

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

Cross-section measurement of Higgs Boson production in association with a Z boson,

decaying via WW^* , using the ATLAS detector

at the LHC at $\sqrt{s} = 13.6$ TeV

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

Mayuri Prabhakar Kawale

Norman, Oklahoma

2025

Cross-section measurement of Higgs Boson production in association with a Z boson,

decaying via WW^* , using the ATLAS detector

at the LHC at $\sqrt{s} = 13.6$ TeV

A DISSERTATION APPROVED FOR THE
HOMER L. DODGE DEPARTMENT OF PHYSICS AND
ASTRONOMY

BY THE COMMITTEE CONSISTING OF

Dr. Michael G. Strauss, Chair

Dr. Braden Abbott

Dr. Howard Baer

Dr. Alberto Marino Valle

Dr. Keith Strevett

Acknowledgements

Thank you would be a gross understatement to express my gratitude to all the people because of whom I was able to pursue and complete my Ph.D. journey. I would like to begin with the first person, Dr. Sile Nic Chormaic, who believed in me and gave me the opportunity to work in her group at OIST, Japan, for two summers during my undergraduate studies. It opened the doors to a new world of research and exploration for me. I am extremely thankful to her.

I wasn't meant to do a Ph.D. in experimental high-energy physics; I had always worked on experimental optics. But as I started exploring options, I met Dr. Michael Strauss, and I knew he would be an excellent mentor, and he has been just that. He had faith in me more than I had in myself, and that kept me pushing forward even when I doubted myself the most. I am extremely grateful to him for his continuous support, guidance, and encouragement throughout my Ph.D. journey. His expertise and insights have been invaluable in my growth as a researcher. I will cherish all the discussions we had, both academic and non-academic.

I would also like to thank my committee members: Dr. Braden Abbott, for your support and guidance throughout my Ph.D. journey; Dr. Howard Baer, I have always admired your dedication to teaching and research, and I am grateful for the opportunity to learn from you; Dr. Alberto Marino Valle, I had joined your research group first before switching to high-energy physics, and I am thankful for the guidance and support you provided during that time as well as being part of my committee; Dr. Keith Strevett, thank you for accepting to be part of my committee in the later stage of my Ph.D. and for your support. I would like to thank the Department of Energy (DOE) for their generous funding which made this research possible. I am grateful to the Department of Physics and Astronomy at the University of Oklahoma for providing a supportive research environment and resources necessary for my work.

This work would not have been possible without the support and collaboration of my colleagues in the ATLAS experiment. I would like to thank all the Run 3 VH group members for their help and valuable discussions. Kyle and Callum, thank you for our weekly informal discussions to study various aspects of the analysis. I couldn't have asked for better colleagues to work with, who made this journey enjoyable. Kyle has worked with me closely on various aspects of the analysis; in particular, the fake estimation studies presented in this thesis are mainly based on his contributions. Callum, you have been my go-to person for all technical aspects of the analysis, and I have learned a lot from you.

Matt H., thank you for developing the neural network framework used in this analysis. Robin and Matt, thank you for all your help in the overall analysis framework. Matt, your help in setting up the statistical analysis framework was invaluable. Michiel, thank you for guiding me through the initial stages of the analysis and helping me understand the physics behind it. I am grateful to Ahmed Markhoos for his help in understanding the machine learning techniques used in this analysis.

Finally, I would like to thank my family and friends for their unwavering support and encouragement throughout this journey. Veena, you have been a wonderful roommate and friend throughout these years. I have enjoyed our numerous discussions on anything and everything, and I will cherish our early morning train rides and hikes in Switzerland. Bijay, thank you for being a great friend and always being there to listen and advise me on numerous occasions. Going on my first hike in Oklahoma and making momos for the first time with you are memories I will always cherish. To my siblings, Madhuri, Rahul, and Rohit, you are my pillars of strength. I have overcome a lot of personal challenges during this journey because of your constant support and belief in me. To my friend and husband, Umang, you have been my rock through it all. Your love, patience, and understanding have kept me going. I am truly blessed to have you in my life, and I am grateful for your presence. I am glad to have taken my entire educational journey, from bachelor's to Ph.D., with you by my side. Lastly, the most important people without whom I wouldn't be here today, my parents. Aai (Jyoti) and Pappa (Prabhakar), your shared passion for educating your children has led me to this point. I will be forever grateful for all your sacrifices.

*To the remarkable women in my life,
Prabhavati (grandmother), Jyoti (mother) and Madhuri (sister),
who inspired me to dream more than I was allowed to.*

*To my nieces, Siya and Veda,
may you always be curious, kind and fearless.*

Abstract

The Higgs boson discovery in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider at CERN made the Standard Model (SM) predictions complete. However, the Standard Model also comes with its own flaws, it does not explain phenomenon like matter-antimatter asymmetry and existence of dark matter and dark energy in the universe. Studying the properties of existing Standard Model particles helps us to better understand their limitations. Higgs boson is responsible in providing masses to other fundamental particles via Higgs Mechanism and Yukawa couplings. Studying the various production and decay modes of Higgs boson will help us understand the interactions of these fundamental particles precisely.

In this thesis, Higgs boson production in association with a Z boson and its decay via WW^* channel into a fully leptonic final state is discussed. The data collected during Run 3 of the ATLAS detector from 2022 and 2023 is used. Since this signal process is limited by statistics, advanced machine learning algorithms, in particular, deep neural networks have been deployed to discriminate between signal and any background processes affecting this analysis. Results obtained from statistical treatment using the maximum likelihood estimator show potential discovery when combining the results with the VH analysis in Run 2+3. Prospects are also discussed which will help achieve the discovery potential and make cross-section measurements in the transverse momentum bins of Z boson.

Keywords: Higgs boson, ATLAS experiment at CERN, Deep Neural Networks, Statistical Analysis

Contents

Abstract	vii
List of Figures	x
List of Tables	xv
List of Acronyms	xix
Introduction	1
1 The Standard Model of Particle Physics	4
1.1 Groups and Gauge Theories	6
1.2 Quantum Electrodynamics	7
1.3 Quantum Chromodynamics	8
1.4 Weak Interactions and Electroweak Theory	9
1.5 Spontaneous Symmetry Breaking	11
1.6 The Higgs Mechanism	12
1.7 The Higgs Boson Physics	13
1.7.1 Production Modes	13
1.7.2 Decay Modes	14
2 The ATLAS Detector at LHC	16
2.1 The Large Hadron Collider	16
2.1.1 Proton injection chain	17
2.1.2 Luminosity	18
2.2 The ATLAS detector	21
2.2.1 The Coordinate System	22
2.2.2 The Magnet System	23
2.2.3 The Inner Detector	23
2.2.4 Calorimeters	27
2.2.5 Muon Spectrometer	30
2.2.6 Forward Detectors	33
2.2.7 Trigger and Data Acquisition System	34
3 Object Reconstruction and Selection	37
3.1 Tracks	37
3.2 Primary Vertex	37
3.3 Electrons	39
3.3.1 Identification	39
3.3.2 Isolation	40

3.4	Muons	40
3.4.1	Identification	41
3.4.2	Isolation	41
3.5	Jets	42
3.5.1	Calibration	43
3.5.2	Jet Vertex Tagging	43
3.5.3	Flavor Tagging	44
3.6	Missing Transverse Momentum	44
3.7	Overlap Removal	45
3.8	Object Selection	45
4	Data and Monte Carlo Samples	47
4.1	Data	47
4.2	Monte Carlo Simulations	48
4.2.1	Signal Samples	50
4.2.2	Background Samples	50
5	Data-Driven Estimation of Fakes	52
5.1	Fake Estimation Control Regions	54
5.1.1	$3l + 1X$ Control Region	54
5.1.2	$2l + 2X$ Control Region	56
5.2	Determining the Fake Factors	58
5.2.1	$2l + 1Y$ Control Region	58
5.3	Systematic Uncertainties from Fake Factors	61
5.3.1	Prompt Lepton Subtraction Uncertainty	61
5.3.2	MC Non-closure Uncertainty	62
6	Neural Networks	64
6.1	Overview of Neural Networks	64
6.2	Hyperparameter Optimization	68
6.3	Model Validation	68
6.4	Feature Importance	69
7	Analysis Strategy	70
7.1	Analysis Overview	70
7.2	Background Contribution	72
7.3	Event Selection	73
7.4	Neural Network Training	76
7.4.1	Input Variables	76
7.4.2	Hyperparameter Tuning	77
7.4.3	Architecture	82
7.4.4	Model Validation	82
7.4.5	Feature Ranking	86
7.4.6	Output	87

7.5	Control and Validation Regions	87
7.5.1	ZZ Control Region	88
7.5.2	ZZ Validation Region	94
7.5.3	$t\bar{t}Z$ Validation Region	98
8	Systematic Uncertainties	102
8.1	Experimental Systematic Uncertainties	102
8.2	Theoretical Uncertainties	105
9	Statistical Analysis	107
9.1	Basic Definitions	107
9.1.1	Hypothesis Testing	107
9.1.2	Parameter of Interest	107
9.1.3	Nuisance Parameters	108
9.1.4	Test Statistic	108
9.1.5	Likelihood Function	108
9.1.6	Confidence Intervals	108
9.1.7	Significance and p-value	109
9.2	Maximum Likelihood Estimation	109
9.3	Test Statistics and p-values	111
10	Expected Results	113
10.1	ZH 4ℓ channel	113
10.2	Combined VH in Run 3	115
10.3	Combined VH in Run 2 and Run 3	115
11	Conclusion	120
A	Bibliography	122

List of Figures

1.1	Tabular representation of particles of the Standard Model along with their properties such as mass, charge, and spin [1].	5
1.2	Schematic representation of parity and handedness: (a) Handedness [2] (b) Parity is conserved (c) Parity is violated [3].	10
1.3	Schematic representation of the Higgs potential for $\mu^2 > 0$ showing the “Mexican hat” shape [4].	13
1.4	Leading-order Feynman diagrams for the main Higgs boson production modes at the LHC: (a) gluon-gluon fusion (ggF), (b) vector boson fusion (VBF), (c) Higgs Strahlung (ZH/WH), and (d) $t\bar{t}H/b\bar{b}H$ production [5]. .	14
1.5	(a) Higgs boson production cross-section as a function of center-of-mass energy ($\sqrt{s}$) at the LHC for a Higgs mass of 125 GeV. (b) Branching ratios of various Higgs boson decay channels as a function of Higgs mass [6]. . . .	15
2.1	The LHC Machine and it’s four experiments displayed with their geographical locations.	16
2.2	Schematic layout of the LHC interaction regions.	17
2.3	Schematic layout of the CERN accelerator complex. The injector chain for accelerating Protons is LINAC, the PS Booster (PSB), the Proton Synchrotron(PS) and the Super Proton Synchrotron (SPS) [7].	18
2.4	Luminosity-weighted, mean number of interactions per crossing collected during 2022-2024 Run 3 data taking at $\sqrt{s}=13.6$ TeV is shown [8].	19
2.5	Total integrated luminosity for the time period of 2022-2024 is shown. Luminosity delivered by the LHC is shown in green, recorded by ATLAS is shown in yellow and the blue color represents the luminosity that is good to use for physics analyses [9].	20
2.6	Schematic of the ATLAS detector subsystems for the Run 3 configuration [10].	21
2.7	Illustration of the magnetic field lines of the ATLAS detector. The green lines correspond to the magnetic field by the central solenoid and the blue circles correspond to the magnetic field by the toroidal magnets [11].	24
2.8	Layout of the ID and it’s sub-detectors is shown. Each of the sub-detector has a barrel component and two endcap components [12].	25
2.9	Transverse view of the ATLAS Inner Detector and illustration of track reconstruction by the sub-detectors [13].	26
2.10	Illustration of the cut-away view of the ATLAS calorimeter system. The EM calorimeter uses LAr calorimeter in the barrel, endcap and forward region. The hadronic calorimeter uses the tile calorimeter in the barrel region and LAr everywhere else [12].	28

2.11	(a) Schematic of the longitudinal view of the ATLAS EM calorimeter showing the accordion geometry and the three layers in depth. (b) Schematic of the tile calorimeter showing the alternating layers of steel plates and scintillating tiles, the optical fibers and the photomultiplier tubes (PMTs) [12].	29
2.12	Schematic of the ATLAS muon spectrometer showing the arrangement of the different detector technologies [12].	31
2.13	Schematic of the New Small Wheel (NSW) showing the arrangement of the Micromegas and sTGC detectors [12].	32
2.14	Schematic of the ATLAS forward detectors showing the LUCID, ALFA, AFP, and ZDC detectors [12].	34
2.15	Schematic of the ATLAS Trigger and Data Acquisition (TDAQ) system showing the Level-1 (L1) trigger, High-Level Trigger (HLT), and data flow [12].	35
3.1	Schematic representation of the particle interactions with different components of the ATLAS detector [14]. The dashed lines correspond to no interaction with that detector component.	38
3.2	Schematic representation of track parameters with respect to the perigee [15].	39
3.3	Electron identification efficiency as a function of transverse momentum (p_T) (left) and pseudorapidity (η) (right) for different identification working points (WPs) [16].	40
3.4	Efficiency of combined muon reconstruction requiring a minimum of four active layers in the Micromegas (MM) and small-strip Thin Gap Chambers (sTGC) detectors, or standalone muons traversing the NSW on the C side of the ATLAS detector. Reduced efficiencies are attributed to detector or readout challenges encountered during the data-taking [17].	41
3.5	Muon reconstruction and identification efficiency for the three WPs during the 2023 data-taking period as a function of (a) transverse momentum (p_T), (b) pseudorapidity (η), and (c) azimuthal angle (ϕ). The samples used in these measurements are from $t\bar{t}$ events and produced using the POWHEG+PYTHIA8 generator [18].	42
3.6	Schematic representation of the particle flow algorithm [19].	43
3.7	Schematic representation of the jet energy calibration steps [20].	43
4.1	Schematic representation of the Monte Carlo simulation process. The red blob represents the hard-scattering event with the structure around it represents the parton showering. The blue color are the incoming partons from the protons involved in the hard-scattering process. The purple color represents the secondary hard scattering events. The light green blobs correspond to the hadronization process, and the darker green represent the hadron decays. The yellow lines are related to the soft photons emitted during the process [21].	49

5.1	The invariant mass distribution of the SFOS lepton pair ($m_{\ell_2\ell_3}$) in the $3\ell+1X$ CR for (a) e -fakes in the 1-SFOS category, (b) μ -fakes in the 1-SFOS category, (c) e -fakes in the 2-SFOS category, and (d) μ -fakes in the 2-SFOS category. The distributions are shown after applying the preselection and the $\Delta R_{\ell_0\ell_1}$ requirements. The uncertainties shown include only statistical uncertainties.	56
5.2	The invariant mass distribution of the SFOS lepton pair ($m_{\ell_2\ell_3}$) in the $2\ell + 2X$ control region for (a) 1SFOS $e\mu$, (b) 2SFOS $e\mu$, (c) 1SFOS ee , (d) 2SFOS ee , (e) 1SFOS $\mu\mu$, and (f) 2SFOS $\mu\mu$ categories. The distributions are shown after applying the preselection and the ΔR_{01} cut. The fake contribution is estimated from data by subtracting the prompt contribution from MC simulation. The uncertainties shown are statistical only.	59
5.3	Distribution of (a) $m_{\ell\ell}^Z$ and (b) m_T^W in the $2\ell + Y$ CR. Only statistical uncertainties are shown.	61
6.1	(a) Schematic of a biological neuron (b) A perceptron with multiple inputs, weights, bias and a non-linear activation function with single output (c) A deep neural network with multiple hidden layers [22].	65
7.1	Feynman diagram illustrating the signal process $ZH \rightarrow ZWW^* \rightarrow lll\nu l\nu$. .	71
7.2	Illustration of the spin correlation of the Higgs boson decay products in the rest frame of the Higgs boson. The black arrows represent the direction of momentum whereas the red arrows represent the spin direction of the particles.	71
7.3	Input variable distributions used for training neural-networks in the 1-SFOS SR. Uncertainties shown are MC statistical uncertainties.	78
7.4	Input variable distributions used for training neural-networks in the 2-SFOS SR. Uncertainties shown are MC statistical uncertainties.	79
7.5	Correlation matrices of the 1-SFOS input variables for ZH and background processes.	80
7.6	Correlation matrices of the 2-SFOS input variables for ZH and background processes.	81
7.8	Distributions of the NN output scores observed from the 5-fold cross-validation technique in 1-SFOS category.	83
7.7	ROC curves observed from the 5-fold cross-validation technique in 1-SFOS category.	84
7.9	ROC Curves for ZH against background node in 1-SFOS category for the final NN trained.	84
7.10	ROC curves observed from the 5-fold cross-validation technique in 2-SFOS category.	85
7.11	Distributions of the NN output scores observed from the 5-fold cross-validation technique in 2-SFOS category.	85
7.12	ROC Curves for ZH against background node in 2-SFOS category for the final NN trained.	86

7.13	Feature ranking for the 1-SFOS signal region. The features are ranked in descending order of their impact on the NN output.	86
7.14	Feature ranking for the 2-SFOS signal region. The features are ranked in descending order of their impact on the NN output.	87
7.15	Neural Network output distribution for the 2-SFOS signal region.	88
7.16	Neural Network output distribution for the 2-SFOS signal region.	89
7.17	The NN output score distribution in the ZZ CR for the 2-SFOS category. .	91
7.18	Plots showing the Data-MC comparison of various kinematic variables in the ZZ control region in 2-SFOS events: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0\ell_1}$ (f) $m_{\ell_2\ell_3}$. The uncertainties contain both statistical and systematic contributions. The statistical uncertainties are indicated by error bars on the data points and systematic uncertainties are shown with hatching in histograms and a yellow band in ratio plots.	92
7.19	Plots showing the Data-MC comparison of various kinematic variables in the ZZ control region in 2-SFOS events: (a) $p_T^{4\ell}$ (b) $m_{4\ell}$ (c) E_T^{miss} (d) N_{jets} (e) $\Delta\phi_{\ell_0\ell_1}^{\text{boost}}$ (f) $\Delta\phi_{\ell_0\ell_1,\text{MET}}$. The uncertainties contain both statistical and systematic contributions. The statistical uncertainties are indicated by error bars on the data points and systematic uncertainties are shown with hatching in histograms and a yellow band in ratio plots.	93
7.20	Distribution of $m_{\tau\tau}$ in the 1-SFOS category after b -jet veto. Negative value corresponds to the events where no $\tau\tau$ pair was produced.	94
7.21	Distributions of observables in the ZZ validation region in 1-SFOS events: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0\ell_1}$ (f) $m_{\ell_2\ell_3}$. The statistical uncertainties are indicated by error bars whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.	96
7.22	Distributions of observables in the ZZ validation region in 1-SFOS events: (a) $p_T^{4\ell}$ (b) $m_{4\ell}$ (c) E_T^{miss} (d) N_{jets} (e) $\Delta\phi_{\ell_0\ell_1}^{\text{boost}}$ (f) $\Delta\phi_{\ell_0\ell_1,\text{MET}}$. The statistical uncertainties are indicated by error bars whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots. . . .	97
7.23	Distribution of observables in the $t\bar{t}Z$ validation region: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0\ell_1}$ (f) $m_{\ell_2\ell_3}$. The statistical uncertainties are shown as error bars on the data points whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.	100
7.24	Distribution of observables in the $t\bar{t}Z$ validation region: (a) $p_T^{4\ell}$ (b) $m_{4\ell}$ (c) E_T^{miss} (d) N_{jets} (e) $\Delta\phi_{\ell_0\ell_1}^{\text{boost}}$ (f) $\Delta\phi_{\ell_0\ell_1,\text{MET}}$. The statistical uncertainties are shown as error bars on the data points whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.	101
9.1	Graphical representation of confidence intervals for a standardized normal distribution with mean ($\mu = 0$) and standard deviation ($\sigma = 1$)[23].	109
10.1	Post-fit distributions of the NN scores in the ZH 1-SFOS SR and 2-SFOS SR.	113
10.2	Profile likelihood scan of the signal strength for the ZH 4ℓ channel.	114

10.3	Pulls and uncertainty breakdowns for all NPs corresponding to the $ZH\ 4\ell$ channel.	116
10.4	Correlation matrix of the uncertainties for the $ZH\ 4\ell$ channel. The correlation plot has a 5% cutoff to show the most relevant correlations.	117
10.5	Profile likelihood scan of the signal strength for the combined VH analysis using Early Run 3 data (2022-2023).	117
10.6	Correlation matrix of the uncertainties for the combined VH analysis using Early Run 3 data (2022-2023). The correlation plot has a 10% cutoff to show the most relevant correlations.	118
10.7	Profile likelihood scan of the signal strength for the combined VH analysis using Run 2 (2015-2018) and Early Run 3 (2022-2023) data.	118
10.8	Correlation matrix of the uncertainties for the combined VH analysis using Run 2 (2015-2018) and Early Run 3 (2022-2023) data. The correlation plot has a 20% cutoff to show the most relevant correlations.	119

List of Tables

1.1	Fundamental forces in nature along with their strengths, mediator particles, and corresponding theories [24]. Here, the strength depends on the source and the distance between the interacting particles. The values given are approximate and meant for comparison purposes.	4
1.2	Major decay channels of the Higgs boson along with their branching ratios for a Higgs mass of approximately 125 GeV [25].	15
3.1	Selection criteria for electrons and muons used in the ZH analysis.	45
3.2	Jets selection criteria used in the ZH analysis.	46
4.1	Summary of the single-lepton trigger requirements used in this analysis for 2022 and 2023 data-taking periods.	48
4.2	Summary of the Good Run Lists (GRLs) used in this analysis for 2022 and 2023 data-taking periods.	48
4.3	Summary of the signal MC samples used in this analysis along with their generators, cross-sections (σ), and precision levels for the highest order calculations.	50
5.1	Requirements of ID and Anti-ID leptons used in defining the prompt and non-prompt leptons in the ZH analysis. The Veto ID selection is only applied to the Anti-ID leptons and refers to the inverted isolation or impact parameter ID-requirements.	53
5.2	Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + e$ -fake CR for the 1-SFOS category. Only statistical uncertainties are included.	54
5.3	Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + \mu$ -fake CR for the 1-SFOS category. Only statistical uncertainties are included.	54
5.4	Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + e$ -fake CR for the 2-SFOS category. Only statistical uncertainties are included.	55
5.5	Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + \mu$ -fake CR for the 2-SFOS category. Only statistical uncertainties are included.	55
5.6	Event yields in the $2l + 2X$ control region for the 1-SFOS category for the $e\mu$ fakes. The uncertainties shown are statistical only.	57
5.7	Event yields in the $2l + 2X$ control region for the 2-SFOS category for the $e\mu$ fakes. The uncertainties shown are statistical only.	57
5.8	Event yields in the $2l + 2X$ control region for the 1-SFOS category for the ee fakes. The uncertainties shown are statistical only.	57

5.9	Event yields in the $2l + 2X$ control region for the 2-SFOS category for the ee fakes. The uncertainties shown are statistical only.	57
5.10	Event yields in the $2l + 2X$ control region for the 1-SFOS category for the $\mu\mu$ fakes. The uncertainties shown are statistical only.	57
5.11	Event yields in the $2l + 2X$ control region for the 2-SFOS category for the $\mu\mu$ fakes. The uncertainties shown are statistical only.	58
5.12	Event yields observed in data and expected from MC processes in the $2\ell+1Y$ CR. The uncertainties shown are statistical only.	60
5.13	Event yields observed in data and expected from MC processes in the WZ CR. The uncertainties shown are statistical only.	60
5.14	Summary of the fake factor values measured in the $2\ell+1Y$ CR for electrons and muons using data-driven technique.	60
5.15	Summary of the prompt lepton subtraction systematic uncertainties on the fake factors for electrons and muons. The "Up" and "Down" columns represent the fake factor values obtained by varying the WZ yield by $\pm 30\%$ in the Z +jets control region. The 'Difference' columns show the relative differences between the nominal fake factor and the varied fake factors. The uncertainties on nominal fake factors due to statistical errors are also shown.	62
5.16	Fake factors calculated using MC simulation in the $2\ell+1Y$ CR for electrons and muons. Sherpa MC samples are used for the Z +jets process.	62
5.17	Summary of the non-closure systematic uncertainties on the fake factors for electrons and muons. The second row for each flavor shows the estimated fakes in the SR using MC-based fake factors. The 'Difference' column shows the relative difference between the nominal fakes using data-driven fake factors and the fakes estimated using MC-based fake factors.	63
5.18	Summary of all the systematic uncertainties associated with the fake factor method for electrons and muons. The total uncertainty is calculated by adding the individual uncertainties in quadrature.	63
7.1	Summary of the selection criteria defining the signal regions for the ZH analysis before applying the neural network selection.	74
7.2	Cutflow for the ZH pre-selection and up to cuts on NN output. Uncertainties shown are purely statistical. "Fakes(DD)" is the data-driven fakes estimate (section 5). "Signif." is the expected significance.	75
7.3	List of input variables used for training the neural networks in both 1SFOS and 2SFOS categories, along with their descriptions.	77
7.4	Set of hyperparameters and their corresponding search space used for tuning the neural networks for both 1-SFOS and 2-SFOS categories.	82
7.5	Summary of the final hyperparameter values selected for the neural network training in the 1-SFOS and 2-SFOS SRs.	83
7.6	Summary of the selections defining the ZZ enriched regions in the 1-SFOS and 2-SFOS categories. These requirements are applied after the b-veto selection from Table 7.1.	89

7.7	Table showing the expected and observed event yields in the ZZ control region for the 2-SFOS category. The uncertainties shown are statistical only.	90
7.8	Table showing the expected and observed event yields in the ZZ validation region for the 1-SFOS category. The uncertainties shown are statistical only. No normalization factor is applied.	95
7.9	Selection criteria defining the $t\bar{t}Z$ VR.	98
7.10	Expected and observed event yields in the $t\bar{t}Z$ validation region. The uncertainties shown are statistical only.	99
8.1	Summary of various nuisance parameters with their short descriptions and sources of the experimental systematic uncertainties considered in this analysis.	104
8.2	List of theory systematic uncertainties considered in the ZH channel. All uncertainties are given as relative variations on the yields. The yields are assumed to be taken from the signal region unless specified otherwise. Uncertainties from multiple sources are added in quadrature. N.B.: $q\bar{q} \rightarrow ZH$ and $gg \rightarrow ZH$ are combined simply as ZH	106
10.1	Expected results for the ZH 4ℓ channel.	114
10.2	Expected results for the combined VH analysis using Early Run 3 data (2022-2023).	115
10.3	Expected results for the combined VH analysis using Run 2 (2015-2018) and Early Run 3 (2022-2023) data.	115

List of Acronyms

ACT	ATLAS Common Tracking
AEOT	Alignment Effect On Track
AFP	ATLAS Forward Proton
ALFA	Absolute Luminosity For ATLAS
ALICE	A Large Ion Collider Experiment
AMVF	Adaptive Multi-Vertex Finder
ANN	Artificial Neural Network
ATLAS	A Toroidal LHC ApparatuS
BDT	Boosted Decision Tree
BSM	Beyond Standard Model
CB	Combined
CERN	European Organization for Nuclear Research
CM	Center of Mass
CMS	Compact Muon Solenoid
CR	Control Region
CSC	Cathode Strip Chamber
CT	Calorimeter-Tagged
DFOS	Different Flavor Opposite Sign
DNN	Deep Neural Network
DQM	Data Quality Monitoring
EM	Electromagnetic
EW	ElectroWeak
FCal	Forward Calorimeter
FF	Fake Factor
fJVT	forward Jet Vertex Tagging
ggF	gluon-gluon Fusion
GRL	Good Run List

HEC	Hadronic End-cap Calorimeter
HLT	High-Level Trigger
IBL	Insertable B-layer
ID	Inner Detector
IP	Interaction Point
IQR	Inter Quartile Range
JER	Jet Energy Resolution
JES	Jet Energy Scale
JVT	Jet Vertex Tagging
L1	Level-1
LAr	Liquid Argon
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LINAC	Linear Accelerator
LO	Leading Order
LUCID	LUMinosity Cherenkov Integrating Detector
LS	Long Shutdown
MAE	Mean Absolute Error
MC	Monte Carlo
MDT	Monitored Drift Tube
MET	Missing Transverse Energy
MLE	Maximum Likelihood Estimator
MS	Muon Spectrometer
MSE	Mean Squared Error
MVA	Muti-Variate Analysis
NLO	Next-to-Leading Order
NN	Neural Network
NNLO	Next-to-Next-to-Leading Order
NP	Nuisance Parameter

NSW	New Small Wheel
OR	Overlap Removal
P4	Four Momentum
PDF	Parton Distribution Function
PF	Particle Flow
PMG	Physics Modelling Group
PMT	Photomultiplier Tube
POI	Parameter of Interest
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PV	Primary Vertex
QCD	Quantum ChromoDynamics
QED	Quantum ElectroDynamics
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
RPC	Resistive Plate Chamber
SCT	Semi-conductor Tracker
SGD	Stochastic Gradient Descent
SF	Scale Factor
SFOS	Same Flavor Opposite Sign
SM	Standard Model
SPS	Super Proton Synchrotron
SR	Signal Region
ST	Segment-Tagged
STXS	Simplified Template Cross-Section
TDAQ	Trigger and Data Acquisition
TGC	Thin Gap Chamber
TRT	Transition Radiation Tracker
VBF	Vector Boson Fusion

VBS	Vector Boson Scattering
VEV	Vacuum Expectation Value
VH	Higgs Strahlung
VR	Validation Region
WP	Working Point
ZDC	Zero Degree Calorimeter

Introduction

The quest to understand the fundamental constituents of matter and the forces that govern their interactions has been a central focus of physics for centuries. The Greek philosopher Democritus first coined the term “atom” around 400 BC, suggesting that matter is composed of small, indivisible particles called “atomos”. This laid the foundation for our modern understanding of the building blocks of matter. In the early 20th century, the experiment by J.J. Thomson led to the discovery of the electron [26], followed by Ernest Rutherford’s discovery of the atomic nucleus [27]. These experiments showed that atoms are not indivisible, and something more fundamental than the atom exists.

With the advent of quantum mechanics, our understanding of atomic structure and behavior evolved significantly. Numerous theories and models were developed to describe the interactions of subatomic particles, leading to the formulation of the Standard Model of particle physics in the latter half of the 20th century. The Standard Model (SM) of particle physics [28] is the most comprehensive theory to date that describes the fundamental particles and their interactions. Particle accelerators were built to test the predictions of the SM, leading to the discovery of various particles, including quarks, leptons, and gauge bosons. The discovery of the Higgs boson in 2012 at the largest particle accelerator in the world, the Large Hadron Collider (LHC) [29, 30], by the ATLAS [31] and CMS collaborations [32], marked a significant milestone in our understanding of the fundamental particles and their underlying mechanisms.

The SM, although successful in describing the experimental observations, is not without its limitations. It doesn’t explain the imbalance of matter and antimatter in the universe, the nature of dark matter and dark energy, and the integration of gravity with the other fundamental forces. One of the most important aspects of doing research in high energy physics experiment involves studying the properties and interactions of the existing particles predicted by the Standard Model. In particular, the Higgs boson plays a crucial role as it is responsible for giving mass to other fundamental particles through the Higgs mechanism [33, 34, 35]. Studying the production and decay of the Higgs boson provides valuable insights into its properties and help validate the predictions of the Standard Model or reveal potential deviations that could indicate new physics.

The production of the Higgs boson in association with a vector boson V (W or Z) as well as Higgs decaying via $H \rightarrow WW^*$ channel are observed independently with high significance. However, the VH production with Higgs decaying via $H \rightarrow WW^*$ channel has only reached evidence level in previous measurements by studying the leptonic final

states using the Run 2 data [36]. It was found that the WH production with $H \rightarrow WW^*$ decay with 3 leptons in the final state and ZH production with $H \rightarrow WW^*$ decay with 4 leptons in the final state were the most sensitive channels. These two channels were also found to be statistically limited. Therefore, with the increased luminosity and center-of-mass energy in Run 3, the WH and ZH channels with $H \rightarrow WW^*$ decay are studied again expecting to improve the significance by reducing the statistical uncertainties, especially by combining the Run 2 and Run 3 data.

This thesis focuses on the measurement of the Higgs boson production cross-section in association with a Z boson, decaying via the $H \rightarrow WW^*$ decay channel using the ATLAS detector [37] at the LHC at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV. The analysis uses data collected in 2022 and 2023 during Run 3 of the LHC and corresponds to an integrated luminosity of 56.1 fb^{-1} . Fully leptonic final state with four charged leptons (electrons or muons) and missing transverse energy from neutrinos is considered [38].

The structure of this thesis follows a logical progression through the various aspects of the analysis and is organized as follows:

- Chapter 1 provides a theoretical framework for the Standard Model of particle physics, including a discussion of the Higgs mechanism and the production and decay modes of the Higgs boson.
- Chapter 2 describes the ATLAS detector and its various subsystems, highlighting the key components relevant to this analysis.
- Chapter 3 discusses the reconstruction and identification of physics objects, such as electrons, muons, and missing transverse energy, which are crucial for the event selection.
- Chapter 4 outlines the data and Monte Carlo samples used in the analysis, along with the trigger strategy employed to select events of interest.
- Chapter 5 details the strategy applied for estimating contribution from fake processes in the signal region.
- Chapter 6 presents a detailed description of the machine learning technique, in particular neural networks used in this analysis to enhance the signal-to-background discrimination.
- Chapter 7 gives an overview of the analysis strategy, the background processes considered, and the event selection criteria used to define the signal region, control regions, and validation regions. The neural network training and final architecture implemented are also discussed in this chapter.
- Chapter 8 describes the systematic uncertainties considered in the analysis and their impact on the final results.
- Chapter 9 presents the statistical analysis performed to extract the signal strength which helps to either confirm the SM predictions or indicate new physics.

- Chapter 10 discusses the expected results obtained from this analysis, including the measured signal strength and the corresponding uncertainties.
- Chapter 11 summarizes the key findings of the analysis and discusses potential avenues for future research in this area.

1 The Standard Model of Particle Physics

Fundamental particles are the building blocks of the universe; they cannot be broken down into smaller components [39]. They interact through fundamental forces, which govern their behavior and interactions. Numerous theoretical models were developed in the 1960s to explain the interactions of fundamental particles. The “standard model of gauge interactions” emerged in the 1970s as the most capable model to describe the experimental data available at the time, eventually becoming the “Standard Model” (SM) of particle physics [28].

The SM is a quantum field theory in which particles are quantized excitations of underlying fields. It is based on the principles of special relativity and quantum mechanics, with particle interactions described by the exchange of force-carrying particles. It explains three of the four fundamental forces: the electromagnetic force, the weak nuclear force, and the strong nuclear force. Gravity is very weak compared to the other forces at the particle level and is not included in the Standard Model. Table 1.1 summarizes the fundamental forces in nature in decreasing order of their strength.

Force	Strength	Mediator Particle(s)	Theory
Strong	10	Gluon (g)	Chromodynamics
Electromagnetic	10^{-2}	Photon (γ)	Electrodynamics
Weak	10^{-13}	W and Z Bosons ($W^\pm, Z^0$)	Flavordynamics
Gravitational	10^{-42}	Graviton (hypothetical)	Geometrodynamics

Table 1.1: Fundamental forces in nature along with their strengths, mediator particles, and corresponding theories [24]. Here, the strength depends on the source and the distance between the interacting particles. The values given are approximate and meant for comparison purposes.

There are two main types of particles in the Standard Model: fermions and bosons. Fermions are half-integer spin particles that make up matter, while bosons are integer spin particles that mediate forces. Fermions consist of quarks and leptons, which are organized into three generations based on their masses [40]. There are six flavors of quarks: up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Independent quarks have not been observed due to a phenomenon known as color confinement [41]. Quarks combine to form hadrons, which are composite particles such as baryons (made of three quarks) and mesons (made of one quark and one antiquark). Protons and neutrons are examples of baryons that make up atomic nuclei. Leptons include the electron (e), muon (μ), tau (τ), and their corresponding neutrinos (ν_e, ν_μ, ν_τ). Each fermion has a corresponding antiparticle with the same mass but opposite charge.

The bosons in the Standard Model include the photon (γ), the W and Z bosons,

the gluons (g), and the Higgs boson (H). Photons are carriers of the electromagnetic force, while the W and Z bosons mediate the weak nuclear force, which is responsible for processes like beta decay. Quark interactions happen through the exchange of gluons, which are the mediators of the strong force. The Higgs boson, discovered in 2012 at the Large Hadron Collider (LHC) [31, 32], is associated with the Higgs field, which gives mass to other fundamental particles through the Higgs mechanism [42]. Figure 1.1 shows the particles and their properties in the Standard Model.

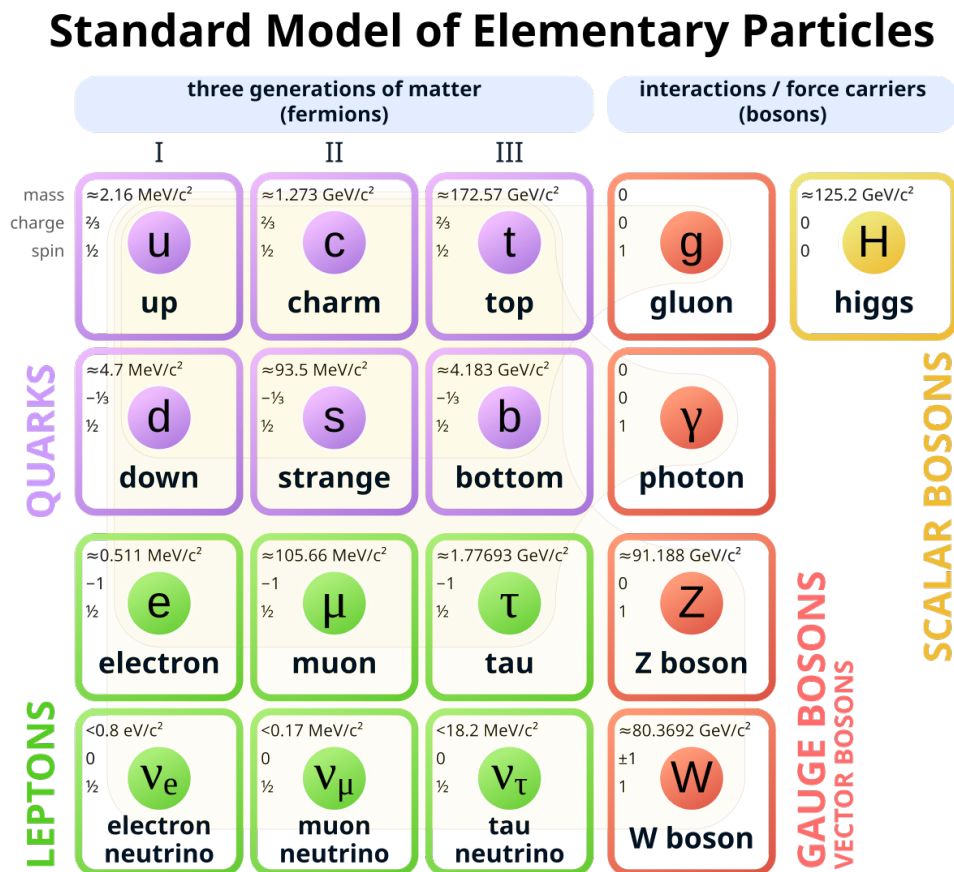


Figure 1.1: Tabular representation of particles of the Standard Model along with their properties such as mass, charge, and spin [1].

Although the Standard model has been extremely successful in explaining a wide range of experimental results, it has limitations. For starters, it does not incorporate gravity, which is described by general relativity. It cannot explain the matter-antimatter imbalance in the universe, nor does it account for dark matter and dark energy, which are believed to make up a significant portion of the universe's mass-energy content [43]. These are some of the puzzle pieces that are missing from the Standard Model, and suggest the existence of physics beyond the Standard Model. In this chapter, the theoretical framework of the Standard Model is discussed, with definition of groups and gauge theories in Section

1.1, followed by detailed discussions on Quantum Electrodynamics (QED) in Section 1.2, Quantum Chromodynamics (QCD) in Section 1.3, weak interaction and electroweak theory in Section 1.4. This follows with a discussion on spontaneous symmetry breaking in Section 1.5 and Higgs mechanism in Section 1.6. Various production and decay modes of the Higgs boson are discussed in Section 1.7.

1.1. Groups and Gauge Theories

Group is a mathematical concept of abstract algebra to describe a set of elements. The elements of the set satisfies the group axioms under the operation such as closure, associativity, identity, and invertibility. These groups can be abelian which satisfies the commutative property or non-abelian which does not. Noether's theorem relates symmetries to conservation laws in physics, stating that for every symmetry there is a conservation law and vice versa [24]. For example, the momentum of a system is conserved under spatial translation symmetry or invariance, while rotational symmetry leads to conservation of angular momentum. The Standard Model (SM) of particle physics is built upon the principles of symmetry and group theory, which provide the mathematical framework for understanding the fundamental forces and particle interactions.

Hermann Weyl first introduced the concept of gauge invariance in 1918 while attempting to unify electromagnetism and gravity [44]. Gauge theory refers to the invariance of the Lagrangian under local transformations of certain symmetry groups. The Lagrangian density ($\mathcal{L}$) for a field ϕ describes the dynamics of the field. The Euler-Lagrange equation for the field ϕ is derived from the principle of least action and is given by:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0 \quad (1.1)$$

where ∂_μ represents the four-gradient operator corresponding to spacetime coordinates. Gauge invariance of this Lagrangian density leads to the introduction of gauge bosons and conservation laws associated with the symmetries. In gauge theories, the symmetry groups are typically Lie groups, which are continuous groups characterized by smooth transformations. The SM is a gauge theory based on the $SU(3) \times SU(2) \times U(1)$ symmetry group where each Lie group is associated with a specific gauge interaction described below:

- **U(1) Gauge Group:** This group is associated with the electromagnetic interaction and is the basis of Quantum Electrodynamics (QED). The U(1) gauge symmetry leads to the existence of the photon as the gauge boson mediating the electromagnetic force. The invariance under U(1) transformations ensures that the electric charge is conserved.
- **SU(2) Gauge Group:** The SU(2) group is related to the weak interaction and forms part of the electroweak theory, which unifies the electromagnetic and weak forces. The SU(2) gauge symmetry gives rise to three gauge bosons: the W^+ , W^- , and Z bosons, which mediate weak interactions. The invariance under SU(2) trans-

formations leads to the conservation of weak isospin.

- **SU(3) Gauge Group:** This group describes the strong interaction and is the foundation of Quantum Chromodynamics (QCD). The SU(3) gauge symmetry leads to the existence of eight gluons as the gauge bosons that mediate the strong force between quarks. The invariance under SU(3) transformations ensures the conservation of color charge.

1.2. Quantum Electrodynamics

As discussed earlier, in QFT the Lagrangian of particles is described using their corresponding fields and particles are interpreted as excitations of these fields. The Lagrangian of a spinor (spin-1/2) field (Dirac field) ψ with mass m is given by:

$$\mathcal{L} = i(\hbar c)\bar{\psi}\gamma^\mu\partial_\mu\psi - (mc^2)\bar{\psi}\psi \quad (1.2)$$

where $\bar{\psi} = \psi^\dagger\gamma^0$ is the Dirac adjoint, γ^μ are the gamma matrices. This Lagrangian is invariant under a global phase transformation of the field ($\psi \rightarrow \psi' = e^{-iq\theta/\hbar c}\psi$), where θ is a constant. This invariance leads to the conservation of electric charge according to Noether's theorem. However, under a local gauge transformation i.e.,

$$\psi \rightarrow \psi' = e^{-iq\theta(x)/\hbar c}\psi \quad (1.3)$$

where $\theta(x)$ is a function of spacetime coordinates, the Lagrangian is not invariant. Here, q is the charge of the particle. To ensure invariance, the extra term needs to be canceled out by introducing a new vector field A_μ that transforms as $A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu\lambda(x)$, where $\lambda(x) = -(\hbar c/q)\theta(x)$. Thus, the new Lagrangian becomes:

$$\mathcal{L} = i(\hbar c)\bar{\psi}\gamma^\mu\partial_\mu\psi - (mc^2)\bar{\psi}\psi - q(\bar{\psi}\gamma^\mu\psi)A_\mu \quad (1.4)$$

The extra term in the end corresponds to the interaction between the Dirac field and the vector field. The Lagrangian of a free vector field A_μ , also known as Proca Lagrangian, is given by:

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{m_A c}{\hbar}\right)^2 A^\mu A_\mu \quad (1.5)$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the field strength tensor. This Lagrangian is gauge invariant only if the mass of the vector field $m_A = 0$.

The transformation of Dirac field $\psi \rightarrow \psi' = e^{i\theta}\psi$ is of the form $\psi \rightarrow \psi' = U\psi$, where $U = e^{i\theta}$ is a unitary matrix. Thus, invoking local gauge invariance of a Dirac field under U(1) symmetry group necessitates the introduction of a massless vector field that interacts with the Dirac field. The combined Lagrangian for the Dirac field and the vector field is given by:

$$\mathcal{L} = [i(\hbar c)\bar{\psi}\gamma^\mu\partial_\mu\psi - (mc^2)\bar{\psi}\psi] - \left[\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\right] - q(\bar{\psi}\gamma^\mu\psi)A_\mu \quad (1.6)$$

This represents the Lagrangian of Quantum Electrodynamics (QED) [24], where the vector field A_μ is identified as the electromagnetic field (photon) that mediates the electromagnetic interaction between charged particles i.e. fermions.

The strength of these interactions is characterized by the fine-structure constant α , which is defined as:

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \quad (1.7)$$

where e is the charge of the electron. As the value of α is small, $\alpha \ll 1$, QED is treated as a perturbative theory. This helps in estimating physical observable using approximate methods when an exact solution is not possible in interaction theories.

1.3. Quantum Chromodynamics

In 1964, Murray Gell-Mann and George Zweig independently proposed the quark model [45, 46], which introduced three flavors of quarks (u , d , s) as spin-1/2 particles. Combining the spin and flavor, a SU(6) symmetry group was formulated to explain their dynamics [47]. However, this posed the problem of symmetric wavefunction for baryons made of three identical quarks, a clear violation of Fermi statistics. This was resolved by introducing a new quantum number called color charge [48, 49, 50]. Three more flavours of quarks (c , b , t) were later discovered bringing the total to six flavours. Each quark flavor has three color states: red, green, and blue. For a given flavor, all the three colors have the same mass. Building on the success of Weyl's gauge theory for QED, Yang and Mills extended the concept of gauge invariance to non-Abelian symmetry groups in 1954 [51]. This development laid the foundation for Quantum Chromodynamics (QCD) which is invariant under the non-Abelian SU(3) gauge group.

The Lagrangian for any quark flavor with three color states, each of mass m , is given by:

$$\mathcal{L} = i\hbar c \bar{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \bar{\psi} \psi \quad (1.8)$$

where ψ is a three-component vector in color space such that

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}, \quad \bar{\psi} = (\bar{\psi}_r, \bar{\psi}_g, \bar{\psi}_b) \quad (1.9)$$

and ψ_r , ψ_g , ψ_b represent the red, green, and blue color states respectively. This is similar to the Dirac Lagrangian discussed in Section 1.2, thus it is invariant under global U(3) transformation where $\psi \rightarrow \psi' = U\psi$. Here, U can be expressed as $U = e^{iH}$, where H is a Hermitian 3×3 matrix. Here, H can be further decomposed as $H = \theta_0 I + \theta_a T_a$, where θ_0 and θ_a are real parameters, I is the identity matrix and T_a represent the Gell-Mann matrices which are the generators of the SU(3) group [24]. Under a local gauge transformation, an additional term appears in the Lagrangian which can be canceled out

by introducing eight new vector fields G_μ^a (where $a = 1, 2, \dots, 8$) that transform as

$$G_\mu^a \rightarrow G_\mu^{a'} = G_\mu^a + \partial_\mu \lambda^a(x) + g_s f^{abc} \lambda^b(x) G_\mu^c \quad (1.10)$$

where $\lambda^a(x)$ are functions of spacetime coordinates, g_s is the strong coupling constant, and f^{abc} are the structure constants of the SU(3) group. Thus, the new Lagrangian becomes:

$$\mathcal{L} = i\hbar c \bar{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \bar{\psi} \psi - g_s (\bar{\psi} \gamma^\mu T_a \psi) G_\mu^a \quad (1.11)$$

The extra term in the end corresponds to the interaction between the quark field and the vector fields.

Combining the Lagrangian of the quark field with that of the vector fields, the QCD Lagrangian is given by:

$$\mathcal{L} = [i\hbar c \bar{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \bar{\psi} \psi] - \left[\frac{1}{4} G^{\mu\nu a} G_{\mu\nu}^a \right] - g_s (\bar{\psi} \gamma^\mu T_a \psi) G_\mu^a \quad (1.12)$$

where $G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c$ is the field strength tensor for the gluon fields. Here, the vector fields G_μ^a are identified as the gluon fields that mediate the strong interaction between quarks.

The vector fields of gluons interact with each other, as seen from the expression of the field strength tensor $G_{\mu\nu}^a$. This is a key difference between QCD and QED, where the photon fields do not self-interact. These self-interactions arise from the non-Abelian nature of the SU(3) gauge symmetry. Thus, any effort to pull the quarks apart results in the generation of new quark-antiquark pairs from the vacuum, leading to the formation of new hadrons. This is why free quarks are not observed in nature and this phenomenon is known as color confinement. On the other hand, at very high energies or short distances, the strong coupling constant g_s becomes small, allowing quarks to behave as nearly free particles. This property is known as asymptotic freedom. This allows perturbative calculations to be performed in high-energy regimes, similar to QED.

1.4. Weak Interactions and Electroweak Theory

The existence of the nuclear weak force is well established through various phenomena such as beta decay, where a neutron decays into a proton, an electron, and an antineutrino. However, the weak interaction is not described by Quantum Electrodynamics (QED) alone. It is rather part of a larger framework known as the electroweak theory. The electroweak theory unifies the weak interactions and the electromagnetic interactions as one. Developed by Sheldon Glashow, Abdus Salam, and Steven Weinberg in the 1960s, the electroweak theory is gauge invariant under the SU(2) x U(1) symmetry group [52, 53, 54]. The SU(2) group corresponds to the weak isospin symmetry, while the U(1) group corresponds to the weak hypercharge symmetry. The gauge invariance of the electroweak theory leads to the prediction of four gauge bosons that mediate the electromagnetic and weak interactions.

These bosons are the photon (γ), which mediates the electromagnetic interaction, and the W^+ , W^- , and Z bosons, which mediate the weak interaction.

The weak interaction is unique in that it can change the flavor of quarks at the interaction vertex, hence the weak interaction theory is also referred to as “flavordynamics”. Unlike other forces, the weak interactions apply to all fermions, including leptons and quarks. Another distinctive feature of the weak interaction is its violation of parity symmetry [55] observed in Wu’s experiment [56] in 1957. Parity refers to the particle behavior under spatial inversion (i.e., mirror reflection) and related to its intrinsic handedness or chirality. The two handedness states are left-handed and right-handed as shown in Figure 1.2. In weak interactions, only left-handed fermions and right-handed antifermions participate.

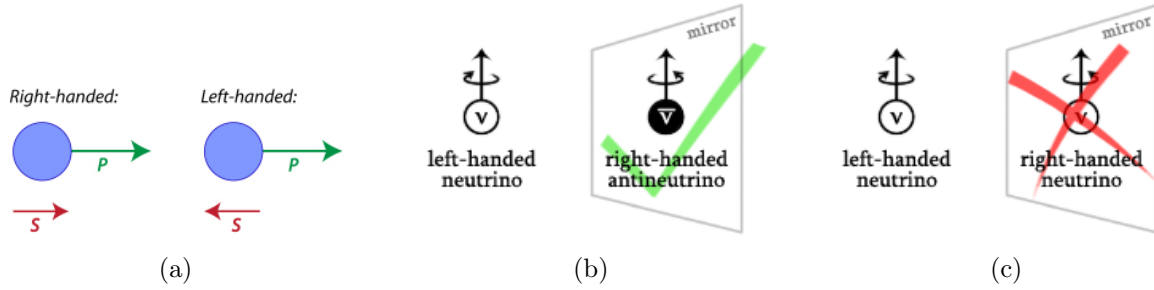


Figure 1.2: Schematic representation of parity and handedness: (a) Handedness [2] (b) Parity is conserved (c) Parity is violated [3].

The Lagrangian for the electroweak theory is given by:

$$\mathcal{L} = \bar{\psi}_L i \gamma^\mu D_\mu \psi_L + \bar{\psi}_R i \gamma^\mu D_\mu \psi_R - \frac{1}{4} W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (1.13)$$

where ψ_L and ψ_R are the left-handed and right-handed fermion fields respectively, $W_{\mu\nu}^a$ is

$$\begin{aligned} W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g \epsilon^{abc} W_\mu^b W_\nu^c, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\ D_\mu &= \partial_\mu + ig W_\mu^a \frac{\tau^a}{2} + ig' Y B_\mu, \end{aligned} \quad (1.14)$$

where g and g' are the coupling constants for the SU(2) and U(1) gauge groups respectively, τ^a are the Pauli matrices, and Y is the weak hypercharge operator.

Under U(1) gauge transformation, the B_μ field transforms as:

$$B_\mu \rightarrow B'_\mu = B_\mu + \partial_\mu \lambda(x) \quad (1.15)$$

where $\lambda(x)$ is a function of spacetime coordinates. Similarly, under SU(2) gauge transfor-

mation, the W_μ^a fields transform as:

$$W_\mu^a \rightarrow W_\mu^{a'} = W_\mu^a + \partial_\mu \alpha^a(x) + g \epsilon^{abc} \alpha^b(x) W_\mu^c \quad (1.16)$$

where $\alpha^a(x)$ are functions of spacetime coordinates, g is the weak coupling constant, and ϵ^{abc} are the structure constants of the SU(2) group.

The gauge fields introduced above are massless. However, the W and Z bosons are known to be massive from experimental observations. This mass generation is explained through the Higgs mechanism in which spontaneous symmetry breaking occurs, as discussed in Section 1.6. But before the spontaneous symmetry breaking occurs, the physical gauge bosons are obtained through the following linear combinations of the original gauge fields:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \\ Z_\mu &= W_\mu^3 \cos \theta_W - B_\mu \sin \theta_W, \\ A_\mu &= W_\mu^3 \sin \theta_W + B_\mu \cos \theta_W, \end{aligned} \quad (1.17)$$

where θ_W is the weak mixing angle (Weinberg angle). Here, $W_\mu^\pm$ are the charged W boson fields, Z_μ is the neutral Z boson field, and A_μ is the photon field, where all the fields are still massless at this stage.

1.5. Spontaneous Symmetry Breaking

Spontaneous symmetry breaking (SSB) is a phenomenon in which the ground state (vacuum state) of a system does not exhibit the same symmetry as the underlying laws governing the system (Lagrangian) [24]. This is demonstrated using a massless scalar field ϕ whose Lagrangian density is given by:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \frac{1}{2}\mu^2 \phi^2 - \frac{1}{4}\lambda \phi^4 \quad (1.18)$$

Here, μ^2 and λ are constants with $\lambda > 0$ for stability. $\mathcal{L}$ can be rewritten as $T - V(\phi)$, where T is the kinetic term and $V(\phi)$ is the potential energy term. The vacuum state of the system corresponds to the minimum of the potential energy, which is given by:

$$V(\phi) = -\frac{1}{2}\mu^2 \phi^2 + \frac{1}{4}\lambda \phi^4 \quad (1.19)$$

When $\mu^2 < 0$, the potential has a single minimum at $\phi = 0$, which is symmetric under the transformation $\phi \rightarrow -\phi$. However, when $\mu^2 > 0$, there are two degenerate minima at $\phi = \pm \sqrt{\mu^2/\lambda}$. In this case, the system can spontaneously choose either of the minima as its ground state, thereby breaking the original symmetry of the Lagrangian.

Choosing either of the minima, leads to a new field definition:

$$\eta \equiv \phi \pm \sqrt{\frac{\mu^2}{\lambda}} \quad (1.20)$$

Substituting this into the Lagrangian results in:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 - \sqrt{\lambda \mu^2} \eta^3 - \frac{1}{4} \lambda \eta^4 + \text{constant} \quad (1.21)$$

Notice the appearance of a mass term (second term in Equation 1.21) for the new field η , indicating that the field acquires mass through spontaneous symmetry breaking.

1.6. The Higgs Mechanism

Higgs mechanism is a process by which gauge bosons in the $SU(2) \times U(1)$ electroweak theory acquire mass through spontaneous symmetry breaking [33, 34, 35]. This mechanism introduces a doublet of complex scalar fields for the underlying $SU(2)$ symmetry, defined as follows:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = 1/\sqrt{2} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.22)$$

where ϕ^+ and ϕ^0 are complex scalar fields, and ϕ_1, ϕ_2, ϕ_3 , and ϕ_4 are real scalar fields. ϕ^+ is electrically charged, while ϕ^0 is electrically neutral. The Lagrangian density for the Higgs field is given by:

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad (1.23)$$

where D_μ is the covariant derivative given in Equation 1.14 and $V(\Phi)$ is the Higgs potential defined as:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.24)$$

Here, μ^2 and λ are constants with $\lambda > 0$ for stability. When $\mu^2 < 0$, the potential has a minimum at zero, but when $\mu^2 > 0$, the potential exhibits a “Mexican hat” shape with a ring of degenerate minima at non-zero values of Φ as shown in Figure 1.3.

Choosing a specific vacuum expectation value (VEV) for the Higgs field modifies the Lagrangian, leading to mass terms for the W and Z bosons:

$$m_W = \frac{1}{2} g v, \quad m_Z = \frac{1}{2} \sqrt{g^2 + g'^2} v \quad (1.25)$$

where $v = \sqrt{-\mu^2/\lambda}$ is the VEV of the Higgs field, and g and g' are the coupling constants for the $SU(2)$ and $U(1)$ gauge groups respectively. The photon remains massless due to the unbroken $U(1)$ symmetry.

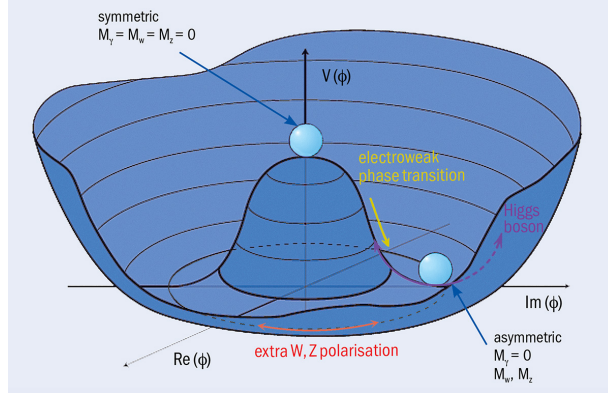


Figure 1.3: Schematic representation of the Higgs potential for $\mu^2 > 0$ showing the “Mexican hat” shape [4].

1.7. The Higgs Boson Physics

The Higgs boson was discovered in 2012 at the Large Hadron Collider (LHC) [31, 32]. It was proposed in the 1960s as part of the Higgs mechanism and is associated with the Higgs field, which gives mass to other particles through spontaneous symmetry breaking [33, 34, 35]. Its discovery was a monumental achievement in particle physics, confirming the last missing piece of the Standard Model. Studying the properties of the Higgs boson, including its production and decay modes, is essential to test the Standard Model and explore potential new physics. This analysis focuses on the measurement of Higgs boson production in association with a Z boson, decaying via $H \rightarrow WW^*$ channel. The following sections discuss the relevant production and decay modes.

1.7.1 Production Modes

The Higgs boson can be produced through several mechanisms, but there are four main production modes at the LHC:

- **Gluon-gluon fusion (ggF):** This is the dominant production mode, where two gluons from the colliding protons create a virtual loop of quarks that then produces a Higgs boson. As the coupling strength of the Higgs boson to any particles is proportional to their mass, the top quark loop provides the largest contribution to this process.
- **Vector boson fusion (VBF):** In this process, two quarks from the colliding protons emit W or Z bosons, which fuse together to produce a Higgs boson. This process is characterized by the presence of two high-energy jets in the final state, which are produced by the scattered quarks. This is the second most common production mode at the LHC.
- **Higgs Strahlung (VH):** In this mode, a Higgs boson is produced in association with a W or Z boson. There is also an off-shell contribution to this process, where

a virtual W or Z boson decays into a Higgs boson and another W or Z boson. This production mode is particularly important for studying the couplings of the Higgs boson to vector bosons. This production mode is the focus of this analysis.

- **$t\bar{t}H/b\bar{b}H$ production:** In this process, a Higgs boson is produced in association with a top quark-antiquark pair ($t\bar{t}H$) or a bottom quark-antiquark pair ($b\bar{b}H$).

Figure 1.4 shows the leading-order Feynman diagrams for these production modes.

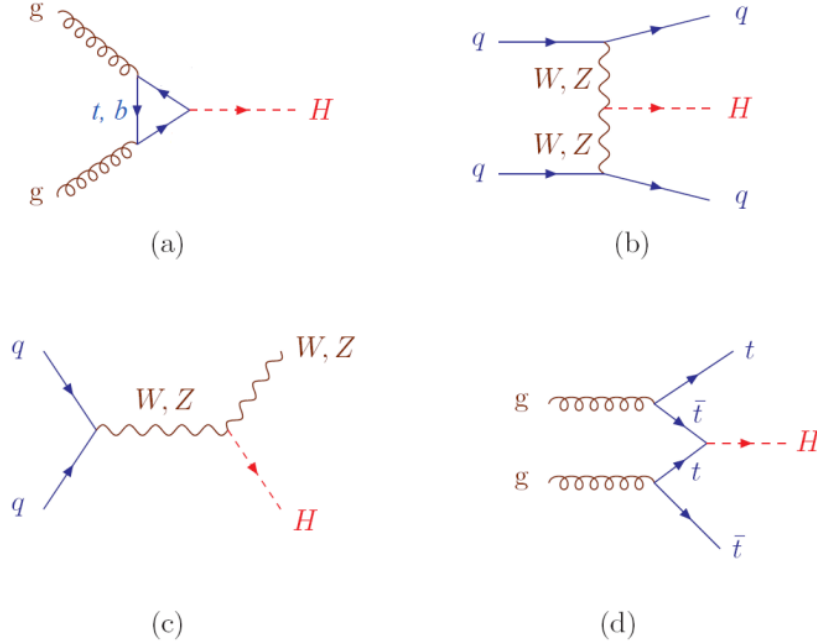


Figure 1.4: Leading-order Feynman diagrams for the main Higgs boson production modes at the LHC: (a) gluon-gluon fusion (ggF), (b) vector boson fusion (VBF), (c) Higgs Strahlung (ZH/WH), and (d) $t\bar{t}H/b\bar{b}H$ production [5].

1.7.2 Decay Modes

The Higgs boson can decay through several channels, with the specific decay modes and their branching ratios depending on its mass. The primary decay channel of the Higgs boson with a mass of approximately 125 GeV is into a pair of bottom quarks ($H \rightarrow b\bar{b}$), which has a branching ratio of about 58%. The bottom quark being heavy, the Higgs boson couples more strongly to it, making this decay mode the most probable. The next most significant decay channel is into a pair of W bosons ($H \rightarrow WW^*$), with a branching ratio of around 21%. As one of the W bosons is off-shell it is not as kinematically favored as the $b\bar{b}$ channel, but it is still a significant decay mode. The remaining decay channels like $H \rightarrow ZZ^*$, $H \rightarrow \tau^+\tau^-$, $H \rightarrow gg$, and $H \rightarrow \gamma\gamma$ have smaller branching ratios, ranging from a few percent to less than one percent as shown in Table 1.2. In this analysis, the $H \rightarrow WW^*$ decay channel is studied, where both W bosons decay leptonically into

charged leptons (e or μ) and neutrinos (ν). This provides cleaner signatures in the detector compared to hadronic decay modes, despite the lower branching ratio, as hadronic decays suffer from large QCD backgrounds.

Decay Channel	Branching Ratio (%)
$H \rightarrow b\bar{b}$	58.2
$H \rightarrow WW^*$	21.4
$H \rightarrow gg$	8.2
$H \rightarrow \tau^+\tau^-$	6.3
$H \rightarrow ZZ^*$	2.6
$H \rightarrow \gamma\gamma$	0.23
$H \rightarrow Z\gamma$	0.15

Table 1.2: Major decay channels of the Higgs boson along with their branching ratios for a Higgs mass of approximately 125 GeV [25].

Figure 1.5 (a) shows the production cross-section of the Higgs boson at a mass of 125 GeV as a function of the center-of-mass energy ($\sqrt{s}$) of the colliding protons at the LHC. Figure 1.5 (b) shows the branching ratios of the various decay channels of the Higgs boson as a function of its mass.

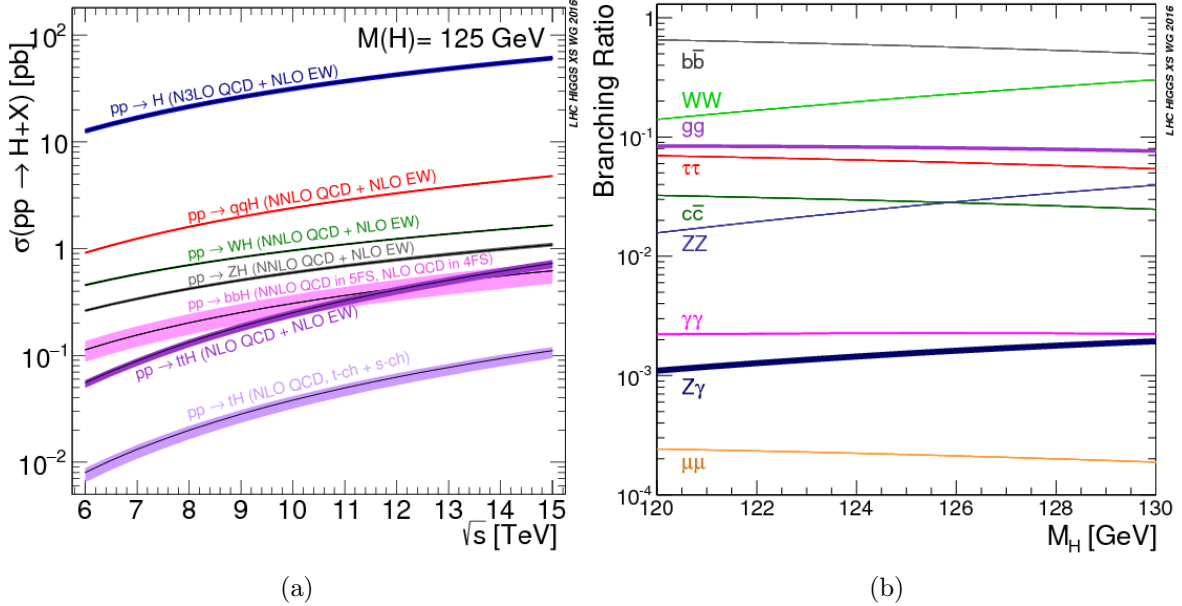


Figure 1.5: (a) Higgs boson production cross-section as a function of center-of-mass energy ($\sqrt{s}$) at the LHC for a Higgs mass of 125 GeV. (b) Branching ratios of various Higgs boson decay channels as a function of Higgs mass [6].

2

The ATLAS Detector at LHC

The Large Hadron Collider (LHC) [29, 30] operated by the European Organization for Nuclear Research (CERN) [57] is the most powerful particle accelerator and collider in the world. It consists of an underground tunnel about 100 meters in depth and a circumference of approximately 27 kilometers spanning across the Swiss and French borders near Geneva, Switzerland as illustrated in Figure 2.1. Protons and ions are accelerated inside this tunnel at speeds more than 99.9999% the speed of light. Only the data collected from proton-proton collisions is considered for the purpose of this analysis.

In this chapter, key LHC technologies (Section 2.1) and the ATLAS detector (Section 2.2), which is the experimental apparatus used in this analysis, is discussed.

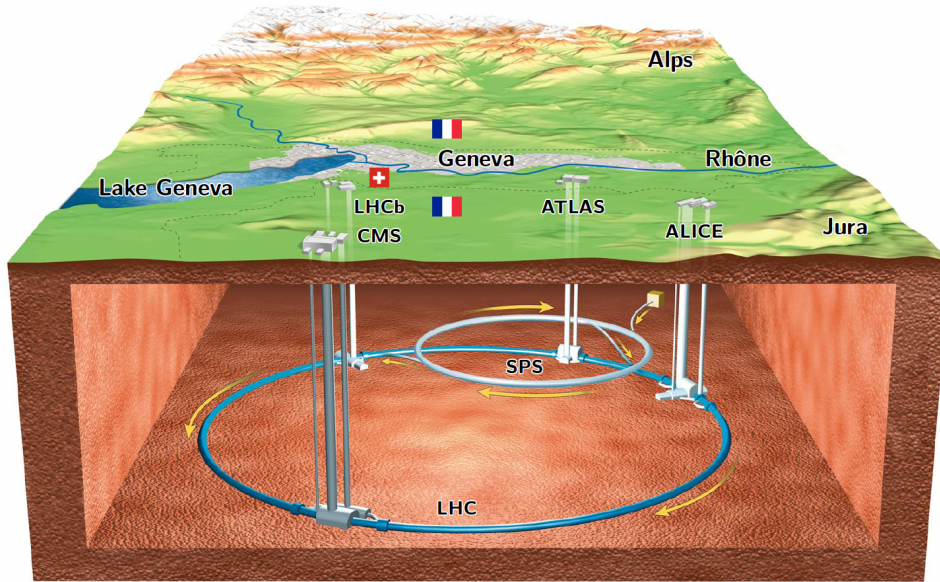


Figure 2.1: The LHC Machine and its four experiments displayed with their geographical locations.

2.1. The Large Hadron Collider

The LHC accelerator was designed to collide protons at a center of mass energy of $\sqrt{s} = 14$ TeV and a peak luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, and to collide heavy ions at 5.5 TeV per nucleon pair and luminosity of $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. It consists of two counter-rotating beams of particles, travelling in separate pipes, with a total of eight interaction regions as shown in Figure 2.2. Beam crossing occurs at four of these interaction regions which houses a detector

each: ATLAS[58], CMS[59], LHCb[60] and ALICE[61]. The beam crossings take place at the center of these detectors. No beam crossings occur at the remaining interaction regions, these are equipped with Radio-Frequency systems, collimation systems and the beam dumping system.

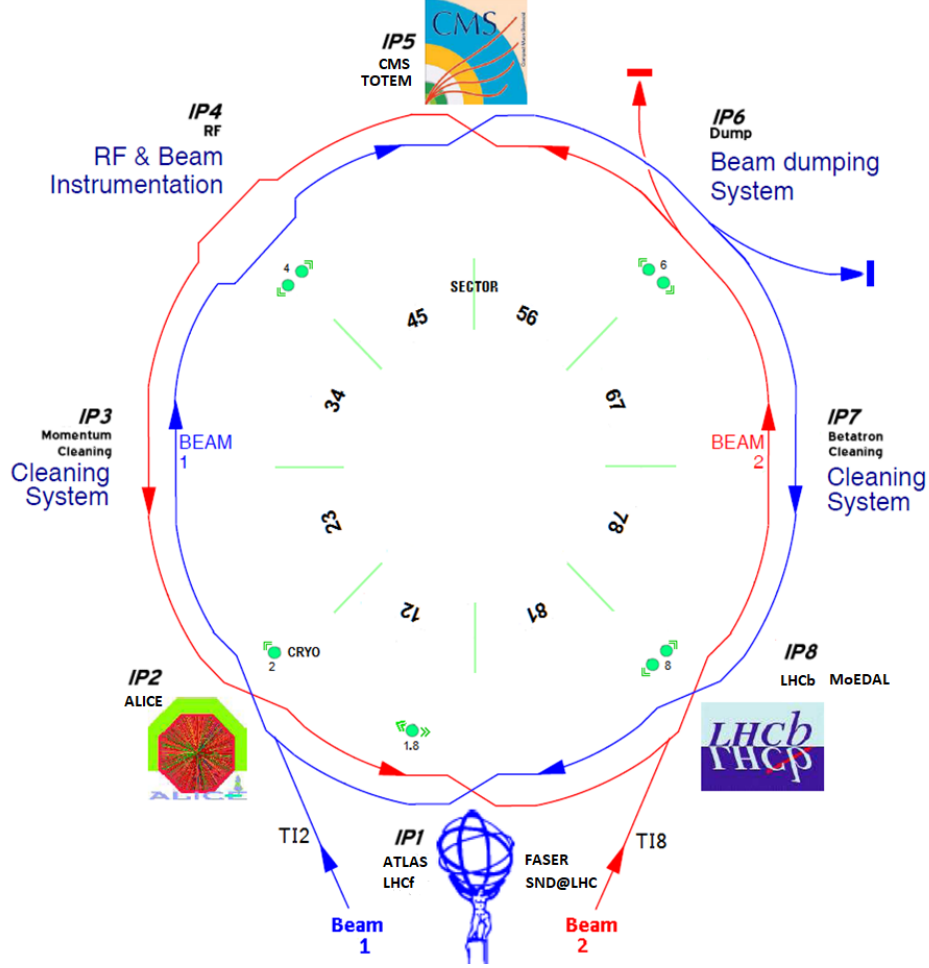


Figure 2.2: Schematic layout of the LHC interaction regions.

2.1.1 Proton injection chain

The path of protons inside the accelerator chain is shown in Figure 2.3. This journey begins at the LINAC4 [62] linear accelerator where hydrogen anions are boosted to an energy of 160 MeV. These ions are injected into the Proton Synchrotron Booster [63]. Unlike its predecessor LINAC2, LINAC4 can adjust beam parameters to match the Booster's acceptance to reduce losses. The hydrogen ions are stripped off of their electrons during injection to leave only protons. This booster accelerates protons to 2 GeV before injecting into the Proton Synchrotron(PS) [64]. The PS accelerates these further to 26 GeV and transfers them to the Super Proton Synchrotron(SPS) [65]. The SPS accelerates protons up to 450 GeV before injecting them into the LHC. The LHC operates at center of mass

energy of $\sqrt{s} = 13.6$ TeV during Run 3 (2022-2025) and the analysis described in this documentation uses data collected in 2022 and 2023.

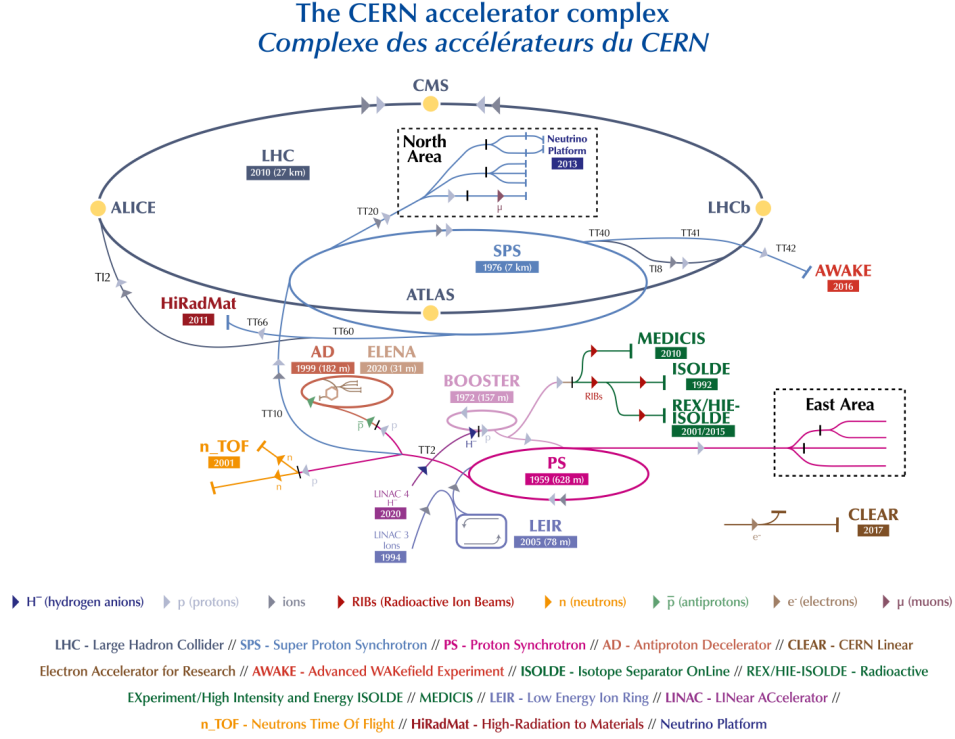


Figure 2.3: Schematic layout of the CERN accelerator complex. The injector chain for accelerating Protons is LINAC, the PS Booster (PSB), the Proton Synchrotron(PS) and the Super Proton Synchrotron (SPS) [7].

2.1.2 Luminosity

The two key parameters in any collider experiment are the center of mass energy E_{CM} and luminosity. E_{CM} is associated with energy available to produce heavy particles and it depends upon the type of collisions. For the same energy of the incoming particle, the E_{CM} is much higher for colliding particles compared to fixed target collisions, where one particle is at rest. Integrated luminosity refers to the total number of events during the entire data-taking period.

The rate (R) of a particular process with cross-section σ during the collisions at the LHC is given as:

$$R = \frac{dN}{dt} = \sigma_{\text{event}} \mathcal{L}_b \quad (2.1)$$

where N_{event} is the number of interactions, σ_{event} is the cross-section for a given event and $\mathcal{L}_b$ is the proportionality constant, also called as the instantaneous luminosity produced by a single colliding bunch (of protons). The unit of luminosity is $\text{cm}^{-2}\text{s}^{-1}$. For inelastic

collisions in the LHC, $\mathcal{L}_b$ can be written as:

$$\mathcal{L}_b = \frac{R}{\sigma_{\text{inel}}} = \frac{\mu f}{\sigma_{\text{inel}}} \quad (2.2)$$

where μ is the number of collisions per bunch crossing, f is the revolution frequency of the protons.

Thus, adding up all such bunch crossings gives us the value of total instantaneous luminosity $\mathcal{L}$ as:

$$\mathcal{L} = N_b \frac{\langle \mu \rangle f}{\sigma_{\text{inel}}} \quad (2.3)$$

where N_b is the number of bunches, $\langle \mu \rangle$ is the average number of interactions per crossing and f is the revolution frequency in the accelerator ring.

Multiple inelastic collisions occur with each bunch crossing. The collision that has the largest momentum transfer is known as the hard-scatter and the rest of them are called pile-up. Therefore, μ is also referred to as the pile-up parameter. The average number of interactions recorded per crossing during the Run 3 are shown in Figure 2.4. Pile-up can be further classified into two categories: in-time and out-of-time pileup [66]. When the additional collisions occur due to same bunch crossing, then it is called in-time pileup. If the additional collisions occur before or after the collision of interest, then they are referred to as out-of-time pileup.

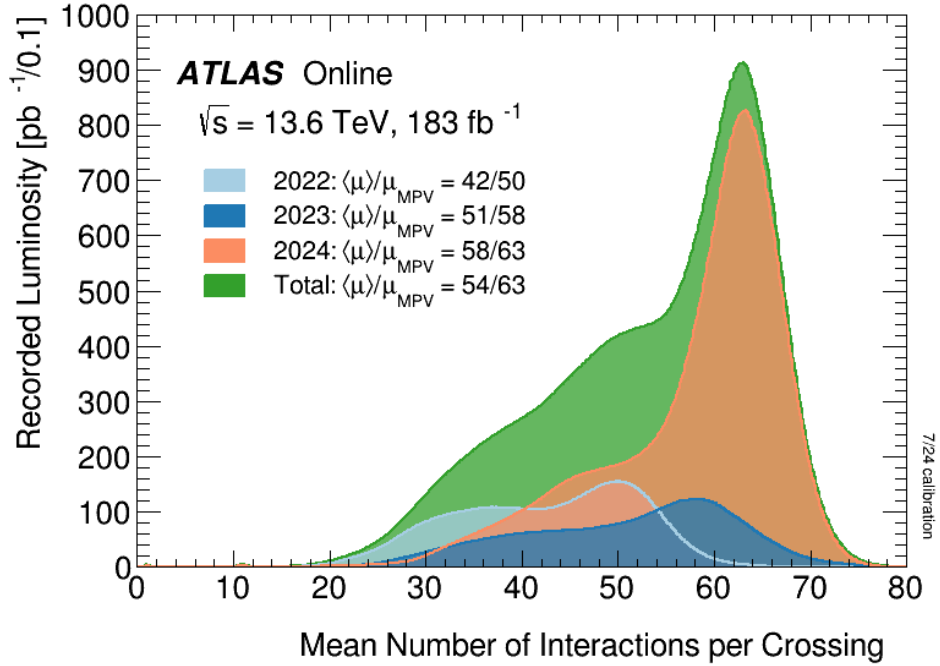


Figure 2.4: Luminosity-weighted, mean number of interactions per crossing collected during 2022-2024 Run 3 data taking at $\sqrt{s} = 13.6$ TeV is shown [8].

The value of the instantaneous luminosity also depends on beam parameters corresponding to the beam profile, crossing angle between the colliding beams, offsets when the beams don't collide head-on, etc. [67]. For beams with Gaussian density distribution, the luminosity $\mathcal{L}$ is given as:

$$\mathcal{L} = \frac{N_b N_1 N_2 f}{4\pi\beta^*\epsilon} F \quad (2.4)$$

where N_b is the number of bunches, N_1 and N_2 are number of protons per bunch, f is the revolution frequency in the accelerator ring, β^* is the beta function at the collision point, ϵ is the beam emittance and F is the reduction factor that accounts for crossing angle geometry [67].

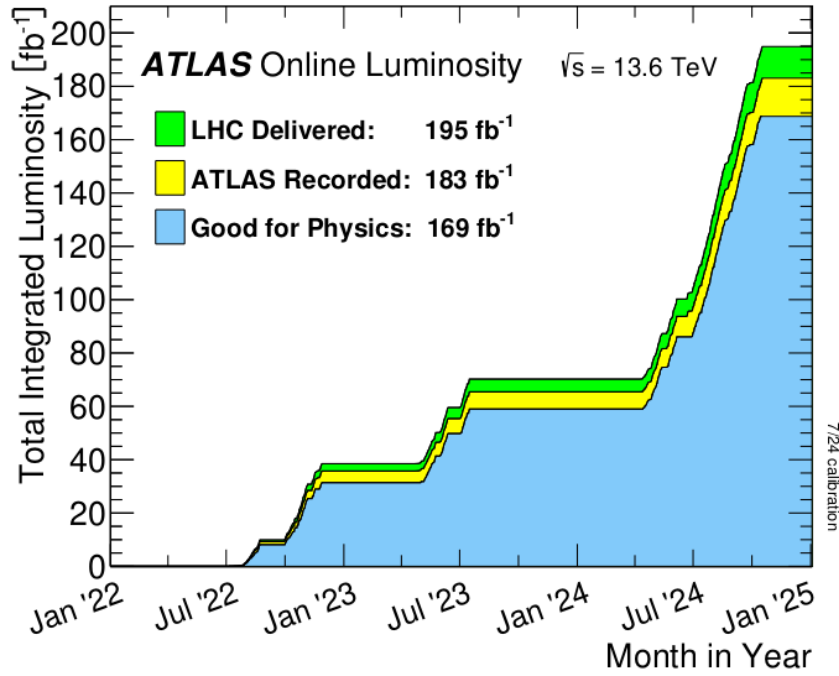


Figure 2.5: Total integrated luminosity for the time period of 2022-2024 is shown. Luminosity delivered by the LHC is shown in green, recorded by ATLAS is shown in yellow and the blue color represents the luminosity that is good to use for physics analyses [9].

Even though $\mathcal{L}$ is important, the beam is not always stable. What really matters is knowing the integrated luminosity (L). It is given by:

$$L = \int_0^T \mathcal{L}(t) dt \quad (2.5)$$

where T is the total run time, excluding the dead time. The integrated luminosity calculations consider all runs of the data-taking period, including LHC injection, acceleration and preparation of collisions. All ATLAS physics analysis ends up using what are called

the ‘good run lists’ (GRLs). GRL consists of runs with stable beam collisions, when data quality criterion were satisfied [68]. The total integrated luminosity recorded by the ATLAS and forms the GRL is shown in the Figure 2.5.

2.2. The ATLAS detector

The ATLAS (A Toroidal LHC ApparatuS) detector [37] is the largest volume particle detector ever constructed. It is 46 meters long, 25 meters high, 25 meters wide, weighs about 7000 tons and sits in a cavern 100 meters underground. It is one of the two general purpose detectors at the LHC. The other being CMS. The particle collisions happen at the center of the detector and the new particles produced after collision fly in all directions. These particles interact with the various sub-detector systems of the ATLAS detector. The detector has cylindrical symmetry and provide nearly 4π solid angle coverage around the interaction point.

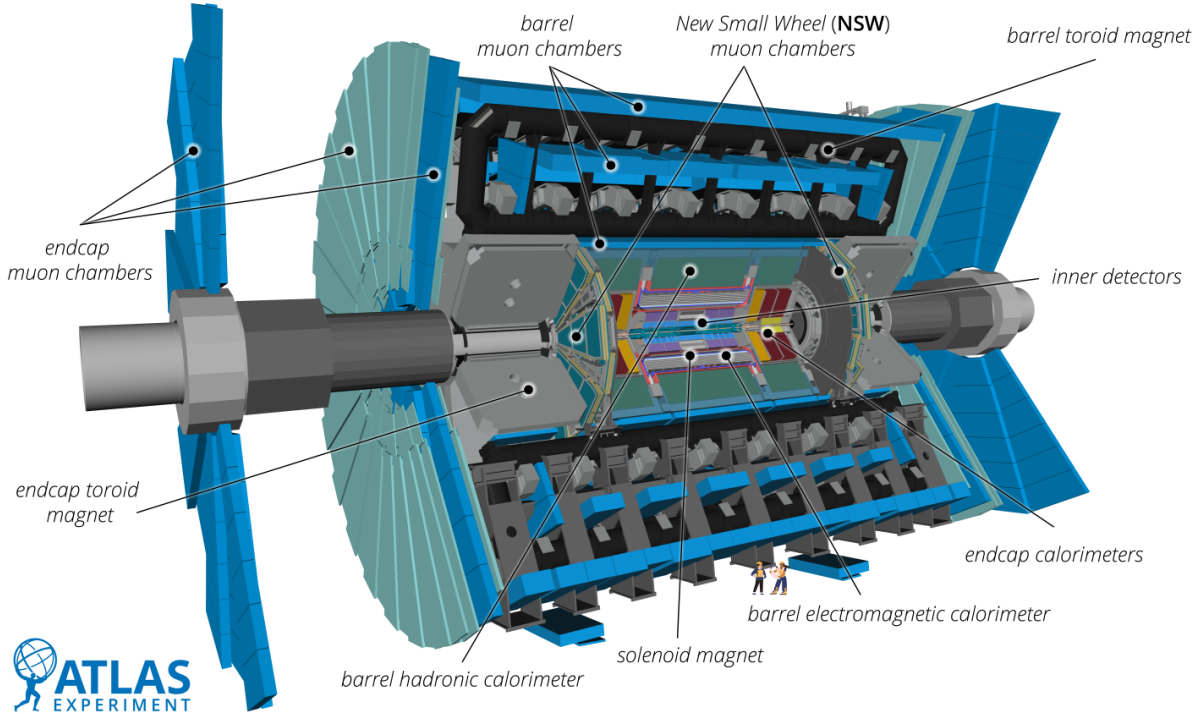


Figure 2.6: Schematic of the ATLAS detector subsystems for the Run 3 configuration [10].

The ATLAS detector was designed to study a wide range of physics processes like precision measurements of the SM, Higgs boson studies, searches for new particles or phenomenon BSM. To study these, the three main sub-detectors come into play: the inner detector (ID), the calorimeters and the muon spectrometer. The ID system is the closest to the beam pipe and is used to track charged particles. The calorimeter system is outside the ID and is used to measure the energy of particles. The muon spectrometer is

the outermost sub-detector and is used to identify and measure the momentum of muons. These sub-detectors are surrounded by a complex magnet system which is used to bend the trajectories of charged particles, allowing for the measurement of their momentum. Figure 2.6 shows a schematic of the ATLAS detector and its sub-systems for the Run 3 configuration.

These sub-systems are arranged in layers around the interaction point and are complemented by the trigger and data acquisition system (TDAQ) and the forward detectors. The TDAQ system is responsible for selecting interesting events from the vast amount of data produced by the LHC collisions, while the forward detectors are designed to measure particles produced at small angles relative to the beam axis. The sub-detectors, forward detectors and TDAQ system are described in the following sections. First, the coordinate system used by the ATLAS detector is described in Section 2.2.1, followed by the magnet system in Section 2.2.2. Next, the inner detector is described in Section 2.2.3, followed by the calorimeter system in Section 2.2.4, the muon spectrometer in Section 2.2.5, the forward detectors in Section 2.2.6 and finally the TDAQ system in Section 2.2.7.

2.2.1 The Coordinate System

The ATLAS detector uses a right-handed coordinate system. The origin of this coordinate system is at the interaction point. The positive x -axis is directed from the origin towards the center of the LHC ring. The positive y -axis is pointing upwards and the z -axis is along the beam direction. The side of the detector on the positive side of the z -axis is called ‘side A’ and the one on the negative side is called ‘side C’. The polar angle θ is defined with respect to the z -axis, and the azimuthal angle ϕ is defined in the x - y plane. Using pseudorapidity (η) instead of the polar angle θ to define particle geometry is more convenient as the detector has cylindrical symmetry. The pseudorapidity η is defined in terms of the polar angle θ as:

$$\eta = -\ln \left(\tan\left(\frac{\theta}{2}\right) \right) \quad (2.6)$$

For massive objects, rapidity (y) is used instead of pseudorapidity. For massless particles, it is equivalent to pseudorapidity. The rapidity y is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (2.7)$$

where E is the energy of the particle and p_z is the momentum of the particle along the z -axis. The differences in rapidity are Lorentz invariant under boosts along the beam axis.

In general, the particle geometry is described in the (r, ϕ, η) coordinate system, where r is the radial distance from the beam axis. The (η, ϕ) coordinate system is rotationally symmetric around the beam axis and the angular distance ΔR between two objects in this space is given by:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.8)$$

where $\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angle between the two objects, respectively.

The transverse momentum p_T , transverse energy E_T and transverse mass m_T of a particle are defined in the plane perpendicular to the beam axis. The transverse momentum p_T is given by:

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (2.9)$$

where p_x and p_y are the momentum components of the particle along the x and y axes, respectively. The transverse energy E_T is defined as:

$$E_T = \sqrt{m^2 + p_T^2} \quad (2.10)$$

where m is the mass of the particle. The transverse energy (E_T) is equivalent to the transverse mass m_T as:

$$m_T = \sqrt{E^2 - p_z^2} = \sqrt{m^2 + p_T^2} \quad (2.11)$$

The transverse quantities are particularly useful in hadron collider experiments like ATLAS, as they are invariant under boosts along the beam axis and provide a measure of the energy and momentum of particles in the plane perpendicular to the beam direction.

2.2.2 The Magnet System

In the ATLAS detector, the magnet system is useful to bend the trajectories of charged particles, allowing for the measurement of their momentum. The magnetic field is also useful to distinguish between charged and neutral particles. The ATLAS magnet system consists of four superconducting magnets: one solenoid and three toroidal magnets. The solenoid magnet surrounds the inner detector and provides a magnetic field of 2T in the barrel region. The toroidal magnets consists of a barrel toroid and two end-cap toroids. The barrel toroid provides a magnetic field of up to 4T and is useful to measure muon momentum. The end-cap toroids also provide a magnetic field up to 4T and extends the magnetic field to particles escaping the detector at shallow angles to the beam pipe. The magnetic field lines of the ATLAS detector are as shown in Figure 2.7.

The solenoid magnet is made up of niobium-titanium superconducting wires and is cooled to temperature of about 4.5 K. The solenoid magnet is 5.8 meters long and has a diameter of 2.56 meters. The toroidal magnets in the barrel region consists of eight coils and is also made up of niobium-titanium superconductor. The end-cap toroids consists of two coils on each side of the detector. The toroidal magnets operate at a temperature of 4.7 K.

2.2.3 The Inner Detector

The ID [69] is the first detector to interact with the particles produced during collision. Due to its vicinity from the beam axis, it receives a high influx of particles and thus

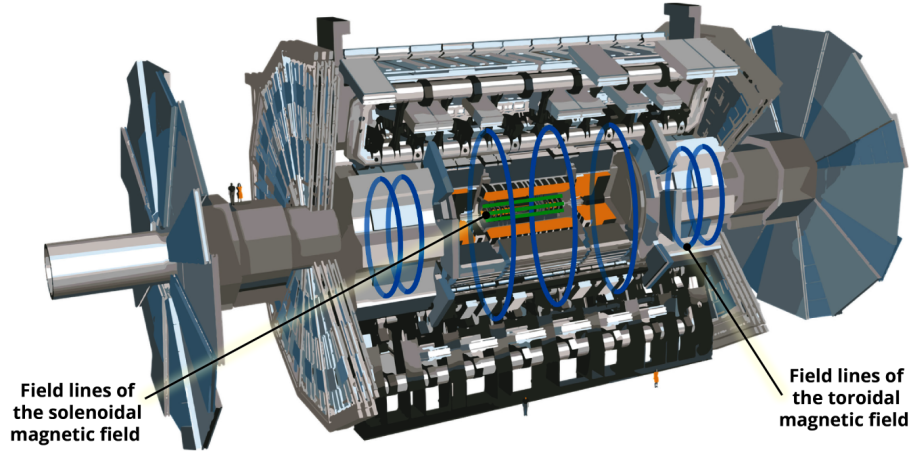


Figure 2.7: Illustration of the magnetic field lines of the ATLAS detector. The green lines correspond to the magnetic field by the central solenoid and the blue circles correspond to the magnetic field by the toroidal magnets [11].

requires high granularity of the sensors. The ID is useful for vertex identification, track reconstruction of charged particles and to identify electrons. As discussed earlier in section 2.2.2, the ID is surrounded by a 2T magnetic field in its barrel region which is useful to track the trajectories of charged particles.

The ID detector consists of three main sub-detectors: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). Each of these are layered in the barrel region with cylindrical symmetry, in the increasing order of their radii. Each of these sub-detector have three components: one barrel and two end-cap. In the barrel region, the pixel detector and the semiconductor tracker provides a coverage of $|\eta| < 2.5$. The TRT provides a coverage of $|\eta| < 2.0$. In the end-cap region, the pixel detector and the semiconductor tracker provides a coverage of $|\eta| < 2.5$. The layout of the inner detector consisting of the three sub-detectors with their barrel and end-cap components is shown in Figure 2.8.

Due to their high proximity to the beam axis, the pixel detector and SCT are made up of silicon semiconductor sensors that offer high resolution, but add a huge material cost. The TRT provides a cost-effective particle tracking and identification by using the straw drift-tube detector which is useful for particle reconstruction and electron identification. Together, these three sub-detectors provide high precision tracking and vertexing capabilities for charged particles. The pixel detector and SCT together form the ATLAS tracking system and while the TRT provides additional tracking information and electron identification capabilities. The ID covers a radial distance from 33.25 mm to 1082 mm from the beam axis and a length of 3512 mm along the beam axis as shown in Figure 2.9. These sub-detectors are described in the following sections.

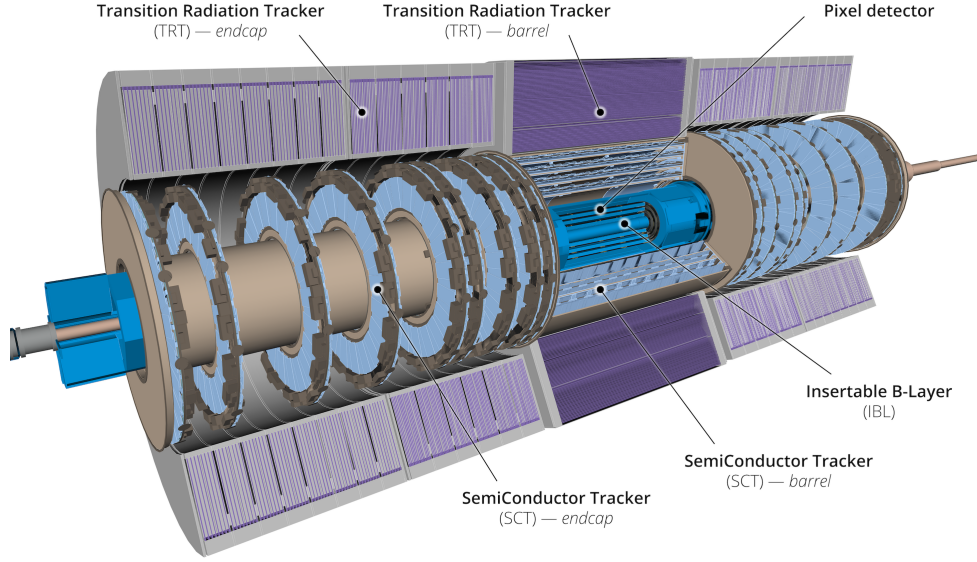


Figure 2.8: Layout of the ID and it's sub-detectors is shown. Each of the sub-detector has a barrel component and two endcap components [12].

Pixel Detector

The pixel detector [70] is the innermost detector of the ATLAS tracking system. It is useful in recognizing the primary and secondary vertices of the particle interaction, tracking and identification of charged particles, calculating the transverse impact parameter and tagging the b-jets.

It is built using silicon semiconductor technology. Each pixel sensor is connected to a read-out chip through bump-bonding. The pixel detector was originally designed to have three layers in the barrel region and three disks on each side of the end-cap region. These three layers are called B-layer, layer 1 and layer 2 in increasing order of their radii. It was designed to operate at peak luminosity of upto $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. For peak luminosity above that, the silicon sensors experience radiation related energy loss and thus have a deteriorating read-out efficiency for the pixel layers. During the long shut down 1 (LS1) detector upgrade, a fourth innermost layer called the insertable B-layer (IBL) [71] was added to enhance the performance of the pixel detector. The IBL is located at a distance of about 3.3 cm from the beam pipe as shown in Figure 2.9.

The pixel detector configuration for Run 3 consists of four concentric cylindrical layers in the barrel region and three end-cap disks located on both sides 'A' and 'C' of the detector, orthogonal to the beam axis. Together, it consists of 92 million pixel sensors. The IBL alone has about 12 million pixels. In the IBL region, the pixel sensor is $50 \mu\text{m} \times 250 \mu\text{m}$ and the rest of the pixels have a dimension of $50 \mu\text{m} \times 400 \mu\text{m}$. These pixel sensors provide high precision to identify the origin and momentum of the particle. It provides accuracies of $10 \mu\text{m}$ in the $r\phi$ space and $115 \mu\text{m}$ along the z -axis both in the barrel and disc regions [12].

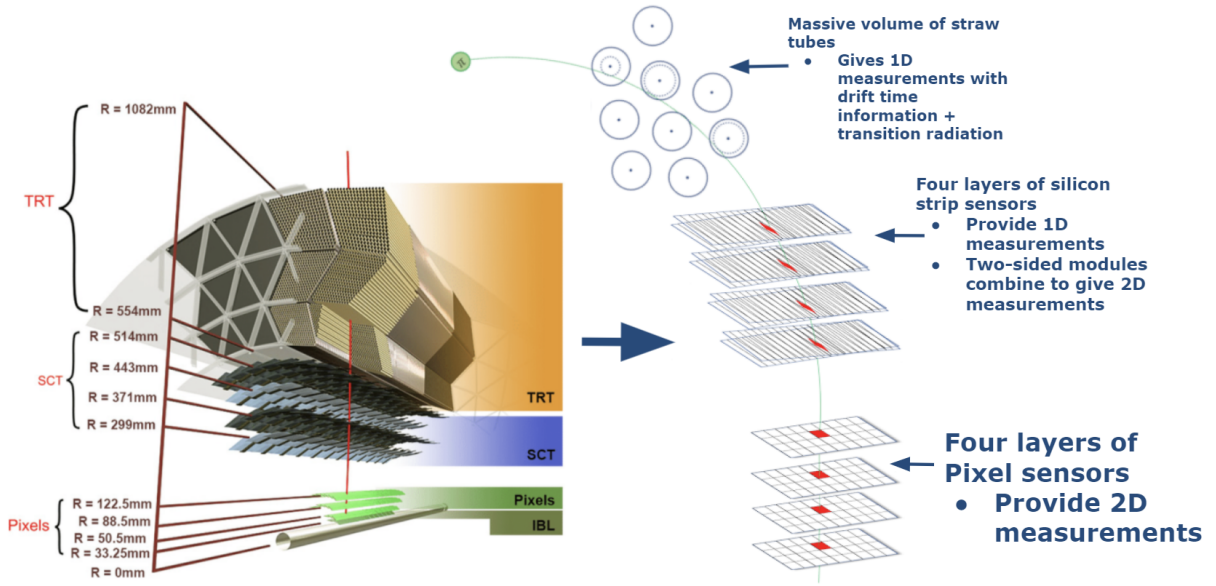


Figure 2.9: Transverse view of the ATLAS Inner Detector and illustration of track reconstruction by the sub-detectors [13].

The particles are tracked by the pixel detector by studying the interaction of charged particles with the semiconductor material which is doped. When a charged particle interacts with a doped silicon layer, it creates an electron-hole pair which produces an electrical signal in the presence of a bias voltage. This electrical signal is recorded as a hit and is useful to track the particle position. The read-out system consists of front-end electronics that amplify and digitize the signals from the pixel sensors and a data acquisition system that collects and processes the data. Each particle track provides four space points along the barrel layer and three along the end-cap disk depending upon the trajectory.

The Semiconductor Tracker

The SCT is the second layer of the ATLAS inner detector and envelops the pixel detector. It is made up of silicon strip detectors and uses the same technology as the pixel detector. Both the pixel detector and SCT are immersed in a 2T magnetic field which helps to track the charged particles. Apart from tracking charged particles, it is useful for impact parameter measurement, pattern recognition and vertex reconstruction. It provides a coverage of $|\eta| < 2.5$ both in the barrel and end-cap region.

It consists of four layers in the barrel region and nine layers in the end-cap region. Each layer consists of silicon strip modules. Each module consists of two pairs of silicon strip sensors glued back to back at an angle of 40 mrad to provide a 2D measurement of the particle position. In the barrel region, one set of strips in each layer is parallel to the beam-axis. In the endcap region, the strips are arranged radially and the mean pitch between the strips is $80\text{ }\mu\text{m}$ [72]. The SCT consists of a total of 4088 silicon strip modules

which altogether has about 6.3 million readout channels. It provides a resolution of $17\text{ }\mu\text{m}$ ($r\phi$) x $580\text{ }\mu\text{m}$ (z/r) in both the barrel and end-cap regions. Each track provides 4 space points in the barrel region and 9 space points in the end-cap region.

The Transition Radiation Tracker

The TRT [73] forms the outermost layer of the inner detector. It consists of about 300,000 straw tubes filled with a gaseous mixture. Each straw tube has a diameter of 4 mm and is made up of a conductive material. In the barrel region, these straws are parallel to the beam pipe and divided into two at the $\eta=0$ mark. In the end cap region, these are placed radially in the wheels. During Run 1 and 2, the straw tubes were filled with xenon-based gas mixture. Due to multiple gas leaks during these two runs, it was cost-effective to replace it with argon-based gas mixture in the entire barrel region and few wheels of the end-cap region [74].

The TRT uses the drift-time measurement technique to track charged particles. When a charged particle passes through the straw tubes, it ionises the gas mixture. The extra electrons produced during the ionization process are drifted towards the anode wire at the center of each tube. The drift time of these electrons is used to calculate the position of the incoming particle. The TRT provides a coverage of $|\eta| < 2.0$ both in the barrel and end-cap region. It provides a resolution of $130\text{ }\mu\text{m}$ in the $r\phi$ space both in the barrel and end-cap region. Each particle produces more than 30 hits in the TRT [12].

The TRT is also useful for particle identification. When a charged particle passes through the boundary with two different dielectric constants, it emits an electromagnetic radiation, also called the transition radiation. The energy of this radiation depends on the gamma factor γ of the particle. Since, $\gamma = \frac{E}{m}$ depends on the mass of the incoming particle, this is useful to distinguish between various particles. In particular, the TRT is useful for electron identification.

2.2.4 Calorimeters

The ATLAS calorimetry system surrounds the inner detector and provides a coverage of $|\eta| < 4.9$. These calorimeters aid in measuring the energy deposited by the incoming particles like electrons, photons and hadrons. It is also useful in determining their position, in jet reconstruction and in determining the missing transverse energy in an event. All ATLAS calorimeters are sampling calorimeters. These contain alternative layers of active medium that measures the energy deposition and an absorbing layer which tries to stop the particle. When a high energy particle enters the calorimeter, it interacts with the absorbing material and produces a shower of secondary particles. These secondary particles further interact with the absorbing material and produce more particles, resulting in a cascade effect, thus generating what we call a particle shower. Every particle shower is unique and thus helps in identifying the type of incoming particle.

There are two main types of calorimeters: Electromagnetic (EM) calorimeter and hadronic calorimeter. In general, the depth of the calorimeter is designed to contain

the particle shower such that it does not reach the muon system surrounding it. The *radiation length* (X_0) and the *interaction length* (λ) are used to determine the thickness of electromagnetic and hadronic calorimeter respectively. The *radiation length* is the average distance travelled by an electron before it loses $1/e$ of its initial energy. The *interaction length* is the average distance travelled by a hadron before it undergoes another inelastic collision with the nucleus of the interacting medium.

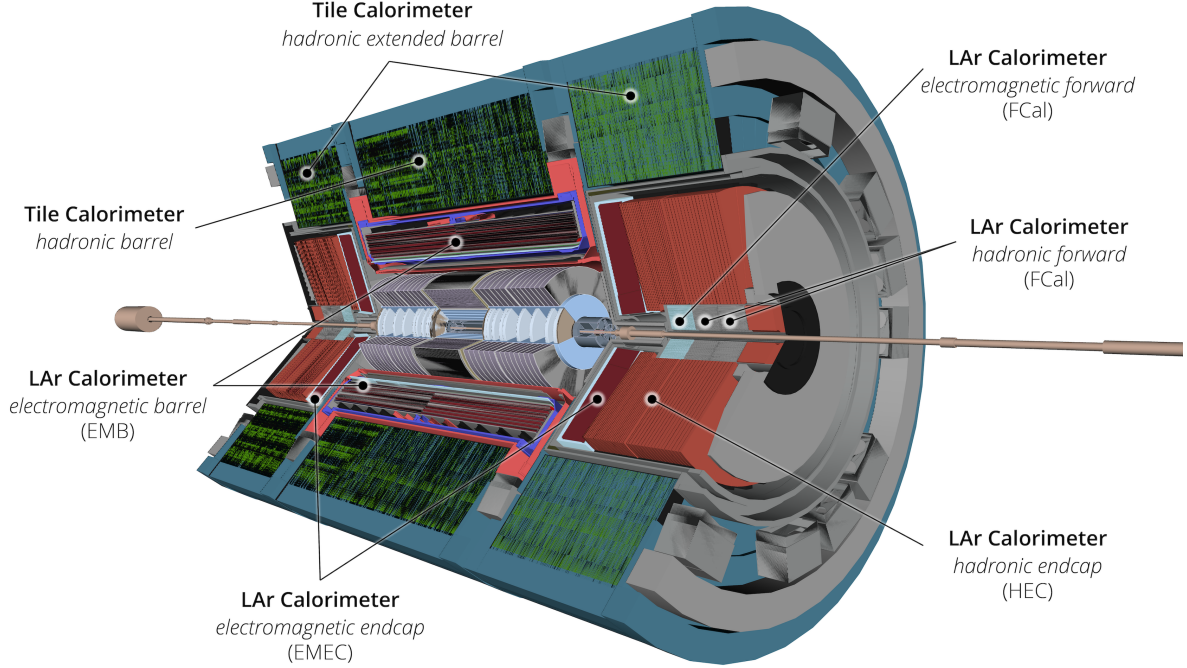


Figure 2.10: Illustration of the cut-away view of the ATLAS calorimeter system. The EM calorimeter uses LAr calorimeter in the barrel, endcap and forward region. The hadronic calorimeter uses the tile calorimeter in the barrel region and LAr everywhere else [12].

The layout of the EM and the hadronic calorimeters in the barrel region, end-cap region and forward region is shown in Figure 2.10. The EM calorimeter has a thickness of more than $22X_0$ in the barrel region and more than $24X_0$ in the endcap region. The hadronic calorimeter has an interaction length of about 9.7λ . Including the thickness of the support system, the total interaction length of hadronic calorimeter is about 11λ . These calorimeters are described in the following sections.

Electromagnetic Calorimeter

The EM calorimeter is a lead-liquid argon (LAr) sampling calorimeter with an accordion geometry. It consists of a barrel part covering the pseudorapidity region $|\eta| < 1.475$ and two end-cap parts covering the region $1.375 < |\eta| < 3.2$. The accordion geometry allows a full ϕ symmetry and fast readout which is crucial for the online event-selection system i.e.

the trigger system. The EM calorimeter is segmented into three layers in depth as shown in Figure 2.11a. The layers are named as front layer, middle layer and back layer in the order of their increasing depth from the interaction point [12].

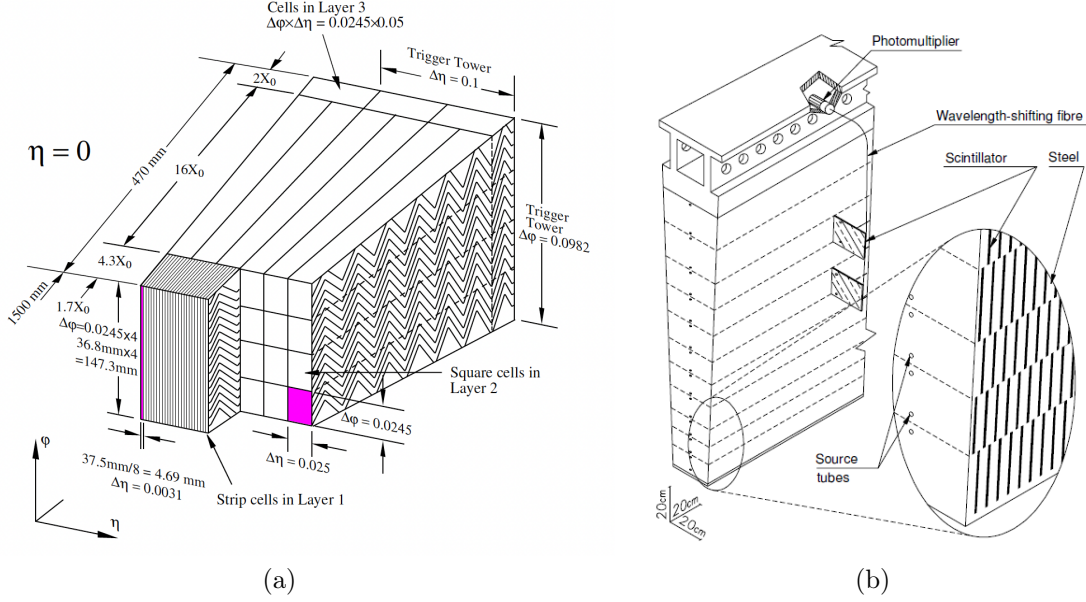


Figure 2.11: (a) Schematic of the longitudinal view of the ATLAS EM calorimeter showing the accordion geometry and the three layers in depth. (b) Schematic of the tile calorimeter showing the alternating layers of steel plates and scintillating tiles, the optical fibers and the photomultiplier tubes (PMTs) [12].

The first layer, also called the *front layer*, has a high granularity in the η direction (up to 0.003) and is used for precise measurement of the shower position and for π^0 rejection. The second layer, also called the *middle layer*, collects the bulk of the energy deposited by the shower and has a granularity of 0.025×0.025 in $\eta \times \phi$. The third layer, also called the *back layer*, is used to correct for leakage of energy from the calorimeter and has a granularity of 0.05×0.025 in $\eta \times \phi$.

In front of the EM calorimeter, there is a thin presampler layer covering the region $|\eta| < 1.8$ which is used to correct for energy loss in the material located upstream of the calorimeter. The total thickness of the EM calorimeter is more than $22X_0$ in the barrel region and more than $24X_0$ in the endcap region. The EM calorimeter is designed to measure the energy of electrons and photons with high precision. The energy resolution of the EM calorimeter is calculated using:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.12)$$

where E is the energy deposited in the calorimeter in GeV, the first term is the stochastic term, the second term is the constant term, and the third term is the noise term. The

calorimeter design values for the parameters are $a = 10\%$, $b = 170$ MeV without pile-up and $c = 0.7\%$ [75].

Hadronic Calorimeter

The hadronic calorimeter is designed to measure the energy of hadrons like pions, protons, neutrons and also jets (Chapter 3). It is a sampling calorimeter with two different technologies used in different regions of the detector. The hadronic calorimeter in the barrel region ($|\eta| < 1.7$) is a scintillating tile calorimeter with steel as the absorbing material. The hadronic calorimeter in the end-cap region ($1.5 < |\eta| < 3.2$) and the forward region ($3.1 < |\eta| < 4.9$) is a liquid argon (LAr) calorimeter with copper and tungsten as the absorbing materials respectively. The layout of the hadronic calorimeter is shown in Figure 2.10. The tile calorimeter consists of three cylindrical barrels: the central barrel covering the region $|\eta| < 1.0$ and two extended barrels covering the region $0.8 < |\eta| < 1.7$. Each barrel is segmented into 64 modules in the ϕ direction. Each module consists of alternating layers of steel plates and scintillating tiles. The scintillation light produced in the tiles is collected by wavelength-shifting fibers and read out by photomultiplier tubes (PMTs). The layout of the tile calorimeter is shown in Figure 2.11b.

The LAr hadronic end-cap calorimeter (HEC) consists of two wheels on each side of the detector. Each wheel is segmented into eight modules in the ϕ direction. Each module consists of alternating layers of copper plates and LAr gaps. The LAr hadronic forward calorimeter (FCal) consists of three modules on each side of the detector. Each module consists of a matrix of copper or tungsten rods embedded in LAr. The FCal is designed to withstand the high radiation and particle flux in the forward region.

The total thickness of the hadronic calorimeter is about 9.7λ . Including the thickness of the support system, the total interaction length of hadronic calorimeter is about 11λ . The energy resolution of the hadronic calorimeter is given by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus c \quad (2.13)$$

where E is the energy in GeV deposited by the incoming particle in the calorimeter, the first term is the stochastic term and the second term is the constant term. The noise term (b) from Eq. 2.12 becomes negligible at high energies. The parameter values are $a = 45\%$ after software compensation and $c = 1.6\%$ [76].

2.2.5 Muon Spectrometer

The muon spectrometer [12] is the outermost layer of the ATLAS detector and is designed to detect and measure the momentum of muons. Muons are charged leptons that can penetrate through the inner detector and calorimeters due to their high mass and low interaction cross-section. The muon spectrometer covers the pseudorapidity region $|\eta| < 2.7$ and is composed of several types of detectors arranged in layers around the beam axis. The muon spectrometer was designed to provide a momentum resolution of about 10%

for muons with transverse momentum of 1 TeV. The combination of the different detector technologies in the muon spectrometer allows for efficient muon identification and precise momentum measurement.

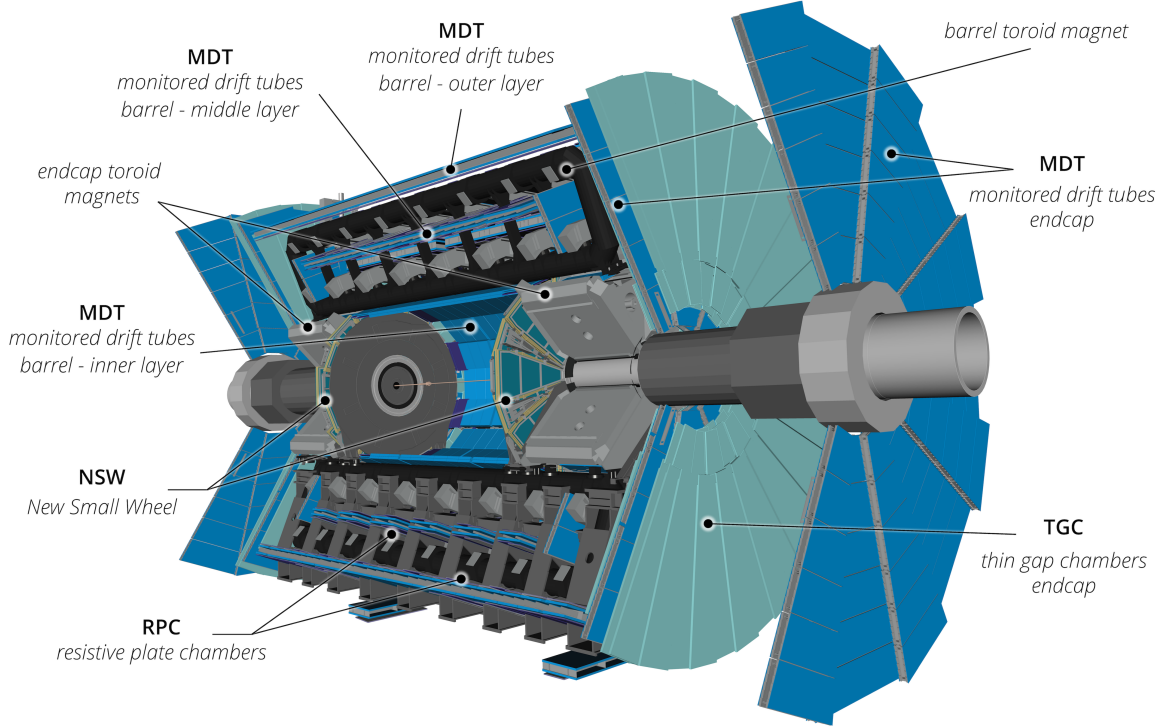


Figure 2.12: Schematic of the ATLAS muon spectrometer showing the arrangement of the different detector technologies [12].

The main components of the muon spectrometer include the Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs), Resistive Plate Chambers (RPCs), and Thin Gap Chambers (TGCs). The endcap region of the muon spectrometer consisted of CSCs in the innermost layer during Run 1 and Run 2. These were replaced with the New Small Wheel (NSW) during the LHC long shutdown 2 (LS2) to improve the performance in the high-luminosity environment of the LHC. Layout of the various components of the muon spectrometer is shown in Figure 2.12.

Monitored Drift Tubes (MDTs)

The MDTs are the primary precision tracking detectors in the muon spectrometer and are used to measure the position of muon tracks with high accuracy. They consist of cylindrical tubes filled with a gas mixture and a central wire that collects the ionization produced by passing muons. The MDTs are arranged in layers around the beam axis and provide a spatial resolution of about 80 micrometers.

Resistive Plate Chambers (RPCs)

The RPCs are fast detectors used for triggering and timing measurements in the barrel region of the muon spectrometer. They consist of two parallel resistive plates with a gas gap in between, and they produce a fast electrical signal when a muon passes through. The RPCs have a time resolution of about 1 nanosecond and are arranged in layers to provide coverage in the barrel region.

Thin Gap Chambers (TGCs)

The TGCs are used in the end-cap region of the muon spectrometer for triggering and timing measurements. They are multi-wire proportional chambers with a thin gas gap and provide fast signals for muon detection. The TGCs have a time resolution of about 2 nanoseconds and are arranged in layers to provide coverage in the end-cap region.

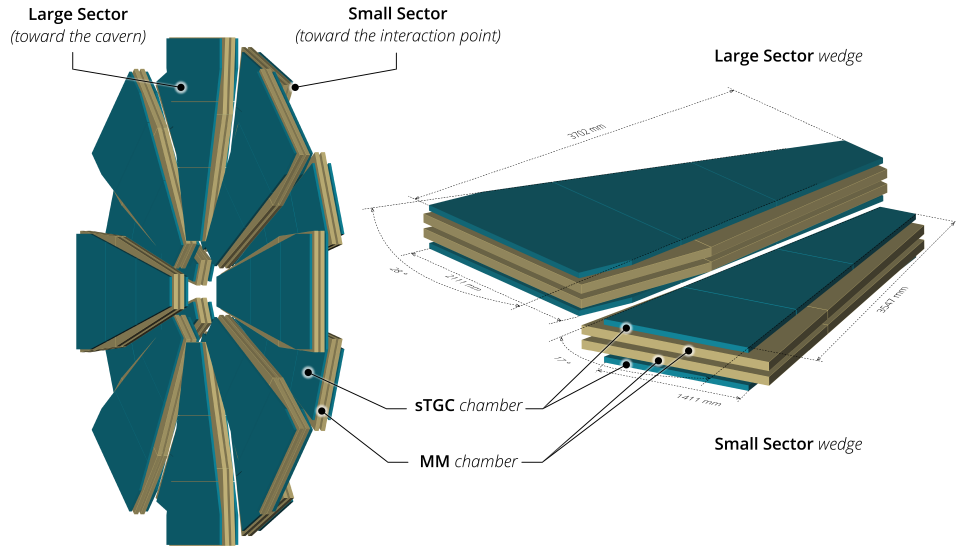


Figure 2.13: Schematic of the New Small Wheel (NSW) showing the arrangement of the Micromegas and sTGC detectors [12].

New Small Wheel (NSW) Upgrade

The New Small Wheel (NSW) is designed to improve the performance of the muon spectrometer in the high-luminosity environment of the LHC. It consists of two types of detectors: Micromegas and small-strip Thin Gap Chambers (sTGCs). The Micromegas detectors provide high spatial resolution, while the sTGCs offer fast timing capabilities. It is located at a distance of about 7.4 meters from the interaction point and covers the pseudorapidity region $1.3 < |\eta| < 2.7$. The NSW upgrade also includes improvements to the readout electronics and data acquisition system to handle the increased data rates in the high-luminosity environment. The design of the New Small Wheel is shown in Figure 2.13.

2.2.6 Forward Detectors

The ATLAS forward detectors are designed to measure particles produced at small angles relative to the beam axis. They provide unique insights into the event topology and help in identifying rare processes that may not be easily detectable by the central detectors alone.

There are four forward detectors in ATLAS [12]: the LUCID (LUminosity Cerenkov Integrating Detector), the ALFA (Absolute Luminosity For ATLAS) detectors, the AFP (ATLAS Forward Proton) detectors, and the ZDC (Zero Degree Calorimeter). The AFP was installed during the LHC long shutdown 1 (LS1) and is operational since Run 2. The layout of the forward detectors is shown in Figure 2.14. These detectors are described in the following sections.

LUCID

The LUCID detector is located at a distance of about 17 meters from the interaction point (IP) on both sides of the ATLAS detector. It is designed to measure the luminosity of the LHC by detecting charged particles produced in the forward region. The LUCID detector consists of two arrays of Cerenkov tubes filled with a gas mixture, which produce light when charged particles pass through them. The light is then detected by photomultiplier tubes (PMTs) and used to calculate the luminosity.

ZDC

The ZDC is located at a distance of about 140 meters from the IP on both sides of the ATLAS detector. It is designed to measure neutral particles, such as neutrons and photons, produced in the forward region. The ZDC consists of a series of tungsten and quartz layers that produce Cerenkov light when neutral particles pass through them. The light is then detected by PMTs and used to measure the energy and position of the neutral particles.

AFP

The AFP detectors are located at a distance of about 210 meters from the IP on both sides of the ATLAS detector. They are designed to detect protons that remain intact after a collision and are scattered at small angles. The AFP detectors consist of silicon pixel detectors for precise position measurements and Cerenkov detectors for timing measurements. The AFP detectors are used to study diffractive events and to search for new physics phenomena.

ALFA

The ALFA detectors are located at a distance of about 240 meters from the IP on both sides of the ATLAS detector. They are designed to measure the absolute luminosity of the LHC by detecting protons scattered at small angles. The ALFA detectors consist of Roman pots, which are movable devices that can be inserted close to the beamline to

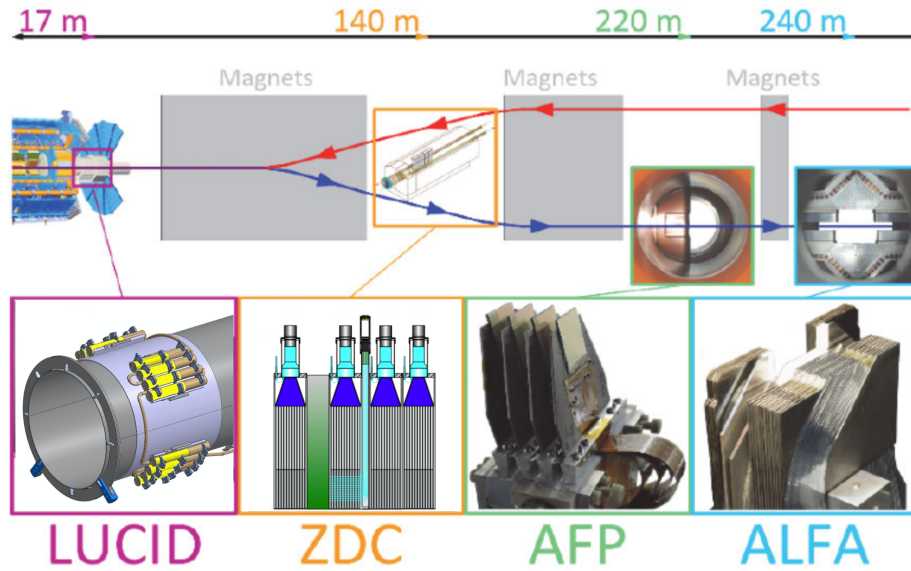


Figure 2.14: Schematic of the ATLAS forward detectors showing the LUCID, ALFA, AFP, and ZDC detectors [12].

detect scattered protons. Each ALFA detector contains scintillating fiber trackers that provide precise position measurements of the scattered protons.

2.2.7 Trigger and Data Acquisition System

The ATLAS trigger and data acquisition (TDAQ) system [77] is designed to select “interesting” events from the vast amount of data produced by the LHC collisions. Any event that contains information about important physics processes or signatures of rare events are deemed as “interesting”. The events are selected through a predefined selection which depends on the goals of the physics analyses, kinematics of the event objects like p_T of an electron, etc. The TDAQ system consists of two main levels: the Level-1 (L1) trigger and the High-Level Trigger (HLT) as shown in Figure 2.15. The L1 trigger is a hardware-based system that makes a decision within 2.5 microseconds based on information from the calorimeters and muon detectors. The L1 trigger reduces the event rate from 40 MHz to about 100 kHz. The HLT is a software-based system that runs on a large computing farm and makes a decision within 200 milliseconds based on the full event. The HLT further reduces the event rate to about 1 kHz, which is then recorded for offline analysis.

The TDAQ system is designed to be flexible and adaptable to different running conditions and physics goals. The trigger menu, which defines the selection criteria for the L1 and HLT, can be modified based on the luminosity, pile-up, and physics priorities. The TDAQ system also includes a data quality monitoring (DQM) system that continuously monitors the performance of the detectors and the trigger system. The DQM system provides real-time feedback to the operators and allows for quick identification and correction

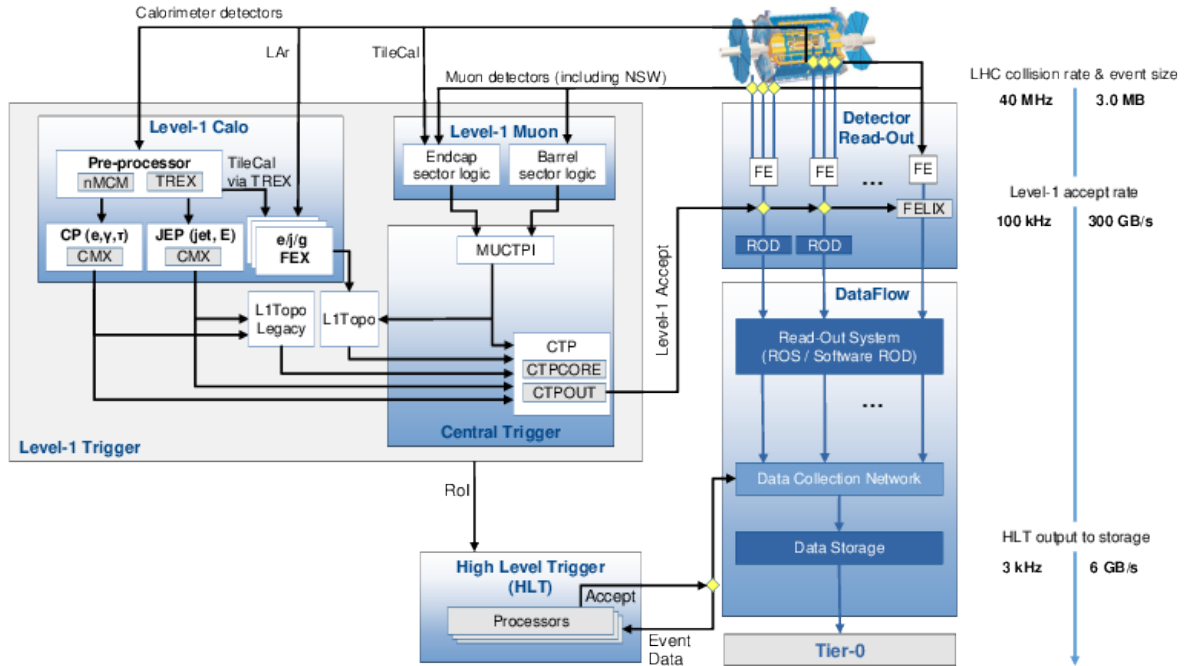


Figure 2.15: Schematic of the ATLAS Trigger and Data Acquisition (TDAQ) system showing the Level-1 (L1) trigger, High-Level Trigger (HLT), and data flow [12].

of any issues.

The upgrades to the TDAQ system [12] during the LHC Run 2 and Run 3 include the installation of new hardware and software components to improve the performance and flexibility of the system. The L1 trigger has been upgraded with new electronics and algorithms to handle the increased data rates and pile-up in the high-luminosity environment of the LHC. The HLT has been upgraded with new computing resources and software frameworks to improve the processing speed and efficiency. These upgrades help to maintain the performance of the TDAQ system in the high-luminosity environment of the LHC and improve the overall physics reach of the ATLAS experiment.

3 | Object Reconstruction and Selection

In high-energy physics experiments like ATLAS, the particles produced in proton-proton collisions are not directly observed. Instead, they interact with the detector components, leaving signals that are reconstructed into physics objects such as electrons, muons, jets, photons, and missing transverse energy (MET). Figure 3.1 illustrates the various physics objects and their interactions with different parts of the ATLAS detector. This chapter outlines the reconstruction and selection criteria for the objects used in this analysis i.e., electrons, muons, jets, and MET. Primary vertex and track reconstruction are also discussed as they are fundamental to accurate object reconstruction. Although tau leptons are considered in the analysis, their reconstruction and selection details are not included in this chapter as their leptonic decay products (electrons and muons) would be covered under the respective sections. Finally, the overlap removal procedure is described to avoid double counting of objects.

3.1. Tracks

Tracks correspond to the trajectories of charged particles as they pass through the inner detector (ID) of ATLAS. These charged particles leave deposits of energy which are clustered as hits in the pixel detector, semiconductor tracker (SCT), and transition radiation tracker (TRT). Particle tracks are reconstructed by connecting these hits using sophisticated algorithms. The tracks are characterized by various parameters with respect to the perigee, which is the point of closest approach to the beamline. The parameters include d_0 and z_0 which represent the transverse and longitudinal impact parameters, respectively, ϕ and θ which denote the azimuthal and polar angles, respectively, and q/p which is the charge-to-momentum ratio of the particle [78]. These parameters are illustrated in Figure 3.2.

The track reconstruction algorithm starts with seed hits in the pixel and SCT detectors, which are then extended to form track candidates. A Kalman filter [79] is used to refine the track parameters by iteratively adding hits and updating the track fit. Fake tracks are suppressed using an ambiguity solving stage which uses a scoring system based on the track quality [80]. A final fit is performed using a global χ^2 minimization to obtain the best estimate of the track parameters [78].

3.2. Primary Vertex

The primary vertex (PV) represents the point of interest where the proton-proton collision occurs. This is identified as a point in space where at least two tracks with $p_T > 500$ MeV

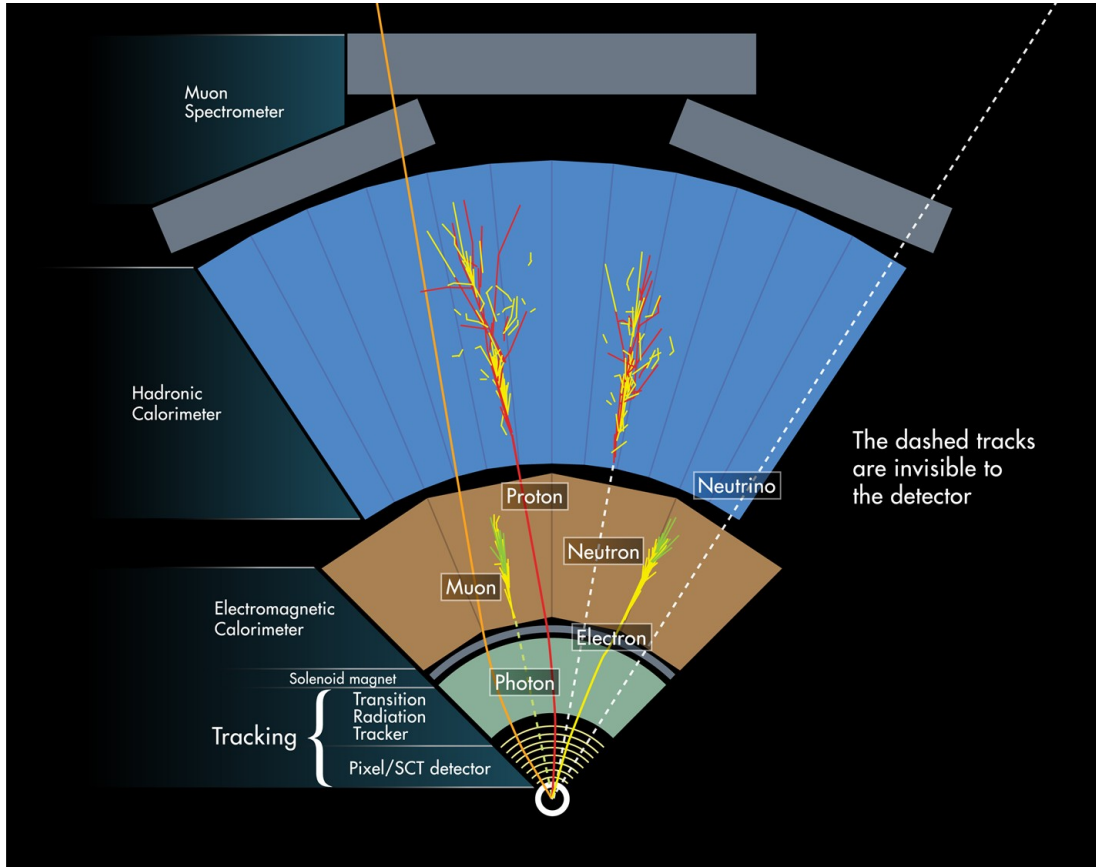


Figure 3.1: Schematic representation of the particle interactions with different components of the ATLAS detector [14]. The dashed lines correspond to no interaction with that detector component.

intersect. When multiple vertices are reconstructed in an event, the PV is selected as the vertex with the highest $\sum p_T^2$ of its associated tracks. This selection criterion is based on the assumption that the hard-scatter interaction, which produces high- p_T particles, will have a higher $\sum p_T^2$ compared to pile-up interactions. Several vertices are possible when reconstructing tracks from multiple interactions in the same bunch crossing.

In Run 3, the reconstruction of offline primary vertices is performed using the ATLAS Common Tracking (ACT) framework [81]. To enhance the efficiency of identifying primary vertices, especially in high pile-up scenarios, two advanced methods are employed: the Gaussian track density seed finder (GS) and the adaptive multi-vertex finder (AMVF) [82]. The GS method locates potential vertex positions by detecting peaks in the distribution of track densities along the beamline. Subsequently, the AMVF method simultaneously fits multiple vertices and assigns tracks to the vertex they are most likely associated with, based on their compatibility. This combined approach significantly improves the precision of vertex reconstruction under challenging conditions.

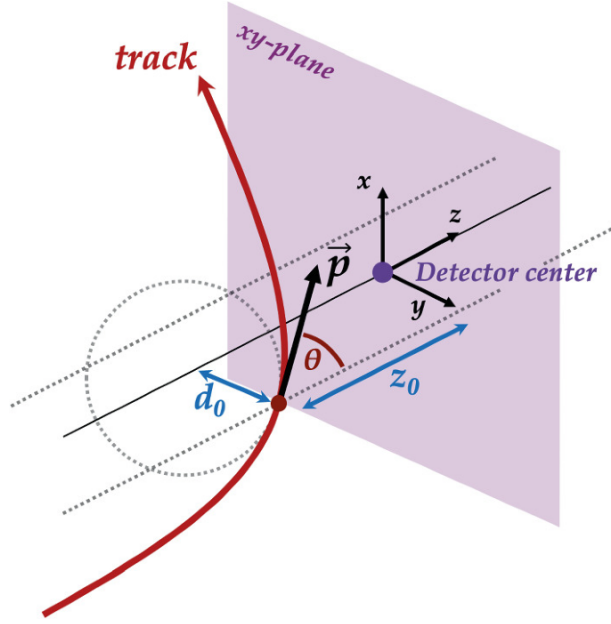


Figure 3.2: Schematic representation of track parameters with respect to the perigee [15].

3.3. Electrons

Electron reconstruction begins with the identification of energy deposits (clusters) in the electromagnetic (EM) calorimeter. These clusters are called TopoClusters, which are formed by grouping neighboring calorimeter cells with significant energy deposits. These clusters are then matched to tracks reconstructed in the inner detector (ID) to form electron candidates. The matching of tracks to clusters is used to build superclusters, which account for energy losses due to bremsstrahlung radiation [83]. The supercluster algorithms [84] replace the traditional cluster algorithms used in Run 1 and Run 2, which were based on fixed-size sliding windows [85] that has fixed sizes of clusters. On the other hand, superclusters dynamically adjust their size and shape based on the energy distribution of the electron shower, allowing for better energy resolution and improved identification of electrons, especially in dense environments. Further selection criteria are applied to “identify” and “isolate” electrons from background processes as described below.

3.3.1 Identification

The identification of electrons involves distinguishing true electrons from background sources such as jets and photon conversions. This is achieved using a multivariate analysis (MVA) approach that combines various discriminating variables related to the shower shape in the EM calorimeter, track quality, and track-cluster matching. The MVA is trained using simulated samples of electrons and background processes to optimize the separation power. Different working points (WPs) are defined based on the MVA output, such as “Loose”, “Medium”, and “Tight”, which correspond to different levels of background rejection and signal efficiency [83]. Figure 3.3 (a) shows the electron identification efficiency

as a function of transverse momentum (p_T) and Figure 3.3 (b) shows the pseudorapidity (η) for different identification WPs.

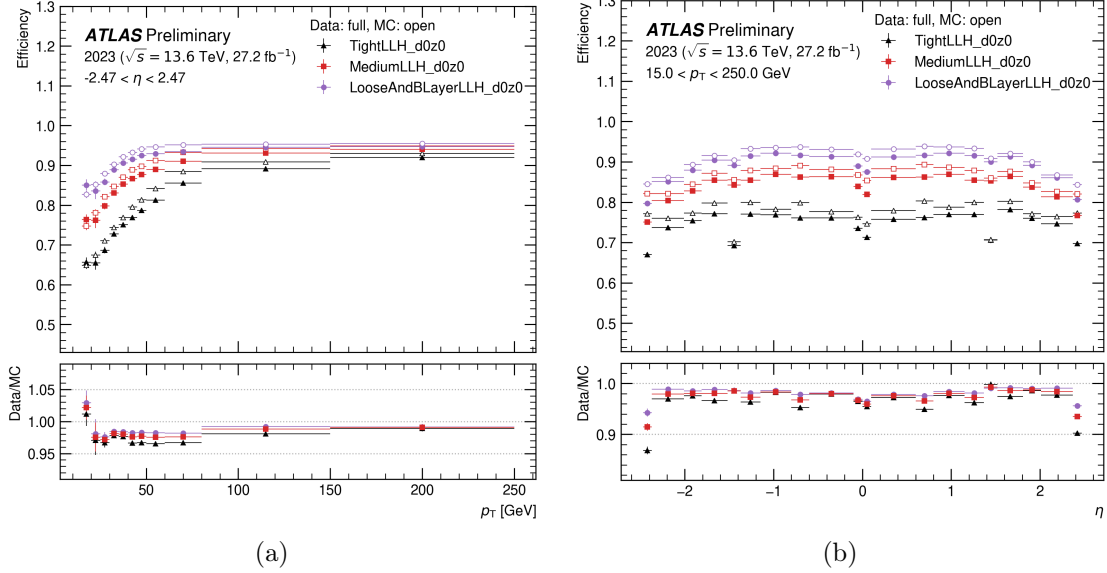


Figure 3.3: Electron identification efficiency as a function of transverse momentum (p_T) (left) and pseudorapidity (η) (right) for different identification working points (WPs) [16].

3.3.2 Isolation

The isolation of electrons is crucial to reduce background from non-prompt electrons originating from heavy-flavor decays or misidentified jets. Isolation criteria are based on the amount of additional energy or track activity in a cone around the electron candidate. Isolation working points (WPs) such as “Loose”, “Tight”, and “Gradient” are defined based on the sum of transverse energy in the calorimeter and the sum of transverse momentum of tracks within a cone of a specified radius around the electron [83].

3.4. Muons

Muon reconstruction begins with segment finding in the muon spectrometer (MS) using hits from the various muon chambers. These segments are then combined with tracks reconstructed in the inner detector (ID) to form muon candidates. Any misalignments of the muon chambers affect the accuracy of muon momentum measurements, this is handled by applying Gaussian constraints on the ensemble of hits used in the track fit [86]. An event model object called as Alignment Effect on Track (AEOT) is used to implement these constraints. A global χ^2 track fit including the AEOTs is performed to minimize the impact of misalignments on the muon momentum resolution [78].

Several reconstruction algorithms are used to reconstruct muons in ATLAS. “Combined” (CB) muons combine tracks from both the ID and MS using an outside-in approach, where the MS track is extrapolated inwards to match with the ID track. An alternative inside-out approach is used in “Segment-Tagged” (ST) muons, where ID tracks are extrapolated outwards to match with segments in the MS. “Calorimeter-Tagged” (CT) muons use energy deposits in the calorimeter matched to ID tracks. Finally, “Extrapolated” (ME) muons are reconstructed using only the MS information and extrapolating the track back to the interaction point [87]. Further enhancements in muon reconstruction have been made with the introduction of the new small wheel (NSW) detectors in Run 3. Figure 3.4 shows high efficiency of muon reconstruction with a requirement of at least four active layers in the NSW detectors [17].

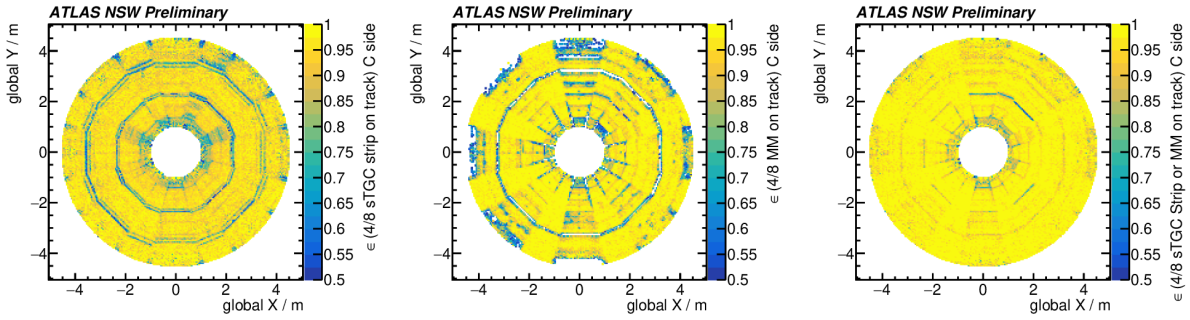


Figure 3.4: Efficiency of combined muon reconstruction requiring a minimum of four active layers in the Micromegas (MM) and small-strip Thin Gap Chambers (sTGC) detectors, or standalone muons traversing the NSW on the C side of the ATLAS detector. Reduced efficiencies are attributed to detector or readout challenges encountered during the data-taking [17].

3.4.1 Identification

Muon identification aims to distinguish true muons from background sources such as pion and kaon decays or in-flight decays of charged hadrons. The in-flight decays can be easily suppressed as its distinctive topology fits poorly with the muon track hypothesis. Depending upon the needs of the analysis, different identification working points (WPs) such as “Loose”, “Medium”, and “Tight” are defined based on the quality of the track fit, the number of hits in the ID and MS, and the compatibility between the ID and MS tracks [88]. Figure 3.5 shows the muon identification efficiency as a function of transverse momentum (p_T), pseudorapidity (η), and azimuthal angle (ϕ) for different identification WPs for the 2023 data-taking period.

3.4.2 Isolation

Muon isolation is essential to measure the activity around the muon candidate. This helps to reduce background contributions based on the analysis needs. To quantify the isolation,

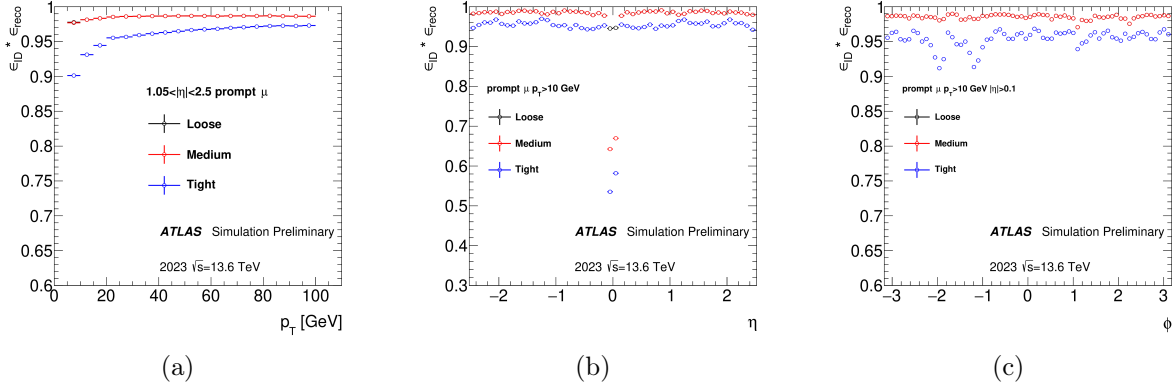


Figure 3.5: Muon reconstruction and identification efficiency for the three WPs during the 2023 data-taking period as a function of (a) transverse momentum (p_T), (b) pseudorapidity (η), and (c) azimuthal angle (ϕ). The samples used in these measurements are from $t\bar{t}$ events and produced using the POWHEG+PYTHIA8 generator [18].

variables based on calorimeter and track information are used. Muon isolation is quantified using two primary variables. The track-based isolation variable ($p_T^{\text{varcone30}}$) calculates the total transverse momentum of tracks within a cone of size $\Delta R = \min\left(\frac{10 \text{ GeV}}{p_T^\mu}, 0.3\right)$ around the muon, excluding the track of the muon itself. The calorimeter-based isolation variable ($E_T^{\text{topocone20}}$) measures the sum of transverse energy in topological clusters within a cone of radius $\Delta R = 0.2$ around the muon, after removing the energy contribution from the muon. Additionally, ratios of these isolation variables are employed to effectively distinguish prompt muons from background sources [88].

3.5. Jets

Although jets are not part of the final state in this analysis, they play a vital role in determining the missing transverse momentum (MET) and in the overlap removal procedure. Therefore, a brief overview of jet reconstruction, calibration, jet vertex tagging (JVT), and flavor tagging is provided in this section. The jet reconstruction process uses particle flow (PF) algorithm which combines information from the topological clusters in the calorimeter and tracks from the inner detector (ID) [19]. The energy of topological clusters that do not have matching tracks are removed. The algorithm determines the probability of a given PF object belongs to one or more clusters. Figure 3.6 shows the track selection, cluster association, and energy subtraction steps of the particle flow algorithm.

The PF objects serve as inputs to the jet reconstruction algorithms. Various jet algorithms are available based on their sensitivity to non-perturbative effects such as hadronization and underlying event activity, pile-up conditions, and detector response [89]. In this analysis, anti- k_t algorithm [90] with a radius parameter of $R = 0.4$ is used to cluster the PF objects into jets. The anti- k_t algorithm is preferred due to its infrared

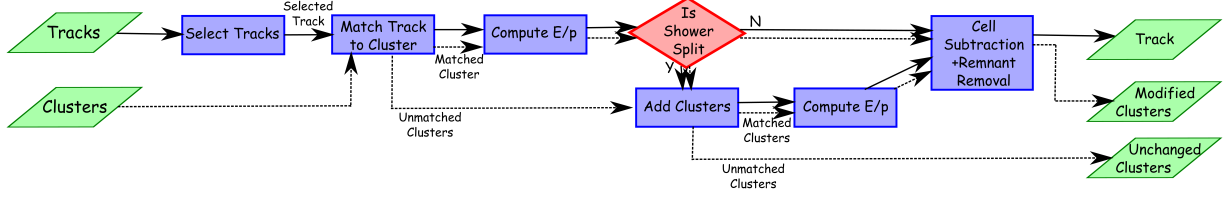


Figure 3.6: Schematic representation of the particle flow algorithm [19].

and collinear safety, which ensures that the jet reconstruction is robust against soft and collinear emissions [89]. The FASTJET package [91] is used to implement the jet clustering algorithms.

3.5.1 Calibration

Jet energy calibration is crucial to correct for detector effects and ensure accurate jet energy measurements. The calibration process involves several steps, including pile-up correction, absolute energy scale calibration, and flavor-dependent corrections. These steps with their descriptions are shown in Figure 3.7.

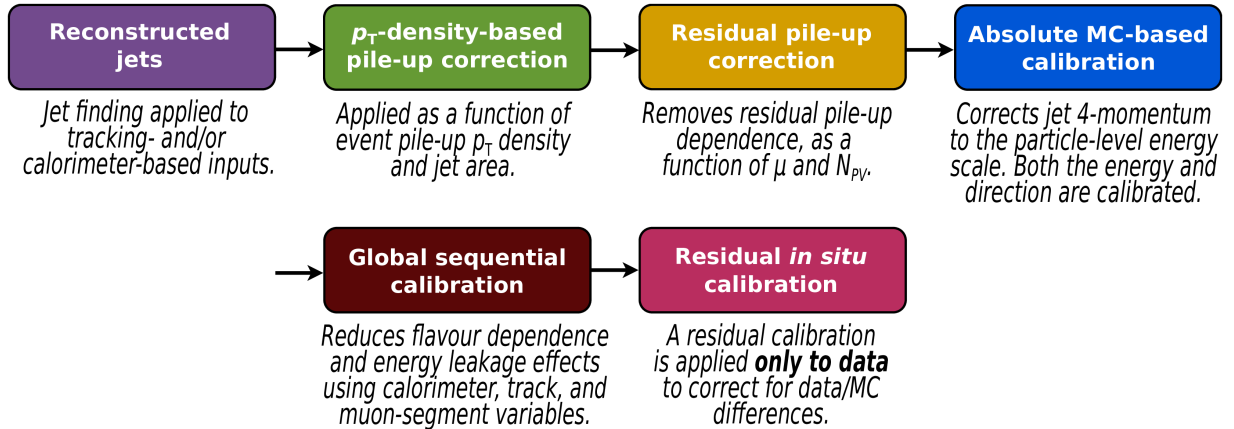


Figure 3.7: Schematic representation of the jet energy calibration steps [20].

3.5.2 Jet Vertex Tagging

Jet Vertex Tagging (JVT) is a technique used to suppress jets originating from pile-up interactions. It utilizes information from tracks associated with the jet and their compatibility with the primary vertex (PV). The JVT algorithm combines variables such as the fraction of jet p_T carried by tracks associated with the PV and the number of tracks from the PV to compute a JVT score. A threshold is applied to the JVT score to classify jets as originating from the PV or from pile-up [92]. A forward JVT (fJVT) algorithm is also used to tag forward jets (with $|\eta| > 2.5$) originating from pile-up interactions [93].

3.5.3 Flavor Tagging

Flavor tagging is the process of identifying jets originating from specific quark flavors, such as b -quarks or c -quarks. This is achieved using multivariate algorithms that combine various discriminating variables related to the properties of tracks and secondary vertices within the jet. The GN2 b -tagging algorithm [94] is used in this analysis to identify b -jets. The algorithm uses a deep neural network (DNN) architecture to combine information from low-level track variables, secondary vertex properties, and jet kinematics to produce a b -tagging score. Different working points (WPs) are defined based on the b -tagging score, such as “Loose”, “Medium”, and “Tight”, which correspond to different levels of background rejection and signal efficiency.

3.6. Missing Transverse Momentum

Any particles that do not interact with the detector, such as neutrinos or potential dark matter candidates, will not be directly detected. Neutrinos from W boson decays contribute to missing transverse momentum in this analysis and their presence can be inferred by measuring the imbalance in the transverse momentum of the detected particles. This imbalance is referred to as missing transverse momentum (MET). The MET is calculated as the negative vector sum of the transverse momenta of all reconstructed objects in the event, including electrons, muons, jets, and photons. Additionally, a soft term is included to account for low- p_T particles that are not associated with any reconstructed objects [95]. The missing transverse momentum ($\vec{p}_T^{\text{miss}}$) is given by the equation:

$$\vec{p}_T^{\text{miss}} = - \underbrace{\left(\sum_{\text{electrons}} \vec{p}_T^e + \sum_{\text{muons}} \vec{p}_T^\mu + \sum_{\text{taus}} \vec{p}_T^\tau + \sum_{\text{jets}} \vec{p}_T^{\text{jet}} + \sum_{\text{photons}} \vec{p}_T^\gamma \right)}_{\text{Hard term}} - \underbrace{\sum_{\text{unused tracks}} \vec{p}_T^{\text{track}}}_{\text{Soft term}} \quad (3.1)$$

where the hard term includes contributions from all the reconstructed objects and the soft term accounts for the tracks not associated with any of these objects but linked to the primary vertex (PV). The magnitude of the missing transverse momentum vector denoted as E_T^{miss} is given by:

$$E_T^{\text{miss}} = |\vec{p}_T^{\text{miss}}| \quad (3.2)$$

and its azimuthal angle is given by:

$$\phi^{\text{miss}} = \tan^{-1} \left(\frac{p_y^{\text{miss}}}{p_x^{\text{miss}}} \right) \quad (3.3)$$

where p_x^{miss} and p_y^{miss} are the x and y components of the missing transverse momentum vector, respectively.

3.7. Overlap Removal

The overlap removal (OR) procedure is essential to avoid double counting of physics objects that may be reconstructed as a single object and address issues arising from close proximity of objects in the detector [96]. To remove overlaps between different physics objects, quantification of their spatial separation is important. Thus, the OR procedure involves applying selection criteria using the angular distance ΔR between objects. The OR applied in this analysis include electron-electron, muon-muon, electron-muon, electron-jet, and muon-jet overlaps. The specific criteria used for overlap removal are given in the Section 3.8 below.

3.8. Object Selection

The selection criteria for the physics objects vary based on the analysis requirements. This section highlights the criteria used in the ZH analysis. Primary vertices, as described in Section 3.2, are required to have at least two associated tracks with $p_T > 500$ MeV and when multiple vertices are present, the one with the highest $\sum p_T^2$ is selected. Track parameters are used to ensure that the leptons originate from the primary vertex (PV). The transverse impact parameter significance $|d_0|/\sigma_{d_0}$ is required to be less than 3 for muons and less than 5 for electrons. The absolute value of the projection of the longitudinal impact parameter along the beamline, $|z_0 \sin(\theta)|$, is required to be less than 0.5 mm for both electrons and muons. Both leptons have a minimum p_T of 10 GeV.

Compared to Run 2 analysis, leptons from the “crack region” between the barrel and endcap EM calorimeters ($1.37 < |\eta| < 1.52$) are now included in the analysis. This is because the analysis is statistics limited and the effect of lepton quality is less significant. Adding these leptons increases the signal events by approximately 5%. The identification working points (WPs) for electrons and muons are set to “MediumLH” and “Medium”, respectively. The isolation WPs for both electrons and muons are set to “Loose_VarRad”. The lepton selection criteria are summarized in Table 3.1.

Criteria	Electrons	Muons
p_T [GeV]	> 10	> 10
$ \eta $	< 2.47	< 2.5
Identification WP	MediumLH	Medium
Isolation WP	Loose_VarRad	Loose_VarRad
$ d_0 /\sigma_{d_0}$	< 5	< 3
$ z_0 \sin(\theta) $ [mm]	< 0.5	< 0.5

Table 3.1: Selection criteria for electrons and muons used in the ZH analysis.

Anti- k_t jets with a radius parameter of $R = 0.4$ are used in this analysis. Jets are required to have a minimum p_T of 20 GeV and $|\eta| < 4.5$. In general, jet counting is

performed using jets with $p_T > 30$ GeV and jets with p_T between 20 GeV and 30 GeV are used for b -jet counting. The JVT requirement is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$ to suppress pile-up jets. JVT > 0.5 corresponding to “Tight” working point is used in this analysis. The forward JVT (fJVT) is applied to jets in the forward region ($2.5 < |\eta| < 4.5$) with p_T between 20 GeV and 60 GeV. The fJVT score < 0.5 , corresponding to “Loose” working point, is used to tag forward pile-up jets. The GN2 b -tagging algorithm with a 85% efficiency working point with a minimum score of two is used to identify b -jets. The jet selection criteria are summarized in Table 3.2.

Criteria	Jets
p_T [GeV]	> 20
$ \eta $	< 4.5
JVT score	> 0.5
fJVT score	< 0.5
GN2 b -tagging score	> 2

Table 3.2: Jets selection criteria used in the ZH analysis.

Calculation of missing transverse momentum (MET) is discussed in Section 3.6. An observable based on MET, called as MET significance is constructed. This observable is useful to distinguish events with genuine MET from those with fake MET arising due to misreconstructed objects.

Finally, to avoid double counting of physics objects, the overlap removal (OR) procedure described in Section 3.7 is applied as follows.

- For electron and muon overlaps, an electron is removed if it shares the same ID track as a combined muon. However, if the muon is calorimeter-tagged, the muon is removed instead.
- For electron-electron overlaps sharing the same track or overlapping in the second layer cluster, the electron with lower p_T is removed.
- For electron-jet overlaps, if a jet is within $\Delta R < 0.2$ of an electron, the jet is removed. For remaining jets within $\Delta R < 0.4$ of an electron, the electron is removed.
- For muon-jet overlaps, if a jet is within $\Delta R < 0.2$ of a muon, the jet is removed if it has less than three associated tracks with $p_T > 500$ MeV. Additionally, if $(p_T^\mu / p_T^{\text{jet}}) > 0.5$ and $(p_T^\mu / \sum_{\text{tracks}} p_T^{\text{track}}) > 0.7$ for $p_T^{\text{track}} > 500$ MeV, the jet is removed. For remaining jets within $\Delta R < 0.4$ of a muon, the muon is removed.

4 Data and Monte Carlo Samples

Any analysis in high energy physics requires data collected from the detector as well as simulated Monte Carlo (MC) samples. The MC samples are generated using various event generators that simulate the underlying physics processes, followed by a detailed simulation of the detector response. Some commonly used event generators include PYTHIA [97], HERWIG [98, 99], SHERPA[100], MADGRAPH [101], and POWHEG-BOX [102, 103]. PYTHIA, HERWIG and SHERPA[100] are general-purpose event generators that use different parton showering and hadronization models, while MADGRAPH and POWHEG-BOX are primarily used for generating hard scattering processes at leading order (LO) and next-to-leading order (NLO) accuracy, respectively.

In this chapter, the datasets from the ATLAS detector used in this analysis are described first (Section 4.1), followed by a description of the MC simulation process and the MC samples (Section 4.2) that are used to model various physics processes relevant to this analysis.

4.1. Data

The data used in this analysis were collected by the ATLAS detector during the Run 3 of the LHC during 2022 and 2023 at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV. This represents the pp collision data with a bunch spacing of 25 ns. The total integrated luminosity of the dataset used in this analysis corresponds to 56.1 fb^{-1} with an uncertainty of 2% [104].

All events of interest are selected using a two-level trigger system described in Section 2.2.7. For the ZH analysis, a single-lepton trigger is used to select events. For the L1 trigger, the minimum transverse momentum (p_T) threshold for electrons varies based on the data-taking period. For 2022 data, it is set to 22 GeV, while for 2023 data, it is increased to 26 GeV. For muons, the L1 p_T threshold is consistently set at 14 GeV for both 2022 and 2023 data. For the HLT, the electron can pass the trigger if it has $p_T > 26$ GeV with “tight” identification and “loose” isolation criteria or $p_T > 60$ GeV with “medium” identification without any isolation requirement or $p_T > 140$ GeV with “loose” identification without any isolation requirement. Electrons have the same HLT criteria for both 2022 and 2023 data. For muons, a same HLT criteria is used for both 2022 and 2023 data. The muon can pass the trigger if it has $p_T > 24$ GeV with “medium” identification or $p_T > 50$ GeV with no additional identification or isolation requirement. Table 4.1 summarizes the trigger requirements used in this analysis.

Not all recorded events are suitable for physics analysis. Therefore, a set of data quality criteria are applied to ensure that only events recorded during stable beam condi-

Lepton	Level-1 Trigger	High Level Trigger
2022		
e	L1EM22VHI	HLT_e26_lhtight_ivarloose
	L1EM22VHI	HLT_e60_lhmedium
	L1EM22VHI	HLT_e140_lhloose
μ	L1MU14FCH	HLT_mu24_ivarmedium
	L1MU14FCH	HLT_mu50
2023		
e	L1eEM26M	HLT_e26_lhtight_ivarloose
	L1eEM26M	HLT_e60_lhmedium
	L1eEM26M	HLT_e140_lhloose
μ	L1MU14FCH	HLT_mu24_ivarmedium
	L1MU14FCH	HLT_mu50

Table 4.1: Summary of the single-lepton trigger requirements used in this analysis for 2022 and 2023 data-taking periods.

tions and when all relevant detector components were functioning properly are considered. These criteria are defined in the form of Good Run Lists (GRLs) [68]. After applying the GRLs, the total integrated luminosity of the dataset used in this analysis is 56.1 fb^{-1} . The GRLs used in this analysis are summarized in Table 4.2.

Year	Good Run List
2022	data22_13p6TeV.periodAllYear_DetStatus-v134-pro28-10_MERGED_PHYS_StandardGRL_All_Good_25ns_ignore_TRIGLAR.xml
2023	data23_13p6TeV.periodAllYear_DetStatus-v133-pro31-11_MERGED_PHYS_StandardGRL_All_Good_25ns_ignoreTRIG_JETCTPIN.xml

Table 4.2: Summary of the Good Run Lists (GRLs) used in this analysis for 2022 and 2023 data-taking periods.

4.2. Monte Carlo Simulations

The Monte Carlo simulation process involves several key steps to accurately model the physics processes and detector response. The first step is the generation of hard-scattering events, which represent the primary interaction between the incoming particles i.e. partons in this analysis. Partons are the quarks and gluons that make up the protons involved in the collision. This is followed by the modeling of underlying events, which include additional soft interactions between the remnants of the colliding protons. Next, parton showering is simulated, where the high-energy quarks and gluons produced in the hard-scattering process radiate additional particles, leading to a cascade of lower-energy partons. These partons then undergo hadronization, a process in which they combine to form color-neutral hadrons. The unstable hadrons subsequently decay into stable particles, which

are the ones ultimately detected. In addition to these processes, pile-up interactions are simulated. These are additional proton-proton collisions that occur in the same or nearby bunch crossings as the primary interaction. Pile-up is an important aspect to consider, as it can significantly affect the detector response and the reconstruction of the events.

The next step involves simulating the response of the ATLAS detector to the particles produced in the event. This includes modeling the interaction of particles with the various components of the detector, such as the calorimeters, trackers, and muon systems. The simulation is performed using the GEANT4 [105] framework, which provides a detailed and accurate representation of the detector geometry and materials. Finally, the simulated events are reconstructed using algorithms that mimic those applied to real data. This ensures that the simulated events can be directly compared to the observed data, allowing for a robust analysis of the physics processes under study. Figure 4.1 provides a schematic representation of the MC event generation and simulation process.

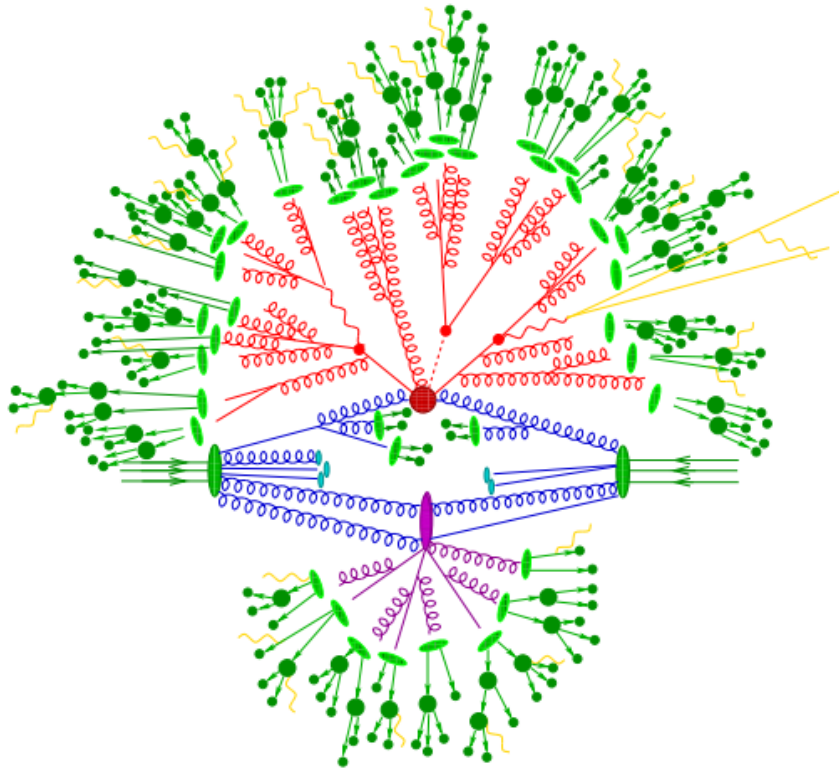


Figure 4.1: Schematic representation of the Monte Carlo simulation process. The red blob represents the hard-scattering event with the structure around it represents the parton showering. The blue color are the incoming partons from the protons involved in the hard-scattering process. The purple color represents the secondary hard scattering events. The light green blobs correspond to the hadronization process, and the darker green represent the hadron decays. The yellow lines are related to the soft photons emitted during the process [21].

The MC samples used to model the signal and background processes relevant to this

analysis are described in the following subsections.

4.2.1 Signal Samples

The signal process considered in this analysis is the associated production of a Higgs boson with a Z boson (ZH), where the Higgs boson decays via the $H \rightarrow WW^* \rightarrow lll\nu l\nu$ channel. The signal MC samples are generated using the POWHEG-BOX generator which provides next-to-leading order (NLO) accuracy in perturbative QCD. The parton showering and hadronization are simulated using PYTHIA8 [97] with the A14 tune [106]. The samples corresponding to Higgs boson decaying via WW^* channel are generated using the PDF4LHC21 [107] parton distribution functions (PDFs). The cross-sections are matched to the next-to-next-to-leading order (NNLO) accuracy in QCD and next-to-leading order (NLO) accuracy in electroweak (EW) corrections [108]. The Higgs boson mass is assumed to be 125 GeV in all signal samples. Table 4.3 summarizes the signal MC samples used in this analysis.

Process	Generator	σ [pb]	Precision
Signal			
$qq \rightarrow Z(\rightarrow inc)H(\rightarrow WW \rightarrow \ell\nu\ell\nu)$	POWHEG+PYTHIA8	0.8397	NNLO
$gg \rightarrow Z(\rightarrow inc)H(\rightarrow WW \rightarrow \ell\nu\ell\nu)$	POWHEG+PYTHIA8	0.0598	NNLO
Background			
$qq \rightarrow ZZ/Z\gamma^* \rightarrow 4\ell$	SHERPA 2.2.16	1.3312	NLO
$qq \rightarrow ZZ \rightarrow 4\ell 2j(\text{VBS})$	SHERPA 2.2.14	0.0148	LO
$gg \rightarrow ZZ \rightarrow 4\ell$	POWHEG+PYTHIA8	0.0284	LO
$tWZ(\rightarrow \ell\ell)$	MADGRAPH5+PYTHIA8	0.0168	NLO
$t\bar{t}Z(\rightarrow ee)$	MADGRAPH5+PYTHIA8	0.0415	NLO
$t\bar{t}Z(\rightarrow \mu\mu)$	MADGRAPH5+PYTHIA8	0.0415	NLO
$t\bar{t}Z(\rightarrow \tau\tau)$	MADGRAPH5+PYTHIA8	0.0411	NLO
$WWZ \rightarrow 4\ell 2\nu$	SHERPA 2.2.16	0.0019	NLO
$WZZ \rightarrow 5\ell 1\nu$	SHERPA 2.2.16	0.0002	NLO
$ZZZ \rightarrow 4\ell 2\nu$	SHERPA 2.2.16	0.0004	NLO
$ZZZ \rightarrow 6\ell$	SHERPA 2.2.16	0.000015	NLO
$qq \rightarrow Z(\rightarrow inc)H(\rightarrow \tau\tau)$	POWHEG+PYTHIA8	0.8396	NLO
$gg \rightarrow Z(\rightarrow inc)H(\rightarrow \tau\tau)$	POWHEG+PYTHIA8	0.0598	NLO
$t\bar{t}H$ dilepton	POWHEG+PYTHIA8	0.0589	NLO

Table 4.3: Summary of the signal MC samples used in this analysis along with their generators, cross-sections (σ), and precision levels for the highest order calculations.

4.2.2 Background Samples

The main background processes for this analysis include diboson production (ZZ , $Z\gamma^*$), top-quark associated production ($t\bar{t}Z$, tWZ), triboson production (WWZ , WZZ , ZZZ),

and other Higgs boson decay channels (ZH with $H \rightarrow \tau\tau, t\bar{t}H$). The ZZ and $Z\gamma^*$ processes are simulated using the SHERPA 2.2.16 generator [100] for qq -initiated events, while POWHEG+PYTHIA8 is employed for gg -initiated events. The cross-section calculations achieve NLO accuracy in QCD for qq -initiated processes, whereas the VBS and gg -initiated processes are calculated at LO precision. The POWHEG+PYTHIA8 samples utilize the NNPDF2.3 [109] PDF set.

The $t\bar{t}Z$ and tWZ processes are modeled using the MADGRAPH5_AMC@NLO [101] generator, which provides NLO accuracy in QCD and employs the NNPDF3.0 [110] PDF set. Parton showering and hadronization for these samples are performed with PYTHIA8 using the A14 tune. Triboson processes (WWZ, WZZ, ZZZ) are generated with SHERPA 2.2.16 at NLO accuracy in QCD. The other Higgs boson decay channels, including ZH with $H \rightarrow \tau\tau$ and $t\bar{t}H$, are produced using POWHEG+PYTHIA8 at NLO accuracy in QCD. The $ZH(\rightarrow \tau\tau)$ samples also rely on the NNPDF2.3 PDF set. Table 4.3 provides a summary of the background MC samples used in this analysis.

5 Data-Driven Estimation of Fakes

Fakes are objects that are misidentified as something that they are not. In the context of this analysis, the fakes refer to fake leptons which do not originate from the primary interaction vertex and are objects that are misidentified as leptons. They can arise from various sources, such as hadrons, jets misidentified as leptons, photon conversions, or non-prompt leptons from heavy-flavor decays. Accurately estimating the contribution of fake leptons is crucial for reducing background contamination and improving the sensitivity of the analysis. For the ZH analysis, a data-driven method is used over relying on MC simulation to estimate the fake lepton background. The two main reasons being limited statistics in MC samples and mismodeling of fake leptons in MC.

The ZH analysis focuses on events with four leptons in the final state, where the fake lepton background arises primarily from events containing one or two fake leptons. Contributions from events with more than two fake leptons are negligible and are therefore not considered. The events with one fake lepton mainly originate from WZ/γ^* and $tZ/t\bar{t}W$ processes, while events with two fake leptons primarily come from Z +jets and $tW/t\bar{t}$ processes. To estimate the fake lepton background in the signal region (SR), a data-driven method called the fake factor (FF) is employed.

The fake factor (FF) method involves measuring the rate at which non-prompt or misidentified leptons pass the lepton selection criteria in a control region (CR) enriched with fake leptons. This rate, known as the fake factor, is then applied to events in the signal region (SR) that contain loose leptons (leptons that fail the tight selection criteria but pass a looser set of criteria) to estimate the number of fake leptons in the SR. This is illustrated in Figure

The fake factor method consists of two main steps:

1. Fake Control Regions: Control regions are defined that are enriched in fake leptons. The contribution of prompt leptons in these regions is subtracted using MC simulation, thus deriving the fake contribution from data.
2. Fake Factors: The fake factors are measured in the control regions by calculating the ratio of the number of “ID” (tight) leptons to the number of “Anti-ID” (loose) leptons.

There are two main fake control regions: one with exactly three prompt leptons and one fake lepton ($3l + 1X$ CR) and another with exactly two prompt leptons and two fake leptons ($2l + 2X$ CR). The prompt lepton (l) follows the selection criteria used in the signal region, also called as “ID”, while the fake lepton (X) follows a looser selection criteria, also called as “Anti-ID”. The criteria for defining ID and Anti-ID leptons are described in Table [5.1](#).

	Muons		Electrons	
Requirement	ID	Anti-ID	ID	Anti-ID
p_T (GeV)	> 10		> 10	
Identification	Medium	Loose	LHMedium	LHVeryLoose
Isolation	Loose_VarRad	Not applied	Loose_VarRad	Not applied
$ \eta $	< 2.5		< 2.47	
$ z_0 \sin(\theta) $ (mm)	< 0.5		< 0.5	
$ d_0 $ significance	< 3	< 15	< 5	
Veto ID	–	Yes	–	Yes

Table 5.1: Requirements of ID and Anti-ID leptons used in defining the prompt and non-prompt leptons in the ZH analysis. The Veto ID selection is only applied to the Anti-ID leptons and refers to the inverted isolation or impact parameter ID-requirements.

The contribution of fakes from the $3l + 1X$ and $2l + 2X$ CRs are extrapolated to the SR using a fake factor (FF) for each lepton flavor (electron or muon). A dedicated control region ($2l + Y$ CR) enriched with Z +jets events is used to measure the fake factors. This CR contains exactly two same-flavor opposite-sign (SFOS) leptons forming a Z boson candidate and at least one additional lepton (Y) which may or may not be a fake lepton. The fake factor is defined as the ratio of the number of ID leptons to the number of Anti-ID leptons in this CR, after subtracting the prompt lepton contribution using MC simulation.

The contribution of fakes in a given CR is calculated using the data yield in that CR after subtracting the prompt lepton contribution estimated from MC simulation. It is given by:

$$N_{\text{fakes}}^{\text{CR}} = N_{\text{data}}^{\text{CR}} - N_{\text{prompt}}^{\text{CR}} \quad (5.1)$$

where $N_{\text{data}}^{\text{CR}}$ is the number of events observed in data and $N_{\text{prompt}}^{\text{CR}}$ is the number of events with four prompt leptons estimated from MC simulation. The contribution of fakes in the SR is then estimated by applying the fake factors to the fakes contributions from the $3l + 1X$ and $2l + 2X$ CRs. It is given by:

$$N_{\text{fakes}}^{\text{SR}} = N_{\text{fakes}}^{3l+1X \text{ CR}} \times FF - N_{\text{fakes}}^{2l+2X \text{ CR}} \times FF \times FF \quad (5.2)$$

where FF is the fake factor measured in the $2l + Y$ CR. Notice that the contribution from the $2l + 2X$ CR is subtracted to avoid double counting of events with two fake leptons in the $3l + 1X$ CR [111]. The events with two fake leptons are counted in the $3l + 1X$ CR when one of the fake leptons passes the ID criteria and is counted as a prompt lepton.

The FF is measured separately for electrons and muons, thus Equation 5.2 can be expanded to account for different lepton flavors in the final state as follows:

$$N_{\text{fakes}}^{\text{SR}} = N_{\text{fakes}}^{3l+e \text{ CR}} \times FF_e + N_{\text{fakes}}^{3l+\mu \text{ CR}} \times FF_\mu - N_{\text{fakes}}^{2l+ee \text{ CR}} \times FF_e^2 - N_{\text{fakes}}^{2l+\mu\mu \text{ CR}} \times FF_\mu^2 - N_{\text{fakes}}^{2l+e\mu \text{ CR}} \times FF_e \times FF_\mu \quad (5.3)$$

where FF_e and FF_μ are the fake factors for electrons and muons, respectively.

In the following sections, the details of the fake control regions (Section 5.2), fake factor measurements (Section 5.2), and the associated systematic uncertainties (Section 5.3) are discussed.

5.1. Fake Estimation Control Regions

5.1.1 $3l + 1X$ Control Region

As discussed earlier, the $3l + 1X$ control region (CR) contains events with three ID leptons and one Anti-ID lepton. The Anti-ID lepton passes the selection described in Table 5.1 and has inverted isolation or impact parameter requirements compared to ID leptons. This ensures that the $3l + 1X$ CR is orthogonal to the signal region (SR).

Out of the three ID leptons, the leading pair of leptons constitutes a same-flavor opposite-sign (SFOS) pair. If multiple SFOS ID lepton pairs are present, the pair with invariant mass closest to the Z boson mass is selected. The leptons forming the SFOS pair are labelled as l_2 and l_3 , while the remaining ID lepton and the Anti-ID lepton are labelled as l_0 and l_1 , following the convention for the SR. The estimated contribution of fakes in this CR is calculated using Equation 5.1.

The expected yields from the MC simulation and the observed data yield in the $3l + 1X$ CR are shown in Table 5.2 and Table 5.3 for the electron and muon channels in the 1SFOS category, respectively.

$\mathcal{L} = 56fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
$3l+e$: preselection	266.88 ± 0.87	2.67 ± 0.05	17.99 ± 0.30	3.29 ± 0.07	0.85 ± 0.03	1.98 ± 0.05	458	182.32 ± 21.42	76.06 ± 1.98
1-SFOS	10.44 ± 0.20	1.28 ± 0.04	9.37 ± 0.22	1.74 ± 0.05	0.42 ± 0.02	1.10 ± 0.04	70	55.02 ± 8.37	36.50 ± 1.08
$\Delta R_{01} > 0.2$	10.44 ± 0.20	1.28 ± 0.04	9.36 ± 0.22	1.74 ± 0.05	0.42 ± 0.02	1.10 ± 0.04	70	55.03 ± 8.37	36.49 ± 1.08
$N_{b-jet} = 0, m_{\tau\tau} < 5TeV$	10.14 ± 0.20	1.22 ± 0.04	1.52 ± 0.08	0.49 ± 0.03	0.41 ± 0.02	0.16 ± 0.01	53	40.58 ± 7.28	30.65 ± 0.91

Table 5.2: Comparison of observed event yields from data and expected event yields from MC simulation in the $3l + e$ -fake CR for the 1-SFOS category. Only statistical uncertainties are included.

$\mathcal{L} = 56fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
$3l+\mu$: preselection	73.66 ± 0.44	1.04 ± 0.03	7.60 ± 0.19	1.46 ± 0.05	0.99 ± 0.03	1.29 ± 0.04	159	80.55 ± 12.62	16.64 ± 0.90
1-SFOS	6.47 ± 0.13	0.48 ± 0.02	3.53 ± 0.13	0.75 ± 0.03	0.50 ± 0.02	0.74 ± 0.03	38	29.05 ± 6.17	9.13 ± 0.66
$\Delta R_{01} > 0.2$	6.47 ± 0.13	0.47 ± 0.02	3.51 ± 0.13	0.75 ± 0.03	0.50 ± 0.02	0.72 ± 0.03	38	29.08 ± 6.17	9.13 ± 0.66
$N_{b-jet} = 0, m_{\tau\tau} < 5TeV$	6.29 ± 0.13	0.45 ± 0.02	0.67 ± 0.06	0.25 ± 0.02	0.49 ± 0.02	0.12 ± 0.01	26	18.39 ± 5.10	7.56 ± 0.59

Table 5.3: Comparison of observed event yields from data and expected event yields from MC simulation in the $3l + \mu$ -fake CR for the 1-SFOS category. Only statistical uncertainties are included.

The e -fakes and μ -fakes contributions in the $3l + 1X$ CR for the 1-SFOS category are approximately 79% and 76%, respectively. Comparing the estimated fake yields derived from data, after subtracting the contributions from prompt backgrounds, with the fake

yields obtained directly from MC simulation reveals a significant discrepancy for both e -fakes and μ -fakes. This highlights the limitations of the MC simulation in accurately modeling fake processes and underscores the need for a data-driven approach to estimate fakes in the H analysis.

The yields in the $3\ell + 1X$ CR for the 2-SFOS category are shown in Table 5.4 and Table 5.5 for the electron and muon channels, respectively. The e -fakes and μ -fakes contributions are approximately 33% and 43%, respectively, in the 2-SFOS category. The reduced fake contribution in the 2-SFOS category compared to the 1-SFOS category is primarily due to the ZZ background, which dominates the prompt background yields in the 2-SFOS category. A similar disagreement between the data-driven fake estimate and the MC prediction is observed in the 2-SFOS category as well.

$\mathcal{L} = 56fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	$t\bar{t}H$	Data	Data-Prompt	fakes (MC)
$3\ell + e$: preselection	266.88 ± 0.87	2.67 ± 0.05	17.99 ± 0.30	3.29 ± 0.07	0.85 ± 0.03	1.98 ± 0.05	458	182.32 ± 21.42	76.06 ± 1.98
2-SFOS	256.44 ± 0.85	1.39 ± 0.03	8.62 ± 0.21	1.56 ± 0.05	0.43 ± 0.02	0.89 ± 0.03	388	127.30 ± 19.72	39.56 ± 1.66
$\Delta R_{01} > 0.2$	256.22 ± 0.85	1.39 ± 0.03	8.60 ± 0.21	1.56 ± 0.05	0.43 ± 0.02	0.88 ± 0.03	387	126.53 ± 19.69	39.49 ± 1.66
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	250.25 ± 0.84	1.34 ± 0.03	1.33 ± 0.08	0.45 ± 0.03	0.42 ± 0.02	0.13 ± 0.01	363	110.41 ± 19.07	35.43 ± 1.62
$m_{4\ell} \notin [115, 130]GeV$	246.35 ± 0.84	1.33 ± 0.03	1.34 ± 0.08	0.44 ± 0.03	0.42 ± 0.02	0.12 ± 0.01	343	94.34 ± 18.54	33.18 ± 1.46
$10 GeV < m_{\ell_0\ell_1} < 50 GeV$	35.47 ± 0.33	0.24 ± 0.02	0.41 ± 0.04	0.10 ± 0.01	0.13 ± 0.01	0.06 ± 0.01	48	12.00 ± 6.94	12.81 ± 1.19

Table 5.4: Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + e$ -fake CR for the 2-SFOS category. Only statistical uncertainties are included.

$\mathcal{L} = 56fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	$t\bar{t}H$	Data	Data-Prompt	fakes (MC)
$3\ell + \mu$: preselection	73.66 ± 0.44	1.04 ± 0.03	7.60 ± 0.19	1.46 ± 0.05	0.99 ± 0.03	1.29 ± 0.04	159	80.55 ± 12.62	16.64 ± 0.90
2-SFOS	67.19 ± 0.42	0.56 ± 0.02	4.06 ± 0.14	0.71 ± 0.03	0.49 ± 0.02	0.56 ± 0.03	121	51.50 ± 11.01	7.51 ± 0.61
$\Delta R_{01} > 0.2$	66.90 ± 0.42	0.56 ± 0.02	4.04 ± 0.14	0.71 ± 0.03	0.49 ± 0.02	0.55 ± 0.03	121	51.79 ± 11.01	7.51 ± 0.61
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	65.25 ± 0.42	0.54 ± 0.02	0.76 ± 0.06	0.23 ± 0.02	0.48 ± 0.02	0.06 ± 0.01	111	44.44 ± 10.54	6.54 ± 0.57
$m_{4\ell} \notin [115, 130]GeV$	63.27 ± 0.41	0.53 ± 0.02	0.73 ± 0.06	0.23 ± 0.02	0.46 ± 0.02	0.05 ± 0.01	101	36.44 ± 10.06	5.92 ± 0.53
$10 GeV < m_{\ell_0\ell_1} < 50 GeV$	14.80 ± 0.20	0.14 ± 0.01	0.26 ± 0.03	0.08 ± 0.01	0.23 ± 0.01	0.02 ± 0.00	26	10.72 ± 5.10	2.91 ± 0.33

Table 5.5: Comparison of observed event yields from data and expected event yields from MC simulation in the $3\ell + \mu$ -fake CR for the 2-SFOS category. Only statistical uncertainties are included.

The invariant mass distribution of the SFOS lepton pair ($m_{\ell_2\ell_3}$) in the $3\ell + 1X$ CR is shown in Figure 5.1 for both e -fakes and μ -fakes in the 1-SFOS and 2-SFOS categories. These distributions are presented after applying the preselection and the $\Delta R_{\ell_0\ell_1}$ requirements. The data points represent the observed data after subtracting the contributions from four-prompt lepton processes, while the stacked histograms represent the contributions from various MC fake processes. As expected, there is a mismatch between the data and MC fake estimation, as the fake processes are not well-modeled in MC simulation. The distributions provide insight into the expected fake background composition that would affect the signal region. For the e -fake cases, the fake background around the Z boson mass peak is primarily driven by WZ/γ^* processes in the 1-SFOS category. In the 2-SFOS category, contributions from Z +jets and $Z\gamma$ processes are also significant. For the μ -fake cases, the dominant contribution around the Z boson mass peak appears to be from Z +jets processes, which exhibit large uncertainties, along with contributions

from WZ/γ^* and top-related processes in both categories. In the 2-SFOS the disagreement between data and MC fakes is improved after applying the requirement on $m_{\ell_0\ell_1}$, particularly for the e -fake cases.

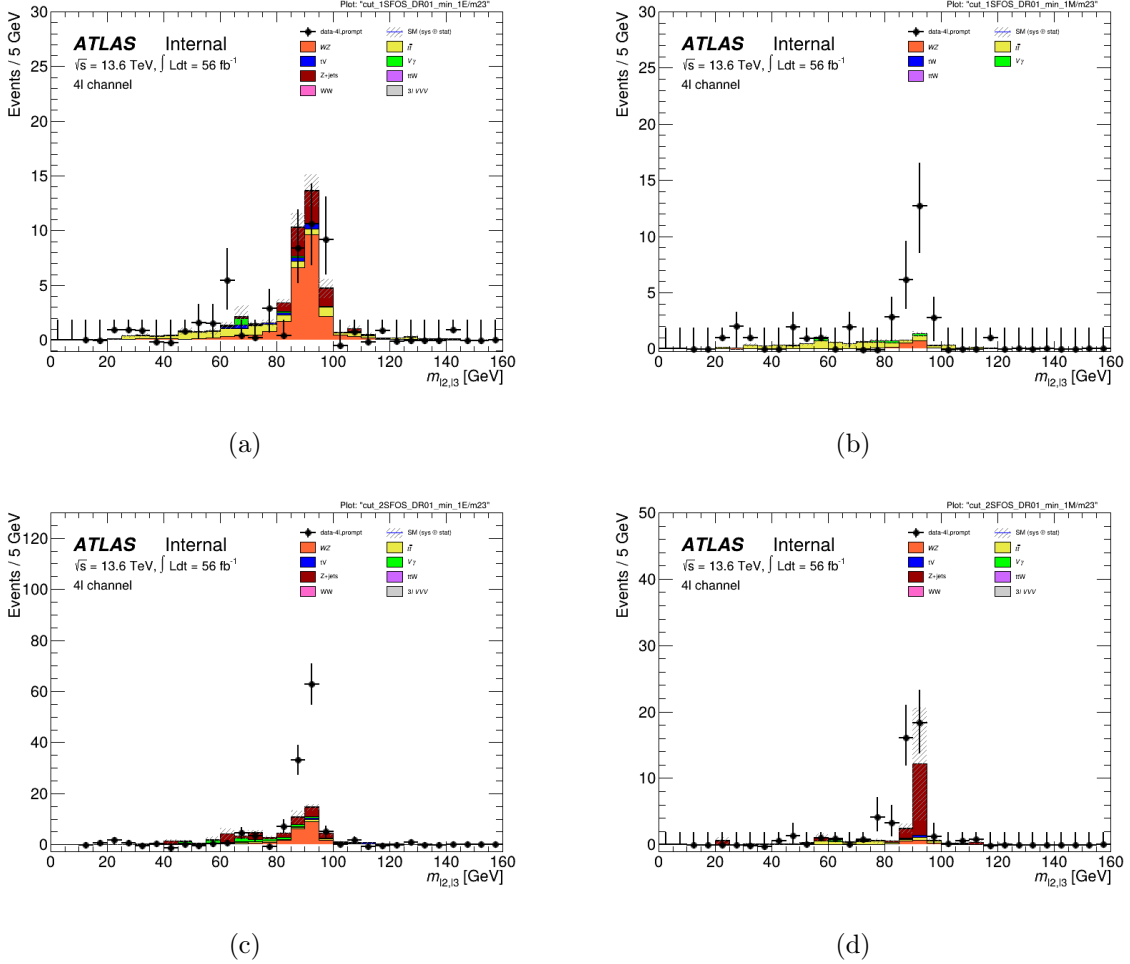


Figure 5.1: The invariant mass distribution of the SFOS lepton pair ($m_{\ell_2\ell_3}$) in the $3\ell + 1X$ CR for (a) e -fakes in the 1-SFOS category, (b) μ -fakes in the 1-SFOS category, (c) e -fakes in the 2-SFOS category, and (d) μ -fakes in the 2-SFOS category. The distributions are shown after applying the preselection and the $\Delta R_{\ell_0\ell_1}$ requirements. The uncertainties shown include only statistical uncertainties.

5.1.2 $2l + 2X$ Control Region

The $2l + 2X$ CR is designed to select events with two leptons passing the ID criteria and two leptons passing the Anti-ID criteria. The two ID leptons form a SFOS pair and are represented by ℓ_2 and ℓ_3 . Table 5.6 and Table 5.7 show the event yields for the $e\mu$ fakes, Table 5.8 and Table 5.9 show the event yields for the ee fakes, and Table 5.10 and Table 5.11 show the event yields for the $\mu\mu$ fakes in the $2l + 2X$ CR for the 1SFOS and 2SFOS

categories, respectively. This CR has about 88% fake lepton purity in both the SFOS categories for the initial stages of the selection and more events compared to the $3l + 1X$ CR. As seen in the $3l + 1X$ CR, a disagreement between fakes estimated from data and MC simulation is observed.

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
2l+e μ : preselection	14.63 ± 0.22	0.19 ± 0.01	2.03 ± 0.10	0.32 ± 0.02	0.25 ± 0.01	0.44 ± 0.02	137	121.17 ± 11.71	65.56 ± 2.47
1-SFOS	4.21 ± 0.15	0.11 ± 0.01	1.37 ± 0.08	0.21 ± 0.02	0.21 ± 0.01	0.27 ± 0.02	100	94.98 ± 10.00	45.85 ± 2.22
$\Delta R_{01} > 0.2$	4.19 ± 0.15	0.11 ± 0.01	1.36 ± 0.08	0.21 ± 0.02	0.21 ± 0.01	0.27 ± 0.02	100	95.01 ± 10.00	45.80 ± 2.22
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	4.05 ± 0.15	0.11 ± 0.01	0.28 ± 0.04	0.09 ± 0.01	0.21 ± 0.01	0.06 ± 0.01	71	66.49 ± 8.43	33.20 ± 2.00

Table 5.6: Event yields in the $2l + 2X$ control region for the 1-SFOS category for the $e\mu$ fakes. The uncertainties shown are statistical only.

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
2l+e μ : preselection	14.63 ± 0.22	0.19 ± 0.01	2.03 ± 0.10	0.32 ± 0.02	0.25 ± 0.01	0.44 ± 0.02	137	121.17 ± 11.71	65.56 ± 2.47
2-SFOS	10.42 ± 0.17	0.08 ± 0.01	0.65 ± 0.06	0.10 ± 0.01	0.04 ± 0.00	0.17 ± 0.01	37	26.19 ± 6.09	19.71 ± 1.09
$\Delta R_{01} > 0.2$	10.38 ± 0.17	0.08 ± 0.01	0.65 ± 0.06	0.10 ± 0.01	0.04 ± 0.00	0.17 ± 0.01	37	26.23 ± 6.09	19.71 ± 1.09
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	10.09 ± 0.16	0.07 ± 0.01	0.15 ± 0.03	0.04 ± 0.01	0.04 ± 0.00	0.03 ± 0.01	32	21.74 ± 5.66	13.13 ± 0.91
$m_{4\ell} \notin [115, 130]GeV$	9.73 ± 0.16	0.07 ± 0.01	0.14 ± 0.03	0.04 ± 0.01	0.03 ± 0.00	0.02 ± 0.01	29	19.10 ± 5.39	12.25 ± 0.90
$10 GeV < m_{\ell_0\ell_1} < 50 GeV$	2.32 ± 0.08	0.02 ± 0.00	0.05 ± 0.02	0.02 ± 0.00	0.02 ± 0.00	0.01 ± 0.00	13	10.61 ± 3.61	6.14 ± 0.68

Table 5.7: Event yields in the $2l + 2X$ control region for the 2-SFOS category for the $e\mu$ fakes. The uncertainties shown are statistical only.

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
2l+ee: preselection	30.98 ± 0.30	0.32 ± 0.02	2.50 ± 0.11	0.40 ± 0.02	0.11 ± 0.01	0.38 ± 0.02	315	282.80 ± 17.75	117.48 ± 5.84
1-SFOS	0.90 ± 0.05	0.15 ± 0.01	1.03 ± 0.07	0.17 ± 0.02	0.03 ± 0.00	0.19 ± 0.02	41	39.56 ± 6.40	25.53 ± 1.27
$\Delta R_{01} > 0.2$	0.90 ± 0.05	0.15 ± 0.01	1.04 ± 0.07	0.17 ± 0.02	0.03 ± 0.00	0.19 ± 0.02	41	39.56 ± 6.40	25.53 ± 1.27
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	0.87 ± 0.05	0.14 ± 0.01	0.15 ± 0.03	0.05 ± 0.01	0.03 ± 0.00	0.03 ± 0.01	20	18.88 ± 4.47	13.55 ± 0.97

Table 5.8: Event yields in the $2l + 2X$ control region for the 1-SFOS category for the ee fakes. The uncertainties shown are statistical only.

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
2l+ee: preselection	30.98 ± 0.30	0.32 ± 0.02	2.50 ± 0.11	0.40 ± 0.02	0.11 ± 0.01	0.38 ± 0.02	315	282.80 ± 17.75	117.48 ± 5.84
2-SFOS	30.07 ± 0.30	0.17 ± 0.01	1.47 ± 0.09	0.24 ± 0.02	0.09 ± 0.01	0.19 ± 0.02	274	243.23 ± 16.56	91.96 ± 5.70
$\Delta R_{01} > 0.2$	29.42 ± 0.30	0.17 ± 0.01	1.43 ± 0.09	0.24 ± 0.02	0.09 ± 0.01	0.18 ± 0.02	273	242.91 ± 16.53	91.88 ± 5.70
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	28.66 ± 0.29	0.17 ± 0.01	0.25 ± 0.04	0.08 ± 0.01	0.09 ± 0.01	0.03 ± 0.01	248	218.98 ± 15.75	76.26 ± 5.43
$m_{4\ell} \notin [115, 130]GeV$	27.73 ± 0.29	0.17 ± 0.01	0.25 ± 0.04	0.07 ± 0.01	0.09 ± 0.01	0.03 ± 0.01	221	192.92 ± 14.87	67.71 ± 5.20
$10 GeV < m_{\ell_0\ell_1} < 50 GeV$	6.27 ± 0.14	0.04 ± 0.01	0.09 ± 0.02	0.02 ± 0.01	0.03 ± 0.00	0.02 ± 0.00	113	106.63 ± 10.63	37.32 ± 4.28

Table 5.9: Event yields in the $2l + 2X$ control region for the 2-SFOS category for the ee fakes. The uncertainties shown are statistical only.

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
2l+ $\mu\mu$: preselection	7.49 ± 0.14	0.06 ± 0.01	0.56 ± 0.05	0.09 ± 0.01	0.14 ± 0.01	0.20 ± 0.02	44	36.02 ± 6.63	21.66 ± 1.05
1-SFOS	0.20 ± 0.02	0.02 ± 0.00	0.22 ± 0.03	0.03 ± 0.01	0.01 ± 0.00	0.09 ± 0.01	7	6.66 ± 2.65	10.80 ± 0.76
$\Delta R_{01} > 0.2$	0.20 ± 0.02	0.02 ± 0.00	0.22 ± 0.03	0.03 ± 0.01	0.01 ± 0.00	0.09 ± 0.01	7	6.66 ± 2.65	10.80 ± 0.76
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5TeV$	0.19 ± 0.02	0.02 ± 0.00	0.04 ± 0.01	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	6	5.76 ± 2.45	9.80 ± 0.74

Table 5.10: Event yields in the $2l + 2X$ control region for the 1-SFOS category for the $\mu\mu$ fakes. The uncertainties shown are statistical only.

The invariant mass distribution of the two ID leptons ($m_{\ell_2\ell_3}$) in the $2l + 2X$ CR for the 1SFOS and 2SFOS categories are shown in Figure 5.2. The fakes in observed

$\mathcal{L} = 56 fb^{-1}$	ZZ	VVV	$t\bar{t}Z$	tWZ	$ZH \rightarrow \tau\tau$	ttH	Data	Data-Prompt	fakes (MC)
$2l+\mu\mu$: preselection	7.49 ± 0.14	0.06 ± 0.01	0.56 ± 0.05	0.09 ± 0.01	0.14 ± 0.01	0.20 ± 0.02	44	36.02 ± 6.63	21.66 ± 1.05
2-SFOS	7.30 ± 0.14	0.04 ± 0.01	0.34 ± 0.04	0.06 ± 0.01	0.13 ± 0.01	0.11 ± 0.01	37	29.37 ± 6.08	10.86 ± 0.72
$\Delta R_{01} > 0.2$	7.28 ± 0.14	0.04 ± 0.01	0.34 ± 0.04	0.06 ± 0.01	0.13 ± 0.01	0.11 ± 0.01	37	29.38 ± 6.08	10.86 ± 0.72
$N_{b\text{-jet}} = 0, m_{\tau\tau} < 5 TeV$	7.10 ± 0.14	0.04 ± 0.01	0.09 ± 0.01	0.02 ± 0.00	0.12 ± 0.01	0.02 ± 0.00	34	26.70 ± 5.83	9.80 ± 0.68
$m_{4\ell} \notin [115, 130] GeV$	6.77 ± 0.14	0.04 ± 0.01	0.08 ± 0.01	0.01 ± 0.00	0.12 ± 0.01	0.02 ± 0.00	29	22.04 ± 5.39	9.09 ± 0.66
$10 GeV < m_{\ell_0\ell_1} < 50 GeV$	2.32 ± 0.08	0.01 ± 0.00	0.04 ± 0.01	0.00 ± 0.00	0.08 ± 0.01	0.01 ± 0.00	16	13.57 ± 4.00	4.64 ± 0.47

Table 5.11: Event yields in the $2l + 2X$ control region for the 2-SFOS category for the $\mu\mu$ fakes. The uncertainties shown are statistical only.

events is evaluated by subtracting the prompt contribution from MC backgrounds. The distributions are shown after applying the preselection and the ΔR_{01} cut and it is shown for each of the three cases: $e\mu$, ee , and $\mu\mu$. For the 1SFOS category, the main contribution to fakes comes from the Z +jets process where a jet is misidentified as a lepton, followed by the contribution from top processes. In the 2SFOS category, the dominant contribution is also from the Z +jets process and the remaining contribution is from processes like $V\gamma$ and top processes.

5.2. Determining the Fake Factors

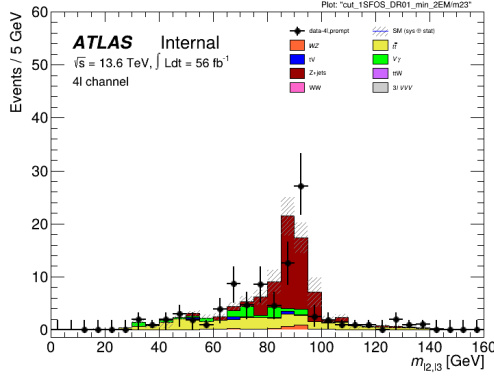
5.2.1 $2\ell + 1Y$ Control Region

As discussed in Section 5, the $2\ell + 1Y$ CR is defined to determine the extrapolation factor which when multiplied by the number of events in the $3\ell + 1X$ CR and $2\ell + 2X$ CR using Equation 5.2 gives the estimated number of fake lepton events in the SR. To this end, the $2\ell + 1Y$ CR is defined to be orthogonal to both the SR and other CRs used in the fake lepton estimation. It is orthogonal to the SR and CRs as it contains only three leptons in the final state. Two of these leptons are required to pass the ID selection criteria and the third lepton (Y) can either pass the ID or the Anti-ID selection criteria. The total charge of the three leptons is ± 1 and the events pass the preselection criteria defined in Section 7.3.

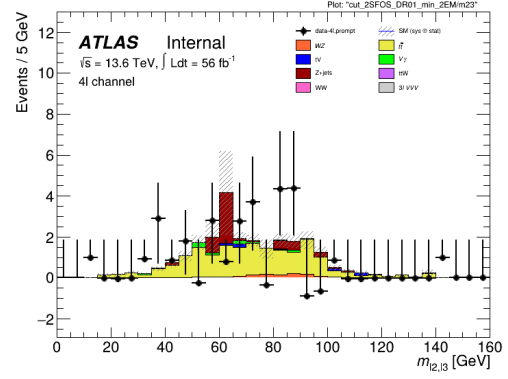
The two ID leptons are required to form a SFOS pair and have $p_T > 25$ GeV. If the ID leptons are electrons, then their invariant mass must be in the range $[80, 100]$ GeV and if they are muons, then the invariant mass must be in the range $[80, 100]$ GeV. This ensures that the ID leptons are consistent with originating from a Z boson decay. If more than one SFOS pair can be formed, then the pair with invariant mass closest to the Z boson mass is selected.

The main contribution to the events with two SFOS pairs in the final state arises from the WZ process with prompt leptons. To suppress this background, a cut on the transverse mass of the W boson is applied. The transverse mass is calculated using the Anti-ID lepton and the missing transverse energy in the event as follows:

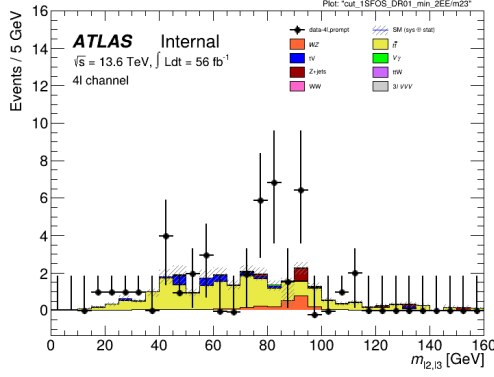
$$m_T^W = \sqrt{2E_T^{\text{fake}} E_T^{\text{miss}} (1 - \cos(\Delta\phi))} \quad (5.4)$$



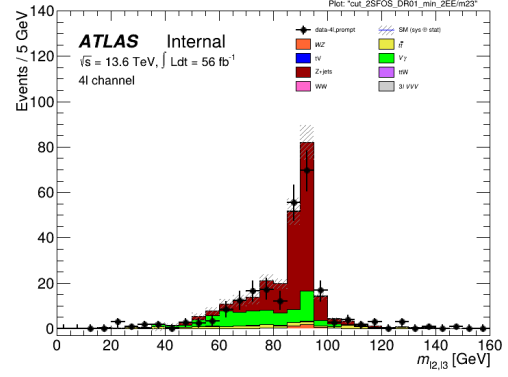
(a)



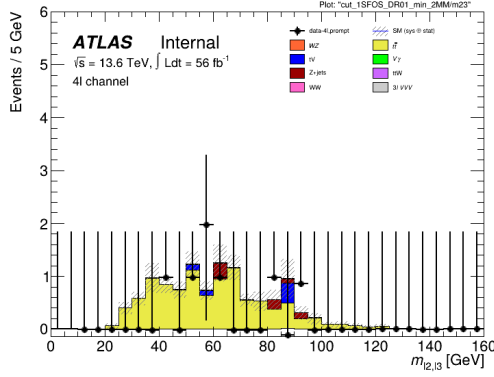
(b)



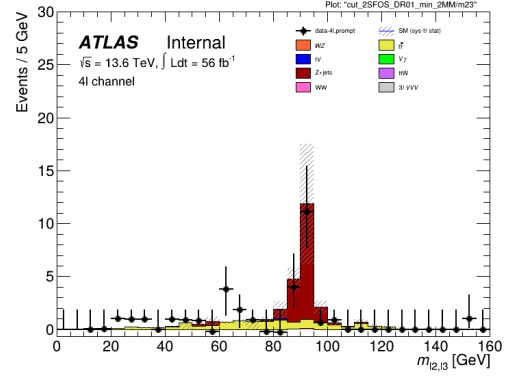
(c)



(d)



(e)



(f)

Figure 5.2: The invariant mass distribution of the SFOS lepton pair ($m_{l_2 l_3}$) in the $2l + 2X$ control region for (a) 1SFOS $e\mu$, (b) 2SFOS $e\mu$, (c) 1SFOS ee , (d) 2SFOS ee , (e) 1SFOS $\mu\mu$, and (f) 2SFOS $\mu\mu$ categories. The distributions are shown after applying the preselection and the ΔR_{01} cut. The fake contribution is estimated from data by subtracting the prompt contribution from MC simulation. The uncertainties shown are statistical only.

where E_T^{fake} is the transverse energy of the fake object and $\Delta\phi$ is the azimuthal angle between the fake object and the missing transverse energy vector. A lower cut of 50 GeV is applied on m_T^W to suppress the WZ background.

The expected yields from MC processes and the observed yield in data in the $2\ell + 1Y$ CR are shown in Table 5.12. The dominant contribution in this CR arises from the Z +jets process when Y is an Anti-ID lepton. Another major contribution comes from the WZ process. In order to correctly estimate the contribution from WZ , a dedicated CR is defined by inverting the m_T^W cut. The yields in this WZ CR are shown in Table 5.13. A normalization factor of 1.04 ± 0.01 is derived from this CR using the ratio of observed data to the expected yield from prompt MC processes.

$\mathcal{L} = 56 fb^{-1}$ (Early Run3)	Z +jets	$V + \gamma$	WZ	ZZ	VVV	Top	Backgrounds ($V + \gamma$, VV , Top)	Data	Fakes	Data/MC
Channel Selection	398520.07 $\pm$ 904.36	103439.18 $\pm$ 318.94	22019.21 $\pm$ 19.95	14602.45 $\pm$ 11.72	146.90 $\pm$ 0.27	51872.52 $\pm$ 57.74	192080.27 $\pm$ 324.95	591823	399742.73 $\pm$ 835.11	1.00 $\pm$ 0.00
Ill final state	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
All lep $p_T > 10$ GeV	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
Overlap: $V\gamma$ gamma/ V jets	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
$m_{\ell} > 10$ GeV	328812.55 $\pm$ 795.87	84639.49 $\pm$ 279.23	19580.97 $\pm$ 15.79	9334.77 $\pm$ 9.59	117.81 $\pm$ 0.22	45987.87 $\pm$ 54.20	159660.92 $\pm$ 285.04	456082	296421.08 $\pm$ 733.03	0.93 $\pm$ 0.00
Z -tagging	178059.60 $\pm$ 517.17	15454.37 $\pm$ 121.50	14496.86 $\pm$ 13.42	3778.19 $\pm$ 5.48	40.95 $\pm$ 0.16	5189.26 $\pm$ 17.00	38959.62 $\pm$ 123.54	220003	181043.38 $\pm$ 485.04	1.01 $\pm$ 0.00
Z lepton $p_T > 25$ GeV	140911.86 $\pm$ 475.39	12379.23 $\pm$ 109.96	11128.35 $\pm$ 11.70	2944.28 $\pm$ 4.84	31.00 $\pm$ 0.14	3649.99 $\pm$ 14.16	30132.86 $\pm$ 111.59	173152	143019.14 $\pm$ 430.82	1.01 $\pm$ 0.00
TransMassW	135181.90 $\pm$ 470.70	11413.12 $\pm$ 108.67	3749.66 $\pm$ 6.79	2398.38 $\pm$ 2.58	8.56 $\pm$ 0.07	2028.64 $\pm$ 10.76	19598.36 $\pm$ 109.51	156041	136442.64 $\pm$ 409.92	1.01 $\pm$ 0.00
fake type: electron	109430.18 $\pm$ 406.75	10954.66 $\pm$ 105.86	1748.54 $\pm$ 4.64	876.97 $\pm$ 2.52	3.91 $\pm$ 0.05	866.94 $\pm$ 6.95	14451.02 $\pm$ 106.22	125505	111053.98 $\pm$ 369.85	1.01 $\pm$ 0.00
fake e η cut	93554.01 $\pm$ 382.65	10234.62 $\pm$ 103.19	1646.49 $\pm$ 4.49	826.22 $\pm$ 2.45	3.69 $\pm$ 0.05	783.02 $\pm$ 6.57	13494.04 $\pm$ 103.53	104574	91079.96 $\pm$ 339.55	0.98 $\pm$ 0.00
ID cuts	5061.94 $\pm$ 107.56	1071.09 $\pm$ 34.18	1391.33 $\pm$ 4.13	617.37 $\pm$ 1.92	3.15 $\pm$ 0.04	137.58 $\pm$ 2.24	3220.52 $\pm$ 34.56	9240	6019.48 $\pm$ 102.15	1.12 $\pm$ 0.02
Anti-ID cuts	88492.06 $\pm$ 367.22	9163.53 $\pm$ 97.37	255.16 $\pm$ 1.76	208.85 $\pm$ 1.53	0.54 $\pm$ 0.02	645.44 $\pm$ 6.18	10273.52 $\pm$ 97.59	95334	85060.48 $\pm$ 323.82	0.97 $\pm$ 0.00
fake type: muon	25751.72 $\pm$ 236.89	458.47 $\pm$ 24.56	2001.12 $\pm$ 4.96	1521.41 $\pm$ 3.82	4.65 $\pm$ 0.05	1161.70 $\pm$ 8.21	5147.33 $\pm$ 26.64	30536	25388.67 $\pm$ 176.77	0.99 $\pm$ 0.01
fake μ η cut	25371.63 $\pm$ 235.49	449.38 $\pm$ 24.42	1984.97 $\pm$ 4.93	1511.23 $\pm$ 3.81	4.61 $\pm$ 0.05	1143.06 $\pm$ 8.14	5093.25 $\pm$ 26.49	30139	25045.75 $\pm$ 175.61	0.99 $\pm$ 0.01
ID cuts	4256.31 $\pm$ 108.58	81.27 $\pm$ 10.72	1775.57 $\pm$ 4.68	1099.84 $\pm$ 3.00	4.07 $\pm$ 0.05	227.89 $\pm$ 3.20	3188.65 $\pm$ 12.49	7634	4445.35 $\pm$ 88.26	1.03 $\pm$ 0.02
Anti-ID cuts	19711.60 $\pm$ 200.83	342.69 $\pm$ 20.12	208.64 $\pm$ 1.55	410.65 $\pm$ 2.35	0.54 $\pm$ 0.02	859.41 $\pm$ 7.25	1821.92 $\pm$ 21.57	20895	19073.08 $\pm$ 146.15	0.97 $\pm$ 0.01

Table 5.12: Event yields observed in data and expected from MC processes in the $2\ell + 1Y$ CR. The uncertainties shown are statistical only.

$\mathcal{L} = 56 fb^{-1}$ (Early Run3)	Z +jets	$V + \gamma$	WZ	ZZ	VVV	Top	Backgrounds ($V + \gamma$, VV , Top)	Data	Fakes	Data/MC
Channel Selection	398520.07 $\pm$ 904.36	103439.18 $\pm$ 318.94	22019.21 $\pm$ 19.95	14602.45 $\pm$ 11.72	146.90 $\pm$ 0.27	51872.52 $\pm$ 57.74	192080.27 $\pm$ 324.95	591823	399742.73 $\pm$ 835.11	1.00 $\pm$ 0.00
Ill final state	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
All lep $p_T > 10$ GeV	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
Overlap: $V\gamma$ gamma/ V jets	376443.85 $\pm$ 879.39	98229.30 $\pm$ 311.11	21619.69 $\pm$ 19.80	10543.30 $\pm$ 10.69	118.62 $\pm$ 0.23	47655.15 $\pm$ 55.34	178166.06 $\pm$ 316.79	578464	400297.94 $\pm$ 823.91	1.04 $\pm$ 0.00
$m_{\ell} > 10$ GeV	328812.55 $\pm$ 795.87	84639.49 $\pm$ 279.23	19580.97 $\pm$ 15.79	9334.77 $\pm$ 9.59	117.81 $\pm$ 0.22	45987.87 $\pm$ 54.20	159660.92 $\pm$ 285.04	456082	296421.08 $\pm$ 733.03	0.93 $\pm$ 0.00
Z -tagging	178059.60 $\pm$ 517.17	15454.37 $\pm$ 121.50	14496.86 $\pm$ 13.42	3778.19 $\pm$ 5.48	40.95 $\pm$ 0.16	5189.26 $\pm$ 17.00	38959.62 $\pm$ 123.54	220003	181043.38 $\pm$ 485.04	1.01 $\pm$ 0.00
Z lepton $p_T > 25$ GeV	140911.86 $\pm$ 475.39	12379.23 $\pm$ 109.96	11128.35 $\pm$ 11.70	2944.28 $\pm$ 4.84	31.00 $\pm$ 0.14	3649.99 $\pm$ 14.16	30132.86 $\pm$ 111.59	173152	143019.14 $\pm$ 430.82	1.01 $\pm$ 0.00
WZ CR (TransMassW > 50 GeV)	536.86 $\pm$ 20.13	119.61 $\pm$ 6.18	6844.93 $\pm$ 9.18	488.72 $\pm$ 1.43	20.83 $\pm$ 0.11	535.91 $\pm$ 4.43	8010.00 $\pm$ 12.01	8850	840.00 $\pm$ 94.84	1.04 $\pm$ 0.01

Table 5.13: Event yields observed in data and expected from MC processes in the WZ CR. The uncertainties shown are statistical only.

Lepton flavor	Fake factor
Electron	0.071
Muon	0.233

Table 5.14: Summary of the fake factor values measured in the $2\ell + 1Y$ CR for electrons and muons using data-driven technique.

The fake factor is defined as the ratio of the number of fakes that pass the ID criteria to the number of fakes that pass the Anti-ID criteria. It is calculated as follows:

$$FF = \frac{N_{\text{fakes}}^{\text{ID, CR}}}{N_{\text{fakes}}^{\text{Anti-ID, CR}}} = \frac{N_{\text{data}}^{\text{ID, } 2\ell+1Y} - N_{\text{prompt}}^{\text{ID, } 2\ell+1Y}}{N_{\text{data}}^{\text{Anti-ID, } 2\ell+1Y} - N_{\text{prompt}}^{\text{Anti-ID, } 2\ell+1Y}} \quad (5.5)$$

The prompt contributions come from MC processes like WZ/γ^* , $t\bar{t}W/tZ$ and triboson processes with three prompt leptons. The value of the fake factor is measured separately for electrons and muons and is summarized in Table 5.14. The invariant mass distributions of the lepton pair forming the SFOS pair and the transverse mass distribution of the W boson is shown in Figure 5.3.

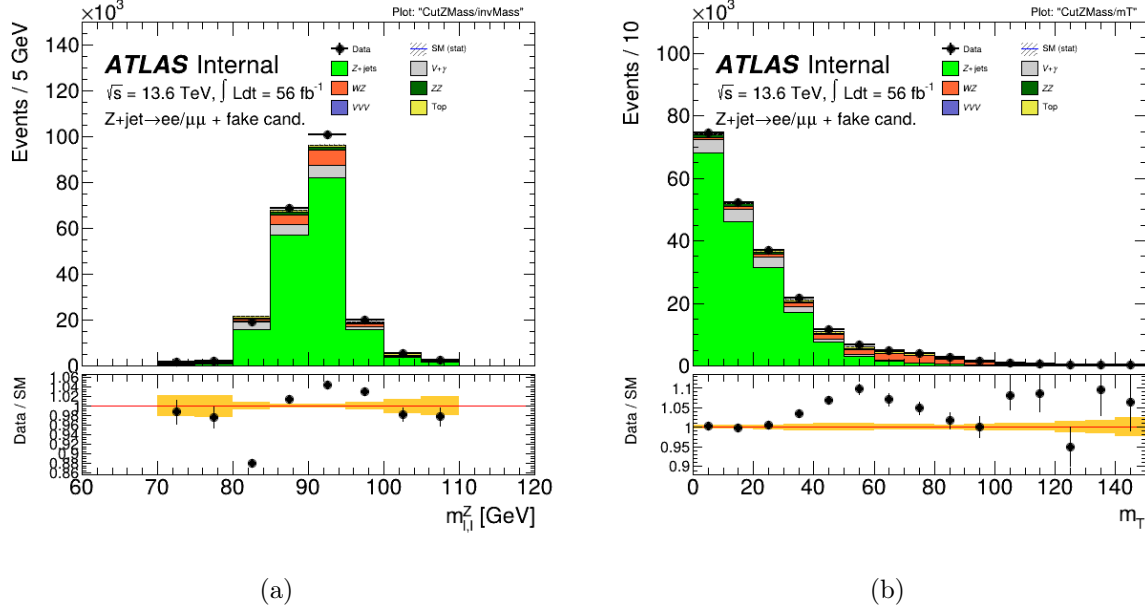


Figure 5.3: Distribution of (a) $m_{\ell\ell}^Z$ and (b) m_T^W in the $2\ell + Y$ CR. Only statistical uncertainties are shown.

5.3. Systematic Uncertainties from Fake Factors

In the ZH analysis, systematic uncertainties associated with the fake factor method are considered to account for potential variations in the fake factor measurements and their impact on the estimated fake lepton background in the SR. The main sources of systematic uncertainties related to the fake factors include:

- **Statistical Uncertainty:** The statistical uncertainty in the fake factor measurement arises from the limited number of events in the fake CRs.
- **Uncertainty from prompt lepton subtraction**
- **Extrapolation Uncertainty** also known as non-closure uncertainty

5.3.1 Prompt Lepton Subtraction Uncertainty

The prompt lepton subtraction uncertainty accounts for the potential mismodeling of prompt lepton contributions in the Z +jets ($2\ell + 1Y$) CR used for measuring the fake

factors. The main contribution to this comes from the WZ process. To estimate this uncertainty, the yields of the WZ process from MC simulation are varied up and down by 30% in the Z -jets CR and the new fake factors are calculated using Equation 5.5. The relative difference between the nominal (summarized in Table 5.14) fake factor and the fake factors calculated with the varied WZ yields are calculated and the largest relative difference is taken as the systematic uncertainty due to prompt lepton subtraction. These uncertainties are found to be 6.9% for electrons and 11.7% for muons. Table 5.15 summarizes the prompt lepton subtraction uncertainties.

$\mathcal{L} = 56 \text{ fb}^{-1}$	Nominal	Stat. Error	Up	Difference	Down	Difference
FF(e)	0.071 ± 0.001	1.7%	0.066	-6.9%	0.076	+6.8%
FF(μ)	0.233 ± 0.005	2.1%	0.206	-11.7%	0.260	+11.6%

Table 5.15: Summary of the prompt lepton subtraction systematic uncertainties on the fake factors for electrons and muons. The "Up" and "Down" columns represent the fake factor values obtained by varying the WZ yield by $\pm 30\%$ in the Z +jets control region. The 'Difference' columns show the relative differences between the nominal fake factor and the varied fake factors. The uncertainties on nominal fake factors due to statistical errors are also shown.

5.3.2 MC Non-closure Uncertainty

The MC Non-closure uncertainty is also known as extrapolation uncertainty as it accounts for potential differences in the estimation of fakes in the SR due to the extrapolation factor or FF applied to the fake CRs. Fake factors based on MC simulation are calculated and these are then used to estimate the fakes contribution in the SR. Table 5.16 summarizes the fake factors calculated using MC simulation. The estimated fakes from MC-based fake factors are then compared to the fakes from data-driven fake factors to assess the level of non-closure. The relative difference between these two estimates is taken as the non-closure uncertainty. This uncertainty is found to be 8.4% for electrons and 19.2% for muons. Table 5.17 shows the summary of the non-closure uncertainties.

FF flavor	MC-based FF
Electron	0.057
Muon	0.216

Table 5.16: Fake factors calculated using MC simulation in the $2\ell + 1Y$ CR for electrons and muons. Sherpa MC samples are used for the Z +jets process.

$\mathcal{L} = 56 \text{ fb}^{-1}$	WZ events	Difference
4 ℓ SR (fake electron)	1.51 ± 0.11	-
(3 ℓ + e CR) $\times$ FF $_e$	$(24.36 \pm 0.43) \times 0.057 = 1.39 \pm 0.02$	8.4%
4 ℓ SR (fake muon)	0.30 ± 0.07	-
(3 ℓ + μ CR) $\times$ FF $_e$	$(1.72 \pm 0.11) \times 0.216 = 0.37 \pm 0.02$	19.2%

Table 5.17: Summary of the non-closure systematic uncertainties on the fake factors for electrons and muons. The second row for each flavor shows the estimated fakes in the SR using MC-based fake factors. The ‘Difference’ column shows the relative difference between the nominal fakes using data-driven fake factors and the fakes estimated using MC-based fake factors.

Table 5.18 shows a summary of all the systematic uncertainties associated with the fake factor method.

Uncertainty	Electrons	Muons
FF statistics	1.7%	2.1%
FF MC subtraction	6.9%	11.7%
Non-closure test	8.4%	19.2%
Total	11.0%	22.6%

Table 5.18: Summary of all the systematic uncertainties associated with the fake factor method for electrons and muons. The total uncertainty is calculated by adding the individual uncertainties in quadrature.

6 | Neural Networks

Using machine learning algorithms [112] is common in most high energy physics experiments as it helps to identify complex patterns hidden in the large datasets which might be missed by traditional techniques. In Run 2, boosted decision trees (BDTs) [113] were used to separate the signal events from background. In the Run 3, BDTs are replaced with artificial neural networks (ANNs) [114] as they can overcome the limitations of BDTs.

Boosted decision trees (BDTs) have certain limitations, such as their inefficiency in capturing higher-order feature interactions, their tendency to create non-smooth decision boundaries, and the requirement for a large number of trees to achieve good performance. In contrast, artificial neural networks (ANNs) are capable of addressing these limitations, provided they are properly tuned and sufficient data is available. Tuning a neural network involves optimizing its architecture (Section 6.1) and hyperparameters (Section 6.2) to achieve the best possible performance on a given task. With the increasing volume of data collected in each experimental run, neural networks (NNs) become a natural choice over BDTs due to their scalability and ability to model complex, non-linear relationships in the data.

In this chapter, an overview of neural networks is provided in Section 6.1 to define the basic concepts and overall architecture of a neural network, followed by the optimization techniques used to tune the neural network in Section 6.2, and model evaluation and validation techniques in Section 6.3. The final neural network architecture used in this analysis is discussed in Section 7.4.3.

6.1. Overview of Neural Networks

An artificial neural network is a machine learning algorithm inspired from a biological neural network. A biological neuron as shown in Figure 6.1 (a) receives signals from other neurons through dendrites. If the sum of these inputs is more than a certain threshold, the neuron ‘fires’ i.e. transmits signal along its axon. These signals are then passed to other neurons through axon terminal.

A perceptron is a simple artificial neural network that takes multiple inputs and gives a single binary output based on the input weights and bias (negative threshold) as shown in Figure 6.1 (b). If the weighted sum of weights is greater than a certain threshold, the perceptron outputs 1 and 0 otherwise. In general, a neural network has three types of layers: the input layer, the output layer and hidden layer(s). The input layer consists of input variables (neurons), the neurons in the middle layer are neither input nor output and constitute a hidden layer. The output layer consists of output neurons, whose values in the case of binary classification are interpreted as the probability that a given event

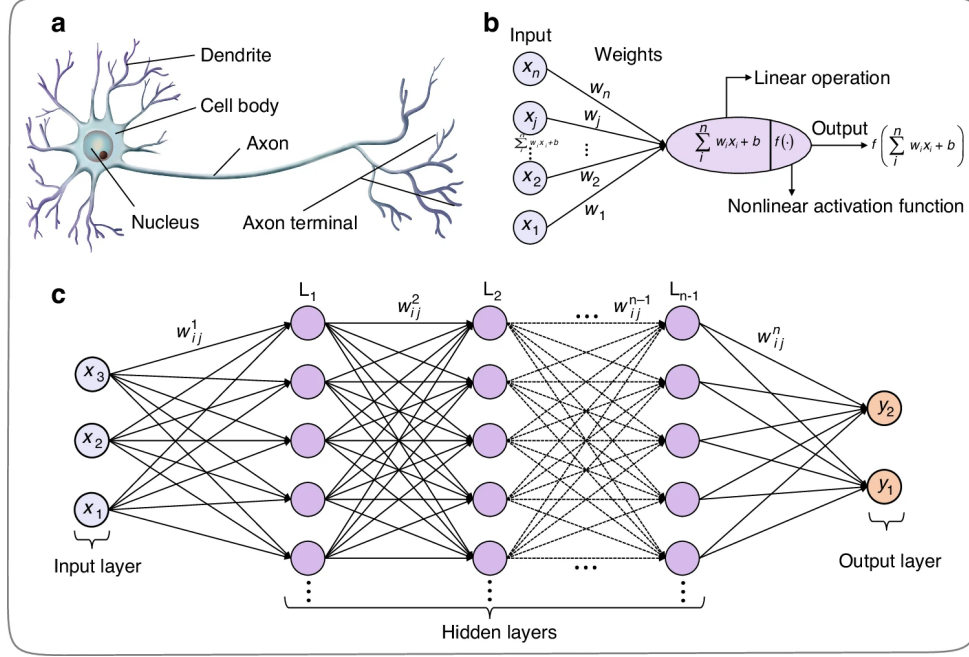


Figure 6.1: (a) Schematic of a biological neuron (b) A perceptron with multiple inputs, weights, bias and a non-linear activation function with single output (c) A deep neural network with multiple hidden layers [22].

(set of input variables) belongs to a specific class (signal or background). A deep neural network is the neural network with multiple hidden layers as shown in Figure 6.1 (c). These neural networks are also called as feed-forward neural network as the information is fed from one layer to other in a forward direction.

The weights and bias can be adjusted based on the decision model to solve a specific problem. Small changes in weights and bias can lead to large changes in the output when using a perceptron model. To overcome this, a non-linear activation function is used to introduce non-linearity to the model. Let's represent the sum of weighted inputs as z .

$$z = \sum (w_i x_i) + b = w \cdot x + b \quad (6.1)$$

where w_i are the weights, x_i are the inputs and b is the bias. The output of a neuron can then be represented as $a = f(z)$, where f is the activation function which is differentiable. The weights and biases are determined based on the output predictions and the true labels (if present) using a cost/loss function. The process of determining the weights and biases is called as training the neural network. The training can be done using supervised learning (with labeled data) or unsupervised learning (without labeled data).

A loss function is defined to calculate the accuracy of the prediction. Some common loss functions are mean squared error (MSE), cross-entropy loss, Huber loss and mean absolute Error (MAE). For binary classification problems, binary cross-entropy loss is commonly used and for multi-class classification problems, categorical cross-entropy loss

is used. This analysis uses a binary classification problem to classify events as signal or background. The binary cross-entropy loss function is defined as:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (6.2)$$

where N is the number of samples, y_i is the true label and $\hat{y}_i$ is the predicted probability for sample i . The loss function is used to optimize the neural networks by adjusting the weights during training to minimize the prediction error. This change is propagated backwards, from the outer layer towards the inner layer and called as backward propagation [115]. The loss function is minimized as a function of weights and bias using a stochastic gradient descent (SGD) algorithm [116]. The well-known SGD variants like Adam [117], Nadam [118] are explored in this analysis.

The varying scales, weights of input features can affect the training. To prevent any bias towards a feature due to uneven scaling, a common scaling is implemented on all input features. Some of the common scaling features include:

- Min-Max Scaling: Scales the features to a fixed range, usually $[0, 1]$, using the formula:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (6.3)$$

where x_{min} and x_{max} are the minimum and maximum values of the feature.

- Standardization (z-score normalization): Scales the features to have zero mean and unit variance:

$$x' = \frac{x - \mu}{\sigma} \quad (6.4)$$

where μ is the mean and σ is the standard deviation of the feature.

- Robust Scaling: Scales the features using the median and interquartile range, making it robust to outliers:

$$x' = \frac{x - \text{median}}{\text{IQR}} \quad (6.5)$$

where IQR is the interquartile range (75th percentile - 25th percentile).

- Normalization: Scales the features to have a unit norm:

$$x' = \frac{x}{\|x\|} \quad (6.6)$$

where $\|x\|$ is the norm of the feature vector.

The choice of activation function can also significantly impact the performance of a neural network. Some commonly used activation functions are sigmoid, hyperbolic tangent (tanh), Rectified Linear Unit (ReLU), Leaky ReLU, Swish and Softmax as given below:

- Sigmoid: $f(z) = \frac{1}{1+e^{-z}}$
- Tanh: $f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$
- ReLU: $f(z) = \max(0, z)$
- Leaky ReLU: $f(z) = \max(0.01z, z)$
- Swish: $f(z) = \frac{z}{1+e^{-z}}$
- Softmax: $f(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$ for each output neuron

The sigmoid and tanh functions are smooth and differentiable, making them suitable for binary classification problems. However, they can suffer from the vanishing gradient problem, where gradients become very small during backpropagation, leading to slow learning. ReLU is computationally efficient and helps mitigate the vanishing gradient problem, but it can suffer from the "dying ReLU" problem, where neurons can become inactive and only output zero. Leaky ReLU addresses this issue by allowing a small, non-zero gradient when the input is negative. Swish is a newer activation function that has shown promise in improving performance in deep networks. Softmax is typically used in the output layer as it converts the output values generated by the final layer into probabilities that sum to one.

Hyperparameters discussed in detail in Section 6.2, are model independent parameters, useful to optimize the training process. The learning rate is a hyperparameter that controls the step size at each iteration while moving towards a minimum of the loss function during training. A smaller learning rate results in slower convergence, while a larger learning rate may cause the model to overshoot the minimum. A batch refers to a subset of the training dataset that is processed together in one forward and backward pass. Batch size then represents the number of events in each batch. Training on batches (rather than the entire dataset) helps in efficient computation and can improve generalization. An epoch is one complete pass through the entire training dataset. During each epoch, the model sees every sample in the dataset once, and the weights are updated accordingly.

In this analysis, a supervised learning approach is used to train the neural network to classify events as signal or background. The truth labels from the Monte Carlo (MC) simulation are used to identify signal and background events. The input features to the neural network are kinematic and topological variables that help to distinguish between signal and background events. The training process involves feeding the input data through the network, calculating the output, comparing it to the truth label using a loss function, and adjusting the weights and biases to minimize the loss. This process is repeated iteratively over multiple epochs until the model converges to a satisfactory level of accuracy. The training is performed using Keras [119] with the TensorFlow [120] backend. The architecture of a neural network including the number of layers, number of neurons in each layer, the activation functions used, the loss function, and the optimization algorithm for the neural networks generated in the *ZH* analysis are discussed in Section 7.4.3.

6.2. Hyperparameter Optimization

Hyperparameters in a neural network architecture are configuration variables that are set before the training process begins. They control the learning process. Unlike model parameters, which are learned during training (such as weights and biases), hyperparameters are not updated by the learning algorithm. Common hyperparameters include the learning rate, batch size, number of epochs, number of layers, number of neurons per layer, activation functions, optimizer type, dropout rate, and regularization coefficients. The choice of hyperparameters can significantly affect the performance and generalization ability of a neural network.

The input variables are transformed using z -score normalization to ensure equal contribution from all input features to the learning process and to help accelerate the convergence of gradient descent function during training. The `StandardScaler` [121] function of *scikit-learn* Python programming package is used for this transformation. The background weights are transformed such that they add upto unity, preventing bias towards either class and improving the robustness of the model.

Hyperparameter optimization is performed using the Hyperband algorithm [122] from *KerasTuner* [123], which efficiently searches the hyperparameter space to identify configurations that yield optimal model performance. Hyperband leverages early-stopping and adaptive resource allocation to evaluate more hyperparameter combinations in less time, making it well-suited for neural network tuning in this analysis. The hyperparameters tuned include the learning rate, batch size, number of epochs, number of hidden layers, number of neurons per layer, activation function, optimizer type and dropout rate after each layer.

6.3. Model Validation

The final neural network model needs to be assessed for overfitting as well as performance. Area under the receiver operating characteristic (ROC) curve is used to measure the performance of the model using the validation dataset. To check for overfitting, a k -fold cross validation technique is applied where $k = 5$ is chosen for the purpose of this analysis. This technique divides the dataset into k independent and equal sized sets. The model is trained iteratively using $k-1$ sets and tested using the remaining one. The entire dataset is used once for testing. The performance of each set is measured using evaluation metrics like accuracy. This ensures robustness and stability of the model, and prevents issues like overfitting.

The final neural network can be used to separate signal from background using an appropriate discriminant. In the ZH analysis, ZZ is the only dominant background, thus using the NN output score as a discriminant makes sense to distinguish between the signal and background.

6.4. Feature Importance

Each feature (input variable) used in training the neural network has a different level of importance or impact on the model's predictions. Certain features are expected to have higher impact due to their kinematics properties. For e.g. in ZH analysis, kinematic variables corresponding to the l_0, l_1 leptons coming from the Higgs decay should have higher impact to separate the signal from the background. Understanding the impact of each feature helps in making informed decisions about feature selection and engineering, ultimately leading to improved model performance.

The features can be ranked based on their order of importance. To calculate the impact of each feature, the features are varied individually by ± 1 standard deviation (σ) from their original values and their effect on the neural network output is observed. Higher the deviation in the NN output, higher is the impact of that feature on the model's predictions. The features can then be ranked in descending order of their impact.

7 Analysis Strategy

In this chapter, the analysis strategy to measure the Higgs boson production cross-section in the associated production mode with a Z boson, decaying via $H \rightarrow WW^*$ decay channel using the ATLAS detector at the LHC at $\sqrt{s} = 13.6$ TeV is discussed. Partial Run 3 data (2022-2023) corresponding to an integrated luminosity of 56.2 fb^{-1} is used for this analysis. The analysis strategy includes an overview of the analysis (Section 7.1), followed by the discussion of various background processes that can mimic the signal signature (Section 7.2). The event selection criteria is described in Section 7.3, neural network training in Section 7.4 and finally the control and validation regions are discussed in Section 7.5.

7.1. Analysis Overview

The ZH analysis considers the final state where the Z boson decays into a pair of leptons and the Higgs boson decays leptonically via $H \rightarrow WW^*$ decay channel, as represented by the Equation 7.1:

$$ZH \rightarrow ZWW^* \rightarrow ll\nu l\nu \quad (7.1)$$

The final state consists of four leptons i.e. electrons and muons, including those from the leptonic decays of τ particles. The leptons from Z decay form a same-flavor opposite-sign (SFOS) lepton pair, the remaining two leptons from the WW^* decay can either form a SFOS pair or a different-flavor opposite-sign (DFOS) pair. The leptons from the Z boson decay are labelled as ℓ_2 and ℓ_3 in the decreasing order of their transverse momentum p_T , while the leptons from the Higgs boson decay are labeled as ℓ_0 and ℓ_1 in the decreasing order of their p_T . The leptons ℓ_2 and ℓ_3 are chosen such that their invariant mass is closest to the Z boson mass among all the SFOS lepton pairs in the event. Henceforth, the labels used in the analysis are consistent with this notation. This is illustrated using a Feynman diagram for the signal process in Figure 7.1.

This final state is also characterized by the missing transverse momentum (E_T^{miss}) arising from the neutrinos produced in the leptonic decays of the W bosons. This helps to distinguish the signal events from the main background process $ZZ \rightarrow 4\ell$, which does not have any real missing transverse momentum. Apart from that, the leptons from the Higgs boson decay tend to have a low invariant mass and small angular separation due to the spin correlation of the Higgs boson decay products in the Higgs boson rest frame. The Higgs boson is a scalar particle with spin-0, which means that it does not have any intrinsic angular momentum. When the Higgs boson decays into two W bosons, the spins of the W bosons are correlated in such a way that they tend to be emitted in the same direction in the Higgs boson frame. Thus, the leptons produced from the decay of the W bosons also

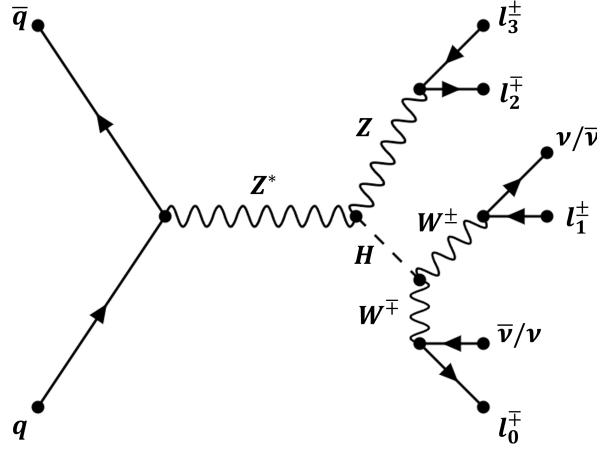


Figure 7.1: Feynman diagram illustrating the signal process $ZH \rightarrow ZWW^* \rightarrow ll\nu l\nu$.

tend to be emitted in the same direction, resulting in a small angular separation between them. This is illustrated with the help of Figure 7.2.

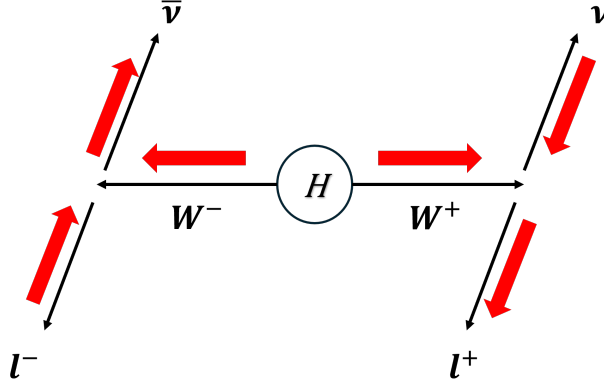


Figure 7.2: Illustration of the spin correlation of the Higgs boson decay products in the rest frame of the Higgs boson. The black arrows represent the direction of momentum whereas the red arrows represent the spin direction of the particles.

The invariant mass of the two leptons (ℓ_1 and ℓ_2) emitted from W boson decays can be reconstructed as follows:

$$m_{\ell_1\ell_2}^2 = (E_{\ell_1} + E_{\ell_2})^2 - (\vec{p}_{\ell_1} + \vec{p}_{\ell_2})^2 \quad (7.2)$$

$$= (E_{\ell_1} - p_{\ell_1})^2 + (E_{\ell_2} - p_{\ell_2})^2 + 2E_{\ell_1}E_{\ell_2} - 2|\vec{p}_{\ell_1}||\vec{p}_{\ell_2}|\cos\Delta\phi \quad (7.3)$$

$$= 2E_{\ell_1}E_{\ell_2}(1 - \cos\Delta\phi) \quad (\text{for } |\vec{p}_{\ell_i}| \gg m_{\ell_i}) \quad (7.4)$$

where, E_{ℓ_1} , E_{ℓ_2} are the energies of the two leptons, $\vec{p}_{\ell_1}$, $\vec{p}_{\ell_2}$ are the three momentum vectors and $\Delta\phi$ is the azimuthal angle between them. Equation 7.4 holds true for events when the three momentum of leptons is much higher compared to their masses, which occurs in such highly energetic events studied in this analysis. Thus, the smaller value of $\Delta\phi$

implies smaller value of the invariant mass of the leptons from Higgs boson decay.

Three different regions are defined in this analysis: the signal region (SR), control region (CR), and validation region (VR). The SR is optimized to maximize the signal-to-background ratio and is used to extract the signal yield. The CRs are enriched in specific background processes and are used to constrain the background normalizations in the final fit. The VRs are designed to be kinematically similar to the SR but with relaxed selection criteria, and are used to validate the background estimation methods. The VRs are not used in the final fit but provide an important cross-check of the analysis strategy. Various CRs and VRs defined in this analysis are discussed in detail in Section 7.5.

This analysis consists of two SRs based on the number of SFOS lepton pairs in the final state, namely 1-SFOS and 2-SFOS signal regions. As the 2-SFOS events are more prone to ZZ background, defining two separate SRs helps to improve the signal sensitivity by reducing the background contributions in each SR and achieving better segregation of signal from background. Various processes contributing to the background in each SR are discussed in detail in Section 7.2. The event selection criteria for the SRs are discussed in detail in Section 7.3. An additional cut on the output of a neural network (NN) is applied in both SRs to further enhance the signal-to-background ratio. Each SR has a dedicated NN trained and the training procedure is described in Section 7.4.

The uncertainties affecting this analysis can be broadly classified into experimental uncertainties and theoretical uncertainties. These are discussed in detail in Chapter 8. This is followed by the discussion of the statistical treatment of the analysis in Chapter 9 and finally the results are presented in Chapter 10.

7.2. Background Contribution

The background processes contributing to the $ZH \rightarrow ZWW^* \rightarrow \ell\ell\nu\ell\nu$ contains four prompt leptons and real or fake missing transverse momentum in the final state. The largest background contribution in this analysis comes from the $ZZ^* \rightarrow 4\ell$ process with real or fake missing transverse momentum, referred to as ZZ hereafter. The real missing transverse momentum comes from $Z \rightarrow \tau^+\tau^-$ decays where the τ decays leptonically producing neutrinos. The fake missing transverse momentum comes from the mismeasurement of jets and/or leptons or pileup. The final state of ZZ decaying leptonically generally contains two pairs of SFOS leptons except when at least one of the Z 's decay via $Z \rightarrow \tau^+\tau^-$ channel to produce a DFOS lepton pair. Thus, the ZZ background contributes to both 2SFOS and 1SFOS categories in this analysis.

Variables that are most important to distinguish between the signal events and ZZ background events include the invariant mass of ℓ_0 and ℓ_1 ($m_{\ell_0\ell_1}$) and the azimuthal angle separation between ℓ_0 and ℓ_1 ($\Delta\phi_{\ell_0\ell_1}$). The signal events tend to have lower values of $m_{\ell_0\ell_1}$ due to the presence of neutrinos in the final state, whereas the ZZ background have higher values, close to the Z boson mass. The value of $\Delta\phi_{\ell_0\ell_1}$ in the signal events is typically smaller compared to the ZZ background events. In the case of ZZ , the leptons are more

back-to-back, resulting in larger values of $\Delta\phi_{\ell_0\ell_1}$.

Other background processes that can mimic the signal signature include top processes such as $t\bar{t}Z$ and tWZ , whose final state contains a b -jet along with leptons and missing transverse momentum. The contribution from these processes can be reduced by vetoing events with b -jets in the final state. There is an irreducible background contribution from tri-boson processes with $4\ell 2\nu$ in the final state, such as WWZ and ZZZ . Processes like ZH and $t\bar{t}H$ with $H \rightarrow \tau\tau$ also contribute to the background when all particles decay leptonically, but their contribution can be reduced significantly by using the selection criteria on the kinematic variables and b -jet veto respectively.

Another important contribution comes from fake processes. Fake processes include events where there are two or three prompt leptons in the final state and remaining leptons are fake leptons. Fake leptons imply objects that are misidentified as leptons, such as jets or non-prompt leptons from heavy-flavor hadron decays. The main sources of fake leptons in this analysis are WZ/γ^* and tZ/tWZ with three prompt leptons and Z +jets and $t\bar{t}/tW$ with two prompt leptons in the final state. Fakes and their contributions are estimated using data-driven techniques as described in detail in Chapter 5.

7.3. Event Selection

An event refers to a single pp collision recorded by the ATLAS detector. Among all the events recorded, only those events that are of interest for the analysis are selected by applying a series of selection criteria. This process is known as event selection. In the ZH analysis, a single lepton trigger as discussed in Section 4.1 is used to select events online. The trigger requirements are already summarized in Table 4.1. Further selection criteria are applied offline to define signal regions (SRs), control regions (CRs), and validation regions (VRs) as discussed below.

The selected events are required to have at least one primary vertex reconstructed as discussed in Section 3.2. The events passing these conditions are then subjected to further selection criteria, called pre-selection. The pre-selection is common for all regions i.e. SRs, CRs, and VRs in this analysis. In this section, only the pre-selection and cuts defining the SRs are discussed. Cuts corresponding to the CRs and VRs will be discussed in Section 7.5.

The events are required to have exactly four reconstructed leptons (electrons or muons) with $p_T > 10$ GeV and the total leptonic charge must be zero. The leading lepton in the event is required to have $p_T > 27$ GeV to ensure that the leading lepton can pass the single-lepton trigger. There should be at least one same flavor opposite-sign (SFOS) lepton pair and all SFOS lepton pairs in the event should have an invariant mass greater than 12 GeV to reject background contributions from low-mass Drell-Yan, leptonic decays of γ , and low mass mesons. If there is a different flavor opposite-sign (DFOS) lepton pair in the event, it must have an invariant mass greater than 10 GeV to minimize contributions from $Z\gamma^*$ and Z + quarkonia with low mass τ pairs.

Following the pre-selection, the events are categorized into different signal regions based on the number of SFOS lepton pairs in the final state. If there are no SFOS lepton pairs in the final state, the events are rejected. Both SRs share common cuts to suppress various background contributions. Mainly, the angular separation between the two leptons ℓ_0 and ℓ_1 is required to be greater than 0.1 to remove overlapping leptons, which would otherwise lead to a double fake contribution in the $2\ell + 2X$ fake CR (Section 5.1.2). An upper cut on the invariant mass of $\tau\tau$ pair is applied following the Run 2 analysis, to reject ZZ background in 1-SFOS SR where one of the Z decays into a $\tau\tau$ pair. To reduce contribution from the top processes, a b -jet veto is applied in both SRs.

ZH analysis signal selections		
	1SFOS	2SFOS
Pre-Selection	4 isolated leptons ($p_T > 10$ GeV) Total lepton charge = 0 Leading p_T lepton: $p_T > 27$ GeV All SFOS lepton pairs: $m_{\ell\ell} > 12$ GeV All DFOS lepton pairs: $m_{\ell\ell} > 10$ GeV	
	One SFOS lepton pair	Two SFOS lepton pairs
Radial separation cut	$\Delta R_{\ell_0\ell_1} > 0.1$	
Large $m_{\tau\tau}$ cut	$m_{\tau\tau} < 5$ TeV	
Top veto	$N_{b\text{-jet}} = 0$	
$H \rightarrow ZZ$ rejection		$m_{4\ell} < 115$ GeV or $m_{4\ell} > 130$ GeV
H mass cut		$10 \text{ GeV} < m_{\ell_0\ell_1} < 50 \text{ GeV}$

Table 7.1: Summary of the selection criteria defining the signal regions for the ZH analysis before applying the neural network selection.

In the 2-SFOS SR, two additional cuts are applied. The invariant mass of the four leptons in the final state are required to be outside the Higgs mass window by 5 GeV, assuming the Higgs mass is 125 GeV. This ensures that the background contribution from $H \rightarrow ZZ$ is rejected. To further suppress the contribution of ZZ in the 2-SFOS SR, the invariant mass of the ℓ_0, ℓ_1 is required to be within 10 and 50 GeV. Following the pre-selection and application of these cuts, the events in both SRs are further subjected to a neural network selection. The details of neural network training and final architecture implemented is discussed in Section 7.4 below. The summary of all the selection criteria defining the SRs before applying the neural network selection is provided in Table 7.1. Table 7.2 shows the expected yields at varying pre-selection stages and for various cuts in both the SRs before applying the neural network selection.

$\sqrt{s} = 13.6 \text{ TeV}, \mathcal{L} = 56 \text{ fb}^{-1} \text{ (EarlyRun3)}$	$qq \rightarrow ZZ$	$gg \rightarrow ZZ$	VBS	VVV	$t\bar{t}Z\&WZ$	$t\bar{t}H$	Fakes(DD)	bkg	ZH	Signif.	Data	Data/MC
Apply weights	2290.44 $\pm$ 3.79	243.94 $\pm$ 0.35	35.34 $\pm$ 0.26	16.57 $\pm$ 0.12	114.83 $\pm$ 0.70	11.51 $\pm$ 0.12	33224.30 $\pm$ 216.53	35936.92 $\pm$ 216.57	12.74 $\pm$ 0.13	0.04 $\pm$ 0.00	9362	0.26 $\pm$ 0.00
Remove large weight	2290.44 $\pm$ 3.79	243.94 $\pm$ 0.35	35.34 $\pm$ 0.26	16.57 $\pm$ 0.12	114.83 $\pm$ 0.70	11.51 $\pm$ 0.12	33224.30 $\pm$ 216.53	35936.92 $\pm$ 216.57	12.74 $\pm$ 0.13	0.04 $\pm$ 0.00	9362	0.26 $\pm$ 0.00
4 ID leptons	1293.18 $\pm$ 2.02	185.21 $\pm$ 0.30	21.08 $\pm$ 0.20	12.47 $\pm$ 0.10	70.09 $\pm$ 0.54	5.40 $\pm$ 0.08	15.45 $\pm$ 4.32	1602.88 $\pm$ 4.82	8.71 $\pm$ 0.11	0.22 $\pm$ 0.00	1873	1.16 $\pm$ 0.03
Jet Clean.	1293.18 $\pm$ 2.02	185.21 $\pm$ 0.30	21.08 $\pm$ 0.20	12.47 $\pm$ 0.10	70.09 $\pm$ 0.54	5.40 $\pm$ 0.08	15.45 $\pm$ 4.32	1602.88 $\pm$ 4.82	8.71 $\pm$ 0.11	0.22 $\pm$ 0.00	1873	1.16 $\pm$ 0.03
Trigger Sel.	1279.53 $\pm$ 2.01	184.88 $\pm$ 0.30	21.01 $\pm$ 0.19	12.46 $\pm$ 0.10	70.09 $\pm$ 0.54	5.39 $\pm$ 0.08	15.86 $\pm$ 4.24	1589.13 $\pm$ 4.74	8.70 $\pm$ 0.11	0.22 $\pm$ 0.00	1843	1.15 $\pm$ 0.03
Trigger Match	1224.52 $\pm$ 1.95	181.50 $\pm$ 0.30	20.40 $\pm$ 0.19	12.32 $\pm$ 0.10	68.89 $\pm$ 0.54	5.26 $\pm$ 0.08	16.58 $\pm$ 4.06	1529.48 $\pm$ 4.55	8.49 $\pm$ 0.10	0.22 $\pm$ 0.00	1765	1.15 $\pm$ 0.03
$m_{t\bar{t}}^{SFOS} > 12 \text{ GeV}$	1126.42 $\pm$ 1.87	179.98 $\pm$ 0.30	19.50 $\pm$ 0.19	12.26 $\pm$ 0.10	67.71 $\pm$ 0.53	5.12 $\pm$ 0.08	22.30 $\pm$ 3.31	1433.28 $\pm$ 3.85	8.30 $\pm$ 0.10	0.22 $\pm$ 0.00	1540	1.07 $\pm$ 0.03
$m_{t\bar{t}}^{SFOS} > 10 \text{ GeV}$	1122.76 $\pm$ 1.87	179.67 $\pm$ 0.29	19.43 $\pm$ 0.19	12.23 $\pm$ 0.10	67.52 $\pm$ 0.53	5.07 $\pm$ 0.08	22.96 $\pm$ 3.18	1429.64 $\pm$ 3.75	8.21 $\pm$ 0.10	0.22 $\pm$ 0.00	1533	1.07 $\pm$ 0.03
$m_{t\bar{t}}^{lead} > 27 \text{ GeV}$	1120.66 $\pm$ 1.87	179.58 $\pm$ 0.29	19.42 $\pm$ 0.19	12.23 $\pm$ 0.10	67.51 $\pm$ 0.53	5.07 $\pm$ 0.08	23.07 $\pm$ 3.18	1427.53 $\pm$ 3.74	8.21 $\pm$ 0.10	0.22 $\pm$ 0.00	1531	1.07 $\pm$ 0.03
1-SFOS	11.81 $\pm$ 0.19	0.89 $\pm$ 0.02	0.22 $\pm$ 0.02	5.51 $\pm$ 0.07	34.00 $\pm$ 0.38	2.72 $\pm$ 0.06	6.94 $\pm$ 1.50	62.09 $\pm$ 1.56	4.06 $\pm$ 0.07	0.50 $\pm$ 0.01	0	0.00
$\Delta R_{01} > 0.2$	11.81 $\pm$ 0.19	0.89 $\pm$ 0.02	0.22 $\pm$ 0.02	5.51 $\pm$ 0.07	34.00 $\pm$ 0.38	2.72 $\pm$ 0.06	6.95 $\pm$ 1.50	62.10 $\pm$ 1.56	4.05 $\pm$ 0.07	0.50 $\pm$ 0.01	0	0.00
$N_{b-jet} = 0$	11.54 $\pm$ 0.19	0.87 $\pm$ 0.02	0.20 $\pm$ 0.02	5.34 $\pm$ 0.07	5.45 $\pm$ 0.14	0.38 $\pm$ 0.02	5.27 $\pm$ 1.25	29.05 $\pm$ 1.27	3.99 $\pm$ 0.07	0.70 $\pm$ 0.02	0	0.00
$m_{T\bar{T}} < 57 \text{ GeV}$	11.52 $\pm$ 0.19	0.87 $\pm$ 0.02	0.20 $\pm$ 0.02	5.31 $\pm$ 0.07	5.39 $\pm$ 0.14	0.38 $\pm$ 0.02	5.28 $\pm$ 1.25	28.95 $\pm$ 1.27	3.98 $\pm$ 0.07	0.70 $\pm$ 0.02	0	0.00
2-SFOS	1108.85 $\pm$ 1.86	178.69 $\pm$ 0.29	19.19 $\pm$ 0.19	6.72 $\pm$ 0.07	33.51 $\pm$ 0.37	2.34 $\pm$ 0.05	16.13 $\pm$ 2.81	1365.44 $\pm$ 3.41	4.15 $\pm$ 0.07	0.11 $\pm$ 0.00	0	0.00
$\Delta R_{01} > 0.2$	1108.40 $\pm$ 1.86	178.69 $\pm$ 0.29	19.16 $\pm$ 0.19	6.72 $\pm$ 0.07	33.50 $\pm$ 0.37	2.33 $\pm$ 0.05	16.15 $\pm$ 2.81	1364.96 $\pm$ 3.40	4.14 $\pm$ 0.07	0.11 $\pm$ 0.00	0	0.00
$N_{b-jet} = 0$	1084.16 $\pm$ 1.85	176.85 $\pm$ 0.29	18.19 $\pm$ 0.18	6.51 $\pm$ 0.07	5.40 $\pm$ 0.14	0.31 $\pm$ 0.02	14.99 $\pm$ 2.70	1306.40 $\pm$ 3.29	4.05 $\pm$ 0.07	0.11 $\pm$ 0.00	0	0.00
$m_{T\bar{T}} < 57 \text{ GeV}$	1082.58 $\pm$ 1.85	176.44 $\pm$ 0.29	18.16 $\pm$ 0.18	6.47 $\pm$ 0.07	5.38 $\pm$ 0.14	0.31 $\pm$ 0.02	14.90 $\pm$ 2.69	1304.24 $\pm$ 3.29	4.05 $\pm$ 0.07	0.11 $\pm$ 0.00	0	0.00
$m_{4\ell} \notin [115, 130] \text{ GeV}$	1074.15 $\pm$ 1.84	175.99 $\pm$ 0.29	14.34 $\pm$ 0.16	6.46 $\pm$ 0.07	5.35 $\pm$ 0.14	0.30 $\pm$ 0.02	12.42 $\pm$ 2.58	1289.00 $\pm$ 3.19	3.98 $\pm$ 0.07	0.11 $\pm$ 0.00	0	0.00
$10 \text{ GeV} < m_{t\bar{t}} < 50 \text{ GeV}$	116.86 $\pm$ 0.57	6.59 $\pm$ 0.06	1.36 $\pm$ 0.05	0.75 $\pm$ 0.03	0.85 $\pm$ 0.06	0.10 $\pm$ 0.01	1.88 $\pm$ 1.25	128.38 $\pm$ 1.37	2.96 $\pm$ 0.06	0.26 $\pm$ 0.01	0	0.00

Table 7.2: Cutflow for the ZH pre-selection and up to cuts on NN output. Uncertainties shown are purely statistical. “Fakes(DD)” is the data-driven fakes estimate (section 5). “Signif.” is the expected significance.

7.4. Neural Network Training

The signal and background events in this analysis are separated using a neural network (NN) classifier. Two separate NNs are trained for the 1-SFOS and 2-SFOS signal regions (SRs) respectively. A binary classification approach is used where the signal events are labeled as 1 and the background events as 0. This is preferred over a multi-class classification approach as both SRs are dominated by a single background process i.e. ZZ^* background. The training is performed using the Keras [119] library with TensorFlow [120] backend. The training is done using simulated events for the signal and background processes described in Chapter 4. To ensure that events used for training are not the same as those on which the network is applied, the simulated events are split into two equal halves based on the event number. The NN generated from even (event number) events is applied to odd events and vice versa. The plots shown in the following sections are corresponding to networks trained using even events. No apparent difference was observed when the plots were produced using networks trained on odd events. Finally, the events used for training are selected using the event selection criteria described in Section 7.3. The entire process of NN training is already discussed in Chapter 6. The following subsections provide details specific to the NNs used in this analysis.

7.4.1 Input Variables

The input variables used for training the neural networks are chosen so that they provide maximum discrimination between the signal and background processes. These are the same set of variables as the ones used in the Run 2 version of this analysis [124]. The selected input variables are listed in Table 7.3 along with their descriptions. The variables are either reconstructed objects as defined in Chapter 3 or derived variables calculated from these objects.

Although the final state of signal events in both 1-SFOS and 2-SFOS categories do not contain any jets, the N_{jet} variable will help to discriminate against background processes like $t\bar{t}Z$ and VBS processes. The origin of leptons are different in signal and background processes, which is reflected in the kinematic distributions of leptons. Therefore, the transverse momenta of the four leptons are included as input variables. The missing transverse energy is a crucial variable to distinguish signal events from backgrounds like ZZ and $Z\gamma$ with little to no missing energy. The invariant mass of $\ell_0\ell_1$ lepton pair is highly effective in separating signal from ZZ background in the 2-SFOS category, where this pair is expected to originate from the off-shell Z^* boson in the background process and from the Higgs boson in the signal process. The invariant mass of $\ell_2\ell_3$ lepton pair is also included to help discriminate against backgrounds that do not contain a Z boson in the final state.

The azimuthal angle between ℓ_0 and ℓ_1 denoted by $\Delta\phi_{\ell_0\ell_1}^{\text{boost}}$ in the Higgs boson rest frame, is included to exploit the spin correlations in the signal process as discussed in Section 7.1. The azimuthal angle between the dilepton system (ℓ_0, ℓ_1) and the missing transverse energy is another variable to distinguish between signal and ZZ background.

Variable	Description
N_{jet}	Number of jets
$p_T^{\ell_i} (i = 0, 1, 2, 3)$	Transverse momentum of leptons
E_T^{miss}	Missing transverse energy
$m_{\ell_0 \ell_1}$	Invariant mass of leptons ℓ_0 and ℓ_1
$m_{\ell_2 \ell_3}$	Invariant mass of leptons ℓ_2 and ℓ_3
$\Delta\phi_{\ell_0 \ell_1}^{\text{boost}}$	Azimuthal angle between ℓ_0 and ℓ_1 in the Higgs rest frame
$\Delta\phi_{(\ell_0 \ell_1, E_T^{\text{miss}})}$	Azimuthal angle between the dilepton system (ℓ_0, ℓ_1) and E_T^{miss}
$m_{\tau\tau}$	Invariant mass of the tau pair
$p_T^{4\ell}$	Transverse momentum of the four-lepton system
$m_{4\ell}$	Invariant mass of the four-lepton system

Table 7.3: List of input variables used for training the neural networks in both 1SFOS and 2SFOS categories, along with their descriptions.

The transverse momentum and invariant mass of the four-lepton system are also included as input variables to capture the overall kinematic properties of the event. The invariant mass of the τ pair is used to reduce contributions from ZZ events where one of the Z bosons decays into a τ lepton pair that subsequently decays leptonically. The value of $m_{\tau\tau}$ is obtained using the collinear approximation method [125] which is based on two assumptions: (i) neutrinos and the visible decay products of the τ leptons are collinear, and (ii) the missing transverse energy in the event is solely due to the neutrinos from the τ decays.

A total of 13 input variables are selected. Figure 7.3 and Figure 7.4 shows the distributions of these input variables for both signal and background processes in the 1-SFOS and 2-SFOS categories. Correlation between these input variables is studied to ensure that each variable provides unique information for the classification task and to avoid redundancy. Figure 7.5 and Figure 7.6 show the correlation matrices for the input variables in the 1-SFOS and 2-SFOS categories, respectively.

7.4.2 Hyperparameter Tuning

Following the procedure outlined in Section 6.2, hyperparameter optimization is conducted to identify the optimal configuration for the neural networks used in this analysis. The hyperparameters tuned include the batch size, number of epochs, number of folds (k) for cross-validation, learning rate, number of hidden layers, number of neurons per layer, dropout rate after each layer, activation function, optimizer type, loss function and weight (kernel) initialization method.

The hyperparameter search space is defined with reasonable ranges for each hyper-

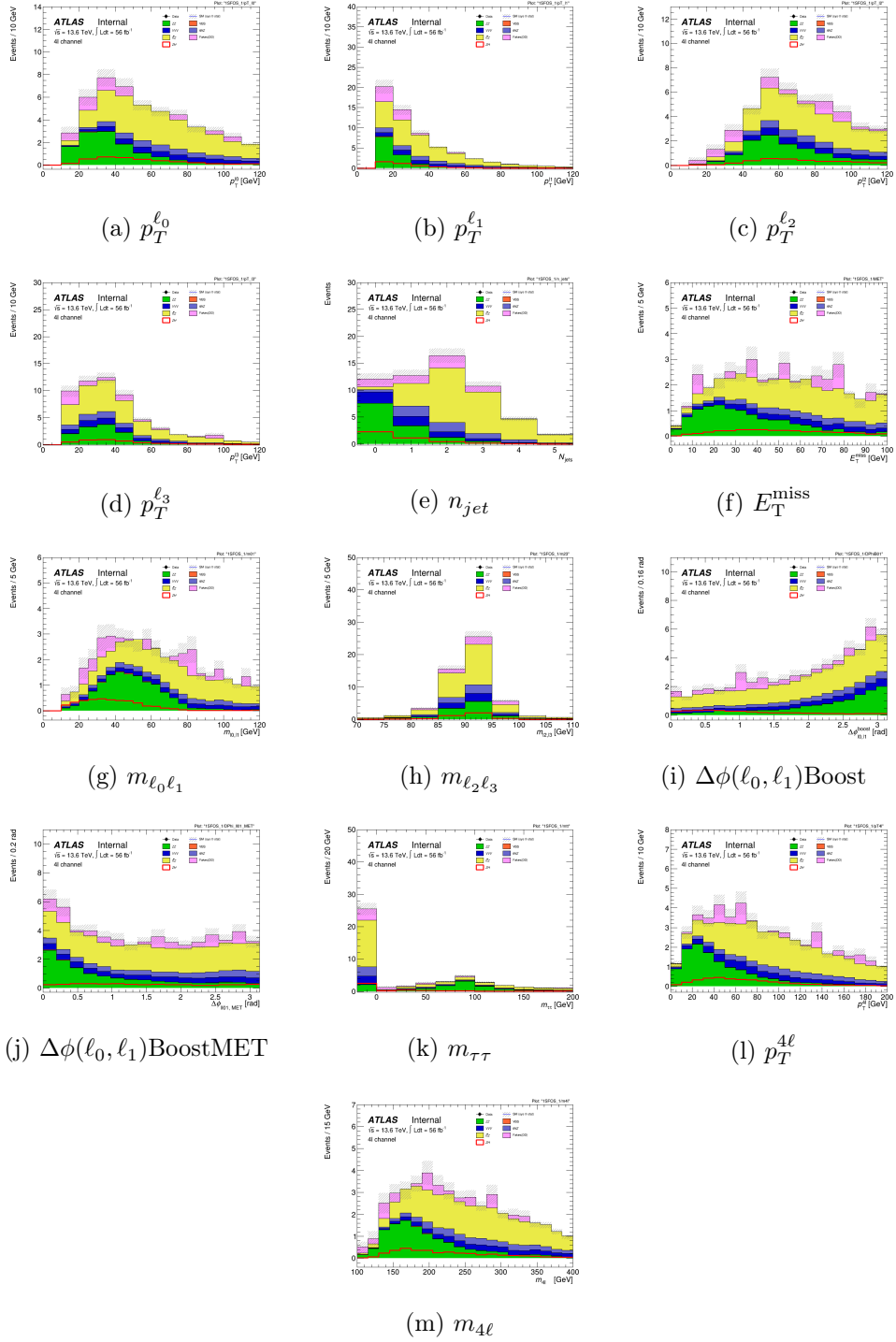


Figure 7.3: Input variable distributions used for training neural-networks in the 1-SFOS SR. Uncertainties shown are MC statistical uncertainties.

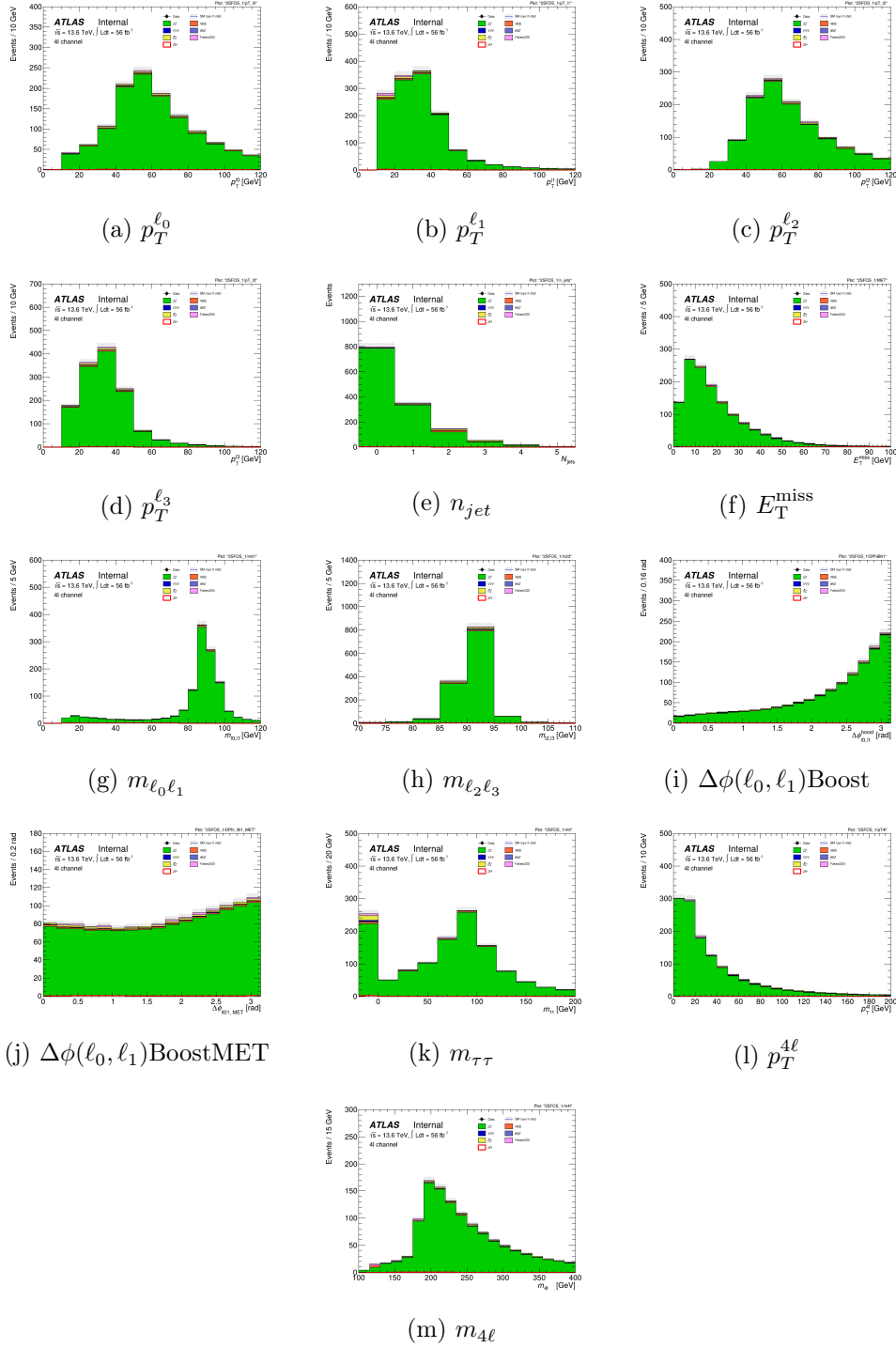
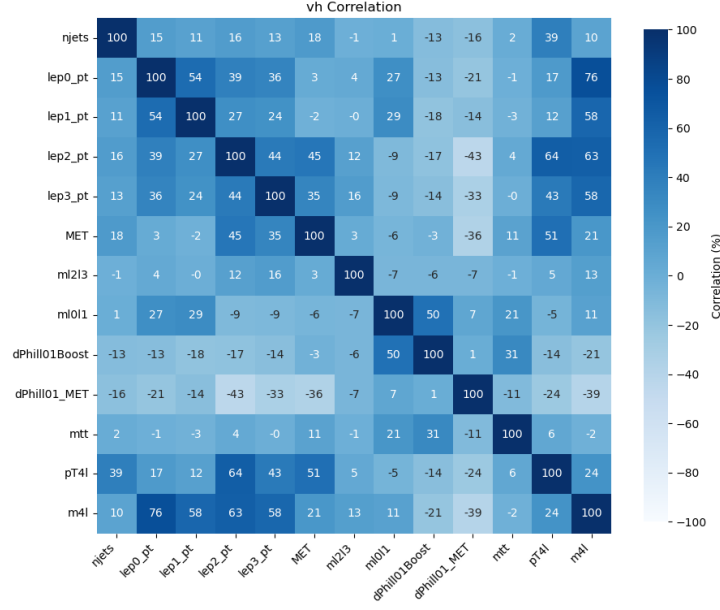
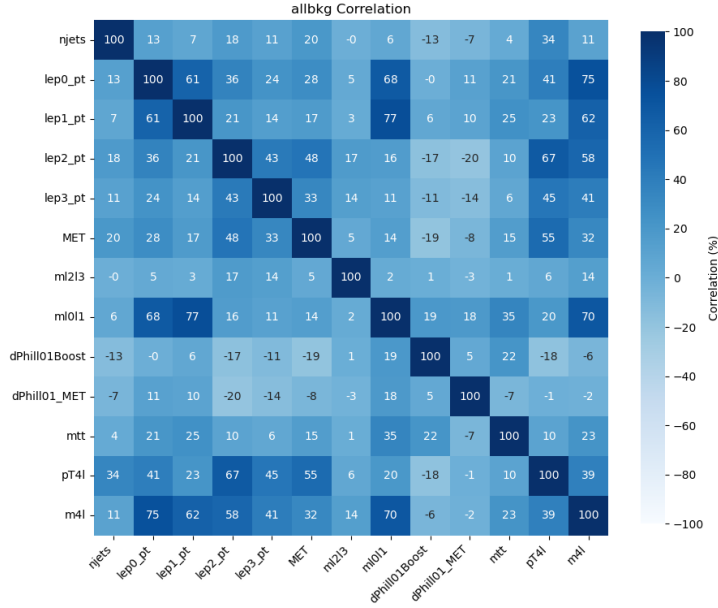


Figure 7.4: Input variable distributions used for training neural-networks in the 2-SFOS SR. Uncertainties shown are MC statistical uncertainties.

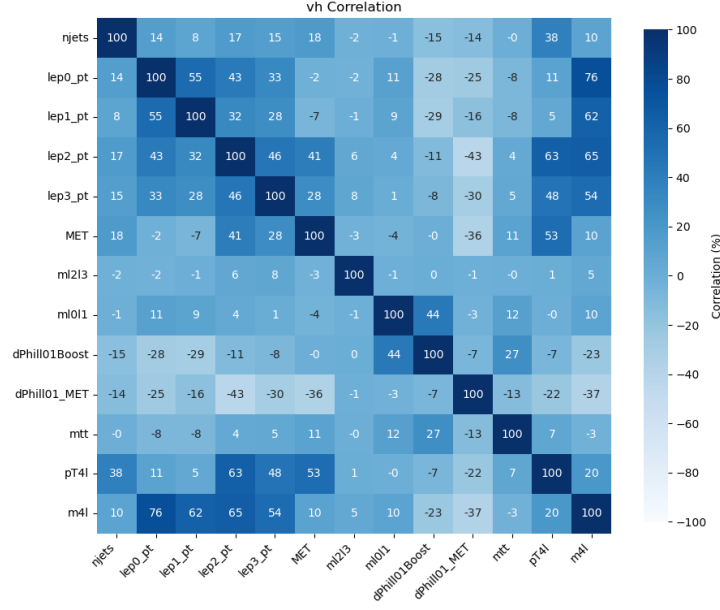


(a) ZH correlation matrix

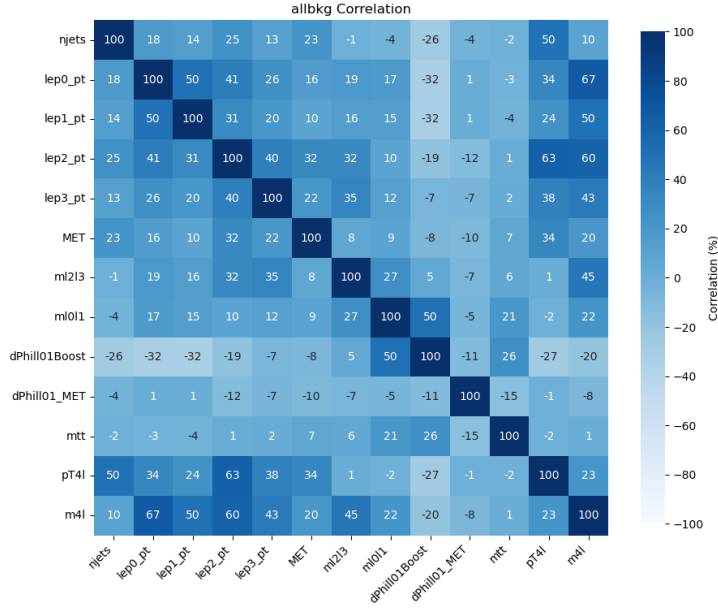


(b) Background correlation matrix

Figure 7.5: Correlation matrices of the 1-SFOS input variables for ZH and background processes.



(a) ZH correlation matrix



(b) Background correlation matrix

Figure 7.6: Correlation matrices of the 2-SFOS input variables for ZH and background processes.

Hyperparameter	Search Space
Batch Size	[1024, 2048, 4096]
Epochs	300
Folds (k)	5
Learning Rate	$[1e^{-5}, 1e^{-6}, 1e^{-7}]$
Hidden Layers	[3, 6]
Neurons per Layer	[1, 2, ..., 256]
Dropout Rate	[0, 0.1, 0.2, 0.3, 0.4, 0.5]
Activation Function	[ReLU, Tanh, Swish]
Optimizer Type	[Adam, Nadam]
Loss Function	Categorical Cross-Entropy
Kernel Initialization	[Glorot Uniform, Glorot Normal]

Table 7.4: Set of hyperparameters and their corresponding search space used for tuning the neural networks for both 1-SFOS and 2-SFOS categories.

parameter and well-known functions are considered for activation functions, optimizers, and loss functions. The tuning process involves training multiple models with different hyperparameter combinations and evaluating their performance using cross-validation to ensure robustness. The dataset is split into training (80%), validation (10%), and test (10%) sets to facilitate this process. Table 7.4 summarizes the hyperparameter search space used in both 1-SFOS and 2-SFOS categories.

7.4.3 Architecture

As discussed earlier, two separate neural networks are trained for the 1-SFOS and 2-SFOS categories. The architecture of the neural networks is finalized based on the hyperparameter tuning results. The final values of the hyperparameters selected for both categories are summarized in Table 7.5.

7.4.4 Model Validation

The final neural network model is validated for overfitting and performance using the techniques discussed in Section 6.3. The area under the receiver operating characteristic (ROC) curve is used to measure the performance of the model using the validation dataset. Overfitting is checked using a 5-fold cross-validation technique. The ROC curves and the distributions for the overfitting checks for 1-SFOS category are shown in Figures 7.7 and 7.8, respectively. The ROC curves from the 5-fold cross-validation technique show consistent performance across all folds. No overfitting is observed as the distributions for training and testing samples are in strong agreement. The ROC curves corresponding to the final model using the full training dataset are shown in Figure 7.9.

Similarly, the ROC curves and the distributions for the overfitting checks for 2-SFOS category are shown in Figures 7.10 and 7.11, respectively. Once again a consistent

Hyperparameter	1SFOS	2SFOS
Batch Size	1024	1024
Epochs	300	300
Folds (k)	5	5
Learning Rate	$1e^{-5}$	$1e^{-5}$
Hidden Layers	6	6
Neurons per Layer	(252, 92, 255, 255, 208, 116)	(198, 251, 180, 117, 179, 48)
Dropout Rate	(0.3, 0.1, 0.2, 0.2, 0, 0.3)	(0.2, 0.1, 0.1, 0.3, 0, 0.2)
Activation Function	(ReLU, Tanh, ReLU, ReLU, Tanh, Swish)	(Swish, ReLU, Tanh, Swish, Tanh, Tanh)
Optimizer Type	Nadam	Adam
Loss Function	Categorical Cross-Entropy	Categorical Cross-Entropy
Kernel Initialization	Glorot Normal	Glorot Normal

Table 7.5: Summary of the final hyperparameter values selected for the neural network training in the 1-SFOS and 2-SFOS SRs.

performance is observed across all folds, with no signs of overfitting. The distributions for training and testing samples show strong agreement. Full training dataset is used to obtain the final model and the corresponding ROC curves are shown in Figure 7.12.

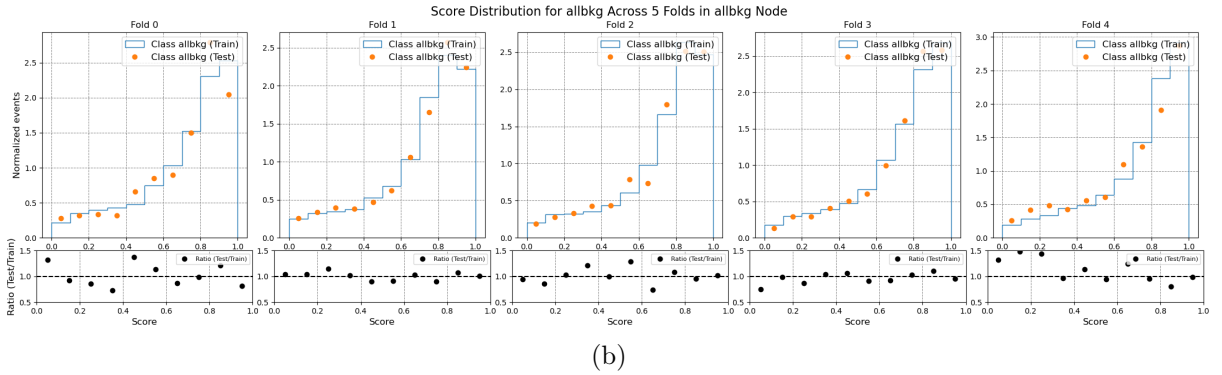
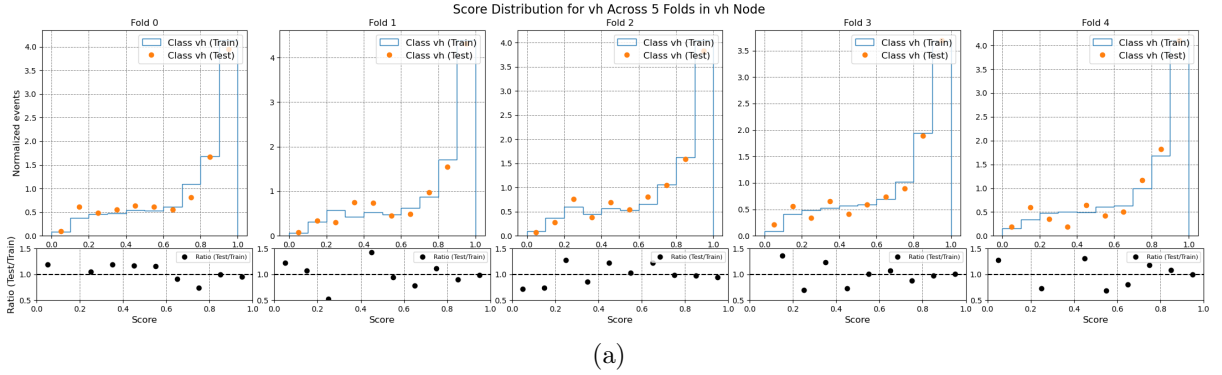


Figure 7.8: Distributions of the NN output scores observed from the 5-fold cross-validation technique in 1-SFOS category.

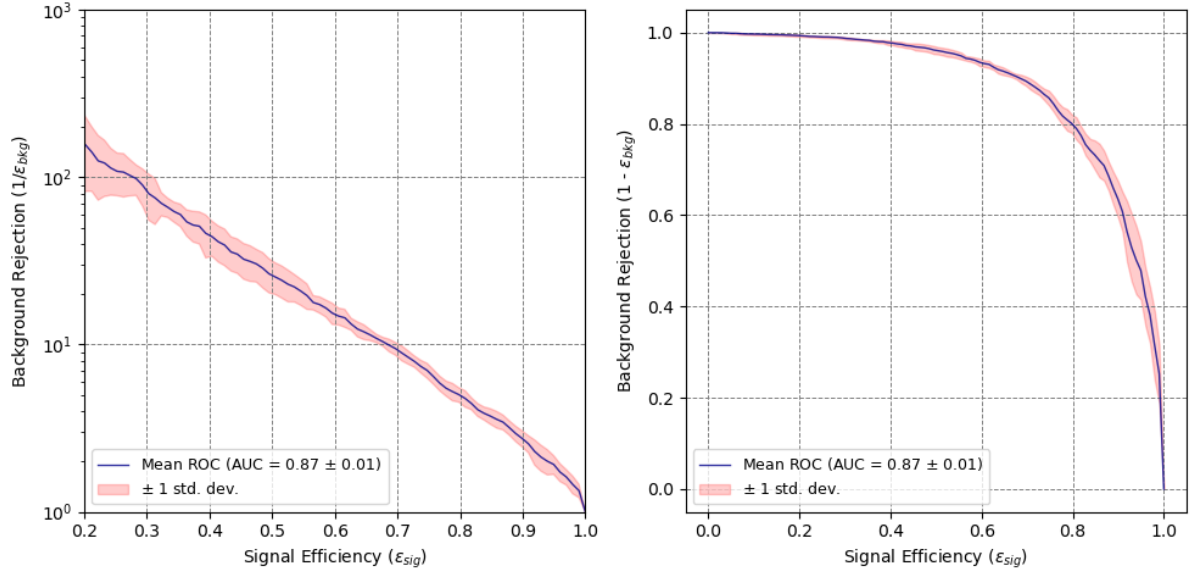


Figure 7.7: ROC curves observed from the 5-fold cross-validation technique in 1-SFOS category.

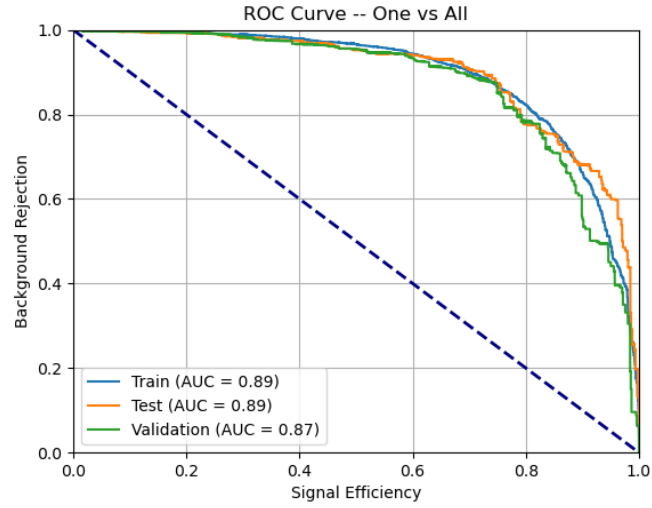


Figure 7.9: ROC Curves for ZH against background node in 1-SFOS category for the final NN trained.

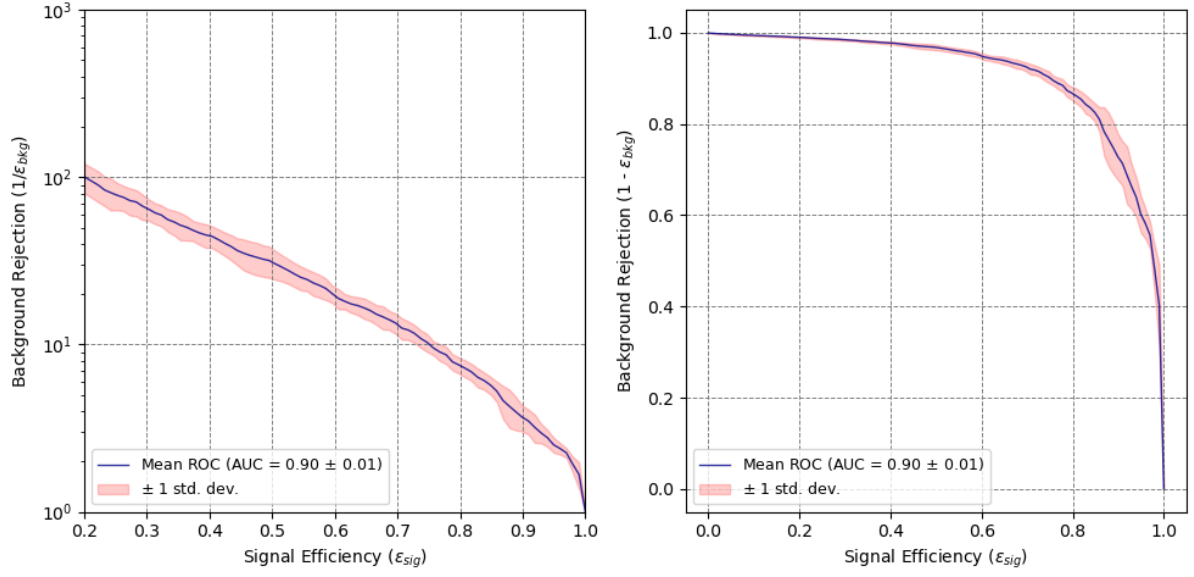


Figure 7.10: ROC curves observed from the 5-fold cross-validation technique in 2-SFOS category.

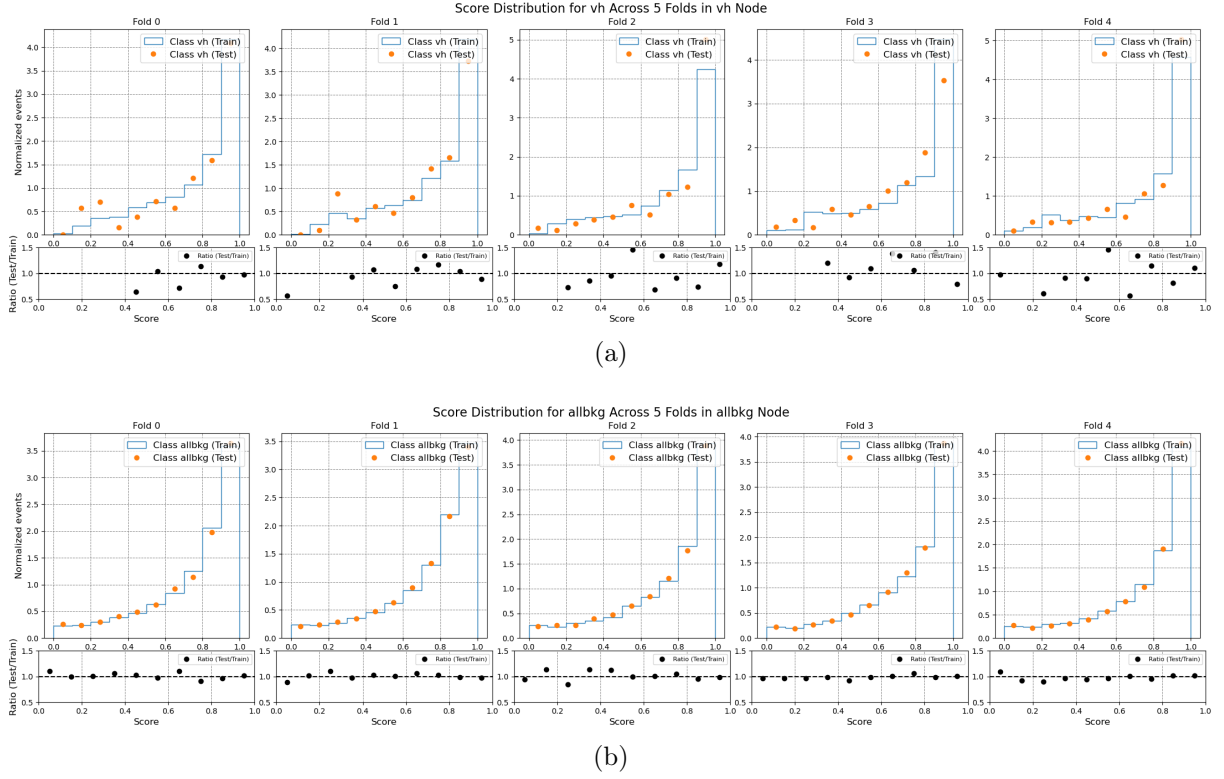


Figure 7.11: Distributions of the NN output scores observed from the 5-fold cross-validation technique in 2-SFOS category.

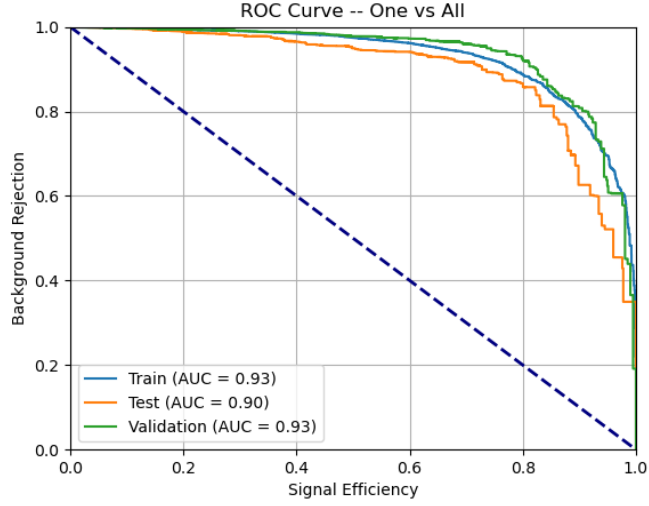


Figure 7.12: ROC Curves for ZH against background node in 2-SFOS category for the final NN trained.

7.4.5 Feature Ranking

As discussed in Section 6.4, each feature is varied individually by ± 1 standard deviation (σ) and their effect on the neural network output score is quantified to calculate the impact of each feature. This is done for both the SRs separately. Figures 7.13 and 7.14 show the feature ranking for the 1-SFOS and 2-SFOS SRs respectively.

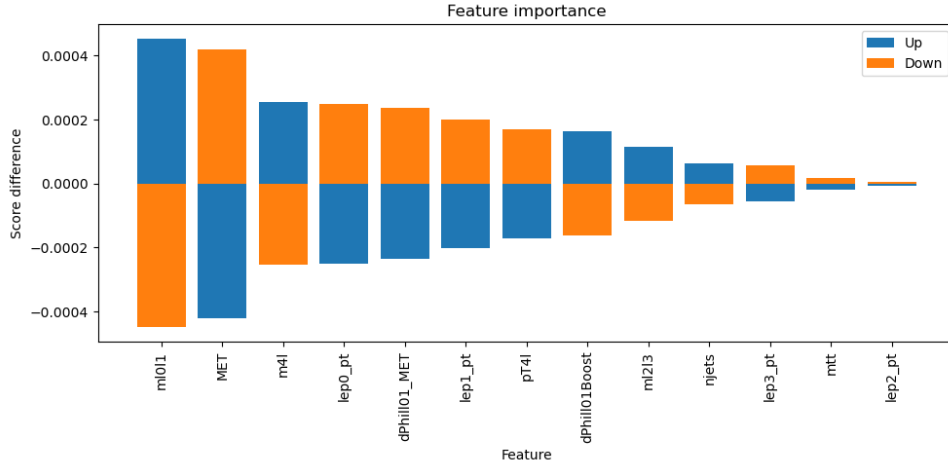


Figure 7.13: Feature ranking for the 1-SFOS signal region. The features are ranked in descending order of their impact on the NN output.

In the 1-SFOS SR, the most important features are the invariant mass of leptons ℓ_0 and ℓ_1 , followed by the missing transverse energy (E_T^{miss}). In the 2-SFOS SR, the most important feature is the missing transverse energy (E_T^{miss}), followed by the transverse

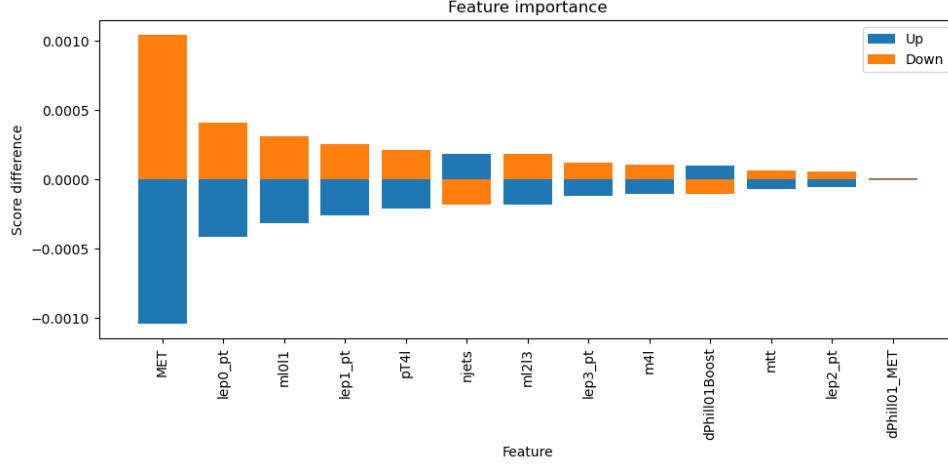


Figure 7.14: Feature ranking for the 2-SFOS signal region. The features are ranked in descending order of their impact on the NN output.

momentum of lepton ℓ_0 and then the invariant mass of leptons ℓ_0 and ℓ_1 . This is consistent with the expectations based on the kinematic distributions discussed in Section 7.4.1.

7.4.6 Output

As discussed in Section 7.4, the neural networks are trained separately for the 1-SFOS and 2-SFOS SRs. Each of the SR further has two separate networks trained for the even and odd events respectively. Applying the even event trained network to the odd events and vice versa gives the NN output score for all events in the SR. The NN output is chosen as the final discriminant for the signal extraction fit as both the SRs are dominated by only one dominant background i.e. ZZ^* background. The remaining backgrounds are small in comparison. Figures 7.15 and 7.16 shows the NN output distributions for the 1-SFOS and 2-SFOS SRs using all events.

7.5. Control and Validation Regions

To understand the contribution from dominant background process, it is useful to create a dedicated control region that is rich in the dominant background events and is orthogonal to the signal region. Orthogonality implies complete isolation and independence from the signal region. This is achieved by applying selection criteria that are distinct from those used in the signal region. The control region is utilized to validate the Monte Carlo (MC) simulation of the background process and to derive normalization factors that can be applied to the background estimation in the signal region.

Section 7.5.1 describes the control region for the ZZ process, which is the dominant background in this analysis. In addition to the control region, validation regions are established to further assess the accuracy of the MC modeling for specific background

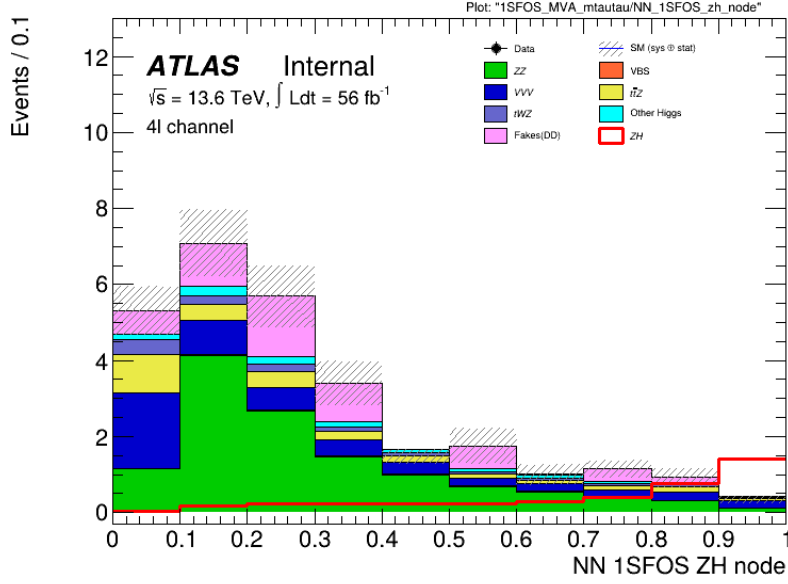


Figure 7.15: Neural Network output distribution for the 2-SFOS signal region.

processes. These validation regions are also designed to be orthogonal to both the signal and control regions, ensuring that they provide an independent check of the background estimations. Sections 7.5.2 and 7.5.3 describe the validation regions for the ZZ and $t\bar{t}Z$ processes, respectively.

7.5.1 ZZ Control Region

The ZZ control region (CR) is designed to be enriched in ZZ events, which constitute the dominant background in this analysis. It contributes approximately around 50% and 97% of the total background in the 1-SFOS and 2-SFOS categories, respectively, as shown in Table 7.7. The main source of ZZ events in this analysis is the production of two Z bosons via quark-antiquark annihilation ($qq \rightarrow ZZ$) and gluon-gluon fusion ($gg \rightarrow ZZ$). 90% of the ZZ events originate from the $qq \rightarrow ZZ$ process, while the remaining 10% come from the $gg \rightarrow ZZ$ process. Additionally, a smaller contribution to this process comes from the production of ZZ in association with jets ($ZZjj$) also known as vector boson scattering (VBS). This account for less than 1% of the total ZZ background in both categories.

A CR enriched in ZZ events for the 2-SFOS category is established, henceforth referred to as the ZZ CR. The ZZ CR uses the same pre-selection criteria as the signal region as summarized in Table 7.1 and shares common selection cuts as the 2-SFOS SR upto the b -jet veto. Following the b -jet veto, additional cuts are applied to increase the purity of ZZ events in the CR. Table 7.6 summarizes the cuts defining the ZZ CR. The invariant mass of the ℓ_2 and ℓ_3 leptons is required to be within 10 GeV of the Z boson mass to select events originating from Z boson decays. An additional requirement of $m_{\ell_0\ell_1} > 50$ GeV is imposed to reject potential ZH signal events, ensuring orthogonality with

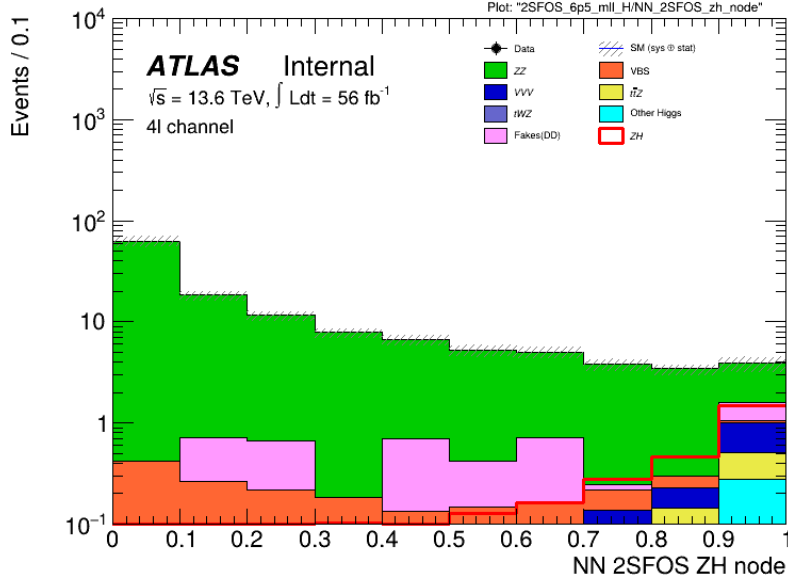


Figure 7.16: Neural Network output distribution for the 2-SFOS signal region.

Processes	Type	Category	Z selection	ZH rejection
$qq \rightarrow ZZ, gg \rightarrow ZZ$	CR	2-SFOS	$ m_{\ell_2 \ell_3} - m_Z < 10 \text{ GeV}$	$m_{\ell_0 \ell_1} > 50 \text{ GeV}$
	VR	1-SFOS	$m_{\tau\tau} > 60 \text{ GeV}$	

Table 7.6: Summary of the selections defining the ZZ enriched regions in the 1-SFOS and 2-SFOS categories. These requirements are applied after the b-veto selection from Table 7.1.

the signal region. The value of $m_{\ell_0 \ell_1}$ is typically lower for ZH signal events.

Table 7.7 shows the expected and observed event yields in the ZZ CR for the 2-SFOS category. The ZZ CR has purity of approximately 98%, with negligible contributions from other background processes, mainly from fake processes that contribute around 1%. A profile likelihood fit is performed to extract the normalization factor for the ZZ background using data from the 2-SFOS category of ZZ CR as discussed in Section 9.2. The NN output score distribution in the 2-SFOS category of the ZZ CR is presented in Figure 7.17. The data is found to be in good agreement with the background predictions within the uncertainties. The distributions of various kinematic variables in the ZZ CR for the 2-SFOS categories are shown in Figure 7.18 and Figure 7.19.

$\sqrt{s} = 13.6 \text{ TeV}, \mathcal{L} = 56 \text{ fb}^{-1} \text{ (EarlyRun3)}$	$qq \rightarrow ZZ$	$gg \rightarrow ZZ$	VBS	VVV	$t\bar{t}Z$	$t\bar{t}H$	Fakes(DD)	bkg	ZH	Data	Data/MC
Pre-selection	1120.66 $\pm$ 1.87	179.58 $\pm$ 0.29	19.42 $\pm$ 0.19	12.23 $\pm$ 0.10	56.34 $\pm$ 0.52	11.17 $\pm$ 0.13	23.07 $\pm$ 3.18	1427.53 $\pm$ 3.74	8.21 $\pm$ 0.10	1531	1.07 $\pm$ 0.03
2-SFOS	1108.85 $\pm$ 1.86	178.69 $\pm$ 0.29	19.19 $\pm$ 0.19	6.72 $\pm$ 0.07	28.00 $\pm$ 0.36	5.52 $\pm$ 0.09	16.13 $\pm$ 2.81	1365.44 $\pm$ 3.41	4.15 $\pm$ 0.07	1447	1.06 $\pm$ 0.03
$\Delta R_{01} > 0.2$	1108.40 $\pm$ 1.86	178.69 $\pm$ 0.29	19.16 $\pm$ 0.19	6.72 $\pm$ 0.07	27.98 $\pm$ 0.36	5.52 $\pm$ 0.09	16.15 $\pm$ 2.81	1364.96 $\pm$ 3.40	4.14 $\pm$ 0.07	1447	1.06 $\pm$ 0.03
$N_{b\text{-jet}} = 0$	1084.16 $\pm$ 1.85	176.85 $\pm$ 0.29	18.19 $\pm$ 0.18	6.51 $\pm$ 0.07	3.92 $\pm$ 0.14	1.48 $\pm$ 0.05	14.99 $\pm$ 2.70	1306.40 $\pm$ 3.29	4.05 $\pm$ 0.07	1380	1.05 $\pm$ 0.03
$m_{\ell_1 \ell_2} > 50 \text{ GeV}, m_{\ell_2 \ell_3} - m_Z < 10 \text{ GeV}$	930.15 $\pm$ 1.72	168.11 $\pm$ 0.29	12.78 $\pm$ 0.15	5.31 $\pm$ 0.06	2.85 $\pm$ 0.12	1.15 $\pm$ 0.04	10.76 $\pm$ 2.19	1131.20 $\pm$ 2.81	0.92 $\pm$ 0.03	1157	1.02 $\pm$ 0.03

Table 7.7: Table showing the expected and observed event yields in the ZZ control region for the 2-SFOS category. The uncertainties shown are statistical only.

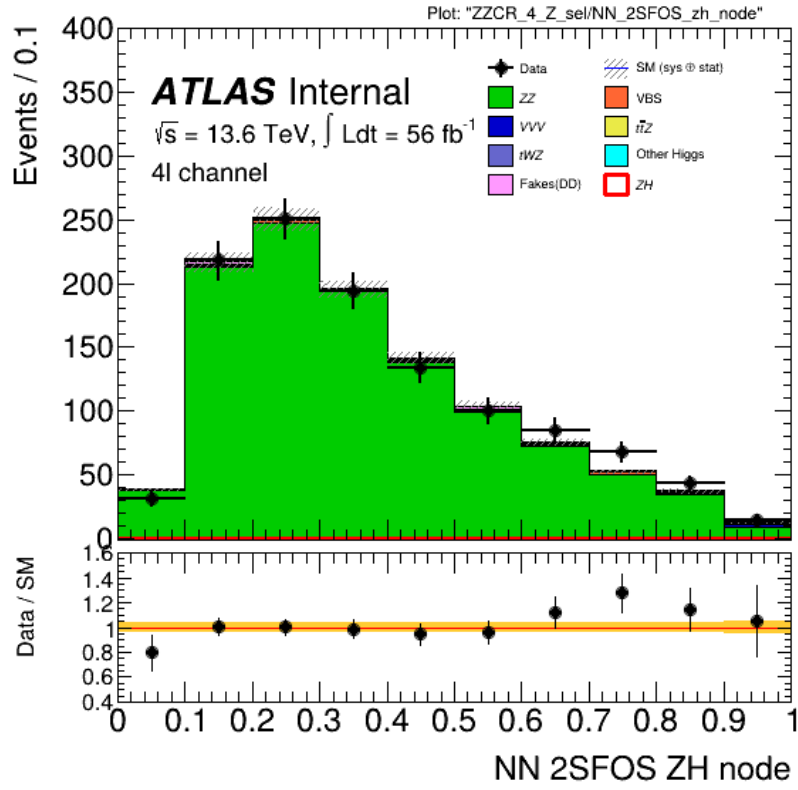


Figure 7.17: The NN output score distribution in the ZZ CR for the 2-SFOS category.

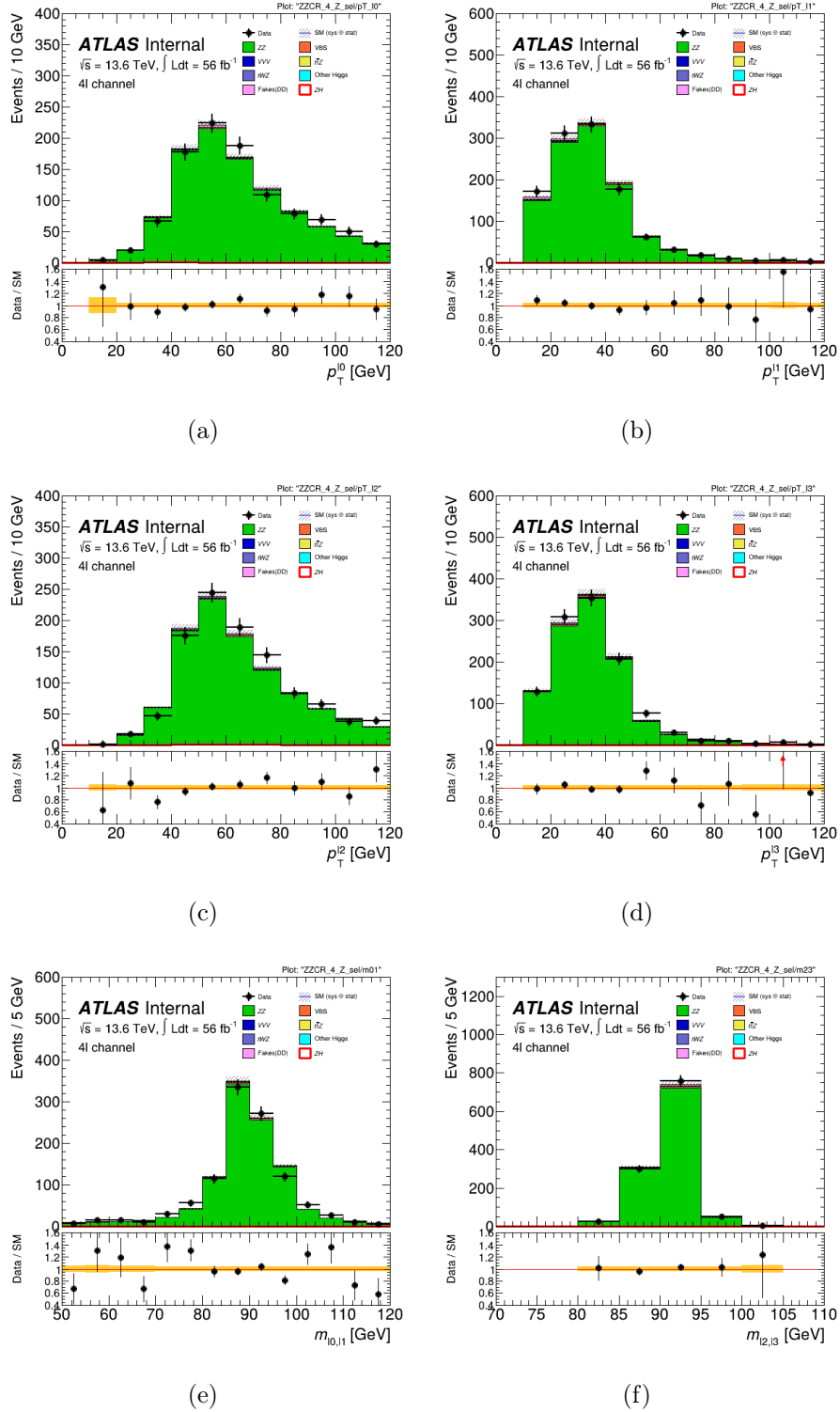
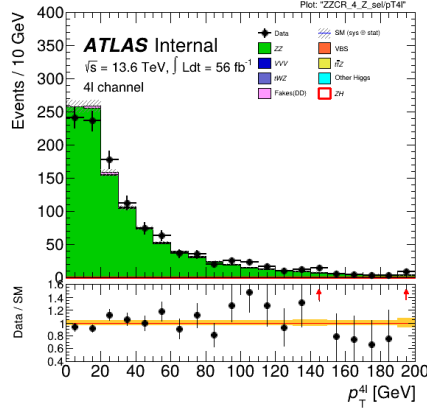
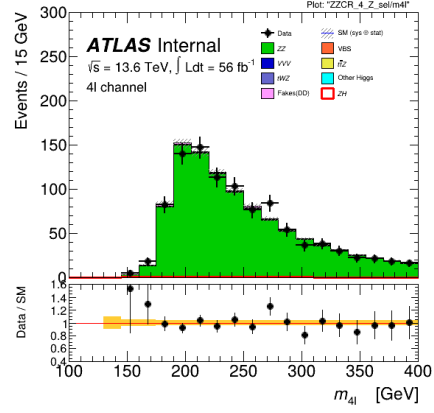


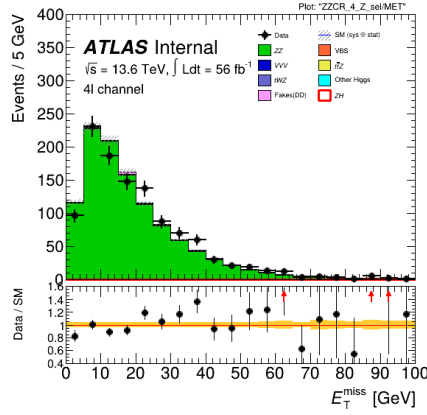
Figure 7.18: Plots showing the Data-MC comparison of various kinematic variables in the ZZ control region in 2-SFOS events: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0\ell_1}$ (f) $m_{\ell_2\ell_3}$. The uncertainties contain both statistical and systematic contributions. The statistical uncertainties are indicated by error bars on the data points and systematic uncertainties are shown with hatching in histograms and a yellow band in ratio plots.



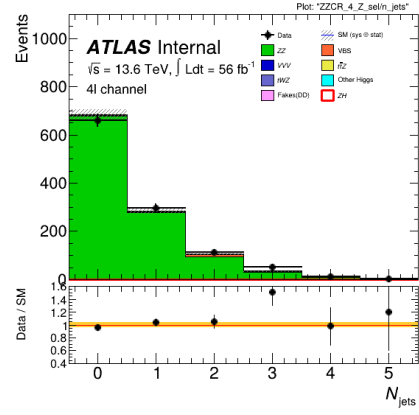
(a)



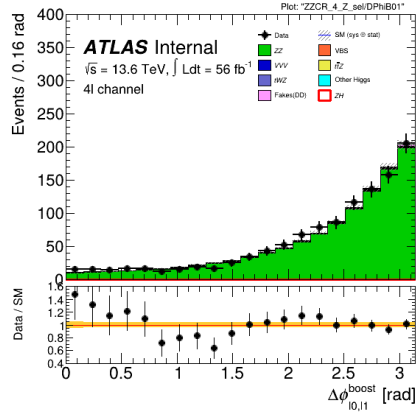
(b)



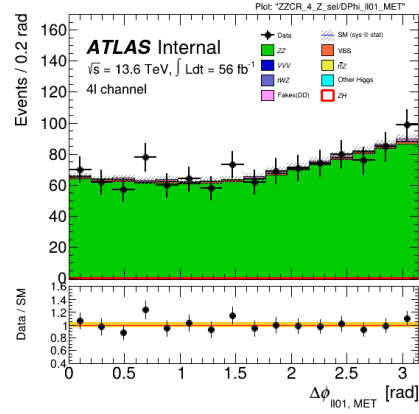
(c)



(d)



(e)



(f)

Figure 7.19: Plots showing the Data-MC comparison of various kinematic variables in the ZZ control region in 2-SFOS events: (a) $p_T^{4\ell}$ (b) $m_{4\ell}$ (c) E_T^{miss} (d) N_{jets} (e) $\Delta\phi_{\ell_0\ell_1}^{\text{boost}}$ (f) $\Delta\phi_{\ell_0\ell_1, \text{MET}}$. The uncertainties contain both statistical and systematic contributions. The statistical uncertainties are indicated by error bars on the data points and systematic uncertainties are shown with hatching in histograms and a yellow band in ratio plots.

7.5.2 ZZ Validation Region

As discussed in Section 7.1, the ZZ process also contributes significantly to the background in the 1-SFOS category. This happens when one of the Z bosons decays into a pair of τ leptons, which subsequently decay leptonically. The pair of leptons originating from the τ decays can form a pair of DFOS leptons while the other Z boson decays into a pair of SFOS leptons. The selection criteria for the ZZ validation region (VR) are similar to those of the 1-SFOS signal region, except for the requirement on the invariant mass of the hypothetical $\tau\tau$ pair ($m_{\tau\tau}$). In the 1-SFOS signal region, an upper cut on $m_{\tau\tau}$ is applied to reject ZZ background events where one of the Z bosons decays into a $\tau\tau$ pair. In contrast, the ZZ VR requires $m_{\tau\tau}$ to be greater than 60 GeV to enrich the region with ZZ events. This selection effectively suppresses potential signal events, which typically exhibit lower values of $m_{\tau\tau}$. Figure 7.20 shows the distribution of $m_{\tau\tau}$ in the 1-SFOS SR after vetoing the b -jet. The selection criteria for the ZZ VR are summarized in Table 7.6.

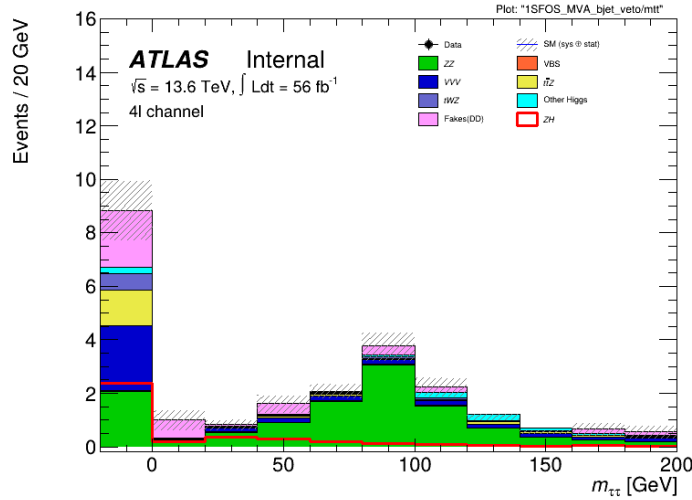


Figure 7.20: Distribution of $m_{\tau\tau}$ in the 1-SFOS category after b -jet veto. Negative value corresponds to the events where no $\tau\tau$ pair was produced.

Table 7.8 shows the contributions from various background processes and data in the ZZ VR for the 1-SFOS category. This region has a purity of approximately 60% for the ZZ events. The contribution from signal events is on the order of the MC uncertainty on the total background events, thus making the signal contamination negligible. The distributions of various kinematic variables in the ZZ VR for the 1-SFOS category are shown in Figure 7.21 and Figure 7.22. In Section 7.5, VR is defined as orthogonal to both the SR and CR. However, in this case, the ZZ VR is not completely orthogonal to the signal region, but given the negligible signal contamination, it is still important to validate the ZZ background modeling in this region.

$\sqrt{s} = 13.6 \text{ TeV}, \mathcal{L} = 56 \text{ fb}^{-1} \text{ (EarlyRun3)}$	$qq \rightarrow ZZ$	$gg \rightarrow ZZ$	VBS	VVV	$t\bar{t}Z$	$t\bar{t}H$	Fakes(DD)	bkg	ZH	Data	Data/MC
Pre-selection	1120.66 $\pm$ 1.87	179.58 $\pm$ 0.29	19.42 $\pm$ 0.19	12.23 $\pm$ 0.10	56.34 $\pm$ 0.52	11.17 $\pm$ 0.13	23.07 $\pm$ 3.18	1427.53 $\pm$ 3.74	8.21 $\pm$ 0.10	1531	1.07 $\pm$ 0.03
1-SFOS	11.81 $\pm$ 0.19	0.89 $\pm$ 0.02	0.22 $\pm$ 0.02	5.51 $\pm$ 0.07	28.35 $\pm$ 0.37	5.65 $\pm$ 0.09	6.94 $\pm$ 1.50	62.09 $\pm$ 1.56	4.06 $\pm$ 0.07	84	1.27 $\pm$ 0.14
$\Delta R_{01} > 0.2$	11.81 $\pm$ 0.19	0.89 $\pm$ 0.02	0.22 $\pm$ 0.02	5.51 $\pm$ 0.07	28.34 $\pm$ 0.37	5.65 $\pm$ 0.09	6.95 $\pm$ 1.50	62.10 $\pm$ 1.56	4.05 $\pm$ 0.07	84	1.27 $\pm$ 0.14
$N_{b\text{-jet}} = 0$	11.54 $\pm$ 0.19	0.87 $\pm$ 0.02	0.20 $\pm$ 0.02	5.34 $\pm$ 0.07	3.86 $\pm$ 0.14	1.59 $\pm$ 0.05	5.27 $\pm$ 1.25	29.05 $\pm$ 1.27	3.99 $\pm$ 0.07	49	1.48 $\pm$ 0.22
$m\tau\tau > 60 \text{ GeV}$	8.17 $\pm$ 0.16	0.56 $\pm$ 0.02	0.12 $\pm$ 0.01	2.56 $\pm$ 0.05	1.77 $\pm$ 0.09	0.68 $\pm$ 0.03	1.73 $\pm$ 0.80	15.71 $\pm$ 0.82	0.73 $\pm$ 0.03	27	1.64 $\pm$ 0.33

Table 7.8: Table showing the expected and observed event yields in the ZZ validation region for the 1-SFOS category. The uncertainties shown are statistical only. No normalization factor is applied.

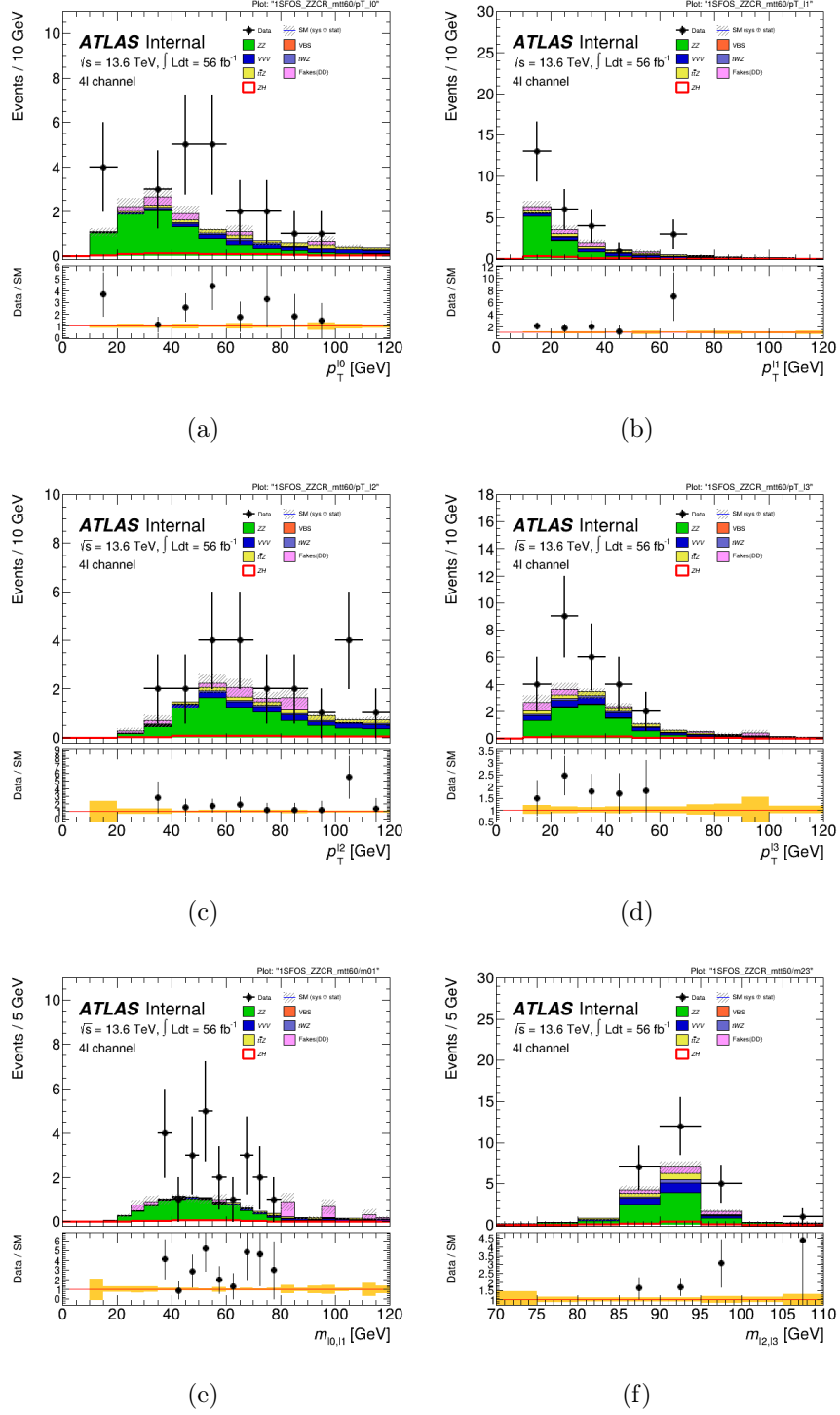


Figure 7.21: Distributions of observables in the ZZ validation region in 1-SFOS events: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0\ell_1}$ (f) $m_{\ell_2\ell_3}$. The statistical uncertainties are indicated by error bars whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.

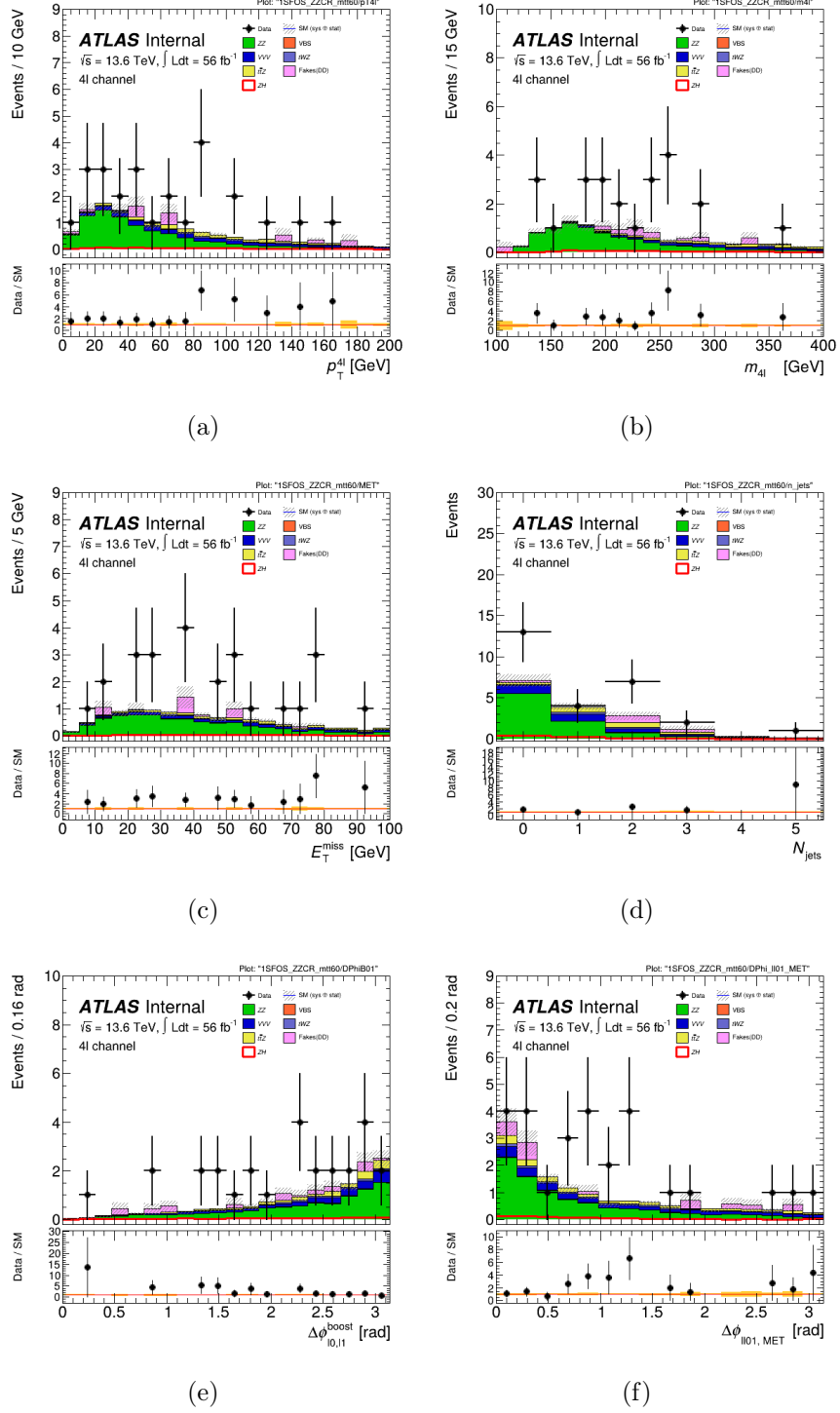


Figure 7.22: Distributions of observables in the ZZ validation region in 1-SFOS events: (a) p_T^{4l} (b) m_{4l} (c) E_T^{miss} (d) N_{jets} (e) $\Delta\phi_{l0,l1}^{boost}$ (f) $\Delta\phi_{l0,l1,MET}$. The statistical uncertainties are indicated by error bars whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.

7.5.3 $t\bar{t}Z$ Validation Region

The $t\bar{t}Z$ process is another significant background in the analysis, particularly in the 1-SFOS category, before applying the b -jet veto. To validate the modeling of the $t\bar{t}Z$ background, a dedicated validation region (VR) is defined. To select events enriched in $t\bar{t}Z$ while ensuring orthogonality to the signal region, the $t\bar{t}Z$ requires events to have at least two jets, with at least one of them being b -tagged. This selection effectively suppresses potential signal events, which typically do not contain b -tagged jets due to the b -jet veto applied in the signal region. There is a requirement of at least one SFOS lepton pair to have invariant mass within 10 GeV of the Z boson mass (m_Z). This ensures that the selected events are consistent with the decay of a Z boson into a pair of leptons. If there is an additional SFOS lepton pair, then invariant mass of the l_0, l_1 lepton pair is required to be outside the Z mass window by 10 GeV to reject the contribution from potential 2-SFOS ZZ events. In addition, the invariant mass $m_{l_0 l_1}$ is required to be greater than 40 GeV if it is a SFOS lepton pair. The selection criteria for the $t\bar{t}Z$ VR are summarized in Table 7.9.

Table 7.9: Selection criteria defining the $t\bar{t}Z$ VR.

Validation Region	Selection
$t\bar{t}Z$	$N_{jet} \geq 2$ $N_{b\text{-tag}} \geq 2$ require at least one SFOS pair within 10 GeV of m_Z if there is another SFOS pair, require $ m_{\ell_0 \ell_1} - m_Z \geq 10$ GeV $m_{\ell_0 \ell_1} > 40$ GeV for 2-SFOS events

Table 7.10 shows the contributions from the MC background processes and data in the $t\bar{t}Z$ VR. This region has a purity of approximately 73% for the $t\bar{t}Z$ events, followed by the contribution from tWZ events which is around 10%. The contribution from signal events is less than 0.2% of the total events due to the requirements on number of jets and b -tagged jets, thus making the signal contamination negligible. The ratio of data to MC comes out to be 1.22 ± 0.17 , considering only the MC statistical uncertainties. Due to the b -jet veto in the signal region, the contribution of $t\bar{t}Z$ in the signal region is significantly suppressed. In the Table 7.2, which shows the event yields before applying the NN score cut in both the 1-SFOS and 2-SFOS signal regions, the contribution from $t\bar{t}Z$ is second lowest after the final cut in each SR. Therefore, the $t\bar{t}Z$ VR is not included in the profile likelihood fit to extract the normalization factor for the $t\bar{t}Z$ background. Figure 7.23 and Figure 7.24 shows the distributions of various kinematic variables.

$\sqrt{s} = 13.6 \text{ TeV}, \mathcal{L} = 56 \text{ fb}^{-1} \text{ (EarlyRun3)}$	$qq \rightarrow ZZ$	$gg \rightarrow ZZ$	VBS	VVV	$t\bar{t}Z$	tWZ	$t\bar{t}H$	Fakes(DD)	bkg	ZH	Data	Data/MC
Pre-selection	1120.66 ± 1.87	179.58 ± 0.29	19.42 ± 0.19	12.23 ± 0.10	56.34 ± 0.52	11.17 ± 0.13	5.07 ± 0.08	23.07 ± 3.18	1427.53 ± 3.74	8.21 ± 0.10	1531	1.07 ± 0.03
$N_{b-jets} \geq 1$	28.81 ± 0.26	2.57 ± 0.04	1.23 ± 0.06	0.49 ± 0.02	48.04 ± 0.48	8.02 ± 0.11	4.34 ± 0.07	2.68 ± 1.13	96.18 ± 1.27	0.19 ± 0.02	102	1.06 ± 0.11
$N_{jets} \geq 2$	16.59 ± 0.16	1.49 ± 0.03	1.13 ± 0.05	0.35 ± 0.02	42.36 ± 0.45	5.57 ± 0.09	3.83 ± 0.07	1.99 ± 1.02	73.30 ± 1.13	0.11 ± 0.01	85	1.16 ± 0.13
$ m_{\ell\ell_3} - m_Z < 10 \text{ GeV}$	15.55 ± 0.15	1.45 ± 0.03	1.03 ± 0.05	0.32 ± 0.02	35.57 ± 0.41	5.07 ± 0.09	1.23 ± 0.04	2.35 ± 0.95	62.56 ± 1.05	0.09 ± 0.01	72	1.15 ± 0.14
$ m_{\ell\ell_1} - m_Z > 10 \text{ GeV}$	4.53 ± 0.08	0.33 ± 0.01	0.38 ± 0.03	0.28 ± 0.02	33.33 ± 0.40	4.77 ± 0.08	1.18 ± 0.04	1.79 ± 0.86	46.58 ± 0.96	0.09 ± 0.01	59	1.26 ± 0.17
$m_{\ell\ell_1} > 40 \text{ GeV}$	3.59 ± 0.07	0.30 ± 0.01	0.23 ± 0.03	0.26 ± 0.02	31.64 ± 0.39	4.52 ± 0.08	1.02 ± 0.04	1.86 ± 0.83	43.43 ± 0.92	0.06 ± 0.01	53	1.22 ± 0.17

Table 7.10: Expected and observed event yields in the $t\bar{t}Z$ validation region. The uncertainties shown are statistical only.

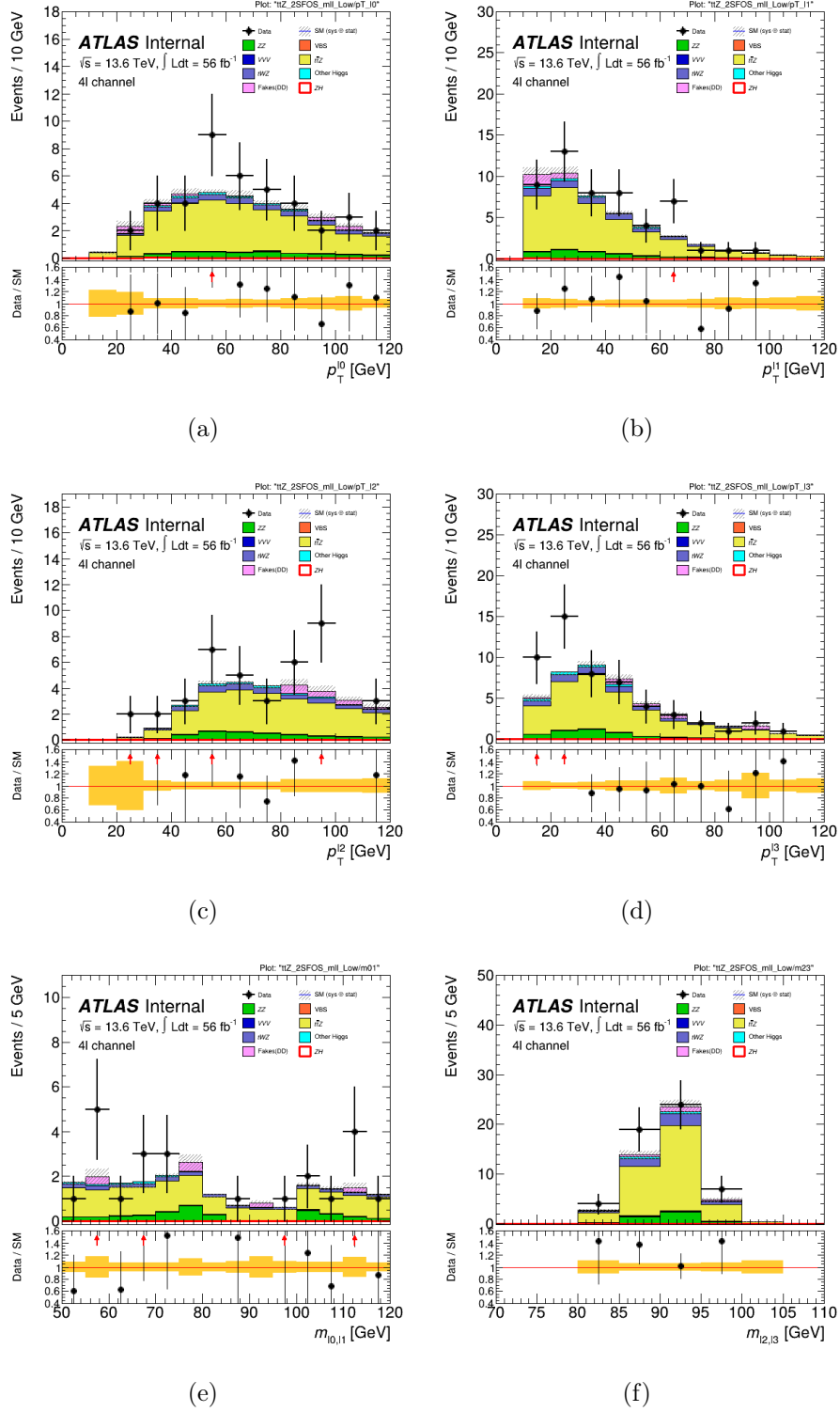


Figure 7.23: Distribution of observables in the $t\bar{t}Z$ validation region: (a) $p_T^{\ell_0}$ (b) $p_T^{\ell_1}$ (c) $p_T^{\ell_2}$ (d) $p_T^{\ell_3}$ (e) $m_{\ell_0 \ell_1}$ (f) $m_{\ell_2 \ell_3}$. The statistical uncertainties are shown as error bars on the data points whereas systematic uncertainties are shown in hatching in histograms and a yellow band in ratio plots.

8 Systematic Uncertainties

Systematic uncertainties can arise from both experimental and theoretical sources, and they can impact both the signal and background estimations. The experimental uncertainties stem from the mismodelling of the detector response, inefficiencies in object reconstruction and uncertainties in luminosity measurement as discussed in Section 8.1. These uncertainties can affect the normalization and shape of the signal and background distributions used in the analysis. Theoretical uncertainties, on the other hand, arise from inaccuracies in modelling the signal and background processes using Monte Carlo (MC) simulations and uncertainties in calculating the cross-section due to missing higher-order corrections as discussed in Section 8.2. These systematic uncertainties are incorporated into the final statistical analysis as nuisance parameters, which are varied within their uncertainties during the fitting procedure. The impact of each systematic uncertainty on the final measurement is evaluated by varying the corresponding nuisance parameter and assessing the change in the signal strength or cross-section measurement, discussed in Chapter 9.

8.1. Experimental Systematic Uncertainties

Experimental systematic uncertainties correspond to uncertainties arising from the detector performance, object reconstruction, and calibration. As discussed earlier, these uncertainties can affect both the normalization and shape of the signal and background distributions. Although there are many sources of experimental systematic uncertainties, they can be broadly categorized into two types:

1. Four-vector (P4) uncertainties: These uncertainties stem from inaccuracies in the measurement of the energy and momentum of reconstructed objects. They directly affect the four-momentum of objects such as leptons, jets, and missing transverse energy (E_T^{miss}), influencing kinematic distributions. The ± 1 standard deviation variations of the nuisance parameters associated with these uncertainties are used in calculating the uncertainties.
2. Scale factor (SF) uncertainties: These uncertainties arise from scale factors used to correct for discrepancies between data and simulation in object reconstruction, identification, and isolation efficiencies. They ensure that the simulation accurately reflects the performance observed in data. The uncertainties are derived from the weight variations applied to simulated events.

A dedicated combined performance group at ATLAS derives these uncertainties. The main sources of experimental systematic uncertainties considered in this analysis are discussed below and also summarized in Table 8.1.

- **Luminosity:** The uncertainty in the integrated luminosity measurement affects the overall normalization for all processes. The luminosity uncertainty is determined using dedicated calibration procedures and is provided by the ATLAS luminosity working group. For this analysis, a luminosity uncertainty of 2.0% is considered for the combination of 2022 and 2023 data [104].
- **Pile-up reweighting:** Pile-up reweighting is used to correct for differences in the pile-up distribution between data and simulation. The uncertainties in the pile-up modeling are typically estimated by varying the pile-up reweighting factors within their uncertainties and assessing the impact on the signal and background estimations.
- **Leptons:** The uncertainties associated with lepton (electron and muon) reconstruction, identification, isolation efficiencies, energy/momentum scale, and resolution are considered.
- **Jets:** Although jets are not explicitly used in the event selection for this analysis, they can still affect the missing transverse energy (E_T^{miss}) calculation. The jet energy scale (JES) and jet energy resolution (JER) uncertainties are considered along with jet vertex tagging (JVT) and forward JVT (fJVT) uncertainties.
- **Flavour Tagging:** The uncertainties corresponding to the identification of jets originating from b -quarks (b -tagging) and the misidentification rates of light-flavor jets as b -jets (mistag rate) can impact the background estimations, particularly for processes involving top quarks are considered.
- **Missing transverse energy modeling:** The uncertainties related to the track-based soft term in the E_T^{miss} calculation are considered. The track-based soft term accounts for low-momentum particles not associated with reconstructed objects.

The total experimental uncertainty is obtained by combining the individual uncertainties in quadrature, taking into account correlations where applicable.

Systematic uncertainty	Short description	Type
Events		
Luminosity	uncertainty on total integrated luminosity	SF
Pileup Reweighting	uncertainty on pileup reweighting	SF
Electrons		
EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	SF
EL_EFF_TriggerEff_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	SF
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty	SF
EL_EFF_ID_CorrUncertaintyNP (0 to 15)	ID efficiency uncertainty split in 16 components	SF
EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP (0 to 17)	ID efficiency uncertainty split in 18 components	SF
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty	SF
EG_SCALE_ALL	energy scale uncertainty	P4
EG_SCALE_AF3		P4
EG_RESOLUTION_ALL	energy resolution uncertainty	P4
EG_RESOLUTION_AF3	energy resolution uncertainty	P4
Muons		
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty	SF
MUON_EFF_TrigSystUncertainty	reconstruction and ID efficiency uncertainty for muons with $p_T > 15$ GeV	SF
MUON_EFF_RECO_STAT		SF
MUON_EFF_RECO_STAT_LOWPT		SF
MUON_EFF_RECO_SYS		SF
MUON_EFF_RECO_SYS_LOWPT		SF
MUON_EFF_ISO_STAT		SF
MUON_EFF_ISO_SYS	isolation efficiency uncertainty	SF
MUON_EFF_TTVA_STAT	track-to-vertex association efficiency uncertainty	SF
MUON_EFF_TTVA_SYS		SF
MUON_CB	momentum resolution uncertainty from combined system	P4
MUON_SCALE	momentum scale uncertainty	P4
MUON_SAGITTA_DATASTAT	charge dependent momentum scale uncertainty	P4
MUON_SAGITTA_RESBIAS		P4
MUON_SAGITTA_GLOBAL		P4
MUON_SAGITTA_PTEXTRA		P4
Jets		
JET_EffectiveNP_Detector (1 to 2)	energy scale uncertainty from the in situ analyses splits into 8 components	P4
JET_EffectiveNP_Mixed (1 to 3)		P4
JET_EffectiveNP_Modelling (1 to 4)		P4
JET_EffectiveNP_Statistical (1 to 6)		P4
JET_EtaIntercalibration_Modelling	energy scale uncertainty on eta-intercalibration (modeling)	P4
JET_EtaIntercalibration_TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)	P4
JET_EtaIntercalibration_NonClosure_0p2_PreRec	energy scale uncertainty on eta-intercalibrations (non-closure)	P4
JET_Pileup_OffsetMu	energy scale uncertainty on pile-up (mu dependent)	P4
JET_Pileup_OffsetNPV	energy scale uncertainty on pile-up (NPV dependent)	P4
JET_Pileup_PtTerm	energy scale uncertainty on pile-up (p_T term)	P4
JET_Pileup_RhoTopology	energy scale uncertainty on pile-up (density ρ)	P4
JET_Flavor_Composition	energy scale uncertainty on flavour composition	P4
JET_Flavor_Response	energy scale uncertainty on samples' flavour response	P4
JET_BJES_Response	energy scale uncertainty on b -jets	P4
JET_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high- p_T jets	P4
JET_JER_DataVsMC_MC16	energy resolution uncertainty, each for both MC d pseudo-data	P4
JET_JER_EffectiveNP (1 to 11, 12restTerm)		P4
JET_JERUnc_Noise_PreRec	JVT efficiency uncertainty	P4
JET_JESUnc_Noise_PreRec		P4
JET_JESUnc_VertexingAlg_PreRec		P4
JET_InSitu_NonClosure_PreRec		P4
JET_FCal_MC23_NonClosure_PreRec	fJVT efficiency uncertainty	P4
JET_NNJvtEfficiency		SF
JET_fJvtEfficiency		SF
Flavour Tagging		
FT_EFF_Eigen_B (0 to 2)	b -tagging efficiency uncertainties	SF
FT_EFF_Eigen_C (0 to 3)		SF
FT_EFF_Eigen_Light (0 to 4)		SF
MET		
MET_SoftTrk_ResoPara	track-based soft term related longitudinal resolution uncertainty	P4
MET_SoftTrk_ResoPerp	track-based soft term related transverse resolution uncertainty	P4
MET_SoftTrk_Scale	track-based soft term related longitudinal scale uncertainty	P4

Table 8.1: Summary of various nuisance parameters with their short descriptions and sources of the experimental systematic uncertainties considered in this analysis.

8.2. Theoretical Uncertainties

Theoretical uncertainties stem from the limitations in the theoretical models and calculations used to predict the signal and background processes. The theoretical uncertainties considered in this analysis include several components. For the ZH signal process, uncertainties from parton distribution functions (PDF), the strong coupling constant (α_s) and quantum chromodynamics (QCD) scale variations are calculated based on the recommendations from the Physics Modelling Group (PMG) [126]. The PDFs describe the momentum distribution of quarks and gluons inside the proton, while the QCD scale variations account for the uncertainties due to missing higher-order corrections in perturbative QCD calculations. The PDF4LHC21 [107] set is used to derive the PDF uncertainties and is calculated using the Hessian method which involves evaluating the quadrature sum of differences between the nominal PDF set and each of the error sets (40 in total). The QCD scale uncertainties are estimated by varying the renormalization (μ_R) and factorization (μ_F) scales up and down by a factor of 2 and 0.5 with respect to the nominal value. The α_s uncertainties occur due to the uncertainty in the value of the strong coupling constant, α_s , which affects the prediction of cross-sections and event kinematics. The parton shower modelling uncertainty for signal process is evaluated by comparing the nominal sample generated using PYTHIA8 with an alternative sample generated using HERWIG7.

For the $q\bar{q} \rightarrow ZZ$ background, the PDF, α_s , and QCD scale uncertainties are calculated using the same PMG recommendations as for the signal process except that the Sherpa samples NNPDF3NLO PDF set is used. The parton shower modelling uncertainty is calculated using the resummation and jet merging scale uncertainties. A two point variation is used for the resummation scale (MUQ) which is varied by a factor of 2 relative to the nominal value. The jet merging scale (CKKW) is varied using two values of 15 GeV and 30 GeV while the nominal value is 20 GeV. For the $q\bar{q} \rightarrow ZZ$ background, QCD scale uncertainty is evaluated using the same method as for the signal process. An additional Higgs interference uncertainty will be calculated for the $gg \rightarrow ZZ$ background to account for the interference between the $gg \rightarrow H^* \rightarrow ZZ$ process and the $gg \rightarrow ZZ$ continuum process. This uncertainty is estimated by varying the interference term within its theoretical uncertainties, which are typically derived from higher-order calculations.

For the triboson (VVV) background, PDF, α_s and QCD scale uncertainties are calculated as described earlier. In addition, uncertainties due to electroweak corrections are also considered. Finally, normalization uncertainties are applied to minor background processes such as $t\bar{t}Z$ and $t\bar{t}H$. These are taken as $\frac{\Delta\sigma_{t\bar{t}V}}{\sigma_{t\bar{t}V}} = 13\%$ [127] and $\frac{\Delta\sigma_{t\bar{t}H}}{\sigma_{t\bar{t}H}} = 10\%$ [128], respectively, based on the theoretical cross-section calculations and experimental measurements.

The theoretical uncertainties considered in this analysis and a summary of their impact on the signal and background yields is provided in Table 8.2.

Systematic uncertainty	Region	Value [%]
<i>Signal Processes</i>		
<i>ZH</i>		
PDF Uncertainty	1-SFOS SR	± 1.5
	2-SFOS SR	± 1.5
α_s Uncertainty	1-SFOS SR	∓ 1.5
	2-SFOS SR	∓ 0.5
QCD Uncertainty	1-SFOS SR	± 6.4
	2-SFOS SR	± 6.7
Parton shower modelling	1-SFOS SR	± 0.1
	2-SFOS SR	∓ 1.6
<i>Background Process</i>		
<i>qqZZ</i>		
PDF Uncertainty	1-SFOS SR	± 0.1
	2-SFOS SR	± 0.2
	ZZ CR	± 0.1
α_s Uncertainty	1-SFOS SR	± 1.3
	2-SFOS SR	± 1.3
	ZZ CR	± 1.3
QCD Uncertainty	1-SFOS SR	± 7.9
	2-SFOS SR	± 8.2
	ZZ CR	± 7.5
Resummation scale	1-SFOS SR	∓ 7.7
	2-SFOS SR	∓ 0.6
	ZZ CR	∓ 3.4
Jet merging scale	1-SFOS SR	∓ 9.3
	2-SFOS SR	± 0.6
	ZZ CR	∓ 1.1
<i>ggZZ</i>		
QCD Uncertainty	1-SFOS SR	± 14.6
	2-SFOS SR	± 17.3
	ZZ CR	± 14.9
<i>VVV</i>		
PDF Uncertainty	SRs and ZZ CR	± 0.1
α_s Uncertainty	SRs and ZZ CR	± 1.1
QCD Uncertainty	1-SFOS SR	± 17.3
	2-SFOS SR	± 16.6
	ZZ CR	± 17.9
Electroweak Uncertainty	1-SFOS SR	± 1.4
	2-SFOS SR	± 1.0
	ZZ CR	± 1.5
<i>ttZ</i>		
Normalization Uncertainty	SRs and ZZ CR	± 13.0
<i>ttH</i>		
Normalization Uncertainty	SRs and ZZ CR	± 10.0

Table 8.2: List of theory systematic uncertainties considered in the ZH channel. All uncertainties are given as relative variations on the yields. The yields are assumed to be taken from the signal region unless specified otherwise. Uncertainties from multiple sources are added in quadrature. N.B.: $q\bar{q} \rightarrow ZH$ and $gg \rightarrow ZH$ are combined simply as ZH .

9 Statistical Analysis

In particle physics experiments, any analysis typically involves the comparison of observed data with theoretical predictions to extract meaningful results. The conclusions drawn from such analysis are based on whether the observed data is consistent with the predictions of the Standard Model (SM) or if there are deviations that could indicate new physics phenomena. Statistical methods are employed to quantify the significance of any observed deviations and to estimate the parameters of interest. These methods include hypothesis testing, parameter estimation, and the construction of a test statistic to interpret the results. In this chapter, definitions of key statistical concepts and tools used in the analysis are provided (Section 9.1), along with an overview of the statistical method employed (Section 9.2) and the test statistics (Section 9.3) used in this analysis.

9.1. Basic Definitions

9.1.1 Hypothesis Testing

When performing a statistical analysis, we typically start with two competing hypotheses: the null hypothesis (H_0) and the alternative hypothesis (H_1). For any analysis, the null hypothesis represents the default assumption, which is often the SM prediction or background only scenario, while the alternative hypothesis represents a deviation from the null hypothesis which is signal+background hypothesis. The goal of hypothesis testing is to determine whether the observed data is consistent with the null hypothesis or if there is sufficient evidence to reject it in favor of the alternative hypothesis.

9.1.2 Parameter of Interest

The parameter of interest (POI) is the specific quantity that is estimated or tested in the analysis. In this analysis, the parameter of interest is the signal strength (μ), which quantifies the deviation of the observed signal from the expected SM prediction. It is defined as:

$$\mu = \frac{\sigma_{\text{obs}}}{\sigma_{\text{SM}}} \quad (9.1)$$

where σ_{obs} is the observed cross-section and σ_{SM} is the predicted cross-section from the SM. A value of $\mu = 1$ indicates that the observed signal is consistent with the SM prediction, while $\mu > 1$ or $\mu < 1$ indicates an excess or deficit of events, respectively.

9.1.3 Nuisance Parameters

Nuisance parameters are parameters that are not of direct interest but must be accounted for in the analysis. They often represent systematic uncertainties, such as detector calibration, background modeling, or theoretical uncertainties (Chapter 8). Incorporating nuisance parameters into the analysis allows for a more accurate estimation of the parameter of interest and helps to ensure that the results are robust against systematic variations.

9.1.4 Test Statistic

Anything that can be calculated from the data and used to perform hypothesis testing is called a test statistic. A test statistic is a function of the observed data that summarizes the information relevant to the hypothesis being tested. The choice of test statistic depends on the specific analysis and the hypotheses being tested. In this analysis, the profile likelihood ratio is used as the test statistic.

9.1.5 Likelihood Function

The likelihood function is a fundamental concept in statistical inference, representing the probability of observing the data given a set of parameters. In the context of particle physics, the likelihood function is often constructed based on the expected signal and background distributions, as well as the observed data. The parameters of interest, such as signal strength or cross-section, can be estimated by maximizing the likelihood function. When the observable being measured is a continuous variable, the likelihood function is typically constructed using probability density functions for a given theoretical model. It is represented mathematically as:

$$\mathcal{L}(\theta) = f(x|\theta) \tag{9.2}$$

where $\mathcal{L}(\theta)$ is the likelihood function, $f(x|\theta)$ is the probability density function of observing the data x given the model parameters θ . The model parameters θ consist of both the parameter of interest (Section 9.1.2) and nuisance parameters (Section 9.1.3).

9.1.6 Confidence Intervals

A confidence interval is a range of values that is likely to contain the true value of a parameter of interest with a specified level of confidence. For a standardized normal distribution (mean of 0 and standard deviation of 1), a 68% confidence interval corresponds to the range $[-1, 1]$, while a 95.45% confidence interval corresponds to the range $[-2, 2]$, as shown in Figure 9.1.

Following the central limit theorem [129], as the sample size increases, the distribution of the sample mean approaches a normal distribution, regardless of the underlying distribution. Any given normal distribution can be easily transformed using the z -score

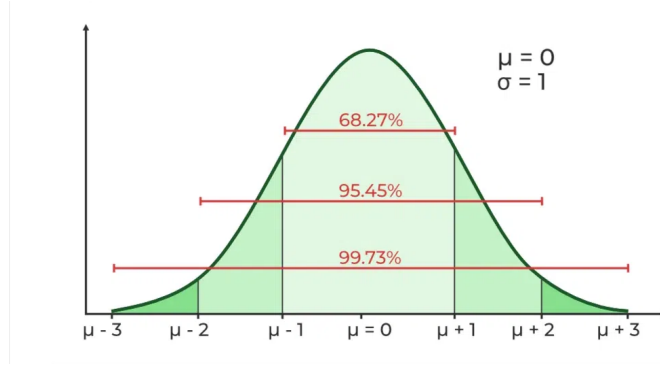


Figure 9.1: Graphical representation of confidence intervals for a standardized normal distribution with mean ($\mu = 0$) and standard deviation ($\sigma = 1$)[23].

formula discussed in Section 6.1, allowing for the application of confidence interval calculations based on the standardized normal distribution.

9.1.7 Significance and p-value

Significance is directly related to the confidence level. It is often expressed in terms of the number of standard deviations (σ) away from the mean of a normal distribution. The p-value measures the probability of the observed data assuming the null hypothesis is true. It is inversely related to the significance value; a smaller p-value indicates a higher significance level.

In particle physics experiments, a 5σ significance level is often used as a threshold for claiming a discovery, corresponding to a confidence level of approximately 99.99994% and a p-value of approximately 3×10^{-7} . This is commonly referred to as the “golden standard” [130].

9.2. Maximum Likelihood Estimation

The maximum likelihood estimator (MLE) is the set of parameter values that maximize the likelihood function. In this analysis, the MLE is used to determine the best-fit values of the signal strength (μ) and nuisance parameters (θ) by maximizing the likelihood function $\mathcal{L}(\mu, \theta)$ [131]. The likelihood function is defined as the product of the probability density functions for the observed data in the signal region (SR) and control regions (CRs), given the expected contributions from signal and background processes. The likelihood function can be expressed as:

$$\mathcal{L}(\mu, \theta) = \prod_{i=1}^{N_{\text{SR}}} \prod_{j=1}^{N_{\text{bins}}} P(n_{ij} | \mu s_{ij}(\theta) + b_{ij}(\theta)) \prod_{c=1}^{N_{\text{CR}}} P(n_c | \mu s_c(\theta) + b_c(\theta)) \prod_{k=1}^{N_{\text{NP}}} \pi_k(\theta_k) \quad (9.3)$$

where μ is the signal strength (POI), θ represents the set of nuisance parameters, N_{SR} is the number of signal regions, N_{bins} is the number of bins in each signal region, N_{CR} is the number of control regions, N_{NP} is the number of nuisance parameters, P is the Poisson probability function, and π_k is the prior probability density for the nuisance parameter θ_k . $s_{ij}(\theta)$, $b_{ij}(\theta)$, $s_c(\theta)$, and $b_c(\theta)$ represent the expected signal and background yields in the bins of signal regions and control regions, respectively, as functions of the nuisance parameters θ . Whereas n_{ij} and n_c are the observed number of data events in the signal and control regions respectively.

In Equation 9.3, the first product runs over all signal regions and their respective bins, while the second product runs over all control regions. The third product runs over all nuisance parameters, incorporating their constraints into the likelihood function. The prior probability density function π_k is typically modeled as a standardized Gaussian distribution, representing the constraint on the nuisance parameter θ_k while the Poisson probability function $P(n|\lambda)$ gives the probability of observing n events given an expected yield λ .

The expected yields for the signal and background processes depend on the nuisance parameters. The nuisance parameters affect both the normalization and shape of the signal and background distributions [132] as discussed in Chapter 8. When considering the effects on normalization, the expected yields can be expressed as:

$$n(\theta) = n_0(1 + \sigma)^\theta \quad (9.4)$$

where n_0 is the nominal expected yield, σ is the relative uncertainty associated with the nuisance parameter, and θ is the nuisance parameter itself, typically modeled as a standard normal variable (mean 0, standard deviation 1). For shape uncertainties, the expected yields can be expressed as:

$$n(\theta) = n_0(1 + \sigma\theta) \quad (9.5)$$

The MLE can then be expressed as:

$$\mathcal{L}(\hat{\mu}, \hat{\theta}) = \arg \max_{\mu, \theta} \mathcal{L}(\mu, \theta) \quad (9.6)$$

where $(\hat{\mu}, \hat{\theta})$ are the values of the signal strength and nuisance parameters that maximize the likelihood function.

The profile likelihood function is obtained by maximizing the likelihood function with respect to the nuisance parameters θ for each fixed value of the signal strength μ . This results in a function $\mathcal{L}_p(\mu)$ that depends only on the parameter of interest μ :

$$\mathcal{L}_p(\mu) = \max_{\theta} \mathcal{L}(\mu, \theta) \quad (9.7)$$

9.3. Test Statistics and p-values

To quantify the presence of signal in the data, a test statistic based on the profile likelihood ratio is used. The profile likelihood ratio is defined as:

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

where $\hat{\mu}$ and $\hat{\theta}$ are the unconditional MLEs of the signal strength and nuisance parameters that find the global maximum of the likelihood function, and $\hat{\hat{\theta}}$ is the conditional MLE of the nuisance parameters for a fixed value of the signal strength μ . The $\mathcal{L}(\mu, \hat{\hat{\theta}})$ is also referred to as the profile likelihood function as it "profiles" a fixed value of μ over the nuisance parameters θ . As the denominator is the global maximum, the value of the numerator never exceeds it, thus the value of $\lambda(\mu)$ never exceeds 1.

As the value of $\lambda(\mu)$ is always between 0 and 1, it is hard to interpret the results in terms of probability distributions. To address this, the test statistic q_μ is defined as:

$$q_\mu = -2 \ln \lambda(\mu) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (9.9)$$

The test statistic q_μ follows a chi-square distribution under the null hypothesis, allowing for the calculation of p-values and confidence intervals for the signal strength μ .

The p-value for a given value of the signal strength μ is calculated as:

$$p_\mu = \int_{q_{\mu, \text{obs}}}^{\infty} f(q_\mu | \mu) dq_\mu \quad (9.10)$$

where $q_{\mu, \text{obs}}$ is the observed value of the test statistic and $f(q_\mu | \mu)$ is the probability density function of the test statistic under the hypothesis of a signal strength μ . The significance of the observed signal is quantified using the p-value. A low p-value indicates a significant deviation from the null hypothesis, suggesting the presence of a signal. The significance is often expressed in terms of the number of standard deviations (σ) from the mean of a standard normal distribution, using the relation:

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

where Φ^{-1} is the inverse of the cumulative distribution function of the standard normal distribution. The expected significance is calculated by generating pseudo-experiments (asimov data) under the null hypothesis and computing the distribution of the test statistic.

The final results are presented in terms of the measured signal strength μ , along with the corresponding uncertainties, and significance (Section 10). These results provide

a comprehensive statistical interpretation of the analysis and its sensitivity to the Higgs boson signal in the ZH production with $H \rightarrow WW^*$ decay channel.

10 Expected Results

The results presented in this thesis are termed as “expected results” because they are derived using asimov data [133]. Asimov dataset is an artificial dataset construction by assuming null hypothesis holds true and setting the nuisance parameters to their nominal values. This dataset is used to estimate the sensitivity of the analysis to the signal process before looking at the actual data. Using asimov data helps to avoid potential biases that may arise from statistical fluctuations in the real data. It also helps to optimize the analysis strategy and selection criteria to maximize the expected sensitivity to the signal process.

10.1. ZH 4ℓ channel

Expected results for the ZH 1-SFOS and 2-SFOS channels are presented in this section. The NN output score is used as the final discriminant in both the signal regions and is used to perform a statistical fit to extract the signal strength (μ_{ZH}). After applying the statistical fits, post-fit distributions of the NN score distributions are plotted by optimizing the binning that maximizes the sensitivity of the analysis. Figures 10.1(a) and 10.1(b) show the post-fit distributions of the NN scores in the 1-SFOS and 2-SFOS SRs respectively. A profile likelihood scan of the signal strength is shown in Figure 10.2.

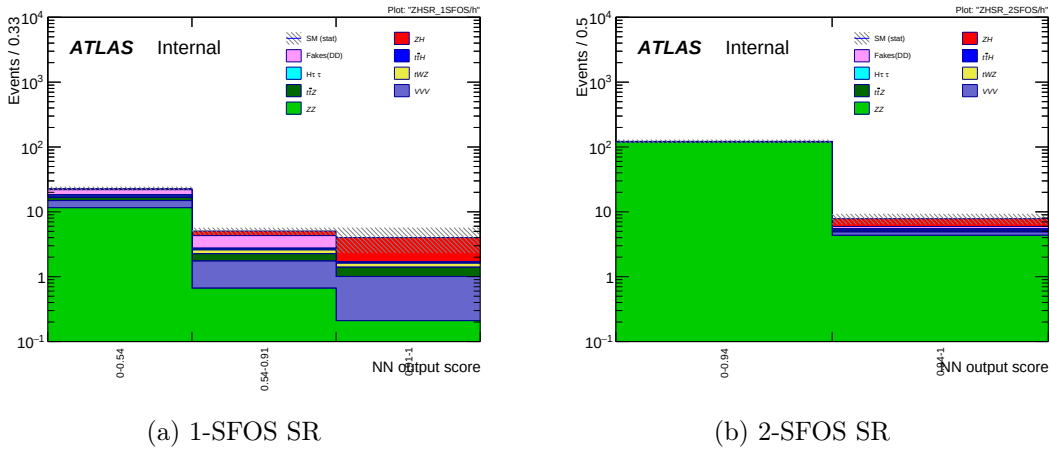


Figure 10.1: Post-fit distributions of the NN scores in the ZH 1-SFOS SR and 2-SFOS SR.

Table 10.1 shows the expected signal strength and the expected significance for the single ZH fit with 4ℓ 1-SFOS and 2-SFOS SRs combined together.

1-SFOS SR, 2-SFOS SR, ZZ CR	
μ_{ZH}	$1.00^{+0.73}_{-0.73}(tot)$
ATLAS norm ZZ	$1.00^{+0.10}_{-0.10}(tot)$
Expected Significance (σ)	1.71

Table 10.1: Expected results for the ZH 4ℓ channel.

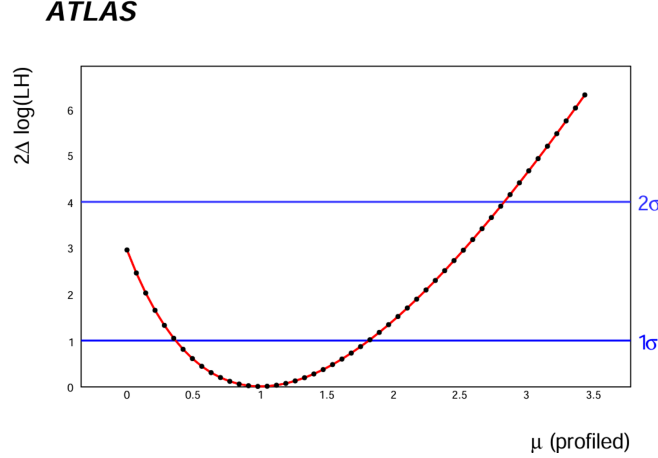


Figure 10.2: Profile likelihood scan of the signal strength for the ZH 4ℓ channel.

The uncertainty breakdowns and parameter pulls for each nuisance parameter (NP) in the full asimov fit is shown in Figure 10.3. The yellow band corresponds to uncertainty contribution to the POI, whereas the blue line with bars correspond to the pulls and constraints on each NP. Pulls are deviations of fitted NPs from their expected values, normalized by their uncertainties. The pulls are calculated for individual NPs. Any missing values for pulls correspond to the grouped NPs. In general, the NPs follow Gaussian constraints and hence their nominal value is expected to be around 0 for a standard Gaussian distribution. NPs corresponding to event counting, like the gamma NPs or normalization NPs, are constrained by Gamma distribution. Their nominal value is set to 1 in the RooFit package used for statistical fits in this analysis [134].

The uncertainty breakdown and parameter pulls help in understanding the impact that the statistical and systematic uncertainties have on the expected results. Two types of fits are performed for each NP: a fully unconditional fit where all the parameters are floating and a partial conditional fit where the NPs are fixed to their respective nominal values. The correlation matrix of the uncertainties is shown in Figure 10.4. The statistical uncertainties are still the dominant source of uncertainty in this analysis. So, there is still room for improvement as more data is added in the future.

10.2. Combined VH in Run 3

The expected results for the combined VH analysis with $H \rightarrow WW^* \rightarrow l\nu l\nu$ decay channel using the Early Run 3 data (2022-2023) are presented in this section. This combines the results from the WH analysis with three leptons in the final state [135] and the ZH analysis with four leptons in the final state described in this thesis.

The expected signal strength and the expected significance for the combined VH analysis are shown in Table 10.2. The uncertainties from both channel are correlated while performing the combined fit. The profile likelihood scan of the signal strength for the combined VH analysis is shown in Figure 10.5. The correlation matrix of the uncertainties for the combined VH analysis is shown in Figure 10.6. The statistical uncertainties are still the dominant source of uncertainty in combined VH analysis as well.

Combined VH (Early Run 3)	
μ_{VH}	$1.00^{+0.49}_{-0.49}(tot)$
Expected Significance (σ)	2.42

Table 10.2: Expected results for the combined VH analysis using Early Run 3 data (2022-2023).

10.3. Combined VH in Run 2 and Run 3

This section presents the expected results for the combined VH analysis with $H \rightarrow WW^*$ decay channel with both Run 2 data (2015-2018) and Early Run 3 data (2022-2023). The Run 2 VH analysis with $H \rightarrow WW^*$ decay channel contains 2 lepton, 3 lepton and 4 lepton final states [36]. The correlation scheme for the uncertainties is adapted from the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis [136].

The expected signal strength and the expected significance for the combined VH analysis with Run 2 and Early Run 3 data are shown in Table 10.3. The profile likelihood scan of the signal strength for the same is shown in Figure 10.7. Figure 10.8 shows the correlation matrix of the uncertainties involved in this combined fit.

Combined VH (Run 2 + Early Run 3)	
μ_{VH}	$1.00^{+0.21}_{-0.21}(tot)$
Expected Significance (σ)	5.34

Table 10.3: Expected results for the combined VH analysis using Run 2 (2015-2018) and Early Run 3 (2022-2023) data.

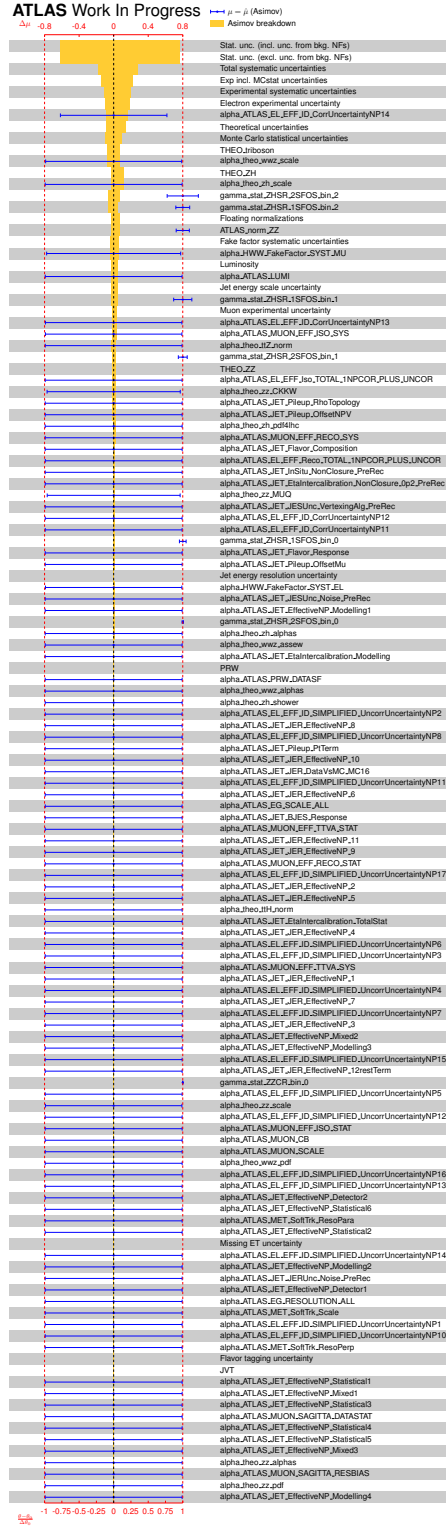


Figure 10.3: Pulls and uncertainty breakdowns for all NPs corresponding to the ZH 4ℓ channel.

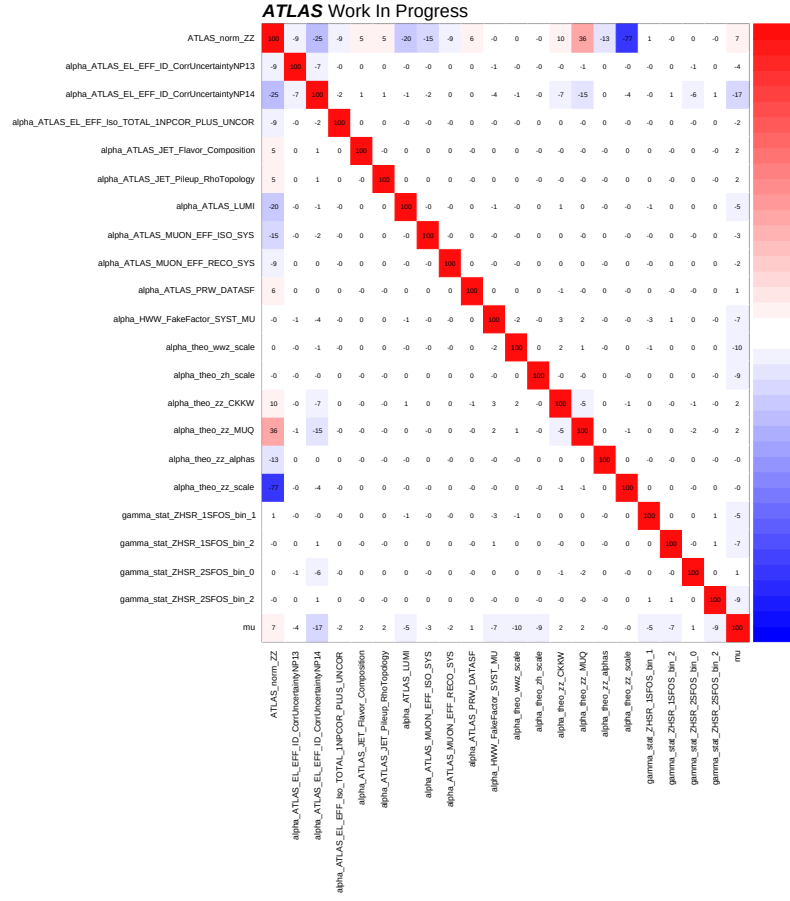


Figure 10.4: Correlation matrix of the uncertainties for the ZH 4ℓ channel. The correlation plot has a 5% cutoff to show the most relevant correlations.

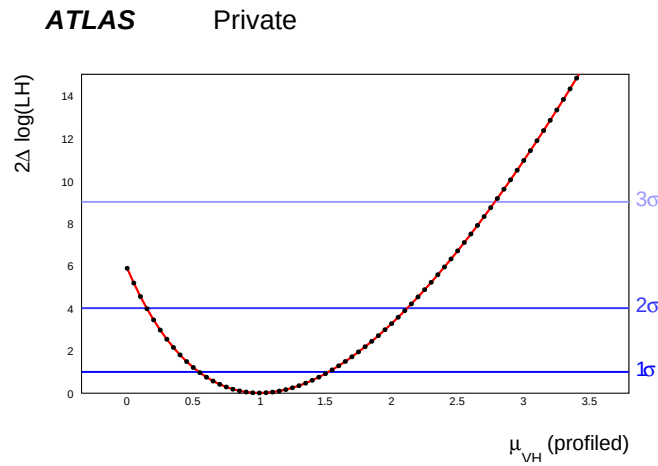


Figure 10.5: Profile likelihood scan of the signal strength for the combined VH analysis using Early Run 3 data (2022-2023).

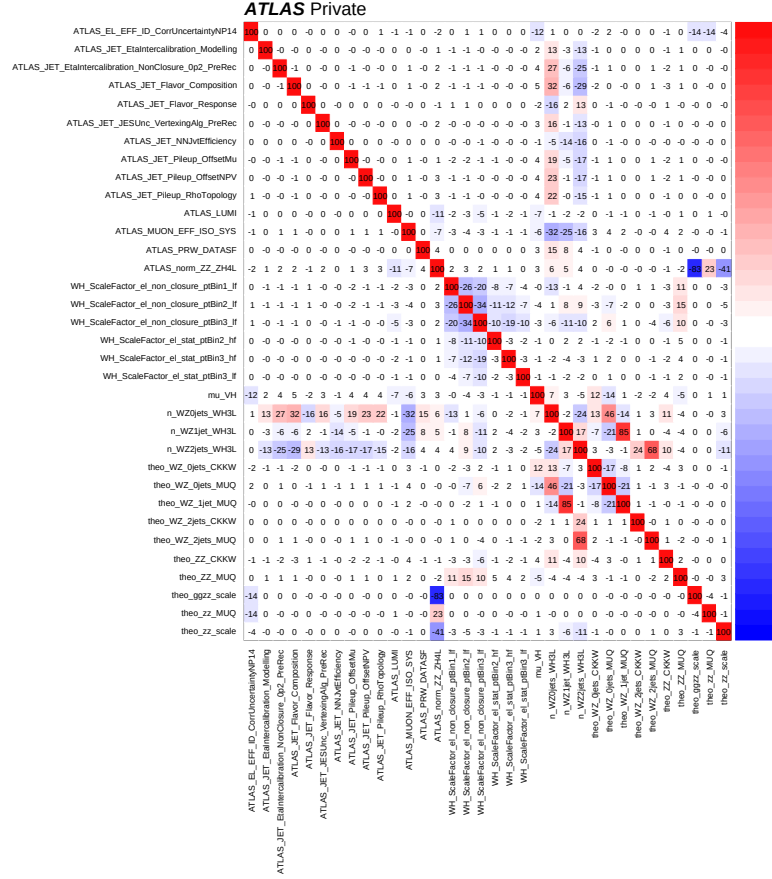


Figure 10.6: Correlation matrix of the uncertainties for the combined VH analysis using Early Run 3 data (2022-2023). The correlation plot has a 10% cutoff to show the most relevant correlations.

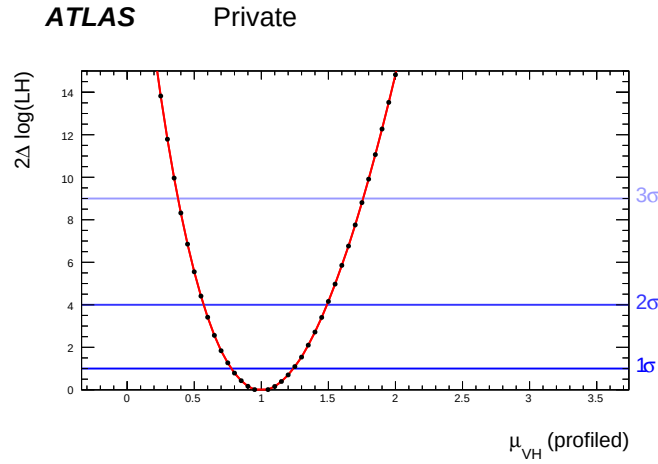


Figure 10.7: Profile likelihood scan of the signal strength for the combined VH analysis using Run 2 (2015-2018) and Early Run 3 (2022-2023) data.

11 Conclusion

This thesis describes a study focused on the search for the Standard Model Higgs boson in events where it is produced alongside a Z boson. The Higgs boson is analyzed in its decay channel to two leptons and two neutrinos via two intermediate W bosons, and the Z boson also decays leptonically, resulting in a final state with four leptons and missing transverse energy. The analysis utilizes proton-proton collision data at a center-of-mass energy of 13.6 TeV, corresponding to an integrated luminosity of 56.1 fb^{-1} , collected during the early Run 3 period in 2022 and 2023.

The study employs advanced techniques, including neural networks, to enhance the discrimination between signal and background events. The analysis is conducted in two signal regions, categorized based on the number of same-flavor opposite-sign lepton pairs (1-SFOS and 2-SFOS). A control region enriched in ZZ events is also defined to better constrain the dominant background.

The expected results indicate a signal strength of $\mu_{ZH} = 1.00^{+0.73}_{-0.73}$ and an expected significance of 1.71σ for the combined 1-SFOS and 2-SFOS signal regions. This combined with the WH analysis with three leptons in the final state from using the Early Run 3 data (2022-2023) results in an expected signal strength of $\mu_{VH} = 1.00^{+0.49}_{-0.49}$ and an expected significance of 2.42σ . All of this combined with the Run 2 VH analysis with $H \rightarrow WW^*$ decay channel results in an expected signal strength of $\mu_{VH} = 1.00^{+0.21}_{-0.21}$ and an expected significance of 5.34σ . These findings contribute to the ongoing efforts to understand the properties of the Higgs boson and its interactions with other particles. An internal note has been prepared for the ATLAS collaboration documenting the analysis and the results [38] and a paper for peer-reviewed publication is being prepared.

Future work will involve adding more data from Run 3 and the preliminary results after including the 2024 data are already showing significant improvement in the expected signal strength and significance. The analysis will also include the Simplified Template Cross-Section (STXS) measurements for the VH production with $H \rightarrow WW^*$ decay channel. This involves finding the cross-section measurements as a function of the gauge boson transverse momentum (p_T^V) which will help to better understand the Higgs boson production mechanism and its couplings to other particles.

Appendices

A Bibliography

- [1] Wikipedia. Standard Model. https://en.wikipedia.org/wiki/Standard_Model. Accessed: 2025-10-26. (Cited on page(s): xi, 5).
- [2] ATLAS Experiment at CERN. Handedness. Accessed: 2025-10-26. (Cited on page(s): xi, 10).
- [3] Suresh Emre. There are no right-handed neutrinos, 2014. (Cited on page(s): xi, 10).
- [4] CERN Courier. The thrill of the chase, 2022. (Cited on page(s): xi, 13).
- [5] Nguyễn Anh Kỳ and Nguyễn Thị Hồng Vân. Latest results on the Higgs boson discovery and investigation at the ATLAS-LHC. *J. Phys.: Conf. Ser.*, 627(1):012011, 2015. (Cited on page(s): xi, 14).
- [6] C. Grojean. Higgs Physics, 2016. (Cited on page(s): xi, 15).
- [7] Ewa Lopienska. The CERN accelerator complex, layout in 2022, 2022. General Photo. (Cited on page(s): xi, 18).
- [8] ATLAS collaboration. Public ATLAS Online Luminosity Plots for Run-3 of the LHC. (Cited on page(s): xi, 19).
- [9] ATLAS collaboration. Data Quality Efficiency Plots. (Cited on page(s): xi, 20).
- [10] Riccardo Maria Bianchi and ATLAS Collaboration. ATLAS experiment schematic or layout illustration, 2022. General Photo. (Cited on page(s): xi, 21).
- [11] Ana Maria Rodriguez Vera and Joao Antunes Pequeno. ATLAS Detector Magnet System, 2021. General Photo. (Cited on page(s): xi, 24).
- [12] ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3. *JINST*, 19(05):P05063, 2024. (Cited on page(s): xi, xii, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36).
- [13] ATLAS Collaboration. General components of the inner detector, 2024. General Photo. (Cited on page(s): xi, 26).
- [14] Joao Pequeno and Paul Schaffner. How ATLAS detects particles: diagram of particle paths in the detector, 2013. (Cited on page(s): xii, 38).
- [15] Maria Mironova. *Experimental Particle Physics with the ATLAS Detector*, pages 23–53. Springer Nature Switzerland, Cham, 2023. (Cited on page(s): xii, 39).

- [16] ATLAS Collaboration. Measurement of Electron Identification efficiency using $\sqrt{s}=13.6$ TeV pp collision recorded in 2023 at the ATLAS detector, 2025. (Cited on page(s): [xii](#), [40](#)).
- [17] ATLAS Collaboration. ATLAS Muon New Small Wheel performance plots, 2023. (Cited on page(s): [xii](#), [41](#)).
- [18] ATLAS Collaboration. Muon reconstruction and identification performance of the ATLAS detector using 2022 and 2023 LHC pp collision data at $\sqrt{s} = 13.6$ TeV, 2024. (Cited on page(s): [xii](#), [42](#)).
- [19] Morad Aaboud et al. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77(7):466, 2017. (Cited on page(s): [xii](#), [42](#), [43](#)).
- [20] Georges Aad et al. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021. (Cited on page(s): [xii](#), [43](#)).
- [21] Stefan Höche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pages 235–295, 2015. (Cited on page(s): [xii](#), [49](#)).
- [22] Fu et al. Optical neural networks: progress and challenges. *Light: Science and Applications*, 13(263), 2024. (Cited on page(s): [xiii](#), [65](#)).
- [23] geeksforgeeks. Python - normal distribution in statistics. (Accessed: 2025-10-15). (Cited on page(s): [xiv](#), [109](#)).
- [24] David J Griffiths. *Introduction to elementary particles*. Physics textbook. Wiley, New York, NY, 2008. (Cited on page(s): [xvi](#), [4](#), [6](#), [8](#), [11](#)).
- [25] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. *CERN Yellow Rep. Monogr.*, 2:1–869, 2017. (Cited on page(s): [xvi](#), [15](#)).
- [26] J. J. Thomson. Xl. cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897. (Cited on page(s): [1](#)).
- [27] E. Rutherford. Lxxix. the scattering of α and β particles by matter and the structure of the atom. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(125):669–688, 1911. (Cited on page(s): [1](#)).
- [28] Sarah Charley. Six fabulous facts about the Standard Model. *Symmetry Magazine*, 2021. (Cited on page(s): [1](#), [4](#)).
- [29] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008. (Cited on page(s): [1](#), [16](#)).

- [30] G. Arduini et al. LHC Upgrades in preparation of Run 3. *JINST*, 19(05):P05061, 2024. (Cited on page(s): [1](#), [16](#)).
- [31] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012. (Cited on page(s): [1](#), [5](#), [13](#)).
- [32] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012. (Cited on page(s): [1](#), [5](#), [13](#)).
- [33] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. (Cited on page(s): [1](#), [12](#), [13](#)).
- [34] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966. (Cited on page(s): [1](#), [12](#), [13](#)).
- [35] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. (Cited on page(s): [1](#), [12](#), [13](#)).
- [36] Georges Aad et al. Measurements of the production cross-sections of a Higgs boson in association with a vector boson and decaying into WW^* with the ATLAS detector at $\sqrt{s} = 13$ TeV. *JHEP*, 08:034, 2025. (Cited on page(s): [2](#), [115](#)).
- [37] ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. Also published by CERN Geneva in 2010. (Cited on page(s): [2](#), [21](#)).
- [38] Michiel Jan Veen, Callum McCracken, Matthew Leary Harris, Kyle Amirie, Mayuri Prabhakar Kawale, William Axel Leight, Benedict Tobias Winter, Bernd Stelzer, Carsten Burgard, Christian Schmitt, Matthew Joseph Basso, Mike Strauss, Oliver Stelzer-Chilton, Robin Hayes, Stephane Willocq, Yulei Zhang, and Peter Krieger. Measurement of the Higgs boson production cross section in association with a vector boson and decaying into WW^* with the ATLAS detector st $\sqrt{s} = 13.6$ TeV. Technical report, CERN, Geneva, 2025. (Cited on page(s): [2](#), [120](#)).
- [39] Ahren Sadoff and Lora Hine. The Standard Model. (Cited on page(s): [4](#)).
- [40] Matthew R. Francis. The mystery of particle generations, 2015. (Cited on page(s): [4](#)).
- [41] Frank Wilczek. Qcd made simple, August 2000. (Cited on page(s): [4](#)).
- [42] Ahmed Abokhalil. The Higgs Mechanism and Higgs Boson: Unveiling the Symmetry of the Universe, 6 2023. (Cited on page(s): [5](#)).
- [43] Katrina Miller. Beyond the Standard Model, 2022. (Cited on page(s): [5](#)).

- [44] H. Weyl. Gravitation and electricity. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1918:465, 1918. (Cited on page(s): 6).
- [45] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964. (Cited on page(s): 8).
- [46] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. *CERN Report*, 1 1964. (Cited on page(s): 8).
- [47] M. A. B. Bég, B. W. Lee, and A. Pais. Su(6) and electromagnetic interactions. *Phys. Rev. Lett.*, 13:514–517, Oct 1964. (Cited on page(s): 8).
- [48] O. W. Greenberg. Spin and unitary-spin independence in a paraquark model of baryons and mesons. *Phys. Rev. Lett.*, 13:598–602, Nov 1964. (Cited on page(s): 8).
- [49] O. W. Greenberg and Daniel Zwanziger. Saturation in triplet models of hadrons. *Phys. Rev.*, 150:1177–1180, Oct 1966. (Cited on page(s): 8).
- [50] O. W. Greenberg. The origin of quark color, 2015. (Cited on page(s): 8).
- [51] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954. (Cited on page(s): 8).
- [52] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961. (Cited on page(s): 9).
- [53] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. (Cited on page(s): 9).
- [54] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968. (Cited on page(s): 9).
- [55] Andrzej K. Wroblewski. The downfall of parity: The revolution that happened fifty years ago. *Acta Phys. Polon. B*, 39:251–264, 2008. (Cited on page(s): 10).
- [56] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. (Cited on page(s): 10).
- [57] CERN. Our history. <https://home.cern/about/who-we-are/our-history>. Accessed: 2025-09-30. (Cited on page(s): 16).
- [58] Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. (Cited on page(s): 17).
- [59] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. (Cited on page(s): 17).
- [60] LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. (Cited on page(s): 17).

- [61] ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. (Cited on page(s): [17](#)).
- [62] Vretenar, et al. *Linac4 design report*, volume 6 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020. (Cited on page(s): [17](#)).
- [63] van der Stok, P D V. Multiturn injection into the CERN proton synchrotron booster. Technical report, CERN, Geneva, 1981. (Cited on page(s): [17](#)).
- [64] Burnet, et al. *Fifty years of the CERN Proton Synchrotron*. CERN Yellow Reports: Monographs. CERN, Geneva, 2011. (Cited on page(s): [17](#)).
- [65] Doble, et al. The Super Proton Synchrotron (SPS): A Tale of Two Lives. *Adv. Ser. Direct. High Energy Phys.*, 27:135–177, 2017. (Cited on page(s): [17](#)).
- [66] Z Marshall. Simulation of Pile-up in the ATLAS Experiment. Technical report, CERN, Geneva, 2014. (Cited on page(s): [19](#)).
- [67] Herr, Werner and Muratori, B. Concept of luminosity, 2006. (Cited on page(s): [20](#)).
- [68] ATLAS collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ tev using the atlas detector at the lhc. *The European Physical Journal C*, 83(10), 2023. (Cited on page(s): [21](#), [48](#)).
- [69] ATLAS collaboration. *ATLAS inner detector*. Technical design report. ATLAS. CERN, Geneva, 1997. (Cited on page(s): [23](#)).
- [70] Norbert Wermes and G Hallewel. ATLAS pixel detector. Technical report, CERN, Geneva, 1998. (Cited on page(s): [25](#)).
- [71] B. Abbott et al. Production and Integration of the ATLAS Insertable B-Layer. *JINST*, 13(05):T05008, 2018. (Cited on page(s): [25](#)).
- [72] ATLAS Collaboration. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014. (Cited on page(s): [26](#)).
- [73] M Capéans-Garrido. The transition radiation tracker of the ATLAS experiment. *IEEE Trans. Nucl. Sci.*, 51:994–1000, 2004. (Cited on page(s): [27](#)).
- [74] ATLAS Collaboration. The ATLAS Trigger System for LHC Run 3 and Trigger performance in 2022. *JINST*, 19:P06029, 2024. (Cited on page(s): [27](#)).
- [75] ATLAS collaboration. Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions. *Eur. Phys. J. C*, 70:723–753, 2010. Comments: 31 pages, 34 figures, accepted in EPJC. (Cited on page(s): [30](#)).
- [76] C. Adloff et al. Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques. *JINST*, 7:P09017, 2012. (Cited on page(s): [30](#)).

- [77] Georges Aad et al. The ATLAS trigger system for LHC Run 3 and trigger performance in 2022. *JINST*, 19(06):P06029, 2024. (Cited on page(s): 34).
- [78] ATLAS collaboration. Software and computing for Run 3 of the ATLAS experiment at the LHC. *Eur. Phys. J. C*, 85:234, 2025. (Cited on page(s): 37, 40).
- [79] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987. (Cited on page(s): 37).
- [80] M. Aaboud et al. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2. *Eur. Phys. J. C*, 77(10):673, 2017. (Cited on page(s): 37).
- [81] Xiacong Ai et al. A Common Tracking Software Project. *Comput. Softw. Big Sci.*, 6(1):8, 2022. (Cited on page(s): 38).
- [82] ATLAS collaboration. Development of ATLAS Primary Vertex Reconstruction for LHC Run 3. Technical report, CERN, Geneva, 2019. (Cited on page(s): 38).
- [83] Georges Aad et al. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *JINST*, 14(12):P12006, 2019. (Cited on page(s): 39, 40).
- [84] ATLAS collaboration. Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-022>. (Cited on page(s): 39).
- [85] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms. Technical report, CERN, Geneva, 2008. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-LARG-PUB-2008-002>. (Cited on page(s): 39).
- [86] William Leight, Pierre-Francois Giraud, Peter Kluit, Jochen Meyer, Edward Moyse, and Alan Poppleton. New Fitting Concept in ATLAS muon tracking for the LHC Run-2. *EPJ Web Conf.*, 214:06006, 2019. (Cited on page(s): 40).
- [87] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 76(5):292, 2016. (Cited on page(s): 41).
- [88] Georges Aad et al. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 81(7):578, 2021. (Cited on page(s): 41, 42).

- [89] Simone Marzani, Gregory Soyez, and Michael Spannowsky. *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, volume 958. Springer, 2019. (Cited on page(s): 42, 43).
- [90] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008. (Cited on page(s): 42).
- [91] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012. (Cited on page(s): 43).
- [92] ATLAS collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical report, CERN, Geneva, 2014. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-018>. (Cited on page(s): 43).
- [93] Matthew Henry Klein, Francesco Rubbo, and Ariel Schwartzman. Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets. Technical report, CERN, Geneva, 2015. (Cited on page(s): 43).
- [94] Georges Aad et al. Transforming jet flavour tagging at ATLAS, 5 2025. (Cited on page(s): 44).
- [95] Georges Aad et al. The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp collisions. *Eur. Phys. J. C*, 85(6):606, 2025. (Cited on page(s): 44).
- [96] D Adams, C Anastopoulos, A Andreazza, M Aoki, L Asquith, M Begel, F Bernlochner, U Blumenschein, A Bocci, S Cheatham, W Davey, P-A Delsart, P-O DeViveiros, A Dewhurst, D Duschinger, F Filthaut, P Francavilla, F Garberson, S Head, A Henrichs, A Hoecker, M Kagan, B Kersevan, TJ Khoo, B Lenzi, D Lopez Mateos, B Malaescu, Z Marshall, T Martin, C Meyer, A Morley, W Murray, M zur Nedden, R Nicolaidou, S Pagan Griso, G Pasztor, P Petroff, C Pizio, R Polifka, X Poveda, R Reece, F Ruehr, F Salvatore, R Sandstroem, T Scanlon, D Scheirich, S Schramm, A Schwartzman, K Suruliz, M Sutton, E Thompson, M Tripiana, A Tuna, S Viel, M Vincet, I Vivarelli, M WIELERS, A Wildauer, and Z Zinonos. Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014. Technical report, CERN, Geneva, 2014. (Cited on page(s): 45).
- [97] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. (Cited on page(s): 47, 50).
- [98] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J. C*, 58:639–707, 2008. (Cited on page(s): 47).

- [99] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76(4):196, 2016. (Cited on page(s): [47](#)).
- [100] Enrico Bothmann et al. Event Generation with Sherpa 2.2. *SciPost Phys.*, 7(3):034, 2019. (Cited on page(s): [47](#), [51](#)).
- [101] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. (Cited on page(s): [47](#), [51](#)).
- [102] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. (Cited on page(s): [47](#)).
- [103] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. (Cited on page(s): [47](#)).
- [104] ATLAS Collaboration. Preliminary analysis of the luminosity calibration for the ATLAS 13.6 TeV data recorded in 2023. Technical report, CERN, Geneva, 2024. (Cited on page(s): [47](#), [103](#)).
- [105] S. Agostinelli et al. GEANT4 - A Simulation Toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003. (Cited on page(s): [49](#)).
- [106] ATLAS collaboration. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, 2014. (Cited on page(s): [50](#)).
- [107] Richard D. Ball et al. The PDF4LHC21 combination of global PDF fits for the LHC Run III. *J. Phys. G*, 49(8):080501, 2022. (Cited on page(s): [50](#), [105](#)).
- [108] Alexander Karlberg et al. Ad interim recommendations for the Higgs boson production cross sections at $\sqrt{s} = 13.6$ TeV. Technical report, CERN, 2 2024. (Cited on page(s): [50](#)).
- [109] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, and Juan Rojo. Parton distributions with QED corrections. *Nucl. Phys. B*, 877(2):290–320, 2013. (Cited on page(s): [51](#)).
- [110] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. (Cited on page(s): [51](#)).
- [111] Fabio Cardillo, Otilia Anamaria Ducu, Julien Maurer, Claudia Merlassino, Joseph Reichert, Erich Varnes, Pepijn Johannes Bakker, and Marcus Morgenstern. Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC. Technical report, CERN, Geneva, 1999. (Cited on page(s): [53](#)).
- [112] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014. (Cited on page(s): [64](#)).

- [113] Yann Coadou. Boosted decision trees, 3 2022. In: Machine Learning in High Energy Physics, World Scientific, pp. 23-60. (Cited on page(s): [64](#)).
- [114] Michael Nielsen. Neural Networks and Deep Learning, 2019. (Cited on page(s): [64](#)).
- [115] D. et al. Rumelhart. Learning representations by back-propagating errors. *Nature*, 533-536(323), 1986. (Cited on page(s): [66](#)).
- [116] Sebastian Ruder. An overview of gradient descent optimization algorithms, 2017. (Cited on page(s): [66](#)).
- [117] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. (Cited on page(s): [66](#)).
- [118] Timothy Dozat. Incorporating nesterov momentum into adam, 2016. (Cited on page(s): [66](#)).
- [119] François Chollet et al. Keras, 2015. (Cited on page(s): [67](#), [76](#)).
- [120] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. (Cited on page(s): [67](#), [76](#)).
- [121] StandardScaler function. Accessed: 2025-10-08. (Cited on page(s): [68](#)).
- [122] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization, 2018. (Cited on page(s): [68](#)).
- [123] Tom O’Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al. Keras Tuner, 2019. (Cited on page(s): [68](#)).
- [124] Toni et al. Baronecelli. Measurement of the Higgs boson production cross section in association with a vector boson and decaying into WW^* with the ATLAS detector at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2021. (Cited on page(s): [76](#)).
- [125] T. Plehn, D. Rainwater, and D. Zeppenfeld. Method for identifying $\vec{H} \tau \vec{\tau} e^\pm \mu^\mp p_T$ at the cern lhc. *Phys. Rev. D*, 61:093005, Apr 2000. (Cited on page(s): [77](#)).
- [126] Twiki. Pmgsystematicuncertaintyrecipes. (Version r29 of 7/6/2024). (Cited on page(s): [105](#)).

- [127] Twiki. Crosssectionnlottv. (Version r16 of 3/30/2020). (Cited on page(s): [105](#)).
- [128] Twiki. Ad interim 13.6tev cross sections. (Version r6 of 13/3/2024). (Cited on page(s): [105](#)).
- [129] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.*, 9(1):60–62, 1938. (Cited on page(s): [108](#)).
- [130] CERN. Why do physicists mention “five sigma” in their results? (Accessed: 2025-10-15). (Cited on page(s): [109](#)).
- [131] Kyle Cranmer. Practical Statistics for the LHC. *In: Proceedings, 2014 European School of High-Energy Physics, pp. 267-308*, pages 267–308, 2015. (Cited on page(s): [109](#)).
- [132] Georges Aad et al. A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery. *Nature*, 607(7917):52–59, 2022. [Erratum: *Nature* 612, E24 (2022)]. (Cited on page(s): [110](#)).
- [133] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. [Erratum: *Eur.Phys.J.C* 73, 2501 (2013)]. (Cited on page(s): [113](#)).
- [134] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical report, New York U., New York, 2012. (Cited on page(s): [114](#)).
- [135] Howard Baer and James D. Wells. Trilepton Higgs signal at hadron colliders. *Phys. Rev. D*, 57:4446–4452, 1998. (Cited on page(s): [115](#)).
- [136] Di Gregorio et al. Search for Higgs boson pair production in the $HH \rightarrow b\bar{b}\gamma\gamma$ final state with the full Run 2 + partial Run 3 pp collision data collected by the ATLAS Experiment. Technical report, CERN, Geneva, 2024. (Cited on page(s): [115](#)).