ATLAS Collaboration, ““Jet reconstruction and performance using particle flow with the ATLAS Detector””, The European Physical Journal C **77** (2017).
- ⁹³M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, Journal of High Energy Physics **2008**, 063–063 (2008).
- ⁹⁴“JetVertexTagger for Run-2 in reco and analysis”, ATLAS Twiki.
- ⁹⁵*Tagging and suppression of pileup jets with the ATLAS detector*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-018> (CERN, Geneva, 2014).
- ⁹⁶A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla, *TMVA - Toolkit for Multivariate Data Analysis*, 2007.
- ⁹⁷“*JVTCalibration*”.
- ⁹⁸“*Tagger Recommendations for Release 21*”.
- ⁹⁹ATLAS Collaboration, “Performance of b-jet identification in the atlas experiment”, Journal of Instrumentation **11**, P04008 (2016).

- ¹⁰⁰*Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013> (CERN, Geneva, 2017).
- ¹⁰¹C. Varni (ATLAS), *Tracking and flavour-tagging performance at ATLAS*, tech. rep. (CERN, Geneva, 2021).
- ¹⁰²*Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-003> (CERN, Geneva, 2017).
- ¹⁰³B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, “Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC”, *Journal of High Energy Physics* **2015**, 10.1007/jhep02(2015)075 (2015).
- ¹⁰⁴D. Krohn, J. Thaler, and L.-T. Wang, “Jet trimming”, *Journal of High Energy Physics* **2010**, 10.1007/jhep02(2010)084 (2010).
- ¹⁰⁵N. Asbah, S. P. Berlendis, M. Cristinziani, F. Deliot, M. Faraj, E. K. Filmer, S. Hu, P. Jackson, K. Johns, A. Juste Rozas, M. H. Klein, A. X. Y. Kong, M. E. Leblanc, L. Li, X. Li, R. E. Lindley, R. Madar, J. Magro, C. Nellist, O. O. Oncel, N. Orlando, G. Panizzo, D. Paredes Hernandez, R. Peters, M. Pinamonti, J. Qian, Y. Qin, A. Quadt, L. Rustige, P. Sabatini, L. Serkin, E. Shabalina, F. Sohns, A. Sonay, E. Varnes, Z. Zheng, M. Murin, A. Sharma, and M. Morii, *Search for $t\bar{t}\bar{t}$ Standard Model Production in the Single Lepton and Opposite Sign Dilepton Final State in proton-proton collisions at $\sqrt{s} = 13$ TeV using full Run 2 data set with the ATLAS Detector*, tech. rep. (CERN, Geneva, 2020).
- ¹⁰⁶ATLAS Collaboration, “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV”, *The European Physical Journal C* **78**, 10.1140/epjc/s10052-018-6288-9 (2018).
- ¹⁰⁷R. Harlander, M. Krämer, and M. Schumacher, *Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach*, 2011.
- ¹⁰⁸M. van Beekveld, A. Kulesza, and L. M. Valero, *Threshold resummation for the production of four top quarks at the LHC*, 2022.

- ¹⁰⁹de Florian et. al. (LHCHiggsCrossSectionWorkingGroup), *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports: Monographs, 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG> (CERN, Geneva, 2017).
- ¹¹⁰ATLAS Collaboration, “Measurement of $W^\pm Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *The European Physical Journal C* **79**, 10.1140/epjc/s10052-019-7027-6 (2019).
- ¹¹¹ATLAS Collaboration, “Measurement of the production cross-section of a single top quark in association with a z boson in proton—proton collisions at 13 TeV with the ATLAS detector”, *Physics Letters B* **780**, 557–577 (2018).
- ¹¹²F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, “ $t\bar{t}H$ associated production at the LHC”, *The European Physical Journal C* **77**, 10.1140/epjc/s10052-017-4601-7 (2017).
- ¹¹³ATLAS Collaboration, “Performance of the ATLAS trigger system in 2015”, *The European Physical Journal C* **77**, 10.1140/epjc/s10052-017-4852-3 (2017).
- ¹¹⁴“*JetEtMiss Recommendations for Release 21 MC16 (2015-2018)*”.
- ¹¹⁵ATLAS Collaboration, “Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *The European Physical Journal C* **75**, 10.1140/epjc/s10052-014-3190-y (2015).
- ¹¹⁶ATLAS Collaboration, “Search for $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ production in a final state with one lepton or two opposite-sign leptons using the full Run 2 pp collisions data at $\sqrt{s} = 13$ TeV”, tech. rep. ANA-EXOT-2019-26-INT2 (2023).
- ¹¹⁷*Jet reclustering and close-by effects in ATLAS run II*, tech. rep., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2017-062> (CERN, Geneva, 2017).
- ¹¹⁸ATLAS Collaboration, “Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Journal of High Energy Physics* **2019**, 10.1007/jhep04(2019)046 (2019).

- ¹¹⁹ATLAS Collaboration, “Measurements of top-quark pair differential and double-differential cross-sections in the l +jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector”, *The European Physical Journal C* **79**, 10.1140/epjc/s10052-019-7525-6 (2019).
- ¹²⁰G. V. Moustakides and K. Basioti, *Training neural networks for likelihood/density ratio estimation*, 2019.
- ¹²¹E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, “Scaling patterns for QCD jets”, *Journal of High Energy Physics* **2012**, 10.1007/jhep10(2012)162 (2012).
- ¹²²P. C. Bhat, “Multivariate analysis methods in particle physics”, *Annual Review of Nuclear and Particle Science* **61**, 281–309 (2011).
- ¹²³P. A. Flach, “ROC Analysis”, in *Encyclopedia of Machine Learning*, edited by G. W. C. Sammut (Springer, Boston, MA), pp. 869–875.
- ¹²⁴E. Alpaydin, *Machine Learning: the New AI* (MIT Press, 2016).
- ¹²⁵B. Denby, “Neural networks and cellular automata in experimental high energy physics”, *Computer Physics Communications* **49**, 429–448 (1988).
- ¹²⁶J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, *Neural Message Passing for Quantum Chemistry*, 2017.
- ¹²⁷C. Birch-Sykes, *Update: HBSM $ttH/A \rightarrow 4\text{tops } 1\text{LOS}$ channel (MVA classification)*, 2022.
- ¹²⁸Coadou, Yann, “Boosted decision trees and applications”, *EPJ Web of Conferences* **55**, 02004 (2013).
- ¹²⁹M. Fey and J. E. Lenssen, *Fast Graph Representation Learning with PyTorch Geometric*, May 2019.
- ¹³⁰P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, C. Gulcehre, F. Song, A. Ballard, J. Gilmer, G. Dahl, A. Vaswani, K. Allen, C. Nash, V. Langston, C. Dyer, N. Heess, D. Wierstra, P. Kohli, M. Botvinick, O. Vinyals, Y. Li, and R. Pascanu, *Relational inductive biases, deep learning, and graph networks*, 2018.
- ¹³¹G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *The European Physical Journal C* **71**, 10.1140/epjc/s10052-011-1554-0 (2011).
- ¹³²W. Verkerke and D. Kirkby, *The roofit toolkit for data modeling*, 2003.
- ¹³³K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, “HistFactory: A tool for creating statistical models for use with RooFit and RooStats”,

¹³⁴S. Mitra, $t\bar{t} + X$ and $t(\bar{t}) + X$ measurements in ATLAS and CMS, 2023.

¹³⁵CMS Collaboration, “Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13\text{TeV}$ ”, The European Physical Journal C **80**, 10.1140/epjc/s10052-019-7593-7 (2020).

¹³⁶LS3 schedule change, 2022.