University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Doctoral Dissertations | Dissertations and Theses

June 2022

MEASUREMENT OF THE FIDUCIAL CROSS SECTION FOR VECTOR-BOSON-FUSION PRODUCTION OF THE HIGGS BOSON IN THE WW DECAY CHANNEL WITH THE ATLAS DETECTOR

Guy Rosin
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Elementary Particles and Fields and String Theory Commons

Recommended Citation

Rosin, Guy, "MEASUREMENT OF THE FIDUCIAL CROSS SECTION FOR VECTOR-BOSON-FUSION PRODUCTION OF THE HIGGS BOSON IN THE WW DECAY CHANNEL WITH THE ATLAS DETECTOR" (2022). *Doctoral Dissertations*. 2566.
https://doi.org/10.7275/28666281 https://scholarworks.umass.edu/dissertations_2/2566

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

MEASUREMENT OF THE FIDUCIAL CROSS SECTION FOR VECTOR-BOSON-FUSION PRODUCTION OF THE HIGGS BOSON IN THE $WW^* \to \ell\nu\ell\nu$ DECAY CHANNEL WITH THE ATLAS DETECTOR

A Dissertation Presented

by

G. ROSIN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2022

Physics

© Copyright by G. Rosin 2022

All Rights Reserved

DocuSign Envelope ID: 237497B6-1242-4A22-8933-D8EA3024250F

MEASUREMENT OF THE FIDUCIAL CROSS SECTION FOR VECTOR-BOSON-FUSION PRODUCTION OF THE HIGGS BOSON IN THE $WW^* \to \ell\nu\ell\nu$ DECAY CHANNEL WITH THE ATLAS DETECTOR

A Dissertation Presented

by

G. ROSIN

Approved as to style and content by:

DocuSigned by:
carlo dallapiccola
62E950BE63A340A...

Carlo Dallapiccola, Chair

DocuSigned by:
Verena Martinez Outschoorn
9BDE97E90CA74C3...

Verena I. Martinez Outschoorn, Member

DocuSigned by:
Grant Wilson
CFE9E7EDAEE14A2...

Grant Wilson, Member

DocuSigned by:
Scott Hertel
212E3DEA847F4CE...

Scott Hertel, Member

DocuSigned by:
Anthony Dinsmore
0580C55317A840D...

Anthony D. Dinsmore, Chair of the Faculty
Physics

ACKNOWLEDGMENTS

I would like to start off by thanking my thesis advisor Carlo Dallapicolla, who was endlessly supportive throughout my years in Umass. I would like to also thank the other professors in the Umass ATLAS group, Stephane Willocq, Ben Brau, Verena Martinez Outschoorn, and Rafael Coelho Lopes de Sa, who were great sources of knowledge and insight.

I would like to thank the scientists at Brookhaven National Lab. First, my supervisor, Stefania Stucci, who went out of her way to be a mentor and friend, and constantly sought to help me improve as a scientist and be a source of comfort when I was feeling overwhelmed. Dave Lynn who was a role model, always encouraging everyone around him, and providing honest feedback when it was necessary. Lastly, technician Russ Burns, who was always open when he was needed.

Alessandro Tricoli, who was a source of guidance, providing mentorship and wisdom with patience and understanding. Thank you to Gaetano Barone, the post-doc on the analysis, who was always available with a large breadth of knowledge on any subject we were struggling in. I want to thank the rest of the graduate students on the analysis team, Jiayi Chen, Sagar Addepalli and Chara Kitsaki. We worked well together and accomplished a lot.

My fellow graduate students in the Umass physics department: Dale Abbott, Mikayla Vessella, Daria Atkinson, Jeff Gertler, Alyssa Conway, Ebru Dogan, Chris

Nedlik, Sarah Zuraw and Sam Krishnamurthy. Everyone in the department was quick to offer help when it was needed and an open hand of friendship.

I want to thank my partner Julie, who's love and support got me through rough times in classes, in work and in life.

I want to thank my mom, who was the rock I could always rely on and my brother who had an open ear when I wanted and a laugh when I needed it.

Finally, thank you to my dad, who was my lodestar all my life. You believed in me, especially when I didn't. May you rest in peace.

ABSTRACT

MEASUREMENT OF THE FIDUCIAL CROSS SECTION FOR VECTOR-BOSON-FUSION PRODUCTION OF THE HIGGS BOSON IN THE $WW^* \to \ell\nu\ell\nu$ DECAY CHANNEL WITH THE ATLAS DETECTOR

MAY 2022

G. ROSIN

B.A., BOSTON UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Carlo Dallapiccola

This doctoral thesis presents a measurement of the fiducial and differential cross section of vector boson fusion produced Higgs boson.The measurement is taken in the $H \to WW* \to \ell\nu\ell\nu$ channel with 139 fb^{-1} of data. Proton-proton collisions from the Large Hadron Collider at $\sqrt{s} = 13$ TeV were recorded by the ATLAS detector. New analysis techniques using boosted decision trees with a statistical fit are introduced to accurately estimate backgrounds in the signal region. The fiducial cross section is measured to be 1.7 ± 0.42fb. The differential cross section was measured for 13 kinemtic variables. No significant deviations from standard model predictions are observed.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS . . . . . **iv**

ABSTRACT . . . . . **vi**

LIST OF TABLES . . . . . **xi**

LIST OF FIGURES . . . . . **xv**

CHAPTER

1. INTRODUCTION **1**

2. THEORY **3**

2.1 standard model 3

2.1.1 Particles and their interactions 3
2.1.2 Quantum Electrodynamics 6
2.1.3 Quantum Chromodynamics 9
2.1.4 Electroweak 10
2.1.5 Spontaneous Symmetry Breaking 11

2.2 LHC Physics 15

2.2.1 Higgs Physics 18
2.2.2 Whats Next? 21

3. ATLAS DETECTOR **23**

3.1 The ATLAS detector 24

4. INNER TRACKER **30**

4.1 silicon detectors 32
4.2 layout 33

5. DATA ACQUISITION **39**

5.1 stave test 43

6. MODULE QUALITY CONTROL **50**

7. ABCSTAR IRRADIATION **60**

7.1 Physics Model 62
7.2 Results 66

8. HWW VBF CROSS SECTION **71**

8.0.1 Motivations 72
8.0.2 Analysis Strategy Overview 73

9. PHYSICS OBJECTS **78**

9.1 Tracking 78
9.2 Leptons 79
9.3 Electrons 79

9.3.1 preselection 80
9.3.2 Identification 80

9.4 Muons 80
9.5 Jets 81

9.5.1 Particle Flow 81
9.5.2 Jet Selection 82
9.5.3 b-tagging 83

9.6 Missing Transverse Energy 83
9.7 Overlap removal 84

10. EVENT SELECTIONS **86**

10.1 Pre-selection 88
10.2 Signal Region Selection 92
10.3 Z+Jets Control Region 97
10.4 Top Validation Region 100
10.5 ggF Control and Validation region 103
10.6 W+jets Control Region 107

11. FIDUCIAL REGION **115**

11.1 Fiducial and inclusive phase space 115
11.2 Detector And Acceptance Corrections 118

12.BACKGROUNDS 119

12.1 Summary Of Backgrounds 119
12.2 Fake Lepton Backgrounds 122
12.3 Zjets Fake Factor Definition 124

12.3.1 Pileup Dependence 128

12.3.1.1 jet dependence 132

12.3.2 jet dependence 135
12.3.3 Fake Factor Corrections 139

12.4 Fakes estimation in the W+jets control region 145

12.4.1 Closure Test 146

13.CONTROL REGION DISCRIMINANTS 155

13.1 Gradient Boosted Decision Tree 156

13.1.1 Top WW Discriminant 157

13.2 GGF Discriminant 161
13.3 VBF Discriminant 163
13.4 N-1 Bdts 167

14.SYSTEMATIC UNCERTAINTY 170

14.1 Experimental Uncertainties 170
14.2 Theory Uncertainty 171

14.2.0.1 VBF Uncertainties 172
14.2.0.2 GGF Uncertainties 174

14.2.1 Top Uncertainties 177
14.2.2 WW Uncertainties 181

14.2.2.1 Non EWK component 181
14.2.2.2 EWK component 185

14.2.3 $Z \to \tau\tau$ Uncertainties 191

15. UNFOLDING **194**

15.1 Introduction 194
15.2 Iterative Bayesian unfolding 195

15.2.1 Statistical Uncertainty 196
15.2.2 Unfolding Bias 198
15.2.3 Coverage 198

15.3 Tikhonov Regularization 201

15.3.1 Bias Tests 202

15.4 Response Matrices 203

15.4.1 Efficiency and Purity 203

15.5 Fakes and Misses 209

15.5.1 Observables Binning 209

16. STATISTICAL TREAMENT **215**

16.1 Introduction 215
16.2 Normalization Parameters 215
16.3 likelihood procedure 217
16.4 statistical fit 219

16.4.1 Sample specific uncertainties 220
16.4.2 smoothing 221

16.5 Asimov fit results 222
16.6 Differential Statistical Analysis 234

16.6.1 Bias Test with Madgraph+Herwig signal samples 235
16.6.2 Floating background parameters scheme 236
16.6.3 Asimov Results 236

17. RESULTS **249**

17.0.1 Inclusive Results 249
17.0.2 Differential Results 250

18. CONLUSION **266**

BIBLIOGRAPHY **268**

LIST OF TABLES

Table **Page**

2.1 Run 1 ($\sqrt{s} = 8$ TeV) experimental measurements of Higgs production cross sections (left) and decay branching ratios (right) 19

7.1 Results of the fit parameters for each of the ABCStars irradiated 68

8.1 Kinematic cuts on the defined phase space regions in the analysis. Definitions of m_T and m_T^2 are given the following chapters on observables. 75

8.2 Kinematic cuts on the defined phase space regions in the analysis. The requirements that are modified with respect to the SR are highlighted in bold. 75

10.1 Background sources, their relative contributions in a preselection region and in the SR (after all cuts), the background estimation methods, and the generators used for modeling. 86

10.2 Pre-Selection cuts shared with ggF analysis 88

10.3 Cutflow in the pre-selection region. 92

10.4 list of signal region cuts 93

10.5 Cutflow of signal region cuts 93

10.6 Cutflow of Drell Yann control region cuts 97

10.7 Cutflow of Top validation region cuts 100

10.8 Cutflow of WW validation region cuts 103

10.9 ggF region definitions 106

10.10 Cutflow of ggF validation region cuts 106

10.11 Cutflow of W+Jets control region cuts 114

11.1 Definition of Fiducial Phase Space for VBF $h \to ww \to \ell\nu\ell\nu$ in the $e - \mu$ decay channel. 116

11.2 Definition of the Inclusive Phase Space for VBF H production. 117

11.3 C-factor and Acceptance factor. 118

12.1 Backgrounds to the measurement in the final state with two charged leptons and neutrinos 120

12.2 Requirements for fully identified and anti-identified leptons. 123

12.3 Fake factor application in the SR. When the dilepton trigger fires, for the fake leptons Z+jets fake factors are applied. However, for the case where a single-lepton trigger fires a di-jet and/or a Z+jets fake factors are used depending which of the leptons fires the trigger. If one of the leptons passes the "anti-id" selections and is the one which fires a single-lepton trigger, a di-jet fake factor is used, whereas for the case where the real lepton fires the single-lepton trigger, the nominal Z+jets fake factor is used instead for the fake lepton. If both leptons in the event are fakes and a single-lepton trigger is fired, a di-jet fake factor for the one which fires the trigger and for the other the nominal Z+jets fake factor are applied. The dijet trigger apply looser lepton quality criteria than those used for the "anti-id" selection. Thus, the dilepton triggers do not cause a trigger bias which is corrected for via the dijet fake factors. 125

12.4 Cutflow of fake-lepton id region. 127

12.5 Full cutflow of the Z+jets fake factor analysis with the additional requirement of at least one jet in the final state. The normalization factor that is derived in the WZ control region is applied at some cut stages and affects the WZ yield as well as the total background and fake yield. 137

12.6 Summary of fit results for fake factors as a function of number of jets or distance between the fake lepton and near by jet. 140

12.7 Events in the W+jets control region for which only the fake candidate fired a single-lepton trigger out of total events. 140

12.8 Events in the double-fakes control region in which both leptons must be "anti-ID" for which the dilepton trigger did not fire. 144

12.9 Muon and Electron CF in three bins of p_T . The nominal CF is derived using Powheg and the systematic uncertainty is obtained by comparing with MadGraph 146

12.10 Cutflow for W+jets Powheg MC events in the signal region (left) and in W+jets CR (right). Uncertainties are statistical. The statistical uncertainties from the fake factors and correction factors are also accounted. 149

12.11 Cutflow for W+jets MadGraph MC events in the signal region (left) and in W+jets CR (right). Uncertainties are statistical. The statistical uncertainties from the fake factors and correction factors are also accounted. 150

13.1 Summarized cuts for VBF signal region, Z+jets control region, top validation region, and WW validation region 156

13.2 Summarized cuts for the W+jets control region and ggF control regions (1, 2, 3) 156

13.3 bdt parameters used for the training of the bdt of top + WW vs. other processes. 158

13.4 Ranking of the variables used as input to the training of the bdt of top + WW vs. other processes. 160

13.5 bdt parameters used for the training of the bdt of ggF vs. other processes. 161

13.6 Ranking of the variables used as input to the training of the bdt trained in the ggF Control Region. 164

13.7 bdt parameters used for the training of the bdt of VBF vs. other processes. 164

13.8 Ranking of the variables used as input to the training of the bdt VBF vs Top+WW. 167

13.9 List of N-1 bdts and the removed variables 168

14.1 Summary of the experimental systematic uncertainties considered. 172

14.2 A summary of theory systematics of all processes, including the types of theory uncertainty and their calculation methods. 173

14.3 VBF theory uncertainties impact on VBF yields in the signal,control and validation regions 174

14.4 ggF theory uncertainties impact on ggF yields in the signal,control and validation regions 177

14.5 top theory uncertainties impact on $t\bar{t}$ yields in the signal,control and validation regions 180

14.6 top theory uncertainties impact on single top yields in the signal, control and validation regions 181

14.7 Breakdown of WW theory uncertainties derived from reco-level samples and their impact on the WW yields in SR and CRs/VRs. 185

14.8 Breakdown of WW theory uncertainties derived from truth-level samples and their impact on the WW yields in SR and CRs/VRs. 185

14.9 Breakdown of EWK WW theory uncertainties their impact on the WW yields in SR and CRs/VRs. 189

14.10 Breakdown of $Z \to \tau\tau$ theory uncertainties their impact on the Zjets yields in SR and CRs/VRs. 191

15.1 Binning optimized for the distributions to unfold. Inclusive bins have no upper bound 214

16.1 summary of parameters of interest in the fit 217

16.2 summary of fit regions in the inclusive case 219

16.3 Floating parameter scheme used for the Top+WW normalisation. 236

16.4 Floating parameter scheme used for the Top+WW normalization. 237

LIST OF FIGURES

Figure **Page**

2.1 List of the particles in the standard model. Top left are quarks that interact strongly, weakly, electromagnetically. Bottom left are leptons which interact weakly and electromagnetically. Bottom row are neutrinos that interact weakly. Each of the forces is mediated by one of the gauge bosons on the right 4

2.2 The sombrero potential of the Higgs field. Any of the local minima breaks the U(1) rotational symmetry 13

2.3 Parton density functions for quarks and gluons in a proton. The uncertainty of the PDF represents significant theoretical uncertainties in LHC experiments. The x-axis represents the fraction of momentum, the y axis is the scattering probability in pp collision 16

2.4 Cartoon of proton-proton collisions. The interaction is described by the collision of the specific quarks in the protons that hard scatter 16

2.5 Feynman diagrams of Higgs production modes. a is gluon-gluon fusion (ggF), b is vector boson fusion (VBF), c is Higgs plus vector boson (VH), d is Higgs plus quark anti quark pair, (qqH, most commonly ttH) 18

2.6 Cross section of Higgs production modes as a function of center of mass energy of the proton proton collision 19

2.7 Current experimental measurements of higgs cross sections of various processes, with uncertainties, relative to standard model expectations 20

2.8 Discovery of the Higgs boson (in the $\gamma\gamma$ channel) through a "bump" of excess events above the standard model background at 125 GeVs 22

3.1 Diagram of the LHC, with the SPS feeding beam. The collision points are where the detectors, ATLAS,CMS,ALICE,LHCb, are setup 25

3.2 The ATLAS detector. The middle of the detector is the proton beam line 25

3.3 A cross-sectional diagram of the ATLAS detector. The inner most layer, ID, measures tracks of particles. The EM calorimeters causes photons and electrons to shower, the tile calorimeter causes hadrons like protons and neutrons to shower, which allows us to identify them. Muons reach the Muon Spectrometer in the outer layer. Neutrinos (and potential BSM particles) are invisible to the detector and are simply missing energy 26

3.4 ATLAS Inner Detector 27

3.5 EM liquid argon and hadronic tile carlorimeters 28

3.6 ATLAS muon spectrometer 29

4.1 Expected shape of the Inner Tracker. The fine inner layer of detectors are called pixels, the coarser outer layer of detectors are called strips. The barrel region will have rectangular staves placed in cylindrical layers, the endcap region will have detectors placed in a petal geometry. 30

4.2 The expected luminosity of the High Luminosity LHC (right) compared with the current luminosity (left). The HL-LHC luminosty is expected to be 7-10 times larger 31

4.3 Expected timeline of LHC upgrades, leading to the High Luminosity LHC (pre-pandemic expectations) 31

4.4 Example of doped silicon lattice. N doped lattice contains an element with one more valence electron, p doped contains elements with one less valence electron 32

4.5 Ionizing particle through a silicon detector. The ionizing particle creates an electron hole pair in the depletion region. The electrons drift onto the charge collection strips at the top. 33

4.6 ITK strips and pixels layout 34

4.7 Barrel staves and endcap petals. The staves are laid out in a rectangular geometry, the petals are laid out like a flower in the endcaps. Both use the same module readout systems and same support structures. 34

4.8 ITk Strips Barrel Stave 35

4.9 End of substructure (EOS) card for staves 35

4.10 Module power board, contains the DC-DC converter, HV multiplexer and AMAC control chip 36

4.11 High voltage bus scheme for modules on a stave, where each four modules share a bus and can be turned on and off with the HV mux 36

4.12 Silicon Strip Module diagram. The hybrid and power board are glued onto the silicon strips, the readout chips ABC and HCC, are then glued on top of the hybrids 37

4.13 ABCStar schematics. Data is digitized with a one bit discriminator, then buffered in the L0 buffer 38

4.14 data flow from modules to stave to EOS card to data acquisition system. The EOS card collects data, provides the clock (TTC) and power, and reads out the module sensors (DCS) 38

5.1 Threshold scan 40

5.2 strobe delay values for each of the ten chips in a hybrid. The chosen delay values 40

5.3 Strobe delay for a single channel 41

5.4 Three point gain test for a single channel 41

5.5 effect of trim range on DAQ tests. The Vt50 (top plot) of all the channels are brought into alignment. The gain (middle plots) and noise of the channels (bottom plot) remain unaffected 42

5.6 cold chuck for module testing 43

5.7 first full electrical stave produced at BNL. The end of substructure card is on the bottom , connected to the fiber optic connection. The microscope was used to find broken or mis-bonded wire bonds 44

5.8 Stave DCS setup. Data from the fiber optics is read out by a GLIB FPGA 44

5.9 EOS card with GBTx chip that drives the optical signal 45

5.10 Comparison of output noise between stave and single module tests 45

5.11 Comparison of Input noise between stave and single module tests 46

5.12 Comparison of the gain between stave and single module tests 46

5.13 Noise increase in the stave test relative to single module tests 47

5.14 Confocal microscope used for module metrology 48

5.15 Results of module metrology, module bowing is within 100 μm 49

6.1 BNL/Harvard four module cold box 51

6.2 Schematic of electrical connections for the four module box 52

6.3 Outline of Software Framework for QC Tests 53

6.4 Setup for cold box DAQ 54

6.5 Temperature sensor outputs from the first DAQ test 54

6.6 Top plot is the first cycle where the trim range was run, the second plot is the second thermal cycle where only strobe delay and three point gain were run. The times each test took is on the x axis. SD is strobe delay, 3PG is thee point gain, 32T-3PG is three point gain with low statistics. 55

6.7 Cooldown and warm up rate for the four module box with the SPS chiller. Warmup is slightly faster as expected but its close to 20 minutes each. 55

6.8 Sensor recordings from thermal cycle 56

6.9 Vacuum pressure vs. temperature for vacuum chucks in the four module box. The path on the left is the plot as the box cools down, the path on the right is the plot as the box warms up 57

6.10 DAQ test result for four module box with all four modules and the full thermal cycle at the warm and cool part of the cycle. The top plot is the Vt50 of each channel, the two middle plots is the estimate of the gain, the bottom plot is input noise 58

6.11 Distribution of the input noise for all four modules in the warm and cold part of the thermal cycle. Room temperature expectations are 600-700 ENC, so the noise in the warm up cycle 59

6.12 Distribution of the output noise for all four modules in the warm and cold part of the thermal cycle 59

6.13 Average Input noise for all modules for each of the ten thermal cycles 59

7.1 Bump in the digital current vs. total ionizing dose 60

7.2 Setup of the four ABCStar boxes relative to the Co-60 source. 61

7.3 Orientation of each ABCStar chip relative to the source (The HCCStar is a seperate co-operating experiment) 62

7.4 Schematics of the electrical connections between the boxes containing the ABCStar chips, DAQ readout, thermo-electric chiller (all in the source room) with the controls outside the source room 63

7.5 Diagram of parasitic transistors in a CMOSFET on an isolation layer 64

7.6 Diagram of the process of positive states trapped in the interface of silicon/ silicon oxide isolation, creating parasitic transistors 65

7.7 Results of RAL annealing studies. The bump is the initial TID bump, the other lines show the bump doesn't return 66

7.8 ABCStar irradation results of the digital leakage current vs. total ionizing dose 67

7.9 ABC130 irradation results from 2016 (bottom) and 2017 (top) campaigns 67

7.10 Model fit to the digital Leakage current vs. total ionizing dose data 68

7.11 ABC130 irradiation results from 2016 (bottom) and 2017 (top) campaigns 69

7.12 Leakage current fall for different temperatures 69

7.13 TID Fall time vs. Temperature 70

8.1 Feynamn diagrams of Higgs production to WW decay 73

10.1 Distribution of observables after preselection with statistical error 89

10.2 Distribution of observables after preselection with statistical error 90

10.3 Distribution of observables after preselection with statistical error 91

10.4 Distribution of observables after signal region cuts with statistical error 94

10.5 Distribution of observables after signal region cuts with statistical error 95

10.6 Distribution of observables after signal region cuts with statistical error 96

10.7 Distribution of observables after DY region cuts with statistical error 98

10.8 Distribution of observables after DY region cuts with statistical error 99

10.9 Distribution of observables after Top validation region cuts with statistical error 101

10.10 Distribution of observables after Top validation region cuts with statistical error 102

10.11 Distribution of observables after WW validation region cuts with statistical error 104

10.12 Distribution of observables after WW validation region cuts with statistical error 105

10.13 Distribution of observables after ggF control region cuts with statistical error 107

10.14 Distribution of observables after ggF control region cuts with statistical error 108

10.15 Distribution of observables after ggF control region cuts with statistical error 109

10.16 Distribution of observables after ggF validation region 2 region cuts with statistical error 110

10.17 Distribution of observables after ggF validation region 2 region cuts with statistical error 111

10.18 Distribution of observables after ggF validation region 2 region cuts with statistical error 112

10.19 Distribution of observables after ggF validation region 3 region cuts with statistical error 113

10.20 Distribution of observables after ggF validation region 3 region cuts with statistical error 113

12.1 Feynamn diagram for Drell Yan process with associated QCD jets 121

12.2 Distributions of the invariant mass of the reconstructed Z-boson candidate. The shape of the distribution agrees nicely between data and MC. At the stage of Z-boson identification (left), the normalization disagrees slightly. After applying a cut on the transverse mass $m_T^W < 50$ GeV(right) the normalization agrees nicely. Normalization factors are not applied. The stacked histograms are background MC processes not including Z+fake. The measured data is shown in black points. The blue points are the data-driven Z+fake estimate. They are calculated by taking the difference between data and the stacked MC processes. The agreement between data and MC is seen by comparing the blue fake estimate to green Z+jets MC. 126

12.3 Distributions in the WZCR show good modeling. A cut of $m_T > 50$ GeV(left) and the requirement that the fake candidate lepton is ID (right) are applied for the WZ CR. 127

12.4 Transverse momentum and pseudorapidity distributions of the fake candidate. The top four plots are electron fake candidates and the bottom four are muon fake candidates. Both the ID (left) and the AntiID (right) selections are shown. The stacked histograms are background, i.e. non-Z+jets MC processes. The difference between the data yield and the stacked histograms constitutes the estimate for Z+jets. The fake estimate is also shown in blue markers and is compared to the non-stacked Z+jets MC prediction. 129

12.5 Fake factors derived in the three-lepton selection. The top two plots show the fake factor (left) and its relative uncertainties (right) as a function of p_T for electrons. The middle two plots show the muon fake factors in the low-η region and the bottom two plots the same in the high-η region. Stat denotes the statistical uncertainties on the fake factors while EW denotes the uncertainty due to the electroweak backgrounds that enter to the Z+jets fake factor estimate. 130

12.6 The pileup (μ) distributions for each campaign separately and for the full Run 2 data. For c16a, the average number of interactions per bunch crossing is used, while for c16d and c16e the actual number of interactions per bunch crossing is used 131

12.7 Electron fake factors in data (left) and MC (right) as functions of the pile-up for each p_T bin, while integrated in η. Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainies in the data-driven metod, the error bands represent the total uncertainties of the data-driven technique. A polynomial fit is performed and the fit parameters are displayed in the insets to assess any possible pileup dependence. 133

12.8 Muon fake Factors in data (left) and MC (right) as functions of the pile-up for each p_T bin, while integrated in η. Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainies in the data-driven metod, the error bands represent the total uncertainties of the data-driven technique. A polynomial fit is performed and the fit parameters are displayed in the insets to assess any possible pileup dependence. 134

12.9 Number of jets in selected Z+jets events in Monte Carlo and in data. Their ratios are used to calculate the weights that correct the MC distributions for jet production mismodeling. 135

12.10 Leading lepton, subleading lepton, and leading jet p_T distributions after the Z-tagging cut (where the fake lepton is selected). Left:before the MC-reweighting, Right: after the MC-reweighting 136

12.11 Muon fake factors as a function of the number of jets estimated with the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique. 137

12.12 Electron fake factors as a function of the number of jets estimated with the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique. 138

12.13 Electron fake factors as a function of the distance between the "fake" lepton and its closest jet, in the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique. 138

12.14 Muon fake factors as a function of the distance between the "fake" lepton and its closest jet, in the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique. 139

12.15 Electron (top) and Muon (bottom) fake factors with trigger bias. The effects of the trigger biases are expected to be different in each campaign due to differences in the trigger configurations and updates in the online reconstruction and calibration of the leptons, thus they are estimated separately for each year. Only statistical uncertainties are shown. 141

12.16 Kinematic distributions in the double-fakes control region. Data points (black markers) describe events with two "anti-id" leptons. Distributions correspond to MC events with two real leptons passing the "anti-id" selection. Their difference (blue markers) represent the estimate of the double fakes. Fake factors are not applied in these plots. 142

12.17 Kinematic distributions in the double-fakes control region. Data points (black markers) describe events with two "anti-id" leptons. Distributions correspond to MC events with two real leptons passing the "anti-id" selection. Their difference (blue markers) represent the estimate of the double fakes. Fake factors are not applied in these plots. 143

12.18 Correction factors derived from MC sample with the Powheg generator, electrons (left) muons (right) 145

12.19 Distributions in the fakes control region. The data (black markers) are shown together with electroweak processes in MC (blue histograms). Their difference (blue markers) represents the estimate of the fake background. No fake factors are applied. 147

12.20 Distributions in the fakes control region. The data (black markers) are shown together with electroweak processes in MC (blue histograms). Their difference (blue markers) represents the estimate of the fake background. No fake factors are applied. 148

12.21 Distributions at preselection level for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs applied (red) using Powheg MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable. 151

12.22 Distributions after the full selection for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs applied (red) using Powheg MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable. 152

12.23 Distributions after the full selection for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs (red) using MadGraph MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable. 153

12.24 Distributions at preselection level for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CRs (red) using MadGraph MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable. 154

13.1 Correlations between in the input variables for the singal and background samples for WW+Top bdt 159

13.2 Signal and background spectrums of the input variables to WW+Top bdt 159

13.3 (Left) comparison of the train and test samples for the WW+Top bdt. The discrepancy between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process 160

13.4 Correlations between in the input variables for the singal and background samples for ggF bdt 162

13.5 Signal and background spectrums of the input variables to ggF bdt 162

13.6 (Left) comparison of the train and test samples for the ggF bdt. The discrepency between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process 163

13.7 Correlations between in the input variables for the singal and background samples for vbf bdt 165

13.8 Signal and background spectrums of the input variables to vbf bdt 166

13.9 (Left) comparison of the train and test samples for the vbf bdt. The discrepency between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process 166

13.10 Signal to background ratio in each bin of the vbf bdt. The binning is the one used in the inclusive fit: 0.5-0.7,0.7-0.86,0.86-0.94,0.94-1.0 167

13.11 bdt output distribution of test and train data for vbf signal, top+ww background for $\Delta\phi_{jj}$(left) and $\Delta\phi_{ll}$(right) 168

13.12 bdt output distribution of test and train data for vbf signal, top+ww background for Δy_{jj}(left) and Δy_{ll}(right) 169

13.13 bdt output distribution of test and train data for vbf signal, top+ww background for $\Delta jetpt_{jj}$(left) and Δm_{ll}(right) 169

14.1 Up and down variations on the VBF bdt output distribution for the considered theoretical uncertainties on the VBF signal. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$. 175

14.2 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the VBF signal. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 176

14.3 The shape difference of transverse momentum of the Higgs plus two leading jets system before and after the cut associated with a third jet. 177

14.4 Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the ggF spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$. 178

14.5 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the ggF spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 179

14.6 Up and down variations on the topWW bdt output distribution for the considered theoretical uncertainties on the single top spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 182

14.7 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the $t\bar{t}$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 183

14.8 Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the $t\bar{t}$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$. 184

14.9 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the WW spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 186

14.10 Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the WW spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$. 187

14.11 Up and down variations on the top+WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 188

14.12 Up and down variations on the top+WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 189

14.13 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 190

14.14 Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the $Z \to \tau\tau$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$. 192

14.15 Up and down variations on the ggF BDT output distribution for the considered theoretical uncertainties on the $Z \to \tau\tau$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown in the ggF control region. 193

15.1 Statistical uncertainty propagated through unfolding vs. bayesian Iteration in Higgs Pt 197

15.2 Truth Cross Section vs. Unfolded Cross Section for three Bayesian iterations. 68% of the data falls within the bands 199

15.3 Profile of Systematic Bias vs Variation in Truth Content. The flatter the curve of the profile the less unfolding bias. The optimal parameter is taken when there's no significant bias improvements from more iterations. 200

15.4 Examination of coverage in unfolding for the Baysian unfolding. 201

15.5 Mean cross section bias, from MC toy tests on samples with different true cross section, as a function of the several observables to be unfolded. The red and black continuous (dotted) lines show cross section bias for the bin-by-bin and matrix methods when the truth cross-section is varied within the one sigma of the expected statistical uncertainty. The red and black shaded areas show, respectively, the median statistical uncertainty. 204

15.6 Plots of bias vs. simulated cross section for matrix inversion and Tikhonov Regularization for bins of M_{jj} 205

15.7 Examination of coverage for matrix inversion and Tikhonov Regularization for M_{jj} 206

15.8 Migration matrices for each observable on the left, with the corresponding number of non-fiducial fakes on the right. 207

15.9 Purity and efficiency per bin for each observable to be unfolded. 208

15.10 "Fake" events, defined as reco-level events in the SR falling outside of the fiducial region at truth, as a function of kinematic observables at reco-level. 210

15.11 "Fake" events, defined as reco-level events in the SR falling outside of the fiducial region at truth, as a function of kinematic observables at reco-level. 211

15.12 "Miss" events, defined as truth-level events in the fiducial region falling outside of the signal region at reco, as a function of kinematic observables at truth-level. 212

15.13 "Miss" events, defined as truth-level events in the fiducial region falling outside of the signal region at reco, as a function of kinematic observables at truth-level. 213

16.1 Profile likelihood scan of σ^{fid}. Dashed curve shows the scan with only statistical uncertainty; solid curve shows the scan with all nuisance parameters floating. 224

16.2 Norm parameters and uncertainties from inclusive Asimov fit. Inner error bars are stat only errors. Mu wwTop bin0 is the norm parameter in the signal region for Top+WW, wwTop3 is the parameter in the ggF control region 225

16.3 Pull plot of the nuisance parameters from the inclusive fit. 226

16.4 Correlation matrix of the floating parameters from the inclusive fit. 227

16.5 Impact of the highest ranked systematics on the fiducial cross section 228

16.6 Impact of the highest ranked systematics on the ggF normalization parameter 229

16.7 Impact of the highest ranked systematics on the Zjets normalization parameter 230

16.8 Impact of the highest ranked systematics on the Top+WW ggFCR normalization parameter 231

16.9 Impact of the highest ranked systematics on the Top+WW SR normalization parameter 232

16.10 Phase space distribution of samples in the inclusive asimov fit. 233

16.11 Fit results (green) using the nominal model with MGH7 Asimov dataset, and comparison to the MGH7 (red) and nominal (black) predictions. The fit result (green) error represents the MC statistical uncertainty. 235

16.12 Post-fit differential cross section as a function of M_{jj} on an Azimov data set and comparison with theoretical expectations. 238

16.13 Left: Fit result of the normalization parameters in the measurement of the differential cross section as a function of M_{jj} with statistics-only and statistics+systematic uncertainties. Right: Correlation between normalization parameters (including the per-bin signal cross section) in the the fit that includes statistics and systematic uncertainties. 239

16.14 Distribution for Signal Region 1 (SRVBF) for each M_{jj} differential bin 240

16.15 Distribution for Signal Region 2 (CRTop) for each M_{jj} differential bin 241

16.16 pull plots of all the nuisance parameters in the M_{jj} fit. 242

16.17 Ranking of systematic impacts on Mjj cross section bin 0. 243

16.18 Ranking of systematic impacts on Mjj cross section bin 1. 244

16.19 Ranking of systematic impacts on Mjj cross section bin 2. 245

16.20 Ranking of systematic impacts on Mjj cross section bin 3. 246

16.21 Ranking of systematic impacts on Mjj cross section bin 4. 247

16.22 Ranking of systematic impacts on Mjj cross section bin 5. 248

17.1 Distribution of the two control regions and signal regions after fit. Errors are stat only in this case. Top to bottom, CRGGF, CRZjets, Signal Region 2, Signal Region 1 250

17.2 (top)Normalization values for background MC from inclusive fit, (bottom) Correlation matrix of fit parameters 251

17.3 Nuisance parameters, central value and error 252

17.4 Mjj Full stat+sys fit to data 253

17.5 DYll Full stat+sys fit to data 254

17.6 DPhijj Full stat+sys fit to data 255

17.7 Higgs Pt Full stat+sys fit to data 256

17.8 Jet0 Pt Full stat+sys fit to data 257

17.9 jet1 pt Full stat+sys fit to data 258

17.10 lep0 pt Full stat+sys fit to data 259

17.11 lep1 pt Full stat+sys fit to data 260

17.12 Ptll Full stat+sys fit to data 261

17.13 DPhill Full stat+sys fit to data 262

17.14 DYjj Full stat+sys fit to data 263

17.15 Mll Full stat+sys fit to data 264

17.16 Cos(θ*) Full stat+sys fit to data 265

CHAPTER 1

INTRODUCTION

The Large Hadron Collider (LHC) is a proton-proton accelerator located at CERN in Geneva, Switzerland. It is one of the largest science endeavors in human history, in terms of number of scientists involved, number of countries in the collaboration, physical size of the apparatus, and total funding. The LHC is a 27 kilometer ring on the border of Switzerland and France, that accelerates protons to a center of mass energy of 14 TeV. The accelerated protons collide at four points in the beam line with each collision point surrounded by detectors. The detectors give us a close view of the interactions between the fundamental particles, and will hopefully allow us to find new physics that explain many still mysterious facets of the universe.

This thesis presents work on the ATLAS detector, the detector placed at the first collision point in the LHC. The first chapter of this thesis presents background on the physics of the standard model and potentially new physics we hope to understand from LHC data. The second chapter presents an overview of the hardware of the ATLAS detector and plans for upgrades. The next four chapters present the work I did in Brookhaven National Lab on research and development of silicon detectors for the upgraded Inner Tracker in the ATLAS detector. It will focus on data acquisition tests and quality control of silicon modules. The bulk of the thesis, chapters 6 onwards, will present a measurement of the fiducial cross section of a Higgs boson, produced via vector boson fusion, decaying into WW and then a final state of two opposite sign leptons using run 2 data from the ATLAS detector. The $H \to WW \to \ell\nu\ell\nu$ analysis measures both the total fiducial cross section and differential cross section

as a function of thirteen important observables in the process. The uncertainty on the final measurement is 24%, a decrease from the 50% uncertainty of the run 1 measurement. While this analysis did not find disagreement with the standard model expectations, it facilitates large combined measurements with other Higgs analysis groups and paves the way for further studies using the increased data from future LHC runs [7].

CHAPTER 2

THEORY

2.1 standard model

The standard model of particle physics is one of the most successful theories in modern physics, confirmed with increasing accuracy with every experiment conducted to date. The standard model describes all the known particles in the universe and three forces, electromagnetic, weak and strong force, that allow particles to interact. Despite the repeated success of the standard model in experiment, there's still many mysteries that it cannot explain, such as how gravity fits into the theory or the origins of dark matter. Further, the standard model has many "ad hoc" components, such as the masses of particles, which have no theoretical motivations but are simply experimentally measured parameters. For these reasons scientists are searching for a deeper theory beyond the standard model to explain these mysteries.

2.1.1 Particles and their interactions

Particles defined in the standard model are smallest possible unit of matter. They are point like, indivisible and have no internal structure. The most commonly known such particles are electrons and photons. Protons and neutrons, which make up the nucleus of atoms are composite matter made up of other fundamental particles. Particles are distinguished by their unique quantum numbers and have an anti-particle counterpart with the same mass but opposite charge and quantum numbers.

One of the quantum numbers that a particle takes on is intrinsic angular momentum or "spin".Particles that have integer value of spin, 0, 1, 2, etc, are called

"Bosons" and particles that have half integer values, $\frac{1}{2}$, $\frac{3}{2}$, etc, are called "Fermions." Interactions between standard model particles are mediated by exchange of bosons between the interacting particles.

Fermions make up all the observable matter in the universe, they group together to create protons, neutrons, atoms and so on. Fermions are split into two types, leptons and quarks. Leptons can interact through the electromagnetic and weak force, and they have charge of $1.6 \cdot 10^{-19}$ Couloumbs, or equivalently 1e. By convention leptons have negative charge and anti-leptons have positive charge. Each lepton has a neutrally charged neutrino associated with it, the electron and electron neutrino for example. There are three generations of leptons with increasing mass, summarized in figure 2.1.

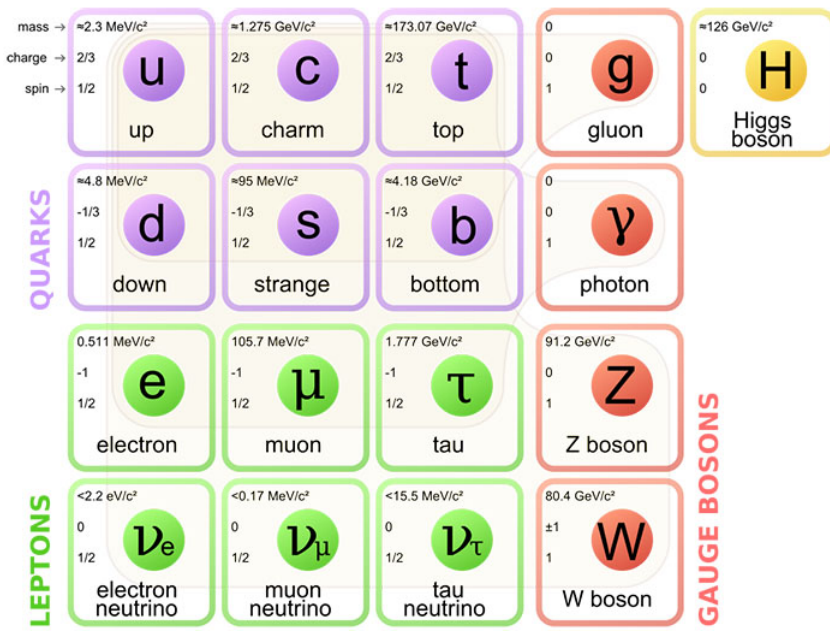


Figure 2.1. List of the particles in the standard model. Top left are quarks that interact strongly, weakly, electromagnetically. Bottom left are leptons which interact weakly and electromagnetically. Bottom row are neutrinos that interact weakly. Each of the forces is mediated by one of the gauge bosons on the right

The other category of fermions are quarks.Quarks can interact electromagnetically, and through the weak and strong force. As with leptons, quarks are divided into

generations. Each generation has two quarks, one positively charged with a charge of either $+\frac{2}{3}$e or -$\frac{1}{3}$e. Stable matter is made up of the lightest quarks from the first generation, the "up" and "down" quark. Protons are a composite of two up quarks and one down quark, giving them a charge of +1e, neutrons consist of two down quarks and one up quark, giving them zero charge total. Why all quarks in nature are found in composite particles (such as protons and neutrons) and never by themselves is one of the yet unexplained phenomena in physics.

There are four gauge bosons and one scalar boson, the Higgs, in the standard model. Each of the gauge boson corresponds to one of the three fundamental forces of nature. As stated before, gravity is notably absent from the standard model. The electromagnetic interaction between charged particles in mediated by an exchange of a photon, which has itself has no charge and spin 1. The electromagnetic force is responsible for keeping electrons in orbit around atomic nuclei, and bonding atoms to form molecules. The strong force mediates the interaction between quarks through a gluon, and is responsible for bounding the quarks of protons and neutrons together in the atomic nucleus. The weak force is mediated by a neutral Z boson and a charged W^+/W^- boson. The Z boson acts similarly to a photon, mediating weak interactions like scattering between particles. The W boson carries a charge so allows particles to decay into other particles allowed by conservation laws. The Higgs boson gives mass to other particles through spontaneous symmetry breaking, described in the following sections.

Photons are massless, spin 1 vector bosons that couple to all particles with an electric charge, leptons, quarks and the W boson. Neutrinos and the other bosons are neutral.

Gluons mediate the strong force, are again massless spin 1 vector bosons, that couple to particles with a "color charge" as an analogue to the electric charge. While electric charge can take on two values, which are arbitrarily assigned positive or

negative, color charge can take on three distinct values, which are assigned as red green or blue. Quarks and gluons carry a color charge, allowing gluons to couple to themselves and other quarks. Gluons and quarks are never isolated, but are confined to a bound state of other quarks that is overall color neutral. That is there exists a Red, Green and Blue quark in the bound state ("white") or a color quark with its anti color.

The Z boson acts similar to a photon, allowing for particles with a weak interaction, which are all quarks, leptons and neutrinos to interact. However the Z and W bosons have a mass (90 and 80 GeV respectively), which means the interaction cross section is much smaller at low energies compared to photons, hence "weak force". The W boson have either a positive or negative electric charge, which allows flavor violating interactions such as a muon to electron decay [2].

2.1.2 Quantum Electrodynamics

Quantum electrodynamics describes the interactions of charged particles and photons. To derive the equations of motions, one begins with the Lagrangian for a free charged particle of mass m:

$$\mathcal{L} = i\psi\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{2.1}$$

where ψ is the dirac spinor, and γ^{μ} is a four vector of gamma matrices from the dirac equation. The gamma matrices are the following:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$
$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.2}$$

And the equation uses the Einstein summation notation where $\gamma^\mu \partial_\mu \psi$ implies a sum over all the indices μ

The Lagrangian in the standard model is required to be gauge invariant, meaning the dynamics of the system are identical under gauge transformation. In the case of QED, the solution ψ is invariant under a phase rotation of angle θ

$$\psi \to e^{i\theta}\psi \tag{2.3}$$

More generally, the angle could have a spatial or time dependence, so that

$$\psi(x) \to e^{i\theta(x)}\psi(x) \tag{2.4}$$

Plug in a phase rotation into the Lagrangian we have:

$$\partial_\mu(e^{i\theta(x)}\psi) = e^{i\theta}\partial_\mu\psi + ie^{i\theta}\partial_\mu\theta \tag{2.5}$$

Which is not invariant. A covariant derivative is defined that transforms correctly under rotation such that

$$D_\mu\psi \to e^{i\theta}D_\mu\psi \tag{2.6}$$

To make this work, another field A_μ is added into the derivative term to cancel out the extra term from the derivative.

$$D_\mu = \partial_\mu - ieA_\mu \tag{2.7}$$

Where e is the electric charge, added with hindsight to clean up the final equation. The Lagrangian becomes

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu\psi A_\mu \tag{2.8}$$

Finally, the kinetic energy of this field can be found from from electromagnetic tensor $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, so the Lagrangian becomes

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.9}$$

Which is the full QED Lagrangian. If A_μ had a mass m, the Lagrangian would no longer be invariant under gauge transformation, so we conclude the new field is massless and couples to the charged particles with coupling strength proportional to the electric charge, which exactly describes the photon field. Starting from a free particle, the requirement that the dynamics of the particle be invariant under gauge transformation naturally derives the existence of photons that couple to that particle, a great triumph of QED.

To find the equations of motion, the Euler-Lagrange equation is applied to the QED Lagrangian, for the electromagnetic field A_μ

$$\partial_\nu \frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)} = \frac{\partial\mathcal{L}}{\partial A_\mu} \tag{2.10}$$

Evaluating the derivatives:

$$\partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} = \partial_\nu F^{\mu\nu} \tag{2.11}$$

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = e\bar{\psi}\gamma^\mu\psi \tag{2.12}$$

So then the equation of motion is

$$\partial_\nu F^{\mu\nu} = e\bar{\psi}\gamma^\mu\psi \tag{2.13}$$

Which is a quantized version of the covariant form of Maxwell's equations from classical electromagnetism.

2.1.3 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong interaction, so named because the strong charge comes in three possible values, Red, Green, Blue (which are just labels analogous to positive and negative labels for electric charge and have nothing to do with the classical idea of color). The QCD Lagrangian for a free fermion of mass m is:

$$\mathcal{L} = \bar{q}_j(i\gamma_\mu\partial_\mu - m)q_j \tag{2.14}$$

Where j denotes one of three possible colors, so the total wave function is then the combination of these $q(x) = q_{Red}q_{Blue}q_{Green}$. Gauge invariance is required in all three colors so that the gauge transformation is

$$q \rightarrow q_{red}e^{i\alpha(x)} + q_{Blue} * e^{i\beta(x)} + q_{green}e^{i\gamma(x)} + \text{mixing terms} \tag{2.15}$$

Which is a total of 8 possible vector fields. This can be written in terms of matrices T_a, where aranges from 0-8 so that the gauge transformation reduces to

$$q(x) \rightarrow e^{i\alpha^a T^a} q(x) \tag{2.16}$$

Take the taylor expansion of the rotation, so we consider only the linear term, so the rotation becomes, as with the QED case

$$\begin{aligned} q(x) &\rightarrow [1 + i\alpha^a T^a] q(x) \\ \partial_\mu q(x) &\rightarrow [1 + i\alpha^a T^a] \partial_\mu q(x) + iT^a q \partial_\mu \alpha^a \end{aligned} \tag{2.17}$$

Which again breaks the gauge invariance of the lagrangian, which is again fixed by adding a new field, in this case eight new fields for each T^a,G^a_μ which gives the QCD Lagrangian:

$$\mathcal{L} = \bar{q}(i\gamma^\mu \partial_\mu - m)q - g\bar{q}\gamma^\mu T_a q G^a_\mu - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} \tag{2.18}$$

Where g represents the coupling constant of the strong force, and $G^a_{\mu\nu}$ is the strong force field tensor, analogous to the electromagnetic field tensor $F^{\mu\nu}$, is defined as

$$G^a_{\mu\nu} = \partial_\mu A^a_\mu + g \cdot f^{abc} A^b_\mu A^c_\nu \tag{2.19}$$

Where f^{abc} is the structure constant of QCD, and A^a_μ are the gluon fields, where μ is the four vector index and a , b, c are indices that run 1 through 8. For the same reason as QCD, the gluon fields are massless but differently, they have a color charge therefore couple to themselves.

2.1.4 Electroweak

The weak interaction is mediated by two boson, W and Z, where the Z boson is neutral and the W boson carries a charge. Unlike QED and QCD, these bosons

are massive, which is explained by symmetry breaking of the weak force. The weak interaction has been unified with the electromagnetic interaction, giving a singular "electroweak" interaction described below.

Electroweak interaction is unified as SU(2)× U(1) gauge group (from empirical observations), where QED was U(1) group and QCD was SU(3), which gives rise to thee gauge fields W_1, W_2, W_3 from the SU(2) group called the weak isospin group and B boson from the U(1) group called the weak hypercharge. The fields are at this point massless before symmetry breaking, and the weak Lagrangian can be formed using the following formulation, for the kinetic energy of the fields

$$\mathcal{L} = -\frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - -\frac{1}{4}B^{\mu\nu}B_s\mu\nu \tag{2.20}$$

and for the kinetic term of the fermions, we again define a covariant derivative with the fields

$$D_\mu = \partial_\mu - i\frac{g}{2}YB_\mu - i\frac{g}{2}T_jW_\mu^j \tag{2.21}$$

Where Y is the weak hypercharge and T_j are components of the weak isospin. The electromagnetic charge Q, arises from linear combination of the T_3 component of the isospin and weak hypercharge, $Q = T_3 + \frac{1}{2}Y$. This gives us the kinetic term for fermions in the Lagrangian:

$$\mathcal{L} = \bar{Q}_j i\gamma^\mu D_\mu Q_j + \bar{\psi}i\gamma^\mu D_\mu\psi \tag{2.22}$$

2.1.5 Spontaneous Symmetry Breaking

If we leave the lagrangian at that, the force carrier bosons of the weak force, Z and W, would be massless for the same arguments as in QED. Experimentally we know thats not the case. In fact they are the third and fourth most massive particles in the standard model respectively. To solve this dilemma we introduce Spontaneous symmetry breaking, which will bring about the Higgs boson.

Going back to the U(1) Lagrangian of field ϕ, with the same gauge invariance and covariant derivative introduced in the section on QED, we have this results

$$\mathcal{L} = (\partial^\mu + ieA^\mu)\phi^*(\partial^\mu - ieA^\mu)\phi - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2 - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.23}$$

Where if $\mu^2 > 0$ the Lagragian becomes the QED Lagrangian for a scalar charged particle. If $\mu^2 < 0$, then the potential V($\phi^*\phi$)=$\mu^2\phi^*\phi - \lambda(\phi^*\phi)^2$ takes the form in figure 2.2, which is sometimes called the sombrero potential. Any one of the minima in the bottom of the sombrero breaks the potential rotational symmetry. A perturbative expansion around the minima of the new potential gives:

$$\phi(x) = \sqrt{\frac{1}{2}(\nu + \eta(x) + i\xi(x))} \tag{2.24}$$

Substituting this intro the Lagrangian gives

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\xi)^2 + \frac{1}{2}(\partial_\mu\eta)^2 - \nu^2\lambda\eta^2 + \frac{1}{2}e^2\nu_\mu^2 - e\nu A_\mu\partial^\mu\xi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.25}$$

From the lagrangian three particles come out, ξ, a vector boson A_μ with mass $m_A = e\nu$ and a scalar field η with mass $\sqrt{2\lambda\nu^2}$. The particle ξ however corresponds to an extra degree of freedom which can be removed with a gauge transformation:

$$\phi \to \sqrt{\frac{1}{2}(\nu + h(x))e^{i\frac{\theta(x)}{\nu}}} \tag{2.26}$$

$$A_\mu \to A_\mu + \frac{1}{e\nu}\partial_\mu\theta \tag{2.27}$$

Which makes the ξ field disappear in the langrangian.

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)^2 - \lambda\nu^2h^2 + \frac{1}{2}e^2\nu^2A_\mu^2 - \lambda\nu h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}e^2A_\mu^2h^2 + \nu e^2A_\mu^2h - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.28}$$

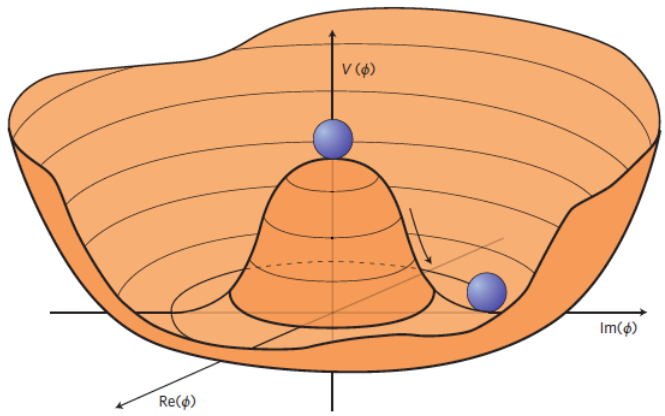


Figure 2.2. The sombrero potential of the Higgs field. Any of the local minima breaks the U(1) rotational symmetry

So now the lagrangian contains a massive vector boson A_μ. In addition a massive scalar boson h comes out, which will correspond to the Higgs boson.

We can now take this formalism and apply it to the electroweak SU(2)×U(1) Lagrangian. Define ϕ as a doublet of complex scalar fields:

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.29}$$

with the SU(2) *times* U(1) lagrangian

$$\mathcal{L} = (D^\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \tag{2.30}$$

and then define the following covariant derivatives

$$D_\mu = \partial_\mu + ig \frac{\tau_a}{2} W^a_\mu \tag{2.31}$$

As before, doing a taylor expansion around the rotation

$$\phi \to (1 + i\alpha(x) \frac{\tau}{2}) \phi \tag{2.32}$$

which gives the potential

$$V(\phi) = mu^2\phi^\dagger\phi + +\lambda(\phi^\dagger\phi)^2 \tag{2.33}$$

And as before if $\mu^2 < 0$ then there's a non-zero vacuum expectation energy. We want the vacuum to be invariant under U(1) so that the photon remains massless and neutrally charged, and so the minima to break the electroweak symmetry is the following:

$$\phi_0 = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{2.34}$$

Substituting the vacuum expectation value gives the following results

$$\phi_0 = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ \sqrt{\nu + H(x)} \end{pmatrix} \tag{2.35}$$

Which leaves only the Higgs field and the vector boson.

The mass of the vector bosons can be derived from expanding parts of the lagrangian:

$$|(-ig\frac{\tau}{2}W^a_\mu - i\frac{g'}{2}B_\mu)\phi|^2 = \frac{1}{8}\phi_0 = \sqrt{\frac{1}{2}}\begin{pmatrix} gW^3_\mu + g'B_\mu & g(W^1_\mu - iW^2_\mu) \\ g(W^1_\mu + iW^2_\mu) & igW^3_\mu + g'B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{2.36}$$

The actual W^+ and W^- fields are related to the component of weak isospin as such:

$$W^\pm = \frac{(W^1 \pm W^2)}{\sqrt{2}} \tag{2.37}$$

so substituting that back in to the lagrangian we get

$$(\frac{1}{2}\nu g)^2 W^+_\mu W^{-,\mu} + \frac{1}{8}(W^3_\mu B_\mu)\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} W^{\mu 3} \\ B^\mu \end{pmatrix} \tag{2.38}$$

And so the mass of the W boson comes out from the first term in the expression, $M_W = \frac{1}{2}\nu g$. The photon and Z boson are related to the weak isospin and weak hypecharge:

$$A_\mu = \frac{g' W^3_\mu + g B_\mu}{\sqrt{g^2 + g'^2}} \tag{2.39}$$

$$Z_\mu = \frac{g W^3_\mu - g' B_\mu}{\sqrt{g^2 + g'^2}} \tag{2.40}$$

with the mass of the photon $M_A = 0$ and mass of the Z boson $M_Z = \frac{1}{2}\sqrt{g' + g}$

This is a complete theory that agrees with the empirical observations that the weak interaction is mediated by a neutral Z , a W^- and W^+ massive vector bosons. Incidental to this theory is another prediction of another scalar particle, the Higgs boson, which also has a mass that has to be empirically derived. In the full standard model lagrangian,the Higgs mechanism gives rise to the fermion mass, which is left as a free parameter in the model and has to be empirically measured. In 2012 the Higgs boson was discovered in the LHC at a mass of 125 GeV, with the combined data of the ATLAS and CMS detectors at the 5σ level, finally completing this picture of the standard model [2].

2.2 LHC Physics

This section will describe how particle physics discoveries and measurements are made at the LHC. The LHC collides protons, initially at a center of mass energy of 7 TeV, later ramped up to 13 TeV in the run 2 data. Protons are composite particles, made up of two up quarks and a down quark, giving it +1 electric charge in total. At high energies, protons act as a bag of particles, with particles popping into existence at a probability described by a parton density function, shown in figure 2.3. Figure 2.4 shows a diagram of a proton-proton collisions. The "hard" collision comes from the interacting quarks or gluons. The hardness of a collision refers to the total momentum transfer. The soft collision is referred to in the diagram as the

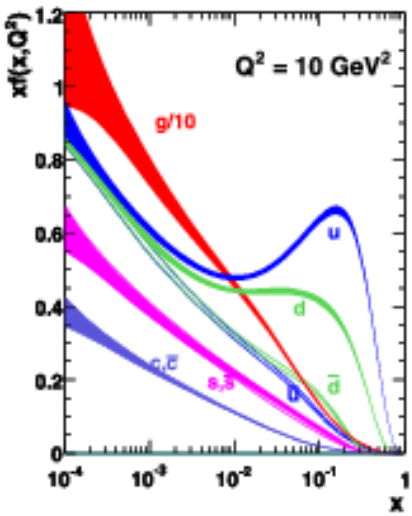


Figure 2.3. Parton density functions for quarks and gluons in a proton. The uncertainty of the PDF represents significant theoretical uncertainties in LHC experiments. The x-axis represents the fraction of momentum, the y axis is the scattering probability in pp collision

underlying event. Parton scattering is the most common hard process at the LHC by far due to the high density of gluons in the proton and the large coupling strength of the strong force compared to the electroweak force. Due to quark confinement, one

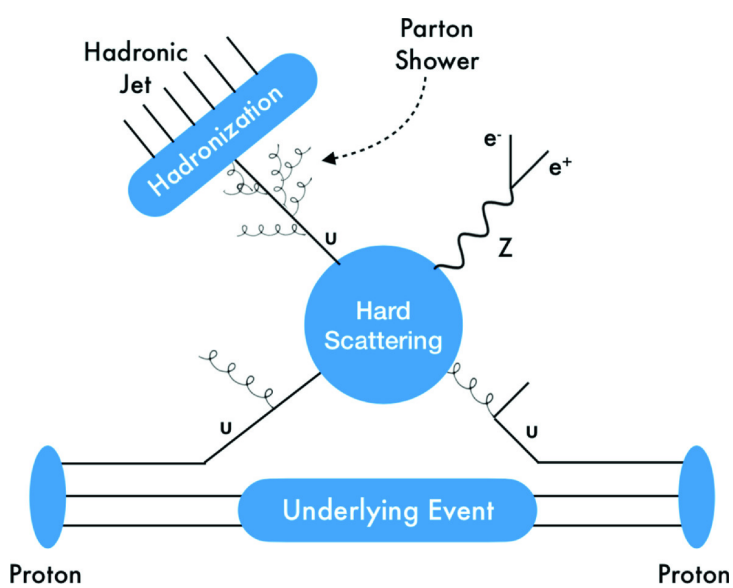


Figure 2.4. Cartoon of proton-proton collisions. The interaction is described by the collision of the specific quarks in the protons that hard scatter

cannot directly detect individual quarks or gluons. Quarks and gluons always appear in a color neutral bound state. Quark confinement is one of the poorly understood phenomenon of particle physics and is an area of active research. Instead, signatures of quarks appear as a "jet" of high energy, tightly collimated particles. Algorithms to reconstruct different signatures in the detector as a single jet are described in later chapters.

The cross section of an event is a measurement with units of area that predicts the probability that will occur. Cross sections in particle physics are measured in units of Barns, which is equivalent to 10^{-28}m^2. The unit "barn" comes from the Manhattan project, where it was used as a code name for the cross sectional area of a fast moving neutron from a nuclear fission reaction. Because a barn is very large compared to typical cross sections in LHC physics, it is more common to use picobarns (pb=10^{-12} barns) or femtobarns (fb=10^{-15} barns).

The instantaneous luminosity of the collider is the number of particle collisions (N) per area per second.

$$\mathcal{L} = \frac{dN}{\sigma dt} \tag{2.41}$$

The integrated luminosity (integrated over time) is the total number of collisions per unit area over the entire run. "Collisions" is a dimensionless number, which gives the integrated luminosity units of inverse area. If a particular process has a cross section σ_p, the number of events we would expect to observe of that process is straightforwardly:

$$N = \int \mathcal{L} dt \times \sigma_p \tag{2.42}$$

The integrated luminosity for run 2, where the data in this analysis originates, is 139 fb^{-1}.

2.2.1 Higgs Physics

The most common (largest cross section) Higgs production modes are gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a vector boson (VH), or associated production with a heavy quark (qqH). The Feynman diagrams of the production modes are shown in figure 2.5. The run 1 ($\sqrt{s}$ = 8 TeV) cross section of all the production modes are shown in table 2.1. It is evident that Higgs production is dominated by ggF, with VBF as a distant second by roughly an order of magnitude. Figure 2.6 shows the expected cross section of each production mode as a function of the center of mass energy of the collision, with the shaded outline indicating theoretical uncertainty. Figure 2.7 shows the current experimental uncertainty in the measurement of each production mode broken down by the decay mode of the Higgs [11].

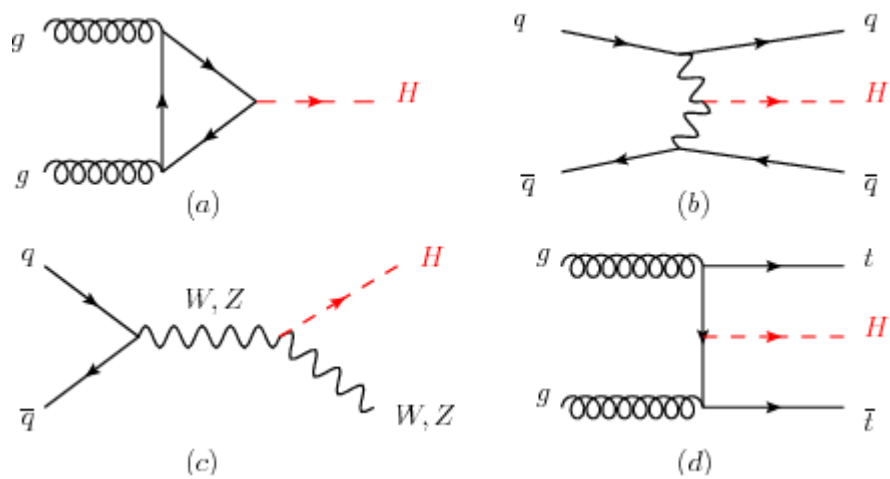


Figure 2.5. Feynman diagrams of Higgs production modes. a is gluon-gluon fusion (ggF), b is vector boson fusion (VBF), c is Higgs plus vector boson (VH), d is Higgs plus quark anti quark pair, (qqH, most commonly ttH)

For the VBF process, two quarks are produced which are are then reconstructed through their associated high energy jets in the forward region. The Higgs boson appears between these two jets, a topology that we take advantage of in the analysis to identify a VBF Higgs.

Production process	Cross section [pb] $\sqrt{s}$ = 7 TeV	$\sqrt{s}$ = 8 TeV
ggF	15.0 ± 1.6	19.2 ± 2.0
VBF	1.22 ± 0.03	1.57 ± 0.04
WH	0.573 ± 0.016	0.698 ± 0.018
ZH	0.332 ± 0.013	0.412 ± 0.013
bbH	0.155 ± 0.021	0.202 ± 0.028
ttH	0.086 ± 0.009	0.128 ± 0.014
tH	0.012 ± 0.001	0.018 ± 0.001
Total	17.4 ± 1.6	22.3 ± 2.0

Decay channel	Branching ratio [%]
$H \to b\bar{b}$	57.1 ± 1.9
$H \to WW^*$	22.0 ± 0.9
$H \to gg$	8.53 ± 0.85
$H \to \tau\tau$	6.26 ± 0.35
$H \to c\bar{c}$	2.88 ± 0.35
$H \to ZZ^*$	2.73 ± 0.11
$H \to \gamma\gamma$	0.228 ± 0.011
$H \to Z\gamma$	0.157 ± 0.014
$H \to \mu\mu$	0.022 ± 0.001

Table 2.1. Run 1 ($\sqrt{s}$ = 8 TeV) experimental measurements of Higgs production cross sections (left) and decay branching ratios (right)

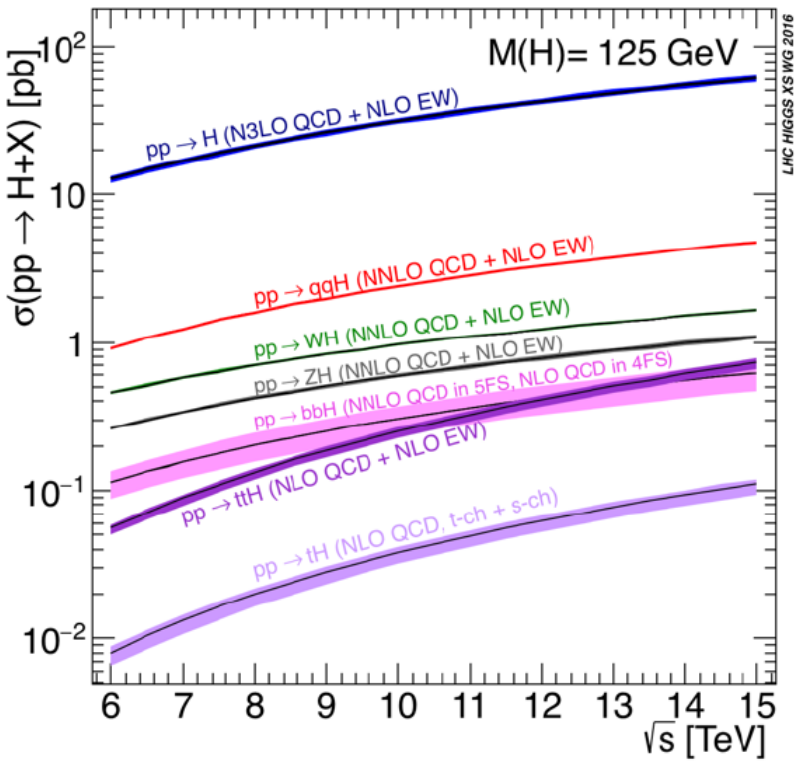


Figure 2.6. Cross section of Higgs production modes as a function of center of mass energy of the proton proton collision

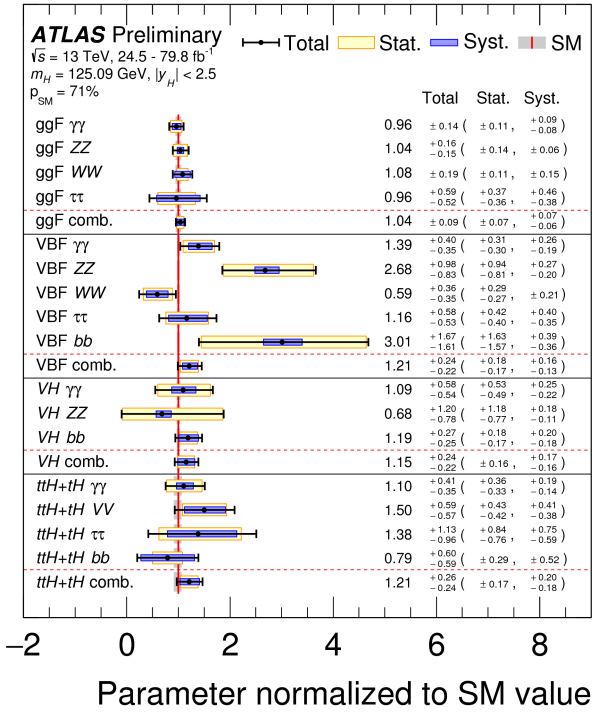


Figure 2.7. Current experimental measurements of higgs cross sections of various processes, with uncertainties, relative to standard model expectations

Table 2.1 shows the expected Higgs decay branching ratios. The Higgs boson couples to mass, so the more massive particles that are still lower mass than the Higgs, b quarks, are the most likely decays. The first discovery of the Higgs boson was however made in the combination of ZZ, $\gamma\gamma$ and WW* channels, which has a lower branching ratio. This is because the combined mass of two W bosons is more than the Higgs mass, which necessitates an off-shell W. The WW channel has less backgrounds, as its not contaminated by QCD jets from hard proton-proton scattering [11]. It was not until 2018 that a discovery of Higgs decaying to bb was achieved at the 5 σ level.

2.2.2 Whats Next?

The discovery of the Higgs boson in 2012 was the last major piece of the standard model puzzle, but there's still a significant amount of empirical observations that the standard model cannot explain. In fact, the particles of the standard model explain about 5% of the observable universe. Dark matter, first discovered by Vera Rubin studying anomalous galactic rotation curves, makes up about 20% of the energy in the universe. The rest is attributed to dark energy, the energy that causes the accelerating expansion of the universe. The standard model of particle physics has nothing to say about 95% of the universe, which means there must be physics beyond the standard model (BSM) to explore.

As mentioned before, the standard model doesn't incorporate gravity. Some scientists speculate of a "graviton" particle, a quanta of the gravitational force analogous to the photon as a quanta of the electromagnetic force, but that hasn't been discovered. Other mysteries include the masses of neutrinos, the strong CP problem, and matter anti-matter asymmetry. The strong CP problem, shortly, is that the strong force appears to not violate Charge-Parity symmetry, even though the weak force does, without a fundamental explanation in the standard model. Matter anti-matter

asymmetry is the question of why does the universe seem to contain much more matter than anti-matter, even though the laws of physics are symmetric between the two.

There are two paradigms when searching for new physics, direct and indirect search. Direct search at the LHC seeks to find new particles by a measurement of excess events above standard model backgrounds, the bump seen in figure 2.8 for example when the Higgs was discovered. Indirect search seeks to find new physics by measuring standard model quantities, such as the cross section of a Higgs boson decay, precisely enough to see if the experimental quantity disagrees with theory prediction, indicating BSM physics contribution to the process. Since the Higgs boson discovery, direct searches of new physics have not yielded any results. Indirect searches have had somewhat more success, notably the muon g-2 experiment which is closely approaching 5σ deviation from standard model expectations [1], and a recent paper from LHCb that found a violation of lepton universality from b-quark decays, at the 3.1 σ level, which is not a discovery but an interesting result. The goal of this analysis is to determine if a precision measurement of the Higgs decay can indirectly reveal interactions between the Higgs boson and new BSM physics.

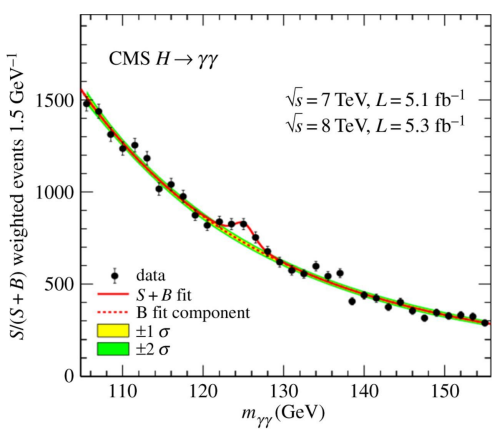


Figure 2.8. Discovery of the Higgs boson (in the $\gamma\gamma$ channel) through a "bump" of excess events above the standard model background at 125 GeVs

CHAPTER 3

ATLAS DETECTOR

The Large Hadron Collider (LHC) is a proton-proton storage ring operating at CERN, and for its 9 years of operation, it has been the world's highest energy particle collider. During LHC operation thus far, protons have collided with increased center-of-mass approaching the design energy of 14 TeV. Instantaneous luminosity has also successively increased, surpassing design instantaneous luminosity of $1024cm^{-2}s^{-1}$ in 2018 to reach $1024cm^{-2}s^{-1}$. The overall data recorded in the ATLAS detector totals more than 1016 collisions. Operation of the LHC has led to the discovery of the Higgs boson and some of the most precise measurements of its properties including the coupling of the Higgs boson to bottom quarks, W and Z bosons, photons and tau leptons. The LHC has also facilitated searches for new physics over a wide parameter space, setting confidence level exclusion limits on masses of supersymmetric particles like squarks, gluinos and neutralinos. The LHC can run continuously for a few years before detector components need to be repaired and replaced. The schedule of data-taking consists of long periods of data accumulation (Run 1 and Run 2) paired with long shutdown periods. The LHC is set to begin Run 3, in which the design center-of-mass energy should be reached, in 2021. Following Run 3, detector upgrades will be installed during a long shutdown. The High-Luminosity LHC (HL-LHC) will begin colliding protons with ten times the luminosity in 2027.

Acceleration of protons start with the Linac-2. This injects protons at 50 MeV into the Proton Synchronous Booster (PSB) where they are further accelerated to 1.4 GeV. In the Proton Synchrotron (PS), the protons are separated into bunches with a spacing

of 25 ns and are further accelerated to 25 GeV before being extracted to the Super Proton Synchrotron (SPS), where they reach 450 GeV. Finally the bunches of protons enters the LHC, where they are accelerated to their final energy of 6.5 TeV. Linac 2, PSB, PS, and SPS were all operational accelerators before the LHC era though each had to be majorly upgraded to handle the energy and beam intensity required for LHC collisions.The LHC layout mimics that of the Large Electron Positron collider (LEP) that was housed in the same tunnels. Figure 3.1 shows the positioning of each experiment at the LHC as well as injection systems and other features. Once proton bunches enter the LHC in two opposing beams, they are accelerated with radio frequency (RF) systems. Located at Point 4 in the LHC schematic, the system consists of 16 RF cavities operating at twice the frequency of the SPS injector. RF cavities are metallic chambers containing oscillating electromagnetic fields; in the LHC this oscillation frequency is 400 MHz. The tuning of this frequency ensures that protons of the ideal energy are not accelerated further and simply maintain their momentum while particles arriving in an RF cavity slightly before or after will be decelerated or accelerated toward the ideal proton energy. This acceleration process can also be used to split beams of protons into discrete bunches, and this is first done with RF cavities in the PS. After proton bunches have circled the LHC approximately 1 million times, peak energy is reached and collisions can commence.

3.1 The ATLAS detector

The ATLAS (A Toroidal LHC ApparatuS) detector diagrammed in 3.2 is a large cylindrical detector centered around one of the collision points in the LHC. The detector has many instruments to examine the particles and interactions from the collision. Figure 3.3 shows how different standard model particles are detected in each layer ATLAS.

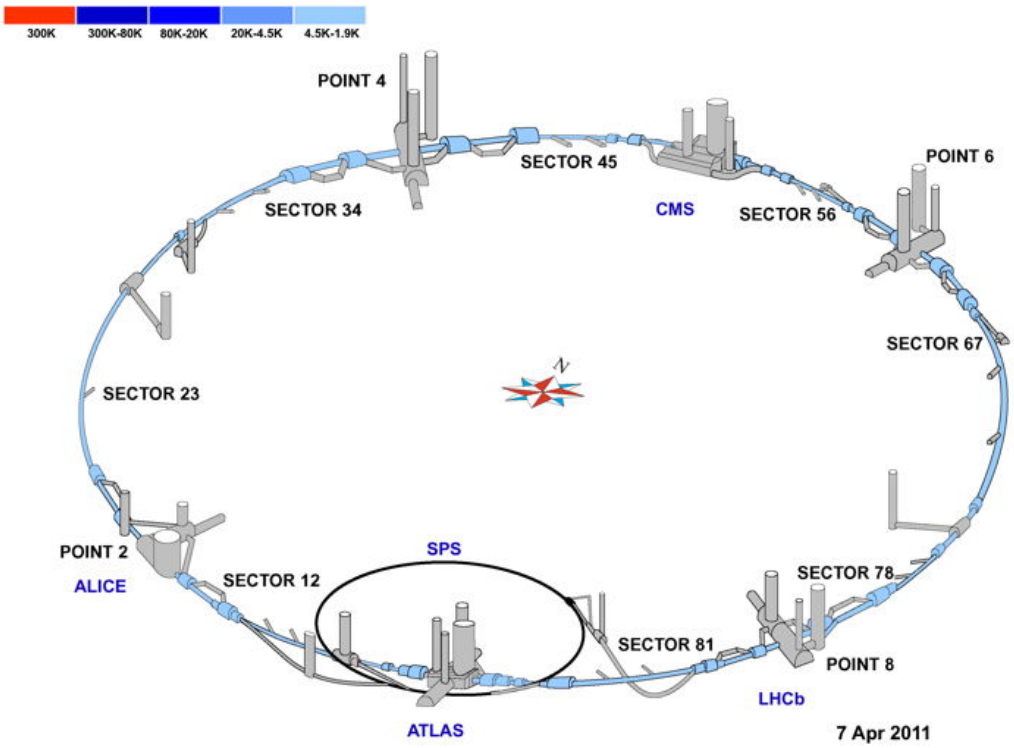


Figure 3.1. Diagram of the LHC, with the SPS feeding beam. The collision points are where the detectors, ATLAS,CMS,ALICE,LHCb, are setup

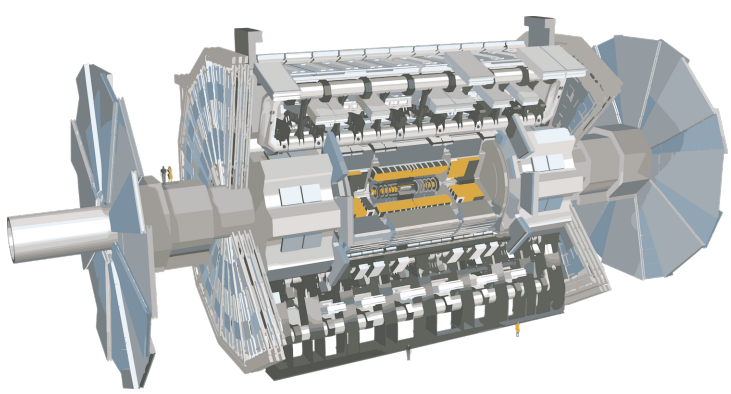


Figure 3.2. The ATLAS detector. The middle of the detector is the proton beam line

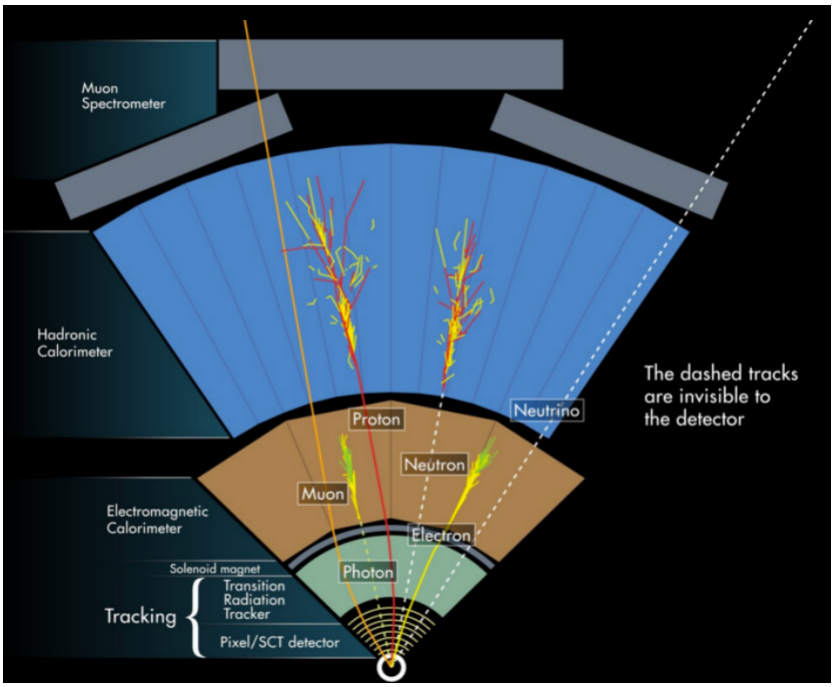


Figure 3.3. A cross-sectional diagram of the ATLAS detector. The inner most layer, ID, measures tracks of particles. The EM calorimeters causes photons and electrons to shower, the tile calorimeter causes hadrons like protons and neutrons to shower, which allows us to identify them. Muons reach the Muon Spectrometer in the outer layer. Neutrinos (and potential BSM particles) are invisible to the detector and are simply missing energy

The most sensible coordinate system to describe particles in the ATLAS detector are cylindrical coordinates. The center of the detector, where the protons are most likely to have collided is the interaction point (IP) and is the designated origin of the coordinate system. The distance from the IP in the transverse plane is d_0 parameter, and ϕ is the azimuthal angle around the beam line. The angle in the longitudinal plane is θ. When analyzing data it is more common to look at the pseudorapidity η, which is defined as $-ln(tan(\theta/2))$, since we are largely interested in smaller values for the polar angle. The final parameter for detected particles is their momentum p, which is calculated from the radius of curvature as they travel through the magnetic field in the inner detector. The transverse momentum $p_T = p \cdot sin(\theta)$ is often more interesting because the total momentum of the particles in the transverse plane must be 0, since the protons collide head-on in the direction of the beamline, so therefore the total p_T of a single collision must cancel out. The inner most part of the ATLAS detector is the

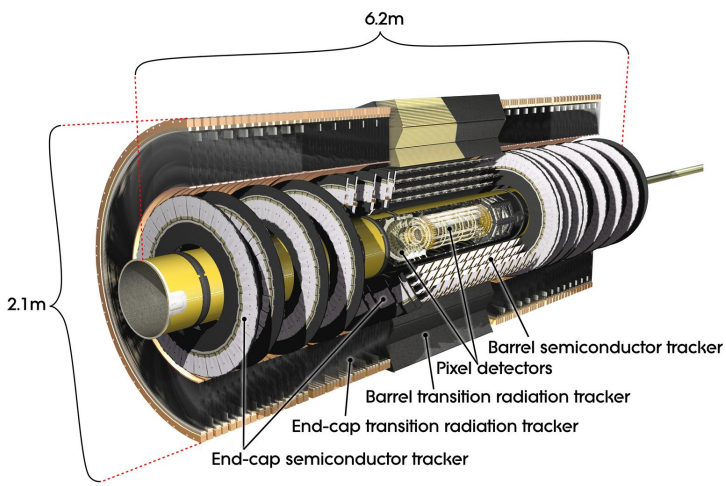


Figure 3.4. ATLAS Inner Detector

Inner Detector (ID), which is used to reconstruct the track of charged particles from the IP. It has a coverage of$|\eta| < 2.5$ and is bathed in a 2 Telsa solenoida magnetic field. The magnetic field causes the tracks of particles to curve, with the radius of curvature proportional to their momentum. The inner most part of the inner detector, the

Insertable-B layer (IBL) added in 2014, consists of four layers silicon pixel detectors. Outside the IBL are the semi-conductor trackers (SCT), which consists of four layers of silicon strip detectors. The outermost part of the inner detector is the Transition Radiation Tracker (TRT), which consists of 50,000 drift tube detectors in the barrel layer and 250,000 tubes in the endcaps.

After the Inner Detector are the calorimeters. The first layer of calorimeters are the EM-calorimeters, made of layers of liquid argon and lead. EM-calorimeters measure the energy of electrons and photons. Lead absorbs the particle which coverts them to a shower of low energy particles that ionize the liquid argon. Electrons lose energy through Bremsstrahlung radiation, while photons lose energy from pair production. Surrounding the liquid argon detectors are the hadronic calorimeters(H-Cal). The h-cals are made of layers of steel and plastic scintillators, Hadrons such as protons and neutrons lose energy in the steel and create a shower of particles picked up by the plastic scintillators and measured with photomultiplier tubes. The calorimeters have a coverage of $0 < |\eta| < 4.9$

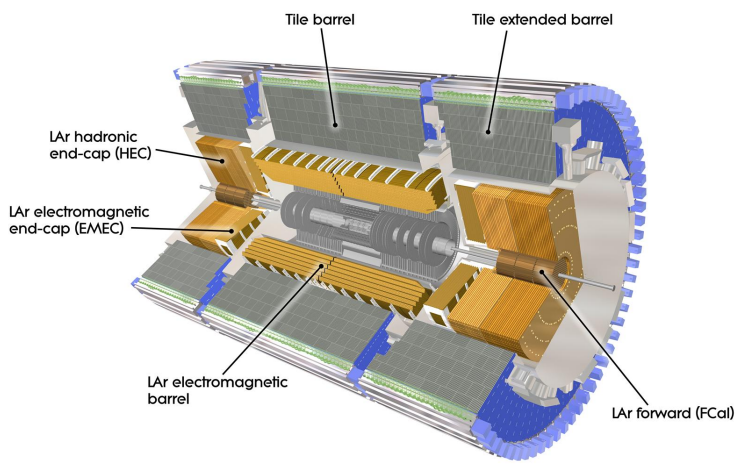


Figure 3.5. EM liquid argon and hadronic tile carlorimeters

Finally the Muon Spectrometer (MS) is outermost layer of ATLAS. Due to their larger mass, muons go through the EM calorimeters without losing much energy to

Bremsstrahlung radiation, so are detected in the MS roughly 10m from the IP. The MS consists of three layers of Monitored Drift Tube chambers covering $|\eta| < 2.7$ and an inner layer of Cathode Strip Chambers with $|\eta| > 2.0$ in the small wheel of the endcap. In addition, it includes trigger chambers that contain 3 layers of Resistive Plate Chambers and 3 layers of Thin Gap Chambers. Muons are a critical measurement in ATLAS, as they can be measured cleanly, with smaller uncertainty and background contamination compared to electrons and photons, providing a useful method of precision measurements.

The ATLAS detector contains a central solenoid magnet, which provides a 2 Tesla magnetic field in the Inner Detector. As discussed, this magnetic field curves the path of charged particles in the ID, allowing us to measure their momentum by the radius of curvature. The inner solenoid magnet is able to achieve a relatively strong field at 4.5 cm thickness using an embedding of niobium-titanium superconductor wires. The MS is embedded in a toroidal magnet, which allows the measurement of the muon's momentum. There are three distinct toroids, one at the barrel and one in each of the two endcaps, reaching a magnetic field of 3.5 Tesla. The magnets operate a temperature of about 4.5^0 K

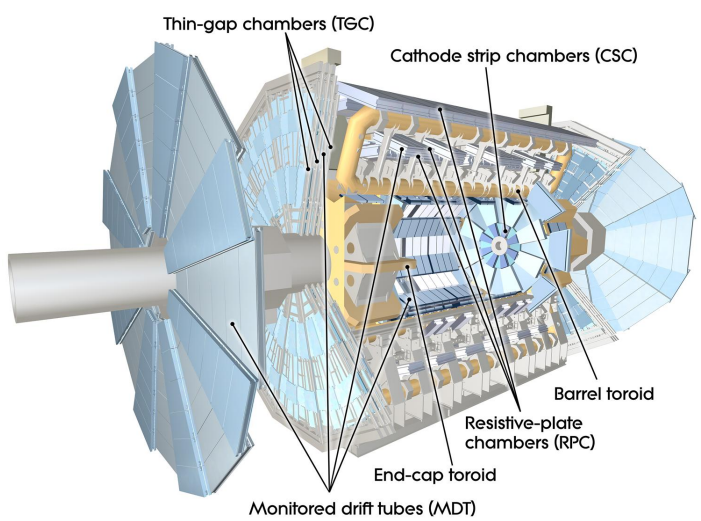


Figure 3.6. ATLAS muon spectrometer

CHAPTER 4

INNER TRACKER

During the planned shutdown in 2024, the LHC will be upgraded to the high luminosity LHC (HL-LHC). At the core of the ATLAS detector upgrade for the HL-LHC is a new sub-detector for tracking charged particles, the Inner Tracker (ITk), figure 4.1. The upgrade of the tracker system is necessary to cope with the larger event collision rate, the larger pileup and the higher radiation doses expected at the HL-LHC compared to those experienced at the LHC. Figure 4.2 shows a comparison between the pileup in the current run of the LHC vs. the expected pileup in the HL-LHC. The expected timeline (as of 2018) is shown in figure 4.3, with the HL-LHC installation between 2024 and 2026 and data taking with high luminosity starting after 2026.

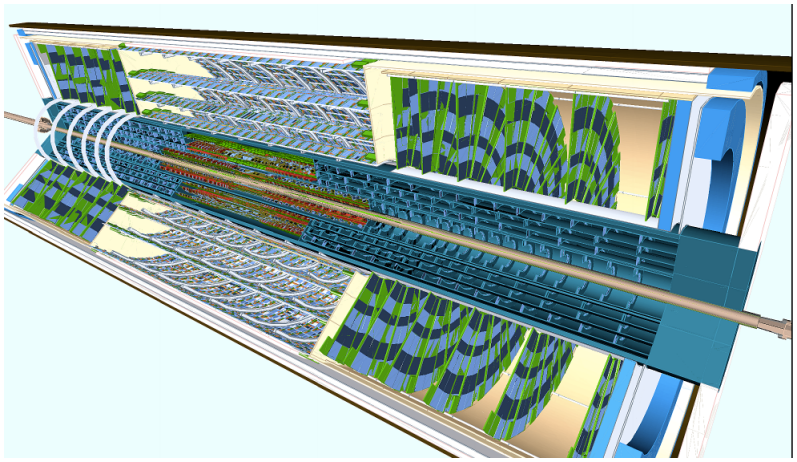


Figure 4.1. Expected shape of the Inner Tracker. The fine inner layer of detectors are called pixels, the coarser outer layer of detectors are called strips. The barrel region will have rectangular staves placed in cylindrical layers, the endcap region will have detectors placed in a petal geometry.

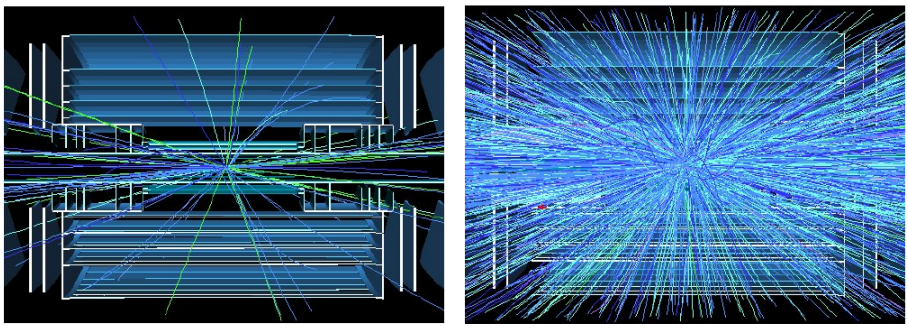


Figure 4.2. The expected luminosity of the High Luminosity LHC (right) compared with the current luminosity (left). The HL-LHC luminosty is expected to be 7-10 times larger

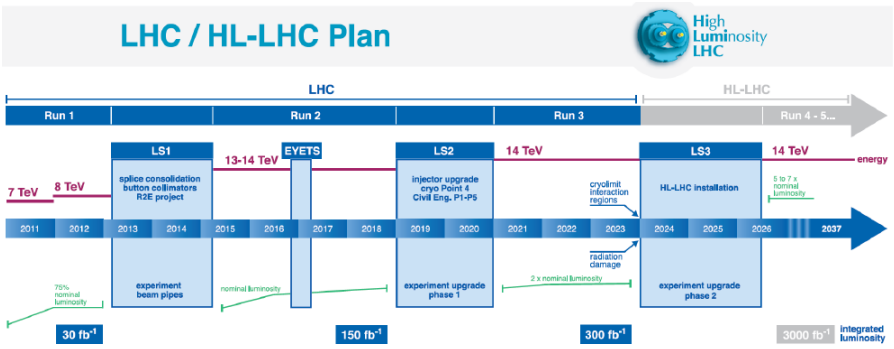


Figure 4.3. Expected timeline of LHC upgrades, leading to the High Luminosity LHC (pre-pandemic expectations)

4.1 silicon detectors

A silicon strip behaves similar to a diode, one side is p-doped and the opposite side is n-doped. Doping silicon is when a foreign element is injected injected into the silicon lattice, as shown in figure 4.4. Silicon has four valence electrons, so p-doped silicon has an element with three valence electrons, n-doped silicon is injected with an element of five valence electrons. Note that p and n-doped are still neutral, it does not imply an access of positive or negative charges. Electrons can jump from n-doped to p-doped elements, leaving a "hole" behind that can drift around the silicon lattice acting like a positive charge.

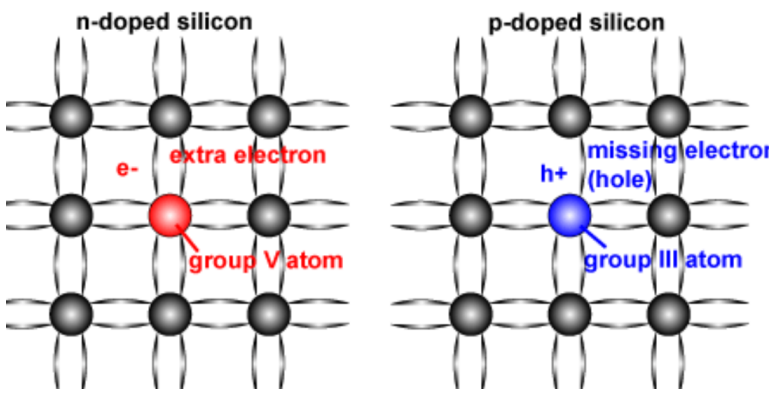


Figure 4.4. Example of doped silicon lattice. N doped lattice contains an element with one more valence electron, p doped contains elements with one less valence electron

When voltage is applied in the forward direction (forward bias) the electrons combine with holes and current flows with little resistance above a threshold voltage. When voltage is applied the other direction (reverse bias), a depletion region between the p-doped and n-doped sides prevents any significant current from flowing.

silicon detectors make use of the depletion layer from reverse biased detector. As shown in figure 4.5. When an ionizing particle moves through the depletion layer, it knocks a valence electron in the silicon lattice, creating an electron hole pair. Due to the reverse bias the electron and hole drift in opposite direction, and the access

charge is collected on the ends of the detector which creates an analogue signal that is then read by the readout electronics.

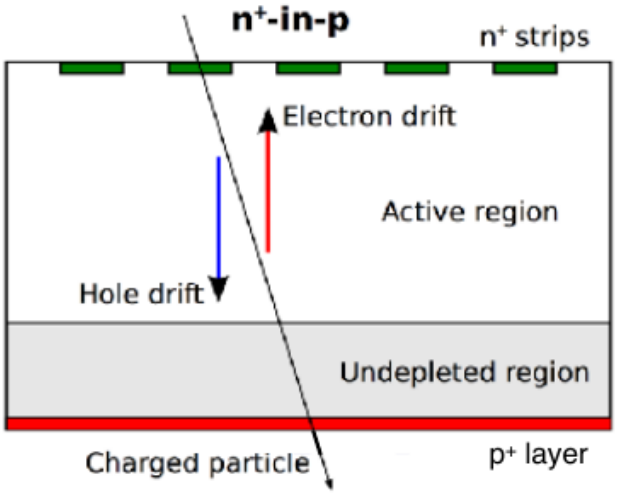


Figure 4.5. Ionizing particle through a silicon detector. The ionizing particle creates an electron hole pair in the depletion region. The electrons drift onto the charge collection strips at the top

4.2 layout

ITk will be an all-silicon detector, with silicon pixel detectors in the inner layer and silicon strips in the outer layers. Figure 4.6 shows a diagram of the current layout, with five pixel layers and four strip layers. The barrel will be made up of rectangular staves shown in figure 4.7 and the endcaps will use a petal geometry shown in figure 4.7. The barrel strip detector will contain four layers of cylindrically placed staves, each of which will provide mechanical support to 28 silicon modules. Each module contains a row of 10 readout chips.[6]

Stave layout is shown in figure 4.8. The body of the stave is composed of carbon fiber, in a honeycomb pattern to minimize material, and a titanium cooling tube for CO_2 cooling. A polyamide bus tape, which covers both faces of the stave, contains the electrical connections to provide power, clock and data readout to the modules.

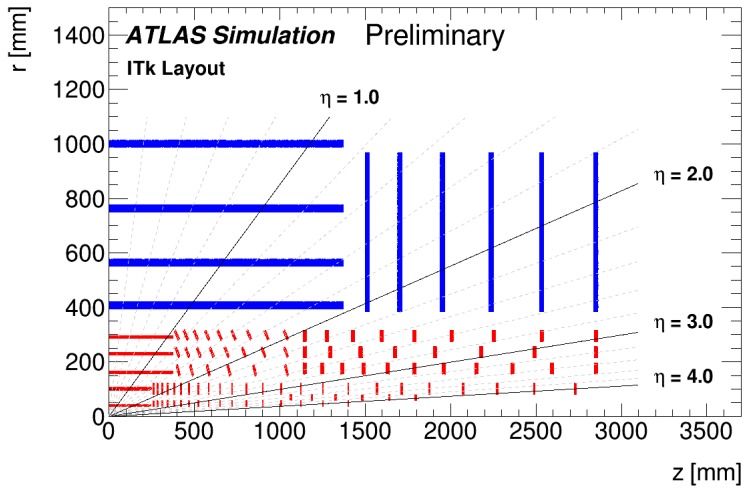


Figure 4.6. ITK strips and pixels layout

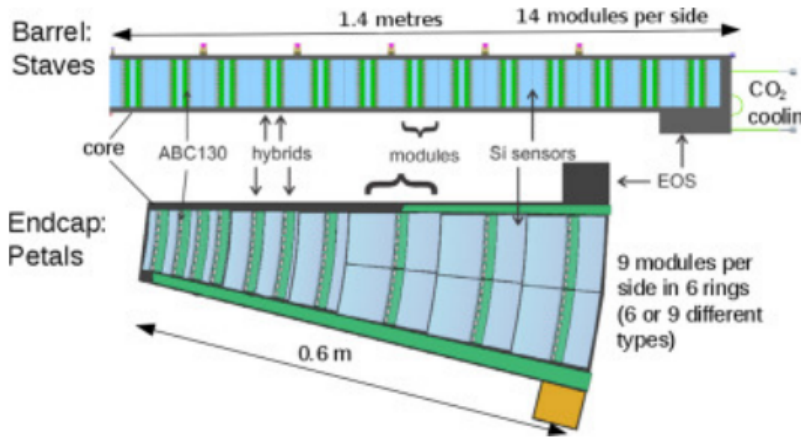


Figure 4.7. Barrel staves and endcap petals. The staves are laid out in a rectangular geometry, the petals are laid out like a flower in the endcaps. Both use the same module readout systems and same support structures.

The data readout and the clock to the stave are provided by an end of substructure (EOS) card, shown in figure 4.9. The data is converted from an electrical to an optical signal with a GBTz chip on the EOS card and readout by the data acquisition system (DAQ) from the EOS through fiberoptic cables.

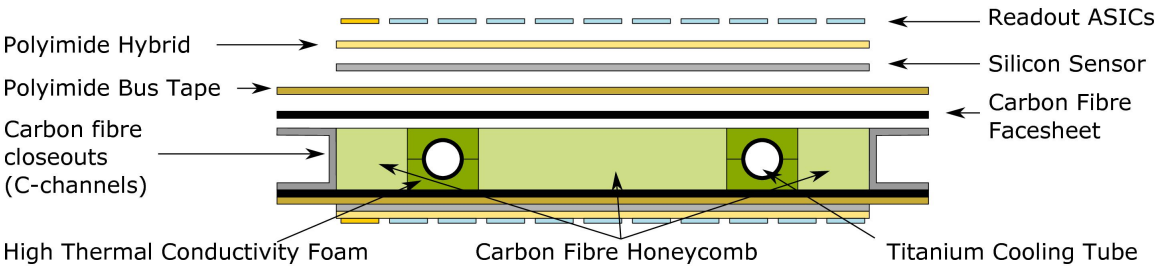


Figure 4.8. ITk Strips Barrel Stave

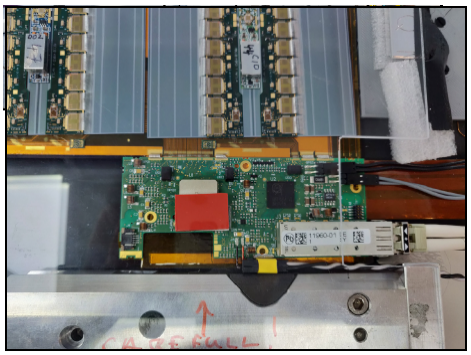


Figure 4.9. End of substructure (EOS) card for staves

Each silicon strip module shown in figure 4.12 contains three printed circuit boards (PCB) glued on top of the silicon strip sensor. The center PCB is the power board, and on the PCBs on both sides of the power board are the Hybrids. The power board, shown in figure 4.10, contains a DC-DC buck converter, high voltage multiplexer, and an AMAC (autonomous monitor and control) chip.The DC-DC converter is used convert 11 volt input voltage to the approximately 1.5 volts needed by the module electronics, to save cable space. The input voltage for the module electronics is termed

low voltage (LV), the voltage to reverse bias the sensor is termed high voltage (HV). The high voltage is planned to operate at 700 volts.

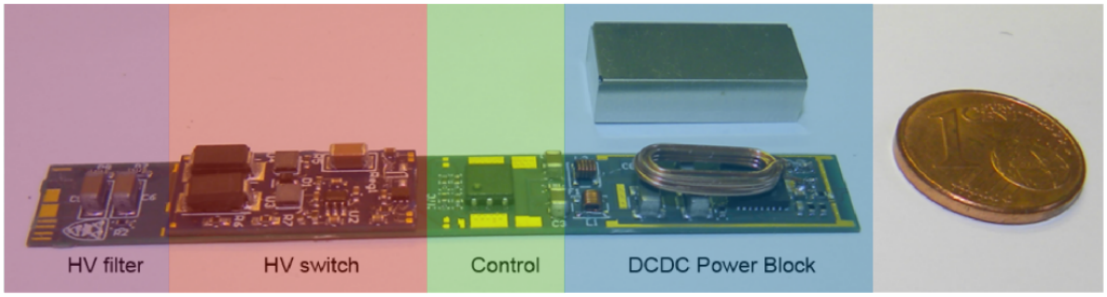


Figure 4.10. Module power board, contains the DC-DC converter, HV multiplexer and AMAC control chip

The high voltage multiplexer uses radiation resistant Panasonic GaNFET. To save cable space, each four modules are on the same high voltage bus, shown in figure 4.11 schematics, instead of each one a singular power supply. The HV multiplexer is used to turn off the high voltage to a single module on the HV bus without disrupting the other modules.

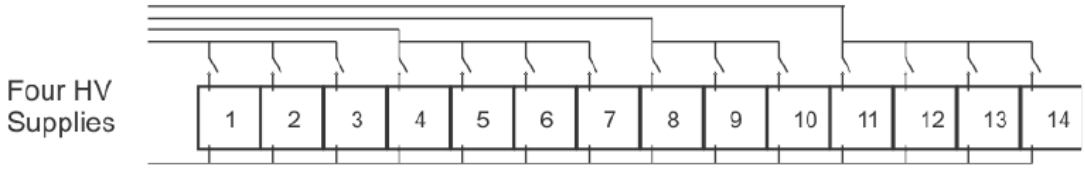


Figure 4.11. High voltage bus scheme for modules on a stave, where each four modules share a bus and can be turned on and off with the HV mux

Finally, the AMAC is used to programmatically control the low voltage and high voltage on the module and readout the module sensors which include thermistors, humidity sensors and module current measurement. The AMAC communicates through

an I^2C like protocol with the data control systems (DCS). Quality control tests will be synchronized with the module sensors read by the AMAC. [9] .

A hybrid is populated with ten readout chips called Atlas Binary Chip (ABC) and one Hybrid Control Chip (HCC). The final version of the chips is the ABCStar and HCCStar set. The ABCStars are wirebonded to the silicon sensor. A signal from the chips is shaped by analogue circuits then digitized with a one-bit discriminator. The data from the ten ABCStars on the hybrid is organized within the HCCStar, which is then read out by ITk data acquisition system. The diagram of the ABCs is shown in figure 4.13, with analogue signal shaping, discriminant and low level triggers. Figure 4.14 shows the overall data flow scheme from the stave modules to the EOS to the data acquisition [6].

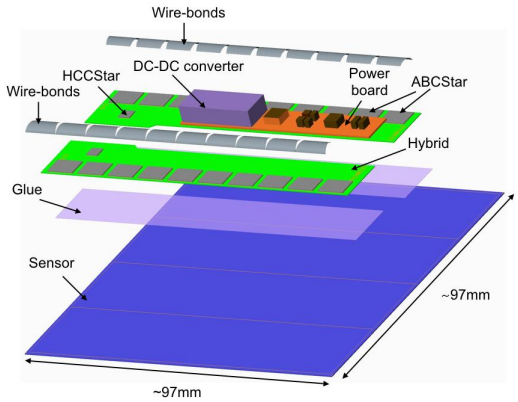


Figure 4.12. Silicon Strip Module diagram. The hybrid and power board are glued onto the silicon strips, the readout chips ABC and HCC, are then glued on top of the hybrids

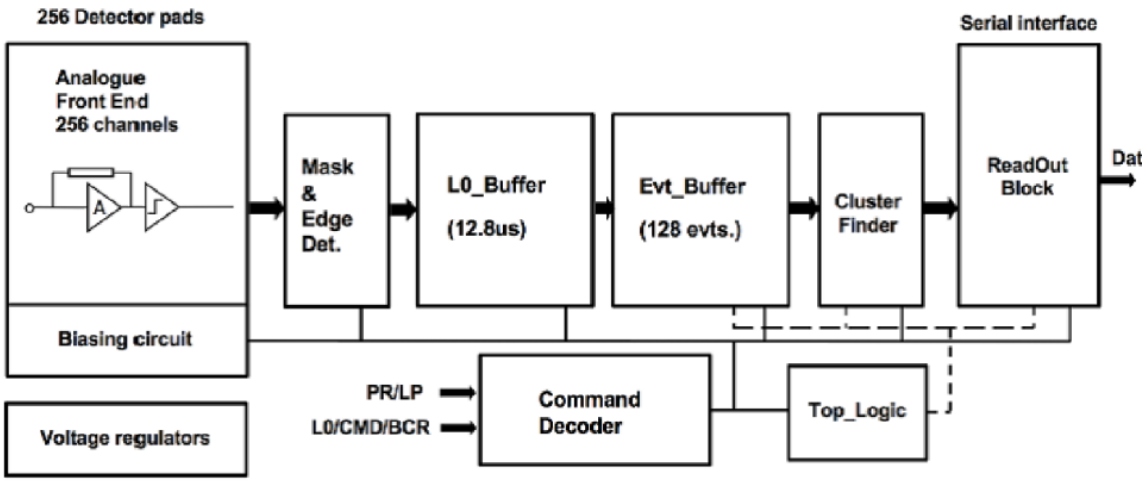


Figure 4.13. ABCStar schematics. Data is digitized with a one bit discriminator, then buffered in the L0 buffer

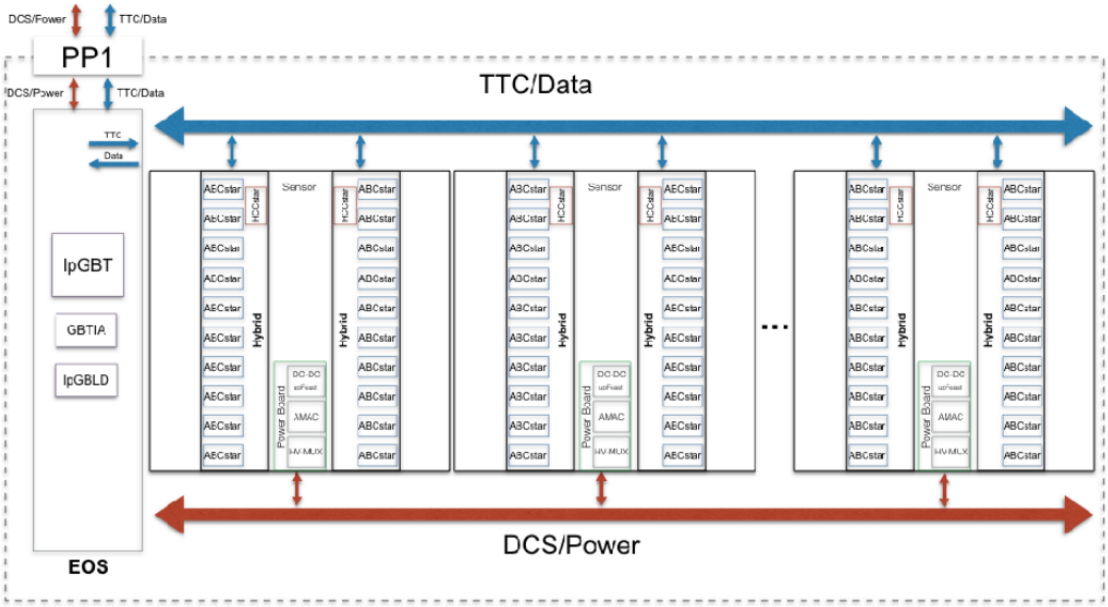


Figure 4.14. data flow from modules to stave to EOS card to data acquisition system. The EOS card collects data, provides the clock (TTC) and power, and reads out the module sensors (DCS)

CHAPTER 5

DATA ACQUISITION

The data acquisition tests are used to determine if any of the channels of the modules are malfunctioning. Malfunction could come from different sources, such as a broken wire bond, dead channel, cross talk between channels and outside noise interference.

The fundamental test of the ABC chip is the threshold scan which produces an s-curve plot seen in figure 5.1. In a threshold test, the settings of the chip's discriminator for each channel is tested from zero to several hundred millivolts, and the occupancy of the channel is measured for a test charge from a capacitor on the channel. For low discriminator values the occupancy is hundred percent, because all hits pass the discriminator. Conversely, for high values the occupancy is zero percent and at some threshold value the occupancy smoothly transitions from one hundred to zero. Without noise the transition would be instantaneous, but because of noise in the system the transition is spread out giving the s-curve shape shown in figure 5.1. Then an error function is fitted to the data, the threshold value at the mid point of the error function is the signal level of the channel is called the vt50 point and the width of the error function is the output noise of the channel in millivolts.

The next three tests performed on the chip are a strobe delay, three point gain and trim range. Strobe delay is the time delay between the command to inject a calibration charge and the actual injection, to synchronize with the clock speed of the readout. Figure 5.3 shows a plot from the ABCStar strobe delay test where the occupancy is plotted as a function of the channel delay. The plot has a symmetric

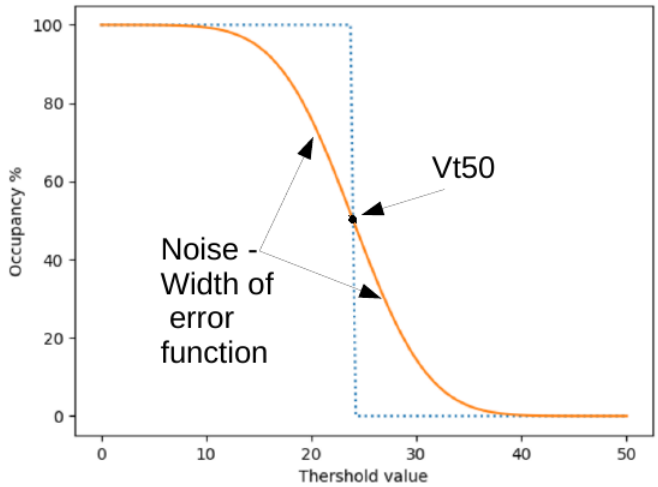


Figure 5.1. Threshold scan

"top hat" shape where a short delay is to early too read out data and a long delay is too late. The delay value is taken from between the 25% to 75% threshold. Figure 5.2 shows the strobe delay levels for all the chips on a single hybrid, where the line indicates the chosen delay levels.

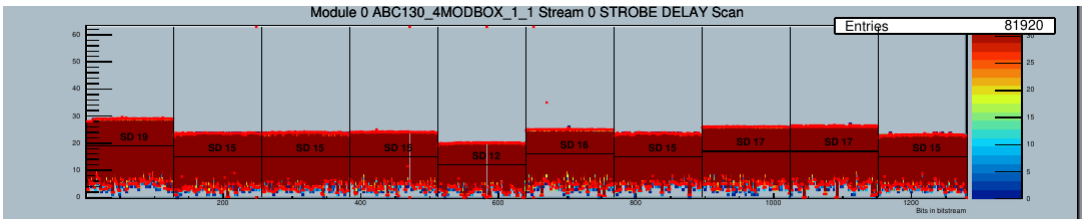


Figure 5.2. strobe delay values for each of the ten chips in a hybrid. The chosen delay values

A three point gain, show in figure 5.4, is a measurement of the gain of each channel. Gain is the signal level of the channel (in volts) for a given input charge. Typically a linear approximation is used to model the relationship between the size of the input charge and the voltage output of the channel, and so three points is sufficient. For more accurate probing of nonlinearity, ten points are be used. The test measures the

gain by taking three test charges at regular intervals (typically 0.5 fC, 1fC, and 1.5 fC) and measuring the vt50 for each charge. The slope of the signal level vs. charge is the gain of the channel in $\frac{mV}{fC}$.

The input noise, a gain-independent noise measurement, of the channel is the output noise (width of the s-curve) divided by the gain, measured in Equivalent Noise Charge (ENC). The expected noise for a working ABC chip is around 600 ENC. Dead channels or broken wire bonds are identified when the input noise is either too low (indicating a broken wire bond) or too high (indicating a possible interference or other problem in the channel). A broken wire bond is identified with an input noise of approximately 400 ENC for a given channel

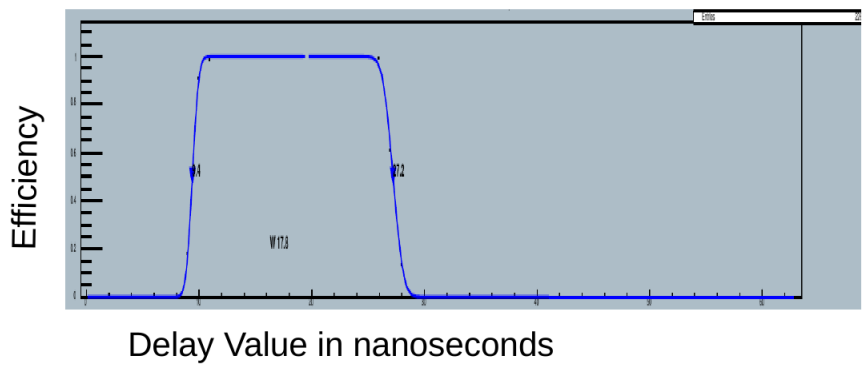


Figure 5.3. Strobe delay for a single channel

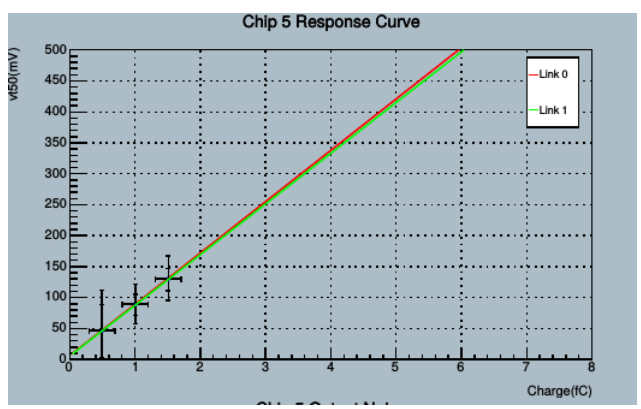


Figure 5.4. Three point gain test for a single channel

A trim range scan adjusts a five bit digital analogue converter (DAC) on each channel to align the vt50 point of all the channels as closely as possible. Trim range is done so that all the channels give the same signal for the same charge (but not necessarily the same noise, this is still on channel by channel basis). Figure 5.5 shows the test values for all the channels of a single hybrid, before and after trimming. The Vt50 points (top plot) are calibrated at test charge, as the plot shows they are brought to alignment. The noise and gain are unaffected by the trim scan.

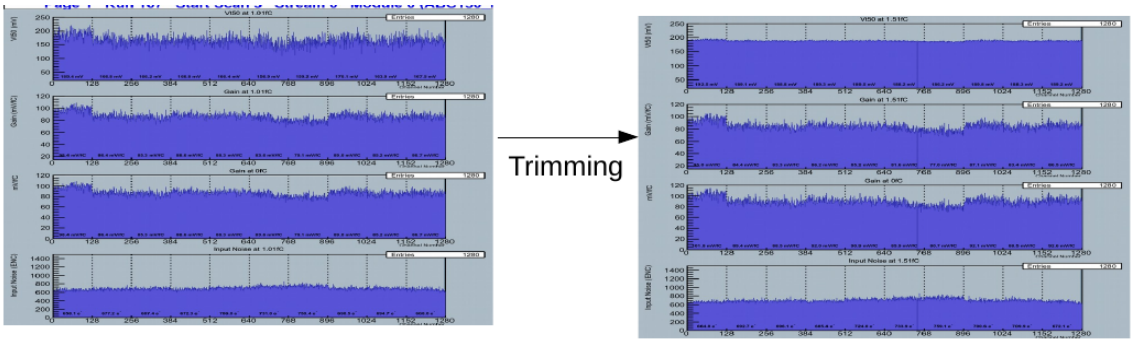


Figure 5.5. effect of trim range on DAQ tests. The Vt50 (top plot) of all the channels are brought into alignment. The gain (middle plots) and noise of the channels (bottom plot) remain unaffected

Finally the voltage vs. current curve is calculated for the sensor. As the sensor acts similar to a diode, the voltage is reverse biased, which gives a sensor current on the order of microamps.There's a roughly linear correlation between the sensor current and the applied voltage until the breakdown voltage. Malfunctioning sensors will often breakdown at to low voltages, below 1000 volts.

Single modules assembled at BNL,SCIPP, and LBNL are first tested in a single module box before being glued to staves. The parts of the single module box include the cold chuck shown in figure 5.6 connected to a water chiller and an HV plate, electrically insulated but in thermal contact with the chuck through a ceramic plate. The single module box is used to test the electrical characteristics of modules before being assembled on staves.

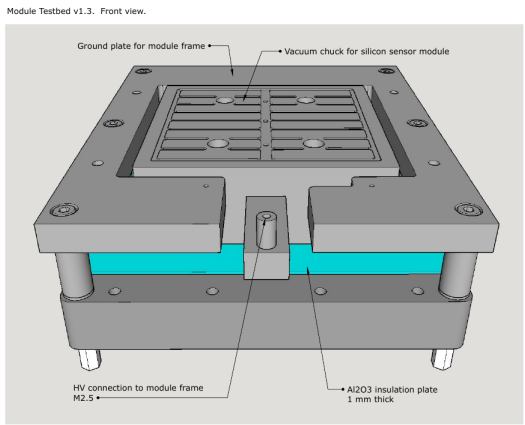


Figure 5.6. cold chuck for module testing

5.1 stave test

The first full assembled stave in BNL is shown in figure 5.7. The stave was tested with the DAQ tests described in the previous section and the results were compared to a single module tests with the cold chuck.The setup is shown in figure 5.8, with a Keithley HV power supply, Sorensen LV power supply and GLIB field programmable gate array (FPGA) in the center to readout the fiber optics. The EOS card for the stave is shown in figure 5.9, contains a GBTx chip to convert the electrical signal from the readout electronics. The tests showed effectiveness of the assembly line, producing modules to gluing onto the local support to wirebonding. The stave contained twelve total modules, 6 produced at LBNL and produced in SCIPP.

The stave was tested with a sensor bias 400 volts, and kept at 15° C. The results of the test are shown in figures 5.11 (input noise) and 5.10 (output noise), and 5.12 (gain), comparing the output noise and input noise between the stave tests and the single module box tests. The plot shows the average noise and gain for every 128 channles (one half of an ABC chip).On first order, the noise and gain of all the channels are comparable between the single module tests and the stave tests, which proves the effectiveness of the stave production assembly line. The gain is about 10%

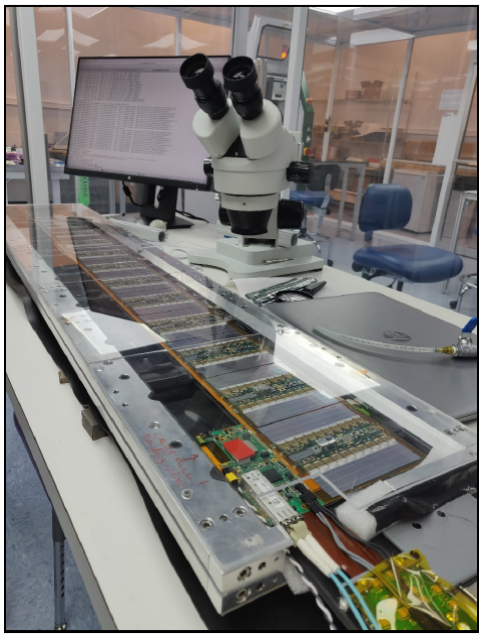


Figure 5.7. first full electrical stave produced at BNL. The end of substructure card is on the bottom , connected to the fiber optic connection. The microscope was used to find broken or mis-bonded wire bonds

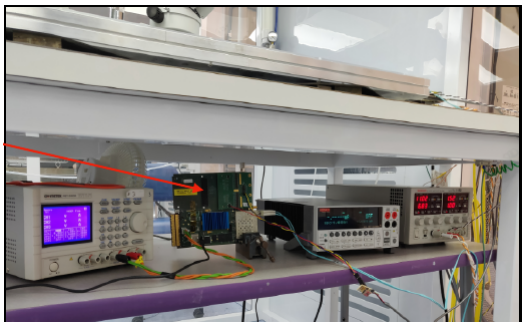


Figure 5.8. Stave DCS setup. Data from the fiber optics is read out by a GLIB FPGA

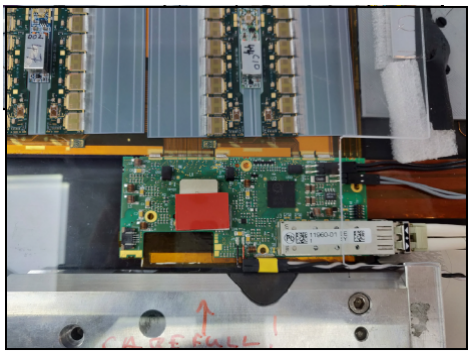


Figure 5.9. EOS card with GBTx chip that drives the optical signal

lower on the stave than on single module tests on a cooling chuck, for reasons that are still under investigation.

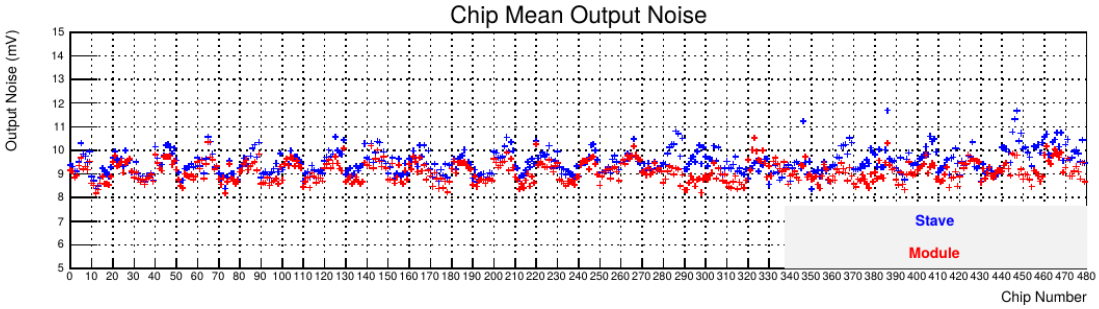


Figure 5.10. Comparison of output noise between stave and single module tests

There is systematic increase in every ten channels because the channels directly glued to the Hybrid have a slightly higher noise due to the capacitance between the strips and the backplane of the Hybrid. The output noise on the stave is comparable to the output noise on the single module box (only 5% difference). The gain on the stave is 10% lower than the single module tests. As a consequence the input noise is 10% higher on the stave, which is larger than expected and is an issue under investigation.

The plot in figure 5.13 shows a noise bump in some selected channels. The bump is from channels near the DC-DC converter, another issue discovered during stave

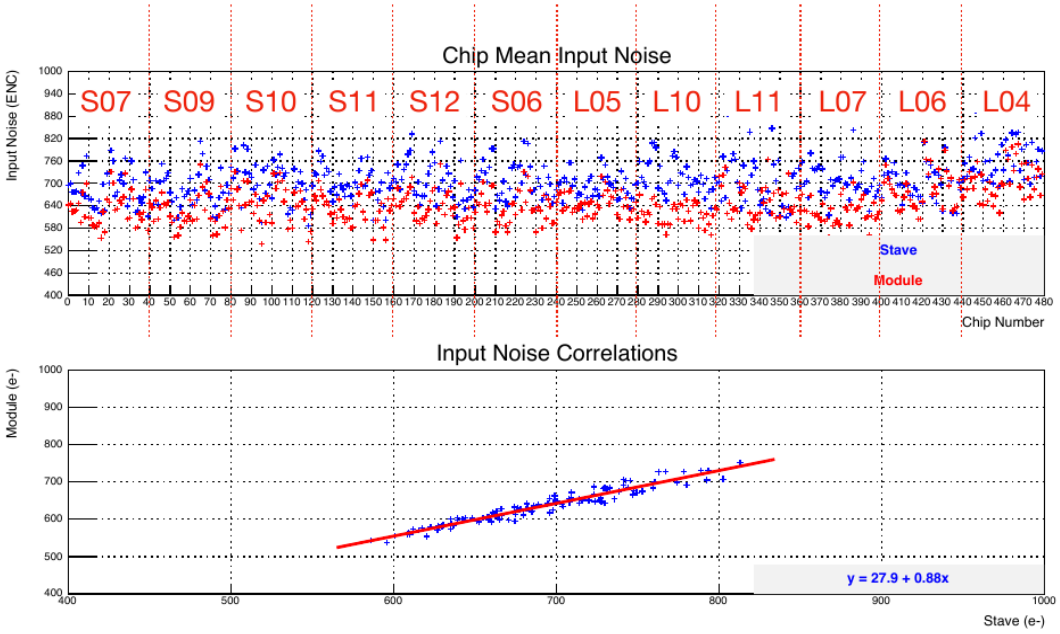


Figure 5.11. Comparison of Input noise between stave and single module tests

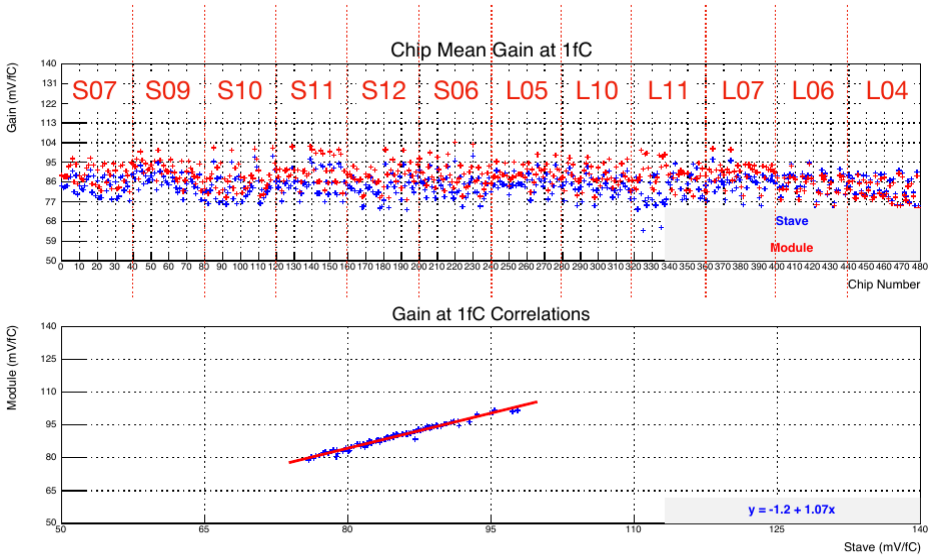


Figure 5.12. Comparison of the gain between stave and single module tests

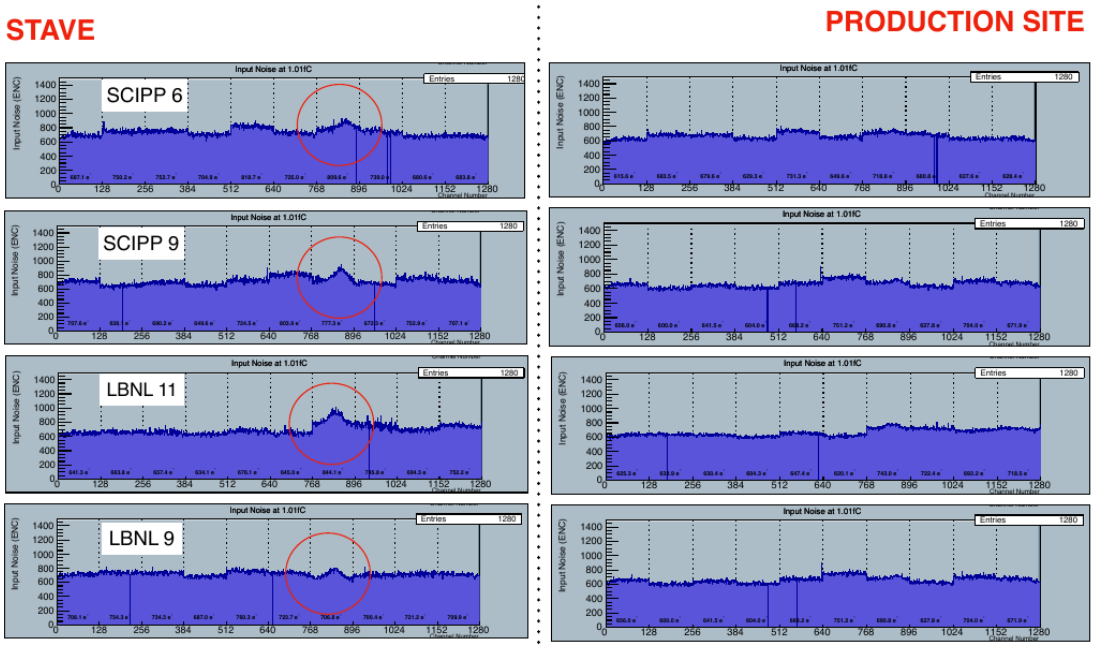


Figure 5.13. Noise increase in the stave test relative to single module tests

testing. This was a recently discovered issue in the design of the Faraday cage around the DC-DC converter which the next design iteration will address. The bump does not appear on the single module tests because eddy currents in the aluminum high voltage plate cancel out the electromagnetic field leaking from the DC-DC converter.

Finally figure 5.15 shows metrology conducted on the sensors. Metrology studies the shape of the sensor at the micrometer scale to see how the shape changes before and after gluing the sensors and determine that the height of the glue is within spec. Its done using a confocal microscope (Keyence LT-9030) shown in figure 5.14. The bowing of the sensor is on the order of 82 μm,and the shape difference is within expected tolerance.

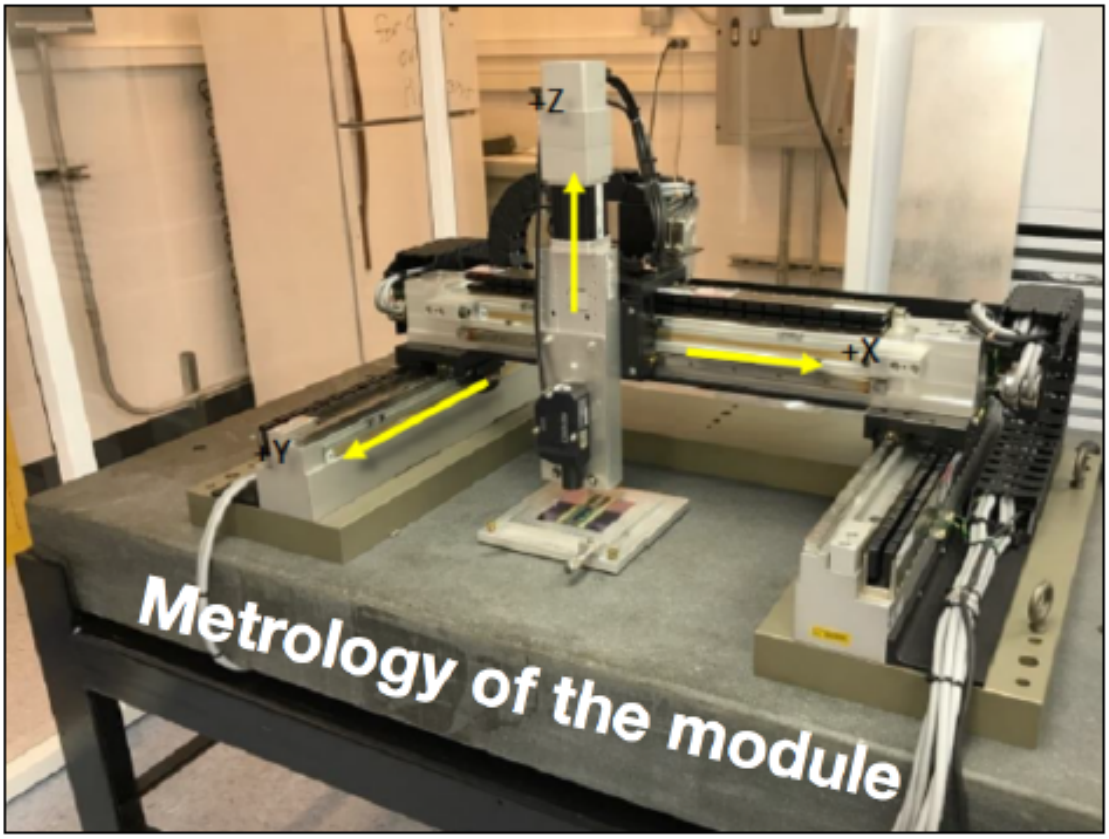


Figure 5.14. Confocal microscope used for module metrology

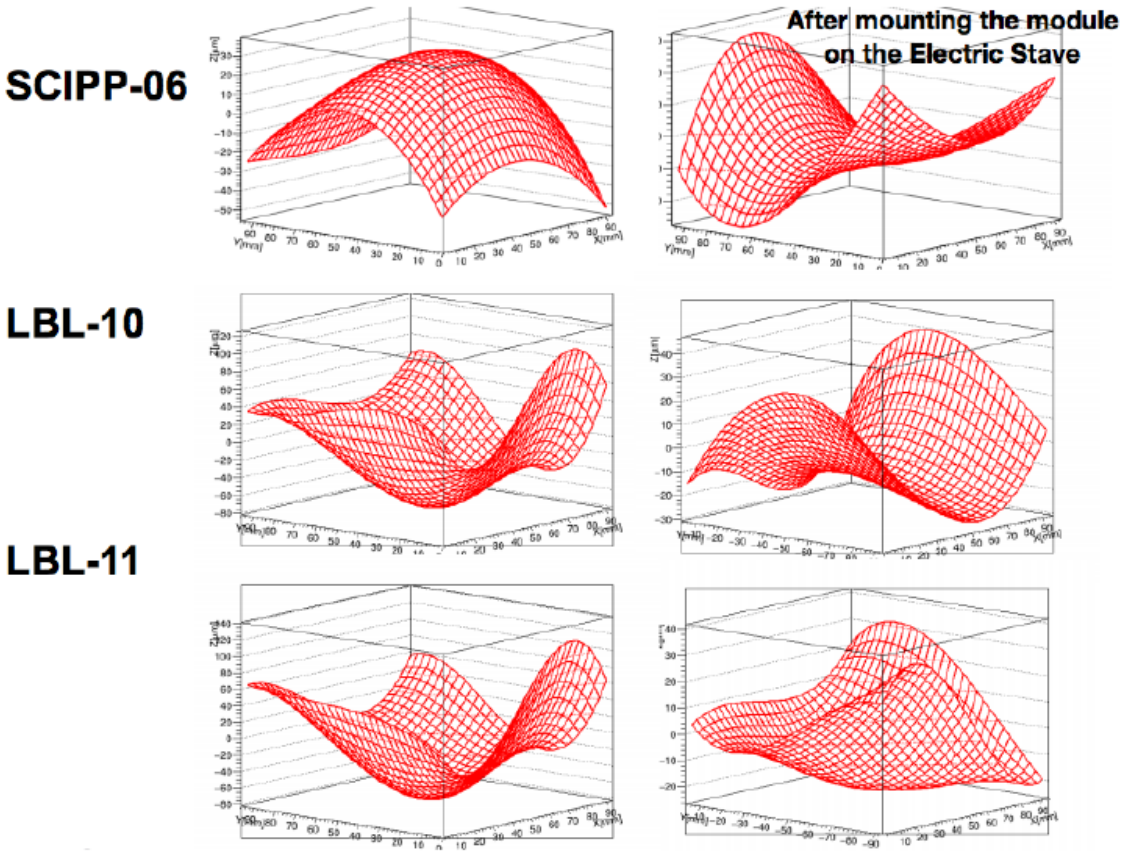


Figure 5.15. Results of module metrology, module bowing is within 100 μm

CHAPTER 6

MODULE QUALITY CONTROL

Quality control for modules is used to ensure high and constant electrical and mechanical quality of the modules and determine if the modules will survive the expected 10 year lifetime in the ATLAS detector. The QC procedure will consist of electrical, mechanical and functional tests that look for sensor deformation, glue shape changes, electrical characteristic of sensors, i.e: I-V and C-V curves and test of the front-end electronic.

The plan for module quality control during production will involve electrical tests, thermal cycling, and HV stability tests. Modules will be cycled from 40° C to -35° C ten times to simulate wear and tear over the lifetime of the detector. Following the ten cycles, the high voltage will be left on for two hours to measure the stability of the current through the sensor. The thermal cycle will begin strobe delay, three point gain and trim range to determine trim DAC values for warm and cold phases of the cycle.

To conduct QC tests, BNL and Harvard University developed a cold box shown in figure 6.1 to test multiple modules at once and thermal cycle them. The cold box contains four cooling chucks similar to the cooling chuck used in the single module box.

The electrical connections to the box are outlines in figure 6.2. The four modules share a common high voltage connections, a decision driven by limited Keithley HV supplies. Module high voltage can be controlled by the HV multiplexer on the power board. The tests described below used the raspberry pi I^2C buses to control the

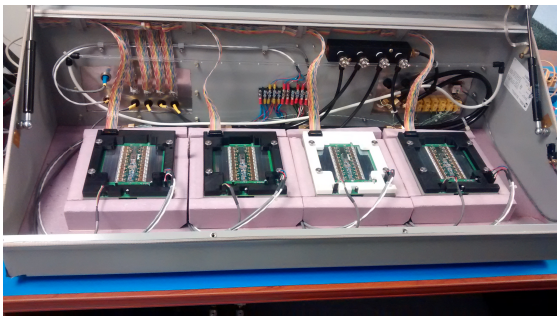
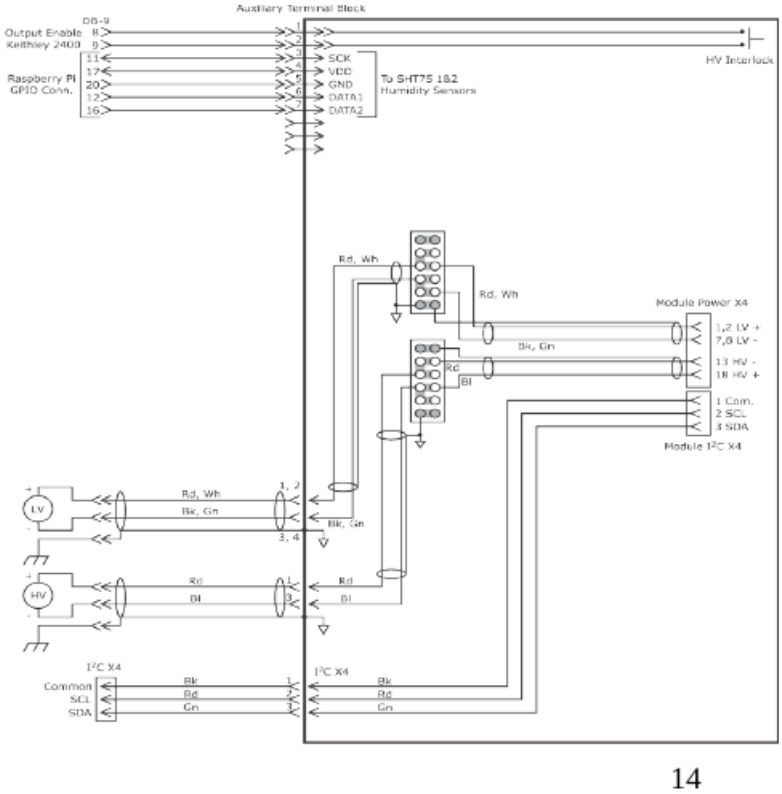


Figure 6.1. BNL/Harvard four module cold box

AMAC and read out the sensors. The newer version of the AMAC, uses a costume I^2C like protocol that will be read out by the FPGA. The lid of the box had an HV interlock to kill high voltage if the box is opened. The temperature for the first tests was read out by thermocouples using a TC-08 that were in thermal contact with the sensor. During production the plan is to use thermistors on the hybrids to measure the temperature, read out by the AMAC. Finally, the box was grounded to be a Faraday cage (but is made up of steel which has poor conductivity).

The cooling system will be controlled by a Raspberry Pi, a small single board computer, acting as a server with the user interacting remotely. The Raspberry Pi will be connected to an analog to digital converter to read out environmental sensors from the Raspberry Pi's general input output (GPIO) pins. A sensor will be needed to measure the humidity inside the box, the ambient temperature, the nitrogen flow rate and the vacuum pressure in the chucks. The Pi will use USB to serial connections to communicate with the slow controls and the chiller. I developed the software framework that will monitor the environmental sensors as well as control the chiller and the power supplies and run data acquisition tests in synchronization with the temperature sensors on the strip module. Figure 6.3 shows an outline for the the software framework described above. One of the biggest challenges we will face is to set up the system to run the tests sufficiently quickly to test several modules per day. There are many choke points that can slow down the testing cycle such as the time



14

Figure 6.2. Schematic of electrical connections for the four module box

for the chiller to cool down and how fast the high voltage can be ramped to determine the IV curve.

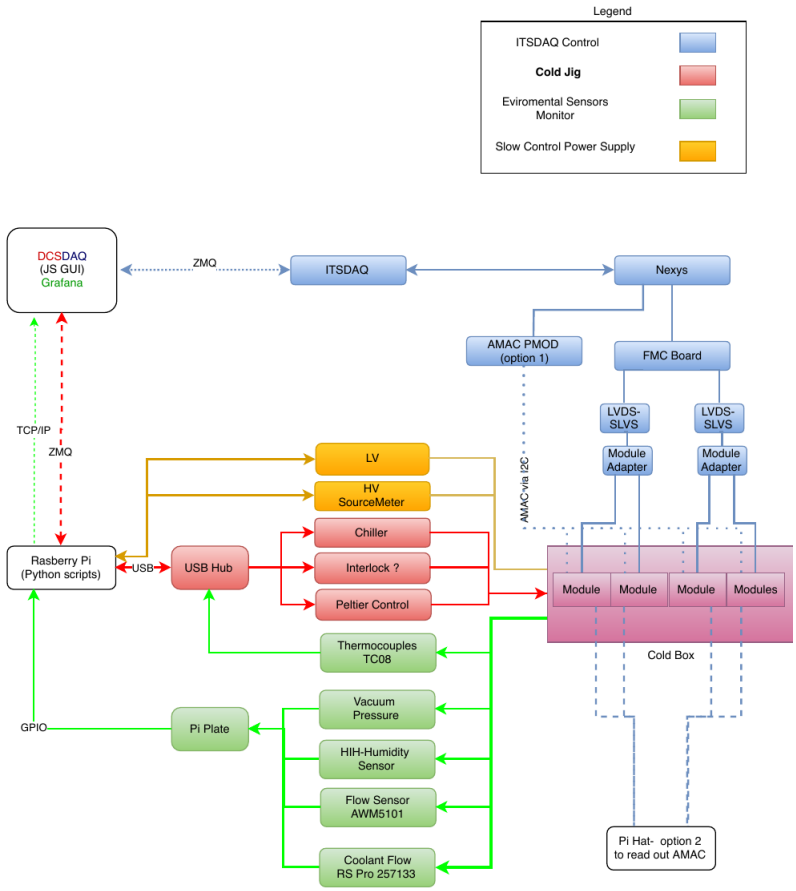


Figure 6.3. Outline of Software Framework for QC Tests

The setup for the data acquisition system is shown in figure 6.4. The enclosure was 3D printed at BNL. On the top left is a Nexys FPGA connected to the central computer through an ethernet connection. The DAQ commands are run through the FPGA and data is histogramed then sent to the computer through a UDP protocol. On the bottom is the raspberry pi, with a I^2C multiplexer to run the DCS. Finally on the left are two Low-voltage differential signaling (LVDS) to Scalable Low-Voltage Signaling (SLVS) converter, each one reading out two modules (so four modules in total).

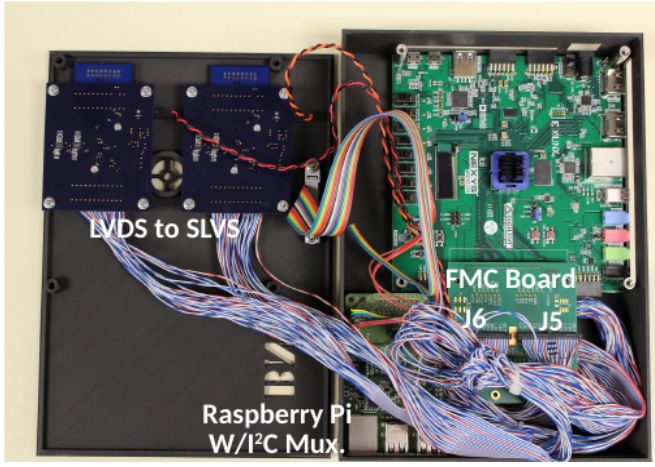


Figure 6.4. Setup for cold box DAQ

The first test of the thermal cycle with two modules is shown in figures 6.5 and 6.6. Figure 6.5 shows the temperatures over four cycles, with a time range done in the first cycle in the hot and cold portion, and the trim range configurations used in the following cycles. Figures 6.6 shows the estimated time for DAQ tests in each portion of the cycle. Strobe delay takes 10 minutes, three point gain 20 minutes, trimRange 35 minutes and the low statistics 3 point gain around 5 minutes. The goal is to run 12 cycles over 10 hours. Figure 6.7 shows the temperature rate for cooldown and warmup. In both cases it's around 20 minutes, so to keep the whole test under 12 hours, the tests in total should take about 10 minutes, which necessitates low statistics tests.

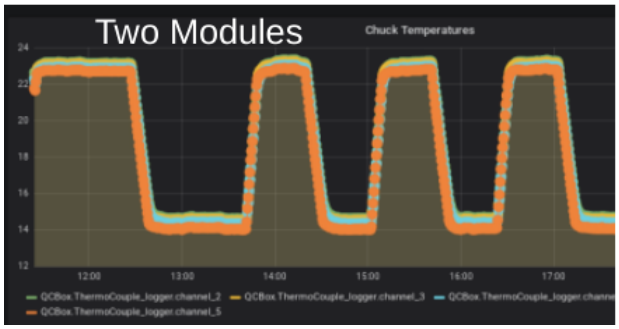


Figure 6.5. Temperature sensor outputs from the first DAQ test

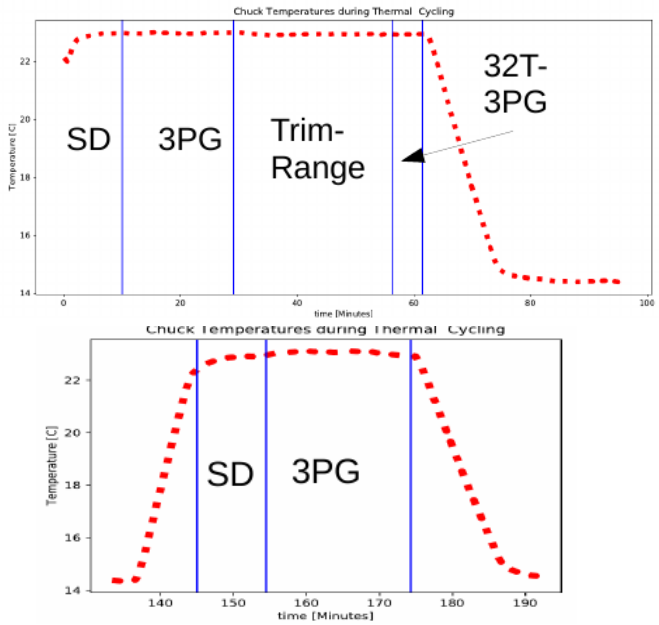


Figure 6.6. Top plot is the first cycle where the trim range was run, the second plot is the second thermal cycle where only strobe delay and three point gain were run. The times each test took is on the x axis. SD is strobe delay, 3PG is thee point gain, 32T-3PG is three point gain with low statistics.

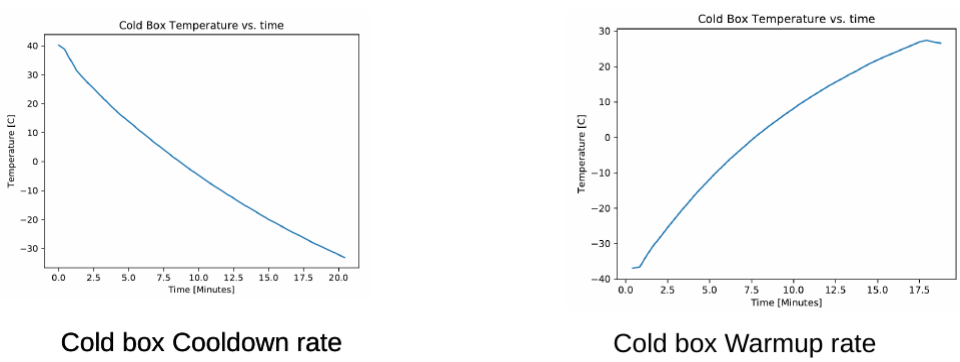


Figure 6.7. Cooldown and warm up rate for the four module box with the SPS chiller. Warmup is slightly faster as expected but its close to 20 minutes each.

figure 6.8 is the sensor recordings from the first QC run with four module in the BNL box with ten thermal cycles and an HV stability test. The top left plot is the air temperature, which gets as cold as 0° C. The two plots on the bottom left are the thermocouple temperature in thermal contact with the cooling chuck, used to automate the thermal cycle. The top right plot is the box relative humidity, which consistently stayed below 5%. The goal is less than 10% humidity to avoid frosting. The middle left plot is the high voltage stability test. Finally the bottom right plot is vacuum pressure used to hold down the chuck. The vacuum pressure goes down by approximately 2 psi at cold temperature because the sensor shape deforms and lifts off the vacuum plates slightly. Figure 6.9 shows the vacuum pressure vs. temperature. There's a hysteresis effect where the cooldown shape deformation is different than the warm up shape deformation.

Figure 6.8. Sensor recordings from thermal cycle

Figure 6.10 shows the results of the DAQ tests for the first thermal cycle for a single module.

Figures 6.11 and 6.12 shows the input and output noise distribution of all the modules in the DAQ tested conducted in the warm part of the cycle vs. the cold part

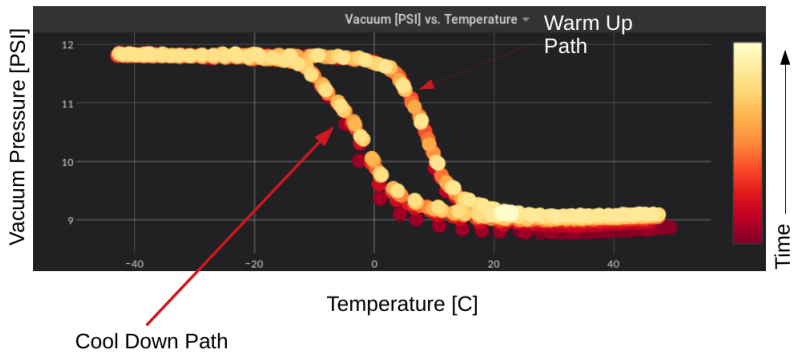
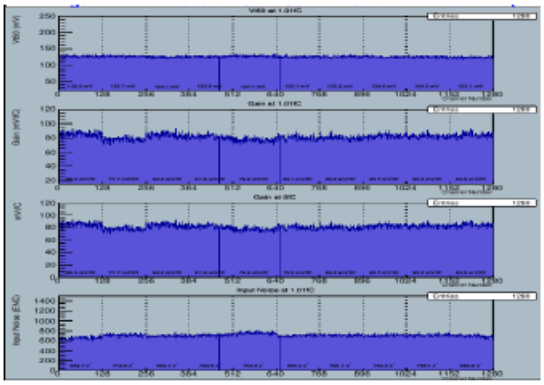
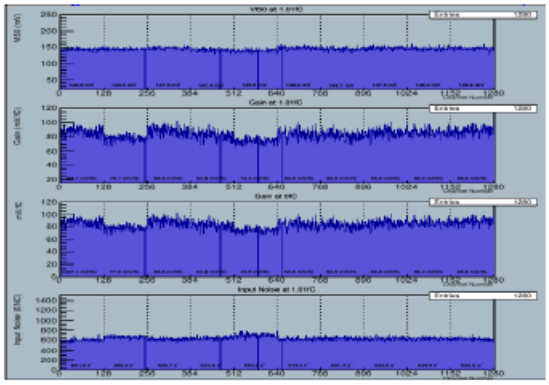


Figure 6.9. Vacuum pressure vs. temperature for vacuum chucks in the four module box. The path on the left is the plot as the box cools down, the path on the right is the plot as the box warms up

of the cycle. The distribution of noise looks healthy in both cases, with lower noise from tests conducted when the modules are cold, as expected. Finally figure 6.13 shows the average noise of all the modules over the course of the 10 thermal cycles.



Three Point Gain at +40C



Three Point Gain at -35 C

Figure 6.10. DAQ test result for four module box with all four modules and the full thermal cycle at the warm and cool part of the cycle. The top plot is the Vt50 of each channel, the two middle plots is the estimate of the gain, the bottom plot is input noise

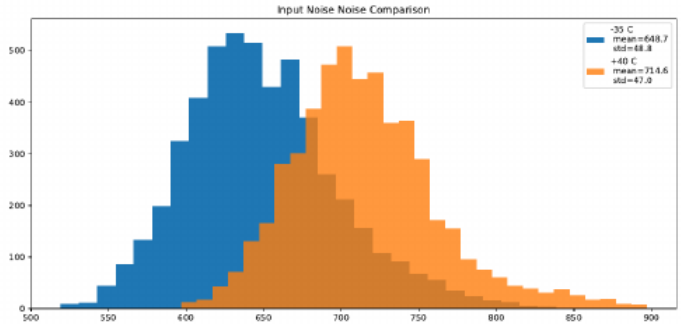


Figure 6.11. Distribution of the input noise for all four modules in the warm and cold part of the thermal cycle. Room temperature expectations are 600-700 ENC, so the noise in the warm up cycle

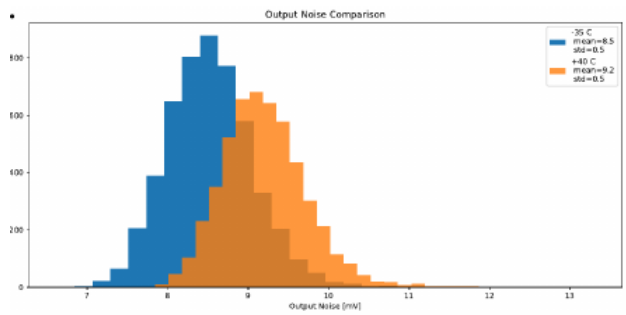


Figure 6.12. Distribution of the output noise for all four modules in the warm and cold part of the thermal cycle

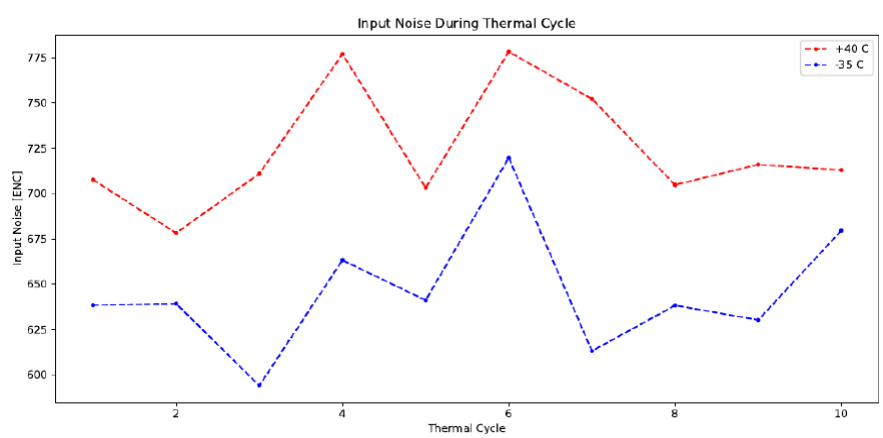


Figure 6.13. Average Input noise for all modules for each of the ten thermal cycles

CHAPTER 7

ABCSTAR IRRADIATION

One of the poorly understood effects of radiation on 130nm CMOS technology is that when exposed to ionizing radiation, they exhibit an increase in the leakage current, followed by a gradual decrease towards baseline after a peak, making a 'bump' shape as shown in figure 7.1. The understanding of this Total Ionizing Dose (TID) induced leakage current is crucial to the design of the ATLAS ITk Strip Detector due to its implication on thermal management and mechanical stability, which relies heavily on the expected power consumption of the modules. Characterizing the TID bump helps determine the safety factor needed for module power consumption [3] .

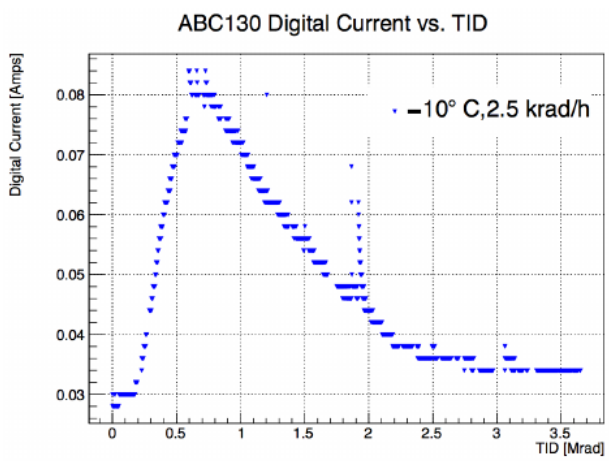


Figure 7.1. Bump in the digital current vs. total ionizing dose

In the ATLAS detector, the ABCstar chips must withstand dose rates of total ionizing radiation in the range of 0.6 krad·h^{-1} - 2.5 krad·h^{-1} from LHC collisions

and will be operated at sub-zero temperatures. To characterize the effects of TID on the leakage current at typical HL-LHC conditions, the leakage current from four ABCstar chips was measured utilizing the ^{60}Co gamma source at BNL in a controlled temperature environment.

Each ABCStar was read out with a single chip board set inside a container designed for the experiment. The container is insulated with Styrofoam and contained a thermo-electric in thermal contact with the chip which was kept at -10° C for three of the boxes and 0° C for the fourth box. The orientation of the boxes is shown in figure 7.3. The closest to the radioactive source received 2.5 krad·h^{-1} dosage, the second closest received 1.1 krad·h^{-1} and the two furthest away each received 0.6 krad·h^{-1} dosage.

Figure 7.2. Setup of the four ABCStar boxes relative to the Co-60 source.

Temperature and humidity were measured with an SHT85 sensor, which is not radiation proof. To protect the sensors, a radiation resistant extension in the box is made to house the sensors while being in environmental contact with the rest of the radiation box. The humidity is kept low by circulating nitrogen through the box and the extension that contains the sensors. I wrote software controls to readout the temperature and humidity sensor through Raspberry pi I2C, control the thermo-electric chiller through serial connection, log the digital and analogue current from

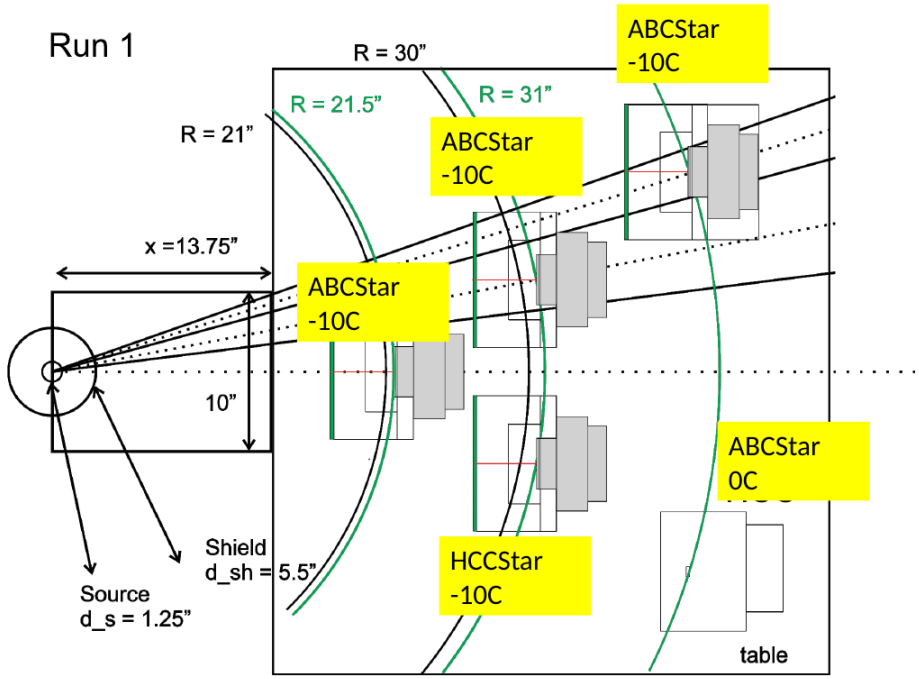


Figure 7.3. Orientation of each ABCStar chip relative to the source (The HCCStar is a seperate co-operating experiment)

the ABCStar and provide software interlock. The analogue and digital current of the ABCStar are measured through the single chip board. The software interlock shuts the power to the chip if humidity persists above ten percent. The box design and the electrical connections are outlined in figure 7.4

7.1 Physics Model

A simple field effect transistor (FET) is shown in figure 7.5. It consists of two positively doped ends, called the source and the drain, separated by a negatively doped gate. Absent an applied voltage, a depletion layer between the positive and negatively doped portions prevents any current from following. When a voltage is applied to the gate, the depletion layer decreases and the current can flow between the drain and source. In this way transistors can act as switches to turn current on

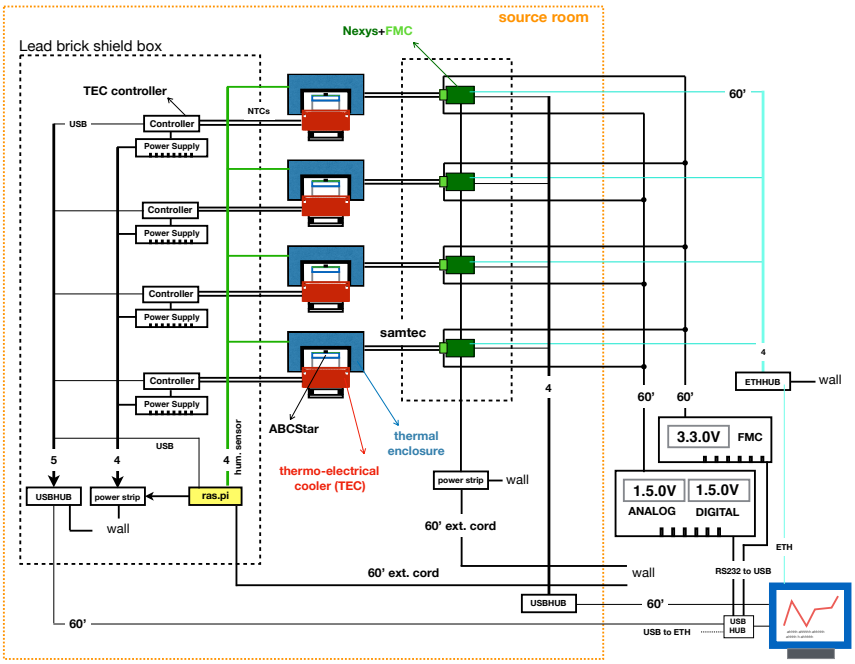


Figure 7.4. Schematics of the electrical connections between the boxes containing the ABCStar chips, DAQ readout, thermo-electric chiller (all in the source room) with the controls outside the source room

and off in digital electronics [3]. In the CMOS 130 nm chip, the array of transistors are isolated by a shallow silicon oxide layer, called the shallow trench isolation (STI).

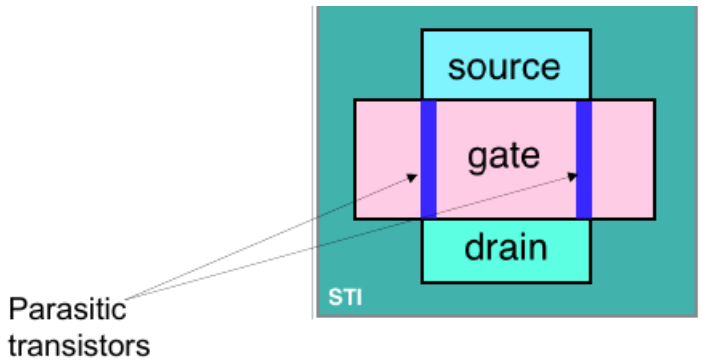


Figure 7.5. Diagram of parasitic transistors in a CMOSFET on an isolation layer

To analyze the results we worked on a simplified physics model of the system. Ionizing radiation in the silicon create electron / hole pairs in the atomic lattice. The holes act like positive states and can drift. Holes created through radiation migrate to the silicon oxide interface in the isolation layer and create a path for the current to flow between source and drain that bypasses the gate causing leakage current. This path is called a parasitic transistor. The parasitic transistors are eventually canceled out after some time by electrons which build up on the opposite side of the silicon oxide interface that are trapped from radiation induced traps in the silicon band gap. This gives the bump shape to the leakage current.This is outlined in figure 7.6 where the positive states float and a trapped in the interface

In our model, the leakage current in these chips is the number of positive states (holes) in the interface (n_h) minus the number of negative states (n_e) in the interface, minus the threshold voltage (n_{thr}), squared, as shown in the equation below.

$$\begin{aligned} I_{Leakage} &= k \cdot (n_h - n_e - n_{thr})^2 \quad if \quad n_h - n_e - n_{thr} > 0 \\ I_{Leakage} &= 0 \quad else \end{aligned} \tag{7.1}$$

$$\frac{dn_h}{dt} = A_h \cdot (N_h - n_h) - P \cdot n_h \tag{7.2}$$

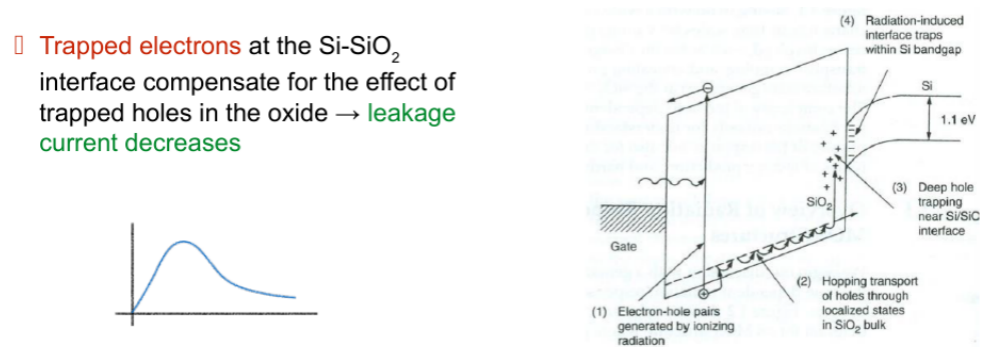


Figure 7.6. Diagram of the process of positive states trapped in the interface of silicon/ silicon oxide isolation, creating parasitic transistors

$$\frac{dn_e}{dt} = A_e \cdot (N_e - n_e) \tag{7.3}$$

$$\begin{aligned} I_{Leakage} &= k \cdot \left(\tfrac{N_h A_h}{A_h + P}(1 - e^{-x(A_h+P)}) - A_e(1 - e^{-x \cdot A_e}) - n_{thr}\right)^2 \quad if \quad n_h - n_e - n_{thr} > 0 \\ I_{Leakage} &= 0 \quad else \end{aligned} \tag{7.4}$$

The differential equations in equation 2 describe the growth in positive and negative states over time. The first equation describes the rate that trapped positive states increase. The first term rate of growth, the rate of hole creation times the number of available hole traps. N_h is the number of hole traps at the silicon oxide interface, n_h is the number of holes currently trapped, P_h is annealing probability and A_h is overall normalization. The second term is the annealing term where trapped holes get released due to thermal excitation and quantum tunneling. The second equation for the growth in negative states is the same except without the annealing term. Experiments conducted in Rutherford Appleton Lab (RAL) (figure 7.7) show that the pre-irradiated chips don't exhibit the TID bump. The experiments at RAL baked ABC Chips in the oven at 80 degrees for 3 months then irradiated the chips, shown in the colored dashed lines. As the plot shows, the TID bump did not return,

this leads us to believe that negative states don't anneal. The differential equations for the growth in positive and negative charges is analytically solvable, the solution is shown in the equation above. The equation has five unknown parameters that we can study through fits to the ABCStar data.

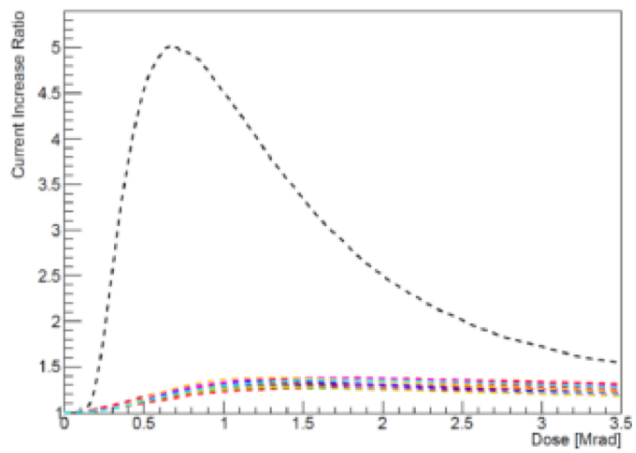


Figure 7.7. Results of RAL annealing studies. The bump is the initial TID bump, the other lines show the bump doesn't return

7.2 Results

The results of the study conducted in 2018 on the ABCStar chips are shown in figure 7.8. In addition, figure 7.9 show studies done on the prototype ABC130 chips using similar setups in 2016 and 2017. The current resolution of the 2018 study is noticeably better due to more precise instruments.

The results are consistent with expectations. The larger dose has larger total leakage current at the peak of the TID bump. The total dose where the bump peaks, around 600 Mrad, is consistent between the different dose and temperature levels. Finally the higher temperature shows lower total leakage current at the peak for the same radiation dose, because of more frequent annealing of the positive states from higher temperature.

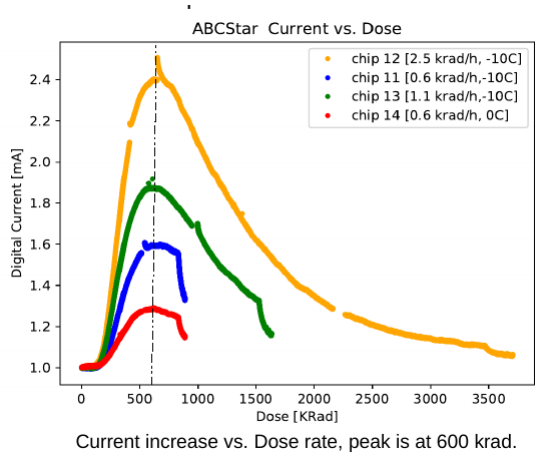


Figure 7.8. ABCStar irradation results of the digital leakage current vs. total ionizing dose

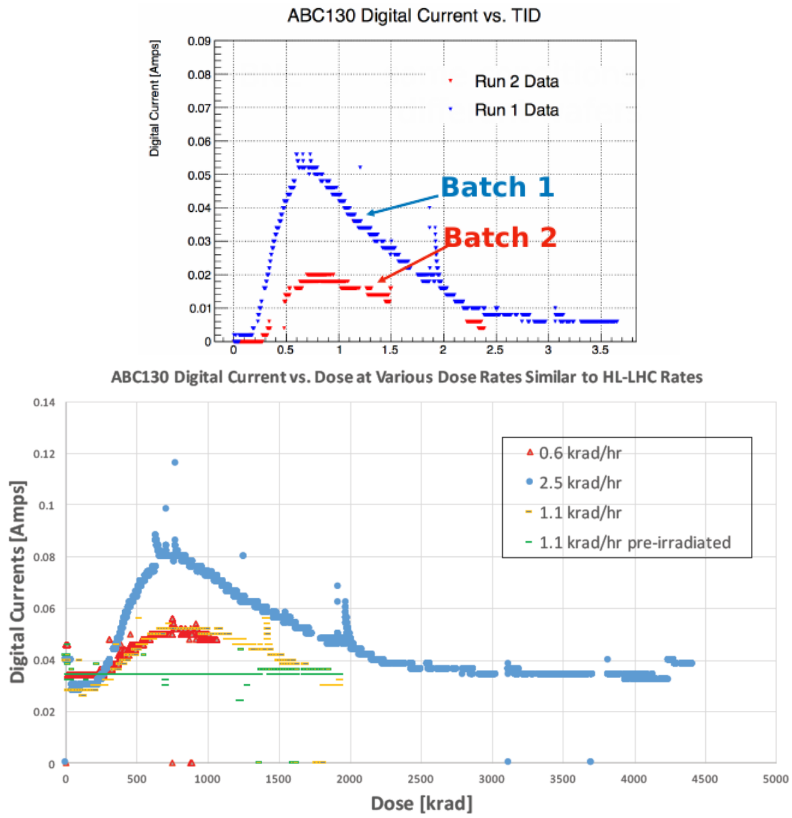


Figure 7.9. ABC130 irradation results from 2016 (bottom) and 2017 (top) campaigns

The fits and parameters are shown in 7.10 and table 7.1. The interpretation of the fit parameters is still ongoing. We hope that this data combined with data from RAL ABCStar data will allow us to characterize the curves as functions of temperature and dose rate.

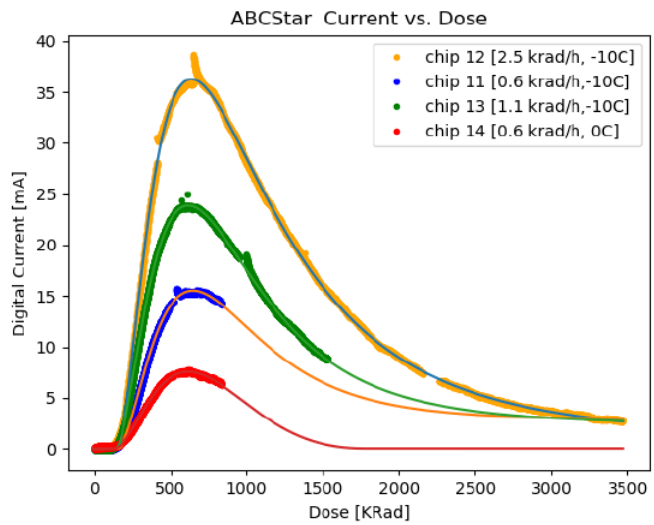


Figure 7.10. Model fit to the digital Leakage current vs. total ionizing dose data

Fit result

Chip	Dose Rate[krad/h]	Temperature [C]	C5 [√mA]	C1 [√mA]	C2[days]	C3[√mA]	C4[days]
12	2.5	-10	8.12	19.9	3.77	10.6	19.34
13	1.1	-10	5.74	21.02	10.445	13.71	25.88
11	0.6	-10	3.95	396.01	27.74	289.37	28.5
14	0.6	0	2.45	392.7	40.77	392.26	42.18

Table 7.1. Results of the fit parameters for each of the ABCStars irradiated

One of the challenges is significant difference in the total peak between different batches of chips. Figure 7.9 shows the peak for different batches in the ABC130 results. Further studies done in RAL shown in figure 7.11 also show significant differences. The difference in the peak can be up to a factor of two. This adds an addition complication to characterize the bump, as more statistics are needed for each batch.

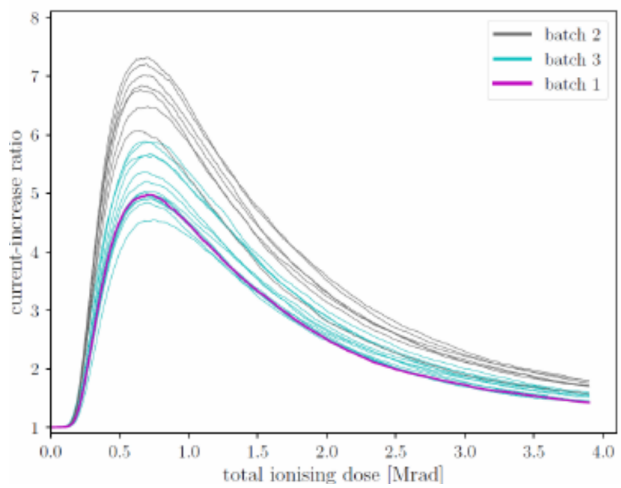


Figure 7.11. ABC130 irradiation results from 2016 (bottom) and 2017 (top) campaigns

Lastly, we looked at the annealing time for positive states. Figure 7.12 shows the decaying current after the source was turned off. Figure 7.13 shows the time constant for the exponential fall off as a function of temperature. The time constants is approximately 1-3 days and inversely proportional to temperature, as expected in the model.

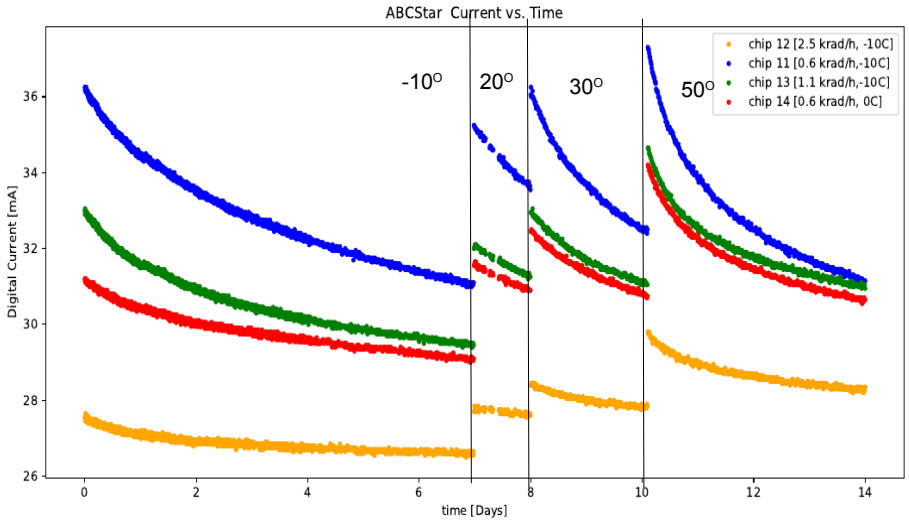


Figure 7.12. Leakage current fall for different temperatures

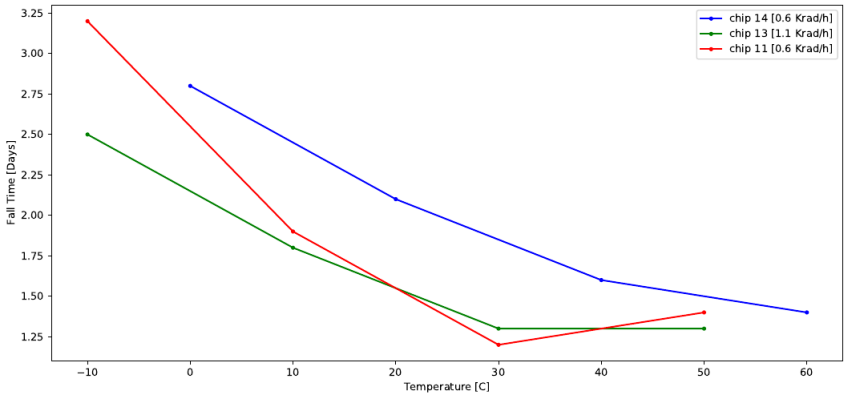


Figure 7.13. TID Fall time vs. Temperature

CHAPTER 8

HWW VBF CROSS SECTION

The analysis presents a measurement of the decay of the standard model Higgs boson to two W bosons with a final state of two opposite sign leptons and two neutrinos. A standard model Higgs boson candidate is identified by starting from data with two isolated charged opposite-sign and different flavor leptons ($e\mu$, μe), and missing transverse energy MET due to the presence of neutrinos in the final state. Kinematic properties of the dilepton system are also used to identify the SM Higgs boson. The SM Higgs boson with spin zero decays in two W bosons with spins anti-aligned. This spin configuration propagates to the W's decay products through the conservation of angular momentum and the structure of the weak interactions. Due to the spin correlations, the angle between the charged leptons is relatively small. The VBF production mode is identified by requiring at least two associated highly energetic jets (*tagging* jets) originating from the initial state quarks. It is distinguished from other production modes, most notably the gluon-gluon fusion (ggF) mode, by kinematic requirements. For example a veto on the W leptons outside the rapidity region defined by the two taggging jets (outside lepton veto or OLV), and a veto on jets produced inside the rapidity region defined by the two taggging jets (central jet veto or CJV).

The main background processes that contribute to the $h \to WW \to l\nu l\nu$ analysis are: $t\bar{t}$, WW, WZ, $W\gamma*$, ZZ, W + jets and $Z \to ee, \mu\mu, \tau\tau$. Top and WW have two W bosons in the final state, similar to the signal. The presence of jets, especially those tagged as b-jets, is used to discriminate against top. Spin correlation kinematics

are used to reject WW. Two or more real leptons are produced by the $Z\tau\tau$ and non-WW electroweak background (WZ, $W\gamma*$, ZZ). These have a smaller cross section but more signal-like kinematics because of the softer subleading lepton, particularly in the case of $W\gamma*$. The W + jets process has a high cross section and passes the signal region selection because a jet is mis-reconstructed as an isolated lepton [13].

8.0.1 Motivations

There are two predominant Higgs production channels in the LHC: gluon-gluon Fusion (ggF) and Vector Boson Fusion (VBF), shown in Figure 8.1. The ggF process is the dominant Higgs production mode and consists of two gluons fusing into a quark loop that then produces a Higgs. The Higgs-quark coupling drives this production mode. VBF is the secondary production channel and forms a Higgs when two quarks each radiate a massive weak boson. These bosons fuse together to create a Higgs, thus this decay is mediated by HWW coupling. The two interacting quarks hadronize to form two jets which exhibit a distinct energetic signature in the forward and backward regions. Physicists at ATLAS and CMS have measured the strength of the Higgs couplings to other particles and thus far all are compatible with the SM Higgs boson. However, the experimental uncertainty on the VBF production cross-section is about 30% in Run-1 analyses. In Run-2 the sensitivity of the measurement to anomalies will be improved due to the larger data set, more Monte Carlo statistics, and more sophisticated analysis techniques including multivariate analysis tools such as deep neural networks and boosted decision trees.

This analysis looks at the the opposite flavor channel to avoids a large background from $Z \to ee$ and $Z \to \mu\mu$ processes. These backgrounds require tight cuts on the invariant dilepton mass (Mll) which limits total possible signal events in the signal region [7].

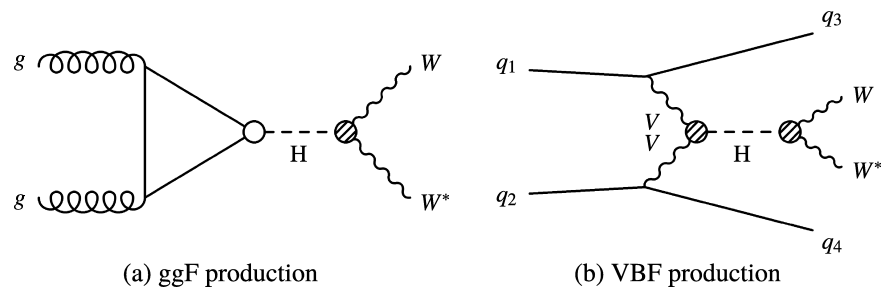


Figure 8.1. Feynamn diagrams of Higgs production to WW decay

8.0.2 Analysis Strategy Overview

The full run-2 data set is analyzed, corresponding to the 2015-2018 data sets at $\sqrt{s} = 13$ TeV proton-proton centre-of-mass energy, and to an integrated luminosity of 139 fb^{-1}. A pre-selection is applied on recorded events to select Higgs boson production processes in which the Higgs boson decays into a W boson pair that in turn decays into the fully leptonic channel with two opposite flavor leptons, i.e. an electron and a muon, and neutrinos in the final state.

The main background processes $t\bar{t}$, single-top production, dibosons, Z+jets, also referred to as Drell-Yan (DY), fakes, where jets are misreconstructed as leptons such W+jets and QCD multijet production, as well as other Higgs production modes, such as ggF, VH, ttH. Major backgrounds ($t\bar{t}$,, single-top production, dibosons, Z+jets, fakes and ggF Higgs boson production), are estimated in a data driven way, while all other backgrounds (mainly Higgs production modes other than ggF and VBF) are estimated by MC.

After preselection several additional selections are applied to define a VBF signal region (SR), and several control regions (CRs) that are enriched in the major background processes: DY-CR and ggF-CR. The top backgrounds are estimated from a fit within the signal region. In the SR kinematic requirements such as outside lepton veto (OLV) and central jet veto (CJV) are applied to select VBF candidate events, the DY-CR exploits kinematic variables sensitive to the different spin configuration

of DY and Higgs decay events. In addition, the contamination from the fake background in SR and CRs is estimated in data using a dedicated sample enriched in fake leptons. All the signal, control and validation regions are described in tables 8.1 and 8.2, where validation regions are used to determine if the Monte Carlo modeling matches the data.

To extract the fiducial and differential cross sections, a simultaneous fit of SR and CRs to data is carried out to estimate the amount of $t\bar{t}$, VV,DY and ggF background events based on data as well as to extract the signal yield. In the fit the signal strengths of the VBF signal, diboson+top ($t\bar{t} + Wt$), DY and ggF processes are left floating. The signal strength of top is considered in combination with VV, since it was determined that there's not enough discrimination power between them. The signal yield is extracted by multiplying the signal strength by the theoretical cross section and is then corrected for experimental efficiency (C-factor) and integrated luminosity (L) to finally extract the signal cross section that is compared to theoretical predictions.

Boosted Decision Trees (BDTs) are used to discriminate between the various processes in the statistical fit. For the signal region, a two-dimensional BDT is trained to discriminate the signal from $t\bar{t}$ and diboson (VBF bdt) and discriminate $t\bar{t}$ and diboson from all other processes (Top+WW bdt).The signal region is then split into two parts, signal region 1 with VBF bdt > 0.5 and signal region 2 with VBF bdt < 0.5. In signal region 1 the fit is done over a projection onto the VBF bdt, in the signal region 2 the fit is done over a projection onto the Top+WW bdt.

For the ggF background estimation one CR is used and a dedicated BDT (ggF BDT) is trained to increase the separation between ggF and all other processes, including VBF. The ggF BDT output distribution is used in the simultaneous fit, i.e. no cut is applied on this the ggF BDT output. Finally the fit in the Drell-Yan control

region is done using the transverse Higgs mass, m_T, which was determined to have sufficient discriminatory power compared to a dedicated Zjets bdt.

For each differential cross section measurement a different BDT is trained by removing from the list of input variables the measured kinematic variable and those that are highly correlated to it, called N-1 bdts in following sections. Experimental and theoretical uncertainties are propagated into the simultaneous fit as nuisance parameters and are profiled by the fit.

VBF SR	**Z+jets CR**	**Top VR**	***WW* VR**
2-lepton	2-lepton	2-lepton	2-lepton
$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)
b-veto (< 1)	b-veto (< 1)	**b-veto (>= 1)**	b-veto (< 1)
M_{ll} > 10 GeV	M_{ll} > 10 GeV	M_{ll} > 10 GeV	M_{ll} > 10 GeV
CJV = True (20 GeV jets)	CJV = True (20 GeV jets)	CJV = True (20 GeV jets)	CJV = True (20 GeV jets)
OLV = True	OLV = True	OLV = True	$\boldsymbol{m_T}$ **> 130 GeV**
$Z \rightarrow \tau\tau$ veto	$\boldsymbol{Z \rightarrow \tau\tau}$ **selection**	$Z \rightarrow \tau\tau$ veto	$\boldsymbol{m_{T2}}$ **> 160 GeV**
$\Delta\phi_{ll}$ < 1.4			
M_{jj} > 450 GeV	M_{jj} > 450 GeV		
Δy_{jj} > 2.1	$\boldsymbol{M_{ll}}$ **< 80 GeV**		

Table 8.1. Kinematic cuts on the defined phase space regions in the analysis. Definitions of m_T and m_T^2 are given the following chapters on observables.

***W*+jets CR**	**ggF CR1**	**ggF VR2**	**ggF VR3**
1 ID lepton, 1 anti-ID	2-lepton	2-lepton	2-lepton
$\boldsymbol{p_{\mathrm{T}}^{\mathrm{miss}}}$ **> 20 GeV**			
$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	$\geq$ 2-jets (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)	**< 2 jets** (p_{T} >30 GeV, $\lvert\eta\rvert$ < 4.5)
b-veto (< 1)	b-veto (< 1)	b-veto (< 1)	b-veto (< 1)
M_{ll} > 10 GeV	M_{ll} > 10 GeV	M_{ll} > 10 GeV	M_{ll} > 10 GeV
CJV = True (20 GeV)	**CJV = False and OLV = True** *or*	**CJV = False and OLV = False**	
OLV = True	**CJV = True and OLV = False**		
$Z \rightarrow \tau\tau$ veto	$Z \rightarrow \tau\tau$ veto	$Z \rightarrow \tau\tau$ veto	$Z \rightarrow \tau\tau$ veto
	$\Delta\phi_{ll}$ < 1.4	$\Delta\phi_{ll}$ < 1.4	$\Delta\phi_{ll}$ < 1.4
M_{jj} > 450 GeV			
Δy_{jj} > 2.1			

Table 8.2. Kinematic cuts on the defined phase space regions in the analysis. The requirements that are modified with respect to the SR are highlighted in bold.

Differential cross sections are measured as a function of kinematic observables unfolded to particle level in the fiducial phase defined in the next chapter. Here follows the list of 13 distributions that are measured and unfolded to particle level:

- Higgs boson transverse momentum (P_H^T)
- Leading and sub-leading jet transverse momenta (P_{j0}^T, P_{j1}^T)
- Leading and sub-leading lepton transverse momenta (P_{l0}^T, P_{l1}^T)
- di-lepton transverse momentum (P_{ll}^T)
- di-lepton rapidity ($|\Delta y_{ll}|$)
- di-jet rapidity (between the two leading jets) ($|\Delta y_{jj}|$)
- di-lepton azimuthal angle ($|\Delta\phi_{ll}|$)
- di-jet azimuthal angle ($\Delta\phi_{jj}$)
- di-lepton invariant mass (M_{ll})
- di-jet invariant mass (M_{jj})
- $\cos(\theta^*)$=$\tanh((\eta^+ - \eta^+)/2)$ defined by the two leading charged lepton directions relative to the beam direction in a frame where the two leptons are back-to-back in the $r - \theta$ plane.

The distributions of the observables are discussed in the section on final results. Potential new physics become apparent in high mass and high momentum region of the observables.

Finally the detector level cross section estimate is unfolded to particle level, meaning effects from detector inefficiencies and migrations between detector and fiducial level bins is accounted for. The unfolding is implemented in the same step as the simultaneous fit, by embedding migration matrices into the likelihood definition. The binning of the unfolded distributions are optimized based on the data statistics, efficiency, purity, systematic contributions, model bias as well as to match the bin boundaries of other Higgs and SM electroweak differential measurements.

The particle-level differential cross sections as well as the fiducial cross section are compared to theoretical predictions calculated at different orders of perturbative QCD and electroweak corrections.

CHAPTER 9

PHYSICS OBJECTS

The analysis uses "Physics objects", high level objects corresponding to physical processes that are reconstructed from raw detector data. The primary physics objects of interest are final state particles of $H \rightarrow WW* \rightarrow l\nu l\nu$,are electrons, muons and neutrinos. The VBF production mode also has two associated jets in the final state. Neutrinos are invisble to the detector, so have to be indirectly measured by missing transverse momentum, abbreviated MET or E^T_{Miss}. Tracks are reconstructed in the inner detector, and particles are identified in the Calorimeter. Photons have no tracks in the Inner Detector and therefore a defined strictly by energy deposits in the calorimeter. Photons deposit energy in calorimeter through pari production.

9.1 Tracking

Track reconstruction through the inner detector is essential to identifying physics objects in the subsequent detectors. The track of a particle is described with five parameters,

- d_0 The closest distance of the track to the primary vertex from the pp collision in the tranverse plane.
- z_0 the closest distance of the track in the longitudinal plane
- ϕ the angle of the track in the transverse plane
- θ the angle of the track in the longitudinal plane

- $\frac{q}{p_T}$ The charge of the particle over the transverse momentum

Due to the cylindrical symmetry of the detector, only the traverse and longitudinal impact parameters need to be specified. The momentum of the track is measured by the radius of curvature in the ID due to the solenoidal magnetic field [64].

9.2 Leptons

Leptons are required to originate from a primary vertex that has at least two associated tracks, with a P_T of at least 500 MeV. In the case of multiple primary vertices, the one with the largest total track P_T, defined as the square sum of all associated track momenta, is chosen. The selection criteria described in this section are chosen to reduce W+jets and QCD backgrounds in the analysis, as well as lepton isolation requirements. Isolation describes the amount of detector activity in the vicinity of the lepton candidate.

Finally at the preselection level, the two leptons in the event must have opposite sign and opposite flavor ($e\mu, \mu e$). Opposite sign veto reduces background from diboson processes.

9.3 Electrons

Electrons are identified in the electromagnetic (EM) calorimeters from clusters of energy deposits. Electrons lose energy due to bremsstrahlung radiation which cascades into a shower of particles that deposit characteristic cluster of energy in the calorimeters which are used to identify electrons. The electrons identified in the EM calorimeter are matched to a track in the inner detector, to create combined electrons. Combined electron identification has four possible working points, very loose, loose, medium and tight. The working point is chosen to minimize mid-identification while having sufficient efficiency. Further, an overlap removal algorithm is applied to overlapping clusters in the calorimeter.

Electron tracks are chosen so that the significance of traverse impact parameter, $\frac{|d_0|}{\sigma_{d_0}}$ is less than 5. The longitudinal impact parameter multiplied by $sin(\theta)$ of the track, $z_0 sin(\theta)$ is less than 0.5 mm.

9.3.1 preselection

Initially, electrons need to pass the Very Loose identification criteria, the traverse and longitudinal impact parameter requirement, have a P_T of at least 10 GeV and $|\eta| < 2.47$.

after overlap removal, electrons must pass the Tight working point for electrons with $P_T < 25$ GeV and medium working point for electrons with $P_T > 25$. The transition region between the barrel and endcap in the EM calorimeter is also excluded, $< 1.37|\eta| < 1.52$. Finally, FCTight isolation working point was chosen. FCTight working point corresponds to a cut on the total energy of the topological cluster reconstructed in the calorimeters, and the sum of the momenta of all the tracks in the cone. The values of these cuts can be found in [].

9.3.2 Identification

Electron identification requires separating electron energy deposits in the EM calorimeter from energy deposits from hadronic jets or converted photons. There are three types of variables used to discriminate: variables belonging to the electron tracks, variables describing the electromagnetic shower from energy loss in the EM Calorimeter, and variables matching the reconstructed EM clusters to the electron track [65].

9.4 Muons

Muons are reconstructed using tracks in the Inner Detector and Muon Spectrometer. An overall fit on the hits from the ID and MS is combined with energy loss in the calorimeters to create Combined Muon objects. As with Electrons, there are

four identification working points: Loose, Medium,Tight, and High-P_T. Each working point has a tradeoff of lower efficiency but higher background rejection[66].

In the initial selection, muons are considered if they have $P_T > 10$ GeV, pass the Loose identification working point and have $|z_0 sin(\theta)| < 1.5$ mm and d_0 significance < 15

After the initial selection, muons must pass the Tight identification working point. The Tight working point has an efficiency, evaluated on $t\bar{t}$ MC, of 89.9% on low energy muons and 91.8% on high energy muons. Additionally, Muons must have $\eta < 2.4$, and impact parameter significance $\frac{|d_0|}{\sigma_{d_0}} < 3$ and $|z_0 sin(\theta)| < 0.5$ mm. Finally, it must pass isolation point FCTight, which differentiates signal Muons from background processes like semileptonic decays [67]

9.5 Jets

A highly collimated shower of hadrons originating from the same particle is called a "Jet". The original particle is identified by the characteristics of the resulting jets from its decay products. Jets are most often are reconstructed from topological clusters in the calorimeter (topo-cluster) with information from the ID and MS to reduce background from pileup. A jet energy scale (JES) factor is calculated to convert calorimeter data to particle energy measurements.

9.5.1 Particle Flow

Jets used in run-2 analysis are reconstructed using particle flow algorithm. The particle flow algorithm enumerates all the reconstructed particles in a jet by matching their topo-clusters and ID tracks. The inputs to particle flow are a set of tracks and a list of topo-clusters. The algorithm then attempts to match each track to a single topo-cluster in the calorimeter. The expected energy in the calorimeter deposited by the particle that also create the associated track is computed based on the topo-

cluster position and associated track momentum. In the case that a particle deposited energy in more than one topo-cluster, the algorithm adds more topo-clusters to the track/cluster system to recover the full shower.

Once all the topo-clusters have been identified, the next step is energy subtraction. In the subtraction step, rings in η,ϕ space are created around the associated track extrapolated to the calorimeter, in each successive calorimeter layer. The rings are wide enough to contain at least one Calorimeter cell. Starting from the ring with the highest energy density, if the energy in all the cells contained in the ring is less then $\langle E_{dep} \rangle$, the expected energy deposited from a track based on its momentum, then the cells are removed and $\langle E_{dep} \rangle$ is reduced by the total energy in the cells. If the energy in the ring is greater than the expected energy leftover, the energy in the remaining cells is scaled down to reach the expected energy, and the algorithm halts. The set of tracks and topo-clusters represent all the reconstructed particles in the event [70].

9.5.2 Jet Selection

For this analysis, jets are reconstructed using the anti-k_t algorithm, with a radius parameter R=0.4. The anti-k_t algorithm takes the topo-clusters from particle flow algorithm, within $|\eta| < 2.5$, and recursively clusters particles together in a cone-like shape within a ΔR in η,ϕ space and a given range in momentum space. The two leading jets are called "tagging" jets. After the jets are built, a sequence of corrections are applied to calibrate the jets to particle-level energy scale.

At preselection, jets are required to have a $p_T > 20$ GeV and $|\eta| < 4.5$. In addition, jets are required to satisfy Jet Vertex Tagger (JVT) requirements. JVT likelihood is a variable used to distinguish Pile up jet. The JVT tight working point is used, where JVT¿0.5 from PFlow jets between $20 < p_T < 60$. GeV

In the final selection, Jets must have $p_T > 30$. Jets that pass preselection but not the final selection are called sub-threshold jets [68].

9.5.3 b-tagging

B-jets are jets induced by a b quark, that pass the jet preselection criteria. Jets are identified as b-jets using a multivariate algorithm called DL1r b-tagging. DL1r is a neural-network based jet flavor tagging algorithm that takes PFlow track parameters as inputs and identifies jet as either from a b-quark, c-quark or light quark with a probability 0-1. For this analysis, b-jets must satisfy 85% efficiency.

9.6 Missing Transverse Energy

From momentum conservation, the vector sum of the transverse momenta of all the particles in a single collision should be zero. Often that is not the case, and this leftover momentum is termed missing transverse energy (MET) or E_T^{miss}. The presence of missing energy indicates either particles that can't be directly observed, such as neutrinos or potential BSM particles, or particles that avoided detection due to poor reconstruction. Missing energy is important in VBF interactions because decay has two opposite flavor neutrinos present. Missing energy is used in the calculation of the Higgs energy, the Higgs transverse mass m_T and the invariants mass of the tau tau vertex, $m_{\tau\tau}$, used in cutting out $Z \to \tau\tau$ events.

Two types of missing energy are typically used, one simply E_T^{miss} and $E_T^{miss,track}$. Both use track-based soft term (TST) reconstruction. TST E_T^{miss} uses tracks in the inner detector not associated with reconstructed hard objects. Hard objects are fully reconstructed and calibrated particle and jets. Fully reconstructed muons combine a MS track with ID track. Electrons and taus are reconstructed from combined calorimeter signal and ID tracks, while photons and jets are reconstructed purely from calorimeter signal [72]. The total E_T^{miss} term is thus

$$E_T^{miss} = -\sum_{hard} p_{x,y} - \sum_{soft} p_{x,y} \tag{9.1}$$

Where the momentum in the x and y direction are calculated separately and then summed in quadrature.

Finally, another variation, object-level $E_T^{miss,sig}$ is used in the analysis.$E_T^{miss,sig}$ can discriminate between missing energy from weakly interacting particles such as neutrinos and missing energy from detector effects such as inefficiencies and limited resolutions.$E_T^{miss,sig}$ is calculated according to:

$$E_T^{miss,sig} = \frac{E_T^{miss}}{\sqrt{\sum H_T}} \tag{9.2}$$

where $\sum H_T$ is the estimated resolution on missing energy calculation. A large value for $E_T^{miss,sig}$ indicates higher likelihood the missing energy is the result of weakly interacting particles.

9.7 Overlap removal

Overlap removal is a procedure in which multiple analysis objects are reconstructed from the same detector signals. The latest recommendation on overlap removal (OR) are followed in the analysis. The official ASG tool is used for the overlap removal [73]. Only production-level preselected objects among electrons, muons and jets are considered. The individual overlap removal is done in the following steps:

- electron-electron: If there is an overlapping second layer cluster (3x5) or a shared track, the lower p_T electron is removed.

- electron-muon: If a combined muon shares an ID track with an electron, the electron is removed. If a calo-tagged muon shares an ID track with an electron, the muon is removed.

- electron-jet: The jet is removed if $\Delta R(jet, e) < 0.2$. For any surviving jets, the electron is removed if $\Delta R(jet, e) < min(0.4, 0.04 + 10GeV/p_T^e)$.

- muon-jet: If$\Delta R(jet, \mu) < 0.2$ then the jet is removed if the jet has less than three associated tracks with $p_T > 500$ MeV or both of the following conditions are met: the p_T ratio of the muon and jet is larger than $0.5(p_T^{\mu}/p_T^{jet} > 0.5)$ and the ratio of the muon p_T to the sum of p_T of tracks with $pT > 500$ MeV associated to the jet is larger than 0.7. For any surviving jets, if $\Delta R(jet, \mu) < min(0.4, 0.04 + 10GeV/p_T^{\mu})$, the muon is removed.

The specific working point chosen from those that are recommended is the so-called "b-tag aware" overlap removal. In this scheme, a jet is always saved from removal if it is b-tagged.

CHAPTER 10

EVENT SELECTIONS

Vector Boson Fusion signal events are characterized by two energetic jets, with large separation in rapidity that accompany the Higgs boson and its decay products. Since the Higgs was produced through the fusion of electro-weak bosons there is no QCD interaction between the jets and Higgs or any of the decay products. In this analysis the backgrounds are estimated with several different methods. The largest backgrounds are estimated through dedicated control regions, the smaller backgrounds are estimated through just Monte Carlo simulations. Estimation methods are listed in table 10.1

Background	size at preselction	size in SR	Estimation Method
WW	$\sim 8\%$	$\sim 20\%$	MC + BDT
Top	$\sim 73\%$	$\sim 61\%$	MC + BDT
ggF Higgs	$< 1\%$	$\sim 2\%$	MC + BDT
W+jets	$\sim 2.5\%$	$\sim 4\%$	Data-driven
$Z \to \tau\tau$	$\sim 16\%$	$\sim 7\%$	Data+MC ($Z \to \tau\tau$ CR)
$W\gamma$,$W\gamma^*$	$\sim 1\%$	$\sim 2.5\%$	MC only
ttH,VH	$<\sim 1\%$	$<\sim 1\%$	MC only

Table 10.1. Background sources, their relative contributions in a preselection region and in the SR (after all cuts), the background estimation methods, and the generators used for modeling.

Top quark production is suppressed in the SR by a b-jet veto requirement, and can be studied in a subregion of the SR that is depleted of signal events and is defined by a boosted decision tree discriminant. The subregion for studying top and diboson backgrounds is defined by a cut on the VBF bdt distribution at < 0.5, and is referred to in the rest of this paper as Signal Region 2 (SR2). The cut on the VBF bdt

provides an enriched region of signal-like top and diboson events. A pure validation region is also defined to cross-check the MC modeling of top-quark production.

The Z+jets background is suppressed by kinematic requirements and is constrained in a dedicated control region and ultimately normalized in the signal region . The ggF Higgs production mode is particularly difficult to separate from the VBF production mode, therefore several ggF-enriched regions have been studied to constrain it in data, and ultimately its normalization is determined in a single ggF CR. The studies concluded that one single ggF CR is sufficient to reliably constrain the ggF background. The W+jets events constitute a background when a jet is miss-identified ('fakes') as a lepton. This background is estimated with data-driven techniques, as described in the section 'fake' estimation, and is then fixed in the simultaneous fit of all processes in SR and CRs.

The analysis employs boosted decision trees (BDT) trained to discriminate between samples in the SR, and one more BDTs trained in the ggF CR. The Signal region employs a 2D BDT, where the first dimension is a discriminator to classify VBF events vs. backgrounds and the second dimension discriminates between Top or WW events vs. all other processes. The ggF control region uses a BDT to discriminate between ggF events vs. all other events. See section on BDTs for details.

This section details the selection requirements used to define signal, control and validation regions. Tables and figures in this section show event yields in data (for the unblinded regions) and in MC simulations at detector level. The MC predictions are normalized to their default normalizations. No additional normalization factors are applied.

10.1 Pre-selection

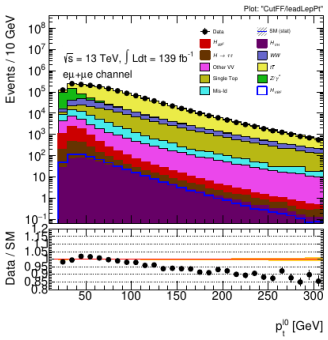
This analysis uses the same preselection as the 2016 VBF $H \to WW \to \ell\nu\ell\nu$ analysis. The cuts are listed in table 10.2, using the same observable definition in the previous section.

Pre-Sel Cut	Description
Channel Selection	$e\mu/\mu e$ channel,event weights applied
Trigger Selection	single lepton and dilepton triggers
Trigger Matching	Applying Trigger Weights
Data Driven muon/electron	W+jets split into electron and muon fake flavors
Cut Jet Cleaning	Pass loose jet cleaning
Cut Vγ/Z+Jets overlap	Eliminate Z+jets events that overlap with Vγ
Only two leptons	Require exactly 2 leptons (in case of fakes make sure 1 ID/1 Anti-ID)
Lead Lepton p_T	$p_T^{lead} > 22$GeV
Sublead Lepton p_T	$p_T^{sublead} > 15$GeV
Opposite sign leptons	Require zero dilepton charge
$m_{\ell\ell}$ cut	$m_{\ell\ell} > 10$ GeV
Fake Factor Implementation	W+jets weight applied

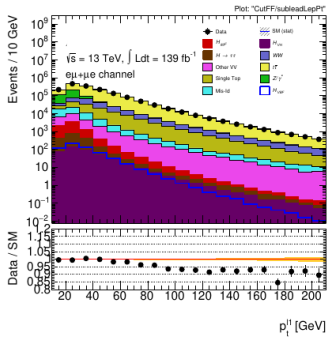
Table 10.2. Pre-Selection cuts shared with ggF analysis

Yields for all samples as well as data for pre-selection cuts are shown in Table 10.3.

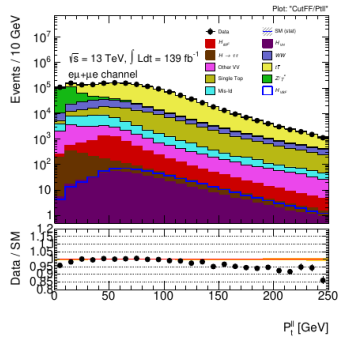
The plots in Figs. 10.1, 10.2 and 10.3 show kinematic distributions after all preselection cuts with fake factors applied for fake estimation. These plots only include the MC stat uncertainty and show good modeling with data within a 5-10% level over a variety of kinematic variables including those used as input to the Top+WW vs. VBF BDT. Discrepancies greater than 10% are only visible in tails of few distributions, with no significant impact on the analysis due to the small statistics in those regions. Theory and experimental systematic uncertainties, which are at the 10-20% level or greater, are not shown here.



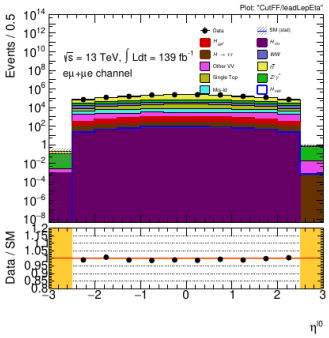
(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$



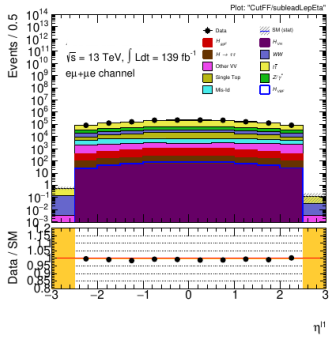
(b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$



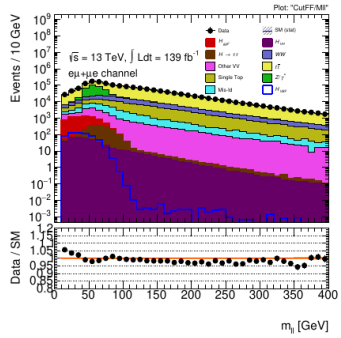
(c) p_{T}^{ll}



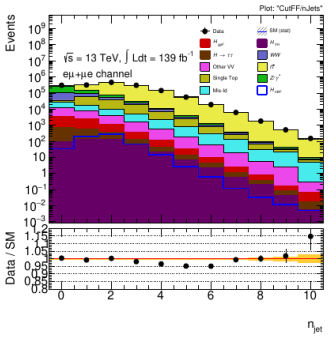
(d) lead $\eta(\ell)$



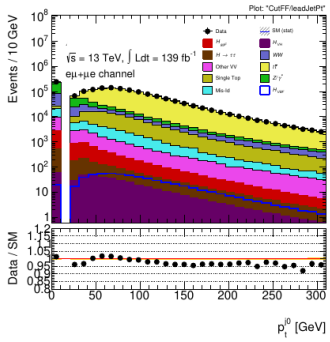
(e) sublead $\eta(\ell)$



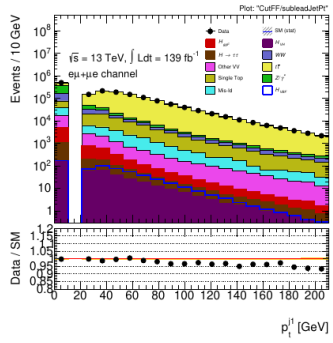
(f) M_{ll}



(g) n_{jets}

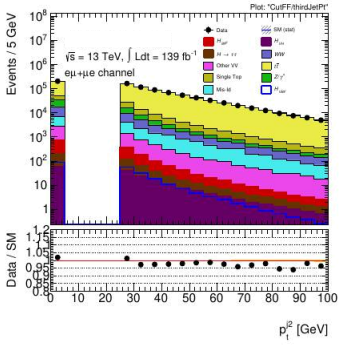


(h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$

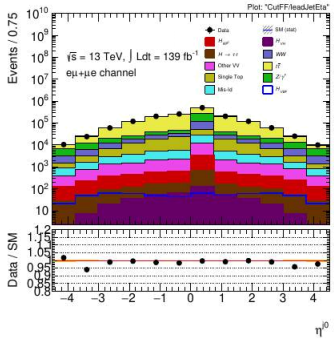


(i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

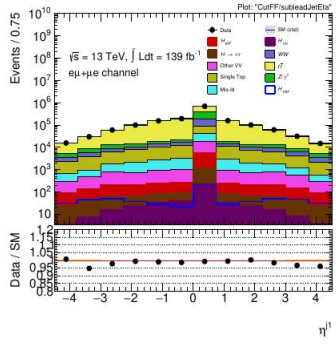
Figure 10.1. Distribution of observables after preselection with statistical error



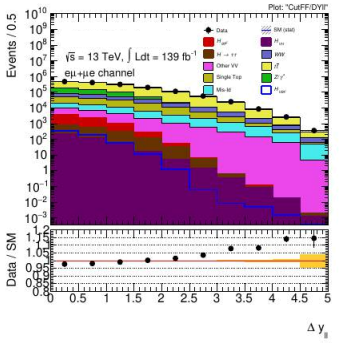
(a) third jet p_{T}



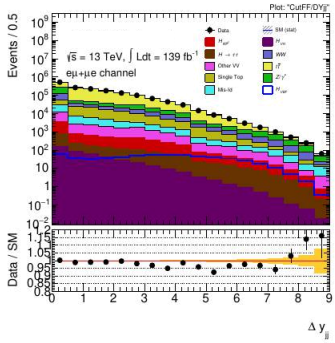
(b) lead η(jet)



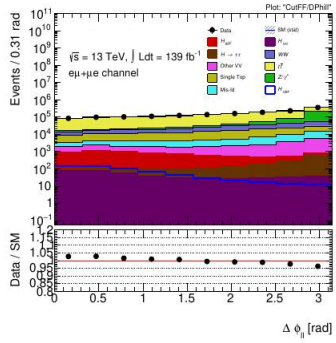
(c) sublead η(jet)



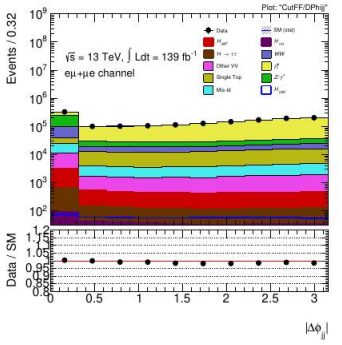
(d) Δy_{ll}



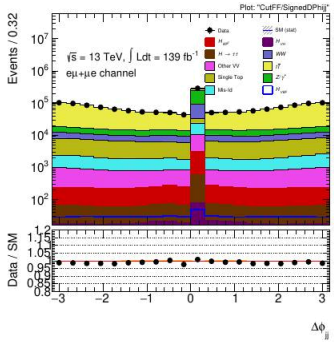
(e) Δy_{jj}



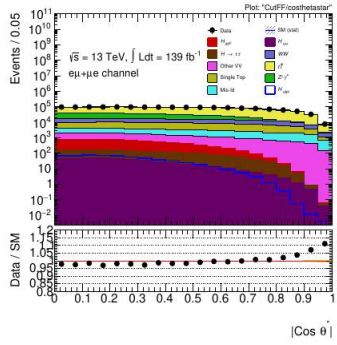
(f) $\Delta\phi_{ll}$



(g) $\Delta\phi_{jj}$

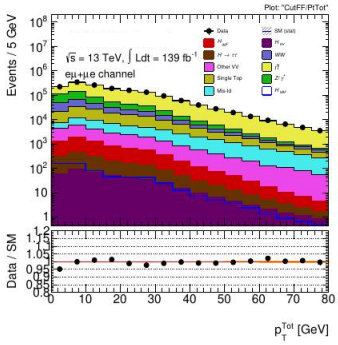


(h) Signed-$\Delta\phi_{jj}$

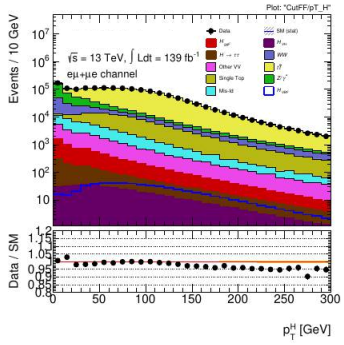


(i) $\cos\theta^*$

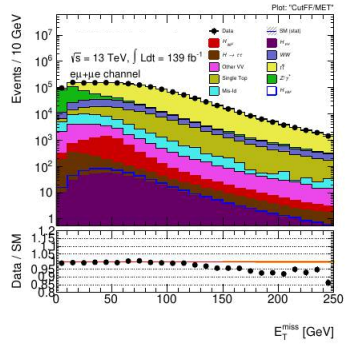
Figure 10.2. Distribution of observables after preselection with statistical error



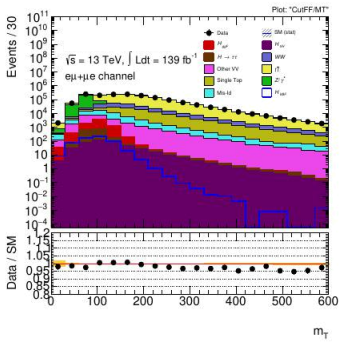
(a) $p_{\mathrm{T}}^{\mathrm{tot}}$



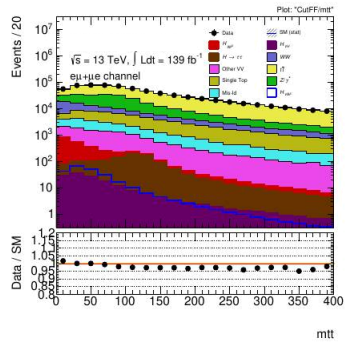
(b) p_{T}^{H}



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



(d) m_{T}



(e) $m_{\tau\tau}$

Figure 10.3. Distribution of observables after preselection with statistical error

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
Channel Selection	619.73±0.76	5912.42 ± 8.09	1869.22 ± 2.69	130744.71 ± 121.33	22149.14 ± 132.73	1215040.34± 243.62	273581.31 ± 471.56	5112304.94± 4305.85	6761602.08± 4342.17	0.24 ± 0.00	4374979	0.65 ± 0.00
Trigger Selection	619.73±0.76	5912.42 ± 8.09	1869.22 ± 2.69	130744.71 ± 121.33	22149.14 ± 132.73	1215040.34± 243.62	273581.31 ± 471.56	5112304.94± 4305.85	6761602.08± 4342.17	0.24 ± 0.00	4374979	0.65 ± 0.00
Trigger Matching	602.71±0.74	5697.95 ± 7.82	1809.71 ± 2.62	128358.45 ± 119.31	21087.93 ± 126.96	1192099.51± 239.59	260696.26 ± 447.19	5303655.16± 4230.14	6913404.96± 4264.02	0.23 ± 0.00	4352644	0.63 ± 0.00
W+jets flavour split muon	602.71±0.74	5697.95 ± 7.82	1809.71 ± 2.62	128358.45 ± 119.31	21087.93 ± 126.96	1192099.51± 239.59	260696.26 ± 447.19	4899580.58± 3754.15	6509330.38± 3792.29	0.24 ± 0.00	4352644	0.67 ± 0.00
W+jets flavour split electron	602.71±0.74	5697.95 ± 7.82	1809.71 ± 2.62	128358.45 ± 119.31	21087.93 ± 126.96	1192099.51± 239.59	260696.26 ± 447.19	3919098.18± 3121.04	5528847.99± 3166.81	0.26 ± 0.00	4352644	0.79 ± 0.00
Jet Cleaning	602.71±0.74	5697.95 ± 7.82	1809.71 ± 2.62	128358.45 ± 119.31	21087.93 ± 126.96	1192099.51± 239.59	260696.26 ± 447.19	3919098.18± 3121.04	5528847.99± 3166.81	0.26 ± 0.00	4352644	0.79 ± 0.00
Overlap: Vgamma/Vjets	602.71±0.74	5697.95 ± 7.82	1809.71 ± 2.62	128358.45 ± 119.31	21087.93 ± 126.96	1192099.51± 239.59	260696.26 ± 447.19	3919098.18± 3121.04	5528847.99± 3166.81	0.26 ± 0.00	4352644	0.79 ± 0.00
Only two Leptons	600.90±0.74	5693.00 ± 7.81	1792.12 ± 2.60	128163.65 ± 119.26	19507.46 ± 125.74	1177783.16± 238.25	260303.18 ± 444.93	3914948.92± 3100.14	5508191.48± 3145.74	0.26 ± 0.00	4323749	0.78 ± 0.00
$p_t^{lead} > 22$ GeV	600.90±0.74	5693.00 ± 7.81	1792.12 ± 2.60	128163.65 ± 119.26	19507.46 ± 125.74	1177783.16± 238.25	260303.18 ± 444.93	3914948.92± 3100.14	5508191.48± 3145.74	0.26 ± 0.00	4323749	0.78 ± 0.00
$p_t^{sublead} > 15$ GeV	598.61±0.74	5662.30 ± 7.79	1784.55 ± 2.60	127974.96 ± 119.18	19409.61 ± 124.67	1175875.71± 238.06	258912.26 ± 444.25	3913337.04± 3099.53	5502956.42± 3144.99	0.26 ± 0.00	4322240	0.79 ± 0.00
Remove High Eventweight Zjets Samples	598.61±0.74	5662.30 ± 7.79	1784.55 ± 2.60	127922.47 ± 115.82	19422.15 ± 118.32	1175875.71± 238.06	253469.93 ± 312.18	3989048.25± 2888.74	5573185.35± 2920.01	0.25 ± 0.00	4322240	0.78 ± 0.00
OS Leptons	598.60±0.74	5662.30 ± 7.79	1784.52 ± 2.60	127922.25 ± 115.82	19419.93 ± 118.32	1175868.71± 238.06	253469.16 ± 312.18	3983509.56± 2887.46	5567636.43± 2918.74	0.25 ± 0.00	4319096	0.78 ± 0.00
$M_{\ell\ell} > 10$ GeV	598.45±0.74	5660.63 ± 7.79	1784.44 ± 2.60	127918.50 ± 115.81	19414.43 ± 118.30	1175831.89± 238.06	253459.69 ± 312.16	3983236.06± 2887.40	5567305.64± 2918.68	0.25 ± 0.00	4318965	0.78 ± 0.00
SF: Z Veto	598.45±0.74	5660.63 ± 7.79	1784.44 ± 2.60	127918.50 ± 115.81	19414.43 ± 118.30	1175831.89± 238.06	253459.69 ± 312.16	3983236.06± 2887.40	5567305.64± 2918.68	0.25 ± 0.00	4318965	0.78 ± 0.00
Leptons ID, singleFakes 1 anti-ID,1 ID, doubleFakes 2*anti-ID	591.57±0.73	5593.39 ± 7.74	1763.44 ± 2.58	126352.68 ± 115.12	19183.18 ± 117.07	1162127.44± 236.67	250269.90 ± 310.22	1163533.77± 1482.59	2728823.80± 1541.87	0.36 ± 0.00	1585575	0.58 ± 0.00
Apply QCD FF weight	591.57±0.73	5593.39 ± 7.74	1763.44 ± 2.58	126352.68 ± 115.12	19183.18 ± 117.07	1162127.44± 236.67	250269.90 ± 310.22	519203.80 ± 1201.95	2084493.83± 1274.35	0.41 ± 0.00	1585575	0.76 ± 0.00
Apply fake factor	591.57±0.73	5593.39 ± 7.74	1763.44 ± 2.58	126352.68 ± 115.12	19183.18 ± 117.07	1162127.44± 236.67	250269.90 ± 310.22	35395.54 ± 222.65	1600685.58± 478.38	0.47 ± 0.00	1585575	0.99 ± 0.00

Table 10.3. Cutflow in the pre-selection region.

10.2 Signal Region Selection

After pre-selection cuts, additional cuts are applied to select a region of phase space with high signal to background ratio. The cuts for the signal region include at least 2 jets with p_T greater than 30 GeV, a b-veto using the DL1r b-tagging algorithm, central jet veto (CJV) and outside lepton veto detailed in previous selection. Additionally after studies of signal yields, a cut on $M_{jj} > 450$ GeV, $\Delta\phi_{\ell\ell} < 1.4$ and $\Delta y_{jj} > 2.1$ are also applied. These requirements further purify the signal region against a range of backgrounds, most notably top-quark, and minimize the theoretical extrapolation. The cut values reported above have been chosen as they improve the signal significance while they retain high signal statistics. Signal Region cuts are listed in table 10.4. The signal and background efficiency, the total number of events removed removed by each cut, is shown in table 10.5. Figures 10.4,10.5, and 10.6 show the distribution of obsevables after all signal region cuts are applied.

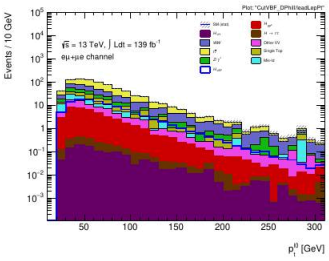
s

Signal region cut	Description
2-jet (30,30)	Require at least 2 jets with $p_{\mathrm{T}} \geq 30$ GeV
b-veto	Use DL1 efficiency b-tag reject events with b-jets/apply b-tag weight
Blinding (VBF)	Blinding condition on data (cut on VBF vs. Top+WW BDT)
CJV (20 GeV)	Cut events with a third central rapidity jet $p_{\mathrm{T}} > 20$ GeV
OLV bool	Leading lepton η required to be between two η 2 leading jets
$Z \rightarrow \tau\tau$ veto	$m_{\tau\tau} < m_Z - 25$ GeV
Δy_{jj} cut	$\Delta y_{jj} > 2.1$
m_{jj} cut	$m_{jj} > 450$ GeV
$\Delta\phi_{ll}$ cut	$\Delta\phi_{ll} < 1.4$

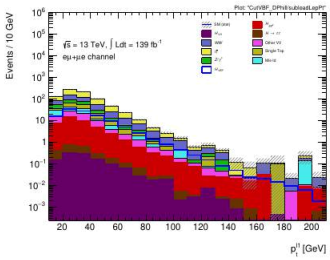
Table 10.4. list of signal region cuts

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
2-jet (30,30)	363.24 ± 0.57	1054.94 ± 3.06	542.18 ± 1.36	25110.51 ± 30.37	4813.96 ± 56.05	920667.89 ± 203.58	25856.34 ± 81.62	11720.52 ± 160.37	989766.33 ± 279.11	0.37 ± 0.00	975285	0.99 ± 0.00
b-veto	320.75 ± 0.54	908.56 ± 2.84	434.60 ± 1.21	21730.53 ± 28.76	3788.26 ± 51.13	64042.42 ± 57.42	21691.11 ± 75.97	3978.86 ± 65.22	116574.34 ± 129.51	0.94 ± 0.00	109428	0.94 ± 0.00
$Z \rightarrow \tau\tau$ veto	275.52 ± 0.50	793.52 ± 2.66	174.76 ± 0.73	12588.68 ± 22.72	2033.41 ± 40.60	38825.59 ± 44.57	7665.25 ± 50.49	2466.94 ± 49.87	64548.15 ± 95.89	1.08 ± 0.00	61311	0.95 ± 0.00
blinding (2-jet)	275.52 ± 0.50	793.52 ± 2.66	174.76 ± 0.73	12588.68 ± 22.72	2033.41 ± 40.60	38825.59 ± 44.57	7665.25 ± 50.49	2466.94 ± 49.87	64548.15 ± 95.89	1.08 ± 0.00	0	0.00
CJV (20GeV)	218.54 ± 0.44	581.16 ± 2.29	130.22 ± 0.63	9043.26 ± 19.65	1415.24 ± 32.69	26539.80 ± 37.26	5466.81 ± 42.67	1721.38 ± 41.14	44897.86 ± 79.76	1.03 ± 0.00	0	0.00
OLV bool	170.42 ± 0.39	160.42 ± 1.21	29.76 ± 0.30	1859.19 ± 9.28	347.38 ± 15.45	6035.59 ± 17.79	1279.55 ± 21.55	314.61 ± 18.21	10026.50 ± 37.93	1.70 ± 0.01	0	0.00
$M_{jj} > 200$ GeV	163.79 ± 0.38	109.83 ± 0.99	15.59 ± 0.22	1315.43 ± 8.02	216.88 ± 9.37	3567.26 ± 13.73	825.48 ± 16.61	211.52 ± 14.62	6262.00 ± 28.83	2.06 ± 0.01	0	0.00
Di Jet Separation, DYjj > 2.1	161.36 ± 0.38	101.05 ± 0.96	13.78 ± 0.20	1175.51 ± 7.89	179.87 ± 8.62	2991.35 ± 12.70	760.07 ± 16.02	189.40 ± 13.69	5411.03 ± 27.26	2.18 ± 0.01	0	0.00
$M_{jj} > 450$ GeV	127.30 ± 0.34	47.39 ± 0.65	7.18 ± 0.14	592.61 ± 5.65	86.50 ± 5.90	1034.49 ± 7.60	357.27 ± 10.85	92.63 ± 8.99	2218.07 ± 17.99	2.68 ± 0.01	0	0.00
DPhill < 1.4	114.35 ± 0.32	39.00 ± 0.59	2.50 ± 0.08	252.07 ± 3.31	42.39 ± 3.48	423.42 ± 4.84	79.42 ± 5.12	46.98 ± 6.15	885.80 ± 10.53	3.76 ± 0.02	0	0.00

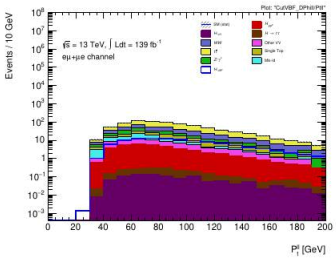
Table 10.5. Cutflow of signal region cuts



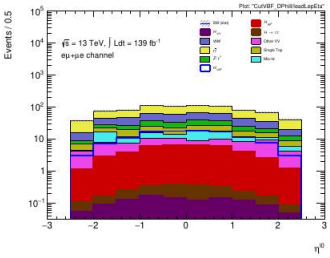
(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$



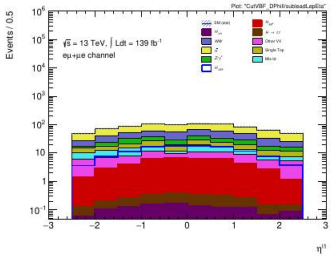
(b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$



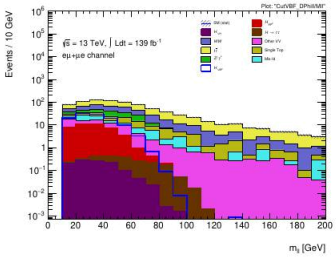
(c) p_{T}^{ll}



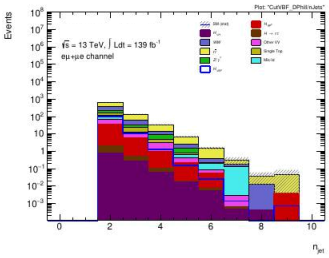
(d) lead $\eta(\ell)$



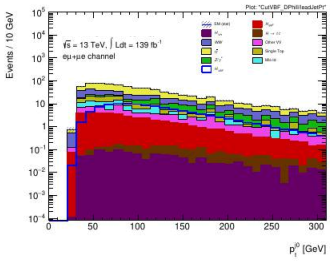
(e) sublead $\eta(\ell)$



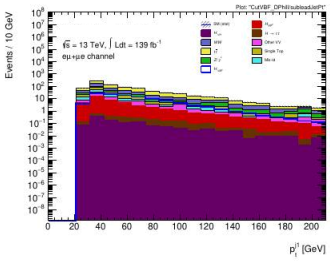
(f) M_{ll}



(g) n_{jets}

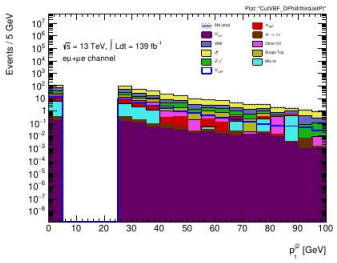


(h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$

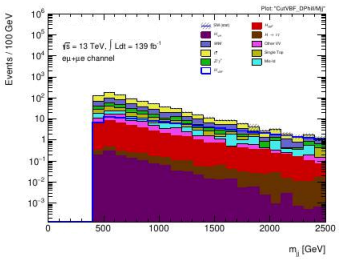


(i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

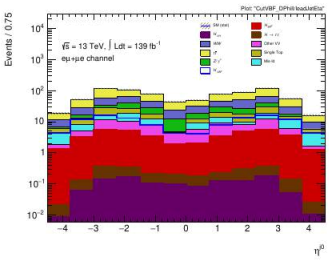
Figure 10.4. Distribution of observables after signal region cuts with statistical error



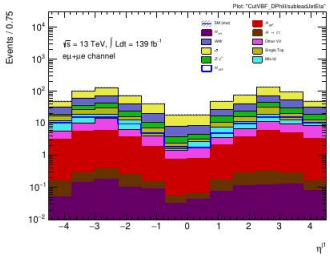
(a) third jet p_{T}



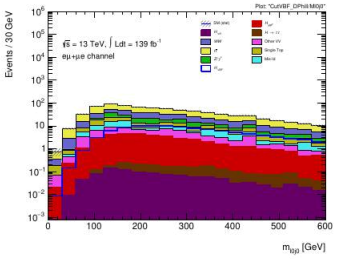
(b) M_{jj}



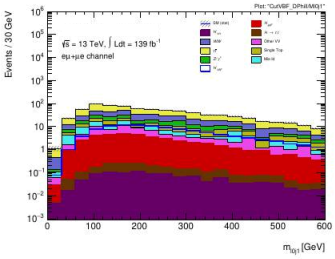
(c) lead η(jet)



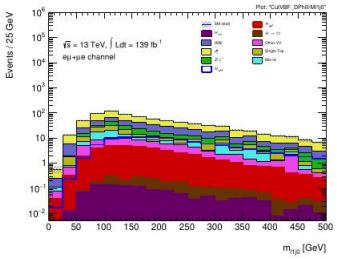
(d) sublead η(jet)



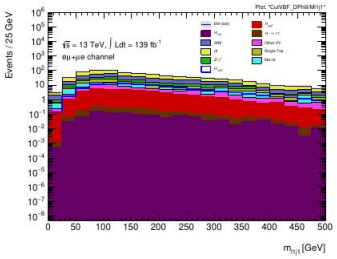
(e) $M_{l_0 j_0}$



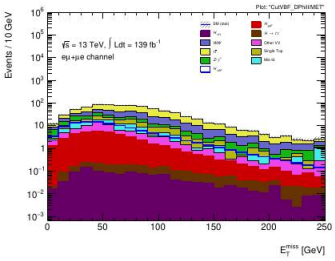
(f) $M_{l_0 j_1}$



(g) $M_{l_1 j_0}$

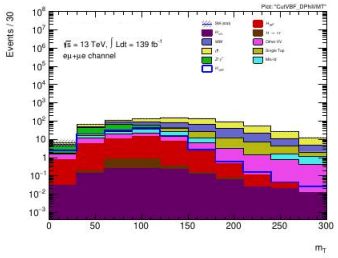


(h) $M_{l_1 j_1}$

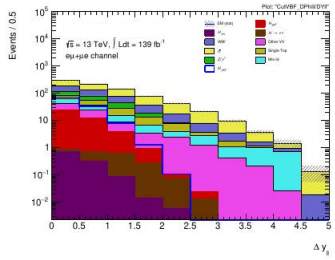


(i) $E_{\mathrm{T}}^{\mathrm{miss}}$

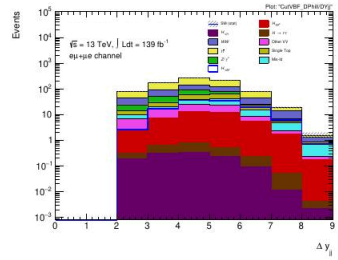
Figure 10.5. Distribution of observables after signal region cuts with statistical error



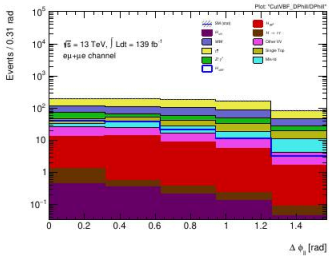
(a) m_{T}



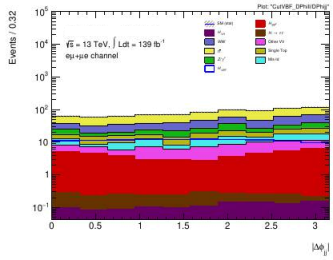
(b) Δy_{ll}



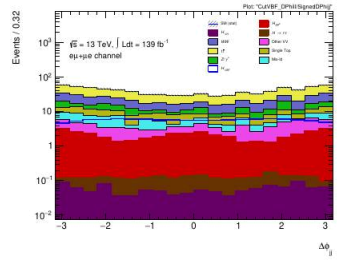
(c) Δy_{jj}



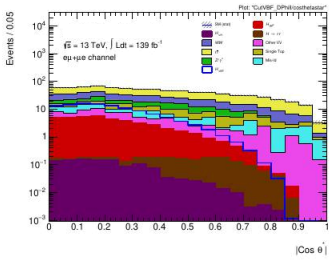
(d) $\Delta\phi_{ll}$



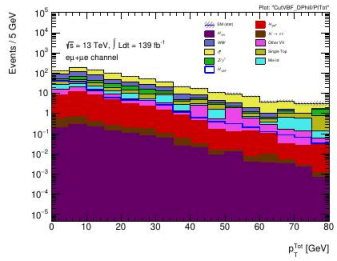
(e) $\Delta\phi_{jj}$



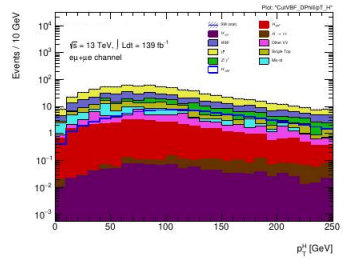
(f) Signed-$\Delta\phi_{jj}$



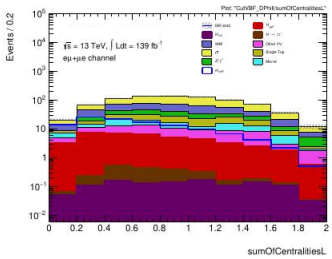
(g) $\cos\theta^*$



(h) $p_{\mathrm{T}}^{\mathrm{tot}}$



(i) p_{T}^{H}



(j) sum of Centralities

Figure 10.6. Distribution of observables after signal region cuts with statistical error

10.3 Z+Jets Control Region

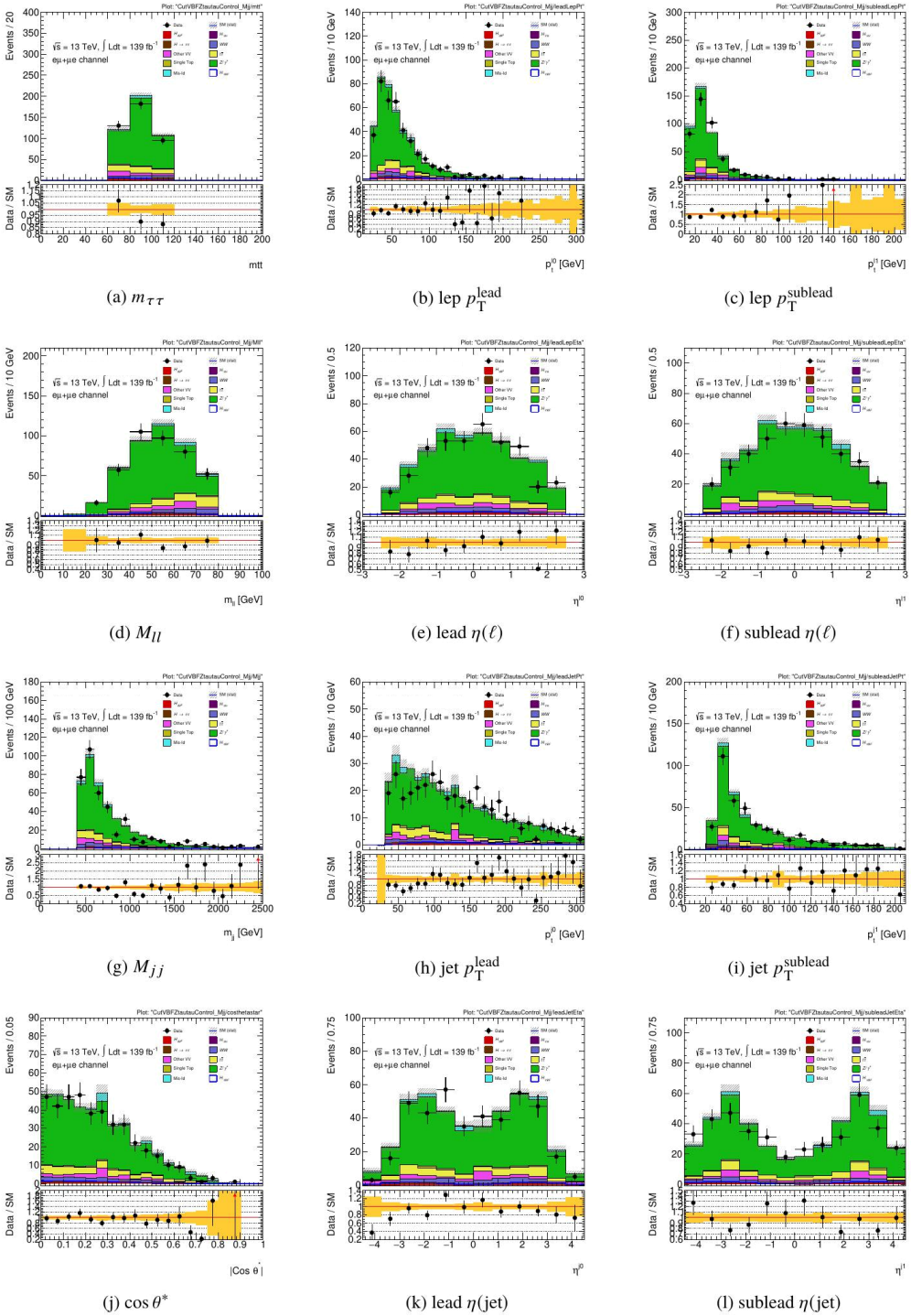
The Drell-Yan background, also referred to as Z+jets, consists of events in which two incoming partons produce two jets in the final state and a Z boson which then decays to two leptons. This background is dominated by $Z \to \tau\tau$ events in which τ leptons decay to electron/muons and neutrinos, since this analysis focuses on the electron-muon channel for the signal selection. Thus, in these background events there are typically two jets and two leptons in the final state as well as missing energy from neutrinos, all of which are also featured in the signal VBF Higgs events.

To discrimnate between $Z \to \tau\tau$ and signal events, the kinematics of the reconstructed τ-τ system and the on shell mass of the Z boson are used. A cut on $m_{\tau\tau}$, where only events within 25 GeV of the Z boson mass (91 GeV) are accepted, is applied to the Zjets control region, which increases the purity from 19% to 52% of the total number of events in the phase space. Table 10.6 show the cuts on the Zjets control region. The cut on Mjj, OLV and CJV are identical to the signal region, and an additional cut on $m_{ll} < 80$ GeV is applied. The cuts on Δy_{jj} and $\Delta\phi_{ll}$ from the signal region are not applied to increase statistics.

Figures 10.7 and 10.8 show good agreement between Monte Carlo simulations and data for various distributions. The discrepancy between data and MC is consistent with statistical fluctuations.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
$Z \to \tau\tau$ CR: $\lvert m_{\tau\tau} - m_Z \rvert < 25$, bVeto	26.92 ± 0.16	66.78 ± 0.75	101.09 ± 0.60	1856.99 ± 8.23	680.49 ± 21.86	5640.17 ± 17.13	8771.15 ± 39.93	440.18 ± 25.39	17556.84 ± 55.49	0.20 ± 0.00	16365	0.93 ± 0.01
$Z \to \tau\tau$ CR: $M_{ll} < 80$ GeV	26.26 ± 0.15	64.53 ± 0.74	83.34 ± 0.54	603.57 ± 4.54	465.82 ± 18.43	1691.93 ± 9.20	8539.07 ± 38.54	229.93 ± 20.48	11678.19 ± 48.48	0.24 ± 0.00	10789	0.92 ± 0.01
$Z \to \tau\tau$ CR: CJV< 20 GeV	20.68 ± 0.14	46.25 ± 0.63	63.35 ± 0.47	430.33 ± 3.92	348.76 ± 16.31	1140.88 ± 7.62	6379.53 ± 34.57	182.88 ± 17.63	8591.96 ± 42.97	0.22 ± 0.00	7922	0.92 ± 0.01
$Z \to \tau\tau$ CR: OLV	15.90 ± 0.12	11.55 ± 0.32	20.81 ± 0.24	93.85 ± 1.86	74.21 ± 10.13	289.58 ± 3.87	1386.46 ± 16.63	20.80 ± 7.77	1897.25 ± 21.40	0.36 ± 0.00	1830	0.96 ± 0.02
Mjj > 450 GeV	12.71 ± 0.11	3.81 ± 0.18	10.09 ± 0.13	27.11 ± 1.09	22.19 ± 4.27	44.83 ± 1.53	318.95 ± 7.13	11.66 ± 4.14	438.62 ± 9.47	0.60 ± 0.01	407	0.90 ± 0.05

Table 10.6. Cutflow of Drell Yann control region cuts



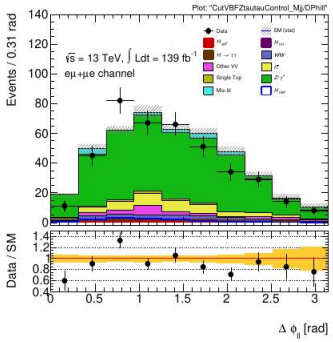
(a) $m_{\tau\tau}$ (b) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (c) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$

(d) M_{ll} (e) lead $\eta(\ell)$ (f) sublead $\eta(\ell)$

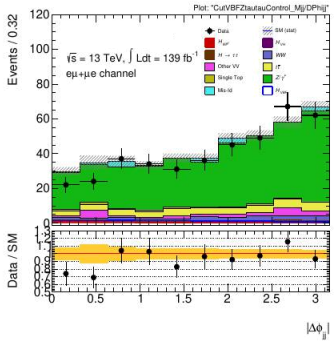
(g) M_{jj} (h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

(j) $\cos\theta^*$ (k) lead η(jet) (l) sublead η(jet)

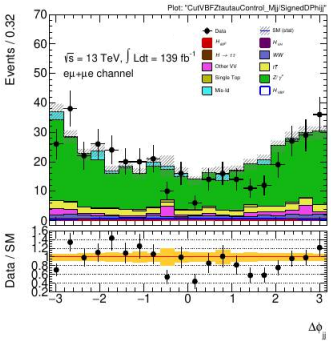
Figure 10.7. Distribution of observables after DY region cuts with statistical error



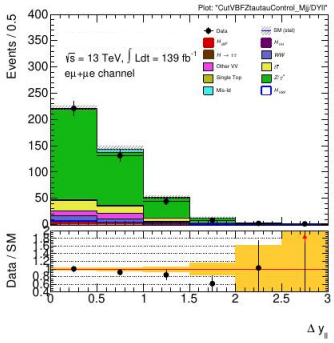
(a) $\Delta\phi_{ll}$



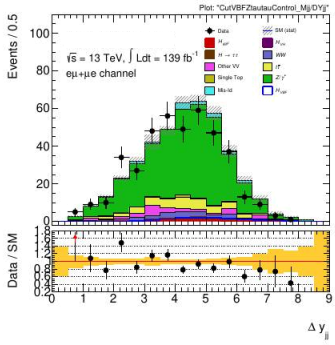
(b) $\Delta\phi_{jj}$



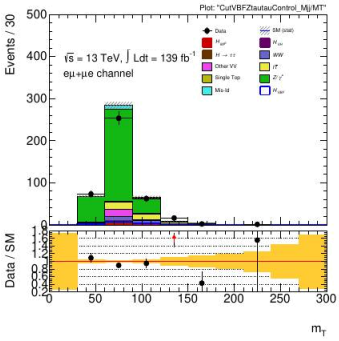
(c) Signed-$\Delta\phi_{jj}$



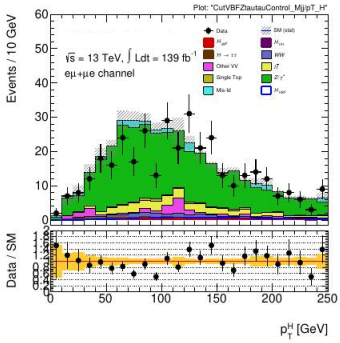
(d) Δy_{ll}



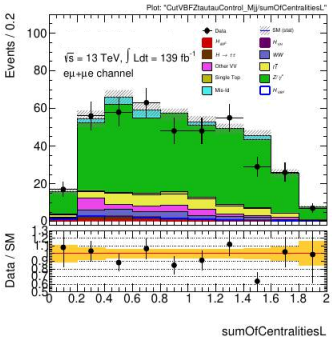
(e) Δy_{jj}



(f) m_{T}



(g) p_{T}^{H}



(h) sum of Centralities

Figure 10.8. Distribution of observables after DY region cuts with statistical error

10.4 Top Validation Region

The top background consists of two major components, Wt and $t\bar{t}$ events. Wt process is where final-state leptons originate from the W boson decays and the jets from the b-quark from the top-quark decay. The top background is dominated by $t\bar{t}$ and is the largest background in the signal region, composing about 60% of the total background. A dedicated BDT is trained to distinguish top backgrounds and signal Higgs events.

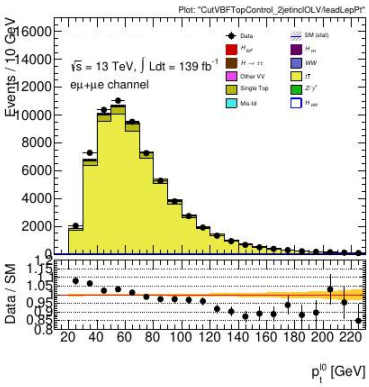
The top validation region (top VR) is used to check the MC modeling agreement with data. Top VR is defined with selection requirements similar to the signal regions, except instead of b-jet veto, a b-jet is required. The one b-jet requirements brings the top VR close to the SR but with significant top purity, 97%. The cuts to the top VR and resulting yields are shown in table 10.7. Figures 10.9 and 10.10 show good agreement between top MC and data, within the expected statistical uncertainty.

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
$n_{b-jets} >= 1$	42.48 ± 0.20	145.78 ± 1.14	107.81 ± 0.62	3379.38 ± 9.83	1050.34 ± 23.09	861090.32 ± 196.43	4174.49 ± 29.91	7523.65 ± 146.54	877471.77 ± 248.17	0.05 ± 0.00	865857	0.99 ± 0.00
$Z \rightarrow \tau\tau$ veto	36.44 ± 0.18	126.06 ± 1.06	55.38 ± 0.41	1873.24 ± 7.64	544.18 ± 17.21	527400.35 ± 152.89	1487.48 ± 19.41	4791.28 ± 117.84	536277.96 ± 194.92	0.05 ± 0.00	534463	1.00 ± 0.00
CJV < 20 GeV	26.88 ± 0.16	82.48 ± 0.86	38.61 ± 0.34	1189.53 ± 6.20	354.87 ± 14.85	360414.30 ± 126.88	973.47 ± 16.35	2897.79 ± 96.60	365951.05 ± 161.11	0.04 ± 0.00	364605	1.00 ± 0.00
OLV	18.64 ± 0.13	19.99 ± 0.43	6.82 ± 0.14	202.67 ± 2.72	62.17 ± 7.23	65108.81 ± 54.35	200.93 ± 8.86	400.08 ± 41.13	66001.47 ± 69.17	0.07 ± 0.00	66168	1.00 ± 0.00

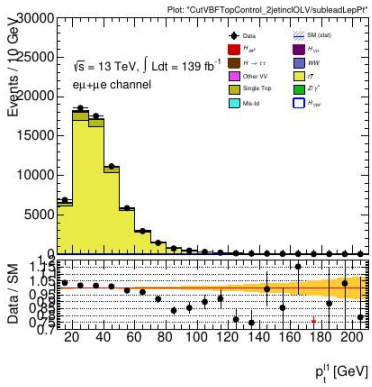
Table 10.7. Cutflow of Top validation region cuts

The WW background consists of QCD and electroweak WW+jets processes. The QCD-induced events are highly dominating in regions of phase space with low energy scale, such as low jet p_T and low m_{jj}, whereas in regions that are particularly sensitive to vector boson fusion and scattering processes, the electroweak WW production gives a contribution close to the one by QCD WW.

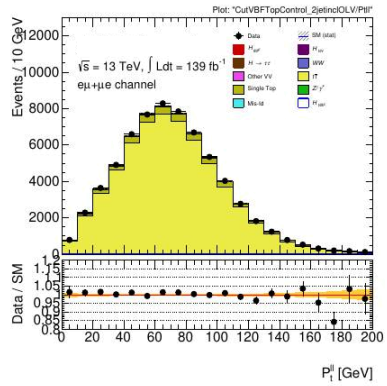
This background is estimated along with the top background in the SR via the top+WW BDT, due to the similarities between the two sets of processes in terms of final states and kinematics as well as due to the difficulty in defining a pure WW control region without top contamination. A WW validation region is defined to demonstrate the WW MC modeling agreement with data. The WW validation region



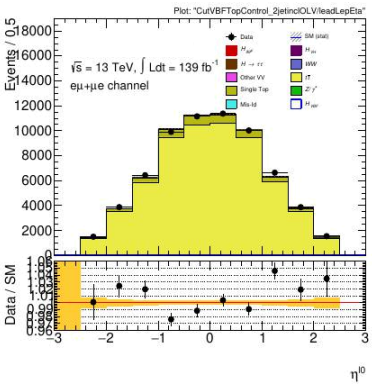
(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$



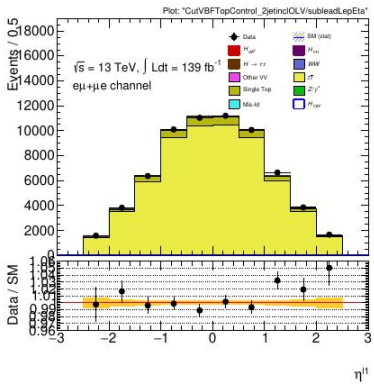
(b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$



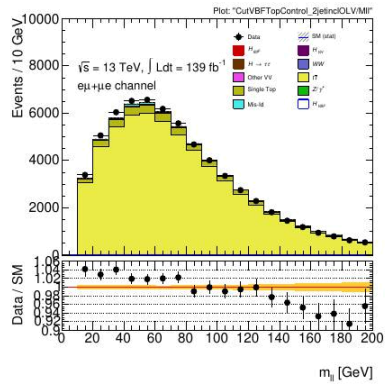
(c) p_{T}^{ll}



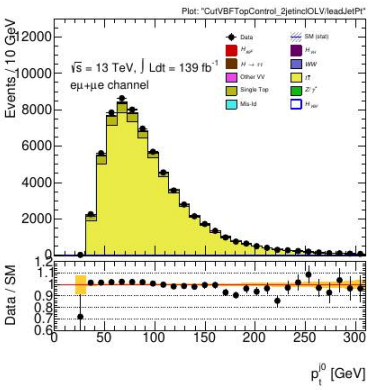
(d) lead $\eta(\ell)$



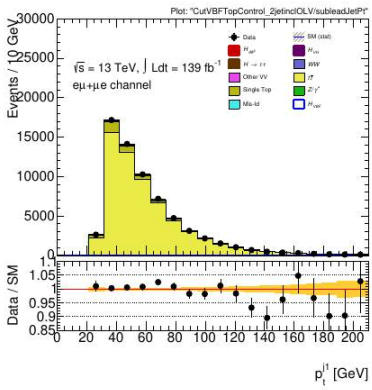
(e) sublead $\eta(\ell)$



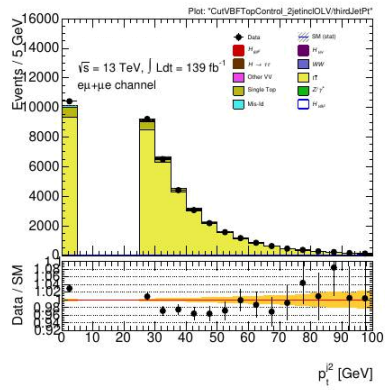
(f) $m_{\ell\ell}$



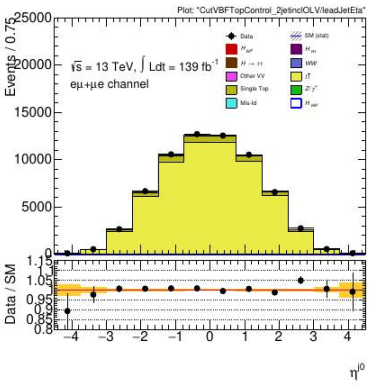
(g) jet $p_{\mathrm{T}}^{\mathrm{lead}}$



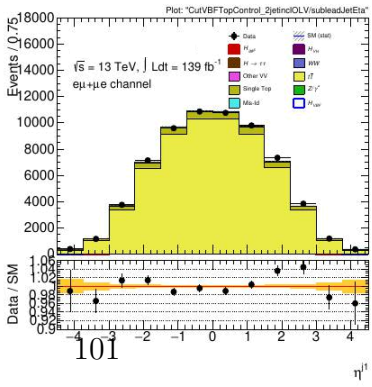
(h) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$



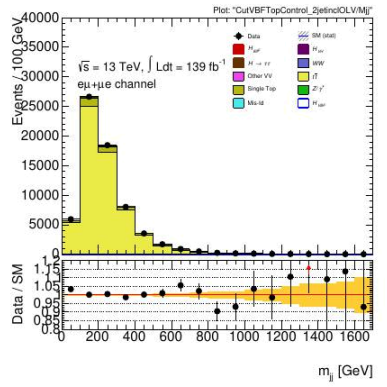
(i) third jet p_{T}



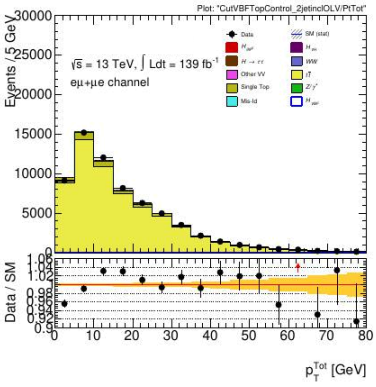
(j) lead η(jet)



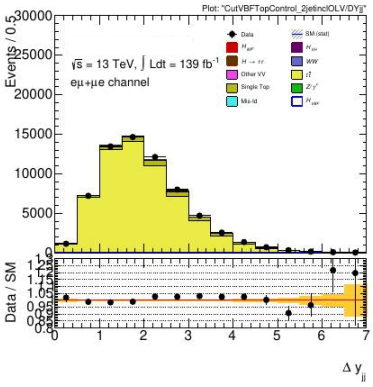
(k) sublead η(jet)



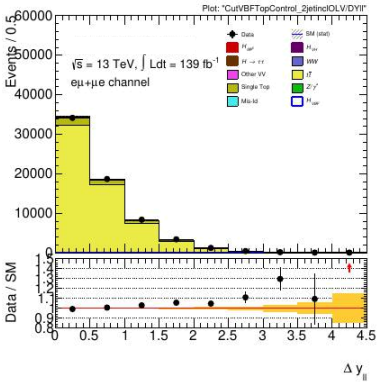
(l) m_{jj}



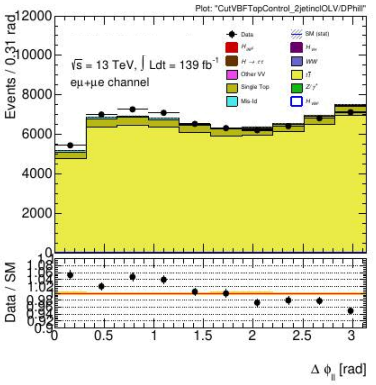
(a) $p_{\mathrm{T}}^{\mathrm{tot}}$



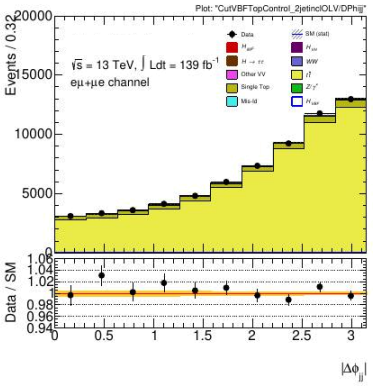
(b) Δy_{jj}



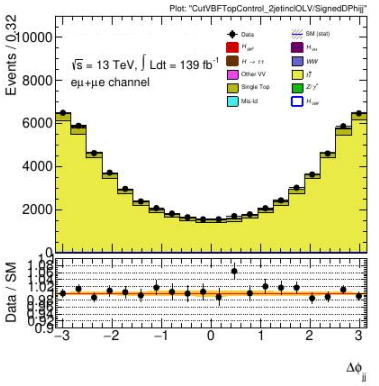
(c) $\Delta y_{\ell\ell}$



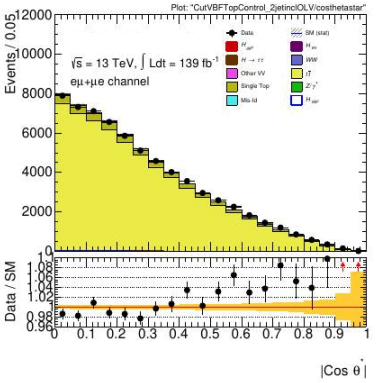
(d) $\Delta\phi_{\ell\ell}$



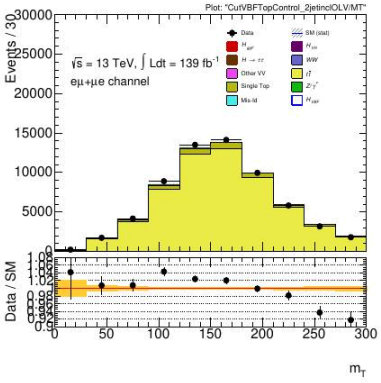
(e) $\Delta\phi_{jj}$



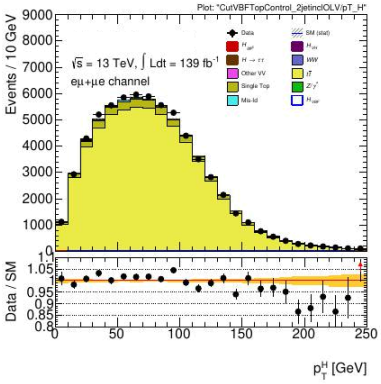
(f) Signed-$\Delta\phi_{jj}$



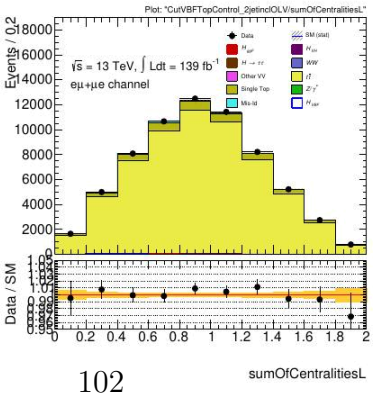
(g) $\cos\theta^*$



(h) m_{T}



(i) p_{T}^{H}



(j) sum of Centralities

definition consists of at least 2 jets, a b-jet veto, central jet veto, jet p_T threshold of 20 GeV and a cut on $m_T > 130$ GeV, not present in the signal region. The region has a purity of 37% and cutflow efficiency in table 10.8. The agreement between data and MC simulations is shown in Figure 10.11 and 10.12 for several distributions. Discrepancies are typically less than 5-10%, and those that are more than 10% are only in tails of the distributions compatible with statistical fluctuations.

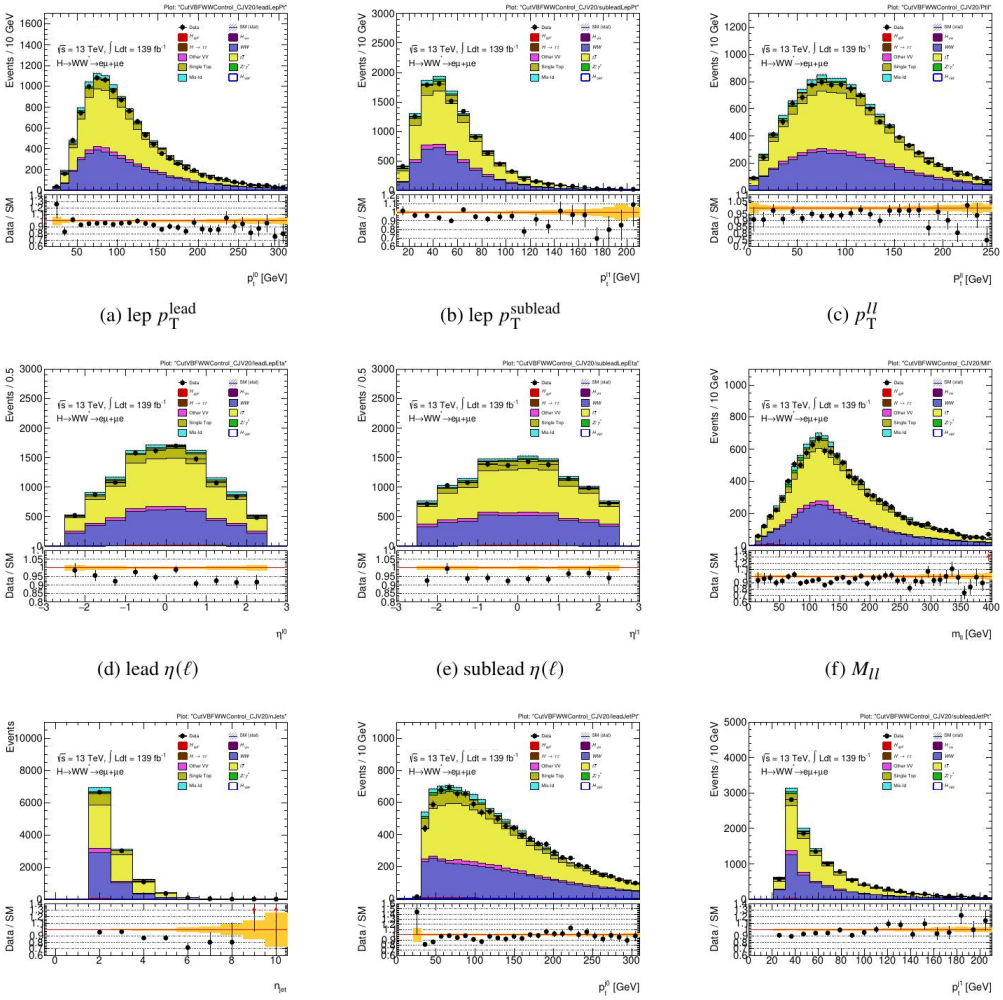
$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
M_T>130 GeV	32.16 ± 0.17	130.04 ± 1.09	59.29 ± 0.44	16680.73 ± 24.87	1763.85 ± 22.78	50511.72 ± 51.19	806.51 ± 16.08	1924.64 ± 46.89	71876.78 ± 78.84	0.12 ± 0.00	68076	0.95 ± 0.00
M_{T2}>160 GeV	18.75 ± 0.13	50.01 ± 0.66	17.07 ± 0.24	6891.90 ± 14.33	668.51 ± 13.73	11746.69 ± 25.15	288.97 ± 9.16	538.03 ± 24.17	20201.18 ± 41.17	0.13 ± 0.00	18605	0.92 ± 0.01
CJV (20GeV)	13.52 ± 0.11	32.35 ± 0.53	11.15 ± 0.20	4638.36 ± 11.95	437.30 ± 11.54	6519.11 ± 19.24	174.97 ± 7.62	345.45 ± 18.69	12158.70 ± 32.47	0.12 ± 0.00	11204	0.92 ± 0.01

Table 10.8. Cutflow of WW validation region cuts

The is an overall offset is visible in all distributions and it is estimated to be 7.5% by comparing the integral of the data and MC distributions in this VR. The impact of an offset of 7.5% on the signal cross section measurements was studied and it was found at most 0.5% effect on differential cross section measurements. The impact on the signal region is relatively small because the contamination of WW processes in the SR is about 45% of the top-quark background, and WW and top-quark backgrounds are normalized together in the fit to data. In the final fit, data-driven approaches are used for estimating these background contributions.

10.5 ggF Control and Validation region

A single ggF control region is used to estimate the total ggF contribution in the signal region and two validation regions are used to validate the MC-data agreement. Originally an ABCD like method was used to estimate the ggF contributions using all three regions, but further studies show that a single ggF region has sufficient statistical power to estimate the ggF contribution.

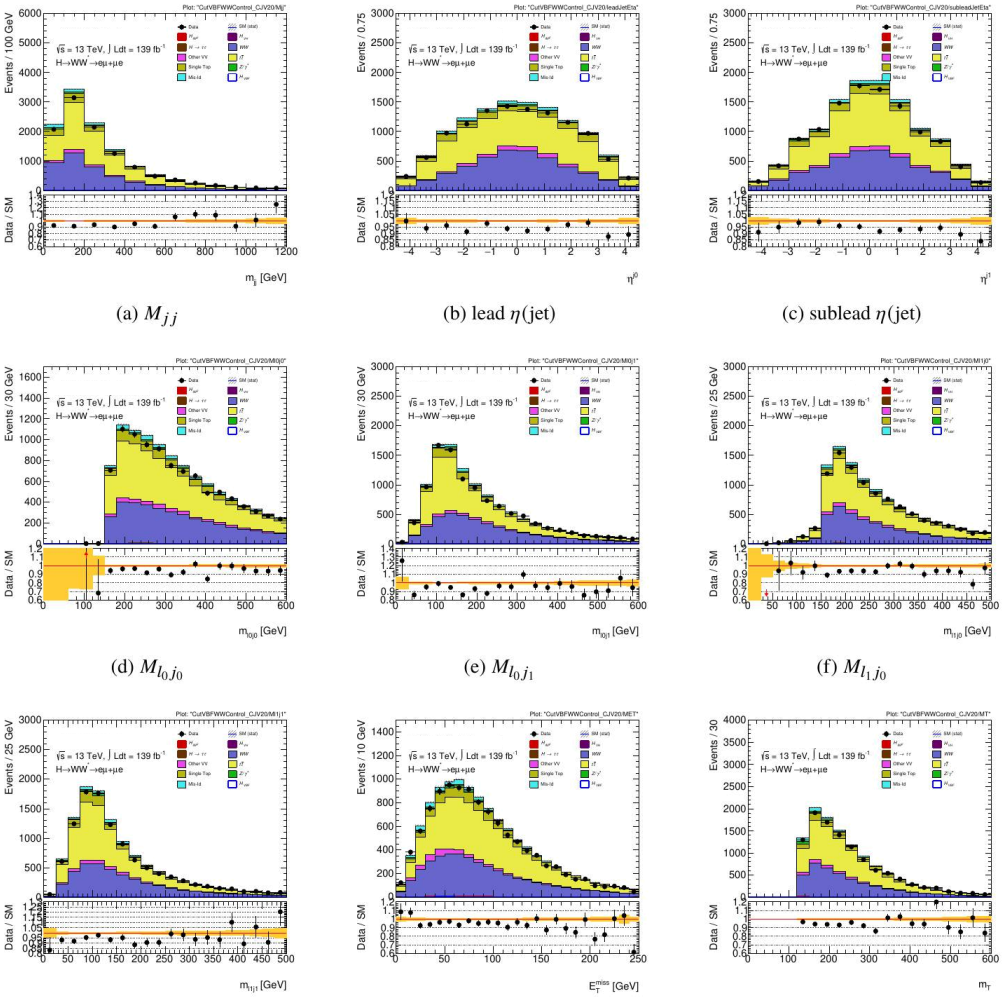


(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) p_{T}^{ll}

(d) lead $\eta(\ell)$ (e) sublead $\eta(\ell)$ (f) M_{ll}

(g) n_{jets} (h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

Figure 10.11. Distribution of observables after WW validation region cuts with statistical error



(a) M_{jj} (b) lead η(jet) (c) sublead η(jet)

(d) $M_{l_0 j_0}$ (e) $M_{l_0 j_1}$ (f) $M_{l_1 j_0}$

(g) $M_{l_1 j_1}$ (h) $E_{\mathrm{T}}^{\mathrm{miss}}$ (i) m_{T}

Figure 10.12. Distribution of observables after WW validation region cuts with statistical error

The control region is chosen to minimize both the statistical and the modeling uncertainties, in particular those originating from the modeling of higher order QCD corrections. The definition of the ggF control and validation regions is inspired from a subdivision of the samples based on the jet multiplicity. The regions use the same preselection cuts as the signal region with additional cuts listed in table 10.9

Region	cuts
ggF-CR	N jet > 1, b veto, $Z \to \tau\tau$ veto, CJV = True and OLV = False, or CJV = False and OLV = True, and $\Delta\phi_{ll} < 1.4$
ggF-VR2	N jet > 1, b veto, $Z \to \tau\tau$ veto, CJV = False and OLV = False, or CJV = False and OLV = True, and $\Delta\phi_{ll} < 1.4$
ggF-VR3	N jet < 2, b veto, $Z \to \tau\tau$ veto, and $\Delta\phi_{ll} < 1.4$

Table 10.9. ggF region definitions

Table 10.10 shows the cutflow efficiencies.

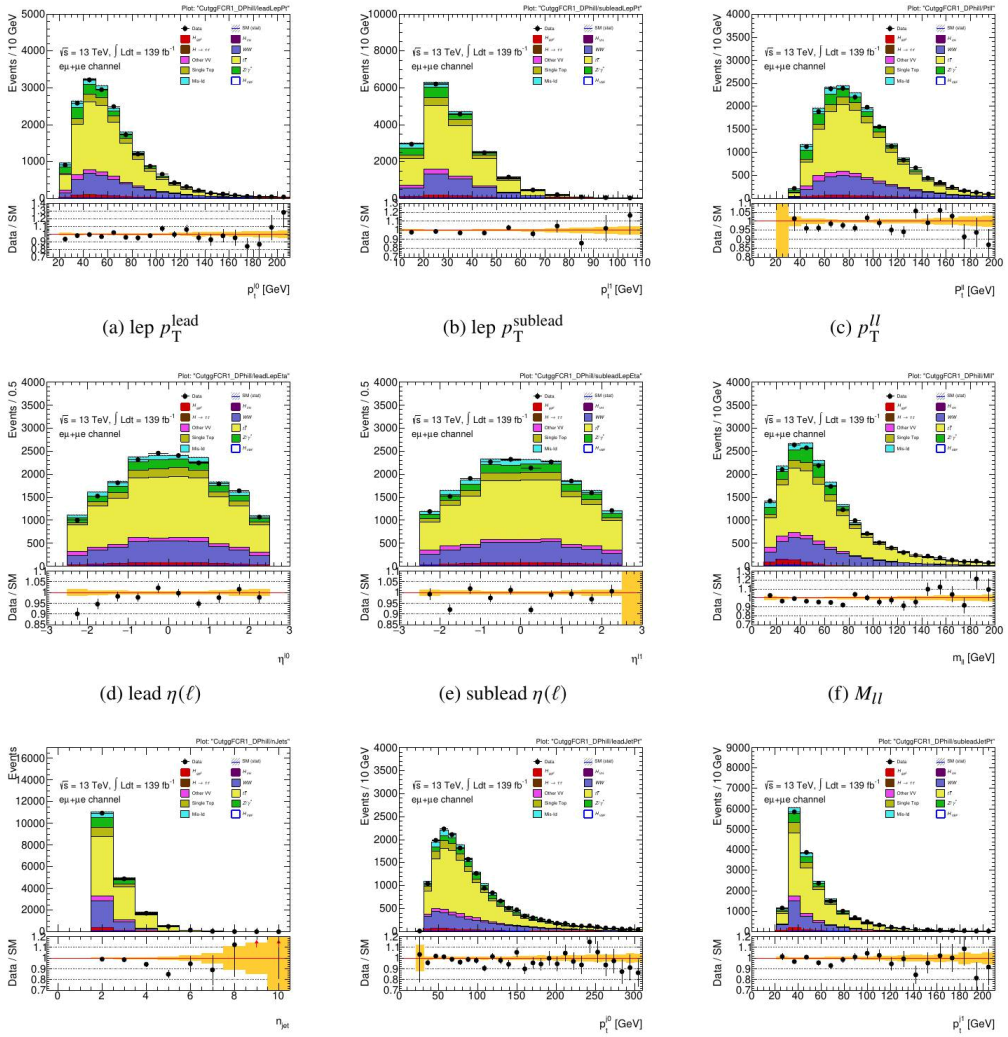
$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	H_{VBF}	H_{ggF}	VH+Htt	WW	Other VV	Top	Zjets	Mis-Id	Total Bkg	Significance	Data	Data/MC
ggF CR1	86.35 ± 0.28	452.30 ± 2.00	78.15 ± 0.47	3762.47 ± 11.55	705.00 ± 27.41	11477.59 ± 24.20	1447.73 ± 21.02	813.18 ± 27.25	18736.41 ± 51.57	0.63 ± 0.00	18228	0.97 ± 0.01
ggF VR2	7.37 ± 0.08	93.15 ± 0.89	16.11 ± 0.21	1029.85 ± 5.40	200.27 ± 11.86	3100.25 ± 12.31	460.10 ± 10.30	229.36 ± 16.02	5129.09 ± 26.17	0.10 ± 0.00	4824	0.94 ± 0.01
ggF VR3	152.51 ± 0.37	2625.91 ± 5.37	102.91 ± 0.56	19718.62 ± 46.59	2554.69 ± 42.53	13084.82 ± 29.93	1859.32 ± 40.53	2914.05 ± 43.18	42860.32 ± 91.71	0.74 ± 0.00	45254	1.05 ± 0.01

Table 10.10. Cutflow of ggF validation region cuts

The expected purities of the ggF Higgs boson production in the ggF-CR1 and the ggF-VR2 and ggF-VR3 regions are 2.5%, 1.9% and 6% respectively. The ggF-VR3 is the sample with the largest statistics, but it is also the sample with a kinematic selection that is farther away from the SR, as it requires events with 0 or 1 jets. The ggF-CR1 is the sample defined with kinematic requirements that are closer to those in the SR as only one of the OLV or CJV cuts fail at a time. The number of events in the ggF-CVR2 is about 26% of the number of events in the ggF-CR1 and therefore has little statistical power compared to the other two samples. Figure 10.13, 10.14 and 10.15 show the MC-data agreement in the ggF control regions

Figure 10.16, 10.17 and 10.18 show the MC-data agreement in the ggF control region.

Figure 10.19 and 10.20 show the MC-data agreement in the ggF validation region 3.

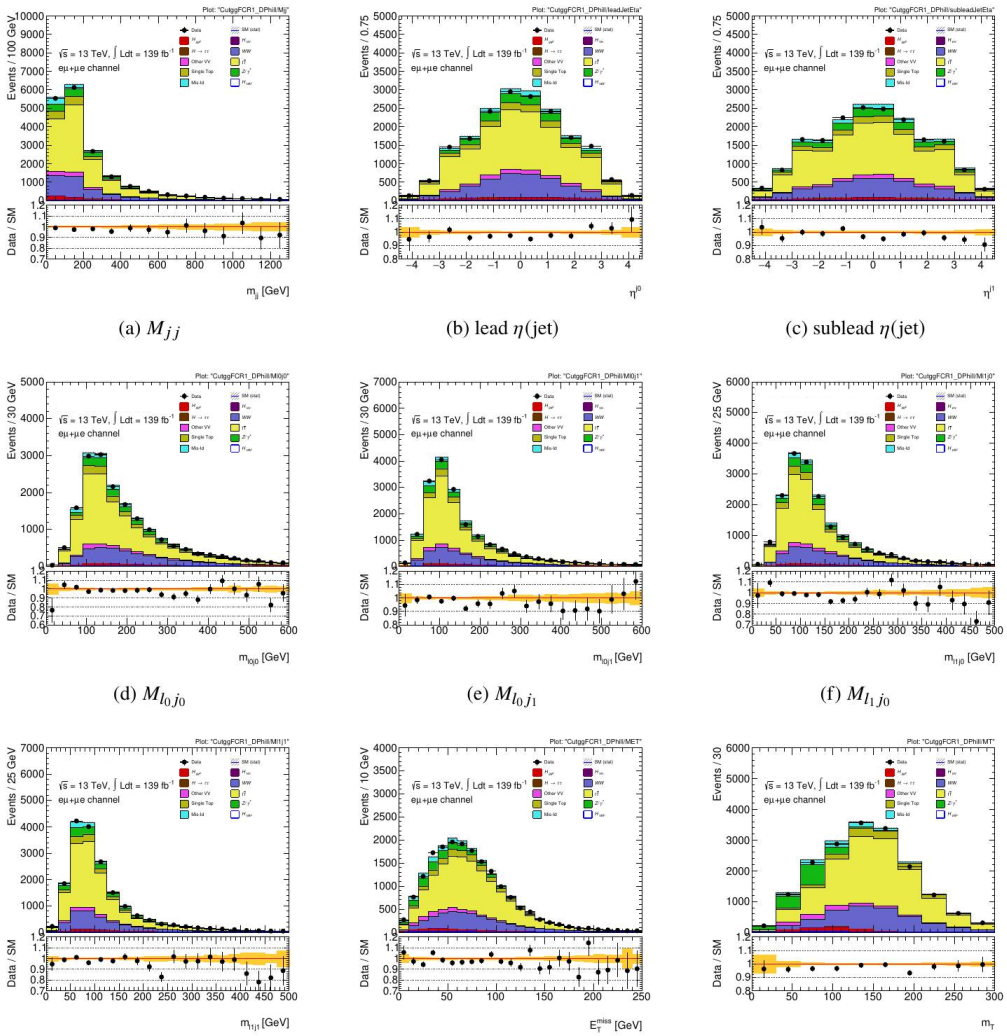


(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) p_{T}^{ll}

(d) lead $\eta(\ell)$ (e) sublead $\eta(\ell)$ (f) M_{ll}

(g) n_{jets} (h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

Figure 10.13. Distribution of observables after ggF control region cuts with statistical error

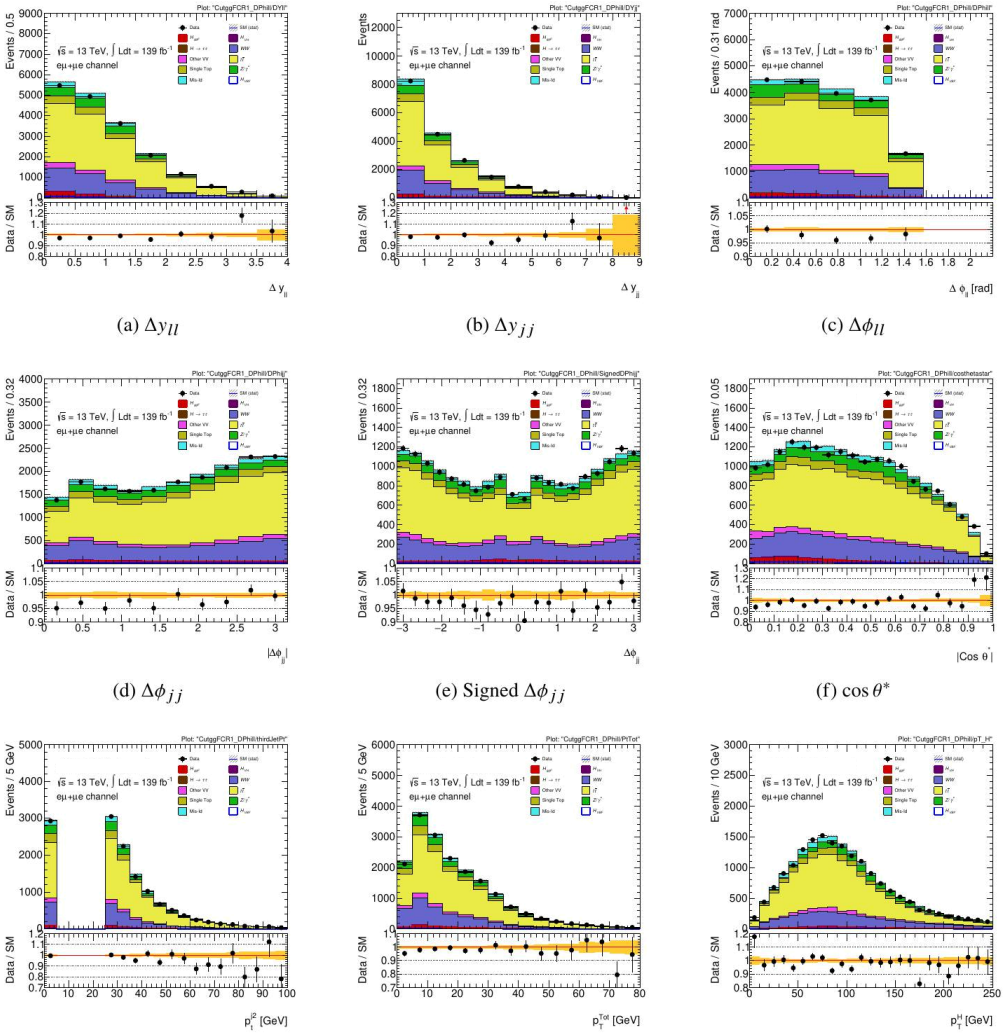


(a) M_{jj} (b) lead η(jet) (c) sublead η(jet)

(d) $M_{l_0 j_0}$ (e) $M_{l_0 j_1}$ (f) $M_{l_1 j_0}$

(g) $M_{l_1 j_1}$ (h) $E_{\mathrm{T}}^{\mathrm{miss}}$ (i) m_{T}

Figure 10.14. Distribution of observables after ggF control region cuts with statistical error

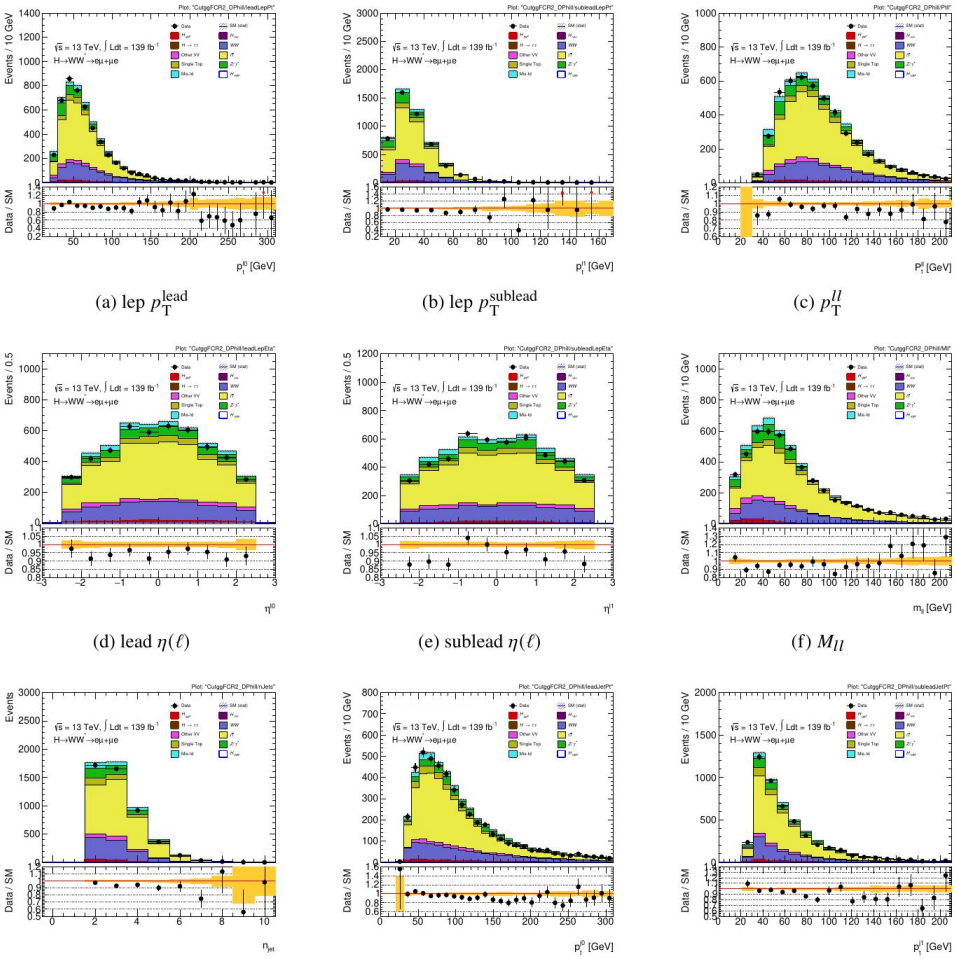


(a) Δy_{ll} (b) Δy_{jj} (c) $\Delta\phi_{ll}$

(d) $\Delta\phi_{jj}$ (e) Signed $\Delta\phi_{jj}$ (f) $\cos\theta^*$

(g) third jet p_{T} (h) $p_{\mathrm{T}}^{\mathrm{tot}}$ (i) p_{T}^{H}

Figure 10.15. Distribution of observables after ggF control region cuts with statistical error

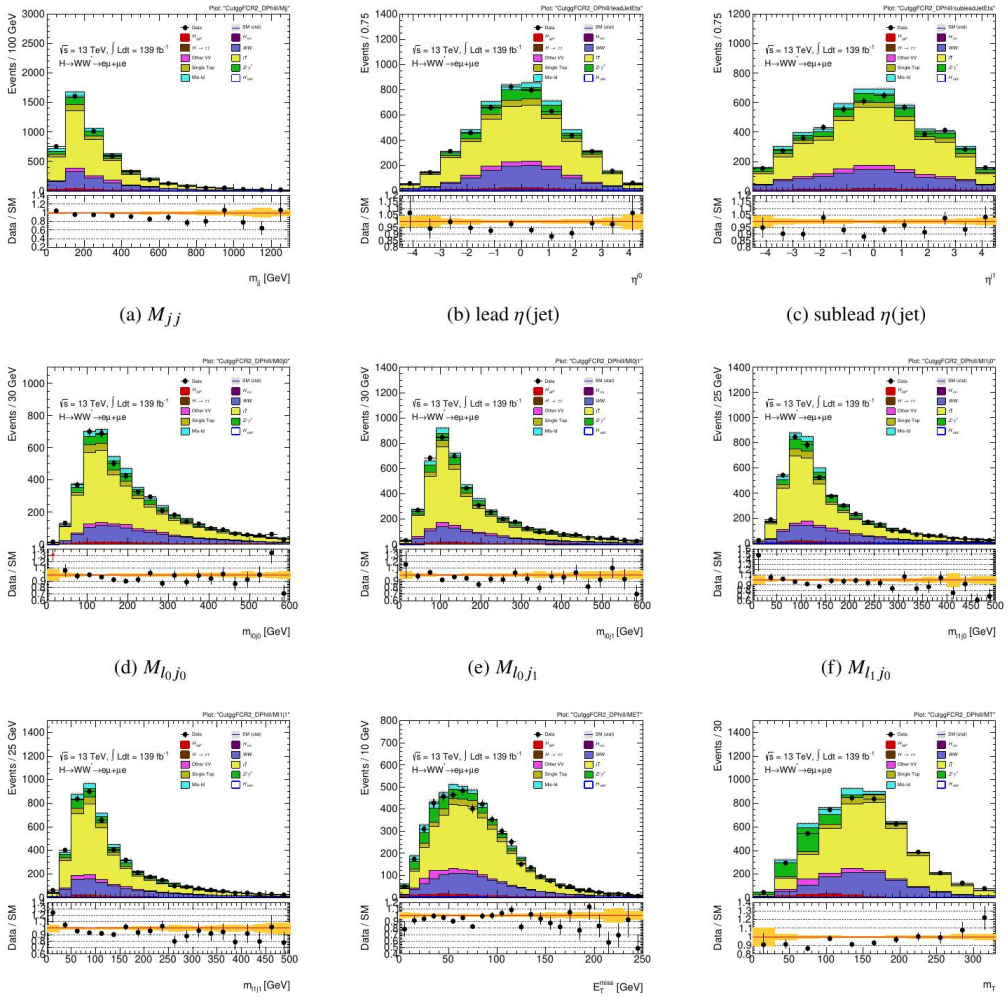


(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) p_{T}^{ll}

(d) lead $\eta(\ell)$ (e) sublead $\eta(\ell)$ (f) M_{ll}

(g) n_{jets} (h) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (i) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$

Figure 10.16. Distribution of observables after ggF validation region 2 region cuts with statistical error

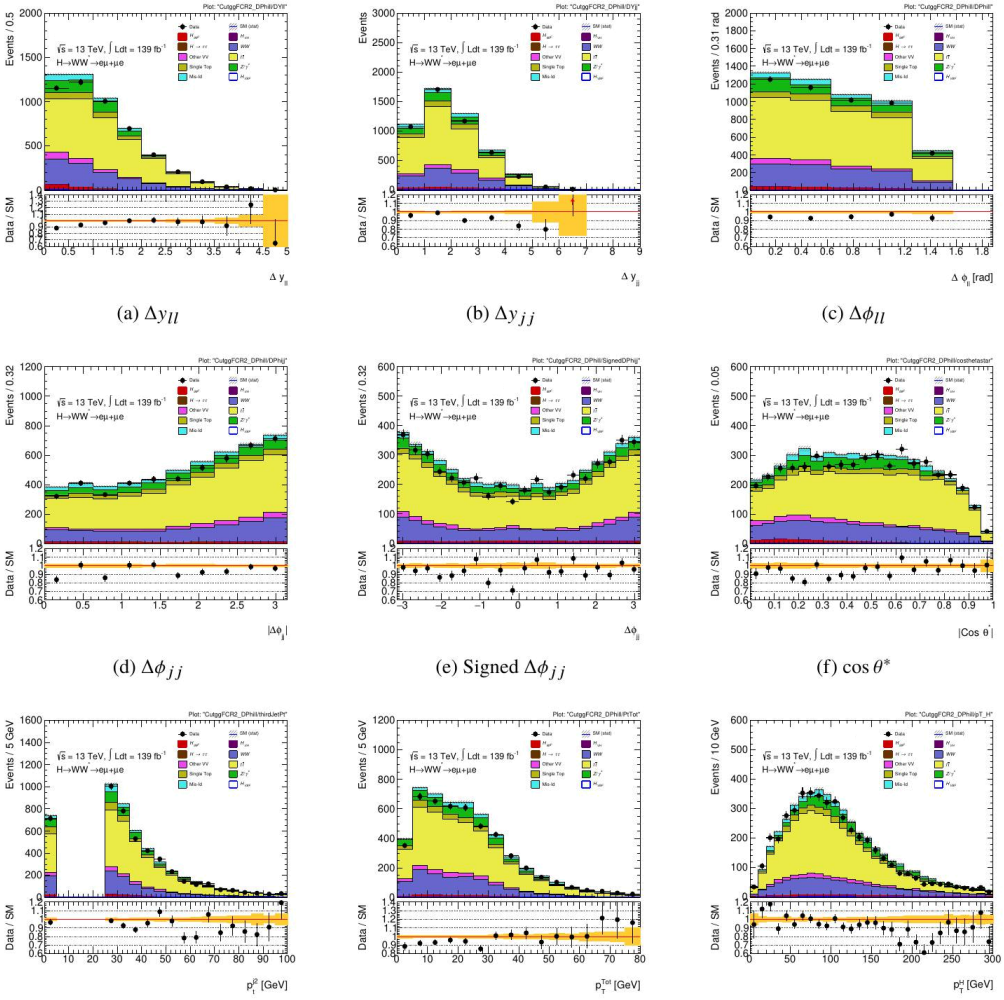


(a) M_{jj} (b) lead η(jet) (c) sublead η(jet)

(d) $M_{l_0 j_0}$ (e) $M_{l_0 j_1}$ (f) $M_{l_1 j_0}$

(g) $M_{l_1 j_1}$ (h) $E_{\mathrm{T}}^{\mathrm{miss}}$ (i) m_{T}

Figure 10.17. Distribution of observables after ggF validation region 2 region cuts with statistical error

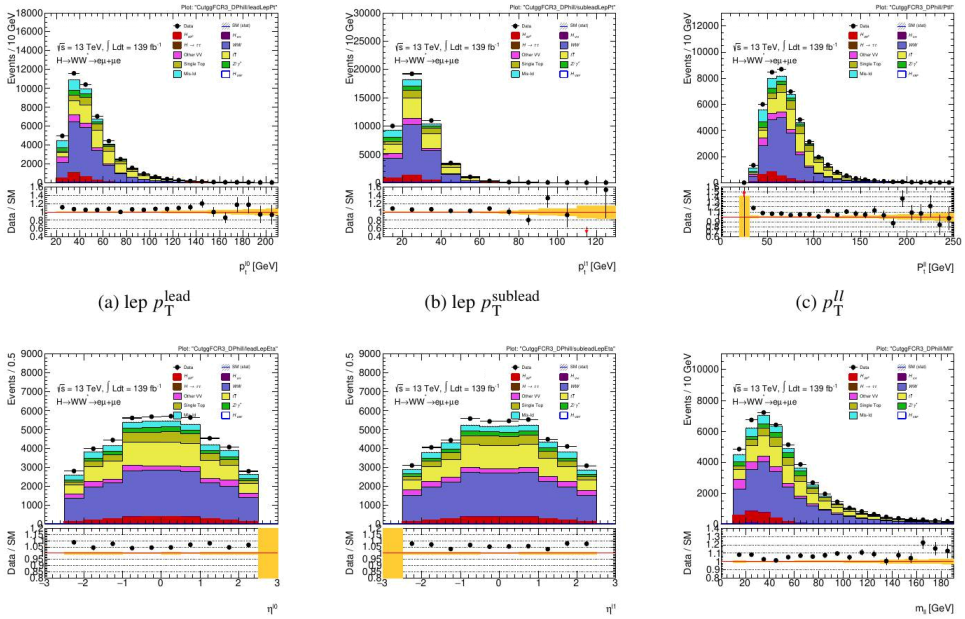


(a) Δy_{ll} (b) Δy_{jj} (c) $\Delta\phi_{ll}$

(d) $\Delta\phi_{jj}$ (e) Signed $\Delta\phi_{jj}$ (f) $\cos\theta^*$

(g) third jet p_{T} (h) $p_{\mathrm{T}}^{\mathrm{tot}}$ (i) p_{T}^{H}

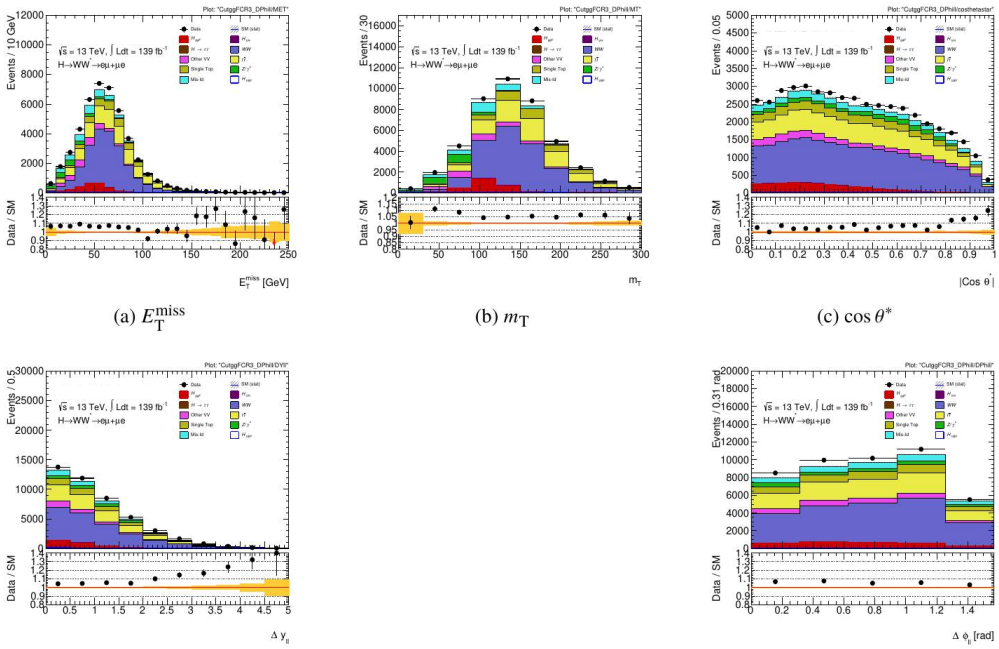
Figure 10.18. Distribution of observables after ggF validation region 2 region cuts with statistical error



(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) p_{T}^{ll}

(d) lead $\eta(\ell)$ (e) sublead $\eta(\ell)$ (f) M_{ll}

Figure 10.19. Distribution of observables after ggF validation region 3 region cuts with statistical error



(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) m_{T} (c) $\cos\theta^*$

(d) Δy_{ll} (e) $\Delta\phi_{ll}$

Figure 10.20. Distribution of observables after ggF validation region 3 region cuts with statistical error

10.6 W+jets Control Region

The W+jets control region is used to estimate the contribution of jets from W+jets processes that are misidentified as leptons, also known as "fakes". The W+jets region has the same kinematic requirements as the signal region but with an additional requirement of one identified ("id") lepton and one anti-identified lepton ("anti-id") instead of two id leptons in the signal region. The method to estimate fake lepton contribution is described in the following chapter. The cutflow is shown in table 10.11. The fake purity in this region is 76% after all cuts are applied

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Total Bkg	data	fakes	fake purity(%)
Channel Selection	8012632.06 ± 2327.10	13124937	5112304.94 ± 4305.85	38.95 ± 0.03
Trigger Selection	8012632.06 ± 2327.10	13124937	5112304.94 ± 4305.85	38.95 ± 0.03
Trigger Matching	7754276.84 ± 2199.12	13057932	5303655.16 ± 4230.14	40.62 ± 0.03
W+jets flavour split muon	5672759.42 ± 1876.51	10572340	4899580.58 ± 3754.15	46.34 ± 0.04
W+jets flavour split electron	3498327.82 ± 1524.29	7417426	3919098.18 ± 3121.04	52.84 ± 0.05
Jet Cleaning	3498327.82 ± 1524.29	7417426	3919098.18 ± 3121.04	52.84 ± 0.05
Overlap: Vgamma/Vjets	3498327.82 ± 1524.29	7417426	3919098.18 ± 3121.04	52.84 ± 0.05
Only two Leptons	3448417.08 ± 1499.17	7363366	3914948.92 ± 3100.14	53.17 ± 0.05
p_t^{lead} > 22 GeV	3448417.08 ± 1499.17	7363366	3914948.92 ± 3100.14	53.17 ± 0.05
p_t^{sublead} > 15 GeV	3447273.96 ± 1498.83	7360611	3913337.04 ± 3099.53	53.17 ± 0.05
Remove High Eventweight Events	3447325.12 ± 1489.49	7360611	3913285.88 ± 3095.03	53.17 ± 0.05
Remove High Eventweight Zjets Samples	3371562.75 ± 992.08	7360611	3989048.25 ± 2888.74	54.19 ± 0.04
OS Leptons	3370094.44 ± 991.88	7353604	3983509.56 ± 2887.46	54.17 ± 0.04
$M_{\ell\ell}$ > 10 GeV	3370035.94 ± 991.87	7353272	3983236.06 ± 2887.40	54.17 ± 0.04
SF: Z Veto	3370035.94 ± 991.87	7353272	3983236.06 ± 2887.40	54.17 ± 0.04
Leptons ID, singleFakes 1 anti-ID,1 ID, doubleFakes 2*anti-ID	764479.23 ± 519.68	1928013	1163533.77 ± 1482.59	60.35 ± 0.09
p_T^{miss} > 20 GeV	527004.88 ± 361.71	1169615	642610.12 ± 1140.37	54.94 ± 0.11
2-jet (30,30); no FF	385167.42 ± 196.07	656338	271170.58 ± 833.53	41.32 ± 0.14
b-veto	63869.88 ± 142.74	197677	133807.12 ± 466.96	67.69 ± 0.28
CJV (20GeV)	44256.34 ± 121.29	137560	93303.66 ± 390.22	67.83 ± 0.34
OLV bool	9437.81 ± 59.38	27999	18561.19 ± 177.55	66.29 ± 0.75
$Z \to \tau\tau$ veto	5237.79 ± 42.66	18022	12784.21 ± 140.86	70.94 ± 0.94
Mjj>200 GeV	3269.09 ± 32.93	11883	8613.91 ± 113.88	72.49 ± 1.17
DYjj>2.1	2851.94 ± 32.06	10879	8027.06 ± 109.12	73.78 ± 1.23
Mjj>450 GeV	1162.08 ± 19.87	4821	3658.92 ± 72.22	75.90 ± 1.85

Table 10.11. Cutflow of W+Jets control region cuts

CHAPTER 11

FIDUCIAL REGION

11.1 Fiducial and inclusive phase space

The experimental measurement of VBF h $\rightarrow WW \rightarrow \ell\nu\ell\nu$ production in the the eμ decay channel is reported as a fiducial cross section defined as follows:

$$\sigma^{\text{fid}}_{\text{VBF}h\rightarrow ww\rightarrow l\nu l\nu} = \frac{N_{\text{obs}} - N_{\text{bkg}}}{C \times \mathcal{L}}, \tag{11.1}$$

where $\mathcal{L}$ is the integrated luminosity, N_{obs} is the observed number of events, N_{bkg} is the estimated number of background events and C is a factor that accounts for detector inefficiencies and resolution effects. The latter is defined as the ratio of the number of reconstructed signal events after the final selection with electrons or muons in the final state (including electrons or muons from τ-decays) to the number of signal events generated in the fiducial region where only direct decays of W bosons to electrons and muons are considered.

The VBF $W \rightarrow WW \rightarrow \ell\nu\ell\nu$ cross section is evaluated in the fiducial phase space of the $e - \mu$ decay channel, as defined in Table 11.1 for the standard model Higgs of 125 GeV. The definition of the fiducial phase space is chosen to be as close as possible to the kinematic phase space at detector level in order to minimize theoretical extrapolation and model dependence.

Electrons and muons are required at particle level to stem from one of the W bosons produced in the hard scatter and are defined at the "dressed level", i.e. the momenta of photons emitted in a cone R = 0.1 around the lepton direction are added to the charged lepton momentum after QED final-state radiation. Final-state

particles, defined as those with a proper lifetime τ corresponding to $c\tau \geq 10$ mm, are clustered into jets (referred to as particle-level jets) using the same algorithm as for detector-level jets, i.e. the anti-k$_t$ algorithm with radius parameter R = 0.4. The selected charged leptons and neutrinos from W boson decays are not included in the jet clustering. The fiducial phase space at particle level does not make any requirement on jets containing b-quarks. The missing transverse momentum is defined at particle level as the transverse component of the vectorial sum of non-interacting particles. Its magnitude is denoted in Table 11.1 as MET. If a jet is within ΔR=0.4 from the selected electron or muon, the event is not considered [15].

Category	Fiducial Requirement	Cut Value
Leptons ($e+\mu$, $\mu+e$)	$\|Eta\|$	< 2.5
	$P_T$lead	$> 22GeV$
	$P_T$sublead	$> 15GeV$
	OS and OF pair	Yes
	ΔR(ℓ,ℓ)	> 0.1
	$\Delta\phi_{\ell\ell}$	< 1.4
	mll	$> 10GeV$
Neutrinos	met	$> 20GeV$
Jets (anti-k$_t$ R=0.4)	pTj	$> 30GeV$
	$\|\eta_j\|$	< 4.5
	N_{jets}	≥ 2
	$N_{\text{b-jet}}$	< 1
	M_{jj}	$> 450GeV$
	DY_{jj}	> 2.1
lepton-jet separation	ΔR(ℓ,jet)	> 0.4
VBF	OLV	Yes
	CJV	Yes

Table 11.1. Definition of Fiducial Phase Space for VBF $h \to ww \to \ell\nu\ell\nu$ in the $e-\mu$ decay channel.

The fiducial cross section is extrapolated in a broader phase space, referred to as the "inclusive" phase space in the following. An acceptance correction factor A is

used to extrapolate the fiducial cross section to an inclusive cross section, according to the following formula:

$$\sigma_{VBF}^{\text{incl}} = \frac{\sigma_{\text{VBF}}^{\text{fid}}}{A}, \tag{11.2}$$

Theoretical calculations are used to estimate the acceptance correction A from the fiducial to the inclusive phase space, and the extrapolated cross section is referred to in the following as inclusive cross section. The theoretical cross section is calculated at NNLO in QCD and NLO at EW, and the prediction given is

$$\sigma_{VBF,H\to WW\to\ell\nu\ell\nu}^{theory} = 85.8 \pm 2.2 \text{ fb} \tag{11.3}$$

The definition of the inclusive phase space is summarized in Table 11.2 [18]. The inclusive cross section is extracted for Higgs production in the vector boson fusion mode with two associated jets in the final state.

Category	Inclusive Requirement	Cut Value
Jets (anti-k$_t$ R=0.4)	pT_j	$\geq 30GeV$
	N_{jets}	≥ 2

Table 11.2. Definition of the Inclusive Phase Space for VBF H production.

The differential cross sections are determined for kinematic variables unfolded to particle-level in the same fiducial phase space as in Table 11.1. The unfolding procedure corrects for migrations between bins in the distributions during the reconstruction of events, and applies fiducial as well as reconstruction efficiency corrections. The fiducial corrections take into account events which are reconstructed in the signal region, the reconstruction efficiency corrects for events inside the fiducial region which are not reconstructed in the signal region due to detector inefficiencies. Tests with MC simulation demonstrate that the method is successful in retrieving the particle-

level distribution in the fiducial region from the reconstructed distribution at detector level in the signal region.

11.2 Detector And Acceptance Corrections

The C factor described in the previous section, is defined as $N_{detector}/N_{truth}$, the estimated yield in the reconstruction level Monte Carlo over the estimated yield in the truth level Monte Carlo. The C-factor value is reported in table 11.3

The acceptance factor is defined as $A = N_{fid}^{truth}/N_{inc}^{truth} = \sigma_{fid}^{truth}/\sigma_{inc}^{truth}$,where the inclusive cross section is taken from the Higgs Cross Section Working Group for Yellow Book Report 4 [16], and the fiducial cross section is calculated at truth level at NLO in QCD and EW, including the theoretical uncertainties described in the following chapters. The results of the fiducial cross section estimates is

$$\sigma_{fid}^{truth} = 2.07 \pm 0.09 \text{ fb} \tag{11.4}$$

And the acceptance factors for three inclusive phase spaces are in table 11.3 [17].

C-factor	0.448
Acceptance(VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$)	0.024 ± 0.001
Acceptance(VBF $H \rightarrow WW^*$)	0.0026 ± 0.0001
Acceptance(VBF H)	0.00055 ± 0.00003

Table 11.3. C-factor and Acceptance factor.

CHAPTER 12

BACKGROUNDS

12.1 Summary Of Backgrounds

Among the $H \to WW$ production modes, the final state with the highest purity at the LHC occurs when each W boson decays leptonically, $W \to \ell\nu$, where ℓ is an electron or muon and ν is a neutrino. A background process has a final state indistinguishable from the final state of the signal process (again, two opposite flavor leptons and two neutrinos). Backgrounds that produce the exact same final particles are irreducible. Other backgrounds have a physically different final state but give similar enough signatures to the signal process that the final state is misreconstructed as a signal process. For example a W+jets decay contains a jet in the final state that is mis-identified as a lepton, this is called the "fakes" background. The major backgrounds are explored in table 12.1.

The major backgrounds are standard model processes that produce a WW part that then decays leptonically, the most significant of which are the SM continuum WW, $t\bar{t}$, and single top production. The other notable backgrounds are other Higgs production models, such as ggF, and Drell Yan events [75].

The most significant background are from top quark decays. A pair of top quarks ($t\bar{t}$) decay leptonically through a WW production which gives identical leptons and missing energy signature. In addition to $t\bar{t}$ background, single top, associated with a b quark or W boson can also decay leptonically. These events can be suppressed with the b-jet cut on b- hadron decays. $t\bar{t}$ remains the most significant background due to inefficiencies in the b-jet tagging.

Name	Process	Feature(s)
WW	WW	Irreducible
Top quarks		
$t\bar{t}$	$t\bar{t} \to Wb\,W\bar{b}$	Unidentified b-quarks
t	tW	Unidentified b-quark
	$t\bar{b}$, $tq\bar{b}$	q or b misidentified as ℓ; unidentified b-quarks
Misidentified leptons (Mis-id)		
Wj	$W+$jet(s)	j misidentified as ℓ
jj	Multijet production	jj misidentified as $\ell\ell$; misidentified neutrinos
Other dibosons		
VV	$W\gamma$	γ misidentified as e
	$W\gamma^*$, WZ, $ZZ \to \ell\ell\ell\ell$	Unidentified lepton(s)
	$ZZ \to \ell\ell\nu\nu$	Irreducible
	$Z\gamma$	γ misidentified as e; unidentified lepton
Drell-Yan (DY)		
$e^+e^-/\mu\mu$	$Z/\mathrm{DY} \to ee,\ \mu\mu$	Misidentified neutrinos
$\tau^+\tau^-$	$Z/\mathrm{DY} \to \tau\tau \to \ell\nu\nu\,\ell\nu\nu$	Irreducible

Table 12.1. Backgrounds to the measurement in the final state with two charged leptons and neutrinos

In the SM continuum WW process, a WW pair does not originate from the Higgs decay, but is produced from q - q, q - g or g - g interactions. The subsequent leptonic decay of the pair can generate the same signature as the signal events. The kinematics between the SM WW production and $H \to WW$ production differ in so far as the leptons direction, i.e. in the $H \to WW$ case the leptons tends to be produced at a smaller opening angle due to the spin correlations of the W-pair originating from the decay of the Higgs boson that is a scalar particle in the SM. This background process can be suppressed with cuts on the opening angle between the two leptons.

The Drell-Yan (DY) process is the dominant source of events with two leptons from the Z boson decay, i.e. Z+jets. Drell Yan process is when a quark and anti-quark annihilate in the S channel, producing a lepton-anti-lepton pair, mediated by a virtual photon or a Z boson (as is the case for the dominant background in this analysis). The process is associated with high-energy QCD jets emitted from the initial state quarks, for example in the Feynman diagram in figure 12.1 Because

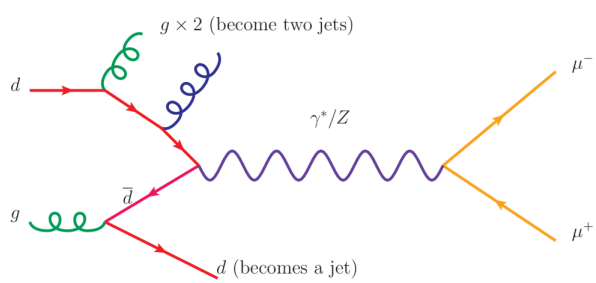


Figure 12.1. Feynamn diagram for Drell Yan process with associated QCD jets

of the leptonic final states and the high energy jets, Z+jets represents a significant irreducible background to the analysis. The DY background events typically feature small lepton momenta, and missing energy either aligned with one lepton or within the sector enclosed by the leptons, and these features can be exploited to reduce this background contamination the signal region.

The final significant background are other Higgs production modes, largely dominated by ggF. Events with a Higgs boson decaying in to a WW pair produced via the ggF mode constitute a challenging background in the case in which two jets are produced in association with the Higgs boson. The kinematic configuration of the jets and leptons can be used to distinguish these two production modes. Other Higgs production modes, such as associated VH and ttH, also contribute to the background but give a very small contribution due to their small production cross sections with respect to the VBF signal.

The backgrounds with the highest level of contamination in the VBF SR are modeled using Monte Carlo samples normalized to data, and include theoretical uncertainties on the extrapolation from the normalization region (CR) to the signal region, and on the shape of the distribution used in the likelihood fit. These include $t\bar{t}$, WW, DY and ggF Higgs production. For backgrounds that give a subdominant contribution, both the shape and the overall normalization is estimated from MC simulations. For the W+jets and multijet backgrounds, the high rates and the uncertainties in modeling misidentified leptons a data-driven model of the kinematic distributions is used, as described in the following section.

12.2 Fake Lepton Backgrounds

The fake lepton background arises from non-prompt leptons produced either by the decay of hadrons containing a "heavy flavor" quark or by the fragmentation of a quark or gluon.

This type of background is estimated using data in a dedicated sample enriched with jets, selected such that the requirements are as close as possible to those used for the signal region selection, except that one or both of the leptons "anti-identified", i.e. the lepton is reconstructed such as passing looser identification and isolation criteria than the leptons selected in the signal region, and fails the SR lepton identification re-

quirements. These criteria are listed in table 12.2. The probability of mis-identifying a jet as a lepton is extrapolated from the jet-enriched data sample to the signal region using "fake factors", defined as the ratio of identified to anti-identified leptons. Lepton properties that commonly affect the fake facto, are the transverse momentum p_T and pseudo-rapidity. Therefore the computation of fake factors is done as a function of these two variables in a jet-enriched data sample, either di-jet or Z+jets. The contamination in the jet-enriched data sample from the other background processes are subtracted before the fake factor calculation [79].

Electron		Muon	
identified	anti-identified	identified	anti-identified
$p_T > 15$ GeV		$p_T > 15$ GeV	
$\|\eta\| < 2.47$,excluding $1.37 < \|\eta\| < 1.52$		$\|\eta\| < 2.45$	
$\|z_0 \sin\theta\| < 0.5$ mm		$\|z_0 \sin\theta\| < 0.5$ mm	
$\|d_0\|/\sigma(d_0) < 5$		$\|d_0\|/\sigma(d_0) < 3$	$\|d_0\|/\sigma(d_0) < 15$
Pass LHTight if $p_T < 25$ GeV Pass LHMedium if $p_T > 25$ GeV Pass FCTight isolation	Pass LHLoose	Pass Quality Tight Pass FCTight isolation	Pass Quality Medium
AUTHOR = 1			
	Veto against identified electron		Veto against identified muon

Table 12.2. Requirements for fully identified and anti-identified leptons.

The Z+jets fake factors are used as for the estimation of the fake contribution from W+jets processes. The di-jet fake factors are used to estimate the background originating from the W+jets events in which the fake lepton is the only object firing the trigger. Table 29 gives a detailed overview of the strategy followed for the fake factor application in the signal region. This section provides a description of the fake-factor method used to estimate the W+jets background.

The analysis signal region requires two leptons passing a stringent lepton identification selection (called "id" leptons), therefore background events from misidentified leptons can be classified as "single-fake" or "double-fake". The single-fake backgrounds have one real lepton from a W and one misidentified lepton, so come from the W+jets processes. The "double fake" background denotes events with both lep-

tons misidentified, from QCD multi-jet processes. The sample of events in the signal region can then be expressed as

$$N_{id+id} = N^{EW}_{id+id} + N^{W+jets}_{id+id} + N^{QCD}_{id+id} \tag{12.1}$$

Where N^{EW}_{id+id} term denotes all processes with two real leptons (including the signal) and the last two terms denote the single- and double fake backgrounds. To estimate the fake backgrounds a W+jets control sample is defined with one id lepton and the other lepton being anti-identified;

$$N_{id+anti-id} = N^{EWMC}_{id+anti-id} + N^{W+jets}_{id+anti-id} + N^{QCD}_{id+anti-id} \tag{12.2}$$

The $N^{EWMC}_{id+anti-id}$ term is a two real leptons process that contaminate the W+jets region, which is estimated from Monte Carlo. It is then subtracted from the data in the region to obtain the number of events with fake backgrounds only. The QCD background is accounted for explicitly in the signal region, so it is subtracted off to avoid double counting it when estimating fake backgrounds. The size of the QCD multi-jet contribution may be estimated in a region defined by two anti-id leptons.

As the fake factors are derived in a Z+jets sample but applied to a W+jets region, it is important to assess any difference in fake composition between these two processes. The fake factors are corrected from Z+jets to opposite charge W+jets based on predictions from simulation. The fake factor ratio, denoted "correction factor" is computed in MC, separately for muons and electrons.

12.3 Zjets Fake Factor Definition

The fake factor is measured as a ratio of events in a three-lepton selection:

$$FF = \frac{N^{i,i,i} - N^{i,i,i}_{non_z jets}}{N^{i,i,a} - N^{i,i,a}_{non_z jets}} \tag{12.3}$$

	Triggger Fired	
Leptons	dileptons	single leptons
1 anti-ID	1 Z+jets FF	1 dijet or 1 Z+jets FF
2 anti-ID	2 Z+jets FF	1 dijet and 1 Z+jets FF

Table 12.3. Fake factor application in the SR. When the dilepton trigger fires, for the fake leptons Z+jets fake factors are applied. However, for the case where a single-lepton trigger fires a di-jet and/or a Z+jets fake factors are used depending which of the leptons fires the trigger. If one of the leptons passes the "anti-id" selections and is the one which fires a single-lepton trigger, a di-jet fake factor is used, whereas for the case where the real lepton fires the single-lepton trigger, the nominal Z+jets fake factor is used instead for the fake lepton. If both leptons in the event are fakes and a single-lepton trigger is fired, a di-jet fake factor for the one which fires the trigger and for the other the nominal Z+jets fake factor are applied. The dijet trigger apply looser lepton quality criteria than those used for the "anti-id" selection. Thus, the dilepton triggers do not cause a trigger bias which is corrected for via the dijet fake factors.

where the superscript 'i' stands for an "id" and 'a' for an "anti-id" lepton.

A sample sufficiently pure in Z+fake lepton is considered and events with three leptons with p_T >15 GeV are selected. Two of the leptons need to be of the same flavour and opposite charge to qualify as originating from the Z decay. They are identified with the same criteria as "id" criteria in Table 12.2 with the exception that they only need to pass FCLoose isolation and a "loose-with-b-layer" electron or "Medium" muon quality working point. The invariant-mass windows of [80, 110] GeV and [70, 110] GeV are required if the two leptons are electrons or muons respectively. If multiple lepton combinations satisfy this requirement, the pair with the invariant mass closest to the Z boson is chosen. The remaining lepton is the fake candidate. When applied to a Z+jets MC sample, this algorithm correctly assigns leptons to the Z boson in about 99% of all cases. The reconstructed Z-boson peak can be seen in Figure 12.2.

To reject events with leptonic WZ decays, the following cut is made:

$$m_T^W = \sqrt{E_T^{miss} \cdot E_T^{fake.}(1 - cos(\phi))} < 50GeV \tag{12.4}$$

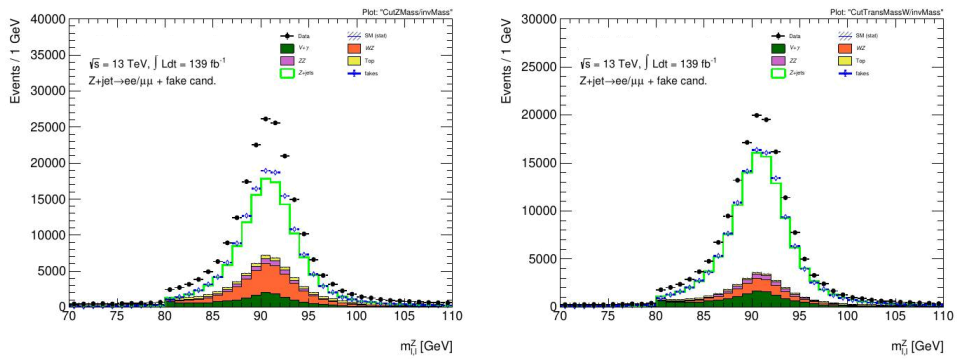


Figure 12.2. Distributions of the invariant mass of the reconstructed Z-boson candidate. The shape of the distribution agrees nicely between data and MC. At the stage of Z-boson identification (left), the normalization disagrees slightly. After applying a cut on the transverse mass $m_T^W < 50$ GeV(right) the normalization agrees nicely. Normalization factors are not applied. The stacked histograms are background MC processes not including Z+fake. The measured data is shown in black points. The blue points are the data-driven Z+fake estimate. They are calculated by taking the difference between data and the stacked MC processes. The agreement between data and MC is seen by comparing the blue fake estimate to green Z+jets MC.

where ϕ is the angle between the missing energy and the fake candidate. The pseudo-rapidity of the fake candidate is restricted to $|\eta| < 2.5$ for muons and $|\eta| < 2.47$ for electrons, which corresponds to the lepton acceptance region in the main analysis. Then, for each fake lepton type, electron and muon, two regions are constructed depending on whether the fake lepton passes or fails the "id" cuts. The region with $m_T^W > 50$ GeV, where the fake candidate is required to pass the "id" criteria, is used as a control region for the WZ process. The region is very pure and a normalization factor of 0.99 ± 0.01 is extracted, shown in figure 12.3. A full cutflow can be found in Table 12.4. A reasonable agreement is seen between the MC estimate of Z+jets events and its data-driven estimate obtained by substracting the non-Z+jets MC events from the selected events in data.

To reject events with leptonic WZ decays, the following cut is made:

$$m_T^W = \sqrt{E_T^{miss} \cdot E_T^{fake.}(1 - cos(\phi))} < 50 \text{ GeV} \tag{12.5}$$

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Z+jets Powheg	$V+\gamma$	WW	WZ	ZZ	Top	VVV	Total Bkg	Data	fakes	ZFake purity(%)
Channel Selection	246181.21 ± 268.27	105001.82 ± 743.46	1002.92 ± 8.44	49267.44 ± 1727.21	14923.47 ± 114.48	86488.80 ± 62.61	240.09 ± 0.94	256924.54 ± 1884.96	605458	348533.46 ± 2039.25	57.57 ± 0.34
lll final state	246181.21 ± 268.27	105001.82 ± 743.46	1002.92 ± 8.44	49267.44 ± 1727.21	14923.47 ± 114.48	86488.80 ± 62.61	240.09 ± 0.94	256924.54 ± 1884.96	605458	348533.46 ± 2039.25	57.57 ± 0.34
All lep p_T > 15 GeV	246181.21 ± 268.27	105001.82 ± 743.46	1002.92 ± 8.44	49267.44 ± 1727.21	14923.47 ± 114.48	86488.80 ± 62.61	240.09 ± 0.94	256924.54 ± 1884.96	605458	348533.46 ± 2039.25	57.57 ± 0.34
Overlap: Vgamma/Vjets	246181.21 ± 268.27	105001.82 ± 743.46	1002.92 ± 8.44	49267.44 ± 1727.21	14923.47 ± 114.48	86488.80 ± 62.61	240.09 ± 0.94	256924.54 ± 1884.96	605458	348533.46 ± 2039.25	57.57 ± 0.34
Z-tagging	139471.86 ± 198.06	18307.04 ± 317.06	125.51 ± 2.91	32270.51 ± 45.33	5995.85 ± 26.42	10694.83 ± 20.85	86.15 ± 0.44	67479.87 ± 322.06	214484	147004.13 ± 564.10	68.54 ± 0.30
TransMassW	125456.68 ± 187.76	15585.31 ± 296.36	45.58 ± 1.87	10126.93 ± 23.07	4267.69 ± 22.15	4059.09 ± 12.90	21.99 ± 0.22	34106.59 ± 298.37	161198	127091.41 ± 500.22	78.84 ± 0.37
fake lepton ID or AntiID	63284.26 ± 136.45	4592.08 ± 163.39	20.73 ± 1.22	9726.48 ± 22.73	4036.29 ± 21.04	3515.93 ± 11.96	21.14 ± 0.21	21912.65 ± 166.74	86206	64293.35 ± 337.65	74.58 ± 0.47
fake type: electron	34343.60 ± 96.99	3850.64 ± 142.84	12.72 ± 1.01	4375.13 ± 16.19	1849.83 ± 11.37	801.10 ± 5.37	9.42 ± 0.14	10898.84 ± 144.31	41125	30226.16 ± 248.90	73.50 ± 0.71
fake el η cut	27149.91 ± 86.45	3417.59 ± 135.42	9.93 ± 0.93	4131.62 ± 15.89	1742.41 ± 10.48	741.56 ± 5.16	8.93 ± 0.14	10052.03 ± 136.85	34003	23950.97 ± 229.63	70.44 ± 0.78
Scale factors				NF = 0.99 ± 0.01				NF = 0.99 ± 0.01		NF = 0.99 ± 0.01	
ID cuts	2429.00 ± 26.18	789.22 ± 63.23	1.24 ± 0.31	3367.53 ± 14.08	1272.17 ± 7.08	277.19 ± 2.63	7.40 ± 0.12	5714.74 ± 65.21	7996	2281.26 ± 110.68	28.53 ± 1.42
Scale factors				NF = 0.99 ± 0.01				NF = 0.99 ± 0.01		NF = 0.99 ± 0.01	
Anti-ID cuts	24720.92 ± 82.39	2628.37 ± 119.75	8.69 ± 0.88	736.64 ± 7.13	470.23 ± 7.72	464.37 ± 4.44	1.53 ± 0.06	4309.83 ± 120.29	26007	21697.17 ± 201.19	83.43 ± 0.93
fake type: muon	28940.67 ± 95.98	741.44 ± 79.33	8.00 ± 0.69	5351.35 ± 15.95	2186.46 ± 17.70	2714.83 ± 10.68	11.72 ± 0.16	11013.81 ± 83.52	45081	34067.19 ± 228.16	75.57 ± 0.62
fake mu η cut	28059.83 ± 94.46	692.08 ± 77.92	7.65 ± 0.69	5240.16 ± 15.77	2142.10 ± 17.54	2637.08 ± 10.51	11.45 ± 0.16	10730.52 ± 82.09	43638	32907.48 ± 224.45	75.41 ± 0.63
Scale factors				NF = 0.99 ± 0.01				NF = 0.99 ± 0.01		NF = 0.99 ± 0.01	
ID cuts	1429.37 ± 21.21	26.49 ± 9.87	1.01 ± 0.24	4349.23 ± 14.46	1364.69 ± 11.69	387.70 ± 3.28	9.54 ± 0.15	6138.65 ± 21.30	7464	1325.35 ± 88.98	17.76 ± 1.21
Scale factors				NF = 0.99 ± 0.01				NF = 0.99 ± 0.01		NF = 0.99 ± 0.01	
Anti-ID cuts	26630.46 ± 92.05	665.59 ± 77.29	6.65 ± 0.64	856.10 ± 6.03	777.41 ± 13.08	2249.38 ± 9.98	1.92 ± 0.07	4557.04 ± 79.26	36174	31616.96 ± 206.05	87.40 ± 0.73

Table 12.4. Cutflow of fake-lepton id region

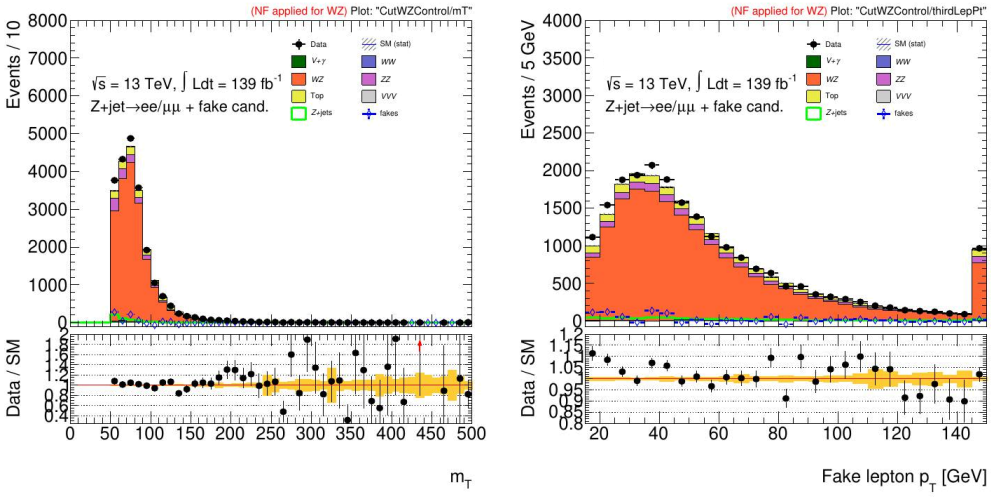


Figure 12.3. Distributions in the WZCR show good modeling. A cut of $m_T > 50$ GeV(left) and the requirement that the fake candidate lepton is ID (right) are applied for the WZ CR.

The distributions of the electron and muon kinematic variables of the fake candidate leptons are shown in Figure 12.4. The Z+jets prediction agrees well with the data-driven estimate of the fake yield in the "id" selection in both electron and muon channels. In the "anti-id" selections of electron fake candidates, the MC slighty overestimates the fake yield, while the opposite occurs in the case of the muon "anti-id" selection.

To reject events with leptonic WZ decays, the following cut is made:

$$m_T^W = \sqrt{E_T^{miss} \cdot E_T^{fake.}(1 - cos(\phi))} < 50GeV \tag{12.6}$$

The uncertainties on the Z+jets fake factors have contributions from systematic and statistical sources. The systematic uncertainty comes from the non-Z+jets (EW background) MC processes which are subtracted from the data yield. The fake factors and their statistical uncertainties are shown in Figure 12.5 for the two lepton types. The plots also summarize the systematic uncertainty from the electroweak subtraction normalization.

12.3.1 Pileup Dependence

The high-luminosity produced by the LHC leads to multiple proton-proton interactions per beam crossing in ATLAS, known as pile-up. As a result, every physics object is affected by pile-up, from additional energy contributions in jets to the misreconstruction of background. The properties and multiplicity of these collisions depend on the luminosity as well as the kinematics of the final state objects. The possible dependence of the Z+jets fake factor estimation related to pile-up was investigated. The pile-up distribution for each campaign separately as well as for the full Run 2 data is presented in Figure 12.6

In order to model the pile-up conditions correctly, the μ data scale factor of 1/1.03 has been applied before reweighting the μ distribution in Monte Carlo to match. For

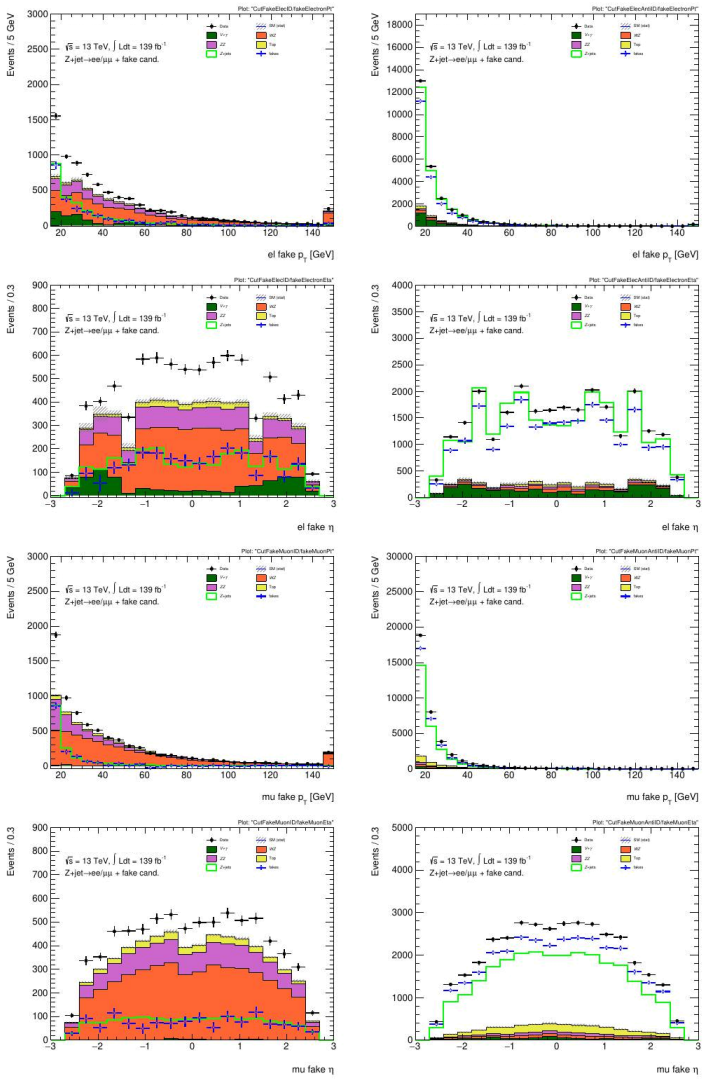


Figure 12.4. Transverse momentum and pseudorapidity distributions of the fake candidate. The top four plots are electron fake candidates and the bottom four are muon fake candidates. Both the ID (left) and the AntiID (right) selections are shown. The stacked histograms are background, i.e. non-Z+jets MC processes. The difference between the data yield and the stacked histograms constitutes the estimate for Z+jets. The fake estimate is also shown in blue markers and is compared to the non-stacked Z+jets MC prediction.

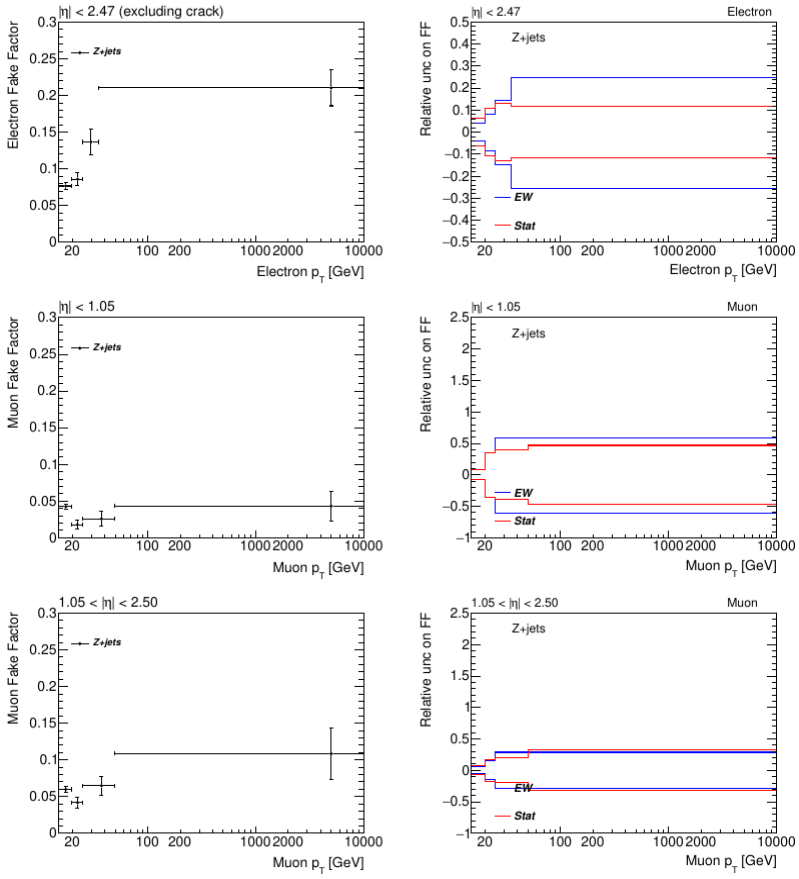
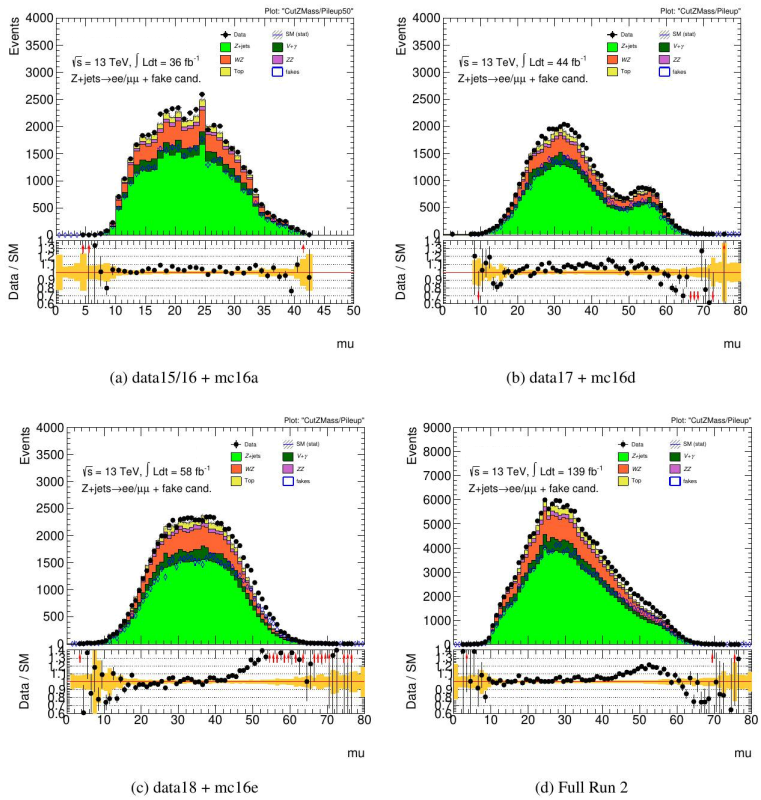


Figure 12.5. Fake factors derived in the three-lepton selection. The top two plots show the fake factor (left) and its relative uncertainties (right) as a function of p_T for electrons. The middle two plots show the muon fake factors in the low-η region and the bottom two plots the same in the high-η region. Stat denotes the statistical uncertainties on the fake factors while EW denotes the uncertainty due to the electroweak backgrounds that enter to the Z+jets fake factor estimate.



(a) data15/16 + mc16a

(b) data17 + mc16d

(c) data18 + mc16e

(d) Full Run 2

Figure 12.6. The pileup (μ) distributions for each campaign separately and for the full Run 2 data. For c16a, the average number of interactions per bunch crossing is used, while for c16d and c16e the actual number of interactions per bunch crossing is used

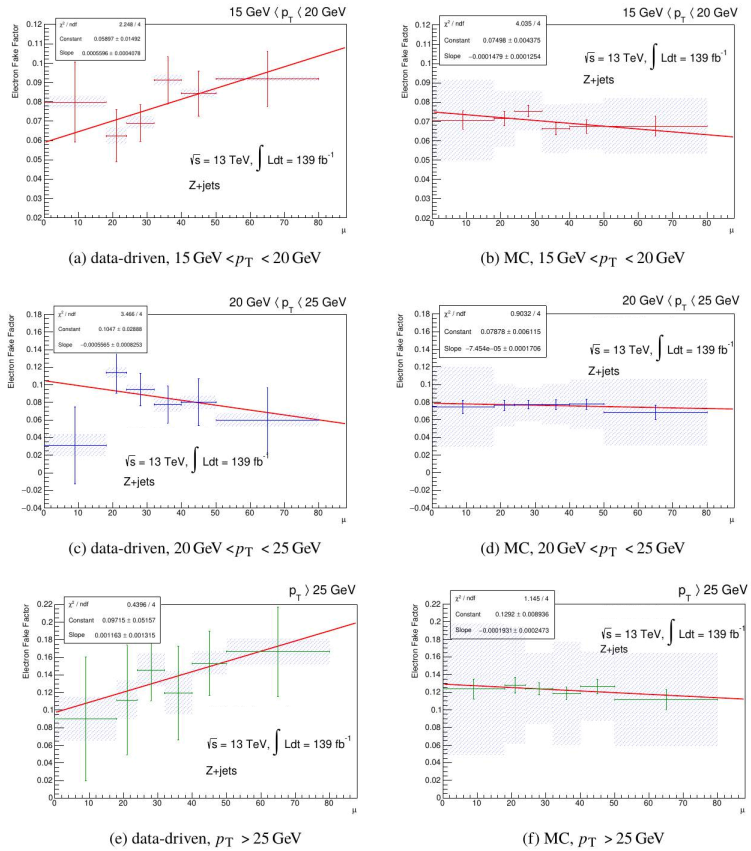
data taken in 2015 and 2016, the pileup (μ) averaged over all bunches in data is used resulting in a good data/MC agreement. For the 2017 and 2018 data, the actual μ for each bunch in data is used in order to reduce the statistical dilution of the samples. For 2017 period, 10% excess in data is observed and at high μ values there is a deficit in data reaching 30%.

The fake factors estimated both with the data-driven technique and using solely Z+jets MC Powheg events, are shown along with their uncertainties in Figures 12.7 and 12.8 for electrons and muons respectively. The MC FF estimation was performed to assure the validity of the data-driven method. The fake factors as functions of pileup are fit with a polynomial function to assess any possible pileup dependence. Given the statistical errors, the slopes derived from such fits are compatible with zero. For the low p_T electrons and muons the fake factors are stable across pile-up, while for the higher p_T the fake factors are dominated by the statistical uncertainties. Therefore the results do not support any strong evidence of pileup dependence so as to introduce a systematic error in the overall fake factor uncertainty.

12.3.1.1 jet dependence

The final state of the VBF signal constitutes of two leptons and two jets, while the fake factors as described in the previous section are measured in a three-lepton event selection without any requirements on jets. The lepton which does not satisfy the requirements for a Z decay candidate is considered as the "fake". The aim of the study is to investigate whether any detector activity is near the fake lepton. Therefore, an additional requirement has been inserted, i.e at least one jet with $p_T > 30$ GeV and $|\eta| < 4.5$, with the fake factors being estimated as functions of the number of jets in the event and the distance between the "fake" lepton and its closest jet.

With the jets insert in the event selection, significant discrepancies on MC and data have been observed on the kinematic variables. To overcome the MC mis-



(a) data-driven, 15 GeV $< p_{\mathrm{T}} <$ 20 GeV

(b) MC, 15 GeV $< p_{\mathrm{T}} <$ 20 GeV

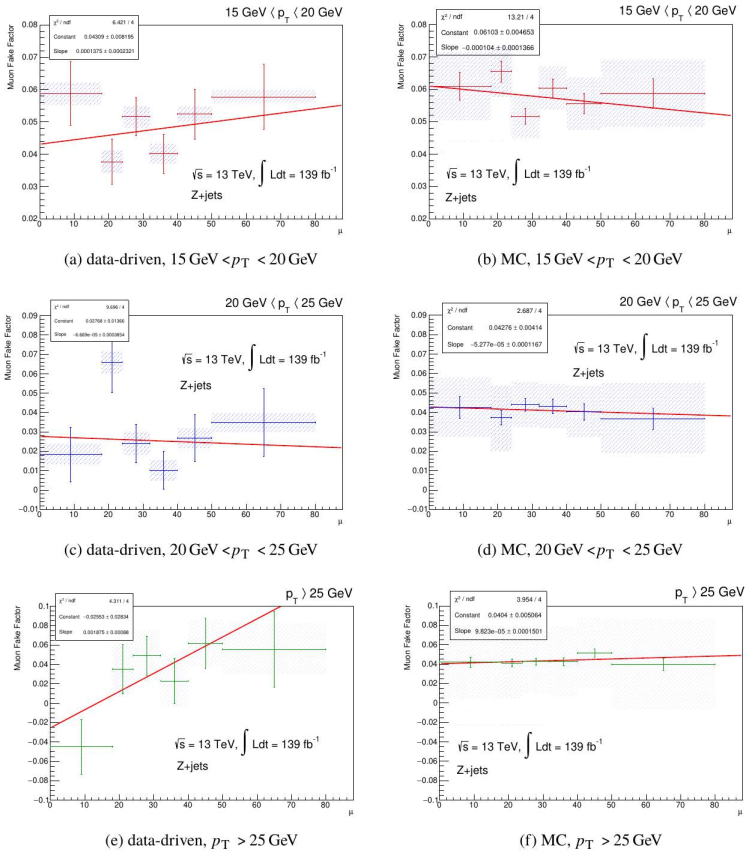
(c) data-driven, 20 GeV $< p_{\mathrm{T}} <$ 25 GeV

(d) MC, 20 GeV $< p_{\mathrm{T}} <$ 25 GeV

(e) data-driven, $p_{\mathrm{T}} >$ 25 GeV

(f) MC, $p_{\mathrm{T}} >$ 25 GeV

Figure 12.7. Electron fake factors in data (left) and MC (right) as functions of the pile-up for each p_T bin, while integrated in η. Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainies in the data-driven metod, the error bands represent the total uncertainties of the data-driven technique. A polynomial fit is performed and the fit parameters are displayed in the insets to assess any possible pileup dependence.



(a) data-driven, 15 GeV $< p_{\mathrm{T}} <$ 20 GeV

(b) MC, 15 GeV $< p_{\mathrm{T}} <$ 20 GeV

(c) data-driven, 20 GeV $< p_{\mathrm{T}} <$ 25 GeV

(d) MC, 20 GeV $< p_{\mathrm{T}} <$ 25 GeV

(e) data-driven, $p_{\mathrm{T}} >$ 25 GeV

(f) MC, $p_{\mathrm{T}} >$ 25 GeV

Figure 12.8. Muon fake Factors in data (left) and MC (right) as functions of the pile-up for each p_T bin, while integrated in η. Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainies in the data-driven metod, the error bands represent the total uncertainties of the data-driven technique. A polynomial fit is performed and the fit parameters are displayed in the insets to assess any possible pileup dependence.

modeling, the Z+jets Powheg events got reweighted to data. The weights (Figure 12.9) are measured at the level of the "id" leptons on the jets distributions and then are applied to the Z+jets MC events, event-by-event. Figure 12.10 show the p_T and $|\eta|$ of the three leptons and the leading jet before and after the MC reweighting. A substantial improvement on the ratios is visible after applying the weights. Such a MC-reweighting procedure only affects the MC modeling of jet production and the fake factor estimation in MC, but does not affect the data-driven estimate of the fake factors. The full cutflow of this study can be found in Table 12.5.

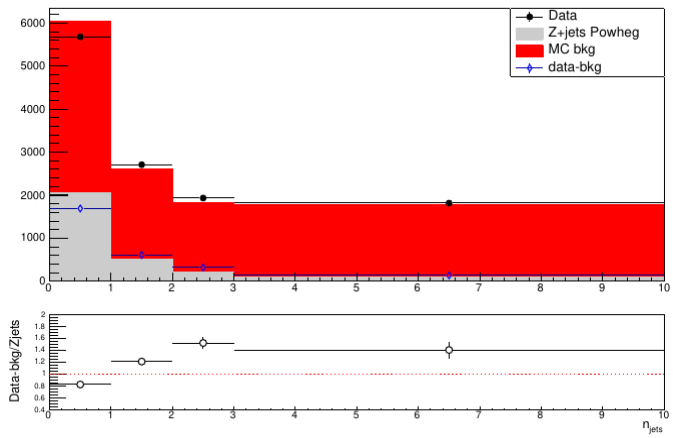


Figure 12.9. Number of jets in selected Z+jets events in Monte Carlo and in data. Their ratios are used to calculate the weights that correct the MC distributions for jet production mismodeling.

12.3.2 jet dependence

The fake factors as a function of the number of jets in the event are shown in Figures 12.11 and 12.12 for electrons and muons respectively. A polynomial fit up to two jets is used as this event configuration is most representantive of the event selection in the signal region. The fake factors as a function of the distance between the "fake" lepton and the leading jet are shown in Figures 12.13 and 12.14 for electrons and muons respectively.

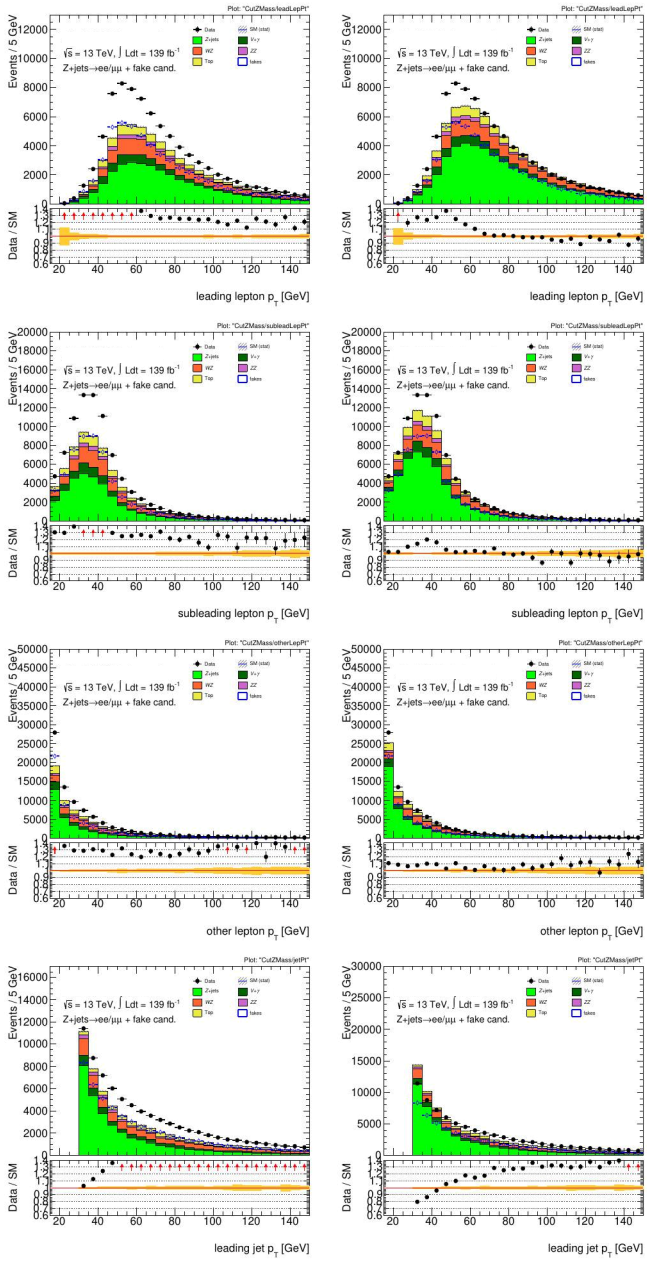


Figure 12.10. Leading lepton, subleading lepton, and leading jet p_T distributions after the Z-tagging cut (where the fake lepton is selected). Left:before the MC-reweighting, Right: after the MC-reweighting

$\sqrt{s} = 13 TeV$, $\mathcal{L} = 139 fb^{-1}$ (Full Run 2)	Z+jets Powheg	$V+\gamma$	WW	WZ	ZZ	Top	VVV	Total Bkg	Data	fakes	ZFake purity(%)
Channel Selection	243999.56 ± 265.69	103576.26 ± 733.79	990.92 ± 8.35	48878.83 ± 1712.65	14773.62 ± 113.44	85411.82 ± 61.84	238.23 ± 0.93	253869.68 ± 1867.72	605458	351588.32 ± 2023.33	58.07 ± 0.34
lll final state	243999.56 ± 265.69	103576.26 ± 733.79	990.92 ± 8.35	48878.83 ± 1712.65	14773.62 ± 113.44	85411.82 ± 61.84	238.23 ± 0.93	253869.68 ± 1867.72	605458	351588.32 ± 2023.33	58.07 ± 0.34
All lep $p_T > 15$ GeV	243999.56 ± 265.69	103576.26 ± 733.79	990.92 ± 8.35	48878.83 ± 1712.65	14773.62 ± 113.44	85411.82 ± 61.84	238.23 ± 0.93	253869.68 ± 1867.72	605458	351588.32 ± 2023.33	58.07 ± 0.34
Overlap: Vgamma/Vjets	243999.56 ± 265.69	103576.26 ± 733.79	990.92 ± 8.35	48878.83 ± 1712.65	14773.62 ± 113.44	85411.82 ± 61.84	238.23 ± 0.93	253869.68 ± 1867.72	605458	351588.32 ± 2023.33	58.07 ± 0.34
lead jet $p_T > 30$ GeV $\|eta\| < 4.5$	105991.70 ± 219.45	38089.89 ± 504.96	594.95 ± 4.90	28862.95 ± 1709.47	7175.34 ± 99.02	77788.18 ± 57.99	152.32 ± 0.73	152663.64 ± 1786.19	325250	172586.36 ± 1875.02	53.06 ± 0.58
Z-tagging	48680.68 ± 145.27	6553.90 ± 200.36	59.30 ± 1.32	14600.21 ± 24.90	2704.82 ± 15.18	7039.57 ± 16.35	48.86 ± 0.32	31006.66 ± 203.14	86652	55645.34 ± 357.65	64.22 ± 0.47
TransMassW	40677.76 ± 131.70	5114.79 ± 182.77	20.08 ± 0.77	4760.24 ± 13.06	1858.94 ± 12.31	2774.15 ± 10.36	12.76 ± 0.16	14540.97 ± 183.94	61191	46650.03 ± 308.26	76.24 ± 0.59
all three leptons ID or AntiID	20235.65 ± 94.63	1499.90 ± 100.22	9.18 ± 0.51	4568.62 ± 12.85	1758.11 ± 12.09	2385.67 ± 9.55	12.24 ± 0.16	10233.71 ± 102.21	33313	23079.29 ± 209.19	69.28 ± 0.73
fake type: electron	10738.38 ± 66.90	1186.98 ± 86.37	5.40 ± 0.40	2087.68 ± 9.46	964.43 ± 6.35	535.13 ± 4.15	5.73 ± 0.10	4785.36 ± 87.22	15351	10565.64 ± 151.52	68.83 ± 1.13
fake el η cut	8334.75 ± 59.10	1034.45 ± 81.27	4.12 ± 0.37	1972.51 ± 9.30	908.59 ± 6.22	493.85 ± 3.98	5.41 ± 0.10	4418.93 ± 82.14	12641	8222.07 ± 139.24	65.04 ± 1.24
Scale factors				NF = 0.94 ± 0.01				NF = 0.94 ± 0.01		NF = 0.94 ± 0.01	
ID cuts	766.89 ± 17.85	246.03 ± 44.92	0.20 ± 0.12	1537.15 ± 8.11	680.70 ± 4.07	163.46 ± 1.60	4.54 ± 0.09	2632.07 ± 45.85	3464	831.93 ± 74.61	24.02 ± 2.19
Scale factors				NF = 0.94 ± 0.01				NF = 0.94 ± 0.01		NF = 0.94 ± 0.01	
Anti-ID cuts	7567.86 ± 56.34	788.43 ± 67.73	3.93 ± 0.36	322.93 ± 3.35	227.89 ± 4.71	330.39 ± 3.65	0.88 ± 0.04	1674.44 ± 68.08	9177	7502.56 ± 117.52	81.75 ± 1.54
fake type: muon	9497.27 ± 66.93	312.92 ± 50.83	3.78 ± 0.32	2480.93 ± 8.69	793.67 ± 10.29	1850.54 ± 8.60	6.51 ± 0.12	5448.36 ± 53.28	17962	12513.64 ± 144.23	69.67 ± 0.96
fake mu η cut	9127.57 ± 65.53	292.45 ± 50.41	3.61 ± 0.31	2427.83 ± 8.55	776.73 ± 10.21	1793.33 ± 8.44	6.39 ± 0.11	5300.33 ± 52.82	17303	12002.67 ± 141.75	69.37 ± 0.97
Scale factors				NF = 0.94 ± 0.01				NF = 0.94 ± 0.01		NF = 0.94 ± 0.01	
ID cuts	421.67 ± 13.96	1.15 ± 1.15	0.26 ± 0.09	1890.12 ± 7.34	487.75 ± 6.90	205.08 ± 1.89	5.31 ± 0.10	2589.66 ± 10.31	3001	411.34 ± 55.74	13.71 ± 1.87
Scale factors				NF = 0.94 ± 0.01				NF = 0.94 ± 0.01		NF = 0.94 ± 0.01	
Anti-ID cuts	8705.90 ± 64.03	291.30 ± 50.40	3.34 ± 0.30	399.34 ± 3.33	288.98 ± 7.53	1588.26 ± 8.23	1.08 ± 0.05	2572.30 ± 51.73	14302	11729.70 ± 130.30	82.01 ± 1.14

Table 12.5. Full cutflow of the Z+jets fake factor analysis with the additional requirement of at least one jet in the final state. The normalization factor that is derived in the WZ control region is applied at some cut stages and affects the WZ yield as well as the total background and fake yield.

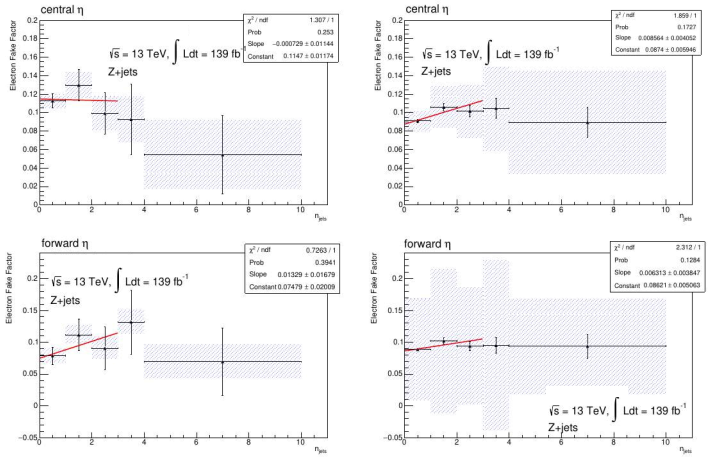


Figure 12.11. Muon fake factors as a function of the number of jets estimated with the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique.

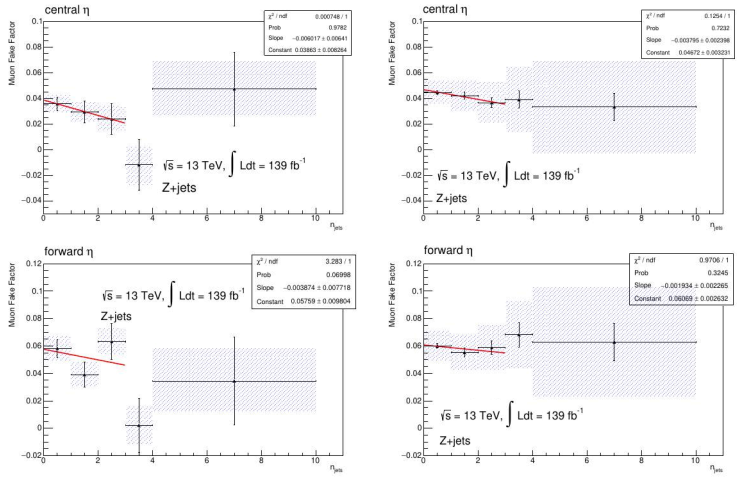


Figure 12.12. Electron fake factors as a function of the number of jets estimated with the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique.

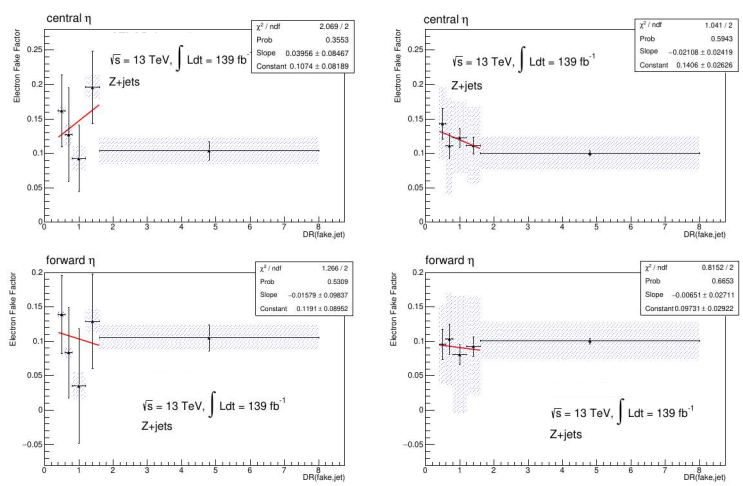


Figure 12.13. Electron fake factors as a function of the distance between the "fake" lepton and its closest jet, in the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique.

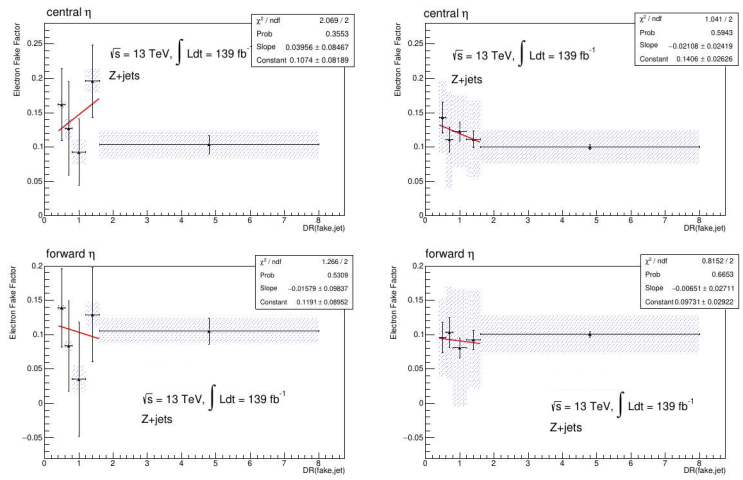


Figure 12.14. Muon fake factors as a function of the distance between the "fake" lepton and its closest jet, in the data-driven method (left) and in Monte Carlo (right). Statistical uncertainties are shown as error bars in all plots. The EW subtraction uncertainties are shown as error bands in the data-driven method. In the case of the MC estimation, to facilitate the comparison between the MC-based estimate and the uncertainties in the data-driven method, the error bands represent the total uncertainties of the data-driven technique.

Most relevant to see if there is any overlap between the 'fake' lepton and a near-by jet that may case a dependence on the fake factor estimation is selecting jets close to the 'fake' lepton. Therefore, the polynomial fit is restricted to low ΔR values. The fit results including all bins in all fits compared to the restricted selected regions are presented in Table 12.6. The slopes, in both data-driven and MC methods, are compatible with zero with acceptable fit quality. From the results, there is no evidence of fake factor dependence on the presence of jets in the events or on the presence of a jet nearby the fake lepton candidate.

12.3.3 Fake Factor Corrections

Trigger bias The fake factors in the di-jets enriched region complements those performed in the Z+jets control region. To maximize efficiency, both dilepton and single-lepton triggers are used in the analysis. The single-lepton triggers have tighter requirements than the "anti-id" definition. When an event is collected because an

			Data-driven		MC	
		Fitting range	Slope ($\times 10^{-2}$)	χ^2/ndf	Slope ($\times 10^{-2}$)	χ^2/ndf
njets	Electron (central η)	full	-0.755 ± 0.610	0.564	0.606 ± 0.265	1.932
		partial	-0.073 ± 1.144	1.307	0.856 ± 0.405	1.859
	Electron (forward η)	full	0.321 ± 0.859	0.703	0.342 ± 0.241	1.340
		partial	1.329 ± 1.679	0.726	0.631 ± 0.385	2.312
	Muon (central η)	full	-0.734 ± 0.742	1.448	-0.257 ± 0.138	0.261
		partial	-0.602 ± 0.641	0.001	-0.380 ± 0.240	0.125
	Muon (forward η)	full	-0.962 ± 0.672	2.149	0.008 ± 0.155	0.935
		partial	-0.387 ± 0.772	3.283	-0.193 ± 0.226	0.971
ΔR(fake,jet)	Electron (central η)	full	6.440 ± 6.458	1.366	1.159 ± 0.965	1.866
		partial	3.956 ± 8.467	1.034	-2.108 ± 2.419	0.521
	Electron (forward η)	full	-0.038 ± 0.973	0.430	-1.607 ± 1.991	0.918
		partial	-1.579 ± 9.837	0.633	-0.651 ± 2.711	0.408
	Muon (central η)	full	-0.148 ± 0.472	0.236	-0.020 ± 0.135	0.066
		partial	-1.785 ± 4.452	0.285	-0.384 ± 1.399	0.065
	Muon (forward η)	full	-0.843 ± 0.838	0.342	0.031 ± 0.177	0.877
		partial	1.354 ± 6.096	0.449	1.323 ± 1.765	1.025

Table 12.6. Summary of fit results for fake factors as a function of number of jets or distance between the fake lepton and near by jet.

"anti-id" lepton fired a single-lepton trigger, the extrapolation to the signal region will be biased if a Z+jets fake factor is applied. To avoid this, a dedicated set of fake factors is derived by requiring the "anti-id" lepton to pass a single-lepton trigger. However Z+jets control region does not yield sufficient statistics for a reliable estimation, and instead they are determined in a di-jets enriched region. The proportion of the fake lepton candidates which fired the single-lepton triggers is shown in Table 12.7.

	$e\mu$	μe
e-Fakes	4.70 / 83.18	6.23 / 277.89
μ-Fakes	0.35 / 878.56	34.40 / 221.48

Table 12.7. Events in the W+jets control region for which only the fake candidate fired a single-lepton trigger out of total events.

Events which have exactly one fake lepton candidate with $p_T > 15$ GeV and at least one jet with $p_T > 25$ GeV are selected while their azimuthal separation should meet the requirement of $\Delta\phi_{lep,lead-jet} < 2.5$ in order to enforce the lepton-jet topology (the kinematic phase space of the dijet events are leptons recoiling away from a jet). To suppress the W+jets processes, events are required to have $E_T^{Miss} < 30GeV$ whereas in the lepton- E_T^{Miss} system the requirement on the transverse mass of $m_T < 60$ GeV is also inserted. The events that pass these selection criteria are finally split into the usual "id/anti-id" cuts. The fake lepton event yield is

determined by subtracting from the total data-observed yield the MC-predicted yield of all other processes (ie, W+jets, top, Z+jets, single photon) that produce leptons and pass the event selection criteria. The EW subtraction uncertainty is propagated to the fake lepton estimate. Figure 12.15 shows the "triggered" fake factors, calculated in the dijet-enriched sample in data, for each data-taking period separately. These fake factors are then used in the main analysis only when the fake lepton fires a single-lepton trigger. When a dilepton trigger is fired, the nominal fake factors, as estimated in Z+jets sample, are used.

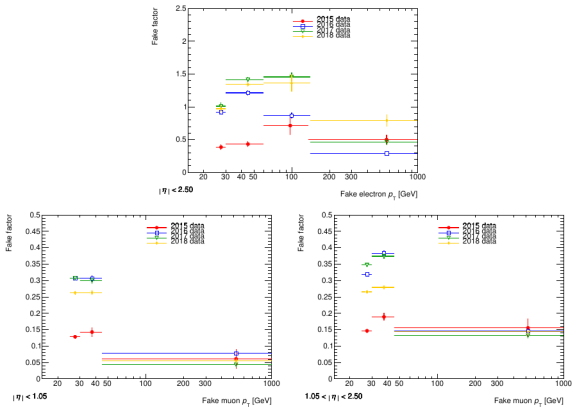
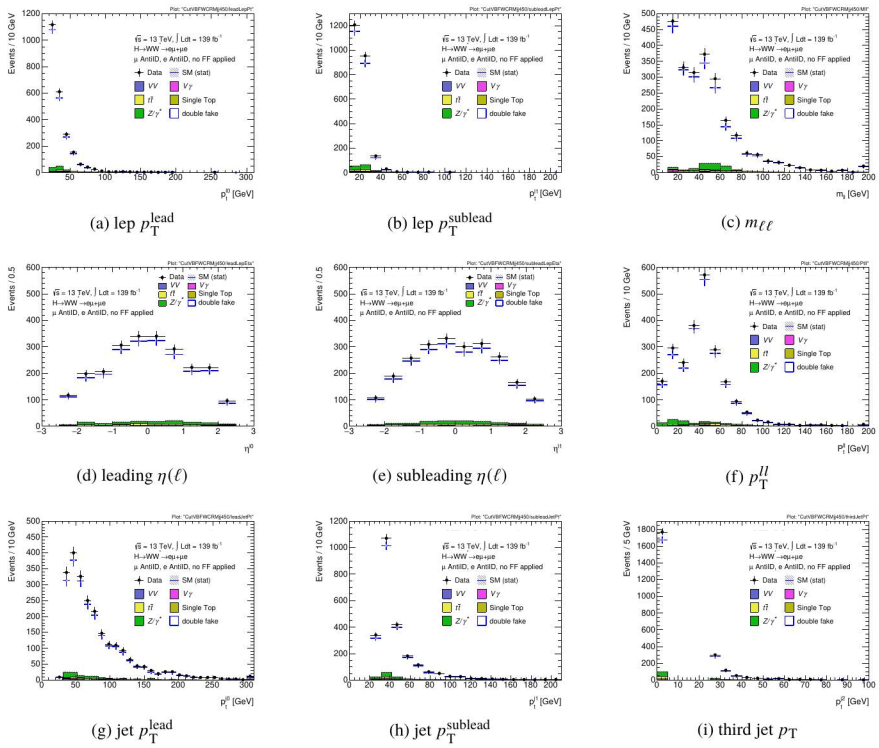


Figure 12.15. Electron (top) and Muon (bottom) fake factors with trigger bias. The effects of the trigger biases are expected to be different in each campaign due to differences in the trigger configurations and updates in the online reconstruction and calibration of the leptons, thus they are estimated separately for each year. Only statistical uncertainties are shown.

Double Fakes Correction: Contributions to the fake lepton background in the signal region not only originate from events with a fake electron or events with a fake muon, but also from events where both leptons are fakes. The latter set of events is dominated by QCD processes that give rise to events with two jets, so-called "double fakes". To evaluate this correction, a double-fake control region is used, where both leptons in the event are required to pass "anti-id" criteria only, while MC processes with two real leptons are subtracted. Fake factors corresponding to two "anti-id"

leptons in the selection are then applied. Kinematic distributions in this control region before fake factor application are shown in Figures 12.16 and 12.17 in which we can see the high purity of the "double fake" control region. Table 12.8 shows the estimated yields of double-fake contribution to the estimation of fakes from the W+jets process in the VBF SR, and the corresponding corrections on the fakes yields, which are found to be approximately 25%.



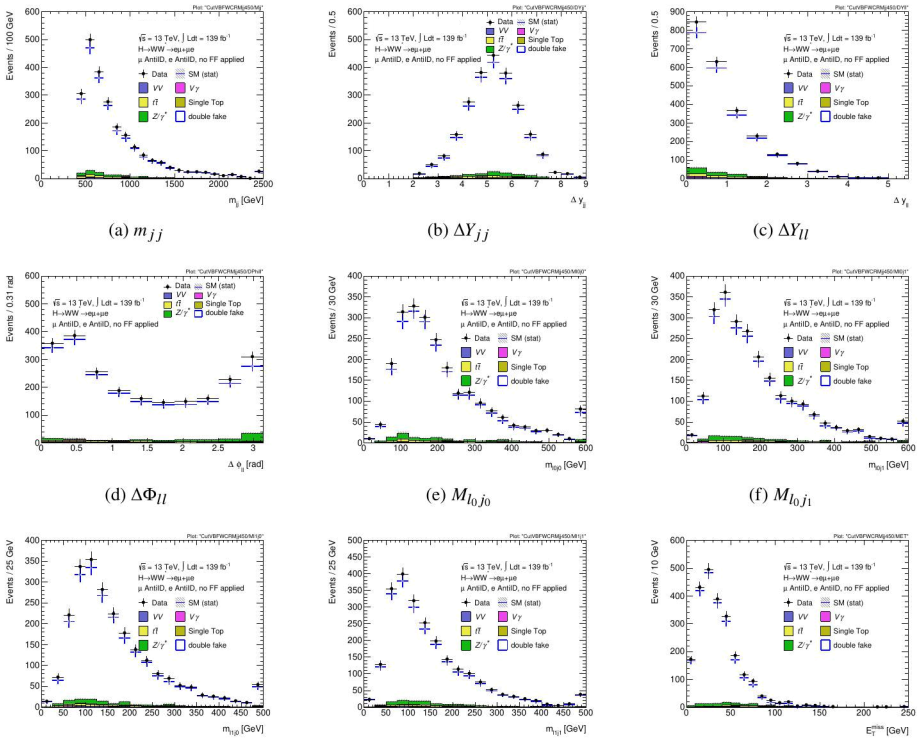
(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) $m_{\ell\ell}$

(d) leading $\eta(\ell)$ (e) subleading $\eta(\ell)$ (f) p_{T}^{ll}

(g) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (h) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$ (i) third jet p_{T}

Figure 12.16. Kinematic distributions in the double-fakes control region. Data points (black markers) describe events with two "anti-id" leptons. Distributions correspond to MC events with two real leptons passing the "anti-id" selection. Their difference (blue markers) represent the estimate of the double fakes. Fake factors are not applied in these plots.

The number of events with two fake candidates where the di-lepton trigger did not fire is given in Table 12.8



(a) m_{jj} (b) ΔY_{jj} (c) ΔY_{ll}

(d) $\Delta\Phi_{ll}$ (e) $M_{l_0 j_0}$ (f) $M_{l_0 j_1}$

(g) $M_{l_1 j_0}$ (h) $M_{l_1 j_1}$ (i) $E_{\mathrm{T}}^{\mathrm{miss}}$

Figure 12.17. Kinematic distributions in the double-fakes control region. Data points (black markers) describe events with two "anti-id" leptons. Distributions correspond to MC events with two real leptons passing the "anti-id" selection. Their difference (blue markers) represent the estimate of the double fakes. Fake factors are not applied in these plots.

channel	Dilepton Trigger Did not Fire / Total
$e\mu$	62.63 / 1278.32
μe	312.63 / 919.49
$e\mu\mu e$	375.26 / 2197.81

Table 12.8. Events in the double-fakes control region in which both leptons must be "anti-ID" for which the dilepton trigger did not fire.

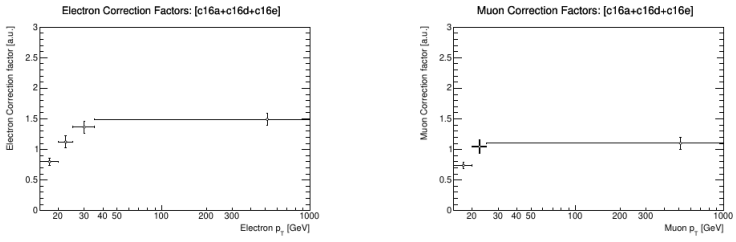
Correction Factors: The data-driven fake factor method described so far assumes that the fake factor F derived in a three-lepton Z+fake selection is applicable to a two-lepton W+fake selection. Differences in event kinematics and production mechanism between W+jets and Z+jets processes, and differences in the selections in their respective event samples lead to differences in characteristics of the fake lepton kinematics, as well as of the remainder of the event. Although it is reasonable to assume that the remainder of the event has a negligible impact on the fake lepton efficiency, the characteristics of the fake lepton may cause important differences in the fake contribution, so the fake factor must be corrected for potential sample dependences. A correction can be made by comparing fake factors measured in Z+jets and W+jets Monte Carlo. To derive these fake factors, the same selections are applied to the MC samples as are used in data. Additionally, one (two) leptons in the event are truth-matched to the W (Z) boson, and the other lepton is identified as the fake. The ratio of these fake factors gives a correction factor or CF,

$$CF = \frac{F^W_{MC}}{F^Z_{MC}} \tag{12.7}$$

The correction factor is then multiplied by the fake factor derived in the three-lepton Z+jets selection in data to provide a fake factor that can be applied to a W+jets sample in data,

$$F^W_{data} = CF \cdot F^Z_{data} \tag{12.8}$$

Because the correction factors produced by this method are calculated using Monte Carlo, they are generator- dependent. This is evaluated as a source of systematic uncertainty by comparing the correction factors calculated with different generators. The nominal correction factors are evaluated in samples generated with Powheg because it was found to have the best statistical precision. The systematic uncertainty is taken as the difference between the central values calculated with Powheg and MadGraph. The electron correction factors use the same p_T binning as the fake factors while the muon correction factors combine the two highest p_T bins into one bin above 25 GeV for better statistics and because in the two highest bins the correction factors were found to be consistent. The dependence on $|\eta|$ was investigated and was found negligible so the correction factors are not binned in $|\eta|$. The correction factors are shown in Figure 12.18. The central values and the statistical and systematic uncertainties on them are summarized in table 12.9 for electrons and muons respectively.



(a) Electron CF (b) Muon CF

Figure 12.18. Correction factors derived from MC sample with the Powheg generator, electrons (left) muons (right)

12.4 Fakes estimation in the W+jets control region

The W+jets control region and the cutflow is defined in the previous section. Distributions in Figs. 12.19, 12.20 show the electroweak background processes in the W+jets control region along with data. The W+jets background, shown in the plots,

Electron Correction Factor			
15 – 20 GeV	20 – 25 GeV	25 – 35 GeV	35 – 1000 GeV
0.80 ± 0.06(stat) ± 0.03(syst)	1.12 ± 0.10(stat) ± 0.33(syst)	1.36 ± 0.10(stat) ± 0.01(syst)	1.49 ± 0.10(stat) ± 0.30(syst)

Muon Correction Factor		
15 – 20 GeV	20 – 25 GeV	25 – 1000 GeV
0.74 ± 0.05(stat) ± 0.02(syst)	1.04 ± 0.11(stat) ± 0.01(syst)	1.10 ± 0.10(stat) ± 0.12(syst)

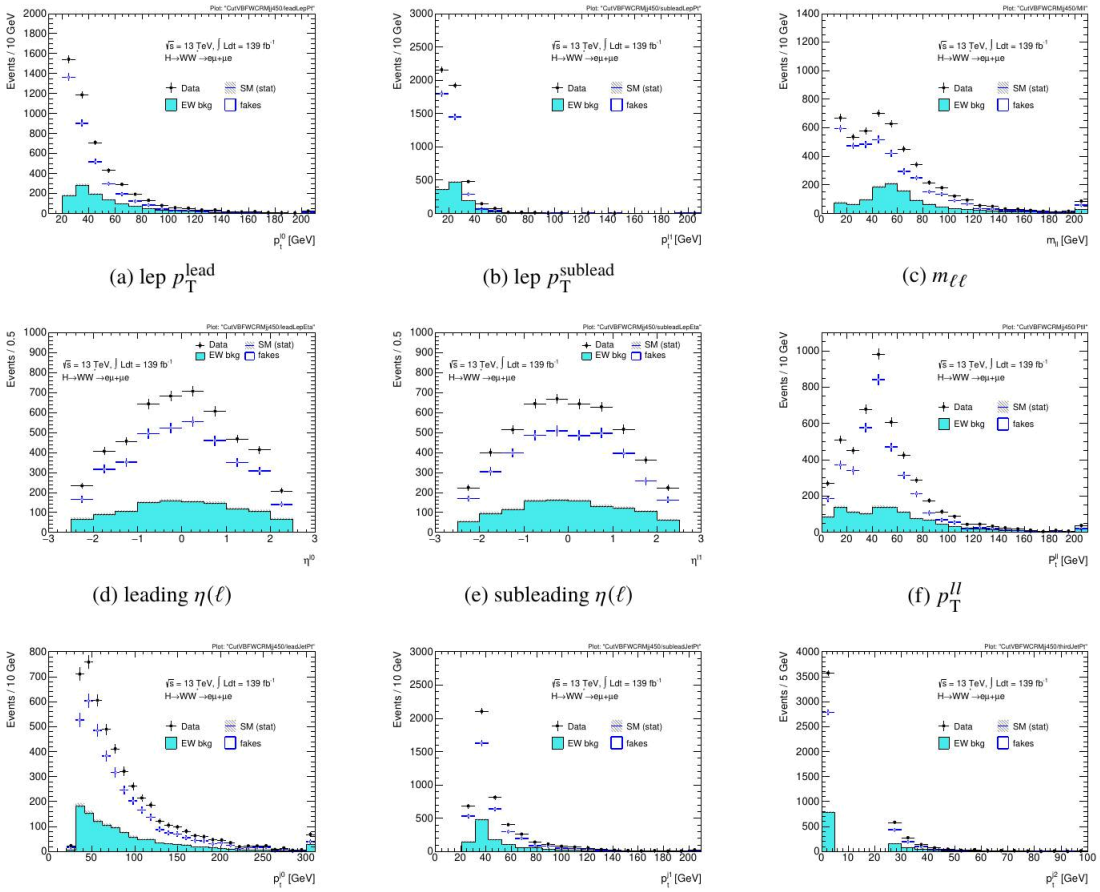
Table 12.9. Muon and Electron CF in three bins of p_T . The nominal CF is derived using Powheg and the systematic uncertainty is obtained by comparing with MadGraph

is taken as the difference between data and EW background. The EW background subtracted from data in the W+jets control region consists of $V\gamma$, diboson, top, and Z+jets events. The EW substracted data sample is then used to estimate fakes contaminating the signal region, after fake factors are applied. Generally, a high fake purity of W+jets events is observed in the W+jets CR.

12.4.1 Closure Test

Closure tests for the fake estimation method were performed in order to assess any potential bias (i.e. non-closure) of the data-driven fake factor method as reported in the previous sub section. These studies were performed in a previous phase space which is differentiated from the absence of cuts on $M_{jj} > 450$ GeV and $\Delta\phi_{ll} < 1.4$.

The closure tests were performed on Monte Carlo simulations using Powheg and MadGraph W+jets samples. Closure is tested by comparing W+jets MC events in the signal region to W+jets events in the W control region, which mimics the signal region except for a cut selecting 1 'id' and 1 'anti-id' lepton, with fake factors applied, as calculated in the Z+jets control region. Since the fake factors are meant to directly translate 'id'/'anti-id' events to 'id'/'id' these numbers should match within statistical variations. The fake factors were derived from the same Monte Carlo generator (using a Z+jets sample) as the one used for the W+jets process. The correction factors (CF) were derived for each Monte Carlo generator independently by taking the ratio of the W+jets and Z+jets fake factors. These correction factors were then applied to the

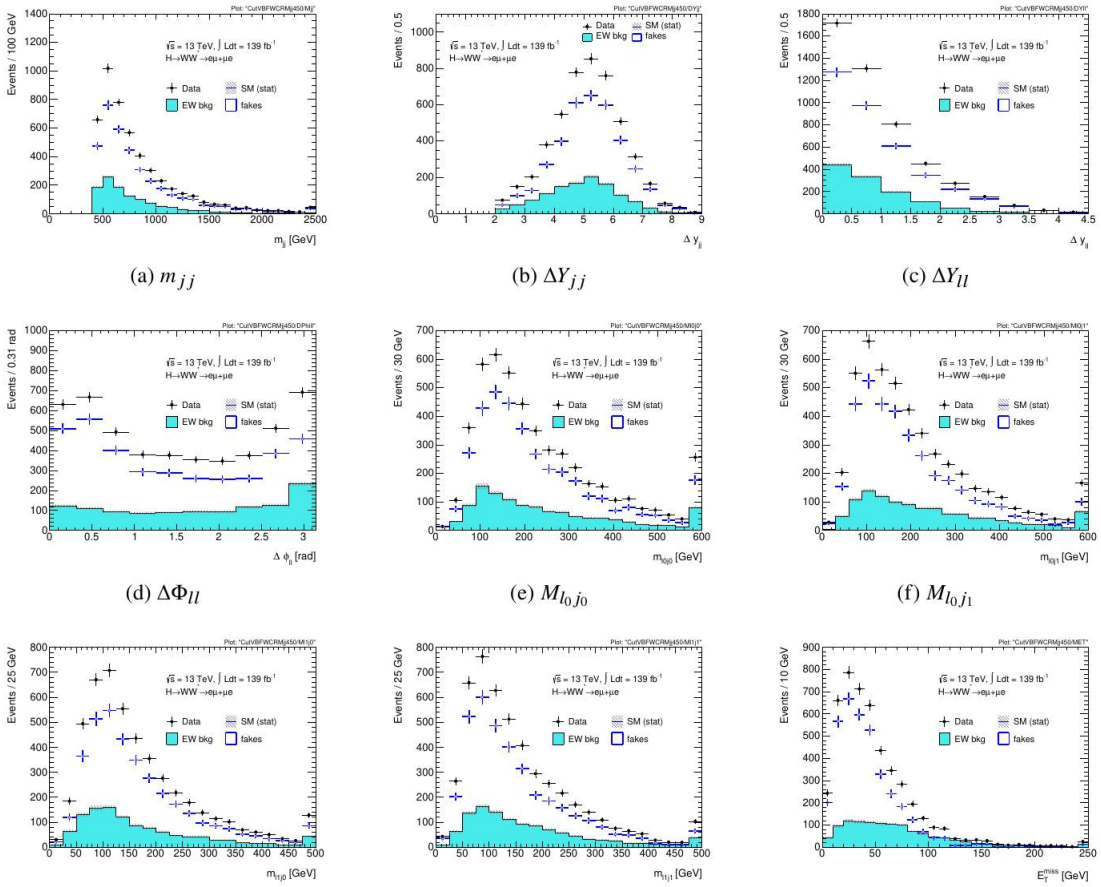


(a) lep $p_{\mathrm{T}}^{\mathrm{lead}}$ (b) lep $p_{\mathrm{T}}^{\mathrm{sublead}}$ (c) $m_{\ell\ell}$

(d) leading $\eta(\ell)$ (e) subleading $\eta(\ell)$ (f) p_{T}^{ll}

(g) jet $p_{\mathrm{T}}^{\mathrm{lead}}$ (h) jet $p_{\mathrm{T}}^{\mathrm{sublead}}$ (i) third jet p_{T}

Figure 12.19. Distributions in the fakes control region. The data (black markers) are shown together with electroweak processes in MC (blue histograms). Their difference (blue markers) represents the estimate of the fake background. No fake factors are applied.



(a) m_{jj} (b) ΔY_{jj} (c) ΔY_{ll}

(d) $\Delta\Phi_{ll}$ (e) $M_{l_0 j_0}$ (f) $M_{l_0 j_1}$

(g) $M_{l_1 j_0}$ (h) $M_{l_1 j_1}$ (i) $E_{\mathrm{T}}^{\mathrm{miss}}$

Figure 12.20. Distributions in the fakes control region. The data (black markers) are shown together with electroweak processes in MC (blue histograms). Their difference (blue markers) represents the estimate of the fake background. No fake factors are applied.

Z+jets fake factor to get finally a fake factor that can be applied to the W+jets control region.

In the signal region selection using Powheg W+jets MC events the total estimated fakes amount to about 67 events, while in the W+jets control region with the fake factors and correction factors applied about 65 fake events are estimated (Table 12.10). The closure is therefore within 3%, well within the total statistical uncertainties.

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Wjets Powheg
Channel Selection	429029.54 ± 2088.97
Trigger Selection	429029.54 ± 2088.97
Trigger Matching	414966.14 ± 2024.45
Jet Cleaning	414966.14 ± 2024.45
Overlap: Vgamma/Vjets	414966.14 ± 2024.45
Only two Leptons	414915.22 ± 2024.34
p_t^{lead} > 22 GeV	414915.22 ± 2024.34
$p_t^{sublead}$ > 15 GeV	414783.01 ± 2023.96
OS Leptons	414736.98 ± 2023.76
$M_{\ell\ell}$ > 10 GeV	414718.18 ± 2023.72
SF: Z Veto	414718.18 ± 2023.72
Leptons ID	14757.01 ± 381.36
2-jet (30,30) fJVT	949.68 ± 96.08
b-veto	820.17 ± 88.07
blinding (2-jet)	820.17 ± 88.07
CJV (20GeV)	618.70 ± 76.42
OLV bool	121.78 ± 33.08
$Z \to \tau\tau$ veto	85.51 ± 27.96
Mjj>200	67.30 ± 24.86
DYjj>2.1	67.30 ± 24.86

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Wjets Powheg
Channel Selection	429029.54 ± 2088.97
Trigger Selection	429029.54 ± 2088.97
Trigger Matching	414966.14 ± 2024.45
Jet Cleaning	414966.14 ± 2024.45
Overlap: Vgamma/Vjets	414966.14 ± 2024.45
Only two Leptons	414915.22 ± 2024.34
p_t^{lead} > 22 GeV	414915.22 ± 2024.34
$p_t^{sublead}$ > 15 GeV	414783.01 ± 2023.96
OS Leptons	414736.98 ± 2023.76
$M_{\ell\ell}$ > 10 GeV	414718.18 ± 2023.72
SF: Z Veto	414718.18 ± 2023.72
SingleFakes 1 anti-ID,1 ID	194343.95 ± 1357.06
Apply fake factor	15288.07 ± 1039.00
p_T^{miss} > 20 GeV	13228.06 ± 904.38
2-jet (30,30) fJVT	868.63 ± 82.45
b-veto	757.13 ± 72.01
CJV (20GeV)	562.89 ± 57.07
OLV bool	140.20 ± 19.45
$Z \to \tau\tau$ veto	95.04 ± 13.45
Mjj>200	69.01 ± 11.27
DYjj>2.1	65.06 ± 10.83

Table 12.10. Cutflow for W+jets Powheg MC events in the signal region (left) and in W+jets CR (right). Uncertainties are statistical. The statistical uncertainties from the fake factors and correction factors are also accounted.

In the signal selection using MadGraph W+jets MC events the total estimated fakes amount to about 171, while in the W+jets control region with the fake factors and correction factors applied about 173 fake events are estimated (Table 12.11). The closure is therefore within 1.2 %, well within total statistical uncertainties.

The closure is also studied as a function of several distributions to study any possible shape biases from the fake factor method. The plots in Fig. 12.21 show comparisons between Powheg MC events in the signal region (black) and those derived in the W+jets control region with the fake factor method (red) after the full selection. In order to test potential biases on a larger dataset, the same distributions are plot after preselection cuts in Fig. 12.22. The results do not show any clear evidence of

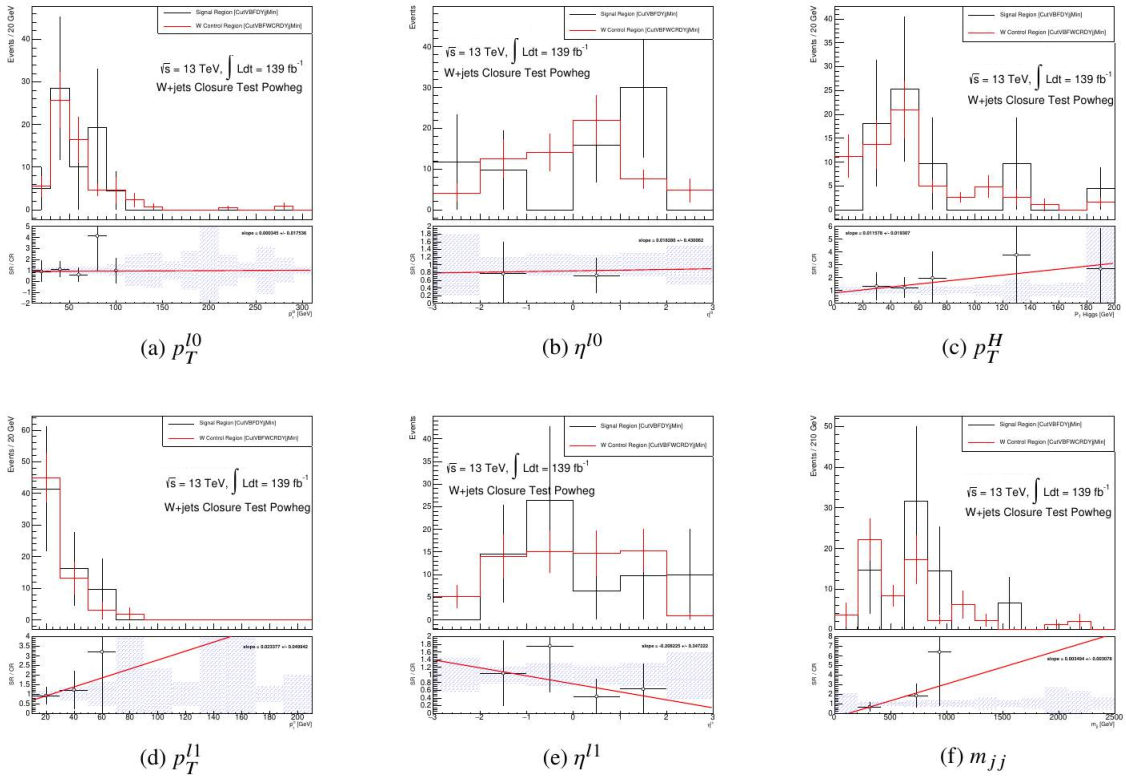
$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Wjets MadGraph
Channel Selection	576388.36 ± 2447.50
Trigger Selection	576388.36 ± 2447.50
Trigger Matching	558441.72 ± 2354.41
Jet Cleaning	558441.72 ± 2354.41
Overlap: Vgamma/Vjets	558441.72 ± 2354.41
Only two Leptons	558329.53 ± 2354.37
p_t^{lead} > 22 GeV	558329.53 ± 2354.37
$p_t^{sublead}$ > 15 GeV	557872.38 ± 2353.94
OS Leptons	557580.25 ± 2353.73
$M_{\ell\ell}$ > 10 GeV	557564.30 ± 2353.72
SF: Z Veto	557564.30 ± 2353.72
Leptons ID	22108.58 ± 529.18
2-jet (30,30) fJVT	3209.79 ± 95.94
b-veto	2675.45 ± 90.34
blinding (2-jet)	2675.45 ± 90.34
CJV (20GeV)	1939.87 ± 75.44
OLV bool	367.08 ± 36.99
$Z \to \tau\tau$ veto	229.30 ± 34.36
Mjj>200	186.59 ± 33.21
DYjj>2.1	170.81 ± 32.96

$\sqrt{s}$ = 13TeV, $\mathcal{L}$ = 139fb^{-1} (Full Run 2)	Wjets MadGraph
Channel Selection	576388.36 ± 2447.50
Trigger Selection	576388.36 ± 2447.50
Trigger Matching	558441.72 ± 2354.41
Jet Cleaning	558441.72 ± 2354.41
Overlap: Vgamma/Vjets	558441.72 ± 2354.41
Only two Leptons	558329.53 ± 2354.37
p_t^{lead} > 22 GeV	558329.53 ± 2354.37
$p_t^{sublead}$ > 15 GeV	557872.38 ± 2353.94
OS Leptons	557580.25 ± 2353.73
$M_{\ell\ell}$ > 10 GeV	557564.30 ± 2353.72
SF: Z Veto	557564.30 ± 2353.72
SingleFakes 1 anti-ID,1 ID	243273.90 ± 1460.28
Apply fake factor	22684.74 ± 3484.24
$p_{\mathrm{T}}^{\mathrm{miss}}$ > 20 GeV	19824.90 ± 3023.67
2-jet (30,30) fJVT	3297.59 ± 476.78
b-veto	2761.77 ± 403.51
CJV (20GeV)	2045.10 ± 298.37
OLV bool	394.47 ± 59.13
$Z \to \tau\tau$ veto	244.64 ± 37.90
Mjj>200	182.80 ± 29.23
DYjj>2.1	172.77 ± 27.89

Table 12.11. Cutflow for W+jets MadGraph MC events in the signal region (left) and in W+jets CR (right). Uncertainties are statistical. The statistical uncertainties from the fake factors and correction factors are also accounted.

statistically significant non-closure effects, and any possible non-closure effects are well within the uncertainties estimated in the data driven method. Therefore, the non-closure is not accounted for as a systematic contribution in the fake estimation method.

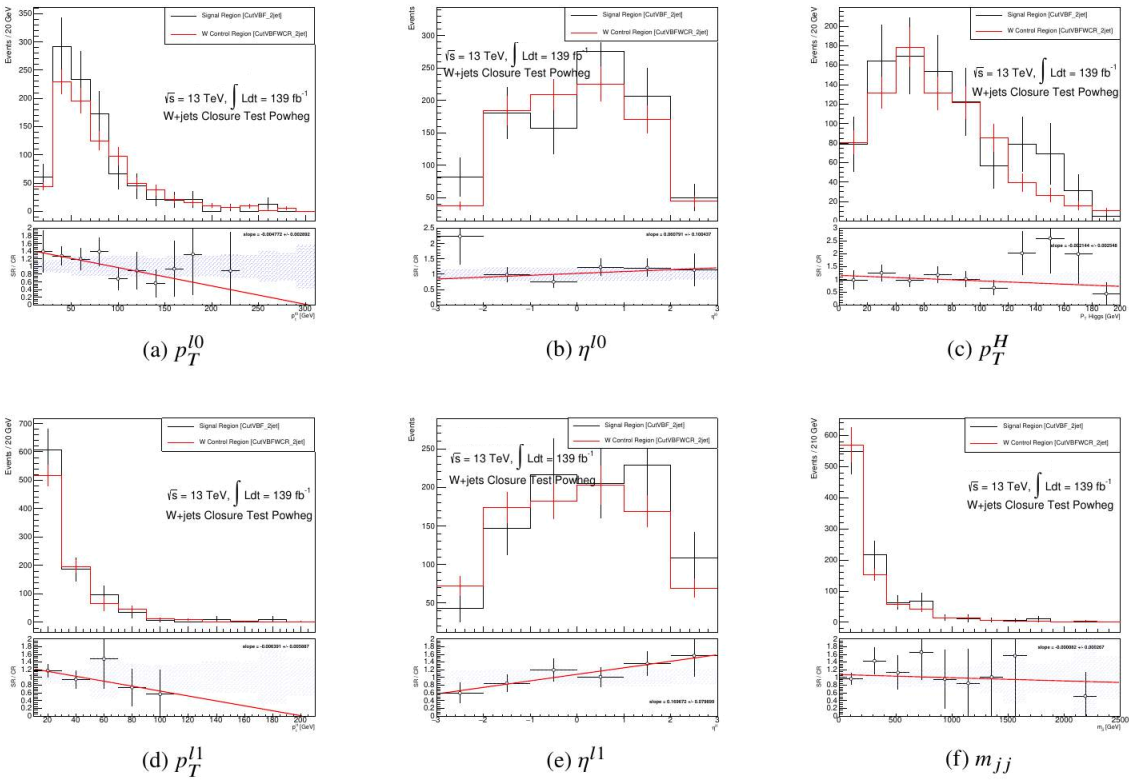
The closure test on distributions is also carried out with MadGraph MC samples which have larger statistics than Powheg. The plots in Fig. 12.23 show comparisons between MadGraph events estimated as fakes in the signal region (black) and those estimated in the W+jets control region with the fake factor method (red). In order to test potential biases with a larger dataset, the same distributions are plot after preselection cuts in Fig. 12.24. Also in this study with MadGraph MC samples, no significant non-closure effects are observed and not beyond the quoted data-driven fake uncertainties, therefore no systematic uncertainty is included to account non-closure bias.



(a) p_T^{l0} (b) η^{l0} (c) p_T^H

(d) p_T^{l1} (e) η^{l1} (f) m_{jj}

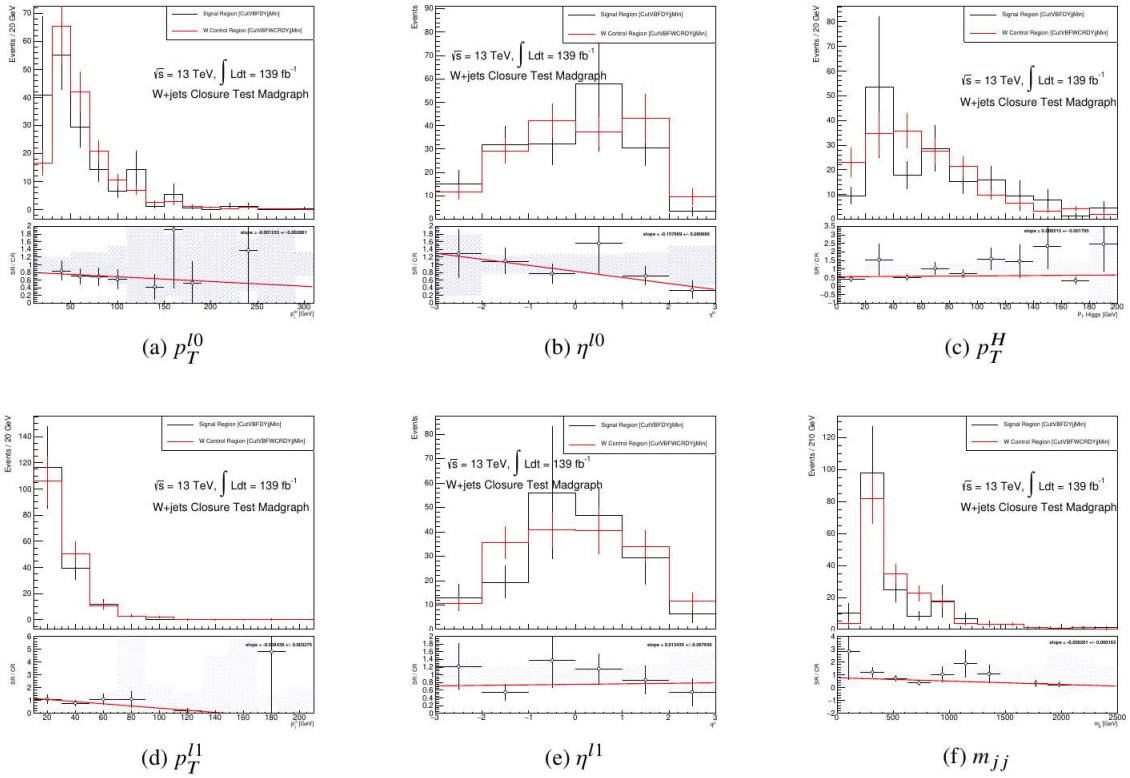
Figure 12.21. Distributions at preselection level for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs applied (red) using Powheg MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable.



(a) p_T^{l0} (b) η^{l0} (c) p_T^{H}

(d) p_T^{l1} (e) η^{l1} (f) m_{jj}

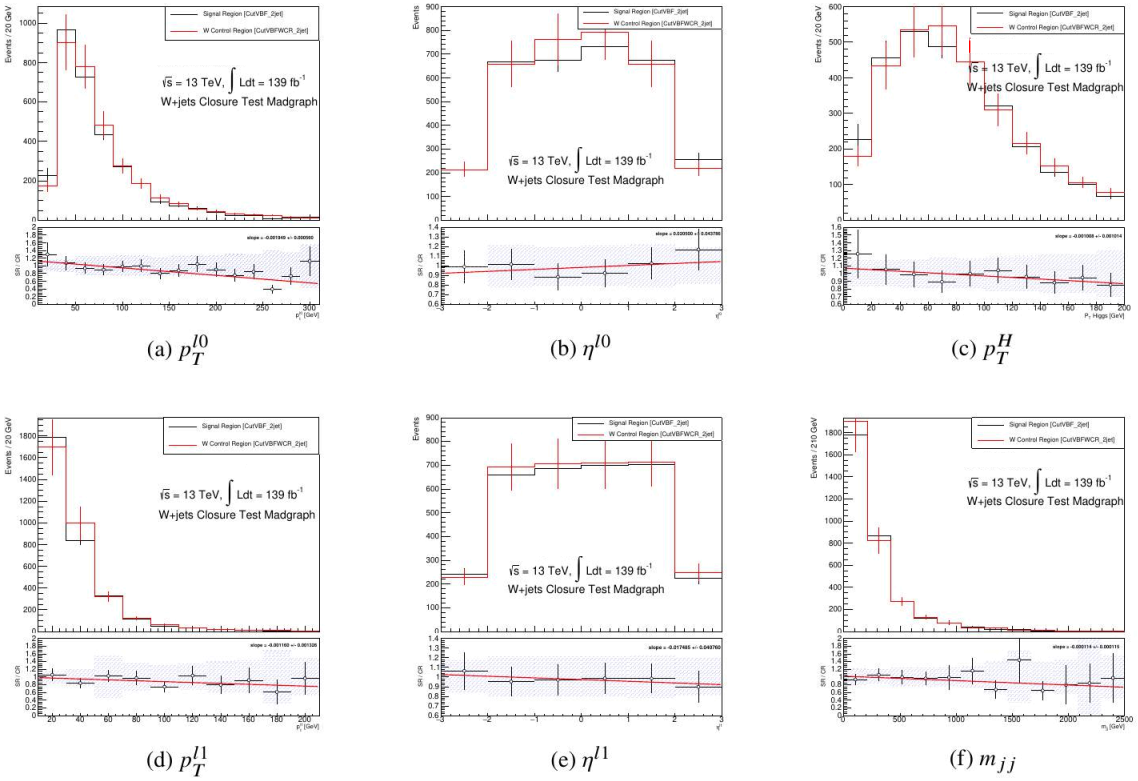
Figure 12.22. Distributions after the full selection for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs applied (red) using Powheg MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable.



(a) p_T^{l0} (b) η^{l0} (c) p_T^H

(d) p_T^{l1} (e) η^{l1} (f) m_{jj}

Figure 12.23. Distributions after the full selection for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CFs (red) using MadGraph MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable.



(a) p_T^{l0} (b) η^{l0} (c) p_T^H

(d) p_T^{l1} (e) η^{l1} (f) m_{jj}

Figure 12.24. Distributions at preselection level for W+jets MC samples in the signal region (black) and W+jets control region with FFs and CRs (red) using MadGraph MC. Statistical uncertainties are shown as error bars. The uncertainties estimated in the data-driven method are shown in error bands. The results of the liner fit of the ratio of the two distributions are shown for each observable.

CHAPTER 13

CONTROL REGION DISCRIMINANTS

The analysis uses multiple dedicated boosted decision trees (bdt) to isolate each significant process from all other processes. The bdts include a Z+jets bdt,ggF bdt and VBF bdt. In addition a bdt is trained to separate Top and WW backgrounds, and a separate bdt to distinguish Top and WW from all other processes.

In the SR, a dedicated selection requirement is added to further reduce the Z+jets background based on the Z+jets bdt. For this bdt, several input variables based on the missing transverse momentum are used, and a cut is applied on the bdt output.

A simultaneous fit of SR and CRs to data is carried out to estimate the amount of $t\bar{t}$, DY and ggF background events based on data as well as extract the signal yield. In the fit the signal strengths of the VBF signal, the $t\bar{t}$, WW and ggF processes are left floating, and the signal strength of $t\bar{t}$ is considered in combination with WW. For the ggF background estimation, three CRs are used with a dedicated ggF bdt for each region. The ggF, VBF and Top WW bdts distributions are used in the simultaneous fit, i.e. no cut is applied on these bdts.

Experimental and theoretical uncertainties are propagated into the simultaneous fit as nuisance parameters and are profiled by the fit.

An overview of the signal, control, and validation regions is shown in Table 1 below. Top and WW validation regions are defined to demonstrate MC modeling of each process in an enriched region, while in the simultaneous fit both of them are estimated directly in the signal region. In Table 2, three ggF control regions

are defined to estimate ggF production Higgs events and a Z+jets control region estimates Drell-Yan background in a region orthogonal to the signal region.

VBF SR	**Z+jets CR**	**Top VR**	**WW VR**
2-lepton	2-lepton	2-lepton	2-lepton
2-jet (30,30)	2-jet (30,30)	2-jet (30,30)	2-jet (30,30)
b-veto (< 1)	b-veto (< 1)	b-veto ($= 1$)	b-veto (< 1)
CJV (20 GeV)	CJV (20GeV)	CJV (20 GeV)	CJV (20 GeV)
OLV bool	OLV bool	OLV bool	$m_T > 130$GeV
Ztt veto	Ztt selection	Ztt veto	$mT_{1jet} > 160$GeV
M_{jj} cut (200GeV)	M_{jj} cut (200GeV)		
DY_{jj} cut (2.1)	DY_{jj} cut (2.1)		
bdt $Zjets$ cut (< 0.5)	bdt $Zjets$ cut (> 0.5)		

Table 13.1. Summarized cuts for VBF signal region, Z+jets control region, top validation region, and WW validation region

W+jets CR	**ggF CR1**	**ggF CR2**	**ggF CR**
1 ID lepton, 1 anti-ID	2-lepton	2-lepton	2-lepton
2-jet (30,30)	2-jet (30,30)	2-jet (30,30)	< 2 jets
b-veto (< 1)	b-veto (< 1)	b-veto (< 1)	b-veto (< 1)
CJV (20 GeV)	CJV>= 20 and OLV= 1 *or*	CJV>= 20 and OLV! = 1	
OLV bool	CJV< 20 and OLV! =1	Ztt veto	
Ztt veto	Ztt selection		
M_{jj} cut (200 GeV)			
DY_{jj} cut (2.1)			
bdt $Zjets$ cut (< 0.5)			

Table 13.2. Summarized cuts for the W+jets control region and ggF control regions (1, 2, 3)

The signal yield is extracted by multiplying its signal strength by the theoretical cross section and is then corrected for experimental efficiency (C-factor) and integrated luminosity to finally extract the signal cross section to be compared to theoretical predictions.

13.1 Gradient Boosted Decision Tree

A decision tree is a collection of successive cuts designed to classify events as signal-like or background-like. A leaf contains all the events left over after all the cuts in a path of the tree are applied. A given signal event is correctly identified if it is placed in a leaf where the majority of events are signal events. and vice-versa for background events.

After the initial tree is built another tree is grown to better separate the signal and background events misidentified by the first tree. This proceeds iteratively until there is a collection of a specified number of trees, in a process known as boosting. There are two main boosting algorithms, adaptive and gradient boosting. In adaptive boosting the data is reweighed every time a new tree is created so that incorrectly classified data is given more weight in the next training. In the gradient boosting case, the error from the previous iteration is calculated for every data point, and a new tree is trained to predict the error made by the previous tree. A weighted average based on the learning rate is taken from all these trees to form a bdt output discriminant with values ranging from -1 to 1, where 1 is signal-like data and -1 in background like-data. The parameters of the bdt tunable by hand, such as the depth of the tree, total number of cuts and number of trees in the ensemble, are called hyperparamters and the total accuracy is evaluated over a large range of values for each hyperparameter.

Unlike decision trees, boosted decision trees (bdt) are robust to overfitting with larger number of trees. Overfitting is when the classifier performs very well on the data used in training but much worse when classifying new data. To check if the classifier overfit, the data is split in half between a training set and testing set, and the hyperparameters of the discriminate is chosen such that the testing set has performance comparable to the training set.

13.1.1 Top WW Discriminant

A bdt is trained using $e\mu+\mu e$ events after the VBF selection and all signal regions cuts so that the phase space in which the bdt is trained is exactly the same as the one where it is applied in the final fit. Since the top-quark and WW processes are very difficult to disentangle from one another as they share the same final state, the bdt is trained to discriminate top-quark and WW (top+WW) processes against VBF, ggF, Z+jets, and $V\gamma$ events. The MC statistics used in the training are half those available after all signal region cuts, while the other half of the statistics is then used to test the bdt training. This corresponds to about 90,000 un-weighted WW and top events and about 115,000 raw VBF, ggF, Z+jets and $V\gamma$ events. This training is carried out using weighted MC events to best account for overall event distributions and include about 2000 total weighted top and WW events and about 550 weighted events from other processes. The optimal parameters were found through a scan of reasonable values and the final set is summarized in Table 13.3.

Parameter	value
Boosting Algorithm	Gradient
Max Tree Depth	22
Number Of Trees	200
Minimum number of events per leaf	5%
Number of cuts	7

Table 13.3. bdt parameters used for the training of the bdt of top + WW vs. other processes.

This bdt utilizes eight lepton and jet kinematic variables to distinguish between signal and background events. These include δY_{jj} , the combination of lepton/jet mass M_{l0j0} for leading lepton and jet , $\Delta\phi_{ll}$, m_T , η_{j0} , η_{j1} , $\Delta\phi_{jj}$, and sum of centralities (L). Plots in figures 13.1 and 13.2 show the distribution and correlation of input variables.

Table 13.4 shows the ranking, i.e. importance, of the variables used as input to the bdt training, and we can see that m_T is dominating.

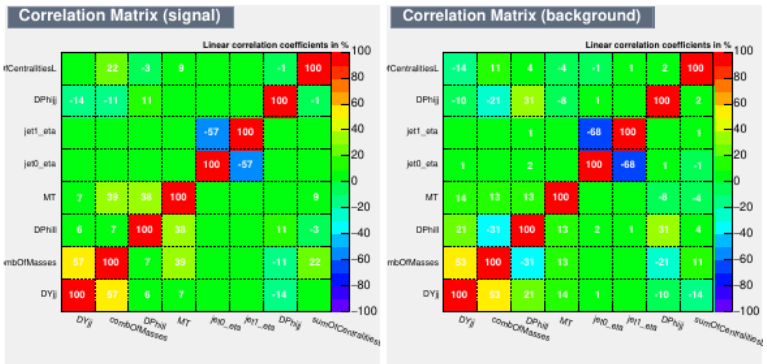


Figure 13.1. Correlations between in the input variables for the singal and background samples for WW+Top bdt

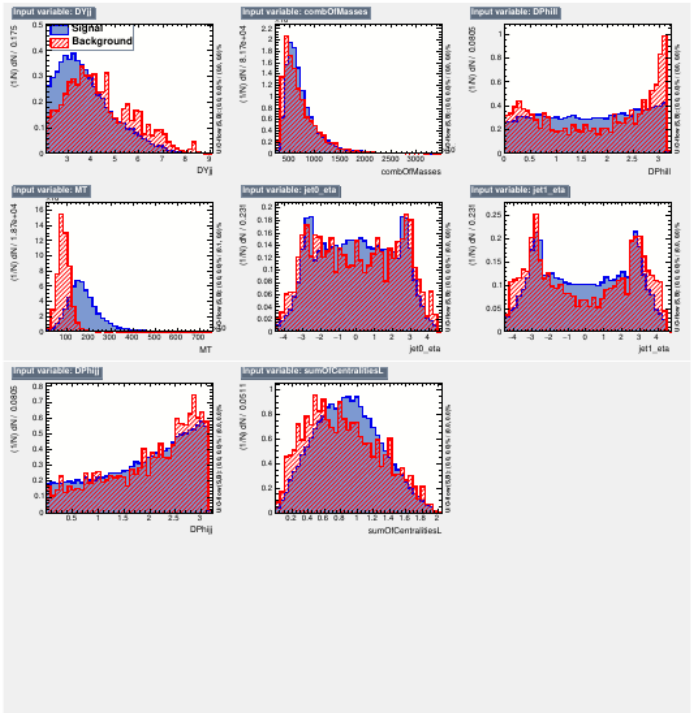


Figure 13.2. Signal and background spectrums of the input variables to WW+Top bdt

The bdt training successfully separates top+WW processes from other processes. In order to quantify the discrimination, an integrated-ROC is calculated through TMVA for weighted normalized samples and an optimal value of 0.924 is found. Comparisons between the test and training, as seen for instance in Fig. 13.3 (left), show that the bdt is un-biased: differences between testing and training distributions of the bdt would imply overtraining, or that the bdt uses too many parameters on too few events. For signal and background, KS-test values of 0.263 and 0.311 are found, and such values show no evidence of over-training. The plot in Fig. 13.3 (right) shows bdt output distribution for all processes in the SR.

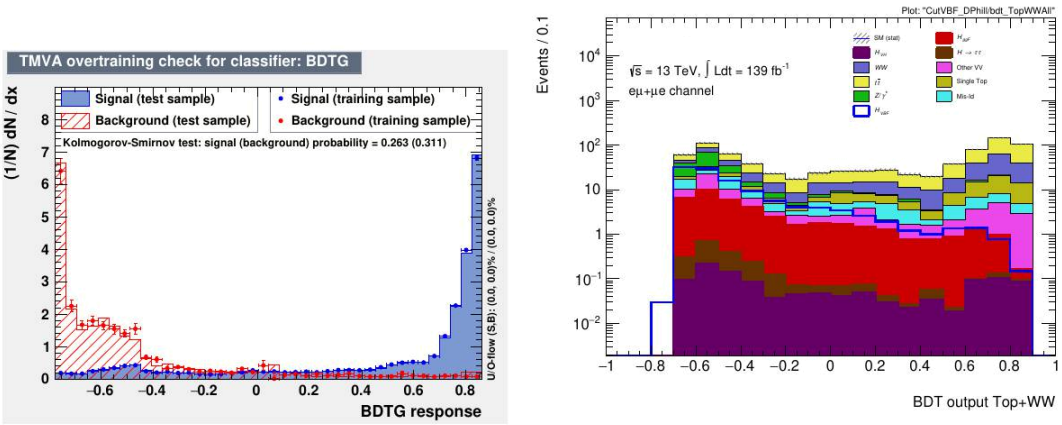


Figure 13.3. (Left) comparison of the train and test samples for the WW+Top bdt. The discrepancy between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process

This bdt output distribution is used in the signal region to take advantage of high significance for top+WW events in the uppermost bins of the distribution. Since this bdt is trained and applied in the signal region, the modelling of input variables is not directly tested, however, modeling at the pre-selection level for each of these variables as well as in the top validation region show no evidence of any significant mis-modelling.

Variable	Importance
m_T	0.64
ΔY_{jj}	0.11
$\Delta\phi_{ll}$	0.08
$\Delta\phi_{jj}$	0.05
Σ Centralities	0.04
Mass Combination M_{j0l0}	0.04
η_{j1}	0.03
η_{j0}	0.01

Table 13.4. Ranking of the variables used as input to the training of the bdt of top + WW vs. other processes.

13.2 GGF Discriminant

The multivariate discriminant used for the ggF-CR1 is a bdt trained using $e\mu+\mu e$ events that pass all ggF-CR1 cuts. The training is for ggF events against VBF signal and all other backgrounds. Half the MC data is used for training and the other half for testing. This corresponds to about 8,000 ggF events and around 250,000 events of signal and other background processes. Events are trained with MC weights applied to accurately represent the sample composition, corresponding to 130 weighted ggF events and 10,000 weighted events from other processes. The final set of hyper parameters is summarized in Table 13.5.

Parameter	value
Boosting Algorithm	Gradient
Max Tree Depth	10
Number Of Trees	600
Minimum number of events per leaf	5%
Number of cuts	7

Table 13.5. bdt parameters used for the training of the bdt of ggF vs. other processes.

This bdt utilizes eight lepton and jet kinematic variables to distinguish between signal and background events. These include δY_{ll} , $\Delta\phi_{ll}$, m_T , M_{ll} $\Delta\phi_{jj}$, E_T^{Miss},

Jet P_T^{lead} and Jet $P_T^{sublead}$. Plots in figures 13.4 and 13.5 show the distribution and correlation of input variables.

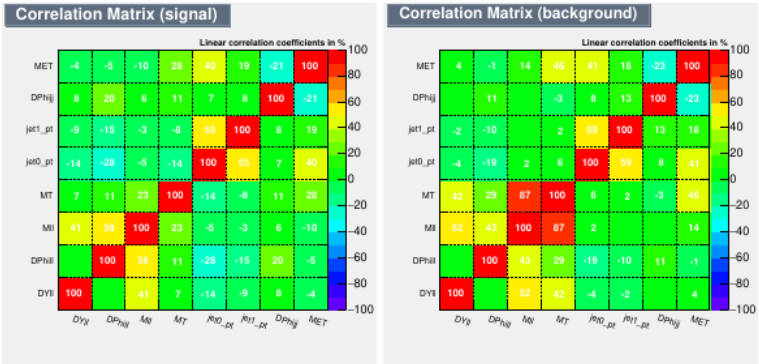


Figure 13.4. Correlations between in the input variables for the singal and background samples for ggF bdt

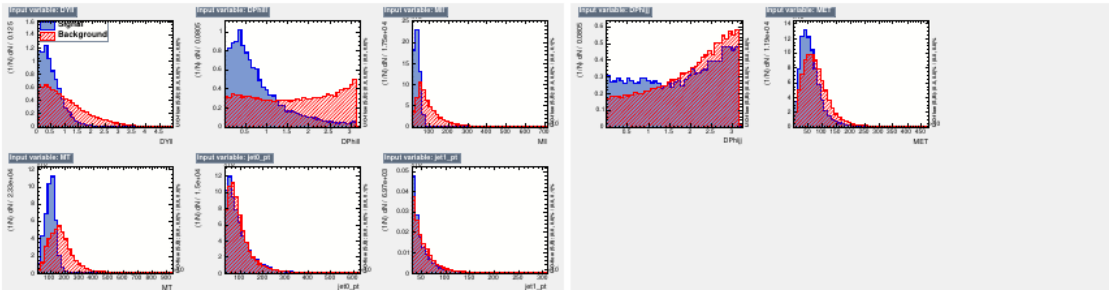


Figure 13.5. Signal and background spectrums of the input variables to ggF bdt

The optimal value of the integrated-ROC calculated is 0.898. Differences between testing and training samples would imply overtraining, or that the bdt uses too many parameters on too few events. The comparison between the test and training samples in Fig. 13.6 shows that the bdt is un-biased, as we can see that the testing and trainings samples are very similar. Additionally, a Kolmogorov-Smirnov test is performed to measure if the test and training distributions differ significantly, and no evidence of over-training is present (values 0.191, 0.049 for signal, background, respectively).Figure 13.6 shows the signal to background discrimination of the bdt. Finally table 13.6 shows the importance of each feature.

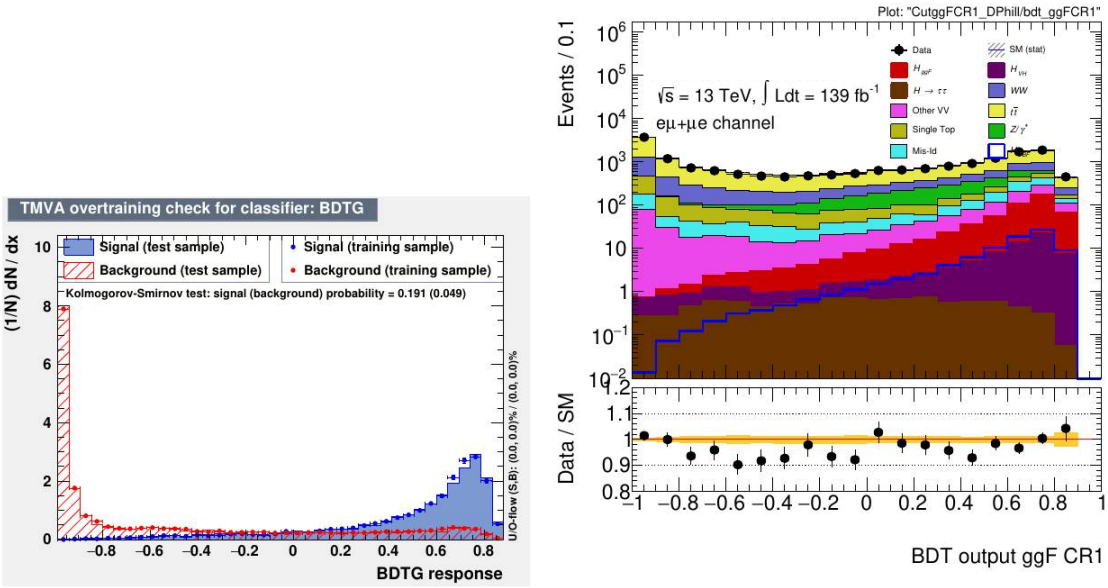


Figure 13.6. (Left) comparison of the train and test samples for the ggF bdt. The discrepency between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process

Variable	Importance
m_T	0.25
$\Delta\phi_{ll}$	0.24
Δy_{ll}	0.17
ΔM_{ll}	0.16
$\Delta\phi_{jj}$	0.12
E_T^{Miss}	0.04
P_{j0}^T	0.00
P_{j1}^T	0.00

Table 13.6. Ranking of the variables used as input to the training of the bdt trained in the ggF Control Region.

13.3 VBF Discriminant

The most useful bdt for this analysis is trained to discriminate between the signal process (VBF produced Higgs boson) and the Top and WW process. This bdt is trained using $e\mu + \mu e$ events after the VBF selection and the signal regions cuts including that on n_{jets} , b-veto, OLV, CJV, $M_{jj} > 200$ GeV and $\Delta y_{jj} > 4.5$. In this way, the phase space in which we train the bdt is broader than the one where we apply it. The training includes only the $t\bar{t}$, Wt, WW as well as other diboson process, as backgrounds, and the VBF signal. The diboson processes are dominated by WW events. The MC statistics used in the training are half of those available after all signal region cuts (as the other half of the event is then used to test the training). This corresponds to 90,000 un-weighted WW and top events, and 100,000 raw VBF events. The training includes MC weights on events to best account for the expected process composition. There are 2000 total weighted top and WW events used in the training, and 80 weighted VBF events. The final set of parameters are listed in table 13.7.

The inut variables to the VBF bdt include ΔY_{ll} , $\Delta\phi_{ll}$, m_T , M_{ll}, $\Delta\phi_{jj}$, Jet P_T^{lead}, Jet $P_T^{sublead}$, Sum of Centralities, ΔY_{ll}, M_{jj}, η_{j0} and η_{j1}. Figures 13.7 and 13.8 shows the distribution and correlation of the input features. Figure 13.9 shows how well the bdt discriminates between signal and the backgrounds. The integrated ROC curve

Parameter	value
Boosting Algorithm	Gradient
Max Tree Depth	22
Number Of Trees	400
Minimum number of events per leaf	5%
Number of cuts	7

Table 13.7. bdt parameters used for the training of the bdt of VBF vs. other processes.

has a value of 0.96. Additionally, a Kolmogorov-Smirnov (KS) test is performed to measure if the the test and training bdt output distributions differ significantly. If the two distributions are random samples of the same parent distribution, the KS-test would give a uniformly distributed value between zero and one (or an average value of 0.5). The closer to 0.5 the KS-test, the greater likelihood the curves come from the same parent, however this calculation is heavily skewed toward lower values so any value above zero (or not very close to zero, on the order of order 10^{-4}) can be considered not indicative of overtraining. For signal and background we find KS-test values of 0.107 and 0.154, and so no evidence of over-training. Tabel 13.8 shows the ranking, i.e. the importance of the variables used as input to the VBF vs Top+WW bdt. Figure 13.10 shows the signal-to-background ratio in the signal region as a function of the bdt used for the.measurement of the inclusive cross section.

13.4 N-1 Bdts

For the differential fits, a new bdt is trained for each observable used as an input variable in the original vbf bdt, where that observable, and highly correlated observables, are removed as an input variable. These N-1 bdts are used as the discriminants in the differential fit in order to not bias the results of the fit. The removed variables have at least 40% correlation with the observable. Table 13.9 shows which variables are removed for which N-1 bdt. Figures 13.11 through 13.13 show the discrimination

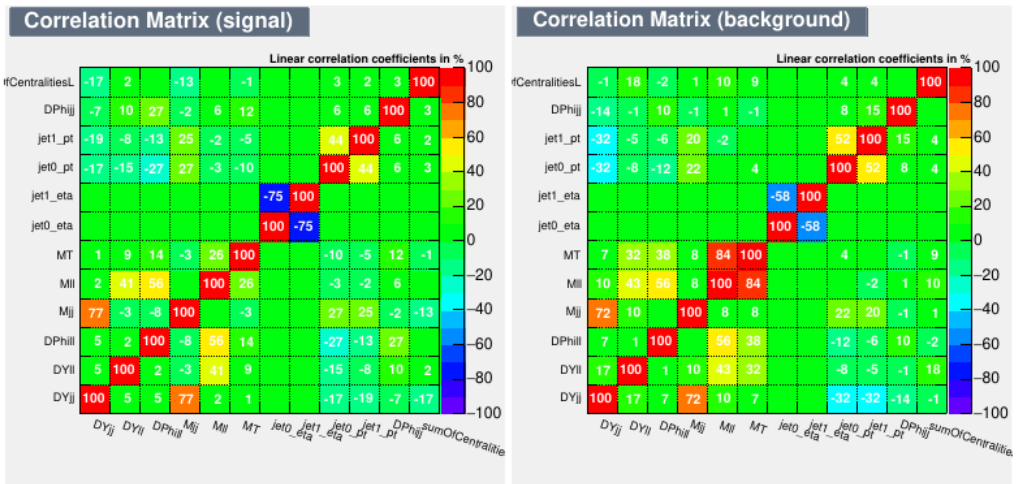


Figure 13.7. Correlations between in the input variables for the singal and background samples for vbf bdt

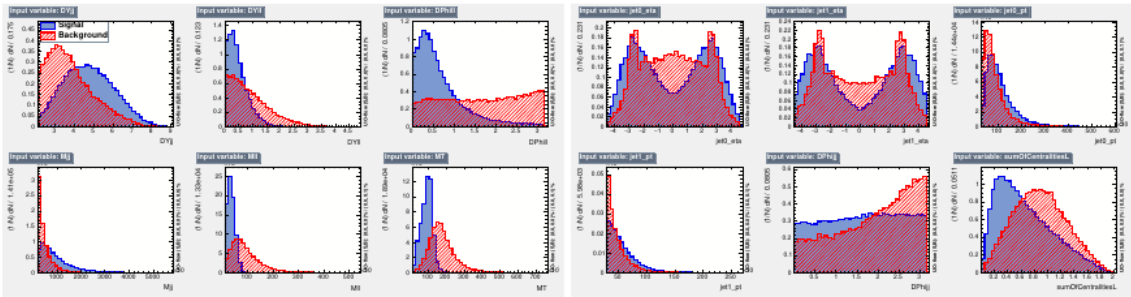


Figure 13.8. Signal and background spectrums of the input variables to vbf bdt

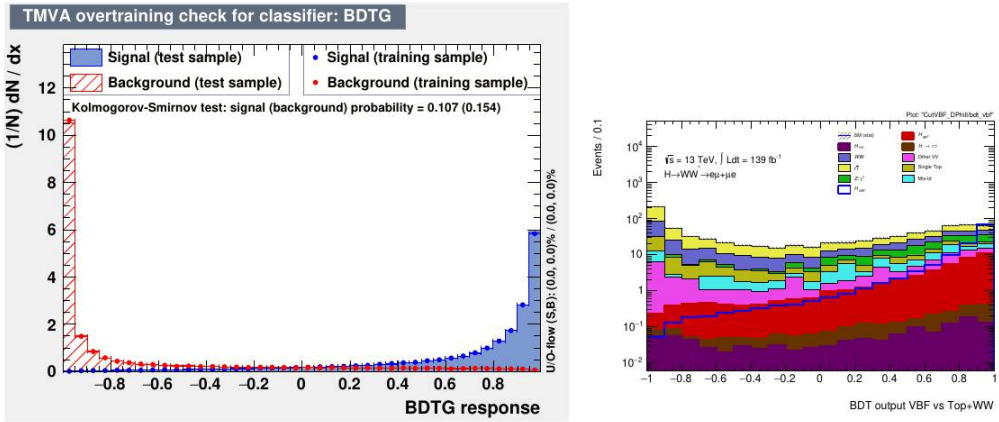


Figure 13.9. (Left) comparison of the train and test samples for the vbf bdt. The discrepency between the train and test samples is small indicating theres not overfitting.On the right is the expected data distribution for the bdt broken down by process

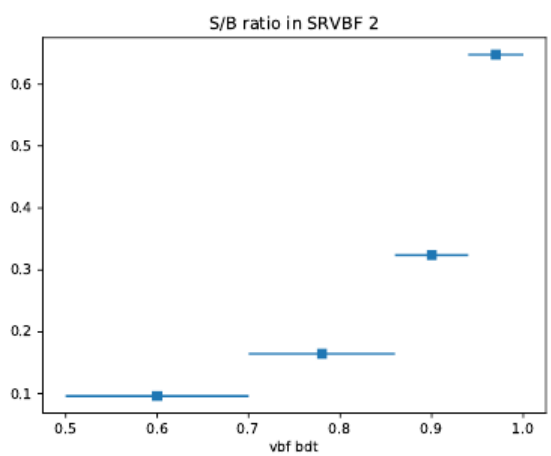


Figure 13.10. Signal to background ratio in each bin of the vbf bdt. The binning is the one used in the inclusive fit: 0.5-0.7,0.7-0.86,0.86-0.94,0.94-1.0

Variable	Importance
m_T	0.14
$\Delta\phi_{ll}$	0.12
Δy_{jj}	0.10
ΔM_{jj}	0.09
ΔM_{ll}	0.09
P^T_{j0}	0.08
Σ Centralities	0.08
Δy_{jj}	0.07
$\Delta\eta_{j0}$	0.07
$\Delta\eta_{j1}$	0.07
P^T_{j1}	0.06
$\Delta\phi_{jj}$	0.05

Table 13.8. Ranking of the variables used as input to the training of the bdt VBF vs Top+WW.

power of each N-1 bdt. The figures show that the N-1 bdts don't lose significant discriminatory power compared to the full vbf bdt.

Variable	Removed Variables from vbf bdt
Signed $\Delta\phi_{jj}$	$\lvert\Delta\phi_{jj}\rvert$
$\Delta\phi_{ll}$	$\Delta\phi_{ll}$ and m_{ll}
Δy_{jj}	Δy_{jj} and m_{jj}
Δy_{ll}	Δy_{ll} and m_{ll}
P^T_{j0}	P^T_{j0} and P^T_{j1}
P^T_{j1}	P^T_{j0} and P^T_{j1}
m_{jj}	Δy_{jj} and m_{jj}
m_{ll}	Δy_{ll},$\Delta\phi_{ll}$, and m_{ll}

Table 13.9. List of N-1 bdts and the removed variables

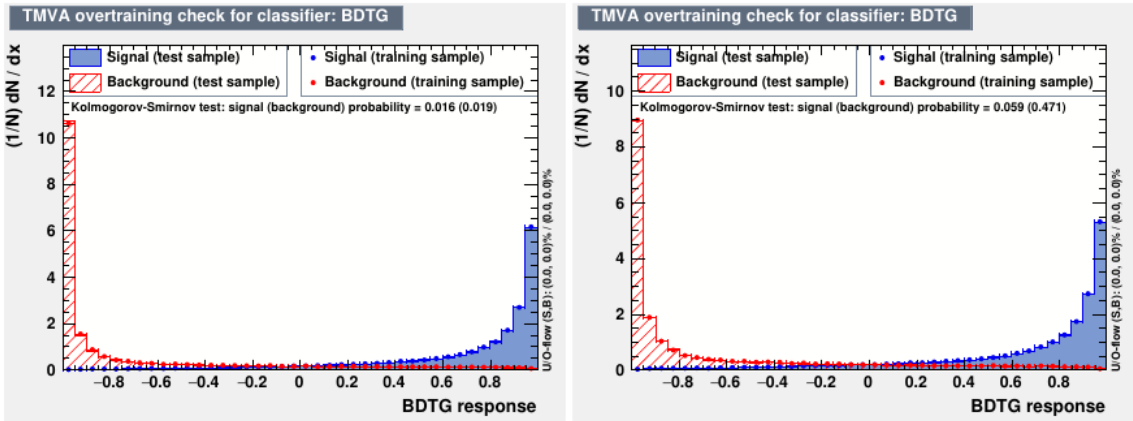


Figure 13.11. bdt output distribution of test and train data for vbf signal, top+ww background for $\Delta\phi_{jj}$(left) and $\Delta\phi_{ll}$(right)

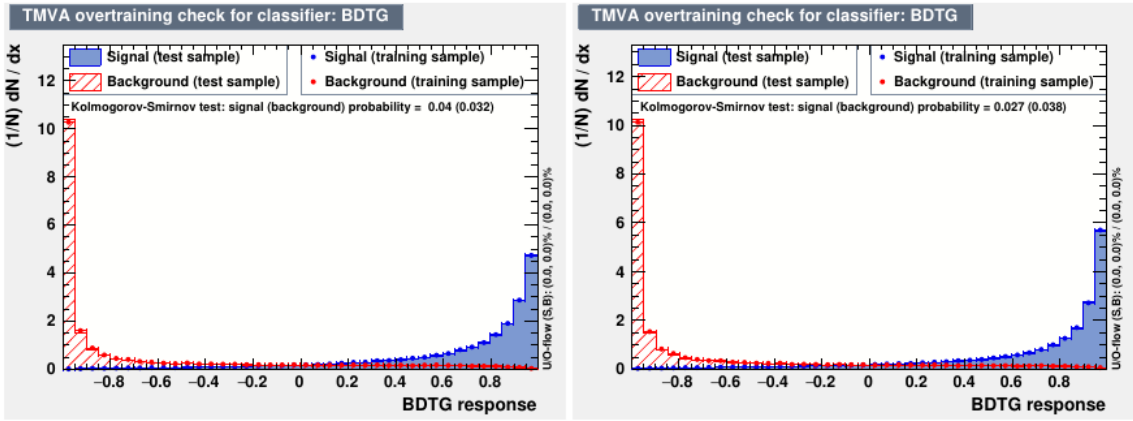


Figure 13.12. bdt output distribution of test and train data for vbf signal, top+ww background for Δy_{jj}(left) and Δy_{ll}(right)

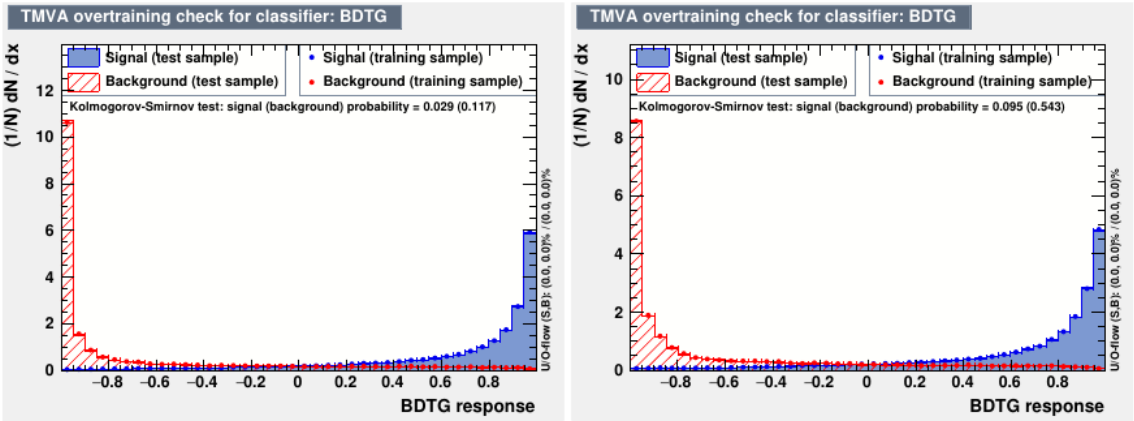


Figure 13.13. bdt output distribution of test and train data for vbf signal, top+ww background for $\Delta jetpt_{jj}$(left) and Δm_{ll}(right)

CHAPTER 14

SYSTEMATIC UNCERTAINTY

The uncertainty on the final measurement are the squared sum of uncertainty due statistical fluctuations and uncertainty due to limitations of the detectors or theoretical understanding of the underlying physics. Several sources of systematics are investigated for both the fiducial and the differential cross section measurements. This section discusses the sources of systematic uncertainties and their impacts on the estimated cross sections. The sources of systematic uncertainties can be divided into two main categories, experimental and theoretical. Experimental uncertainties are due to the limited abilities of the ATLAS detectors, calibration and reconstruction algorithms. Theoretical uncertainties are from limit the modeling of physics processes.

14.1 Experimental Uncertainties

Several types of experimental systematics uncertainties affect this analysis. Each of these are calculated for this analysis using the latest recommendations from the Combined Performance (CP) groups, for muons, electrons, jet/E , Trigger, flavor tagging, and pile-up. Table 55 summarizes the experimental uncertainties that are considered in the analysis. Here follows a brief description of the sources of experimental systematics, and the corresponding nuisance parameters used in the analysis, as recommended by the CP groups. The standard set of experimental systematics is made of 67 four-momentum nuisance parameters and 59 efficiency scale factor nuisance parameters for each muon, electrons and jet uncertainties.

The most significant uncertainties come from the jet energy resolution, jet energy scale and MET calculations. MET is calculated by taking the negative sum of all the momentum from the event. The total vector momentum should end up as zero so leftover momentum is invisible particles such as neutrinos (or new physics) or misreconstructed particles which leads to systematic uncertainty. The jet uncertainty comes from limited resolution of detectors when reconstructing the total energy of the jets. Jet energy scale uncertainty comes from uncertainty in measuring the scale factor to convert detector readings to energy deposits in the calorimeter.

Jet and lepton uncertainties are evaluated by testing the number of events in the final state affected by shifting energy/momentum in the MC simulation by a scale factor. This is done by recalculating the weights applied to MC to match data after shifting the energy/momentum spectrum from the nominal. The shifted spectrum gives the shape and normalization uncertainty. Finally the overall integrated luminosity uncertainty is 1.7% from the full run 2 dataset.x

The full list of systematic uncertainties are listed in table 14.1.

14.2 Theory Uncertainty

Theory systematics are calculated per sample, for Signal VBF, $Z \to \tau\tau$, top,WW and ggF samples. Theory systematics for the other backgrounds is negligible. The systematics are described in table 14.2 and in detail in the following section.The most significant uncertainties are estimates of the parton density function (PDF), the strong coupling constant, α_s estimates of the QCD scale parameter, matching uncertainty between QCD prediction and parton showers, and generator uncertainties. The uncertainties are evaluated in one of multiple ways. For two point systematics, such as generator uncertainties, an alternate generator is used to calculate the expected yield from the nominal one. The expected uncertainty is then calculated by reflecting the yield from the alternate generator across the nominal one, which gives an envelope

Systematic uncertainty	Short description
	Event
Luminosity	uncertainty on total integrated luminosity
PRW_DATASF	uncertainty on pileup reweighting
	Electrons
EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_TriggerEff_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty
EL_EFF_ID_CorrUncertaintyNP (0 to 15)	ID efficiency uncertainty splits in 16 components
EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP (0 to 17)	ID efficiency uncertainty splits in 18 components
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty
EG_SCALE_ALL	energy scale uncertainty
EG_SCALE_AF2	
EG_RESOLUTION_ALL	energy resolution uncertainty
	Muons
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty
MUON_EFF_TrigSystUncertainty	
MUON_EFF_RECO_STAT	reconstruction and ID efficiency uncertainty for muons with $p_T > 20$ GeV
MUON_EFF_RECO_SYS	
MUON_EFF_RECO_STAT_LOWPT	reconstruction and ID efficiency uncertainty for muons with $p_T < 20$ GeV
MUON_EFF_RECO_SYS_LOWPT	
MUON_ISO_STAT	isolation efficiency uncertainty
MUON_ISO_SYS	
MUON_TTVA_STAT	track-to-vertex association efficiency uncertainty
MUON_TTVA_SYS	
MUON_ID	momentum resolution uncertainty from inner detector
MUON_MS	momentum resolution uncertainty from muon system
MUON_SCALE	momentum scale uncertainty
MUON_SAGITTA_RHO	charge dependent momentum scale uncertainty
MUON_SAGITTA_RESBIAS	
	Jets
JET_EffectiveNP_Detector (1 to 2)	detector related energy scale uncertainty
JET_JER_EffectiveNP_ (1 to 11)	
JET_JER_EffectiveNP_12RestTerm	energy resolution uncertainty
JET_JER_DataVsMC_MC16	
JET_BJES_Response	energy scale uncertainty on b-jets
JET_EffectiveNP_Mixed (1 to 3)	energy resolution uncertainty
JET_EffectiveNP_Modelling (1 to 4)	energy scale uncertainty on eta-intercalibration (modeling)
JET_EffectiveNP_Statistical (1 to 6)	statistical resolution uncertainty
JET_EtaIntercalibration_NonClosure_highE	
JET_EtaIntercalibration_NonClosure_negEta	energy scale uncertainty on eta-intercalibrations (non-closure)
JET_EtaIntercalibration_NonClosure_posEta	
JET_EtaIntercalibration_TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)
JET_EtaIntercalibration_Modelling	energy scale uncertainty on eta-intercalibrations (modeling)
JET_Pileup_OffsetMu	energy scale uncertainty on pile-up (mu dependent)
JET_Pileup_OffsetNPV	energy scale uncertainty on pile-up (NPV dependent)
JET_Pileup_PtTerm	energy scale uncertainty on pile-up (pt term)
JET_Pileup_RhoTopology	energy scale uncertainty on pile-up (density ρ)
JET_Flavor_Composition	energy scale uncertainty on flavour composition
JET_Flavor_Response	energy scale uncertainty on samples' flavour response
JET_PunchThrough_MC16	energy scale uncertainty for punch-through jets
JET_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high-p_T jets
JET_JvtEfficiency	JVT efficiency uncertainty
FT_EFF_Eigen_B (0 to 2)	b-tagging efficiency uncertainties ("BTAG_MEDIUM"): 3 components for b jets, 3 for c jets and 4 for light jets
FT_EFF_Eigen_C (0 to 2)	
FT_EFF_Eigen_Light (0 to 3)	
FT_EFF_extrapolation	b-tagging efficiency uncertainty on the extrapolation to high-p_T jets
FT_EFF_extrapolation_from_charm	b-tagging efficiency uncertainty on tau jets
	MET
MET_SoftTrk_ResoPara	track-based soft term related longitudinal resolution uncertainty
MET_SoftTrk_ResoPerp	track-based soft term related transverse resolution uncertainty
MET_SoftTrk_Scale	track-based soft term related longitudinal scale uncertainty

Table 14.1. Summary of the experimental systematic uncertainties considered.

of uncertainty. The ranking of the uncertainty impact on the final measurement after the statistical fit is given in the following chapters [80].

14.2.0.1 VBF Uncertainties

The VBF theoretical uncertainties include parton shower (PS), PDF, α_s , QCD scale variations and matching scheme. Parton shower uncertainties are derived from comparison between different parton shower generators (Powheg+Pythia8 and Powheg+Herwig7)

process	uncertainty	variations	method
VBF	QCD scale	7-point	envelope
	PDF	PDF4LHC15_nlo_30_pdfas(0-30)	symmetric Hessian $\sqrt{\sum_k (X_k - X_0)^2}$
	α_s	PDF4LHC15_nlo_30_pdfas(31-32)	average
	matching	Powheg+Herwig7, MG5_aMC@NLO+Herwig7	single variation
	PS/UE	Powheg+Pythia8, Powheg+Herwig7	single variation
ggF	QCD scale	7-point	envelope
	QCD scale	Stewart-Tackmann	Stewart-Tackmann
	PDF	PDF4LHC15_nlo_30_pdfas(0-30)	symmetric Hessian $\sqrt{\sum_k (X_k - X_0)^2}$
	α_s	PDF4LHC15_nlo_30_pdfas(31-32)	average
	PS/UE	Powheg+Pythia8, Powheg+Herwig7	single variation
top	QCD scale	7-point	envelope
	PDF	NNPDF30_nlo_as_0118	standard deviation $\sqrt{\frac{1}{N}\sum_i (X_i - X_0)^2}$
	matching	Powheg+Pythia8, MG5_aMC@NLO+Pythia8	single variation
	PS/UE	Powheg+Pythia8, Powheg+Herwig7	single variation
	ISR	Var3cUp, Var3cDown	up down variation
	FSR	muRfac= 2.0, muRfac= 0.5	up down variation
	Top/Wt interference	DS, DR scheme	single variation
WW (reco -level)	QCD scale	7-point	envelope
	PDF	NNPDF30_nnlo_as_0118	standard deviation $\sqrt{\frac{1}{N}\sum_i (X_i - X_0)^2}$
	α_s	NNPDF30_nnlo_as_0117, NNPDF30_nnlo_as_0119	up down variation
WW (truth -level)	merging	ckkw15, ckkw30	up down variation
	resummation	QSF down, QSF up	up down variation
	PS/UE	Powheg+Pythia8, Powheg+Herwig++	single variation
	PS recoil scheme	CSSKIN	single variation
Z$\tau\tau$	QCD scale	7-point	envelope
	PDF	NNPDF30_nnlo_as_0118	standard deviation $\sqrt{\frac{1}{N}\sum_i (X_i - X_0)^2}$
	α_s	NNPDF30_nnlo_as_0117, NNPDF30_nnlo_as_0119	up down variation
	generator	Sherpa2.2.1+Pythia8, MG5_aMC@NLO+Pythia8	single variation

Table 14.2. A summary of theory systematics of all processes, including the types of theory uncertainty and their calculation methods.

and matching uncertainties from the comparison between different matrix element simulation (Powheg+Herwig7 and Madgraph5+Herwig7). These uncertainties constitute the largest theoretical uncertainties in the signal region. The variation is obtained by internal re-weighting within the generation of the nominal Powheg+Pythia8 sample, which is then used to calculate for the PDF, α_s and QCD scale uncertainties. The α_s uncertainties are accounted for by varying the α_s value by ± 0.001 with respect to its central value of 0.118. QCD scale uncertainties include variations of factorization and renormalization scales by factors of 2 and 0.5, and are found with

an 7-point envelope scheme, in which unphysical diagonal terms, e.g. renormalization scale multiplied by 2.0 and factorization scale multiplied by 0.5, are not considered.

systematics and their impact in signal yields in the signal region, control/validation regions are shown in Table 14.3.

These uncertainties are examined in Figs. 14.1 (signal region 2) and 14.2 (signal region 1) for any shape effects on the BDT output used in the final fit. The plots in the signal region show very little shape effects.

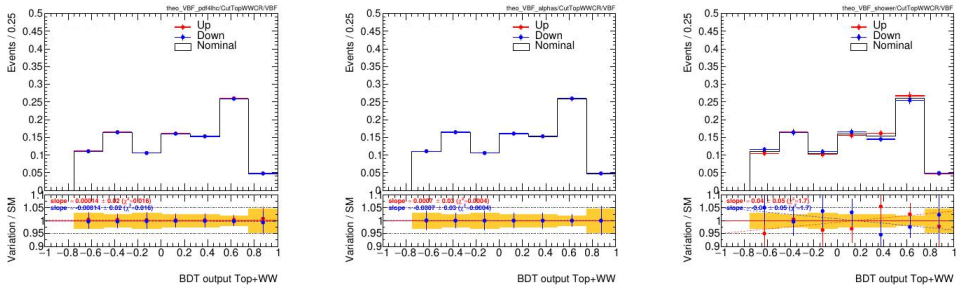
Fit regions	Impact high / low (+/- stat) [%]							
	VBF_SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_VBF_pdf4lhc (VBF)	1.8	1.8	1.7	3.4	1.7	1.7	1.7	1.7
theo_VBF_alphas (VBF)	0.4	0.5	0.3	1.0	0.4	0.5	0.5	0.7
theo_VBF_scale (VBF)	0.4 / -0.3	1.0 / -1.1	0.4 / -0.7	1.9 / -1.3	0.6 / -0.3	0.7 / -0.9	1.0 / -0.9	0.5 / -0.3
theo_VBF_generator (VBF)	9.4 ± 1.2	13.6 ± 3.9	5.5 ± 3.3	14.9 ± 3.1	8.9 ± 3.0	4.3 ± 1.1	0.8 ± 3.6	-7.1 ± 0.8
theo_VBF_shower (VBF)	-9.8 ± 0.7	-1.2 ± 2.4	-10.0 ± 1.9	-15.5 ± 1.5	1.6 ± 2.0	8.7 ± 0.8	15.1 ± 2.8	4.0 ± 0.6

Table 14.3. VBF theory uncertainties impact on VBF yields in the signal,control and validation regions

14.2.0.2 GGF Uncertainties

The ggF parton shower uncertainties use the same prescription described for the VBF samples. The ggF theoretical uncertainties also include PDF, and α_s . PDF and α_s uncertainties are derived with the same strategy as the VBF process. The α_s uncertainties are accounted for by varying the α_s value by ± 0.001 with respect to its central value of 0.118.

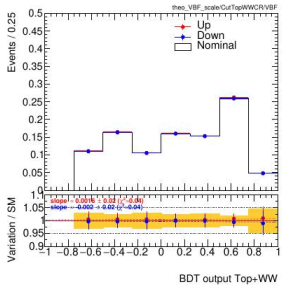
For the evaluation of the QCD scale uncertainty, the Stewart-Tackmann (ST) method is used. The uncertainty calculated from this method is reported in Table 14.4. Since the ST method depends on a certain spectrum of the Higgs transverse momentum and on the transverse momentum of the Higgs plus leading two jets system for jet multiplicity greater than one, the central jet veto (CJV) might modify the spectrum significantly. The shape difference of transverse momentum of the Higgs



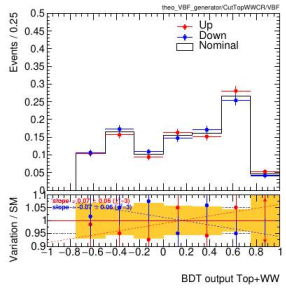
(a) PDF uncertainty

(b) α_s uncertainty

(c) Shower uncertainty

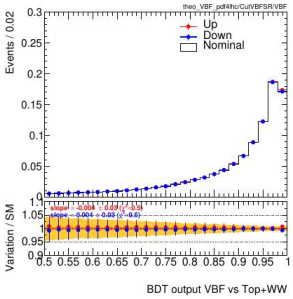


(d) QCD scale uncertainty

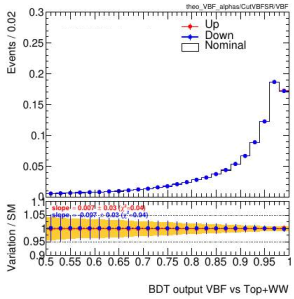


(e) Generator uncertainty

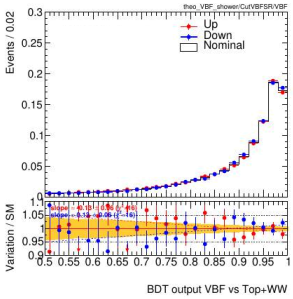
Figure 14.1. Up and down variations on the VBF bdt output distribution for the considered theoretical uncertainties on the VBF signal. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$.



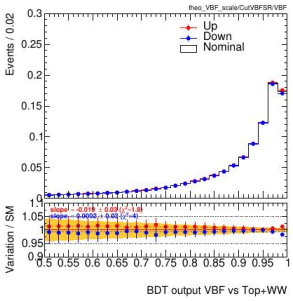
(a) PDF uncertainty



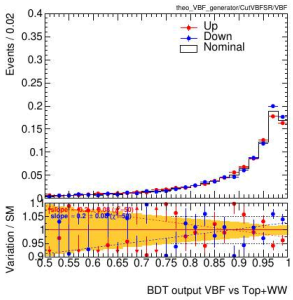
(b) α_s uncertainty



(c) Shower uncertainty



(d) QCD scale uncertainty



(e) Generator uncertainty

Figure 14.2. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the VBF signal. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.

plus two leading jets system is studied before and after the cut associated with the third jet in figure 14.3 The difference is small and therefore we conclude the ggF control region phase space definition does not introduce large extrapolation and the ST method is valid for QCD scale uncertainty evaluation. The BDT output of VBF and ggF-CR1 of the uncertainties used in the fit are shown in figure 14.4 and figure 14.5

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_VR2	ggF_VR3
theo_ggf_pdf (ggF)	1.5	1.8	1.7	2.2	2.0	1.9	2.0	2.3
theo_ggf_alphas (ggF)	3.3	3.2	3.3	3.4	3.2	3.2	3.2	2.0
theo_ggf_QCDscale_ggH (ggF)	2.9	3.9	3.9	6.8	6.8	7.3	7.5	4.4
theo_ggf_QCDscale_ggH_res (ggF)	2.6	3.2	3.5	6.6	6.4	7.0	7.2	1.9
theo_ggf_QCDscale_ggH_ptH_m01 (ggF)	2.6	3.3	3.0	4.3	4.1	4.5	4.5	0.5
theo_ggf_QCDscale_ggH_m12 (ggF)	0.9	0.1	3.0	10.1	9.4	10.7	11.6	-2.0
theo_ggf_QCDscale_ggH_qm_t (ggF)	1.9	0.9	3.6	1.6	2.2	1.5	2.1	0.1
theo_ggf_QCDscale_ggH_pTH (ggF)	3.5	1.5	6.2	2.3	5.1	3.3	4.1	0.6
theo_ggf_QCDscale_ggH_vbf2j (ggF)	10.6	6.4	8.3	2.5	1.7	1.3	1.1	0.1
theo_ggf_QCDscale_ggH_vbf3j (ggF)	-4.7	-0.7	-4.1	-0.4	0.3	0.5	0.6	-0.0
theo_ggf_shower (ggF)	-18.5 ± 2.1	-12.9 ± 3.9	-20.0 ± 5.7	-5.2 ± 2.9	-3.8 ± 2.3	4.7 ± 0.6	19.4 ± 1.5	-4.3 ± 0.3
theo_ggf_QCDscale_7pt_envelope (ggF)	11.6 / -11.7	15.8 / -8.5	13.4 / -12.5	12.7 / -9.1	12.0 / -12.3	11.0 / -11.2	13.1 / -12.1	1.5 / 0.0

Table 14.4. ggF theory uncertainties impact on ggF yields in the signal,control and validation regions

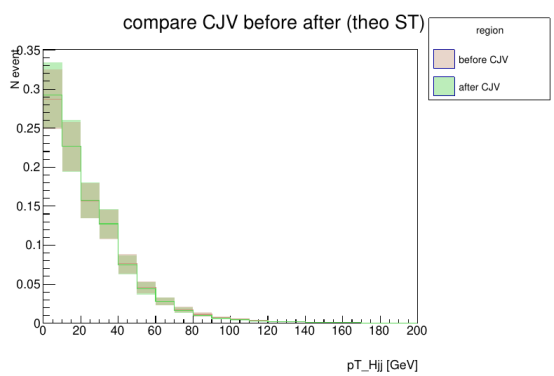
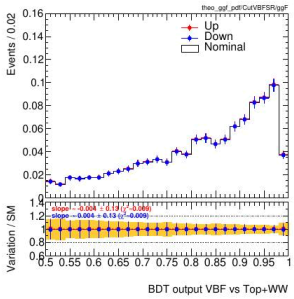


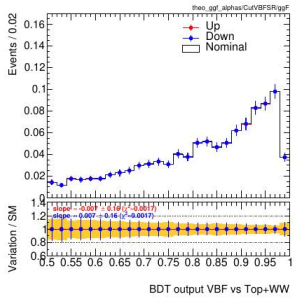
Figure 14.3. The shape difference of transverse momentum of the Higgs plus two leading jets system before and after the cut associated with a third jet.

14.2.1 Top Uncertainties

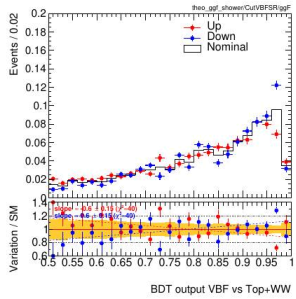
The top background consists of $t\bar{t}$ and Wt (single-top) processes where $t\bar{t}$ is the dominant background by far. Theoretical uncertainties are produced for both single



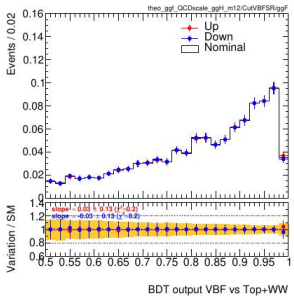
(a) PDF uncertainty



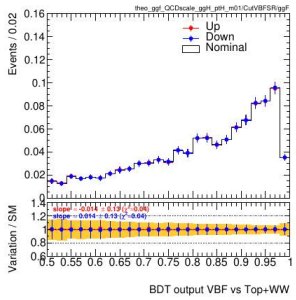
(b) α_s uncertainty



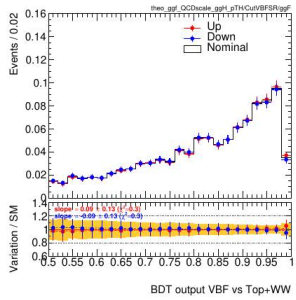
(c) Shower uncertainty



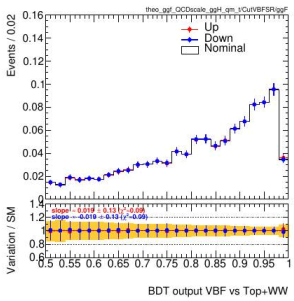
(d) QCD scale uncertainty (m12)



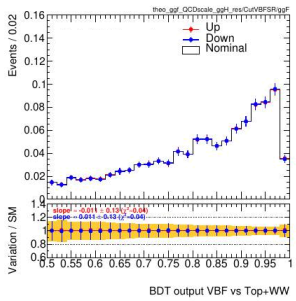
(e) QCD scale uncertainty (ptH_m01)



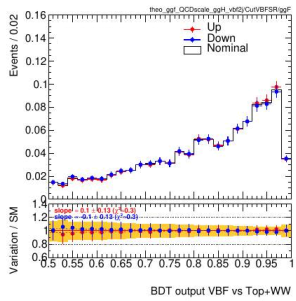
(f) QCD scale uncertainty (pTH)



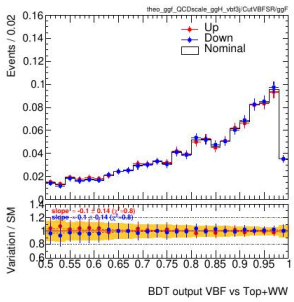
(g) QCD scale uncertainty (qm_t)



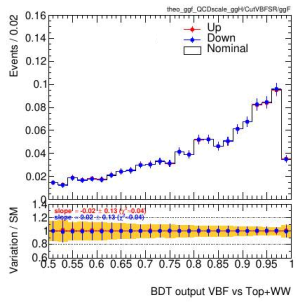
(h) QCD scale uncertainty (res)



(i) QCD scale uncertainty (vbf2j)

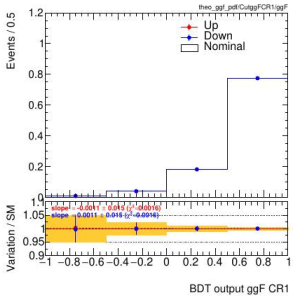


(j) QCD scale uncertainty (vbf3j)

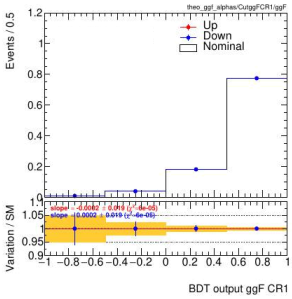


(k) QCD scale uncertainty (ggH)

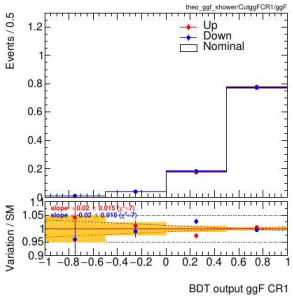
Figure 14.4. Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the ggF spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$.



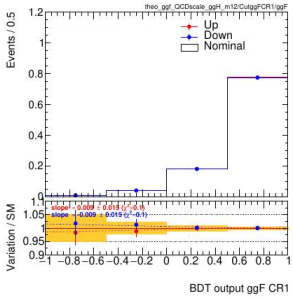
(a) PDF uncertainty



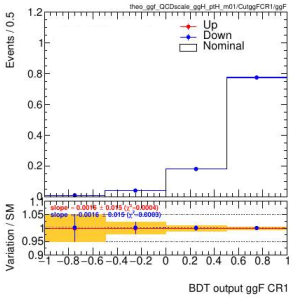
(b) α_s uncertainty



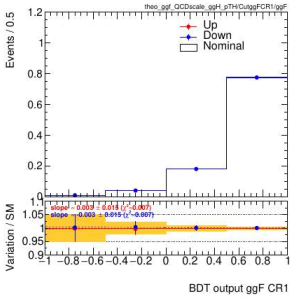
(c) Shower uncertainty



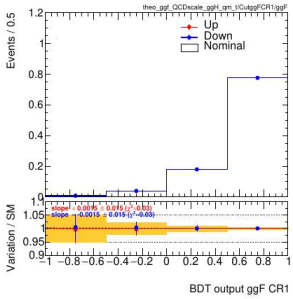
(d) QCD scale uncertainty (m12)



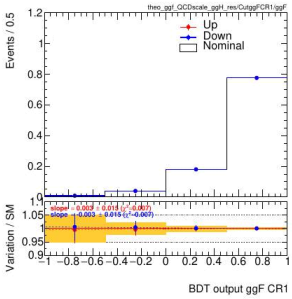
(e) QCD scale uncertainty (ptH_m01)



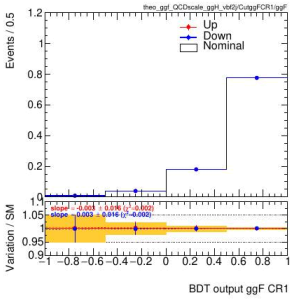
(f) QCD scale uncertainty (pTH)



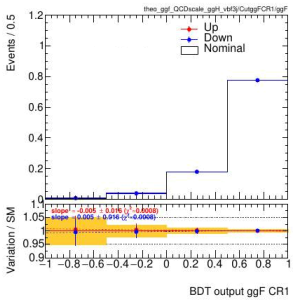
(g) QCD scale uncertainty (qm_t)



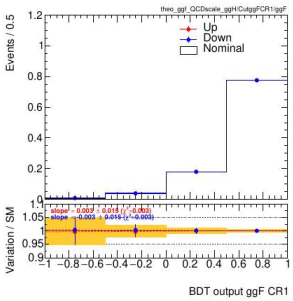
(h) QCD scale uncertainty (res)



(i) QCD scale uncertainty (vbf2j)



(j) QCD scale uncertainty (vbf3j)



(k) QCD scale uncertainty (ggH)

Figure 14.5. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the ggF spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.

top and $t\bar{t}$ samples independently and both include those from a variety of sources including hard scatter generation, parton shower, factorization and renormalization scales, initial state radiation (ISR), final state radiation (FSR), and PDFs. For parton shower uncertainties, the same scheme used in ggF and VBF samples is used (i.e. comparisons between Powheg+Pythia8 and Powheg+Herwig7). Hard scatter matching uncertainties are similarly derived from comparing the nominal generation with the one by Madgraph5 as alternative generator. QCD scale uncertainties from the 7-point envelope of renormalization and factorization variations. ISR and FSR uncertainties are determined through an up and down variation weights from Pythia8. PDF uncertainty is from the NNPDF30_nlo variations in the Powheg+Pythia8 sample. The $t\bar{t}$ sample does not include NNLO corrections since the normalization of the background is derived in the fit to data, and the application of such reweighing shows spurious effects on the top modeling, specifically in the MC-matching uncertainty, in SR and Top VR. The Wt events contain one additional systematic uncertainty - interference. This is derived from removing $t\bar{t}$ -like contributions at the level of amplitude (DR) or matrix elements (DS). Table 14.5 shows theory uncertainty from $t\bar{t}$ events in the signal, control and validation regions while Table 14.6 shows the same for Wt. While Wt uncertainties are large, the background has a minimal contribution to the analysis so these do not have any large effect on overall results.

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_ttbar_scale (ttbar)	16.9 / -14.3	16.3 / -14.3	17.7 / -14.7	11.5	18.2 / -15.1	14.7 / -13.1	15.9 / -13.8	12.5 / -11.9
theo_ttbar_shower (ttbar)	-3.3 ± 3.1	0.9 ± 1.8	-3.3 ± 4.5	-5.9 ± 0.1	4.1 ± 0.4	3.8 ± 0.3	10.4 ± 0.6	2.0 ± 0.3
theo_ttbar_matching (ttbar)	14.3 ± 6.3	10.7 ± 3.5	23.9 ± 9.4	-1.4 ± 0.2	-0.2 ± 0.8	9.1 ± 0.6	5.8 ± 1.2	9.6 ± 0.6
theo_ttbar_isr (ttbar)	-0.6 / 0.5	-0.6 / 0.5	1.5 / -1.6	-0.6 / 0.5	-0.6 / 0.6	0.1 / -0.1	1.4 / -1.2	-0.6 / 0.6
theo_ttbar_fsr (ttbar)	2.2 / -3.8	0.0 / -0.8	3.8 / -13.7	1.1 / -1.7	-1.9 / 3.5	-3.8 / 5.6	-5.2 / 7.7	-5.1 / 8.1
theo_ttbar_pdf (ttbar)	2.3	2.1	2.2	1.5	1.7	1.7	1.7	1.5

Table 14.5. top theory uncertainties impact on $t\bar{t}$ yields in the signal,control and validation regions

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_wt_scale (wt)	5.4 / -5.2	6.7 / -6.1	8.0 / -6.9	5.2 / -5.0	7.3 / -6.4	5.8 / -5.4	6.1 / -5.6	4.4 / -4.5
theo_wt_shower (wt)	-2.6 ± 14.1	12.6 ± 9.3	-27.4 ± 16.2	4.6 ± 1.0	27.7 ± 2.2	30.7 ± 2.0	50.6 ± 4.9	18.5 ± 1.1
theo_wt_matching (wt)	-7.5 ± 20.0	20.0 ± 12.9	-26.0 ± 22.6	18.0 ± 1.4	19.9 ± 2.8	18.9 ± 2.5	18.4 ± 5.6	18.6 ± 1.4
theo_wt_isr (wt)	-1.7 / 3.4	1.1 / -1.3	-1.1 / 0.4	-0.5 / 0.4	-1.2 / 0.9	-0.9 / 1.0	-0.7 / 0.4	-1.0 / 0.8
theo_wt_fsr (wt)	-5.0 / -15.1	2.0 / 5.3	8.2 / -18.0	2.0 / -4.1	-1.0 / -0.4	-3.0 / 4.4	-2.8 / 6.1	-3.1 / 5.2
theo_wt_interference (wt)	13.2 ± 14.9	-0.8 ± 7.4	39.1 ± 30.0	3.3 ± 0.9	-10.0 ± 1.5	-2.3 ± 1.5	-3.1 ± 3.0	6.3 ± 0.9
theo_wt_pdf (wt)	2.3	2.2	2.5	1.9	2.0	1.9	1.9	1.9

Table 14.6. top theory uncertainties impact on single top yields in the signal, control and validation regions

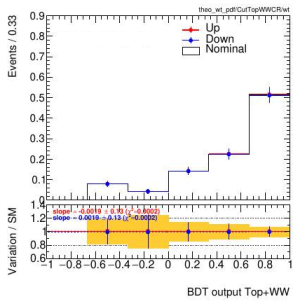
These uncertainties are examined in Fig. 14.6 for any shape effects on the BDT output used in the final fit. The plots show very little statistics to discern any shape effects. Figure 14.7 shows very little shape effects from theoretical uncertainties on the far more numerous $t\bar{t}$ backgrounds. Figure 14.8 shows some shape effects from theoretical uncertainties in the part of the Signal Region not very populated by $t\bar{t}$ backgrounds.

14.2.2 WW Uncertainties

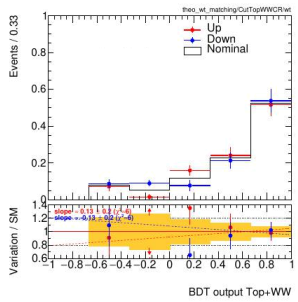
14.2.2.1 Non EWK component

Some of the WW background theoretical uncertainties are derived from reco-level samples: PDF, α_s , and QCD scale variations; others are derived from truth-level samples: merging, resummation, parton shower and recoil scheme uncertainties. In the latter case, the samples are only available at truth-level, so the full event selections was replicated at truth-level. For the most part, this is accomplished by replacing the variables used in the event selection with their truth-level equivalents. PDF and α_s uncertainties are derived from the NNPDF30_nnlo MC replica sets and QCD uncertainties from the 7-point scale variations envelope.

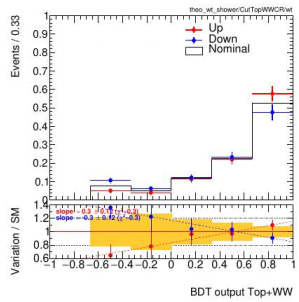
From the truth-level samples, parton shower uncertainty is derived from a comparison between Powheg+Pythia8 and Powheg+Herwig++. Uncertainty from merging scale is calculated from samples with CKKW merging scale at 15 GeV and 30 GeV,



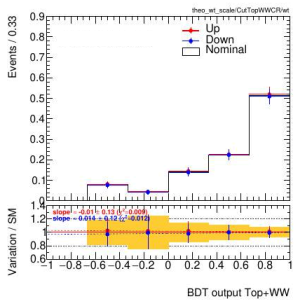
(a) PDF uncertainty



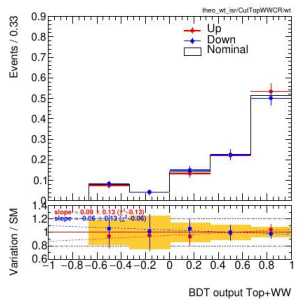
(b) Matching uncertainty



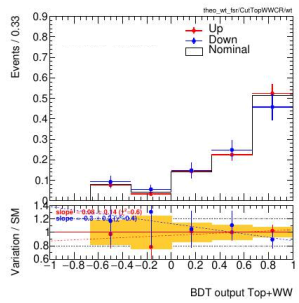
(c) Shower uncertainty



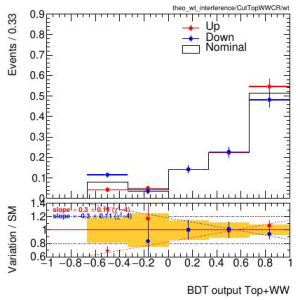
(d) QCD scale uncertainty



(e) ISR uncertainty

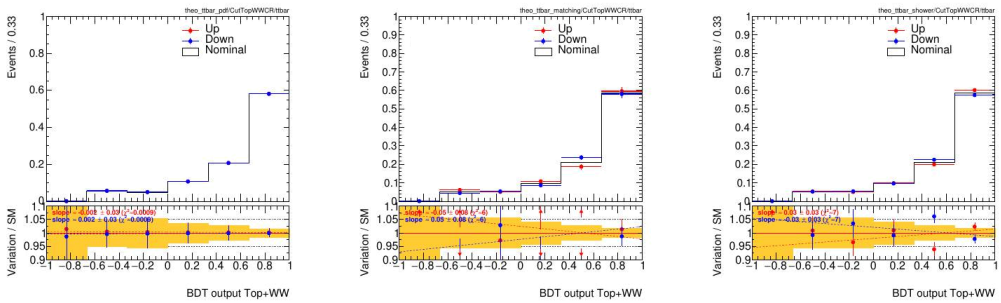


(f) FSR uncertainty

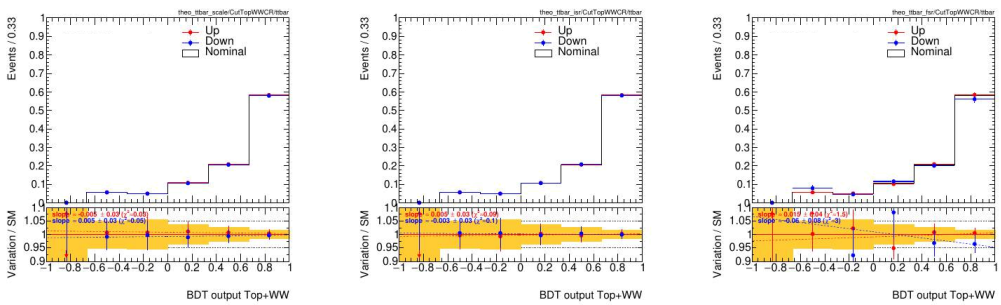


(g) Interference uncertainty

Figure 14.6. Up and down variations on the topWW bdt output distribution for the considered theoretical uncertainties on the single top spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.

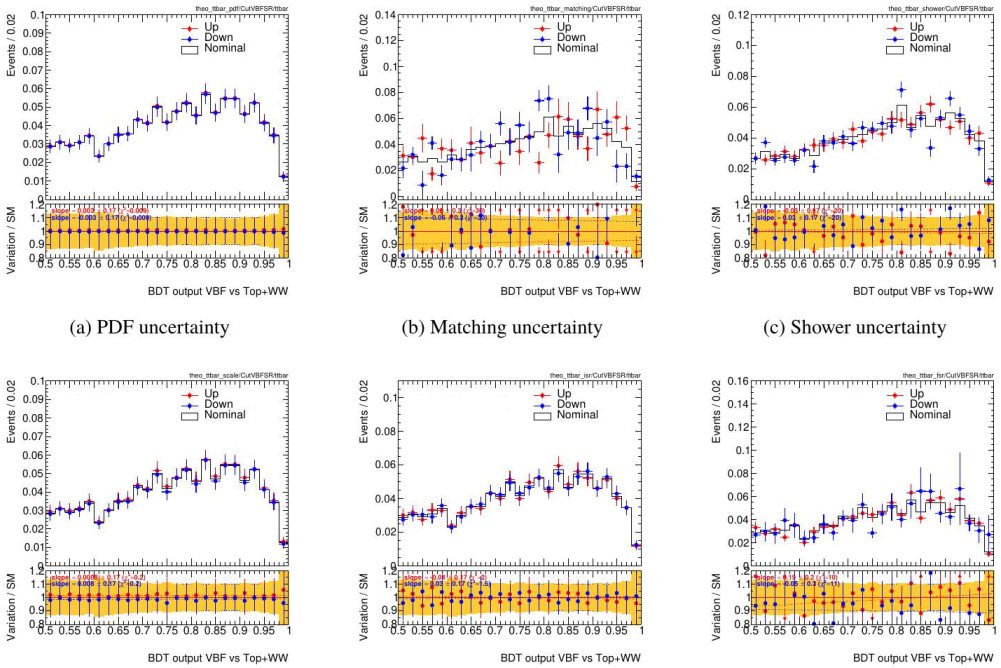


(a) PDF uncertainty (b) Matching uncertainty (c) Shower uncertainty



(d) QCD scale uncertainty (e) ISR uncertainty (f) FSR uncertainty

Figure 14.7. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the $t\bar{t}$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.



(a) PDF uncertainty (b) Matching uncertainty (c) Shower uncertainty

(d) QCD scale uncertainty (e) ISR uncertainty (f) FSR uncertainty

Figure 14.8. Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the $t\bar{t}$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$.

and resummation uncertainty from samples with resummation scale of parton shower multiplied by 0.5 and 2.0. Alternative parton shower recoil scheme (CSSKIN) is used to generate the recoil scheme uncertainty. The WW theory systematics within the signal region, control/validation regions are shown in Table 14.7 and 14.8.These uncertainties are examined in Figs. 14.9 and 14.10 for any shape effects on the BDT output used in the final fit. The plots show no shape effects within MC stat uncertainty.

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_ww_scale (WW)	20.7 / -14.8	21.9 / -15.3	22.8 / -15.6	26.6 / -19.2	26.6 / -18.9	26.8 / -19.1	30.8 / -21.0	4.5 / -4.6
theo_ww_pdf (WW)	1.0	0.9	1.1	1.0	0.9	1.0	0.9	1.7
theo_ww_alphas (WW)	1.9	2.0	1.9	2.3	2.3	2.3	2.4	1.2

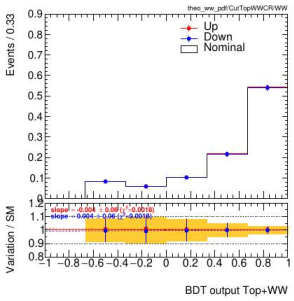
Table 14.7. Breakdown of WW theory uncertainties derived from reco-level samples and their impact on the WW yields in SR and CRs/VRs.

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_ww_CKKW (WW_Truth)	-10.8 ± 5.2	-6.0 ± 3.4	9.7 ± 8.0	-3.7 ± 5.3	-4.1 ± 0.7	-4.1 ± 0.5	-7.5 ± 0.8	-2.2 ± 0.3
theo_ww_QSF (WW_Truth)	-4.8 ± 5.1	-11.1 ± 3.1	-5.5 ± 6.8	-7.3 ± 4.8	-4.6 ± 2.1	-4.3 ± 0.4	-3.5 ± 0.8	-3.1 ± 0.4
theo_ww_shower (WW_Truth)	0.2 ± 2.9	3.7 ± 2.2	-1.0 ± 4.4	-3.6 ± 4.3	-7.0 ± 0.6	8.1 ± 0.4	25.8 ± 1.2	-2.4 ± 0.1
theo_ww_CSSKIN (WW_Truth)	3.1 ± 5.1	-3.9 ± 3.1	-3.8 ± 6.4	-12.8 ± 4.6	-0.9 ± 0.7	-4.8 ± 0.4	-4.8 ± 0.8	0.8 ± 0.4

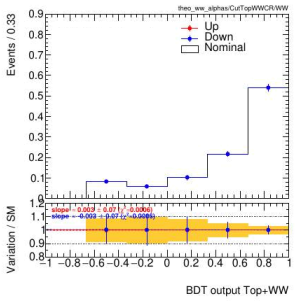
Table 14.8. Breakdown of WW theory uncertainties derived from truth-level samples and their impact on the WW yields in SR and CRs/VRs.

14.2.2.2 EWK component

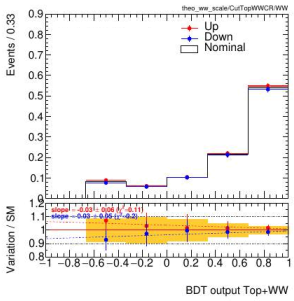
Theoretical uncertainties on electroweak WW background include PDF, α_s, and QCD scale variation. Additionally, the parton shower uncertainty is derived from a comparison between Madgraph5+Pythia8 and Madgraph5+Herwig7. The effect of PDF uncertainty was evaluated using the standard deviation of 100 NNPDF replicas around the nominal. Similarly, the effect of α_s was evaluated using three replicas with different settings of α_s . A normalization uncertainty of $\pm 15\%$ is applied in all regions as the EW scale uncertainty. The main effect of missing EW scales is expected



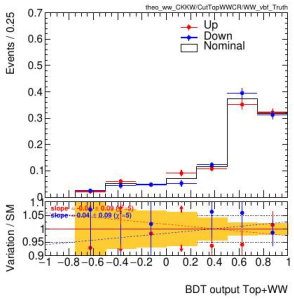
(a) PDF uncertainty



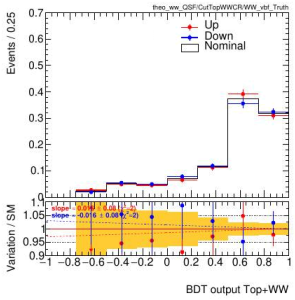
(b) α_s uncertainty



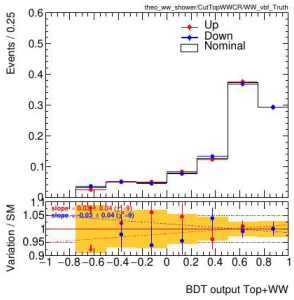
(c) QCD scale uncertainty



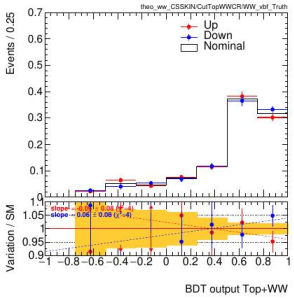
(d) Merging uncertainty



(e) Resummation uncertainty

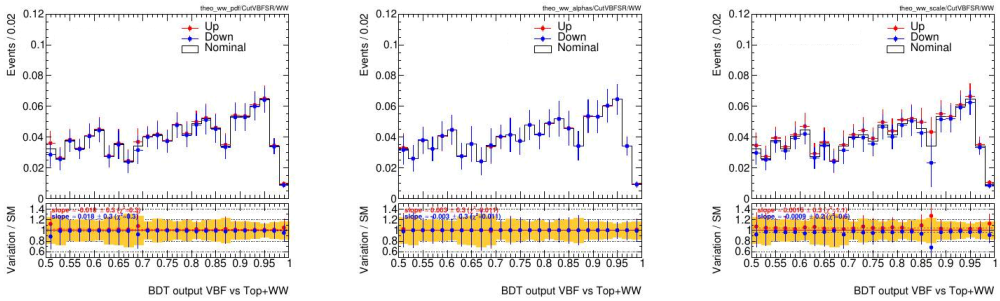


(f) Shower uncertainty



(g) PS recoil uncertainty

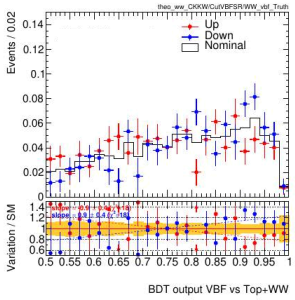
Figure 14.9. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the WW spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.



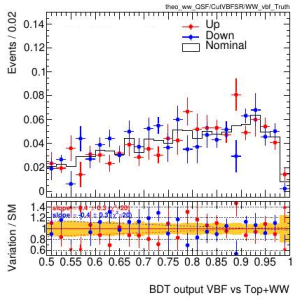
(a) PDF uncertainty

(b) α_s uncertainty

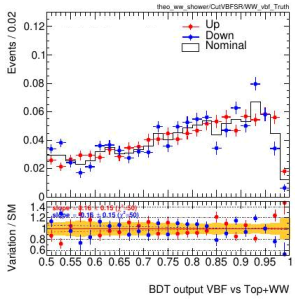
(c) QCD scale uncertainty



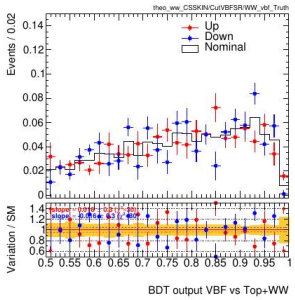
(d) Merging uncertainty



(e) Resummation uncertainty



(f) Shower uncertainty



(g) PS recoil uncertainty

Figure 14.10. Up and down variations on the VBF BDT output distribution for the considered theoretical uncertainties on the WW spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF > 0.5$.

to come from the NLO EW correction. This correction has not yet been computed for the opposite-sign EW WW process. However, it stems from logarithmic corrections that are very consistent across VB samples. As such, the so-called "leading-log approximation" can be used to derive an estimate of its size. This approximation has been used in previous analysis, for which the full NLO EW corrections have also been derived. It generally gives a result that agrees within 1-2% with the full calculation. More details of the calculation are given in the reference. The EW WW theory systematics within the signal region, control/validation regions are shown in Table 14.9.

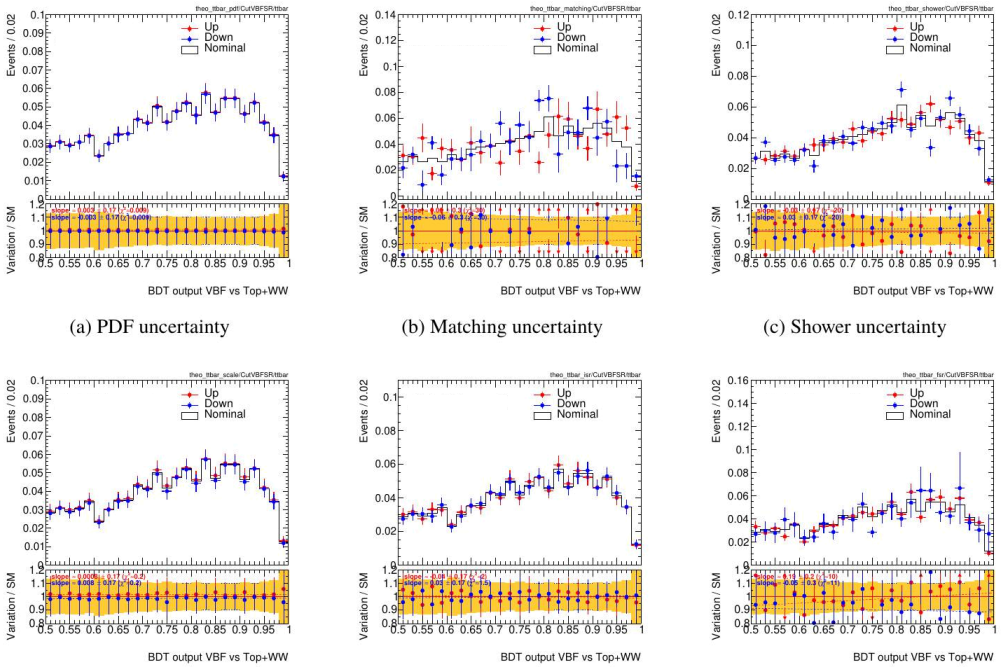
These uncertainties are examined in Figs. 14.11, 14.13, and. 14.13 for any shape effects on the BDT output used in the final fit. The plots show no significant shape effects within MC stat uncertainty.

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_ewww_scale (EWWW)	7.4 / -6.5	7.4 / -6.5	7.0 / -6.2	3.3 / -3.4	7.4 / -6.5	3.8 / -3.6	3.3 / -3.2	1.5 / -1.8
theo_ewww_pdf (EWWW)	1.5	1.5	1.7	1.2	1.6	1.4	1.5	1.4
theo_ewww_alphas (EWWW)	0.1	0.1	0.1	0.4	0.0	0.1	0.1	0.3
theo_ewww_shower (EWWW)	4.4 ± 2.0	6.6 ± 1.4	4.9 ± 3.6	-0.9 ± 2.1	3.2 ± 0.5	-2.6 ± 0.9	-7.1 ± 2.1	1.5 ± 1.4

Table 14.9. Breakdown of EWK WW theory uncertainties their impact on the WW yields in SR and CRs/VRs.

14.2.3 $Z \to \tau\tau$ Uncertainties

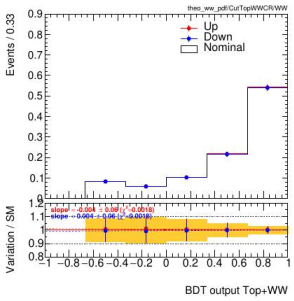
The $Z \to \tau\tau$ uncertainties include PDF, α_s , QCD scale, and generator variations. PDF and α_s uncertainties are calculated from internal re-weighting within the nominal Sherpa samples using the NNPDF30_nnlo PDF sets. QCD uncertainties include the 7-point variation of renormalisation and factorization scales. Generator uncertainties are derived from a comparison of the baseline generator Sherpa with Madgraph5, both interfaced with Pythia8. The Madgraph5 samples suffer from low statistics manifested as large shape uncertainties. The $Z \to \tau\tau$ theory systematics within the signal region, control/validation regions are shown in Table 14.10. These uncertainties



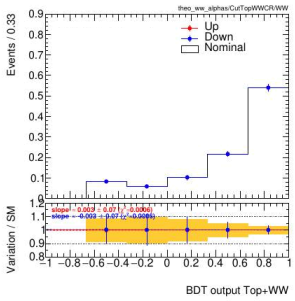
(a) PDF uncertainty (b) Matching uncertainty (c) Shower uncertainty

(d) QCD scale uncertainty (e) ISR uncertainty (f) FSR uncertainty

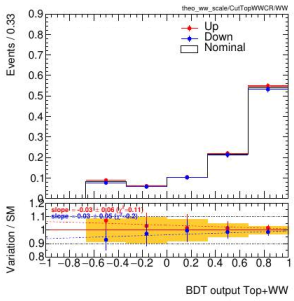
Figure 14.11. Up and down variations on the top+WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.



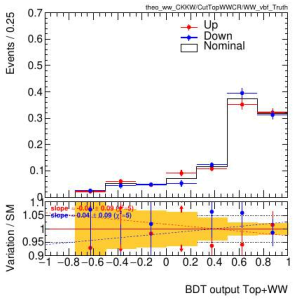
(a) PDF uncertainty



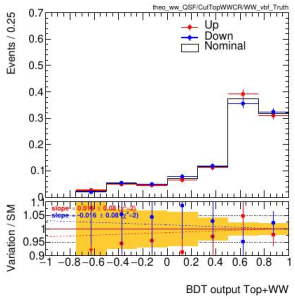
(b) α_s uncertainty



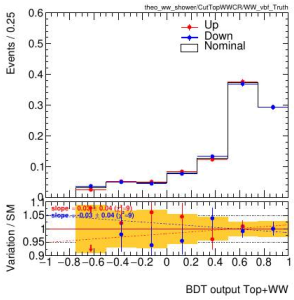
(c) QCD scale uncertainty



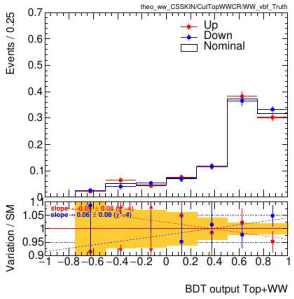
(d) Merging uncertainty



(e) Resummation uncertainty

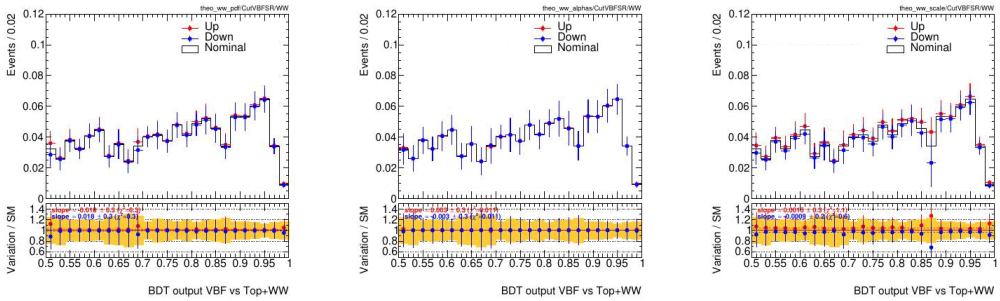


(f) Shower uncertainty



(g) PS recoil uncertainty

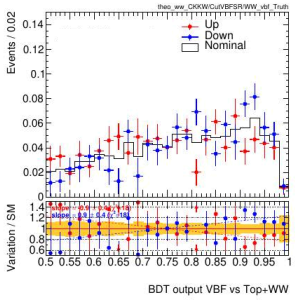
Figure 14.12. Up and down variations on the top+WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.



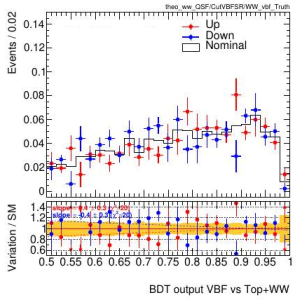
(a) PDF uncertainty

(b) α_s uncertainty

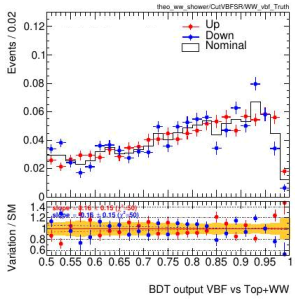
(c) QCD scale uncertainty



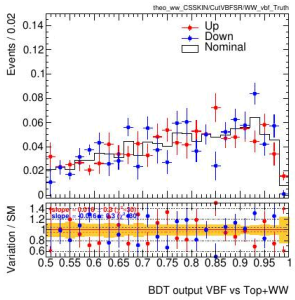
(d) Merging uncertainty



(e) Resummation uncertainty



(f) Shower uncertainty



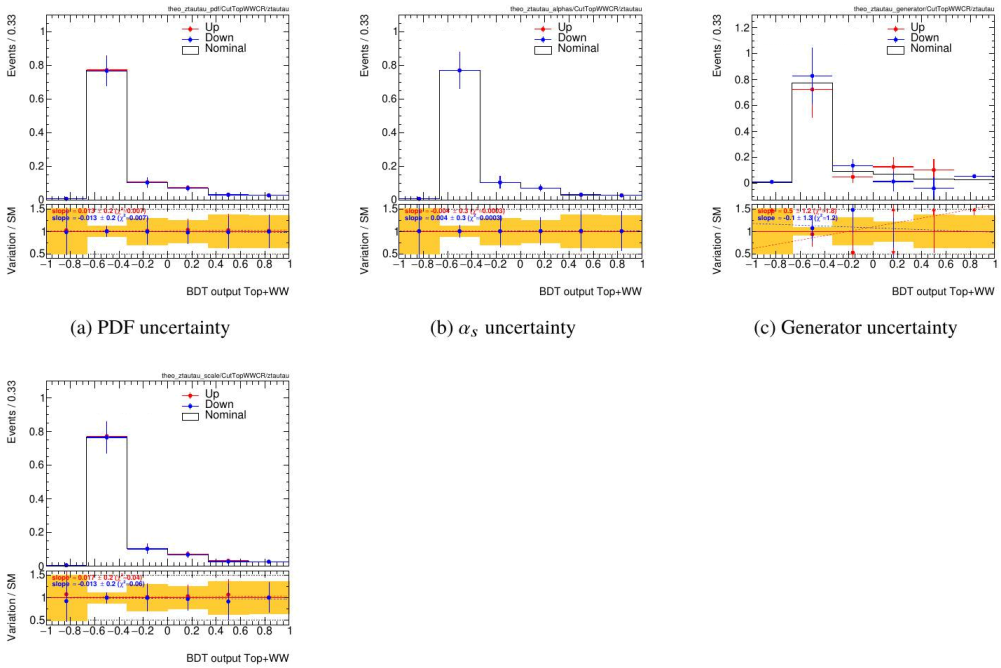
(g) PS recoil uncertainty

Figure 14.13. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the diboson spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.

are examined in Figs. 14.14 and. 14.15 for any shape effects on the BDT output used in the final fit.

Fit regions	Impact high / low (+/- stat) [%]							
	SR	TopWW_CR	Zjets_CR	Top_VR	WW_VR	ggF_CR1	ggF_CR2	ggF_CR3
theo_ztautau_scale (ztautau)	19.9 / -14.2	19.8 / -12.8	12.7 / -8.9	8.5 / -11.4	4.6 / -5.6	23.8 / -15.8	32.8 / -20.8	7.7 / -7.0
theo_ztautau_pdf (ztautau)	1.0	0.9	1.1	1.6	1.2	0.9	0.9	1.4
theo_ztautau_alphas (ztautau)	1.9	1.8	1.6	2.6	1.6	2.3	2.6	1.5
theo_ztautau_generator (ztautau)	37.5 ± 26.7	50.4 ± 39.6	49.7 ± 12.2	24.3 ± 13.3	-3.6 ± 14.3	8.8 ± 4.0	-24.8 ± 5.5	-7.9 ± 3.9

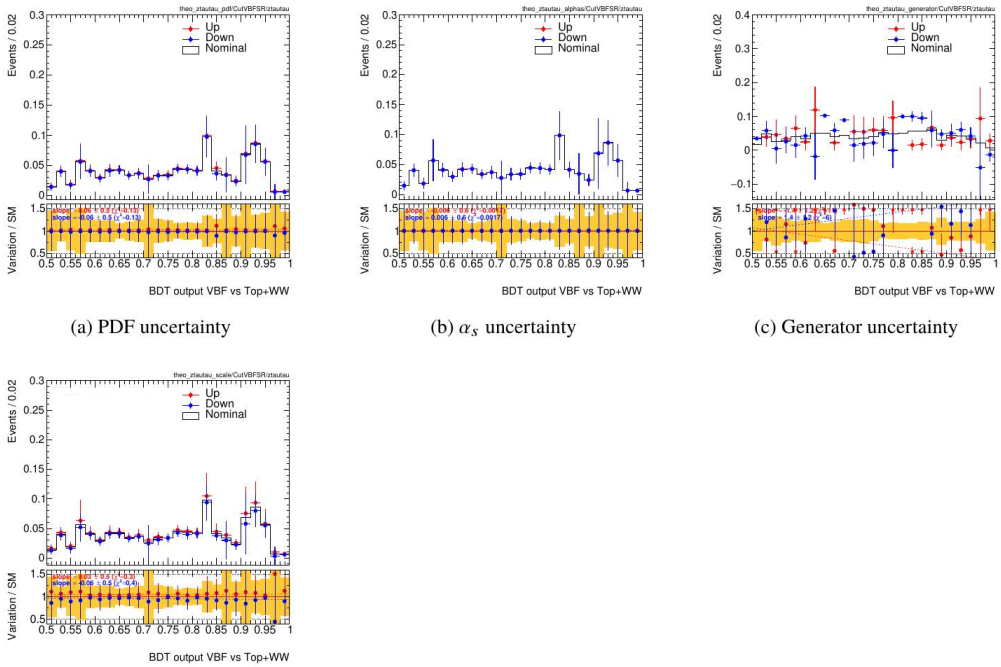
Table 14.10. Breakdown of $Z \to \tau\tau$ theory uncertainties their impact on the Zjets yields in SR and CRs/VRs.



(a) PDF uncertainty (b) α_s uncertainty (c) Generator uncertainty

(d) QCD scale uncertainty

Figure 14.14. Up and down variations on the top +WW vs All BDT output distribution for the considered theoretical uncertainties on the $Z \to \tau\tau$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown after all signal region cuts and $DVBF < 0.5$.



(a) PDF uncertainty

(b) α_s uncertainty

(c) Generator uncertainty

(d) QCD scale uncertainty

Figure 14.15. Up and down variations on the ggF BDT output distribution for the considered theoretical uncertainties on the $Z \to \tau\tau$ spectrum. The nominal sample is shown in black and slopes are calculated for up and down variations to display any potential linear effect on shapes. The yellow band in the ratio plot shows the MC stat uncertainty of the nominal sample. Distributions are shown in the ggF control region.

CHAPTER 15

UNFOLDING

15.1 Introduction

The unfolding process corrects for detector effect such as inefficiencies and limited resolution in the distribution of kinematic variables. The corrected cross section measurements refer to the fiducial kinematic region, which corresponds to the same phase space of the detector level distributions, both for the jets and the boson decay products.

The most straightforward way to unfold the detector distribution is by inverting the migration matrix. That is, if the reconstruction level bins are represented by vector $\vec{R}$ and truth level bins are represented by vector $\vec{T}$, then a response matrix M is calculated from the Monte Carlo simulations such that

$$M\vec{T} = \vec{R} \tag{15.1}$$

And we simply obtain the truth level through matrix inversion:

$$\vec{T} = M^{-1}(\vec{v} - \vec{b}) \tag{15.2}$$

Where $\vec{v}$ is the observed data, and $\vec{b}$ is the estimated background, and so $\vec{R} = \vec{v} - \vec{b}$.

The challenge comes from the possibility that the response matrix M is ill-conditioned, which means small fluctuations due to statistical errors in the input vector $\vec{R}$ will be amplified and cause large unphysical fluctuations in the solution for $\vec{T}$. To correct for

that possibility, the unfolding procedure may require regularization. Two main methods of regularization are presented here; iterative bayesian unfolding and Tikhonov regularization. It was ultimately decided that the response matrices for all our obersevables were sufficiently well conditioned that regularization was not necessary. The final unfolding method, the matrix likelihood procedure, is discussed in the next chapter.

15.2 Iterative Bayesian unfolding

In bayesian iteration, an apriori assumption that the final distribution behaves like the simulation is used as a basis [8]. From there we want to determine the probability that a bin i in the detector level distribution would lead to bin j in the true distribution. According to Bayes theorem, the probability of an event ending up in bin i from the truth-level distribution, while being in bin j in the detector-level distribution, is given by:

$$P(T_i|R_j) = \frac{P(R_j|T_i) \cdot P_0(T_i)}{\sum_k P(R_j|T_k) \cdot P_0(T_k)} \tag{15.3}$$

The total bin content,n , of truth bin i is then given by:

$$n(T_i) = \frac{1}{\epsilon_i} \sum_j n(R_j) \cdot P(T_i|R_j) \tag{15.4}$$

Where $n(R_J)$ is the total content of bin j in the detector-level distribution and ϵ is the efficiency of the truth bin i. Efficiency is the ratio of events that pass both truth-level and detector-level selections divided by number of events that only pass the truth selection.

For the algorithm we produce the response matrix M_{ij} from equation (15.3):

$$M_{ij} = \frac{P(R_j|T_i) \cdot P_0(T_i)}{\sum_k P(R_j|T_k) \cdot P_0(T_k) \cdot \sum_l^{n(R)} P(R_l|T_i)} \tag{15.5}$$

where the last term $\sum_l^{n(R)} P(R_l|T_i)$ is the efficiency factor discussed earlier. The truth distribution is then simply given by:

$$n(T_i) = \sum_j M_{ij} \cdot n(R_j) \tag{15.6}$$

The migration matrix is trained with MC simulation with the same final cuts in the detector-level spectrum and truth-level spectrum. The more iterations the more the final unfolded distribution behaves like data and less like the MC prior.

15.2.1 Statistical Uncertainty

The statistical uncertainty on the unfolded distribution is propagated from the reconstruction level through randomly reweighing each bin in the reconstruction distribution and using the same migration matrix to solve for the reweighed truth distribution. The random variations experiments are repeated one thousand times. Each variation taken randomly from a gaussian distribution with a center of the original content of each bin and with a standard deviation equal to the statistical uncertainty of each bin. For each of the one thousand toys the difference in the weighted truth content and the unfolded distribution is one event in the histogram.

The statistical uncertainty is the variance in the gaussian histogram of each bin. Figure 15.1 shows the estimated uncertainty on each unfolded distribution propagated through three bayesian iterations.

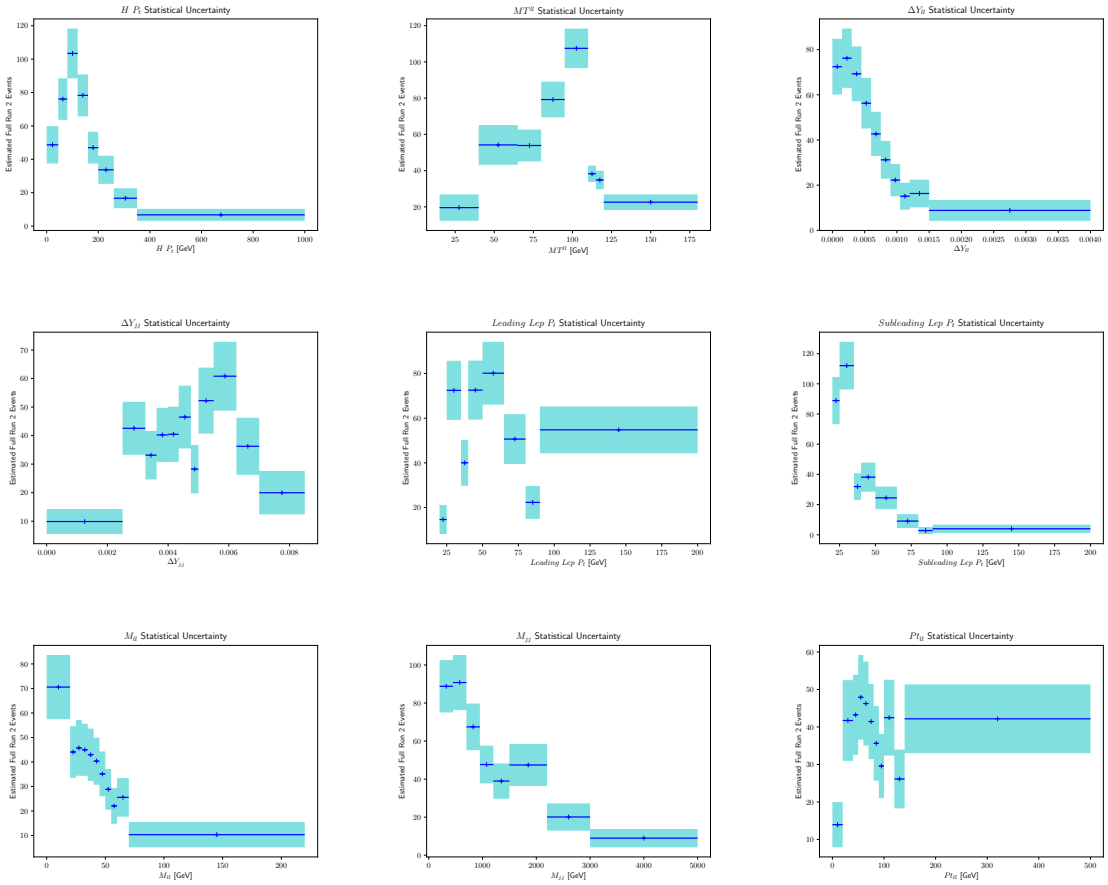


Figure 15.1. Statistical uncertainty propagated through unfolding vs. bayesian Iteration in Higgs Pt

15.2.2 Unfolding Bias

The purpose of the closure test is to check the bias of the unfolded distribution. Any regularization method trained on Monte Carlo will bias the unfolded distribution towards the Monte Carlo data, potentially hiding new physics in the real data.

To measure the bias, a new truth distribution is created by varying the bins in the original distribution by a random number drawn from a gaussian centered on the nominal truth content with a width equal to the uncertainty on the bin. The corresponding reconstruction-level distribution is estimated by folding the fluctuated truth distribution using the migration matrix trained from MC. The folded reco distribution is unfolded using iterative bayesian method. The bias is the difference between the unfolded distribution and the fluctuated truth distribution. Figure 15.2 shows the relation between the unfolded bin values on the y-axis and the fluctuated truth values on the x-axis for Higgs pt distribution.

The method of evaluating systematic uncertainties can be used to find the optimal normalization parameter for unfolding, which is the number of iterations in the Baysian algorithm in this case. Figure 15.3 shows the profiles for the 2D histograms with several iterations. The larger slope indicates a larger bias towards the Monte Carlo. More iterations reduce the maximum bias for large fluctuations.

15.2.3 Coverage

The concept of coverage was introduced by Mikael Kuusela, a measurement of how much an unfolded distribution is biased towards the truth Monte Carlo. Coverage was examined by taking the same reweighing algorithm and in addition to reweighing the truth distribution and calculating a reconstruction distribution, the folded detector-level distribution is then reweighed and unfolded. The resulting unfolded distribution minus the truth distribution gives the Gaussian distributions shown in Figure 15.4. If

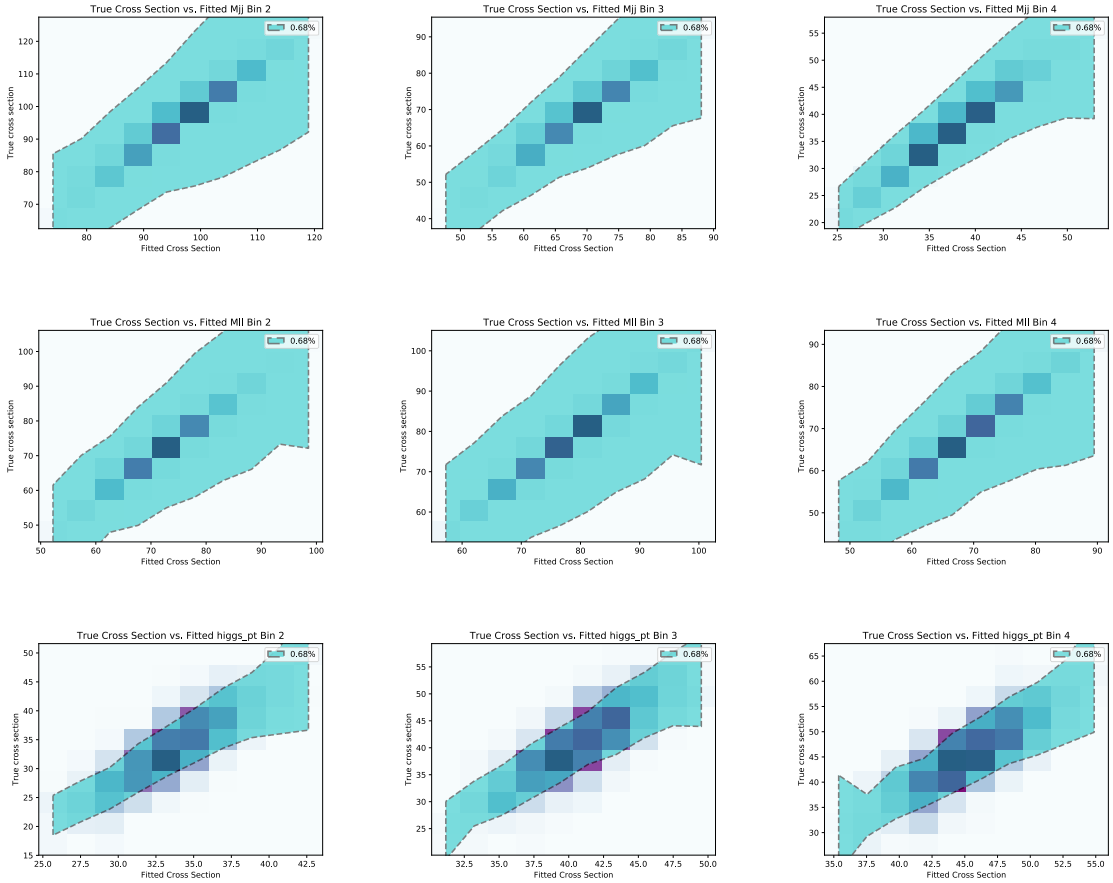


Figure 15.2. Truth Cross Section vs. Unfolded Cross Section for three Bayesian iterations. 68% of the data falls within the bands

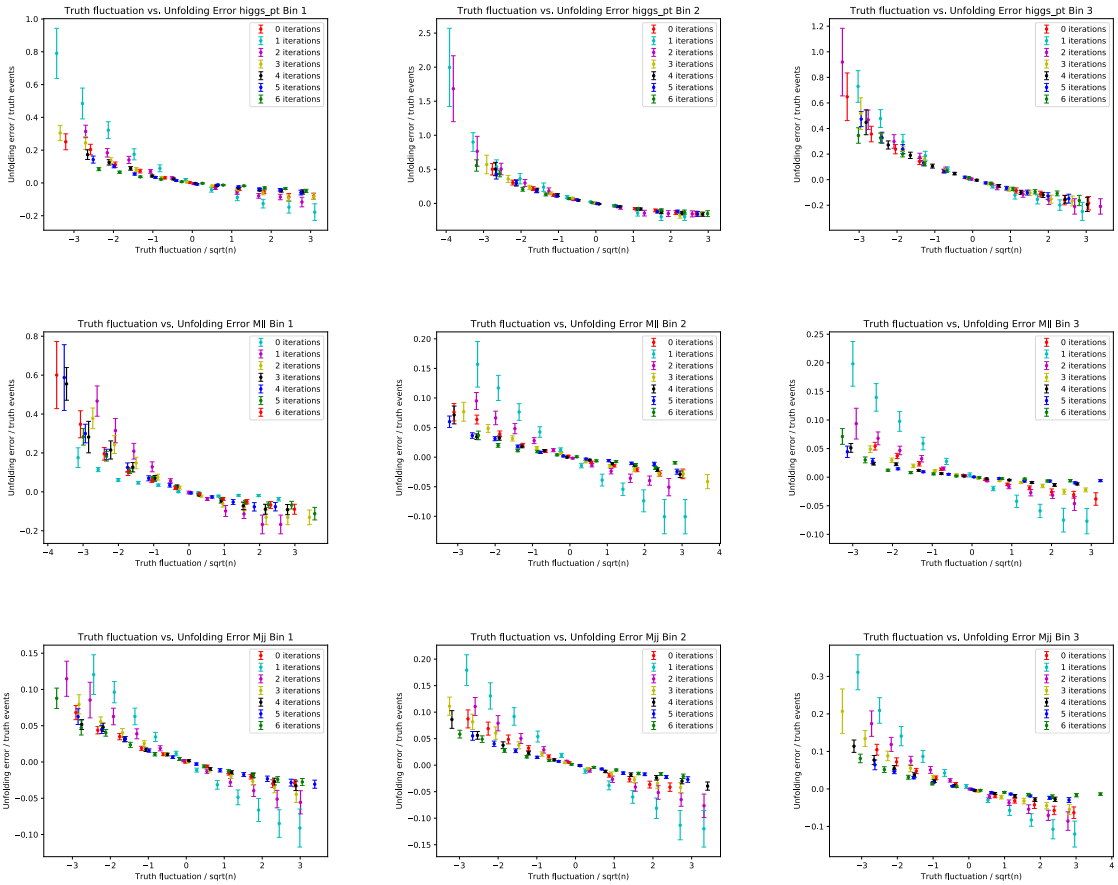


Figure 15.3. Profile of Systematic Bias vs Variation in Truth Content. The flatter the curve of the profile the less unfolding bias. The optimal parameter is taken when there's no significant bias improvements from more iterations.

the coverage is sufficient then the width of the plots divided by the expected unfolding uncertainty is equal to 1. From the plots they are indeed close to 1.

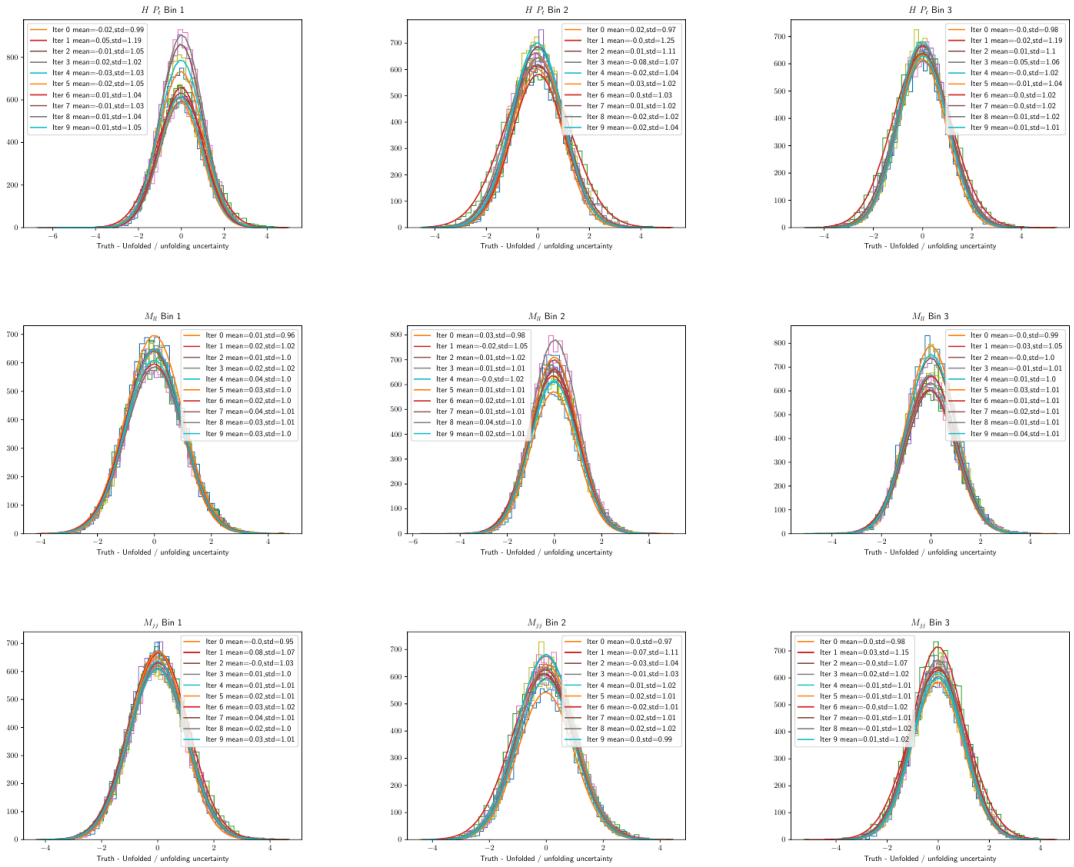


Figure 15.4. Examination of coverage in unfolding for the Baysian unfolding.

15.3 Tikhonov Regularization

The second regularization method was Tikhonov Regularization, where a constraint term is added to the likelihood. The extra term is built on the assumption that the final distribution must be smooth and follow the Monte Carlo, thus gives gives the likelihood function a "penalty" for large discontinuities in the second deriva-

tives, and a penalty for disagreement with the Monte Carlo expectation. The penalty term takes the following form:

$$\tau \sum_{i=2}^{n-1} ((\frac{\sigma_{i+1}}{\sigma_{i+1,\text{truth}}} - \frac{\sigma_i}{\sigma_{i,\text{truth}}}) - (\frac{\sigma_i}{\sigma_{i,\text{truth}}} - \frac{\sigma_{i-1}}{\sigma_{i-1,\text{truth}}}))^2 \tag{15.7}$$

Where the parameters τ is tunable by hand and controls the amount of regularization. σ_i is the cross section in bin i and $\sigma_{i,truth}$ is the expected cross section from the Monte Carlo. This term can then be used to suppress large oscillations that may come from an ill-conditioned response matrix. This method of regularization was studied with bias tests similar to the unfolding method. In the end no benefit was seen over the matrix likelihood method.

15.3.1 Bias Tests

As with the iterative bayesian tests, the statistical and systematic bias from unfolding was tested for matrix inversion, bin by bin unfolding and matrix inversion with regularization. For each toy, the truth-level cross section in each bin of the tested distribution is varied within a range corresponding to the expected statistical uncertainty of the bin. For each variation, a detector-level pseudo-dataset is obtained by folding the truth-level dataset with detector effects using the migration matrix of the corresponding distribution. All backgrounds and all regions participating in the nominal fit are included here, as if it was a fit on data.

Figure 15.6 shows the bias for each bin in M_{jj} as a function of the fluctuated truth cross section. The bias here is defined as unfolded cross section from the fit minus the true cross section. The dotted lines indicate cross sections that fall within $\pm 1\sigma$ of the nominal cross sections. As expected, matrix inversion has the smallest bias, followed by matrix inversion with regularization then bin by bin unfolding has the largest bias. For completeness, figure 15.7 shows the coverage from the three methods, over M_{jj} bins

Finally figure 15.5 shows the bias at 1σ for each bin of each observable compared with the expected statistical uncertainty. The solid black line represents the bias from matrix inversion, solid red line from bin by bin unfolding and the blue line from matrix inversion and regularization. The faded background represents the statistical uncertainty. Again, the bias from matrix inversion is the smallest of the three methods, and the statistical uncertainty from all three methods is comparable. Since the fluctuations from matrix likelihood method aren't amplified with respect to the bin by bin unfolding, the unfolding problem is well conditioned for all the unfolding observables, and so regularization is not needed.

15.4 Response Matrices

Response matrix is a plot of truth events vs. reconstruction events for a distribution. Response matrix shows how much events migrate between bins due to detector effects. Figure 15.8 shows the response matrices for kinematic distributions. The X-axis is the reconstructed distribution and the Y-axis is the truth distribution. The jet based observables tend to be less diagonal than lepton observables because of the limited jet resolution of the detector. Similarly momentum based observables are less diagonal than positional observables (based on the azimuthal angle or rapidity). The least diagonal (and least well-conditioned as a result) observable is the measurement of the Higgs traverse momentum, because of reconstruction uncertainties and the involvement of MET uncertainty in calculating the Higgs momentum from the final state particles.

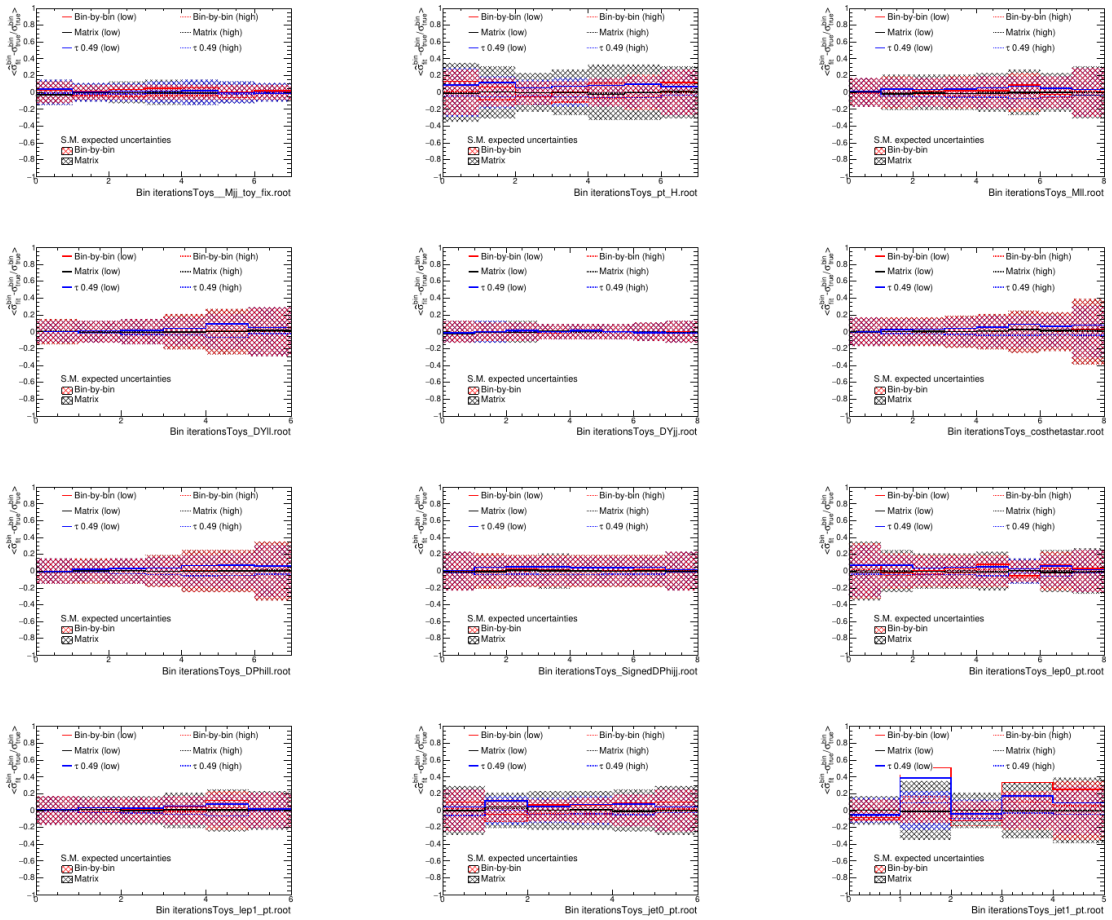


Figure 15.5. Mean cross section bias, from MC toy tests on samples with different true cross section, as a function of the several observables to be unfolded. The red and black continuous (dotted) lines show cross section bias for the bin-by-bin and matrix methods when the truth cross-section is varied within the one sigma of the expected statistical uncertainty. The red and black shaded areas show, respectively, the median statistical uncertainty.

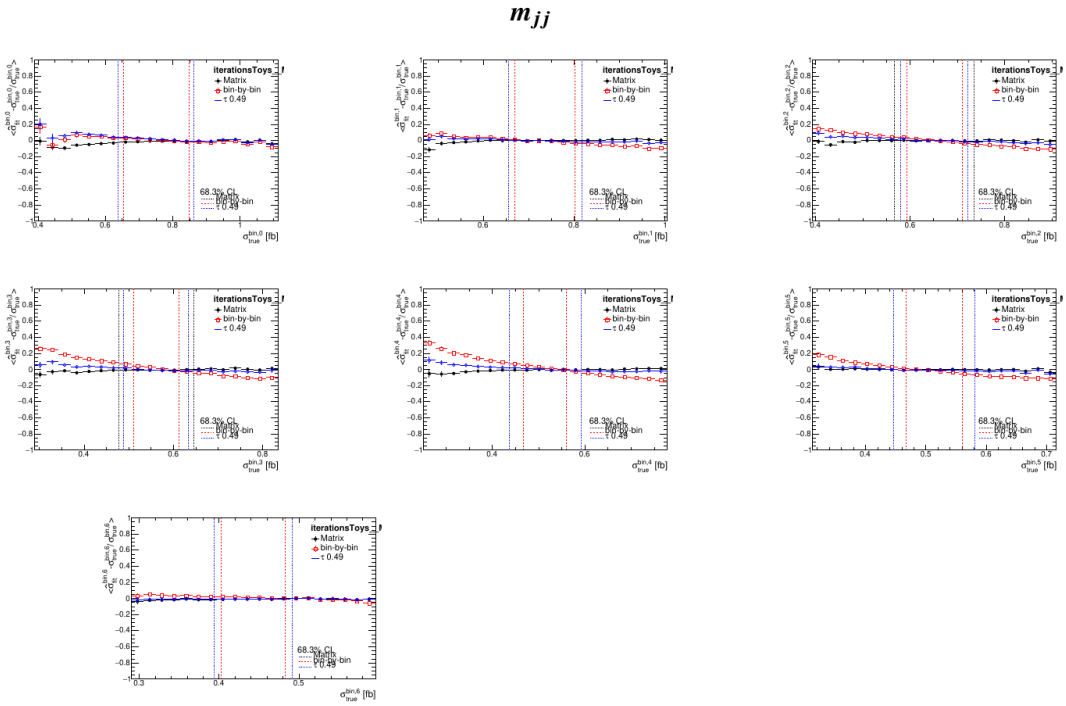


Figure 15.6. Plots of bias vs. simulated cross section for matrix inversion and Tikhonov Regularization for bins of M_{jj}

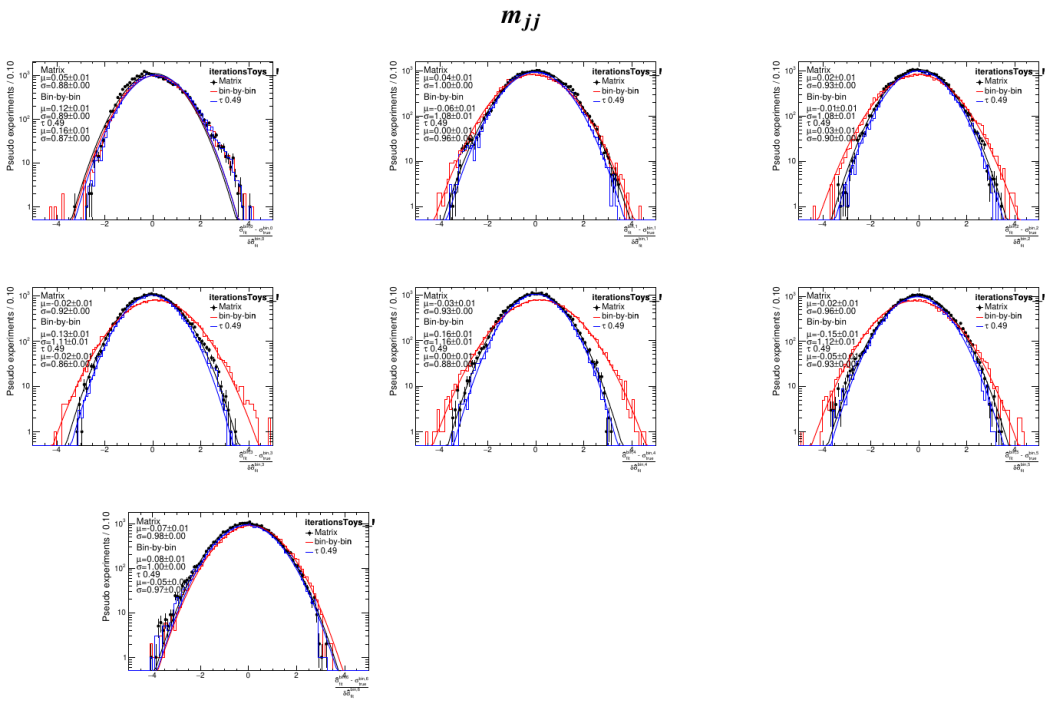


Figure 15.7. Examination of coverage for matrix inversion and Tikhonov Regularization for M_{jj}

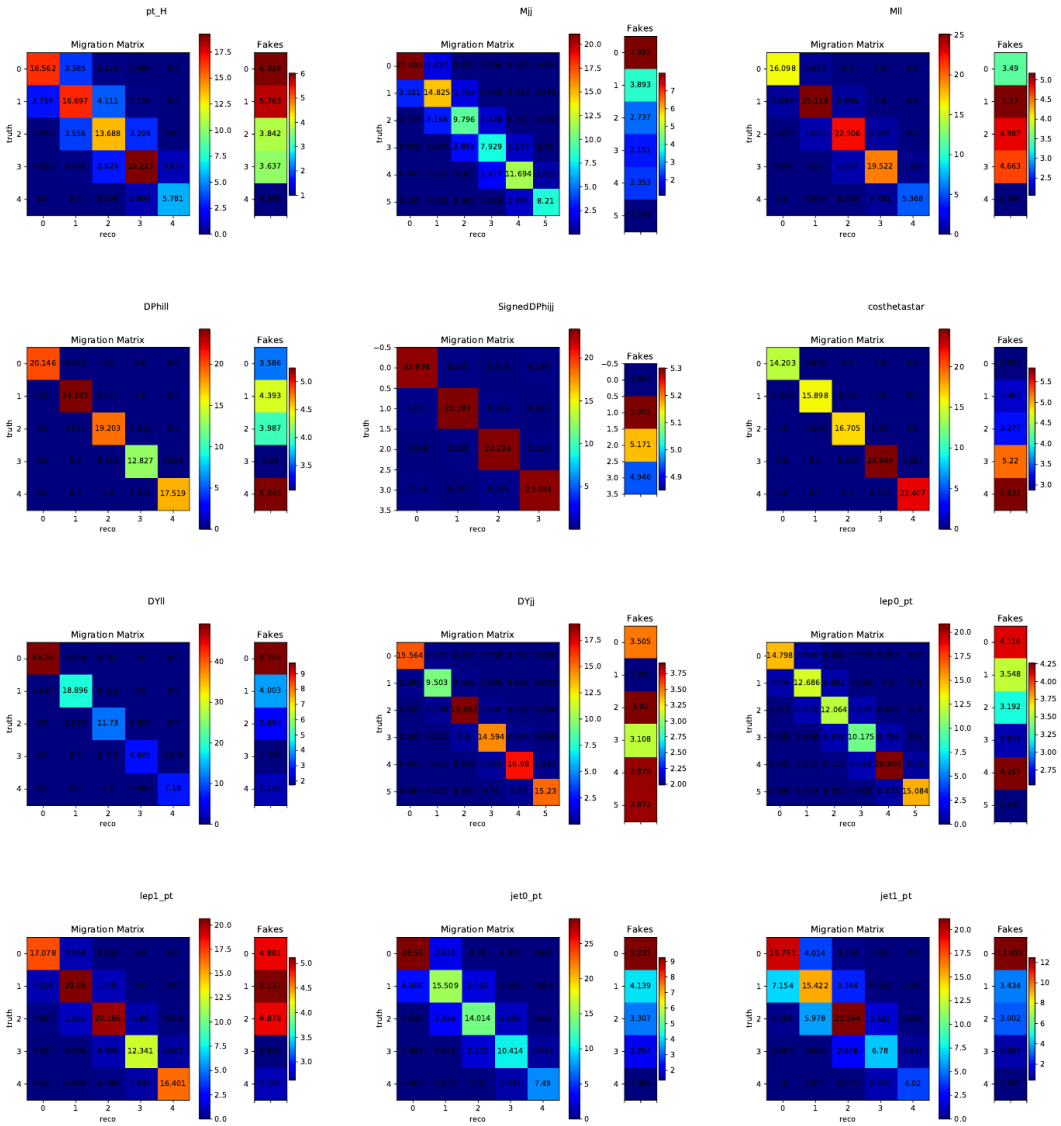


Figure 15.8. Migration matrices for each observable on the left, with the corresponding number of non-fiducial fakes on the right.

15.4.1 Efficiency and Purity

Efficiency and purity are measurements of bin-to-bin migrations between the detector-level histograms and particle-level histograms. They are given by the following formulas:

$$P_i = \frac{N_i^{\text{Truth\&Reco}}}{N_i^{\text{Reco}}} \tag{15.8}$$

$$\epsilon_i = \frac{N_i^{\text{Truth\&Reco}}}{N_i^{\text{Truth}}} \tag{15.9}$$

Where P_i is the purity in bin i in the reconstruction histogram, ϵi is the efficiency of bin i in the truth histogram. $N_i^{\text{Truth\&Reco}}$ is the number of reco events with a truth match in bin i and N_i is the total number of events in each respective truth and reco histograms.

Purity and efficiency plots of all observables are shown in Figure 15.9 for the optimized binning chosen in each distribution to be unfolded. Positional variables (Phi, rapidity) have full diagonal matrices, thus high efficiency. The variables based on measurements of Jets (Jet PT , Mjj, Higgs PT) have larger off diagonal elements thus lower efficiency. The aim of the binning choice was for each bin to have at least 60-70% efficiency.

15.5 Fakes and Misses

The so-called "fakes", as outlined in previous sections, are reconstruction-level events that fall inside of the signal region at detector-level but outside of the fiducial region at truth level, while the so-called "Misses" are truth events that fall inside of the fiducial phase space but outside the signal region at reconstruction level. Figures 15.10 and 15.11 show the fakes in each bin of the reconstruction-level distributions, while Figures 15.12 and 15.13 show the misses as a function of truth-level distributions. The total amount of fakes in reco-level data is 17%, total amount of misses is 63%.

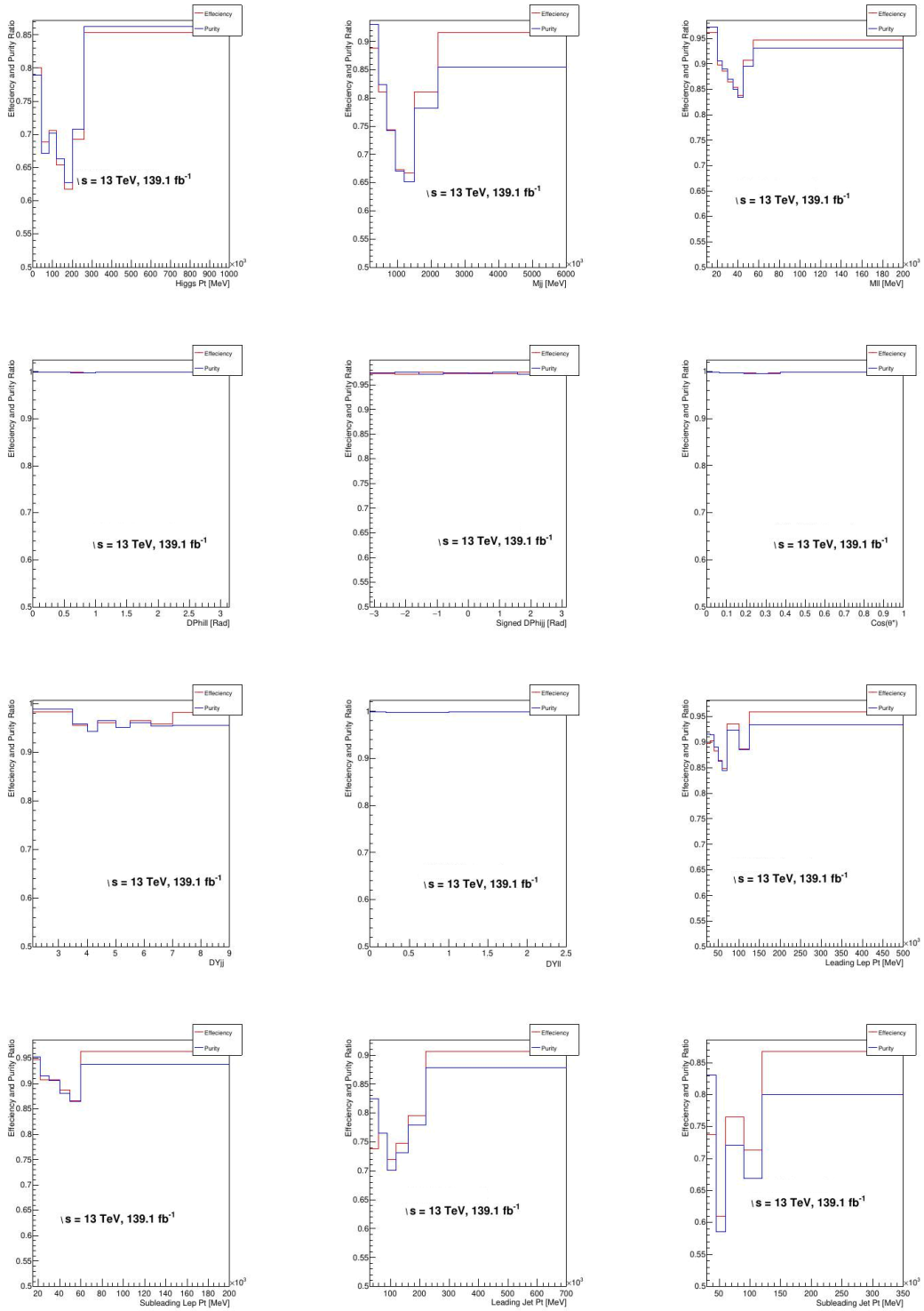


Figure 15.9. Purity and efficiency per bin for each observable to be unfolded.

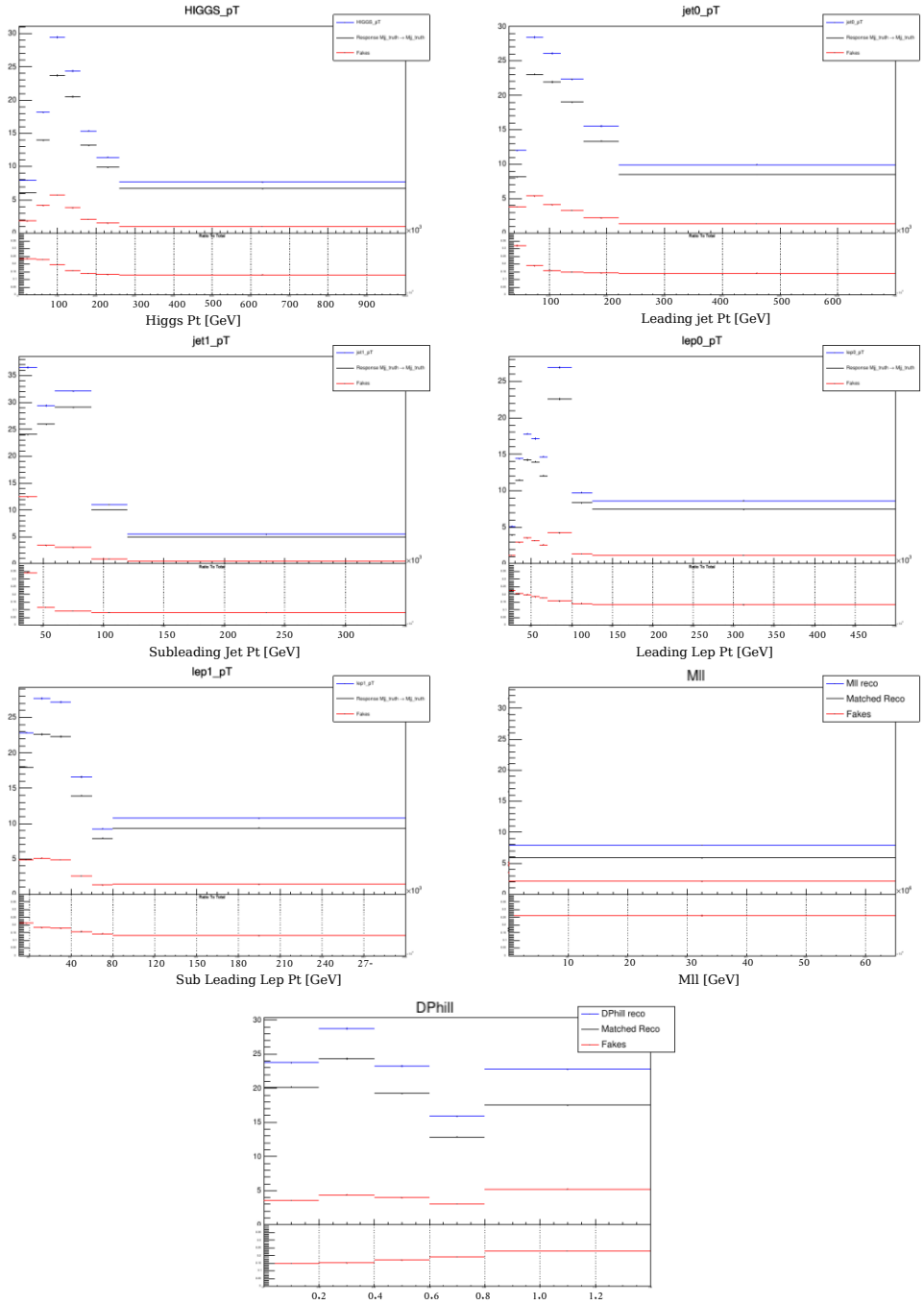


Figure 15.10. "Fake" events, defined as reco-level events in the SR falling outside of the fiducial region at truth, as a function of kinematic observables at reco-level.

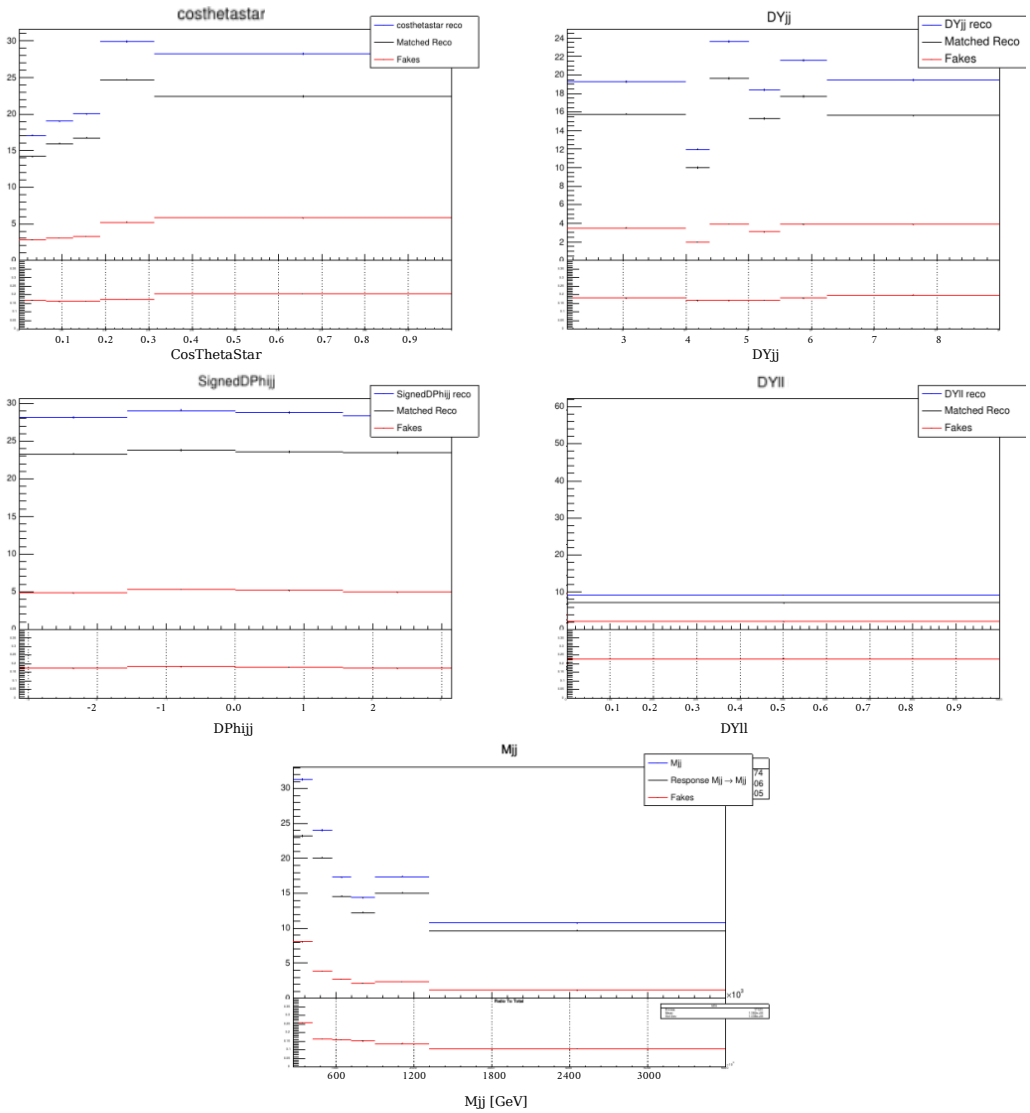


Figure 15.11. "Fake" events, defined as reco-level events in the SR falling outside of the fiducial region at truth, as a function of kinematic observables at reco-level.

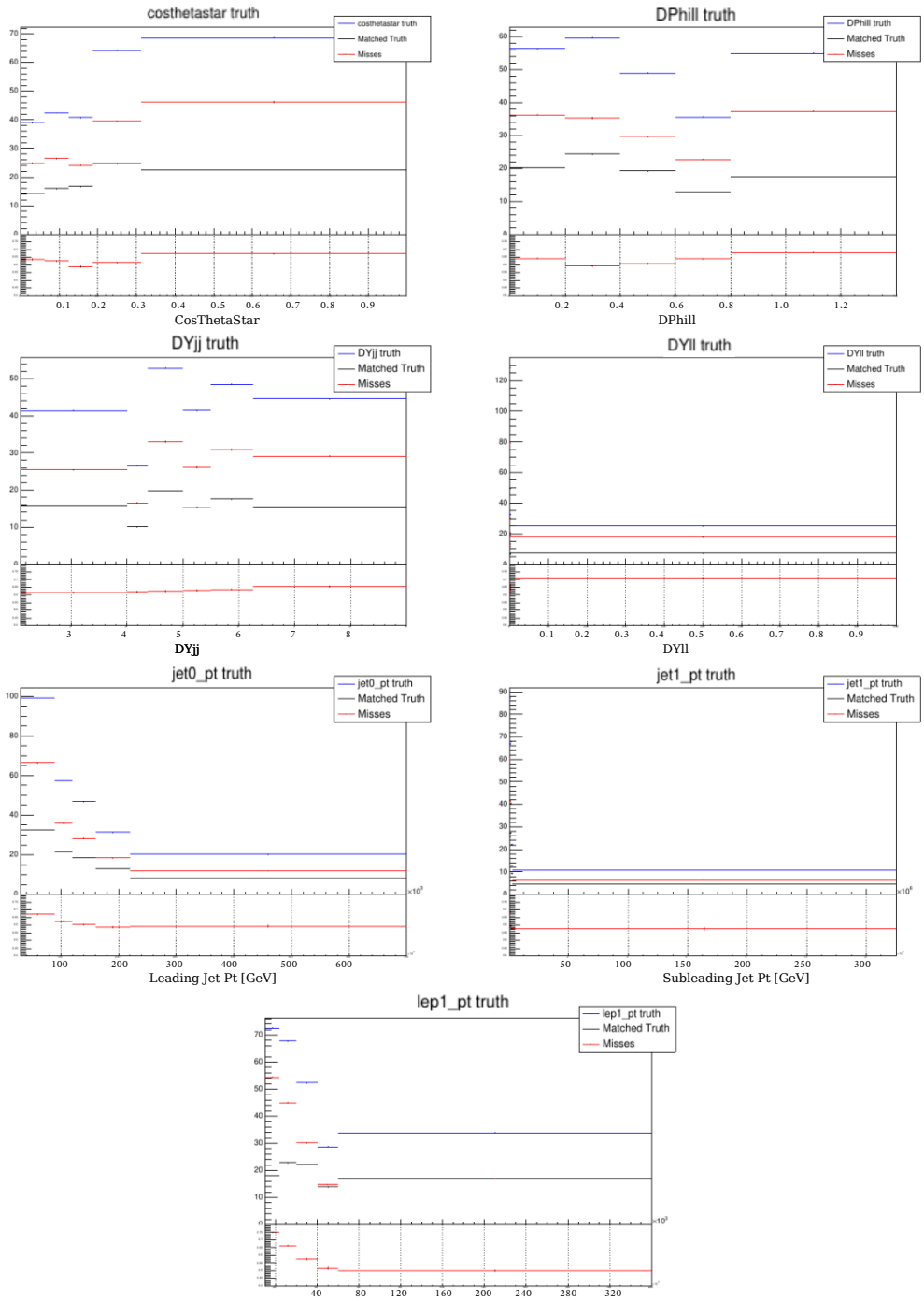


Figure 15.12. "Miss" events, defined as truth-level events in the fiducial region falling outside of the signal region at reco, as a function of kinematic observables at truth-level.

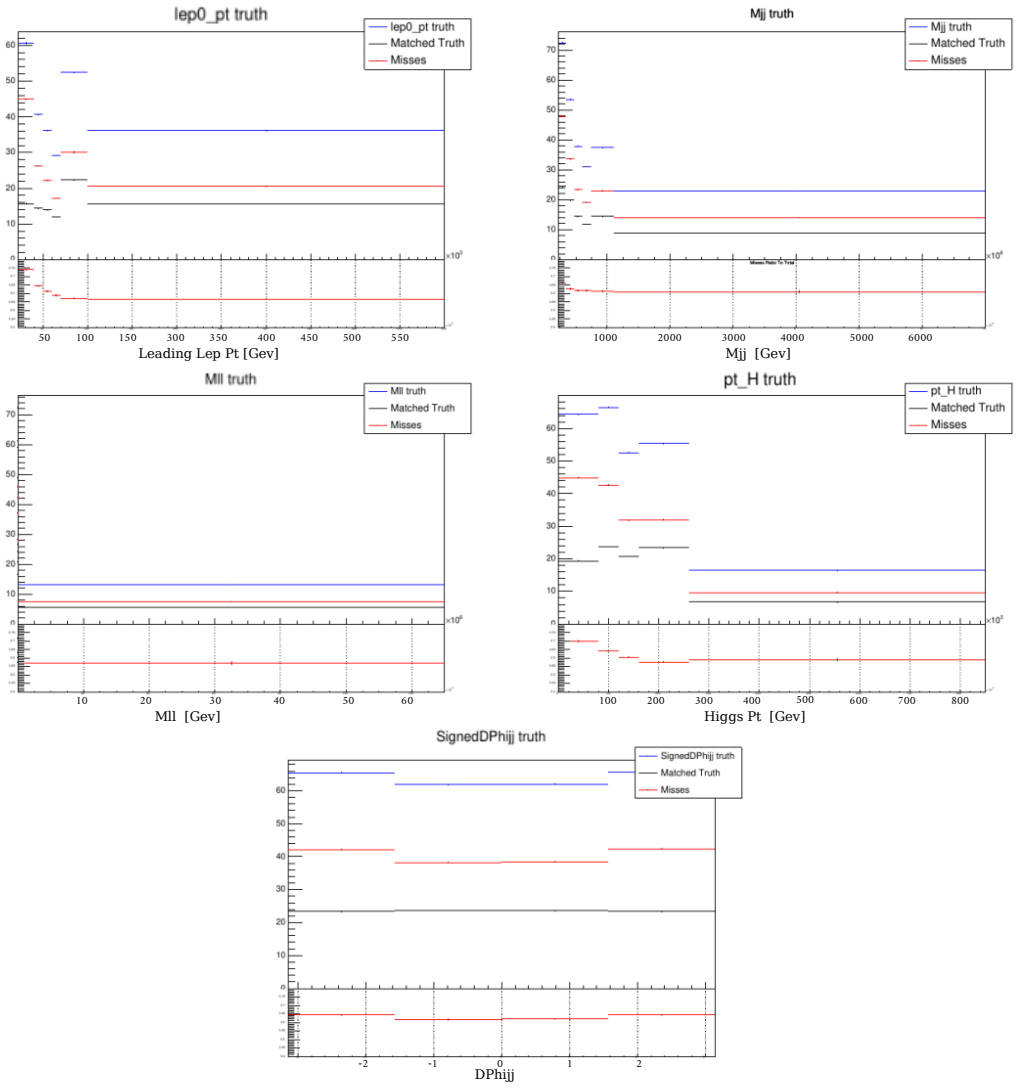


Figure 15.13. "Miss" events, defined as truth-level events in the fiducial region falling outside of the signal region at reco, as a function of kinematic observables at truth-level.

15.5.1 Observables Binning

Bin edges were optimized to find a balance between a sufficient number of bins to retain the features of the observable shapes and purity and efficiency values as high as possible. In this optimization process we strove to get purity and efficiency greater than 60%, and to ensure that at least 10 events are expected for the VBF signal in each bin. The bin edges have also been agreed with other Higgs analyses for which differential measurements are available or planned in the immediate future to allow possible combinations, such as the ggF Higgs differential analysis, the Higgs to ZZ differential analysis, as well as the SM WW cross section measurements. The differential binning choices are shown in table 15.1

Observable	Bin Edges
M_{ll}	10,30,40 55,Inclusive
M_{jj}	450,700,950,1200,1500,2200,Inclusive
$Cos(\theta^*)$	0,0.0625,0.125,0.1875,0.3125,1.0
Δy_{ll}	0,0.4,0.6,0.8,1.0,Inclusive
Δy_{jj}	2.1,4.0,4.375,5,5.5,6.25,9.0
$\Delta\phi_{ll}$	0,0.2,0.4,0.6,0.8,1.4,π
$\Delta\phi_{jj}$	$-\pi$,$-\frac{\pi}{2}$,0,$\frac{\pi}{2}$, π
Higgs p_{T}	0,80,120,160,200,260,850
$p_{\mathrm{T}}^{\ell\ell}$	0,60,80,100,140,10000
Leading Lepton p_{T}	22,40,50,60,70,100,500
Subleading Lepton p_{T}	15,22,30,40,50,200
Leading Jet p_{T}	30,90,120,160,220,700
Subleading Jet p_{T}	30,45,60,90,120,Inclusive

Table 15.1. Binning optimized for the distributions to unfold. Inclusive bins have no upper bound

CHAPTER 16

STATISTICAL TREAMENT

16.1 Introduction

The signal cross section is extracted through a simultaneous fit over the signal region, the ggF control region and Zjets control region. As discussed in previous sections, the discriminant in the ggF control region is a dedicated ggF BDT. The discriminant in the Zjets control region is the kinematic variable m_T between 40 and 120 GeV, defined in the section on observables. The signal region is split into two parts, Signal Region 1 (SR1) and Signal Region 2 (SR2), split along the vbf dedicated bdt (VBF BDT), where in SR1 $VBFBDT > 0.5$ and SR2 is defined as VBF BDT< 0.5, such that SR1 is enriched with signal events and SR2 is enriched in Top and WW backgrounds. The fit over SR1 is performed over the VBF BDT and the fit in SR2 is performed over the TopWW BDT. The binning for each of the control and signal regions is optimized for $\frac{S}{B}$ ratio.

16.2 Normalization Parameters

For each of the significant backgrounds, the shape of the distribution is taken from Monte Carlo and the overall normalization is extracted from the norm parameter in the fit to data. The norm parameterization for each background is as follows:

Standard Model WW: A common template for the QCD and the EW WW processes is built. The normalization is constrained in the simultaneous fit. The shape of this background process is taken from MC. A single floating parameter together

with the top-quark background is used in the fit for the Signal Region, and a separate floating parameter for the GGF control region

$t\bar{t}$+ Wt: The $t\bar{t}$ and Wt processes are combined. Similarly to the SM WW normalization, the top-quark background is normalized as a single floating parameter together with the SM WW background in the signal region and separate norm parameter in the GGF control region , with their relative composition fixed by the MC predictions. The shape of this background process is taken from MC.

Z+jets: The QCD and EW contribution to the $Z \to \tau\tau$ production are merged in a single template. The shape of this background is taken from Monte Carlo. The systematics are propagated via extrapolation factors (acceptance corrections) estimated by the signal region over control region ratio: we vary the yields estimated in the signal region and in the control region, and normalize to the integral of the yield in these regions. The theory systematics are estimated on the QCD component only, which is dominant, and are assumed to be uncorrelated.

ggF: The $H \to WW$ contribution from gluon gluon fusion plus two jets leaking into the VBF signal region is estimated using data to minimize the dependance on the prediction of higher order corrections in the ggF cross section. The ggF normalization of a MC-based template in the signal region is kept floating in the fit, with its constraint coming from the discrimination power of the BDT discriminant, since the constraint from data in the SR on ggF is sparse, due to poor statistics.

W+jets: A data-driven method is used to determine both the normalization and the shape. The electron and muon fakes are considered together, as well as the double fake correction. Various systematic uncertainties are considered on the fake factor computation. Several systematic uncertainties are considered: W+jets sample composition, the subtraction of processes with real leptons in the anti-ID selection, and statistical uncertainties on the determination of the fake factors and double-fake correction on the fakes estimation. The different sources of systematic uncertainties

are assumed to be uncorrelated between each other as well as uncorrelated between the fake factor for electrons and muons. For the statistical uncertainty of the fake factor, the variations of the bins of the fake factor are considered as independent variations.

Other VV: This component includes WZ/γ^* , ZZ, $W\gamma$, and $Z\gamma$. The shape of this background category is taken from Monte Carlo, as well as the normalization for all components of this group of backgrounds.

The normalization parameters are summarized in table 16.1

Parameter	Description
σ_{VBF}	Estimated Cross section in fiducial region (in fb)
μ_{wwTop}	Normalization of Top and Diboson backgrounds in Signal Region
μ_{ggf}	Normalization of ggF in signal region+ggF control region
μ_{Zjets}	Normalization of Zjets in signal region+Zjets control region
μ_{Zjets_CRGGF}	Normalization of Zjets in ggF control region
μ_{wwTop_CRGGF}	Normalization of Top + WW in ggF control region

Table 16.1. summary of parameters of interest in the fit

16.3 likelihood procedure

The likelihood function is defined is the product of Poisson distributions for the total number of events and the number of expected events in each bin. The compound probability density function of each binned distribution is the product of likelihood of each bin The liklihood function takes the form of the equation below, as a function of signal strengh μ_s and background strength μ_b

$$\mathcal{L}(\mu_s, \mu_b) = P(N|\mu_s + \mu_b \cdot b_{SR}^{exp}) \times P(M|\mu_b \cdot b_{CR}^{exp}) \tag{16.1}$$

Where P denotes a poisson distribution, N is the number of events observed in data in the signal region, M is the number of events in data observed in the control regions, b_{SR}^{exp} is the expected number of events background from MC simulation in the signal

region and b_{CR}^{exp} is the expected number of events from MC in the control regions. μ_s and μ_b are the normalization parameters for signal and background in the fit. The total cross section is defined by the following

$$\sigma_{fiducial} = \frac{\mu_s}{L \cdot A \cdot C} \tag{16.2}$$

Where L is the total integrated luminosity, C is the correction for reconstruction efficiencies and A is the acceptance factor, which in the fiducial case is 1. The C factor is calculated by taking the ratio of truth level MC to reco level MC.

Nuisance parameters are added to the likelihood function to account for additional uncertainties apart from statistical. For an uncertainty θ an additional constraint that takes the form $N(\bar{\theta}|\theta)$ for measurement $\bar{\theta}$ given some uncertainty parameterized by nuisance parameter θ. Uncertainties from limited MC statistics are accounted for with a poisson constraint (called gamma parameters) that takes the form:

$$P(\bar{\theta}|\theta, \lambda) = \frac{(\theta\lambda)^{\bar{\theta}} e^{-\theta\lambda}}{\bar{\theta}!} \tag{16.3}$$

Where λ is the nominal value of the measurements. Systematic uncertainties are given the form of a unit gaussian:

$$G(\bar{\theta}|\theta, 1) = \frac{1}{\sqrt{2\pi}} e^{\frac{(\bar{\theta}-\theta)^2}{2}} \tag{16.4}$$

The events in the $e\mu$ and μe channels are summed together and the likelihood is defined as the product of Poissonian over the BDT distributions in the signal region, multiplied by the product of Poissonians over the BDT distribution bins in the control regions. The normalization parameter μ_b that was introduced in the simple case above and it is now split into several background sources outlined in table 16.1 The minor backgrounds which do not have floating normalization factors as well as data-driven

W+jets estimate are added to the Poisson expectations. The total likelihood function thus takes the following complete form:

$$\mathcal{L}(\mu,\vec{\theta}) = \prod_{j=0}^{N_BDT_bins} P(N_j|\mu_{sj}+\sum_{bckg}\mu_b^n b^{nj})\times \prod_{k=0}^{CR_bins} P(N_k|\mu_{sk}+\sum_{bckg}\mu_b^n b^{nks})\times\prod_{i=0}^{N_{\theta i}} N(\bar{\theta}_i|\theta_i) \tag{16.5}$$

where b^{nj} is the background source of type n (i.e. ggF, WW, top-quark and Z to tau tau) in bin j with corresponding normalization parameter μ_b .

The signal and background expectations are functions of the nuisance parameters θ. These functions are parametrized by a sixth order polynomial piecewise interpolation between the impacts of the nuisance parameters to each given quantity they affect at the values of $\pm 1\sigma$. This is done such as that the response of s and b to each θ is factorized from the nominal value of the expected rate., that is $s = s_0 \times \prod \nu(\theta)$, where $\nu(\theta)$ depends on the systematics source. In general, all the systematics affect both the shape of the discriminant and overall normalization. The normalization impact of the systematics is normalized to all the sum of all regions modeled by that parameter.

16.4 statistical fit

The fit is performed simultaneously over all 4 regions. Table 16.2 shows the binning in each region and which discriminant is used.

Region	Discriminant	Binning
Signal Region 1	VBF BDT	4 dynamic
Signal Region 2	TopWW BDT	4 equal sized
CRZjets	M_T	5 equal sized
CRGGF	GGF BDT	4 equal sized

Table 16.2. summary of fit regions in the inclusive case

for the signal region, the 4 dynamic bins of the VBF BDT are 0.5-0.7,0.7-0.86,0.86-0.94,0.94-1.0. It was chosen to insure a reasonable signal to background ratio in all

4 bins. For the Zjets control region, the variable m_T was found to have sufficient discriminatory power compared with a dedicated zjets BDT.

Systematic uncertainties are considered in the likelihood function with a gaussian constraint nuisance parameter described in the previous section. There are two types of systematic uncertainties, ones that affect the overall normalization and ones that affect the shape of the distributions. For major backgrounds, but types of uncertainties are considered. For minor backgrounds, such as Higgs production modes other than ggF, only the normalization is considered as there are to few MC statistics for the shape. Individual sources of systematics uncertainty are considered uncorrelated. Correlation of a given systematic uncertainty is maintained across processes and channels. The normalization scheme for systematic uncertainties on different samples is described in detail in the following section.

16.4.1 Sample specific uncertainties

Uncertainties impacting the signal: All uncertainties impacting the signal are normalized to the signal region, thus to the integral of SR1 and SR2 in the fiducial phase-space. Experimental uncertainties impact the numerator of the C-factor (fiducial efficiency in each bin), while modeling uncertainties impact both the numerator and denominator of the C-factor (fiducial efficiency in each bin). Acceptance uncertainties are considered only on the bin-to-bin migration in the signal region events, and on the extrapolations from the signal region to all other regions where VBF contributions exists.

Uncertainties affecting top-quark and WW backgrounds: Since these backgrounds carry a floating normalization for several regions, two cases are to be distinguished. A) For cases where the normalization is determined from data, i.e. in the SR and the ggF CRs, each component of the uncertainty is normalized to sum of the Top+WW expectation in each respective region. This means that only bin-to-bin or region-

to-region (SR1 to SR2) effects are considered. Since these samples carry a common normalization factor, each uncertainty impacts the total template by the relative fraction of the sample composition. The same definition is used for both experimental and modeling uncertainties. All modeling uncertainties are treated as uncorrelated between the components of the pdf. Each common source of experimental uncertainty that impacts all discriminants used in the analysis is treated as correlated between each category or bin of the likelihood. B) For cases where this background is determined from simulation, i.e. in the Z+jets CR, the yields are normalized to the MC expectation, hence the uncertainties over the extrapolation to the full phase-space is used.

Uncertainties affecting Z+jets: Similar to the Top and WW case, Zjets uncertainties in the Zjets control region and the signal regions are normalized to the sum of the two regions. The systematics in the CRGGF control regions are normalized to that region only.

ggF uncertainties: In the CRGGF region, the uncertainty is normalized to the sample in that region. In all other regions the uncertainties are normalized to the MC expectation of the full phase-space.

16.4.2 smoothing

For some systematic uncertainties, the limited MC statistics cause large, unphysical bin to bin statistical fluctuations. To mitigate the impact of this for bins with smaller MC statistics, less than 10^4 events, the distribution is smoothed with a gaussian kernel approximation. For systematic uncertainties, RooKeysPdf is used to build a gaussian kernel density approximation from the raw event counts, which is then binned into the binning described in the previous section. For theory systematics, the histograms are smoothed at bin level by replacing each bin content with the sum of all the bin contents with gaussian weights.

To further reduce the amount of spurious fluctuations of systematic uncertainties in case of small MC- statistics some regularization of the variations is assumed. The following procedure is used. For each bin of the template if a given uncertainty fluctuates the yield more than 2σ than the average difference in all bins but the current one between the variation and the nominal, then the uncertainty in this bin is replaced by the interpolated effect between the two neighboring bins. The sign of the impact in this case is determined by the weighted average of the two neighboring bins, keeping the derivative of the underlying distribution. To note that the 2σ threshold is judged to be conservative, based on the fact that no single distribution contains more than 9 bins. Detector uncertainties are symmetrized if their one-sided impact is higher than 30% with respect to the other side. All uncertainties are symmetrized if their normalization impact is less than 0.5%. These criteria are applied to all types of templates, them being either raw or from the Gaussian -Kernel estimation. However, since the Gaussian-Kernel estimates are regularized prior to this, they rarely satisfy any of the above criteria.

16.5 Asimov fit results

The expected performance of the fit has been studied using the Asimov dataset. The Asimov dataset replaces the ensemble testing performed with Monte Carlo pseudo-experiments with a single "representative" dataset which returns the true value for each estimated parameter. The Asimov dataset allows studying the constraints on the nuisance parameters that could be obtained with the expected data distributions and statistical uncertainties. From the fit it was found that the expected fiducial cross section is:

$$\sigma^{fid} = 2.07^{+0.45}_{-0.45}\text{fb} = 2.07^{+0.35}_{-0.35}(\text{stat})\text{fb}^{+0.28}_{-0.28}(\text{sys})\text{fb} \tag{16.6}$$

where the total uncertainty is obtained from the fit while letting all parameters and nuisance parameters free to float and the statistical component is obtained by fitting the same dataset and fixing all nuisance parameters to their maximum likelihood estimate obtained from the fit while letting all parameters free. The profile likelihood scan of σ^{fid} is shown in figure 16.1. Figure 16.2 shows the results of the floating parameters in the fit with their uncertainty The correlation matrix between the parameters is in figure 16.4. The large correlations between the Top+WW parameters in the signal region and ggF control region comes from high impact systematics that are correlated across the regions, such as $t\bar{t}$ matching, $t\bar{t}$ shower and WW scale. The parameter of interest, the fiducial cross section, is only weekly correlated with the other parameters, which indicates that the weight of systematic uncertainties in the final result is small regardless of systematic uncertainty on the leading backgrounds. This is a result of the high discrimination in the signal region and the normalization of all backgrounds to data.

The results of the nuisance parameters are shown in figure 16.3 As expected nuisance parameters corresponding to systematic uncertainties are all centered on zero. The majority of the parameters (especially the ones related to the detector performance) are not constrained with respect to their prior uncertainties.

Figure 16.5 shows the impact of each nuisance parameter to the fiducial cross section after repeating the fit while fixing the parameters to their upper and lower maximum likelihood interval. The proceeding figures from 16.6 to 16.9 show the impact of the systematics on each of the fit paramaters.

Figure 16.10 shows the post fit distribution of the signal and background samples in the two control regions and the signal region.

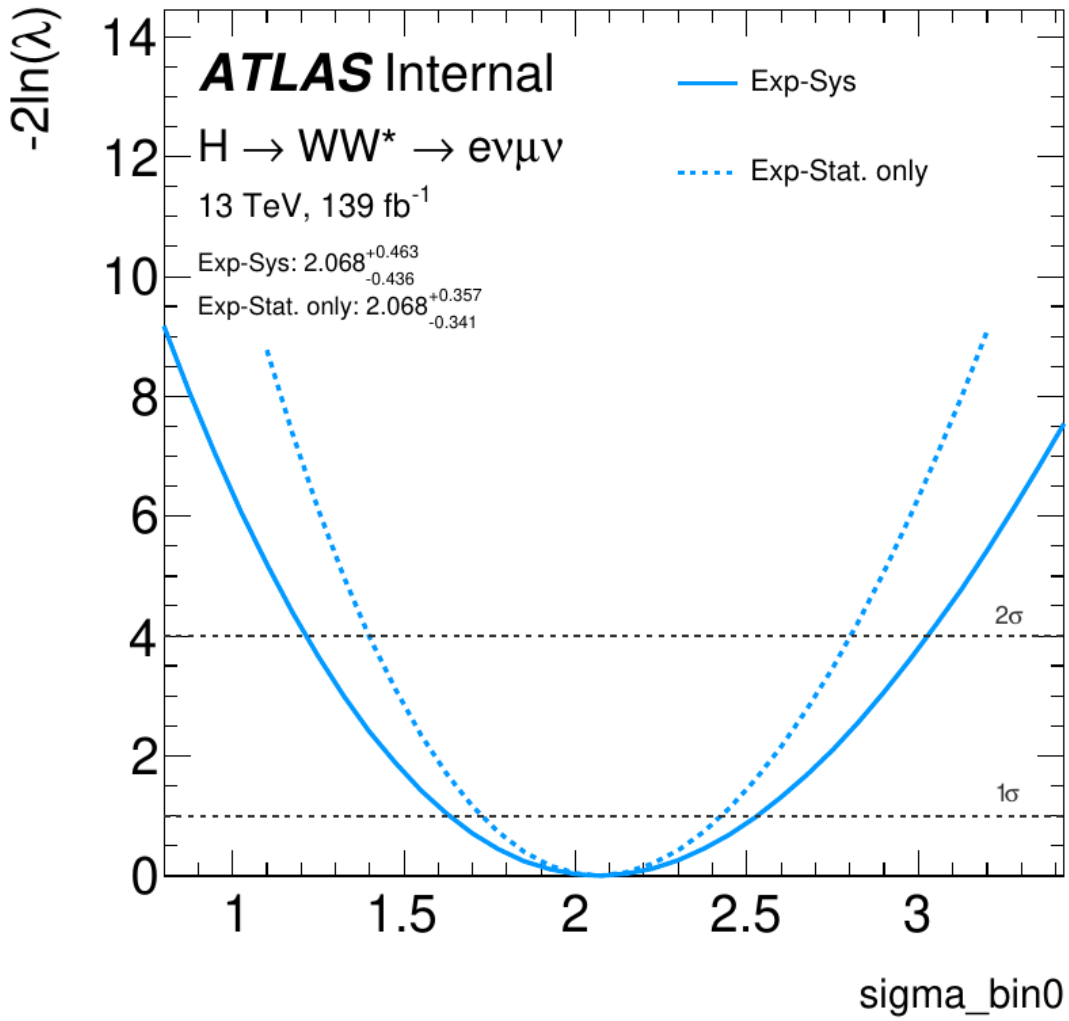


Figure 16.1. Profile likelihood scan of σ^{fid}. Dashed curve shows the scan with only statistical uncertainty; solid curve shows the scan with all nuisance parameters floating.

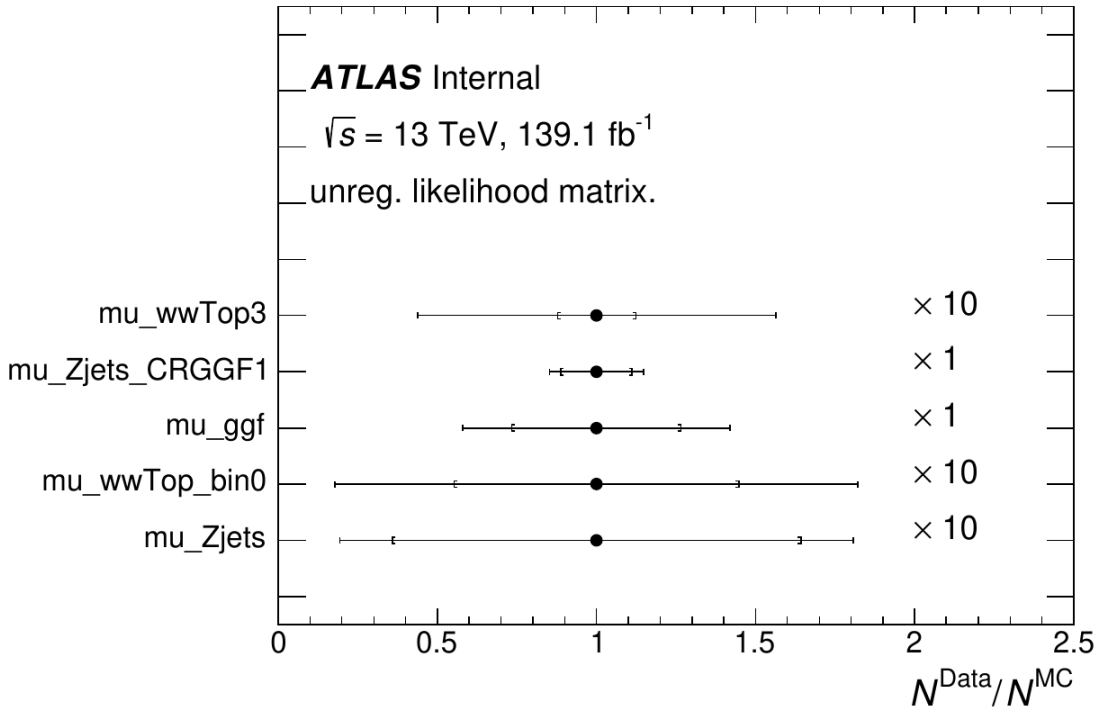


Figure 16.2. Norm parameters and uncertainties from inclusive Asimov fit. Inner error bars are stat only errors. Mu wwTop bin0 is the norm parameter in the signal region for Top+WW, wwTop3 is the parameter in the ggF control region

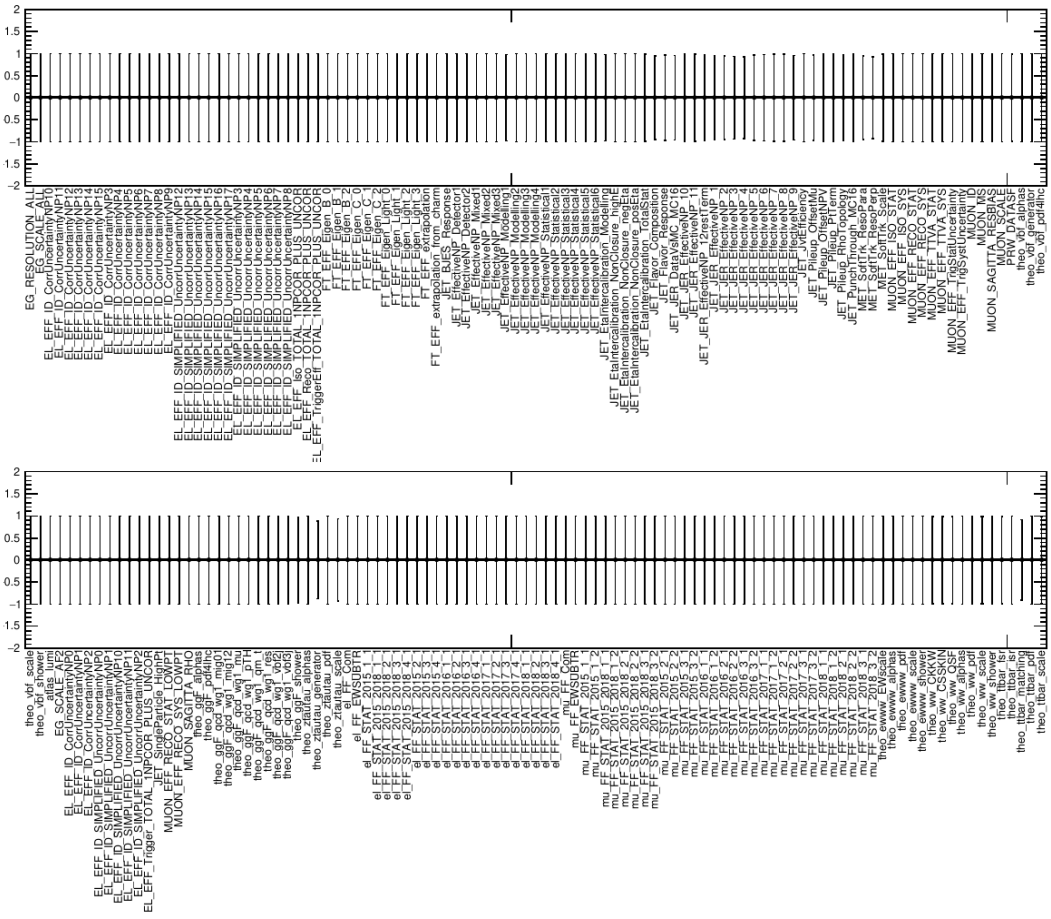


Figure 16.3. Pull plot of the nuisance parameters from the inclusive fit.

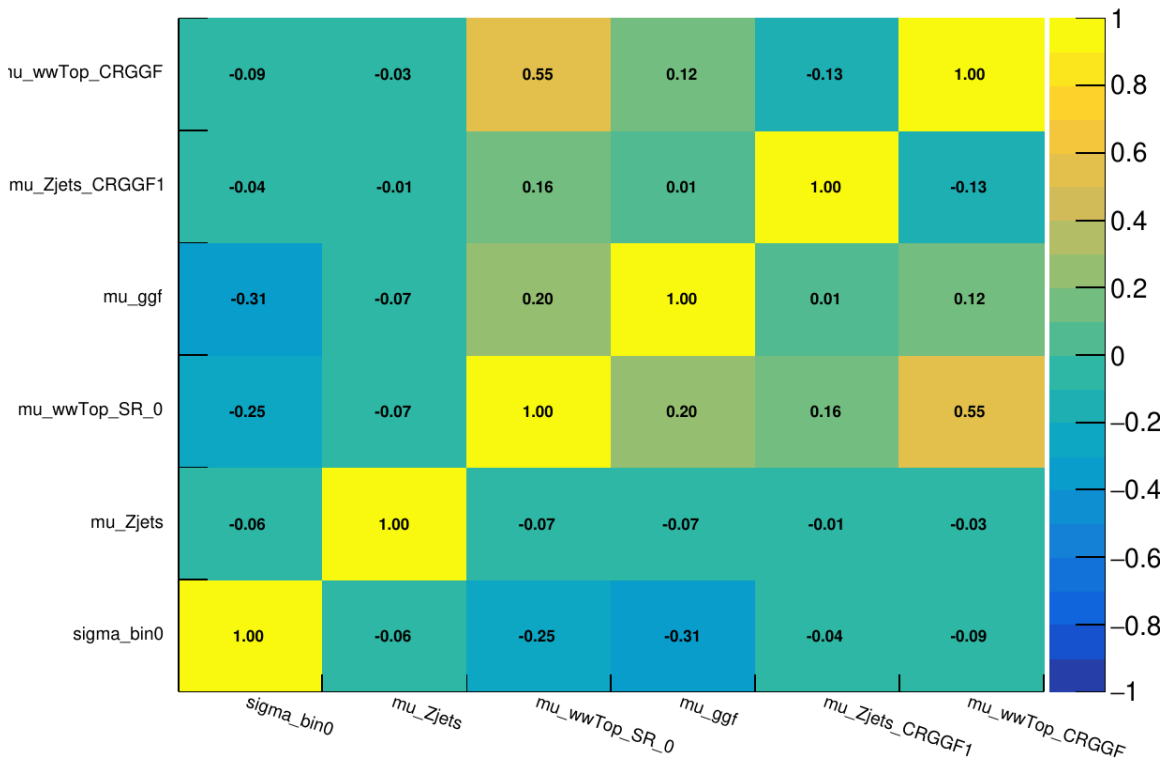


Figure 16.4. Correlation matrix of the floating parameters from the inclusive fit.

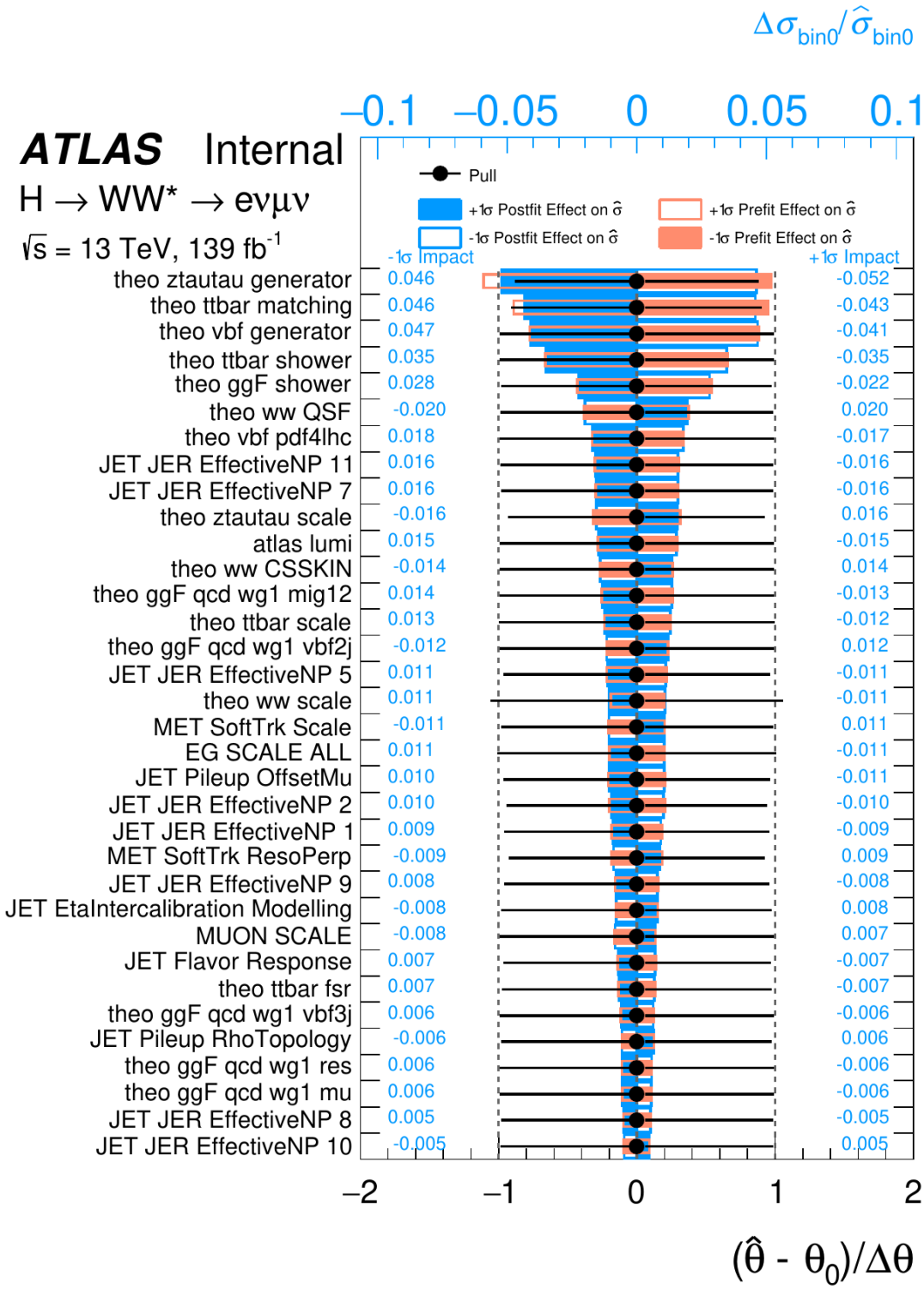


Figure 16.5. Impact of the highest ranked systematics on the fiducial cross section

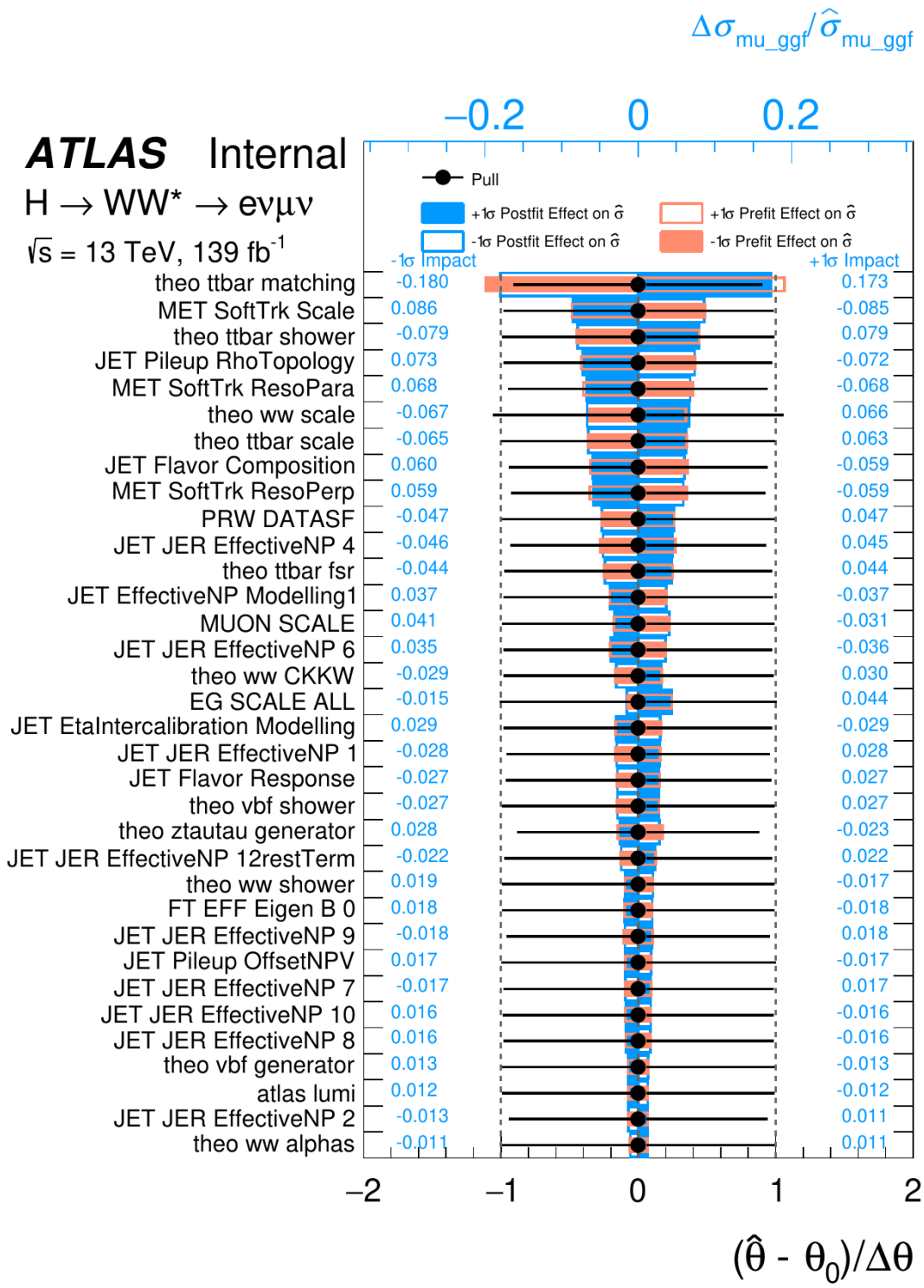


Figure 16.6. Impact of the highest ranked systematics on the ggF normalization parameter

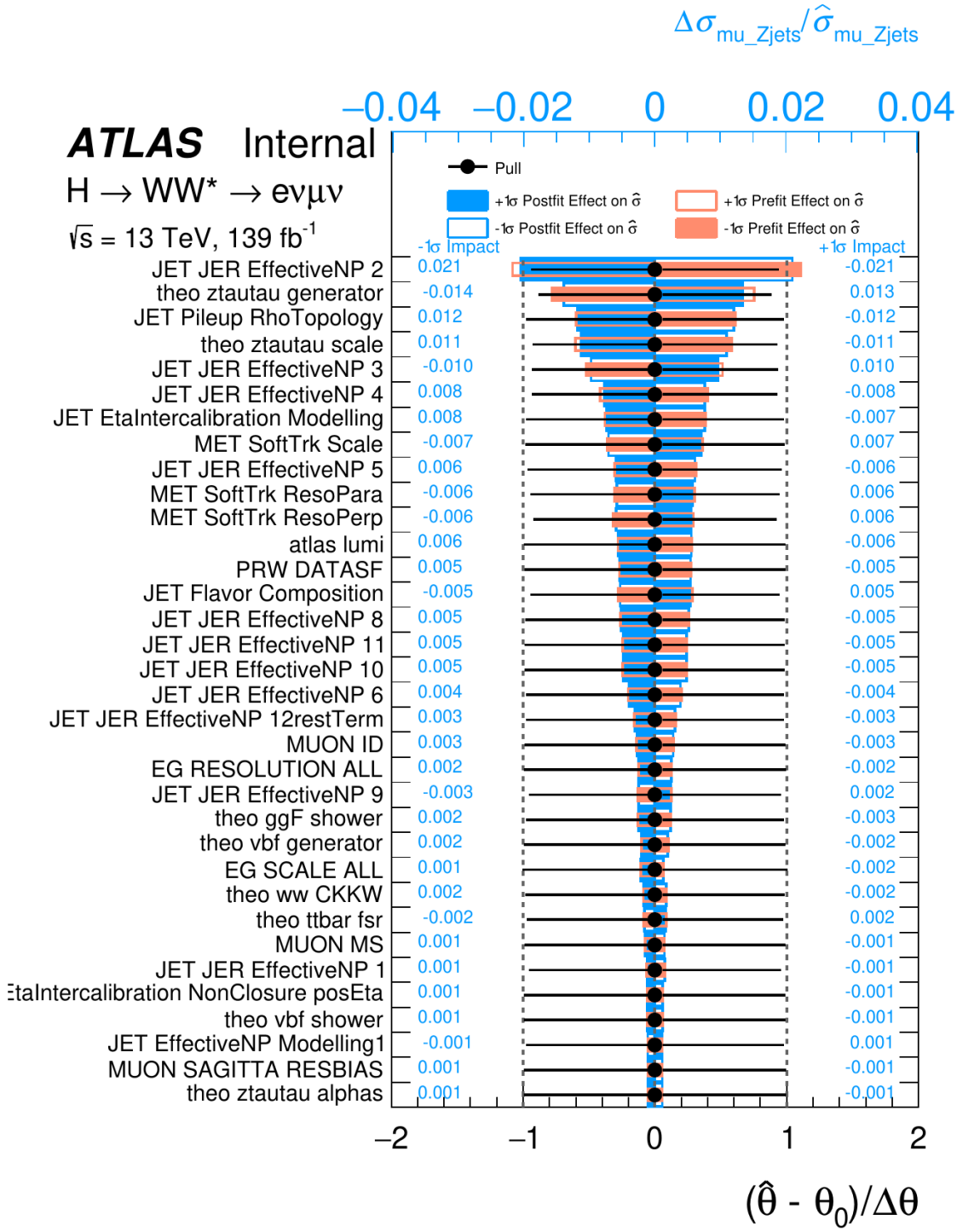


Figure 16.7. Impact of the highest ranked systematics on the Zjets normalization parameter

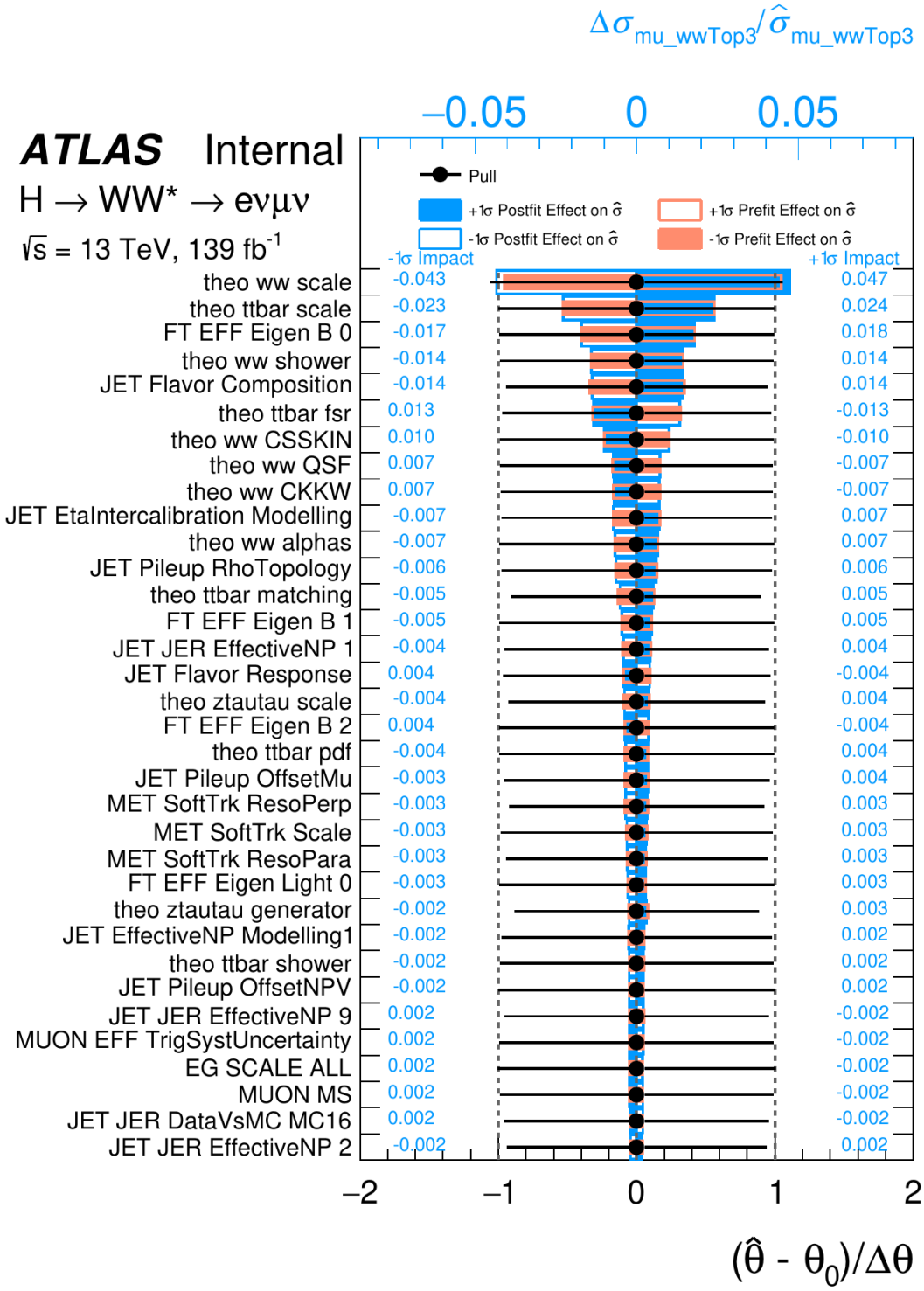


Figure 16.8. Impact of the highest ranked systematics on the Top+WW ggFCR normalization parameter

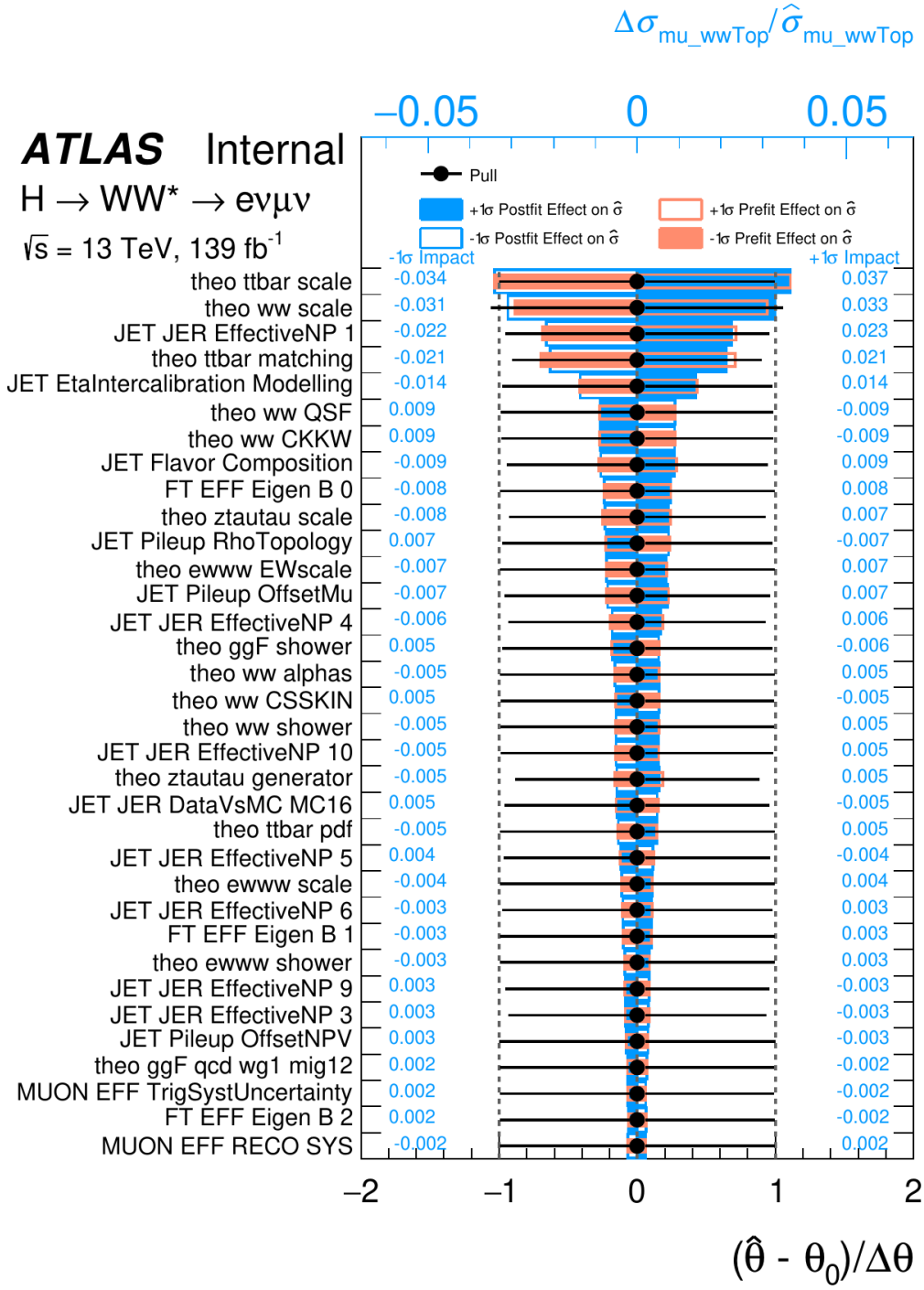
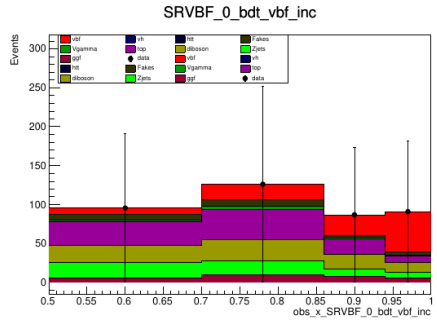
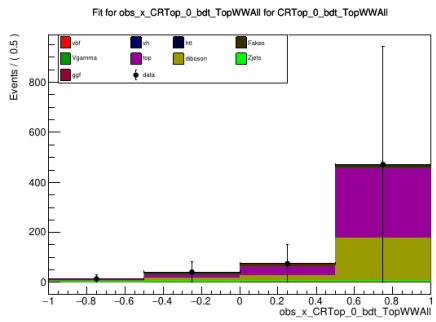
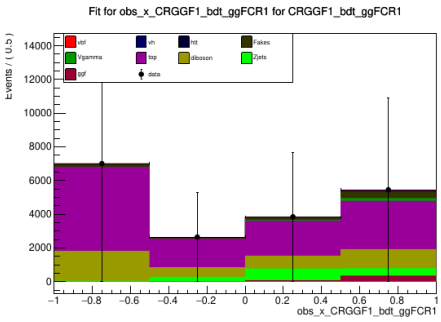


Figure 16.9. Impact of the highest ranked systematics on the Top+WW SR normalization parameter

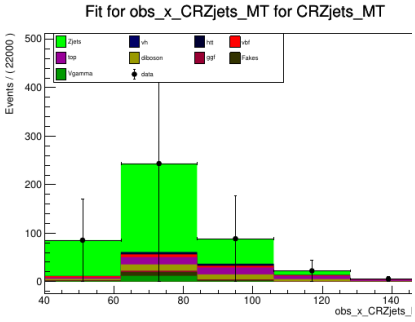




(a) Signal Region



(b) ggF-CR1 Control Region



(c) Z+jets Control Region

Figure 16.10. Phase space distribution of samples in the inclusive asimov fit.

16.6 Differential Statistical Analysis

The analysis for the differential case is nearly identical to the inclusive case described above, where instead of a single signal region, the signal region is split up by the differential bins. The main challenges are decreased statistics in each differential bin compared with the inclusive case and accounting for bin to bin migrations between the differential bins. The low statistics some differential bins is mitigated by merging the discriminant binning in some cases. For Signal Region 1, the dynamics inclusive binning is used, and neighboring bins are combined if the expected VBF in a bin is less than 1 . For Signal Region 2, the discriminant is rebinned from 4 to 2 bins if any process is 0 in the four bin case.

The strategy for dealing with bin to bin migrations is by the matrix-likelihood method. This method is similar to a matrix inversion-style of unfolding, however, the detector response matrix is embedded into the likelihood such as its minimization accounts for bin-to-bin migrations, as well as migrations from and to the integrated fiducial region. The signal parameter for differential bin i takes the form:

$$\mu_{si} = f_i \cdot L \cdot \sigma + \sum_{j}^{diff_bins} L \cdot A_{ij} \cdot C_{ij} \cdot \sigma^j \tag{16.7}$$

Where f_i is the fraction of events outside the fiducial region but reconstructed into the signal region in that bin, L is the integrated luminosity, $A_{i,j}$ is again the acceptance and C_{ij} is the element of the migration matrix.

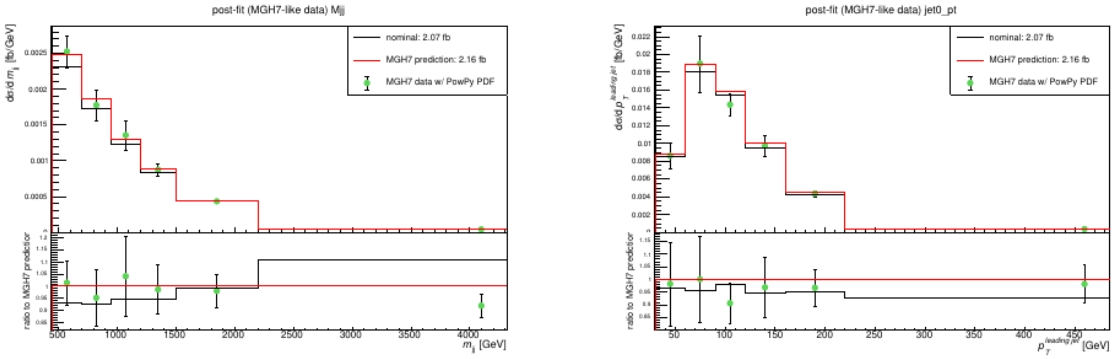
Systematic uncertainties are modeled as described in the previous section for both signal and background. Some differences are to be found in the treatment of uncertainties impacting the signal. Instead of propagating the uncertainties to a single factor, as done for the integrated fiducial cross section, they are propagated to each element of C_{ij}. In practice, this yields no difference in the treatment of the detector related uncertainties, as the denominator in Ci, j remains unchanged and the normalization is carried over all the differential bins of the signal region. For modeling

uncertainties instead, the impact to C_{ij} is evaluated by varying both the truth level and reconstruction level MC and generating a varied migration matrix for each signal sample systematic. s

16.6.1 Bias Test with Madgraph+Herwig signal samples

Possible modelling biases were tested using an Asimov data that is created from Madgraph5_NLO+Herwig7 (MGH7) sample. This sample and Powheg+Herwig7 are used for evaluating the signal matching uncertainty. Since some BDT input variables have shape difference between the MGH7 sample and the nominal (Powheg+Pythia8), such a test is used to search for any bias in the fitting model used in the nominal sample in the case of the pseudo-data presents different features, i.e. MGH7-like.

The bias for a few observables are shown in Fig. 16.11. An estimate of the bias was obtained by comparing the MGH7 prediction and the stat-only post-fit result using the MGH7 Asimov dataset. The p-value of the fit is 0.81 for M_{jj} and 0.94 for the leading jet p_T, therefore the measurement is not sensitive to any bias in the analysis within the MC statistical uncertainties of the fitting templates.



(a) (b)

Figure 16.11. Fit results (green) using the nominal model with MGH7 Asimov dataset, and comparison to the MGH7 (red) and nominal (black) predictions. The fit result (green) error represents the MC statistical uncertainty.

16.6.2 Floating background parameters scheme

In the differential fit, each of the differential bins needs a Top+WW normalization parameter. The possible schemes are either one global normalization parameter for all the bins, a unique parameter for each differential bin or shared parameter between differential bins. The third option is a good compromise between more accurate fit results and not too many floating parameters which makes it harder for the fit to converge. The normalization scheme for differential Top+WW parameters is described in table. 16.3. Every two or three bins share a normalization parameter, unless the expected shape of the Top+wW is significantly different, in which case the scheme is adjusted.

Observable	Range and bins with independent Top+WW normalisations
m_{jj}	One normalisation for every two bins.
Leading Lepton p_{T}	One normalisation for every two bins.
Leading Jet p_{T}	One normalisation for every two bins.
Sub Leading Lepton p_{T}	One normalisation for first 3 bins and one parameter for the last 2.
Sub Leading Jet p_{T}	One normalisation for first 3 bins and one parameter for the last 2.
Δy_{ll}	One normalisation for first 3 bins and one parameter for the last 2.
Δy_{jj}	One normalisation for every two bins.
m_{ll}	One normalisation for first 3 bins and one parameter for the last 2.
Higgs p_{T}	One parameter for first bin, one parameter shared between the next 3 bins and one parameter for the last bin
$\Delta\phi_{jj}$	One normalisation for the first two bins, one paramter for the last two bins.
$\Delta\phi_{ll}$	One normalisation for first 3 bins and one parameter for the last 2.
$Cos(\theta*)$	One normalisation for each every bin.

Table 16.3. Floating parameter scheme used for the Top+WW normalisation.

16.6.3 Asimov Results

The results for M_{jj} fit is presented here. Figure 16.12 shows the expected differential cross section as a function of M_{jj}, table 16.4 shows the same results, with statistical and systematic uncertainty noted. Figure 16.13 shows the values and uncertainties of the normalization parameters. Except for Top+WW in the signal region, the normalization parameters are identical to the inclusive case. For M_{jj} there are 3

Top+WW norm parameters, shared between every two bins. The left plot shows the correlation matrix between the normalization parameters. The correlation between the cross section parameters is due to bin to bin migrations, the correlation between the Top+WW parameters in the signal region and the ggF control region, as with the inclusive case, is due to large systematic uncertainties in $t\bar{t}$ and WW. 16.14 and 16.15 show the bdt distribution in signal region 1 and signal region 2 respectively of each of the differential bins in M_{jj}.

The results of the nuisance parameters are shown in figure 16.16. Since the nuisance parameters are constrained by a unit Gaussian, if the fit converged and we have good closure the nuisance parameter will have a nominal value of 0 and width of 1. Finally figures 16.17 to 16.22 show the rankings of the most important systematic uncertainties and their impact on the cross section measurements.

Bin xsec	value [fb]	stat error [fb]	stat+sys error [fb]
σ_0	0.58	0.39	0.47
σ_1	0.43	0.31	0.34
σ_2	0.31	0.26	0.29
σ_3	0.25	0.21	0.22
σ_4	0.31	0.17	0.18
σ_5	0.19	0.10	0.10

Table 16.4. Floating parameter scheme used for the Top+WW normalization.

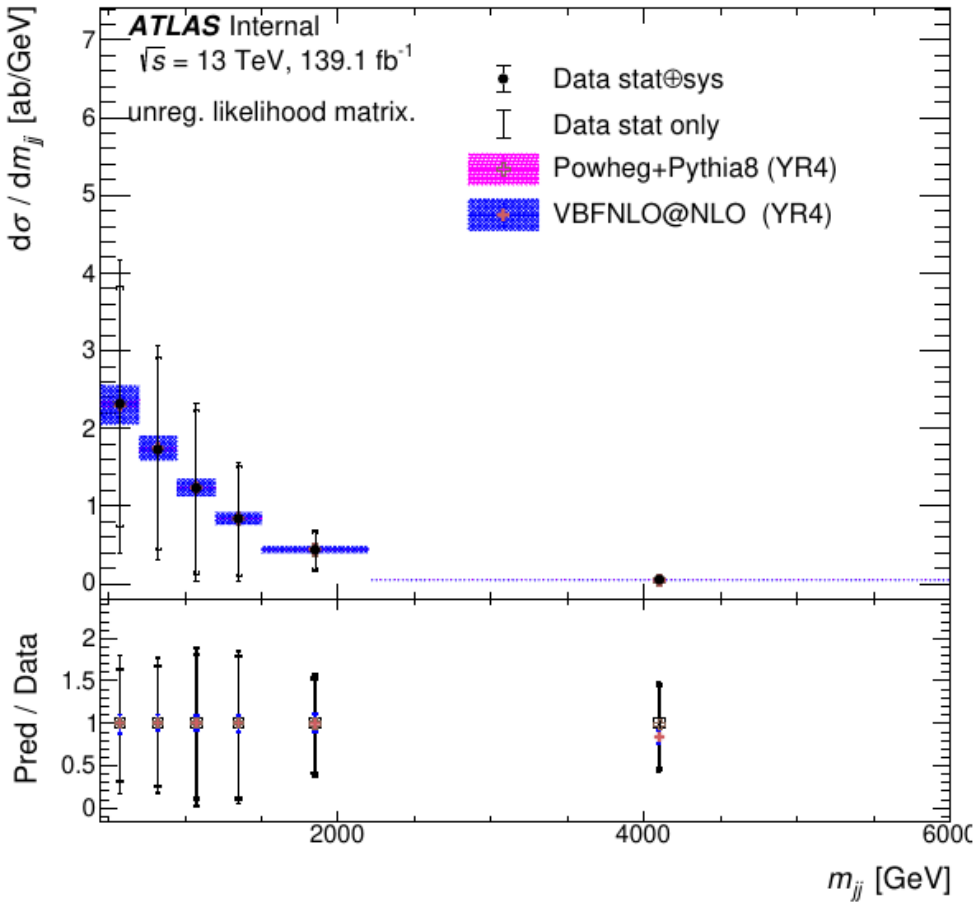


Figure 16.12. Post-fit differential cross section as a function of M_{jj} on an Azimov data set and comparison with theoretical expectations.

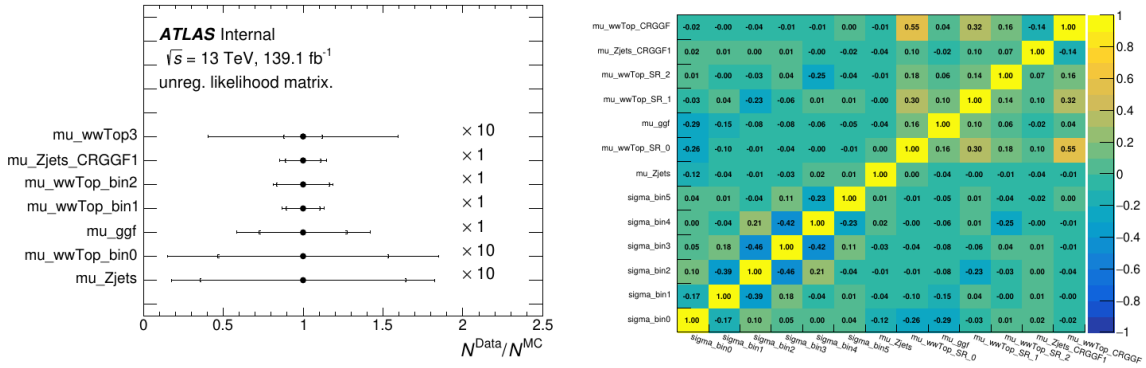


Figure 16.13. Left: Fit result of the normalization parameters in the measurement of the differential cross section as a function of M_{jj} with statistics-only and statistics+systematic uncertainties. Right: Correlation between normalization parameters (including the per-bin signal cross section) in the the fit that includes statistics and systematic uncertainties.

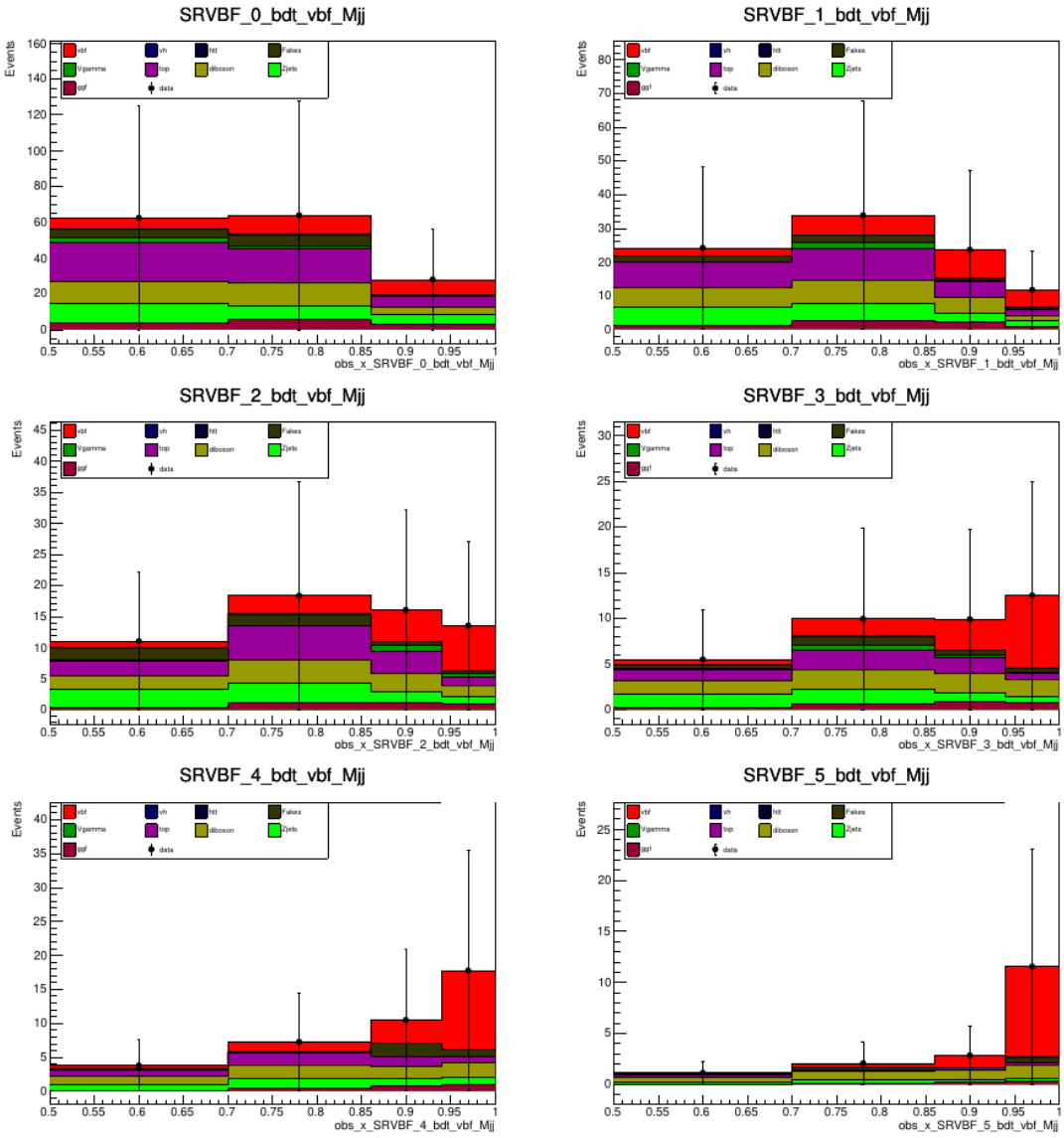


Figure 16.14. Distribution for Signal Region 1 (SRVBF) for each M_{jj} differential bin

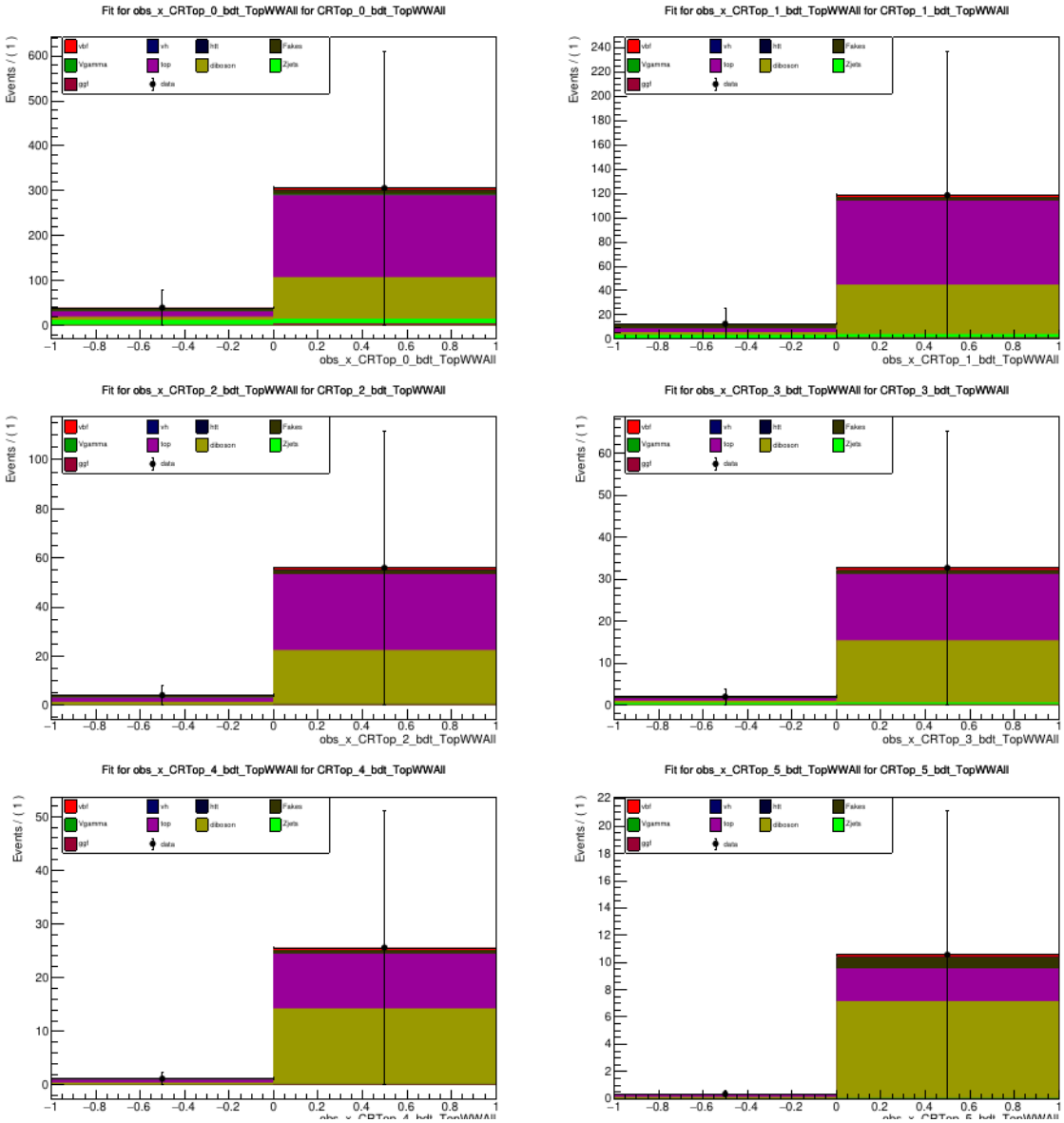


Figure 16.15. Distribution for Signal Region 2 (CRTop) for each M_{jj} differential bin

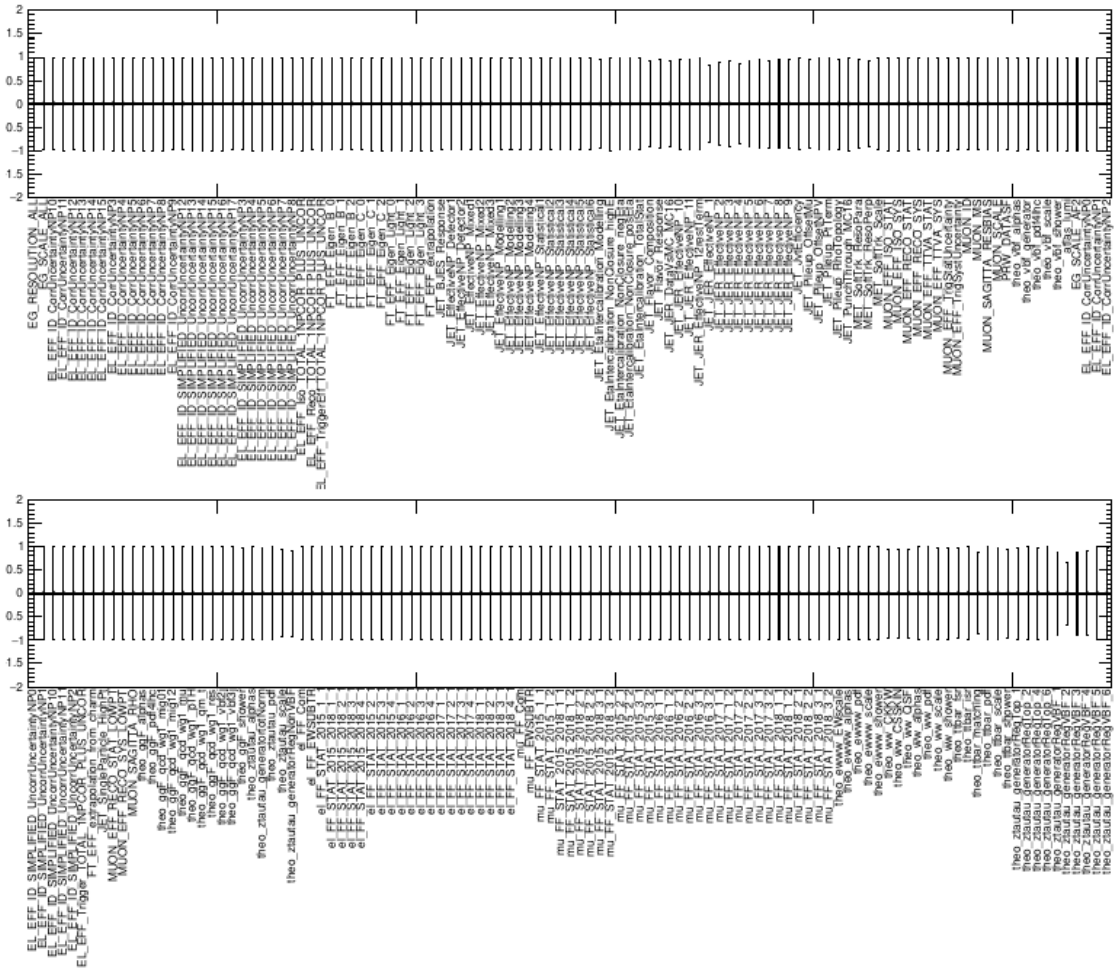


Figure 16.16. pull plots of all the nuisance parameters in the M_{jj} fit.

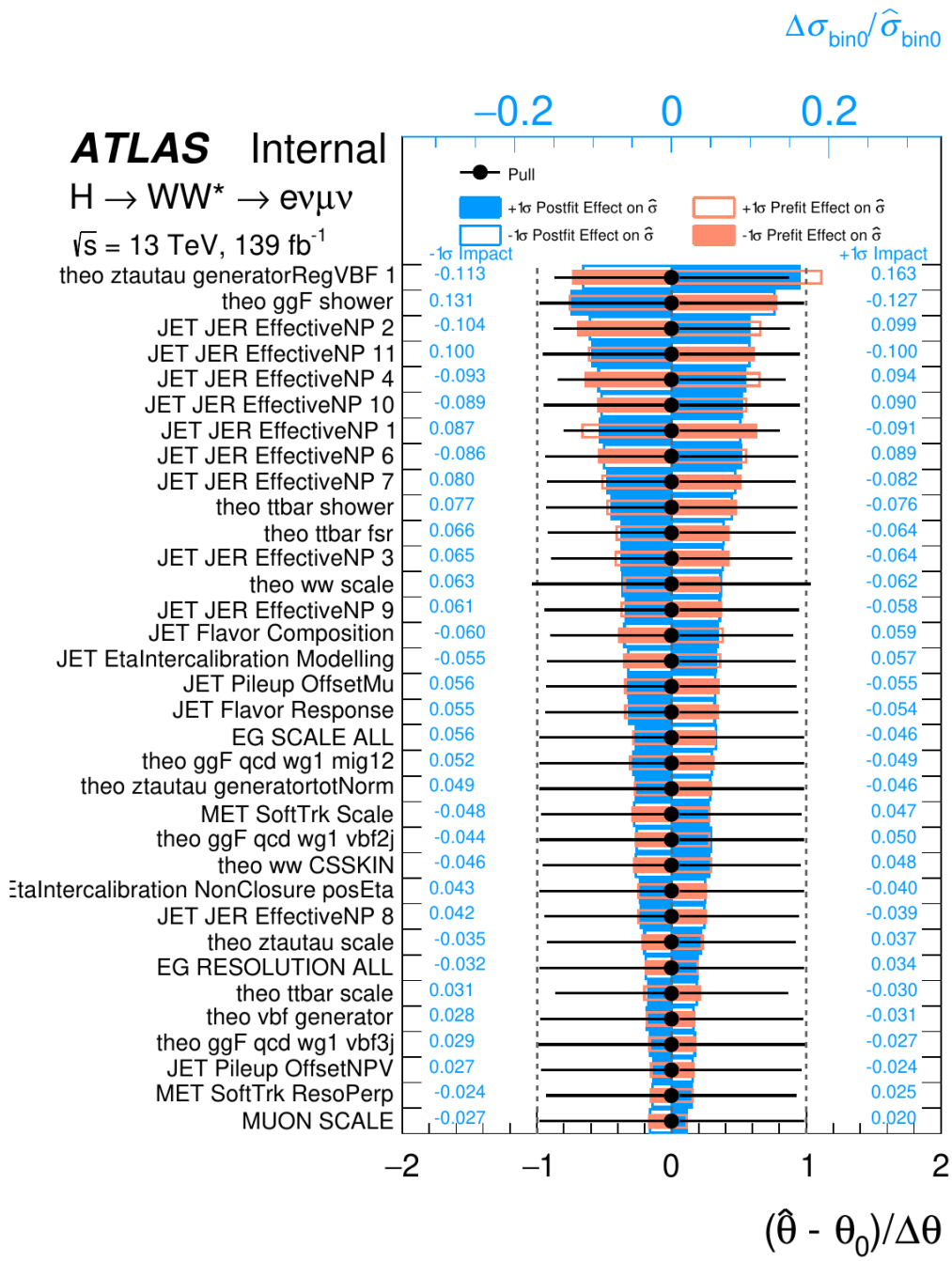


Figure 16.17. Ranking of systematic impacts on Mjj cross section bin 0.

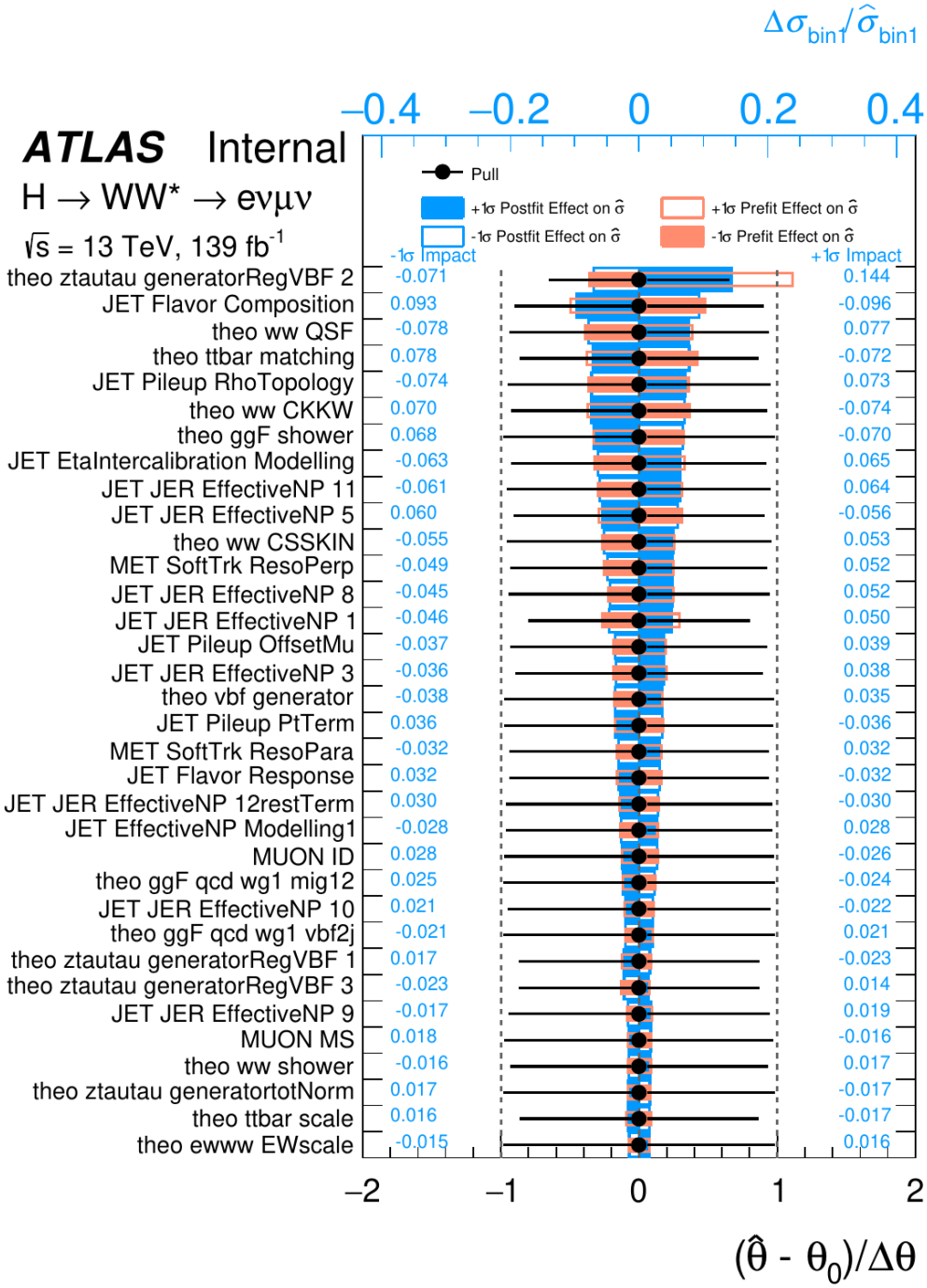


Figure 16.18. Ranking of systematic impacts on Mjj cross section bin 1.

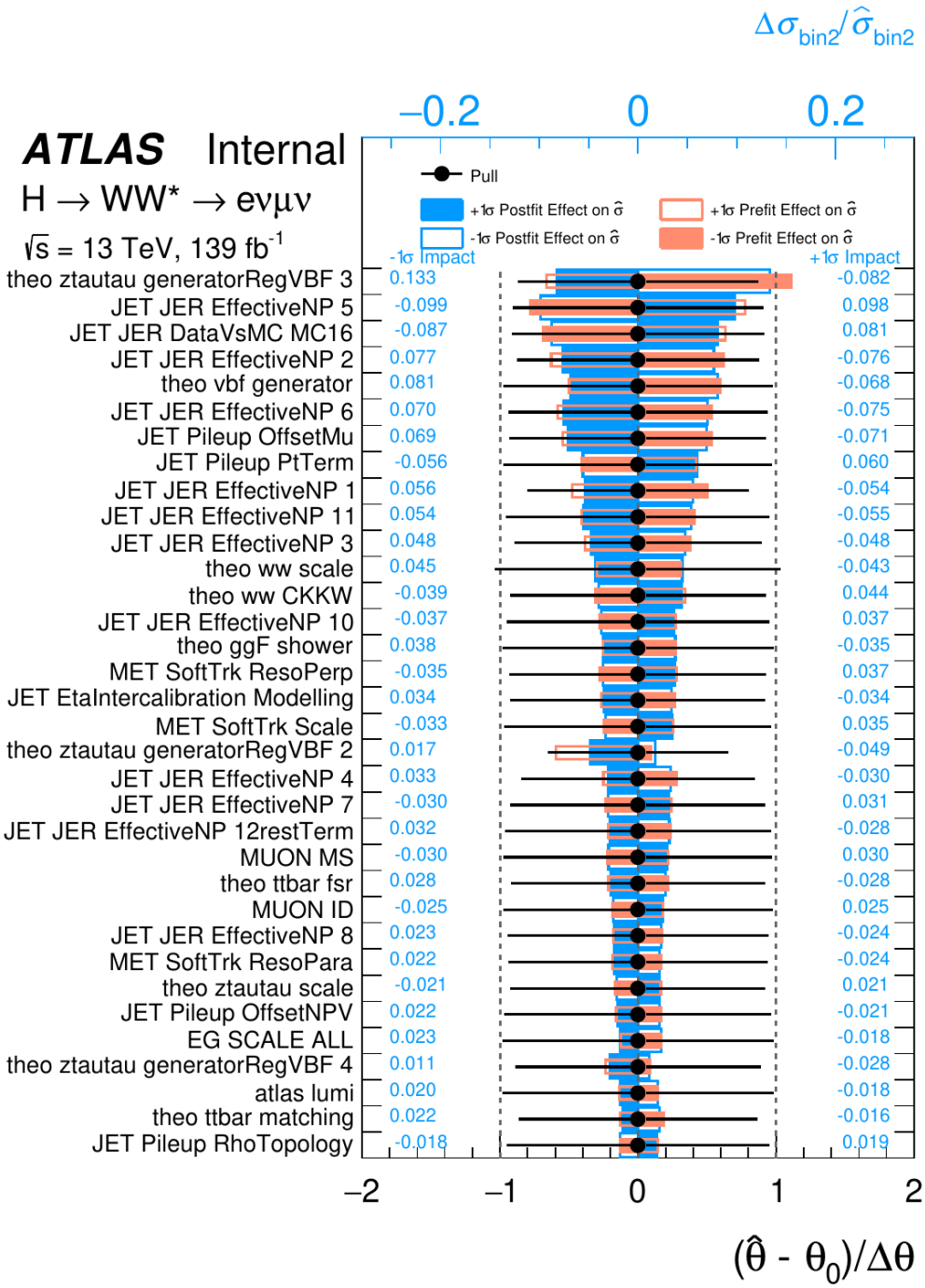


Figure 16.19. Ranking of systematic impacts on Mjj cross section bin 2.

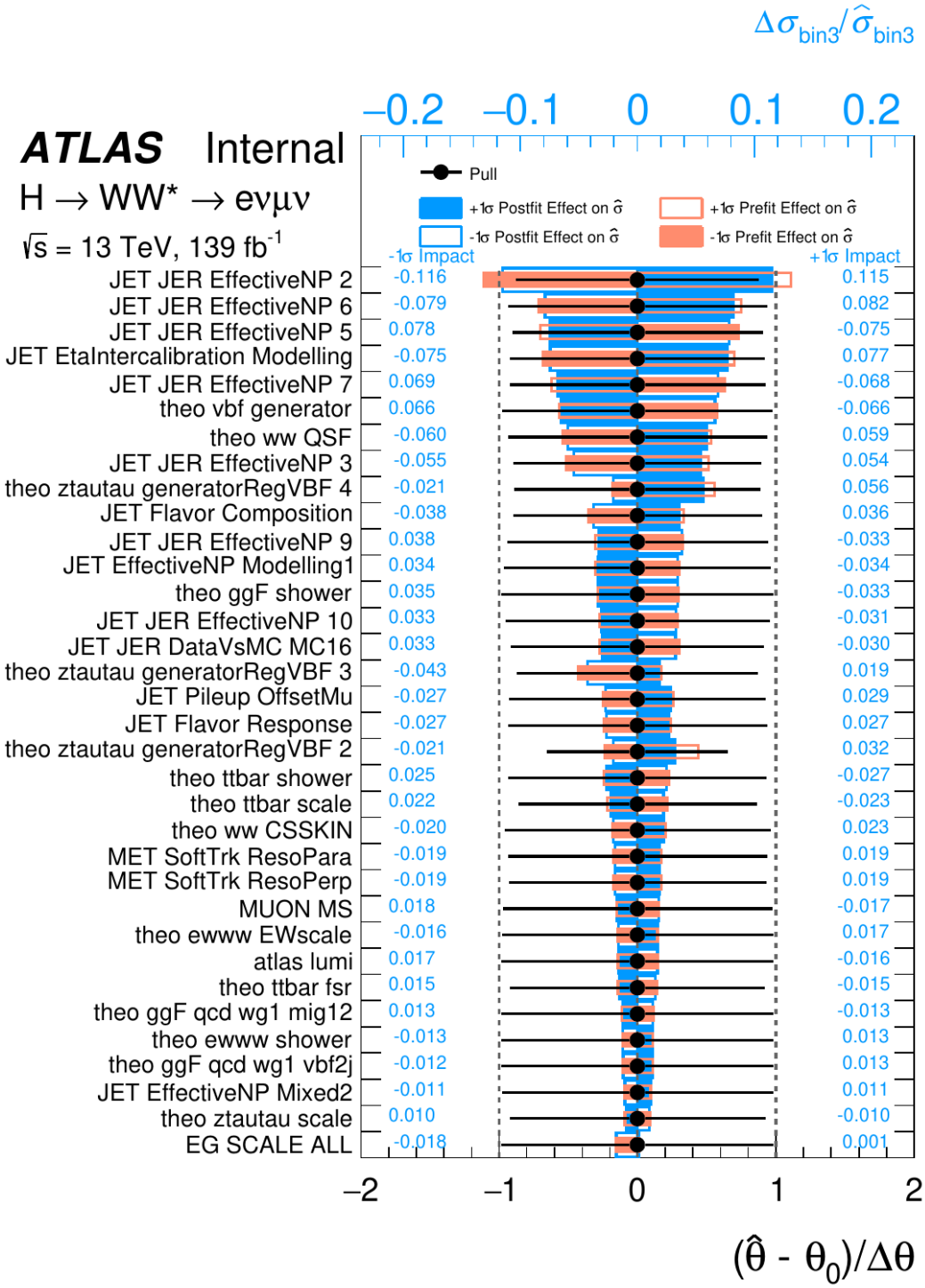


Figure 16.20. Ranking of systematic impacts on Mjj cross section bin 3.

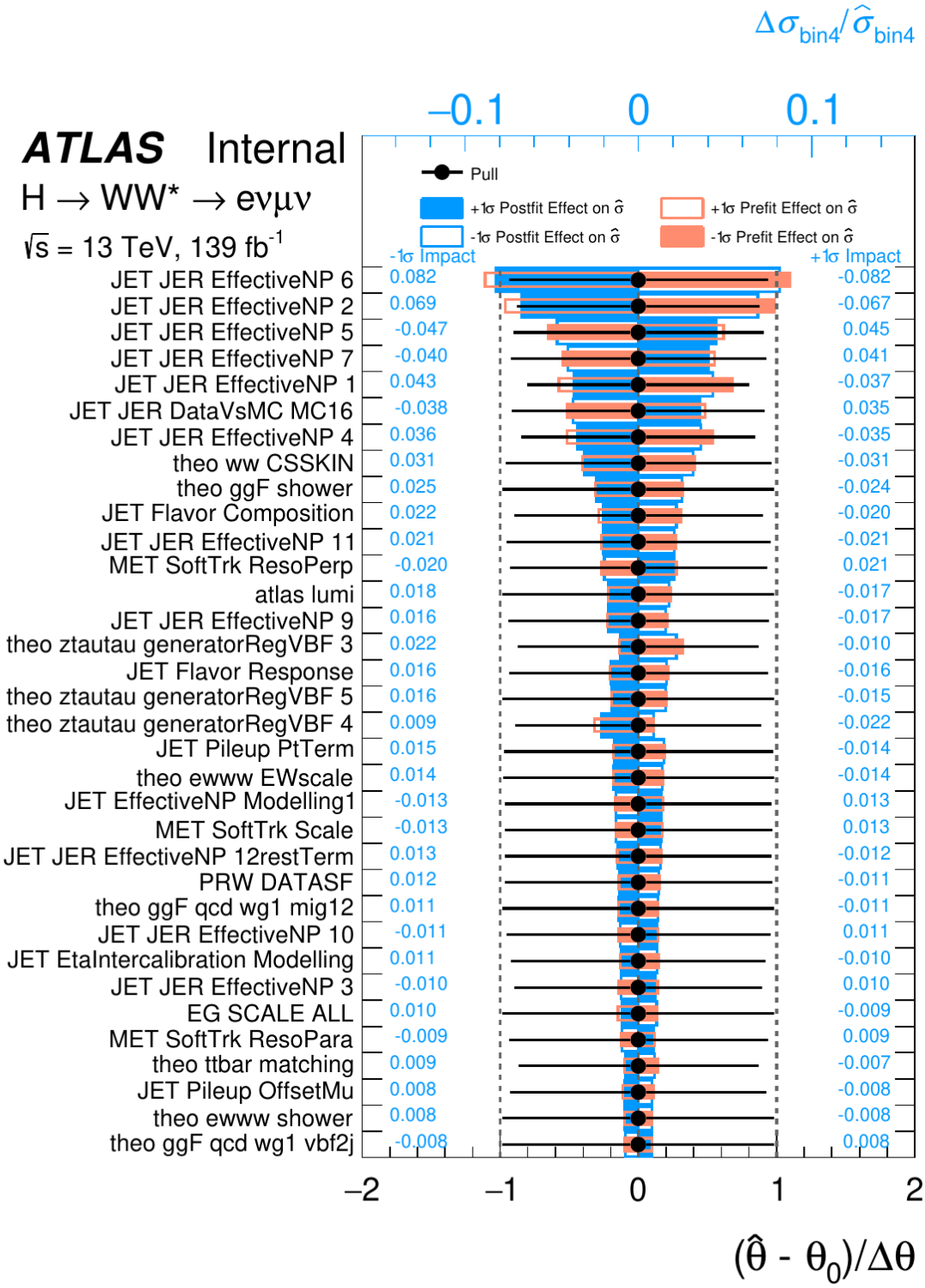


Figure 16.21. Ranking of systematic impacts on Mjj cross section bin 4.

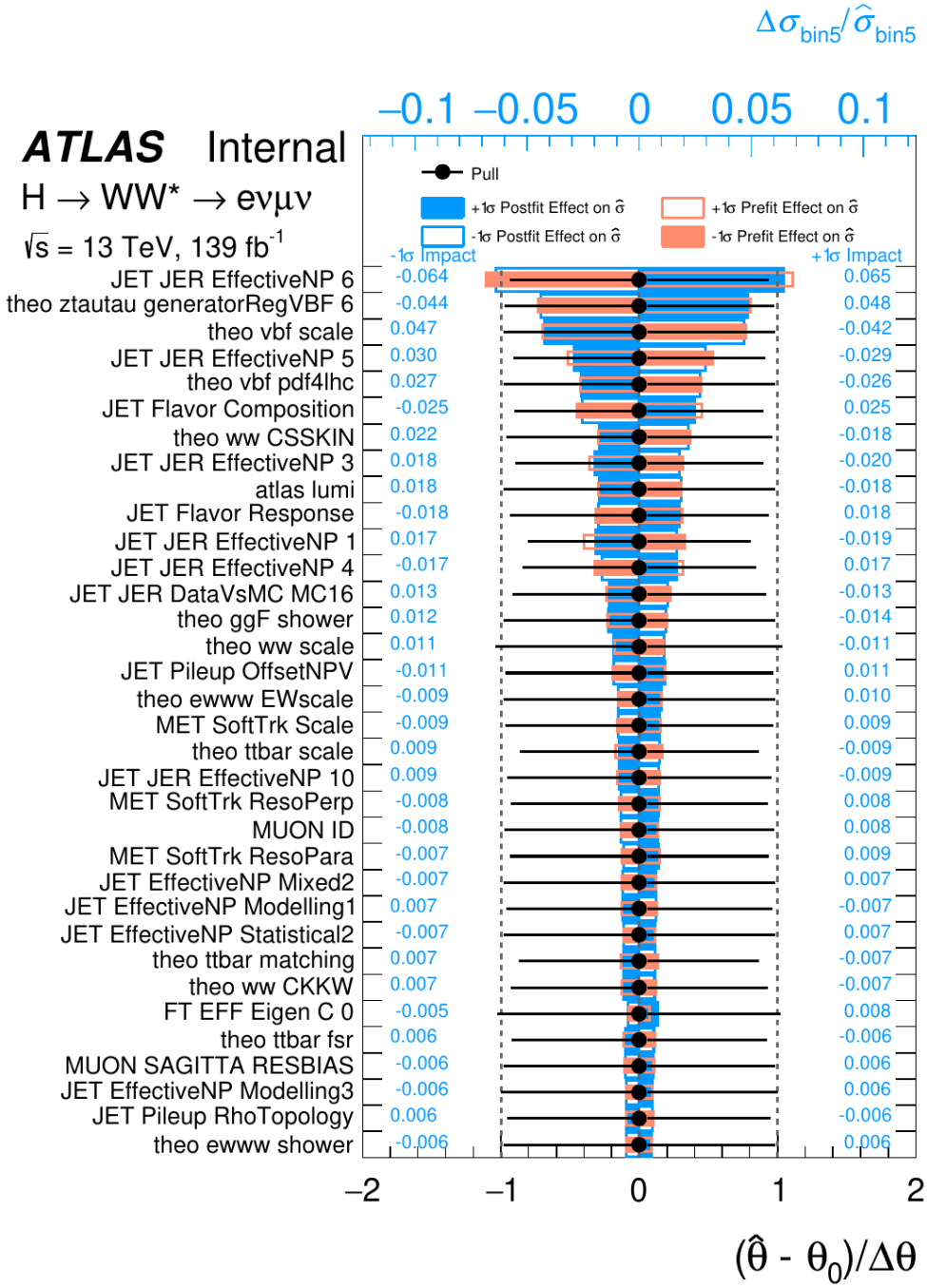


Figure 16.22. Ranking of systematic impacts on Mjj cross section bin 5.

CHAPTER 17

RESULTS

After validating the method with the Asimov results, the data was unblinded with the advise and consent of the Higgs Committee. This section will show the results of the statistical fit outlined in the previous section, with unblinded data from the LHC used for inputs and the full statistical and systematic uncertainties.

17.0.1 Inclusive Results

The result for the total cross section in the fiducial region is shown below. This cross section was extracted from the statistical fit described in the previous section, corrected by a C factor derived from the ratio of truth level to reconstruction level Monte Carlo.

$$\sigma^{fid} = 1.70\text{fb} \pm 0.42(\text{total}) \pm 0.36(\text{stat}) \pm 0.22(\text{sys}) \pm 0.20(\text{theory}) \pm 0.11(\text{exp}) \quad (17.1)$$

These results agree with the standard model prediction of 2.07 fb, at the 1 sigma level. The results exclude the null hypothesis that the VBF $H \to WW^* \to \ell\nu\ell\nu$ process is absent from data at 4.7 σ significance. Figure 17.1 shows the post-fit distributions of the discriminant in each of the signal and control regions. Figure 17.2 shows the results for all the normalization parameters and the correlation between the parameters. The correlations are close to what was observed in the Asimov results. As expected, the results are still dominated by statistical uncertainty. Future

analysis will take advantage of these results to combine with other Higgs production processes such as ggF to obtain an overall cross section with more statistics and lower uncertainty. Finally figure 17.3 shows the values of the nuisance parameters extracted from the fit.

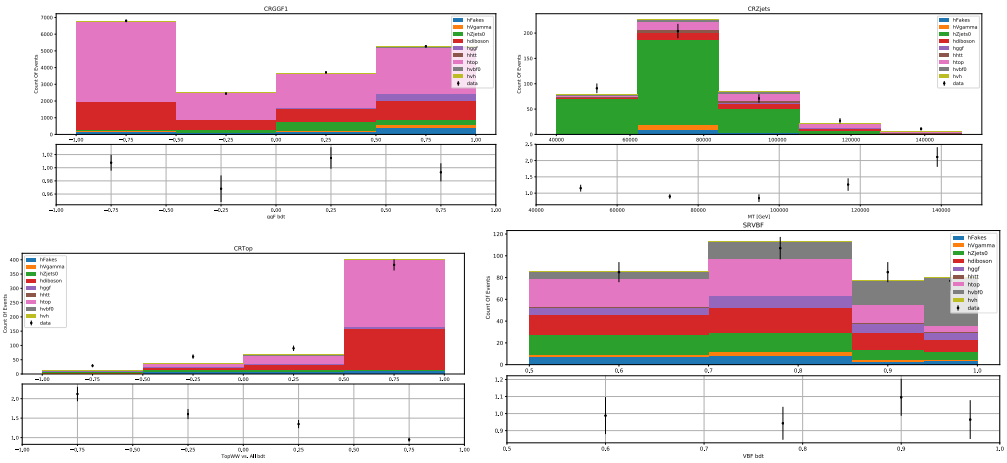


Figure 17.1. Distribution of the two control regions and signal regions after fit. Errors are stat only in this case. Top to bottom, CRGGF, CRZjets, Signal Region 2, Signal Region 1

17.0.2 Differential Results

This section looks at the differential cross section as a function of thirteen observables in the $H \to WW^* \to \ell\nu\ell\nu$ process outlined in the introductory section. The results are in agreement with standard model expectations. Each section contains the expected differential cross section per bin compared with the expectations from Monte Carlo, the correlation between the fitting parameters and the normalization values derived from the fit.

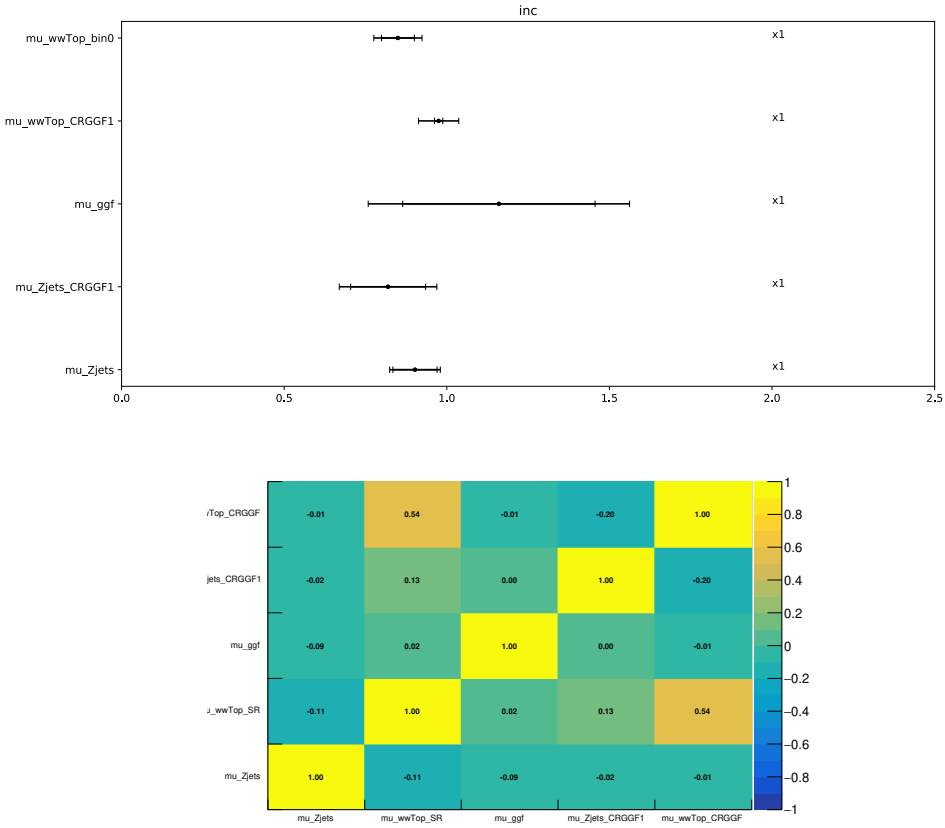


Figure 17.2. (top)Normalization values for background MC from inclusive fit, (bottom) Correlation matrix of fit parameters

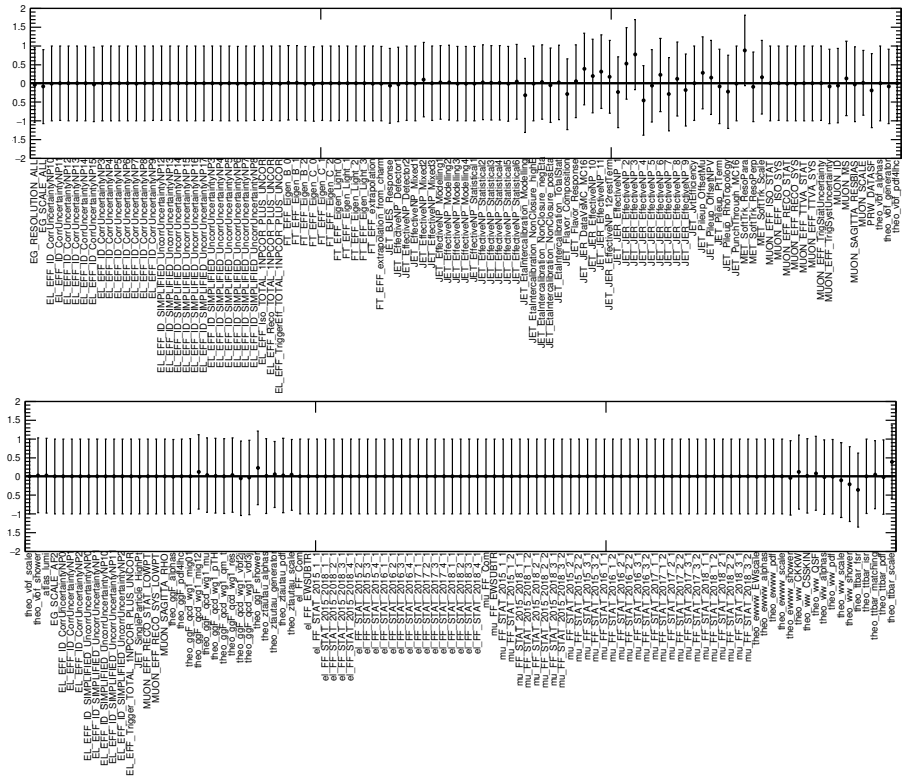


Figure 17.3. Nuisance parameters, central value and error

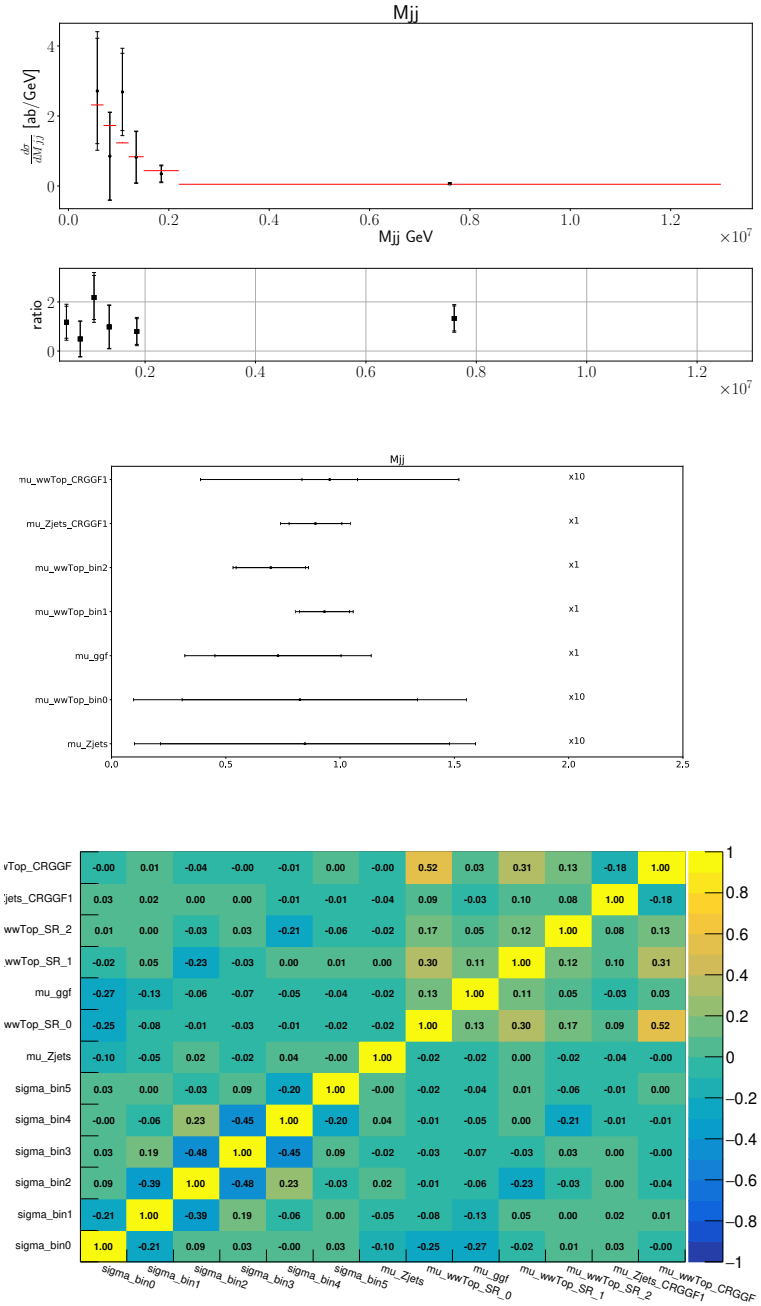


Figure 17.4. Mjj Full stat+sys fit to data

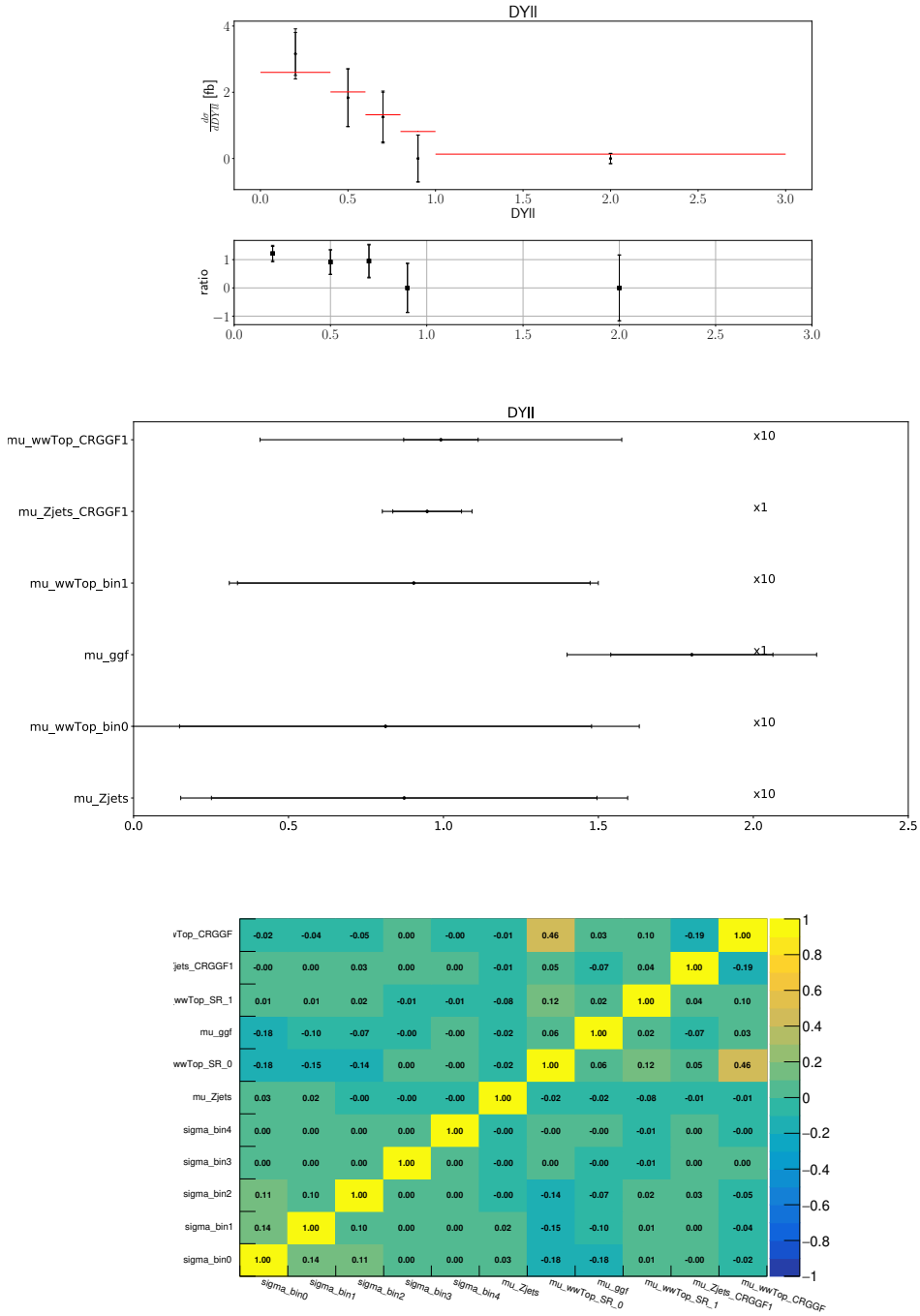


Figure 17.5. DYll Full stat+sys fit to data

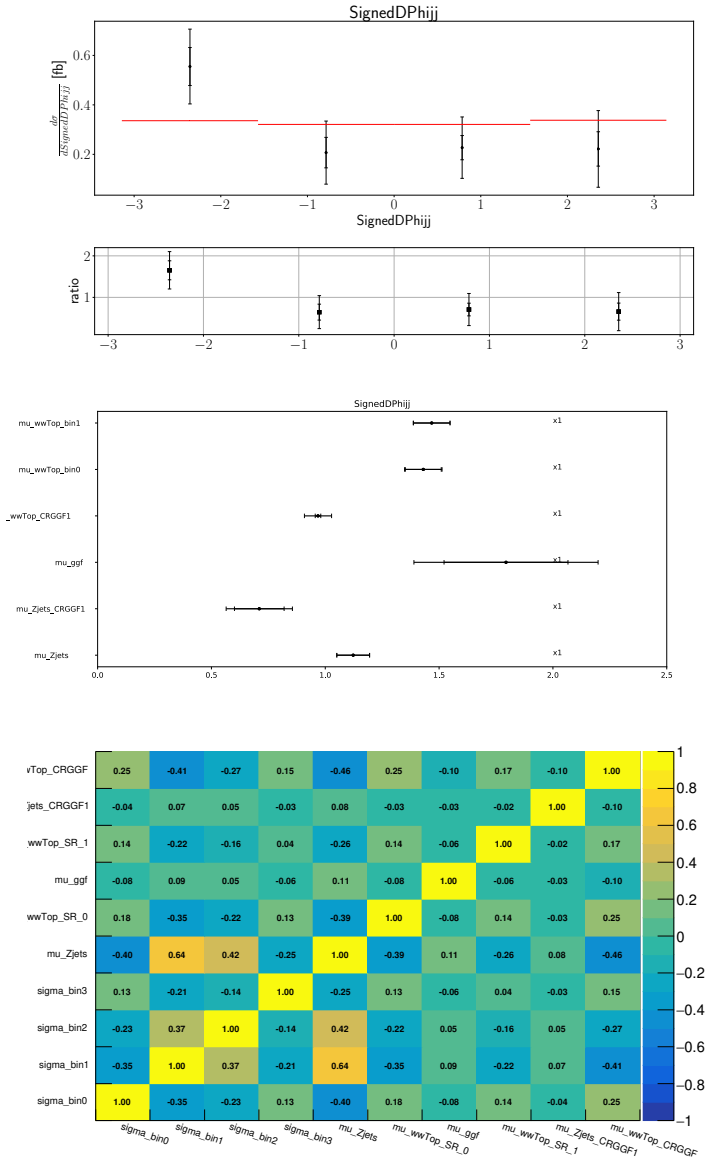


Figure 17.6. DPhijj Full stat+sys fit to data

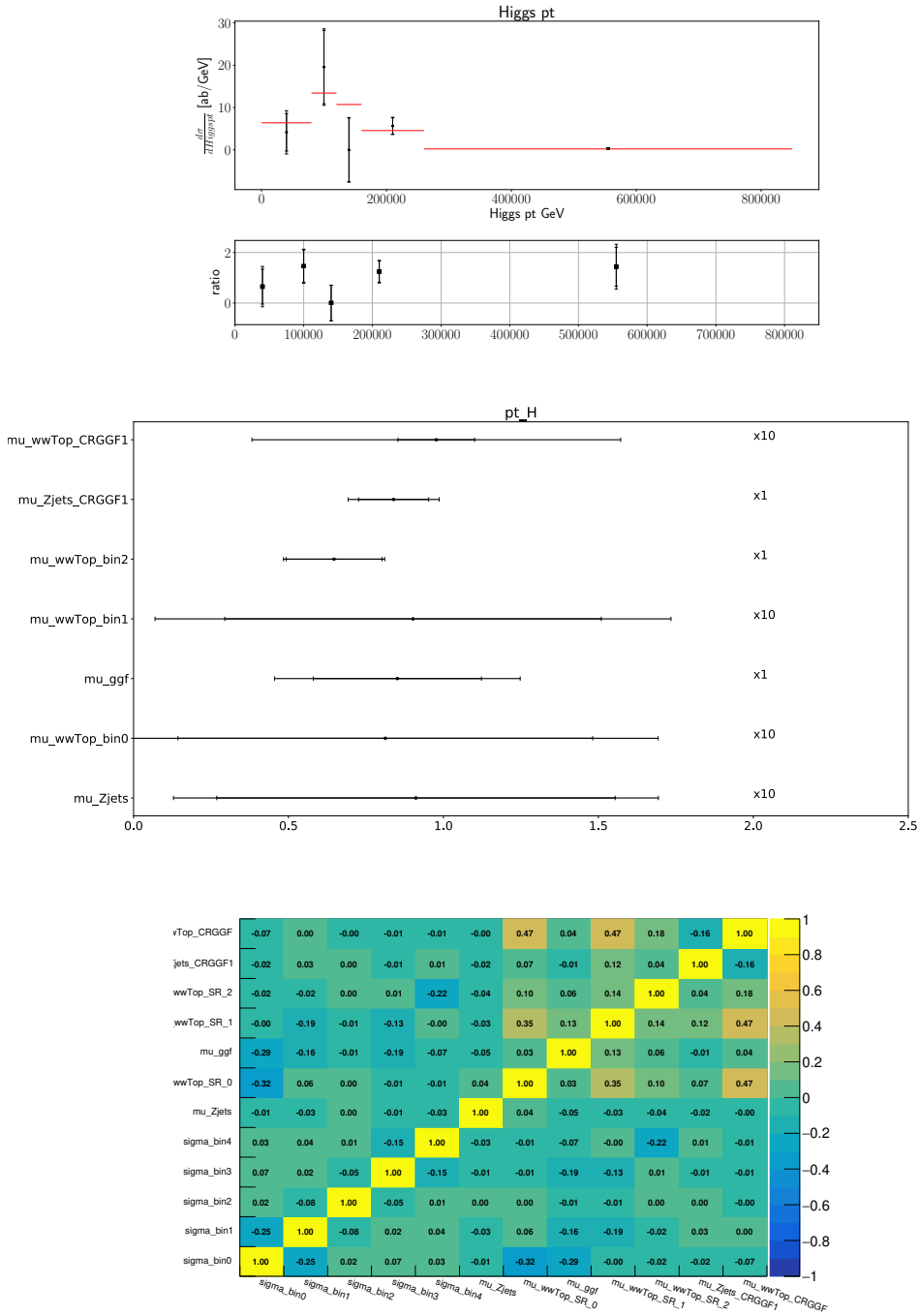


Figure 17.7. Higgs Pt Full stat+sys fit to data

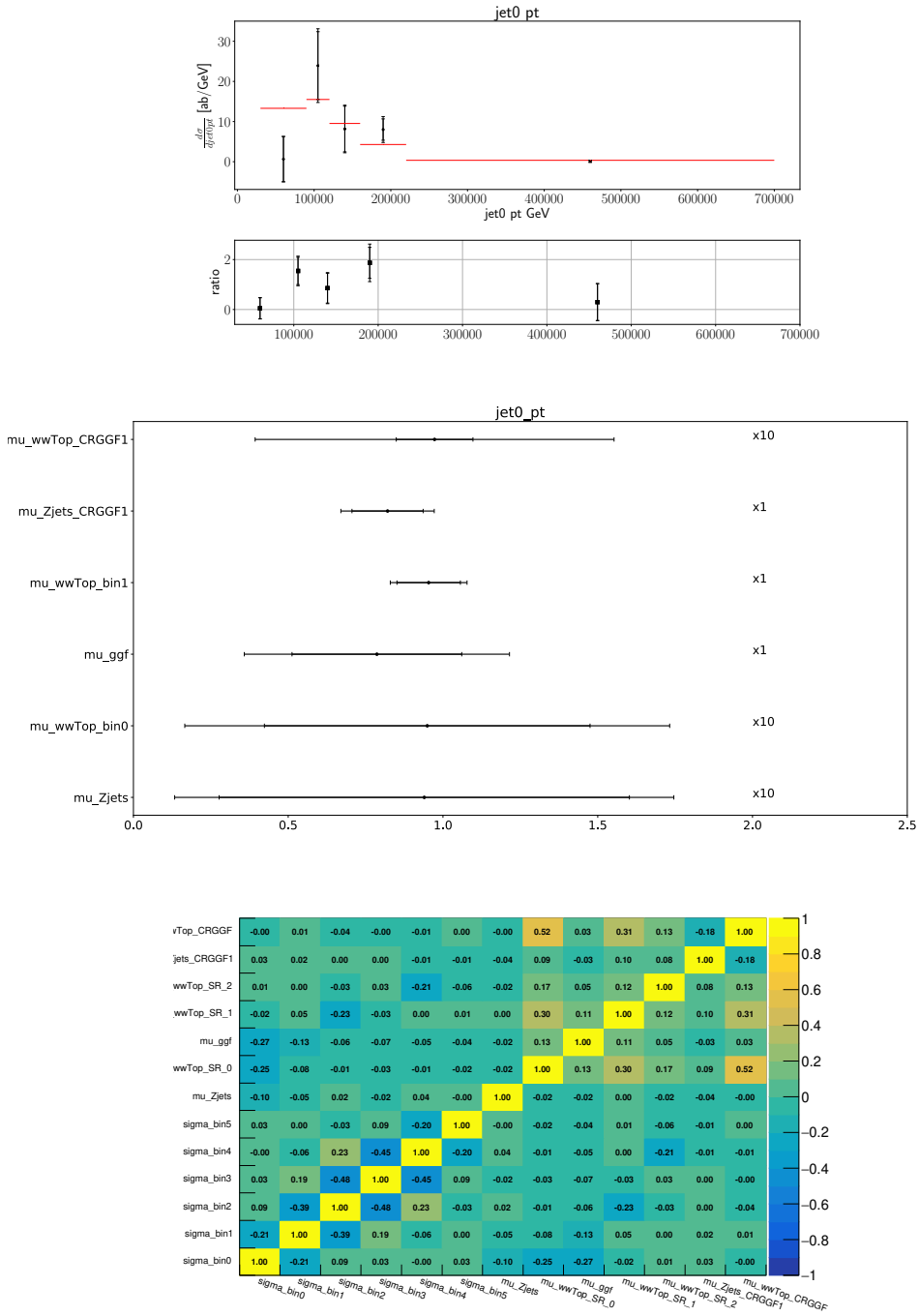


Figure 17.8. Jet0 Pt Full stat+sys fit to data

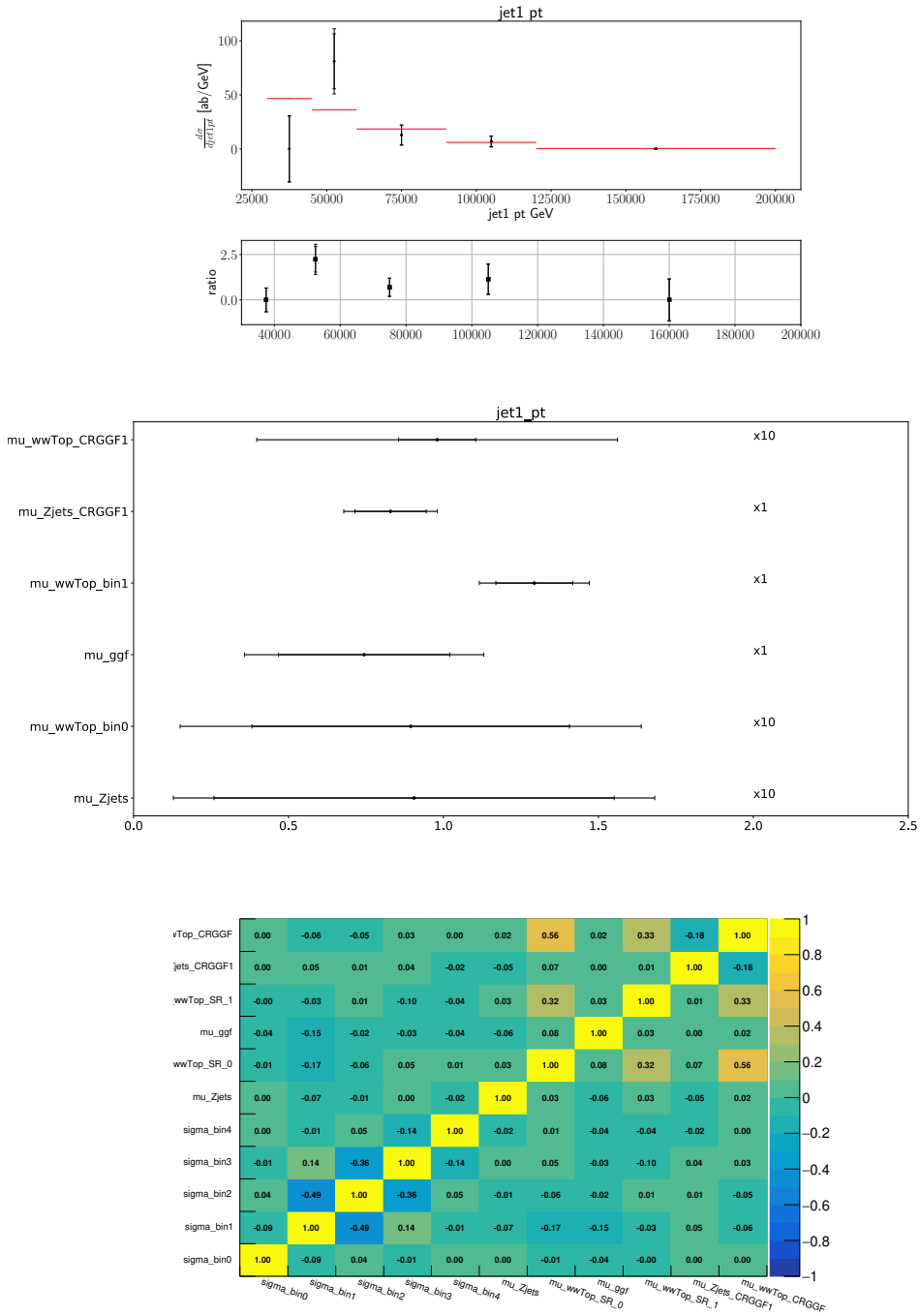


Figure 17.9. jet1 pt Full stat+sys fit to data

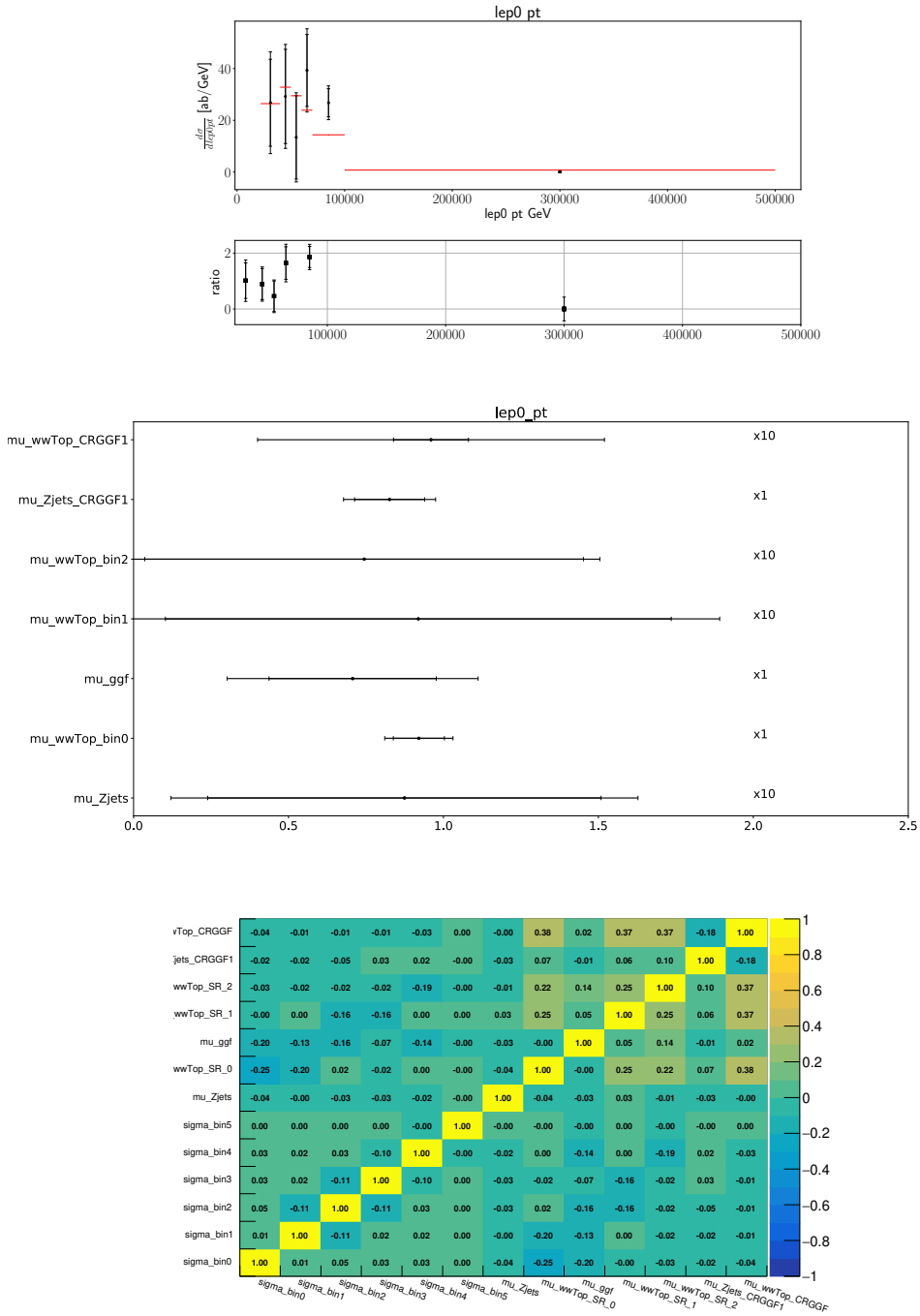


Figure 17.10. lep0 pt Full stat+sys fit to data

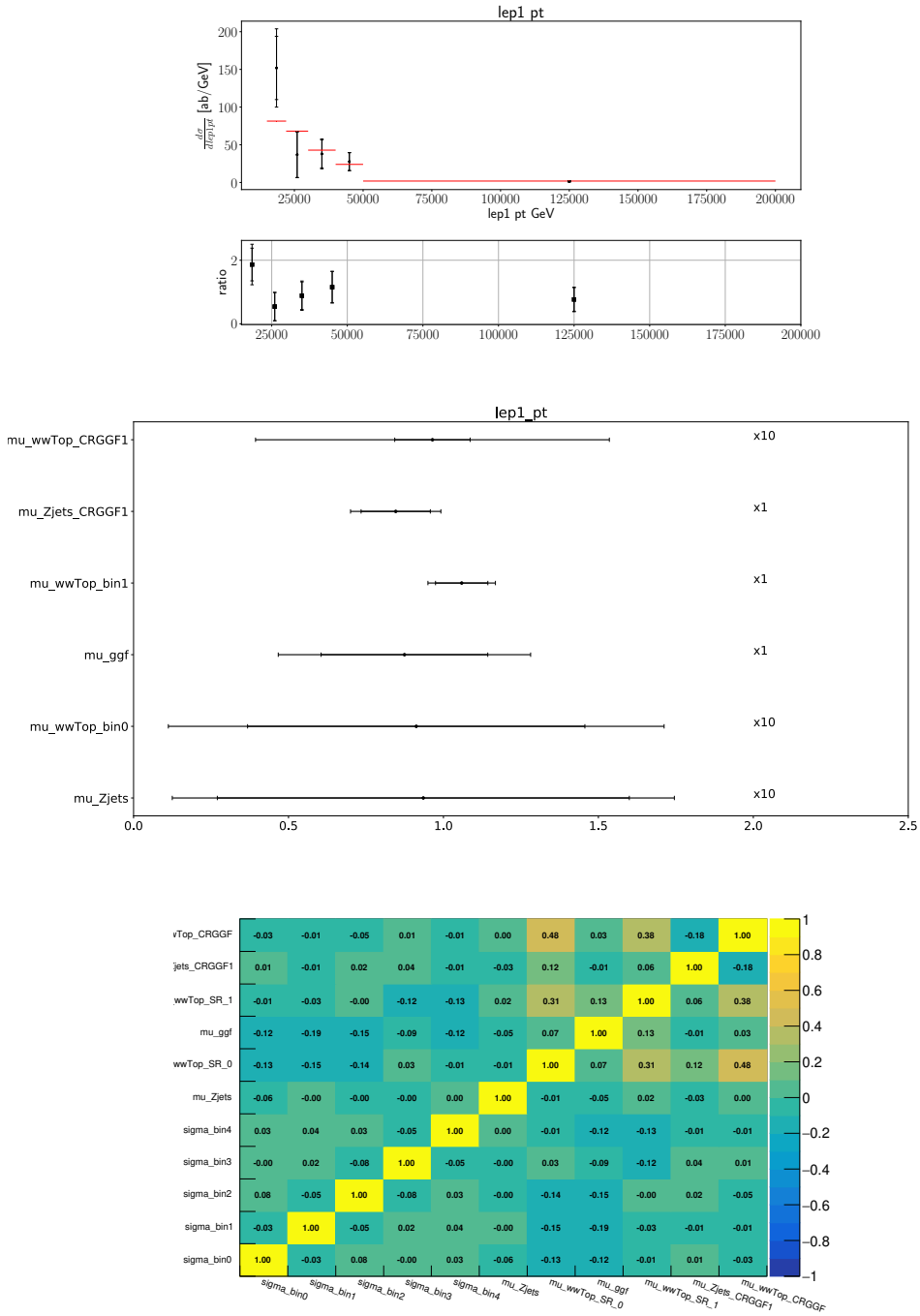


Figure 17.11. lep1 pt Full stat+sys fit to data

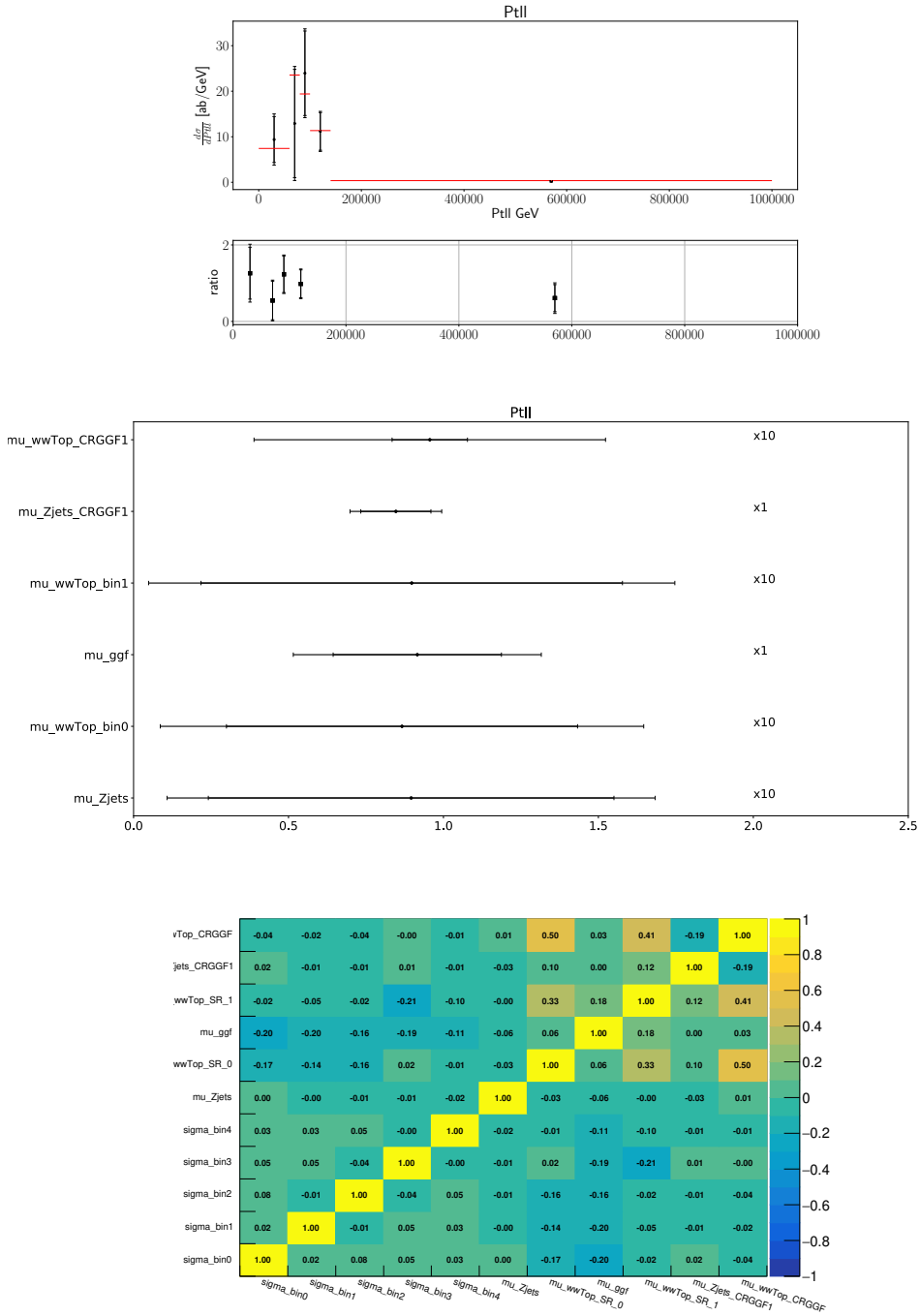


Figure 17.12. Ptll Full stat+sys fit to data

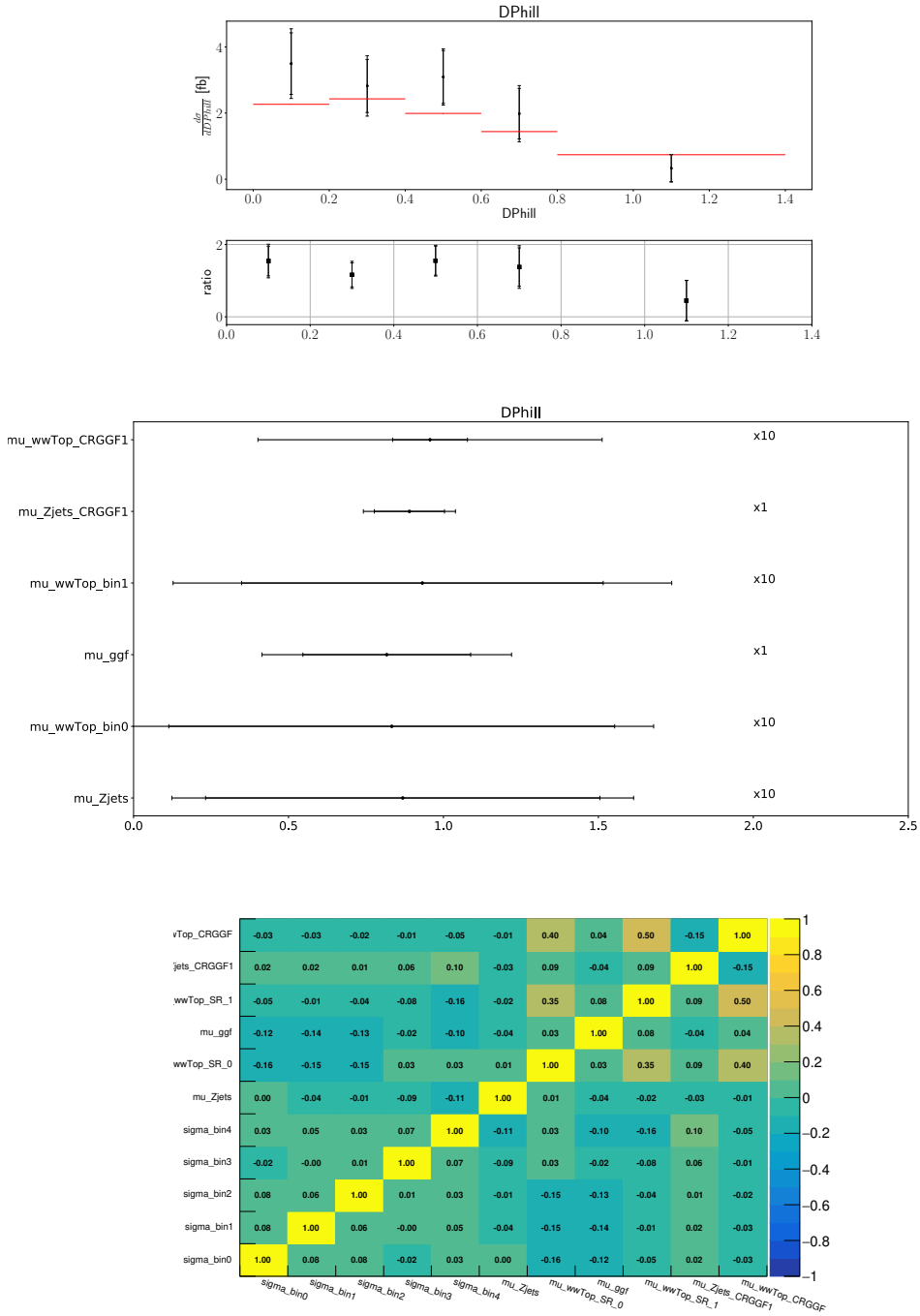


Figure 17.13. DPhill Full stat+sys fit to data

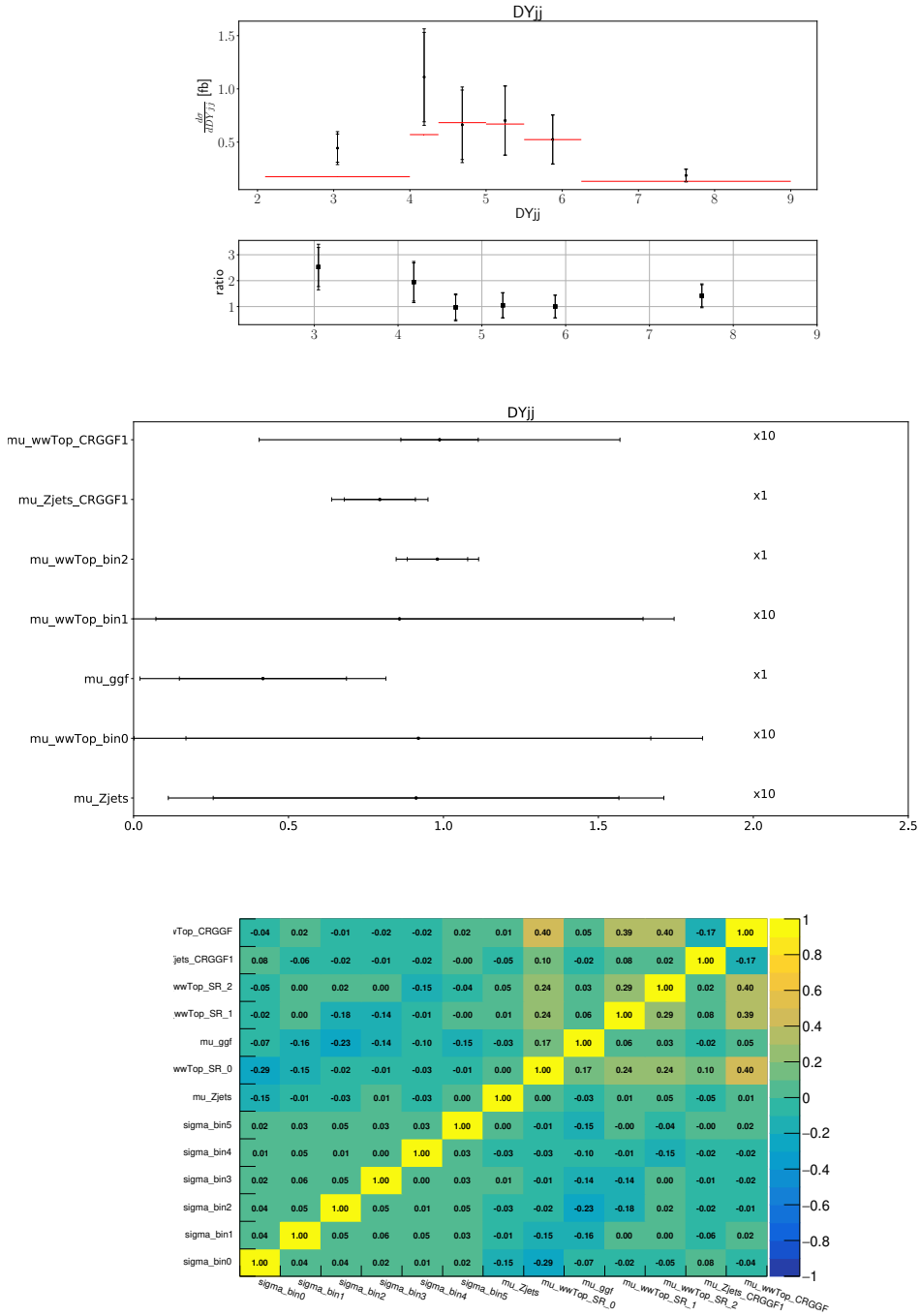


Figure 17.14. DYjj Full stat+sys fit to data

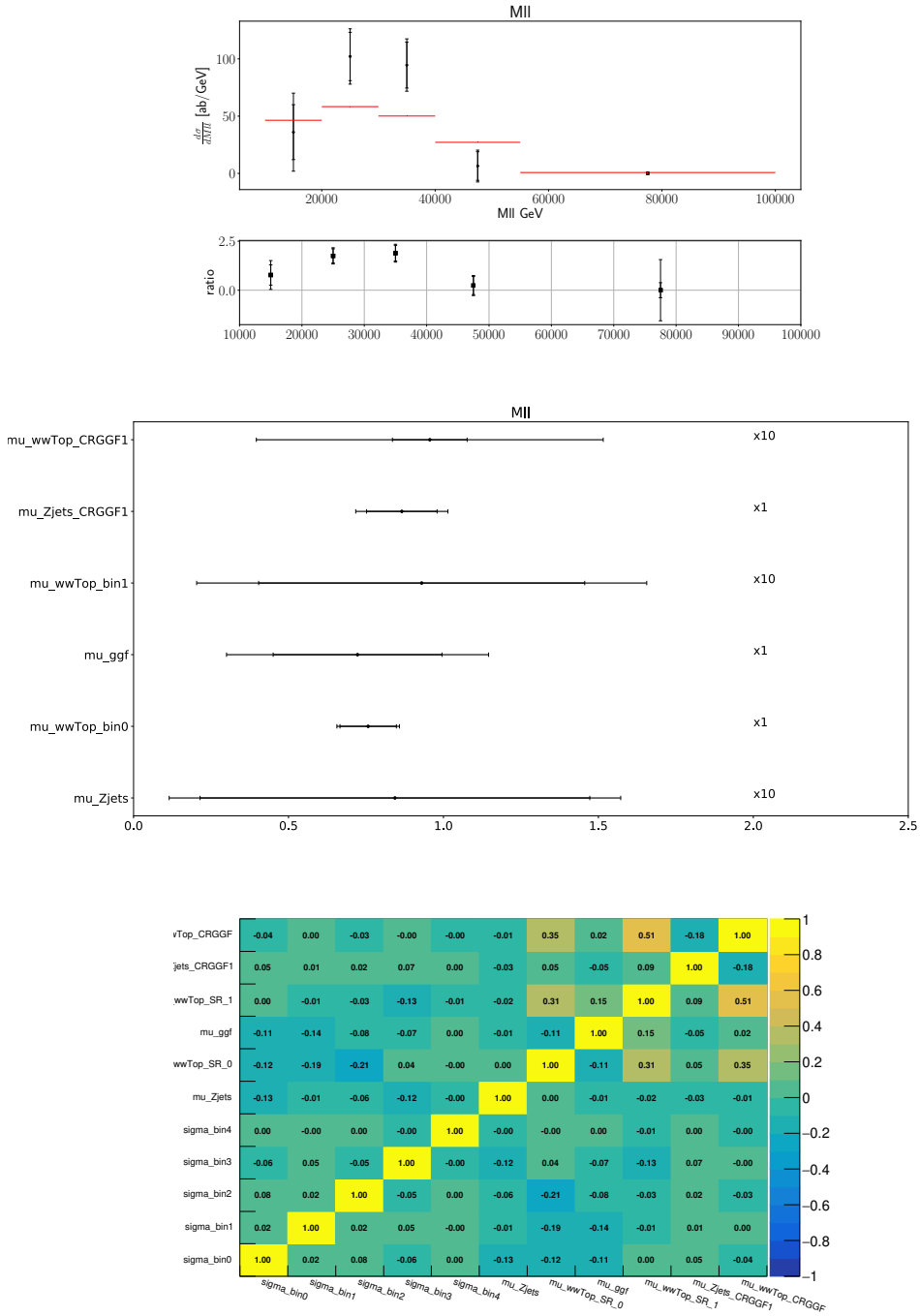


Figure 17.15. Mll Full stat+sys fit to data

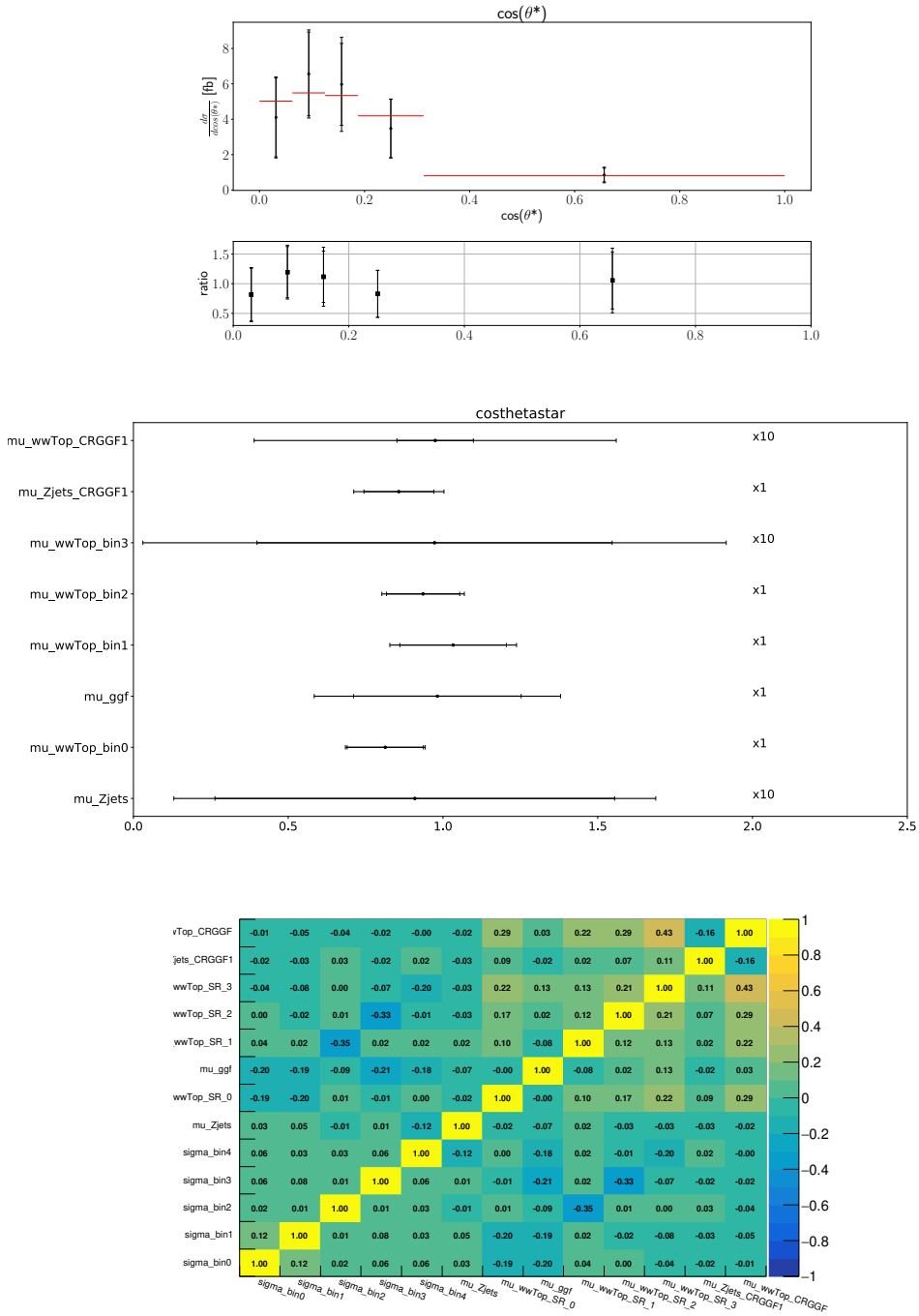


Figure 17.16. Cos(θ*) Full stat+sys fit to data

CHAPTER 18

CONLUSION

Progress on ATLAS upgrades and analysis of the cross section of Higgs boson to WW decay are presented. Brookhaven National Lab (BNL) will be responsible for producing about a quarter of the silicon strip modules in the united states, running quality control tests on all US produced modules and assembling staves. Harvard and BNL collaborated on a cold box to conduct QC tests on four modules. The first iteration of the box with the software framework was successfully tested with four ABC130 modules. Results from ABCStar radiation show the expected bump in the leakage current from ionizing radiation.

The results of the radiation tests agree with expectations from the physics model. Higher dosage leads to more leakage current, because the annealing positive states are quickly replaced with newly created positive states. Higher temperature leads less leakage current, because positive states anneal faster. The Peak of the leakage current occurs at 600 Krad regardless of dosage or temperature. This is true for low dosage, around 1Krad/hour. For high dosage, such as the results from RAL at the order 1 Mrad/hour, the peak is closer 1 Mrad. Calculating a consistent value for the total leakage current as a function of temperature and dosage is difficult because of large batch to batch variations, requiring larger statistics per batch.

The measurement of the fiducial cross section of a VBF produced Higgs boson decay to WW to opposite flavor leptons yielded results in agreement with standard model expectations, with approximate 24% uncertainty in the final measurement, compared to 50% uncertainty in the previous measurement. The cross section as

a function of thirteen important observables was also presented, with no significant disagreement with the standard model expectations.

The VBF production mode is rare, with only 7% of all the Higgs bosons produced at the LHC. This makes getting large statistics for this production mode difficult. The uncertainty is still dominated by statistical uncertainty, so the methods developed in this analysis will be an important starting point for future iterations in high luminosity LHC which will benefit from an order of magnitude more data. Theory uncertainty is dominated by the $Z \to \tau\tau$ process, which will benefit, even possibly in the near future, from increased Monte Carlo statistics. In addition, a combined measurement with the ggF and other Higgs processes will give further possibilities for probing new physics with precision measurements.

The significant backgrounds in the analysis are from standard model WW,$t\bar{t}$, Drell Yan, ggF Higgs, and fake leptons. Fake leptons are estimated in a data driven way in the W+Jets control region using the ratio of leptons identified correctly with strict requirements and loose requirements. The other major standard model backgrounds are estimated in the signal region with a statistical fit, aided by multivariate discriminants. This method resulted in an estimated fiducial cross section of 1.7 ± 0.42 fb, and 4.7σ significance, compared to 1.8σ significance of the previous measurement. This agrees with the standard model expectation of 2.07 fb.

The analysis, with increased precision, finds no significant deviations from the standard model. Further reductions in uncertainty are expected soon in run 3 and beyond, both because of higher statistics and increasingly sophisticated usage of new algorithms like multivariate discriminants.

BIBLIOGRAPHY

[1] A. Keshavarzi, K. Khaw, T. Yoshioka, *Muon g-2: A review*, 2021, arXiv:2106.06723 [hep-ex]

[2] Griffiths, D. J,(2008) *Introduction to elementary particles; 2nd rev. version*, Wiley. x

[3] K. Dette, for the ATLAS Pixel Collaboration, *Total Ionising Dose effects in the FE-I4 front end chip of the ATLAS pixel IBL detector.*, JINST 11 (2016) C11028

[4] M. Backhaus, *Parametrization of the radiation induced leakage current increase of NMOS transistors*, 2016, doi:10.1088/1748-0221/12/01/P01011, https://arxiv.org/abs/1610.01887

[5] K. Bierwagen, U. Blumenschein, A. Quadt, *Bayesian Unfolding*, Georg-August-University Göttingen, https://indico.cern.ch/event/107747 contributions/32648/attachments/24323/35007/bierwagen.pdf

[6] ATLAS Collaboration,*Technical Design Report for the ATLAS ITk-Strips Detector*, ATL-COM-UPGRADE-2016-033, 2016. https://cds.cern.ch/record/2220769

[7] ATLAS Collaboration,*Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV*, Aad, G., Abbott, B. et al. J. High Energ. Phys. 2016:45. doi:10.1007/JHEP08(2016)045 . https://arxiv.org/pdf/1606.02266v2.pdf

[8] G. D'Agostini, *Improved iterative Bayesian unfolding*,2010, https://arxiv.org/abs/1010.0632

[9] S.Stuci, *ATLAS ITk Strip Detector for High Luminosity LHC*, 2018,https://cds.cern.ch/record/2626900?

[10] Campanelli, M., Pilkington, A., Robinson, J.E.M. , *Measurement of the production and azimuthal decorrelation of dijets with a jet veto in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, 2018, https://cds.cern.ch/record/1473139/files/ATL-COM-PHYS-2012-1220.pdf?

[11] ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at* $\sqrt{s} = 7$ *and 8 TeV in the ATLAS experiment*, Eur. Phys. J. C76 (2016) 6, DOI: 10.1140/epjc/s10052-015-3769-y

[12] S. Chatrchyan et al., *Observation of a New Boson with Mass Near 125 GeV in pp Collisions at* $\sqrt{s} = 7$ *and 8 TeV*, GJHEP 06 (2013) 081, arXiv: 1303.4571 [hep-ex]

[13] A. M. Sirunyan et al.,*Measurements of properties of the Higgs boson decaying to a W boson pair in pp Collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Lett. B 791 (2019) 96, arXiv: 1806.05246 [hep-ex]

[14] G. Aad et al.,*Observation and measurement of Higgs boson decays to WW* with the ATLAS detector*,Phys. Rev. D 92 (2015) 012006, arXiv: 1412.2641 [hep-ex]

[15] A. Banfi, S. Redford, M. Vesterinen, P. Waller and T. R. Wyatt,*Optimisation of variables for studying dilepton transverse momentum distributions at hadron colliders*,Eur. Phys. J. C 71 (2011) 1600, arXiv: 1009.1580 [hep-ex] (cit. on p. 12).

[16] S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group, ed.* by S. Heinemeyer,CERN Yellow Reports,2013, url: http://cds.cern.ch/record/1559921

[17] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D 98 (3 2018) 030001, https://link.aps.org/doi/10.1103/PhysRevD.98.030001

[18] D. de Florian et al.,*Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*,2/2017 (2016), arXiv: 1610.07922 [hep-ph]

[19] K. Arnold et al.,*Vbfnlo: A parton level Monte Carlo for processes with electroweak bosons*,Computer Physics Communications 180 (2009) 16611670, issn: 0010-4655,url: http://dx.doi.org/10.1016/j.cpc.2009.03.006

[20] G. P. Lepage, *VEGAS - an adaptive multi-dimensional integration program, tech. rep. CLNS-447*,Cornell Univ. Lab. Nucl. Stud., 1980, url: http://cds.cern.ch/record/123074

[21] J. Butterworth et al.,*Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report, 2014*,arXiv: 1405.1067 [hep-ph]

[22] *Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector*, tech. rep. ATL-PHYS-PUB-2019-007, CERN, 2019,url: https://cds.cern.ch/record/2665057

[23] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D 93 (3 2016) 033006, url: https://link.aps.org/doi/10.1103/PhysRevD.93.033006

[24] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*,The European Physical Journal

C 75 (2015), issn: 1434-6052, url: http://dx.doi.org/10.1140/epjc/s10052-015-3397-6.

[25] R. D. Ball et al., Parton distributions for the LHC run II, Journal of High Energy Physics 2015 (2015), issn: 1029-8479, url: http://dx.doi.org/10.1007/JHEP04(2015)040

[26] K. Assamagan et al., *The ATLAS Monte Carlo Project*, tech. rep. ATL-SOFT-INT-2010-002, CERN, 2010, url: https://cds.cern.ch/record/1239901 (cit. on p. 36).

[27] P. Nason and C. Oleari,*NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*,JHEP 02 (2010) 037, arXiv: 0911.5299 [hep-ph]

[28] M. Ciccolini, A. Denner and S. Dittmaier,*Strong and electroweak corrections to the production ofHiggs + 2-jets via weak interactions at the LHC*, Phys. Rev. Lett. 99 (2007), arXiv: 0707.0381 [hep-ph]

[29] P. Bolzoni et al.,*Higgs production via vector-boson fusion at NNLO in QCD*,Phys. Rev. Lett. 105 (2010), arXiv: 1003.4451 [hep-ph]

[30] J. Butterworth et al.,*PDF4LHC recommendations for LHC Run II*, J. Phys. G43 (2016),arXiv: 1510.03865 [hep-ph]

[31] Harland-Lang, L. A., Martin, A. D., Motylinski, P. and Thorne, R. S.,*Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C 75 (2015) 204,url: https://doi.org/10.1140/epjc/s10052-015-3397-6

[32] A. Djouadi, J. Kalinowski and M. Spira,*HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput. Phys. Commun. 108 (1998) 56, arXiv: hep-ph/9704448

[33] A. Bredenstein, A. Denner, S. Dittmaier and M. Weber,*Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \to WW/ZZ \to 4$ fermions*, JHEP 02 (2007) 080, arXiv: hep-ph/0611234

[34] 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, arXiv: 1405.0301 [hep-ph]

[35] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*,JHEP 11 (2004) 040, arXiv: hep-ph/0409146

[36] 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,arXiv: 1002.2581 [hep-ph]

[37] J. Butterworth et al.,*PDF4LHC recommendations for LHC Run II*, J. Phys. G 43 (2016) 023001, arXiv: 1510.03865 [hep-ph]

[38] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*,JHEP 11 (2007) 070, arXiv: 0709.2092 [hep-ph]

[39] K. Hamilton, P. Nason, E. Re and G. Zanderighi,*NNLOPS simulation of Higgs boson production*,JHEP 10 (2013) 222, arXiv: 1309.0017 [hep-ph]

[40] K. Hamilton, P. Nason and G. Zanderighi,*Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, JHEP 05 (2015) 140, arXiv: 1501.04637 [hep-ph]

[41] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*,JHEP 10 (2012) 155, arXiv: 1206.3572 [hep-ph]

[42] J. M. Campbell et al., *NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM*, JHEP 07 (2012) 092, arXiv: 1202.5475 [hep-ph]

[43] K. Hamilton, P. Nason, C. Oleari and G. Zanderighi,*Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, JHEP 05 (2013) 082, arXiv: 1212.4504 [hep-ph]

[44] S. Catani and M. Grazzini,*An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, Phys. Rev. Lett. 98 (2007) 222002, arXiv: hep-ph/0703012

[45] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*x, JHEP 05 (2016) 058, arXiv: 1602.00695 [hep-ph]

[46] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, Phys. Rev. Lett. 114 (2015) 212001, arXiv: 1503.06056 [hep-ph]

[47] F. Dulat, A. Lazopoulos and B. Mistlberger,*iHixs 2 — Inclusive Higgs cross sections*, Comput. Phys. Commun. 233 (2018) 243, arXiv: 1802.00827 [hep-ph]

[48] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, JHEP 11 (2009) 088, arXiv: 0909.3420 [hep-ph]

[49] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*,Phys. Lett. B 679 (2009) 467, arXiv: 0907.2997 [hep-ph]

[50] R. V. Harlander, H. Mantler, S. Marzani and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, Eur. Phys. J. C 66 (2010) 359, arXiv: 0912.2104 [hep-ph]

[51] A. Pak, M. Rogal and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, JHEP 02 (2010) 025, arXiv: 0911.4662 [hep-ph]

[52] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders*,Phys. Lett. B 670 (2008) 12, arXiv: 0809.1301 [hep-ph]

[53] M. Bonetti, K. Melnikov and L. Tancredi, *Higher order corrections to mixed QCD-EWcontributions to Higgs boson production in gluon fusion*, Phys. Rev. D 97 (2018) 056017, [Erratum: Phys.Rev.D 97, 099906 (2018)], arXiv: 1801.10403 [hep-ph]

[54] M. Ciccolini, S. Dittmaier and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*,Phys. Rev. D 68 (2003) 073003, arXiv: hep-ph/0306234

[55] O. Brein, A. Djouadi and R. Harlander,*NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. B 579 (2004) 149, arXiv: hep-ph/0307206

[56] T. Sjöstrand, S. Mrenna and P. Skands,*A brief introduction to PYTHIA 8.1*, Computer Physics Communications 178 (2008) 852, issn: 0010-4655, url: http://dx.doi.org/10.1016/j.cpc.2008.01.036

[57] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*,Nuclear Physics B 737 (2006) 73, issn: 0550-3213, url: http://dx.doi.org/10.1016/j.nuclphysb.2005.12.022

[58] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, Journal of High Energy Physics 2011 (2011), issn: 1029-8479, url: http://dx.doi.org/10.1007/JHEP11(2011)064

[59] F. Cascioli et al.,*Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production*, JHEP 01 (2014) 046, arXiv: 1309.0500 [hep-ph]

[60] S. Alekhin, K. Melnikov and F. Petriello, *Fixed target Drell-Yan data and NNLO QCD fits of parton distribution functions*, Phys. Rev. D 74 (5 2006) 054033, url: https://link.aps.org/doi/10.1103/PhysRevD.74.054033

[61] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through O(αS4)*,Physical Review Letters 110 (2013), issn: 1079-7114, url: http://dx.doi.org/10.1103/PhysRevLett.110.252004

[62] ATLAS Collaboration,*The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, url: https://cds.cern.ch/record/2206965

[63] A. Aggarwal et al., *Optimization of the $H \to WW*$ analysis using the full Run 2 dataset*, tech. rep. ATL-COM-PHYS-2020-303, CERN, 2020,url: https://cds.cern.ch/record/2716687

[64] ATLAS Collaboration,*Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2019-015, 2019, url: https://cds.cern.ch/record/2670380

[65] ATLAS Collaboration,*Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST 14 (2019) P12006, arXiv: 1908.00005 [hep-ex]

[66] ATLAS Collaboration, *Muon reconstruction performance in early* $\sqrt{s} = 13$ *TeV data*,ATL-PHYS-PUB-2015-037, 2015, url: https://cds.cern.ch/record/2047831

[67] ATLAS Collaboration,*Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13TeV$, Eur. Phys. J. C 76 (2016) 292, arXiv: 1603.05598 [hep-ex]

[68] M. Cacciari, G. P. Salam and G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP 04 (2008) 063, arXiv: 0802.1189 [hep-ph]

[69] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C 72 (2012) 1896,arXiv: 1111.6097 [hep-ph]

[70] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C 77 (2017) 466, arXiv: 1703.10485 [hep-ex]

[71] 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,arXiv: 1703.09665 [hep-ex]

[72] B. Brunt et al., *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, tech. rep. ATL-COM-PHYS-2016-407, CERN, 2016, url: https://cds.cern.ch/record/2149445

[73] Will Buttinger, *Updated Overlap Removal Working Points*, tech. rep.x 2015, url: https://indico.cern.ch/event/457238/

[74] T. Plehn, D. L. Rainwater and D. Zeppenfeld, *A method for identifying $H \to \tau\tau \to e\mu$ + missing $p(T)$ at the CERN LHC*, Phys.Rev. D61 (2000) 093005, arXiv: hep-ph/9911385 [hep-ph]

[75] G. Aad et al.,*Observation and measurement of Higgs boson decays to WW* with the ATLAS detector*, Phys. Rev. D 92 (2014) 012006, url: https://cds.cern.ch/record/1975394

[76] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, Phys. Lett. B463 (1999) 99, arXiv: hep-ph/9906349 [hep-ph]

[77] G. Aad et al., *Observation and measurement of Higgs boson decays to WW* with the ATLAS detector*, Phys. Rev. D 92 (1 2015) 012006, url: https://link.aps.org/doi/10.1103/PhysRevD.92.012006

[78] A. Aggarwal et al., *Measurements of the Higgs boson production cross section via ggF and VBF in $H \to WW* \to \ell\nu\ell\nu$ with 139^{-1}fb at $\sqrt{s}$ = 13 TeV*, tech. rep. ATL-COM-PHYS-2020-302, CERN, 2020,url: https://cds.cern.ch/record/2716457

[79] F. Cardillo et al., *Tools for estimating fake/non-prompt lepton backgrounds in ATLAS*, tech. rep. ATL-COM-PHYS-2019-1071, CERN, 1999,url: https://cds.cern.ch/record/2686919

[80] I. W. Stewart and F. J. Tackmann, *Theory uncertainties for Higgs mass and other searches using jet bins*, Phys. Rev. D 85 (3 2012) 034011, url: https://link.aps.org/doi/10.1103/PhysRevD.85.034011

[81] A. Denner, S. Dittmaier, P. Maierhöfer, M. Pellen and C. Schwan, *QCD and electroweak corrections to WZ scattering at the LHC*, Journal of High Energy Physics 2019 (2019), issn: 1029-8479, url: http://dx.doi.org/10.1007/JHEP06(2019)067

[82] B. Biedermann, A. Denner and M. Pellen, *Large Electroweak Corrections to Vector-Boson Scattering at the Large Hadron Collider*, Physical Review Letters 118 (2017), issn: 1079-7114, url: http://dx.doi.org/10.1103/PhysRevLett.118.261801

[83] A. Denner, R. Franken, M. Pellen and T. Schmidt, *NLO QCD and EW corrections to vector-boson scattering into ZZ at the LHC*, Journal of High Energy Physics 2020 (2020), issn: 1029-8479, url: http://dx.doi.org/10.1007/JHEP11(2020)110

[84] G. Aad et al.,*Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4l decay channel at* $\sqrt{s} = 13\ TeV$, Eur. Phys. J. C 80 (2020) 942,xarXiv: 2004.03969 [hep-ex]