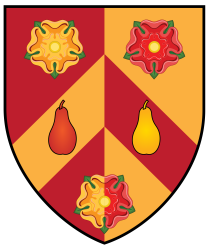
Search for heavy ZZ resonances in the $\ell^+\ell^-\nu\bar{\nu}$ channel using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC.

Mariyan Petrov
Wolfson College, Oxford





Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term, 2018

Search for heavy ZZ resonances in the $\ell^+\ell^-\nu\bar{\nu}$ channel using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC.

Mariyan Petrov
Wolfson College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term, 2018

Abstract

This thesis provides a description of the search for new resonances in event signatures consisting of two Z bosons, where one Z boson decays to a opposite-sign same-flavour lepton pair $\ell^+\ell^-$ (ℓ = electron or muon), while the other Z boson decays to neutrinos which are experimentally quantified by a large missing transverse momentum. The description of the search includes the analysis selection criteria, the background estimation and the exclusion limits on the production of new physics. The analysis is divided into two signal regions, based on the resonant production mode i.e. gluon-gluon fusion or vector boson fusion. The result of the $H \to ZZ \to \ell\ell\nu\nu$ search is combined with the result from the $H \to ZZ \to 4\ell$ analysis to arrive at better constrains of heavy Higgs boson production and to set exclusion limits on two higgs doublet models. Additionally, upper limits are derived for spin-2 Kaluza-Klein graviton resonances for masses between 600 GeV and 2 TeV in the context of the Randall-Sundrum model. The analysis search uses data gathered from proton-proton collisions at a center-of-mass energy of 13 TeV, corresponding to a total integrated luminosity of 36.1 fb^{-1} collected by the ATLAS detector during the 2015+2016 data runs at the Large Hadron Collider. No statistically significant excess over the background prediction is found in the 2015+2016 data.

Acknowledgements

I would like to start by expressing my sincere thanks to my supervisor, Professor Daniela Bortoletto for her guidance throughout the course of this thesis. It has been an honour to be a part of the $\ell\ell\nu\nu$ group at Oxford and to be your student.

Of course any long journey in science and life has many inspiring people in it and I would like to mention them here. My Master's supervisor, Dr. Nikolas Kauer, and Professor Pedro Teixeira-Dias set me on the path of particle physics, for which I am eternally grateful. Professor Ian Shipsey, Professor Cigdem Issever, Dr. Todd Huffman, Professor Tony Weidberg and Dr. Farrukh Azfar provided me with support without which I wouldn't have progressed in my first year in Oxford as fast as I did.

I owe a huge amount to my mentor whilst at CERN, Giacomo Artoni, without whom I would have been completely lost through my later research years. During my beginning steps and later years in ATLAS I had invaluable help from Teng Jian Khoo, Claire Lee, Till Eifert, Doug Schaefer and Bo Liu. I would like to thank the members of the $4\ell+\ell\ell\nu\nu$ analysis group at ATLAS, in particular Monica Trovatelli, Yusheng Wu, Hideki Okawa, Eleni Mountricha, R.D. Schaffer and Sarah Heim who had critical insight and excellent management skills without which nothing could be accomplished. To my brother-in-arms, Wen Guo, I would like to thank for aiding me in those crucial moments when everything was hanging in the balance. In addition, the analysis would not have been made possible without the help of Yuri Naryshkin, Artem Basalaev, Cong Geng, Gabriele Sabato, Alison Elliot, Kenji Hamano and Xiangyang Ju.

During my years as a DPhil student I have met many people who brought me both moments of tremendous joy and true inspiration. In no particular order I would like to thank the wonderful and colourful characters of Luigi Vigani, Luigi Marchese, Francesco Guili, Miha Zgubic, Kathrin Becker, Lydia Beresford, Katherine Pachal, Jesse Liu, Mike Nelson, Jon Burr, Will Fawcett, Luca Ambroz, Cecilla Tosciri, Gabija Zemaityte, Santiago Paredes, Maria Giovanna Foti, Martina Pili, Aaron O'Neill, Kla Karava, and Cees Carels. I would also like to thank Jack Wright for his patience with me during our time in CERN. Furthermore, the staff at Oxford, especially Sue Geddes and Kim Proudfoot, have been exceptional and helped me a lot during my stay.

I would like to express my gratitude to a dear friend of mine, Nurfikri Norjoharuddeen, whose boundless imagination and tireless pursuit of knowledge has always inspired me to be better. You have helped my during some difficult moments in my life and for that I am eternally indebted to you.

Lastly, I would like to thank my parents, my brother and sister, and my relatives without whom I would never had the courage to follow my dreams and be at the place where I am today.

Preface

The studies outlined in this thesis, have been produced in a collaborative effort with other ATLAS members. The goal of this thesis is to present my personal contribution to this research topic. I performed several optimization methods for the analysis, estimated a few of the background contributions and their uncertainties, and produced the exclusion limits of the analysis.

In chapter 4, my studies in the missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) sector are presented. My personal contributions consist of multiple studies on the observed and expected $E_{\mathrm{T}}^{\mathrm{miss}}$ performance in the ATLAS detector via various data/MC and diagnostic plots. I prepared a poster for LHCC [1], some of the public plots of $E_{\mathrm{T}}^{\mathrm{miss}}$ [2,3] and provided studies and figures for the 2015 $E_{\mathrm{T}}^{\mathrm{miss}}$ paper [4]. I am a contact editor for the $E_{\mathrm{T}}^{\mathrm{miss}}$ conference note [5].

In chapter 5, an overview of the $H \to ZZ \to \ell\ell\nu\nu$ analysis search is provided. I was involved with the studies of the vector boson fusion (VBF) Higgs boson production mode but I also made several contributions in the gluon gluon fusion (ggF) sector of the analysis. I ran an optimization procedure to better isolate the VBF topology from various background processes, including the reduction of the ggF signal yield. I studied the modelling of the analysis variables, and reduced the contribution of pile-up jets in the VBF event signal region. I introduced a variable, $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$, to suppress Z + jets background production. Additionally, I showed that it improves the expected exclusion limits. I worked towards the estimation of the background contribution of the ZZ and Z + jets processes and their experimental/theoretical uncertainties in the VBF region.

In chapter 7, I performed the limit setting for the ggF and VBF $H \to ZZ \to \ell\ell\nu\nu$ analysis search. I made ranking plots of systematics and pre-/post-fit plots for different mass points. I also derived the limit setting for a spin-2 Graviton interpretation. I was involved in the combination of the $4\ell + \ell\ell\nu\nu$ analyses by preparing the analysis inputs for the combination framework implemented by the $H \to ZZ \to 4\ell$ analysis. The results of the combination are presented in the $4\ell + \ell\ell\nu\nu$ high mass paper [6].

In chapter 8, I produce and discuss performance plots for $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated with an alternative

jet reconstruction method and its benefits for the analysis. In addition, I am also involved in the studies of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, for which I have produced various data/MC comparisons. I am a contact editor for an upcoming $E_{\mathrm{T}}^{\mathrm{miss}}$ significance conference note [7].

All of the figures were produced from my own framework unless otherwise indicated in the captions. Please note that in this thesis figures with the label 'ATLAS' have been published in papers. Any results which were not produced by the $H \to ZZ \to \ell\ell\nu\nu$ analysis team or are published in papers are cited appropriately. Natural units are utilised throughout, $c = \hbar = 1$.

Contents

1 Introduction **1**

2 The Standard Model and its extensions **3**

2.1 The Standard Model of Particle Physics 3

2.2 Electro-weak symmetry breaking 5

2.3 Properties of the Higgs boson 8

2.3.1 SM Higgs boson production 8

2.3.2 SM Higgs boson decay 10

2.3.3 Higgs particle for large values of the Higgs boson mass 11

2.4 Extensions of the Standard Model 12

2.4.1 Shortcomings of the SM 12

2.4.2 Beyond the Standard model 12

2.5 Two Higgs Doublet Models 13

2.6 Gravitons 14

2.7 Run I results 15

2.7.1 $H \to ZZ \to \ell\ell\nu\nu$ analysis 15

2.7.2 Combination with other $H \to ZZ$ decay modes 15

3 The ATLAS Experiment **20**

3.1 The Large Hadron Collider 20

3.2 The ATLAS detector 22

3.3 Inner Detector 24

3.4 Calorimeter system 27

3.4.1 Electromagnetic calorimeters 28

3.4.2 Hadron calorimeters 28

3.5 Muon Spectrometer 30

3.6 Trigger systems 32

3.6.1 Single lepton triggers examples 34

3.7 Data 34

3.7.1 Luminosity 34

3.7.2 ATLAS Data 37

3.8 ATLAS event data model and simulation . . . 38

4 Physics object reconstruction 40

4.1 Tracks . . . 41

4.2 Physics objects . . . 41

4.2.1 Electrons . . . 42

4.2.2 Muons . . . 43

4.2.3 Jets . . . 44

4.2.4 Overlap removal . . . 48

4.3 Missing Transverse Momentum ($\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$) . . . 48

4.3.1 $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction . . . 49

4.3.2 $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ performance evaluation in the di-lepton final state . . . 52

4.3.3 Alternative $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction . . . 60

5 Analysis methodology 65

5.1 The significance of the $H \to ZZ \to \ell\ell\nu\nu$ final state . . . 65

5.2 Analysis Strategy . . . 66

5.2.1 Object Selection . . . 67

5.2.2 Event Pre-Selection . . . 69

5.2.3 Trigger Menus . . . 69

5.2.4 $H \to ZZ \to \ell\ell\nu\nu$ event selection . . . 70

5.2.5 $qq \to qqH$ signal region . . . 73

5.2.6 $gg \to H$ signal region . . . 75

5.2.7 Effects of selections . . . 76

5.3 MC Samples . . . 76

5.3.1 Signal samples . . . 76

5.3.2 Background samples . . . 77

5.4 Modelling . . . 78

5.4.1 $H \to ZZ \to \ell\ell\nu\nu$ analysis observables . . . 79

5.4.2 VBF-related observables . . . 80

5.4.3 Modelling of $m_{\mathrm{T}}^{\mathrm{ZZ}}$ in the $H \to ZZ \to \ell\ell\nu\nu$ signal region . . . 89

6 Background estimation and systematic uncertainties 91

6.1 Background estimation . . . 91

6.1.1 Overview . . . 91

6.1.2 ZZ background . . . 92

6.1.3 WZ background . . . 98

6.1.4 $t\bar{t}, WW, Wt$, and $Z \to \tau\tau$ backgrounds . . . 103

6.1.5 Z+jets background . . . 108

6.1.6 W+jets, ttV, tri-boson backgrounds . . . 119
6.2 Systematic Uncertainty . . . 120
6.2.1 Experimental uncertainties . . . 121
6.2.2 Theoretical uncertainties . . . 126

7 Results **128**
7.1 Yields in the signal region . . . 128
7.2 Statistical interpretation . . . 131
7.2.1 Likelihood function . . . 132
7.2.2 Nuisance parameters . . . 133
7.2.3 Fitting procedure . . . 134
7.2.4 Analysis inputs . . . 135
7.3 Test statistic and upper limits . . . 135
7.3.1 Observed exclusion limits . . . 136
7.4 Nuisance parameter pulls and rankings . . . 137
7.5 Post-fit impact on the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape and yield . . . 140
7.6 Combination with 4ℓ . . . 146
7.6.1 $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ analysis and results . . . 149
7.6.2 Observed excesses . . . 151
7.6.3 Statistical interpretation and combination . . . 152
7.6.4 Combined upper limits . . . 153
7.6.5 Ranking of systematic uncertainties . . . 154
7.6.6 Exclusion limits on 2HDMs . . . 154
7.6.7 Upper limits for Graviton production . . . 159

8 Future analysis developments **160**
8.1 PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ . . . 160
8.1.1 The particle-flow technique . . . 161
8.1.2 The particle-flow jets within the $E_{\mathrm{T}}^{\mathrm{miss}}$ framework . . . 162
8.1.3 PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ in $H \to ZZ \to \ell\ell\nu\nu$ analysis . . . 162
8.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance . . . 164

9 Conclusions and Outlook **169**

A More $\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$ distributions **171**

B 2D significance scans **173**

C Signal and background distributions **175**

D MC samples **177**

E Modelling of $p_{\mathrm{T}}^{\mathrm{Z}}$ **179**

F Efficiency of applied selections on signal and background **181**

G Pile-up jet events **185**

H 3ℓ control region plots **187**

I Efficiency factor binned in p_{T}, η **188**

J Z + jets bias checks **189**

K One side band $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ modelling **191**

L Theoretical Uncertainties **193**

M Upper limits **194**

N Denominator of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance **197**

Bibliography **200**

List of Figures

2.1 Summary of standard Model particles 5
2.2 Higgs potential 7
2.3 Higgs particle production cross section evolution 9
2.4 Higgs boson production modes 10
2.5 Higgs boson branching fraction for different decay modes 11
2.6 The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distributions in the final signal region in the ggF category. 16
2.7 Upper limits on ggF/VBF Higgs production cross section for $\ell\ell\nu\nu$ analysis . 16
2.8 Combined upper limits on ggF/VBF Higgs production cross section as a function of m_{H} 17
2.9 Exclusion contours in the 2HDM for $\tan\beta$ vs m_{H} 18
2.10 Exclusion contours in the 2HDM for $\tan\beta$ vs $\cos(\beta-\alpha)$ 19

3.1 CERN Hydgrogen source 21

3.2 CERN accelerator complex . . . 22
3.3 ATLAS detector . . . 23
3.4 ATLAS Inner detector . . . 25
3.5 Inserable B-Layer . . . 26
3.6 ATLAS Calorimeter systems . . . 27
3.7 Electromagnetic barrel calorimeter and the pre-sampler. . . . 29
3.8 Tile barrel calorimeter design . . . 30
3.9 Muon spectrometer system and toroidal magnet system. . . . 31
3.10 Trigger efficiency rates for $Z \to ee$ . . . 35
3.11 Trigger efficiency rates for $Z \to \mu\mu$ . . . 36
3.12 Integrated luminosity as a function of time for different periods . . . 38
3.13 Pile-up profile in run II, measured with the ATLAS detector . . . 39

4.1 Schematic of the anti-k_t algorithm for a radius parameter $R = 1.0$. . . . 46
4.2 Inclusive $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ . . . 53
4.3 Inclusive soft $E_{\mathrm{T}}^{\mathrm{miss}}$ component distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ . . . 53
4.4 Inclusive jet $E_{\mathrm{T}}^{\mathrm{miss}}$ component distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ . . . 54
4.5 Inclusive lepton $E_{\mathrm{T}}^{\mathrm{miss}}$ component distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ . . . 54
4.6 $E_{\mathrm{T}}^{\mathrm{miss}}$ diagnostics: $E_{\mathrm{T}}^{\mathrm{miss}}$ scale vs $p_{\mathrm{T}}^{\mathrm{Z}}$ . . . 55
4.7 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the number of primary vertices . . . 56
4.8 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the N_{PV} in a jet veto category . . . 57
4.9 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the average number of interactions $\langle\mu\rangle$ . . . 57
4.10 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the $\sum E_T$ . . . 58
4.11 $E_{\mathrm{T}}^{\mathrm{miss}}$ tail fraction . . . 59
4.12 $E_{\mathrm{T}}^{\mathrm{miss}}$ comparison for all definitions in data . . . 61
4.13 $E_{\mathrm{T}}^{\mathrm{miss}}$ comparison for all definitions in Z + jets . . . 62
4.14 RMS resolution in data for different reconstruction algorithms . . . 63
4.15 $E_{\mathrm{T}}^{\mathrm{miss}}$ tail fraction . . . 64
4.16 $E_{\mathrm{T}}^{\mathrm{miss}}$ response as a function of Z boson p_{T} . . . 64

5.1 Scans of the significance, Z, for increasing $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cut-values in the $H \to ZZ \to ee\nu\nu$ (left) and $H \to ZZ \to \mu\mu\nu\nu$ (right) channels. . . . 73
5.2 Variation of the combined expected limits with increasing $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cut for signal mass points between 300 GeV and 1 TeV. . . . 74
5.3 Kinematic distributions that are used for event selections n the $H \to ZZ \to ee\nu\nu$ channel after applying the Z mass window cut (1) . . . 81
5.4 Kinematic distributions that are used for event selections in the $H \to ZZ \to \mu\mu\nu\nu$ channel after applying the Z mass window cut (1) . . . 82

5.5 Kinematic distributions that are used for event selections in the $\ell\ell\nu\nu$ analyses in the $ee\nu\nu$ (top) and $\mu\mu\nu\nu$ (bottom) channels after applying the Z mass window cut in the $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ and Number of $b-$jets distribution 83
5.6 Transverse mass distribution before any analysis specific selections 84
5.7 Kinematic distributions that are used for event selections in the $\ell\ell\nu\nu$ analyses in the $ee\nu\nu$ channel after applying Z mass window cut (2) 85
5.8 Kinematic distributions that are used for event selections in the $\ell\ell\nu\nu$ analyses in the $\mu\mu\nu\nu$ channel after applying Z mass window cut (2) 86
5.9 Significance scan of the jet p_{T} after all inclusive selections for a range of signal mass points. 87
5.10 Jet pseudorapidity before/after PU jet removal 88
5.11 Removal of pile-up jets 89
5.12 Transverse mass distribution in the signal region before any corrections . . . 90

6.1 Feynman diagram for the production of the irreducible ZZ background . . . 93
6.2 The derived NNLO QCD k-factors binned in m_{ZZ} for the $q\bar{q} \to ZZ$ process. 94
6.3 Detector-level distribution of the transverse mass $m_{\mathrm{T}}^{\mathrm{ZZ}}$ in the signal region with applied $k-$factor correction. 95
6.4 The m_{jj} and $|\Delta\eta_{jj}|$ comparison distributions between `Powheg` and `Sherpa` simulated samples. 96
6.5 The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distributions of `Sherpa` and `Powheg` samples with added uncertainties in the combined $H \to ZZ \to \ell\ell\nu\nu$ channel. 98
6.6 Differences in the modelling of jet distributions in different MC samples. . . 99
6.7 Data/MC modelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the 3ℓ control region. 101
6.8 Comparison between data and simulation after a $|m_Z - m_{e\mu}| < 15$ GeV cut in the $e\mu$ Control Region for the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. 105
6.9 $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape for the $e\mu$ background in the High Mass SR for electrons 107
6.10 $m_{\mathrm{T}}^{\mathrm{ZZ}}$ systematic shape for the $e\mu$ background in the High Mass SR 107
6.11 Scheme of the boolean ABCD method. 110
6.12 MC bias $\frac{N_A * N_D}{N_B * N_C}$ dependence on additional $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cuts for high mass analyses. 113
6.13 Bias ratio $(1+\epsilon_i)/(1+\epsilon_0)$ dependence on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ selections in var2$'$ for data and MC. 114
6.14 Scheme of the one sideband method. A selection is reversed to construct the sideband region. 116
6.15 m_{T}^{ZZ} distributions in the signal region with the full selection (black) and looser selection (red) applied. 118

6.16 m_{T}^{ZZ} distributions in the loose signal and sideband region for the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. Figure produced from an analysis collaborator. . 118
6.17 Data-driven estimation of the final m_{T}^{ZZ} distributions in the High Mass analysis for the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. 119

7.1 Each background contribution present in the $H \to ZZ \to \ell\ell\nu\nu$ signal region is visualized in the form of a pie chart. 130
7.2 Transverse mass distribution in the ggF signal region 130
7.3 95% C.L. limits on $\sigma\times\mathrm{BR}(H \to ZZ)$ for a narrow width heavy Higgs boson produced in ggF channel as a function of its pole mass. 138
7.4 95% C.L. limits on $\sigma\times\mathrm{BR}(H \to ZZ)$ for a narrow width heavy Higgs boson produced in VBF channel as a function of its pole mass 139
7.5 The nuisance parameter pulls and ranking in the case of fitting ggF signal + background model to data. 142
7.6 The nuisance parameter pulls and ranking in the case of fitting VBF signal + background model to data. 143
7.7 Correlation matrix of nuisance parameters and signal strength, for signal mass points of 300 GeV. 144
7.8 Correlation matrix of nuisance parameters and signal strength, for signal mass points of 600 GeV 145
7.9 Pre-fit (left) and Post-fit (right) plots (comparison between data and predictions (signal is added on top of SM background)) in the $ee\nu\nu$ final state. . . 147
7.10 Pre-fit (left) and Post-fit (right) plots (comparison between data and predictions (signal is added on top of SM background)) in the $\mu\mu\nu\nu$ final state. . . 148
7.11 Four-lepton invariant mass distribution for the ggF-enriched category and the VBF-enriched category. 152
7.12 The upper limits at 95% confidence level on the cross section times branching ratio for the ggF production mode($\sigma_{\mathrm{ggF}} \times BR(H \to ZZ)$) and VBF production mode ($\sigma_{\mathrm{VBF}} \times BR(H \to ZZ)$) 155
7.13 The exclusion contour in the 2HDM Type-I and Type-II models for $\cos(\beta - \alpha) = -0.1$ 157
7.14 The exclusion contour in the 2HDM Type-I and Type-II models for $m_{\mathrm{H}} = 200$ GeV 158
7.15 Limits on $\sigma \times BR(G_{\mathrm{KK}} \to ZZ)$ for a KK graviton produced with $k/\bar{M}_{Pl} = 1$. 159

8.1 Schematic of PFlow object reconstruction 161
8.2 Loose PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ 163
8.3 Loose PFlow jet $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ 163

8.4 Loose PFlow soft $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with $\mathcal{L} = 36.1\mathrm{fb}^{-1}$ 164
8.5 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance run II definition, data/MC 167
8.6 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance run I definition, data/MC 167

A.1 The distribution of ϕ^{miss}, E_x^{miss} and E_y^{miss} between data and MC 171
A.2 The distribution of ϕ^{miss}, E_x^{miss} and E_y^{miss} between data and MC 172

B.1 Scans of the significance Z for different cut values of m_{jj} and $|\Delta\eta_{jj}|$. 173
B.2 Scans of the significance Z for different cut values of m_{jj} and $|\Delta\eta_{jj}|$. 174

C.1 Distribution for the $|\Delta\eta_{jj}|$ variable 175
C.2 Distribution for the m_{jj} variable 176

E.1 Correction factor applied to $p_{\mathrm{T}}^{\mathrm{Z}}$ distribution to correct mis-modelling. 179
E.2 Z boson distribution after application of data/MC mis-modelling correction . 180

F.1 The signal acceptance versus resonance mass, for ggF heavy Higgs, for VBF heavy Higgs and for Graviton. Only the statistical uncertainty is shown. . . 181

G.1 Removal of PU jets using higher p_{T} requirement 185
G.2 Removal of pile-up jets 186

H.1 Data/MC modelling of the $p_{\mathrm{T}}, \eta, \phi$ of the third lepton in the three lepton control region. 187

J.1 Events in each region on $E_{\mathrm{T}}^{\mathrm{miss}}$ vs $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ plane, all regions combined. $H \to ZZ \to ee\nu\nu$-channel. $H \to ZZ \to ee\nu\nu$ channel. 189
J.2 Events in each region on $E_{\mathrm{T}}^{\mathrm{miss}}$ vs $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ plane, (a) - (d) for corresponding regions. 190

K.1 m_{T}^{ZZ} distributions in the sideband ($E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} < 0.4$) regions after the full selection applied for the $H \to ee\nu\nu$ channel. 191
K.2 m_{T}^{ZZ} distributions in the sideband ($E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} < 0.4$) regions after the full selection applied for the $H \to ZZ \to \mu\mu\nu\nu$ channel. 192

L.1 Theoretical systematic uncertainties on the WZ background due to QCD scale and PDF variation, as a function of the transverse mass of the WZ resonance. 193
L.2 Theoretical systematic uncertainties on the ZZ background due to QCD scale and PDF variation, as a function of the transverse mass of the ZZ resonance. 193

N.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance denominator, data/MC 197

N.2 Resolutions for the objects entering in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The plot shows the resolutions for jets, combined muons (CB μ), electrons, photons not converted to electron-positron pairs (Unc.γ) and τ-leptons hadronically decaying (Hadr.τ) in one charged hadron, zero neutral hadrons and a non detectable τ-neutrino ($1p0n$) or in three charged hadrons, one or more neutral hadrons and a non detectable τ-neutrino ($3pXn$). The curves for jets include the contribution from the additional jets produced by the pile-up, which gives the staggered shape for $|\eta| = 4.5$. Figure produced from $E_{\mathrm{T}}^{\mathrm{miss}}$ significance expert. 198
N.3 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance denominator, data/MC 198
N.4 ROC curve for $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV 199
N.5 ROC curve for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV in a di-jet system 199

List of Tables

4.1 Overlap removal criteria adopted following the ATLAS harmonisation scheme. PU jets check means "JVT> 0.59 if jet has $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$". . . 48

5.1 Summary of object selections used in the $\ell\ell\nu\nu$ analysis. 68
5.2 Summary of events pre-selections used in the $\ell\ell\nu\nu$ analysis. 69
5.3 Trigger summary in $\ell\ell\nu\nu$ analysis 70
5.4 Summary of inclusive events selections used in the $\ell\ell\nu\nu$ analysis. 72
5.5 Summary of the VBF events selections used in the $\ell\ell\nu\nu$ analysis. 75
5.6 Summary of the ggF events selections used in the $\ell\ell\nu\nu$ analysis. 76
5.7 Summary of the signal+background samples of used in the $H \to ZZ \to \ell\ell\nu\nu$ analysis. 79

6.1 Expected background contributions to the total background, their estimation and their corresponding control region. 92
6.2 Estimated $qq \to ZZ$ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and theoretical uncertainties. . . 94
6.3 Estimated $gg \to ZZ$ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and $k-$factor uncertainties. . . . 95

6.4 Selection efficiencies between Powheg and Sherpa with addition of $m_{jj} > 550$ GeV and $|\Delta\eta_{jj}| > 4.4$. 97
6.5 Estimated $qq \to ZZ$ background yields in the VBF category. The statistical uncertainty and systematic uncertainty on the yield are derived from the impact of experimental uncertainties plus the modelling systematic. 97
6.6 Relative efficiency of VBF selections with respect to the $H \to ZZ \to \ell\ell\nu\nu$ analysis selections. 99
6.7 Estimated $gg \to ZZ$ background yields in the VBF category. 99
6.8 The expected number of background and observed number of events in the 3ℓCRs with cuts. 101
6.9 Estimated WZ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and $k-$factor uncertainties. 102
6.10 Inefficiency SF of the WZ background. 102
6.11 Estimated WZ background yields in the VBF category. 103
6.12 Closure of the $e\mu$ backgrounds estimation. 106
6.13 Estimates in ggF category with binned efficiency. 106
6.14 Data-driven $e\mu$ estimate in the VBF category. 108
6.15 $e\mu$ yields in N_{jets} region and VBF region without the Z-mass requirement. . 109
6.16 Summary of the Z + jets background estimation using the ABCD method with statistical and systematic uncertainties. 111
6.17 The systematic uncertainties are due to the methodology uncertainty and the uncertainty on the subtraction of non-Z backgrounds from MC. 112
6.18 Expected yields for the Z + jets background in the VBF category, as obtained with Sherpa sample. 115
6.19 Expected yields for the Z + jets background at various selection levels with loose Z + jets quality requirements. 115
6.20 TFs obtained from region with loosened E_T^{miss}/H_T cut. Errors shown are statistical only. 116
6.21 Estimated Z + jets background yields in the VBF category. 116
6.22 Estimated Z+jets background yields with the one sideband method. 118
6.23 W + jets data-driven estimation in the electron and muon channel as obtained from the data-driven technique. 120
6.24 Triboson estimation as obtained from simulation. 120
6.25 Impact of various uncertainty sources on the physics processes involved in this analysis. 120
6.26 The summary of experimental and theory systematic uncertainties 122

6.27 The fractional experimental uncertainties on the estimation of $qqZZ$ background. 123
6.28 The fractional experimental uncertainties on the predicted yields of the heavy Higgs signal process with m_{H} =600GeV. 123
6.29 Theory uncertainties of PDF, QCD and PS variation on the acceptance of ggF signals from mass point 300 GeV to 1400 GeV in ggF category. 126
6.30 Theory uncertainties of PDF, QCD and PS variation on acceptance of VBF signals from mass point 300 GeV to 1400 GeV in VBF category. 126

7.1 The observed data and MC expected (scaled to 36.1 fb^{-1}) yields in $H \to ZZ \to ee\nu\nu$ and $H \to ZZ \to \mu\mu\nu\nu$ signal regions after full ggF event selections. 131
7.2 The observed data and MC expected (scaled to 36.1 fb^{-1}) yields in $H \to ZZ \to ee\nu\nu$ and $H \to ZZ \to \mu\mu\nu\nu$ signal regions after full VBF selections. . 131
7.3 List of nuisance parameters considered in the limit setting. 141
7.4 Signal and signal+background yields for the 300 GeV, 600 GeV and 1 TeV ggF mass points before and after the fit. 146
7.5 Signal and signal+background yields for the 300 GeV, 600 GeV and 1 TeV VBF mass points before and after the fit. 146
7.6 Summary of the event selection requirements. For the 4ℓ analysis, the two lepton pairs are denoted as m_{12} and m_{34}. 150
7.7 $\ell^+\ell^-\ell^+\ell^-$ search: Number of expected and observed events for $m_{4\ell} > 130$ GeV, together with their statistical and systematic uncertainties, for the ggF- and VBF-enriched categories. 151
7.8 Impact of the leading systematic uncertainties on the predicted signal event yield which is set to the expected upper limit, expressed as a percentage of the cross section for the ggF (left) and VBF (right) production modes at $m_{\mathrm{H}} = 300$, 600, and 1000 GeV. 156

D.1 Heavy Higgs signal in ggF process. 177
D.2 Summary of the heavy Higgs signal samples of vector boson fusion process used in this analysis. 177
D.3 Summary of the RS graviton samples used in this analysis. 178
D.4 Cross sections at NLO in perturbation theory for ZZ background 178
D.5 Cross sections at NLO in perturbation theory and filter efficiency for WZ and WW `Powheg` + `Pythia8` samples at $\sqrt{s} = 13$ TeV. 178

F.1 Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $qq \to ZZ \to \ell\ell\nu\nu$ in a m_{ZZ} window between 550 and 650 GeV. 182

F.2 Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the WZ background in a m_{ZZ} window between 550 and 650 GeV. 182
F.3 Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the Z + jets. 183
F.4 Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $gg \to H \to ZZ$ sample ($m_{\mathrm{H}} = 600$ GeV). 183
F.5 Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $qq \to qqH$ ($m_{\mathrm{H}} = 600$ GeV) sample. 184

I.1 Efficiency factor values in bin of p_{T} and η from MC. Errors contain statistical uncertainty only. 188
I.2 Efficiency factor values in bin of p_{T} and η from Data. Errors contain statistical uncertainty only. 188

M.1 The observed and expected 95% CL upper limits on the gluon fusion production cross-section of an additional narrow-width heavy Higgs boson 194
M.2 The observed and expected 95% CL upper limits on the vector boson fusion production cross-section of an additional narrow-width heavy Higgs boson . 195
M.3 The observed and expected 95% CL upper limits on the production cross-section of a RS Graviton 196

N.1 Resolution for the objects entering in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for a representative kinematic value for each object. 197

Glossary

2HDM Two Higgs Doublet Model : a Higgs sector extension model in which an extra Higgs doublet is introduced.

ATLAS A Toroidal LHC ApparatuS : a general purpose particle detector located at the LHC.

BSM Beyond the Standard Model : theories of particle physics that postulate new particles and/or interactions, which can explain some of the deficiencies of the Standard Model.

CMS Compact Muon Solenoid : a general purpose particle physics detector based at the LHC.

CSC Cathode Strip Chamber : component of the MS used in the precision measurements of muon tracks.

DxAOD Derivation of xAOD : suited for analysis, 1% of the xAOD size.

EBH Englert-Brout-Higgs: the mechanism of generating masses through spontaneous symmetry breaking.

ECAL : Electro-magnetic Calorimeter.

EM : Electro-magnetic.

EMEC Electro-magnetic Endcap part of the LAr, $1.375 < |\eta| < 3.2$.

EMTopo EM Topological cluster: three dimensional reconstruction of particle showers in the calorimeters, implemented using a topogical clustering algorithm.

EW Electroweak force : force carried out by massive mediator gauge boson particle W, Z..

FCAL Forward calorimetry part of the LAr, $3.1 < |\eta| < 4.9$.

FSR Final state radiation : Emission of radiation, in the forms of photons, from particles in the signal final state.

ggF Gluon-gluon fusion : production of Higgs boson through gluon-gluon interaction via triangle diagram of quarks.

GRL Good runs list : a configuration set for different data periods consisting of data quality flags which form a mechanism to reflect the overall data quality status of the ATLAS sub-detectors..

HEC Hadronic Endcap part of the LAr, $1.5 < |\eta| < 3.2$.

HLT High Level Trigger : the final stage of the ATLAS trigger. Events that pass HLT are recorded.

IBL Insertable B-Layer : the ATLAS pixel tracking system located closest to the beamline.

id Identification : a requirement placed on objects measurement by the ATLAS detector using discriminating variables to computationally select a specific type of particle with a high efficiency and reject backgrounds to the particle with a large rejection rate. Various levels of identification are used, commonly referred to as *loose*, *medium* and *tight* depending on the identification efficiency.

ID Inner Detector of ATLAS: consisting of the Pixel (and IBL in Run-2), SCT and TRT subsystems.

JER Jet Energy Resolution : the precision at which the energy of any given jet can be resolved.

JES Jet Energy Scale : a calibration applied to reconstructed jets in order to correct for detector effects.

JVT Jet-Vertex-Tagger : Run-II algorithm, based on multivariate methods to supress PU jets.

L1 Level 1 : the first tier of the ATLAS trigger system, utilising calorimeter and MS information.

LAr : Liquid Argon Calorimeter.

LEP : The Large Electron Positron collider.

LHC : The Large Hadron Collider.

LO Leading Order : a scattering calculation taking into account the lowest order Feynman diagram.

MC Monte Carlo : a random sampling technique used to generate events.

MDT Monitored Drift Tube : component of the MS used in the precision measurements of muon tracks.

MS Muon Spectrometer : the outer ATLAS tracking subsystem, optimised for the tracking of muons.

N3LO Next-to-Next-to-Next-to-Leading Order : a scattering calculation taking into account the lowest, next-to lowest, and next-to-next-to-lowest order Feynman diagrams.

NLO Next-to-Leading Order : a scattering calculation taking into account the lowest and next to lowest order Feynman diagrams.

NNLO Next-to-Next-to-Leading Order : a scattering calculation taking into account the lowest and next to lowest, and next-to-next-to-lowest order Feynman diagrams.

NP Nuisance parameter: Parametrisation of the systematic uncertainties in the likelihood fit.

NWA Narrow width approximation : an approximation for the Higgs boson width, $\gamma_H << m_{\mathrm{H}}$, which renders the Higgs resonance to be narrow.

OR : Overlap removal.

pdf Probability density function : specify the probability of a random variable to fall within a particular range of values.

PDF Parton Distribution Function : the momentum distribution functions of the partons inside the proton.

PFlow Particle Flow: An object reconstruction technique, where the tracks in the ID are matched to calorimeter deposits.

POI Parameter of interest: In particle physics analyses, it is an important model independent parameter to determine, as it quantifies the signal observed over the expectation from simulation.

PS Parton showering : simulation of the hadronisation stage of an event reaction.

PU Pile-Up : the average number of proton-proton interactions per bunch crossing.

QCD Quantum Chromodynamics : quantum field theory describing the strong interaction.

QED Quantum Electrodynamics : quantum field theory describing the electro-magnetic interaction.

ROC Receiver operating characteristic : predicits the ability of variable to accept or reject a certain effect as different discrimination thresholds are varied.

RoI Region of Interest : an area of the ATLAS detector containing L1 reconstructed objects that pass the L1 trigger.

RPC Resistive-Plate Chamber : component of the MS used in the muon trigger.

RS : Randall-Sundrum.

SCT Semi-Conductor Tracker : one of the ATLAS tracking subsystems, consisting of silicon strip modules.

SF : Scale factor.

SM Standard Model : a quantum field theory describing the known matter particles and their interactions through the electroweak and strong forces.

SUSY Supersymmetry : a candidate extension of the SM in which a new symmetry between bosons and fermions is introduced.

TF : Transfer factor.

TGC Thin-Gap Chamber : component of the MS used in the muon trigger.

TRT Transition Radiation Tracker : one of the ATLAS tracking subsystems, consisting of inert gas filled straw tubes.

TST Track Soft Term : a term in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation that consists of the sum of the transverse momenta of ID tracks not associated to any reconstructed objects..

VBF Vector boson fusion : production of Higgs boson through quark-quark interaction via di-boson diagram.

vev Vacuum expectation value : The average expectation value of a quantum field in vacuum.

xAOD Analysis Object Data : physics output such as electrons, muons, tracks etc.

Chapter 1

Introduction

The Higgs particle, held as the holy grail of elementary high energy particle physics, was independently observed by two large scale scientific collaborations at the Large Hadron Collider (LHC): ATLAS and CMS [8, 9]. The discovered Higgs boson is direct evidence of the existence of the Higgs field, the quantum field which is responsible for mass generation of gauge bosons and fermions, the particles which make up the Standard Model (SM) of particle physics. The Higgs boson mass (m_h) was measured to be $125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{syst})$ GeV [10]. ATLAS and CMS have confirmed that the particle properties, spin, parity and couplings are, within the experimental uncertainties [11–13], consistent with the prediction of the SM Higgs particle. A recent report based on combination of ATLAS and CMS measurements of the Higgs coupling to other SM particles can be found here [14].

The observation of the Higgs boson has confirmed a key SM prediction and completed the final missing ingredient of the theory. However, several crucial questions are still unresolved [15]. One example is the so-called "hierarchy problem" which can be examined within the context of the SM Higgs boson [16]. More specifically, a quadratic dependence on the energy scale is found when considering the virtual quantum corrections to m_h. If the scale extends to the Planck energy scale $\mathcal{O}(10^{18}$ GeV) then the quantum corrections are supposed to be much larger than the mass of the Higgs. In fact, *any* new physics which couples to the Higgs must produce a quadratic scale dependence. Within the SM the only way to mitigate the dependence is accomplished by adjusting the bare parameter values ("fine-tuning") to the new physics scale. This fine-tuning of parameters is unnatural as no known physical processes can produce it other that just pure accident.

One possible solution of the hierarchy problem is through the existence of extra particles which share similar properties as their SM equivalents. With this in mind, a natural cancellation of the dependence can occur. The theory of super-symmetry (SUSY) provides such a framework i.e. one where each SM particle has a super-partner equivalent. Therefore SUSY

theories are able to form such a natural cancellation mechanism [17]. In addition, SUSY shows that the coupling of three of the forces of nature (Strong, Weak and Electro-Magnetic) to unify at a single energy scale.

However, a single Higgs field by itself cannot generate the mass of all SUSY particles. The presence of at least one additional Higgs field is required. The simplest extension of the Higgs sector is through the Two Higgs Doublet Models (2HDMs) [18], which introduce a second complex scalar field. The 2HDM predicts in total the existence of five Higgs particles. Two of the five bosons are spin-0 neutral scalars, one is a pseudo-scalar and the remaining two are charged scalars. The observed SM Higgs can be identified with one of the neutral scalar Higgs particles from the 2HDM.

Alternatively, the large quantum correction to the Higgs boson mass, when considering the Planck scale, can be resolved by the existance of extra dimensions. More specifically, a compactified extra dimension can result in an apparent separation between the TeV and Planck scale while in reality the actual scale separation is small, which results in a tiny m_h correction [19, 20].

The aim of this thesis report is to summarize the searches for an additional heavy Higgs boson, using the $H \to ZZ \to \ell\ell\nu\nu$ decay mode, where one Z boson decays on-shell to a pair of leptons ($Z \to \ell\ell$), while the other decays to a neutrino pair ($Z \to \nu\bar{\nu}$). As neutrinos are weakly interacting, they leave no detectable signature in the ATLAS detector, and the only way to measure their presence is through the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$. This search is focused in the high mass range between 300 GeV and 1.5 TeV, with 36.1 fb^{-1} of integrated luminosity collected by the ATLAS experiment during the 2015 and 2016 data period runs at the LHC.

A brief overview of the theoretical building blocks for heavy Higgs boson searches is presented in chapter 2. After this, chapter 3, describes the LHC experiment and the ATLAS detector. The reconstruction of physics objects (such as electrons, muons etc.), by the ATLAS experiment, are described in chapter 4 with a focus on $E_{\mathrm{T}}^{\mathrm{miss}}$, due to its importance in the $H \to ZZ \to \ell\ell\nu\nu$ analysis and contribution of the author to it. The description of the analysis is performed in chapter 5, with details on the search selections, cut optimization schemes, background estimations and systematic uncertainties. The results in the signal regions, their upper limis, nuisance parameters, combination with $H \to ZZ \to 4\ell$ analysis and spin-2 interpretations are summarized in chapter 7. Future developments, which may offer improvement in the analysis search, are briefly discussed in chapter 8.

Chapter 2

The Standard Model and its extensions

The Standard Model SM of elementary particle physics was built from years of theoretical and experimental research [21]. It describes the set of particles and their interactions, which constitute the matter content of the observable universe. The SM provides the most accurate explanation of particle interactions at the currently highest experimentally achievable energy scales.

Despite being enormously successful in explaining a plethora of phenomena, a set of physics processes remain unexplained. A class of models known as Beyond the SM (BSM) aim to provide a description of these phenomena [22].

A brief overview of the SM is given in section 2.1. The generation of mass within the SM by the Higgs field is introduced in section 2.2. The properties of the quanta of the Higgs field in the SM are described in section 2.3. The problems of the SM and its extensions are outlined in section 2.4. One of the most popular models for BSM physics in the Higgs sector, 2HDM, is overviewed in section 2.5. An alternative BSM model, the Randall-Sundrum model, which offers a description of gravity at small scales, is explored in section 2.6. Finally the results from Run I of the LHC, with 20.3 fb^{-1} of data collected by ATLAS at $\sqrt{s} = 8$ TeV, are summarized in section 2.7.

2.1 The Standard Model of Particle Physics

The matter and forces within the SM are viewed as a physical manifestation of quantum fields [23, 24]. The quanta of these fields are point-like particles which have a range of properties such as mass, charge, spin etc. Matter is made from spin-1/2 particles called fermions while forces are transmitted due to spin-1 particles called vector bosons. The

fermion content consists of quarks (q) and leptons (ℓ) arranged in three generations of charge, flavour and mass. Quarks have a fractional electric charge and an additional charge called colour. The strong force is mediated between quarks via massless vector bosons called gluons, g, which interact with the quarks and between themselves via their colour charge. The quarks and gluons are confined in hadrons such as the proton because of asymptotic freedom [25]. The weak force is mediated by three massive bosons: two electrically charged $W^{\pm}$ bosons and a neutral Z^0 boson. The mediator of the EM force is the electrically neutral massless photon, γ. The contents of the SM theory are summarised in Figure 2.1.

The interactions in the SM arise from the invariance of its Lagrangian density[1], $\mathcal{L}$, as a result of a local gauge transformation [26]. The integral over the space-time coordinates ($x^{\mu} = (x, y, z, t)$) of the Lagrangian function forms the action $\mathcal{S} = \int \mathcal{L} d^4x$ of a system. From the principle of least action, i.e. $\delta\mathcal{S} = 0$, follow the equations of motion, which describe the physical consequences of the fields contained in the $\mathcal{L}$ [27]. A local gauge transformation implies that the transformation acts differently on different space-time points, in contrast to a global gauge transformation which acts the same for all points. The requirement for a Lagrangian to be invariant under a local gauge operation demands the addition of extra terms, which correspond to the interactions between matter particles mediated by quanta of the gauged fields.

The three forces in the SM: strong, weak and EM, are generated by $SU(3)_C \times SU(2)_L \times U(1)_Y$ local gauge symmetries, where $U(N)$ refers to a group of $N \times N$ unitary matrices with $SU(N)$ being a subset with determinant equal to unity. The strong force is generated by a $SU(3)_C$ symmetry where C refers to the colour charge. The electro-weak (EW) force, i.e. the unification between weak and EM forces, is generated by the $SU(2)_L \times U(1)_Y$ gauge symmetry. L refers to weak isospin coupling to left-handed particles only, where the left-handed particle chirality state refers to the particle spin being opposite to the particle's momentum. Y refers to the weak hypercharge, which is the sum of the electric charge operator and the third component of the weak isospin.

The SM is also invariant under CPT symmetry where C stands for charge, P for parity and T for time reversal. More importantly, the CP symmetry is preserved in strong and EM interactions but violated in weak interactions. The violation of CP symmetry is one of the required Sakharov conditions for the observed matter-antimatter imbalance [28–30].

An important problem in the formulation of the SM was due to the masses of the W and Z gauge bosons. The breakthrough was provided by spontaneous symmetry breaking of the $SU(2) \times U(1)$ gauge group [31–33].

[1] A Lagrangian density, simply referred to as Lagrangian, represents a function of the fields and their derivatives.

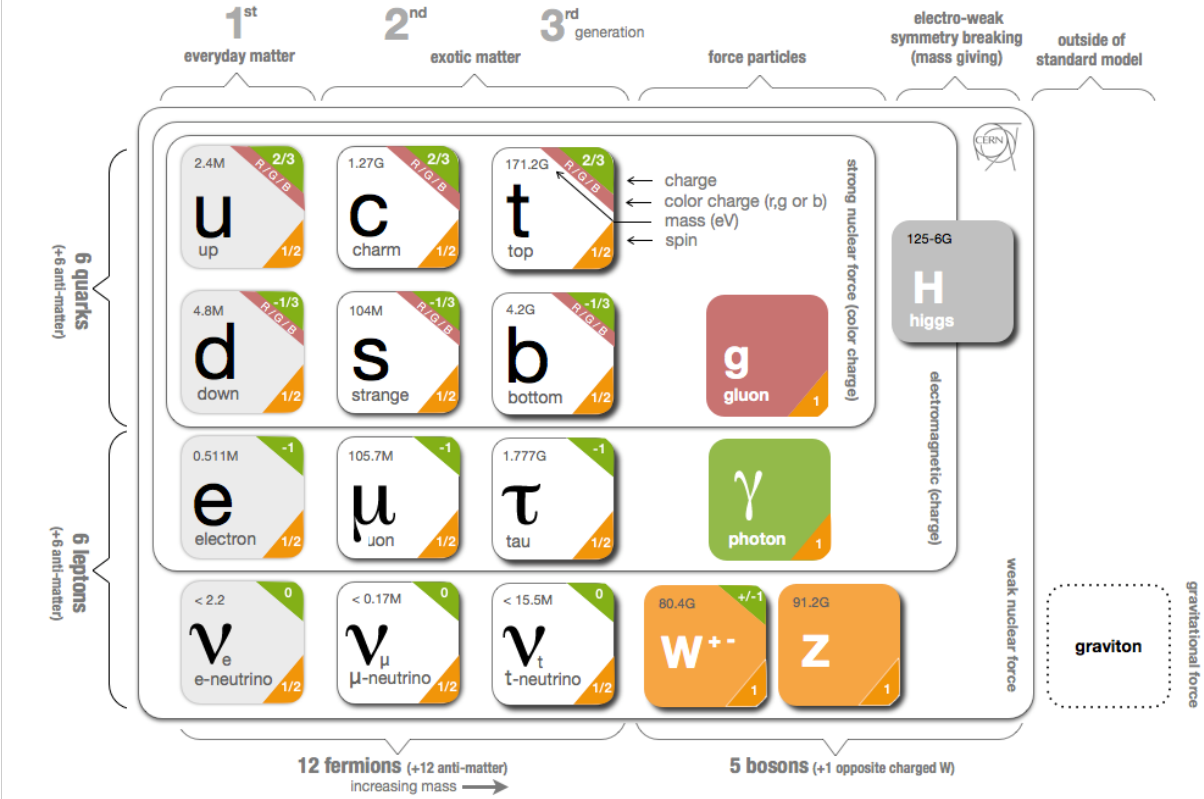


Figure 2.1: This figure [34] shows the elementary particles contained in the Standard Model, together with a summary of their basic properties.

2.2 Electro-weak symmetry breaking

The $SU(2) \times U(1)$ gauge symmetry faced a major obstacle in the SM formulation. The W and Z bosons have non-zero masses, which would violate the gauge invariance from which the interactions between particles are built. By comparison the most spectacular success of gauge theories [35], quantum electrodynamics (QED), is mediated by a massless photon where the Lagrangian naturally generates the photon interaction term. The incorporation of massive bosons was later realised to be possible via spontaneous electro-weak symmetry breaking.

The solution of the mass term problem originated from considerations of the energy of the lowest vacuum state. If the vacuum state had instead a non-zero vacuum expectation value (vev), then the symmetry of theory can be *spontaneously* broken rendering the theory with the needed mass terms without violating the local gauge principle. In the SM framework, the spontaneous symmetry breaking of the electro-weak $SU(2) \times U(1)$ gauge symmetry is

able to successfully generate mass terms for the gauge bosons via the Englert-Brout-Higgs (EBH) mechanism [31–33]. The EBH mechanism introduces a new quantum field, called the Higgs field, where SM particles acquire mass if they couple to it. The quanta of the Higgs field is the Higgs boson, further discussed in section 2.3.

A consequence of electro-weak symmetry breaking is the presence of massless bosons, referred to as Goldstone bosons [36]. The Goldstone bosons are unphysical states, as no known observation has found their existence. The final piece of the EBH mechanism was the proof that the Goldstone bosons can be completely removed from the theory with an appropriate choice of gauge. As the gauge is irrelevant to the physical consequences of a theory, the Goldstone bosons were no longer a part of the theoretical formulation, leaving the single Higgs field and its couplings to the massive gauge bosons.

A simple example of the EBH mechanism is the spontaneous symmetry breaking of the rotational $U(1)$ symmetry. The scalar part of the Lagrangian is given by:

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \tag{2.1}$$

where ϕ is a complex scalar doublet, $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ made from two real fields, ϕ_1 and ϕ_2. The simplest potential field $V(\phi)$ which gives rise to mass terms follows:

$$V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4, \mu^2 < 0, \lambda > 0 \tag{2.2}$$

where the parameter μ is a field mass term and λ is self-coupling term. For the mechanism to generate a real mass, μ^2 must be negative and λ positive. The shape of the potential term is presented in Figure 2.2. The potential has an unstable state at $V(\phi) = 0$, while the stable ground state of minimum potential is located where the value of ϕ is such that $\partial V/\partial\phi$ is zero. The value of ϕ at the ground state is equal to:

$$|\phi|^2 = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \tag{2.3}$$

where v is the non-zero vev of the field ϕ. The transition to the stable ground state generates the Higgs field term in the Lagrangian. However, the transition can occur at any point in the ground state. Therefore the ground state is degenerate and this choice about the ground state gives the Goldstone boson. The scalar field ϕ can then be written as:

$$\phi = \frac{1}{\sqrt{2}}(v + \eta + i\xi) \tag{2.4}$$

where the two fields η and ξ are seen as the excitations around the vacuum state. The field η corresponds to the Higgs field, with a mass term $m_\eta = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}$ while the ξ field gives rise to a massless Golstone spin-0 particle. The Lagrangian is then written as:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) + \frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) - V(\eta, \xi). \tag{2.5}$$

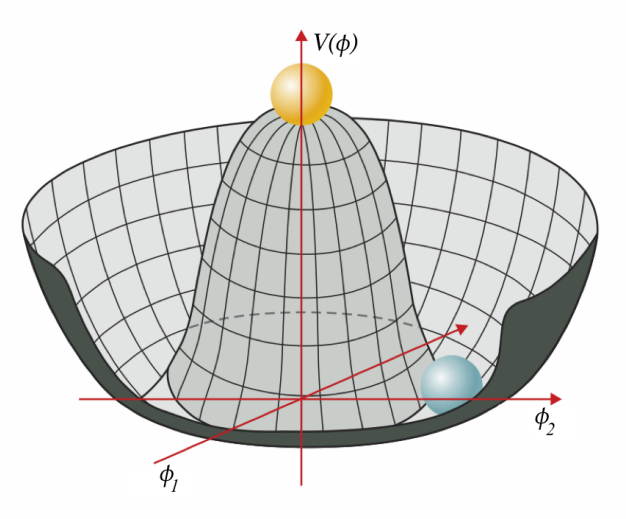


Figure 2.2: The shape of potential for a complex doublet field, ϕ, as a function of the real component fields ϕ_1 and ϕ_2.

In the EBH mechanism, the $U(1)$ local gauge symmetry is spontaneously broken, which then generates a mass term for a gauge field B. The Lagrangian reads:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \tag{2.6}$$

where $F^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu$ is the kinetic term of the field B and $D_\mu\phi$ is the covariant derivative of the field ϕ. A covariant derivative leaves the Lagrangian gauge invariant, where $D_\mu = \partial_\mu + igB_\mu$, where g is a coupling constant. Substituting the field ϕ from equation 2.4 transforms the Lagrangian into:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2 + \frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}g^2v^2B_\mu B^\mu + gvB_\mu(\partial^\mu\epsilon) - V \tag{2.7}$$

where V is the potential term containing the interaction terms between the Higgs field and the gauge field B. A simple substitution, $B'_\mu(x) = B_\mu(x) + \frac{1}{gv}\partial_\mu\xi(x)$ leaves equation 2.7 gauge invariant. This translates into a gauge choice of $\phi'(x) = e^{-i\xi(x)/v}\phi(x)$, which is called the unitary gauge. This gauge choice shows how to eliminate the Goldstone bosons from the Lagrangian.

In the context of the SM, the EBH mechanism is applied to the $SU(2)\times U(1)$ gauge symmetry. Spontaneous symmetry breaking leaves three Goldstone bosons, which are identified as the longitudinal polarisation degrees of the $W^\pm$ and Z^0 bosons, which acquire a mass term. The mass terms for each boson, is related to the vev of the Higgs field by:

$$m_W = \frac{1}{2}g_W v, m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} = \frac{m_W}{\cos(\theta_W)}, m_A = 0, m_h = v\sqrt{2\lambda} \tag{2.8}$$

where g_W and g' are gauge couplings of the weak fields, and m_W, m_Z, m_A, m_h are the masses of the W, Z, photon and Higgs particles, respectively. The mass terms for the fermions are

generated similarly by electroweak symmetry breaking through the Yukawa Lagrangian. The mass terms for leptons are related by $m_\ell = g_\ell v/\sqrt{2}$ and quarks $m_q = g_q v/\sqrt{2}$ (see [37]).

2.3 Properties of the Higgs boson

The EBH mechanism requires the presence of an additional quantum field, the Higgs field η, which other particles interact with to gain mass. The quantisation of η results in the production of the quanta of the Higgs field. Therefore, the Higgs particle provides direct evidence of the existence of the Higgs field and therefore the EBH mechanism. The couplings to other fields are predicted by the SM but the energy scale is not, leaving the mass parameter of the Higgs quanta to be determined by experimental measurement. The SM Higgs boson is a spin-0 electrically neutral scalar boson, while the photon, $W^\pm$ and Z are spin 1 particles [37]. It is CP-even, which renders it invariant under charge and parity transformation. The vev corresponding to the SM Higgs particle is equal to 246 GeV.

The search for the Higgs boson started over 40 years ago when the SM was developed. Before the observation in 2012, a number of constrains from theoretical and experimental measurements were placed on the Higgs boson mass, m_h. Theoretical constraints from electroweak fits constrained m_h to be below 200 GeV [38]. Direct constraints from LEP showed a lower bound of $m_h > 114.4\,\text{GeV}$ at 95% CL [39].

The Higgs boson was observed in 2012 independently by the ATLAS and CMS collaborations. The observed mass was $m_h = 125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{sys})\,\text{GeV}$ [40] and its spin was found to be consistent with the prediction of the SM [41].

2.3.1 SM Higgs boson production

At the Large Hadron Collider, the Higgs boson is produced via four physical mechanisms: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a gauge boson (VH), and associated production with a top-quark pair ($t\bar{t}H$). Figure 2.3 shows the evolution of the SM Higgs particle production cross section for the various production modes as function of m_h. At the LHC, the two largest production modes are the ggF and VBF, shown in Figure 2.4, which are also the main focus of this thesis. The remaining two production modes are discussed in [37] and evidence in Run II of the LHC is given in [42,43].

The ggF process proceeds via a triangular diagram of virtual quarks, as the Higgs boson does not couple directly to gluons. The quark triangle is dominated by top quarks, and to a smaller extent the bottom quark. The top quark mass in the triangle is approximated to be very large, which allows the computation of the cross section to next-to-next-leading

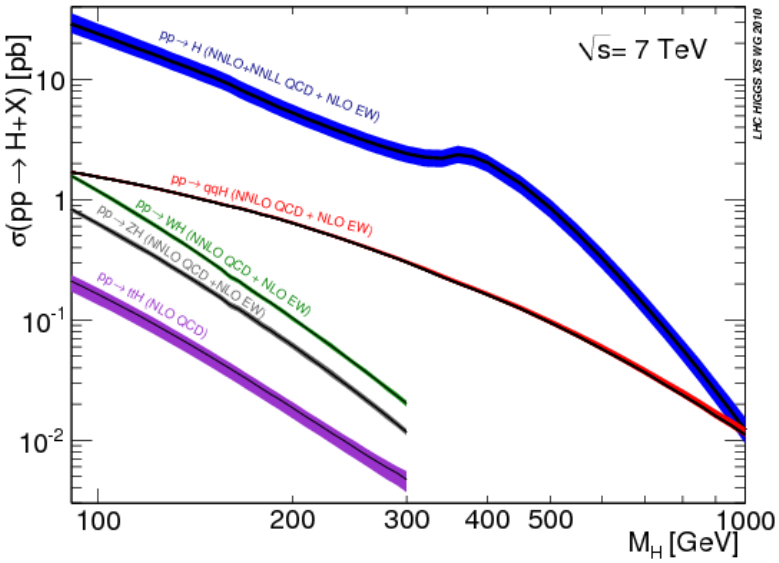


Figure 2.3: The evolution of the production cross section of the Higgs boson as a function of m_h for the $pp \to H$ (ggF) and $pp \to qqH$ (VBF) production modes at the LHC. Figure taken from [44].

order (NNLO) and next-to-next-to-next-leading order (N3LO) in α_s at $m_h = 125$ GeV. The next-to-leading order (NLO) corrections increase the leading order (LO) cross section prediction by 80%, while the NNLO increases the cross section by a further 30%. NLO EW correction increase the cross section by 5%. The recently added N3LO corrections only enhance the production cross section by 3% [45]. At $\sqrt{s} = 13$ TeV, the ggF production cross section makes up 87% of the total Higgs boson production cross section. The total ggF production cross section at 13 TeV for a 125 GeV Higgs particle is:

$$\sigma_{ggF}^{\text{total}} = 48.6^{+2.2}_{-3.3}(\text{theory}) \pm 1.6\text{pb}(\text{PDF} + \alpha_s). \tag{2.9}$$

The VBF process proceeds by scattering of two initial state quarks mediated by the exchange of W or Z bosons and the Higgs particle is produced by the interaction between the weak bosons. The two initial state quarks, are scattered into the forward and backward detector regions, where they give rise to two highly energetic jets. These unique topological characteristics can be used to select VBF events. Because the emission of the W boson is governed by the weak interaction, the VBF production cross section is much smaller than ggF and it accounts for only 7 % of the total cross section. The total VBF production cross section is 3.748 pb^{-1} for $m_h = 125\,\text{GeV}$ at $\sqrt{s} = 13\,\text{TeV}$.

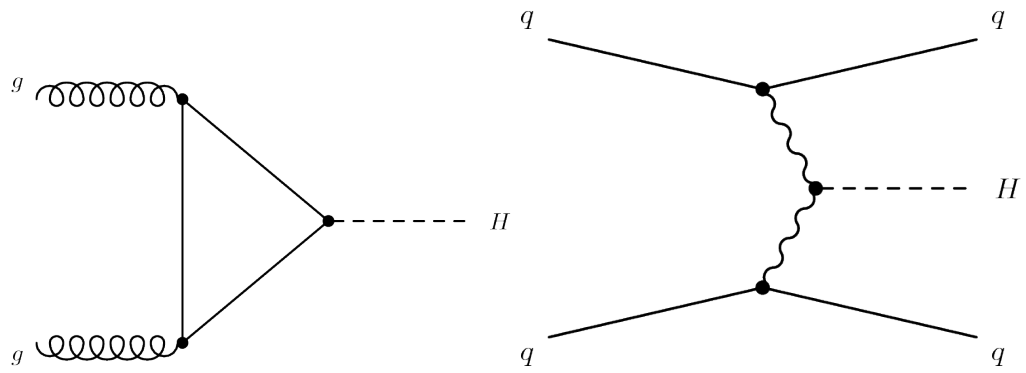


Figure 2.4: The two largest Higgs boson production modes at the LHC where ggF contributes 87% while VBF contributes 7% to the total Higgs particle production cross section.

2.3.2 SM Higgs boson decay

The Higgs particle is too heavy to be stable for a long time, and quickly decays to particles with smaller mass which convert the initial rest mass energy into momentum. Since the Higgs boson couplings to SM is proportional to their respective particle masses, the Higgs boson prefers to decay into the heaviest pair which are allowed by the available phase space. In experimental particle physics, the probability of a particle to decay to a different set of particle is quantified by the branching fraction. The branching fraction for each of the Higgs boson decay modes is predicted by SM and it varies with m_h since the couplings are dependent on it. Figure 2.5 shows the variation of branching fraction as a function of m_h.

At $m_h = 125\,\text{GeV}$, a wide range of decay modes are available to the Higgs quanta. The branching fraction is dominated by $H \to b\bar{b}$ decays, with about 58% where the second largest decay mode $H \to WW$ makes up 21% of the branching fraction. The rest are shared between $H \to \tau^+\tau^-$ (6%), $H \to ZZ$ (3%) among others. The $H \to \gamma\gamma$ process, has a branching ratio of 0.23%. In 2012, the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay modes were used to discover the Higgs boson as they provide the cleanest signal measurement against the large QCD background. The $H \to b\bar{b}$ search, with data from the 2016 runs showed an excess of signal events of 3.5 σ [42].

The total decay width of the Higgs particle, Γ_h, varies as a function of m_h. Γ_h is a sum of the partial widths, Γ_i, of the different decay modes. The branching fraction of each decay mode is defined as Γ_i/Γ_h. In the SM, the Higgs boson at 125 GeV, has a small total decay width of $\Gamma_h = 4.07$ MeV relative to the LHC energy resolution.

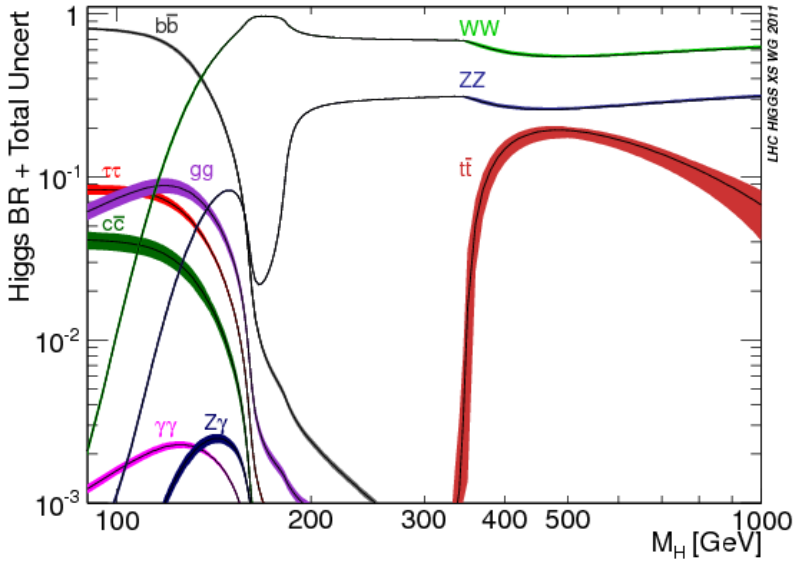


Figure 2.5: The variation of the Higgs boson branching fractions as a function of m_h at the LHC. Figure taken from [44].

2.3.3 Higgs particle for large values of the Higgs boson mass

The aim of the $H \to ZZ \to \ell\ell\nu\nu$ analysis is to look for a Higgs resonance at large values of the Higgs boson mass, m_{H}. The heavy Higgs quanta can be expected to share similar properties with the SM Higgs boson i.e. spin-0 electrically neutral CP-even scalar and to be produced via the ggF and VBF mechanisms. Justification for this assumption in theoretical models is provided by 2HDMs in section 2.5.

An important distinction between the heavy Higgs resonances and SM case, for large masses, occurs with the total decay width Γ_H. At large m_{H}, the decay width of the SM Higgs boson is proportional to m_{H} as:

$$\Gamma(H \to WW + ZZ) \sim 0.5 \text{ TeV}[m_{\mathrm{H}}/\text{TeV}]^3. \tag{2.10}$$

The decay width is thus very large and subsequently the SM Higgs particle is no longer a narrow peak but rather a very broad distribution. Interestingly the SM Higgs boson also has an off-shell signal contribution [46].

In contrast, the width of the heavy Higgs resonance is assumed to be negligible, which follows from the narrow-width-approximation (NWA). For on-shell mass states the accuracy of the NWA is thoroughly discussed in [47]. The NWA renders the heavy Higgs particle to behave much in the same way as the SM Higgs boson at the resonance mass. However, this approximation can breakdown if quantum mechanical interference effects between the heavy Higgs particle and the non-resonant SM diboson background are present, which would then

modify the background line-shape, the kinematic distributions and production cross sections. Such fears can be cleared away with the application of the $H \to ZZ \to \ell\ell\nu\nu$ analysis selections, which have been demonstrated to suppress all interference contributions [48].

2.4 Extensions of the Standard Model

The SM is one of the most successful theories of particle physics, with ability to predict various phenomena across a range of energies, but it does not provide a clear treatment to several important problems and therefore is far away from being considered a theory of *everything*.

2.4.1 Shortcomings of the SM

One of the most glaring points which renders the SM only an approximate theory is the amount of free parameters present in it. In total 19 parameters are determined by experiment, which includes the mass parameters of all fermions and Higgs boson, the CKM mixing angles, CP violation phase, the gauge couplings and Higgs vev [49].

As mentioned in the introduction, the hierarchy problem states that the Higgs boson mass has a quadratic scale divergence through heavy fermion loops, which, in the SM context, can only be explained by a fine-tuning of the tree parameters with a precision of 10^{-15} [50]. It seems very unlikely that such a process can happen purely by accident.

The hierarchy problem alludes to another outstanding problem, the nature of gravity. The SM can be seen as an effective field theory, valid in the electroweak energy scale, but breaking down near Planck scales where the effects of gravity become important. No unified quantum mechanical description of gravity within the SM exists [51].

Observations from astronomy has indicated that the visible matter constitutes only $\sim 5\%$ of the total content of the universe. About 27 % of the content is made from non-luminous dark matter which has so far been only observed through the cosmic microwave background and velocity distributions of galaxies. The rest, $\sim 68\%$, is made from dark energy, a mysterious form of energy thought to cause the accelerated expansion of the universe [52]. None of these effects are predicted or explained by the SM.

2.4.2 Beyond the Standard model

Many models have been developed over time to solve some of the problems in the SM. Proposed solutions of the hierarchy problem invoke supersymmetric theories which provide

a natural cancellation mechanism to the quadratic scale dependence in the Higgs boson mass. SUSY theories require the existence of additional particles, called "superpartners" of their SM equivalent [53]. Additional Higgs fields are also required to provide mass terms to the full spectrum of particles predicted by SUSY. The most simple mechanism of extra Higgs fields is given by 2HDMs.

Other models, such as the Randall-Sundrum model [19] have offered solution to the weakness of gravity by proposing compactification of extra dimensions which bridge the gap between the Planck and electro-weak scale.

2.5 Two Higgs Doublet Models

The 2HDM is an extension of the SM, where a second complex scalar field is added to the SM Lagrangian [54]. The two complex Higgs doublets, denoted by ϕ_1 and ϕ_2, modify the potential, V, which (assuming CP conservation) is given by:

$$V = m_{11}^2\phi_1^\dagger\phi_1 + m_{22}^2\phi_2^\dagger\phi_2 - m_{12}(\phi_1^\dagger\phi_2 + \phi_2^\dagger\phi_1) + \frac{\lambda_1}{2}(\phi_1^\dagger\phi_1)^2 + \frac{\lambda_2}{2}(\phi_2^\dagger\phi_1)^2 + \lambda_3\phi_1^\dagger\phi_1\phi_2^\dagger\phi_2 + \lambda_4\phi_1^\dagger\phi_2\phi_2^\dagger\phi_1 + \frac{\lambda_5}{2}((\phi_1^\dagger\phi_2)^2 + (\phi_2^\dagger\phi_1)^2) \tag{2.11}$$

where the m_{ab}, λ_c parameters are real. After spontaneous symmetry breaking, five Higgs bosons appear: CP-even neutral scalar Higgs particles, h and H, a CP-odd [2] neutral pseudoscalar A and two charged scalar partners $H^\pm$ [55]. The mass terms in the lagrangians for each of the five Higgs boson is given by a mass matrix. A suitable basis is chosen where those mass matrices can be diagonalised. It is found that one angle, α, is needed to diagonalise the two charge neutral scalar terms, and another angle β to diagonalise the two charged scalars and the pseudoscalar Higgs particles. The angle α is known as the mixing angle between the two CP-even neutral Higgses, while the angle β is the mixing between the charged scalars and the pseudoscalar. An important definition in 2HDM is [56, 57]:

$$\tan(\beta) = v_2/v_1, \tag{2.12}$$

where v_1, v_2 are the vevs for the two doublet fields, ϕ_1, ϕ_2 and the total vev is $v = \sqrt{v_1^2 + v_2^2} = 246\,\text{GeV}$. In addition, the angle $(\beta - \alpha)$ is directly related to the CP-even neutral Higgs boson coupling strengh to the gauge bosons. They are given by:

$$g_{\text{hVV}} = g_{\text{h}_{\text{SM}}\text{VV}}\sin(\beta-\alpha), g_{\text{HVV}} = g_{\text{h}_{\text{SM}}\text{VV}}\cos(\beta-\alpha). \tag{2.13}$$

The other free parameters of the theory are Higgs masses m_h, m_H, m_A, $m_{H^\pm}$ and mass term m_{12}^2 responsible for mixing of the two Higgs doublet states.

[2] CP-odd pseudoscalars have a -1 eigenvalue after CP transformation as oppose to CP-even which have $+1$.

The "alignment limit" i.e. $\cos(\beta - \alpha) = 0$ is where the H decouples from the vector gauge bosons VV ($V = W, Z$) and h is indistinguishable from the SM Higgs. The 2HDM are experimentally studied near the alignment limit where $\cos(\beta - \alpha)$ has a small non-zero value and h behaves very close to the SM prediction. Additionally the 2HDMs are studied for masses where the Higgs decay widths are small which renders the interference effects with the non-resonanant ZZ backgrounds to be negligible.

A problematic issue in 2HDMs is the presence of flavour changing neutral currents (FCNC) at tree level. FCNC occur because the two Yukawa matrices, i.e. the matrices which contain the couplings of the two Higgs fields to the fermion generations, do not have a general basis where they are both diagonal. The off-diagonal terms introduce FCNC in the fermion generations which are not experimentally observed. To ensure the absence of FCNC certain discrete CP symmetries are required to be preserved. This results in two general types of 2HDM: type I where all right handed quarks couple to the same Higgs doublet and type II where all right-handed up-type (u, c, t) couple to one of the Higgs doublet, and the right-handed down-type (d, s, b) quarks couple to the other. SUSY models require type II 2HDMs, but type I is also studied on its own.

2.6 Gravitons

The Randall-Sundrum (RS) model was developed in order to tackle the hierarchy problem [19, 20]. The hierarchy problem concerns the large gap between the masses of the SM and the Planck mass, M_{pl}. The RS model proposes that the universe is a 5D anti-de Sitter (AdS_5) space i.e. a manifold with negative curvature which has two branes (boundaries where certain particles and interactions can exist). The 4D "Tevbrane" is where the SM scale is located while the "Planckbrane" is located at the Planck scale. The large hierarchy difference of scales is a result of a warped extra dimension, where the scale is given by $\text{TeV}/M_{\text{pl}} \sim e^{-k\pi R}$ where k is the curvature and R is the proper size of the warped extra dimension. According to the RS model the weakness of gravity at the TeV scale is a result of the small overlap of the graviton wave function from the warped extra dimension near the Tevbrane. Gravity is mediated through Kaluza-Klein (KK) gravitons [58] which have a spin-2 mode with masses and couplings at the TeV scale.

The RS model coupled with KK gravitons makes specific predictions which are studied at LHC energies [59, 60]. In particular, the decay width of the lightest KK resonance to a pair of Z bosons is $\Gamma(G \to ZZ) \sim 0.06 \times (k/\bar{M}_{\text{pl}})^2 \times m_{\text{G}}$ [61] where m_{G} is the mass of the KK graviton and $\bar{M}_{\text{pl}}$ is the reduced Planck mass ($\bar{M}_{\text{pl}} = M_{\text{pl}}/\sqrt{8\pi}$). Values of $k/\bar{M}_{\text{pl}}$ range between 0.5 and 2, with $k/\bar{M}_{\text{pl}}$ commonly set to unity. This is valid in the model since

higher curvature terms can be suppressed, provided $k/\bar{M}_{\mathrm{pl}} < 2.9$ [61].

The graviton can be created via ggF and VBF production in a similar manner to the Higgs. In the $H \to ZZ \to \ell\ell\nu\nu$ analysis the graviton is only studied in the ggF category.

2.7 Run I results

2.7.1 $H \to ZZ \to \ell\ell\nu\nu$ analysis

In run I, the $H \to ZZ \to \ell\ell\nu\nu$ analysis was performed for the ggF and VBF production modes in the $ZZ \to ee\nu\nu$ and $ZZ \to \mu\mu\nu\nu$ final states in the m_{H} range between 240 GeV and 1 TeV. The analysis used a total integrated luminosity of 20.3 fb^{-1} of 8 TeV data, collected by ATLAS in 2012 [62]. The ggF category adopted the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ as a discriminant between signal and background, while the VBF category used a single bin count due to lack of statistics for proper shape modelling.

The definition of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ analysis observable is given in equation 5.1. The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution in the ggF category is shown in Figure 2.6. The total number of observed events in data in the ggF (VBF) region is 309 (4) with an expected number of events of $301.3 \pm 8.4 \pm 61.7$ ($1.37 \pm 0.21 \pm 0.78$). No significant excess was found and consequently upper limits were placed on each production mode. The combined ($ee\nu\nu + \mu\mu\nu\nu$) upper limits for ggF and VBF production are shown in Figure 2.7.

2.7.2 Combination with other $H \to ZZ$ decay modes

A combination was performed between the various final states of the $H \to ZZ$ decay mode [63]. The specific decay chains of the $H \to ZZ$ mode in question are: $ZZ \to 4\ell, ZZ \to \ell\ell\nu\nu, ZZ \to \ell\ell qq, ZZ \to qq\nu\nu$. Most of these final states (with the exception of the 4ℓ) are mainly used in the regions with large m_{H}, where they are sensitive to heavy Higgs boson production.

The $ZZ \to 4\ell$ final state uses the invariant mass of the di-lepton pair, $m_{4\ell}$, as the discriminant variable between signal and background. The search for heavy Higgs boson production has a mass range between 200 GeV and 1 TeV. The 4ℓ analysis consists of two signal regions, ggF and VBF, where ggF is split into sub-categories depending on the lepton flavour.

The $ZZ \to \ell\ell qq$ decay mode consists of a di-lepton pair from an on-shell Z boson and two jets from the other Z boson decay. The discriminating variable is the invariant mass, $m_{\ell\ell jj}$, formed from the lepton and jets pairs. The signal regions are divided into resolved-jet

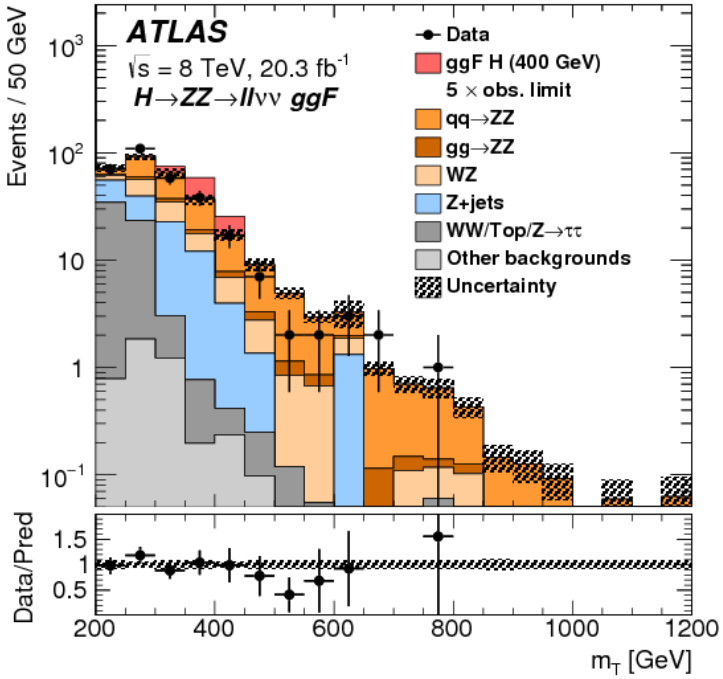


Figure 2.6: The m_{T}^{ZZ} distributions in the final signal region in the ggF category. The data-driven estimations are used where applicable. For comparison, the ggF signal is also shown with mass of 400 GeV. Figure is taken from [62].

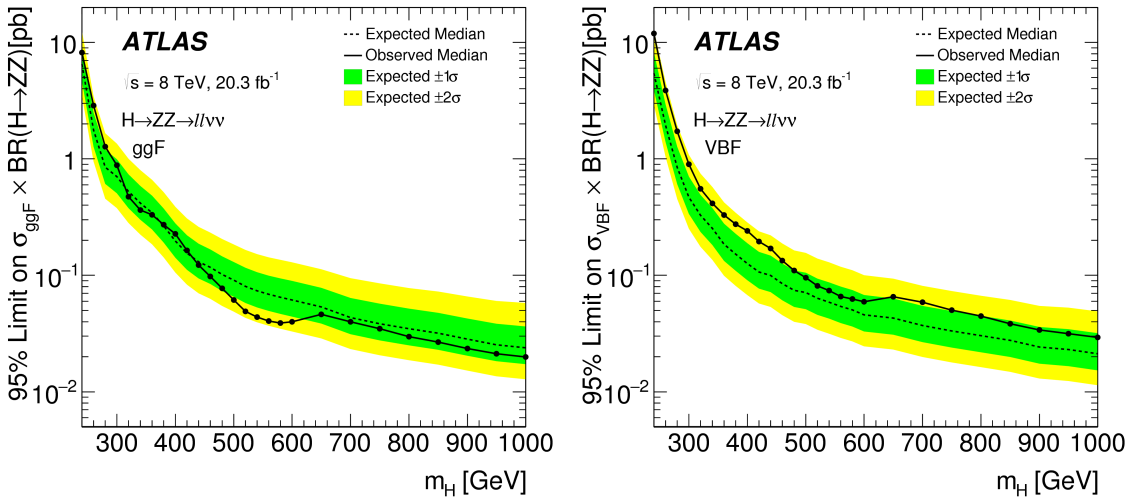


Figure 2.7: 95% CL upper limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ for a heavy Higgs decay to $\ell\ell\nu\nu$. The left plot is set for the ggF signal while the right plot is set for the VBF signal. Figures taken from [62].

and merged-jet categories, making up the ggF channel, and a single category for the VBF channel. The range of the $m_{\ell\ell jj}$ distribution is between 200 GeV and 1.5 TeV.

The $ZZ \to qq\nu\nu$ channel uses a transverse mass variable, m_{T}, defined from the $E_{\mathrm{T}}^{\mathrm{miss}}$ and

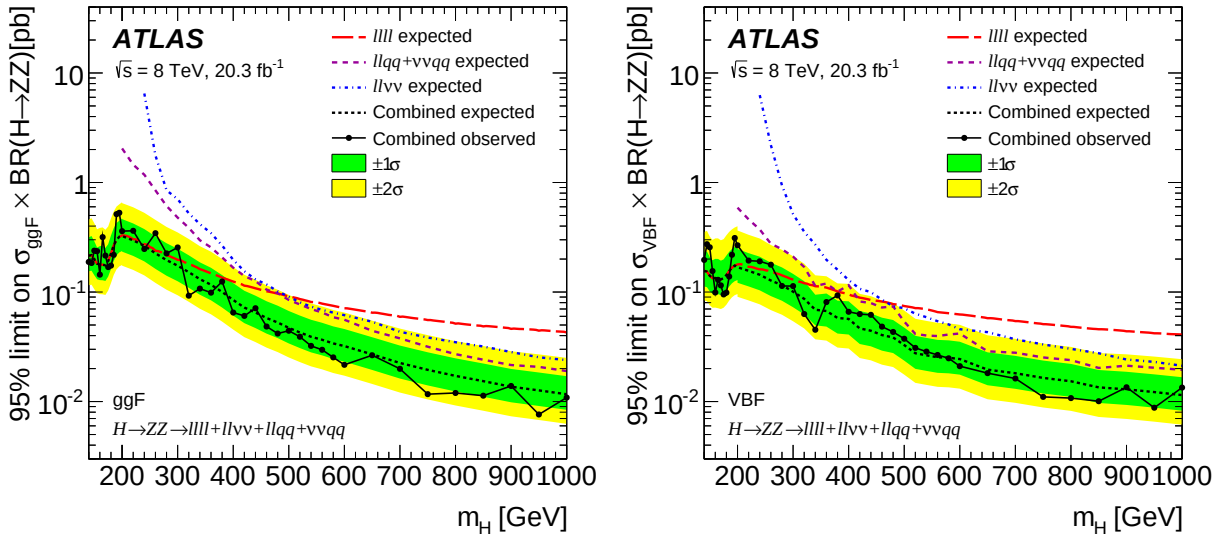


Figure 2.8: 95% CL upper limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ as a function of m_{H}, resulting from the combination of all of the searches in the (a) ggF and (b) VBF channels. The solid black line and points indicate the observed limit. The dashed black line indicates the expected limit and the bands the 1-σ and 2-σ uncertainty ranges about the expected limit. The dashed coloured lines indicate the expected limits obtained from the individual searches; for the $\ell\ell qq$ and $qq\nu\nu$ searches, only the combination of the two is shown as they share control regions. Figures are taken from [63].

di-jet system as the discriminating variable. The ggF category splits the un-tagged and tagged version, which requires exactly two $b-$tagged jets. The m_{T} range covered by the two channels is between 300 GeV and 1.2 TeV.

The results are reported as combined upper limits on the ggF and VBF production cross sections, which are derived from the combined fit across all signal regions. In Figure 2.8, the combined observed and combined expected upper limits from all four decay modes are plotted, and the individual 4ℓ, $\ell\ell\nu\nu$, $\ell\ell qq + qq\nu\nu$ expected limits are overlaid. As expected the 4ℓ is seen to dominate the sensitivity at low m_{H} values while in the intermediate and high mass ranges the sensitivity is provided by the $\ell\ell\nu\nu$ and $\ell\ell qq + qq\nu\nu$ searches. The combination of all four measurements produces improved observed and expected limits due the constraint of correlated uncertainties between different analyses. The observed upper limit is found to be compatible with the expected.

Following the discussion in section 2.5, the exclusion limits for observation of type I and type II 2HDMs are produced. As mentioned before the studies on 2HDM are characterized by a few parameters such as $\tan(\beta)$, $\cos(\beta - \alpha)$ and m_{H}. The $\tan(\beta)$ parameter is related to the ratio of vevs for each field and provides constraints on the mixing between the charged scalars and the pseudoscalar while $\cos(\beta - \alpha)$ is linked to the coupling of the heavy Higgs to the gauge bosons. All of these studies are performed under the assumption of a narrow

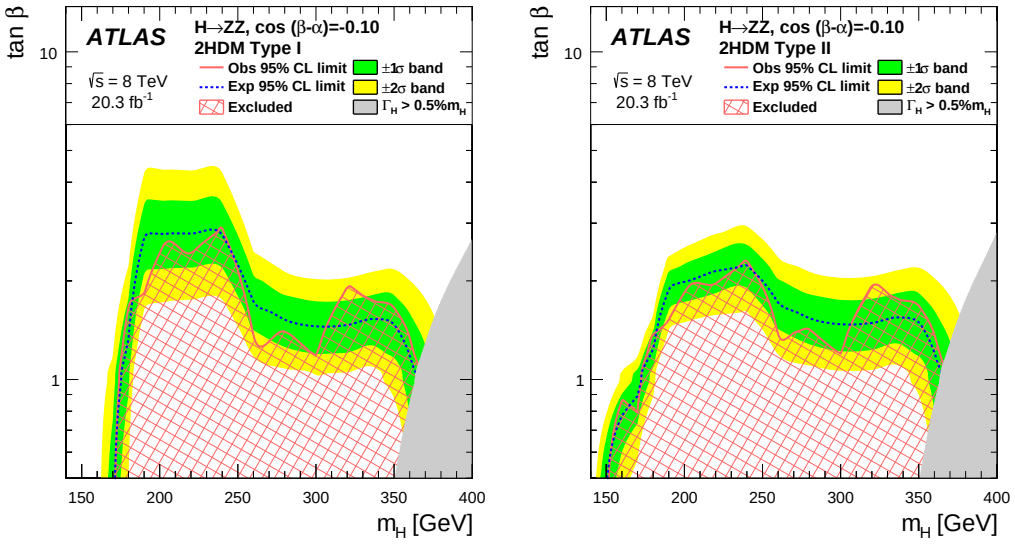


Figure 2.9: 95% CL exclusion contours in the 2HDM (left) Type-I and (right) Type-II models near the alignment limit $\cos(\beta-\alpha) = -0.1$, shown as a function of the parameters $\cos(\beta-\alpha)$ and $\tan\beta$. The red hashed area shows the observed exclusion, with the solid red line denoting the edge of the excluded region. The dashed blue line represents the expected exclusion contour and the shaded bands the 1-σ and 2-σ uncertainties on the expectation. The vertical axis range is set such that regions where the light Higgs boson couplings are enhanced by more than a factor of three from their SM values are avoided. Figures are taken from [63].

width Higgs boson with m_{H} greater than 150 GeV. For a broader discussion of 2HDMs at the LHC run II the reader is referred to [64].

In Figure 2.9 the $\tan(\beta)$ is scanned for an m_{H} range between 150 GeV and 400 GeV for both type I and type II 2HDM models. The study is produced near the alignment limit where the value of $\cos(\beta-\alpha)$ is set to -0.1 i.e. within the restraint given for SM-like h (where $\sin(\beta-\alpha) \geq 0.99 \rightarrow |\cos(\beta-\alpha)| \sim 0.1$). The crossed red lines indicate the excluded regions at 95% confidence level while the grey band indicates regions where the Higgs width cannot be approximated by the narrow width assumption. In the range of $200 \lesssim m_H \lesssim 250$ GeV the region with $\tan(\beta) \sim 2$ is excluded at 95% CL. Small $\tan(\beta)$ values ($\lesssim 1$) are excluded for the entire m_{H} range[3].

The Higgs mass is fixed at $m_{\mathrm{H}} = 200$ GeV, and the exclusion contours are plotted for $\tan(\beta)$ against $\cos(\beta-\alpha)$ in Figure 2.10. This figure indicates that for regions at negative $\cos(\beta-\alpha)$ the exclusion region covers almost the entire range at low $\tan(\beta)$. For positive $\cos(\beta-\alpha)$ below 0.5, low $\tan(\beta)$ values can be excluded and for $\cos(\beta-\alpha) > 0.6$, the exclusion contours are able to exclude more regions at high $\tan(\beta)$.

[3] The main sensitivity in these studies comes from the 4ℓ searches, with $\ell\ell\nu\nu$ contributing a certain amount after 300 GeV.

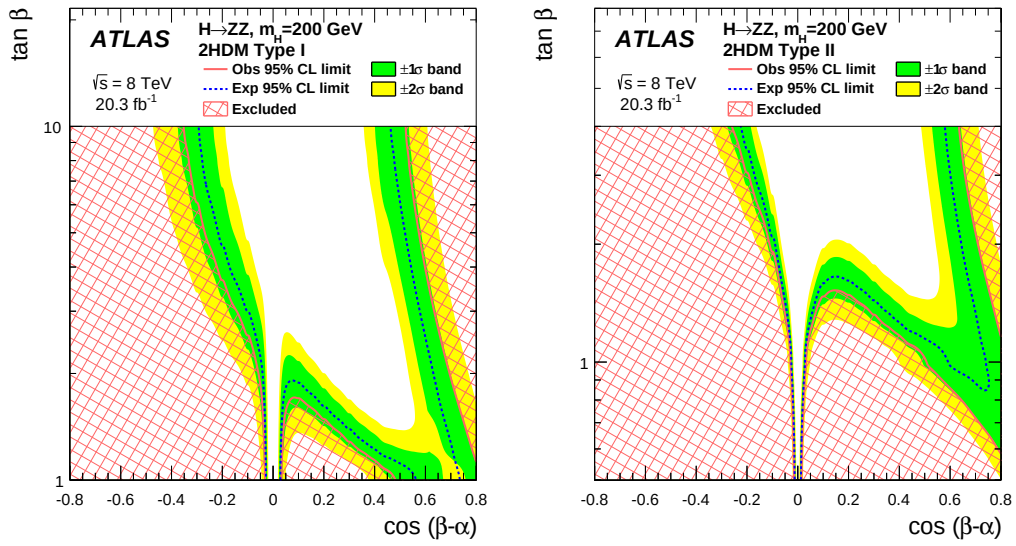


Figure 2.10: 95% CL exclusion contours in the 2HDM (left) Type-I and (right) Type-II models for $m_{\mathrm{H}} = 200$ GeV, shown as a function of the parameters $\cos(\beta - \alpha)$ and $\tan\beta$. The red hashed area shows the observed exclusion, with the solid red line denoting the edge of the excluded region. The dashed blue line represents the expected exclusion contour and the shaded bands the 1-σ and 2-σ uncertainties on the expectation. The vertical axis range is set such that regions where the light Higgs couplings are enhanced by more than a factor of three from their SM values are avoided. Figures taken from [63].

Chapter 3

The ATLAS Experiment

ATLAS, short for "A Toroidal LHC ApparatuS", is one of the largest and most complex particle physics detectors ever built. It serves as a microscope to probe the behaviour of interactions on the very small scales, $\sim 10^{-18}$ m. ATLAS, along with the CMS detector, performs a wide range of particle physics measurements, to test the validity of the currently accepted standard model of particle physics and to search for new unexplained physics phenomena.

This chapter starts by describing the Large Hadron Collider (LHC) facility, its structure and function in section 3.1. Then an overview of the ATLAS detector is presented in section 3.2. Afterwards, each of main hardware subcomponents are introduced, the inner detector (ID) in section 3.3, the calorimeter system in section 3.4 and the muon spectrometer (MS) in section 3.5. An important software component of the ATLAS detector, the triggering system, in presented in section 3.6. The data used in this thesis is summarised in section 3.7. Lastly, a few details on the detector simulation infrastructure are provided in section 3.8.

3.1 The Large Hadron Collider

The LHC facility houses an accelerator complex which collides protons with protons accelerated to energies of 6.5 TeV per proton beam [65] resulting in a centre of mass energy, $\sqrt{s}$, of 13 TeV. It spans 27 km in diameter and it is located roughly 100 meters below ground at the French-Swiss border near Geneva. The proton beams consist of packed short-pulse bunches which contain roughly 10^{11} protons, which form a bunch train. The protons are made by stripping the outer electron of hydgrogen atoms in the duoplasmatron source shown in Figure 3.1.

The proton acceleration goes through several cycles, which are outlined schematically in Figure 3.2. The first step is a linear accelerator, which accelerates the protons up to 50 MeV.



Figure 3.1: Source of hydrogen atoms at the LHC.

Direct beam injection into the ring section, without an initial proton boost is not stable due to residual magnetic fields present in the ring cavities. The injection proceeds to the booster ring, where the proton beams achieve an energy of 1.4 GeV and are passed to the proton synchroton, where the energy is increased to 26 GeV. The penultimate step is the super proton synchroton, where the beams gain an energy of 450 GeV. The last step is the injection into the main LHC ring, where the energy in each beam, E_b, goes up to 6.5 TeV, providing $\sqrt{s} = 2 \times E_b = 13$ TeV.

The acceleration of protons in the LHC is achieved via radio frequency superconducting cavities which use EM fields, oscillating at a frequency of 400 MHz. A total of 1232 dipole magnets are used, in order to bend the protons through the acrs of the LHC. Two separate beam tubes and two different magnetic fields gradients accelerate and bend the protons, which circulate in opposite directions. The dipole magnets achieve a field strength above ~ 7 T, which is generated by a current of 11.8 kA through superconducting niobium-titanium (NbTi) coils cooled down to 1.9 K using superfluid helium. 858 quadrupole magnets focus and bunch together the protons to prevent the dispersion of the proton beams. In order for the cryogenic systems to be isolated from the room-temperature parts of the machine, a high insulating vacuum ($\sim 10^{-6}$ mbar) is required. A high vacuum ($\sim 10^{-11}$ mbar) is also required in order to increase the mean lifetime of proton beams.

There are four collision points throughout the LHC, where each of the four main detectors are placed. ATLAS and CMS are general purpose detectors which record all collisions of interest while the LHCb detector looks for events with B-hadrons and ALICE studies busy multi track events produced by either heavy ion or proton collisions.

Proton-proton colliders do not suffer from certain physics related problems which can plague other accelerator machines such as the LEP experiment, which collided electrons with

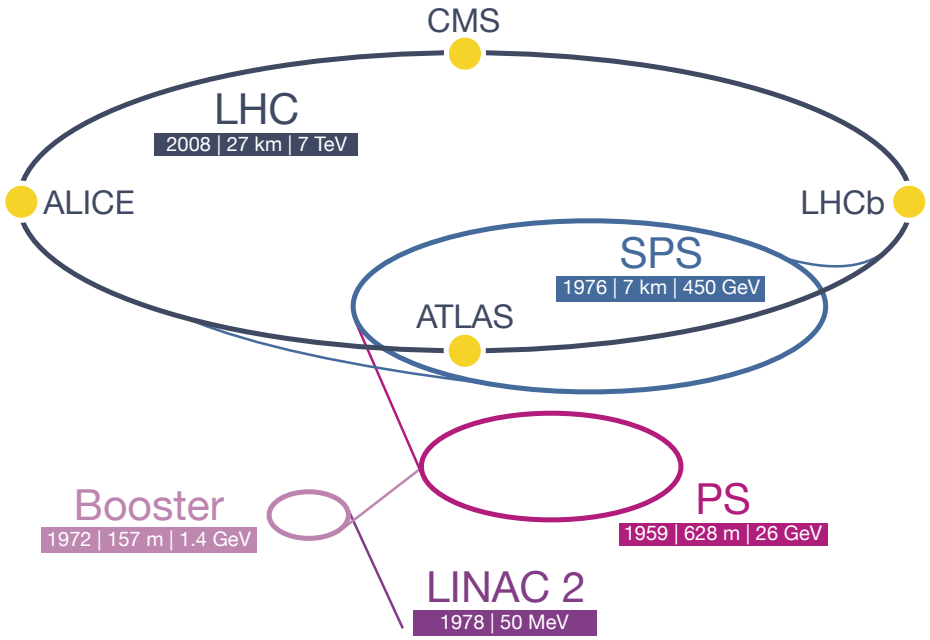


Figure 3.2: This figure shows a diagram of the CERN accelerator complex. The speed and energy of the injected protons is increased by a chain of accelerators: a linear accelerator (Linac), the Booster, the Proton Synchotron (PS), and the Super Proton Synchotron (SPS) accelerators, before entering the LHC ring. The protons are then accelerated further in opposite directions, before colliding inside the ALICE, CMS, LHCb and ATLAS detectors. This figure is adapted from [66].

positrons and had very large energy losses due to the emission of synchrotron radiation. Another example is the Tevatron, which collided protons with anti-protons and it wasn't able to achieve a very high luminosity because of difficulties in antiproton production.

However, *pp* machines have other physics related problems. Whereas in LEP, the energy of the initial state is known to a good precision, at the LHC the initial energy state between the colliding quarks and gluons is unknown as they carry some fraction of the proton energy. Their initial energy state is described by parton distribution functions (PDF) [67].

3.2 The ATLAS detector

ATLAS is a general purpose detector which extends 44 m in length and 25 m in height and consists of sub-detectors arranged in an multi-layered cylindrical configuration. Its total combined mass reaches nearly 7000 tonnes and it containes 100 million readout channels, which process particle collisions at a rate of 40 MHz, to keep in-time with the fast 25 ns bunch crossings. ATLAS is able to precisely reconstruct the position, momentum and energy of charged and neutral particles emerging from the interaction point, which allows for precise

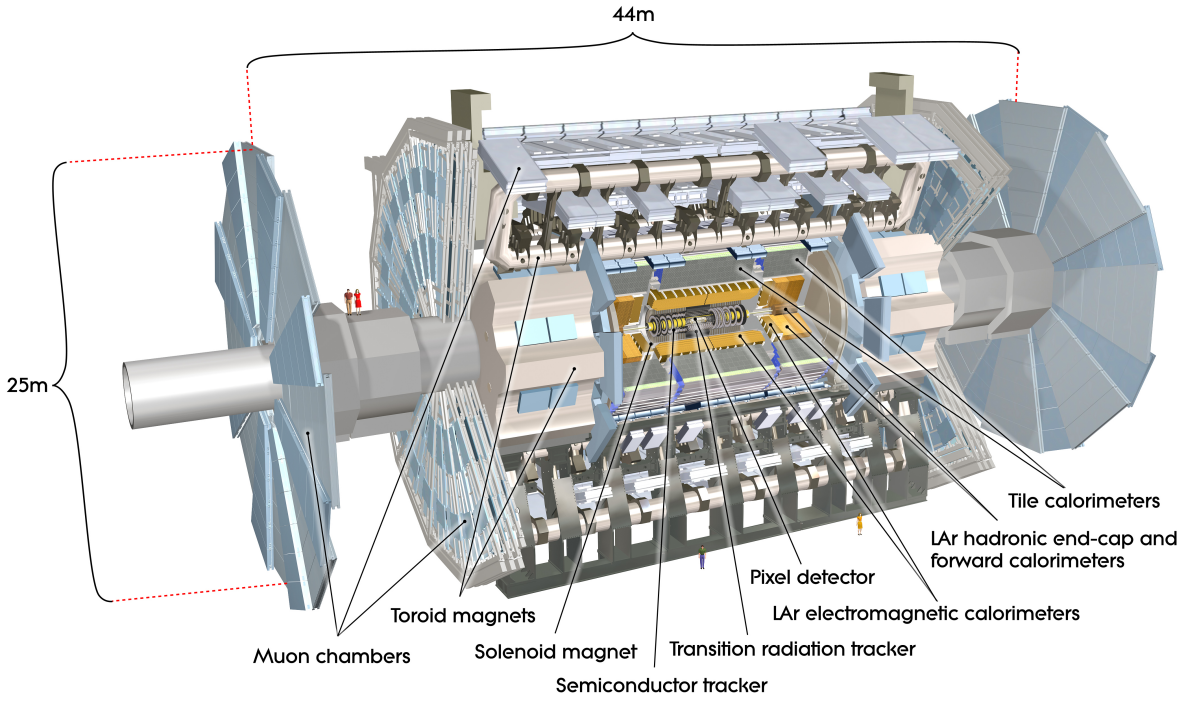


Figure 3.3: The ATLAS detector and its sub-detectors are shown. This figure is taken from [69].

studies of the properties of high energy physics processes [68]. The detector has a near complete 4π coverage, which provides reconstruction of the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, which is used as an indicator of the energy carried away by weakly interacting particles such as the neutrinos. The purpose of the ATLAS experiment lie in both testing the currently accepted theory of particle physics i.e. the SM to higher precision than previous detectors of its kind achieved and to search for new physics phenomena. Searching for new physics is done by looking for an excess number of events in data over the SM background expectation.

A schematic view of the ATLAS detector is available in Figure 3.3. ATLAS uses a right-handed coordinate system whose origin is at the center of the detector. The $+x$ coordinate points inwards, towards the center of the LHC ring, the $+y$ coordinate points upwards from the beam pipe, and the $+z$ coordinate points in the anti-clockwise direction along the beam line. In ATLAS it is more practical to use cylindrical coordinates, with r representing the distance from the $z-$axis origin, the azimuthal angle, ϕ, measured from the $x-y$ plane. In particle physics, an additional coordinate is used to denote the angle of a particle relative to a beam axis. This coordinate can be used for both massive and massless particles, but in this thesis, only the massless (and Lorentz invariant) definition is used. This is the pseudorapidity, η, defined as a function of the polar angle, θ as:

$$\eta = -\ln\left(\tan(\theta/2)\right). \tag{3.1}$$

Here the polar angle is defined from the $z-$axis, with positive and negative z represented by

$\theta = 0, \pi$. In the $x-y$ plane, the particle momentum and energy are given by their transverse components p_{T} and E_T and their position by η. Angular separation between particles is provided in the $\eta - \phi$ plane by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The detector is divided into three general regions: the central (barrel) region and two forward (end-cap) regions. The central region extends up to $|\eta| < 2.4$, while the forward region extends between $2.4 < |\eta| < 4.9$. ATLAS consists of four general detector subsystems to accurately track and capture the emerging particles from the event collision: the inner detector (ID) is used for tracking of charged particles, the EM calorimeters capture electrons and photons which produce EM showers, the hadronic calorimeter captures all hadron particles and finally the muon spectrometer (MS) are used to track the only surviving particle type, the muon. ATLAS uses two separate magnetic systems: a superconducting solenoid coil, which generates a solenoidal field of 2T strength in the ID and a toroidal field in the muon systems. Three toroidal magnets, one in barrel and two in the endcaps, are used to generate toroidal field. In the ID, the solenoidal field is very uniform allowing for very precise measurements while the toroidal magnetic fields have varying strengths across different parts of the detector resulting in a complicated configuration [70].

Due to the extremely high collision rate, at a frequency of 40 MHz, dedicated trigger systems are employed to record events which might show interesting physics. The total available rate of recording is 1 kHz, so the trigger systems must reject on the average of 10^6. The ATLAS trigger system is divided into two general levels: the Level-1 (L1) trigger and high level trigger (HLT). The L1 trigger reduces the rate down to $\sim$ 100 kHZ by using information from the calorimeter and muon detectors. HLT compromises of an L2 trigger and special event filters, which makes use of more precise information from the calorimeters, ID and MS tracking info.

3.3 Inner Detector

The ATLAS ID offers a highly accurate track and vertex reconstruction for all charged particles emerging from the interaction point. The precise track reconstruction is very important as it allows to measure both the position of the particle and from the track curvature determine its charge and momentum. The ID is made up from three main components: the silicon pixel detector, the semiconductor tracker (SCT) layers and the transition radiation tracker (TRT). A cut-away schematic of the ID is shown in Figure 3.4.

The silicon pixel detector [71] uses cutting-edge technologies to provide precise reconstruction of the event collision. It is made up of 1744 sensors, with size 19×63 mm^2 and thickness 250 μ m. Each sensor contain 47268 pixels with varying sizes. 90% of the pixels have a size

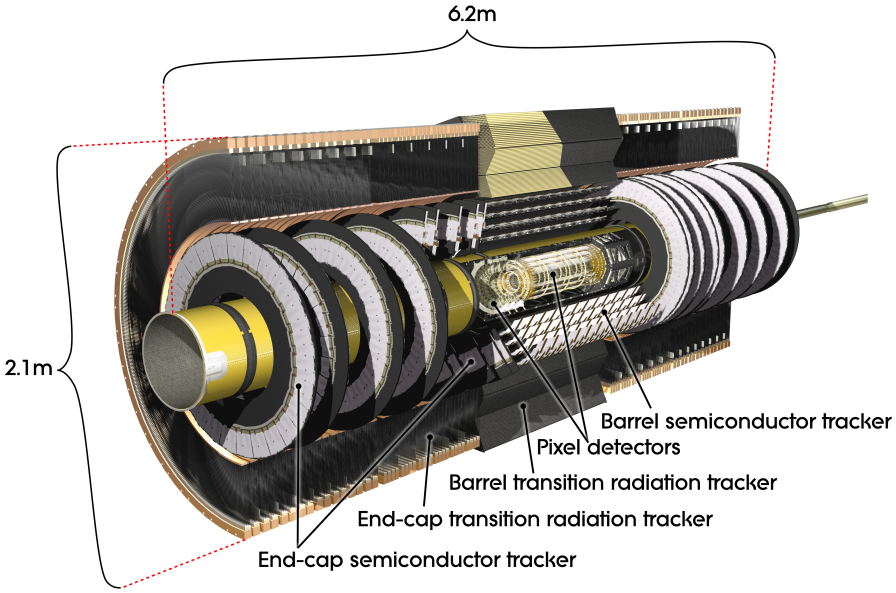


Figure 3.4: A cut-away view of the ATLAS inner detector used in Run-1. Figure taken from [68].

of 50×400 μm while the 10% front end modules are 50×600 μm. This gives a total of 80 million readout channels.

When the charged particles emerge from the interaction point and enter the pixel sensors, they create electron-hole pairs which drift under the influence of a bias voltage ($\sim$ 150 V) to the cathode and anode parts of the sensor. The signal is then amplified in the readout chip to produce an observable signal. The precision mostly depends on the pixel dimensions, with current precision $\sim$ 10 μm.

An additional layer was added to the pixel detector between the beam pipe and the innermost pixel (B-) layer during the long shutdown of the LHC between 2012 and 2015, which is called the insertable (B-) layer (IBL) [72]. The IBL [72] consists of cylindrical support structures around the beam pipe, known as staves, which contain both planar and 3D silicon sensors. The average distance between the sensors and beam pipe is 33mm. The total number of staves is 14, each with size of $64 \times 2\text{cm}^2$ and containing 32 individual chips, where each chip has 26880 pixels. The schematic view of the IBL is shown in Figure 3.5. The main considerations for introduction of the IBL were the degradation in precision measurements of the impact parameters from the pixel B-Layer due to the intensive radiation environment [73].

The SCT detectors provide precise track reconstruction of the charged particle paths through it. The charge of the particle is determined from the direction of bending of the track under the solenoidal field, and the momentum is obtained from the amount of bending under the field. The SCT is placed after the pixel detector and consists of 4088 modules of silicon strips,

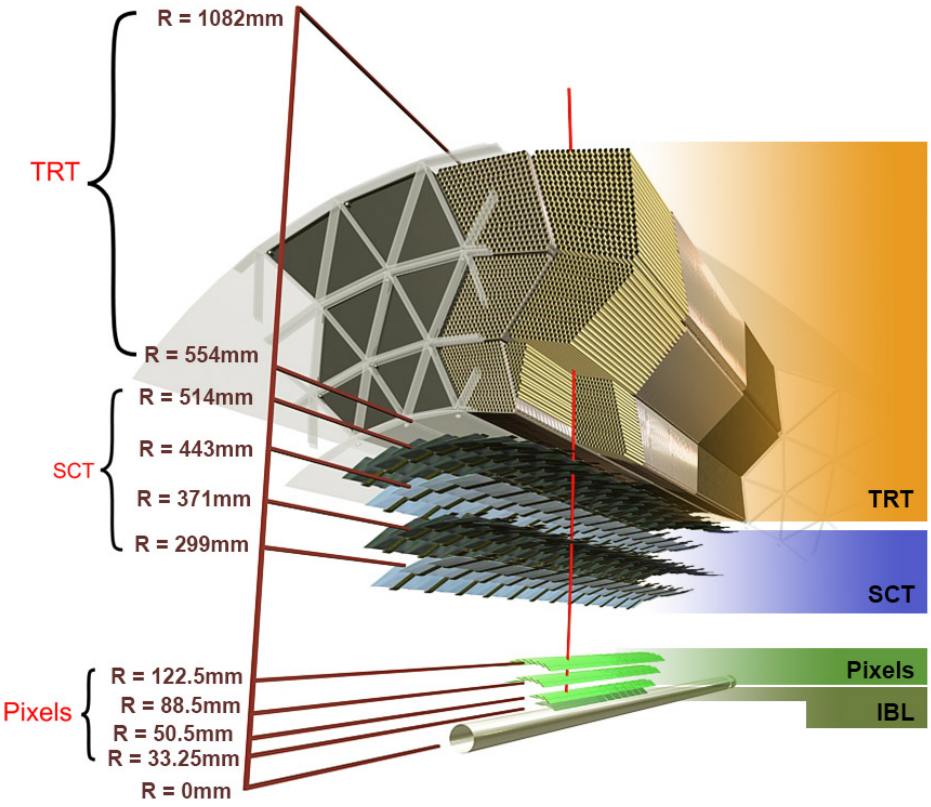


Figure 3.5: The sub-detectors of the inner detector are shown for the barrel region, together with their distance from the beam line. From the beam line outwards, the sub-detectors are: the Insertable B-layer (IBL), the pixel detector, the semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). This figure is taken from [73].

in four concentric barrel layers and two end-cap regions. Each module is made from strips with 80 μm width, with two sensors on each side chained together to provide 768 strips of total length $\sim$ 12 cm. Back-to-back sensors are sticked together at a stereo angle of 40 mrad to determine the space-points of particle tracks and resolve ambiguities in measurements. The SCT has an overall number of $\sim$ 6.3 million readout channels.

The TRT consists of straw drift tubes which provide a detailed reconstruction of the charged particle tracks and discrimination between different particle types [74]. The straw tubes are designed with an anode wire in the middle of the tube and a mixture of $Xe/CF_4/CO_2$ gases, which have a high absorption rate of the transition photons. Transition photons are produced when charged particles pass through mediums with two different di-electron constants. The TRT is divided into a barrel region ($|\eta| < 1.0$) and two endcap regions ($1.0 < |\eta| < 2.0$). The barrel contains a total of 52 544 straws where each straw has a length of 144 cm. They are placed parallel to the beam axis and spread across 96 modules and are split into 32 modules to fill any gap regions in the azimuth ϕ. Each end-cap contain 122 880 straw tubes, each with a length of 37 cm arranged in a wheel geometry. The total readout channel number is twice the number of straws because each end of the anode wires is read out separately.

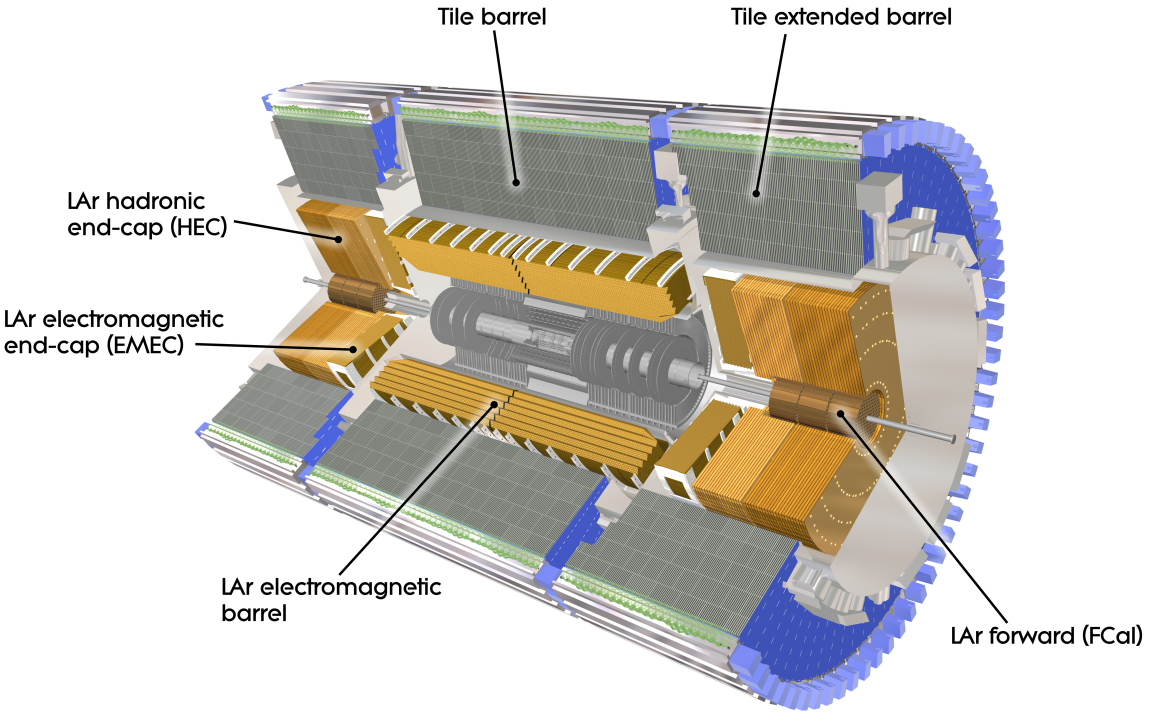


Figure 3.6: The layout and sub-detectors of the calorimeter systems are shown. The calorimeters lie outside of the inner detector and the solenoidal magnet system. The electromagnetic calorimeter consists of the LAr barrel calorimeter, the inner LAr end-cap calorimeters, and the first layer of the forward calorimeters. The hadronic calorimeter consists of the tile barrel calorimeter, the tile extended barrel calorimeter, the outer LAr end-cap calorimeters, and the second two layers of the forward calorimeters. This figure is taken from [77].

3.4 Calorimeter system

Reconstruction and capture of the particles is achieved through the ATLAS calorimeter subsystems, shown in Figure 3.6. Electrons, positrons and photons are exclusively captured in liquid argon (LAr) sampling calorimeters [75]. Hadrons are reconstructed using the tile calorimeter in the central barrel region and LAr calorimeters in the forward end-cap regions. The total coverage for all hadronic objects reaches $|\eta| < 4.9$ [76]. The particles going through the calorimeters deposit their energy via ionisation showers i.e. charged particles traversing material in the calorimeters (known as absorber material) generate an electromagnetic shower through bremsstrahlung and pair production which continues until the photon and electron energy become too low. The complete reconstruction of the shower provides the energy of the original parent particle, and the shower profile can provide additional information on the original particle type.

3.4.1 Electromagnetic calorimeters

The LAr sampling calorimeter is made from interleaved lead/stainless steel absorbers, which develop the EM shower, and sub-detectors which measure the energy of the particles in the shower. These sub-detectors consist of electrodes submerged in liquid argon, cooled to a temperature of 88K. When a charge particle passes through, it frees the ions and electrons in the liquid argon which drift under the influence of the electric field between the electrodes. The charge is collected on the electrodes and amplified to provide a large signal gain. The collected charge is proportional to the energy of the charged particles [78]. The choice of liquid argon is motivated by several factors such as precise energy resolution, the stability of argon (due to being a noble gas), radiation hardness and abundance.

The design of the LAr calorimeter system is divided into a central barrel chamber and two EM end-cap chambers (EMEC). The LAr barrel chamber spans $|\eta| < 1.475$ and consists of two half-barrels with an accordion shaped geometry giving a uniform ϕ-coverage. The LAr central barrel is divided into three layers, the first with fine segmentation to record precise information of the shower position, while the second and third layer are coarser to gather the shower energy. A liquid argon sub-detector is placed before the LAr barrel, known as a pre-sampler which measures the energy from electrons and photons before the particles enter the calorimeters and is used to estimate the lost energy in the calorimeter system. The central barrel geometry is presented in Figure 3.7. The region $1.37 < |\eta| < 1.52$ is typically removed from most analyses, as it is a transitional region between the barrel and end-cap region. It contains detector support structures which are not equipped for precise energy measurements. The EMEC has an accordion geometry, similar to the barrel chamber, but rotated by 90° to have full coverage.

3.4.2 Hadron calorimeters

The calorimeters used to reconstruct the hadronic activity are composed of the tile barrel calorimeter, the hadronic end-cap calorimeter (HEC) and the forward calorimeter (FCAL).

The tile calorimeter [79] extends to $|\eta| < 1.7$ and is positioned outside of the LAr calorimeter. It consists of three separate cylinders, one central barrel spanning $|\eta| < 1.0$ and two barrels between $0.8 < |\eta| < 1.7$ on each side. Plastic scintillators are used as an active component [80] to reconstruct the shower while the absorber material is iron. Each barrel chamber is segmented into 64 "wedges" (modules) in the azimuth ϕ, which results in a $\Delta\phi$ granularity of 0.1 radians. Three dimensional cells are formed from the ϕ, η and radial segments. The tiles are arranged into three layers, where the first two layers have a cell size of $0.1 \times 0.1(\Delta\eta \times \Delta\phi)$ while the third has a coarser granularity $0.2 \times 0.1(\Delta\eta \times \Delta\phi)$. The

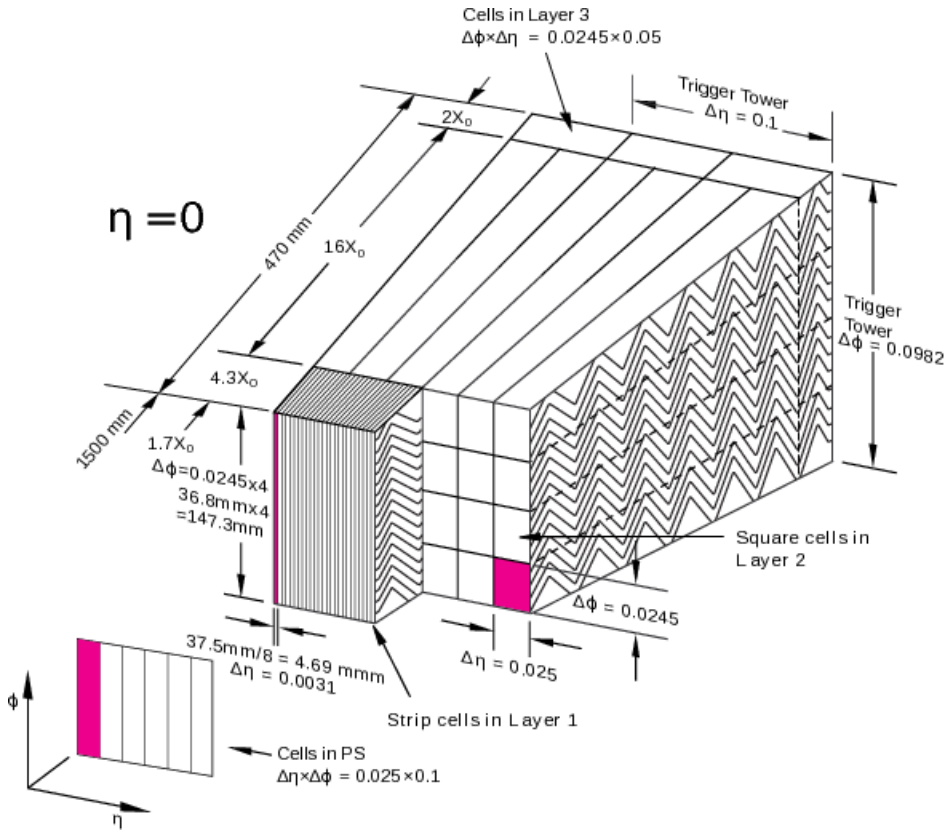


Figure 3.7: This diagram shows the layout and segmentation of the electromagnetic barrel calorimeter and the pre-sampler. The dimensions of the calorimeter cells and trigger towers are illustrated. This figure is taken from [75].

coarser granularity for the tiles, in comparison to the LAr calorimeter, is due to the longer shower profiles of the hadrons. A schematic of the tile calorimeter is shown in Figure 3.8. Wavelength shifting fibres are attached to either cell edge, which transmit the collected light from the passing charged particles through the tiles[1]to photomultiplier tubes where the signal is amplified. The energy of the particle is proportional to the intensity of the collected light [81].

Beyond the barrel is the HEC, which extends between $1.5 < |\eta| < 3.2$. The HEC compromises two wheels on each side of the barrel chamber. Each wheel uses flat (instead of accordion geometry) parallel copper plates which act as the shower absorber material and the active material is liquid argon.

The FCAL [82] is located between $3.1 < |\eta| < 4.9$ and is constructed differently than the LAr calorimeters. The different design is motivated by considerations of radiation robustness. The FCAL is constructed from three cylindrical parallel plate modules, the first made from copper and used for mostly EM measurements, while the latter two are made from

[1]When a particle passes through the scintillation tiles, it produces excitations in the atoms, which emit light when they de-excite.

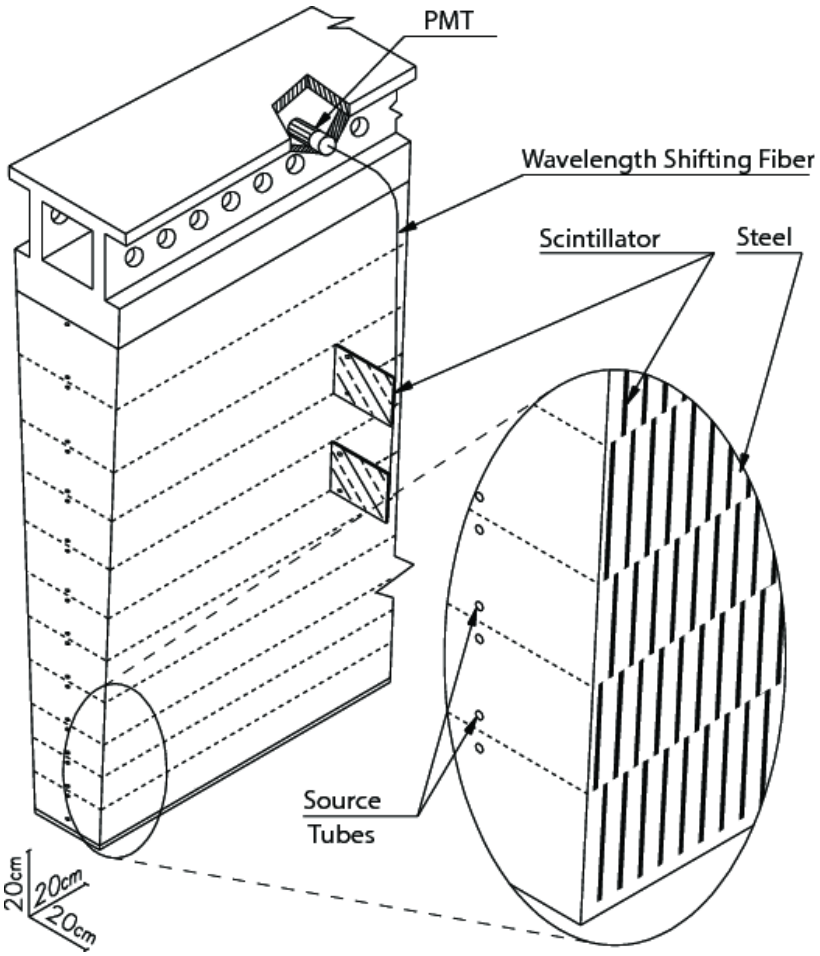


Figure 3.8: This diagram shows the layout of the tile barrel calorimeter, including its alternating steel and scintillator structure and the light detection system. In the detector, this section would be positioned parallel to the beam line, with particles entering from the bottom of the section. This figure is taken from [79].

tungsten and used for reconstruction of hadronic activity. The FCAL uses absorber structure, consisting of rods and cylindrical tubes made from electrodes, with narrow liquid argon gap between the tube and rod. Liquid argon is used to ensure a fast readout time and reduce buildup of ions.

3.5 Muon Spectrometer

The muon spectrometer (MS), shown in Figure 3.9, forms the outermost detector in ATLAS [83]. The spectrometer is split into a barrel section ($|\eta| < 1.05$) and two endcap regions, ($1.05 < |\eta| < 2.7$). The MS is made from monitored drift tubes (MDTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs) and thin gap chambers (TGCs). A toroidal magnetic field is provided by three superconducting air-core toroidal magnets, which give a 0.5 T field strength in the barrel and 1 T in the endcaps. With these design specifications,

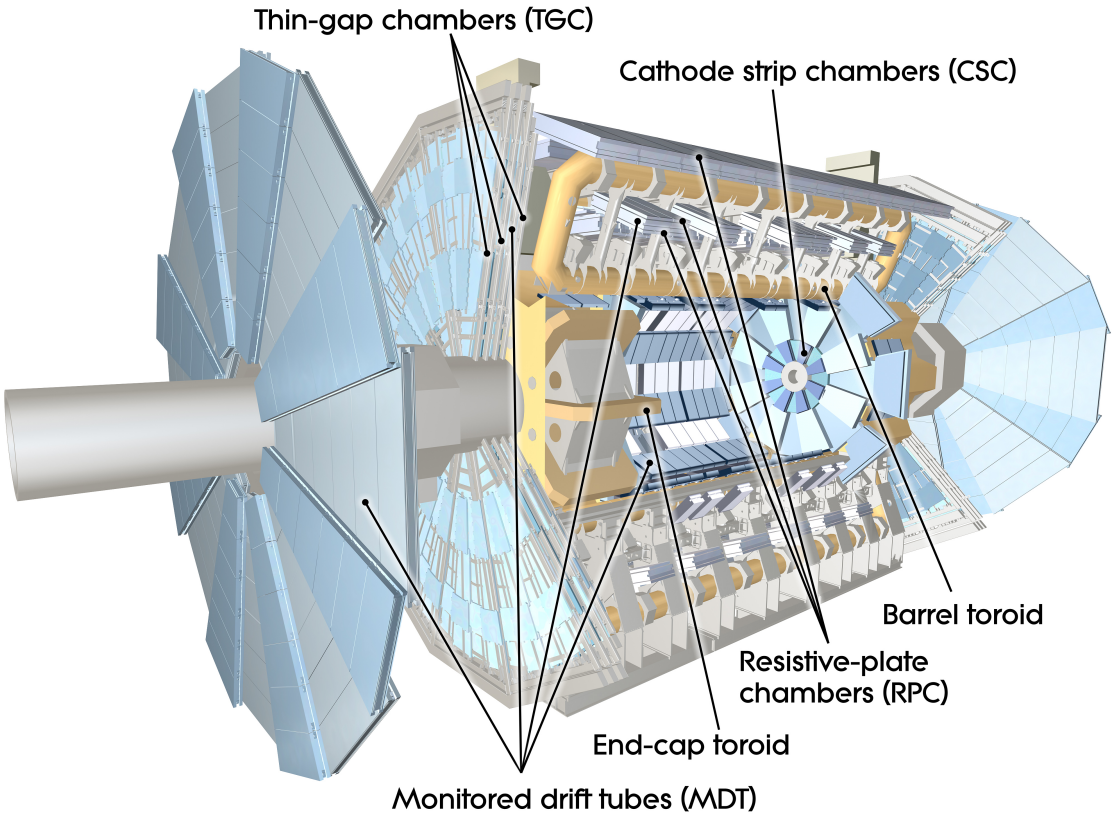


Figure 3.9: The sub-detectors of the muon spectrometer are shown, as well as the toroidal magnet system. The Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC) provide precision position information. The Resistive-plate Chambers (RPC) and Thin-gap Chambers (TGC) provide coarse position information and trigger capabilities. This figure is taken from [84].

the MS is able to provide a relative resolution of 3% over a wide p_{T} scale, reaching 10% for $p_{\mathrm{T}} = 10$ GeV.

The MDTs [85] are made from pressured drift tubes, with diameter ~ 30 mm which holds Ar/CO_2 in 93%/7% proportion at 3 atmospheric bars. The drift tube chambers are arranged in three sequential layers. The ionised electrons are collected at a central tungsten-rhenium wire operating at a 3080 V. There are 1088 drift tube chambers which cover an area of 5500 mm^2. Each chamber provides $\sim 6-8$ separate η measurements of the muon path. The spatial resolution per hit is 80 μm with alignment precision of 30 μm.

The CSCs [86] are multi-wire proportional chambers (MWPCs) with cathodes segmented into orthogonal strips, which allow the measurement of both spatial coordinates. The CSCs are divided into two disks with eight chambers, where each chamber has four planes which provide four independent η, ϕ readouts. The operating voltage in each chamber is 1900 V with a gas mixtures Ar/CO_2 in 80%/20% proportion. The CSCs are placed in the end-cap region, $2.0 < |\eta| < 2.7$, since they can sustain a higher hits rate and achieve good time

resolution. In addition, they provide a good spatial measurement with a hit resolution of ~ 60 μm and an alignment precision of ~ 60 μm.

The RPCs ($|\eta| < 1.05$) [87] are constructed from three concentric cylindrical layers, each containing parallel electrode plates with a gaseous mixture. When the muon particle passes through the detector, an ionisation avalanche is created which is then collected at each electrode. The electric field strength is 4.9 kV/mm, which allows a timing resolution of 5 ns. The first two layers are closer in distance than the remaining layers which allows triggering on low p_{T} muons ($\sim 6 - 9$ GeV) in the first two layers and triggering on high p_{T} muons ($\sim 9 - 35$ GeV) for the last two layers.

The TGCs [88] are placed in end-cap region ($1.05 < |\eta| < 2.4$). They use MWPCs similarly in design to the CSCs. The large voltage potential (2900 V), results in a large electric field strength, which combined with the small wire-to-wire separation leads to a good time resolution. The wire width varies from ~ 11 mm $- 56$ mm as a function of η to achieve the desired granularity. In total the 9 wheels include the TGCs and layers from the MDTs. The TGCs achieve a spatial hit resolution of $2 - 6$ mm in the radial coordinate and $3 - 7$mm in the azimuth.

Overall the spatial resolution from the RPCs and the TGCs is of the order of $5-10$ mm while the timing precision is $\sim 15 - 25$ ns, enough for fast triggering and reliable beam-crossing identification.

3.6 Trigger systems

The trigger system is responsible for selecting an event for further processing and studies in analyses. It is essential for recording the data of interest, as the LHC delivers a high rate of pp collisions, most of which are low energy scatterings.

In run II, the ATLAS trigger systems were upgraded to handle the large 40 MHz collision rate and the short 25 ns bunch spacing [89]. The trigger systems are divided into two general categories: the Level-1 (L1) trigger and high level trigger (HLT). The L1 trigger consists of hardware eqiupment which performs quick event decisions based on calorimeter and muon trigger information. The HLT, in run I, consisted of a Level-2 trigger which required partial event data, and event filter farms which use the full detector information. In run II, the two elements were combined into a single farm, to reduce hardware and software complications and to achieve better sharing of resources. The output rate from the L1 trigger is 100 kHz, while the longer processing of HLT allows the rate to be reduced to 1 kHz [90].

Triggers exist for all reconstructed objects such as the $e^{\pm}, \mu^{\pm}, \tau^{\pm}$, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. In ad-

dition, minimum bias triggers are employed, which collect events using the minimum bias trigger scintillators [91]. Events recorded by the minimum bias triggers are used for studies in hadronisation, diffraction, total cross section measurements and non-perturbative QCD interactions [92].

The L1 trigger is hardware-based with a fast decision rate ($\sim 2.5\ \mu s$) and is comprised from two trigger algorithms, one of which uses calorimeter granularity information (L1 Calo) while the other gathers information from the muon detectors (L1 Muon). The L1 Calo trigger algorithm implements a sliding window technique [93] to group calorimeter cell towers in "trigger towers", which have dimensions of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. The trigger towers identify maxima in the transverse energy E_T. The regions of interest (RoI) corresponding to candidate objects are defined around the maxima. The RoI is passed to the HLT for regional reconstruction in the algorithms.

The RoIs for electrons is a 2×2 trigger tower cluster in the EM calorimeter for which the sum of transverse energy in at least one of the four pairs of nearest neighbour towers is greater than a certain threshold energy value. The thresholds can vary in different regions, due to the different detector energy responses. The tower energies are calibrated at the EM scale, which fully reconstructs the energy deposited by electrons.

The RoIs are modified in the case for jets since hadronic showers are not always perfectly reconstructed. The RoI for jets are either a 4×4 or 8×8 tower windows which sum both EM and hadronic transverse energies which must pass a certain energy value, and in addition must consist of a 2×2 tower core which has a local maximum. The different sizes of tower windows perform different tasks, with larger windows able to reconstruct more of the hadronic energy, while smaller windows better separate jets.

For the L1 Muon algorithm, as briefly mentioned in section 3.5, the RPCs and TGCs are used to trigger on muon objects. Three hits are required either in the RPCs or the TGCs to trigger a muon. Exception is made for low p_T muons in the RPCs which require only two hits. In run II, a fourth layer of the RPC was brought online to recover some trigger acceptance from losses due to support structures. Additional chambers were also installed to further minimise losses from acceptance holes.

The events which pass the L1 triggers are then processed in the HLT systems, which use tracking information from the inner detector, finer granularity calorimeter information and more precise spatial measurements from the muon detectors. The HLT reconstruction proceeds in either of two ways, either the RoIs from L1 are used or the full detector information is used. If the event passes the HLT then the events are stored for offline reconstruction, calibration and analysis.

Triggers are also classified into un-prescaled and prescaled. Un-prescaled triggers record all events which pass the given set of triggers. The prescaled triggers instead reject a fraction of potentially interesting events. Prescaling triggers ensures that they don't exceed their bandwidth limit.

3.6.1 Single lepton triggers examples

The triggers which are used in the analysis are the single lepton triggers. They consists of the lowest un-prescaled triggers on electron and muon objects plus lepton triggers at large energies to recover some acceptance losses.

In 2015, the recommended single electron trigger was called "`e24_lhmedium_L1EM20VH`". It stands for a 24 GeV requirement on the electron transverse energy, E_{T}^{e}, a "Medium" electron object identification (outlined in section 4.2.1), a L1 trigger with an electromagnetic cluster with $E_{\mathrm{T}} > 20$ GeV, a hadronic veto for an excess of energy in the hadronic calorimeter behind the electromagnetic cluster and $\eta-$dependent E_{T} selection to compensate for inactive detector material. At higher energies, losses in trigger efficiency are partially recovered by a trigger which requires a E_{T} cut above either 60 (120) GeV with medium (loose) electron identification.

The lowest un-prescaled trigger for muons is called "`mu20_iloose_L1MU15`", which requires a p_{T} greater 20 GeV, with loose isolation $I < 12\%$, where $I = \sum_{\mathrm{track}} p_{\mathrm{T}}|_{\Delta R=0.2}/p_T$ is the ratio of the sum of all track p_{T} around a cone of $\Delta R = 0.2$ of the muon candidate and p_{T} is the transverse momentum of the muon. The "`L1MU15`" is a L1 muon object with a p_{T} greater than 15 GeV. To recover some efficiency loss in the high momentum region, a complementary HLT trigger, "`mu50`", with a p_{T} cut of 50 GeV is used.

The performance of single lepton triggers in 2016 data is examined in Figure 3.10 for electrons and Figure 3.11 for muons. The trigger menus are provided in [94].

3.7 Data

3.7.1 Luminosity

The number of events produced from the collisions, N_{events}, is directly proportional to the cross section of a given process. The luminosity is the factor which reflects the amount of particles that are collided at the LHC [97] and it is the factor which connects the cross section for a given process to give the total number of observed events. One can view the cross section as some unknown quantity, fixed by the natural processes and the luminosity as

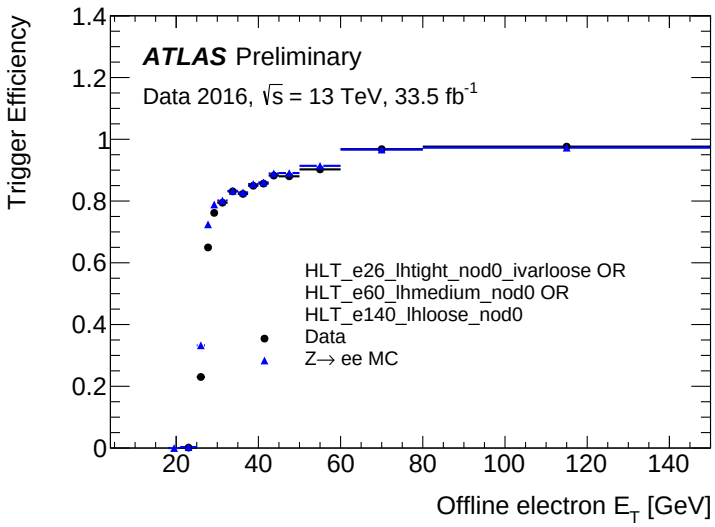


Figure 3.10: Efficiency of the logical OR between e26_lhtight_nod0_ivarloose, e60_lhmedium_nod0 and e140_lhloose_nod0 triggers as a function of the offline electron candidate's transverse energy (E_T). The Level-1 trigger requires an isolated electromagnetic cluster with $E_T > 22$ GeV. The e26_lhtight_nod0_ivarloose trigger requires an electron candidate with $E_T > 26$ GeV satisfying the likelihood-based tight identification without applying transverse impact parameter requirements but applying variable-size cone isolation, the e60_lhmedium_nod0 trigger requires $E_T > 60$ GeV and likelihood-based medium identification, and the e140_lhloose_nod0 trigger requires $E_T > 140$ GeV and likelihood-based loose identification. The offline reconstructed electron is required to pass a likelihood-based tight identification and be isolated. The efficiencies were measured with a tag-and-probe method using $Z \to ee$ decays in data and Monte Carlo. The error bars show the statistical uncertainties. Figure is taken from [95].

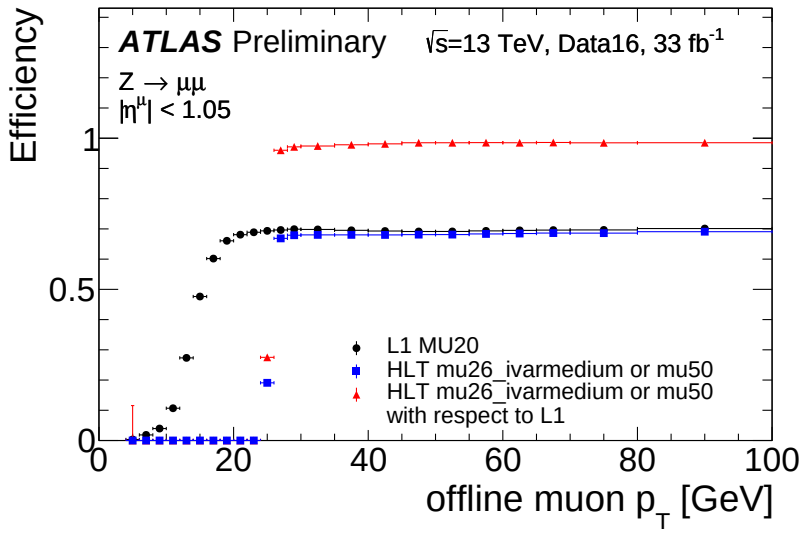


Figure 3.11: Absolute efficiency of Level 1 (L1) MU20 trigger and absolute and relative efficiencies of the OR of mu26_ivarmedium with mu50 High Level triggers (HLT) plotted as a function of p_{T} of offline muon candidates in the barrel detector region. The efficiency is computed with respect to offline isolated muon candidates which are reconstructed using standard ATLAS software and are required to pass "Medium" quality requirement. The MU20 trigger requires that a candidate passed the 20 GeV threshold requirement of the L1 muon trigger system. The mu26_ivarmedium trigger is seeded by the MU20 trigger and is required to satisfy a 26 GeV HLT threshold and to pass a medium isolation selection computed using inner detector tracks reconstructed online by the HLT within a cone with a variable size which depends on the p_{T} of the muon. The mu50 trigger is seeded by MU20 at L1 and is required to satisfy a p_{T} threshold of 50 GeV. The efficiency is measured using a tag-and-probe method with $Z \to \mu\mu$ candidates, with no background subtraction applied, in 13 TeV data from 2016 with 25 ns LHC bunch spacing. Only statistical data uncertainties are shown. Figure is taken from [96].

the factor quantifying the LHC machine performance. The total number of observed events is defined as:

$$\overbrace{N_{\text{events}}}^{observed} = \overbrace{\mathcal{L}}^{machine} \times \overbrace{\sigma_{\text{process}}}^{nature} \tag{3.2}$$

where $\mathcal{L}$ is the total integrated luminosity and σ is the predicted cross section for a given process. The total luminosity is found by integrating the instantaneous luminosity, L, over a certain time interval, $\Delta t = t_1 - t_0$, $\mathcal{L} = \int_{t_0}^{t_1} L dt$. The instantaneous luminosity[2] is defined as the measured luminosity per event collision and is given by:

$$L = \frac{N_b n_1 n_2 f_r}{4\pi\sigma_x\sigma_y} F \tag{3.3}$$

where N_b is the number of bunches in a beam, n_i is the number of protons in a bunch, f_r is the frequency of revolution, $\sigma_{x,y}$ are the horizontal and vertical beam profiles and F is a reduction factor due to the crossing angle between the two beams. At the LHC, the instantaneous luminosity had reached orders of $L \sim 10^{+34}\text{cm}^{-2}\text{s}^{-1}$ during the 2016 and 2017 data run periods.

Most of the parameters in the luminosity are accurately measured from various beam monitors. These include the revolution rate and beam intensity parameters. The measurement of the beam profiles is achieved through van der Meer (vdM) scans [98]. The vdM scans consist of moving the two *pp* beams across each other step-wise in the transverse plane. The rate of collisions is measured along the different points which provides an effective determination of the beam height. The reduction factor and beam profiles are measured in each detector experiment using sub-detectors, known as luminometers. In ATLAS [99], the luminosity is measured from multiple luminometers using multiple algorithms. This is motivated by the varying degrees of sensitivity to different experimental effects, e.g. PU, beam-induced backgrounds, acceptance etc.

3.7.2 ATLAS Data

The data presented in this thesis was collected by the ATLAS detector between 2015 and 2016 data run periods at the LHC at a center of mass energy, $\sqrt{s}$, of 13 TeV with a 25 ns bunch spacing. The total delivered integrated luminosity at ATLAS was 42.7 fb^{-1}, however due to losses during reconstruction and periods where certain components of the ATLAS detector were switched off, the total luminosity available for physics analyses is 36.1 fb^{-1}. In 2017, the LHC exceeded its expected goal of 45 fb^{-1} and delivered a total luminosity of $\sim$ 50 fb^{-1}. The luminosity for all LHC run periods is shown in Figure 3.12. It indicates the changing of the beam parameters, which result in a higher luminosity.

[2]This is a simplified formula which requires two beams equal in size which are Gaussian distributed.

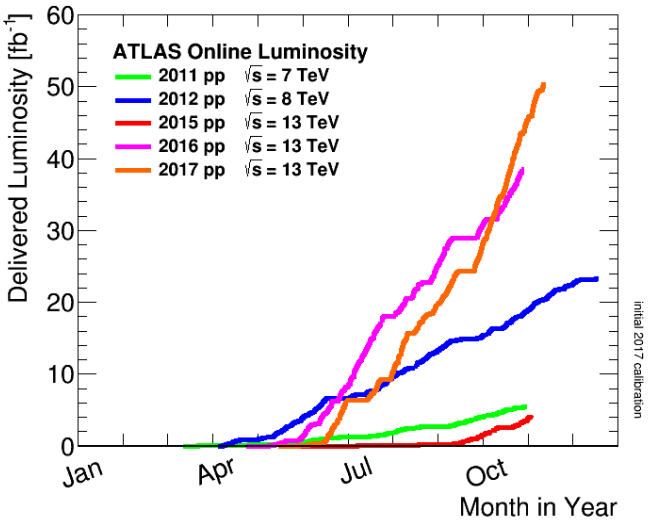


Figure 3.12: Integrated luminosity as a function of time for different periods collected by the ATLAS detector. Figure is taken from [100].

However, there is a delicate trade-off between the increase in luminosity and the increase in "pile-up" (PU). In pp collisions, a hard scatter vertex signifies an event of interest, but due to the QCD nature of these interactions, additional vertices are formed from multiple soft QCD interactions. These soft vertices, in addition to the hard scatter vertex, form the PU events. The PU is classified into two categories, based on its origins. The first category is "in-time" PU, which is caused by multiple soft collisions within the same event bunch crossing. The second is "out-of-time" PU which is caused by collisions from different bunch crossings and longer detector read-out times. The in-time PU is examined through the number of reconstructed primary events, N_{PV} [101], while out-of-time PU can be quantified by the average number of interactions, $\langle\mu\rangle$ [102]. The PU profile is shown in Figure 3.13 for the 2015 and 2016 periods. As can be observed, the subsequent change in beam conditions to increase the luminosity has also affected the amount of PU resulting in a total average μ of ~ 24.

3.8 ATLAS event data model and simulation

For run II, the ATLAS collaboration redesigned its event data model (EDM) for analyses to reduce the complexity of the event data structures [103]. The common data format adopted is xAODs which provide the information on the complete event-level reconstruction. For most analyses, a small subset of the event information is kept, typically $\mathcal{O} \sim (1-2\%)$ of the original xAOD size in derivations (DxAODs), which are then slimmed to contain only the relevant analysis variables. Both data and simulated samples follow the xAOD file format.

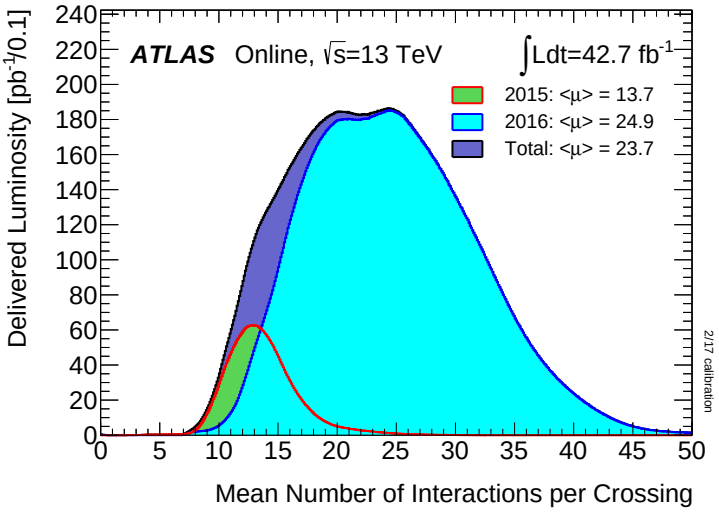


Figure 3.13: Pile-up profile in run II, measured with the ATLAS detector. Figure taken from [100].

The simulation of interactions is achieved through monte-carlo (MC) generators which calculate the matrix elements for scattering via certain processes to a given precision. The parton kinematics are described by parton distribution functions (PDFs) while parton showers (PS), hadronisation and underlying events are provided by showering algorithms.

The MC samples are linked with the virtual reconstruction of ATLAS through the simulation infrastructure [104]. The `GEANT4` framework [105] is used to simulate the detector response of the particles passing through ATLAS. PU interactions are simulated from inelastic pp collisions and overlaid on the event simulation.

The calculation of scattering matrix amplitudes and generation of the hard scatter process can proceed via different general purpose MC generators [106], such as `Powheg` [107] and `Sherpa` [108]. The MC generators are interfaced with the `Pythia8` [109] event generators, which are responsible for the simulation of PS, underlying events and hadronisation. All these processes are computed with a dedicated set of ATLAS tuned parameters such as A2 [110], which depend on the specific samples considered. Various sets of PDF are considered and used for different samples [111]. The mean number of interactions, $\langle\mu\rangle$, is re-weighted to the observed distribution, to correct the mis-modelling between data and simulation.

Chapter 4

Physics object reconstruction

The Higgs particle is short-lived and decays into particles which carry the energy and momentum of the Higgs resonance. In this thesis, the focus is on the Higgs boson decay to two Z vector bosons ($H \to ZZ$). The Z bosons also quickly decay to lighter particles, some of which can be reconstructed by the ATLAS detector.

This particular final state consits of two Z bosons, one of which decays to opposite sign same-flavour leptons (e^+e^- or $\mu^+\mu^-$), and the other whose decay product are neutrinos ($\nu\bar{\nu}$). The neutrino flavour can consist of any of the three lepton types (ν_e, ν_μ, ν_τ). The search is performed in an inclusive jet category which means that events with jets are also accepted into the final state. Photons and $\tau-$leptons are not selected for this final state but are reconstructed as jets.

The charged lepton pairs are reconstructed using track and calorimeter information while jets are constructed solely from calorimeter clusters. Suppression of PU is provided by the accurate vertex resolution of the ATLAS ID. Neutrinos don't leave any visible signatures in the detectors and thus can be only quantified by the amount of missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$).

Since the charged objects use tracking information, the basics behind track and vertex reconstruction are given in section 4.1. The details behind the calibration and quality selections are provided in section 4.2.1 for electrons, section 4.2.2 for muons and section 4.2.3 for jets. Removal of overlap between the various objects is covered in section 4.2.4. Finally the details behind the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction and performance are discussed in section 4.3.

4.1 Tracks

The tracks of charged particles are found through recorded hits in the ID [112] by using an "inside-out" tracking algorithm [113]. The overall coverage provided by all components of the ID extends to $|\eta| < 2.5$. The tracks are valuable as inputs for the physics objects ($e^{\pm}, \mu^{\pm}$ etc.) reconstruction, combined with information from the calorimeters or the MS. "Loose" track quality selections require a minimum track p_{T} of 400 MeV and a minimum number of 7 recorded hits in the silicon detector. Further tracking requirements are applied on the number of shared modules and number of holes[1].

The point at which an interaction between particles occurs is referred to as the interaction vertex. The "primary" vertex is the point where the highest energy interaction happens. It is located by an algorithm designed to find the vertex with largest sum of momenta ($\sum p_T^2$) of tracks. A single hard "primary" collision is defined per event collision, while the other "softer" interactions come from PU. An important distinction from PU are secondary vertices which can be found at some distance from the origin of the primary interaction point. They may originate from a few sources such as long-lived $B-$hadrons decays or photon conversions.

Every track position, with respect to the axis along the beam-line (z), is determined by two impact parameters. The transverse impact parameter, d_0, quantifies the distance of closest approach on the transverse plane to the beam-line, while the longitudinal parameter, z_0, is the distance of closest approach along the beam-line [114]. The resolution of the impact parameters is sensitive to the track reconstruction, which is provided by the individual measurements in the silicon detector. This is inherently sensitive to the alignment of the ID. Mis-alignments of the ID are thoroughly investigated to minimize any potential measurement biases on the resolution [115, 116].

4.2 Physics objects

Track and calorimeter objects which have passed a particular set of reconstruction and identification selections are referred to as "physics" objects. They are classified by their particle type i.e. $e^{\pm}, \mu^{\pm}$ or as a shower of hadrons known as "jets". They are used in high energy physics analyses to reconstruct events.

[1] A shared module is either a hit in the pixel detector, used by more than one track, or two shared hits in the same SCT layer. A hole is intersection of reconstructed track trajectory with a sensitive detector element which does not result in a hit.

4.2.1 Electrons

Electron reconstruction is performed in three steps [117]. The first step is a sliding window algorithm which searches for an electron "seed" where size of the window is 3×5 in units of 0.025×0.025 in $\Delta\phi \times \Delta\eta$. The algorithm constructs clusters around the seed, which contain the shower of secondary particles. The efficiency of sliding window method is $\geq 99\%$ for E_{T} greater than 15 GeV. The second step is track reconstruction, which is performed using pattern recongnition and global track fits. The final step is to match the track candidate to the seed cluster. If more than one track is in very close proximity to the seed cluster, the closest matching track is found by using a track-cluster distance metric. The energy of the electron is provided by the fully calibrated cluster [118], while the η, ϕ description is provided by the matched track.

In the analysis, electrons are accepted with a minimum E_{T} greater than 20 GeV with $|\eta| < 2.47$, including electrons reconstructed in the crack region. The pseudorapidity selection is applied on the electron cluster.

The identification of electrons from the primary interaction relies on sequential and multivariate (MVA) algorithms in order to reject electrons from photon conversions, hadronic jets or semileptonic hadron decays. The "signal" electrons are obtained from a set of likelihood-based identification criteria (LHid). The parameters used in the likelihood concern the longitudinal and lateral calorimeter shower profiles, the track properties and the track-to-cluster matching. The "medium" LHid criteria are chosen since they operate at a 90% efficiency for electrons with $E_{\mathrm{T}} \approx 40$ GeV and improve at higher energy.

To further reject electrons from secondary sources, isolation requirements are employed. A calorimeter isolation parameter is based on a fixed cone radius, dR, of 0.2 around an electron cluster. Corrections are added for shower leakage and PU noise. A track isolation parameter in the ID is obtained by a variable cone size around an electron track. The varible cone size is achieved thanks to the small cell size of the tracker. The cone radius varies as a function of the p_{T}, $dR = \min(0.2, 10\ \mathrm{GeV}/p_{\mathrm{T}})$. The "Loose" track and calorimeter isolation operating point provides a flat 99% signal electron efficiency [117]. Triggering on electron objects is based on a likelihood discriminant optimized with reconstructed shower shapes and track variables during the data taking process. The discriminant is adjusted to correct for PU effects.

A tag-and-probe method in $Z \to ee$ (and $J/\psi \to ee$) events estimates the overall efficiencies of the reconstruction, identification, isolation and trigger methods. In the tag-and-probe method, one of the electrons, called tag, must satisfy certain strict objects requirements and the probe electron can satisfy only certain loose pre-id criteria. The efficiency factors are

measured from the probe electron. The reconstruction efficiency is obtained from the ratio of the number of "good" track quality (7 hits in the ID) reconstructed electrons to the total number of reconstructed electrons ECAL. The id efficiency is found from the ratio of the identified (passing the LHid) electrons to the total number of reconstructed electrons. The trigger efficiency is determined from the ratio of the triggered electrons over all identified electrons. The isolation efficiency is established by the ratio of the number of electrons passing the identification and isolation requirement to the ones only passing the identification requirement. The trigger efficiency, for a given trigger, is measured with respect to the offline id algorithm and the loose isolation working point.

Cuts on the impact parameter are implemented to associate the electrons to the hard-scatter vertex and to remove contributions from photon conversions and secondary vertices. The transverse impact parameter significance, d_0/σ_{d_0}, is required to be less than 5. The selection on the longitudinal impact parameter is translated into $\Delta z_0 \times \sin(\theta) < 0.5$ mm[2], where Δz_0 is the longitudinal parameter z_0 calculated w.r.t. the primary vertex and θ is the polar angle of the track [114, 119].

4.2.2 Muons

Muons are reconstructed through global track fits from the ID and MS [120]. First, the tracks are independently reconstructed from recorded hits in the tracker and spectrometer. Then the track from the MS is extrapolated and matched with an ID track, which forms a combined track. The reconstruction efficiency is found using a tag-and-probe method in $Z \to \mu\mu$ events for $|\eta|$ less than 2.5. The efficiency is found to be $\sim 99\%$ for p_{T} greater than 5 GeV [120].

To identify muons from the hard scatter interaction and reject ones from secondary processes such as pion and kaon decay, a selection is placed on the significance of q/p:

$$q/p \text{ significance} = \frac{|(q/p)_{\mathrm{ID}} - (q/p)_{\mathrm{MS}}|}{\sqrt{\sigma(q/p)_{\mathrm{ID}}{}^2 + \sigma(q/p)_{\mathrm{MS}}{}^2}} \tag{4.1}$$

where q/p is the muon charge divided by the momentum, measured either in the ID or MS and $\sigma(q/p)$ is the error associated with q/p. It provides a good separation between genuine and fake muons and a stable selection efficiency across wide η and p_{T} ranges. Selections are also made on the number of hits in the ID and MS to ensure a good momentum measurement. Combined track algoritms may add or remove MS hits in the global track fit to obtain the best possible resolution.

[2]The $\Delta z_0 \times \sin(\theta)$ instead of Δz_0 is used to avoid rejecting tracks with an expected larger in the forward detector region.

The identification category of muons is called "medium", where muons must satisfy specific hit requirements in the ID and MS and must have a q/p significance of less than seven [120]. The medium category minimizes the systematics linked to the reconstruction and calibration of muons. The majority (99.5%) of muons within this category consist of combined muons while the remaining can be reconstructed in the wider acceptance region, $2.5 < |\eta| < 2.7$. In the $H \to ZZ \to \ell\ell\nu\nu$ analysis the wider acceptance region is not used since the muon pseudorapidity requirement is less than 2.5.

Isolation selections on muons further suppress fakes from semileptonic heavy flavour quark decays. For a calorimeter-based isolation, a fixed cone of $dR = 0.2$ is used around the muon, after subtracting the muon energy deposit and correcting for PU. The ratio of the transverse energy of reconstructed muon calorimeter clusters to the muon momentum is required to be less than 0.2 for the muon to be isolated. For the track-based isolation, the scalar sum of the transverse momenta of all the tracks with $p_{\mathrm{T}} > 1$ GeV in cone of size $dR = \min(10\ \mathrm{GeV}/p_{\mathrm{T}}, 0.3)$ is used. Here, for the muon to be considered isolated, the ratio of the scalar p_{T} track sum to the total muon p_{T} must be less that 0.3. The "Loose" isolation criteria are found to have a constant 99% selection efficiency across η and p_{T} ranges [120].

An additional recontruction efficiency, called track-to-vertex-association (TTVA), is applied due to modelling of the impact parameters d_0 and z_0 between data and simulation [121].

During the 2016 data taking period, a new effect emerged in the muon reconstruction, the so-called "sagitta bias" which originated from local misalignments of the ID [122]. The bias on the track sagitta causes a charge-dependent bias of the reconstruction of the muon momentum. A parametrized (η, ϕ) correction is applied to every muon, which reduces the sagitta bias effect to per-mille level at the Z boson mass scale and it improves the $m_{\mu\mu}$ resolution by 1-5%. Systematic uncertainties are estimated from simulation.

The analysis requires that muons have transverse momentum greater than 20 GeV and pseudorapidity within the confines of the ID tracker ($|\eta| < 2.5$). Impact parameter cuts are applied, similarly to the electron case, with $\Delta z_0 \sin(\theta) < 0.5$ mm, but with a transverse impact parameter significance less than 3 [114, 119].

4.2.3 Jets

Quarks and gluons are bound in hadrons and mesons and can never be observed isolated. At the LHC, when quarks and gluons emerge from the interaction point they quickly undergo *hadronisation*, which forms a collimated beam of particles, mostly consisting of hadrons.

The jets are constructed using the anti-k_t sequential recombination algorithm [123] with a

radius parameter, R, of 0.4. The anti-k_t is implemented via `FastJet` [124]. The inputs come from topological cell clusters, which are built from neighbouring calorimeter cells, calibrated at the electromagnetic (EM) energy scale[3], with energy significantly above the noise levels [125].

The anti-k_t algorithm is soft-resilient, infrared and co-linear safe. Soft-resilient means that particles with low transverse momentum (i.e. "soft") do not modify the shape of the jet. It is infrared and co-linear safe because soft perturbative objects are iteratively clustered back to the particles with high transverse momentum (i.e. "hard"). The anti-k_t algorithm is widely adopted by ATLAS for different studies which involve jets. The jet scheme relies on the distance parameter between constitutents objects in the jet to be:

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right)\left(\frac{\Delta_{ij}}{R}\right)^2 \tag{4.2}$$

where $\Delta_{ij}^2 = (\Delta y_{ij})^2 + (\Delta\phi_{ij})^2$ is the sum of squares of the difference in rapidities (Δy) and difference in azimuthal angle ($\Delta\phi$) and $k_{t,i}$ is the transverse momentum for the i constituent. In addition, for each constituent, a beam distance is also defined:

$$d_{iB} = 1/k_{T,i}^2. \tag{4.3}$$

A compiled set of distances is formed from d_{ij} and d_{iB}. Afterwards the algorithm finds the minimum of the d_{ij} and d_{iB} values. If d_{ij} is smaller, then i and j are clustered together by combining their four-momentum and all distances are recompiled. If instead the d_{iB} is smaller than d_{ij}, then the constituent is called a jet. The algorithm repeats these steps on the remaining input objects until no more constituents are left.

The shape of the jet follows from the definition of d_{ij}. Soft objects will have a larger separation between themself than the separation between hard and soft ones. The reason is due to the $1/p_{\mathrm{T}}^2$ factor present in the distance metric. This results in soft particles tending to cluster around a single hard object. If no other hard objects are present within $2R$, the resulting jet will have a conical shape with radius R. If other hard objects are present, then the object with the largest p_{T} will have a conical shape and the other will be clipped. For general cases of comparable momenta, both can be clipped. The effect of the anti-k_t algorithm on the shape of jets is seen in Figure 4.1. The four-vectors of all jet constituents are combined to provide the jet four-momenta properties such as energy, momentum and mass.

After the jet reconstruction step, the jet energy scale (JES) is calibrated according to the EM+JES scheme [126]. It consists of five correction/calibration stages:

[3]The EM scale corresponds to the correct reconstructed energy of electrons and photons but does not have any corrections for the loss of signal for hadrons due to the non-compensating nature of the calorimeters.

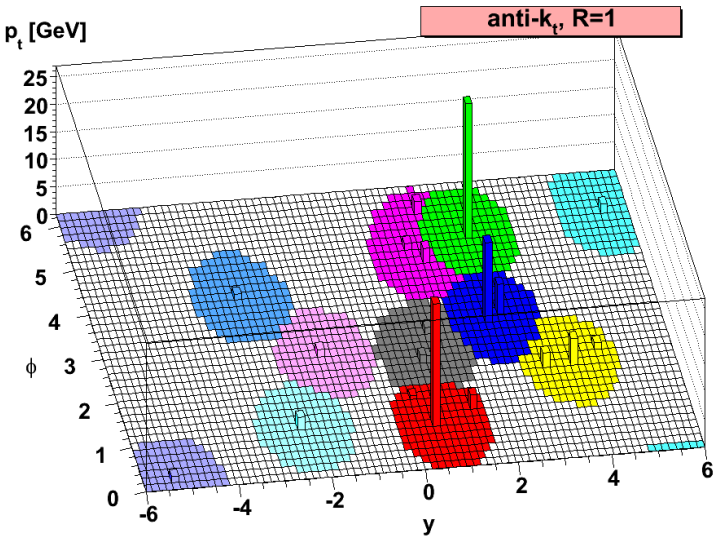


Figure 4.1: Schematic of the anti-k_t algorithm for a radius parameter $R = 1.0$. A consequence of the algorithm is the formation of conical jet boundaries. Figure is taken from [123].

- Origin correction which recalculates the jet four-momentum from the hard-scatter vertex and not the detector center.
- An area-based PU subtraction and residual corrections obtained from simulation.
- Absolute MC-based jet energy scale calibration to correct the four-momentum of jets to its particle-level energy scale.
- Global sequential calibration which improves jet energy reconstruction.
- Residual *in situ* correction, which is applied to data derived using only well-measured reference objects (e.g. electrons and photons).

There are 80 systematic uncertainties on the JES, which are related to the various assumptions during the calibration steps, such as the PU modelling, multi-jet modelling and jet composition. Uncertainties also come from assumptions in the *in situ* calibration step, which are related with simulation, sample statistics, event topologies and object uncertainties. The JES uncertainties in the analysis are introduced through a strong reduction scheme, which reduces the 80 systematics to a more manageable set of five [127].

After the reconstruction and calibration steps, the energy resolution of jets is measured. This is accomplished by two different in situ methods as described in [128]. The overall agreement with data is at the level of 10 %.

Further removal of PU jets is crucial in reducing the size of the systematic uncertainties. In run I, a jet-vertex-fraction (JVF) requirement was employed to reject PU jets, however

it was found to be dependent on the number of recontructed vertices (N_{PV}) [129]. In run II, a jet-vertex-tagger (JVT) MVA technique is implemented which provides an improved rejection of PU jets without the dependency on the event activity. The JVT is constructed as a two-dimensional likelihood, based on a $k-$nearest neighbor MVA, from a N_{PV} corrected JVF and a variable proportional to the ratio of the scalar sum of tracks, $\sum p_{\mathrm{T}}$, to the total calibrated jet p_{T}. The JVT tagging algorithm is only applied to jets with p_{T} below 60 GeV and $|\eta|$ less than 2.4. The output JVT value has a range between 0 (for soft PU jet) and 1 (for hard-scatter jet). The default requirement selects jets with JVT greater than 0.59, which provides a PU rejection efficiency of 92% and an observed fake rate of 2%. JVT uses track information, which requires the matching of tracks with calorimeter jets, which is performed by ghost track association, where the ID tracks are clustered into a jet as *ghost particles* [130]. A ghost particle is a soft particle with infinitely small momentum. The JVT efficiency corrections factors are p_{T} dependent for jets between 20 and 60 GeV. They vary between 3.7% for the lowest jet p_{T} bin and 0.9% for the highest jet p_{T} bin. The efficiencies are estimated using a tag-and-probe method, where a Z boson is tagged and the leading jet recoiling against the Z is the probe [129].

Additional jet quality criteria are imposed to remove jets from non-collision processes such as beam induced background, cosmic-ray showers and random calorimeter noise bursts [131]. More specifically, the selections are applied to calorimeter variables to suppress detector noise while for non-collision backgrounds, requirements are placed on energy ratio variables, which measure the energy in a certain detector component. The "Loose" jet category provides a selection efficiency $\geq 99\%$ for p_{T} greater than 20 GeV [132].

Jets are also categorized based on their origin. "Light" flavour jets originate from u and d quarks as opposed to "Heavy" flavour jets which come from c and b quarks. In particular, the $b-$quarks can form long-lived $B-$hadrons which travel some distance before decaying into jets. The resolution of the pixel detector allows to accurately reconstruct the decay from the secondary vertex, thereby providing an effective $b-$jets identification method, known as $b-$jet tagging [133]. Other $b-$tagging algorithms are also employed, which use cuts on the impact parameters or which further exploit the topological structure of $B-$hadron decays. A boosted decision tree MVA technique combines all the $b-$tagging algorithms to obtain a $b-$jet discriminant. Operating points are defined by a single cut value on the discriminant output distribution and are chosen to provide a specific b-jet efficiency on an inclusive $t\bar{t}$ sample. Correction factors are applied to the simulated event samples to compensate for differences between data and simulation in the b-tagging efficiency for b, c and light-jets [134–136]. The factors for b-jets are derived from $t\bar{t}$ events containing two leptons. The corrections are consistent with unity and their associated uncertainties are at the level of a few percent over most of the jet p_{T} spectrum [137].

The analysis setup selects jets with p_{T} greater than 20 GeV, $|\eta| < 4.5$ and JVT greater than 0.59 for jets with transverse momentum below 60 GeV.

4.2.4 Overlap removal

The different objects can be found to overlap between each other e.g. an electron can be reconstructed by both the electron and jet clustering algorithms. The removal of object overlap seeks to minimize any cases where one physical entity can be reconstructed as two separate objects, e.g. an electron can be reconstructed by both the electron and jet clustering algorithms, or where two objects are close-by and difficult to distinguish. It is an issue which affects all analyses.

Table 4.1 summaries the overlap removal strategy adopted by the $H \to ZZ \to \ell\ell\nu\nu$ analysis which follows the recommendations from the ATLAS harmonisation scheme [138]. The measure of distance between objects used in overlap removal is $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2}$ where $\Delta y(\Delta\phi)$ is the rapidity (azimuthal) difference between the two objects. Overlap removal is handled differently in $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, which is discussed in section 4.3.

	Reference objects	Criteria
Remove jets	electrons	$\Delta R_{e-\mathrm{jet}} < 0.2$
	muons	$\Delta R_{\mu-\mathrm{jet}} < 0.2$ if $N_{\mathrm{Trk}}(\mathrm{jet}) < 3$ **OR** $(p_{\mathrm{T}}^{jet}/p_{\mathrm{T}}^{\mu} < 2$ **AND** $p_{\mathrm{T}}^{\mu}/\Sigma_{\mathrm{Trk}} p_{\mathrm{T}} > 0.7)$
Remove electrons	jets	$0.2 < \Delta R_{e-\mathrm{jet}} < 0.4$ **AND** PU jets check
	muons	share the same ID track
Remove muons	jets	$\Delta R_{\mu-\mathrm{jet}} < 0.4$ **AND** PU jets check

Table 4.1: Overlap removal criteria adopted following the ATLAS harmonisation scheme. PU jets check means "JVT> 0.59 if jet has $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$".

4.3 Missing Transverse Momentum ($\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$)

In its most basic form, the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is nothing more than the imbalance of momentum in the transverse plane of the detector due to invisible particles. As such, it has a direction and magnitude, which depend on the reconstruction performance of all other objects in an event. The ATLAS detector is specifically designed to have a ϕ coverage of 4π in order to reconstruct

the missing momentum. The $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is a global detector performance variable as it uses most or all of the available detector parts in its reconstructions. Therefore the studies of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ performance also translate into an overall performance study of the detector itself. More importantly, the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is also a key variable in the search of new physics in the high energy sector [139].

4.3.1 $\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$ reconstruction

The $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is obtained from all reconstructed particles in the detector. Two general terms form the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ definition:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{hard}) + \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{soft}) \tag{4.4}$$

where the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$(hard) is formed from all detector measured and calibrated physics objects while the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ (soft) term is formed from unassociated tracks. The $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from the transverse components of all objects found in the event and the full definition can be written as:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = \overbrace{\underbrace{-\sum_{\text{selected electrons}} \mathbf{p}_{\mathrm{T}}^{e}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},e}} \underbrace{-\sum_{\text{accepted photons}} \mathbf{p}_{\mathrm{T}}^{\gamma}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\gamma}} \underbrace{-\sum_{\text{accepted } \tau\text{-leptons}} \mathbf{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\tau_{\mathrm{had}}}} \underbrace{-\sum_{\text{selected muons}} \mathbf{p}_{\mathrm{T}}^{\mu}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss},\mu}} \underbrace{-\sum_{\text{accepted jets}} \mathbf{p}_{\mathrm{T}}^{\mathrm{jets}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss,jets}}}}^{\text{hard term}} \overbrace{\underbrace{-\sum_{\text{unused tracks}} \mathbf{p}_{\mathrm{T}}^{\mathrm{track}}}_{\mathbf{E}_{\mathrm{T}}^{\mathrm{miss,soft}}}}^{\text{soft term}} \tag{4.5}$$

where each $E_{\mathrm{T}}^{\mathrm{miss}}$ term is calculated as the negative vector sum of all the measured object momenta.

The length of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector gives the magnitude of the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$:

$$E_{\mathrm{T}}^{\mathrm{miss}} = |\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \tag{4.6}$$

where E_x^{miss} and E_y^{miss} are the x, y components of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$. The direction of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector is given in terms of the azimuthal angle:

$$\phi^{\mathrm{miss}} = \tan^{-1}(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}}). \tag{4.7}$$

The hard term, as mentioned above, are the identifiable entities in the detector setup, which the event reconstruction process associates to physics objects. The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction associates entities to objects in the order: electrons ($e^{\pm}$), photons (γ), hadronically decaying tau leptons (τ), jets and lastly muons ($\mu^{\pm}$). Any tracks, unused in the hard term reconstruction are added to the soft term[4].

[4]This can happen if a track fails a certain p_{T} or identification requirement.

4.3.1.1 Hard $\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$ term

Electrons are constructed from cluster deposits in the EM calorimeters and matched with ID tracks. In the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, electrons with "medium" quality criteria are used, with a p_{T} requirement greater than 10 GeV. The electrons in the transition region between the barrel and end-cap calorimeters ($1.37 < |\eta| < 1.52$) are excluded and instead reconstructed as jets. The acceptance goes up to the edges of the tracker at $|\eta| < 2.47$.

Muons are formed from a track matched in the ID with a segment or track from the MS. Muons, which enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, are required to have "medium" quality criteria with a p_{T} cut greater than 20 GeV extending to $|\eta| < 2.5$.

Photons and hadronically decaying taus are not explicitly used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation but instead enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation as jets. They have a small net contribution to the overall $E_{\mathrm{T}}^{\mathrm{miss}}$.

Jets are formed from topological clusters of calorimeter cells using the anti-k_t algorithm with a radius parameter of 0.4. They are calibrated using the EM+JES scheme [126]. Jets within the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction follow the same kinematic selections and JVT quality criteria as outlined in section 4.2.3.

Jets are rejected, to avoid double-counting, if they are found to overlap with other physics objects. Additionally a partial overlap removal treatment may also be performed in order to better suppress PU jets. In these cases a PU jet can survive JVT requirement if its track is found close to an electron or muon track [5].

The overlap removal between electrons and jets is based on the ratio of the electron energy to the jet energy $k_E = E_e^{\mathrm{EM}}/E_{\mathrm{jet}}^{\mathrm{EM}}$. For $k_E \leq 50\%$ jets are included in the jet term, with a scaling of $1 - k_E$ applied to their energy. For $k_E > 50\%$, the jet is rejected, while the tracks associated with jet are added to the soft term.

The situation for muons is more complex as it has more sources of overlapping with jets. A muon can overlap with a PU jet, which can lead to JVT mis-tagging, and result in PU jets that enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. This can happen if the p_{T} of the overlapping muon ID track is large enough to increase the JVT value beyond the cut value. An additional effect is due to a stochastic process connected with large calorimeter muon energy losses. The deposited muon energy is reconstructed as a hard jet pointing to the primary vertex. This results in double counting of the muon p_{T} in $E_{\mathrm{T}}^{\mathrm{miss}}$, since the calibrated muon has already been corrected for this effect. Additional muon-jet overlaps can develop due to muons emitting high energy photons at small angles with respect to the muons direction. These photons are then reconstructed as jets since the muon track violates photon isolation. The muon

reconstruction does not recover the momentum of the emitted photon, to the jet is added to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

Rejecting muon-jet overlaps requires a set of selections, which are described in [4]. Briefly, to reject PU jet contributions, the ID track p_{T} of the muons must be at least 80% of the sum of transverse momenta of all ID tracks connected with the jet, the overall transverse momentum of the jet must not be too large in comparison to the muon track momentum and the total number of tracks associated to the jet from the primary vertex is less than five. To identify photons from emissions, a set of selections is imposed on the jets found close to muon ID tracks. The majority of the deposited jet energy must be in the ECAL and the shower profile must be narrow, indicating a dense EM shower. Further selections are placed on the amount of tracks from the primary vertex and fraction of transverse momentum carried by jet tracks to the one from the muon.

The only difference in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction outline above and the one used in the $H \to ZZ \to \ell\ell\nu\nu$ analysis is in the id and p_{T} requirement of the lepton objects. In the $H \to ZZ \to \ell\ell\nu\nu$ analysis the electrons and muons are "loose" with a p_{T} requirement of 7 GeV. This is done in order to not significantly bias the WZ background due to the presence of a third "loose" quality lepton.

4.3.1.2 Track-based soft term (TST)

In run I of the LHC, the soft term was built from calorimeter deposits but in the rich PU environment of run II, the large residual PU contamination was shown to severely degrade the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution [139]. As a result, in run II, the soft term is built exclusively from tracks, which offers an improved PU rejection and a better $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution.

The TST is reconstructed from unidentified tracks which have no applied calibration. The inclusion of the TST is very important as it aides in capturing the hadronic recoil from low energy interactions.

Tracks which contribute to the soft term must have a minimum p_{T} greater than 400 MeV, $|\eta|$ less than 2.5 and impact parameter selections of $|d_0| < 1.5$ mm, and $|z_0 \sin(\theta)| < 1.5$ mm. Partial overlap removal is applied to the soft term, which includes removal of tracks found with $\Delta R(\text{track,electron}) < 0.05$ or tracks associated with muons and tracks ghost-associated with jets.

Due to the nature of the reconstruction method of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ (soft), two important caveats must be noted: first, it is missing the soft neutral particles since they do not leave any tracks in the ID and second, only charged particles within the region covered by the ID ($|\eta| < 2.5$) are reconstructed. These two effects have an impact on the overall $E_{\mathrm{T}}^{\mathrm{miss}}$ performance.

The performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is evaluated in events which contain opposite sign, same flavour lepton pair, with invariant mass $m_{\ell\ell}$ within 15 GeV of the Z mass (= 91.2 GeV).

The process considered as "signal" in this final state is the $Z(\rightarrow \ell\ell)$+jets. The genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ for Z + jets events is expected to be negligible and any significant deviations are driven by jet mis-reconstruction or PU noise. The Z + jets signal is simulated using the `Sherpa` generator, which calculates the matrix element to NNLO in the strong coupling constant, α_s. The parton-level matrix element is interfaced with the `Pythia8` PS algorithm, which creates the PS and underlying event. In addition to Z + jets, other processes can have a di-lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ signature, such as top ($t\bar{t}$) and diboson (WZ, ZZ, WW). These events are considered as backgrounds to the Z + jets. The backgrounds are simulated with via `Powheg` + `Pythia8` generator. The signal and backgrounds are normalized to an integrated luminosity of 36 fb^{-1}.

In addition to the Z mass window requirement, the di-lepton event selection also requires each lepton to have a minimum p_{T} greater than 20 GeV and pass the "medium" id quality criteria. The leading lepton is required to have a p_{T} higher than 30 GeV, as explained in section 5.2.3. Both leptons are required to pass the minimum single lepton trigger requirements outlined in section 3.6.

4.3.2 $\mathbf{E_T^{miss}}$ performance evaluation in the di-lepton final state

4.3.2.1 Modelling between data and simulation

The reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in the di-lepton final state is given in Figure 4.2. The Z + jets signal dominates the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region until $\sim$ 100 GeV, where the top background becomes the dominant contribution. Diboson events contribute more significantly at the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Discrepancies on the order of $20 - 25$ % are found for events in the range between 50 and 150 GeV, which are covered by experimental uncertainties on the background, shown in the shaded band. The band consist of systematics on the constituents which enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, TST and PU modelling [140]. The largest contribution to the uncertainty band are the combined JES systematics [126]. The distribution of the azimuth, ϕ^{miss}, x and y components of $E_{\mathrm{T}}^{\mathrm{miss}}$ are given in Figure A.1 (A.2) for the $Z \rightarrow ee$ ($Z \rightarrow \mu\mu$) channel.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is further studied by examining the individual component terms which enter into the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term is given in Figure 4.3. Mis-modelling is present due to the uncalibrated tracks which enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. The systematic

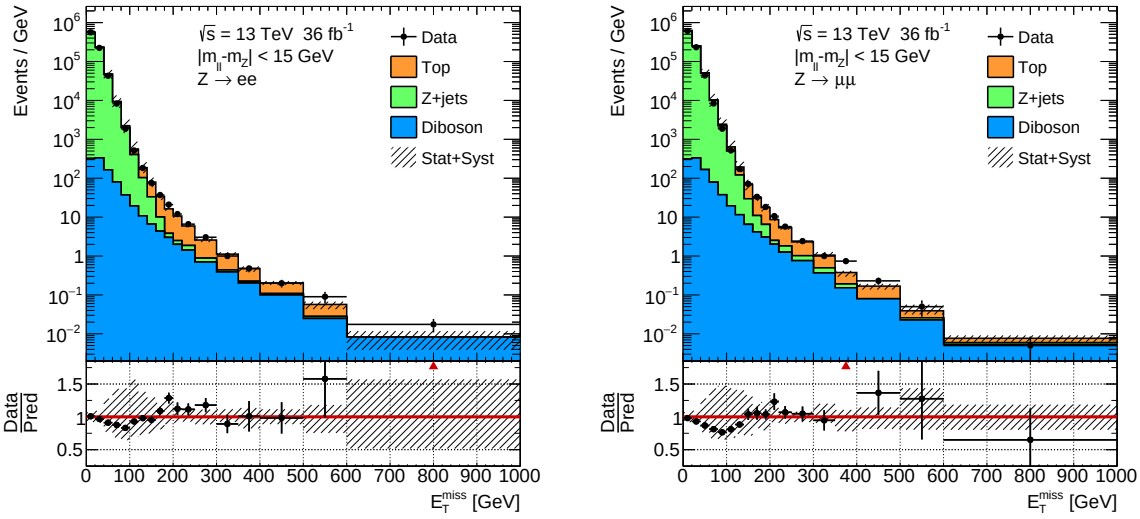


Figure 4.2: The inclusive $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the ee and $\mu\mu$ channels are shown. All simulated samples are scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

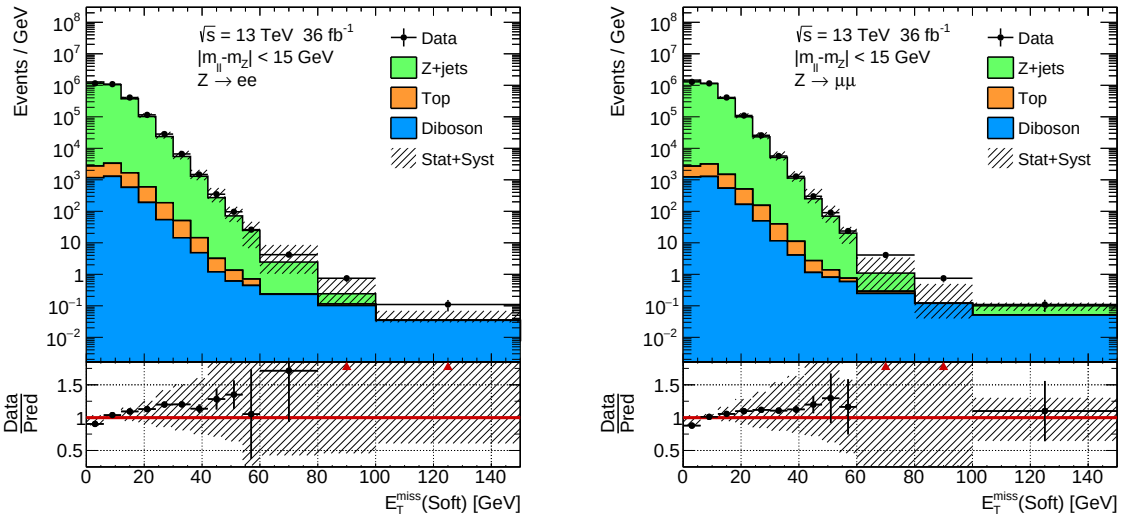


Figure 4.3: The plot on the left (right) shows the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the ee ($\mu\mu$) channel. All simulated samples are scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

uncertainty band is dominated by the scale systematic uncertainties of the TST. The hard $E_{\mathrm{T}}^{\mathrm{miss}}$ term consists of lepton and jet components. The jet term is shown in Figure 4.4 while the lepton term is given in Figure 4.5. They are shown to have an overall good agreement between data and simulation.

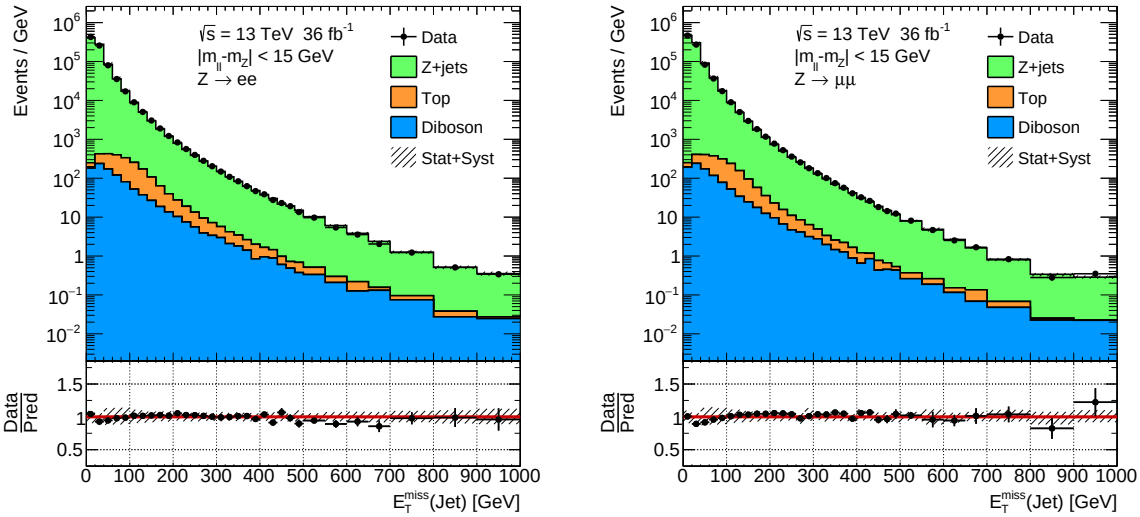


Figure 4.4: The plot on the left (right) shows the jet $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the ee ($\mu\mu$) channel. All simulated samples are scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s}$ = 13 TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

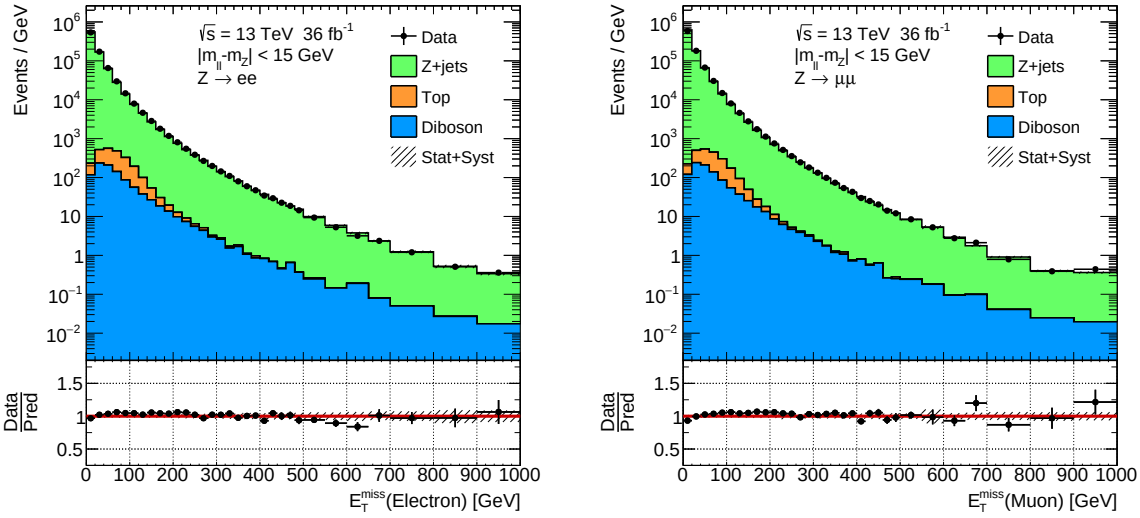


Figure 4.5: The plot on the left (right) shows the lepton $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the ee ($\mu\mu$) channel. All simulated samples are scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s}$ = 13 TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

4.3.2.2 Diagnostics

Diagnostic plots are valuable in closely studying the $E_{\mathrm{T}}^{\mathrm{miss}}$ performance both in data and simulation. The diagnostics consist of plots indicating the behaviour of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution. The $E_{\mathrm{T}}^{\mathrm{miss}}$ scale shows an overall energy scale balance between all reconstructed objects in an event. For a ideal detector the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ should perfectly balance against all other constituents. However, for a real detector, various detector inefficiencies, reconstruction

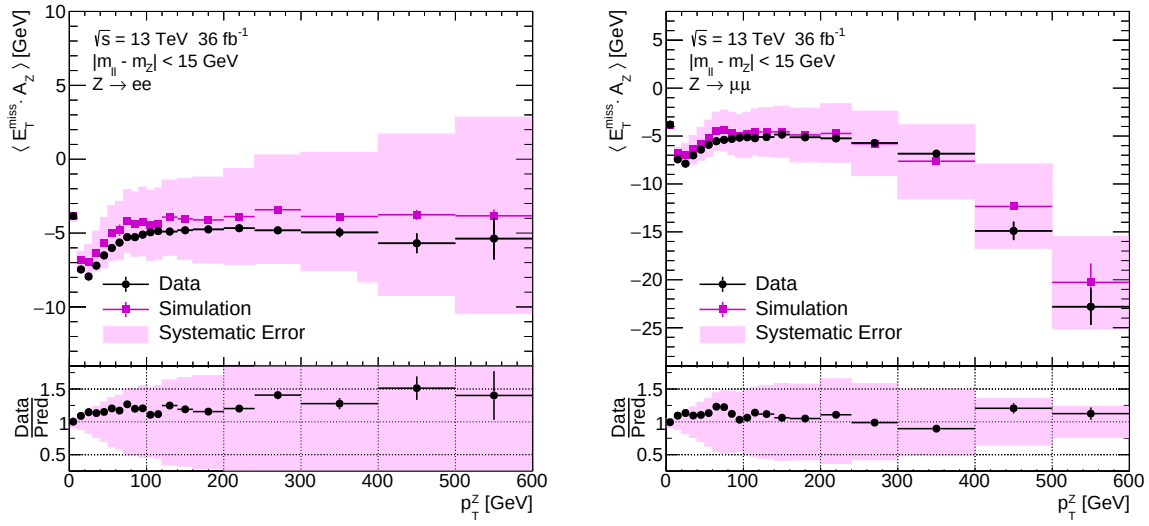


Figure 4.6: Inclusive $E_{\mathrm{T}}^{\mathrm{miss}}$ scale vs $p_{\mathrm{T}}^{\mathrm{Z}}$ for the $Z \to ee(\mu\mu)$ channel with integrated luminosity up to 36 fb^{-1} at $\sqrt{s} = 13$ TeV. Here simulation includes the signal $Z(\ell\ell)$+jets and the background processes. The shaded band is formed from the statistical and experimental systematic errors on the MC.

algorithms or missing objects give rise to imbalances which are reflected in the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution shows the overall measured resolution and is studied as a function of event parameters such as N_{PV}. Deviations from the approximate Gaussian behaviour of the resolution can appear due to events with large energy event activity.

4.3.2.2.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ recoil The recoil of the $E_{\mathrm{T}}^{\mathrm{miss}}$ against the $p_{\mathrm{T}}^{\mathrm{Z}}$ is useful in the studies of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale. In $Z \to \ell\ell$ events, the p_{T} of the Z boson is a measure for the overall energy scale of the process. The definition of $E_{\mathrm{T}}^{\mathrm{miss}}$ recoil is the projection of the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector on an axis $\mathbf{A}_Z$ defined by the transverse momentum of the Z boson i.e.:

$$\mathbf{A}_Z = \frac{\mathbf{p}_{\mathrm{T}}^{\ell^+} + \mathbf{p}_{\mathrm{T}}^{\ell^-}}{\left|\mathbf{p}_{\mathrm{T}}^{\ell^+} + \mathbf{p}_{\mathrm{T}}^{\ell^-}\right|} = \frac{\mathbf{p}_{\mathrm{T}}^{Z}}{p_{\mathrm{T}}^{Z}}\,. \tag{4.8}$$

The projection is sensitive of the hadronic recoil against the Z boson since:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \mathbf{A}_Z \simeq -\left(p_{\mathrm{T}}^{Z} + \left(\sum \mathbf{p}_{\mathrm{T}}^{\mathrm{jets}} + \sum \mathbf{p}_{\mathrm{T}}^{\mathrm{tracks}}\right) \cdot \mathbf{p}_{\mathrm{T}}^{Z}/p_{\mathrm{T}}^{Z}\right) \tag{4.9}$$

because $\sum \mathbf{p}_{\mathrm{T}}^{\mathrm{leptons}} \sim \mathbf{p}_{\mathrm{T}}^{Z}$ for events containing a Z boson. The average of the projection, $\langle \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \mathbf{A}_Z \rangle$, is indicative of the overall $E_{\mathrm{T}}^{\mathrm{miss}}$ scale. The projection can be produced in data as well since since the genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ for the $Z \to \ell\ell$ final state is very small.

In Figure 4.6, $\langle \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \mathbf{A}_Z \rangle$ is plotted against $p_{\mathrm{T}}^{\mathrm{Z}}$. The projection shows a non-linear behaviour for both data and simulation. In the low $p_{\mathrm{T}}^{\mathrm{Z}}$ region, the Z boson does not perfectly balance against the soft hadronic recoil ($\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \sim p_{\mathrm{T}}^{\mathrm{Z}} - \sum \mathbf{p}_{\mathrm{T}}^{\mathrm{tracks}}$). This happens because of uncalibrated tracks and missing contributions to the track-based soft term due to soft neutral

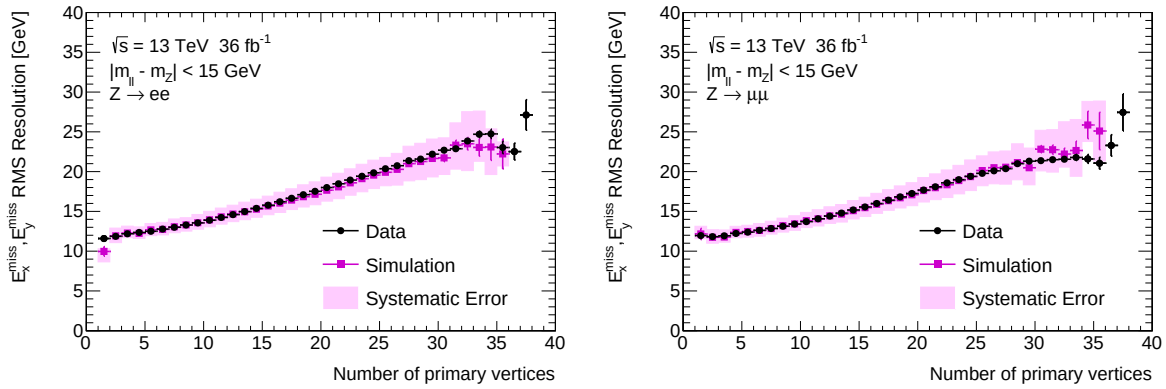


Figure 4.7: $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the number of primary vertices with integrated luminosity up to 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

particles, charged particles in the forward region and particles within $|\eta| < 2.5$ but with p_{T} less than 400 MeV which are rejected. For $p_{\mathrm{T}}^{\mathrm{Z}}$ greater than 20 GeV the $E_{\mathrm{T}}^{\mathrm{miss}}$ recoil improves since the Z boson starts to balance against the calibrated jets in the hard term. An overall negative offset at ~ -5 GeV remains due to underestimated soft neutral contribution.

The difference between simulation and data is covered by the systematic uncertainties. The differences are due to object calibration in data and simulation and different reconstruction biases. In the high $p_{\mathrm{T}}^{\mathrm{Z}}$ region, the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale in the $Z \to \mu\mu$ channel degrades due to residual muon-jet overlap present in both data and MC. Studies are on-going to suppress this residual overlap, however, at the time of writing no recommendation has been released to improve the performance in this region.

4.3.2.2.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is measured by the root-mean-square (RMS) width of the E_x^{miss} and E_y^{miss} components of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$. The RMS provides sensitivity for gaussian tails, which are lost for gaussian fits of the $E_{\mathrm{T}}^{\mathrm{miss}}$ components.

To estimate the impact on the resolution due to in-time (i.e. per bunch collision) PU activity, the RMS of the combined E_x^{miss} and E_y^{miss} distributions is plotted in bins of the number of reconstructed vertices (N_{PV}) in the event. In Figure 4.7 the evolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of N_{PV} is shown. The degradation in the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution follows as a consequence of residual calorimeter noise fluctuations through the jet term. While PU effects are mostly suppressed, the fluctuations introduce greater mis-measurement of the jet energy resolution, leading to a poor jet p_{T} measurement. Overall, the agreement between data and simulation is good and most dis-agreements fall within the systematics uncertainty band. For events without any jets, the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is mostly insensitive to PU activity because of the suppression of tracks from PU vertices in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are implemented. The

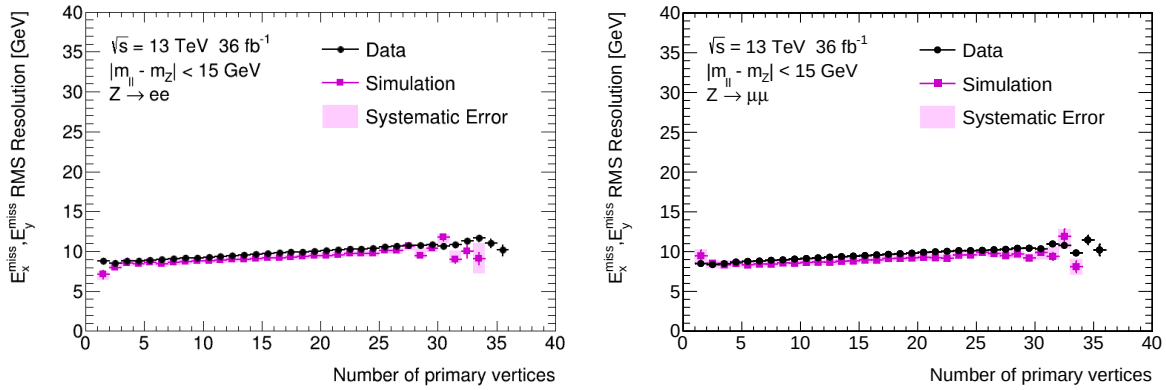


Figure 4.8: $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution plotted as a function of the number of primary vertices in a jet veto category with integrated luminosity up to 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

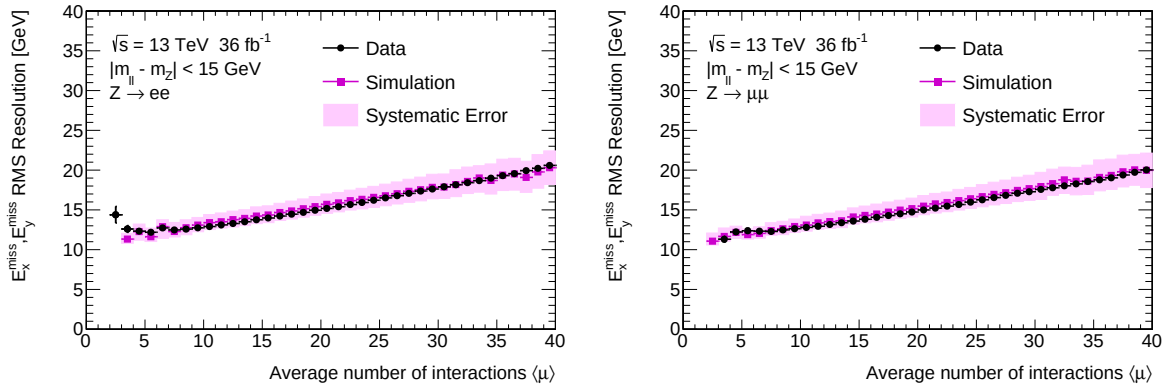


Figure 4.9: $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution vs the average number of interactions $\langle\mu\rangle$ for the $Z \to ee(\mu\mu)$ channel with integrated luminosity up to 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is presented in Figure 4.8. Some differences between data and simulation arise due to residual PU jet inclusion resulting in broadening of the E_x^{miss},E_y^{miss} distributions. A small degradation of the resolution with increasing PU can also be observed for both data and MC. This is caused by merging of vertices as more soft PU interactions are added to the soft term [141].

Out-of-time event activity exist due to the short LHC bunch-crossing time window. The average number of interactions, $\langle\mu\rangle$, taken as the average of the number of reconstructed vertices per event over a two-minute window is used to quantify the out-of-time PU. Large dis-agreements in the $\langle\mu\rangle$ range can indicate imperfect modelling of the PU description provided by the minimum bias studies. In Figure 4.9 the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of $\langle\mu\rangle$ for data and simulation is presented. Here, as in the N_{PV} case, the overall agreement between data and simulation is good. Any disagreement is covered by the systematic uncertainties.

The resolution can also be measured as a function of the overall energy in an event which is

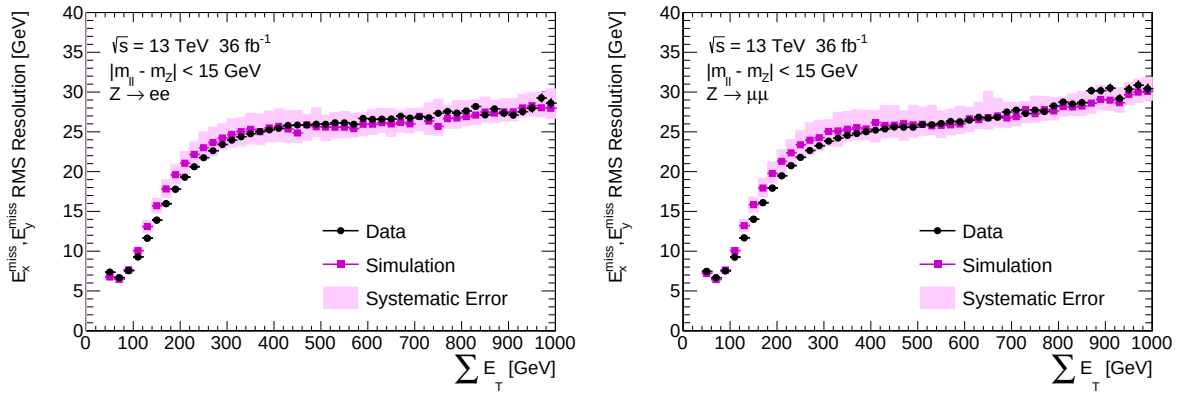


Figure 4.10: $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution vs $\sum E_T$ distributions for the $Z \to ee(\mu\mu)$ channel with integrated luminosity up to 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band is formed from the statistical and experimental systematic errors on the MC.

quantified by the scalar p_{T} sum (ΣE_{T}) of all reconstructed constituents. The scalar sum is defined by:

$$\sum E_{\mathrm{T}} = \sum_{\substack{\text{selected}\\ \text{electrons}}} p_{\mathrm{T}}^{e} + \sum_{\substack{\text{accepted}\\ \text{photons}}} p_{\mathrm{T}}^{\gamma} + \sum_{\substack{\text{accepted}\\ \tau\text{-leptons}}} p_{\mathrm{T}}^{\tau_{\mathrm{had}}} + \sum_{\substack{\text{selected}\\ \text{muons}}} p_{\mathrm{T}}^{\mu} + \sum_{\substack{\text{accepted}\\ \text{jets}}} p_{\mathrm{T}}^{\mathrm{jet}} + \sum_{\substack{\text{unused}\\ \text{tracks}}} p_{\mathrm{T}}^{\mathrm{track}}. \tag{4.10}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution against the ΣE_{T} is shown in Figure 4.10. The dependency for low ΣE_{T} ($\sim 100 - 200$) GeV originates from leptons and soft term. In these regions the resolution is proportional to the lepton resolution, and the width of the E_x^{miss},E_y^{miss} distributions evolves with $\propto p_{\mathrm{T}}^{\ell}$. For high ΣE_{T}, the width is dominated by the p_{T} resolution of the jets. As the jets are reconstructed from the calorimeters, the resolution scales as $\sigma_{x(y)}^{\mathrm{miss}} \propto \sqrt{\sum E_{\mathrm{T}}}$. Discrepancies between data and simulation are mostly covered by systematic uncertainties. In the low ΣE_{T} bins, the dis-agreement indicates a better resolution estimate in data than simulation due to imprecise modelling in MC of low jet activity events.

4.3.2.2.3 $E_{\mathrm{T}}^{\mathrm{miss}}$ tails Large values of $E_{\mathrm{T}}^{\mathrm{miss}}$ can be caused by genuine physics processes such as production of invisible or weakly interacting particles. However, they can also be generated by mis-reconstruction of physics objects. The E_x^{miss}, E_y^{miss} distributions can develop non-Gaussian profiles, which create larger tails in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Such behaviour can be caused by object resolutions with non-Gaussian shapes and/or event-by-event reconstruction fluctuations. The enhanced tails in $E_{\mathrm{T}}^{\mathrm{miss}}$ are estimated by the fraction of events containing $E_{\mathrm{T}}^{\mathrm{miss}} > E_{\mathrm{T}}^{\mathrm{miss,threshold}}$:

$$f_{\mathrm{tail}} = \frac{1}{\mathcal{H}} \int_{E_{\mathrm{T}}^{\mathrm{miss,threshold}}}^{\infty} h(E_{\mathrm{T}}^{\mathrm{miss}}) dE_{\mathrm{T}}^{\mathrm{miss}}, \text{ where } \mathcal{H} = \int_{0}^{\infty} h(E_{\mathrm{T}}^{\mathrm{miss}}) dE_{\mathrm{T}}^{\mathrm{miss}} \tag{4.11}$$

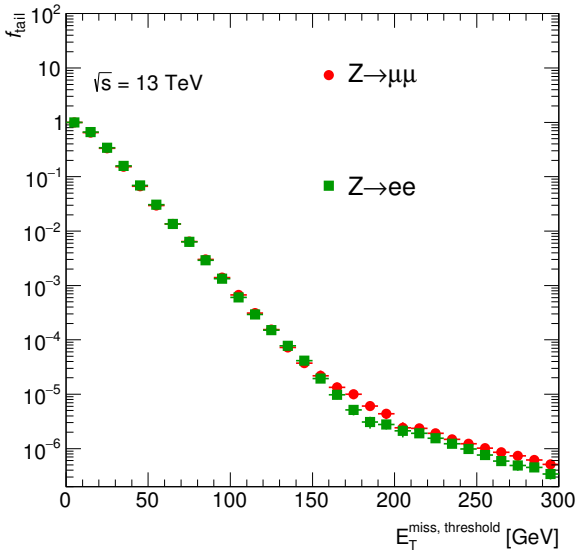


Figure 4.11: $E_{\mathrm{T}}^{\mathrm{miss}}$ tail fraction for Z + jets events. Slight degradation in the $E_{\mathrm{T}}^{\mathrm{miss}}$ tail for the muons with respect to the electrons reflects the poorer object reconstruction at high p_{T}.

where $h(E_{\mathrm{T}}^{\mathrm{miss}})$ is the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution and $E_{\mathrm{T}}^{\mathrm{miss,threshold}}$ is a threshold value, which is used to estimate the tail fraction f_{tail}. Improvements in $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution are reflected as a decrease in the tail fraction f_{tail} for fixed $\mathcal{H}$. This can indicate improvements in the TST reconstruction due to removal of vertices with poor quality. Figure 4.11 shows the tail fraction for $Z \to \ell\ell$+jets events for a varying threshold range between 0 and 300 GeV. Overall good agreement exists between the $Z \to ee$ and $Z \to \mu\mu$ channels except at large threshold values. In the large $E_{\mathrm{T}}^{\mathrm{miss}}$ regions, the dominant $E_{\mathrm{T}}^{\mathrm{miss}}$ contribution is not from PU, but from inefficiencies associated with the constituent reconstruction. The muon reconstruction worsens due to poorer resolution measurement at large p_{T}.

4.3.2.2.4 $\boldsymbol{E_{\mathrm{T}}^{\mathrm{miss}}}$ systematics The systematics for the hard $E_{\mathrm{T}}^{\mathrm{miss}}$ terms are due to the uncertainties in the measurements of the objects entering in this term. The TST is entirely associated with the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction process and specific uncertainties are computed for it.

The uncertainties on the TST are found from projections of the $\mathbf{p}_{\mathrm{T}}^{\mathrm{soft}}$ $(= -\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{soft}))$ onto $\mathbf{p}_{\mathrm{T}}^{\mathrm{hard}}$ $(= -\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{hard}))$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties encompass systematics on the scale and resolution. The systematics on the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale are obtained from the maximal difference between data and simulation for the average of the parallel projection, $\langle \mathcal{P}_{||} \rangle$, as a function of $p_{\mathrm{T}}^{\mathrm{hard}}$. The systematic uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is obtained from the maximal difference between data and simulation in the variances (RMS^2) of the parallel and perpen-

dicular projections, defined as:

$$\mathrm{RMS}_{\parallel}^2 = \langle(\mathcal{P}_{\parallel})^2\rangle - \langle\mathcal{P}_{\parallel}\rangle^2 \text{ and } \mathrm{RMS}_{\perp}^2 = \langle(\mathcal{P}_{\perp})^2\rangle \tag{4.12}$$

where $\langle\mathcal{P}_{\perp}\rangle = 0$. All the uncertainties are estimated in 0 jet bins where the soft term is expected to point in the direction of the hadronic recoil. A complete discussion is provided in [4].

4.3.3 Alternative $\mathbf{E_T^{miss}}$ reconstruction

It is natural to research for improvements to the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithm discussed thus far. Such studies are discussed here, but first a brief overview on the important details of the currently used $E_{\mathrm{T}}^{\mathrm{miss}}$ are given.

The current $E_{\mathrm{T}}^{\mathrm{miss}}$ definition can be categorized as "Loose" as it contains the loosest selections on the jets (i.e. a $p_T > 20$ GeV requirement). It can also be further split by the technique of reconstructing jets i.e. the anti-k_t electromagnetic topological (EMTopo) cluster algorithm.

An extension of the loose $E_{\mathrm{T}}^{\mathrm{miss}}$, called "Tight", is made by requiring the jets in the forward detector region ($|\eta| > 2.4$) to have a p_{T} cut greater than 30 GeV. The aim of the tight $E_{\mathrm{T}}^{\mathrm{miss}}$ is to reject the PU contribution present in the low jet p_{T} spectrum in the forward region. While offering an improvement to the resolution, it can exclude objects which help in balancing the $E_{\mathrm{T}}^{\mathrm{miss}}$ recoil against the Z boson. Therefore a degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale is expected.

Another alternative to the current definition is to reconstruct $E_{\mathrm{T}}^{\mathrm{miss}}$ purely from tracks, without taking into account the cluster deposits in the calorimeter. It makes use of the excellent vertex resolution offered by ATLAS ID. The track $E_{\mathrm{T}}^{\mathrm{miss}}$ provides a more robust PU rejection and renders the resolution insensitive to the increasing PU activity. However, this method also shares the same caveats as the TST, namely neutral particles are missed and coverage is bound by the limits of the ID tracking chamber volume ($|\eta| < 2.5$).

As was already mentioned, the jets are solely reconstructed from EMTopo clusters. However, a different technique has recently become available, specifically, the particle-flow (PFlow) method. The PFlow approach associates tracks in the ID to topological clusters in the calorimeters. This method suppresses deposits from PU in the calorimeter. In addition, the momentum of tracks is used if the tracker resolution is smaller than the calorimeter resolution. The PFlow jets are reconstructed with the anti-k_t algorithm with $R = 0.4$ parameter using as inputs charged particles originating from the primary vertex and neutral PFlow objects. A full discussion of the PFlow algorithm is provided in [142].

The PFlow reconstruction offers a reduced size of JER and PU. A JVT selection of 0.2 is

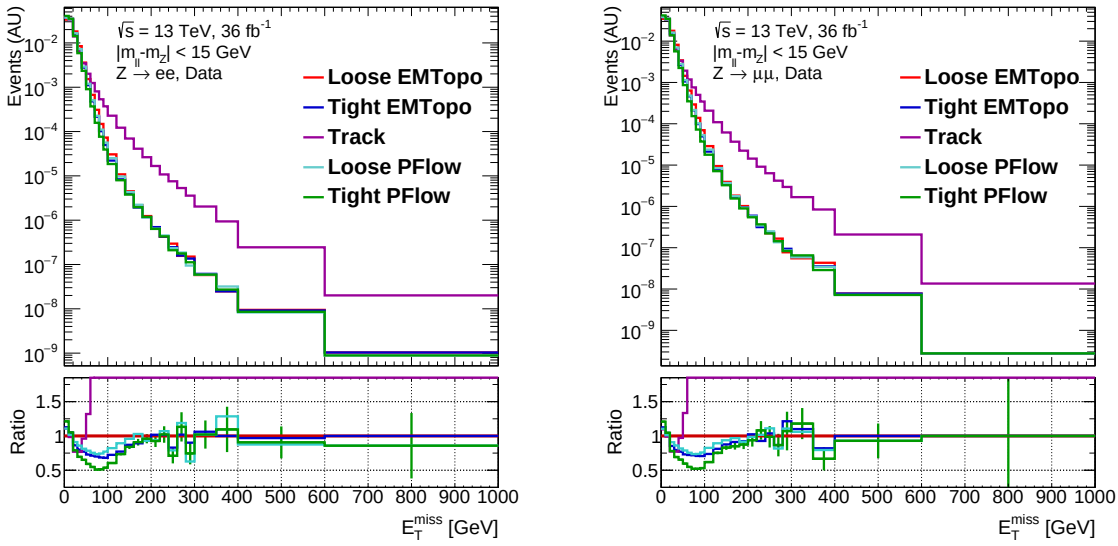


Figure 4.12: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution comparisons in data for all the available reconstruction options. The ratio is formed by dividing the various $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions by the loose EMTopo $E_{\mathrm{T}}^{\mathrm{miss}}$.

applied for jet with p_{T} less than 60 GeV and $|\eta| < 2.4$ to further suppress additional PU interactions.

The PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ follow similar definitions to the EMTopo case, with a jet p_{T} requirement greater than 20 GeV for the loose PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ case and forward jet p_{T} higher than 30 GeV for the tight PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ case. Currently only a small JVT selection of 0.2 is applied to reject certain soft hadronic PU contributions. Further details on the PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are covered in 8.1.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the different definitions are shown in Figure 4.12 in data. The greatest improvement occurs for tight PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ below 100 GeV by 50 % due to the better PU rejection. This conclusion is also supported by the comparison of the Z + jets MC, which is shown in Figure 4.13. The $E_{\mathrm{T}}^{\mathrm{miss}}$ tails in data are comparable between most definitions as expected.

4.3.3.1 Performance of the different reconstruction algorithms

All of the proposed $E_{\mathrm{T}}^{\mathrm{miss}}$ variants come with pros and cons, which are reflected in their performance. In Figure 4.14, the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution degradation as a consequence of the increasing PU and jet activity is compared for all the proposed $E_{\mathrm{T}}^{\mathrm{miss}}$ definitions. The best resolution measurement can be found in the tight category of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms because of the removal of low p_{T} forward jets. Additionally, the PFlow algorithm perform better than the EMTopo for both loose and tight definitions. Track $E_{\mathrm{T}}^{\mathrm{miss}}$ is insensitive to

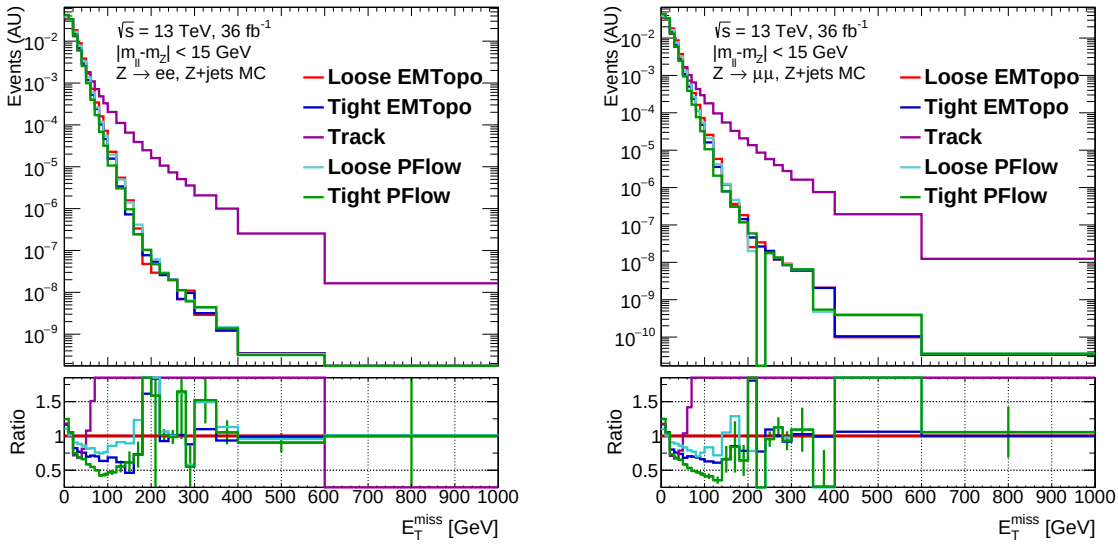


Figure 4.13: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution comparisons in Z + jets MC for all the available reconstruction options. The ratio is formed by dividing the various $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions by the loose EMTopo $E_{\mathrm{T}}^{\mathrm{miss}}$.

PU but has large resolution for increasing jet activity due to the loss of neutral particles in high p_{T} jets. This deficit is also the cause of the poor resolution at low N_{PV}.

The algorithmic improvements are also reflected in the tail fraction of the $E_{\mathrm{T}}^{\mathrm{miss}}$. In Figure 4.15, for $E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds between 40-150 GeV, the tail fraction decreases due to the better pile-up rejection. After 150 GeV, similar mis-reconstruction inefficiencies exist for all algorithms.

Improving the resolution comes at a cost, which is the degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale. Rejecting the low p_{T} forward jets results in an increased bias, as the imbalance between the Z boson and the hadronic recoil increases. Figure 4.16 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ projection on Z axis as a function of the $p_{\mathrm{T}}^{\mathrm{Z}}$. For momenta less than 90 GeV, the loose EMTopo and PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ provide the best performance, however the situation changes for large values of $p_{\mathrm{T}}^{\mathrm{Z}}$. In that region, the least bias occurs for the loose EMTopo definition[5]. The tight selections have a small harmful effect on the scale, consistent with the requirement of removing forward jets which increases the reconstruction bias. The poor performance of the track $E_{\mathrm{T}}^{\mathrm{miss}}$ definition is due to neglecting the Z boson recoil against neutral objects and missing contributions outside of the tracking volume.

[5] The degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale in the PFlow reconstruction is still a matter of further studies with Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ performance sub-groups currently investigating the PFlow jet calibration in $Z \to \ell\ell$ events.

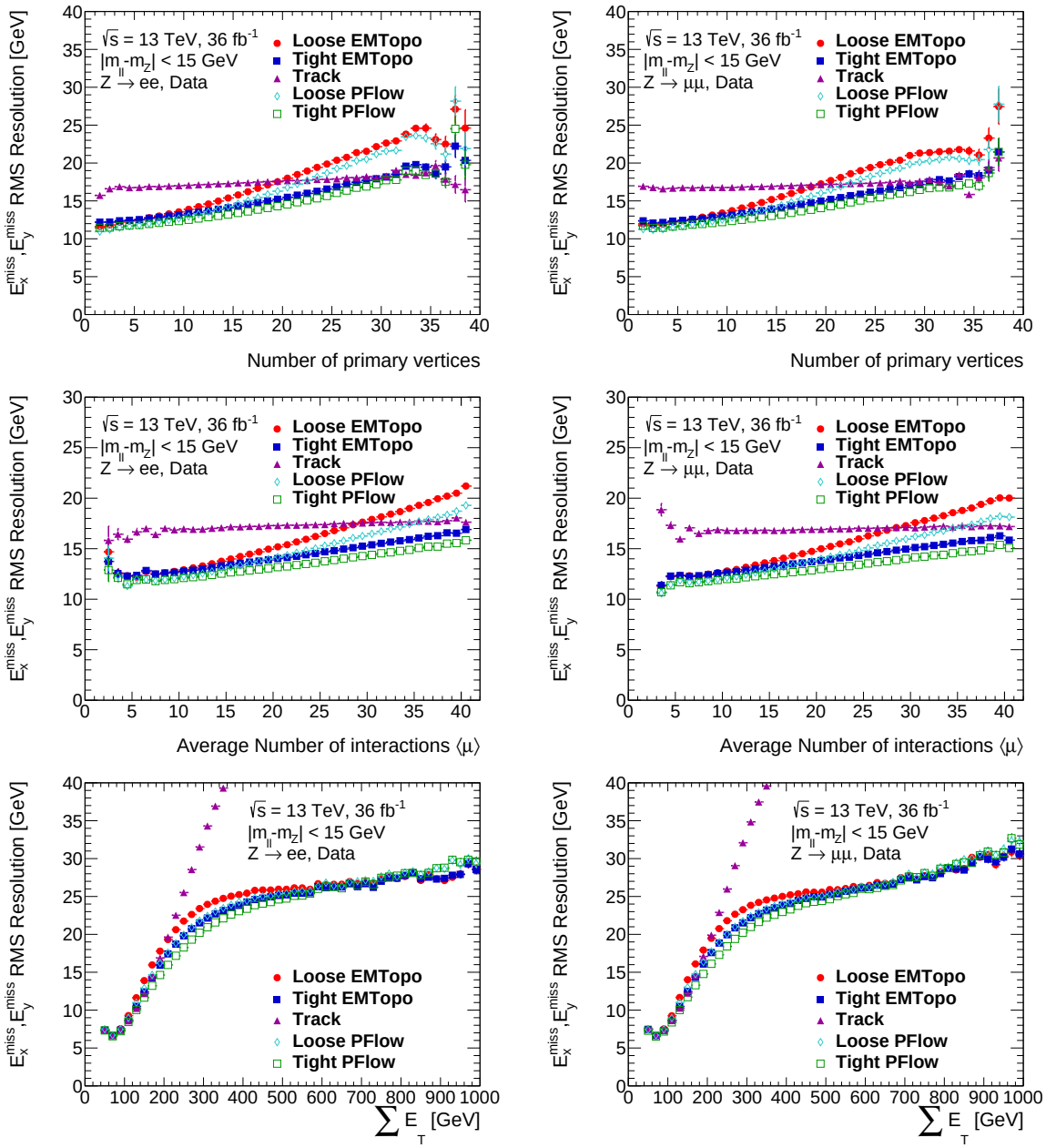


Figure 4.14: RMS resolution in data as a function of the number of primary vertices, average number of interactions and $\sum E_{\mathrm{T}}$ for different reconstruction algorithms with integrated luminosity up to 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV.

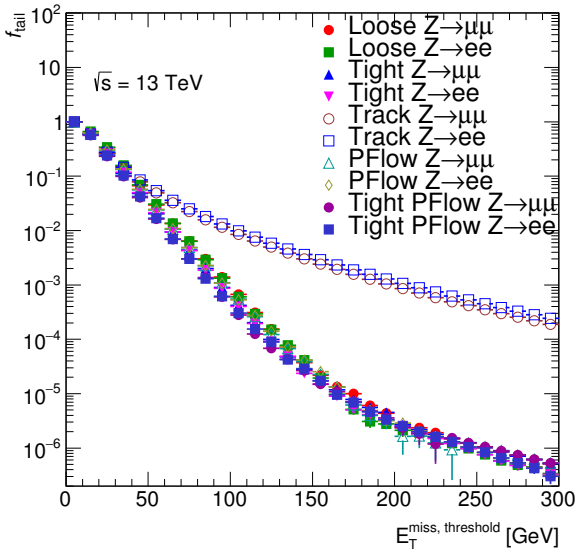


Figure 4.15: $E_{\mathrm{T}}^{\mathrm{miss}}$ tail fraction for $Z \to ee(\mu\mu)$+jets events for all $E_{\mathrm{T}}^{\mathrm{miss}}$ definitions.

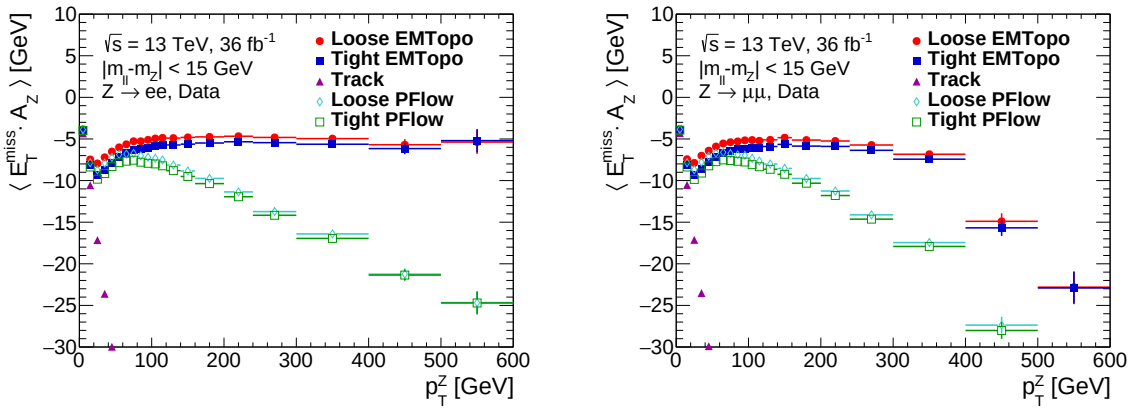


Figure 4.16: $E_{\mathrm{T}}^{\mathrm{miss}}$ response in data as a function of the Z boson p_{T} with integrated luminosity up to 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV.

Chapter 5

Analysis methodology

This chapter presents the search strategy for the existance of a Higgs particle in the mass range between 300 GeV and 1.5 TeV. The search signature looks for a Higgs boson which decays to two Z bosons where one of the Z bosons decays to a lepton pair, $\ell^+\ell^-(= e, \mu)$, while the other Z boson decays to neutrinos. Events with a lepton-decaying Z boson are found with a window selection on the mass of the Z boson. The neutrinos are invisible and are detected only from the amount of missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$. Since the neutrinos cannot be reconstructed, the invariant mass of the event, m_{ZZ}, cannot be determined. Instead a proxy variable, the transverse mass, $m_{\mathrm{T}}^{\mathrm{ZZ}}$, seeks to approximate the behaviour of m_{ZZ}.

The $H \to ZZ \to \ell\ell\nu\nu$ analysis is studied with 36.1 fb^{-1} data, recorded in pp collision at $\sqrt{s} = 13$ TeV using the ATLAS detector. This is the updated version of the studies carried out during run I [62].

The importance of the $\ell\ell\nu\nu$ final state is motivated in section 5.1. A detailed description of the analysis strategy, including all of the relevant selections and optimization methodologies, is placed in section 5.2. Information on the simulated samples used for signal and background processes is provided in 5.3. Distributions of the analysis variables are presented in section 5.4, including an overview of selection optimizations and signal region modelling.

5.1 The significance of the $H \to ZZ \to \ell\ell\nu\nu$ final state

The $H \to ZZ \to \ell\ell\nu\nu$ analysis relies on the topology defined by an opposite-sign same flavour lepton pair ($\ell^+\ell^-,\ \ell = (e, \mu)$) + large $E_{\mathrm{T}}^{\mathrm{miss}}$ signature. The leptons in the pair can be either opposite sign electrons or muons. This rather simple event state is very powerful in the searches of Higgs and BSM physics [143–145].

In the context of Higgs searches with Z bosons, the benefits of $\ell^+\ell^- + E_{\mathrm{T}}^{\mathrm{miss}}$ event signature

follow from the $H \to ZZ$ decay which contributes $\sim 1/3$ to the Higgs branching fraction when m_{H} is larger than 200 GeV (see Figure 2.5). The branching fraction of lepton pair production is small, with 3.4% per pair. However, since the Z can decay in any of the three neutrino flavours $(\nu_e, \nu_\mu, \nu_\tau)$, the fraction is increased by ~ 21 %, which provides an enhanced rate of $\ell\ell\nu\nu$ event production. Other decay chains, such as $H \to ZZ \to qq\nu\nu$ or $H \to ZZ \to qq\ell\ell$ also contain sizeable event yields, since the $Z \to qq$ process has a branching fraction of $\sim 69\%$. However, these final states suffer from QCD background contamination and mainly benefit from improved jet resolution measurements at masses above ~ 1 TeV. The strength of the $H \to ZZ \to \ell\ell\nu\nu$ analysis is in the intermediate Higgs mass range, as shown in Figure 2.7, where the sensitivity is mostly limited by certain theoretical and detector related uncertainties (see Table 7.8).

5.2 Analysis Strategy

The analysis defines the search strategy employed to enhance the sensitivity to the signal of interest. The signal topology is defined by the modes of production i.e. gluon-gluon fusion (ggF), $gg \to H$, and vector boson fusion (VBF), $qq \to qqH$. Each signal exists for a range of Higgs boson mass points between 300 GeV and 1.5 TeV. A separate search for spin-2 resonances is also carried out. The details behind the generation of the signal samples are given in section 5.3.1.

A number of SM produced processes also have a di-lepton + $E_{\mathrm{T}}^{\mathrm{miss}}$ event signature and can appear as a genuine signal events. These signatures are considered as "backgrounds" to the Higgs resonance signal. They are determined from data or simulation depending on their particular event properties. The backgrounds mostly consist of diboson (ZZ, WZ, WW), top $(t\bar{t}, Wt$, single-$t)$ and Z + jets ($Z \to (ee, \mu\mu, \tau\tau)$+jets) production. The samples which simulate these SM processes are presented in section 5.3.2, while their estimation is the subject of section 6.1.

The analysis seeks to find events which contain two opposite sign and same flavour leptons originating from a Z boson, and in addition a large $E_{\mathrm{T}}^{\mathrm{miss}}$ signature. Events containing a third lepton (either electron or muon) are vetoed, to reduce WZ production. The Z boson is selected by a window cut around the Z mass, m_{Z}. A cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$ is placed to ensure events have a genuine missing momentum and suppress any $E_{\mathrm{T}}^{\mathrm{miss}}$ due to mis-reconstruction. Angular selections are applied to select events with two leptons in the same hemisphere and a back-to-back topology with $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$. Additional cuts are enforced to limit the contribution of certain background processes.

The strategy consists of several selection stages. The first stage is the object selection (sec-

tion 5.2.1) which presents a summary of the kinematical and quality requirements on electrons, muons and jets in order to be considered for further study. Afterwards events are preselected (section 5.2.2) where events in data are removed if found to have spurious effects, such as those containing erroneous detector measurements. Event selection requirements based on trigger flags are briefly discussed in section 5.2.3. The final step is the analysis event selection (section 5.2.4), where the main part of the Higgs search is described.

The study discussed so far is inclusive of both Higgs boson production modes meaning that no distinctive features have been exploited to separate events produced from $gg \to H$ and $qq \to Hqq$ interactions. Defining two separate phase space regions for each mode is achieved through selections on the topology of the leading and sub-leading jet system. Events which pass these requirements are classified into a VBF signal region (section 5.2.5), while events which fail them are placed into a ggF signal region (section 5.2.6). Both categories look for excesses of events in data over the predicted rate in the individual lepton channel modes, $ee\nu\nu$ and $\mu\mu\nu\nu$, and in their combination, $\ell\ell\nu\nu$.

The discriminating variable between signal and background processes is the transverse mass of the ZZ system, $m_{\mathrm{T}}^{\mathrm{ZZ}}$. The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ is defined as:

$$m_{\mathrm{T}}^{\mathrm{ZZ}} \equiv \sqrt{\left[\sqrt{m_{\mathrm{Z}}^2 + |\mathbf{p}_{\mathrm{T}}^{\mathrm{Z}}|^2} + \sqrt{m_{\mathrm{Z}}^2 + |\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}|^2}\right]^2 - \left[\mathbf{p}_{\mathrm{T}}^{\mathrm{Z}} + \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}\right]^2} \tag{5.1}$$

where m_{Z} is the mass of the Z boson, fixed at 91.2 GeV, $\mathbf{p}_{\mathrm{T}}^{Z}$ is transverse momentum of the Z boson, obtained from $\mathbf{p}_{\mathrm{T}}^{\mathrm{Z}} = \mathbf{p}_{\mathrm{T}}^{\ell\ell} = \mathbf{p}_{\mathrm{T}}^{\ell_1} + \mathbf{p}_{\mathrm{T}}^{\ell_2}$, and $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ indicates the amount of missing momentum, determined from equation 4.5. The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ variable is a proxy of the m_{ZZ} since the invariant mass cannot be obtained as the energy and momentum of the individual neutrinos cannot be reconstructed.

5.2.1 Object Selection

The first stage of the analysis consists of the selections on physics objects. The requirements of the electrons, muons and jets are described in sections 4.2.1, 4.2.2 and 4.2.3, respectively. The two leading p_{T} leptons which pass all object selections are classified as "signal", while any extra leptons are classified as "veto". The removal of object reconstruction overlap is summarized in section 4.2.4. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is built from loose quality electrons, muons and anti-k_t EMTopo jets with a radius parameter, ΔR, equal to 0.4. Selections on constituents inside the $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm are discussed further in section 4.3. Table 5.1 presents all the prerequisites on the objects in order to be processed further.

Object	Object Selection
Electrons	$E_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.47$
	"Medium" object quality condition
	$\lvert d_0/\sigma_{d_0}\rvert < 5$ and $\lvert z_0 \cdot \sin(\theta)\rvert < 0.5$ mm
	"Loose" isolation criteria
	OR between e and μ and OR between e and jets
	Signal (veto) e has $E_{\mathrm{T}} > 20$ (7) GeV and LH "Medium" ("Loose")
Muons	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.5$
	Reconstructed using combined algorithm
	$\lvert d_0/\sigma_{d_0}\rvert < 3$ and $\lvert z_0 \cdot \sin(\theta)\rvert < 0.5$ mm
	"Loose" isolation criteria
	OR between μ and jets
	Signal (veto) μ has $p_{\mathrm{T}} > 20$ (7) GeV and "Medium" ("Loose") id criteria
Jets	anti-k_t EMTopo reconstruction algorithm with $\Delta R = 0.4$
	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 4.5$
	JVT > 0.59 (jets with $p_{\mathrm{T}} < 60$ GeV, $\lvert\eta\rvert < 2.4$)
	OR between jets and μ
	"LooseBad" quality criteria
$E_{\mathrm{T}}^{\mathrm{miss}}$	"Loose" e and μ ($p_{\mathrm{T}} > 7$ GeV) before OR
	"Loose" jets ($p_{\mathrm{T}} > 20$ GeV and JVT > 0.59 for $p_{\mathrm{T}} < 60$ GeV, $\lvert\eta\rvert < 2.4$)
	Soft term built from tracks in ID

Table 5.1: Summary of object selections used in the $\ell\ell\nu\nu$ analysis.

5.2.2 Event Pre-Selection

A set of event quality criteria must be satisfied before any analysis selections are done. The criteria consist of selections to remove spurious events in data, due to detector inefficiencies.

First, the events must be a part of the good runs list (GRL) [146, 147] otherwise they are discarded. The GRL selects events which use all detector components, which renders the total integrated luminosity, for the 2015+2016 data taking period, to 36.1 fb^{-1}. In addition, all events are required to be fully reconstructed to remove detector mis-measurements. Hits in the SCT are checked to ensure events are not affected by problems in SCT readout. LAr and Tile quality flags are checked in order to protect against noise bursts and data corruption. Events are also required to contain a primary vertex, by finding a hard-scatter vertex with the largest $\sum p_T^2$ with at least two outgoing associated tracks.

Low quality jets give rise to large fake $E_{\text{T}}^{\text{miss}}$ that results in a poor description of the tail of the $E_{\text{T}}^{\text{miss}}$ distribution. In order to remove jets originating from non-collision effects, such as hardware problems, cosmic-ray showers or beam related backgrounds, a set of jet quality criteria is applied. Events which do not pass the loose jet quality selections are rejected [148].

The event pre-selections used in the $H \to ZZ \to \ell\ell\nu\nu$ analysis are summarised in Table 5.2.

Required to be in the GRL.
Not flagged as being incomplete.
SCT error requirement i.e. SCT hits are recorded.
LAr and Tile error requirements
Required to have at least one hard-scattering vertex with at least two associated tracks.
Pass “Loose” jet quality criteria.

Table 5.2: Summary of events pre-selections used in the $\ell\ell\nu\nu$ analysis.

5.2.3 Trigger Menus

Before moving to the description of the analysis, an important overview of the trigger menus must be made. The menus and single lepton triggers were described in section 3.6. In 2015, the combination of lepton triggers in the $H \to ZZ \to \ell\ell\nu\nu$ analysis was investigated and found to always exceed 95% efficiency [149]. Due to the sharp turn-on curve of single lepton trigger efficiencies, a $p_{\text{T}} > 30$ GeV requirement on the leading in p_{T} lepton is applied, which

has been proved to keep the high signal acceptance in the analysis. The lowest un-prescaled triggers changed between 2015 and 2016 due to the increase in the instantaneous luminosity, L. In 2016, L reached the LHC design value of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. The performance of the 2016 triggers are presented in Figure 3.10 (Figure 3.11) for the $ee\nu\nu$ ($\mu\mu\nu\nu$) channel. The full trigger flags implemented in the analysis are provided in Table 5.3.

	Trigger selection
Single Muon	mu20_iloose_L1MU15 OR mu50 (2015) mu24_ivarmedium OR mu50 (2016) mu26_ivarmedium OR mu50 (2016, after $10^{34}\text{cm}^{-2}\text{s}^{-1}$)
Single Electron	e24_lhmedium_L1EM20VH OR e60_lhmedium OR e120_lhloose (2015) e24_lhtight_nod0_ivarloose OR e60_lhmedium_nod0 (2016) OR e140_lhloose_nod0 (2016) e26_lhtight_nod0_ivarloose OR e60_lhmedium_nod0 (2016, after $10^{34}\text{cm}^{-2}\text{s}^{-1}$) OR e140_lhloose_nod0 (2016, after $10^{34}\text{cm}^{-2}\text{s}^{-1}$)

Table 5.3: Trigger requirement in $H \to ZZ \to \ell\ell\nu\nu$ analysis in 2015 and 2016 data periods. The single lepton triggers are un-prescaled in the 2015 and 2016 data taking period. The minimum trigger threshold changed in 2016, following the increase in luminosity.

5.2.4 $H \to ZZ \to \ell\ell\nu\nu$ event selection

The selections consist of specific kinematic requirements on the $H \to ZZ \to \ell\ell\nu\nu$ final state in order to separate signal events from various background processes. A summary of all applied analysis selections is provided in Table 5.4.

The set of event selections to isolate the Higgs signal starts with trigger requirements on the lepton objects. For the event to survive, the leptons are needed to fire at least one of the trigger flags, which are defined in section 5.2.3. Events are then required to contain at least two opposite-charge leptons of the same flavour and satisfy the medium id criteria. If a third lepton is present then the event is vetoed to reduce contamination from the WZ background. The lepton with a higher value of p_{T} from the di-lepton pair, must have a transverse momentum greater than 30 GeV due to the sharp turn-on curve of the single lepton triggers.

To select events which contain a Z boson, the invariant mass of the lepton pair, $m_{\ell\ell}$, must

have a value close to Z mass. More specifically, $m_{\ell\ell}$ is required to be within 15 GeV of m_{Z}, $|m_{\ell\ell} - m_{\mathrm{Z}}| < 15$ GeV, where m_{Z} is equal to 91.187 GeV. The Z mass condition reduces the non-Z backgrounds such as $t\bar{t}$ and WW.

Since the signal signature has a pair of neutrinos in the final state, a large $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is also applied. This requirement aides in restraining the Z + jets background, which peaks for low $E_{\mathrm{T}}^{\mathrm{miss}}$ values. The threshold for events to be admitted in the signal region is 120 GeV, where the majority of events are expected to have a genuine missing transverse momentum.

Since the leptons are produced from a Z boson with non-zero transverse momentum, they are expected to be oriented in the same relative η and ϕ direction. More precisely, the two leptons are required to have an angular separation, $\Delta R_{\ell\ell}$, less than 1.8, where $\Delta R_{\ell\ell} = \sqrt{\Delta\eta_{\ell\ell}{}^2 + \Delta\phi_{\ell\ell}^2}$.

In a signal signature, the two Z bosons, produced from the decay of a heavy Higgs particle, fly in opposite directions. This translates into a condition where the expected separation between the leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ is very large. In particular, the azimuthal angle between the lepton pair and $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\mathrm{miss}})$ must be greater than 2.7. This condition further suppresses backgrounds such as Z + jets.

Due to the specific topology of the signal event, the Z boson p_{T} can be expected to recoil against the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$, provided no initial or final state radiation is included. Imbalances in the recoil are quantified by the fractional p_{T}, $|{E_{\mathrm{T}}}^{\mathrm{miss,jet}} - {p_{\mathrm{T}}}^{\ell\ell}|/{p_{\mathrm{T}}}^{\ell\ell}$, where $E_{\mathrm{T}}^{\mathrm{miss}}$ is modified by the vectorial sum of the transverse momentum of jets, ${E_{\mathrm{T}}}^{\mathrm{miss,jet}} = |\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} + \Sigma\mathbf{p}_{\mathrm{T}}(\mathrm{jet})|$. This is done in order to reject events originating from the Z + jets background. The fractional p_{T} cut is required to be less than 20 %.

Further reduction of Z + jets contamination is accomplished with a set of quality criteria. The first requirement is applied on the minimum distance between any jet with p_{T} greater than 100 GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$, min $\Delta\phi(p_{\mathrm{T}}^{\mathrm{jet}} > 100$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}})$. It is required to have a value greater than 0.4. This suppresses the contributions of mis-measured jets in the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction. The second selection uses $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$, where H_{T} is defined as the scalar sum of the transverse momenta of the jets and leptons. The Z + jets events tend to have large H_{T} values without a correspondingly large $E_{\mathrm{T}}^{\mathrm{miss}}$ and thus cluster at low $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$, in distinction to the signal. A threshold value of 0.4 was found to be satisfactory as further described in section 5.2.4.1.

Lastly, to decrease the contamination from the $t\bar{t}$ background, a veto on any b-tagged jets is applied. For the jets to be b-tagged, the multivariate algorithm output, MV2c10 [133], is required to have a value greater than the cut-value of 0.1758. This cut-value corresponds to a 85% efficiency working point of correctly tagging a b-jet (WP) in a $t\bar{t}$ sample.

The single lepton (electron and muon) triggers are required
Event should contain a same-flavor, opposite-charge lepton pair (muons or electrons)
$p_{\mathrm{T}}^{\ell,\text{leading}} > 30$ GeV, $p_{\mathrm{T}}^{\ell,\text{sub-leading}} > 20$ GeV
No additional 'veto' leptons are required (with $p_{\mathrm{T}} > 7$ GeV)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV
$E_{\mathrm{T}}^{\text{miss}} > 120$ GeV
$\Delta R_{\ell\ell} < 1.8$ GeV
$\Delta\phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\text{miss}}) > 2.7$ GeV
Minimal $\Delta\phi(p_{\mathrm{T}}^{\text{jet}} > 100$ GeV$, E_{\mathrm{T}}^{\text{miss}}) > 0.4$
$\lvert {E_{\mathrm{T}}}^{\text{miss,jet}} - {p_{\mathrm{T}}}^{\ell\ell}\rvert / {p_{\mathrm{T}}}^{\ell\ell} < 0.2$
$E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}} > 0.4$, where $H_{\mathrm{T}} = \sum p_{\mathrm{T}}^{\text{lep}} + \sum p_{\mathrm{T}}^{\text{jet}}$
No b-tagged jet (85% WP, $p_{\mathrm{T}} > 20$ GeV)

Table 5.4: Summary of $H \to ZZ \to \ell\ell\nu\nu$ events selections.

5.2.4.1 Discussion of $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ selection

The $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ selection was introduced to the analysis during the later stages of the 2016 data taking period when increased PU activity resulted in more Z + jets events. The $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ variable provides discrimination between events with non-negligible $E_{\mathrm{T}}^{\text{miss}}$ value, caused by object mis-reconstruction, and genuine $E_{\mathrm{T}}^{\text{miss}}$ events. The purpose of the selection on the $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ is therefore to act as a quality cut, to target events with fake $E_{\mathrm{T}}^{\text{miss}}$ originating from Z + jets.

A cut-based optimization strategy was carried out with a statistical significance metric defined as:

$$Z = \sqrt{2((S+B)\log(1+S/B) - S)} \tag{5.2}$$

where S and B are the signal and background yields. This formula holds in regions of low and high statistics provided the background uncertainty is small [150]. For regions with large background yields, the significance reduces to the standard $\sim S/\sqrt{B}$. Plots of the significance for different $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ cut thresholds are provided in Figure 5.1 for different Higgs mass points. For $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}} > 0.4$ the significance is unaltered, with a reduced number of Z + jets events as given in Table F.3. A final check was performed by computing the change in the expected exclusion limits with increasing the threshold $E_{\mathrm{T}}^{\text{miss}}/H_{\mathrm{T}}$ cut-values.

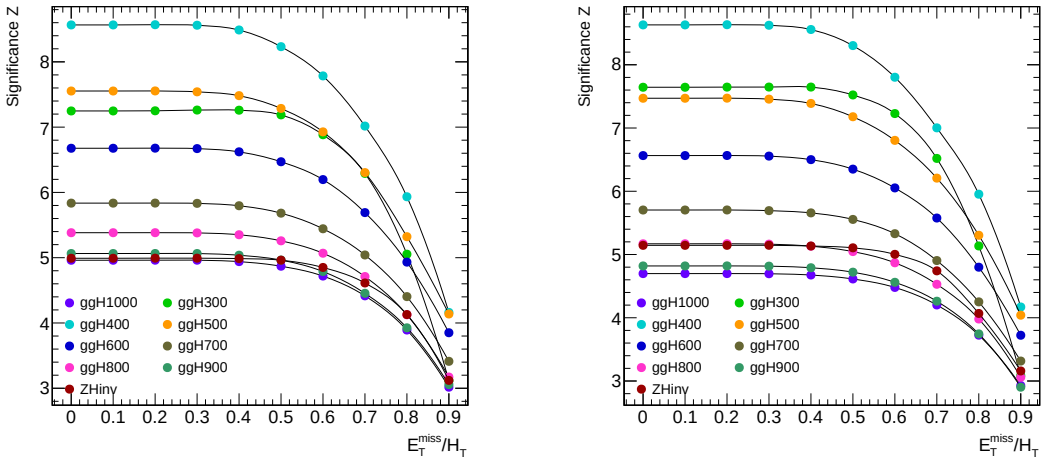


Figure 5.1: Scans of the significance, Z, for increasing $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cut-values in the $H \to ZZ \to ee\nu\nu$ (left) and $H \to ZZ \to \mu\mu\nu\nu$ (right) channels. All ggF signal samples between 300 GeV and 1 TeV are considered.

Figure 5.2 shows that for a cut-value greater than 0.4 (blue curve) the limits improve in the low m_{H} range. At large masses the limit is comparable with the no cut on $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ case.

5.2.5 $qq \to qqH$ signal region

The production of Higgs boson was outlined for the SM framework in section 2.3.1. The heavy Higgs boson is presumed to follow the same modes as the SM case, which are shown as Feynman diagrams in Figure 2.3.

The selections discussed so far have been made inclusive of the channels of Higgs boson production, without exploiting any specifics of their respective topologies. Isolating the VBF mode requires the application of tight selections on variables formed from information on the leading and sub-leading jets in the event. They are imposed in order to suppress background processes and remove overlap with the ggF Higgs signal. The selections are summarized in Table 5.5.

The first selection is applied on the number of jets, N_{jets}. They must be at least two jets to be able to form a di-jet system. The system itself is formed from the highest and second highest p_{T} jets. Thus the N_{jets} cut forms the basis of a very loose VBF phase space.

The next requirement is imposed on the transverse momentum of jets, which has to be greater than 30 GeV. It was introduced to combat an increase of PU jets in the forward $|\eta|$ regions (see section 5.4.2.2).

In the di-jet topology, the invariant mass of two jet system, m_{jj}, is valuable in discriminating

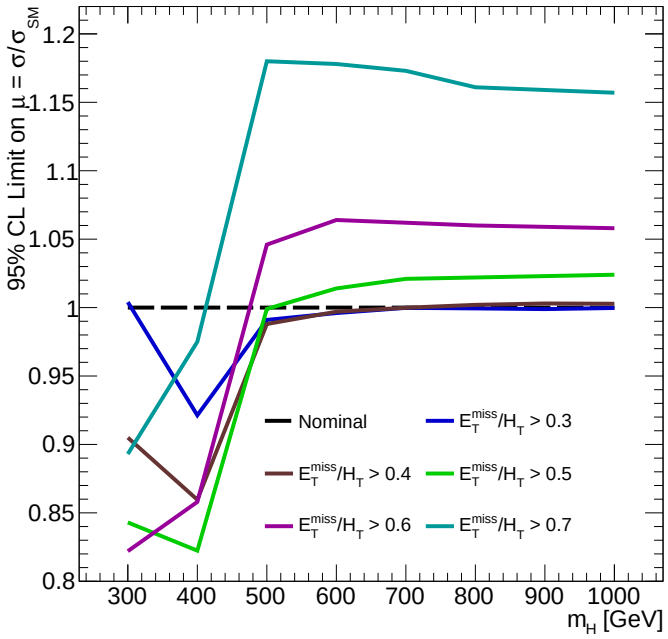


Figure 5.2: Variation of the combined expected limits with increasing $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cut for signal mass points between 300 GeV and 1 TeV. The plot shows the ratio of CL limits of the signal strength parameter, μ with respect to the nominal case. The nominal refers to the situation with no applied cut on $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$.

between signal and background production. This is due to the fact that signal events are expected to be highly energetic which is a consequence of the hard-scatter between the initial state quarks. The m_{jj} is obtained from the sum of four-momenta in spherical coordinates $(p_{\mathrm{T}}, \eta, \phi, m)$. In the analysis setup the m_{jj} must be greater than 550 GeV for an event to be considered as signal. The second parameter in the di-jet system is the magnitude of the difference between jet pseudorapidities, $|\Delta\eta_{jj}| = |\eta_{j_1} - \eta_{j_2}|$. A large cut on $|\Delta\eta_{jj}|$ ensures that the jets are in the forward part of the calorimeter. In the analysis search, the $|\Delta\eta_{jj}|$ is necessary to have a value greater than 4.4. The optimization of the selections on mjj and $|\Delta\eta_{jj}|$ are studied in section 5.2.5.1.

5.2.5.1 Optimization of VBF di-jet cut-values

In run I, the di-jet VBF selections were chosen to have values $m_{jj} > 550$ GeV and $|\Delta\eta_{jj}| > 4.4$, as described in [62]. In run II, the requirements on the di-jet system were re-visited in order to determine if an improved signal sensitivity can be achieved.

The optimization strategy for the di-jet topology is performed using a two-dimensional scan of the m_{jj} and $\Delta\eta_{jj}$ cut-values. The statistical significance, Z, is used to highlight cut-values where more sensitivity to the signal is obtained. The definition of Z follows from

Require a cut on the number of jets, $N_{\text{jets}} \geq 2$.
Minimum jet p_{T} requirement, $p_{\text{T}}^{\text{jet}} > 30$ GeV, discussed in 5.4.2.2
Di-jet mass selections, $m_{jj} > 550$ GeV
Di-jet pseudorapidity separation, $\lvert\Delta\eta_{jj}\rvert > 4.4$

Table 5.5: Summary of VBF events selections used in the $\ell\ell\nu\nu$ analysis.

equation 5.2. To better isolate the event topology the $H \to ZZ \to \ell\ell\nu\nu$ analysis selections have been applied beforehand. The VBF signal yield is estimated by MC samples which are described in section 5.3. The total background contribution includes the diboson, top and Z + jets processes. The ggF signal is also added, provided it matches with the VBF mass point, since the ggF process is also considered as a background in the VBF phase space.

The two-dimensional (m_{jj} vs $|\Delta\eta_{jj}|$) scans over the significance are given in Figure **??** for masses of 400, 700 and 1000 GeV. The significance is calculated from the cumulative distributions of the two-dimensional signal and background histograms after all of the analysis criteria are applied. The greatest sensitivity to VBF signals is expectedly found for large cut-values of m_{jj} and $|\Delta\eta_{jj}|$. Certain m_{jj} cut-values, indicated by large regions in red, greatly reduce the background statistics to a point where any selection on $|\Delta\eta_{jj}|$ is irrelevant. As a consequence it is difficult to accurately set unique cut-based values for all mass points. Selecting a loose set of selections on the di-jet variables has the adverse effect of introducing more ggF signal events, as shown in Figures C.1, C.2, where the signals have been scaled for clarity.

The choice of cut-values on the di-jet variables was reached with consideration of the balance in statistical precision (also necessary for background estimation) and ggF signal inclusion. Stricter selections on m_{jj} and $|\Delta\eta_{jj}|$ would result in a very low event yield, insufficient for proper exclusion limits. Looser selections would add more ggF signal yield, leaving large theoretical uncertainties and reducing the sensitivity to the VBF signal. The run I selections are seen, by the Z scans and di-jet distributions, to provide sufficient quality for further studies of the VBF phase space. Therefore the analysis follows the run I selections on the di-jet system.

5.2.6 $gg \to H$ signal region

The ggF signal region is created to be orthogonal to the VBF phase space. This is done to remove the overlap between the signal region defined from the $H \to ZZ \to \ell\ell\nu\nu$ event

cuts and the VBF topology. The ggF phase space consists of events passing all of the event requirements plus any events which fail one or more VBF selections. The cuts which isolate the ggF topology are given in Table 5.6.

All $H \to ZZ \to \ell\ell\nu\nu$ event selections applied.
Any events which fail any one of VBF requirements is added to ggF region.

Table 5.6: Summary of ggF events selections used in the $\ell\ell\nu\nu$ analysis.

5.2.7 Effects of selections

Each of the analysis selections impact the signals and backgrounds yields in various proportions. The effects on the ZZ diboson background, are given in Table F.1 in a m_{ZZ} window between 550 and 650 GeV. The $H \to ZZ \to \ell\ell\nu\nu$ selections overall only reduce the number of ZZ events by 20 %. This is due to the ZZ process sharing similar event kinematics as the signal of interest. Another diboson background, WZ, is presented in Table F.2 for the same m_{ZZ} cut. The WZ process is rejected with an efficiency of 70 % after all ggF analysis cuts. The reduction with each analysis selection for the Z + jets background is presented in Table F.3. It is lowered by more than 99 % with the application of the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement.

The effects of the analysis selection on two signal samples are listed in Table F.4 (Table F.5) for a ggF (VBF) signal for a Higgs mass of 600 GeV. The inclusive selections reduce the ggF (VBF) signal events by $\sim 30\%$ ($\sim 40\%$), a consequence from the $\Delta\phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\mathrm{miss}})$ selection. The application of the VBF analysis selections reduce the ggF signal contribution by more than 98%.

5.3 MC Samples

MC generators, as mentioned in section 3.8, simulate the signal and background processes present in the $H \to ZZ \to \ell\ell\nu\nu$ signature. The specific set of generators are summarized in Table 5.7 for all MC samples.

5.3.1 Signal samples

Three groups of signal samples were generated, two for a spin-0 heavy Higgs interpretation based on the mode of production and one for a spin-2 RS graviton interpretation.

The simulation of Higgs events is obtained from the Powheg-Box v2 [151] generator, where m_{H} is set to the heavy resonance mass. The samples are created separately for each mode of Higgs production. The Powheg generator contains matrix elements calculated to next-to-leading order (NLO) in the strong coupling constant α_s, also referred by as QCD. The generator is interfaced with Pythia 8.212 [152], which simulates the decay chains to $ZZ \to \ell\ell\nu\nu$ and the parton showering (PS). The Powheg generator calculates the parton kinematics with the CT10 [153] PDF. The Powheg + Pythia8 is computed with the AZNLO set of tuned parameters [154] which use the CTEQ6L1 PDF [155]. The signal samples use the narrow width approximation (NWA), as discussed in section 2.3.3. The samples are simulated in the mass range between between 300 GeV and 1.6 TeV in steps of 100 GeV for $m_{\mathrm{H}} < 1$ TeV and steps of 200 GeV for $m_{\mathrm{H}} > 1$ TeV. In section 7.6.4, a morphing technique is adopted to generate signal yields for any value of m_{H} [156]. The Higgs production cross section is summarized in Table D.1 (D.2) for $gg \to H$ ($qq \to qqH$) process for all samples used in the analysis. The VBF production cross section is adopted from ggF, but scaled by 1/10 since the number of VBF events is expected to be an order of magnitude less than ggF.

The samples which model the KK gravitons within the RS context are generated by MadGraph5_aMC@NLO [157] at leading order (LO) in α_s. The NNPDFv2.3 at LO is used to simulate the parton distribution functions [158, 159] used by MG5_aMC@NLO + Pythia8 and the A14 tune [160]. The sample simulates the decay chain through to two Z bosons for the dilepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ final state. The ratio of the curvature of the extra dimension, k, to the reduced Planck mass, $\bar{M}_{\mathrm{pl}}$, is chosen to be unity [59, 61] which sets the width of the resonance to be 6% of the graviton mass, m_{G}. The samples are produced in the mass range between 600 GeV and 2 TeV in steps of 200 GeV. They are processed through a fast detector simulation, which makes use of parametrisation of calorimeter response [161]. The production cross section of the KK gravitons are summarized in Table D.3.

5.3.2 Background samples

The main backgrounds in the heavy Higgs search are formed from diboson, top and Z + jets processes. Various MC generators can simulate them to different degrees of accuracy.

The ZZ diboson background is divided based on the modes of production into $qq \to ZZ$ and $gg \to ZZ$. The $qq \to ZZ$ process is generated by Powheg-Box v2 to NLO in QCD and interfaced with Pythia8. The PDF set is provided by CT10nlo. Next-to-next-leading order (NNLO) in QCD and NLO electro-weak corrections (EW) are added as a function of the invariant mass, m_{ZZ}, of the ZZ process [162, 163]. In the VBF category, a Sherpa sample validated the Powheg estimation, since Sherpa evaluates one parton at NLO QCD and three partons at LO QCD. The matrix element of the $gg \to VV, V = W, Z$ process is calculated

from the gg2VV v3.1.6 generator [164] at LO QCD. To remove WW interference the charged leptons flavour is made to differ from the neutrino flavour. The $ZZ \to llvv$ decay chain is simulated with Powheg + Pythia8 using the CT10 PDF and the A14 tune of parameters. Higher-order α_s corrections are calculated in massless quark loops [165–167] in the heavy-top approximation ($m_t \to \infty$) [168], which also includes the $gg \to H^* \to ZZ$ reaction [169]. The corrections are accounted by scaling the yield by a factor of 1.7 and adding a 60 % theoretical uncertainty. The cross sections for the ZZ MC are given in Table D.4. Further description of the $k-$factors for the ZZ process are provided in section 6.1.2.

The remaining diboson processes, WW and WZ, are simulated with Powheg-Box at NLO in QCD with the CT10nlo PDF set and Pythia8 generator. The WZ sample considers the $WZ \to \ell\nu\ell^+\ell^-$ decay chain while the WW background accounts the $WW \to \ell\nu\ell\nu$ process. In the VBF category, a Sherpa sample is used due to the better modelling of jets. The cross section for each process is listed in Table D.5.

The $t\bar{t}$, single-t and Wt backgrounds are modelled with the Powheg-Box v2.0 using the CT10 PDF. The PS simulation is handled by Pythia v6.428 [170, 171] with the Perugia set of tuned parameters [172]. The event properties of charm and bottom decays are modelled by EvtGen 1.2.0 [173]. The mass of the top quark is set to be 172.5 GeV.

The Z + jets process was simulated using Sherpa, with the NNPDFv3.0 NNLO PDF set [174] with dedicated PS tuning by the Sherpa authors. The Z + jets is normalized to cross section calculated to NNLO in α_s [175]. The samples are sliced by phase space cuts defined by the maximum between the H_T and the transverse momentum of the Z boson in order to generate sufficient statistics.

Other backgrounds which are included are W + jets simulated with Powheg + Pythia8 , ttW and ttZ produced from MG5_aMC@NLO + Pythia8, and triboson ($VVV, V = W, Z$) backgrounds simulated with Sherpa [108].

5.4 Modelling

The modelling of kinematic variables (analysis observables) is checked before any $H \to ZZ \to \ell\ell\nu\nu$ selections are applied, except for the $|m_{\ell\ell} - m_{\mathrm{Z}}| < 15$ GeV window cut. This is a standard procedure to ensure simulations provide a good description of the backgrounds in the wider phase space. An important note, which must be emphasised, is that these plots do **not** reflect the behaviour in the signal regions e.g. the Z + jets process before any specific analysis requirements has $\sim 10^7$ events, while after all $H \to ZZ \to \ell\ell\nu\nu$ cuts, it makes up only $\sim 6\%$ of the total background.

Process	Generators	PDF	Tune
$gg \to H \to ZZ \to \ell\ell\nu\nu$	Powheg + Pythia8	CT10	AZNLO CTEQ6L1
$qq \to qqH \to qq\ell\ell\nu\nu$	Powheg + Pythia8	CT10	AZNLO CTEQ6L1
$G \to ZZ \to \ell\ell\nu\nu$	MG5_aMC@NLO + Pythia8	NNPDF23LO	A14 NNPDF23LO
$qq \to ZZ \to \ell\ell\nu\nu$	Powheg + Pythia8	CT10nloME	AZNLO CTEQ6L1
$gg \to ZZ \to \ell\ell\nu\nu$	gg2VV + Pow Pythia8	CT10	A14 NNPDF23LO
$WZ \to \ell\nu\ell\ell$	Powheg + Pythia8	CT10nloME	AZNLO CTEQ6L1
$WW \to \ell\nu\ell\nu$	Powheg + Pythia8	CT10nloME	AZNLO CTEQ6L1
Z + jets	Sherpa	NNPDF30NNLO	—
$t\bar{t}$, single-t, Wt	Powheg + Pythia6	CT10	Perugia 2012
W + jets	Powheg + Pythia8	CT10nloME	AZNLO CTEQ6L1
$ttV(V = W, Z)$	MG5_aMC@NLO + Pythia8	NNPDF23LO	A14 NNPDF23LO
$VVV(V = W, Z)$	Sherpa	CT10	—

Table 5.7: Summary of the signal+background samples of used in the $H \to ZZ \to \ell\ell\nu\nu$ analysis.

The data/MC comparison of kinematical variables is given in section 5.4.1. The observables which are used in the VBF search are provided in section 5.4.2.

5.4.1 $H \to ZZ \to \ell\ell\nu\nu$ analysis observables

The set of observables which comprise the $H \to ZZ \to \ell\ell\nu\nu$ analysis were outlined in section 5.2.4. In this section the level of agreement between data and simulation is checked.

The major backgrounds after the Z mass requirement are the Z + jets, top and diboson (WZ, WW, ZZ) processes. Single top, Wt, W + jets and triboson backgrounds are combined into a category called "Other". Two signal samples, $gg \to H$ and $qq \to qqH$ at $m_{\mathrm{H}} =$ 600 GeV, are also added, for comparison purposes.

In Figure 5.3 the invariant mass of the lepton pair, $m_{\ell\ell}$, the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, the ΔR distance between the two leptons, the $\Delta\phi$ separation between the Z boson and $E_{\mathrm{T}}^{\mathrm{miss}}$, fractional p_{T} difference and minimum $\Delta\phi$ separation between any jet with p_{T} $>$ 100 GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$ are given for the $H \to ZZ \to ee\nu\nu$ channel. The corresponding kinematic plots for the $H \to ZZ \to \mu\mu\nu\nu$ channel are shown in Figure 5.4. The $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ and number of $b-$jets, after the Z mass window cut, are presented in Figure 5.5 for both channels.

The total integrated luminosity, $\mathcal{L}$, corresponds to $\sim$ 36 fb^{-1} collected in the 2015+2016 data run periods at the LHC. The MC samples are normalized to the cross section, which are provided in section 5.3, and scaled to $\mathcal{L}$. The shaded band indicates the statistical and experimental detector systematic uncertainties on the MC expectation. The largest

contribution to the uncertainty band is comprised of the JES, $E_{\mathrm{T}}^{\mathrm{miss}}$ and PU systematics.

An overall good agreement between data and simulation is observed for the event observables, with any discrepancies covered by the available uncertainty.

5.4.1.1 Modelling of $p_{\mathrm{T}}^{\mathrm{Z}}$ and $m_{\mathrm{T}}^{\mathrm{ZZ}}$

The transverse momentum of the Z boson, $p_{\mathrm{T}}^{\mathrm{Z}}$, is not correctly modelled in the inclusive distributions. This results in improper modelling of the variables related to the $p_{\mathrm{T}}^{\mathrm{Z}}$, such as the $\Delta R_{\ell\ell}$ and $m_{\mathrm{T}}^{\mathrm{ZZ}}$. The mis-modelling of $p_{\mathrm{T}}^{\mathrm{Z}}$ is fixed by implementing binned correction factors to the Z + jets distribution. The correction is derived as the ratio of data (subtracted with non-Z backgrounds) to Z + jets MC. The correction factors are shown in Figure E.1, which then corrects the $p_{\mathrm{T}}^{\mathrm{Z}}$ as presented in Figure E.2.

The distribution of the transverse mass, as defined in equation 5.1, is shown in Figure 5.6 after the Z mass cut and after the correction to the $p_{\mathrm{T}}^{\mathrm{Z}}$. A larger tail is present in the $H \to ZZ \to \mu\mu\nu\nu$ channel due to the degradation in resolution for high p_{T} muons.

5.4.2 VBF-related observables

The VBF analysis selects on observables sensitive to the jet kinematics in the event. The VBF requirements are placed on top of the $H \to ZZ \to \ell\ell\nu\nu$ selections. As stated in section 5.4.1, it still useful to study the performance of the kinematic variables before any conditions are placed (other than the Z mass window). The modelling of the number of jets, N_{jets}, leading jet transverse momentum, di-jet mass, m_{jj}, and di-jet pseudorapidity, $|\Delta\eta_{jj}|$, is presented in Figure 5.7 after the Z mass window selection in the $H \to ZZ \to ee\nu\nu$ channel. The corresponding plot for the $H \to ZZ \to \mu\mu\nu\nu$ channel is presented in Figure 5.8.

Overall the agreement between data and MC is acceptable with most discrepancies covered by the experimental systematic uncertainties. The N_{jets} is seen to agree well for jet multiplicities ≤ 6, consistent with the NNLO prediction of `Sherpa`. The kinematics of the leading p_{T} jet are predicted to NLO in QCD. However, as observed in the kinematic distributions, certain reconstruction effects were found to cause a mis-modelling in di-jet system kinematics. This is the focus of section 5.4.2.2.

5.4.2.1 Modelling of di-jet kinematics

In the variables related to the di-jet kinematics it is evident that disagreements exist between data and simulation. They are covered by the JES uncertainties, however certain peculiarities

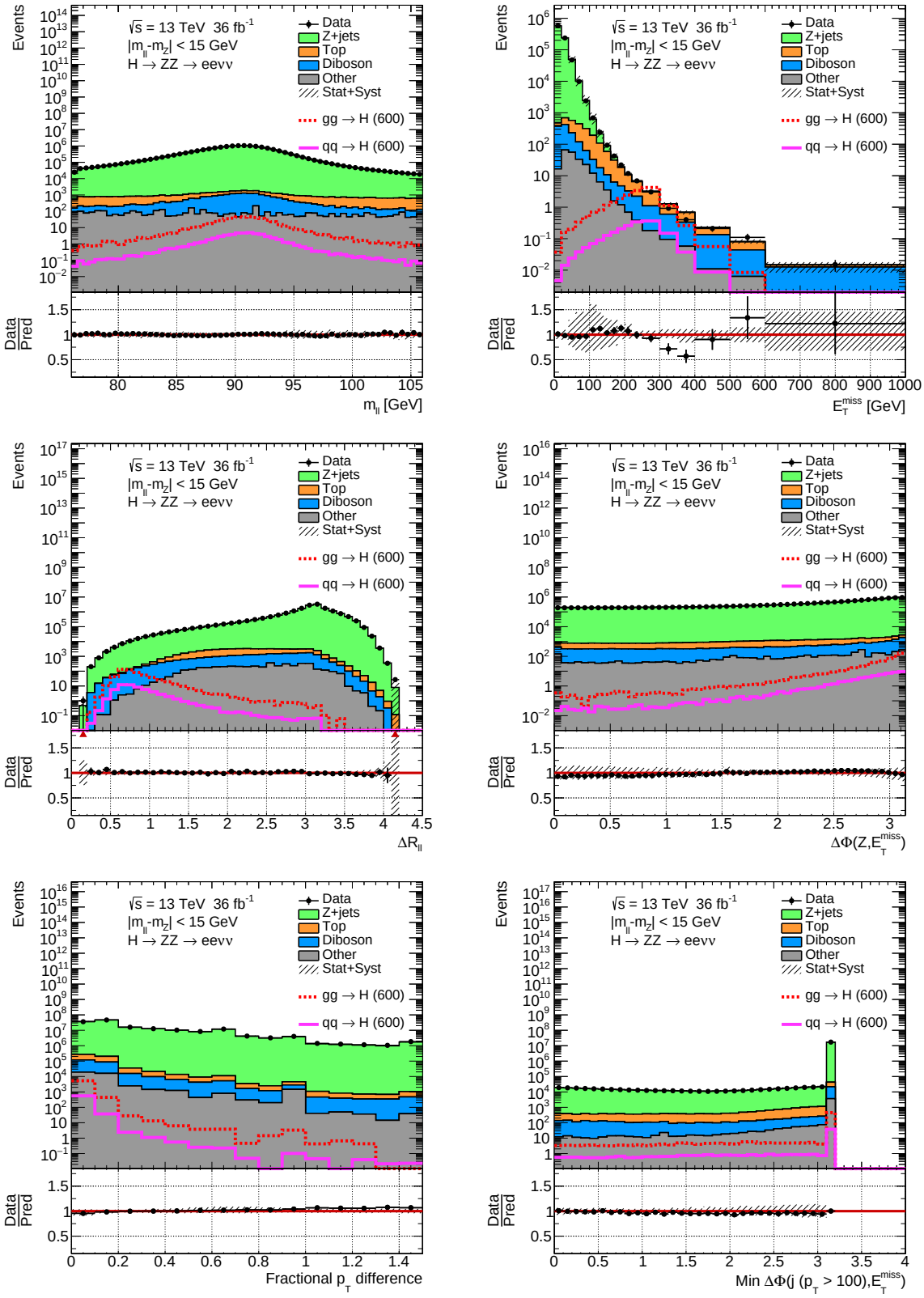


Figure 5.3: Kinematic distributions that are used for event selections in the $H \to ZZ \to ee\nu\nu$ channel after applying the Z mass window cut: (top left) Z mass, (top right) $E_{\mathrm{T}}^{\mathrm{miss}}$, (center left) the distance between leptons, $\Delta R_{\ell\ell}$, (center right) $\Delta\Phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\mathrm{miss}})$,(bottom left) frac.$p_{\mathrm{T}}$ diff.,(bottom right) min $\Delta\Phi(p_{\mathrm{T}} > 100$ GeV$, E_{\mathrm{T}}^{\mathrm{miss}})$. The ratio plot is between data and the total combined background MC expectation. MC samples are normalized to their cross section values and re-scaled to the data integrated luminosity. The error band in the ratio plot shows the statistical and experimental systematic uncertainty on the MC expectation.

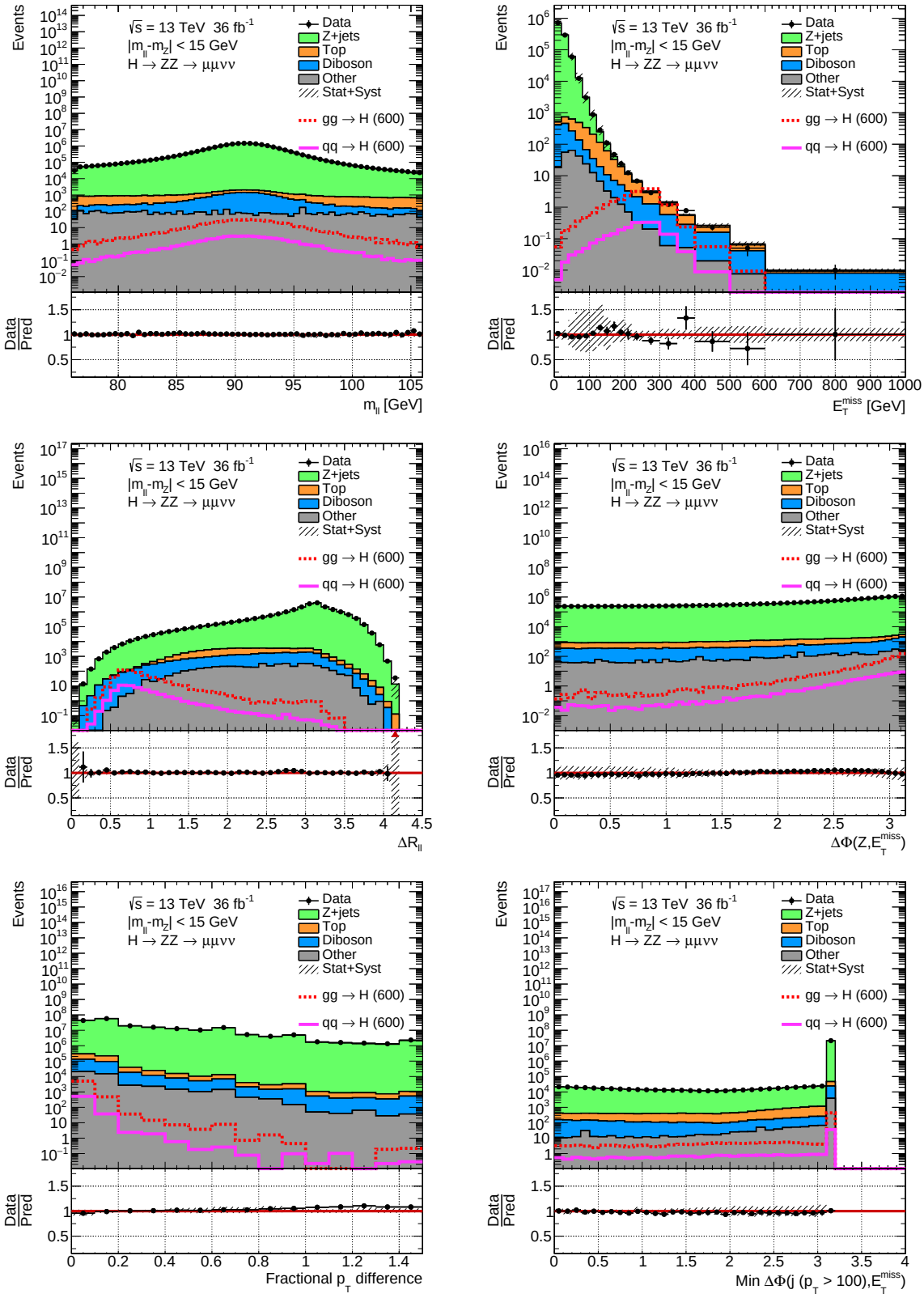


Figure 5.4: Kinematic distributions that are used for event selections in the $H \to ZZ \to \mu\mu\nu\nu$ channel after applying the Z mass window cut: (top left) Z mass, (top right) $E_{\mathrm{T}}^{\mathrm{miss}}$, (center left) the distance between leptons, $\Delta R_{\ell\ell}$, (center right) $\Delta\Phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\mathrm{miss}})$,(bottom left) frac.$p_{\mathrm{T}}$ diff.,(bottom right) min $\Delta\Phi(p_{\mathrm{T}} > 100$ GeV$, E_{\mathrm{T}}^{\mathrm{miss}})$. The ratio plot is between data and the prediction from the total background MC expectation. MC samples are normalized to their cross section values and re-scaled to the data integrated luminosity. The error band in the ratio plot shows the statistical and experimental systematic uncertainty on the MC expectation.

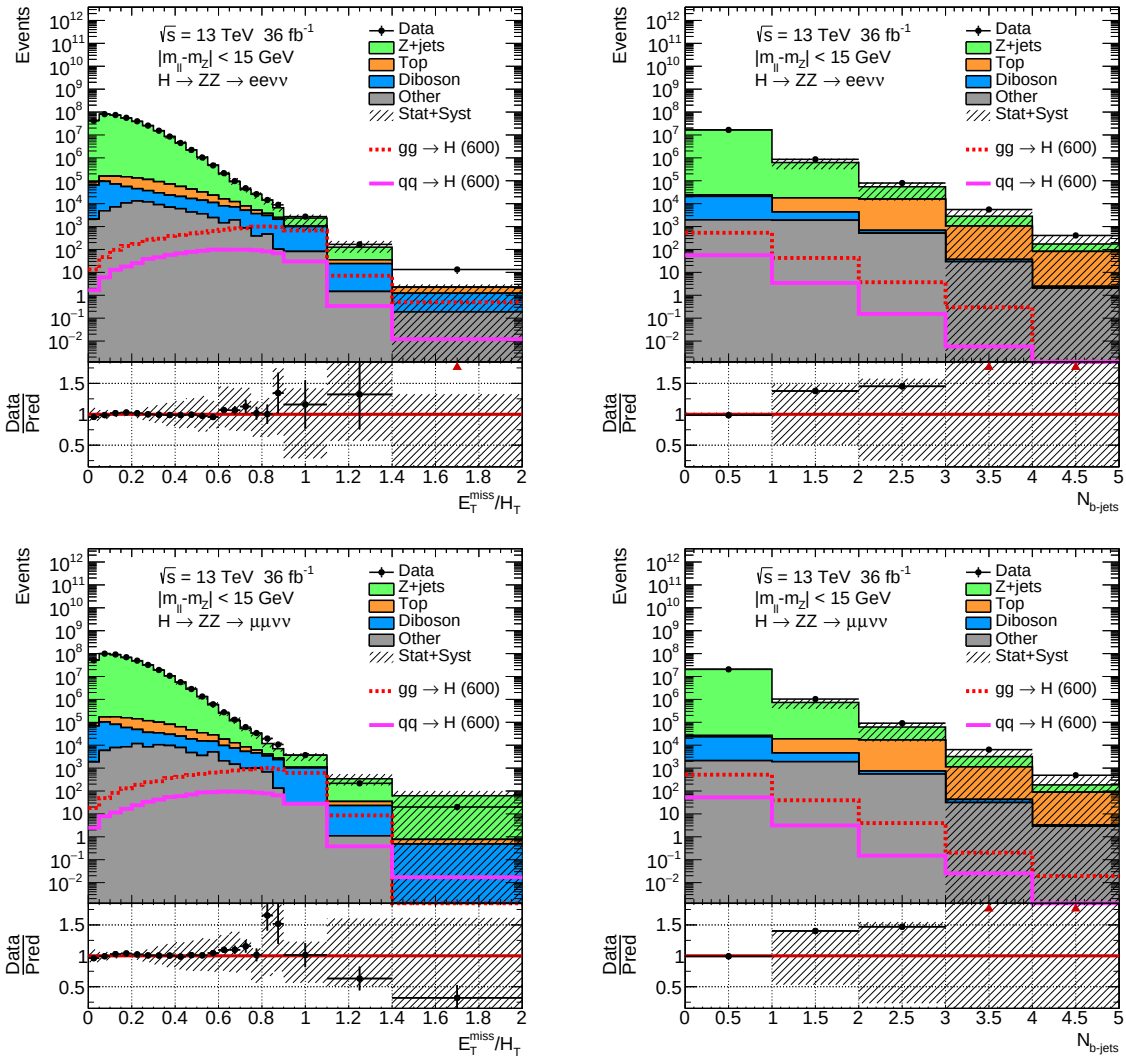


Figure 5.5: Kinematic distributions that are used for event selections in the $H \to ZZ \to ee\nu\nu$ (top) and $H \to ZZ \to \mu\mu\nu\nu$ (bottom) channels after applying the Z mass window cut in the $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ and Number of $b-$jets distribution. The ratio plot is between data and the prediction from the total background MC expectation. MC samples are normalized to their cross section values and re-scaled to the data integrated luminosity. The error band in the ratio plot shows the statistical and experimental systematic uncertainty on the MC expectation.

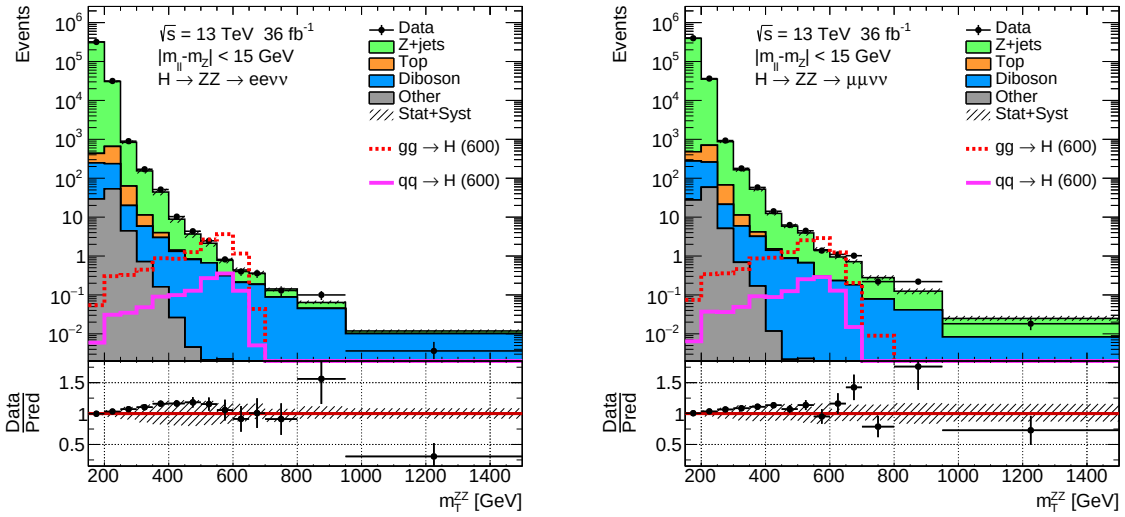


Figure 5.6: $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution in the $H \to ZZ \to ee\nu\nu$ (left) and $H \to ZZ \to \mu\mu\nu\nu$ (right) channels after applying the Z mass window cut.

can be observed in the ratio plots in Figures 5.7 and 5.8 in the $|\Delta\eta_{jj}|$ distributions ($5.5 < |\Delta\eta_{jj}| < 6.0$). The origin of this effect was not entirely understood at the time of the analysis and no official recommendation was provided on how to treat it. A suggestion to use the bunch crossing number in order to reject events from first few bunch train crossings was considered but was not pursued since it would remove more than 10 % of the available data. Studies were carried out on how to better suppress their contribution.

The jets in $5.5 < |\Delta\eta_{jj}| < 6.0$ were isolated and found to largely cluster in the low jet p_{T} spectrum ($\sim$ 20 GeV) and in a specific part of the detector region ($|\eta_j| \sim 2.9$). This motivated the addition of a selection on the p_{T} of jets, which survive the inclusive and $N_{\mathrm{jets}} \geq 2$ selections. The p_{T} cut-value was chosen to be 30 GeV after an optimization procedure was carried out, which is described in section 5.4.2.2.

5.4.2.2 Jet p_{T} cut-value optimization

The optimization procedure for the p_{T} requirement on jets is similar in nature to the method described in section 5.2.4.1. The $H \to ZZ \to \ell\ell\nu\nu$ selections were first applied and the total background included the contribution of the ggF signal. The significance, Z, scan for different jet p_{T} cut-values is shown in Figure 5.9. The best cut-value to select jets is 30 GeV as it preserves the greatest sensitivity to the VBF signal while excluding more of the problematic jets. The rise of the significance for momenta around 40-50 GeV (60-70 GeV) in the $H \to ZZ \to ee\nu\nu$ ($H \to ZZ \to \mu\mu\nu\nu$) channel was found to be caused by negative weights in the `Sherpa` Z + jets MC which lower the overall background yield, giving the

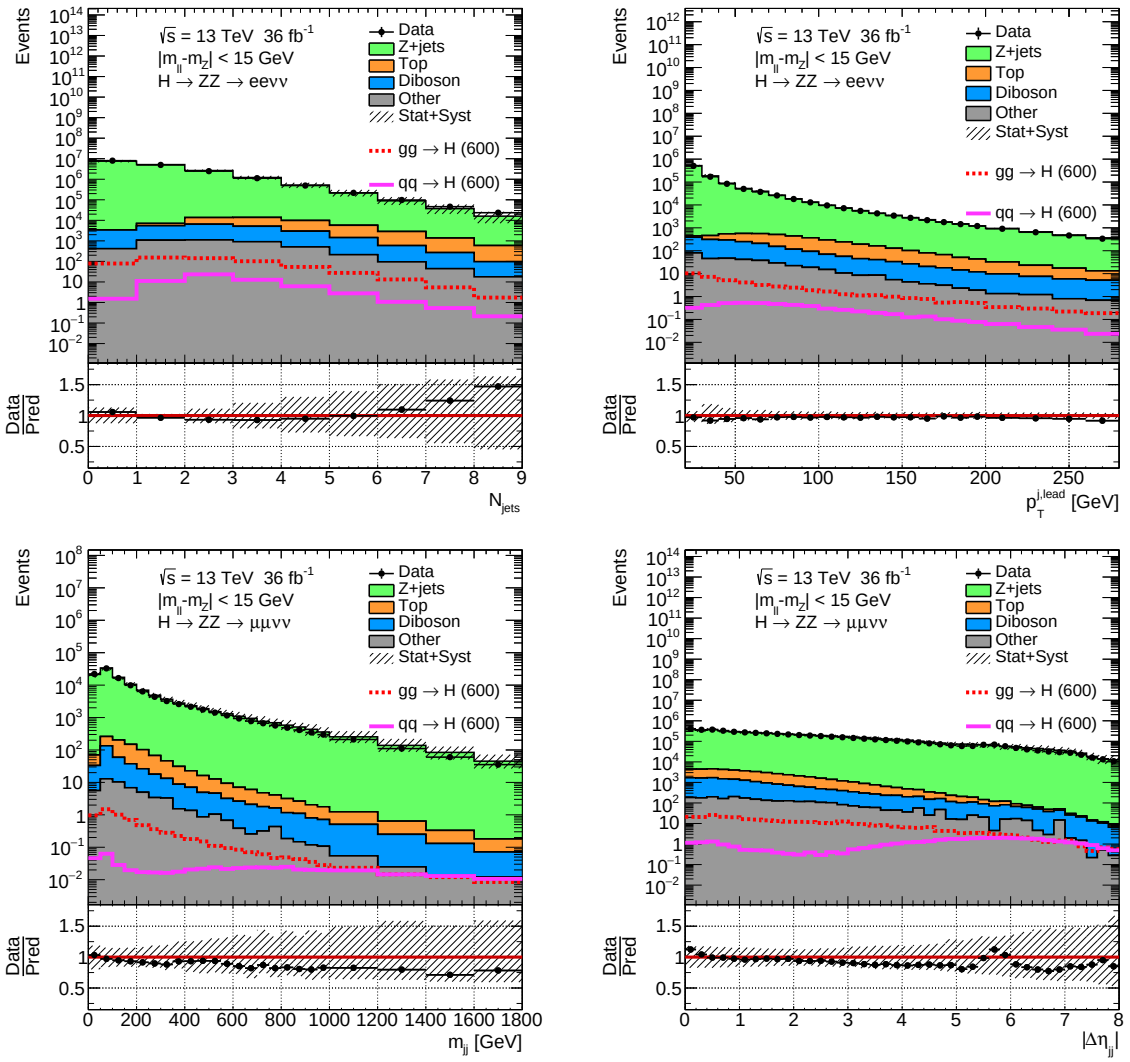


Figure 5.7: Kinematic distributions that are used for event selections in the $H \to ZZ \to ee\nu\nu$ channel after applying Z mass window cut:(top left) N_{jets}, (top right) $p_{T,\text{jet}}^{\text{lead}}$, (bottom left) m_{jj} and (bottom right) $|\Delta\eta_{jj}|$. The bottom plot shows the ratio between data and predicted total background expectation. MC samples are normalized to their cross section values and re-scaled to the data integrated luminosity. The error band in the ratio plot shows the statistical and experimental systematic uncertainty on the MC expectation.

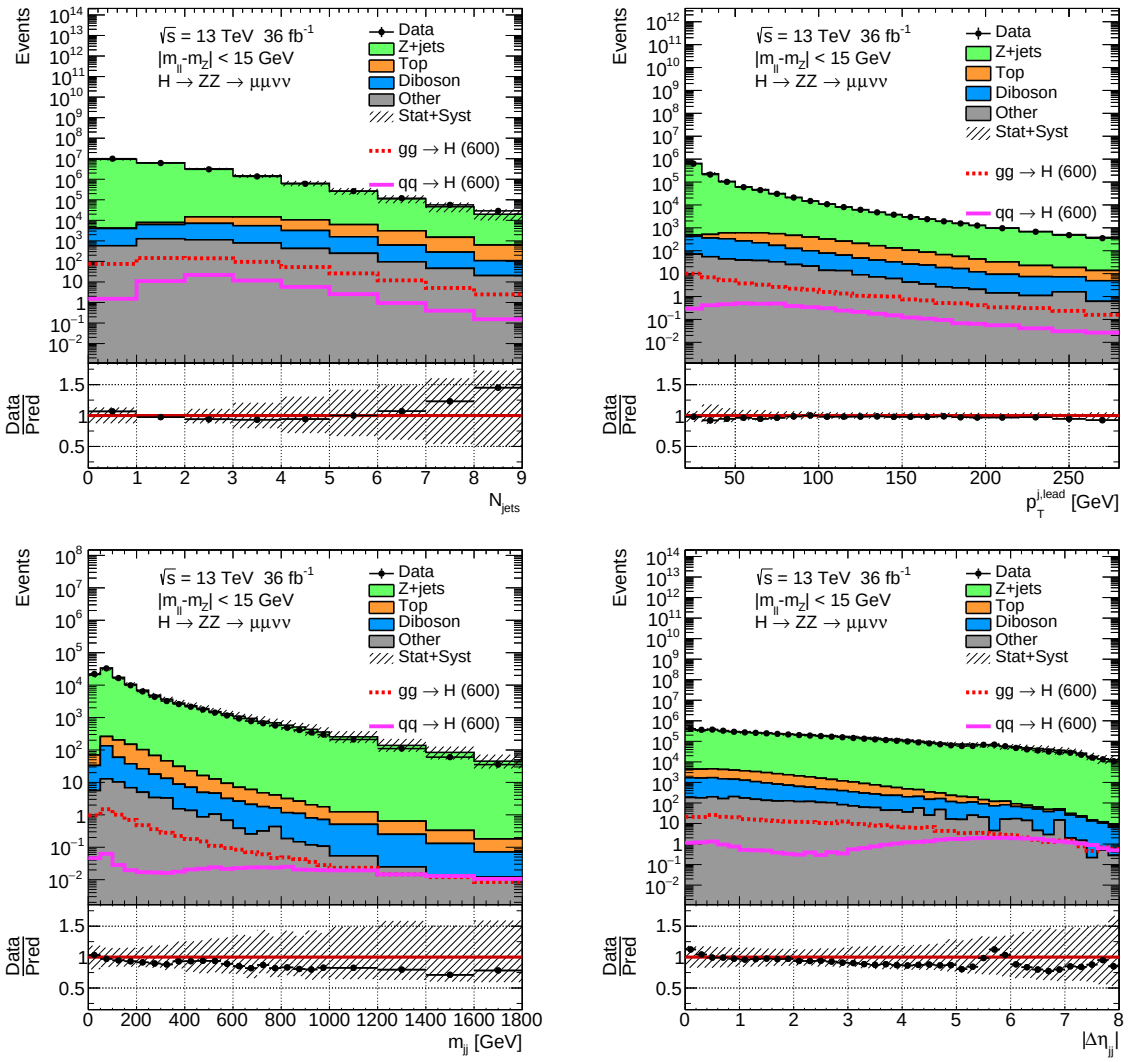


Figure 5.8: Kinematic distributions that are used for event selections in the $H \to ZZ \to \mu\mu\nu\nu$ channel after applying Z mass window cut:(top left) N_{jets}, (top right) $p_{T,\text{jet}}^{\text{lead}}$, (bottom left) m_{jj} and (bottom right) $|\Delta\eta_{jj}|$. The bottom plot shows the ratio between data and predicted total background expectation. MC samples are normalized to their cross section values and re-scaled to the data integrated luminosity. The error band in the ratio plot shows the statistical and experimental systematic uncertainty on the MC expectation.

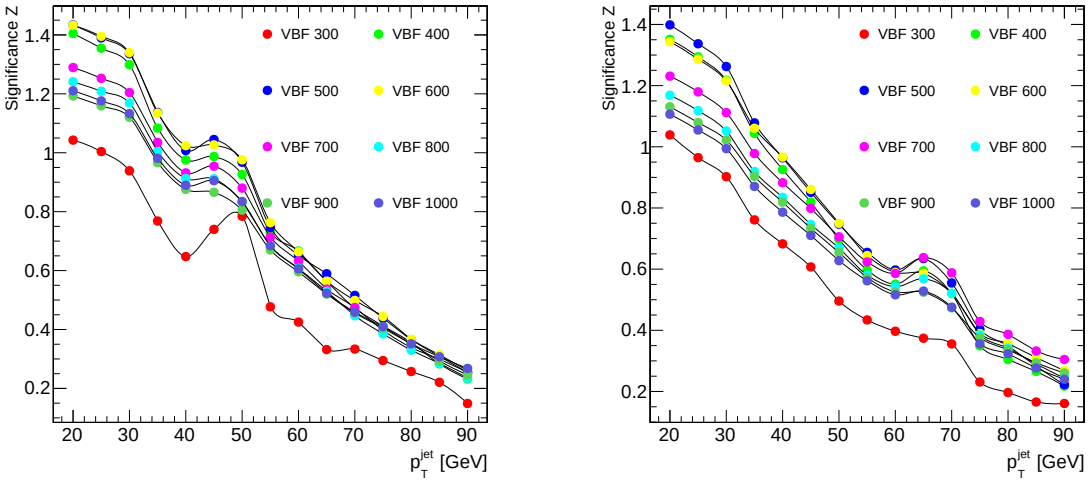


Figure 5.9: Significance scan of the jet p_{T} after all inclusive selections for a range of signal mass points. The left plot is for the $H \to ZZ \to ee\nu\nu$ case while the right is for the $H \to ZZ \to \mu\mu\nu\nu$ case. Sensitivity is lowered at $p_{\mathrm{T}} = 30\,\mathrm{GeV}$, but with a greater reduction in the data/MC mis-modelling.

appearance of an increase in sensitivity.

The modelling in data improved as a result of the introduction of the new requirement. Figure G.1 shows the di-jet kinematic distributions after selecting jets with a higher p_{T} cut.

5.4.2.3 Removal of the PU jets

The cause of the mis-modelling was later understood to be caused by saturation in the readout electronics of the inner wheel of the LAr end-cap calorimeter in the region $2.5 < |\eta| < 3.2$. The saturation originated at the end of the first bunch train[1] and the next few crossings from the next train where the electronics is recovering. Two effects occur as a result, the first being a lower response for the first two bunch crossings, which was found to have negligible effect in jets and the second effect resulting in a large noise burst, which gives many PU jets in a small fraction of the data.

These effects are not simulated in MC. This can be clearly observed in the η distribution of the sub-leading in p_{T} jets in Figure 5.10. In the forward regions the JVT is not available to reject them.

Studies were carried out on cluster level parameters, which examined their behaviour in the early and later stages of the bunch crossing collisions. A few key variables were chosen

[1] The proton beam is made from separate packed bunches of protons. The different bunches are arranged in bunch trains with 25 ns spacing.

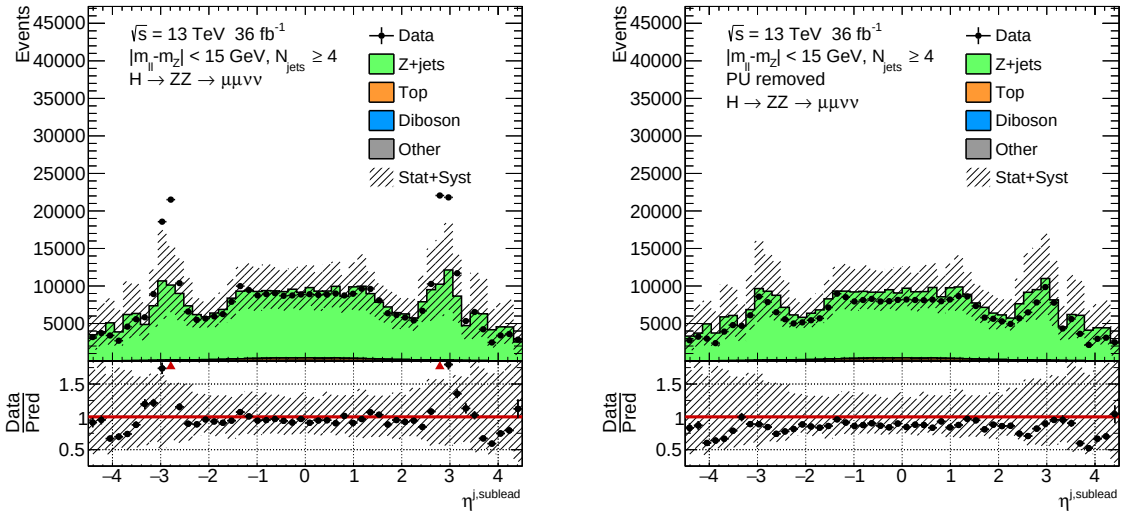


Figure 5.10: Left: The sub-leading jet pseudorapidity for events with a Z mass selection and $N_{\text{jets}} > 4$ with PU jets in the $H \to ZZ \to \mu\mu\nu\nu$ channel. Right: The sub-leading jet pseudo-rapidity for events with a Z mass selection and $N_{\text{jets}} > 4$ with removed PU jets in the $H \to ZZ \to \mu\mu\nu\nu$ channel.

to suppress the PU jets. Among these are the LAr cluster EM energy, $E^{\text{EM}}_{\text{T,LAr}}$, and quality, $\langle Q_{\text{LAr}} \rangle$, which is defined as the energy-squared weighted average of the LAr cell quality [132]. The proposed selections reject events with more than 20 clusters which have $2.5 < |\eta| < 3.2$, $E^{\text{EM}}_{\text{T,LAr}} > 500$ MeV and $\langle Q_{\text{LAr}} \rangle > 0.2$. The consequence of these selections on the jet pseudorapidity are viewed in Figure 5.10. The effect of the removal of the PU jets in the di-jet system for the $H \to ZZ \to ee\nu\nu$ channel are given in Figure 5.11. The plots are shown for m_{jj} and $|\Delta\eta_{jj}|$ before and after PU jet reduction. The modelling between data and simulation is observed to have a greater compatibility for the di-jet invariant mass and pseudorapidity after the removal of the PU events in data. This conclusion is also reached based on the same plots in the $H \to ZZ \to \mu\mu\nu\nu$ channel. Figure G.2 presents the events which are flagged as due to PU jets. The impact of these new requirements is a reduction of only 1.4 % (1.7%) for data in the $H \to ZZ \to ee\nu\nu$ ($H \to ZZ \to \mu\mu\nu\nu$) final state after the Z mass window cut with respect to the nominal selections. The events in data, in the VBF signal region, were checked and no differences were found.

Unfortunately the proper treatment of these jets came after the analysis was unblinded and finalised so the incorporation of these selections is going to happen for future analyses. The selections are already applied in the new production of xAODs and the detector setup was slightly changed to avoid the appearance of these jets in 2017 data.

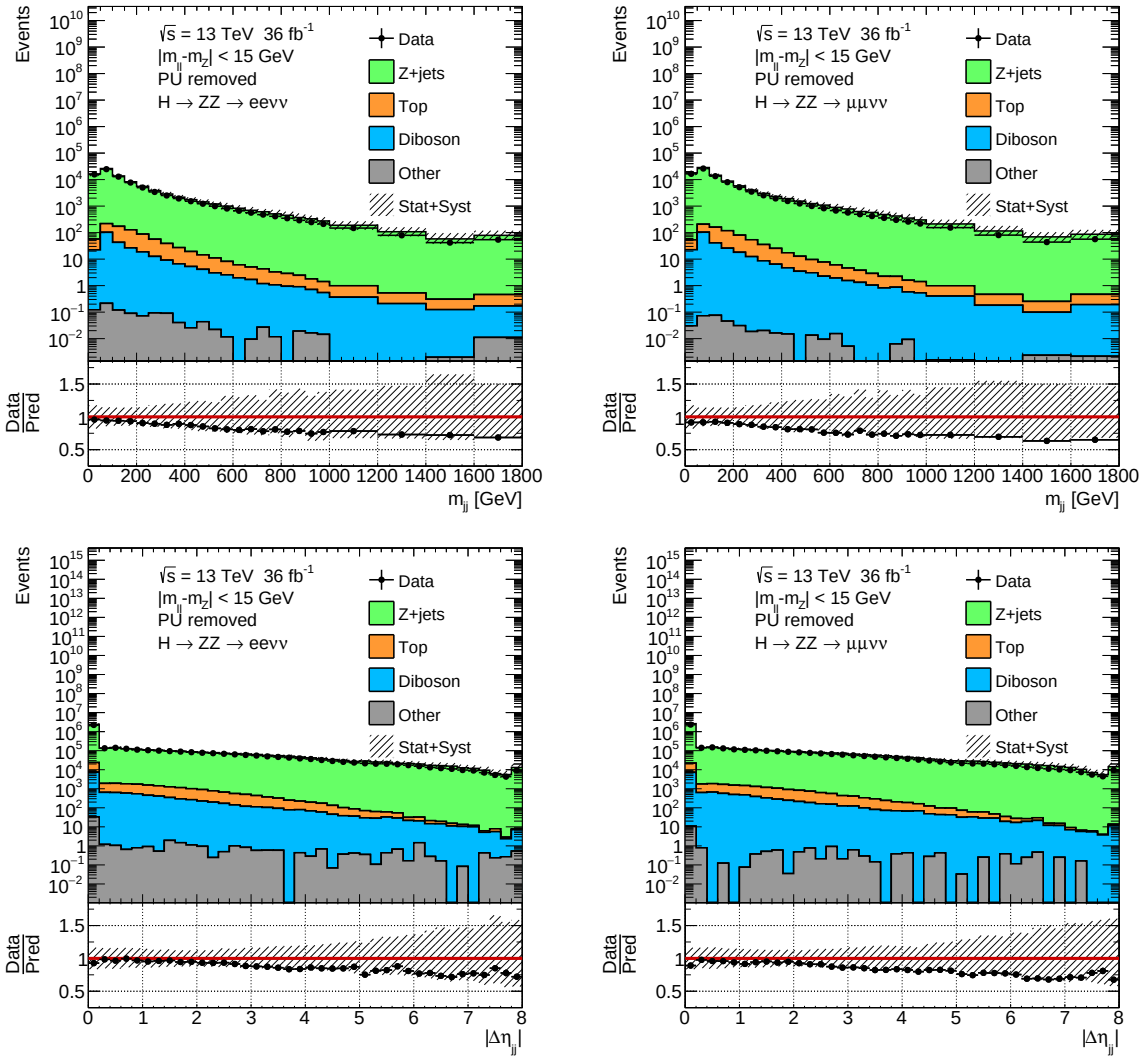


Figure 5.11: The removal of the PU jets in the $Z \to ee$ channel, following the selections released by the ATLAS Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ group. The overall reduction in data combined from the $ee\nu\nu$ and $\mu\mu\nu\nu$ channels is $\sim 3\%$.

5.4.3 Modelling of $m_{\mathrm{T}}^{\mathrm{ZZ}}$ in the $H \to ZZ \to \ell\ell\nu\nu$ signal region

The comparison between data and MC for the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution was given after the Z mass window selection in Figure 5.6. Since $m_{\mathrm{T}}^{\mathrm{ZZ}}$ is the discriminant observable, it is worthwhile to examine its behaviour in the $H \to ZZ \to \ell\ell\nu\nu$ signal region which is formed from the analysis selections in Table 5.4.

Figure 5.12 demonstrates the modelling of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution in signal region for data, all available backgrounds and two signal samples. A few subtleties are present in these plots:

- The data distribution is normally not shown, to prevent biases in analysis studies
- The MC distributions are entirely estimated from simulation, they have been only

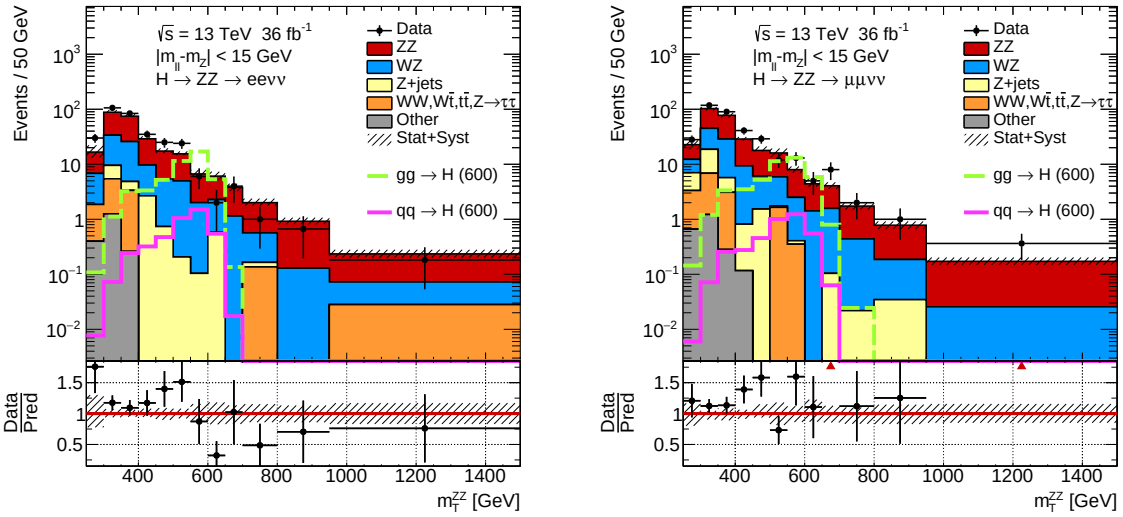


Figure 5.12: Transverse mass distribution in the signal region before any data-driven corrections are applied to the background, left: $H \to ZZ \to ee\nu\nu$ channel, right: $H \to ZZ \to \mu\mu\nu\nu$ channel. This figure is to be used as reference guide for the changes that will caused to the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distributions by the background estimation.

scaled to 36 fb^{-1}

- Diboson backgrounds are split to ZZ and WZ. The WW background is merged with single-t, Wt and $t\bar{t}$. Other backgrounds include W + jets and triboson.

Figure 5.12 is meant as reference for the changes that will be caused by the data-driven corrections on the backgrounds and their impact on the transverse mass. This figure presents **only** the backgrounds derived purely by MC . The limits on the production of Higgs bosons are obtained from the distributions provided in Figure 7.2 in section 7.1.

One thing to note in Figure 5.12 is the small excess of events in data found in the region between 450 and 500 GeV. It is found to be consistent for both lepton channels. As will be seen later (see Figure 7.2), the excess is lowered because of the increase of the WZ and Z + jets yields after the data-driven correction.

The VBF category lacks statistics for a proper description of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape, which limits the analysis to a "cut-and-count" study, where only the number of events left after all selections are considered. The comparison between MC-driven and data-driven background estimations are provided in Table 7.2. The MC yield is comparable to the result obtain data-driven corrections.

Chapter 6

Background estimation and systematic uncertainties

The methods for estimating the background contributions in the signal regions are summarized in section 6.1. The sources of theoretical and experimental systematic uncertainties are discussed in section 6.2.

6.1 Background estimation

The backgrounds in the $H \to ZZ \to \ell\ell\nu\nu$ analysis are determined from a mix of MC and data-driven techniques. Each background estimation is succinctly summarized in section 6.1 and each contribution to the total background prediction is given in Table 6.1.

6.1.1 Overview

- The $ZZ \to \ell^+\ell^-\nu\bar{\nu}$ process forms the largest contribution to the total background in the ggF and VBF signal regions. The normalization and shape of the ZZ background transverse mass distribution is obtained purely from simulation. Corrections from higher order processes are taken from theoretical calculations [176].

- The production of $WZ \to \ell\nu\ell^+\ell^-$ events has the second largest background contribution. One of the leptons is mis-reconstructed, thereby 'faking' the di-lepton final state. The normalization of the WZ background is taken from a three-lepton control region while shape is extracted from MC.

- The $t\bar{t}, Wt, WW$ and $Z \to \tau\tau$ backgrounds are determined in a control region enriched in $e\mu$ events. Both the background shape and normalization are estimated in data. Difference in reconstruction efficiencies between electrons and muons is also accounted.

Background	m_{T}^{ZZ} Estimation	Control region	ggF %	VBF %
$ZZ \to \ell^+\ell^-\nu\bar{\nu}$	MC (yield+shape*)	N/A	$\sim 58\%$	$\sim 45\%$
$WZ \to \ell\nu\ell^+\ell^-$	Data (yield), MC (shape*)	3ℓ	$\sim 32\%$	$\sim 29\%$
Z + jets	Data (yield), MC$\to$data (shape*)	ABCD	$\sim 6\%$	$\sim 17\%$
$WW/Wt/t\bar{t}/Z \to \tau\tau$	Data (yield), MC$\to$data (shape*)	$e\mu$	$\sim 3\%$	$\sim 9\%$
W + jets	Data (yield), MC (shape*)	Fake Factor	$< 1\%$	$\sim 0\%$
Other	MC (yield+shape*)	N/A	$< 1\%$	$< 1\%$

Table 6.1: Expected background contributions to the total background, their estimation and their corresponding control region. * For VBF category, only a single bin fit is used.

- The Z + jets background uses a data-driven "ABCD" method, where control regions are formed from a pair of mostly uncorrelated boolean variables, to obtain the yield in the signal region. The shape of the Z + jets m_{T}^{ZZ} distribution is obtained from a one-sideband method.

- The W + jets process can mimic the di-lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ signature through jets which fake leptons. The yield is extracted from a control region enriched in jets. The shape is taken from MC

- Other backgrounds, include ttW, ttZ and tri-boson backgrounds, which contribute very little to the final states and are estimated solely from simulation.

- In the VBF category the yields are extrapolated through "loose" VBF control regions. The shape is not considered since statistics for a proper shape description is insufficient.

6.1.2 ZZ background

The non-resonant continuum ZZ process is an "irreducible" background as it contains the same final state signature and kinematics as the signal. Therefore a precise set of cuts to reduce its contribution is difficult to obtain. Data measurements of ZZ backgrounds by ATLAS and CMS suffer from large experimental uncertainties [177]. A better prediction of the ZZ process is obtained through theoretical calculations [176]. The measurement of the ZZ cross section in run I is given in [178].

The SM ZZ event originates from two sources. The first is quark-quark interactions, as shown for tree level in Figure 6.1. The second originates from loop-induced gluon-gluon interactions with QCD corrections which increase the production cross section by $\mathcal{O}(50 - 100\%)$ depending on the scale [165].

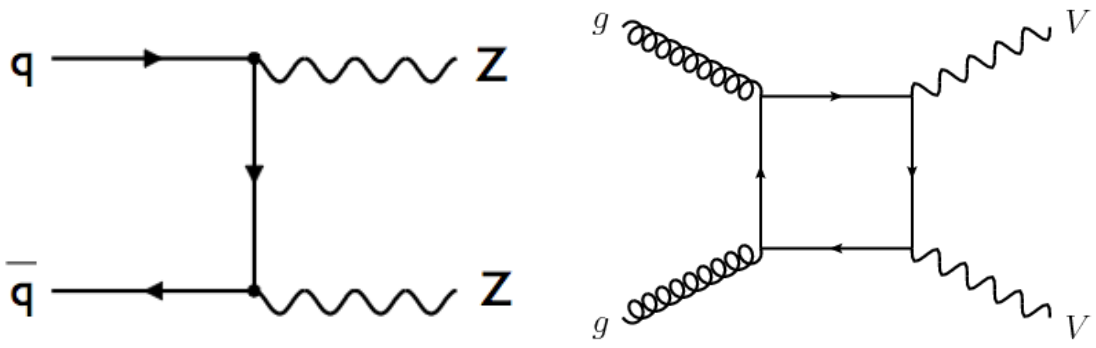


Figure 6.1: Left plot: tree-level diagram for the production of the irreducible $qq \to ZZ$ background. Right plot: loop-induced diagram for the ZZ process through gluon-gluon interaction.

6.1.2.1 Estimation in the ggF category

6.1.2.1.1 ***qq*** $\to$ ***ZZ*** Quark-quark interaction to two Z bosons is simulated by the `Powheg` + `Pythia8` generator [107], which models QCD corrections to NLO and the EW corrections to LO. PS and final state radiation (FSR) are obtained from `Pythia8` [109]. Higher-order processes, which modify the production cross section, are incorporated through a k-factor correction [179–181]. The $k-$factor is calculated for NNLO QCD and NLO EW interactions.

The NNLO QCD $k-$factor is defined as:

$$k = \frac{\sigma_{\mathrm{NNLO}}(pp \to ZZ) - \sigma_{\mathrm{LO}}(gg \to ZZ)}{\sigma_{\mathrm{NLO}}(pp \to ZZ)} \tag{6.1}$$

where $\sigma_{\mathrm{NNLO}}(pp \to ZZ)$ ($\sigma_{\mathrm{NLO}}(pp \to ZZ)$) is the calculated cross section for the $pp \to ZZ$ process at NNLO (NLO) and $\sigma_{\mathrm{LO}}(gg \to ZZ)$ is the cross section for the $gg \to ZZ$ process at LO. The $gg \to ZZ$ cross section is subtracted in the $k-$factor definition since this process is separately estimated. The correction evolves with m_{ZZ} and therefore is applied on an event-per-event basis as a function of m_{ZZ}.

The NNLO QCD k-factor is obtained in a phase space region where only a Z mass window requirement is applied. The k-factor is shown in Figure 6.2 as a function of m_{ZZ}. Two different versions of the $k-$factor are presented, the "old" k-factor was calculated for the $ZZ \to 4\ell$ final state and extrapolated to the $ZZ \to \ell\ell\nu\nu$ phase space while the "new" k-factor is explicitly calculated for a $ZZ \to \ell\ell\nu\nu$ final state [176]. Another important difference is the optimization of the mass slicing of m_{ZZ}. This was done with the aim to achieve sufficient statistical precision and to avoid artificial statistical fluctuations in the k-factor. The effects on the transverse mass distribution with the old and new k-factor corrections are shown in Figure 6.3. The "new" $k-$factor modifies the ZZ background by 10 % across the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ range.

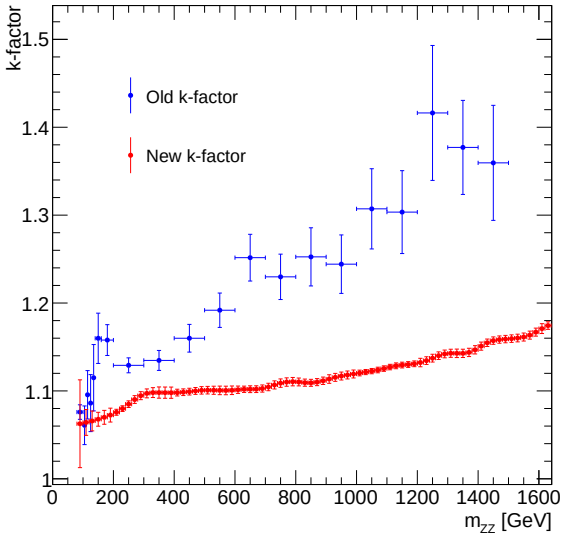


Figure 6.2: The derived NNLO QCD k-factors binned in m_{ZZ} for the $q\bar{q} \to ZZ$ process. Also shown for comparison are the old k-factors used in the 2015 analysis, which were calculated from the charged four-lepton channel [176].

The uncertainty on the acceptance from the choice of PDF is computed from the variation of 52 eigenvectors of `CT10nlo`. The uncertainty on the acceptance from the choice of renormalization and factorization scales (μ_R, μ_F) is estimated from the maximum deviation of the acceptance based on the variation of the scales. The theoretical uncertainties on ZZ are further discussed in section 6.2.2.

The NLO EW k-factor correction is calculated differentially in $m_{\mathrm{T}}^{\mathrm{ZZ}}$ with the full analysis selections applied. It corrects double counting of FSR from leptons since this has been already applied when generating the MC sample.

The yield of the $qq \to ZZ$ process in the ggF signal region is provided in Table 6.2 for both lepton modes. The statistical uncertainty is derived in MC while the systematic uncertainties originate from detector and theory uncertainties.

$qq \to ZZ$ ggF Category	
$ee\nu\nu$	161.2 ± 2.6 (stat) ± 18.6 (syst)
$\mu\mu\nu\nu$	163.8 ± 2.5 (stat) ± 18.9 (syst)

Table 6.2: Estimated $qq \to ZZ$ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and theoretical uncertainties.

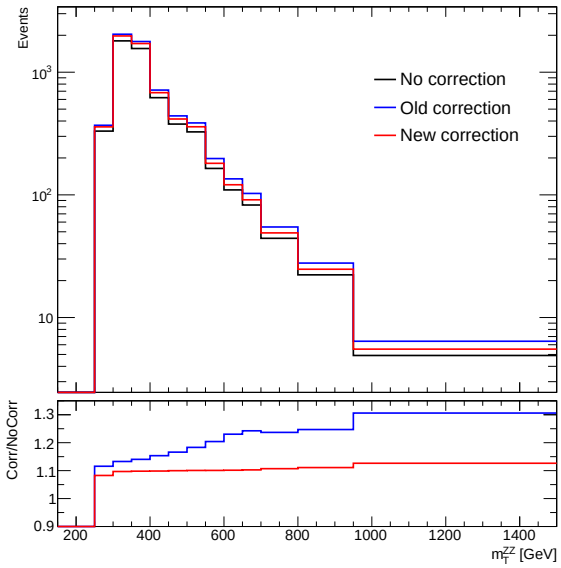


Figure 6.3: Detector-level distribution of the transverse mass $m_{\mathrm{T}}^{\mathrm{ZZ}}$ in the signal region with applied $k-$factor correction. The black histogram shows the case without a correction, the blue histogram shows results with the older k-factor correction, while the red histogram is done with the updated k-factor.

6.1.2.1.2 ***gg → ZZ*** As stated before, the $gg \to ZZ$ background, is produced through quark loops with large QCD corrections which modify the production cross section. The k-factor for higher order corrections is a single value since the differential cross section has a mild dependence on m_{ZZ}. Renormalization and factorization scales were varied between m_{Z} and $2m_{\mathrm{Z}}$ which increases the cross section from 67% to 95% [176]. Based on this, a flat k-factor of 1.7 was applied with an assigned conservative 60 % uncertainty. The $gg \to ZZ$ process is simulated through a MC sample created by the `gg2VV` generator [46, 48].

The $gg \to ZZ$ background yield is presented in Table 6.3 for both lepton modes. The statistical uncertainty is obtained from `gg2VV` while the systematic uncertainties reflects the uncertainty on the $k-$factor.

$gg \to ZZ$ ggF Category	
$ee\nu\nu$	16.0 ± 0.4 (stat) ± 9.4 (syst)
$\mu\mu\nu\nu$	15.9 ± 0.4 (stat) ± 9.4 (syst)

Table 6.3: Estimated $gg \to ZZ$ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and $k-$factor uncertainties.

6.1.2.2 Estimation in the VBF category

6.1.2.2.1 $qq \rightarrow ZZ + 2$jets The simulation and measurements of QCD and EW induced production of $qq \rightarrow ZZ + 2$jets events is described in [182]. Unfortunately, certain MC modelling problems are encountered in their description. The `Powheg` event generator, as stated before, simulates the leading p_{T} jet to LO in QCD while the kinematics of the sub-leading p_{T} jet is generated by the PS algorithm. This affects the modelling performance of the di-jet system and consequently the acceptance in the VBF region.

An alternative sample, created by a `Sherpa` generator, was initially investigated since it offers a proper description of the jet dynamics. However, this sample was unable to correctly predict the expected background contribution since negative MC weights reduced the $qq \rightarrow ZZ + 2$jets yields to unphysical levels. The `Sherpa` sample instead assesses the modelling differences with jets from the `Powheg` sample. In Figure 6.4, the modelling of the shapes of m_{jj} and $|\Delta\eta_{jj}|$ distributions are presented for each MC sample with statistical and systematic uncertainties from `Powheg` added on top. Both samples are observed to agree within the uncertainties and therefore can be further examined.

Modelling corrections on the normalization were explored, such as re-weighting of the `Powheg` N_{jets} observable to match with the corresponding distribution from the `Sherpa` sample. The yield was found to increase between 3% and 11%, which is within the `Powheg` statistical uncertainty. For simplicity, the yield provided from `Powheg` is taken without any corrections (other than the $k-$factor).

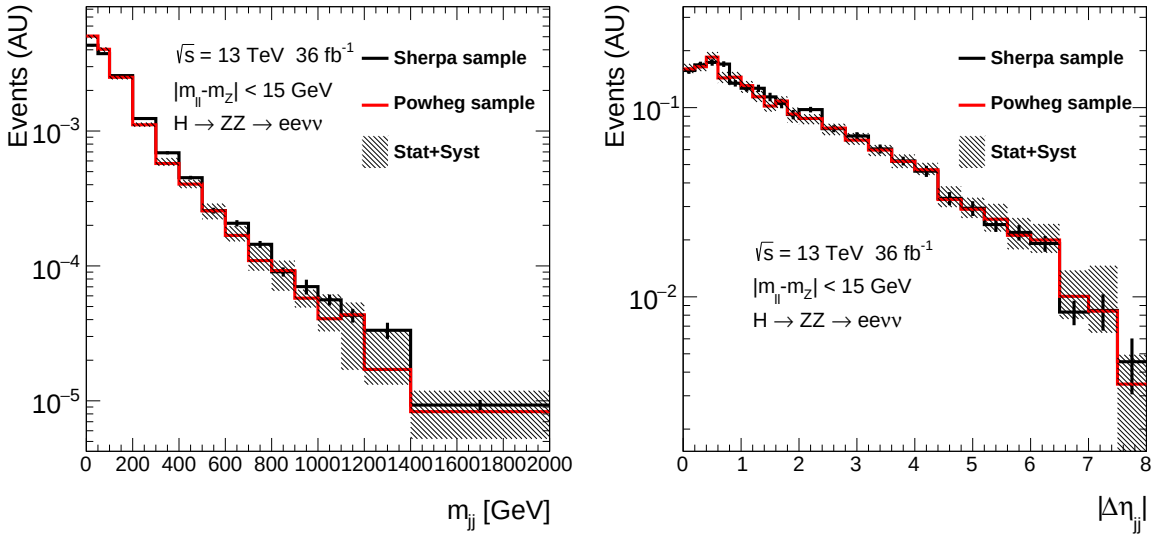


Figure 6.4: The m_{jj} and $|\Delta\eta_{jj}|$ comparison distributions between `Powheg` and `Sherpa` simulated samples. Both agree withing the statistical+systematic uncertainties.

Generator	$ee\nu\nu$	$\mu\mu\nu\nu$
`Powheg` ϵ	5.89%	6.35%
`Sherpa` ϵ	5.77%	5.72%

Table 6.4: Selection efficiencies between `Powheg` and `Sherpa` with addition of $m_{jj} > 550$ GeV and $|\Delta\eta_{jj}| > 4.4$.

$qq \to ZZ$ VBF Category	
$ee\nu\nu$	0.78 ± 0.19 (stat) ± 0.45 (syst)
$\mu\mu\nu\nu$	0.85 ± 0.16 (stat) ± 0.35 (syst)

Table 6.5: Estimated $qq \to ZZ$ background yields in the VBF category. The statistical uncertainty and systematic uncertainty on the yield are derived from the impact of experimental uncertainties plus the modelling systematic.

The modelling uncertainty on the acceptance is estimated from studies of the selection efficiency. It is defined as the ratio of the number of events after a Z mass window selection and with at least two jets to the number of events with full VBF selections, $\epsilon = N_{\text{Zmass}+N_{\text{jets}}+\text{VBF}}/N_{\text{Zmass}+N_{\text{jets}}}$. The differences between the selection efficiencies give a modelling uncertainty of 10%. The efficiencies are provided in Table 6.4. The modelling of the transverse mass distribution is compared between the two generators in Figure 6.5. The compatibility between the MC samples was found to be at a satisfactory level.

Discrepancies between data and the prediction in `Sherpa` are observed for the 13 TeV data in VBF Z + jets events [183]. For $m_{jj} \approx 550$ GeV a disagreement of $\sim 20\%$ is seen. In addition, the 8 TeV results in W + jets measurements find a similar difference [184]. From these considerations, a conservative 20% modelling uncertainty was added, giving an overall VBF $qq \to ZZ$ modelling systematic of $\approx 22\%$.

The yield in the $qq \to ZZ$ background in the VBF signal region is summarized in Table 6.5. It is obtained from the `Powheg` generator with added $k-$factor corrections. Statistical uncertainties are taken directly from MC while systematic uncertainties are combined from detector and VBF modelling systematics.

6.1.2.2.2 $gg \to ZZ + 2$jets The `gg2VV` sample lacks sufficient statistics in the VBF category. The yield of the $gg \to ZZ$ background is determined by using the outcome from the ggF category times a simulation-based transfer factor (TF) i.e.

$$N^{\text{VBF}}_{gg\to ZZ} = N^{\text{ggF}}_{gg\to ZZ} \times \text{TF} = N^{\text{ggF}}_{gg\to ZZ} \times (N^{\text{VBF}}/N^{\text{ggF}}) \tag{6.2}$$

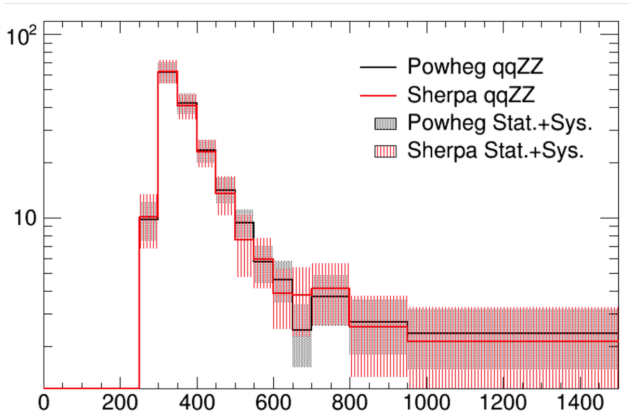


Figure 6.5: The m_{T}^{ZZ} distributions of `Sherpa` and `Powheg` samples with added uncertainties in the combined $H \to ZZ \to \ell\ell\nu\nu$ channel.

The TF is derived from the ratio of yields between the VBF and ggF phase space with certain selections removed to reduce the statistical uncertainty of the TF. The modelling uncertainty on the TF is estimated from two alternative MC samples, one generated with `Sherpa` and a `Powheg` on-shell Higgs resonance sample. The `Sherpa` sample was checked and found to lack events in the signal region so instead the on-shell Higgs boson sample is used to compare the effect of the jet modelling on the TF. As the Higgs production sample is generated in the 4ℓ final state, one of the lepton pairs is replaced by a pair of neutrinos, resulting in $E_{\mathrm{T}}^{\mathrm{miss}}$.

In Table 6.6 the difference in the TF is presented. The nominal efficiency is taken as 1.4%, with variation of up to 50%. The relative efficiency drops in the `gg2VV` due to different number of jets in the on-shell Higgs resonance region. The modelling of the jet kinematics for the various samples is given in Figure 6.6. The 50% modelling uncertainty is combined with the 60% $k-$factor uncertainty from the LO $gg \to ZZ$ expectation. From these considerations, a conservative uncertainty estimate of 100% is assigned. The number of weighted $gg \to ZZ$ events in the VBF signal region is summarized in Table 6.7.

6.1.3 WZ background

The second largest diboson background is the $WZ \to \ell\nu\ell^+\ell^-$ process. The requirement of exactly two opposite charge and same flavour leptons greatly suppresses the number of WZ events. However, the WZ process can mimic the di-lepton + $E_{\mathrm{T}}^{\mathrm{miss}}$ signature if one of the leptons is mis-reconstructed. The normalization of the WZ background in data is obtained from a dedicated control region. The control region is defined by relaxing the cut on the number of leptons in order obtain an enriched in WZ region. The shape of the m_{T}^{ZZ} distribution is obtained from simulation. The experimental and theoretical (due to running

Relative VBF efficiency w.r.t inclusive	gg2VV	Sherpa	on-shell Higgs
After inclusive selections	1.4%	0.7%	—
After inclusive selections, but not $E_{\mathrm{T}}^{\mathrm{miss}}$ cut	1.4%	0.6%	—
Before inclusive selections	1.6%	1.3%	0.8%
Before inclusive selections but requiring on-shell Higgs region	0.85%	1.31%	0.8%

Table 6.6: Relative efficiency of VBF selections with respect to the $H \to ZZ \to \ell\ell\nu\nu$ analysis selections.

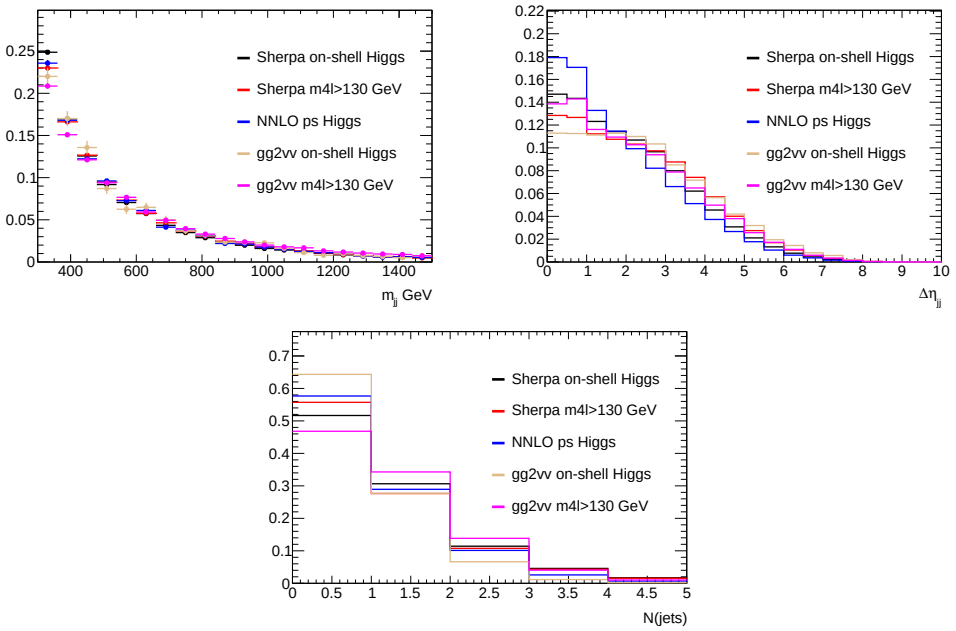


Figure 6.6: Differences in the modelling of jet distributions in different MC samples: (top left) invariant mass (m_{jj}) of the two leading jets, (top right) $|\Delta\eta_{jj}|$ of the two leading jets (bottom center) and the number of selected jets (N_{jets}). Figures produced from an analysis collaborator.

$gg \to ZZ$ VBF Category	
$ee\nu\nu$	0.23 ± 0.23 (stat.+syst.)
$\mu\mu\nu\nu$	0.23 ± 0.23 (stat.+syst.)

Table 6.7: Estimated $gg \to ZZ$ background yields in the VBF category. The statistical uncertainty and systematic uncertainty on the yield are derived from the impact of experimental uncertainties.

of QCD scale, and variation of PDFs) uncertainties is also accounted.

6.1.3.1 Background estimation in the ggF category

The WZ yield is derived from a three lepton control region, obtained by loosening of the third lepton veto and then extrapolating to the $H \to ZZ \to \ell\ell\nu\nu$ phase space. A scale factor (SF) correction is calculated from the observed differences between data and MC which are applied to the WZ simulation yield in the signal region. The yield in the signal region is found by:

$$N^{\text{data}}_{2\ell,\text{SR}} = N^{\text{MC}}_{2\ell,\text{SR}} \times \left(\frac{N^{\text{data}}_{3\ell,\text{CR}}}{N^{\text{MC}}_{3\ell,\text{CR}}}\right) = N^{\text{data}}_{3\ell,\text{CR}} \times \left(\frac{N^{\text{MC}}_{2\ell,\text{SR}}}{N^{\text{MC}}_{3\ell,\text{CR}}}\right) \tag{6.3}$$

where $N^{\text{data}}_{2\ell,\text{SR}}$ is the number of events in the data in the $\ell\ell + E_{\text{T}}^{\text{miss}}$ signal region and the SF is defined as $\text{SF} = N^{\text{data}}_{3\ell,\text{CR}}/N^{\text{MC}}_{3\ell,\text{CR}}$. A simulation-based TF between the 2-lepton and 3-lepton regions is defined as $N^{\text{MC}}_{2\ell,\text{SR}}/N_{3\ell,\text{CR}}$.

As mentioned earlier, the control region is obtained by the removal of the third lepton veto, but it is not the sole requirement. Further selections are made, to better isolate the WZ topology from other backgrounds. The third lepton satisfies the same set of selections as the two leptons, including a p_{T} cut greater than 20 GeV. The Z mass window selection is also placed, to select events containing a Z boson. A transverse mass variable, m_{T}^{W}, is defined from the $E_{\text{T}}^{\text{miss}}$ and the p_{T} of the third lepton. The formula follows the same definition as the m_{T}^{ZZ} (eq. 5.1) with p_{T}^{Z} replaced by the third lepton transverse momentum. The m_{T}^{W} is required to be greater than 60 GeV to reduce the number of Z + jets events. The contamination of $t\bar{t}$ events is restrained by a veto on $b-$jets. This selection yields a WZ purity[1] in the control region of about 90%.

The control region is split into four categories based on the lepton flavour: $ee+e, ee+\mu, \mu\mu+e, \mu\mu+\mu$. The scale factors are estimated in each category from which a single combined SF is derived, defined as the average SF from all four regions. Table 6.8 presents the SF numbers for each region and the total number of expected and observed WZ background events. The non-WZ backgrounds are then subtracted from the data. The combined SF, which is used to normalize the WZ events, is found to be $1.29 \pm 0.03(\text{stat.}) \pm 0.07(\text{sys.})$. The comparison between data and simulation is presented in Figure 6.7 for the $E_{\text{T}}^{\text{miss}}$ distribution in the 3ℓ control region. The modelling of the p_{T}, η, ϕ between data and simulation of the third lepton is provided in Figure H.1.

The MC prediction scaled by the data-driven SF is provided in Table 6.9. The statistical uncertainty on the WZ SF is found from the statistical error due to the amount of data in

[1] The purity of a given background in a control region is defined by the ratio of the expected background to the total expected background in that region.

channel	Expected	Observed	scale factor
ee + e	632.2 ± 5.0	958	1.39 ± 0.05
ee + μ	673.7 ± 5.2	941	1.25 ± 0.05
$\mu\mu$ + e	702.7 ± 5.3	993	1.26 ± 0.06
$\mu\mu$ + μ	797.6 ± 5.7	1121	1.28 ± 0.04
Total	2806.1 ± 10.6	4013	1.29 ± 0.03

Table 6.8: The expected number of background and observed number of events in the 3 lepton control region $m_{\mathrm{T}}^{W} > 60$ GeV after applying the b-jet veto region. The last column shows the SF. It is obtained by subtracting other backgrounds (in MC) from the observed yields in data, and taking the ratio between this and the expected number of events from the WZ MC. Number of observed events is for an integrated luminosity of 36.1 fb^{-1}. Errors are statistical only.

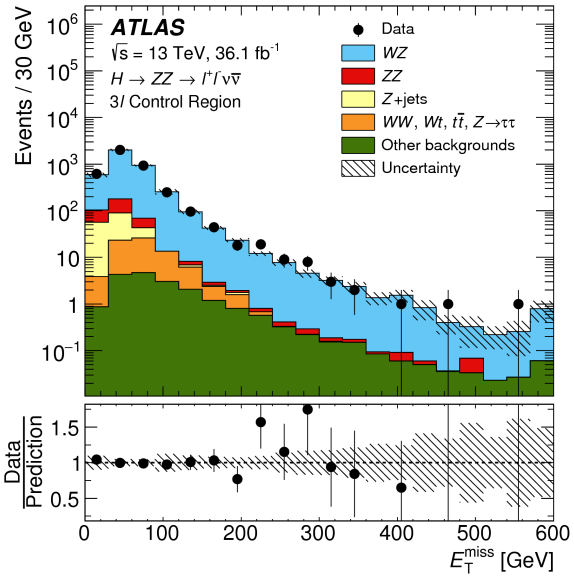


Figure 6.7: Data/MC modelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the 3ℓ control region. Figure is taken from [6].

the WZ control region. The systematic uncertainty is evaluated using the MC-driven TF. The QCD scale and PDF uncertainties have been found to have a negligible impact on the TF ($< 0.2\%$). The detector uncertainties were evaluated and found to be on the order of $\sim 4\%$.

6.1.3.2 Inefficiency SF for WZ events

An additional SF is applied to the WZ backgrounds. This is needed to account for the inefficiency of selecting exclusively di-lepton events in the signal region and missing the third lepton. To estimate the inefficiency SF, two main conditions are imposed on the third

lepton: a) it is mis-identified or not reconstructed, and b) it is within the boundary of the ATLAS ID.

The inefficiency SF is calculated by using the efficiencies in data and simulation for finding the third lepton in each of the four possible lepton combinations, following the formula:

$$\mathrm{SF} = \frac{1 - \epsilon^{\mathrm{data}}}{1 - \epsilon^{\mathrm{MC}}} \tag{6.4}$$

where ϵ can either be the efficiency in data or MC of entering the signal region. Table 6.10 presents the impact of the inefficiency SF on the total WZ background yield in the various lepton channels. The inefficiency SF on the WZ background contributes $\sim 3-4\,\%$ when the third lepton is a muon, and 1% when it is an electron. This difference is added as an additional systematic to the WZ background. The impact of the SF on the number of WZ events is estimated when the largest SF is multiplied by the percent yields of the W decay fraction ($W \to \mu\nu\,(20\,\%)$ and $W \to e\nu\,(15\,\%)$).

WZ ggF Category	
$ee\nu\nu$	92.8 ± 2.2 (stat) ± 4.3 (syst)
$\mu\mu\nu\nu$	99.5 ± 2.3 (stat) ± 3.2 (syst)

Table 6.9: Estimated WZ background yields in the ggF category. The statistical error is obtained from the MC generator uncertainty. The systematics on the yield is derived from the impact of detector and $k-$factor uncertainties.

	no 'anti-ID' SF	with 'anti-ID' SF	% Increase in Yield
$\mu\mu + \mu$	1770	1959	+10%
$ee + \mu$	1745	2098	+20%
$\mu\mu + e$	1176	1251	+6%
$ee + e$	1200	1270	+5%
Total	5891	6578	+12%

Table 6.10: The yields in MC which shows the effect of the 'anti-ID' SF. The biggest impact of applying the inefficiency SF occurs when the third lepton is a muon, which is due to slightly larger reconstruction efficiencies which makes discrepancies between data and MC more evident.

6.1.3.3 Background estimation in the VBF category

The di-jet analysis cuts ($m_{jj} > 550\,\text{GeV}$, $|\Delta\eta_{jj}| > 4.4$) leave very little available statistics to estimate the WZ SF. The WZ yield is obtained from a TF calculated in a loose VBF region multiplied with the outcomes from the ggF estimation. The loose VBF defined by requiring at least two jets which must have a p_{T} greater than 30 GeV. An alternative MC sample, generated by `Sherpa`, is used because of the better jet description. The systematic uncertainties, similar to the ggF case, are determined from the simulated TF. They evaluate the variation of the WZ yield due to the detector systematics and QCD, PDF uncertainties. The systematic is 5% (29%) from the experimental (theoretical) sources. An additional systematic, at the level of $\sim 11\%$ is added, due to modelling differences between `Powheg` and `Sherpa` generators, resulting in different TF values. The WZ TF from `Sherpa` was estimated to be $0.95 \pm 0.04(\text{stat}) \pm 0.28(\text{syst})$ and the corresponding yields are given in Table 6.11. No inefficiency SF are considered for the VBF signal region.

WZ VBF Category	
$ee\nu\nu$	0.63 ± 0.03 (stat) ± 0.19 (syst)
$\mu\mu\nu\nu$	0.66 ± 0.03 (stat) ± 0.19 (syst)

Table 6.11: Estimated WZ background yields in the VBF category. The statistical uncertainty and systematic uncertainty on the yield are derived from the impact of experimental uncertainties.

6.1.4 $t\bar{t}, WW, Wt$, and $Z \to \tau\tau$ backgrounds

The four processes $t\bar{t}, WW, Wt$ and $Z \to \tau\tau$ have a unique property which allows them to be estimated in a single control region. They have the third largest contribution, making up $\sim 7\%$ of the total background. The data-driven estimation of these sets of backgrounds is performed via an $e\mu$ control region which is further described in 6.1.4.1.

In the $\ell\ell\nu\nu$ analysis these backgrounds are mostly suppressed by specific analysis criteria:

- $t\bar{t}$ production is suppressed by requiring a veto on b-jets.
- WW background is reduced by the Z mass window requirement and the large $E_{\text{T}}^{\text{miss}}$ cut.
- The Wt background has a small production cross section and doesn't have a large impact on the total background.
- $Z \to \tau\tau$ is reduced by the Z mass cut plus a few analysis selections.

6.1.4.1 $e\mu$ control region

All four backgrounds have decay modes which resemble the di-lepton plus $E_{\mathrm{T}}^{\mathrm{miss}}$ signature. However, they have a unique feature in the lepton sector which is a consequence of lepton universality. Specifically the backgrounds approximately decay in a $1:1:2$ proportion in the $ee:\mu\mu:e\mu$ channels. This motivates the use of a dedicated control region in the $e\mu$ phase space from which the number of events in the di-lepton final state can be determined.

The $e\mu$ control region contains opposite sign and opposite flavour leptons. In addition, the control region offers a natural suppression of processes with Z boson production. Any additional leptons with loose id and $p_{\mathrm{T}} > 7$ GeV which may enter the $e\mu$ region are vetoed to reduce WZ production. A set of extra requirements are imposed in order to better approximate the $H \to ZZ \to \ell\ell\nu\nu$ signal region and reduce contamination from other backgrounds. In particular a 15 GeV Z mass window selection is applied on the invariant mass formed from the $e\mu$ pair, $|m_{e\mu} - m_{\mathrm{Z}}| < 15$ GeV. The full $e\mu$ control region is formed when the rest of the $H \to ZZ \to \ell\ell\nu\nu$ analysis selections are applied.

The modelling between data and simulation of observables, similarly to section 5.4.1, is compared after the Z similar requirement. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the $e\mu$ control region is given in Figure 6.8. Excesses are present at $E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV which lies outside of the control region. The purity of $e\mu$ events is greater than 95% while the remaining contribution originates from diboson processes.

Using the $e\mu$ number of events obtained in the control region, the number of ee ($\mu\mu$) events in the signal region, $N_{ee,\mathrm{SR}}^{e\mu}$ ($N_{\mu\mu,\mathrm{SR}}^{e\mu}$), is found from:

$$N_{ee,\mathrm{SR}}^{e\mu} = \frac{1}{2} \times \epsilon \times N_{e\mu}^{\mathrm{data,sub}}, N_{\mu\mu,\mathrm{SR}}^{e\mu} = \frac{1}{2} \times \frac{1}{\epsilon} \times N_{e\mu}^{\mathrm{data,sub}} \tag{6.5}$$

where $N_{e\mu}^{\mathrm{data,sub}} = N_{e\mu}^{\mathrm{data}} - N_{\mathrm{sub}}^{\mathrm{other}}$ is the number of data events in $e\mu$ control region after subtracting the non-$e\mu$ background ($N_{\mathrm{sub}}^{\mathrm{other}} = \sum_i^{\text{non-}e\mu} N_i$). Since the electron and muon objects have different reconstruction, an additional factor, ϵ, is included in the estimation to reflect change in acceptance due to the different reconstruction object efficiencies. This factor is defined as:

$$\epsilon^2 = \frac{N_{ee}^{\mathrm{Zmass}}}{N_{\mu\mu}^{\mathrm{Zmass}}} \tag{6.6}$$

where $N_{ee}^{\mathrm{Zmass}}(N_{\mu\mu}^{\mathrm{Zmass}})$ are the number of events in the $ee(\mu\mu)$ channel after the Z mass window cut.

The ϵ factor, as defined in equation 6.6 is "unbinned", as no additional requirements are placed on the number of events. However, this is not exactly a correct treatment since the performance of the detector reconstruction varies in different η regions and for different

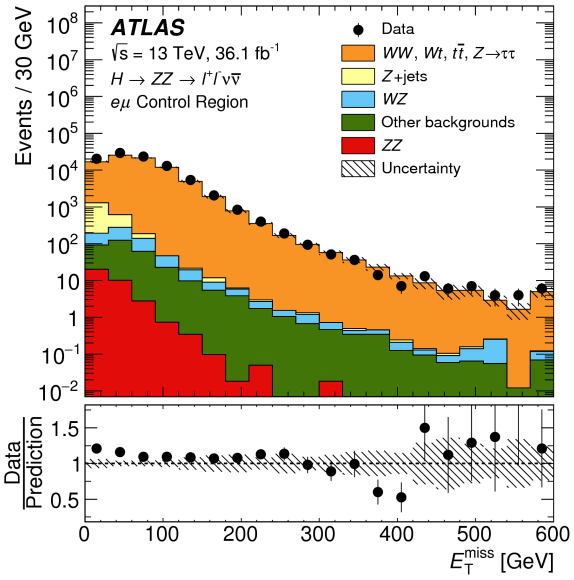


Figure 6.8: Comparison between data and simulation after a $|m_Z - m_{e\mu}| < 15$ GeV cut in the $e\mu$ control Region for the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Statistical and systematic errors are included in the error band of the plots. Excesses are below $E_{\mathrm{T}}^{\mathrm{miss}}$ of 120 GeV and therefore occur outside of the control region. Figure is taken from [6].

momenta ranges. Therefore a more accurate assessment of the efficiency factor is done through the "binned" efficiency factor, which is defined as:

$$\epsilon^2 = N^{\mathrm{Zmass}}_{e^1_{(p_T,\eta)}e^2_{(p_T,\eta)}} / N^{\mathrm{Zmass}}_{\mu^1_{(p_T,\eta)}\mu^2_{(p_T,\eta)}}. \tag{6.7}$$

where $N^{\mathrm{Zmass}}_{e^1_{(p_T,\eta)}e^2_{(p_T,\eta)}}$ ($N^{\mathrm{Zmass}}_{\mu^1_{(p_T,\eta)}\mu^2_{(p_T,\eta)}}$) are the number of events in the ee ($\mu\mu$) channel after the Z mass window requirement from two dimensional binned p_{T} and η distribution. The p_{T} binning is: $p_T < 44$ GeV, 44 GeV $< p_T <$ 52 GeV, 52 GeV $< p_T <$ 2 TeV. The binning in η is defined in four different regions: BB, BE, EB and EE where B stands for barrel and E for end-cap. The ordering of the regions is related to the p_{T} of the leading and sub-leading leptons e.g. BE means that the leading lepton is in the barrel while the sub-leading is in the end-cap. Sufficient statistics for the determination of the efficiency factor was ensured by optimization of the binning. The efficiency factor in bins of p_{T} and η are summarized in Table I.2 (I.1) for data (MC).

6.1.4.2 Closure in simulation

The $e\mu$ background estimation method is validated with a "closure test" i.e. by checking if the outlined procedure produces compatible outcomes with the naive prediction. Closure is achieved if the derived yield from simulation is compatible with the expected number of events in the signal region. The number of $e\mu$ events for each case are shown in Table 6.12.

Channel	$N_{e\mu}^{est.}$	$N_{e\mu}^{MC}$
$ee\nu\nu$	11.90 ± 1.51	10.19± 1.54
$\mu\mu\nu\nu$	13.52 ± 1.63	14.34 ± 2.37

Table 6.12: Closure of the $e\mu$ backgrounds estimation. $Ne\mu^{est}$ is the number of $e\mu$ events in SRs estimated using the binned efficiency factors while $Ne\mu^{MC}$ is the expected event yields for the $e\mu$ background. Numbers are normalised to an integrated luminosity of 36.1 fb^{-1}. Errors contain the statistical uncertainty only.

In this closure test, the binned ϵ factor is used. The results agree within the statistical uncertainties with the remaining non-closure in the central values taken as a source of systematic uncertainty. It is 6% for in the $ee\nu\nu$ channel and 14% in the $\mu\mu\nu\nu$ channel.

6.1.4.3 Yields in ggF signal region

The yields are estimated on a event-per-event basis using the binned efficiency factor. The unbinned efficiency factor is not used in the final estimate. The results for the ggF signal region are shown in Table 6.13 for both data and simulation including statistical and systematic uncertainties. The systematic uncertainties are taken from the amount of non-closure.

Channel	ggF Data Driven estimate	MC with Binned (p_{T},η)
$ee\nu\nu$	9.2 ± 2.2 (stat) ± 1.4 (syst)	9.9 ± 1.6 (stat) ± 2.4 (syst)
$\mu\mu\nu\nu$	10.7 ± 2.5 (stat) ± 0.9 (syst)	14.0 ± 2.5 (stat) ± 2.7 (syst)

Table 6.13: Background estimates in ggF category with binned efficiency. The data driven and MC methods are both shown including statistical and systematic uncertainties.

6.1.4.4 $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape and systematics

The shape of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution is estimated using data in the $e\mu$ control region. The $e\mu$ $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape is normalized according to the data-driven yields. Figure 6.9 shows the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape for both data and simulation. The overall agreement is good, but the tail of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution in data lacks sufficient statistics for a proper description. A shape uncertainty, estimated from MC, is added instead.

The shape systematic is split into a core and tail parts. The core is estimated from the up and down variations of the shape, obtained by varying the binned ϵ factor by $\pm 1\sigma$. The

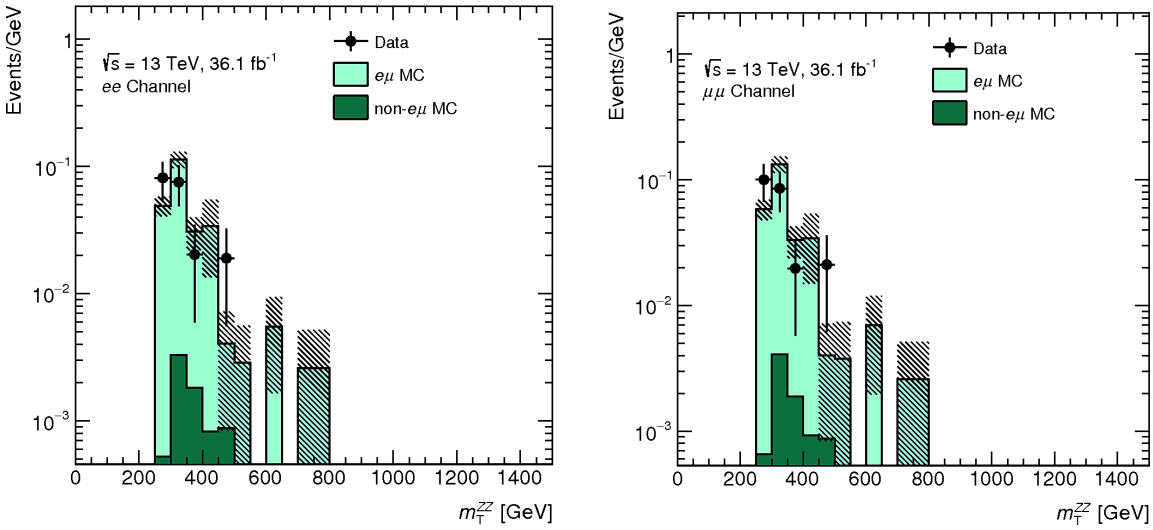


Figure 6.9: $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape for the $e\mu$ background in the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. The error bands include statistical uncertainty only. Figure produced from analysis collaborator.

impact is 1 %, averaging from all bins. The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ tail is taken from the $e\mu$ MC, which is then attached to the core as a tail systematic on the upward shape variation. The systematics are then normalized to the data driven estimate. The nominal and derived up/down shape systematics are shown in Figure 6.10.

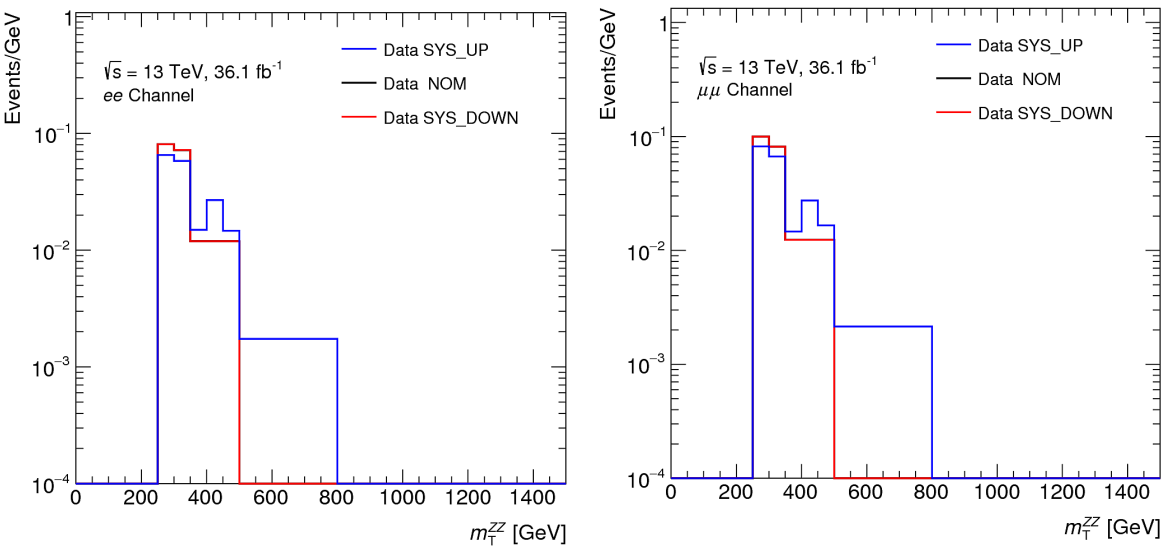


Figure 6.10: $m_{\mathrm{T}}^{\mathrm{ZZ}}$ systematic up(blue), down(red) and nominal(black) shape for the $e\mu$ background in the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channel. Figure produced from analysis collaborator.

6.1.4.5 Estimation in the VBF category

In the VBF category, the analysis requirements are found to be too restrictive to obtain an estimate of the $e\mu$ number of events. The last selection after which any events survive is $N_{\text{jets}} \geq 2$ which are 1.23 ± 0.73 (1.74 ± 1.03) for the $ee\nu\nu$ ($\mu\mu\nu\nu$) channel.

The estimation proceeds by extrapolation to the VBF phase space by using a TF derived from simulation. It is calculated from the ratio of the yield after the $N_{\text{jets}} \geq 2$ cut and the yield after the full VBF selections. To limit the statistical uncertainty the TF is derived from the combination between both lepton channels. The value of the MC based TF is found to be 0.131 ± 0.082. The $e\mu$ VBF yields are given in Table 6.14.

The systematic uncertainty is computed from the experimental and modelling uncertainties on the TF. The experimental uncertainties are propagated from the uncertainties on the individual objects such as JES, $E_{\text{T}}^{\text{miss}}$ and the lepton reconstruction uncertainties. These contribute $\sim 60\%$ to the systematic uncertainty. The modelling uncertainty of the VBF selections is taken from a region, where the Z mass requirement is relaxed in order to increase the available statistics. The number of $e\mu$ entries without Z mass cut is given in Table 6.15. The maximal TF difference between data and MC, obtained from the m_{jj} and $|\Delta\eta_{jj}|$ distributions, are taken as the modelling uncertainty. This results in a 53% systematic on the TF. The statistical uncertainty on the MC-based TF is propagated from the data-driven estimate at the $N_{\text{jets}} \geq 2$ selection step and the statistical uncertainty on the $N_{\text{jets}} \to \Delta\eta_{jj}$ selection efficiency. The uncertainty on the $e\mu$ estimate is found to be $\sim 92\%$.

$e\mu$ VBF Category	
$ee\nu\nu$	0.16±0.14 (stat) ±0.15 (syst)
$\mu\mu\nu\nu$	0.23±0.20 (stat) ±0.21 (syst)

Table 6.14: Data-driven $e\mu$ estimate in the VBF category. The yields are obtained by multiplying the data-driven estimates after the $N_{\text{jets}} \geq 2$ selection by the TF for the $ee\nu\nu$ and $\mu\mu\nu\nu$ channels separately. The statistical uncertainties are propagated through the errors of the TF. The systematic uncertainties are estimated through the experimental and modelling uncertainties of TF (53%)

6.1.5 Z+jets background

The Z + jets process can falsely appear as a genuine event signature due to the mis-reconstruction or mis-calibration of jets which then produce a non-negligible $E_{\text{T}}^{\text{miss}}$. The core of the $E_{\text{T}}^{\text{miss}}$ distribution occurs at low values since low energy jets are mis-reconstructed

Source	$N_{\text{jets}} \geq 2$ region	VBF region	TF
Data	40	2	0.05
$e\mu$ MC	23.89 ± 2.2	1.16 ± 0.4	0.0486 ± 0.0172

Table 6.15: $e\mu$ yields in N_{jets} region and VBF region without the Z-mass requirement.

more often than highly energetic ones. The production cross section of the $Z \to \ell\ell$ process is 1869 pb [185] which results in $\mathcal{O}(10^7)$ events after the Z mass requirement.

In the $H \to ZZ \to \ell\ell\nu\nu$ analysis the Z + jets background can be reduced by more than 99% when a sufficiently high $E_{\text{T}}^{\text{miss}}$ cut is imposed. Additional selections are applied to further remove the remaining Z + jets events in the $E_{\text{T}}^{\text{miss}}$ tail without affecting the signal contribution.

The data-driven normalization and its uncertainty are obtained from a modified ABCD method. The Z + jets transverse mass shape is found from a one-sideband control region. Shape systematics are assigned from differences with the shape in a looser signal region. In the VBF region a MC-based TF is used to derive the yield from the data-driven estimate in the ggF region.

The Z + jets process in the $H \to ZZ \to \ell\ell\nu\nu$ signal region forms a minor contribution to the total background. However, the signal sensitivity depends, in part, on the large associated systematic uncertainties which makes the estimation of the Z + jets background an important priority of the analysis.

6.1.5.1 ABCD Methodology

To arrive at a data-driven estimate of the background process, a control region is specified which is enriched in Z + jets events. Other backgrounds are subtracted from data to limit contamination in the estimation. The ABCD method specifies 4 different sectors, one being the signal sector (A), where the analysis selections hold, and three sideband sectors (B,C,D) where an event is accepted if it fails those selections. The method is typically specified in a two dimensional phase space, defined by two uncorrelated analysis observables. The remaining analysis selections are applied beforehand, to be as close as possible to the signal region. The number of events in A, can then simply be found from:

$$N_A^{\text{SR}} = N_C^{\text{CR}} \times \frac{N_B^{\text{CR}}}{N_D^{\text{CR}}} \tag{6.8}$$

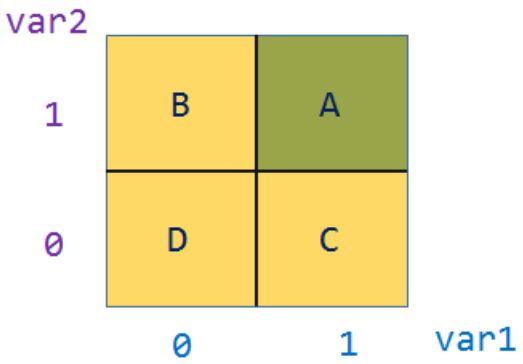


Figure 6.11: Scheme of the boolean ABCD method. Signal region A is defined by both variables equal true, representing all selections passed. For region B,D the event failed either the first or second boolean cut, while for region C the event failed both boolean cuts.

where N_A^{SR} is the number of background events in the signal region A, while N_X^{CR} are the number of events in the sideband control regions (X = B,C,D). The non-Z backgrounds are subtracted from data in the sideband regions before applying this formula.

In previous iterations of the analysis, the ABCD method used the $\Delta\phi(p_{\mathrm{T}}^{\mathrm{Z}}, E_{\mathrm{T}}^{\mathrm{miss}})$ and the fractional p_{T} cuts to define the four regions, with A being the region where both selections are passed while B, C and D contain events which fail one or both of these requirements.

6.1.5.2 Shortcomings and modification of the ABCD method

The simple version of the ABCD method as outlined above is no longer sufficient. The reason lies in the increased suppression of Z + jets events due to the introduction of the $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cut. Many attempts to construct a suitable control region failed, either from lack of statistics or from non-closure due to high correlation between observables. This prompted the need for a modified definition of the background estimation, the outcome of which is called the "boolean approach to ABCD".

In the boolean re-definition, the two variables (var1 and var2) consist of two boolean variables formed from several selections on analysis observables. In this way, if any event fails anyone of the multiple selections, it is placed in one of the sideband regions, increasing the overall event population. The scheme is presented graphically in Figure 6.11.

The choice of the boolean variables presented a challenge, as it was difficult to find the best set with minimal correlation. The set of boolean variables that are chosen is:

$$\mathrm{var1} = E_{\mathrm{T}}^{\mathrm{miss}} > 120 \text{ GeV AND } E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} > 0.4 \tag{6.9}$$

Z + jets ggF Category	
$ee\nu\nu$	16.9 ± 0.63 (stat) $^{+6.84}_{-11.41}$ (syst)
$\mu\mu\nu\nu$	19.7 ± 0.59 (stat) $^{+17.11}_{-9.42}$ (syst)

Table 6.16: Summary of the Z + jets background estimation using the ABCD method with statistical and systematic uncertainties. Systematic uncertainties come from the level of correlation in MC, estimated by difference between N_A/N_C [MC] and N_B/N_D [MC], experimental uncertainties on this difference and subtraction of non-Z + jets backgrounds.

$$\begin{aligned} \text{var2} = |{E_{\text{T}}}^{\text{miss,jet}} - p_{\text{T}}^{\ell\ell}|/p_{\text{T}}^{\ell\ell} < 0.2 \text{ AND } \Delta\phi(p_{\text{T}}^{Z}, E_{\text{T}}^{\text{miss}}) > 2.7 \\ \text{AND } \Delta R_{\ell\ell} < 1.8 \text{ AND } N_{b-jets} = 0. \end{aligned} \tag{6.10}$$

However, this set of variables still suffer from the presence of large correlation. The source of the correlation originates from events in the low $E_{\text{T}}^{\text{miss}}$ regions (see Figure J.1, J.2). To reduce the amount of correlation two additional requirements are added to the second boolean:

$$\text{var2}' = \text{var2 AND } E_{\text{T}}^{\text{miss}} > 30 \text{ GeV AND } E_{\text{T}}^{\text{miss}}/H_{\text{T}} > 0.1 \tag{6.11}$$

The selections were optimized to reduce the size of systematics uncertainties.

The data-driven yields of the Z + jets background in the ggF category are presented in Table 6.16. In the ABCD control region the non-Z backgrounds are subtracted from data. To obtain a better estimate the ZZ background is corrected to NNLO QCD and the WZ yield is scaled to the data-driven result. Statistical errors due to the MC subtraction are considered. The systematic uncertainties and their derivation are discussed in the next section.

6.1.5.3 Systematic Uncertainties

The modified ABCD approach has two main sources of uncertainty:

- Non-Z background subtraction
- ABCD estimation: which includes the non-closure (correlation induced bias) as quantified by the ϵ, the experimental systematics on ϵ and uncertainties from the $E_{\text{T}}^{\text{miss}}$ and $E_{\text{T}}^{\text{miss}}/H_{\text{T}}$ cut-values in var2$'$.

The up and down systematic uncertainties are kept anti-symmetric to reflect more accurately the level of accuracy and to avoid putting an over-conservative estimation on the

ABCD systematic. The size of the systematic uncertainties for each source is summarized in Table 6.17.

The bias in the ABCD method results from $N_A/N_C \neq N_B/N_D$ which can be quantified by:

$$1 + \epsilon = \frac{N_A \times N_D}{N_B \times N_C}. \tag{6.12}$$

where ϵ is the amount of non-closure of the ABCD estimation and is directly linked to the level of dis-agreement between N_A/N_C and N_B/N_D.

The bias, in data, cannot be determined since the yield in region A cannot be obtained without unblinding the signal region. In simulation, however, the value of ϵ is known since N_A can be estimated from MC. The value of the bias is taken as a source of uncertainty of the ABCD methodology. An additional source of uncertainty of the ABCD estimation are the size of the experimental systematic uncertainty on ϵ.

Another uncertainty source, as mentioned in the last section, originates from the selections on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ in var2$'$. Varying the cut-values produces different values of ϵ. The change in the size of the correlation induced bias is presented in Figure 6.12, which is derived by fixing one of the selections (either the $E_{\mathrm{T}}^{\mathrm{miss}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$) to their nominal value and varying the remaining requirement threshold in var2$'$ by 40 %. The empty circle marker represents the nominal value ($E_{\mathrm{T}}^{\mathrm{miss}} > 30, E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} > 0.1$) while the black circles represent the different ϵ values for the various ABCD regions.

To validate the studies on the bias in data the ratios of $(1 + \epsilon)$ are studied. The ratios are examined for both data and MC because the dependence on N_A is eliminated since the yield in the signal region is constant for any ABCD estimation. In the ratios the same subset of

	$ee\nu\nu$-channel	$\mu\mu\nu\nu$-channel
Methodology	$^{+30.8\%}_{-66.2\%}$	$^{+84.2\%}_{-42.8\%}$
Non-Z subtraction	26.3%	21.3%
Total systematic	$^{+40.5\%}_{-67.5\%}$	$^{+86.9\%}_{-47.8\%}$

Table 6.17: The systematic uncertainties are due to the methodology uncertainty and the uncertainty on the subtraction of non-Z backgrounds from MC. The methodology uncertainty includes the correlation-induced bias $1 - \frac{N_C \times N_B}{N_D \times N_A}$ described in the text and the uncertainty associated with the additional $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cuts, which is estimated by varying these cuts by 40% from their nominal values. Both sources yields roughly equal contributions to the methodology uncertainty.

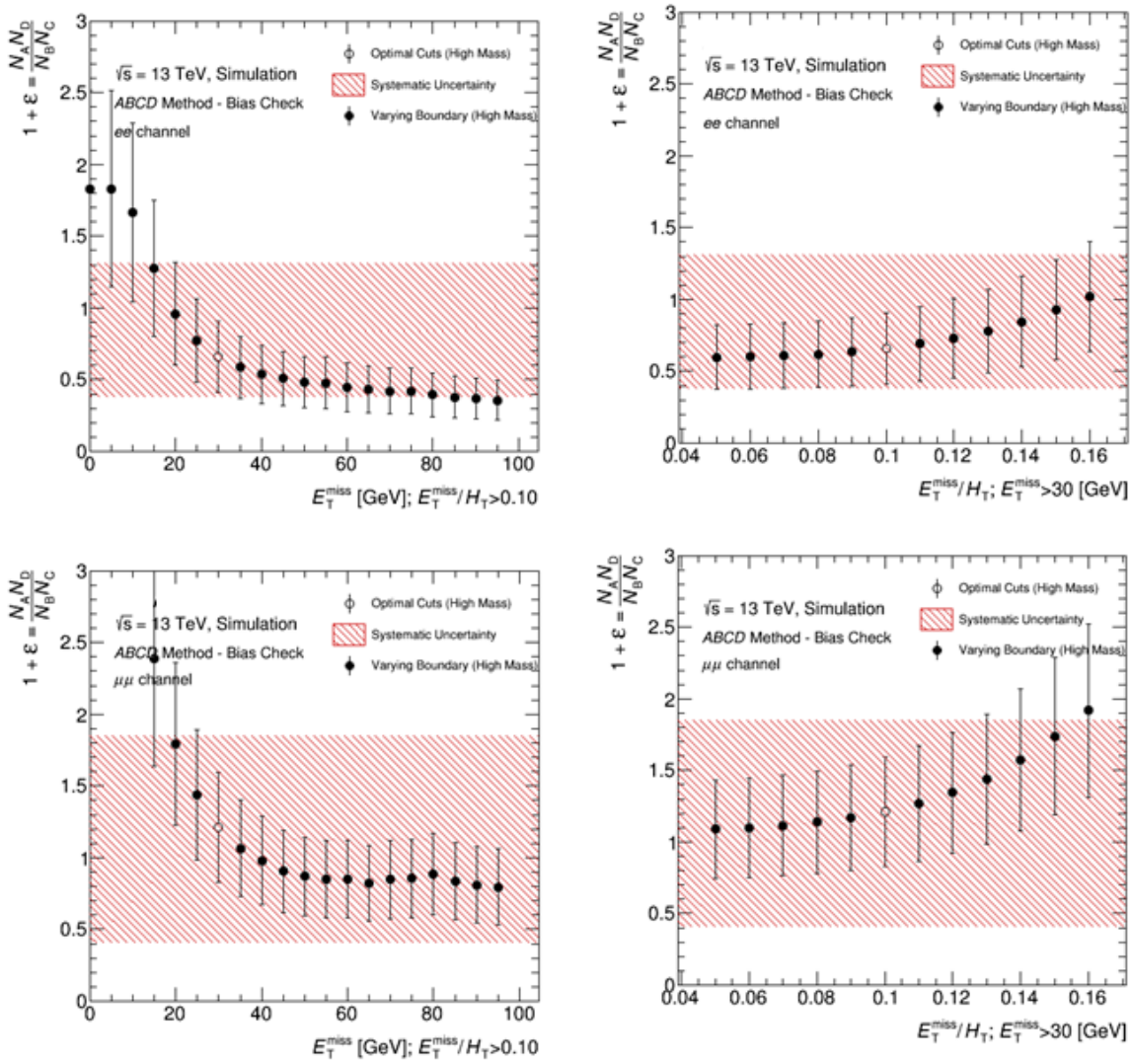


Figure 6.12: MC bias $\frac{N_A \times N_D}{N_B \times N_C}$ dependence on additional $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ cuts in the second boolean variable. The top plots are made in the $H \to ZZ \to ee\nu\nu$ channel while the bottom are in $H \to ZZ \to \mu\mu\nu\nu$ channel. The shaded band shows the combined systematic uncertainty, while the uncertainties on the markers are statistical. Figure produced from an analysis collaborator.

variables ($E_{\mathrm{T}}^{\mathrm{miss}}$,$E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$) is varied. The bias ratios are then:

$$\frac{1+\epsilon_i}{1+\epsilon_0} = \frac{N_C^0 \times N_B^0/N_D^0}{N_C^i \times N_B^i/N_D^i} \tag{6.13}$$

where ϵ_i is the bias in the i^{th} ABCD region and the 0^{th} region is the nominal. This is plotted in Figure 6.13. For both regions of low and high correlation (where the non-closure is greater) an overall consistency between data and MC is found.

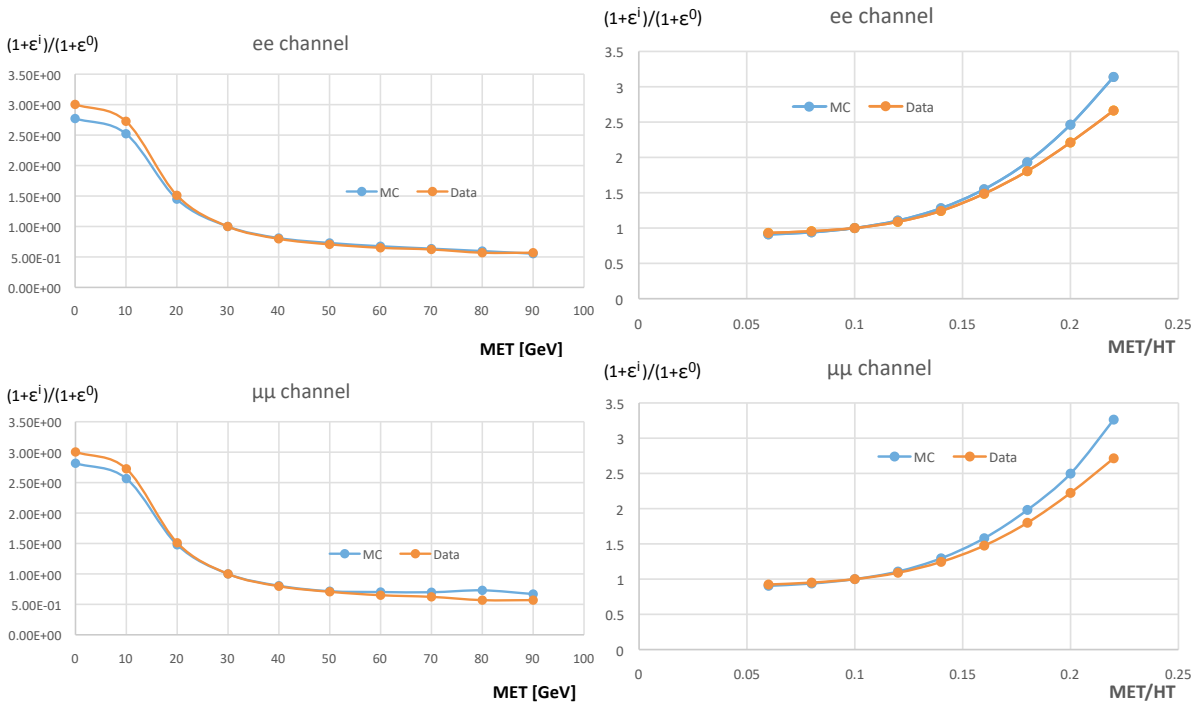


Figure 6.13: Bias ratio $(1+\epsilon_i)/(1+\epsilon_0)$ dependence on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ selections in var2′ for data and MC. The top plot is given for $ee\nu\nu$ channel while the bottom is given for $\mu\mu\nu\nu$ channel. An overall small dis-agreement is present between data and simulation in regions of high correlation. Figure produced from an analysis collaborator.

6.1.5.4 $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape using modified ABCD method

The Z + jets transverse mass shape, used in the derivation of the Higgs production exclusion limits, is not obtained from the ABCD method but instead from an alternative data-driven estimation, as outlined in section 6.1.5.6. The data-driven shape gives a more accurate representation of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution while the ABCD shape and its uncertainties are used as additional systematics.

The ABCD is defined with a looser signal region in order to extract the Z + jets shape from MC. This involves loosening a few selections, including the $b-$jet veto and minimum $\Delta\phi(p_{\mathrm{T}}^{\mathrm{jet}} > 100$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}})$ cuts. The yields are scaled using the data-driven ABCD method while the systematic on the shape is taken from experimental uncertainties.

6.1.5.5 Estimation in the VBF category

The expected yield from the Z + jets background which is present inside the VBF signal region is small and is accompanied by large fluctuations. The weighted MC yields and raw entries in the $ee\nu\nu$ and $\mu\mu\nu\nu$ channels are reported in Table 6.18. The estimates in the signal region are 0.03 ± 0.29 (0.18 ± 0.28) for the $ee\nu\nu(\mu\mu\nu\nu)$ channels with very large statistical uncertainties because the simulated sample includes few events with large event weights.

Selection	$ee\nu\nu$ Raw [Weighted]	$\mu\mu\nu\nu$ Raw [Weighted]
$\ell\ell\nu\nu$	315 [9.32 ± 3.34]	337 [16.72 ± 5.61]
$\ell\ell\nu\nu + N_{\text{jets}} \geq 2$ ($p_{\text{T}}^{\text{jet}} > 30\,\text{GeV}$)	114 [5.85 ± 1.63]	120 [4.94 ± 1.72]
$\ell\ell\nu\nu$ + VBF cuts	6 [0.02 ± 0.29]	5 [0.18 ± 0.28]

Table 6.18: Expected yields for the Z + jets background in the VBF category, as obtained with `Sherpa` sample. Both raw and weighted yields are given. Errors correspond to the statistical uncertainty only.

Selection	$ee\nu\nu$ Raw [Weighted]	$\mu\mu\nu\nu$ Raw [Weighted]
$\ell\ell\nu\nu$, loose quality	3776 [103.2 ± 5.8]	4074 [118.5 ± 8.4]
$\ell\ell\nu\nu+$ $N_{\text{jets}} \geq 2$ ($p_{\text{T}}^{\text{jet}} > 30\,\text{GeV}$), loose quality	2060 [62.2 ± 4.0]	2261 [66.3 ± 5.5]
$\ell\ell\nu\nu$ + VBF cuts, loose quality	54 [1.99 ± 0.49]	73 [2.56 ± 0.69]

Table 6.19: Expected yields for the Z + jets background at various selection levels with loose Z + jets quality requirements, as obtained with the `Sherpa` MC sample. The loose quality requirement is defined as having no $\Delta\phi$(jet, $E_{\text{T}}^{\text{miss}}$), fractional p_{T}, $b-$jet veto cuts, and $E_{\text{T}}^{\text{miss}}/H_{\text{T}}$ must be greater than 0.3 to increase statistics. Both raw and weighted yields are given. Errors correspond to the MC statistical uncertainty only.

The yield for $H \to ZZ \to ee\nu\nu$ is further reduced due to negative MC weights.

The Z + jets yields in the VBF category are obtained in a similar way to other backgrounds, namely by using a TF to extrapolate the ggF estimate to the VBF phase space. As very little events remained after all $H \to ZZ \to \ell\ell\nu\nu$ selections, a few of the Z + jets cleaning requirements are loosened to limit the statistical error. In particular, the minimum $\Delta\phi(p_{\text{T}}^{\text{jet}} > 100$ GeV, $E_{\text{T}}^{\text{miss}})$, fractional p_{T} and $b-$jet veto cuts were dropped. However, the loosened region still suffered from limited statistics, so an additional requirement was relaxed, namely the $E_{\text{T}}^{\text{miss}}/H_{\text{T}}$ cut was reduced from 0.4 to 0.3. This also ensured a meaningful value of the TF. In Table 6.19, the change in the yields is given for the loose $H \to ZZ \to \ell\ell\nu\nu$ and the loose VBF categories.

The TF is computed from the ratio of the loose VBF yield to the one from the $H \to ZZ \to \ell\ell\nu\nu$ region . The resulting value is reported in Table 6.20. A separate TF, obtained from a Z + jets sample produced from a `MadGraph5` MC generator, is also reported in order to obtain a useful comparison on the accuracy of the estimate and to assign a modelling uncertainty. The value of the `MadGraph5` TF is found to be consistent with the one obtained from `Sherpa`. Since the two numbers agree within statistical precision, a modelling uncertainty was not

Generator	$ee\nu\nu$	$\mu\mu\nu\nu$
Madgraph	0.0173 ± 0.0105	0.0139 ± 0.0141
`Sherpa`	0.0192 ± 0.0049	0.0216 ± 0.0060

Table 6.20: TFs obtained from region with loosened E_T^{miss}/H_T cut. Errors shown are statistical only.

Z + jets VBF Category	
$ee\nu\nu$	0.33 ± 0.08 (stat) ± 0.33 (syst)
$\mu\mu\nu\nu$	0.43 ± 0.12 (stat) ± 0.43 (syst)

Table 6.21: Estimated Z + jets background yields in the VBF category. Statistical and systematic uncertainties are shown. The systematic uncertainties are derived from the impact of experimental uncertainties.

assigned.

The estimate of the Z + jets background in the VBF region is summarized in Table 6.21. An extra systematic, in addition to the uncertainty on the TF, is obtained from the estimate of the ABCD bias ϵ. It contributes $\sim$ 50% to the total error. This is coupled with a $\sim$ 70% systematic from the experimental uncertainties in the VBF region. Conservatively, the overall uncertainty is set to 100 % on the VBF Z + jets yield.

6.1.5.6 One sideband method

The transverse mass shape from the ABCD methodology is not entirely reliable as it lacks sufficient statistics and is derived from simulation. In contrast, a single sideband method, as defined below, extracts the shape of the Z + jets $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution from data in a control region which can be easily extrapolated to the signal region. This method also provides a normalization value, which, while consistent with the ABCD result, has a higher central value and a large background uncertainty.

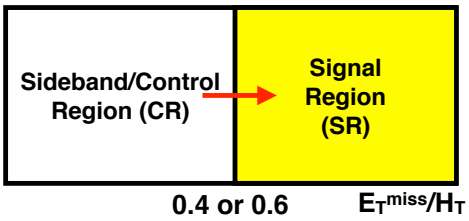


Figure 6.14: Scheme of the one sideband method. A selection is reversed to construct the sideband region.

The sideband background estimation relies on creating a control region by reversing one of the analysis selections cuts, to obtain a region with increased Z + jets production. The scheme is shown in Figure 6.14, which leads to:

$$N_{\mathrm{SR}}^{\mathrm{est}} = N_{\mathrm{CR}}^{\mathrm{data,sub}} \times \left(\frac{N_{\mathrm{SR}}^{\mathrm{MC}}}{N_{\mathrm{CR}}^{\mathrm{MC}}} \right) \tag{6.14}$$

where $N_{\mathrm{SR}}^{\mathrm{est}}$ is the estimated number of events in the signal region and $N_{\mathrm{CR}}^{\mathrm{data,sub}}$ is the number of events in data in the sideband control region, subtracted for non-Z backgrounds. The TF, $N_{\mathrm{SR}}^{\mathrm{MC}}/N_{\mathrm{CR}}^{\mathrm{MC}}$, is fully derived from simulation and extrapolates the data yield in the sideband region to the yield in the signal region.

The $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ observable is chosen since it provides good discrimination between the Z + jets and other processes. The one sideband region is defined by the reversal of the $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ requirement. The modelling between data and MC of the $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ observable is given in Figure K.2 in the sideband control regions. The prediction is overall consistent with data.

The final data-driven yields in the signal region are given in Table 6.22. This result is larger than the prediction from simulation, because of the greater number of data entries.

6.1.5.7 Data-driven estimation of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape

To obtain the lineshape of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution all analysis selections can be applied. However, this leaves insufficient statistics for a proper description of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ tail. Similar to the VBF case, certain analysis requirements are dropped to gain statistics i.e. fractional p_{T} difference, minimum $\Delta\phi(p_{\mathrm{T}}^{\mathrm{jet}} > 100\ \mathrm{GeV}, E_{\mathrm{T}}^{\mathrm{miss}})$ and $b-$jet veto. To validate this approach, the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ in simulation were compared between the loose signal region and the region with full selections, and found to agree within the statistical uncertainties. This is presented in Figure 6.15.

The shape from the loose signal region is compared with the one obtained from loose sideband control region. Both are displayed in Figure 6.16 and for comparison purposes the subtracted non-Z data shape from the loose sideband is also given.

A re-weighting function is extracted between the simulation in the loose signal and loose sideband regions. This function is then applied to the data $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape in the sideband regions. The data-driven Z + jets distribution in the signal region is given in Figure 6.17. The shape uncertainty is determined from the difference between the data-driven one and the shape from simulation in the ABCD method.

Z+jets	Yields
$ee\nu\nu$	18.1 ± 5.2 (stat) ± 5.6 (syst) [10.0 ± 3.3 (stat)]
$\mu\mu\nu\nu$	34.0 ± 8.7 (stat) ± 19.0 (syst) [16.4 ± 5.5 (stat)]

Table 6.22: Estimated Z+jets background yields with the one sideband method. Numbers in brackets are from the MC expectation for comparison. Statistical uncertainty and systematic uncertainty are shown (estimation described in the text).

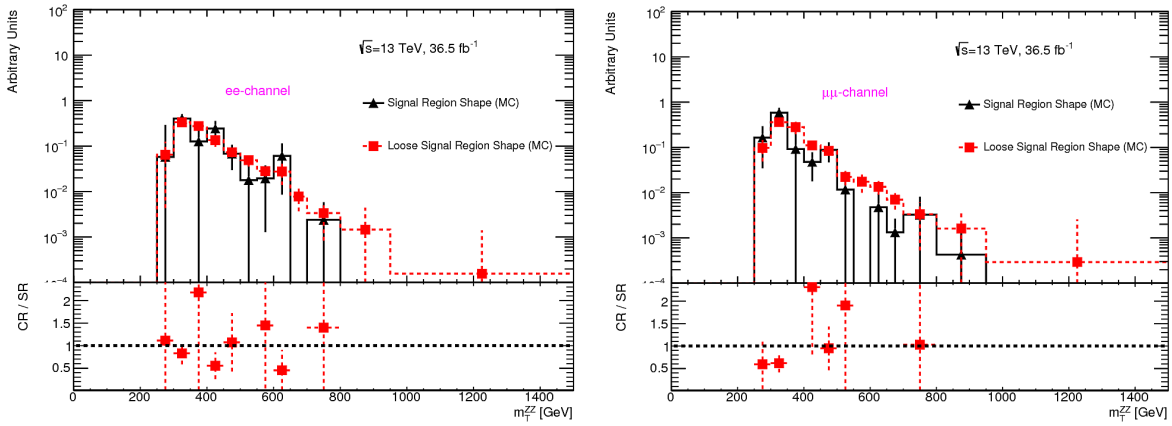


Figure 6.15: m_{T}^{ZZ} distributions in the signal region with the full selection (black) and looser selection (red) applied in the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. Figure produced from an analysis collaborator.

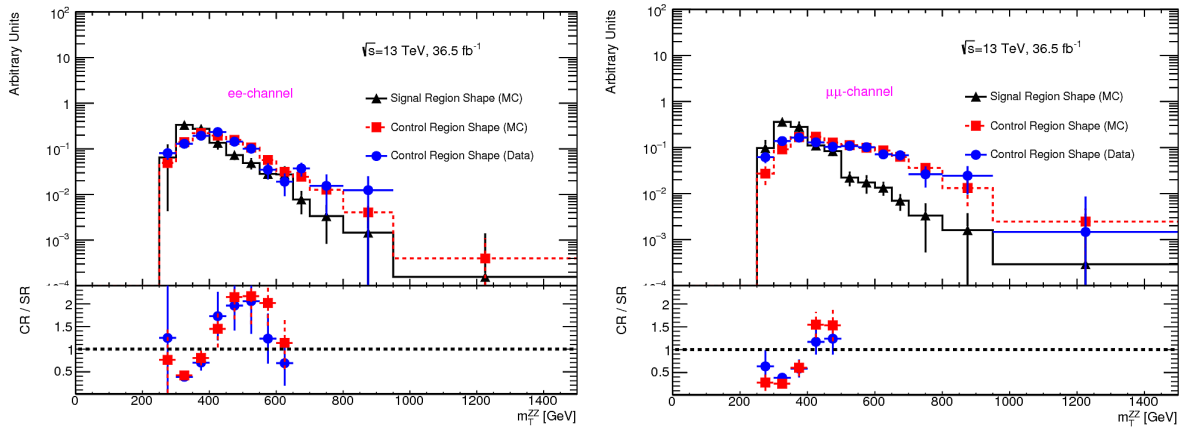


Figure 6.16: m_{T}^{ZZ} distributions in the loose signal and sideband region for the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. Figure produced from an analysis collaborator.

6.1.5.8 Systematic uncertainties

The main source of uncertainty in the one sideband method comes from the ratio $N_{\mathrm{SR}}^{\mathrm{MC}}/N_{\mathrm{CR}}^{\mathrm{MC}}$. The major impact of experimental systematics are the JES and jet response. They contribute 29% (55%) in the $ee\nu\nu$ ($\mu\mu\nu\nu$) channel. The subtraction of the non-Z backgrounds contribute 11% (8.4%) to the $ee\nu\nu$ ($\mu\mu\nu\nu$) channel for the total uncertainty. The statistical uncertainty is propagated from data and the $N_{\mathrm{SR}}^{\mathrm{MC}}/N_{\mathrm{CR}}^{\mathrm{MC}}$ term.

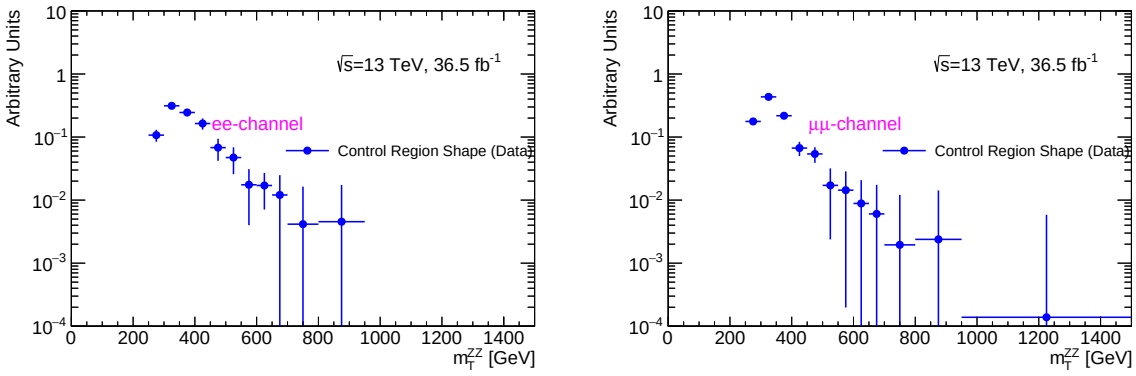


Figure 6.17: Data-driven estimation of the final m_{T}^{ZZ} distributions in the $ee\nu\nu$ (left) and $\mu\mu\nu\nu$ (right) channels. Figure produced from an analysis collaborator.

6.1.6 W+jets, ttV, tri-boson backgrounds

W + jets, ttV and VVV (where $V = W, Z$) processes contribute very little to the total background in the $H \to ZZ \to \ell\ell\nu\nu$ phase space. They provide $\sim \mathcal{O}(1\%)$ in the ggF topology and contain negligible contribution for the VBF topology. The number of W + jets events is found through a data-driven "fake" factor method while the remaining backgrounds are obtained from simulation.

The yield of the W + jets background is derived by scaling the number of events in data, estimated from a W + jets control region, by a fake factor. The fake factor quantifies the probability for a jet to be mis-identified as a lepton. More precisely, it is defined as the ratio of the probabilities for a jet to pass the full lepton object selections to the probability of a jet passing a lepton-like selection i.e. where the jet is allowed to fail either identification or isolation. The fake-factor is further binned in the lepton p_T and η to obtain a greater precision. The estimation is carried out for data and separately for the simulated W + jets process. For consistency tests, the Z + jets background is also included.

The control region is defined from a tag and probe method, by having one "good" lepton (i.e. passing all lepton selections) and one "bad" lepton (i.e. failing either the identification or isolation requirement). The reconstructed bad lepton is seen as likely to be originating from a jet. The $\ell\ell\nu\nu$ event selection is applied on top. The non-W backgrounds are subtracted after which the yield is scaled by the fake factor. The data-driven W + jets estimate is given in Table 6.23 for both lepton channels. The systematic uncertainty on data is taken from differences between the fake factor in W + jets and Z + jets estimation along with uncertainties associated with subtraction of the non-W backgrounds.

The ttV and tri-boson background yields, along with their uncertainties and $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape, are derived purely from simulation. The samples are generated by `Sherpa`. The number of

Channel	Yield
$ee\nu\nu$	$0.38 \pm 0.1 \pm 0.04$
$\mu\mu\nu\nu$	$1.27 \pm 1.53 \pm 0.75$

Table 6.23: W + jets data-driven estimation in the electron and muon channel as obtained from the data-driven technique. The systematic uncertainties are estimated using the differences in the W + jets and Z + jets MC-based fake-factors and the uncertainty associated with the subtraction of MC based non-W backgrounds.

Channel	Yield
$ee\nu\nu$	$1.12 \pm 0.04 \pm 0.09$
$\mu\mu\nu\nu$	$1.03 \pm 0.04 \pm 0.08$

Table 6.24: ttV and tri-boson background contributions to the ggF signal region. The yield, statistical and systematic uncertainties are obtained from simulation.

background events in the signal region are provided in Table 6.24.

6.2 Systematic Uncertainty

The uncertainties that impact the analysis come from event reconstruction, theory calculation, data-driven estimation of backgrounds as well as luminosity determination. For the signal processes, the reconstruction, acceptance and theory uncertainties play roles in the determination of $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape while for backgrounds, estimated from simulation, both yield and shape can be affected by reconstruction and theory uncertainties as well as the luminosity uncertainty. For backgrounds estimated with data, the yields are fully determined in data, while the shapes are taken from data where available. Table 6.25 summarizes the implementations of various types of uncertainties on the physics processes involved in this analysis.

Process	Lumi.	Reconstruction	Data-Driven	Theory
Signal	-	Acceptance, Shape	-	Acceptance
ZZ	Yield	Yield, Shape	-	Yield, Shape
WZ	-	Shape	Yield	Shape
Z + jets	-	Shape	Yield, Shape	-
$t\bar{t}$ /WW/Wt/$Z \to \tau\tau$	-	-	Yield, Shape	-
W + jets	-	-	Yield, Shape	-

Table 6.25: Impact of various uncertainty sources on the physics processes involved in this analysis.

6.2.1 Experimental uncertainties

This section describes the detector systematics of simulated signals and backgrounds. The experimental systematics encompass the uncertainties on the different objects i.e. electrons (6.2.1.1), muons (6.2.1.2), jets (6.2.1.3), including flavour tagging (6.2.1.4) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (6.2.1.5) but also include luminosity (6.2.1.6) and PU (6.2.1.7). Table 6.26 provides a summary description of each systematic variation. The largest experimental uncertainties for the largest background in the ggF signal region are provided in Table 6.27. In addition, in Table 6.28 the largest uncertainties for the 600 GeV ggF signal are given.

6.2.1.1 Electron uncertainties

The reconstruction, calibration and derivation of efficiency corrections for electrons was detailed in section 4.2.1. The uncertainties on the electron objects are estimated from the energy scale and the electron efficiency factors.

The uncertainties on the resolution and scale are provided by the ATLAS $e\gamma$ group [186]. The analysis is only mildly sensitive to the energy scale of electrons for which only a special simplified correlation model of the uncertainties is used. In this scheme all effects are considered fully correlated in detector η and consequently summed in quadrature. With this simplification only two systematic variations are provided, one for the scale and one for the resolution.

Uncertainties are assigned on the reconstruction, identification, isolation and trigger efficiency factors which reflect the observed dis-agreement between data and simulation [117].

Overall, the electron systematic uncertainties have a small impact on the signal sensitivity.

6.2.1.2 Muon uncertainties

The reconstruction, calibration and derivation of efficiency corrections for muons were detailed in section 4.2.2. The uncertainties on the muon objects are estimated from several sources which are provided by the ATLAS muon combined performance (MCP) group [187].

Systematics are added to the resolution smearing of the muon track in the ID and MS chambers, due to small mis-alignments present between them. A scale systematic is assigned from variation of the muon momentum scale. Two additional scale systematic variations are added, due to the muon sagitta bias [122]. The first is for the resolution, based on the residual charge-dependent bias after correction while the second is on ρ, based on recombination of corrections on the Z scale.

Systematic variation	Description	Section
Luminosity		
TotalLumi	Total integrated Luminosity	6.2.1.6
Pile up reweighting		
PRW_DATASF	pileup μ profile reweighting	6.2.1.7
Electrons		
EG_RESOLUTION_ALL	electron resolution smearing	6.2.1.1
EG_SCALE_ALL	electron energy calibration	6.2.1.1
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	electron identification efficiency	6.2.1.1
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	electron reconstruction efficiency	6.2.1.1
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	electron isolation efficiency	6.2.1.1
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	electron trigger efficiency	6.2.1.1
Muons		
MUON_ID	muon ID track smearing	6.2.1.2
MUON_MS	muon MS track smearing	6.2.1.2
MUON_SCALE	muon momentum calibration	6.2.1.2
MUON_EFF_STAT	muon reconstruction efficiency	6.2.1.2
MUON_EFF_SYS	muon reconstruction efficiency	6.2.1.2
MUON_EFF_STAT_LOWPT	muon reconstruction efficiency	6.2.1.2
MUON_EFF_SYS_LOWPT	muon reconstruction efficiency	6.2.1.2
MUON_EFF_TrigStatUncertainty	muon trigger efficiency	6.2.1.2
MUON_EFF_TrigSystUncertainty	muon trigger efficiency	6.2.1.2
MUON_ISO_STAT	muon isolation efficiency	6.2.1.2
MUON_ISO_SYS	muon isolation efficiency	6.2.1.2
MUON_TTVA_STAT	muon TTVA efficiency	6.2.1.2
MUON_TTVA_SYS	muon TTVA efficiency	6.2.1.2
MUON_SAGITTA_RESBIAS	muon sagitta bias	6.2.1.2
MUON_SAGITTA_RHO	muon sagitta bias	6.2.1.2
Jets		
JET_GroupedNP_[1-3]	Jet energy scale	6.2.1.3
JET_EtaIntercalibration_NonClosure	Jet energy scale	6.2.1.3
JET_JER_SINGLE_NP	Jet energy resolution	6.2.1.3
JET_JvtEfficiency	Jet jvt efficiency	6.2.1.3
Flavor tagging		
FT_EFF_B_systematics	b-jet tagging efficiency	6.2.1.4
FT_EFF_C_systematics	c-jet tagging efficiency	6.2.1.4
FT_EFF_Light_systematics	light-jet tagging efficiency	6.2.1.4
FT_EFF_extrapolation	b-jet tagging efficiency with high pt extrapolation	6.2.1.4
FT_EFF_extrapolation_from_charm	c-jet tagging efficiency with high pt extrapolation	6.2.1.4
$E_{\mathrm{T}}^{\mathrm{miss}}$		
MET_SoftTrk_Scale	$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term scale	6.2.1.5
MET_SoftTrk_ResoPara	$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term resolution	6.2.1.5
MET_SoftTrk_ResoPerp	$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term resolution	6.2.1.5
Data-driven, Theory and VBF systematics		
SigPDF	Signal PDF systematic	6.2.1.8
SigQCD	Signal QCD scale systematic	6.2.1.8
SigSHW	Signal shower scale systematic	6.2.1.8
QCDscale_VV	$qq \to ZZ$ QCD scale systematic	6.2.1.8
pdf_qq	$qq \to ZZ$ PDF systematic	6.2.1.8
VBF_qqZZ_Model	$qq \to ZZ$ VBF modelling systematic	6.2.1.8
ggZZ_Theo	$gg \to ZZ, k-$factor uncertainty	6.2.1.8
WZ_ScaleStat	WZ scale statistical uncertainty	6.2.1.8
WZ_ScaleSys	WZ scale systematic uncertainty	6.2.1.8
WZ_ScaleInEff	WZ inefficiency systematic uncertainty	6.2.1.8
emu_Stat	$e\mu$ statistical uncertainty	6.2.1.8
emu_Sys	$e\mu$ systematic uncertainty	6.2.1.8
emu_Shape	$e\mu$ shape uncertainty	6.2.1.8
ABCD_Stat	Z + jets ABCD method statistical uncertainty	6.2.1.8
ABCD_Sys	Z + jets ABCD method systemaitc uncertainty	6.2.1.8
ABCD_Shape	Z + jets ABCD method shape uncertainty	6.2.1.8
Wjets_Stat	W + jets statistical uncertainty	6.2.1.8
Wjets_Sys	W + jets systematic uncertainty	6.2.1.8

Table 6.26: The summary of experimental and theory systematic uncertainties

Systematic Uncertainty	$ee\nu\nu$ channel		$\mu\mu\nu\nu$ channel	
	up (%)	down (%)	up (%)	down (%)
JET_GroupedNP_1	0.98	-1.56	1.34	-1.02
JET_JvtEfficiency	1.06	-1.05	1.03	-1.02
EG_SCALE_ALL	1.01	-0.72	0.02	-0.03
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	1.75	-1.74	0.00	0.00
MUON_EFF_SYS__1	0.00	0.00	1.82	-1.80

Table 6.27: The fractional experimental uncertainties on the estimation of $qq \to ZZ$ background. Presented are the largest uncertainties affecting this background from experimental sources. They include the uncertainties on the JES, jet JVT efficiency, $e\gamma$ scale and ID efficiency and muon reconstruction efficiency. The uncertainties are shown separately for $ee\nu\nu$ and $\mu\mu\nu\nu$ channels, and the sign of the uncertainties reflects the direction to which the corresponding variation of uncertainty can change the estimated yields.

Systematic Uncertainty	$ee\nu\nu$ channel		$\mu\mu\nu\nu$ channel	
	up (%)	down (%)	up (%)	down (%)
JET_GroupedNP_1	1.70	-1.72	1.75	-2.14
JET_JvtEfficiency	1.63	-1.60	1.62	-1.59
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	1.89	-1.88	0.00	0.00
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	1.56	-1.56	0.00	0.00
FT_EFF_Light_systematics	1.13	-1.12	1.11	-1.09
PRW_DATASF	0.50	-0.71	0.70	-1.10

Table 6.28: The fractional experimental uncertainties on the predicted yields of the heavy Higgs signal process with m_{H} =600GeV. Presented are the largest uncertainties affecting this background from experimental sources. They include the uncertainties on the JES, jet JVT efficiency, $e\gamma$ ID and isolation efficiency, light jet flavour tagging and PU uncertainties. The uncertainties are shown separately for $ee\nu\nu$ and $\mu\mu\nu\nu$ channels, and the sign of the uncertainties reflects the direction to which the corresponding variation of uncertainty can change the estimated yields.

Uncertainties are assigned to the reconstruction, identification, and isolation efficiencies, similarly to the electron. A systematic is attached to the track-to-vertex-association efficiency. Two different components of the systematics variations are used for each correction, the statistical and systematic error on the scale factors. Extra systematics were added for the low p_{T} component parts of the scale factors [120].

Systematic uncertainties were added for the trigger scale factor, similarly to the reconstruction case with both statistical and systematic uncertainties on the trigger efficiency factor.

6.2.1.3 Jet uncertainties

The reconstruction, calibration and derivation of JVT efficiency corrections for jets was detailed in section 4.2.3. The jet systematics constitute the largest contribution to the

object detector systematics. They are provided by the official ATLAS Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ performance group [188].

The derivation of the JES uncertainty is described in [126]. The JES uncertainty in the analysis is obtained from a strongly reduced model of nuisance parameters, recommended for most analysis studies. An extra uncertainty is added, due to an observed dis-agreement in the $2.4 < \eta < 2.5$ region. A one sided systematic is provided for the jet resolution and energy smearing. Lastly, a systematic is provided for the efficiency of the JVT selection.

6.2.1.4 Flavor tagging uncertainties

The envelope method for the flavour systematics considers three uncertainties for the heavy b, c and light flavour jets. They correspond to the uncertainties on the tagging efficiencies for the three jet types. Supplementary systematics are added for jets with very large p_{T} (> 300 GeV) where a data correction cannot be applied and the efficiency is exclusively estimated from simulation [137].

6.2.1.5 $\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$ uncertainties

The only uncertainties provided by $E_{\mathrm{T}}^{\mathrm{miss}}$ group [189] are coming from the TST. The uncertainties on other $E_{\mathrm{T}}^{\mathrm{miss}}$ terms are propagated from their respective object uncertainties. Three uncertainties on the soft terms are provided, two for the resolution and one for the scale [190]. The resolution uncertainties are taken from the data/MC differences of the variances of the projections parallel and perpendicular to p_T^{hard}. The scale systematic is taken for the data/MC difference of the soft term projection along the parallel component of p_T^{hard}.

6.2.1.6 Luminosity uncertainty

The luminosity uncertainty is $\pm 3.2\%$ for the combined 2015 and 2016 dataset. The uncertainty is partially correlated across both run periods [191]. The luminosity error applied to both signal and background simulated samples.

6.2.1.7 PU uncertainty

The PU uncertainty is deduced from variation of $\langle\mu\rangle$ which changes the PU re-weighting scheme. The differences between the up variation and the down variation with respect to the nominal are assigned as a systematic [192].

6.2.1.8 Data-driven and VBF uncertainties

For data-driven backgrounds, statistical and systematic uncertainties are obtained from the differences in yields between data and simulation. The shape is extracted from data if enough statistical precision is left, otherwise a shape from MC is used.

In the VBF phase space, the statistical and systematic uncertainties are derived from MC based transfer factors. Modelling systematics, if not negligible, are found from comparisons between different monte-carlo generators.

The $qq \to ZZ$ and $gg \to ZZ$ backgrounds are deduced from simulation, with experimental systematics obtained from the detector uncertainties. In the VBF region, an additional systematic is taken from the modelling of the VBF acceptance from the `Powheg` and `Sherpa` generators. More details are described in section 6.1.2.

The WZ background uses a data-driven statistical and systematic uncertainty on the scale factor, which is used to fix yield differences between MC and data. The uncertainties are then translated into a systematic on the background yields. An extra uncertainty is provided by the inefficiency scale factor for mis-reconstruction the third lepton. For the VBF analysis, the systematics are inferred from the transfer factor with an added modelling uncertainty for the `Powheg` and `Sherpa` generators. The WZ estimation and uncertainties are covered in section 6.1.3.

The $WW, t\bar{t}, Wt, Z \to \tau\tau$ backgrounds use a data-driven method in an $e\mu$ control region. The systematic is added from the non-closure of the $e\mu$ background estimation method and statistical uncertainty on the efficiency factor, ϵ. The $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution is derived in data while the shape systematic is taken from differences obtained by changing ϵ by $\pm 1\sigma$. For the VBF category a transfer factor is used to scale the yield while the systematics are found from MC and generator differences. The details are covered in section 6.1.4.

The yield of the Z + jets background is determined via a data-driven method. The uncertainty on the Z + jets yield is in all taken from four separate sources. The overall uncertainty on the signal region yield is on the order of 70%. In the VBF category, similar to the other backgrounds, an MC estimated transfer factor is used to derive the yields in the VBF category. Modelling uncertainties were not added as they were found to be negligible. The Z + jets description is covered in section 6.1.5.

The W + jets background is obtained from data in section 6.1.6, but it contributes $\sim 1\%$ to the overall background. Other backgrounds are taken from simulation.

For Higgs and Graviton signal samples, only the experimental detector uncertainties are considered.

6.2.2 Theoretical uncertainties

Theoretical uncertainties reflect the particular choice of various parameters such as MC generator types, PDFs, QCD scales and PS algorithms. They are added for processes estimated fully or partially from simulation i.e. the signal and the ZZ/WZ backgrounds.

The theory uncertainties on the Higgs samples stem from the particular event generation parameters. The QCD scale uncertainty is derived by variations of the renormalization scale, μ_R, and factorization scale, μ_F from the nominal $\mu_R = \mu_F = m_{ZZ}$ between 0.5 and 2.0. The variation forms an envelope over the mass range which is used to determine the uncertainty for each hypothetical Higgs mass. The PDF uncertainty is determined by evaluating the

	PDF		QCD		PS	
	ee (%)	$\mu\mu$ (%)	ee (%)	$\mu\mu$ (%)	ee (%)	$\mu\mu$ (%)
NW300 GeV	$+0.9-0.7$	$+1.2-1.4$	$+0.5-0.5$	$+0.6-0.8$	$+2.6-2.6$	$+9.0-9.0$
NW400 GeV	$+0.7-0.9$	$+0.8-0.9$	$+0.4-0.5$	$+0.4-0.4$	$+1.0-1.0$	$+3.0-9.0$
NW500 GeV	$+0.6-0.8$	$+0.6-0.8$	$+0.4-0.4$	$+0.3-0.4$	$+1.0-1.0$	$+2.5-2.5$
NW600 GeV	$+0.5-0.7$	$+0.5-0.7$	$+0.3-0.4$	$+0.3-0.4$	$+0.8-0.8$	$+2.5-2.5$
NW700 GeV	$+0.5-0.8$	$+0.4-1.1$	$+0.3-0.4$	$+0.2-0.3$	$+0.2-0.2$	$+0.6-0.6$
NW800 GeV	$+0.4-0.6$	$+0.6-0.6$	$+0.2-0.3$	$+0.2-0.3$	$+1.7-1.7$	$+0.7-0.7$
NW900 GeV	$+0.3-0.6$	$+0.3-0.5$	$+0.2-0.3$	$+0.2-0.3$	$+0.8-0.8$	$+0.1-0.1$
NW1000 GeV	$+0.3-0.5$	$+0.3-0.7$	$+0.2-0.2$	$+0.2-0.2$	$+1.0-1.0$	$+0.6-0.6$
NW1200 GeV	$+0.2-0.3$	$+0.2-0.4$	$+0.2-0.2$	$+0.1-0.2$	$+0.2-0.2$	$+1.0-1.0$
NW1400 GeV	$+0.1-0.2$	$+0.2-0.3$	$+0.1-0.2$	$+0.1-0.2$	$+0.4-0.4$	$+3.2-3.2$

Table 6.29: Theory uncertainties of PDF, QCD and PS variation on the acceptance of ggF signals from mass point 300 GeV to 1400 GeV in ggF category.

	PDF		QCD		PS	
	ee (%)	$\mu\mu$ (%)	ee (%)	$\mu\mu$ (%)	ee (%)	$\mu\mu$ (%)
NW300 GeV	$+0.0-0.0$	$+2.7-0.0$	$+1.4-1.4$	$+1.4-0.0$	$+2.6-2.6$	$+9.0-9.0$
NW400 GeV	$+0.5-0.2$	$+0.6-0.0$	$+0.3-0.3$	$+1.0-0.8$	$+1.0-1.0$	$+3.0-3.0$
NW500 GeV	$+0.5-0.2$	$+0.4-0.2$	$+0.2-0.3$	$+0.5-0.5$	$+1.0-1.0$	$+2.5-2.5$
NW600 GeV	$+0.4-0.2$	$+0.5-0.2$	$+0.3-0.4$	$+0.3-0.3$	$+0.8-0.8$	$+2.5-2.5$
NW700 GeV	$+0.3-0.6$	$+0.3-0.4$	$+0.3-0.8$	$+0.1-0.3$	$+0.2-0.2$	$+0.6-0.6$
NW800 GeV	$+0.4-0.3$	$+0.3-0.2$	$+0.3-0.6$	$+0.1-0.1$	$+1.7-1.7$	$+0.7-0.7$
NW900 GeV	$+0.4-0.2$	$+0.3-0.3$	$+0.3-0.7$	$+0.3-0.7$	$+0.8-0.8$	$+0.1-0.1$
NW1000 GeV	$+0.3-0.3$	$+0.5-0.2$	$+0.7-1.3$	$+0.7-1.3$	$+1.0-1.0$	$+0.6-0.6$
NW1200 GeV	$+0.3-0.3$	$+0.5-0.2$	$+0.7-1.3$	$+0.7-1.3$	$+1.0-1.0$	$+0.6-0.6$
NW1400 GeV	$+0.6-1.0$	$+2.1-0.7$	$+1.8-0.8$	$+5.3-1.8$	$+0.4-0.4$	$+3.2-3.2$

Table 6.30: Theory uncertainties of PDF, QCD and PS variation on acceptance of VBF signals from mass point 300 GeV to 1400 GeV in VBF category.

envelope formed from the variation of the different PDF sets from the nominal `CT10nlo` PDF. The PS uncertainty is found from changing the different parameters of the `AZNLO` tune such as the multi-parton interaction and final-state radiation models in the `Pythia8` setup. Variations of the shape were found to be negligible. The effects of the theory systematics on the signal samples are summarized in Table 6.29 (6.30) for the ggF(VBF) signal. For the Graviton signal, the PDF uncertainty on the `NNPDF23LO` is evaluated and its effects are found to be negligible. The theoretical uncertainties on the Gravitons signal sample are not applied as they are negligible in size to the statistical uncertainty in data, which dominates the overall sensitivity.

For the $qq \to ZZ$ background, the PDF and QCD uncertainties are found from particle level samples produced from the `Powheg` generator. The evaluation method is the same as described above for the signal samples. The average size of the PDF uncertainty is 2.4% and QCD scale systematic is 3.6%. Shape systematics are found to be small ($< 1\%$). The PS systematic is determined from the comparison of the `Pythia8` and `Herwig` generators [193]. The variation of the PDF and QCD systematic as a function of the transverse mass is shown in Figure L.2.

The dominant theoretical systematics for $gg \to ZZ$ is due to the $k-$factor uncertainty yielding a 60% uncertainty on the acceptance which is applied to the whole $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution.

The WZ background yield is data-driven, while the shape is taken purely from simulation. The PDF and QCD theory shape uncertainties are small ($< 1\%$), much smaller than the uncertainties in the data-driven estimate. The effect of the PDF and QCD systematic on the m_{WZ} is shown in Figure L.1.

Chapter 7

Results

The strategy, modelling, background estimation and their associated uncertainties, described in chapter 5 lead to the next important step in the analysis, the interpretation of the results from the signal region.

The number of events derived by data-driven methods and simulation are summarized in section 7.1. The statistical framework used to perform the fit of the data to signal hypothesis is described in section 7.2. In the absence of statistically significant excesses, exclusion upper limits are placed on the ggF and VBF Higgs production cross section, which are detailed in section 7.3. The impact of the systematic uncertainties on the parameter-of-interest (POI), μ, is explored in section 7.4. The constraints from the fitting procedure on the transverse mass distributions and yields are showed in section 7.5.

A combination between the $H \to ZZ \to 4\ell$ and $H \to ZZ \to \ell\ell\nu\nu$ searches was performed with 2015+2016 data to increase the sensitivity to Higgs production. The $H \to ZZ \to 4\ell$ analysis strategy and results, combination framework and derived exclusion limits are overviewed in section 7.6, which also includes an interpration on spin-2 RS graviton production.

7.1 Yields in the signal region

The description behind the estimation of the yields in the signal region were described earlier in this thesis. The search strategy for a Higgs boson with large mass in the $H \to ZZ \to \ell\ell\nu\nu$ decay mode and the selections applied to isolate the Higgs signal are defined in section 5.2. The data, corresponding to 36.1 fb^{-1} integrated luminosity, was recorded by the ATLAS detector in the 2015 and 2016 LHC run period, more details can be found in section 3.7. The signal yields are taken from the simulated samples for either the spin-0 Higgs interpretation or spin-2 graviton interpretation. The background contributions in the analysis are a mixture

of data-driven and/or simulation based yields and shape. The specifics behind the production of monte-carlo simulation samples are summarized in section 5.3. The details for each estimation methodology are covered in section 6.1.

The discriminant in the $H \to ZZ \to \ell\ell\nu\nu$ analysis is the transverse mass, $m_{\mathrm{T}}^{\mathrm{ZZ}}$. It is defined in equation 5.1 as a function of $m_{\mathrm{Z}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{Z}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$. The distribution in the ggF topology for all the backgrounds and a single signal ($m_{\mathrm{H}} = 600$ GeV) is given in Figure 7.2 in both lepton channels.

The individual background yields, which make up the total background contribution in the signal region are visualised as a pie chart in Figure 7.1. The separate processes can be seen to make up a different percentage of the total background depending on the event topology. The breakdown by percentige for each of the processes are provided in Table 6.1.

The largest SM process in the signal region is the irreducible ZZ background. The $qq \to ZZ$ is the largest mode of production of these events, contributing $\sim 84\%$ to the ZZ production cross section, while the rest originate from the virtual loop $gg \to ZZ$ process. The yield and $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape of the ZZ background are derived from simulation, with an applied $k-$factor to correct for higher-order processes which modify to the naive tree level prediction. The ZZ process contributes $\sim 60\%$ in the ggF and $\sim 45\%$ in the VBF categories, respectively.

The second largest background is the WZ background, which mimics the $\ell\ell\nu\nu$ signature through the mis-reconstruction of the third lepton. Its yield is estimated from data, in a three lepton control region, while its transverse mass shape is taken from MC. Its overall contribution is $\sim 30\%$ in the signal regions.

The next major SM decay mode is the Z + jets background, which can resemble the di-lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ state if jet mis-reconstruction is present. If a jet (or jets) are mis-reconstructed or mis-calibrated a non-negligible $E_{\mathrm{T}}^{\mathrm{miss}}$ value can be formed, which may survive the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection requirement. Several quality selections are applied to minimize this effect, since the Z + jets process has large associated systematic uncertainties. The yield is derived from a modified ABCD method which uses Boolean variables to define a control region. The transverse mass shape is obtained from data in a one side-band method.

The four processes, $t\bar{t}/WW/Wt/$ and $Z \to \tau\tau$, are estimated in the signal region from an $e\mu$ control region. Both the background yield and $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape are determined from data.

The W + jets background is present in the $\ell\ell\nu\nu$ final state through jets mis-reconstructed as leptons. The yield is extracted from the number of data events in a W + jets control region which is scaled by fake factor. Other processes i.e. ttV and tri-boson backgrounds are determined by simulation.

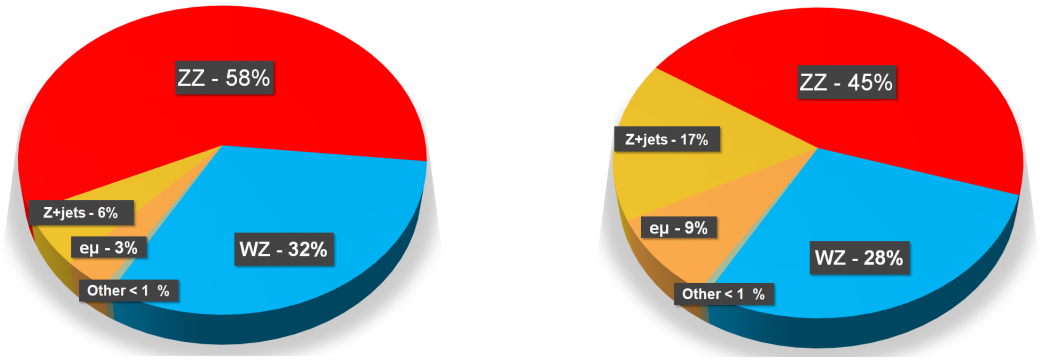


Figure 7.1: Each background contribution present in the $H \to ZZ \to \ell\ell\nu\nu$ signal region is visualized in the form of a pie chart. The left chart contains the background proportions in the ggF category while the right presents the proportions in the VBF category.

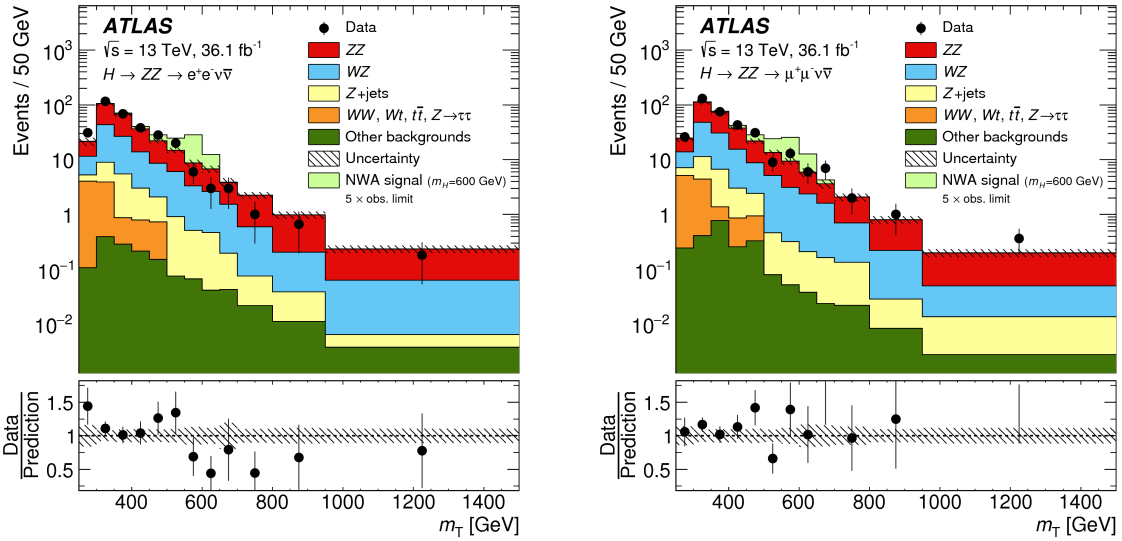


Figure 7.2: Transverse mass distribution in the analysis search. Left: $H \to ZZ \to eenn$ channel, Right: $H \to ZZ \to \mu\mu\nu\nu$ channel. The backgrounds are determined following the description in Section 6.2 and the last bin includes the overflow. The simulated $m_{\mathrm{H}} = 600$ GeV signal is normalized to a cross section corresponding to five times the observed limit. The error bars on the data points indicate the statistical uncertainty and markers are drawn at the bin center. The systematic uncertainty in the prediction is shown by the hatched band. The lower panels show the ratio of data to prediction. Figure is taken from [6].

The signal region yields for the ggF part of the analysis are presented in Table 7.1 while the yields in the VBF part are shown in Table 7.2.

No evidence of a statistically significant excess is found in the $H \to ZZ \to \ell\ell\nu\nu$ analysis.

	$ee\nu\nu$ channel		$\mu\mu\nu\nu$ channel	
Data	324		357	
	Simulation	Data-driven	Simulation	Data-driven
H($m_H = 300\,\text{GeV}$)	46.6 ±1.27±2.92	-	51.9 ±1.4 ±5.36	-
H($m_H = 600\,\text{GeV}$)	74.22±0.71±3.74	-	70.74 ±0.68±4.03	-
H($m_H = 1\,\text{TeV}$)	6.43 ±0.06±0.37	-	5.62 ±0.05±0.32	-
Graviton($m_G = 600\,\text{GeV}$)	155.3±2.66±8.11	-	150.4 ±2.6 ±7.32	-
$qq \to ZZ$	161.2 ± 2.6 ± 18.6	-	163.8 ± 2.5 ± 18.9	-
$gg \to ZZ$	16.0 ± 0.4 ± 9.4	-	15.9 ± 0.4 ± 9.4	-
WZ	71.5 ± 1.7 ± 4.5	92.8 ± 2.2 ± 4.3	75.6 ± 1.8 ± 4.4	99.5 ± 2.3 ± 3.2
$Z \to \tau\tau$/Top/WW/Wt	9.9 ± 1.6 ± 2.4	9.2 ± 2.2 ± 1.4	14.0 ± 2.5 ± 2.7	10.7 ± 2.5 ± 0.9
Z+jets	8.9 ± 3.2 ± 8.1	16.9 ± 0.6 ± 11.4	18.6 ± 5.6 ± 7.5	19.7 ± 0.6 ± 17.2
W+jets	0 ± 0 ± 0	0.4 ± 0.1 ± 0.04	0 ± 0 ± 0	1.3 ± 1.5 ± 0.8
Other	1.12 ± 0.04 ± 0.09	-	1.03 ± 0.04 ± 0.08	-
Total Bkg.	268.6 ±4.8 ±22.9	297.2±4.1±24.2	289.0 ±6.9± 23.0	311.5±4.6± 27.4

Table 7.1: The observed data and MC expected (scaled to 36.1 fb^{-1}) yields in $ee\nu\nu$ and $\mu\mu\nu\nu$ signal regions after full ggF event selections. The first and second errors represent the statistical and systematic uncertainties, respectively. The total background prediction is given in the last row.

	$ee\nu\nu$ channel		$\mu\mu\nu\nu$ channel	
Data	4		5	
	Simulation	Data-driven	Simulation	Data-driven
H($m_H = 300\,\text{GeV}$)	0.12 ±0.01 ±0.03	-	0.11 ±0.02 ± 0.03	-
H($m_H = 600\,\text{GeV}$)	2.34 ±0.04 ±0.14	-	2.13 ±0.04 ± 0.13	-
H($m_H = 1\,\text{TeV}$)	3.7 ±0.32 ±0.20	-	3.1 ±0.21 ± 0.23	-
$qq \to ZZ$	0.78 ± 0.19 ± 0.45	-	0.85 ± 0.16 ± 0.35	-
$gg \to ZZ$	0.23 ± 0.00 ± 0.23	-	0.23 ± 0.00 ± 0.23	-
WZ	0.66 ± 0.21 ± 0.29	0.63 ± 0.03 ± 0.19	0.69 ± 0.14 ± 0.23	0.66 ± 0.03 ± 0.19
$Z \to \tau\tau$/Top/WW/Wt	0.41 ± 0.31 ± 0.26	0.16 ± 0.14 ± 0.15	0.40 ± 0.33 ± 0.73	0.23 ± 0.20 ± 0.21
Z+jets	0.03 ± 0.29 ± 0.25	0.33 ± 0.08 ± 0.33	0.18 ± 0.28 ± 0.66	0.43 ± 0.12 ± 0.43
W+jets	0 ± 0 ± 0	0 ± 0 ± 0.	0 ± 0 ± 0	0. ± 0. ± 0.
Other	0.02 ± 0.01 ± 0.00	-	0.01 ± 0.00 ± 0.01	-
Total Bkg.	2.13 ± 0.51 ± 0.69	2.15 ± 0.36 ± 0.65	2.36 ± 0.48 ± 1.09	2.41 ± 0.26 ± 0.66

Table 7.2: The observed data and MC expected (scaled to 36.1 fb^{-1}) yields in $ee\nu\nu$ and $\mu\mu\nu\nu$ signal regions after full VBF selections. Estimates only have statistical errors shown for MC and data-driven estimations are shown with statistical and systematic uncertainties, respectively. The total background prediction is given in the second to last row.

7.2 Statistical interpretation

Most BSM analysis searches at the LHC involve counting experiments of the observed number of data events, n, which are compared with a background-only hypothesis and with

a signal+background hypothesis. The signal strength parameter, μ, represents the signal rate for any hypothetical new signal production modes. In the current scenario it represents a heavy Higgs boson decaying to two Z bosons. The $\mu = 0$ value corresponds to the background-only hypothesis, while $\mu = 1$ corresponds to the nominal signal hypothesis. The expected total number of events is then:

$$\mu S + B \tag{7.1}$$

where $S(B)$ is the total number of predicted signal (background) events.

In the $H \to ZZ \to \ell\ell\nu\nu$ analysis the statistical package used to obtain the limits is the `TRexFitter` framework [194] developed in `HistFactory`[195].

7.2.1 Likelihood function

In its most basic form, the likelihood function is just the conditional probability of a certain observed quantity e.g. the total number of observed events, n_{tot}, given some unknown parameters $(\mu, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ are a vector of nuisance parameters as will be explained in section 7.2.2. Quite often n_{tot} is a Poisson random variable with a certain mean value, ν, which follows a Poisson probability distribution:

$$\text{Pois}(n_{\text{tot}}|\nu) = \frac{(\nu)^{n_{\text{tot}}}}{n_{\text{tot}}!} \times e^{-\nu}. \tag{7.2}$$

The likelihood function for the VBF analysis will have the form of 7.2 as it only contains the measurement for the total number of entries. However for the ggF analysis an additional term is added to the likelihood function to include the shape information from the m_{T}^{ZZ} distribution in addition to the normalization from n_{tot}. The binned term follows a multinominal distribution, according to:

$$\frac{n_{\text{tot}}!}{n_1!...n_N!}\left(\frac{\nu_1}{n_{\text{tot}}}\right)^{n_1}...\left(\frac{\nu_N}{n_{\text{tot}}}\right)^{n_N}, \tag{7.3}$$

where N is the total number of bins. Since the n_{tot} is a Poisson random variable with a certain mean ν then the distribution can be approximated to:

$$\frac{n_{\text{tot}}!}{n_1!...n_N!}\left(\frac{\nu_1}{\nu}\right)^{n_1}...\left(\frac{\nu_N}{\nu}\right)^{n_N}. \tag{7.4}$$

Combining 7.1, 7.2 and 7.4 the total likelihood function then becomes

$$\mathcal{L}(\mu, \boldsymbol{\theta}) \sim \text{Pois}(n_{\text{tot}}|\mu S + B) \prod_{b \in \text{bins}}^{N} \left[\frac{\mu\nu_b^S(\boldsymbol{\theta}) + \nu_b^B(\boldsymbol{\theta})}{\mu S + B}\right]^{n_b}, \tag{7.5}$$

where n_b are the number of observed events and $\nu_b^S(\nu_b^B)$ are the expected number of signal (background) events in the bin b.

In practice, it is much simpler to work with the logarithm of the likelihood function, or log-likelihood, as it reduces the complicated product in $\mathcal{L}$ to a sum of all terms. Since the log function is monotonically increasing, the maximum/minimum of the function is unchanged.

7.2.2 Nuisance parameters

Adding systematic uncertainties in the likelihood is accomplished through a set of nuisance parameters (NPs) $\boldsymbol{\theta}$. The NPs are parametrized using unit gaussian functions ($\sigma = 1, \mu = 0$) centered on their nominal expectation values, $\widetilde{\boldsymbol{\theta}}$:

$$G(\widetilde{\theta}|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(\widetilde{\theta}-\theta)^2/2}. \tag{7.6}$$

These gaussian functions constrain $\boldsymbol{\theta}$ to their assigned values within their given uncertainties. The signal and background number of events is dependent on the systematic uncertainties and therefore on the NPs. The modification of the signal and backgrounds is:

$$S = S_0 \times \prod_{i=1}^{c} \beta(\theta_i), B = B_0 \times \prod_{i=1}^{c} \beta(\theta_i) \tag{7.7}$$

where S_0, B_0 are the nominal expectation values and c is the total number of uncertainties. The systematics are further divided into two general categories: normalization uncertainties and normalization+shape uncertainties. The former type of uncertainty only affect the overall yield and therefore are assigned a simpler prescription, namely $\beta(\theta) = c^{\theta}$ where c^{θ} is a log-normal distribution with c determined by measuring β at $\theta = \pm 1$. For the latter version of the uncertainties, the normalization and shape components are split into separate contributions such that variations on the shape do not effect the expected rate. The shape uncertainties are incorporated by using a vertical interpolation method through a truncated gaussian [156, 196]. This involves producing three templates per NP from the variation of the nominal estimate by $\pm 1\sigma$. More specifically, $\beta_{\text{shape}}(\theta) = 1 + \epsilon\theta$, where ϵ is obtained at $\theta = \pm 1$, is constrained by a unit gaussian. The truncation is added to avoid negative values of $\beta_{\text{shape}}(\theta)$. The statistical uncertainties from data-driven or MC estimated backgrounds are added as an auxiliary measurement $\widetilde{\theta}$ with expected number of events $a\theta$, where a is constant taken at the nominal value of $\widetilde{\theta}$. The likelihood function then becomes:

$$\mathcal{L}(n|\mu, \boldsymbol{\theta}) = \text{Pois}(n_{\text{tot}}|\mu S + B) \prod_{b \in \text{bins}}^{N} \left[\frac{\mu \nu_b^S(\boldsymbol{\theta}) + \nu_b^B(\boldsymbol{\theta})}{\mu S + B} \right]^{n_b} \times \prod_{i=1}^{c} G(\widetilde{\theta}|\theta) \tag{7.8}$$

where $\prod_{i=1}^{c} G(\widetilde{\theta}|\theta)$ are auxiliary measurements of the NPs incorporated in the likelihood through a product of unit gaussians.

7.2.3 Fitting procedure

The likelihood method allows to obtain the NPs by scanning for values which maximize the likelihood. In this manner, the maximum likelihood estimators ($\hat{\mu}$, $\hat{\boldsymbol{\theta}}$) are obtained, which represent the best fit values to the observation. The NPs broaden the likelihood distribution which reflects the available amount of information about a certain parameter. The constraint on the NPs is determined from a conditional fit, where the condition is given by the data observed in the signal region. The conditional parameters, $\hat{\hat{\boldsymbol{\theta}}}$, are defined to maximize the likelihood given a certain μ value, which is allowed to float in the fit. However, the POI is given several constraints. First, $\hat{\mu} \geq 0$, to ensure a physical value for μ and second, $\hat{\mu} \leq \mu$, to provide an upper limit on the production of new physics given the observation in data. For the hypothesis tests and upper limits, the signal strength is defined as:

$$\mu = \frac{\sigma \times \mathrm{BR}(H \to ZZ)}{\sigma_{\mathrm{SM}} \times \mathrm{BR}_{\mathrm{SM}}(H \to ZZ)} \tag{7.9}$$

where σ_{SM} and $\mathrm{BR}_{\mathrm{SM}}(H \to ZZ)$ are the production cross-section and branching fraction, respectively, for a SM Higgs decaying to a pair of Z bosons. For the specific analysis search topologies, the signal strength parameter becomes:

$$\mu_{\mathrm{ggF}} = \frac{\sigma_{\mathrm{ggF}} \times \mathrm{BR}(H \to ZZ)}{\sigma_{\mathrm{SM,ggF}} \times \mathrm{BR}_{\mathrm{SM}}(H \to ZZ)}, \; \mu_{\mathrm{VBF}} = \frac{\sigma_{\mathrm{VBF}} \times \mathrm{BR}(H \to ZZ)}{\sigma_{\mathrm{SM,VBF}} \times \mathrm{BR}_{\mathrm{SM}}(H \to ZZ)}. \tag{7.10}$$

The likelihoods of the separate $H \to ZZ \to ee\nu\nu$ and $H \to ZZ \to \mu\mu\nu\nu$ channels are simultaneously fitted to produce a total combined limit, which correlates the systematic uncertainties in each channel. It is essentially adding two observations together, which reduces the error from each individual measurement. The resulting combined fit and derived upper limit are more sensitive than each channel individually. Separate lepton channel fits are also produced as an important cross-check.

In the analysis, the fit procedure is performed independently for the ggF and VBF categories. In the $4\ell + \ell\ell\nu\nu$ combination (discussed in section 7.6), both event topologies from the $H \to ZZ \to \ell\ell\nu\nu$ search are combined with their corresponding partners from the $H \to ZZ \to 4\ell$ analysis and their systematic uncertainties are correlated.

The ggF topology consists of a binned transverse mass distribution which provides shape information in contrast to the VBF category where only a single bin category is defined. The choice of a single bin in the VBF category is motivated by the lack of statistical precision and incorrect shape modelling in the VBF phase space.

7.2.4 Analysis inputs

The inputs for the $H \to ZZ \to \ell\ell\nu\nu$ analysis are made up of one dimensional histograms of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution for the data, signal and background samples. For the ggF category the range of the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution extends up to 1.8 TeV, while the VBF category consists of a single $m_{\mathrm{T}}^{\mathrm{ZZ}}$ bin. The systematic histograms for the MC-based signals and backgrounds, which involve the experimental systematic uncertainties are outlined in section 6.2.

7.3 Test statistic and upper limits

The compatibility between data and the hypothesis in question is gauged through the test statistic. The definition of the test statistic is inspired by the Neyman-Person lemma [197], and is given by:

$$\widetilde{q}_{\mu}^{\mathrm{obs}} = -2 \times \ln\left(\frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}\right), \text{given } 0 \leq \hat{\mu} \leq \mu \tag{7.11}$$

where the denominator is the unconditional likelihood fit for the estimators $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$, which maximize $\mathcal{L}$. The numerator is the conditional likelihood fit, which contains the conditional NPs, $\hat{\hat{\boldsymbol{\theta}}}$, which maximizes $\mathcal{L}$ for a given μ. The likelihood ratio tests the compatibility between the observed value, $\hat{\mu}$, and the one given by a certain hypothesis μ. If the likelihood ratio approaches unity then $\widetilde{q}_{\mu} \to 0$ which implies that a good agreement between hypothesis and data is found and is non-zero for poorer agreement. The condition $\hat{\mu} \leq \mu$ translates into test hypothesis which looks for excesses in data over background.

The test statistic, $\widetilde{q}_{\mu}^{\mathrm{obs}}$, follows a probability density function (pdf), $f(\widetilde{q}_{\mu}|\mu^{'})$, for any given value of $\mu^{'}$. The pdfs can be computed using toy MC pseudo-datasets, but due to the computational intensity a useful asymptotic limit approximation is used to define a calculable formula [150].

The upper limits on a production process mean the maximum production rate supported by observation at a certain statistical level of confidence. They are derived with a modified frequentist confidence interval (CL_S) method [150, 198]. In the CL_S method, the compatibility of the background-only hypothesis, is tested against the signal+background hypothesis for different μ values. The limits are placed on μ and then defined as limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ using equation 7.9.

With the determination of the pdfs for the test-statistic, the $p-$values can be obtained. The $p-$value is the number which examines the compatibility between the observed test statistic,

$\widetilde{q}_\mu^{\text{obs}}$, with the hypothesis in question. For the signal+background hypothesis, it is:

$$p_\mu = P\left(\widetilde{q}_\mu \geq \widetilde{q}_\mu^{\text{obs}} | S+B\right) = \text{CL}_{S+B} = \int_{\widetilde{q}_\mu^{\text{obs}}}^{\infty} f(\widetilde{q}_\mu|\mu) d\widetilde{q}_\mu \tag{7.12}$$

while for the background-only hypothesis:

$$1-p_b = P\left(\widetilde{q}_\mu \geq \widetilde{q}_\mu^{\text{obs}} | B\right) = \text{CL}_B = \int_{\widetilde{q}_\mu^{\text{obs}}}^{\infty} f(\widetilde{q}_\mu|\mu=0) d\widetilde{q}_\mu \tag{7.13}$$

The CL_S is computed from:

$$\text{CL}_S = \frac{\text{CL}_{S+B}}{\text{CL}_B} = \frac{p_\mu}{1-p_b}. \tag{7.14}$$

If $\text{CL}_S \leq \alpha$, where α is the signficance level, then the signal is rejected with $(1-\alpha)$ confidence level. The conventional choice for α is 0.05, which leaves a 95% confidence level on the upper exclusion limits for μ. The 95% CL_S upper limits are obtained by finding the value of μ which gives $\text{CL}_S(\mu) = 0.05$.

Upper limits are placed on two kinds of test statistics, the observed ($\widetilde{q}_\mu^{\text{obs}}$) and the expected ($\widetilde{q}_\mu^{\text{A}}$). The limit on the expected test statistic shows the overall sensitivity of a given analysis. The limits are similarly calculated from the CL_S, but whereas the observed integrated over $\widetilde{q}_\mu^{\text{obs}}$, the expected integrates from the median significance $\text{med}[Z_\mu] = \sqrt{\widetilde{q}_\mu^{\text{A}}}$. The expected test statistic has a similar definition to the observed:

$$\widetilde{q}_\mu^{\text{A}} = -2 \times \ln\left(\frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\mu', \hat{\boldsymbol{\theta}})}\right), \text{given } 0 \leq \hat{\mu} \leq \mu. \tag{7.15}$$

The distribution of $\widetilde{q}_\mu^{\text{A}}$ is generated by an Asimov data-set, which is a pseudo-data-set created with a background-only hypothesis. $\pm 1\sigma$ and $\pm 2\sigma$ error bands, which correspond 68% and 95% confidence intervals are computed.

7.3.1 Observed exclusion limits

The exclusion limits on the ggF production cross section for a NWA Higgs in the range from 300 GeV to 1.4 TeV in m_{H} are presented in Figure 7.3 for both the individual $H \to ZZ \to ee\nu\nu$ and $H \to ZZ \to \mu\mu\nu\nu$ channels and their combination. Table M.1, in the appendix, contains the upper limits for each mass point. The corresponding exclusion limits on the VBF production cross section are provided in Figure 7.4 and their values for each mass point are shown in Table M.2.

To understand the behaviour of the limits in the ggF signal region, the attention is focused on the m_{T}^{ZZ} distribution in Figure 7.2. In the region between 400 and 500 GeV, in the

$H \to ZZ \to ee\nu\nu$ channel, a slight excess can be seen, which translates into the disagreement between observed and expected exclusion limits at $m_{\mathrm{H}} \sim 500$ GeV. For higher values of m_{H}, the observed limit is smaller than the expected limit due to MC predicting a higher yield than the one observed in the data in the high m_{T} bins. The situation in the $H \to ZZ \to \mu\mu\nu\nu$ channel is reversed since the total background yield is slightly less than the total number events in data. Overall the disagreements with the expected limit partially cancel out in the combined ggF limit. There, the observed and expected limits are compatible within $\pm 1\sigma$ of the error bands.

In the case of the VBF exclusion limits no shape information on the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ is available, as the analysis is strictly a single bin fit. Therefore the background contribution for each mass point is the same and it is only the signal contribution which is varied. There is an excess of data over the total background in both channels, as shown in Table 7.2. However, they are not statistically significant as the number of events predicted by MC are consistent within $\pm 1\sigma$ with data.

7.4 Nuisance parameter pulls and rankings

Assesing the impact of the systematics uncertainties is performed through examination of the effects of the NPs, $\boldsymbol{\theta}$ on the fit. The variation on the signal strength, μ, caused by the NPs is also investigated, to determine its effect on the signal/background yields and the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution.

In Table 7.3, all the nuisance parameters are listed and a brief description is given on how they are shared amongst the different background processes. Each reconstruction uncertainty is correlated for samples purely based in monte-carlo simulation with the theoretical uncertainty also accounted. The luminosity uncertainty is also accounted for each process, except the signal. Uncertainties from data-driven estimation only apply to their specific backgrounds. Uncertainties from statistical fluctuations are applied per bin for MC derived processes only.

The NPs are random variables, which are gaussian distributed. Their impact on the fit is checked through:

- the pulls of the NPs,
- the pre-/post-fit impact on μ
- ranking of systematics.

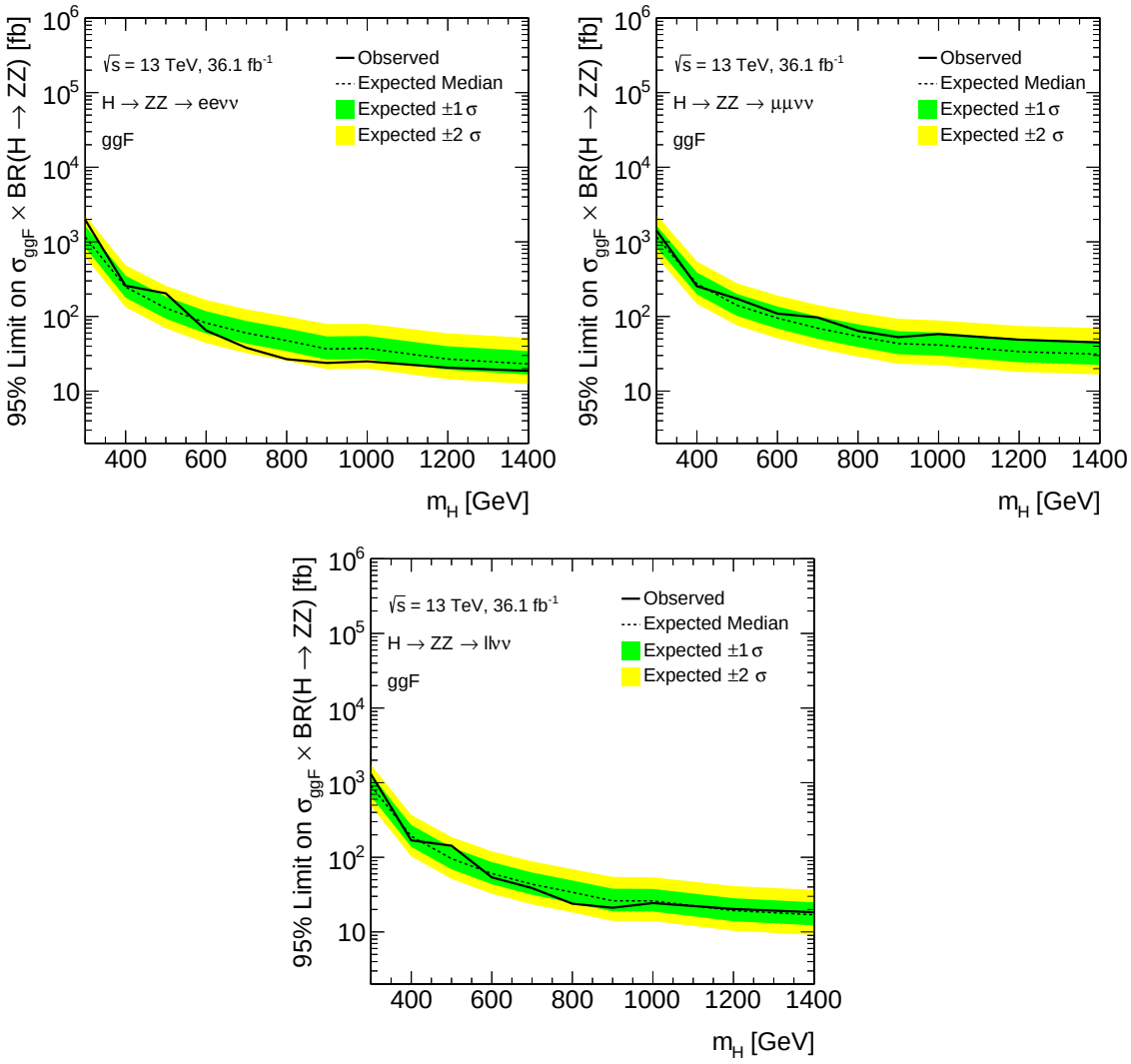


Figure 7.3: 95% C.L. limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ for a narrow width heavy Higgs boson produced in ggF channel as a function of its pole mass. The limits are derived using events in ee (top left), $\mu\mu$ (top right) and combined channels (bottom). The observed and expected limits are derived with considering the full uncertainties discussed in previous sections. The limits are derived for mass points between 300GeV and 1TeV with a 100GeV interval. Two additional mass points are added, 1.2TeV and 1.4TeV. And limit lines between each mass points are interpolated using an exponential function, or equivalently to say, linearly interpolated given a logarithmic y-axis. **Both the statistical and systematic uncertainties are considered in the limit setting**.

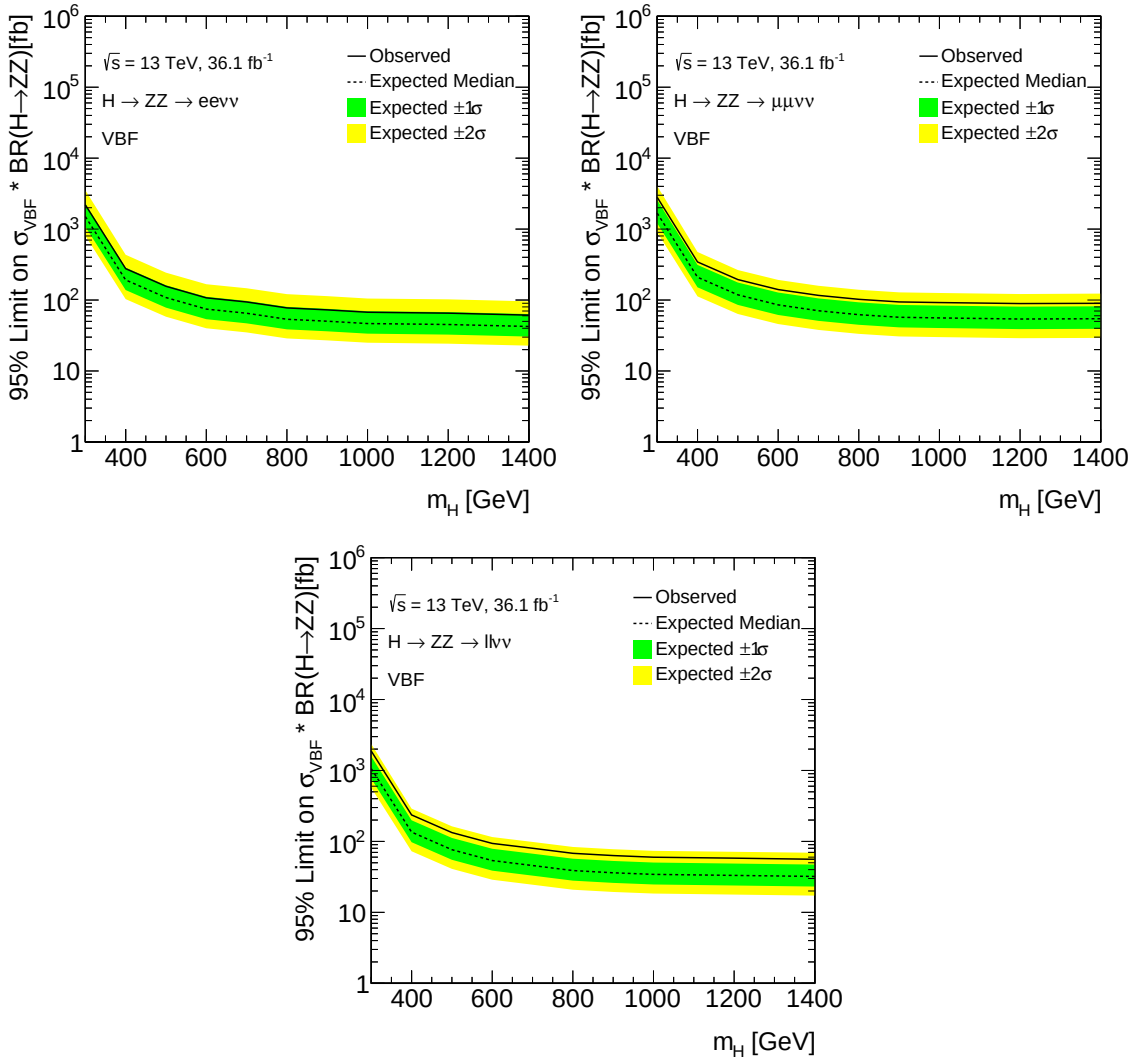


Figure 7.4: 95% C.L. limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ for a narrow width heavy Higgs boson produced in VBF channel as a function of its pole mass. The limits are derived using events in ee (top left), $\mu\mu$ (top right) and combined channels (bottom). The observed and expected limits are derived with considering the full uncertainties discussed in previous sections. The limits are derived for mass points between 300GeV and 1TeV with a 100GeV interval. And limit lines between each mass points are interpolated using an exponential function, or equivalently to say, linearly interpolated given a logarithmic y-axis. **Both the statistical and systematic uncertainties are considered in the limit setting**.

The pull of a NP represents the difference between the nominal NP value, θ_0, and the estimator value, $\hat{\theta}$, normalized by the constraint on the nominal NP, $\Delta\theta$. The normalization also ensures that the constraint takes on ± 1 values. The NP constraint is calculated during the fit which determines if the nuisance parameter is over-constrained by the fit i.e. if two or more auxiliary measurements place constraints on a particular systematic.

The pre-fit impact on μ is obtained by taking the NP at its value from the estimator, and shifting it by ± 1 and re-doing the fit. The change in the signal strength, $\Delta\mu$, is a consequence of the ± 1 NP variation. The post-fit impact on μ is obtained by fixing the NP at the estimator value, and shifting the value of μ by $\pm 1\sigma$. The ranking of the NPs are listed in descending order in the post-fit impact on μ.

The rankings of systematics are built for each Higgs mass point using the best fit performed to data. In Figure 7.5 the ranking of systematics are presented for several m_{H} points to follow the change of dominant systematics across the m_{H} range. Only the ten largest systematics are given. Systematics with a variation less than 0.1% do not have an impact on the limit and are removed during the fit. The equivalent plot from VBF is shown in Figure 7.6.

Lastly a correlation matrix of the signal strength and nuisance parameters, post-fit, is shown in Figure 7.7 for a 300 GeV mass and Figure 7.8 for a 600 GeV mass. These plots display correlation between systematics which leads to the change in the post-fit yields for the Higgs signal in question.

7.5 Post-fit impact on the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ shape and yield

As outlined in the last section, the values of the signal strength and NPs change during the fit. These in turn change the expected yields of the signal and background. Studying the differences between pre-fit and post-fit provides important insight of the effect of the likelihood fits and auxiliary measurements on the $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distribution.

The value on the POI are set to the their best fit value from the unconditional fit, $\hat{\mu}$. In Figure 7.9, the pre-fit and post-fit $m_{\mathrm{T}}^{\mathrm{ZZ}}$ distributions are shown for the $H \to ZZ \to ee\nu\nu$ channel for three different signal mass points. The expected contribution in the ratio plot, also takes into account the value of the signal yield, with the nominal signal hypothesis assumed. In the post-fit, the fit reduces the signal yield and increases the uncertainty band to cover any dis-agreements between data and the expected background. Figure 7.10 shows the corresponding scenario for the $H \to ZZ \to \mu\mu\nu\nu$ final state.

The effect of the fit is also examined through the change in the expected yields in the signal regions. In Table 7.4, the yields for the signal, S and signal+background, $S + B$ are given

Systematic uncertainty	Type
EG_RESOLUTION_ALL	electron reconstruction and correlated for all MC-based processes
EG_SCALE_ALL	
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	
EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	
MUONS_ID	muon reconstruction and correlated for all MC-based processes
MUONS_MS	
MUONS_SCALE	
MUON_EFF_STAT	
MUON_EFF_STAT_LOWPT	
MUON_EFF_SYS	
MUON_EFF_SYS_LOWPT	
MUON_EFF_TrigStatUncertainty	
MUON_EFF_TrigSystUncertainty	
MUON_ISO_STAT	
MUON_ISO_SYS	
FT_EFF_B	b-jet reconstruction and correlated for all MC-based processes
FT_EFF_C	
FT_EFF_Light	
FT_EFF_extrapolation	
FT_EFF_extrapolation_from_charm	
JET_EtaIntercalibration_NonClosure	jet reconstruction and correlated for all MC-based processes
JET_GroupedNP_1	
JET_GroupedNP_2	
JET_GroupedNP_3	
JET_JER_SINGLE_NP	
JvtEfficiency	
MET_SoftTrk_ResoPara	$E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction and correlated for all MC-based processes
MET_SoftTrk_ResoPerp	
MET_SoftTrk_Scale	
PRW_DATASF	PU reconstruction and correlated for all MC-based processes
LUMI	luminosity and correlated for all MC-based processes
SigQCD	theory signal systematics
SigPDF	
SigSHW	
VBF_qqZZ_Model	modelling uncertainty for $qqZZ$ background in VBF category
BkgQCD	theory systematics ($qqZZ$, WZ) and correlated for $qqZZ$ and WZ, only shape unc. for WZ
BkgPDF	
BkgSHW	
WZScale	background-WZ and correlated for all MC-based processes
ggZZTheo	theory-$ggZZ$ k-factor and only for $ggZZ$
ABCD_sys	Z+jets background and correlated for all MC-based processes
ABCD_stat	
ABCD_shape	
emu_sys	$e\mu$ background and correlated for all MC-based processes
emu_stat	
fakefactor_sys	W+jets background and correlated for all MC-based processes
ggZZMCStatChnXBinY	MC statistics, one nuisance parameter per bin
qqZZMCStatChnXBinY	
SigMCStatChnXBinY	
WZMCStatChnXBinY	
ZjetsMCStatChnXBinY	
emuMCStatChnXBinY	

Table 7.3: List of nuisance parameters considered in the limit setting.

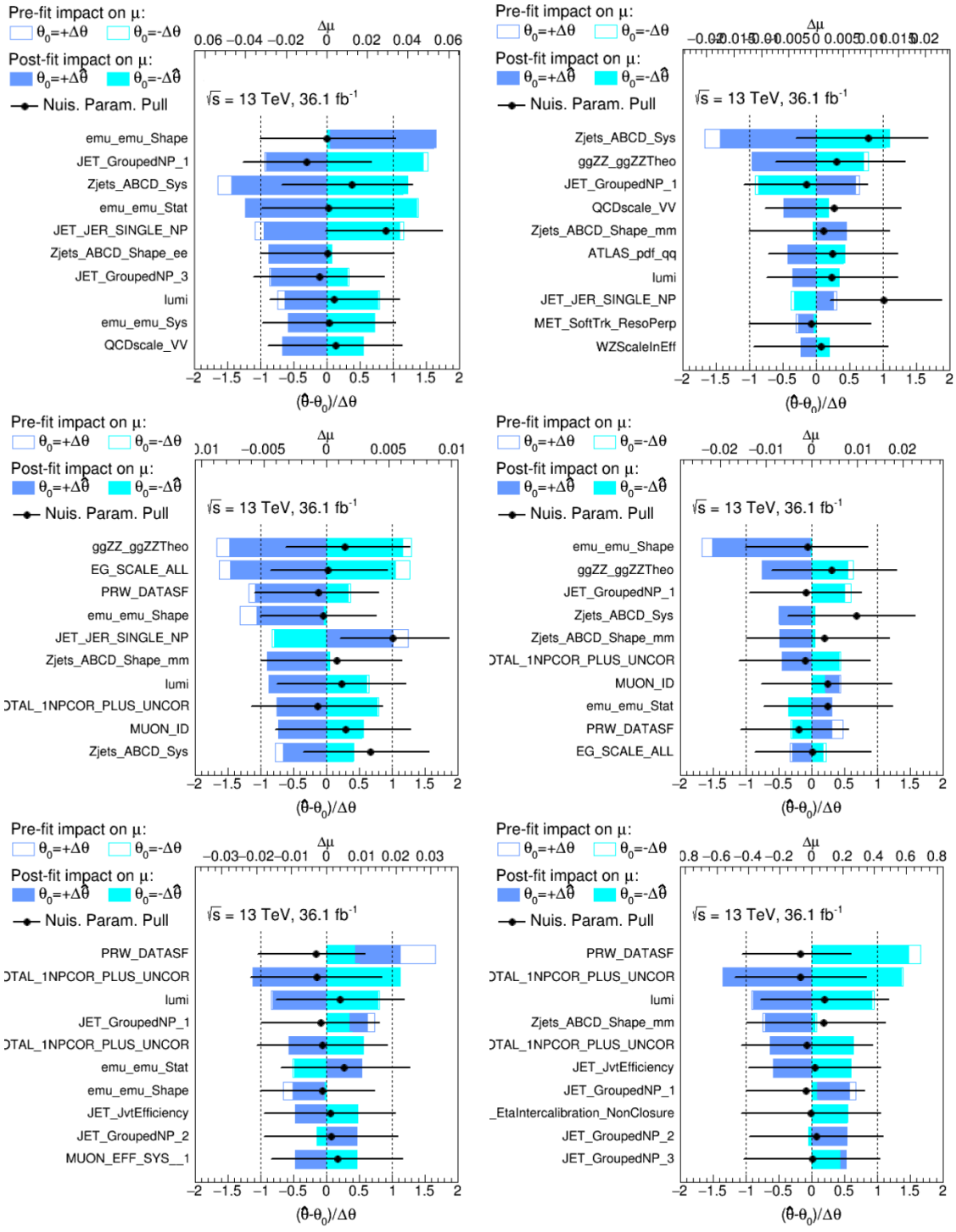


Figure 7.5: The nuisance parameter pulls and ranking in the case of fitting ggF signal + background model to data. ggF Signal mass points 300/400GeV (top left/right), 600/800GeV (middle left/right), and 1/1.2TeV (bottom left/right) are shown. Dots with error bars give the pre-fit (black) and post-fit (red) values and errors of nuisance parameters (pull plot), and the bottom x-axis should be used to check the pull plots. Shaded areas show the impact on the signal strength by changing each nuisance parameter to $\pm 1\sigma$ according to their values and errors in the pre-fit (yellow) and post-fit (blue) scenarios (ranking plot), and the top x-axis should be used to check the ranking. Systematics are pruned if effect is smaller than 1%, and only the top 10 systematic variations are kept.

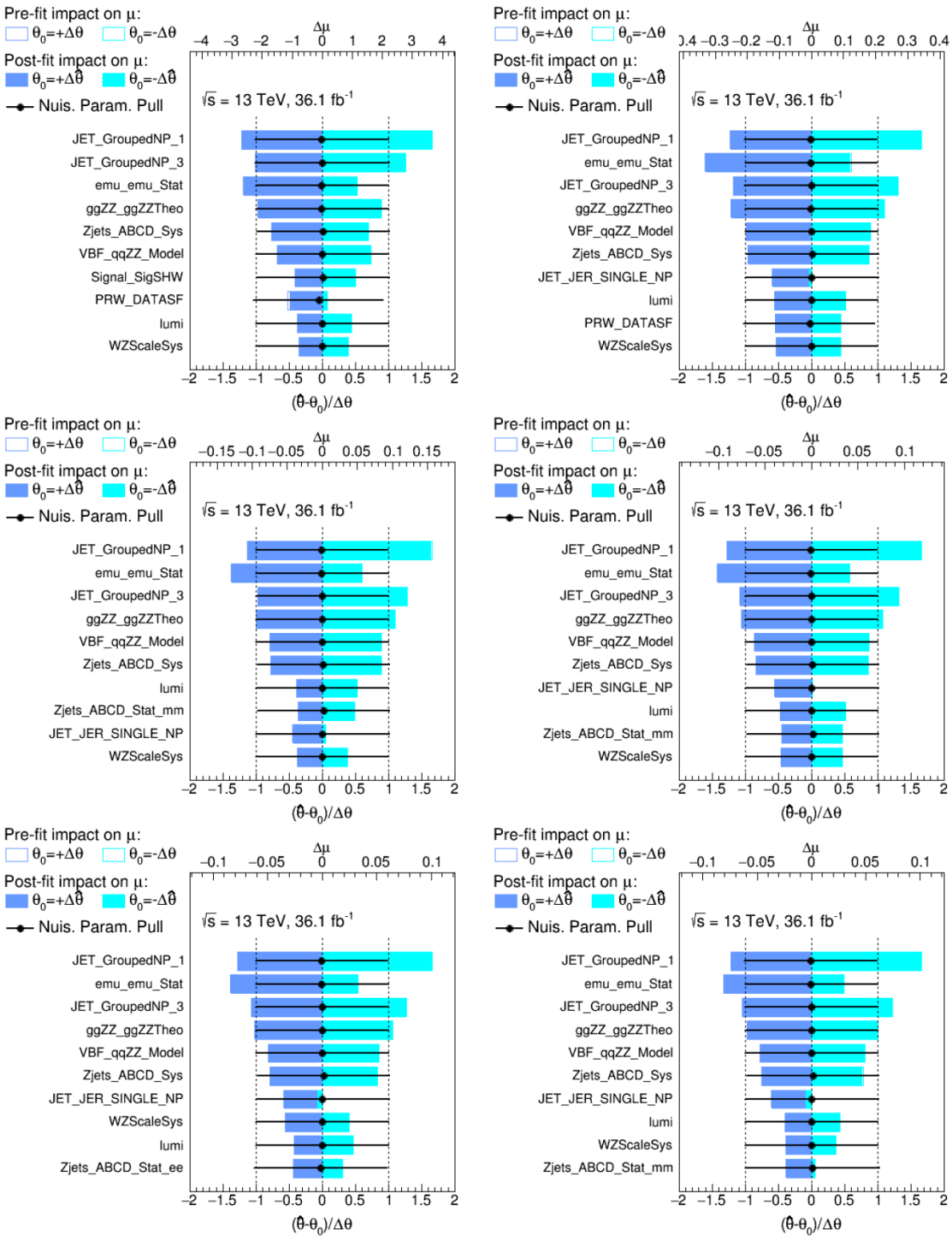


Figure 7.6: The nuisance parameter pulls and ranking in the case of fitting VBF signal + background model to data. VBF Signal mass points 300/400GeV (top left/right), 600/800GeV (middle left/right), and 1/1.2TeV (bottom left/right) are shown. Dots with error bars give the pre-fit (black) and post-fit (red) values and errors of nuisance parameters (pull plot), and the bottom x-axis should be used to check the pull plots. Shaded areas show the impact on the signal strength by changing each nuisance parameter to $\pm 1\sigma$ according to their values and errors in the pre-fit (yellow) and post-fit (blue) scenarios (ranking plot), and the top x-axis should be used to check the ranking. Systematics are pruned if effect is smaller than 1%, and only the top 10 systematic variations are kept.

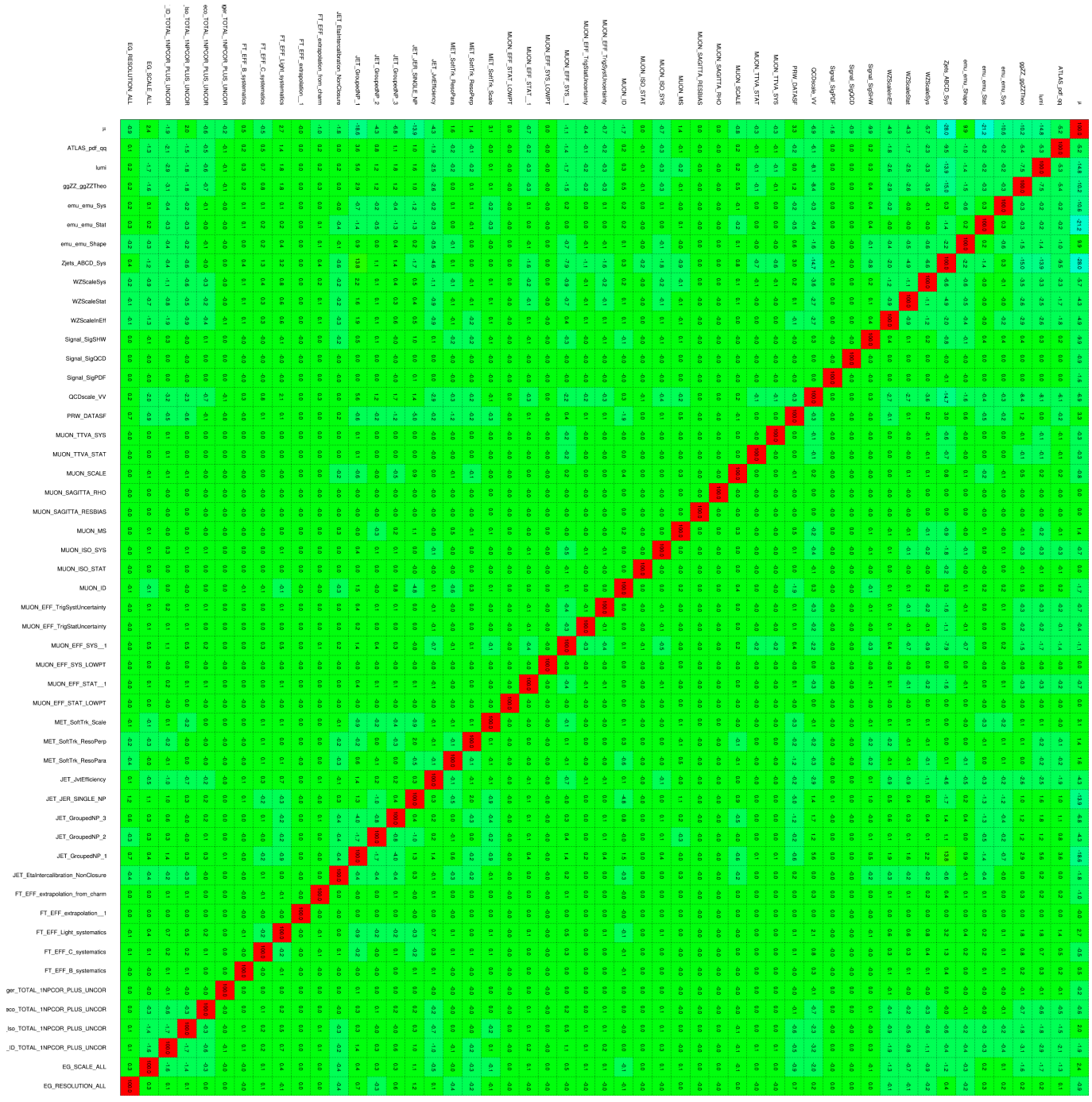


Figure 7.7: Correlation matrix of nuisance parameters and signal strength, for signal mass points of 300 GeV.

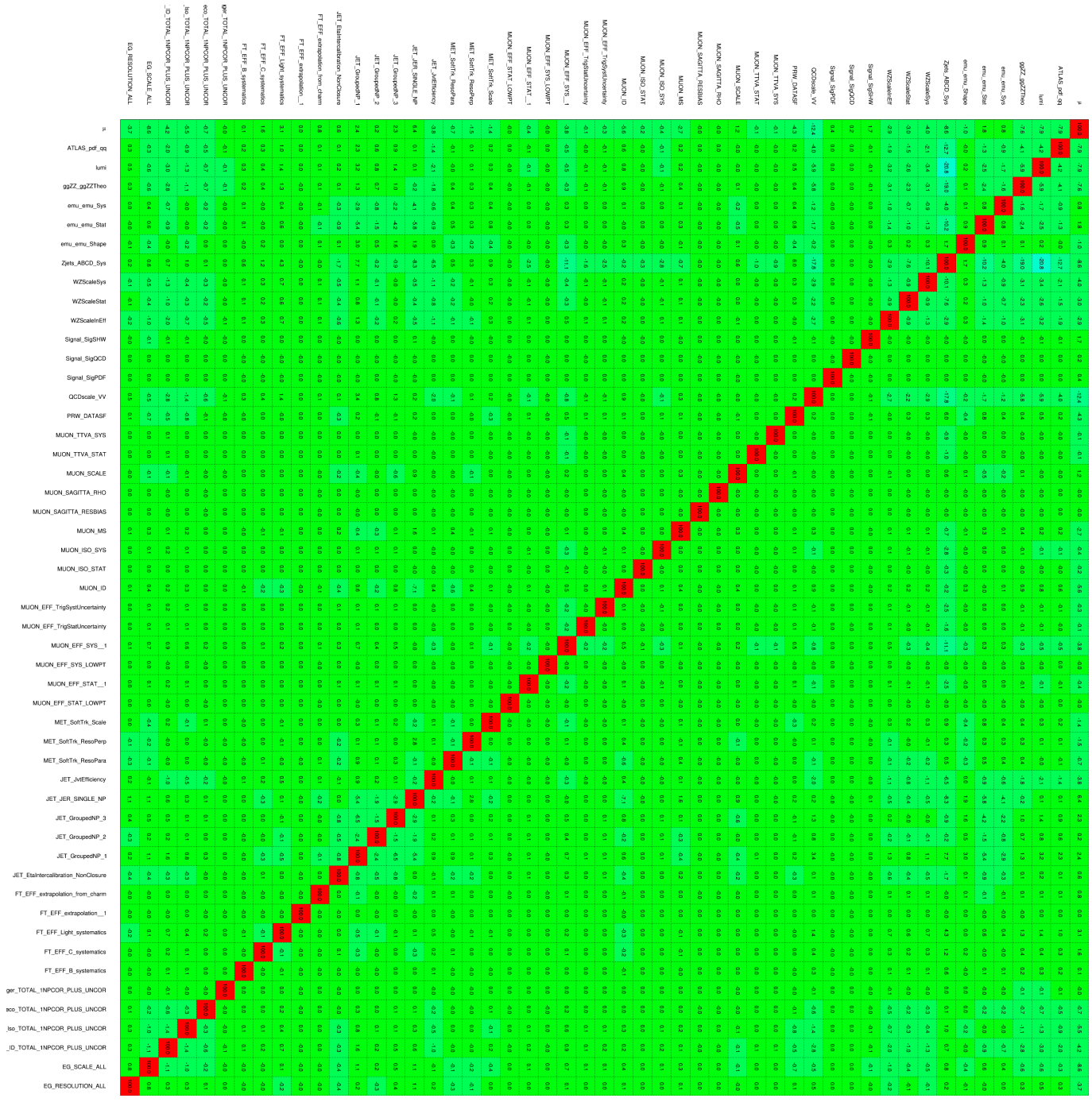


Figure 7.8: Correlation matrix of nuisance parameters and signal strength, for signal mass points of 600 GeV

ggF Mass	Channel	Signal Pre-fit	Signal Post-fit	$S+B$ Pre-fit	$S+B$ Post-fit
$m_{\mathrm{H}} = 300.0$ GeV	$\ell\ell\nu\nu$	98.5 ± 6.0	15.4 ± 12.3	711 ± 24	682 ± 16
$m_{\mathrm{H}} = 600.0$ GeV	$\ell\ell\nu\nu$	144.9 ± 4.5	-5.6 ± 6.3	757 ± 29	658 ± 17
$m_{\mathrm{H}} = 1$ TeV	$\ell\ell\nu\nu$	12 ± 0.4	-2.4 ± 4.4	624 ± 29	660 ± 15
Data	$\ell\ell\nu\nu$	672			

Table 7.4: Signal and signal+background yields for the 300 GeV, 600 GeV and 1 TeV ggF mass points before and after the fit.

VBF Mass	Channel	Signal Pre-fit	Signal Post-fit	$S+B$ Pre-fit	$S+B$ Post-fit
$m_{\mathrm{H}} = 300.0$ GeV	$\ell\ell\nu\nu$	0.23 ± 0.03	4.4 ± 1.5	4.7 ± 0.8	8.9 ± 1.9
$m_{\mathrm{H}} = 600.0$ GeV	$\ell\ell\nu\nu$	4.4 ± 0.2	4.4 ± 1.6	8.9 ± 0.9	8.9 ± 1.7
$m_{\mathrm{H}} = 1$ TeV	$\ell\ell\nu\nu$	6.8 ± 0.4	4.4 ± 1.8	10.4 ± 0.7	8.9 ± 1.8
Data	$\ell\ell\nu\nu$	9			

Table 7.5: Signal and signal+background yields for the 300 GeV, 600 GeV and 1 TeV VBF mass points before and after the fit.

pre-fit and post-fit in the ggF signal region for the 300 GeV, 600 GeV and 1 TeV mass points. The effect of the fit is to reduce the signal yield to the point where the $S+B$ value post-fit is compatible with the observation in data. It can be seen that the signal post-fit can take on negative values, due to the specific range scan chosen for μ during the fit. These negative values are consistent with zero due to the large uncertainty. The uncertainty on $S+B$ is smaller post-fit due to the constraints from the auxiliary measurements, which enter in the NPs.

Table 7.5 shows the effect of the fit on the expected yields in the VBF signal region. The effects post-fit on the signal contributions vary between the three mass points. For the $m_{\mathrm{H}} = 300$ GeV, the signal post-fit is scaled up, to ensure $S+B$ is comparable with the expectation from data. The yield for $m_{\mathrm{H}} = 600$ GeV remains the same, but with an inflated uncertainty. For the $m_{\mathrm{H}} = 1$ TeV, the post-fit yield is scaled down. Overall, post-fit the uncertainties on the signal and $S+B$ increase significantly, reflecting the higher statistical uncertainty present in this region.

7.6 Combination with 4ℓ

The combination procedure between $H \to ZZ \to 4\ell$ and $H \to ZZ \to \ell\ell\nu\nu$ analysis involves a simultaneous fit between all the signal regions present in both searches. The four lepton signature is separated based on the production modes into ggF and VBF categories. The ggF category is made from four signal regions: $4e, 2e2\mu, 2\mu2e, 4\mu$. The VBF category consists of

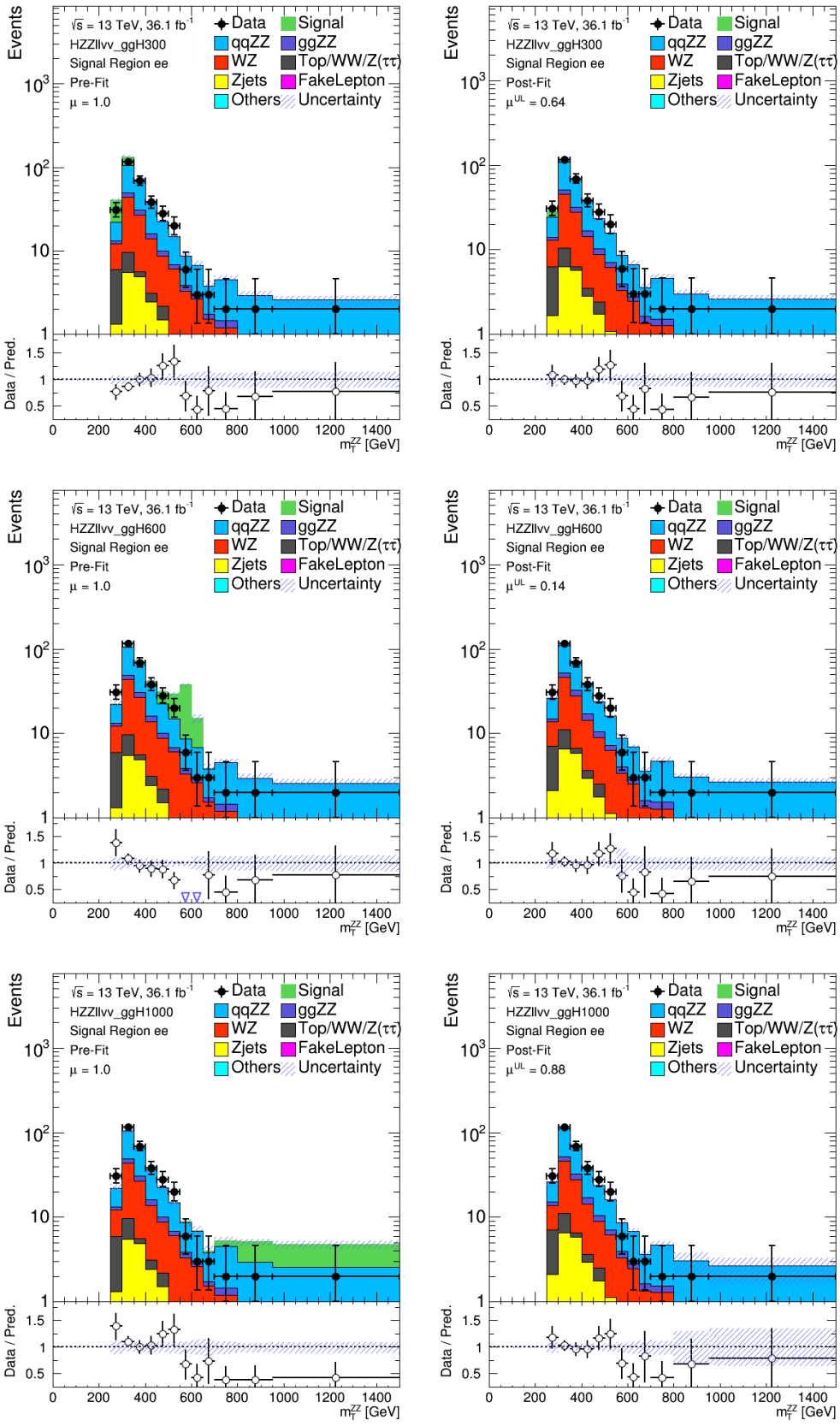


Figure 7.9: Pre-fit (left) and Post-fit (right) plots (comparison between data and predictions (signal is added on top of SM background)) in the $H \to ZZ \to ee\nu\nu$ final state. Both $ee\nu\nu$ and $\mu\mu\nu\nu$ channels are considered in the same fit i.e. combined fit.

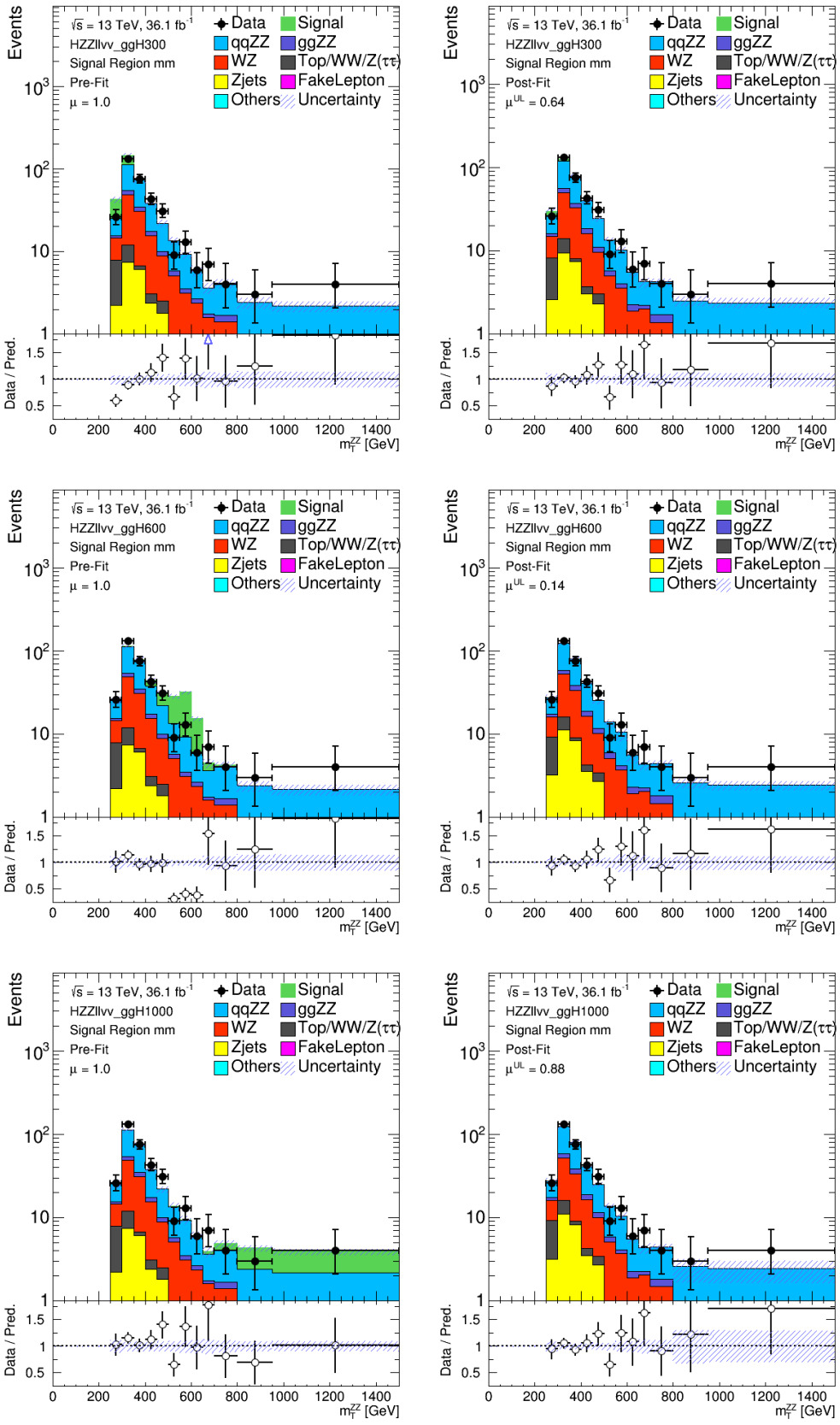


Figure 7.10: Pre-fit (left) and Post-fit (right) plots (comparison between data and predictions (signal is added on top of SM background)) in the $H \to ZZ \to \mu\mu\nu\nu$ final state. Both $ee\nu\nu$ and $\mu\mu\nu\nu$ channels are considered in the same fit (combined fit).

a single signal region, which contains all possible final states. The discriminating variable is the $m_{4\ell}$ distribution. The analysis is mainly sensitive to new physics production in the $200 - 400$ GeV.

The analysis selections for the $H \to ZZ \to 4\ell$, along with the yields and modelling in the signal region, are summarized in section 7.6.1. Discussion of some analysis results are found in section 7.6.2. The combination strategy between the $H \to ZZ \to 4\ell$ and $H \to ZZ \to \ell\ell\nu\nu$ is outlined in section 7.6.3. The combined upper limits for a narrow width Higgs are presented in section 7.6.4. The impact of the correlated systematic uncertainties on the POI is given in section 7.6.5. The derived exclusion limits for 2HDMs are given in section 7.6.6. Lastly, exclusion limits on KK gravitons within the RS model are provided in section 7.6.7.

The limits are derived within a custom statistical framework built by the $H \to ZZ \to 4\ell$ group [199].

7.6.1 $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ analysis and results

The object and event selections are summarized in Table 7.6. The leptons are classified into four categories dependent on the lepton flavour: $4\mu, 4e, 2e2\mu$. They are selected using a combination of lepton triggers, single-,di- and tri-lepton triggers. Overall the trigger efficiency is 98% for signal events. The lepton candidates are made from two lepton-quadruplets, which contain same-flavour and opposite sign lepton pairs. Each lepton entering a quadruplet must satisfy a set of p_{T} requirements. Ambiguity between quadruplets is resolved by checking the closeness of the pair masses m_{12}, m_{34} to the Z boson mass value. J/ψ contamination is supressed by requiring the mass of each lepton pairs to be greater than 5 GeV. $Z-$mass constraints are applied to both $Z-$candidates to improve the invariant mass resolution. In the VBF category, at least two jets are required with $m_{jj} > 400$ GeV and $|\Delta\eta_{jj}| > 3.3$. The ggF category is orthogonally defined to the VBF event region.

The $H \to ZZ \to 4\ell$ signal region is of high purity, as the cross sections for most SM processes in this phase space are small. This leaves the largest contribution to the total background to be the irreducible ZZ process. The total background contributions are separated into a $qq \to ZZ$ ($\sim 85\%$) component, a $gg \to ZZ$ ($\sim 10\%$) component, small residual contributions from Z + jets and $t\bar{t}$ processes ($\sim 2\%$) and other small contributions such as triboson. The ZZ process is estimated from simulation, while the Z + jets and $t\bar{t}$ are obtained from data-driven methods. The Z + jets and $t\bar{t}$ are supressed by applying impact parameter requirements as well as track- and calorimeter-based isolation requirements on the leptons.

$qq \to ZZ$ is simulated through a `Sherpa` sample, whose matrix element contains NLO QCD

Physics Objects	
Electrons	
Loose Likelihood quality electrons with hit in innermost layer, $E_{\mathrm{T}} > 7$ GeV and $\lvert\eta\rvert < 2.47$	
Interaction point constraint: $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm (if ID track is available)	
Muons	
Loose identification with $p_{\mathrm{T}} > 5$ GeV and $\lvert\eta\rvert < 2.7$	
Calo-tagged muons with $p_{\mathrm{T}} > 15$ GeV and $\lvert\eta\rvert < 0.1$, segment-tagged muons with $\lvert\eta\rvert < 0.1$	
Stand-alone and silicon-associated forward restricted to the $2.5 < \lvert\eta\rvert < 2.7$ region	
Combined, stand-alone (with ID hits if available) and segment-tagged muons with $p_{\mathrm{T}} > 5$ GeV	
Interaction point constraint: $\lvert d_0\rvert < 1$ mm and $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm (if ID track is available)	
Jets	
anti-k_T jets with *bad-loose* identification, $p_{\mathrm{T}} > 30$ GeV and $\lvert\eta\rvert < 4.5$	
Jets with $p_{\mathrm{T}} < 60$ GeV and $\lvert\eta\rvert < 2.4$ are required to pass the pile-up jet rejection	
at the 92% working point (JVT score > 0.59).	
***b*-tagging**	
Previously selected jets with $\lvert\eta\rvert < 2.5$ passing the `MV2_c10` algorithm at its 70% working point	
overlap removal	
Jets within $\Delta R < 0.2$ of an electron or $\Delta R < 0.1$ of a muon are removed	
Event Selection	
Quadruplet Selection	- Require at least one quadruplet of leptons consisting of two pairs of same-flavour opposite-charge leptons fulfilling the following requirements: - minimum p_{T} cuts for three leading leptons in the quadruplet: 20, 15 and 10 GeV - Maximum one calo-tagged or stand-alone muon or silicon-associated forward per quadruplet - Select best quadruplet (per channel) to be the one with the (sub)leading dilepton mass (second) closest the Z mass - Leading di-lepton mass requirement: $50 < m_{12} < 106$ GeV - Sub-leading di-lepton mass requirement: $12 < m_{34} < 115$ GeV - $\Delta R(\ell, \ell') > 0.10$ (0.20) for all same (different) flavour leptons in the quadruplet - Remove quadruplet if alternative same-flavour opposite-charge di-lepton gives $m_{\ell\ell} < 5$ GeV
Isolation	- Contribution from the other leptons of the quadruplet is subtracted - Muon track isolation ($\Delta R <= 0.30$): $\Sigma p_{\mathrm{T}}/p_{\mathrm{T}} < 0.15$ - Muon calorimeter isolation ($\Delta R = 0.20$): $\Sigma E_{\mathrm{T}}/p_{\mathrm{T}} < 0.30$ - Electron track isolation ($\Delta R <= 0.20$) : $\Sigma E_{\mathrm{T}}/E_{\mathrm{T}} < 0.15$ - Electron calorimeter isolation ($\Delta R = 0.20$) : $\Sigma E_T/E_T < 0.20$
Impact Parameter Significance	- Apply impact parameter significance cut to all leptons of the quadruplet - For electrons: $d_0/\sigma_{d_0} < 5$ - For muons: $d_0/\sigma_{d_0} < 3$
Vertex Selection	- Require a common vertex for the leptons: - χ^2/ndof < 6 for 4μ and < 9 for others.

Table 7.6: Summary of the event selection requirements. For the 4ℓ analysis, the two lepton pairs are denoted as m_{12} and m_{34}.

modelling for the 0-1 jet bins and LO modelling for the 2-3 jet bins. Additional corrections are applied for NLO EW diagrams. The $gg \to ZZ$ is also taken from a `Sherpa` sample, which models models the 0-1 jet bins at LO, and applies a 60% uncertainty from NLO corrections. The $m_{4\ell}^{ZZ}$ shape is modelled by an empirical formula, described in [200]. The Z + jets and $t\bar{t}$ contributions are obtained separately for $\ell\ell\mu\mu$ and $\ell\ell ee$ final states. In the $\ell\ell\mu\mu$ state, three control regions are used to estimate the light and heavy contribution of the Z + jets and $t\bar{t}$ background, by either inverting the some of the impact parameter selections or by inverting some of the isolation requirements. A fit over all the regions is then used to extrapolate to the signal region with transfer factors. In the $\ell\ell ee$ final state a dedicated "$3\ell + X$" control region is used, which estimates the contribution from photon conversions, jets mis-identification and electrons from heavy flavour decays. More details can be found in [201].

7.6.1.1 Results

The yields of the high mass signal regions are presented in Table 7.7 and the invariant mass distributions are shown in Figure 7.11. Excesses of data over the prediction are seen in the ggF category at $m_{4\ell} \sim 240$ GeV and $m_{4\ell} \sim 700$ GeV. The excess at 240 GeV occurs mostly within the $4e$ channel while the excess at 700 GeV is across all channels.

Table 7.7: $\ell^+\ell^-\ell^+\ell^-$ search: Number of expected and observed events for $m_{4\ell} > 130$ GeV, together with their statistical and systematic uncertainties, for the ggF- and VBF-enriched categories.

Process	ggF-enriched categories			VBF-enriched category
	4μ channel	$2e2\mu$ channel	$4e$ channel	
ZZ	297 ± 1 ± 40	480 ± 1 ± 60	193 ± 1 ± 25	15 ± 0.1 ± 6.0
ZZ (EW)	1.92 ± 0.11 ± 0.19	3.36 ± 0.14 ± 0.33	1.88 ± 0.12 ± 0.20	3.0 ± 0.1 ± 2.2
Z + jets/$t\bar{t}/WZ$	3.7 ± 0.1 ± 0.8	7.8 ± 0.1 ± 1.1	4.4 ± 0.1 ± 0.8	0.37 ± 0.01 ± 0.05
Other backgrounds	5.1 ± 0.1 ± 0.6	8.7 ± 0.1 ± 1.0	4.0 ± 0.1 ± 0.5	0.80 ± 0.02 ± 0.30
Total background	308 ± 1 ± 40	500 ± 1 ± 60	203 ± 1 ± 25	19.5 ± 0.2 ± 8.0
Observed	357	545	256	31

7.6.2 Observed excesses

In the $H \to ZZ \to 4\ell$ analysis, two excesses are observed in data. One at 240 GeV and one at 700 GeV. In the ggF production, the local significance is 3.6σ. The corresponding global significance is 2.2σ which is calculated individually for each excess in the mass range between 200 GeV and 1.2 TeV. For the $H \to \ell\ell\nu\nu$ analysis, no significant excesses have been found. Since the analysis is more sensitive in the high mass range, the excess at 700 GeV

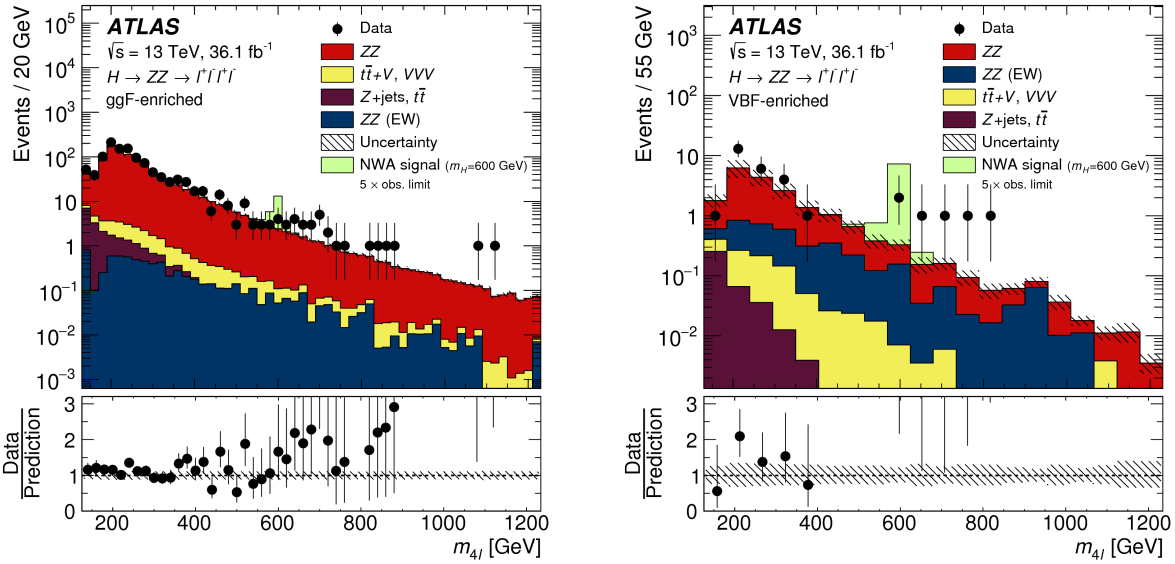


Figure 7.11: Four-lepton invariant mass distribution for the ggF-enriched category (left) and the VBF-enriched category (right). The last bin includes the overflow. The error bars on the data points indicate the statistical uncertainty, while the systematic uncertainty on the prediction is shown by the hatched band. The bottom part of the figures shows the ratio of data over expectation. Figure is taken from [6].

present in 4ℓ is excluded at 95% confidence level. At 240 GeV, the $H \to \ell\ell\nu\nu$ search is not sensitive and the excess cannot be excluded.

In the combination between 4ℓ and $\ell\ell\nu\nu$, the excess at 700 GeV has a global significance of less than 1σ and a local significance of 2σ. Since no significant excess is found, upper limits are placed on the production cross section for ggF and VBF modes.

7.6.3 Statistical interpretation and combination

7.6.3.1 $H \to ZZ \to 4\ell$ statistical interpretation

The four lepton search uses a similar likelihood function to the $H \to ZZ \to \ell\ell\nu\nu$ analysis with several differences. In the dilepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ analysis, a binned likelihood is used for the ggF category, while for the four lepton mode pdfs model the signal and background shapes of the $m_{4\ell}$ distributions. The signal pdf, $f_S(x_e)$ is constructed for a variety of m_{H} values. This is achieved with the application of a smoothing algorithm, which also removes statistical fluctuations [202]. For the background pdf, $f_B(x_e)$, the irreducible ZZ backgrounds are taken from simulation and the same smoothing algorithm is applied. The data-driven backgrounds separately estimate the shape and normalization in dedicated control regions. Again a smoothing algorithm is applied to obtain the pdf. The NPs for the signal and backgrounds are constrained by Gaussian functions, similar to the $\ell\ell\nu\nu$ case. The likelihood function in

the $H \to ZZ \to 4\ell$ analysis then takes the form:

$$\mathcal{L}(n|\mu,\boldsymbol{\theta}) = \text{Pois}(n|\mu S + B)\left[\prod_{e=1}^{n}\frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B}\right] \times \prod_{n=1}^{N} N(\widetilde{\theta}|\theta). \tag{7.16}$$

The fitting procedure is the same as described in section 7.2.3. Two signal strengths are used for the ggF and VBF categories. The upper limits on the signal strengths μ are similarly derived from finding a μ value which yields a CL_S of 0.05.

7.6.3.2 $4\ell + \ell\ell\nu\nu$ combination

For a combination of two $H \to ZZ$ searches, the likelihood functions from each analysis are merged. The fit is then performed simultaneously across all the signal regions but separately for the ggF and VBF production. This is done as the ratio of production via these two processes in absence of a specific mechanism. The procedure for combination of Higgs searches is outlined in [203, 204], and is implemented with RooFit [205] and RooStats [206].

The main point in the combination is the correlation of systematic uncertainties. Uncorrelated uncertainties are given the simpler prescription of auxiliary measurements and applied as in equations (7.8, 7.16). A correlated uncertainty is modelled in the combined fit by using a NP common to all searches. These correlated NPs are thus assumed to have a 100 % correlation. The assumption of total correlation is applied to the signal strengths μ_{ggF} and μ_{VBF}. Before the combination, certain NPs with very small impact on the normalization are pruned. For signals, among the correlated NPs are the theoretical signal systematics for PDF, QCD and PS. For backgrounds, the list of correlated systematics include QCD scale and PDF uncertainties for the ZZ backgrounds, as well as experimental uncertainties on the luminosity and JES uncertainties. Lepton NPs which are not pruned, are taken to have a 100% correlation.

7.6.4 Combined upper limits

The combined upper limits on the signal strength parameters, μ_{ggF} and μ_{VBF}, from the combined $4\ell + \ell\ell\nu\nu$ searches are derived in a similar manner as described in section 7.3 i.e. by finding the value of μ which gives a CL_S of 0.05. The combined upper limits for ggF and VBF regions are given in Figure 7.12. In addition, the individual limits from the 4ℓ and $\ell\ell\nu\nu$ channels are also shown for comparison purposes.

In the $\ell\ell\nu\nu$ analysis, the Higgs masses are generated for specific masses. However, for the combination they are used as in inputs to a moment-morphing technique [156] in order to obtain m_{T}^{ZZ} distributions for any mass point.

From the upper limits plots, several conclusions are immediately reached: at low masses the 4ℓ dominates the overall sensitivity and the excess at $m_{4\ell}$ makes a clear impact at m_{H} $\sim$ 250 GeV where the $\ell\ell\nu\nu$ search is not sensitive. The observed CL_S limit is seen to be consistent with the expected limit within $\pm 1\sigma$ except at 700 GeV, where the excess in the 4ℓ channel is present. For high masses, the $\ell\ell\nu\nu$ dominates the sensitivity up to 1.2 TeV. The last mass point, at 1.4 TeV, is not considered in the combination as the 4ℓ analysis does not have any statistical precision.

Overall, the results in run II are seen to improve the exclusion limits obtained from run I, shown in Figure 2.8, with the ratios of parton luminosities taken into account, by factors of two to seven depending on the mass point [6].

7.6.5 Ranking of systematic uncertainties

The effect of the systematic uncertainties depends on the mass point and event category. At $m_{\mathrm{H}} = 300$ GeV in the ggF category, the luminosity uncertainty, the $\ell\ell\nu\nu$ Z + jets data-driven systematic uncertainty, the PS and $e\mu$ statistical uncertainty have the highest impact on the cross section. PU and JES systematics are important for low to intermediate masses. For high masses electron isolation efficiencies and QCD scale uncertainties become important. For VBF production, at low m_{H} the dominating systematics are JES, theoretical uncertainties on the ZZ background, PU and Z + jets uncertainties. At high m_{H}, Z + jets and $e^{\pm}\mu^{\mp}$ statistical uncertainties have a greater impact. The top four leading systematics for a range of mass points are summarized in Table 7.8. The reason for the greater impact of the luminosity is due to it entering as a NP in the simulated signal and background events. It can vary between 4% to 7% across the mass range depending on the amount of signal and background present.

7.6.6 Exclusion limits on 2HDMs

The theory behind 2HDM was introduced in section 2.5 and the exclusion limit results from run I, are presented in section 2.7.

Briefly summarising some of the important points in the studies of 2HDM:

- the 2HDM exclusion limits are limited to a region where the validity of the narrow width approximation for the Heavy Higgs H holds
- in the calculation of the limits the rates of ggF and VBF signal production are given by the 2HDM theoretical prediction

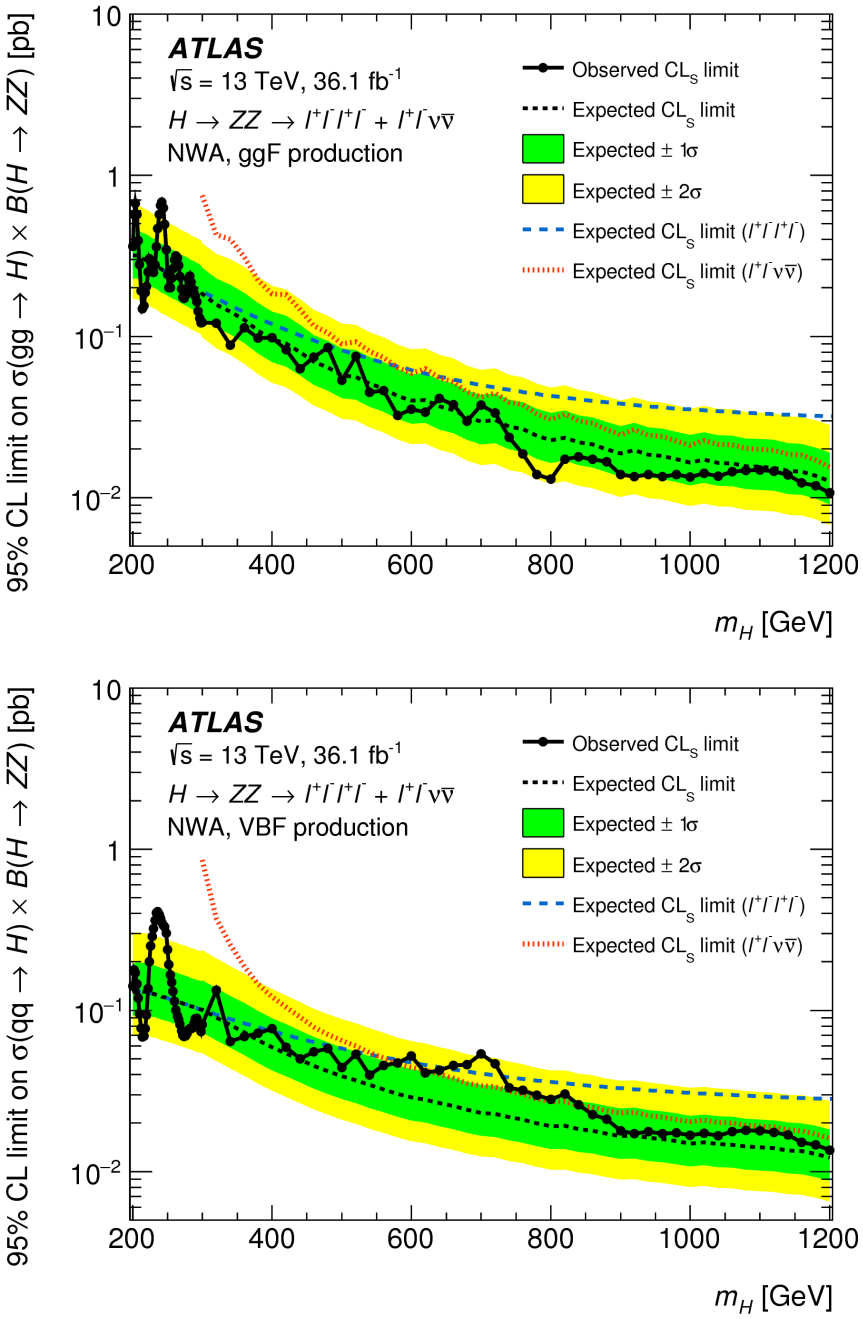


Figure 7.12: The upper limits at 95% confidence level on the cross section times branching ratio for the ggF production mode($\sigma_{\text{ggF}} \times BR(H \rightarrow ZZ)$) and VBF production mode ($\sigma_{\text{VBF}} \times BR(H \rightarrow ZZ)$). The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limits. The dashed coloured lines indicate the expected limits obtained from the individual searches. Figure is taken from [6].

ggF category		VBF category	
Systematic	Effect [%]	Systematic	Effect [%]
$m_{\mathrm{H}} = 300$ GeV			
Luminosity	4	Parton Showering	9
Z + jets systematic	3.3	JES	4
Parton Showering	3.2	Luminosity	4
$e\mu$ stat. unc.	3.2	$qq \to ZZ$ QCD Scale	4
$m_{\mathrm{H}} = 600$ GeV			
Luminosity	6	Parton Showering	6
Pile-up reweighting	5	JES	6
Z + jets systematic	4	Pile-up reweighting	6
$qq \to ZZ$ QCD Scale	3.1	Luminosity	4
$m_{\mathrm{H}} = 1$ TeV			
Luminosity	4	Parton Showering	6
$qq \to ZZ$ QCD Scale	2.3	JES	5
JVT	1.9	Z + jets modelling	4
Z + jets systematic	1.8	Luminosity	4

Table 7.8: Impact of the leading systematic uncertainties on the predicted signal event yield which is set to the expected upper limit, expressed as a percentage of the cross section for the ggF (left) and VBF (right) production modes at $m_{\mathrm{H}} = 300$, 600, and 1000 GeV.

- the exclusion limits are presented for $\tan(\beta)$ vs $\cos(\beta - \alpha)$ at a fixed m_{H} and $\tan(\beta)$ vs m_{H}

The exclusion limit for $\tan(\beta)$ against the Higgs m_{H} parameter is presented in Figure 7.13. The $\cos(\beta - \alpha)$ is fixed at -0.1 to be close to the alignment limit where there is sensitivity to the production of H. The $\ell\ell\nu\nu$ analysis contributes to the sensitivity of the 4ℓ search in the region beyond 300 GeV. The corresponding run I plots are shown in Figure 2.9. The exclusion limits are seen to have improved from run I by a factor of 2 or 3, depending on the mass, for type I 2HDM. An improvement by 50% for type II 2HDM is also observed. The only place where the observed limit has performed worsed than the expected is in the region of the excess at ~ 240 GeV.

In Figure 7.14 the $\tan(\beta)$ vs $\cos(\beta - \alpha)$ is shown for a fixed $m_{\mathrm{H}} = 200$ GeV, where the black line indicates the observed exclusion, the dashed blue line is the expected exclusion and red cross lines indicate the already excluded regions. The choice of Higgs mass is motivated by the smallness of the Higgs width at that energy. The corresponding run I plot is presented in Figure 2.10. The limits have significantly improved for both type I and type II 2HDM.

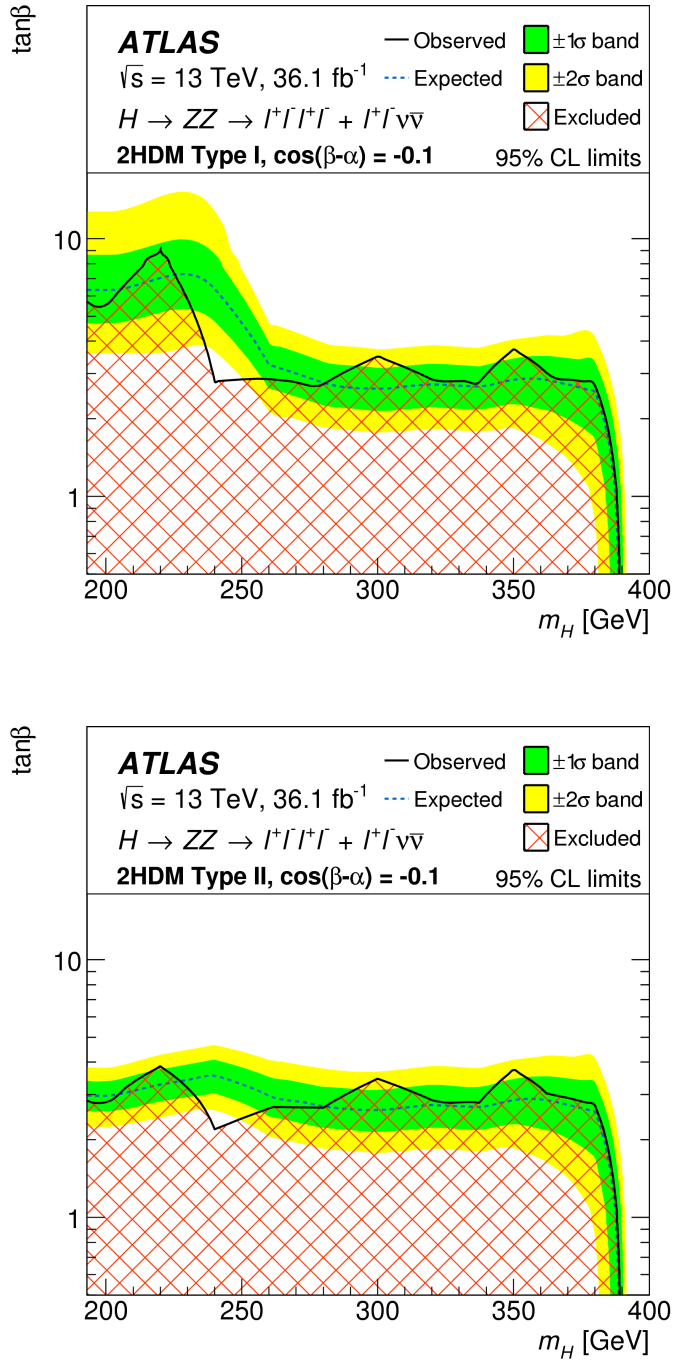


Figure 7.13: The exclusion contour in the 2HDM Type-I and Type-II models for $\cos(\beta-\alpha) = -0.1$, shown as a function of the heavy scalar mass m_H and the parameter $\tan\beta$. The green and yellow bands represent the $\pm1\sigma$ and $\pm2\sigma$ uncertainties on the expected limits. The hatched area shows the observed exclusion. Figure is taken from [6].

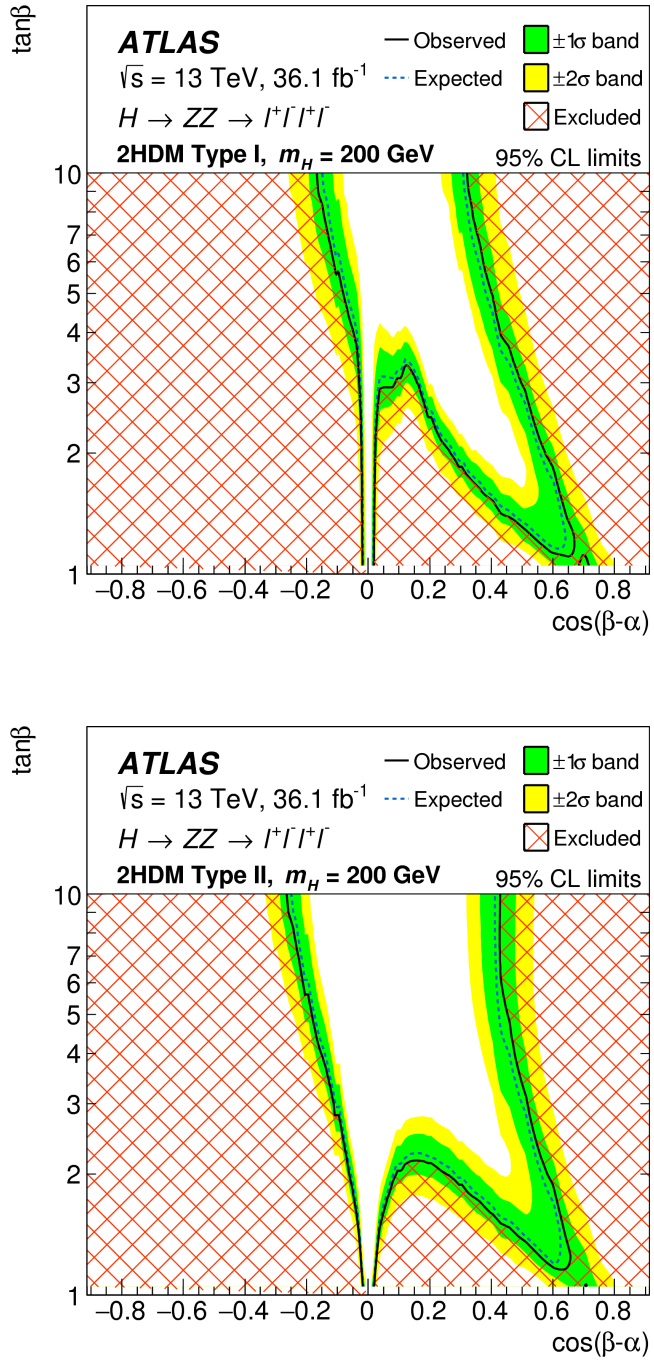


Figure 7.14: The exclusion contour in the 2HDM Type-I and Type-II models for $m_{\mathrm{H}} = 200$ GeV shown as a function of the parameters $\cos(\beta - \alpha)$ and $\tan\beta$. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limits. The hatched area shows the observed exclusion. Figure is taken from [6].

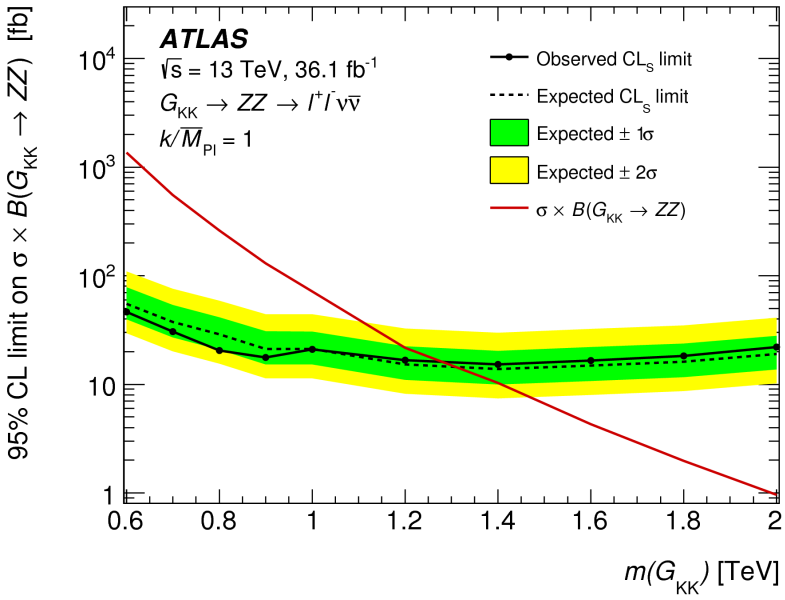


Figure 7.15: Limits on $\sigma \times BR(G_{\mathrm{KK}} \to ZZ)$ for a KK graviton produced with $k/\bar{M}_{\mathrm{pl}} = 1$. The green and yellow bands give the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties of the expected limits. The predicted production cross section as a function of the G_{KK} mass $m(G_{\mathrm{KK}})$ is shown by the red solid line. Figure is taken from [6].

7.6.7 Upper limits for Graviton production

The main signal interpretation so far has been on spin-0 Higgs resonances. However, other signal interpretations are also possible as motivated in section 2.6, where the RS model with spin-2 KK gravitons was briefly discussed. The analysis setup for the spin-2 graviton interpretation is exactly the same as the spin-0 equivalent.

Graviton samples are produced for the graviton mass, m_{G}, in the range between 600 GeV and 2 TeV. The samples are generated for a $k/\bar{M}_{\mathrm{pl}}$ factor of unity which leaves $\Gamma(G \to ZZ) \sim 6\% \times m_{\mathrm{G}}$, as detailed in section 5.3.1.

In the combination scheme, the $H \to ZZ \to 4\ell$ search is not very sensitive above 600 GeV and was found to have a negligible effect on the exclusion limits of the lightest KK gravitons. Therefore only the limits from $H \to ZZ \to \ell\ell\nu\nu$ search are presented. In Figure 7.15, the combined upper limits on $\sigma \times \mathrm{BR}(G_{\mathrm{KK}} \to ZZ)$ for a KK graviton decaying to $ZZ \to \ell\ell\nu\nu$ are presented. Overlaid on top is the theoretical line for a KK graviton with $k/\bar{M}_{\mathrm{pl}} = 1$. From the combined limits the lightest KK resonance is excluded up to 1.3 TeV at 95% confidence level. The observed and expected limits by numbers are detailed in Table M.3.

Chapter 8

Future analysis developments

During the course of run II several interesting developments took place, which promise to bring further improvements in many analyses. An important observable where the performance can be improved upon is the $E_{\mathrm{T}}^{\mathrm{miss}}$. The alternative reconstruction algorithms, outlined in section 4.3.3, have already shown that a more precise estimate of $E_{\mathrm{T}}^{\mathrm{miss}}$ is possible. The reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$, using particle-flow (PFlow) jets, and their data/MC agreement is presented in section 8.1.

Another approach, stems from the object level $E_{\mathrm{T}}^{\mathrm{miss}}$ significance which aims to provide a more effective discrimination between Z + jets and events with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$. The significance uses the resolution of objects which enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction to form the total error on the $E_{\mathrm{T}}^{\mathrm{miss}}$. This is discussed in section 8.2.

8.1 PFlow $\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$

The PFlow algorithm was briefly mentioned in section 4.3.3. Here a more complete discussion of the algorithm and the incorporation into the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is provided.

Jets, up to this point, have been considered in the context of the anti-k_t EMTopo reconstruction as introduced in section 4.2.3. This technique solely supports the exclusive reconstruction of jet kinematics from the calorimeter system. Information from tracks are only implemented in the jet calibration and JVT to supress PU. The PFlow method, instead, directly tries to match jet tracks with their calorimeter deposits. In this manner the track component can be incorporated into the structure of jets with the added benefit of improved resolution and natural PU suppression. This happens because the PFlow construction is also able to decide which detector component information i.e. calorimetry, inner detector or both, to use for maximum precision. Only jets are reconstructed by the PFlow method since leptons are accurately associated to tracks and no significant improvements in resolution can

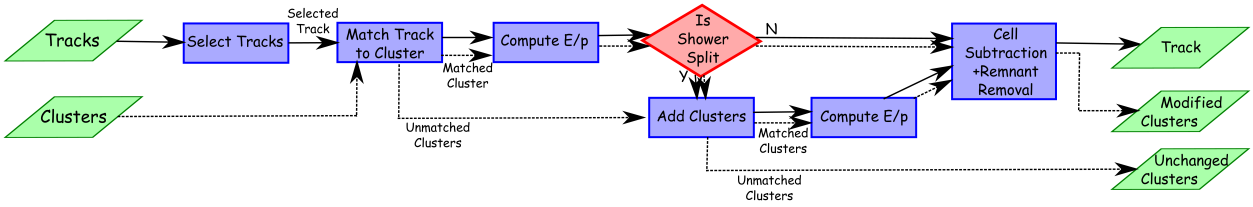


Figure 8.1: Schematic of PFlow object reconstruction. Figure is taken from [142].

be achieved.

8.1.1 The particle-flow technique

The methodology of constructing jets from tracks and calorimeter deposits is schematically described in Figure 8.1. Initially tracks are selected according to certain quality criteria and then an attempt is made to match a track to a single topological cluster in the calorimeter, after which the expected energy in the calorimeter is computed based on its position on track momentum. The algorithm then proceeds to evaluate the probability that the particle energy was deposited in multiple clusters and decides if multiple clusters are needed to recover the full jet cluster energy. Cell-by-cell subtraction is performed from a set of matched clusters to remove their energy deposits and avoid double counting. If it is found that the energy in the system is consistent with the expected fluctuations of a signal from a single particle then the cluster remnants are removed.

The track quality criteria consist of nine hits in the silicon detectors, no missing pixel hits [140], $p_{\mathrm{T}}^{\mathrm{track}} > 0.5$ GeV and $|\eta| < 2.5$. The track-to-cluster matching is aided by defining a distance metric $\Delta R'$:

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2} \tag{8.1}$$

where σ_η, σ_ϕ are computed as the standard deviation of displacements of the topological cluster constituent cells in η, ϕ with respect to its barycenter. A base selection applied to clusters is to have $E^{\mathrm{clus}}/p^{\mathrm{trk}} > 0.1$, where E^{clus} is the cluster energy and p^{trk} is the track momentum. An attempt is then made to match the pre-selected cluster cell to a track using the $\Delta R'$ metric. The correct cluster is almost always found within a small $\Delta R'$ distance. If no pre-selected cluster is matched, then the particle is assumed not to form a cluster in the calorimeter.

Recovering the splitting of showers is achieved by selecting a discriminant, sensitive to this effect. The discriminant is defined as the significance value between the expected energy and the energy of the matched cluster. Next, the cell-by-cell subtraction is performed, by using rings in $\eta - \phi$ space. Remnant cell energy, which can potentially originate from random

fluctuations (such as PU), is subtracted or kept if found above a certain threshold.

The calorimeters jets are still obtained with the anti-k_t clustering algorithm with a radius parameter of 0.4. The clustering algorithm uses both charged and neutral PFlow constituents as inputs.

8.1.2 The particle-flow jets within the $E_{\mathrm{T}}^{\mathrm{miss}}$ framework

The PFlow jets in the context of $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction form the contributions of the PFlow $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$(jets) and $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$(soft) terms. The soft term is constructed from the charged PFlow tracks which are not associated to any well-identified physics object and which survive to the track-to- vertex association. The track quality requirements for objects to be considered in the soft term is:

- $p_{\mathrm{T}} > 0.5$ GeV,
- $|z_0| < 2.0$ mm,
- ΔR(track,electron) > 0.05.

The size of the PFlow soft $E_{\mathrm{T}}^{\mathrm{miss}}$ systematics are of similar order to the EMTopo based soft $E_{\mathrm{T}}^{\mathrm{miss}}$ as they follow a similar definition [207].

The improvements in the reconstruction of the missing transverse momentum are observed in the ability to limit the amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ generated by poor jet resolution and calibration. This behaviour can be deduced from Figure , where the PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ is compared with the EMTopo definition for the Z + jets MC. The spread of the core $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is reduced due to the better measurement of low p_{T} jets.

The performance of the total PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ is presented in Figure 8.2. The dis-agreements between data and simulation behave in similar fashion to the EMTopo definition. The discrepancies are not fatal since they are covered by the experimental detector systematics. For events in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, which contain genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ events, the performance of the both definitions are of similar order. The performance of the PFlow jet (soft) term is presented in Figure 8.3 (8.3). The distributions are found to be compatible between data and simulation.

8.1.3 PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ in $H \to ZZ \to \ell\ell\nu\nu$ analysis

Figure 8.2 shows the state of the $E_{\mathrm{T}}^{\mathrm{miss}}$ performance after the object and event pre-selections outlined in section 5.2, so the studies of the PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ are very close to realization within

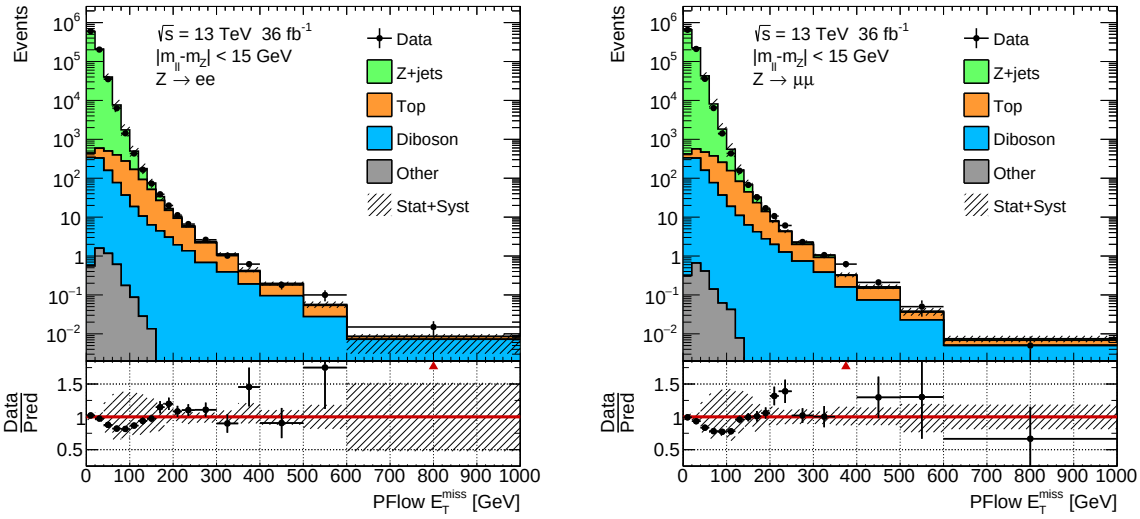


Figure 8.2: The PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the $Z \to ee$ (left) and $Z \to \mu\mu$ (right) channels are shown. MC is scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band contains the statistical and systematic errors on the MC.

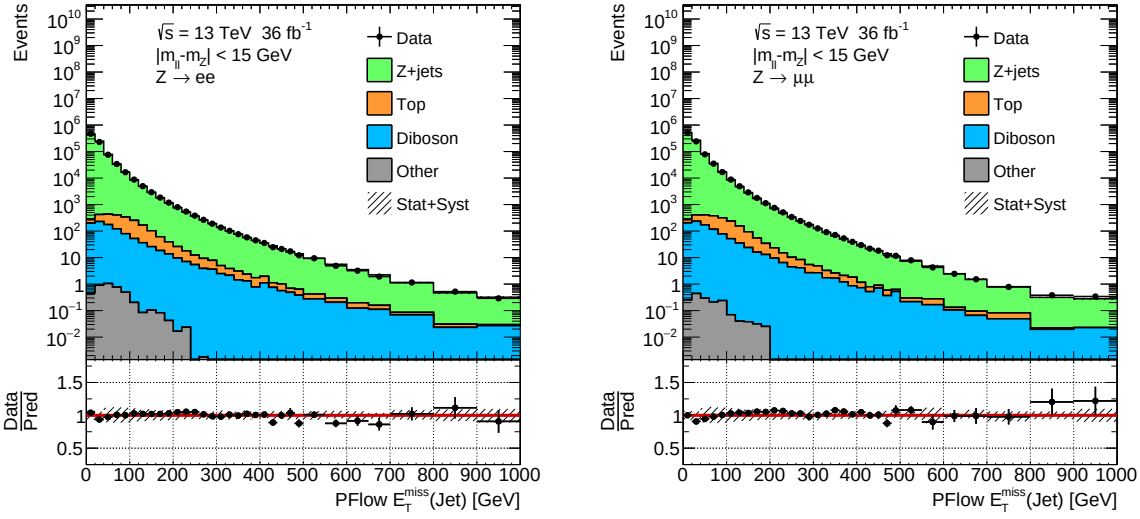


Figure 8.3: The PFlow distribution for the $Z \to ee$ (left) and $Z \to \mu\mu$ (right) channels are shown for the jet $E_{\mathrm{T}}^{\mathrm{miss}}$ component. MC is scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band contains the statistical and systematic errors on the MC.

the context of the $H \to ZZ \to \ell\ell\nu\nu$ search. In terms of improvement a first guess is the better rejection of the Z + jets background. This can translate to better exclusion limits for Higgs mass points in the range between 300 and 500 GeV. In addition, some sensitivity to the 300 GeV signal may be recovered if the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement has a lower threshold. However, if all selections are re-optimised, an improvement in sensitivity for all mass points maybe seen.

The PFlow $E_{\mathrm{T}}^{\mathrm{miss}}$ definition has yet to be thoroughly researched before being considered as an alternative to the current $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction in the $H \to ZZ \to \ell\ell\nu\nu$ searches. Although

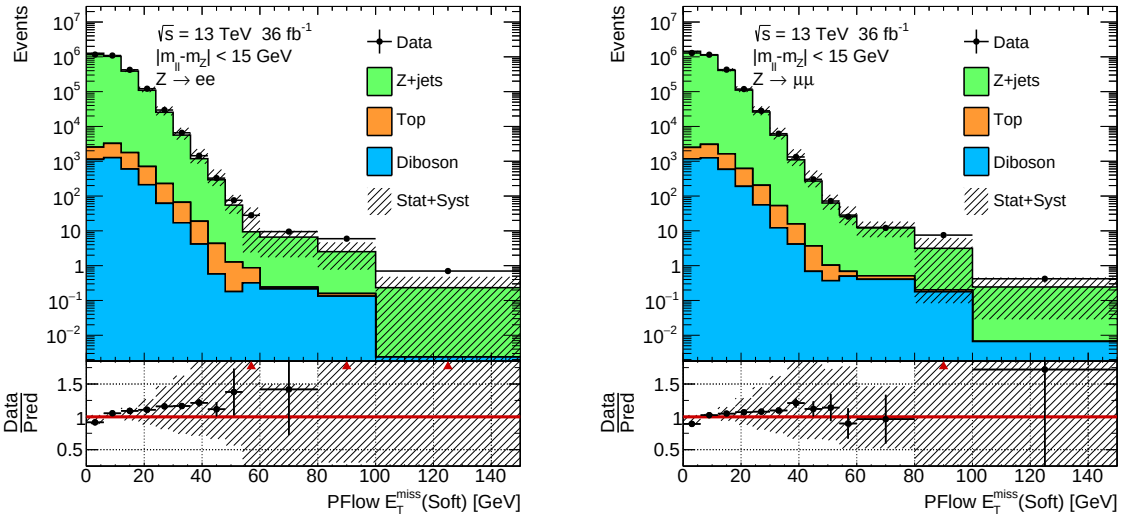


Figure 8.4: The PFlow distribution for the $Z \to ee$ (left) and $Z \to \mu\mu$ (right) channels are shown for the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ component. MC is scaled to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV. The shaded band contains the statistical and systematic errors on the MC.

the reduction in resolution, shown in Figure 4.14, is very promising, the severe degradation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale, presented in Figure 4.16, leaves still much to be desired.

8.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ significance

The value of adopting the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance[1], $\mathcal{S}$, relies on identification and suppression of events where the value of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ originates from mis-reconstruction and detector inefficiencies. In run I of the LHC, $\mathcal{S}$ was defined by $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\Sigma E_{\mathrm{T}}}$, which is known as the event based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance. The ΣE_{T} is derived from the scalar sum of all objects in an event, as given by equation 4.10. During run I, the error on the $E_{\mathrm{T}}^{\mathrm{miss}}$ was described by $\sqrt{\Sigma E_{\mathrm{T}}}$, since the energy resolution approximately develops as $\sigma/E \sim a/\sqrt{E}$ as the $E_{\mathrm{T}}^{\mathrm{miss}}$ was entirely dependent on the calorimeter deposits.

However, in run II, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed with tracks in addition to calorimeter clusters and $\sqrt{\Sigma E_{\mathrm{T}}}$ does not fully reflect the resolution of the tracks as the error evolves with the momentum and depends on the direction of the track. Each object has two p_{T} resolutions, one defined parallel to the direction of the object, $\sigma_{p_{\mathrm{T}}}$, and one perpendicular to the direction, given by $p_{\mathrm{T}} \times \sigma_{\phi}$ where σ_{ϕ} is the angular resolution of the transverse plane. The relative resolution for all objects in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction for a given p_{T} and η are given in Table N.1. The evolution of the relative resolution as a function of p_{T} in certain η ranges are visualized in Figure N.2.

[1] The significance of any variable X is defined as the ratio of the variable to its corresponding error i.e. $X/\Delta X$.

A proper definition, which takes into consideration both track and calorimeter information, is provided by a log-likelihood approach [7, 208]. The log-likelihood approach consist of the ratio of likelihoods between the hypothesis that the real transverse momentum carried by invisible particles, $\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}}$, is zero against the hypothesis that $\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}}$ is non-zero. The definition of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is:

$$\mathcal{S}^2 = 2 \ln \left(\frac{\max_{\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} \neq 0} \mathcal{L}(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} | \mathbf{p}_{\mathrm{T}}^{\mathrm{inv}})}{\max_{\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} = 0} \mathcal{L}(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} | \mathbf{p}_{\mathrm{T}}^{\mathrm{inv}})} \right). \tag{8.2}$$

where numerator and denominator are calculated by maximizing the parameters of the likelihood given $\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} \neq \mathbf{0}$ and $\mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} = \mathbf{0}$, respectively. $\mathcal{S}$ expresses the likelihood of that the real value of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is equal to the sum of the measured value provided that the genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ is zero. It is characteristic of the Z + jets events in the $E_{\mathrm{T}}^{\mathrm{miss}}$ tail to have a large resolution error which would result in a relatively small value of $\mathcal{S}$. Large values of $\mathcal{S}$ indicate that the $E_{\mathrm{T}}^{\mathrm{miss}}$ is more likely to originate from invisible particles.

Under the assumption that the measurement of each object is independent and the uncertainties to be described by Gaussian pdfs with covariance matrix, $\mathbf{V}_i$, the likelihood becomes a two dimensional Gaussian:

$$\mathcal{L}(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} | \mathbf{p}_{\mathrm{T}}^{\mathrm{inv}}) \propto \exp \left(\left(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} - \mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} \right)^{\mathrm{T}} \left(\sum_i \mathbf{V}_i \right)^{-1} \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} - \mathbf{p}_{\mathrm{T}}^{\mathrm{inv}} \right) \right). \tag{8.3}$$

Therefore $\mathcal{S}$ becomes a χ^2 distribution, with two degrees of freedom:

$$\mathcal{S}^2 = 2 \ln \left(\frac{\mathcal{L}(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} | \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})}{\mathcal{L}(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} | \mathbf{0})} \right) = \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \right)^{\mathrm{T}} \left(\sum_i \mathbf{V}_i \right)^{-1} \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} \right), \tag{8.4}$$

where the index i corresponds to each of the reconstructed objects that enters the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation and $\mathbf{V}_i$ is the corresponding covariance matrix. The χ^2 definition expresses the dependencies between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the individual resolutions in the covariance matrix. The covariance matrix for each hard object can be defined in a coordinate system where one axis is aligned with the measured p_{T} for each object, with $\mathbf{p}_{\mathrm{T,i}} = (p_{\mathrm{T},i} \cos(\phi_i), p_{\mathrm{T},i} \sin(\phi_i))^{\mathrm{T}}$:

$$\mathbf{V}_i = \begin{pmatrix} \sigma^2_{p_{\mathrm{T},i}} & 0 \\ 0 & {p_{\mathrm{T},i}}^2 \sigma^2_{\phi_i} \end{pmatrix}, \tag{8.5}$$

were the $p_{\mathrm{T},i}$ and ϕ_i measurements are uncorrelated. The covariance matrix for the TST is defined as:

$$\mathbf{V}_{\mathrm{soft}} = \begin{pmatrix} \sigma^2_{\mathrm{soft}} & 0 \\ 0 & \sigma^2_{\mathrm{soft}} \end{pmatrix}, \tag{8.6}$$

where the soft term variance, σ^2_{soft}, is found from the RMS of E_x^{miss} in $Z \to \mu\mu$ events with an applied jet veto. The width is mainly due to the soft term resolution, which is found

to be 8.9 GeV. The total covariance matrix is defined, in the $x-y$ coordinate system, by matrix rotation:

$$\mathbf{V}_{xy} = \sum_i R^{-1}\left(\phi_i\right)\, \mathbf{V}_i\, R\left(\phi_i\right) + \mathbf{V}_{\text{soft}} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{pmatrix}, \tag{8.7}$$

where $R\left(\phi_i\right)$ is the two dimensional rotation matrix in the azimutal direction of each object.

The rotation is performed in either the longitudinal and transverse direction to the $\mathbf{E}_{\text{T}}^{\text{miss}}$ which gives:

$$\mathcal{S}^2 = \left(E_{\text{T}}^{\text{miss}}, 0\right) \begin{pmatrix} \sigma_L^2 & \rho_{LT}\sigma_L\sigma_T \\ \rho_{LT}\sigma_L\sigma_T & \sigma_T^2 \end{pmatrix}^{-1} \begin{pmatrix} E_{\text{T}}^{\text{miss}} \\ 0 \end{pmatrix}, \tag{8.8}$$

where the total covariance matrix is given by the rotation in the angle of the reconstructed $E_{\text{T}}^{\text{miss}}$, i.e. $\phi(E_{\text{T}}^{\text{miss}})$:

$$\mathbf{V}_{L,T} = \begin{pmatrix} \sigma_L^2 & \rho_{LT}\,\sigma_L\,\sigma_T \\ \rho_{LT}\,\sigma_L\,\sigma_T & \sigma_T^2 \end{pmatrix} = R\left(\phi(E_{\text{T}}^{\text{miss}})\right)\, \mathbf{V}_{x,y}\, R^{-1}\left(\phi(E_{\text{T}}^{\text{miss}})\right), \tag{8.9}$$

and σ_L^2, σ_T^2 are the variances in the longitudinal and tranverse directions to the $E_{\text{T}}^{\text{miss}}$, and ρ_{LT} is the correlation factor of the longitudinal L and transverse T measurements. The $E_{\text{T}}^{\text{miss}}$ significance in equation (8.8) is found to be:

$$\mathcal{S}^2 = \frac{\left|E_{\text{T}}^{\text{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}. \tag{8.10}$$

$\mathcal{S}$ carries the meaning of the $E_{\text{T}}^{\text{miss}}$ significance i.e. where the reconstucted $E_{\text{T}}^{\text{miss}}$ is in the numerator, and the information on the uncertainty propagation and directions with correlations of the objects are in the denominator. This definition has a close relationship with the event based significance.

The object based $E_{\text{T}}^{\text{miss}}$ significance, as given in equation (8.10), is plotted in Figure 8.5 for the 2015+2016 data and MC simulated processes which were first introduced in section 4.3. The Z + jets background dominates at low values ($\mathcal{S} < 10$) as expected from the $E_{\text{T}}^{\text{miss}}$ distribution. The agreement between data and prediction is seen to be very good and within the experimental uncertainty band. The uncertainty bands are obtained from the experimental object systematics and PU.

Interestingly, the tail in the $E_{\text{T}}^{\text{miss}}$ significance is found to be larger in the $Z \to ee$ channel rather than the $Z \to \mu\mu$ channel. This would imply that the objects are resolved better in the $Z \to ee$ channel. This behaviour is not seen with the event based significance (see Figure 8.6) where both lepton channels are found to perform in a similar fashion. The only possible origin for this discrepancy is in the denominator of the significance, $\sigma_L\sqrt{1-\rho^2}$, which is presented in Figure N.1. It is clear that the denominator, driven by the object resolutions, has less tails in the $Z \to ee$ channel. The reason for this is obvious from

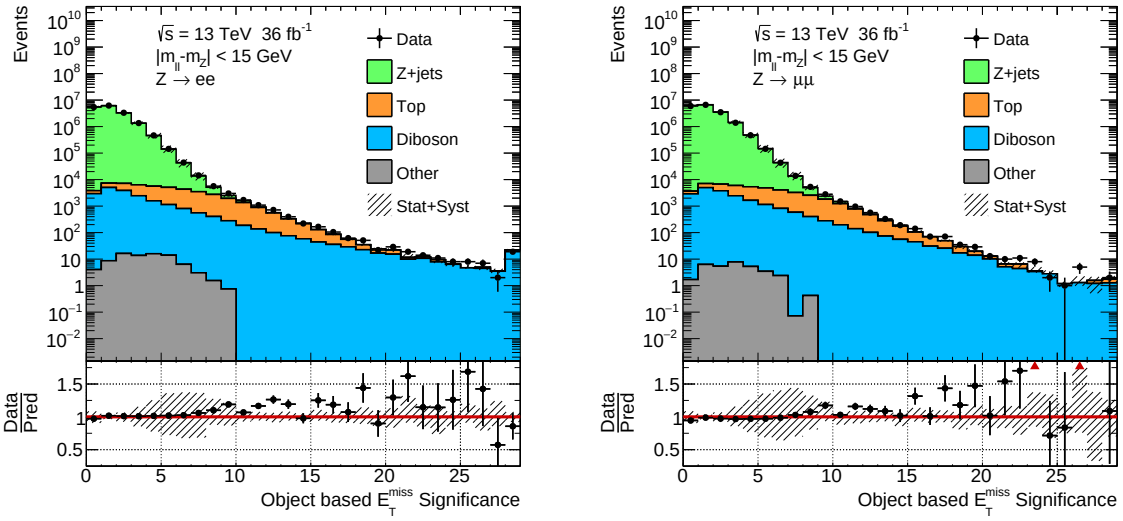


Figure 8.5: Data vs prediction for the object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance for an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV.

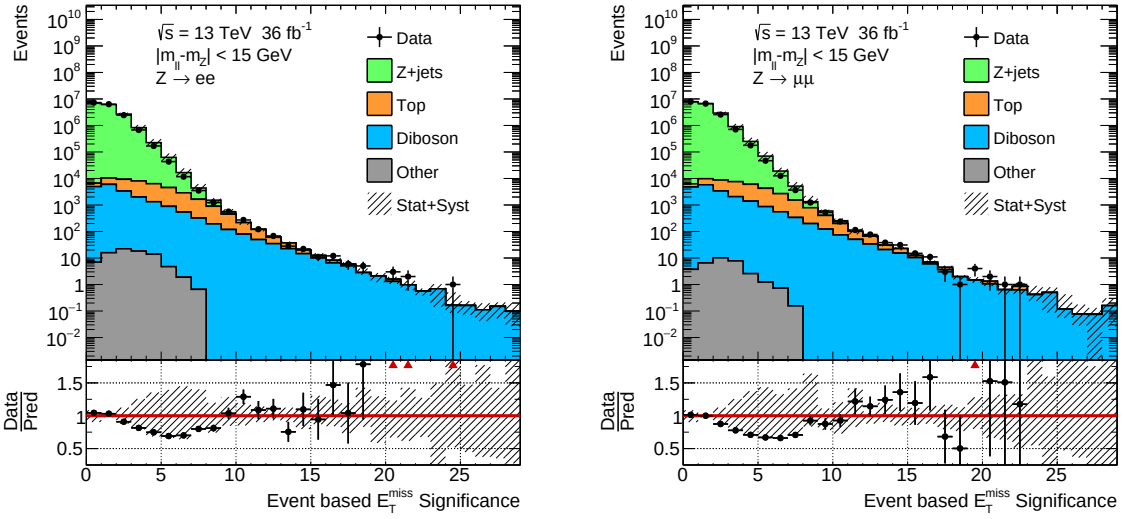


Figure 8.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, following the run I convention, between data and MC for an integrated luminosity of 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV.

Figure N.2, where the poor resolution of muon tracks at high p_{T} is seen. The resolution of the electrons at large p_{T} is much smaller because they are reconstructed in the calorimeters. The denominator, from the event based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, is shown in Figure N.3. It is clear that it fails to fully account for the resolution effects.

The performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is evaluated using the $Z \to ee$ and $ZZ \to ee\nu\nu$ event topologies. The $Z \to ee$ process contains little genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ while the $ZZ \to ee\nu\nu$ process has large missing momentum due to the invisible neutrinos. The performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is obtained through ROC curves which consider $Z \to ee$ as background and $ZZ \to ee\nu\nu$ as signal. In Figure N.4, the ROC curves for the $E_{\mathrm{T}}^{\mathrm{miss}}$, object based and

event based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance definitions are provided with a preliminary selection of $E_{\mathrm{T}}^{\mathrm{miss}}$ > 100 GeV. The expected performance of the object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is spectacular in this topology i.e. a working point with 80 % signal efficiency results in nearly 100 % $Z +$ jets background rejection. The object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is also the best performer in a di-jet topology as shown in Figure N.5.

The incorporation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is considered as a possibility for the $H \to ZZ \to \ell\ell\nu\nu$ searches. As shown above, the object based significance performs very well in topologies very similar to the analysis search. The modelling between data and simulation has good levels of agreement for both lepton channels. In addition, the object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can provide near perfect $Z +$ jets background rejection with only a minor impact on the signal efficiency. Studies within the analysis group are currently in discussion phase as the lack of man-power prevents further development in this area.

Chapter 9

Conclusions and Outlook

At the end of 2017, the halfway point of run II of the LHC has been reached and passed. With close to 80 fb^{-1} gathered, the current schedule towards the 120 fb^{-1} goal at the end of run II looks very promising. In 2016 and 2017, the LHC achieved its goal of reaching the design instantaneous luminosity, L, of $10^{+34}\text{cm}^{-2}\text{s}^{-1}$. This increase in luminosity has allowed for expanded and refined searches in the high energy regions where potential new physics can reside.

There have been interesting developments in run II, such as the excess observed in $\gamma\gamma$ searches in 2015 data, which have led to renewed interest from theoreticians in searches at the high energy frontier [209]. Similarly, the current searches undertaken may provide hints that the SM may be incomplete at the TeV scale. As was mentioned in section 7.6.2, two excesses were observed in the $H \to ZZ \to 4\ell$ searches at 240 and 700 GeV. This is where the $\ell\ell\nu\nu$ analysis played a key role, as it provided a precise measurement in the high mass region. The result was the elimination of the large excess at 700 GeV.

This thesis has summarized the search for an additional Higgs boson in the $\ell\ell\nu\nu$ final state in the first two data run periods of run II programme of the LHC. No statistically significant excess above the standard model background expectation was observed in the high mass $\ell\ell\nu\nu$ analysis. The studies presented in this thesis have led to the suppression of the Z + jets background, by optimizing an additional selection criteria using $E_{\text{T}}^{\text{miss}}/H_{\text{T}}$, in turn this has brought improvements on the expected limits. Another important point was the optimisation of VBF selections and the inclusion of the VBF signal region into the main analysis. Cleaning of PU jets was also implemented and the background estimation for the all major backgrounds was described. The exclusion limits for spin-0 Higgs production were produced and shown to have an improvement, from their run I results. In addition, an alternative signal interpretation was performed, namely the spin-2 RS Gravitons, which is a new result for the $\ell\ell\nu\nu$ analysis.

The total integrated luminosity by the end of run II is expected to be 120 fb^{-1}, i.e. ~ 3 time more than the luminosity used in this thesis (~ 36.1 fb^{-1}). The exclusion limits, in the current state of the $\ell\ell\nu\nu$ analysis, are expected to improve by an average of 46 % in the ggF category and 18 % in the VBF category.

However, with several promising developments, as outlined in chapter 8, the sensitivity can be expected to improve further. By incorporating either the PFlow $E_{\text{T}}^{\text{miss}}$ or $E_{\text{T}}^{\text{miss}}$ significance, the backgrounds generated by fake $E_{\text{T}}^{\text{miss}}$ can be more effectively suppressed, reducing the uncertainty on the overall background contribution.

Other developments are expected, one being in the form of a data-driven estimation of the largest background, $qq \to ZZ$. This can aid in reducing the uncertainty from the $k-$factors and result in a more accurate estimate.

Further studies in the VBF search can take place with an introduction of a shape-based analysis, once sufficient data has been collected.

Similarly, in the 4ℓ analysis, the excesses at 240 and 700 GeV will be further studied with more data. A future combination with the 4ℓ analysis will also be considered, to provide more accurate exclusion limits in the high mass regime, where the $\ell\ell\nu\nu$ analysis will dominate the analysis sensitivity.

Appendix A

More $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions

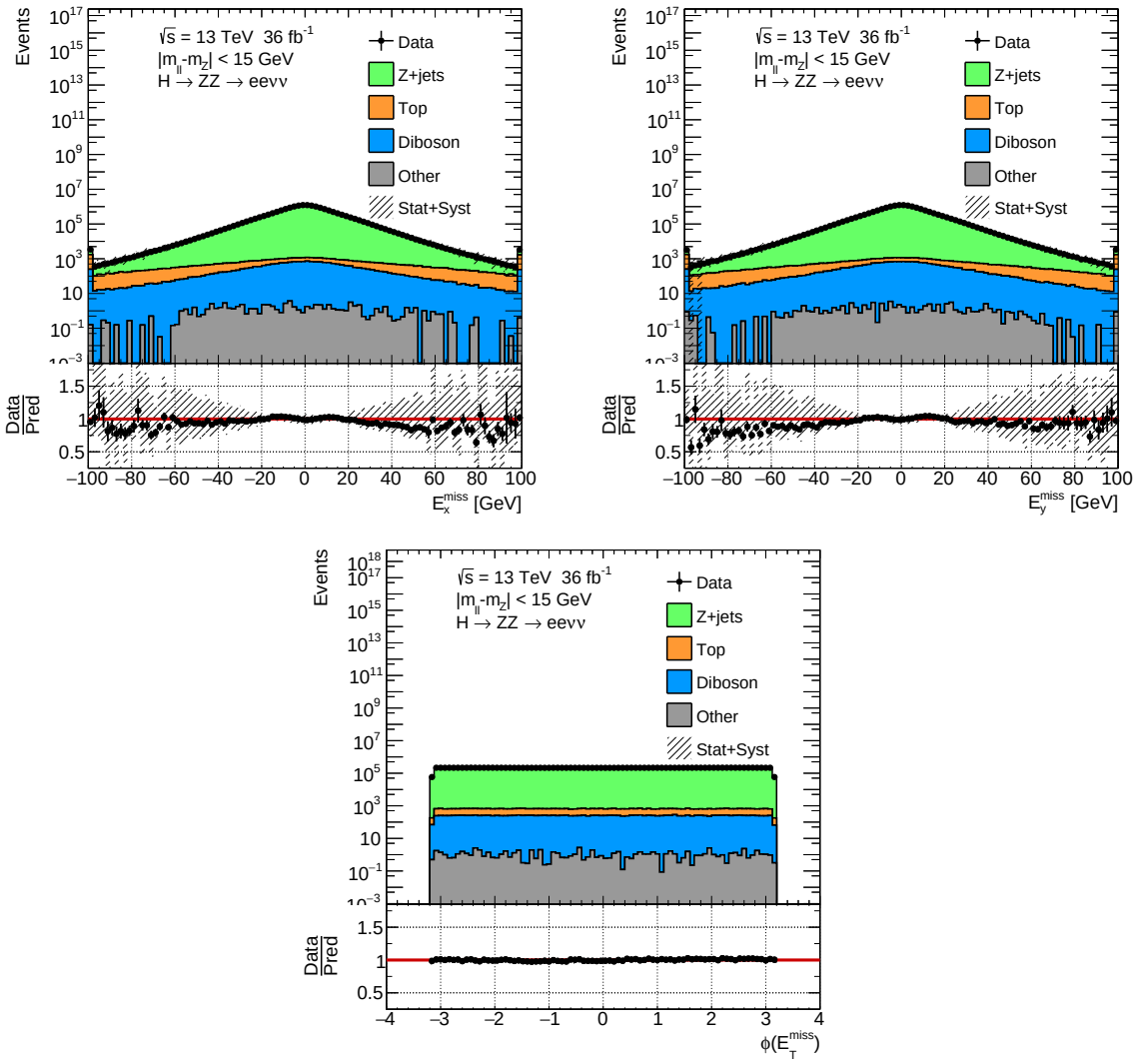


Figure A.1: The distribution of ϕ^{miss}, E_x^{miss} and E_y^{miss} between data and MC for an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV.

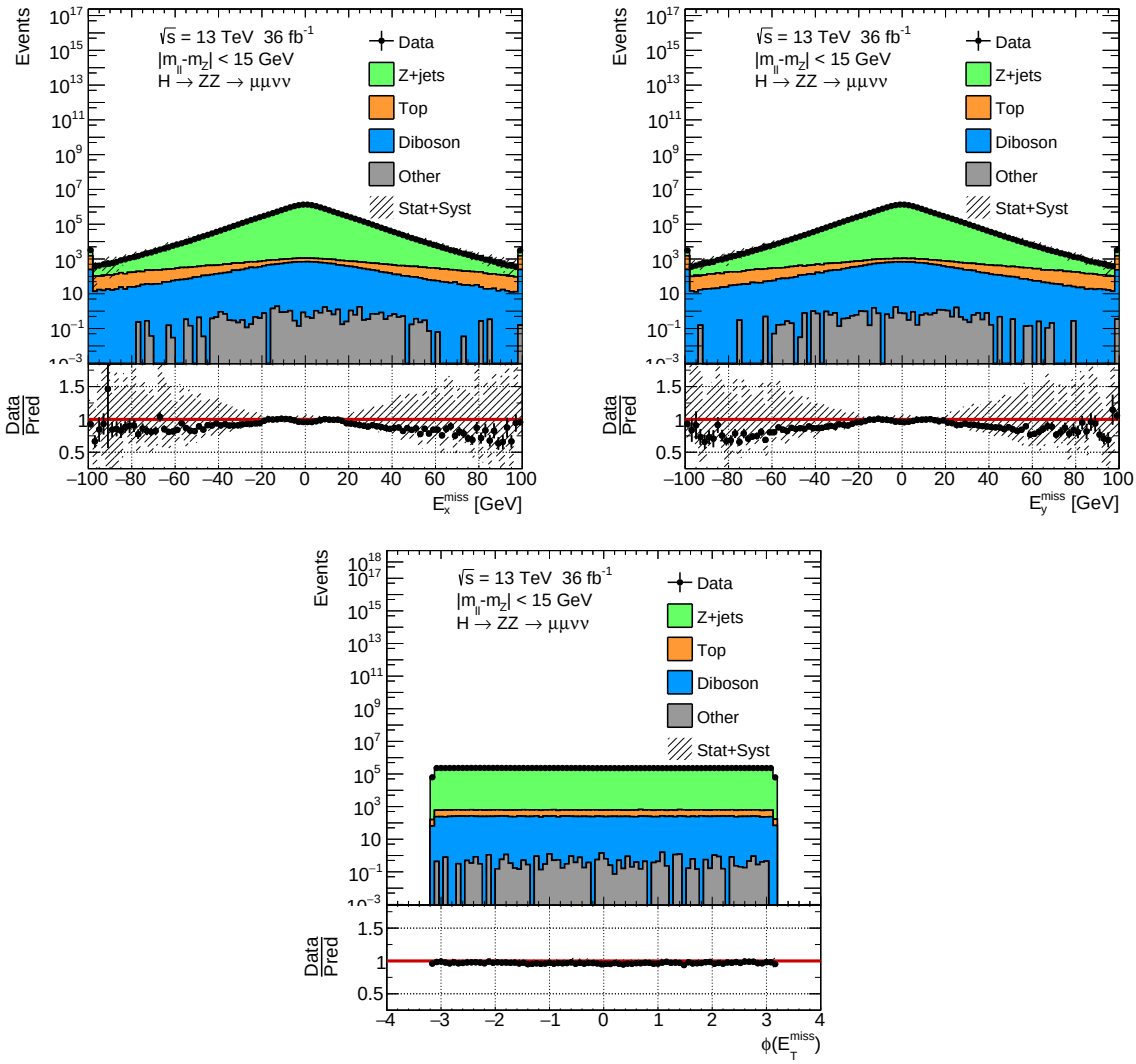


Figure A.2: The distribution of ϕ^{miss}, E_x^{miss} and E_y^{miss} between data and MC for an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV.

Appendix B

2D significance scans

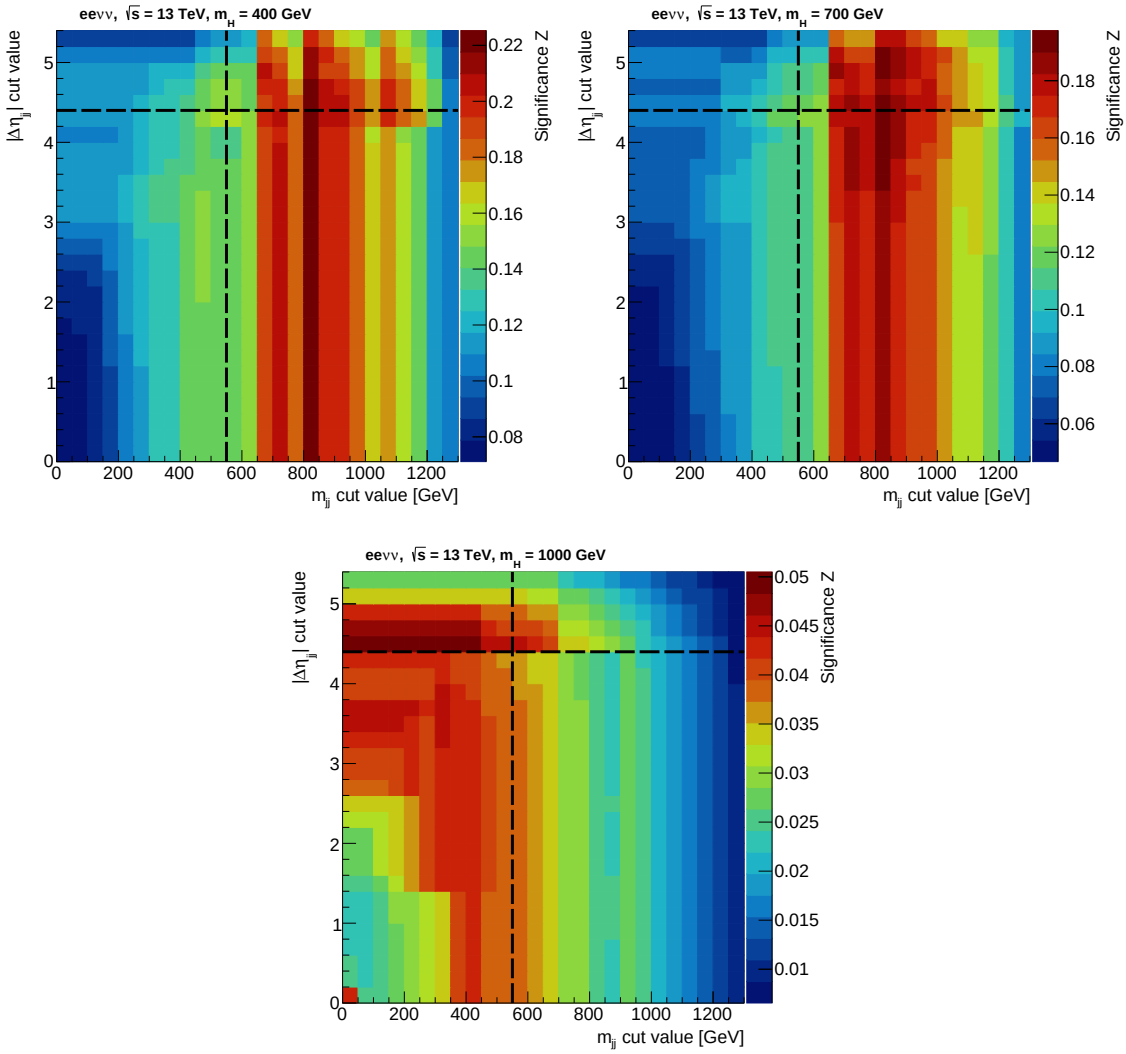


Figure B.1: Scans of the significance Z for different cut values of m_{jj} and $|\Delta\eta_{jj}|$. Black dashed lines indicate the VBF cuts of $m_{jj} > 550$ GeV and $|\Delta\eta_{jj}| > 4.4$.

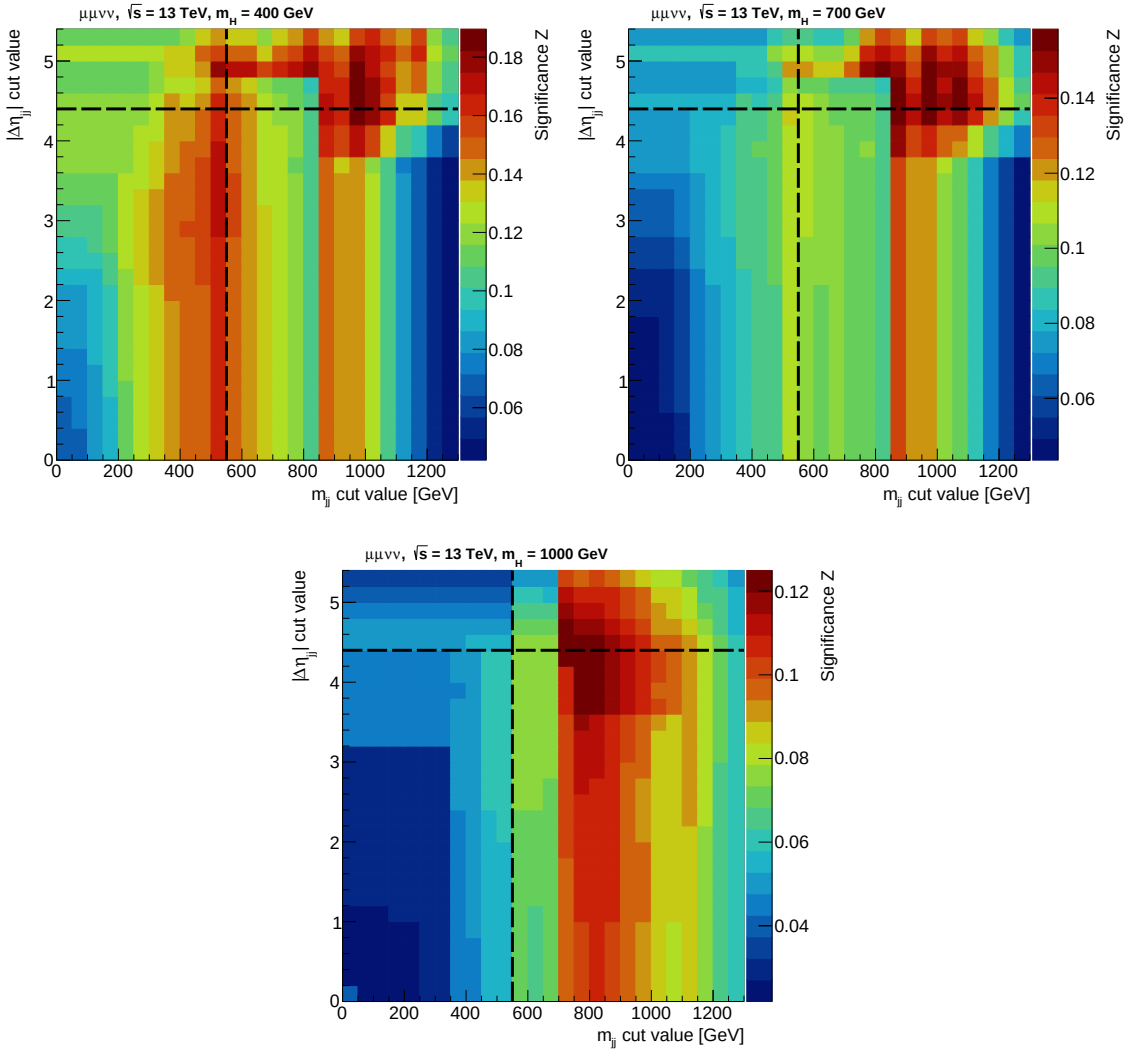


Figure B.2: Scans of the significance Z for different cut values of m_{jj} and $|\Delta\eta_{jj}|$. Black dashed lines indicate the VBF cuts of $m_{jj} > 550$ GeV and $|\Delta\eta_{jj}| > 4.4$.

Appendix C

Signal and background distributions

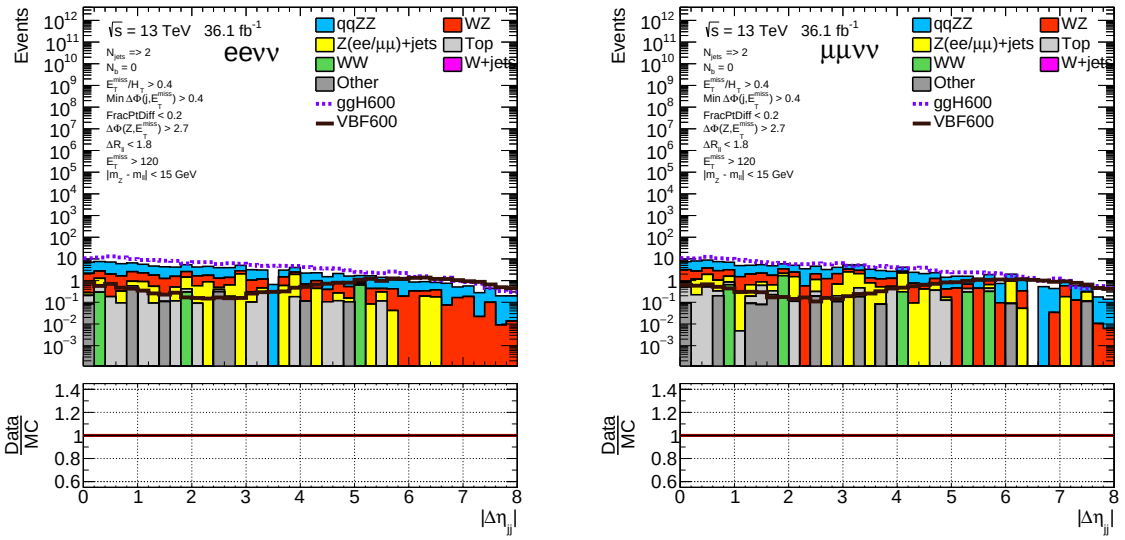


Figure C.1: Distribution for the $|\Delta\eta_{jj}|$ variable. ggF and VBF signals are scaled by an order of 5 for clarity.

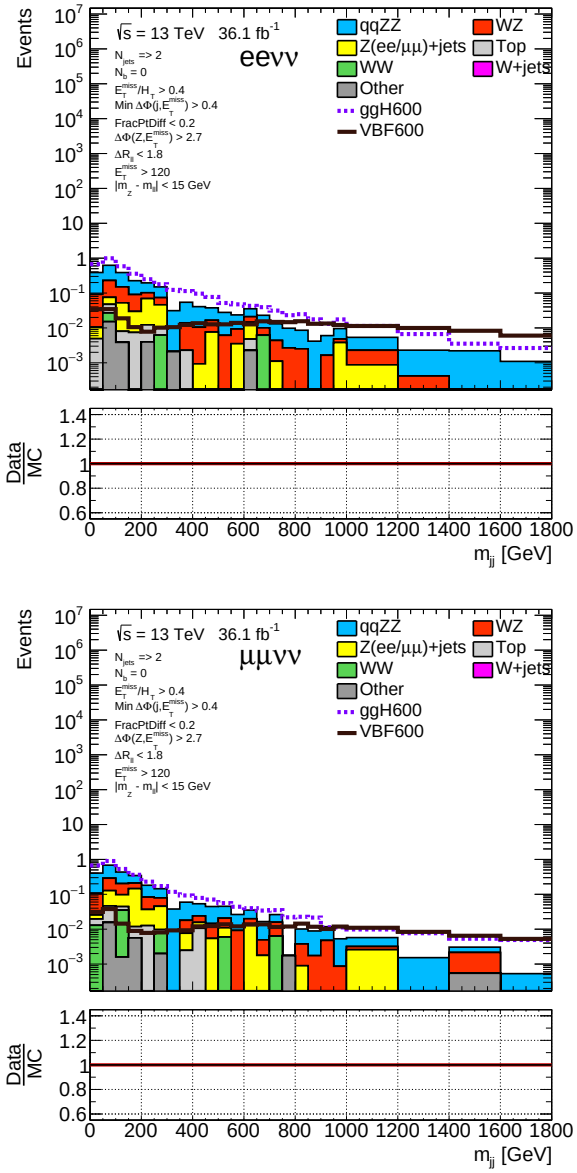


Figure C.2: Distribution for the m_{jj} variable. ggF and VBF signals are scaled by an order of 5 for clarity.

Appendix D

MC samples

MCID	Process	Generators	PDF	Events	Filter eff.	Cross section
341350	ggH300NW	Powheg + Pythia8	CT10	99600	0.44477	6.646 pb
341351	ggH400NW	Powheg + Pythia8	CT10	100000	0.44411	6.356 pb
341352	ggH500NW	Powheg + Pythia8	CT10	100000	0.44467	3.077 pb
341353	ggH600NW	Powheg + Pythia8	CT10	98000	0.44463	1.376 pb
341354	ggH700NW	Powheg + Pythia8	CT10	100000	0.44610	640.57 fb
344172	ggH750NW	Powheg + Pythia8	CT10	99000	0.44517	446.36 fb
341355	ggH800NW	Powheg + Pythia8	CT10	99800	0.44549	315.39 fb
341356	ggH900NW	Powheg + Pythia8	CT10	100000	0.44543	163.61 fb
341357	ggH1000NW	Powheg + Pythia8	CT10	99600	0.44499	88.953 fb
341358	ggH1200NW	Powheg + Pythia8	CT10	100000	0.44598	26.437 fb
341359	ggH1400NW	Powheg + Pythia8	CT10	100000	0.44521	9.546 fb
341360	ggH1600NW	Powheg + Pythia8	CT10	99800	0.44448	3.778 fb

Table D.1: Summary of the heavy higgs signal samples of gluon gluson fusion process used in this analysis. The cross sections are for Standard Model Higgs Boson of that certain mass

MCID	Process	Generators	PDF	Events	Filter eff.	Cross section
341368	VBFH300NW	Powheg + Pythia8	CT10	99600	0.44484	664.6 fb
341369	VBFH400NW	Powheg + Pythia8	CT10	99800	0.44539	635.6 fb
341370	VBFH500NW	Powheg + Pythia8	CT10	100000	0.44621	307.7 fb
341371	VBFH600NW	Powheg + Pythia8	CT10	96000	0.44497	137.6 fb
341372	VBFH700NW	Powheg + Pythia8	CT10	100000	0.44497	64.057 fb
341373	VBFH800NW	Powheg + Pythia8	CT10	99400	0.445	31.539 fb
341374	VBFH900NW	Powheg + Pythia8	CT10	99800	0.4446	16.361 fb
341375	VBFH1000NW	Powheg + Pythia8	CT10	100000	0.44513	8.8953 fb
341376	VBFH1200NW	Powheg + Pythia8	CT10	100000	0.4443	2.6437 fb
341377	VBFH1400NW	Powheg + Pythia8	CT10	100000	0.44543	0.9546 fb
341378	VBFH1600NW	Powheg + Pythia8	CT10	98000	0.44569	0.3778 fb

Table D.2: Summary of the heavy Higgs signal samples of vector boson fusion process used in this analysis.

MCID	Process	Generators	PDF	Events	Filter eff.	Cross section
303335	RS_G_ZZ_llvv_c10_m0600	MadGraphPythia8	NNPDF23LO	30000	1	31.41 pb
303336	RS_G_ZZ_llvv_c10_m0800	MadGraphPythia8	NNPDF23LO	30000	1	5.915 pb
303337	RS_G_ZZ_llvv_c10_m1000	MadGraphPythia8	NNPDF23LO	28000	1	1.644 pb
303338	RS_G_ZZ_llvv_c10_m1200	MadGraphPythia8	NNPDF23LO	30000	1	0.5686 pb
303339	RS_G_ZZ_llvv_c10_m1400	MadGraphPythia8	NNPDF23LO	30000	1	0.2251 pb
303340	RS_G_ZZ_llvv_c10_m1600	MadGraphPythia8	NNPDF23LO	30000	1	0.09837 pb
303341	RS_G_ZZ_llvv_c10_m1800	MadGraphPythia8	NNPDF23LO	29000	1	0.04606 pb
303342	RS_G_ZZ_llvv_c10_m2000	MadGraphPythia8	NNPDF23LO	30000	1	0.0228 pb

Table D.3: Summary of the RS graviton samples used in this analysis.

Process	Generator	σ (pb)	Filter efficiency
$qq \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$	POWHEG+Pythia8	0.91795	1.0
$gg \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$	POWHEGgg2vv+Pythia8	0.05187	0.66248
$qq \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$	SHERPA	12.465	1.0

Table D.4: Cross sections at NLO in perturbation theory and filter efficiency for ZZ `Powheg` + `Pythia8` samples at $\sqrt{s} = 13$ TeV. Processes listed refer to both quark-quark and gluon-gluon production modes. It must be noted that the qq cross section values cannot be compared by simply accounting for the difference in the branching ratios among the various final states; in fact the samples are generated with $m_{\ell\ell} > 4$ GeV which gives access to the Drell-Yan diagrams in the 4ℓ final state, diagrams that do not contribute to the $\ell^+\ell^-\nu\nu$ final state.

Process	Generator	σ (pb)	Filter efficiency
$WZ \to \ell\nu\ell^+\ell^-$	POWHEG+Pythia8	4.4625	1.0
$WZ \to \ell\nu 2\nu$	POWHEG+Pythia8	2.7778	1.0
$WW \to \ell\nu\ell\nu$	POWHEG+Pythia8	10.631	1.0
$WZ \to \ell\nu\ell^+\ell^-$	SHERPA	4.583	1.0

Table D.5: Cross sections at NLO in perturbation theory and filter efficiency for WZ and WW `Powheg` + `Pythia8` samples at $\sqrt{s} = 13$ TeV.

Appendix E

Modelling of $p_{\mathrm{T}}^{\mathrm{Z}}$

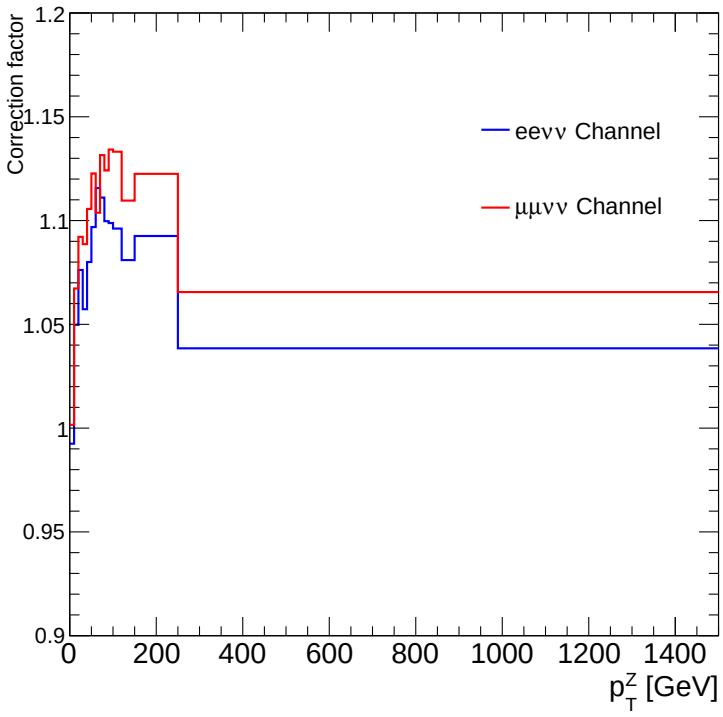


Figure E.1: Correction factor applied to $p_{\mathrm{T}}^{\mathrm{Z}}$ distribution to correct mis-modelling between data and simulation in the inclusive distributions.

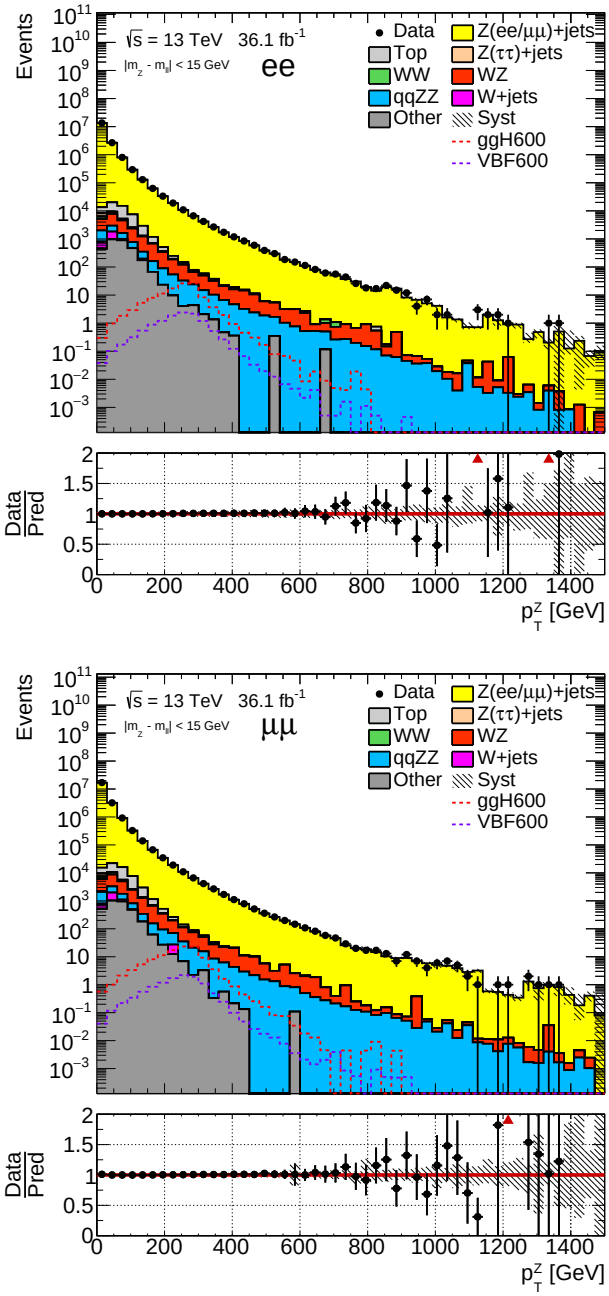


Figure E.2: Z boson distribution after application of data/MC mis-modelling correction. Left: $ee\nu\nu$ channel, Right: $\mu\mu\nu\nu$ channel.

Appendix F

Efficiency of applied selections on signal and background

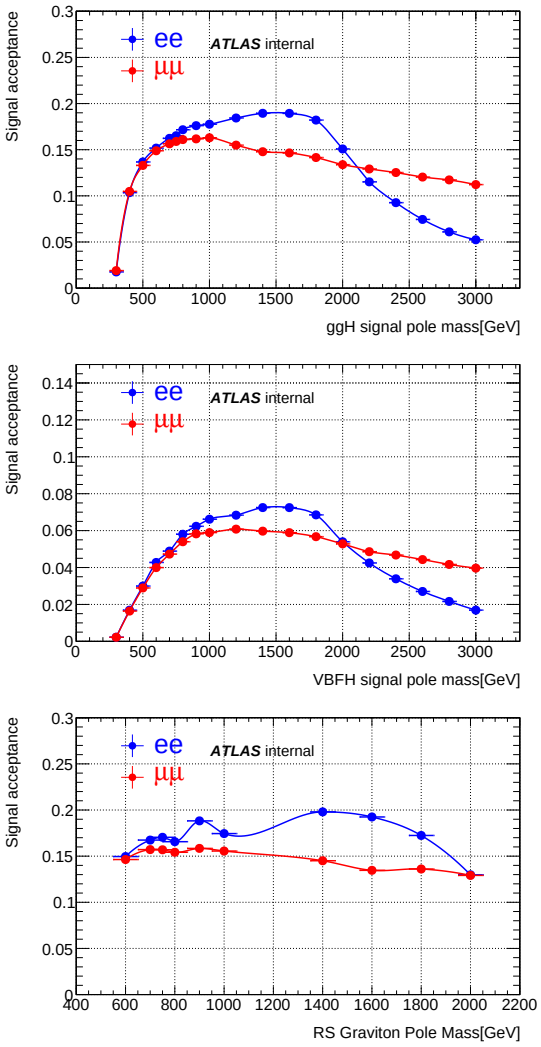


Figure F.1: The signal acceptance versus resonance mass, for ggF heavy Higgs, for VBF heavy Higgs and for Graviton. Only the statistical uncertainty is shown.

Selection	$ee\nu\nu$			$\mu\mu\nu\nu$		
	Yield	Abs(%)	Rel(%)	Yield	Abs(%)	Rel(%)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV	381.00	100.00	100.00	379.00	100.00	100.00
$E_T^{\text{miss}} > 120$ GeV	381.00	100.00	100.00	377.00	99.47	99.47
$\Delta R < 1.8$	380.00	99.74	99.74	377.00	99.47	100.00
$\Delta\phi(Z, E_T^{\text{miss}}) > 2.7$	333.00	87.40	87.63	335.00	88.39	88.86
$\lvert E_T^{\text{miss}}(\text{withjets}) - p_{\text{T}}^{\ell\ell}\rvert / p_{\text{T}}^{\ell\ell} < 0.2$	327.00	85.83	98.20	334.00	88.13	99.70
$min(\Delta\phi(jet, E_T^{\text{miss}})) > 0.4$	315.00	82.68	96.33	322.00	84.96	96.41
$E_T^{\text{miss}}/H_T > 0.4$	310.00	81.36	98.41	313.00	82.59	97.20
$N_b = 0$	299.00	78.48	96.45	293.00	77.31	93.61
$N_j \geq 2$	117.00	30.71	39.13	101.00	26.65	34.47
$p_T^{j_1}, p_T^{j_2} > 30$ GeV	35.00	9.19	29.91	37.00	9.76	36.63
$M_{jj} > 550$ GeV	2.00	0.52	5.71	5.00	1.32	13.51
$\lvert\Delta\eta_{jj}\rvert > 4.4$	1.00	0.26	50.00	4.00	1.06	80.00

Table F.1: Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $qq \to ZZ \to \ell\ell\nu\nu$ in a m_{ZZ} window between 550 and 650 GeV.

Selection	$ee\nu\nu$			$\mu\mu\nu\nu$		
	Yield	Abs(%)	Rel(%)	Yield	Abs(%)	Rel(%)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV	500.00	100.00	100.00	486.00	100.00	100.00
$E_T^{\text{miss}} > 120$ GeV	285.00	57.00	57.00	289.00	59.47	59.47
$\Delta R < 1.8$	282.00	56.40	98.95	288.00	59.26	99.65
$\Delta\phi(Z, E_T^{\text{miss}}) > 2.7$	232.00	46.40	82.27	234.00	48.15	81.25
$\lvert E_T^{\text{miss}}(\text{withjets}) - p_{\text{T}}^{\ell\ell}\rvert / p_{\text{T}}^{\ell\ell} < 0.2$	227.00	45.40	97.84	232.00	47.74	99.15
$min(\Delta\phi(jet, E_T^{\text{miss}})) > 0.4$	191.00	38.20	84.14	197.00	40.53	84.91
$E_T^{\text{miss}}/H_T > 0.4$	169.00	33.80	88.48	172.00	35.39	87.31
$N_b = 0$	141.00	28.20	83.43	150.00	30.86	87.21
$N_j \geq 2$	75.00	15.00	53.19	97.00	19.96	64.67
$p_T^{j_1}, p_T^{j_2} > 30$ GeV	44.00	8.80	58.67	44.00	9.05	45.36
$M_{jj} > 550$ GeV	2.00	0.40	4.55	4.00	0.82	9.09
$\lvert\Delta\eta_{jj}\rvert > 4.4$	2.00	0.40	100.00	4.00	0.82	100.00

Table F.2: Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the WZ background in a m_{ZZ} window between 550 and 650 GeV.

Selection	$ee\nu\nu$			$\mu\mu\nu\nu$		
	Yield	Abs(%)	Rel(%)	Yield	Abs(%)	Rel(%)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV	17717383.00	100.0000	100.00	19583119.00	100.0000	100.00
$E_T^{\text{miss}} > 120$ GeV	29640.00	0.1673	0.17	40551.00	0.2071	0.21
$\Delta R < 1.8$	20688.00	0.1168	69.80	30361.00	0.1550	74.87
$\Delta\phi(Z, E_T^{\text{miss}}) > 2.7$	10621.00	0.0599	51.34	17499.00	0.0894	57.64
$\lvert E_T^{\text{miss}}(\text{withjets}) - p_{\text{T}}^{\ell\ell}\rvert / p_{\text{T}}^{\ell\ell} < 0.2$	10050.00	0.0567	94.62	16667.00	0.0851	95.25
$min(\Delta\phi(jet, E_T^{\text{miss}})) > 0.4$	2436.00	0.0137	24.24	3380.00	0.0173	20.28
$E_T^{\text{miss}}/H_T > 0.4$	1133.00	0.0064	46.51	1239.00	0.0063	36.66
$N_b = 0$	316.00	0.0018	27.89	337.00	0.0017	27.20
$N_j \geq 2$	220.00	0.0012	69.62	228.00	0.0012	67.66
$p_T^{j_1}, p_T^{j_2} > 30$ GeV	115.00	0.0006	52.27	119.00	0.0006	52.19
$M_{jj} > 550$ GeV	9.00	0.0001	7.83	6.00	0.0000	5.04
$\lvert\Delta\eta_{jj}\rvert > 4.4$	7.00	0.0000	77.78	5.00	0.0000	83.33

Table F.3: Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the Z + jets.

Selection	$ee\nu\nu$			$\mu\mu\nu\nu$		
	Yield	Abs(%)	Rel(%)	Yield	Abs(%)	Rel(%)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV	21402.00	100.00	100.00	21217.00	100.00	100.00
$E_T^{\text{miss}} > 120$ GeV	19995.00	93.43	93.43	19773.00	93.19	93.19
$\Delta R < 1.8$	19772.00	92.38	98.88	19530.00	92.05	98.77
$\Delta\phi(Z, E_T^{\text{miss}}) > 2.7$	16788.00	78.44	84.91	16528.00	77.90	84.63
$\lvert E_T^{\text{miss}}(\text{withjets}) - p_{\text{T}}^{\ell\ell}\rvert / p_{\text{T}}^{\ell\ell} < 0.2$	16627.00	77.69	99.04	16333.00	76.98	98.82
$min(\Delta\phi(jet, E_T^{\text{miss}})) > 0.4$	16262.00	75.98	97.80	15987.00	75.35	97.88
$E_T^{\text{miss}}/H_T > 0.4$	15756.00	73.62	96.89	15470.00	72.91	96.77
$N_b = 0$	14812.00	69.21	94.01	14540.00	68.53	93.99
$N_j \geq 2$	7380.00	34.48	49.82	7213.00	34.00	49.61
$p_T^{j_1}, p_T^{j_2} > 30$ GeV	2991.00	13.98	40.53	2816.00	13.27	39.04
$M_{jj} > 550$ GeV	364.00	1.70	12.17	359.00	1.69	12.75
$\lvert\Delta\eta_{jj}\rvert > 4.4$	325.00	1.52	89.29	332.00	1.56	92.48

Table F.4: Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $gg \to H \to ZZ$ sample ($m_{\text{H}} = 600$ GeV).

Selection	$ee\nu\nu$			$\mu\mu\nu\nu$		
	Yield	Abs(%)	Rel(%)	Yield	Abs(%)	Rel(%)
$\lvert m_Z - m_{\ell\ell}\rvert < 15$ GeV	21438.00	100.00	100.00	20921.00	100.00	100.00
$E_T^{\text{miss}} > 120$ GeV	19937.00	93.00	93.00	19375.00	92.61	92.61
$\Delta R < 1.8$	19643.00	91.63	98.53	19034.00	90.98	98.24
$\Delta\phi(Z, E_T^{\text{miss}}) > 2.7$	15333.00	71.52	78.06	14844.00	70.95	77.99
$\lvert E_T^{\text{miss}}(\text{withjets}) - p_{\text{T}}^{\ell\ell}\rvert / p_{\text{T}}^{\ell\ell} < 0.2$	15196.00	70.88	99.11	14709.00	70.31	99.09
$min(\Delta\phi(jet, E_T^{\text{miss}})) > 0.4$	14603.00	68.12	96.10	14133.00	67.55	96.08
$E_T^{\text{miss}}/H_T > 0.4$	13774.00	64.25	94.32	13401.00	64.06	94.82
$N_b = 0$	13129.00	61.24	95.32	12749.00	60.94	95.13
$N_j \geq 2$	9197.00	42.90	70.05	8809.00	42.11	69.10
$p_T^{j_1}, p_T^{j_2} > 30$ GeV	5792.00	27.02	62.98	5461.00	26.10	61.99
$M_{jj} > 550$ GeV	4357.00	20.32	75.22	4065.00	19.43	74.44
$\lvert\Delta\eta_{jj}\rvert > 4.4$	4251.00	19.83	97.57	3962.00	18.94	97.47

Table F.5: Cutflow table presenting absolute event counts as well as relative and absolute efficiencies for the $qq \to qqH$ ($m_{\text{H}} = 600$ GeV) sample.

Appendix G

Pile-up jet events

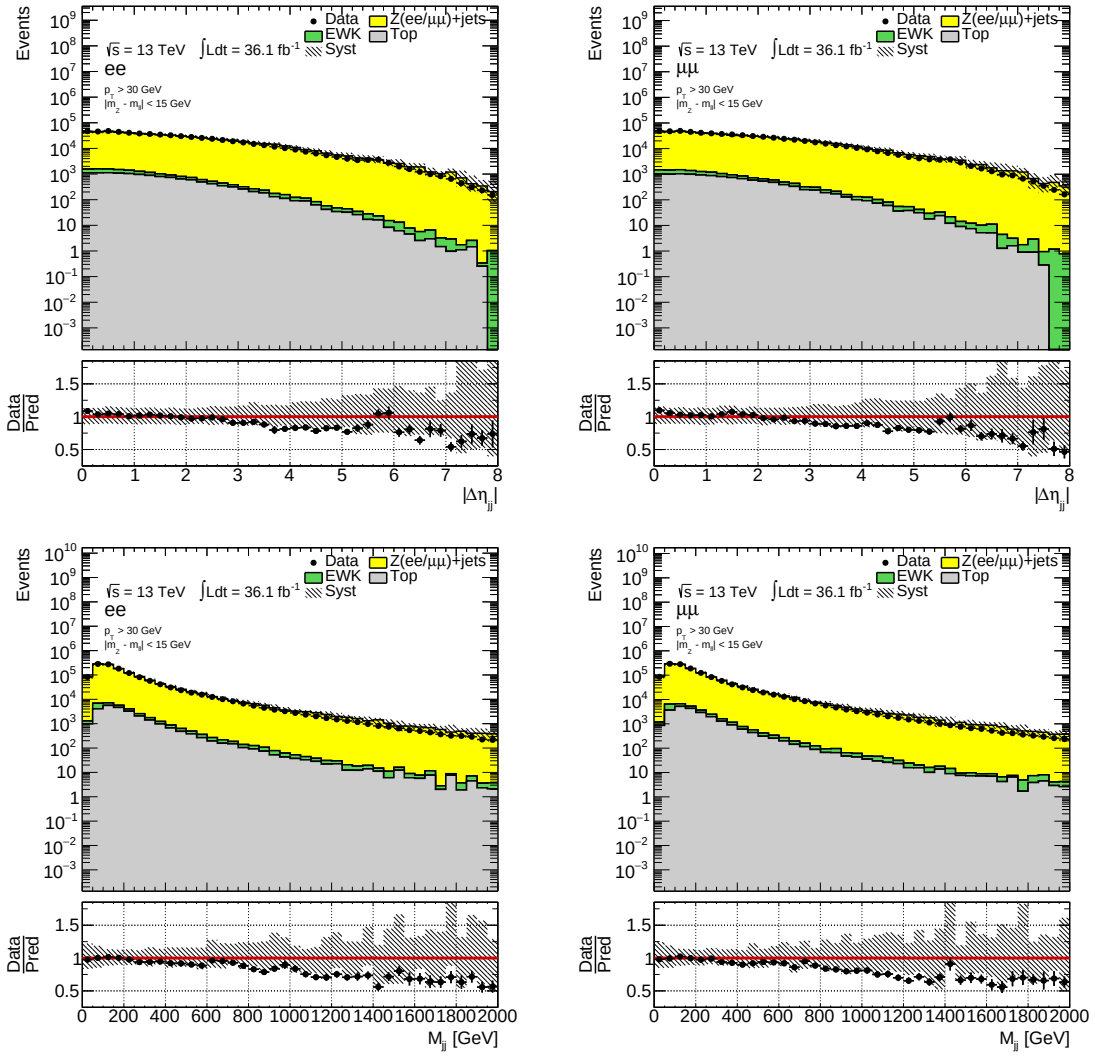


Figure G.1: Removal of low p_{T} PU jets by raising the minimum jet p_{T} requirement from 20 to 30 GeV. Agreement between data and simulation is seen to improve in the problematic regions.

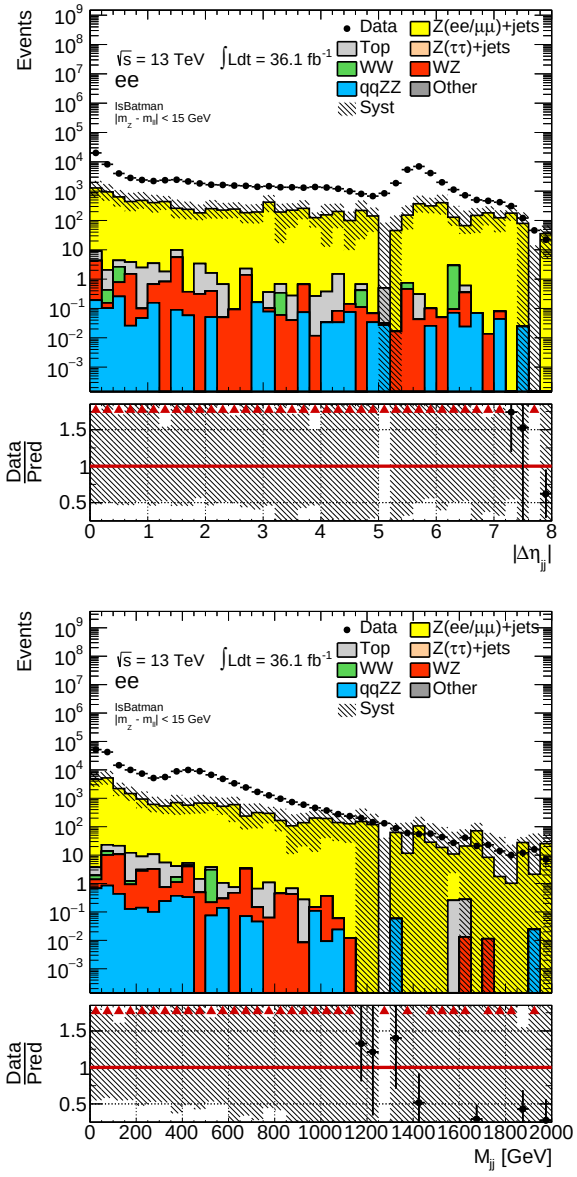


Figure G.2: The removal of the PU jets in the *ee* channel, following the official selections released by the ATLAS Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ group. The overall reduction in data combined from the $ee\nu\nu$ and $\mu\mu\nu\nu$ channels is $\sim 3\%$.

Appendix H

3ℓ control region plots

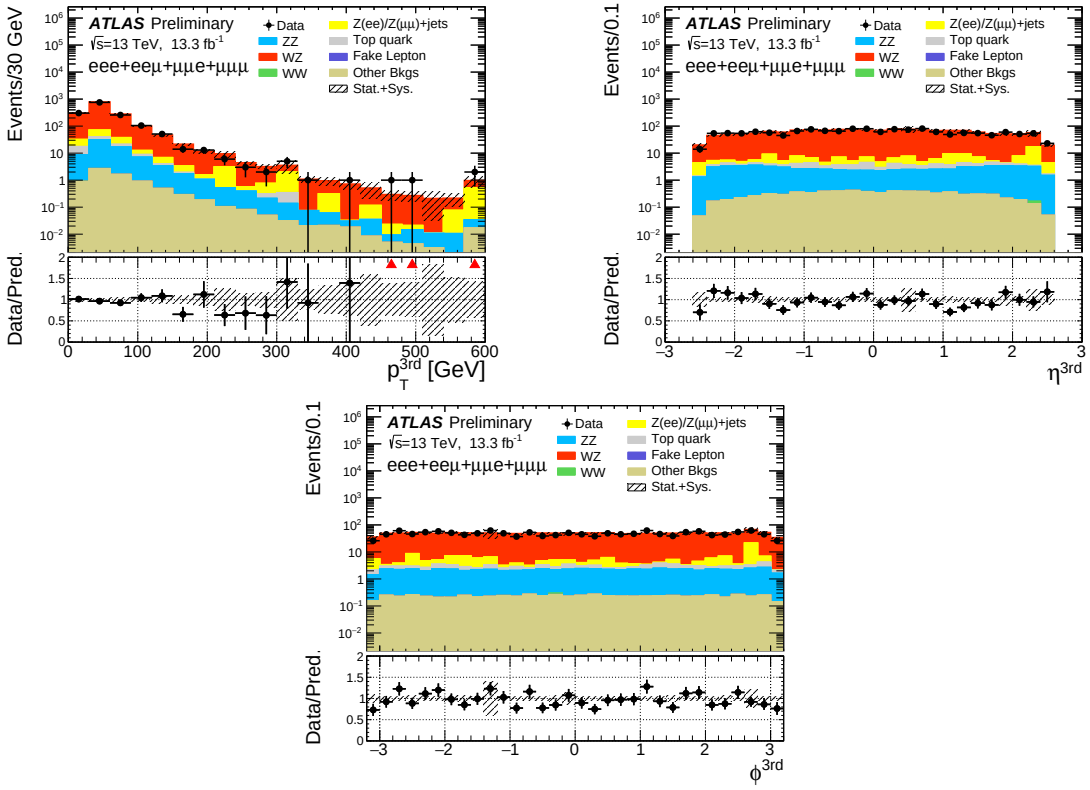


Figure H.1: Data/MC modelling of the $p_{\mathrm{T}}, \eta, \phi$ of the third lepton in the three lepton control region. All individual lepton channels are combined. Figures produced from an analysis collaborator.

Appendix I

Efficiency factor binned in p_{T}, η

Lepton p_{T} (lead, sub-lead)	BB	BE	EB	EE
(30-44,20-44)	0.9518 ± 0.0015	0.8785 ± 0.0018	0.8509 ± 0.0018	0.8386 ± 0.0044
(44-52,20-44)	0.9757 ± 0.0016	0.8522 ± 0.0031	0.8334 ± 0.0030	0.8120 ± 0.0024
(44-52,44-52)	0.9402 ± 0.0029	0.7791 ± 0.0074	0.7820 ± 0.0076	0.7347 ± 0.0034
(52-2000,20-44)	1.0029 ± 0.0021	0.8685 ± 0.0040	0.8662 ± 0.0038	0.8009 ± 0.0032
(52-2000,44-52)	0.9730 ± 0.0139	0.8815 ± 0.0108	0.8717 ± 0.0106	0.7810 ± 0.0077
(52-2000,52-2000)	1.0434 ± 0.0037	0.9208 ± 0.0064	0.9165 ± 0.0067	0.8576 ± 0.0064

Table I.1: Efficiency factor values in bin of p_{T} and η from MC. Errors contain statistical uncertainty only.

Lepton p_{T} (lead, sub-lead)	BB	BE	EB	EE
(30-44,20-44)	0.9469 ± 0.0004	0.8745 ± 0.0005	0.8429 ± 0.0004	0.8291 ± 0.0012
(44-52,20-44)	0.9686 ± 0.0004	0.8435 ± 0.0007	0.8226 ± 0.0006	0.7985 ± 0.0006
(44-52,44-52)	0.9380 ± 0.0007	0.7742 ± 0.0019	0.7748 ± 0.0019	0.7248 ± 0.0009
(52-2000,20-44)	0.9980 ± 0.0005	0.8715 ± 0.0008	0.8537 ± 0.0008	0.7898 ± 0.0007
(52-2000,44-52)	0.9808 ± 0.0014	0.8605 ± 0.0030	0.8648 ± 0.0030	0.7639 ± 0.0018
(52-2000,52-2000)	1.0331 ± 0.0019	0.9024 ± 0.0037	0.9180 ± 0.0038	0.8395 ± 0.0028

Table I.2: Efficiency factor values in bin of p_{T} and η from Data. Errors contain statistical uncertainty only.

Appendix J

Z + jets bias checks

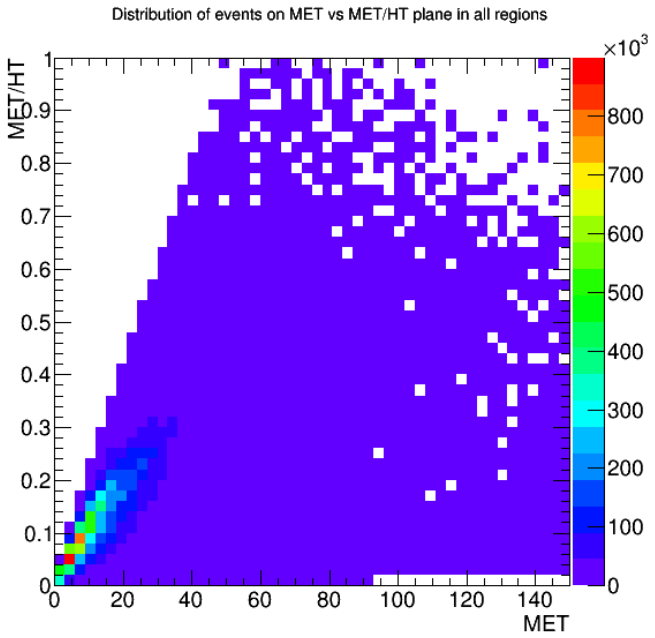


Figure J.1: Events in each region on $E_{\mathrm{T}}^{\mathrm{miss}}$ vs $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ plane, all regions combined. $H \to ZZ \to ee\nu\nu$-channel. $H \to ZZ \to ee\nu\nu$ channel. Figure produced from an analysis collaborator.

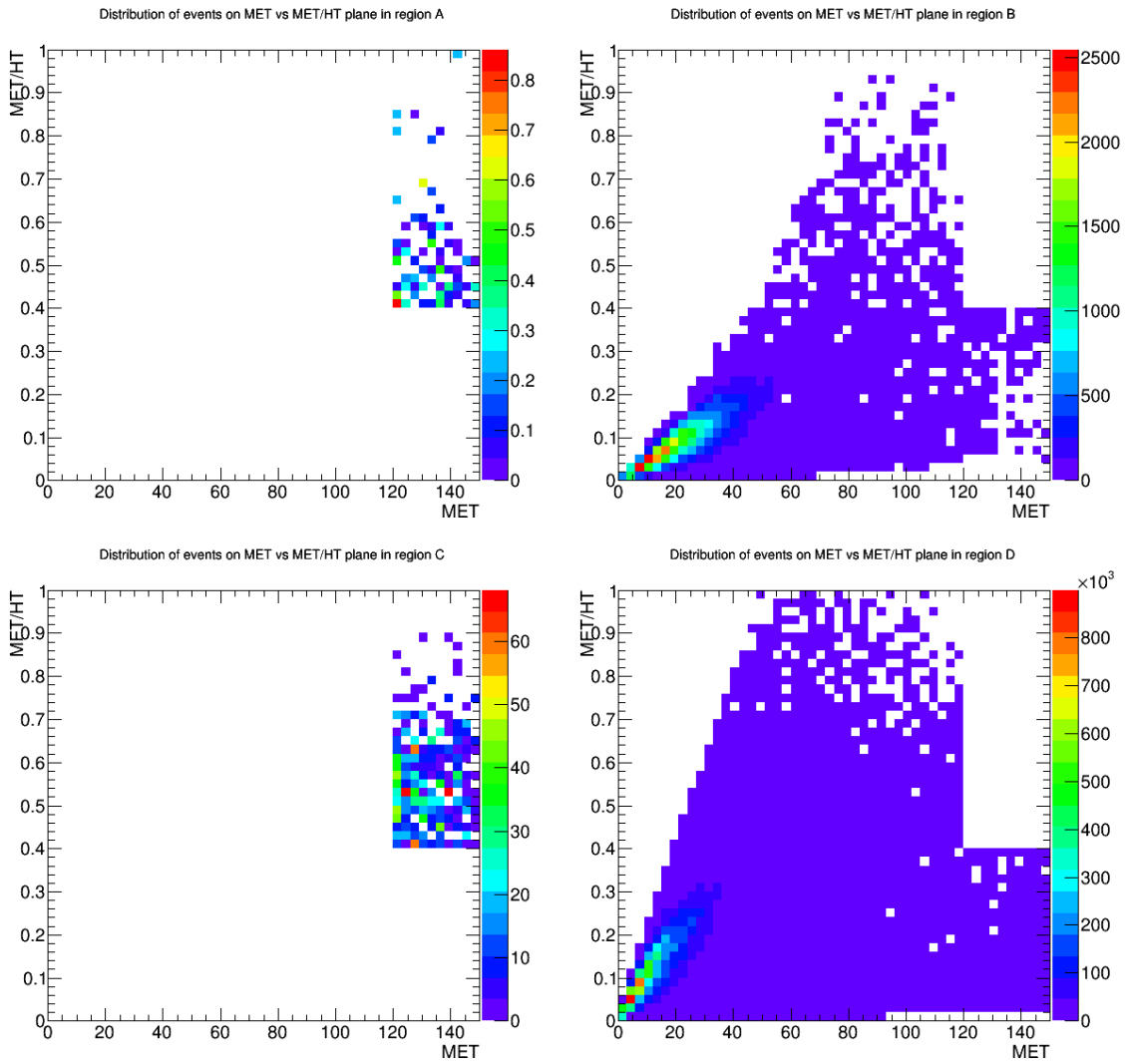


Figure J.2: Events in each region on $E_{\mathrm{T}}^{\mathrm{miss}}$ vs $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ plane, (a) - (d) for corresponding regions. Highly correlated Z + jets bulk is easily seen in B and D control regions. $H \to ZZ \to ee\nu\nu$-channel. Figure produced from an analysis collaborator.

Appendix K

One side band $E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}}$ modelling

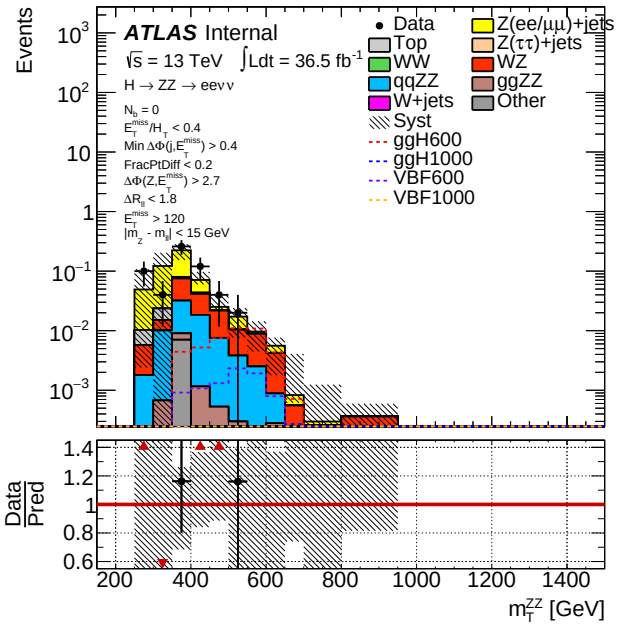


Figure K.1: m_{T}^{ZZ} distributions in the sideband ($E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} < 0.4$) regions after the full selection applied for the $H \to ZZ \to ee\nu\nu$ channel. The error band in the ratio plot contains the systematic uncertainty, while the error bars on the points contain the data + MC statistical uncertainty.

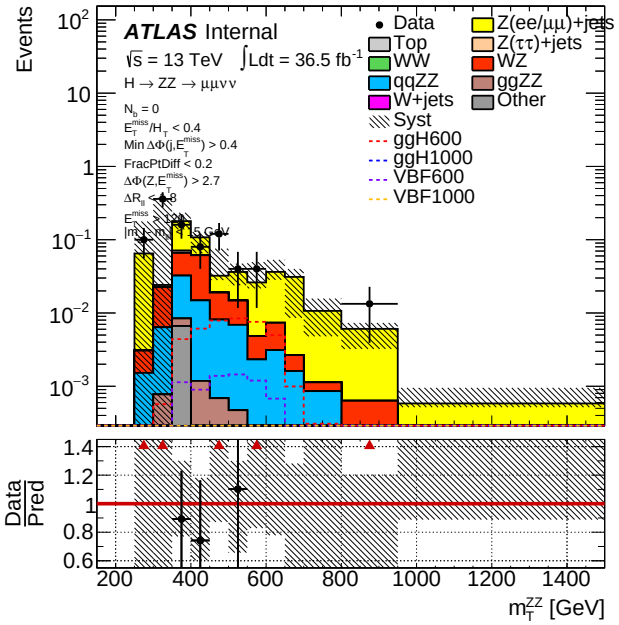


Figure K.2: m_{T}^{ZZ} distributions in the sideband ($E_{\mathrm{T}}^{\mathrm{miss}}/H_{\mathrm{T}} < 0.4$) regions after the full selection applied for the $H \to ZZ \to \mu\mu\nu\nu$ channel. The error band in the ratio plot contains the systematic uncertainty, while the error bars on the points contain the data + MC statistical uncertainty.

Appendix L

Theoretical Uncertainties

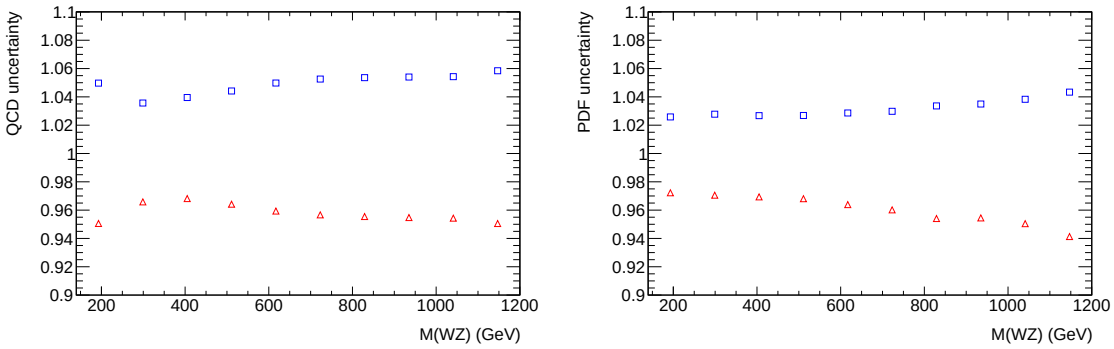


Figure L.1: QCD and PDF scale included shape uncertainties as a function of the mass of the WZ resonance. Figure produced from an analysis collaborator.

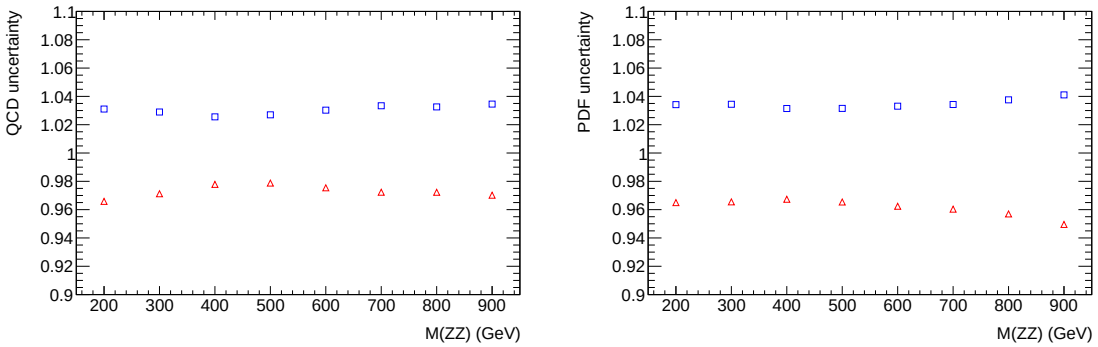


Figure L.2: Theoretical systematic uncertainties on the ZZ background due to QCD scale and PDF variation, as a function of the transverse mass of the ZZ resonance. Figure produced from an analysis collaborator.

Appendix M

Upper limits

Limits on ggF produced heavy Higgs				
Mass Point [GeV]	Expected Limits [fb]	Exp+2σ, Exp+1σ	Exp-1σ, Exp-2σ	Observed Limits
ee channel				
300.0 GeV	1159.26	2259.26 , 1637.31	835.31 , 622.20	1968.06
400.0 GeV	248.45	486.63 , 352.86	179.02 , 133.35	258.23
500.0 GeV	129.72	257.88 , 184.90	93.47 , 69.62	204.01
600.0 GeV	81.90	166.66 , 117.68	59.01 , 43.96	64.96
700.0 GeV	59.99	125.00 , 86.96	43.23 , 32.20	38.06
800.0 GeV	47.50	100.23 , 69.25	34.23 , 25.50	26.77
900.0 GeV	36.73	79.29 , 54.00	26.46 , 19.71	23.82
1000.0 GeV	37.34	79.57 , 54.69	26.90 , 20.04	25.04
1200.0 GeV	26.84	59.41 , 39.87	19.34 , 14.41	20.59
1400.0 GeV	23.09	51.64 , 34.42	16.64 , 12.39	18.66
$\mu\mu$ channel				
300.0 GeV	1160.43	2298.49 , 1644.86	836.15 , 622.83	1411.22
400.0 GeV	274.35	547.03 , 389.23	197.69 , 147.25	254.58
500.0 GeV	140.74	278.04 , 200.01	101.41 , 75.54	172.62
600.0 GeV	95.17	191.19 , 136.08	68.57 , 51.08	109.28
700.0 GeV	69.63	142.29 , 100.20	50.17 , 37.37	96.85
800.0 GeV	54.10	112.89 , 78.47	38.98 , 29.04	64.14
900.0 GeV	43.32	93.09 , 63.56	31.22 , 23.25	52.80
1000.0 GeV	41.49	88.53 , 60.80	29.90 , 22.27	58.10
1200.0 GeV	33.74	74.62 , 50.03	24.31 , 18.11	48.86
1400.0 GeV	31.15	69.82 , 46.46	22.44 , 16.72	44.79
combined channel				
300.0 GeV	863.56	1679.65 , 1218.76	622.25 , 463.50	1320.04
400.0 GeV	191.93	371.53 , 270.52	138.29 , 103.01	168.34
500.0 GeV	96.01	187.45 , 135.89	69.18 , 51.53	142.67
600.0 GeV	60.62	120.53 , 86.39	43.68 , 32.54	53.60
700.0 GeV	43.65	88.28 , 62.60	31.45 , 23.43	38.49
800.0 GeV	34.04	69.60 , 49.01	24.53 , 18.27	23.79
900.0 GeV	26.17	54.75 , 38.01	18.86 , 14.05	21.01
1000.0 GeV	26.12	53.99 , 37.77	18.82 , 14.02	24.27
1200.0 GeV	19.33	41.24 , 28.32	13.93 , 10.38	20.26
1400.0 GeV	16.99	37.67 , 24.98	12.24 , 9.12	18.28

Table M.1: The observed and expected 95% CL upper limits on the gluon fusion production cross-section of an additional narrow-width heavy Higgs boson times the branching fraction of Higgs to ZZ decay $\sigma\times\mathrm{BR}(H \to ZZ)$, shown for mass points of 300GeV, 400GeV, 500GeV, 600GeV, 700GeV, 750GeV, 800GeV, 900GeV, 1TeV, 1.2TeV and 1.4TeV , respectively. Limits are derived using ee, $\mu\mu$ and combined ($ee + \mu\mu$) channels, as indicated in the table. **Both the statistical and systematic uncertainties are considered in the limit setting**.

Limits on VBF produced heavy Higgs				
Mass Point [GeV]	Expected Limits [fb]	Exp+2σ, Exp+1σ	Exp-1σ, Exp-2σ	Observed Limits
ee channel				
300.0 GeV	1502.12	3529.72 , 2279.05	1082.36 , 806.23	2196.04
400.0 GeV	191.88	434.51 , 286.42	138.26 , 102.99	276.83
500.0 GeV	108.40	243.14 , 161.61	78.11 , 58.18	156.12
600.0 GeV	74.50	167.07 , 111.07	53.68 , 39.98	107.32
700.0 GeV	65.12	146.55 , 97.22	46.92 , 34.95	93.87
800.0 GeV	53.57	121.03 , 80.13	38.60 , 28.75	77.39
900.0 GeV	50.28	113.47 , 75.19	36.23 , 26.99	72.63
1000.0 GeV	46.61	104.98 , 69.63	33.58 , 25.02	67.26
1200.0 GeV	45.18	102.25 , 67.61	32.56 , 24.25	65.28
1400.0 GeV	42.38	95.86 , 63.41	30.54 , 22.75	61.23
$\mu\mu$ channel				
300.0 GeV	1676.13	4041.18 , 2568.31	1207.75 , 899.62	2806.53
400.0 GeV	209.70	475.85 , 312.72	151.10 , 112.55	344.27
500.0 GeV	118.48	264.77 , 176.50	85.37 , 63.59	194.34
600.0 GeV	85.46	192.13 , 127.52	61.58 , 45.87	140.47
700.0 GeV	70.58	158.79 , 105.41	50.86 , 37.88	116.25
800.0 GeV	62.14	139.56 , 92.80	44.78 , 33.35	102.34
900.0 GeV	57.15	128.26 , 85.33	41.18 , 30.68	94.02
1000.0 GeV	55.95	125.70 , 83.57	40.32 , 30.03	92.13
1200.0 GeV	53.99	121.47 , 80.70	38.91 , 28.98	89.18
1400.0 GeV	54.50	123.25 , 81.60	39.27 , 29.25	89.87
combined channel				
300.0 GeV	1059.51	2359.44 , 1577.09	763.44 , 568.67	1876.57
400.0 GeV	135.28	289.80 , 198.40	97.48 , 72.61	236.17
500.0 GeV	76.40	163.13 , 111.94	55.05 , 41.01	133.18
600.0 GeV	53.71	115.19 , 78.79	38.70 , 28.83	93.68
700.0 GeV	45.70	98.31 , 67.16	32.93 , 24.53	80.08
800.0 GeV	38.83	83.55 , 57.05	27.98 , 20.84	67.92
900.0 GeV	36.11	77.58 , 53.02	26.02 , 19.38	63.16
1000.0 GeV	34.30	73.66 , 50.35	24.71 , 18.41	59.87
1200.0 GeV	33.17	71.40 , 48.72	23.90 , 17.80	58.07
1400.0 GeV	32.06	69.17 , 47.14	23.10 , 17.21	55.98

Table M.2: The observed and expected 95% CL upper limits on the vector boson fusion production cross-section of an additional narrow-width heavy Higgs boson times the branching fraction of Higgs to ZZ decay $\sigma\times$BR($H \to ZZ$), shown for mass points of 300GeV, 400GeV, 500GeV, 600GeV, 700GeV, 800GeV, 900GeV, 1TeV, 1.2TeV and 1.4TeV, respectively. Limits are derived using ee, $\mu\mu$ and combined ($ee+\mu\mu$) channels, as indicated in the table. **Both the statistical and systematic uncertainties are considered in the limit setting**.

Limits on Randall-Sundrum Graviton				
Mass Point [GeV]	Expected Limits [fb]	Exp+2σ, Exp+1σ	Exp-1σ, Exp-2σ	Observed Limits
ee channel				
600.0	76.65	155.70 , 110.11	55.23 , 41.14	52.00
700.0	52.51	108.75 , 75.97	37.84 , 28.18	30.15
800.0	40.40	85.12 , 58.89	29.11 , 21.69	22.89
900.0	29.26	63.42 , 43.09	21.08 , 15.70	19.54
1000.0	30.57	65.72 , 44.90	22.03 , 16.41	22.08
1200.0	21.11	47.18 , 31.46	15.21 , 11.33	17.37
1400.0	18.69	41.94 , 27.89	13.47 , 10.03	15.85
1600.0	19.92	44.83 , 29.75	14.35 , 10.69	17.08
1800.0	22.65	51.09 , 33.86	16.32 , 12.15	19.49
2000.0	29.71	67.64 , 44.45	21.41 , 15.95	25.61
$\mu\mu$ channel				
600.0	83.30	167.30 , 119.12	60.03 , 44.71	101.07
700.0	58.96	121.22 , 85.04	42.48 , 31.64	82.30
800.0	45.89	96.20 , 66.70	33.07 , 24.63	55.39
900.0	35.94	77.33 , 52.77	25.89 , 19.29	45.02
1000.0	33.63	72.49 , 49.44	24.23 , 18.05	46.56
1200.0	27.03	60.39 , 40.25	19.47 , 14.51	38.46
1400.0	25.59	57.65 , 38.22	18.44 , 13.73	36.19
1600.0	28.35	64.00 , 42.39	20.43 , 15.21	39.96
1800.0	28.53	64.45 , 42.68	20.56 , 15.31	40.21
2000.0	29.60	67.43 , 44.30	21.33 , 15.89	41.63
combined channel				
600.0	55.18	109.31 , 78.52	39.76 , 29.62	46.46
700.0	37.55	75.77 , 53.82	27.06 , 20.16	30.56
800.0	28.84	58.92 , 41.52	20.78 , 15.48	20.49
900.0	21.15	44.33 , 30.76	15.24 , 11.35	17.67
1000.0	21.14	44.04 , 30.64	15.23 , 11.35	20.97
1200.0	15.21	32.72 , 22.36	10.96 , 8.16	16.64
1400.0	13.79	29.79 , 20.31	9.94 , 7.40	15.26
1600.0	14.91	32.24 , 21.95	10.74 , 8.00	16.51
1800.0	16.16	34.97 , 23.80	11.65 , 8.67	18.20
2000.0	19.03	41.17 , 28.02	13.71 , 10.21	22.02

Table M.3: The observed and expected 95% CL upper limits on the production cross-section of a Randall-Sundrum Graviton times the branching fraction of Graviton to ZZ decay $\sigma\times$BR($G* \rightarrow ZZ$), shown for mass points between 600GeV and 2TeV. Limits are derived using ee, $\mu\mu$ and combined ($ee + \mu\mu$) channels, as indicated in the table. **Both the statistical and systematic uncertainties are considered in the limit setting**.

Appendix N

Denominator of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance

Object	Kinematic	Relative resolution	
		Parallel	Perpendicular
Electrons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.7%	0.4%
Photons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.9%	0.4%
Hadronic τ	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	5.5% – 6.7%	1%
Jets	$p_{\mathrm{T}} = 20$ GeV, $\eta = 0$	22%	4.6%–7.1%
	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	7%	1.1%–1.6%
Muons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	2–4%	0.1%
Track Soft Term	TST=100 GeV	8.9%	8.9%

Table N.1: Resolution for the objects entering in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for a representative kinematic value for each object.

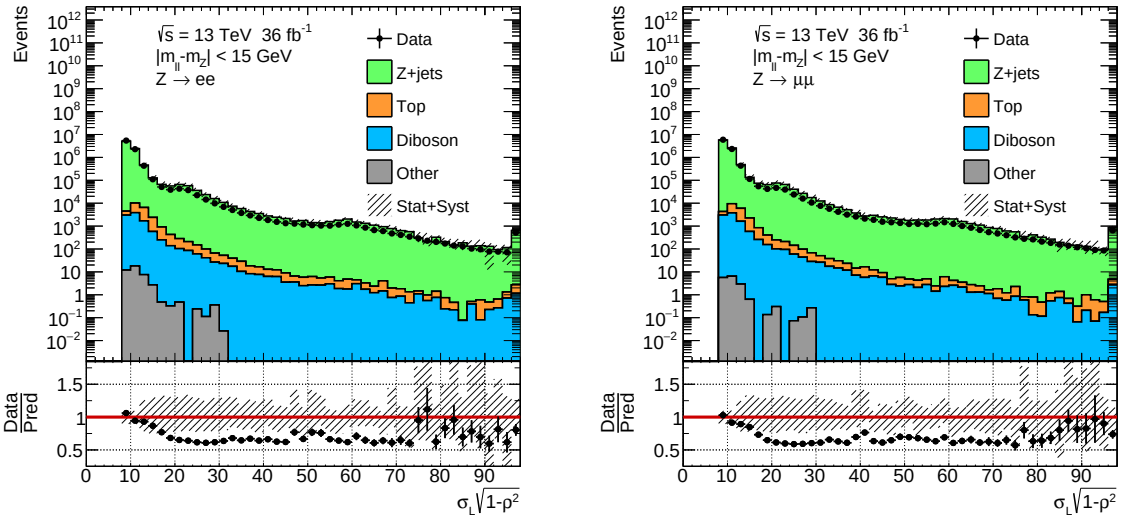


Figure N.1: Denominator of object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, data/MC comparison for an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV.

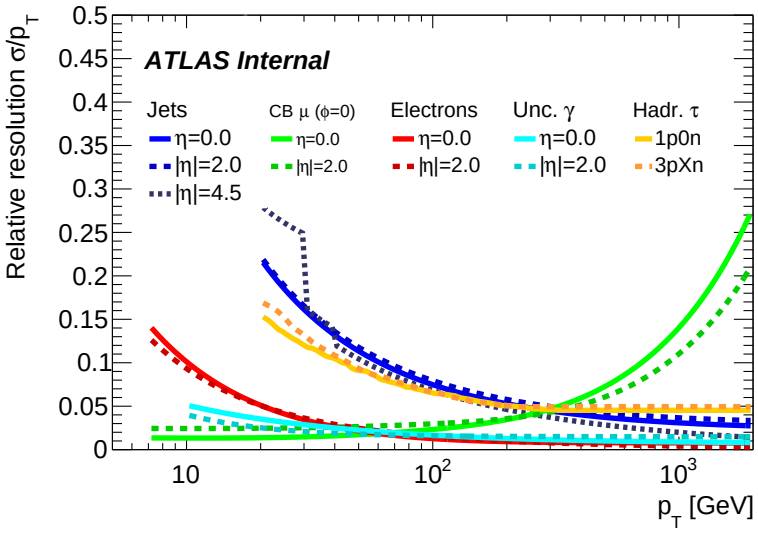


Figure N.2: Resolutions for the objects entering in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The plot shows the resolutions for jets, combined muons (CB μ), electrons, photons not converted to electron-positron pairs (Unc.γ) and τ-leptons hadronically decaying (Hadr.τ) in one charged hadron, zero neutral hadrons and a non detectable τ-neutrino ($1p0n$) or in three charged hadrons, one or more neutral hadrons and a non detectable τ-neutrino ($3pXn$). The curves for jets include the contribution from the additional jets produced by the pile-up, which gives the staggered shape for $|\eta| = 4.5$. Figure produced from $E_{\mathrm{T}}^{\mathrm{miss}}$ significance expert.

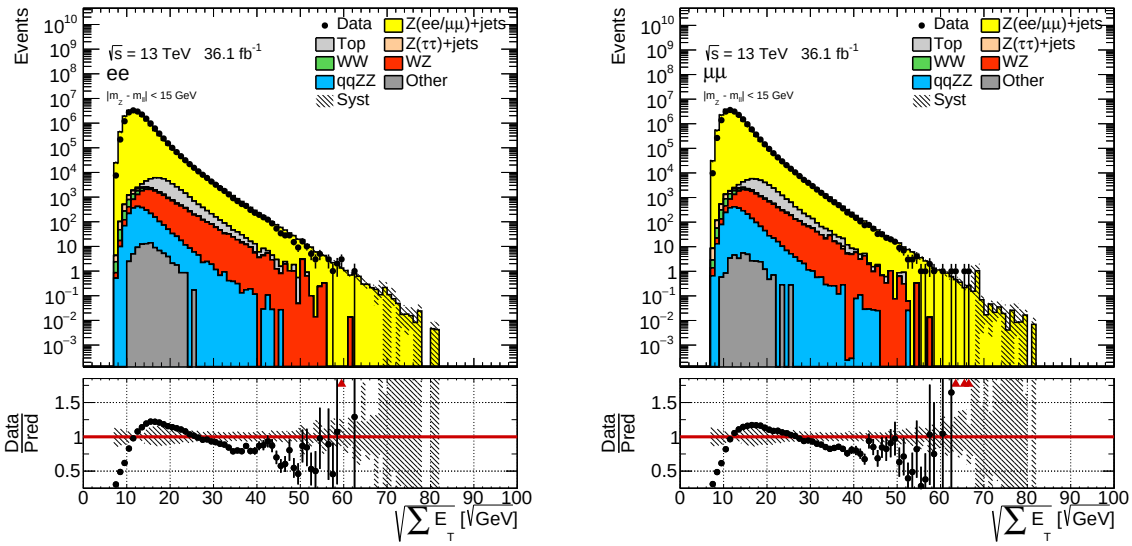


Figure N.3: Denominator of event based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, data/MC comparison for an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV.

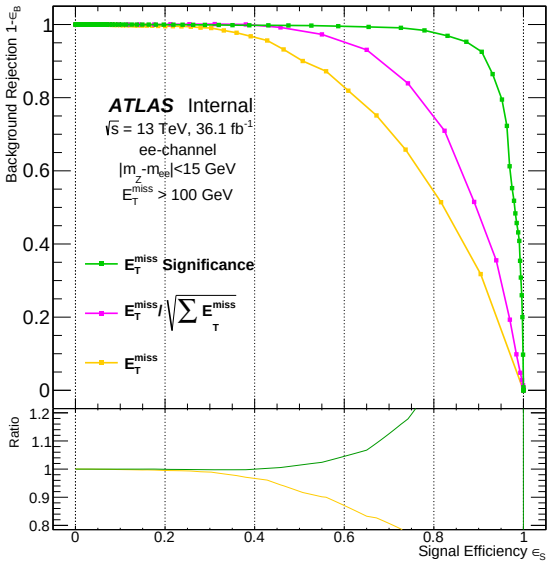


Figure N.4: ROC curve after the Z mass window selection and $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 100 GeV. Figure produced from $E_{\mathrm{T}}^{\mathrm{miss}}$ significance expert.

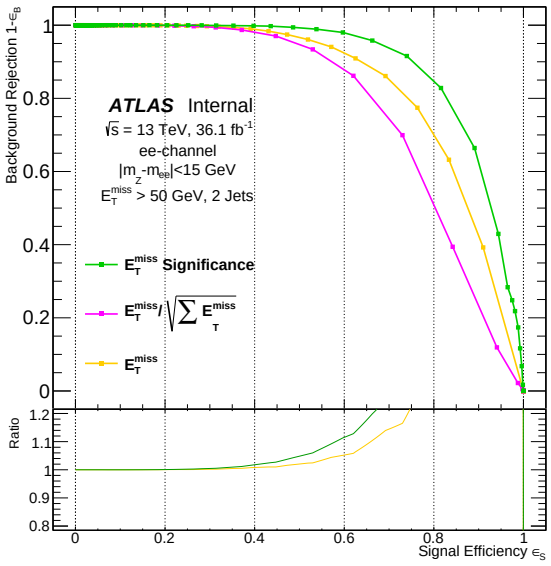


Figure N.5: ROC curve after the Z mass window selection and $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 50 GeV and $N_{\mathrm{jets}} = 2$. Figure produced from $E_{\mathrm{T}}^{\mathrm{miss}}$ significance expert.

Bibliography

[1] M. Petrov, *Reconstruction and performance of Missing Transverse Energy with 3.21/fb of data collected by the ATLAS detector*, , CERN, Geneva, Feb, 2016. https://cds.cern.ch/record/2134946.

[2] M. Petrov, *Public plots: Missing transverse energy performance and systematic uncertainties using the full 2015 dataset*, Feb, 2016. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-003/.

[3] M. Petrov, *Public plots: Missing Transverse Momentum Distribution and Performance in 2016 data*, July, 2016. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-008/.

[4] B. Brunt et al., *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, , CERN, Geneva, Apr, 2016. https://cds.cern.ch/record/2149445.

[5] ATLAS Collaboration, E_T^{miss} *performance in the ATLAS detector using 2015-2016 LHC p-p collisions*, ATLAS-CONF-2018-023, CERN, Geneva, Jun, 2018. https://cds.cern.ch/record/2625233.

[6] ATLAS Collaboration, *Search for heavy ZZ resonances in the* $\ell^+\ell^-\ell^+\ell^-$ *and* $\ell^+\ell^-\nu\bar{\nu}$ *final states using proton proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, arXiv:1712.06386 [hep-ex].

[7] D. Schaefer, D. M. Portillo Quintero, P. Francavilla, C. M. Macdonald, M. Petrov, and E. Tolley, *Object based Missing Transverse Momentum Significance in the ATLAS Detector*, , CERN, Geneva, Dec, 2017. https://cds.cern.ch/record/2294922.

[8] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[9] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].

[10] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at* $\sqrt{s} = 7$ *and* 8 *TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, arXiv:1503.07589 [hep-ex].

[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. C **76** (2016) 6, arXiv:1507.04548 [hep-ex].

[12] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. C **75** (2015) 212, arXiv:1412.8662 [hep-ex].

[13] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B **726** (2013) 120, arXiv:1307.1432 [hep-ex].

[14] ATLAS and CMS Collaborations, *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*, JHEP **08** (2016) 045, arXiv:1606.02266 [hep-ex].

[15] S. Troitsky, *Unsolved problems in particle physics*, Phys. Usp. **55** (2012) 72–95, arXiv:1112.4515 [hep-ph]. [Usp. Fiz. Nauk182,77(2012)].

[16] C. Csáki and P. Tanedo, *Beyond the Standard Model*, in *Proceedings, 2013 European School of High-Energy Physics (ESHEP 2013): Paradfurdo, Hungary, June 5-18, 2013*, pp. 169–268. 2015. arXiv:1602.04228 [hep-ph].

[17] N. Sakai, *Naturalness in Supersymmetric Guts*, Z. Phys. C **11** (1981) 153.

[18] P. Fayet, *Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino*, Nucl. Phys. B **90** (1975) 104–124.

[19] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].

[20] L. Randall and R. Sundrum, *An Alternative to compactification*, Phys. Rev. Lett. **83** (1999) 4690–4693, arXiv:hep-th/9906064 [hep-th].

[21] S. Weinberg, *The Making of the standard model*, Eur. Phys. J. **C34** (2004) 5–13, arXiv:hep-ph/0401010 [hep-ph]. [,99(2005)].

[22] J. D. Lykken, *Beyond the Standard Model*, in *CERN Yellow Report CERN-2010-002, 101-109*. 2010. arXiv:1005.1676 [hep-ph].

[23] T. W. B. Kibble, *The Standard Model of Particle Physics*, 2014. arXiv:1412.4094 [physics.hist-ph].

[24] S. F. Novaes, *Standard model: An Introduction*, in *Particles and fields. Proceedings, 10th Jorge Andre Swieca Summer School, Sao Paulo, Brazil, February 6-12, 1999*, pp. 5–102. 1999. arXiv:hep-ph/0001283 [hep-ph].

[25] D. J. Gross, *Asymptotic freedom and the emergence of QCD*, in *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992*, pp. 199–232. 1992. arXiv:hep-ph/9210207 [hep-ph].

[26] A. Pich, *The Standard model of electroweak interactions*, in *The Standard model of electroweak interactions*, pp. 1–49. 2008. arXiv:0705.4264 [hep-ph]. [,1(2007)].

[27] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics (3rd Edition)*. Addison-Wesley, 3 ed., June, 2001. http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0201657023.

[28] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Soviet Physics Uspekhi **34** (1991) no. 5, 392. http://stacks.iop.org/0038-5670/34/i=5/a=A08.

[29] M. Blanke, *Introduction to Flavour Physics and CP Violation*, in *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*, pp. 71–100. 2017. arXiv:1704.03753 [hep-ph].

[30] A. G. Cohen, A. De Rujula, and S. L. Glashow, *A Matter - antimatter universe?*, Astrophys. J. **495** (1998) 539–549, arXiv:astro-ph/9707087 [astro-ph].

[31] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[32] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[33] G. Guralnik, C. Hagen, and T. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585–587.

[34] A. Purcell, *Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules*, Aug, 2012. BUL-NA-2012-269.

[35] S. G. Karshenboim, *Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants*, Phys. Rept. **422** (2005) 1–63, arXiv:hep-ph/0509010 [hep-ph].

[36] J. Goldstone, *Field theories with Superconductor solutions*, Il Nuovo Cimento (1955-1965) **19** (Jan, 1961) . https://doi.org/10.1007/BF02812722.

[37] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys. Rept. **457** (2008) 1–216, arXiv:hep-ph/0503172 [hep-ph].

[38] P. B. Renton, *Electroweak fits and constraints on the Higgs mass*, in *Proceedings, 32nd International Conference on High Energy Physics (ICHEP 2004): Beijing, China, August 16-22, 2004*, pp. 564–567. 2004. arXiv:hep-ph/0410177 [hep-ph].

[39] S. Dittmaier and M. Schumacher, *The Higgs Boson in the Standard Model - From LEP to LHC: Expectations, Searches, and Discovery of a Candidate*, Prog. Part. Nucl. Phys. **70** (2013) 1–54, arXiv:1211.4828 [hep-ph].

[40] Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** (2016) no. 10, 100001.

[41] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013) 120–144, arXiv:1307.1432 [hep-ex].

[42] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP **12** (2017) 024, arXiv:1708.03299 [hep-ex].

[43] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, arXiv:1712.08891, CERN, Geneva, Dec, 2017.

[44] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph].

[45] LHC Higgs Cross Section Working Group Collaboration, D. e. a. de Florian, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, FERMILAB-FN-1025-T, Oct, 2016. https://cds.cern.ch/record/2227475.

[46] N. Kauer and G. Passarino, *Inadequacy of zero-width approximation for a light Higgs boson signal*, JHEP **08** (2012) 116, arXiv:1206.4803 [hep-ph].

[47] C. F. Uhlemann and N. Kauer, *Narrow-width approximation accuracy*, Nucl. Phys. **B814** (2009) 195–211, arXiv:0807.4112 [hep-ph].

[48] N. Kauer, *Interference effects for $H \to WW/ZZ \to \ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ searches in gluon fusion at the LHC*, JHEP **12** (2013) 082, arXiv:1310.7011 [hep-ph].

[49] C. D. Froggatt and H. B. Nielsen, *Trying to understand the standard model parameters*, Surveys High Energ. Phys. **18** (2003) 55–75, arXiv:hep-ph/0308144 [hep-ph].

[50] D. R. T. Jones, *Comment on "Bare Higgs mass at Planck scale"*, Phys. Rev. **D88** (2013) no. 9, 098301, arXiv:1309.7335 [hep-ph].

[51] I. Brock and T. Schörner-Sadenius, *Physics at the Terascale*. Wiley, 2011. https://books.google.co.uk/books?id=a_a5ZzYzG6sC.

[52] M. Klasen, M. Pohl, and G. Sigl, *Indirect and direct search for dark matter*, Prog. Part. Nucl. Phys. **85** (2015) 1–32, arXiv:1507.03800 [hep-ph].

[53] A. Salam and J. A. Strathdee, *Supersymmetry and Nonabelian Gauges*, Phys. Lett. B **51** (1974) 353–355.

[54] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, Phys. Rept. **459** (2008) 1–241, arXiv:hep-ph/0503173 [hep-ph].

[55] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1–102, arXiv:1106.0034 [hep-ph].

[56] H. E. Haber and D. O'Neil, *Basis-independent methods for the two-Higgs-doublet model. II. The Significance of tanβ*, Phys. Rev. **D74** (2006) 015018, arXiv:hep-ph/0602242 [hep-ph]. [Erratum: Phys. Rev.D74,no.5,059905(2006)].

[57] J. F. Gunion, T. Han, J. Jiang, and A. Sopczak, *Determining tan beta with neutral and charged Higgs bosons at a future e + e- linear collider*, Phys. Lett. **B565** (2003) 42–60, arXiv:hep-ph/0212151 [hep-ph].

[58] V. H. Satheesh Kumar and P. K. Suresh, *Gravitons in Kaluza-Klein theory*, arXiv:gr-qc/0605016 [gr-qc].

[59] A. Oliveira, *Gravity particles from Warped Extra Dimensions, predictions for LHC*, arXiv:1404.0102 [hep-ph].

[60] A. Das and S. SenGupta, *Lightest Kaluza–Klein graviton mode in a back-reacted Randall–Sundrum scenario*, Eur. Phys. J. **C76** (2016) no. 8, 423, arXiv:1506.05613 [hep-th].

[61] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, *Warped Gravitons at the LHC and Beyond*, Phys. Rev. **D76** (2007) 036006, arXiv:hep-ph/0701186 [hep-ph].

[62] H. Okawa, J. Price, J. Vossebeld, L. Xu, Z. Zhao, and B. Zhou, *Search for a Higgs boson in the $H \to ZZ \to \ell\ell\nu\nu$ Channel*, , CERN, Geneva, Apr, 2014. https://cds.cern.ch/record/1693161.

[63] ATLAS Collaboration, *Search for an additional, heavy Higgs boson in the $H \to ZZ$ decay channel at $\sqrt{s} = 8$ TeV in pp collision data with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 45, arXiv:1507.05930 [hep-ex].

[64] J. Bernon, J. F. Gunion, H. E. Haber, Y. Jiang, and S. Kraml, *Scrutinizing the alignment limit in two-Higgs-doublet models: $m_h = 125$ GeV*, Phys. Rev. D **92** (Oct, 2015) 075004.

[65] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. https://cds.cern.ch/record/782076.

[66] F. Kohn, *Measurement of the charge asymmetry in top quark pair production in pp collision data at $\sqrt{s} = 7$ TeV using the ATLAS detector*. PhD thesis, U. Gottingen (main), 2012. arXiv:1204.0952 [hep-ex].

[67] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[68] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[69] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar, 2008. https://cds.cern.ch/record/1095924.

[70] M. Aleksa, F. Bergsma, P. A. Giudici, A. Kehrli, M. Losasso, X. Pons, H. Sandaker, P. S. Miyagawa, S. W. Snow, J. C. Hart, and L. Chevalier, *Measurement of the ATLAS solenoid magnetic field*, Journal of Instrumentation **3** (2008) no. 04, P04003. http://stacks.iop.org/1748-0221/3/i=04/a=P04003.

[71] F. Hartmann, *Silicon tracking detectors in high-energy physics*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **666** (2012) no. Supplement C, 25 – 46. Advanced Instrumentation.

[72] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010. https://cds.cern.ch/record/1291633.

[73] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, PoS **EPS-HEP2015** (2015) 261, arXiv:1608.07850 [physics.ins-det].

[74] ATLAS Collaboration, E. Hines, *Performance of Particle Identification with the ATLAS Transition Radiation Tracker*, in *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011, Providence, USA, August 9-13, 2011*. 2011. arXiv:1109.5925 [physics.ins-det].

[75] ATLAS Collaboration, *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*, in *Proceedings, 3rd International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA 2013): Marseille, France, June 23-27, 2013*. 2013. arXiv:1306.6756 [physics.ins-det]. https://inspirehep.net/record/1240499/files/arXiv:1306.6756.pdf.

[76] C. W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. **75** (Oct, 2003) 1243–1286.

[77] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, Mar, 2008. https://cds.cern.ch/record/1095927.

[78] ATLAS LAr Collaboration, Z. Meng, *Performance of the ATLAS liquid argon calorimeter*, in *Physics at the LHC2010. Proceedings, 5th Conference, PLHC2010, Hamburg, Germany, June 7-12, 2010*, pp. 406–408. 2010. https://inspirehep.net/record/891081/files/meng.pdf.

[79] B. Sotto-Maior Peralva, *Calibration and Performance of the ATLAS Tile Calorimeter*, in *Proceedings, International School on High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013): Rio de Janeiro, Brazil, March 17-24, 2013*. 2013. arXiv:1305.0550 [physics.ins-det].

[80] J. E. Mdhluli, B. Mellado, and E. Sideras-Haddad, *Neutron irradiation and damage assessment of plastic scintillators of the Tile Calorimeter*, Journal of Physics: Conference Series **802** (2017) no. 1, 012008. http://stacks.iop.org/1742-6596/802/i=1/a=012008.

[81] F. Carrió, V. Castillo, A. Ferrer, L. Fiorini, Y. Hernández, E. Higón, B. Mellado, L. March, P. Moreno, R. Reed, C. Solans, A. Valero, and J. A. Valls, *The sROD module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator*, Journal of Instrumentation **9** (2014) no. 02, C02019. http://stacks.iop.org/1748-0221/9/i=02/a=C02019.

[82] A. Artamonov et al., *The ATLAS Forward Calorimeter*, Journal of Instrumentation **3** (2008) no. 02, P02010. http://stacks.iop.org/1748-0221/3/i=02/a=P02010.

[83] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331068.

[84] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, Mar, 2008. https://cds.cern.ch/record/1095929.

[85] *Monitored Drift Tubes Chambers for Muon Spectroscopy in ATLAS*, ATL-MUON-94-044. ATL-M-PN-44, CERN, Geneva, Jun, 1994. https://cds.cern.ch/record/685844.

[86] T. Argyropoulos et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE Trans. Nucl. Sci. **56** (2009) 1568–1574.

[87] G. Cattani and the RPC group, *The Resistive Plate Chambers of the ATLAS experiment: performance studies*, Journal of Physics: Conference Series **280** (2011) no. 1, 012001. http://stacks.iop.org/1742-6596/280/i=1/a=012001.

[88] K. Nagai, *Thin gap chambers in ATLAS*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **384** (1996) no. 1, 219 – 221. http://www.sciencedirect.com/science/article/pii/S0168900296010650. BEAUTY '96.

[89] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018. ATLAS-TDR-023, Sep, 2013. https://cds.cern.ch/record/1602235. Final version presented to December 2013 LHCC.

[90] *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, CERN, Geneva, Mar, 2016. https://cds.cern.ch/record/2136007.

[91] A. Sidoti, *Minimum Bias Trigger Scintillators in ATLAS Run II*, Journal of Instrumentation **9** (2014) no. 10, C10020. http://stacks.iop.org/1748-0221/9/i=10/a=C10020.

[92] R. E. Kwee, *Minimum Bias Trigger in ATLAS*, ATL-DAQ-PROC-2010-055, CERN, Geneva, Dec, 2010. https://cds.cern.ch/record/1312194.

[93] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. https://cds.cern.ch/record/1099735.

[94] Y. Nakahama, C. Leonidopoulos, and A. Ruiz-Martinez, *Trigger Menu Table in 2017 for* 1.7×10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, , CERN, Geneva, Oct, 2017. https://cds.cern.ch/record/2288213. Aiming to make this table public for RRB in 2017 mid-October and future conferences, followed by the menu pub note after the 2017 data taking finishes.

[95] A. Ruiz-Martinez, F. Monticelli, J. Hoya, and S. D. Jones, *Electron and photon trigger efficiency plots using the full 2016 dataset*, , CERN, Geneva, Mar, 2017. https://cds.cern.ch/record/2254973.

[96] L. Dell'Asta and S. M. Shaw, *Muon trigger public plots with full 2016 dataset*, , CERN, Geneva, Jan, 2017. https://cds.cern.ch/record/2242824.

[97] W. Herr and B. Muratori, *Concept of luminosity*, . https://cds.cern.ch/record/941318.

[98] S. van der Meer, *Calibration of the effective beam height in the ISR*, CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. https://cds.cern.ch/record/296752.

[99] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** (2016) no. 12, 653, arXiv:1608.03953 [hep-ex].

[100] ATLAS Collaboration, *Luminosity and Pile-up ATLAS Public plots in Run II*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[101] ATLAS Collaboration, F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, ATL-PHYS-PROC-2016-163. 12, CERN, Geneva, Oct, 2016. https://cds.cern.ch/record/2222390.

[102] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. **C76** (2016) no. 11, 581, arXiv:1510.03823 [hep-ex].

[103] A. Buckley et al., *Implementation of the ATLAS Run 2 event data model*, Journal of Physics: Conference Series **664** (2015) no. 7, 072045. http://stacks.iop.org/1742-6596/664/i=7/a=072045.

[104] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[105] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[106] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145–233, arXiv:1101.2599 [hep-ph].

[107] C. Oleari, *The POWHEG-BOX*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 36–41, arXiv:1007.3893 [hep-ph].

[108] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[109] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820 [hep-ph].

[110] *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1474107.

[111] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[112] *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037683.

[113] T. Cornelissen et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. https://cds.cern.ch/record/1020106.

[114] ATLAS Tracking CP group, *Impact parameters*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackingDC14#Impact_parameters_z0_d0_definiti.

[115] ATLAS Collaboration, S. Marti-Garcia, *Alignment of the ATLAS Inner Detector*, . https://cds.cern.ch/record/2238429.

[116] ATLAS Collaboration, *ATLAS Inner Detector Alignment Performance with February 2015 Cosmic Ray Data*, ATL-PHYS-PUB-2015-009, 2015. https://cds.cern.ch/record/2008724.

[117] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016. https://cds.cern.ch/record/2157687.

[118] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[119] ATLAS Tracking CP group, *Track-to-vertex association*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TrackingCPMoriond2017#Track_to_Vertex_Association.

[120] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. **C76** (2016) no. 5, 292, arXiv:1603.05598 [hep-ex].

[121] ATLAS Muon CP group, *Track-to-vertex association*, . https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesMC15#Muon_track_to_vertex_association.

[122] ATLAS Collaboration, *Measurement of the Higgs boson mass in the* $H \to ZZ^* \to 4\ell$ *and* $H \to \gamma\gamma$ *channels with* $\sqrt{s}$*=13TeV pp collisions using the ATLAS detector*, ATLAS-CONF-2017-046, CERN, Geneva, Jul, 2017. https://cds.cern.ch/record/2273853.

[123] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[124] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[125] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[126] 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. **D96** (2017) no. 7, 072002, arXiv:1703.09665 [hep-ex].

[127] ATLAS Jet CP group, *Jet uncertainties*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertainties20152016Data20p7.

[128] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ *TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) no. 3, 2306, arXiv:1210.6210 [hep-ex].

[129] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. https://cds.cern.ch/record/1700870.

[130] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv:0802.1188 [hep-ph].

[131] ATLAS Collaboration, *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run*, JINST **8** (2013) P07004, arXiv:1303.0223 [hep-ex].

[132] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037702.

[133] ATLAS Collaboration, G. Aad et al., *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) no. 04, P04008, arXiv:1512.01094 [hep-ex].

[134] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2018-006, 2018. https://cds.cern.ch/record/2314418.

[135] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in* $t\bar{t}$ *Events Using a Likelihood Approach with the ATLAS Detector*, 2018. https://cds.cern.ch/record/2306649.

[136] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using* $t\bar{t}$ *events at* $\sqrt{s} = 13$ *TeV*, arXiv:1805.01845.

[137] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2160731.

[138] D. Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, . https://cds.cern.ch/record/1700874.

[139] ATLAS Collaboration, A. Policicchio, *BSM searches in ATLAS*, ATL-PHYS-PROC-2015-010, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2001836.

[140] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, , CERN.

[141] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV*, ATL-PHYS-PUB-2015-026, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037717.

[142] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. **C77** (2017) no. 7, 466, arXiv:1703.10485 [hep-ex].

[143] ATLAS Collaboration, *Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in 8 TeV pp collisions with the ATLAS detector*, Eur. Phys. J. **C75** (2015) no. 7, 318, arXiv:1503.03290 [hep-ex]. [Erratum: Eur. Phys. J.C75,no.10,463(2015)].

[144] CMS Collaboration, *Search for new physics in events with opposite-sign leptons, jets, and missing transverse energy in pp collisions at 7 TeV*, Phys. Lett. **B718** (2013) 815–840, arXiv:1206.3949 [hep-ex].

[145] A. Gupta, R. Primulando, and P. Saraswat, *A New Probe of Dark Sector Dynamics at the LHC*, JHEP **09** (2015) 079, arXiv:1504.01385 [hep-ph].

[146] ATLAS GRL group, *Good Runs List for 2015 data*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DataMCForAnalysis2015.

[147] ATLAS GRL group, *Good Runs List for 2016 data*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DataMCForAnalysis2016.

[148] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, , CERN, Geneva, May, 2015. https://cds.cern.ch/record/2016323.

[149] *Study of ll+MET final state in HZZ group with proton-proton collisions at $\sqrt{s}$ = 13 TeV: objects definition and background estimation*, , CERN, Geneva, Oct, 2015. https://cds.cern.ch/record/2058246. Authors: the ll+MET HZZ group.

[150] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

[151] P. Nason and G. Zanderighi, *W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2*, Eur. Phys. J. **C74** (2014) no. 1, 2702, arXiv:1311.1365 [hep-ph].

[152] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[153] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[154] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv:1406.3660 [hep-ex].

[155] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[156] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, Nucl. Instrum. Meth. **A771** (2015) 39–48, arXiv:1410.7388 [physics.data-an].

[157] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[158] S. Carrazza, S. Forte, and J. Rojo, *Parton Distributions and Event Generators*, in *Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*, pp. 89–96. 2013. arXiv:1311.5887 [hep-ph].

[159] R. D. Ball et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, Nucl. Phys. **B838** (2010) 136–206, arXiv:1002.4407 [hep-ph].

[160] *ATLAS Run 1 Pythia8 tunes*, ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1966419.

[161] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1300517.

[162] F. Cascioli et al., *ZZ production at hadron colliders in NNLO QCD*, Phys. Lett. **B735** (2014) 311–313, arXiv:1405.2219 [hep-ph].

[163] M. Grazzini, S. Kallweit, and D. Rathlev, *ZZ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, Phys. Lett. **B750** (2015) 407–410, arXiv:1507.06257 [hep-ph].

[164] N. Kauer, C. O'Brien, and E. Vryonidou, *Interference effects for $H \to W\,W \to \ell\nu q\overline{q}'$ and $H \to ZZ \to \ell\overline{\ell}q\overline{q}$ searches in gluon fusion at the LHC*, JHEP **10** (2015) 074, arXiv:1506.01694 [hep-ph].

[165] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, *QCD corrections to ZZ production in gluon fusion at the LHC*, Phys. Rev. **D92** (2015) no. 9, 094028, arXiv:1509.06734 [hep-ph].

[166] J. M. Campbell, R. K. Ellis, M. Czakon, and S. Kirchner, *Two loop correction to interference in $gg \to ZZ$*, JHEP **08** (2016) 011, arXiv:1605.01380 [hep-ph].

[167] S. Alioli, F. Caola, G. Luisoni, and R. Röntsch, *ZZ production in gluon fusion at NLO matched to parton-shower*, Phys. Rev. **D95** (2017) no. 3, 034042, arXiv:1609.09719 [hep-ph].

[168] K. Melnikov and M. Dowling, *Production of two Z-bosons in gluon fusion in the heavy top quark approximation*, Phys. Lett. **B744** (2015) 43–47, arXiv:1503.01274 [hep-ph].

[169] C. S. Li, H. T. Li, D. Y. Shao, and J. Wang, *Soft gluon resummation in the signal-background interference process of $gg(\to H^*) \to ZZ$*, JHEP **08** (2015) 065, arXiv:1504.02388 [hep-ph].

[170] ATLAS Collaboration, *Simulation of top quark production for the ATLAS experiment at sqrt(s) = 13 TeV*, ATL-PHYS-PUB-2016-004, CERN, Geneva, Jan, 2016. https://cds.cern.ch/record/2120417.

[171] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[172] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[173] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155.

[174] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[175] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput. Phys. Commun. **182** (2011) 2388–2403, arXiv:1011.3540 [hep-ph].

[176] HZZ Group, *Theoretical Uncertainties Occurring In $H \to ZZ$ analyses*, . https://cds.cern.ch/record/2232040.

[177] W. Buttinger et al., *Measurement of the ZZ production cross section in pp collisions at 8 TeV using ZZ->2l2nu and ZZ->4l decay channels with the ATLAS detector*, . https://cds.cern.ch/record/1976407.

[178] ATLAS Collaboration, *Measurement of the ZZ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV using the $ZZ \to \ell^-\ell^+\ell'^-\ell'^+$ and $ZZ \to \ell^-\ell^+\nu\bar{\nu}$ channels with the ATLAS detector*, JHEP **01** (2017) 099, arXiv:1610.07585 [hep-ex].

[179] M. Grazzini, S. Kallweit, S. Pozzorini, D. Rathlev, and M. Wiesemann, *W^+W^- production at the LHC: fiducial cross sections and distributions in NNLO QCD*, Journal of High Energy Physics **2016** (2016) no. 8, 140.

[180] M. Grazzini, S. Kallweit, and D. Rathlev, *Status of diboson production in NNLO QCD*, in *Proceedings, 12th International Symposium on Radiative Corrections (Radcor 2015) and LoopFest XIV (Radiative Corrections for the LHC and Future Colliders): Los Angeles, CA, USA, June 15-19, 2015*. 2016. arXiv:1601.06036 [hep-ph].

[181] B. Biedermann, A. Denner, S. Dittmaier, L. Hofer, and B. Jäger, *Electroweak Corrections to $pp \to \mu^+\mu^-e^+e^- + X$ at the LHC: A Higgs Boson Background Study*, Phys. Rev. Lett. **116** (Apr, 2016) 161803.

[182] CMS Collaboration, A. M. Sirunyan et al., *Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. **B774** (2017) 682–705, arXiv:1708.02812 [hep-ex].

[183] J. Iturbe, J. Crane, A. Johnson, S. Lammers, A. Pilkington, D. Price, T. Wyatt, *OPEN PRESENTATION: Measurement of the cross section for electroweak*

production of di-jets in association with a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector, . https://indico.cern.ch/event/630683/contributions/2548473/attachments/1444710/2225355/JuliaIturbe_VBFZjj_OpenPresentation.pdf.

[184] The ATLAS Collaboration, *Measurements of electroweak Wjj production and constraints on anomalous gauge couplings with the ATLAS detector*, arXiv:1703.04362 [hep-ph].

[185] *Measurement of W and Z Boson Production Cross Sections in pp Collisions at root s = 13 TeV in the ATLAS Detector*, ATLAS-CONF-2015-039, CERN, Geneva, Aug, 2015. https://cds.cern.ch/record/2045487.

[186] ATLAS Egamma Combined Performance Group, *Electron recommendations*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LatestRecommendationsElectronIDRun2.

[187] ATLAS Muon Combined Performance Group, *Muon recommendations*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15#MCP_recommendations.

[188] ATLAS JetEtMiss Performance Group, *Usage of Jets in analyses: rebuilding and systematics*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtmissRecommendations2016.

[189] ATLAS JetEtMiss Performance Group, *Usage of Etmiss in analyses: rebuilding and systematics*, . https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EtmissRecommendationsRel20p7.

[190] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037700.

[191] ATLAS Luminosity group, *Luminosity uncertainty recommendation*, . https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics#2016_13_TeV_proton_proton_Morion.

[192] W. Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, , CERN, Geneva, May, 2015. https://cds.cern.ch/record/2014726.

[193] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. **C76** (2016) no. 4, 196, arXiv:1512.01178 [hep-ph].

[194] ATLAS Higgs Group, *TRexFitter (TtHFitter) framework*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TtHFitter.

[195] ROOT Collaboration Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with*

RooFit and RooStats, CERN-OPEN-2012-016, New York U., New York, Jan, 2012. https://cds.cern.ch/record/1456844.

[196] *Analysis of HWW ggF and VBF production modes with 20 fb-1 and 5fb-1 of data collected with the ATLAS detector at 7 and 8 TeV*, , CERN, Geneva, May, 2014. https://cds.cern.ch/record/1702035.

[197] G. Cowan, *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. https://books.google.co.uk/books?id=ff8ZyW0nlJAC.

[198] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693–2704.

[199] ATLAS 4 lepton Group, *Four lepton framework*, . https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HiggsZZRunIIWorkspaces.

[200] G. Artoni et al., *Search for heavy Higgs bosons in the four-lepton decay channel with the ATLAS detector using 35 /fb of 13 TeV data*, , CERN, Geneva, Nov, 2016. https://cds.cern.ch/record/2231587.

[201] S. H. Abidi et al., *Event selection and background estimation in the $H \to ZZ^{(*)} \to 4\ell$ channel at $\sqrt{s} = 13\,TeV$ - Moriond 2017 Analysis*, , CERN, Geneva, Nov, 2016. https://cds.cern.ch/record/2231576.

[202] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198–207, arXiv:hep-ex/0011057 [hep-ex].

[203] ATLAS Collaboration, G. Aad et al., *Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D86** (2012) 032003, arXiv:1207.0319 [hep-ex].

[204] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, *Procedure for the LHC Higgs boson search combination in Summer 2011*, CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011. https://cds.cern.ch/record/1379837.

[205] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116 [physics]. [,186(2003)].

[206] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003 [physics.data-an].

[207] E. Tolley, D. Schaefer, T. J. Khoo, M. H. Klein, M. Petrov, M. Testa, F.-y. Tsai, M. Valente, C. J. Young, and M. Scornajenghi, *E_T^{miss} Performance in the ATLAS detector using 2016 LHC p-p Collisions*, ATL-COM-PHYS-2017-1732, CERN, Geneva, Dec, 2017. https://cds.cern.ch/record/2294891.

[208] CMS Collaboration, *Missing transverse energy performance of the CMS detector*, JINST **6** (2011) P09001, arXiv:1106.5048 [physics.ins-det].

[209] R. Garisto, *Editorial: Theorists React to the CERN 750 GeV Diphoton Data*, Phys. Rev. Lett. **116** (Apr, 2016) 150001. https://link.aps.org/doi/10.1103/PhysRevLett.116.150001.