

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

Legacy Analysis of Standard Model Higgs boson
in the $H \rightarrow WW^* \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ decay channel from pp collisions
at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

Bijay Shrestha

Norman, Oklahoma

2024

Legacy Analysis of Standard Model Higgs boson
in the $H \rightarrow WW^* \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ decay channel from pp collisions
at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC

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

BY THE COMMITTEE CONSISTING OF

Dr. Michael G. Strauss, Chair

Dr. Phillip Gutierrez

Dr. Lance Lobban

Dr. Howard Baer

Dr. Mukremin Kilic

Acknowledgements

Looking back on the thirty years of education I've been privileged to receive; I realize it's impossible to name all the people who have supported and inspired me along the way. From my first steps guided by my parents' loving hands to the rigorous tutelage navigated alongside peers and mentors, each stage has positively impacted my journey. However, I would like to extend special appreciation to those who have played a key role in shaping my doctoral experience.

Firstly, I would like to express my gratitude to my supervisor, Dr. Michael Strauss. I am grateful for your unwavering support, guidance, and understanding throughout my doctoral journey. I would also like to express my appreciation for your attention to detail along with the constructive feedback in guiding me throughout the research. It has been an honor having you as my advisor, mentor, role model, and friend. I will forever cherish the memory of our hike in the treacherous winter conditions of the Jura mountains.

I am deeply grateful to the Department of Energy (DOE) for their generous funding and support, which made this research possible. I would also like to thank the *HWW* couplings analysis team, the *b*-jet trigger team, and the OU Physics and Astronomy Department. I am thankful for the collaborative environment, which included thorough and engaging physics discussions, as well as the constant backing I needed. I gained significant knowledge from each of you, and it was a pleasure to work with all of you.

I extend my heartfelt thanks to my undergraduate advisor, Dr. Hye-Young Kim, for trusting me and giving me the opportunity to conduct my first research. The valuable lessons you taught me throughout the years have helped me become a good student, researcher, and teaching assistant.

Last but not least, I would like to thank my parents, whose sacrifices throughout the years have been the bedrock of my academic journey. Thank you for never stopping to encourage me and for always believing in me. Equally, I am profoundly thankful to my fiancée, Maura, whose patience and understanding throughout my doctoral studies have been extraordinary. This thesis is as much a product of your support as it is of my efforts. Thank you.

Abstract

The 2012 discovery of a Standard Model-like Higgs boson at CERN’s Large Hadron Collider (LHC) marked a pivotal moment in high-energy physics. In the years that have followed, the Higgs boson’s properties and role within the Standard Model have been refined through increasingly precise measurements of this particle in a variety of final states and kinematic regimes.

This thesis contributes to this ongoing exploration by presenting a legacy analysis of the Standard Model Higgs boson production cross sections via gluon-gluon fusion, focusing on the previously unexplored 2-jet same-flavor leptonic final state, utilizing the $H \rightarrow WW^*$ decay mode. The analysis leverages data from proton-proton collisions at $\sqrt{s} = 13$ TeV, collected by the ATLAS detector during the LHC’s Run 2, comprising an integrated luminosity of 139 fb^{-1} . This work implements Deep Neural Network techniques to enhance signal identification and background suppression, rather than using m_T as the discriminant variable. The innovative DNN-based approach presented here aims to deepen the understanding of the Higgs sector, optimize the ggF production process analysis, and provide new insights into the interactions of the Higgs boson with the Standard Model. Employing DNN as the discriminant variable in the ggF channels significantly improved background suppression. The inclusion of the ggF 2-jet same-flavor leptonic state improved the cross-section sensitivity by 15% in the ggF inclusive study and by 25% in the ggF 2jet channel.

Keywords: Higgs boson | Gluon-gluon fusion | Deep Neural Networks

Contents

1	Introduction	1
1.1	General Overview	1
1.2	Structure of this manuscript	2
2	Theoretical Framework	3
2.1	Standard Model of Particle Physics	3
2.1.1	Lie Groups	4
2.1.2	Gauge Symmetries	5
2.1.3	Strong, Weak, and Electromagnetic Interactions	5
2.1.4	Spontaneously broken symmetries	7
2.1.5	The Brout-Englert-Higgs (BEH) mechanism	9
2.1.6	Electroweak Symmetry breaking	11
2.2	Measurement Predictions	13
2.2.1	Cross sections	13
2.2.2	Parton density functions	14
2.2.3	Underlying events	15
2.2.4	Parton shower and Hadronization	15
2.2.5	Pile-up	16
2.2.6	Luminosity	17
2.3	Monte Carlo Simulation	18
2.4	Statistical Data Analysis	20
2.4.1	Parameter of interest (PoI)	20
2.4.2	Nuisance parameter	21
2.4.3	Asimov data set	21
2.5	Higgs physics at the Large Hadron Collider (LHC)	22
2.5.1	Production modes of the Higgs boson	22
2.5.2	Decay channels	24
3	Experimental Setup: ATLAS	26
3.1	Large Hadron Collider	26
3.1.1	LHC Hardware Configuration	27
3.1.2	The path of protons	28
3.1.3	Experimental facilities located in CERN	29
3.2	The ATLAS Detector	30
3.2.1	The Inner Detector	31
3.2.2	Calorimeter	35
3.2.3	The Muon Spectrometer	37
3.2.4	Trigger	38
3.2.5	Event Reconstruction	39

3.2.6	Detector Simulation	45
3.3	The Worldwide LHC Computing Grid	46
3.4	Highlights of Higgs bosons measurements	46
4	$H \rightarrow WW^* \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ Physics	49
4.1	Introduction	49
4.2	Data and Monte Carlo Samples	51
4.2.1	Data Samples	51
4.2.2	Monte Carlo Samples	53
4.3	Object Identification and Selection	59
4.3.1	Electrons and Muons	59
4.3.2	Jets and Missing Transverse Momentum	60
4.3.3	Overlap Removal	60
4.4	Observables	61
4.4.1	Topological Observables	61
4.4.2	Background Observables	62
4.4.3	Other observables	63
4.5	Event Selection	64
4.6	STXS Categorization	67
4.7	Background Estimation	70
4.7.1	Top quark Background	71
4.7.2	Zll Background	71
4.7.3	STXS background estimation	72
4.7.4	Normalization factors	72
4.8	Misidentified Leptons	75
4.8.1	The Fake Factor Method	76
4.8.2	W + jets Control Region	77
4.8.3	Measurements of the fake factors in Z +jets event	78
4.8.4	Measurements of the fake factors in di-jet events	78
4.9	Data Analysis Techniques	80
4.9.1	Maximum Likelihood Method	80
4.9.2	Boosted Decision Tree	82
4.9.3	Deep Neural Networks	83
5	Systematic Uncertainties	92
5.1	Systematic Uncertainties	92
5.1.1	Experimental Uncertainties	92
5.1.2	Theoretical Uncertainties	95
5.2	Statistical Uncertainties	97
5.3	ggF Statistical Treatment	98
5.4	Zll mis-modeling of study	105
6	Results	108
6.1	STXS Measurements	108

6.2 Inclusive Measurements	108
7 Conclusions	117
A Bibliography	136

List of Figures

2.1	Standard Model of Particle Physics, including the respective properties such as charge, spin, and mass	3
2.2	The potential energy density ($\mathcal{V}$) as a function of ϕ_1 and ϕ_2 . At left, the potential with $\mu^2 > 0$; while at right, for the case with $\mu^2 < 0$	9
2.3	The potential $V(\phi)$ for $\mu^2 < 0$ and $\lambda^2 > 0$. The minimum defined in $\langle\phi\rangle = \frac{v}{\sqrt{2}}$ and the reparametrization defined as $\phi(x) = \frac{v+h(x)}{\sqrt{2}}e^{i\frac{\chi(x)}{v}}$	10
2.4	The parton distribution functions for the proton at $Q = 2$ GeV (left) and $Q = 100$ GeV (right) for quarks as well as gluons, where Q is the probe energy.	15
2.5	The two harmonization model. The Lund String (left) and Cluster (right) models.	16
2.6	The total integrated luminosity produced and recorded between 2015 to 2018. (b) The average number of interactions per bunch crossing during the data taking.[1]	18
2.7	Illustration of a pp collision as simulated by a Monte Carlo event generator. The large red circle indicates the primary hard-scatter process.[2] The emitted particles originating from partons in both the initial and final state are depicted in blue and red. The light green blobs represent the parton hadronization, while the dark green blobs represent the subsequent decay of the hadrons. Yellow lines represent the emission of low-energy photons. The blue and purple ovals represent the beam remnant and underlying event, respectively.	19
2.8	The main modes of the Higgs boson production at LHC.	22
2.9	The predicted SM Higgs boson production cross sections as a function of the center-of-mass energy. Line widths represent the respective theory uncertainties.[3]	23
2.10	The Feynman diagrams representing the main Higgs boson decays.	24
2.11	The predicted branching ratios for the SM Higgs decay modes near $m_H = 125$ GeV.[3]	25
3.1	Display of the LHC machine and its four main experiments.	27
3.2	The CERN accelerator complex, showing the LINAC2, the Booster, the PS and the SPS, the LHC, and other experimental setups [4]	28
3.3	Schematic display of the octagonal structure of the LHC with interaction points and facilities	29
3.4	A Cut-away view of the ATLAS detector to show the scale of the detector [5]	31
3.5	Schematic view of the ATLAS inner detector in Run I [6]	32
3.6	Schematic view of the insertable b-layer (IBL).	33
3.7	Schematic view of the structure of the inner detector barrel. [6]	34
3.8	Schematic view of Calorimeter system of ATLAS. [7]	35

3.9	Layout of an ECAL barrel module. The three distinct layers can be seen, each containing a different degree of granularity, which is specified. The $\delta\phi$ values are given in rad.	36
3.10	Cut-away view of the muon system.[8]	37
3.11	ATLAS trigger system and Data Acquisition (DAQ) system.[9]	39
3.12	Schematic view of the different signatures of particles traversing the detector.[10]	40
3.13	Diagram of a charged-particle track and the associated parameters describing its trajectory [11].	41
3.14	Illustration of the path of an electron [12].	42
3.15	MV2c10 BDT output for different jet flavors evaluated with $t\bar{t}$ events.	44
3.16	The invariant mass spectrum from the diphoton channel, as presented on July 4, 2012. The data points show the measured distribution, with each event weighted by the signal-to-background ratio for each sub-category of events. The solid red line represents the result of fitting the signal combined with the background, while the dashed red line represents the background.	47
3.17	The measured coupling strength as a function of the mass of different SM particles [13].	48
4.1	Experimentally measured decay branching fractions of W bosons [14].	50
4.2	Experimentally measured decay branching fractions of τ -leptons.	50
4.3	Spin correlations in the $H \rightarrow WW$ decay mode. The thin black arrows indicate the particles' directions of motion, while the thicker blue arrows indicate their spin projections.	51
4.4	The trigger efficiencies plotted against the p_T for different configurations in the 0 jet ggF (top left), one jet ggF (top right), and VBF (bottom) analyses. These efficiencies are calculated after preselection and the condition of having a leading lepton with a p_T greater than 22 GeV.	53
4.5	Event displays of candidate Higgs bosons from the (a) ggF $N_{\text{jet}} = 0$, (b) ggF $N_{\text{jet}} = 1$, (c) ggF-enriched $N_{\text{jet}} \geq 2$, and (d) VBF-enriched $N_{\text{jet}} \geq 2$ two analysis regions. Each display illustrates a cross-sectional and longitudinal view of the ATLAS detector. The final state particles are the electron (yellow line), the muons (turquoise line), the missing transverse momentum (white arrow), and the reconstructed jets (red and green cones). Yellow and red squares indicate the presence of energy deposits in the ECAL and HCAL and the reconstructed tracks by green lines.	56
4.6	The main background processes of the HWW analyses: (a) WW , (b) $t\bar{t}$, (c) Drell-Yan (DY) and (d) non-prompt background.	57
4.7	Overview of the categorization of events based on jet multiplicity.	65
4.8	Illustration of Central Jet Veto (Purple) and Outside Lepton veto (Teal).	66
4.9	Detailed illustration of Legacy run II HWW STXS measurement.	68
4.10	Detailed illustration of Legacy run II HWW STXS measurement.	69

4.11	Schematic setup with one signal contribution (red) and several background contributions (blue, green, and yellow) with one signal region (SR) and two control regions (CR1 and CR2). Factors that normalize the green and yellow background contributions to the two control regions are measured simultaneously, and then β_i is applied to all regions to change the background yields.	74
4.12	Feynman diagrams of signal (a) and W + jets (b) and QCD (c) backgrounds that contain fake leptons.	75
4.13	Results of the fake factor calculations, including statistical uncertainties. .	80
4.14	The parameter space spanned by the input variables (right) is divided by a decision tree (left).	83
4.15	The architecture of a Deep Neural Network.	84
4.16	Example of activation functions.	85
4.17	a) the BCE loss for 2-jet DNN, b) Accuracy for the 2-jet DNN, and c) loss vs significance plot	91
5.1	An analysis of the ranking distribution of nuisance parameters in the ggF fit with an Asimov dataset. The yellow and red bands account for their contribution to the total uncertainty in the analysis. This contribution is calculated as the quadratic difference between the uncertainty on μ in the main fit with all nuisance parameters and a fit where the nuisance parameter in question has been set to its best-fit value.	102
5.2	Post-fit Correlations between the fit parameters in the 2jet SF analysis Only nuisance parameters with correlations greater than 10% are shown for visibility.	103
5.3	Expected sensitivities for the parameters of interest for different configurations.	104
5.4	Distributions of p_t^H in Z_{ll} CR Z +jets in 2-jet SF channels.	105
5.5	Heat-maps showing the correlation between the DNN ggF output and the transverse momentum (p_t^H) for (a) ggF 0-200 GeV p_T^H (b) ggF 200-300 GeV p_T^H , and (c) ggF GT300 GeV p_T^H	106
5.6	1D plots of p_t^H for each DNN bin	106
5.7	1D plots of the DNN for each p_t^H bin	107
6.1	Relative Cross-section μ -fit uncertainty. All POIs are floating, and all channels are considered (DF and 2jSF).	109
6.2	Stacked histograms of the DNN ggF output for different regions based on the transverse momentum of the Higgs boson candidate (p_T^H) in the SF channel. (a) shows the region with p_T^H between 0 and 200 GeV, (b) with p_T^H between 200 and 300 GeV, and (c) with p_T^H greater than 300 GeV. The stacked contributions represent various background processes, and the bottom panel shows the ratio of data to the Standard Model (SM) prediction.	114

6.3	Stacked histograms of the DNN ggF output for control regions based on the transverse momentum of the Higgs boson candidate (p_T^H) in the SF channel. (left) shows the Zll control region with p_T^H between 0 and 200 GeV and (right) shows the top control region with p_T^H between 0 and 200 GeV. . . .	115
6.4	Inclusive distribution of m_T in the a) ggF 2j SF and b) VBF 2j SF regions. The bottom panel displays the ratio of the data yields to SM. Here, normalization factors are applied.	116
1	a) Total Time per Event (ms) vs. b -jet algorithms and b) Total Chain Time (s) vs. b -jet chains	123
2	CPU Time per Event of TrigFTF__bjet with 9659 events	124
3	Automation process followed by cost-monitoring script	125
4	Observing the change in the CPU usage by TrigFTF__bjet with changing parameters.	126
5	Di-lepton transverse momentum distributions after the preselection.	128
6	Transverse momentum of leading lepton distributions for 2 jets at different m_{ll} slices.	129
7	Transverse mass distribution without and with the $p_T^{t\ell} > 40\text{GeV}$ cut.	130
8	Distributions of $\Delta\phi_{\ell\ell}, m_{\ell\ell}, m_{jj}, m_T, p_T^{\text{tot}}, p_T^{H^{JJ}}, p_T^{\ell\ell}$ and $p_T^{\ell\ell}$ in the inclusive partial unblinded ggF signal region. The contributions from each STXS partial unblinded region received the corresponding background normalizations. . .	131
9	Distributions of $\Delta\phi_{\ell\ell}, m_{\ell\ell}, m_{jj}, m_T, p_T^{\text{tot}}, p_T^{H^{JJ}}, p_T^{\ell\ell}$ and $p_T^{\ell\ell}$ in the inclusive partial unblinded VBF signal region. The contributions from each STXS partial unblinded region received the corresponding background normalizations. . .	132
10	ggF DNN output distribution in the 3 ggF 2-jet SF Zll CRs and VRs with the binning used in the statistical analysis. The data is partially unblinded in the low DNN. NFs are applied.	133
11	ggF DNN output distribution in the 3 ggF 2-jet SF top VRs with the binning used in the statistical analysis. The data is partially unblinded in the low DNN. NFs are applied.	133

List of Tables

4.1	Minimum p_T requirements used at the each trigger. Letters “T”, “M” and “L” denote the Tight, Medium, and Loose electron identification requirements, respectively. The letter “i” indicates an isolation requirement to improve rejection power against non-prompt leptons.	52
4.2	Overview of Monte Carlo generators employed for simulating the signal and background processes in the analysis, alongside the associated cross sections used for normalization.	54
4.3	Summary of physics processes and causes of $\ell\ell + E_T^{miss}$ (+jets) signature and estimation methods [15].	58
4.4	Electron selections used in the analysis.	59
4.5	Muon selection used in the analysis.	60
4.6	Event selection criteria used to define the signal regions (SR).	66
4.7	Event selection criteria used to define the control regions (CR) and WW validation region (VR).	71
4.8	Summary of the criteria for fully identified (ID) and anti-identified (Anti-ID) electrons (left) and muons (right).	77
4.9	The input variables used for training the DNN models.	87
4.10	DNN bin edges for signal regions in the ggF and VBF 2j same flavor channel	89
4.11	Different DNN binning considered before choosing the above DNN binning scheme for ggF 2jSF	90
5.1	Summary of the experimental systematic uncertainty sources considered in the analysis	94
5.2	Summary of the theoretical systematic uncertainties considered in the analysis.	96
5.3	Asimov fit results for signal strengths and background normalization factors in the ggF 2j same flavor channel.	101
5.4	Expected sensitivities for the parameters of interest for different configurations (table form).	104
6.1	The NFs which scale the respective yields of the predicted background processes in the SRs.	109
6.2	Event yields of data and MC predictions for signal and background processes in each SR.	109
6.3	The expectation of the uncertainties on the signal strengths using the Asimov data set.	110
6.4	The observed values of the uncertainties on signal strengths using the full Run 2 data set.	110

6.5	Breakdown of the largest contributions to the total uncertainty in different channels; combined (ggF+VBF), ggF only, and VBF only, with the individual uncertainties being grouped together.	111
6.6	Signal and background MC event count per ggF STXS signal region and CR for 2jSF channel after applying background normalization factors. . . .	112
6.7	Signal and background MC event count per VBF STXS signal region and CR for 2jSF channel after applying background normalization factors. . . .	113

1 Introduction

1.1. General Overview

The pursuit to understand the fundamental constituents of nature and the underlying principles that govern their interactions has been a cornerstone of scientific inquiry throughout human history. The notion of these elemental building blocks of matter traces back to ancient Greek civilization, where it was referred to as 'atomos,' signifying its indivisible nature. The field of science has made significant progress since this ancient atomic philosophy, consistently advancing by building upon past accomplishments, formulating novel theories, and subjecting them to rigorous experimentation.

In contemporary physics, the Standard Model (SM) is a testament to our current understanding of elementary particles and their interactions. Formulated in the second half of the 20th century, the SM is known for its exceptional predictive capability and has been tested by many experiments. However, in its infancy, the SM had a significant flaw; the particles described by the SM were massless, which contradicted experimental evidence. Moreover, any effort to resolve this discrepancy in a makeshift manner would lead to the loss of certain symmetries that were the foundation of the theory.

In 1964, this issue was resolved by employing the concept of spontaneous symmetry breaking. This mathematical method explains how a system, initially characterized by perfect symmetry, ultimately selects one particular state, thereby breaking the symmetry. Robert Brout, François Englert, Peter Higgs, and three other theorists independently proposed a mechanism (BEH mechanism) that could explain why certain particles have mass. The masses of the particles in this mechanism are generated by their interaction with the Higgs field. Furthermore, this mechanism also predicted the existence of a scalar boson, namely the Higgs boson.

Numerous efforts were made to discover this elusive particle and examine the proposed mechanism. Particle accelerators are the primary instrument used to investigate this theory. These devices accelerate particles and cause them to collide either with a target or with each other. Due to its large mass and low production cross-section, the discovery of the Higgs boson required the construction of a powerful particle accelerator.

The Large Hadron Collider (LHC) located at the European Organization for Nuclear Research (CERN) in Geneva is currently the largest and most powerful accelerator where protons are accelerated to unprecedented energies and brought into collision within four separate detectors located around the LHC ring. On July 4th, 2012, the ATLAS [16] and CMS [17] detectors made a long-awaited announcement, revealing the discovery of the Higgs boson. This discovery was significant as it not only completed the missing component of the Standard Model of Particle Physics but also hinted at the possibility of

newer particles and unresolved issues that the SM does not address.

Aside from explicitly searching for new particles, precise measurements of the theory's predictions are vital. Precise measurements of the Higgs mechanism, such as coupling with fermions and bosons, can reveal Beyond Standard Model (BSM) effects. So, in 2015 and 2016, CERN carried out Run II with pp collisions at a center-of-mass energy of 13 TeV. This manuscript comprehensively examines pp collision data obtained during 2015 and 2016. The study presents a legacy analysis of one of the primary methods of Higgs production at the LHC, namely gluon-gluon fusion (ggF). Subsequently, the Higgs boson decays into a pair of W bosons, which further decay into two charged leptons, with a particular emphasis on the previously unexplored 2-jet same-flavor leptonic final state. The rationale for choosing this decay mode over others is due to its relatively high branching ratio [15, 18]. Furthermore, this subsequent leptonic decay mode results in final states that are clear and identifiable.

Here, a legacy analysis refers to research efforts that build on data collected in earlier runs of experiments. These analyses aim to improve sensitivity and maximize the scientific output of the dataset by revisiting previous analyses [19] with improved techniques and methodologies or theoretical models or addressing unanswered questions. Legacy analyses are important as they enable scientists to extract more precise results without the need for additional data collection.

1.2. Structure of this manuscript

The content of this thesis is structured in three distinct parts: In the first part of this thesis, Chapter 2 provides a detailed introduction to the theoretical framework with an emphasis on the Higgs mechanism. It also gives an overview of the different decay channels and production modes of the Higgs boson, as well as the primary findings from the search for the Higgs boson at the LHC.

In the second part, Chapter 3 introduces the experimental setup, i.e., the Large Hadron Collider (LHC) and the ATLAS Experiment. It also explores the data processing techniques used to collect and prepare the dataset.

In the last part of the thesis, Chapters 4 to 7 introduce the general analysis performed to quantify Higgs boson production in the $H \rightarrow WW^*$ that further decays leptonically. It also presents the study results using the pp collision data collected during 2015 and 2016. Chapter 7 concludes this thesis by summarizing the findings and providing a perspective on future research.

2 Theoretical Framework

This chapter introduces the Standard Model of Elementary Particle Physics (SM) and provides a theoretical framework on which the study is based. The equations in this chapter employ natural units, where physical action is measured in terms of the reduced Planck constant $\hbar$, and velocities are measured in terms of c .

2.1. Standard Model of Particle Physics

The Standard Model describes nature in terms of elementary particles and their interactions. The development of the Standard Model brought coherence to our understanding of the fundamental particles and their interactions in general [20, 21, 22, 23]. Figure 2.1 depicts a widely accepted portrayal of the particle composition within the Standard Model. The Standard Model describes and categorizes all known elementary particles into distinct types: Spin- $\frac{1}{2}$ fermions, spin-1 gauge bosons, and the spin-0 Higgs boson. The figure also shows the so-far unobserved graviton, which is not part of the SM [24, 25, 26, 27, 28, 29].

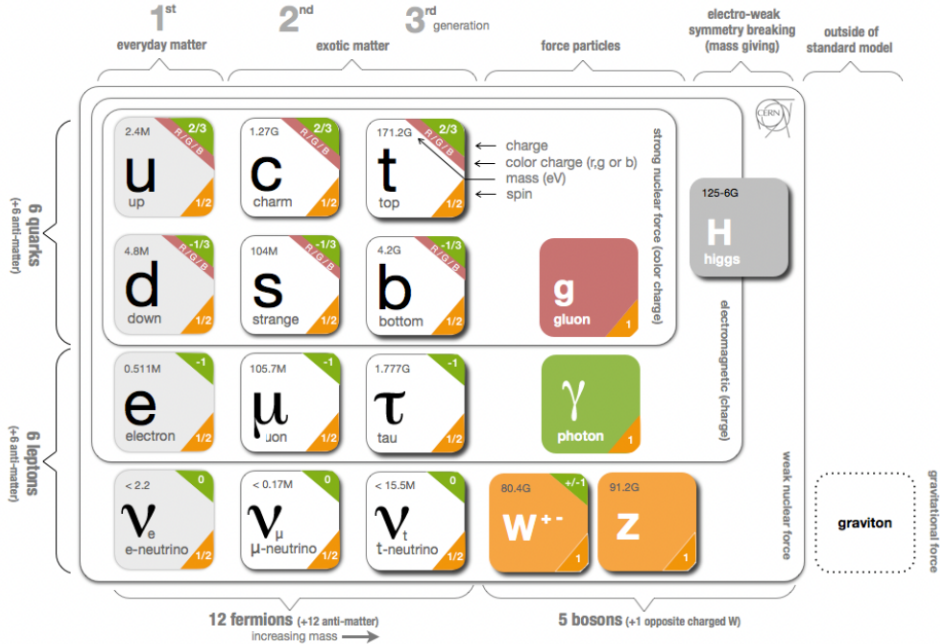


Figure 2.1: Standard Model of Particle Physics, including the respective properties such as charge, spin, and mass

Spin- $\frac{1}{2}$ fermions, the building blocks of matter, come in three generations, which differ primarily in mass. Fermions are further divided into quarks and leptons based on their propensity to interact with the strong forces. Leptons comprise electrons, muons, taus, and their corresponding neutrinos. Leptons interact through the weak force and, if charged, can interact via the electromagnetic (EM) force. On the other hand, quarks comprise up, down, charm, strange, top, and bottom. These quarks interact via the strong force as well as the weak and EM forces.

The Standard Model is a theoretical framework that combines relativistic gauge theories to describe electroweak and the strong interactions. This theory can provide a quantitative explanation of three of the four fundamental forces: the strong, weak, and electromagnetic forces. The fourth interaction is gravity, which is not currently accounted for by the SM. The spin 1-gauge bosons, also known as vector bosons, namely gluons, photons, and $W^\pm$, Z^0 bosons, are the mediators of the strong, electromagnetic, and weak interactions between particles, respectively. Spin-0 Higgs bosons give mass to the fermions and mediating bosons $W^\pm$, Z^0 via Yukawa coupling and electroweak symmetry breaking, respectively.

The SM describes the interactions of the particles using a mathematical formalism known as Quantum Field Theory (QFT). In QFT, the states or local excitations of quantized fields that permeate space and time represent the elementary particles. The dynamics of a system, i.e., the motion and interactions in these fields, are governed by a mathematical quantity known as the Lagrangian. The Lagrangian of the SM is invariant under continuous internal transformations of the group $SU(3) \otimes SU(2) \otimes U(1)$. This invariance is referred to as gauge invariance and is critical for ensuring that the theory is renormalizable. A brief introduction to these ideas will be discussed in the following sections.

2.1.1 Lie Groups

Group Theory is a branch of mathematics that studies a set of elements along with a particular operation that follows specific rules, known as a group. An example of a group is the set of integers (positive and negative whole numbers) under the addition operation. This set satisfies all the properties of a group, namely, closure (the operation of combining two elements always results in an element that is also in the group), the existence of an identity element, the existence of an inverse for each element, and associativity (the order of combination of elements does not matter). These groups provide a structured way to study the symmetries, transformations, and structures.

A Lie group is a smooth manifold that satisfies the group properties and has differentiable group operations. Consider the rotation group in three-dimensional space, often denoted as $SO(3)$. It consists of all possible rotations around any axis. This group is a Lie group because the rotations are continuous and smooth operations, and the set of rotations forms a smooth manifold. Quantum Field Theory utilizes several Lie groups, including the unitary groups $U(N)$, the special unitary groups $SU(N)$, the orthogonal groups $O(N)$,

and the special orthogonal groups $SO(N)$.

2.1.2 Gauge Symmetries

The particles' kinematics and dynamics are characterized by corresponding Lagrangian density ($\mathcal{L}$). The interactions between particles are determined by various terms in the Lagrangian, which are derived from the principle of gauge symmetry. In QFT, a particle is characterized by a field $\psi(x)$, where x represents the four space-time coordinates. The gauge symmetry can be classified as global when the Lagrangian associated with a particle remains unchanged under transformation.

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha^A T^A} \psi(x) \quad (2.1)$$

where T^A are the generators of a specific group of transformations, while A represents the indices across all dimensions of the transformation group, and α^A is a global phase transformation, meaning it remains constant throughout all space-time. Furthermore, through Noether's theorem [30], this symmetry implies the existence of a conserved current.

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha^A(x) T^A} \psi(x) \quad (2.2)$$

However, in the SM, the dynamics of all particles are dictated by local gauge symmetry. This symmetry refers to the property of the Lagrangian remaining unchanged (invariance) under local gauge group transformation. Hence, the global phase transformation (α^A) is now space and time-dependent, i.e., $\alpha^A(x)$, also known as coupling strength, making it no longer invariant over gauge transformations.

By enforcing this local gauge invariance, it becomes necessary to introduce new massless vector fields to restore the Lagrangian invariance. These supplementary local gauge fields have spin-1 and account for the particles that mediate the interaction. The excitations of the gauge fields correspond to particles referred to as “gauge bosons.” The Lagrangian of the standard model incorporates twelve gauge fields, consisting of eight for the $SU(3)$ generators, three for the $SU(2)$ generators, and one for the $U(1)$ generator.

2.1.3 Strong, Weak, and Electromagnetic Interactions

In theory, this group $SU(3) \otimes SU(2) \otimes U(1)$ provides a framework to understand particles and their interactions. When coupled with quarks, $SU(3)$ is the gauge group of Quantum Chromodynamics (QCD), accurately representing the strong interaction. The eight $SU(3)$ gauge fields correspond to the various colored states of the gluon. Similarly, W^+ , W^- , and Z^0 are the three gauge fields of $SU(2)$, representing the weak interactions, and γ is the gauge field of $U(1)$, representing the electromagnetic interactions.

The Lagrangian of quantum chromodynamics (QCD) for free quark color fields q is

given by

$$\mathcal{L}_{QCD} = \bar{q}_i (i\gamma^\mu \mathcal{D}_\mu - m) q_i - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.3)$$

where $\mathcal{D}_\mu = \partial_\mu + ig_s T^a G_\mu^a$ is the associated covariant derivative and T^a are the eight generators of the SU(3) group. The indices “i” represent the three color charges (red, green, and blue), and the gluon field strength tensor is

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s \sum_{b,c=1}^8 f_{abc} G_\mu^b G_\nu^c \quad (2.4)$$

The term $\bar{q}_i (i\gamma^\mu \mathcal{D}_\mu - m) q_i$ represents both the free propagation of quarks and interactions with gluons. $G_{\mu\nu}^a$ can be attributed to the non-vanishing commutations between the Gell-Mann matrices, resulting in the formation of self-coupling among the gluon fields. The excitations of this $G_{\mu\nu}^a$ fields correspond to the gluons, which are massless due to their origin from an unbroken local gauge symmetry.

The Lagrangian of quantum electrodynamics (QED) of a fermion with charge q is given by

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\not{\partial} - m)\psi - q\bar{\psi}\gamma^\mu\psi A_\mu \quad (2.5)$$

where $\not{\partial} \equiv \gamma^\mu \partial_\mu$, $\bar{\psi} \equiv \psi^\dagger \gamma^0$, γ^μ are Dirac matrices and μ are the indices of the four space-time coordinates (0 to 3). ψ is a four-component Dirac spinor which describes a free-charged fermion. In order to be invariant under the local gauge transformation, the Maxwell tensor (electromagnetic field strength tensor), $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is introduced. A_μ is a gauge field, and its excitations correspond to the photons, which are massless due to their origin from an unbroken U(1) gauge symmetry. Moreover, the term $q\bar{\psi}\gamma^\mu\psi A_\mu$ encodes the interaction between the two fields A_μ and ψ .

The Lagrangian for electroweak theory before the symmetry breaking is given by

$$\mathcal{L}_{EW} = \sum_\psi i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (2.6)$$

with the SU(2) gauge boson fields

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk} W_\mu^j W_\nu^k \quad (2.7)$$

and the U(1) gauge boson

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.8)$$

This Lagrangian is invariant under SU(2) $\times$ U(1) gauge transformations. However, in order to maintain gauge invariance, it is necessary to include these gauge bosons and

fermions to be massless, which is in drastic contrast with the observed massive $W^\pm$ and Z^0 bosons.

The masses of the fermions present an additional challenge because left and right-handed quark and lepton helicity states couple differently with the weak gauge fields. In the SM, the distinction between left and right-handed fermions is considered by treating them as separate fields with distinct couplings. However, including a fermion mass term in the Lagrangian would result in the coupling of these different fields, consequently leading to the violation of SU(2) gauge invariance. Hence, this violation of gauge invariance would result in the fermions being massless, which directly contradicts the experimentally determined masses of the fermions. This issue within the electroweak theory is known as the mass problem.

2.1.4 Spontaneously broken symmetries

Both aforementioned problems can be resolved by employing a concept known as “spontaneous symmetry breaking.” Typically, a physical system is considered to have a spontaneously broken symmetry when the Lagrangian governing its dynamics is invariant under a symmetry transformation, but its ground state does not. In QFT, the energy ground state is the vacuum of space.

The concept involves incorporating additional quantum fields with spin 0, also known as scalar fields, into the theory. The scalar fields are incorporated into the system using a unique type of interaction, which ensures that the lowest energy state does not correspond to zero values of the fields. Although the Lagrangian maintains the symmetry during interaction, the ground state of the field will inevitably break this symmetry. So, even though a particular state does not exhibit symmetry, the Lagrangian still preserves the gauge invariance. Thus, the symmetry is referred to as “spontaneously broken” in this context.

The Goldstone model [31] is an excellent example to illustrate how symmetry-violating mass terms that violate symmetry in a theory can be introduced. The Goldstone model employs a U(1) global invariant approach, and the Lagrangian is expressed as:

$$\mathcal{L}_{\text{Goldstone}} = (\partial^\mu \phi^*(x)) (\partial_\mu \phi(x)) - \mu^2 |\phi(x)|^2 - \lambda |\phi(x)|^4 \quad (2.9)$$

where μ^2 and λ are arbitrary real parameters, and the complex scalar $\phi(x)$ is given by

$$\phi(x) = \frac{1}{\sqrt{2}} (\phi_1(x) + i\phi_2(x)) \quad (2.10)$$

Figure 2.2 displays the potential of the field, denoted as $\mathcal{V}(\phi) = -\mu^2 |\phi(x)|^2 - \lambda |\phi(x)|^4$. To ensure a lower bound on the ground state, the value of λ must be positive. Furthermore, in the case, where μ^2 is positive, the function $\mathcal{V}$ has only one minimum at $\phi(0) = 0$. If μ^2

is negative, the system will exhibit a local maximum at $\phi(0) = 0$, and the potential $\mathcal{V}(\phi)$ is minimized at

$$\phi_0 = \sqrt{\frac{-\mu^2}{2\lambda}} e^{i\theta} \quad (2.11)$$

where θ represents the angle that determines the direction in the complex plane ϕ . In this instance, it is evident that the system does not possess $U(1)$ invariance. Since the Lagrangian is invariant under global transformations, choosing an arbitrary direction is unimportant. Therefore, when $\theta = 0$ is selected, it corresponds to the ground state.

$$\phi_0 = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}} \quad (2.12)$$

is purely real, where $v = -\mu^2/\lambda$ is the vacuum expectation value.

Two fields $\sigma(x)$ and $\eta(x)$ around the vacuum are then introduced so the new complex scalar is defined by

$$\phi(x) = \frac{1}{\sqrt{2}}(v(x) + \sigma(x) + i\eta(x)) \quad (2.13)$$

The potential $\mathcal{V}$ exhibits quadratic growth when the σ field is excited. This excitation corresponds to a displacement in the radial plane with $\phi_2 = 0$. On the other hand, the η field remains unchanged, representing a displacement along the valley of minima as depicted. Consequently, the associated quantum excitations are characterized as massless. The η boson is linked to the breaking of $U(1)$ symmetry and is commonly referred to as the Goldstone boson. The Goldstone boson's lack of mass can be confirmed by formulating the free Lagrangian of the model.

$$\mathcal{L}_{\text{Goldstone}}^0 = \frac{1}{2} (\partial^\mu \sigma(x)) (\partial_\mu \phi(x)) + \frac{1}{2} (2\lambda v^2) \sigma^2(x) + \frac{1}{2} (\partial^\mu \eta(x)) (\partial_\mu \eta(x)) \quad (2.14)$$

Consequently, two bosons with spin 0 have emerged: the σ boson with a mass of $m_\sigma = \sqrt{2\lambda v^2}$, and the η boson, which remains massless due to the absence of a $\eta^2(x)$ term.

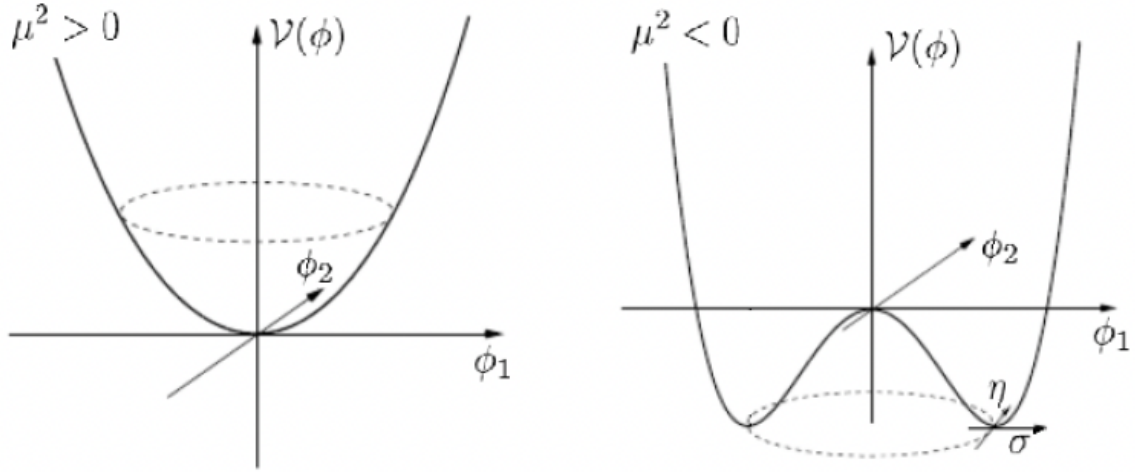


Figure 2.2: The potential energy density ($\mathcal{V}$) as a function of ϕ_1 and ϕ_2 . At left, the potential with $\mu^2 > 0$; while at right, for the case with $\mu^2 < 0$.

2.1.5 The Brout-Englert-Higgs (BEH) mechanism

In 1964, three separate groups independently applied the Goldstone theorem to local gauge theory. BEH mechanism [32, 33, 34, 35, 36, 37] is the process by which a vector boson acquires mass without destroying the gauge invariance of the Lagrangian. The Higgs boson is the massive spin-0 boson associated with the scalar field $\sigma(x)$, which couples with the electroweak $SU(2) \times U(1)$ gauge fields.

In order to show how massive bosons can appear, consider this toy example of spontaneous symmetry breaking.

The free Lagrangian of a “universe containing nothing but photons” is given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.15)$$

where $F_{\mu\nu}$ is the electromagnetic field strength tensor. It is invariant under the local gauge transformation

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{\partial}{\partial x_\mu}\eta(x) \quad (2.16)$$

where $\eta(x)$ is some arbitrary function.

Now consider the scenario where the photons in this toy universe possess a nonzero mass. As observed in the Lagrangian, the local gauge symmetry is disrupted when an explicit mass term is added without considering its implications.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu \quad (2.17)$$

This Lagrangian is no longer invariant under the transformation. However, to overcome

this, one can add a new scalar field ϕ that carries a charge $-e$, interacts with itself as well as with photons, and is subject to a global potential, $V(\phi) = -\mu^2|\phi|^2 + \lambda^2|\phi|^4$.

The Lagrangian with this field could be rewritten as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \left| \left(\frac{\partial}{\partial x_\mu} - ieA_\mu(x) \right) \phi \right|^2 - V(\phi) \quad (2.18)$$

This is invariant under the local gauge transformation

$$\begin{aligned} A_\mu(x) &\rightarrow A_\mu(x) - \frac{\partial}{\partial x_\mu} \eta(x) \\ \phi(x) &\rightarrow e^{ie\eta(x)} \phi(x). \end{aligned} \quad (2.19)$$

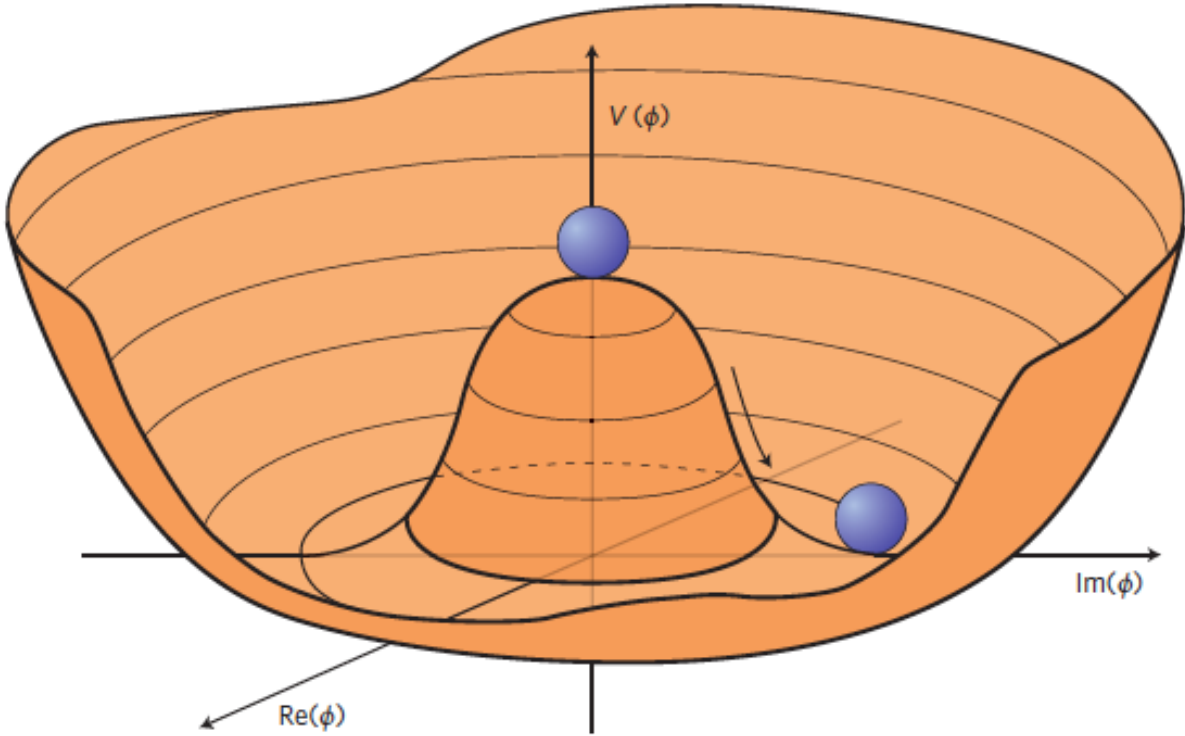


Figure 2.3: The potential $V(\phi)$ for $\mu^2 < 0$ and $\lambda^2 > 0$. The minimum defined in $\langle \phi \rangle = \frac{v}{\sqrt{2}}$ and the reparametrization defined as $\phi(x) = \frac{v+h(x)}{\sqrt{2}} e^{i\frac{\chi(x)}{v}}$.

By reparametrizing,

$$\phi(x) = \frac{v+h(x)}{\sqrt{2}} e^{i\frac{\chi(x)}{v}},$$

where, $v = \sqrt{\frac{\mu^2}{\lambda}}$ and the new Lagrangian is

$$\begin{aligned}\mathcal{L} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(ev)^2 A_\mu A^\mu - \lambda v^2 h^2 - \lambda v h^3 \\ & -\frac{1}{4}\lambda h^4 + \frac{1}{2}e^2 A_\mu A^\mu h^2 + ve^2 A_\mu A^\mu h + \frac{1}{2}\partial_\mu h \partial^\mu h.\end{aligned}\tag{2.20}$$

In this way, the Lagrangian incorporates two key elements: $m = ev$ representing the mass term, and a physical Higgs field $h(x)$ with a mass of $m = \sqrt{2\lambda}v$, and it still satisfies the local gauge invariance. This process by which a vector boson acquires mass without violating the gauge invariance of the Lagrangian is known as the Higgs mechanism.

2.1.6 Electroweak Symmetry breaking

The electroweak model utilizes the Higgs mechanism to develop a theory involving three large bosons, namely $W^\pm$ and Z , along with a photon with no mass. To achieve this objective, the gauge symmetry is disrupted through the inclusion of a self-interacting complex scalar field known as the Higgs field. The most basic option for the Higgs field is a complex SU(2) isospin doublet of scalar fields. This doublet has no color charge, implying that the SU(3) sector remains unaffected by this field. The Lagrangian which describes this Higgs field is given by

$$\mathcal{L}_H = D^\mu \phi^\dagger(x) D_\mu \phi(x) - \mu^2 \phi^\dagger(x) \phi(x) - \lambda \left(\phi^\dagger(x) \phi(x) \right)^2\tag{2.21}$$

where the covariant derivative is

$$D^\mu = \partial^\mu + \frac{1}{2}ig\tau_i W_i^\mu + \frac{1}{2}ig'Y B^\mu\tag{2.22}$$

Here, λ is positive in order to have vacuum stability, and μ^2 is negative to break the electroweak symmetry. The Higgs field is then chosen in a way that results in the breaking of the SU(2) $\times$ U(1) gauge symmetry and the loss of invariance under U(1)_EM transformations.

$$\phi = \frac{1}{2} \begin{pmatrix} 0 \\ v + \sigma(x) \end{pmatrix},$$

then the Higgs Lagrangian becomes

$$\mathcal{L}_H = \mathcal{L}_H^0 + \mathcal{L}_0^{\text{mass}} + \mathcal{L}_{HH}^{\text{Int}} + \mathcal{L}_{HW}^{\text{Int}} + \mathcal{L}_{HZ}^{\text{Int}}\tag{2.23}$$

where

$$\begin{aligned}
\mathcal{L}_H^0 &= \frac{1}{2} (\partial^\mu \sigma) (\partial_\mu \sigma) - \frac{1}{2} m_H^2 \sigma^2, \\
\mathcal{L}_V^0 &= m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu, \\
\mathcal{L}_{HH}^{\text{Int}} &= -\lambda v \sigma^3 - \frac{1}{4} \lambda \sigma^4, \\
\mathcal{L}_{HW}^{\text{Int}} &= \frac{1}{2} v g^2 \sigma W_\mu^+ W^{-\mu} + \frac{1}{4} g^2 \sigma^2 W_\mu^+ W^{-\mu}, \\
\mathcal{L}_{HZ}^{\text{Int}} &= \frac{v g^2}{4 \cos \theta_W} \sigma Z_\mu Z^\mu + \frac{g^2}{8 \cos \theta_W} \sigma^2 Z_\mu Z^\mu.
\end{aligned}$$

Here, the parameters were replaced by the particles' masses:

$$\begin{aligned}
v &= \frac{2m_W}{g} = \frac{2m_Z}{\sqrt{g^2 + g'^2}}, \\
\lambda &= -\frac{\mu^2}{v^2} = \frac{m_H^2}{2v^2} = \frac{g^2 m_H^2}{8m_W^2}.
\end{aligned}$$

The first term of the Lagrangian represents the mass of the Higgs boson ($m_H^2 = -2\mu^2$). The second term gives the mass of the vector bosons $W^\pm$ and Z , while the third term corresponds to the triple and quartic self-coupling. The last two terms describe the Higgs couplings with the vector bosons.

However, the Higgs mechanism alone does not produce the masses of fermions. In order to obtain the masses of the fermions, each fermion meant to acquire a mass is allowed to interact with the Higgs field via Yukawa coupling without disrupting the gauge symmetry.

$$\mathcal{L}_Y = - \left(\lambda_f \frac{\nu}{\sqrt{2}} \right) \bar{f}_L f_R \quad (2.24)$$

where f represents a particular quark or charged lepton and λ_f is its Yukawa coupling parameters. the fermion mass is given by,

$$m_f = \lambda_f \frac{v}{\sqrt{2}} \quad (2.25)$$

This proportionality indicates that the Higgs boson has a stronger coupling with heavier fermions than lighter fermions. Therefore, the final Lagrangian of the SM can be

written as

$$\begin{aligned}
\mathcal{L}_{SM} = & \underbrace{\frac{1}{4}W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G_{\mu\nu}^\alpha G_{\alpha}^{\mu\nu}}_{\text{kinetic energies and self-interactions of the gauge bosons}} \\
& + \underbrace{\bar{L}\gamma^\mu \left(i\partial_\mu - \frac{1}{2}g\tau \cdot W_\mu - \frac{1}{2}g'YB_\mu \right) L + \bar{R}\gamma^\mu \left(i\partial_\mu - \frac{1}{2}g'YB_\mu \right) R}_{\text{kinetic energies and electroweak interactions of fermions}} \\
& + \underbrace{\frac{1}{2} \left| i\partial_\mu - \frac{1}{2}g\tau \cdot W_\mu - \frac{1}{2}g'YB_\mu \right| \phi|^2 - V(\phi)}_{W^\pm, Z, \gamma \text{ and Higgs masses and couplings}} + \underbrace{(G_1 \bar{L}\phi R + G_2 \bar{L}\phi_c R + \text{h.c.})}_{\text{fermion masses and couplings to Higgs}}
\end{aligned} \tag{2.26}$$

2.2. Measurement Predictions

The rest of this chapter provides a concise overview of how the theoretical framework can be tested experimentally using particle detectors such as ATLAS. It also introduces the phenomenology of the Higgs boson production and decay modes of the Higgs boson, with an emphasis on $H \rightarrow WW^*$.

2.2.1 Cross sections

Protons are composite particles consisting of two up quarks and one down quark. The valence quarks interact through gluons, which can subsequently generate quark-antiquark pairs. Consequently, a pp collision involves multiple interactions among the constituents of the protons, accompanied by a spectrum of four-momentum transfers.

The cross-section (σ) is a measure of the likelihood for a particular process to occur during any proton-proton collision. The cross-section is a measure of the “target area” where a particular interaction can take place. As the cross-section increases, the probability of interaction also increases. The cross-section for some specific process can be calculated using the Feynman rules, which are determined by the Lagrangian of the underlying theory. A cross-section for a quantum mechanical transition $ij \rightarrow X$ is given by

$$\sigma_{ij \rightarrow X} = \int \mathcal{P} d\Omega \tag{2.27}$$

and,

$$\hat{\sigma}_{ij \rightarrow X} = \sum_{n=0}^{\infty} \alpha_S^{m+n} \hat{\sigma}_{ij \rightarrow X}^{(n)} \tag{2.28}$$

where $\mathcal{P}$ is the transition probability. Moreover, m is the number of QCD vertices for the

simplest Feynman diagram of a process $ij \rightarrow X$, and $\hat{\sigma}_{ij \rightarrow X}^{(n)}$ is the n^{th} order cross-section in the strong coupling constant, α_S . Here, the strong coupling constant, α_S , is large at low energy scales (long distances) but small at high energy scales (short distances). Therefore, high-momentum transfer interactions generating particles like the Higgs boson can be at the LHC are calculated as a perturbative series in terms α_S . The first non-zero term, i.e., $n=0$, in the series is regarded as leading-order (LO) and offers a preliminary estimate of the cross-section. The following terms are referred to as next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO), and so forth, and incorporate supplementary corrections.

Each term in the expansion can be regarded as a compilation of diagrams, commonly known as Feynman diagrams, which collectively contribute to the overall process and possess an identical number of internal vertices. Utilizing the Feynman rules, derivable from the theory's fundamental Lagrangian, one can convert each diagram into a matrix element $\mathcal{M}$, which is subsequently employed to find the corresponding cross-section. The effect of missing higher-order terms in a fixed-order calculation can be estimated by considering its residual dependence on the renormalization scale, μ_R .

2.2.2 Parton density functions

The pp cross-section of an interaction is a convolution of the partonic cross-section with the respective parton distribution functions (PDFs) of each colliding proton.

$$\sigma(p_1, p_2) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij}(\mu_F^2, \mu_R^2) \quad (2.29)$$

where f_i term is the PDF of the parton i in one proton, x_i is the momentum fraction of parton i .

PDFs describe the distribution or likelihood of finding partons (quarks and gluons) within a hadron, providing information about the momentum and flavor of these constituents. PDFs aim to describe the non-perturbative characteristics of hadrons that cannot be computed analytically but can be obtained by conducting comprehensive data fits at various scales. The DGLAP evolution equations [38, 39, 40, 41] govern the relationship between the energy scale and the dependence of the PDFs. The PDF set is the collection of PDFs that is obtained as a result.

PDFs cannot be predicted from first principles, which necessitates obtaining them for common hadron types through a combined fit to numerous individual experimental measurements. Numerous collaborations offer PDF sets that vary in experimental data, parameterization models, treatment of theoretical parameters and values, and uncertainties. These collaborations encompass CTEQ [42], MRST2007 [43], NNPDF [44], and PDF4LHC [45]. Figure 2.4 shows the latest outcomes of the proton PDFs calculated by the CTEQ collaboration.

PDFs are instrumental in calculating the cross-section of a parton-level process.

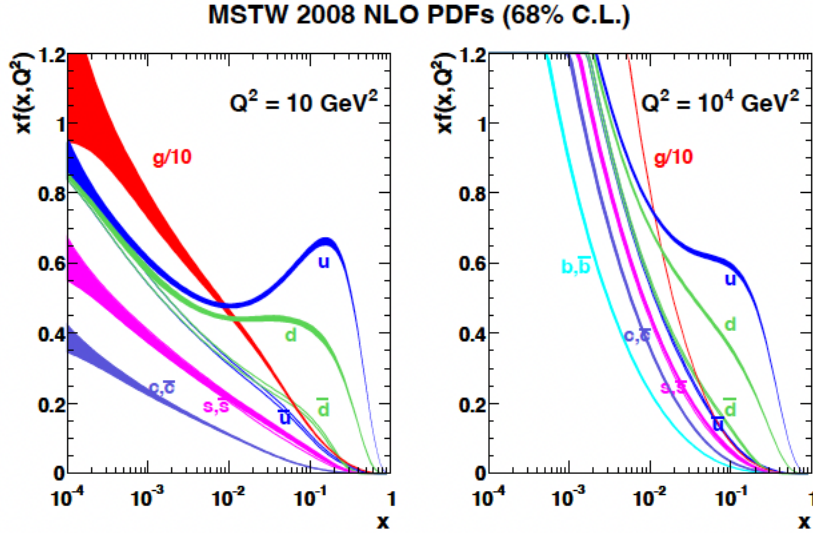


Figure 2.4: The parton distribution functions for the proton at $Q = 2$ GeV (left) and $Q = 100$ GeV (right) for quarks as well as gluons, where Q is the probe energy.

2.2.3 Underlying events

Additional interactions and processes can occur alongside the primary hard scattering process during a particle collision, depending on the density profile of the proton. Moreover, in some collisions, more than one parton (quark or gluon) from each colliding proton may interact, leading to additional scattering and secondary processes, known as Multiple Parton Interactions (MPI). These secondary processes, along with MPIs and the fragmentation of beam remnants, are commonly known as the underlying event (UE). ATLAS typically calculates the underlying event by studying energy deposition and tracking activity in regions of the detector that are less affected by the hard scatter, using the transverse region method or jet area subtraction techniques while applying pileup corrections. The energy scale of the UE is lower than that of the hard scatter and is typically characterized by a uniform distribution of soft activity in the detector. However, correctly calculating the underlying events typically requires using Monte Carlo generators.

2.2.4 Parton shower and Hadronization

Parton Showers is the process of simulating the radiation of additional particles during a high-energy collision in particle physics. During a particle collision, highly energetic partons may emit additional QCD particles, such as quarks and gluons. The parton shower is typically simulated using a Monte Carlo generator, which operates on a per-event basis and is often combined with the underlying event.

Hadronization refers to when quarks and gluons, which are elementary particles, combine to form hadrons, such as protons and neutrons. At a scale near Σ_{QCD} around 200 MeV, due to the confining nature of QCD, the particles with color charge (gluons and

quarks) create color-neutral composites, hadrons. The process is referred to as hadronization and cannot be calculated using perturbative methods due to the high values of the strong coupling constant at low momenta. Event generators use two hadronization models:

- Lund string fragmentation model [46]: This model is based on the assumption that QCD shows Coulomb-like behavior at a short distance but transitions to linear behavior at a longer distance due to gluon self-attraction. A strong color field induces the creation of $q\bar{q}$ pairs, resulting in hadronization. It emphasizes the role of one-dimensional strings, their breaking, and subsequent fragmentations in creating color-neutral hadrons.
- Cluster model [47]: Unlike the Lund String model, the Cluster model does not involve the concept of one-dimensional strings. Instead, based on the idea of pre-confinement, it focuses on the formation of clusters of quarks and gluons that eventually result in color-neutral hadrons.

The choice between these models often depends on the specific characteristics of the experimental data being analyzed. It can impact various detector effects, such as jet response, heavy-flavor tagging, and lepton isolation.

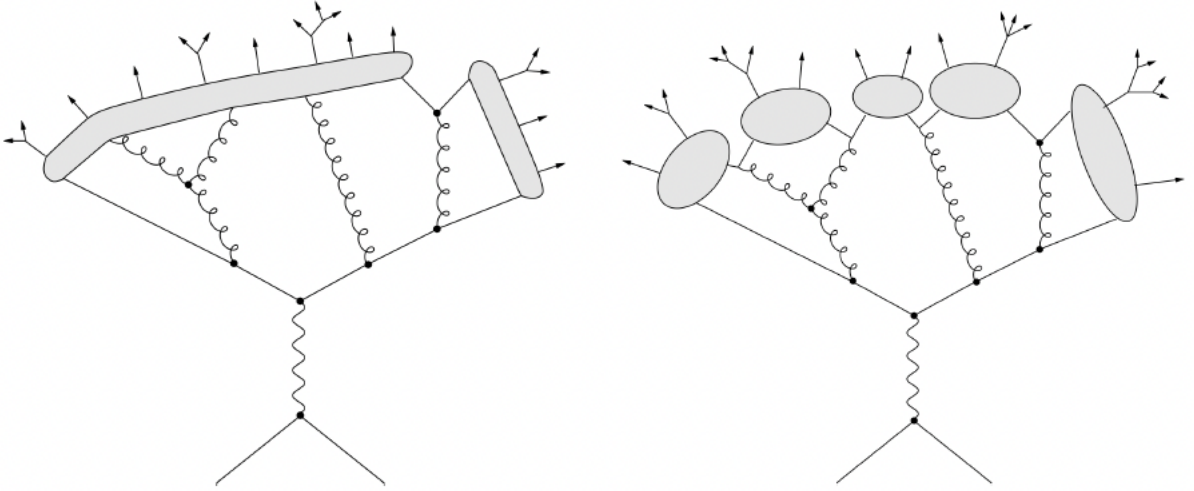


Figure 2.5: The two hadronization model. The Lund String (left) and Cluster (right) models.

2.2.5 Pile-up

Detectors are designed to measure the products of individual collision events precisely. However, multiple pp collisions can occur in quick succession in a particle detector, known as a pile-up. This results in the overlapping of signals from different collision events, making it challenging to distinguish and accurately reconstruct the hard scatter process of interest. The occurrence of pile-up events is a consequence of the high collision rates and

luminosities. There are two distinct categories of pile-up: in-time pile-up and out-of-time pile-up.

In-time pile-up refers to the occurrence of multiple proton-proton collisions happening simultaneously within the detector during the same bunch crossing. The particles generated in these additional pile-up events hinder the reconstruction of the primary event of interest. The excess energy accumulated in the detector due to pile-up will impact the measured energies in the calorimeter and significantly impact the lepton isolation and the measurement of jet energies. Out-of-time pile-up happens when signals from previous bunch crossings interfere with the current bunch crossing. This type of pile-up mainly affects ECAL measurements. While out-of-time pile-ups can cause a decrease in energy resolution, it is generally a less significant issue compared to in-time pile-ups. Understanding and dealing with pile-ups is instrumental in improving the precision of measurement. Ways to deal with the pile-ups are described in later chapters.

2.2.6 Luminosity

The instantaneous luminosity, denoted as $\mathcal{L}$, is a parameter that defines the characteristics of an accelerator and depends on its construction features:

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \quad (2.30)$$

where n_1 and n_2 are the particles contained in the bunches colliding at a frequency f , and the σ_x and σ_y are the transverse beam widths. Moreover, the integrated luminosity L is defined as the time integral of the instantaneous luminosity:

$$L = \int \mathcal{L} dt \quad (2.31)$$

Physics experiments aim to maximize the event rate ($\frac{dN}{dt}$) achievable.

$$\frac{dN}{dt} = \mathcal{L} \times \sigma \quad (2.32)$$

The relationship between the cross-section and the number of expected events in an experiment is given by the equation.

$$N = \sigma L = \int \mathcal{L} dt \quad (2.33)$$

Hence, higher values of luminosities correspond to higher amounts of data collected and are often expressed in units of ‘barns’ (1 barn = 10^{-24}cm^{-2}). Figure 2.6 shows the total luminosity collected at $\sqrt{s} = 13 \text{ TeV}$, which provides the majority of the data analyzed. The analysis is performed using the full Run 2 data set, corresponding to an integrated luminosity of $140 \text{ fb}^{-1} \approx 1.1 \times 10^{16} \text{ pp collisions}$.

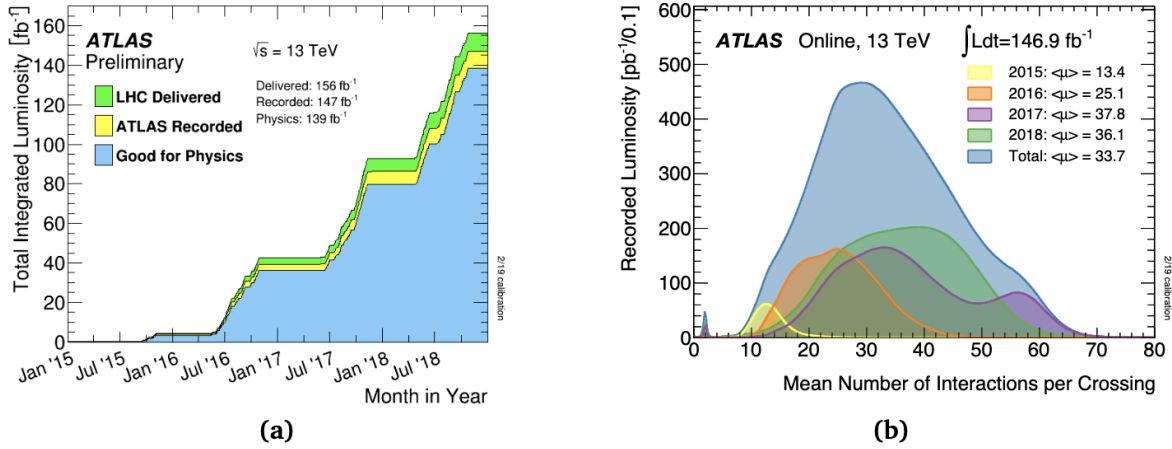


Figure 2.6: The total integrated luminosity produced and recorded between 2015 to 2018. (b) The average number of interactions per bunch crossing during the data taking.[1]

2.3. Monte Carlo Simulation

Monte Carlo (MC) event generators are utilized to provide accurate predictions and simulations of pp collision events. These generators have been calibrated using observed data from multiple measurements, resulting in a high level of agreement with data collected by the ATLAS detector. Several programs with distinct algorithmic implementations of particle physics are used to simulate a full collision event. The following MC event generators are employed in this thesis:

- PYTHIA [48] is a multi-purpose event generator capable of calculating the partonic cross-section, parton shower, hadronization, and underlying event. The parton shower simulation involves the arrangement of color-charged dipoles in order of their transverse momentum. The Lund string model is employed to simulate hadronization.
- HERWIG [49] is utilized to compare the underlying event, hadronization, and parton shower as simulated by PYTHIA. This is because HERWIG uses an angular-ordered shower for the parton shower, while the cluster model is used for the hadronization process.
- SHERPA [50] is specifically designed to simulate full events unlike PYTHIA and HERWIG. It uses the same technique as PYTHIA to simulate the parton shower, whereas it uses the cluster model for hadronization. SHERPA is favored over the previous two generators for the partonic cross-section, as it computes at next-to-leading order (NLO) instead of LO by the previous two.
- MADGRAPH [51] computes partonic cross-sections at either LO or NLO. It is commonly interfaced with PYTHIA.

- PoWHEG [52] simulates partonic cross-section at NLO or higher levels. Like MADGRAPH, it is interfaced to PYTHIA to simulate UEPS.

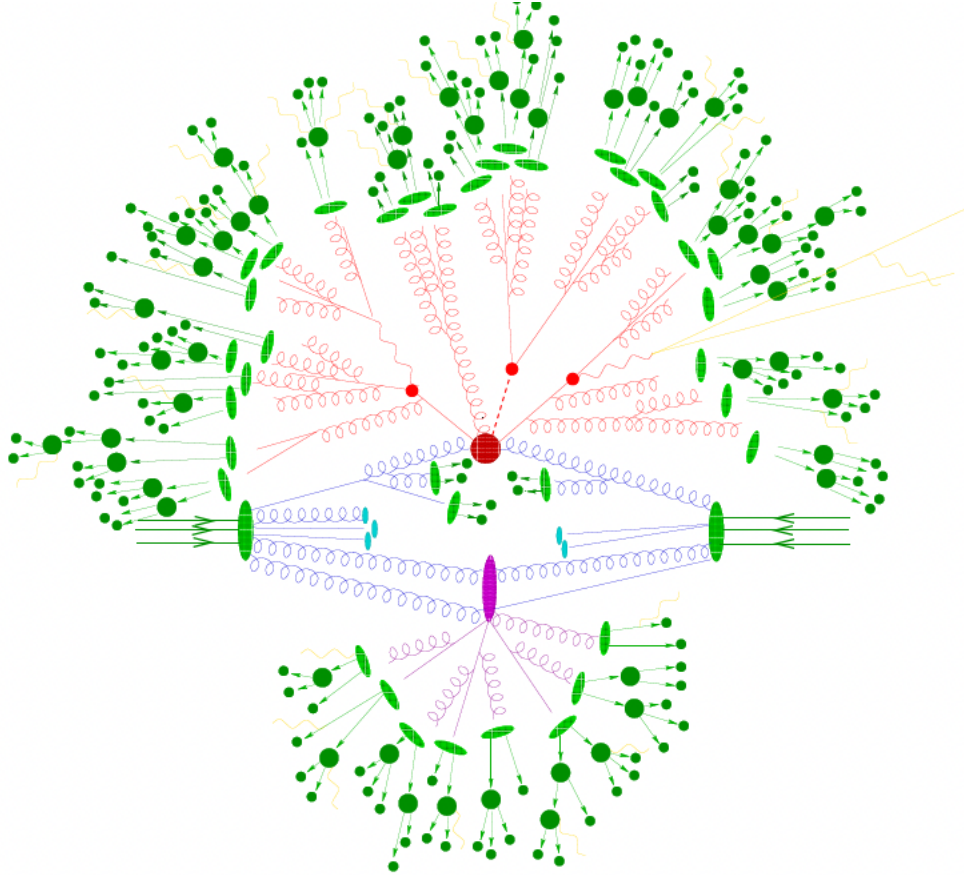


Figure 2.7: Illustration of a pp collision as simulated by a Monte Carlo event generator. The large red circle indicates the primary hard-scatter process.[2] The emitted particles originating from partons in both the initial and final state are depicted in blue and red. The light green blobs represent the parton hadronization, while the dark green blobs represent the subsequent decay of the hadrons. Yellow lines represent the emission of low-energy photons. The blue and purple ovals represent the beam remnant and underlying event, respectively.

A complication arises when combining a partonic cross section determined at next-to-leading order (NLO) or higher with a parton shower simulation, as this can lead to double-counting of emissions. NLO calculations include leading-order (LO) diagrams, virtual corrections via loop diagrams, and additional emissions, which may overlap with emissions simulated by the parton shower. To address this issue, two approaches, matching and merging, are employed. Matching avoids double-counting by modifying the NLO calculation to depend on the specific parton shower used, as implemented in MADGRAPH5_aMC@NLO [51]. Alternatively, in programs like PoWHEG, the emission with the highest momentum is simulated in the partonic cross section, and parton shower probabilities are adjusted

accordingly.

Parton showers are generally designed for low-momentum collinear emissions, with high-momentum emissions better handled by fixed-order calculations. Double-counting can also occur when merging exclusive partonic cross-sections with varying emission multiplicities. To mitigate this, SHERPA employs algorithms like CKKW-L and MEPS@NLO [53, 54, 55, 56], which correct the two highest parton shower emissions to align with the partonic cross section. Thresholds are introduced to determine whether a hadronic activity is modeled using the partonic cross section or the parton shower.

2.4. Statistical Data Analysis

The process of hypothesis testing is used to assess various aspects of a recently identified particle. Two hypotheses are compared with the observed data. The null or background-only hypothesis, represented as H_0 , posits that the prediction of the Standard Model is accurate and any observed deviation is solely due to random chance. On the other hand, the alternative hypothesis or signal plus background hypothesis, represented as H_1 , posits that the observed deviations are a result of a non-random cause.

In order to measure the agreement between the observed data and the prediction made under either hypotheses, a test statistic t is used. The p -value, under either of the two hypotheses, is the probability of measuring deviations from the predictions that are at least as large as the deviations that were actually observed. In particle physics, the p -value is transformed into an equivalent significance, which is defined as the number of standard deviations above the mean of a Gaussian distributed variable with a probability equal to p . Two specific scenarios are particularly interesting. A standard deviation of 5σ with respect to the expectation value is regarded as a discovery, which is equivalent to $p < 2.87 \times 10^{-7}$. The p -value enables the rejection of the hypothesis under consideration if the calculated p -value is lower than this threshold.

If the hypotheses are parametric, the test statistic is constructed using likelihood functions ($\mathcal{L}(\mu, \vec{\theta}|x)$). The likelihood function represents the probability that a specific set of parameter values accurately represents the true values of the model, given a dataset x obtained from that model. Here, a conceptual differentiation is established between two categories of parameters.

2.4.1 Parameter of interest (PoI)

A PoI (denoted by μ) refers to a parameter, typically a free parameter in H_1 , but is either absent or fixed to a particular value in H_0 . In this study, the signal strengths are PoIs since the measurement's primary goal is to make statistical inferences about them.

2.4.2 Nuisance parameter

A nuisance parameter (denoted by θ) is a parameter that is commonly found in both the H_1 and H_0 . It is not directly measured but must be modeled and is constrained to specific ranges using prior distribution functions. These are parameters that are included in order to account for systematic uncertainties.

Here, large values of $\mathcal{L}$ are associated with situations in which the observation of the given data set is more probable. And the test statistics t in terms $\mathcal{L}$ of is given by:

$$t_\mu = -\ln \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}$$

and the profile likelihood ratio $\Lambda(\mu)$ is then defined as

$$\Lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}$$

where, $\hat{\mu}, \hat{\theta}$ represent unconditional maximum likelihood estimator. These estimators correspond to the values of μ and θ that maximize the likelihood function in absolute terms. Alternatively, $\hat{\hat{\theta}}(\mu)$ represents the maximum likelihood estimator of θ conditional on μ . In other words, the function $\hat{\hat{\theta}}(\mu)$ maximizes the values of θ for a specific value of μ .

The profile likelihood function is used to test the null and alternate hypotheses. The parameter $\Lambda(\mu)$ is constrained within the interval $[0, 1]$, with $\Lambda(\mu) \sim 1$ denoting the highest level of concordance between the tested value of μ and the observed dataset.

2.4.3 Asimov data set

The Asimov data set [57] is a data set that can be used in particle physics, particularly to test and validate the statistical model. It is the ideal expected data under a specific hypothesis, which is typically the null hypothesis. This dataset is beneficial for estimating the anticipated sensitivity of an experiment, enabling researchers to make predictions regarding the extent to which a specific hypothesis can be tested without statistical fluctuations. In this analysis, the statistical model is developed blindly, i.e., the statistical analysis does not take into account the observed data in the SRs. The process involves generating the Asimov dataset, refining and validating the statistical analysis, and subsequently utilizing the observed data to extract the final measurement results. The Asimov data set yields are determined by combining the theoretical predictions of the signal and background processes considered during the analysis. Consequently, the pulls of NPs are zero, and the values of the POI and background normalizations are unity, by construction. In order to construct a robust statistical model, estimate the expected precision of the measurement, and comprehend the performance of the statistical model without introduc-

ing cognitive bias into the measurements, it is crucial to carry out the statistical analysis using the Asimov data set rather than the observed data set.

2.5. Higgs physics at the Large Hadron Collider (LHC)

The ATLAS and CMS collaboration announced the detection of the Higgs boson on July 4, 2012, during Run I of the LHC [16, 17]. This discovery resulted in Francois Englert and Peter Higgs being awarded the Nobel Prize in Physics in 2013. The Large Hadron Collider (LHC) is the most powerful particle collider ever constructed. Chapter 3 provides a more comprehensive explanation of the Large Hadron Collider and its experiments. This section discusses the main Higgs production modes and their decay channels at the LHC.

2.5.1 Production modes of the Higgs boson

The Large Hadron Collider (LHC) can generate the Standard Model Higgs bosons in proton-proton collisions through various mechanisms. Figure 2.8 shows the leading-order Feynman diagrams and Figure 2.9 shows the relative contributions of these mechanisms at various center-of-mass energies.

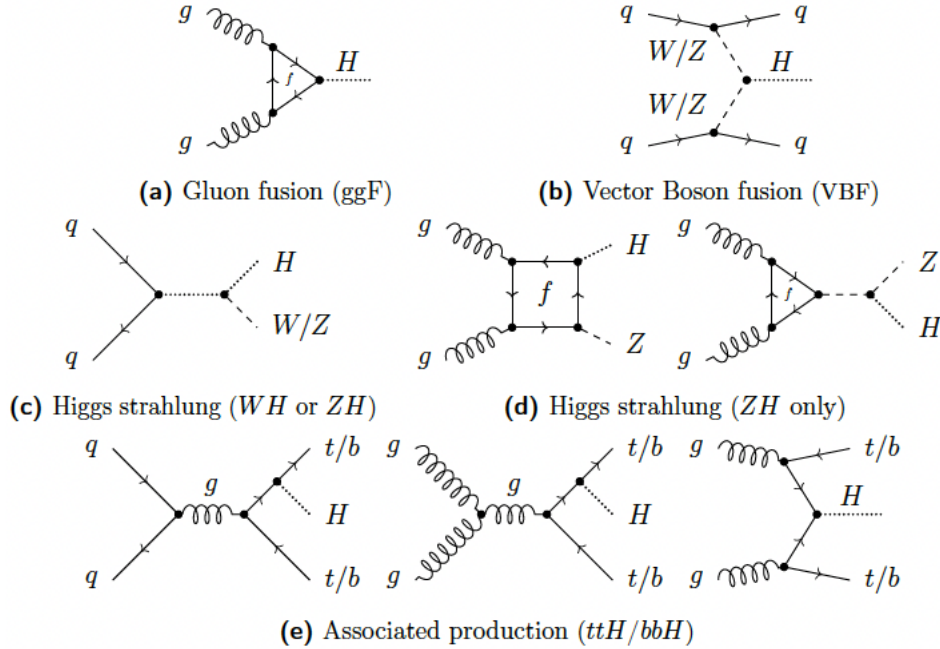


Figure 2.8: The main modes of the Higgs boson production at LHC.

- **Gluon-gluon fusion (ggF):** This is the dominant process for producing Higgs bosons at the LHC. The mechanism involves the fusion of two incoming gluons mediated by a virtual fermion loop. The primary contributions arise from loops involving the top quark, the heaviest fermion.

- **Vector boson fusion (VBF):** The second largest production mode involves the fusion of two vector bosons, such as W or Z bosons, resulting in the final state qqH . Here, these vector bosons are produced by fusing two incoming quarks. The two outgoing, energetic quark jets in the forward region are targeted for event selection for this production mode.

- **Higgs Strahlung (VH):** A Higgs boson is produced in association with a vector boson, such as a W or Z boson. The likelihood of this production mode is lower than VBF; it can be targeted by searching events with an additional vector boson in the final state.

- **Top quark associated production (ttH):** This process is produced in association with a top quark and antiquark pair, which, in turn, is emitted from the gluon splitting. Due to the low production cross section and less accessible final state signature, the measurement of this mode is challenging.

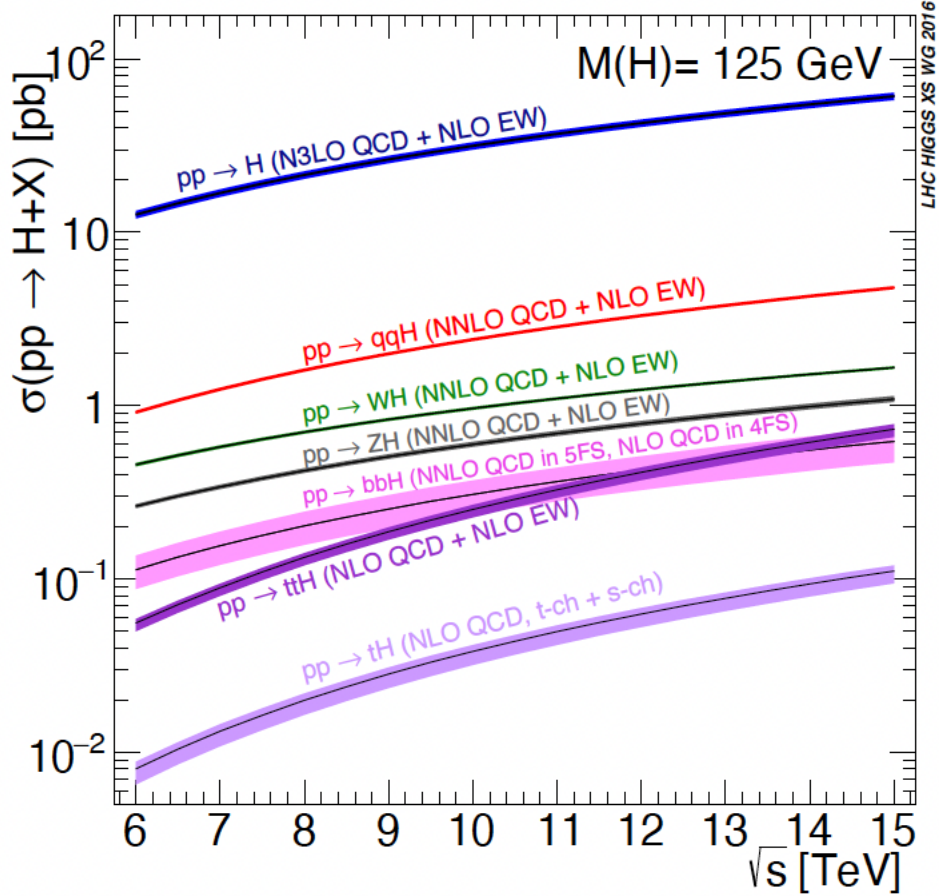


Figure 2.9: The predicted SM Higgs boson production cross sections as a function of the center-of-mass energy. Line widths represent the respective theory uncertainties.[3]

The ggF and VBF production processes have been confirmed to meet the threshold of 5 sigma, as evidenced by the latest measurements conducted by the ATLAS and CMS collaborations.

2.5.2 Decay channels

Due to its relatively high mass, the Higgs boson is very short-lived, with a predicted decay width of 4.1 MeV. Therefore, in order to detect events containing a Higgs boson, the final state particles of its decay modes must be examined since they will decay before interacting with the detector material. As illustrated in Figure 2.10, Higgs bosons decay to various final states. Moreover, Figure 2.11 shows the leading branching fractions of Higgs boson decays. $H \rightarrow b\bar{b}$ decay mode is the dominant decay mode for a Higgs boson with a mass of $m_H = 125$ GeV, accounting for about 58% of the total decay events. However, measuring Higgs boson decays to this specific final state presents challenges because it is exclusively hadronic, offering minimal distinctive features to set it apart from other occurrences observed in a hadronic collision. In other words, it would be contaminated with a huge QCD background.

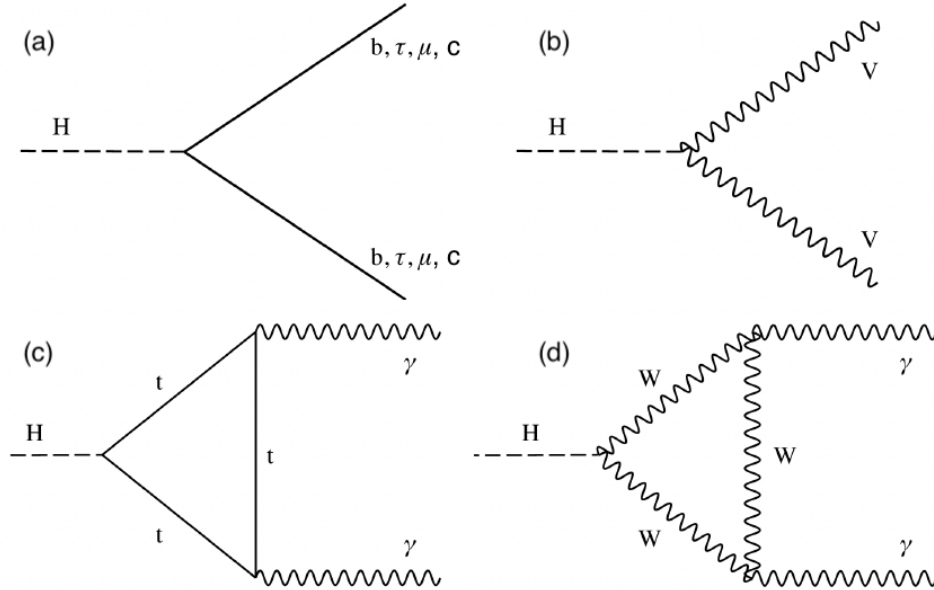


Figure 2.10: The Feynman diagrams representing the main Higgs boson decays.

The second most common decay mode is $H \rightarrow WW^*$, where one W boson is off-shell given the measured Higgs boson mass, occurring about 21% of the time. This decay channel does not encounter the issue of $H \rightarrow b\bar{b}$, as the W bosons can undergo leptonic decay. The final state with leptons has a clean experimental signature, resulting in high signal-to-background ratios. This decay is the focus of this thesis.

Furthermore, the decay process into a pair of tau leptons does not exhibit any distinctive characteristics to distinguish it from backgrounds in a hadronic collision. Decays

involving gluons and $c\bar{c}$, similar to $H \rightarrow b\bar{b}$, are challenging to distinguish from the background events at the hadron collider. Even though the decay to pairs of Z bosons or diphotons has a small branching ratio, it can be achieved with great efficiency thanks to the clean signature. The decays $H \rightarrow Z\gamma$ and $\mu\mu$ are extremely rare, requiring more integrated luminosity.

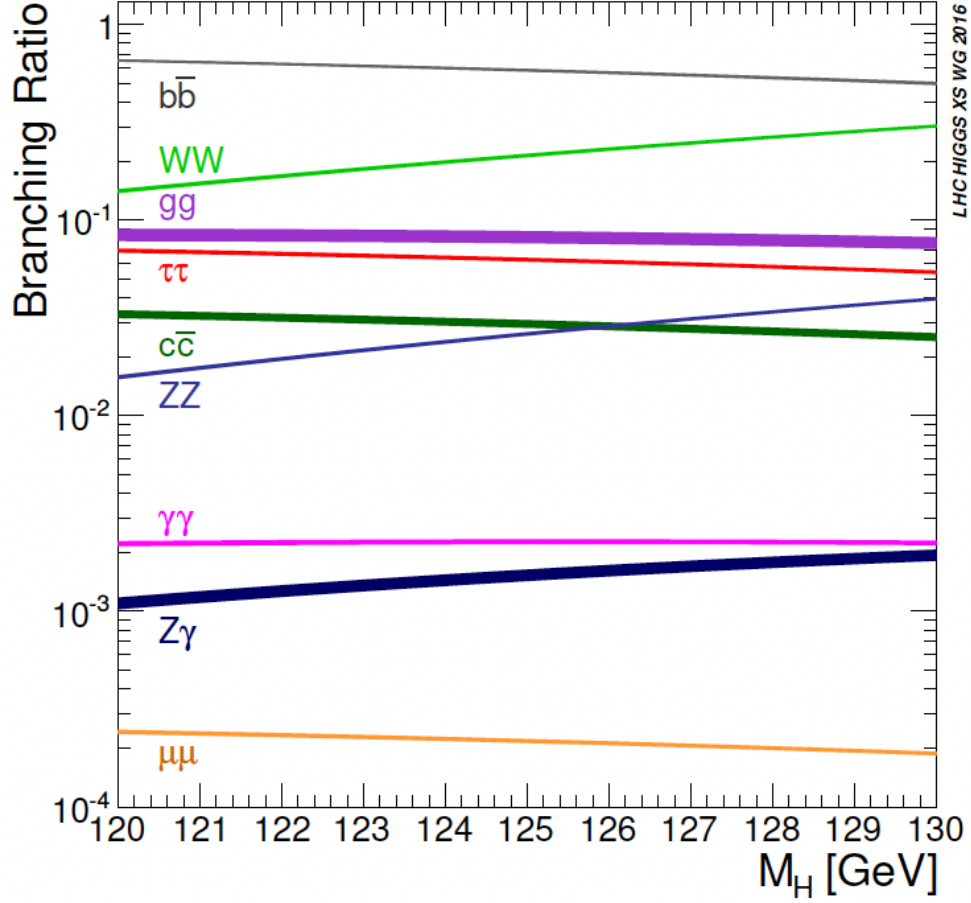


Figure 2.11: The predicted branching ratios for the SM Higgs decay modes near $m_H = 125$ GeV.[3]

3 | Experimental Setup: ATLAS

Particle colliders are crucial in generating controlled experiments to observe and understand particles and their interactions. The Large Hadron Collider (LHC) shown in Figure 3.1 is the largest and the most powerful particle collider ever built, capable of colliding beams of protons or heavy ions at a high center of mass energies. It is operated by the European Council for Nuclear Research (CERN) and is located near Geneva, Switzerland.

Four major detectors are strategically positioned at the interaction points of the LHC. These detectors reconstruct events precisely by identifying the particles generated in the collisions and by analyzing particle types, their kinematics, and energies. The dataset used for the results presented in this thesis was collected with the ATLAS detector, one of the main experiments at the LHC.

This chapter introduces the layout and functionality of the LHC and ATLAS detectors. For a detailed description of the hardware aspects, refer to the original technical design reports [58, 59, 60, 61] and reports on upgrades [62, 63, 64]. The final section of this chapter also provides an overview of the measurements of the Standard Model at the LHC. Additionally, it highlights the achievements of the ATLAS and CMS collaboration in discovering the Higgs boson in 2012.

3.1. Large Hadron Collider

The Large Hadron Collider (LHC) is a two-ring, superconducting accelerator and collider housed inside a 26.7 km long circular tunnel buried underground at an average depth of 100 m. It sits across the border between Switzerland and France, near Geneva, amidst the Jura mountains and Lake Geneva. The LHC is situated in the same tunnel that formerly accommodated the Large Electron Positron Collider (LEP), which was disassembled around 2001. It can collide proton pairs with a high center-of-mass energy of 13.6 TeV as of 2023. The LHC is a synchrotron that guides and accelerates the proton beams to their maximum energy in opposite directions in separate beam pipes. It then directs them to collide head-on at the center of four detectors positioned along the ring: ATLAS, CMS, ALICE, and LHCb.

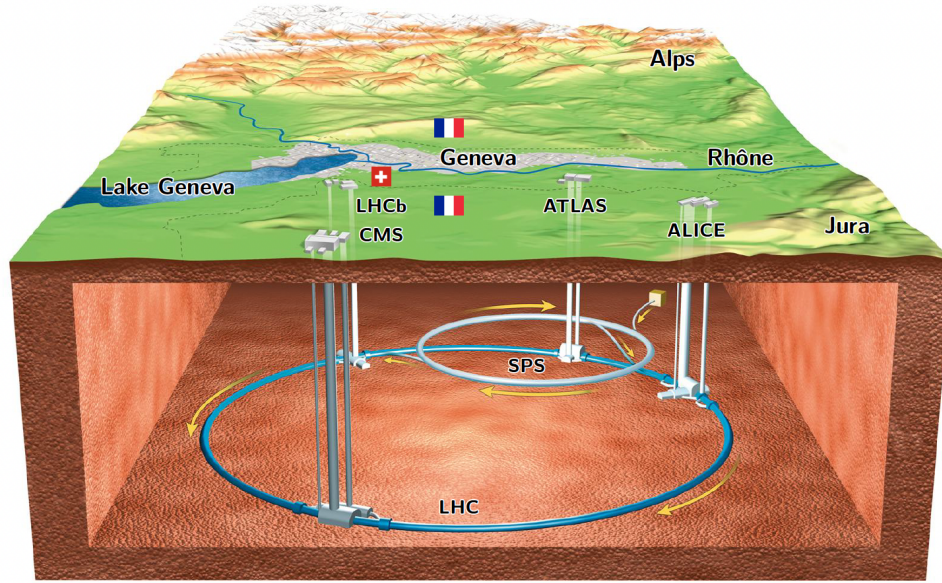


Figure 3.1: Display of the LHC machine and its four main experiments.

3.1.1 LHC Hardware Configuration

The LHC is partitioned into octands. One octand houses the acceleration system, which consists of superconducting radiofrequency cavities. Two octands contain the beam cleaning systems, which removes particles with large momentum offset or betatron amplitudes. One octand houses the beam abort systems, which terminate each LHC run and direct the remaining particles into beam dumps. The remaining four octands serve as the interaction points where the particles collide.

In order to bend the protons around the ring, powerful superconducting dipole magnets are employed to generate the magnetic field. Each 15 m-long superconducting dipole magnet is made out of cables composed of Niobium-titanium (NbTi) filaments. A series of about 1200 of these dipole magnets are cooled to a temperature of 1.9 K with superfluid helium and can reach magnetic fields of up to 8.33 T. The magnetic field maintains the hadrons in their circular orbit. Smaller quantities of various dipole, quadrupole, sextupole, and octupole magnets are used to focus the beam and make minor corrections to the beam direction.

The protons are accelerated and packed in distinct bunches by subjecting them to an oscillating longitudinal electric field in 16 LHC radio frequency (RF) cavities [65], which also act to restore energy losses. Each bunch contains about 100 billion protons and is separated by 25 ns in time, providing a bunch collision rate of 40 MHz.

3.1.2 The path of protons

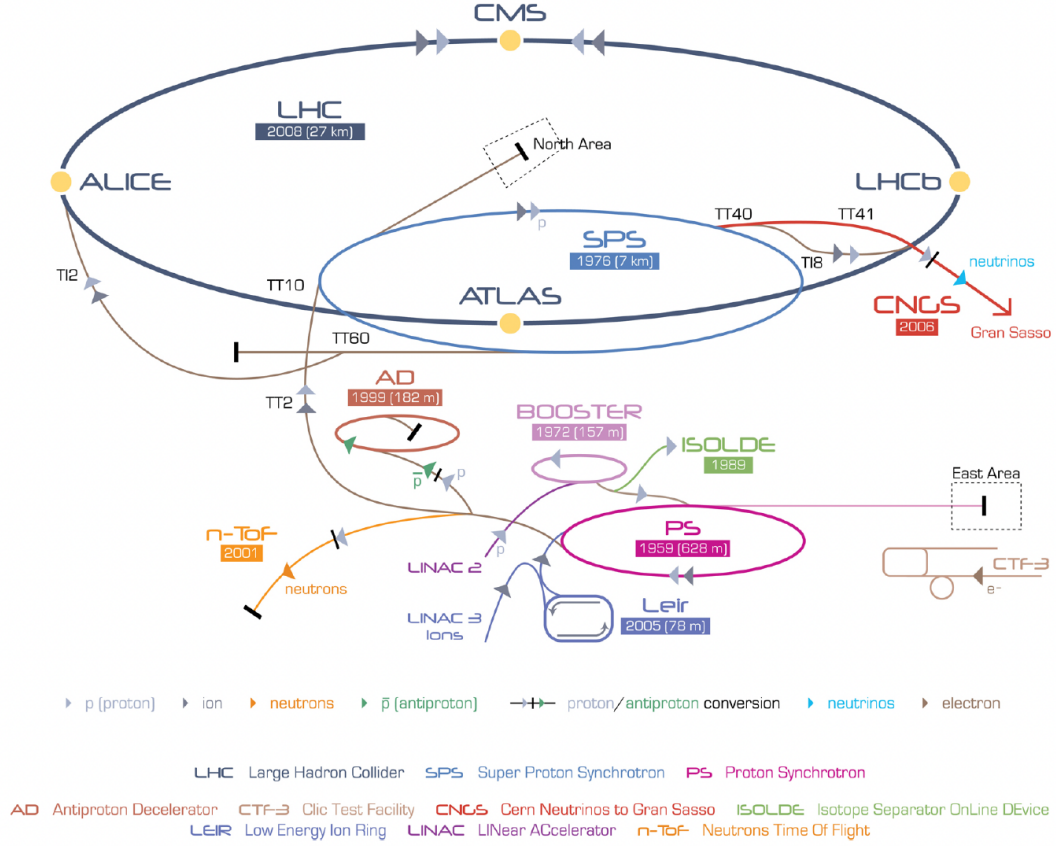


Figure 3.2: The CERN accelerator complex, showing the LINAC2, the Booster, the PS and the SPS, the LHC, and other experimental setups [4]

A schematic of the CERN accelerator complex is shown in Figure 3.2. The proton journey starts in the LHC by feeding hydrogen atoms at a precisely controlled rate into a LINear ACcelerator, LINAC2, where the electrons are stripped off to leave hydrogen nuclei, i.e., proton. These protons are then accelerated to 50 MeV. They are then fed into a synchrotron called the Booster, which increases the energy to 1.4 GeV. In order to maximize the intensity of the beam, the packet is distributed among four stacked rings. Subsequently, they are recombined and sent into the Proton Synchrotron (PS), increasing their energy to 26 GeV. The point of transition is reached, where the energy added to the protons by the oscillating electric field cannot translate into any significant or noticeable increase in the velocity and is primarily a mass increase of the proton.

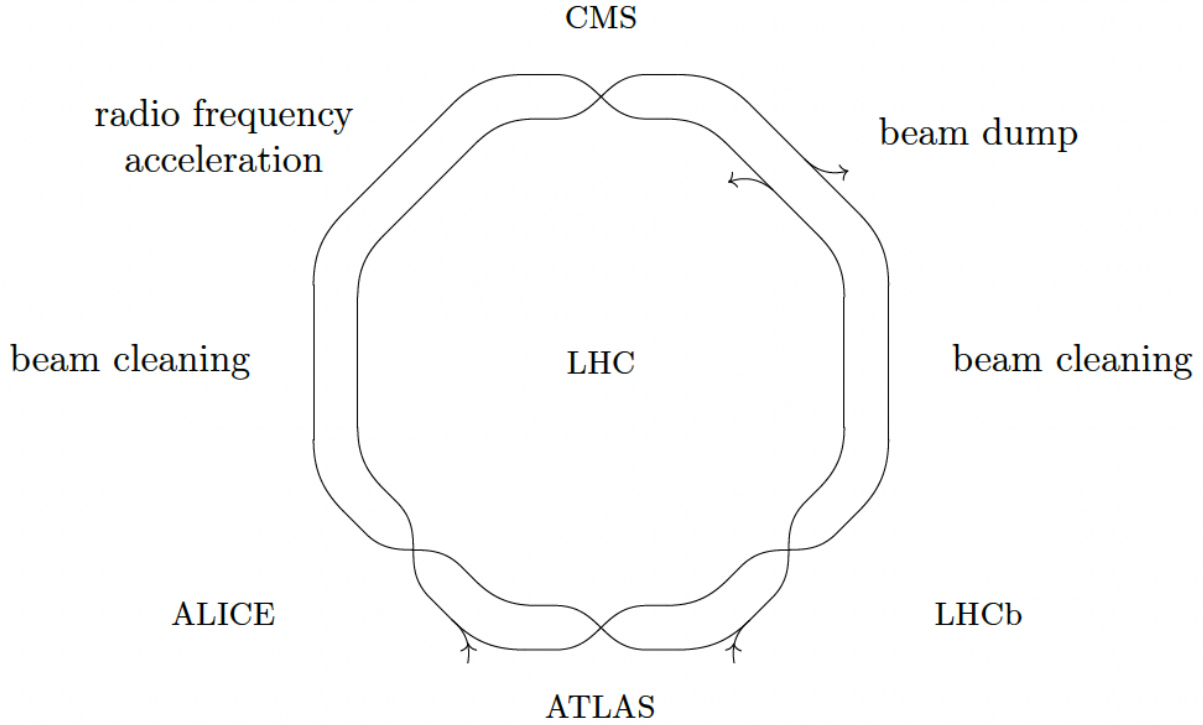


Figure 3.3: Schematic display of the octagonal structure of the LHC with interaction points and facilities

Afterward, the proton beams are injected into the Super Proton Synchrotron (SPS), where their energy is around 450 GeV, and then launched into the orbit of LHC. The LHC has two vacuum pipes containing proton beams circulating in opposite directions. Using sophisticated kickers, incoming packets are synchronized with those already in circulation. Here, each proton has an energy of nearly 7 TeV. The counter-rotating beams crossover in four detector caverns where they are made to collide. The detectors track the debris of these collisions. Each LHC operation takes place in fill cycles, ideally lasting around 10 hours, until the beam deteriorates so much that it is more efficient to dump and refill. Up to 2808 bunches are injected into the LHC from SPS under normal conditions. In 2016, the LHC delivered stable beams to the experiments for about 60% of its operational time, delivering 40 fb^{-1} to the ATLAS and CMS experiments in a year.

3.1.3 Experimental facilities located in CERN

- The ATLAS and CMS detectors [66] are general-purpose detectors designed to achieve precise SM physics measurements and search possible signs of new physics, especially beyond the Standard Model (BSM) physics independently. Even though the purpose of both detectors is the same, they differ in design to cross-check any discovery. In July 2012, the Higgs boson was discovered using these detectors.

- The LHCb experiment [67] is designed to measure CP violation and rare decays of bottom and charm hadrons.
- The ALICE experiment [68] is designed to study the physics of quark-gluon plasma at high energy density and temperature in heavy ion collisions. This creates energy densities similar to those immediately after the Big Bang.

Other small experimental facilities, such as TOTEM [69], LHCf [70], FASER, and MoEDAL[71], specialize in measuring other physics phenomena like cross-sections of scattering and magnetic monopoles.

3.2. The ATLAS Detector

The ATLAS (A Toroidal LHC Apparatus) detector [5] is a general-purpose particle detector located at the LHC collision point closest to the main CERN site. The purpose of the detector is to investigate various interaction processes, including precisely measuring the SM predictions, measuring the properties of the Higgs boson, and exploring BSM processes. The ATLAS detector must accurately identify and reconstruct all the physics objects generated during a hadronic collision to achieve this objective.

Figure 3.4 shows a cut-away view of the ATLAS detector, the largest particle detector ever built, with a length of 44 m, a diameter of 25 m, and a weight of 7000 tons. The sensors are strategically positioned to cover almost the entire solid angle surrounding the interaction point. It provides almost omnidirectional particle identification and precision measurement of energy and momentum of particles released from the collision. The ATLAS structure is multilayered, onion-like; each layer serves a specific function. The beam pipe is positioned along the cylindrical axis of symmetry of the detector. This section briefly overviews these concentric subsystems, including their function, design, characteristics, and the mechanisms involved in data taking.

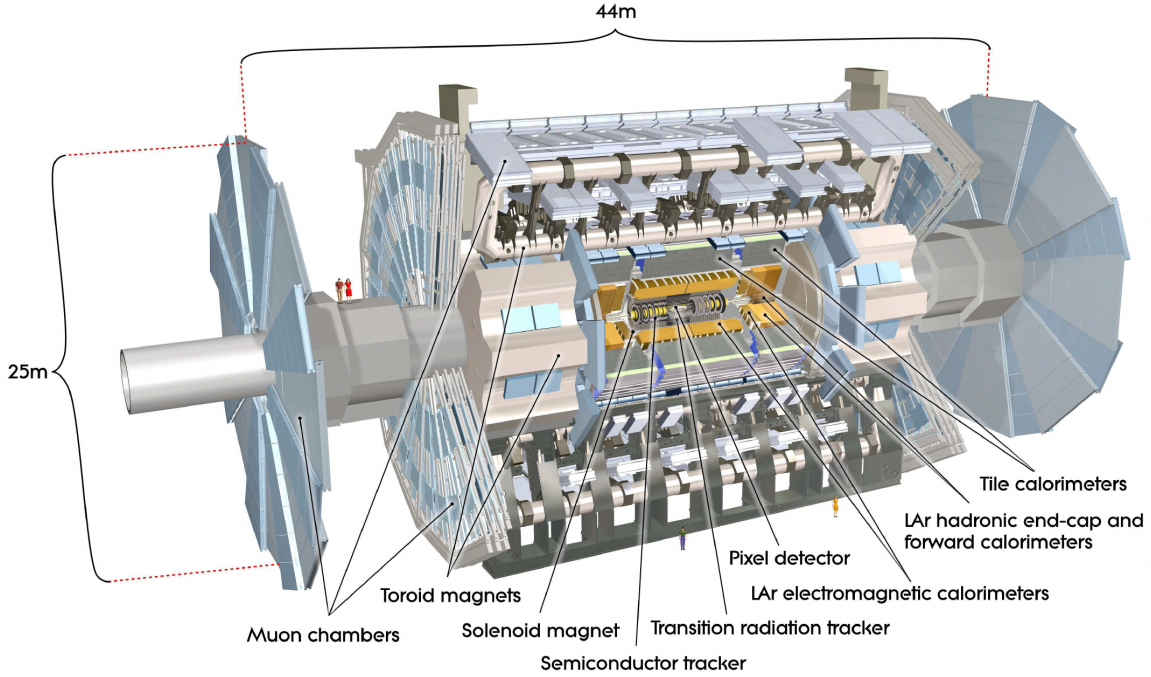


Figure 3.4: A Cut-away view of the ATLAS detector to show the scale of the detector [5]

The ATLAS detector uses a right-handed Cartesian coordinate system originating at the nominal interaction point in the center of the detector. The z -axis is tangential to the LHC ring. The x -axis is perpendicular to that, pointing radially inward towards the center of the LHC ring, while the y -axis is directed vertically upwards. The azimuthal angle ϕ , ranging from 0 to 2π , is measured around the beam axis against the x -axis. The polar angle $\theta \in [0, \pi]$ is measured relative to the beam axis. However, using the pseudorapidity $\eta = -\ln(\tan(\frac{\theta}{2}))$ is more common. Transverse momenta and energies are defined within the $x - y$ plane. The separation in kinematic space between two physics objects can be measured using

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

3.2.1 The Inner Detector

The Inner Detector (ID) is the closest subdetector to the interaction point and measures the trajectories or tracks of charged particles produced from collision with high spatial resolution. Furthermore, measuring these trajectories in a magnetic field measures the particle's momentum. Figure 3.5 shows a schematic view of the inner detector. The ID comprises the Pixel Detector, the Semiconductor Tracker (SCT), and the transition radiation tracker (TRT). These subdetectors are built using either silicon sensors or straw drift tubes. The ID is immersed in a 2T axial magnetic field provided by a superconducting

solenoid magnet. These subsystems are divided into a barrel and two endcaps covering the central and forward regions. Figure 3.7 shows a cross-sectional view of the ID while a track moves through the barrel region of each of the three sub-detectors.

Silicon Trackers

Silicon trackers operate on the principle that when charged particles pass through the doped silicon sensors, they produce electron-hole pairs (by setting valence electrons free) that can be collected using an applied electric field. The charge is stored within the sensor, allowing for the precise determination of the particle's position. The silicon tracker covers the range of $|\eta| < 2.5$.

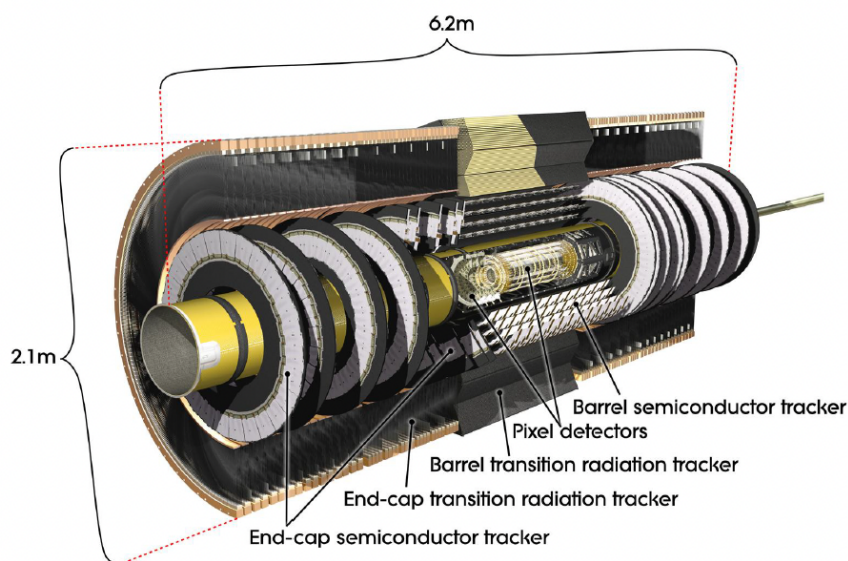


Figure 3.5: Schematic view of the ATLAS inner detector in Run I [6]

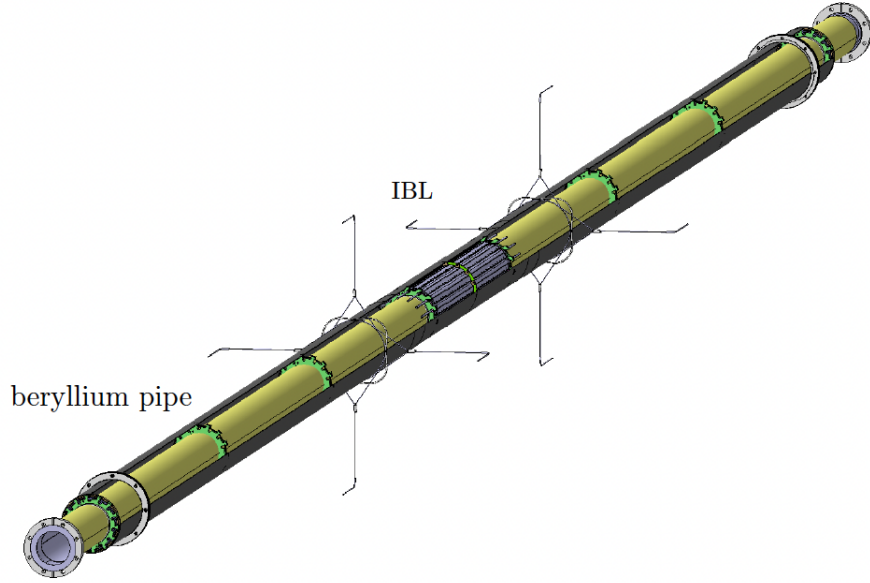


Figure 3.6: Schematic view of the insertable b-layer (IBL).

- The Insertable b-layer (IBL), depicted in Figure 3.6, is the detector element closest to the beampipe (around 3cm from the beam). Installed between 2013 and 2015, this addition introduces an extra 12 million readout channels. Consequently, it enhances the precision of vertex reconstruction, especially for decay vertices of heavy hadrons.
- The Pixel Detector consists of three barrel layers and three endcap layers of semiconductor material. The device comprises over 80 million discrete readout channels, each measuring $50 \times 400 \mu\text{m}^2$. They provide measurements of vertices with the finest granularity, a position resolution of $10 \times 115 \mu\text{m}$ across the entire pseudorapidity range.
- The SemiConductor Tracker (SCT) encloses the Pixel detectors, which consist of another four-barrel layer and nine endcap layers. Each SCT layer consists of a double layer of silicon strips, which are askew by 40 mrad with respect to each other to allow two-dimensional readout. Each double layer measures a discrete space point. SCT consists of a total of 6.3 million readout channels. The measurements at each SCT layer precisely determine the position of charged particles in the $r - \phi$ direction with an accuracy of $17 \mu\text{m} \times 580 \mu\text{m}$.

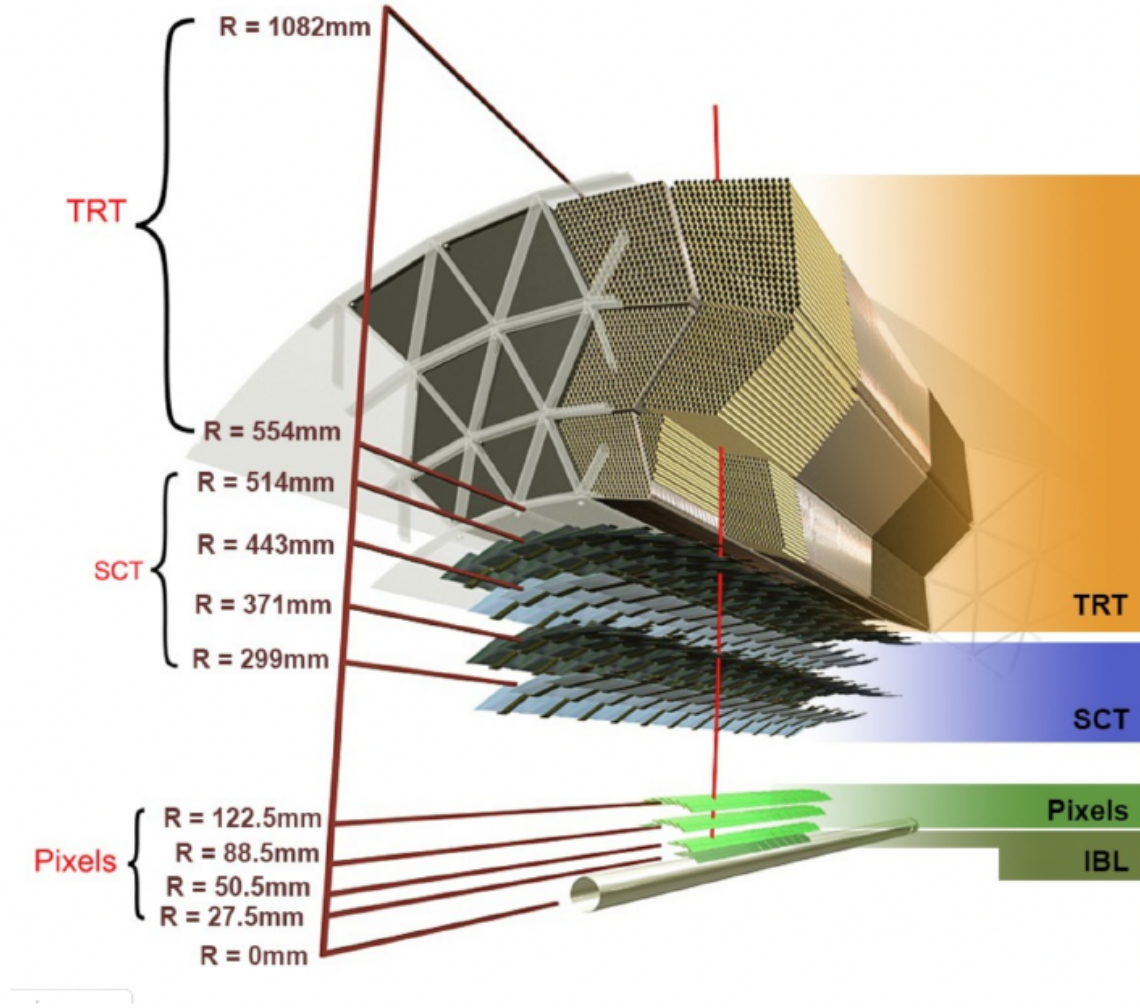


Figure 3.7: Schematic view of the structure of the inner detector barrel. [6]

Transtion Radiation Tracker (TRT)

The TRT is the outermost and the largest layer of the ID, and it operates based on the principle that when a charged particle passes between two media with different dielectric constants, it emits electromagnetic radiation. This radiation causes the gas mixture inside the straw to become ionized. An applied electric field propels the freed electrons toward the wire at the center of the straw, where they are registered.

The TRT comprises around 300,000 straw drift tubes that are 4 mm in diameter and filled with a Xenon-based gas mixture. The straw tubes are interleaved with material in which electrons create transition radiation, and the gas mixture is xenon-based to detect transition radiation efficiently. TRT provides position measurements with an accuracy of $130\text{ }\mu\text{m}$ per straw, only in $R - \phi$ plane and a pseudorapidity coverage of $|\eta| < 2.0$. The intensity of the transition radiation emitted by a charged particle can be used to determine its Lorentz factor. This can be used to measure the particle's mass, which can be used in

particle identification.

The discrete space points, as measured by the pixel detector and SCT combined with the numerous consecutive $R - \phi$ measurements by the TRT, offer extensive data for pattern recognition to identify and precisely reconstruct the tracks of charged particles.

3.2.2 Calorimeter

The calorimeter system shown in Figure 3.8 is designed to fully absorb electromagnetic and hadronic showers to measure their energy, utilizing various technologies. Most particles enter the calorimeter and initiate particle showers. Every secondary particle releases energy and generates further particles until all the energy is absorbed. The shower's form and dimension depend on the energy and the type of the primary particle. It has a coverage extending to $|\eta| < 4.9$ and prevents particles from reaching the muon system. The ATLAS calorimeter is a sampling calorimeter, meaning there are alternating active and passive layers (absorbers), measuring only a fraction of the energy of a particle by the active layers. The total energy is then deduced after careful calibration of the detector.

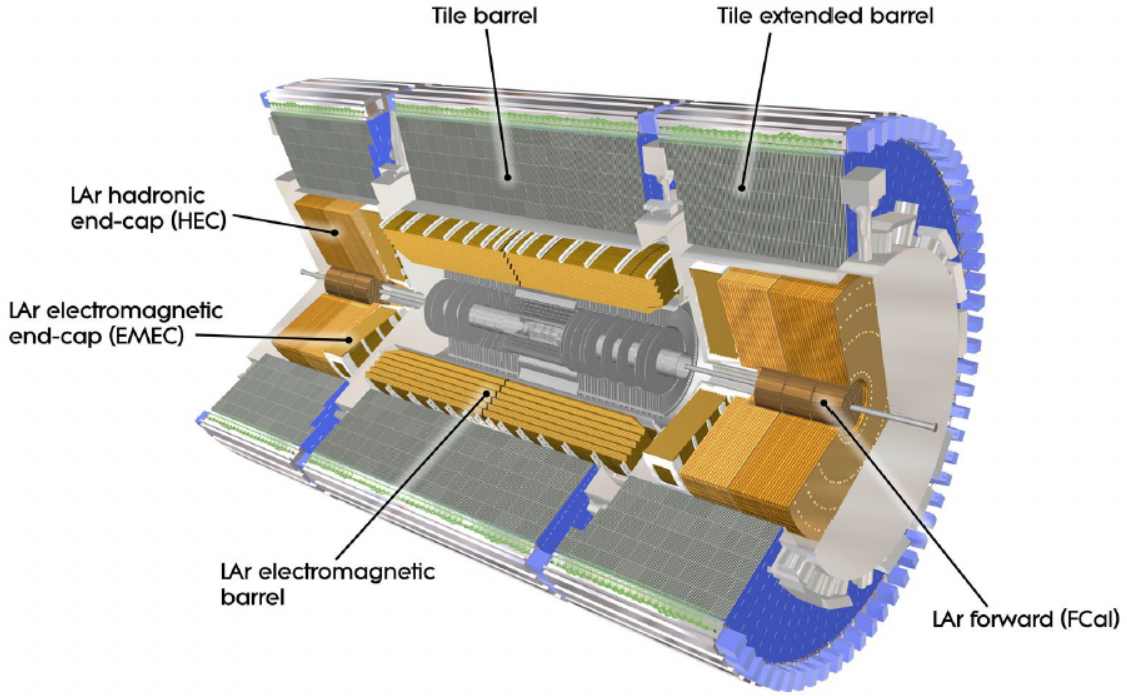


Figure 3.8: Schematic view of Calorimeter system of ATLAS. [7]

The ATLAS calorimeter comprises the electromagnetic (ECAL) and the hadronic calorimeter (HCAL). A cut-away view of the calorimeter is shown in the Figure 3.9.

- The ECAL is a highly granular inner layer that accurately measures the energy of electrons and photons. The ECAL uses liquid-argon (LAR) as an active layer and lead absorber plates as a passive layer in an accordion-like geometry. It provides complete and uniform ϕ coverage and fast readouts. The ECAL's barrel extends to $|\eta| < 1.475$, and the endcap's inner region and outer wheel extend to $|\eta| < 2.5$ and $|\eta| < 3.2$, respectively. The ECAL consists of around 175,000 readout channels. The electromagnetic calorimeter consists of three designated layers. The first layer, the presampler, has a very high granularity, distinguishing between photons and neutral pions and between electrons and charged pions. The second layer is much coarser and is used to contain the bulk of the energy deposition. The third layer also has a coarse resolution and is used to assess the majority of the energy measurement.

Figure 3.9: Layout of an ECAL barrel module. The three distinct layers can be seen, each containing a different degree of granularity, which is specified. The $\delta\phi$ values are given in rad.

- The HCAL around the ECAL is responsible for effectively stopping hadronic showers and coarsely measuring their energy so that the total energy-momentum balance of the collision can be calculated. The hadronic calorimeter comprises three sections: a tile calorimeter, hadronic endcap calorimeters, and forward calorimeters. The tile calorimeter uses scintillating plastic tiles as the active layer and steel as the absorber. The hadronic endcap shares the LAr supply with ECAL for the active layer and uses copper as the passive layer. Here, Tile and Hadronic endcap calorimeters are designed to measure the energy of hadrons. Meanwhile, the forward calorimeter provides both electromagnetic and hadronic measurements. The Forward Calorimeter shares the LAr supply similar to the hadronic end-cap but uses tungsten or copper as the passive layers.

3.2.3 The Muon Spectrometer

The Muon Spectrometer (MS) envelops the calorimeter and occupies the largest share of volume. The MS detects and gives precise muon momentum as they pass the detector. The configuration of the MS is shown in Figure 3.10. The MS relies on its large superconducting toroidal magnet to bend the muons. The magnets are cooled with liquid helium, providing 0.5 to 2T magnetic field strength.

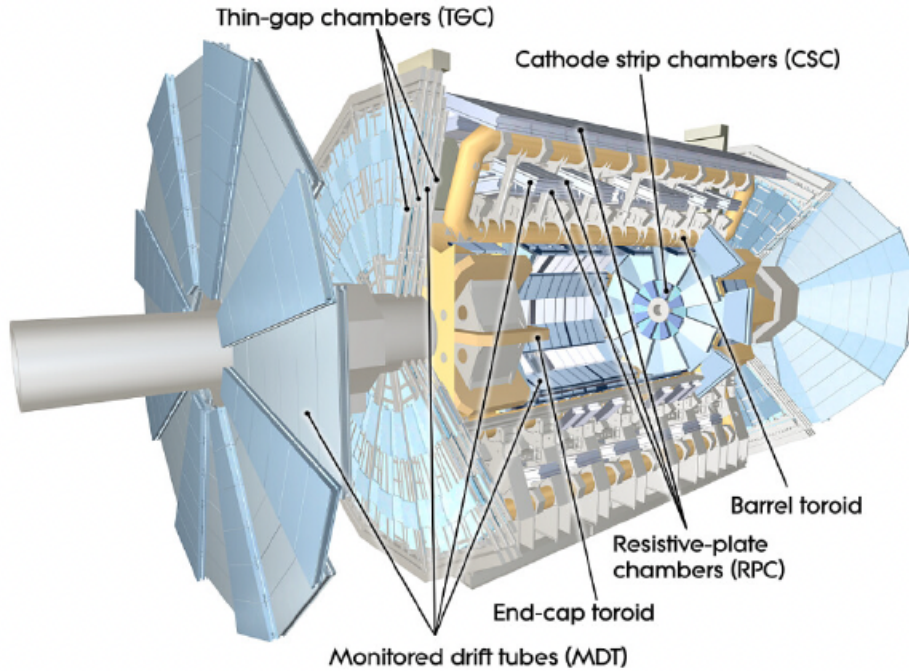


Figure 3.10: Cut-away view of the muon system.[8]

The muon chambers are structured in an octagonal symmetry. There are four distinct types of muon chamber systems, each operating on the principle that when a charged

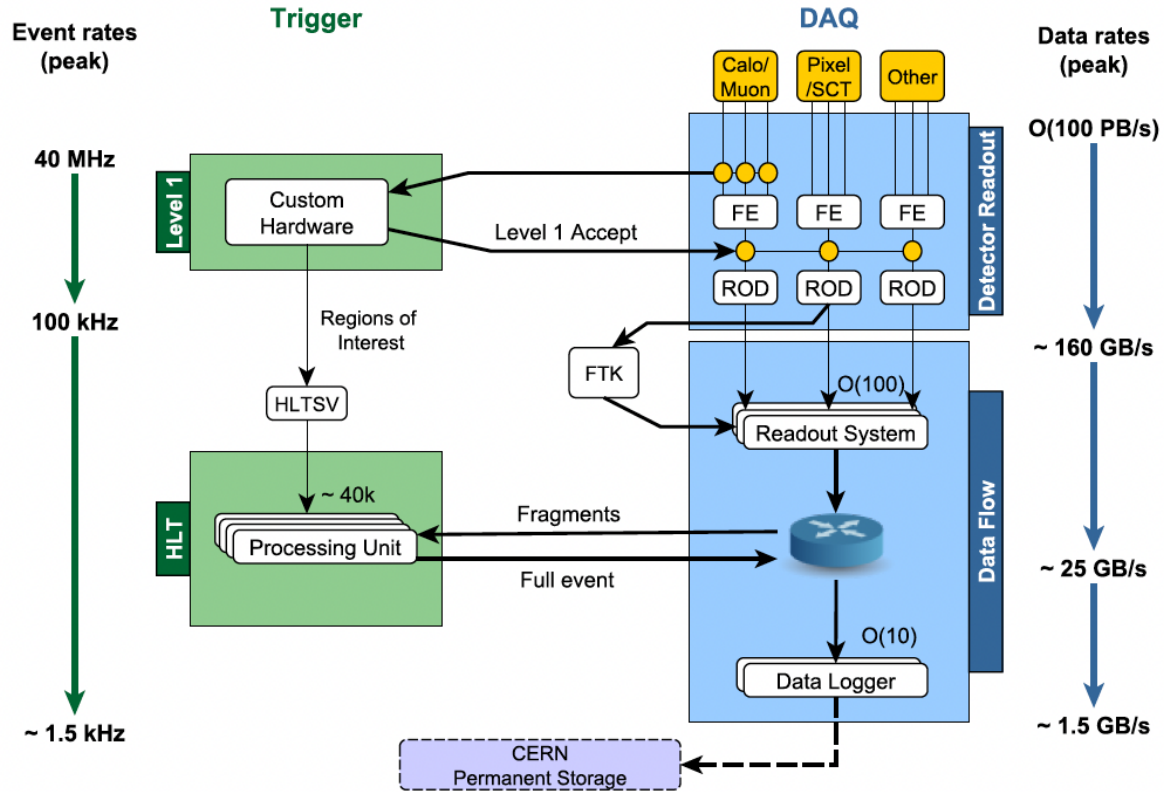
particle passes through, it ionizes the gas, which can then drift charges to oppositely charged electrodes, where they can be measured. Muon tracks are reconstructed with an absolute value of pseudorapidity (η) between 1.6 and 2.7. In the region where the absolute value of η is between 1.4 and 1.6, the bending is achieved through a combination of the barrel and end-cap fields. Eight layers of Monitored Drift Tubes (MDT) are used to measure the positions of muons across most of the η -range with an average resolution of $80\ \mu\text{m}$. Cathode Strip Chambers (CSC) are utilized within the $2 < |\eta| < 2.7$ range. The muon trigger chambers span the pseudorapidity range $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are utilized in the central section, while Thin Gap Chambers (TGC) are employed in the outer regions for fast triggering.

3.2.4 Trigger

As previously stated, the LHC operates at a bunch crossing frequency of 40 MHz. A significant number of proton-proton interactions can occur during each bunch crossing. The production rate of pp collisions in the ATLAS interaction region greatly exceeds the rate at which they can be recorded. This is due to a lack of disk storage spaces for all the events and the delay caused by the limited speed at which the detector can read and process information. In order to tackle this problem, the ATLAS trigger system is designed using an optimized network of hardware and software. Trigger systems efficiently and promptly determine which events will be recorded.

The ATLAS trigger system achieves data filtering in two stages. The first consists of a hardware-based Level-1 (L1) trigger, which reduces the data flow from 40 MHz to 100 kHz. The second stage is a software-based High-Level Trigger (HLT), which reduces the L1 output rate to 1 kHz.

- The L1 trigger consists of fast, tailor-made electronic devices that determine regions of interest (RoI) based on signals from the calorimeter and trigger chambers in the muon systems. The L1 trigger utilizes integrated circuits to execute event selections based on complex quantities, such as missing transverse energy.
- The HLT is a software-based layer implemented in an onsite computer farm with 40k processing units. Using the RoIs provided by L1 and the complete readout from the detector, specific algorithms are executed in a series of steps to extract features and carry out high-quality and full event reconstructions.



HLT: High Level Trigger – FTK: Fast Tracker – DAQ: Data Acquisition
 ROD: Read-Out Device – HLTSV: HLT Supervisor – FE: Front End

Figure 3.11: ATLAS trigger system and Data Acquisition (DAQ) system.[9]

After any trigger selects the events, a data acquisition (DAQ) system manages the data flow from the detector to the disk, which is written to the disk and reconstructed. Figure 3.11 shows a diagram of trigger operations and data flow during the LHC Run 2.

3.2.5 Event Reconstruction

Particle reconstruction, or simply reconstruction, is a process of converting the basic signals recorded by the detector into sets of measurements associated with particles generated in the collision. The ATLAS reconstruction software uses various algorithms to identify and separate energy depositions and tracks from individual particles or particle jets captured by the detector. The result of a reconstruction algorithm is a set of derived measurements corresponding to a specific particle's characteristics, which later can be used in physics analyses. Multiple layers of reconstruction are employed, where the output of one reconstruction algorithm is frequently utilized as the input for another.

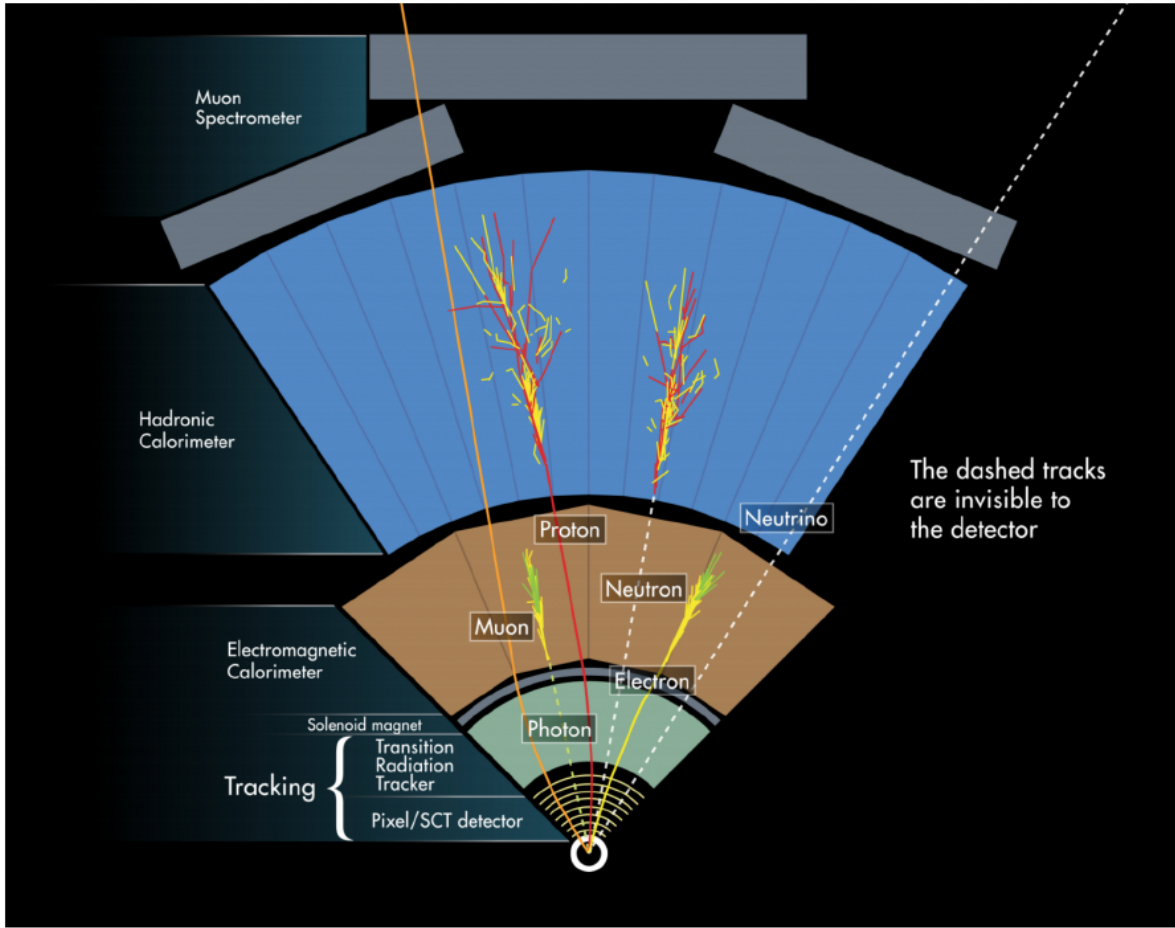


Figure 3.12: Schematic view of the different signatures of particles traversing the detector.[10]

- **Track Reconstruction:** The initial stage of particle reconstruction is tracking. It involves identifying the trajectories of individual charged particles produced from the interaction point, resulting in the reconstruction of “tracks.” Particle’s transverse momenta can then be calculated by looking into the radius of their curvature and the applied magnetic field. Track finding is performed using measurements made by ID and MS as a charged particle passing through these detectors would deposit energy along their paths. Reconstructed tracks are critical for many aspects of particle reconstruction and identification [72, 73].
- **Vertexing:** A primary vertex is an interaction point where two beam protons collide. The reconstruction of vertices involves the alignment of intersecting track groups. As the average number of interactions per bunch crossing (pile-up) increases, identifying vertices becomes more difficult. So, new algorithms are being developed [74, 75].

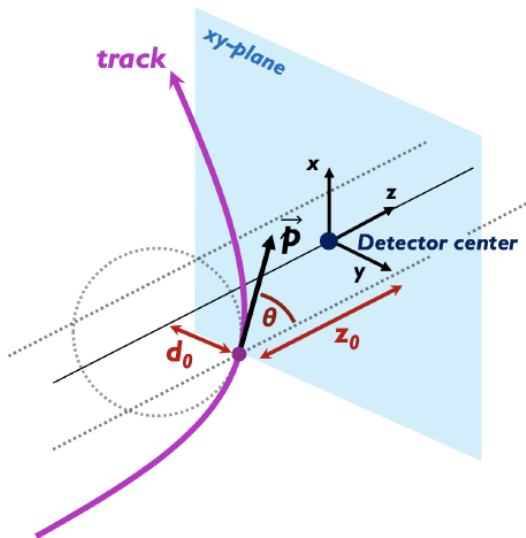


Figure 3.13: Diagram of a charged-particle track and the associated parameters describing its trajectory [11].

- Calorimeter clusters:** These are the other basic inputs of particle reconstruction. When interacting particles pass through the calorimeter, they create showers of particles. The energy of these particles is then measured across multiple calorimeter cells. Cluster reconstruction is a process that links together nearby cells and assigns them to individual particles, allowing for the measurement of the energy of the original particle. Electromagnetic particles like electrons and photons tend to generate concentrated and compact showers, primarily confined within the electromagnetic (EM) calorimeter. Particles interacting through strong forces, such as pions and kaons, will generate wide showers that can penetrate deeply into the hadronic calorimeter. The algorithms can determine whether the clusters are primarily electromagnetic or hadronic by considering their location and energy density.
- Electrons:** Electron reconstruction is relatively clean and involves matching reconstructed energy clusters found in the ECAL to tracks reconstructed in the ID. An electron leaves a track in the ID and initiates an electromagnetic shower in the ECAL, and almost all of the electron's energy is absorbed before reaching the hadronic calorimeter. However, this signature suffers from large backgrounds from other charged particles, which algorithms can handle. Tracks with a loose correlation with an electromagnetic (EM) cluster are refitted using the Gaussian Sum Filter (GSF) algorithm [76]. This algorithm considers the non-linear effects of bremsstrahlung radiation and significantly enhances the accuracy of their transverse momentum (p_T) and impact parameter measurements. Several working points (WPs) are established for electron identification to accommodate the requirements of various analyses. A likelihood-based method, which considers various discriminating variables, is employed to achieve this objective. This method offers three levels of identification,

each corresponding to different efficiencies in identifying particles with increasing background rejection [77].

- **LooseLH** corresponds to 96% identification efficiency,
- **MediumLH** corresponds to 94% identification efficiency, and
- **TightLH** corresponds to 88% identification efficiency.

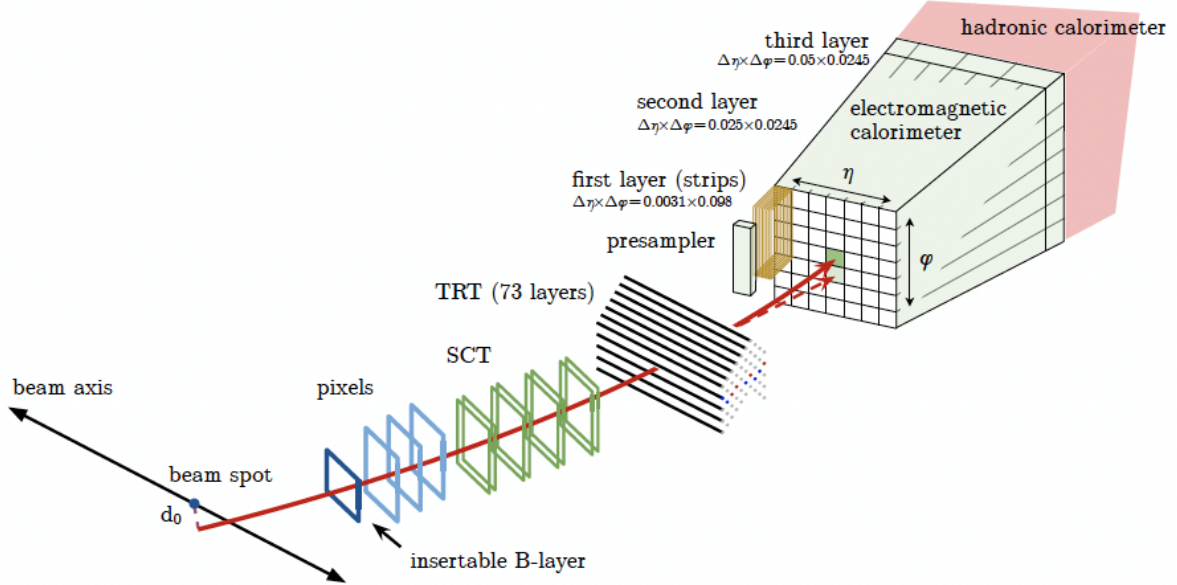


Figure 3.14: Illustration of the path of an electron [12].

- **Muons:** Muons are one of the simplest particles to identify. The tracks in the ID are reconstructed to match those in the MS. Muons also deposit minimal energy in the ECAL and HCAL. Since the calorimeters absorb the other interacting particles before reaching the MS, muons are distinguished solely based on the fact that they made it to the MS. Muons are identified by a set of criteria designed to reject secondary muons originating from pion and kaon decays [78]. There are four options available for selecting muons:

- The Loose identification criteria optimize muon reconstruction efficiency while guaranteeing reliable muon tracks.
- The Medium criteria, which is the default selection for muons in ATLAS, reduce systematic uncertainties related to muon reconstruction and calibration.
- The Tight criteria reduce the occurrence of fake muon rate in order to maximize the sample purity.
- The high- p_T selection optimizes the accuracy of momentum measurement for tracks with a transverse momentum exceeding 100 GeV.

- **Lepton Isolation:** Leptons generated through the decay of W and Z bosons tend to possess high momentum, typically exceeding 15 GeV. These muons are also produced in isolation, with minimal surrounding detector activity. When identifying isolated leptons, the energy of the reconstructed tracks and clusters near the reconstructed lepton is commonly required not to exceed a certain threshold.

This isolation requirement effectively suppresses muons generated from background processes. These criteria can include isolating tracks by setting a cutoff on the total scalar momentum of all particles originating from the same primary vertex within a specific cone size around the lepton candidate, or isolating can be done based on calorimeter measurements by setting a threshold on the sum of energy depositions around the cluster associated with the lepton candidate. The different working points for the lepton isolation are:

- Loose and Tight isolation provide isolation efficiencies of 99% and 95%, respectively for all leptons independent of their p_T .
 - Gradient and GradientLoose isolation provide isolation efficiencies linear in lepton p_T .
 - LooseTrackOnly isolation has an efficiency of 99% for all leptons independent of their p_T and applies only track isolation criteria.
- **Jets:** Jet reconstruction is another essential component for event reconstruction. Strongly interacting particles with high energy, such as quarks and gluons, produced as final state particles in a pp collision undergo a process called parton showering and hadronization. During hadronization, these particles transform into sprays of particles with a common orientation, known as a jet. These jets can be identified as a collection of energy deposits in the calorimeter and corresponding tracks in the ID.

The jets used for the analysis presented in this thesis are reconstructed employing the anti- k_t algorithm [79] from topological calorimeter clusters. The identification of a high transverse momentum jet suggests the existence of a quark or gluon in the final state. The four momenta of the resulting jets are determined by summing the four momenta of their associated constituents, which are assumed massless. The jet energies undergo calibration and correction. The jet energy scale and resolution uncertainties are often the largest source of experimental uncertainty [80].

As the multiplicity of pp collision per bunch crossing rises, pile-up poses a specific difficulty in reconstructing jets from the hard scatter. Specialized algorithms are designed to detect pile-up jets. The jet vertex tagger (JVT) is a multi-variable likelihood method to suppress jets from pile-up events [81].

- **Heavy Flavor Tagging:** The jet reconstruction algorithms typically cannot identify the specific parton responsible for initiating a given jet. The jets that are initiated by b -quarks are an exception. These b -jets can be identified from the precise vertexing of the tracks in the jet. Hadrons flavored with bottom quarks have

a relatively long lifespan, typically on the scale of millimeters. Jets produced by b -quarks, known as “ b -jets,” can be distinguished by analyzing the reconstructed tracks associated with the jet, i.e., calculating the distance of the decay vertex from the origin of the jet. The process, known as “ b -tagging,” categorizes jets as b -jets if they exhibit multiple tracks consistent with originating from a particle with a long lifespan. Highly advanced multivariate techniques are utilized to carry out the b -tagging process, offering various working points. The analysis in this thesis uses the MV2c10 algorithm, which employs a high-level boosted decision tree (BDT) discriminant. This discriminant is based on several lower-level taggers that utilize essential quantities such as track impact parameters or secondary vertices [82]. An example of the BDT output from the MV2c10 algorithm is shown in Figure 3.15.

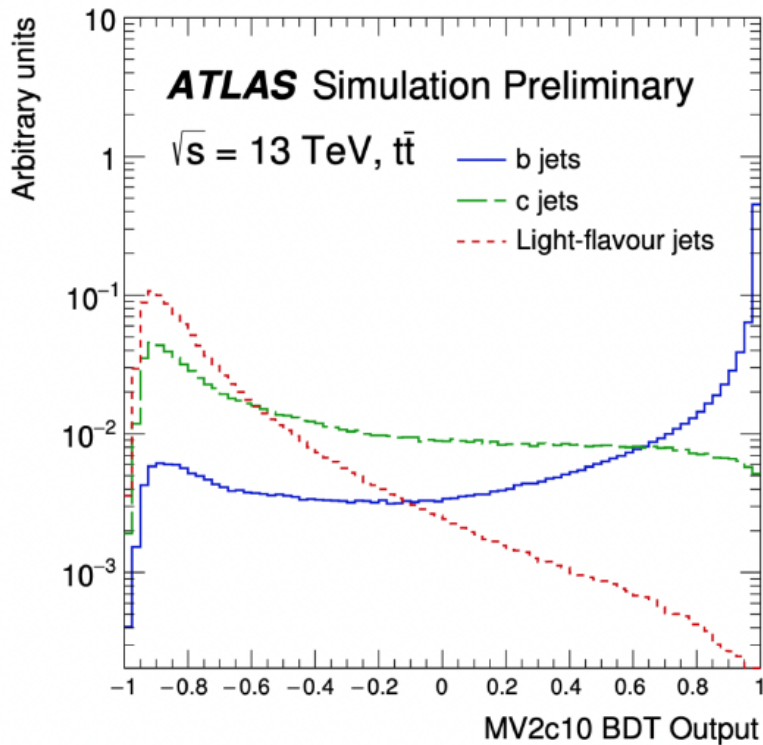


Figure 3.15: MV2c10 BDT output for different jet flavors evaluated with $t\bar{t}$ events.

- **Missing Transverse Momentum:** Particles that only interact via the weak force, such as neutrinos, can be detected by observing an overall transverse momentum imbalance. Except for neutrinos, all SM particles generated in a given interaction will either release their energy in the calorimeters or be detected by the MS.

Since the initial transverse momentum of the colliding protons should be zero, any observed momentum imbalance suggests the existence of non-interacting particles, i.e., neutrinos. This imbalance is known as the “missing transverse energy” or E_T^{miss} . There are multiple methods to determine the E_T^{miss} . The simplest iteration of E_T^{miss} is

determined by adding the transverse momentum (p_T) values of all the reconstructed calorimeter clusters and any reconstructed muons.

However, due to the imperfect resolution of all detectable and reconstructed physics objects and the high pile-up severely affects the E_T^{miss} resolution. When adding up multiple clusters, the inherent accuracy and the non-Gaussian distribution of the detector’s response result in a significant uncertainty in the reconstructed E_T^{miss} . A more accurate estimation of E_T^{miss} can be achieved by aggregating the transverse momenta of objects at higher levels. Energy measurements for identified leptons, photons, and jets are enhanced through specialized calibrations tailored to each particle type [83, 84].

3.2.6 Detector Simulation

The particles emitted from the primary collision point in a real data event will interact with the detector, releasing their energy and generating additional particles during their journey. Hence, it is important to simulate these effects using Monte Carlo. Monte Carlo samples are produced utilizing the worldwide LHC grid to evaluate the consistency between the measured data and theoretical predictions. The detector response to the Monte-Carlo-generated events must be accurately simulated using specialized detector simulation software. This software considers the particle measurement and identification inefficiencies caused by the detector’s geometry and material. The simulation data can be processed using the same reconstruction software as the real experimental data to obtain comparable results. Numerous Monte Carlo generators are available, each specializing in specific types of physics processes. After their generation, these events undergo a series of processing steps considering factors such as initial and final state radiation, hadronization, and showering.

Using the model of the ATLAS detector, the toolbox GEANT4 (GEometry ANd Tracking) [85] simulates the interaction of the particles produced in the event with the detector material. Atlfast-II (or AFII) simulation could be used to make multiple approximations to speed up the computation.

Moreover, data events will typically consist of multiple pp collisions. These overlapping signals from different pp are challenging issues for reconstruction or pile-up. The particles generated in these additional pile-up events hinder the reconstruction of the primary event of interest. This has a large effect on measured calorimeter energies, lepton isolation energies, and measurement of jet energies. Monte Carlo simulations need to be produced to address the additional pile-up. After tuning to data measurements, PYTHIA8 simulates the additional proton-proton interactions by employing soft QCD processes. Each pp interaction is simulated independently and subsequently merged before being digitized.

The detector’s response (“digits”) is ultimately obtained by converting these simulated interactions by dedicated digitization software. Following this, the reconstruction of simulated events is carried out in the same manner as the reconstruction of actual data events. The modeling of detector noise involves measuring the rates in the data for a

specific readout technology and then storing the average noise level in a database.

3.3. The Worldwide LHC Computing Grid

Since the computational resources must be globally accessible, as most LHC researchers are not physically present at CERN, a distributed computing infrastructure was built, referred to as the Worldwide LHC Computing Grid (WLCG or Grid). This system facilitates the computational capabilities for all the LHC experiments. It offers a range of services to process and store their data and generate Monte Carlo (MC) samples. The Grid is organized in a multi-tiered system, wherein the tiers are categorized based on their computational capabilities, storage capacity, and internet connectivity speed. Tier-0 (T0) is responsible for recording and saving all the raw data from the LHC. T0 also performs a first reconstruction of hits and objects, which are then passed down to Tier-1 (T1). T1 are the national centers (13 countries) where a portion of RAW and object data is stored, and large-scale reprocessing is performed and distributed to Tier-2. T2 are regional groups (laboratories and universities) that provide capacities for MC simulation, data analysis, and production of data in AODs. The University of Oklahoma is a T2 center.

3.4. Highlights of Higgs bosons measurements

On July 4, 2012, the ATLAS and CMS collaborations independently announced the detection of a signal in the data above the background noise having a 5σ significance excess at approximately 125 GeV. This excess was observed while analyzing the invariant mass spectrum of the diphoton decay mode using the Run I data, as shown in Figure 3.16. The bump suggests the existence of a novel particle that decays into two photons.

This graph provides additional insights into the newly discovered particle. According to the charge conservation principle, this newly discovered particle must have no charge. Additionally, since it decays into two massless bosons (photons), the particle must be a boson, and its spin cannot be one, restricting the possibilities to spin 0 or 2. The inclusion of additional data and complex analyses of the angular correlations among the decay products indicated that the newly discovered particle is a scalar boson.

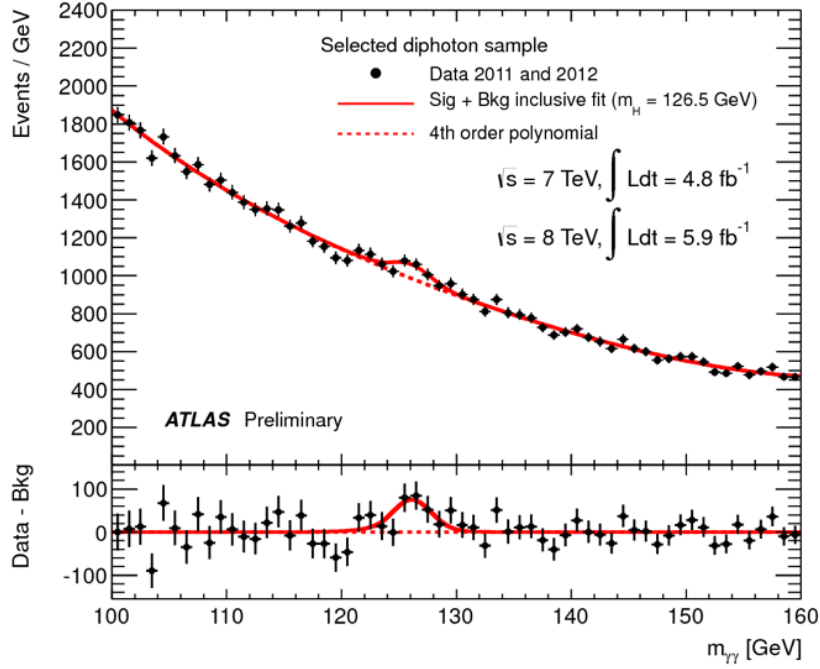


Figure 3.16: The invariant mass spectrum from the diphoton channel, as presented on July 4, 2012. The data points show the measured distribution, with each event weighted by the signal-to-background ratio for each sub-category of events. The solid red line represents the result of fitting the signal combined with the background, while the dashed red line represents the background.

The validity of the SM Higgs hypothesis can be examined by analyzing the Higgs coupling to fermions. As mentioned in the earlier chapter, the Yukawa term in the SM Lagrangian establishes a linear relationship between the coupling strengths of fermions and their masses, while the W and Z bosons with the square of their masses. Using the k -framework, we can quantify the coupling strength of the Higgs boson to individual particles. A comparison of the SM prediction and the fit of the measurement of Higgs coupling with other SM particles is shown in Figure 3.17. It clearly shows that the strength of the interaction depends upon the mass of the particle: the greater the mass, the stronger its interaction with the Higgs field. This suggests the Higgs boson observed in 2012 is indeed responsible for the BEH mechanism of electroweak symmetry breaking.

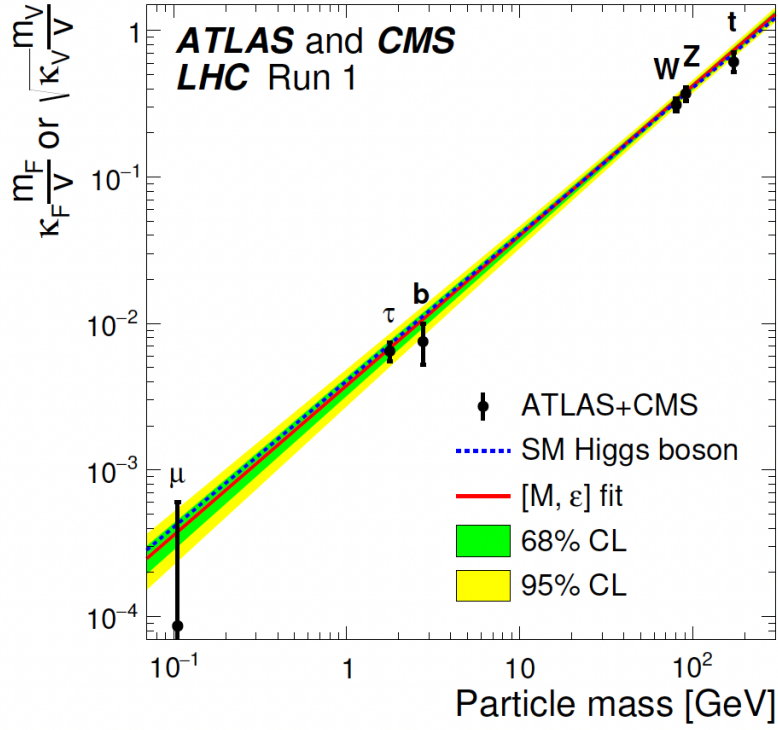


Figure 3.17: The measured coupling strength as a function of the mass of different SM particles [13].

Due to its heavy mass and small production cross-section, sophisticated analysis techniques and powerful accelerators were required to produce and study the Higgs boson. With its ability to collide protons at high center-of-mass energy, the LHC has collected substantial evidence that strengthens the Standard Model's fundamental principles. However, the data taken in 2015 and 2016 after the upgrade provide an even more precise analysis of the SM Higgs mechanisms and the search for effects beyond the SM in the Higgs sector, which constitutes the main focus of the remaining part of the thesis.

4 $H \rightarrow WW^* \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ Physics

This chapter offers a comprehensive overview of $H \rightarrow WW^*$ physics. The rationale for utilizing the WW final state is outlined, and the analysis strategy for this decay mode, based on $\sqrt{s} = 13$ TeV with the Run II data, is presented. The data, Monte Carlo samples, and object identification and selections used in the analysis are described in sections 4.3 and 4.4. Moreover, the chapter discusses the background processes associated with the analysis and how they are estimated in section 4.5, along with a general summary of the different data-driven methods and data analysis techniques introduced in section 4.6.

4.1. Introduction

As discussed in Section 2.4.2, the $H \rightarrow WW^*$ decay channel has the second highest branching ratio at $m_H = 125 \pm 0.24$ GeV [86]. The only decay channel with a higher branching fraction is the decay to a pair of bottom quarks. However, identifying events of this nature ($H \rightarrow bb$) is exceptionally challenging due to the exclusively hadronic outcome, which has extremely large background processes. Due to this multijet background, events arising from the dominant ggF production mode are difficult to distinguish. This constraint does not apply to the events that decay through the $H \rightarrow WW^*$ channel [87].

Due to the extremely short lifespan, only the decay products of the W bosons can be measured. These decays can be classified as hadronic (final states containing quarks) and leptonic (final states containing a lepton and neutrino pair). Each W boson decays to leptons with a total branching ratio of 32.59 % or to hadrons with a branching ratio of 67.41% [14]. However, the signal extraction for fully hadronic decays is challenging due to the presence of QCD background, despite their higher occurrence. On the other hand, the fully leptonic final states have a cleaner signature for the trigger because many of the physics processes leading to the leptonic final states typically involve electroweak interactions, which have a smaller cross-section than the QCD processes.

The signature comprises two isolated, high p_T leptons with opposite charges and a sizable amount of missing transverse energy, E_T^{miss} , originating from the neutrinos. Therefore, the $H \rightarrow WW^* \rightarrow l\nu l'\nu$ final state is chosen for this study. Figure 4.1 shows the branching ratio of all relevant decay modes of $H \rightarrow WW^*$.

Of the charged leptons, electrons are stable, and muons are stable enough to be directly measured by the detector. However, tau leptons undergo further decay. The decay modes with the branching ratio of tau particles are illustrated in Figure 4.2. The total branching ratio of decays to lighter leptons is approximately 35.2%, with the remainder being hadronic decays [14]. Therefore, in the context of this analysis, “leptons” (l) refer only to light leptons, electrons (e), or muons (m).

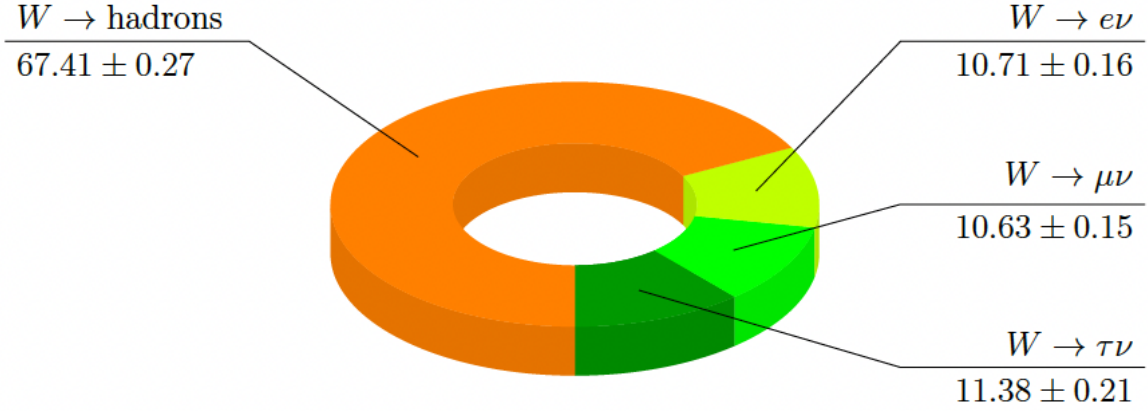


Figure 4.1: Experimentally measured decay branching fractions of W bosons [14].

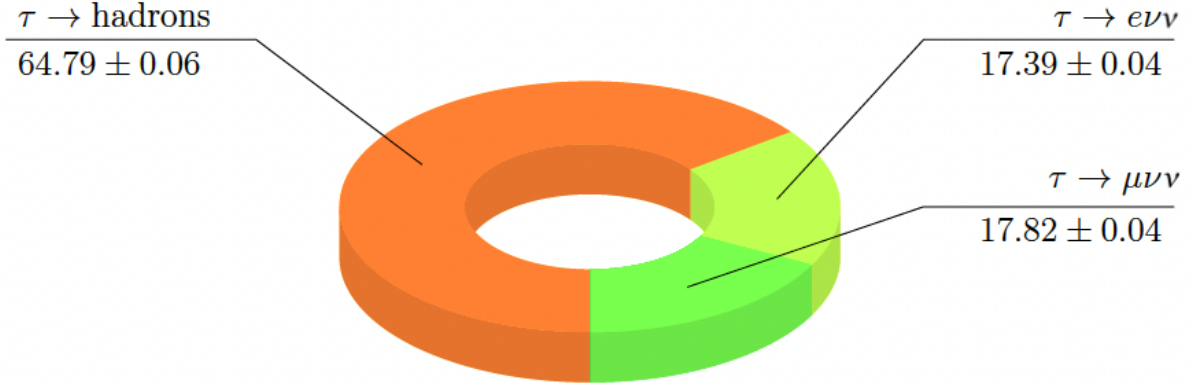


Figure 4.2: Experimentally measured decay branching fractions of τ -leptons.

In analyses involving two leptons in the final state, it is common practice to label the leptons based on their transverse momenta. The lepton with the highest transverse momentum is referred to as the “leading” lepton (l_1), while the lepton with the second highest transverse momentum is referred to as the “subleading” lepton (l_2). Hence, there are four lepton flavor configurations, namely, two same-flavor (SF) configurations (ee and $\mu\mu$) and two different-flavor (DF) configurations ($e\mu$ and μe). Events featuring two leptons of the SF are more likely through a pair production involving a Z/γ^* boson (“Drell-Yan” (DY) process) [88]. The DY process has a cross-section larger than that of the Higgs boson signal. Hence, in previous studies, the final state leptons with different flavors were chosen as the events featuring the DF offer a significantly higher signal sensitivity. A thorough analysis of the SF modes was not entirely conducted using Run II data.

Moreover, the analysis is divided into four distinct categories, depending on the number of jets produced, known as jet multiplicity (N_{jets}). These categories include events with no jets, events with exactly one jet targeting ggF production mode, and events with

at least two jets targeting both ggF and VBF production modes. This categorization is done because the ggF process generates jets solely through parton radiation from the initial state partons. In contrast, VBF production is characterized by the presence of two energetic jets that are widely separated in rapidity. Since the SM Higgs boson is a scalar boson, the spin correlations among the particles in the final state (the two charged leptons) are emitted in the same hemisphere, while the two neutrinos are cast in the opposite hemisphere, shown in Figure 4.3) can be effectively utilized in the process of event selection.

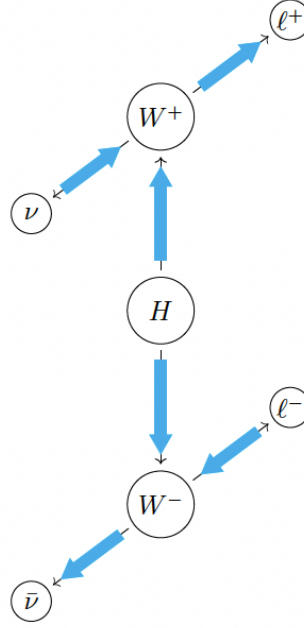


Figure 4.3: Spin correlations in the $H \rightarrow WW$ decay mode. The thin black arrows indicate the particles' directions of motion, while the thicker blue arrows indicate their spin projections.

4.2. Data and Monte Carlo Samples

4.2.1 Data Samples

The analyses utilize a data set corresponding to a total integrated luminosity of 140 fb^{-1} (36.6 fb^{-1} in 2015-2016, 44.6 fb^{-1} in 2017 and 58.8 fb^{-1} in 2018) of pp collision data taken at a center-of-mass energy of 13 TeV by the ATLAS detector. The criteria for selecting data events for this analysis, including a combination of single Lepton triggers and one $e - \mu$ di-lepton trigger and associated transverse momentum requirements, are provided in Table 4.1.

Lepton	Level-1 trigger	High-level trigger
Year 2015		
e	20 GeV	24M OR 60M OR 120L GeV
μ	15 GeV 20i OR 50 GeV	
$e\mu$	e : 15 GeV, μ : 10 GeV	e17_lhloose_mu14
Year 2016		
e	20 GeV	24 Ti OR 60M OR 140L GeV
μ	15 GeV	24 i OR 50 GeV
$e\mu$	e : 15 GeV, μ : 10 GeV	e17_lhloose_mu14
Year 2016: after D4		
e	20 GeV	26 T i OR 60M OR 140L GeV
μ	15 GeV	26 i OR 50 GeV
$e\mu$	e : 15 GeV, μ : 10 GeV	e17_lhloose_mu14

Table 4.1: Minimum p_T requirements used at the each trigger. Letters “T”, “M” and “L” denote the Tight, Medium, and Loose electron identification requirements, respectively. The letter “i” indicates an isolation requirement to improve rejection power against non-prompt leptons.

The single Lepton triggers capture most events, whereas the Dilepton trigger allows for the additional acceptance of low p_T events. An additional dilepton trigger with an even lower p_T threshold was used, which slightly improved overall trigger efficiency. This information is illustrated in Figure 4.4, which displays the trigger efficiencies for the lepton triggers with the lowest p_T thresholds used in this analysis. Since the trigger efficiency as a function of lepton p_T is described by a steep increase in trigger efficiency around the trigger E_T threshold, it is necessary to correct the simulated samples to consider the trigger efficiency accurately. For the same flavor channel, only single-lepton triggers are used in order to simplify the measurement (at a cost of $O(10\%)$ signal yield loss).

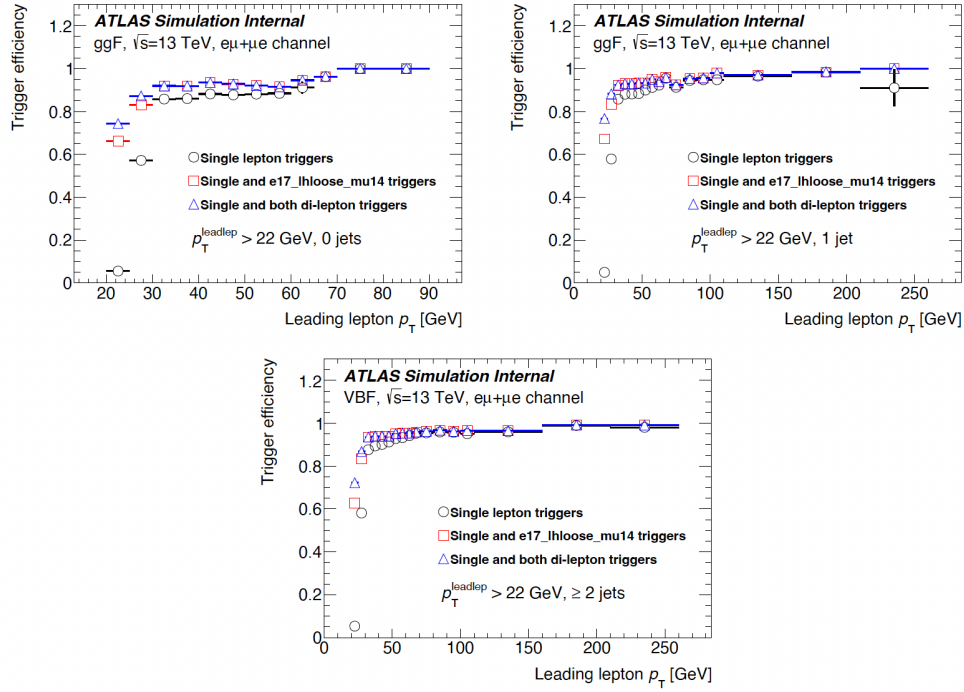


Figure 4.4: The trigger efficiencies plotted against the p_T for different configurations in the 0 jet ggF (top left), one jet ggF (top right), and VBF (bottom) analyses. These efficiencies are calculated after preselection and the condition of having a leading lepton with a p_T greater than 22 GeV.

4.2.2 Monte Carlo Samples

The theory prediction is produced through Monte Carlo (MC) samples for both the signal and background. The signal and background processes are modeled using Monte Carlo generators, documented in Table 4.2. With the exception of SHERPA [50], for most processes, separate programs are used to generate the hard scattering process and to model the parton showering (PS), hadronization, and the underlying event (UE). The signal and some of the background processes in the final three steps utilize PYTHIA8.210 [48], PYTHIA8.186 [89], or PYTHIA6.428. HERWIG [49] is employed for the PS and hadronization in assessing systematic uncertainties, while the UE is simulated using the JIMMY[90].

The CT10 PDFs are utilized for the hard scattering process in POWHEG-BOX v2 [91], whereas the NNPDF23LO PDF set is employed for MAD-GRAPH5 [92]. The AZNLO [93] tune is utilized for the diboson and signal processes, while the A14 tune is employed for other processes. The hard scattering NLO predictions, as calculated by SHERPA2.1.1 [94], utilize the CT10 PDF set along with a specifically tailored set of parameters from the parton shower developed by the Sherpa authors [95]. Pile-up is modeled using PYTHIA [96]. Every event produced undergoes a detector simulation based on GEANT4 [85], and the standard ATLAS reconstruction is applied to collision data.

The ggF production mode is simulated with NNLO accuracy in QCD using the

POWHEG-BOX v2 and normalized to a cross-section calculated with NNNLO accuracy in QCD. VBF production mode, on the other hand, is simulated using POWHEG-BOX v2, incorporating NLO accuracy in QCD and normalized to a cross-section computed with NLO accuracy in QCD, incorporating an approximate NNLO correction.

Process	Generator	$(\sigma\mathcal{B})(\text{pb})$	Precision $\sigma_{\text{incl.}}$	analysis
ggF, $H \rightarrow WW$	POWHEG + PYTHIA8	10.4	NNLO	
VBF, $H \rightarrow WW$	POWHEG + PYTHIA8	0.808	NNLO	
WH, $H \rightarrow WW$	POWHEG + PYTHIA8 (MiNLO [137])	0.293	NNLO	
ZH, $H \rightarrow WW$	POWHEG + PYTHIA8 (MiNLO [137])	0.189	NNLO	
inclusive $W \rightarrow \ell\nu$	POWHEG + PYTHIA8	6.02×10^4	NNLO	WH
inclusive $Z/\gamma^* \rightarrow \ell\ell$ ($m_{\ell\ell} \geq 40\text{GeV}$)	MADGRAPH	6.04×10^3	NNLO	
inclusive $Z/\gamma^* \rightarrow \ell\ell$ ($40 \geq m_{\ell\ell} \geq 10\text{GeV}$)	MADGRAPH	8.01×10^3	NNLO	
$(W \rightarrow \ell\nu)\gamma$ ($p_T^\gamma > 10\text{GeV}$)	SHERPA	453	LO	
$(Z \rightarrow \ell\ell)\gamma$ ($p_T^\gamma > 10\text{GeV}$)	SHERPA	175	LO	
$t\bar{t}$ leptonic	POWHEG + PYTHIA6	87.6	NNLO+NNLL	
Wt leptonic	POWHEG + PYTHIA6	7.55	NLO	
$t\bar{t}W/Z$	MADGRAPH	0.62	NLO	WH
tZ non-all-hadronic	MADGRAPH	0.24	LO	WH
$q\bar{q}/g \rightarrow WW \rightarrow \ell\nu\ell\nu$	POWHEG + PYTHIA8	11.2	NNLO	ggF, VBF
$Z^{(*)}Z^{(*)} \rightarrow 2\ell 2\nu$ ($m_{\ell\ell} \geq 4\text{GeV}$)	POWHEG + PYTHIA 8	0.925	NLO	VBF
$q\bar{q}/g \rightarrow 2\ell 2\nu$	SHERPA	12.8	NLO	WH
$gg \rightarrow 2\ell 2\nu$	SHERPA	0.72	NLO	
$q\bar{q}/g \rightarrow \ell\nu\ell\ell$	SHERPA	11.9	NLO	
$q\bar{q}/g, gg \rightarrow \ell\ell\ell\ell$	SHERPA	11.5	NLO	
$WZ \rightarrow q\ell\ell\ell$	SHERPA	3.76	NLO	WH
$ZZ \rightarrow q\ell\ell\ell$	SHERPA	2.36	NLO	WH
EW $WW + 2$ jets ($\ell\nu\ell\nu$)	SHERPA	0.012	LO	VBF
EW $WZ + 2$ jets ($\ell\nu\ell\ell$)	SHERPA	0.038	LO	WH
EW $ZZ + 2$ jets ($\ell\ell\ell\ell$)	SHERPA	0.116	LO	WH
EW $q\bar{q} \rightarrow (Z \rightarrow \tau\tau)q\bar{q}$	SHERPA	2.54	LO	VBF
$WWW(3\ell 3\nu)$	SHERPA	8.34×10^{-3}	NLO	WH
$WWZ(4\ell 2\nu, 2\ell 4\nu)$	SHERPA	5.16×10^{-3}	NLO	WH
$WZZ(5\ell 1\nu, 3\ell 3\nu)$	SHERPA	1.07×10^{-3}	NLO	WH
$ZZZ(6\ell 0\nu, 4\ell 2\nu, 2\ell 4\nu)$	SHERPA	3.16×10^{-4}	NLO	WH

Table 4.2: Overview of Monte Carlo generators employed for simulating the signal and background processes in the analysis, alongside the associated cross sections used for normalization.

Signal

The signal processes are the ggF and VBF production modes, with the Higgs boson ($m_H = 125\text{ GeV}$) decaying into two W bosons, which subsequently decay into two leptons in the final state ($H \rightarrow WW^* \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ decay mode). Other Higgs boson production mechanisms such as the Higgs-strahlung process, $t\bar{t}H$, $b\bar{b}H$, and tHx are disregarded due to their small contribution, while the presence of $H \rightarrow \tau\tau$ decays is acknowledged as background as it is fixed to SM predictions.

The two final state leptons resulting from the $H \rightarrow WW^*$ process display distinct kinematics compared to leptons generated by background processes. As briefly mentioned in Section 4.1, in the rest frame of the Higgs boson (spin-0 boson), the W bosons (spin-1 boson) resulting from its decay have opposite spins and are aligned in opposite directions to maintain the conservation of angular momentum. The leptons and neutrinos resulting from the decay of a W boson are emitted in a direction opposite to the spin of the W

boson, while anti-leptons and anti-neutrinos are emitted in the same direction as the spin of the W boson. This is because a W boson interacts exclusively with left-handed fermions (or right-handed antifermions) in the SM. As a result, a small opening angle is observed between the lepton pair (hence, a small invariant mass), whereas lepton pairs produced by background processes typically have a wider distribution of opening angles and invariant masses. Figure 4.5 shows the ATLAS's event displays of reconstructed event candidates for $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ in each analysis category.

Furthermore, a transverse mass can be calculated to differentiate the Higgs boson signal from background processes by utilizing the known transverse momentum components of the final state particles and the missing transverse momentum.

$$m_T = \sqrt{\left(E_T^{\ell\ell} + E_T^{\text{miss}}\right)^2 - \left(\mathbf{p}_T^{\ell\ell} + \mathbf{E}_T^{\text{miss}}\right)^2} \quad (4.1)$$

where $E_T^{\ell\ell} = \sqrt{m_{\ell\ell}^2 + |\mathbf{p}_T^{\ell\ell}|^2}$; $m_{\ell\ell}$ and $\mathbf{p}_T^{\ell\ell}$ refer to the invariant mass of the lepton pair and the vector sum of the lepton p_T . m_T is a lower bound on m_H . This observable is important to measure the Higgs boson signal, as numerous background processes show a distinct distribution in m_T than the Higgs boson signal.

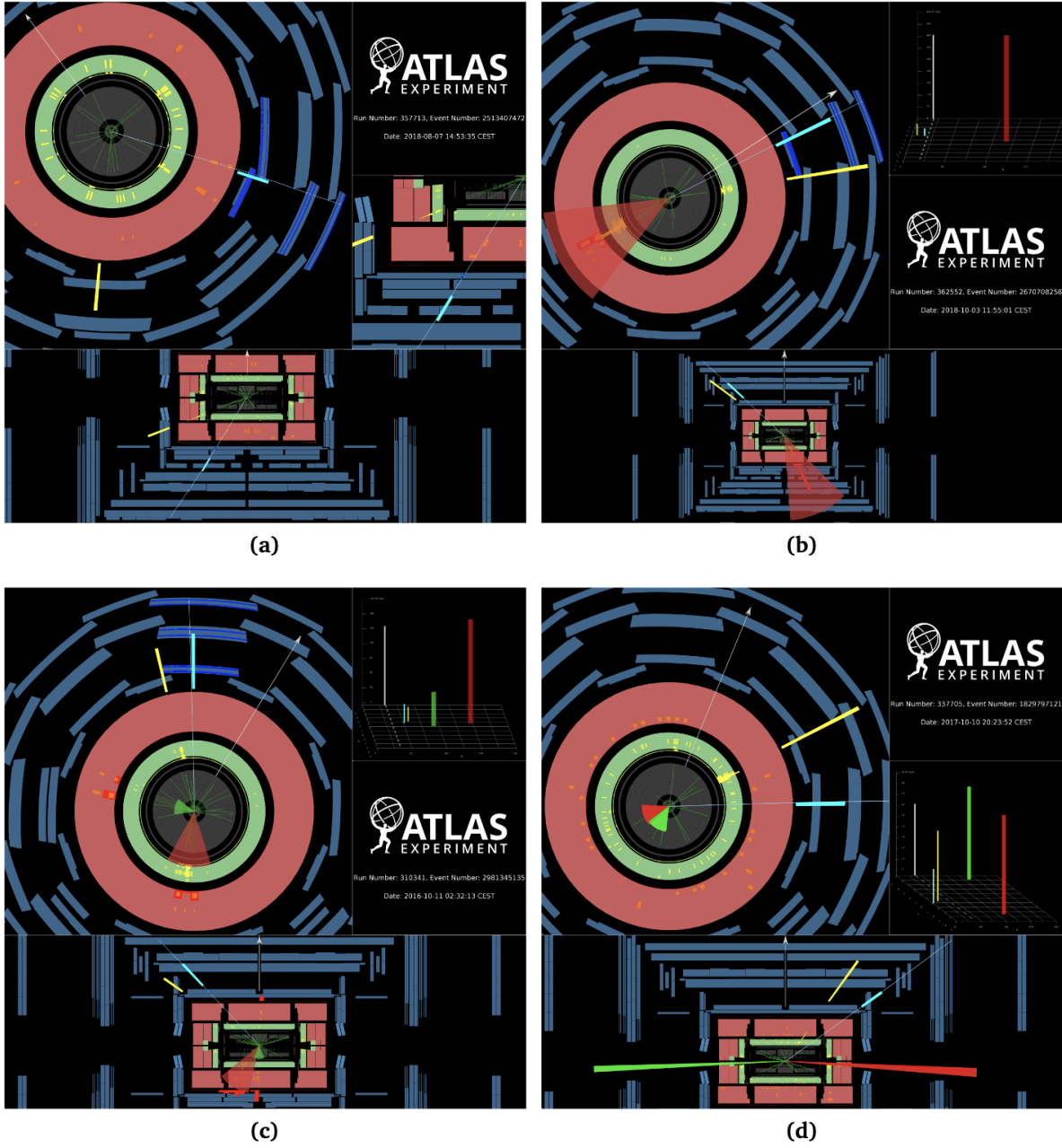


Figure 4.5: Event displays of candidate Higgs bosons from the (a) ggF $N_{\text{jet}} = 0$, (b) ggF $N_{\text{jet}} = 1$, (c) ggF-enriched $N_{\text{jet}} \geq 2$, and (d) VBF-enriched $N_{\text{jet}} \geq 2$ analysis regions. Each display illustrates a cross-sectional and longitudinal view of the ATLAS detector. The final state particles are the electron (yellow line), the muons (turquoise line), the missing transverse momentum (white arrow), and the reconstructed jets (red and green cones). Yellow and red squares indicate the presence of energy deposits in the ECAL and HCAL and the reconstructed tracks by green lines.

Background

The primary sources of the SM backgrounds in the analyses presented here consist of events arising from the production of top-quarks ($t\bar{t}$, Wt), pairs and triplets of electroweak bosons (WW or Non-resonant WW), W or Z bosons produced alongside hadronic jets, and pure QCD multijet events.

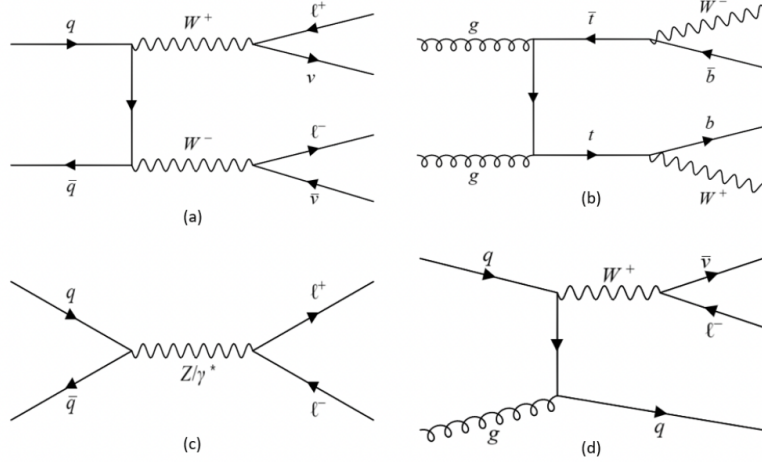


Figure 4.6: The main background processes of the HWW analyses: (a) WW , (b) $t\bar{t}$, (c) Drell-Yan (DY) and (d) non-prompt background.

- **Drell-Yan (DY) process** is the main source of background for the same-flavor categories due to its significantly high cross-section. Furthermore, a DY process can appear as a real event when a Z/γ^* boson undergoes decay into two leptons with opposite charges, and a mis-reconstruction of energies in the detector occurs results in a fake detection of E_T^{miss} . DY process is generated using SHERPA2.2.1 and is divided based on the maximum value between H_T and p_T^V . In addition, the $Z \rightarrow \tau\tau$ process applies an extra filter on the lepton (or hadron) p_T to enhance the population of the analysis phase space. The provided samples have been normalized to cross-sections calculated with NNLO accuracy.
- **Nonresonant W^+W^- production** refers to the irreducible background where a pair of W bosons are produced through the scattering of $q\bar{q}$ or $gg \rightarrow WW$. Although the final state is the same, the background is reduced because the W bosons from the signal are produced from the decay of a spin-0 particle, so their spins are antiparallel. The WW background is produced independently for the $qq \rightarrow WW$ and $gg \rightarrow WW$ production modes. The main contribution among the two is the $qq \rightarrow WW$ process, which is produced using SHERPA2.2.2, with the matrix elements being computed for up to one extra parton at NLO accuracy and up to three extra partons at LO accuracy. The SHERPA 2.1.1 simulation models the loop-induced process of $gg \rightarrow WW$, with the inclusion of zero or one additional jet and normalized to the NLO cross-section.

- **The top quark background** arises from the production of $t\bar{t}$ pairs and single top quark production. The former is generated through gluon fusion or by the scattering of $q\bar{q}$, while the latter is produced by the tW in either the s-channel or t-channel. The top quark undergoes decay with a branching ratio of 98% into a W boson and a b-quark. Therefore, the top background exhibits the same final state of the signal along with b -jets. Top-quark pair production is generated by utilizing POWHEG within the POWHEG-BOX framework. A filter is implemented, mandating that the W bosons undergo lepton decay. The samples are normalized to cross-sections calculated with NNLO+NNLL accuracy. The single top production is generated using POWHEG-BOX 2.0 and incorporates EVTGEN 1.2.0 as an additional tool to improve the modeling accuracy of bottom and charm hadron decays.
- **The non-prompt lepton background** refers to the occurrence of a single W boson being produced together with a jet, which has a high probability of being mistakenly identified as a lepton, thus potentially being misinterpreted as a signal event. The likelihood of this misidentification depends on the momentum of the jet.

Name	Physics process	Cause of $\ell\ell + E_T^{miss}(+jets)$ signature	Estimation method
Other Higgs boson (Other H)			
VH production	$pp \rightarrow VH, V = \{W, Z\}$	hadronic V: irreducible; leptonic V: lepton(s) not reconstructed	Theoretical predictions
Inclusive ttH	$pp \rightarrow ttH$	untagged b-jets	
ggF , VBF production of $H \rightarrow \tau\tau$	$H \rightarrow \tau\tau$	irreducible	
Non-resonant WW			
$qqWW$	$qq \rightarrow WW$	irreducible	Data-driven (CR)
$ggWW$	$gg \rightarrow WW$	irreducible	Theoretical prediction
Top quarks ($t\bar{t}/Wt$)			
$t\bar{t}$	$t\bar{t} \rightarrow WbWb$	untagged b-jets	Data-driven (CR)
Wt	$\{Wt, tb/tqb\}$	untagged b-jets; q or b reconstructed as ℓ	
Mis-identified leptons (Mis-Id)			
Wj	$W + \text{jet(s)}$	j reconstructed as ℓ	Data-driven (fake factor)
jj	QCD multi-jet	jj reconstructed as $\ell\ell$; fake E_T^{miss}	
Non-WW diboson and triboson			
Other VV (V)	$W\gamma$	γ reconstructed as e	Theoretical predictions
	$W\gamma^*, WZ, ZZ \rightarrow \ell\ell\ell$	lepton(s) not reconstructed	
	$ZZ \rightarrow \ell\ell\nu\nu$	irreducible	
	$Z\gamma$	γ reconstructed as e ; lepton not reconstructed	
	WWW, WWZ, ZZZ	leptons not reconstructed	
Drell-Yan (Z/γ^*)			
$\tau\tau$	$Z/\gamma^* \rightarrow \tau\tau \rightarrow \ell\nu\ell\nu$	irreducible	Data-driven (CR)
$ee/\mu\mu$	$Z/\gamma^* \rightarrow ee/\mu\mu$	fake E_T^{miss}	

Table 4.3: Summary of physics processes and causes of $\ell\ell + E_T^{miss}(+jets)$ signature and estimation methods [15].

Other backgrounds include the associated production of a photon with a vector boson ($W\gamma$, $W\gamma^*$, and $Z\gamma$) when the photon generates a pair of leptons, as well as the production of dibosons (WZ and ZZ) and tribosons (VVV) when the leptons are emitted beyond the detector's range or misidentified. The events with $W\gamma$ and $Z\gamma$ are modeled using SHERPA 2.2.2 with NLO accuracy. These simulations impose a minimum p_T requirement of 7 GeV for the γ particle and to be at a distance $\Delta R > 0.1$ from any lepton. The WZ and ZZ processes are produced using SHERPA 2.1. The $W + \text{jets}$ process is estimated using a data-driven method. However, Monte Carlo samples are still utilized to estimate the uncertainty in the sample's composition.

4.3. Object Identification and Selection

Events must possess a primary vertex that contains a minimum of two associated tracks, each with a $p_T > 400$ MeV. When multiple vertices are reconstructed in a single event, the vertex with the largest sum of square transverse track momenta is chosen. The electron, muon, τ_{had} decay, jet, and E_T^{miss} objects are determined based on the outputs of the previously described reconstruction algorithms in Section 3.2.5.

4.3.1 Electrons and Muons

To reduce the number of misidentified leptons, criteria on the isolation are required. The electrons used in the analysis are required to have a $E_T > 15$ GeV and pass the “MediumLH” or “TightLH” selection as defined in Section 3.2 for electrons with a transverse energy greater or smaller than 25 GeV, respectively. Electron candidates must be within the acceptance of the ID ($|\eta| < 2.47$), excluding the transition region $1.37 < |\eta| < 1.52$. A criterion (Author = 1) is implemented to select electrons that were firmly reconstructed as electrons, significantly diminishing the W background by nearly half while preserving a high signal efficiency.

The recently implemented improved tight prompt lepton veto PLIVTight diminishes the yields of misidentified leptons by a factor of two compared to the previously used FCTight or Tight_VarRad isolation working points. Despite a reduction in signal efficiency of up to 10%, employing PLIVTight is advantageous for this analysis. Additionally, the quality of the modeling in validation regions that are enhanced with misidentified leptons, such as regions with two leptons with the same electric charge, is benefitted by the utilization of the PLIVTight working point.

The muons are obtained via the combined muon definition. Muon candidates must pass the Medium selection criteria, must be within the acceptance of the ID ($|\eta| < 2.5$), and have $p_T > 15$ GeV.

All leptons to be considered must come from the hard-scatter primary vertex by requiring the absolute value of the longitudinal impact parameter to satisfy $|z_0 \sin \theta| < 0.5$ mm. Furthermore, the significance of the transverse impact parameter is required to be $|d_0|/\sigma_{d_0} < 5$ ($|d_0|/\sigma_{d_0} < 3$) for electrons (muons), where, where d_0 and z_0 are the transverse and longitudinal impact parameters of the lepton track with respect to the primary vertices, and σ_{d_0} is the uncertainty on d_0 . The leptons must also be well isolated, using Gradient and GradientLoose criteria.

channel	Electron ID	η range	Author	Isolation	Impact parameter
0/1 -jet 2-jet	LHTight LHMedium	$(\eta < 1.37) \parallel (1.52 < \eta < 2.47)$	1	PLIVTight	$ z_0 \sin \theta < 0.5$ mm, $ d_0 /\sigma_{d_0} < 5$

Table 4.4: Electron selections used in the analysis.

p_T range	Muon ID	Isolation	Impact parameter
$> 15\text{GeV}$	"Tight"	PLIVTight	$ z_0 \sin \theta < 0.5 \text{ mm}, d_0 /\sigma_{d_0} < 3$

Table 4.5: Muon selection used in the analysis.

4.3.2 Jets and Missing Transverse Momentum

The jets are reconstructed utilizing the anti- k_t algorithm, employing a distance parameter of $R = 0.4$, as specified in Section 3.2. Jets with $|\eta| < 2.4$ must have a $p_T > 25 \text{ GeV}$, while jets outside this range must have a $p_T > 30 \text{ GeV}$. Jets with a pseudorapidity $|\eta| > 4.5$ are excluded. To suppress pile-up, jets with $p_T < 60 \text{ GeV}$ and $(|\eta|) < 2.4$ must have a calibrated JVT variable greater than 0.64.

Jets containing b-hadrons are identified using the MV2c10 b -tagging algorithm, which operates with an 85% efficiency working point and is subsequently referred to as b -jets. In each signal region definition, a b -jet veto is applied for jets with $p_T > 20 \text{ GeV}$ to suppress top-quark background. Jets in the range $20 \text{ GeV} < p_T < 30 \text{ GeV}$ are referred to as "sub-threshold" because they fall below the nominal p_T requirement defining the jet categories of the analysis and are also included in the b -veto to improve the suppression of the top backgrounds.

Leptons that deposit significant energy in the calorimeter may be reconstructed as jets. Furthermore, jets may be reconstructed as leptons originating from the hard scatter. Hence, jets are eliminated if they fall within a cone of size $\Delta R < 0.2$ of an electron candidate or have fewer than three associated tracks and are within a cone of size $\Delta R < 0.2$ of a muon candidate. However, if a jet that has three or more linked tracks is located within a cone of size $\Delta R < 0.4$ of a muon candidate or if any jet is situated within $0.2 < \Delta R < 0.4$ of an electron candidate, the corresponding lepton candidate is eliminated instead.

This analysis utilizes the track soft term missing transverse energy E_T^{miss} and the track-based missing transverse momentum p_T^{miss} , as defined in Section 3.2.5. The reconstructed missing transverse energy (E_T^{miss}) is determined using the physics objects that have undergone overlap removal and the energy that is not attributed to these objects after pile-up suppression.

4.3.3 Overlap Removal

The overlap removal (OR) in the analysis deals with the issue of duplication and isolation, which involves handling two separate but nearby objects. Hence, a strong correlation exists between the elimination of overlap and the identification of objects. The following steps are taken to remove the object overlaps, with the rejected objects being excluded from further consideration.

- Electron-muon: If a muon emits a hard photon and the energy deposited in the calorimeter matches the muon track, it can be identified as an electron. The electron

is eliminated if a muon and an electron are detected together ($\Delta R < 0.01$) and share the same identification track. The overlapping leptons with the same flavor are addressed during the reconstruction process.

- **Electron-jet:** Electrons can be reconstructed as a jet due to the energy deposits they leave in the calorimeter. The jet is discarded if its $\Delta R(jet, e) < 0.2$ to eliminate these duplicates. If the distance between a jet and an electron, $\Delta R(jet, e) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^e)$, then the electron is removed. This is because it is probable that the electron is a genuine decay product (either light or heavy flavor) of the jet.
- **Muon-jet:** If $\Delta R(jet, \mu) < 0.2$, then the jet will be removed if the jet has fewer than three associated tracks with a $p_T > 500 \text{ MeV}$. If the value of $\Delta R(jet, \mu) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^\mu)$, then the muon is eliminated from any surviving jets. This is because the muon is probably a genuine heavy flavor decay product of the jet.

4.4. Observables

Various analytic quantities can be derived from the pp collision event on an event-by-event basis (observables). In a collection of numerous events, constraints on the distributions of observables can lead to a significant distinction between signal and background processes. In this analysis, numerous observables are defined and described below. They are classified into three categories:

4.4.1 Topological Observables

These observables are defined to exploit the expected kinematics of the Higgs boson signal process.

- $\Delta\phi_{\ell\ell}$: The azimuthal angle between the two leptons, $\Delta\phi_{\ell\ell}$, will be smaller for the Higgs boson signal than for background processes, such as non-resonant WW . This is because of the spin correlations between the two leptons for the Higgs boson signal process.
- $m_{\ell\ell}$: The invariant mass of the two leptons, $m_{\ell\ell}$, is expected to be smaller for the Higgs boson signal than for background processes for the same reason as $\Delta\phi_{\ell\ell}$.
- m_T : The transverse mass of the lepton pair, m_T , is defined as

$$m_T = \sqrt{\left(E_T^{\ell\ell} + E_T^{\text{miss}}\right)^2 - \left|\vec{p}_T^{\ell\ell} + \vec{E}_T^{\text{miss}}\right|^2} \quad (4.2)$$

where $E_T^{\ell\ell} = \sqrt{|\vec{p}_T^{\ell\ell}|^2 + m_{\ell\ell}^2}$, $\vec{p}_T^{\ell\ell}$ is the vector sum of the lepton transverse momenta, and E_T^{miss} (with magnitude E_T^{miss}) is the missing transverse momentum.

4.4.2 Background Observables

These observables are particularly effective in rejecting specific background events.

- $m_{\tau\tau}$: The invariant mass of the τ -lepton pair in $Z/\gamma^* \rightarrow \tau\tau$ events is reconstructed using the collinear approximation method [97]. This method assumes the τ -leptons are highly boosted such that the decay products are collinear to the τ -lepton directions and that the neutrinos are the only source of E_T^{miss} in the event. This observable is effective at reducing background from Z/γ^* events, as well as $H \rightarrow \tau\tau$ events.
- $N_{b\text{-jet}}$: The number of jets with $p_T > 20\text{GeV}$ and $|\eta| < 2.5$ that are identified as b -jets by the b -tagging algorithm DL1r. $N_{b\text{-jet}}$, is effective in reducing the background from top quark events.
- $p_T^{\ell\ell}$: The vector sum of the two highest transverse momentum leptons is generally substantial for Higgs boson signal events, where the two leptons oppose the two neutrinos. This observable is proficient in diminishing Z/γ^* events, wherein the leptons are generally oriented back-to-back. This is one of the most important observable for this thesis.
- $\Delta\phi(\ell\ell, E_T^{\text{miss}})$: The azimuthal angle between the two highest- p_T leptons and E_T^{miss} is typically back-to-back for Higgs boson signal events and most other background events. Small values of this observable most likely represent mis-measured objects.
- $\max(m_T^\ell)$: The transverse mass of each lepton is defined as

$$m_T^{\ell_i} = \sqrt{2p_T^{\ell_i} E_T^{\text{miss}} (1 - \cos \Delta\phi(\ell_i, E_T^{\text{miss}}))} \quad (4.3)$$

where i represents either of the two highest- p_T leptons. This quantity is typically small for Z/γ^* and multi-jet background processes but will be larger for at least one lepton in the Higgs boson signal process due to the presence of at least one real W boson. It is effective to place a lower bound on the maximum value of m_T^ℓ for the two leptons.’

- m_{T2} : The observable m_{T2} extends the concept of m_T to decay processes with identical decay branches, a and b , originating from the same parent particles, in order to determine the mass of the parent particle. It can be applied to events originating from dileptonic $t\bar{t}$ decays, as well as dileptonic non-resonant WW decays, in order to estimate the top quark mass or W boson mass. Generally, m_{T2} is defined as

$$m_{T2} = \min_{p_T^a + p_T^b = E_T^{\text{miss}}} \left[\max(m_T^a, m_T^b) \right] \quad (4.4)$$

The m_{T2} observable is used in the estimation of the non-resonant WW background yields in the ggF-enriched $N_{\text{jet}} \geq 2$ category.

4.4.3 Other observables

These observables are beneficial in identifying ggF-enriched and VBF-enriched $N_{\text{jet}} \geq 2$ events and are utilized to formulate signal-rich or background-rich SRs and CRs or serve as training for the DNN.

- $p_{\text{T}}^{\text{tot}}$: The total transverse momentum is defined as the summation leptons, missing transverse energy, and the total momentum of all the jets that pass the object selection requirements. Large values of this observable are obtained for background events with significant soft gluon radiation. This radiation forms jets that are typically below the $p_{\text{T}} = 30\text{GeV}$ threshold for jet counting in this analysis, resulting in an imbalance of the observed transverse momentum in the event.
- Δy_{jj} : The presence of two forward, color-disconnected jets in VBF Higgs boson signal events means there will be a large difference in the rapidity of these jets, Δy_{jj} , with respect to background processes and ggF signal events, where additional hadronic activity is typically very central in the detector.
- $m_{\ell_{ij_i}}$: The invariant masses of four combinations of leptons and jets, $m_{\ell_{ij_i}}$, with $i = \{1, 2\}$, with 1 and 2 denoting the two highest $-p_{\text{T}}$ objects, are used as input to the DNN.
- $E_{\text{T}}^{\text{miss}}$ significance: The observable $E_{\text{T}}^{\text{miss}}$ significance is defined, event-by-event, as a ratio of likelihood functions, quantifying the level of agreement with the hypothesis that the measured $\mathbf{E}_{\text{T}}^{\text{miss}}$ is consistent with zero true $\mathbf{E}_{\text{T}}^{\text{miss}}$, denoted $p_{\text{T}}^{\text{inv}}$. Large values of $E_{\text{T}}^{\text{miss}}$ significance imply the measured $\mathbf{E}_{\text{T}}^{\text{miss}}$ is not consistent with the resolution of physics objects entering the $\mathbf{E}_{\text{T}}^{\text{miss}}$ calculation, and that the event is more likely to have involved unidentifiable particles. This observable is used as an input variable to the DNN and is helpful in suppressing top-quark background events.
- Central jet veto (CJV): The jet with the highest p_{T} above 30 GeV that is located between the two VBF jets in pseudorapidity is defined as the central jet. Using this jet, the central jet veto (CJV) is defined as

$$\text{CJV} = \begin{cases} 1 & \text{no central jet with } p_{\text{T}} > 30\text{GeV} \text{ in the event.} \\ 0 & \text{there is a central jet with } p_{\text{T}} > 30\text{GeV} \text{ in the event.} \end{cases}$$

- m_{jj} : Given the distinctive forward nature of the two VBF jets, kinematic arguments imply the invariant mass of the two jets, m_{jj} , is expected to be larger than the invariant mass of two jets in the background and ggF Higgs boson signal events.
- Outside Lepton Veto (OLV) Similarly to the CJV above, the OLV determines whether one or both leptons are located within the rapidity gap of the two VBF jets and is defined as

$$\text{OLV} = \begin{cases} 1 & \text{both leptons located within rapidity gap.} \\ 0 & \text{one or both leptons not located within rapidity gap.} \end{cases}$$

- $\eta_\ell^{\text{centrality}}$: The location of the two leptons in the detector may be defined with respect to the two VBF jets on the η -axis as:

$$\eta_\ell^{\text{centrality}} = \sum_i 2 \cdot \left| \frac{\eta_{\ell_i} - \bar{\eta}}{\eta_{j_0} - \eta_{j_1}} \right|$$

where $\bar{\eta} = (\eta_{j_0} + \eta_{j_1})/2$ is the average pseudorapidity of the two VBF jets, j_0 and j_1 , and ℓ_i is one of the leptons. This observable is used as an input to the DNN.

- p_T^H : The p_T of Higgs boson candidates, p_T^H , is defined at detector-level as the vector sum of the p_T of the lepton pair and E_T^{miss} i.e.

$$p_T^H = \left| \mathbf{p}_T^{\ell\ell} + \mathbf{E}_T^{\text{miss}} \right|$$

- p_T^{Hjj} : Similarly to p_T^H , p_T^{Hjj} is defined using detector-level inputs as the vector sum of the p_T of the lepton pair, E_T^{miss} , and the p_T of the two highest- p_T jets, j_0 and j_1 , i.e.

$$p_T^{Hjj} = \left| \mathbf{p}_T^{\ell\ell} + \mathbf{E}_T^{\text{miss}} + \mathbf{p}_T^{j_0} + \mathbf{p}_T^{j_1} \right|$$

4.5. Event Selection

Measurements of the Higgs boson are conducted by categorizing events according to selection criteria on kinematic quantities (observables) derived from the four-momenta of final state particles, which are measured and identified by the detector. Following the selection of events featuring two high- p_T leptons of opposite charge and same flavor (the preselection), four analysis categories are established according to jet multiplicity as shown in Figure 4.7. This is executed due to the varying background compositions in each jet-binned category and the distinct jet multiplicities associated with the ggF and VBF production modes of the Higgs boson.

The $N_{\text{jet}} = 0$ and 1 categories target the measurement of the ggF production mode, and the $N_{\text{jet}} \geq 2$ category is divided into ggF-enriched and VBF-enriched categories. This manuscript will only focus on the $N_{\text{jet}} \geq 2$ categories targeting the ggF production mode. The primary analysis can be found in [19] to get a complete picture of the study of $H \rightarrow WW^*$.

This category was selected for the thesis based on the previous discussion. The

ggF process only generates jets via parton radiation originating from the initial state partons, while jets serve as a characteristic element of the production mechanism for VBF. Therefore, the highest signal yield for the ggF production mode is expected for lower jet multiplicity ($N_{\text{jets}} < 2$). In previous studies, the ggF-enriched $N_{\text{jet}} \geq 2$ category was combined with the VBF-enriched $N_{\text{jet}} \geq 2$ category as $N_{\text{jet}} \geq 2$ category. So, the ggF-enriched $N_{\text{jet}} \geq 2$ category was not studied as a distinct category.

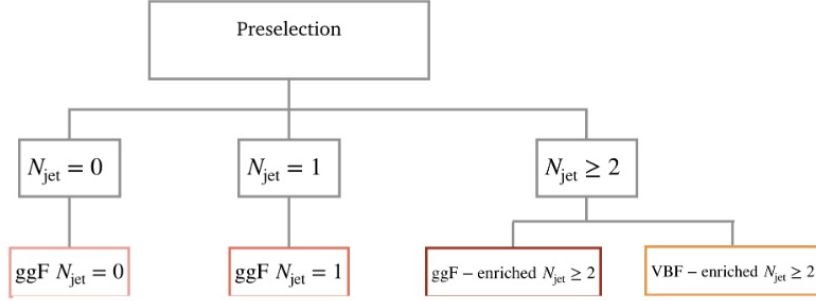


Figure 4.7: Overview of the categorization of events based on jet multiplicity.

Selection criteria for the distribution of observables are optimized for each analysis category to enhance the event yield of the Higgs boson signal process. There are three types of phase spaces used in this analysis. They are the signal region (SR), the control region (CR), and the validation region (VR). The event yield resulting from a defined set of selection criteria is termed the signal region when it focuses on the Higgs boson signal or the control region when it targets a particular background process. In the SR, the signal's contribution to the background is highest, so this region is used in the statistical fit to find the parameter of interest (POI). The CR, on the other hand, is meant to have a fairly pure contribution from a key background. Hence, its main purpose is to make sure that the background process is well-modeled and understood. The utilization of CRs allows for the concurrent measurement of background process rates alongside the Higgs boson signal, thereby diminishing the reliance of background rates in signal-enriched regions on predictions derived from Monte Carlo simulations. The VRs do the same thing as the CR. The only difference is that the CR is used in the statistical fit to limit the normalization parameter of the background, while the VRs' job is to make sure that the data and MC agree and validate the fit over the SR and CR. A Deep Neural Network (DNN) discriminates the signal from background processes, which is trained on all background processes, including the VBF Higgs boson signal, to assess a score that quantifies the likelihood of an event being more ggF-like or more background-like.

By requiring events to fail either the central jet veto (CJV) or outside lepton veto (OLV), the two categories can be forced to be mutually exclusive. Here, CJV rejects events with jets above the $p_T > 30$ GeV, and OLV ensures the two charged leptons within the two-tag jet. This criterion is designed to retain events that are part of the signal process, as the ggF processes can produce events with central jets and tend to produce fewer leptons in the final state. VBF processes, on the other hand, are characterized by

the presence of two high-energy forward jets at very large rapidities and typically involve the exchange of neutral vector bosons (Z bosons). The leptonic decay of these bosons can contribute to background events. With higher luminosity, an independent study of the ggF-enriched $N_{\text{jet}} \geq 2$ category, especially in the same flavor (SF) channel, is necessary for the legacy analysis.

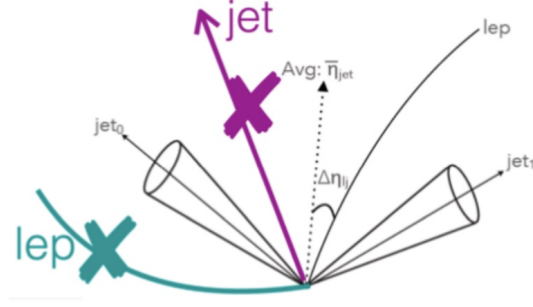


Figure 4.8: Illustration of Central Jet Veto (Purple) and Outside Lepton veto (Teal).

A summary of the event selections for the ggF-enriched $N_{\text{jet}} \geq 2$ category with SF channel is provided in Table 4.6.

Category	$N_{\text{jet}, (p_T > 30\text{GeV})} \geq 2\text{ggF(SF)}$
Preselection	Two isolated, same-flavor leptons ($\ell = e, \mu$) with opposite charge $p_T^{\text{lead}} > 22\text{GeV}, p_T^{\text{sublead}} > 15\text{GeV}$ $m_{\ell\ell} > 12\text{GeV}$ $E_T^{\text{miss}} \text{ significance} > 4, p_T l l > 40\text{GeV}$
Background Rejection	$N_{b\text{-jet}, (p_T > 20\text{GeV})} = 0$ $m_{\tau\tau} < m_Z - 25\text{GeV}$
$H \rightarrow WW^* \rightarrow ee/\mu\mu$ topology	$m_{\ell\ell} < 55\text{GeV}$ $\Delta\phi_{\ell\ell} < 1.8$ VH veto : $ m_{jj} - 85 > 15\text{GeV}$ or $\Delta y_{jj} > 1.2$ VBF veto: fail central jet veto or fail outside lepton veto

Table 4.6: Event selection criteria used to define the signal regions (SR).

The initial sample of events must meet the preselection criteria for data quality and triggering and must consist of exactly two leptons, as previously discussed, with the same flavors and opposite charges. Furthermore, the leading lepton must have a $p_T > 22$ GeV, while the subleading lepton must have a $p_T > 15$ GeV. For the analysis of categories focusing on the ggF production mode, selection criteria of a dilepton invariant mass $m_{\ell\ell} > 12$ GeV is implemented to suppress any low mass meson resonances. This is a tighter selection compared to other channels to reduce Υ and similar resonances. The b -jet veto rejects b -tagged jets with $p_T > 20$ GeV, subsequently suppressing the top quark production background.

As mentioned earlier, the ggF-enriched $N_{\text{jet}} \geq 2$ category events must fail either the CJV or OLV. Furthermore, $V(\rightarrow qq)H$ production is suppressed by rejecting events in a region defined by $|m_{jj} - 85| \leq 15\text{GeV}$ and $\Delta y_{jj} \leq 1.2$. By requiring $m_{ll} < 55\text{ GeV}$, combined with the small angle of lepton opening with the kinematics of a Higgs boson, effectively decreases WW and the Drell-Yan backgrounds and also accounts for off-shell contributions. The $\Delta\phi_{\ell\ell}$ selection is applied in the ggF-enriched $N_{\text{jet}} \geq 2$ category, which reduces the remaining DY background while retaining most of the signal and also accounts for spin correlations contributions. A requirement of $m_{\tau\tau} < m_Z - 25\text{ GeV}$ significantly reduces the remaining DY contribution and is applied. Furthermore, $p_T^{\ell\ell}$ provides a handle over the unique $H \rightarrow W^\pm W^\mp$ topology, large values of $p_T^{\ell\ell}$ corresponding to a scenario when the leptons and the neutrinos are emitted in opposite hemispheres. E_T^{miss} significance ($E_T^{\text{miss}} \text{ significance} \leq E_T^{\text{miss}}/10\text{GeV}$) is also added to suppress any Zll contributions.

A statistical analysis is conducted on a selected and categorized set of events involving two distinct measurements of the Higgs boson production cross sections. The initial measurement is inclusive, simultaneously assessing the ggF and VBF production rates of the Higgs boson through regions classified by jet multiplicity. For the inclusive measurement, all the ggF-enriched $H \rightarrow WW^*$ events are considered ggF signals. A single parameter of interest is linked to the ggF contribution. The same is done for VBF - all the $H \rightarrow WW^*$ VBF events are considered VBF signals. Higgs contributions other than $H \rightarrow WW^*$ ggF and VBF are small in the SRs and are considered backgrounds.

The second is the STXS measurement, which quantifies ggF and VBF production of the Higgs boson across distinct kinematic fiducial regions. The areas designated for the STXS measurement are subdivisions of the SRs utilized for the inclusive measurement, determined by observables, including the reconstructed p_T of the Higgs boson candidate and the reconstructed invariant mass of the two highest p_T jets.

4.6. STXS Categorization

Multiple kinematic fiducial regions are utilized to optimize the cross-section measurements. The fiducial region is divided into distinct and non-overlapping partitions, known as STXS (Simplified Template Cross Sections) bins [98, 99, 100, 101]. These mutually exclusive categories of phase space, called bins, are created based on various kinematic distributions derived from the physics objects in the fiducial volume and designed to be theoretically meaningful and experimentally distinguishable. Simply, STXS is a structured way of classifying events into bins based on kinematic properties such as transverse momentum, rapidity, and other properties of particles. The framework was developed as a complementary method for measuring fiducial differential cross-sections and as an evolution in signal strength measurements.

The purpose of defining the STXS bins is to enhance the usefulness and accuracy of STXS measurements. This is achieved by designing the bins to:

- maximize the experimental sensitivity to events that involve Higgs bosons by pro-

viding finely-grained measurements,

- minimize uncertainties in extrapolating measurements to regions that depend on theoretical prediction,
- allow the combination of measurements from various decay modes and different experiments,
- isolate possible BSM effects.

The measurement is executed through profile likelihood unfolding: in the likelihood fit, parameters of interest are established that adjust the signal yields of the templates associated with the specific truth bins. To enhance sensitivity, signal regions are delineated to closely align with the truth-level categorization. Thus, cross-contamination between the truth-level and reco-level bins can be mitigated.

To meet these requirements, the bins are determined based on global event quantities, such as the reconstructed transverse momentum of the Higgs boson (p_T^H), the number of reconstructed jets (N_{jet}), and the invariant mass of two jets (m_{jj}). Furthermore, as more data is collected, the STXS bins can be subdivided into smaller sub-bins to refine the measurement. This technique of refining the granularity of the STXS bins is called a “stage”, where each stage corresponds to a distinct set of partitions within the fiducial region. The extent to which the production cross-sections are divided is primarily determined by the current precision of the measurements using the available dataset. The $H \rightarrow WW^*$ analysis signal samples have been split according to the Stage 1.2 STXS framework [101]. Unlike the VBF, there is no m_{jj} splitting in the ggF channel because the sensitivity is not sufficient.

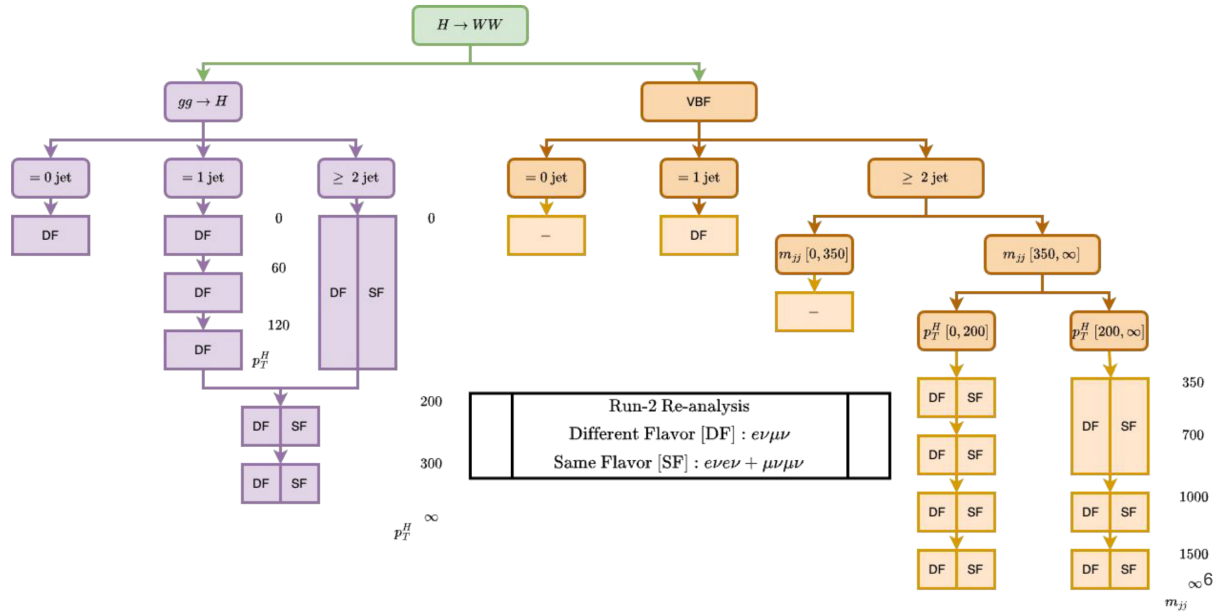


Figure 4.9: Detailed illustration of Legacy run II HWW STXS measurement.

However, current measurements focusing on specific modes of the Higgs boson production and decays do not have enough sensitivity to define many kinematic regions precisely using reconstructed particle objects. Several STXS bins are merged to form Reduced STXS Stage 1.2 to address this limitation. The ggF template process for stage 1.2 is characterized by the $gg \rightarrow H$ process, along with higher-order QCD and electroweak corrections. This process can be categorized into several mutually exclusive truth bins concerning the four-momentum of the Higgs truth particle and the corresponding truth-level jets. STXS for the ggF-enriched $N_{\text{jet}} \geq 2$ category applies three STXS bins based on the transverse momentum of the Higgs boson ($p_T^H = |\mathbf{p}_T^{\ell\ell} + \mathbf{E}_T^{\text{miss}}|$) with the same signal region splitting:

- $p_T^H < 200$ GeV selection,
- $200 \text{ GeV} < p_T^H < 300$ GeV selection,
- $p_T^H < 300$ GeV selection.

Figure 4.10 illustrates the relative contributions of the various merged STXS bins across all reconstructed SRs. This analysis has reco-level signal regions in which the signal is decomposed into truth categories. Each of these regions is used by the corresponding truth STXS category signal component and thus drives the sensitivity to these truth POIs in the fit. In each instance, the target categories yield the most significant contribution to the respective SRs designed to select them.

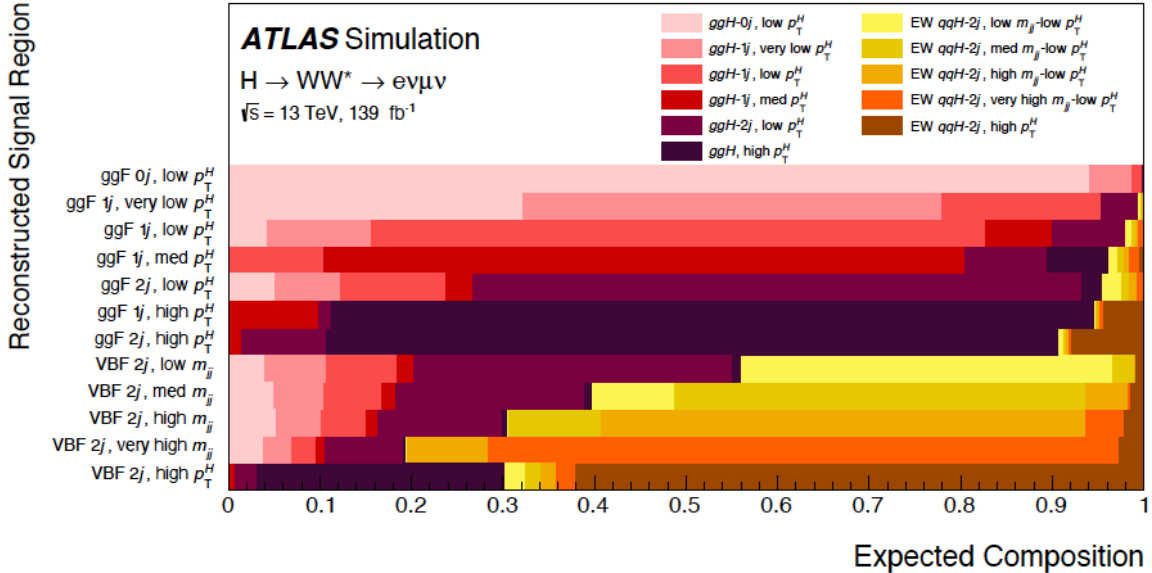


Figure 4.10: Detailed illustration of Legacy run II HWW STXS measurement.

4.7. Background Estimation

Various methods are used to address the wide range of background processes that affect the final signal region (SR) yields which include WW , top ($t\bar{t}$ and Wt), WZ , $W\gamma^{(*)}$, ZZ , W +jets, QCD, $Z \rightarrow \tau\tau$, and $Z \rightarrow ee/\mu\mu$ events [102]. These backgrounds can be classified into different categories based on their properties in the final state, enabling them to satisfy the signal region selection criteria.

- The final state of $t\bar{t}$, Wt , and WW includes two W bosons, which is similar to the signal. The existence of b -tagged jets can be used to reject processes that involve top quarks.
- VV processes ($Z/\gamma^* \rightarrow \tau\tau$, and the "Non- WW diboson" processes WZ , $W\gamma^{(*)}$, and ZZ) have a small cross-section and can be accurately represented by the Monte Carlo generator and theoretical predictions of the cross-section. So, no control regions (CR) are used for these backgrounds.
- The production of W + jets and multijets through QCD processes has a high cross section. These events become relevant in the signal region when a jet gives rise to an object that is reconstructed as an isolated lepton. These backgrounds with misidentified leptons (non-prompt leptons from heavy-flavor hadron decay or jets misidentified as prompt leptons) are handled by fully data-driven fake factor methods.

This analysis utilizes multiple control regions (a subset of the data enriched in the background process being estimated and with low expected signal) to estimate the normalization factors by full inclusion to the likelihood fit of important background contributions from the data. Every control region and validation region starts from the preselection. CRs are constructed to study and understand the characteristics of background processes that mimic the signal of interest in the analysis, while VRs are used to validate the models and ensure that the analysis techniques are not biased by overfitting the specific characteristics of the control or signal regions. A CR for a specific background process is usually defined by inverting or modifying one of the event selection requirements of the SR.

CR	$N_{\text{jet}} = 2$
top quark	$n_{b\text{-jets}} = 1$
	$\Delta\phi_{\ell\ell} < 1.8$
	$m_{\tau\tau} < m_Z - 25 \text{ GeV}$
	$m_{ll} < 55 \text{ GeV}$
	$p_{\text{T}}^{\ell\ell} > 40 \text{ GeV}$
$Z \rightarrow ll$	$n_{b\text{-jets}} = 0$
	$ m_{\ell\ell} - m_Z < 15 \text{ GeV}$
	$\Delta\phi_{\ell\ell} < 1.8$
	$p_{\text{T}}^{\ell\ell} > 40 \text{ GeV}$
	CJV30 and OLV = False
VR	$N_{\text{jet}} = 2$
$qq \rightarrow WW$	$n_{b\text{-jets}} = 0$
	$m_{\ell\ell} > 80 \text{ GeV}$
	$m_{\text{T}2} > 165 \text{ GeV}$
$Z \rightarrow \tau\tau$	$n_{b\text{-jets}} = 0$
	$m_{\ell\ell} < 55 \text{ GeV}$
	$ m_{\tau\tau} - m_Z < 25 \text{ GeV}$

Table 4.7: Event selection criteria used to define the control regions (CR) and WW validation region (VR).

4.7.1 Top quark Background

The top quark backgrounds are associated with the $t\bar{t}$ and Wt processes. Top backgrounds are expected to contribute more to $N_{\text{jet}} \geq 2$ and $N_{\text{jet}} = 1$ channels than at $N_{\text{jet}} = 0$ channel by virtue of having b -jets. Therefore, by inverting the b -jet veto, the CRs are orthogonal to the SRs. However, in the ggF-enriched $N_{\text{jet}} \geq 2$ category, the top quark CR is defined by excluding events with b -jets and ensuring the orthogonality to the SR and WW CR. This is achieved by imposing two conditions: $m_{\ell\ell} > 80 \text{ GeV}$ and $m_{\text{T}2} < 165 \text{ GeV}$. The top quark CRs have pre-fit top quark process purities of 71 % for ggF enriched $N_{\text{jet}} \geq 2$, and 97% for VBF-enriched $N_{\text{jet}} \geq 2$.

4.7.2 Zll Background

The Zll control region for the ggF 2jSF analysis is intended to verify the background estimation, especially from the Drell-Yan processes, in which a Z boson decays into two leptons ($Z \rightarrow ll$). The $Z \rightarrow \ell\ell$ background is one of the leading contributions in the same-flavor channel. This control region is defined by selecting events featuring two jets and two identical-flavor leptons (electrons or muons) possessing an invariant mass aligned with the Z boson mass range. The events outside the Z mass region must have $|m_{ll} - m_Z| > 15 \text{ GeV}$ to reduce the non- Z background contributions inside the Z mass region (the resonant nature of this background). Since it does not produce real $E_{\text{T}}^{\text{miss}}$ (as compared to $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$), missing transverse momentum is mostly incurred via mismeasurements

of additional jets, making this background reducible via minimum requirements on E_T^{miss} and notably its significance E_T^{miss} Significance. Therefore, $\text{METsig} > 4.0$ is used to reject $Z \rightarrow ee/mm$ background and $p_T^{\ell\ell} > 40$ GeV to mitigate the mis-modeling of Z -jets. Despite these measures, this background is the main reason for the deterioration of the sensitivity of the same-flavor channel with respect to the cleaner different-flavor signature.

4.7.3 STSX background estimation

The methodology employed for estimating background processes in the STXS measurement is based on the same principle that was used for selecting the SR. The CRs for the background processes in each analysis category serve as the initial reference for the STXS CRs. The STXS CRs incorporate additional selection criteria for p_T^H or m_{jj} , ensuring that the background rate in an STXS SR is standardized using a corresponding NF from a dedicated CR that also has the same selection criteria for p_T^H or m_{jj} .

4.7.4 Normalization factors

A control region (CR) for a particular background process is typically established by inverting or altering one of the event selection criteria of the signal region (SR), thereby creating a newly defined area that is enriched in the specific background process of interest and maintains kinematic similarity to the SR. Since the background is usually much more dominant than the signal, this control region can often be set up so that the signal contribution is negligible in that area. A normalization factor (NF), denoted by β_i , for a specific background process, i , measured in a control region (CR), is defined as the ratio of the background-subtracted data yields to the event yields of process i predicted by theory for that region. The NFs are used to normalize the rate of the respective background processes in distributions and yield tables.

$$\beta_i = \frac{N_{\text{data}}^{\text{CR}} - N_{\text{bkg}}^{\text{CR}}}{N_{i,\text{theo}}^{\text{CR}}}, \quad (4.5)$$

This normalization factor is applied to the signal region. The estimated yield of a process i in the signal region N_i^{SR} , is derived using the corresponding theory prediction $N_{i,\text{theo}}^{\text{SR}}$ as

$$N_i^{\text{SR}} = \beta_i \times N_{i,\text{theo}}^{\text{SR}} \quad (4.6)$$

$N_{i,\text{theo}}^{\text{SR}}$ denotes the event yield of a process i in the signal region as forecasted by Monte Carlo simulation. Equation 4.3 can be altered as follows:

$$\begin{aligned}
N_i^{\text{SR}} &= \frac{N_{\text{data}}^{\text{CR}} - N_{\text{bkg}}^{\text{CR}}}{N_{i,\text{theo}}^{\text{CR}}} \times N_{i,\text{theo}}^{\text{SR}} \\
&= \left(N_{\text{data}}^{\text{CR}} - N_{\text{bkg}}^{\text{CR}} \right) \times \frac{N_{i,\text{theo}}^{\text{SR}}}{N_{i,\text{theo}}^{\text{CR}}} \\
&= \left(N_{\text{data}}^{\text{CR}} - N_{\text{bkg}}^{\text{CR}} \right) \times \alpha_i^{\text{SR}}
\end{aligned} \tag{4.7}$$

where α_i^{SR} denotes an extrapolation factor. As α_i^{SR} represents a ratio of event yields forecasted through Monte Carlo simulation. Numerous systematic uncertainties arising from various experimental and theoretical sources are canceled in the ratio. The cancellation is most significant when the SR and CR have similar kinematics, which explains why the event selection requirements for defining CRs should only be changed minimally. If the background composition in the SR is different from the one in the CR or the MC generator mismodels the ratio, the NF will not model this change accurately. This approach, however, only corrects the total normalization of background yields using the NF; any kinematic mis-modeling in the signal region remains uncorrected.

The defining requirements of a CR are determined by ensuring the theoretical uncertainty in the extrapolation factor, α_i^{SR} , is minimized. It is also important to ensure that selection criteria don't create unnatural trends in the way that observables are distributed. The background NFs and the signal normalizations (signal strengths) in the statistical analysis are determined. However, before performing the full statistical analysis, it is possible to estimate the values of the NFs using only the data from the CRs and a matrix inversion.

Here, the matrix inversion method is a subsequent advancement of the basic normalization factor technique. When estimating all i processes using the basic normalization factor method from i distinct control regions, the results will be skewed due to the contamination of each control region by other backgrounds. This issue can be expressed in matrix notation as

$$n_{\text{bkg}}^{\text{SR}} = \begin{pmatrix} n_1^{\text{SR}} & \cdots & n_i^{\text{SR}} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \end{pmatrix} \tag{4.8}$$

The same normalization factors β_i link the estimates within the control regions, specifically,

$$\begin{pmatrix} n_{\text{obs.}}^{\text{CR1}} \\ \vdots \\ n_{\text{obs.}}^{\text{CR}i} \end{pmatrix} = \begin{pmatrix} n_1^{\text{CR1}} & \cdots & n_i^{\text{CR1}} \\ \vdots & \ddots & \vdots \\ n_1^{\text{CR}i} & \cdots & n_i^{\text{CR}i} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \end{pmatrix} \tag{4.9}$$

where, $n_{\text{obs.}}^{\text{CR}i} = N_{\text{data}}^{\text{CR}i} - N_{\text{bkg}}^{\text{CR}i}$. Hence,

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \end{pmatrix} = \begin{pmatrix} n_1^{\text{CR1}} & \dots & n_i^{\text{CR1}} \\ \vdots & \ddots & \vdots \\ n_1^{\text{CRi}} & \dots & n_i^{\text{CRi}} \end{pmatrix}^{-1} \cdot \begin{pmatrix} n_{\text{obs.}}^{\text{CR1}} \\ \vdots \\ n_{\text{obs.}}^{\text{CRi}} \end{pmatrix} \quad (4.10)$$

Therefore, the consistency issue of the basic normalization factor method can be resolved through a simple matrix inversion. However, a significant limitation of this solution is that uncertainty propagation is no longer straightforward, as the variables involved in the matrix inversion possess associated uncertainties. A straightforward and dependable method to estimate uncertainty involves utilizing random matrices generated within the bounds of their input uncertainties. The mean and standard deviation of the resulting distributions in the normalization factors β are subsequently employed for further calculations. Systematic uncertainties can be propagated by assessing the method for each variation individually and comparing the outcomes. A toy example illustrating a 2×2 matrix is presented in Figure 4.11.

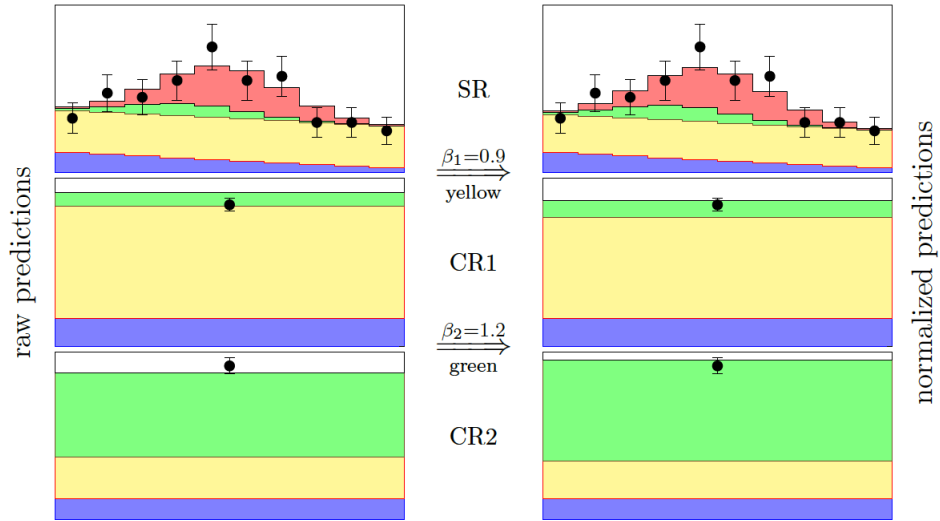


Figure 4.11: Schematic setup with one signal contribution (red) and several background contributions (blue, green, and yellow) with one signal region (SR) and two control regions (CR1 and CR2). Factors that normalize the green and yellow background contributions to the two control regions are measured simultaneously, and then β_i is applied to all regions to change the background yields.

NFs for WW , $Z\tau\tau$, and top quark are taken from the 2-jet different flavor channel and Zll from the 2-jet same flavor channel. This was done to mitigate contamination from Zll processes and to utilize more accurate background estimates. The decision is further justified by the similarity in jet kinematics between the 2jDF and 2jSF channels, both involving two jets, which results in comparable behavior of backgrounds such as WW , top, and $Z\tau\tau$ regarding jet multiplicity and transverse momentum. Moreover, the 2jSF

channel demonstrates higher sensitivity to Zll processes due to the predominance of the same-flavor leptons in this background. Consequently, deriving Zll NFs from the 2jSF channel allows for a better constraint of this specific background, enhancing the overall accuracy of the analysis.

4.8. Misidentified Leptons

Misidentification is an important source of background for physics analyses that employ particle-level identification criteria. Although there have been vigorous attempts to optimize the identification of objects, it is impossible to eliminate the occurrence of misidentification completely. One of the main reasons for using physics signatures involving leptons in the final state is the large reduction of QCD multijet backgrounds achieved through the efficient lepton identification criteria. The level of jet suppression in ATLAS is 10^{-5} [103]. Although it is unlikely for a jet to be mistakenly classified as a lepton, the large production cross-section of QCD multijets at the LHC means that some jets will inevitably be misidentified as leptons, and the fraction of such events is so small that the Monte Carlo simulations cannot accurately model it.

These objects are commonly known as “fake-leptons” or simply just “fakes”. Even some real leptons are also encompassed within this broad category if they arise from secondary processes, such as semi-leptonic heavy flavor decays, charged hadrons, or photon conversions, as these leptons were not produced in isolation as the result of a high-energy electroweak process. A primary lepton produced in isolation from a hard electroweak scattering event is called a prompt lepton. These observations are applicable to both electrons and muons. However, the rate of fake electrons is significantly greater than that of fake muons. This discrepancy is primarily because muons are identified using additional measurements in the dedicated muon system.

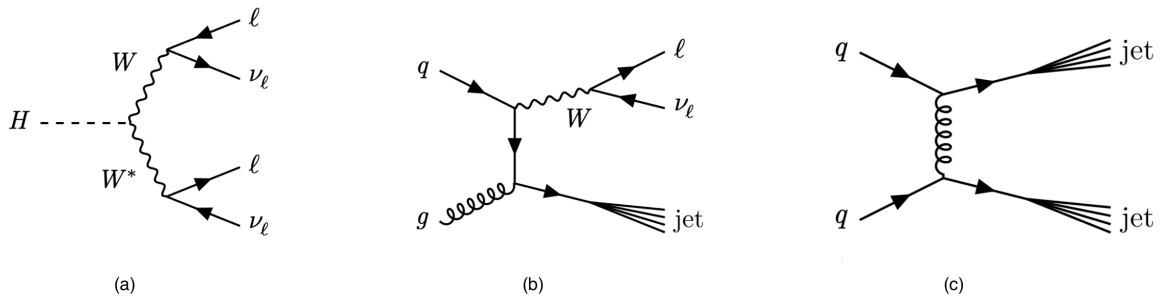


Figure 4.12: Feynman diagrams of signal (a) and W + jets (b) and QCD (c) backgrounds that contain fake leptons.

Figure 4.12 shows the backgrounds that exhibit a signature similar to the HWW signal: W + jets contain one real lepton and one fake lepton, whereas QCD features multiple fake leptons and in $V\gamma^*$ photon fakes as lepton due to photon conversion. In the

di-lepton analyses discussed in this thesis, the majority of misidentified leptons originate from W +jet events in which one prompt lepton is produced when the W boson decays leptonically, accompanied by one or more jets where one of the jets is misidentified as a second prompt lepton. In addition, despite much lower contribution, pure QCD processes can still contribute where two jets are misidentified as prompt leptons. Quantifying this particular background using the measured data is crucial, as the Monte Carlo simulation may not accurately model the misidentification rate due to the need for precise modeling of detector effects. The "fake factor" method is a data-driven technique that models background caused by particle misidentification.

4.8.1 The Fake Factor Method

The fundamental idea of the fake factor method is straightforward: choose a control region (CR) of the background being estimated and then use an extrapolation factor, referred to as the “fake factor”, to relate these events and the background in the signal region (SR) [102]. The method is data-driven, assuming the CR is chosen based on data and the fake factor is determined in a separate fake-enriched data sample using the same definitions as for the CR. The fake factor measures the ratio of background events misidentified according to a specific set of criteria to those misidentified as passing the complete particle selection in the signal region.

The total number of events in the selected signal region with two fully identified leptons can be expressed as

$$N_{\text{id+id}}^{\text{SR}} = N_{\text{id+id}}^{\text{p}} + N_{\text{id+id}}^{\text{f}} \quad (4.11)$$

where the superscripts ‘p’ and ‘f’ denote prompt or fake, respectively.

In order to determine the background contribution from fake leptons $N_{\text{id+id}}^{\text{f}}$, a CR is established with an enhanced rate of fake leptons. Events in this CR must have one lepton satisfying the signal selection criteria and one “anti-identified” or “anti-id” lepton that fails the signal selection but satisfies less restrictive criteria. Table 4.8 outlines the less restrictive anti-identified categories.

ID electron	Anti-ID electron	ID muon	Anti-ID muon
$p_T > 15\text{GeV}$ $ \eta < 2.47$, excluding $1.37 < \eta < 1.52$ $ z_0 \sin \theta < 0.5 \text{ mm}$ Pass LHTight if $p_T < 25\text{GeV}$ Pass LHMedium if $p_T > 25\text{GeV}$ $ d_0 /\sigma(d_0) < 5$ FixedCutTrackCone 40 isolation if $p_T < 25\text{GeV}$ Pass Gradient isolation if $p_T > 25\text{GeV}$ AUTHOR =1	Pass LHLoose	$p_T > 15\text{GeV}$ $ \eta < 2.45$ $ z_0 \sin \theta < 0.5 \text{ mm}$ Pass Quality Tight $ d_0 /\sigma(d_0) < 3$ $E_T^{\text{varcone } 30}/p_T < 0.09$ $p_T^{\text{varcone } 30}/p_T < 0.06$	Pass Quality Medium $ d_0 /\sigma(d_0) < 15$ Veto against identified muon
	Veto against identified electron		

Table 4.8: Summary of the criteria for fully identified (ID) and anti-identified (Anti-ID) electrons (left) and muons (right).

The isolation cuts are removed, and events containing fully identified leptons are rejected in the Anti-ID definition. For electrons, the less restrictive criterion called LHLoose identification is used. For muons, the d_0 significance cut is loosened to enhance the fake contribution. In addition to this change, the events in the CR must satisfy the complete event selection criteria, with the anti-identified lepton being considered fully identified. Similar to the signal region, the total number of events in the W +jets control region can be expressed as

$$N_{\text{id+id}}^{W+\text{jetsCR}} = N_{\text{id+id}}^f + N_{\text{id+id}}^p \quad (4.12)$$

where id represents Anti-ID leptons, and the fake factor is now defined as:

$$f_{\text{id}}^{\text{id}} \equiv \frac{N_{\text{id}}}{N_{\text{id}}} \quad (4.13)$$

Thus, the fake contribution in the SR can be determined as,

$$\begin{aligned}
 N_{\text{id+id}}^f &= N_{\text{id+id}}^f \times f_{\text{id}}^{\text{id}} \\
 &= \left(N_{\text{id+id}}^{\text{data}} - N_{\text{id+id}}^{\text{EWC}} \right) \times f_{\text{id}}^{\text{id}}
 \end{aligned} \quad (4.14)$$

4.8.2 W + jets Control Region

Validation of the fake background estimate is performed in a control region defined by applying the full selection criteria, except requiring ID + Anti-ID leptons instead of ID + ID leptons, and is enriched in W +jets events. This loosening of the identification criteria for one of the leptons increases the contribution in which one lepton originates from a fake

or non-prompt source. However, the sample still includes electroweak backgrounds such as Z +jets, diboson, and top processes, which must be subtracted. Events with an Anti-ID electron are called " e -fake," while events with an Anti-ID muon are called " μ -fake." When they are combined, they are commonly known as either "fakes" or "Mis-Id."

4.8.3 Measurements of the fake factors in Z +jets event

Z + jets events pose an excellent dataset for quantifying the fake factor due to its higher cross-section. Z + jets events are characterized by events having exactly three loosely defined reconstructed leptons, each with $p_T > 15$ GeV. This second measurement of the fake factor is employed to evaluate systematic uncertainties associated with this background estimation technique.

To begin, a comprehensive Z selection is carried out, which necessitates the two of the leptons satisfy the following criteria: -

- The presence of two leptons with the same flavor and opposite charge, each with a $p_T > 25$ GeV.
- The leptons should pass the identification requirements for both electrons and muons, as listed in Table 4.8.
- The invariant mass of the lepton pair falls within the range of $70\text{GeV} < m_{\ell\ell} < 110\text{GeV}$

The data set used to measure the fake factor is constructed by imposing the following criteria:

- The presence of exactly one loosely identified lepton.
 - $p_T > 15$ GeV and
 - For muons, $|\eta| < 2.5$ and for electrons, $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$.
- The additional lepton must be back-to-back with the Z candidate, by imposing $\Delta\phi_{Z\ell} > 2.5$.
- The third lepton is later classified as the fake candidate.
- ZW Veto: Events are rejected if the transverse mass of the system consisting of the third lepton and the missing transverse energy exceeds 50 GeV.

The fake factor is then calculated as the ratio between the number of identified and anti-identified leptons as a function of their transverse momentum and pseudorapidity.

4.8.4 Measurements of the fake factors in di-jet events

A di-jet data sample is used to determine fake factors to cross-check with Z +jets fake factors and address cases involving a trigger bias in the estimation. They are calculated as a function of p_T and $|\eta|$ of the anti-identified object. A dataset is generated from

the low- p_T , single-lepton prescaled trigger streams to assess fake factors. Typically, the anti-identified lepton refers to the lepton with the lower p_T that did not fire the trigger.

The nominal fake factors apply to all events in which the trigger mechanism selected the leading lepton. These triggers are specifically designed to avoid introducing any trigger bias, as they have a looser criterion than the Anti-ID lepton definition. To increase the contribution from dijet events and suppress the background contributions from electroweak processes, the following selections are imposed:

- Exactly one fake candidate object, a loosely identified lepton with $p_T > 15\text{GeV}$
- At least one jet with $p_T > 22\text{GeV}$
- both being separated by an azimuthal angle of $\Delta\phi > 2.5$
- $p_T^{\text{miss}} < 30\text{GeV}$
- $m_T^\ell < 60\text{GeV}$

However, there is a possibility that the anti-identified lepton is the only triggered lepton in the event and, consequently, was selected by one of the un-prescaled single-lepton triggers used in the analysis. Applying the nominal fake factor to these events introduces a slight trigger bias in the background estimation, as the trigger selection interferes with the definition of the fake candidate object. To mitigate this trigger bias, separate fake factors are derived for these events using the dataset selected based on the standard triggers. These dedicated fake factors are called triggered fake factors, which use the analysis triggers rather than prescaled ones. The triggered fake factors are exclusively applied to events where the anti-identified lepton is activated while the identified lepton remains inactive. The nominal fake factors are utilized for all other events.

The fake enriched sample obtained through the selections may contain minor contributions from hard electroweak scattering processes, potentially leading to contamination by real prompt leptons. The contributions from the processes $W + \text{jets}$, $Z + \text{jets}$, and $t\bar{t}$ are estimated using Monte Carlo and subtracted from the sample. The statistical uncertainties that arise from the limited size of these Monte Carlo samples are subsequently applied to the resulting fake factors.

Due to potential changes in detector conditions and prescales, the fake factors may also change. Thus, the data events used to calculate the fake factors are weighed based on the prescale of the trigger employed to gather them. This process guarantees that the fake factors correspond to un-prescaled data in the specific region where they are implemented. Once the selection criteria are implemented, the remaining events are classified as either identified or anti-identified. Once the electroweak background has been removed, these two choices form the numerator and denominator in the fake factor.

Finally, if we want to apply the Fake factor evaluated from $Z + \text{jets}$ sample to $W + \text{jets}$ samples, we need to include a correction factor. This Correction Factor is defined by $CF = \frac{FF^{W+\text{jets}}}{FF^{Z+\text{jets}}}$ and is taken from MC. POWHEG is used as the generator for Correction Factors. All things multiplied give us the yield in the signal region.

$$N_{\text{misld}}^{SR} = F F_{\text{data-driven}}^{Z+\text{jets}} \times C F_{MC} \times N_{\text{misld}}^{CR} \quad (4.15)$$

Figure 4.13 compares the counterfeit factors obtained from Z + jets events and those estimated in di-jet events.

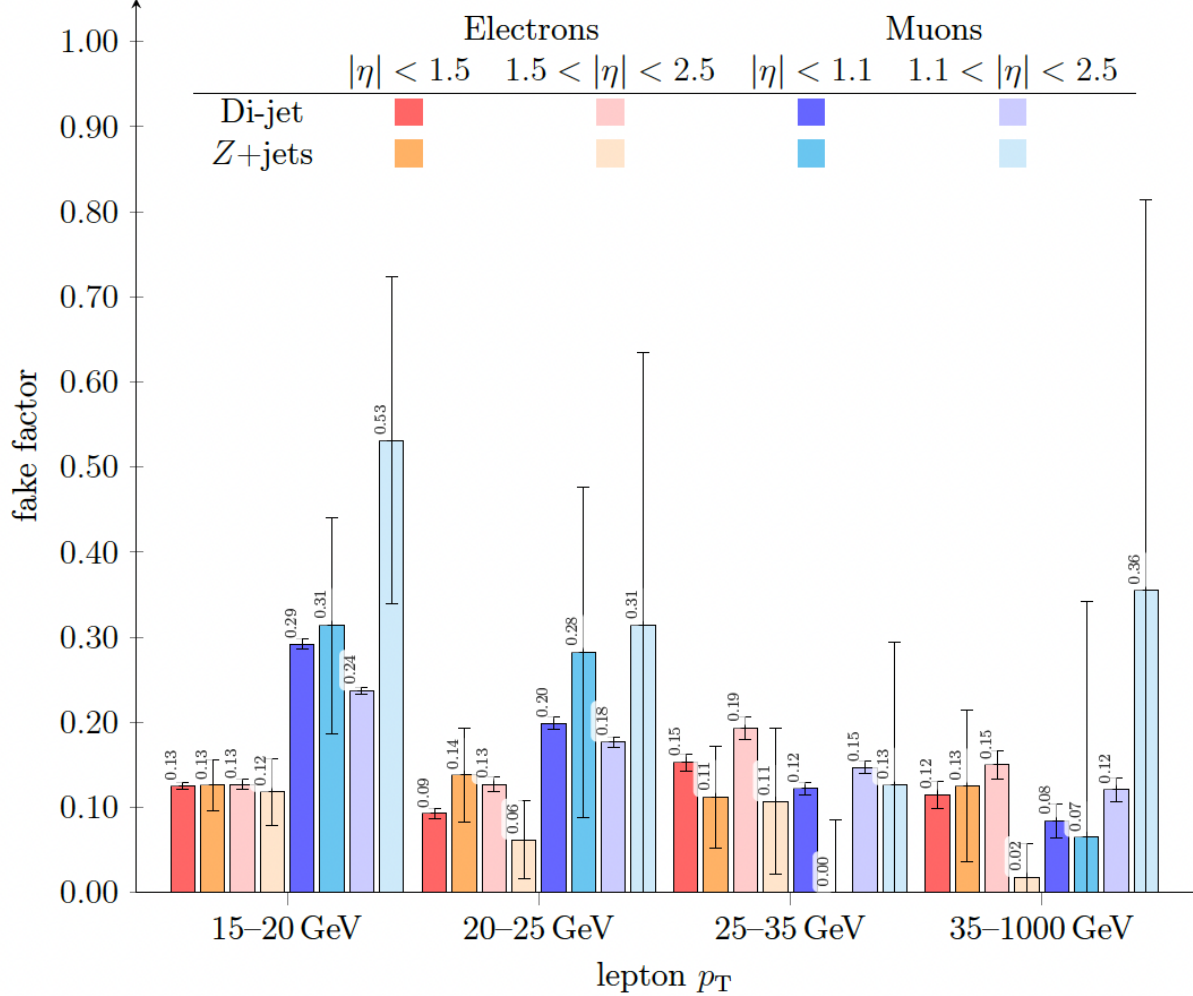


Figure 4.13: Results of the fake factor calculations, including statistical uncertainties.

4.9. Data Analysis Techniques

4.9.1 Maximum Likelihood Method

Introduced in Section 2.4, the likelihood function for a parametric function is a joint probability of observing the given data set x given the parameters μ , and θ , where μ is the set of parameters of interest and θ is the set of nuisance parameters (NPs).

$$\mathcal{L}(\mu, \theta) = P(x; \mu, \theta) \quad (4.16)$$

Here, the likelihood function ($\mathcal{L}$ is itself, not a probability, as it is not normalized to unity. However, larger values of ($\mathcal{L}$ suggest scenarios under which the observation of the given data set is more likely.

In this analysis, the likelihood function is constructed as the product of Poisson terms, one for each bin in all SRs and one for each CR. The full likelihood function is defined by a product of four terms:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = & \prod_i^{N_{\text{SRs}}} \prod_b^{N_{\text{bins}}} \text{Pois} \left(N_{ib} \mid \boldsymbol{\mu} \cdot \mathbf{S}_{ib}(\boldsymbol{\theta}) + \sum_k^{N_{\text{bkg}}} \beta_k \cdot B_{kib}(\boldsymbol{\theta}) \right) \\ & \times \prod_l^{N_{\text{CRs}}} \text{Pois} \left(N_l \mid \boldsymbol{\mu} \cdot \mathbf{S}_l(\boldsymbol{\theta}) + \sum_k^{N_{\text{bkg}}} \beta_k \cdot B_{kl}(\boldsymbol{\theta}) \right) \\ & \times \prod_t^{\theta} \text{Gaus}(\vartheta_t \mid \theta_t) \times \prod_b^{N_{\text{bins}}} \text{Pois}(\xi_b \mid \zeta_b \cdot \theta_b). \end{aligned} \quad (4.17)$$

The function of each term is as follows:

- The first term represents the product of Poisson functions that describe the expected Higgs boson signal and k background yields, $\mathbf{S}(\boldsymbol{\theta})$ and $B_k(\boldsymbol{\theta})$, in each bin (b) of the SRs (i), given the observed data (N_{ib}). The Poisson distribution can be expressed in the form $\text{Pois}(x; \nu) = \frac{\nu^x e^{-\nu}}{x!}$. The anticipated yield of the Higgs boson signal can be expressed as $S_{jib} = \mathcal{L}_{ib} \times (\sigma_j \cdot \mathcal{B}_{H \rightarrow WW*})^{\text{SM}} \times ((A \cdot \epsilon)_{jib}^{\text{SM}})$ for production mode j . The expected Higgs boson signal yield is normalized by the signal strengths while the yields of background processes estimated using a dedicated CR are normalized by the β_k , which are otherwise set to unity.
- The second term represents the multiplication of Poisson functions that describe the expected Higgs boson signal and k background yields. The CRs are used to assess the normalization of the background processes that are the focus of each CR, ensuring that only the overall event yields are considered.
- The third term introduces NPs, $\boldsymbol{\theta}$, into the likelihood function that is used to constrain an associated systematic uncertainty with a Gaussian function.
- The fourth term represents the product of Poisson functions incorporated into the model to account for the finite sample size of the MC samples.

4.9.2 Boosted Decision Tree

Before the deep neural network techniques, the discriminatory power was maximized by training a specialized Boosted Decision Tree (BDT) to provide a clean signal region [104, 105]. Boosted Decision Tree (BDT) is the most common ML model used at ATLAS and CMS for signal discrimination. BDT classifiers are multivariate classifiers that exhibit superior classification performance compared to linear models or simple sets of cuts. In particular, they are generally more effective in utilizing sets of potentially correlated variables with non-linear relationships that offer minimal separation individually. The BDT classifiers are based on eight discriminating variables for each object, and they are categorized as either SM signal or background. The variables have been selected because they exploit properties of the $H \rightarrow W^\pm W^\mp$ decay, such as $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, and m_T , for the ggF analysis, or Δy_{jj} , m_{jj} , $p_T^{\text{tot.}}$, $\sum_{\ell,j} m_{\ell j}$, and η_ℓ for the characteristic VBF signature. They are trained with data labeled as signals or backgrounds to adapt them to their application. The BDT classifiers are subsequently employed to classify unseen data.

A decision tree iteratively divides the parameter space spanned by the input variables into mutually exclusive regions. Using a greedy algorithm, the cuts are selected to minimize the yield-weighted mean of the Gini impurities, I , in the resulting nodes.

$$I = 2p \cdot (1 - p), \quad \text{and} \quad p = \frac{\text{signal yield}}{\text{total yield}} \quad (4.18)$$

The selection of each cut is entirely determined by the training data in the corresponding node, which becomes increasingly sparse with each iteration. Large decision trees are susceptible to adjusting to statistical fluctuations in the training data. Consequently, stopping conditions are implemented, such as those determined by a node's training data yield.

Typically, a single small decision tree does not effectively utilize the available information. Multiple weak classifiers, such as small decision trees, are combined to create powerful classifiers through boosting. After each tree is constructed, adaptive boosting modifies the weights of the objects in the training data. The weights are increased for objects misclassified by the previous tree, which is why the next tree prioritizes them. The classification performance of each tree is used to determine the weighting. Gradient boosting employs a loss function, L , to evaluate the classification performance via

$$\sum_i L(y_i, f(\mathbf{x}_i)) \quad (4.19)$$

Where y_i is the label of object "i" (+1 for signal and -1 for background), $\mathbf{x}_i$ are that object's variable values, and $f(\mathbf{x}_i) \in [-1, 1]$ is the response of the BDT classifier.

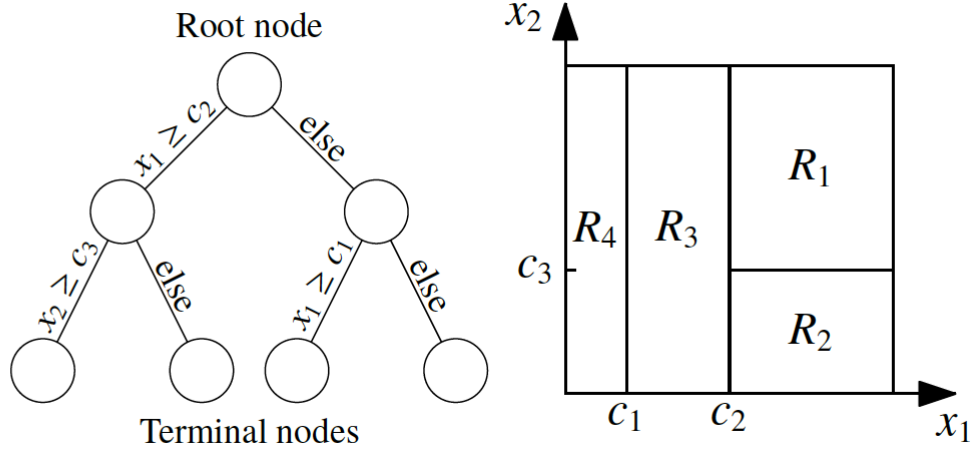


Figure 4.14: The parameter space spanned by the input variables (right) is divided by a decision tree (left).

Even though BDTs provide good discriminatory power, employing a Deep Neural Network (DNN) presents numerous benefits. DNN can autonomously acquire intricate, non-linear correlations within the data, enhancing the distinction between signal and background areas. They are especially adept at handling large datasets, utilizing their scalability to improve discrimination capability. Deep neural networks offer versatility by managing high-dimensional data and incorporating diverse feature types without extensive manual feature engineering. Furthermore, if trained adequately, they can generalize more effectively to novel data, potentially resulting in clearer signal regions with diminished background interference. The capacity to optimize via backpropagation and additional methods further improves the model's efficacy.

4.9.3 Deep Neural Networks

Previous studies of the $H \rightarrow WW^*$ decay used the m_T of the Higgs boson as a discriminant variable for the ggF channel to enhance the signal sensitivity and suppress the background events. This study uses the Deep Neural Network technique as discriminant variables to separate ggF, VBF, and background. Neural Networks have emerged as a crucial computational and statistical tool for analyses like the $H \rightarrow WW^*$ analysis that is driven by empirical data. The domain of neural networks encompasses a wide range of approaches but can be categorized into three primary learning paradigms: supervised, unsupervised, and reinforcement learning. The supervised learning uses predefined truth labels, which the network can verify against. In contrast, unsupervised learning lacks such labels, and the network's objective is to learn and extract features from the dataset rather than minimize the inaccuracy. Reinforcement learning lacks truth labels, but it distinguishes itself from unsupervised learning by implementing a reward system instead of relying on cross-checking. Supervised neural networks are employed in the current analysis. In supervised learning, the desired output is known in advance but blinded from the neural

network. Hence, the objective of the neural network is to predict the blinded variable of interest based on input variables that lack an apparent correlation to the output variable.

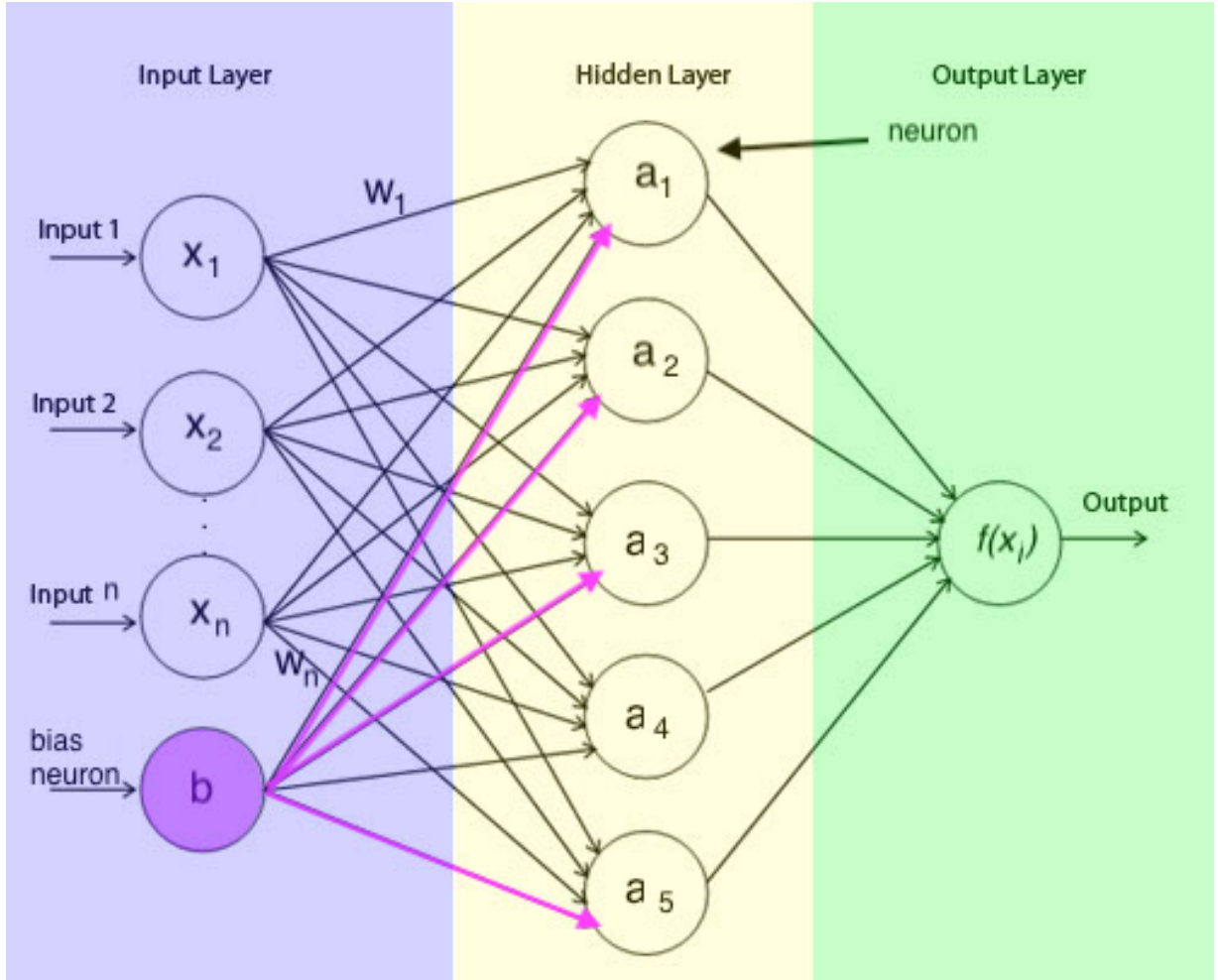


Figure 4.15: The architecture of a Deep Neural Network.

The architecture of the Deep Neural Network is illustrated in Figure 4.15, which consists of layers in which every neuron in a particular layer has a connection to every neuron in the previous layer. In a sequential layer structure, the initial layer is referred to as the input layer, which receives the input variables that the neural network needs to learn from. The final layer of the neural network outputs a variable or parameter of interest. There are one or more hidden layers between the input layer and the output layer, which aim to increase the learning capacity of the neural network by introducing additional independent variables. By giving a different weight w_i for each of the connections between the neurons, it is possible to attribute a value to each neuron as

$$y = f(z); z = \sum_i w_i x_i + b \quad (4.20)$$

where y is the output of the neuron, x_i are the neuron values in the previous layer (inputs), and b is the neuron bias. For a simple neural network such as this one, the trainable parameters are the weights and biases only. The neural network iteratively adjusts these parameters during the training process in order to minimize the discrepancy between its predictions and the target values.

Since the derivative of this neuron function remains constant, it is impossible to determine which weight in the input neurons can yield a more accurate prediction. Hence, the neurons' output must be non-linear to broaden the range of functions that the model can depict. The introduction of non-linearity is achieved by using activation functions in each neuron. Figure 4.16 displays several activation functions.

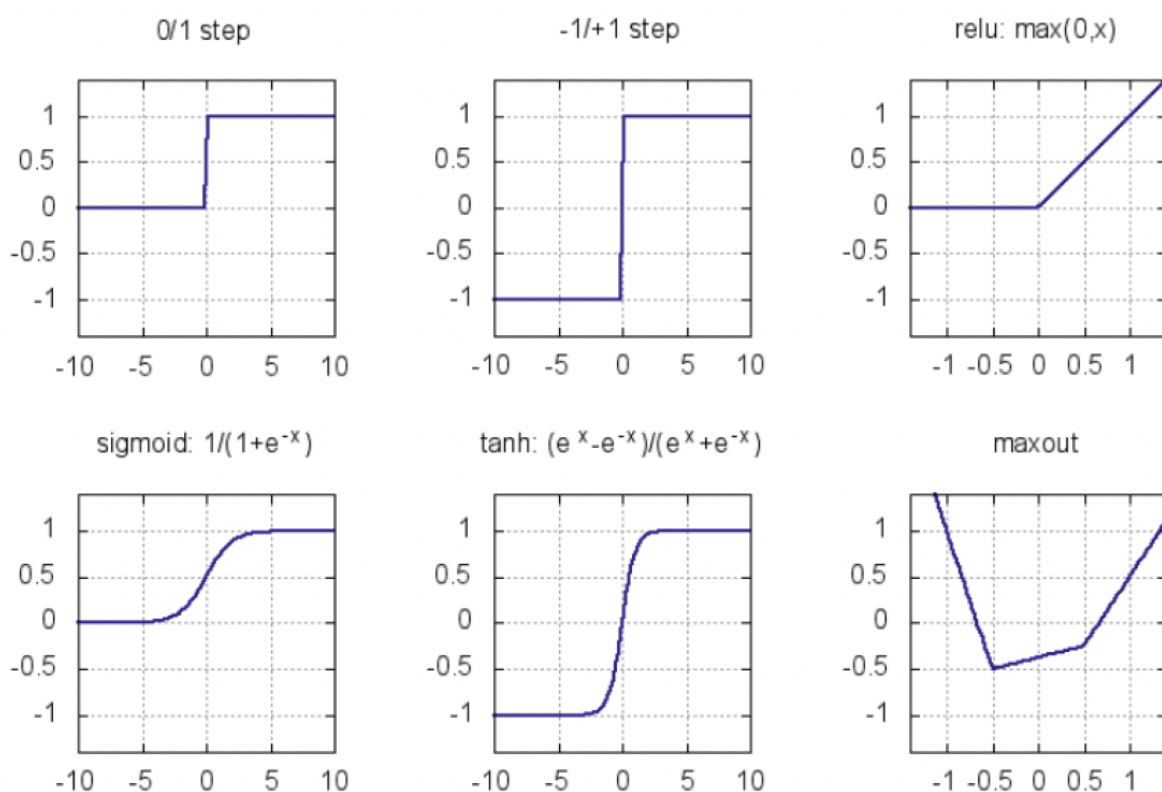


Figure 4.16: Example of activation functions.

In the initial phases of neural network development, the activation function was commonly represented as a Heaviside step function, meaning that the neuron operated in a binary manner of 0 or 1. However, this step function cannot produce multiple output values and causes the neuron to fire for any non-zero input, regardless of the signal strength. On the other hand, the Sigmoid and hyperbolic tangent activation functions have slight changes in prediction for extremely high or low values of x_i , which gives rise to the vanishing gradient problem. Rectified Linear Units (ReLU) activation is employed in the hidden

layers of the DNN model employed in this analysis, as it allows faster convergence. On the other hand, the output layer employs the softmax activation function with two outputs:

$$\text{Relu}(z) = \begin{cases} 0 & \text{if } z_i < 0 \\ z & \text{if } z_i > 0 \end{cases}, \quad (4.21)$$

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (4.22)$$

Here, $\{i, j\}$ represents the shape of the layer output and the shape of the layer input, respectively. Dropout layers were incorporated to mitigate the risk of overfitting by randomly scaling the non-zero inputs within the model by a ratio of $1/(1-\text{rate})$. Only the training set is impacted by the dropout layers; the validation and test sets are unaffected. An architecture of the DNN is shown in the appendix.

So, the DNN is a machine learning method composed of multiple non-linear operations in a series of fully connected layers to progressively extract higher-level features from the variable input to recognize data patterns. The DNN method can be used for classification analysis, categorizing a given dataset into different classes, or regression analysis, making forecasts based on the data. In this particular scenario, the DNN is employed to differentiate and categorize signal processes from background processes.

The DNN is constructed and trained using the KERAS interface and TensorFlow library [106, 107]. The DNN aims to differentiate the ggF processes from other events, both the background and the other production modes. The misidentified events are excluded because they are estimated using a data-driven approach. The DNN is trained using events that have undergone the b -jet veto. Each weight and bias must be trained sequentially from the input layer to the output layer in order to produce an accurate final result.

Each Monte Carlo event has an associated weight that is crucial in training the neural network. However, there is a challenge because NLO cross-section approximations lead to negatively weighted events. In order to prevent the training data from getting smaller, the absolute MC weight is applied rather than discarding these negative events. Another issue arises from the imbalanced weight sum between the ggF process and other processes, resulting in the DNN favoring the tags with higher weight sums. Training weights for ggF-tagged events and other processes are equated to fix this.

The variables shown in Table 4.9 are used as the input variables for the DNN. The same DNN training scheme is used in the 2jDF channel, with the same variables. p_T^l is not used to reduce dependency on mis-modelling. The training process exploits three spatial symmetries to reduce input complexity: rotational symmetry in the ϕ plane and mirror symmetry in the ϕ and η planes.

Type	#	Symbol	Description
Background rejection			
$t\bar{t}$ and $Z+$	1	p_T^{tot}	Total transverse momentum
jets	2	E_T^{Missing} Sig	Significance of the Missing Transverse Energy
Signal selection			
$H \rightarrow WW^*$			
	3	$\Delta\phi_{jj}$	Angle between dijets
	4	$\Delta\phi_{ll}$	Angle between dileptons
	5	m_{ll}	Dilepton mass, sum of two leptons
	6	m_T	Transverse mass
2jSF topology			
	7	Δy_{jj}	Difference in rapidity between jets
	8	m_{jj}	Dijet mass, sum of two jets
	9	$\eta_{l1}^{\text{Centrality}}$	Measure of how central lepton 1
	10	$\eta_{l2}^{\text{Centrality}}$	Measure of how central lepton 2
	11	$p_T(J_1)$	Transverse Momentum of Jet 1
	12	$p_T(J_2)$	Transverse Momentum of Jet 2
	13	$p_T(J_3)$	Transverse Momentum of Jet 3
	14	$m_{\ell 1, j1}$	Combined mass lepton 1 and jet 1
	15	$m_{\ell 1, j2}$	Combined mass lepton 1 and jet 2
	16	$m_{\ell 2, j1}$	Combined mass lepton 2 and jet 1
	17	$m_{\ell 2, j2}$	Combined mass lepton 2 and jet 2
	18	$\Delta R_{\ell 1, j1}$	Distance lepton 1 and jet 1
	19	$\Delta R_{\ell 1, j2}$	Distance lepton 1 and jet 2
	20	$\Delta R_{\ell 2, j1}$	Distance lepton 2 and jet 1
	21	$\Delta R_{\ell 2, j2}$	Distance lepton 2 and jet 2

Table 4.9: The input variables used for training the DNN models.

The performance of a machine learning algorithm is evaluated by a loss function ($\mathcal{L}$). The loss metric measures the level of discrepancy in the truth level quantities to the neural network output after each optimization iteration. The learning process is focused on optimizing this loss function by finding minima in its gradient or maximum likelihood fits.

If the connections in the neural network are made in the direction mentioned, this would represent a feed-forward neural network. The most popular method for training feed-forward neural networks is the back-propagation algorithm.

Each network weight is optimized in the Gradient Descent algorithms using a rule given by

$$w_i^t = w_i^{t-1} - \alpha \frac{\partial \mathcal{L}(w_i^{t-1})}{\partial w_i} \quad (4.23)$$

The analysis uses the Nesotrov Adaptive Moment (NADAM) optimizer, which utilizes adaptive learning rate and momentum-based techniques, to minimize the Binary Cross Entropy (BCE) loss function. BCE is represented in the following form:

$$\mathcal{L}_{BCE}(y^p, y^t) = \frac{1}{N} \sum_{i=1}^n [(y_i^t - 1) \log(1 - y_i^p) - y_i^t \log(y_i^p)] \quad (4.24)$$

where y^t and y^p represent the truth and predicted outputs of the neural network, respectively. The BCE loss function is a logarithm-based metric used to compare a pair of labels. The factors $(y_i^t - 1)$ and y_i^t guarantee a positive loss when $y_i^t = (0, 1)$ and the predicted value falls within the range of $y_i^p \in \{0, 1\}$.

Moreover, in the analysis, a bidirectional wrapper is employed with the Recurrent Neural Network (RNN) layers to address the problem of non-preferential ordering. A bidirectional wrapper processes the feature vector in both forward and backward temporal directions; thus, at each time t , the hidden state is expected to be cognizant of its past and future. The outputs from these two directions can be integrated through various straightforward operations such as addition, multiplication, and concatenation.

Splitting the data set

Memorization, over-fitting, and under-fitting are typical issues faced by neural networks. Memorization occurs when a neural network, rather than learning the phase space's features, begins duplicating the input features. Over-fitting and under-fitting are more general behaviors that occur when a neural network begins to oversimplify or overgeneralize the space's features. To mitigate the aforementioned concerns, the data set is split into three parts: Training, Validation, and Testing. A DNN with two output nodes is trained. A ggF node serves as the discriminant to perform the statistical analysis for the ggF SRs and a VBF node serves this purpose for the VBF SRs. The training incorporates events that meet the ggF same-flavor pre-selection criteria, along with the b-veto selection, and encompasses all significant backgrounds, excluding the data-driven W +jets estimation. The training set is utilized to optimize the network, while the validation set primarily assesses the optimized network's performance when confronted with foreign data. Ultimately, the testing set is employed to validate the generalization.

Dividing the dataset into multiple training, validation, and testing sets aids in validating network predictions; however, it may lead to suboptimal results due to incomplete training on the entire dataset. A solution is to cross-validate the network by augmenting it with the union of the non-overlapping sets, specifically the validation or test sets.

DNN Binning

The DNN binning for each signal region is optimized, and the methodology aligns with other channels. The signal events are uniformly distributed throughout the DNN spectrum, whereas the background events are concentrated in the lower DNN bins. The DNN binning for each SR is optimized and the method is consistent with other channels. The signal events are relatively evenly distributed across the DNN spectrum while the background events are clustered at low DNN bins. Table 4.10 delineates the categorization of Deep Neural Network (DNN) outputs into distinct bins according to production mode, transverse momentum (p_T^H) intervals, and dijet invariant mass (m_{jj}) intervals to differentiate signal regions in Higgs analysis. The table delineates the bin edges for DNN outputs, which categorize events into distinct signal regions according to the neural network’s prediction scores. The DNNs are partially unblinded in the low DNN regions to further verify the background modeling.

prod. mode	p_T^H range [GeV]	m_{jj} range [GeV]	DNN bin edges
ggF	[0, 200]	-	{0.00, 0.71, 0.81, 1.00}
	(200, 300]	-	{0.00, 0.70, 0.79, 0.85, 1.00}
	(300, ∞]	-	{0.00, 0.70, 0.79, 1.00}
VBF	[0, 200]	(350, 700]	{0.00, 0.52, 0.71, 1.00}
		(700, 1000]	{0.0, 0.68, 0.83, 1.00}
		(1000, 1500]	{0.00, 0.84, 1.00}
		(1500, ∞]	{0.00, 0.87, 0.94, 1.00}
VBF	(200, ∞]	(350, 1000]	{0.00, 0.57, 0.76, 1.00}
		(1000, 1500]	{0.00, 0.79, 0.94, 1.00}
		(1500, ∞]	{0.00, 0.85, 1.00}

Table 4.10: DNN bin edges for signal regions in the ggF and VBF 2j same flavor channel

Other DNN binning schemes were considered, but the above scheme (Table 4.10) was selected as it provided a better statistical sensitivity by optimizing how the data is grouped. The goal is to ensure that each bin has enough events for meaningful analysis and avoids bins with very low statistics that could lead to large uncertainties. It is tested that larger bin multiplicities or other binning methods do not lead to a noticeably ($> O(2\%)$) better sensitivity.

STXS Bins	0 – 200	200 – 300	Gt300
Coarse Binning	{61, 71, 77, 81, 85}	{70, 79, 85}	{77}
Even Binning	{56, 66, 72, 77, 80, 83, 86}	{64, 74, 79, 83, 86}	{69, 77, 84}
Odd Binning	{56, 66, 72, 77, 80, 83, 86}	{64, 74, 79, 83, 86}	{72, 82}

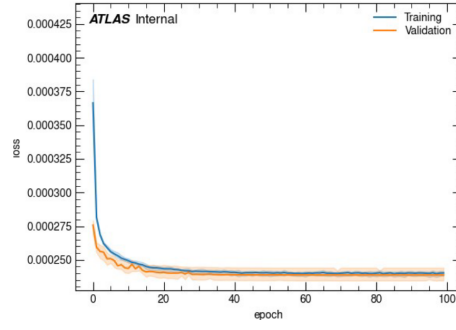
Table 4.11: Different DNN binning considered before choosing the above DNN binning scheme for ggF 2jSF

DNN binning is crucial for enhancing signal purity and statistical evaluation. The DNN allocates a score reflecting the probability of an event being a signal instead of background noise. Rather than considering the DNN output as a continuous variable, it is categorized into bins, facilitating enhanced control over the signal-to-background ratio within each bin. The DNN bin values are established using statistical methods and physical considerations to enhance the signal-to-background separation while preserving sensitivity across different regions of the DNN output. To further verify the modeling, several kinematic distributions are plotted for the inclusive partially unblinded ggF and VBF regions shown in the Appendix.

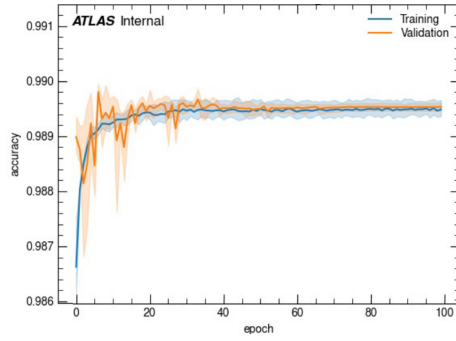
Results

Figure 4.17 illustrates the Binary Cross-Entropy (BCE) loss curve for both the 2-jet Deep Neural Networks (DNNs).

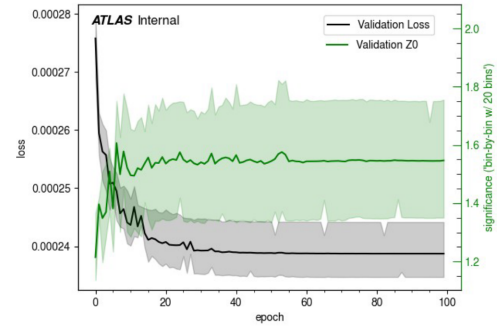
The plots illustrate the BCE loss and accuracy progression for the training and validation datasets across 100 epochs. The training and validation curves for the BCE loss plot exhibit a similar trajectory, characterized by a pronounced initial decline followed by stabilization at a minimal loss value (≈ 0.0002). Similarly, both curves for the accuracy plot stabilize at approximately 0.989 for each set. Minor fluctuations in validation accuracy occur due to dataset variations and stochastic behavior during training. The convergence of these curves further indicates that the model is well-regularized and not overfitting.



(a)



(b)



(c)

Figure 4.17: a) the BCE loss for 2-jet DNN, b) Accuracy for the 2-jet DNN, and c) loss vs significance plot

5 | Systematic Uncertainties

Reconstructing the original energy and momenta of all particles in an event with absolute precision is an unattainable task. Hence, the measurements are subject to uncertainty due to various sources. There are two distinct categories in which these uncertainties can be classified: a statistical component and a systematic component. The first category of uncertainties emerges from fluctuation in the observed dataset and expected event yields due to limited sample sizes in Monte Carlo simulations. The second category of uncertainties arises from the reconstruction and calibration of objects within the detector (experimental uncertainties) and the modeling of signal and background processes in the simulation (theoretical uncertainties). The estimation of systematic uncertainties is performed in all analysis regions. This chapter provides an overview of the statistical and systematic uncertainties that impact the analysis. Sections 5.1 and 5.2 outline the experimental and theoretical uncertainties, respectively, whereas Section 5.3 outlines the statistical uncertainties.

5.1. Systematic Uncertainties

As mentioned previously, systematic uncertainties can be divided into two categories: uncertainties related to detector and reconstruction effects and uncertainties related to theoretical predictions of the signal and main backgrounds.

5.1.1 Experimental Uncertainties

Experimental uncertainties pertain to the inherent limitations in accurately representing the response of a detector and accurately reconstructing and calibrating objects with absolute precision. The uncertainties are derived by dedicated performance groups not affiliated with this specific analysis and encompass a wide range of sources. The performance groups' uncertainties are transmitted to each MC-simulated sample. These variations can potentially alter an object's four-momentum, resulting in the movement of events between different regions and bins. Additionally, they can also modify the efficiency scale factor of an event or object. Both scenarios will result in changes in the expected acceptance of events. The $\pm 1\sigma$ variation of the uncertainty source is considered to be the difference between the nominal yield and the varied yield caused by a single source of experimental uncertainty.

The complete list of the experimental uncertainties considered in the analysis is shown in Table 5.1. The systematic uncertainties for electrons and muons are derived separately. The electron-related systematics include uncertainties on their reconstruction, identification, isolation, trigger efficiencies, and associated energy scale and resolution.

The muon-related systematics are similar, except with an additional uncertainty on the efficiency of track-to-vertex association (TTVA) impact parameter cuts.

The systematic uncertainties on the jet energy scale, which are commonly measured as a function of p_T and η , occur during the calibration of the jet energy scale (JES) and jet energy resolution (JER). Each step of the JES calibration involves various sources of uncertainty, such as pileup removal, in-situ calibration, η -intercalibration, flavor composition, and punch-through jets. The uncertainties in the JER can be attributed to the various in-situ methods employed to calculate it.

The reconstruction of E_T^{miss} and p_T^{miss} is subject to systematic uncertainties that stem from both the hard objects and the soft term. The calculation of E_T^{miss} and p_T^{miss} takes into account the calibrated electron, muon, and jet objects as inputs. Consequently, the systematic uncertainties associated with these objects are taken into account in the calculation of E_T^{miss} and p_T^{miss} . The systematic uncertainties on the soft term are calculated by utilizing $Z \rightarrow \mu\mu$ events and the variable $\vec{p}_T^{\text{hard}}$, which represents the sum of the hard objects in the vector p_T . The scale and resolution of the soft term are influenced by uncertainties that are associated with both the longitudinal and transverse components of this quantity.

The total integrated luminosity uncertainty for the full Run 2 data set is estimated to be 1.7%. The uncertainty in luminosity measurement is determined by employing a specialized detector known as LUCID-2.

Systematic uncertainty	Short description	syst. type
Event		
Luminosity	uncertainty on total integrated luminosity	SF
Pileup Reweighting	uncertainty on pileup reweighting	SF
Electrons		
EL EFF Reco Total INPCOR PLUS UNCOR	reconstruction efficiency uncertainty	SF
EL EFF ID CorrUncertaintyNP (0 to 14)	ID efficiency uncertainty split in 15 components	SF
EL EFF ID SIMPLIFIED UncorrUncertaintyNP (0 to 15)	ID efficiency uncertainty split in 16 components	SF
EL EFF Iso Total INPCOR PLUS UNCOR	isolation efficiency uncertainty	SF
EL EFF Trigger Total INPCOR PLUS UNCOR	trigger efficiency uncertainty	SF
EG SCALE ALLCORR		P4
EG SCALE E4SCINTILLATOR		P4
EG SCALE LARTEMPERATURE EXTRA2015PRE	energy scale uncertainty	P4
EG SCALE LARTEMPERATURE EXTRA2016PRE		P4
EG SCALE LARCALIB EXTRA2015PRE		P4
EG RESOLUTION ALL	energy resolution uncertainty	P4
Muons		
MUON EFF STAT		SF
MUON EFF SYS	reconstruction and ID efficiency uncertainty for muons with $p_T > 15$ GeV	SF
MUON ISO STAT		SF
MUON ISO SYS	isolation efficiency uncertainty	SF
MUON TTVA STAT		SF
MUON TTVA SYS	track-to-vertex association efficiency uncertainty	SF
MUON EFF TrigStatUncertainty		SF
MUON EFF TrigSystUncertainty	trigger efficiency uncertainty	SF
MUON SCALE	momentum scale uncertainty	P4
MUON ID	momentum resolution uncertainty from inner detector	P4
MUON MS	momentum resolution uncertainty from muon system	P4
Jets		
JET 21NP JET EffectiveNP 1		P4
JET 21NP JET EffectiveNP 2		P4
JET 21NP JET EffectiveNP 3		P4
JET 21NP JET EffectiveNP 4	energy scale uncertainty from the in situ analyses split into eight components	P4
JET 21NP JET EffectiveNP 5	energy scale uncertainty from the in situ analyses split into components	P4
JET 21NP JET EffectiveNP 6		P4
JET 21NP JET EffectiveNP 7		P4
JET 21NP JET EffectiveNP 8stestTerm		P4
JET 21NP JET EtaIntercalibration Modeling	energy scale uncertainty on eta-intercalibration (modeling)	P4
JET 21NP JET EtaIntercalibration TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)	P4
JET 21NP JET EtaIntercalibration NonClosure	energy scale uncertainty on eta-intercalibrations (non-closure)	P4
JET 21NP JET Pileup OffsetMu	energy scale uncertainty on pileup (mu dependent)	P4
JET 21NP JET Pileup OffsetNPV	energy scale uncertainty on pileup (NPV dependent)	P4
JET 21NP JET Pileup PtTerm	energy scale uncertainty on pileup (pt term)	P4
JET 21NP JET Pileup RhoTopology	energy scale uncertainty on pileup (density ρ)	P4
JET 21NP JET Flavor Composition	energy scale uncertainty on flavor composition	P4
JET 21NP JET Flavor Response	energy scale uncertainty on sample flavor response	P4
JET 21NP JET BJES Response	energy scale uncertainty on b -jets	P4
JET 21NP JET PunchThrough MC15	energy scale uncertainty for punch-through jets	P4
JET 21NP JET SingleParticle HighPt	energy scale uncertainty from the behaviour of high- p_T jets	P4
JET JER SINGLE NP	energy resolution uncertainty	P4
JET JvtEfficiency	JVT efficiency uncertainty	SF
FT EFF Eigen B	b -tagging efficiency uncertainties ("BTAG MEDIUM"): 3	SF
FT EFF Eigen C	b -tagging efficiency uncertainties ("BTAG MEDIUM"): 3	SF
FT EFF Eigen L	components for b jets, 4 for c jets and 5 for light jets	SF
FT EFF Eigen extrapolation	b -tagging efficiency uncertainty on the extrapolation to high- p_T jets	SF
FT EFF Eigen extrapolation from charm	b -tagging efficiency uncertainty on tau jets	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 5.1: Summary of the experimental systematic uncertainty sources considered in the analysis

5.1.2 Theoretical Uncertainties

The theoretical uncertainties arise from the inherent challenge of accurately predicting and modeling the signal and background processes, thereby introducing such uncertainties into the measurements. Determining theoretical uncertainties typically relies on the absolute expected yields of a process for the most significant processes across all analysis regions and on how the sample is normalized. Regarding signals, these uncertainties also include the migrations between the signal regions. The uncertainties for background processes normalized using dedicated control regions are derived from the variations observed in the extrapolation from the control to signal regions. The uncertainty factors considered for each process encompass variations in the QCD scale, uncertainties in modeling the parton shower and underlying event (PS/UE), and variations in the PDF set. Table 5.2 provides a summary of the theoretical uncertainty sources that are taken into account in this analysis, along with a description of their estimation and approximate values.

Estimating theoretical uncertainty associated with the absence of higher orders in perturbative calculations (QCD scale uncertainty) is achieved by adjusting the renormalization and factorization scales during the process. The generation of theoretical predictions in PDF sets is subject to uncertainty from multiple sources. It is necessary to propagate experimental uncertainty in the fit of the PDF and the uncertainty arising from the assumed functional form of the PDF through measurements.

Typically, distinct software applications are employed to compute the hard scatter and the modeling of the UEPS. Uncertainty arises when two programs are interfaced. One potential source of uncertainty arises when attempting to eliminate the double-counting of jets during the merging of events with different jet multiplicities. Another uncertainty arises when attempting to interface an NLO or NNLO calculation with a parton shower. The uncertainties associated with this source are generally determined by fixing the program used for modeling the UEPS while changing the software employed for calculating the hard scatter.

Uncertainty also arises from implementing different algorithms by PYTHIA, HERWIG, and SHERPA in the parton shower and hadronization model (UEPS). Typically, UEPS uncertainty is determined through a hard scatter calculation.

process	uncertainty	Updated	Already included in statistical analysis for		
			0/1j DF	2 j DF	2 j SF
ggF	QCD scale	No	Yes	Yes	Yes
	matching	Yes	n/a	Yes	Yes
	PS/UE	Yes	Yes	Yes	Yes
	PDF+ α_s	No	Yes	Yes	Yes
VBF	QCD scale	No	Yes	Yes	Yes
	matching	Yes	Yes	Yes	Yes
	PS/UE	Yes	Yes	Yes	Yes
	PDF + α_s	No	Yes	Yes	Yes
Other Higgs	normalization	No	Yes	Yes	Yes
$qqWW$ (non-EW)	matching (CKKW)	No	Yes	Yes	Yes
	PS (QSF, CSSKIN)	No	Yes	Yes	Yes
	PDF + α_s	Yes	Yes	Yes	Yes
	QCD scale (7-point)	Yes	Yes	Yes	Yes
	EW correction	New	Yes	Yes	Yes
$ggWW$ (non- H)	$gg \rightarrow WW$ scale	No	Yes	Yes	Yes
EW WW	QCD scale	No	Yes	Yes	Yes
	PDF + α_s	No	Yes	Yes	Yes
	Shower	No	Yes	Yes	Yes
	EW correction	No	Yes	Yes	Yes
top	QCD scale (7-point)	No	Yes	Yes	Yes
	radiation (ISR/FSR)	No	Yes	Yes	Yes
	PDF+ α_s	No	Yes	Yes	Yes
	$t\bar{t}$ matching (pThard)	Yes	Yes	Yes	Yes
	$t\bar{t}$ PS/UE (Py8 vs H7)	No	Yes	Yes	Yes
	$t\bar{t}$ reweighting	New	Yes	Yes	Yes
	Wt matching (pThard)	Yes	Yes	Yes	Yes
	Wt interference (DS vs DR)	Yes	Yes	Yes	Yes
	Wt hdamp (dyn. scale)	New	Yes	Yes	Yes
	Wt PS/UE (Py8 vs H7)	Yes	Yes	Yes	Yes
Z+jets	QCD scale	Yes	Yes	Yes	Yes
	PDF+ α_s	Yes	Yes	Yes	Yes
	CKKW and QSF	Yes	Yes	Yes	Yes
	EW correction	New	Yes	Yes	Yes
	p_T spectrum	New	Yes	Yes	Yes
WZ/ZZ	QCD scale, PDF and α_s	Yes	Yes	Yes	Yes
	CKKW, QSF, CSSKIN	Yes	Yes	Yes	Yes
	EW correction	New	Yes	Yes	Yes
$V \gamma$	QCD scale, PDF and α_s	New	Yes	Yes	Yes
	CKKW, QSF, CSSKIN (from WZ)	New	Yes	Yes	Yes
triboson	Generic overall norm	No	Yes	Yes	Yes

Table 5.2: Summary of the theoretical systematic uncertainties considered in the analysis.

5.2. Statistical Uncertainties

The data and expected signal and background events in the SRs and CRs are subjected to a statistical analysis using the formalism of maximum likelihood estimation, as discussed in the previous chapter. This approach enables the formulation of meaningful statements and inferences regarding the SM. The analysis is conducted by developing a statistical model using the HISTFITTER software [108], which is derived from the ROOFIT software. The statistical model in this analysis is based on a binned likelihood function, which is defined by a set of parameters, $\alpha = (\mu, \theta)$, where μ represents signal strengths (parameter of interest) and θ represents a vector of nuisance parameters (NPs).

The likelihood function, as described in Chapter 4, is a statistically robust way that is implemented in this analysis. The expected yields of the signal and background processes (N) are defined in relation to an NP, $N = N(\theta_t)$, associated with a systematic uncertainty. The likelihood function includes various systematic uncertainties, resulting in a parameterization that can be expressed as follows:

$$N(\theta) = N_0 \times \prod_t \nu(\theta_t) \quad (5.1)$$

where N_0 represents the nominal expected yield of the process, and $\nu(\theta_t)$ is a function of θ_t . The exact form in which $\nu(\theta_t)$ is parameterized depends upon the systematic uncertainty linked to the NP, as well as whether an NP alters the yields of a process, the shape of the discriminating variable, or both. The computation involves determining the event yields that correspond to the $\pm 1\sigma$ variations of each systematic uncertainty.

A likelihood fit is initially conducted using an Asimov dataset, which is generated by setting $\mu = 1$. Here, the Asimov dataset is performed for each channel, each with simultaneous extraction of signal strengths (μ) and background normalization factors (NF). This test aims to assess the expected significance of the ggF signal and the expected uncertainties associated with its strength. The anticipated significance of the ggF signal is determined to be $Z = 5.3$. Subsequently, a likelihood test is conducted utilizing the complete observed dataset.

The estimation of the POIs and NPs is determined by maximizing the likelihood function, which is constructed corresponding to the setup. Then, the HISTFITTER minimizes the negative log-likelihood function, $-\ln \mathcal{L}(\mu, \theta)$. Here, the likelihood is parametrised by signal strength (μ)/ POI and $NPs(N(\tilde{\theta} | \theta))$ and is given by:

$$\mathcal{L}(\mu_i, \mu_j, \mu_b) = \prod_{\text{SRskCRs bins}} P\left(N \mid \mu_j \text{ S}_{i,\text{ggF}} + \mu_j \text{ S}_j, \text{VBF} + \sum_{\text{backgrounds}} \mu_k b_k\right) \times \prod_{i=1}^{N_\theta} N(\tilde{\theta} \mid \theta) \quad (5.2)$$

where $P(N \mid v)$ is the Poisson distribution, μ_k are the background normalization factors

(NF), b_k the expected background yields, $s_{i,\text{ggF}}$ and $s_{j,\text{VBF}}$ are the expected ggF and VBF yields per STXS bin, and μ_i and μ_j are the signal strength parameters per STXS bin. For backgrounds other than those normalized from control regions, $\mu_k = 1$. Some of the signal strength parameters are fixed to $\mu_i = 1$ when performing channel-exclusive fits. For example, the parameters representing ggF and VBF signal for ≥ 2 jets are fixed.

This minimizing of log-likelihood is commonly known as the fit, which is computationally simpler. The parameters in the likelihood function that yield values and uncertainties are commonly known as best-fit or post-fit.

5.3. ggF Statistical Treatment

The statistical methodology employed to analyze and integrate the results of each category has been formulated by the ATLAS and CMS Collaborations within the framework of the LHC Higgs Combination Group. Two hypotheses are evaluated to identify a new signal process: the null hypothesis, which posits only background, and the alternative hypothesis, which includes both signal and background.

The anticipated quantity of occurrences in each category within each bin can be articulated as

$$n_i = \mu \cdot s_i + b_i \quad (5.3)$$

where s_i denotes the anticipated signal yield, b_i represents the expected background yield, and μ_i signifies the signal strength to be quantified. If $\mu = 0$, the expectation aligns with the background-only hypothesis, whereas $\mu = 1$ corresponds to the nominal signal hypothesis.

A Poisson distribution represents the anticipated yield in each bin. The overall probability of observing n_i events for each bin i is the product of the respective Poisson probabilities.

$$\text{Poisson}(n_i | \mu \cdot s + b) = \prod_i \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \quad (5.4)$$

The systematic uncertainties are denoted by distinct nuisance parameters θ . Each nuisance is characterized by a prior distribution function $p(\hat{\theta} | \theta)$, where $\hat{\theta}$ denotes the corresponding default value of each nuisance parameter. Log-normal distributions characterize the predominant nuisance parameters, whereas flat priors were employed for the statistical uncertainties arising from the Top and WW background measurements in data control regions. Furthermore, the correlation among systematic uncertainties is considered. The likelihood function is defined by

$$\mathcal{L}(\text{data} \mid \mu, \theta) = \text{Poisson}(\text{data} \mid \mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta} \mid \theta) \quad (5.5)$$

The test statistic q_μ is employed to evaluate a proposed μ value derived from the profile likelihood ratio.

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data} \mid \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} \mid \hat{\mu}, \hat{\theta})} \quad (5.6)$$

The likelihood function for the fit is primarily formulated as the product of Poisson terms, corresponding to each bin in all signal regions and each control region. Furthermore, systematic uncertainties are considered through $\pm 1\sigma$ variations when feasible and are represented as nuisance parameters, each accompanied by a specific constraint term.

The signal and background yields can fluctuate as functions of the nuisance parameters $\vec{\theta}$. They are specifically parameterized as $s = s_0 \times \prod \nu(\theta)$, where the response to each nuisance parameter is separated from the nominally anticipated rate. Four primary sources of uncertainty are identified, each characterized by a distinct form for $\nu(\theta)$: a flat rate source, a shape source, a statistical source, and a normalization source. Flat rate sources are expressed as $\nu_{\text{flat}}(\theta) = \kappa^\theta$, where κ is ascertained by evaluating ν_{flat} at $\theta = \pm 1$. In this instance, the constraint term $C(\theta)$ is a standard Gaussian. Systematics that influence rate and shape are initially divided into a distinct flat component and a pure shape component, ensuring that alterations to the pure shape component do not impact the overall expected rate. The pure shape component employs vertical linear interpolation to assess variation, with its constraint term represented by a truncated Gaussian. The statistical uncertainties are represented by a Poisson constraint $P(\tilde{\theta} \mid \theta\lambda)$, indicating an auxiliary measured event count $\tilde{\theta}$ with an anticipated count of $\theta\lambda$. The constraint term for a normalization source parallels that of a statistical source, with the distinction that the observed and expected event counts in the pertinent control region are utilized.

The two categories of experimental systematics outlined in section 5.1 can influence the fit in fundamentally distinct manners. The systematics of the efficiency scale factor solely alter the weight of each event, maintaining a complete correlation with the nominal yields, whereas the four-momentum systematics can result in events transitioning into or out of a signal or control region. In Monte Carlo events with low statistics, this may also result in artificially large and unphysical fluctuations. To enhance fit stability and decrease convergence time, a procedure known as "pruning" is employed, whereby certain sufficiently minor uncertainties are disregarded. Initially, rate uncertainties are excluded from the fit for a specific sample in a region if its variation is less than 0.5%. Subsequently, any shape uncertainty for which no individual bin in a region exceeds 0.5% is eliminated from the fit.

The results in this thesis for the comprehensive Run II analysis of the $H \rightarrow WW$ decay are initially blinded; that is, the fit is conducted without examining the data in

the signal region to avoid biasing the analysis strategy by the data. Consequently, a likelihood fit is initially conducted using an Asimov dataset created by setting $\mu = 1$ to assess the anticipated ggF signal significance and the expected uncertainties regarding the ggF signal strength. This is a straightforward approach to assessing signal sensitivity before data examination.

Subsequently, a likelihood fit will be conducted utilizing the complete observed dataset. Figure 5.1 presents the ranking of the nuisance parameter's influence on the signal strength, and Figure 5.2 illustrates the correlations among the nuisance parameters using the Asimov dataset.

Estimating Analysis Sensitivity with an Asimov dataset

The fits are studied using the Asimov dataset approach, which replaces Monte Carlo pseudo-experiments with a single representative dataset generated under nominal parameter values. This method provides insights into sensitivity, constraints on nuisance parameters, and correlations between fit parameters. In these fits, constrained nuisance parameters should be near 0, and floating normalizations, including statistical uncertainty parameters, should be close to 1.

Estimating effects of systematic uncertainties

This analysis quantifies the effects of systematic uncertainties using several methods. Unless stated otherwise, the $\pm 1\sigma$ variations of nuisance parameters are considered. "Variation" refers to changes in event yield within a bin, typically evaluated before fitting to understand the statistical model.

Two key methods assess how nuisance parameters affect the parameter(s) of interest: *breakdown* and *impact*. The *breakdown* method evaluates the combined effect of an uncertainty source by comparing fits with and without that source, using quadratic subtraction to measure its influence. The *impact* method measures the effect of individual nuisance parameters by observing changes in the parameter(s) of interest when fixing a parameter to its best-fit value plus or minus its uncertainty. While *breakdowns* represent combined effects and *impacts* provide granular, directional insights, both are used for ranking nuisance parameters and may differ in cases of significant parameter correlations.

An Asimov fit is performed by combining the two ggF and VBF $N_{\text{jet}} \geq 2$ same-flavor channels to test the expected signal strength for each of the STXS POIs. The fit is μ -style. The fit results summary is shown in Table 5.3. A summary of the fit normalization factors including their uncertainties is also shown in the table. The parameter correlations of $\geq 20\%$ in Figure 5.2 and the detailed nuisance parameter impacts in Figure 5.1.

prod. mode	p_T^H range [GeV]	m_{jj} range [GeV]	stat+sys error
ggF	[0, 200]		$1.00^{+3.76}_{-3.67}$
	(200, 300]		$1.00^{+2.32}_{-2.11}$
	(300, ∞)	-	$1.00^{+3.45}_{-3.62}$
sig prod. mode	background	p_T^H range	stat+sys error
ggF	Zll	[0,200]	$1.00^{+0.40}_{-0.26}$
		(200, ∞)	$1.00^{+0.48}_{-0.28}$

Table 5.3: Asimov fit results for signal strengths and background normalization factors in the ggF 2j same flavor channel.

Moreover, the current result of the fit in terms of normalized cross-section values and their uncertainties are shown or different configurations are shown in Table 5.4 and Figure 5.3 with all systematic uncertainties included. The figure shows how including the same flavor channel has impacted the signal strength. By comparing the combined fit of the 2-jet different-flavor channel and the same-flavor channel (in black bars) to the 2-jet different-flavor channel only (in red bars), the improvement in the sensitivity by including the same-flavor channel can be observed. Especially at high p_T^H regions, the sensitivity of SF channels is visible. Similarly looking at the tabular format, the uncertainty has dropped significantly. Further work is being conducted on the fits to include the observed fit and compare it with the Asimov fits.

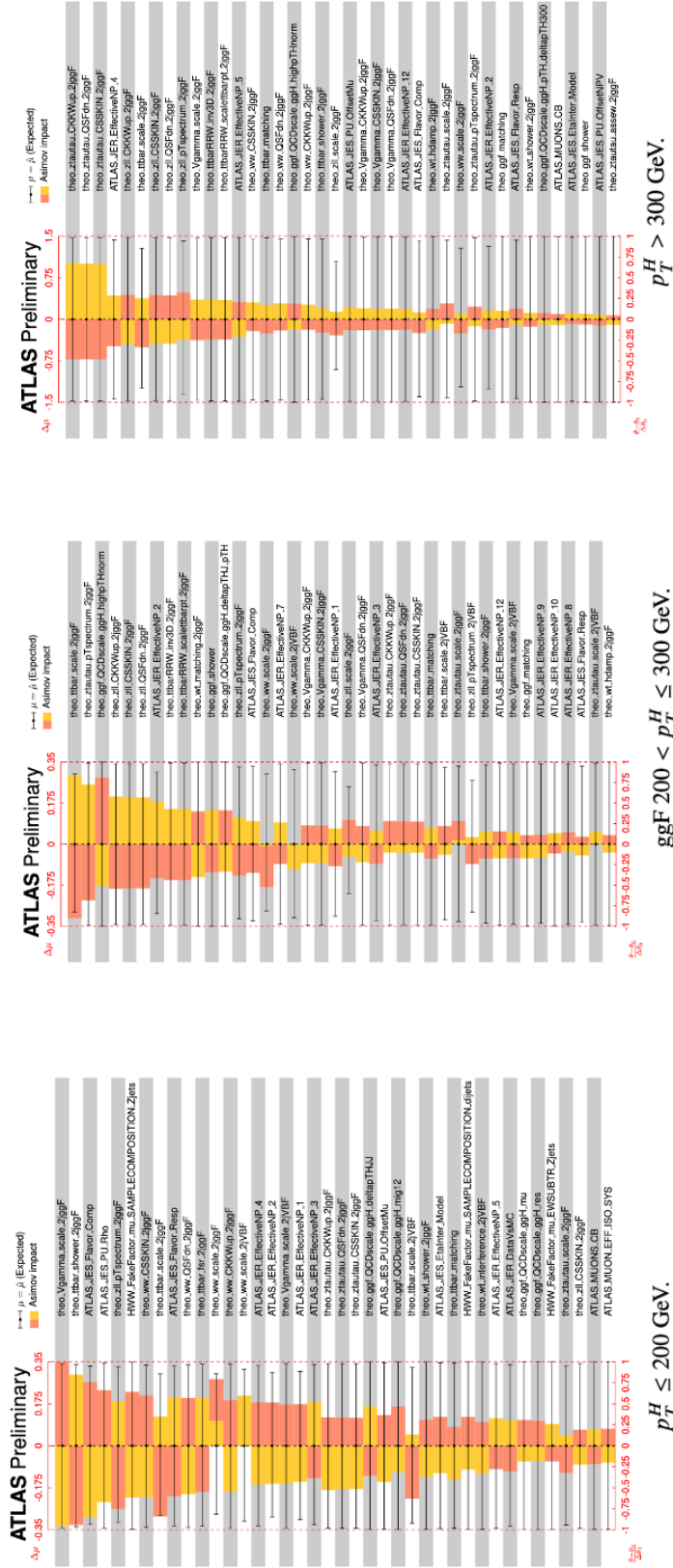
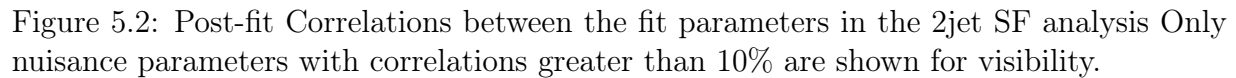


Figure 5.1: An analysis of the ranking distribution of nuisance parameters in the ggF fit with an Asimov dataset. The yellow and red bands account for their contribution to the total uncertainty in the analysis. This contribution is calculated as the quadratic difference between the uncertainty on μ in the main fit with all nuisance parameters and a fit where the nuisance parameter in question has been set to its best-fit value.



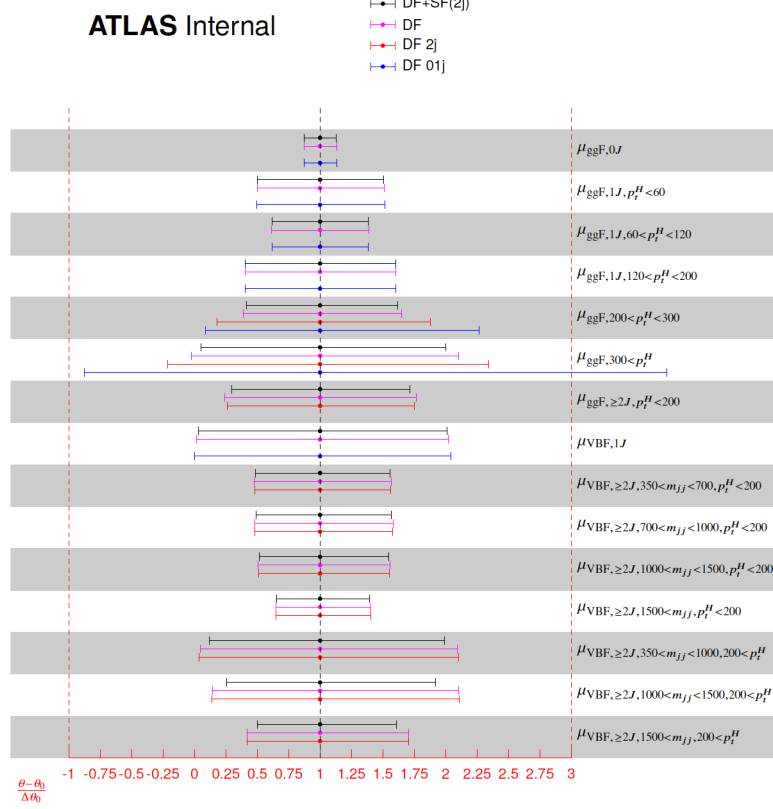


Figure 5.3: Expected sensitivities for the parameters of interest for different configurations.

Contribution	Uncertainty		Uncertainty		Uncertainty		Uncertainty	
	DF 01j		DF 2j		DF		DF+SF(2)	
$\mu_{\text{ggF},0J}$	-0.1303	0.1349			-0.1297	0.1341	-0.1291	0.1332
$\mu_{\text{ggF},1J,p_t^H < 60}$	-0.5081	0.5181			-0.5016	0.5110	-0.4988	0.5071
$\mu_{\text{ggF},1J,60 < p_t^H < 120}$	-0.3851	0.3857			-0.3896	0.3903	-0.3826	0.3853
$\mu_{\text{ggF},1J,120 < p_t^H < 200}$	-0.5993	0.6053			-0.6000	0.6064	-0.5966	0.6033
$\mu_{\text{ggF},200 < p_t^H < 300}$	-0.9117	1.2696	-0.8213	0.8819	-0.6125	0.6520	-0.5891	0.6216
$\mu_{\text{ggF},300 < p_t^H}$	-1.8759	2.7631	-1.2172	1.3454	-1.0271	1.1044	-0.9481	1.0017
$\mu_{\text{ggF},\geq 2J,p_t^H < 200}$			-0.7401	0.7509	-0.7594	0.7669	-0.7062	0.7150
$\mu_{\text{VBF},1J}$	-1.0028	1.0436			-0.9834	1.0239	-0.9723	1.0113
$\mu_{\text{VBF},\geq 2J,350 < m_{jj} < 700,p_t^H < 200}$			-0.5233	0.5659	-0.5282	0.5701	-0.5163	0.5574
$\mu_{\text{VBF},\geq 2J,700 < m_{jj} < 1000,p_t^H < 200}$			-0.5195	0.5814	-0.5217	0.5834	-0.5089	0.5675
$\mu_{\text{VBF},\geq 2J,1000 < m_{jj} < 1500,p_t^H < 200}$			-0.4926	0.5560	-0.4940	0.5580	-0.4860	0.5469
$\mu_{\text{VBF},\geq 2J,1500 < m_{jj},p_t^H < 200}$			-0.3518	0.4044	-0.3547	0.4071	-0.3501	0.3980
$\mu_{\text{VBF},\geq 2J,350 < m_{jj} < 1000,200 < p_t^H}$			-0.9637	1.1046	-0.9541	1.0960	-0.8831	0.9908
$\mu_{\text{VBF},\geq 2J,1000 < m_{jj} < 1500,200 < p_t^H}$			-0.8629	1.1121	-0.8563	1.1037	-0.7458	0.9237
$\mu_{\text{VBF},\geq 2J,1500 < m_{jj},200 < p_t^H}$			-0.5843	0.7079	-0.5830	0.7068	-0.5001	0.6105

Table 5.4: Expected sensitivities for the parameters of interest for different configurations (table form).

5.4. Zl mis-modeling of study

SHERPA Z +jets has been known to show mis-modeling in certain ATLAS analyses, especially in regions with higher jet multiplicities or high transverse momenta. Studies focusing on various kinematic distributions and observables have observed discrepancies between SHERPA’s predictions and data. These discrepancies are often linked to SHERPA’s treatment of QCD effects, including parton showers and hadronization, with the uncertainty between SHERPA 2.2.X and 2.2.1 generally assumed to account for them. More information is provided in Appendix B. Although SHERPA is a robust tool, its constraints in these domains may occasionally result in inaccuracies when predicting certain aspects of Z +jets events. The shape of p_T^{ll} in the distribution shown in Figure 5.4 clearly suggests that Z jets MC has mis-modeling.

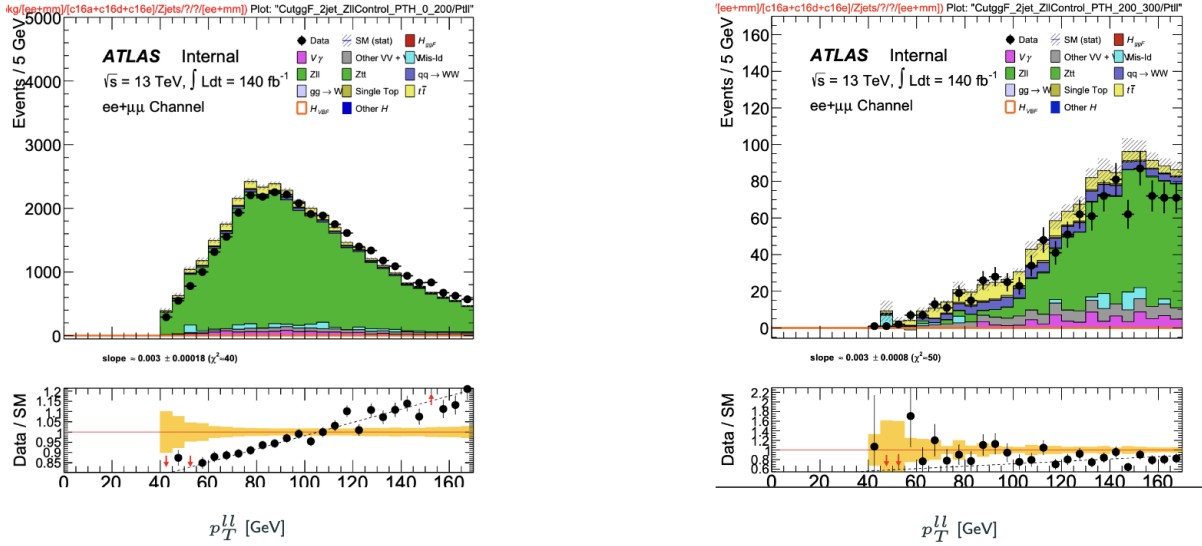


Figure 5.4: Distributions of p_T^{ll} in Z_{ll} CR Z +jets in 2-jet SF channels.

Hence, the correlation between p_T^{ll} in Z_{ll} CR and DNN was investigated (Figure 5.5) to assess the model’s effectiveness in minimizing sensitivity to the SHERPA Z_{ll} mis-modeling. Correlation functions of -0.1637 for ggF 0-200 p_T^H , 0.03333 for ggF 200-300 p_T^H , and 0.1349 for ggF GT300 p_T^H suggest that there is very weak correlation. Hence, we can confirm that mis-modeling in p_T^{ll} is irrelevant to the DNN modeling. Further, 1D plots of p_T^{ll} for each DNN bin and 1D plots of the DNN for each p_T^{ll} bin (Figure 5.6 and Figure 5.7) were plotted to see the nuances and confirm the claim.

A few mitigating strategies have already been implemented, such as using $p_T^{ll} > 40\text{GeV}$ cut on 2jSF channel, shown in Appendix B, additional Z jets p_T shape theory systematics (comparison of nominal Z jets sample with the old Sherpa2.2.1 Z jets sample shows that this covers the mis-modeling) and checks for background mis-modeling in CRs for both the DNN and kinematic distributions.

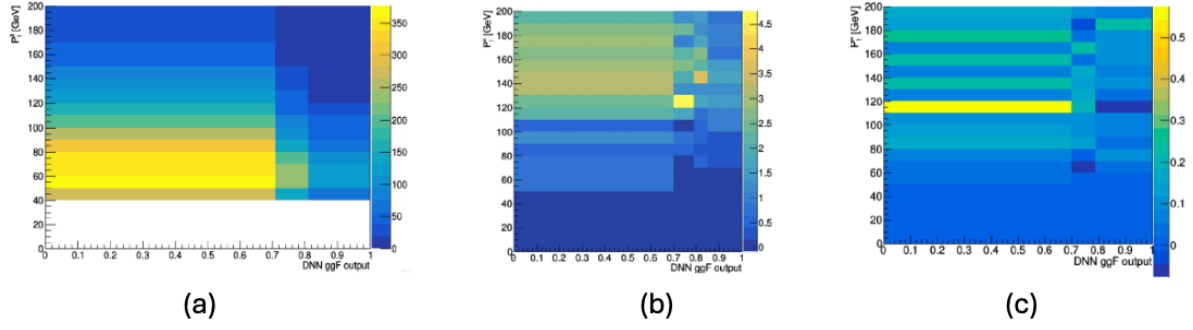


Figure 5.5: Heat-maps showing the correlation between the DNN ggF output and the transverse momentum (p_t^H) for (a) ggF 0-200 GeV p_T^H , (b) ggF 200-300 GeV p_T^H , and (c) ggF GT300 GeV p_T^H .

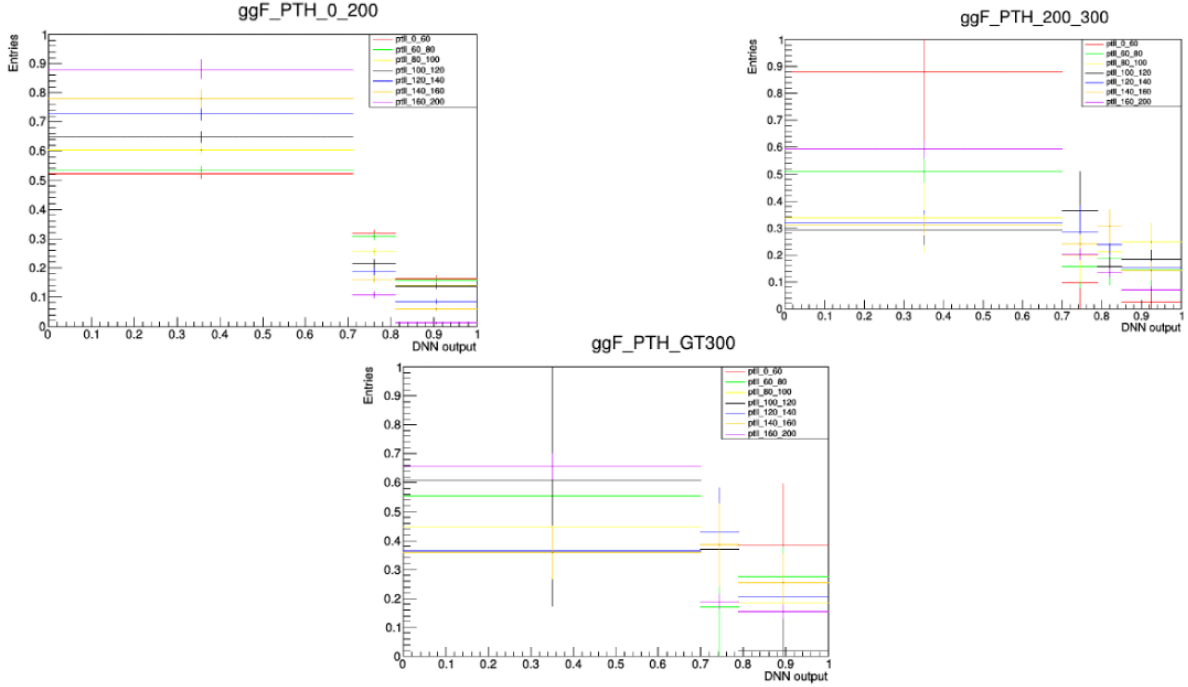


Figure 5.6: 1D plots of p_t^H for each DNN bin

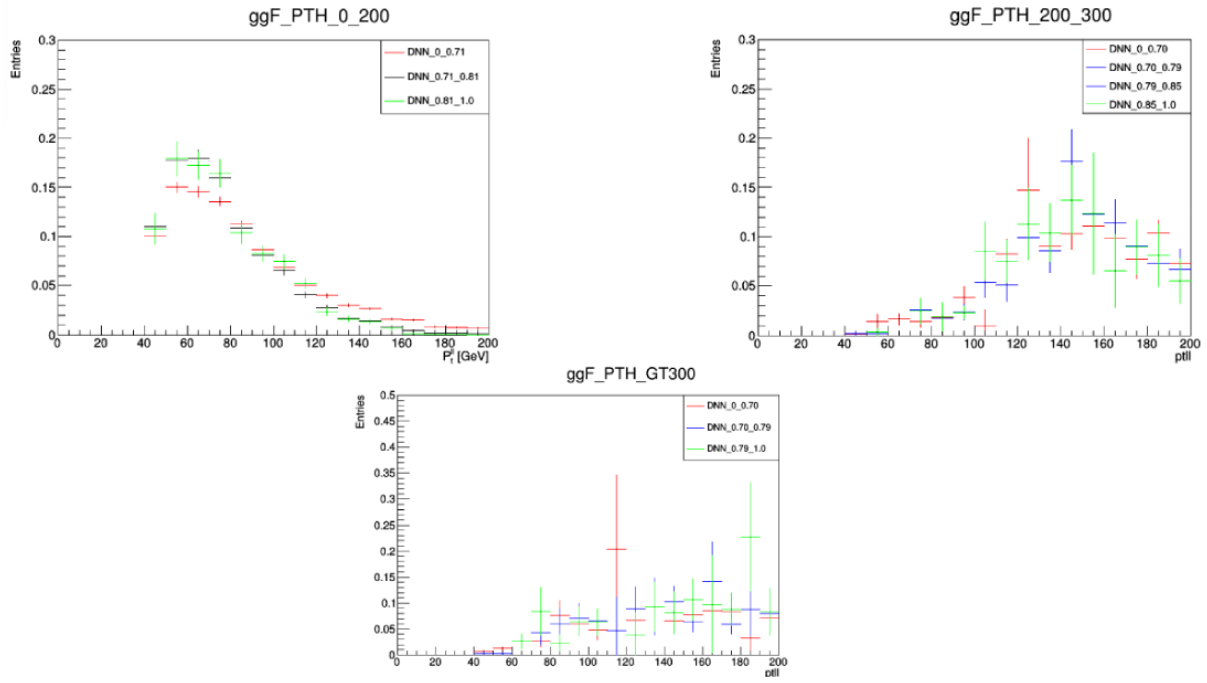


Figure 5.7: 1D plots of the DNN for each p_t^{ll} bin

6 Results

Upon executing the minimization of the likelihood function for the inclusive and STXS measurements, the values and uncertainties of the POIs are obtained, along with important distributions of observables within the analysis regions, forming the results of the measurements. The findings of the STXS measurement are outlined in Section 6.1, while the outcomes of the inclusive measurements (the 2 POI and 1 POI fits) are detailed in Section 6.2. The results in this chapter are post-fit, as they are derived following the minimization of the likelihood function. They denote the aggregate impact of systematic and statistical uncertainties across all regions in the likelihood function, including their intercorrelations, alongside background and signal processes normalized in accordance with likelihood minimization.

6.1. STXS Measurements

In the STXS measurement, multiple cross-sections are simultaneously measured. We were able to achieve a finer STXS 1.2 measurement with more bins and higher sensitivity. The relative cross-section error of each measured STXS bin, normalized by the respective SM prediction, is illustrated in Figure 6.1. The number on the left box represents the relative cross-section error in the previous study and the number on the right box represents the relative cross-section error in the legacy analysis. Several influential changes were made, such as the inclusion of the same flavor 2-jet channel, the addition of the VBF 1-jet channel, a lower fake count due to tighter lepton identification and isolation, and dedicated DNNs for all channels. In the ggF-enriched 2 jets channel, 25% improvements are seen due to the addition of same-flavor channels as seen in the figure.

6.2. Inclusive Measurements

The improvement in the relative cross-section error in the inclusive regions is also illustrated in Figure 6.1. The ggF-inclusive region saw an improvement of 15% compared to the previous study and the VBF-inclusive region saw an improvement of 23%. The NFs for background processes estimated with CRs used to scale the respective predictions of background processes in each SR are listed in Table 6.1. The NFs are consistent with unity within their respective uncertainties. For ggF, in the lower p_T^H range (0-200 GeV), the NF uncertainty was $1.0_{+0.412299}^{-0.232008}$ and in the higher p_T^H regime (GT200 GeV), the uncertainty was $1.0_{+0.475528}^{-0.28179}$. The yields of the signal and background processes in each SR are shown in Table 6.2. The largest deviation is only seen in the ggF GT300 where observed exceeds predicted by 1.4σ ; otherwise, no significant deviation is seen in other regions.

measurement. The legacy analysis employs identical signal regions (SRs) and control regions (CRs) as those utilized in the STXS measurement. Despite these constraints, the sensitivity achieved is comparable to that of inclusive SRs and CRs, primarily due to the implementation of DNNs. For the simultaneous measurement of the ggF and VBF production modes of the Higgs boson (the 2 POI fit), the POIs are μ_{ggF} and μ_{VBF} .

Tables 6.3 and 6.4 provide a comprehensive analysis of the uncertainties associated with the signal strength measurement in Higgs boson production using Asimov and full Run II data. These are preliminary results and more studies will be conducted. The total uncertainty in both measurements is categorized into various sources using the breakdown method, encompassing statistical uncertainty, experimental systematic uncertainty (Exp. sys.), systematic uncertainty from the modeling of the Higgs boson signal process (Sig.), and systematic uncertainty from the modeling of background processes (Bkg.). The statistical uncertainty remained consistent, whereas the experimental systematic uncertainty demonstrated an improvement, particularly evident in the ggF. The signal uncertainty exhibited a notable enhancement, especially for the VBF. Furthermore, a substantial improvement was observed in the background uncertainty.

	Total	Statistical	Exp. sys.	Bkg.	Sig. theory
μ_{ggF}	0.14	0.06	0.08	0.07	+0.07/ − 0.06
μ_{VBF}	+0.24/ − 0.21	0.14	+0.08/ − 0.07	+0.09/ − 0.08	+0.16/ − 0.11

Table 6.3: The expectation of the uncertainties on the signal strengths using the Asimov data set.

	Total	Statistical	Exp. sys.	Bkg.	Sig. theory
μ_{ggF}	+0.13/ − 0.12	0.06	0.06	0.05	+0.07/ − 0.06
μ_{VBF}	+0.21/ − 0.19	0.14	+0.08/ − 0.07	0.05	+0.12/ − 0.11

Table 6.4: The observed values of the uncertainties on signal strengths using the full Run 2 data set.

The cutflow displaying estimated event yields from different processes is shown in Table 6.6 for ggF and Table 6.7 for VBF. At the final cut level, the data/MC ratio is in good agreement with unity, given the uncertainty. Here, NFs are applied to the cutflows. Moreover, a breakdown of the largest contributions to the total uncertainty in the cross-sections are shown in Table 6.5.

Source	$\frac{\Delta\sigma_{\text{ggF+VBF}} \cdot \mathcal{B}_{H \rightarrow WW^*}}{\sigma_{\text{ggF}^* + \text{VBF}} \cdot \mathcal{B}_{H \rightarrow W^*}} [\%]$	$\frac{\Delta\sigma_{\text{ggF}} \cdot \mathcal{B}_{H \rightarrow WW^*}}{\sigma_{\text{gg}} \cdot \mathcal{B}_{H \rightarrow WW^*}} [\%]$	$\frac{\Delta\sigma_{\text{VBF}} \cdot \mathcal{B}_{H \rightarrow WW}}{\sigma_{\text{VBF}} \cdot \mathcal{B}_{H \rightarrow W}} [\%]$
Data statistical uncertainties	4.6	5.1	15
Total systematic uncertainties	9.5	11	18
MC statistical uncertainties	3.0	3.8	4.9
Experimental uncertainties	5.2	6.3	6.7
Flavor tagging	2.3	2.7	1.0
Jet energy scale	0.9	1.1	3.7
Jet energy resolution	2.0	2.4	2.1
$E_{\text{T}}^{\text{miss}}$	0.7	2.2	4.9
Muons	1.8	2.1	0.8
Electrons	1.3	1.6	0.4
Fake factors	2.1	2.4	0.8
Pileup	2.4	2.5	1.3
Luminosity	2.1	2.0	2.2
Theoretical uncertainties	6.8	7.8	16
ggF	3.8	4.3	4.6
VBF	3.2	0.7	12
WW	3.5	4.2	5.5
Top	2.9	3.8	6.4
$Z \tau\tau$	1.8	2.3	1.0
Other VV	2.3	2.9	1.5
Other Higgs	0.9	0.4	0.4
Background normalizations	3.6	4.5	4.9
WW	2.2	2.8	0.6
Top	1.9	2.3	3.4
$Z\tau\tau$	2.7	3.1	3.4
Total	10	12	23

Table 6.5: Breakdown of the largest contributions to the total uncertainty in different channels; combined (ggF+VBF), ggF only, and VBF only, with the individual uncertainties being grouped together.

$\sqrt{s} = 13.7\text{ TeV}$, $\mathcal{L} = 140.9\text{ fb}^{-1}$ (full Run 2)	VBF [125 GeV]	ggF [125 GeV]	Other VV	Other VV	VVV	V_γ	$t\bar{t}$	Single Top	Zee	Zu	M=Id	AUBg	H	Data (MC) (kg-Sig)
$2\text{-jet } (90\text{-}20)\text{ DVT}$	122.19 ± 0.33	335.97 ± 1.67	625.78 ± 0.99	625.78 ± 0.99	189.19 ± 2.88	1114.02 ± 14.87	42870.02 ± 128.17	28869.63 ± 18.76	285496.58 ± 586.97	3874.33 ± 30.69	1760.98 ± 127.46	782587.49 ± 640.11	1083.98 ± 1.97	741360
b-veto	107.11 ± 0.31	283.60 ± 1.53	181.23 ± 0.69	181.23 ± 0.69	141.26 ± 2.47	11296.44 ± 156.24	26907.51 ± 32.67	3370.04 ± 6.83	228110.88 ± 554.76	3107.76 ± 27.80	900.02 ± 93.95	280673.61 ± 586.02	571.95 ± 1.71	254604
$m_T < 55\text{ GeV}$	79.29 ± 0.27	214.03 ± 1.31	44.69 ± 0.34	44.69 ± 0.34	17.26 ± 0.87	980.40 ± 48.63	4073.42 ± 12.59	480.55 ± 2.58	7854.98 ± 73.51	473.87 ± 10.67	185.85 ± 22.23	13803.89 ± 91.07	338.91 ± 1.40	16492
VBF orthogonal: full CIV < 30 full OAV	18.39 ± 0.13	131.90 ± 1.04	21.97 ± 0.34	21.97 ± 0.34	10.21 ± 0.70	320.02 ± 30.03	2180.52 ± 9.69	293.09 ± 1.99	4597.20 ± 51.43	286.30 ± 4.97	113.67 ± 16.08	9329.20 ± 64.46	171.68 ± 1.07	9735
SR DNN 2jet PTH 0-200	12.79 ± 0.11	102.08 ± 0.93	16.17 ± 0.24	16.17 ± 0.24	6.74 ± 0.59	175.02 ± 32.56	2260.35 ± 9.43	271.08 ± 1.93	4442.46 ± 51.14	148.50 ± 3.73	99.52 ± 15.53	8561.28 ± 63.69	131.31 ± 0.96	0
SR DNN 2jet PTH 200-300	3.82 ± 0.06	21.06 ± 0.39	3.72 ± 0.10	3.72 ± 0.10	1.66 ± 0.26	33.21 ± 5.33	143.69 ± 2.26	18.20 ± 0.48	130.03 ± 5.25	82.19 ± 2.26	11.40 ± 3.63	550.82 ± 9.07	28.60 ± 0.41	0
SR DNN 2jet PTH GT200	1.69 ± 0.04	8.45 ± 0.24	1.78 ± 0.07	1.78 ± 0.07	0.81 ± 0.26	10.19 ± 1.53	44.61 ± 1.20	6.03 ± 0.28	24.11 ± 1.34	40.29 ± 1.09	1.75 ± 2.01	194.09 ± 3.45	11.74 ± 0.25	0
$m_T - m_{\Delta\phi}$	0.51 ± 0.03	3.55 ± 0.15	0.58 ± 0.03	0.58 ± 0.03	0.25 ± 0.05	6.52 ± 0.84	59.81 ± 1.41	6.03 ± 0.28	24.11 ± 1.34	40.29 ± 1.09	1.75 ± 2.01	194.09 ± 3.45	11.74 ± 0.25	0
$m_T - m_{\Delta\phi} < 15$	0.51 ± 0.03	3.55 ± 0.15	0.58 ± 0.03	0.58 ± 0.03	0.25 ± 0.05	6.52 ± 0.84	59.81 ± 1.41	6.03 ± 0.28	24.11 ± 1.34	40.29 ± 1.09	1.75 ± 2.01	194.09 ± 3.45	11.74 ± 0.25	0
$p_{\text{th}} > 40\text{ GeV}$	0.51 ± 0.02	1.28 ± 0.10	25.22 ± 0.26	25.22 ± 0.26	17.72 ± 0.75	2044.46 ± 55.34	2400.35 ± 9.76	303.52 ± 2.05	56983.39 ± 132.93	31.32 ± 1.64	372.41 ± 28.31	6505.90 ± 148.97	27.03 ± 0.28	10957
METSig-4	0.51 ± 0.02	1.28 ± 0.10	25.22 ± 0.26	25.22 ± 0.26	17.72 ± 0.75	2044.46 ± 55.34	2400.35 ± 9.76	303.52 ± 2.05	56983.39 ± 132.93	31.32 ± 1.64	372.41 ± 28.31	6505.90 ± 148.97	27.03 ± 0.28	10957
VH veto	0.49 ± 0.02	1.09 ± 0.09	20.88 ± 0.24	20.88 ± 0.24	15.26 ± 0.70	1890.44 ± 53.25	2183.32 ± 9.32	272.54 ± 1.94	50751.63 ± 125.53	27.99 ± 1.55	354.27 ± 26.77	5895.08 ± 139.60	22.47 ± 0.26	90281
VBF orthogonal: full CIV < 30 full OAV	0.45 ± 0.01	12.99 ± 0.19	4.97 ± 0.19	4.97 ± 0.19	2.29 ± 0.35	1267.55 ± 18.42	1506.51 ± 8.83	187.48 ± 1.62	36970.84 ± 108.81	9.32 ± 0.46	253.87 ± 23.48	4270.76 ± 120.75	13.52 ± 0.20	42342
ZKTH ggF 2jet PTH 0-200	0.05 ± 0.01	0.45 ± 0.06	3.38 ± 0.09	3.38 ± 0.09	4.59 ± 0.35	220.11 ± 11.36	246.41 ± 2.94	29.74 ± 0.62	2852.05 ± 17.08	11.56 ± 0.78	28.76 ± 5.14	4119.69 ± 21.02	3.88 ± 0.11	4084
ZKTH ggF 2jet PTH 200-300	0.05 ± 0.01	0.45 ± 0.06	3.38 ± 0.09	3.38 ± 0.09	4.59 ± 0.35	220.11 ± 11.36	246.41 ± 2.94	29.74 ± 0.62	2852.05 ± 17.08	11.56 ± 0.78	28.76 ± 5.14	4119.69 ± 21.02	3.88 ± 0.11	4084
ZKTH ggF 2jet PTH GT200	0.05 ± 0.01	0.45 ± 0.06	3.38 ± 0.09	3.38 ± 0.09	4.59 ± 0.35	220.11 ± 11.36	246.41 ± 2.94	29.74 ± 0.62	2852.05 ± 17.08	11.56 ± 0.78	28.76 ± 5.14	4119.69 ± 21.02	3.88 ± 0.11	4084

Table 6.6: Signal and background MC event count per ggF STXS signal region and CR for 2jSF channel after applying background normalization factors.

$\sqrt{s} = 137\text{ eV}$, $\mathcal{L} = 140\text{ nb}^{-1}$ (Full Run 2)	VBF [125 GeV]	ggF [125 GeV]	Other H	EWWW	gWW	Other VV	VVV	V ₁	Single Top	Zeeuu	Ztt	Mt-tt	Alleg	H	Data	Data MC(1kg-Sig.)
$Z/\gamma \rightarrow \mu\mu$ (30 GeV)	107.11 ± 0.31	283.91 ± 1.53	181.25 ± 0.69	462.46 ± 1.39	41.81 ± 1.81	5460.98 ± 11.82	141.26 ± 5.17	1186.44 ± 116.27	18906.51 ± 32.67	3273.04 ± 6.83	3874.13 ± 1.70	4750.08 ± 1.70	520297.03 ± 1.70	1683.05 ± 1.71	1679.00	0.86 ± 0.00
$Z \rightarrow \tau\tau$ veto	87.08 ± 0.28	237.81 ± 1.41	94.41 ± 0.49	6250.59 ± 16.34	226.32 ± 1.10	296.70 ± 1.51	3660.06 ± 9.05	87.18 ± 1.95	18906.51 ± 32.67	2773.04 ± 6.83	3107.76 ± 17.80	843.07 ± 72.93	182771.06 ± 586.02	571.95 ± 1.71	254604	0.82 ± 0.00
$M_t < 70\text{ GeV}$	78.48 ± 0.26	140.82 ± 1.07	23.78 ± 0.24	1337.50 ± 6.81	56.38 ± 0.51	53.24 ± 0.67	233.23 ± 2.29	15.31 ± 0.83	4105.05 ± 12.59	463.79 ± 2.51	431.34 ± 13.79	182.57 ± 21.79	14751.90 ± 91.65	243.08 ± 1.13	14852	0.99 ± 0.01
$p_{Tl} > 40\text{ GeV}$	73.68 ± 0.26	131.37 ± 1.03	21.33 ± 0.23	1210.72 ± 6.43	51.14 ± 0.49	48.00 ± 0.64	210.14 ± 2.15	14.04 ± 0.80	3604.99 ± 11.96	427.60 ± 2.41	337.70 ± 9.27	132.72 ± 18.39	12688.45 ± 72.68	226.37 ± 1.09	13852	1.01 ± 0.01
$\Delta\phi_{l\bar{l}} > 150^\circ$	73.68 ± 0.26	131.37 ± 1.03	21.33 ± 0.23	1210.72 ± 6.43	51.14 ± 0.49	48.00 ± 0.64	210.14 ± 2.15	14.04 ± 0.80	3604.99 ± 11.96	427.60 ± 2.41	337.70 ± 9.27	132.72 ± 18.39	12688.45 ± 72.68	226.37 ± 1.09	13852	1.01 ± 0.01
CR	50.20 ± 0.23	49.88 ± 0.65	5.72 ± 0.12	346.77 ± 8.88	33.16 ± 0.40	20.60 ± 0.42	65.25 ± 1.28	3.23 ± 0.37	1237.05 ± 7.02	135.28 ± 1.47	78.99 ± 2.81	46.13 ± 10.28	4312.24 ± 45.27	114.81 ± 0.70	4097	1.04 ± 0.02
OLV bnd	51.83 ± 0.22	24.54 ± 0.45	2.39 ± 0.07	149.22 ± 2.76	29.31 ± 0.37	9.77 ± 0.29	27.69 ± 0.86	1.69 ± 0.26	371.38 ± 3.51	56.90 ± 0.88	740.30 ± 37.02	21.87 ± 16.46	1550.41 ± 26.67	78.76 ± 0.51	0	0.00
ggHbbg, 2l MJT 350, 700 PTH 0 200	15.56 ± 0.11	11.96 ± 0.32	0.95 ± 0.05	81.69 ± 1.99	7.58 ± 0.19	5.64 ± 0.22	14.93 ± 0.59	0.80 ± 0.18	259.09 ± 3.05	35.51 ± 0.69	571.06 ± 22.38	15.35 ± 1.15	1043.30 ± 24.91	26.47 ± 0.54	0	0.00
ggHbbg, 2l MJT 700, 1000 PTH 0 200	7.97 ± 0.08	3.37 ± 0.18	0.35 ± 0.02	18.89 ± 1.09	4.49 ± 0.15	1.79 ± 0.12	4.16 ± 0.38	0.13 ± 0.07	30.99 ± 1.40	8.48 ± 0.34	125.12 ± 9.79	1.96 ± 2.28	231.03 ± 10.67	11.89 ± 0.30	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	7.97 ± 0.08	3.37 ± 0.18	0.35 ± 0.02	18.89 ± 1.09	4.49 ± 0.15	1.79 ± 0.12	4.16 ± 0.38	0.13 ± 0.07	30.99 ± 1.40	8.48 ± 0.34	125.12 ± 9.79	1.96 ± 2.28	231.03 ± 10.67	11.89 ± 0.30	0	0.00
ggHbbg, 2l MJT 350, 700 PTH 0 200	7.66 ± 0.08	1.10 ± 0.10	0.09 ± 0.01	4.77 ± 0.64	4.72 ± 0.15	0.76 ± 0.08	1.50 ± 0.32	0.08 ± 0.05	9.88 ± 0.62	1.70 ± 0.15	30.00 ± 1.32	2.55 ± 1.69	63.94 ± 5.31	8.85 ± 0.13	0	0.00
ggHbbg, 2l MJT 350, 700 PTH 0 200	5.89 ± 0.07	4.46 ± 0.18	0.64 ± 0.04	27.41 ± 0.71	3.20 ± 0.12	0.32 ± 0.05	4.66 ± 0.28	0.44 ± 0.14	4.46 ± 0.24	0.65 ± 0.09	26.47 ± 1.70	11.42 ± 0.75	125.78 ± 3.03	11.17 ± 0.26	0	0.00
ggHbbg, 2l MJT 1000, 1500 PTH 0 200	3.77 ± 0.06	0.78 ± 0.07	0.14 ± 0.01	3.95 ± 0.34	1.59 ± 0.09	0.05 ± 0.02	0.69 ± 0.09	0.08 ± 0.06	4.48 ± 0.38	0.65 ± 0.09	3.17 ± 0.45	2.63 ± 0.36	19.67 ± 1.07	4.69 ± 0.10	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJT 1500, 1600 PTH 0 200	3.44 ± 0.07	0.42 ± 0.05	0.15 ± 0.01	2.62 ± 0.24	2.34 ± 0.11	0.06 ± 0.02	0.27 ± 0.09	0.42 ± 0.20	2.26 ± 0.28	0.42 ± 0.07	2.16 ± 0.46	1.33 ± 0.18	14.83 ± 0.78	6.02 ± 0.09	0	0.00
ggHbbg, 2l MJ																

The distribution of the DNN output in the ggF-enriched $N_{jet} \geq 2$ SR is depicted in Figure 6.2, with the background normalization factors used and the binning used in the statistical analysis. To avoid bias, a blinding strategy is implemented to verify the background modeling by partially unblinding the DNNs in the low DNN regions. The first half of the DNN bins or the first $\frac{N-1}{2}$ bins are unblinded. The higher DNN bins, where the signal is expected to dominate, remain blinded until the analysis is finalized with an exception, In the ggF p_T^H 0 – 200 GeV SR, the first two bins of a total of three DNN bins are unblinded due to the fact that the second DNN bin in this SR has a small s/b . Figure 6.3, shows the DNN plots for the control regions for ggF p_T^H 0-200 GeV. Zll events tend to have low DNN scores and therefore the distributions are typically one bin. Other p_T^H regions are in the appendix.

The DNN output is a score in the range of 0 to 1. Events with scores closer to 1 comprise more of the signal, while events with scores closer to 0 are dominated by background events. In general, the modeling between data and MC is good.

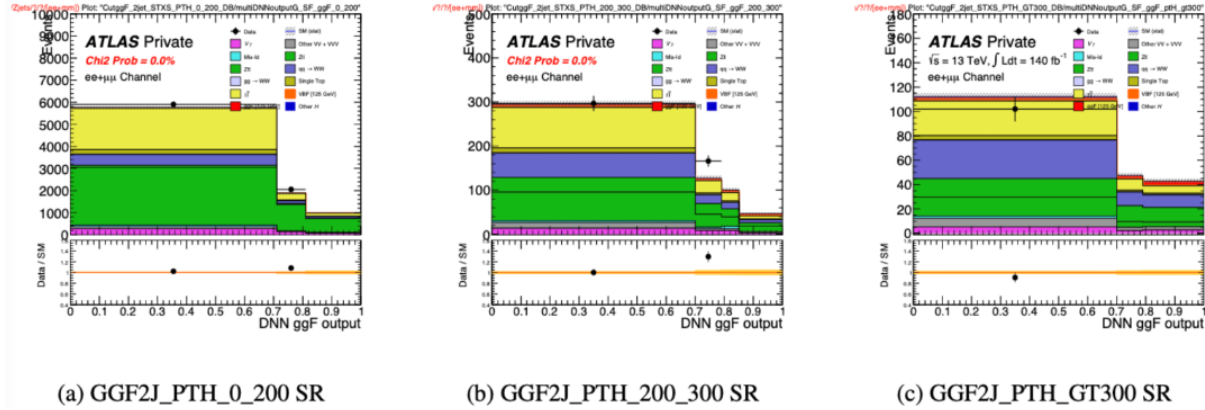


Figure 6.2: Stacked histograms of the DNN ggF output for different regions based on the transverse momentum of the Higgs boson candidate (p_T^H) in the SF channel. (a) shows the region with p_T^H between 0 and 200 GeV, (b) with p_T^H between 200 and 300 GeV, and (c) with p_T^H greater than 300 GeV. The stacked contributions represent various background processes, and the bottom panel shows the ratio of data to the Standard Model (SM) prediction.

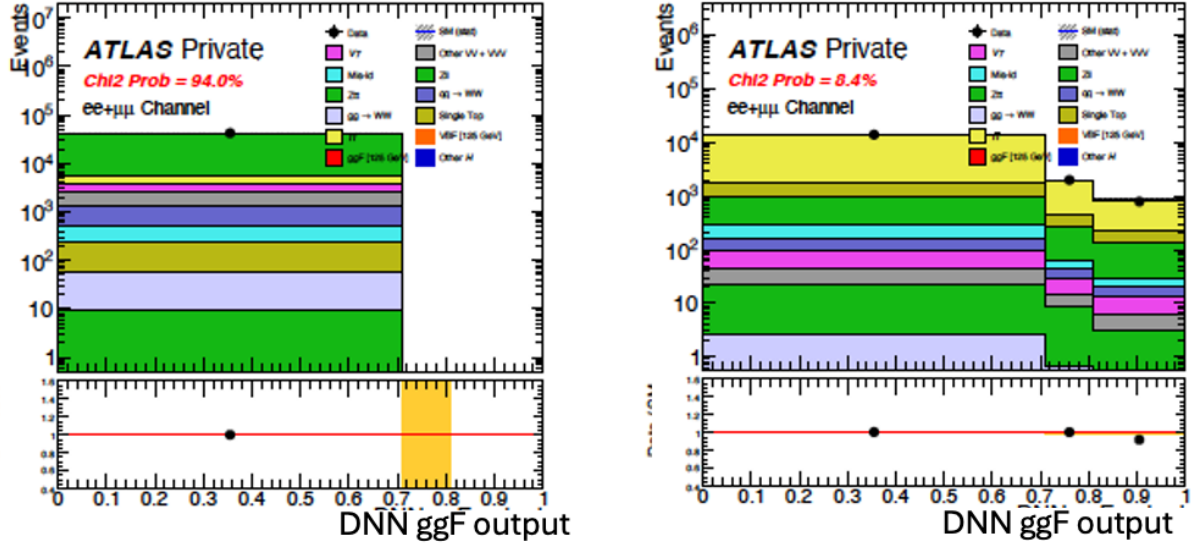
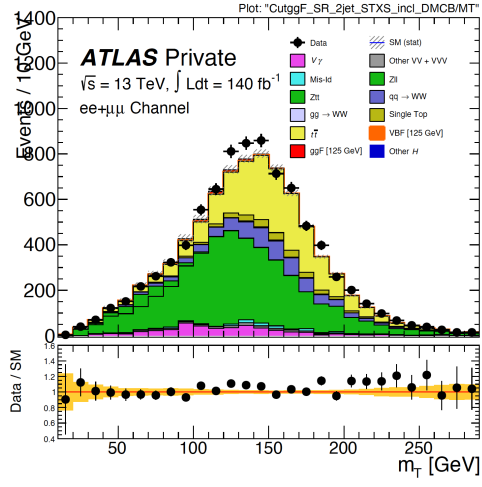
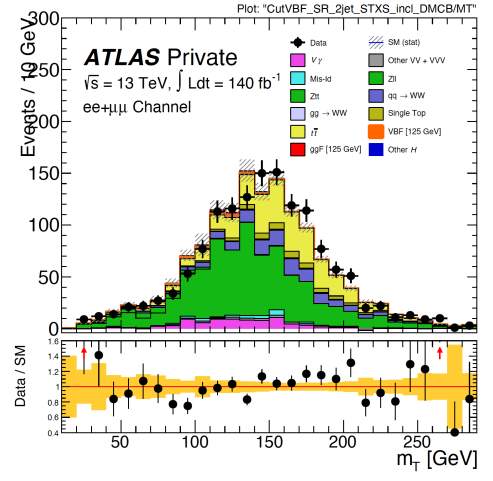


Figure 6.3: Stacked histograms of the DNN ggF output for control regions based on the transverse momentum of the Higgs boson candidate (p_T^H) in the SF channel. (left) shows the Zll control region with p_T^H between 0 and 200 GeV and (right) shows the top control region with p_T^H between 0 and 200 GeV.

To verify the modeling, several kinematic distributions are plotted for the inclusive unblinded ggF SR regions. Here, normalization factors are applied. In general, the MC properly describes the data. Figure 6.4 shows the m_T distributions for the ggF and VBF two jets with the same flavor regions. More inclusive plots are shown in the appendix. The modeling between data and MC is good.



(a)



(b)

Figure 6.4: Inclusive distribution of m_T in the a) ggF 2j SF and b) VBF 2j SF regions. The bottom panel displays the ratio of the data yields to SM. Here, normalization factors are applied.

7 Conclusions

This thesis presents comprehensive legacy measurements of the ggF production cross-section of the Higgs boson using the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay mode. The study has yielded significant improvements and insights into measuring Higgs boson properties and production mechanisms. The measurements are based on the complete LHC Run 2 data set, which was collected by the ATLAS detector from pp collisions generated at $\sqrt{s} = 13\text{TeV}$ between 2015 and 2018. This data set corresponds to an integrated luminosity of 139fb^{-1} . The signal-to-background discrimination of individual regions is enhanced by categorizing the events selected for analysis based on jet multiplicity and production mode. Dedicated CRs are used to measure the significant contamination of the selected event yields from background processes, such as non-resonant WW , top quark, and Z/γ^* processes, to constrain their normalization within each jet-binned region. In the ggF analysis regions, an advanced analysis technique involving 2D-DNN is employed as the discriminating observable to separate ggF, VBF, and background processes. It has replaced the previous method of using m_T . Employing DNN as the discriminant variable in ggF channels significantly improved background suppression.

Notably, the inclusion of the Same Flavor 2-jet channel and the addition of the VBF 1-jet channel have expanded the scope of the analysis. These advancements have been facilitated by several influential changes in the legacy analysis compared to previous studies. Key improvements in the analysis include a 15% enhancement in sensitivity for the ggF inclusive study and a 25% enhancement for a finer Simplified Template Cross Section (STXS) measurement in the 2jet Channel.

Our findings confirm that the observed mis-modeling in p_t^H is irrelevant to the DNN modeling, providing confidence in the robustness of our analysis techniques. However, challenges remain, particularly in modeling Z +jet processes in 1/2-jet channels, especially in the 2-jet Same Flavor channel. While the cause of this mis-modeling is yet to be fully understood, we have implemented a few strategies already, such as using $p_T^H > 40\text{GeV}$ cut on 2jSF channel, additional Z jets p_T shape theory systematics, and checks for background mis-modeling in CRs for both the DNN and kinematic distributions. To facilitate a smooth unblinding process, we will partially unblind the low DNN bins in the 2-jet Signal Regions while keeping the high DNN bins, which are more pure in signal and drive signal sensitivity, blinded. This approach allows us to scrutinize the background processes, mainly clustered in low DNN bins, regarding modeling accuracy.

Looking ahead, several future considerations emerge from this work. The potential implementation of STXS 1.3, featuring a possible 4-bin $\Delta\phi_{jj}$ within p_T^H/m_{jj} VBF bins, offers opportunities for more detailed measurements. The upcoming Run-3 combination naturally lends itself to the inclusion of CP/EFT measurements, though measuring finer bins may prove challenging. These include updating METsig samples with full DNN

re-training, optimizing binning, and cleaning systematics, particularly in areas with constraints such as JER, FT, and theory systematics. The planned partial unblinding after the METsig update will be cautiously approached, mindful of the observed mis-modeling issues. The implementation of electron DNN identification aims to provide better fake rejection and potentially improve same-sign modeling, addressing current challenges in lepton identification and background estimation. These anticipated improvements in both analysis techniques and computational frameworks position us to achieve even greater precision and sensitivity in future Higgs boson measurements.

In conclusion, this thesis has demonstrated significant advancements in the precision and scope of Higgs boson measurements. While challenges remain, particularly in background modeling, the improvements in analysis techniques, coupled with the ongoing refinements and prospects outlined, position us to gain even deeper insights into Higgs boson physics in the coming years. The work presented here not only enhances our understanding of fundamental particle interactions but also paves the way for future discoveries and refinements in high-energy physics.

Appendices

A | Cost-monitoring of b -jet triggers

This appendix serves to document additional contributions made to ATLAS projects beyond the primary $H \rightarrow WW^*$ measurement that forms the core of this thesis. These supplementary efforts, while not central to the main body of research, represent significant contributions to the broader ATLAS collaboration and deserve acknowledgment. A key component of these additional contributions was a qualification task undertaken with the b -jet trigger group. This task encompassed two main responsibilities:

- **CPU Consumption Analysis:** A comprehensive analysis was conducted on the computational demands of various algorithms and their components essential for online b -tagging, including track finding, vertex finding, and b -tagging algorithms.
- **Automated Cost Monitoring System:** An innovative system was developed to automate the monitoring of b -jet trigger costs. This system was designed to operate on a nightly basis, providing crucial and timely information on trigger performance.

This section presents a comprehensive account of the methodologies employed and significant findings that emerged from the cost-monitoring of b -jet triggers in the ATLAS experiment.

Context and Motivation

The transition of ATLAS to a multi-threaded environment (AthenaMT) necessitated a reevaluation of computational costs associated with various trigger algorithms. B -jet triggers, known for their high resource demands, were a primary focus of this analysis. Two key factors underscore the importance of this investigation:

- The critical role of b -jet events in the search for novel physics and the evaluation of known processes.
- The significant CPU consumption of b -jet trigger algorithms, particularly those involving fast-tracking (e.g., `TrigFastTrackFinder_bjet`) and precision tracking (e.g., `TrigAmbiguitySolver_bjet`) for primary and secondary vertex reconstruction and b -tagging evaluation.

While substantial optimizations were implemented during LHC Run 2, notably through the efforts of Katharine Leney and Stephen Sekula in 2017 and 2018, CPU usage remained a concern for b -jet triggers. The shift to AthenaMT, coupled with changes to EMPFlow jets and DL1r tagging, necessitated a new cost analysis to assess the current state of b -jet trigger performance.

Project Structure

The project was structured into three main tasks:

- Tool Development and Comparative Analysis
 - Identification and documentation of appropriate cost-monitoring tools, as previous tools were either non-functional or outdated.
 - Comparative analysis of CPU usage across various b-jet trigger algorithms and chains, as well as non-b-jet trigger algorithms.
- Automation of Cost-Monitoring
 - Development of an automated system to generate histograms and Excel files on a nightly basis.
 - Development of an automated system to generate histograms and Excel files on a nightly basis, enabling regular assessment of the impact of algorithm modifications on CPU usage.
- Parameter Sensitivity Analysis
 - Investigation of the impact of parameter variations ($\min_{p_T}$, ROI size including eta and phi widths) on computational cost.
 - This analysis served two purposes: a) Validation of the cost monitoring tools and Athena Release Tester (ART) codes. b) Establishment of a framework for future parameter optimization.

The subsequent sections will delve into the methodologies employed for each task and present the key findings that emerged from this comprehensive analysis of b-jet trigger costs in the ATLAS multi-threaded environment.

Task1: Finding the tools for cost-monitoring:

Initial efforts were hampered by a lack of documentation and understanding of appropriate cost-monitoring programs. Previously utilized tools proved incompatible with 2018 and 2022 data sets. After consultation with other Trigger groups, `Trig_reco_tf.py` was identified as the optimal tool for our cost-monitoring requirements. `Trig_reco_tf.py` functions as a wrapper between the ATLAS software (e.g., Athena) and the production system, facilitating the coordinated execution of multiple programs necessary for cost monitoring.

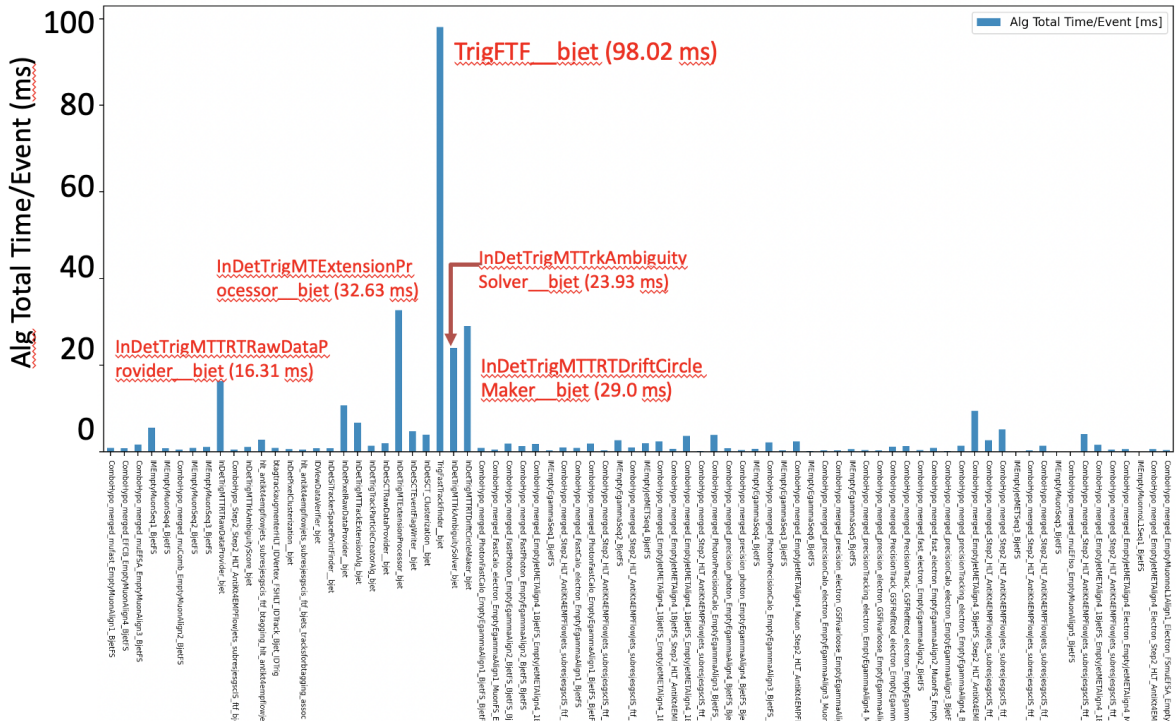
```
1  Trig_reco_tf.py --inputBS_RD0File=$INPUT --athenaopts "BSRD0toRAW:--imf
   --threads=8--concurrent-events=8 --nprocs=1" --maxEvents 1000 --
   autoConfiguration everything --beamType collisions --runNumber 431885 --
   costopts "all:--monitorChainAlgorithm"--outputNTUP_TRIGCOSTFile="
   TrigCostRoot_Results.root" --precommand
2  """setMenu='Physics_pp_run3\_{v1}';doL1Sim=True;rewriteLVL1=True;""" --loglevel
   DEBUG --prodSysBSRD0 True --ignoreErrors False 2>&1 | tee transforms.log
```

The input data used for testing is data22_13p6TeV.00431885.physics_EnhancedBias.merge.RAW._lb0545._SFO-15._0001.1, which is also used by other ART tests. The execution of Trig_reco_tf.py over the input data gives the root file (TrigCostRoot_Results.root), which contains histograms and a metadata tree. Subsequently, a post-processing script (CostAnalysisPostProcessing.py) is applied to the root file to create CSV tables and a JSON metadata tree.

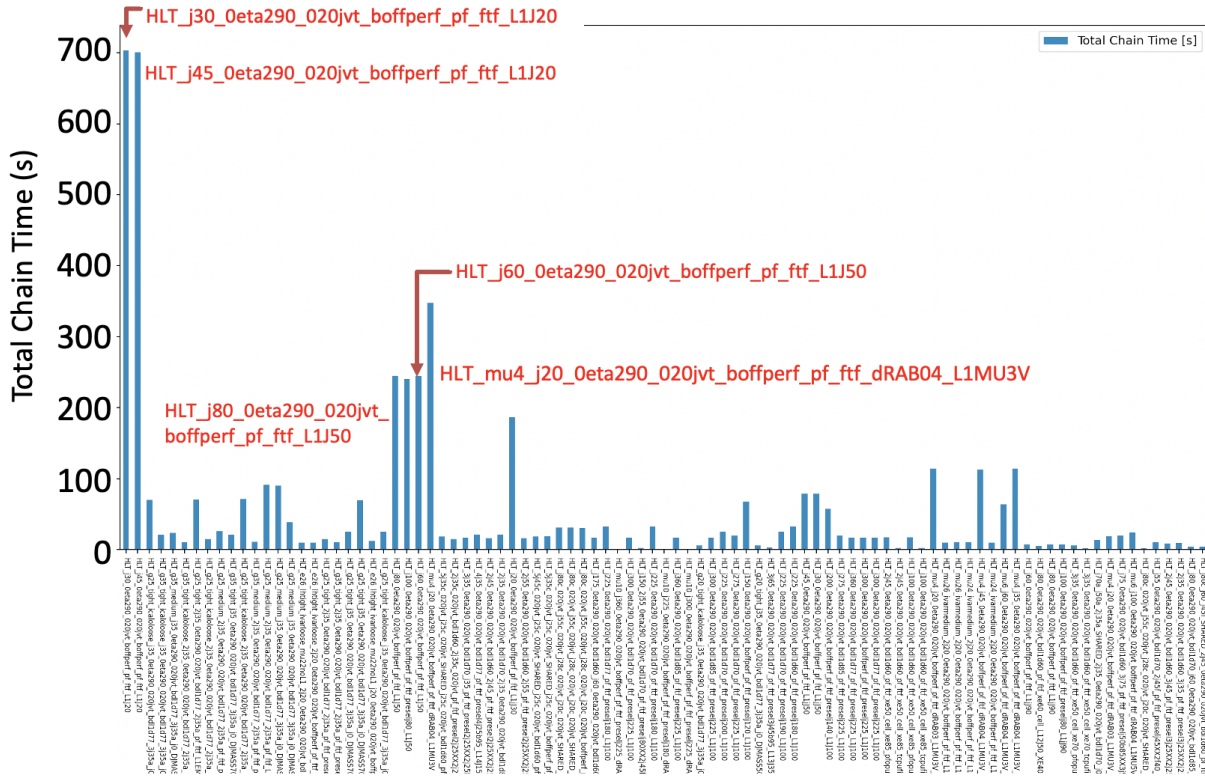
Analysis of the generated Excel files allowed for investigation and comparison of CPU usage across various b -jet trigger algorithms and chains, as well as non b -jet trigger algorithms. This comparative analysis serves as a crucial first step towards optimization, potentially enhancing the efficiency of b -jet trigger software.

Figure 1 (a) shows the algorithm total time per call [ms] for b -jet algorithms. Out of all the b -jet algorithms, TrigFastTrackFinder__bjet (TrigFTF__bjet) dominated CPU cost, which took a mean algorithm total time/call of 98.02 ms, which uses more than three times as much CPU time as the next-most CPU-demanding algorithm. Algorithms that followed after that were InDetTrigMTEExtensionProcessor__bjet and InDetTrigMT-TrkAmbiguitySolver__bjet, among others. These findings align with previous research on jet triggers, which identified fast-tracking and certain precision-tracking algorithms as the most computationally expensive components.

B -jet Chains were also investigated, as shown in Figure 1 (b). HLT_j20_0eta290_020jvt_boffperf_pf_ftf_L1jJ30 was the chain with the most total chain time (s). This particular chain was selected for a single chain-based study, with the potential for more comprehensive investigation in future work.



(a)



(b)

Figure 1: a) Total Time per Event (ms) vs. b -jet algorithms and b) Total Chain Time (s) vs. b -jet chains

To identify features potentially overlooked due to limited data, an extended cost-monitoring analysis was conducted using a larger event sample. The CPU Time per Event for the TrigFTF__bjet algorithm shows a right-skewed distribution with a long tail with some events extending up to 1000 ms, with a mean of 55.61 ms. Similar distributional features were observed in other algorithms, suggesting a consistent pattern across the b-jet trigger system.

Moreover, a deeper inspection of Figure 1 a) and Figure 2 revealed an apparent discrepancy in the algorithm total time per event for TrigFTF__bjet. Hence, the CPU time/event ratio of two different algorithms from the two runs, i.e., 263 events and 9659 events, was calculated. The ratio of CPU time per event between the two algorithms remained consistent across both runs, indicating that the cost-monitoring tool is functioning as intended. This consistency provides validation for our methodology and tools. The initial apparent discrepancy in absolute values can be attributed to the use of different computing machines at different times for the two runs. When properly normalized, this discrepancy is resolved, further confirming the reliability of our cost-monitoring approach.

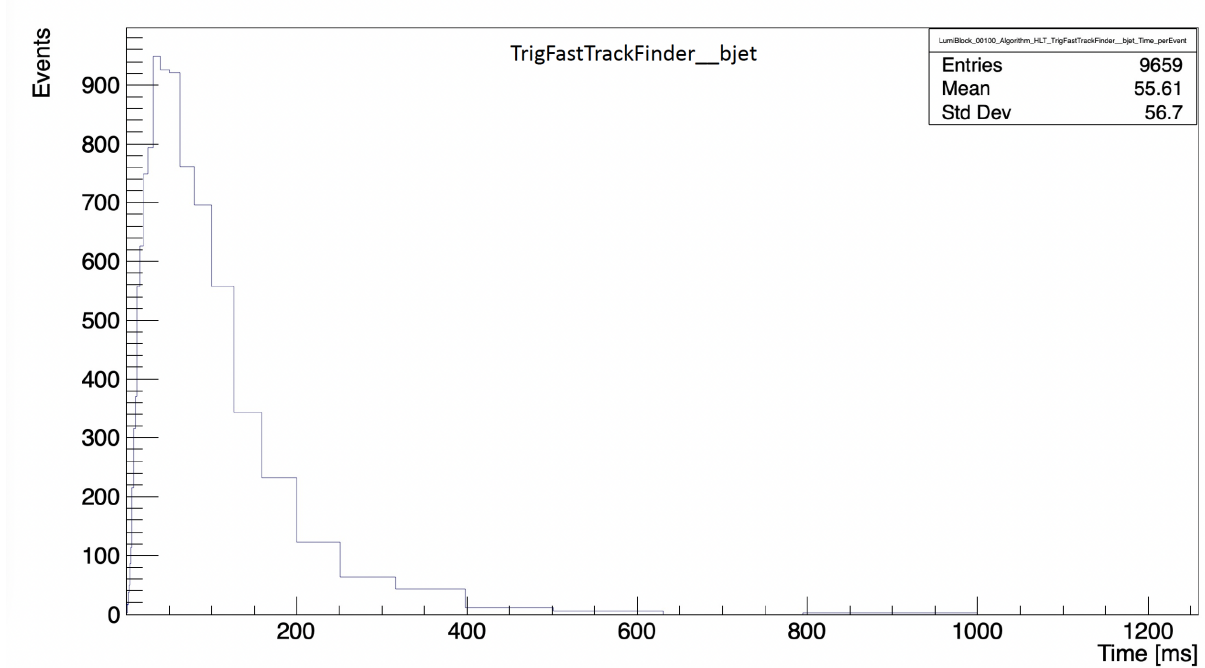


Figure 2: CPU Time per Event of TrigFTF__bjet with 9659 events

Task2: Automating the process:

The second major task of this project involved the development of an Athena Release Tester (ART) system to automate the cost-monitoring process on a nightly basis. The goal of the task is to regularly monitor the trigger performance to promptly identify and address any issues in the software and to provide consistent and comparable data over time. The automated system consists of several key components: a script file (cost_automation-

new.sh) to set up the Athena environment and automate the Trig_reco_tf.py, and CostAnalysisPostProcessing.py to produce the necessary root files and excel files. The only input for the script file is the path to the data sample and the Lumiblock value of the sample. The script also runs read_csv.py and grapher.c, which produces comparison graphs (similar to Figure 1) and histograms from the root file, respectively.

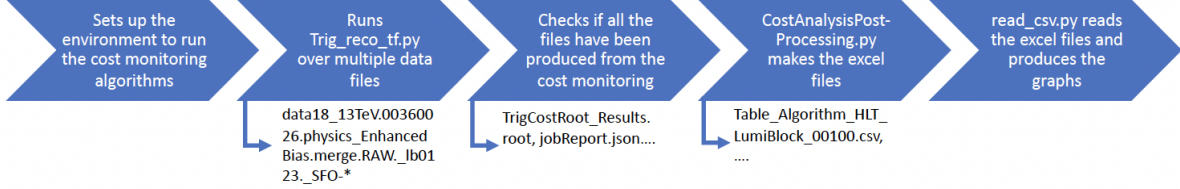


Figure 3: Automation process followed by cost-monitoring script

In order to obtain this script file, the GitLab repository is <https://gitlab.cern.ch/bshresth/cost-monitoring.git>.

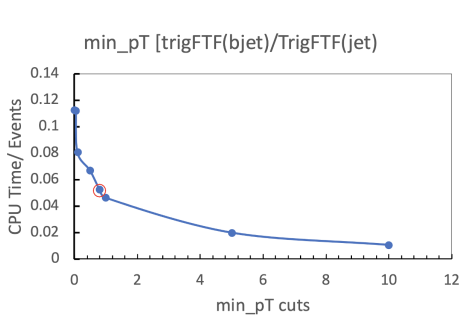
Task 3: Testing the process:

Following the implementation of the automated system, a series of tests were conducted to validate the functionality of the cost-monitoring tools. The primary method of validation involved analyzing the effects of parameter changes on CPU Time/Event. The validation process focused on the effects of three key parameters: $\min_{p_T}$, etawidth, and phiwidth on CPU Time/Event. The standard value for the $\min_{p_T}$ was 0.8 GeV, and the two widths were set to 0.4. Changes were implemented to these three parameters in the ConfigSettings__bjet class of ConfigSettings.py located in TrigInDetConfig, one parameter change at a time. Here, the algorithm that was studied was TrigFTF__bjet and was normalized with TrigFTF__jet.

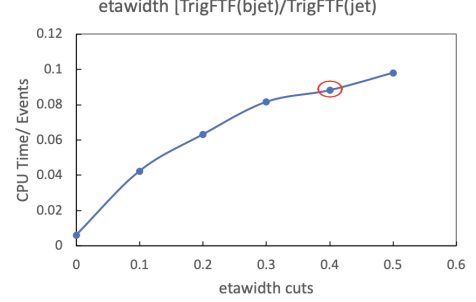
The relationship between $\min_{p_T}$ and CPU Time/Event aligned with theoretical expectations, as illustrated in Figure 4. An increase in the $\min_{p_T}$ cut-off corresponded to a decrease in CPU Time. This can be attributed to the reduction in the number of tracks included in the analysis at higher $\min_{p_T}$ values, thereby decreasing the volume of data to be processed. Contrary to expectations, initial tests of phi halfwidth and eta halfwidth showed minimal impact on CPU time usage. Despite testing a wide range of values (0.0001 to 10), CPU Time/Event fluctuated around a constant value, suggesting that the changes in phiwidth and etawidth were not being effectively implemented.

After consultation with Tim Martin, Katharine Leney, and Dan Guest, it was determined that the width changes in ConfigSettings.py were being ignored due to an issue in the Bjet menu code, likely introduced during a recent update. An alternative approach was devised, implementing width changes directly in ViewCreatorCentredOnJetWithPV-ConstraintROITool.h. This method produced the expected results for width variations, as shown in the accompanying figure. The observed trend aligns with theoretical expectations: larger regions of interest (ROI) result in more jets being accounted for, increasing

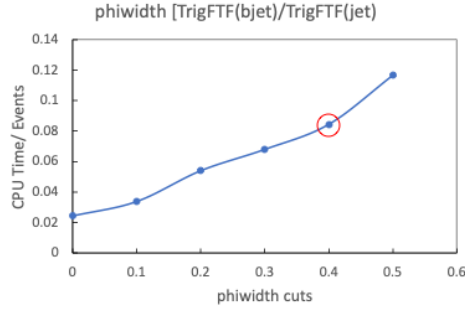
the volume of data to be processed and, consequently, the required CPU time for analysis. In these figures, the red circle denotes the standard values set for the parameters.



(a) $\min_{p_T}$ cuts vs. CPU usage



(b) etawidth cuts vs. CPU usage



(c) phiwidth cuts vs. CPU usage

Figure 4: Observing the change in the CPU usage by TrigFTF__bjet with changing parameters.

Tuning these parameters, as Katherine and Steve have done before, would improve CPU usage for the new multi-threaded environment. Only minimal gains in CPU can be obtained by raising the track p_T threshold from 0.8 GeV. Hence, the $\min_{p_T}$ cut is reasonably optimized. The impact of varying the eta and phi width of chains on the trigger efficiencies is also minimal as the time here is in [ms]. A further study on efficiencies could be performed to look at the effects of changing etawidth.

****Update:** The issue where width parameters set in ConfigSettings.py were being ignored by BjetMenuSequences.py has been resolved. This improvement allows for direct manipulation of both eta and phi widths through ConfigSettings.py, streamlining future optimization efforts.

Conclusion

In a nutshell, using the automation script file, a routine analysis of CPU usage can be performed in the AthenaMT environment. This can be useful to achieve a reduction in CPU cost that is appropriate for avoiding trigger prescaling without significantly compromising the efficiency of trigger chains and alerting if any discrepancies arise in CPU usage. The graphs of CPU usage obtained by changing the parameters suggest that the parameters are reasonably optimized. The fast-tracking algorithms and some precision tracking algorithms used in ambiguity solving and track extension consume the most CPU Time in the *b*jet trigger. If these algorithms were made more efficient, the CPU consumption of the *b*jet trigger would be greatly reduced. The goal of this project going forward is to optimize the algorithms and parameters as Katherine and Stephen previously did.

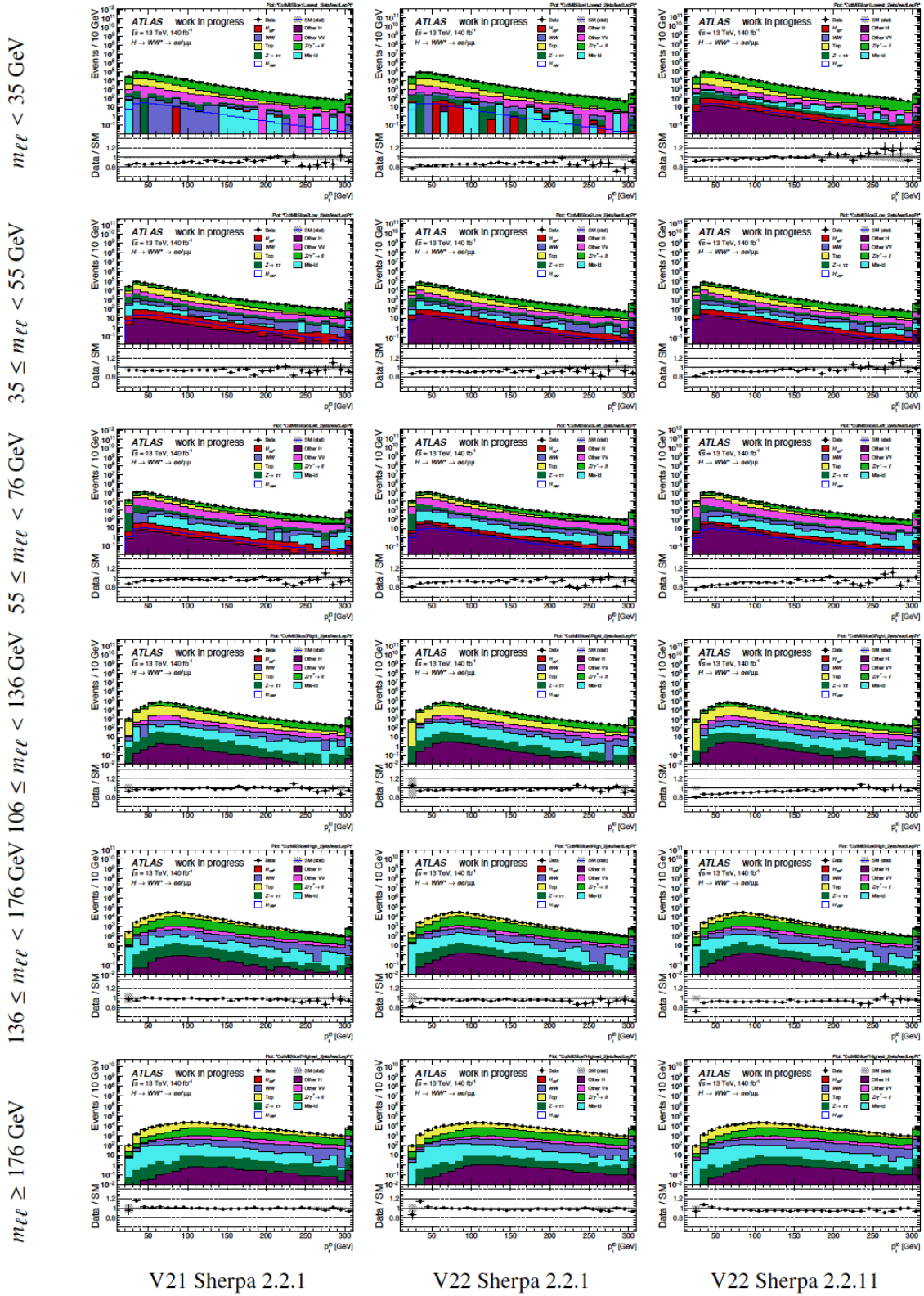


Figure 6: Transverse momentum of leading lepton distributions for 2 jets at different $m_{l\bar{l}}$ slices.

The data/MC discrepancy can be partially mitigated by excluding events with $p_T^{\ell\ell} < 40\text{GeV}$. This cut reduces the mis-modelling of Sherpa 2.2.11 to a greater degree than for Sherpa 2.2.1. Still, the behavior of this discrepancy is rather complex as the data/MC ratio depends not only on jet multiplicity and lepton- p_T but also on the di-lepton invariant mass $m_{\ell\ell}$ [140]. This means that the extrapolation from control regions defined on the Z-peak to signal regions defined off the Z-peak might be sizeable. In the analysis, the effect of normalizing $Z \rightarrow \ell\ell$ in dedicated on-peak control regions is tested in the full statistical fit. Ultimately $Z \rightarrow \ell\ell$ control regions are used only for the $p_T^H > 200\text{GeV}$ categories, where the $p_T^{\ell\ell}$ mismodelling is much less pronounced. $Z \rightarrow \ell\ell$ in the lower- p_T^H control regions is normalized in the lowest, background-rich bins of the DNN in the signal regions.

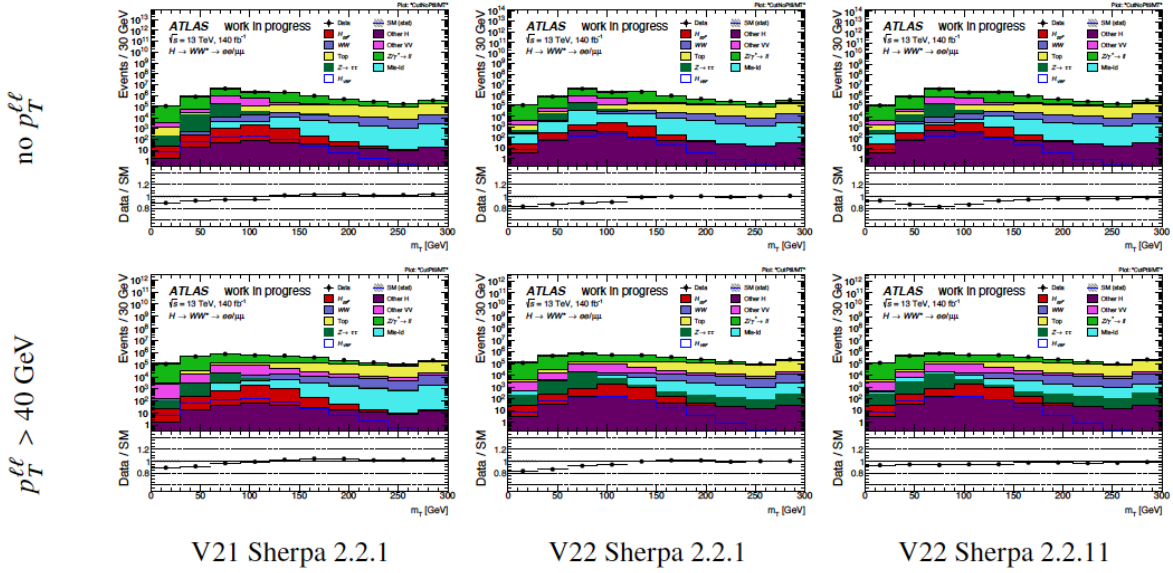


Figure 7: Transverse mass distribution without and with the $p_T^{\ell\ell} > 40\text{GeV}$ cut.

Kinematic Distributions for 2jSF

To further verify the modeling, several kinematic distributions are plotted for the inclusive partially unblinded ggF and VBF regions, as shown in Figure 8 and 9, respectively. In general, the MC properly describes the data.

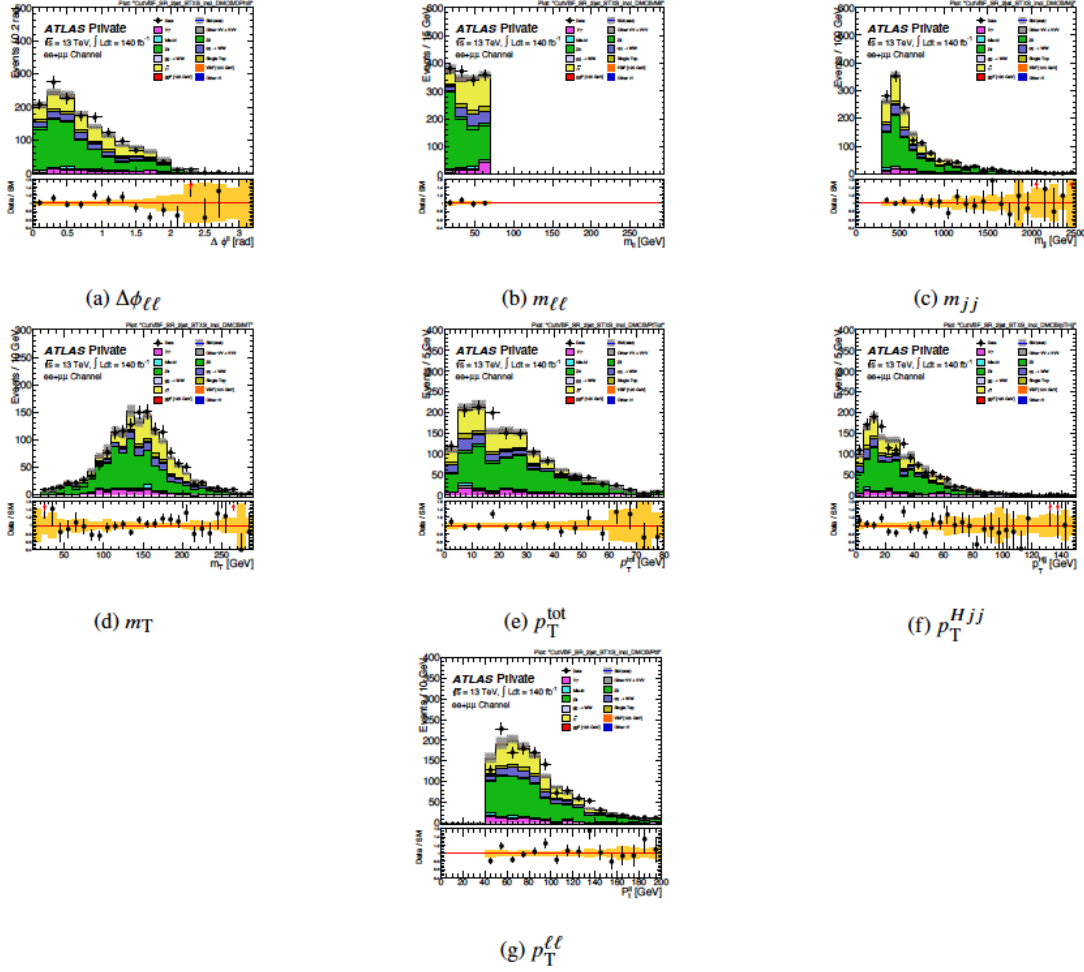
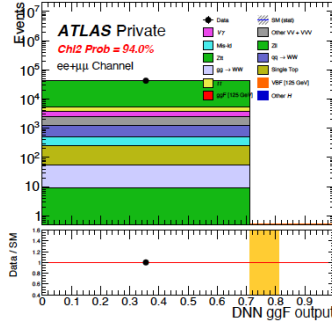


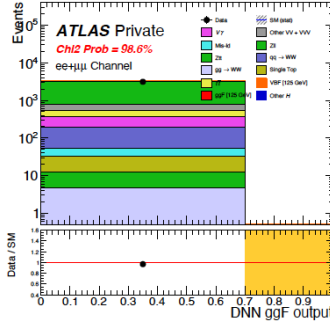
Figure 9: Distributions of $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, m_{jj} , m_T , p_T^{tot} , p_T^{HJJ} and $p_T^{\ell\ell}$ in the inclusive partial unblinded VBF signal region. The contributions from each STXS partial unblinded region received the corresponding background normalizations.

ggF DNN output distribution for the Control Region

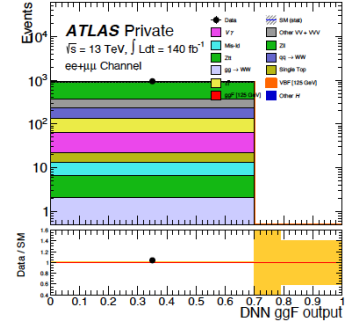
The DNN distributions in the Zll and top CRs or VRs are shown in Figure 10 and Figure 11 for each ggF category. The top validation regions for ggF are defined with the same selections as the SR, except for the selection of exactly one b-tagged jets. In general, theMC properly describes the data.



(a) GGF2J_PTH_0_200 topVR

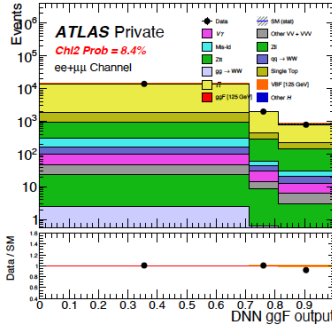


(b) GGF2J_PTH_200_300 topVR

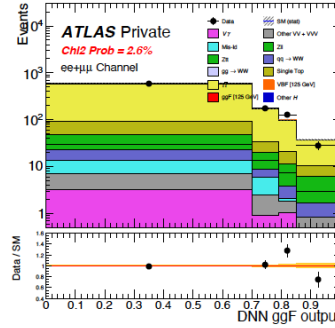


(c) GGF2J_PTH_GT300 topVR

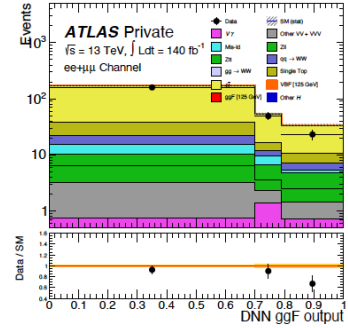
Figure 10: ggF DNN output distribution in the 3 ggF 2-jet SF Zll CRs and VRs with the binning used in the statistical analysis. The data is partially unblinded in the low DNN. NFs are applied.



(a) GGF2J_PTH_0_200 topVR



(b) GGF2J_PTH_200_300 topVR



(c) GGF2J_PTH_GT300 topVR

Figure 11: ggF DNN output distribution in the 3 ggF 2-jet SF top VRs with the binning used in the statistical analysis. The data is partially unblinded in the low DNN. NFs are applied.

1. AFII - Atlas Fast II
2. ALICE - A large ion collider experiment
3. ATLAS - **A Toroidal LHC ApparatuS**
4. BSM - beyond Standard Model
5. CL - Confidence level
6. CMS - Compact muon solenoid
7. CP - combined performance
8. CR - control region(s)
9. DY - Drell-Yan
10. eCal - electromagnetic calorimeter
11. (f)JVT - (forward) Jet Vertex Tagger
12. FSR - final state radiation
13. GRL - good runs lists
14. hCal - hadronic calorimeter
15. HLT - High Level Trigger
16. IBL - insertable b -layer
17. ID - inner detector
18. IP - interaction point
19. ISR - initial state radiation
20. JER - Jet energy resolution
21. JES - Jet energy scale
22. L1 - Level 1 trigger
23. LAr - liquid argon
24. LHC - Large Hadron Collider
25. LHCb - Large Hadron Collider beauty
26. LO - leading order
27. MC - Monte-Carlo simulation
28. ME - Matrix element
29. MS - muon spectrometer

- 30. NF - normalization factor(s)
- 31. NLO - next-to-leading order
- 32. NNLO - next-to-next-to-leading order
- 33. N³LO - next-to-next-to-next-to leading order
- 34. OR - overlap removal
- 35. PFlow - particle flow
- 36. PMG - Physics Modeling Group
- 37. pp - proton–proton
- 38. PS - Parton shower
- 39. PU - pile-up
- 40. PV - primary vertex
- 41. SCT - semiconductor tracker
- 42. SF - same-flavor
- 43. SF - scale factor
- 44. SM - Standard Model
- 45. SR - signal region(s)
- 46. TRT - transition radiation tracker
- 47. VEV - Vacuum expectation value
- 48. WP - working point(s)

A Bibliography

- [1] ATLAS Collaboration. Luminosity public results run 2.
- [2] S. Höche. Introduction to parton-shower event generators, 2015.
- [3] LHC Higgs Cross Section WG. Lhc higgs cross section wg picture gallery.
- [4] J. Haffner. The cern accelerator complex. complexe des accélérateurs du cern, 2013. General Photo.
- [5] ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *JINST*, 3:S08003, 2008.
- [6] J. Pequeno. Computer generated image of the atlas inner detector, 2008.
- [7] J. Pequeno. Computer generated image of the atlas calorimeter, 2008.
- [8] J. Pequeno. Computer generated image of the atlas muons subsystem, 2008.
- [9] ATLAS Collaboration. Run 2 public daq plots.
- [10] S. Mehlhase. Atlas detector slice (and particle visualisations), 2021.
- [11] ATLAS Collaboration. Updates of the atlas tracking event data model (release 13), 2007.
- [12] ATLAS Collaboration. Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 lhc proton-proton collision data at $\sqrt{s} = 13$ tev. *Eur. Phys. J. C*, 79:639, 2019.
- [13] ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 2008.
- [14] M. Veltman. *Diagrammatica: The Path to Feynman Diagrams*. Cambridge University Press, 1994.
- [15] ATLAS Collaboration. Observation and measurement of higgs boson decays to ww^* with the atlas detector. *Phys. Rev. D*, 92:012006, 2015.
- [16] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Phys. Lett. B*, 716:1, 2012.
- [17] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Phys. Lett. B*, 716:30, 2012.

- [18] ATLAS Collaboration. Measurements of the higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 tev in the atlas experiment. *Eur. Phys. J. C*, 76(6), 2016.
- [19] ATLAS Collaboration. Measurements of gluon-gluon fusion and vector-boson fusion higgs boson production cross-sections in the $h \rightarrow ww^* \rightarrow e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Phys. Lett. B*, 789:508, 2019.
- [20] S. L. Glashow. Partial-symmetries of weak interactions. *Nucl. Phys.*, 22:579, 1961.
- [21] A. Salam. Weak and electromagnetic interactions. In *Conf. Proc. C*, page 367, 1968.
- [22] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19(21):1264, 1967.
- [23] G. 't Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44:189, 1972.
- [24] A. Pich. The standard model of electroweak interactions. In *2010 European School of High Energy Physics*, 2012.
- [25] J. Ellis. Higgs physics. In *2013 European School of High-Energy Physics*, 2015.
- [26] A. Djouadi. The anatomy of electro-weak symmetry breaking i: The higgs boson in the standard model. In ., 2008.
- [27] G. Altarelli. Collider physics within the standard model: a primer. In ., 2013.
- [28] S. F. Novaes. Standard model: An introduction. In *10th Jorge Andre Swieca Summer School: Particle and Fields*, 1999.
- [29] M. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2013.
- [30] E. Noether. Invariante variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, page 235, 1918.
- [31] J. Goldstone, A. Salam, and S. Weinberg. Broken symmetries. *Phys. Rev.*, 127(3):965, 1962.
- [32] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13(9):321, 1964.
- [33] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132, 1964.
- [34] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13(16):508, 1964.
- [35] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13(20):585, 1964.

- [36] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145(4):1156, 1966.
- [37] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155(5):1554, 1967.
- [38] V. N. Gribov and L. N. Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438, 1972.
- [39] L. N. Lipatov. The parton model and perturbation theory. *Yad. Fiz.*, 20:181, 1974.
- [40] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B*, 126:298, 1977.
- [41] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP*, 46:641, 1977.
- [42] J. et al. Pumplin. New generation of parton distributions with uncertainties from global qcd analysis. *JHEP*, 07:012, 2002.
- [43] A. Sherstnev and R. S. Thorne. Parton distributions for lo generators. *Eur. Phys. J. C*, 55:553, 2008.
- [44] R. D. et al. Ball. Parton distributions with lhc data. *Nucl. Phys. B*, 867:244, 2013.
- [45] J. et al. Butterworth. Pdf4lhc recommendations for lhc run ii. *J. Phys. G*, 43:023001, 2016.
- [46] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Phys. Rep.*, 97(2–3):31–145, 1983.
- [47] G. Marchesini and B. R. Webber. Simulation of qcd jets including soft gluon interference. *Nucl. Phys. B*, 238:1, 1984.
- [48] T. et al. Sjöstrand. An introduction to pythia 8.2. *Comput. Phys. Commun.*, 191:159, 2015.
- [49] J. et al. Bellm. Herwig 7.0/herwig++ 3.0 release note. *Eur. Phys. J. C*, 76:196, 2016.
- [50] T. et al. Gleisberg. Event generation with sherpa 1.1. *JHEP*, 02:007, 2009.
- [51] J. et al. Alwall. 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.
- [52] S. et al. Alioli. A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *JHEP*, 06:043, 2010.

- [53] S. et al. Höche. A critical appraisal of nlo+ps matching methods. *JHEP*, 09:049, 2012.
- [54] S. et al. Höche. Qcd matrix elements + parton showers: The nlo case. *JHEP*, 04:027, 2013.
- [55] S. et al. Catani. Qcd matrix elements + parton showers. *JHEP*, 11:063, 2001.
- [56] S. et al. Höche. Qcd matrix elements and truncated showers. *JHEP*, 05:053, 2009.
- [57] G. et al. Cowan. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. Erratum: *Eur. Phys. J. C* 73 (2013) 2501.
- [58] ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3:S08002, 2008.
- [59] TOTEM Collaboration. The totem experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08007, 2008.
- [60] MoEdAL Collaboration. The moedal experiment at the lhc - a new light on the terascale frontier. *Journal of Physics: Conference Series*, 631(1):012014, 2015.
- [61] S. Myers. The lep collider, from design to approval and commissioning. John Adams' Lecture, CERN, Geneva, 1991.
- [62] Fermilab. Design report tevatron 1 project, 1984. FERMILAB-DESIGN-1984-01.
- [63] F. Bordry. Lhc smashes targets for 2016 run.
- [64] ATLAS Collaboration. The atlas collaboration.
- [65] L. Evans and P. Bryant. Lhc machine. *JINST*, 3:S08001, 2008.
- [66] CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3:S08004, 2008.
- [67] LHCb Collaboration. The lhcb detector at the lhc. *Journal of Instrumentation*, 3(08):S08005, 2008.
- [68] ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3:S08002, 2008.
- [69] TOTEM Collaboration. The totem experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08007, 2008.
- [70] LHCf Collaboration. The lhcf detector at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08006, 2008.
- [71] MoEDAL Collaboration. The moedal experiment at the lhc - a new light on the terascale frontier. *Journal of Physics: Conference Series*, 631(1):012014, 2015.

- [72] ATLAS Collaboration. Performance of the atlas track reconstruction algorithms in dense environments in lhcb run 2. *Eur. Phys. J. C*, 77:673, 2017.
- [73] ATLAS Collaboration. Early inner detector tracking performance in the 2015 data at $\sqrt{s} = 13$ tev, 2015. ATL-PHYS-PUB-2015-051.
- [74] ATLAS Collaboration. Reconstruction of primary vertices at the atlas experiment in run 1 proton-proton collisions at the lhcb. *Eur. Phys. J. C*, 77:332, 2017.
- [75] ATLAS Collaboration. Vertex reconstruction performance of the atlas detector at $\sqrt{s} = 13$ tev, 2015. ATL-PHYS-PUB-2015-026.
- [76] ATLAS Collaboration. Improved electron reconstruction in atlas using the gaussian sum filter-based model for bremsstrahlung, 2012. ATLAS-CONF-2012-047.
- [77] ATLAS Collaboration. Electron identification measurements in atlas using $\sqrt{s} = 13$ TeV data with 50 ns bunch spacing, 2015. ATL-PHYS-PUB-2015-041, CERN, Geneva, Sep, 2015.
- [78] ATLAS Collaboration and H. Herde. Muon reconstruction performance in atlas at run 2, 2015. ATL-PHYS-PROC-2015-121, CERN, Geneva, Oct, 2015.
- [79] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 04:063, 2008.
- [80] ATLAS Collaboration. Jet energy measurement with the atlas detector in proton-proton collisions at $\sqrt{s} = 7$ tev. *Eur. Phys. J. C*, 73(3):2304, 2013.
- [81] ATLAS Collaboration. Tagging and suppression of pileup jets with the atlas detector, 2014. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014.
- [82] ATLAS Collaboration. Optimisation of the atlas b -tagging performance for the 2016 lhcb run, 2016. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016.
- [83] ATLAS Collaboration. Performance of missing transverse momentum reconstruction for the atlas detector in the first proton-proton collisions at $\sqrt{s} = 13$ TeV, 2015. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015.
- [84] ATLAS Collaboration. Expected performance of missing transverse momentum reconstruction for the atlas detector at $\sqrt{s} = 13$ TeV, 2015. ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul, 2015.
- [85] S. et al. Agostinelli. Geant4: A simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [86] P. Artoisenet et al. A framework for higgs characterisation. *JHEP*, 11:043, 2013.

- [87] CMS Collaboration. Search for a standard model higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method. *Eur. Phys. J. C*, 75(6):251, 2015.
- [88] J. Goldstone, A. Salam, and S. Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, 1962.
- [89] LHCf Collaboration. The lhcf detector at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08006, 2008.
- [90] S. Myers. The lep collider, from design to approval and commissioning. john adams’ lecture, 1991. CERN, Geneva, 1991.
- [91] Fermilab. Design report tevatron 1 project, 1984. FERMILAB-DESIGN-1984-01, FERMILAB, 1984.
- [92] F. Bordry. Lhc smashes targets for 2016 run.
- [93] ATLAS Collaboration. Particle identification performance of the atlas transition radiation tracker, 2011. ATLAS-CONF-2011-128, CERN, Geneva, Sep, 2011.
- [94] ATLAS Collaboration. The atlas inner detector commissioning and calibration. *Eur. Phys. J. C*, 70:787–821, 2010.
- [95] ATLAS Collaboration. A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using lh run 1 pp -collision data with the atlas detector.
- [96] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ tev using the atlas detector. *Eur. Phys. J. C*, 76:581, 2016.
- [97] T. et al. Plehn. A method for identifying $h \rightarrow \tau\tau \rightarrow e^\pm\mu^\pm +$ missing p_t at the cern lh. *Phys.Rev. D*, 61:093005, 2000.
- [98] D. et al. de Florian. Handbook of lh higgs cross sections: 4. deciphering the nature of the higgs sector, 2016.
- [99] M. et al. Boggia. The higgstools handbook: a beginners guide to decoding the higgs sector. *Journal of Physics G: Nuclear and Particle Physics*, 45:065004, 2018.
- [100] J. R. et al. Andersen. Les houches 2015: Physics at tev colliders standard model working group report, 2016.
- [101] N. et al. Berger. Simplified template cross sections - stage 1.1, 2019.
- [102] M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet. Atlas insertable b-layer technical design report, 2010. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Sep, 2010.

- [103] P. Czodrowski. The atlas trigger system: Ready for run 2. CDS:2039543.
- [104] L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.
- [105] Jerome H. Friedman. Stochastic gradient boosting. *Comput. Stat. Data Anal.*, 38(4):367–378, 2002.
- [106] Martín Abadi et al. Tensorflow: Large-scale machine learning on heterogeneous systems, 2015. (cit. on pp. 67, 95).
- [107] F. Chollet et al. Keras, 2015. (cit. on pp. 67, 95).
- [108] M. et al. Baak. Histfitter software framework for statistical data analysis. *Eur. Phys. J. C* 75, 153, 2015.