OBSERVATION AND MEASUREMENT OF THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH A VECTOR BOSON AND DECAYING TO A PAIR OF BOTTOM QUARKS WITH THE ATLAS DETECTOR AT LHC

Approved by:

Dr. Stephen Sekula
Associate Professor of Physics, SMU

Dr. Fredrick Olness
Professor of Physics, SMU

Dr. Robert Kehoe
Professor of Physics, SMU

Dr. Charles Young
Senior Scientist, SLAC

CERN-THESIS-2020-225
27/10/2020

OBSERVATION AND MEASUREMENT OF THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH A VECTOR BOSON AND DECAYING TO A PAIR OF BOTTOM QUARKS WITH THE ATLAS DETECTOR AT LHC

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Physics

by

Peilong Wang

B.S., Physics, University of Science and Technology of China

December 19, 2020

Copyright (2020)

Peilong Wang

All Rights Reserved

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor Dr. Stephen Sekula for the advisory and support of my research at Southern Methodist University (SMU), Argonne National Laboratory (ANL), Organisation européenne pour la recherche nucléaire (CERN), and SLAC National Accelerator Laboratory (SLAC).

I would like to express my appreciation to Dr. Fredrick Olness, Dr. Robert Kehoe, and Dr. Charles Young for acting as my dissertation committee.

I would like to thank Dr. Ryszard Stroynowski for the financial support and the wisdom on physics shared with me; Dr. Roberto Vega and Dr. Kent Hornbostel for the wonderful lectures given on Quantum Mechanics and Quantum Field Theory; and Dr. Jingbo Ye for his great support of my hardware work in Opto-Electronics Lab at SMU and SLAC.

I am grateful to Dr. Jinlong Zhang, Dr. Jeremy Love and Dr. James Proudfoot for the guidance and advisory on my project at ANL; to Dr. Valerio Dao, Dr. Tatsuya Masubuchi and Dr. Francesco Lo Sterzo for the advice and help on my physics analysis at CERN; and to Dr. Dong Su, Dr. Charles Young and Dr. Zijun Xu for offering me the opportunity and support to work on the detector upgrade project at SLAC.

I acknowledge the U.S. Department of Energy for the grant support of my research, and in particular, US-ATLAS for the grant to relocate to SLAC.

I also thank my parents for their great understanding and patience.

Last but not least, many thanks to everyone who has helped me in the past years while whose name does not appear here.

Wang, Peilong B.S., Physics, University of Science and Technology of China

Observation and Measurement of the Higgs Boson Produced in Association with a Vector Boson and Decaying to a Pair of Bottom Quarks with the ATLAS Detector at LHC

Advisor: Dr. Stephen Sekula
Doctor of Philosophy degree conferred December 19, 2020
Dissertation completed December 19, 2019

A search of the Higgs boson decaying into $b\bar{b}$ pair using the WH and ZH associated production is performed with the ATLAS detector. The analyzed data were collected in 2015, 2016, 2017 and 2018 LHC Run-2 at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 139 fb^{-1}. Final states that contains 0, 1, and 2 charged leptons (electrons or muons, denoted as ℓ) are considered, to target the processes of $Z \to \nu\nu$, $W \to \ell\nu$ and $Z \to \ell\ell$. Object reconstruction, event selection, and signal and control region definition are included. Systematic uncertainty and statistical analysis used to extract the final results are presented. This analysis observes the production $pp \to (W, Z)H$ in $H \to b\bar{b}$ channel at a significance of 6.7 standard deviations (where the expectation is 6.7 standard deviations) and represents the most precise observation to date of this physics process. Cross-sections of VH (WH and ZH) with $V \to$ leptons and $H \to b\bar{b}$ are measured as a function of the vector boson transverse momentum in the kinematic fiducial volumes and found to be consistent with the Standard Model predictions.

TABLE OF CONTENTS

LIST OF FIGURES xii

LIST OF TABLES xxi

CHAPTER

1 Introduction 1

2 Theoretical Framework and Higgs phenomenology 4

2.1. Theoretical Framework 4

2.1.1. Elementary Particles 6

2.1.2. Gauge Theory of Elementary Particles 6

2.1.2.1. (Special) Relativistic Fields 7

2.1.2.1.1 Electromagnetic Field 7

2.1.2.1.2 A Scalar (Spin-0) Field 8

2.1.2.1.3 A Spinor (Spin-$\frac{1}{2}$) Field 8

2.1.2.1.4 A Vector (Spin-1) Field 9

2.1.2.1.5 Summary 9

2.1.2.2. Yang-Mills Theory and Chromodynamics 10

2.1.2.2.1 Yang-Mills Theory 10

2.1.2.2.2 Chromodynamics 13

2.1.2.3. Spontaneous Symmetry-breaking and the Higgs Mechanism 14

2.1.2.3.1 Spontaneous Symmetry-breaking 14

2.1.2.3.2 The Higgs Mechanism 16

2.1.2.4. The Lagrangian of a Theoretical Model 19

2.2. Higgs Phenomenology 21
2.2.1. Higgs Production at LHC 21
2.2.2. Higgs Decay Channels 23
3 The LHC and ATLAS Detector 26
3.1. Large Hadron Collider (LHC) 26
3.1.1. LHC Magnet 28
3.1.1.1. Dipole and Quadrupole 28
3.1.1.2. Superconductivity 30
3.1.2. Beam Injection and Beam Dump 31
3.2. The ATLAS Detector 31
3.2.1. Physics Requirements 33
3.2.2. Magnet System 35
3.2.3. Inner Detector 36
3.2.4. Calorimetry 37
3.2.5. Muon Spectrometer 38
3.2.6. Luminosity Monitoring Detectors 38
3.2.7. Trigger and Data Acquisition System 40
3.3. My Contribution to the ATLAS Detector Upgrade 42
3.3.1. Contribution to the Faster TrcKer (FTK) Project 42
3.3.1.1. FLIC and its Associated ATCA Blade 44
3.3.1.2. FLIC Data Quality Monitoring Software 44
3.3.2. Contribution to the Inner Tracker (ITk) Project 49
3.3.2.1. Data Transmission Chain of the ITk Pixel Detector 49
3.3.2.2. Test of the GBCR ASIC 51

3.3.2.3. Test of the Pixel Detector Data Transmission Chain 53

4 Data and Monte Carlo 55

4.1. Data 55

4.2. Monte Carlo Samples 56

4.2.1. VH Signal 58

4.2.2. Vector Boson + Jets 60

4.2.3. Top-Pair Production 61

4.2.4. Single-top 63

4.2.5. Diboson 64

5 Physics Object Reconstruction 67

5.1. Trigger Selection 67

5.1.1. Recommended Trigger Chains 68

5.2. Physics Object Reconstruction 69

5.2.1. Lepton 70

5.2.1.1. Electron 70

5.2.1.2. Muon 71

5.2.1.3. Hadronic Tau 73

5.2.2. Jet 74

5.2.2.1. Standard Jet Collections 75

5.2.2.2. b-tagging 76

5.2.2.2.1 Truth Tagging and Hybrid Truth Tagging 78

5.2.2.3. b-jet Energy Correction 78

5.2.2.3.1 Muon-in-jet Correction 79

5.2.2.3.2 PtReco Correction 79

5.2.2.3.3 Kinematic Fit 80
5.2.3. Overlap Removal 83
5.2.4. Missing Transverse Momentum 85
6 Event Selection 86
6.1. Event Selection 86
6.1.1. 0-lepton 87
6.1.2. 1-lepton 87
6.1.3. 2-lepton 88
6.2. Signal and Control Region 89
6.3. Regions for Cross-section Measurement 93
7 $t\bar{t}$ Rejection and Tau-veto in the 1-lepton Channel 96
7.1. $t\bar{t}$ Rejection in the 1-lepton Channel 96
7.2. Tau-veto in the 1-lepton Channel 108
8 $VHbb$ Analysis 114
8.1. Multivariate Discriminants 114
8.1.1. Boosted Decision Tree 114
8.1.2. Input Variables 119
8.1.3. Set-up and Training 122
8.2. Systematic Uncertainties 124
8.2.1. Experimental Uncertainty 125
8.2.2. Modeling Uncertainty 126
8.2.2.1. V+jet 128
8.2.2.2. $t\bar{t}$ 128
8.2.2.2.1 $t\bar{t}$ shape uncertainty (shape comparison method) 129

8.2.2.3. Single Top-quark 130
8.2.2.4. Diboson Production 133
8.2.2.5. Multi-jet Background in the 1-lepton Channel 133
8.2.2.6. VH Signal 134
8.3. Statistical Analysis 138
8.3.1. Hypothesis 138
8.3.2. Likelihood Function 139
8.3.3. Hypothesis Testing 140
8.3.4. Fit to Data 142
9 Results 144
9.1. Signal Strength Measurements 144
9.1.1. VH Signal Strength 144
9.1.2. Dijet-mass Cross-check 152
9.1.3. Diboson Cross-check 154
9.2. Cross-section Measurements 156
10 Conclusion 159
APPENDIX
A $t\bar{t}$ Rejection in the 1-lepton Channel 162
B Tau-veto in the 1-lepton Channel 164
B.1. Cut-flow 164
B.2. Over-training Check 166
B.3. ROC Curves 168
B.4. MVA Input Variable Distribution Comparison 170
B.5. Correlation Matrices 181

B.6. Evaluation in the region of $150 < p_{\mathrm{T}}^{V} < 250$ GeV 184

B.7. Evaluation in the region of $p_{\mathrm{T}}^{V} > 250$ GeV 187

C $t\bar{t}$ Modeling: Shape Uncertainties Using the Shape Comparison Method 190

C.1. Fits for shape uncertainties 190

C.1.1. 0-lepton fits 190

C.1.2. 1-lepton fits 193

D Pre-fit Data-MC Comparison in the 1-lepton Channel 197

BIBLIOGRAPHY 202

LIST OF FIGURES

Figure Page

2.1 A picture of elementary particles. The Higgs boson is located in the center, with quarks in the top-left, leptons in the bottom-left and mediators in the right [15]. 5

2.2 (left) Potential $V(\phi) = \mu^2(\phi^\dagger\phi) + \frac{\lambda}{4}(\phi^\dagger\phi)^2$ with symmetry at ground states; (right) potential $V(\phi) = -\mu^2(\phi^\dagger\phi) + \frac{\lambda}{4}(\phi^\dagger\phi)^2$, where the symmetry is broken. 15

2.3 Diagram of the Higgs production at LHC: (a) gluon-gluon fusion, (b) vector-boson fusion, (c) associated production with a vector boson, and (d) associated production with top quarks [23]. 21

2.4 Cross sections of the Higgs boson production as a function of the center-of-mass energy for proton-proton collisions. The theoretical uncertainties are represented as bands [24]. 22

2.5 The branching ratios of the Higgs boson decays near $m_H = 125$ GeV . The theoretical uncertainties are represented as bands [24]. 24

3.1 Facilities for the accelerator at CERN [27]. 26

3.2 (left) The "two-in-one" design of the dipole magnets; (right) cross-section of dipole magnets in a cryostat [26]. 29

3.3 (left) Two quadrupoles work together to keep the protons tightly bunched in the direction transverse to the z-axis; (middle) picture of a quadrupole magnet; (right) the current flow of a quadrupole [36]. 29

3.4 (left) Cross-section of a superconducting cable; (middle) picture and cross-section of a strand; (right) the distribution of the conductor in the dipole coil cross-section [26]. 30

3.5 Schematic layout of beam dumping system elements around LHC Point 6 (distance in m) [26]. 31

3.6 Overview of the ATLAS detector [37]. 32

3.7 The ATLAS experiment location at the LHC and the coordinate system of ATLAS detector are shown in the figure. The side-A of the detector is defined as the side of positive z and side-C is the side of negative z. This picture is modified from Ref. [39]. 33

3.8 Geometry of the ATLAS magnet system windings [37]. 35

3.9 (left) Picture of the barrel Inner Detector traversed by a charged track; (right) plan view of a quarter-section of the Inner Detector showing each of the sub-detector elements [37]. 36

3.10 (left) View of the ATLAS calorimetry; (right) cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in different calorimeters [37]. 37

3.11 (left) Cross-section of the barrel muon system; (right) cross-section of the muon system in a plane containing the beam axis. Note: the three-letter acronym denotes the location and type of the muon detector installed. Refer to Ref. [37] for details of the notation. 39

3.12 (left) Sketch of the LUCID-2 detector; (right-top) Hamamatsu R760 PMT (quartz window 10 mm diameter); (right-middle) Hamamatsu R760 PMT (quartz window 7 mm diameter); (right-bottom) Bundles of LUCID-2 read-out fibers. [55] 40

3.13 Schematic layout of the ATLAS trigger and data acquisition system in Run 2 [59]. 41

3.14 Illustration of the FLIC data quality monitoring software running on the ATCA blade. 43

3.15 Test-stand for the FLIC DQM software test at ANL. `USTB13` and `USTB17` are external computer names. 43

3.16 The structure of the FLIC DQM software. FLIC DQM software is a parallel program. While the DQM software keeps receiving events and buffer them, standalone threads access the buffered events, decode them and publish the results to Online Histogram (OH). 45

3.17 The UDP datagram protocol is used for the associated ATCA blade to receive data. It is designed to be a single direction communication, which means that FLIC boards can only send data and the ATCA blade can only receive data. 45

3.18 The column in the left is the data format that the ATCA blade received from the FLIC board; the column in the right is the data format decoded on the ATCA blade. Different blocks of the data (assembly header, record header, track header, record trailer, etc.) are marked with different background colors for the convenience of comparison. 47

3.19 FLIC DQM software publishes histograms directly to Online Histogram (OH). . 48

3.20 The histograms published to OH by the FLIC DQM software can be viewed from the TDAQ control panel. The box in red at the top-right corner shows a histogram of the χ^2 of the monitored tracks; the box in red at the bottom-right corner shows a list of monitored histograms published. 48

3.21 (left) A schematic layout of the ITk detector. The Pixel detector is marked in red color; (right) the total ionizing dose for the Pixel detector in the HL-LHC. 50

3.22 The data transmission chain for the ITk Pixel detector. 50

3.23 The test-stand for the GBCR (version 1) ASIC test at SMU. 52

3.24 The test-stand for the GBCR (version 1) ASIC test at SMU; Test results of the GBCR (version 1) ASIC at SMU. 52

3.25 The data transmission chain test at SLAC. The Bit-Error-Rate of the system achieves a level of less than 10^{-10}. 54

4.1 (left) The total integrated luminosity delivered by LHC and recorded by ATLAS during Run-2 at $\sqrt{s} = 13$ TeV; (right) the distribution of the mean number of interactions per bunch crossing after weighted by luminosity. Data are recorded at stable beams (special runs and machine commissioning periods are included as well) [62]. 55

4.2 A diagram showing the physics processes of a typical proton-proton collision [68]. 57

5.1 Illustration of the physics of a b-jet [154]. 77

5.2 The structure of the MV2 b-tagging algorithm. 77

5.3 The PtReco correction factor for the b-jets with muon is marked as red and for those without muon is marked as blue. The jets after muon-in-jet correction are denoted as "OneMu". 80

5.4 Comparison of the m_{bb} resolution with the muon-in-jet correction, PtReco correction and Kinematic Fit. All corrections are applied to the jet energy scale. The simulated $qqZH$ events in the 2-jet category, $p_T^Z > 150$ GeV region of the 2-lepton channel are shown [134]. 82

5.5 The cut-flow diagram of the overlap-removal procedure. 84

6.1 Signal yield distribution of ΔR between the two selected b-jets as a function of p_T^V in the 1-lepton channel are shown in the 2-tag 2-jet (a) and 2-tag 3-jet (b) categories. The black lines demonstrate the upper and lower continuous cuts used to categorize the events into the signal and control regions [159]. 89

6.2 (Top) migration matrix of the expected signal yield between the truth-p_{T}^V-regions (x-axis) and reconstructed-p_{T}^V-regions (y-axis); (bottom) migration matrix of the signal fraction (with respect to the reconstructed-p_{T}^V-region) between the truth-p_{T}^V-regions (x-axis) and reconstructed-p_{T}^V-regions (y-axis). Regions with the signal yield below 0.1 or fraction below 0.1% are ignored [159]. 95

7.1 The diagrams are $t\bar{t}$ decaying (a) dileptonically, (b) semi-leptonically, (c) to a leptonic τ, (d) to a hadronic τ and (e) to $\tau\tau$. Diagram (f) is WH signal. . . 96

7.2 After the selection of 2-tag 2-jet category in the $p_{\mathrm{T}}^V > 150$ GeV region of the 1-lepton channel, each $t\bar{t}$ event is categorized according to decay mode (in truth-level) and then filled into this histogram. 97

7.3 The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 3-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^V > 150$ GeV) after the distributions are normalized to unity. 97

7.4 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel. 102

7.5 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel. 103

7.6 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel. 104

7.7 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel. 105

7.8 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel. 106

7.9 Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel. 107

7.10 The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 2-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^{V} > 150$ GeV) after the distributions are normalized to unity. 110

7.11 The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 3-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^{V} > 150$ GeV) after the distributions are normalized to unity. 110

7.12 (a): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel; (b): the comparison of binned background MVA distributions with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel; (c): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} > 250$ GeV region of 2-tag 2-jet category of the 1-lepton channel; (d): the comparison of binned background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} > 250$ GeV region of 2-tag 2-jet category of the 1-lepton channel. 112

7.13 (a): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel; (b): the comparison of binned background MVA distributions with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel; (c): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} > 250$ GeV region of 2-tag 3-jet category of the 1-lepton channel; (d): the comparison of binned background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} > 250$ GeV region of 2-tag 3-jet category of the 1-lepton channel. 113

8.1 An example of a decision tree. The root node is represented as a Rectangle; each (internal) node is represented as a circle; and each leaf is represented as a triangle. Each of them contains a different number of signal and background events. Applied selection cuts are shown at the top of the root node and (internal) nodes. 116

8.2 Illustration of *AdaBoost* algorithm. AdaBoost grows a forest of stumps. The training samples are reweighted using the previous classifier before a new stump is grown. The errors that the previous stump makes influence how the following stump is made. 118

8.3 Over-training checks for the 2-jet (left) and 3-jet (right) category of p_{T}^{V} >150 GeV region trained on odd-number MC samples in the 1-lepton channel. The ROC curve compares the training data-set with a statistically independent testing data-set 124

8.4 $m_{b\bar{b}}$ shape variations used in the 0-/1-lepton channel for the 75GeV $< p_T^V <$ 150GeV and $p_T^V >$ 150GeV regions. Figures (a-d) show the systematic variations selected from the all MC-to-MC comparisons summarized in figures (e-h). 131

8.5 p_T^V shape variations used in the 0-/1-lepton channel for the 2-/3-jet categories. Figures (a-b) show the systematic variations selected from the all MC-to-MC comparisons summarized in figures (c-d). 132

8.6 (left) Illustration of the p-value obtained from an observed value of the test statistic $t_{\mu,\mathrm{obs}}$. (right) Illustration of the relation between the significance Z and the p-value for a standard normal distribution $\varphi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$ [195]. . . . 142

9.1 The BDT$_{VH}$ output post-fit distributions in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channels for 2-b-tag 2-jet events, for the $150 < p_{\mathrm{T}}^{\mathrm{V}} < 250$ GeV (left) and $p_{\mathrm{T}}^{\mathrm{V}} > 250$ GeV (right) $p_{\mathrm{T}}^{\mathrm{V}}$ regions. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_h = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.02$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal ($\mu = 1.02$) and background is shown in the lower panel. The BDT$_{VH}$ output distributions are shown with the binning used in the global likelihood fit. Figures are from the publication of the $VHbb$ analysis of ATLAS [159]. 149

9.2 Best values of the signal strength $\mu_{VH}^{b\bar{b}}$ for the 0-, 1- and 2- lepton channels and their combination in the unconditional fit to the Run 2 data of the three channels combined. The (black) total observed uncertainty is quoted together with its decomposition in the (green) statistical component, and systematic component. In this plot the uncertainty due to background scale factors is included in the statistical component [109]. 150

9.3 The fitted values of the Higgs boson signal strength μ_{VH}^{bb} for $m_h = 125$ GeV for the WH and ZH processes and their combination. The individual μ_{VH}^{bb} values for the $(W/Z)H$ processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently. The probability of compatibility of the individual signal strengths is 71%. Figures are from the publication of the $VHbb$ analysis of ATLAS [159]. 150

9.4 The distribution of m_{bb} in data after subtracting all backgrounds except the WZ and ZZ diboson processes, as obtained with the dijet-mass analysis. The contributions from all lepton channels, $p_{\mathrm{T}}^{\mathrm{V}}$ regions and number-of-jets categories are summed and weighted by their respective S/B ratios, with S being the total fitted signal and B the total fitted background in each region. The expected contribution of the associated WH and ZH production of a SM Higgs boson with $m_h = 125$ GeV is shown scaled by the measured signal strength ($\mu = 1.17$). The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band. Figure is from the publication of the $VHbb$ analysis of ATLAS [159]. 153

9.5 The fitted values of the VZ signal strength μ_{VZ}^{bb} for the WZ and ZZ processes and their combination. The individual μ_{VZ}^{bb} values for the WZ and ZZ processes are obtained from a simultaneous fit with the signal strengths for each of the WZ and ZZ processes floating independently. The probability of compatibility of the individual signal strengths is 27%. Figure is from the publication of the $VHbb$ analysis of ATLAS [159] 155

9.6 Measured VH, $V \to$ leptons cross-sections times the $H \to b\bar{b}$ branching fraction in the reduced STXS scheme. Figure is from the publication of the $VHbb$ analysis of ATLAS [159]. 157

1.1 After the 2t2j selection of 150ptv region in 1-lep and then **cut on BDT score greater than 0 (∼90% signal efficiency)**, each ttbar event is asked what its decay mode is (in truth-level) and then filled here. dileptonic (e-e, μ-μ, e-μ): 10.0%; semileptonic (jet-e, jet-μ): 60.8%; hadronictau (e-τ, μ-τ): 23.1%; leptonictau (jet-τ): 4.5%; τ-τ: 1.6%. 163

1.2 After the 2t3j selection of 150ptv region in 1-lep and then **cut on BDT score greater than 0 (∼90% signal efficiency)**, each ttbar event is asked what its decay mode is (in truth-level) and then filled here. dileptonic (e-e, μ-μ, e-μ): 6.9%; semileptonic (jet-e, jet-μ): 69.2%; hadronictau (e-τ, μ-τ): 17.5%; leptonictau (jet-τ): 5.0%; τ-τ: 1.3%. 163

2.1 Over-training Check 166

2.2 BDT Efficiency 167
2.3 ROCComparison_1MinusBgEff_default_TauVeto 168
2.4 ROCComparison_RejFactor_default_TauVeto 168
2.5 ROCRatio_default_TauVeto_1MinusBgEff 169
2.6 ROCRatio_default_TauVeto_RejFactor 169
2.7 MVA Comparison 170
2.8 MVA Input Variable Comparison – mBB 171
2.9 MVA Input Variable Comparison – pTV 172
2.10 MVA Input Variable Comparison – MET 173
2.11 MVA Input Variable Comparison – Mtop 174
2.12 MVA Input Variable Comparison – dPhiLBmin 175
2.13 MVA Input Variable Comparison – dPhiVBB 176
2.14 MVA Input Variable Comparison – dRBB 177
2.15 MVA Input Variable Comparison – dYWH 178
2.16 MVA Input Variable Comparison – mTW 179
2.17 MVA Input Variable Comparison – pTB1 180
2.18 MVA Input Variable Comparison – pTB2 181
2.19 Correlation Matrix – Background 182
2.20 Correlation Matrix – Signal 183
2.21 MVA 150_250ptv 184
2.22 MVA TrafoD 150_250ptv 185
2.23 ROC Curve 150_250ptv 186
2.24 MVA 250ptv 187
2.25 MVA TrafoD 250ptv 188
2.26 ROC Curve 250ptv 189

3.1 $m_{b\bar{b}}$ shape variation in 0-lepton for the 2-/3-jet category. ISR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7. 191

3.2 E_T^{miss} shape variation in 0-lepton for the 2-/3-jet category. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7. 192

3.3 $m_{b\bar{b}}$ shape variation in 1-lepton for the 75GeV $< p_T^V <$ 150GeV regime. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7. 194

3.4 $m_{b\bar{b}}$ shape variation in 1-lepton for the $p_T^V >$ 150 GeV regime. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7. 195

3.5 p_T^V shape variation in 1-lepton channel, inclusive of $p_T^V >$ 75GeV. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7. 196

4.1 The pre-fit 150 GeV$< p_T^W <$ 250 GeV m_{bb} distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109]. 198

4.2 The pre-fit $p_T^W >$ 250 GeV m_{bb} distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109]. 199

4.3 The pre-fit 150 GeV$< p_T^W <$ 250 GeV p_T^W distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109]. 200

4.4 The pre-fit $p_T^W >$ 250 GeV p_T^W distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109]. 201

LIST OF TABLES

Table Page

2.1 Summary of the 19 parameters in the Lagrangian of the Standard Model (natural units are used) [22]. 20

2.2 The Higgs production cross sections for $m_H = 125$ GeV at a center-of-mass energy of 13 TeV at LHC [24]. 22

3.1 General performance goals of ATLAS detector. "$\oplus$" indicates a quadratic sum. The energy resolution of calorimeters can generally defined as $\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$, where the first term is the "stochastic" term, the second term is the "noise" term, and the third term is the "constant" term. The units for E and p_{T} are in GeV [37]. 34

4.1 A table summarizing the Monte Carlo generators used for the simulation of different processes. 58

4.2 A table summarizing the VH signal samples generated for the systematic uncertainty estimation. Note: μ_R is the renormalization scale and μ_F is the factorization scale [82–93]. 59

4.3 A table summarizing the V+jets samples generated for the systematic uncertainty estimation. 61

4.4 A table showing the heavy flavor filters used in the V+jets MC sample generation [109]. 62

4.5 A table summarizing the $t\bar{t}$ samples generated for the systematic uncertainty estimation. Note: "Var3c" is the variation of strong coupling in the initial state shower affecting the description of the $t\bar{t}$ gap fraction, the dijet decorrelation and the Z-boson transverse momentum. 63

4.6 A table summarizing the single-top samples generated for the systematic uncertainty estimation. 65

4.7 A table summarizing the diboson samples generated for the systematic uncertainty estimation. 66

5.1 The triggers used in the 0-, 1- and 2-lepton channels. 68

5.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in the VHbb analysis during the 2015-2018 data-taking period [134]. 69

5.3 Single-electron triggers used in the VHbb analysis during the 2015-2018 data-taking period [134]. 69

5.4 Single-muon triggers used in the VHbb analysis during the 2015-2018 data-taking period [134]. 69

5.5 Electron selection requirement. Note: LHLoose and LHTight are two criteria of the electron likelihood identification; FCLoose and FixedCutHigh-PtCaloOnly are the operating points of the electron isolation which have a fixed requirement either on the calorimeter or the track isolation variables (or both) [109]. 72

5.6 Muon selection requirements [109]. 73

5.7 Table caption text 75

5.8 b-jet energy corrections applied in each channel of the $VHbb$ analysis. 79

6.1 Summary of the event selection and categorization in the 0-, 1- and 2-lepton channels. Table is from the publication of the $VHbb$ analysis of ATLAS [159]. 90

6.2 Cuts defining the high and low ΔR control region [109]. 91

6.3 The split of VH signal events into signal and control regions for each channel and each p_{T}^{V} and number of jets category. Numbers are given as the percentage of total signal events in that given analysis category [109]. 92

6.4 A table summarizing the STXS regions and the corresponding reconstructed-p_{T}^{V}-regions. The current analysis is not sensitive to the regions WH, $p_{\mathrm{T}}^{W,\,\mathrm{t}} < 150$ GeV and ZH, $p_{\mathrm{T}}^{Z,\,\mathrm{t}} < 75$ GeV, and their cross-sections are fixed to the SM prediction within their theoretical uncertainties [159]. 94

7.1 The percentage of $t\bar{t}$ processes in the 2-tag 2-jet region of 1-lepton (with $p_{\mathrm{T}}^{V} >$ 150 GeV). 100

7.2 The percentage of $t\bar{t}$ processes in the 2-tag 3-jet region of 1-lepton (with $p_{\mathrm{T}}^{V} >$ 150 GeV). 100

7.3 Strategies of $t\bar{t}$ events rejection in the 1-lepton channel. 101

7.4 The cut-flow of signal and backgrounds in 1-lepton 2-jet region. 111

7.5 The cut-flow of signal and backgrounds in 1-lepton 3-jet region. 111

8.1 A table showing the regions that BDTs are trained on. 115

8.2 Variables used to train the multivariate discriminant in each lepton channel [109]. 121

8.3 Explanation of the MVA input variables in each lepton channel. Table is modified from Ref. [109]. 122

8.4 Hyper-parameters used for the BDT training. Exceptions for the 1-lepton VH and diboson training are given in brackets [109]. 124

8.5 Summary of all shape uncertainties for the $t\bar{t}$ process with short descriptions and the name of the corresponding nuisance parameters. 130

8.6 Summary of the systematic uncertainties in the background modeling for Z + jets, W + jets, $t\bar{t}$, single top-quark, and multi-jet production. 'ME' represents the matrix element generator variation and 'PS' represents the parton shower generator variation. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159]. 135

8.7 Summary of the systematic uncertainties in signal modeling. 'PS/UE' represents the parton shower/underlying event. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159]. 136

8.8 Summary of the systematic uncertainties in the background modeling for diboson production. 'PS/UE' represents the parton shower/underlying event. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. When extracting the $(W/Z)Z$ diboson production signal yield, as the normalizations are unconstrained, the normalization uncertainties are removed. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159]. 137

9.1 Factors applied to the nominal normalization of the $t\bar{t}$, $W + \mathrm{HF}$ and $Z + \mathrm{HF}$ backgrounds, as obtained from the global likelihood fit to the 13 TeV data for the nominal multivariate analysis. The errors represent the combined statistical and systematic uncertainties. 146

9.2 The Higgs boson signal, background and data yields for each signal region category in the 0- and 1-lepton channels after the full selection. The signal and background yields are normalized to the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection. 147

9.3 The Higgs boson signal, background and data yields for each signal region category in the 2-lepton channel after the full selection. The signal and background yields are normalized to the results of the global likelihood fit. The top background is derived from $e\mu$-CR data. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection. 148

9.4 Breakdown of the contributions to the uncertainty in μ_{VH}^{bb} for the VH, WH and ZH signal strength measurements. The sum in quadrature of the systematic uncertainties attached to the categories differs from the total systematic uncertainty due to correlations. Table is from the publication of the $VHbb$ analysis of ATLAS [159]. 151

9.5 Best-fit values and uncertainties for the VH, $V \to$ leptons cross-section times the $H \to b\bar{b}$ branching fraction, in the reduced STXS scheme. The SM predictions for each region, computed using the inclusive cross-section calculations and the simulated event samples are also shown. The total systematic uncertainty, equal to the difference in quadrature between the total uncertainty and the statistical uncertainty, differs from the sum in quadrature of the Th. sig., Th. bkg., and Exp. systematic uncertainties due to correlations. All leptonic decays of the V bosons (including those to τ-leptons, $\ell = e, \mu, \tau$) are considered. Table is from the publication of the $VHbb$ analysis of ATLAS [159]. 158

10.1 Estimation of the uncertainties in the signal strength measurement of the VH, WH and ZH production for Run-3. In this table, the statistical uncertainties and experimental uncertainties are divided by $\sqrt{2}$ due to the increase of the integrated luminosity from Run-2 to Run-3. The theoretical and modeling uncertainties are kept the same. The total uncertainties are the sum in quadrature of the statistical uncertainties and systematic uncertainties. The table to be compared with Run-2 is Table 9.4. 161

2.1 Background cut-flow in 2J region 164

2.2 Signal cut-flow in 2J region 164

2.3 Background cut-flow in 3J region 165

2.4 Signal cut-flow in 3J region 165

2.5 Significance (after trafoD) 186

2.6 ROC integral 186

2.7 Significance (after trafoD) 189

2.8 ROC integral 189

Simple is natural.

CHAPTER 1

Introduction

Particle physics is focused on the study of the nature of the particles that constitute matter and forces. Since the 1950s, particle physicists have been able to describe the structure of matter and the interaction of particles using a series of elegant equations which are now referred to as the Standard Model [1–3]. The Standard Model (SM) has been able to describe three of the four known fundamental forces and classify all known elementary particles. To explain why particles except the gluon and photon have mass, theories of the Standard Model were developed and a new massive scalar particle was predicted in the 1960s. To confirm the existence of this particle, many experiments were conducted from the 1980s onward.

The Large Hadron Collider constructed at CERN was aimed at the discovery of this particle and exploration of the limits of the Standard Model. After about 10 years of effort in both construction and analysis, this particle, called the Higgs boson [4–7], was discovered by ATLAS and CMS [8, 9], two general-purpose detectors at LHC, at a mass of 125 GeV in 2012. Since then, the focus has been to test the various Higgs production and decay modes predicted by the Standard Model.

The dominant decay of the Higgs boson is $H \to b\bar{b}$ which has a theoretical branching ratio of $\sim 58\%$ [10]. Due to the large contribution from multi-jet background, the associated production of the Higgs boson with a W or Z has the largest sensitivity for studying $H \to b\bar{b}$ decay[1], given that the leptons from the W or Z can efficiently trigger the signal events and remove the multi-jet background (c.f. Section 8.2.2.5). This analysis also has the best

[1] Although gluon-gluon fusion has a higher cross-section, it is difficult to be triggered by the detector.

sensitivity for measuring the WH and ZH production and can be used to constrain the decay width of the Higgs boson.

This thesis describes the observation and the final and most sensitive measurement of $H \to b\bar{b}$ in the VH production using the ATLAS detector in Run-2 of LHC. All 139 fb^{-1} of proton-proton (pp) collision data collected at a center-of-mass energy of 13 TeV from 2015 to 2018 are used in this thesis (c.f. Chapter 9). Systematic uncertainties are reduced through the updated experimental uncertainties and background modeling uncertainties (c.f. Section 8.2). Events are selected in 0-lepton, 1-lepton and 2-lepton channels, based on the number of charged leptons (electrons or muons) from the decay of associated vector boson (c.f. Chapter 6). Boosted Decision Trees (BDTs), built from the kinematic variables of the selected events, are used to maximize the sensitivity of this analysis (c.f. Section 8.1). BDT outputs are combined using a binned maximum-likelihood fit for extraction of signal strength and background normalizations (c.f. Section 8.3). The signal extraction method is validated with the dijet-mass analysis and the diboson (VZ, $Z \to b\bar{b}$) analysis (c.f. Chapter 9).

In this analysis, my contribution is on the reduction and modeling of top-pair background processes, which remained a leading source of background in this analysis even after the observation of $H \to b\bar{b}$ in 2018. The modeling of top-pair background is crucial, given that this is a systematics-dominated analysis (c.f. Section 4.2.3 and 8.2.2.2). Top quarks decay 99% of the time to W bosons, and those decay 11% of the time to tau leptons. I explored a tau-veto as a means to reduce this background and I led this effort in this analysis which has shown a great potential to reject the top-pair background (c.f. Chapter 7). In addition, I contributed to the development of the top-pair Monte Carlo filter used in this analysis, which enriches the number of top-pair events passing the selection of each channel despite the challenge to computing and storage resources within ATLAS (c.f. Section 4.2.3). My major technical contribution to the ATLAS experiment was work on the test of the FTK-to-LVL2-Interface-Crate (FLIC) board and the development of FLIC data monitoring software

(c.f. Section 3.3.1) used for the Fast TracKer project (FTK) [11] . The FTK was a hardware track finder designed for the ATLAS trigger system to reconstruct the tracks from the inner tracker. In further support of the ATLAS upgrade, I worked on the test of the Gigabit Cable Receiver ASIC version 1 (GBCRv1) and the data transmission test for the ATLAS Inner Tracker (ITk) Pixel detector (c.f. Section 3.3.2). The ATLAS ITk detector is a charged particle tracking detector under development for the High-Luminosity LHC [12].

This thesis describes the VH with $H \to b\bar{b}$ analysis, highlighting with details my contributions to this important ATLAS measurement and my work on ATLAS technical projects.

CHAPTER 2

Theoretical Framework and Higgs phenomenology

This chapter explains some of the key ideas in the development of the theoretical framework, as well as the Higgs production and decay mechanism at the Large Hadron Collider.

2.1. Theoretical Framework

Quantum field theory is a theoretical framework developed in the 20th century to explain the existence of elementary particles and 3 of the 4 fundamental interactions. It combines the view of quantum mechanics, special relativity and classical field theory. Its predictions have been validated by experiments repeatedly. Quantum field theory is a gauge theory which means the Lagrangian is invariant under local transformation (in a Lie group[1]). Maxwell's theory for electrodynamics is the earliest theory that has a gauge symmetry. To have a theoretical model explaining the strong interactions, the concept of gauge theory in Abelian groups (quantum electrodynamic theory) was extended to non-Abelian groups by Chen-Ning Yang and Robert Mills[2] in 1954 [13]. In 1961, Sheldon Glashow extended the electroweak unification models [1]. Then in 1967, the Higgs mechanism was incorporated into Glashow's electroweak interaction model by Steven Weinberg and Abdus Salam, which gives it its modern form [2, 14].

[1]In mathematics, a Lie group is a (1) smooth manifold (2) obeying the group properties and (3) whose group operations are differentiable. Lie groups play an important role in the symmetry of physical systems.

[2]Yang-Mills theory was first established in 1954 but was criticized by Wolfgang Pauli because gauge invariance requires the quanta of the Yang-Mills field be massless and it could not be explained at that time. The theory was left unnoticed until the 1960s when the concept was introduced that particles can acquire masses through the symmetry breaking process in massless theories. Later, we shall see that Yang-Mills theory naturally fits into the $SU(2) \times U(1)$ electroweak symmetry structure and the $SU(3)$ strong interaction symmetry structure.

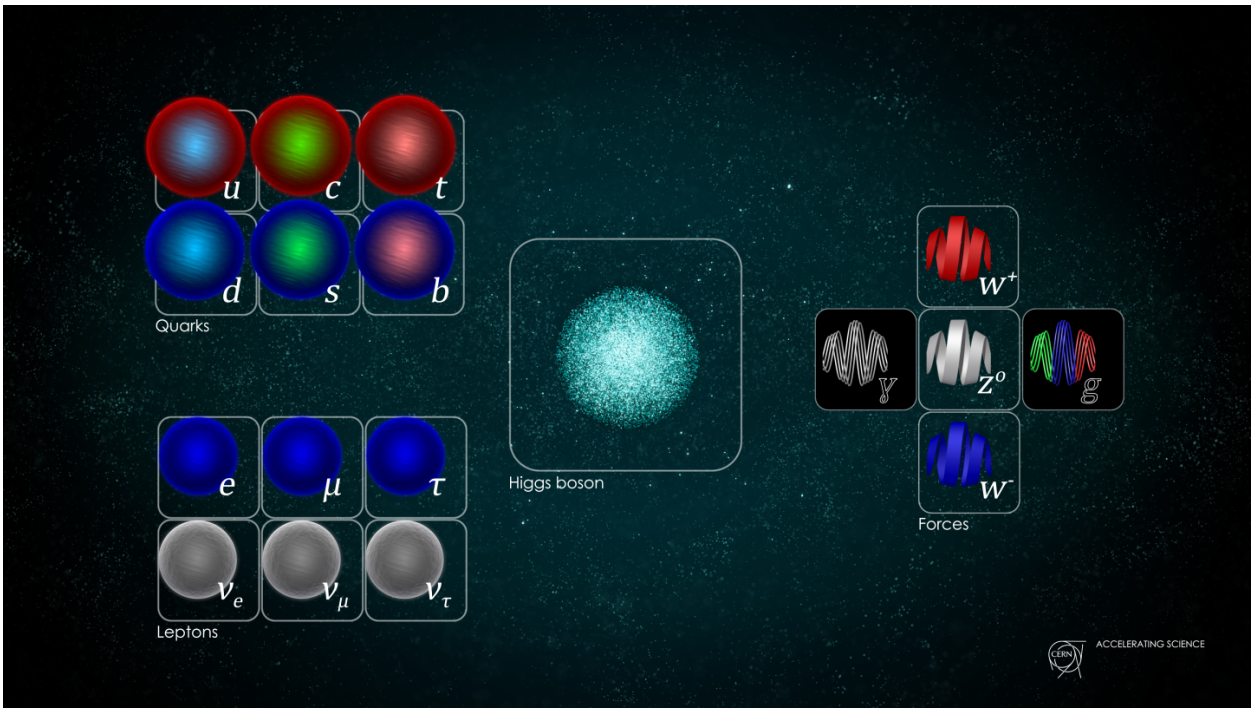


Figure 2.1: A picture of elementary particles. The Higgs boson is located in the center, with quarks in the top-left, leptons in the bottom-left and mediators in the right [15].

2.1.1. Elementary Particles

In our modern theory, evidence indicates that matter consists of particles and the interaction of these particles is achieved by exchanging virtual quanta called "mediators" (Yukawa's theory). These matter particles (which don't have inner structure) and mediator particles are assumed to be elementary (as shown in Fig. 2.1). Particles are distinguished by a few properties, such as spin[3], charge and mass. Based on spin, particles with half-integer spin are called fermions and particles with integer spin are called bosons. Based on charge, there are particles and antiparticles; from Dirac's relativistic quantum theory, there exists a corresponding antiparticle with the opposite electric charge and same mass, for every kind of particle. Particles also interact with each other. Fermions participating in strong interactions (having color charge) are called quarks; fermions only participating in electroweak interactions (having no color charge) are called leptons; bosons carrying forces are mediators, and are formally called gauge bosons; bosons with spin-0 are called scalar bosons. At this moment, there's only one fundamental scalar boson observed – the Higgs boson – which generates the masses of leptons and quarks. In addition, based on the mass, decay and interaction behavior, fermions are grouped into 3 generations. All of these will be explained in more detail in the following sections of this chapter.

2.1.2. Gauge Theory of Elementary Particles

As a successful field theory, gauge theory has been able to explain the dynamics of elementary particles. In this section, the development of gauge theory for elementary particle physics is presented.

[3]In quantum mechanics, spin is an intrinsic form of angular momentum of elementary particles [16].

2.1.2.1. (Special) Relativistic Fields

In the (special) relativistic theory, space and time coordinates are equally treated. Thus, the Euler-Lagrange equations can be written in four-dimension as

$$\partial_\mu \left(\frac{\partial L}{\partial \left(\partial_\mu \phi_i \right)} \right) = \frac{\partial L}{\partial \phi_i}, \tag{2.1}$$

where $\partial_\mu \phi$ is defined as $\partial_\mu \phi_i \equiv \dfrac{\partial \phi_i}{\partial x^\mu}$. The *Lagrangian* is a function that describes the state of a dynamic system in terms of position coordinates and their time derivatives. It is defined as the difference between the kinetic energy and potential energy of the system. Thus if the Lagrangian of a field is known, its field equation can be easily derived.

2.1.2.1.1 Electromagnetic Field

Starting from an easy one, the Lagrangian for an electromagnetic field with source $j^\mu \equiv (\rho_{\mathrm{em}}, \boldsymbol{j}_{em})$ is generally shown as

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\nu A_\nu, \tag{2.2}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. With the Euler-Lagrangian equations, it yields

$$\partial_\mu F^{\mu\nu} = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu, \tag{2.3}$$

where $\partial^2 = \partial_\mu \partial^\mu = \partial_t^2 - \nabla^2$, which is the Maxwell equation for the electromagnetic potential[4] A^ν. A^ν describes a massless vector field and it is the field of photon if quantized in the gauge theory[5].

[4]If let $\partial_\mu j^\mu = 0$, it becomes the current continuity equation. The charge is conserved.

[5]For the gauge transformation and quantization of A^ν, refer to Ref. [17].

2.1.2.1.2 A Scalar (Spin-0) Field

The Lagrangian for a spin-0 scalar field can be constructed as

$$L = \frac{1}{2}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - \frac{1}{2} m^2 \phi^2. \tag{2.4}$$

When applying the Euler-Lagrange equation, we obtain

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{2.5}$$

This is so called Klein-Gordon equation. The solution ϕ is a real scalar field and corresponds to a particle of spin-0 and mass m. As we shall see later, the Higgs field is a scalar field and satisfies the Klein-Gordon equation.

2.1.2.1.3 A Spinor (Spin-$\frac{1}{2}$) Field

A spinor field describes a particle with half-integer spin. The Lagrangian for a spin-$\frac{1}{2}$ field can be shown to be

$$L = \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi. \tag{2.6}$$

Here ψ and $\overline{\psi}$ are treated independently. Applying the Euler-Lagrange equation to $\overline{\psi}$, we obtain the Dirac equation[6]

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi = 0, \tag{2.7}$$

which of the solution describes the field of a particle of spin-$\frac{1}{2}$ and mass m. Fermion fields satisfy the Dirac equation.

[6] If the Euler-Lagrange equation is applied to ψ, it yields the adjoint of the Dirac equation $\overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right) = 0$, where $\overline{\psi} \equiv \psi^\dagger \gamma^0$. γ^μ are Pauli matrices.

2.1.2.1.4 A Vector (Spin-1) Field

If we have a four-vector field A^μ, we can construct the Lagrangian

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu. \tag{2.8}$$

The Euler-Lagrange equation yields

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu + m^2 A^\nu = 0. \tag{2.9}$$

This is Proca equation the solution of which describes the field of a particle of spin-1 and mass m. If we set $m = 0$, we obtain the electromagnetic field equation in empty space – Maxwell's equations. We'll see later that the weak interaction fields of Z and $W^\pm$ satisfy the Proca equation (and are not massless).

2.1.2.1.5 Summary

In relativistic field theory, the field equation of a physics system is obtained from its Lagrangian. The Lagrangian of a particular system is not unique; constants and terms can be added to or multiplied by the Lagrangian (as far as terms cancels out after applying the Euler-Lagrange equation)[7]. For the field with a source (e.g. charge), similar to the electromagnetic field, the general spin-0, spin-$\frac{1}{2}$ and spin-1 field follows the current conservation law [18].

[7]The solutions to the Euler-Lagrange equation describe the evolution of the physical system. Thus, parameters which cancels out after applying the Euler-Lagrange equation don't affect the physics results.

2.1.2.2. Yang-Mills Theory and Chromodynamics

Gauge theory demands a local transformation invariance (gauge invariance) of the field by its definition[8]. Unfortunately, the Lagrangian of Dirac equation is invariant under a global phase transformation, but not a local phase transformation. Thus, to expand the gauge theory (to electro-weak theory and chromodynamics), we have to convert a globally invariant Lagrangian into a locally invariant one. One key milestone is the Yang-Mills Theory. It was started from the symmetry group of $SU(2)$[9] and extended to $SU(3)$, the symmetry group of quantum chromodynamics (QCD).

2.1.2.2.1 Yang-Mills Theory

If we construct the fields of two spinors, ψ_1 and ψ_2 (with identical mass) without considering interactions, the Lagrangian is

$$L = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.10}$$

where we can combine ψ_1 and ψ_2 into a two-component column vector (the same logic applies to the adjoint spinors $\overline{\psi_1}$ and $\overline{\psi_2}$)

$$\psi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}; \quad \overline{\psi} = \begin{pmatrix} \overline{\psi_1} & \overline{\psi_2} \end{pmatrix}. \tag{2.11}$$

[8]Gauge theory requires that its Lagrangian does not change (is invariant) under local transformations so that the underlying physics does not change after the transformation. Local transformation invariance is a stricter requirement than global transformation invariance [17].

[9]The $SU(2)$ group was initially introduced to explain the isotopic spin and isotopic gauge invariance. Later, it was proven to be more successful when used to develop the electroweak unification and quantum chromodynamics. The symmetry group of the electroweak interactions is $SU(2) \times U(1)$. The symmetry group of the electromagnetic interactions, $U(1)$, is a subgroup of $SU(2) \times U(1)$.

Equation 2.10 then represents the sum of two Dirac Lagrangian. Since ψ_1 and ψ_2 have global invariance, L admits a global invariance as well. Thus, the transformation on ψ can be expressed as

$$\psi(x) \to \psi'(x) = e^{i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}/2}\psi(x), \tag{2.12}$$

where parameters $\boldsymbol{\alpha}$ are independent of x (that's why it is called a global transformation)[10].

With the global transformation invariance, the choice of which two base states to use is a matter of convention. However, the choice cannot be made independently at all space-time points, only globally (at all space-time points). This seems to be an unpleasant limitation and suggested an inconsistency with the localized field concept, which motivated Yang and Mills to replace this global (space-time independent) phase transformation by the local (space-time dependent) one[11]

$$\psi(x) \to \psi'(x) = e^{ig\boldsymbol{\tau}\cdot\boldsymbol{\alpha}(x)/2}\ \psi(x), \tag{2.13}$$

in which the phase parameter $\boldsymbol{\alpha}(x)$ are functions of $x^\mu \equiv (t, \boldsymbol{x})$ [19].

To make the Lagrangian invariant under local transformation, the derivative of ψ transforms as

$$\partial_\mu\psi(x) \to \partial_\mu\psi'(x) = e^{ig\boldsymbol{\tau}\cdot\boldsymbol{\alpha}(x)/2}\ \partial_\mu\psi(x) + i\frac{g}{2}\boldsymbol{\tau}\cdot\partial_\mu\boldsymbol{\alpha}(x)\ e^{ig\boldsymbol{\tau}\cdot\boldsymbol{\alpha}(x)/2}\ \psi(x), \tag{2.14}$$

Then, to cancel the second offending term, we can define the covariant derivative $D_\mu \equiv \partial_\mu + i\frac{g}{2}\boldsymbol{\tau}\cdot\boldsymbol{A_\mu}$, and assign a transformation rule to the gauge fields $\boldsymbol{A^\mu} = (A_1^\mu,\ A_2^\mu,\ A_3^\mu)$ such that

$$D_\mu\psi(x) \to e^{ig\boldsymbol{\tau}\cdot\boldsymbol{\alpha}(x)/2}\ D_\mu\psi(x). \tag{2.15}$$

[10]The three matrices $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ are Hermitian Pauli matrices. They are labeled as $\boldsymbol{\tau}$ to distinguish them from the mathematically identical $\boldsymbol{\sigma}$ matrices which are associated with the real spin degree of freedom [19].

[11]g is inserted in the exponential to make this analogous to the electromagnetic $U(1)$ case, $\psi(x) \to e^{iq\chi(x)}\psi(x)$. g will be the coupling strength [19].

Then the Lagrangian will be invariant. By using the infinitesimal local $SU(2)$ transformation with parameter $\boldsymbol{\epsilon}(x)$[12], the transformation rule for $\boldsymbol{A^\mu}$ [19] can be deduced as

$$\begin{aligned} \boldsymbol{A}^\mu \to \boldsymbol{A}'^\mu &= \boldsymbol{A}^\mu + \delta \boldsymbol{A}^\mu \\ &= \boldsymbol{A}^\mu - \partial^\mu \boldsymbol{\epsilon}(x) - g[\boldsymbol{\epsilon}(x) \times \boldsymbol{A}^\mu]. \end{aligned} \tag{2.16}$$

Therefore, the new Lagrangian is

$$\begin{aligned} L &= \overline{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi \\ &= \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi - g\overline{\psi}\gamma^\mu \boldsymbol{\tau} \cdot \boldsymbol{A}_\mu \psi. \end{aligned} \tag{2.17}$$

It is invariant under the local transformation. Lastly, due to the introduction of three new vector fields $\boldsymbol{A}^\mu$, their free Lagrangian $L_A = -\frac{1}{4}\boldsymbol{F}\mu\nu \cdot \boldsymbol{F}^{\mu\nu}$ has to be added, which results in the complete Lagrangian

$$L = \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi - g\overline{\psi}\gamma^\mu \boldsymbol{\tau} \cdot \boldsymbol{A}_\mu \psi - \frac{1}{4}\boldsymbol{F}\mu\nu \cdot \boldsymbol{F}^{\mu\nu}. \tag{2.18}$$

Now it describes the Dirac fields of two equal mass particles interacting via three massless vector gauge fields.

Yang-Mills theory was initiated to provide an explanation for a different phenomenology. But it is proven to be greatly successful when applied to the development of electro-weak unification and quantum chromodynamics after the introduction of the idea that particles acquire mass through symmetry breaking (which is described in Sec. 2.1.2.3).

[12] Since this infinitesimal transformation is local, $\boldsymbol{\epsilon}(x)$ are functions of x.

2.1.2.2.2 Chromodynamics

Experimental evidence suggested that quarks are subject to a force with three kinds of charge. These are named "red," "green," and "blue," since all observed quark matter states involve combinations of charge that are overall neutral ("colorless"). Although various flavors of quark[13] carry different masses, the three colors of the same flavor are all supposed to have the same mass. Thus the free Lagrangian for the same flavor can be written in the same format with Eq. 2.10,

$$L = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.19}$$

but with

$$\psi \equiv \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}, \overline{\psi} = \begin{pmatrix} \overline{\psi_r} & \overline{\psi_b} & \psi_g \end{pmatrix}. \tag{2.20}$$

Following the procedure explained above for the Yang-Mills Lagrangian, we obtain

$$L = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi - g_s\overline{\psi}\gamma^{\mu}\boldsymbol{\lambda}\cdot\boldsymbol{A}_{\mu}\psi - \frac{1}{4}\boldsymbol{F}\mu\nu\cdot\boldsymbol{F}^{\mu\nu}. \tag{2.21}$$

where gauge fields $\boldsymbol{A}_{\mu}$ has eight terms[14]. It describes three equal-mass Dirac fields (three possible color charge states of a given flavor) interacting via eight massless vector gauge fields (gluons). L now has local $SU(3)$ gauge transformation invariance. The eight color currents of the Dirac fields can be obtained as

$$\boldsymbol{j}^{\mu} \equiv g_s\left(\overline{\psi}\gamma^{\mu}\boldsymbol{\lambda}\psi\right). \tag{2.22}$$

[13] A flavor is a specie of elementary particles. Six flavors of quarks are discovered so far.

[14] $\boldsymbol{\lambda}$ (named Gell-Mann matrices) are eight 3×3 traceless Hermitian matrices used to study the strong interactions. g_s is the strong interaction coupling constant.

They act as the source of the color fields $\boldsymbol{A}^{\mu}$, similar to the case where the electric currents act as the source of the electromagnetic fields. This description of quark interactions via the gluon field is consistent with all experimental evidence.

2.1.2.3. Spontaneous Symmetry-breaking and the Higgs Mechanism

The prompt from global transformation invariance to local transformation invariance works great with the electromagnetic and strong interactions. But to use it to explain the weak interactions, the fact that gauge fields are required to be massless in the above approaches, becomes an obstacle[15]. Spontaneous symmetry-breaking and the Higgs mechanism gracefully solve this problem.

2.1.2.3.1 Spontaneous Symmetry-breaking

Spontaneous symmetry-breaking is a process by which the symmetry of a physical system at normal states is broken at its ground/vacuum states. To illustrate the idea of spontaneous symmetry breaking, let's take a system with global U(1) symmetry as an example[16].

The Lagrangian of a complex scalar field system can be written as

$$L = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \tag{2.23}$$

where $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, $\phi^\dagger = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)$ and $V(\phi) = \mu^2(\phi^\dagger\phi) + \frac{\lambda}{4}(\phi^\dagger\phi)^2$. L is invariant at global U(1) transformation $\phi \to e^{-i\alpha}\phi$. If we further rewrite $V(\phi)$ as $V(\phi) = \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{\lambda}{16}(\phi_1^2 + \phi_2^2)^2$, the potential can be shown as in Fig. 2.2. Clearly, the minimum of $V(\phi)$

[15]The mass term in Proca Lagrangian is not gauge invariant. Although photon and gluons are massless, but that's not the case for W and Z bosons.

[16]More details about the spontaneously broken local U(1) symmetry and local SU(2) $\times$ U(1) symmetry, refer to Ref. [19].

is uniquely at $\phi = 0$. When we quantize this model (usually treat small oscillations of the field about the minimum as approximately harmonic), we find that there are two degrees of freedom with the same mass μ (degeneracy) at the ground/vacuum states.

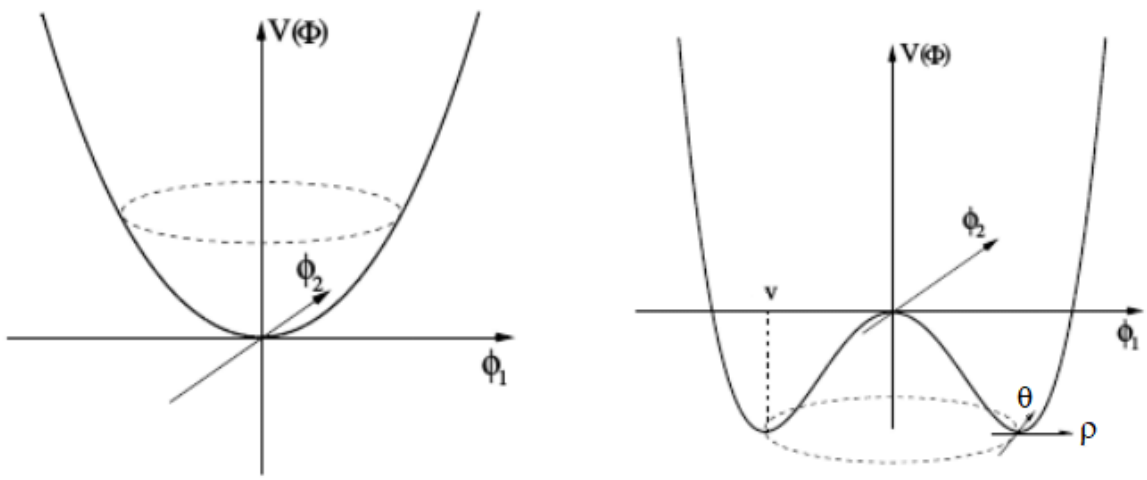


Figure 2.2: (left) Potential $V(\phi) = \mu^2(\phi^\dagger\phi) + \frac{\lambda}{4}(\phi^\dagger\phi)^2$ with symmetry at ground states; (right) potential $V(\phi) = -\mu^2(\phi^\dagger\phi) + \frac{\lambda}{4}(\phi^\dagger\phi)^2$, where the symmetry is broken.

We can break the symmetry by changing the sign of μ^2 in $V(\phi)$ (Goldstone model, see Ref. [19]). The Lagrangian becomes

$$L = (\partial_\mu\phi)^\dagger (\partial^\mu\phi) + \mu^2 \left(\phi^\dagger\phi\right) - \frac{\lambda}{4}\left(\phi^\dagger\phi\right)^2, \tag{2.24}$$

which still has global U(1) symmetry but the ground state is shifted (see Fig. 2.2). The minimum occurs at $(\phi^\dagger\phi) = \frac{1}{2}(\phi_1^2 + \phi_2^2) = \frac{2\mu^2}{\lambda}$. By defining $v \equiv \frac{2|\mu|}{\sqrt{\lambda}}$, we can obtain $|\phi| = \frac{v}{\sqrt{2}}$.

In order to quantize this model, we need to expand $\phi(x)$ in normal modes about the minimum. For a more clear picture, $\phi(x)$ can be rewritten in the polar coordinates at each x,

$$\phi(x) = \frac{1}{\sqrt{2}}\,\rho(x)\,e^{i\theta(x)/v}, \tag{2.25}$$

where $\rho(x)$ gives the radial degree of freedom and $\theta(x)$ gives the cylindrical degree of freedom[17]. Since the potential is at the minimum at $\rho = v$ for any θ, we can choose the starting vacuum as $\rho = v$ and $\theta = 0$, and rewrite $\rho(x)$ as $\rho(x) = v + h(x)$ with $h = 0$ at the minimum. Therefore, after promoting $\theta(x)$ and $h(x)$ to fields, Lagrangian becomes

$$L = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{2}(\partial_\mu \theta)(\partial^\mu \theta) - \mu^2 h^2 + \frac{\mu^4}{\lambda} + \text{cubic and quartic terms}, \tag{2.26}$$

where the potential is $V = \mu^2 h^2 - \dfrac{\mu^4}{\lambda}$ + cubic and quartic terms. This shows a different particle spectrum – the θ-mode is massless (Goldstone boson) and the h-mode has a mass of $\sqrt{2}\mu$. The vacuum expectation value of the field is $\dfrac{v}{\sqrt{2}}$, which is not invariant under the symmetry group. U(1) symmetry is no longer visible in the spectrum[18] [20].

2.1.2.3.2 The Higgs Mechanism

The spontaneous symmetry-breaking above is triggered by setting $\mu^2 \to -\mu^2$ by hand. In this section, we will see how the Higgs mechanism provides a simple model for investigating what happens when a gauge symmetry is broken. To explain it, let's go through the breaking process of the local U(1) symmetry.

[17] v is included so that θ has the same dimension (mass) as ρ and ϕ.

[18] θ and h are orthogonal to the previous degenerate ground states.

The Lagrangian of a scalar field with global U(1) symmetry has been shown in Eq. 2.23. To make it invariant under local phase transformation,

$$\phi(x) \to \phi'(x) = e^{-i\alpha(x)}\phi(x), \tag{2.27}$$

we can replace ∂_μ by covariant derivative D_μ ($D_\mu \equiv \partial_\mu + iqA_\mu$), which then requires the potential transforms as

$$A^\mu \to A'^\mu = A^\mu + \frac{1}{q}\partial^\mu \alpha(x). \tag{2.28}$$

A free Lagrangian term of A^μ ($L_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$) has also to be added for the gauge transformation, similar to Sec. 2.1.2.2.1. Thus, all these result in

$$L = (D_\mu\phi)^\dagger (D^\mu\phi) - \mu^2 \left(\phi^\dagger\phi\right) - \frac{1}{4}\lambda \left(\phi^\dagger\phi\right)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.29}$$

If we break the symmetry by taking $\mu^2 \to -\mu^2$, then $\phi(x)$ can be rewritten in the polar coordinates as

$$\phi(x) = \frac{1}{\sqrt{2}}\,\rho(x)\,e^{i\theta(x)/v} = \frac{1}{\sqrt{2}}\,(v + h(x))\,e^{i\theta(x)/v}, \tag{2.30}$$

but now with local U(1) symmetry transformation. In this case, the phase of $\phi(x)$ is completely arbitrary, since any change in α of Eq. 2.27 can be compensated by an appropriate transformation on A^μ in Eq. 2.28, and L remains the same. It means that the θ field which serves as the phase of $\phi(x)$ in Eq. 2.30, can be made vanish if we choose a special gauge for A^μ. Since the degree of freedom of θ cannot simply disappear, to track where this θ field go, we can look at the field equation for A^ν[19]

$$\partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu. \tag{2.31}$$

[19] j^ν is the electromagnetic current (similar to Eq. 2.3).

The electromagnetic current can also be expressed as $j^\nu = iq(\phi^\dagger \partial^\nu \phi - (\partial^\nu \phi^\dagger)\phi) - 2q^2 A^\nu \phi^\dagger \phi$[20], and $\phi(x)$ is parameterized in the polar coordinate (Eq. 2.30). Thus, it leads to

$$j^\nu = -v^2 q^2 (A^\nu - \frac{\partial^\nu \theta}{vq}) + \text{quadratic and cubic terms}, \tag{2.32}$$

where the higher order terms represent interactions. If bringing in Eq. 2.31 and only the linear terms in Eq. 2.32 is retained, we obtain

$$\partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = -v^2 q^2 (A^\nu - \frac{\partial^\nu \theta}{vq}). \tag{2.33}$$

The left side of the equation is invariant under the transformation of A^μ in Eq. 2.28. So if we take the definition

$$A'^\nu = A^\nu - \frac{\partial^\nu \theta}{vq}, \tag{2.34}$$

then the equation for A'^ν will be

$$\partial^2 A'^\nu - \partial^\nu \partial_\mu A'^\mu = -v^2 q^2 A'^\nu \quad \text{or} \quad (\partial^2 + v^2 q^2) A'^\nu - \partial^\nu \partial_\mu A'^\mu = 0, \tag{2.35}$$

which is the Proca equation describing a free vector field of mass vq. Thus, that special gauge transformation for A^μ is found (that is Eq. 2.34). Under this gauge transformation, the Goldstone field θ is incorporated to the massive vector field A'^ν after being "swallowed" by the massless gauge field A^ν. We can also obtain $\alpha(x) = -\theta(x)/v$, after comparing Eq. 2.34 with Eq. 2.28, which leads to that ϕ is real $\phi = \frac{1}{\sqrt{2}}(v + h)$. Then, the final Lagrangian is

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2}(q^2 v^2) A_\mu A^\mu + \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \mu^2 h^2 + \text{interactions}, \tag{2.36}$$

[20] This current is obtained through in $j^\nu = -\frac{\partial L}{\partial A_\nu}$.

which describes a spin-1 field A^μ of mass vq and a scalar field h of mass $\sqrt{2}\mu$. h is also known as the Higgs field.

2.1.2.4. The Lagrangian of a Theoretical Model

The development of the gauge theory results in many theoretical models. Eq. 2.37 shows the Lagrangian of a model[21] which is currently used to describe 3 of the 4 interactions, and commonly referred to as the "Standard Model" [21]. It is observed to be theoretically self-consistent and has demonstrated great success in providing testable experimental predictions, although it is not a total theory of nature and in at least one place is demonstrably incomplete (the mass of Fermions called "neutrinos").

$$
\begin{aligned}
L \;=\; & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{8}tr\left(\boldsymbol{W}_{\mu\nu}\boldsymbol{W}^{\mu\nu}\right) - \frac{1}{2}tr\left(\boldsymbol{G}_{\mu\nu}\boldsymbol{G}^{\mu\nu}\right) && \text{U(1), SU(2) and SU(3) gauge terms} \\
& + \left(\overline{\nu}_L, \overline{e}_L\right)\widetilde{\sigma}^\mu iD_\mu \begin{pmatrix}\nu_L \\ e_L\end{pmatrix} + \overline{e}_R\sigma^\mu iD_\mu e_R + \overline{\nu}_R\sigma^\mu iD_\mu\nu_R + (h.c.) && \text{lepton dynamical term} \\
& -\frac{\sqrt{2}}{v}\left[\left(\overline{\nu}_L, \overline{e}_L\right)\phi M^e e_R + \overline{e}_R\overline{M}^e\overline{\phi}\begin{pmatrix}\nu_L \\ e_L\end{pmatrix}\right] && \text{electron, muon, tau mass term} \\
& -\frac{\sqrt{2}}{v}\left[\left(-\overline{e}_L, \overline{\nu}_L\right)\phi^* M^\nu\nu_R + \overline{\nu}_R\overline{M}^\nu\phi^T\begin{pmatrix}-e_L \\ \nu_L\end{pmatrix}\right] && \text{neutrino mass term} \\
& + \left(\overline{u}_L, \overline{d}_L\right)\widetilde{\sigma}^\mu iD_\mu \begin{pmatrix}u_L \\ d_L\end{pmatrix} + \overline{u}_R\sigma^\mu iD_\mu u_R + \overline{d}_R\sigma^\mu iD_\mu d_R + (h.c.) && \text{quark dynamical term} \\
& -\frac{\sqrt{2}}{v}\left[\left(\overline{u}_L, \overline{d}_L\right)\phi M^d d_R + \overline{d}_R\overline{M}^d\overline{\phi}\begin{pmatrix}u_L \\ d_L\end{pmatrix}\right] && \text{down, strange, bottom mass term} \\
& -\frac{\sqrt{2}}{v}\left[\left(-\overline{d}_L, \overline{u}_L\right)\phi^* M^u u_R + \overline{u}_R\overline{M}^u\phi^T\begin{pmatrix}-d_L \\ u_L\end{pmatrix}\right] && \text{up, charm, top mass term} \\
& + \left(\overline{D_\mu\phi}\right)D^\mu\phi - m_h^2\left[\overline{\phi}\phi - \frac{v^2}{2}\right]^2\frac{1}{2v^2} && \text{Higgs dynamical and mass term}
\end{aligned}
\tag{2.37}
$$

[21] In the Lagrangian of the Standard Model, $(h.c.)$ represents the Hermitian conjugate of the previous terms. $\overline{\psi} = (h.c.)\psi = \psi^\dagger = \psi^{*T}$; ϕ is a complex Higgs field with 2 components. More about the explanation on the Standard Model Lagrangian can be found at Ref. [21].

The Standard Model has 19 parameters (listed in Table 2.1 along with their currently established values) whose values are not predicted and must therefore be determined by experiment; this allows for self-consistency checks of the Standard Model, as multiple independent measurements can constrain the same parameter and thus offer the possibility of observing discrepancies. One of the goals in the analysis of this thesis is to measure the rate at which the Higgs boson decays to a pair of bottom quarks, a test of the Standard Model's predictions of its Yukawa coupling to fermions.

Description	Parameter	Input value from experiment
Mass of electron	m_e	510.9989461(31) keV
Mass of muon	m_μ	105.6583745(24) MeV
Mass of tau	m_τ	1.77686(12) GeV
Mass of up quark	m_u	2.2 MeV
Mass of down quark	m_d	4.7 MeV
Mass of strange quark	m_s	95 MeV
Mass of charm quark	m_c	1.275 GeV
Mass of bottom quark	m_b	4.18 GeV
Mass of top quark	m_t	173.0 GeV
CKM 12-mixing angle	θ_{12}	13.1°
CKM 23-mixing angle	θ_{23}	2.4°
CKM 13-mixing angle	θ_{13}	0.2°
CKM CP-violating phase	δ	0.995
Gauge coupling for U(1)	g_1 or g'	0.357
Gauge coupling for SU(2)	g_2 or g	0.652
Gauge coupling for SU(3)	g_3 or g_s	1.221
QCD vacuum angle	θ_{QCD}	~ 0
Higgs vacuum expectation value	v	246.2196(2) GeV
Mass of Higgs	m_H	125.18 GeV

Table 2.1: Summary of the 19 parameters in the Lagrangian of the Standard Model (natural units are used) [22].

2.2. Higgs Phenomenology

The behavior of the Higgs boson can be predicted based on the theoretical framework described above. In this section, the production and decay modes of the Higgs boson at LHC is presented.

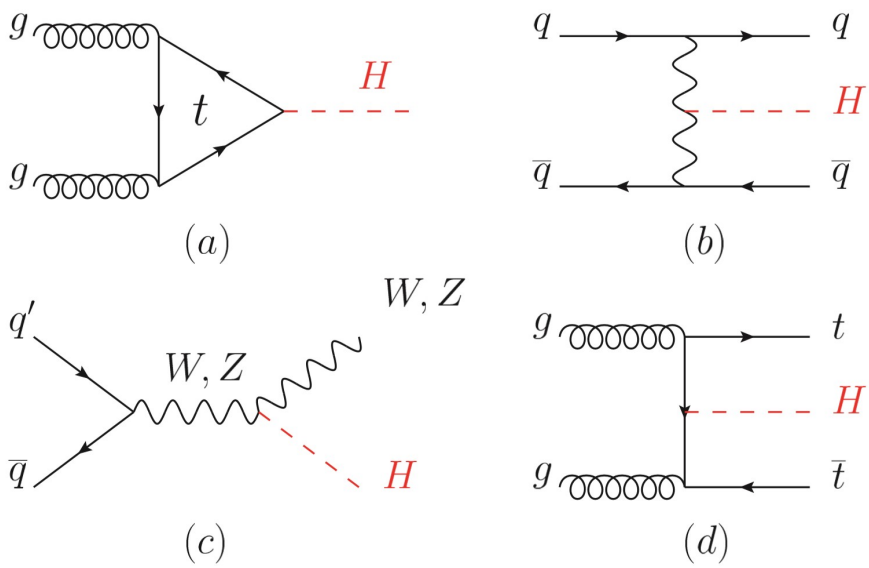


Figure 2.3: Diagram of the Higgs production at LHC: (a) gluon-gluon fusion, (b) vector-boson fusion, (c) associated production with a vector boson, and (d) associated production with top quarks [23].

2.2.1. Higgs Production at LHC

In the proton-proton collision machine, the Higgs boson has four main production modes used in the measurements: gluon-gluon fusion (ggF), vector-boson fusion (VBF), associated production with a vector boson (VH), and associated production with top quarks (ttH). Their Feynman diagrams are shown in Fig. 2.3.

In the ggF process, gluons don't couple directly to the Higgs boson, but through a loop where virtual quarks are exchanged. Since the coupling of the Higgs boson to fermions is

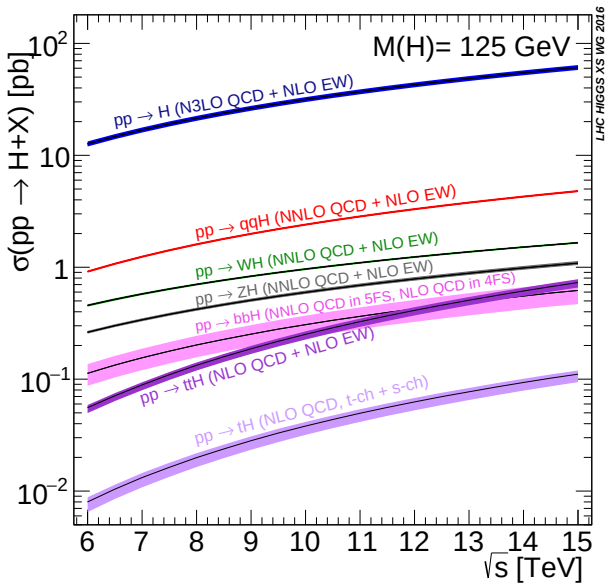


Figure 2.4: Cross sections of the Higgs boson production as a function of the center-of-mass energy for proton-proton collisions. The theoretical uncertainties are represented as bands [24].

Higgs Production	**Cross Section (pb)**
ggF	$48.6^{+5\%}_{-5\%}$
VBF	$3.78^{+2\%}_{-2\%}$
WH	$1.37^{+2\%}_{-2\%}$
ZH	$0.88^{+5\%}_{-5\%}$
$t\bar{t}H$	$0.50^{+9\%}_{-13\%}$

Table 2.2: The Higgs production cross sections for $m_H = 125$ GeV at a center-of-mass energy of 13 TeV at LHC [24].

proportional to the particle's mass[22], the loop process is more likely for heavy quarks (top or bottom quarks). ggF dominates the Higgs production at LHC (see Fig. 2.4 and Table 2.2).

In the VBF process, two fermions collide in exchange of a W or Z boson, and the Higgs boson radiates from this vector boson. The scattered quarks result into two hard jets in the forward regions of the detector. VBF is the second largest cross-section process in the Higgs production at LHC.

In the VH production, a fermion and an antifermion collide and form a virtual W or Z boson. When this virtual boson carries enough energy, it can radiate a Higgs boson. The VH process has the third largest cross-section at LHC. The full kinematic information of the decaying products (leptons) provided by the VH production can further help suppress the large QCD backgrounds[23].

In the ttH process, there are two gluons colliding, with each decaying to a top-antitop quark pair. A top quark and an antitop quark from each pair form a Higgs boson together. The ttH process has the fourth largest cross-section.

The VH production mode provides a clean environment for studying $H \to b\bar{b}$ by using the leptons from the associated W or Z boson for triggering. This thesis is focused on the study of Higgs through the VH production.

2.2.2. Higgs Decay Channels

The previous measurements on the Higgs mass give a result of $m_H = 125.18 \pm 0.16$ GeV [25]. With the Higgs mass m_H treated as an input to the theoretical model, the

[22]The coupling of the Higgs boson to bosons (mediators) is proportional to the square of the particle's mass.

[23]QCD background refers to the large number of hadronic jets initiated by quarks and gluons at the LHC.

prediction on the branching ratio of the Higgs decay channels can be made, as shown in Fig. 2.5.

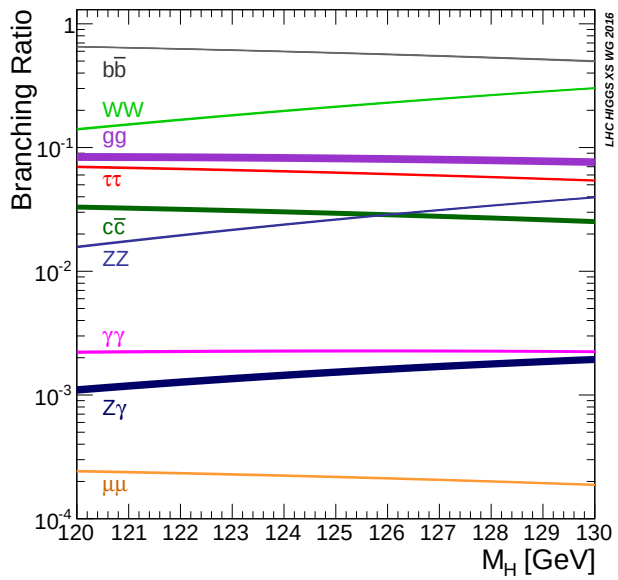


Figure 2.5: The branching ratios of the Higgs boson decays near $m_H = 125$ GeV . The theoretical uncertainties are represented as bands [24].

Different decay products of the Higgs boson result in different features in the decay channel. For example, the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ are known to be golden channels. Since all particles in their final state can be well reconstructed, m_H can be measured with excellent resolution. The $H \to W^+W^- \to \ell^+\nu_\ell\ell^-\bar{\nu}_\ell$ and $H \to \tau^+\tau^-$ suffer from the energy loss due to the presence of neutrinos and large backgrounds; thus, their m_H resolutions are relatively poor.

The $H \to b\bar{b}$ has the largest branching ratio among all decay channels. It is mostly measured through the VH production where the leptons from the W or Z are used for triggering the signal events and rejecting the multi-jet backgrounds[24]. The W boson can

[24]The multi-jet background arises from the pp collision where a large number of jets are generated.

be reconstructed from its leptonic decay $W \to \ell\nu$ ($\ell = e$ or μ[25]), while the Z boson can be reconstructed from the decay of $Z \to e^+e^-$, $\mu^+\mu^-$ or $\nu\bar{\nu}$. The Higgs boson is reconstructed from the two b-tagged jets. Challenges on measuring $H \to b\bar{b}$ remain in the b-tagging efficiency, b-jet momentum and energy resolution, and the background estimation (modeling).

Details of the measurement on $H \to b\bar{b}$ with the VH production mode are presented in the following chapters of this thesis.

[25]The case that W boson decaying to τ lepton is not considered in the analysis of this thesis.

CHAPTER 3

The LHC and ATLAS Detector

This chapter briefly describes the Large Hadron Collider and the ATLAS detector. My contribution to the ATLAS upgrade is presented at the end of this chapter.

3.1. Large Hadron Collider (LHC)

The Large Hadron Collider [26] is a proton-proton circular collider located at the border of Switzerland and France near Geneva. It's an underground ring with a circumference of 27 km. It currently delivers proton-proton collisions at a center-of-mass of 13 TeV, which is the highest energy available at any facility in the world.

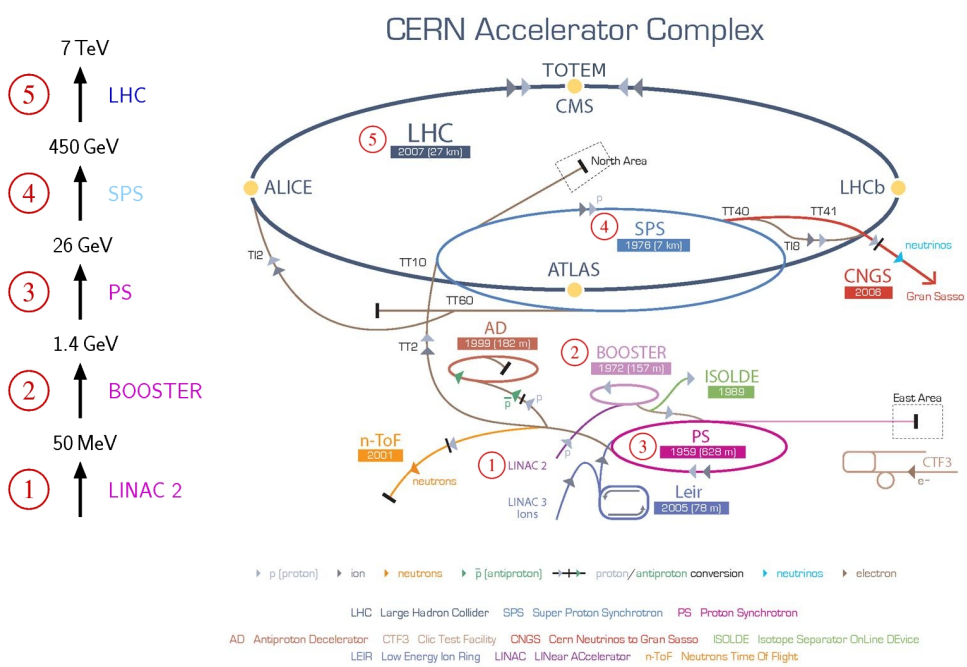


Figure 3.1: Facilities for the accelerator at CERN [27].

In the LHC, a proton is accelerated through several stages (see Fig. 3.1). First, it is accelerated to $0.3c$ (c is the speed of light) using a LINAC – a linear accelerator. Then, by going through the SPS (Super Proton Synchrotron), the proton is further accelerated to $0.8c$ and is then injected to the larger ring — the LHC. In the LHC, the proton is accelerated to $0.999999991c$ at 7 TeV, at which point the protons are ready to collide.

The acceleration in LHC is done using radio-frequency (RF) waves, similar to other circular accelerators [28]. The RF wave is tuned at a certain frequency so that the protons inside the LHC tunnel are accelerated along the wave troughs. The waves are designed to be at 400 MHz (2.5 ns) and protons are filled at every 10th trough, which leads to the pp collisions at roughly every 25 ns.

The LHC is a two-ring accelerator, which means that the two proton beams[1] that collide with each other are accelerated in separate rings. This is because the collision particles are proton-proton and the same bending magnets and RF structure cannot be used to accelerate both beams. Proton-proton collisions are used[2] because it's easier to achieve the luminosity requirement[3].

The purpose of building the LHC is mainly for the confirmation of the Higgs boson in the Standard Model and searching for physics beyond the Standard Model [29]. These have been leading questions that most physicists have been concerned with since the 1980s. The current LHC physics program is mainly focused on the precision measurement of the Higgs physics and Standard Model, as well as searches for evidence of physics beyond the Standard Model [30].

[1] A beam in the LHC is not a continuous string of particles but is divided into chunks a few centimeters long. Each chunk is called a bunch. The bunches cross at the collision point every 25 ns.

[2] Antiprotons are more difficult to produce in the large quantities needed for the LHC.

[3] The (instantaneous) luminosity is defined as $\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt}$, which is the number of events at a certain time interval divided by the interaction cross-section.

There are several large experiments on the LHC ring: ATLAS [31], CMS [32], ALICE [33], and LHCb [34]. ATLAS and CMS are two general-purpose detectors, of which the main goal was to discover and study the Higgs boson. ALICE is a heavy-ion experiment and LHCb is focused on B physics (studying the properties of B hadrons — hadrons containing at least one bottom quark). ATLAS and CMS, both aiming at a peak luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$, are located at Point 1 and Point 5 (referring to the interaction points of the counter-circulating proton beams intersecting with each other in the LHC, see Fig. 3.1). LHCb is at Point 8 where the luminosity is $\mathcal{L} = 10^{32}\text{cm}^{-2}\text{s}^{-1}$. ALICE is at Point 2 and only runs when the LHC conducts heavy-ion collisions instead of proton collisions [26].

The design and construction of the LHC have proven to be excellent, and many of the physicists and engineers devoted a lot of effort and overcame many difficulties to make this possible [35]. Some of the design features are summarized below.

3.1.1. LHC Magnet

Space limitations in the tunnel and the need to restrain costs have led to the adoption of the "two-in-one" design for almost all the superconducting magnets of LHC. This design accommodates the windings for the two proton beams in a common cold mass and cryostat, and the magnetic flux circulates in the opposite direction through the two beams (see Fig. 3.2). Of course, this makes the magnet structure complicated.

3.1.1.1. Dipole and Quadrupole

Bending the proton beam is achieved through the magnetic field of dipole magnets. The LHC ring accommodates 1232 main dipoles. The dipolar magnetic field is created by superconductive currents on the two sides of the beam pipe. Since the current direction of

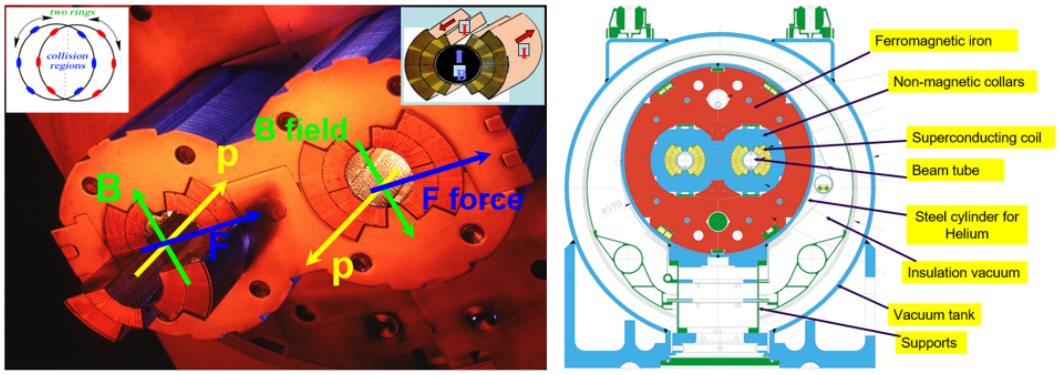


Figure 3.2: (left) The "two-in-one" design of the dipole magnets; (right) cross-section of dipole magnets in a cryostat [26].

each side is opposite to the other, a single magnetic field perpendicular to the pipe axis is created (see Fig. 3.2).

Protons are all positively charged particles, and the repulsive forces make the beam diverge. So it's necessary to focus the beam. This is achieved by quadrupole magnets. With half of the quadrupoles constraining the beam width and the other constraining the beam height (see Fig. 3.3), the protons are constantly adjusted toward the center of the beam pipe where the greatest number of collisions is to occur. The LHC consists of 392 main quadrupoles.

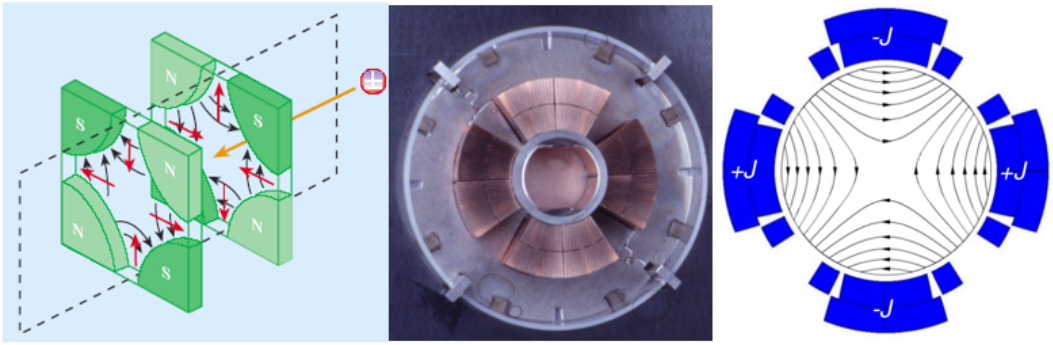


Figure 3.3: (left) Two quadrupoles work together to keep the protons tightly bunched in the direction transverse to the z-axis; (middle) picture of a quadrupole magnet; (right) the current flow of a quadrupole [36].

Also, given that the protons in a single RF trough have a spread of momenta and each follows a slightly different trajectory, 3232 sextupole magnets are used to overcome this problem and further focus the beam.

3.1.1.2. Superconductivity

The LHC relies on superconductivity to generate the required magnetic field. The dipole magnet is a winding made from special superconducting cables (Rutherford cables), containing 28 strands in the inner coil layers and 36 strands in the outer coil layers. Each strand of the cable is made of thousands of Niobium-titanium (Nb-Ti) filaments of 6 µm in diameter for the inner and 7 µm in diameter for the outer layer of the coil, and surrounded by a stabilizer (typically copper). The strands are twisted to reduce inner-strand coupling currents and also provide more mechanical stability (see Fig. 3.4). The cables are cooled to 2 K by superfluid helium for operation.

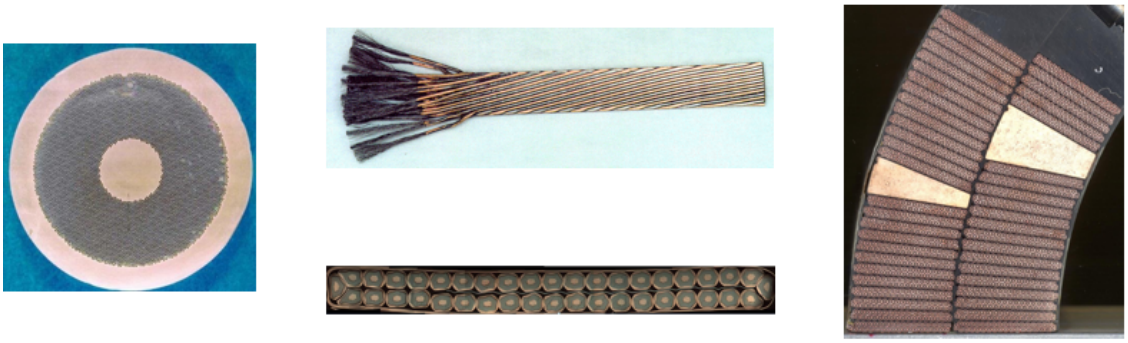


Figure 3.4: (left) Cross-section of a superconducting cable; (middle) picture and cross-section of a strand; (right) the distribution of the conductor in the dipole coil cross-section [26].

3.1.2. Beam Injection and Beam Dump

Injection of the proton beams into the LHC is achieved through the injection insertions in Points 2 and 8. The transfer line TI 2 (see Fig. 3.1) delivers the beam to $\sim$ 150 m before Point 2 for injection to Ring 1; TI 8 (see Fig. 3.1) takes the beam to $\sim$ 160 m before Point 8 for injection to Ring 2.

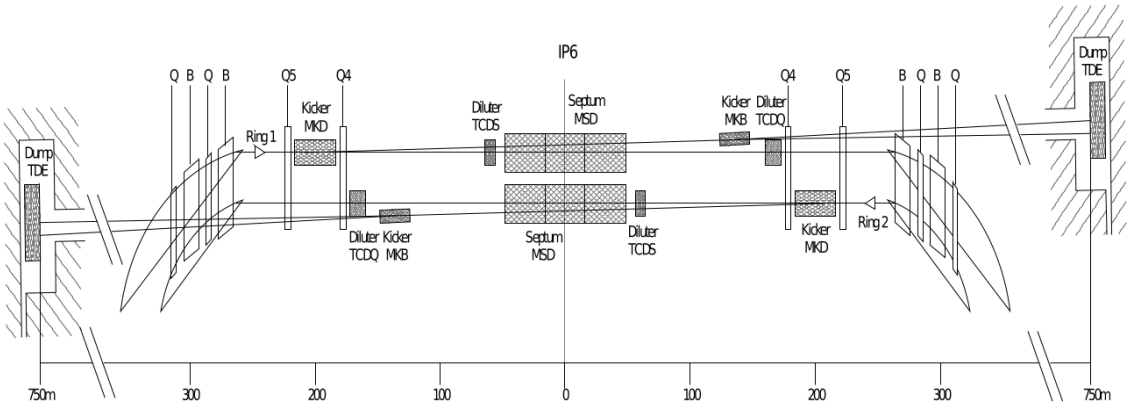


Figure 3.5: Schematic layout of beam dumping system elements around LHC Point 6 (distance in m) [26].

The dedicated beam dumping system (see Fig. 3.5) of the LHC is sited in Point 6. The system can kick the beam from each ring in a loss-free way, to an external absorber positioned sufficiently far away allowing the beam to go through the diluter system (so as not to overheat the absorber material). Given the energy stored in the LHC beam (362 MJ per beam at 7 TeV beam energy), the dumping system must meet extremely high-reliability criteria to make sure that the LHC is running safe and stable and the data-recording of the detectors is not interrupted frequently.

3.2. The ATLAS Detector

ATLAS (A Toroidal LHC ApparatuS) [38] is a general-purpose detector built for probing *pp* and heavy-ion collisions (see Fig. 3.6). The nominal interaction point is defined as the

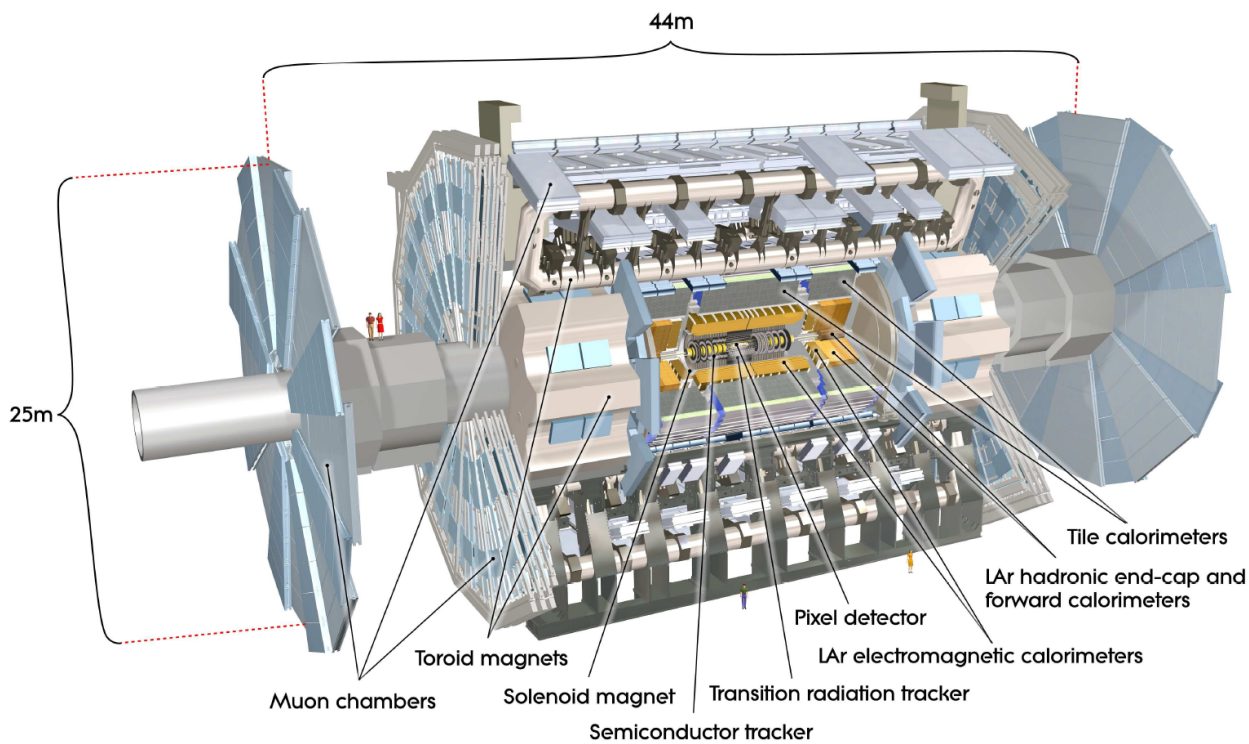


Figure 3.6: Overview of the ATLAS detector [37].

origin of the coordinate system, while x, y, z-axis are defined as in Fig. 3.7. The azimuthal angle ϕ is defined as the angle around the beam axis (z-axis), and the polar angle θ is defined as the angle from the beam axis. Pseudorapidity is normally used in hadron colliders to describe the angle relative to the beam axis for highly relativistic particles, due to that it is much easier to estimate than the rapidity[4]. Its definition is $\eta = -\ln\tan(\theta/2)$. ΔR is defined as the distance in the pseudorapidity-azimuthal angle space which equals to $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The rich physics provided by LHC has been used as a benchmark to build the ATLAS detector. The Higgs boson is produced at ~ 0.5 Hz at $\sqrt{s} = 13$ TeV; a range of particularly important production and decay processes can be studied. The top quark is produced at a rate of ~ 10 Hz at LHC, which provides an opportunity to measure its coupling and spin.

[4]The rapidity, as a measure for relativistic velocity, is defined as $y = 1/2\ln[(E + p_z)/(E - p_z)]$.

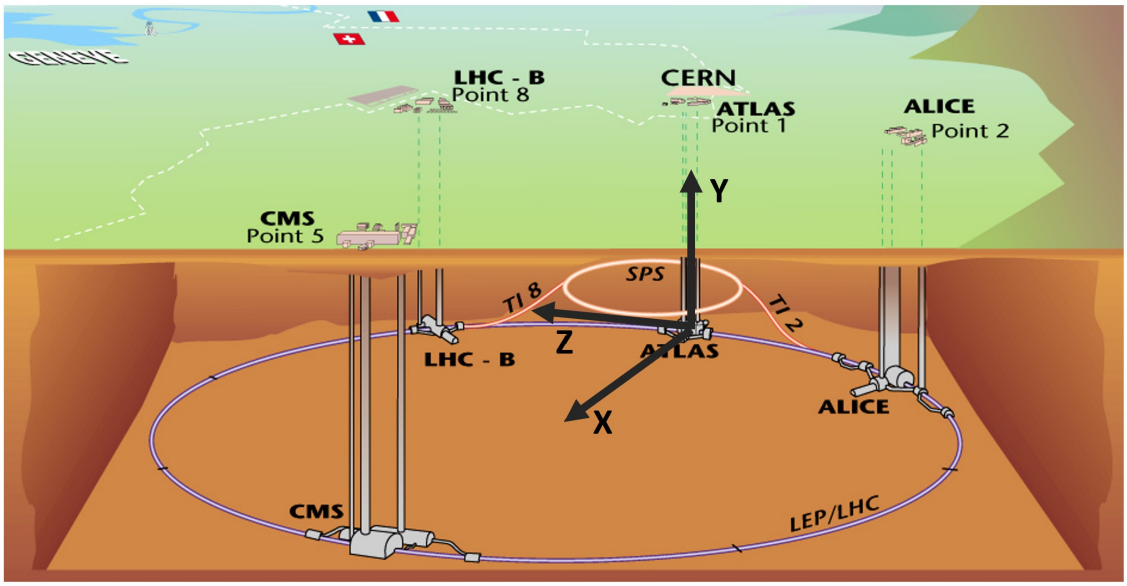


Figure 3.7: The ATLAS experiment location at the LHC and the coordinate system of ATLAS detector are shown in the figure. The side-A of the detector is defined as the side of positive z and side-C is the side of negative z. This picture is modified from Ref. [39].

The high luminosity and increasing interaction cross-section with beam energy at LHC also enable further precision tests of QCD, electroweak interaction, and flavor physics.

The final states of the $VHbb$ analysis are leptons, jets, and neutrinos. The need to reconstruct and identify charged leptons, quark jets as well as missing momentum from neutrinos, requires the use of nearly every key ATLAS system. Thus, the design, construction, and operation of the ATLAS detector are very important to the $VHbb$ analysis.

3.2.1. Physics Requirements

The ATLAS detector is required to survive at the LHC designed luminosity and cover the physics phenomena at the TeV scale as discussed in Sec. 2.2. The nature of pp collision produces a large amount of QCD jets over the rare processes (like the $H \to b\bar{b}$ decay), which challenges the particle identification of the detector. Thus, it imposes high standards on the detector design (summarized below) and the performance goal (shown in Table 3.1).

- Radiation-hard sensors and electronics, and high granularity for particle identification.
- Large acceptance in pseudorapidity, and full azimuthal angle coverage.
- Good charged-particle reconstruction in the inner tracker, and good vertexing for the tagging of τ-leptons and b-jets.
- Good calorimetry for e/γ identification and energy resolution, and jet and missing transverse momentum reconstruction.
- Good muon detector for muon identification and momentum resolution.
- Efficient triggering for low transverse momentum[5] objects and good rejection of background.

Detector component	**Required resolution**	**η coverage**	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	σ_{p_T}/p_T=10% at p_T = 1 TeV	± 2.7	± 2.4

Table 3.1: General performance goals of ATLAS detector. "$\oplus$" indicates a quadratic sum. The energy resolution of calorimeters can generally defined as $\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$, where the first term is the "stochastic" term, the second term is the "noise" term, and the third term is the "constant" term. The units for E and p_{T} are in GeV [37].

[5]The momentum of an object in the transverse plane is often denoted as p_{T}.

3.2.2. Magnet System

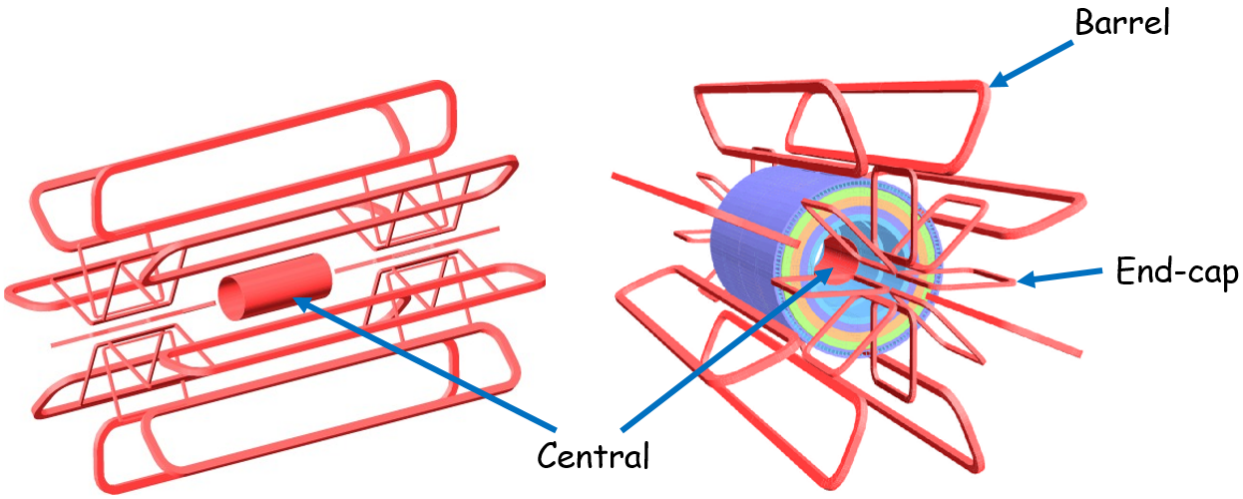


Figure 3.8: Geometry of the ATLAS magnet system windings [37].

The magnetic field is a widely used tool to measure the momentum of charged particles. When a charged particle travels in a magnetic field, the direction of its velocity can be changed by the Lorentz force. Thus its trajectory will be a curve (whether or not the magnetic field is homogeneous). Given that the magnetic field can be mapped at each point of the tracking detectors, the hits left by the charged particle can be used to reconstruct its trajectory and then measure its momentum. Taking the classical case (velocity $\ll c$) for example, while a charged particle executing circular motion in a plane perpendicular to a magnetic field, its momentum can be measured by $p = qBR$, where q is the charge, B is the magnetic field strength and R is the radius of the orbit.

The ATLAS magnet system consists of 1 solenoid and 3 toroids (1 barrel toroid and 2 end-cap toroids). The magnetic system provides the magnetic field over a volume of $\sim$ 12000 m^3 (defined as the region in which the field exceeds 50 mT), and stores energy of 1.6 GJ in total [40]. The general layout of the magnetic system is shown in Fig. 3.8.

The central solenoid is aligned to the beam axis and provides a 2 T axial magnetic field for the inner tracker. Its layout was optimized by reducing the material thickness in front of the calorimeter ($\sim$ 0.66 radiation length[6], which affects the calorimeter's performance). The flux of the magnetic field is returned by the steel of the ATLAS hadronic calorimeter and its girder structure [41]. The barrel toroid provides the magnetic field (0.5 T) to fill the cylindrical volume surrounding the calorimeters, where the muon detectors of the central regions are. It consists of 8 coils encased in the individual racetrack-shaped, stainless-steel vacuum vessel. The two end-cap toroids generate the magnetic field (1 T) required for optimizing the bending power in the end-cap regions of the muon spectrometer system [42].

3.2.3. Inner Detector

The ATLAS Inner Detector (ID) is designed to provide robust pattern recognition, excellent momentum resolution and vertices measurements for charged tracks above a given p_T threshold (nominally 0.5 GeV), within the central region ($|\eta|< 2.5$). It also provides electron identification for an energy range of 0.5 - 150 GeV within the region of $|\eta|< 2.0$ [43].

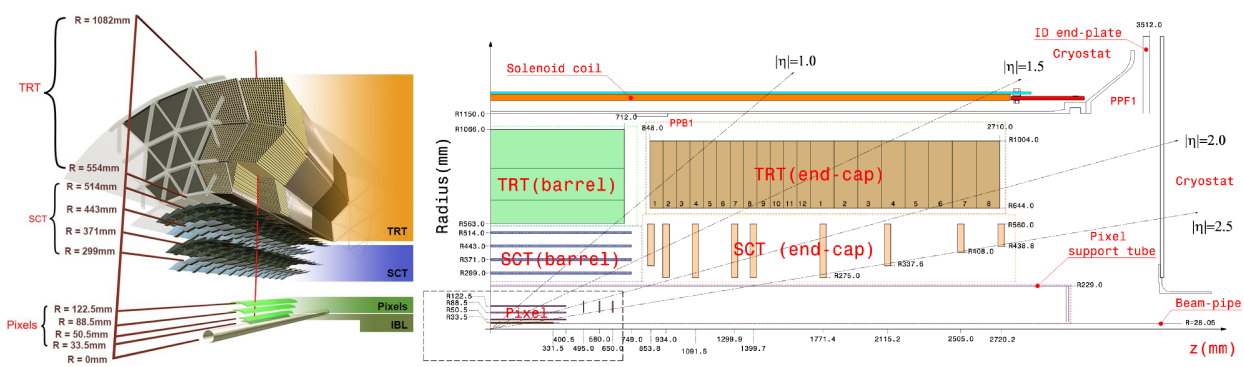


Figure 3.9: (left) Picture of the barrel Inner Detector traversed by a charged track; (right) plan view of a quarter-section of the Inner Detector showing each of the sub-detector elements [37].

[6]The radiation length is a characteristic of a material, related to the energy loss of high energy particles electromagnetically interacting with it.

The ID consists of three sub-detectors, as shown in Fig. 3.9. At inner radii, the space points from silicon pixel layers (Pixel and IBL) and stereo pairs of silicon microstrip (SCT) layers are used for the high-resolution pattern recognition and reconstruction of charged tracks. At larger radii, the transition radiation tracker (TRT) uses many layers of gaseous straw tube elements to record the hits of tracks and provide electron identification complementary to that of the calorimeter [44].

3.2.4. Calorimetry

The ATLAS calorimeters consist of several sampling detectors of full ϕ coverage and symmetry around the beam axis (see Fig. 3.10). The calorimeters closest to the beam-line are housed in three cryostats, one in the barrel and two in the end-caps. The barrel electromagnetic calorimeter is housed in the barrel cryostat; the forward calorimeters, each end-cap containing an electromagnetic end-cap calorimeter (EMEC) [45], a hadronic end-cap calorimeter (HEC) [46] located behind the EMEC, and a forward calorimeter (FCal) [47] , are housed in the relative end-cap cryostats. All these calorimeters use liquid argon as the active detector medium which is chosen for its intrinsic linear behavior, the stability of response over time, and intrinsic radiation-hardness [48].

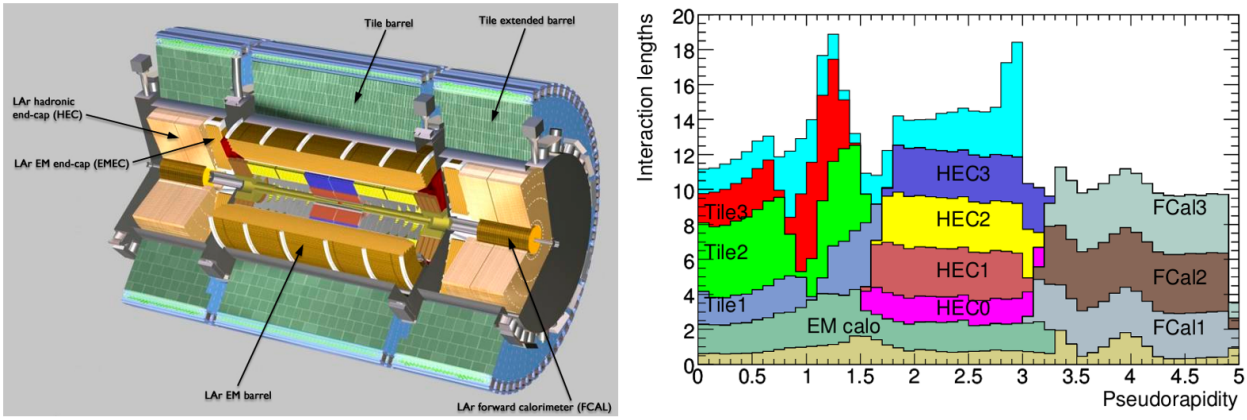


Figure 3.10: (left) View of the ATLAS calorimetry; (right) cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in different calorimeters [37].

The outer hadronic calorimeter consists of 1 central barrel and 2 extended barrels. Scintillator tiles are chosen as the sampling medium, and steel is chosen as the absorber medium for the hadronic calorimeter. This is to provide the maximum radial depth with the least cost of ATLAS [49].

3.2.5. Muon Spectrometer

The muon spectrometer is designed to detect muons exiting the calorimeters and measure their momentum in the region of $|\eta| < 2.7$. It also allows the data acquisition system to trigger on these muons in the region of $|\eta| < 2.4$. The goal is to have a standalone p_{T} resolution of $\sim 10\%$ for 1 TeV tracks [50].

The precise momentum measurement is performed by the Monitored Drift Tube chambers (MDT), which combine high measurement accuracy, predictability of mechanical deformations, and simplicity of construction (see Fig. 3.11). Their coverage on pseudorapidity is $|\eta| < 2.7$. MDT chambers consist of three to eight layers of drift tubes, which operate at an absolute pressure of 3 bar, and achieve an average resolution of 80 µm per tube. [51].

In the forward region ($2 < |\eta| < 2.7$), Cathode-Strip Chambers (CSC) are used in the inner-most tracking layer due to their higher rate capability and time resolution. CSCs are multi-wire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows both coordinates to be measured from the induced-charge distribution. Each CSC chamber has a resolution of 40 µm in the bending plane and ~ 5 mm in the transverse plane [52].

For the triggering, Resistive Plate Chambers (RPC) were selected to trigger on muon tracks [53] in the barrel region ($|\eta| < 1.05$); in the end-cap region ($1.05 < |\eta| < 2.4$), Thin Gap Chambers (TGC) were chosen for this purpose [54].

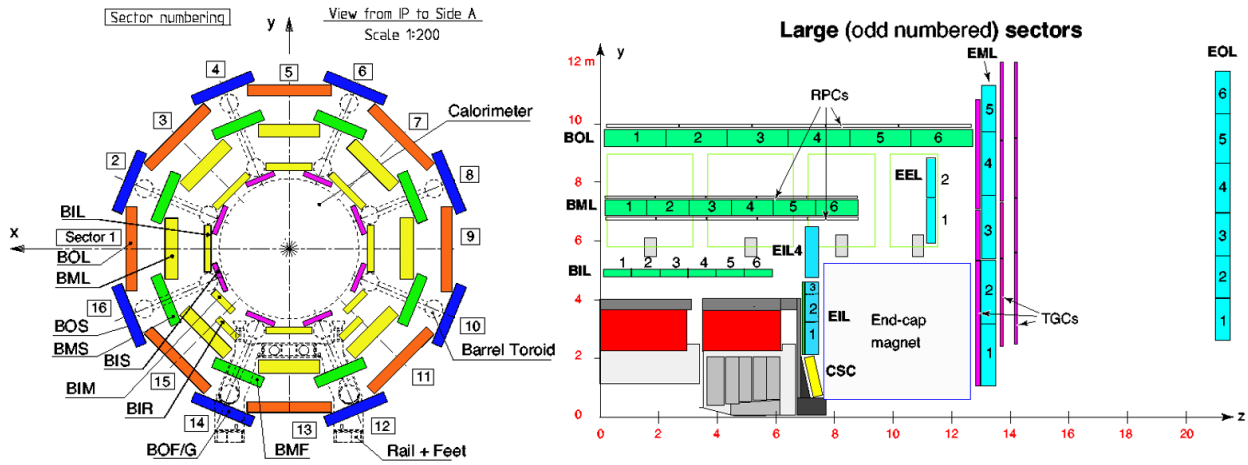


Figure 3.11: (left) Cross-section of the barrel muon system; (right) cross-section of the muon system in a plane containing the beam axis. Note: the three-letter acronym denotes the location and type of the muon detector installed. Refer to Ref. [37] for details of the notation.

3.2.6. Luminosity Monitoring Detectors

The luminosity of the collider is a key input to the interpretation of the measurement in the physics analysis, especially the uncertainty on that measurement. LUCID-2 (the new LUminosity measurement using Cherenkov Integrating Detector) [55] is used in ATLAS to measure the luminosity and its associated uncertainty. It is located at a distance of ± 17 m from the interaction point, near the TAS (Target Absorber Secondaries) collimator.

The LUCID-2 detector consists of two subdetectors using PMTs to collect the Cherenkov light[7]. One subdetector is made of 16 PMTs grouped by four and installed around the beam pipe in the ATLAS TAS shield region. The quartz window of the PMT itself acts as a Cherenkov radiator. Hamamatsu R760 PMTs with sensitive window diameters of 10 mm and 7 mm are used (see Fig. 3.12). The 7mm diameter ones are to keep the detector acceptance low duet to the increased particle production rate in Run-2 [55]. The other subdetector uses 4 bundles of quartz fibers as Cherenkov radiator. These bundles are read out by PMTs (one

[7]Cherenkov light is emitted when a charged particle passes through a dielectric medium at a speed greater than the phase velocity of light in that medium.

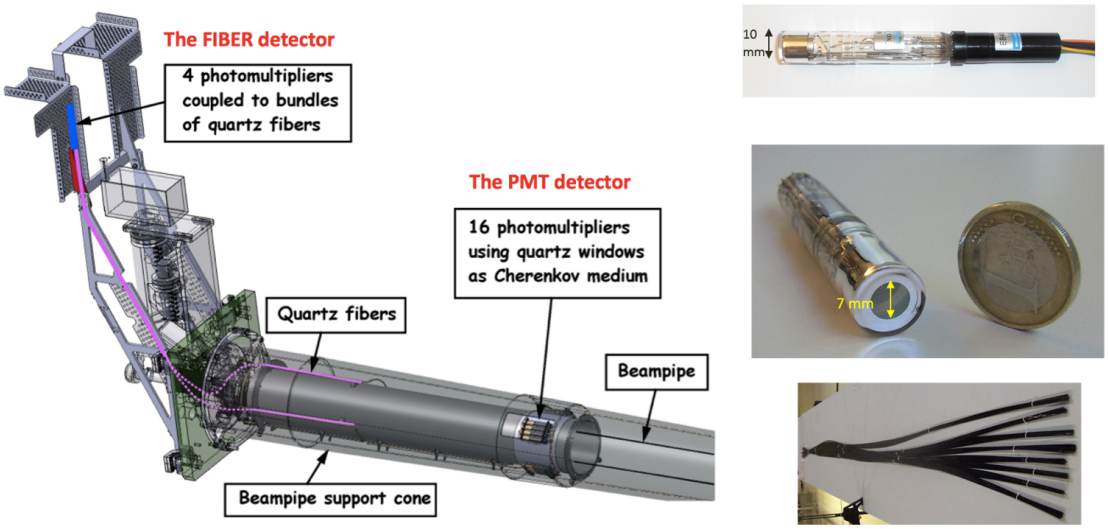


Figure 3.12: (left) Sketch of the LUCID-2 detector; (right-top) Hamamatsu R760 PMT (quartz window 10 mm diameter); (right-middle) Hamamatsu R760 PMT (quartz window 7 mm diameter); (right-bottom) Bundles of LUCID-2 readout fibers. [55]

per bundle) placed on top of the ATLAS shielding (see Fig. 3.12). All the 20 readout PMTs are inserted in mu-metal cylinders to shield from stray magnetic fields.

The van der Meer method [56] is used to measure the luminosity using the LUCID-2 detector. By varying the separation of two colliding beams, the mean number of pp interaction per bunch as a function of the separation can be acquired. Then the total number of protons in each colliding bunch can be determined from the measurement of the beam currents [57].

3.2.7. Trigger and Data Acquisition System

The trigger system is responsible for deciding whether to keep a given collision event or not for later study. It consists of 2 levels of event selection: Level-1 (L1) and the High-Level Trigger (HLT). The DAQ (data acquisition system) receives the events data from the detector specific readout electronics. After accepted by the L1 trigger, events are buffered in the Read-Out System (ROS) and processed by the HLT. The HLT uses full granularity and

precision of calorimeter and muon chamber data, as well as the data from the inner detector, for event building in RoIs or the whole detector. After accepted by the HLT, events are transferred to local storage at the experimental site and exported to the Tier-0 facility at CERN's computing center for offline reconstruction [58].

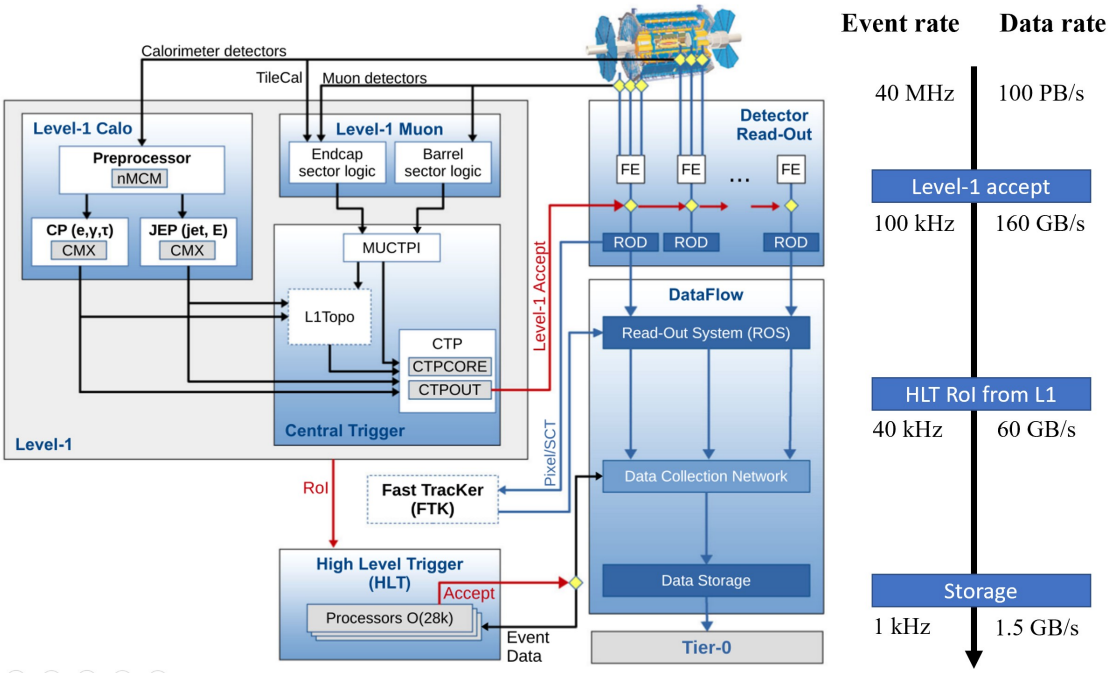


Figure 3.13: Schematic layout of the ATLAS trigger and data acquisition system in Run 2 [59].

3.3. My Contribution to the ATLAS Detector Upgrade

The ATLAS detector has been constantly upgraded to improve its capability in selecting required physics events and reconstructing physics objects. My contributions to the ATLAS detector upgrade are mainly on two projects,

- Fast TracKer (FTK): test of FLIC FPGA and its associated ATCA blade for the data quality monitoring, and development of the data quality monitoring software;
- Inner Tracker (ITk): test of the GBCR ASIC, and test of the data transmission chain for the ITk Pixel detector.

I will explain these projects and my contributions in detail in the following sections.

3.3.1. Contribution to the Faster TrcKer (FTK) Project

The Fast TracKer (FTK) system was designed to provide global ID track reconstruction at the L1 trigger rate using lookup tables stored in custom association memory chips for the pattern recognition. The implementation of the FTK system was believed to be able to improve the Higgs candidate event selection, lepton isolation, jet flavor tagging, and primary vertex finding.

FTK-to-LVL2-Interface-Crate (FLIC) is the last stage board in the FTK system. It receives data from the upstream board, replaces the local coordinates of recorded tracks by the global ones, and organizes the data format for the HLT input. The data FLIC receives have been fully fitted by previous FTK boards. Thus, one thing that FLIC was designed for is to use its associated ATCA blade[8] to monitor the data quality of the FTK output data.

[8]The ATCA blade is a computer specially designed for the ATCA crate. The ATCA specifications can be seen in Ref. [60].

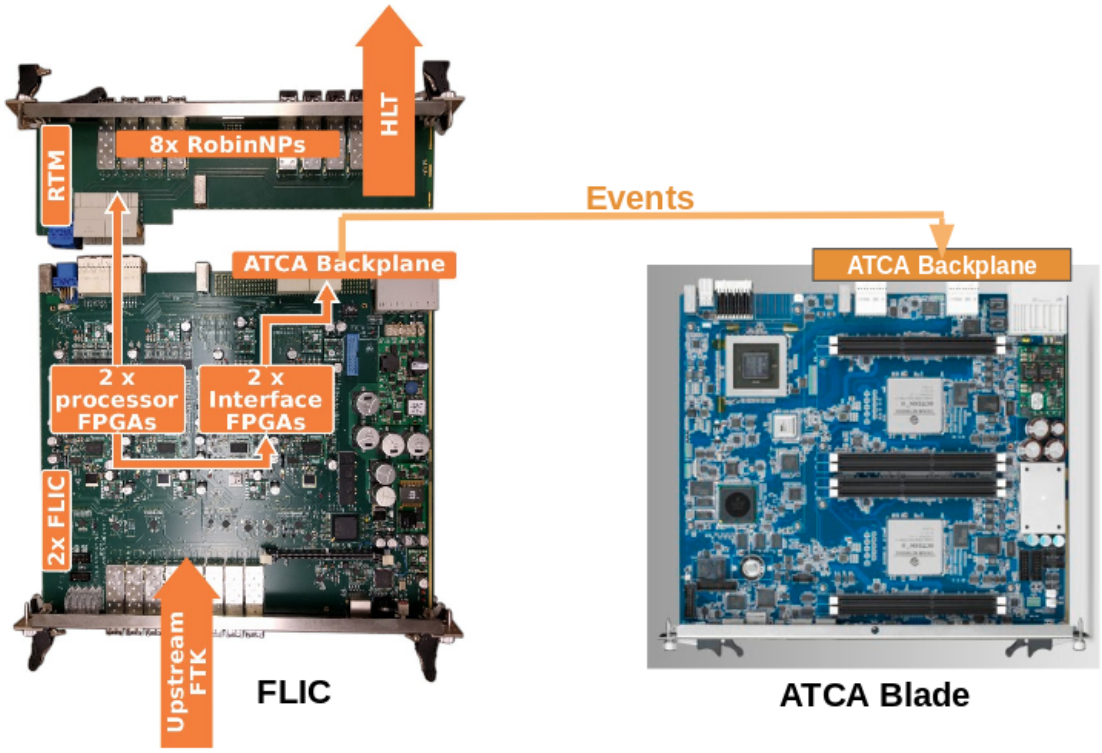


Figure 3.14: Illustration of the FLIC data quality monitoring software running on the ATCA blade.

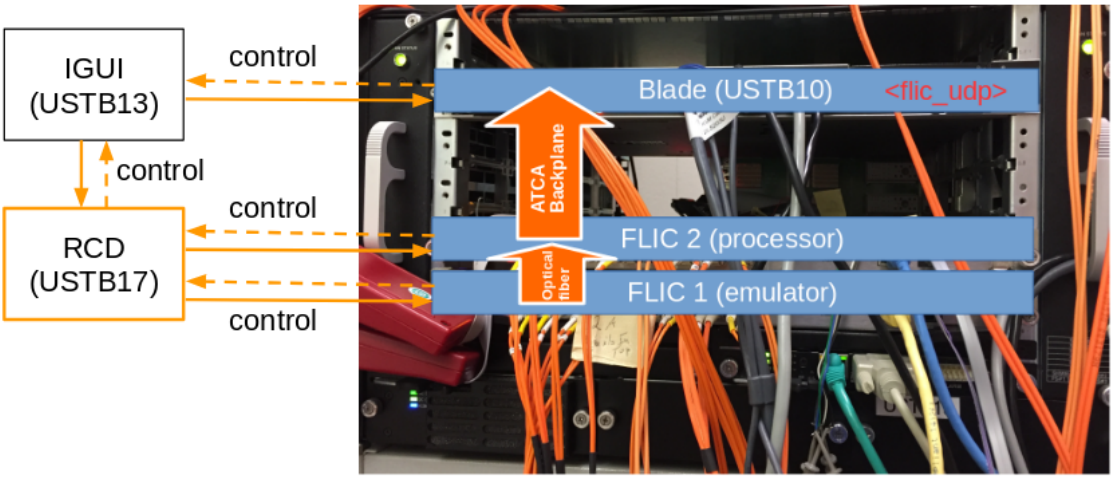


Figure 3.15: Test-stand for the FLIC DQM software test at ANL. `USTB13` and `USTB17` are external computer names.

The associated ATCA blade receives a copy of data from the FLIC board while the FLIC is processing the data and sending them to the HLT (see Fig. 3.14).

In this project, I tested and tuned the FLIC data quality monitoring system (FLIC and its associated ATCA blade), and wrote thousands of lines of C++ code for the data quality monitoring software, which prior to my work did not exist for the FTK. This software would allow us to monitor the FTK output data and publish quality-check histograms to Online Histogram service (OH).

3.3.1.1. FLIC and its Associated ATCA Blade

A test-stand was set up at Argonne National Lab (ANL) for the test of FLIC data quality monitoring (DQM) hardware and the data quality monitoring software, as shown in Fig. 3.15. In the test-stand, one FLIC board acts as the emulator of the upstream board sending data to the second FLIC board. The second FLIC board processes the data received and then sends a copy of the data to the ATCA blade through the ATCA crate backplane. Two external computers are used to control the emulator board, the FLIC board, and the ATCA blade through the Ethernet cable. The FPGAs on the processor FLIC are carefully configured so that they can send a copy of data through the Ethernet connection of the ATCA backplane in correct format.

3.3.1.2. FLIC Data Quality Monitoring Software

The DQM software runs on the ATCA blade. I designed the structure as shown in Fig. 3.16. It first receives the fragments of an event sent from the FLIC board and combines them into a complete event. This event is then buffered to the 8 GB RAM of the ATCA blade. At the same time, parallel threads access the buffered event and decode them into

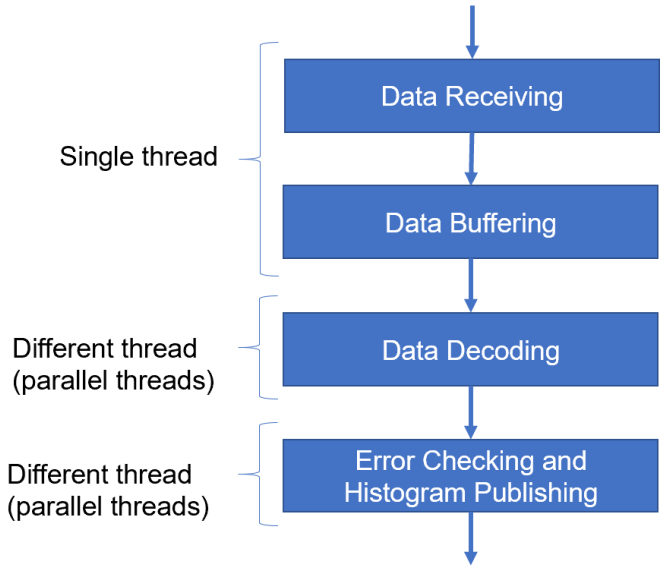


Figure 3.16: The structure of the FLIC DQM software. FLIC DQM software is a parallel program. While the DQM software keeps receiving events and buffer them, standalone threads access the buffered events, decode them and publish the results to Online Histogram (OH).

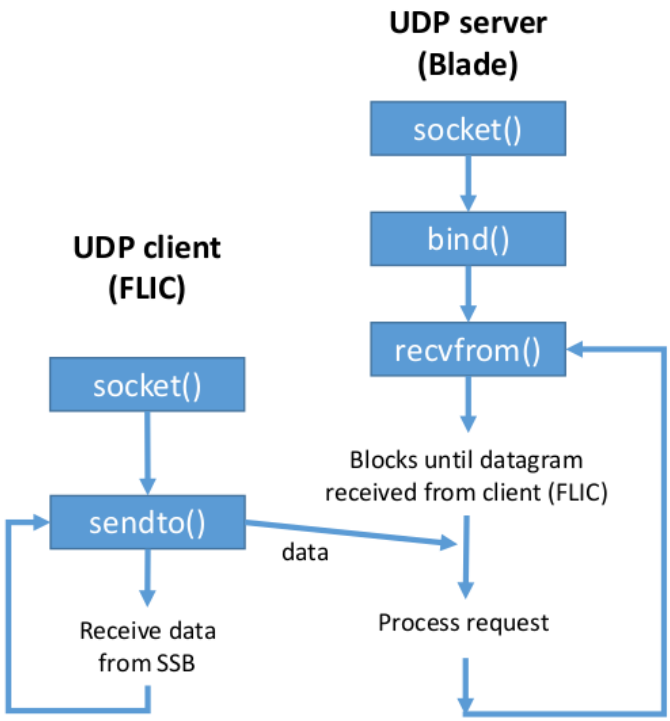


Figure 3.17: The UDP datagram protocol is used for the associated ATCA blade to receive data. It is designed to be a single direction communication, which means that FLIC boards can only send data and the ATCA blade can only receive data.

the standard C++ storage structure. Those events decoded are further processed by other parallel threads for error-checking and histogram publishing.

The UDP protocol[9] is used on the ATCA blade and FLIC board for data communication, as shown in Fig. 3.17. The single-direction communication is applied to avoid that the ATCA blade stops the FLIC board processing upstream data. The raw data format of FTK output is decoded by the FLIC DQM, as illustrated in Fig. 3.18. The 8 GB RAM can buffer a large number of events which can be written to TDAQ computer at Point 1 when the FTK system stops working accidentally. This mechanism significantly helps debug the possible problems of the FTK system. The process of FLIC DQM software publishing histograms is shown in Fig. 3.19. FLIC DQM software directly publishes histograms to OH where the run-control shifter can view them from the TDAQ control panel (as shown in Fig. 3.20). Checks on the header and trailer of the FTK data, and the error flag and debug block are made; relative warning and error (if found in the data) will be published to the TDAQ control panel.

The FLIC DQM software I developed had been tested at Point 1 in Run-2. Due to the serious design and technical challenges, the whole FTK project was unfortunately canceled. Valuable experience and lessons were learned from all of these processes.

[9]UDP, User Datagram Protocol, is a protocol in the transport layer. The main difference between the UDP protocol and the TCP/IP protocol is that UDP protocol doesn't ask for feedback when sending a datagram, while TCP/IP protocol asks for feedback when sending a datagram. More technical details can be referred to Ref. [61].

Bit => Word	15	14	13	12	11	10	09	08	07	06	05	04	03	02	01	00
AH01	C/F	RSVD			LEVEL 1 ID[11..0]											
AH02	# of words following this word															
AH03	INPUT BITMASK								FRAGMENT BITMASK							
AH04	FRGMENT ID 0				FRGMENT ID 1				FRGMENT ID 2				FRGMENT ID 3			
AH05	FRGMENT ID 4				FRGMENT ID 5				FRGMENT ID 6				FRGMENT ID 7			
ET01	1	0	1	1	1	1	1	0	User Tag						EV ST	CP FLG
RH05	0	Run Number [30:16]														
RH06	Run Number [15:0]															
RH07	Extended Level 1 ID [31:16]															
RH08	Extended Level 1 ID [15:0]															
RH09	Reserved															
RH10	Reserved				BCID [11:0]											
RH11	Reserved															
RH12	Reserved								Level 1 Trigger Type [7:0]							
RH13	Reserved								Detector Event Type [7:0]							
RH14	Reserved												TIM[3:0]			
TH1	Res [3:0]			L	B				D				A			
TH2	SECTOR NUMBER															
TH3	Reserved				TF#				Res	Tower Number						
TH4	Reserved								Layer Map [11:0]							
TH5	Reserved								ROAD ID [23:16]							
TH6	ROAD ID [15:0]															
TH7	TRACK CHISQ															
TH8	TRACK D0															
TH9	TRACK Z0															
TH10	TRACK COTTH															
TH11	TRACK PHI0															
TH12	TRACK CURV															
IBLa	0	COL WIDTH			COL COORDINATE [11:0]											
IBLb	s	ROW_WIDTH			ROW COORDINATE [11:0]											
PL0a	0	COL WIDTH			COL COORDINATE [11:0]											
PL0b	s	ROW_WIDTH			ROW COORDINATE [11:0]											
PL1a	0	COL WIDTH			COL COORDINATE [11:0]											
PL1b	s	ROW_WIDTH			ROW COORDINATE [11:0]											
PL2a	0	COL WIDTH			COL COORDINATE [11:0]											
PL2b	s	ROW_WIDTH			ROW COORDINATE [11:0]											
SAx0	0	HIT2 WIDTH			ReV	HIT2 COORDINATE [10:0]										
SSt0	0	HIT1 WIDTH			ReV	HIT1 COORDINATE [10:0]										
SAx1	0	HIT2 WIDTH			ReV	HIT2 COORDINATE [10:0]										
SSt1	0	HIT1 WIDTH			ReV	HIT1 COORDINATE [10:0]										
SAx2	0	HIT2 WIDTH			ReV	HIT2 COORDINATE [10:0]										
SSt2	0	HIT1 WIDTH			ReV	HIT1 COORDINATE [10:0]										
SAx3	0	HIT2 WIDTH			ReV	HIT2 COORDINATE [10:0]										
SSt3	0	HIT1 WIDTH			ReV	HIT1 COORDINATE [10:0]										
RTF1	E				0				D				A			
RTF2	Length of Debug Block															
RTFD(1)	debug information															
RTFD(N)	debug information															
RTF3	E				0				D				F			
RTF4	Length of Debug Block															
RTF5	L1ID [31:16]															
RTF6	L1ID [15:0]															
RTF7	Error Flag [31:16]															
RTF8	Error Flag [15:0]															
RTF9	Reserved															
RTF10	Reserved															
RTF11	Reserved															
RTF12	Reserved															
AF01	Fixed value 0xE0F1															
AF02	Fixed value 0xE0F2															
AF03	Fixed value 0xE0F3															

Repeated for as many fragments as are available to assemble

Repeated for as many tracks as present in record

```
============start of an assembly header===========
Number of word following this word:          340
+++++++++++++start of a record header+++++++++++++
Checksum:                                    passed
Mystery number:                              668
RunNumber:                                   305419896
ExtendedLevel1ID:                            4
Bunch Crossing ID:                           4
L1TriggerType:                               162
Detector Event Type:                         68
TIM:                                         8
++++++++++++++++end of a record header+++++++++++++
---------------start of a track header------------
Link ID:                                     0
Sector Number :                              44
Tower Number:                                5
Track Fitter Number:                         0
Layermap[11:0] :
  11SCT 10SCT 09SCT 08SCT 07SCT 06SCT 05SCT 04SCT 03PIX 02PIX 01PIX 00IBL
     0     1     0     0     0     1     0     0     0     0     1     0
RoadID:                                      8704265
chi sq :                                     -0.011864
track d0:                                    4.152344
track z0:                                    -0.000131
track cotth:                                 0.000563
track phi:                                   0.010208
track curv:                                  3.226562
---------------end of a track header-------------
+++++++++++++start of a record trailer++++++++++++
Begin trailer debug block :                  e0da
Length of debug block | Indeterminate[3:0] : 0
End debug block :                            e0df
Length of debug block | Indeterminate[3:0] : 0
L1ID:                                        4
ErrorFlag :                                  -286335523
Module data word count (checksum):           329
End of fragment word:
                                             e0f0
                                             a5a5
                                             5a5a
                                             0e0f
++++++++++++++++end of a record trailer+++++++++++++
```

Figure 3.18: The column in the left is the data format that the ATCA blade received from the FLIC board; the column in the right is the data format decoded on the ATCA blade. Different blocks of the data (assembly header, record header, track header, record trailer, etc.) are marked with different background colors for the convenience of comparison.

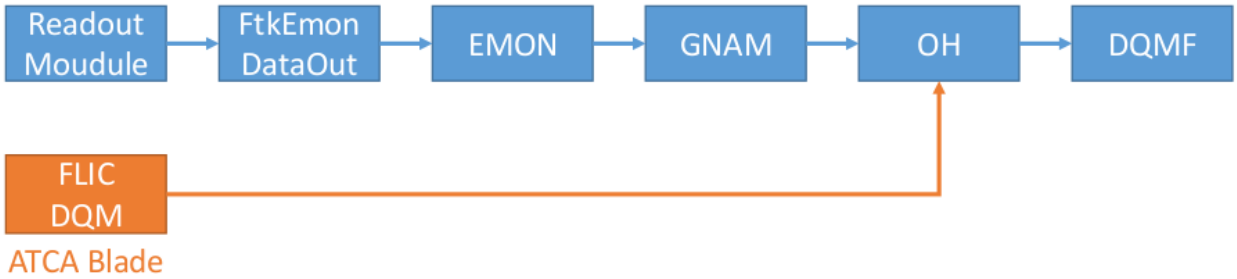


Figure 3.19: FLIC DQM software publishes histograms directly to Online Histogram (OH).

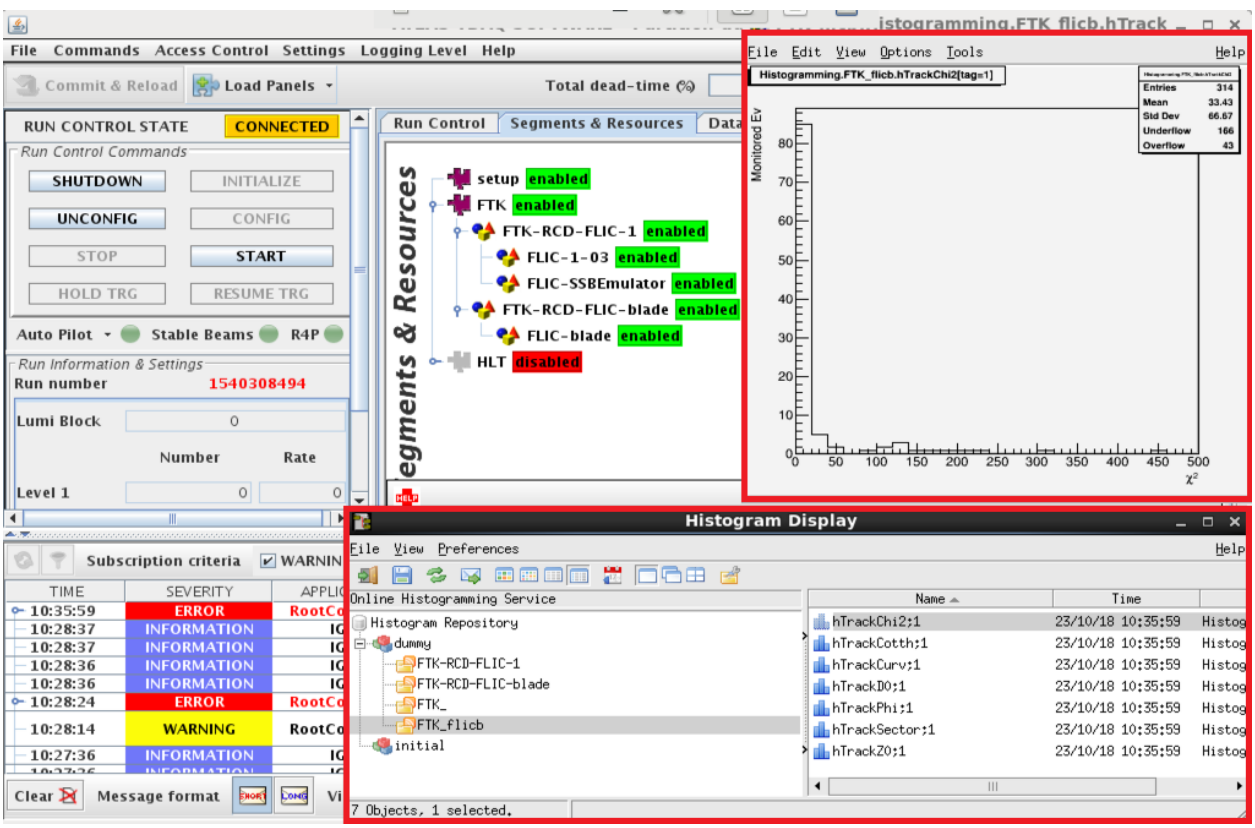


Figure 3.20: The histograms published to OH by the FLIC DQM software can be viewed from the TDAQ control panel. The box in red at the top-right corner shows a histogram of the χ^2 of the monitored tracks; the box in red at the bottom-right corner shows a list of monitored histograms published.

3.3.2. Contribution to the Inner Tracker (ITk) Project

The ATLAS Inner Tracker (ITk) detector (see Fig. 3.21) is a charged particle tracking detector designed for the High Luminosity LHC (HL-LHC), where there will be about 200 inelastic *p-p* collisions per beam crossing. It will replace the current tracking system with better performance on tracking and vertexing as well as physics object reconstruction. The goal of it is to operate over the HL-LHC program during which the integrated luminosity will reach ~ 3000 fb^{-1}. As we know, laser diodes are used to convert the electric signals of the readout chip to optical signals for the Pixel detector. Since the radiation level in ITk will increase by roughly an order of magnitude (see Fig. 3.21), the laser diodes cannot operate close to the Pixel module anymore. Thus, they have to be placed a few meters (3 to 6 meters) away from the Pixel module, and Flex and Twinax cables[10] are used to transmit the electric signal from the readout module to the Opto-conversion laser diodes.

3.3.2.1. Data Transmission Chain of the ITk Pixel Detector

The whole chain for the Pixel detector data transmission is shown in Fig. 3.22. A downlink is used to send the control command at 160 Mbps, and several uplinks are used to transmit the data to DAQ at 1.28 Gbps. In the downlink, the VTRx+ (Versatile TransReceiver) module receives the control command from the back-end through optical fibers and converts optical signals to electric signals before sending them to the lpGBT (Low-Power Gigabit Transceiver) and GBCR (Gigabit Cable Receiver) module. The GBCR module then sends the control signal through its pre-emphasis[11] which drives the Twinax and Flex cable. The readout chip receives the control signal and is therefore configured. In the uplink, the

[10]The Flex cable is made of a flexible printed circuit board (PCB), pictures of which are shown in Fig. 3.22. The Twinax Cable is a cable with two inner conductors, which is good for differential signal transmission.

[11]Pre-emphasis and equalizer are two ways to improve the signal quality due to the high-frequency loss. Pre-emphasis boosts the high-frequency signal before the transmission over the cable, while equalizer renders the high-frequency signal after the transmission over cable.

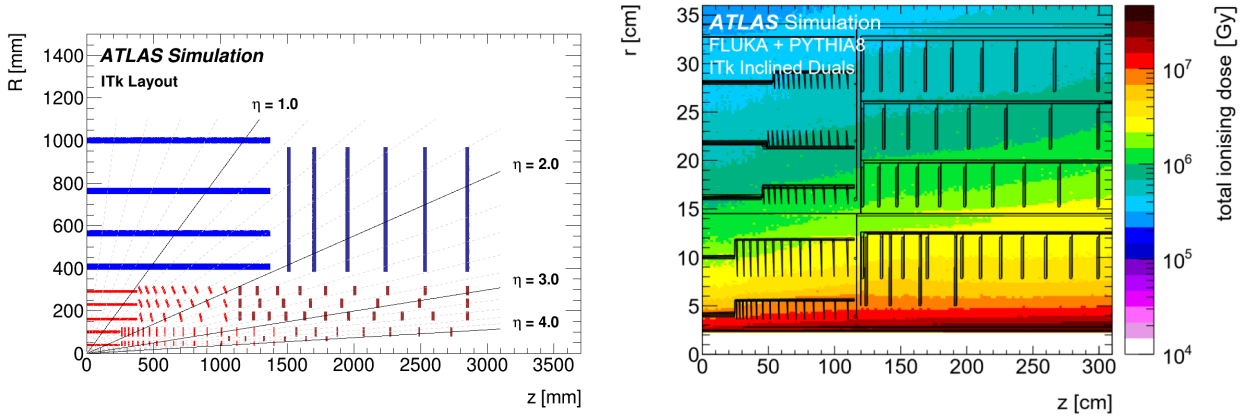


Figure 3.21: (left) A schematic layout of the ITk detector. The Pixel detector is marked in red color; (right) the total ionizing dose for the Pixel detector in the HL-LHC.

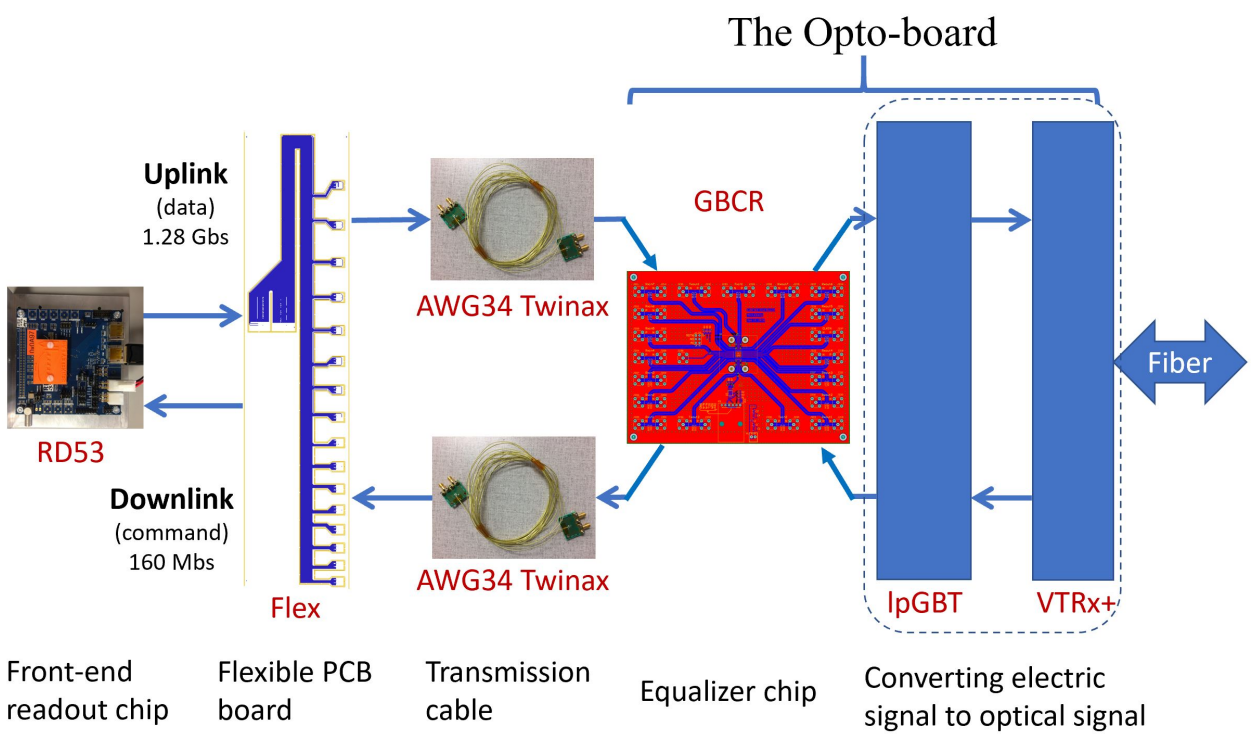


Figure 3.22: The data transmission chain for the ITk Pixel detector.

readout chip sends the data through the Flex cable and Twinax cable to the GBCR module. The GBCR compensates for the signal loss using its equalizer after receiving the data and then sends the data to lpGBT and VTRx+ for optical signal conversion.

In this project, I tested the performance of GBCR version 1 ASIC and wrote the LabVIEW program for automatizing the test. After that, I relocated to SLAC to integrate and test the whole data transmission chain for the Pixel detector.

3.3.2.2. Test of the GBCR ASIC

The test of GBCRv1 ASIC was conducted at SMU. The focus was on the performance of GBCRv1 ASIC[12]. The test-stand at SMU is shown in Fig. 3.23. The signal generator injects a clean pulse signal to a 5-meter Twinax cable. Due to the high-frequency loss in the cable, we can barely see the original signal (see Fig. 3.24). After receiving the signal, GBCR compensates for the high-frequency loss in the cable using its adjustable equalizer. The recovered signals are tested on jitter, signal sensitivity, and Bit-Error-Rate[13], and well meet the requirement. The LabVIEW program configures the GBCR ASIC by changing its register values. It scans the register value and records the eye diagram for the selection of the best working point. The results are presented in Fig. 3.24 that the equalizer of GBCRv1 well recovers the signal loss in the 5 meter Twinax cable at a speed of 5.12 Gbps.

[12]The GBCRv1 was primarily designed for recovering the signal in uplink at 5.12 Gbps. The speed of uplink was later changed to 1.28 Gbps in GBCRv2. The pre-emphasis didn't exist for the downlink of GBCRv1, although it was later added for GBCRv2.

[13]Bit-Error-Rate (BER) is defined as $\frac{\text{Number of error bits receive}}{\text{Number of total bits receive}}$. Pseudo-random bit sequences (PRBS) are normally used in the BER measurement. Random bit sequence means that the next bit randomly either "0" or "1".

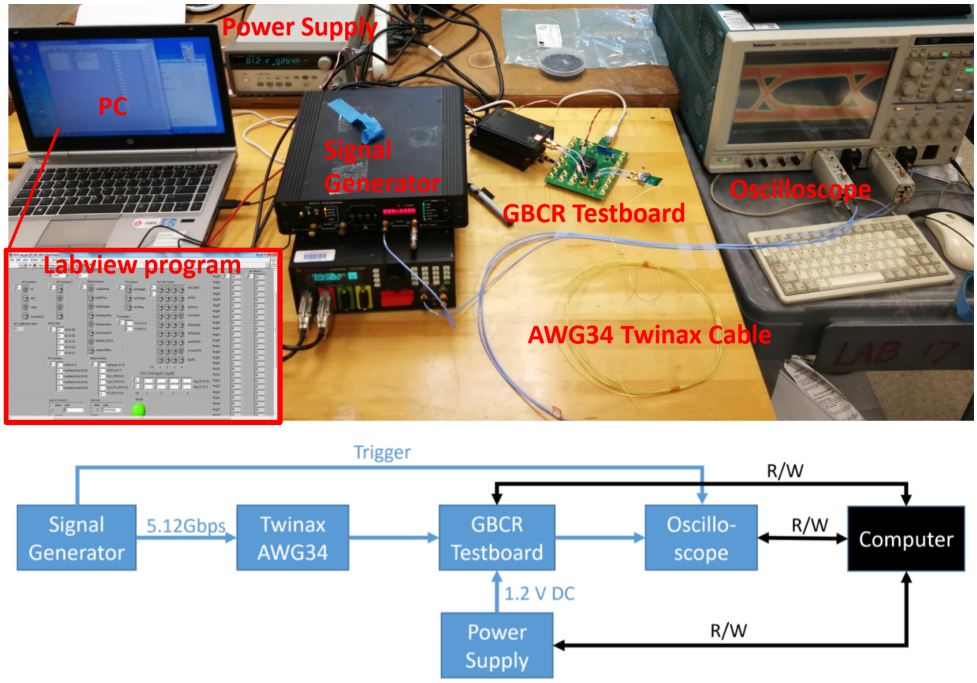


Figure 3.23: The test-stand for the GBCR (version 1) ASIC test at SMU.

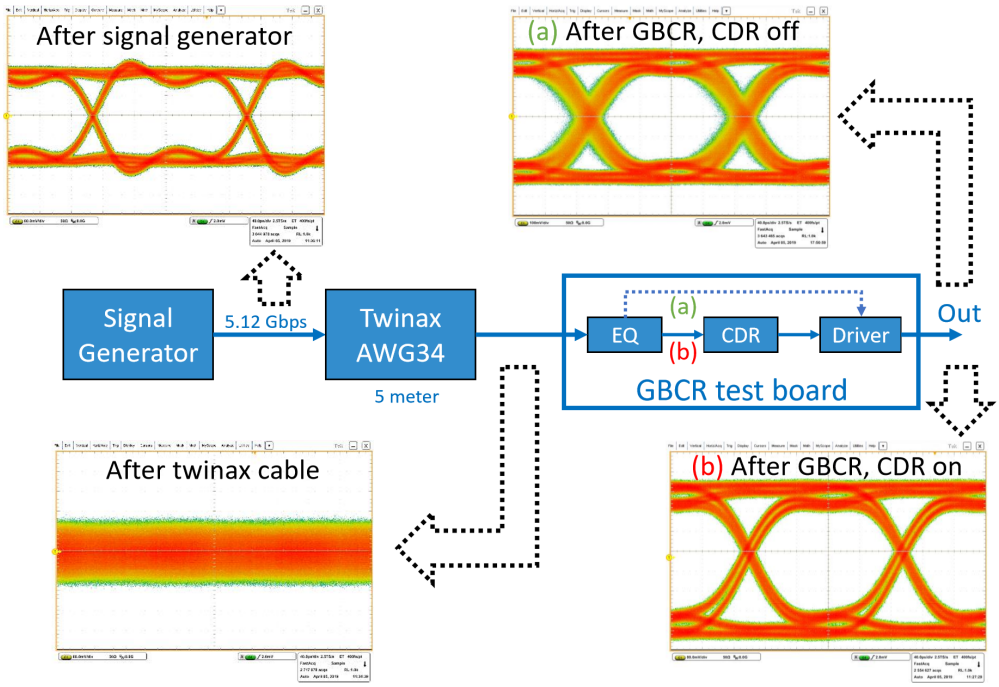


Figure 3.24: The test-stand for the GBCR (version 1) ASIC test at SMU; Test results of the GBCR (version 1) ASIC at SMU.

3.3.2.3. Test of the Pixel Detector Data Transmission Chain

The test at SLAC emphasizes the system-level test after the integration of the whole data transmission chain. A test-stand based on Fig. 3.22 was built. The goal of the test is to make sure that the readout chip can be well configured by the command from the DAQ[14] and the data can be well received on the DAQ under an acceptable Bit-Error-Rate. Each component of the transmission chain was adjusted accordingly to achieve this goal. Cables, connectors, and adapter boards were optimized and carefully designed for less signal loss. DAQ boards were upgraded with new hardware and new firmware to help improve the data transmission stability. A commercial pre-emphasis was added to the downlink for the command signal to reach the front-end readout chip with good quality. This test achieved a Bit-Error-Rate of less than 10^{-10} in the data transmission system. It gave valuable feedback to the GBCR ASIC version 2 design where the pre-emphasis in the downlink is added back for better quality of signal transmission.

[14] An FPGA Mezzanine Card (FMC) card plugged in the Xilinx KCU 105 board serves as the DAQ system.

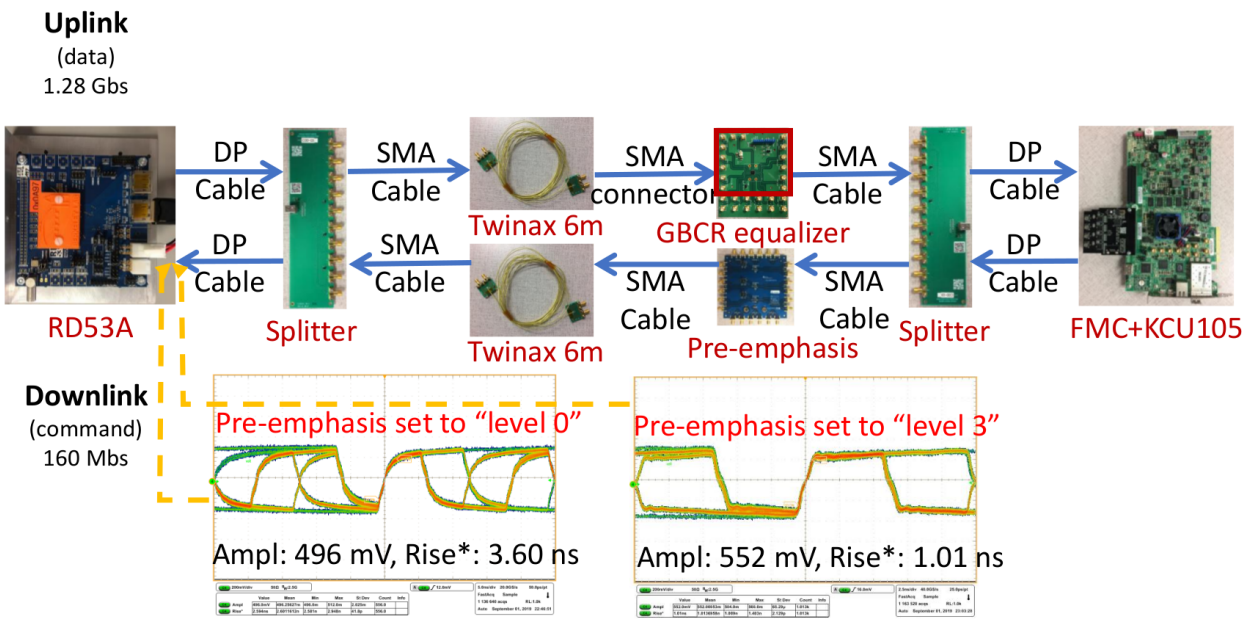


Figure 3.25: The data transmission chain test at SLAC. The Bit-Error-Rate of the system achieves a level of less than 10^{-10}.

CHAPTER 4

Data and Monte Carlo

This chapter describes the data and Monte Carlo samples used in the VHbb analysis. I contributed to the data-taking by taking trigger and run-control shifts, and to the $t\bar{t}$ filter study by generating Monte Carlo samples for the test.

4.1. Data

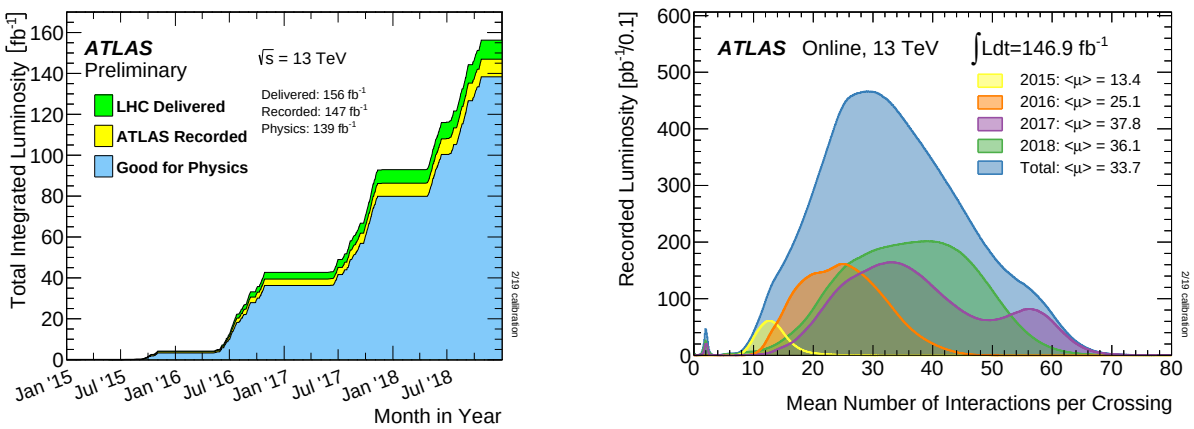


Figure 4.1: (left) The total integrated luminosity delivered by LHC and recorded by ATLAS during Run-2 at $\sqrt{s} = 13$ TeV; (right) the distribution of the mean number of interactions per bunch crossing after weighted by luminosity. Data are recorded at stable beams (special runs and machine commissioning periods are included as well) [62].

The data-set used in this analysis is the full Run 2 pp collision data recorded at a center-of-mass of 13 TeV during 2015-2018 data-taking. The integrated luminosity of the data qualified as being "good for physics" (explained in the next paragraph) is 139 fb^{-1}. The bunch spacing of the LHC is 25 ns. The distribution of the mean number of interactions per

crossing weighted by luminosity, denoted μ[1], is shown in Fig. 4.1 (right). The luminosity is primarily measured using the LUCID-2 detector [63], and the uncertainty of the integrated luminosity of Run-2 data-set is measured to be 1.7% [57] (c.f. Section 3.2.6).

The data-taking of the ATLAS detector starts after the declare of "stable beam" by LHC, which indicates that stable pp collisions are achieved. An "run" in ATLAS refers to a data-set taken by the detector in a continuously recording period. It can be further divided into luminosity blocks (LB), a period of $\sim 60\,\mathrm{s}$. In an LB, instantaneous luminosity, configurations of detector and trigger, and data quality conditions are considered constant[2]. A set of XML [66] files containing the LBs qualified for use in physics analysis, named "Good Run List", is generated after filtering out the data of anomalous conditions such as a magnet that is off, etc. The integrated luminosity of "good for physics" is calculated from Good Run List.

4.2. Monte Carlo Samples

Monte Carlo (MC) techniques are widely used to evaluate difficult integrals or to sample random variables with complicated probability density functions. At LHC, experiments require a sound understanding of signal and background processes. Fully exclusive[3] hadronic final states are required for the input to Geant4 [67] for the processes in the detector to be fully simulated. Thus, to fill in the gap between the parton level computations and the hadronic final states, Monte Carlo event generators are developed to simulate the parton showering, hadronizing, and particle decay processes (see Table 4.1 for a summary of available MC generators).

[1] At any given time interval, the actual number of interactions is Poisson distributed. The mean number of interactions per crossing is calculated as mean of the Poisson distribution of the number of interactions per crossing of each proton bunch [62].

[2] Detector and trigger status, configuration and other time-dependent information such as calibration constants, called detector "conditions," [64] are stored in the ATLAS conditions data base [65] during data-taking.

[3] In particle physics, "exclusive" means that one counts only processes with given and well defined particles in the final states; "inclusive" means that one selects all processes that include the specified products.

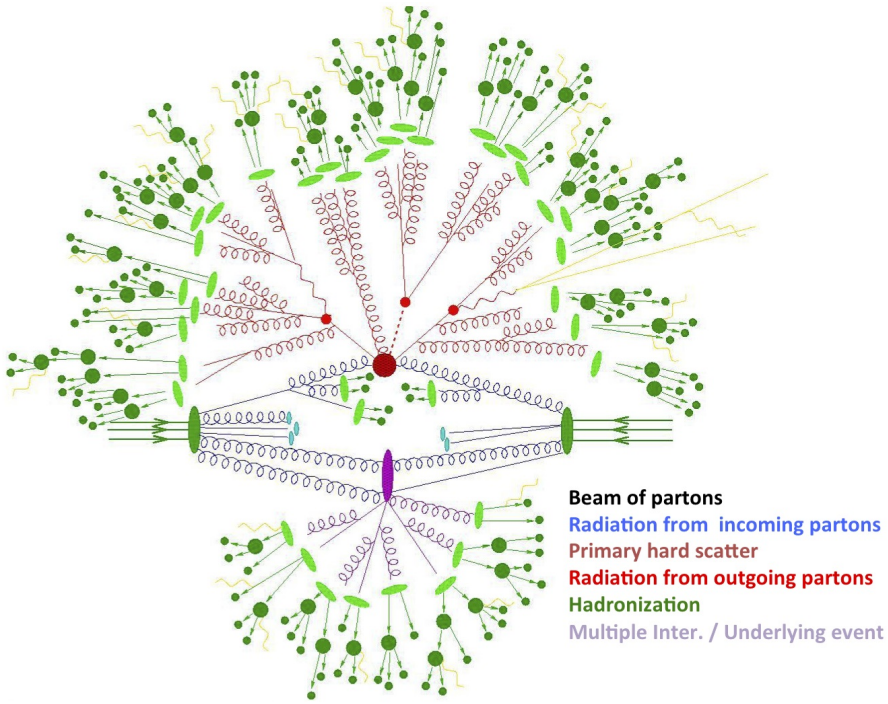


Figure 4.2: A diagram showing the physics processes of a typical proton-proton collision [68].

As shown in Fig. 4.2, the process of a proton-proton collision is complicated. It is therefore divided into a few steps to be simulated. First, before entering the hard-scattering process, incoming charged particles can radiate. This is called initial-state radiation (ISR). The scattering of the partons with large momentum transfer (either an elastic scatter or an inelastic process such as the creation of a system of large mass) while colliding, is called hard-scattering. Based on the Feynman rules, we can calculate the final states for a few numbers of particles (leading-order (LO) or next-to-leading order (NLO)) from the initial state. The emission of the final state particles is called final-state radiation (FSR).

Then, starting from the quarks and gluons final states, based on its short-distance and short-time fluctuation (from the matrix elements), the process is evolved (to have more quarks and gluons). This evolutionary process is called the parton shower. Although the colored quarks and gluons can be considered free during the collision, subsequently color

interactions will organize them into colorless hadrons; this is called hadronization. Some hadrons decay extremely quickly and many measurements from previous experiments have contributed to the hadron decay model. The detector only observes the hits left by the particles during hadron decay; these hits can be reconstructed as a bunch of tracks and neutral particles from the same vertex, which is called a "jet." However, in addition to the hard-scattering process (and its associated ISR and FSR), there are also other processes happening during the proton-proton collision; these processes together are called the "underlying event" (UE).

Hard process	Pythia [69]	Herwig [70]	Sherpa [71]	Powheg [72]	MadGraph5_aMC@NLO [73]		
Matrix Element calculator	Pythia	Herwig	Sherpa	Powheg	MadGraph5_aMC@NLO		
Parton shower	Pythia	Herwig	Sherpa				
NLO + ME matching & merging			Sherpa	Powheg	MadGraph5_aMC@NLO		
Underlying event	Pythia	Herwig	Sherpa				
Hadronization	Pythia	Herwig	Sherpa				
Hadron and tau decay		Herwig	Sherpa			EvtGen [74]	TAUOLA [75]

Table 4.1: A table summarizing the Monte Carlo generators used for the simulation of different processes.

4.2.1. VH Signal

The VH signal processes have 2 main production modes: $qq \to ZH$ and WH, as well as $gg \to ZH$. There are 3 main decay modes: $ZH \to \nu\nu b\bar{b}$, $ZH \to llb\bar{b}$, $WH \to l\nu b\bar{b}$, where $l = e$ or μ. All qq-initiated production processes are simulated using the Powheg generator [76] with the Multiscale Improved NLO procedure (MiNLO) [77, 78], interfaced to Pythia 8 [79] applying the AZNLO tune[4] [80] with NNPDF3.0[5] [81]. For the gg-initiated ZH process, Powheg (LO QCD) was interfaced to Pythia 8 applying the AZNLO tune with NNPDF3.0 for both matrix element (ME) and parton shower (PS). Alternative MC

[4]In the Monte Carlo simulations, to achieve the best description of the data, the parameters of simulations are adjusted. This is called a "tune."

[5]Parton Distribution Functions (PDFs), which describes the momentum distribution of the partons within the proton, are necessary inputs to almost all theory predictions for the hadron colliders.

samples are generated to check the signal modeling and study systematic uncertainties. The configurations of the MC samples are summarized in Table 4.2.

Systematic	VH signal	ME generation and matching	Parton shower	Hadronization	Underlying event
nominal	$qq \to WH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, NNPDF3.0 (NLO)		
	$qq \to ZH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, NNPDF3.0 (NLO)		
	$gg \to ZH$	Powheg (LO QCD) NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, NNPDF3.0 (NLO)		
Parton shower and underlying event	$qq \to WH$	Powheg + MiNLO NNPDF3.0 (NLO)	Herwig 7 H7UE tune, MMHT2014 (LO)		
	$qq \to ZH$	Powheg + MiNLO NNPDF3.0 (NLO)	Herwig 7 H7UE tune, MMHT2014 (LO)		
	$gg \to ZH$	Powheg + LO QCD NNPDF3.0 (NLO)	Herwig 7 H7UE tune, MMHT2014 (LO)		
Multiple Parton Interaction	$qq \to WH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, MPI cut-off: 1.91-2.05 GeV		
	$qq \to ZH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, MPI cut-off: 1.91-2.05 GeV		
FSR	$qq \to WH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, $\mu_R = 0.5$, $\mu_F = 0.5$; or $\mu_R = 2$, $\mu_F = 2$		
	$qq \to ZH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, $\mu_R = 0.5$, $\mu_F = 0.5$; or $\mu_R = 2$, $\mu_F = 2$		
ISR	$qq \to WH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, ISR cut-off: 0.5-3.0 GeV Primordial k_T: 0.5-2.5 GeV [80]		
	$qq \to ZH$	Powheg + MiNLO NNPDF3.0 (NLO)	Pythia 8 AZNLO tune, ISR cut-off: 0.5-3.0 GeV Primordial k_T: 0.5-2.5 GeV		

Table 4.2: A table summarizing the VH signal samples generated for the systematic uncertainty estimation. Note: μ_R is the renormalization scale and μ_F is the factorization scale [82–93].

All samples are normalized to the best theoretical prediction of the cross-section of corresponding processes. The cross-sections for the qq-initiated WH and ZH processes are calculated at NNLO in QCD and NLO in EW [82–88]. The cross-section for the gg-initiated ZH process is calculated at NLO in QCD, with the resummation of next-to-leading loga-

rithmic (NLL) soft gluon terms included [89–93]. The WH signal samples are normalized to the production cross-section of 1.37 pb. The ZH signal samples are normalized to 0.88 pb.

4.2.2. Vector Boson + Jets

The production a W or Z boson in association with jets is one of the main backgrounds for all lepton channels. The V+jet processes are simulated with Sherpa 2.2 [94–98] interfaced with NNPDFs [99].

Sherpa is able to model the large jet multiplicities through the combination of different ME with different parton multiplicities (maximally two extra partons included in ME at NLO, and 3 or 4 extra partons included in QCD at LO). A CKKM extended merging scheme[6] with a merging scale of $Q_{cut} = 20$ GeV[7] is used to merge different parton multiplicities [102]. The internal parton shower and underlying event models of Sherpa are used (Catani-Seymour dipole factorization formalism[8] and Lund multiple interaction model[9]).

V+jet samples are split according to the p_T^V and H_T, by introducing a cut at generation level and producing different slices in $\max(p_T^V,\ H_T)$ (p_T^V is the transverse momentum of the vector boson; H_T is the scalar sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$, and the p_{T} values of the lepton and all selected jets.). Since the final states require one, two or more b-tagged jets, to increase the MC statistics in this specific heavy-flavor enriched phase-space, different flavor filters are

[6]CKKW merging scheme is a method to combine the QCD ME and PS during the simulation of hadronic final states. It is named after the four authors [100].

[7]The shower is terminated when the virtual mass-square of the partons have fallen to the hadronization scale, $q^2 = Q_{cut}^2$ [101].

[8]The jet cross-section is calculated to the next-to-leading order in perturbative QCD for accurate predictions. The Catani-Seymour dipole factorization formalism uses an imaginative dipole formalism to construct a completely general algorithm for next-to-leading order calculations of arbitrary jet quantities in arbitrary processes [103].

[9]The Lund multiple interaction model is a model developed for the hadronic events with specifically the perturbative parton-parton scattering framework extended to the low p_{T} region [104].

used to select the flavor composition of the jets produced in association with the Vector boson. The filters used for the V+jets samples are shown in Table 4.3.

V+jets samples using MadGraph 5 [73] interfaced to Pythia 8 are also generated for estimation of the modeling uncertainties. MadGraph 5 provides a LO QCD description of the parton shower and the underlying events, merging with matrix-element calculations with different parton multiplicities. The CKKW-L merging scheme [105, 106] was applied with a merging scale of $Q_{cut} = 30$ GeV, and NNPDF2.3 (LO) was used for the ME calculation.

The cross-section of the single boson is calculated at next-to-next-to-leading order (NNLO) QCD [107]. A scale factor[10] of $k_{NNLO}^{QCD} = 0.9702$ is used to scale the $W(l\nu)$+jets samples to the NNLO prediction, and $k_{NNLO}^{QCD} = 0.9751$ for $Z(ll)$+jets samples. The k-factor of $Z(\nu\nu)$+jets samples is obtained from the correction from the $Z(ll)$ to $Z(\nu\nu)$ process [108].

Systematic	ME generation and matching	Parton shower	Hadronization	Underlying event	Filter
nominal	Sherpa 2.2, NNPDF3.0 (NLO), $\mu_{Q2} = 1$, $Q_{cut} = 20$ GeV	Sherpa 2.2, NNPDF3.0 (NLO), $\mu_R = 1$, $\mu_F = 1$			BFilter, CFilterBVeto, CVetoBVeto
variation	MadGraph 5, NNPDF2.3 (LO), $Q_{cut} = 30$ GeV	Pythia 8.2 A14 tune, NNPDF2.3 (LO), `pTdef` = 2, `pTHard` = 0, $\mu_R = 1$, $\mu_F = 1$			BFilter, CFilterBVeto, CVetoBVeto

Table 4.3: A table summarizing the V+jets samples generated for the systematic uncertainty estimation.

4.2.3. Top-Pair Production

Powheg [110] and Pythia 8 were used as the default MC generator for the $t\bar{t}$ processes. NNPDF3.0 (NLO) was used in Powheg for ME generation, interfaced to Pythia 8 applying A14 tune [111] with NNPDF2.3 (LO) for PS, hadronization, and UE [112]. `hdamp` is the resummation damping factor to control the ME/PS matching in Powheg (which regulates

[10] The cross-section of the process is theoretically calculated at NNLO QCD. However, the Matrix Element used in the Monte Carlo generator is at NLO. Thus, the ratio of the NNLO and NLO cross-section, defined as k-factor, is used to scale the Monte Carlo sample to NNLO prediction.

Filter	Description
BFilter	at least 1 b-hadron with p_T >0 GeV and $\|\eta\|$<4
	at least 1 b-hadron with p_T >5 GeV and $\|\eta\|$<2.9
CFilterBVeto	at least 1 c-hadron with p_T >4 GeV and $\|\eta\|$<3
	veto events which pass the BFilter
CVetoBVeto	veto events which pass the BFilter or the CFilterBVeto

Table 4.4: A table showing the heavy flavor filters used in the V+jets MC sample generation [109].

the high-p_{T} radiation). It is found to give the best description by setting it to 1.5 m_{top}, when comparing the MC samples with 8 TeV and 13 TeV data. `pTdef` and `pTHard` are the parameters controlling the merging between POWHEG and PYTHIA through the use of vetoed showering[11]. They are varied with `hdamp` = 1.5 m_{top}, because they are strongly correlated with the p_{T} definition used for ISR and FSR, and the procedure to calculate the matching scale. `pTdef` = 2 and `pTHard` = 0[12] are found to be optimal [112].

To increase the $t\bar{t}$ events in the selected phase space, samples are filtered at the event generation stage using truth level information. The "non-all-had" filter requires events to have at least one W boson decay leptonically. The "dilepton" filter requires both W bosons decay leptonically. In addition, MC samples with filters specific to the selection of each channel are used to reduce the MC statistical uncertainties. In the 1-lepton channel, filters are applied based on the number of leptons, the lepton kinematics and the p_T^W (the transverse momentum of the W boson) at generation level. It is required to have exactly 1 final state

[11]Both POWHEG and PYTHIA are based on a combined evolution of ISR and FSR in p_{T}-related "hardness" variables but the hardness definition differ. The mismatch between POWHEG-hardness and PYTHIA-hardness can be minimized if PYTHIA shower knows the POWHEG-hardness criterion and value. Then PYTHIA can fill the missing phase space regions through vetoed showering: let the shower sweep over the full phase space, using its PYTHIA-hardness ordering, and use the POWHEG-hardness to veto those emissions that POWHEG should already have covered [72].

[12]`pTdef` is the definition for the POWHEG-hardness criterion; `pTdef` = 2 means that the PYTHIA-hardness definitions are used. `pTHard` is the value for the POWHEG-harness criterion; `pTHard` = 0 means that the value in the "SCALUP" member (of the LHA/LHEF class) is used.

lepton with its p_{T} greater than 20 GeV and $|\eta|$ less than 3. The filtered sample is split in the slices of [0-100, 100-200, >200] GeV according to the p_{T} of the W boson. The development of this dedicated $t\bar{t}$ filter was initiated by SMU and then expanded with the work of others in the $VHbb$ analysis group.

Alternative MC samples with different configurations are generated to access the systematic uncertainty on the processes of initial state radiation, ME generation and matching, and hadronization. The summary of the configuration of alternative samples is shown in Table 4.5.

The value of the $t\bar{t}$ cross-section is 831.76 pb at a center-of-mass energy of 13 TeV for the top quark mass of 172.5 GeV. It is calculated at NNLO in QCD, with the resummation of next-to-next-to-leading logarithmic (NNLL) [113–119] soft gluon terms[13] included. All $t\bar{t}$ samples are normalized to the NNLO + NNLL cross-section.

Systematic	ME generation and matching	Parton shower	Hadronization	Underlying event	Filter
nominal	Powheg, NNPDF3.0 (NLO), `hdamp` $= 1.5 \cdot m_{top}$	Pythia 8.2 A14 tune, NNPDF2.3 (LO), `pTdef` $= 2$, `pTHard` $= 0$, $\mu_R = 1$, $\mu_F = 1$			non-all-had dilepton
ISR low variation	Powheg, NNPDF3.0 (NLO), `hdamp` $= 1.5 \cdot m_{top}$	Pythia 8.2 A14 tune (Var3c down), NNPDF2.3 (LO), `pTdef` $= 2$, `pTHard` $= 0$, $\mu_R = 2$, $\mu_F = 2$			non-all-had dilepton
ISR high variation	Powheg, NNPDF3.0 (NLO), `hdamp` $= 3 \cdot m_{top}$	Pythia 8.2 A14 tune (Var3c up), NNPDF2.3 (LO), `pTdef` $= 2$, `pTHard` $= 0$, $\mu_R = 0.5$, $\mu_F = 0.5$			non-all-had dilepton
hard scattering generation and matching	MadGraph5_aMC@NLO , NNPDF 3.0 (NLO)	Pythia 8.2 A14 tune, NNPDF2.3 (LO), `pTdef` $= 2$, `pTHard` $= 0$, $\mu_R = 1$, $\mu_F = 1$			non-all-had dilepton
hadronization model	Powheg, NNPDF3.0 (NLO), `hdamp` $= 1.5 \cdot m_{top}$	Herwig 7.0 H7UE tune, MMHT2014 (LO),			non-all-had dilepton

Table 4.5: A table summarizing the $t\bar{t}$ samples generated for the systematic uncertainty estimation. Note: "Var3c" is the variation of strong coupling in the initial state shower affecting the description of the $t\bar{t}$ gap fraction, the dijet decorrelation and the Z-boson transverse momentum.

[13]Uncertainties of PDFs and α_S were calculated using the PDF4LHC prescription [120]. PDF sets of MSTW2008 68% CL NNLO [121, 122], CT10 NNLO [123, 124] and NNPDF2.3 5f FFN [99] PDF sets, were added in quadrature to the scale uncertainty.

4.2.4. Single-top

The default single-top (s-, t- and Wt-channel) samples are generated using POWHEG with NNPDF3.0, interfaced to PYTHIA 8 applying the A14 tune with NNPDF2.3. PYTHIA 8 is further interfaced to EvtGen [125] to simulate the decays of heavy flavor particles. Samples in the s-channel [126] and t-channel are generated with a leptonic filter (where the W boson decays leptonically), while the Wt-channel [127] is generated both applying a dileptonic filter (both W bosons decay leptonically) and inclusively. Single-top samples don't make use of the POWHEG resummation damping parameter `hdamp`; therefore its variations are not considered for these samples.

There are also alternative samples in s-, t- and Wt-channels generated to assess the single-top modeling. The configurations of the alternative samples are shown in Table 4.6. A different matrix element calculation in the Wt-channel was used to generate additional MC samples as well (using the diagram subtraction (DS) scheme, rather than the diagram removal (DR)[14] scheme used in the nominal sample).

The single-top samples used in this analysis are normalized to the cross-section calculated at higher orders, specified for each channel. The cross-section of the t-channel is $\sigma_t = 136.02$ pb for top quark and $\sigma_{\bar{t}} = 80.95$ pb for anti-top quark. The cross-section of the s-channel is $\sigma_t = 6.35$ pb for top quark and $\sigma_{\bar{t}} = 3.97$ pb for anti-top quark. They are calculated at NLO QCD [130,131] for a top quark mass of 172.5 GeV. The cross-section of the Wt-channel process is $\sigma_t = 71.7$ pb for top quark plus anti-top quark processes. It is calculated at NNLO (MSTW2008 NNLO PDF [121,122]), for a top quark mass of 172.5 GeV. The samples with filters are normalized to the value of cross-section multiplied by the leptonic branching ratio.

[14] In DR, the real contribution to Wt-channels $\mathcal{R}$ is defined by eliminating the $t\bar{t}$ contribution matrix element $\mathcal{M}^{t\bar{t}}$ from the generic amplitude $\mathcal{M}$: $\mathcal{R}^{\mathrm{DR}} = \frac{|\mathcal{M}-\mathcal{M}^{t\bar{t}}|^2}{2s}$, where s is squared center-of-mass energy. In DS, the full squared amplitude $\mathcal{M}$ is kept but a local counter-term $\mathcal{C}^{\mathrm{SUB}}$ is subtracted from it to suppress the $t\bar{t}$ contribution at the cross-section level: $\mathcal{R}^{\mathrm{DS}} = \frac{|\mathcal{M}|-\mathcal{C}^{\mathrm{SUB}}}{2s}$ [127].

Systematic	ME generation and matching	Parton shower	Hadronization	Underlying event	Filter
Nominal	POWHEG NNPDF3.0 (NLO)	PYTHIA 8 A14 tune, NNPDF2.3 (LO), $\mu_R = 1$, $\mu_F = 1$			Wt-: dileptonic, inclusive t-, s-: leptonic
Low radiation (t-, s-, Wt-channel)	POWHEG CT10 PDF (NLO)	PYTHIA 6 PERUGIA2012 tune, CTEQ6L1PDF (LO), $\mu_R = 2$, $\mu_F = 2$			Wt-: dileptonic, inclusive t-, s-: leptonic
High radiation (t-, s-, Wt-channel)	POWHEG CT10 PDF (NLO)	PYTHIA 6 PERUGIA2012 tune, CTEQ6L1PDF (LO), $\mu_R = 0.5$, $\mu_F = 0.5$			Wt-: dileptonic, inclusive t-, s-: leptonic
ME calculation (Wt-channel)	POWHEG DS scheme CT10 PDF (NLO)	PYTHIA 6 PERUGIA2012 tune, CTEQ6L1PDF (LO)			Wt-: dileptonic, inclusive
Parton shower (t-, Wt-channel)	POWHEG CT10 PDF (NLO)	HERWIG ++ [128] CTEQ6L1-UE-EE-5 tune [129] CTEQ6L1PDF (LO)			Wt-: dileptonic, inclusive t-: leptonic
Matrix element (t-, Wt-channel)	MADGRAPH5_aMC@NLO CT10f4 PDF	HERWIG ++ CTEQ6L1-UE-EE-5 tune CTEQ6L1PDF (LO)			Wt-: dileptonic, inclusive t-: leptonic

Table 4.6: A table summarizing the single-top samples generated for the systematic uncertainty estimation.

4.2.5. Diboson

Diboson processes can be produced through either qq or gg processes. Some diboson processes contribute significantly to the 3 channels of the $VHbb$ analysis: $ZZ \to b\bar{b}\nu\bar{\nu}$, $ZW \to b\bar{b}l\nu$, $ZZ \to b\bar{b}ll$. Other diboson processes can give small contributions as well, by mistagging if a jet from W decay is mistagged, or one of the leptons fails to be reconstructed in $ZW \to \nu\bar{\nu}q\bar{q}$, $ZW \to llq\bar{q}$, and $WW \to l\nu q\bar{q}$.

The default MC generator for the diboson processes is SHERPA 2.2, using NNPDF3.0 for both ME calculation and PS simulation. SHERPA provides a combination of different matrix elements with different parton multiplicities to model the large jet multiplicities. The CKKW-L merging scheme [105, 106] was used in the ME calculation with a merging scale of $Q_{cut} = 20$ GeV. Alternative samples were generated to determine the modeling uncertainties. The configurations of those samples are summarized in Table 4.7.

SHERPA provides an NLO calculation for the diboson cross-sections for 0 or 1 extra partons and a LO calculation for 2+ extra partons in hard scattering. Each diboson sample is normalized to its SHERPA predicted cross-section.

Systematic	ME generation and matching	Parton shower	Hadronization	Underlying event
Nominal	SHERPA, NNPDF3.0 (NLO), Q_{cut} = 20 GeV	SHERPA NNPDF2.3 (LO)		
Scale variation	SHERPA, NNPDF3.0 (NLO), Q_{cut} = 20 GeV	SHERPA, NNPDF2.3 (LO) μ_R = 0.5 or 2, μ_F = 0.5 or 2		
Matrix element & parton shower	POWHEG CT10 PDF (NLO)	PYTHIA 8 AZNLO tune CTEQ6L1PDF (LO)		
Parton shower	POWHEG CT10 PDF (NLO)	HERWIG ++, CTEQ6L1-UE-EE-5 tune CTEQ6L1PDF (LO)		

Table 4.7: A table summarizing the diboson samples generated for the systematic uncertainty estimation.

Before fully utilizing the data, the Monte Carlo samples here are used to build the reconstruction algorithms, model the physics processes, and improve the statistical analysis techniques. The procedure and configuration of the MC simulation explained in this chapter establish a solid base for the systematic uncertainty estimation described in Sec. 8.2.2.

CHAPTER 5

Physics Object Reconstruction

This chapter describes the trigger used in the VHbb analysis, as well as the reconstruction of lepton, jet and missing transverse momentum.

5.1. Trigger Selection

The final state of the 0-lepton channel ($\nu\nu qq$) basically consists of $E_{\mathrm{T}}^{\mathrm{miss}}$ due to neutrinos and jets from the Higgs decay. To select the $\nu\nu qq$ process based on this signature, the lowest unprescaled[1] $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers were used during each data collection period from 2015 to 2018, taking into account the balance of trigger efficiency[2] and simplicity. The $E_{\mathrm{T}}^{\mathrm{miss}}$ at trigger level is determined only based on the energy measured in the calorimeter.

The final state of the 1-lepton channel ($\ell\nu qq$) consists of one lepton plus $E_{\mathrm{T}}^{\mathrm{miss}}$ and jets. Therefore, the electron final-state of the 1-lepton channel used the lowest unprescaled single-electron triggers (no $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement) in each data-taking period. Due to the inefficiencies in single-muon triggers, $E_{\mathrm{T}}^{\mathrm{miss}}$-only triggers used in the 0-lepton channel were applied to the selection of the 1-lepton muon channel.

The 2-lepton channel final state ($\ell\ell qq$) consists of two leptons and jets. Thus, to select those $e^{\pm}e^{\mp}/\mu^{\pm}\mu^{\mp}$ events, the single-electron triggers mentioned before and the lowest unprescaled single-muon triggers were used in each data-taking period.

[1]Prescaling is used to reduce the trigger output rate for certain stages (L1 item or HLT trigger chains). A trigger chain without a prescale applied is called "unprescaled".

[2]The trigger efficiency is defined as: $\epsilon_{\mathrm{trigger}} = \frac{N_{\mathrm{trigger}}}{N_{\mathrm{offline}}}$, where N_{trigger} is the number of triggered object candidates (like electron) and N_{offline} is the number of identified and reconstructed object candidates that would have been selected exclusively by offline approaches. Both N_{trigger} and N_{offline} are counted from the same number of events [132].

Channel	Final state	Trigger
0-lepton	$\nu\nu qq$	lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger
1-lepton	$l\nu qq$	lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for μ lowest single-electron trigger for e
2-lepton	$llqq$	lowest single-electron trigger and lowest single-muon trigger

Table 5.1: The triggers used in the 0-, 1- and 2-lepton channels.

5.1.1. Recommended Trigger Chains

The event selections implemented at the HLT are referred to as chains, since a sequence of selection algorithms are chained together. The complete set of all trigger chains is called "trigger menu." Each chain specifies a set of L1 trigger seeds which if present will activate the chain. The L1 trigger menu consists of 512 single items and combinations, and the HLT menu is composed of O(1000) chains [133].

The trigger menu is usually named in the form of `HLT_<HLT objects and thresholds>_L1<L1 OBJECTS AND THRESHOLDS>`. For instance, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, named `HLT_xe70_L1XE50`, means that this trigger chain is seeded using the region of interest identified by `L1_XE50`[3] calorimeter trigger, calibrated at the EM scale with a threshold of 50 GeV for $E_{\mathrm{T}}^{\mathrm{miss}}$ at L1 and a threshold of 70 GeV at HLT. The trigger chains of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, single-electron trigger and single-muon trigger used in Run 2 are shown in Table 5.2, Table 5.3 and Table 5.4.

[3] $E_{\mathrm{T}}^{\mathrm{miss}}$ is denoted in the trigger as `XE`. The uppercase is used in the L1 trigger and the lowercase is used in the HLT trigger.

Trigger Name	Data-taking Period	Threshold (GeV)
HLT_xe70_L1XE50	2015	70
HLT_xe90_mht_L1XE50	2016 (A-D3)	90
HLT_xe110_mht_L1XE50	2016 ($\geq$ D4)	110
HLT_xe110_pufit_L1XE55	2017	110
HLT_xe110_pufit_xe70_L1XE50	2018	110

Table 5.2: $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in the VHbb analysis during the 2015-2018 data-taking period [134].

Trigger Name	Data-taking Period	Threshold (GeV)
HLT_e24_lhmedium_L1EM20VH	2015	24
HLT_e60_lhmedium	2015	60
HLT_e120_lhloose	2015	120
HLT_e26_lhtight_nod0_ivarloose	2016 – 2018	26
HLT_e60_lhmedium(_nod0)	2016 – 2018	60
HLT_e140_lhloose(_nod0)	2016 – 2018	140
HLT_e300_etcut	2018	300

Table 5.3: Single-electron triggers used in the VHbb analysis during the 2015-2018 data-taking period [134].

Trigger Name	Data-taking Period	Threshold (GeV)
HLT_mu20_iloose_L1MU15	2015	20
HLT_mu50	2015 – 2018	60
HLT_mu26_ivarmedium	2016 – 2018	26

Table 5.4: Single-muon triggers used in the VHbb analysis during the 2015-2018 data-taking period [134].

5.2. Physics Object Reconstruction

When particles are produced from pp collisions and travel outward, hits are recorded by the tracking detectors and energy is deposited in the calorimeters. The tracking algorithms use the hits from pixel layers as seeds and apply a staged pattern-recognition approach, which includes a loose track candidates search giving a number of combinatorial track candidates, followed by a stringent ambiguity-solver comparing and rating the individual tracks[4]. The interaction vertices are reconstructed from the fit to at least 2 selected tracks. The one with the largest $\sum p_T^2$ of associated tracks is selected as the primary vertex. The tracks, primary vertex and the signals of calorimeters are then used together for the reconstruction of momentum, mass and charges, from which the identity of the particles can be worked out. Those reconstructed particles with their measured physics quantities are called "physics objects."[5] Leptons, photons and jets are examples of physics objects.

5.2.1. Lepton

5.2.1.1. Electron

An electron is generally defined as an object consisting of a cluster built from the energy deposits in the calorimeter and a matched track (tracks) from the inner detector [135]. The reconstruction algorithm first selects the cluster of energy deposits measured in topologically connected EM and hadronic calorimeter cells [136], denoted topo-clusters. It then performs a refit of the Inner Detector tracks to account for the bremsstrahlung[6] and matches them to

[4]In detail, there are inside-out, outside-in, and some special algorithms used for the track reconstruction.

[5]A physics object is a representation of a physical particle or system of particles (e.g. jets and $E_{\mathrm{T}}^{\mathrm{miss}}$) that uses a combination of detector information to stand-in for the real particle or system.

[6]When an electron interacts with the material of the Inner Detector, it radiates a bremsstrahlung photon. A significant amount of energy can be lost by the electron due to bremsstrahlung. The photon radiated from the electron can convert to an electron and a positron whose interactions with the detector material can generate multiple tracks in the inner tracker.

the selected topo-clusters. The algorithm also builds conversion vertices out of the refitted tracks and matches them to the selected topo-clusters. The matched topo-clusters are used to seed the electron supercluster[7] building. After applying the initial position corrections and energy calibrations to the resulting superclusters, the supercluster-building algorithm matches tracks to the electron superclusters (and conversion vertices to the photon superclusters). Since an object can be reconstructed as both an electron and a photon due to the similarity of their electromagnetic interactions, an ambiguity resolution is performed to remove overlaps between object candidate lists. The final electrons (and photons) are then built and their energy calibrated[8]. The electron identification is performed using a likelihood-based method. It allows the discriminating variables to separate electrons (or photons) from the background and make the quality cuts for physics analyses [137].

The "loose" electrons are used in the $VHbb$ analysis, which requires the electrons p_{T} larger than 7 GeV, $|\eta|$ in the range of 2.47, and that the "loose" likelihood criteria[9] is passed. A requirement on the impact parameter[10] is applied to reject tracks from pile-up[11] and the loose track isolation (FCLoose) is applied to reduce jet-faking electrons. In the WH 1-lepton channel, tighter electron selections – "medium" electron identification and "tight" likelihood isolation, are used for the reduction of multi-jet background [135].

---

[7]A supercluster is a dynamic, variable-size cluster to recover low energy photons radiated due to bremsstrahlung interactions in the Inner Detector, and connect them to their associated electron or converted photon [137].

[8]The electron energy calibration is performed with a MC-based MVA, trained on sensitive cluster variables, like the energy in different layers of calorimeters. The energy is then corrected with data driven techniques [137].

[9]The electron identification efficiencies are are 93%, 88% and 80% on average for the "loose," "medium," and "tight" operating points. To reduce the background contamination in the selected data, probe electrons are required to satisfy a "very loose" requirement on the likelihood discriminant. This requirement rejects $\sim$ 95% of the background [135].

[10]d_0 is defined as the transverse parameter relative to the beam-line (BL).

[11]For every 25 ns, there's a proton beam-crossing in the LHC. There are multiple p-p collisions during a beam crossing but only the one whose primary vertex has the largest sum of p_{T} is triggered and selected. Thus, additional particles resulting from the previous p-p collisions can be collected during the time window of the selected one; those tracks resulting from them are called pile-up tracks.

Electron Selection	p_{T}	η	Identification	d_0^{sig} w.r.t. BL	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
$VH - loose$	>7 GeV	$\lvert\eta\rvert < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	FCLoose
$ZH - signal$	>27 GeV	$\lvert\eta\rvert < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	FCLoose
$WH - signal$	>27 GeV	$\lvert\eta\rvert < 2.47$	LH Tight	< 5	< 0.5 mm	FixedCutHighPtCaloOnly

Table 5.5: Electron selection requirement. Note: LHLoose and LHTight are two criteria of the electron likelihood identification; FCLoose and FixedCutHighPtCaloOnly are the operating points of the electron isolation which have a fixed requirement either on the calorimeter or the track isolation variables (or both) [109].

5.2.1.2. Muon

Muons are first reconstructed independently in ID and MS. All information in ID and MS are then used together to reconstruct the muon tracks for physics analyses. In ID, muons are reconstructed like any other charged particles [138, 139]. In MS, muon reconstruction starts with a search for patterns among hits throughout the spectrometer. Segments are formed by fitting the close hits in the same chamber to the trajectory. Muon track candidates are then built by fitting together hits to form segments in different layers[12]. Four muon types are defined based on which subdetectors (ID, MS or calorimeters) are used in the reconstruction. The primary type used in ATLAS is called "combined muon", which is reconstructed from global refit using the hits from both ID and MS [140].

In the $VHbb$ analysis, muons are required to have $|\eta| < 2.7$, $p_{\mathrm{T}} > 7$ GeV and small impact parameters. "Loose" muons are selected using a generous set of quality criteria that select 96.7% of true muons in the low momentum region ($4 < p_{\mathrm{T}} < 20$ GeV) and 98.1% in the high momentum region ($20 < p_{\mathrm{T}} < 100$ GeV)[13], as determined in simulation [140], and a loose track isolation working point. In the 1-lepton channel, the "medium" muon identification criteria and the "tight" track isolation working point are used [109]. This has a selection

[12]The same segments can be used to build several track candidates in the beginning. Later an overlap removal algorithm selects the best assignment to a single track, or allows for the segments to be shared between two tracks.

[13]Four muon identification selections ("loose," "medium," "tight" and "high-p_{T}") are provided for the needs of different physics analyses [140].

efficiency of 95.5% on true muons in the low-p_{T} region and 96.1% in the high-p_{T} region, determined in simulation [140].

Muon Selection	p_{T}	η	Identification	d_0^{sig} w.r.t. BL	$\lvert\Delta z_0 \sin\theta\rvert$	Isolation
$VH-loose$	>7 GeV	$\lvert\eta\rvert<2.7$	Loose quality	<3	<0.5 mm	FixedCutLoose
$ZH-signal$	>27 GeV	$\lvert\eta\rvert<2.5$	Loose quality	<3	<0.5 mm	FixedCutLoose
$WH-signal$	>25 GeV(27 GeV for 75< p_{T}^{V}<150 GeV)	$\lvert\eta\rvert<2.5$	Medium quality	<3	<0.5 mm	FixedCutHighPtTrackOnly

Table 5.6: Muon selection requirements [109].

5.2.1.3. Hadronic Tau

Over 90% of hadronic tau decays occur through just five dominant decay modes, which yield one or three charged hadrons, up to two neutral pions and a tau neutrino [141]. The hadronic tau reconstruction algorithm is seeded by calorimeter energy deposits which have been reconstructed as individual jets. Such jets are formed from topo-clusters using anti-k_t algorithm [142] with a spacing parameter of $R = 0.4$, using topo-clusters as inputs. To seed a hadronic tau candidate, a jet must fulfill the requirement of $|\eta| < 2.5$, $p_{\mathrm{T}} > 20$ GeV, and being outside of $1.37 < |\eta| < 1.52$ (the transition region between the barrel and end-cap electromagnetic calorimeter) [143].

In an event with pile-up, the default primary vertex does not always correspond to the vertex at which the tau lepton was produced. The tau vertex association algorithm uses all the tracks (with $p_{\mathrm{T}} > 1$ GeV) in a cone $\Delta R < 0.2$ around the seed jet direction to identify the primary vertex associated with the tau[14]. The p_{T} of these tracks is summed and the candidate of tau vertex to which the largest fraction of the p_{T} sum is matched is chosen as the tau vertex [144]. This vertex is then used to determine the hadronic tau direction, to associate

[14]Tau vertex is another primary vertex reconstructed using all the tracks in a cone $\Delta R < 0.2$ associated with the tau. So even though a hadronic tau decay only has 1 or 3 tracks, all the tracks in the cone $\Delta R < 0.2$ together are capable of the primary vertex reconstruction.

tracks and to build the coordinate system in which identification variables are calculated. The three-momentum of the hadronic tau is calculated by computing η and ϕ of barycenter[15] of the topo-clusters of the jet seed, assuming a mass of zero for each constituent. The four-momentum of all clusters in the region $\Delta R < 0.2$ around the barycenter are recalculated using the tau vertex coordinate system and summed to have the hadronic tau direction. The hadronic tau mass is defined to be zero [143]. The energy of the hadronic tau is then calibrated using a dedicated schemes.

The tau identification algorithm uses Boosted Decision Tree (BDT)[16] to reject backgrounds from quark- and gluon-initiated jets [146, 147]. The BDT for the tau candidates associated with one and three tracks are trained separately [148]. Three working points ("loose," "medium" and "tight") corresponding to three tau identification efficiencies are provided[17]. In the VHbb analysis described in this thesis, the "medium" working point of the tau identification is used for the tau-veto study in Sec. 7.2 [149].

5.2.2. Jet

Quarks and gluons manifest themselves in the detector as a directional spray of tracks and calorimeter deposits, known as jets. This is true of b-quark initiated jet (b-jets) produced by the Higgs boson in our signal processes and the jets initiated from other quarks or gluons originating from the background processes. Jets play an important role in identifying event topologies and kinematics consistent with our signal processes [134].

[15]The barycenter is computed as the η and ϕ (weighted by energy) of all calorimeter cells within the topo-clusters [145].

[16]Definition of BDT and details about how it works can be seen at Sec. 8.1.1.

[17]The efficiency is designed to be independent of p_{T}. The target efficiencies are 60%, 55% and 45% for the generated 1-track "loose," "medium" and "tight" working points, and 50%, 40% and 30% for the corresponding generated 3-track target efficiencies [149].

5.2.2.1. *Standard Jet Collections*

The primary jet collection used in this analysis is reconstructed from topological calorimeter cells (topo-clusters) [150] using the anti-k_t algorithm. Topo-clusters are built from neighboring calorimeter cells containing a significant energy above a noise threshold that is estimated from measurement of calorimeter electronic noise and simulated pile-up noise. The energies of calorimeter cells are measured at the electromagnetic energy scale, and corresponds to the energy deposited by electromagnetically interacting particles. Each cluster[18] is passed as an input to the anti-k_t algorithm with a radius parameter of $R = 0.4$ to form a jet; this collection is referred to as an AntiKt4EMTopoJet (see Table 5.7) [152].

Jet Category	Selection Requirements
Forward Jets	jet cleaning $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ $2.5 \leq \lvert\eta\rvert < 4.5$
Signal Jets	$p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $\lvert\eta\rvert < 2.5$ jet cleaning JVT > 0.59 for $\lvert\eta\rvert < 2.4$ and > 0.11 for $2.4 < \lvert\eta\rvert < 2.5$ if ($p_{\mathrm{T}} < 120\,\mathrm{GeV}$)

Table 5.7: `AntiKt4EMTopoJets` selection requirements[19] The jet cleaning is applied via the JetCleaningTool, that removes events in regions corresponding to hot calorimeter cells [109].

Jet from non-collision backgrounds or noise in the calorimeters are removed using the jet cleaning criteria [152]. Jets in the central region ($|\eta| < 2.5$, named Signal Jets) are required to have $p_{\mathrm{T}} > 20$ GeV. Jets in the forward region ($2.5 < |\eta| < 4.5$, named Forward Jets) are required to have $p_{\mathrm{T}} > 30$ GeV. A likelihood-based discriminant named the Jet Vertex Tagger (JVT) [144, 153] is used to remove jets with $p_{\mathrm{T}} < 120$ GeV and $|\eta| < 2.5$, which arise from

[18] Those topo-clusters are three-dimensional, massless and positive-energy [151].

pile-up. The JVT score is initially calculated on uncalibrated jets during the reconstruction and is recalculated for each calibrated jet for further selection [109].

5.2.2.2. b-tagging

The jets from the Higgs boson in our signal process should originate from b-quarks. It is therefore necessary to distinguish b-jets from light-flavor jets (u-, d-, s-quark and gluon) and c-jets, to maximize signal sensitivity. This procedure is commonly referred to as b-tagging, as illustrated in Fig. 5.1 [134].

The standard algorithm for b-tagging used in the analysis is the MV2 algorithm (see Fig. 5.2) [155]. This algorithm combines the output of the low-level tagging algorithms, such as the log-likelihood ratios from IP2D and IP3D[20] [156], secondary vertices from SV1[21] [157], and information about the relationship between the primary and secondary vertices from JetFitter[22] [158]. Those information together with the p_{T} and $|\eta|$ of the jets are input to a boosted decision tree to produce final discriminant. The output of MV2 algorithm uses a score between `-1` and `1` to give the likelihood for a jet to be a real b-jet. The algorithm is trained on a sample composed of $t\bar{t}$ events (with at least one lepton from leptonic decay of a W) and hadronically decaying Z' events to compensate the steeply falling $t\bar{t}$ spectrum above 250 GeV. The training has been performed with b-jets as signal and a mixture of light-jets and c-jets as background [134]. The average efficiency of MV2 is tuned to be 70% for b-jets in the $t\bar{t}$ MC samples, corresponding to light-jet mis-identification efficiency of 0.3% and c-jet mis-identification efficiency of 12.5% [159].

[20] IP2D and IP3D are two complementary impact parameter-based algorithms. The IP2D tagger makes use of the signed transverse impact parameter significance of tracks to construct a discriminating variable, whereas IP3D uses both the track signed transverse and the longitudinal impact parameter significance in a two-dimensional template to account for their correlation [156].

[21] SV1 is a secondary vertex tagging algorithm which reconstructs a single displaced secondary vertex in a jet [157].

[22] JetFitter is a topological multi-vertex algorithm which exploits the topological structure of weak b- and c-hadron decays inside the jet and tries to reconstruct the full b-hadron decay chains [158].

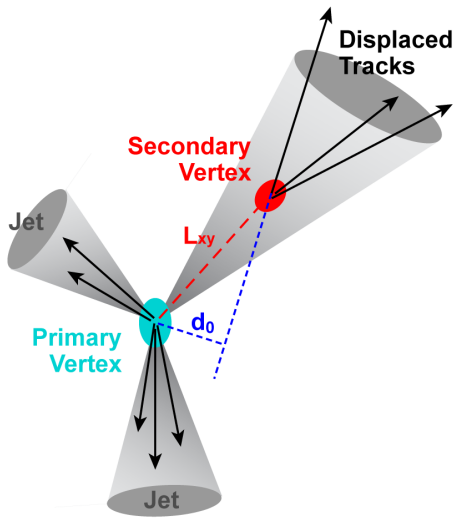


Figure 5.1: Illustration of the physics of a b-jet [154].

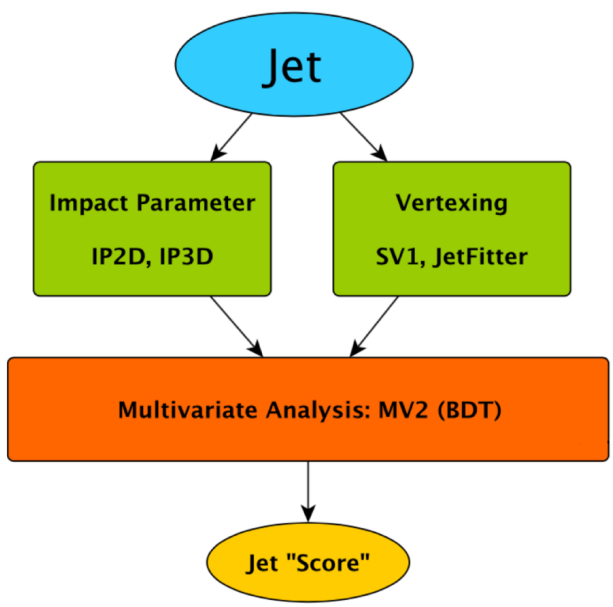


Figure 5.2: The structure of the MV2 b-tagging algorithm.

5.2.2.2.1 Truth Tagging and Hybrid Truth Tagging

In the Monte Carlo simulation, jets can be "truth-tagged" to gain in simulation statistics, especially for the processes containing c- or light-jets in the final states. A jet is labeled by its true origin in the simulation (e.g. "b", "c", or "light"). Its p_{T} and η are used to determine the probability it would have been identified as a b-jet using the tagging algorithm. That probability is assigned to the jet as a weight in an event[23]. Events are thus weighted, rather than being eliminated by a strict cut on the b-jet tagging algorithm.

To further reduce the discrepancies between the truth-tagged and direct-tagged events, a hybrid tagging approach is also adopted in this analysis, which applies truth tagging to non-b-jets and direct tagging to b-jets, in an event.

5.2.2.3. b-jet Energy Correction

The b-tagged jets are calibrated with the standard jet energy scale, which restores the jet energy scale to that of truth jets reconstructed at the particle-level energy scale [160]. In addition, corrections based on the unique features of b-jets such as the semileptonic decay, secondary vertex and B hadron mass, are applied to improve the energy measurement scale and resolution. A summary of the corrections used in each channel of the $VHbb$ analysis is shown in Table 5.8.

[23]The probability of a jet being identified as a b-jet is derived from the b-tagging efficiency which depends on the jet p_{T} and η. Thus, the probability of an event requiring 2 b-tagged jets out of 3 jets can be determined through the possible combinations of the b-tagged and non-b-tagged jets.

	muon-in-jet	PtReco	Kinematic Fit
0-lep, 2-jet/3-jet	Yes	Yes	
1-lep, 2-jet/3-jet	Yes	Yes	
2-lep, 2-jet/3-jet	Yes		Yes
2-lep, 4plus-jet	Yes	Yes	

Table 5.8: b-jet energy corrections applied in each channel of the $VHbb$ analysis.

5.2.2.3.1 Muon-in-jet Correction

It is found that about 10% of the b-jets have reconstructed muons which deposit a few GeV of energy in the calorimeter[24]. If a muon of medium quality and $p_{\mathrm{T}} > 5$ GeV is found within a variable cone[25] around the jet axis, the four-momentum of the muon is added to that jet, after subtracting the energy loss in the calorimeter. No isolation criteria is applied. The muon closest to the jet axis is chosen, if more than one muon is found in the cone [161]. The systematic uncertainties on the muon-in-jet correction are demonstrated to be negligible [134].

5.2.2.3.2 PtReco Correction

The unique features of b-jets described before can cause additional resolution effect compared to the standard jets. A p_{T}-dependent correction (PtReco) is used to correct the response to the low p_{T} jets with leptonic or hadronic decays of heavy-flavor hadrons (see Fig. 5.3). The correction factor is derived from the p_{T} ratio of the reconstructed b-tagged jets (with all previous corrections applied) and the corresponding truth jets (produced by clustering final state particles of MC truth record, including neutrinos and muons) [161].

[24]Light jets produce muons from charged pion decay in-flight. Charm and bottom jets produce them from semi-leptonic decays.

[25]The variable cone is defined as $\Delta\mathrm{R}(\mathrm{jet}, \mu) < \min(0.4, 0.04 + 10\ \mathrm{GeV}\ /\ p_T^{\mu})$ [134].

The b-jets with muons and without muons are considered separately; only the b-jets with muons are corrected. The mean value of the p_{T} ratio is calculated in each p_{T} slice and used as the correction factor. The calculation is done in the nominal MC samples and applied to all systematic variation samples including the latest MC b-jet energy scale uncertainty. Therefore, no additional uncertainty is applied [134]. Together with the muon-in-jet correction, the PtReco correction improves the resolution of the dijet mass by $\sim$ 10 - 20% [161].

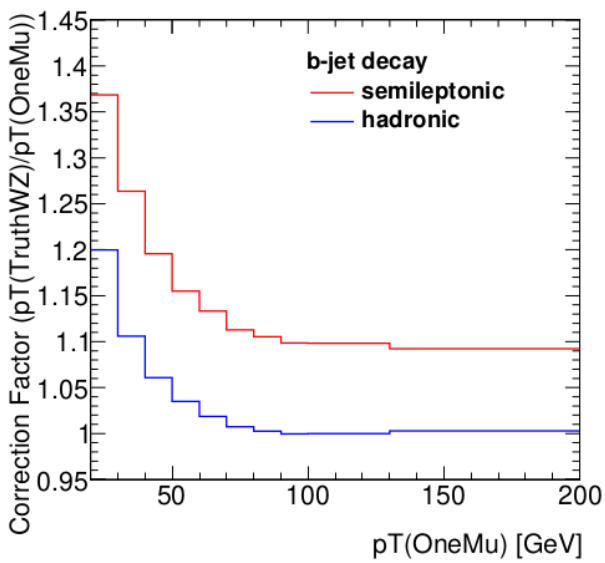


Figure 5.3: The PtReco correction factor for the b-jets with muon is marked as red and for those without muon is marked as blue. The jets after muon-in-jet correction are denoted as "OneMu".

5.2.2.3.3 Kinematic Fit

In the 2-lepton channel, the $Z \to ll$ is fully reconstructed. Since there is no neutrino in such an event and the lepton resolution (typically 1% level) is much better than the jet resolution (typically 10% level), the b-jet energy can be corrected by constraining the $llbb$ system to be balanced in the transverse plane [134]. The Kinematic Fit correction is achieved with a negative log likelihood minimized to obtain the fit value, as the equation

shown below [134]. This log likelihood function consists of a Breit-Wigner constraint on the invariant mass of dilepton, a Gaussian constrain on each component of p_{T} of the $llbb$ system, dedicated transfer functions relating the truth-jet p_{T} to their reconstructed ones (after muon-in-jet correction, but no PtReco correction), and a prior built from the expected truth-jet p_{T} spectrum in ZH events (which has a role similar to the PtReco correction) [162]. This correction improves the resolution of m_{bb} by up to 40% in the 2-lepton channel[26](see Fig. 5.4) [159].

$$-2\ln L = \underbrace{\sum_i \frac{(p_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{fit}})}{\sigma^2} - \sum_j 2\ln L(p_{\mathrm{T}},\ p_{\mathrm{T}}^{\mathrm{fit}})}_{\text{Experimental leptons, jet and MET resolution}} + \underbrace{\frac{\sum_{i,j} p_{\mathrm{x}}^2}{\sigma_{\mathrm{bal}}^2} + \frac{\sum_{i,j} p_{\mathrm{y}}^2}{\sigma_{\mathrm{bal}}^2}}_{\text{Transverse balance system}} + \underbrace{2\ln\{(m_{ll}^2 - m_Z^2)^2 + m_Z^2\Gamma_Z^2\}}_{\text{Z-mass Breit-Wigner constraint}}$$

[26]There is no improvement in 4plus-jet events therefore Kinematic Fit is not used in 4plus-jet events.

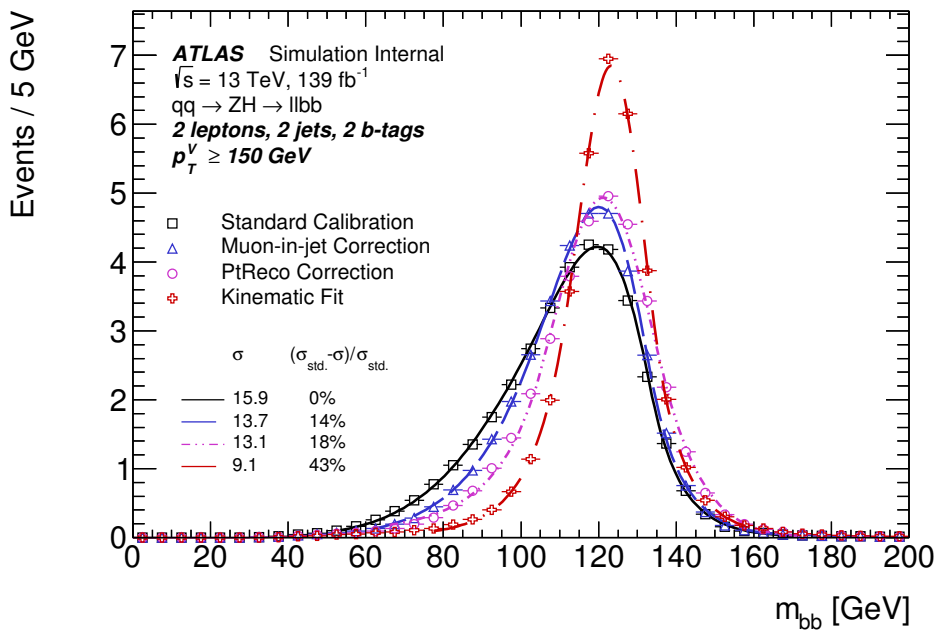


Figure 5.4: Comparison of the m_{bb} resolution with the muon-in-jet correction, PtReco correction and Kinematic Fit. All corrections are applied to the jet energy scale. The simulated $qqZH$ events in the 2-jet category, $p_T^Z > 150$ GeV region of the 2-lepton channel are shown [134].

5.2.3. Overlap Removal

The overlap removal procedure is applied before the reconstruction of missing transverse momentum to avoid any double-counting of objects created from the same tracks and calorimeter deposits within an event. During this procedure, only the pre-selected objects among electrons, muons, τ-leptons, small-R jets ($R = 0.2$) and large-R jets ($R = 1.0$ or larger) are considered. The individual overlap removals are done in the following sequence (Fig. 5.5).

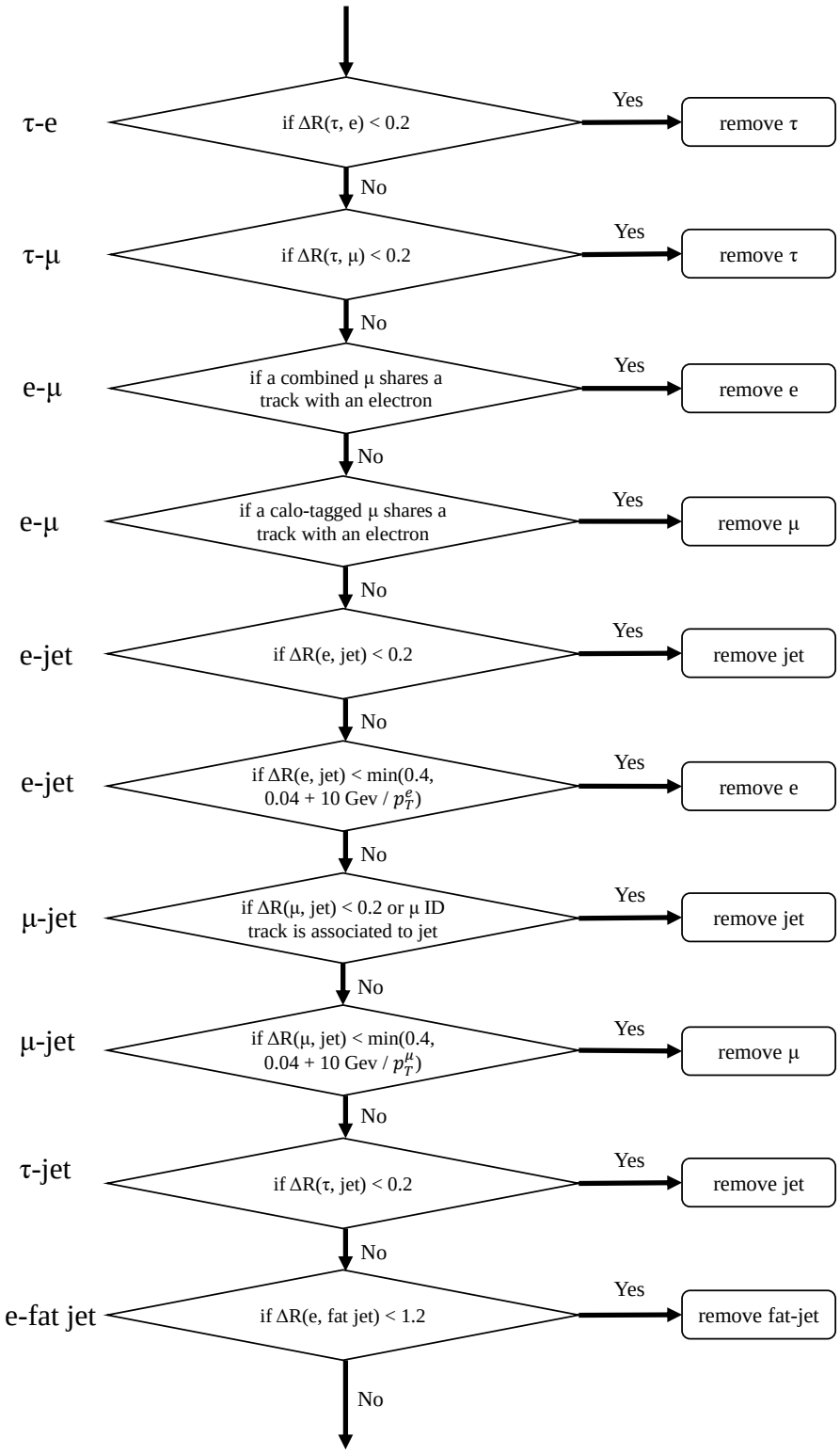


Figure 5.5: The cut-flow diagram of the overlap-removal procedure.

5.2.4. Missing Transverse Momentum

The missing transverse momentum, $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$, plays an important role in the 0-lepton and 1-lepton channels due to the existence of neutrino in $Z \to \nu\nu$ and $W \to \ell\nu$. The $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed as the negative vector sum of the transverse momentum of both the hard term objects and the track-based soft term, as the equation shown below [163]. The hard term contribution arises from muons, electrons, photons, hadronically decaying τ-leptons and jets. The soft term is calculated as the vector sum of the p_{T} of the tracks matched to the primary vertex but not associated to a reconstructed object. Since these tracks are associated with the primary vertex, the track-based soft term is robust against the pile-up [134]. Neutral particle signals from the calorimeter suffer from significant contributions from pile-up and are therefore not included in the soft term [163]. The magnitude of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is referred to as $E_{\mathrm{T}}^{\mathrm{miss}}$ [27].

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = \underbrace{\underbrace{-\sum_{\substack{\text{selected}\\\text{electrons}}} \mathbf{p}_{\mathrm{T}}^{e}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},e}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{photons}}} \mathbf{p}_{\mathrm{T}}^{\gamma}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\gamma}} \underbrace{-\sum_{\substack{\text{accepted}\\\tau\text{-leptons}}} \mathbf{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\tau_{\mathrm{had}}}} \underbrace{-\sum_{\substack{\text{selected}\\\text{muons}}} \mathbf{p}_{\mathrm{T}}^{\mu}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\mu}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{jets}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{jet}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss,jet}}}}_{\text{hard term}} \underbrace{\underbrace{-\sum_{\substack{\text{unused}\\\text{tracks}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{track}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss,soft}}}}_{\text{soft term}} .$$

[27] Another missing transverse momentum, denoted as $\mathbf{E}_{\mathrm{T,\,trk}}^{\mathrm{miss}}$, $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ or softMET, is calculated using only tracks in the inner tracking detector and from the primary vertex [159]. $\mathbf{E}_{\mathrm{T,\,trk}}^{\mathrm{miss}}$ provides a robust estimate of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ while being less sensitive to the pile-up. However, since $\mathbf{E}_{\mathrm{T,\,trk}}^{\mathrm{miss}}$ is only based on tracks left by charged particles, it cannot account for neutral particles at all [134].

CHAPTER 6

Event Selection

This chapter describes the event selection for the 0-, 1- and 2-lepton channel. Also, definitions of signal region and control region, and regions for cross-section measurement are presented.

6.1. Event Selection

The signal events are categorized into 3 channels depending on the number of leptons from the processes of $ZH \to \nu\nu b\bar{b}$, $WH \to \ell\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$. In all three channels, events are required to have exactly two b-tagged jets and the leading b-tagged-jet p_{T} is required to be greater than 45 GeV, given of the signature of $H \to b\bar{b}$. In each channel, events are further split into 2-jet and 3-jet (3-or-more-jet in the 2-lepton channel) categories. In the 0-lepton and 1-lepton channels, only one non-b-jet is allowed in the 3-jet category. This is to remove the $t\bar{t}$ background with four or more jets. In the 2-lepton channel, the 3-or-more-jet category has all the events with the number of non-b-jet greater than or equal to one. This is to increase the signal acceptance by 100% in the 3-or-more-jet category [159].

The reconstructed transverse momentum of the vector boson, p_T^V, further divides the events in the 2-jet and 3-jet categories into more regions. In the 0-lepton channel, p_T^V corresponds to $E_{\mathrm{T}}^{\mathrm{miss}}$; in the 1-lepton channel, it corresponds to the vectorial sum of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the charged-lepton p_{T}; in the 2-lepton channel, it corresponds to the p_{T} of the 2-lepton system. The signal-to-background ratio increases for large p_T^V values [164, 165]. Thus, events in the high-p_T^V region of the 2-jet and 3-jet categories are selected and split into 2 regions – 150 GeV $< p_T^V <$ 250 GeV and $p_T^V >$ 250 GeV. In the 2-lepton channel, an

additional fiducial measurement region is studied via the inclusion of a medium-p_T^V region with 75 GeV $< p_T^V <$ 150 GeV [159].

Event selections specific to each channel are defined in the following subsections.

6.1.1. 0-lepton

In the 0-lepton channel, a specific selection is defined to select events containing a Z decaying to a pair of neutrinos, in addition to the $H \to b\bar{b}$ selection. For this purpose, $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 150 GeV and no "loose" lepton candidates exist in the event. This is motivated by the fact that the offline trigger efficiency for events with a reconstructed $E_{\mathrm{T}}^{\mathrm{miss}} = 150$ GeV is about 90%[1] [109]. Further requirements on the scalar sum of the p_{T} (H_{T}) of the jets in the events are applied to remove a region where the trigger efficiency is dependent on the number of jets in the event[2]. For the 2-jets events, H_{T} must be larger than 120 GeV and for the 3-jet events, H_{T} must be larger than 150 GeV [109].

Due to the fact that the fake high $E_{\mathrm{T}}^{\mathrm{miss}}$ in multi-jet events typically arises from mismeasured jets in the calorimeters and $E_{\mathrm{T}}^{\mathrm{miss}}$ tends to be aligned with the mismeasured jets, a selection on the azimuthal angular difference of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$, jets and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$is applied to remove multi-jet events. The selection imposes four requirements[3], also referred to as "anti-QCD cuts", in Ref. [109, 166]. The values of the cuts are tuned in such a way that the remaining fraction of multi-jet contamination is of the order of 1% of the total background. Therefore it is negligible in the 0-lepton channel [134].

[1] Trigger scale correction factor are applied to the MC events, ranging from 0.95 at offline $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV, to ~ 1 at $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV [159].

[2] The dependency is due to the mis-modeling in the simulation. It was found in the Run-1 analysis as well [109].

[3] (1) $|\Delta\Phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})| < 90°$; (2) $|\Delta\Phi(jet1, jet2)| < 140°$; (3) $|\Delta\Phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, h)| > 120°$; (4) $min[|\Delta\Phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \text{pre-sel. jets})|] > 20°$ for 2 jets, $> 30°$ for 3 jets. Details can be referred to Ref. [109].

6.1.2. 1-lepton

Selections in the 1-lepton channel are implemented to confine events containing a W decaying to a neutrino and an electron or a muon in addition to the $H \to b\bar{b}$ [109]. So an event is required to have exactly 1 "tight" electron with $p_{\mathrm{T}} > 27$ GeV or 1 "tight" muon with $p_{\mathrm{T}} > 25$ GeV. Additional "loose" leptons are vetoed. In the electron sub-channel, a selection of $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV is applied to reduce the multi-jet background. In the muon sub-channel, the same $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers and correction factor as the 0-lepton channel are used. Given that muons are not included in the trigger $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers effectively select on p_T^V and are more efficient than the single-muon triggers in the analysis regions [159].

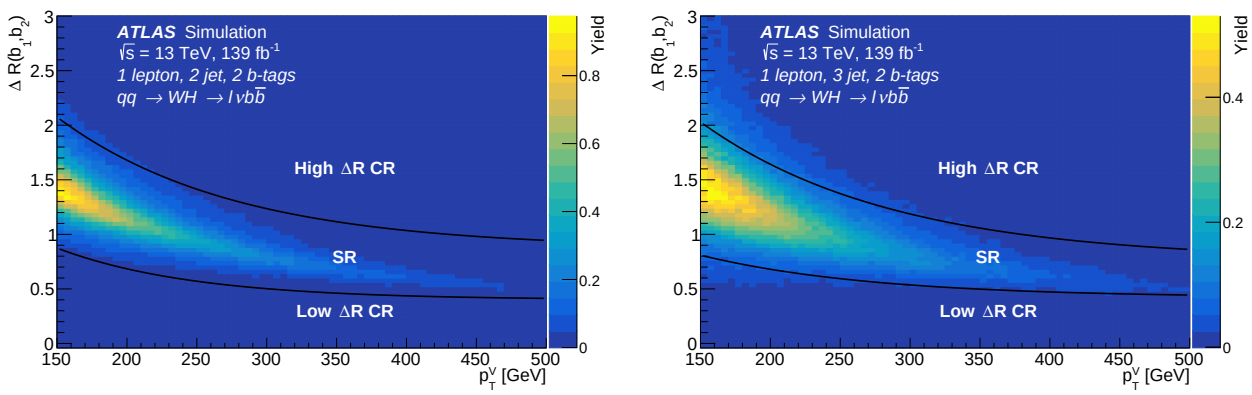
6.1.3. 2-lepton

In the 2-lepton channel, the decay of Z boson to two same flavor leptons (ee or $\mu\mu$) needs to be reconstructed with the $H \to b\bar{b}$ decay. Events are required to have exactly two same-flavor "loose" leptons, at least one of which has $p_{\mathrm{T}} > 27$ GeV and satisfies the electron selection of "ZH-signal" in Table 5.5 or the muon selection of "ZH-signal" in Table 5.6. For the di-muon events, the two muons are additionally required to have opposite charges[4]. The invariant mass of ee or $\mu\mu$ must be $81 < m_{\ell\ell} < 101$ GeV, to suppress the $t\bar{t}$ and multi-jet backgrounds. After all these requirements, the multi-jet background in the 2-lepton channel is negligible [109].

[4]This requirement is not applied to di-electron events due to higher rate of charge mis-identification [167].

6.2. Signal and Control Region

To better constrain the modeling of background processes, a number of signal and control regions have been defined after the event selection. Events are categorized into either signal regions or control regions, using a continuous selection on the ΔR between the two b-tagged jets ($\Delta R(\vec{b_1}, \vec{b_2})$) as a function of $p_{\mathrm{T}}^{\mathrm{V}}$. A lower and upper cut was introduced to outline the signal yield in the $\Delta R(\vec{b_1}, \vec{b_2})$-$p_{\mathrm{T}}^{\mathrm{V}}$ plane, which creates the high and low ΔR control regions [109], as shown in Fig 6.1. The signal events splitting into signal and control regions for each analysis category is summarized in Table 6.3.



(a) (b)

Figure 6.1: Signal yield distribution of ΔR between the two selected b-jets as a function of p_T^V in the 1-lepton channel are shown in the 2-tag 2-jet (a) and 2-tag 3-jet (b) categories. The black lines demonstrate the upper and lower continuous cuts used to categorize the events into the signal and control regions [159].

The upper cut is defined in the 1-lepton channel to keep 95% of the signal in the signal region for 2-jet events, and 85% of the signal for 3-jet events [159], as shown in Table 6.2. This region in the 1-lepton channel is enriched with $t\bar{t}$ and single-top backgrounds[5]. In the 0-lepton channel, the contribution from $t\bar{t}$ is enhanced. With the division into categories

[5]These control regions are first constructed in the 1-lepton channel and then applied to 0-lepton and 2-lepton channels to keep the definition of control region common in all 3 channels.

Table 6.1: Summary of the event selection and categorization in the 0-, 1- and 2-lepton channels. Table is from the publication of the $VHbb$ analysis of ATLAS [159].

Selection	0-lepton	1-lepton		2-lepton
		e sub-channel	μ sub-channel	
Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single lepton	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single lepton
Leptons	0 *loose* leptons	Exactly 1 *tight* electron 0 additional *loose* leptons $p_{\mathrm{T}} > 27$ GeV	Exactly 1 *tight* muon 0 additional *loose* leptons $p_{\mathrm{T}} > 25$ GeV	Exactly 2 *loose* leptons $p_{\mathrm{T}} > 27$ GeV Same-flavor Opposite-sign charges ($\mu\mu$)
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 150 GeV	> 30 GeV	–	–
$m_{\ell\ell}$	–	–	–	81 GeV $< m_{\ell\ell} < 101$ GeV
Jet p_{T}		> 20 GeV for $\lvert\eta\rvert < 2.5$ > 30 GeV for $2.5 < \lvert\eta\rvert < 4.5$		
b-jets		Exactly 2 b-tagged jets		
Leading b-tagged jet p_{T}		> 45 GeV		
Jet categories	Exactly 2 / Exactly 3 jets	Exactly 2 / Exactly 3 jets		Exactly 2 / ≥ 3 jets
H_{T}	> 120 GeV (2 jets), >150 GeV (3 jets)	–		–
$\min[\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{\mathrm{jets}})]$	$> 20^\circ$ (2 jets), $> 30^\circ$ (3 jets)	–		–
$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{bb})$	$> 120^\circ$	–		–
$\Delta\phi(\vec{b_1}, \vec{b_2})$	$< 140^\circ$	–		–
$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$	$< 90^\circ$	–		–
p_{T}^V regions	– 150 GeV $< p_{\mathrm{T}}^V < 250$ GeV $p_{\mathrm{T}}^V > 250$ GeV	– 150 GeV $< p_{\mathrm{T}}^V < 250$ GeV $p_{\mathrm{T}}^V > 250$ GeV		75 GeV $< p_{\mathrm{T}}^V < 150$ GeV 150 GeV $< p_{\mathrm{T}}^V < 250$ GeV $p_{\mathrm{T}}^V > 250$ GeV
Signal regions	$\Delta R(\vec{b_1}, \vec{b_2})$ signal selection			
Control regions	High and low $\Delta R(\vec{b_1}, \vec{b_2})$ side-bands			

along p_T^V, a better constraining of $t\bar{t}$ and single top modeling systematics is achieved [109]. In the 2-lepton channel, this region is enriched with the $Z + \text{HF}$ events. The lower cut is defined to keep 90% of the diboson yield in the signal region in the 1-lepton channel, shown in Table 6.2. This is to ensure sufficient diboson events remain when later conducting the diboson validation analysis [159]. The low ΔR region is enriched with $W + \text{HF}$ events in the 1-lepton channel and $Z + \text{HF}$ events in the 2-lepton channel. In the 0-lepton channel, $W + \text{HF}$ events are enhanced in this region which allows to better constrain $V + \text{HF}$ systematics with the categorization along p_T^V [109]. The plots illustrating the signal and background events are in the pre-fit distributions at Appendix D.

Category	Cut
High ΔR $2-jet$	$\Delta R > 0.87 + e^{1.38-0.00795\times p_T^V}$
High ΔR $3-jet$	$\Delta R > 0.76 + e^{1.33-0.0073\times p_T^V}$
Low ΔR $2-jet$	$\Delta R < 0.40 + e^{0.788-0.01023\times p_T^V}$
Low ΔR $3-jet$	$\Delta R < 0.42 + e^{0.268-0.00809\times p_T^V}$

Table 6.2: Cuts defining the high and low ΔR control region [109].

In addition, an e-μ control region is defined in the 2-lepton to enrich the top events. Since $t\bar{t}$ is flavor symmetric, a high purity $t\bar{t}$ control region can be obtained by requiring the flavor of the dilepton pair to be different (eμ or μe) [109]. The top control region is considered after the event selection; the only difference between the analysis phase space and the top control region is the lepton flavor (kinematics of $t\bar{t}$ between two phase space don't change). The top control region is used to extract a data-driven top background modeling template in the 2-lepton channel in the profile likelihood fit. Plots illustrating the distribution of $t\bar{t}$ in the e-μ control region in the 2-lepton channel are at Appendix D.

VH Signal						
0 lepton						
	2 jet			3 jet		
	Low ΔR CR	SR	High ΔR CR	Low ΔR CR	SR	High ΔR CR
150 GeV$< p_T^Z <$ 250 GeV	1.9%	92.8%	5.0%	2.7%	84.4%	12.9 %
$p_T^Z >$ 250 GeV	1.1%	94.3%	4.6%	3.2%	81.2%	15.5%
1 lepton						
	2 jet			3 jet		
	Low ΔR CR	SR	High ΔR CR	Low ΔR CR	SR	High ΔR CR
150 GeV$< p_T^W <$ 250 GeV	1.4%	93.6%	5.1%	2.6%	82.6%	14.8%
$p_T^W >$ 250 GeV	0.7%	94.6%	4.7%	3.2%	82.3%	14.5%
2 lepton						
	2 jet			$\geq$ 3 jet		
	Low ΔR CR	SR	High ΔR CR	Low ΔR CR	SR	High ΔR CR
75 GeV$< p_T^Z <$ 150 GeV	1.1%	92.6%	6.3%	4.1%	79.5%	16.4%
150 GeV$< p_T^Z <$ 250 GeV	0.8%	93.1%	6.1%	2.8%	73.7%	23.5%
$p_T^Z >$ 250 GeV	0.3%	94.3%	5.4%	2.3%	67.5%	30.2%

Table 6.3: The split of VH signal events into signal and control regions for each channel and each p_{T}^V and number of jets category. Numbers are given as the percentage of total signal events in that given analysis category [109].

6.3. Regions for Cross-section Measurement

Simplified Template Cross Sections (STXS) scheme [168,169] is adopted to define common regions for the Higgs measurements in LHC experiments. It provides fine-grained regions for different Higgs production modes, and allows for the combination in different decay channels and eventually between experiments.

The STXS scheme used in the measurement of VH with $V \to$ leptons and $H \to b\bar{b}$ is summarized in Table 6.4. In this scheme, $qq \to ZH$ and $gg \to ZH$ are treated as a single ZH production since the sensitivity is not enough to distinguish them. All leptonic decays of gauge bosons (including $Z \to \tau\tau$ and $W \to \tau\nu$, which are extrapolated from the electron and muon channel measurements) are used for the STXS definition. Regions in STXS scheme are combined and chosen such that the total uncertainty of the measurement in each region is near or below 100%. The signal acceptance (with the efficiency of the experimental selection included) and distribution (of the discriminating variables) in each STXS region is estimated from the simulated VH samples based on the p_{T}^{V} in the truth level (denoted as $p_T^{V,t}$). The migration matrices on the signal yield between the truth-p_{T}^{V}-region and reconstructed-p_{T}^{V}-region are shown in Fig. 6.2, where the top one describes the migration on the signal yield, and the bottom one describes the migration on the fraction of signal events corresponding to the reconstructed-p_{T}^{V}-region. Since the acceptance in the region of $p_T^{W,t} < 150$ GeV and $p_T^{Z,t} < 75$ GeV is very low (about 0.1% level, refer to Ref. [169]), cross-sections in these regions are constrained to the SM prediction, within their theoretical uncertainties [159]. Since these regions only contribute very little to the selected events, and thus their effects can be neglected.

The experimental value of the production cross-section (with $V \to$ leptons and $H \to b\bar{b}$ branching ratio) in each of the STXS regions is determined by a binned maximum-likelihood fit to the data, details of which are explained in Sec. 8.3.

STXS region		Corresponding reconstructed analysis regions		
Process	$p_{\mathrm{T}}^{V,\,\mathrm{t}}$ interval	Number of leptons	p_{T}^{V} interval	Number of jets
WH	150–250 GeV	1	150–250 GeV	2, 3
WH	$>$ 250 GeV	1	$>$ 250 GeV	2, 3
ZH	75–150 GeV	2	75–150 GeV	2, $\geq$ 3
ZH	150–250 GeV	0	150–250 GeV	2, 3
		2	150–250 GeV	2, $\geq$3
ZH	$>$ 250 GeV	0	$>$ 250 GeV	2, 3
		2	$>$ 250 GeV	2, $\geq$3

Table 6.4: A table summarizing the STXS regions and the corresponding reconstructed-p_{T}^{V}-regions. The current analysis is not sensitive to the regions WH, $p_{\mathrm{T}}^{W,\,\mathrm{t}} < 150$ GeV and ZH, $p_{\mathrm{T}}^{Z,\,\mathrm{t}} < 75$ GeV, and their cross-sections are fixed to the SM prediction within their theoretical uncertainties [159].

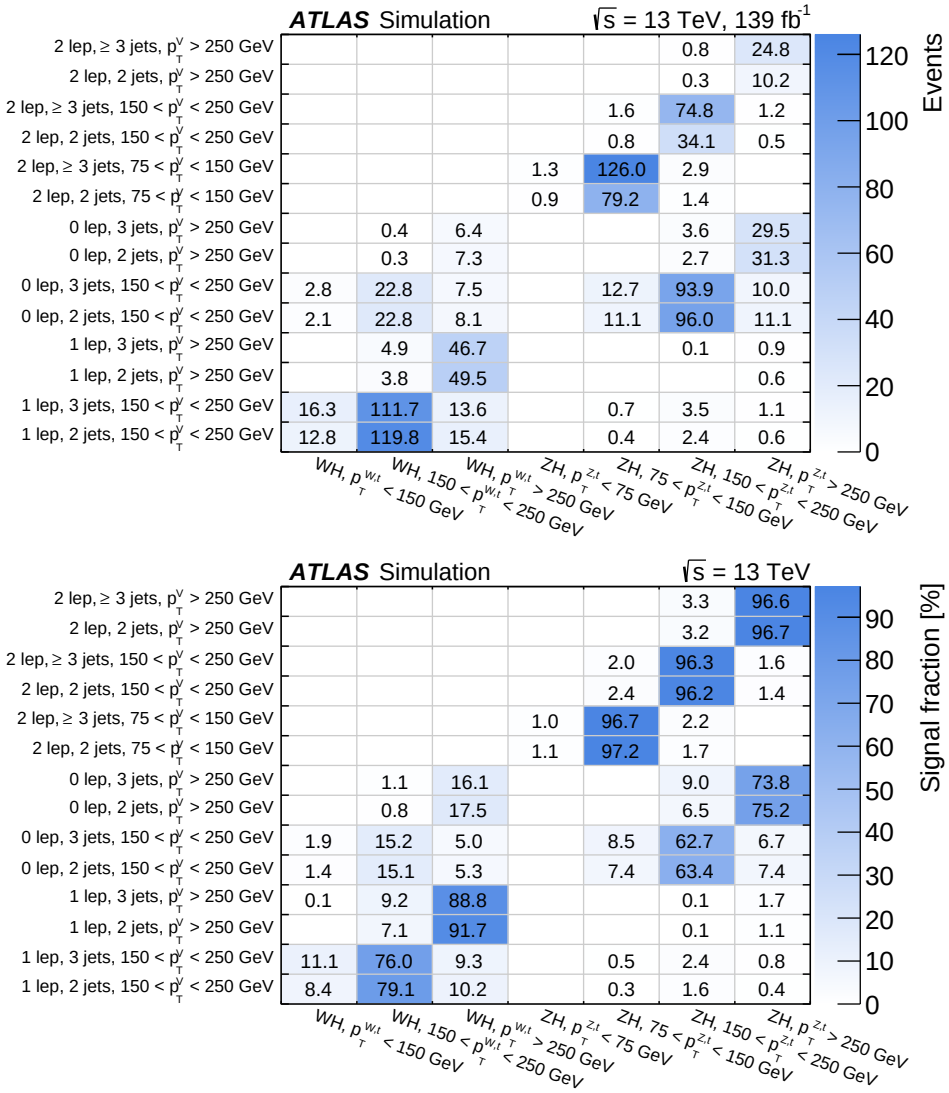


Figure 6.2: (Top) migration matrix of the expected signal yield between the truth-p_{T}^{V}-regions (x-axis) and reconstructed-p_{T}^{V}-regions (y-axis); (bottom) migration matrix of the signal fraction (with respect to the reconstructed-p_{T}^{V}-region) between the truth-p_{T}^{V}-regions (x-axis) and reconstructed-p_{T}^{V}-regions (y-axis). Regions with the signal yield below 0.1 or fraction below 0.1% are ignored [159].

CHAPTER 7

$t\bar{t}$ Rejection and Tau-veto in the 1-lepton Channel

This chapter describes one of my main contributions to the VHbb analysis – the study of the $t\bar{t}$ rejection and tau-veto in the 1-lepton channel.

7.1. $t\bar{t}$ Rejection in the 1-lepton Channel

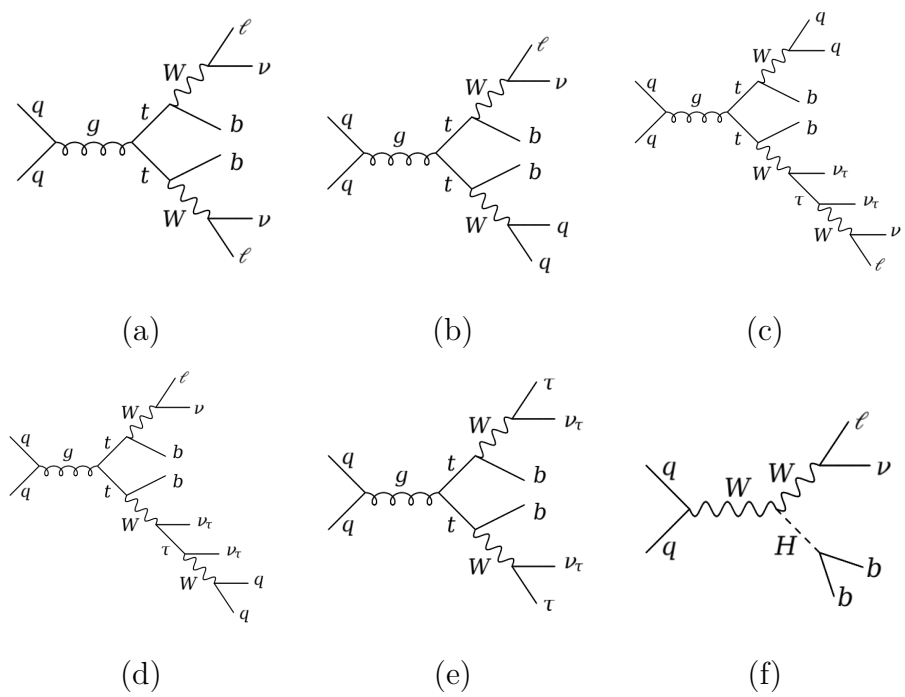


Figure 7.1: The diagrams are $t\bar{t}$ decaying (a) dileptonically, (b) semi-leptonically, (c) to a leptonic τ, (d) to a hadronic τ and (e) to $\tau\tau$. Diagram (f) is WH signal.

Top-pair production is one of the main backgrounds in the 1-lepton channel and has many decay processes, as shown in Fig. 7.1. It is very useful to understand what $t\bar{t}$ decay processes have been selected in the 1-lepton channel, as this will lead us to a better $t\bar{t}$ rejection.

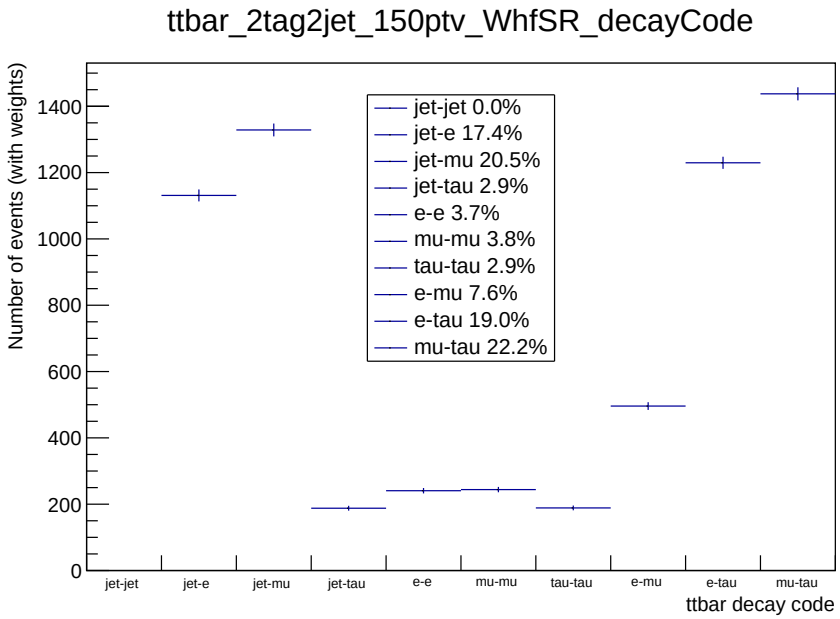


Figure 7.2: After the selection of 2-tag 2-jet category in the $p_{\mathrm{T}}^{V} > 150$ GeV region of the 1-lepton channel, each $t\bar{t}$ event is categorized according to decay mode (in truth-level) and then filled into this histogram.

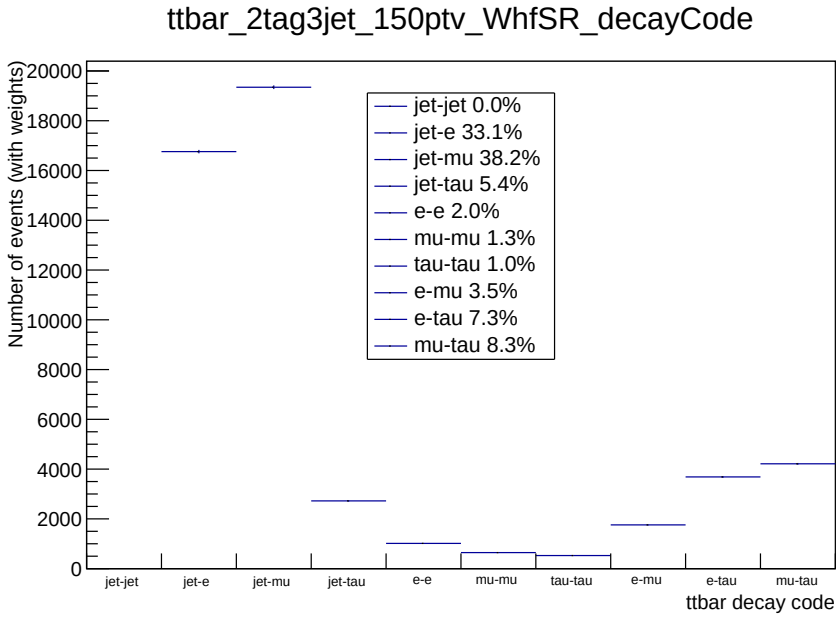


Figure 7.3: The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 3-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^{V} > 150$ GeV) after the distributions are normalized to unity.

The exploration of identifying the $t\bar{t}$ decay processes has been done through a variable named `codeTTbarDecay` which divides the $t\bar{t}$ events into ten categories based on their decays in the truth-level: "jet-jet," "jet-e," "jet-mu," "jet-tau," "e-e," "mu-mu," "tau-tau," "e-mu," "e-tau," "mu-tau." The distribution of the $t\bar{t}$ events with different decay modes after the selection of the 1-lepton channel is shown in Fig. 7.2 for 2-jet category and Fig. 7.3 for 3-jet category. When studying using `codeTTBarDecay`, one issue is that the children of tau lepton in truth-level are not saved in the MC samples. For example, when we get an integer value of 8 from `codeTTBarDecay`, although we can tell that this $t\bar{t}$ event has a decay mode of "e-tau", we don't know whether this tau decays leptonically or hadronically (because the children of tau lepton are not saved). Given that the selection of the 1-lepton channel requires exactly one reconstructed electron or muon, it is reasonable that we assume the tau decays hadronically in this "e-tau" event (otherwise, the number of reconstructed electron or muon will be two), which means that it is considered a hadronic tau. Thus, e-tau and mu-tau events can be categorized as "hadronic tau" events. Although this tau may be a leptonic tau, the possibility is very small[1]. Due to the same reason (the selection of exactly one reconstructed electron or muon in the 1-lepton channel), the tau in the "jet-tau" event is considered a leptonic tau.

Thus, all selected $t\bar{t}$ events can be grouped into "dileptonic," "semileptonic," "hadronic tau," "leptonic tau" and "tau-tau," categories, as summarized in Table 7.1 and Table 7.2. We can learn that the semileptonic decaying $t\bar{t}$ has the largest percentage among all $t\bar{t}$ decay processes after the selection. It is not a surprise, given that the $t\bar{t}$ semileptonic decay has a very similar final state with WH signal process. However, it is also very difficult to reject

[1]Because, for the $t\bar{t}$ e-tau events that pass the 1-lepton channel selection, the number of events where tau decays hadronically is much larger compared to the ones where tau decays leptonically. Therefore, the comparison and conclusion we obtain from this study are acceptable. Of course, there are still very few leptonic taus in the e-tau events categorized as hadronic tau. A more robust way of dealing with this is to use the truth-matching rather than `codeTTBarDecay`. Then we know exactly whether this is a leptonic-tau event or hadronic-tau event. The truth-matching was not yet ready in the analysis framework by the time the $t\bar{t}$ rejection was studied.

the semileptonic $t\bar{t}$ events, especially under the situation that no more than 3 jets are saved in the MC samples generated for the 1-lepton channel in this round of analysis.

The comparison of MVA input variable distributions of different $t\bar{t}$ decay modes are presented in Fig. 7.4, 7.5 and 7.6 for the 2-jet category, and Fig. 7.7, 7.8 and 7.9 for the 3-jet category. From the comparison, we can find that,

- for the variables that have nothing to do with $E_{\mathrm{T}}^{\mathrm{miss}}$, dileptonic (in black) and hadronictau (in green) distributions are similar and semileptonic (in red) and leptonictau (in blue) distributions are similar;
- for the variables related to $E_{\mathrm{T}}^{\mathrm{miss}}$, semileptonic (in red) and leptonictau (in blue) distributions are different.

The reasons can be explained as

- both the leptons in the dileptonic events and the jets in hadronic tau events are very soft (why distributions in black and green are similar), and the decay products of leptonic tau decay are very similar to that of the semileptonic decay, if not considering the number of neutrinos (why distributions in red and blue are similar);
- the decay products of leptonic tau decay and semileptonic decay are not similar anymore if considering the number of neutrinos (why distributions in red and blue are different).

Thus, for different $t\bar{t}$ decay processes, different strategies can be proposed to reject them. These rejection strategies are summarized in Table 7.3.

Process	dileptonic	semileptonic	hadronictau	leptonictau	tau-tau
Percentage in $t\bar{t}$	15.2%	38.0%	41.0%	2.9%	2.9%

Table 7.1: The percentage of $t\bar{t}$ processes in the 2-tag 2-jet region of 1-lepton (with $p_{\mathrm{T}}^{V} >$ 150 GeV).

Process	dileptonic	semileptonic	hadronictau	leptonictau	tau-tau
Percentage in $t\bar{t}$	6.8%	71.3%	15.6%	5.4%	1.0%

Table 7.2: The percentage of $t\bar{t}$ processes in the 2-tag 3-jet region of 1-lepton (with $p_{\mathrm{T}}^{V} >$ 150 GeV).

Hadronic tau	(i) Tau-veto (`nTaus` < 1). (ii) Reconstructed mTW(τ) using p_T^τ and $E_{\mathrm{T}}^{\mathrm{miss}}$.
Dileptonic	Its MVA variable distributions are very similar to the hadronic tau ones. Hopefully, it can benefit from the rejection of hadronic tau.
Leptonic tau	Its MVA variable distributions are different from the semileptonic ones (and WH signal). Maybe the $E_{\mathrm{T}}^{\mathrm{miss}}$ related distributions (like $E_{\mathrm{T}}^{\mathrm{miss}}$, mTW, pTL, dPhi(lep, MET) can play a role in the rejection.
Semileptonic	If four or more jets in an MC event are stored, a four-jet control region can be created for the semileptonic $t\bar{t}$ rejection.
Other ideas	(i) Lepton charge, soft MET, MV2c10 scores for b-jets and other jets, m_{bl}, etc. (ii) Comparison of the distributions in different flavor composition of 2 b-tagged jets.

Table 7.3: Strategies of $t\bar{t}$ events rejection in the 1-lepton channel.

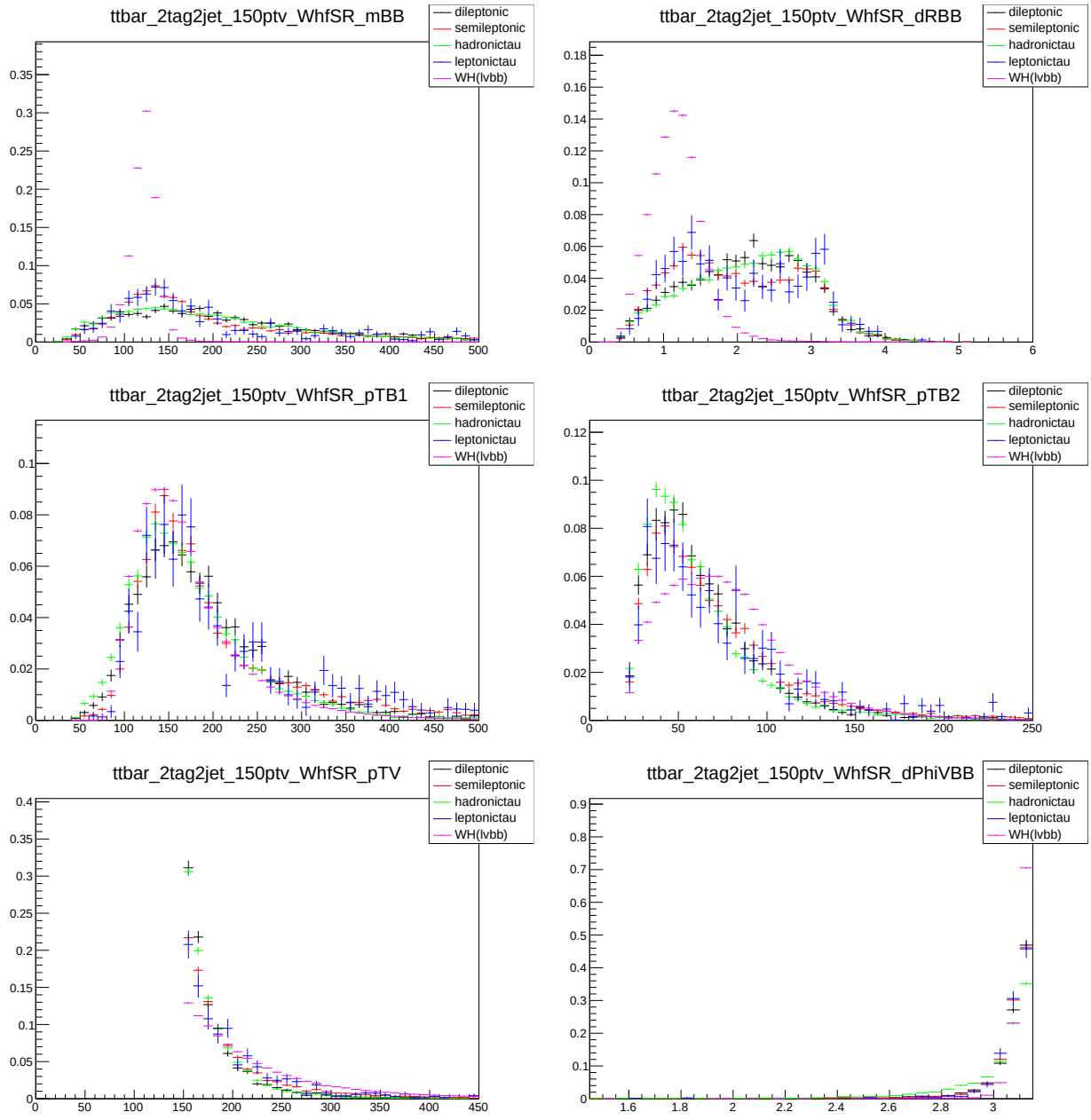


Figure 7.4: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel.

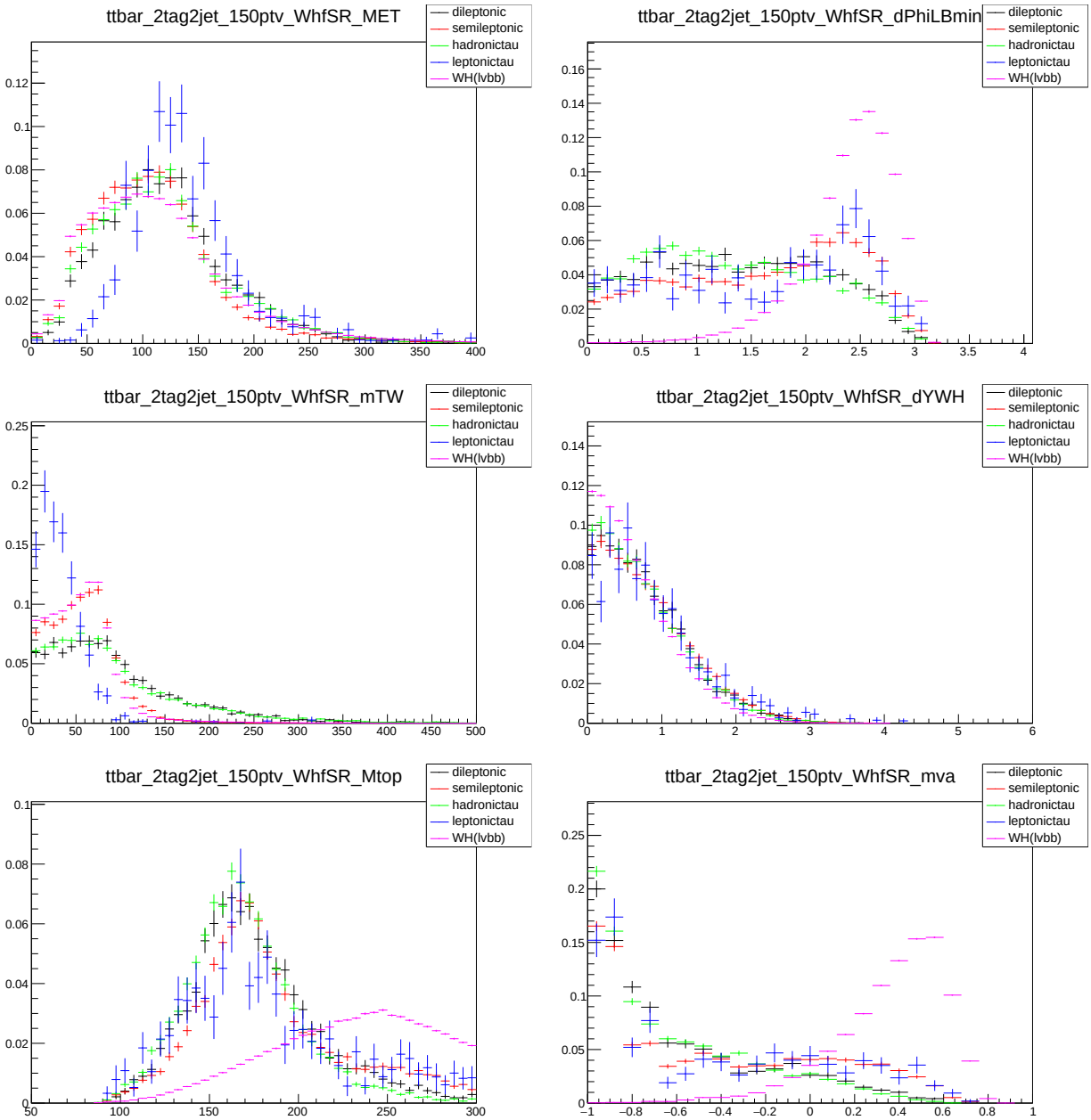


Figure 7.5: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel.

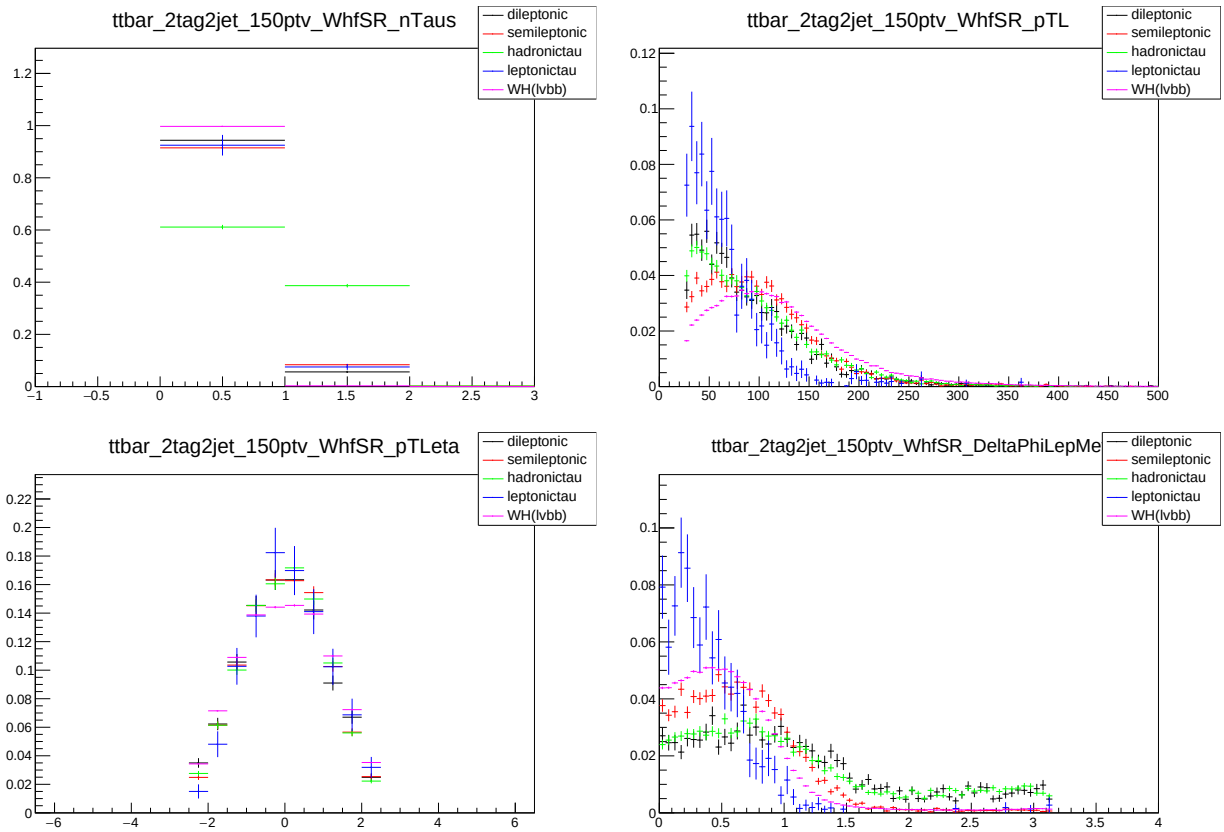


Figure 7.6: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 2-jet region of the 1-lepton channel.

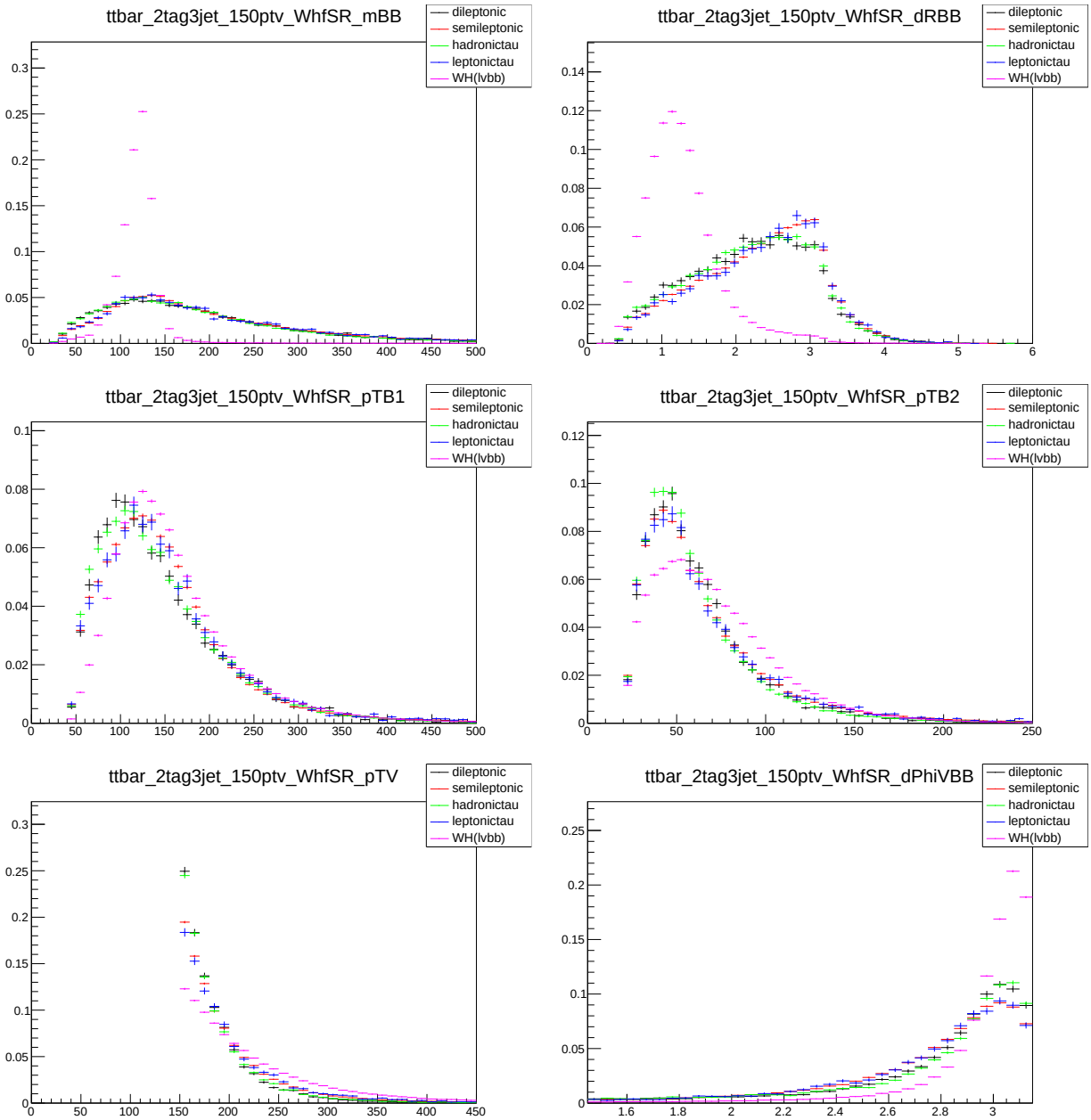


Figure 7.7: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel.

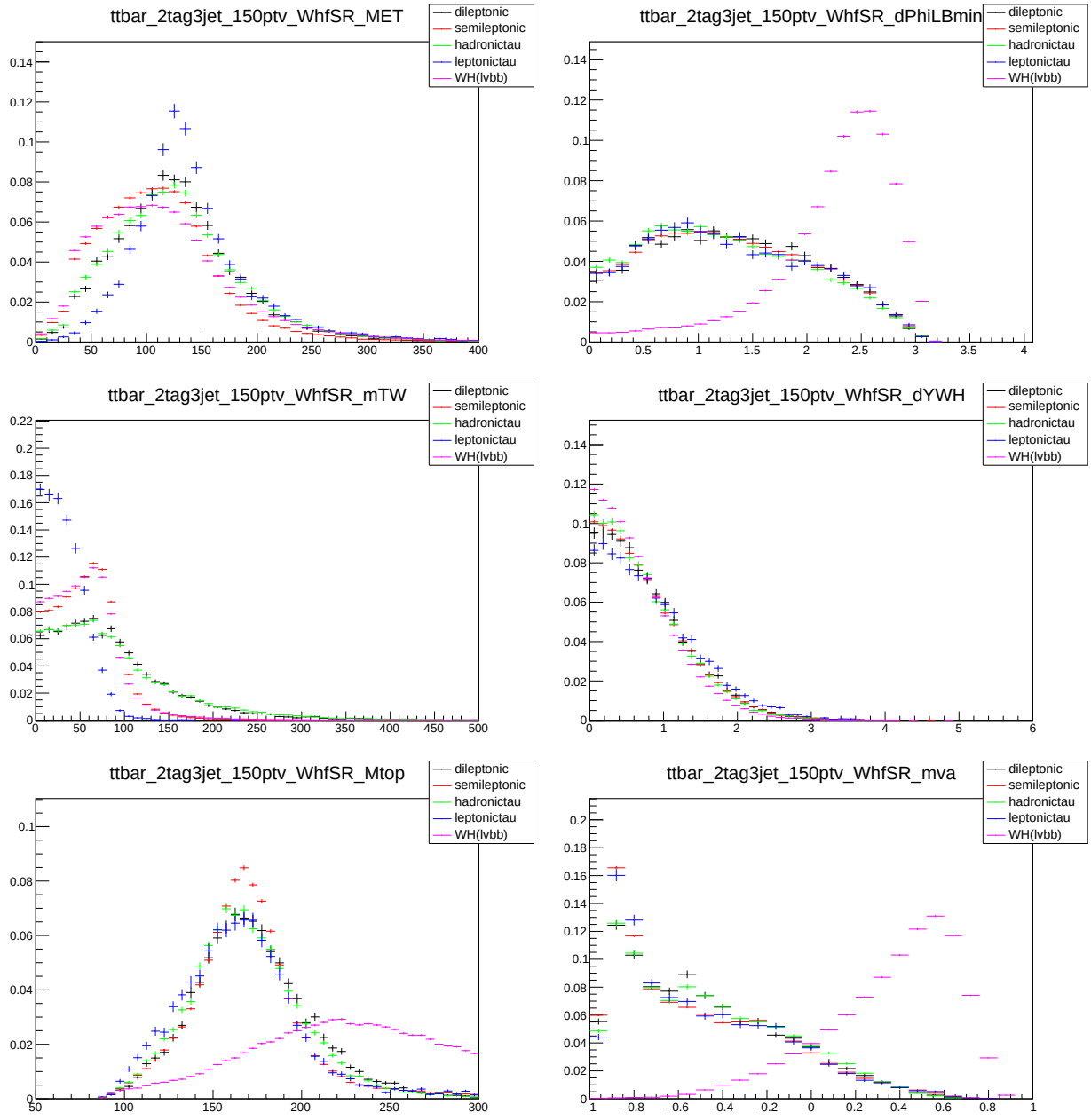


Figure 7.8: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel.

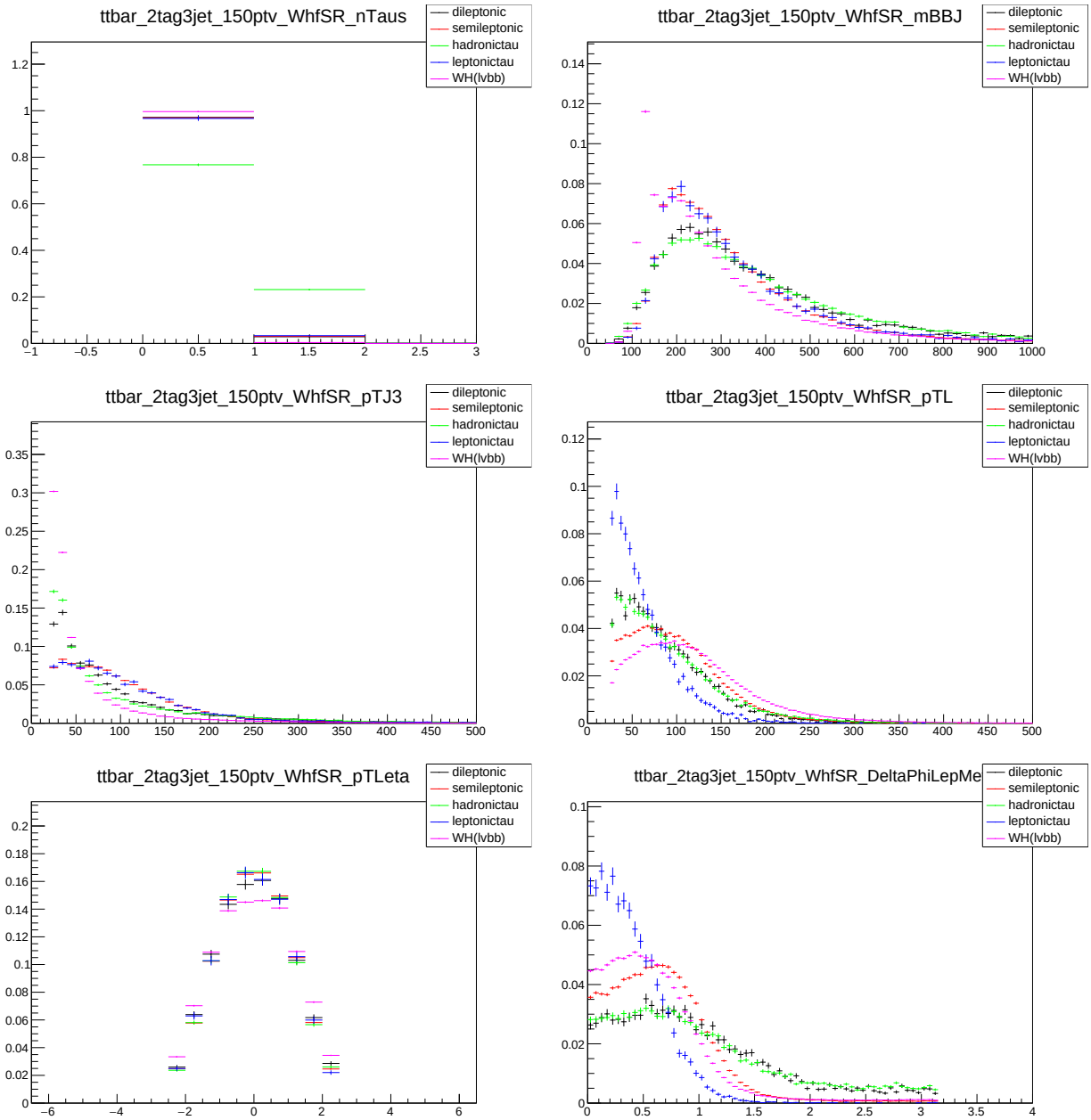


Figure 7.9: Comparison of possible MVA input variable distributions of different $t\bar{t}$ decay modes (and $WH(l\nu bb)$ in the 3-jet region of the 1-lepton channel.

7.2. Tau-veto in the 1-lepton Channel

The $t\bar{t}$ events decaying to a hadronic tau (labeled as "hadronic tau" $t\bar{t}$) has the second largest fraction in all $t\bar{t}$ events after the selection, as shown in Table 7.1 and 7.2. From the hadronic tau reconstruction algorithm (see Sec. 5.2.1.3), a variable `nTaus` denoting the number of reconstructed hadronic tau in an event is available. The `nTaus` distribution on different $t\bar{t}$ decay processes are shown in Fig. 7.10 (for the 2-jet category) and Fig. 7.11 (for the 3-jet category) for comparison. We can clearly see that the `nTaus` distribution of the hadronic tau $t\bar{t}$ events has a large proportion of `nTaus` > 1 events, compared to other $t\bar{t}$ processes. Thus, the hadronic tau events can be simply rejected by applying "tau-veto" (`nTaus` < 1)!

Table 7.4 shows the cut-flow[2] in the 2-jet category of the 1-lepton channel and Table 7.5 shows the cut-flow in the 3-jet category of the 1-lepton channel. After the application of tau-veto, 13.1% (3128 predicted events) of the total MC background events in the 2-jet category and 5.58% (7254 predicted events) of the total background events in the 3-jet category are rejected, which of them are $t\bar{t}$ hadronic tau events, as illustrated in Fig. 7.10 and Fig. 7.11. The VH signal only loses 0.73% (1.53 predicted events) in the 2-jet category and 0.86% (1.74 predicted events) in the 3-jet category after the tau-veto in the 1-lepton channel. The comparisons of BDT distributions with and without tau-veto are presented in Fig. 7.12 and Fig. 7.13 (the distributions are not normalized). They further validate that tau-veto removes not only the low BDT region events but also the high BDT region events for both the 2-jet and 3-jet categories.

The tau-veto study has proven its great power in rejecting the $t\bar{t}$ events in the 1-lepton channel. Although its systematic uncertainty estimation was not yet concluded by the time

[2]In the cut-flow table, "mTop<225 || mBB>75" is the old signal region definition before the introduction of the low and high ΔR signal and control regions. This study was done with the old signal region definition.

the $VHbb$ analysis went out for publication, it had gained a very high priority and is anticipated to be a standard part of the analysis shortly.

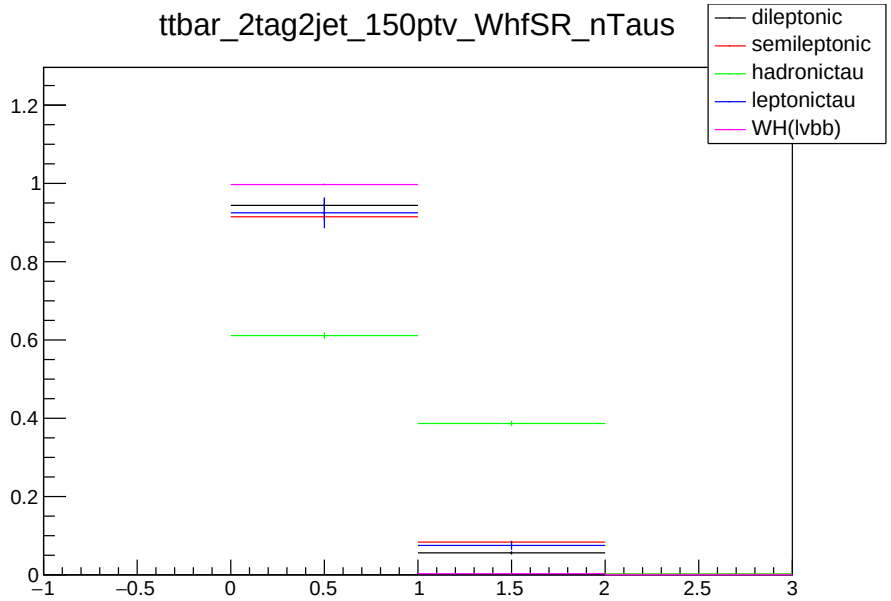


Figure 7.10: The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 2-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^{V} > 150$ GeV) after the distributions are normalized to unity.

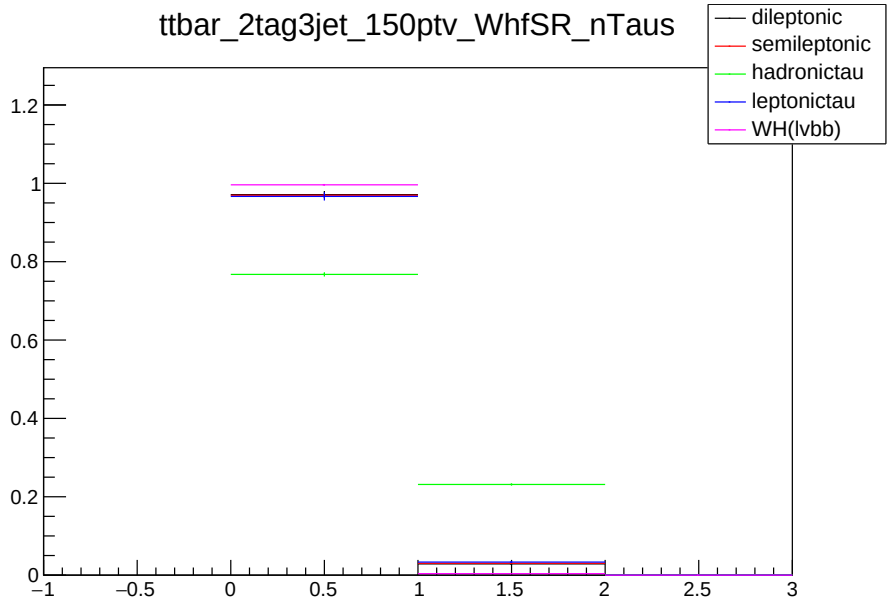


Figure 7.11: The comparison of "nTaus" variable distributions of different decay processes of $t\bar{t}$ in 2-tag 3-jet category of the 1-lepton channel (with $p_{\mathrm{T}}^{V} > 150$ GeV) after the distributions are normalized to unity.

Cut	Bkg2J (sumOfWeights)	Sig2J (sumOfWeights)
all	3.19×10^6	1419.63
nJ>1.5	3.19×10^6	1419.63
nJ<2.5	152901	504.464
pTV>150	25036.2	209.599
mTop<225 || mBB>75	23811.7	208.379
nTaus<1	20683.7	206.853

Table 7.4: The cut-flow of signal and backgrounds in 1-lepton 2-jet region.

Cut	Bkg3J (sumOfWeights)	Sig3J (sumOfWeights)
all	3.19×10^6	1419.63
nJ>2.5	3.04×10^6	915.167
nJ<3.5	515677	454.51
pTV>150	133245	205.993
mTop<225 || mBB>75	130112	202.157
nTaus<1	122858	200.417

Table 7.5: The cut-flow of signal and backgrounds in 1-lepton 3-jet region.

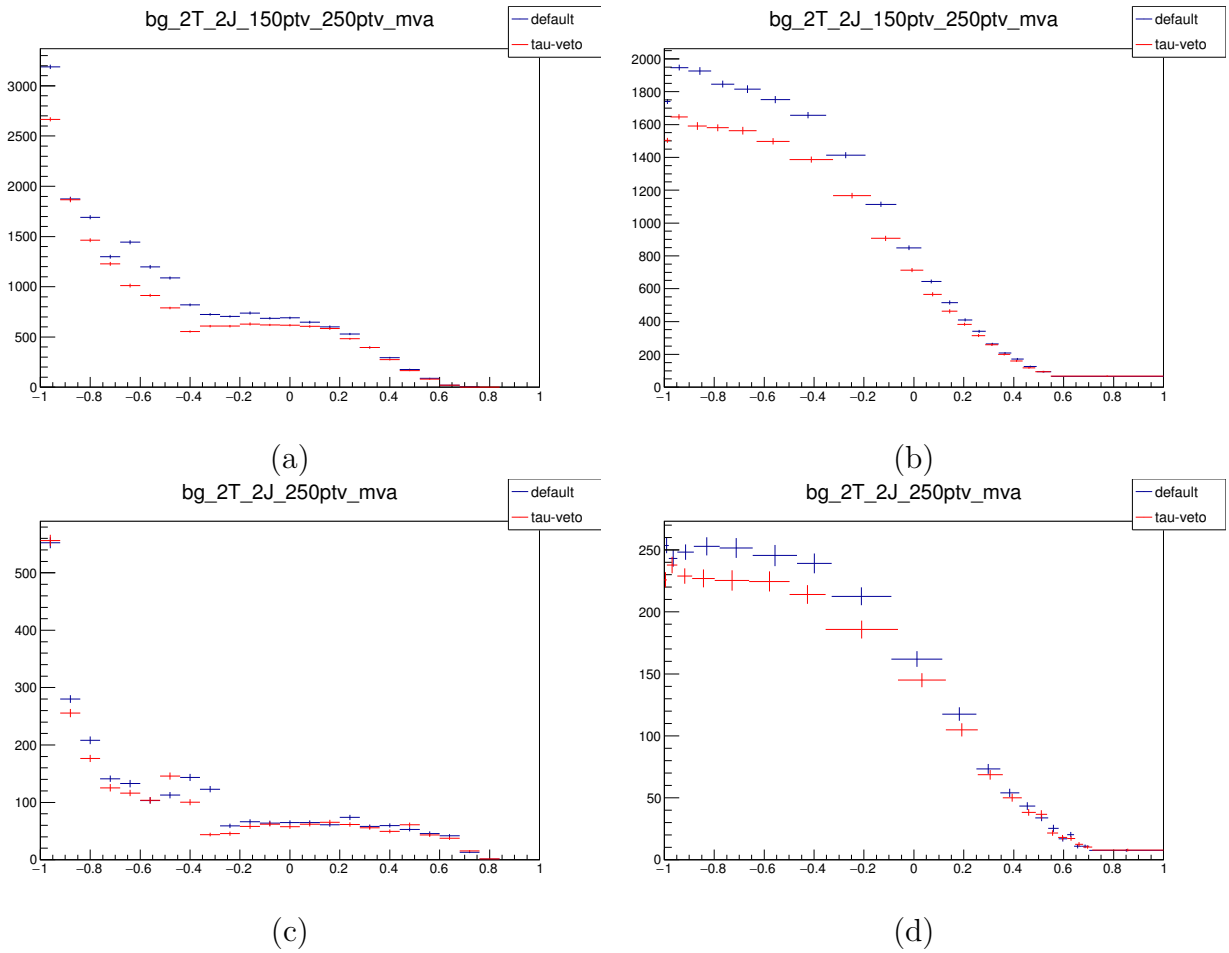


Figure 7.12: (a): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel; (b): the comparison of binned background MVA distributions with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel; (c): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} >$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel; (d): the comparison of binned background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} >$ 250 GeV region of 2-tag 2-jet category of the 1-lepton channel.

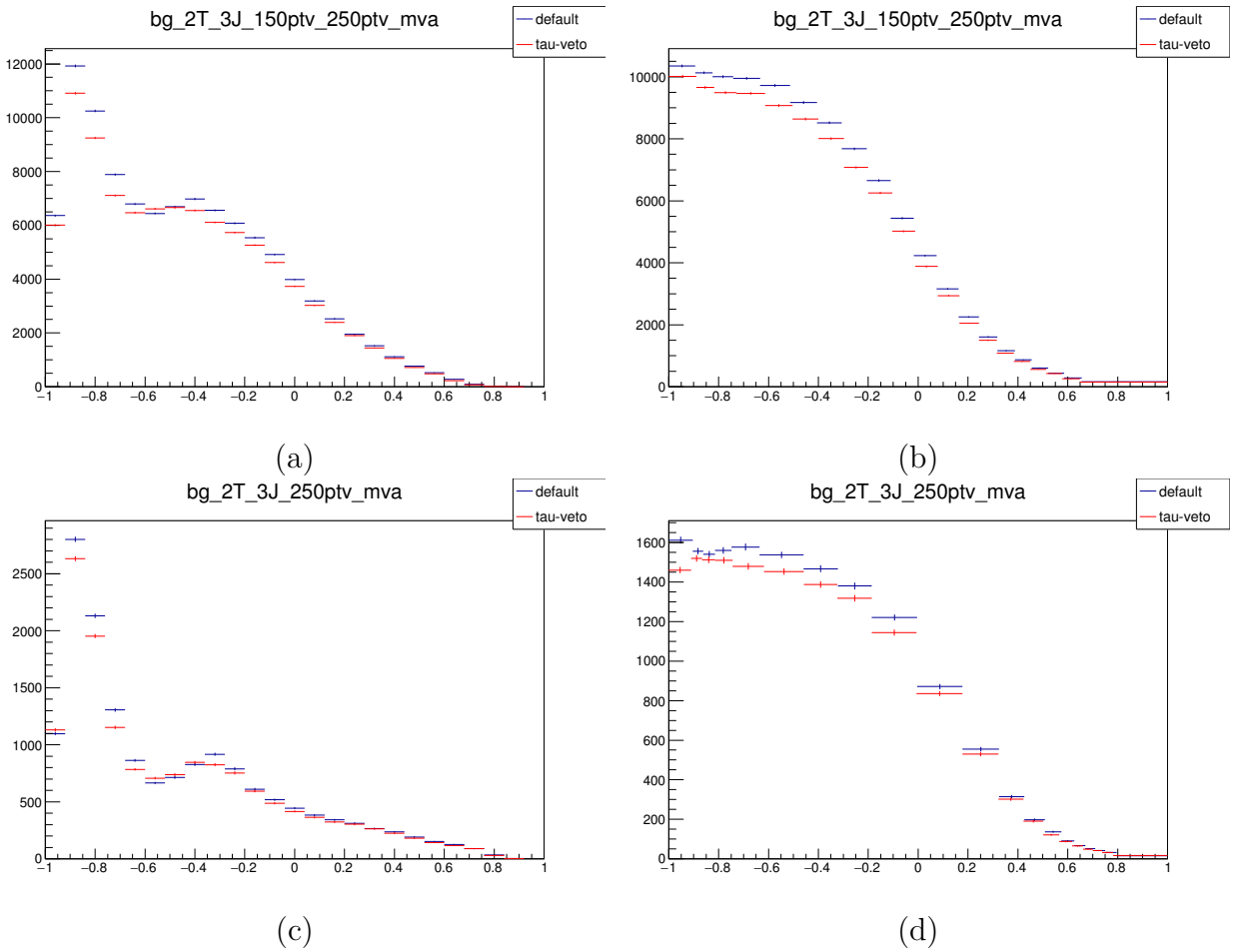


Figure 7.13: (a): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel; (b): the comparison of binned background MVA distributions with and without tau-veto in the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel; (c): the comparison of background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} >$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel; (d): the comparison of binned background MVA distributions (including $t\bar{t}$, W+jets, etc.) with and without tau-veto in the $p_{\mathrm{T}}^{V} >$ 250 GeV region of 2-tag 3-jet category of the 1-lepton channel.

CHAPTER 8

$VHbb$ Analysis

This chapter describes the three critical analysis methods used in the $VHbb$ analysis – multivariate analysis, systematic uncertainty estimation, and statistical analysis. My contribution is on the systematic uncertainty estimation of the $t\bar{t}$ background.

8.1. Multivariate Discriminants

Multivariate analysis (MVA) has been widely used in high energy physics to improve the sensitivity of the analysis. In this analysis, two sets of boosted decision trees (BDTs) are trained using the same input variables. One is the nominal set, denoted as BDT_{VH}, designed to discriminate the VH signal from the background processes; the other is the cross-check set, denoted as BDT_{VZ}, defined to separate the $VZ(Z \to b\bar{b})$ process from the VH signal and other background processes, and to validate the VH analysis [159]. In each set, BDTs are trained in eight regions categorized by the number of leptons, number of jets and range of the reconstructed p_{T}^{V}, as shown in Table 8.1. The 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV and $p_{\mathrm{T}}^{V} >$ 250 GeV regions in each lepton channel and jet category are merged together for the training. This is because the $p_{\mathrm{T}}^{V} >$ 250 GeV region has low training statistics which would easily lead to over-training. The BDT training is also performed in the inclusive phase space of signal and control regions (before the categorization of events based on $\Delta R(\vec{b_1}, \vec{b_2})$ and p_{T}^{V} described in Sec. 6.2 takes place). This is done to enhance the training statistics[1] [109]. The BDT outputs, evaluated in each signal region, are used as final discriminating variables [159].

[1]Studies showed that the training only using signal region events introduces significant levels of over-training without enhancing the performance [109].

0 lepton	2 jets (2 b-tagged)	$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV	BDT_1
	3 jets (2 b-tagged)	$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV	BDT_2
1 lepton	2 jets (2 b-tagged)	$p_T^W > 150$ GeV	BDT_3
	3 jets (2 b-tagged)	$p_T^W > 150$ GeV	BDT_4
2 lepton	2 jets (2 b-tagged)	75 GeV $< p_T^Z <$ 150 GeV	BDT_5
	$\geqslant$3 jets (2 b-tagged)	75 GeV $< p_T^Z <$ 150 GeV	BDT_6
2 lepton	2 jets (2 b-tagged)	$p_T^Z > 150$ GeV	BDT_7
	$\geqslant$3 jets (2 b-tagged)	$p_T^Z > 150$ GeV	BDT_8

Table 8.1: A table showing the regions that BDTs are trained on.

8.1.1. Boosted Decision Tree

In this analysis, the TMVA package of ROOT is included for the use of BDTs. A decision tree is a tool that uses a tree-like model of conditions and their possible consequences to help make decisions. It is one way to display an algorithm that only contains conditional control statements. A decision tree usually contains a sequence of selection cuts applied in a specified order on the discriminating variables of a data-set. The discriminating variables are selected by experimentalists based on their power of separating the signal from backgrounds. An example of decision tree is shown in Fig. 8.2. The very top of the tree is called *Root node*. Each cut splits the data-set into nodes, each of which corresponds to a given number of events, classified as signal or background. A node can be further split by the application of the subsequent cut in the tree. Nodes with no further selection applied on are called *leaves* or *leaf nodes*. Signal or background is usually largely dominant in leaves[2]. Each branch represents a sequence of cuts and the outcome of them[3] [170].

Selection cuts are usually tuned to optimize the split in each node. One way used in the optimization is to maximize the gain of the Gini index for the split of each node. The

[2]A node in which the signal or background is not largely dominant could also be classified as a leaf, and the selection path is then ended. This is to prevent that there are too few events remaining per node, or the total number of identified nodes is too large. There are different criteria in the real situation [170].

[3]Each node represents a statistical hypothesis test on an attribute.

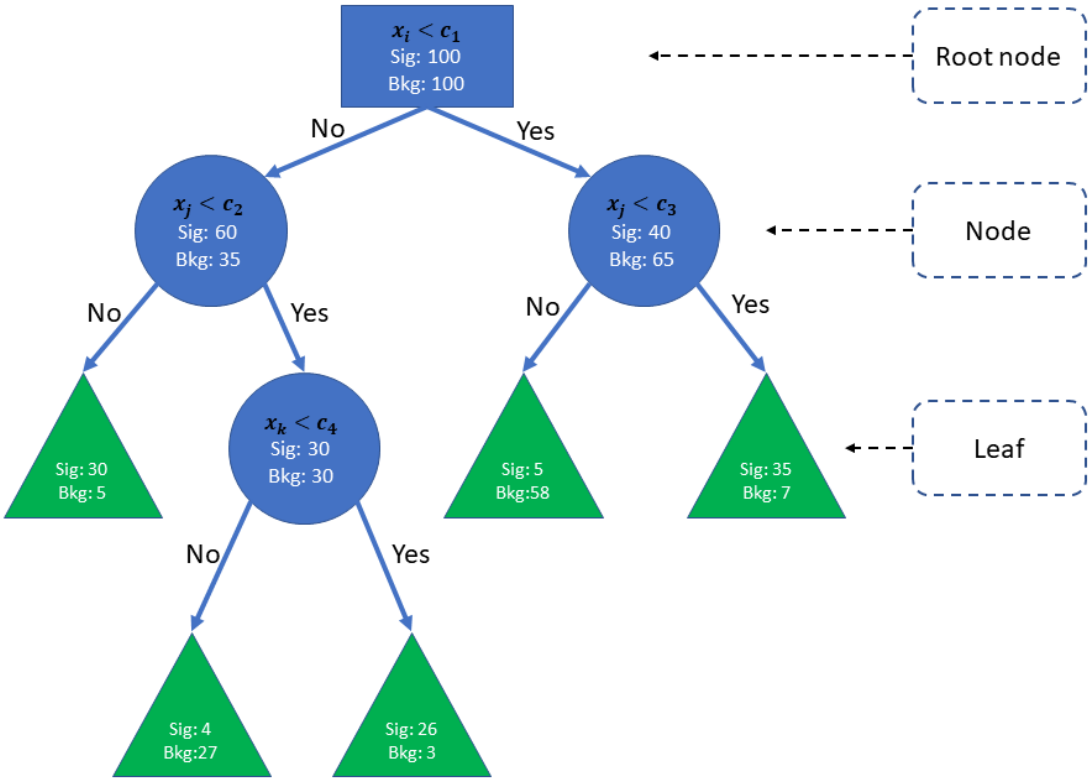


Figure 8.1: An example of a decision tree. The root node is represented as a Rectangle; each (internal) node is represented as a circle; and each leaf is represented as a triangle. Each of them contains a different number of signal and background events. Applied selection cuts are shown at the top of the root node and (internal) nodes.

Gini index is defined as $G = P(1 - P)$, where P is the purity of the node (the fraction of signal events). G is equal to zero for the nodes containing only signal or background events. For a chosen cut, the gain due to the splitting of a node A into nodes B_1 and B_2 is $\Delta G = G(A) - G(B_1) - G(B_2)$. By varying the cut threshold, the maximum gain can be achieved. The one selection cut which best separates the signal and background is usually used at the root of the tree [170].

Although a single decision tree can be optimized as mentioned above, one aspect has prevented it from being the ideal tool for prediction – inaccuracy[4] [172]. A single tree can work accurately with the data used to create it, but it is not flexible when it comes to classifying new samples. *Random forests* [173] is an advanced approach that combines the simplicity of decision trees with flexibility resulting in a vast improvement in the accuracy. The random forest algorithm grows many decision trees in parallel (usually hundreds) from replicas of the training samples obtained by randomly resampling the input data[5]. This results in a wide variety of trees. The variety is what makes the random forest more effective than individual decision trees. The final score of the algorithm is given by an unweighted average of the prediction (0 or 1, *no* or *yes*) by each individual tree. The process of bootstrapping the data plus using the aggregate to make a decision is called "bagging" [172].

The term "boosting" refers to a family of algorithms which converts the weak learners to strong ones. The idea of boosting is to train weak learners sequentially, each trying to correct its predecessor, to improve the model predictions. The weak learners in decision trees are shallow trees, sometimes as small as decision stumps (trees with just one node and two leaves). Figure 8.2 shows an example of how the *AdaBoost* algorithm [174] works. *AdaBoost* first builds a stump which is selected from attributes and has the maximum gain of the Gini index[6]. That stump is trained using the initial weights of the training samples[7]. *AdaBoost* then iteratively grows a forest of stumps and for each successive iteration, the event weights are individually modified. The modified weights of the training events are based on how

[4]The limit of the performance is set by the Neyman-Pearson lemma [170]. When we perform a hypothesis test between two hypotheses H_0: $\theta = \theta_0$ and H_1: $\theta = \theta_1$ using the likelihood-ratio $\Lambda(x)$ with threshold η, which rejects H_0 in favor of H_1 at a significance level of $\alpha = \mathrm{P}\ (\Lambda(\mathrm{x}) \leq \eta \mid H_0)$, where $\Lambda(\mathrm{x}) = \frac{\mathcal{L}(\theta_0 \mid x)}{\mathcal{L}(\theta_1 \mid x)}$ and $\mathcal{L}(\theta \mid x)$ is the likelihood function, the Neyman-Pearson lemma tells us that the likelihood ratio, $\Lambda(x)$, is the most powerful test at significance level α [171].

[5]The trees grown in the random forests are full-sized decision trees. There's no maximum depth required.

[6]The cut threshold of that stump is determined by the gain of the Gini index, as well.

[7]The modifications on samples at each boosting step apply weights w_1, w_2, ..., w_N to each of the training events. All weights are initially set to $w_i = 1/N$, so that in the first step, the classifier is trained on the samples in the usual way [172].

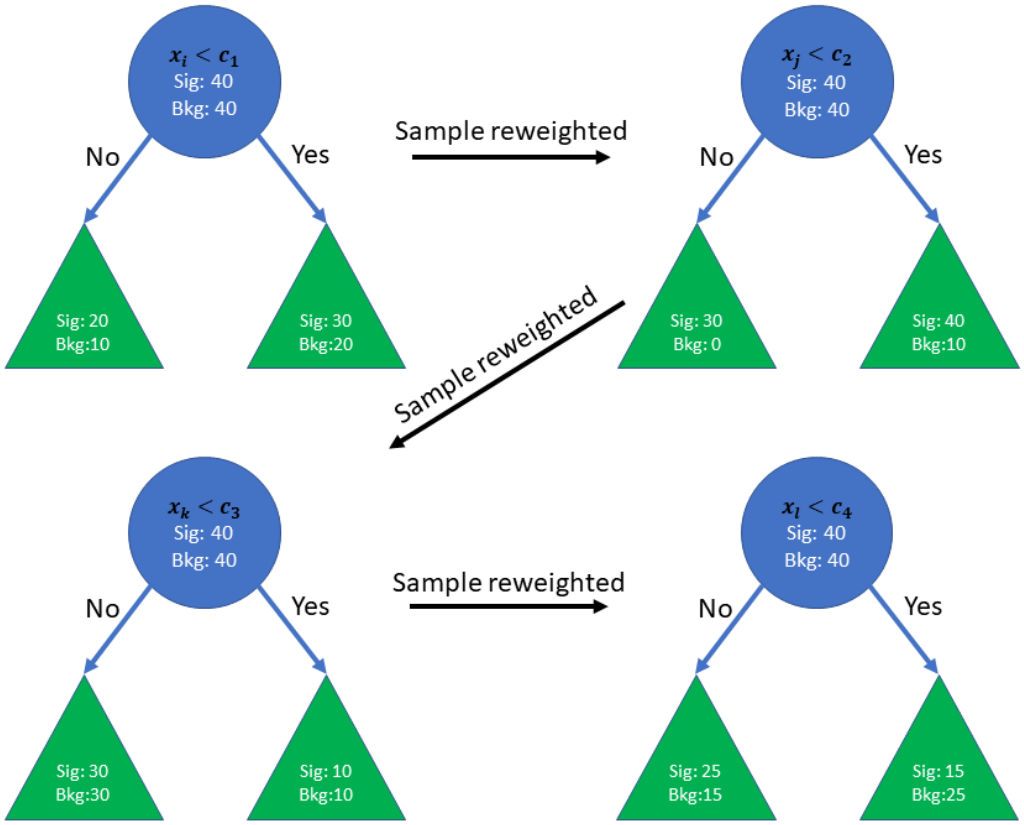


Figure 8.2: Illustration of *AdaBoost* algorithm. AdaBoost grows a forest of stumps. The training samples are reweighted using the previous classifier before a new stump is grown. The errors that the previous stump makes influence how the following stump is made.

well the previous stump classifies the samples, which is measured by the total error of that stump. The total error of that stump is calculated as the sum of the weights associated with the mis-classified events. Those events mis-classified by the previous stump will have their weights increased, whereas those that were classified correctly will have their weights decreased. In this way, as iterations proceed, events that are difficult to classify correctly receive escalating influence[8]. In that way, each successive classifier is forced to focus on those training events that are missed by the previous ones in the sequences [172]. In the end, the

[8]The sample weights can be used to calculate the weighted Gini index to determine which attribute/variable should be the next stump. The weighted Gini index would put more emphasis on correctly classifying the events that were mis-classified by the last stump since those events have larger weights.

output of the final classifier is the weighted average of all stumps in the forest given by each of the steps.

The *Gradient Boost* approach [175] has many similarities with *AdaBoost*. But instead of making a very short tree first, the *Gradient Boost* algorithm starts by making a single leaf. This leaf represents the initial prediction for every individual event and is usually the average value of the variable to be predicted. Then, based on the pseudo-residuals – the difference between the observed values and the predicted values, a new tree is built by *Gradient Boost*. This tree is usually larger than a stump but still has its size restricted. This tree's contribution is scaled to the final prediction with a learning rate. Then another tree is added based on the new pseudo-residuals, and *Gradient Boost* keeps adding trees based on those errors made by the previous trees until it has made the maximum number of trees specified, or additional trees doesn't significantly reduce the size of the residual. Thus, when there's new data available, the prediction can be made by starting the initial predicted value then adding scaled value from the trees in sequence. *Gradient Boost* is an aggressive strategy. At each step, the solution tree is the one whose gradient reduces most. Thus, the tree predictions are analogous to the components of the negative gradient [172].

In the $VHbb$ analysis, *Gradient Boost* was chosen among all other approaches due to its power in separating signal from backgrounds.

8.1.2. Input Variables

In $VHbb$ analysis, MVAs are constructed using a list of discriminating variables. MVAs were initially built with m_{bb} and $\Delta R(\vec{b_1}, \vec{b_2})$. More variables are then added in order after m_{bb} and $\Delta R(\vec{b_1}, \vec{b_2})$, with the one offering the largest improvement insignificance first. The final MVA is constructed when a variable in sequence is added and no further improvement is seen. MVAs are constructed separately in each lepton channel and jet category, as shown

in Table 8.1. The variables used in each channel are listed in Table 8.2 and the explanation can be found in Table 8.3.

There are a few MVA variables newly added for the full Run-2 analysis. The binned MV2c10 distribution[9] was found to improve the sensitivity of the 0-lepton channel by $\sim$ 7% and 1-lepton channel by $\sim$ 10% due to the increased discrimination power against backgrounds where a c-jet or light-flavor jet has been misidentified as a b-jet, especially $W \to cq$ in the $t\bar{t}$ and Wt backgrounds. The track-based $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term[10] was found to increase the VH signal significance by 2-3% depending on the analysis region in the 0-lepton channel [109]. The $\cos\theta(\vec{\ell^{-}}, \vec{Z})$ variable uses the difference in polarization between the ZH signal and Z+jets background and improves the sensitivity in the 2-lepton channel by $\sim$ 7% [159, 178].

[9]The MV2c10 discriminant is grouped into two bins corresponding to efficiencies of 0-60% and 60-70% which are calibrated to data [155, 176, 177].

[10]Soft MET provides additional rejection against $t\bar{t}$ background. Due to the kinematic and detector acceptance, unreconstructed objects (such as leptons or b-jets) will result in a larger soft MET for $t\bar{t}$ compared to the VH signal [159].

Variable	Name	0-lepton	1-lepton	2-lepton
m_{bb}	mBB	✓	✓	✓
$\Delta R(\vec{b_1}, \vec{b_2})$	dRBB	✓	✓	✓
$p_{\mathrm{T}}^{b_1}$	pTB1	✓	✓	✓
$p_{\mathrm{T}}^{b_2}$	pTB2	✓	✓	✓
p_T^V	pTV	✓	✓	✓
$\Delta\phi(\vec{V}, \vec{bb})$	dPhiVBB	✓	✓	✓
binned MV2c10(b_1)	bin_MV2c10B1	✓	✓	
binned MV2c10(b_2)	bin_MV2c10B2	✓	✓	
$\lvert\Delta\eta(\vec{b_1}, \vec{b_2})\rvert$	dEtaBB	✓		
m_{eff}	MEff	✓		
track-based soft $E_{\mathrm{T}}^{\mathrm{miss}}$	softMET	✓		
$E_{\mathrm{T}}^{\mathrm{miss}}$	MET	$\equiv p_T^V$	✓	
$\min[\Delta\phi(\vec{\ell}, \vec{b})]$	dPhiLBmin		✓	
m_{T}^W	mTW		✓	
$\lvert\Delta\mathrm{y}(\vec{W}, \vec{bb})\rvert$	dYWH		✓	
m_{top}	mTop		✓	
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$	METSig			✓
$\lvert\Delta\eta(\vec{V}, \vec{bb})\rvert$	dEtaVBB			✓
$m_{\ell\ell}$	mLL			✓
$\cos\theta(\vec{\ell^-}, \vec{Z})$	cosThetaLep			✓
	Only in 3 Jet Events			
$p_T^{\mathrm{jet}_3}$	pTJ3	✓	✓	✓
m_{bbj}	mBBJ	✓	✓	✓

Table 8.2: Variables used to train the multivariate discriminant in each lepton channel [109].

Variable	Explanation
	commonly used in all lepton channels
m_{bb}	invariant mass of the two b-tagged jets
$\Delta R(\vec{b_1}, \vec{b_2})$	distance in η and ϕ between the two b-tagged jets
$p_{\mathrm{T}}^{\mathrm{b_1}}$	transverse momentum of the b-tagged jet in the dijet system with the higher p_{T}
$p_{\mathrm{T}}^{\mathrm{b_2}}$	transverse momentum of the b-tagged jet in the dijet system with the lower p_{T}
p_{T}^{V}	transverse momentum of the vector boson; reconstructed as $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 0 lepton channel, vectorial sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse momentum of the lepton in the 1 lepton channel and vectorial sum of the transverse momenta of the two leptons in the 2 lepton channel
$\Delta\phi(\vec{V}, \vec{bb})$	distance in ϕ between the vector boson candidate ($E_{\mathrm{T}}^{\mathrm{miss}}$ in the 0 lepton channel, $E_{\mathrm{T}}^{\mathrm{miss}}$ and the lepton in the 1 lepton channel and the di-lepton system in the 2 lepton channel) and the Higgs boson candidate (the dijet system constructed from the two b-tagged jets)
$p_{\mathrm{T}}^{\mathrm{jet_3}}$	transverse momentum of the signal jet with the highest transverse momentum amongst the signal non-b-jets; only used for events with 3 or more signal jets
m_{bbj}	invariant mass of the two b-tagged jets and the signal jet with the highest transverse momentum amongst the signal non-b-jets; only used for events with 3 or more signal jets
	0-lepton specific variables
binned MV2c10($\mathrm{b_1}$)	binned distribution of the MV2c10 b-tagging output score of the leading b-tagged jet. The bins in this distribution are identical with the bins defined for the pseudo-continuous b-tagging working point; more details are given in the object note for this analysis [179].
binned MV2c10($\mathrm{b_2}$)	binned distribution of the MV2c10 b-tagging output score of the sub-leading b-jet tagged. The bins in this distribution are identical with the bins defined for the pseudo-continuous b-tagging working point.
$\lvert\Delta\eta(\vec{b_1}, \vec{b_2})\rvert$	distance in η between the two b-tagged jets
m_{eff}	scalar sum of $E_{\mathrm{T}}^{\mathrm{miss}}$, the p_{T} of the two b-jets and the p_{T} of the third jet if present
track-based soft $E_{\mathrm{T}}^{\mathrm{miss}}$	the vectorial sum of the p_{T} of all tracks in the events that are not associated to any reconstructed event. This variable is identical with the track based soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term defined by the JetETmiss combined performance group, see Ref. [180].
	1-lepton specific variables
binned MV2c10($\mathrm{b_1}$)	binned distribution of the MV2c10 b-tagging output score of the leading b-tagged jet. The bins in this distribution are identical with the bins defined for the pseudo-continuous b-tagging working point.
binned MV2c10($\mathrm{b_2}$)	binned distribution of the MV2c10 b-tagging output score of the sub-leading b-tagged jet. The bins in this distribution are identical with the bins defined for the pseudo-continuous b-tagging working point.
$E_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse energy of the event
$\min[\Delta\phi(\vec{\ell}, \vec{b})]$	distance in ϕ between the lepton and the closest b-tagged jet
m_{T}^{W}	transverse mass of the W boson candidate; defined as $m_{\mathrm{T}}^{W} = \sqrt{2p_T^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos(\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})))}$
$\lvert\Delta\mathrm{y}(\vec{W}, \vec{bb})\rvert$	difference in rapidity between the Higgs boson candidate and W boson candidate
m_{top}	reconstructed mass of the leptonically decaying top quark
	2-lepton specific variables
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$	quasi-significance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, defined as $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$ with S_{T} the scalar sum of the p_{T} of the leptons and jets in the event.
$\lvert\Delta\eta(\vec{Z}, \vec{bb})\rvert$	distance in η between the dilepton and dijet system of the b-tagged jets
$m_{\ell\ell}$	invariant mass of the dilepton system
$\cos\theta(\vec{\ell^-}, \vec{Z})$	angle between the negatively charged lepton in the Z-boson rest frame and the Z-boson flight direction in the rest-frame

Table 8.3: Explanation of the MVA input variables in each lepton channel. Table is modified from Ref. [109].

8.1.3. Set-up and Training

BDTs are trained using all nominal MC samples in this analysis, with the mc16a, mc16d, and mc16e production period combined. The signal template is defined by the VH samples and the background template is defined by the sum of all background samples. In each region, the BDT training is split into two sub-trainings: one is trained with the even event-number events and evaluated with the odd ones; the other is trained with the odd event-number events and evaluated with the even ones. This ensures the orthogonality between the events on which we train and evaluate performance. The final discriminant is built by adding the discriminant distributions of the even and odd events since the physics between the even and odd events are expected to be the same. Because of the b-tagging efficiency, truth-tagging is applied to all events in the BDT training to maximize the MC statistic in the training[11]. This improves training performance and avoids over-training. The long tails of the BDT input variables often distract the BDT by wasting the degrees of freedom on classifying the small number of events. Thus, to increase the stability of the training, the range of the variables is limited to include 99% of the events, with the events beyond those limits artificially set to the extremum value of the selected range. The hyper-parameters used in the BDT training are listed in Table 8.4. The learning rate, number of trees, and the depth of trees are scanned coarsely to have the optimal set-up for this analysis. The final hyper-parameter set-up was chosen based on maximizing sensitivity without significant over-training.

To avoid any bias towards the training data set, BDT over-training was assessed using the training data set and evaluation data set. No significant difference in the BDT output score distribution is found between the training and the evaluation data set. Neither for the signal nor the background template. This is also reflected in the ROC curve and signal and

[11]The MC events entering the profile likelihood fit are hybrid-truth-tagged. It was tested to use hybrid truth tagging in the training. No change in the sensitivity of the MVA was found.

TMVA Setting	Value	Definition
Boost Type	gradient boosting	Boost procedure
Shrinkage	0.5	Learning rate
Separation Type	Gini index	Node separation gain
Prune Method	No Pruning	Pruning method
NTrees	200 (600 for 1-lepton VH)	Number of trees
Max Depth	4 (2 for 1-lepton diboson)	Maximum tree depth
nCuts	100	Number of equally spaced cuts tested per variable per node
nEventsMin	5%	Minimum number of events in a node (% of total events)

Table 8.4: Hyper-parameters used for the BDT training. Exceptions for the 1-lepton VH and diboson training are given in brackets [109].

background efficiencies[12] shown in Fig. 8.3. The diboson analysis retains the BDTs but using the diboson samples as the signal template and VH samples in the background template. All training configurations and hyper-parameter settings remain the same except those given in brackets in Table 8.4.

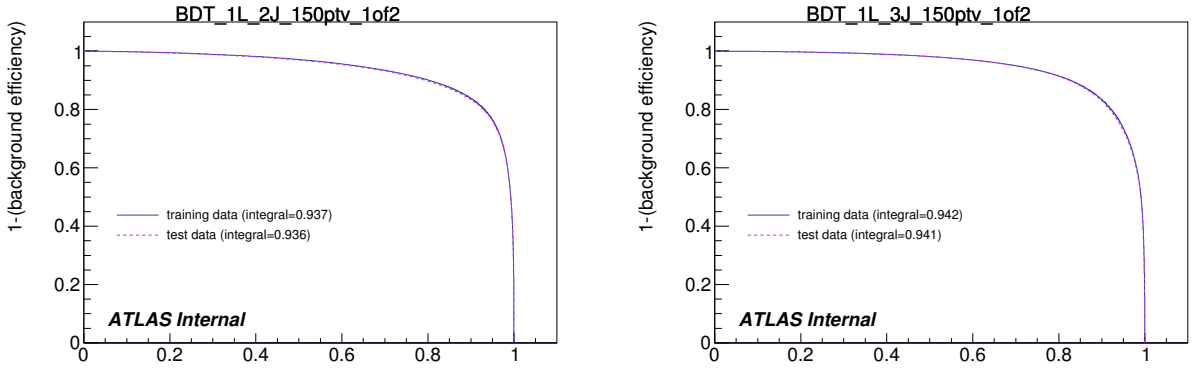


Figure 8.3: Over-training checks for the 2-jet (left) and 3-jet (right) category of p_{T}^{V} >150 GeV region trained on odd-number MC samples in the 1-lepton channel. The ROC curve compares the training data-set with a statistically independent testing data-set

[12]The area under the curve is a measurement of the signal-to-background discrimination performance of the BDT.

8.2. Systematic Uncertainties

The systematic uncertainties in this analysis can be categorized into two categories: one is the experimental uncertainties related to the detector design and physics object reconstruction; the other is the modeling uncertainty related to the theoretical models behind the Monte Carlo generators[13].

8.2.1. Experimental Uncertainty

The experimental uncertainties in this analysis come from the reconstructions of jets, leptons, and $E_{\mathrm{T}}^{\mathrm{miss}}$, as well as the processes of b-tagging, pile-up (re-weighting), and luminosity measurement.

In the jet reconstruction, the main uncertainties are from the jet energy scale (JES) and the jet energy resolution (JER). Many steps are involved in JES calibration and JER modeling, resulting in various uncertainties being included, detailed in [160, 181]. The electrons and muons used in $VHbb$ analysis introduce small uncertainties, most of which come from the uncertainties in reconstruction, identification, isolation, trigger efficiencies, energy scale, and resolution [182, 183]. For the $E_{\mathrm{T}}^{\mathrm{miss}}$, due to the way it is calculated, uncertainties in $E_{\mathrm{T}}^{\mathrm{miss}}$ result from various sources: the energy scale and resolution propagated from jets and leptons, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, soft term tracks, and even the underlying event model [184].

Jet b-tagging is the dominant experimental uncertainty. To correct the mis-modeling of the b-tagging algorithm, b-tagging scale factors are derived from the difference of efficiency measured in data and MC samples. These correction factors are derived separately for b-jets, c-jets, and light-flavor jets, with their uncertainties measured individually and decomposed into independent components [176, 177, 185]. The integrated luminosity in Run-2 is measured

[13]For the modeling using a data-driven method, its uncertainty is dominated by the data statistical uncertainty.

based on the LUCID-2 detector (see Sec. 3.2.6) and its uncertainty is 1.7%, as detailed in Sec. 4.1. To describe the pile-up in data, the average number of interactions per bunch crossing, μ, in the Monte Carlo samples is scaled by 1.03 and uncertainty on the scale factor is included [186].

The experimental uncertainties are summarized in Table 9.4.

8.2.2. Modeling Uncertainty

The modeling uncertainty is used to describe the uncertainties of the theoretical model used in the MC generators to model the real physics interaction processes. Because it is not possible to fully compute all physical processes and compare them with the nominal MC generators (which are believed to best describe the data), alternative MC generators with different assumptions and theoretical models are used in the $VHbb$ analysis to compare with the nominal MC generators. The differences between the nominal and the alternative generators are used to estimate the modeling uncertainties. Given that there are multiple steps in the Monte Carlo simulation and the $VHbb$ analysis, it will be far more complicated if comparing and calculating the differences between the nominal and alternative MC generators at each step. Thus, the differences between MC generators are calculated only once for the relevant variables before input to the final likelihood fit. Those differences are represented by the normalization uncertainty, the relative acceptance uncertainty, and the shape uncertainty, which are explained below.

Although the simulated physics processes in $VHbb$ have been normalized to the most accurate cross-section calculation available, there are still differences in the yield of the events, compared with the data observations of these processes. This difference is called "normalization uncertainty". It is derived from the ratio of the yield between the nominal

and alternative MC samples,

$$\frac{N_{\text{nominal}} - N_{\text{alternative}}}{N_{\text{nominal}}} \times 100\%. \tag{8.1}$$

Since the ratios of the yield between different analysis regions (and also different jet flavor compositions) are used to constrain the likelihood fit, the uncertainties of these ratios can be derived from the double-ratio method using the nominal and alternative MC samples, as in Eq. 8.2 [109]. These uncertainties are called "relative acceptance uncertainty,"

$$\frac{\mathit{Acceptance}[\text{Category}_{\text{A}}(\text{nominal MC})]}{\mathit{Acceptance}[\text{Category}_{\text{B}}(\text{nominal MC})]} \Big/ \frac{\mathit{Acceptance}[\text{Category}_{\text{A}}(\text{alternative MC})]}{\mathit{Acceptance}[\text{Category}_{\text{B}}(\text{alternative MC})]} \,. \tag{8.2}$$

The theoretical models can cause shape differences (the result of entries migrating in variable space) in the kinematic variable distributions, such as p_{T}^{V}, when compared with the real physics. Thus, the shape uncertainties are estimated from comparing the shape distribution of the kinematic variables between the nominal and alternative MC samples. Then, the shape in the alternative samples which has the largest deviation from the shape of the nominal sample is assigned as the "shape uncertainty".

In addition, a new method using a trained BDT (denoted "BDT_S") to categorize the nominal sample from the alternative (using the sample input variables as in BDT_{VH}) is also used for modeling the shape uncertainties [187]. The ratio of BDT_S distributions of the alternative and the nominal samples are used to derive a ratio function for the nominal sample reweighted to match the alternative sample. Thus, the shape differences between the nominal and alternative MC generators are represented by this ratio function, referred to as R_{BDT}[14].

[14]The input variable distribution for the nominal sample after R_{BDT} reweighting is verified to agree well with those in the alternative sample [159].

In this analysis, the shape uncertainties are derived in a way combining the shape comparison method and the BDT reweighting method. The shape uncertainty of p_{T}^{V} is derived using the shape comparison method due to its large impact when being input for the likelihood fit (see Sec. 8.3.2). The shape uncertainties of the MVA input variables other than p_{T}^{V} are derived using the BDT reweighting method[15].

8.2.2.1. V+jet

Given the flavor composition of the two b-tagged jets, the V+jets background can be divided into 3 main components: V+heavy flavor or V+HF ($V + bb$, $V + bc$, $V + b\ell$, $V + cc$), $V + c\ell$, and $V + \ell\ell$, where ℓ denotes lepton. Due to the requirement of two b-tagged jets, there are only $\sim 1\%$ $V + cl$ and $V + ll$ background in each analysis region. Thus, a single normalization uncertainty is sufficient for each of these backgrounds. For the V+HF background, its overall normalization is left free to float in the global likelihood fit, separately for 2-jet and 3-jet category. The relative acceptance uncertainties are derived for the ratio of yields in the 0-lepton and 1-lepton channel for the W+HF background, and in the 0-lepton and 2-lepton channel for the Z+HF background. The uncertainties of the flavor composition and SR-to-CR ratios are done in the same way as in Eq. 8.2. The shape uncertainties of W+HF are derived using the p_{T}^{V} distribution and the BDT_R method, while those of $Z + HF$ are derived using the m_{bb} and p_{T}^{V} from comparisons with data in the m_{bb} side-bands ($m_{bb} <$ 80 GeV or $m_{bb} >$ 140 GeV) after subtracting non Z+jets backgrounds [188].

8.2.2.2. $t\bar{t}$

The modeling of $t\bar{t}$ background in the 0-lepton and 1-lepton channels is where I contributed to the VHbb analysis.

[15]This is achieved by re-weight the nominal p_{T}^{V} distribution to match the alternative one before the BDT_S training [188].

In the 0-lepton and 1-lepton channels, the $t\bar{t}$ normalization parameters are set free to float separately for 2-jet and 3-jet category. The relative acceptance uncertainty is derived from the ratio of yield in the 0-lepton and 1-lepton channel using the double-ratio method as shown in Eq. 8.2. Although the dominant flavor component of the two b-tagged jets is bb in $t\bar{t}$, the uncertainties on the ratio of different components (bb, bc, and other) are estimated [159]. The shape uncertainty in the 0-lepton and 1-lepton channels are derived using the combined method previously described in Sec. 8.2.2. The shape uncertainties derived from the shape comparison method are shown in Sec. 8.2.2.2.1 in detail.

In the 2-lepton channel, due to the existence of the $e\mu$-control region (described in Sec. 6.2), the $t\bar{t}$ background (and single top Wt as well) is modeled with a data-driven method. Since there is almost no difference in the kinematic variable distributions between the $e\mu$-control region and the signal region, the $t\bar{t}$ (and single top Wt) events can be directly used to model those in the signal region, and calculate the normalization and shape uncertainties. The possible normalization difference caused by the lepton trigger, reconstruction and identification between the signal region and the $e\mu$-control region, is corrected by a scale factor α, shown in Eq. 8.3 [188],

$$N^{\mathrm{SR}}_{top,data} = \frac{N^{\mathrm{CR}}_{top,data}}{N^{\mathrm{CR}}_{top,MC}} \times N^{\mathrm{SR}}_{top,MC} = \frac{N^{\mathrm{SR}}_{top,MC}}{N^{\mathrm{CR}}_{top,MC}} \times N^{\mathrm{CR}}_{top,data}, \tag{8.3}$$

where $\alpha \equiv \frac{N^{\mathrm{SR}}_{top,MC}}{N^{\mathrm{CR}}_{top,MC}}$. Then, the statistical uncertainty in the $e\mu$-control region becomes the dominant uncertainty, and most of the systematic uncertainties on the extrapolation parameter cancel, especially the theoretical modeling and jet-related systematic uncertainties.

8.2.2.2.1 $t\bar{t}$ shape uncertainty (shape comparison method)

This subsection describes the shape comparison method I used to derive the shape uncertainties of $t\bar{t}$ (and other backgrounds) process. This method is used before the study of the

BDTr method concludes. Since the VHbb analysis uses the BDT score as the fit discriminant to extract the final results, the p_{T}^{V} and $m_{b\bar{b}}$ distributions, which are the most important underlying physics quantities, are used for the shape uncertainties. The procedure of deriving the shape uncertainties of p_{T}^{V} and $m_{b\bar{b}}$ using the shape comparison method is as follows:

- compare the p_{T}^{V} and $m_{b\bar{b}}$ distributions of the alternative MC samples with the nominal ones;
- a ratio function was fit for each of the distribution comparison;
- the ratio function which has the largest deviation from the nominal is assigned as the shape uncertainty.

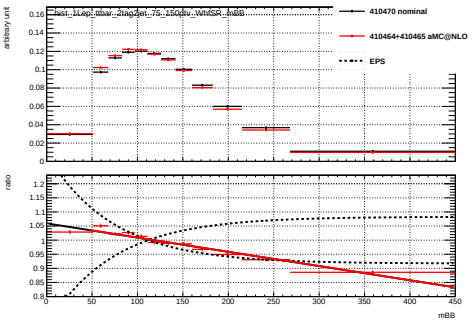
The p_{T}^{V} and $m_{b\bar{b}}$ shape uncertainties derived in the signal regions are shown in Fig. 8.4 and 8.5. The details are documented in Appendix C.1.1 and C.1.2. The nuisance parameters which the p_{T}^{V} and $m_{b\bar{b}}$ shape uncertainty are input to for the likelihood fit are shown in Table 8.5.

8.2.2.3. Single Top-quark

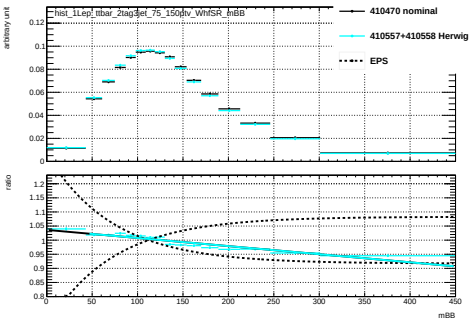
The single-top background has 3 channels: Wt-, t- and s-channel. The s-channel has a small contribution so only a normalization uncertainty is assigned. For the t-channel, uncertainties on normalization, relative acceptance, and shape of the m_{bb} and p_{T}^{V} distributions

Analysis region	Uncertainty	Type	Source	Nuisance Parameter
0,1 lepton	p_T^V shape	shape + norm	aMC@NLO+Pythia 8 (ME, 2-/3-jet)	`TTbarPTV`
2 lepton	p_T^V shape + norm	aMC@NLO+Pythia 8 (ME)	`TTbarPTV_L2`	
0,1 lepton	$m_{b\bar{b}}$ shape	shape + norm	aMC@NLO+Pythia 8 (ME, 2-jet) + Herwig7 (PS, 3-jet)	`TTbarMBB`
2 lepton	$m_{b\bar{b}}$ shape	shape	aMC@NLO+Pythia 8	`TTbarMBB_L2`

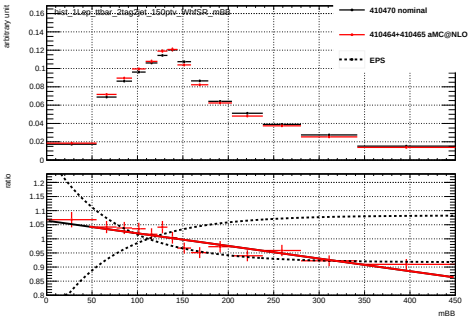
Table 8.5: Summary of all shape uncertainties for the $t\bar{t}$ process with short descriptions and the name of the corresponding nuisance parameters.



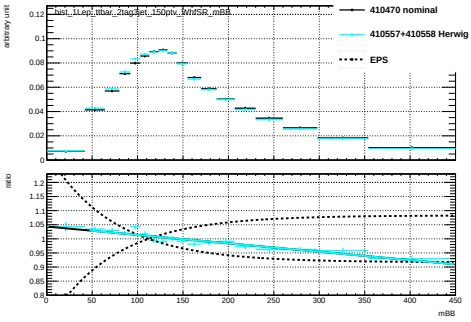
(a) 2-jet, 75 GeV $< p_T^V <$ 150 GeV



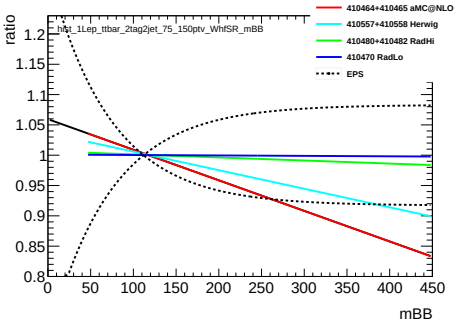
(b) 3-jet, 75 GeV $< p_T^V <$ 150 GeV



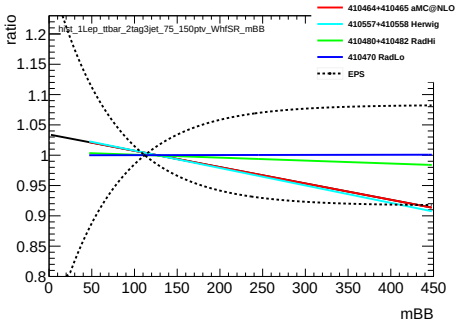
(c) 2-jet, $p_T^V >$ 150 GeV



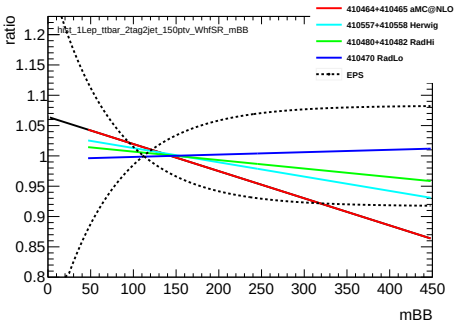
(d) 3-jet, $p_T^V >$ 150 GeV



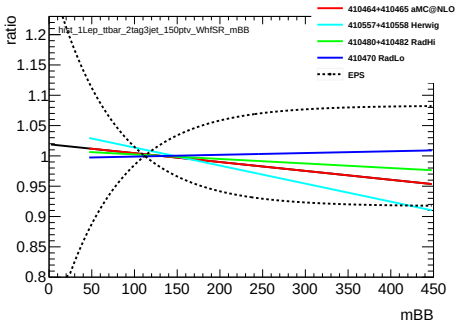
(e) 2-jet, 75 GeV $< p_T^V <$ 150 GeV



(f) 3-jet, 75 GeV $< p_T^V <$ 150 GeV

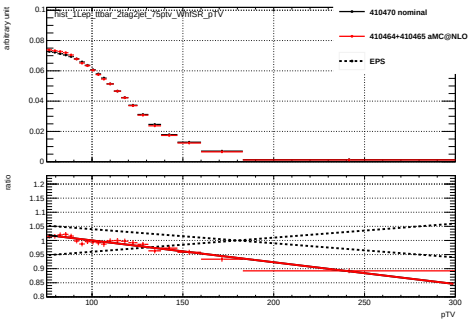


(g) 2-jet, $p_T^V >$ 150 GeV

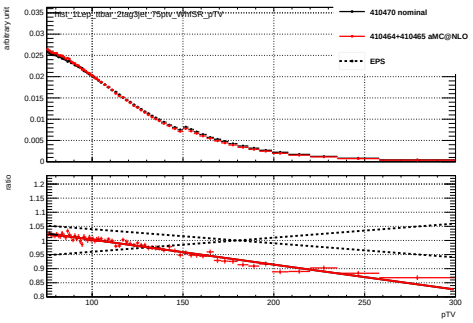


(h) 3-jet, $p_T^V >$ 150 GeV

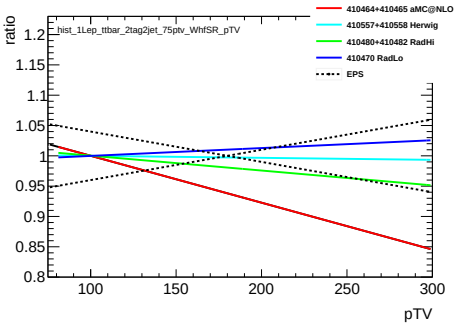
Figure 8.4: $m_{b\bar{b}}$ shape variations used in the 0-/1-lepton channel for the 75GeV $< p_T^V <$ 150GeV and $p_T^V >$ 150GeV regions. Figures (a-d) show the systematic variations selected from the all MC-to-MC comparisons summarized in figures (e-h).



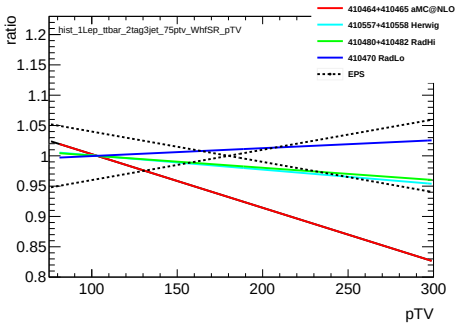
(a) 2-jet, $p_T^V > 150$ GeV



(b) 3-jet, $p_T^V > 150$ GeV



(c) 2-jet, $p_T^V > 150$ GeV



(d) 3-jet, $p_T^V > 150$ GeV

Figure 8.5: p_T^V shape variations used in the 0-/1-lepton channel for the 2-/3-jet categories. Figures (a-b) show the systematic variations selected from the all MC-to-MC comparisons summarized in figures (c-d).

are derived using the alternative sample variation method. For the Wt-channel, the flavor composition of the two b-tagged jets is divided into bb and "other" two categories and their relative acceptance uncertainties and shape uncertainties of m_{bb} and p_{T}^{V} distributions are evaluated separately.

8.2.2.4. Diboson Production

The diboson backgrounds originate from 3 processes: WZ, WW and ZZ. For WW, only normalization uncertainty is assigned because of its small contribution[16]. For WZ, and ZZ, the uncertainties on normalization, relative acceptance and shape uncertainties of m_{bb} and p_{T}^{V} are estimated and shown in Table 8.8.

8.2.2.5. Multi-jet Background in the 1-lepton Channel

The QCD multi-jet background has a large production cross-section at LHC, as discussed in Sec. 2.2. It can arise from the jet energy mis-measurement in the 0-lepton channel or the jet-faked-leptons in the 1-lepton and 2-lepton channel. Due to the selection cuts applied, the multi-jet events in the 0-lepton and 2-lepton channels are negligible. In the 1-lepton channel, the contribution of the multi-jet background is small (level of a few percent). It is modeled from the data by creating a multi-jet control region (MJ-CR), which is enriched with multi-jet events and from which all other simulated backgrounds are subtracted. The MJ-CR is obtained by inversion of the tight isolation cut, with the loose isolation cut still applied.

The requirement on the number of b-tagged jets is loosened from 2 to 1 to reduce the impact of statistical fluctuation by increasing the yield of multi-jet events. The normalization of the multi-jet background is estimated from a template fit to the m_T^W distribution after

[16]The contribution of WW is less than 0.1% of the total background [188].

the nominal selection with the 2 b-tag requirement, where the multi-jet shape is taken from the 1-b-tag MJ-CR and shapes of other backgrounds taken from Monte Carlo simulation. The normalization uncertainty is derived from changes to the m_T^W template distribution and relative yield [189]. Since the MV2c10 scores of the two b-tagged jets are used in the BDT, while one b-tagged jet is required in MJ-CR, a method using the 2D distribution of each jet's MV2c10 score is used to emulate those scores for the one b-tagged multi-jet event [188].

8.2.2.6. VH Signal

The systematic uncertainties in the calculations of the VH production cross-section and the $H \to b\bar{b}$ branching ratio are assigned using the recommendations of the LHC Higgs Cross Section Working Group [92, 93, 190–192]. The m_{bb} shape uncertainties from the QCD scale variation, parton shower/underlying event variation, and PDF+α_S are derived from comparing the relative MC samples or error sets[17]. The uncertainty on the p_{T}^V shape migration originating from the higher-order electro-weak correction is evaluated for the $qq \to VH$ production [24, 169, 193, 194].

For the STXS measurement, uncertainties of the acceptance and shapes of variables are estimated based on the STXS regions. The uncertainties of the theoretical cross-section are not included in the likelihood fit since they only affect the predictions with which they are compared [159].

[17]For the QCD scale uncertainty, the samples used for comparison are the factorization and renormalization scale varied samples. For the PDF+α_S uncertainty, the ones used are the PDF error set PDF4LHC15_30 and the α_S uncertainties which follow the PDF4LHC recommendation for Run-2.

Z + jets	
$Z + ll$ normalization	18%
$Z + cl$ normalization	23%
Z + HF normalization	Floating (2-jet, 3-jet) × ($75\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{V}} < 150\,\mathrm{GeV}, p_{\mathrm{T}}^{\mathrm{V}} > 150\,\mathrm{GeV}$)
$Z + bc$-to-$Z + bb$ ratio	30 – 40%
$Z + cc$-to-$Z + bb$ ratio	13 – 16%
$Z + bl$-to-$Z + bb$ ratio	20 – 28%
SR-to-low ΔR CR ratio	3.8 – 9.9% ($75\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{V}} < 150\,\mathrm{GeV}, p_{\mathrm{T}}^{\mathrm{V}} > 150\,\mathrm{GeV}$)
SR-to-high ΔR CR	2.7 – 4.1% ($75\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{V}} < 150\,\mathrm{GeV}, p_{\mathrm{T}}^{\mathrm{V}} > 150\,\mathrm{GeV}$)
0-to-2 lepton ratio	7%
$p_{\mathrm{T}}^{\mathrm{V}}$	M+S ($75\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{V}} < 150\,\mathrm{GeV}, p_{\mathrm{T}}^{\mathrm{V}} > 150\,\mathrm{GeV}$)
m_{bb}	S ($75\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{V}} < 150\,\mathrm{GeV}, p_{\mathrm{T}}^{\mathrm{V}} > 150\,\mathrm{GeV}$)
W + jets	
$W + ll$ normalization	32%
$W + cl$ normalization	37%
W + HF normalization	Floating (2-jet, 3-jet)
$W + bc$-to-$W + bb$ ratio	15% (0-lepton) and 30% (1-lepton)
$W + cc$-to-$W + bb$ ratio	10% (0-lepton) and 30% (1-lepton)
$W + bl$-to-$W + bb$ ratio	26% (0-lepton) and 23% (1-lepton)
SR-to-CR ratio	3.6-15%
0-to-1 lepton ratio	5%
$p_{\mathrm{T}}^{\mathrm{V}}$	M+S (2-jet, 3-jet)
$\mathrm{R_{BDT}}$	S
$t\bar{t}$ (0+1-lepton channels only)	
$t\bar{t}$ normalization	Floating (2-jet, 3-jet)
0-to-1 lepton ratio	8%
$t\bar{t}$ (flavor composition) bc-to-bb ratio (ME)	7.6 – 8.2% (0-lepton), 1.3 – 3.8% (1-lepton)
$t\bar{t}$ (flavor composition) bc-to-bb ratio (PS)	2.1 – 3.2% (0-lepton), 1.5 – 7.1% (1-lepton)
$t\bar{t}$ (flavor composition) other-to-bb ratio (ME)	2.8 – 6.4% (0-lepton), 3.3 – 5.7% (1-lepton)
$t\bar{t}$ (flavor composition) other-to-bb ratio (PS)	5.6 – 13% (0-lepton), 0.3 – 2.1% (1-lepton)
$p_{\mathrm{T}}^{\mathrm{V}}$	M+S (2-jet, 3-jet)
$\mathrm{R_{BDT}}$ ME variation	M+S (2-jet, 3-jet)
$\mathrm{R_{BDT}}$ PS variation	M+S (0-lepton, 1-lepton)
Single top-quark	
Cross-section	4.6% (s-channel), 4.4% (t-channel), 6.2% (Wt)
Acceptance 2-jet	17% (t-channel), 55% ($Wt(bb)$), 24% (Wt(other))
Acceptance 3-jet	20% (t-channel), 51% ($Wt(bb)$), 21% (Wt(other))
m_{bb}	M+S (t-channel, $Wt(bb)$, Wt(other))
$p_{\mathrm{T}}^{\mathrm{V}}$	M+S (t-channel, $Wt(bb)$, Wt(other))
Multi-jet (1-lepton)	
Normalization	30 – 200% (2-jet), 100% (3-jet)
BDT template	M+S

Table 8.6: Summary of the systematic uncertainties in the background modeling for Z + jets, W + jets, $t\bar{t}$, single top-quark, and multi-jet production. 'ME' represents the matrix element generator variation and 'PS' represents the parton shower generator variation. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159].

	Signal
Cross-section (scale)	0.7% (qq), 25% (gg)
$H \to b\bar{b}$ branching fraction	1.7%
Scale variations in STXS bins	3.0 – 3.9% ($qq \to WH$), 6.7 – 12% ($qq \to ZH$), 37 – 100% ($gg \to ZH$)
PS/UE variations in STXS bins	1 – 5% for $qq \to VH$, 5 – 20% for $gg \to ZH$
PDF+α_S variations in STXS bins	1.8 – 2.2% ($qq \to WH$), 1.4 – 1.7% ($qq \to ZH$), 2.9 – 3.3% ($gg \to ZH$)
m_{bb} from scale variations	M+S ($qq \to VH$, $gg \to ZH$)
m_{bb} from PS/UE variations	M+S
m_{bb} from PDF+α_S variations	M+S
p_T^V from NLO EW correction	M+S

Table 8.7: Summary of the systematic uncertainties in signal modeling. 'PS/UE' represents the parton shower/underlying event. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159].

ZZ	
Normalization	20%
0-to-2 lepton ratio	6%
Acceptance from scale variations	10 – 18%
Acceptance from PS/UE variations for 2 or more jets	6%
Acceptance from PS/UE variations for 3 jets	7% (0-lepton), 3% (2-lepton)
m_{bb} from scale variations	M+S (correlated with WZ uncertainties)
$p_{\mathrm{T}}^{\mathrm{V}}$ from scale variations	M+S (correlated with WZ uncertainties)
m_{bb} from PS/UE variations	M+S (correlated with WZ uncertainties)
$p_{\mathrm{T}}^{\mathrm{V}}$ from PS/UE variations	M+S (correlated with WZ uncertainties)
m_{bb} from matrix-element variations	M+S (correlated with WZ uncertainties)
WZ	
Normalization	26%
0-to-1 lepton ratio	11%
Acceptance from scale variations	13 – 21%
Acceptance from PS/UE variations for 2 or more jets	4%
Acceptance from PS/UE variations for 3 jets	11%
m_{bb} from scale variations	M+S (correlated with ZZ uncertainties)
$p_{\mathrm{T}}^{\mathrm{V}}$ from scale variations	M+S (correlated with ZZ uncertainties)
m_{bb} from PS/UE variations	M+S (correlated with ZZ uncertainties)
$p_{\mathrm{T}}^{\mathrm{V}}$ from PS/UE variations	M+S (correlated with ZZ uncertainties)
m_{bb} from matrix-element variations	M+S (correlated with ZZ uncertainties)
WW	
Normalization	25%

Table 8.8: Summary of the systematic uncertainties in the background modeling for diboson production. 'PS/UE' represents the parton shower/underlying event. In the 'M+S' symbol, 'M' indicates that the shape uncertainty includes a migration effect that allows relative acceptance changes between regions, and 'S' indicates that the uncertainty only acts on the shape in the signal region. When extracting the $(W/Z)Z$ diboson production signal yield, as the normalizations are unconstrained, the normalization uncertainties are removed. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed. The table is from the publication of the $VHbb$ analysis of ATLAS [159].

8.3. Statistical Analysis

The statistical analysis of $VHbb$ is achieved by applying a global likelihood fit to all channels and regions at the same time, with the signal strength and parameters of interest (POI) extracted by maximizing the likelihood. The procedure is explained in the following subsections in detail.

8.3.1. Hypothesis

The *hypothesis test* has been widely used not only in particle physics but also a lot of other fields. In the $VHbb$ analysis, it is used to determine whether the collected data is better described by background-only or background-plus-signal events. Thus, two hypotheses are made:

- H_0, the *null hypothesis*, where the data is background-only;
- H_1, the *alternative hypothesis*, where the data is signal-plus-background.

To express those hypotheses in mathematics, a scale factor, μ, is introduced, such that,

$$n^{\text{expected}}(\mu) = \mu \cdot s^{\text{expected}} + b^{\text{expected}}, \tag{8.4}$$

where $n^{\text{expected}}(\mu)$ is the total expected number of events to be compared with the collected data, s^{expected} is the expected number of signal events, and b^{expected} is the expected number of background events[18]. μ is also referred to as the *signal strength*, defined as the ratio of

[18] s^{expected} and b^{expected} are the statistical expectations of the number of signal and background events [195]. Both of them depend on their theoretical models, which are from the Monte Carlo simulation in most of the cases in $VHbb$.

the measured over the expected signal cross-section, $\mu = \sigma / \sigma_{\mathrm{SM}}$. Then, the hypotheses can be written as[19]:

- H_0, the *null hypothesis*, $\mu = 0$;
- H_1, the *alternative hypothesis*, $\mu \neq 0$.

Before testing those hypotheses, the likelihood function has to be constructed.

8.3.2. Likelihood Function

For each bin of the kinematic variable histogram, the number of events in it follows the Poisson distribution[20]. Then, the global likelihood function can be written as the product of Poisson probability terms over all bins of the input histograms [109],

$$\mathcal{L}(\lambda_i) = \prod_{i \in \text{ bins}} \frac{\lambda_i^{k_i} e^{-\lambda_i}}{k_i!}, \tag{8.5}$$

where λ_i is the expected value of number of events in bin i, which is $n_i^{\text{expected}}(\mu)$, and k_i is the observed value of number of events in bin i, which is n_i^{data}. Thus, given of Eq. 8.4, the global likelihood function can be rewritten as,

$$\mathcal{L}(\mu) = \prod_{i \in \text{ bins}} \frac{(\mu \cdot s_i^{\text{expected}} + b_i^{\text{expected}})^{n_i^{\text{data}}} \cdot e^{-(\mu \cdot s_i^{\text{expected}} + b_i^{\text{expected}})}}{n_i^{\text{data}}!}. \tag{8.6}$$

[19] In most cases, μ is assumed to be $\mu \geq 0$. The hypotheses could be much more complicated depending on the measurement [195].

[20] The Poisson distribution expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event [196]. A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$, if, for $k = 0, 1, 2, \ldots$, the probability density function of X is given by: $f(k; \lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$, where λ is equal to the expected value of X and also to its variance, $\lambda = \mathrm{E}(X) = \mathrm{Var}(X)$ [197].

However, due to the uncertainties of the shape and normalization of the distributions[21], the expected number of events, s_i^{expected} and b_i^{expected}, at bin i is not 100% known. s_i^{expected} and b_i^{expected} are dependent on those uncertainties, and thus can be written as $s_i^{\text{expected}}(\boldsymbol{\theta}_s^{\text{shape}})$[22] and $b_i^{\text{expected}}(\boldsymbol{\theta}_b^{\text{shape}}, \theta_b^{\text{normalization}})$. Those uncertainties are together represented as *nuisance parameters* (NPs), $\boldsymbol{\theta} = (\boldsymbol{\theta}_s^{\text{shape}}, \boldsymbol{\theta}_b^{\text{shape}}, \theta_b^{\text{normalization}})$ [195]. To obtain more information on the normalization and shapes, subsidiary measurements are made, such as the kinematic variables in the control regions and the flavor composition of the two b-tagged jets. These values are implemented as priors of Gaussian or log-normal probability density functions to constrain the relative *nuisance parameters*[23]. The Gaussian or log-normal probability density functions are multiplied with the likelihood functions, and they can be together denoted as $\mathcal{L}_{\text{NP}}(\boldsymbol{\theta})$[24],

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i \in \text{ bins}} \frac{(\mu \cdot s_i^{\text{expected}} + b_i^{\text{expected}})^{n_i^{\text{data}}} \cdot e^{-(\mu \cdot s_i^{\text{expected}} + b_i^{\text{expected}})}}{n_i^{\text{data}}!} \cdot \mathcal{L}_{\text{NP}}(\boldsymbol{\theta}). \tag{8.7}$$

8.3.3. Hypothesis Testing

To test a hypothesized value of μ, the ratio based on the likelihood function in Eq. 8.7 is constructed,

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{8.8}$$

[21]The scale factors from the comparison of data and Monte Carlo simulation enter the likelihood fit standalone from the *nuisance parameter*.

[22]The uncertainty on the signal normalization $\theta_s^{\text{normalization}}$ is not included in s_i^{expected}, since the signal strength μ has existed (whose uncertainty expresses the same physics meaning).

[23]Those *nuisance parameters* follow the Gaussian or log-normal probability density functions, which means that their value having a large deviation from the prior (or central) value is small.

[24]For example, $\mathcal{L}_{\text{NP}}(\boldsymbol{\theta})$ could be $\prod_{j \in \text{ nsyst}} \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-(\widetilde{\theta}_j - \theta_j)^2/2\sigma_j^2}$, where $\widetilde{\theta}_j$ central value of the Gaussian for *nuisance parameter* j, θ_j is the corresponding best fit value, and σ_j is the prior uncertainty of θ_j [198].

where in the numerator, $\hat{\hat{\boldsymbol{\theta}}}$ denotes the value of $\boldsymbol{\theta}$ that maximizes $\mathcal{L}$ for a given μ (conditional likelihood), and in the denominator, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximize the likelihood $\mathcal{L}$ (unconditional likelihood). The maximization happens while fitting to the data and estimating the values of $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$[25]. If the data favors a hypothesized value of μ, $\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})$ and $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})$ will yield close maximized results, and $\lambda(\mu)$ will be close to 1 ($0 \leq \lambda(\mu) \leq 1$).

Conveniently, we can define the statistic, t_μ, for the test of hypothesis,

$$t_\mu = -2\, ln\, \lambda(\mu), \tag{8.9}$$

where the larger incompatibility between the data and the hypothesized μ, the higher value of t_μ is. To quantify the level of incompatibility, the p-value is defined as:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu)\, dt_\mu, \tag{8.10}$$

where $t_{\mu,obs}$ is the value of the statistic t_μ observed from the data, and $f(t_\mu|\mu)$ is the probability density function of t_μ under the hypothesized value of μ[26] [195]. The relations of p-value with respect to $t_{\mu,obs}$ and significance Z[27] are shown in Fig. 8.6.

With the hypothesis $\mu = 0$ in Sec. 8.3.1, if the observed $t_{0,obs}$ is 5σ significance deviated from this hypothesis[28], then it leads to the claim of the discovery of a new signal.

---

[25]This is why λ is a function of μ.

[26]The probability density function is given by the asymptotic formula; more related can be found in Ref. [195].

[27]The measured μ value is extracted from the unconditional fit to the data, where $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})$ is maximized. The significance of μ is evaluated using the likelihood function $\mathcal{L}(\mu, \boldsymbol{\theta})$, with μ changed near the measured value and $\boldsymbol{\theta}$ is refitted. The expected (median) significance of μ can be estimated by a representative data set, Asimov data set. In Asimov data set, all parameters remain at their expected (nominal) value, including the nuisance parameters, although they could possibly be constrained. Asimov data set is normally constructed from the Monte Carlo samples. For more about obtaining experimental sensitivity using Asimov data set, refer to Ref. [195, 198, 199].

[28]If the observed $t_{0,obs}$ is 5σ significance deviated from the hypothesis, it means that the observed data has a high degree of incompatibility with the null hypothesis.

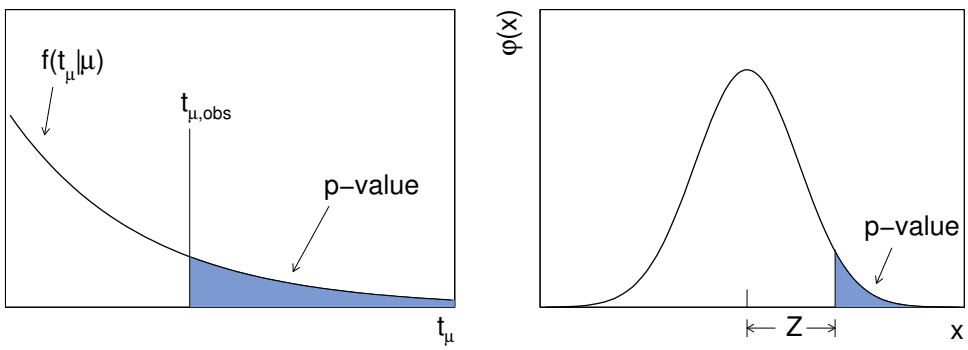


Figure 8.6: (left) Illustration of the p-value obtained from an observed value of the test statistic $t_{\mu,\text{obs}}$. (right) Illustration of the relation between the significance Z and the p-value for a standard normal distribution $\varphi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$ [195].

8.3.4. Fit to Data

The global likelihood fit is performed simultaneously in the 14 signal regions of the three channels, with the 28 control regions input as event yields in all fit configurations. The signal regions and control regions are shown in Table 6.3. The signal and background processes which are described in Sec. 4.2 and Sec. 8.2.2, are included in the fit[29]. All processes modeled by the Monte Carlo are hybrid-truth-tagged, as described in Sec. 5.2.2.2.1.

In the nominal VH analysis, the output distributions of BDT_{VH} are used as input to the fit. Three POI configurations are evaluated for the fit:

(i) a single-POI fit to measure the signal strength μ_{VH}^{bb}[30];

(ii) a two-POI fit to measure the signal strengths of μ_{WH}^{bb} and μ_{ZH}^{bb} simultaneously;

(iii) a five-POI fit to measure the cross-section in the five STXS regions (multiplied by the branching ratios of $H \to b\bar{b}$ and $V \to$ leptons) [159].

[29] As described in Sec. 8.2.2.2, in the 2-lepton channel, $t\bar{t}$ and single-top processes are estimated by the data-driven method and considered as a single template.

[30] This is the signal strength after multiplying the VH production cross-section and the branching ratio into $b\bar{b}$.

To determine the robustness of the results against assumptions, the dijet mass, and diboson analysis are conducted to validate the VH analysis. The dijet mass analysis uses the m_{bb} distribution (instead of the output distributions of BDT_{VH}) as input to the fit. It only has a single-POI fit to measure μ^{bb}_{VH}. The diboson analysis is used to measure the signal strength of WZ and ZZ processes. It uses the output distributions of BDT_{VZ} as input to the fit, where the SM Higgs boson is included as a background process[31]. There are two POI configurations studied:

(i) a single-POI fit to measure the signal strength μ^{bb}_{VZ}[32];

(ii) a two-POI fit to measure μ^{bb}_{WZ} and μ^{bb}_{ZZ} together [159].

The systematic uncertainties are implemented into the fit as *nuisance parameters*, as described in Sec. 8.2. In addition, the procedures of "smoothing" and "pruning" are applied to reduce the statistical fluctuation and simplify the fit. Since independent nuisance parameters have been introduced to account for the Monte Carlo statistical uncertainties, the inflation of the shapes in systematic uncertainties due to limited statistics should be smoothed out. The "smoothing" algorithms merge bins from one extremum to the next until no local extrema are left in the BDT distribution and the statistical uncertainty in each bin is less than 5%[33]. The "pruning" procedure removes the uncertainties which have a negligible effect on the distributions that entered the fit, for example, the normalization uncertainty variation that is less than 0.5%. The detailed steps of "smoothing" and "pruning" are documented in Ref. [109].

[31]The SM Higgs boson process is normalized to the predicted SM cross-section which of the uncertainty is set to 50%. It is believed to be conservative enough to cover previous measurement and uncertainty [159].

[32]This is the combined signal strength of WZ and ZZ diboson processes.

[33]The smoothing procedure is applied to the uncertainties associated to e/γ, $E_{\mathrm{T}}^{\mathrm{miss}}$, muons, taus, JVT, jet, PRW(pile-up reweighting) and multi-jet modeling shapes [109].

CHAPTER 9

Results

The likelihood fit to the data described in Sec. 8.3.4 leads us to the results in this chapter, where we will discuss the measurements of the signal strength and cross-section of the $VHbb$ analysis, as well as its cross-check and validations.

9.1. Signal Strength Measurements

9.1.1. VH Signal Strength

One of the main goal of the $VHbb$ analysis is to extract the signal strength of the process. After fitting to data, the normalization factors of the $t\bar{t}$ and V+jets background are displayed in Table 9.1. The yields of the signal, backgrounds and data for each category in the 0-, 1- and 2-lepton channels are shown in Table 9.2 and 9.3. The BDT distributions of the 2-jet category in each of the lepton channel are shown in Fig. 9.1. For a Higgs boson mass of 125 GeV, the fitted value of the VH signal strength is:

$$\mu_{VH}^{bb} = 1.02^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.}),$$

as shown in Fig. 9.2. Therefore, for the VH production, the background-only hypothesis is rejected with a significance of 6.7 standard deviations, where the expectation is 6.7 standard deviations.

The results of the simultaneous fit to measure the signal strengths separately to the WH and ZH production are presented in Fig. 9.3. The fitted values of the two signal strengths

are:

$$\mu_{WH}^{bb} = 0.95^{+0.27}_{-0.25} = 0.95 \pm 0.18(\text{stat.})^{+0.19}_{-0.18}(\text{syst.}),$$
$$\mu_{ZH}^{bb} = 1.08^{+0.25}_{-0.23} = 1.08 \pm 0.17(\text{stat.})^{+0.18}_{-0.15}(\text{syst.}).$$

The WH and ZH production reject the background-only hypothesis with observed significances of 4.0 and 5.3 standard deviations, with the expected significances of 4.1 and 5.1 standard deviations respectively[1].

The sources of uncertainties on the measurement of the VH, WH and ZH signal strengths are summarized in Table 9.4. The total statistical uncertainty is defined as the uncertainty in μ when all the NPs are fixed to their best-fit values. The total systematic uncertainty is then computed as the difference in quadrature between the total uncertainty in μ and the total statistical uncertainty [159]. For the WH and ZH signal strength measurements, the total statistical and systematic uncertainties are similar in size with the b-tagging, jet, $E_{\mathrm{T}}^{\mathrm{miss}}$, background modeling and signal systematic uncertainties. The statistical uncertainty from simulated event samples has been significantly reduced compared to the previous result, due to the increase on the number of simulated events.

[1]The simultaneous/combined fit has a linear correlation of 2.7% between the WH and ZH production.

Process and Category	Normalization factor
$t\bar{t}$ 2-jet	0.98 ± 0.09
$t\bar{t}$ 3-jet	0.93 ± 0.06
$W + \text{HF}$ 2-jet	1.06 ± 0.11
$W + \text{HF}$ 3-jet	1.15 ± 0.09
$Z + \text{HF}$ 2-jet, $75 < p_{\text{T}}^{\text{V}} < 150$ GeV	1.28 ± 0.08
$Z + \text{HF}$ 3-jet, $75 < p_{\text{T}}^{\text{V}} < 150$ GeV	1.17 ± 0.05
$Z + \text{HF}$ 2-jet, 150 GeV $< p_{\text{T}}^{\text{V}}$	1.16 ± 0.07
$Z + \text{HF}$ 3-jet, 150 GeV $< p_{\text{T}}^{\text{V}}$	1.09 ± 0.04

Table 9.1: Factors applied to the nominal normalization of the $t\bar{t}$, $W + \text{HF}$ and $Z + \text{HF}$ backgrounds, as obtained from the global likelihood fit to the 13 TeV data for the nominal multivariate analysis. The errors represent the combined statistical and systematic uncertainties.

Signal regions	0-lepton		0-lepton		1-lepton		1-lepton	
	$150\,\text{GeV} < p_{\text{T}}^{\text{V}} < 250\,\text{GeV}$		$p_{\text{T}}^{\text{V}} > 250\,\text{GeV}$		$150\,\text{GeV} < p_{\text{T}}^{\text{V}} < 250\,\text{GeV}$		$p_{\text{T}}^{\text{V}} > 250\,\text{GeV}$	
Sample	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet
Z + jets	2846 ± 80	3830 ± 160	338 ± 13	533 ± 23	102 ± 5	207 ± 11	10 ± 1	25 ± 2
W + jets	634 ± 63	1500 ± 110	83 ± 9	220 ± 19	1850 ± 160	4080 ± 270	353 ± 33	935 ± 69
Single top	237 ± 35	770 ± 130	9 ± 2	36 ± 8	990 ± 160	3570 ± 600	70 ± 15	313 ± 66
$t\bar{t}$	1157 ± 76	5470 ± 220	39 ± 5	151 ± 16	4600 ± 210	21030 ± 620	188 ± 19	970 ± 58
Diboson	360 ± 55	333 ± 79	86 ± 13	70 ± 17	229 ± 57	264 ± 83	57 ± 15	68 ± 22
Multi-jet	–	–	–	–	130 ± 47	24 ± 14	16 ± 10	7 ± 8
Total bkg.	5234 ± 63	11910 ± 100	554 ± 15	1011 ± 21	7899 ± 82	29170 ± 170	694 ± 23	2318 ± 42
Signal ($\mu = 1.02$)	147 ± 24	130 ± 22	40 ± 6	33 ± 6	148 ± 24	125 ± 21	52 ± 9	44 ± 7
Data	5397	11875	578	1046	8044	29316	727	2378

Table 9.2: The Higgs boson signal, background and data yields for each signal region category in the 0- and 1-lepton channels after the full selection. The signal and background yields are normalized to the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.

Signal regions	2-lepton		2-lepton		2-lepton	
	$75\,\text{GeV} < p_{\text{T}}^{\text{V}} < 150\,\text{GeV}$		$150\,\text{GeV} < p_{\text{T}}^{\text{V}} < 250\,\text{GeV}$		$p_{\text{T}}^{\text{V}} > 250\,\text{GeV}$	
Sample	2-jet	≥3-jet	2-jet	≥3-jet	2-jet	≥3-jet
Z + jets	5900 ± 100	11630 ± 160	716 ± 19	2499 ± 52	84 ± 3	537 ± 16
W + jets	1 ± 0	6 ± 0	< 1	2 ± 0	< 1	< 1
Top	3193 ± 57	8796 ± 87	52 ± 7	389 ± 19	1 ± 1	15 ± 4
Diboson	283 ± 47	443 ± 78	83 ± 14	169 ± 30	20 ± 4	52 ± 10
Total bkg.	9378 ± 86	20880 ± 130	851 ± 19	3058 ± 44	106 ± 4	605 ± 14
Signal ($\mu = 1.02$)	78 ± 14	106 ± 21	34 ± 6	59 ± 12	10 ± 2	18 ± 3
Data	9463	20927	881	3148	123	614

Table 9.3: The Higgs boson signal, background and data yields for each signal region category in the 2-lepton channel after the full selection. The signal and background yields are normalized to the results of the global likelihood fit. The top background is derived from $e\mu$-CR data. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.

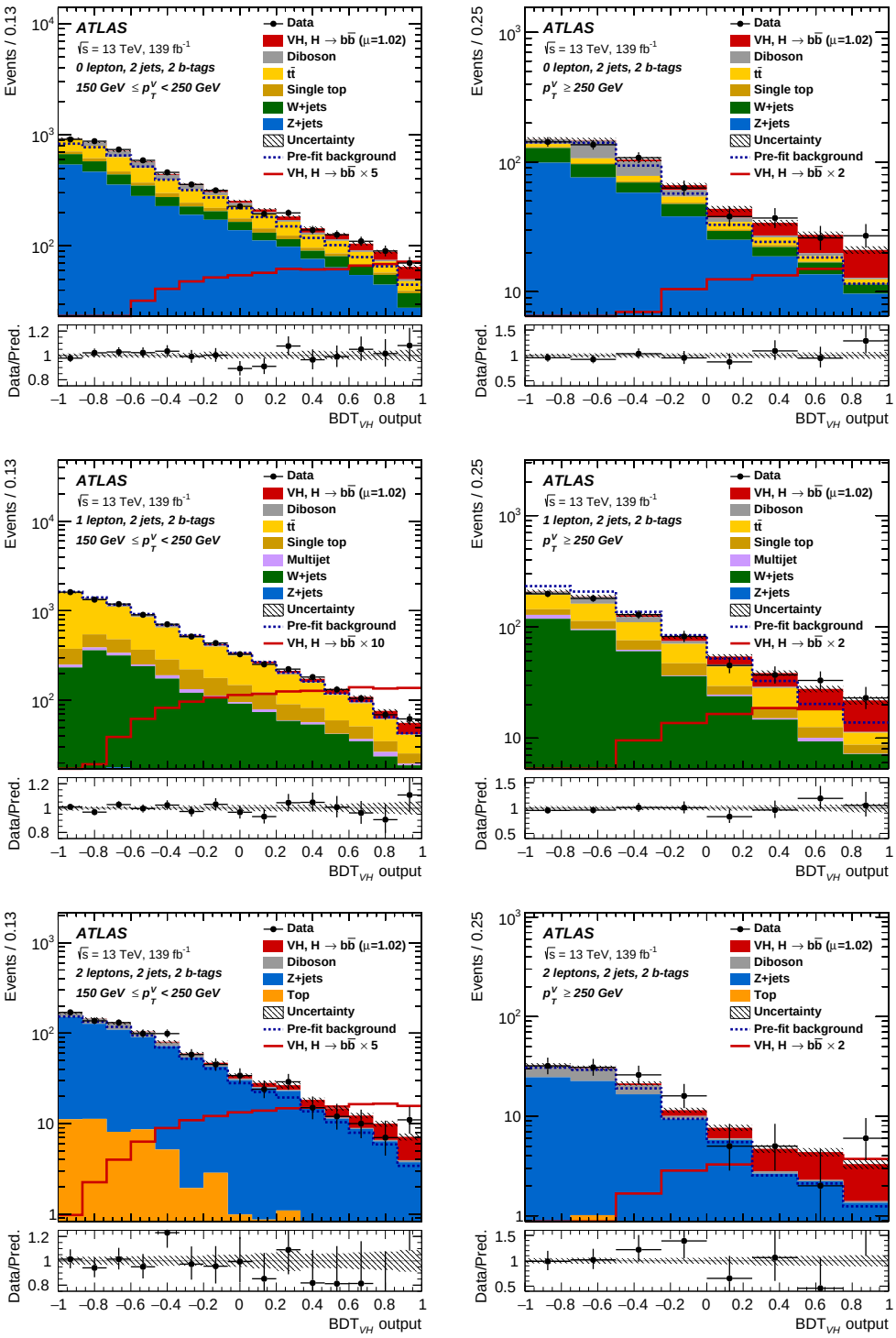


Figure 9.1: The BDT_{VH} output post-fit distributions in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channels for 2-b-tag 2-jet events, for the $150 < p_{\mathrm{T}}^{\mathrm{V}} < 250$ GeV (left) and $p_{\mathrm{T}}^{\mathrm{V}} > 250$ GeV (right) $p_{\mathrm{T}}^{\mathrm{V}}$ regions. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_h = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.02$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal ($\mu = 1.02$) and background is shown in the lower panel. The BDT_{VH} output distributions are shown with the binning used in the global likelihood fit. Figures are from the publication of the $VHbb$ analysis of ATLAS [159].

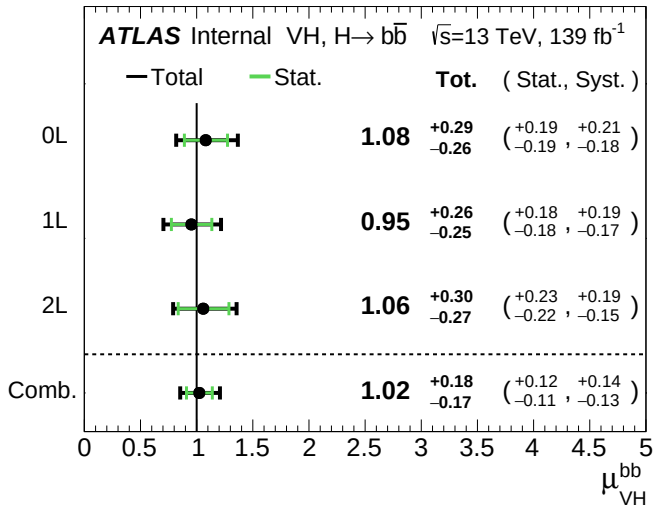


Figure 9.2: Best values of the signal strength $\mu_{VH}^{b\bar{b}}$ for the 0-, 1- and 2- lepton channels and their combination in the unconditional fit to the Run 2 data of the three channels combined. The (black) total observed uncertainty is quoted together with its decomposition in the (green) statistical component, and systematic component. In this plot the uncertainty due to background scale factors is included in the statistical component [109].

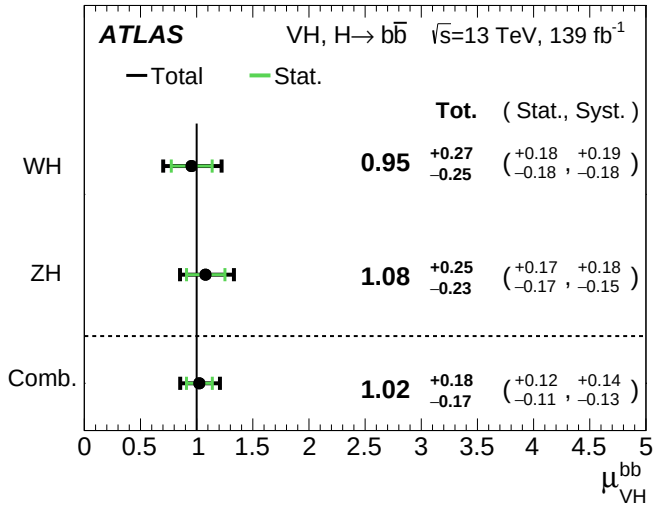


Figure 9.3: The fitted values of the Higgs boson signal strength μ_{VH}^{bb} for $m_h = 125$ GeV for the WH and ZH processes and their combination. The individual μ_{VH}^{bb} values for the $(W/Z)H$ processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently. The probability of compatibility of the individual signal strengths is 71%. Figures are from the publication of the $VHbb$ analysis of ATLAS [159].

Source of uncertainty		σ_μ		
		VH	WH	ZH
Total		0.177	0.260	0.240
Statistical		0.115	0.182	0.171
Systematic		0.134	0.186	0.168
Statistical uncertainties				
Data statistical		0.108	0.171	0.157
$t\bar{t}$ $e\mu$ control region		0.014	0.003	0.026
Floating normalizations		0.034	0.061	0.045
Experimental uncertainties				
Jets		0.043	0.050	0.057
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.015	0.045	0.013
Leptons		0.004	0.015	0.005
b-tagging	b-jets	0.045	0.025	0.064
	c-jets	0.035	0.068	0.010
	light-flavor jets	0.009	0.004	0.014
Pile-up		0.003	0.002	0.007
Luminosity		0.016	0.016	0.016
Theoretical and modeling uncertainties				
Signal		0.072	0.060	0.107
Z + jets		0.032	0.013	0.059
W + jets		0.040	0.079	0.009
$t\bar{t}$		0.021	0.046	0.029
Single top quark		0.019	0.048	0.015
Diboson		0.033	0.033	0.039
Multi-jet		0.005	0.017	0.005
MC statistical		0.031	0.055	0.038

Table 9.4: Breakdown of the contributions to the uncertainty in μ_{VH}^{bb} for the VH, WH and ZH signal strength measurements. The sum in quadrature of the systematic uncertainties attached to the categories differs from the total systematic uncertainty due to correlations. Table is from the publication of the $VHbb$ analysis of ATLAS [159].

9.1.2. Dijet-mass Cross-check

The fitted value of the signal strength in the dijet-mass cross-check after combining all channels is,

$$\mu_{VH}^{bb} = 1.17^{+0.25}_{-0.23} = 1.17 \pm 0.16(\text{stat.})^{+0.19}_{-0.16}(\text{syst.}),$$

which rejects the background-only hypothesis with a significance of 5.5 standard deviations compared to an expected value of 4.9 standard deviations. It is in good agreement[2] with the result of the multivariate analysis in Sec. 9.1.1. The m_{bb} distribution is shown in Fig. 9.4, which is summed over all regions and channels, and weighted by their respective values of the ratio of the yield of fitted Higgs boson signal to background, after subtracting all backgrounds except for the yields of WZ and ZZ processes.

[2]Good agreement is also found between the signal strengths values in the individual channels from the dijet-mass analysis and those from the multivariate analysis.

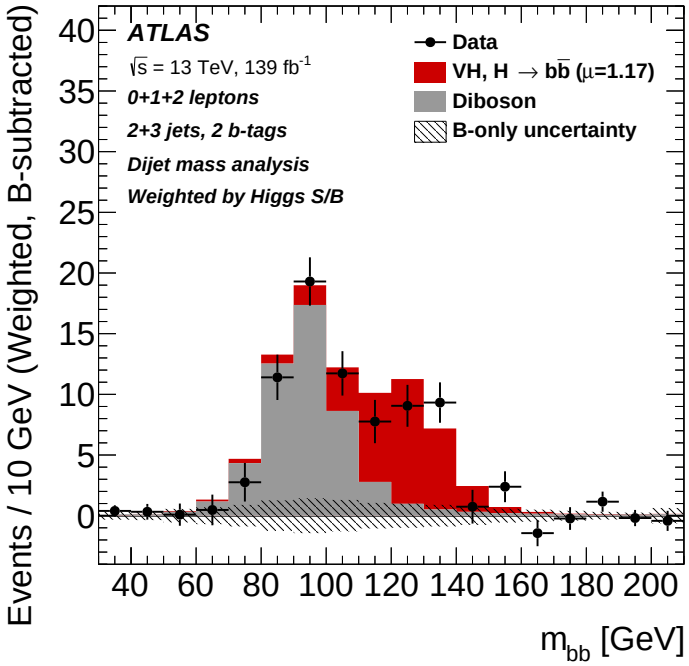


Figure 9.4: The distribution of m_{bb} in data after subtracting all backgrounds except the WZ and ZZ diboson processes, as obtained with the dijet-mass analysis. The contributions from all lepton channels, $p_{\mathrm{T}}^{\mathrm{V}}$ regions and number-of-jets categories are summed and weighted by their respective S/B ratios, with S being the total fitted signal and B the total fitted background in each region. The expected contribution of the associated WH and ZH production of a SM Higgs boson with $m_h = 125$ GeV is shown scaled by the measured signal strength ($\mu = 1.17$). The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band. Figure is from the publication of the $VHbb$ analysis of ATLAS [159].

9.1.3. Diboson Cross-check

As described in Sec. 8.3.4, the measurement of VZ (WZ and ZZ) production uses a multivariate approach. It serves as a validation of the VH analysis. After fit to data, the measured signal strength is

$$\mu_{VZ}^{bb} = 0.93^{+0.15}_{-0.14} = 0.93^{+0.07}_{-0.06}(\text{stat.})^{+0.14}_{-0.12}(\text{syst.}),$$

which well agrees with the standard model prediction. A fit to measure the signal strengths of the WZ and ZZ production modes separately is also performed. The results are shown in Fig. 9.5.

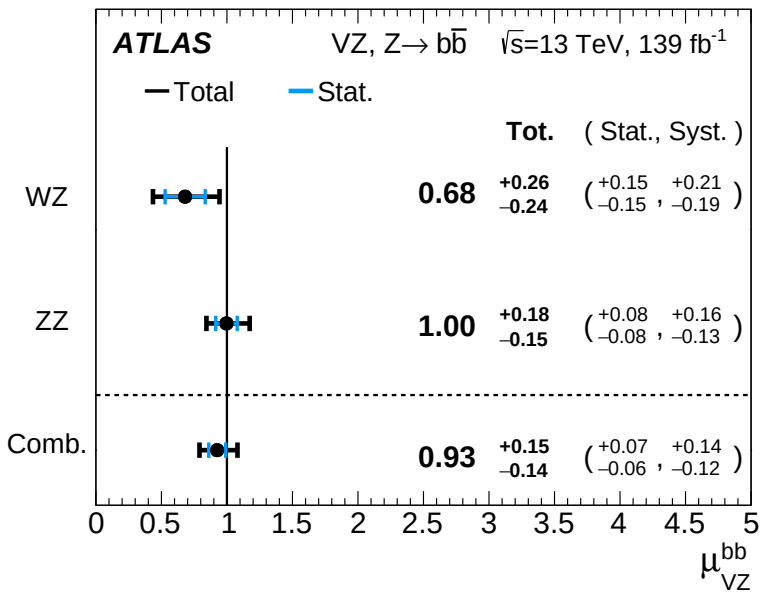


Figure 9.5: The fitted values of the VZ signal strength μ_{VZ}^{bb} for the WZ and ZZ processes and their combination. The individual μ_{VZ}^{bb} values for the WZ and ZZ processes are obtained from a simultaneous fit with the signal strengths for each of the WZ and ZZ processes floating independently. The probability of compatibility of the individual signal strengths is 27%. Figure is from the publication of the $VHbb$ analysis of ATLAS [159]

9.2. Cross-section Measurements

The measured VH cross-section times the $V \to$ and $H \to b\bar{b}$ leptons branching ratios, $\sigma \times B$, are summarized in Table 9.5 and Fig. 9.6 together with the theoretical predictions in the reduced STXS regions. The cross-section measurements are consistent with the Standard Model expectations and their relative uncertainties vary from 30% in the highest p_{T}^{V} regions to 85% in the lowest p_{T}^{V} regions. In all regions, there are large contributions from the background modeling, b-tagging and jet systematic uncertainties. The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty is one of the largest uncertainties in the lowest p_{T}^{V} region of both the WH and ZH measurements. In the ZH measurements, the signal uncertainties also make a sizable contribution due to the limited precision in the theoretical calculation of the $gg \to ZH$ process.

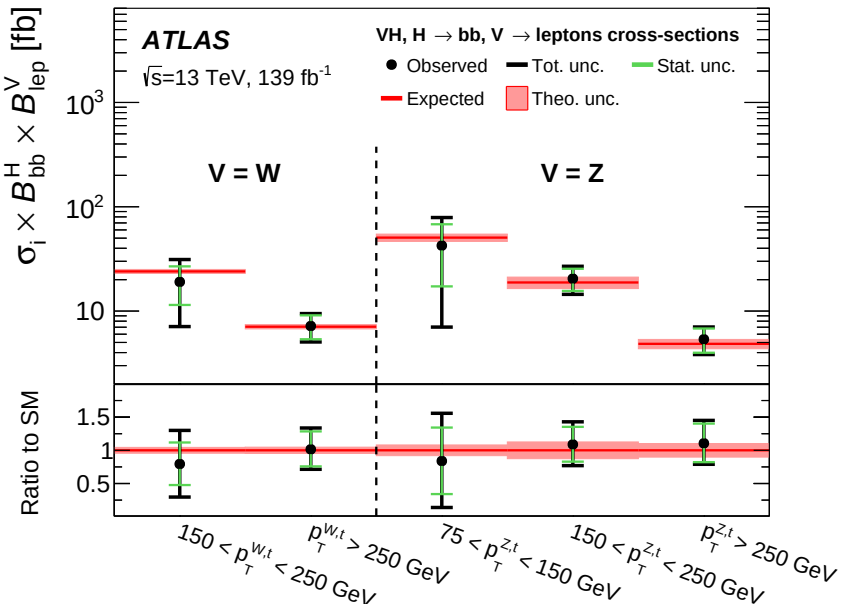


Figure 9.6: Measured VH, $V \to$ leptons cross-sections times the $H \to b\bar{b}$ branching fraction in the reduced STXS scheme. Figure is from the publication of the $VHbb$ analysis of ATLAS [159].

STXS region		SM prediction	Result	Stat. unc.	Syst. unc. [fb]		
Process	$p_{\mathrm{T}}^{V,\,\mathrm{t}}$ interval	[fb]	[fb]	[fb]	Th. sig.	Th. bkg.	Exp.
$W(\ell\nu)H$	150–250 GeV	24.0 ± 1.1	19.0 ± 12.1	± 7.7	± 0.9	± 5.5	± 6.0
$W(\ell\nu)H$	> 250 GeV	7.1 ± 0.3	7.2 ± 2.2	± 1.9	± 0.4	± 0.8	± 0.7
$Z(\ell\ell/\nu\nu)H$	75–150 GeV	50.6 ± 4.1	42.5 ± 35.9	± 25.3	± 5.6	± 17.2	± 19.7
$Z(\ell\ell/\nu\nu)H$	150–250 GeV	18.8 ± 2.4	20.5 ± 6.2	± 5.0	± 2.3	± 2.4	± 2.3
$Z(\ell\ell/\nu\nu)H$	> 250 GeV	4.9 ± 0.5	5.4 ± 1.7	± 1.5	± 0.5	± 0.5	± 0.3

Table 9.5: Best-fit values and uncertainties for the VH, $V \to$ leptons cross-section times the $H \to b\bar{b}$ branching fraction, in the reduced STXS scheme. The SM predictions for each region, computed using the inclusive cross-section calculations and the simulated event samples are also shown. The total systematic uncertainty, equal to the difference in quadrature between the total uncertainty and the statistical uncertainty, differs from the sum in quadrature of the Th. sig., Th. bkg., and Exp. systematic uncertainties due to correlations. All leptonic decays of the V bosons (including those to τ-leptons, $\ell = e, \mu, \tau$) are considered. Table is from the publication of the $VHbb$ analysis of ATLAS [159].

CHAPTER 10

Conclusion

Ever since the Higgs boson was discovered in 2012, the focus of the ATLAS experiment has been shifted to the measurement of its properties. The analysis presented in this thesis has measured the Higgs boson decaying to a pair of b-quarks produced in association with a vector boson, using the p-p collision data of 139 fb^{-1}at a center-of-mass of 13 TeV for the full Run-2 of LHC.

The measured signal strength of the VH signal for a Higgs mass of 125 GeV is $\mu_{VH}^{bb} = 1.02^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.})$, which rejects the background-only hypothesis with a significance of 6.7 standard deviations compared to an expectation of 6.7 standard deviations. It leads to an observation of the VH production in the $H \to b\bar{b}$ channel at the ATLAS experiment[1]. It is also the first time any single experiment has observed $H \to b\bar{b}$ from a single production mechanism without support from additional production mechanisms.

The cross-sections of VH with $V \to$ leptons and $H \to b\bar{b}$ are measured in the simplified STXS scheme. They are consistent with the Standard Model calculations and their uncertainties vary from 30% in the highest p_{T}^{V} regions to 85% in the lowest p_{T}^{V} regions.

The LHC is being upgraded and plans to deliver an integrated luminosity of 300 fb^{-1}at the end of Run-3. Although it is hard to predict what the $VHbb$ results will be by then, the precision of measurements due to the effect of doubled data compared to Run-2 can be estimated. Table 10.1 shows the estimated precision of the signal strength measurements

[1]Since the measured signal strength of ZH signal is $\mu_{ZH}^{bb} = 1.08 \pm 0.17(\text{stat.})^{+0.18}_{-0.15}(\text{syst.})$ which rejects the background-only hypothesis with a significance of 5.3 standard deviations, this analysis also leads to an observation of the ZH production in the $H \to b\bar{b}$ channel.

for Run-3. The statistical uncertainties in those tables are evaluated with the statistical uncertainties in Table 9.4 divided by $\sqrt{2}$. The systematic uncertainties are the sum in quadrature of the experimental uncertainties and the theoretical and modeling uncertainties[2]. Since the study of the experimental uncertainties often relies on the statistics of the selected events in data[3], the experimental uncertainties in Run-3 are simply estimated from those values in Run-2 after divided by $\sqrt{2}$. Then the total uncertainty is the sum in quadrature of the statistical uncertainties and systematic uncertainties.

The estimated total uncertainties in Table 10.1 are slightly smaller than those in Run-2, which means that an increase in statistics alone won't simply improve the precision of measurements in this analysis. The signal and background modeling, as well as relevant theoretical models, need to be optimized, at the same time. In addition, room exists for improvement on the object reconstruction, event selection, and flavor tagging. I have also demonstrated that a veto on hadronic taus can improve background rejection and expect this to be a standard part of the analysis in Run-3.

In Run 2, for the first time, ATLAS and CMS have established the direct observation of the Higgs decaying to quarks. While this process so far fits the pattern first predicted in the Standard Model in the 1960s and 1970s, the uncertainty on this measurement is large. In addition, it is essential to also observe the Higgs decaying to second-generation fermions (e.g. muons or charm quarks) to establish the pattern of such decays predicted by the Standard Model. The work described in this thesis will serve all of these developments as the basis for future discoveries while establishing this observation for the textbooks in particle physics.

[2]The assumption here is that the experimental uncertainties are independent of the theoretical and modeling uncertainties, while in fact, that's not 100% true.

[3]For example, the muon momentum scale and resolution are derived using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decay events. The b-tagging efficiency is calibrated using $t\bar{t}$ events which is enriched in b-jets.

Source of uncertainty in Run-3	σ_μ		
	VH	WH	ZH
Total estimated uncertainty	0.140	0.206	0.187
Statistical	0.081	0.129	0.121
Systematic	0.114	0.161	0.143
Experimental uncertainties	0.069	0.093	0.088
Theoretical and modeling uncertainties	0.091	0.132	0.113

Table 10.1: Estimation of the uncertainties in the signal strength measurement of the VH, WH and ZH production for Run-3. In this table, the statistical uncertainties and experimental uncertainties are divided by $\sqrt{2}$ due to the increase of the integrated luminosity from Run-2 to Run-3. The theoretical and modeling uncertainties are kept the same. The total uncertainties are the sum in quadrature of the statistical uncertainties and systematic uncertainties. The table to be compared with Run-2 is Table 9.4.

APPENDIX A

$t\bar{t}$ Rejection in the 1-lepton Channel

Additional $t\bar{t}$ rejection plots are presented here.

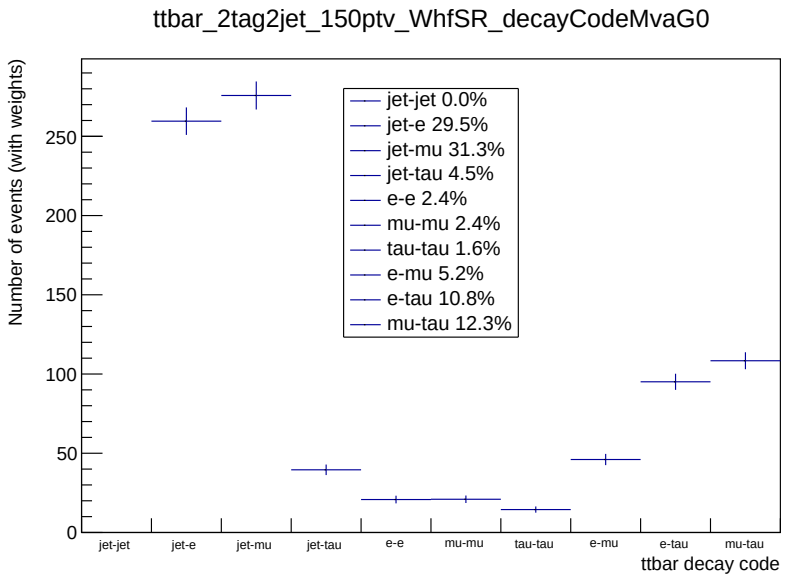


Figure 1.1: After the 2t2j selection of 150ptv region in 1-lep and then **cut on BDT score greater than 0 (∼90% signal efficiency)**, each ttbar event is asked what its decay mode is (in truth-level) and then filled here. dileptonic (e-e, μ-μ, e-μ): 10.0%; semileptonic (jet-e, jet-μ): 60.8%; hadronictau (e-τ, μ-τ): 23.1%; leptonictau (jet-τ): 4.5%; τ-τ: 1.6%.

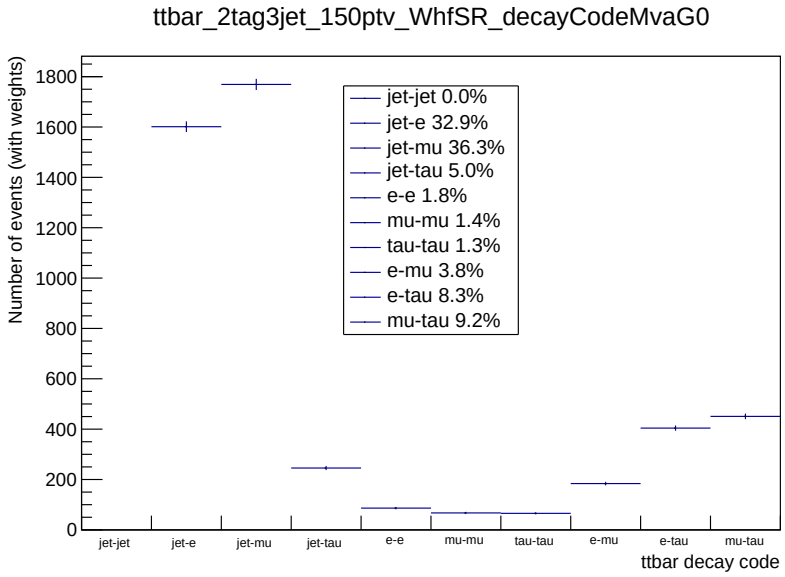


Figure 1.2: After the 2t3j selection of 150ptv region in 1-lep and then **cut on BDT score greater than 0 (∼90% signal efficiency)**, each ttbar event is asked what its decay mode is (in truth-level) and then filled here. dileptonic (e-e, μ-μ, e-μ): 6.9%; semileptonic (jet-e, jet-μ): 69.2%; hadronictau (e-τ, μ-τ): 17.5%; leptonictau (jet-τ): 5.0%; τ-τ: 1.3%.

APPENDIX B

Tau-veto in the 1-lepton Channel

Study on the implementation of tau-veto in the 1-lepton channel is presented here.

B.1. Cut-flow

<finalizeTraining>	Cut flow counts for Background:			
Cut	sumOfWeights	passed/total (sumOfWeights)	nEvents	passed/total (nEvents)
all	3.19E+06	1	2.30E+08	1
nJ>1.500000	3.19E+06	1	229728477	1
nJ<2.500000	152901	0.0478738	57742112	0.251349
pTV>150.000000	25036.2	0.00783894	19375844	0.0843424
Mtop<225 || mBB>75	23811.7	0.00745555	18221363	0.079317
nTaus<1	20683.7	0.00647615	17787338	0.0774277

Table 2.1: Background cut-flow in 2J region

<finalizeTraining>	Cut flow counts for Signal:			
Cut	sumOfWeights	passed/total (sumOfWeights)	nEvents	passed/total (nEvents)
all	1419.63	1	6.43E+06	1
nJ>1.500000	1419.63	1	6428275	1
nJ<2.500000	504.464	0.355349	2125807	0.330696
pTV>150.000000	209.599	0.147643	1500498	0.233422
Mtop<225 || mBB>75	208.379	0.146784	1469344	0.228575
nTaus<1	206.853	0.145709	1442864	0.224456

Table 2.2: Signal cut-flow in 2J region

<finalizeTraining>	Cut flow counts for Background:			
Cut	sumOfWeights	passed/total (sumOfWeights)	nEvents	passed/total (nEvents)
all	3.19E+06	1	2.30E+08	1
nJ>2.500000	3.04E+06	0.952126	171986365	0.748651
nJ<3.500000	515677	0.161461	58128852	0.253033
pTV>150.000000	133245	0.0417197	21403180	0.0931673
Mtop<225 || mBB>75	130112	0.0407386	20385323	0.0887366
nTaus<1	122858	0.0384674	19797725	0.0861788

Table 2.3: Background cut-flow in 3J region

<finalizeTraining>	Cut flow counts for Signal:			
Cut	sumOfWeights	passed/total (sumOfWeights)	nEvents	passed/total (nEvents)
all	1419.63	1	6.43E+06	1
nJ>2.500000	915.167	0.644651	4302468	0.669304
nJ<3.500000	454.51	0.320161	2039766	0.317312
pTV>150.000000	205.993	0.145103	1478434	0.229989
Mtop<225 || mBB>75	202.157	0.142401	1436182	0.223416
nTaus<1	200.417	0.141175	1409979	0.21934

Table 2.4: Signal cut-flow in 3J region

B.2. Over-training Check

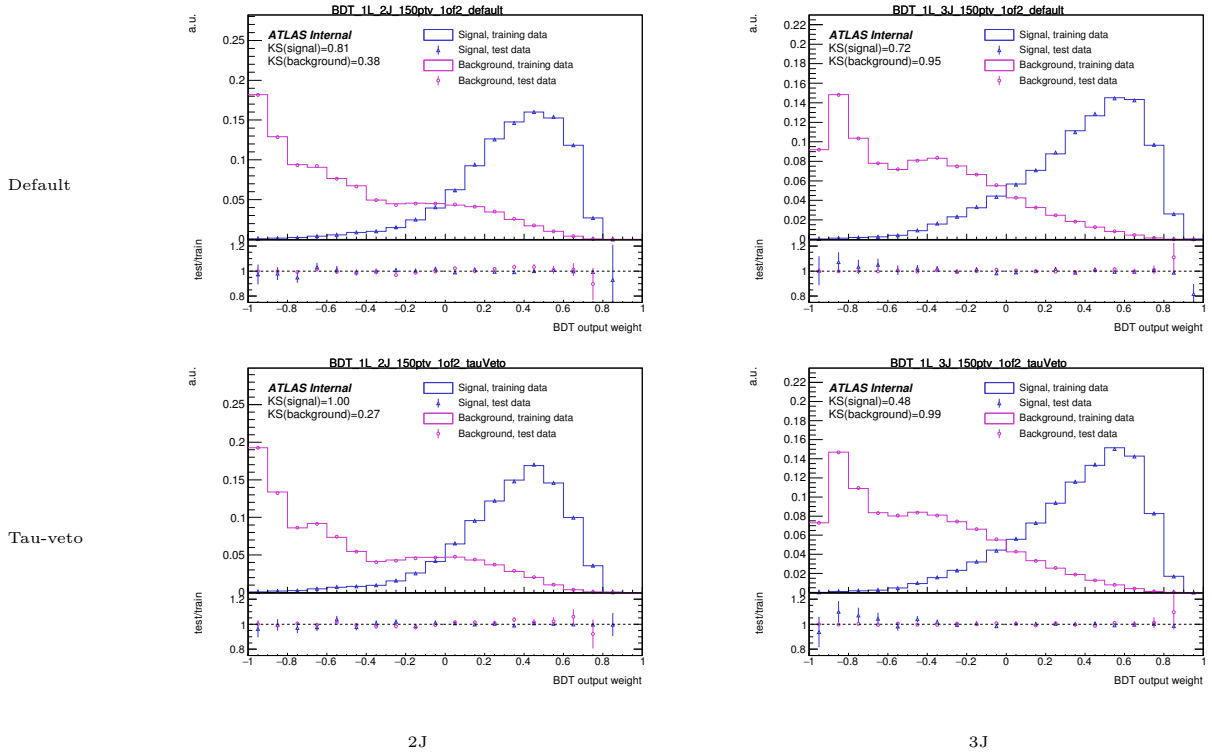


Figure 2.1: Over-training Check

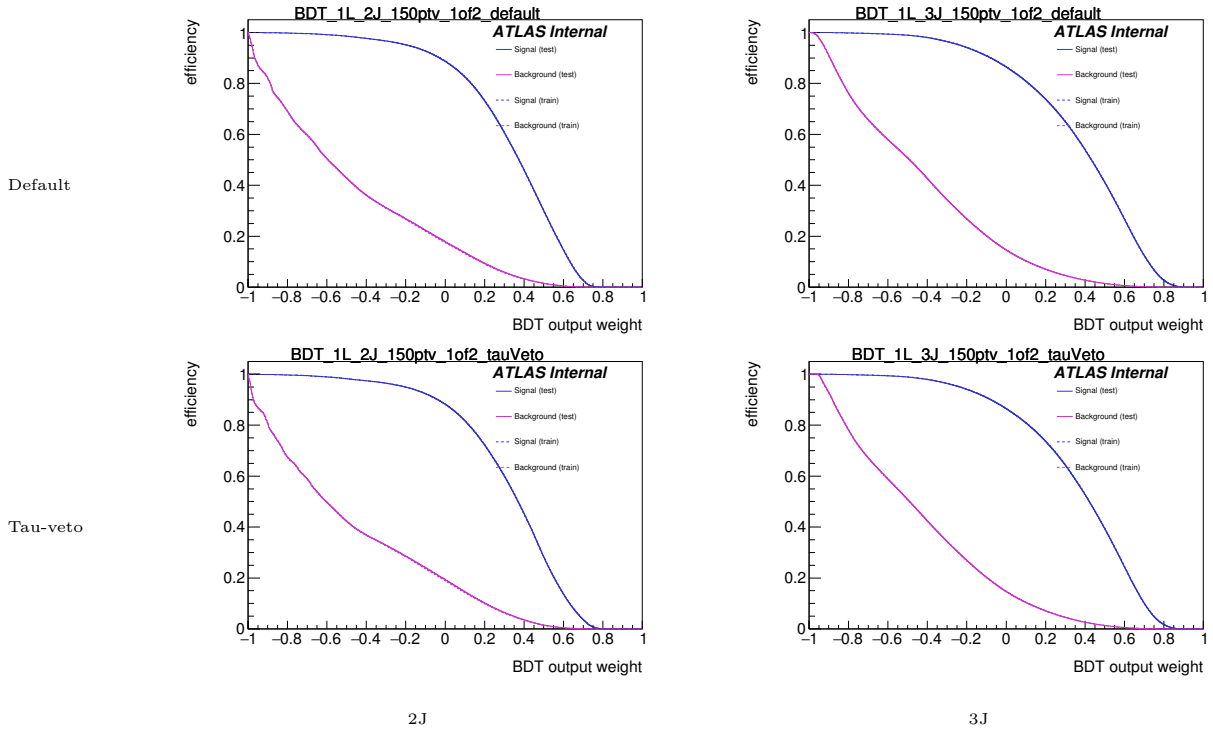
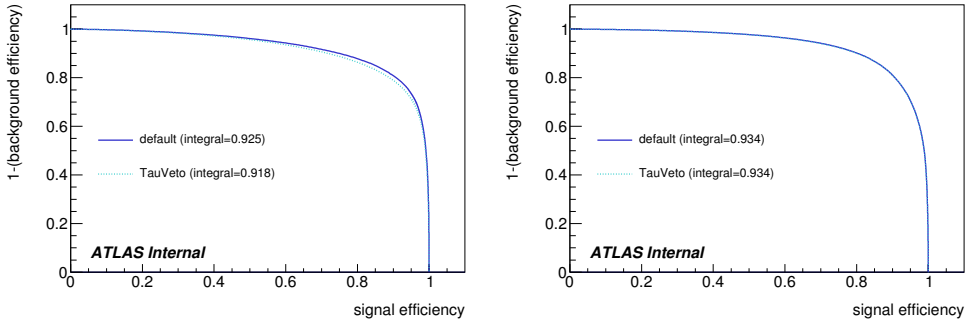


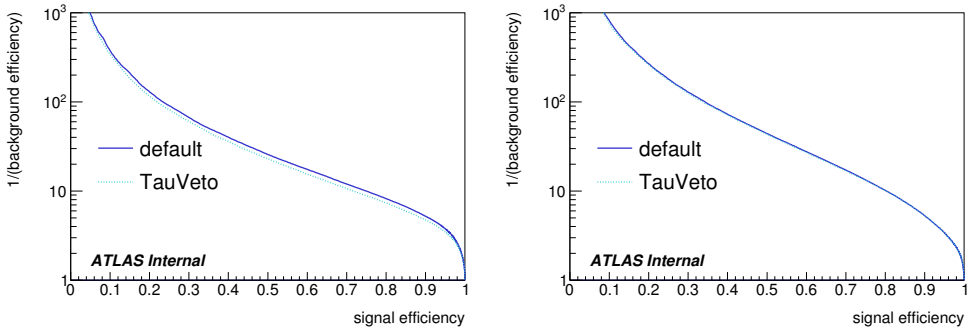
Figure 2.2: BDT Efficiency

B.3. ROC Curves



2J 3J

Figure 2.3: ROCComparison_1MinusBgEff_default_TauVeto



2J 3J

Figure 2.4: ROCComparison_RejFactor_default_TauVeto

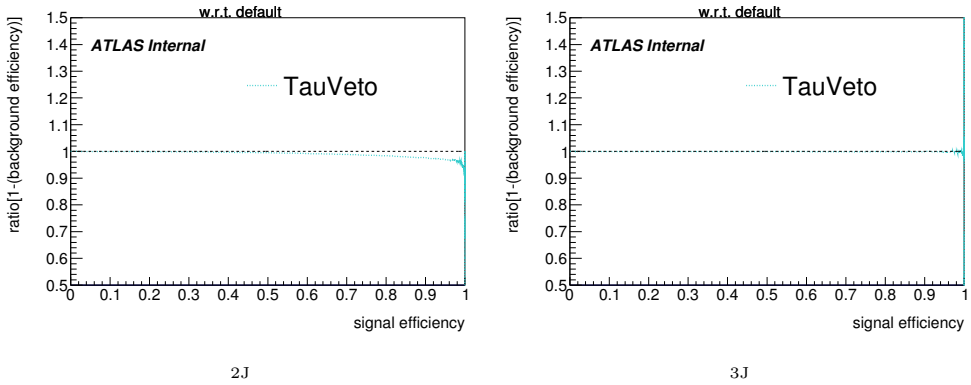


Figure 2.5: ROCRatio_default_TauVeto_1MinusBgEff

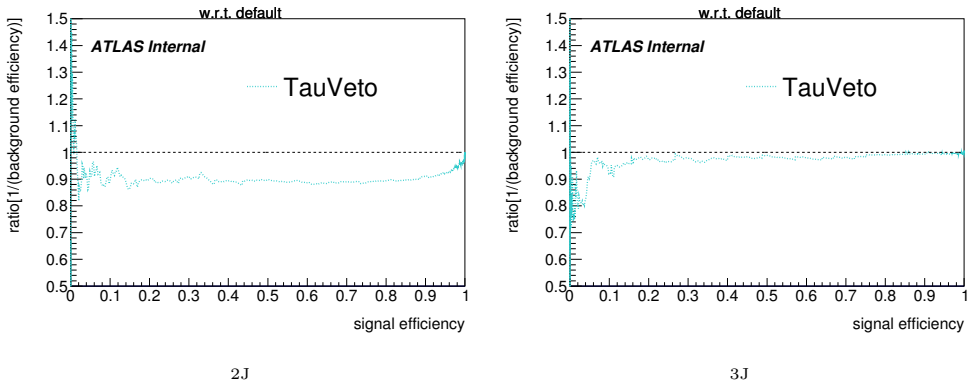


Figure 2.6: ROCRatio_default_TauVeto_RejFactor

B.4. MVA Input Variable Distribution Comparison

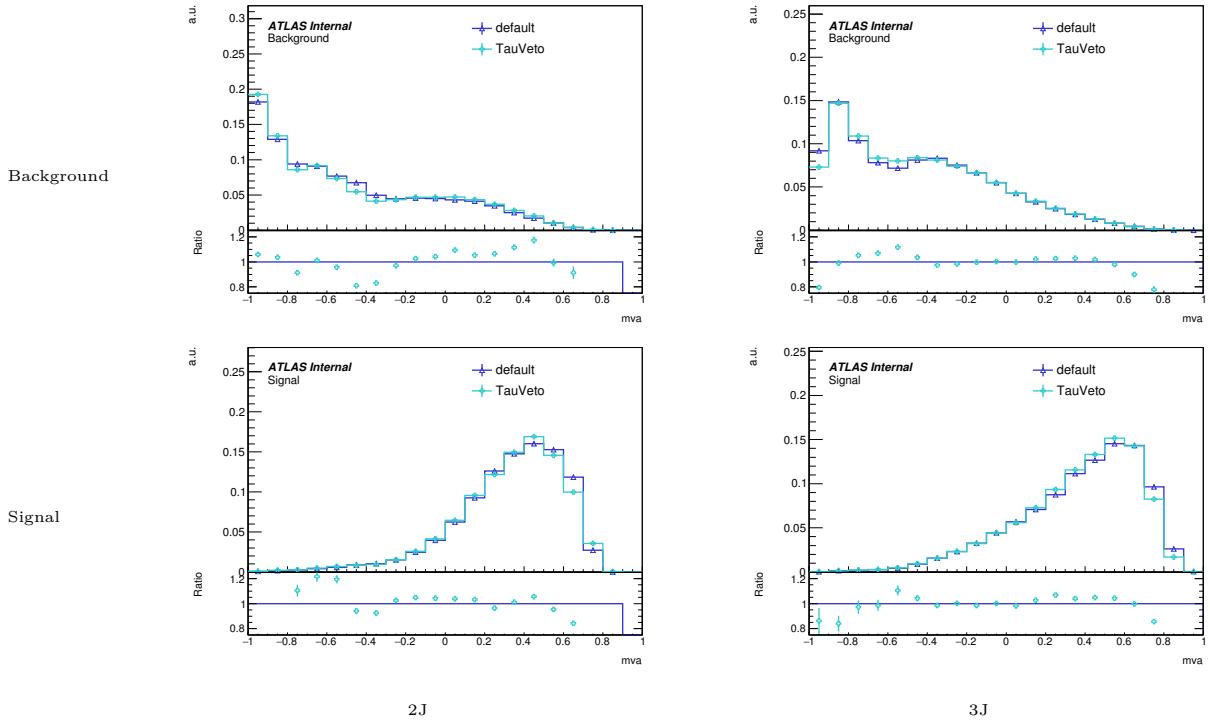


Figure 2.7: MVA Comparison

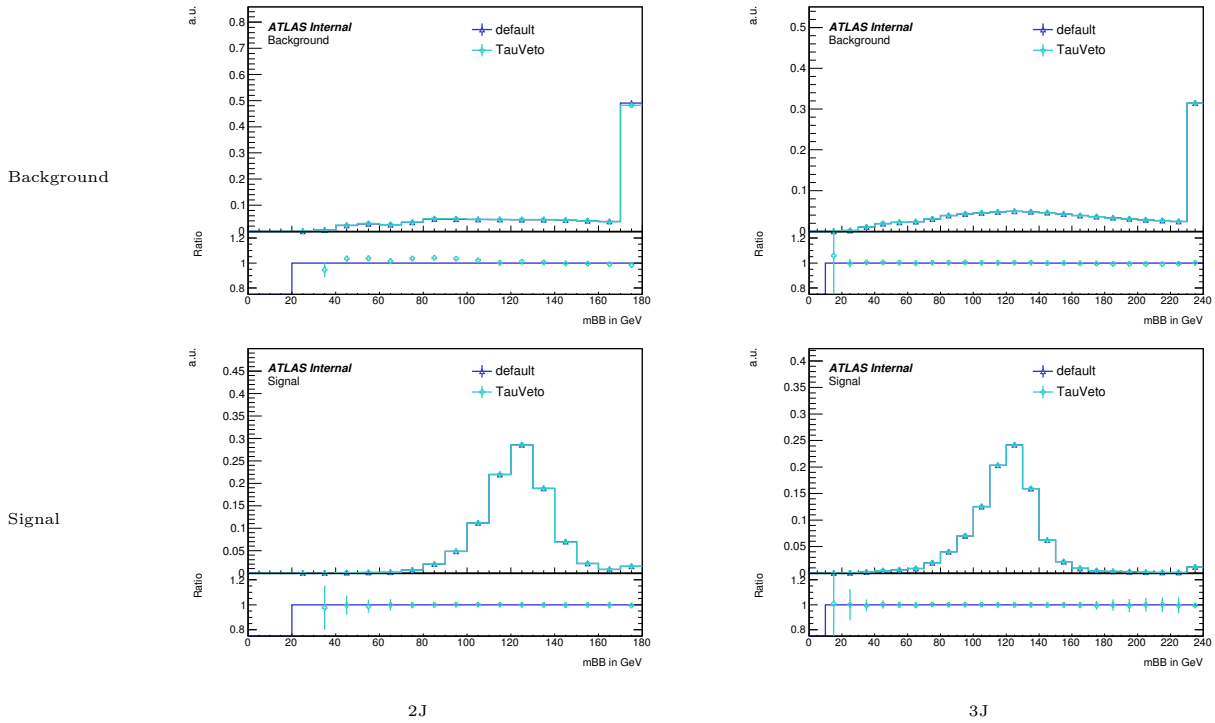


Figure 2.8: MVA Input Variable Comparison – mBB

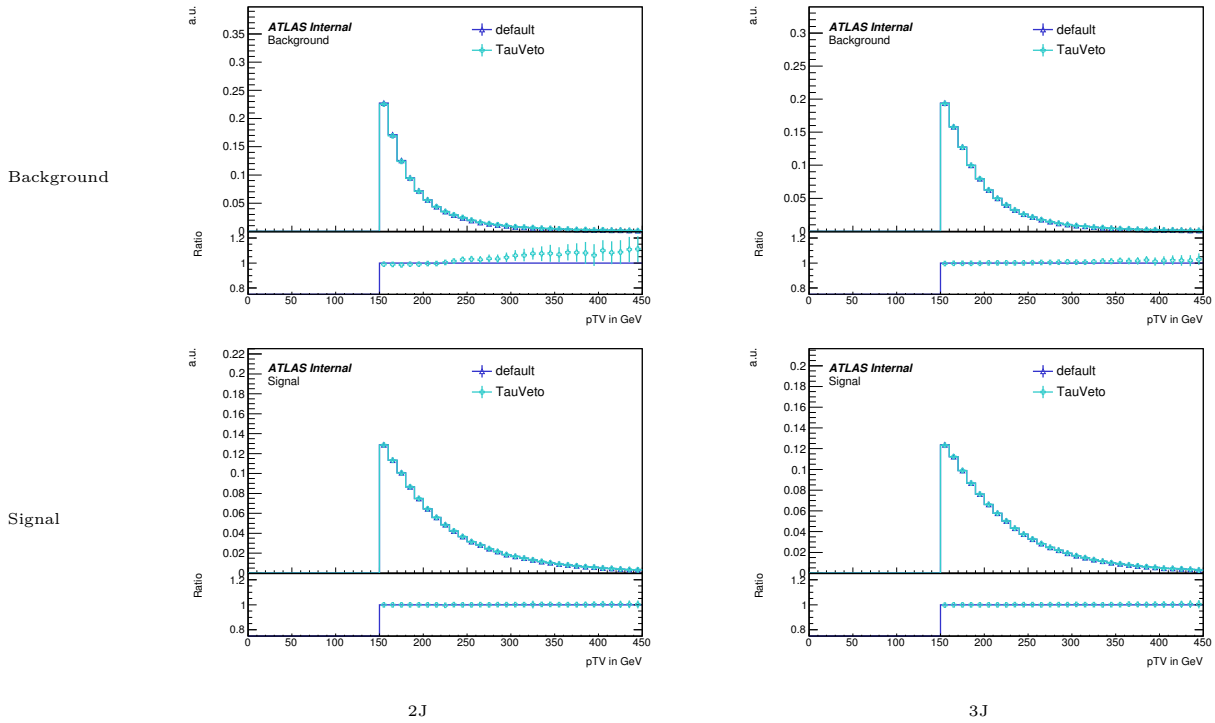


Figure 2.9: MVA Input Variable Comparison – pTV

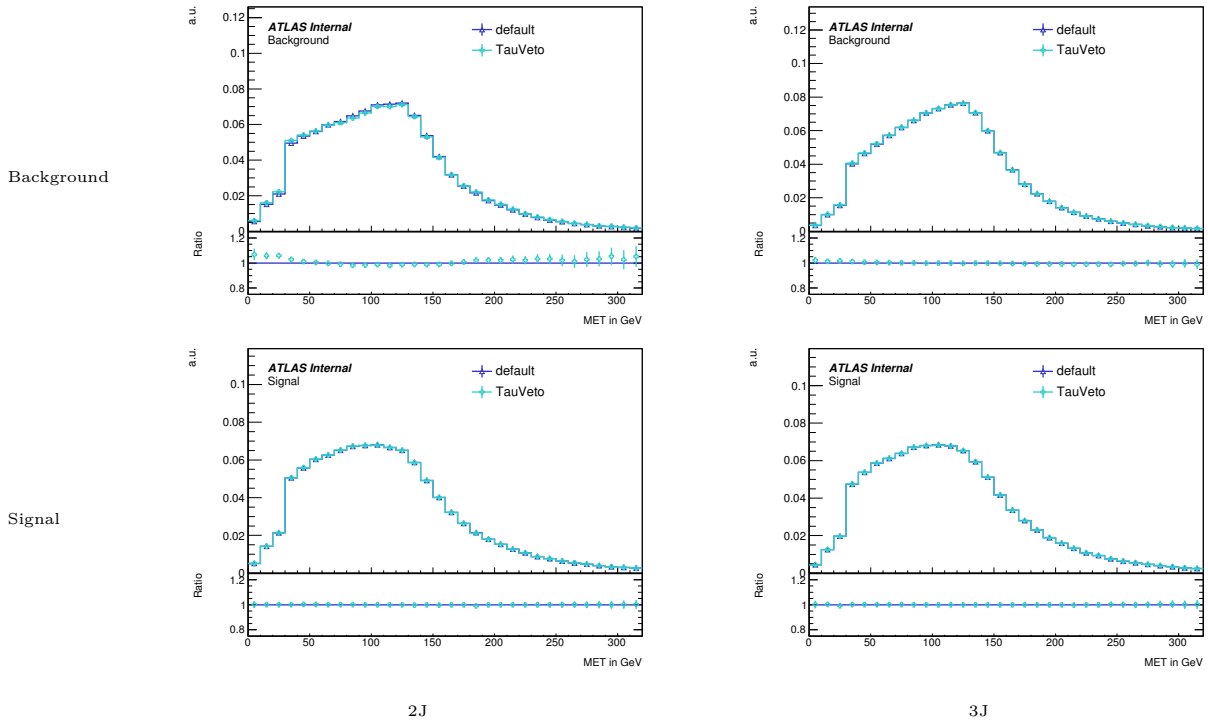


Figure 2.10: MVA Input Variable Comparison – MET

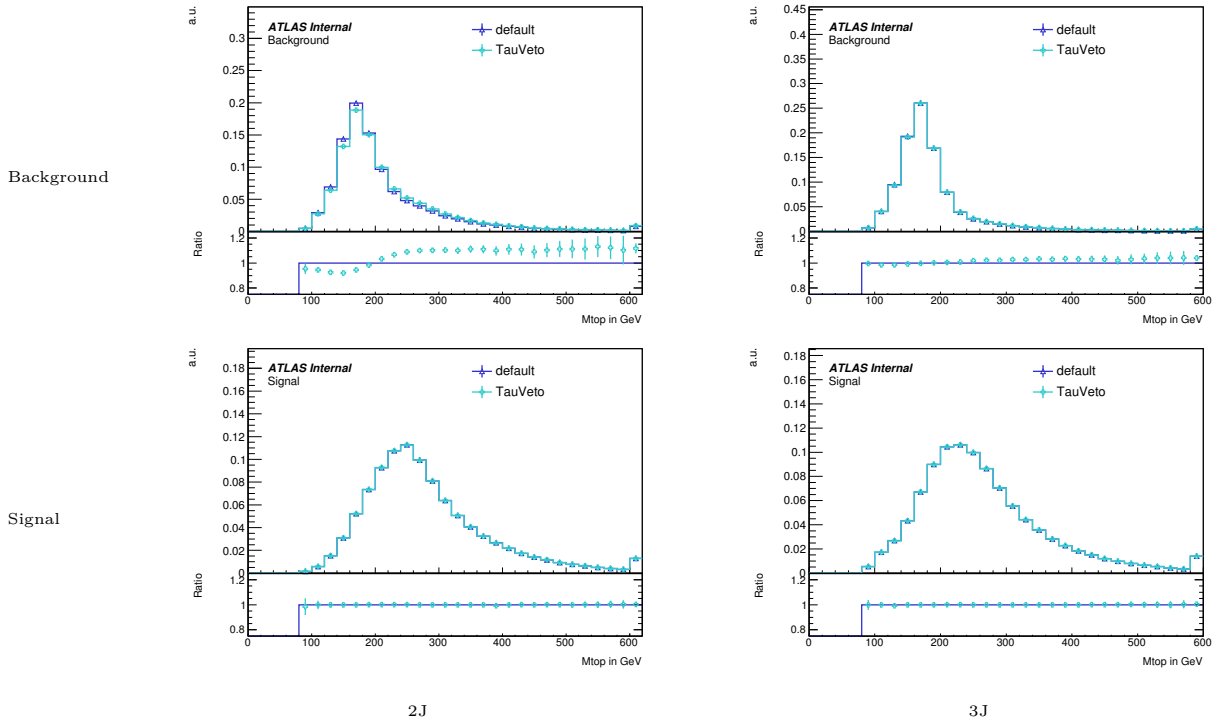


Figure 2.11: MVA Input Variable Comparison – Mtop

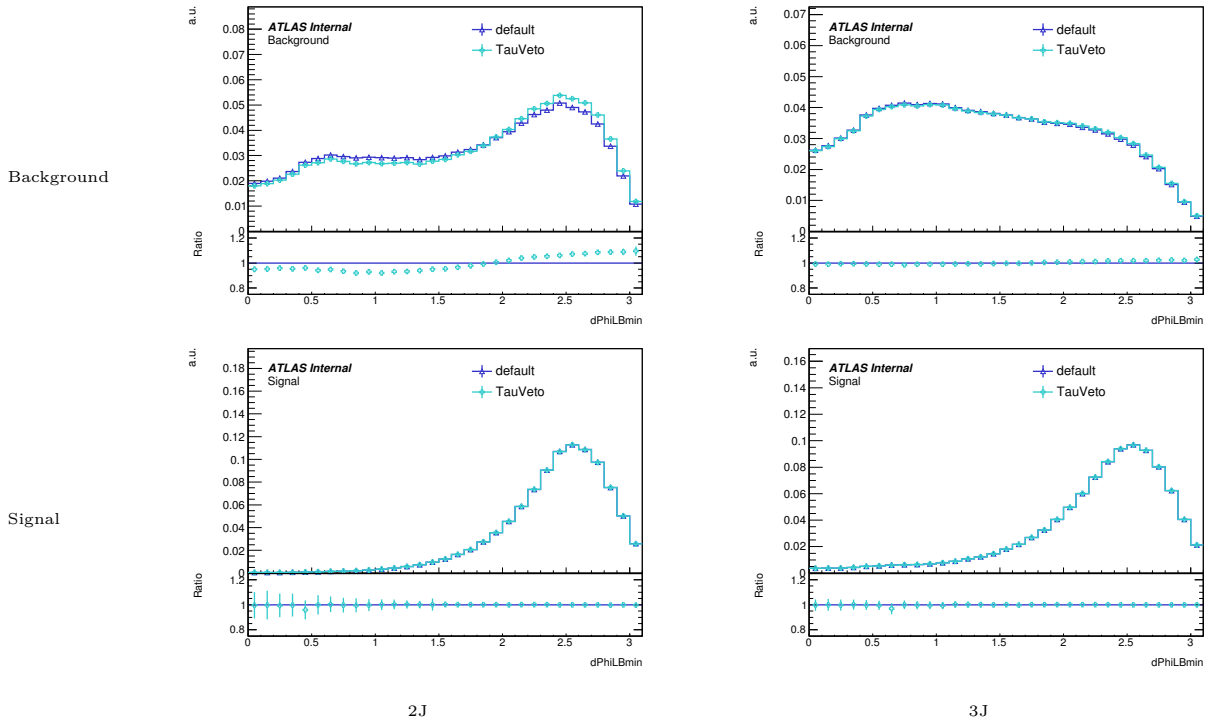


Figure 2.12: MVA Input Variable Comparison – dPhiLBmin

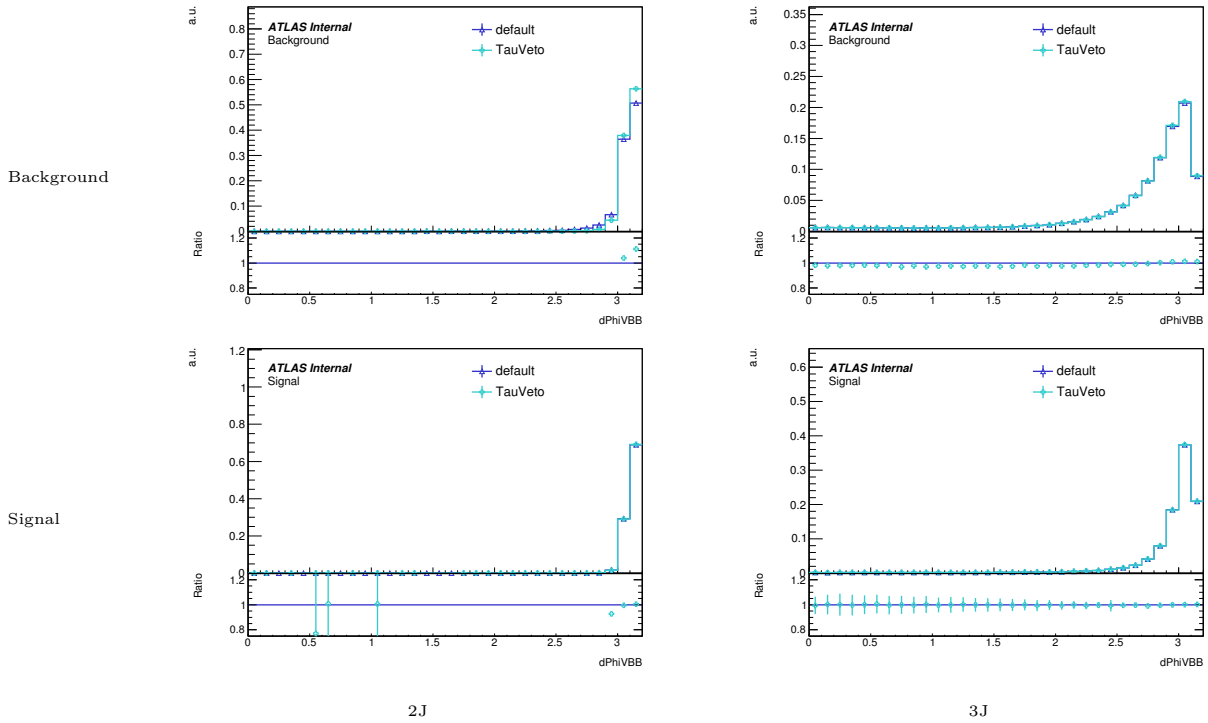


Figure 2.13: MVA Input Variable Comparison – dPhiVBB

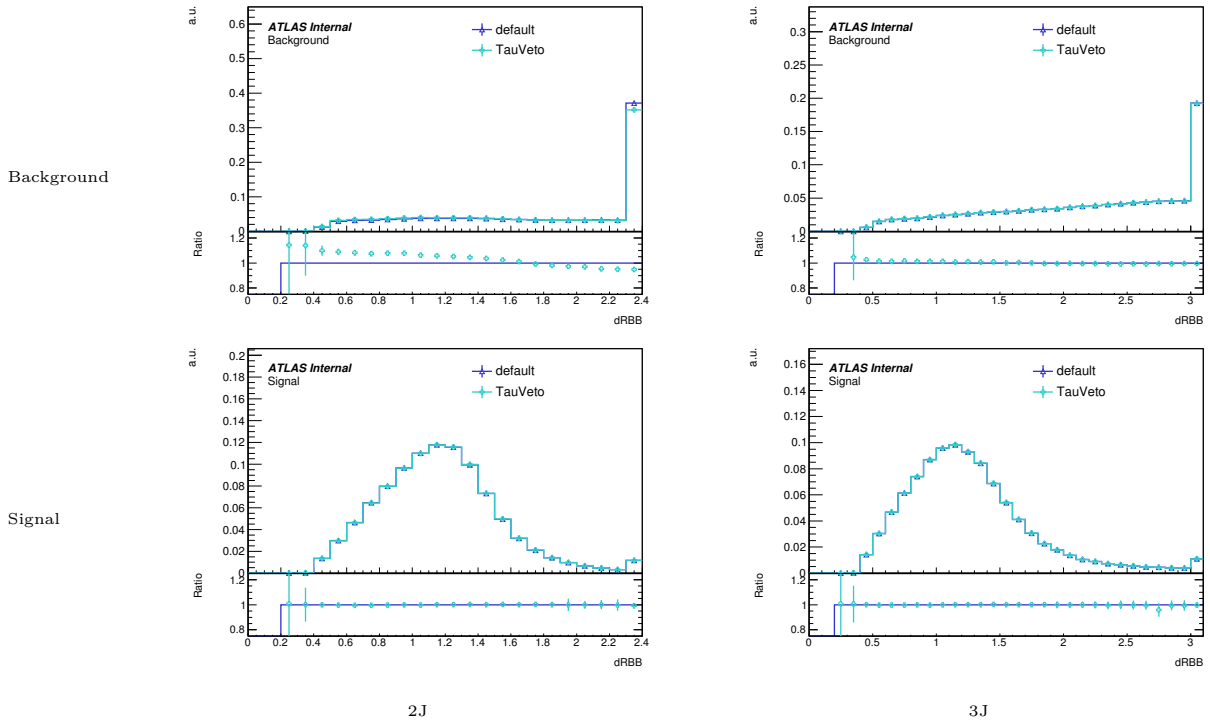


Figure 2.14: MVA Input Variable Comparison – dRBB

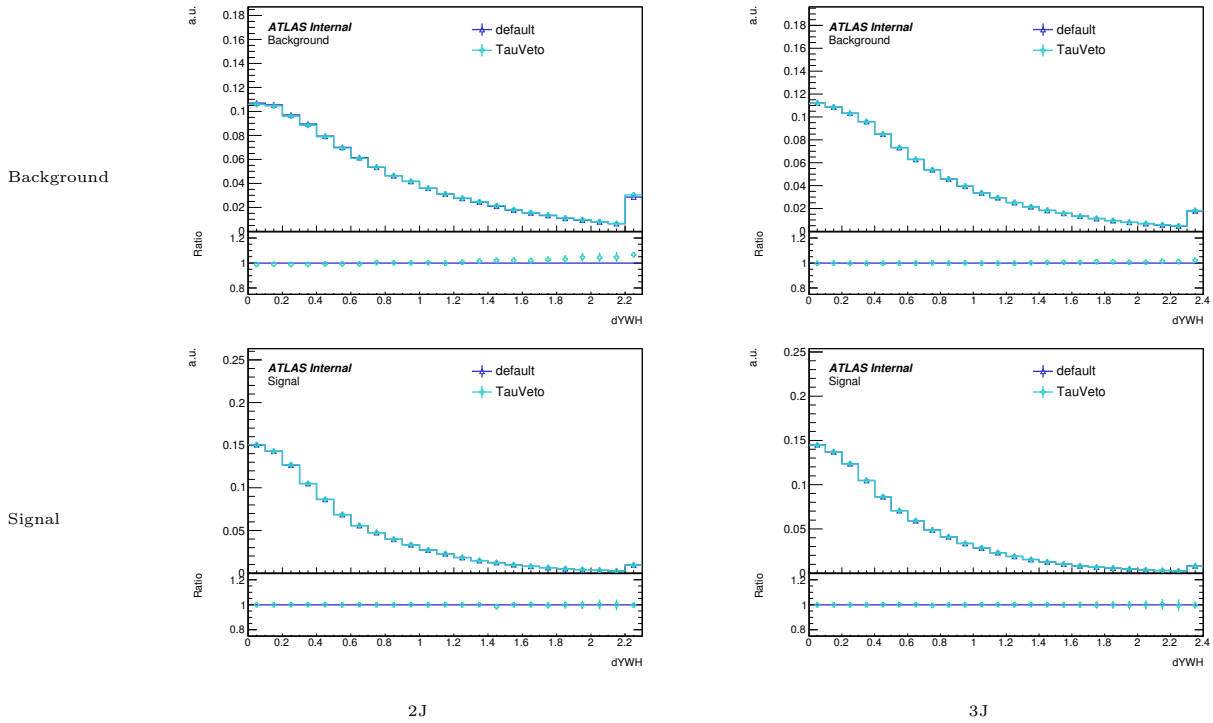


Figure 2.15: MVA Input Variable Comparison – dYWH

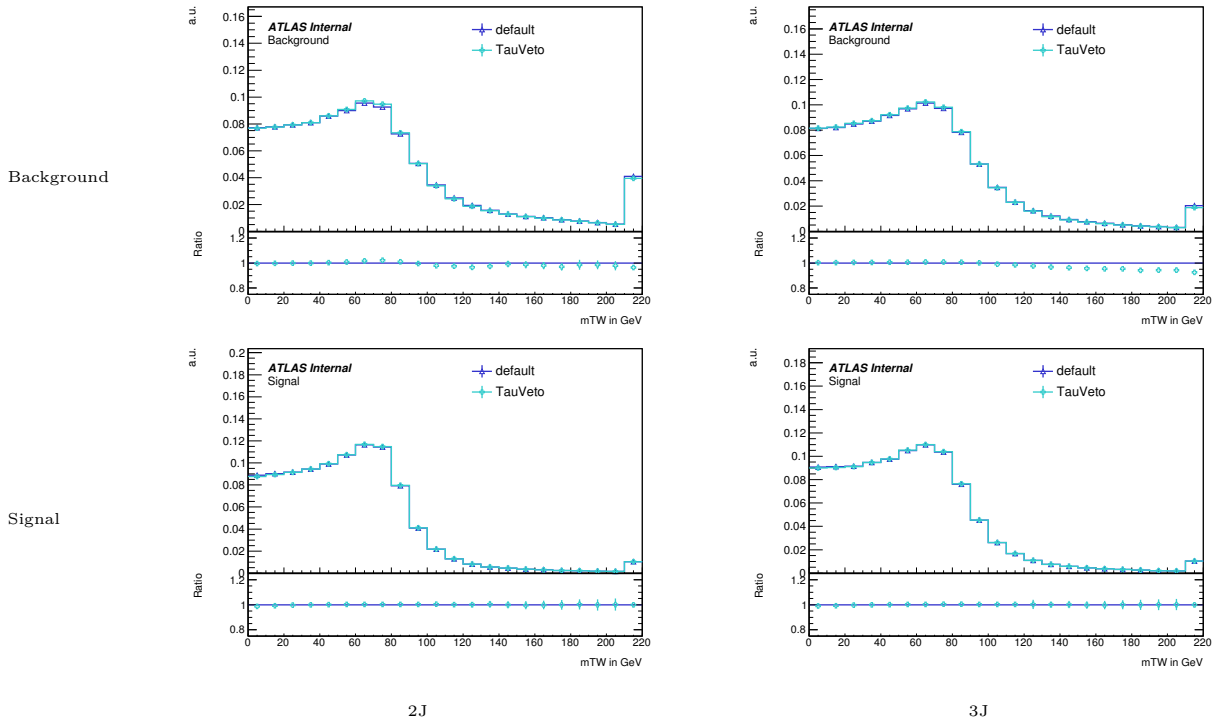


Figure 2.16: MVA Input Variable Comparison – mTW

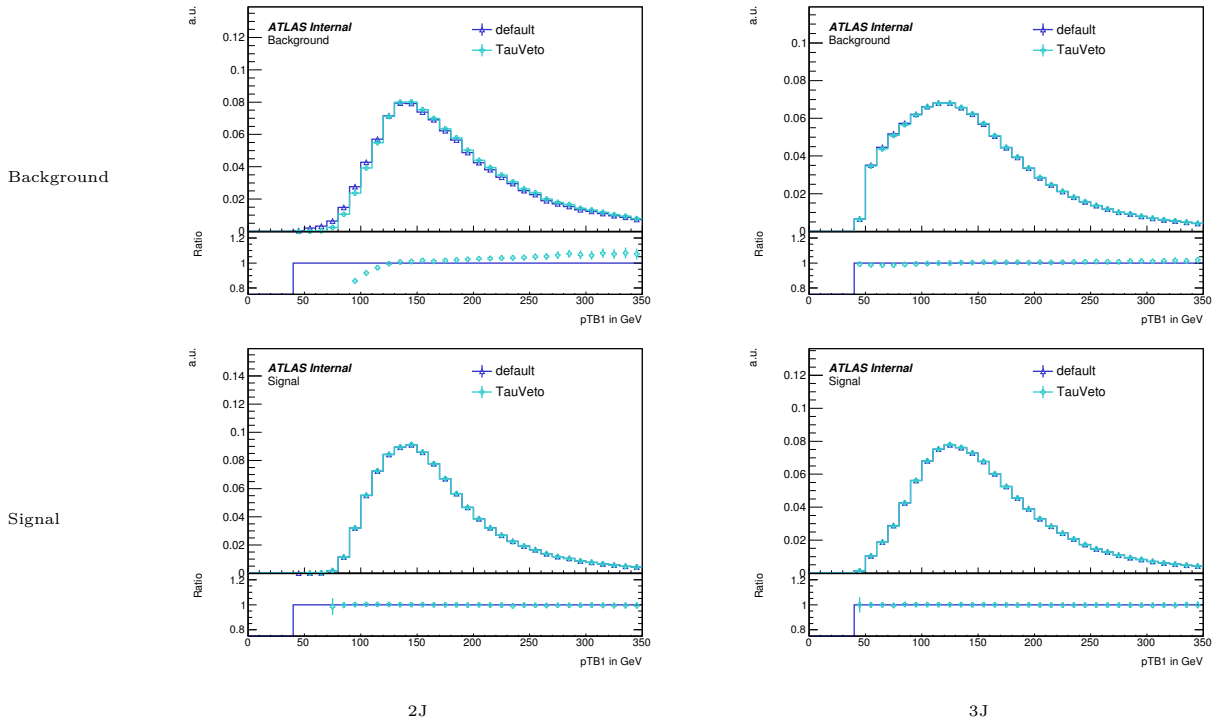


Figure 2.17: MVA Input Variable Comparison – pTB1

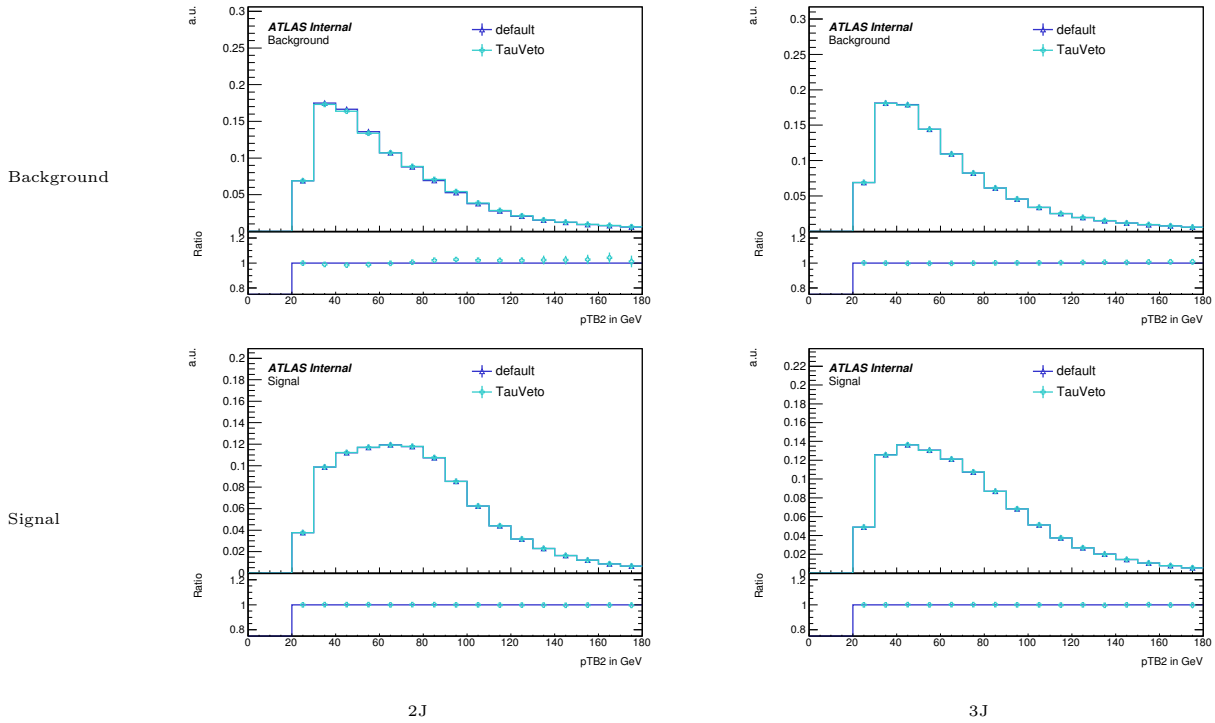


Figure 2.18: MVA Input Variable Comparison – pTB2

B.5. Correlation Matrices

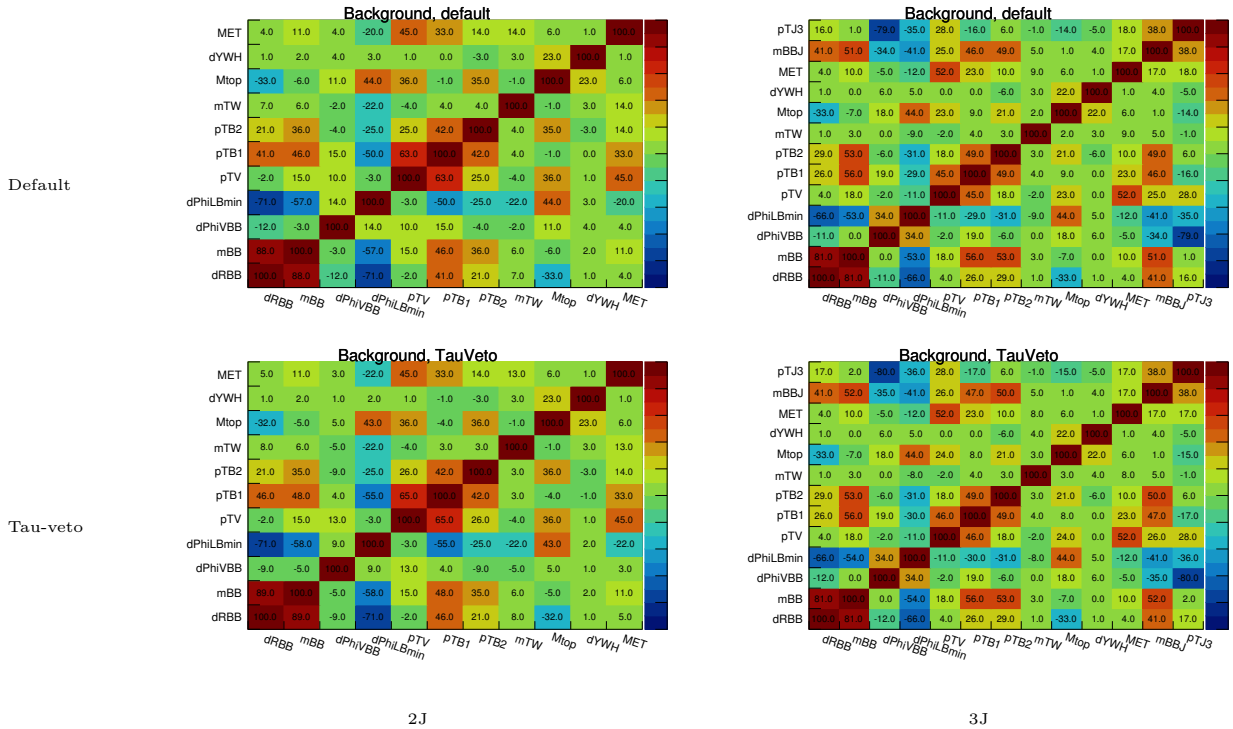


Figure 2.19: Correlation Matrix – Background

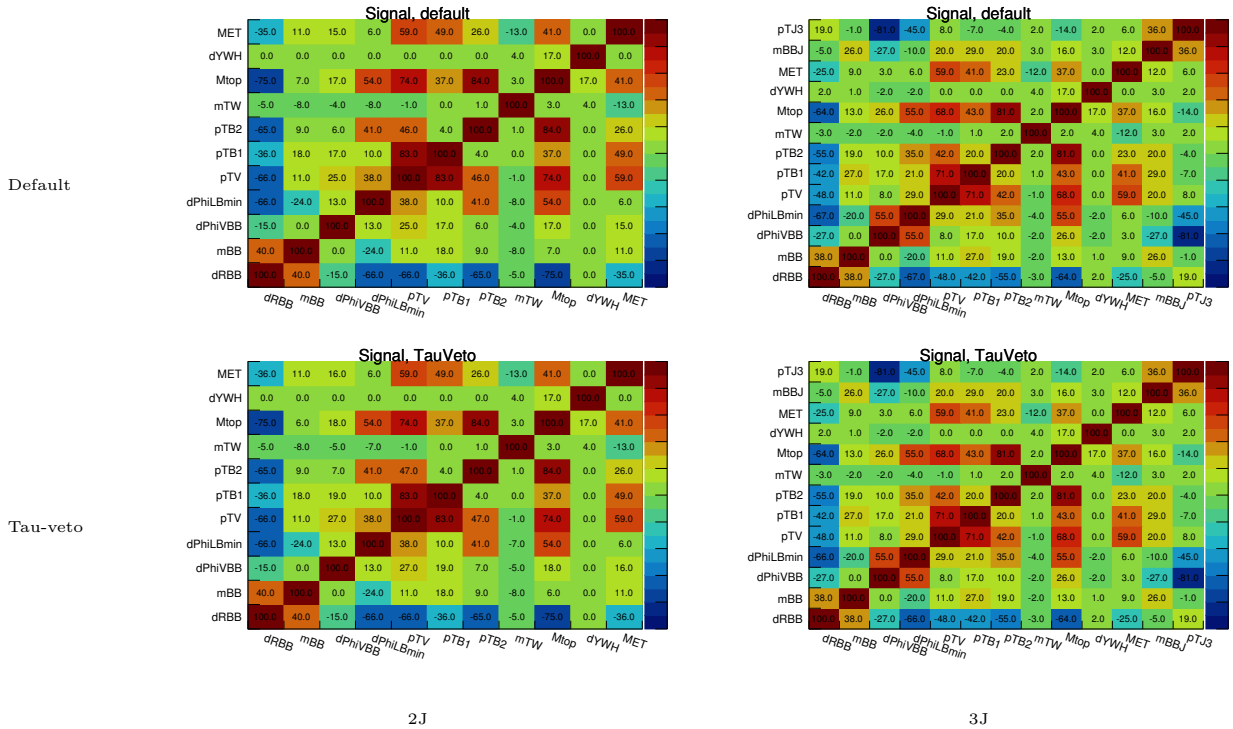


Figure 2.20: Correlation Matrix – Signal

B.6. Evaluation in the region of $150 < p_{\mathbf{T}}^{V} < 250$ GeV

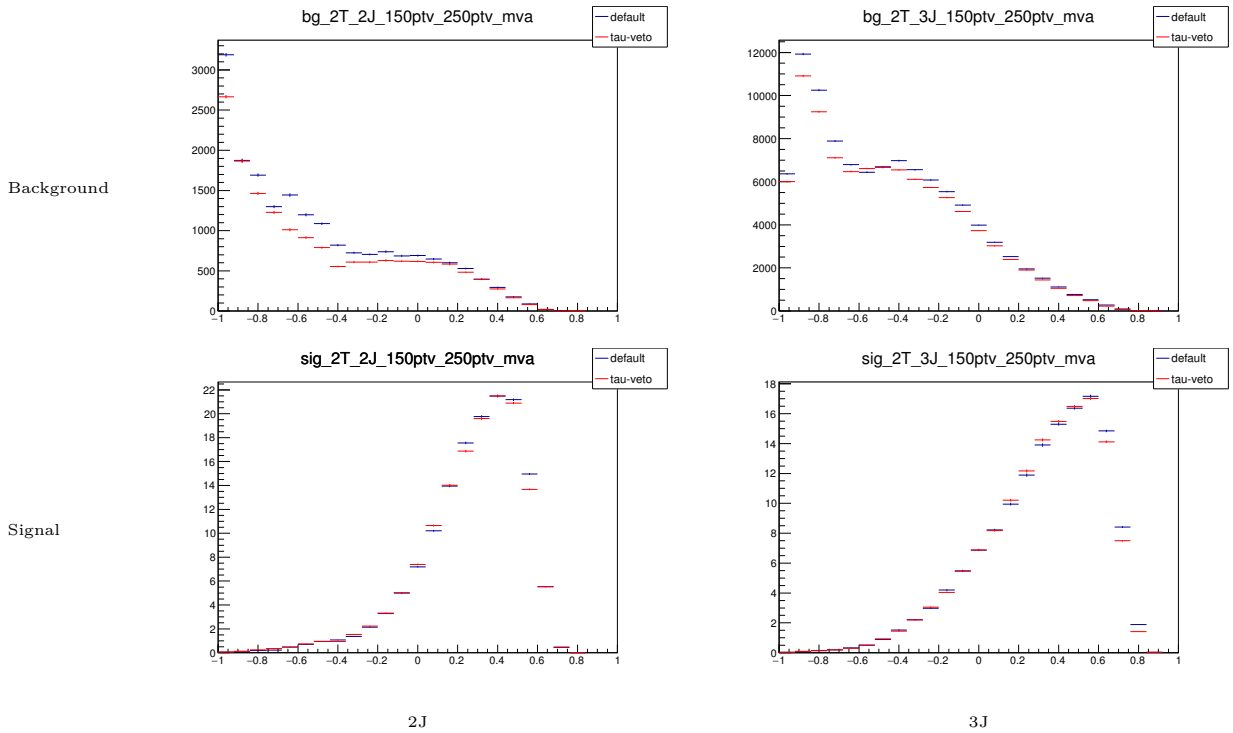


Figure 2.21: MVA 150_250ptv

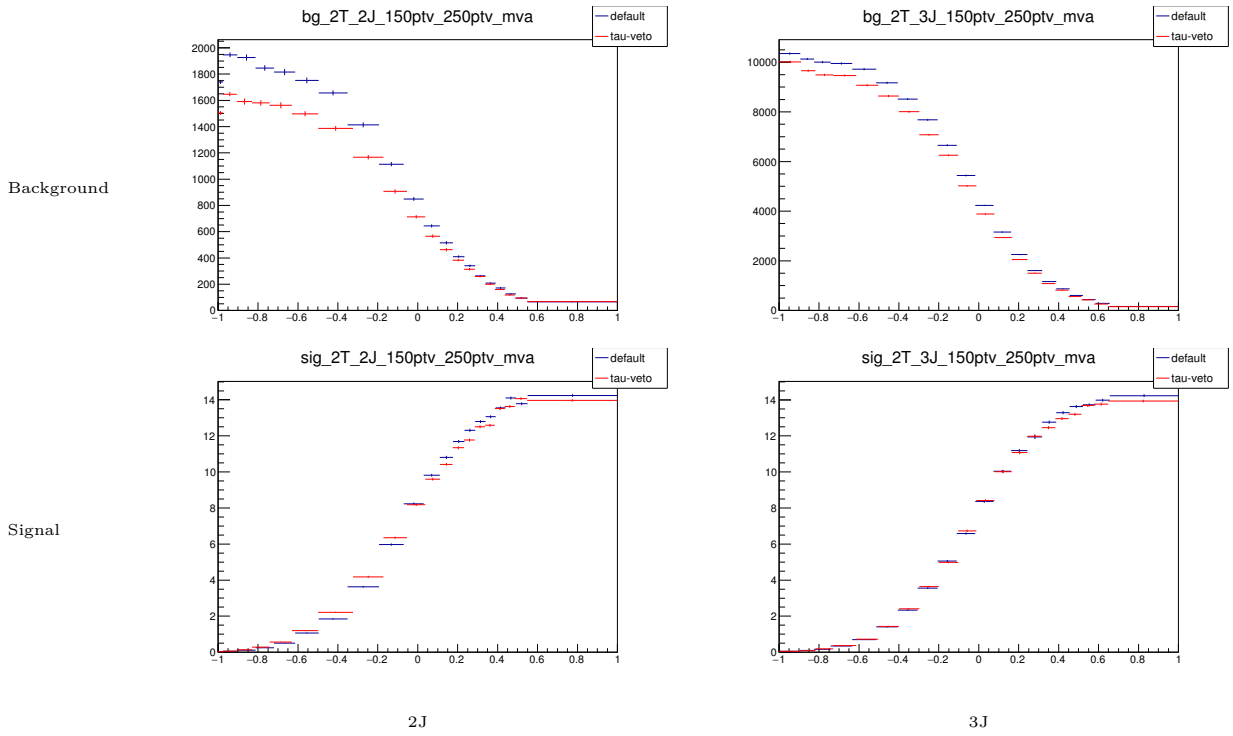


Figure 2.22: MVA TrafoD 150_250ptv

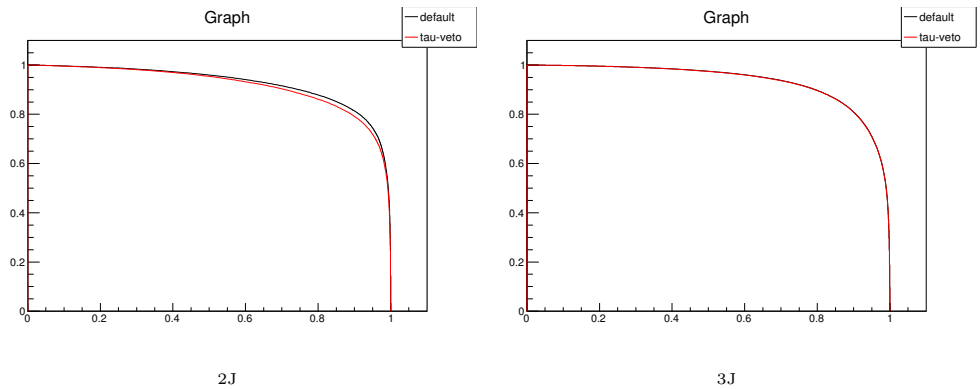


Figure 2.23: ROC Curve 150_250ptv

	default	tau-veto
2T_2J_150ptv_250ptv	3.17014	3.17402
2T_3J_150ptv_250ptv	1.77993	1.79894

Table 2.5: Significance (after trafoD)

	default	tau-veto
2T_2J_150ptv_250ptv	0.925767	0.917221
2T_3J_150ptv_250ptv	0.932634	0.932101

Table 2.6: ROC integral

B.7. Evaluation in the region of $p_{\mathbf{T}}^{V} > 250$ GeV

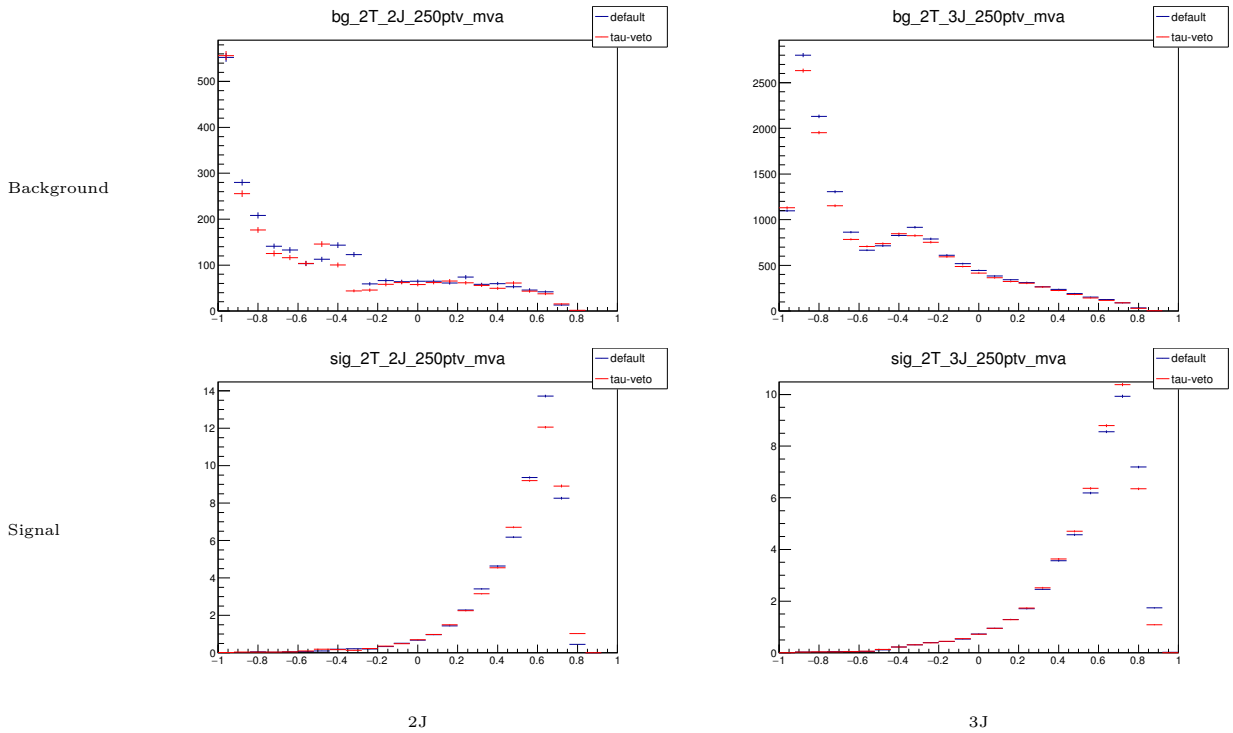


Figure 2.24: MVA 250ptv

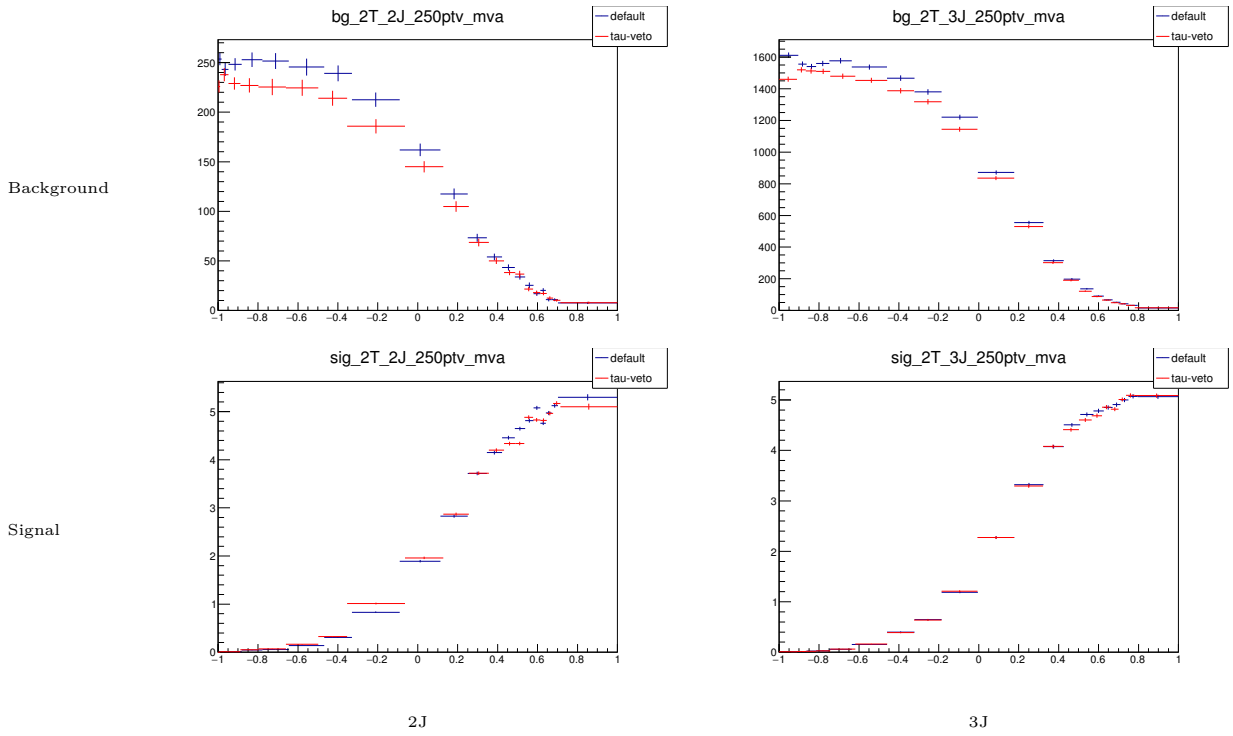


Figure 2.25: MVA TrafoD 250ptv

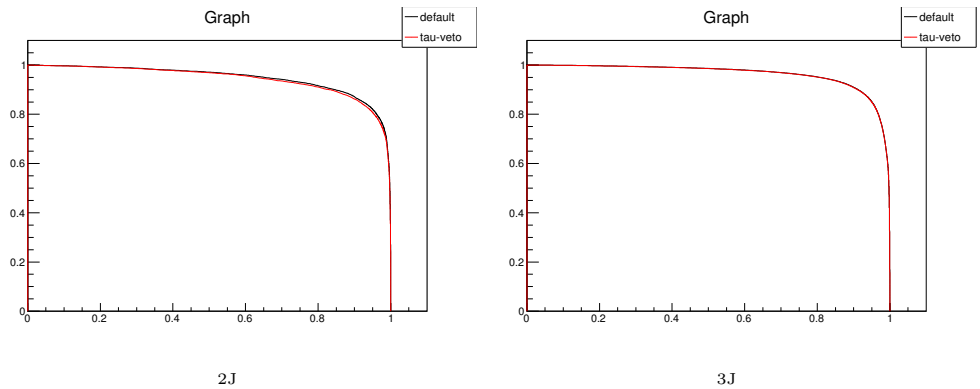


Figure 2.26: ROC Curve 250ptv

	default	tau-veto
2T_2J_250ptv	3.47389	3.44802
2T_3J_250ptv	2.0587	2.03135

Table 2.7: Significance (after trafoD)

	default	tau-veto
2T_2J_250ptv	0.947409	0.943598
2T_3J_250ptv	0.963691	0.963208

Table 2.8: ROC integral

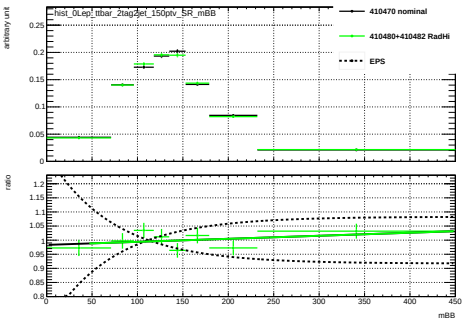
APPENDIX C

$t\bar{t}$ Modeling: Shape Uncertainties Using the Shape Comparison Method

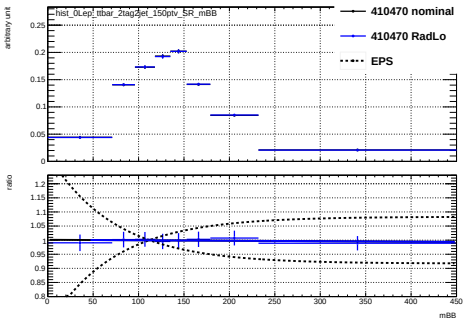
C.1. Fits for shape uncertainties

C.1.1. 0-lepton fits

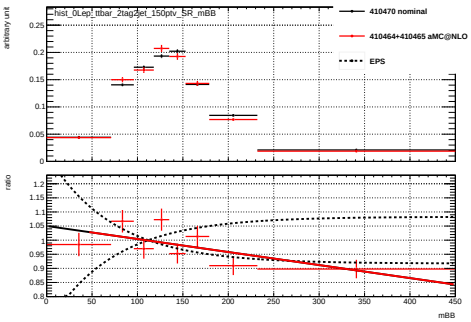
The $m_{b\bar{b}}$ shape variation in 0-lepton 2-jet and 3-jet regions is shown in Fig. 3.1; for the case of ISR variations, no significant shapes are observed. The p_T^V shape variation in 0-lepton 2-jet and 3-jet regions is shown in Fig. 3.2.



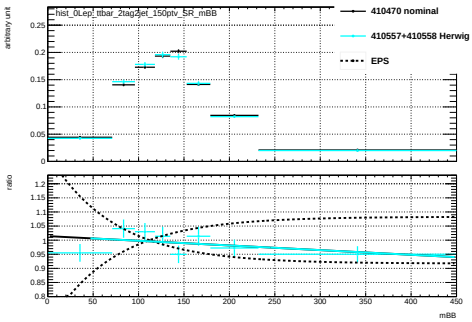
(a) I/FSR variation (up) for 2-jet



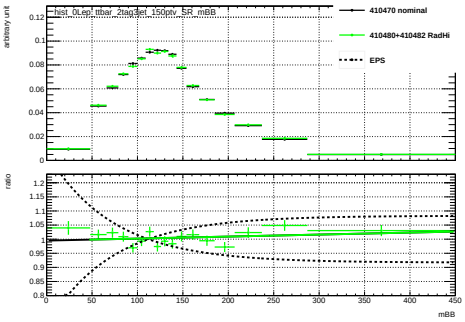
(b) ISR variation (down) for 2-jet



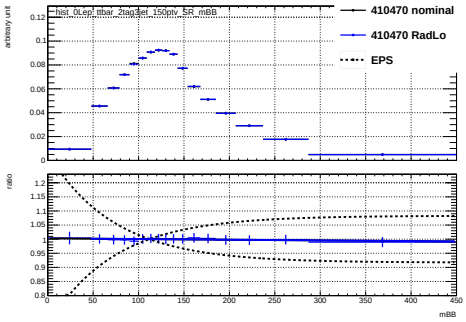
(c) ME variation for 2-jet



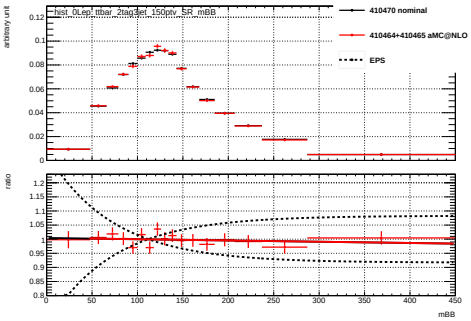
(d) PS variation for 2-jet



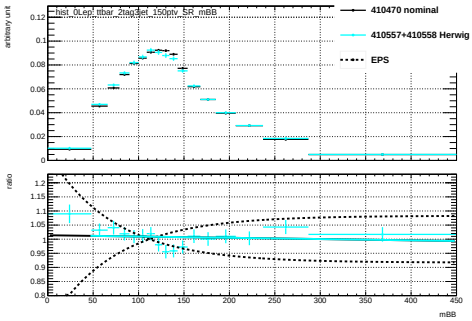
(e) ISR variation (up) for 3-jet



(f) ISR variation (down) for 3-jet

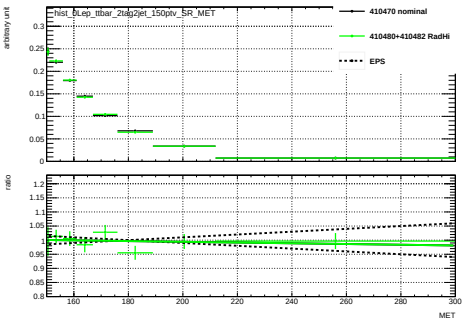


(g) ME variation for 3-jet

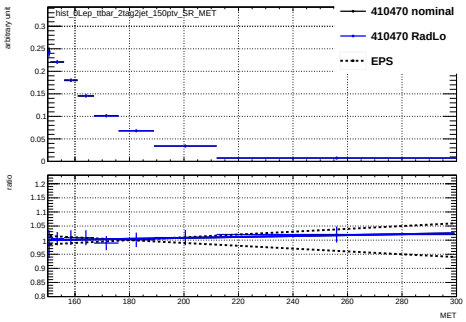


(h) PS variation for 3-jet

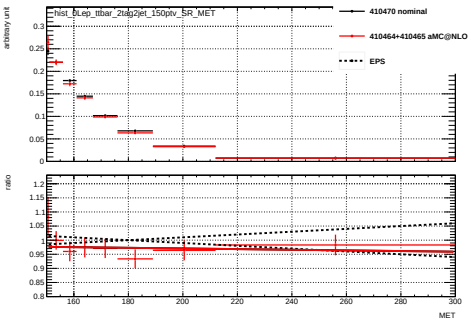
Figure 3.1: $m_{b\bar{b}}$ shape variation in 0-lepton for the 2-/3-jet category. ISR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7.



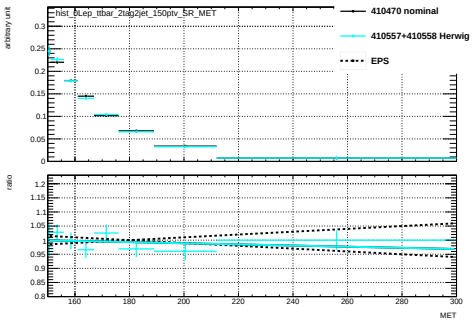
(a) I/FSR variation (up) for 2-jet



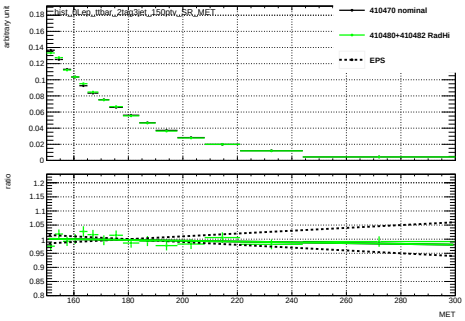
(b) ISR variation (down) for 2-jet



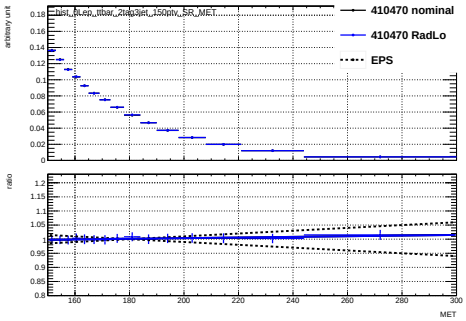
(c) ME variation for 2-jet



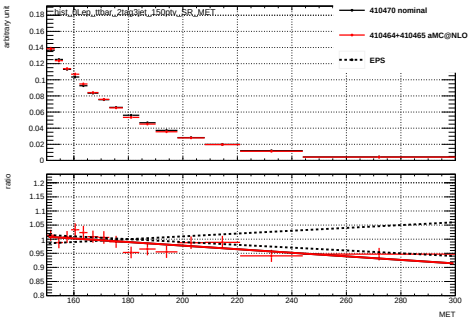
(d) PS variation for 2-jet



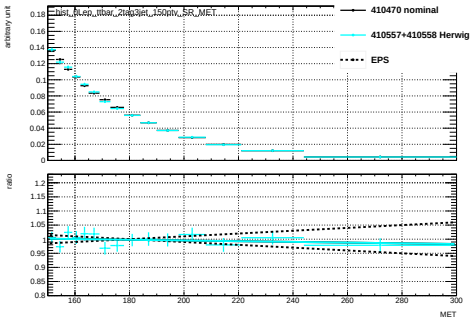
(e) ISR variation (up) for 3-jet



(f) ISR variation (down) for 3-jet



(g) ME variation for 3-jet

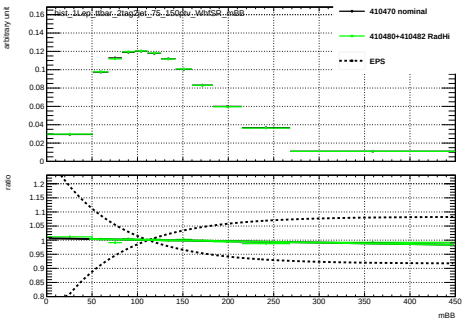


(h) PS variation for 3-jet

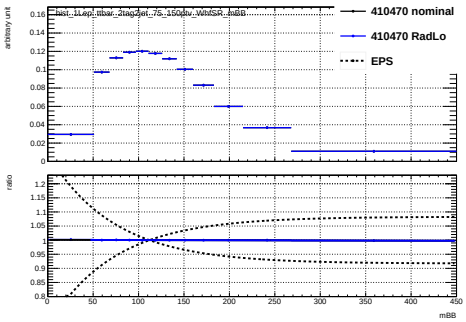
Figure 3.2: E_T^{miss} shape variation in 0-lepton for the 2-/3-jet category. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7.

C.1.2. 1-lepton fits

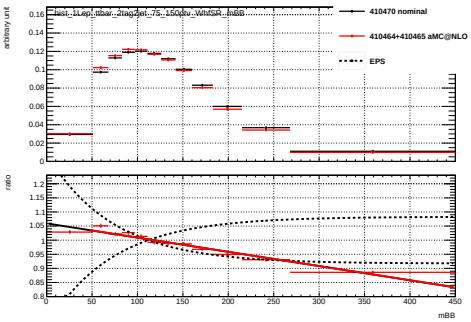
The $m_{b\bar{b}}$ shape variation in 1-lepton 2jet and 3jet regions is shown in Fig. 3.3 and Fig. 3.4; for the case of ISR variations, no significant shapes are observed. The p_T^V shape variation in 1-lepton 2jet and 3jet regions is shown in Fig. 3.5



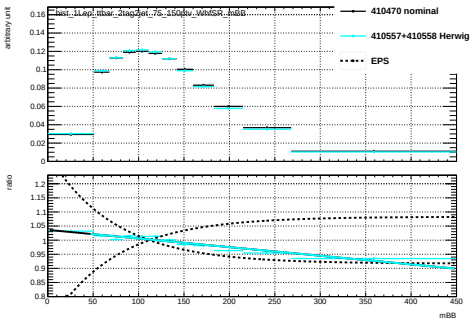
(a) I/FSR variation (up) for 2-jet



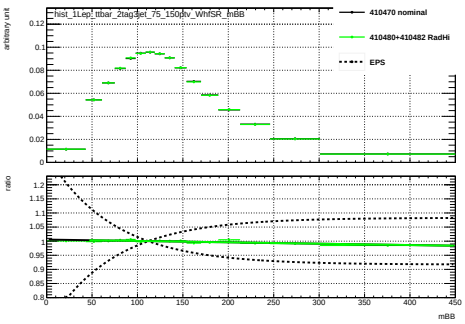
(b) I/FSR variation (down) for 2-jet



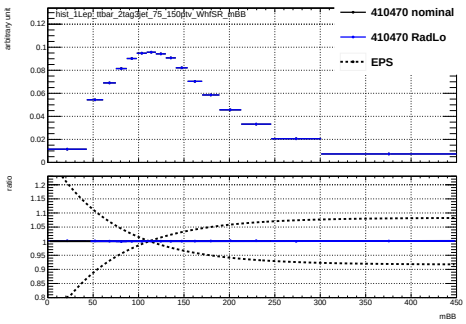
(c) ME variation for 2-jet



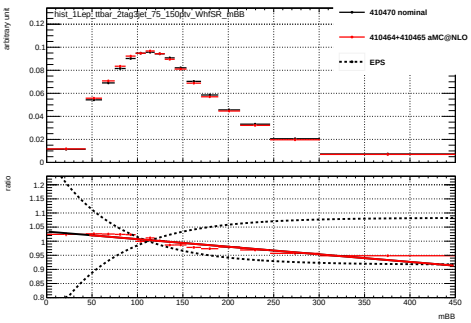
(d) PS variation for 2-jet



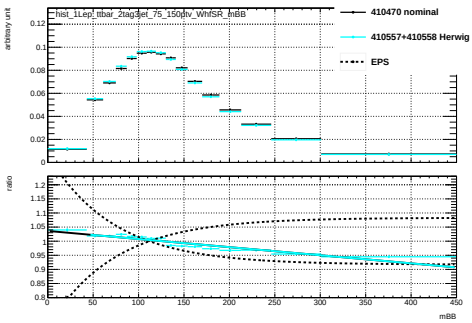
(e) I/FSR variation (up) for 3-jet



(f) I/FSR variation (down) for 3-jet

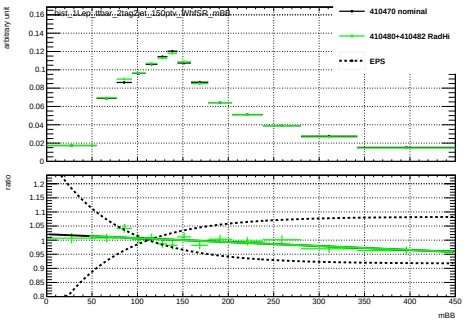


(g) ME variation for 3-jet

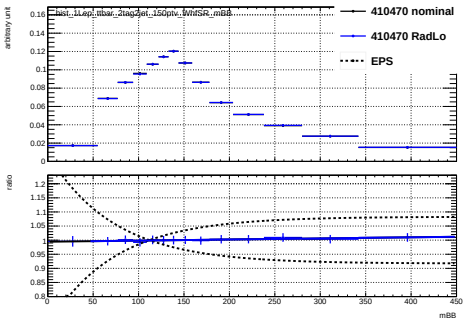


(h) PS variation for 3-jet

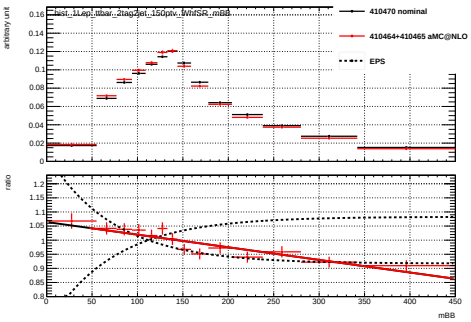
Figure 3.3: $m_{b\bar{b}}$ shape variation in 1-lepton for the 75GeV $< p_T^V <$ 150GeV regime. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7.



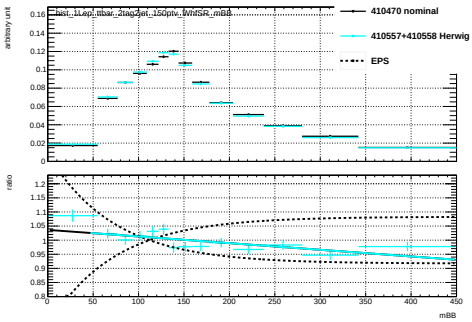
(a) I/FSR variation (up) for 2-jet



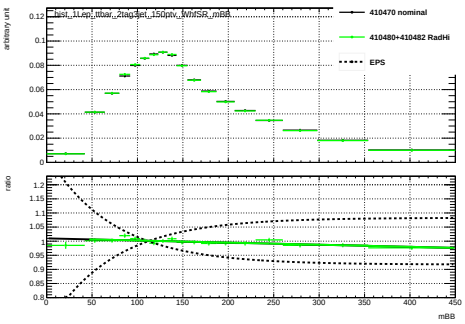
(b) I/FSR variation (down) for 2-jet



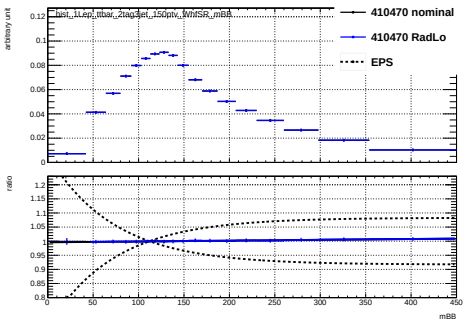
(c) ME variation for 2-jet



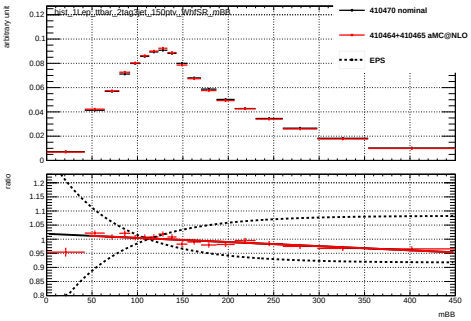
(d) PS variation for 2-jet



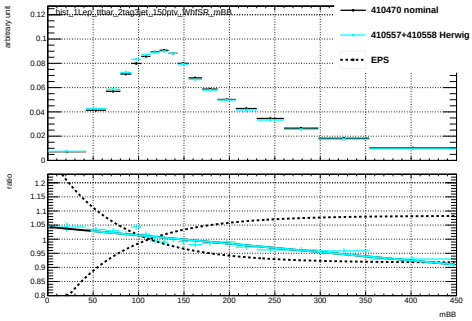
(e) I/FSR variation (up) for 3-jet



(f) I/FSR variation (down) for 3-jet

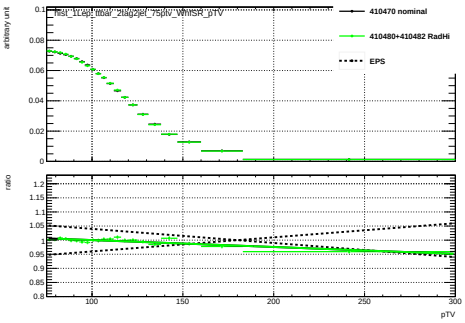


(g) ME variation for 3-jet

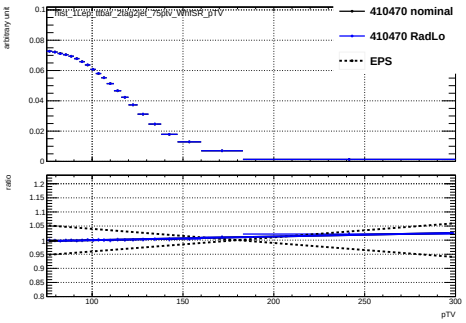


(h) PS variation for 3-jet

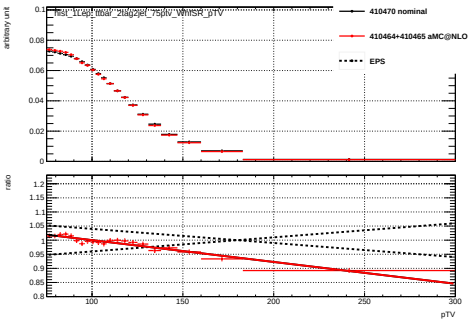
Figure 3.4: $m_{b\bar{b}}$ shape variation in 1-lepton for the $p_T^V > 150$ GeV regime. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7.



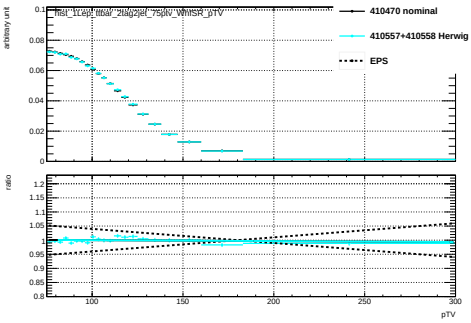
(a) I/FSR variation (up) for 2-jet



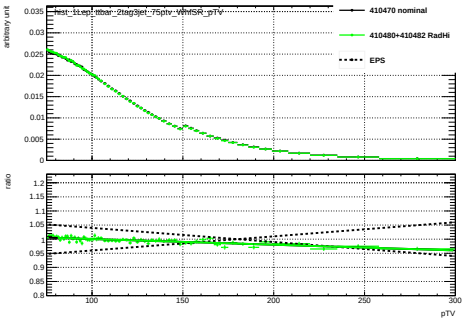
(b) I/FSR variation (down) for 2-jet



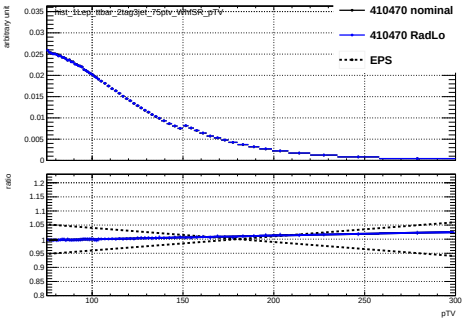
(c) ME variation for 2-jet



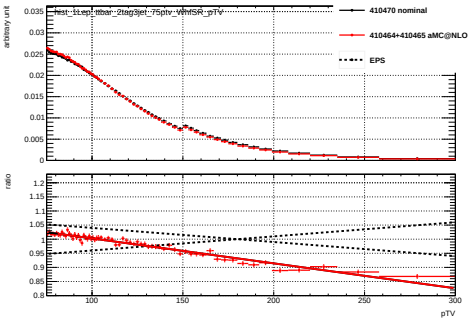
(d) PS variation for 2-jet



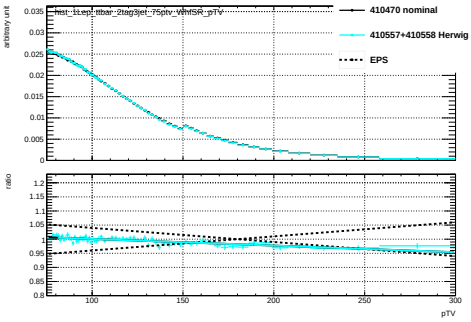
(e) I/FSR variation (up) for 3-jet



(f) I/FSR variation (down) for 3-jet



(g) ME variation for 3-jet



(h) PS variation for 3-jet

Figure 3.5: p_T^V shape variation in 1-lepton channel, inclusive of $p_T^V > 75$GeV. I/FSR variation is the scale variation, ME variation is comparison to aMC@NLO, and PS variation is comparison to Powheg +Herwig 7.

APPENDIX D

Pre-fit Data-MC Comparison in the 1-lepton Channel

This section shows the data-MC comparison plots in the 1-lepton channel before input to the MVA and fit to the template. The m_{bb} and p_T^W variables are shown in terms of the p_T^W regions, jet categories, and signal/control regions. Fig. 4.1 and 4.2 correspond to the distribution of m_{bb}, and Fig. 4.3 and 4.4 correspond to the distribution of p_T^W .

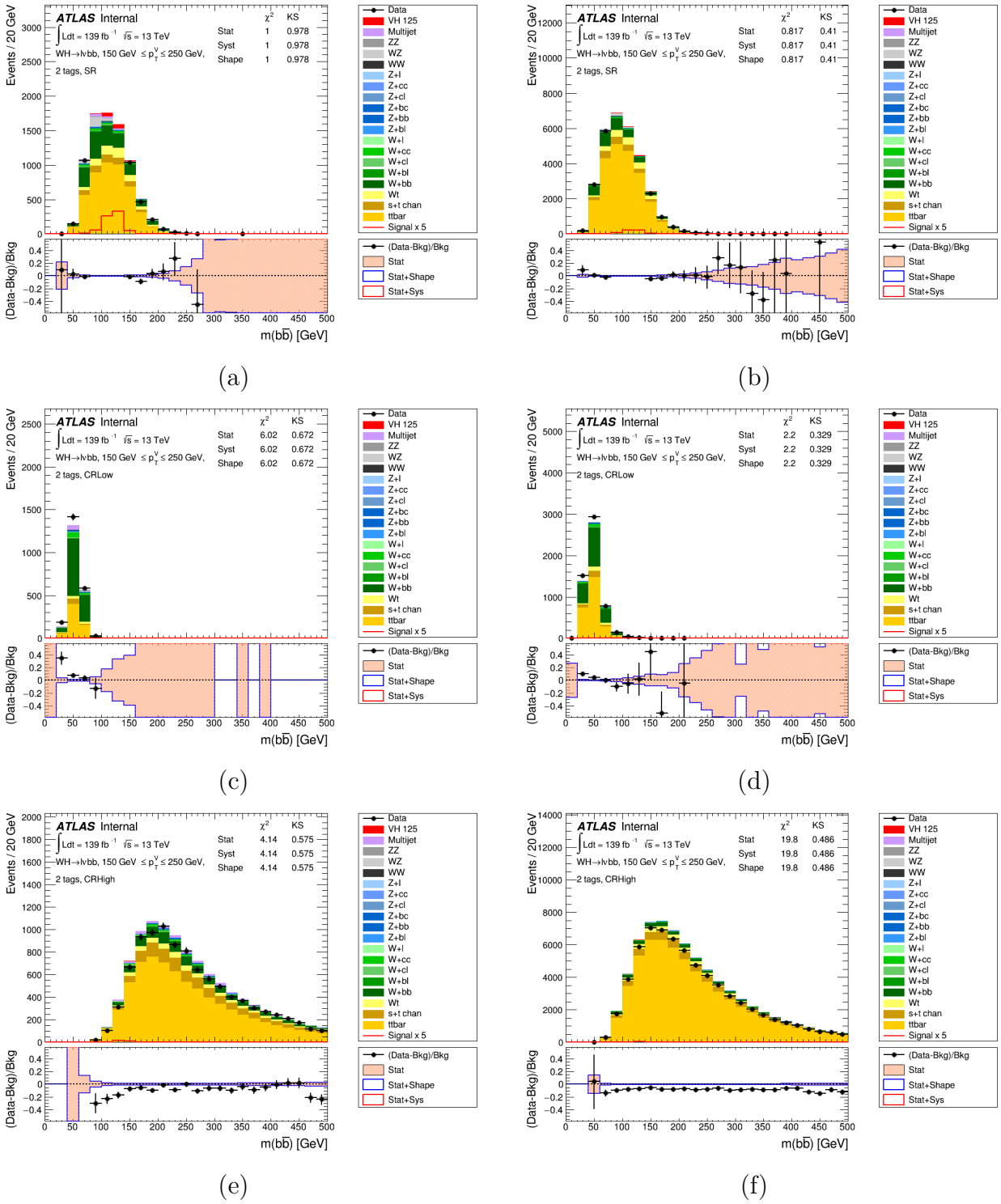


Figure 4.1: The pre-fit 150 GeV$< p_T^W <$ 250 GeV m_{bb} distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109].

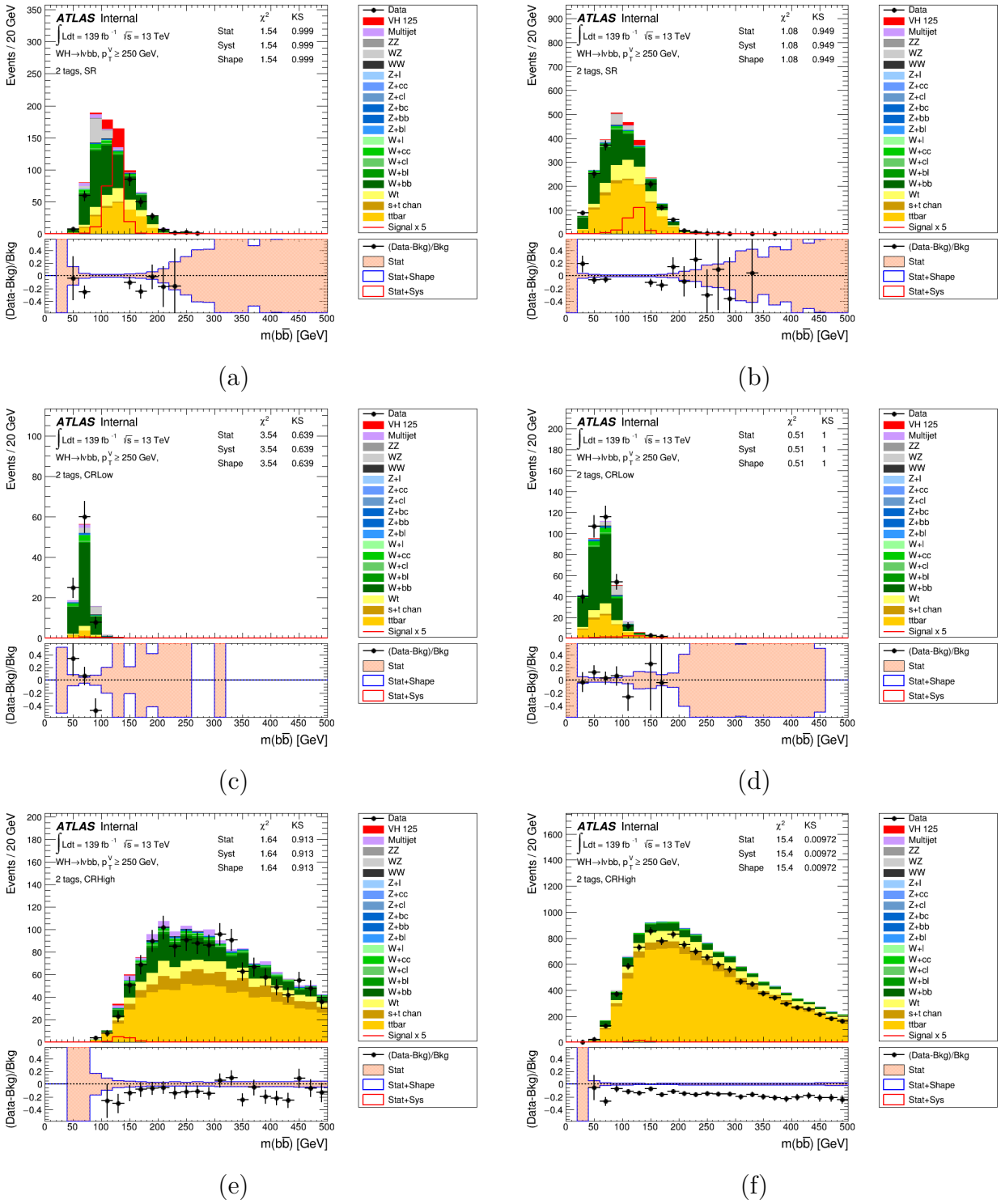


Figure 4.2: The pre-fit $p_T^W > 250$ GeV m_{bb} distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109].

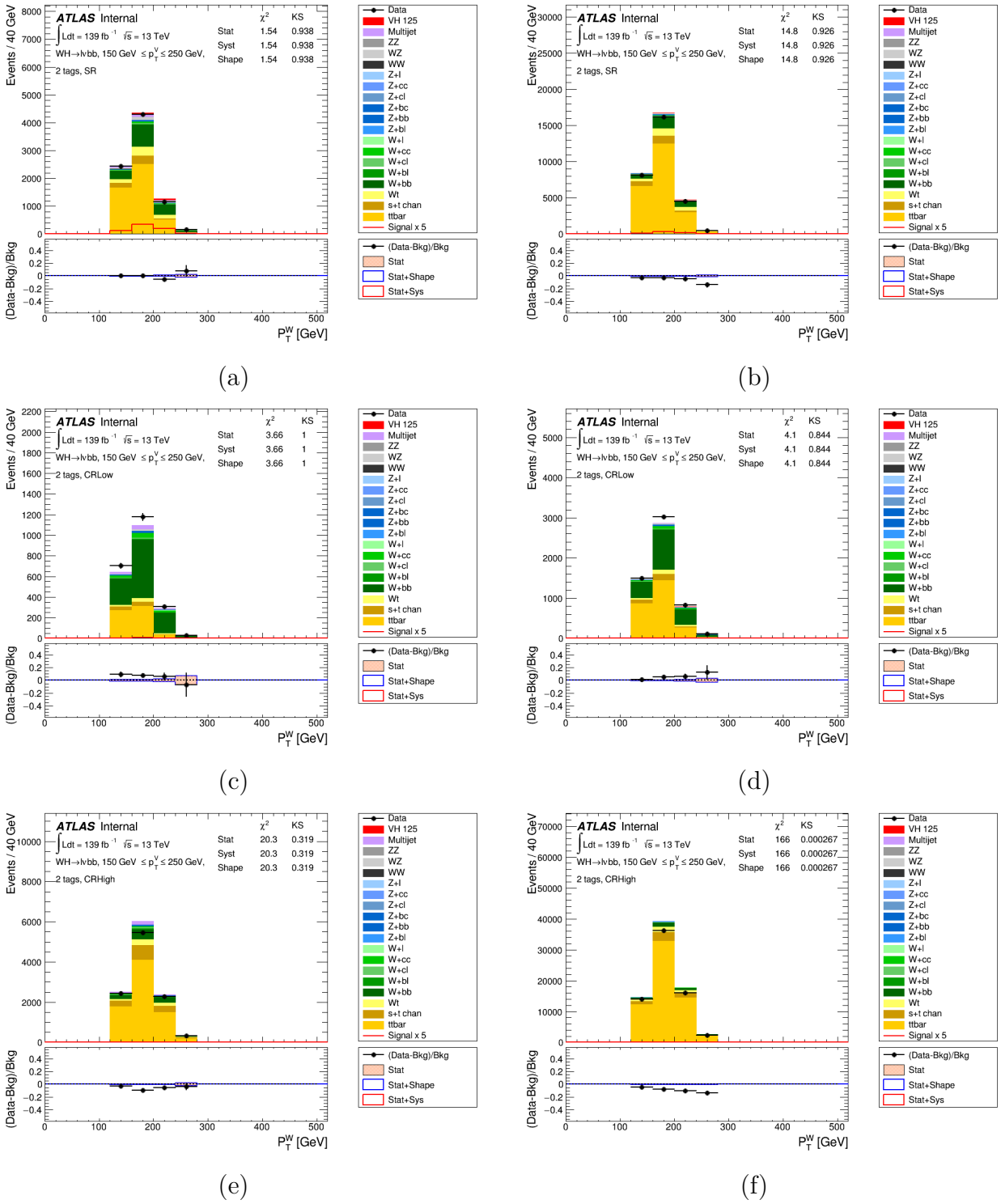


Figure 4.3: The pre-fit 150 GeV$< p_T^W <$ 250 GeV p_T^W distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109].

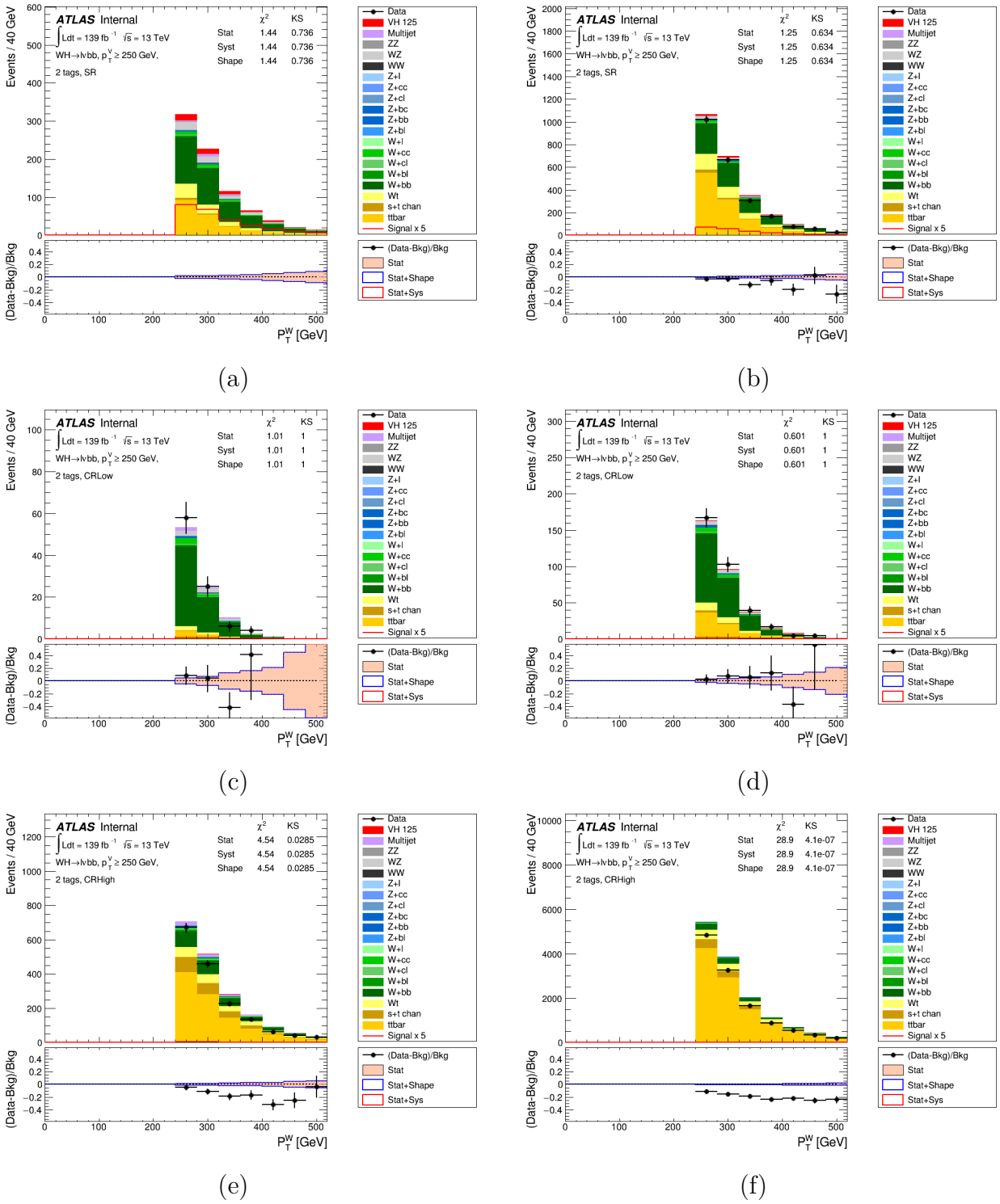


Figure 4.4: The pre-fit $p_T^W > 250$ GeV p_T^W distributions in the 1-lepton channel in the signal region in 2-tag 2-jet (a), 2-tag 3-jet (b), the low ΔR control region in 2-tag 2-jet (c), 2-tag 3-jet (d) and the high ΔR control region in 2-tag 2-jet (e) and 2-tag 3-jet (f). The background and signal samples are normalized to the expected cross-section predictions [109].

BIBLIOGRAPHY

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, *Nuclear Physics* **22** (1961) 579 – 588. 1, 4

[2] S. Weinberg, *A model of leptons*, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. 1, 4

[3] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, *Nuclear Physics B* **44** (1972) 189 – 213. 1

[4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323. 1

[5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509. 1

[6] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. 1

[7] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587. 1

[8] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1, [1207.7214]. 1

[9] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, [1207.7235]. 1

[10] A. Djouadi, J. Kalinowski and M. Spira, *Hdecay: a program for higgs boson decays in the standard model and its supersymmetric extension*, *Computer Physics Communications* **108** (1998) 56 – 74. 1

[11] M. Shochet, L. Tompkins, V. Cavaliere, P. Giannetti, A. Annovi and G. Volpi, *Fast TracKer (FTK) Technical Design Report*, Tech. Rep. CERN-LHCC-2013-007. ATLAS-TDR-021, Jun, 2013. 3

[12] A. Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva, Sep, 2017. 3

[13] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (Oct, 1954) 191–195. 4

[14] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, *Nuovo Cim.* **11** (1959) 568–577. 4

[15] D. Dominguez, *Standard Model. Le modèle standard*, . xii, 5

[16] W. Gerlach and O. Stern, *Der experimentelle nachweis der richtungsquantelung im magnetfeld*, *The European physical journal* **9** (1922) 349–352. 6

[17] I. Aitchison and A. Hey, *Gauge theories in particle physics: A practical introduction. Vol. 1: From relativistic quantum mechanics to QED*. CRC Press, 2012. 7, 10

[18] D. Griffiths, *Introduction to elementary particles*. 2008. 9

[19] I. Aitchison and A. Hey, *Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory*. CRC Press, 2012. 11, 12, 14, 15

[20] K. Hornbostel, *Field Theory and Elementary Particles Lecture Note*. http://www.physics.smu.edu/kjh/7314/15/7314.html, 2016. 16

[21] W. Cottingham and D. Greenwood, *An introduction to the standard model of particle physics*. Cambridge University Press, 4, 2007. 19

[22] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013. xxi, 20

[23] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, *Phys. Rept.* **457** (2008) 1–216, [hep-ph/0503172]. xii, 21

[24] LHC Higgs Cross Section Working Group collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922. xii, xxi, 22, 24, 134

[25] Particle Data Group collaboration, M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi et al., *Review of particle physics*, *Phys. Rev. D* **98** (Aug, 2018) 030001. 23

[26] L. Evans and P. Bryant, *LHC machine*, *Journal of Instrumentation* **3** (aug, 2008) S08001–S08001. xii, 26, 28, 29, 30, 31

[27] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN." Dec, 2008. xii, 26

[28] D. Boussard and T. P. R. Linnecar, *The LHC Superconducting RF System*, Tech. Rep. LHC-Project-Report-316. CERN-LHC-Project-Report-316, CERN, Geneva, Dec, 1999. 27

[29] J. Ellis, *Physics at the LHC*, *European Physical Journal C* **34** (May, 2004) 51–56. 27

[30] B. Heinemann, J. de Blas, M. Cepeda, C. Grojean, F. Maltoni, A. Nisati et al., *Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020*, Tech. Rep. arXiv:1910.11775, Geneva, Oct, 2019. 27

[31] ATLAS Collaboration collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*. LHC Tech. Proposal. CERN, Geneva, 1994. 28

[32] CMS Collaboration collaboration, *Technical proposal*. LHC Tech. Proposal. CERN, Geneva, 1994. 28

[33] ALICE Collaboration collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*. LHC Tech. Proposal. CERN, Geneva, 1995. 28

[34] *LHCb : Technical Proposal*. Tech. Proposal. CERN, Geneva, 1998. 28

[35] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole et al., *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004, 10.5170/CERN-2004-003-V-1. 28

[36] L. Guiraud, "Model of an LHC superconducting quadrupole magnet. Aimant quadripôle supraconducteur." Jan, 2000. xii, 29

[37] ATLAS Collaboration collaboration, ATLAS, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. 437 p. xiii, xxi, 32, 34, 35, 36, 37, 39

[38] ATLAS Collaboration collaboration, A. Airapetian, V. Grabsky, H. Hakopian, A. Vartapetian, B. Dick, F. Fares et al., *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999. 31

[39] J.-L. Caron, "Overall view of LHC experiments.. Vue d'ensemble des experiences du LHC.." May, 1998. xiii, 33

[40] H. H. J. ten Kate, *Atlas superconducting magnet system status*, *IEEE Transactions on Applied Superconductivity* **17** (June, 2007) 1191–1196. 35

[41] ATLAS collaboration, H. H. J. ten Kate, *ATLAS superconducting toroids and solenoid*, *IEEE Trans. Appl. Supercond.* **15** (2005) 1267–1270. 4 p. 36

[42] ATLAS Collaboration collaboration, *ATLAS magnet system: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997. 36

[43] ATLAS Collaboration collaboration, *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997. 36

[44] ATLAS collaboration, G. Aad et al., *The ATLAS Inner Detector commissioning and calibration*, *Eur. Phys. J.* **C70** (2010) 787–821, [1004.5293]. 37

[45] J. Pinfold, J. Soukup, J. Archambault, C. Cojocaru, M. Khakzad, G. Oakham et al., *Performance of the atlas liquid argon endcap calorimeter in the pseudorapidity region 2.5<||<4.0 in beam tests*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **593** (2008) 324 – 342. 37

[46] B. Dowler, J. Pinfold, J. Soukup, M. Vincter, A. Cheplakov, V. Datskov et al., *Performance of the atlas hadronic end-cap calorimeter in beam tests*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **482** (2002) 94 – 124. 37

[47] A. Artamonov et al., *The ATLAS forward calorimeters*, *JINST* **3** (2008) P02010. 37

[48] ATLAS Collaboration collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. 37

[49] ATLAS Collaboration collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. 38

[50] ATLAS Collaboration collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. 38

[51] M. Livan, *Monitored drift tubes in atlas*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **384** (1996) 214 – 218. 38

[52] T. Argyropoulos et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, *IEEE Trans. Nucl. Sci.* **56** (2009) 1568–1574. 38

[53] E. C. Zeballos, I. Crotty, D. Hatzifotiadou, J. L. Valverde, S. Neupane, M. Williams et al., *A new type of resistive plate chamber: The multigap rpc*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **374** (1996) 132 – 135. 38

[54] K. Nagai, *Thin gap chambers in atlas*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **384** (1996) 219 – 221. 38

[55] M. Bruschi, *The new ATLAS/LUCID detector*, Tech. Rep. ATL-FWD-PROC-2015-002, CERN, Geneva, Jun, 2015. xiii, 39, 40

[56] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. 40

[57] ATLAS Collaboration collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, Tech. Rep. ATLAS-CONF-2019-021, CERN, Geneva, Jun, 2019. 40, 56

[58] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos et al., *Performance of the atlas trigger system in 2015*, *The European Physical Journal C* **77** (May, 2017) . 41

[59] A. R. M. and, *The run-2 ATLAS trigger system*, *Journal of Physics: Conference Series* **762** (oct, 2016) 012003. xiii, 41

[60] W3C Recommendation collaboration, *PICMG 3.0 Revision 3.0 AdvancedTCA Base Specification*, 2015. 42

[61] W. R. Stevens, *UNIX Network Programming: Networking APIs: Sockets and XTI*. Prentice Hall PTR, USA, 2nd ed., 1997. 46

[62] ATLAS collaboration, G. Aad et al., *ATLAS data quality operations and performance for 2015-2018 data-taking*, 1911.04632. xiv, 55, 56

[63] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *Journal of Instrumentation* **13** (jul, 2018) P07017–P07017. 56

[64] M. Böhler, M. Borodin, A. Formica, E. Gallas and V. Radescu, *Evolution of ATLAS conditions data and its management for LHC run-2*, *Journal of Physics: Conference Series* **664** (dec, 2015) 042005. 56

[65] M. Verducci, *ATLAS conditions database experience with the LCG COOL conditions database project*, *Journal of Physics: Conference Series* **119** (jul, 2008) 042031. 56

[66] W3C Recommendation collaboration, *Extensible Markup Language (XML) 1.0 (Fifth Edition)*, 2008. 56

[67] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce et al., *Geant4—a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250 – 303. 56

[68] R. Teuscher, *High ET Jet Physics*, *Physics in Collision conference in 2011* (2011) . xiv, 57

[69] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, [1410.3012]. 58

[70] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196, [1512.01178]. 58

[71] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [0811.4622]. 58

[72] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146]. 58, 62

[73] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [1405.0301]. 58, 61

[74] D. J. Lange, *The evtgen particle decay simulation package*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** (2001) 152 – 155. 58

[75] S. Jadach, J. H. Kühn and Z. Was, *Tauola - a library of monte carlo programs to simulate decays of polarized leptons*, *Computer Physics Communications* **64** (1991) 275 – 299. 58

[76] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [1002.2581]. 58

[77] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: multi-scale improved NLO*, *JHEP* **10** (2012) 155, [1206.3572]. 58

[78] G. Luisoni, P. Nason, C. Oleari and F. Tramontano, *$HW^{\pm}/HZ$ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, *JHEP* **10** (2013) 083, [1306.2542]. 58

[79] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [0710.3820]. 58

[80] ATLAS collaboration, G. Aad et al., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **09** (2014) 145, [1406.3660]. 58, 59

[81] NNPDF collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [1410.8849]. 58

[82] M. Ciccolini, S. Dittmaier and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, *Phys. Rev. D* **68** (2003) 073003, [hep-ph/0306234]. xxi, 59

[83] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Phys. Lett.* **B579** (2004) 149–156, [hep-ph/0307206]. xxi, 59

[84] G. Ferrera, M. Grazzini and F. Tramontano, *Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO*, *Phys. Rev. Lett.* **107** (2011) 152003, [1107.1164]. xxi, 59

[85] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, *Eur. Phys. J.* **C72** (2012) 1868, [1111.0761]. xxi, 59

[86] G. Ferrera, M. Grazzini and F. Tramontano, *Higher-order QCD effects for associated WH production and decay at the LHC*, *JHEP* **04** (2014) 039, [1312.1669]. xxi, 59

[87] G. Ferrera, M. Grazzini and F. Tramontano, *Associated ZH production at hadron colliders: the fully differential NNLO QCD calculation*, *Phys. Lett. B* **740** (2015) 51–55, [1407.4747]. xxi, 59

[88] J. M. Campbell, R. K. Ellis and C. Williams, *Associated production of a Higgs boson at NNLO*, *JHEP* **06** (2016) 179, [1601.00658]. xxi, 59

[89] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, *JHEP* **02** (2013) 078, [1211.5015]. xxi, 59, 60

[90] B. Hespel, F. Maltoni and E. Vryonidou, *Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM*, *JHEP* **06** (2015) 065, [1503.01656]. xxi, 59, 60

[91] R. V. Harlander, A. Kulesza, V. Theeuwes and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, *JHEP* **11** (2014) 082, [1410.0217]. xxi, 59, 60

[92] R. V. Harlander, S. Liebler and T. Zirke, *Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-doublet model*, *JHEP* **02** (2014) 023, [1307.8122]. xxi, 59, 60, 134

[93] O. Brein, R. V. Harlander and T. J. E. Zirke, *vh@nnlo - Higgs Strahlung at hadron colliders*, *Comput. Phys. Commun.* **184** (2013) 998–1003, [1210.5347]. xxi, 59, 60, 134

[94] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [0811.4622]. 60

[95] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, [1111.5206]. 60

[96] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, [0808.3674]. 60

[97] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038, [0709.1027]. 60

[98] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, [1207.5030]. 60

[99] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244, [1207.1303]. 60, 63

[100] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD matrix elements + parton showers*, *Journal of High Energy Physics* **2001** (nov, 2001) 063–063. 60

[101] B. Webber, *Parton shower Monte Carlo event generators*, *Scholarpedia* **6** (2011) 10662. 60

[102] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, [1101.2599]. 60

[103] S. Catani and M. H. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett.* **B378** (1996) 287–301, [hep-ph/9602277]. 60

[104] T. Sjostrand and M. van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, *Phys. Rev.* **D36** (1987) 2019. 60

[105] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046, [hep-ph/0112284]. 61, 65

[106] N. Lavesson and L. Lonnblad, *W+jets matrix elements and the dipole cascade*, *JHEP* **07** (2005) 054, [hep-ph/0503293]. 61, 65

[107] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order*, *Phys. Rev. Lett.* **103** (2009) 082001, [0903.2120]. 61

[108] J. Butterworth, E. Dobson, U. Klein, B. Mellado Garcia, T. Nunnemann, J. Qian et al., *Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV*, Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva, Aug, 2010. 61

[109] K. Al Khoury, T. P. Calvet, K. Liu, E. Schopf, S. H. Abidi, L. Ambroz et al., *Measurement of $VH, H \to b\bar{b}$ with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2019-1276, CERN, Geneva, Oct, 2019. xvii, xx, xxi, xxii, xxiii, 62, 72, 73, 75, 76, 87, 88, 89, 91, 92, 114, 120, 121, 122, 124, 127, 139, 143, 150, 198, 199, 200, 201

[110] S. Frixione, P. Nason and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, [0707.3088]. 61

[111] ATLAS Collaboration, "ATLAS Pythia 8 tunes to 7 TeV data." ATL-PHYS-PUB-2014-021, 2014. 61

[112] ATLAS Collaboration, "Studies on top-quark Monte Carlo modelling for Top2016." ATL-PHYS-PUB-2016-020, 2016. 61, 62

[113] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with {NNLL} threshold resummation*, Nuclear Physics B **855** (2012) 695 – 741. 63

[114] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Physics Letters B **710** (2012) 612 – 622. 63

[115] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-level-precision physics at the tevatron: Next-to-next-to-leading order qcd corrections to $q\overline{q} \to t\overline{t}+x$*, Phys. Rev. Lett. **109** (Sep, 2012) 132001. 63

[116] M. Czakon and A. Mitov, *Nnlo corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, Journal of High Energy Physics **2012** (2012) . 63

[117] M. Czakon and A. Mitov, *Nnlo corrections to top pair production at hadron colliders: the quark-gluon reaction*, Journal of High Energy Physics **2013** (2013) . 63

[118] M. Czakon, P. Fiedler and A. Mitov, *Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_S^4)$*, Phys. Rev. Lett. **110** (Jun, 2013) 252004. 63

[119] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, [1112.5675]. 63

[120] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, 1101.0538. 63

[121] A. Martin, W. Stirling, R. Thorne and G. Watt, *Parton distributions for the lhc*, The European Physical Journal C **63** (2009) 189–285. 63, 64

[122] A. Martin, W. Stirling, R. Thorne and G. Watt, *Uncertainties on α_s in global pdf analyses and implications for predicted hadronic cross sections*, The European Physical Journal C **64** (2009) 653–680. 63, 64

[123] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (Oct, 2010) 074024. 63

[124] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, P. Nadolsky et al., *Ct10 next-to-next-to-leading order global analysis of qcd*, Phys. Rev. D **89** (Feb, 2014) 033009. 63

[125] ATLAS Collaboration, "Comparison of Monte Carlo generator predictions for bottom and charm hadrons in the decays of top quarks and the fragmentation of high p_{T} jets." ATL-PHYS-PUB-2014-008, 2014. 64

[126] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, JHEP **09** (2009) 111, [0907.4076]. 64

[127] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. **C71** (2011) 1547, [1009.2450]. 64

[128] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada et al., *Herwig++ physics and manual*, *The European Physical Journal C* **58** (2008) 639–707. 65

[129] M. Seymour and A. Siodmok, *Constraining mpi models using σ_{eff} and recent tevatron and lhc underlying event data*, *Journal of High Energy Physics* **2013** (2013) . 65

[130] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, *Phys. Rev. D* **81** (2010) 054028, [1001.5034]. 64

[131] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, *Phys. Rev. D* **83** (2011) 091503, [1103.2792]. 64

[132] ATLAS collaboration, G. Aad et al., *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J.* **C80** (2020) 47, [1909.00761]. 67

[133] ATLAS collaboration, A. Ruiz-Martinez, *The ATLAS Run-2 Trigger Menu for higher luminosities: Design, Performance and Operational Aspects*, *EPJ Web Conf.* **182** (2018) 02083. 68

[134] L. Ambroz, P. Thompson, S. H. Abidi, K. Al Khoury, S. Argyropoulos, H. Arnold et al., *Object selections for SM Higgs boson produced in association with a vector boson in which $H \to b\bar{b}$ and V decays leptonically with Run-2 data*, Tech. Rep. ATL-COM-PHYS-2019-1207, CERN, Geneva, Sep, 2019. xv, xxii, 69, 74, 76, 79, 80, 81, 82, 85, 87

[135] G. Aad, B. Abbott, D. Abbott, A. A. Abud, K. Abeling, D. Abhayasinghe et al., *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *Journal of Instrumentation* **14** (dec, 2019) P12006–P12006. 70, 71

[136] ATLAS collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J.* **C77** (2017) 490, [1603.02934]. 70

[137] ATLAS Collaboration, "Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach." ATL-PHYS-PUB-2017-022, 2017. 71

[138] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. 72

[139] ATLAS Collaboration, "Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV." ATLAS-CONF-2010-072, 2010. 72

[140] ATLAS collaboration, ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, *Eur. Phys. J.* **C76** (2016) 292, [1603.05598]. 72, 73

[141] ATLAS collaboration, G. Aad et al., *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, *Eur. Phys. J.* **C76** (2016) 295, [1512.05955]. 73

[142] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189]. 73

[143] ATLAS collaboration, G. Aad et al., *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, *Eur. Phys. J.* **C75** (2015) 303, [1412.7086]. 73, 74

[144] ATLAS Collaboration, "Tagging and suppression of pileup jets with the ATLAS detector." ATLAS-CONF-2014-018, 2014. 73, 75

[145] W. Lampl et al., "Calorimeter Clustering Algorithms: Description and Performance." ATL-LARG-PUB-2008-002, 2008. 74

[146] L. Breiman, J. Friedman, R. Olshen and C. Stone collaboration, *Classification and Regression Trees*. Chapman & Hall, 1984. 74

[147] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, *J. Comput. Syst. Sci.* **55** (1997) 119–139. 74

[148] ATLAS Collaboration, "Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC." ATL-PHYS-PUB-2015-045, 2015. 74

[149] ATLAS Collaboration, "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV." ATLAS-CONF-2017-029, 2017. 74

[150] ATLAS collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C73** (2013) 2304, [1112.6426]. 75

[151] ATLAS collaboration, ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 1703.09665. 75

[152] ATLAS Collaboration, "Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector." ATLAS-CONF-2015-029, 2015. 75

[153] ATLAS collaboration, ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, [1510.03823]. 75

[154] Unkown, "b_tagging." CERN Twiki, https://twiki.cern.ch/twiki/pub/Atlas/TapmBjet/b_tagging.png. xiv, 77

[155] ATLAS collaboration, G. Aad et al., *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J.* **C79** (2019) 970, [1907.05120]. 76, 120

[156] ATLAS Collaboration, "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run." ATL-PHYS-PUB-2017-013, 2017. 76

[157] ATLAS Collaboration, "Secondary vertex finding for jet flavour identification with the ATLAS detector." ATL-PHYS-PUB-2017-011, 2017. 76

[158] ATLAS Collaboration, "Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC." ATL-PHYS-PUB-2018-025, 2018. 76

[159] T. P. Scanlon and T. Masubuchi, *Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at* 13 *TeV with the ATLAS detector*, Tech. Rep. ATLAS-COM-CONF-2020-017, CERN, Geneva, Mar, 2020. xv, xvii, xviii, xxii, xxiii, xxiv, 76, 81, 85, 86, 87, 88, 89, 90, 91, 93, 94, 95, 114, 120, 127, 129, 134, 135, 136, 137, 142, 143, 145, 149, 150, 151, 153, 155, 157, 158

[160] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002, [1703.09665]. 78, 125

[161] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, *JHEP* **12** (2017) 024, [1708.03299]. 79, 80

[162] ATLAS collaboration, G. Aad et al., *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, *JHEP* **01** (2015) 069, [1409.6212]. 81

[163] ATLAS collaboration, M. Aaboud et al., *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Eur. Phys. J.* **C78** (2018) 903, [1802.08168]. 85

[164] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [0802.2470]. 86

[165] ATLAS Collaboration, "ATLAS Sensitivity to the Standard Model Higgs in the HW and HZ Channels at High Transverse Momenta." ATL-PHYS-PUB-2009-088, 2009. 86

[166] T. Masubuchi, J. Benitez, A. S. Bell, S. Argyropoulos, H. Arnold, Y. Amaral Coutinho et al., *Search for a Standard Model Higgs boson produced in association with a vector boson and decaying to a pair of b-quarks*, Tech. Rep. ATL-COM-PHYS-2016-1724, CERN, Geneva, Nov, 2016. 87

[167] ATLAS collaboration, M. Aaboud et al., *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV*, *Eur. Phys. J. C* **79** (2019) 639, [1902.04655]. 88

[168] N. Berger et al., *Simplified Template Cross Sections - Stage 1.1*, 1906.02754. 93

[169] ATLAS Collaboration, *Measurement of VH, $H \to b\bar{b}$ production as a function of the vector-boson transverse momentum in* 13 *TeV pp collisions with the ATLAS detector*, *JHEP* **05** (2019) 141, [1903.04618]. 93, 134

[170] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*, vol. 941. Springer, 2017, 10.1007/978-3-319-62840-0. 115, 116, 117

[171] J. Neyman and E. Pearson, *On the problem of the most efficient tests of statistical hypotheses*. Philosophical Transactions of the Royal Society of London Series A. Royal Society, 1933. 117

[172] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer, 2009. 117, 118, 119

[173] L. Breiman, *Random Forests*, *Machine Learning* **45** (2001) 5–32. 117

[174] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, *Journal of Computer and System Sciences* **55** (1997) 119 – 139. 117

[175] L. Breiman, *Prediction games and arcing algorithms*, *Neural Computation* **11** (1999) 1493–1517, [https://doi.org/10.1162/089976699300016106]. 119

[176] ATLAS Collaboration, "Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector." ATLAS-CONF-2018-001, 2018. 120, 125

[177] ATLAS Collaboration, "Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV." ATLAS-CONF-2018-006, 2018. 120, 125

[178] D. Goncalves and J. Nakamura, *Role of the Z polarization in the $H \to b\bar{b}$ measurement*, *Phys. Rev. D* **98** (2018) 093005, [1805.06385]. 120

[179] VHbb analysis group, *Object selections for SM Higgs boson produced in association with a vector boson in which $H \to bb$ and V decays leptonically with Run-2 data: Object support note for VH(bb) 2015+2016+2017 dataset publication*, . 122

[180] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ TeV." ATL-PHYS-PUB-2015-027, 2015. 122

[181] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, *Eur. Phys. J. C* **73** (2013) 2306, [1210.6210]. 125

[182] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **76** (2016) 292, [1603.05598]. 125

[183] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} =$ 13 TeV*, *Eur. Phys. J.* (2019) , [1902.04655]. 125

[184] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903, [1802.08168]. 125

[185] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, [1907.05120]. 125

[186] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **117** (2016) 182002, [1606.02625]. 126

[187] K. Cranmer, J. Pavez and G. Louppe, *Approximating Likelihood Ratios with Calibrated Discriminative Classifiers*, 1506.02169. 127

[188] K. Al Khoury, T. P. Calvet, Y. Ma, E. Schopf, S. H. Abidi, F. Ahmadov et al., *Data/Monte Carlo Modelling Studies for the Standard Model $VH, H \to b\bar{b}$ Analysis*, Tech. Rep. ATL-COM-PHYS-2019-1205, CERN, Geneva, Sep, 2019. 128, 129, 133, 134

[189] ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59, [1808.08238]. 134

[190] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, *CERN-2011-002* (2011) , [1101.0593]. 134

[191] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, *CERN-2012-002* (2012) , [1201.3084]. 134

[192] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, *CERN-2013-004* (2013) , [1307.1347]. 134

[193] ATLAS Collaboration, "Evaluation of theoretical uncertainties for simplified template cross section measurements of V-associated production of the Higgs boson." ATL-PHYS-PUB-2018-035, 2018. 134

[194] J. R. Andersen et al., *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*, 2018. 1803.07977. 134

[195] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (Feb, 2011) . xvii, 138, 139, 140, 141, 142

[196] F. A. Haight, *Handbook of the Poisson distribution [by] Frank A. Haight*. Wiley New York, 1967. 139

[197] Roy Y. and David G., *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers (2nd ed.)*. Hoboken, USA: Wiley, 2014. 139

[198] A. S. Bell, *b-Tagging and Evidence for the Standard Model $H \to b\bar{b}$ Decay with the ATLAS Experiment*, Nov, 2017. 140, 141

[199] A. L. Read, *Presentation of search results: theCLstechnique*, *Journal of Physics G: Nuclear and Particle Physics* **28** (sep, 2002) 2693–2704. 141