Search for a new resonance in the WZ leptonic decay channel in proton-proton collisions using the ATLAS detector

DISSERTATION

for obtaining the academic title

Doctor rerum naturalium

(Dr. rer. nat.)

Submitted to

School of Science

Faculty of Physics

at

Technische Universität Dresden

By

Abhishek Nag

born on 23.10.1993 in India

Dresden, August 2022

Submitted on 11 August 2022

1. Reviewer: Prof. Dr. Michael Kobel
2. Reviewer: Em. Prof. Dr. Charikleia Petridou

Abstract

Although, the Standard Model of particle physics (SM) is currently the best model we have to describe the fundamental particles of the universe, it still has limitations. Some of the known limitations of the SM are in the description of neutrinos, dark matter, the low mass of the Higgs boson which leads to hierarchy and naturalness problems, among others. Therefore, there is a need for physics beyond SM. A simple way to discover new physics and particles are resonance searches. The search for a resonance decaying into dibosons is a powerful probe for electroweak symmetry breaking, one of the predictions at the hearth of the SM.

A search for a heavy resonance decaying into WZ in the fully leptonic channel (electrons and muons) is performed in this work using the proton-proton collision data collected by the ATLAS experiment at the Large Hadron Collider at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 139 fb^{-1}. This search fits the LHC objectives as this search is most sensitive at low masses and complements the searches performed in hadronic channels. Limits on the production cross-section times branching ratio are obtained as a function of the resonance mass for resonances arising from two different models, the Heavy Vector Triplet and Georgi-Machacek model. Two different production modes are considered, the vector boson fusion and the Drell-Yan, on which independent limits are set.

Acknowledgement

I feel overwhelmed as I think of the people who have loved, supported, taught and guided me over the past few years before and during my PhD journey at TU Dresden. I had the pleasure of working in an incredibly friendly and optimistic environment of the ATLAS Collaboration which kept me motivated to learn, to adapt and improve everyday.

I would like to thank my supervisor Michael Kobel for giving me this prestigious opportunity and supporting me with his guidance and encouragement. He has been very kind in showering his wisdom upon me time and again.

I would also like to thank Dr. Suchandra Dutta of Saha Institute of Nuclear Physics, Dr. Ananda Dasgupta and Dr. Ritesh Singh of IISER Kolkata for guiding me during my Masters and supporting me in my pursuit of getting a PhD position in Experimental Particle Physics.

This work would not have been possible without the support of past and current lovely VBS group members in Dresden. Thanks a lot to Carsten Bittrich, Tim Herrmann, Franziska Iltzsche, Orcun Kolay, Joany Manjarres, Jan-Eric Nitschke, Max Stange and Stefanie Todt. Special thanks once again to Carsten and Tim for giving their valuable time and guiding me through the vast pool of codes and knowledge which are available for doing physics analysis in the ATLAS Collaboration. A very special thanks to Joany Manjarres for mentoring and supporting me very closely during my entire thesis work.

I also want to thank my collaborators at CERN, especially Georges Azuelos, Robert Les and Miaoran Lu, who have been very generous and kind and always provided me with all possible help or information that was needed. I look back very proudly on our final work. I want to thank everyone who helped shape this thesis by providing comments and suggestions: Georges Azuelos, Samprikta Das, Tim Herrmann, Orcun Kolay, Joany Manjarres, Jan-Eric Nitschke, Frank Siegert and Max Stange.

Very special thanks to my very close friends Anwesha Das and Mainak Mandal, who were there for me during my highs and lows of my PhD journey in Dresden, talking and discussing over dinner and making me smile. Also, thanks to Bishawjit Mallick and Dipika Das for treating me like their younger brother in a distant land. Also thanks to all my friends in Dresden.

Last but not the least, a very very special thanks to my parents: my lovely mother and my caring father. Thank you Maa and Baba for being my support system all throughout my life, for always motivating and encouraging me in my endeavours. You are my superpower and my suit of armour.

Dedicated to my lovely parents

Contents

Chapter 1

Introduction

"Faith is my sword. Truth is my shield. Knowledge my armor"

Doctor Strange

Elementary Particle Physics is a branch of physics that studies the nature of the irreducibly smallest particles that are the fundamental building blocks of matter and their interactions. It is also called "high energy physics", because many elementary particles do not exist naturally, but are created only during high energy collisions of other particles, as can be detected in particle accelerators. Over the last century, the goal of particle physics has been to identify these fundamental elementary particles and their interactions. Since its formulation in the 1960s and 1970s, the Standard Model of particle physics provides the theoretical foundation to explain these fundamental particles and fields along with their dynamics.

The idea that all matter is fundamentally composed of elementary particles dates from at least the 6th century BC. 400 years ago, Galileo started piecing together the basic principles of reality, what we now call Modern Science. But the questions he was trying to answer are as old as humanity itself. What is the world made of? What are the fundamental building blocks of the universe from which we, stars and everything else is constructed. In the centuries since Galileo, thousand of theories and experiments have peered into smaller and smaller distances, converging on a single picture of the structure of matter called the Standard Model.

The Standard Model of particle physics (SM) is currently the best model we have to describe the fundamental particles and their interactions in the universe. It has quite successfully described and predicted the existence of many particles and phenomena like the Brout-Englert-Higgs mechanism[1, 2] and the resulting Higgs bosons and mass of W and Z bosons as explained in Chapter 2. But there are still some limitations of the theory. The SM doesn't include the Gravitational force which is one of the fundamental forces in nature and it doesn't describe dark matter and dark energy which comprises $\approx 95\%$ of the available mass and energy content of the universe as suggested by cosmological observations. Also, the mass of the Higgs Boson, discovered by the ATLAS and CMS experiment at CERN in 2012 [3, 4] is much lighter than the Planck scale. This is called the Hierarchy problem. Therefore, we need

to look for particles that can show signs of the existence of physics beyond the SM. Over the years, many theories have been formulated to incorporate these shortcomings, like Supersymmetry, but they haven't yet been proven experimentally. In this thesis, resonances are used to search for physics beyond SM.

In Particle Physics, a resonance is a peak observed in the total or differential cross-section as function of energy or mass of a scattering experiment. If there is a new particle that can be described by beyond SM physics, a peak can be observed above the SM contribution in the invariant mass distribution. In the past, resonances have been helpful in discoveries like the Z boson in 1983 at the Super Proton Synchrotron at CERN [5]. In this analysis, a search is performed for a resonant particle which decays to a W and Z boson pair which further decays into leptons, i.e. electrons, muons, taus and neutrinos. The fully leptonic channel is found to be more sensitive to resonances with mass below $\sim$ 1 TeV because of the low background, in spite of the low branching ratio. Proton-proton collisions are performed at the Large Hadron Collider at CERN at the center of mass energy of 13 TeV and the data is collected by the ATLAS detector detailed in Chapter 3. The reconstruction of the objects from the detector output that are used in this analysis are detailed Chapter 4. The full RunII data collected over four years from 2015 to 2018 is being analyzed, detailed in Section 5.1. Also, Monte Carlo (MC) generated sample is produced, described in Section 5.2, to mimic the different SM processes and compared with data to make sure we understand our data correctly.

Searches for a W' in an extended gauge model, decaying to WZ in the fully leptonic mode, at $\sqrt{s} = 8$ TeV with 20 fb^{-1} of data have been performed by the ATLAS [6] and CMS [7] Collaborations. The present analysis extends searches for resonant WZ production, performed by ATLAS in Run 2 of the LHC using pp collision data at $\sqrt{s} = 13$ TeV [8], with 36 fb^{-1} of integrated luminosity. In this thesis, an attempt in search for heavy resonances decaying into W and Z boson which further decay to leptons in the ATLAS detector has been made with larger dataset with respect to previous publications and using Neural Network.

The object and event selection criteria required for this analysis are explained in Chapter 6. In this analysis we are looking for new resonant particles for which we define signal regions enriched in signal like events. The strategy to define the signal regions are detailed in Chapter 7. The SM processes which are the background in this search are explained in Chapter 8 and the uncertainties arising from the modelling of these background processes and from detector effects are described in Chapter 9. A statistical search is performed using data and MC which is described in Chapter 10. The results of this search and conclusion are finally presented in Chapter 11 and Chapter 12.

Chapter 2

Theoretical Foundation

"If we can't accept limitations, we're no better than the bad guys."

Captain America: Civil War

The goal of particle physics is to understand the elementary, tiniest constituents of matter and the fundamental interactions between them. The Standard Model (SM) is a widely accepted theory that has been quite successful in describing coherently most of the particle physics knowledge obtained so far from experiments. But it is probably incomplete, as there are many parameters and there are also some phenomena that the SM has not been able to describe. For example, the description of neutrinos, dark matter, the low mass of the Higgs boson and it does not incorporate gravity.

In this chapter, the foremost vital features of the Standard Model and its predictions are discussed. Then, the limitations and therefore the motivation for a search for physics beyond SM are also presented. Finally, the possible theories that are used as benchmark models for the search for physics beyond the Standard Model in this analysis are described.

2.1 The Standard Model

The Standard Model emerged during the 1960s and 1970s to explain the elemental interactions and particles known in nature. The SM is based on quantum field theory [9] it combines the principles of quantum mechanics and special relativity. At the fundamental level, space is permeated by fields and particles are quantum fluctuations of the fields.

Group theory provides a mathematical framework to describe the symmetries of the Langrangian. The SM must be invariant under transformations under certain symmetry groups. First, Langrangian is invariant under Lorentz transformations, which are rotations or relativistic boosts. The symmetries are used to derive the underlying properties of interactions. Global gauge symmetries lead to conserved quantities, called quantum numbers. Gauge fields follow from local gauge invariance, i.e. the transformation under which invariance occurs, is a function of space-time.

There are four fundamental interactions known as gravity, electromagnetic, weak, and strong interaction. Gravity is best described by the theory of general

Mediator	Interaction	Spin	Charge	Mass
γ	Electromagnetic	1	0	0
$W^{\pm}$	Weak	1	$\pm$ 1	80.379 $\pm$ 0.012 GeV
Z	Weak	1	0	91.1876 $\pm$ 0.0021 GeV
g	Strong	1	0	0
H		0	0	125.10 $\pm$ 0.14 GeV

Table 2.1: Overview of the mediators (bosons) of fundamental interactions in Standard Model.

relativity, which is not part of the Standard Model. The theory of strong interactions, Quantum Chromodynamics (QCD) [10–12] was proposed in 1973 to describe the formation of nuclei. The electroweak theory [13–15], which unifies the electromagnetic and weak interactions, was proposed between 1961-1967.

The SM additionally must satisfy the conditions of unitarity and renormalisability. Unitarity refers to the fact that the sum of all calculated probabilities of a process must be equal to one, while renormalisability refers to the ability to predict physical interactions at all energies.

2.1.1 Particle content

According to the Standard Model, everything in the universe is made up of 12 different types of matter particles, interacting with 3 forces, electromagnetic, weak and strong interactions. Every particle is either a fermion, which is a matter particle, or a boson, which is a force particle or mediator. The distinction between the fermions and bosons lies in the quantum world. Fermions have half-integer spins and must obey the Pauli exclusion principle, this means that two or more identical fermions cannot occupy the same quantum state within a quantum system simultaneously. Bosons, on the other hand, have integer spins and are not obliged to obey the Pauli exclusion principle.

The interactions between fundamental particles are mediated by the quanta of the gauge fields. These mediators are summarised in table 2.1. The electroweak interaction is mediated by a massless photon and three massive gauge bosons, namely, W^+, W^-, and Z, which are linear combinations of the quanta of underlying SU(2) $\otimes$ U(1) symmetry groups. They are called bosons because they obey Bose-Einstein statistics and have integer spin quantum numbers. The strong interaction is initiated by eight gluons. In the SM, all these gauge particles have spin 1 and are referred to as vector bosons. The quantum numbers associated with the SU(2) gauge groups are called weak isospin I and for U(1) are weak hypercharge Y. They are related to the electrical charge by:

$$Q = T_3 + Y \tag{2.1}$$

where T_3 is the third component of the weak isospin. In addition to these gauge bosons, there is another boson in the SM, called the Higgs boson, which has spin 0 and arises due to a mechanism that introduces masses, as explained in 2.1.5.

All matter is made up of spin 1/2 particles, called fermions, which are identified by their electroweak quantum numbers and by whether they participate in

Generation	Family	Name		Charge Q [e]	Mass
1st	Quarks	u	up	2/3	2.16 MeV
		d	down	-1/3	4.67 MeV
	Leptons	e	electron	-1	0.511 MeV
		ν_e	e-neutrino	0	<1.1 eV
2nd	Quarks	c	charm	2/3	1.27 GeV
		s	strange	-1/3	93 MeV
	Leptons	μ	muon	-1	0.1057 GeV
		ν_μ	μ-neutrino	0	<1.1 eV
3rd	Quarks	t	top-quark	2/3	172.9 GeV
		b	bottom-quark	-1/3	4.18 GeV
	Leptons	τ	tau	-1	1.777 GeV
		ν_τ	τ-neutrino	0	<1.1 eV

Table 2.2: Overview of the elementary matter particles, .i.e. fermions, in Standard Model.

strong interactions. These fermions are summarised in table 2.2. All fermions participate in weak interactions in the SM. Fermions that do not participate in strong interactions are color neutral (as color is the charge for strong interaction) and are called *leptons* and those which interact with strong interactions are called *quarks*.

At everyday low energies, everything is reduced to just three matter particles: an *electron* and two species of quarks, namely, the *up* quark and the *down* quark. The protons and neutrons each contain three quarks: the protons have two *up* quarks and a *down* quark, while the neutron has two *down* quarks and an *up* quark. Protons and neutrons together make up the nucleus, electrons and a nucleus make an atom, and a collection of bound atoms together make up the matter. All the beauty and complexity that we see in the world around us can be traced to this same collection of three particles, rearranged over and over in different combinations.

Leptons without an electromagnetic charge are called neutrinos. Neutrinos are extremely light and barely interact with anything else. Millions of neutrinos per cubic meter have been streaming through the universe uninterrupted since the first few seconds after the Big Bang.

So, there are basically four different types of matter particles. Three of them make up all the matter around us and the fourth, electron neutrino. For reasons unknown, nature took this collection of these four particles and made at least two further copies. These have the same exact quantum numbers as the original particle, except that they are heavier, e.g. the *muon* is 200 times heavier than the *electron* while *tau* is almost three and half thousand times heavier. The same generational pattern then repeats for the quarks. There are two heavier versions of the *down* quark, called *strange* and *bottom* quarks, and two heavier versions of the *up* quark called *charm* and *top* quarks. Then there are also two more neutrinos, the muon neutrino (ν_μ), and tau neutrino (ν_τ). The second and third-generation particles are not seen in everyday life. They can be created, but they are unstable and they quickly decay to their first generation of particles. Each fermion (with spin 1/2) has an associated anti-particle, which has the opposite electric charge, opposite colour charge, and opposite weak isospin charge, are not listed in Table 2.2.

2.1.2 Gauge Theory

As mentioned above, the Standard Model is a gauge field theory based on the symmetry group $SU(3)\otimes SU(2)\otimes U(1)$ in the framework of relativistic quantum field theory. In classical physics, symmetry is a transformation that preserves measurable physical properties.

The particles are described as quantum fields where their interactions and propagation are described in the Lagrangian formalism. A detailed explanation of this formalism can be found in [16, 17]. The Euler-Lagrange equations provide a recipe on how to derive the equations of motion for a system given the Lagrangian. The equations of motion for a freely propagating fermion $\phi(\vec{x},t)$ with mass m applied to a Lagrangian density:

$$\mathcal{L} = \bar{\phi}(i\gamma^\mu\partial_\mu - m)\phi \tag{2.2}$$

where γ^μ are the Dirac matrices connecting the spinor representation to the vector representation of the Lorentz group.

The Lagrangian density is invariant under global gauge transformations of the fermion field, i.e., $\phi \rightarrow e^{-i\alpha}\phi$, where α is a real phase factor. According to Noether's theorem, there is an associated conserved quantity for every continuous symmetry and vice versa. For a local gauge transformation, the phase factor α is allowed to depend on the space-time coordinates $x = \{\vec{x},t\}$:

$$\phi(\vec{x},t) \rightarrow e^{-i\alpha(\vec{x},t)}\phi(\vec{x},t) \tag{2.3}$$

To make the Lagrangian in eq. 2.2 invariant under a local gauge transformation, a vector field A_μ, also called gauge field is introduced with the desired transformation properties

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha \tag{2.4}$$

The A_μ couples to fermionic fields $\phi(\vec{x},t)$ with strength e. A covariant derivative is defined which is specific to local gauge transformation by:

$$D_\mu = -\partial_\mu - ieA_\mu \tag{2.5}$$

The symmetry in eq. 2.3 corresponds to $U(1)$ gauge symmetry and the gauge field A_μ can be interpreted as the photon field. To keep gauge invariance of A_μ a new term is included in the Lagrangian:

$$\mathcal{L}_{kinetic} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.6}$$

where, $F^{\mu\nu}$ is the field strength tensor defined as:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2.7}$$

With the above modifications and inclusions, the Lagrangian density in 2.2 becomes:

$$\mathcal{L}_{QED} = \bar{\phi}(i\gamma^\mu D_\mu - m)\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.8}$$

This is the Lagrangian density of quantum electrodynamics (QED). Similarly, the Lagrangian density of quantum chromodynamics (QCD) can be derived

by requiring gauge invariance under local $SU(3)_C$ transformation and the Lagrangian density of electroweak interaction theory can be obtained by requiring symmetry under local $SU(2)_L \otimes U(1)_Y$ symmetry.

2.1.3 Quantum Chromodynamics

The massive spin 1/2 fermions are the fundamental building blocks of matter. The interaction between quarks is described by the theory of strong interactions known as quantum chromodynamics (QCD) and is obtained by requiring symmetry under local $SU(3)_C$ gauge transformation on quark colour field q. The Lagrangian density of free colour fields is:

$$\mathcal{L} = \bar{q}_j(i\gamma^\mu \partial_\mu - m)q_j \tag{2.9}$$

where, $j \in$(red, green, blue) refers to the colour charge of the quark. The local gauge invariance refers to the transformation:

$$q_j \rightarrow e^{i\alpha_a(x)\frac{\lambda_a}{2}} q_j \tag{2.10}$$

where, λ_a with $a = 1, \cdots, 8$ are the 8 linearly independent, traceless Gell-Mann matrices. The λ_a matrices do not commute with each other

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc}\frac{\lambda_c}{2}$$

where, f_{abc} is the structure constant of $SU(3)$. QCD is a non-abelian gauge theory as the generators are non commuting. To make the Lagrangian local gauge invariant, the following covariant derivative is defined

$$D_\mu = \partial_\mu + i g_s \frac{\lambda_a}{2} G^a_\mu \tag{2.11}$$

where, G^a_μ are the gluon fields. This covariant derivative couples the quark field with a coupling strength g_s to the gluon fields via the generators $\lambda_a/2$. There are 8 gluon fields for the phase variation of three quark colour and they transform as:

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g_s}\partial_\mu \alpha_a - f_{abc}\alpha_b G^c_\mu \tag{2.12}$$

The Lagrangian density in eq.2.9 is modified:

$$\mathcal{L}_{QCD} = \bar{q}_j(i\gamma^\mu D_\mu - m)q_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{2.13}$$

where $G^a_{\mu\nu}$ is the field strength tensor defined by:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu$$

The colour states can be combined in different ways to form a colour neutral state. Colour-neutral bound states of quarks are called Hadrons. Mostly the hadrons are either mesons, which is a combination of a quark and an anti-quark, or baryons with three quarks. At high probe energies, these hadrons exhibit a structure of virtual quark-antiquark pairs from vacuum fluctuations and their composition can be determined on a probabilistic basis, summarized in parton distribution functions (PDF).

2.1.4 Electroweak theory

Weak forces are responsible for radioactive beta decay, where, for example, a neutron turns into a proton by emitting an electron and an electron antineutrino. In the 1930s, the field of quantum mechanics was able to produce a relativistic equation of motion for electrons via the Dirac equation, but the nuclear decay processes remained a mystery. Fermi initially tried to explain the beta decay by quantum physics as a direct contact interaction between the four fermions, which was motivated by the apparent extreme short-range interaction hence the name "weak" interactions. But this model only worked at low energies. Meanwhile, quantum electrodynamics was developed with the interaction mediated by photons. In 1957, Julian Schwinger [18] proposed a set of force mediating gauge bosons for the weak interaction. In 1954, Yang and Mills [19] developed a non-Abelian gauge theory based on the $SU(2)$ symmetry, which was extended in 1961 by Glashow [13] to a $SU(2) \otimes U(1)$ group, that should describe both the electromagnetic and weak interactions.

Experimental observations suggest that the weak interaction affects left-handed fermion fields only (and right-handed anti-fermion fields). The electromagnetic and weak interactions are described in a unified electroweak theory with $SU(2)_L \otimes U(1)_Y$ gauge symmetry. The $U(1)_Y$ symmetry can be derived following the example of electrodynamics and $SU(2)_L$ symmetry in isospin space is motivated via the weak interaction. This gauge group has four associated massless gauge bosons $W^a_\mu, a = 1, 2, 3$, and B_μ for the $SU(2)_L$ and $U(1)_Y$ groups respectively as well as the corresponding gauge coupling constants g and g'. The fermion fields are represented by the left-handed doublets χ_L and right-handed singlets ψ_R.

For the first generation leptons and quarks:

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \qquad \text{and} \qquad \chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L , \tag{2.14}$$

$$\psi_R = e_R \,\text{and}\, \nu_{eR} \qquad \text{and} \qquad \psi_R = u_R \,\text{and}\, d_R \tag{2.15}$$

where the corresponding fermions of the second and third generation follow accordingly. The Lagrangian density for these fermion fields is required to be invariant under local gauge transformations corresponding to both the $SU(2)_L$ and $U(1)_Y$ symmetry groups:

$$\chi_L \rightarrow e^{i\beta(x)Y + i\alpha_a(x)\tau_a}\chi_L \tag{2.16}$$

$$\psi_R \rightarrow e^{i\beta(x)Y}\psi_R \tag{2.17}$$

where $\alpha(x)$ and $\beta(x)$ are the phase transformation factors of the $SU(2)_L$ and $U(1)_Y$ symmetry groups. The Pauli matrices $\tau_a, a = 1, 2, 3$ and the weak hypercharge operator Y are the generators of the groups. The gauge-invariant Lagrangian is:

$$\begin{aligned} \mathcal{L}_0 = {} & \bar{\chi}_L \gamma^\mu \Big[i\partial_\mu - g\frac{\tau_a}{2} W^a_\mu + \frac{g'}{2} B_\mu \Big] \chi_L \\ & + \bar{\psi}_R \gamma^\mu \Big[i\partial_\mu + g' B_\mu \Big] \psi_R - \frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \end{aligned} \tag{2.18}$$

where, $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors. In order to build the gauge-invariant kinetic term for the gauge fields, field strengths are introduced as:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.19}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc} W^b_\mu W^c_\nu \tag{2.20}$$

$B_{\mu\nu}$ remains invariant under gauge transformations, while $W_{\mu\nu}$ transforms covariantly. The field strengths $W^a_{\mu\nu}$ contain quadratic terms giving rise to cubic and quartic self-interaction between the vector fields W^a_μ, $a = 1, 2, 3$. The Feynmann diagrams for cubic and quartic interactions are shown in Figure 2.1 [20]. The strength of these interactions is given by the same $SU(2)_L$ coupling g which appears in the fermionic piece of the Lagrangian. The boson interaction part of the Lagrangian is:

$$\mathcal{L}_{Bosons} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a \tag{2.21}$$

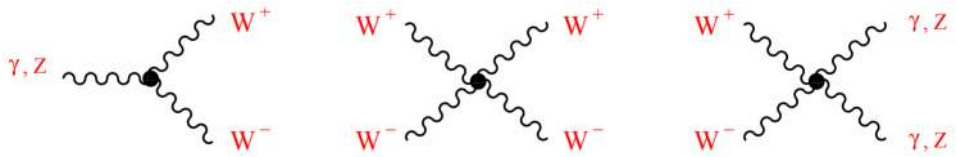


Figure 2.1: Feynmann diagram for gauge boson self-interaction vertices

The term containing the $SU(2)_L$ matrix gives rise to charged-current interactions with the boson field $W^\pm_\mu \equiv (W^1_\mu \mp i W^2_\mu)/\sqrt{2}$ and its complex conjugate. The charged current Lagrangian for a single generation of quarks and leptons is

$$\mathcal{L}_{CC} = \frac{g}{2\sqrt{2}} \Big\{ W^\dagger_\mu [\bar{u}\gamma^\mu (1-\gamma_5) d + \bar{\nu}_e \gamma^\mu (1-\gamma_5) e] + h.c. \Big\} \tag{2.22}$$

The universality of the quark and lepton interactions is now a direct consequence of the assumed gauge symmetry.

The Lagrangian in eq. 2.18 also contains interactions with the neutral gauge fields W^3_μ and B_μ, these bosons are identified as Z and γ. Since both fields are neutral, they can be arbitrarily combined as:

$$\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \tag{2.23}$$

In terms of field Z and γ, the neutral-current Lagrangian is given by

$$\begin{aligned} \mathcal{L}_{NC} = \sum_j \bar{\psi}_j \gamma^\mu \Big\{ & A_\mu \Big[g \frac{\tau_3}{2} \sin\theta_W + g' Y \cos\theta_W \Big] \\ & + Z_\mu \Big[g \frac{\tau_3}{2} \cos\theta_W - g' Y \sin\theta_W \Big] \Big\} \psi_j \end{aligned} \tag{2.24}$$

In order to get QED from the A_μ, the following condition needs to be imposed:

$$g \sin(\theta_W) = g' \cos(\theta_W) = e \quad \text{and} \quad Y = Q - T_3 \tag{2.25}$$

where, $T_3 = \tau_3/2$ is the weak isospin operator associated to $SU(2)_L$, Q is the electromagnetic charge and Y is the hypercharge associated to $U(1)_Y$. The

above equation 2.25 relates the $SU(2)_L$ and $U(1)_Y$ couplings to the electromagnetic coupling, providing the unification of the electroweak interactions.

The Lagrangian in eq. 2.18 and eq. 2.20 describes an interaction between massless fermion fields and massless gauge fields which is in contradiction to the experimental observation of massive fermions and bosons in Tables 2.1 and 2.2.

2.1.5 Higgs Mechanism

The Lagrangian in equation 2.18 describes massless bosons and fermions, while experimentally it has been shown that the physical $W^{\pm}$ and Z bosons should be quite heavy. The mass of fermions also needed to be explained. In order to generate masses, the Brout-Englert-Higgs (BEH) mechanism was proposed in 1964 [1, 2, 21, 22]. This mechanism proposes a new field with a corresponding potential assigned to it. The overall potential and thus the Lagrangian is gauge invariant, but some symmetries are broken in the ground state.

In the BEH mechanism, two scalar complex fields ϕ^+ and ϕ^0 are introduced and combined in a $SU(2)_L$ doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.26}$$

The SM Lagrangian is extended by a new term $\mathcal{L}_{BEH}$

$$\mathcal{L}_{BEH} = (D_\mu \phi)^\dagger (D^\mu \phi) - (\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2) \tag{2.27}$$

where the first term describes the kinematic of the new fields and $V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$ is the new BEH potential. This Lagrangian is invariant under $SU(2)_L \otimes U(1)_Y$ transformations. There is a infinite set of degenerate states with minimum energy when $\mu^2 < 0$ and $\lambda > 0$, satisfying

$$|\phi_0| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}} \tag{2.28}$$

where v is the vacuum expectation value of the BEH field. By choosing a particular solution as a ground state, the symmetry is spontaneously broken, hence the name spontaneous symmetry breaking (SSB). One can freely choose a convenient value, called vacuum state ϕ_0,

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.29}$$

and general expansion around this vacuum state can be transformed to

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.30}$$

The number of degrees of freedom in the ϕ field is reduced from four, for two complex fields ϕ^+ and ϕ^0, to one, for the expansion around the minimum. The choice of the minimum spontaneously breaks the $SU(2)_L \otimes U(1)_Y$ symmetry to the electromagnetic $U(1)_Q$ and this mechanism is called "Electroweak symmetry breaking" (EWSB). According to the Nambu-Goldstone theorem [23,

24], there exists a massless scalar boson, typically referred to as Goldstone bosons, for each of the broken symmetries. The Goldstone bosons are "eaten" by the massless W and Z bosons, giving them a longitudinal degree of freedom and hence a mass. (A massless boson has only transverse polarization, like a photon, but a massive boson also has a longitudinal polarization.)

2.1.5.1 Boson masses and interactions

The covariant derivative in eq. 2.27 couples the scalar multiplet to the $SU(2)_L \otimes U(1)_Y$ gauge bosons. The Lagrangian is expanded around the chosen minimum rather than the Higgs field H to derive the masses of the gauge bosons. The Higgs field H is required to be electrically neutral and the choice of $SU(2)_L$ doublet makes the hypercharge of H be $Y_H = 1$. The kinetic piece of the Lagrangian in eq. 2.27 takes the form:

$$\begin{aligned}(D_\mu\phi)^\dagger(D^\mu\phi) = &\frac{1}{4}(v+H)^2 g^2 W_\mu^\dagger W^\mu \\ &+ \frac{1}{8}(g^2+g'^2)(v+H)^2 Z_\mu Z^\mu\end{aligned} \tag{2.31}$$

The vacuum expectation value of the neutral scalar has generated a quadratic term for the $W^\pm$ and the Z, i.e. those gauge bosons have acquired masses:

$$M_Z \cos\theta_W = M_W = \frac{1}{2}vg \tag{2.32}$$

while for the photon no mass term is introduced, thus:

$$M_A = 0 \tag{2.33}$$

Therefore, a clever way is formulated which gives masses to the weak force mediators by simply adding the isodoublet complex scalar fields 2.26 to the Lagrangian in $SU(2)_L \otimes U(1)_Y$ model. The total Lagrangian also remains invariant under gauge transformations, and, SSB occurs. A new prediction of the BEH mechanism is the existence of a massive scalar boson, the Higgs boson H whose mass was a free parameter and was found experimentally.

In addition to the mass terms, terms with the structure HVV or $HHVV$ with $V = W^\pm, Z$ are derived. These terms describe three-point and four-point interactions between the Higgs field and gauge boson pairs W^+W^- or ZZ. The coupling strength is proportional to the mass of the gauge boson squared, and the Feynman diagrams for these interactions are shown in Figure 2.2 [20]

2.1.5.2 Fermion masses

The fermion mass term has the form:

$$\mathcal{L}_{fermion} = -m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \tag{2.34}$$

The term violates $SU(2)_L$ gauge symmetry since left-handed fermions are doublets and right-handed fermions are singlets. But with the introduction of an additional scalar doublet into the model, additional terms are added to the Lagrangian. The terms describing the interaction between the leptons and the BEH field

$$\mathcal{L}_{YL} = y_l^i(\bar{L}^i\phi e_R^i + \phi^\dagger \bar{e_R^{-i}} L^i) \tag{2.35}$$

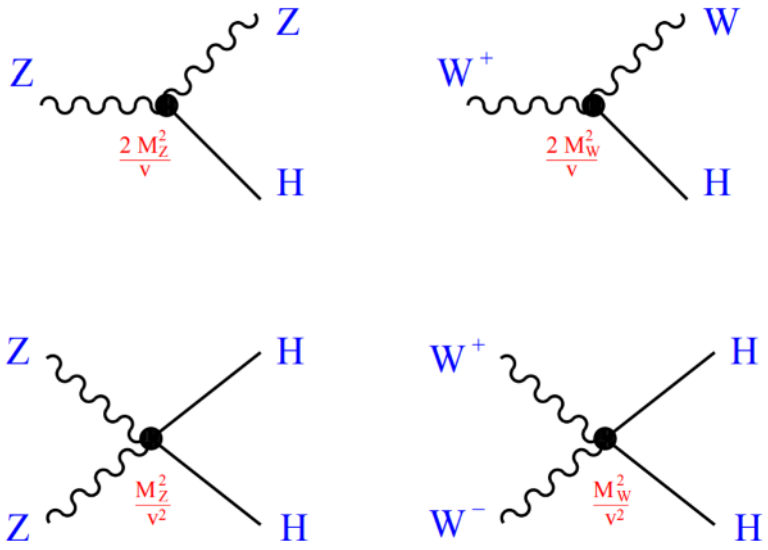


Figure 2.2: Feynmann diagram of the Higgs coupling to the gauge bosons

where, $i = 1, 2, 3$ sums over the generations and y^i are the arbitrary coupling parameters. This makes the Lagrangian now gauge invariant. After SSB, in unitary gauge, this Yukawa Lagrangian gives the lepton mass terms:

$$m_l = y_l \frac{v}{2} \tag{2.36}$$

the coupling strength for lepton and the Higgs boson is given by m_l/v and the free parameters y_l can be chosen to add mass terms for each lepton. Terms for neutrino masses are not included in the model by definition. Additional terms to produce mass terms for quarks

$$\mathcal{L}_{YQ} = Y_d^{jk} \bar{Q}_L^j \phi d_R^k + Y_u^{jk} \bar{Q}_L^j i\tau_2 \phi^* u_R^k + \text{h.c.} \tag{2.37}$$

where $j, k = 1, 2, 3$ sums over the generations and the Yukawa matrices Y_d^{jk} and Y_u^{jk} are complex matrices which give the correct mass term for each quark. Therefore, by the SSB mechanism, one can also get the quark mass terms:

$$m_d = y_d \frac{v}{2}, \quad m_u = y_u \frac{v}{2} \tag{2.38}$$

where y_d and y_u are free parameters.

Therefore, the final Standard Model Lagrangian after obtaining the missing terms for the massive particles can be written as:

$$\mathcal{L}_{SM} = \mathcal{L}_{QED} + \mathcal{L}_{QCD} + \mathcal{L}_{Bosons} + \mathcal{L}_{BEH} + \mathcal{L}_{Yukawa} \tag{2.39}$$

where $\mathcal{L}_{Yukawa} = \mathcal{L}_{YL} + \mathcal{L}_{YQ}$. The Feynmann rules to describe all interactions and their strengths are derived from this Standard Model Lagrangian $\mathcal{L}_{SM}$

2.1.6 Limitations

Although the Standard Model has been a quite successful theory that has been tested by numerous experiments, it is still not the ultimate theory that

describes everything in nature. There are several shortcomings in this model that motivate the search for physics beyond Standard Model. Some of these issues are:

1. The Standard Model doesn't include the gravitational force, which is one of the four forces in nature.There are 2 good reasons why gravity is not included in the SM. First, at a microscopic level, the force of gravity is so weak that it barely has any effect on a single subatomic particle. Second, it is not really known how to incorporate general relativity, which is a classical theory, into the quantum world. We have no idea how to probe into a black hole where quantum gravity effects are at work.

2. The Standard Model describes only 5% of the available mass and energy content of the universe suggested by Cosmological observations. The other 95% is dark matter($\approx 23\%$) and dark energy($\approx 72\%$).

3. The SM (in its initial formulation) predicts massless and stable neutrinos. Recent discoveries have shown that neutrinos oscillate between flavours and it is possible only if they aren't massless.

4. There are many ad-hoc parameters in the SM, like the masses, and number of generations.

5. In SM there is the hierarchy problem which arises due to the low mass of the Higgs. The virtual corrections to the mass of the Higgs boson diverge and therefore very fine-tuning is required to have a low mass for the Higgs. Supersymmetry and Composite Higgs models try to address the hierarchy problem [25].

All these limitations indicate that the SM is not the ultimate theory and therefore we need beyond SM physics to describe nature

2.2 Beyond Standard Model

It is an important and stimulating task to ensure proper communication between theory and experiment, particularly related to the hypothetical TeV scale extensions of the SM. There are interesting and well-motivated inclusive frameworks that are defined as sets of general assumptions about the BSM physics but cannot be interpreted as a single complete model. Within each framework, many models can be constructed with definite predictions. There is no clear justification for selecting one model and one would need all of them to be compared to data. It would be very tedious as this will further require many different models with many free parameters that need to be scanned for new physics analysis.

The search for a general BSM physics can be addressed if we restrict our search to direct experimental observations of new physics which consists of the creation of narrow resonances. Resonance is one of the simplest methods to look for physics beyond the Standard Model (BSM) and to discover new particles. It is described as a statistically significant bump over a smooth SM background. In this thesis, two main production processes for the resonance

are considered, one by quark fusion known as Drell-Yan (DY), and the other by vector boson fusion (VBF). In VBF processes, a quark and/or anti-quark pair scatters through the exchange of an electroweak gauge boson, γ, Z or W, resulting in the production of a heavy resonance.

In this thesis, a search for a $W^{\pm}Z$ resonance produced by either a DY or VBF process is done with both bosons decaying leptonically. Therefore, in the final state, there are three leptons and one neutrino. For the VBF process, there are additional quark (anti-quark) jets pair produced in the opposite direction. Figure 2.3 shows a topology for a VBF process.

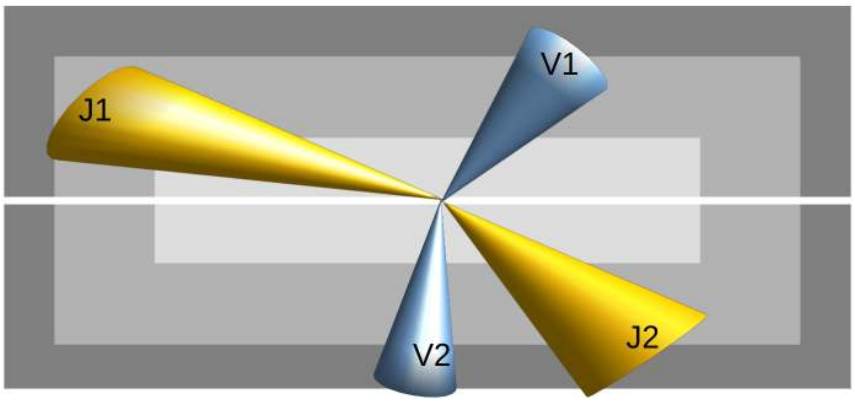


Figure 2.3: Illustration of VBF topology in the LHC. J1 and J2 are the two quark (anti-quark) jets and V1 and V2 are the two bosons.

One BSM model is the so-called "Simplified Model" which has been used in Supersymmetry searches. Another method could be extensions of the present Standard Model that contain scalars in triplet or larger isospin representations and that preserve the custodial symmetry in the scalar sector, like the "Georgi-Machacek" (GM) model. In our analysis, we use these two models as a benchmark and they are detailed in the following sub-sections.

2.2.1 Simplified Lagrangian

In Simplified Lagrangian, the idea is that resonant searches are generally not sensitive to all the details and free parameters of the full model, but they are sensitive to only those parameters that control the mass of the resonance and the interactions involved in its production and decay. Therefore, it is useful to define a simplified description of resonance by a phenomenological Lagrangian where only the mass parameters and relevant couplings are kept. Apart from symmetry constraints, the Simplified Lagrangian does not need to be specific to a definite model, unlike the complete model lagrangian. Its main goal is to provide a phenomenological parameterization of a broad set of explicit models which contain only those terms which are necessary in the resonance description. The experimental results should be presented in the parameter space of the phenomenological Lagrangian and expressed in terms of a likelihood function. The Simplified Lagrangian serves as a bridge between phenomenological parameters and the experimental data. The phenomenological parameters of the model predict experimental observations and inversely, experimental measurements can be translated into constraints on the model parameters of an

explicit model.

The Simplified Lagrangian is constructed to describe only the on-shell resonance production and therefore, some care is required when comparing it with data.

2.2.2 Heavy Vector Triplets

An example of such a Simplified Lagrangian is a simple but well-motivated model of electroweak charged spin one resonance known as Heavy Vector Triplets (HVT) [26]. In this model, in addition to the SM fields and interactions, a real vector $V_\mu^a, a = 1, 2, 3$ in the adjoint representation of $\mathrm{SU}(2)_L$ is considered. This vector describes two charged and one electrically neutral heavy spin one particle with the charge Eigenstate fields:

$$V_\mu^\pm = \frac{V_\mu^1 \mp iV_\mu^2}{\sqrt{2}}, \quad V_\mu^0 = V_\mu^3 \tag{2.40}$$

Similar to [26], the dynamics of the phenomenological Lagrangian can be written as:

$$\begin{aligned}\mathcal{L}_V = &-\frac{1}{4}D_{[\mu}V_{\mu]}^a D^{[\mu}V^{\mu]a} + \frac{m_V^2}{2}V_\mu^a V^{a\mu} \\ &+ ig_V c_H V_\mu^a H^\dagger \tau^a \overleftrightarrow{D}^\mu H + \frac{g^2}{g_V}c_F V_\mu^a J_F^{\mu\, a} \\ &+ \frac{g_V}{2}c_{VVV}\epsilon_{abc}V_\mu^a V_\nu^b D^{[\mu}V^{\nu]c} + g_V^2 c_{VVVHH}V_\mu^a V^{\mu a} H^\dagger H \\ &- \frac{g}{2}c_{VVW}\epsilon_{abc}W^{\mu\nu a}V_\mu^b V^c\nu\end{aligned} \tag{2.41}$$

where,

$$D^{[\mu}V^{\mu]} = D_\mu V_\nu^a - D_\nu V_\mu^a, \quad D_\mu V_\nu^a = \partial_\mu V_\nu^a + g\epsilon^{abc}V_\mu^b V_\nu^c \tag{2.42}$$

and g denotes the $\mathrm{SU}(2)_L$ gauge coupling while $J_F^{\mu\, a}$ represents the SM left handed fermionic current:

$$J_F^{\mu\, a} = \Sigma_f \bar{f}_L \gamma^\mu \tau^a f_L \tag{2.43}$$

The Higgs parameter c_H controls the V interaction with the physical Higgs field and the SM vector bosons W and Z and is responsible for Vector Boson Fusion (VBF) production and decay to vector bosons. Similarly, the parameter c_F describes the interaction with fermions which is responsible for resonant Drell Yan production and its decay to fermions. It can be generalized for different couplings to leptons, light quarks, and third-generation quarks

$$c_F V.J_F \to c_l V.J_l + c_q V.J_q + c_3 V.J_3 \tag{2.44}$$

where J_l, J_q and J_3 are the SM currents for leptons, light quarks and third-generation quarks, respectively. To note, eq. 2.41 consists of two coupling constants, g_V and g. g_V represents the typical strength of V interactions and the dimensionless coefficients "c" parametrize the departure from the typical size. An extra factor of g^2/g_V is included due to fermion couplings. The

parameter c_F, as defined in eq. 2.41, is of order one and c_H is of order one in the strongly coupled scenario but can be lower in the weakly coupled case. In any case, they are never larger than one except for c_3 which can be larger than one in strongly coupled scenarios. The parameter g_V can vary over one order of magnitude from a weakly coupled case with ≈ 1 to a strong limit with $\approx 4\pi$. The parametrization of g_V is useful at the theoretical level but is redundant and could be absorbed into the c parameters.

The other three new parameters c_{VVV}, c_{VVHH} and c_{VVW} do not include vertices of a single V with light SM fields and therefore do not contribute to V decays and therefore not relevant for searches in the LHC. Therefore the phenomenology can be expressed by the four parameters: c_H, c_l, c_q and c_3.

2.2.2.1 Masses

After the EWSB, photon is the only massless state identified as the gauge field affiliated with the unbroken $\mathrm{U}(1)_{em}$ is given by:

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \tag{2.45}$$

where, $\tan\theta_W = g'/g$. g and g' are traded for $e \approx \sqrt{4\pi/137}$ and $\sin\theta_W$. The two neutral mass eigenstates are the SM Z boson and one heavy vector of mass M_0. The charged and neutral mass matrices are connected by custodial symmetry and a generalized custodial relation can be obtained as [26]:

$$m^2_W M^2_\pm = \cos^2\theta_W m^2_Z m^2_0 \tag{2.46}$$

Hierarchy is required as the SM masses should be $m_{W,Z} \approx 100GeV$ which is possible if:

$$\frac{\hat{m}_{W,Z}}{\hat{m}_V} \approx \frac{m_{W,Z}}{M_{\pm,0}} \leq 10^{-1} \ll 1 \tag{2.47}$$

There is some mass mixing of W and Z with W' and Z' which can be important for low masses of M_V, and therefore the model is mostly valid for heavy new vectors. It is found that the charged and neutral heavy vectors are degenerate and thus comparable production rates at LHC.

2.2.2.2 Decay widths

The relevant decay channels are di-lepton, di-quark and di-boson. All the couplings to fermions are guided the parameter $g^2 c_F/g_V$. The partial width to fermions is given by:

$$\Gamma_{V_\pm \to f\bar{f}'} \simeq 2\Gamma_{V_0 \to f\bar{f}'} \simeq N_c[f] \left(\frac{g^2 c_F}{g_V}\right)^2 \frac{M_V}{48\pi} \tag{2.48}$$

where, $N_c[f]$ is the number of colors, which is 3 for di-quark and 1 for di-lepton decays. For heavy gauge bosons, the "Equivalence Theorem" [26] applies: the longitudinal polarization will dominate and the coupling to the Higgs boson, c_H, will be the same as to the SM gauge bosons W and Z. The partial widths of the heavy vectors to longitudinal W and Z is calculated as:

$$\Gamma_{V_0 \to W^+_L W^-_L} \simeq \Gamma_{V_\pm \to W^\pm_L Z_L} \simeq \frac{g^2_V c^2_H M_V}{192\pi} \tag{2.49}$$

2.2.2.3 Production rates

Both the DY and VBF processes are found to be proportional to the partial with $\Gamma_{V\to ij}$ of the inverse decay process:

$$\sigma(pp \to V + X) \propto \Sigma_{i,j\in p}\frac{\Gamma_{V\to ij}}{M_V} \tag{2.50}$$

where, $i, j = q, \bar{q}, W, Z$ denote the colliding partons in the two protons.

2.2.2.4 Explicit Models

Here, three explicit models, Model A, Model B and Model C, of the Simplified Model are described and are used as benchmark models for our analysis as suggested in [26]. Model A describes the vector triplet emerging from an underlying weakly-coupled extension of the SM gauge group [27]. Model B describes the vector triplet considered in the strongly coupled Composite Higgs scenario [28]. Model C describes the vector triplet produced by vector boson fusion. Feynman diagrams for the production and decay to WZ of the HVT is shown in Figure 2.4

2.2.2.4.1 Model A

In this model, the SM gauge theory is extended to $SU(2)_1\times$ $SU(2)_2$ $\times$ $U(1)_Y$. The SM fermions are charged under $SU(2)_1$ and $U(1)_Y$. The SM Higgs doublet is extended by an additional scalar field ϕ, transforming as a real bidoublet $(2,2)_0$ under $SU(2)_1\times$ $SU(2)_2$ with Y=0. It obtains a vacuum expectation value (VEV)

$$\langle\phi\rangle = \begin{pmatrix} f & 0 \\ 0 & f \end{pmatrix} \tag{2.51}$$

This VEV breaks the $SU(2)_1\times$ $SU(2)_2$ gauge symmetry into the SM $SU(2)_L$ gauge group. After mass mixing and field redefinition, the physical states are identified as the SM weak bosons and a triplet of heavy vectors. The couplings for this model are of the order:

$$c_H \sim -g^2/g_V^2 \quad \text{and} \quad c_F \sim 1 \tag{2.52}$$

With such couplings, the partial width of the $V_\pm$ to fermions, Eq 2.48, and gauge bosons, eq. 2.49, become comparable.

2.2.2.4.2 Model B

To address the hierarchy problem, in some models the Higgs boson emerges as a pseudo Nambu-Goldstone boson, and therefore as a light state, from an underlying strong dynamics. They predict the existence of heavy vector resonances. One such example is the Minimal Composite Higgs Model (MCHM) [28], which predicts the existence of heavy vector resonances with electroweak quantum numbers with strong coupling to gauge bosons. Model B is designed to represent the lightest of the new vector resonances described by the Simplified Model. In this model, the couplings to boson and fermions both are comparable of the order one

$$c_H \sim c_F \sim 1 \tag{2.53}$$

With the above choice of couplings, the partial width $\Gamma_{V_\pm}$ to gauge bosons, Eq 2.49, will dominate over to that fermions, eq. 2.48.

2.2.2.4.3 Model C

The Drell-Yan production process of heavy vector resonances requires coupling to fermions and is therefore insensitive to small values of c_F, in the plane of parameter space c_F vs c_H. To investigate a scenario where there is no coupling of the heavy vectors to fermions, a new model is defined, namely, Model C. It is defined such that at this point in parameter space $c_F = 0$. Model C couples to vector bosons and is therefore used for VBF searches. In this analysis, HVT masses are excluded for this Model C with $g_V c_H$ parameters product set to 1.0, 2 and 3.0.

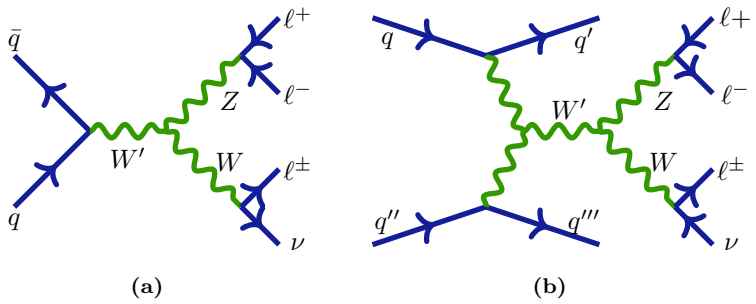


Figure 2.4: Representative Feynman diagrams for heavy resonance production and decay to WZ bosons (a) HVT W' production via Drell-Yan and (b) HVT W' production via Vector Boson Fusion. The subsequent decays to the $\ell^+\ell^-\ell^\pm\nu$ are also shown.

2.2.3 Georgi Machacek Model

Given the uncertainties in Higgs boson coupling measurements, EWSB mechanism with extended Higgs sectors is still allowed. The problems of SM can possibly be explained in the Higgs extended models, such as the Georgi-Machacek (GM) model [29]. The Georgi Machacek model is a minimal extension of the SM Higgs sector which includes isospin higher than 2 [30]. Higgs extensions are subject to this stringent constraint and to satisfy it, the ρ parameter is preserved at the tree level by imposing a global $SU(2)_L \times SU(2)_R$ symmetry on the scalar potential. The ρ parameter is a measure of the relative strengths of neutral and charged-current interaction in four-fermion processes at zero momentum transfer. In SM, the ρ parameter relates the mass of the W and Z boson as:

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} = 1 \tag{2.54}$$

The custodial symmetry ($\rho = 1$) is satisfied by extending the SM Higgs doublet $(\phi^+, \phi^0)^T$ by two isospin triplets: one complex triplet $(\chi^{++}, \chi^+, -\chi^0)^T$ with $Y = 2$ and one real triplet $(\xi^+, \xi^0, -\xi^{++})^T$ with $Y = 0$ [31]. The doublet is written in the form of a bi-doublet Φ and the triplets are combined to form

a bi-triplet χ

$$\Phi = \begin{pmatrix} \phi^{0*} & \phi^{+} \\ -\phi^{+*} & \phi^{0} \end{pmatrix}, \qquad \chi = \begin{pmatrix} \chi^{0*} & \xi^{+} & \chi^{++} \\ -\chi^{+*} & \xi^{0} & \chi^{+} \\ \chi^{++*} & -\xi^{+*} & \chi^{0} \end{pmatrix} \tag{2.55}$$

Thus the field χ describes a doubly charged particle, two singly charged particles and three real neutral particles.

Electroweak symmetry breaking occurs when both Φ and χ develop non-zero VEVs, given by $\langle\phi\rangle = \frac{v_\phi}{\sqrt{2}} I_{2\times2}$ and $\langle\chi\rangle = \frac{v_\chi}{\sqrt{2}} I_{3\times3}$. The $W^\pm$ and Z boson masses receive contribution from both the VEVs:

$$M_W^2 = M_Z^2 \cos^2\theta_H = \frac{1}{4} g^2 (v_\phi^2 + 8\chi^2) \tag{2.56}$$

Ten physical Higgs scalars are identified and classified as: a fiveplet (a quintet) (H_5), a triplet (H_3) and two singlets (h, H). In terms of the fields in 2.55, the custodial fiveplet and triplet states are:

$$\begin{aligned} H^{\pm\pm} &= \chi^{\pm\pm} \\ H_5^{\pm} &= \frac{\chi^{\pm} - \xi^{\pm}}{\sqrt{2}} \\ H_5^{0} &= \frac{\sqrt{2}\xi^{0,r} - \chi^{0,r}}{\sqrt{3}} \\ H_3^{\pm} &= -sin\theta_H \phi^* + cos\theta_H \frac{\chi^{\pm} + \xi^{\pm}}{\sqrt{2}} \\ H_3^{0} &= -sin\theta_H \phi^0 + cos\theta_H \chi^{0,i} \end{aligned} \tag{2.57}$$

where,

$$\begin{aligned} \cos\theta_H &= \frac{v_\phi}{v} \\ \sin\theta_H &= \frac{2\sqrt{2} v_\chi}{v} \end{aligned} \tag{2.58}$$

v being the VEV of the SM Higgs, and the neutral fields are decomposed into real and imaginary parts as:

$$\begin{aligned} \phi^0 &\to \frac{v_\phi}{\sqrt{2}} + \frac{\phi^{0,r} + i\phi^{0,i}}{\sqrt{2}} \\ \chi^0 &\to v_\chi + \frac{\chi^{0,r} + i\chi^{0,i}}{\sqrt{2}} \\ \xi^0 &\to v_\chi + \xi^{0,r} \end{aligned} \tag{2.59}$$

The custodial singlet states are given by:

$$\begin{aligned} h &= \cos\alpha \phi^{0,r} - sin\alpha H_1^{0'} \\ H &= \sin\alpha \phi^{0,r} + cos\alpha H_1^{0'} \end{aligned} \tag{2.60}$$

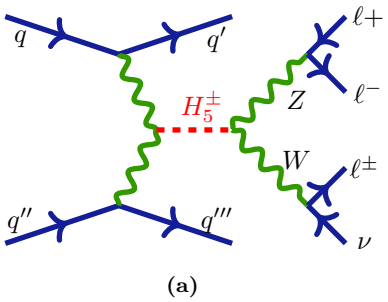
where, α is the mixing angle and

$$H_1^{0'} = \frac{\xi^0 + \sqrt{2}\chi^{0,r}}{\sqrt{3}} \tag{2.61}$$

The fiveplet state H_5 is fermiophobic and couples to vector bosons and are therefore an ideal candidate for VBF searches.

2.2.3.1 H5 Benchmark Plane

The H5 benchmark plane is designed to facilitate searches for the $H_5^{\pm}$ states in vector boson fusion with decays to $W^{\pm}Z$ in the LHC [32, 33]. This benchmark makes sure that the mass of the triplet states m_3 is always larger than the fiveplet states m_5 and the other parameters are chosen accordingly and therefore forbids the Higgs to Higgs decays, i.e. $H_5 \to H_3H_3$ or $H_5 \to H_3V$ are kinematically forbidden and assures a $\mathrm{BR}(H_5 \to VV) = 100\%$. In the H5 benchmark plane, the largest possible parameter region allowed by theoretical constraints is for 200 GeV $< m_5 <$ 3000 GeV. Feynman diagram for the production and decay to WZ of the GM singly charged $H_5^{\pm}$ which is searched for in this analysis is shown in Figure 2.5



(a)

Figure 2.5: Representative Feynman diagrams for heavy resonance production and decay to WZ bosons of GM $H_5^{\pm}$ production via Vector Boson Fusion. The subsequent decays to the $\ell^+\ell^-\ell^{\pm}\nu$ are also shown.

Chapter 3

Experimental setup

"With great power, comes great responsibility."
Spider-Man: No Way Home

The experimental data used for this thesis has been collected at the Atlas detector during 2015, 2016, 2017 and 2018 with full Run 2 of the Large Hadron Collider. Although no studies were performed focusing on the experimental setup, this chapter provides important information on the experimental design and the analysis chain which is instrumental for a better understanding of the investigation of the data. This chapter gives a wide overview and will provide a brief outline of the experimental setup, including the particle accelerator LHC and the Atlas detector.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest particle accelerator, built from 1998 to 2008 by the European Organization for Nuclear Research (CERN) [34] in the existing Large Electron-Positron (LEP) accelerator tunnel. It is located near Geneva, Switzerland, and crosses the boundaries of France. The LHC's primary design goal was to be a discovery machine, to allow for searches of new physics impacts in a plethora of different scenarios. These scenarios fueled the design goal of LHC, making it capable of accelerating protons or heavy ions to unexplored energies and instantaneous luminosity. The LHC can operate in a variety of beam modes, colliding two proton beams, a proton and a Pb ions beams, or two beams of Pb ions. During collisions at these high energies, the constituents of protons or the heavy ions, quarks and gluons, interact with each other providing an accessible energy domain for new particle searches from $\approx$ 100 GeV to few TeV.

The LHC is installed 100 meters under the earth's surface and is spread over a 27km circumference. The accelerator complex is a series of machines that accelerates the energy of the beam progressively. In this context, the LHC is considered as the last element in this chain. The data used in this thesis is produced by proton-proton collisions.

The protons are extracted from a hydrogen bottle at the beginning of the chain. After that, those protons are injected into the linear accelerator (LINAC 2) and reach 50 MeV of energy, followed by the Booster where the proton beam

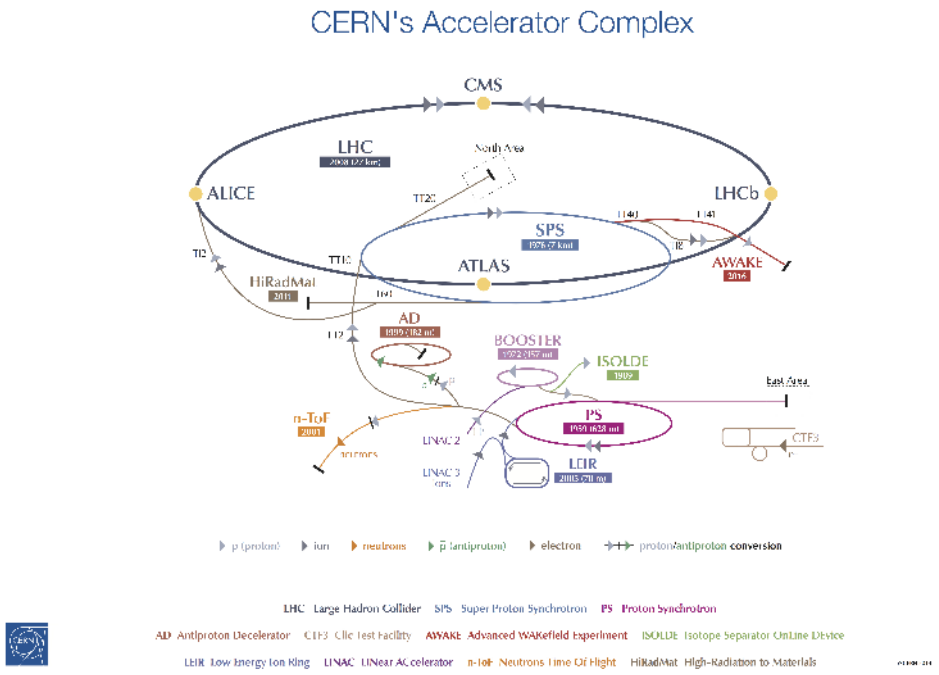


Figure 3.1: Schematic of the CERN accelerator complex including the LHC (left) and the acceleration chain (right). Protons are injected from Linac 2 to Booster. Then PS and SPS accelerate them sequentially before they enter the LHC. For each accelerator, the year of start and the circumference/size are shown. Figure is adapted from [35]

reaches 1.4 GeV. Then they go through the Proton Synchrotron (PS) and lastly to the Super Proton Synchrotron (SPS) to achieve an energy of 450 GeV. Lastly, proton beams are injected into the LHC in two opposite directions and are accelerated to 6.5 TeV. Two adjoining parallel pipes of beams are installed inside the tunnel, and they travel in opposite directions. Figure 3.1 shows the different parts of the LHC.

1232 dipole magnets and 392 quadrupole magnets are used in LHC to hold the beams in their circular orbits and focus the proton beams to collide at four different points, respectively. Higher order multipole magnets (including octupoles, sextupoles, decapoles) are used to rectify the smaller imperfections in the geometrical field. In total, approximately 9,600 magnets are used. With an aim of maximising the number of collisions, to attain high luminosity, the LHC gets 2,808 bunches in every beam, each bunch are separated by 25 ns.

3.1.1 Experiments at the LHC

The proton beams accelerated in opposite directions collide at four different points where detectors are installed to detect the particles generated. The two most significant experiments are A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS) [36]. They are general purpose detectors and adopt a symmetric cylindrical geometry to cover all possible solid angles. As the researchers are working on two significant independent experiments, it allows the scientists to substantiate any finding at the LHC. The other two detectors

are used for different specific purposes. LHCb (LHC beauty) [37] intends to study the CP violation and search for rare B-meson decays by adopting an asymmetric geometrical structure, as the emission of B mesons are especially occurring at small angles to the beams because of their relatively small mass as compared with the high collision energy. ALICE (A Large Ion Collider Experiment) [38] examines the quark-gluon plasma, a state of matter where the gluons and quarks are unrestricted, which is necessary to probe the mechanism of colour confinement and hadron construction.

There are additional three small experiments installed in the LHC to achieve distinct physics goals. The TOTEM (TOTal Elastic and diffractive cross section Measurement) experiment [39] aims at the total cross-section, elastic scattering and diffractive processes of the proton-proton collisions. The LHCf [40] or LHC forward experiment analyses particles produced in the forward part of collisions. Due to its unique placement, particles at almost zero degrees to the proton beam pipe direction can be measured and captured. The MoEDAL (the Monopole and Exotics Detector at the LHC) [41] experiment does direct searches for the exotic particles and magnetic monopole.

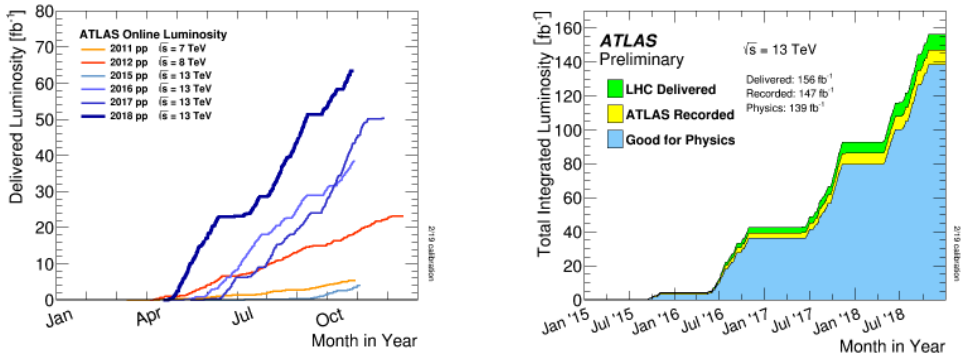
3.1.2 Operation energies and integrated luminosities

The luminosity represents the number of collision events per area per time. The LHC physics operation commenced in November of 2009 and reached 1.18 TeV in every beam [42], surpassing the record of 0.98 TeV in each beam at the Tevatron. In 2010, the LHC could attain 3.5 TeV per beam, delivering center-of-mass energy, $\sqrt{s}$, of 7 TeV [43]. In 2011 it produced an integrated luminosity of 6.1 fb^{-1} at $\sqrt{s} = 7$ TeV and successively in 2012 it delivered 23.1 fb^{-1} at $\sqrt{s} = 8$ TeV [44]. After that, the LHC was closed down in February of 2013 for a couple of years due to a technical stop, preparing to expand the collision energy. On 3rd June 2015, the LHC resumed producing physics data with $\sqrt{s}$ = 13 TeV [45], also called as Run 2. Consequently, in 2016, 2017 and 2018, the machine concentrated on elevating the luminosity and achieved its design instantaneous luminosity of $1.0 \times 10^{34} cm^{-2} s^{-1}$ on 29th June 2016 [46]. The maximal instantaneous luminosity is 2.05×10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which was reached on 2 November 2017 [47]. Run 2 was stopped on 10 December 2018 [48] and underwent a second two year technical stop in 2019-2021 before restarting again for Run 3 in 2022.

The ATLAS experiment has collected a large amount of data due to the successful and stable operation of the LHC. Figure 3.2 shows the data delivered in LHC and recorded in ATLAS in the past years.

3.2 The ATLAS Detector

The ATLAS experiment [49, 50] is one of the largest general-purpose detectors that has been built for investigating a wide range of physics processes, ranging from precision measurements of SM properties to observation of new physics phenomena. It holds a cylindrical form, which weighs 7,000t, is 46m long, and is 25m in diameter. ATLAS encompasses different sub-detectors piled up from inside to outside, to keep track of the momentum, trajectory and power of par-



(a) **(b)**

Figure 3.2: Integrated luminosity in ATLAS during stable beams for different data-taking periods (2011-2018)(a). Integrated luminosity delivered by the LHC (green), recorded by ATLAS (yellow), and certified quality data for physics analytics (blue) at $\sqrt{s} = 13$ TeV from 2015 to 2018 (b). The plateaus correspond to so-called technical stops of the LHC. Plots are taken from ATLAS public website

ticles. At the same time, this information facilitates particle identification and reconstruction. A large magnet system curves the trajectories of the ion thus their momenta can be judicially estimated and evaluated precisely. Particles that are emerging due to the collisions in the transverse direction to the beam can be identified in the barrel part. On the other hand, those particles which are close to the beam are detected by so-called endcaps. The ATLAS comprises four major components that are the calorimeter, the tracker, the magnet system and the muon spectrometer, depicted in a schematic view of the ATLAS detector in Fig 3.3.

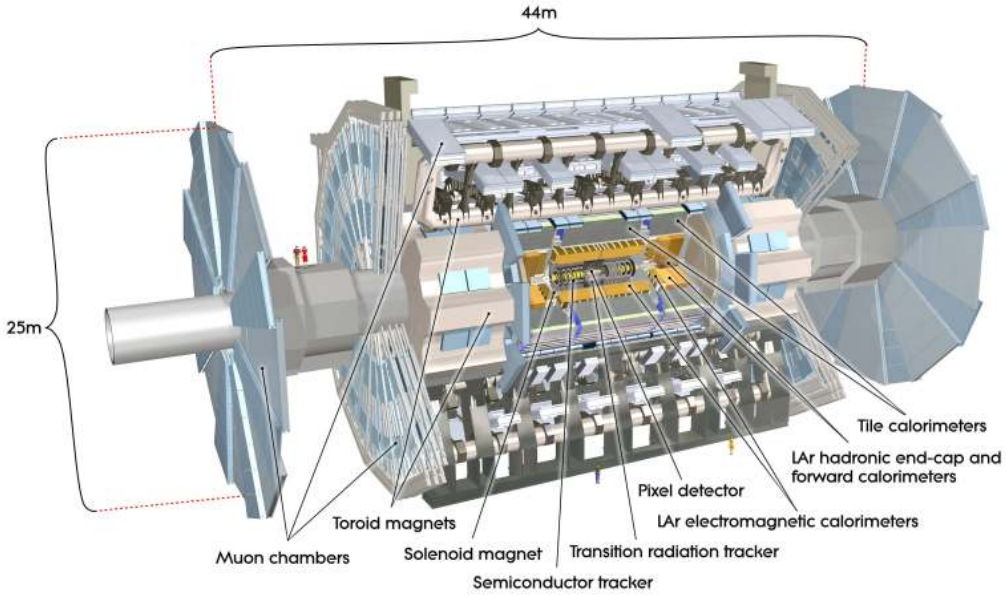


Figure 3.3: "Cut-away view of the ATLAS detector. The dimensions of the detector are $25m$ in height and $44m$ in length. The overall weight of the detector is approximately 7000 tonnes." [49].

3.2.1 Coordinate system

The ATLAS detector uses a right-handed coordinate system as shown in Figure 3.4. The origin is at the pp interaction point (IP) in the core of the ATLAS detector and the z-axis lies in the beam pipe. Moreover, the x-axis arises from the IP to the inside of the LHC ring, and the y-axis proceeds upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane. The azimuthal angle, ϕ, is measured in the $x - y$ plane and the polar angle, θ, is the angle to the z-axis, which is used to define pseudorapidity η = -ln(tan(θ/2)).

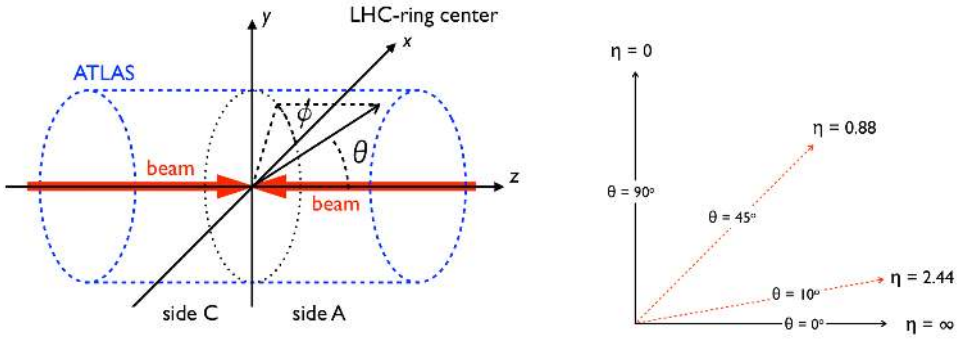


Figure 3.4: The coordinate system used in the ATLAS detector (left) and relation between pseudorapidity and polar angle (right) [51]

The transverse and longitudinal impact parameters, d_0 and z_0, are defined as the point to the closest approach of the trajectory of a track to the primary vertex in the transverse plane and z-direction, respectively.

3.2.2 The Inner Detector

The inner detector (ID) is the central component of the ATLAS detector designed to reconstruct the primary and secondary vertices and charged particle trajectories with high precision over the pseudorapidity range of $|\eta| < 2.5$. It consists of three sub-components that are present in a uniform magnetic field parallel to the beam axis, a pixel detector (Pixel), semiconductor tracker (SCT) and transition radiation tracker (TRT). Each of the sub-components contains two endcaps and a barrel. Figure 3.5 shows the layout of the ATLAS inner detector and the upper part of Figure 3.6 shows the schematic of different sub-components of the ID.

The Pixel Detector

The Pixel detector [50, 52] is the innermost layer containing four concentric cylindrical layers, including one layer of the Insertable B-Layer or IBL [53], around the beam in the barrel and other discs-like three layers encompassing the end cap. It surrounds a pseudorapidity area of $|\eta| < 2.5$. Silicon pixel sensors are employed with the size of $50 \times 250\mu m$ for the IBL and $50 \times 400\mu m$ for the other regions. The intrinsic spatial resolution in $\phi - z$ ($\phi - r$ if endcaps) is $8 \times 40\mu m$ for the IBL and $10 \times 115\mu m$ for the rest. Overall, it covers the space of $31mm < r < 242mm$ and $|z| < 3092mm$. The lower part of Figure 3.6 shows the different layers of the Pixel.

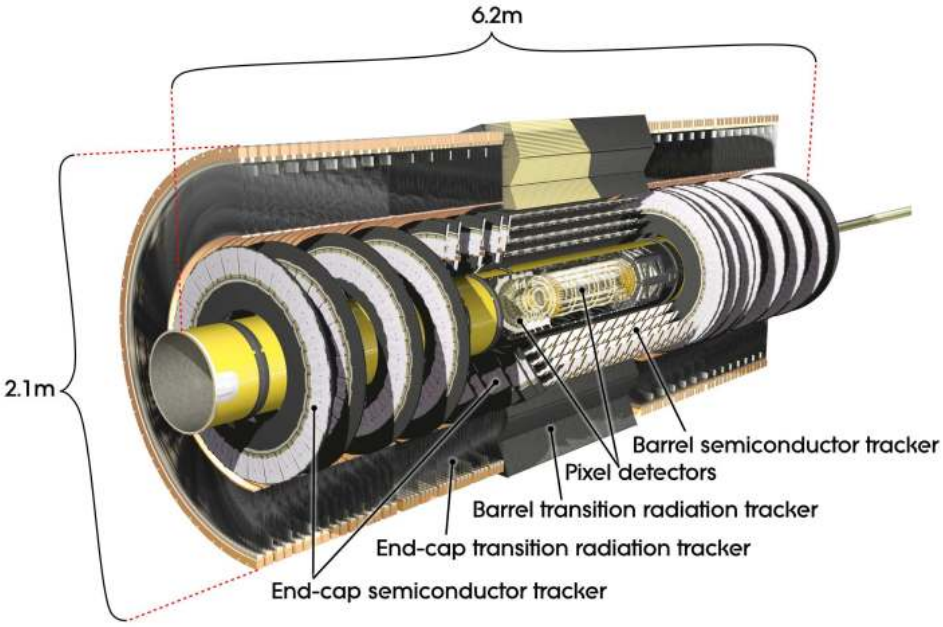


Figure 3.5: "Cut-away view of the ATLAS inner detector." [49].

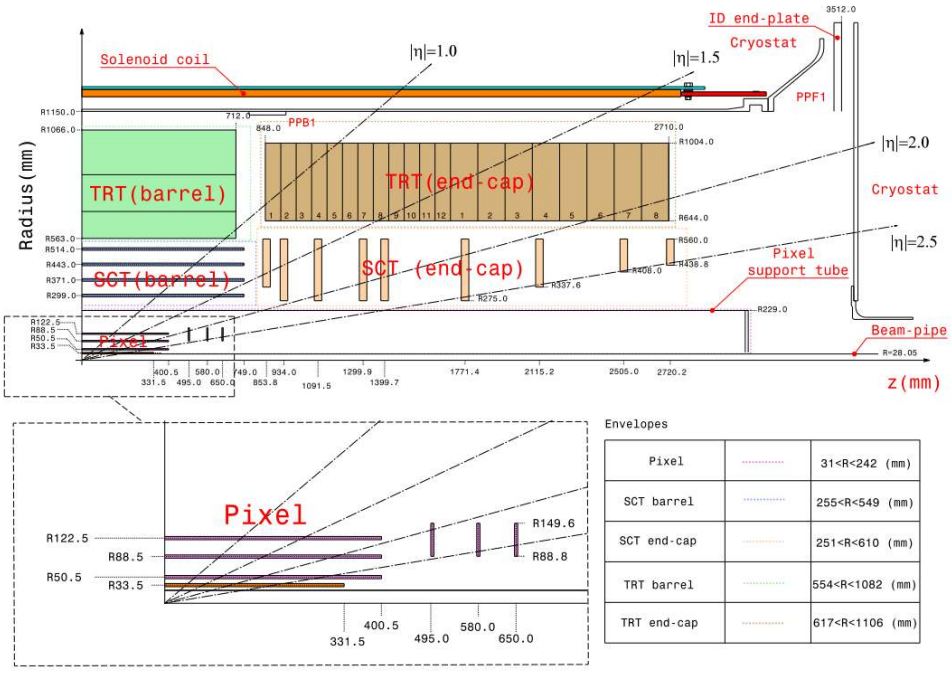


Figure 3.6: The $r - z$ cross-section view of the layout of a quadrant of the ATLAS inner detector for Run 2. The top panel shows the whole inner detector, whereas the bottom-left panel shows a magnified view of the pixel detector region [52].

The SCT

The SCT [50, 52] is a silicon microstrip tracker made up of nine disc layers in the endcap and four cylindrical layers in the barrel. The positions and extents of the elements were optimized, thus any track naturally intersects four

SCT layers, being independent of its pseudorapidity. Both the endcaps and the barrel are constructed by modules that are built mostly by four microstrip sensors. Among them, two microstrip sensors are of the size of $6.36 \times 6.40 cm^2$ which are wire-bonded concurrently to develop a $12.8cm$ long strip, and the other two strips are pasted together with a glue back-to-back at a 40 $mrad$ stereo angle, to deliver a needed z (barrel) or r (endcaps) dimensions. The SCT gains an intrinsic spatial resolution in $\phi - z$ ($\phi - r$ if endcaps) of $17 \times 580\mu m$. The SCT envelops $251mm < r < 610mm$ and $|z| < 2797mm$ as shown in upper part of Figure 3.6.

The TRT

The TRT [50, 52] is a gaseous detector, consisting of straw tubes of 4 mm diameter, each acting as a single drift chamber. It complements the SCT and Pixel in estimating the transverse momentum of charged particles by providing a large number of $(R - \phi)$ measurements. The barrel contains around 50000 straws that are clustered into three rings. In the endcaps, straws are aligned radially and organized in 20 wheels per side, with eight layers in each wheel. The straws contain Xenon-based gas mixture. The TRT improves the momentum resolution of tracks and provides particle identification because of the transition radiation. The resolution of TRT is 130 μm in the r direction within $|\eta| < 2$. It covers the space of $554mm < r < 1106mm$ and $|z| < 2744mm$.

3.2.3 The Calorimeter

The calorimeter system measures the energy of incident charged and neutral particles by absorbing the energy deposition over a massive pseudorapidity range of $|\eta| < 4.9$. Various kinds of calorimeters are used in ATLAS which complements each other, focusing on different particle-matter interactions as shown in Figure 3.7. Showers are made by a cascade of secondary particles in the dense absorber material, providing particle identification and energy measurement. The electromagnetic calorimeter (EMC) determines the energy of particles that interact through the electromagnetic force, like photons and electrons. The hadron calorimeter (HCal), placed directly outside the EMC, measures the energy of hadrons. In addition to the EMC and the HCal, ATLAS calorimetry also incorporates the forward calorimeter (FCal) to cover the forward region.

The EMC

The EMC is a liquid-argon (LAr) sampling detector with a lead absorber. With an accordion geometry, it delivers a full azimuthal range until $|\eta| < 3.2$ without any cracks. It comprises of a barrel part ($|\eta| < 1.475$) and two endcaps ($1.375 < |\eta| < 3.2$) with some overlapping. Also, the barrel is divided into two halves at $z = 0$ for essential cabling and electronics. Every endcap calorimeter is composed mechanically by a couple of coaxial wheels. In the region with $|\eta| < 1.8$, a presampler is placed to correct for the energy loss upstream of the material of the calorimeter. The three (two if without presampler) sampling layers have different granularities. The finer granularity is in the most internal

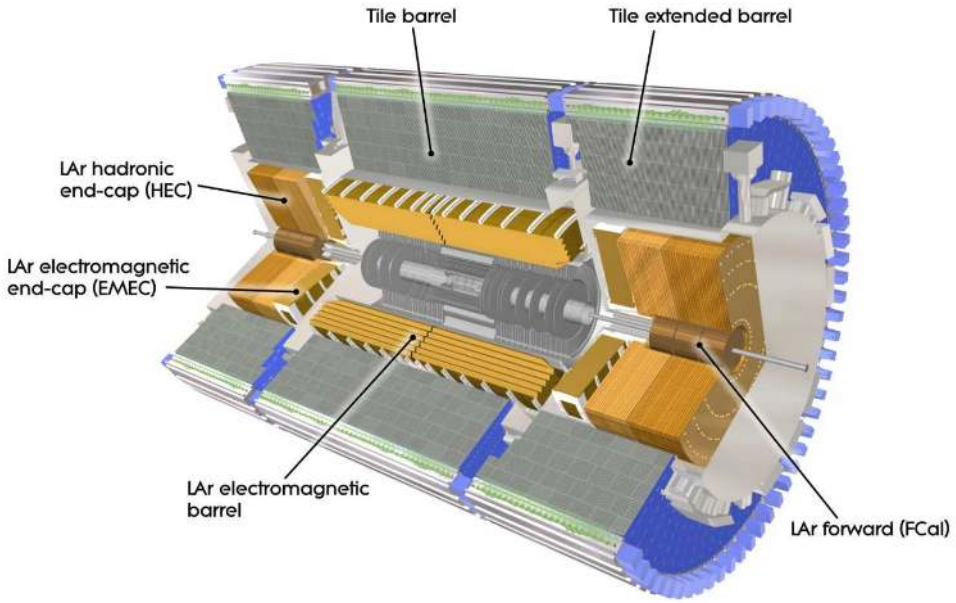


Figure 3.7: View of the ATLAS calorimeter system consisting of the EMC and the HCal. The FCal is in the very forward and backward regions close to the beam pipe. [49]

layer to achieve exceptional resolution and precise spatial measurements while coarser in the outermost layer to detect leakage. The thickness of the EMC is more than 24 radiation lengths (X_0) in the barrel and above 26 X_0 in the endcaps.

The HCal

The HCal employs a tile plastic-scintillators with lead absorber in the barrel ($|\eta| < 1.0$) and two extended barrel parts ($0.8 < |\eta| < 1.7$) covering the endcaps of the other calorimeters. Liquid-Argon (LAr) calorimeters are used in the endcaps with a copper absorber, covering $1.5 < |\eta| < 3.2$.

The FCal

The FCal has three LAr layers and is positioned radially inside of the HCal endcaps in the range $3.1 < |\eta| < 4.9$. The first layer uses a copper absorber to measure the energy of electromagnetic shower measurements. The other two layers use tungsten to provide energy measurements of hadronic showers.

3.2.4 The Muon Spectrometer

The muon spectrometer (MS) [54] is designed to detect muons that pass through the calorimeter without stopping and measure their momenta and charge. It is the outermost part of the ATLAS detector. It incorporates trigger sections and high-precision chambers for bunch-crossing identification and momentum measurement, respectively. The trigger chambers use resistive plate chambers (RPCs) in the barrel and thin gap chambers (TGCs) in the endcaps for fast time response, in the range $|\eta| < 2.4$. The high-precision chambers utilize monitor drift tubes (MDTs) for the majority and Cathode-Strip chambers (CSCs)

for the innermost part of the internal endcap disks, covering $|\eta| < 2.7$. MDTs provide good spatial resolution and CSCs provide good momentum measurements. Figure 3.8 shows a schematic of the ATLAS muon spectrometer.

The trigger chambers employ three layers of wireless gaseous detectors to obtain good timing response, which includes contributions of the backend electronics and signal propagation. The timing details contribute to the muon level 1 trigger. They are positioned in the barrel ($|\eta| < 1.05$), reaching a temporal resolution of approximately $1.5ns$. In the end caps ($1.05 < |\eta| < 2.4$), TGCs are used to deliver a temporal resolution of $4ns$. The trigger chambers also measure the coordinate of both the bending and non-bending planes, which is used for the final muon reconstruction.

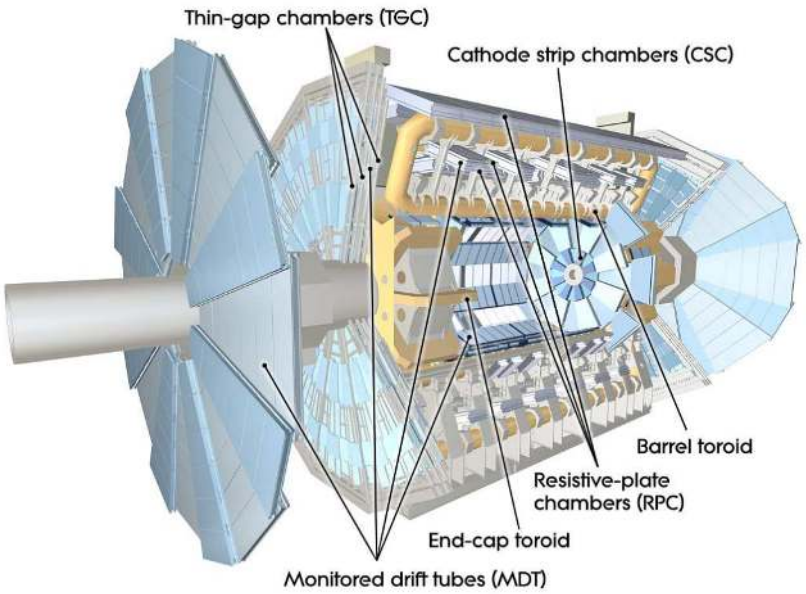


Figure 3.8: "Cut-away view of the ATLAS muon system." [49]

3.2.5 The Magnet system

The ATLAS detector is equipped with a unique hybrid superconducting magnet system [55] that bends the charged particle which is used to calculate the momentum precisely and decide the charge of the particle. It comprises of a central solenoid [56, 57] covering the ID, which provides a 2T axial magnetic field, and three gigantic toroids, two endcaps and one barrel, as a section of the MS which provide a toroidal magnetic field of approximately 1T and 0.5T, respectively [49, 58, 59]. Figure 3.9 shows a geometric view of the ATLAS magnet system. The solenoid is cylindrical in shape with $2.4m$ diameter and a length of $5.3m$. The barrel toroid is $25.3m$ long and has an external diameter of $20.1m$, whereas the end cap toroids have an axial length of $5.0m$ and an outer diameter of $10.7m$. Every toroid is composed of eight coils arranged symmetrically around the beam pipe.

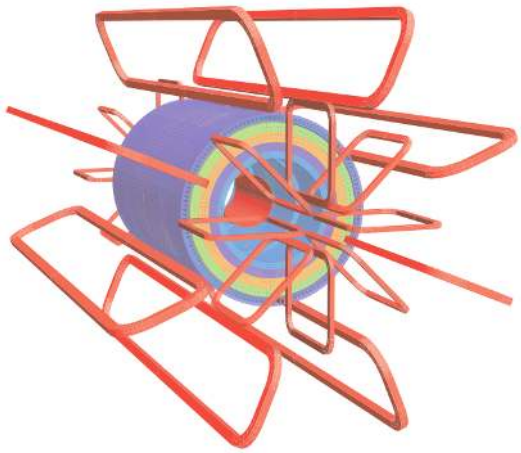


Figure 3.9: "Geometry of magnet windings and tile calorimeter steel. The eight barrel toroid coils, with the end-cap coils interleaved are visible. The solenoid winding lies inside the calorimeter volume. The tile calorimeter is modelled by four layers with different magnetic properties, plus an outside return yoke." [49]

3.2.6 The Trigger system

With the LHC's $25ns$ bunch spacing, the ATLAS detector's bunch-crossing rate is approximately $40MHz$. It is extremely difficult to record each event at this rate. As a result, to select and record important events, a trigger system [60] is essential. The trigger system has been significantly altered for Run 2 to cope with the elevated rate caused by the higher center-of-mass energy and increased luminosity. The Run 2 trigger [61] is made up of two parts: the level-1 (L1) trigger and the high-level trigger (HLT). The L1 trigger is hardware-based and determines the Regions-of-Interest in the detector with a latency of less than 2.5 μs depending on coarse calorimeter and muon information. The HLT is software-based, implementing rapid reconstruction algorithms either in the Regions-of-Interest or the entire event. The L1 trigger reduces the complete event rate from $40MHz$ to $100kHz$, and the HLT further reduces it to $1kHz$.

There are four components to the L1 trigger: the L1 calorimeter trigger system (L1Calo), the central trigger processors (CTP), the L1 muon trigger system (L1Muon), and the L1 topological trigger modules (L1Topo). L1Calo analyses information obtained from the EMC and HCAL, whereas L1Muon influences the responses generated by the MS trigger chambers. When the L1Topo receives the information from both, it quickly calculates event topological quantities, such as angular distance and invariant mass, to enhance the likelihood of hard collision identification. For deriving the final L1 triggering decision, the CTP integrates the information from these three systems. If the L1 trigger obtains the event, the CTP informs the subdetector readout system and passes the event fragment to the HLT, along with the L1 trigger summary.

In addition to running fast reconstruction algorithms, the HLT uses some fundamental requirements relating to electrons, muons, taus, jets, and missing transverse momentum that are relevant for physics analysis.

Chapter 4

Object reconstruction

"This is where our journey begins"

Black Panther

In this analysis, a fully leptonic decay channel for $W^{\pm}Z \rightarrow l\nu ll$ is considered. The fully leptonic channel has lower backgrounds, but because of a lower branching ratio, it is expected to be particularly sensitive to low-mass resonances. Thus, the objects involved in this analysis are electrons, muons, missing transverse energy and jets. To perform this analysis, it is important to reconstruct, identify and measure the energy and direction of these objects. Different objects interact in different ways with the detector. Every object has a salient signature, described in Fig 4.1. The offline reconstruction algorithms utilise all the information of the subdetector to rebuild and recognise these objects. Within the ATLAS collaboration, there are combined performance groups that provide tools and recommendations which is used in the analysis.

The first step is to combine individual hits in the trackers to reconstruct tracks of objects and combine with corresponding energy deposition clusters in the calorimeter. Different objects like electrons, muons and jets are reconstructed from these tracks and clusters. The reconstruction of electrons, muons and jets are described in the Sections 4.1, 4.2 and 4.3, respectively. The reconstruction of the tagged jets is described in sub-section 4.3.1 and the missing transverse energy to account for the undetected neutrino is described in Section 4.4.

4.1 Electrons

Electrons exit with tracks within the ID and form clusters in the EMC. As a result, the electron reconstruction integrates the important information of both the cluster and the track [62]. Electron identification and isolation criteria decide whether a candidate is taken into consideration or not.

Reconstruction and identification

Electrons are reconstructed by a sliding-window algorithm [63] with calorimeter clusters in the region of $|\eta| < 2.47$ [64]. These calorimeter clusters are matched with the ID tracks which are originating from the hard-scatter vertex. This

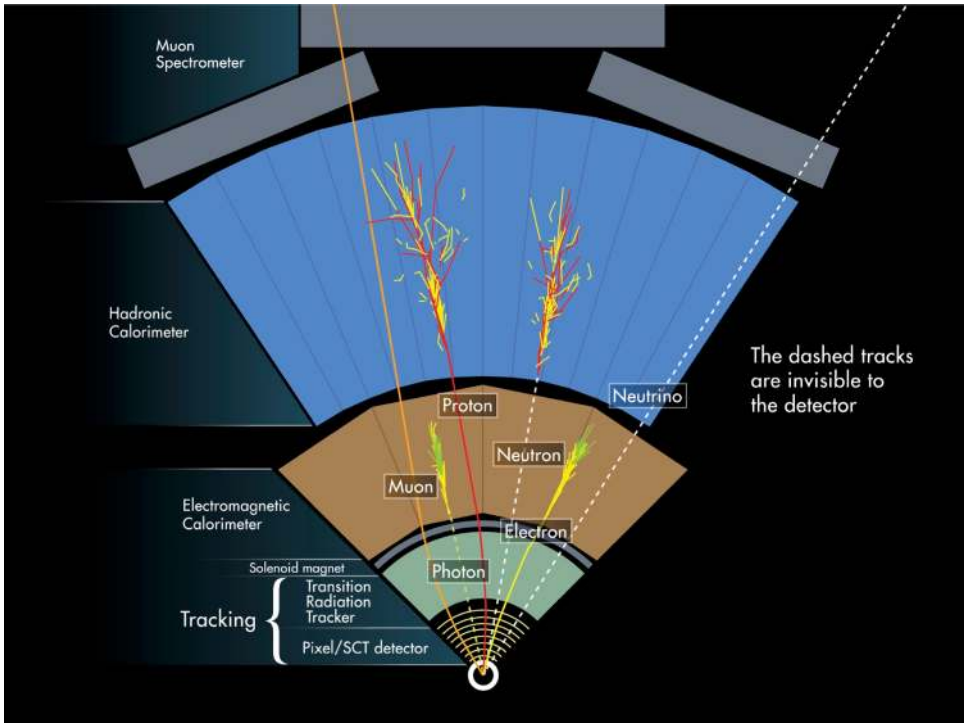


Figure 4.1: Signatures of various particles in the ATLAS detector. Each of them interacts differently with the matter, leaving signals in various subdetectors

matching is based on the angular separation between the extrapolated track position and cluster barycentre to the calorimeter. A tracking algorithm fixes ambiguity in matching by utilising the cluster-track distance and the number of hits in the pixel and first layer of the SCT detectors. Subsequently, the cluster is broadened a little accounting for the energies gathered beyond the sliding window. To improve reconstruction for candidates, especially for electrons which undergo significant energy loss due to bremsstrahlung, the track associated with a cluster, passing loose shower shape requirement, is refitted using a Gaussian-Sum Filter. Finally, the momentum and energy of the electron are calculated using the matched tracks and the calibrated cluster energy, respectively [65]. The momentum direction is acquired from the associated track direction extrapolated to the hard-scatter vertex.

The electron identification algorithm employs a multivariate approach to identify electrons against objects that mimic the signature of an electron, like hadronic jets or electrons from γ conversions. The algorithm is based on a likelihood value depending on observables of tracks and clusters, which incorporates hits in the IBL and the Pixel, the radiation-energy measurement in the TRT, the quality of the track, shower shape, cluster-track matching parameters. Four working points are defined based on the likelihood thresholds, `VeryLoose`, `Loose`, `Medium`, or `Tight` electrons, where the `Tight`(`Medium`) one is a subset of `Medium`(`Loose`). An additional working point is defined, `LooseAndBLayer`, by requiring on a hit in the IBL for electrons passing the `Loose` identification. `LooseAndBLayer` is used for electron identification in this analysis.

Isolation

Electron isolation purifies prompt electron candidates by isolating them spatially from other reconstructed objects. Two different types of discriminating variables are developed to build the prerequisites, $E_{T,\text{coneXY}}$ and $p_{T,\text{coneXY}}$. The $E_{T,\text{coneXY}}$ refers to the integration of E_T of other topological clusters, within the cone of $\Delta R = 0.\text{XY}$ near the candidate electron cluster. On the other hand, the $p_{T,\text{coneXY}}$ is the sum of p_T of different tracks coming from the hard-scatter vertex, within a cone of $\Delta R = \min(0.\text{XY}, 10\ \text{GeV}/p_T)$ near the candidate electron track. Typically for electrons the value of XY is 20. Various different working points are defined considering the requirements on $E_{T,\text{coneXY}}/p_T$ and $p_{T,\text{coneXY}}/p_T$. The two working points used in this analysis are listed in Table 4.1.

Working Point	Calorimeter isolation	Track isolation
FCLoose	$E_{T,\text{cone20}}/p_T < 0.2$	$p_{T,\text{cone20}}/p_T < 0.15$
FCTight	$E_{T,\text{cone20}}/p_T < 0.06$	$p_{T,\text{cone20}}/p_T < 0.06$

Table 4.1: Overview of different electron isolation working points used in this analysis

Calibration

The calibration refers to the correction of the electron energy to account for the energy loss in different layers of the EMC [65, 66]. A multivariate regression algorithm is employed to consider the impact of the material in front of the calorimeter. The inconsistency of the detector response arising from irregular magnetic fields in specific regions of the calorimeter is also corrected. Subsequently, an in-situ method using data of $Z \to ee$ events is employed to correct remaining disagreement in the energy scale and resolution.

4.2 Muons

Muons are reconstructed from tracks and hits as they pass through the ID and MS, respectively [67]. The reconstruction algorithms construct the muon tracks by combining information from the two independent subdetectors. Isolation and identification algorithms are employed to decide which candidate is accepted.

Reconstruction and identification

Muon reconstruction in the ID is identical to the electron reconstruction illustrated earlier. Hough transformation [68] is used in the MS to search for hit patterns to construct track segments in each layer. However, the reconstruction begins from the middle layer, which has two trigger chambers, where the track segments usually contain more hits. Muon track candidates are constructed by extending the segments in the inner and outer layer segments. An MS track can be constructed from a minimum of two segments, excluding the MS barrel endcap transition area, where one high-quality segment is enough. All hits relevant to the candidate are fitted, and the tracks are accepted depending on a global χ^2 fit.

After that, an integrated muon reconstruction considers the information from the ID, calorimeters and MS. In this context, five distinct muon candidate types are feasible [69]:

- *Combined* (CB): Muons reconstructed by doing a combined fit starting from tracks reconstructed in the MS and extrapolating to ID tracks, considering the energy loss in the calorimeters, are identified as CB muons.
- *Inside-out* (IO): IO muons are reconstructed using an additional IO algorithm which extrapolates ID tracks to the MS. The algorithm searches for atleast three MS hits, which could be loosely aligned and does not depend on independently reconstructed MS tracks. Then a combined fit using the ID track, the energy loss in the calorimeters and the hits in the muon system is performed.
- *Segment tagged* (ST): ST Muons are identified if an ID track extrapolated to the MS fulfils an additional requirement of tight angular matching to atleast one segment reconstructed in the MS. ST muons adopt the measured parameters of the associated ID track.
- *Calorimeter tagged* (CT): A trajectory in the ID is identified as a CT muon if the associated energy depositions in the calorimeters are compatible with a minimum ionizing particle. The reconstruction algorithm of CT muons accepts ID tracks a p_T threshold of 5 GeV to avoid large contamination from background contributions at low p_T.
- *StandAlone* or *MS extrapolated* (ME): In the $2.5 < |\eta| < 2.7$ region, where the ID provides no coverage, tracks reconstructed in the MS are extrapolated back to the interaction point.

Thereafter, the muon candidate collection is finalised after resolving overlaps from all these categories. The muon identification algorithm recognizes prompt muons and suppresses fake muons, arising from falsely reconstructed charged muons, by requiring specific criteria on the quality of tracks. Three working points, `Loose`, `Medium` and `Tight`, are defined in ATLAS and delivered for diverse physics analyses.

The `Medium` working point is the default at ATLAS and is designed to minimize the systematic uncertainties associated with muon reconstruction and calibration. Only CB and IO muon types are used within $|\eta| < 2.5$ of the ID range. It requires q/p compatibility (defined in [69]) to be less than seven and to have atleast two precision stations, except in region $|\eta| < 0.1$, where only one precision station is required. "The number of precision stations of a muon is defined as the number of MS stations in which the muon has at least three hits in the MDT or CSC detectors" [69]. The reconstruction efficiency of `Medium` muons is over 95%.

The `Loose` working point is designed to maximize the identification efficiency while maintaining reasonable track quality and accepts all muons passing the `Medium` working point and CT and ST muons in the range $|\eta| <0.1$ are also included. For IO muons with p_T below 7 GeV, only one precision station is accepted in the range $|\eta| <1.3$, if they can also be categorised as ST muons.

The `Tight` working point is designed to maximize the selection purity and also uses only combined and extrapolated muon types. It includes only CB and IO muons passing the `Medium` working point, with at least two precision stations. Additional criteria on q/p compatibility are imposed to exclude contamination from low-p_T muons.

Additional working points, `Low/High-`p_T, are also introduced. The `low-`p_T working point maximizes the selection efficiency of low-p_T muons, while the `high-`p_T working point maximizes the momentum resolution of high-p_T muons, both using all muon types.

Isolation

Muon isolation helps distinguish prompt muons from those coming from decay of jet constituents. Similar to electrons, a track-based quantity $p_{T,\text{coneXY}}$ and calorimeter-based quantity $E_{T,\text{coneXY}}$ are used. The XY values are 20 and 30 for $E_{T,\text{coneXY}}$ and $p_{T,\text{coneXY}}$, respectively. Different working points are defined by requiring different conditions on these quantities. The working points used in this analysis are shown in Table 4.2.

Working Point	Calorimeter isolation	Track isolation
FCLoose	$E_{T,\text{cone20}}/p_T < 0.3$	$p_{T,\text{cone30}}/p_T < 0.15$
FCTight	$E_{T,\text{cone20}}/p_T < 0.06$	$p_{T,\text{cone30}}/p_T < 0.06$

Table 4.2: Overview of different muon isolation working points used in this analysis

Calibration

Only combined muons are used for the estimation of muon calibration [70]. The calibration involves a combined weighted average of momentum and resolution correction of the ID and MS tracks. The correction factors are estimated by comparing $Z \to \mu\mu$ events in data and Monte Carlo simulations, passing the `Medium` identification, and validated by $J/\psi \to \mu\mu$ events

4.3 Jets

Jets are reconstructed as a group of particles coming from the hadronisation of quarks and gluons. They deposit their energy in the calorimeters. At first, calorimeter clusters are rebuilt which represents an entity and then the jet algorithm is used. The jet reconstruction algorithm then clusters collimated objects to create a jet and extrapolates the initial characteristics of the parton.

Reconstruction

Jets are used to tag the VBS event topology, and to veto the b-quark-induced jets (b-jets) in the event. The most renowned jet algorithm is the anti-k_t algorithm [71] which is used by ATLAS in various research. The distance between a couple of objects i and j, d_{ij}, is defined as:

$$d_{ij} = \min\Big(\frac{1}{p_{T,i}^2}.\frac{1}{p_{T,j}^2}\Big)\frac{\Delta R_y^2}{R^2} \tag{4.1}$$

where p_T is the transverse momentum of every object, ΔR_y is the angular distance and R is radius parameter whose value is fixed at 0.4 in this research. Topo-clusters are used as they combine information of the energy deposited in topologically connected calorimeter cells [72].

In this analysis, we use the `AntiKt4EMPFlow` jets which employ the particle flow (PFlow) algorithm. The PFlow algorithm [73] provides a list of tracks and a list of topo-clusters containing both the unmodified topoclusters and a set of new topo-clusters resulting from the energy subtraction procedure. First, well-measured tracks are selected. The algorithm then attempts to match each track to a single topo-cluster in the calorimeter. The expected energy in the calorimeter, deposited by the particle that also created the track, is computed based on the topo-cluster position and the track momentum. It is relatively common for a single particle to deposit energy in multiple topo-clusters. For each track/topo-cluster system, the algorithm evaluates the probability that the particle energy was deposited in more than one topo-cluster. On this basis, it decides if it is necessary to add more topo-clusters to the track/topo-cluster system to recover the full shower energy. The expected energy deposited in the calorimeter by the particle that produced the track is subtracted cell by cell from the set of matched topo-clusters. Finally, if the remaining energy in the system is consistent with the expected shower fluctuations of a single particle's signal, the topo-cluster remnants are removed.

Identification

Numerous observables are constructed to differentiate hard-scatter jets from pile-up (PU) jets. The jet-vertex-fraction (JVF) [74] is defined by the p_T of the tracks which are correlated with a specific jet:

$$\mathrm{JVF} = \frac{\Sigma_i p_T(\mathrm{trk}_i)}{\Sigma p_T(\mathrm{trk}_i) + \Sigma_j p_T(\mathrm{trk}_j)} \tag{4.2}$$

where i runs over all the hard-scatter tracks while j runs over the remaining tracks associated with that jet, such as PU jets. JVF gives an estimate of the probability that the jet comes from a hard-scatter vertex. An alternative quantity, $corrJVF$ is beneficial for minimizing the dependency on number of vertices PU tracks:

$$\mathrm{corrJVF} = \frac{\Sigma_i p_T(\mathrm{trk}_i)}{\Sigma p_T(\mathrm{trk}_i) + \frac{\Sigma_j p_T(\mathrm{trk}_j)}{k.N_{\mathrm{trk}}(\mathrm{PU})}} \tag{4.3}$$

Here, k has a value of 0.01 which is validated, although, the performance is not sensitive to the choice of k [75]. An additional variable is R_{pT} which is described as the ratio of the scalar sum of the p_T of all hard-scatter tracks of the jet and fully calibrated p_T of the jet:

$$R_{pT} = \frac{\Sigma_i p_T(\mathrm{trk}_i)}{p_T(\mathrm{jet})} \tag{4.4}$$

where R_{pT} reaches 0 specifically for PU jets. It provides substantial supplementary information to $corrJVF$. Therefore $corrJVF$ and R_{pT} are combined to make a new discriminant, Jet-vertex-tagger (JVT). JVT is a combined

discriminant using a two-dimensional likelihood algorithm based on k-nearest neighbours. The training sample comprises PU jets and hard-scatter jets with 20 GeV $< p_T <$ 50 GeV and $|\eta| < 2.4$.

Calibration

Jet energies are corrected [76] for detector inhomogeneities, the non-compensating nature of the calorimeter, and the impact of multiple overlapping pp interactions. Correction factors are derived using test beam, cosmic ray, pp collision data, and a detailed `Geant4` detector simulation.

4.3.1 Tagged Jets

The identification of b-jets, jets that originate from the hadronisation of b-quarks to b-hadrons, relies on the properties of the production and weak decay of b-hadrons, which allow us to distinguish them from jets which contain only lighter quarks or jets originating from c-quarks. Generally, the b-hadron comprises approximately 70% of the actual b-quark momentum. Considering its huge mass, the decaying products of a b-hadron may have large transverse momentum and a large opening angle. Most importantly, b-hadrons have a relatively large lifetime τ, of the order of 1.5 ps. As a result, it leads to the creation of a secondary vertex. The process of defining b-jets, called b-tagging, is conducted using multivariate analysis which combines crucial information.

b-tagging algorithm

There are three types of basic algorithms used for b-tagging at ATLAS: impact-parameter-based algorithms, secondary vertex reconstruction algorithms, and decay chain multi-vertex reconstruction algorithms [77, 78]. The impact parameter based algorithm combines large impact parameters of tracks from the b-hadron decay. It constructs two likelihood ratios between b-jet and other flavour jets. One such ratio, IP3D, employs transverse (z_0) and longitudinal (d_0) impact parameter and their correlation, while the other one, IP2D, considers only the transverse impact parameter. The secondary vertex finding algorithm reconstructs a secondary vertex and inspects its features such as the invariant mass of objects, the track numbers and so on. The multi-vertex fit algorithm tries to reconstruct the whole b-hadron decay chain by combining information from multiple vertices. This approach can also resolve b and c-hadron vertices.

Two multivariate algorithms based on different architectures, MV2 and DL1, combine all the output variables of the above three algorithms to build a global discriminant for the jet flavour. The DL1 NN has a multidimensional output corresponding to the probabilities for a jet to be a b-, c- or light-flavour jet. Its topology consists of a mixture of fully-connected hidden layers. The input variables for the DL1 NN are p_T and η of the jets. For each selected track, various track parameters like the transverse (z_0) and longitudinal (d_0) impact parameters, the fraction of transverse momentum carried by the track relative to the jet p_T (p_T^{frac}), the angular distance between the track and the jet-axis (ΔR(track, jet)) and the track likelihood ratios, defined previously as in the IP2D and IP3D algorithms, are fed to the network. In our analysis, we use the

DL1r high-level tagger which is based on an DL1 NN combined with a low-level tagger, RNN.

Working points

With an aim of harmonisation and centralisation of the systematic uncertainty estimation efforts, various b-tagging working points are recommended and provided for different analyses in the ATLAS collaboration. They are described by a single cut on the DL1 output, as shown in Fig. 4.2, so that a certain b-jet efficiency is satisfied.

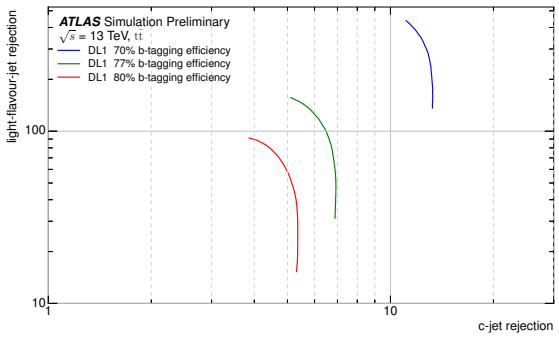


Figure 4.2: DL1 light-flavour vs c-jet rejection for a set of working points corresponding to different values of the b-jet efficiency. The evaluation is performed on $t\bar{t}$ events [77]. Rejection is the probability of mistakenly tagging as a b-jet a jet originating from a light-flavour parton.

4.4 Missing transverse momentum

The missing momentum in the transverse plane is calculated to detect particles which do not leave any track or deposit energy, such as neutrinos. It depends heavily on the cluster, tracks and jet measurement and is an essential part of the reconstruction of the event. The x and y components of the missing transverse momentum, employing calibrated hard-scatter objects, are represented as [79]

$$E_{x,y}^{\text{miss}} = E_{x,y}^{\text{miss},e} + E_{x,y}^{\text{miss},\gamma} + E_{x,y}^{\text{miss},\tau} + E_{x,y}^{\text{miss, jets}} + E_{x,y}^{\text{miss},\mu} + E_{x,y}^{\text{miss, soft}} \tag{4.5}$$

where every object term is the negative vectorial sum of the momenta of the respective calibrated objects in the event. The soft term arises from contributions of objects that derive from hard-scatter vertex but are impossible tp reconstruct as other identified objects. It can be calculated using low-p_T tracks, deriving a track-based soft term, or employing low energy topological clusters, and getting a calorimeter-based soft term. As the tracker is insensitive to the PU tracks and can catch particles that are unable to reach the calorimeters, the track-based soft term is favoured. For this reason, the soft-term can only be calculated in the central region with $|\eta| < 2.5$. From the x and y components, the magnitude can be computed as:

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \tag{4.6}$$

Another important quantity is the scalar sum of transverse momenta of those objects, ΣE_T defined as

$$\Sigma E_T = \Sigma p_T^e + \Sigma p_T^\gamma + \Sigma p_T^\tau + \Sigma p_T^{\text{jets}} + \Sigma p_T^\mu + \Sigma p_T^{\text{soft}} \tag{4.7}$$

Chapter 5

Data and Monte Carlo

"In times of crisis, the wise build bridges while the foolish build barriers."

Black Panther

Testing of theoretical predictions through experiments require a large amount of data prior to the analysis. For this thesis, the proton-proton collision data has been collected from the ATLAS detector at the LHC. Also, simulated data has been generated to mimic the SM and the theoretical models, as described in Section 2.2, which are used to optimize the analysis chain and compare with the collision data.

5.1 Data sample

This analysis is based on the data sample obtained by proton-proton collisions at a center of mass energy of $\sqrt{s} = 13$ TeV by the ATLAS detector at the LHC from 2015 to 2018. Recorded events are grouped into units of luminosity of increasing magnitude during data taking. Luminosity blocks are the smallest unit and mark an interval of stable detector conditions of approximately two minutes length. Events are selected based on data quality flags per luminosity block, using the Good Run List (GRL) [80, 81]. This vetoes the events which occurred during those luminosity blocks where the detector was not fully operational. For this analysis, the used Good Runs Lists are:

- `data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns`
- `data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns`
- `data17_13TeV.periodAllYear_DetStatus-v97-pro21-17_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim`
- `data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim`

The combined integrated luminosity as reported by the ATLAS luminosity calculation tool [82] is 139 fb^{-1} with an uncertainty of 1.7%.

Selected events are required to have triggered either the single-electron or the single-muon trigger [83, 84]. It is required that at least one of the analysis

leptons is responsible for the trigger of the events by using the trigger matching tool. The triggers are summarized in Table 5.1. The name of each trigger provides information about the associated requirements that go along with it. The name consists of a combination of tags that can be interpreted. The tag `eXX` requires electrons with a transverse momentum of at least `XX` GeV. The same naming scheme holds for muons, which are indicated by `mu` instead. The tags like `iloose`, `ivarmedium`, `lhloose`, `lhmedium` and `lhtight` succeeding a lepton requirement indicate isolation and identification requirements applied in the High-Level Trigger (HLT). The tag `nod0` for the electron triggers simply indicates that the impact parameter d_0 information is not used for said reconstruction.

	2015	2016-2018
Single muon	`HLT_mu20_iloose_L1MU15` `HLT_mu50`	`HLT_mu26_ivarmedium` `HLT_mu50`

Single electron	`HLT_e24_lhmedium_L1EM20VH`
	`HLT_e60_lhmedium`
	`HLT_e120_lhloose`
		`HLT_e26_lhtight_nod0_ivarloose`
		`HLT_e60_lhmedium_nod0`
		`HLT_e140_lhloose_nod0`

Table 5.1: Summary of the single lepton HLT triggers that are used in this analysis.

This analysis uses di-lepton filter samples `DAOD_STDM5` derivations (derivation format 5 of the Standard Model physics group). The derivation formats of the Standard Model group, in general, are designed to provide the minimal set of information necessary to apply all corrections by the combined performance groups for all reconstructed particle types and are thus widely usable. The `DAOD_STDM5` format applies a strong skimming to store only events with at least three identified leptons with $p_T > 6$ GeV. The p-tag used is `p4095`, `p4096`, and `p4097`

5.2 Signal Monte Carlo samples

Simulated events of signal and SM background processes with prompt leptons were generated based on Monte Carlo methods. The `Geant4` toolkit [85, 86] within the ATLAS software framework was used to simulate the geometry, material and response of the ATLAS detector and its parts. For some samples, `FastCaloSim` (Fast Calorimeter Simulation) package was used to provide a parametrised simulation of the particle energy response and the energy distribution in the calorimeters because it reduces the time of calorimeter simulation. Table 5.2 summarises the different strategies used in this analysis.

5.2.1 HVT signals

Simulated signal events are used to model the benchmark physics processes and to optimize the selection criteria. HVT signal samples for Drell-Yan production $qq' \to W' \to WZ \to lvll$ were used to interpret results. All lepton flavours were included in $Z \to \ell\ell$ and $W \to \ell\nu$ decays. The samples for Model A were generated with `Madgraph 2.6.5` [87] interfaced to `Pythia 8.186` [88], for

Strategy	Tag	Description
Full `Geant4` Simulation	`Fullsim`	Detailed detector simulation
Atlfast-II	`AF2`	`FastCaloSim`

Table 5.2: Different detector simulation strategies provided in ATLAS and used in this analysis

parton level hadronization, at leading order (LO) in QCD using the model file provided by [26]. The `NNPDF23_lo_as_0130_qed` was used as PDF set and A14 hadronisation tune [89] to create the samples. The HVT model A signal samples are generated with $g_V = 1$ and the parameter c_F is assumed to be the same for all types of fermions for masses of diboson resonances ranging from 250 GeV to 5 TeV. Information about the DY signal samples is summarised in Table 5.3.

DSID	**Mass**	**Events**	**Filter eff.**	**Cross-section [fb]**	**k-factor**
307376	250	60000	1.00	4110	1.00
307377	300	50000	1.00	2260	1.00
307378	400	60000	1.00	717	1.00
302266	500	30000	1.00	283.9	1.00
302267	600	30000	1.00	132.2	1.00
302268	700	18000	1.00	69.23	1.00
302269	800	25000	1.00	39.41	1.00
302270	900	30000	1.00	23.80	1.00
302271	1000	45000	1.00	15.09	1.00
302272	1100	18000	1.00	9.921	1.00
302273	1200	30000	1.00	6.751	1.00
302274	1300	10000	1.00	4.676	1.00
302275	1400	30000	1.00	3.313	1.00
302276	1500	29000	1.00	2.391	1.00
302277	1600	30000	1.00	1.749	1.00
302278	1700	28000	1.00	1.298	1.00
302279	1800	30000	1.00	0.9720	1.00
302280	1900	30000	1.00	0.7350	1.00
302281	2000	18000	1.00	0.5610	1.00
302282	2200	35000	1.00	0.3334	1.00
302283	2400	45000	1.00	0.2028	1.00
302284	2600	30000	1.00	0.1258	1.00
302285	2800	29000	1.00	0.07925	1.00
302286	3000	20000	1.00	0.05036	1.00
302287	3500	25000	1.00	0.01692	1.00
302288	4000	30000	1.00	0.00587	1.00
302289	4500	25000	1.00	0.00200	1.00
302290	5000	30000	1.00	0.00073	1.00

Table 5.3: Summary of the LO HVT signal MC simulation for the Drell-Yan analysis.

For the VBF resonance search signal samples for HVT Model C are generated with `Madgraph 2.6.5` interfaced to `Pythia 8.186` at LO in QCD for masses of diboson resonances ranging from 250 GeV to 2 TeV [26]. The coupling parameters g_V and c_H of the heavy triplet are set to one except the coupling to fermions $g_F = c_F(g^2/g_V)$, where c_F is set to 0. The W and Z decay to all three lepton flavours (e, μ, τ) and no lepton filter is applied. A dijet invariant mass of at least 150 GeV is required in this case at event generation. Information about the LO signal VBF HVT MC samples is summarised in Table 5.4

DSID	Mass	Events	Filter eff.	Cross-section [fb]	k-factor
307730	250	200000	1.00	24.196	1.00
313538	275	200000	1.00	22.7	1.00
307731	300	200000	1.00	10.54	1.00
313539	325	200000	1.00	6.71	1.00
313540	350	200000	1.00	4.524	1.00
313541	375	200000	1.00	3.19	1.00
307732	400	200000	1.00	2.299	1.00
313542	425	200000	1.00	1.73	1.00
313543	450	200000	1.00	1.32	1.00
313544	475	200000	1.00	1.02	1.00
307733	500	200000	1.00	0.7975	1.00
313545	525	200000	1.00	0.639	1.00
313546	550	200000	1.00	0.515	1.00
307734	600	200000	1.00	0.3408	1.00
307735	700	200000	1.00	0.1663	1.00
307736	800	200000	1.00	0.0879	1.00
307737	900	200000	1.00	0.04988	1.00
307738	1000	200000	1.00	0.02961	1.00
307739	1100	200000	1.00	0.01814	1.00
307740	1200	200000	1.00	0.01171	1.00
307741	1300	200000	1.00	0.007621	1.00
307742	1400	200000	1.00	0.005078	1.00
307743	1500	200000	1.00	0.003449	1.00
307744	1600	200000	1.00	0.002391	1.00
307745	1700	200000	1.00	0.001673	1.00
307746	1800	200000	1.00	0.001192	1.00
307747	1900	200000	1.00	0.0008536	1.00
307748	2000	200000	1.00	0.0006183	1.00

Table 5.4: Summary of the LO HVT signal MC simulation for the VBF analysis.

5.2.2 GM signals

For the GM VBF model two sets of signal samples were produced with LO (see Appendix B.1) and next-to-leading order (NLO) accuracy in the mass range 200 to 1000 GeV vetoing W or Z bosons in the s-channel. NLO GM samples are the ones used in the analysis. Samples were produced with `MadGraph 2.7.2` [90] and `Pythia 8.186` for parton shower and hadronisation modelling. All lepton flavours were included in $Z \to \ell\ell$ and $W \to \ell\nu$ decays. The parameters were calculated with `GMCALC` [91], according to the *H5Plane* benchmark [92]. The parameter $\sin(\theta_H)$ was set to 0.5 up to and including mass 800 GeV, and 0.25

above 800 GeV, consistent with existing constraints [33]. The matrix element calculation was done using the `NNPDF30NLO` [93] PDF sets using the A14 tuning parameters and with the dipole recoil shower scheme to prevent the generation of an excess of central jet radiation [94]. The samples used in the analysis together with the $p_T(\mathrm{j})> X$ values used in the generation are summarized in Table 5.5.

DSID	**Mass**	**Events**	$\sin(\theta_H)$	p_T**(j)**	**Filter eff.**	**Cross-section [fb]**
450765	200	200000	0.5	10.	0.4444	23.8
502511	225	200000	0.5	10.	0.4466	20.31
450766	250	200000	0.5	10.	0.4490	17.54
502512	275	200000	0.5	10.	0.456	15.04
450767	300	200000	0.5	10.	0.4605	13.78
502513	325	200000	0.5	10.	0.469	11.33
450768	350	200000	0.5	10.	0.4782	10.54
502514	375	200000	0.5	10.	0.476	9.147
450769	400	200000	0.5	10.	0.4871	8.215
502515	425	200000	0.5	10.	0.483	7.116
450770	450	200000	0.5	10.	0.4926	6.342
502516	475	200000	0.5	10.	0.484	5.291
450771	500	200000	0.5	10.	0.5011	4.979
502517	525	200000	0.5	20.	0.492	3.902
502518	550	200000	0.5	20.	0.501	3.449
502519	600	200000	0.5	20.	0.502	2.782
502520	700	200000	0.5	20.	0.510	1.994
502521	800	200000	0.5	20.	0.514	1.327
502522	900	200000	0.25	20.	0.519	0.2495
502523	1000	200000	0.25	20.	0.522	0.1842

Table 5.5: Summary of GM NLO signal MC simulations with all lepton flavour final states for the VBF analysis.

5.3 SM background simulation

5.3.1 WZ SM background

The $W^{\pm}Z$ SM process is the major background in this analysis. The dominant background WZ-QCD samples were modelled using `Sherpa` 2.2.2 [95], and the process is defined as $pp \to \ell\nu\ell\ell$. A study was performed before the production of this sample using a truth sample where the bosons were considered as on-shell if the invariant mass of the leptons is within the $15 \times \Gamma_{boson}$ window of the boson mass. It was concluded that the Z bosons are produced on-shell while some off-shell W production was allowed as shown in Figure 5.1. This WZ-QCD sample includes processes upto order 4 in the electroweak coupling constant, α_{EW}, upto one jet at NLO and second or third jet at LO and merged with parton shower. An alternative NLO WZ-QCD sample has also been produced to estimate an uncertainty due to generator and parton shower modelling using `MadGraph 5.2.6.5` with FxFx merging [96] of up to two extra jets and hadronization performed with `Pythia 8.186`.

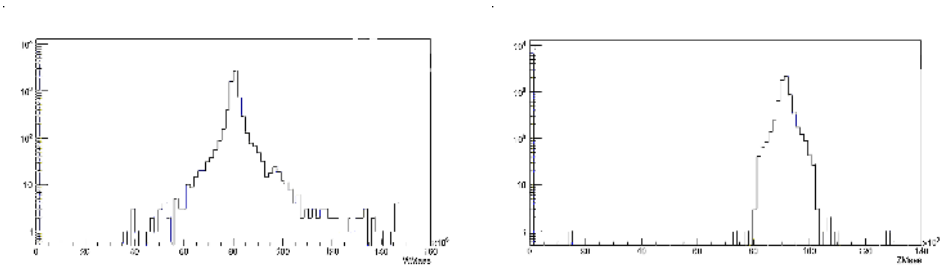


Figure 5.1: Distribution of matrix element $m_{l\nu}$ invariant mass corresponding to W boson (left) and m_{ll} corresponding to Z boson (right) in the fiducial phase-space defined by having minimum 3 leptons, with Z leptons having same flavour opposite charge & $p_T > 15$ GeV and W lepton having $p_T > 20$ GeV, transverse mass of W to be minimum 30 GeV and rapidity separation between the two boson to be more than 2.5. The x-axis represents the mass of the boson in MeV and the y-axis shows the number of events. Here only the channels were the Z and W decay to different lepton flavours are studied (i.e. $ee\mu$, $ee\tau$, $\mu\mu e$, $\mu\mu\tau$, $\tau\tau e$, $\tau\tau\mu$, where the third lepton refers to the W boson) and same flavour channels are set at bin 0.

The WZ SM background also includes a separate $WZjj$-EW sample modelled using `Madgraph 2.7.3` with `Pythia 8.244`. This sample is produced with zero-order in the strong coupling constant, α_s, leading to WZ final states associated with at least two jets originating from electroweak vertices from matrix-element partons. The initial and final states include processes with a b quark, therefore the contributions from the single top process tZj are also included in this sample. An additional sample using the same `MadGraph` 2.7.3 matrix element and `Herwig 7.2.1` for parton shower was produced to estimate an uncertainty due to the parton shower modelling. Samples for small interference between electroweak and QCD WZ production were modelled using `MadGraph` 2.7.3 + `Pythia 8.244` and their contribution is added with the WZ-EWK background.

5.3.2 Other SM backgrounds

The samples for $q\bar{q} \to ZZ \to 4\ell$, $q\bar{q} \to ZZ \to \ell\ell\ \nu\nu$ and triboson are generated with `Sherpa 2.2.2` [97], which provides full modelling of the high-energy pp collisions, including hard-scattering, parton shower, hadronization and underlying event, using matrix elements at NLO accuracy in QCD. The samples include off-shell effects and Higgs boson contributions. The purely electroweak process $q\bar{q} \to ZZjj \to 4\ell jj$ and the $gg \to ZZ$ process were also generated with `Sherpa 2.2.2`. Samples were generated using the `NNPDF3.0NNLO` PDF set [93] and `Sherpa` parton-shower parameter values.

`Madgraph 2.3.3` is used to generate samples for $t\bar{t}V$ at NLO in QCD with `NNPDF3.0NNLO` PDF and interfaced with `Pythia 8.210` using the A14 tune and the `NNPDF2.3LO` [93] PDF set. The non-prompt background samples including $Z\gamma$, $W\gamma$, Drell-Yan $Z \to \ell\ell$, $W \to \ell\nu$ as well as top-quark pairs and single top-quark have been produced to estimate the fake/non-prompt lepton background contribution. Samples with $V\gamma$ in the final state were simulated with the `Sherpa 2.2.4` generator with LO accuracy in QCD for up to three additional parton emissions using the `NNPDF3.0NNLO` PDF set and `Sherpa`

parton-shower parameter values. Drell-Yan $Z \to \ell\ell$ and $W \to \ell\nu$ were produced with POWHEGBOX v1 generator [98–101] at NLO accuracy interfaced with Pythia 8.186 using CTEQ6L1 PDF set [102] for parton shower and CT10NLO PDF set for hard -scattering processes. The top-quark pairs and single top-quark productions samples were generated with the POWHEGBOX v2 generator at NLO with NNPDF2.3LO PDF set interfaced with Pythia 8.230 using A14 tune [89] and the NNPDF2.3LO PDF set. The VVV samples are generated with Sherpa 2.2.2 using NNPDF3.0NNLO PDF set.

The list of SM background MC samples used for background estimate is shown in Table 5.6. The cross-sections include the indicated branching ratios and are used to normalize the samples, multiplied by filter efficiencies to account for any phase-space restrictions applied when generating the samples.

DSID	Process	Generators	PDF	Filter eff.	Cross-section [pb]	k-factor	status
364253	$WZ \to \ell\nu\ell\ell$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	4.583	1.00	nominal
361292	$WZ \to \ell\nu\ell\ell$	Madgraph+Pythia8	A14NNPDF23LO	0.326	1.720	1.00	alternative
361293	$WZ \to \ell\nu\ell\ell$	Madgraph+Pythia8	A14NNPDF23LO	0.326	1.720	1.00	alternative
364739	WZjj + tZ: $WZ \to e^-\nu\mu\mu$, $WZ \to \mu^-\nu ee$	MadGraph+Pythia8	NNPDF3.0NNLO	1.00	0.01543	1.0	nominal
364740	WZjj + tZ: $WZ \to e^+\nu\mu\mu$, $WZ \to \mu^+\nu ee$	MadGraph+Pythia8	NNPDF3.0NNLO	1.00	0.02567	1.0	nominal
364741	WZjj + tZ: $WZ \to \mu^-\nu\mu\mu$, $WZ \to e^-\nu ee$	MadGraph+Pythia8	NNPDF3.0NNLO	1.00	0.00770	1.0	nominal
364742	WZjj + tZ: $WZ \to \mu^+\nu\mu\mu$, $WZ \to e^+\nu ee$	MadGraph+Pythia8	NNPDF3.0NNLO	1.00	0.01281	1.0	nominal
830000	WZjj + tZ: $WZ \to e^-\nu\mu\mu$, $WZ \to \mu^-\nu ee$	MadGraph+Herwig	CT10 (ME), MMHT2014lo68cl (shower/MPI)	1.00	0.0154	1.0	alternative
830001	WZjj + tZ: $WZ \to e^+\nu\mu\mu$, $WZ \to \mu^+\nu ee$	MadGraph+Herwig	CT10 (ME), MMHT2014lo68cl (shower/MPI)	1.00	0.0257	1.0	alternative
830002	WZjj + tZ: $WZ \to \mu^-\nu\mu\mu$, $WZ \to e^-\nu ee$	MadGraph+Herwig	CT10 (ME), MMHT2014lo68cl (shower/MPI)	1.00	0.00770	1.0	alternative
830003	WZjj + tZ: $WZ \to \mu^+\nu\mu\mu$, $WZ \to e^+\nu ee$	MadGraph+Herwig	CT10 (ME), MMHT2014lo68cl (shower/MPI)	1.00	0.0128	1.0	alternative
364250	$q\bar{q} \to ZZ \to \ell\ell\ell\ell$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	1.2523	1.00	nominal
364254	$q\bar{q} \to ZZ \to \ell\ell\nu\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	1.2501	1.00	nominal
364283	$q\bar{q} \to ZZjj \to \ell\ell\ell\ell jj$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.01057	1.00	nominal
345705	$gg \to ZZ \to \ell\ell\ell\ell$ ($m_{4\ell}$<130)	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.0099577	1.00	nominal
345706	$gg \to ZZ \to \ell\ell\ell\ell$ ($m_{4\ell}$>130)	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.010163	1.00	nominal
366140-364149	LO $Z\gamma$	Sherpa 2.2.4	NNPDF3.0NNLO	1.00	46.3	1.00	nominal
361106	$Z \to ee$	Powheg+Pythia8	CTEQ6L1	1.00	1901.2	1.026	nominal
361107	$Z \to \mu\mu$	Powheg+Pythia8	CTEQ6L1	1.00	1901.2	1.026	nominal
361108	$Z \to \tau\tau$	Powheg+Pythia8	CTEQ6L1	1.00	1901.2	1.026	nominal
361100	$W^+ \to e\nu$	Powheg+Pythia8	CTEQ6L1	1.00	11306.0	1.0172	nominal
361101	$W^+ \to \mu\nu$	Powheg+Pythia8	CTEQ6L1	1.00	11306.0	1.0172	nominal
361102	$W^+ \to \tau\nu$	Powheg+Pythia8	CTEQ6L1	1.00	11306.0	1.0172	nominal
361103	$W^- \to e\nu$	Powheg+Pythia8	CTEQ6L1	1.00	8283.1	1.0357	nominal
361104	$W^- \to \mu\nu$	Powheg+Pythia8	CTEQ6L1	1.00	8283.1	1.0357	nominal
361105	$W^- \to \tau\nu$	Powheg+Pythia8	CTEQ6L1	1.00	8282.9	1.0357	nominal
410470	$t\bar{t}$ ($\geqslant 1\ell$)	Powheg+Pythia8	NNPDF2.3LO	0.54382	729.77	1.1398	nominal
410155	$t\bar{t}W$	Madgraph+Pythia8	NNPDF2.3LO	1.00	0.54822	1.096	nominal
410218	$t\bar{t}Z(ee)$	Madgraph+Pythia8	NNPDF2.3LO	1.00	0.036865	1.12	nominal
410219	$t\bar{t}Z(\mu\mu)$	Madgraph+Pythia8	NNPDF2.3LO	1.00	0.036868	1.12	nominal
364242	$WWW \to 3\ell 3\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.0071979	1.00	nominal
364243	$WWZ \to 4\ell 2\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.0017966	1.00	nominal
364244	$WWZ \to 2\ell 4\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.0035467	1.00	nominal
364245	$WZZ \to 5\ell 1\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.00018807	1.00	nominal
364246	$WZZ \to 3\ell 3\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	0.4463	0.00074436	1.00	nominal
364247	$ZZZ \to 6\ell 0\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	1.00	0.00001451	1.00	nominal
364248	$ZZZ \to 4\ell 2\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	0.2245	0.00008652	1.00	nominal
364249	$ZZZ \to 2\ell 4\nu$	Sherpa 2.2.2	NNPDF3.0NNLO	0.4447	0.00017140	1.00	nominal

Table 5.6: Summary of background MC simulation.

Chapter 6

Event and Object Selection

"The hardest choices require the strongest wills."
Avengers: Infinity Wars

The objective of this search is to observe a new physics resonance decaying to WZ. The fully leptonic channel is used. The selection starts by cleaning the ATLAS data and focus on choosing clean objects to build an inclusive WZ phase space.

6.1 Object Selection

For leptons (electrons and muons), four selections are defined: `Baseline`, `Loose`, `tight-`Z and `tight-`W. The `Baseline` leptons are used for overlap removal and to veto events with a fourth lepton. The `Loose` leptons are used for building the fake control region and the matrix method estimation. The `tight-`Z and `tight-`W leptons are assigned to Z and $W^{\pm}$ boson, respectively. It is ensured that `tight-`W leptons also pass the `tight-`Z selection and `tight-`Z leptons pass the `Baseline` selection.

6.1.1 Muons

The momentum of the reconstructed muon candidates used in this analysis are corrected by the `CP::MuonCalibrationAndSmearingTool`.

The `Baseline` muon candidates are required to have p_T>5 GeV (15 GeV for CT muons) and pass the `Loose` working point, described in Section 4.2, of the `MuonSelectionTool` as recommended by the Muon Combined Performance group. The pseudo-rapidity η of each muon candidate is required to fulfil $|\eta|$ <2.7

For muons with $p_T > 300$ GeV the `High-`p_T quality cut is applied. Cosmic muons are excluded by vetoing muons with $d_0 > 1$ mm. Cuts on impact parameters are applied to ensure that the candidates come from the primary vertex. Therefore the tracks assigned to `Baseline` muons have to fulfil $|z_0 \cdot \sin(\theta)| < 0.5$ mm with respect to the primary vertex. `Baseline` muons also need to qualify `FCLoose` isolation criteria.

The `tight-Z` and `tight-W` muons should pass the `Baseline` requirements and p_T >25 GeV. Additionally, the `tight-Z` muons pass `Medium` quality cut,

which accepts only CB and IO muons, while muons are tagged as `tight-W` if they pass the `Tight` quality cut. Both W and Z candidate muons are required to pass `FCTight` isolation working point and also required to have d_0 significance ($|d_0/\sigma(d_0)|$) of less than three, where $\sigma(d_0)$ is the uncertainty of the d_0 measurement.

Additionally, `Loose` muons are defined with the same requirements as the `tight-Z` muons except for the identification and isolation requirements and used for building the fake control region for the matrix method background estimation.

Muon object selection used in the analysis is summarized in Table 6.1.

Muon object selection				
Selection	`Baseline`	`Loose`	`tight-Z`	`tight-W`
$p_{\mathrm{T}} > 5$ GeV (15 GeV for CT muons)	✓	✓	✓	✓
$p_{\mathrm{T}} > 25$ GeV		✓	✓	✓
$\|\eta\| < 2.7$	✓	✓	✓	✓
$\|z_0 \sin\theta\| < 0.5$ mm	✓	✓	✓	✓
cosmic cut ($\|d_0\| < 1$ mm)	✓	✓	✓	✓
$\|d_0/\sigma(d_0)\| < 3$		✓	✓	✓
Loose quality (if p_T >300 GeV High-p_T quality)	✓	✓	✓	✓
FCLoose isolation	✓	✓	✓	✓
μ-jet overlap removal		✓	✓	✓
Medium quality (if p_T >300 GeV High-p_T quality)			✓	
FCTight isolation			✓	
Tight quality (if p_T >300 GeV High-p_T quality)				✓
FCTight isolation				✓

Table 6.1: Muon object selection used in the analysis.

6.1.2 Electrons

The `Baseline` electrons are required to pass `LooseLH+BLayer` identification and `FCLoose` isolation working point. The transverse momentum of `Baseline` electrons is required to be greater than 7 GeV and cluster $|\eta|$ should be less than 2.47. Similar to muons, to ensure that the electron originates from the primary vertex, it is required to have a longitudinal impact parameter $|z_0 \sin(\theta)| < 0.5$ mm.

The `tight-Z/W` electrons are the signal electrons used for the final event selection. All `tight-Z/W` electrons must firstly pass `baseline` selection. The p_{T} threshold of `tight-Z/W` electrons is raised to 25 GeV. Cluster $|\eta|$ region 1.37 to 1.52 is vetoed to avoid the detector crack region. The signal electrons must also pass the transverse impact parameter cut, which means d_0 significance should be less than 5.0. The `tight-Z` electrons are required to pass `MediumLH` identification and `FCTight` isolation working point. For the `tight-W` electrons, the `TightLH` identification working point and `FCTight` isolation working point are used.

The `Baseline` electrons are used for overlap removal and missing transverse

energy calculation. To avoid overlap between signal electrons and muons or jets, `tight-Z/W` electrons must pass the official overlap removal requirement.

Just like muons, the `Loose` criteria is also defined for electrons with the same requirements as the `tight-Z` electron except for the identification and isolation requirement.

The details of electron object selection are summarized in Table 6.2.

Electron object selection				
Selection	`Baseline`	`Loose`	`tight-Z`	`tight-W`
$p_{\mathrm{T}} > 7$ GeV	✓	✓	✓	✓
$p_{\mathrm{T}} > 25$ GeV		✓	✓	✓
$\lvert\eta^{\mathrm{cluster}}\rvert < 2.47$	✓	✓	✓	✓
Exclude $1.37 < \lvert\eta^{\mathrm{cluster}}\rvert < 1.52$		✓	✓	✓
Electron object quality	✓	✓	✓	✓
$\lvert z_0 \sin\theta\rvert < 0.5$ mm	✓	✓	✓	✓
$\lvert d_0/\sigma(d_0)\rvert < 5$		✓	✓	✓
LooseLH+BLayer identification	✓	✓	✓	✓
FCLoose isolation	✓	✓	✓	✓
e-μ and e-e overlap removal		✓	✓	✓
e-jets overlap removal		✓	✓	✓
MediumLH identification			✓	
FCTight isolation			✓	
TightLH identification				✓
FCTight isolation				✓

Table 6.2: Electron object selection used in the analysis.

6.1.3 Jets

Jet candidates are reconstructed using the anti-k_t algorithm with a radius parameter $R = 0.4$ and calibrated using a combination of `JERSmearingTool`, `JERTool` and `JetCalibrationTool` with the latest recommendations. To mitigate the pile-up contamination, jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $\eta < 2.4$ are required to have a jet-vertex-tagger (JVT) score greater than 0.2 [103]. Jets with $p_{\mathrm{T}} < 120\,\mathrm{GeV}$ and $2.5 < \eta < 4.5$ are required to have a forward jet-vertex-tagger (fJVT) score less than 0.5 and $|timing| < 10$ ns [103].

The ATLAS jet flavour tagging algorithm, `DL1r` (Section 4.3.1), is used to suppress $t\bar{t}$ background in the signal region. Out of the possible working points corresponding to different b-tagging efficiencies, the 85% working point (threshold value of 0.1758) is selected as the baseline tagger to achieve high $t\bar{t}$ reduction.

`Baseline` jets are defined as $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 4.5$ which are used to calculate missing transverse momentum and carry out overlap removal. `Baseline` jets passing pile-up removal, b-jet veto, and overlap removal are tagged as `VBS jet` candidates. Table 6.3 summarizes the jets selection.

Jet object selection		
Selection	`Baseline`	`VBS jet`
$p_{\mathrm{T}} > 30$ GeV	✓	✓
$\|\eta\| < 4.5$	✓	✓
Pile-up Removal		✓
veto b-Tagging		✓
μ-jet overlap removal		✓
e-jets overlap removal		✓

Table 6.3: Selection for small-R jets

6.1.4 Missing Transverse Energy

In this analysis, missing transverse energy is reconstructed with the `METMaker-Tool`, provided by ATLAS, using calibrated electrons and muons passing the `Baseline` selection and calibrated jets before any selection. However, it does not include contributions from soft neutral particles and from forward regions ($|\eta| > 2.5$).

6.1.5 Overlapping objects treatment

The overlap between electrons, muons and jets is handled by the `OverlapRemovalTool` with the `Standard` working point as described here [104, 105]. Additional overlap removal is then applied to keep the electron with the highest p_T if more than one electrons share the same ID track and the electron with the highest cluster E_T if two or more electrons with clusters are close to each other. To avoid the overlap between baseline muons and electrons, calorimeter tagged muons which share the same inner detector track as the electron are removed.

6.2 Event Selection

For the measurement of WZ production, candidate events are selected from data recorded by the ATLAS detector and MC simulations, where each boson decays leptonically. An inclusive WZ phase space is defined at the detector level from where the signal regions will be extracted. The fully leptonic final state gives the cleanest signature. All signal regions will be optimized on top of this cut-based selection in the next chapter. WZ leptonic decay to all final states with electrons and muons for all events are considered : $\mu^+\mu^-\mu^\pm, \mu^+\mu^-e^\pm, e^+e^-\mu^\pm, e^+e^-e^\pm$. The different event selection criteria are described in the following and are summarized in Table 6.4.

The event selection starts by requiring quality conditions on the recorded data by rejecting LAr, Tile and SCT corrupted events, as well as incomplete events. Event cleaning in both data and simulated events is done by rejecting events with at least one misidentified jet of non-collision background or detector problems origin [106], with respect to the high-efficiency `Loose` working point of selecting jets [107]. Data events are selected based on data quality flags per

luminosity block, using the Good Run List listed in Section 5.1.

Candidate events are required to have at least one primary vertex with at least two associated tracks with a transverse momentum of $p_T > 500$ MeV. If several primary vertices are found, the primary vertex selected in the analysis is the one with the largest $\Sigma(p_T)^2$ in the event. The presence of an unclean jet in an event can affect many other features of that event, including but not limited to the MET calculation and calorimeter behaviour. Thus a cleaning designed to veto events that have one or more unclean jets is applied. The event must have exactly three leptons passing the `Loose`$^-$ selection criteria, where `Loose`$^-$ is defined as `Loose` as in Section 6.1 but without d_0-Significance cut. This level is defined for reducible background estimation using the matrix method. These leptons are required to have triggered single lepton trigger, enabled during the entire run period, during the readout of the event. The trigger algorithms require the transverse momentum threshold of the leptons in 2015 is 24 GeV for electrons and 20 GeV for muons satisfying a loose isolation requirement based only on ID track information and for 2016 to 2018, the trigger threshold is increased to 26 GeV for both electrons and muons, and tighter isolation requirements were applied, due to the higher instantaneous luminosity.

The event must have two leptons of the same flavour and opposite charge as the Z lepton candidates, with an invariant mass that is consistent with the Z mass: $|M_{ll} - 91.1876| < 20$ GeV. If there are more than one pair of leptons that can form a Z, the pair with invariant mass closest to the PDG mass is taken. The third lepton is then taken as the W lepton candidate. The three `Loose`$^-$ leptons after this requirement has in total four possible combinations, namely $\mu^{\pm}\mu^{\mp}\mu^{\mp}$, $\mu^{\pm}\mu^{\mp}e^{\mp}$, $e^{\pm}e^{\mp}\mu^{\mp}$, and $e^{\pm}e^{\mp}e^{\mp}$, where the notation follows: the first and second lepton means Z lepton candidates and the third lepton means the W lepton candidate and the Z lepton with the same charge as the W lepton is always placed at the second position. The following texts share the same rule as the notation specified here. In order to decrease background from ZZ, events with 4 or more leptons passing the `Soft` criteria are vetoed. `Soft` leptons are defined as `Baseline` leptons passing the overlap removal. Additionally, missing transverse energy in the event must be greater than 25 GeV and the lepton with the highest transverse momentum assigned to any boson to have $p_T > 27$ GeV.

Inclusive event selection	
Event cleaning	Reject LAr, Tile and SCT corrupted events and incomplete events
Primary vertex	Hard scattering vertex with at least two tracks
Trigger	Single lepton trigger
Jet cleaning	pass DFCommonJets_eventClean_LooseBad
ZZ veto	Less than 4 *soft* leptons
N leptons	Exactly three leptons passing the *loose*$^-$ lepton selection
Z candidate	built from Same-Flavor-Opposite-Sign (SFOS) lepton pair with $M_{\ell\ell}$ closest to Z PDG mass
W candidate	built from the third lepton and MET
W, Z selection	Z leptons passing `tight-Z` lepton selection. W leptons passing `tight-W` lepton selection.
Mass window	$\lvert M_{\ell\ell} - M_Z \rvert < 20$ GeV
Missing Energy	$E^T_{\text{miss}} > 25$ GeV
Leading lepton p_{T}	Leading Lepton $p_{\text{T}} > 27$ GeV

Table 6.4: Overview of the analysis event selection.

Chapter 7

Search Strategy

"I think we must learn from our mistakes and do better."
Eternals

In this chapter, the strategy to select those events that could indicate the possible presence of new physics is discussed. It starts from the events passing the inclusive WZ selection described in the previous chapter and are categorized into two distinct search regions in order to discriminate Drell-Yan (DY) and VBF production modes. All final states with three charged leptons and missing transverse momentum from $W^{\pm}Z$ decays are considered for this search analysis.

The Drell-Yan category is defined by cut based selection and described in Section 7.1. For the VBF-enriched category, two analysis chains are developed, one that makes use of an Artificial Neural Network (ANN) to tag VBF events as described in Section 7.2, and a second one is using a simple cut based selection that will be described in Section 7.3 and is used as validation and can be used for future interpretations by theorists. A summary of the signal regions of the analysis is shown in Table 7.3.

7.1 Drell-Yan signal region

The transverse momenta of selected W and Z bosons from a heavy resonance produced essentially at rest in the s-channel, have almost 50% of the resonance mass. The cuts based on the ratio of the boson transverse momentum to the WZ invariant mass, $p_{\mathrm{T}}^{V}/m(WZ)$, are used for the Drell-Yan production mode. Events passing the WZ inclusive selection are required to have $p_{\mathrm{T}}^{V}/m(WZ)$ greater than 0.35 for both bosons to reduce the contribution of the non-resonant WZ production. Figure 7.1 shows the distributions of p_T^Z/M_{WZ} and p_T^W/M_{WZ} for the signal and the WZ SM backgrounds.

The combined detector acceptance and signal selection efficiency ($A \times \epsilon$) of the DY HVT W' selection, relative to the generated signal events, is shown in Figure 7.2. Decays of W and Z bosons into all flavours of leptons are included at event generation. The $A \times \epsilon$ values decrease for resonance masses above approximately 2 TeV due to the collinearity of electrons from highly boosted $Z \to ee$ decays, for which the lepton isolation is less efficient. Figure 7.3 shows the pre-fit distribution for the Drell-Yan signal region.

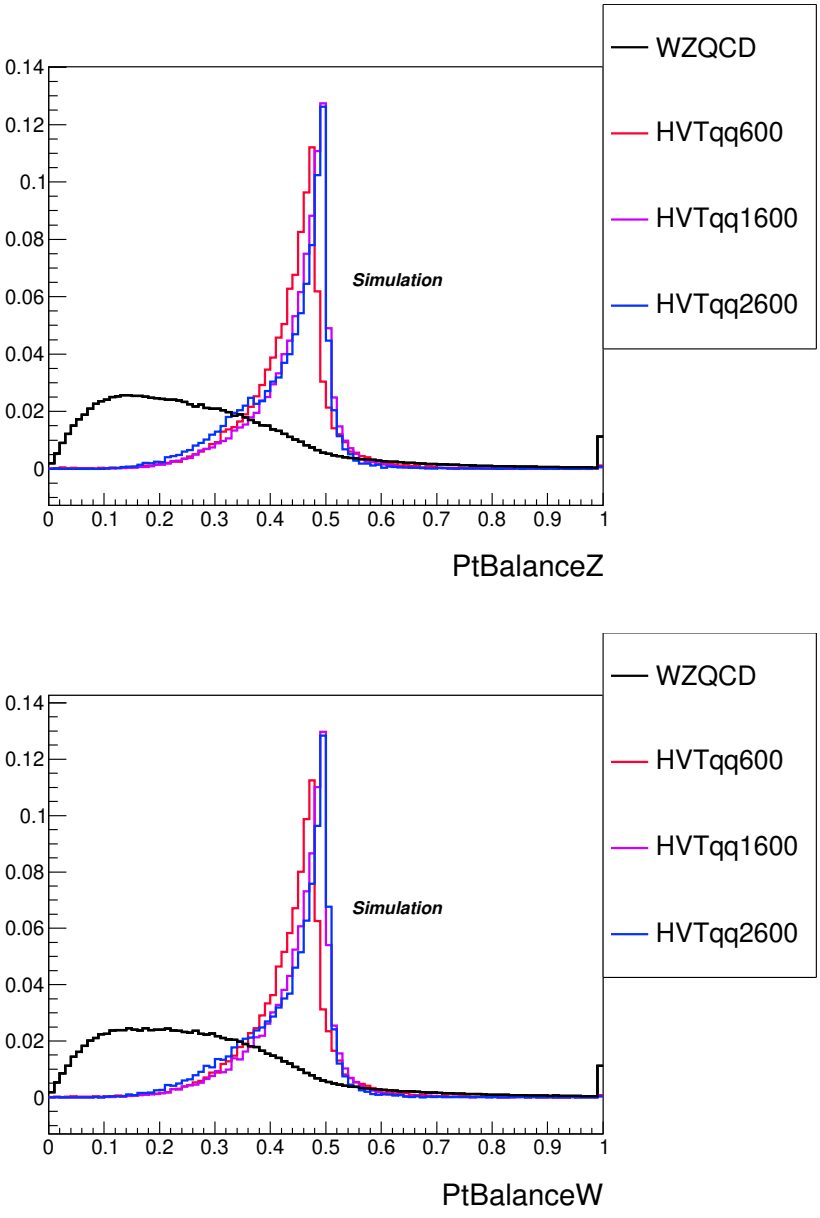


Figure 7.1: Distributions of p_T^Z/M_{WZ} (up) and p_T^W/M_{WZ} (down) for HVT signal mass points and the WZ SM background normalized to unit.

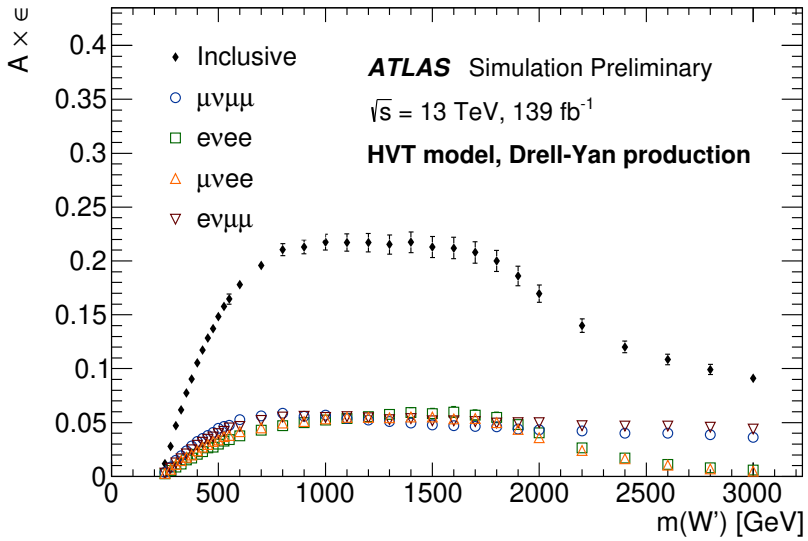


Figure 7.2: The acceptance (A) times efficiency (ϵ) of W' boson selection after the Drell-Yan signal region selection at different mass points for the individual channels $\mu\nu\mu\mu$, $e\nu ee$, $\mu\nu ee$, $e\nu\mu\mu$, and the sum of all channels. The uncertainty includes both statistical and experimental systematic components.

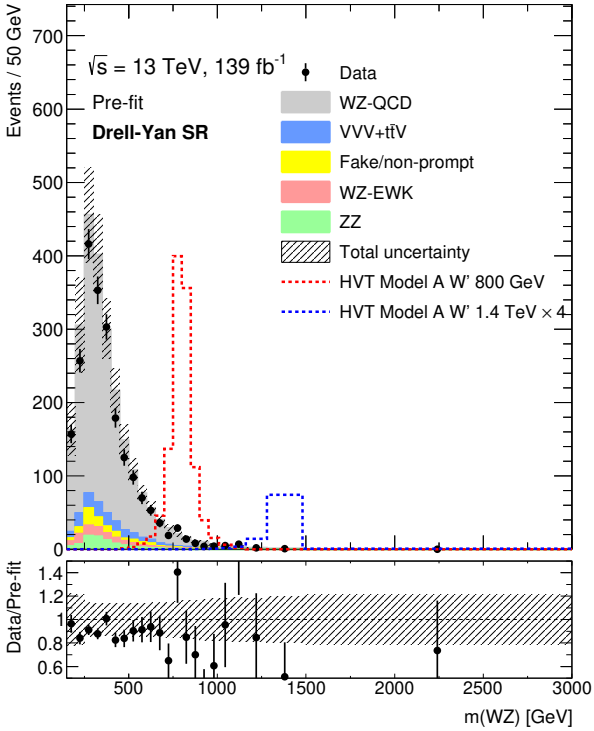


Figure 7.3: Comparisons of the data and the expected background distributions of the WZ invariant mass in the Drell-Yan signal region. The expected distribution from an HVT W' resonance with mass of 800 GeV and 1.4 TeV in the signal region are shown. The bottom panels show the ratios of the data to the pre-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

7.2 Neural Network VBF signal region

The VBF production mode is characterised by two high-energetic jets in the forward detector. The VBF production signature comprises two or more jets with large separation and high di-jet invariant mass resulting from quarks from which a vector boson has been radiated. Therefore, the typical VBF variables include the m_{jj}, $\Delta\phi_{jj}$ and other kinematic variables of the jets together with many other discriminating variables. Selection of events can be performed using Machine learning algorithms to classify events as VBF like signal and background events.

In order to select the signal events, an Artificial Neural Network (ANN) is used. The ANN performs a binary classification task and categorises events as belonging either to a VBF process or to the background. The ANN training is implemented using the `Keras` package [108] running on top of the `TensorFlow` package [109]. The ANN training region is defined by events passing the inclusive selection and requiring events to have at least a pair of jets satisfying the *VBF jets* selection with an invariant mass(m_{jj}) above 100 GeV. The kinematic variables used for training show very similar distributions for both GM and HVT benchmark signals and the training yielded similar results. The ANN is trained in this region with all mass samples of simulated GM $H_5^{\pm}$ events simultaneously as signal, against the SM WZ EWK and QCD events as background. The di-jet invariant mass and the difference in ϕ between the two leading jets for this ANN training region are shown in Fig 7.4.

The list of the input variables used for the ANN optimization is given in Table 7.1. These were chosen on the basis of their impact on the training and such that highly correlated variables are not used simultaneously. The "event centrality" is a measure of the smaller pseudorapidity difference between the most forward jet and the most forward lepton in either hemisphere. It is defined as:

$$\zeta_{\text{Lep}} = \min\{[\min(\eta_{\ell_1}, \eta_{\ell_2}, \eta_{\ell_3}) - \min(\eta_{j_1}, \eta_{j_2})], [\max(\eta_{j_1}, \eta_{j_2}) - \max(\eta_{\ell_1}, \eta_{\ell_2}, \eta_{\ell_3})]\}$$

with ℓ_1, ℓ_2, ℓ_3 being the three leptons from the WZ decay and η_{j_1}, η_{j_2} the leading-p_T and subleading-p_T VBF jets.

Table 7.1: Variables used for ANN training.

Training variable	Definition
m_{jj}	Invariant mass of the two leading-p_T jets
$\Delta\phi_{jj}$	Difference in ϕ of the two leading-p_T jets
η_W, η_Z	Pseudorapidities of the reconstructed gauge bosons
η_{j1}	Leading jet pseudorapidity
ζ_{Lep}	Event centrality
E_T^{miss}	Missing transverse energy
H_T	Scalar p_T sum of the *VBF jets* and the leptons from the WZ decay

ANN training is done using a 4-fold cross-validation technique and the rectified linear unit, or ReLU, is used as an activation function at each node. An overview of the hyperparameters of the ANN training used in this analysis is presented in Table 7.2.

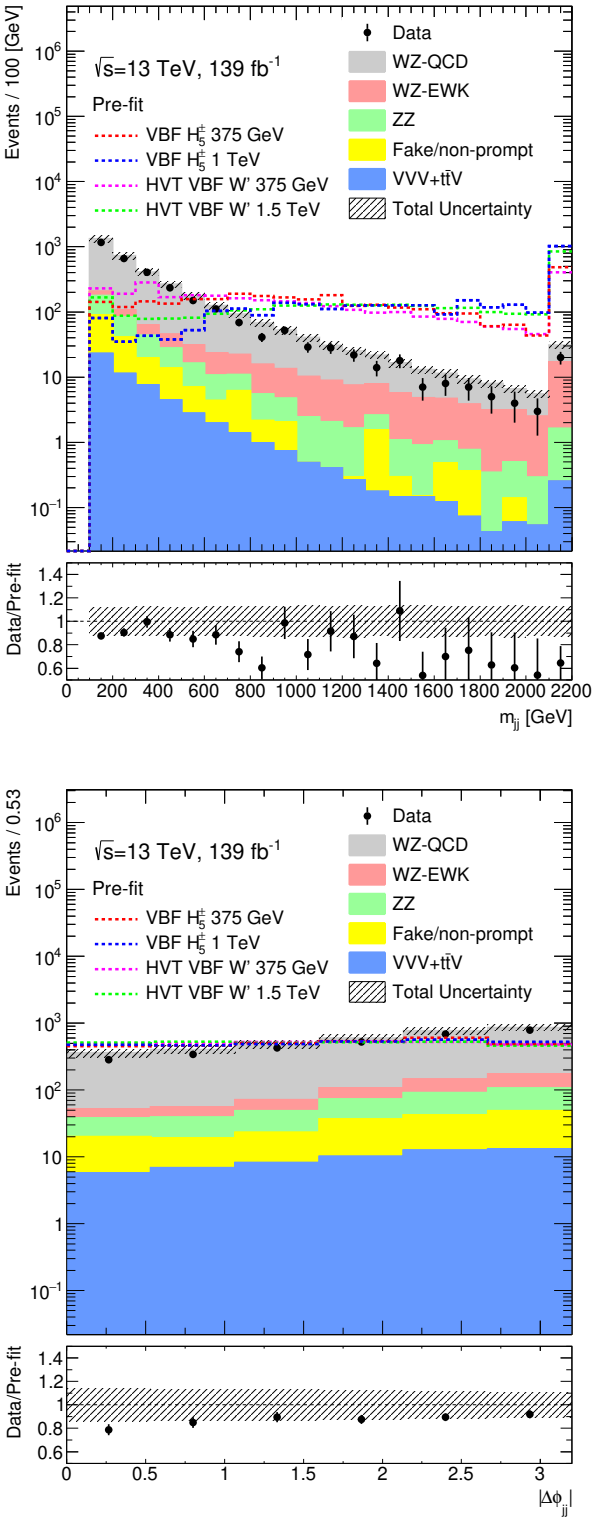


Figure 7.4: Comparisons of the observed data and the background distributions, the dijet invariant mass (top) and the difference in ϕ between the two leading jets (bottom) are shown for events in the ANN training. The signals are normalized to the total data in the training region. The pre-fit uncertainty in the total background prediction is shown as grey bands, combining only statistical and systematic contributions.

Hyperparameter	Value
Epochs	100
Number of Layers	2
Neurons per layer	45
Learning rate	0.028
Dropout	0.2
Nesterov's Momentum [110]	0.7
Folds	4

Table 7.2: Hyperparameters used for MVA selection of GM signals.

After training with all mass samples simultaneously, the cut value on the ANN output score is chosen in such a way that it maximizes the significance for the lowest mass point (200 GeV). The advantage of this approach is that it reduces considerably the training effort and a single signal region can be used. In order to find the optimal cut value on the ANN output, the significance Z, of observing n events given a prediction of b background events with variance σ is used:

$$Z = \sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} \tag{7.1}$$

This formula includes an approximation for Gaussian systematic uncertainties.

An ANN score value of more than 0.82 on the ANN output maximizes the significance and is therefore chosen to define the VBF signal region for both GM and HVT models. After all selection cuts are applied the VBF signal region effectively starts at $m_{jj} > 500$ GeV.

Good modelling by MC simulations is validated by the good description of the ANN score distribution of data in the WZ-QCD control region and VBF signal region shown in Figure 7.5.

The acceptance times efficiency $\mathcal{A} \times \epsilon$ of the ANN-based VBF selection as a function of the mass of the VBF $H_5^{\pm}$ and of the HVT W' boson, relative to the generated signal events, are shown in Figure 7.6. Decays of W and Z bosons into all flavours of leptons are included at event generation. Figure 7.7 shows the pre-fit distribution for the ANN VBF signal region.

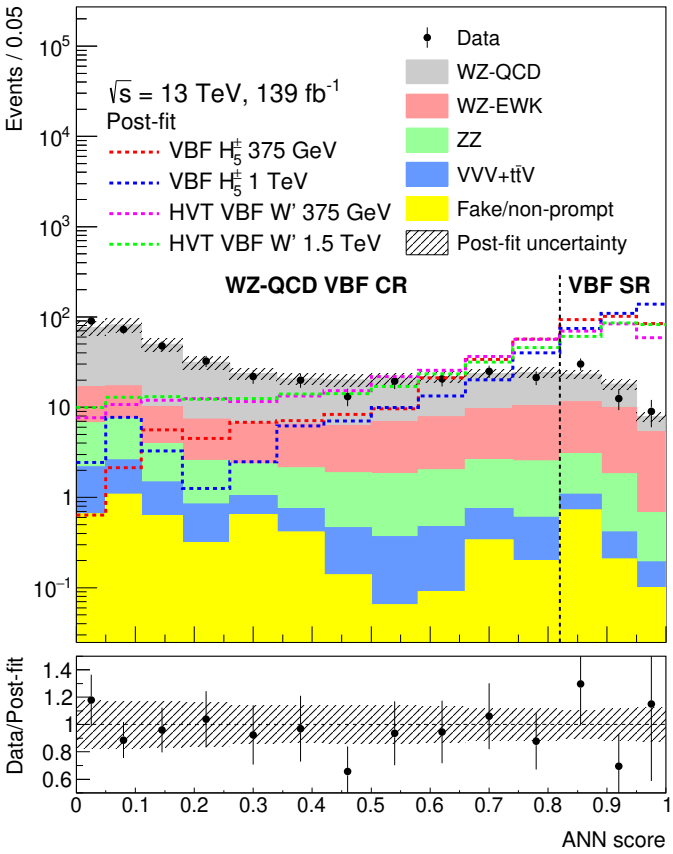


Figure 7.5: Distribution of the ANN discriminant output in the *WZ*-QCD VBF control region and the VBF signal region. The background predictions are obtained through a background-only simultaneous fit to the VBF signal region and the *WZ*-QCD VBF and *ZZ* VBF control regions as described in Section 11.2. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions. The distributions for the HVT VBF model W' and GM $H_5^{\pm}$ simulations are shown normalized to the data integral. The vertical dotted line shows the cut on the ANN output score used to define the VBF signal region.

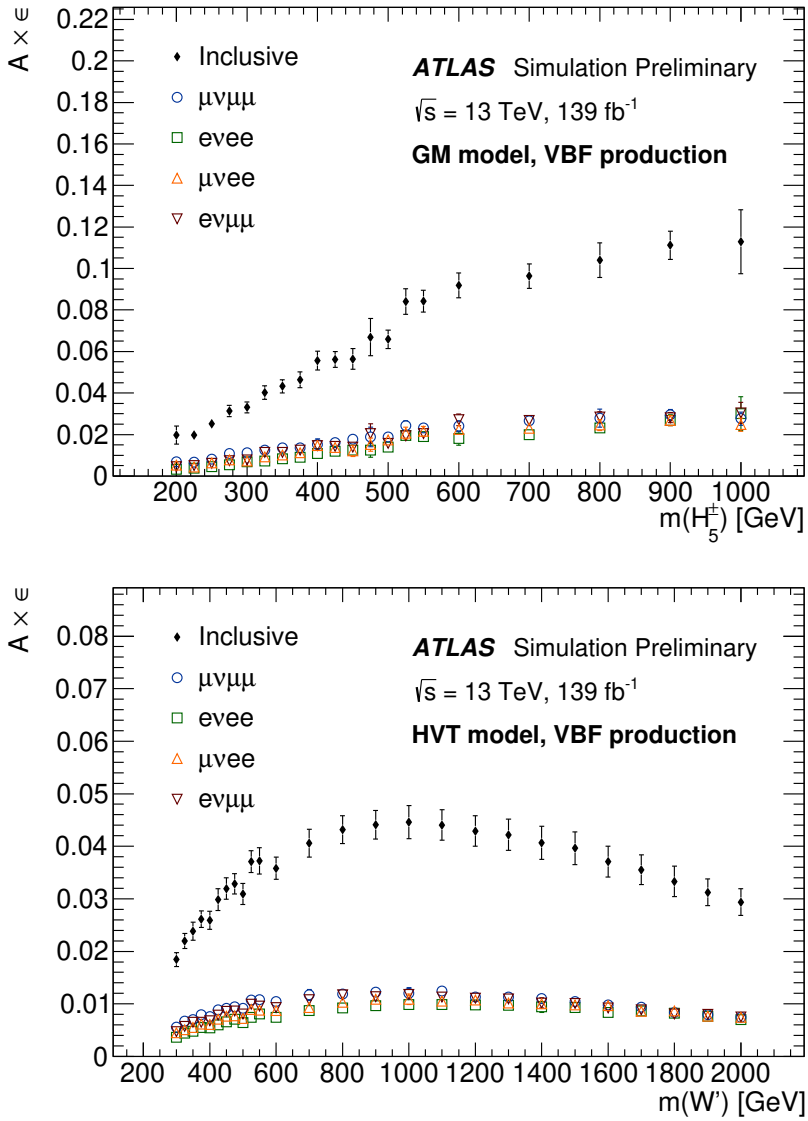


Figure 7.6: The acceptance (A) times efficiency (ϵ) of VBF $H_5^{\pm}$ (top) and HVT W' (bottom) selection after the ANN-based VBF selection at different mass points for the individual channels $\mu\nu\mu\mu$, $e\nu ee$, $\mu\nu ee$, $e\nu\mu\mu$, and the sum of all channels. The uncertainty includes both statistical and experimental systematic components.

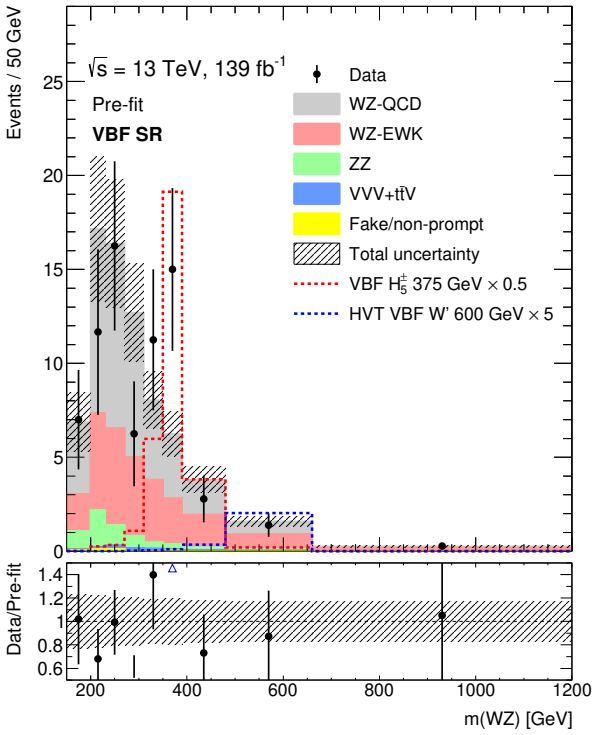


Figure 7.7: Comparisons of the data and the expected background distributions of the WZ invariant mass in the ANN VBF signal region. The expected distributions from an $H_5^{\pm}$ GM model resonance with a mass of 375 GeV and from an HVT W' (model A) of mass 600 GeV are shown in the signal region are shown. The bottom panels show the ratios of the data to the pre-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

7.3 Cut-based VBF signal region

For the vector-boson fusion analysis, a cut-based selection was also developed to facilitate reinterpretations of the results outside the ATLAS Collaboration and to serve in parallel to confirm the stability of the results obtained with the ANN.

The cut-based signal region was defined and optimized by studying the effect on significance by changing requirements on different discriminating variables based on VBF topology [111]. The events in the VBF cut based signal region are required to contain two or more jets, with $p_T > 30$ GeV and $|\eta| < 4.5$, which have significant rapidity separation $|\Delta Y_{jj}| > 3.5$, and with a di-jet invariant mass above 500 GeV.

For further rejection of the top quark pair background, $t\bar{t}+V$ and tZ events, a b-jet veto is applied on the two jets with the largest transverse momentum. Out of the possible working points corresponding to different b-tagging efficiencies, the 85% working point (threshold value of 0.1758) is selected as the baseline. This cut rejects $\sim 50\%$ of these backgrounds.

The acceptance times efficiency $\mathcal{A} \times \epsilon$ of the cut-based VBF selection as a function of the mass of the VBF $H_5^{\pm}$ and of the HVT W' boson, relative to the generated signal events, are shown in Figures 7.8 and 7.9. Decays of W and Z bosons into all flavours of leptons are included at event generation. Figure 7.7 shows the pre-fit distribution for the cut-based VBF signal region.

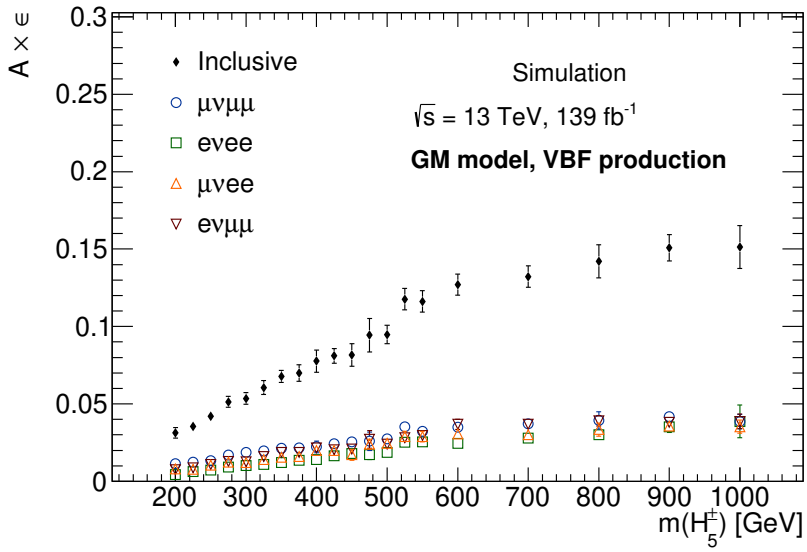


Figure 7.8: The acceptance (A) times efficiency (ϵ) of VBF $H_5^{\pm}$ selection after the cut-based VBF selection at different mass points for the individual channels $\mu\nu\mu\mu$, $e\nu ee$, $\mu\nu ee$, $e\nu\mu\mu$, and the sum of all channels. The uncertainty includes both statistical and experimental systematic components.

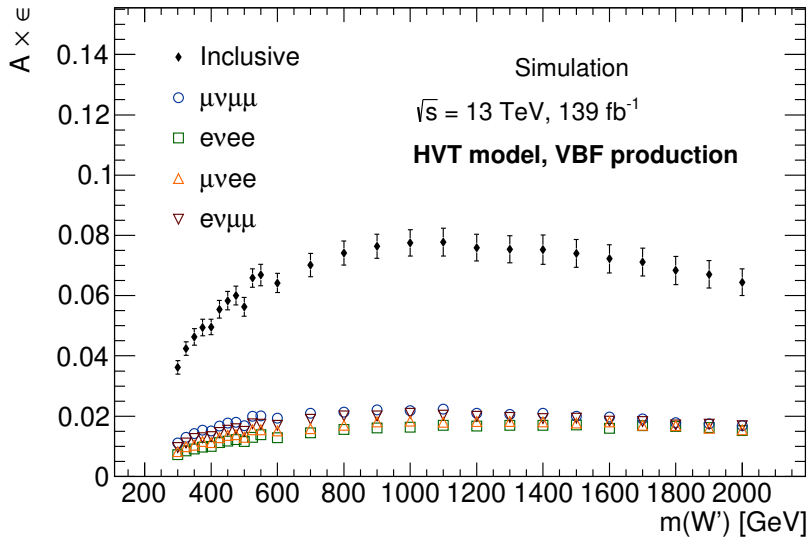


Figure 7.9: The acceptance (A) times efficiency (ϵ) of HVT W' selection after the cut-based VBF selection at different mass points for the individual channels $\mu\nu\mu\mu$, $e\nu ee$, $\mu\nu ee$, $e\nu\mu\mu$, and the sum of all channels. The uncertainty includes both statistical and experimental systematic components.

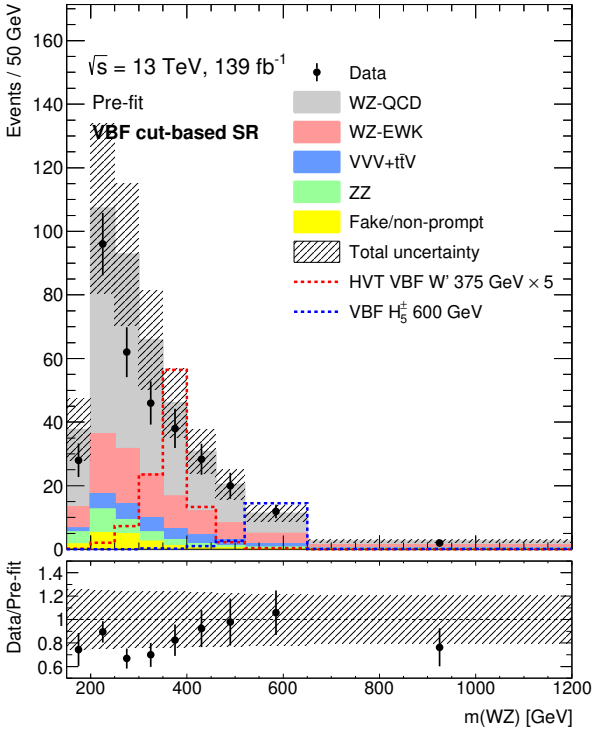


Figure 7.10: Comparisons of the data and the expected background distributions of the WZ invariant mass in the cut-based VBF signal region. The expected distributions from an $H_5^{\pm}$ GM model resonance with a mass of 600 GeV and from an HVT W' (model A) of mass 375 GeV are shown in the signal region are shown. The bottom panels show the ratios of the data to the pre-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

		Signal Regions		
Selection	ANN training	VBF ANN-based	VBF cut-based	Drell-Yan cut-based
$E_T^{miss} > 25$ GeV	✓	✓	✓	✓
4 lepton veto	✓	✓	✓	✓
inclusive WZ	✓	✓	✓	✓
at least 2 VBS jets	✓	✓	✓	
b-jet veto	✓	✓		
Mjj > 100 GeV	✓	✓		
Mjj > 500 GeV			✓	
$\|\Delta Yjj\| > 3.5$			✓	
ANN cut		score>0.82		
Boson p_T imbalance				✓

Table 7.3: Summary of the training and signal regions in the analysis

Chapter 8

Background

"No man can win every battle, but no man should fall without a struggle."

Spiderman Homecoming

In search for physics beyond the SM, background estimation is a very important step in order to tell apart the SM background from new phyics effects. In a resonance search like the one presented in this thesis, the SM background is expected to be a smooth falling distribution and the signal will appear as a peak on top. In this chapter the methods to estimtate the different sources of SM background that contribue to this search are discussed.

The background sources are classified into two groups: events where all lepton candidates are prompt leptons stemming from hard-process (irreducible background) and events where at least one of the candidates leptons is not a prompt lepton (reducible background). Candidates that are not prompt leptons are also called 'fake' leptons. The SM processes that can mimic the WZ signature mainly come from the following sources:

- Non-prompt or *reducible* background: SM processes that have one or two leptons in the final state originating from jets or photons but typically with higher cross-section:
  - $t\bar{t}$ is a substantial background process where both the *top* quarks decay leptonically with additional b-quarks which can be misidentified as leptons,
  - $Z + \gamma$ process where the photon is reconstructed as an electron,
  - other contributing background comes mainly from Z+jets, single top (s- and t-channel production of t-quarks), WW+jets or QCD multijet processes where one or two jets are mis-reconstructed as leptons. This background is due to fake leptons which include pion, kaon and heavy-quark decays to real leptons, in addition to jets misidentified as lepton.
- Prompt or *irreducible* background: SM processes that produce events with atleast three real leptons:
  - The SM WZ-EWK and WZ-QCD process which has exactly the same signature as the signal and where three leptons pass the event selection,

- ZZ process where three leptons pass the event selection, while the fourth lepton is not reconstructed or identified,
- $t\bar{t} + V$ process $V = W, Z$, as top decays through the electroweak interaction into a W boson and a b quark,
- VVV process where V denotes any vector bosons,
- $t + Z$ process,
- double parton scattering (DPS) processes where the WZ signature comes mostly from collisions between two pairs of partons producing single W and Z/γ^* bosons.

Different approaches have been followed to estimate the contribution of the different backgrounds. The contributions from the irreducible backgrounds WZ-QCD, WZ-EWK, ZZ, tribosons VVV and $t\bar{t}V$ (where $V = Z$ or W) are estimated using simulation. The shape of the WZ-QCD and ZZ are estimated from MC while their normalization has been constrained by data in dedicated control regions enriched with those backgrounds, more details in Section 8.1.

The reducible backgrounds originate from $t\bar{t}$, Z+jets, $Z\gamma$ and WW+jets production processes where jets or photons were misidentified as leptons. For both analysis regions, the reducible background normalisations are estimated using a data-driven method and referred to as "Fakes/non-prompt". The method is based on a global matrix which makes use of the differences in object characteristics between real and fake/non-prompt leptons on a statistical basis [112]. To build the matrix the leptons are classified as loose or tight candidates. Tight leptons candidates are the nominal signal leptons. Loose lepton candidates are leptons that do not meet the isolation and identification criteria of signal leptons but satisfy all the other signal leptons selection criteria. The probability for a fake/non-prompt lepton to pass the loose or tight lepton criteria is determined from data, using dedicated control samples enriched in fake/non-prompt leptons. The measured lepton misidentification probabilities are applied to the candidate events in data where at least one and up to two of the lepton candidates are loose. Then, using a matrix inversion, the number of events with at least one misidentified lepton, which represents the amount of reducible background in the selected sample, is obtained. This method is also able to provide shapes of the p_T distributions. The shape in the Drell-Yan and cut-based VBF analysis is obtained from the data-driven method. For the ANN VBF analysis, due to the limited statistics, the shapes are taken from the simulation.

8.1 Irreducible background

The normalization of the irreducible WZ-QCD and ZZ backgrounds are estimated using dedicated control regions which have large contributions from the corresponding processes. These control regions are defined to be kinematically close to the signal regions but orthogonal to them. The normalization of the WZ-QCD and ZZ is extracted by simultaneously fitting the signal and the control regions.

Three control regions (CR) are designed to have high purity, a small statistical uncertainty in terms of the background process of interest and to contain

only a small fraction of the potential signal. The three CRs defined are detailed below and summarised in Table 8.3.

8.1.1 WZ-QCD Drell-Yan control region

The WZ Drell-Yan control region is built by reversing the cut on the variables p_T^Z/m_{WZ} and p_T^W/m_{WZ} used to define the Drell-Yan signal region, i.e., requiring $p_T^Z/m_{WZ} < 0.35$ and $p_T^W/m_{WZ} < 0.35$. Events are also required to pass $p_T^V/m_{WZ} > 0.10$. The $p_T^V/m_{WZ} > 0.10$ requirement is applied to be closer to the signal region. This CR is then by definition orthogonal to the Drell-Yan signal region. Figure 8.1 shows the invariant mass distribution in the CR and the predicted event yields are summarized in Table 8.1. It can be seen that is clearly enriched in WZ processes with 81% purity and small contamination from fake background processes, ZZ and other irreducible backgrounds. This region is included in the Drell-Yan signal region fits to constrain the WZ QCD normalization.

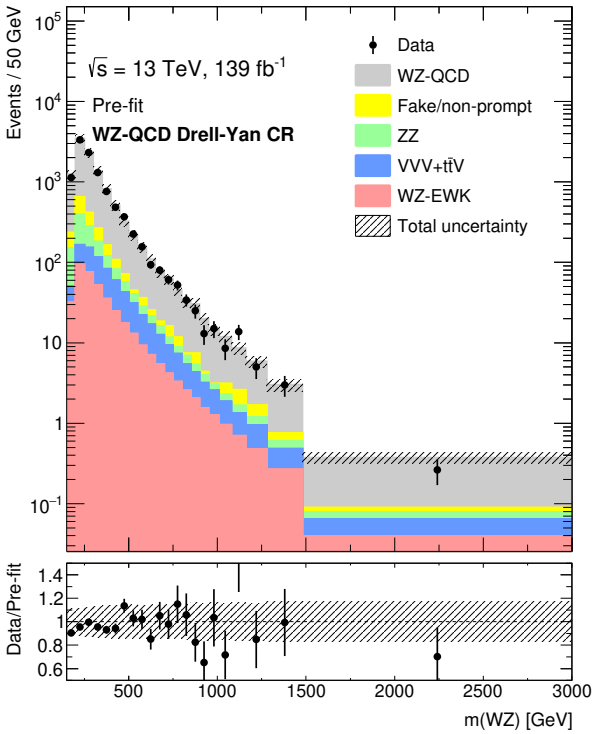


Figure 8.1: Pre-fit WZ invariant mass is shown in the WZ Drell-Yan control region. The distributions are shown here with the optimized variable binning (see Section 10.3.1). All the uncertainties are shown by shaded bands. Theory systematic uncertainties are of the order of 20% and flat over the M_{WZ} range are not shown

8.1.2 WZ-QCD VBF control regions

Two dedicated control regions are built to assess the normalization of the WZ-QCD mediated production in the VBF like phase space with two jets to ensure the orthogonality between the control regions and the two VBF signal regions, i.e cut-based and ANN. The control regions are defined as follows.

8.1.2.1 WZ-QCD VBF control regions for ANN analysis

For the VBF ANN signal regions WZ-QCD control region is defined by requiring events to have $M_{jj} > 500$ GeV to be closer in terms of the VBF topology and the requirement on the ANN score is reverted, i.e., ANN < 0.82. Figure 8.2 on the top shows the data/MC comparison and the predicted event yields are summarized in Table 8.1. The WZ-QCD mis-modelling is visible in this region and will be corrected in the limit setting fit and is fully covered by the theory systematic uncertainties that are of the order of 30% for the WZ-QCD in this region, see details in Section 10.4.2. This VBF control region is enriched by WZ-QCD mediated production (purity above 75%).

8.1.2.2 WZ-QCD VBF control region for cut-based analysis

WZ-QCD VBF control region for cut-based analysis is built by requiring events with at least 2 jets and di-jet mass $M_{jj} > 500$ GeV but failing the rapidity separation of the jets requirement, i.e. $|\Delta Y_{jj}| < 3.5$. Events passing the Drell-Yan signal region selection are also excluded by requiring $p_T^Z/m_{WZ} < 0.35$ and $p_T^W/m_{WZ} < 0.35$. The bottom plot in Figure 8.2 shows the invariant mass distribution in this CR. From the predicted event yields summarized in Table 8.1, it can be clearly seen that this cut-based CR is rich in WZ-QCD mediated production, around 72%, with small contamination from WZ-EWK, ZZ and irreducible background processes. This region is included in the VBF signal region fits to constrain the WZ QCD normalization. A mis-modelling for low WZ invariant mass can be seen in the ratio of Figure 8.2 (bottom). The modelling will be corrected in the fit and is fully covered by the theory systematic uncertainties that are of the order of 30% for the WZ-QCD in this region.

	WZ DY CR	WZ VBF ANN CR	WZ VBF CR
WZ QCD	8810 ± 329	514 ± 40	172 ± 9
WZ EWK	387 ± 14	97 ± 4	29 ± 1
$t\bar{t}$V+VVV	410 ± 83	9 ± 2	18 ± 4
ZZ	623 ± 34	32 ± 4	11 ± 1
Fakes	685 ± 218	14 ± 6	7 ± 3
Total Background	10916 ± 502	667 ± 47	237 ± 15
Data	10522	522	209

Table 8.1: Overview of observed and predicted event yields for an integrated luminosity of 139 fb^{-1} per processes in each WZ QCD control region.

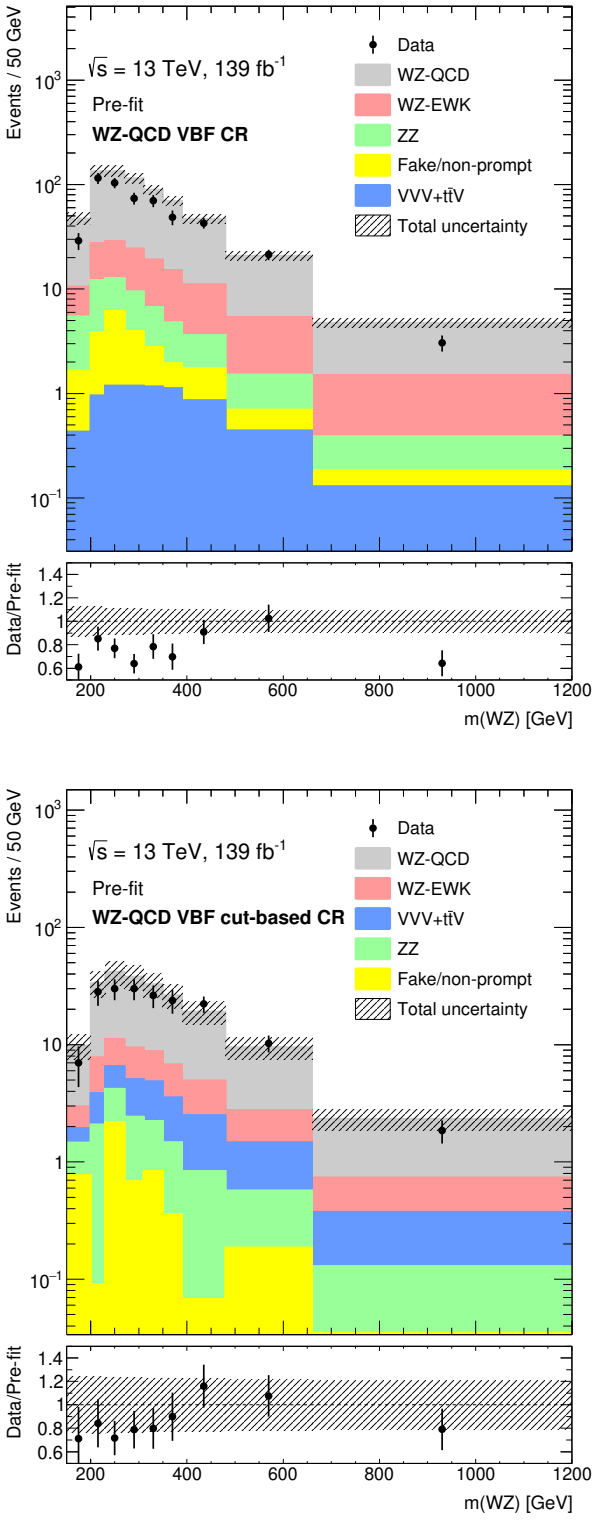


Figure 8.2: The Pre-fit distributions of the WZ invariant mass is shown in the WZ VBF ANN control region (top), cut-based control region (bottom). The distributions are shown here with the optimized variable binning for GM region (see Section 10.3.1). All the uncertainties affecting the shapes are shown by shaded bands. Theory systematic uncertainties are of the order of 30% and flat over the M_{WZ} range.

8.1.3 ZZ control regions

The ZZ background normalization is obtained from two dedicated control regions, one inclusive in terms of jets, named ZZ Drell-Yan control region (ZZ CR), and another where events are required to have at least 2 jets called ZZ VBF control region (ZZJJ CR). Both regions require events to have at least 3 analysis lepton candidates, and at least 1 extra baseline lepton (of any flavour or charge) with p_T>7 GeV and p_T>5 GeV for electrons and muons, respectively. The E_T^{miss} cut applied in the nominal analysis to account for the neutrino is removed for this control region. Since the 4-lepton veto selection is reverted this region is by definition orthogonal to all the signal regions of the analysis.

Figure 8.3 shows the WZ invariant mass distribution in the ZZ CR on the left and the ZZJJ CR on the right. For 8.2 it can be seen that these CR are clearly enriched in ZZ events (93% for ZZ CR and 81% for ZZJJ CR). The ZZ CR region is included in the Drell-Yan signal region fits to constrain the ZZ background normalization, and the ZZJJ CR is included in the VBF signal region fits, for both cut-based and ANN signal regions. A theory uncertainty of 10% in the normalization of the ZZ background includes the PDF and scale effects.

	ZZ CR	$ZZJJ$ CR
WZ QCD	35 ± 2	13 ± 1
WZ EWK	6 ± 0.3	3 ± 0.2
$t\bar{t}$V+VVV	65 ± 13	31 ± 7
ZZ	1305 ± 48	244 ± 38
Fakes	23 ± 9	10 ± 5
Total Background	1434 ± 61	301 ± 43
Data	1554	305

Table 8.2: Overview of observed and predicted event yields for an integrated luminosity of 139 fb^{-1} per processes in ZZ and $ZZJJ$ control region.

	Control Regions				
Selection	ZZ	$ZZjj$	WZ qqF cut-based	$WZ-QCD$ VBF cut-based	$WZ-QCD$ VBF ANN-based
$E_T^{miss} > 25$ GeV			✓	✓	✓
4 lepton veto	fail	fail	✓	✓	✓
inclusive WZ			✓	✓	✓
at least 2 jets		✓	✓	✓	✓
b-jet veto				✓on VBF jets	✓
Mjj > 100 GeV				✓	✓
Mjj > 500 GeV				✓	✓
$\|\Delta Yjj\| > 3.5$				fail	
ANN cut					fail
Boson p_T imbalance			fail	fail	

Table 8.3: Summary of the control regions in the analysis

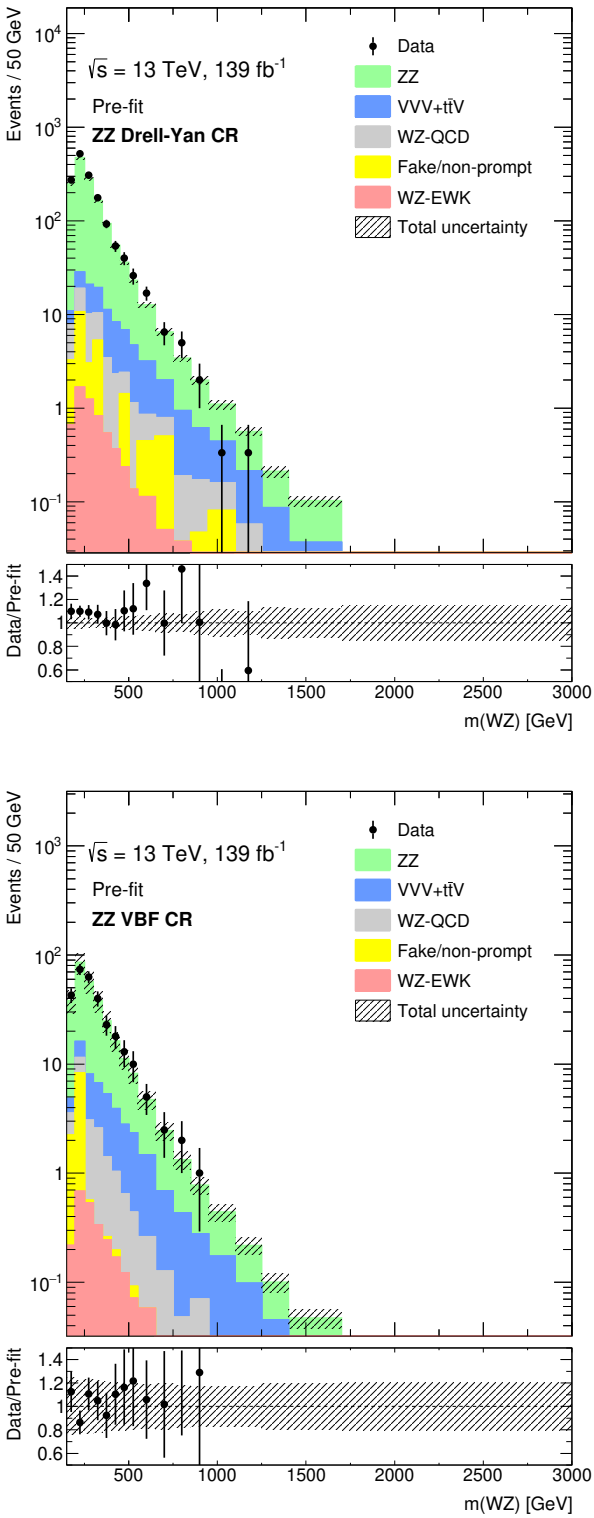


Figure 8.3: The pre-fit distributions of the WZ invariant mass is shown in the inclusive ZZ control region at the top and in the $ZZJJ$ VBS control region at the bottom. The figures show all the uncertainties by shaded bands. A flat 10% theory uncertainty is expected on the ZZ background normalization and is not included in the figure.

8.1.4 $t\bar{t}V$ Validation region

The $t\bar{t}V$ background contribution is the third largest background contribution in all signal regions. Therefore a validation region was constructed to study the $t\bar{t}V$ modelling and normalization.

The $t\bar{t}V$ region is defined by requiring to have at least 2 jets with $M_{jj} > 150$ GeV and at least one b-jet. Since no significant shape miss-modelling was found and in order to keep the fit simple it was decided to not to include this region in the fit and use the theory uncertainties to cover for the observed differences between data and simulation.

Table 8.4 and Figure 8.4 shows the observed and expected number of events in this region, the $t\bar{t}V$ background accounts for ∼40% of the events.

	ttV VR
$t\bar{t}$V	137 ± 30
WZ QCD	123 ± 27
WZ EWK	43 ± 8
ZZ	12 ± 33
Fakes/non-prompt	35 ± 25
Total Background	349 ± 60
Observed	391

Table 8.4: Observed and expected events in the $t\bar{t}V$ validation region

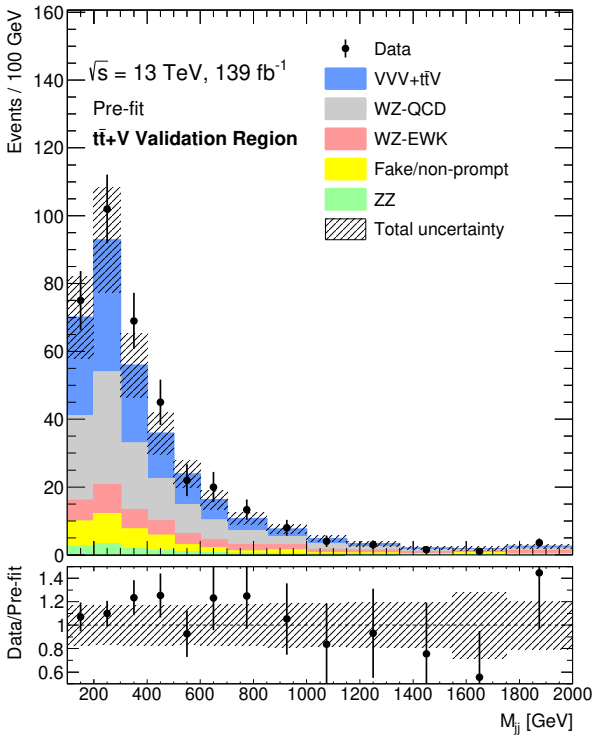


Figure 8.4: Data/MC comparisons in the $t\bar{t}V$ validation region. The error band contains statistical and experimental uncertainties. The theory uncertainty of 13% is added to the band.

Chapter 9

Systematic Uncertainties

"We're the Guardians of the Galaxy."
Guardians of the Galaxy Vol.1

Systematic uncertainties are an important ingredient for quantative comparison of simulated and measured data. The theoretical modelling, object and event reconstruction have an impact on the signal and control regions because they affect the background estimations, signal acceptance and the shape of the distributions of the invariant mass. Depending on the nature of the uncertainty these can be classified into two groups: (a) theoretical uncertainties associated with the MC modelling of both the background and signal processes and (b) experimental uncertainties related to the detector and reconstruction performance.

9.1 Theory uncertainties

Theoretical uncertainties are used to evaluate the WZ-QCD, WZ-EWK and ZZ mis-modelling by the event generators. Studies were performed to evaluate these uncertainties but are not a part of this thesis. For each of these backgrounds, uncertainties due to higher-order QCD corrections are evaluated by varying the renormalization and factorization scales independently by factors of two and one-half. For the WZ-QCD and ZZ backgrounds that have data-driven normalizations only the shape variations on the reconstructed $m(WZ)$ are considered. The uncertainties due to the PDF and the α_s value used in the PDF determination are evaluated using the PDF4LHC prescription [113] for WZ-QCD, WZ-EWK and ZZ background. A parton shower modelling uncertainty in each of the WZ-QCD and WZ-EWK background templates are estimated separately using two MC samples with different shower models.

The theory uncertainties summarised above are evaluated in all the analysis signal and control regions and they are considered as uncorrelated across those regions in the statistical analysis.

An uncertainty of 20% is assigned to the $t\bar{t}V$ and VVV cross-sections [114–116]. It consists of contributions from PDF uncertainties and QCD scale uncertainties.

Uncertainties in the signal acceptances due to PDF and scale choices are also evaluated. These uncertainties are calculated for several resonant mass points,

and for each model, production process and decay. The theory uncertainties on the HVT signals are evaluated to be under 20% for all production modes and they are 30% for the GM model.

9.2 Experimental uncertainties

Experimental uncertainties arise from the luminosity calculation, triggers, and reconstruction and identification of leptons and jets, as well as the computation of the missing transverse energy. For convenience, the individual uncertainties can be grouped and described as follows:

Luminosity The uncertainty in the integrated luminosity is 1.7%. It is derived following a methodology similar to that detailed in Ref. [117]

PRW_DATASF This includes the uncertainty in the pileup re-weighting of MC events is included to cover the uncertainty in the ratio of the predicted and measured inelastic cross-sections [118].

JET Systematic uncertainties affecting the reconstruction and energy calibration of jets are propagated through the analysis. They are the dominant experimental uncertainties in the VBF selection. These include seven jet energy scale parameters (`JET_EffectiveNP`), six eta-intercalibration (`JET_EtaIntercalibration`) (one each for modelling and statistical uncertainties, and four for non-closure), four for pile-up dependence (`JET_Pileup`), and one each for high-momentum jets (`JET_SingleParticle_HighPt`), mismodelling of flavour composition (`JET_Flavor_Composition`) and response (`JET_Flavor_Response`), treatment of b-jets (`JET_BJES_Response`) and punch-through jets (`JET_PunchThrough_MC16`). These also include mis-modelling of efficiency of jet vertex tagging (`JET_JvtEfficiency` and `JET_fJvtEfficiency`)

JET_JER Those due to the jet energy scale and resolution are obtained from simulations and *in-situ* techniques [119].

FT_EFF These include the uncertainties in the b-tagging efficiency and the mistag rate is also taken into account. The effect of jet uncertainties on the expected number of events ranges up to 15% in the VBF selection.

MUON These are the uncertainties due to the muons reconstruction, identification and isolation requirements estimated using tag-and-probe methods in $Z \to \ell\ell$ events in data and simulation [70, 120]. Uncertainties in the efficiencies of muon triggers, muon energy scale and resolution are also included.

EG These are the uncertainties due to electron calibration originating from the mismodelling of the electron energy resolution (`EG_RESOLUTION_ALL`) and enery scale (`EG_SCALE`).

EL_EFF They include the systematic variations due to the effects of electron efficiency measurements in imprecise measurement of identification (`EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR`), isolation (`EL_EFF_Iso_TOTAL_1NPCOR`

_PLUS_UNCOR) and reconstruction (EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR).

MET The uncertainty in the missing transverse momentum measurement is estimated by propagating the uncertainties in the transverse momenta of preselected leptons and jets as well as those in the soft term [79].

FF An uncertainty in the prediction of the fake/non-prompt background is also taken into account as it affects the shape and normalization of the background distributions. The total uncertainty is about 60% (more than 100%) for the Drell-Yan (VBF) selections. It is larger for the VBF selection because of the higher statistical uncertainty.

Chapter 10

Statistical Analysis

"We're in the endgame now."

Avengers: Infinity War

The statistical treatment of this analysis uses a combined profile likelihood fit to binned discriminant distributions in all categories simultaneously. The definition of the likelihood and the configuration of the fit in terms of inputs and parameters are defined in this section.

10.1 Statistical Methodology

To derive the equations for statistical analysis of the data we start from the so-called "marked Poisson model" [121], as shown here:

$$\mathcal{P}(x_1...x_n|\mu) = \text{Pois}(n|\mu S + B) \left[\prod_{e=1}^{n} \frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B} \right] \tag{10.1}$$

where S is the number of signal events and B is the number of background events. A "signal strength" parameter μ is introduced such that $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ corresponds to the signal+background hypothesis. This continuous parameter μ is our parameter of interest. The background and signal "shapes" are denoted by $f_S(x)$ and $f_B(x)$ and note that these are probability density functions (pdf) normalized such that $\int dx f(x) = 1$. This probability model is for obtaining n events in the data where the discriminating variable for event e has a value x_e. We have considered the probability density of obtaining x_e based on the relative mixture $f_S(x)$ and $f_B(x)$ for a given value of μ.

Since we use histograms, we use the binned equivalent of the probability model above and denote the signal and background histograms as ν_b^{sig} and ν_b^{bkg}, where b is the bin index. The bin ν_b and the shape $f(x)$ are related as:

$$f_S(x_e) = \frac{\nu_{b_e}^{\text{sig}}}{S\Delta_{b_e}} \qquad \text{and} \qquad f_B(x_e) = \frac{\nu_{b_e}^{\text{bkg}}}{B\Delta_{b_e}} \tag{10.2}$$

where b_e is the index of the bin containing x_e and Δ_{b_e} is the width of that same bins.

The Likelihood is introduced as the conditional probability for the measurement of a fixed set of data given a theory hypothesis, the likelihood is often written as $\mathcal{L}$(theory). The theory for this analysis comprises signal and background models including μ. The Likelihood can be expressed as:

$$\mathcal{L}(\mu) = \prod_{i=1}^{n} f(x_i; \mu) \tag{10.3}$$

where $f(x_i; \mu)$ is pdf of measurement of n observables x_i given the theory with signal strength μ. The likelihood is evaluated event-by-event for a physical observable such as the invariant mass of the WZ.

10.1.1 Likelihood definition

The framework used for the construction of the Likelihood is called `HistFactory` [121] which employs `RooFit` [122] implementations of fundamental statistical tools and simplifies the construction of complex likelihoods for binned histograms in multiple channels and regions. For the statistical analysis presented here, we use the binned likelihood function which is the product of Poisson probabilities for each bin of the histogram, with θ as additional nuisance parameters (NP) of the theory, times the constraint term of auxiliary measurements:

$$\mathcal{L}(\mu, \theta) = \mathcal{P}(n_{cb}, a_p | \phi_p, \alpha_p, \gamma_b) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb} | \nu_{cb}) \cdot G(L_0 | \lambda, \Delta_L) \cdot \prod_{p \in \mathbb{S}+\Gamma} f_p(a_p | \alpha_p) \tag{10.4}$$

where, following mnemonic index conventions are used:

- $b \in$ bins
- $c \in$ channels
- $p \in$ parameters

$\mathbb{S} = \{\alpha_p\}$ is the set of systematic nuisance parameters, i.e. `ShapeSyst`, Γ is the set of bin-by-bin statistical uncertainties and $f_p(a_p|\alpha_p)$ is a constraint for the nuisance parameter α_p with auxiliary measurement a_p. λ is the nuisance parameter of the true and unknown value of the luminosity, L_0 is the fixed auxiliary measurement of the luminosity value in the"conditional ensemble" and $G(L_0|\lambda, \Delta_L)$ is the constraint on the luminosity.

10.1.2 Profile likelihood ratio method

To test a hypothesized value of μ we consider the profile likelihood ratio [123]:

$$\lambda(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0, \\ \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \end{cases} \tag{10.5}$$

where the numerator is the likelihood function. The quantity $\hat{\hat{\theta}}(\mu)$ and $\hat{\hat{\theta}}(0)$ denotes the value of θ that maximizes $\mathcal{L}$ for the specified μ, i.e., it is the

conditional maximum-likelihood (ML) estimator of θ given strength parameter of μ or 0, respectively. The denominator is the maximised likelihood function, i.e., $\hat{\theta}$ and $\hat{\mu}$ are their ML estimators. A test statistic is defined as

$$q_\mu = \begin{cases} -2ln(\lambda) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2ln\frac{\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathcal{L}(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0, \\ -2ln\frac{\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu \end{cases} \tag{10.6}$$

We can calculate the significance using the asymptotic formula [123]:

$$Z_\mu = \sqrt{q_\mu} \tag{10.7}$$

This test statistic is used to measure the compatibility of the signal+background model with the observed data and for exclusion intervals derived with the CLs [124] method.

The CLs method is motivated by frequentist analysis. The test statistic q is defined such that it increases monotonically for increasing signal-like experiments. The confidence in the signal+background hypothesis is given by the probability that the test statistic is less than or equal to the value observed in the experiment, q_{obs}:

$$CL_{s+b} = P_{s+b}(q \leq q_{\text{obs}}) \tag{10.8}$$

where

$$P_{s+b}(q \leq q_{\text{obs}}) = \int_{-\text{inf}}^{q_{\text{obs}}} \frac{dP_{s+b}}{dq} dq \tag{10.9}$$

and where dP_{s+b} is the pdf of the test-statistic for signal+background model q_μ. Small CL_{s+b} value means high compatibility with background hypothesis. Similarly, the confidence in background only hypothesis is given by:

$$CL_b = P_b(q \leq q_{\text{obs}}) \tag{10.10}$$

where,

$$P_b(q \leq q_{\text{obs}}) = \int_{-\text{inf}}^{q_{\text{obs}}} \frac{dP_b}{dq} dq \tag{10.11}$$

and where dP_b is the pdf of the test-statistic for background-only experiments q_0. The value of CL_b close to 1 means poor compatibility with background-only hypothesis.

The confidence in the signal hypothesis CLs is defined by normalizing the confidence level observed for the signal+background hypothesis, to the confidence level observed for the background-only hypothesis, i.e., $CLs = CL_{s+b}/CL_b$. The signal hypothesis is excluded at the confidence level CL when

$$1 - CLs \leq CL \tag{10.12}$$

The limit set on μ is then translated into a limit on the signal cross-section times branching ratio, $\sigma \times \text{B}$, using the theoretical cross-section and branching ratio for the given signal model.

This simplified asymptotic formula is only valid for a large number of expected events where the distribution of the q_μ is approximated analytically. It

is therefore needed to check if this approximation is within the regime where there are limited events. Toys are used to model the q_μ distribution in a more realistic way and used to validate the limits obtained from the asymptotic formula.

10.2 Limits with Toys

The recommended procedure needed for computing frequentist exclusion limits based on profile likelihood ratio tests are based on toy Monte Carlos which is used to validate the asymptotic formulas. The result of the significance test is a p-value, p_μ. The value of signal strength μ is excluded at 95% confidence level if one finds $p_\mu < 0.05$. strength.

To find the observed upper limit [125]

1. The likelihood function is constructed as in eq. 10.4, $\mathcal{L}(\mu, \theta)$ where μ is the signal strength and θ represents the nuisance parameters.
2. The test statistic q_μ (eq. 10.6) is constructed based on $\lambda(\mu)$ likelihood ratio (eq. 10.5) and the observed test statistic for the tested μ, $q_{\mu,obs}$ is found.
3. Toy Monte Carlo experiments are generated to construct the pdf of q_μ under signal with strength μ, $f(q_\mu|\mu, \hat{\hat{\theta}}(\mu, obs))$, where the $\hat{\hat{\theta}}(\mu, obs)$ is the conditional MLE based on the observed data. The nuisance parameters are fixed to their conditional MLEs for generating the toy Monte Carlo but are constrained by a Gaussian function (unconditional ensemble).
4. From the constructed distribution of q_μ for the signal+background, $f(q_\mu|\mu, \hat{\hat{\theta}}(\mu, obs))$, the p-value of the observation is calculated:
$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu, \hat{\hat{\theta}}(\mu, obs))dq_\mu \tag{10.13}$$
5. By throwing multiple toys the μ_{up}(obs) which satisfies $p_{\mu_{up}} = 5\%$ is found.

To find the median expected upper limit and bands we need the distribution of upper limits from background-only experiments. This requires repeating the above procedure several times for background-only pseudo-experiments [125].

1. Background-only toy MC experiments are generated according to the distribution based on $(\mu = 0, \hat{\hat{\theta}}(0, obs))$. Here $\hat{\hat{\theta}}(0, obs)$ is the conditional MLE based on the observed data. The toy background-only experiments make each have a corresponding pseudo-data set, denoted by $data'$.
2. For each of the background-only pseudo experiments $data'$, the above procedure replacing obs with $data'$ is repeated. The corresponding $\hat{\hat{\theta}}(0, data')$ is calculated and the above procedure is followed to find $\mu_{up}(data')$, histograming of which leads to the distribution of upper-limits from the background-only experiments, $f(\mu_{up}|\mu = 0, \hat{\hat{\theta}}(0, obs))$.
3. The $\mu_{up}(data')$ distribution is drawn for background-only experiments to find its median.

4. To find the ±1 and ±2 sigma bands, the above generated μ_{up} pdf is used to find the 68% and 95% bands respectively.

Figure 10.1 shows the distribution of the upper limits on the cross-section times branching ratio obtained after the fit of 5000 toy datasets. The toy data in the figure was fitted using the 1.2 TeV HVT W' signal produced via VBF as a signal hypothesis. The lines in the figure show the observed and expected limits as well as the 1 and 2 sigma bands.

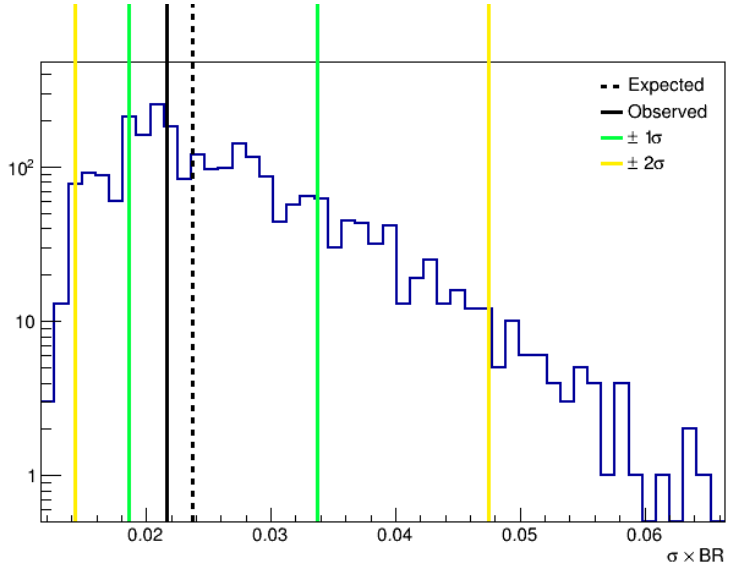


Figure 10.1: The distribution of the upper limits on the cross-section times branching ratio obtained after the fit of 5000 toy datasets for 1.2 TeV HVT W' signal produced via VBF. The black solid line shows the observed limits and the black dashed shows the median for expected limits. The green and yellow lines show the ±1 and ±2 sigma bands, respectively.

10.3 Statistical Analysis strategy

The WZ invariant mass distribution is used as the discriminating variable which is built from the four-vector of the reconstructed W and Z bosons. The longitudinal momentum of the neutrino, p_z, is estimated by solving the equation $E^2 = P^2 + M^2$, using the x and y components of the missing transverse energy, E_T^{miss}, and the four-vector of the charged lepton assigned to the W boson. On solving the quadratic equation, if there are two real solutions, the one with the smaller magnitude is chosen and if there is no real solution, the real part of the solution is chosen [51]. Using this reconstructed neutrino four-vector, the four-vector of the W boson is reconstructed.

A binned maximum-likelihood fit is performed using the reconstructed WZ invariant mass spectrum. Histogram templates of the signal and backgrounds are fitted using the standard package RooStats [126] called within the `ResonanceFinder` package. The number of fake-lepton events in each region is obtained with the data-driven method. The number of events for the $t\bar{t}+V$ and VVV is predicted by the simulation in each region. For simplicity all the decay channels

(electron and muon) are merged together in a single distribution. Additionally to the signal region, the control regions for the WZ QCD mediated productions and the ZZ backgrounds are included.

The numbers of observed and predicted events in each of these regions are described using Poisson probability density functions. There are two extra free parameters considered per signal region: a normalisation scale for the WZ-QCD, and ZZ background. The other backgrounds from $t\bar{t} + V$, and VVV are allowed to vary in the fit within their respective uncertainties of 20%. The experimental systematic uncertainties (Section 9.2) and the MC statistical uncertainties on the expected values are included in the fit as nuisance parameters which are constrained by a Gaussian function with a width corresponding to the size of the uncertainty considered and a Poissonian function, respectively. Correlations between these parameters are also taken into account.

In addition to the experimental uncertainties, theoretical uncertainties (Section 9.1) calculated directly from MC following PMG recommendations are used. PDF and Scale uncertainties are not correlated among processes. The fake background is allowed to vary within the uncertainties estimated by the Matrix Method. Systematic uncertainties inducing changes of less than 1% in the nominal yields are not considered (pruned away).

The product of the various probability density functions forms the likelihood, which the fit maximises by adjusting the input parameters and the nuisance parameters described above. The fit may introduce a negative correlation between the four normalisation scales. The relative uncertainty on the individual contributions may therefore increase, but the sum of the contributions is estimated more precisely: the total background relative uncertainty may then be smaller than the sum in quadrature of the individual components.

Two fit configurations are used, referred to as the Drell-Yan and VBF configurations. The Drell-Yan fits include the Drell-Yan SR, *WZ*-QCD CR and *ZZ*-CR. In the VBF configuration, fits include the VBF-SR, *WZjj*-QCD CR and *ZZjj*-CR. Separate fits are performed for the different models tested and for different resonance mass hypotheses.

The inputs to the fit, for each of the fit configuration, are as follows:

1. The WZ invariant mass distribution in the signal region,
2. The WZ invariant mass distribution in the WZ-QCD and ZZ control regions (each SR has its respective WZ-QCD and ZZ CR),
3. Free-floating normalisation factor for WZ-QCD and ZZ background,
4. Systematic and statistical uncertainties (Chapter 9).

10.3.1 Binning optimization

The question of what binning to choose for the WZ invariant mass distribution to calculate the bin-by-bin significance has no universal answer and to some extend needs to be investigated empirically.

The first step was to check the shape of our signal and for that the signal MC is fitted with Breit-Wigner convoluted with a Crystal-Ball function in order to estimate the width of the resonance; the width of the peak is plotted

as a function of the resonance mass see Figure 10.2. To be able to exploit the signal shape information we will need to choose as many bins as possible. Nevertheless, some constraints on the number of bins we can use come from the statistics available in the signal and the underlying uncertainties in the expected SM backgrounds. Especially when it comes to statistical analysis, a stable fit template is important to allow for a reasonable uncertainty estimate and to mitigate the risk of strong fluctuations in the background template, that could cause numerical problems in the maximum likelihood estimation.

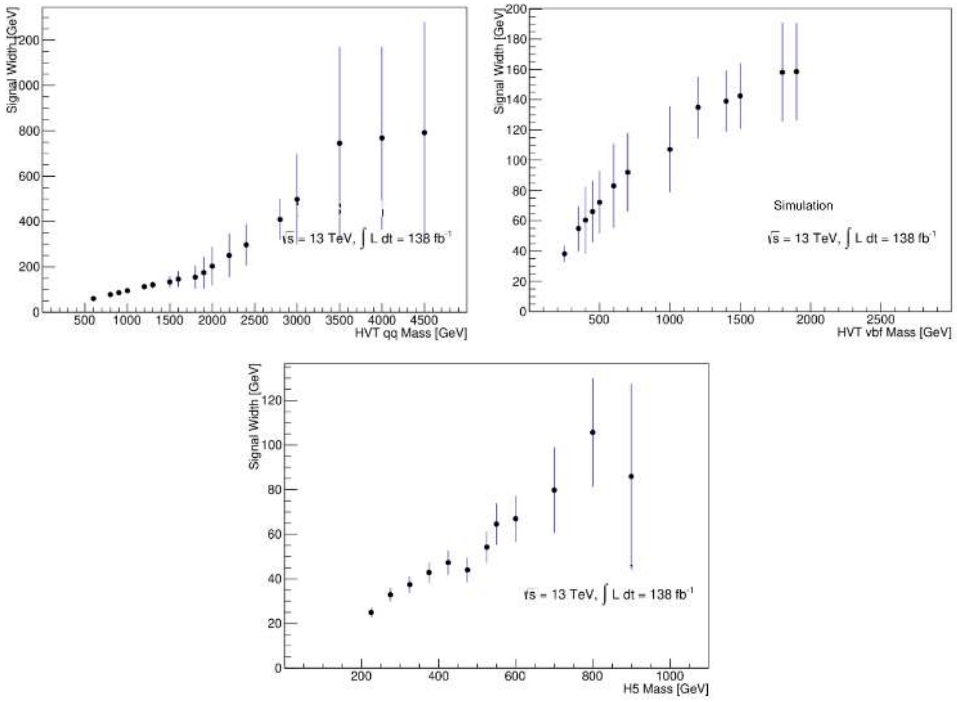


Figure 10.2: Width of the resonance peak as a function of the resonance mass in the signal Drell-Yan and VBF signals.

To study the sensitivity of our fit to the different binning options we have started with a distribution of the WZ invariant mass containing 500 bins of 10 GeV and use this as an input to the binning optimization algorithm. The algorithm for binning optimization we have used takes the input histogram and merge adjacent bin until the following criteria are satisfied:

- when enough statistics are available the bin size respect the expected signal resolution
- minimum background per bin (default = 10, 5 for high mass region): For the asymptotic approximation to work.
- maximum relative background MC uncertainty (default = 0.3): To avoid bins with large MC uncertainties.

The expected limits obtained with different binning options are then compared. A signal+background fit is performed using the signal region only and

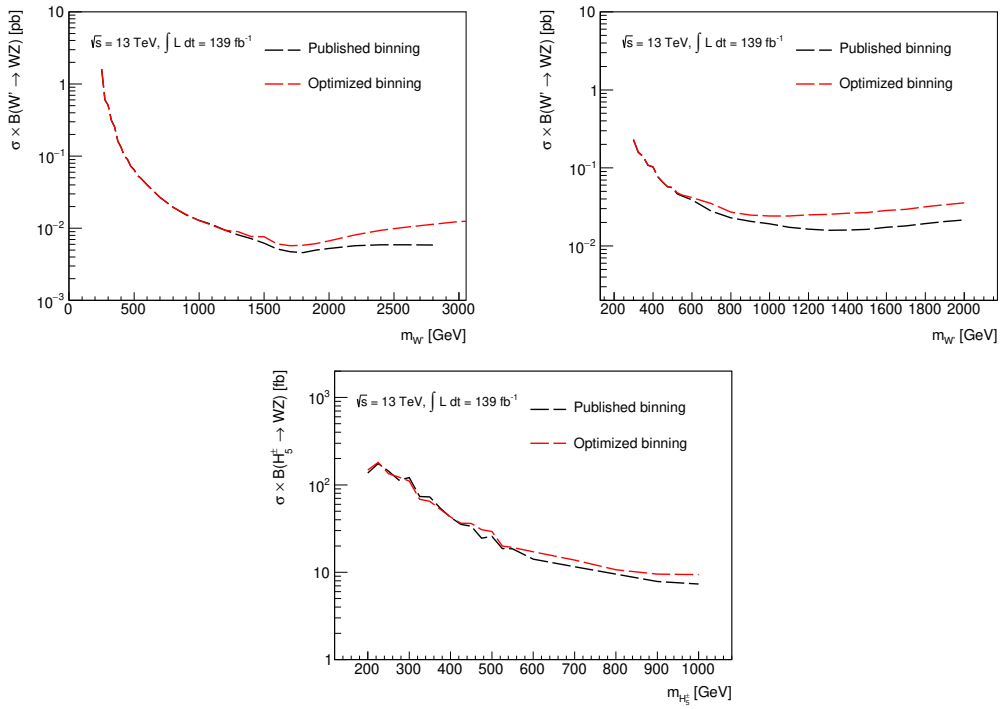


Figure 10.3: Expected limits for different binning of the WZ invariant mass distribution. The plot on the top left shows the DY HVT signal region, on the top right the HVT VBF ANN signal region and the GM VBF ANN signal region at the bottom.

including all object uncertainties and theory uncertainties to find the exclusion limits.

Figure 10.3 shows the expected limits calculated using the asymptotic approximation for the different binning proposals. The so-called optimized binning is chosen as it respects the resonant width and the statistics available. The published binning, used in previous publication [8], do not respect the minimum statistics requirements and in many bins at high mass has less than 1 event expected so the asymptotic limits can not be fully trusted in those bins. The difference in expected limits between the two binning options is not large.

This means two different binning for the VBF signal region (one each for HVT and GM) and another one for the Drell-Yan signal region are defined. The following optimized binning is used for this analysis:

- GM VBF signal region (9 bins): [150,200,230,270,310,350, 390, 480,660,5000]
- HVT VBF signal region (9 bins): [150,200,250,300,350,400,460, 520,650,5000]
- HVT Drell-Yan signal region (22 bins):[150,200,250,300,350,400,450, 500,550,600,650,700,750,800,850,900,950,1010,1080,1160,1280, 1480,5000]

10.4 Irreducible background Control Region only fit

In order to study the Data/MC agreement in the control regions of the analysis, background only fit configurations are performed using data and background predictions from the control regions as defined in Section 8.1 for each of the fit configuration.

10.4.1 Drell-Yan Control region fit

The WZ Drell-Yan control region and the ZZ control region are fitted simultaneously. The WZ-QCD and the ZZ background normalisation are free-floating parameters, while all the other backgrounds are allowed to move within the uncertainties. The nuisance parameter pull of the fit is shown in Figure 10.4, a pull and constrain are visible on the theory scale and parton shower uncertainties of the WZ-QCD background and a smaller pull on the pileup re-weighting uncertainty as well as in the systematic uncertainty introduced by the statistic uncertainty of the control regions used for Fake rate measurement. The post-fit distributions of the ZZ and WZ Drell-Yan control regions are shown in Figure 10.5, the ZZ normalization shows the need for a flat normalization factor of 10%, while the WZ-QCD normalization shows a normalization correction of about 15% for all masses, the scaling factors resulting from the fit are shown in Table 10.1.

μ_{WZQCD}	0.847 ± 0.096
μ_{ZZ}	1.1 ± 0.05

Table 10.1: Scaling factors for the floating normalization for WZQCD and ZZ in simultaneous WZ Drell-Yan and ZZ CR only fit with data

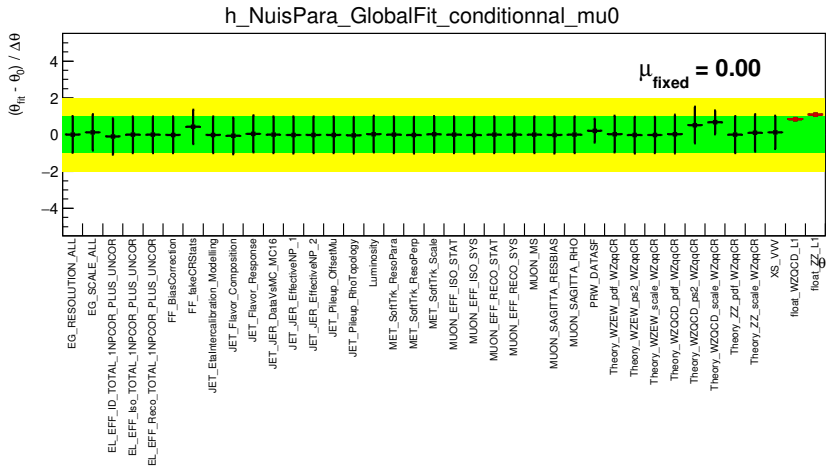


Figure 10.4: Nuisance parameter pulls of the Drell-Yan background only control region fit with data.

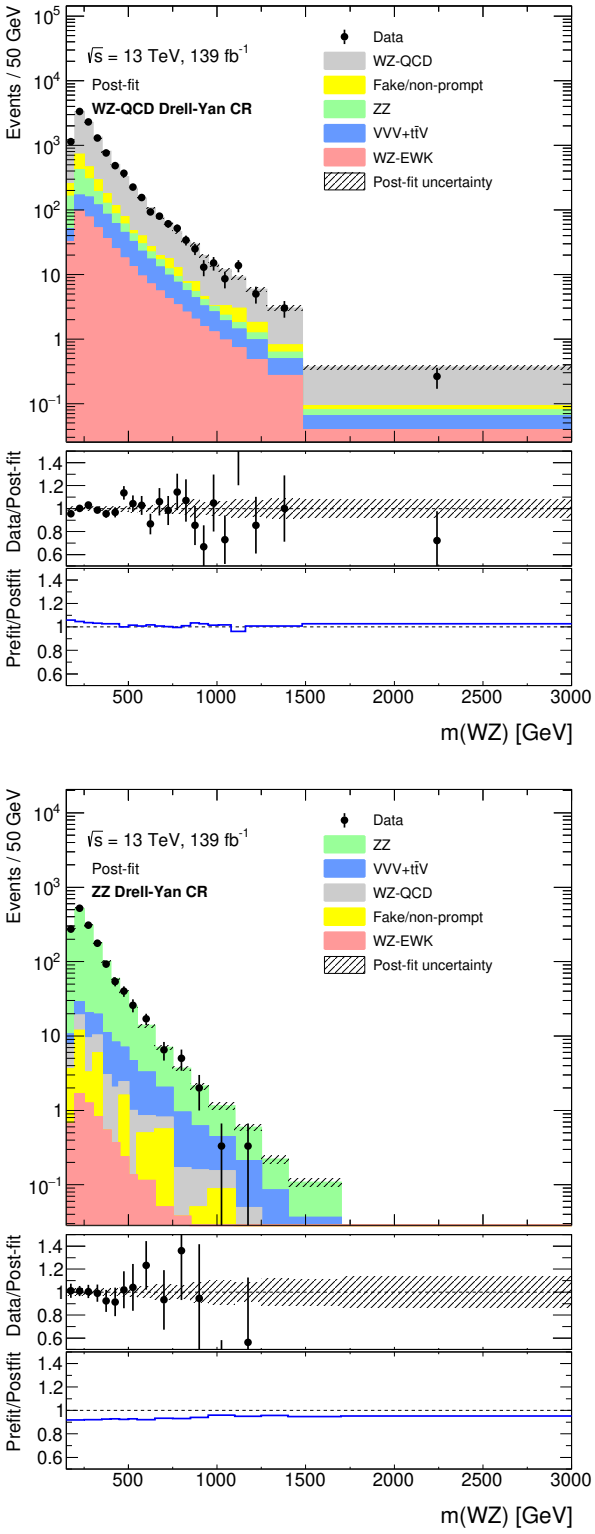


Figure 10.5: The post-fit invariant mass distributions of the WZ Drell-Yan CR and ZZ are shown. The ratio panel in the middle shows the ratio Data/Post-fit background, the post-fit uncertainties are shown as shaded area. The panel on the bottom shows the pre-fit/post-fit ratio.

10.4.2 VBF ANN Control regions fit:

For each of the VBF signal regions simultaneous fits of their respective WZ VBF control region (see Section 8.1.2 for the definitions) and the $ZZJJ$ control region (see Section 8.1.3 for definitions) are performed. The WZ-QCD and the $ZZjj$ backgrounds normalization are free-floating parameters, while all the other backgrounds are allow to move within the uncertainties.

The results for the control region only fits are shown in the following:

10.4.2.1 HVT VBF ANN Control regions only fit:

The nuisance parameter pull of the fit is shown in Figure 10.6, a small constraint in the systematic uncertainty introduced by the statistic uncertainty of the control regions used for Fake rate measurement is observed. The post-fit distributions of the $ZZJJ$ and WZ-QCD ANN VBF control regions are shown in Figure 10.7, the $ZZJJ$ control region shows a good agreement with data, while the WZ-QCD normalization shows a normalization correction of about 28% for all masses, in agreement with the results seen by the VBS WZ analysis [127], the scaling factors resulting from the fit are shown in Table 10.2.

μ_{WZQCD}	0.724 ± 0.099
μ_{ZZJJ}	1.07 ± 0.29

Table 10.2: Scaling factors for the floating normalization of WZ-QCD and ZZ in simultaneous WZ ANN VBF and ZZJJ CR only fit with data

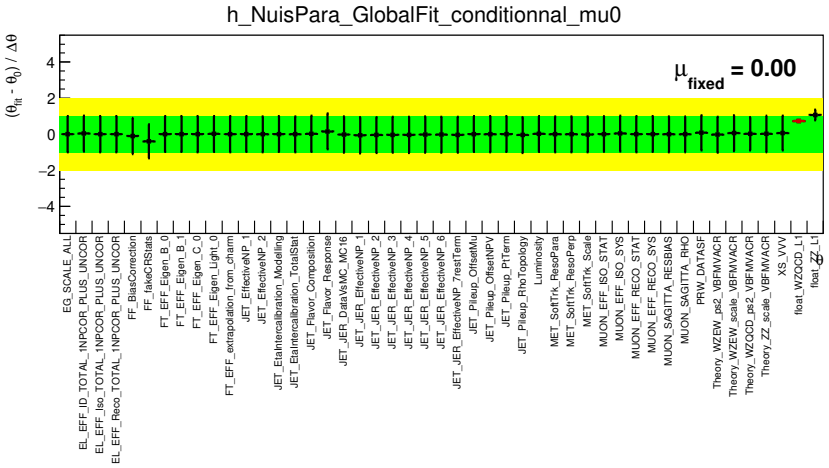


Figure 10.6: Nuisance parameter pulls of the HVT VBF ANN background only control region fit.

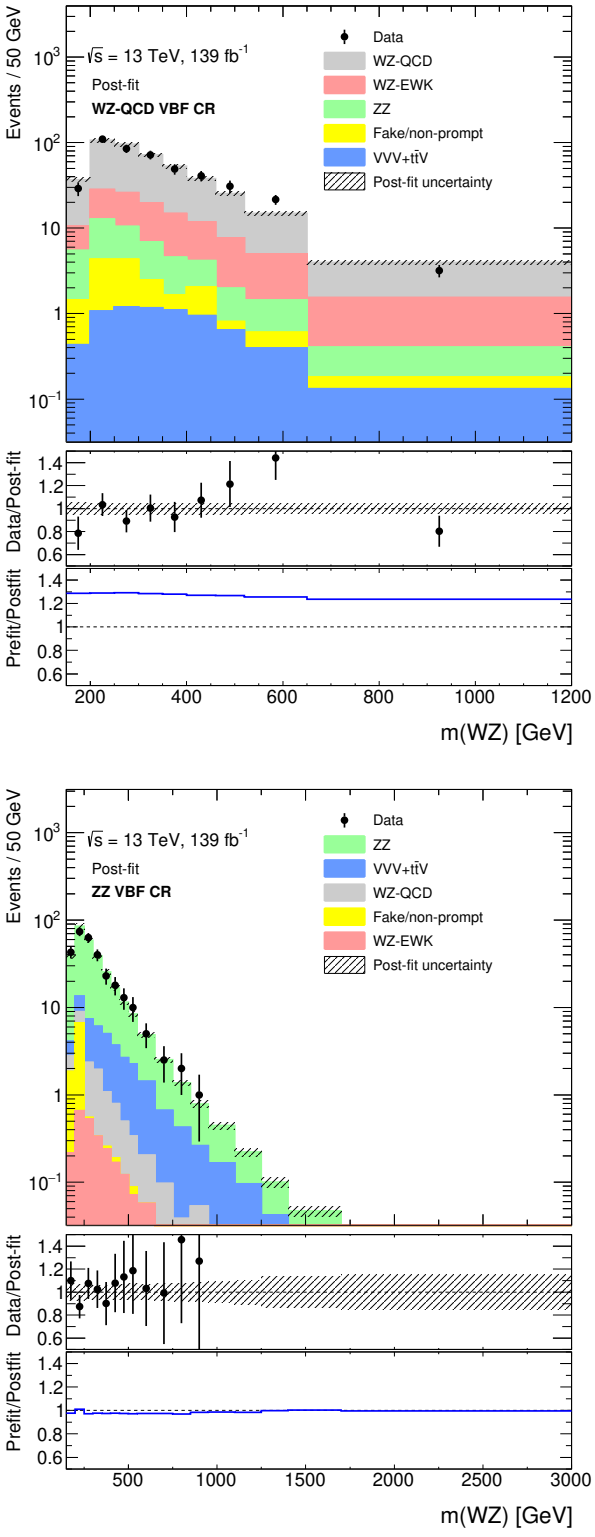


Figure 10.7: The post-fit invariant mass distributions of the WZ VBF ANN CR and ZZJJ are shown. The panel in the middle shows the ratio Data/Post-fit background, the post-fit uncertainties are shown as shaded area. The panel on the bottom shows the pre-fit/post-fit ratio.

10.4.2.2 GM VBF ANN Control regions only fit:

The nuisance parameter pull of the fit is shown in Figure 10.8. Similar to what is seen for the HVT VBF, no significant pulls or overconstraints are visible except for a small constraint in the systematic uncertainty introduced by the statistic uncertainty of the control regions used for fake rate measurement. The post-fit distributions of the $ZZJJ$ and WZ GM ANN VBF control regions are shown in Figure 10.9 where the $ZZJJ$ control region shows a good agreement with data, while the WZ-QCD normalization shows the need for a normalization correction of ∼28%, the scaling factors resulting from the fit are shown in Table 10.3.

μ_{WZQCD}	0.724 ± 0.103
μ_{ZZ}	1.06 ± 0.29

Table 10.3: Scaling factors for the floating normalization for WZQCD and ZZJJ in simultaneous ANN CR and ZZJJ CR only fit with data

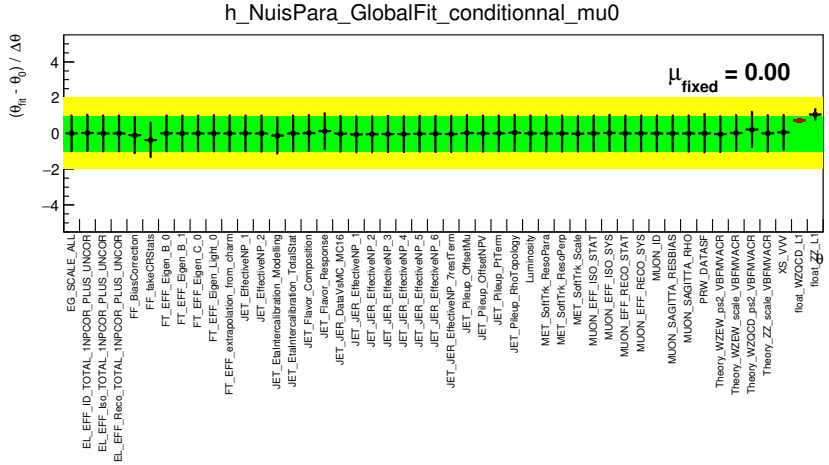


Figure 10.8: Nuisance parameter pulls of the GM VBF ANN background only fit.

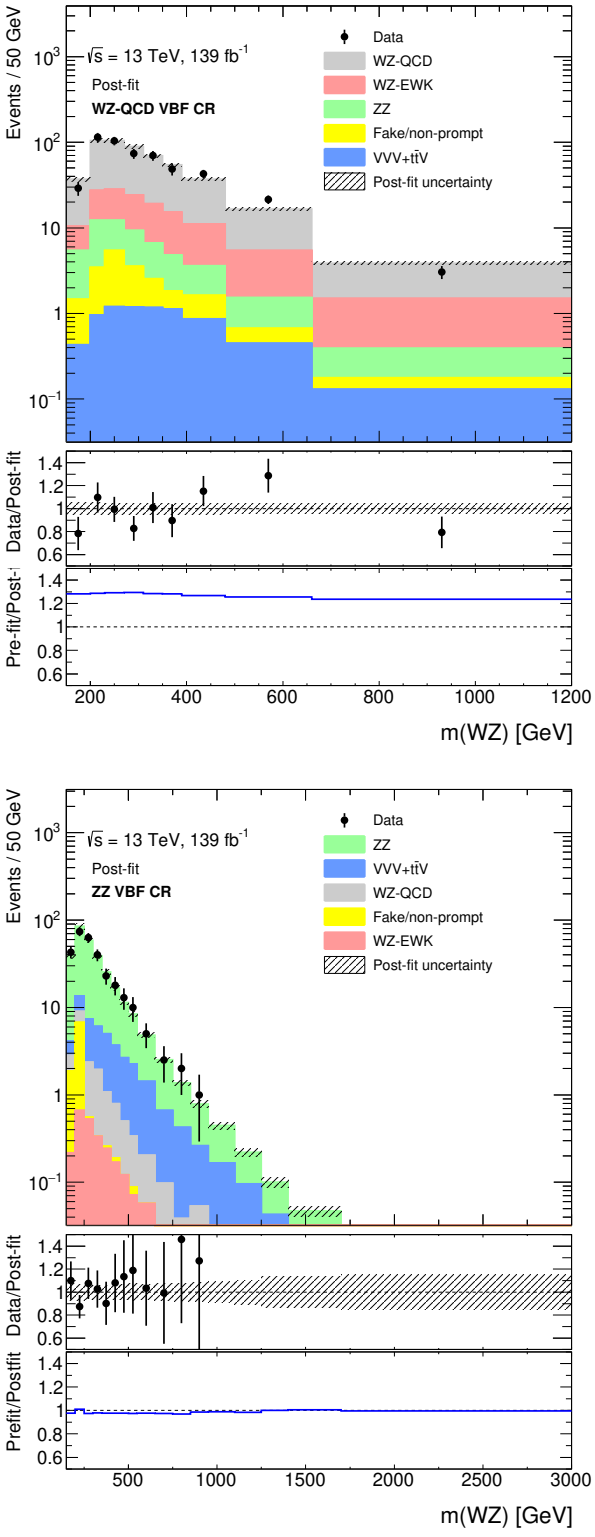


Figure 10.9: The post-fit invariant mass distributions of the WZ VBF ANN CR and $ZZJJ$ are shown. The panel in the middle shows the ratio Data/Post-fit background, the post-fit uncertainties are shown as shaded area. The panel on the bottom shows the pre-fit/post-fit ratio.

10.4.3 VBF cut-based Control regions only fit

The nuisance parameter pull of the fit is shown in Figure 10.10, a small constraint is observed in the systematic uncertainty introduced by the statistic uncertainty of the control regions used for Fake rate measurement. The post-fit distributions of the $ZZJJ$ and cut-based WZ-QCD VBF control regions are shown in Figure 10.11, where the $ZZJJ$ control region shows a good agreement with data, while the WZ-QCD normalization shows a shape and normalization correction up to 15%, the scaling factors resulting from the fit are shown in Table 10.4.

μ_{WZQCD}	0.867 ± 0.275
μ_{ZZ}	1.02 ± 0.28

Table 10.4: Scaling factors for the floating normalization for WZQCD and ZZJJ in simultaneous cut-based VBF CR and ZZJJ CR only fit with data

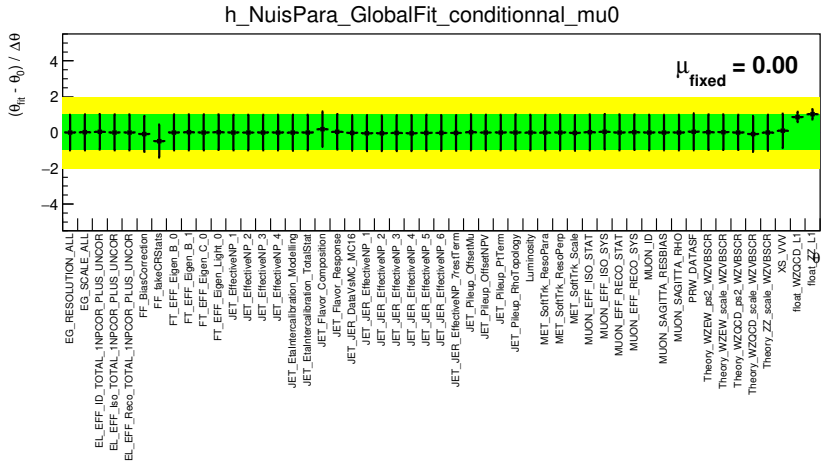


Figure 10.10: Nuisance parameter pulls of the VBF cut-based background only control region fit.

Similar to the VBF ANN control regions fit, the VBF cut-based control regions only fit give the same results for both the GM VBF control region and HVT VBF control region, since they are the same regions with only slightly different binning.

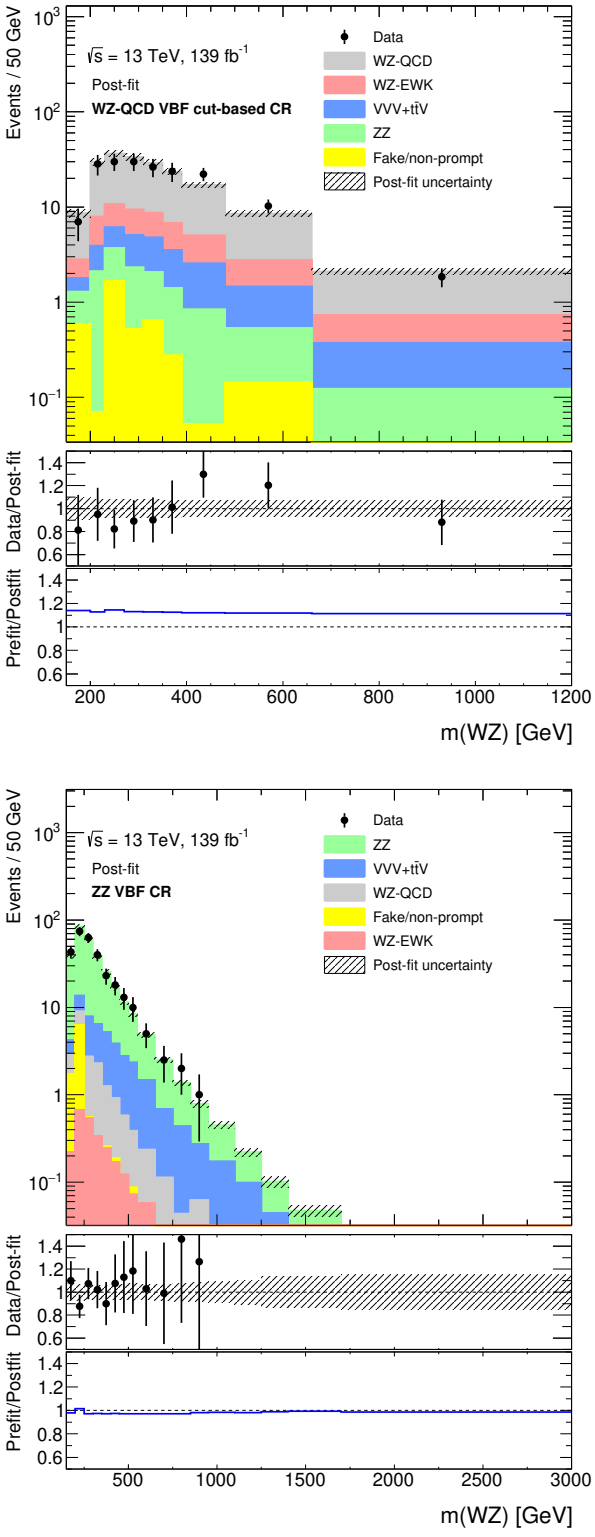


Figure 10.11: The post-fit invariant mass distributions of the WZ VBF cut-based CR and ZZ are shown. The panel in the middle shows the ratio Data/Post-fit background, the post-fit uncertainties are shown as shaded area. The panel on the bottom shows the pre-fit/post-fit ratio.

Chapter 11

Results

"The measure of a person, of a hero, is how well they succeed at being who they are."

Avengers: Endgame

In this Chapter the results of the statistical analysis applied to the selected data in the signal regions are presented. As discussed in Chapter 10, a total of 5 fits are performed, targeting the two production modes Drell-Yan and VBF. For the Drell-Yan analysis, noly the cut-based signal region selection and HVT model are used. For the VBF fit, two models are used, the HVT (spin 1) and GM (spin 0), and two signal regions either using cut-based selection or ANN are performed.

The Drell-Yan fits include the Drell-Yan SR, WZ-QCD CR and ZZ-CR. The VBF configuration fits include the corresponding VBF-SR, corresponding $WZjj$-QCD CR and $ZZjj$-CR. The Drell-Yan configuration is used to search for a W' boson predicted by the HVT model A. Two VBF-fits are performed using each of the VBF configuration: one for the search for a VBF produced W' predicted by the HVT model C, and the second fit for the charged Higgs boson, $H_5^{\pm}$, search as predicted by the GM model. Separate fits are performed for the different models tested and for different resonance mass hypotheses. The fit performed is background-only fit with data, where the signal strength parameter, μ, is fixed at 0.

11.1 Drell-Yan fit

A simultaneous background fit with data was performed in the Drell-Yan configuration by fitting the Drell-Yan signal region, together with the Drell-Yan WZ-QCD and ZZ control regions. The normalization of the WZ-QCD and ZZ backgrounds are also free-floating parameters of the fit.

Figure 11.1 and Table 11.1 shows the yields and uncertainties evaluated after a background only fit to the data in the Drell-Yan signal region and WZ-QCD and ZZ control regions. The uncertainty of the total expected estimate can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources. Table 11.2 shows the value of the normalization parameters for the WZ-QCD and ZZ backgrounds after the fit. The pulls of nuisance parameters for the fit

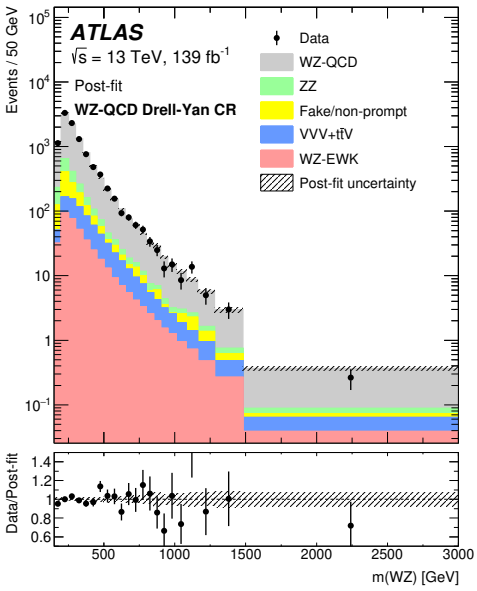
are shown in Figure 11.2.

	DY signal region	DY WZ QCD CR	ZZ CR
WZ-QCD	1734 ± 77	8438 ± 409	30 ± 3
WZ-EWK	89 ± 10	386 ± 41	6 ± 0.3
$VVV + t\bar{t}V$	148 ± 27	413 ± 75	65 ± 12
ZZ	95 ± 5	681 ± 42	1431 ± 42
Fakes/non-prompt leptons	88 ± 49	606 ± 376	20 ± 7
Total background	2155 ± 71	10524 ± 164	1554 ± 47
Observed	2155	10522	1554

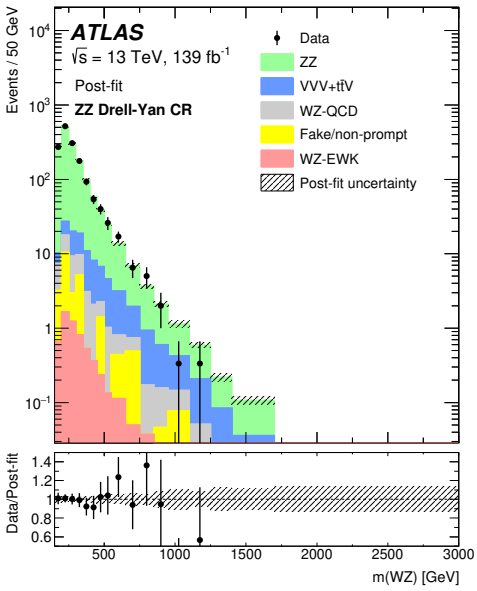
Table 11.1: Expected and observed post-fit yields in the Drell-Yan signal region, the corresponding WZ-QCD and ZZ control regions. The yields and uncertainties are presented after the background-only fit to the data in the Drell-Yan signal region with the corresponding WZ-QCD and ZZ control regions. The uncertainty in the total background estimate is smaller than the sum in quadrature of the individual background contributions due to anti-correlations between the estimates of different background sources.

μ_{WZQCD}	0.88 ± 0.09
μ_{ZZ}	1.10 ± 0.05

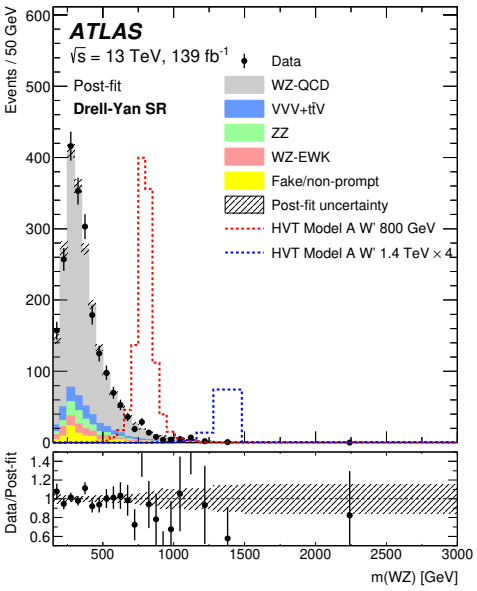
Table 11.2: Scaling factors for the floating normalization for WZ QCD and ZZ in simultaneous Drell-Yan SR, WZ-QCD and ZZ CRs background only fit with data



(a) *WZ*-QCD Drell-Yan control region



(b) *ZZ* control region



(c) Drell-Yan signal region

Figure 11.1: Comparisons of the data and the expected background distributions of the *WZ* invariant mass in the Drell-Yan signal region and its respective *WZ*-QCD and *ZZ* control regions. The background predictions are obtained through a background-only simultaneous fit to the Drell-Yan signal region and the *WZ*-QCD Drell-Yan and *ZZ* control regions. For illustration, the expected distribution from an HVT W' resonance with a mass of 800 GeV and 1.4 TeV, normalized to their predicted cross-section, are shown in the signal region. The bottom panels show the ratios of the data to the post-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

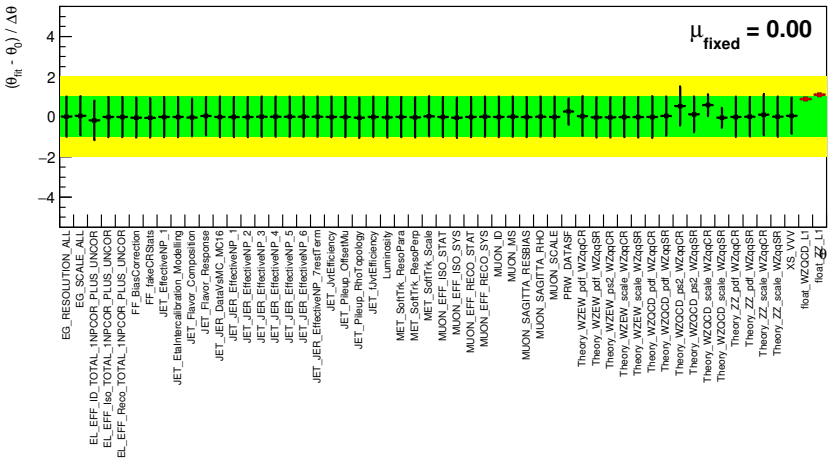


Figure 11.2: Nuisance parameter pulls in the Drell-Yan signal region after background only fit with WZ-QCD and ZZ control regions

11.2 VBF ANN fit

A simultaneous fit was performed in the VBF category by fitting the VBF ANN signal region, and the $WZjj$-QCD ANN and $ZZjj$ control regions. The normalization of the WZ-QCD, and ZZ are free-floating parameters of the fit. Since we have two slightly different binnings, each for the GM model and the HVT VBF model, we do this fit twice. The post fit results are quite similar for the two fits, as expected, because the regions being fit is the same with only sligthly different binning.

11.2.1 HVT VBF ANN model fit

The W' HVT signal hypothesis was fitted in this region with the signal region and $WZjj$-QCD ANN control region having optimised HVT VBF binning.

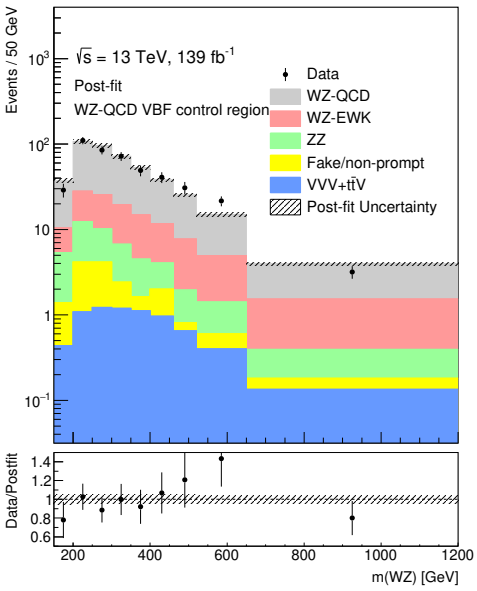
Table 11.3 and Figure 11.3 shows the yields and uncertainties evaluated after a background only fit to the data in the VBF ANN signal region and *WZjj*-QCD ANN and *ZZjj* control regions. The uncertainty on the total expected estimate can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources. Table 11.4 shows the value of the normalization parameters for the WZ-QCD and ZZ backgrounds after the fit. The pulls of nuisance parameters for the fit are shown in Figure 11.4.

	VBF ANN SR	$WZjj$-QCD CR	$ZZjj$ CR
WZ-QCD	28 ± 4	373 ± 24	9 ± 1
WZ-EWK	26 ± 3	97 ± 7	3 ± 0.3
$VVV + t\bar{t}V$	0.9 ± 0.2	10 ± 2	32 ± 7
ZZ	5 ± 1	34 ± 4	253 ± 19
Fakes/non-prompt leptons	0.3 ± 0.7	11 ± 5	7 ± 4
Total background	61 ± 6	525 ± 25	305 ± 20
Observed	67	522	305

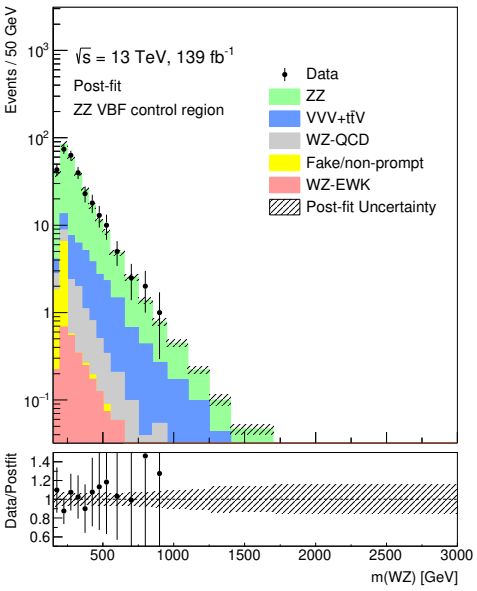
Table 11.3: Expected and observed post-fit yields in the VBF ANN signal region and *WZjj*-QCD ANN and *ZZjj* control regions. The yields and uncertainties are presented after the background-only fit to the data in the VBF ANN signal region with the WZ-QCD VBF ANN and ZZ VBF control regions. The uncertainty in the total background estimate is smaller than the sum in quadrature of the individual background contributions due to anti-correlations between the estimates of different background sources.

$\mu_{WZjjQCD}$	0.70 ± 0.07
μ_{ZZjj}	0.97 ± 0.23

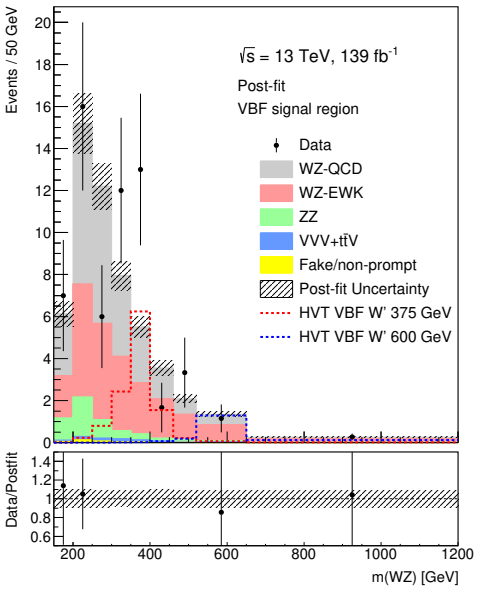
Table 11.4: Scaling factors for the floating normalization for WZ QCD and ZZ in simultaneous VBF ANN SR, $WZjj$-QCD and $ZZjj$ CRs background only fit with data



(a) *WZjj*-QCD control region



(b) *ZZjj* control region



(c) VBF signal region

Figure 11.3: Comparisons of the data and the expected background distributions of the *WZ* invariant mass in the VBF signal region and its respective *WZjj*-QCD ANN and *ZZjj* control regions. The background predictions are obtained through a background-only simultaneous fit to the VBF signal region and the *WZ*-QCD and *ZZ* VBF control regions. For illustration, the expected distribution from a HVT W' model resonance with a mass of 375 GeV and 600 GeV are shown in the signal region. The bottom panels show the ratios of the data to the post-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions. The p-values will be discussed later in section 11.6

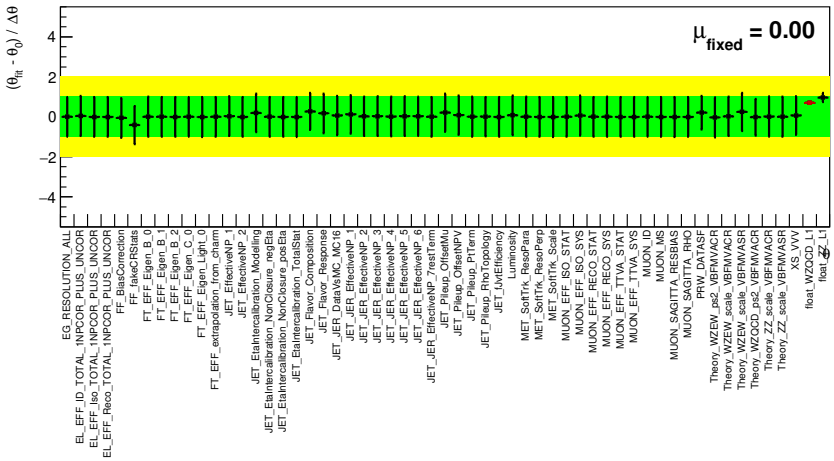


Figure 11.4: Nuisance parameter pulls in the VBF ANN signal region after background only fit with $WZjj$-QCD and $ZZjj$ control regions

11.2.2 GM VBF ANN model fit

The $H_5^{\pm}$ GM signal hypothesis was fitted in this region with the signal region and $WZjj$-QCD ANN control region having optimised GM binning.

Table 11.5 and Figure 11.5 shows the yields and uncertainties evaluated after a background only fit to the data in the VBF ANN signal region and *WZjj*-QCD ANN and *ZZjj* control regions. The uncertainty of the total expected estimate can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources. Table 11.6 shows the value of the normalization parameters for the WZ-QCD and ZZ backgrounds after the fit. The pulls of nuisance parameters for the fit are shown in Figures 11.6.

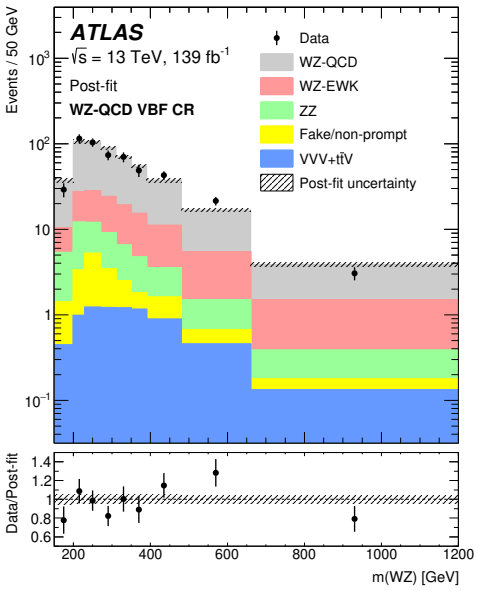
Figure 11.7 shows the background only post-fit distributions of some of the variables used in the ANN training. They all show good agreement between data and simulation.

	VBF ANN SR	$WZjj$-QCD CR	$ZZjj$ CR
WZ-QCD	29 ± 4	373 ± 25	9 ± 1
WZ-EWK	26 ± 3	97 ± 8	3 ± 0.3
$VVV + t\bar{t}V$	0.9 ± 0.2	10 ± 2	32 ± 7
ZZ	5 ± 1	33 ± 5	253 ± 19
Fakes/non-prompt leptons	0.3 ± 0.8	11 ± 5	7 ± 4
Total background	61 ± 6	525 ± 25	305 ± 20
Observed	67	522	305

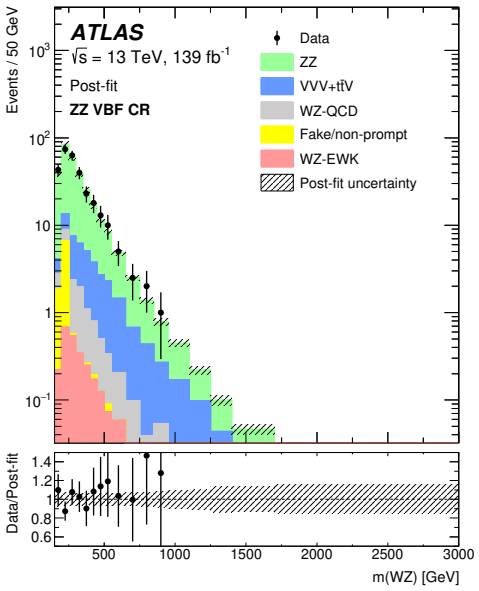
Table 11.5: Expected and observed post-fit yields in the VBF ANN signal region and *WZjj*-QCD ANN and *ZZjj* control regions. The yields and uncertainties are presented after the background-only fit to the data in the VBF ANN signal region with the WZ-QCD VBF ANN and ZZ VBF control regions. The uncertainty in the total background estimate is smaller than the sum in quadrature of the individual background contributions due to anti-correlations between the estimates of different background sources.

$\mu_{WZjjQCD}$	0.70 ± 0.09
μ_{ZZjj}	0.97 ± 0.23

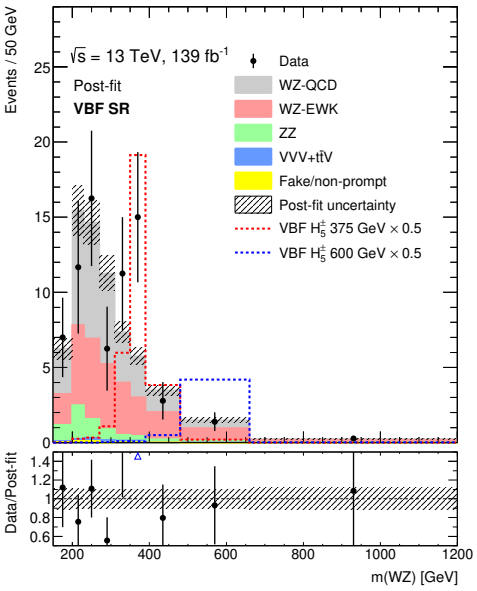
Table 11.6: Scaling factors for the floating normalization for WZ-QCD and ZZ in simultaneous VBF ANN SR, $WZjj$-QCD and $ZZjj$ CRs background only fit with data



(a) *WZjj*-QCD control region



(b) *ZZjj* control region



(c) VBF signal region

Figure 11.5: Comparisons of the data and the expected background distributions of the *WZ* invariant mass in the VBF signal region and its respective *WZjj*-QCD ANN and *ZZjj* control regions. The background predictions are obtained through a background-only simultaneous fit to the VBF signal region and the *WZjj*-QCD and *ZZjj* VBF control regions. For illustration, the expected distribution from a GM $H_5^{\pm}$ model resonance with a mass of 375 GeV and 600 GeV are shown in the signal region. The bottom panels show the ratios of the data to the post-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions. The p-values will be discussed later in section 11.6

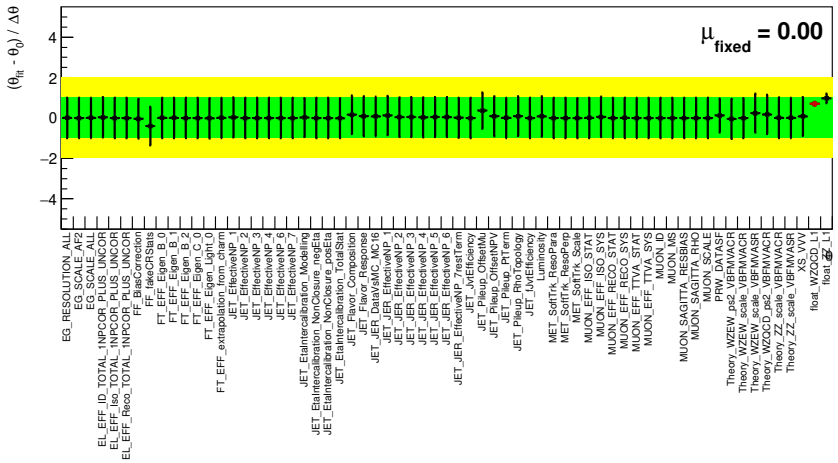


Figure 11.6: Nuisance parameter pulls in the VBF ANN signal region after background only fit with $WZjj$-QCD and $ZZjj$ control regions

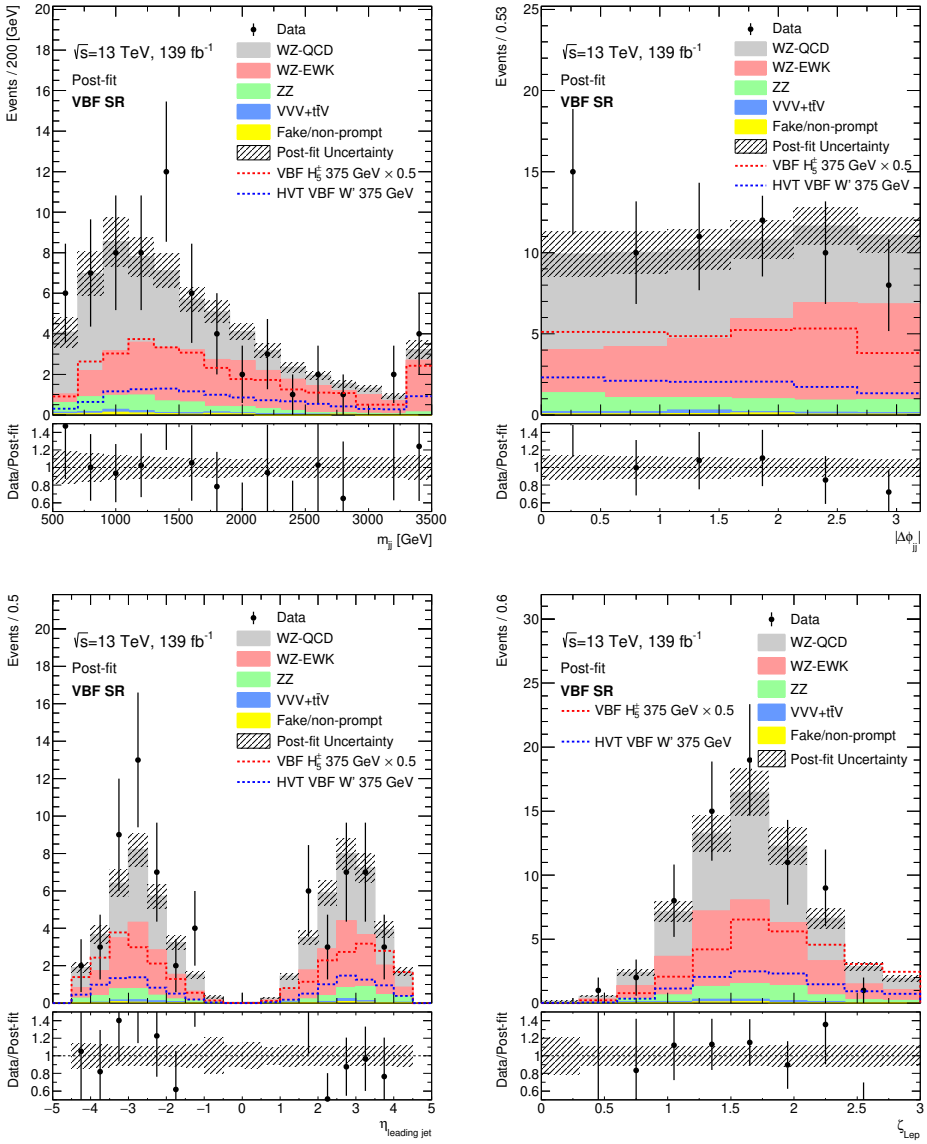


Figure 11.7: Comparisons of the observed data and the expected background distributions in the VBF signal region region for some of the variables used in the ANN training. The background predictions are obtained through a background-only simultaneous fit in the VBF signal region and the *WZ*-QCD and *ZZ* VBF control regions. The fitted background normalization and nuisance parameters have been propagated. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

11.3 VBF cut-based fit

A simultaneous fit was performed in the VBF category by fitting the VBF cut-based signal region, and the $WZjj$-QCD cut-based and $ZZjj$ control regions. The normalization of the WZ-QCD and ZZ background are free-floating parameters of the fit. Since we have two slightly different binnings, each for GM model and HVT VBF model, we do this fit twice. The post fit results are quite similar for the two fits, as expected, because the regions being fit is the same with only sligthly different binning.

11.3.1 HVT VBF cut-based model fit

The W' HVT signal hypothesis was fitted in this region with the signal region and $WZjj$-QCD ANN control region having optimised HVT VBF binning.

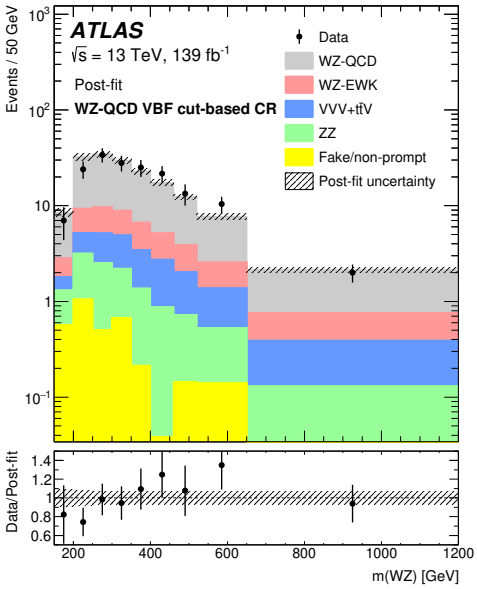
Table 11.7 and Figure 11.8 shows the yields and uncertainties evaluated after a background only fit to the data in the VBF ANN signal region and *WZjj*-QCD ANN and *ZZjj* control regions. The uncertainty of the total expected estimate can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources. Table 11.8 shows the value of the normalization parameters for the WZ-QCD and ZZ backgrounds after the fit. The pulls of nuisance parameters for the fit are shown in Figure 11.9.

	VBF cut-based SR	$WZjj$-QCD CR	$ZZjj$ CR
WZ-QCD	210 ± 24	144 ± 14	10 ± 3
WZ-EWK	100 ± 11	29 ± 2	3 ± 0.3
$VVV + t\bar{t}V$	31 ± 7	19 ± 4	32 ± 6
ZZ	26 ± 4	11 ± 2	253 ± 19
Fakes/non-prompt leptons	14 ± 5	4 ± 2	7 ± 3
Total background	382 ± 22	207 ± 15	305 ± 20
Observed	381	209	305

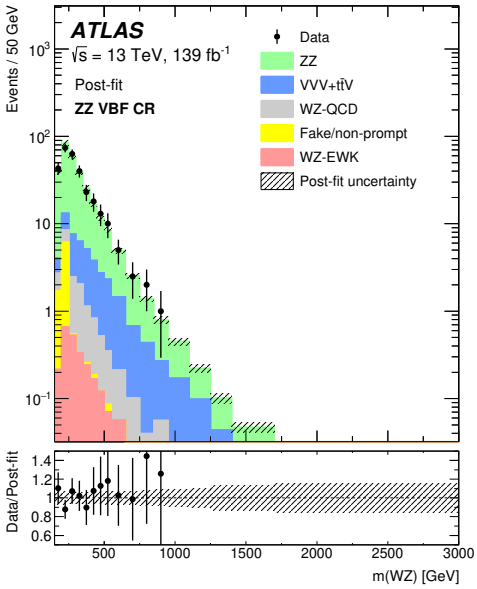
Table 11.7: Expected and observed post-fit yields in the VBF cut-based signal region and *WZjj*-QCD cut-based and *ZZjj* control regions. The yields and uncertainties are presented after the background-only fit to the data in the VBF cut-based signal region with the WZ-QCD VBF cut-based and ZZ VBF control regions. The uncertainty in the total background estimate is smaller than the sum in quadrature of the individual background contributions due to anti-correlations between the estimates of different background sources.

$\mu_{WZjjQCD}$	0.77 ± 0.22
μ_{ZZjj}	1.08 ± 0.27

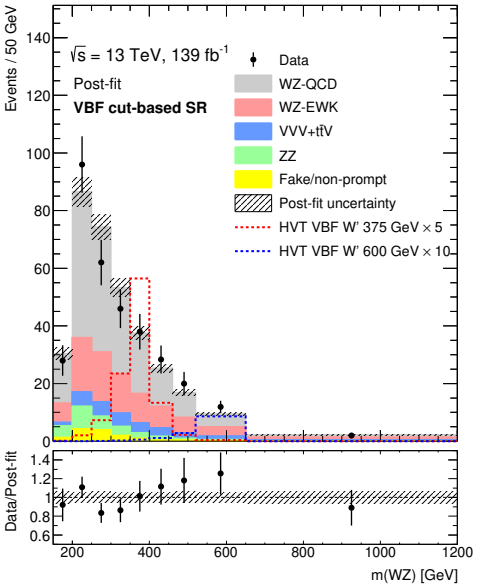
Table 11.8: Scaling factors for the floating normalization for WZ QCD and ZZ in simultaneous VBF cut-based SR, $WZjj$-QCD and $ZZjj$ CRs background only fit with data



(a) *WZjj*-QCD control region



(b) *ZZjj* control region



(c) VBF signal region

Figure 11.8: Comparisons of the data and the expected background distributions of the *WZ* invariant mass in the VBF signal region and its respective *WZjj*-QCD ANN and *ZZjj* control regions. The background predictions are obtained through a background-only simultaneous fit to the VBF signal region and the *WZ*-QCD and *ZZ* VBF control regions. For illustration, the expected distribution from a HVT W' model resonance with a mass of 375 GeV and 600 GeV are shown in the signal region. The bottom panels show the ratios of the data to the post-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

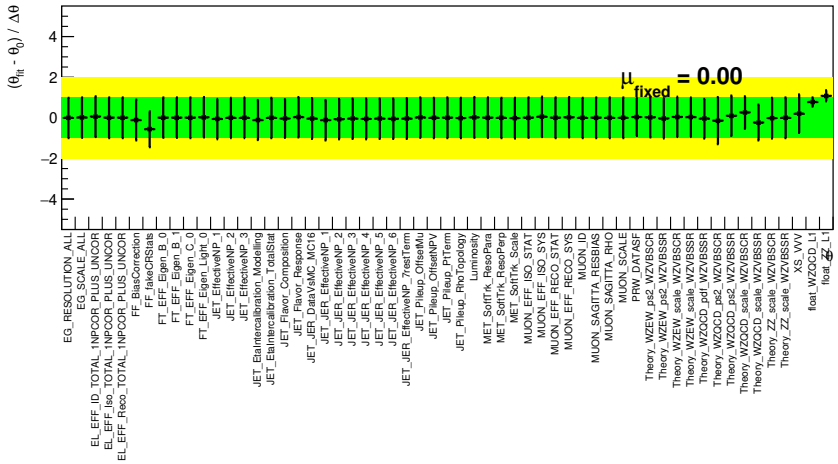


Figure 11.9: Nuisance parameter pulls in the VBF cut-based signal region after background only fit with $WZjj$-QCD and $ZZjj$ control regions

11.3.2 GM VBF cut-based model fit

The $H_5^{\pm}$ GM signal hypothesis was fitted in this region with the cut-based signal region and $WZjj$-QCD cut-based control region having optimised GM binning.

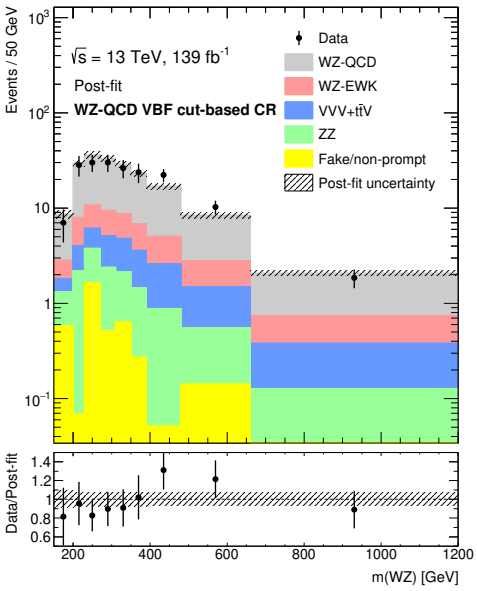
Table 11.9 and Figure 11.10 shows the yields and uncertainties evaluated after a background only fit to the data in the VBF cut-based signal region and *WZjj*-QCD cut-based and *ZZjj* control regions. The uncertainty of the total expected estimate can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources. Table 11.10 shows the value of the normalization parameters for the WZ-QCD and ZZ backgrounds after the fit. The pulls of nuisance parameters for the fit are shown in Figures 11.11.

	VBF cut-based SR	$WZjj$-QCD CR	$ZZjj$ CR
WZ-QCD	210 ± 24	144 ± 15	10 ± 2
WZ-EWK	101 ± 10	29 ± 2	3 ± 0.3
$VVV + t\bar{t}V$	31 ± 7	19 ± 4	32 ± 6
ZZ	27 ± 4	11 ± 2	254 ± 19
Fakes/non-prompt leptons	14 ± 5	4 ± 2	7 ± 3
Total background	382 ± 22	207 ± 15	305 ± 20
Observed	381	209	305

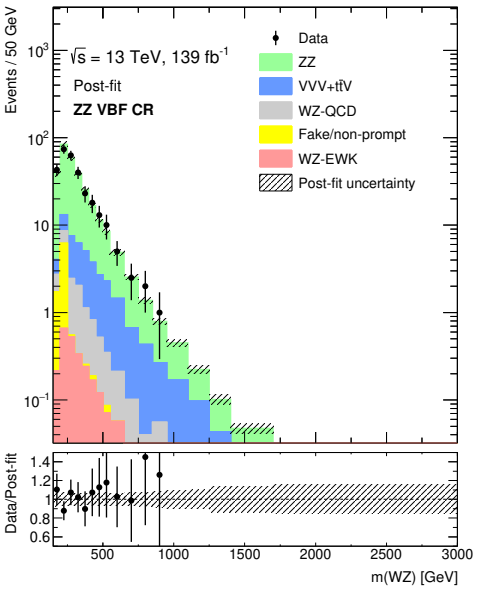
Table 11.9: Expected and observed post-fit yields in the VBF cut-based signal region and *WZjj*-QCD cut-based and *ZZjj* control regions. The yields and uncertainties are presented after the background-only fit to the data in the VBF cut-based signal region with the WZ-QCD VBF cut-based and ZZ VBF control regions. The uncertainty in the total background estimate is smaller than the sum in quadrature of the individual background contributions due to anti-correlations between the estimates of different background sources.

$\mu_{WZjjQCD}$	0.77 ± 0.21
μ_{ZZjj}	1.08 ± 0.26

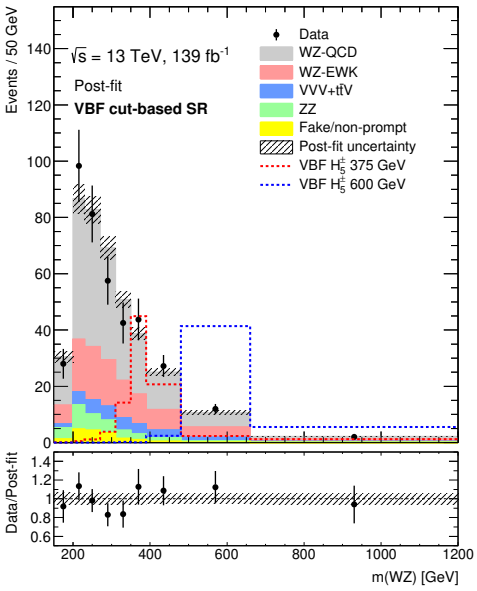
Table 11.10: Scaling factors for the floating normalization for WZ-QCD and ZZ in simultaneous VBF cut-based SR, $WZjj$-QCD and $ZZjj$ CRs background only fit with data



(a) *WZjj*-QCD control region



(b) *ZZjj* control region



(c) VBF signal region

Figure 11.10: Comparisons of the data and the expected background distributions of the *WZ* invariant mass in the VBF signal region and its respective *WZjj*-QCD cut-based and *ZZjj* control regions. The background predictions are obtained through a background-only simultaneous fit to the VBF signal region and the *WZjj*-QCD and *ZZjj* VBF control regions. For illustration, the expected distribution from a GM $H_5^{\pm}$ model resonance with a mass of 375 GeV and 600 GeV are shown in the signal region. The bottom panels show the ratios of the data to the post-fit background predictions. The uncertainty in the total background prediction, shown as grey bands, combines statistical and systematic contributions.

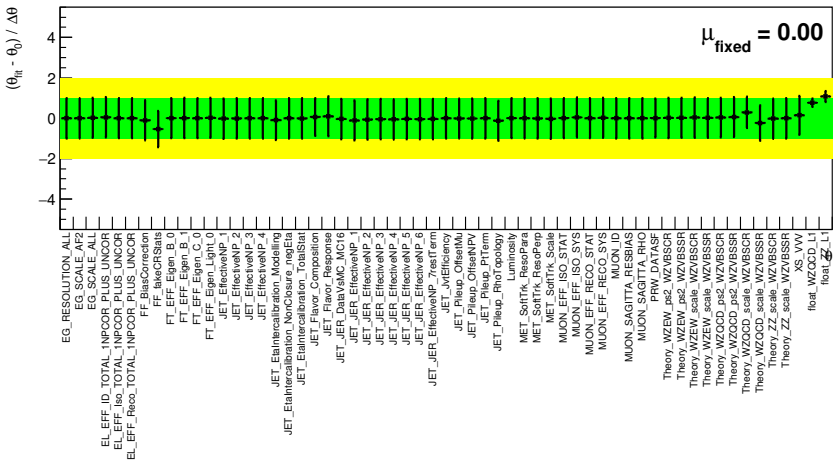


Figure 11.11: Nuisance parameter pulls in the VBF cut-based signal region after background only fit with $WZjj$-QCD and $ZZjj$ control regions

11.4 Exclusion Upper limit on the production of heavy resonances

Due to the absence of the excesses above the SM expectation, we set the 95% CL exclusion limits on the cross-section times branching fraction of the different signal models. Constraints on the production of heavy resonances are derived by repeating the fit to the signal-plus-background hypothesis for different signal models. Upper limits on cross-sections times branching fraction to WZ are calculated using the asymptotic approximation [123].

Nuisance parameter ranking plots are also shown to understand the impact of each nuisance parameters on the signal strength. The impact for an individual nuisance parameter is evaluated by reperforming the fit while fixing the respective nuisance parameter to the up- and down-variations around its best-fit value using its post-fit uncertainties.

For the HVT model search, Figure 11.12 presents the observed and expected limits on $\sigma \times B(W' \to WZ)$ at 95% CL as a function of the W' mass for the HVT model in the Drell-Yan signal region. Masses below 2.4 TeV can be excluded for Model A and 2.5 TeV for Model B. For resonance masses above 2 TeV, the exclusion limits become worse due to the poorer acceptance at high mass (see Figure 7.2). There was no significant gain in limits by splitting by leptonic channels as observed in Appendix E.

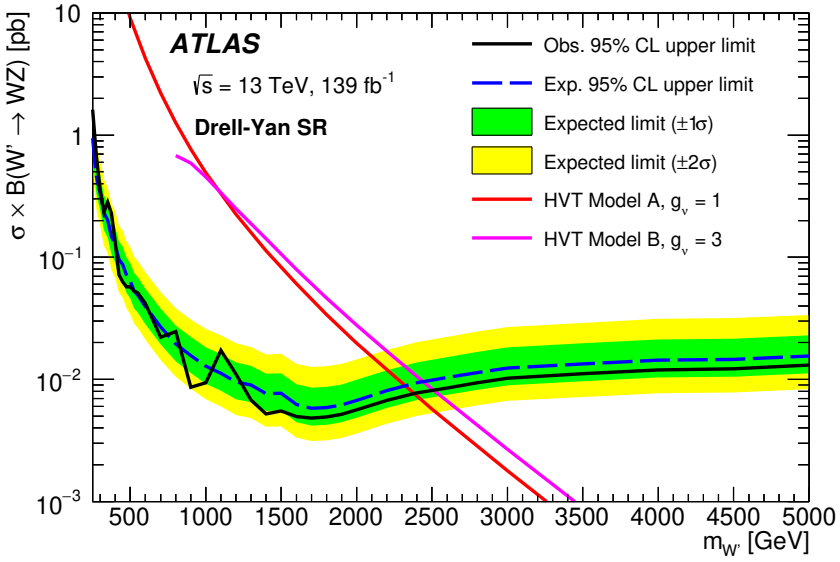


Figure 11.12: Observed and expected 95% CL exclusion upper limits on $\sigma \times B(W' \to WZ)$ for the Drell-Yan production of a W' boson in the HVT model as a function of its mass. The LO theory predictions for HVT Model A with $g_V = 1$ and Model B with $g_V = 3$ are also shown.

The ranking plot in Figure 11.13 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for the Drell-Yan signal and control region fit of signal of mass 500 GeV. The nuisance parameter with the highest impact on the signal strength is the jet flavour uncertainties followed by WZ-QCD theory parton shower uncertainty in the signal region.

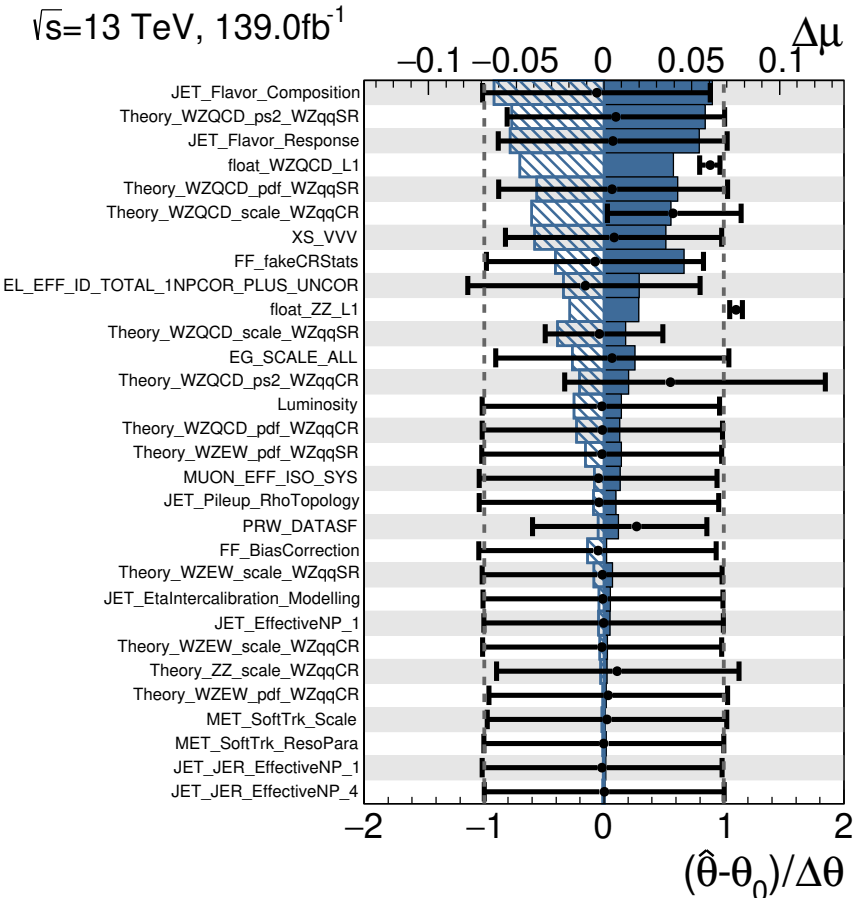


Figure 11.13: Ranking of the nuisance parameters included in the SR+CR fit according to their post-fit impact on the measured signal strength μ using Drell-Yan signal of mass 500 GeV and observed data with the production cross-section of the signal equalling the expected 95% CL upper limits. Only the top 30 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The impact of each nuisance parameter $\Delta\mu$, as shown by the blue coloured bars (upper x-axis), are computed by comparing the nominal best-fit μ with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its post-fit uncertainty $\pm\Delta\hat{\theta}$. The black points show the pulls of the nuisance parameters with respect to their nominal values, θ_0. These pulls and their relative post-fit errors, $(\hat{\theta} - \theta_0)/\Delta\theta$, refer to the lower scale.

Regarding the VBF production mode, the limit on $\sigma \times B(W' \to WZ)$ is shown in Figure 11.14. Because of large mass mixing, which depends on the coupling $g_V c_H$, between the SM gauge bosons and W', the theory curves start at different values of W' mass, close to the limit of validity of the HVT model. Masses below 340 GeV, 500 GeV and 700 GeV can be excluded for the HVT VBF model with $c_F = 0$ and $g_V c_H = 1.0$, 1.5 and 2.0, respectively. The ranking plot in Figure 11.15 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for this fit.

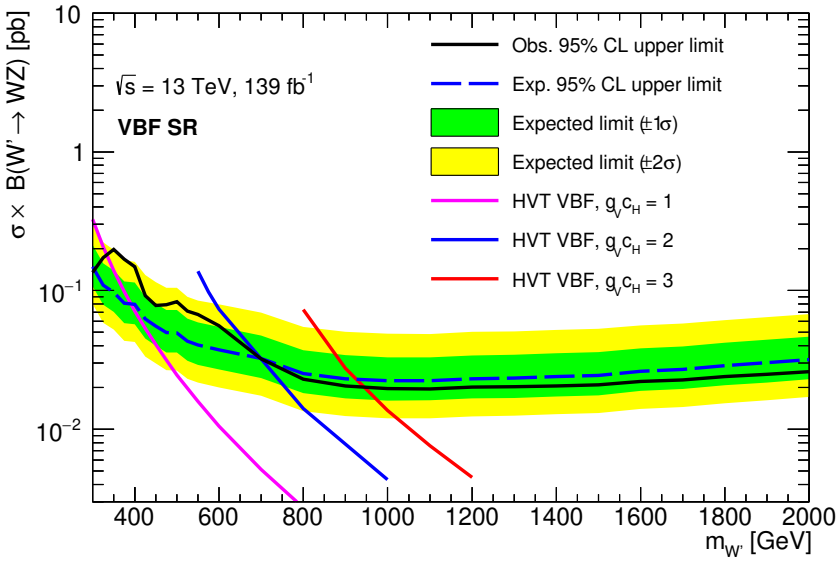


Figure 11.14: Observed and expected 95% CL upper limits on $\sigma \times B(W' \to WZ)$ for the VBF production of a W' boson in the HVT with parameter $c_F = 0$, as a function of its mass. The leading-order theory predictions for HVT VBF model with different values of the coupling parameters g_V and c_H are also shown.

The ranking plot in Figure 11.15 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for the ANN signal and control region fit of VBF produced HVT W' signal of mass 500 GeV. The nuisance parameter with the highest impact on the signal strength is the ZZ background normalization parameter followed by WZ-EWK modelling uncertainty. The jet uncertainties are the next highest ranked nuisance parameters.

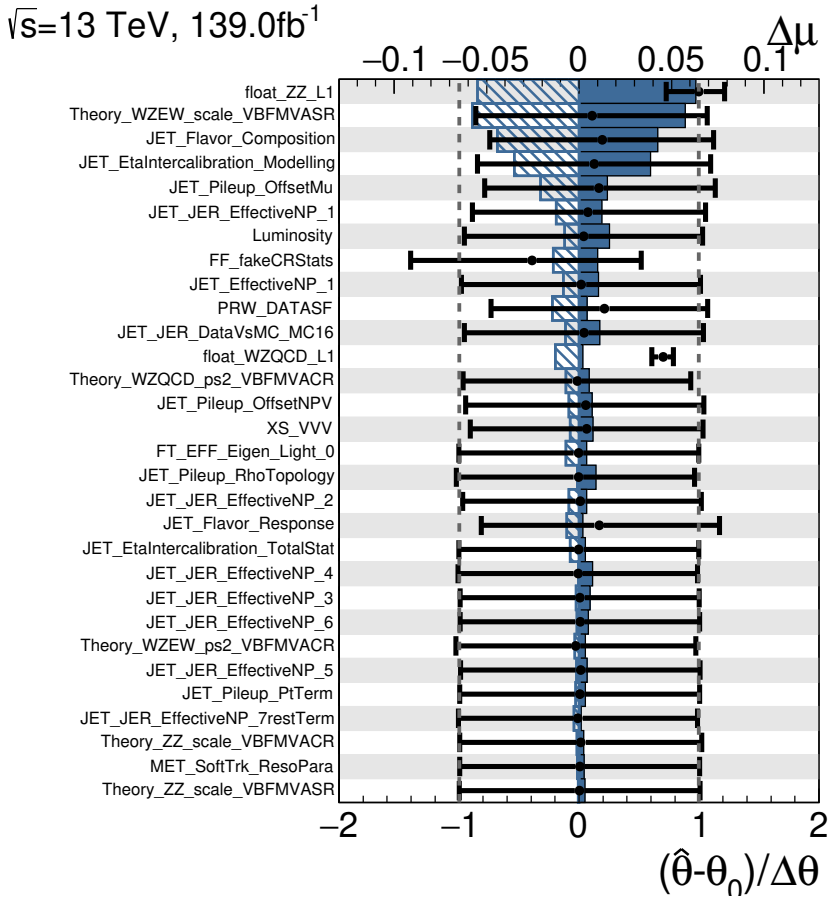
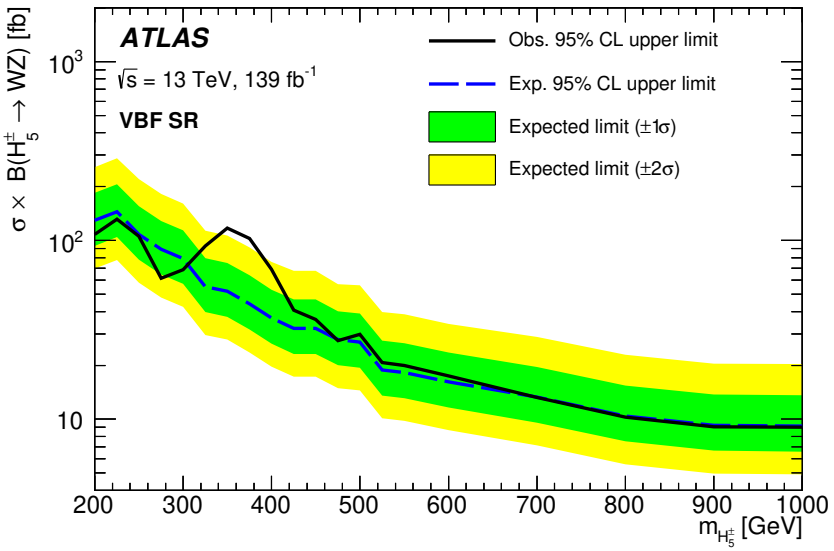
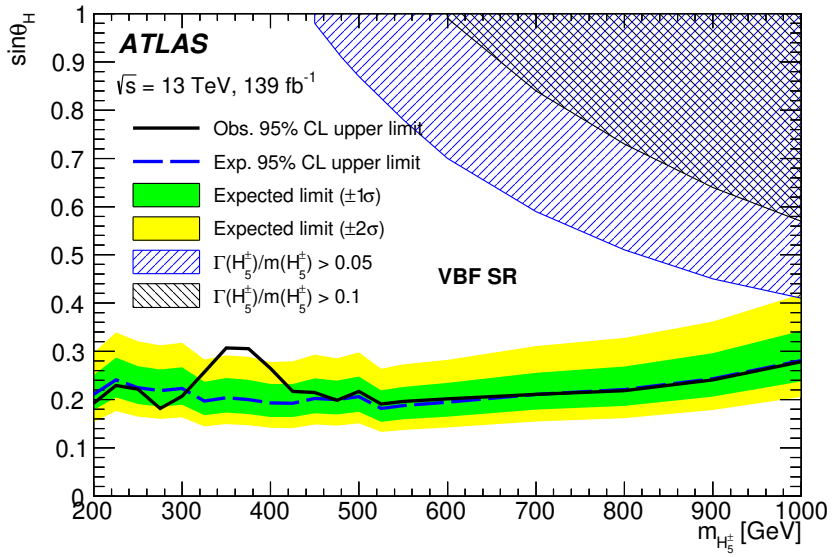


Figure 11.15: Ranking of the nuisance parameters included in the ANN SR+CR fit according to their post-fit impact on the measured signal strength μ using VBF produced HVT W' signal of mass 500 GeV and observed data with the production cross-section of the signal equalling the expected 95% CL upper limits. Only the top 30 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The impact of each nuisance parameter $\Delta\mu$, as shown by the blue coloured bars (upper x-axis), are computed by comparing the nominal best-fit μ with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its post-fit uncertainty $\pm\Delta\hat{\theta}$. The black points show the pulls of the nuisance parameters with respect to their nominal values, θ_0. These pulls and their relative post-fit errors, $(\hat{\theta} - \theta_0)/\Delta\theta$, refer to the lower scale.

For the $H_5^\pm$ GM model search, observed and expected exclusion limits at 95% confidence level (CL) on $\sigma \times B(H_5^\pm \to WZ)$ and on the mixing parameter $\sin\theta_H$ are shown in Figure 11.16. The intrinsic width of the scalar resonance, for $\sin\theta_H = 0.5$, is narrower than the detector resolution in the mass region explored. The shaded regions show the parameter space for which the $H_5^\pm$ width exceeds 5% and 10% of $m_{H_5^\pm}$. The ranking plot in Figure 11.17 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for this fit.



(a)



(b)

Figure 11.16: Observed and expected 95% CL upper limits (a) on $\sigma \times B(H_5^\pm \to WZ)$ and (b) on the parameter $\sin\theta_H$ of the GM model as a function of $m_{H_5^\pm}$. The shaded region shows where the theoretical intrinsic width of the resonance would be larger than 5% or 10% of the mass.

The ranking plot in Figure 11.17 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for the ANN VBF signal and control region fit of GM $H_5^{\pm}$ signal of mass 500 GeV. The nuisance parameter with the highest impact on the signal strength is the WZ-EWK modelling uncertainty followed by jet pileup uncertainty.

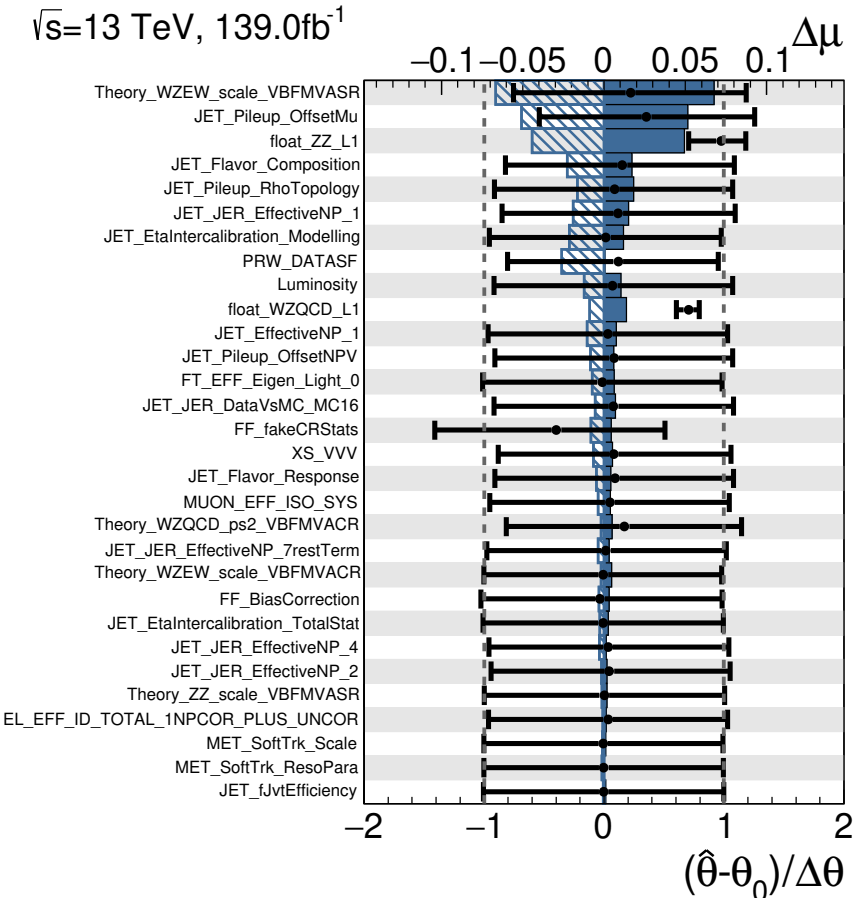


Figure 11.17: Ranking of the nuisance parameters included in the ANN SR+CR fit according to their post-fit impact on the measured signal strength μ using GM $H_5^{\pm}$ signal of mass 500 GeV and observed data with the production cross-section of the signal equalling the expected 95% CL upper limits. Only the top 30 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The impact of each nuisance parameter $\Delta\mu$, as shown by the blue coloured bars (upper x-axis), are computed by comparing the nominal best-fit μ with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its post-fit uncertainty $\pm\Delta\hat{\theta}$. The black points show the pulls of the nuisance parameters with respect to their nominal values, θ_0. These pulls and their relative post-fit errors, $(\hat{\theta} - \theta_0)/\Delta\theta$, refer to the lower scale.

Constraints on the production of heavy resonances are also derived by repeating the fit to the signal-plus-background hypothesis for different signal models using the cut-based analysis.

For the cut-based HVT model search, Figure 11.18 presents the observed and expected limits on $\sigma \times B(W' \to WZ)$ at 95% CL as a function of the W' mass. Masses below 370 GeV, 470 GeV and 590 GeV for HVT VBF production can be excluded for the HVT VBF model with $c_F = 0$ and $g_V c_H = 1.0$, 1.5, and 2.0, respectively. The ranking plot in Figure 11.19 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for this fit.

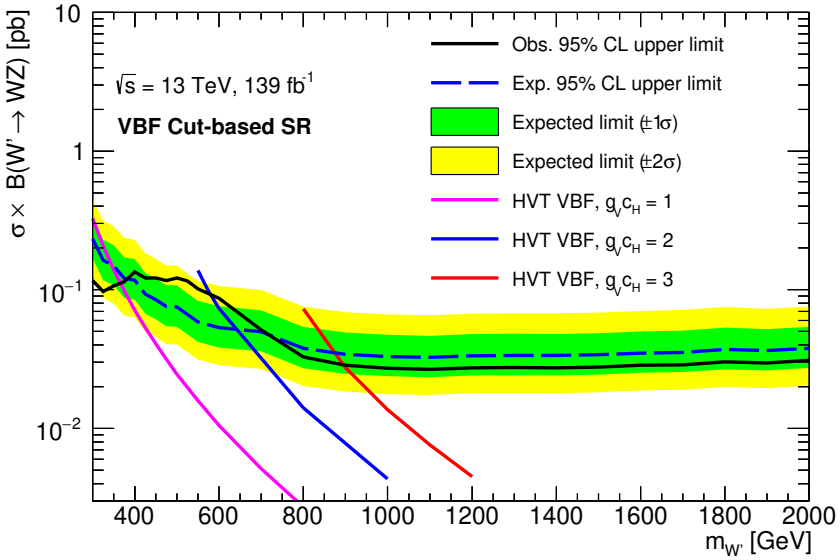


Figure 11.18: Using the cut-based VBF selection, the observed and expected 95% CL upper limits on $\sigma \times B(W' \to WZ)$ for the VBF production of a W' boson in the HVT with parameter $c_F = 0$, as a function of its mass. The LO theory predictions for HVT VBF model with different values of the coupling parameters g_V and c_H are also shown.

The ranking plot in Figure 11.19 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for the cut-based VBF signal and control region fit of VBF produced HVT W' signal of mass 500 GeV. The nuisance parameter with the highest impact on the signal strength is the ZZ background normalization parameter followed by $t\bar{t}V$ and VVV cross-sections uncertainty. The WZ modelling uncertainties in the signal region are the next highest ranked nuisance parameters.

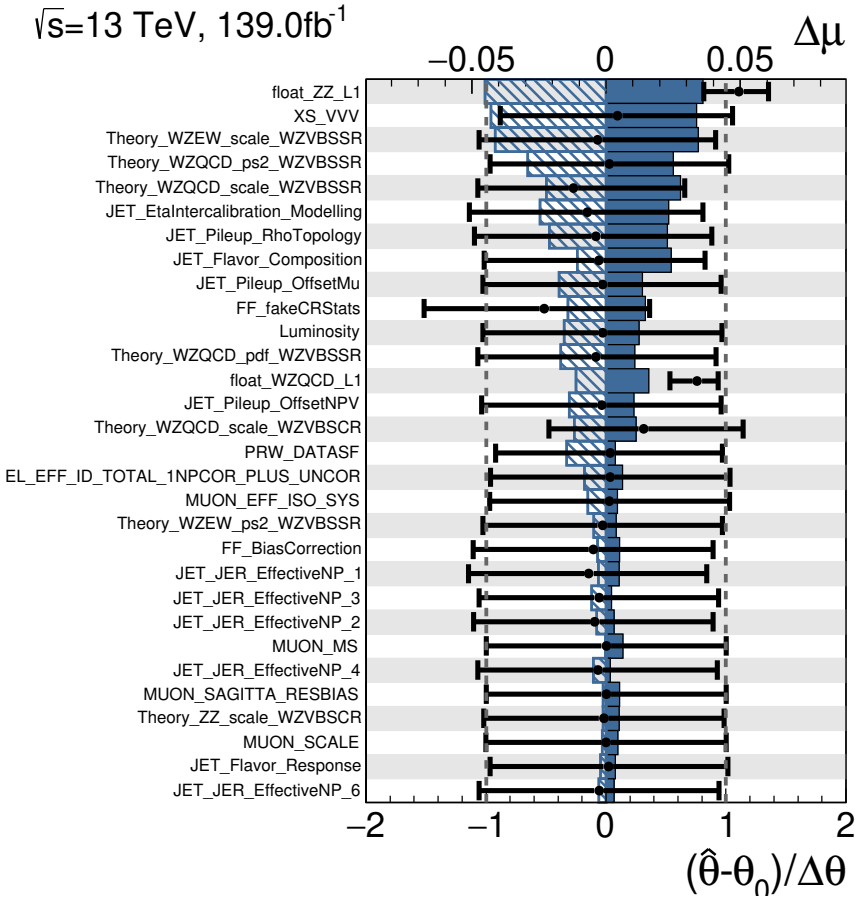
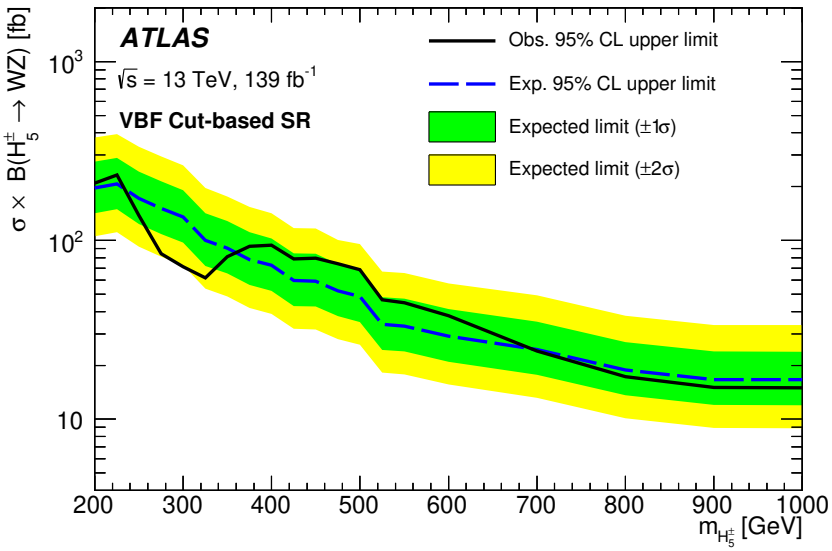
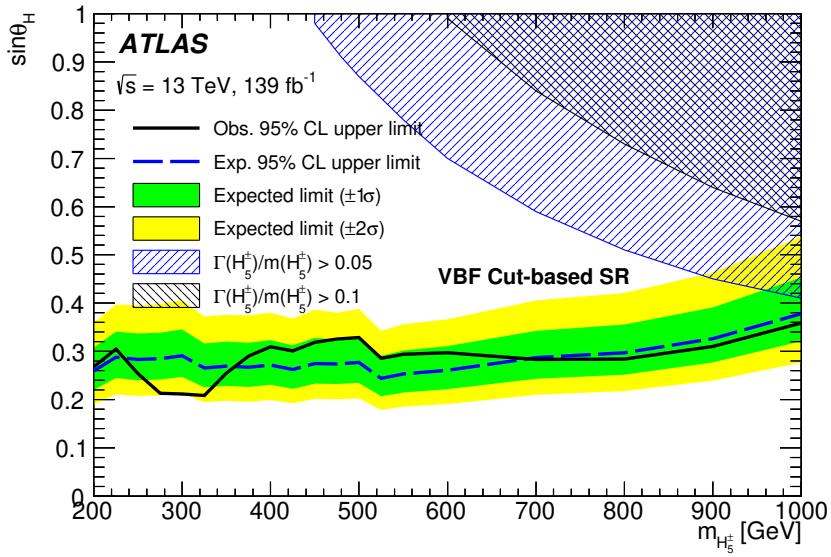


Figure 11.19: Ranking of the nuisance parameters included in the cut-based SR+CR fit according to their post-fit impact on the measured signal strength μ using HVT VBF produced W' signal of mass 500 GeV and observed data with the production cross-section of the signal equalling the expected 95% CL upper limits. Only the top 30 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The impact of each nuisance parameter $\Delta\mu$, as shown by the blue coloured bars (upper x-axis), are computed by comparing the nominal best-fit μ with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its post-fit uncertainty $\pm\Delta\hat{\theta}$. The black points show the pulls of the nuisance parameters with respect to their nominal values, θ_0. These pulls and their relative post-fit errors, $(\hat{\theta} - \theta_0)/\Delta\theta$, refer to the lower scale.

For the cut-based $H_5^{\pm}$ GM search, observed and expected exclusion limits at 95% confidence level (CL) on $\sigma \times B(H_5^{\pm} \to WZ)$ and on the mixing parameter $\sin\theta_H$ are shown in Figure 11.20. The intrinsic width of the scalar resonance, for $\sin\theta_H = 0.5$, is narrower than the detector resolution in the mass region explored. The shaded regions show the parameter space for which the $H_5^{\pm}$ width exceeds 5% and 10% of $m_{H_5^{\pm}}$. The ranking plot in Figure 11.21 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for this fit.



(a)



(b)

Figure 11.20: Using the cut-based VBF selection, the observed and expected 95% CL upper limits (a) on the $\sigma \times B(H_5^{\pm} \to WZ)$ and (b) on the parameter $sin\theta_H$ of the GM Model as a function of $m_{H_5^{\pm}}$. The shaded region shows where the theoretical intrinsic width of the resonance would be larger than 5% or 10% of the mass.

The ranking plot in Figure 11.21 shows the post-fit impact of the nuisance parameters on the measured differential signal strength μ for the cut-based VBF signal and control region fit of GM $H_5^{\pm}$ signal of mass 500 GeV. The nuisance parameter with the highest impact on the signal strength is jet flavour uncertainty followed by the ZZ background normalization parameter.

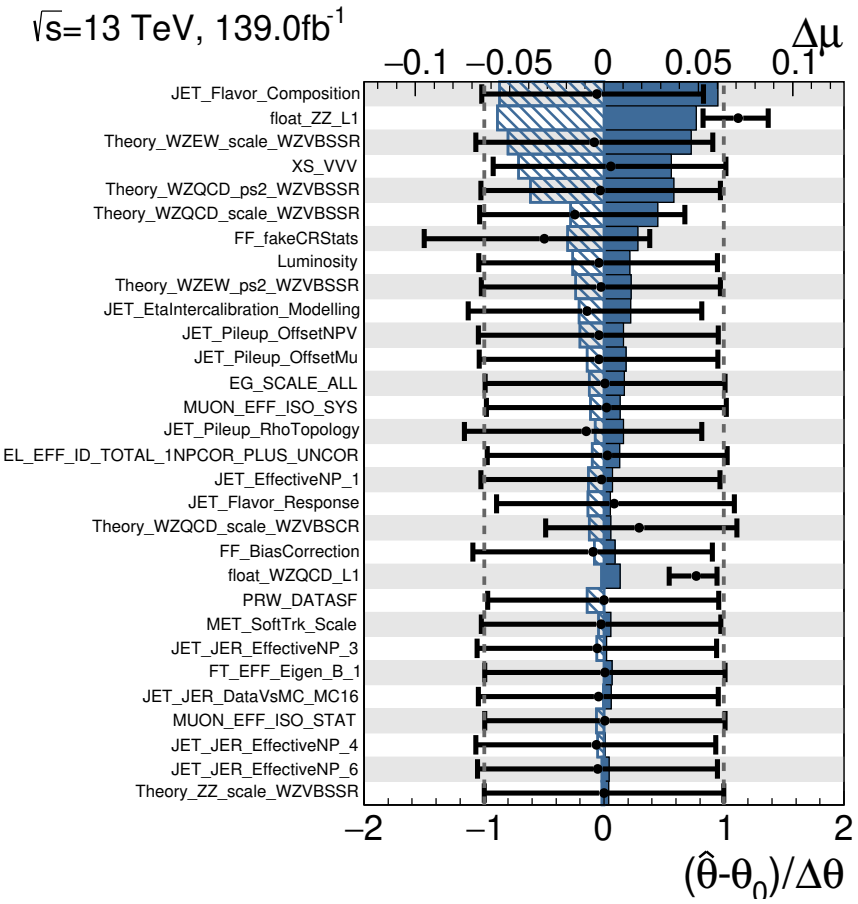


Figure 11.21: Ranking of the nuisance parameters included in the cut-based SR+CR fit according to their post-fit impact on the measured signal strength μ using GM $H_5^{\pm}$ signal of mass 500 GeV and observed data with the production cross-section of the signal equalling the expected 95% CL upper limits. Only the top 30 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The impact of each nuisance parameter $\Delta\mu$, as shown by the blue coloured bars (upper x-axis), are computed by comparing the nominal best-fit μ with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its post-fit uncertainty $\pm\Delta\hat{\theta}$. The black points show the pulls of the nuisance parameters with respect to their nominal values, θ_0. These pulls and their relative post-fit errors, $(\hat{\theta} - \theta_0)/\Delta\theta$, refer to the lower scale.

The expected limits extracted using the cut-based analysis for both models are between 30% and 50% less stringent than the ones extracted using the ANN signal region and therefore justifying the use of ANN in our analysis.

11.5 Impact of systematic uncertainties

The effects of systematic uncertainties on the search are studied for hypothesized signals using a signal-strength parameter μ, which is the ratio of the extracted cross-section to the injected hypothesized signal cross-section. For this study, the signal production cross-section is set to be equal to the expected median upper limits (Section 11.4). The expected relative uncertainties in the best-fit μ value after the maximum-likelihood fit are shown in Table 11.11 for each reference models and mass points: Drell-Yan production of a W' boson in the HVT model with mass $m(W') = 1100$ GeV, and VBF production of a $H_5^{\pm}$ in the GM model with mass $m(H_5^{\pm}) = 375$ GeV as well as W' in the HVT model with $m(W') = 375$ GeV.

The individual sources of systematic uncertainty are combined into fewer background modelling and experimental categories. For signals with higher mass, the data statistical uncertainty is dominant. For VBF production, the uncertainties with the largest impact on the sensitivity of the searches is from jet uncertainties, such as those in the jet energy scale and resolution, followed by the normalization of the irreducible backgrounds WZ-QCD and ZZ. Jet uncertainties are expected to be large since the VBF region requires atleast two jets in the forward region of the ATLAS detector where the coverage of the Inner tracker is limited.. The WZ-QCD and ZZ normalization are limited by the statistics in the control region, but with more data (with Run-3 data) this uncertainty is expected to be reduced. For Drell-Yan production, the dominant uncertainty comes from the sizes of the Monte-Carlo (MC) samples followed by the uncertainties from the theory modelling of the parton shower of the WZ background.

The normalization of the WZ-EWK was not taken from data, here the Scale, PDF and parton shower uncertainties contain shape effects as well as a large normalization component. The modelling of the WZ-EWK was done using leading-order MC, higher accuracy MC modelling will certainty help to reduce this uncertainty. With the observation of WZ-EWK production using the run-2 data, theorist have now a comparison point for their calculations and improvements are expected. The larger data statistics could also allow us to built a WZ-EWK enrich control region. There are also significant contributions from the reducible background shape and normalization, and from the muon and electron identifiction. Uncertainties related to luminosity and pileup also play a relevant role in both signal regions.

Table 11.11: Dominant relative uncertainties in the best-fit signal-strength parameter (μ) for a hypothetical HVT signal of mass $m(W') = 1\,100$ GeV in the Drell-Yan signal region, a GM signal of mass $m(H_5^{\pm}) = 375$ GeV and HVT signal of mass $m(W') = 375$ GeV in the VBF signal region. For this study, the production cross-sections of the signals are set to the expected median upper limits at these two mass values. Uncertainties with smaller contributions are not included.

Source of uncertainty	$\Delta\mu/\mu$ [%]		
	Drell–Yan signal region	VBF signal region	
	$m(W') = 1100$ GeV	$m(H_5^{\pm}) = 375$ GeV	$m(W') = 375$ GeV
WZ-QCD + ZZ normalization	2	11	16
WZ background: parton shower	6	1	4
WZ background: scale, PDF	5	8	8
Fake/non-prompt background	3	1	2
ZZ background: scale, PDF	0.2	<0.1	<0.1
$VVV + t\bar{t}V$ modelling	3	1	1
Electron identification	6	3	4
Muon identification	1	4	5
Jet uncertainty	0.8	16	17
Flavour tagging	0	1	2
Missing transverse momentum	0.2	0.5	0.6
MC statistical uncertainty	10	5	5
Luminosity	2	8	10
Pile-up	0.1	8	7
Total systematic uncertainty	16	22	25
Data statistical uncertainty	54	55	53
Total	56	59	59

11.6 Observed p-Value

The global p-value is calculated by:

$$p_{global} = p_{local} + N_{up} e^{-\frac{1}{2}(Z^2_{local} - Z^2_{ref})} \tag{11.1}$$

where N_{up} is the number of times the observed local p-Value line crosses the reference line,i.e. Z_{ref} which in our case is the 0σ line, in the downward direction.

For Drell-Yan, as shown in Figure 11.22, the W' mass of 1100 GeV shows the most deviation of observed p-value = 0.11 which corresponds to 1.23σ. The calculated global p-value is 1.05 which corresponds to 0σ.

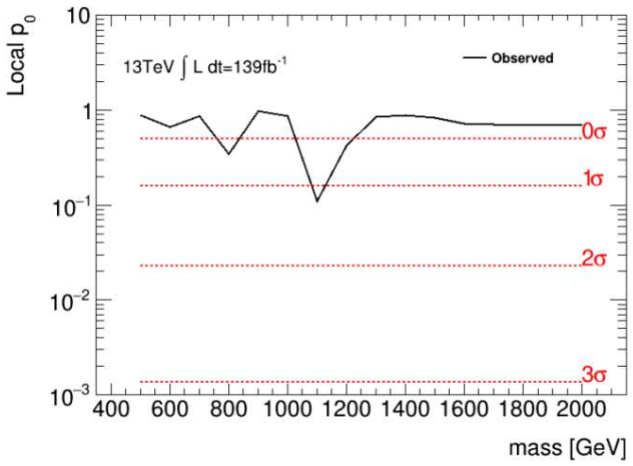


Figure 11.22: Observed local p-value for Drell-Yan signals

For HVT ANN VBF, as shown in Figure 11.23, the W' mass of 375 GeV shows the most deviation of observed p-value = 6.19×10^{-3} which corresponds to 2.5σ. The calculated global p-value is 0.05 which corresponds to 1.65σ.

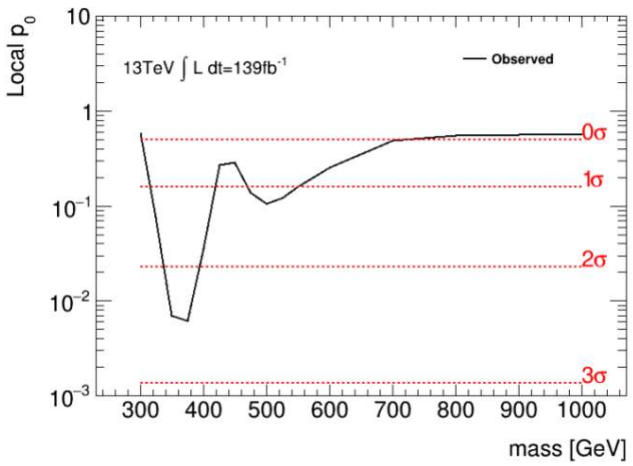


Figure 11.23: Observed local p-value for ANN HVT VBF signals

For GM ANN VBF, as shown in Figure 11.24, the $H_5^{\pm}$ mass of 375 GeV shows the most deviation of observed p-value $= 2.36 \times 10^{-3}$ which corresponds to 2.82σ. The calculated global p-value is 0.06 which corresponds to 1.57σ.

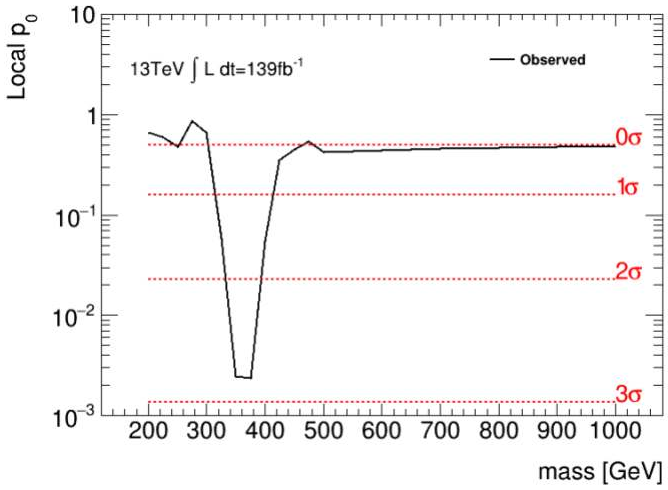


Figure 11.24: Observed local p-value for ANN GM VBF signals

The cut-based VBF analysis doesn't show any excess for both GM and HVT model and is confirmed from the local p-value distribution over GM $H_5^{\pm}$ mass in Figure 11.25.

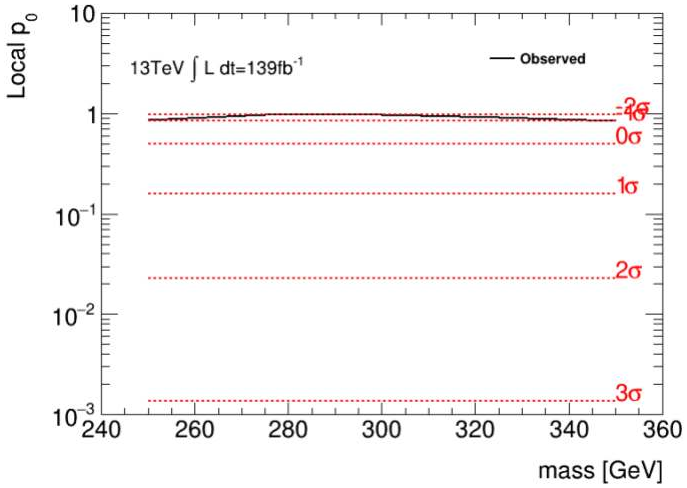


Figure 11.25: Observed local p-value for Cut-based VBF signals

11.7 Toy results

As a test of the asymptotic approximation used in the statistical analysis, limits are also computed with ensembles of pseudo-experiments in all signal and control regions.

The resulting toy limits for all the 3 models are presented. A comparison table is also shown to show the difference in expected toy limits with respect to the Asymptotic limits for Drell-Yan in figures 11.26, and for the ANN models in 11.27 and 11.28 for some of the mass points. The cross-section upper limits obtained by toys agree with the asymptotic in all cases. At higher masses, where event yields become smaller, the discrepancy between the two methods becomes larger, but they remain within 6-10%.

mass	Asymptotic limit	Toy limit	difference
250	0.950181	0.950860	<0.1%
500	0.0623339	0.063620	2%
1000	0.0129439	0.013731	6%
1500	0.00768685	0.008268	7.6%
2000	0.00681217	0.007269	7%
3000	0.0124489	0.013521	8.6%

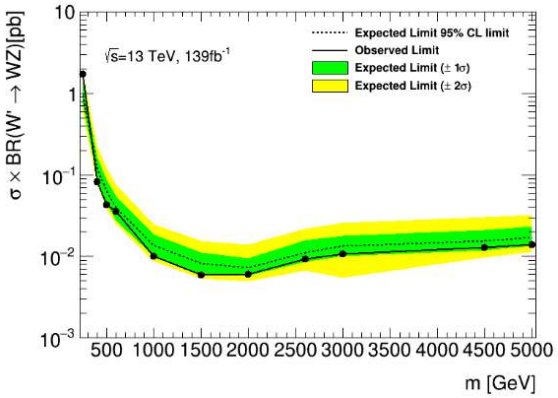


Figure 11.26: Comparison of the expected limits from Asymptotic method and Toys for the HVT Drell-Yan region on the left and expected limits for some of the mass points with Toys on the right. Both of them include the control regions and all the uncertainties

mass	Asymptotic limit	Toy limit	difference
300	0.146758	0.149659	2%
350	0.0962812	0.096781	0.5%
500	0.0492112	0.048951	0.5%
800	0.025146	0.026508	5%
1200	0.0233467	0.023718	1.5%
2000	0.0315885	0.033552	6%

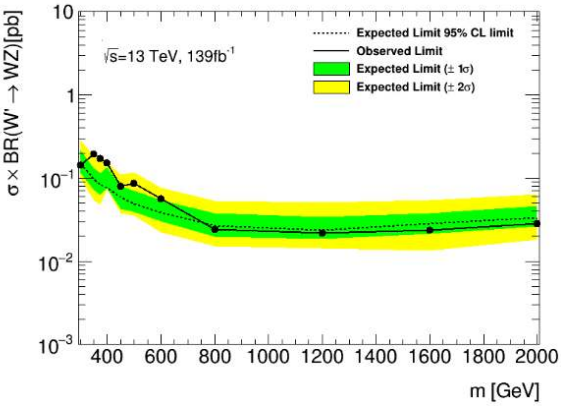


Figure 11.27: Comparison of the expected limits from Asymptotic method and Toys for the HVT VBF ANN region on the left and expected limits for some of the mass points with Toys on the right. Both of them include the control regions and all the uncertainties

mass	Asymptotic limit	Toy limit	difference
200	126.632	128.9771	2%
300	76.7468	79.58783	4%
350	50.8624	52.16574	2.6%
400	36.0504	37.05981	3%
500	26.4627	26.72057	1%
1000	9.00944	9.024011	0.2%

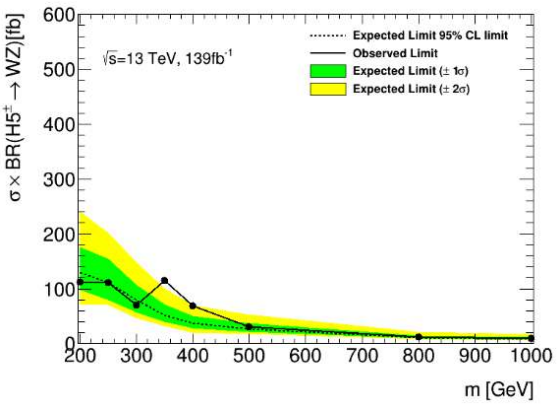


Figure 11.28: Comparison of the expected limits from Asymptotic method and Toys for the GM VBF ANN region on the left and expected limits for some of the mass points with Toys on the right. Both of them include the control regions and all the uncertainties

Chapter 12

Conclusion

"Part of the journey is the end."

Avengers: Endgame

In this thesis a search for resonant WZ production in the fully leptonic final states (electrons and muons) was performed using pp collision data collected by the ATLAS experiment at $\sqrt{s} = 13$ TeV with an integrated luminosity of 139 fb^{-1}. Two different production processes are considered, Drell-Yan, where the resonant particle is produced by quark fusion, and vector-boson fusion (VBF), where a quark and/or anti-quark pair scatters through the exchange of an electroweak gauge boson.

The data in the Drell-Yan selection are found to be consistent with Standard Model predictions. The results are used to derive upper limits at 95% confidence level on the cross-section times branching ratio of the phenomenological Heavy Vector Triplet benchmark Model A and Model B with coupling constant $g_V = 1$ and $g_V = 3$, respectively, as a function of the resonance mass. No evidence of heavy vector resonance production for masses below 2.4 TeV for Model A and 2.5 TeV for Model B was found.

In the case of the VBF production process, limits on the production cross-section times branching ratio to $W^{\pm}Z$ of a hypothetical resonance are obtained as a function of the mass for a heavy vector triplet or for a charged member of the fiveplet scalar in the Georgi-Machacek (GM) model using Artificial Neural Network (ANN) and cut-based is also performed for interpretations in models not considered in this thesis.

The ANN results show a local excess of events over the Standard Model (SM) expectations at a resonance mass of around 375 GeV. The local significances for signals of a heavy vector W' boson or a $H_5^{\pm}$ are 2.5 and 2.8 standard deviations respectively. The respective global significances calculated considering the Look Elsewhere effect are 1.7 and 1.6 standard deviations respectively. With no evidence of heavy vector W' resonance production, masses below 340 GeV, 500 GeV and 700 GeV for HVT VBF production can be excluded for a model with the $g_V c_H$ product of parameters set to 1.0, 2.0 and 3.0, respectively.

In the cut-based VBF results, no excess of events over the SM was observed. With no evidence of heavy vector W' resonance production, masses below 370 GeV, 470 GeV and 590 GeV for HVT VBF production can be excluded for a model with the $g_V c_H$ product of parameters set to 1.0, 2.0 and 3.0,

respectively. As compared to cut-based, the ANN shows improvement in the limits by approximately 40% for the GM model and approximately 30% for the HVT model. The use of ANN has reduced the total background contribution in the signal region by approximately 84%. The main reason for the better limits obtained is the lower background in the signal region due to ANN selection.

In the ATLAS collaboration, several other analyses [128–130] have been published targeting the GM predictions and providing limits on the $\sin(\theta_H)$ parameter. These are shown in Figure 12.1 and summarised in [131], together with our new results. Currently, the limits obtained by the WZ fully-leptonic channel are the most stringent on $\sin(\theta_H)$ over all the mass range.

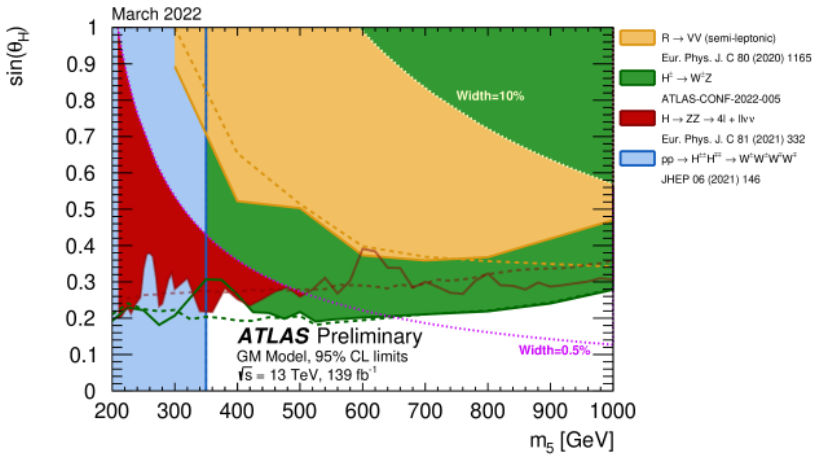


Figure 12.1: Regions of the H5 plane benchmark [92] of the Georgi-Machacek model [30, 132] excluded via direct searches for new, heavy, neutral or charged Higgs bosons. Limits are quoted at 95% CL and are indicated for the data (solid lines) and the expectation from the background model assuming only SM processes (dashed lines). The coloured areas indicate the observed excluded regions in this parameter space. The Higgs bosons that are searched for are produced via VBF for all searches shown here apart from $pp \to H^{\pm\pm}H^{\mp\mp}$ production that proceeds via the Drell-Yan mechanism. The vector-boson fusion calculations are performed with the VBF@NNLO code [133, 134], whereas the Drell-Yan calculation follows Ref. [135]. The branching ratios and the natural widths of the Higgs bosons in the Georgi-Machacek model are calculated using GMCALC [91]. Because the heavy scalar search, $H \to ZZ \to 4l + ll\nu\nu$, and the radion search, $R \to VV$ (semi-leptonic), are valid only for relatively narrow Higgs bosons, dotted lines at 0.5% and 10% denote the reach of the validity of their exclusion areas, respectively. In particular, for the $H \to ZZ \to 4l + ll\nu\nu$ search, the limit on the H5 plane is valid up to about a Higgs boson mass of 500 GeV. The red solid and dashed lines beyond this mass correspond to the observed and expected exclusion, respectively, of the cross-section limits without the width limitation.

CMS has also produced results for the same model with singly or doubly charged Higgs boson decaying to WZ boson pairs or same-sign boson pairs, respectively [136]. CMS has used a cut-based approach for both channels and performed a simultaneous fit is performed with a two-dimensional distribution for the WZ signal region (SR), with 7 bins in the transverse mass of the WZ system (m_T^{WZ}) and 2 bins in the dijet invariant mass distribution (m_{jj}), along with m_{jj} distributions for the nonprompt lepton, tZq and ZZ control regions. No significant excess of events with respect to the standard model background predictions is observed. Despite the larger statitics in CMS analysis, the limits on $\sin(\theta_H)$ obtained by ATLAS using the ANN are better. The limits on the

cross section vs mass of the H^+ are also better in the ATLAS analysis using the ANN.

The cut-based analysis from ATLAS follows closely the selection used by the SM publication [137], CMS analysis in addition to the ATLAS selection cuts includes some extra selections on the lepton Zeppenfeld variable [138] and the hadronic tau veto. After selection, the main background for both is the WZ SM production accounting for approximately 80% of the SR background. The largest difference in background composition comes from the non-prompt background, which in CMS case is about 9% while in ATLAS accounts only for about 2%, even if not dominant the non-prompt lepton background has large uncertainties associated. The ATLAS and CMS expected limits using the cutbased approach agree within one sigma, with ATLAS expected limits in the H^+ cross-section being better than the CMS limits in all the mass spectum. In the case of the $sin(\theta_H)$ limits, CMS exploits the $W^{\pm}W^{\pm}$ and WZ channels, the larger statistics from the two analysis makes the CMS limits sligtly better than the ATLAS cut-based that uses only the WZ channel.

Toys were used as validation of the asymptotic results in this analysis and are found to be in good agreement with a difference of about 6-10% in the high mass range.

No evidence of new physics was found using the WZ resonance search and the Standard Model remains strong. A new chapter for this new physics quest will be written with the Run 3 and High-Luminosity LHC (HL-LHC) data, and maybe by then the SM will show up its limitations to describe WZ data. Various ideas on how to improve the current results in the future are as follows:

- For the search of a heavy resonance that decays $X \to WZ \to l\nu ll$, the acceptance $\times$ efficiency of the analysis decreases for electrons above signal mass 2 TeV. The loss occurs as electrons are boosted and get closer together, when this happens the electrons are removed by the usual electron isolation requirements. Different approaches to recover those events could be tried, such as corrections to isolation cone, or the usage of MVA algorithms dedicated to this cases, in order to reconstruct the two electrons together, similar to "fat jets" [139] these will be "fat electrons" algorithms.
- Use of ANN to create signal regions for both the VBF and Drell-Yan category.
- For the GM model, a combined limits of $W^{\pm}W^{\pm} + WZ$ can be produced following the CMS strategy [136] to get directly a combined fit on the $\sin(\theta_H)$ parameter.
- With more statistics in future, we could probably allow the WZ-EWK component to be a free floating parameter in VBF production fits and then the WZ-EWK contribution be normalized by data.
- With the HL-LHC and the ATLAS upgrades, the High-Granularity Timing Detector (HGTD) [140] and the Inner Tracker [141] (ITk) will allow having better jet measurements in the forward part. This will help pile-jet rejection and a reduction of the systematics associated with the forward jets.

Appendix A

Conventions and Notations

A.1 Natural Units

In order to simplify the notation, it is helpful to use a scheme for units where common variables have simple representations. The customary convention is called "natural units". In this scheme,

$$\hbar = c = 1 \tag{A.1}$$

The unit for energies is not set by this and is commonly written in electronvolt eV. Also masses and momenta can be expressed in this unit.

A.2 Summation Convention

The goal of this convention is to improve notational brevity by removing the Σ operator in special situations. The convention is to skip the Σ operator, when an otherwise undefined index occurs in the superscript and the subscript in a term.

$$\Sigma_{\mu=0}^{3} x_\mu x^\mu =: x_\mu x^\mu \tag{A.2}$$

The allowed range of the index variable can be deduced from the used letter. For indices in the Latin alphabet, the convention is to sum over 1,2, and 3, while indices in the Greek alphabet are summed from 0 to 3.

A.3 Relativistic and Dirac Notation

For relativity, we set $x^0 = t$, $x^1 = x$, $x^2 = y$, $x^3 = z$ and denote the components of the position 4-vector by x^μ, with a Greek index. The components of spatial 3-vectors will be denoted by Latin indices, e.g. $x^i = (x, y, z)$.

We use the usual metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{A.3}$$

resulting in the relativistic derivative ∂^μ:

$$\partial^\mu = \left(\frac{\partial}{\partial t}; -\frac{\partial}{\partial x}; -\frac{\partial}{\partial y}; -\frac{\partial}{\partial z}\right) \tag{A.4}$$

Gamma matrices γ^μ are used similar to Weyl basis, with adjusted sign convention

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{A.5}$$

where σ_i are the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.6}$$

The γ^5 is given by:

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{A.7}$$

Appendix B

GM simulation

B.1 GM LO simulation

Information about the LO samples are summarised in Tables B.1 and B.2 the nominal PDF set used is `NNPDF30_nlo_as_0119_atlas`. As these samples contain only electronic and muonic final states, additional samples were the W decays to a tau and a tau neutrino were produced. Informations about these samples are summarised in Table B.2. For these sampled a minimum p_T of 15 GeV (10 GeV) of the jets (leptons) are required at generator level as well as $|\eta|$ <5 for jets and $|\eta|$ <2.7 for leptons are required.

DSID	Mass	Events	Filter eff.	Cross-section [fb]	k-factor
305028	200	40000	1.00	7.0596	1.00
309501	250	40000	0.77361	7.71	1.00
305029	300	40000	1.00	3.9238	1.00
309501	350	40000	0.77472	4.582	1.00
305030	400	40000	1.00	2.4428	1.00
309501	450	40000	0.78052	3.275	1.00
305031	500	40000	1.00	1.6113	1.00
305032	600	40000	1.00	1.1005	1.00
305033	700	40000	1.00	0.77398	1.00
305034	800	40000	1.00	0.55433	1.00
305035	900	40000	1.00	0.40394	1.00

Table B.1: Summary of the LO GM signal MC simulations with electron and muon final states for the VBF analysis. The mass points 250, 350 and 450 GeV also contain the $ll\tau\nu_\tau$ final states. These samples are not used in the analysis.

DSID	Mass	Events	Filter eff.	Cross-section [fb]	k-factor
307181	200	45000	0.3296	3.521	1.00
307182	300	42000	0.3486	2.149	1.00
307183	400	42000	0.3544	1.356	1.00
307184	500	45000	0.3495	0.802	1.00
307185	600	45000	0.3541	0.5888	1.00
307186	700	45000	0.3613	0.3853	1.00
307187	800	44000	0.3582	0.2775	1.00
307188	900	45000	0.3581	0.2570	1.00

Table B.2: Summary of the LO GM signal MC simulations with $ll\tau\nu_\tau$ final states for the VBS analysis. These samples are not used in the analysis.

Appendix C

WZ-EWK background in VBF signal region fits

The WZ-EWK background is the second largest background in both of the VBF signal regions, accounting for ∼20% in the cut-based VBF signal region and ∼30% for the VBF ANN signal region.

Given this large contribution in the signal region the possibility of including a free normalization factor for the WZ-EWK contribution (μ_{WZEWK}) was studied, by performing an Asimov CR+SR fit and comparing the obtained uncertainties on the WZ-EWK prediction with the pre-fit theory uncertainty of ∼20% (5% PDF, 10% scale and 15% PS).

The background only Asimov fit results in the VBF cut-based signal region is shown in table C.1. The fit includes the inclusive ZZ control region, the WZ-QCD VBF control region and the cut-based VBF ignal region. The following three free normalization parameters are part of the fit the μ_{WZEWK}, μ_{WZQCD} and μ_{ZZ}. It can be seen that the uncertainty on the WZ EWK background is increased to ∼30% when including the μ_{WZEWK} as a free parameter in the signal region, so it is decided that the WZ EWK will be taken from simulation and only allow to move within uncertainties.

	VBF cut-based SR
ZZ	24.9 ± 3.4
Fakes	14.8 ± 3.2
VVV	1.0 ± 0.23
WZ-EWK	104.8 ± 28.8
WZ-QCD	292.5 ± 17.3
$t\bar{t}V$	29.4 ± 4.2
Postfit	467.5 ± 16.0

Table C.1: Background only asimov fits performed using the cut-based VBF signal region. The fit includes the inclusive ZZ CR, the WZ-QCD VBF CR and the VBF SR and three free normalization parameters the μ_{WZEWK}, μ_{WZQCD} and μ_{ZZ}.

Appendix D

ZZ background in VBF signal region fits

The ZZ VBF region defined in Section 8.1.3 could be either used as a Validation Region (VR) to verify that modelling and the assigned uncertainties to the ZZ background in events with two or more jets are reasonable, or used as a Control Region (CR) to constrain the ZZ background in the VBF signal region fits. To decide on the approach to use the following two Asimov fits were performed:

- Option 1 ($ZZJJ$ as control region): Two free-floating normalization parameters ZZ normalization (μ_{ZZ}) and WZ normalization (μ_{WZQCD}) in three regions, the $ZZJJ$, WZ VBF CR and the VBF signal region (either VBF cut-based or ANN).
- Option 2 ($ZZJJ$ not part of the fit): One free-floating normalization parameter for the WZ normalization (μ_{WZQCD}) in two regions, WZ VBF CR and the VBF signal region (either VBF cut-based or ANN). The ZZ background is then allowed to move only within the theory and experimental uncertainties.

The background only Asimov fit results for both options discussed above are shown using the VBF cut-based signal region in table D.1 and for the ANN signal region in table D.2. In both signal regions, the uncertainty on the ZZ background is reduced when including the $ZZJJ$ as a control region, so it is decided that the $ZZJJ$ will be part of the combined VBF fit.

	VBF cut-based SR	
	Option 1 (including $ZZJJ$ CR)	Option 2
ZZ	25.5 ± 2.2	25.5 ± 4.1
Fakes	15.7 ± 3.1	15.7 ± 3.1
VVV	1.0 ± 0.23	1.0 ± 0.23
WZ EW	106.5 ± 15.8	106.5 ± 15.8
WZ QCD	298.8 ± 15.6	298.8 ± 14.5
$t\bar{t}V$	29.5 ± 4.2	29.5 ± 4.2
Postfit	477.1 ± 16.4	477.1 ± 16.5

Table D.1: Background only asimov fits performed using the cut-based VBF signal region with and w.o the $ZZJJ$ as a Control Region. The fit that includes the $ZZJJ$ CR and the μ_{ZZ} as free parameter has a better ZZ post-fit uncertainty than the SR only fit, we decide to use the $ZZJJ$ as a control region. .

	VBF ANN SR	
	Option 1 (including $ZZJJ$ CR)	Option 2
ZZ	6.91 ± 0.83	6.91 ± 1.29
Fakes	0.0 ± 0.0	0.0 ± 0.0
VVV	0.24 ± 0.06	0.24 ± 0.06
WZ EW	31.5 ± 1.5	31.5 ± 1.4
WZ QCD	50.2 ± 4.8	50.2 ± 4.5
$t\bar{t}V$	0.71 ± 0.11	0.71 ± 0.11
Postfit	89.6 ± 4.4	89.6 ± 4.5

Table D.2: Background only asimov fits performed using the ANN VBF signal region with and w.o the $ZZJJ$ as a Control Region. The fit that includes the $ZZJJ$ CR and the ZZ as free parameter has a better post-fit uncertainty than the SR only fit, so we decide to use the $ZZJJ$ as a control region. .

Appendix E

Limit extraction in Drell-Yan region split by decay channels

The Drell-Yan cut-based signal region is the analysis region containing the largest expected data statistics. For this region the possibility of using four categories depending on the bosons decay channels was considered. Figure E.1 shows the background distribution in the four decay channels from left to right $\mu\mu\mu\nu$, $\mu\mu e\nu$, $ee\mu\nu$ and $eee\nu$, the background composition is very similar across all channels. The $Z + e$ channels having an slightly larger background contribution due to Z+jets events. Figure E.2 shows the invariant mass distribution of the WZ pair in the four different channels, shapes are compatible across the WZ decay channels.

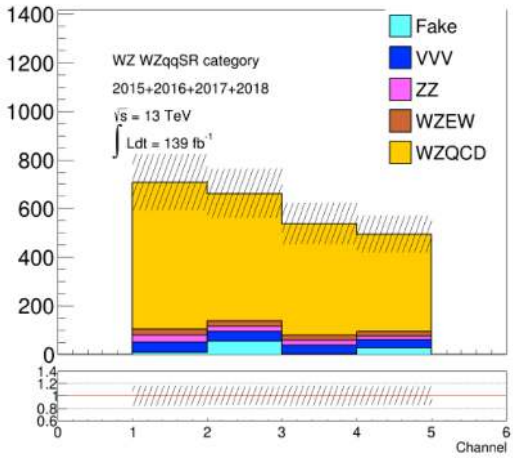


Figure E.1: Yield in the Drell-Yan signal region split by decay channels, where channel 1 is $\mu\mu\mu$, channel 2 is $\mu\mu e$, channel 3 is $ee\mu$ and channel 4 is eee

Limits were also extracted using as input the four WZ invariant mass distributions. The ZZ and WZ QCD control regions were as well binned per decay channel. Figure E.3 shows the extracted limits using all the systematic uncertainties split per channel.

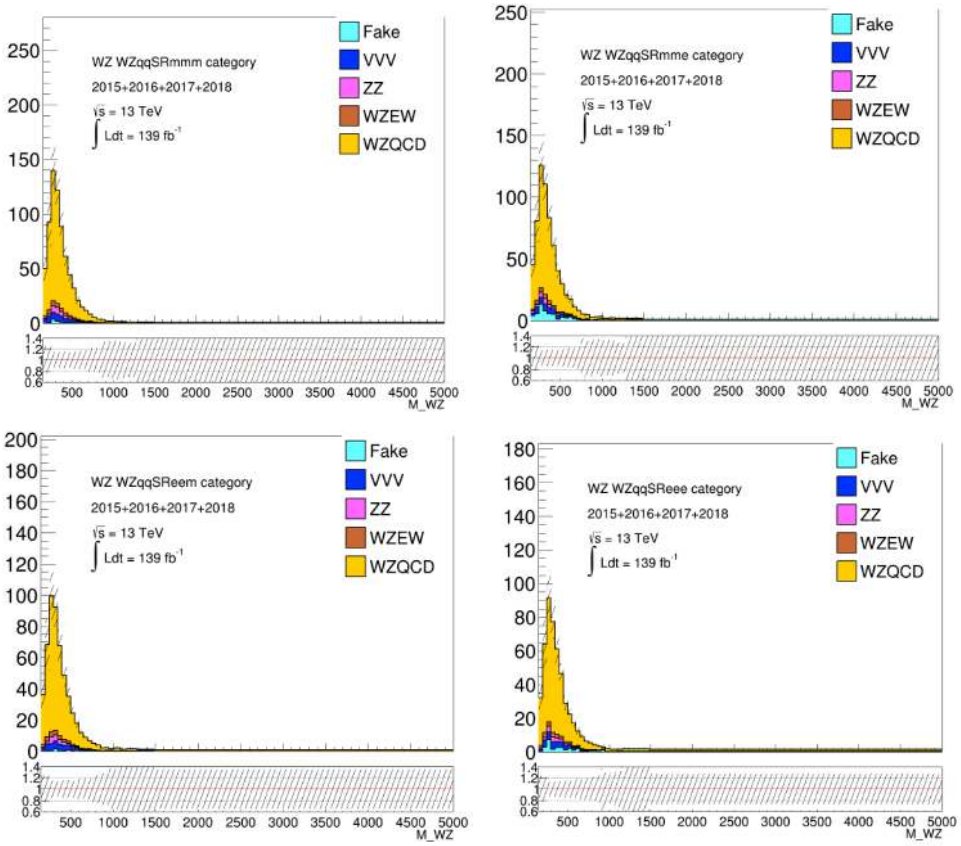


Figure E.2: Invariant mass of the WZ system reconstructed in the $\mu\mu\mu\nu$ top left, $\mu\mu e\nu$ top right, $ee\mu\nu$ bottom left and $eee\nu$ bottom right

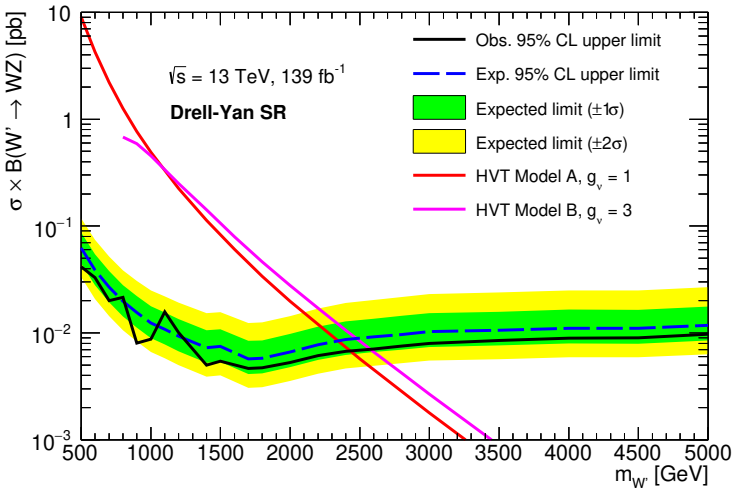


Figure E.3: Expected exclusion limits, using the Drell-Yan signal region split per channel, on $\sigma \times \mathrm{BR}(W' \to WZ)$ signal at 95% C.L confidence level as a function of m_{WZ}.

Figure E.4 shows the comparison of the expected exclusion limits extracted in one case using the total M_{WZ} distribution and in the other case using the M_{WZ} split per decay channel. The comparison plot shows similar performance for both analysis strategies at low mass up to 2 TeV. The visible differences are related to the last bin of the M_{WZ} [1480,inf] TeV the bin with the lowest expected events.

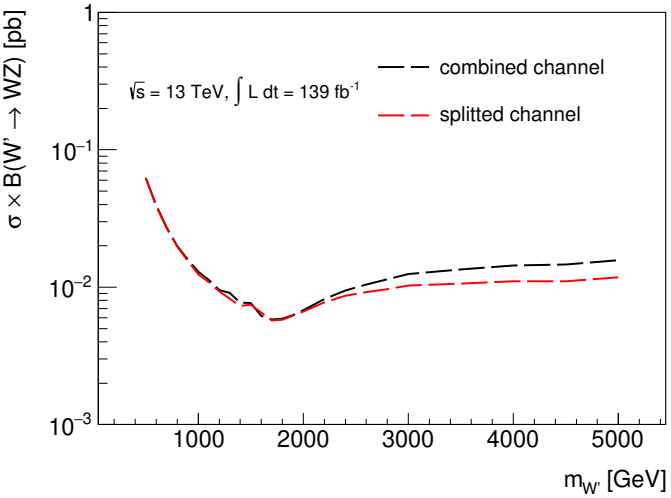


Figure E.4: Comparison of the expected exclusion limits extracted in one case using the total M_{WZ} distribution and in the other case using the M_{WZ} split per decay channel.

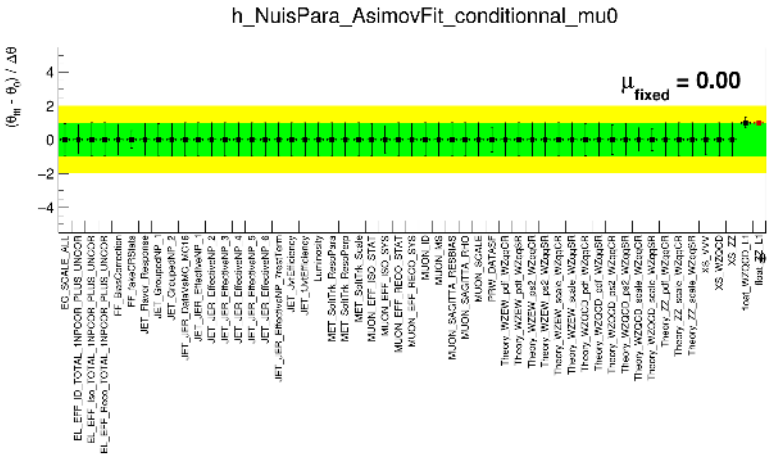


Figure E.5: Nuisance pull plot for background only Asimov fit with all SR and CR split by decay channels

List of Figures

List of Tables

Bibliography

[1] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor (cit. on pp. 1, 10).

[2] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor (cit. on pp. 1, 10).

[3] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 1).

[4] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 1).

[5] J. D. Dowell, *The CERN proton-antiproton collider and the discovery of the W and Z particles*, Sci. Prog. **69** (1984) 257 (cit. on p. 2).

[6] ATLAS Collaboration, *Search for WZ resonances in the fully leptonic channel using pp collisions at sqrt(s) = 8 TeV with the ATLAS detector*, Phys. Lett. B **737** (2014) 223, arXiv: 1406.4456 [hep-ex] (cit. on p. 2).

[7] CMS Collaboration, *Search for New Resonances Decaying via WZ to Leptons in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **740** (2015) 83, arXiv: 1407.3476 [hep-ex] (cit. on p. 2).

[8] ATLAS Collaboration, *Search for resonant WZ production in the fully leptonic final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **787** (2018) 68, arXiv: 1806.01532 [hep-ex] (cit. on pp. 2, 86).

[9] Michael E. Peskin and Daniel V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, 1995, ISBN: 978-0-201-50397-5 (cit. on p. 3).

[10] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. B **47** (1973) 365 (cit. on p. 4).

[11] David J. Gross and Frank Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343, ed. by J. C. Taylor (cit. on p. 4).

[12] H. David Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346, ed. by J. C. Taylor (cit. on p. 4).

[13] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on pp. 4, 8).

[14] Steven Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on p. 4).

[15] Abdus Salam and John Clive Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168 (cit. on p. 4).

[16] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley (1984) (cit. on p. 6).

[17] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press (1995) (cit. on p. 6).

[18] Julian S. Schwinger, *A Theory of the Fundamental Interactions*, Annals Phys. **2** (1957) 407 (cit. on p. 8).

[19] Chen-Ning Yang and Robert L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191, ed. by Jong-Ping Hsu and D. Fine (cit. on p. 8).

[20] A. Pich, "The Standard model of electroweak interactions", *2004 European School of High-Energy Physics*, 2005 1, arXiv: hep-ph/0502010 (cit. on pp. 9, 11).

[21] Peter W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132 (cit. on p. 10).

[22] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585, ed. by J. C. Taylor (cit. on p. 10).

[23] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965 (cit. on p. 10).

[24] Yoichiro Nambu, *Quasiparticles and Gauge Invariance in the Theory of Superconductivity*, Phys. Rev. **117** (1960) 648, ed. by J. C. Taylor (cit. on p. 10).

[25] Kaustubh Agashe, Roberto Contino, and Alex Pomarol, *The Minimal composite Higgs model*, Nucl. Phys. B **719** (2005) 165, arXiv: hep-ph/0412089 (cit. on p. 13).

[26] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, *Heavy Vector Triplets: Bridging Theory and Data*, (2014), arXiv: 1402.4431 [hep-ph] (cit. on pp. 15–17, 43, 44).

[27] Vernon D. Barger, Wai-Yee Keung, and Ernest Ma, *A Gauge Model With Light W and Z Bosons*, Phys.Rev. **D22** (1980) 727 (cit. on p. 17).

[28] Roberto Contino, David Marzocca, Duccio Pappadopulo, and Riccardo Rattazzi, *On the effect of resonances in composite Higgs phenomenology*, JHEP **1110** (2011) 081, arXiv: 1109.1570 [hep-ph] (cit. on p. 17).

[29] Xiao Kang Du, Zhuang Li, Fei Wang, and Ying Kai Zhang, *Explaining The New CDF II W-Boson Mass Data In The Georgi-Machacek Extension Models*, (2022), arXiv: 2204.05760 [hep-ph] (cit. on p. 18).

[30] Howard Georgi and Marie Machacek, *DOUBLY CHARGED HIGGS BOSONS*, Nucl. Phys. B **262** (1985) 463 (cit. on pp. 18, 130).

[31] Katy Hartling, Kunal Kumar, and Heather E. Logan, *Indirect constraints on the Georgi-Machacek model and implications for Higgs boson couplings*, Phys. Rev. D **91** (2015) 015013, arXiv: 1410.5538 [hep-ph] (cit. on p. 18).

[32] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016), arXiv: 1610.07922 [hep-ph] (cit. on p. 20).

[33] Heather E. Logan and Mark B. Reimer, *Characterizing a benchmark scenario for heavy Higgs boson searches in the Georgi-Machacek model*, Phys. Rev. D **96** (2017) 095029, arXiv: 1709.01883 [hep-ph] (cit. on pp. 20, 45).

[34] CERN Accelerating science, URL: https://home.cern (cit. on p. 21).

[35] Julie Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, (2013), General Photo, URL: https://cds.cern.ch/record/1621894 (cit. on p. 22).

[36] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 22).

[37] A. Augusto Alves Jr. et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005 (cit. on p. 23).

[38] K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002 (cit. on p. 23).

[39] G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007 (cit. on p. 23).

[40] O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006 (cit. on p. 23).

[41] James Pinfold et al., *Technical Design Report of the MoEDAL Experiment*, (2009) (cit. on p. 23).

[42] *LHC sets new world record*, URL: https://press.cern/news/press-release/cern/lhc-sets-new-world-record (cit. on p. 23).

[43] *CERN achieves 7 TeV collisions at Large Hadron Collider*, URL: https://physicsworld.com/a/cern-achieves-7-tev-collisions-at-large-hadron-collider/ (cit. on p. 23).

[44] M. Lamont, *Status of the Large Hadron Collider*, J. Phys. Conf. Ser. **347** (2012) 012001, ed. by Serkant Ali Cetin, Peter Jenni, Veysi Erkcan Ozcan, and Vedat Nefer Senoguz (cit. on p. 23).

[45] *LHC Season 2: First physics at 13 TeV to start tomorrow*, URL: https://home.cern/news/news/accelerators/lhc-season-2-first-physics-13-tev-start-tomorrow (cit. on p. 23).

[46] *LHC performance reaches new highs*, URL: https://home.cern/news/news/accelerators/lhc-performance-reaches-new-highs (cit. on p. 23).

[47] *Record luminosity: well done LHC*, URL: https://home.cern/news/news/accelerators/record-luminosity-well-done-lhc (cit. on p. 23).

[48] *LHC Report: Another run is over and LS2 has just begun...*, URL: https://home.cern/news/news/accelerators/lhc-report-another-run-over-and-ls2-has-just-begun?utm_source=Bulletin&utm_medium=Email&utm_content=2018-12-12E&utm_campaign=BulletinEmail (cit. on p. 23).

[49] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003. 437 p (cit. on pp. 23, 24, 26, 28–30).

[50] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex] (cit. on pp. 23, 25–27).

[51] Joany Manjarres Ramos, "WZ Diboson Measurements with the ATLAS experiment at the LHC and Performance of resistive Micromegas in view of HL-LHC applications", Presented 08 Jul 2013, 2013, URL: https://cds.cern.ch/record/1626587 (cit. on pp. 25, 83).

[52] ATLAS Collaboration, *Study of the material of the ATLAS inner detector for Run 2 of the LHC*, JINST **12** (2017) P12009, arXiv: 1707.02826 [hep-ex] (cit. on pp. 25–27).

[53] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, (2010) (cit. on p. 25).

[54] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, (1997) (cit. on p. 28).

[55] Magnet System, URL: https://atlas.cern/Discover/Detector/Magnet-System (cit. on p. 29).

[56] ATLAS Collaboration, *ATLAS central solenoid: Technical design report*, (1997) (cit. on p. 29).

[57] A. Yamamoto et al., *The ATLAS central solenoid*, Nucl. Instrum. Meth. A **584** (2008) 53 (cit. on p. 29).

[58] ATLAS Collaboration, *ATLAS barrel toroid: Technical design report*, (1997) (cit. on p. 29).

[59] ATLAS Collaboration, *ATLAS endcap toroids: Technical design report*, (1997) (cit. on p. 29).

[60] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C **72** (2012) 1849, arXiv: 1110.1530 [hep-ex] (cit. on p. 30).

[61] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 30).

[62] 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** (2019) 639, arXiv: 1902.04655 [physics.ins-det] (cit. on p. 31).

[63] W Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROU LARG-PUB-2008-002: CERN, 2008, URL: https://cds.cern.ch/record/1099735 (cit. on p. 31).

[64] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex] (cit. on p. 31).

[65] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV*, (2016) (cit. on pp. 32, 33).

[66] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, JINST **14** (2019) P03017, arXiv: 1812.03848 [hep-ex] (cit. on p. 33).

[67] ATLAS Collaboration, *Muon reconstruction performance in early $\sqrt{s}$ = 13 TeV data*, (2015) (cit. on p. 33).

[68] Priyanka Mukhopadhyay and Bidyut B. Chaudhuri, *A survey of Hough Transform*, Pattern Recognition **48** (2015) 993, ISSN: 0031-3203, URL: https://www.sciencedirect.com/science/article/pii/S0031320314003446 (cit. on p. 33).

[69] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on p. 34).

[70] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on pp. 35, 76).

[71] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 35).

[72] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv: 2007.02645 [hep-ex] (cit. on p. 36).

[73] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 36).

[74] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, (2013) (cit. on p. 36).

[75] ATLAS Collaboration, *Tagging and suppression of pileup jets*, (2014) (cit. on p. 36).

[76] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2304, arXiv: 1112.6426 [hep-ex] (cit. on p. 37).

[77] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, (2017) (cit. on pp. 37, 38).

[78] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, (2015) (cit. on p. 37).

[79] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex] (cit. on pp. 38, 77).

[80] URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/GoodRunsListsTutorial (cit. on p. 41).

[81] URL: http://atlasdqm.web.cern.ch/atlasdqm/grlgen/ (cit. on p. 41).

[82] URL: https://atlas-datasummary.cern.ch/lumicalc (cit. on p. 41).

[83] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex] (cit. on p. 41).

[84] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [physics.ins-det] (cit. on p. 41).

[85] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 42).

[86] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 42).

[87] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph] (cit. on p. 42).

[88] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 42).

[89] *ATLAS Pythia 8 tunes to 7 TeV data*, (2014) (cit. on pp. 43, 47).

[90] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 44).

[91] Katy Hartling, Kunal Kumar, and Heather E. Logan, *GMCALC: a calculator for the Georgi-Machacek model*, (2014), arXiv: 1412.7387 [hep-ph] (cit. on pp. 44, 130).

[92] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016), arXiv: 1610.07922 [hep-ph] (cit. on pp. 44, 130).

[93] Richard D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on pp. 45, 46).

[94] Alessandro Ballestrero et al., *Precise predictions for same-sign W-boson scattering at the LHC*, Eur. Phys. J. C **78** (2018) 671, arXiv: 1803.07943 [hep-ph] (cit. on p. 45).

[95] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 45).

[96] Rikkert Frederix and Stefano Frixione, *Merging meets matching in MC@NLO*, JHEP **12** (2012) 061, arXiv: 1209.6215 [hep-ph] (cit. on p. 45).

[97] Enrico Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 46).

[98] Paolo Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 47).

[99] Stefano Frixione, Paolo Nason, and Carlo Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 47).

[100] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 47).

[101] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, *NLO vector-boson production matched with shower in POWHEG*, JHEP **07** (2008) 060, arXiv: 0805.4802 [hep-ph] (cit. on p. 47).

[102] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 (cit. on p. 47).

[103] JetVertexTagger, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTagger (cit. on p. 51).

[104] Overlap Removal Tools, URL: https://indico.cern.ch/event/631313/contributions/2683959/attachments/1518878/2373377/Farrell_ORTools_ftaghbb.pdf (cit. on p. 52).

[105] Overlap Removal Tools git repository, URL: https://gitlab.cern.ch/atlas/athena/tree/21.2/PhysicsAnalysis/AnalysisCommon/AssociationUtils (cit. on p. 52).

[106] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, tech. rep. ATLAS-COM-CONF-2015-024, CERN, 2015, URL: https://cds.cern.ch/record/2016323 (cit. on p. 52).

[107] Jet cleaning recommendations, URL: %7Bhttps://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets2015%7D (cit. on p. 52).

[108] Francois Chollet et al., *Keras*, 2015, URL: https://github.com/fchollet/keras (cit. on p. 58).

[109] Martín Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, Software available from tensorflow.org, 2015, URL: https://www.tensorflow.org/ (cit. on p. 58).

[110] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton, "On the importance of initialization and momentum in deep learning", *Proceedings of the 30th International Conference on Machine Learning*, ed. by Sanjoy Dasgupta and David McAllester, vol. 28, Proceedings of Machine Learning Research 3, PMLR, 2013 1139 (cit. on p. 60).

[111] Dennis Kraft, *Signal selection optimization for the Vector Fusion Category in the $WZ \to lvll$ resonance search analysis performed in ATLAS at the LHC*, (), URL: cds.cern.ch/record/2779839/files/CERN-THESIS-2021-118.pdf (cit. on p. 64).

[112] ATLAS Collaboration, *Measurements of $W^{\pm}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings*, Phys. Rev. D **93** (2016) 092004, arXiv: 1603.02151 [hep-ex] (cit. on p. 68).

[113] Jon Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph] (cit. on p. 75).

[114] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, tech. rep., CERN, 2016, URL: http://cds.cern.ch/record/2119986 (cit. on p. 75).

[115] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex] (cit. on p. 75).

[116] ATLAS Collaboration, *Evidence for the production of three massive vector bosons with the ATLAS detector*, Phys. Lett. B **798** (2019) 134913, arXiv: 1903.10415 [hep-ex] (cit. on p. 75).

[117] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep., CERN, 2019, URL: http://cds.cern.ch/record/2677054 (cit. on p. 76).

[118] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, arXiv: 1606.02625 [hep-ex] (cit. on p. 76).

[119] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex] (cit. on p. 76).

[120] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on p. 76).

[121] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012), URL: https://cds.cern.ch/record/1456844 (cit. on pp. 79, 80).

[122] Wouter Verkerke and David Kirkby, *The RooFit toolkit for data modeling*, 2003, URL: https://arxiv.org/abs/physics/0306116 (cit. on p. 80).

[123] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, [Erratum: Eur. Phys. J. C73, 2501 (2013)], arXiv: 1007.1727 [physics.data-an] (cit. on pp. 80, 81, 112).

[124] Alexander L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693 (cit. on p. 81).

[125] Frequentist Limit Recommendation, URL: https://twiki.cern.ch/twiki/pub/AtlasProtected/StatisticsTools/Frequentist_Limit_Recommendation.pdf (cit. on p. 82).

[126] L. Moneta et al., "The RooStats Project", *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, ACAT2010*, Proceedings of Science, 2010, arXiv: 1009.1003 (cit. on p. 83).

[127] ATLAS Collaboration, *Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **793** (2019) 469, arXiv: 1812.09740 [hep-ex] (cit. on p. 89).

[128] ATLAS Collaboration, *Search for heavy diboson resonances in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1165, arXiv: 2004.14636 [hep-ex] (cit. on p. 130).

[129] ATLAS Collaboration, *Search for heavy resonances decaying into a pair of Z bosons in the $\ell^+\ell^-\ell'^+\ell'^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using 139 fb^{-1} of proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 332, arXiv: 2009.14791 [hep-ex] (cit. on p. 130).

[130] ATLAS Collaboration, *Search for doubly and singly charged Higgs bosons decaying into vector bosons in multi-lepton final states with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **06** (2021) 146, arXiv: 2101.11961 [hep-ex] (cit. on p. 130).

[131] ATLAS Collaboration, *Summary plot for the interpretation of heavy Higgs boson searches in the Georgi-Machacek model*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-008: CERN, 2022, URL: http://cds.cern.ch/record/2803996 (cit. on p. 130).

[132] Michael S. Chanowitz and Mitchell Golden, *Higgs boson triplets with Mw=Mzcos w*, Physics Letters B **165** (1985) 105, ISSN: 0370-2693, URL: https://www.sciencedirect.com/science/article/pii/0370269385907002 (cit. on p. 130).

[133] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro, *Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **105** (1 2010) 011801, URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.011801 (cit. on p. 130).

[134] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro, *Vector boson fusion at next-to-next-to-leading order in QCD: Standard model Higgs boson and beyond*, Phys. Rev. D **85** (3 2012) 035002, URL: https://link.aps.org/doi/10.1103/PhysRevD.85.035002 (cit. on p. 130).

[135] Margarete Mühlleitner and Michael Spira, *Note on doubly charged Higgs boson pair production at hadron colliders*, Phys. Rev. D **68** (11 2003) 117701, URL: https://link.aps.org/doi/10.1103/PhysRevD.68.117701 (cit. on p. 130).

[136] CMS Collaboration, *Search for charged Higgs bosons produced in vector boson fusion processes and decaying into vector boson pairs in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 723, arXiv: 2104.04762 [hep-ex] (cit. on pp. 130, 131).

[137] ATLAS Collaboration, *Observation of electroweak* $W^{\pm}Z$ *boson pair production in association with two jets in pp collisions at* $\sqrt{s} = 13$*TeV with the ATLAS Detector*, (2018) (cit. on p. 131).

[138] David L. Rainwater, R. Szalapski, and D. Zeppenfeld, *Probing color singlet exchange in Z + two jet events at the CERN LHC*, Phys. Rev. D **54** (1996) 6680, arXiv: hep-ph/9605444 (cit. on p. 131).

[139] Davison E. Soper and Michael Spannowsky, *Combining subjet algorithms to enhance ZH detection at the LHC*, JHEP **08** (2010) 029, arXiv: 1005.0417 [hep-ph] (cit. on p. 131).

[140] ATLAS Collaboration, *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*, tech. rep., CERN, 2020, URL: https://cds.cern.ch/record/2719855 (cit. on p. 131).

[141] Lingxin Meng, "ATLAS ITk Pixel Detector Overview", *International Workshop on Future Linear Colliders*, 2021, arXiv: 2105.10367 [physics.ins-det] (cit. on p. 131).

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation wurde in der Zeit von August 2017 bis August 2022 im Institut für Kern- und Teilchenphysik unter der wissenschaftlichen Betreuung von Prof. Dr. Michael Kobel angefertigt.

Ich erkenne die Promotionsordnung des Bereichs Mathematik und Naturwissenschaften an der Technischen Universität Dresden von 23.02.2017 an.

Abhishek Nag
Dresden, 11.08.2022