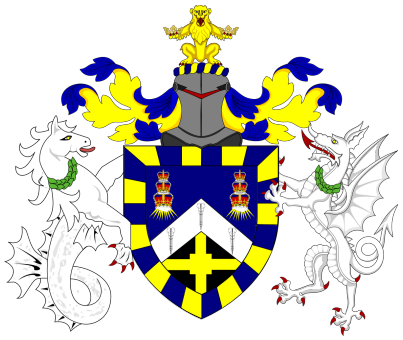
Search for heavy resonances decaying into a Z boson and a Higgs boson with the ATLAS detector



Tong Qiu

Submitted in partial fulfillment of the requirements

of the Degree of Doctor of Philosophy

School of Physical and Chemical Sciences

Queen Mary University of London

September 2022

CERN-THESIS-2022-309
26/09/2022

Statement of originality

I, Tong Qiu, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: Tong Qiu

Date: 08 Aug 2022

Details of collaboration and publications:

The results presented in this thesis is produced in collaboration with the ATLAS collaboration. The work has also been published in [1] and [2].

Abstract

The thesis presents a search for a heavy resonance decaying into a Z boson and a Standard-Model-like Higgs boson (h), with the Z boson decaying into two leptons and the Higgs boson decaying into two b-quarks using data collected from 2015 to 2018 by the ATLAS detector corresponding to an integrated luminosity of 139 fb^{-1}. The search probes the reconstructed invariant or transverse mass distributions of the Zh candidates in a mass range from 220 GeV to 5 TeV. The result of the search is compatible with the Standard-Model prediction, and upper limits at the 95% CL are set on the gluon-gluon fusion and the b-associated production cross sections of a pseudoscalar Higgs boson in the Two-Higgs-Doublet Models and the Drell-Yan production cross sections of a heavy vector boson in the Heavy-Vector-Triplet models. The limits on cross sections are interpreted as constraints on the parameters of various Two-Higgs-Doublet Models. Furthermore, this thesis presents a study of a new analysis approach which uses neural networks to improve the sensitivity of the search. In addition to the resonance search, the thesis also discusses the offline simulation of the ATLAS Level-1 calorimeter trigger.

Acknowledgements

Firstly, I would like to thank my supervisor, Ulla Blumenschein, for guiding me through my PhD. She has patiently answered all my questions and provided detailed feedback on my work. I am grateful for the great opportunities she gave me during my PhD. My research would not be possible without her support.

I would like to thank Dominik Duda, Carlo Dallapiccola, Wade Fisher, Garabed Halladjian, Andreas Honle, Sandra Kortner, Tong Li, Lianliang Ma, Stefan Maschek, Nora Pettersson and Makayla Vessella. It has been my pleasure to work with the ATLAS VH semileptonic analysis team during the past four years. I miss our weekly meetings where we had lots of fruitful discussions. The completion of this thesis would not be possible without collaboration with this team.

I would like to thank members of the ATLAS Round-2 VV+VH semileptonic analysis team, especially Robert Les, Yassine El Ghazali, Maria Mazza, Alessandra Palazzo, Gabriele Chiodini, Antonio Giannini. I enjoyed working in the new analysis team at the end of my PhD and the few in-person meetings we had during my short stay at CERN.

I would like to thank members of the ATLAS L1Calo software teams, especially Alison Elliot, Marcella Bona, Jesal Mandalia, Antonio Costa, Cecilia Tosciri, Harry Cook, Stephen Hillier, Jacob Kempster, Sergi Bosca, Alan Watson and Varsiha Sothilingam. With their help, I was able to contribute to the development of the Athena framework which is crucial for the ATLAS experiment.

I would also like to thank everyone in the QMUL particle physics group for the help and company during my PhD. I would particularly like to thank Jon Hays, Seth Zenz and Teppei Katori for helping me with my progression and providing feedback on my thesis chapters.

In the end, I would like to thank my family and friends for their support and love. Thank you!

Contents

Chapter 1

Introduction

A Higgs boson with a mass of about 125 GeV was discovered by the ATLAS [3] and CMS [4] experiments at the Large Hadron Collider (LHC), and its observed properties are compatible with the Standard-Model predictions. However, the Standard Model does not explain all experimental observations. For example, it does not explain gravity, dark matter or the matter-antimatter asymmetry in the Universe.

One possible extension of the Standard Model are the Two-Higgs-Doublet Models (2HDM) [5], where an additional Higgs doublet exists and the observed Higgs boson is part of an extended Higgs sector. The model predicts the existence of five Higgs-like bosons: two charged ($H^{\pm}$), two CP-even (h, H) and one CP-odd (A), where h is the Standard-Model-like Higgs boson which has been discovered. Many theories beyond the Standard Model, such as supersymmetry, require the existence of a second Higgs doublet. In this thesis, we discuss a search for the pseudoscalar Higgs A decaying into an h boson and a Z boson, with the h boson decaying into a bottom quark-antiquark pair and the Z boson decaying into electrons, muons or neutrinos. The result will be interpreted in four CP-conserving models with no tree-level flavour changing neutral currents. In the type-I model, all fermions couple to one Higgs doublet. In the type-II model, all up-type quarks couple to one Higgs doublet while all down-type quarks and leptons couple to the other one. In the flipped 2HDM, all down-type quarks couple to one Higgs doublet while all up-type quarks and leptons couple to the other one. In the lepton-specific 2HDM, all leptons couple to one Higgs doublet while all quarks couple to

the other one.

A heavy resonance decaying into the same final state is also expected in the Heavy-Vector-Triplet (HVT) [6] model which predicts the existence of a beyond-the-Standard-Model (BSM) heavy neutral vector boson Z'. Many BSM models predict the existence of heavy vector bosons decaying into Standard-Model Higgs bosons. However, the search is difficult because there are many models and free parameters. Since the search is not sensitive to all details of the models, a phenomenological Lagrangian which retains relevant couplings and the Z' mass is proposed. The free parameters of the HVT model can be computed by the specific models analytically. As a result, the results of HVT searches can be translated into different specific models.

Previous searches for the $A/Z' \to Zh$ have been made by the ATLAS [7] experiment using the data collected during 2015-2016 corresponding to an integrated luminosity of 36 fb^{-1}, and a mild excess of about three standard deviations has been observed for the b-associated production of the A boson at 440 GeV. The CMS experiment [8] has also searched for $Z' \to Zh$ using the full Run 2 data corresponding to an integrated luminosity of 137 fb^{-1}, and an excess of about three standard deviations has been observed at 1 TeV.

The thesis describes a search using the Run 2 data collected by the ATLAS detector corresponding to an integrated luminosity of 139 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The HVT signal considered in the search is the Z' boson produced in the Drell-Yan process. The 2HDM signal is produced by two modes: the gluon-gluon fusion which produces an A boson and the b-associated production which produces an A boson and two associated b quarks.

The thesis will first discuss a cut-based analysis which uses a similar analysis strategy as the previous ATLAS search [7]. The analysis sets upper limits on the production cross section of the A boson and the Z' boson. The analysis also calculates exclusion limits on the parameters of the type-I, type-II, flipped and lepton-specific 2HDMs. The thesis will then introduce a new analysis approach which uses deep neural networks, for which the expected upper limits on the production cross section of the A boson were calculated.

In addition to the search for the A boson and the Z' boson, the thesis will also discuss the development of the offline simulation for the ATLAS Level-1 calorimeter triggers. The simulation is part of the ATLAS offline software, and it will be used for the production of the Monte-Carlo samples for Run 3.

The author was the main analyzer of the 2-lepton channel where the Z boson decays into electrons or muons, and he was responsible for studies related to the event selection optimization, data/MC comparison, theory uncertainty calculation, background modelling and statistical interpretation in the 2-lepton channel. The author then collaborated with other members of the analysis team and combined the 2-lepton results with the analysis where the Z boson decays into two neutrinos. The author was also responsible for the signal smearing studies, the 2HDM interpretation and the calculation of pseudo-experiment limits. In addition to the physics analysis, the author also created the HEPData page and the analysis Recast [9].

After that, the author was involved in the second round of the Vh analysis where the same search was performed with new object-reconstruction methods and machine learning. The author was responsible for the production of signal samples and the neural-network studies. He has also contributed to the development of the analysis framework.

The author has contributed to the development and the operation of the ATLAS experiment. He was involved in the development of the Level-1 calorimeter trigger simulation and contributed to the validation of the simulation. The author worked in the ATLAS control room as a shifter at the trigger and run control desks.

The thesis is structured as follows. The theoretical motivation of the analysis is described in Chapter 2. Chapter 3 describes the ATLAS experiment. The development of the ATLAS level-1 calorimeter trigger simulation is described in Chapter 4. Chapter 5 describes how the physical objects such as electrons and jets are reconstructed from detector signals. Chapter 6 describes the 2-lepton channel analysis. Chapter 7 discusses the combined results and the 2HDM interpretation. The analysis with neural networks is described in Chapter 8. Chapter 9 concludes the thesis.

Chapter 2

Theory

This chapter discusses the theoretical motivation of the analysis. It starts with the Standard Model, the theory which describes all known elementary particles and their dynamics and interactions. Despite the fact that the Standard Model agrees with almost all experimental measurement, it is not a complete theory. Section 2.3 discusses the limitations of the Standard Model and some of the BSM theories which address them. It later describes two simple extensions to the Standard Model, the 2HDMs and the HVT Models. The two models are the main focus of this thesis. This chapter also briefly discusses the simulation of proton-proton collision events.

2.1 The Standard Model

The Standard Model describes elementary particles and the interactions between them via the strong force, weak force and electromagnetic force. In the Standard Model, all particles can be categorized into fermions with half-integer spin following the Fermi-Dirac statistics and bosons with integer spin following the Bose-Einstein statistics.

Fermions can be further categorized into leptons and quarks. There are two types of leptons according to the electric charge they carry: charged leptons with $Q = 1$ and neutrinos with $Q = 0$. Leptons fall into three generations: the electron (e^-) and the electron neutrino (ν_e); the muon (μ^-) and the muon neutrino (ν_μ); the tau (τ^-) and the tau neutrino (ν_τ). Each lepton has its corresponding antilepton (e^+, $\bar{\nu_e}$, μ^+, $\bar{\nu_\mu}$, τ^+ and

$\bar{\nu_\tau}$) with the sign of the electric charge reversed. Quarks can be classified into up-type quarks (up u, charm c and top t) carrying electric charge $Q = \frac{2}{3}$ and down-type quarks (down d, strange s and bottom b) carrying electric charge $Q = -\frac{1}{3}$. Each quark also has its corresponding antiquark ($\bar{u}$, $\bar{c}$, $\bar{t}$, $\bar{d}$, $\bar{s}$ and $\bar{b}$) with the sign of the electric charge reversed.

In the Standard Model, the three fundamental interactions are mediated by spin-1 vector bosons. The electromagnetic interaction and the weak interaction are unified as the electroweak interaction which is mediated by the exchange of photons (γ), Z and $W^{\pm}$ bosons. The strong interaction is mediated by the exchange of gluons (g). The Standard Model also contains a spin-0 scalar boson known as the Higgs boson (h), which will be discussed in Section 2.1.3.

The discussion in this section is mainly based on Refs. [10] and [11].

2.1.1 Particles and Fields

The Standard Model describes particles and their interactions based on quantum field theory. In quantum mechanics, the state of a single particle can be described by wave functions of discrete modes. In quantum field theory, each mode is the excitation of a quantum field, and the nth excitation can be interpreted as having n particles. The kinematics of a field (ϕ) is described by the Lagrangian density $\mathcal{L}$, and its equation of motion can be determined using the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0. \tag{2.1}$$

The Lagrangian needs to be invariant under the Lorentz transformation. The Lagrangian of a spin-0 scalar field can be written as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2. \tag{2.2}$$

Its equation of motion is the Klein-Gordon equation

$$(\Box + m^2)\phi = 0. \tag{2.3}$$

A spin-1 vector field satisfies the Proca Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^\mu A_\mu, \tag{2.4}$$

where the field strength tensor $F_{\mu\nu} = \partial_\mu A^\nu - \partial_\nu A^\mu$. When describing electromagnetic fields with zero photon mass, the equation of motion of the Proca Lagrangian is Maxwell's equations in the absence of sources.

A spin-$\frac{1}{2}$ field ψ with four complex components satisfies the Dirac Lagrangian

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \tag{2.5}$$

where γ^μ are the gamma matrices. The equation of motion of the field is the Dirac equation

$$i\gamma^\mu(\partial_\mu\psi) - m\psi = 0. \tag{2.6}$$

2.1.2 Gauge Invariance

The symmetry of the Standard Model is described by group theory. A group is a set of objects $\{g_i\}$ that satisfies the following rules:

- The product of two elements $g_i \times g_j = g_k$ is also an element of the group.
- There is an identity element for which $\mathbb{1} \times g_i = g_i$.
- Each element in the group also has its inverse for which $g_i^{-1} \times g_i = \mathbb{1}$.
- The associative law $(g_i \times g_j) \times g_k = g_i \times (g_j \times g_k)$ holds for all elements in the group.

The symmetry of the Standard Model is described by Lie groups which have a finite number of generators and an infinite number of elements. An element of a Lie group can be written as

$$U = e^{i\theta_\alpha T^\alpha} \cdot \mathbb{1}, \tag{2.7}$$

where θ_α are parameters and T^α are the generators of the group. The commutation relations of Lie group generators are

$$[T^a, T^b] = if^{abc}T^c, \tag{2.8}$$

where f^{abc} are the structure constants of the group. A group is Abelian if $f^{abc} = 0$, otherwise it is non-Abelian. A group is unitary if $U^\dagger U = \mathbb{1}$, and it is a special group if $|U| = 1$. A special unitary group acting on an N-dimensional vector space (SU(N)) has $N^2 - 1$ generators.

Quantum Electrodynamics

The Standard Model is a quantum field theory where the Lagrangian is invariant under the local gauge transformation of the $SU(3)\times SU(2)\times U(1)$ group. In the Standard Model, quantum electrodynamics (QED) is introduced by the $U(1)$ symmetry. Transforming under the $U(1)$ local gauge transformations of angle $\theta(x)$, the spin-$\frac{1}{2}$ field becomes

$$\psi(x) \to \psi'(x) = e^{iQ\theta(x)}, \tag{2.9}$$

where Q is a constant. The Dirac Lagrangian in Equation (2.5) becomes

$$\mathcal{L} \to \mathcal{L}' - Q\bar{\psi}\gamma^\mu\partial_\mu\theta(x)\psi. \tag{2.10}$$

To make the Lagrangian invariant under the local gauge transformation, we can replace ∂_μ with the covariant derivative $D_\mu = \partial_\mu + iQA_\mu$. By adding a new term $(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu})$ which describes the free propagation of the new A_μ field, we have the QED Lagrangian

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu\partial_\mu - m_e)\psi + e\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.11}$$

where m_e is the electron mass and e is the electric charge. In QED, the electromagnetic force is carried by massless photons, and they interact with particles carrying electric charge. The photon itself does not carry electric charge.

Weak Interaction

The SU(2) group describes the symmetry of the weak interaction. The weak interaction involves the chirality of fermions. The chiral operator is defined as $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, and a fermion field can be written as the sum of a left-handed field and a right-handed field,

$$\psi = \psi^L + \psi^R, \tag{2.12}$$

where

$$\psi^R = \frac{1+\gamma^5}{2}\psi, \psi^L = \frac{1-\gamma^5}{2}\psi. \tag{2.13}$$

The left-handed leptons (e_L, μ_L, τ_L, ν_{eL}, $\nu_{\mu L}$, $\nu_{\tau L}$) and the left-handed quarks (u_L, d_L, s_L, c_L, t_L, b_L) can be represented by isospin doublets with weak isospin $I = \frac{1}{2}$:

$$\psi^L_{lepton} = \begin{pmatrix} e_L \\ \nu_{eL} \end{pmatrix}, \begin{pmatrix} \mu_L \\ \nu_{\mu L} \end{pmatrix}, \begin{pmatrix} \tau_L \\ \nu_{\tau L} \end{pmatrix}, \quad \psi^L_{quark} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} s_L \\ c_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}. \tag{2.14}$$

The right-handed leptons and the right-handed quarks can be represented by isospin singlets with $I = 0$:

$$e_R, \mu_R, \tau_R, u_R, s_R, t_R, d_R, c_R, b_R. \tag{2.15}$$

The isospin doublets transform under the $SU(2)$ local gauge transformation while the singlets are unaffected. We define the third component of the weak isospin I^3 to be $\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0$ for left-handed up-type quarks, left-handed down-type quarks, left-handed neutrinos, left-handed charged leptons and right-handed fermions respectively.

Like the treatment for QED, to make the Lagrangian invariant under the $SU(2)$ local gauge transformation, we can replace ∂_μ with the covariant derivative

$$D_\mu = \partial_\mu + ig_W \boldsymbol{T}^\mu \cdot \boldsymbol{W}_\mu(x), \tag{2.16}$$

where g_W is a constant, $\boldsymbol{T}^\mu$ are the generators of the $SU(2)$ group and $\boldsymbol{W}_\mu(x)$ are three new gauge fields. Due to the non-Abelian nature of the SU(2) group, the field strength tensor has to include a gauge boson self-interaction term to make the Lagrangian gauge

invariant

$$W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu - g_W f_{ijk} W_j^\mu \times W_k^\nu. \tag{2.17}$$

The weak gauge fields only couple to fields carrying weak isospin.

Quantum Chromodynamics

Like the weak interaction, quantum chromodynamics (QCD) is also a non-Abelian gauge theory. The $SU(3)$ symmetry group has eight generators, which requires eight gauge fields to make the Lagrangian gauge invariant. The covariant derivative can be written as

$$D_\mu = \partial_\mu + i g_s \boldsymbol{T}^\mu \cdot \boldsymbol{G}_\mu(x), \tag{2.18}$$

where g_s is a constant, $\boldsymbol{T}^\mu$ are the generators of $SU(3)$ group and $\boldsymbol{G}_\mu(x)$ are the new gauge fields. The field strength tensor is

$$G_i^{\mu\nu} = \partial^\mu G_i^\nu - \partial^\nu G_i^\mu - g_s f_{ijk} G_j^\mu \times G_k^\nu. \tag{2.19}$$

The gauge bosons of the strong interaction are massless gluons, which interact with particles carrying colour charges. A quark carries one of the three colour charges (r, g, b), while an antiquark carries one of the three anti-colour charges $(\bar{r}, \bar{g}, \bar{b})$. The non-Abelian nature of the $SU(3)$ group means that gluons also carry colour charges, and they can interact among themselves. The QCD Lagrangian can be written as

$$\mathcal{L}_{QCD} = i\bar{\psi}_c \gamma^\mu (D_\mu - m)\psi_c - \frac{1}{4}\boldsymbol{G}^{\mu\nu}\boldsymbol{G}_{\mu\nu}, \tag{2.20}$$

where

$$\psi_c = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}. \tag{2.21}$$

Because of the non-Abelian nature of QCD and the massless gluons, particles carrying colour charges cannot be isolated. This phenomenon is called colour confinement. Quarks exist in the form of hadrons. The colour singlet states $r\bar{r}$, $g\bar{g}$, and $b\bar{b}$ are called meson

while rbg and $\bar{r}\bar{g}\bar{b}$ statues are called baryons. If two high-momentum quarks travel back to back, quarks and antiquarks will be produced because of colour confinement. This phenomenon is called hadronisation. Quarks and gluons produced at the LHC are observed as collimated streams of particles called hadronic jets.

2.1.3 The Higgs Mechanism and The Electroweak Unification

The kinematics of the gauge bosons are described by the Proca Lagrangian in Equation (2.4), but the mass term $\frac{1}{2}m^2A^\mu A_\mu$ is not invariant under the local gauge transformation. The mass term is not needed for QED and QCD because photons and gluons are massless. However, a mass term is required to describe the weak interaction since its gauge bosons are massive. This is also a problem for the fermion mass since the mass term in Equation (2.5) is also not invariant under the local gauge transformation. In the Standard Model, a Higgs field is introduced to break the $U(1) \times U(2)$ symmetry and give masses to fermions and $W^\pm$, Z bosons.

Spontaneous Symmetry Breaking

In the Standard Model, the Higgs field is a $SU(2)$ doublet

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.22}$$

and its Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{2.23}$$

When $\mu^2 < 0$, the potential term $-\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 = \frac{v^2}{2}$ is minimum at $\phi^\dagger\phi = -\frac{\mu^2}{2\lambda}$. One possible solution for the vacuum is

$$\phi_{vac} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.24}$$

If one expands the Higgs field around the minimum potential, there will be one massive scalar boson and three massless scalar bosons known as Goldstone bosons. The three massless bosons can be removed by applying the unitary gauge, and we have

$$\phi_{vac} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.25}$$

where $h(x)$ is the massive scalar field of the Higgs boson.

Gauge Boson Mass

To make Equation (2.23) invariant under the local gauge transformation of $U(1) \times U(2)$ group, one can replace the partial derivative with the covariant derivative

$$D_\mu = \partial_\mu + \frac{i}{2} g_W \boldsymbol{\sigma} \cdot \boldsymbol{W}_\mu + \frac{i}{2} g' B_\mu, \tag{2.26}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices, g' is a constant and B_μ is the gauge field associated with the $U(1)$ symmetry. The mass term for the $W^{(1,2)}$ fields is $\frac{1}{8} g_W^2 v^2 W^{(1,2)\mu} W_\mu^{(1,2)}$, and the one for the B field and the $W^{(3)}$ fields is

$$\frac{v^2}{8} \begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix} \begin{pmatrix} g_W^2 & -g_W g' \\ -g_w g' & g'^2 \end{pmatrix} \begin{pmatrix} W_\mu^{(3)} \\ B_\mu \end{pmatrix}. \tag{2.27}$$

After diagonalizing the mass matrix, the physical Z boson field and physical photon field (A) can be written as a combination of the B field and the $W^{(3)}$ field:

$$\begin{aligned} A_\mu &= cos\theta_W B_\mu + sin\theta_W W_\mu^{(3)}, \\ Z_\mu &= -sin\theta_W B_\mu + cos\theta_W W_\mu^{(3)}, \end{aligned} \tag{2.28}$$

where $cos\theta_W = \frac{g_W}{\sqrt{g_W^2 + g'^2}}$. The physical $W^\pm$ fields are a combination of the $W^{(1)}$ field and the $W^{(2)}$ field:

$$W_\mu^\pm = \frac{1}{2}(g W_\mu^{(1)} \mp i W_\mu^{(2)}). \tag{2.29}$$

The mass term coefficients for photons, Z bosons and $W^{\pm}$ bosons are 0, $g_W v/2$ and $v\sqrt{g_W^2 + g'^2}/2$ respectively.

Fermion Mass

The Higgs field also generates masses for fermions via the Yukawa Coupling. The following Lagrangian can be constructed which is invariant under the $U(1) \times SU(2)$ local gauge transformation

$$\mathcal{L} = -g_f[\bar{\psi}_f^L \phi \psi_f^R + (\bar{\psi}_f^L \phi \psi_f^R)^\dagger] + g_f[\bar{\psi}_f^L \phi_c \psi_f^R + (\bar{\psi}_f^L \phi_c \psi_f^R)^\dagger], \tag{2.30}$$

where ϕ_c is the conjugate doublet of ϕ and g_f is a constant. The first term in the Lagrangian gives mass to charged leptons and down-type quarks, while the second term gives mass to up-type quarks.

We can write the Yukawa coupling Lagrangian in the mass basis of quarks by diagonalizing the mass matrix for quarks, but this rotation also affects the electroweak terms. The relationship between the mass eigenstates of down-type quarks (d, s, c) and their flavour eigenstates (d', s', c') is

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.31}$$

where V_{CKM} is the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The matrix contains a complex phase which is a source of CP violation. The coupling of the B_μ and $W_\mu^{(3)}$ is unaffected by this rotation, so flavour changing neutral current (FCNC) is suppressed in the Standard Model.

The same can also be done for neutrinos where the flavour eigenstates (ν_e, ν_μ, ν_τ)

can be written as a combination of the mass eigenstates (ν_1, ν_2, ν_3)

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = V_{PMNS} \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{2.32}$$

where V_{PMNS} is the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, which allows neutrinos to change flavours as it propagates through space.

2.2 Modelling of Proton-Proton Collision Events

Quantum field theory can describe proton-proton collision events, which is crucial for understanding the signal and background processes in the LHC experiment. This section will discuss the calculation of cross sections and the Monte Carlo (MC) simulation of proton-proton collision events. The discussion in this section is mainly based on [12] and [13].

2.2.1 Cross Sections and Decay Rates

The cross section of a scattering process is defined as

$$\sigma = \frac{N}{T\Phi}, \tag{2.33}$$

where N is the number of particles scattered in a period of time T and Φ is the incoming flux. The decay rate Γ is the probability of a particle decaying in a unit time. The branching ratio is defined as $\frac{\Gamma_i}{\Gamma_{total}}$ where Γ_{total} is the overall decay rate and Γ_i is the decay rate of a particular process.

To calculate cross sections and decay rates, we define the S-matrix as

$$\langle f | S | i \rangle_{\text{Heisenberg}} = \langle f, \infty | i, -\infty \rangle_{\text{Schrodinger}}, \tag{2.34}$$

which contains information about the time evolution between the initial states and the finals states [12]. The S-matrix is an identity matrix if no interaction happens. We can

use matrix elements $\mathcal{M}$ to describe the amplitude of interaction:

$$\langle f| S - \mathbb{1} |i\rangle = i(2\pi)^4 \delta^4(\Sigma p_i - \Sigma p_f)\mathcal{M}, \tag{2.35}$$

where p_i are the initial-state momenta and p_f are the final-state momenta. The matrix element $\mathcal{M}$ can be calculated perturbatively using Feynman rules. Fermi's Gold Rule states the transition rate is the product of $|M|^2$ and the density of states. The cross section and decay rate are the integral over all possible phase space.

2.2.2 Renormalization

Figure 2.1: The leading order diagram (left) and a diagram with a loop correction (right) for a scattering process. The diagrams are created using the TikZ-Feynman package [14].

Figure 2.1 shows a leading order diagram and a next-to-leading order diagram with a virtual loop correction for a scattering process. To calculate the physical cross section, we need to calculate the sum of matrix elements of all possible diagrams. The calculation of the leading order diagram yields a finite amplitude, but the amplitude of the diagram with the loop correction is infinite. The scattering amplitudes represent the probability of transition, so the infinite amplitude is not physical.

The cause of the problem is that the integral is up to an arbitrarily high energy scale where the Standard Model and the quantum field theory are no longer valid. The infinity can be removed by regularizations, such as applying a cutoff for the integral at a certain momentum Λ, or doing the calculation in $(4 - \varepsilon)$ dimensions. By using the cutoff regulator, the amplitude is a function of Λ, but the differences in amplitudes at different energy scales are independent of Λ. As a result, physical observables have to be measured at certain energy scales in order to make predictions using the theory.

In tree-level theory, the masses, fields and coupling constants in the Lagrangian are

finite. However, they are redefined as infinite bare parameters which absorb the infinity when considering higher-order diagrams. A bare electric charge e_0 can be written as

$$e_0 = (1 + \delta)e_R, \tag{2.36}$$

where δ is an infinite counterterm and e_R is the renormalized charge observed at a certain energy scale. Counterterms can be chosen in a form that removes the divergence in the amplitude.

When defining the counterterms, a constant (renormalization scale) μ_R needs to be introduced to ensure that the counterterm has a meaningful unit. Physical observables are independent of μ_R if they are calculated to all orders. Since one can only calculate observables to a finite order, the value of μ_R is chosen to minimize the contributions from higher-order diagrams. The effect of the contributions from higher orders can be estimated by varying μ_R, and the estimation can be considered as a systematic uncertainty.

Running Coupling Constants

Since the bare parameters in the Lagrangian are independent of the scale μ_R, we have

$$\mu_R \frac{de_0}{d\mu_R} = 0, \tag{2.37}$$

which gives the renormalization group equation:

$$\frac{de_R(\mu_R)}{\beta(e_R(\mu_R))} = \frac{d\mu_R}{\mu_R}, \tag{2.38}$$

where is $\beta(e_R)$ is the QED beta function [15] which depends on e_R. Integrating both sides from the energy scale Q_0 to the energy scale Q_1 gives

$$e^2(Q_2) = \frac{e^2(Q_0)}{1 - \frac{e^2(Q_0)}{6\pi} ln(\frac{Q_1}{Q_0})}. \tag{2.39}$$

The result shows that the observed effective charge increases with the energy scale. This effect is called the running coupling constant.

Like the QED, the coupling constant of the QCD also changes with the energy scale. Due to the non-Abelian nature of the QCD, its coupling constant decreases as the energy scale increases. The perturbation theory is valid only when the coupling constant $\ll 1$. Hence, the perturbation theory for the QCD breaks down at low energy scales where the coupling constant is large.

Infrared Divergence and Collinear Divergence

Figure 2.2: Final-state radiation diagrams for $ee \to \mu\mu$. The diagrams are created using the TikZ-Feynman package [14].

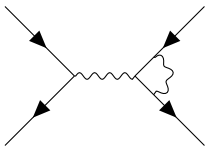


Figure 2.3: The vertex correction diagram for $ee \to \mu\mu$. The diagram is created using the TikZ-Feynman package [14].

Figure 2.2 shows the final-state radiation diagram for the $ee \to \mu\mu$ process where an additional photon is emitted from the final-state muons. The amplitude of the two diagrams diverges if the photon energy approaches zero or the angle between the photon and the muon approaches zero. The two types of divergence are called infrared divergence and collinear divergence. Unlike the ultraviolet divergence discussed in the previous sections, the infrared and collinear divergence cannot be cancelled by calculating the difference in amplitude at different energy scales. The infrared divergence and the collinear divergence also occur when calculating the initial-state radiation where a photon is emitted from the initial-state electrons. The finite physical result can be restored by calculating the total cross section of both diagrams in Figure 2.2 and the vertex correction diagram in Figure 2.3.

The infinite cross section of photon emission is not a problem in practice because particle detectors only have a finite resolution, so the emitted photon cannot be resolved if its energy is too low or the angle between the muon and photon is too small. The final-state muon and its emitted photons with energy and angles below a certain threshold can be defined as a jet.

2.2.3 Parton Distribution Functions

Protons are composite particles where quarks inside interact with each other via gluons, and the gluons also create quark-antiquark pairs. The point-like constituents of protons are called partons. At high energy, proton collisions need to be considered as interactions between partons. Parton distribution functions (PDFs) are the probability of finding a parton carrying a fraction x of the proton's total energy. The cross section of a proton collision process can be written in terms of PDFs ($f(x)$):

$$\sigma = \int_0^1 dx_1 \int_0^1 dx_2 f(x_1) f(x_2) \sigma(x_1 p_1, x_2 p_2), \tag{2.40}$$

where $\sigma(x_1 p_1, x_2 p_2)$ is the interaction cross section, p_1 and p_2 are the momenta of the two protons and x_1 and x_2 are the fractions of momenta the partons carry [13].

During the cross section calculation, infrared divergence and collinear divergence occur because of initial-state radiation. To avoid the infinity, a factorization scale (μ_F) is defined as a cutoff where the cross section is calculated perturbatively if the momenta of emitted particles $|k| > \mu_F$. The emitted particle is considered as part of a redefined PDF $f(x, \mu_F)$ at the energy scale μ_F if $|k| < \mu_F$ [13]. Physics is independent of the arbitrary choice of μ_F, but μ_F does have an effect in practice due to limited orders of the calculation. Like the renormalization scale μ_R, the theory uncertainty related to the choices of μ_F can be estimated by varying μ_F. The PDF can only be determined by measurement because the perturbative QCD calculation fails at low energy scales.

2.2.4 Monte Carlo Simulation

Due to the complexity of cross section calculations, it is difficult to model the proton-proton collision analytically. Instead, the distributions of the signals and backgrounds are estimated by Monte Carlo methods where software tools generate a large number of possible outcomes of proton-proton collisions [13]. The software tools also simulate how the detector responds to the simulated events.

The simulation starts from the calculation of matrix elements of parton interaction processes up to a fixed order. The higher-order effects on the normalization of the distribution can be taken into account by applying a k-factor to the weights of the simulated events.

Parton-shower algorithms approximate the emission of partons from the initial and final states. When the emitted gluon is almost collinear to the mother parton, the process that a parton splits into a parton and a gluon is independent of the rest of the scattering process, so the cross section of a diagram with a parton emission can be factorized as a product of the leading-order diagram's cross section and the probability of the splitting estimated by the DGLAP splitting functions [16]. The successive splittings are simulated by Markov-chain algorithms until a cutoff on the momentum is reached. Splittings are not allowed below the cutoff because the detector can no longer resolve the emitted particles, and the result diverges. After the parton-shower simulation, the partons and gluons form colourless hadrons which eventually enter the detector simulation.

The matrix-element calculation makes a good prediction when the mother parton and the emitted gluon are well separated, while the parton shower describes the splitting better in the collinear limit. When combining the result of matrix-element calculation and parton-shower simulation, the overlapping real emission and virtual correction need to be removed to avoid double counting.

After a hard parton scattering, the remaining protons become unstable and break apart. Multiple-parton interactions happen in a single collision between two protons. Those two processes produce underlying events. In the real LHC environment, pileup happens when multiple proton collisions happen in the same bunch crossing. The underlying event and pileup are also simulated by the Monte Carlo event generators

because they affect the event reconstruction. The Monte Carlo events are reweighted to match the pileup level in data.

The last step of the Monte Carlo simulation is the detector simulation where interaction between the particles and the detector is simulated by GEANT4-based detector simulation [17] [18]. The physical objects are eventually reconstructed from the detector responses using the same software packages for the data reconstruction.

2.3 Beyond The Standard Model

The Standard Model explains all known particles and agrees well with the current experimental results. However, the Standard Model is not the ultimate theory because it does not answer all crucial questions about the universe. It is believed that new physics exists at high energy scales.

2.3.1 Limitations of the Standard Model

Gravity

One of the four fundamental forces in the universe, gravity, is not described by the Standard Model. It is instead described by the theory of general relativity. The effect of gravity is weak, and it can be ignored in current particle physics experiment, but a quantum-gravity theory is needed to explain physics beyond the Planck scale [11].

Matter-Antimatter Asymmetry

The big bang created an equal amount of matter and antimatter, but our current universe is dominated by matter [11]. To create the matter-antimatter asymmetry, baryon-number violation and CP violation are needed. CP violation has been observed in the quark sector, and it is explained by the phase in the CKM matrix. Although never observed, the baryon number can also be violated in the Standard Model by the anomaly where symmetries in the classical Lagrangian are broken when quantum effects are added [12]. More baryon-number violation and CP violation sources are needed to explain the observed matter-antimatter asymmetry in the universe.

The Strong CP Problem

In the Standard Model, a CP violating term in the strong interaction is also allowed. However, experiments show the term is extremely small, and the Standard Model cannot explain why [12].

Dark Matter

Astrophysical measurements suggest that most matter density in the universe exists in the form of dark matter. None of the particles in the Standard Model are a candidate for dark matter due to cosmological theory constraints. New weakly interacting massive particles are one explanation for dark matter [11].

The Hierarchy Problem

As described in Section 2.2.2, particle masses receive loop corrections. The corrections for fermion masses and gauge boson masses are only proportional to the particle mass itself because their masses are protected by custodial symmetries. However, the Higgs mass is not protected in the Standard Model, and it is quadratically sensitive to the cutoff energy scale Λ. To avoid fine-tuning the Higgs mass, the energy scale where new physics occurs must be below 1-2 TeV to ensure that the correction is comparable to the measured Higgs mass of 125 GeV [19].

2.3.2 Two-Higgs-Doublet Models

The 125 GeV Higgs boson discovered at the LHC is compatible with the Standard-Model prediction. However, it could also be part of an extended Higgs sector. In 2HDMs, the Higgs sector consists of two $SU(2)$ doublets, and it predicts the existence of five Higgs-like particles. In this section, the discussion of the 2HDM theory is mainly based on Ref. [5].

Motivation

The 2HDM [5] is motivated by many beyond-the-Standard-model theories. In super-symmetry (SUSY) theories [20], each Standard-Model fermion has its bosonic SUSY

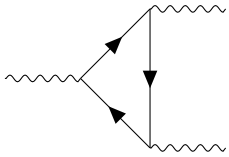


Figure 2.4: The Feynman diagram of a triangular fermion loop which causes anomalies. The diagram is created using the TikZ-Feynman package [14].

partner, and each Standard-Model boson has its fermionic SUSY partner. The SUSY theories avoid the hierarchy problem because SUSY particles' contribution to the mass correction cancels the contribution from their Standard-Model partners. In SUSY, a second Higgs doublet is needed to cancel the triangle anomalies where the axial current of gauge groups is conserved at the tree level, but it is not conserved when considering contributions from the triangular fermion loop shown in Figure 2.4. The violation of gauge symmetries is not allowed because it breaks the Ward identity [21] which cancels an unphysical degree of freedom of photons. In the Standard Model, the gauge anomalies are cancelled by contributions from the three generations of leptons and quarks. In SUSY theories, the Higgs boson's fermionic partner also contributes to the triangular anomalies; thus, a second Higgs doublet is required to provide more Higgs-like particles to cancel its contribution.

The 2HDM is also motivated by the axion theories [22]. The theory provides a solution to the strong CP problem by introducing a new $U(1)$ symmetry which is spontaneously broken. The new Goldstone boson can be rotated into the strong CP phase, creating an axion potential term. The mechanism moves the potential to zero even if the strong CP phase is not initially zero. The excitations of the vacuum potential are axions, and it is a candidate for dark matter. In axion theories, a second Higgs doublet is needed to make the Yukawa coupling invariant under the new symmetry.

The 2HDM can also provide additional sources of CP violation, which helps to explain the matter-antimatter asymmetry observed in the current universe [5].

The 2HDM Lagrangian

The vacuum structure of 2HDMs is complex, and the minimum potentials can be CP-conserving, CP-violating and charge-violating. In order to simplify the study, this

thesis only considers the CP-conserving 2HDMs whose scalar potential is

$$\begin{aligned} V = & m_{11}^2\Phi_1^\dagger\Phi_1 + m_{22}^2\Phi_2^\dagger\Phi_2 - m_{12}^2(\Phi_1^\dagger\Phi_2 + \Phi_2^\dagger\Phi_1) + \frac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2)^2 \\ & + \lambda_3\Phi_1^\dagger\Phi_1\Phi_2^\dagger\Phi_2 + \lambda_4\Phi_1^\dagger\Phi_2\Phi_2^\dagger\Phi_1 + \frac{\lambda_5}{2}[(\Phi_1^\dagger\Phi_2)^2 + (\Phi_2^\dagger\Phi_1)^2], \end{aligned} \tag{2.41}$$

where Φ_1 and Φ_2 are the two complex scalar Higgs doublets, and m_{11}, m_{22}, λ_1 , λ_2, λ_3, λ_4 and λ_5 are real parameters [5]. The ground state which minimizes the scalar potential is

$$\langle\phi_1\rangle_0 = \begin{pmatrix} 0 \\ \frac{\nu_1}{\sqrt{2}} \end{pmatrix}, \quad \langle\phi_2\rangle_0 = \begin{pmatrix} 0 \\ \frac{\nu_2}{\sqrt{2}} \end{pmatrix}. \tag{2.42}$$

where ν_1 and ν_2 are the vacuum expectation values. Expanding about the minimum potential leads to eight free fields, where three of them give mass to the W and Z bosons, and the remaining five are massive Higgs fields. Substituting the decomposed Higgs doublets

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ (\nu_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, \quad a = 1, 2, \tag{2.43}$$

into the scalar potential gives three mass terms:

$$\begin{aligned} \mathcal{L}_{\phi^\pm} &= (\phi_1^-, \phi_2^-) M_\phi \begin{pmatrix} \phi_1^+ \\ \phi_2^+ \end{pmatrix}, \\ \mathcal{L}_\eta &= (\eta_1, \eta_2) M_\eta \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \\ \mathcal{L}_\rho &= (\rho_1, \rho_2) M_\rho \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \end{aligned} \tag{2.44}$$

where $M_\phi^\pm$, M_η and M_ρ are the mass matrices for the charged scalars, the neutral scalars and the pseudoscalars respectively. The M_ρ matrix is diagonalized by the rotation angle α, resulting in two physical fields:

$$\begin{aligned} h &= \rho_1 \mathrm{sin}\alpha - \rho_2 \mathrm{cos}\alpha, \\ H &= -\rho_1 \mathrm{cos}\alpha - \rho_2 \mathrm{sin}\alpha, \end{aligned} \tag{2.45}$$

	type-I	type-II	flipped	lepton-specific
ξ_h^u	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
ξ_h^d	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$
ξ_h^l	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_H^u	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
ξ_H^d	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$
ξ_H^l	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
ξ_A^u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
ξ_A^u	$-\cot\beta$	$\tan\beta$	$\tan\beta$	$-\cot\beta$
ξ_A^u	$-\cot\beta$	$\tan\beta$	$-\cot\beta$	$\tan\beta$

Table 2.1: Values of Yukawa coupling constants of the natural Higgs bosons in the 2HDMs. The table is taken from [5].

where h is the light neutral scalar and H is the heavy neutral scalar. The light neutral scalar h is considered to be the 125 GeV Higgs boson discovered at the LHC. The $M_\phi^\pm$ and M_ρ matrices are diagonalized by the rotation angle β, which leads to three physical fields:

$$\begin{aligned} H^\pm &= \phi_1^\pm \sin\beta - \phi_2^\pm \cos\beta, \\ A &= -\rho_1 \cos\alpha - \rho_2 \sin\alpha, \end{aligned} \tag{2.46}$$

where $H^\pm$ are the charged scalars and A is the pseudoscalar.

One feature of the 2HDM is the existence of FCNCs mediated by the charge-natural Higgs bosons at the tree level because the mass matrices which describe the coupling of fermions to the two Higgs doublets cannot be diagonalized simultaneously. This can be avoided if all fermions of the same charge only couple to one of the Higgs doublets [23] [24]. There are four models [5] defined by the coupling between the Higgs doublets and fermions. In the type-I model, all fermions couple to one Higgs doublet. In the type-II model, all up-type quarks couple to one Higgs doublet while all down-type quarks and leptons couple to the other one. In the flipped 2HDM, all down-type quarks couple to one Higgs doublet while all up-type quarks and leptons couple to the other one. In the lepton-specific 2HDM, all leptons couple to one Higgs doublet while all quarks couple to the other one.

The Yukawa coupling between the fermion f and the natural Higgs bosons is described

by the Lagrangian

$$\mathcal{L} = -\frac{m_f}{v}(\xi_h^f \bar{f} f h + \xi_H^f \bar{f} f H - i\xi_A^f \bar{f}\gamma_5 f A), \tag{2.47}$$

where the values of the three coupling constants ξ_h^f, ξ_H^f and ξ_A^f are given in Table 2.1 [5].

The Phenomenology

The thesis will focus on the $A \to Zh \to l\bar{l}b\bar{b}$ decay channel. Table 2.1 shows that the $h \to b\bar{b}$ decay is not possible at certain α values because the coupling constants of the light natural scalar Higgs h to down-type quarks become zero. In the type-I model, this effect is known as the fermiophobic limit because the coupling of h to all fermions is zero at $\alpha = \frac{\pi}{2}$. The search has no sensitivity in the regions of phase space without the $h \to b\bar{b}$ decay.

Kominis [25] has studied the production of the pseudoscalar A at the LHC. Gluon-gluon fusion is the dominant production mode at the LHC, but its production cross section decreases as tanβ increases. In the type-II and the lepton-specific model, the cross section of the b-associated production increases with tanβ, and it becomes dominant in the regions of phase space with high tanβ values. The cross section of both production modes is not affected by $\cos(\beta - \alpha)$. In addition to the two production modes which are considered in this thesis, it is also possible to search for the t-associated production of the A boson at the LHC. This process has a small cross section, but it has a clear signature due to the two leptons from the top quark decay.

Kominis [25] also studied the decay of the pseudoscalar A. The main decay modes are $A \to b\bar{b}$, $c\bar{c}$, $t\bar{t}$, $\tau^+\tau^-$, Zh, gg, $\gamma\gamma$, $Z\gamma$. The $A \to Zh$ is particularly important when the A boson mass (m_A) is above the Zh production threshold. However, $A \to t\bar{t}$ dominates if the m_A is above the $t\bar{t}$ production threshold. The coupling of the A boson to the Z boson and the h boson is proportional to $\cos(\beta - \alpha)$. At the alignment limit where $\cos(\beta - \alpha) = 0$, the $A \to Zh$ decay is not possible, so the search also has no sensitivity in the regions of phase space near the alignment limit.

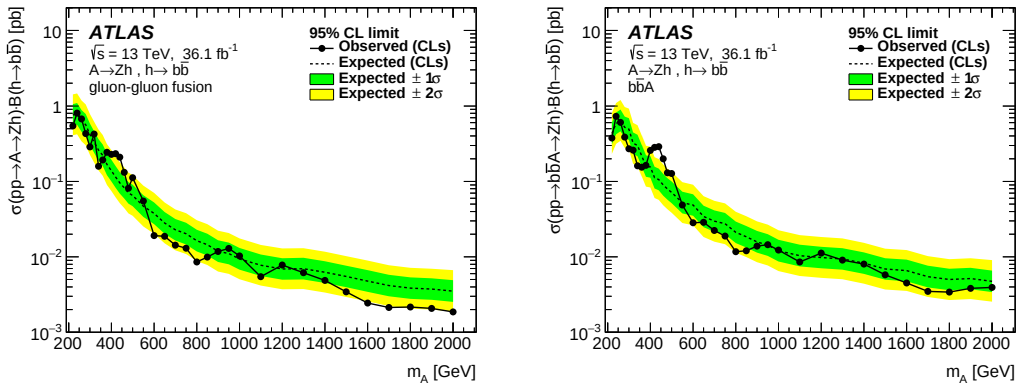


Figure 2.5: The expected and observed upper limits for the gluon-gluon fusion production mode (left) and the b-associated production made (right) on $\sigma(pp \to A \to ZH) \times BR(H \to b\bar{b})$ at the 95% confidence level calculated using the Run 2 data collected by the ATLAS detector corresponding to an integrated luminosity of 36.1 fb^{-1}. The plots are taken from Ref. [7].

Experimental Constraints

The previous ATLAS analysis [7] has searched for the $A \to Zh \to l\bar{l}b\bar{b}$ process using the 13 TeV proton-proton collision data corresponding to an integrated luminosity of 36.1 fb^{-1} in a mass range of 220 GeV to 2 TeV, shown in Figure 2.5. Upper limits at the 95% confidence level have been placed between 2×10^{-3} pb and 8×10^{-1} pb for the gluon-gluon fusion production and between 3×10^{-3} pb and 7×10^{-1} pb for the b-associated production. The observed limits in general agree with the expected limits, except for an above-three-standard-deviation local excess at the 440 GeV in the b-associated production search.

The CMS Collaboration has conducted the same search [26] using 13 TeV proton-proton collision data corresponding to an integrated luminosity of 35.9 fb^{-1} in a mass range of 225 GeV to 1 TeV, shown in Figure 2.6. The observed limits agree with the expected limits. Upper limits of the two production modes at the 95% confidence level have been placed between 1×10^{-2} pb to 1 pb.

Kling, Su and Su [27] have made a comprehensive study about the exclusion of 2HDM parameters by combining the results of all direct 2HDM searches and relevant Standard-Model measurements at the LHC. Figure 2.7 shows the regions excluded by the $A/H \to \gamma\gamma/tt/bb/\tau\tau$, $A \to Zh$, $H \to hh/VV$, $h \to VV$, $h \to AA$ and four tops

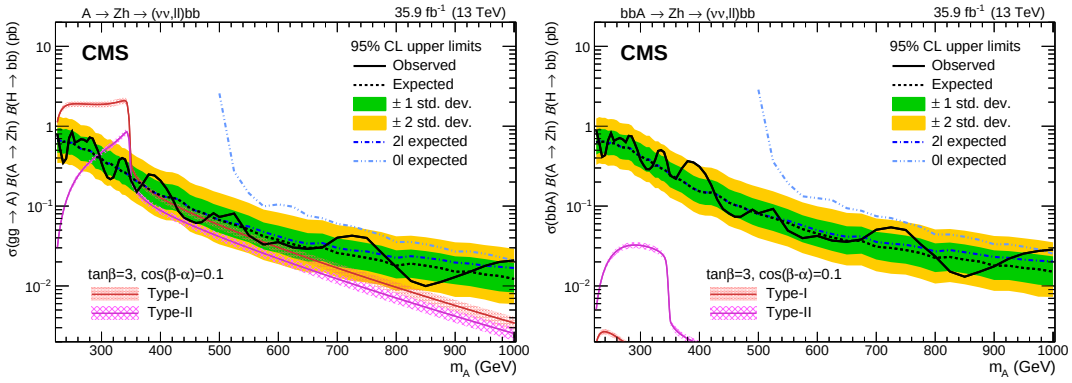


Figure 2.6: The expected and observed upper limits for the gluon-gluon fusion production mode (left) and the b-associated production made (right) on $\sigma(pp \to A \to ZH) \times BR(H \to b\bar{b})$ at the 95% confidence level calculated using the Run 2 data collected by the CMS detector corresponding to an integrated luminosity of 35.9 fb^{-1}. The plots are taken from Ref. [26].

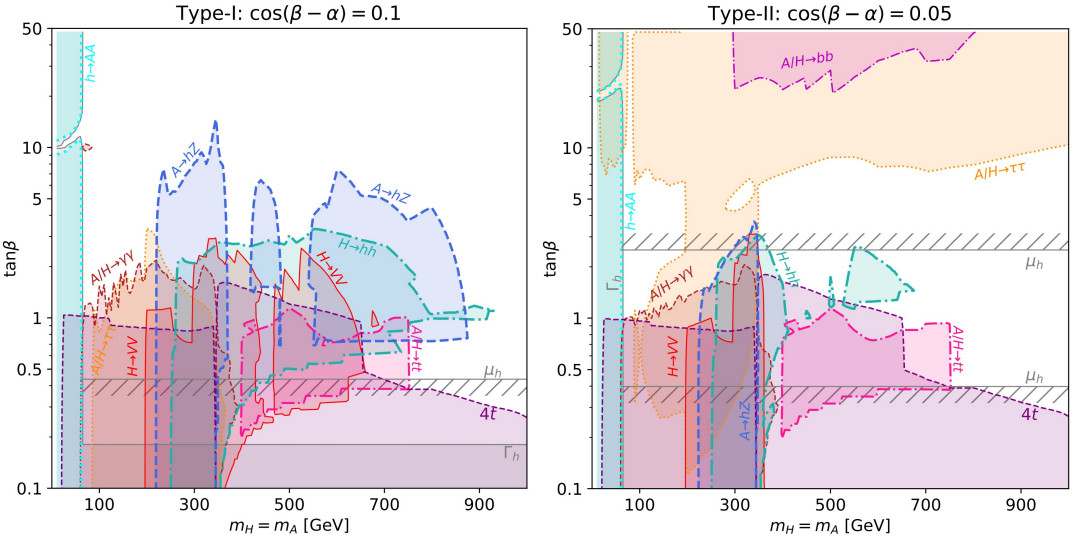


Figure 2.7: The areas excluded by the LHC searches at the 95% confidence level in the tanβ-m_A space for the type-I model at $\cos(\alpha - \beta) = 0.1$ (left) and the type-II model at $\cos(\alpha - \beta) = 0.05$ (right). The area inside the grey lines is excluded by the Standard Model Higgs measurement. The plots are taken from Ref. [27].

searches on the tanβ-m_A/m_H plane. The plots of the type-I model at $\cos(\beta-\alpha)=0.1$ shows that the $A \to Zh$ is one of the most important processes in the 2HDMs searches. However, the plot of the type-II model shows that the process has poor exclusion power at $\cos(\beta-\alpha)=0.05$ where the $A \to \tau\tau$ decay becomes more important because this process is still possible at the alignment limit. The plots also show that the $A \to Zh$ search has strong exclusion power in the m_A range below the $t\bar{t}$ production threshold, and the regions with m_A above $2m_t$ are mainly covered by the $A \to t\bar{t}$ searches. In the plot of the type-I model, some regions above the $t\bar{t}$ production threshold are also excluded by the $A \to Zh$ search. This is because the observed limit is better than the expected limit between 600-800 GeV due to fluctuation, as shown in Figure 2.5. The $A \to hh$ searches have excluded the mass scenario where $m_A < m_h/2$ except for a narrow region at tan$\beta = 10$ where the decay is not possible.

The precision measurement of the Standard-Model Higgs also provides a strong constraint on the 2HDM parameters. Figure 2.8 shows the regions on the tanβ-cos$(\alpha-\beta)$ space excluded by the combined Standard-Model Higgs measurements using the ATLAS data at the centre of mass energy of 13 TeV corresponding to an integrated luminosity up to 79.8 fb^{-1} [28]. The plots show that the current Standard-Model measurements favour the parameter space near the alignment limit where the coupling of the light scalar h is the same as that of the Higgs predicted by the Standard Model. In the type-I model, a wide region surrounding the alignment limit is still allowed. However, this is not the case for the type-II, lepton-specific and flipped models where a narrow region surrounding the alignment limit and a small isolated region at high tanβ are allowed. This is because the coupling strength to light scalar Higgs is not the same for all fermions in the type-II, lepton-specific and flipped models. In the isolated region, the coupling to some fermions is the same as the Standard-Model prediction while the coupling to other fermions is not. The current measurements [28] are not precise enough to exclude those regions.

2.3.3 Heavy Vector Triplets

The experimental search for new physics is difficult because of the presence of a large number of possible models. Moreover, each individual model typically has many free

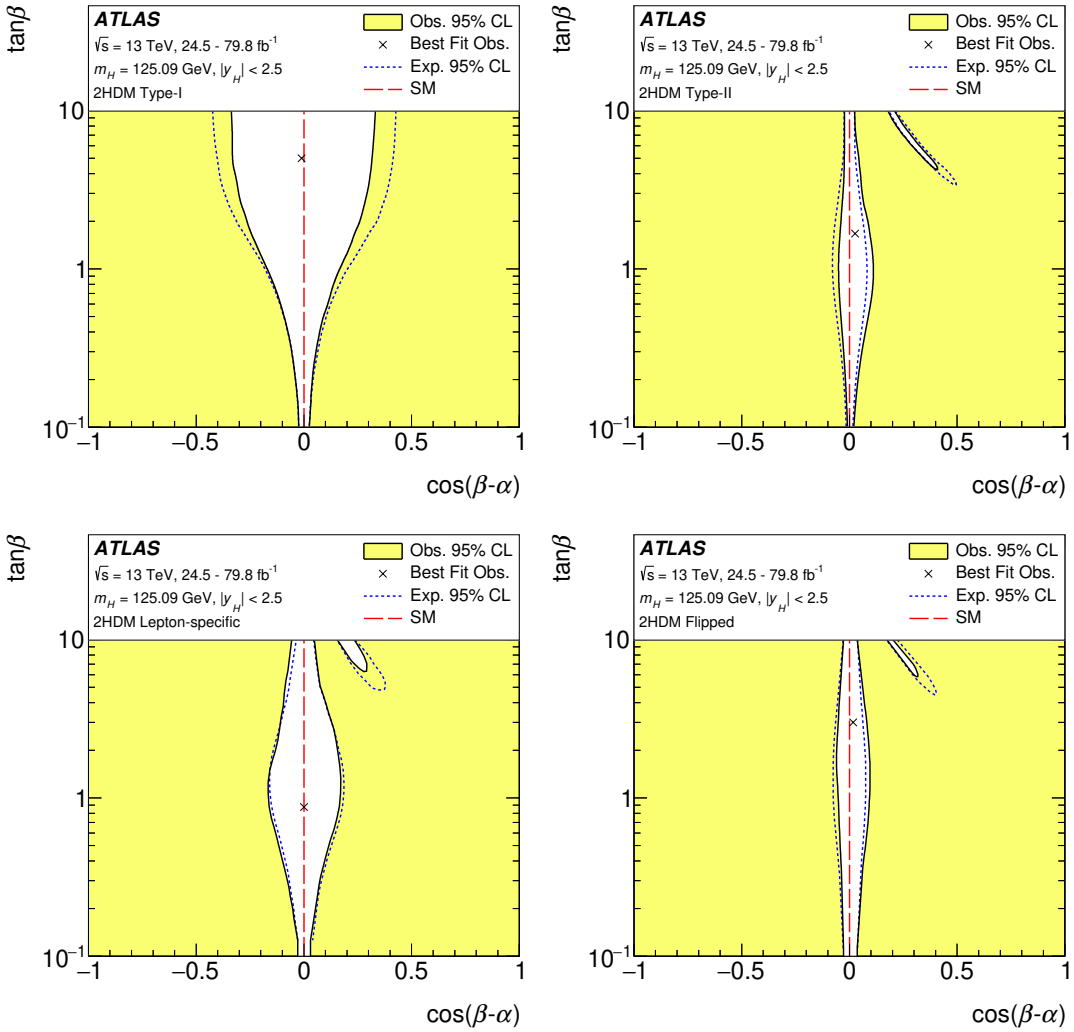


Figure 2.8: The 95% confidence level limits on the 2HDM parameters in the tanβ-cos($\alpha - \beta$) space at $m_H = 125$ GeV for the type-I (upper left), type-II (upper right), lepton-specific (lower left) and flipped models (lower right). The limits are calculated using measurements of the Standard-Model Higgs boson decay by the ATLAS experiment. The plots are taken from Ref. [28].

parameters, and a scan of the full parameter space is difficult. Resonance searches are a simple way to discover new physics because it is only sensitive to the mass and the coupling strength of the new resonance.

The heavy vector triplets model [6] is a phenomenological Lagrangian which can be used to interpret the results of resonance searches. It introduces new heavy vector bosons (W' and Z') that couple to Standard Model particles. Many beyond-the-Standard-Model theories, such as technicolour models [29] [30] and little Higgs models [31], require a new interaction at a high energy scale and predict the existence of the W' bosons and the

Z' boson [19]. The results of HVT resonance searches can be translated into specific models because the free parameters in the HVT Lagrangian can be computed using the details of each model analytically.

In the HVT model, a vector field $V_\mu^a (a = 1, 2, 3)$ of the $SU(2)$ group is introduced and it is described by the Lagrangian [6]

$$\begin{aligned} \mathcal{L} = -\frac{1}{4} D_{[\mu} V_{\mu]}^a D^{[\mu} V^{\mu]a} + \frac{m_V^2}{2} V_\mu^a V^{\mu a} \\ + \frac{i}{2} g_V c_H V_\mu^a H^\dagger \sigma^a (i H^\dagger \sigma^a D^\mu H - i D^\mu H^\dagger \sigma^a H) H + \frac{g^2}{g_V} c_F V_\mu^a J_F^{(\mu a)} \\ + h.c., \end{aligned} \tag{2.48}$$

where g is the Standard Model $SU(2)$ gauge coupling constant, g_V describes the magnitude of the interaction, c_F and c_H are parts of the coupling constants of the heavy vector bosons to the Standard-Model fermions and bosons respectively, $J_F^{(\mu a)}$ is the left-handed fermionic currents and

$$D_{[\mu} V_{\mu]}^a = D_\mu V_\nu^a - D_\nu V_\mu^a, D_\mu V_\nu^a = \partial_\mu V_\nu^a + g \epsilon^{abc} W_\mu^b V_\nu^c. \tag{2.49}$$

The first line of the Lagrangian describes the kinetics of the vector field. The mass m_V is not the physical mass of the particles because V_μ^a is not a mass eigenstate. After diagonalizing the mass matrix, the relationship between the masses of the Standard Model vector bosons (m_Z and m_W) and those of the new vector boson ($m_{W'}$ and $m_{Z'}$) is given by

$$m_W^2 m_{W'}^2 = cos^2 \theta_W m_Z^2 m_{Z'}^2. \tag{2.50}$$

To restore the Standard Model measurement, the Z' boson and the W' bosons must be degenerate in mass. The second line describes the coupling of the Z' boson and the W' bosons to the Standard Model particles. The *h.c.* term in the last line describes the quadratic and cubic coupling of the V_μ^a and it has very little effect on the phenomenology at the LHC.

The most important production modes of the Z' boson and the W' bosons at the LHC are the Drell-Yan process and the vector boson fusion process. In the Drell-Yan

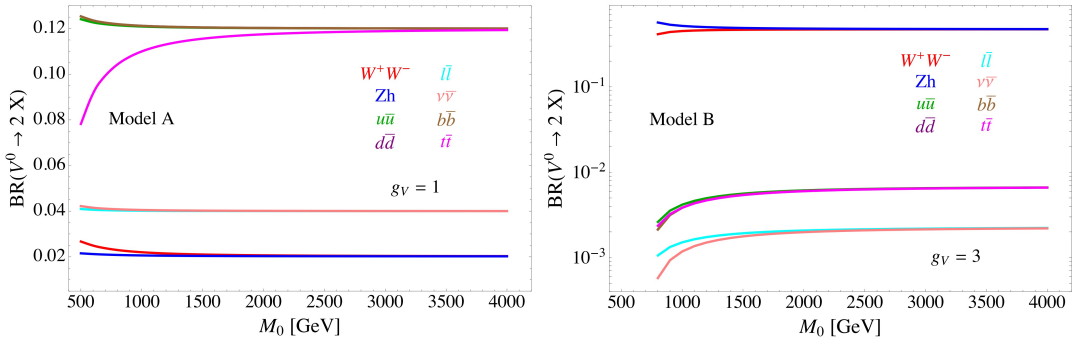


Figure 2.9: The branching ratios of the Z' boson decays in Model A HVT (left) and Model B (right) HVT. The plots are taken from Ref. [6].

process, the Z' boson and the W' bosons are produced via quark-antiquark annihilation. In the vector boson fusion process, the Z' boson and the W' bosons are produced by the fusion of two vector bosons radiated from quarks. The Drell-Yan process is the dominant production mode because the vector boson fusion process is suppressed by the electroweak coupling constant. However, vector boson fusion also becomes important when the fermion coupling vanishes ($c_F \to 0$).

The paper which introduced the HVT model [6] has proposed two benchmark models. In Model A, the coupling of the new vector bosons to the Standard Model fermions and bosons is comparable ($c_H \sim -g^2/g_V^2$ and $c_F \sim 1$). In Model B, the coupling to the Standard Model fermions is suppressed ($c_H \sim c_F \sim 1$). The branching ratios of the Z' boson decaying to Standard Model particles are shown in Figure 2.9. In Model A, all decays are of the same magnitude. In Model B, the decays to Zh and W^+W^- dominate.

The ATLAS Collaboration has searched for the Drell-Yan production of the Z' boson with the Z' boson decaying into W^+W^- [32] or Zh [7]. The upper limits on the cross section are shown in Figure 2.10. The W^+W^- search has probed a mass range of 300 GeV to 5 TeV using the 13 TeV data corresponding to an integrated luminosity of 139 fb^{-1} while the Zh search has probed a mass range of 500 GeV to 5 TeV using the data corresponding to an integrated luminosity of 36.1 fb^{-1}. The observed limits agree with the expected limits in both searches. The W^+W^- search has better sensitivity for both models due to the higher luminosity. It has excluded Model A up to a mass of 3.5 TeV and Model B up to a mass of 3.9 TeV.

The CMS Collaboration has also searched for the two processes [8, 33] using the

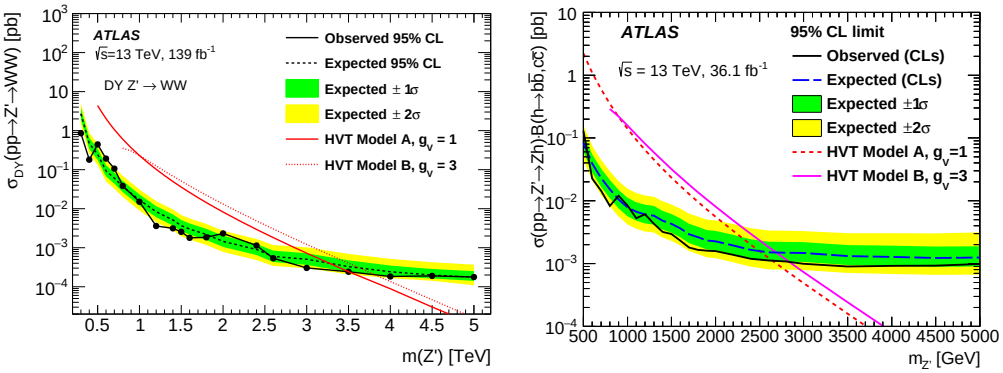


Figure 2.10: The expected and observed upper limits on $\sigma(pp \to Z' \to W^+W^-)$ (left) and $\sigma(pp \to Z' \to Zh) \times BR(H \to b\bar{b}, c\bar{c})$ (right) at the 95% confidence level calculated using the Run 2 data at a centre of mass energy of 13 TeV collected by the ATLAS detector corresponding to an integrated luminosity of 139 fb^{-1} and 36.1 fb^{-1} respectively. The plots are taken from Refs. [7] and [32].

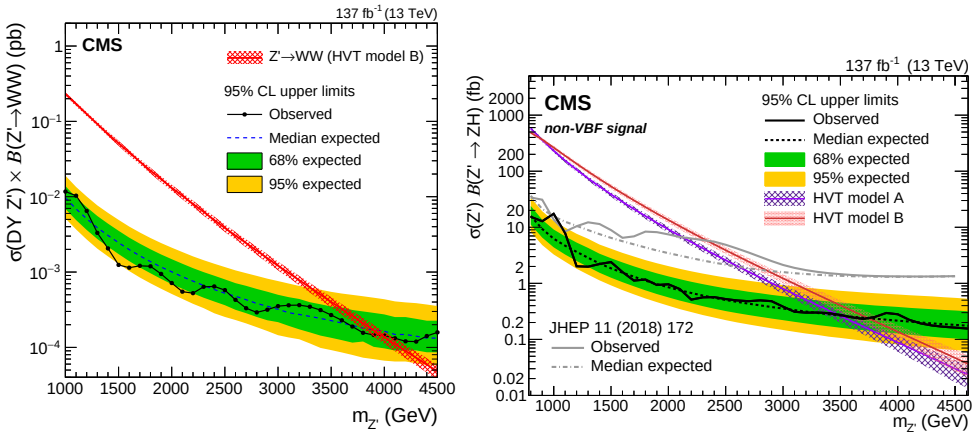


Figure 2.11: The expected and observed upper limits on $\sigma(pp \to Z' \to W^+W^-)$ (left) and $\sigma(pp \to Z' \to Zh)$ (right) at the 95% confidence level calculated using the Run 2 data at a centre of mass energy of 13 TeV collected by the CMS detector corresponding to an integrated luminosity of 137 fb^{-1}. The plots are taken from Refs. [8] and [33].

13 TeV data corresponding to an integrated luminosity of 137 fb^{-1}, which is shown in Figure 2.11. The W^+W^- search covers a mass range of 1 TeV to 4.5 TeV and excludes Model B up to a mass of 4 TeV. The Zh search covers a mass range of 0.8 TeV to 4.6 TeV, and it has excluded Model A up to a mass of 3.5 TeV and Model B up to a mass of 3.7 TeV.

Chapter 3

The ATLAS Experiment

This chapter discusses the experimental setup of the analysis. It first introduces the LHC which accelerates and collides protons. It then describes the ATLAS detector which collects proton-proton collision data.

3.1 The Large Hadron Collider

The LHC [34] is a proton-proton collider operated by the European Organization for Nuclear Research (CERN). Situated on the French-Swiss border near Geneva, it is installed in a 26.7 km circular tunnel lying 45 to 170 m under the ground. It is designed to search for physics beyond the Standard Model by accelerating protons and producing the collision up to a centre of mass energy of 14 TeV. It also accelerates heavy ions to study the quark-gluon plasma.

There are four major experiments at the LHC. The ATLAS [35] and the CMS [36] experiments use general-purpose detectors located at two different collision points to measure Standard-Model processes and search for new physics beyond the Standard Model. The experiments share the same physics goal, but they are operated by two different collaborations independently. The LHCb experiment [37] uses a forward detector close to a collision point for precision measurements and new physics searches in the heavy flavour sector. The ALICE experiment [38] studies the quark-gluon plasma using a detector dedicated to the measurement of heavy ion collisions.

The acceleration of protons happens at several stages [34]. It starts with the Linear Accelerator 2 (LINAC2) where ionized hydrogen is accelerated to 50 MeV. The protons are then accelerated by the Proton Synchrotron Booster, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) to 1.4 GeV, 26 GeV and 450 GeV, respectively. Finally, the beam is injected into the LHC for the final acceleration. In the LHC, the protons travel in opposite directions in two adjacent beam pipes. The protons are accelerated by radio-frequency cavities and they are confined in the LHC ring by superconducting dipole magnets. Quadrupole magnets are also used to collimate the beam and keep it focused for collision.

In the LHC, the protons are grouped into different bunches separated by 25 ns. The number of times (N) that a process happens is

$$N = \sigma \int \mathcal{L}(t)dt, \tag{3.1}$$

where σ is the cross section of the process and $\mathcal{L}(t)$ is the instantaneous luminosity of the collision at time t [10]. The instantaneous luminosity can be written as

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_X\sigma_Y}, \tag{3.2}$$

where f is the frequency of the bunch crossing, n_1 and n_2 are the number of particles in each bunch, and σ_X and σ_Y are the width of the bunches in two orthogonal axes. The account of data the collider produces during a time period T can be described by the integrated luminosity

$$L = \sigma \int_0^T \mathcal{L}(t)dt. \tag{3.3}$$

Run 2 took place during 2015 - 2018, where the LHC collided protons at a centre of mass energy of 13 TeV. The luminosity as a function of time and pileup recorded by the ATLAS detector is shown in Figure 3.1. The LHC delivered data corresponding to a total integrated luminosity of 156 fb^{-1} with 147 fb^{-1} recorded by the ATLAS detector. During data taking, the ATLAS detector sometimes has issues which affect the quality of the data. During the analysis, bad events are filtered out, and the data corresponding

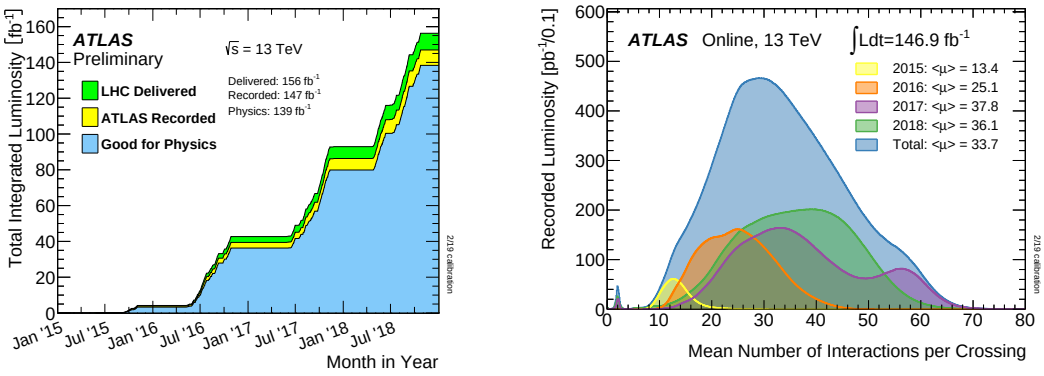


Figure 3.1: The luminosity as a function of time (left) and the pileup recorded by the ATLAS detector (right) during Run 2. The plots are taken from Ref. [39].

to a total integrated luminosity of 139 fb^{-1} has been determined to be good for physics analysis. During 2018, the ATLAS detector recorded the highest average pileup of 36.1 interactions per bunch crossing.

Run 3 [40] is planned to take place during 2022 - 2025, where the LHC will operate at a centre of mass energy of 13.6 TeV and deliver data corresponding to a total integrated luminosity of 300 fb^{-1}. An increased pileup of about 80 interactions per bunch crossing on average is expected.

3.2 The ATLAS Detector

The ATLAS detector [35], which is shown in Figure 3.2, is a general-purpose detector at the LHC that covers almost the full solid angle. The detector consists of several cylindrically symmetric components surrounding the beampipe: the inner detector, the calorimeters and the muon spectrometer. The designed position for proton-proton collisions to take place, also known as the nominal interaction point, is at the centre of the detector. Particles produced by the collision first enter the inner detector where the momenta of charged particles are measured. The particles then enter the calorimeters where they are absorbed by a series of absorbing layers followed by active layers that measure their energy. After that, muons will enter the muon spectrometer, and their momenta and energy will be measured. Neutrinos will finally escape the ATLAS detector and be measured as missing transverse momentum.

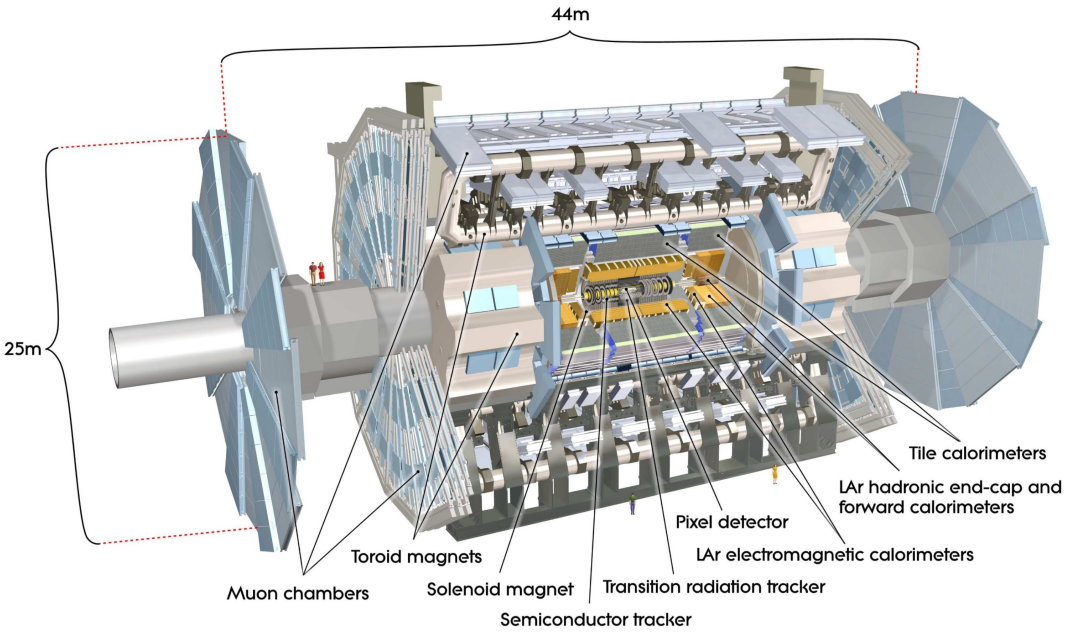


Figure 3.2: Layout of the ATLAS detector. The picture is taken from Ref. [35].

The official definition of the ATLAS coordinate system is used to describe the detector and the kinematics of particles. The origin of the right-handed Cartesian coordinate system is the nominal interaction point, while the z-axis is the direction of the beam. The positive y-axis is perpendicular to the LHC rings and points upward, while the positive x-axis points to the centre of the LHC rings.

Several variables are defined based on the ATLAS coordinate system. The rapidity is defined as:

$$y = \frac{1}{2}\ln[\frac{E + p_z}{E - p_z}], \tag{3.4}$$

where E is the energy of the particle and p_z is the momentum in the z direction. Under Lorentz transformation along the beam direction, the rapidity difference between two particles is invariant, making it a good variable to describe the angular separation in collider physics. However, it is hard to measure rapidity because it requires a precise measurement of E and p_z. As a result, the experiment normally uses a similar quantity called pseudorapidity which is defined as:

$$\eta = -\ln(\tan\frac{\theta}{2}), \tag{3.5}$$

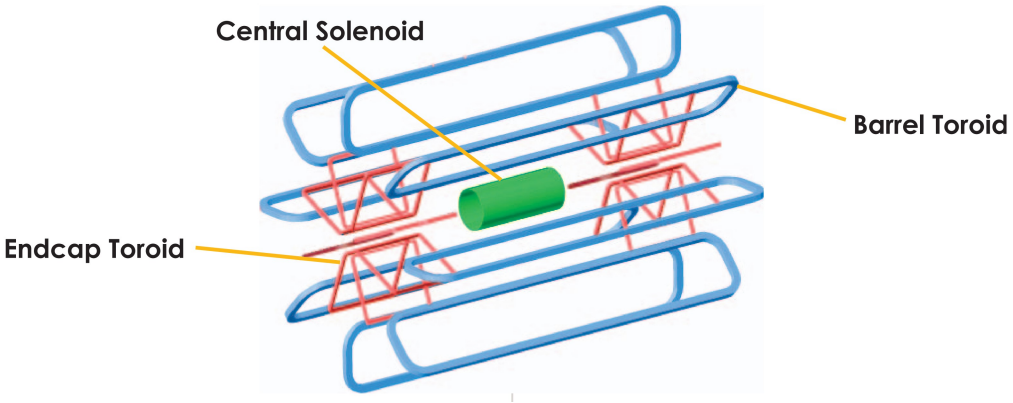


Figure 3.3: Layout of the magnets of the ATLAS detector. The picture is taken from Ref. [41].

where θ is the polar angle. The pseudorapidity is equivalent to the rapidity in the massless limit. The distance between objects is defined as

$$\Delta R^2 = \Delta\eta^2 + \Delta\phi^2, \tag{3.6}$$

where ϕ is the azimuthal angle.

3.2.1 The Magnets

Figure 3.3 shows the magnet system [35] of the ATLAS detector. The magnets generate a magnetic field which enable the detector to determine the momenta of charged particles using their trajectory. The central solenoid provides the magnetic field for the inner detector. With a length of 5.3 m, a diameter of 2.4 m and a thickness of 4.5 cm, it surrounds the inner detector and provides a magnetic field of 2 T. The magnetic field of the muon spectrometer is provided by three toroids in the outermost layer of the detector. The barrel toroid consists of 8 coils and has a length of 25.3 m and a diameter of 20.1 m. It surrounds the muon spectrometer in the barrel region and provides a magnetic field of approximately 0.5 T. The two endcap toroids surround part of the muon spectrometer in the endcap region and provide a magnetic field of approximately 1 T. Each endcap toroid consists of 8 coils and has a length of 5.0 m and a diameter of 10.7 m. The magnets in the ATLAS detector are cooled to approximately 4.7 K to allow superconductivity.

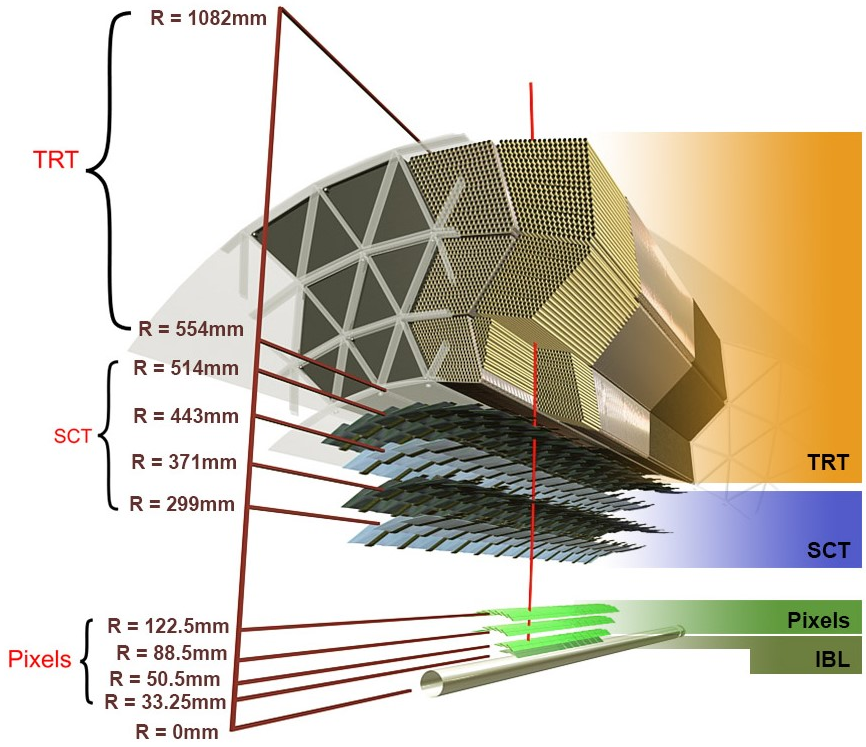


Figure 3.4: The layout of the ATLAS inner detector. The picture is taken from Ref. [35].

3.2.2 The Inner Detector

Figure 3.4 shows the layout of the inner detector (ID) [35]. It consists of four parts: the insertable B-layer (IBL), the pixel detector, the semiconductor tracker (SCT) and the transition-radiation tracker (TRT). The ID is inside a 2 T magnetic field, and the momentum and the charge of charged particles are determined by measuring its trajectory. It provides tracking information for the vertex measurement and contributes to particle identification. The transverse momentum resolution of the ID measurement is

$$\frac{\sigma_{p_t}}{p_t} = 0.05\%\, p_t \text{ GeV} \oplus 1\%, \tag{3.7}$$

where p_t is the transverse momentum of the particle [42].

The pixel detector, the SCT and the IBL are semiconductor detectors. Charged particles create electron-hole pairs as they pass through the semiconductor. The pairs are pulled to the opposite side by an electric field, and the current created can be measured.

The inner-most part of the ID is a pixel layer called the insertable B-layer [43]

which was installed during the first long shutdown of the LHC. It has a pixel size of 50×250 μm^2 in $R-\phi \times z$ and covers a range of $|\eta| < 2.5$. It is installed to compensate for the dead pixels in the pixel detector caused by radiation damage and improve tracking performance.

The layer next to the IBL is the pixel detector which provides the finest granularity after the IBL. It consists of three cylindrical pixel sensor layers in the barrel region and three disks in each endcap region, covering a range of $|\eta| < 2.5$. Each pixel sensor has a minimum pixel size of 50×400 μm^2 in $R-\phi \times z$. The detector has an impact parameter resolution of 10 μm in the $R-\phi$ direction and 115 μm in the z direction.

The next part of the ID is the semiconductor tracker. It consists of 4088 modules of silicon-strip detectors distributed in four cylindrical layers (barrel region) and nine planar discs (each endcap region). It has an accuracy of 17 μm in $R-\phi$ direction and 580 μm in z direction.

The outer-most layer is the transition-radiation tracker which consists of straw tubes filled with xenon-based gas. Charged particles radiate photons when passing through the boundary between two different media. The photons then ionize the gas, which allows measurement. A light particle (like an electron) tends to emit more photons than a heavy particle (like a pion) if they have the same momentum, which enhances the particle identification power of the inner detector. There are 50,000 144 cm long straws in the barrel region and 250,000 39 cm straws in the endcap region, covering a range of $|\eta| < 2$. Unlike other subdetectors, the TRT only provides a $R-\phi$ direction measurement with a resolution of 130 μm.

3.2.3 The Calorimeters

Figure 3.5 shows the ATLAS calorimeters [35]. They consist of electromagnetic calorimeters and hadronic calorimeters, covering a range of $|\eta| < 4.9$. The electromagnetic calorimeters measure the energy of electrons and photons, while the hadronic calorimeters measure the energy of hadrons. The muons and neutrinos generally escape the calorimeters, while most other particles are totally absorbed by the calorimeters. Within the inner detector measurement range, the high resolution of the electromagnetic calorimeter

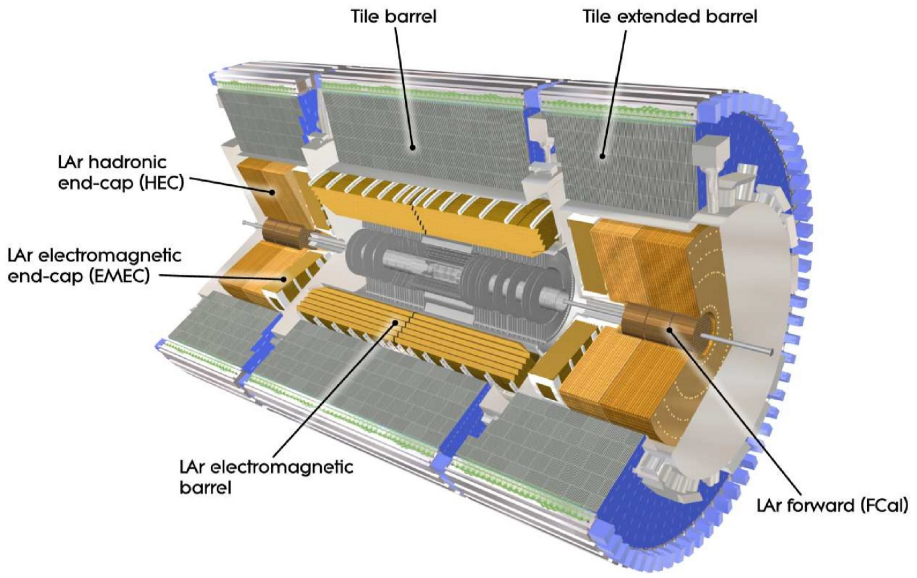


Figure 3.5: The layout of the ATLAS calorimeters. The picture is taken from Ref. [35].

enables the precision measurement of energy for electrons and photons. Beyond the inner detector measurement range, the rest of the calorimeter is used for jet reconstruction and missing transverse energy measurement.

The electromagnetic calorimeter is a sampling calorimeter with liquid argon as the active medium and lead as the absorbing material. Incoming electrons interact with the absorbing medium via bremsstrahlung, and the photons radiated produce electron-positron pairs. Incoming photons also produce electron-positron pairs. The process continues and produces electromagnetic showers. The shower interacts with the active medium via ionization, which allows the energy to be measured. The ATLAS electromagnetic calorimeter consists of a barrel part ($|\eta| < 1.475$) and two endcap parts ($1.375 < |\eta| < 3.2$) with the same energy resolution [44] of

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%. \tag{3.8}$$

The innermost layer of the electromagnetic calorimeter is a LAr presampler layer which is used to correct the energy loss before the particle reaches the calorimeters. It covers a range of $|\eta| < 1.8$ and has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$. The front layer is next to the presampler layer. It has a granularity of $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$ in the

barrel region and the granularity varying from $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$ to 0.025×0.1 in the endcap region. The front layer has the finest granularity to identify photons from pion decays. The middle layer of the electromagnetic calorimeter absorbs most of the energy, and it has a granularity varying from $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ to 0.1×0.1. The back layer has a granularity of $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$. Its main purpose is to absorb the remaining energy and help distinguish between electromagnetic showers and hadronic showers.

Surrounding the electromagnetic calorimeter is the hadronic calorimeter. Incoming hadrons interact with the nuclei of the detector via the strong interaction. Charged hadrons also lose energy via ionization. The particles produced in the process continue to interact with the detector and create hadronic shows. The ATLAS hadronic calorimeter has several different components. The tile calorimeter is a sampling calorimeter using steel as an absorbing material and scintillating tiles as the active medium, covering a range $|\eta| < 1.7$. The endcap region ($1.5 < |\eta| < 3.2$) is covered by the LAr hadronic endcap. Both the LAr hadronic endcap and the tile calorimeter have an energy resolution [45] of

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%. \tag{3.9}$$

The forward region ($3.1 < |\eta| < 4.9$) is covered by the LAr forward calorimeter with an energy resolution [45] of

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%. \tag{3.10}$$

The hadronic calorimeter has a coarser granularity than the electromagnetic calorimeter because the hadronic shower is normally wider than the electromagnetic shower.

3.2.4 The Muon Spectrometer

The outmost part of the ATLAS detector is the muon spectrometer [35], which is shown in Figure 3.6. Muons' Bremsstrahlung radiation in the calorimeter is suppressed due to the high mass, so they generally escape the calorimeter and enter the muon spectrometer. The muon spectrometer is inside a magnetic field which bends the trajectories of muons and enables the measurement of momenta. The magnetic field is provided by the large

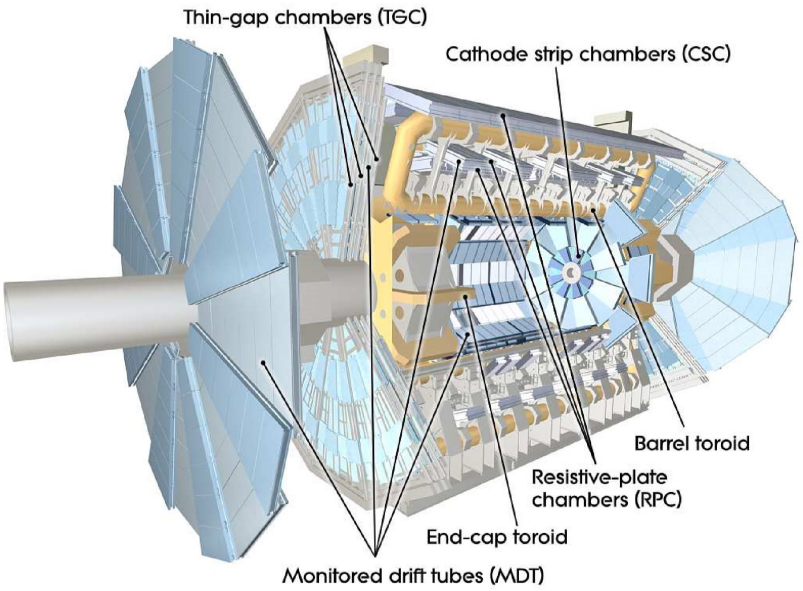


Figure 3.6: The layout of the ATLAS muon spectrometer. The picture is taken from Ref. [35].

barrel toroid in the region $|\eta| < 1.4$ and two endcap magnets in the region $1.6 < |\eta| < 2.7$. The magnetic field in the region $1.4 < |\eta| < 1.6$ is provided by both magnet systems.

The ATLAS muon spectrometer consists of individual muon chambers distributed in four subsections: monitored drift tubes, cathode strip chambers, resistive plate chambers and thin gap chambers. Muons passing through the chambers ionize the gas and produce electrons and ions which can be measured. Monitored drift tubes filled with argon and carbon dioxide cover a range $|\eta| < 2.7$. In the forward region ($2.0 < |\eta| < 2.7$), cathode strip chambers filled with carbon tetrafluoride, argon and carbon dioxide gas are installed to handle the high particle rate because they have good spatial resolution and response time. Monitored drift tubes and cathode strip chambers are both used for precision tracking.

The resistive plate chambers and the thin gap chambers are installed to provide information for triggering due to their fast response time. The resistive plate chambers cover the barrel region of $|\eta| < 1.05$, while the thin gap chambers cover the forward regions of $1.05 < |\eta| < 2.4$.

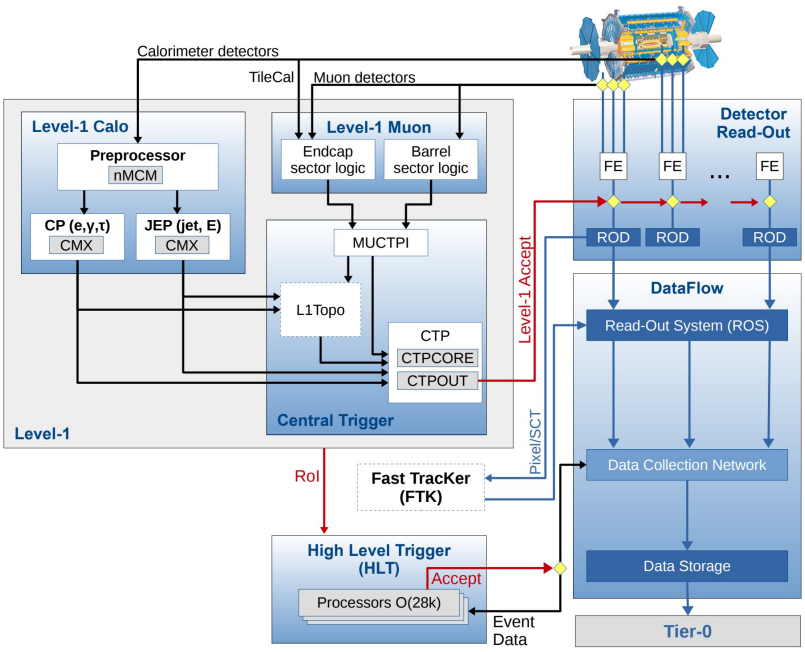


Figure 3.7: The ATLAS trigger system during Run 2. The picture is taken from Ref. [46].

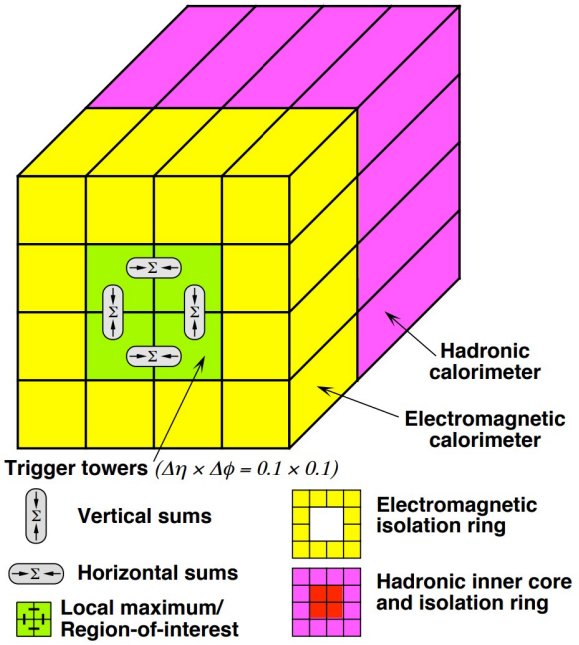


Figure 3.8: The ATLAS L1 calorimeter trigger tower during Run 2. The picture is taken from Ref. [47].

3.2.5 The Trigger System

The LHC produces a large number of collision events, and data collected by the ATLAS detector has to be preselected by the trigger system before storage. The ATLAS trigger system used during Run 2 [46] is shown in Figure 3.7. It consists of the level-1 (L1) trigger [48] and the High Level Trigger (HLT) [49].

Events are first selected by the hardware-based L1 trigger. The L1 calorimeter (L1Calo) trigger [47] identifies objects of interest using the energy deposit in the calorimeters. It looks for calorimeter objects with transverse energy (E_T) above certain predefined thresholds. The analogue signal from the calorimeters is first digitized and calibrated by the preprocessor. The preprocessor's output is then processed by the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP). The CP identifies isolated e/γ and τ objects while the JEP identifies jets, the total transverse momentum and the missing transverse momentum.

The CP algorithms [47] search for objects of interest by scanning all possible 4×4 trigger tower windows. Each trigger tower contains the E_T with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the electromagnetic layer and the hadronic layer, shown in Figure 3.8. To identify isolated objects, the sum of E_T in the 12 cells surrounding the central 2×2 cells in the electromagnetic layer is required to be less than a certain threshold. For e/γ objects, the sum of E_T in the hadronic layer is required to be less than certain thresholds to make sure that the energy is contained in the electromagnetic layer. The isolation is not required if the E_T of the object passes a certain predefined threshold. The sum of E_T in the central 2×2 towers is required to be a local maximum to avoid double counting. In Run 3, the e/γ and τ algorithms have been upgraded to make use of trigger towers with a finer granularity to handle the increased luminosity and pileup. More details about the Run 3 L1Calo system will be discussed in Chapter 4.

The L1 muons trigger (L1Muon) estimates the muons' transverse momenta using the signals from the Thin Gap Chambers and Resistive Plate Chambers. The information about the muon candidates is sent to the Muon-to-Central Trigger Processor Interface (MUCTPI) which counts the numbers of muon candidates passing each of the predefined transverse momentum thresholds.

The Level-1 Topological Processor (L1Topo) receives information from the L1Calo and the L1Muon and makes decisions using global objects such as the missing transverse energy and invariant mass. The final decision of the L1 trigger is made by the Central Trigger Processor (CTP).

L1 triggers reduce the event rate from 40 MHz to about 100 kHz. After the L1 trigger selection, events are further selected by the High Level Trigger. The HLT is software based and runs on CPU farms. It receives Regions of Interest (RoI) in the $\eta - \phi$ space identified by the L1 trigger system and performs regional reconstruction of physics objects using the full detector granularity. The event rate is reduced to about 1 kHz after the HLT selection. Events selected by the HLT are stored for further analysis.

The thresholds applied to objects by the L1 trigger and the HLT are defined in the trigger menu. The ATLAS experiment uses different sets of trigger menus. The primary triggers are used for physics study, while others are used for efficiency measurements and detector calibrations. The Run 2 trigger menus used during 2015, 2016, 2017 and 2018 are summarized in [50], [51], [52] and [53] respectively.

3.3 The ATLAS Phase-I Upgrade

In Run 3, the LHC will deliver data corresponding to a total integrated luminosity of 300 fb^{-1} with the expected average pileup of about 80 interactions per bunch crossing [40]. The detector performance required in Run 3 is beyond the original design specifications. In order to make the ATLAS detector operate efficiently under the new condition, the Phase-I upgrade of the ATLAS detector took place during the long shutdown 2 (LS2) of the LHC.

The L1 muon trigger rate is proportional to the luminosity [54], and the trigger rate expected during Run 3 exceeds the capacity of the ATLAS Trigger and Data Acquisition (TDAQ) system. The Run 2 muon trigger relies on the Big Wheel (thin gas chambers) to trigger on muons in the forward regions. However, studies [55] show that the method has a high fake rate (about 90%) due to radiation generated by the material between the big wheels and the small wheels. The current muon small wheels consist of cathode strip

chambers and monitored drift tubes. To reduce the trigger rate without losing events which are important for physics analysis, a new small wheel [55] with a faster response covering the region $1.3 < |\eta| < 2.7$ has replaced the old small wheel to provide enhanced information for the trigger. With the information from both the new small wheel and the big wheel, the muon trigger will be able to reconstruct the tracks of muons and reject particles which have not originated from the collision point.

During the LS2, the electronics of the LAr calorimeters have been upgraded [56] to provide trigger towers with a finer granularity in the L1Calo. The new information from the LAr calorimeters is crucial for reducing the L1Calo trigger rate without damaging physics results. With the upgraded system, the L1Calo will make use of LAr readouts of the granularity up to $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$. More details about the Phase-I upgrade of the L1Calo will be discussed in Chapter 4.

The TDAQ system has also been updated [57] during LS2. The new trigger algorithms have been designed and implemented to make use of the new readout from the upgraded LAr calorimeters and muon system. The hardware has also been upgraded with the latest technology to increase processing power and data transmission speed.

Chapter 4

The Level-1 Calorimeter Trigger Offline Simulation

This chapter discusses the L1Calo offline simulation for the Phase-I upgrade of the ATLAS TDAQ system [57]. The L1Calo offline simulation is a C++ module which is part of the ATLAS offline software framework (Athena) [58]. It processes the digitized signals from the calorimeters and identifies regions of interest which are then processed by other parts of the trigger system. The L1Calo offline simulation software will be part of the ATLAS detector simulation for the Monte Carlo events production. It also works as a tool to monitor the L1Calo system and validate the L1Calo online simulation and the firmware.

This chapter focuses on the simulation of the Electron Feature Extractor (eFEX) module [59]. The eFEX module is responsible for the clustering and identification of e/γ and τ candidates using sliding-window algorithms. The algorithms calculate jet rejection variables which discriminate between $e/\gamma/\tau$ and jets.

This chapter will first introduce the Run 3 L1Calo system, the eFEX algorithms and the Athena framework. It will then describe the structure of the L1Calo eFEX offline simulation software. After that, it will focus on the software modules designed and developed by the author. Finally, some plots created using the outputs of the framework will be shown.

It should be noted that Section 4.4 and Section 4.5 describe the status of the software

at the time when the author was actively involved in the development. The structure of the simulation is still subject to change as the development work continues.

4.1 Overview of the Run 3 L1Calo System

Figure 4.1 shows the architecture of the upgraded L1Calo system for Run 3 [60]. The L1Calo system accepts digital signals from the liquid-argon (LAr) calorimeter and the tile calorimeter (TileCal). The analogue signals from the LAr calorimeter are digitized by the LAr Trigger Digitizer Boards (LTDB) and then transmitted to the LAr Trigger processing Mezzanine (LATOME) cards where the transverse energy (E_T) is computed. The analogue signals from the TileCal are digitized by the Tile Rear Extension (TREX) modules installed in the Pre-processor.

During Run 2, the L1Calo system has processed analogue trigger tower signals using the Jet Energy Processor and the Cluster Processor. In Run 3, trigger towers are broken down further into SuperCells to provide LAr calorimeter signals with a finer granularity [60]. A SuperCell is a combination of four or eight LAr calorimeter cells. Three new Feature Extractor (FEX) modules are being installed to process the SuperCell inputs. The eFEX identifies e/γ and τ candidates. The Jet Feature Extractor (jFEX) identifies candidate jets, τ, missing transverse energy, sums of transverse energy and forward electrons. The Global Feature Extractor (gFEX) identifies candidate large-R jets as well as missing transverse energy and transverse energy sums. The optical signals from the LATOME and the TREX are routed to the three FEX modules by the Fiber Optic Exchange (FOX) module.

The three FEX modules produce Trigger Object (TOB) words containing information about the selected candidates. They are sent to the Level-1 Topological Trigger Processor (L1Topo) for further processing. The modules used during Run 2 will work in parallel with the new system at the beginning of Run 3 to validate the FEX outputs, and they will be removed once the new system is validated.

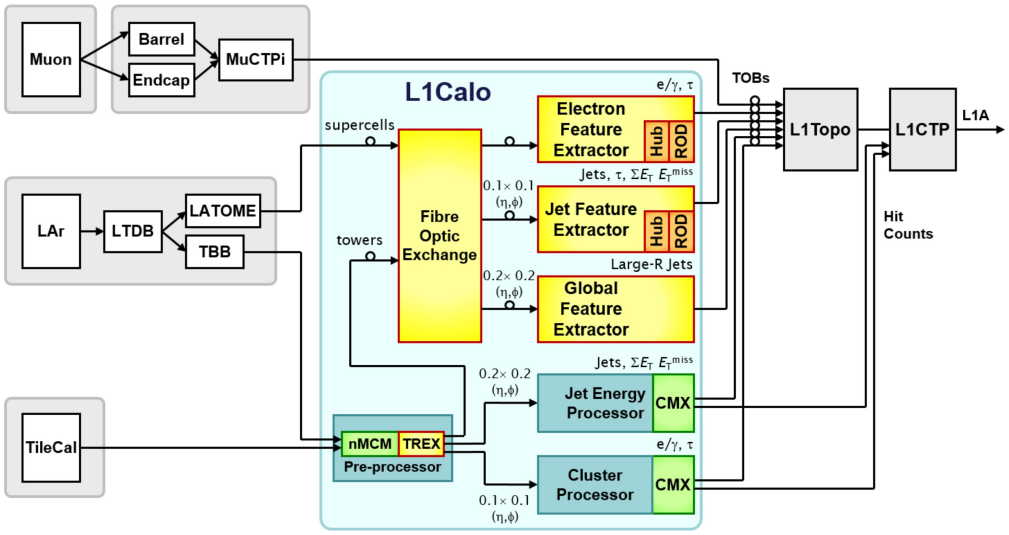


Figure 4.1: Overview of the ATLAS L1Calo system for Run 3. The picture is taken from Ref. [60].

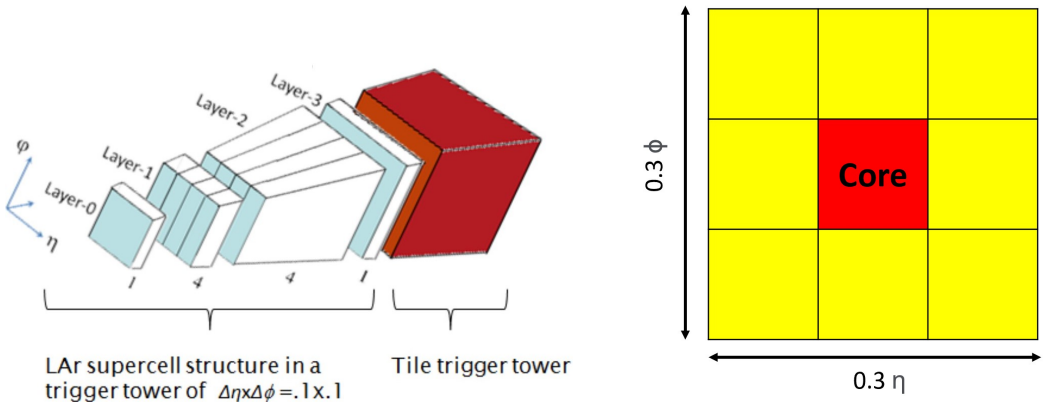


Figure 4.2: Layouts of a trigger tower (left) and a 3×3 tower window for the algorithms (right). The picture on the left is taken from Ref. [59].

4.2 The eFEX Algorithms

This chapter focuses on the eFEX module which uses sliding-window algorithms to search for isolated e/γ or τ energy deposits in the calorimeters. The algorithms process SuperCells contained in trigger towers which are shown in Figure 4.2. Covering a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, each trigger tower consists of four electromagnetic layers and one hadronic layer, where the layer-0 and the layer-3 each contain one SuperCell from the LAr electromagnetic calorimeter, the layer-1 and the layer-2 each contain four SuperCells from the LAr electromagnetic calorimeter, and the hadronic

layer contains one tile trigger tower from the tile calorimeter or one SuperCell from the LAr hadronic calorimeter. For simplicity, we treat both the tile trigger towers and the LAr hadronic calorimeter SuperCells as hadronic layer SuperCells in the rest of the chapter, unless specified otherwise.

The algorithms [60] process the E_T of SuperCells in a 3×3 tower window shown in Figure 4.2, where the red trigger tower at the centre is the primary core tower. To identify isolated e/γ energy deposits, the algorithms first look for the primary seed which is the layer-2 SuperCell with the highest E_T in the primary core tower. It then compares the E_T of two adjacent SuperCells in the ϕ direction, and the one with a greater E_T is defined as the secondary seed. The trigger tower where the secondary seed is located is the secondary core tower. The algorithms will calculate the cluster energy which is defined as

$$E_{\text{Cluster}} = E_{03} + E_{12}, \tag{4.1}$$

where E_{03} is the E_T sum of all layer-0 and layer-3 SuperCells in the primary and the secondary core towers, and E_{12} is the E_T sum of the layer-1 and layer-2 SuperCells in a $\Delta\eta \times \Delta\phi = 0.75 \times 0.2$ area around the primary and the secondary seeds. Finally, the algorithms will calculate three jet rejection variables.

The authors' work involves implementing a jet rejection variable called total cluster width which is defined as:

$$w_s^2 = \frac{\sum E_{ij}(i - i_{\text{seed}})^2}{\sum E_{ij}}, \tag{4.2}$$

where $\sum E_{ij}$ is the E_T sum of all SuperCells in a $\Delta\eta \times \Delta\phi = 0.125 \times 0.3$ area centred at the primary seed, i and j are the SuperCell indices in η and ϕ directions respectively, and i_{seed} is the η index of the primary seed. The value $\sum E_{ij}(i - i_{\text{seed}})^2$ is the E_T sum weighted by the squared distance between a SuperCell and the primary seed in the η direction. The weights $(i - i_{\text{seed}})^2$ are 0, 1 and 4 when the η distance is 0, 0.25 and 0.5 respectively. The e/γ objects produce a narrower shower than jets, so tower windows with w_s^2 below certain thresholds are selected.

The other two jet rejection variables for the identification of e/γ objects are the

hadronic condition and the R_η isolation. The hadronic condition R_{had} is defined as

$$R_{\text{had}} = \frac{E_{\text{had}}}{E_{\text{had}} + E_0 + E_1 + E_2 + E_3}, \tag{4.3}$$

where E_{had} is the sum of all hadronic layer E_T in a tower window, E_0 and E_3 are the E_T sums in $\Delta\eta \times \Delta\phi = 0.1 \times 0.3$ areas centred around the primary core tower in the layer-0 and the layer-3 respectively, and E_1 and E_2 are the sums of E_T in $\Delta\eta \times \Delta\phi = 0.075 \times 0.3$ areas centred around the primary seed in the layer-1 and the layer-2 respectively. The electrons and photons are expected to deposit most of their energy in the electromagnetic calorimeter, so trigger-tower windows with R_{had} below certain thresholds are selected.

The R_η is defined as

$$R_\eta = \frac{E_{\text{E2}} - E_{\text{C2}}}{E_{\text{E2}}}, \tag{4.4}$$

The E_{C2} is the cluster energy calculated with only layer-2 SuperCells, and E_{E2} is the E_T sum of all layer-2 SuperCells in a $\Delta\eta \times \Delta\phi = 0.175 \times 0.3$ area centred around the primary seed. Isolated electrons and photons are expected to have a low R_η, so trigger-tower windows with R_η below certain thresholds are selected.

The thresholds for the three jet rejection variables are optimized by the L1Calo Trigger and Performance group, and they will be provided before the start of Run 3. The requirement of jet rejection variables are only applied when $E_{\text{Cluster}} < 60$ GeV. Objects with high E_{Cluster} automatically pass the selection.

The selected tower windows are delivered to the L1Topo in the form of TOB words. The counts of numbers of the thresholds satisfied are stored in the TOB word as integers for the three jet rejection variables separately. The thresholds are assumed to be in increasing order. Each TOB word also contains data related to the η, ϕ location of the tower window, Cluster Energy, seed location and information about whether the primary seed is the most energetic supercell in the tower.

There are 24 eFEXs covering all SuperCells from the calorimeters, and each eFEX processes an area of $\Delta\eta \times \Delta\phi = 1.8 \times 1.0$. Each eFEX consists of four Field Programmable Gate Arrays (FPGA) [61] with each one processing trigger towers in an area of $\Delta\eta \times \Delta\phi = 0.6 \times 1.0$ [60].

4.3 The Athena Framework

The ATLAS experiment uses the Athena framework [58] for event reconstruction, detector simulation and physics analysis. The Athena framework is based on the GAUDI framework [62] which is also used by the LHCb experiment. The framework uses an object-oriented methodology and the code is mainly based on C++.

The framework is component-based. A component is a C++ object derived from a variety of predefined abstract classes. Components share a group of standard interfaces, which allows them to be controlled easily. The Athena framework contains many predefined components. Some of them are commonly used for physics analysis, while others are utilities for code development. Developers can build applications from predefined components, but they can also design their own components based on various base classes. The most important types of Athena components for the development of the offline simulation are:

- **Algorithms**. An Athena Algorithm is a C++ class derived from the algorithm base classes. Each Algorithm must contain three methods: *initialize()*, *execute()* and *finalize()*. The *initialize()* and the *finalize()* methods are only executed once at the beginning and the end of an Athena application. The *execute()* contains the actual analysis code, and it is called once for every single event in the dataset.

- **Tools**. An Athena Tool is a part of code which operates on data. It is called inside Algorithms. It may be called more than once per event. Unlike Algorithms, Tools do not have a standard interface, which allows it to be customized. However, it normally contains an *initialize()* method called at the beginning of the application.

- **Services**. Services are utilities that facilitate code development. It performs tasks such as displaying error messages, accessing data and creating output files.

An Athena application is configured by the Athena JobOption. A JobOption is a Python script which lists all the Algorithms, Tools and Services required by the analysis. It specifies the execution order of the Algorithms. It may also contain application settings such as output file names. Athena applications are managed by the Athena Application

Manager. The Application Manager creates and executes Algorithms, Tools and Services according to the JobOption.

The ATLAS experiment processes a large amount of data and simulation, and the offline software must run on more than one processor to reduce computation time. Running separate Linux processes is not practical because of the large memory consumption of Athena applications. As a result, objects such as Tools and Services are shared between different threads once created to save memory [63].

4.4 Overview of the L1Calo eFEX Offline Simulation Software

The L1Calo offline simulation software is an object-oriented C++ package of the Athena framework. It is designed to be efficient and thread-safe. The simulation's input is the Athena *CaloCellContainer* and *TriggerTower* objects which are collections of Athena SuperCell and tile trigger tower objects. The code contains independent simulations for the eFEX, jFEX and gFEX. The chapter will focus on the simulation of the eFEX, but the simulation of the jFEX and gFEX also has a similar code structure. The Run 2 legacy modules (JEP and CP) are not included in the simulation. The offline simulation software can be incorporated into the Athena detector simulation, but it can also run independently for L1Calo related studies.

The simulation follows the structure of the firmware closely. The eFEX simulation can be divided into the mapping simulation (*eSuperCellTowerMapper* class and *eFEXSysSim* class), the individual eFEX board simulation (*eFEXSim* class), the individual FPGA simulation (*eFEXFPGA* class), the algorithm simulation (*eFEXegAlgo* class and *eFEXtauAlgo* class) and the TOB output simulation (*eFEXegTOB* class and *eFEXtauTOB* class). It also contains the *eFEXNtupleWriter* Athena Algorithm which monitors the simulation for debugging purposes. The simulation software use the StoreGate Service [64] to store objects in RAM, which makes the object accessible to all classes in the simulation.

Every eFEX module, trigger tower and SuperCell object in the simulation can be

uniquely determined by its ID. The trigger tower ID is a 6-digit decimal integer related to the trigger tower's η and ϕ position in the detector. A SuperCell object can be uniquely identified by an unsigned 64-bit integer hash ID. Each FPGA has an integer ID goes from 0 to 3, and it can also be uniquely determined by pairing with the ID of the eFEX (an integer from 0 to 23) to which it belongs.

The Unified Modeling Language (UML) [65] sequence diagram in Figure 4.3 shows the structure of the eFEX simulation. The simulation is controlled by an Athena JobOptions Python script. For each event, the main part of the simulation, the *eFEXDriver* Athena Algorithm, is called first to load inputs and perform the L1Calo eFEX simulation. After that, an optional Athena Algorithm, *eFEXNtupleWriter*, can be called to save monitoring data for debugging.

At the beginning of the simulation, the *eFEXDriver* first builds an *eTowerContainer* object which is a collection of all trigger towers in the calorimeters. The simulation uses the *eTower* objects to represent individual trigger towers, and the read and write methods for the E_T of SuperCells with noise threshold cut are available in the *eTower* class.

The SuperCell E_T needs be stored in the *eTowerContainer* object because the algorithms will access it to obtain SuperCell E_T in each tower window. The SuperCell E_T of the *eTowerContainer* object is assigned by the *eSuperCellTowerMapper* class where the mapping between SuperCells and trigger towers is made. The SuperCell E_T from the LATOME is stored in the Athena *CaloCellContainer* object, and the mapping is made by associating the IDs of the SuperCells with the trigger tower's position. The description of the LATOME mapping code is available in Section 4.5. The E_T delivered by the TREX is stored in the Athena *TriggerTowerContainer* object, and the mapping is made using the position of SuperCells and that of trigger towers. At the current stage of the development, the Run 2 L1Calo offline simulation module of the Athena framework has to be executed to produce the *TriggerTowerContainer* object prior to the Run 3 simulation. In the future, a new module could be implemented to produce the *TriggerTowerContainer* without the Run 2 simulation. Both the *CaloCellContainer* object and the *TriggerTowerContainer* object are obtained using the StoreGate Service.

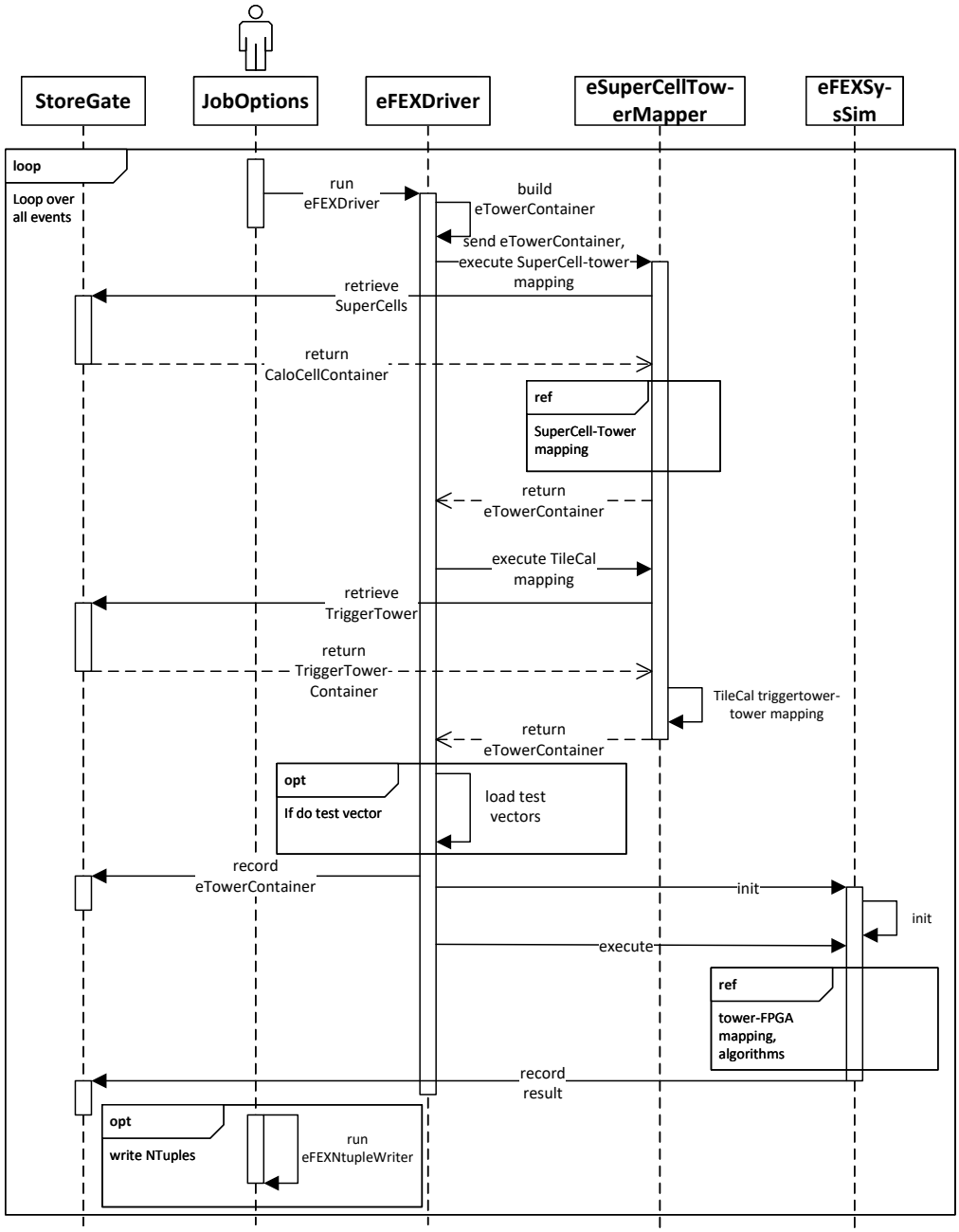


Figure 4.3: UML sequence diagram for the L1Calo eFEX offline simulation software. The reference fragments are in Figure 4.4 and Figure 4.5.

After the SuperCell-tower mapping, an optional step could happen in the *eFEXDriver* where the SuperCell E_T is replaced by an alternative dataset known as the test vector. More details will be discussed in Section 4.5.

After the mapping simulation, the *eTowerContainer* object is stored in RAM using

the StoreGate Service, making it available to all the other classes in the simulation. The *eFEXDriver* then calls the *eFEXSysSim* object to run the tower-FPGA mapping, the eFEX simulation and the algorithm simulation. The *eFEXegTOB* objects produced by the algorithm simulation are stored in RAM via the StoreGate Service. At the end of the simulation, the *eFEXNtupleWriter* could be called to save monitoring data. More details about the *eFEXSysSim* and the *eFEXNtupleWriter* are available in Section 4.5.

4.5 Technical Design of the Mapping, Testing and Validation Modules

The authors' work in the development of the offline simulation software involves the mapping, testing and validation modules. This chapter presents a detailed technical description of the relevant codes.

4.5.1 The Mapping Modules

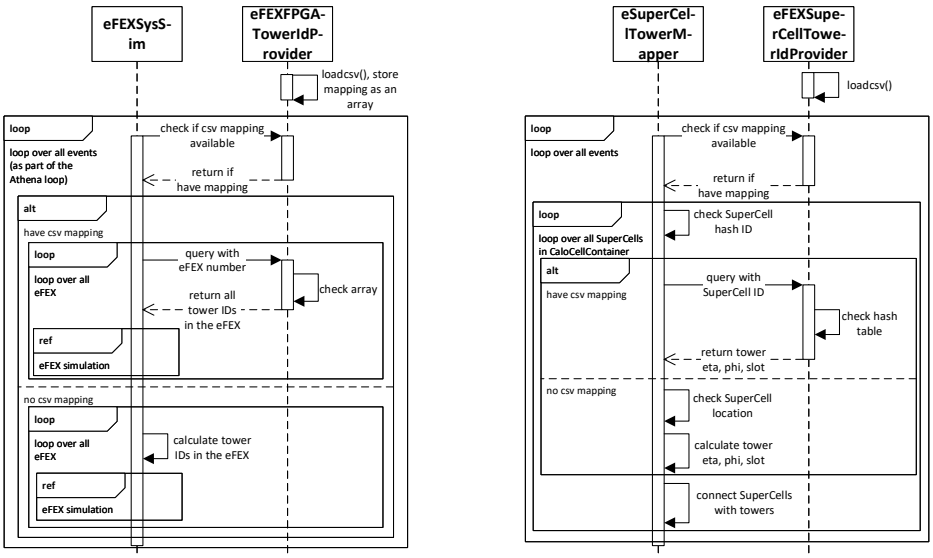


Figure 4.4: UML sequence diagrams for the tower-FPGA mapping code (left) and the SuperCell-tower mapping code (right).

The L1Calo offline simulation uses SuperCells from the LATOME and the TREX as input, and mapping is required to assign the SuperCell E_T to the trigger towers and

make connections between the trigger towers and the FPGA modules. The mapping can be done by hard-coded calculation using the η and ϕ position of the SuperCells, trigger towers and FPGAs. It can also be done using mapping data stored in Comma-Separated Values (CSV) files, allowing changes to be implemented easily in case of possible mapping changes in the future.

The classes which load the CSV files are Athena Tool components, which allows the mapping class to be managed by the Athena ToolHandles [66]. The advantage of using ToolHandles is that the mapping classes are created and managed by the Athena Application Manager, which reduces memory consumption because the mapping object is shared among all threads after its creation. The dedicated mapping classes also improve the readability of the code.

The UML sequence diagram in Figure 4.4 describes the FPGA-Tower mapping code. The mapping class *eFEXFPGATowerIdProvider* loads the CSV file and provides the IDs of all the trigger towers in an eFEX or an FPGA. The *eFEXFPGATowerIdProvider* class also sorts the IDs to match their position in the L1Calo firmware. The CSV file contains information about a trigger tower's ID, η, ϕ and the eFEX and FPGA to which it belongs. The CSV file is preloaded before iterating through all the events. During the preloading, the tower IDs in every FPGA are sorted using the η and ϕ values and stored in RAM as a referential array. The preloading allows the simulation to search for a trigger tower ID using the eFEX ID and the FPGA ID with $\mathcal{O}(1)$ time complexity. The mapping information will later be utilized by the *eFEXSysSim* object if the CSV mapping is available; otherwise, the mapping would be determined by the hard-coded calculation.

The CSV file for the SuperCell-tower mapping contains the SuperCell IDs and the IDs of the trigger towers to which they belong. Before iterating through all the events, the *eFEXSuperCellTowerIdProvider* object loads the trigger tower IDs and stores the mapping information in RAM as a hash table with the SuperCell ID as the key. The mapping information is then utilized by the *eSuperCellTowerMapper* object to assign SuperCell E_T to trigger towers. The *eSuperCellTowerMapper* object will use the hard-coded calculation if the CSV file is not available.

4.5.2 The eFEX, FPGA and Algorithm Modules

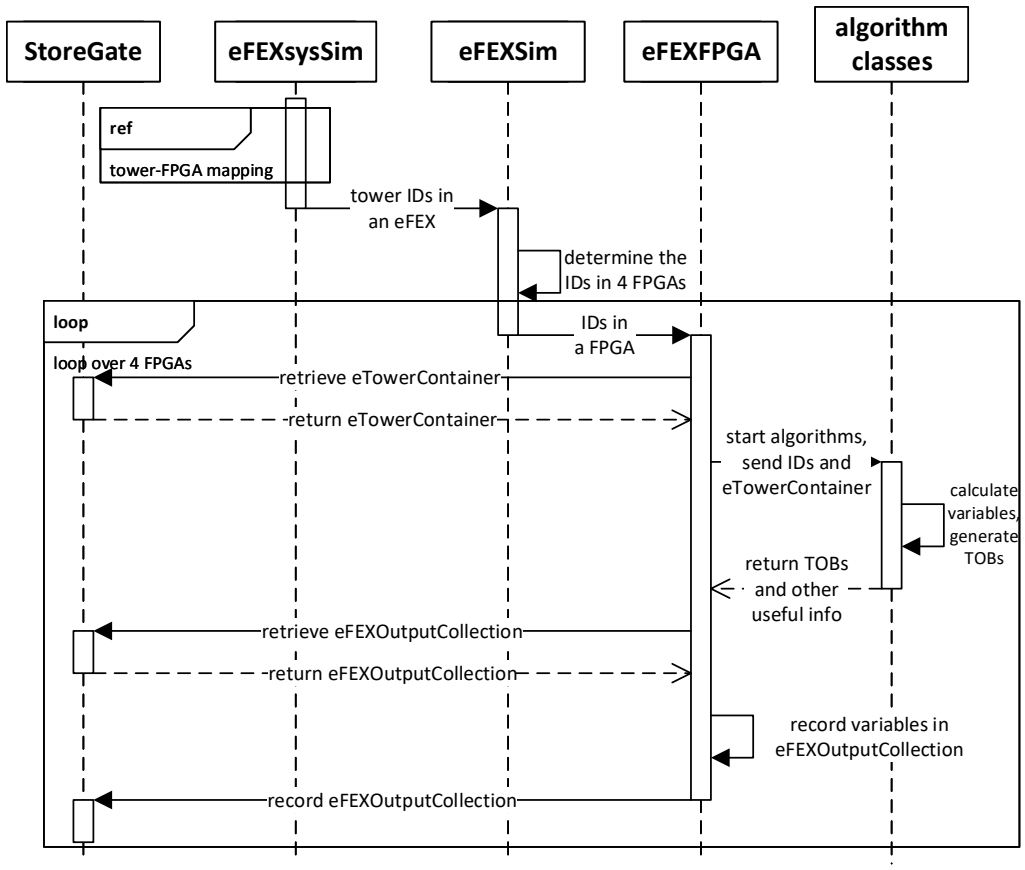


Figure 4.5: UML sequence diagram for eFEX, FPGA and algorithm simulation. The reference fragment for the tower-FPGA mapping is in Figure 4.4

The UML sequence diagram in Figure 4.5 shows the simulation of the eFEX, FPGA and algorithms. The simulation of the eFEX is initiated by the *eFEXsysSim* object by sending IDs of all the trigger towers in an eFEX module to the *eFEXSim* object. The *eFEXSim* object then starts the simulation of four FPGAs by calling the *eFEXFPGA* object which receives the IDs of all the trigger towers in an FPGA. The *eFEXFPGA* object retrieves the *eTowerContainer* via the StoreGate Service and starts the e/γ and τ algorithm simulation by creating 3×3 tower windows and initiating the *eFEXegAlgo* and *eFEXtauAlgo* objects with the corresponding trigger tower IDs in a tower window. The algorithm simulation retrieves the SuperCell E_T within the tower window and

calculates all variables required to create TOBs. The TOB objects are then sent back to the *eFEXFPGA* object, where the relevant information for monitoring is recorded in the *eFEXOutputCollection* object via the StoreGate Service.

4.5.3 Test Modules

The *eFEXNtupleWritter* class is an optional module for testing and debugging purposes. The module accesses the information stored in the *eFEXOutputCollection* object and creates a ROOT tuple file [67] for analysis. The variables contained in the output file include the η, ϕ, sums of E_T and jet rejection variables of all the tower windows in the calorimeters. The class also retrieves the η, ϕ and E_T of truth-level particles which can be used to make comparisons with the TOBs.

The *eFEXNtupleWritter* class and the *eFEXOutputCollection* should only be used at the development stage of the offline simulation software because it slows down the computation and generates unnecessary output files. An option which disables its functionalities without affecting other parts of the simulation is available in the JobOptions file.

4.5.4 Test Vectors for Validation

The L1Calo online software [68] is a standalone framework for the control, testing and monitoring of the L1Calo system. It is also capable of event-by-event simulation of the L1Calo system using test vector input. The test vector is a text file containing the SuperCell E_T of a single FPGA module or that of the whole calorimeter system. With the same input data, the offline simulation software is expected to to produce exactly the same output as the online software. To make comparisons with the online software, the offline simulation needs to accept input in the test vector data format.

As is discussed in Section 4.4, an optional module (*eFakeTower*) is created after the mapping of SuperCells. The module loads the test vector file and replaces the E_T stored in the *TriggerTower* object with the one provided by the test vector file. After that, the framework will be able to perform eFEX, FPGA and algorithm simulation using the same input as the online simulation.

The test vector module will be used in the development stage of the framework for debugging and validation purposes, and its functionalities will be disabled in the final release.

4.6 Results

This section shows the preliminary results of the simulation. The plots in this section were made using the monitoring data created by the *eFEXNtupleWriter* Athena Algorithm. The plots were created using the simulated $Z \to e^+e^-$ and dijet MC samples. The next-to-leading-order (NLO) $Z \to e^+e^-$ events were generated by PowhegBox [69–71], while the parton shower and hadronisation were simulated by Pythia8 [72]. The AZNLO [73] set of parameter tuning and the CTEQ6 [74] PDF set were used. The leading-order (LO) dijet events were generated by Pythia8 with the LO NNPDF 2.3 [75, 76] PDF set and the A14 [77] parameter tuning. The decays of the bottom and charm hadrons in both samples were simulated by EvtGen [78].

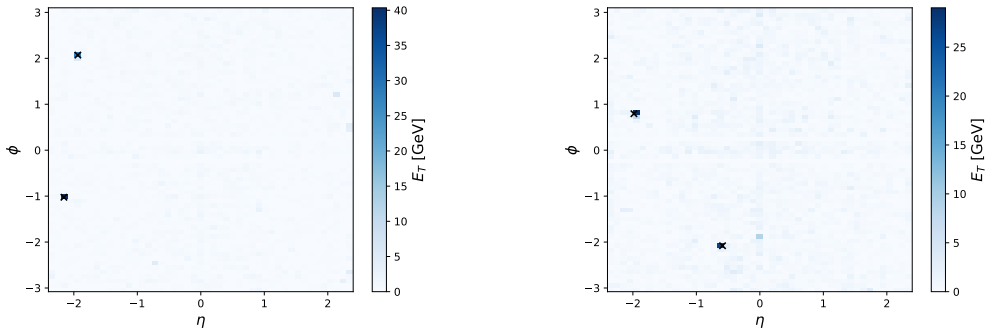


Figure 4.6: The distributions of the electromagnetic-layer E_T of two random $Z \to e^+e^-$ events in the $\eta - \phi$ plane. The black crosses are the locations of the electrons.

Figure 4.6 shows the distributions of the total tower E_T in the four electromagnetic layers for two random $Z \to e^+e^-$ events in the $\eta - \phi$ plane. The plots show that the SuperCell mapping is correct because the two most energetic towers and the two truth electrons are at the same locations.

Figure 4.7 shows the distributions of W_s for the $Z \to e^+e^-$ and dijet events. The plot only shows W_s in trigger-tower windows with $E_{\text{Cluster}} > 10$ GeV. The plot shows that the implementation of the W_s variable works as expected since jets have a higher

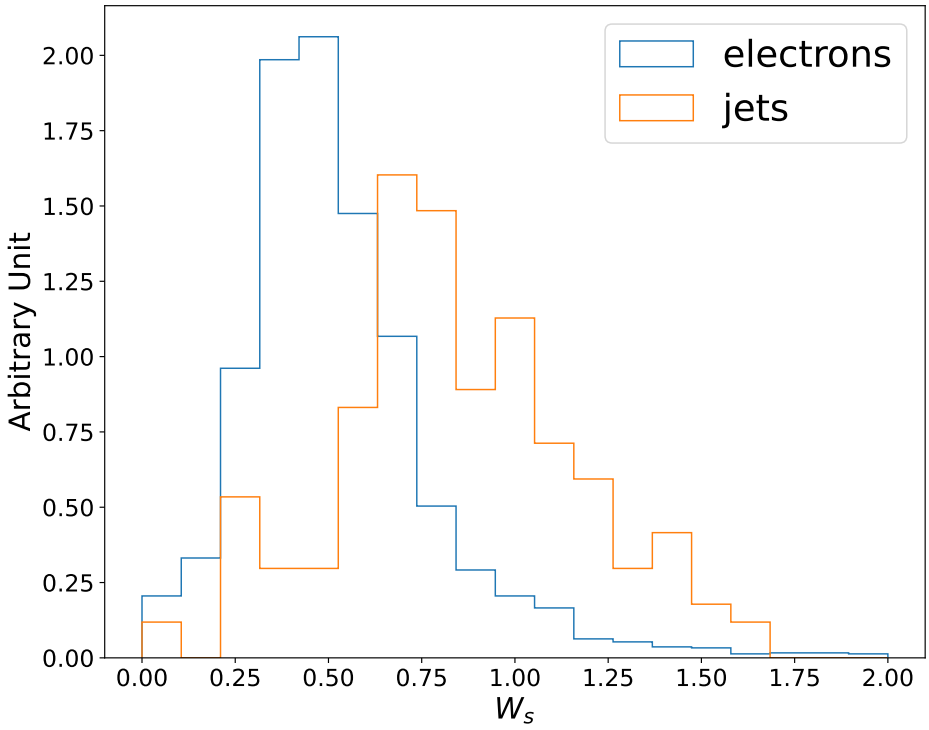


Figure 4.7: The distributions of W_s for the $Z \to e^+e^-$ and dijet events.

W_s value than electrons.

Figure 4.6 and Figure 4.7 show that the distributions of E_T and W_s calculated by the simulation in the Athena framework behave as expected. In addition to the variable distributions, validation also needs to be conducted by comparing the result of the online simulation with that of the Athena-based simulation using the method described in Section 4.5.4. At the time of writing, the validation over the full η/ϕ range is still ongoing, and the simulation is aimed to be ready for use in Run-3 analyses.

Chapter 5

Object Reconstruction

Chapter 3 describes how particles interact with detectors. Before the analysis, physics objects, such as electrons, muons, jets and missing transverse momenta, need to be identified and reconstructed from detector signals. This chapter describes the reconstruction methods for the physics objects used in this thesis.

5.1 Tracks and Primary Vertices

Tracks of charged particles are reconstructed using hits in the inner detector. A track is described by five parameters: $d0$, $z0$, ϕ, θ and q/p. The first two parameters, $d0$ and $z0$, are the transverse and longitudinal impact parameters to the interaction point. ϕ and θ are the azimuthal and polar angles of the track's momentum at the point closest to the interaction point. q/p is the ratio between the charge and the magnitude of the momentum.

There are two approaches to reconstruct tracks [79]. The inside-out algorithm starts from 3-point seeds in the 3D space of the silicon trackers, where seeds are preselected based on variables like the impact parameters to reduce computing time. After that, the trajectory is extrapolated to add hit points, and low-quality track candidates are rejected to remove fake tracks. Finally, hits in the TRT are added, and the track is reconstructed using full information of the inner detector. The tracks are required to have $p_T > 400$ MeV. The inside-out algorithm is mainly used to reconstruct charged

particles originating from the interaction point. To reconstruct particles produced by decays inside the inner detector, such as electrons from photon conversion and particles from b-hadron decays, the outside-in algorithm starts from segments of the TRT and the trajectories are extrapolated inward to add hits from the silicon trackers. It also reconstructs other tracks missed in the inside-out tracking.

Vertices are the interaction points identified using the reconstructed tracks [80]. The reconstruction of vertices starts with a seed which is the global maximum of the track distribution on the z-axis. After that, the vertex's location is determined by fits using the nearby tracks. A seed for a new vertex is created if a track has more than 7σ incompatibility with the existing vertices. The procedure is repeated until all tracks are associated with vertices. The primary vertex is the vertex whose tracks have the highest $\sum p_T^2$.

5.2 Electrons

The thesis uses electrons reconstructed by the method described in [81]. Electrons are reconstructed by matching tracks in the inner detector to the clusters in the calorimeter. If more than one track is matched, electron tracks are selected based on the number of pixel hits and the ΔR separation between track and cluster. If the track of the candidate electron is matched to a secondary vertex, the electron is likely to be from photon conversion. Information about pixel detector hits, the candidate electron's momentum and the presence of the secondary vertex is used to differentiate prompt electrons from photons.

Photons, hadronic jets and decaying hadrons may be misidentified as prompt electrons, so a likelihood-based (LH) discriminator

$$d_L = \frac{L_S}{L_S + L_B} \tag{5.1}$$

is defined for electron identification, where L_S and L_B are the likelihood of a candidate being a prompt electron and the likelihood of a candidate being a process that mimics the signature of a prompt electron, respectively. L_S and L_B are calculated using variables like

the energy deposit in each calorimeter layer, track qualities and track-cluster matching. Three electron identification working points, LH loose, LH medium and LH tight, are defined using cuts on d_L [81].

A signature of prompt electrons is low activity in the detector region surrounding the electron candidate, so candidates also need to satisfy isolation requirements. Two variables, p_T^{cone20} and E_T^{cone20}, are used for the isolation of clusters and tracks. E_T^{cone20} is the sum of transverse energy in a cone of $\Delta R = 0.2$ in the electromagnetic calorimeters centered on the candidate with the transverse energy of the candidate subtracted. p_T^{cone20} is the sum of the tracks' transverse momentum in a cone of $\Delta R = \min(10/p_T, 0.2)$ centered on the candidate with the transverse momentum of the candidate subtracted. Products from boosted particle decays are close to each other, so variable radii are used in the track calculation to take advantage of the fine granularity of the inner detector. Two isolation working points are defined using cuts on E_T^{cone20} and p_T^{cone20}. The FixedCutLoose working point requires $E_T^{\text{cone20}}/p_T < 0.2$ and $p_T^{\text{cone20}}/p_T < 0.15$, while the FixedCutTight working points requires $E_T^{\text{cone20}}/p_T < 0.06$ and $p_T^{\text{cone20}}/p_T < 0.06$.

The requirements $|d_0|/\sigma_{d_0} < 5$ and $|z_0 sin(\theta)| < 0.05$ mm are placed on the impact parameters of the tracks to ensure that the electrons are associated with the primary vertex, where σ_{d_0} is the uncertainty on the transverse impact parameter.

5.3 Muons

Muons are reconstructed using information from the inner detector, the calorimeters and the muon spectrometer [82]. The tracks in the muon spectrometer are reconstructed by a χ^2 fit using hits in the muon spectrometer. There are five methods to reconstruct muon candidates. The combined muons are reconstructed by matching the tracks in the muon spectrometer to the tracks in the inner detector, with the energy loss in the calorimeters taken into account. The inside-out combined muons are reconstructed by extrapolating the inner detector tracks to the muon spectrometer. This method recovers muons in the regions with poor muon spectrometer coverage and low-p_T muons which stop inside the muon spectrometer, because it does not rely on reconstructed tracks in the muon

spectrometer. The muon-spectrometer extrapolated method extrapolates tracks in the muon spectrometer to the interaction point to recover muons in the regions not covered by the inner detector. The segment-tagged muons require at least one segment in the muon spectrometer matched to a track in the inner detector, and the muons are reconstructed using the parameters of inner-detector tracks. The calorimeter-tagged muons are also reconstructed using the parameters of inner-detector tracks. The difference is that the tracks in the inner detector are extrapolated to the calorimeters instead of the muon spectrometer. This method recovers muons in the region with poor muon spectrometer coverage.

Hadronic jets which penetrate the calorimeters and muons from decaying hadrons may be reconstructed as prompt muons. Therefore, identification and isolation requirements are applied.

Variables such as χ^2 of fits and q/p compatibility between the inner detector and the muon spectrometer are used for muon identification. The loose, medium and tight working points are defined with increasing prompt muon purities [82]. In addition to the three working points, a High-p_T working point is defined for muons with $p_T > 100$ GeV to ensure good momentum resolution.

Same as the treatment for electrons, p_T^{cone20} and E_T^{cone20} are also used for the muon isolation. The FixedCutLoose working point requires $E_T^{\text{cone20}}/p_T < 0.3$ and $p_T^{\text{cone20}}/p_T < 0.15$, while the FixedCutTight working points requires $E_T^{\text{cone20}}/p_T < 0.06$ and $p_T^{\text{cone20}}/p_T < 0.06$.

Likes electrons, muons are also required to satisfy $|d_0|/\sigma_{d_0} < 5$ and $|z_0 sin(\theta)| < 0.05$ mm to ensure that the reconstructed muons are associated with the primary vertex.

5.4 Jets

Jets are reconstructed using the energy deposits in the calorimeters. The first step of the reconstruction is to determine the topological clusters [83] of energy deposit in calorimeters. The significance of energy deposit in a calorimeter cell is E_c/σ_c where E_c is the energy deposit in the cell and σ_c is the uncertainty caused by background noise. A

cluster is seeded from high significant cells while mild significant cells nearby are merged to form a cluster.

Jet algorithms [84] are then used to merge clusters into jets. A weighted distance between a cluster i and another cluster j in the $\phi - y$ plane is defined as

$$d_{ij} = \min(k_{T,i}^n, k_{T,j}^n)\frac{\Delta R_{ij}^2}{R^2}, \tag{5.2}$$

where $k_{T,i}^n$ and $k_{T,j}^n$ is the transverse momenta of the cluster i and the cluster j, $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, R is a user defined parameter and n is a parameter specified by the choice of algorithm. The beam distance of the cluster i is defined as

$$d_{iB} = k_{T,i}^n. \tag{5.3}$$

The algorithm compares the two distance. If d_{ij} is smaller, the two clusters are merged. If d_{iB} is smaller, then the cluster is identified as a jet. The algorithm is known as the Cambridge-Aachen algorithm when using an unweighted distance ($n = 0$). When $n = 2$, the algorithm becomes the k_T algorithm which prefers to merge low-p_T clusters. When $n = -2$, the algorithm becomes the anti-k_T algorithm which prefers to merge high-p_T clusters.

Track jets are reconstructed using tracks in the inner detector. The track jets reconstructed by the algorithms discussed above are fixed-radius track jets. The jets from high-p_T particle decays are collimated and overlap with each other. The variable-radius (VR) track jet algorithm [85] can be used to improve the reconstruction, where the radius R in Equation (5.2) depends on p_T:

$$R_{\mathrm{VR}} = \frac{\rho}{k_{T,i}^n} \tag{5.4}$$

where ρ is a parameter to be specified. An upper bound R_{max} and a lower bound R_{min} are defined for the R_{VR} to make sure that jets will not get too large or shrink below the detector resolution. Studies [86] show that the best parameters for the $h \to b\bar{b}$ analysis are $\rho = 30$ GeV, $R_{\mathrm{max}} = 0.4$ and $R_{\mathrm{min}} = 0.02$.

There are multiple methods used to reconstruct jets in the ATLAS experiment. The topological cluster jets are reconstructed using clusters in the calorimeters with the anti-k_T algorithm [83]. The η and ϕ of the jet are the relative position with respect to the primary vertex instead of the centre of the detector. The Particle Flow (PFlow) jets are reconstructed using both tracks and the calorimeter clusters to improve the reconstruction in high pileup environments [87]. It also improves the energy resolution and the mass resolution of the jets. The PFlow algorithm is motivated by the fact that the inner detector has a better momentum resolution than the calorimeters at low p_T. The algorithm first matches tracks to calorimeter clusters. The energy deposit is estimated using the track and subtracted from the clusters. After the subtraction, the cluster is removed if the remaining energy is consistent with an energy fluctuation. Track-CaloCluster (TCC) jets are also reconstructed using both the tracks and the calorimeter clusters [88]. However, it uses the energy of the clusters and the η and ϕ of the tracks. The TCC jet has better performance at high p_T, while the PFlow jet has better performance at low p_T. The Unified Flow Objects (UFOs) combine both algorithms and provide good performance in both the high p_T and the low p_T [89].

The reconstructed jet's energy is corrected to the energy of the truth jet by applying a jet energy scale (JES), which is derived from MC simulation. A correction is also applied to data to account for the difference between data and MC, which reduces the JES uncertainties. A likelihood called Jet Vertex Tagger (JVT) [90] is used to remove jets coming from the pileup. It is defined using information from the calorimeters and the inner detector and applied to jets with $p_T < 60$ GeV and $\eta < 2.4$.

The analysis searches for the $A/Z' \to Zh$ processes with the h boson decaying into two b-jets. The h boson can be reconstructed using two small-R jets which are reconstructed using anti-k_T algorithms with $R = 0.4$. When the h boson is boosted, the two jets from its decay are merged as one large-R jet which is reconstructed by the anti-k_T algorithm with $R = 1$. The analysis also uses track jets which are reconstructed by the VR track jet algorithm. They are used for the b-tagging of large-R jets, which will be discussed in Section 5.5.

Large-R jets are sensitive to the contamination of pileup events, initial-state radiations

and multi-parton interactions, so a jet trimming method [91] is applied, where low-p_T subjets inside the large-R jets are subtracted.

Two additional corrections are applied to the b-tagged jets. A Muon-in-Jet correction [92] is applied to account for the energy mismeasurement caused by the muons from the semileptonic decay of b-quarks and c-quarks, where the muon momentum is added to the jet momentum while the muon energy deposits in the calorimeters are substituted. After the Muon-in-Jet correction, a p_T-based correction (PtReco) which is determined from Vhbb simulations is applied to the small-R jets. The correction improves the di-jet mass resolution.

5.5 b-Tagging

b-tagging is a procedure which identifies jets originating from B-hadrons. B-hadrons have a longer lifetime compared with other hadrons, which allows the reconstruction of a secondary vertex. The MV2c10 b-tagging algorithm [93] is used in this analysis.

The algorithm starts with three basic algorithms. The first algorithm is the impact parameter-based algorithm. The algorithm calculates the impact parameter of the track closest to the primary vertex, then uses a likelihood-based approach to identify b-jets. The impact parameter can be calculated using two different approaches. The IP3D approach uses both transverse impact parameters and longitudinal impact parameters, while The IP2D approach only uses transverse impact parameters. The second algorithm is the secondary vertex finding algorithm which reconstructs the secondary vertex in the jet. The properties of the reconstructed secondary vertex are used to identify b-jets. The third algorithm is the decay chain multi-vertex algorithm, which reconstructs the decay chain of the primary vertex to b-hadron to c-hadron.

A boosted decision tree (BDT) is trained using the outputs of the three basic algorithms and some other kinematic properties of the jet. The BDT is trained on MC samples with b-jets as signals and 90% light jets and 10% c-jets as backgrounds. Jets satisfying the 70% percent efficiency working point are defined as a b-tagged jet.

5.6 Missing Transverse Momentum

Neutrinos are not detected by the ATLAS detector, and they contribute to the momentum imbalance in the transverse plane. The missing transverse momentum (MET) is the negative vector sum of momenta in the transverse plane. The variable is mainly used to describe the momenta of neutrinos, but other effects such as unreconstructed objects and measurement errors also contribute. The variable is calculated using all reconstructed physical objects such as leptons, jets and photons. Other unreconstructed objects whose tracks are associated with the primary vertex are also used.

5.7 Overlap Removal

An object may be reconstructed by different parts of the detector which produces duplicates. Objects which are close to each other may also affect the measurement. The thesis uses the same overlap removal method as the previous ATLAS VH analysis [7]. The following overlap removal procedure for small-R jets is applied:

- If an electron and a muon share the same track, only keep the muon.
- If a jet is within the range $\Delta R < 0.2$ of a surviving electron, remove the jet.
- If an electron is within the range $\Delta R < \min(0.4, 0,04 + 10 \text{ GeV } / p_T^e)$ of a surviving jet, and the jet also satisfies JVT > 0.64, $p_T^{jet} < 50$ GeV and $|\eta| < 2.4$, remove the electron.
- If a jet is within the range $\Delta R < 0.2$ of a surviving muon, and the jet has less than 3 tracks, or if the jet has p_T^{μ}/(number of tracks)> 0.7 and $p_T^{jet}/p_T^{muon} < 2$, remove the jet.
- If a muon is within the range $\Delta R < \min(0.4, 0,04 + 10 \text{ GeV } / p_T^e)$ of a surviving jet, and the jet also satisfies JVT > 0.64, $p_T^{jet} < 50$ GeV and $|\eta| < 2.4$, remove the muon.
- If a tau is within the range $\Delta R < 0.4$ of a jet, remove the jet.

For large-R jets, the last step related to taus is not performed. A large-R jet is removed if it is within the range $\Delta R < 1.2$ of an electron.

Chapter 6

Search for a Heavy Resonance in the $llbb$ Final State

This chapter discusses the search for a heavy resonance (Z' and A) decaying into a Z boson and an h boson in the $l^-l^+b\bar{b}$ final state (2-lepton channel) where l depicts an electron or a muon. The search considers the gluon-gluon fusion production (ggA) and the b-associated production (bbA) of the A boson and the Drell-Yan production of the Z' boson. The search uses the same analysis method for both the ggA and the Z' processes. The analysis method for the bbA process is different due to the two additional b-jets in the final state. The search for the bbA and ggA processes covers a mass range of 220 GeV to 2 TeV, while the search of the Z' boson covers a mass range of 300 GeV to 5 TeV. The results of the 2-lepton channel search are combined with that of the 0-lepton channel ($\nu\bar{\nu}b\bar{b}$ final state), which will be discussed in Chapter 7.

The analysis is a successor to the previous ATLAS Vh analysis [7] using 36.1 fb^{-1} of data, whose results have been summarized in Section 2.3. The new analysis improved the sensitivity of the search with the full Run 2 dataset and new object reconstruction methods. The event selection has been updated to increase the sensitivity to the signals and improve the background modelling. A new background modelling method has also been developed to improve the agreement between the data and the MC simulation.

The analysis method was initially developed using the simulated events, where event selections were designed to increase the purity of the signal events. Control regions with

low signal purity were defined using events that failed the signal-region event selection. A background-modelling method was developed to improve the agreement between the data and the MC simulation. The background modelling was tested using the data in the control regions first. The same background-modelling method was then applied to the signal region for resonance searches.

6.1 Signals, Backgrounds and Monte Carlo Simulation

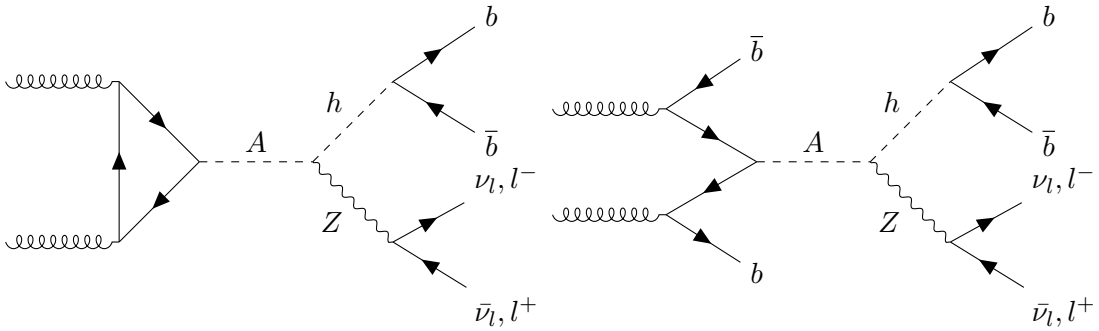


Figure 6.1: The Feynman diagrams of the ggA process (left) and the bbA process (right). The diagrams are created using the TikZ-Feynman package [14].

The Feynman diagrams of the ggA and bbA processes in the final states of $l^-l^+b\bar{b}$ and $\nu\bar{\nu}b\bar{b}$ are shown in Figure 6.1. The LO ggA samples were generated by MadGraph5 2.3.3 and interfaced with Pythia 8.186 [72, 76] using the NNPDF LO 2.3 PDF set [75, 76] and the A14 tuning [77]. The NLO ggA samples were generated by MadGraph5_aMC@NLO 2.3.3 and interfaced with Pythia 8.186 with the CT10F4 NLO PDF set [94] and the A14 tuning.

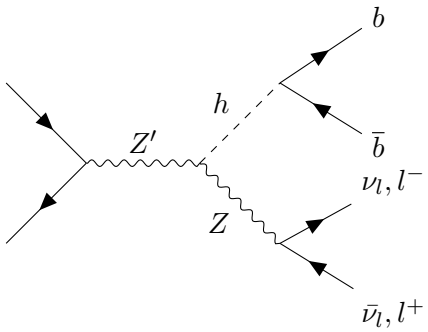


Figure 6.2: The Feynman diagram of the Drell-Yan production of the Z' boson. The diagram is created using the TikZ-Feynman package [14].

Figure 6.2 shows the Drell-Yan production of the Z' boson in the $l^-l^+b\bar{b}$ and $\nu\bar{\nu}b\bar{b}$ final states. The samples were generated at LO accuracy by MadGraph5 2.3.3 and interfaced with Pythia 8.186 using the NNPDF 2.3 LO PDF set and the A14 tuning.

The bbA and ggA samples were simulated as a grid of mass hypothesis between 220 GeV and 2 TeV, while the Z' samples were simulated as a grid between 300 GeV and 5 TeV. A morphing procedure [95] was used to interpolate between the simulated mass points, where the distributions of the reconstructed signal mass m_{VH} and the reconstructed signal transverse mass $m_{T,VH}$ were approximated using the distributions of the adjacent masses from the simulation. Studies [96] show that the morphed distributions give a more conservative result than the simulated distributions. All signal samples were generated with the assumption that the heavy resonance has zero natural width (narrow-width approximation).

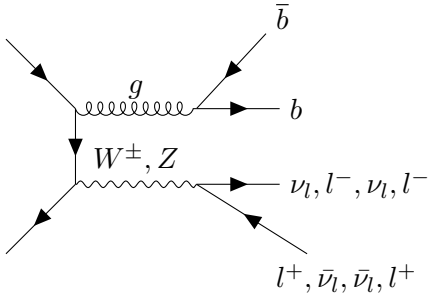


Figure 6.3: The Feynman diagram for the Z+jets background and the W+jets background. The diagram is created using the TikZ-Feynman package [14].

The largest background in this analysis is the Z/W+jets background which is shown in Figure 6.3. The background was generated by Sherpa 2.2.1 [97] using the NNPDF 3.0 NLO PDF set [98], where diagrams with up to two additional parton emissions were simulated at NLO accuracy while diagrams with three or four additional parton emissions were simulated at LO accuracy. The tuning developed by the Sherpa authors was used. A global k-factor was applied to normalize the cross section of the samples to the next-to-next-to-leading order (NNLO) accuracy [99].

The second largest background in the analysis is the $t\bar{t}$ process shown in Figure 6.4. The background was simulated by PowhegBox v2 [69–71] and interfaced with Pythia 8.230 using the NNPDF 3.0 NLO PDF set and the A14 tuning. The cross section of the

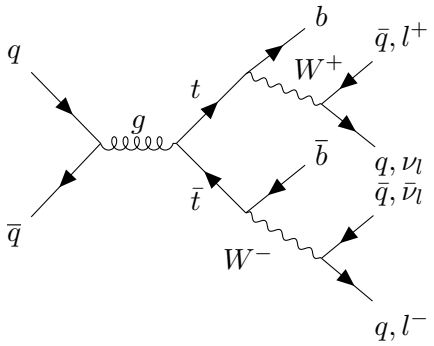


Figure 6.4: The Feynman diagram of the $t\bar{t}$ background. The diagram is created using the TikZ-Feynman package [14].

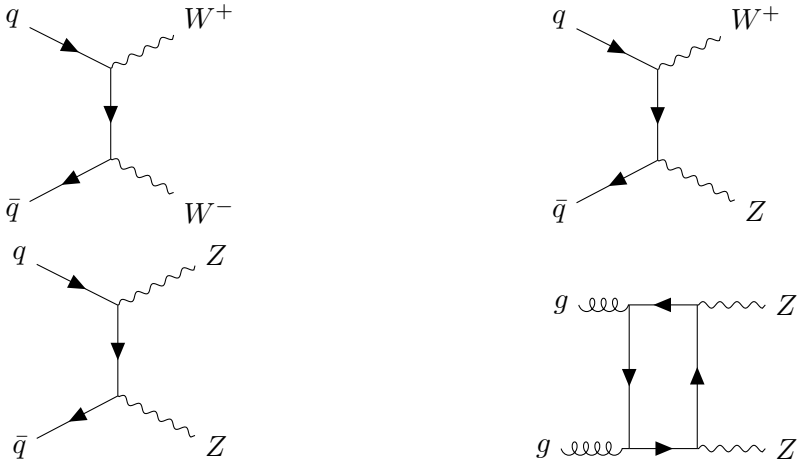


Figure 6.5: The Feynman diagrams of the WW, WZ, ZZ and the loop-induced $ggZZ$ backgrounds. The diagrams are created using the TikZ-Feynman package [14].

sample is normalized to the NNLO accuracy [100–105].

The diboson (WW, WZ, ZZ) backgrounds, shown in Figure 6.5, were simulated by Sherpa 2.2.1 with the NNPDF 3.0 NNLO PDF set, where diagrams with up to one additional parton emission were simulated at NLO accuracy while diagrams with two or three additional parton emissions were simulated at LO accuracy. The loop-induced $ggZZ$ process, shown in Figure 6.5, was simulated at LO by Sherpa 2.2.2 with the NNPDF 3.0 NNLO PDF set.

The s-channel single-top background, the t-channel single-top background and the W-associated single-top background are shown in Figure 6.6. They were simulated by PowhegBox v2 at NLO using the NNPDF 3.0 NLO PDF set, interfaced with Pythia 8.230 with the NNPDF 2.3 LO PDF set and the A14 tuning. The overlap between

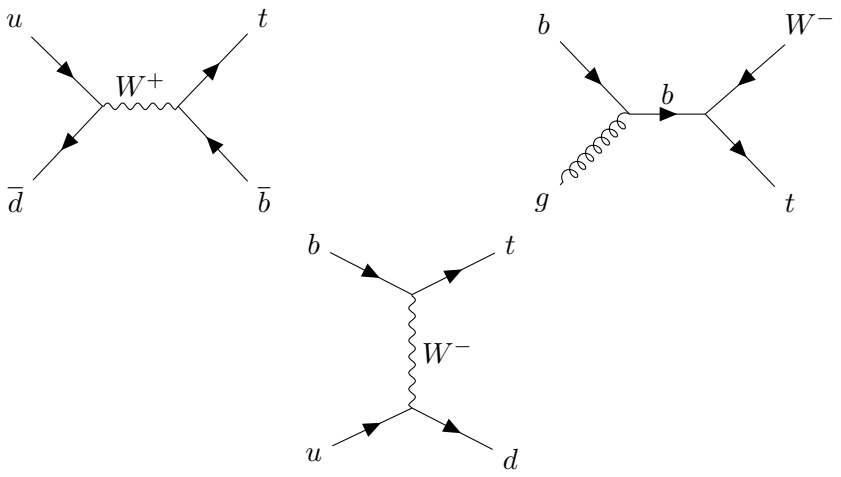


Figure 6.6: The Feynman diagrams of the s-channel single-top background, the t-channel single-top background and the W-associated single-top background. The diagrams are created using the TikZ-Feynman package [14].

the W-associated single-top background and the $t\bar{t}$ background was removed using the diagram removal scheme [106].

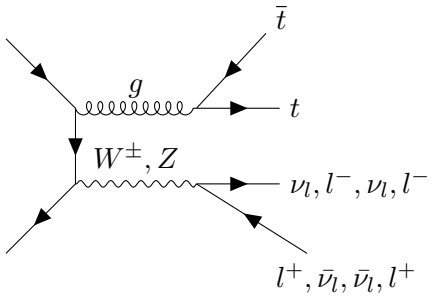


Figure 6.7: The Feynman diagram of the $t\bar{t} + Z$ and the $t\bar{t} + W^{\pm}$ backgrounds. The diagram is created using the TikZ-Feynman package [14].

The $t\bar{t} + Z$ and the $t\bar{t} + W^{\pm}$ backgrounds, shown in Figure 6.7, were generated at NLO accuracy by MadGraph5_aMC@NLO 2.3.3 and interfaced with Pythia 8.186 using the NNPDF LO 2.3 PDF set and the A14 tuning.

The Standard-Model Higgs processes, shown in Figure 6.8, also contribute to the backgrounds of the analysis. The analysis considers the background where the Standard-Model Higgs boson decays into two b-quarks or two c-quarks. The $t\bar{t} + h$ was simulated by MadGraph5_aMC@NLO 2.3.3 and interfaced with Pythia 8.212 using the NNPDF LO 2.3 PDF set and the A14 tuning. The Higgs-strahlung Zh/Wh production was simulated by PowhegBox v2 at NLO with up to one parton emission

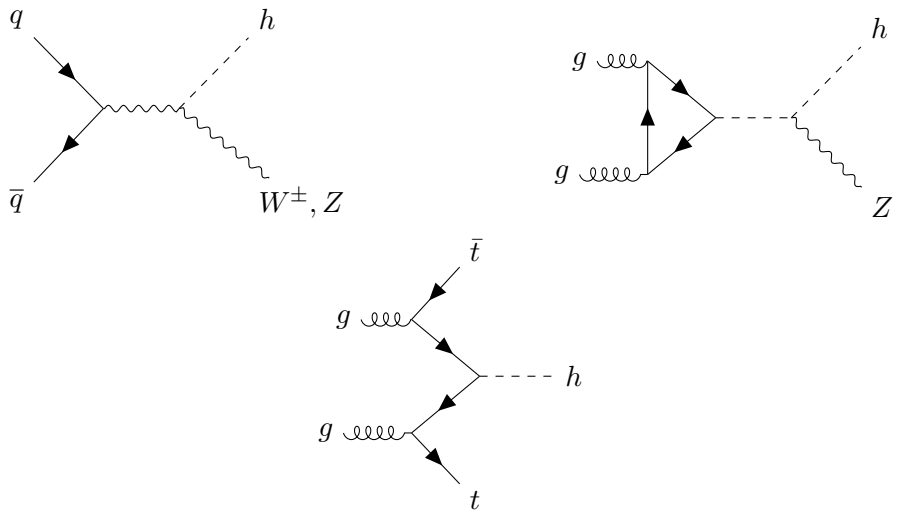


Figure 6.8: The Feynman diagrams of the Higgs-strahlung production (top left), the gluon-fusion Higgs production (top right) and the t-associated Higgs production (bottom). The diagrams are created using the TikZ-Feynman package [14].

and interfaced with Pythia 8.212 using the PDF4LHC PDF set [107] and the AZNLO tuning [73]. The gluon-fusion Zh production was simulated at LO with the same tools as the loop-induced Zh.

6.2 Object Definition

The $A/Z' \to Zh \to l^+l^-b\bar{b}$ decay channel has a final state with two charged leptons and two b-jets. The electrons and muons are reconstructed using the methods described in Chapter 5. The electrons are selected using the LH loose identification working point. The low-p_T muons are selected using the loose working point, while muons with $p_T > 300$ GeV are selected using the High-p_T working point instead. The loose isolation working point is applied to electrons and muons with $p_T < 100$ GeV, while no isolation is required for electrons and muons with $p_T > 100$ GeV. The electrons are required to be within the inner detector's coverage of $|\eta| < 2.47$, and muons are required to be within the muon spectrometer's coverage of $|\eta| < 2.7$.

There are three types of jets in the analysis. Small-R jets are reconstructed from topological clusters with the anti-k_T algorithm with $R = 0.4$. They are sensitive to low-mass signals. When the light Higgs boson is boosted, the two jets from its decay are

merged into one large-R jet. The large-R jets are TCC jets reconstructed with $R = 1$. The third type of jet is the track jet built using the VR track-jet algorithm. It is used for the flavour tagging of large-R jets.

The b-jets are identified using the MV2c10 b-tagging algorithm at the 70% efficiency working point. Small-R jets satisfying $|\eta| < 2.5$ and $p_T > 20$ GeV are called signal jets. The b-tagging is applied to all signal jets and track jets within the range $|\eta| < 2.5$. The jets in the forward regions ($2.5 < |\eta| < 4.5$) are forward jets. The forward jets are not eligible for b-tagging.

6.3 Event Selection and Analysis Strategy

data taking period	electron trigger	muon trigger
2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50
2016 2017 2018	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_mu26_ivarmedium HLT_mu50

Table 6.1: A list of triggers used in the analysis [50–53]. The number following the lepton type (e or mu) is the minimum transverse momentum requirement in GeV. The text following lh or i is the isolation requirement. $nod0$ means that the transverse impact parameter is not used in the electron identification. $etcut$ means that only information about E_T is used in the trigger selection. $L1EM$ and $L1MU$ are the place where the trigger is seeded and the number after this is the minimum L1 p_T requirement. H means using a transverse momentum dependent veto against the energy deposit in the hadronic calorimeter behind the electron. V means using a transverse momentum threshold as a function of η.

Events for the analysis are selected based on the signature of the signal processes. The event selection makes sure that the analysis only uses events with a good reconstruction quality. It also increases the purity of signal events. The same event selections are applied to both the data and the simulated events. This thesis uses event selections and analysis strategies which are similar to the previous ATLAS VH analysis [7].

The data is collected by a combination of single-electron and single-muon triggers shown in Table 6.1. The trigger selects events using a combination of p_T thresholds and isolation requirements. The p_T thresholds used during 2016 - 2018 are higher than

those used during 2015 because of the increased instantaneous luminosity. The object which fires the trigger needs to match one of the reconstructed electrons or muons with $p_T > 27$ GeV.

The analysis only uses data collected when the detector is in good working condition. The quality of the data is checked by experts, and analysis only uses data which is good for physics analysis.

The analysis looks for a Z' boson or an A boson decaying into a Z boson and an h boson, with the Z boson decaying into two electrons or two muons and the h boson decaying into two b-quarks. The analysis only uses events with exactly two electrons or two muons. The Z boson candidate is reconstructed using the two charged-lepton objects. The events are split into the resolved category and the merged category according to the reconstruction method of the h boson. In the resolved category, the h boson candidate is reconstructed by the two leading (in p_T) b-tagged signal jets when there are at least two b-tagged signal jets available. In the case where only one b-tagged signal jet is available, the b-tagged signal jet and a remaining signal jet with the highest p_T are selected. In the merged category, the h boson candidate is reconstructed by one large-R jet with exactly two track jets associated with it, and at least one of the track jets has to be b-tagged. If the event satisfies the requirements for both resolved and merged categories, it is assigned to the resolved category because the resolved reconstruction method has a better signal mass resolution.

Variable	Resolved	Merged
Number of jets	$\geq$2 signal jets	$\geq$1 large-R jet $\geq$ 2 track jets (matched to leading $R = 1.0$ jet)
Leading jet p_T [GeV]	> 45	> 250
m_{bb} [GeV]	100–145	75–145
Leading lepton p_T [GeV]	> 27	> 27
Sub-leading lepton p_T [GeV]	> 20	> 25
Muon charge	opposite	
Lepton flavour	same flavour	
$E_T^{\text{miss}}/\sqrt{H_T}$ [$\sqrt{\text{GeV}}$]	$< 1.15 \ + 8 \times 10^{-3} \cdot m_{Vh}/(1\text{ GeV})$	
$p_{T,\ell\ell}$ [GeV]	$> 20 + 9 \cdot \sqrt{m_{Vh}/(1\text{ GeV}) - 320}$ when $m_{Vh} > 320$	
$m_{\ell\ell}$ [GeV]	[max[40 GeV, $87 - 0.030 \cdot m_{Vh}/(1\text{ GeV})$], $97 + 0.013 \cdot m_{Vh}/(1\text{ GeV})$]	

Table 6.2: The signal region event selection for the $A/Z' \to Zh \to l^+l^-b\bar{b}$ search.

The events are further selected based on the kinematics of the reconstructed objects, which is shown in Table 6.2. The leading small-R jets are required to have $p_T > 45$ GeV, while the leading large-R jets are required to have $p_T > 250$ GeV. In the merged category, the mass of the leading large-R jets is required to be within the Standard-Model Higgs mass windows 75 GeV $< m_{bb} <$ 145 GeV. The requirement is tighter in the resolved category due to its better mass resolution, where the invariant mass of the di-jet object is required to have 100 GeV $< m_{bb} <$ 145 GeV. The leading lepton needs to have $p_T > 27$ GeV in both the resolved and merged categories. The sub-leading lepton p_T is required to be greater than 20 GeV in the resolved category, while it is required to be greater than 27 GeV in the merged category. The Z decays produce two oppositely charged leptons with the same flavour. The muons in an event are required to be oppositely charged. This is not applied to electrons because of the high rate of electron-charge misidentification. The leptons in an event are required to have the same flavour to reduce the $t\bar{t}$ background.

The signal events are expected to have low missing transverse energy (E_T^{miss}) because there are no neutrinos in the final states, so a cut is applied on $E_T^{\text{miss}}/\sqrt{H_{\text{T}}}$, where H_T is the scalar p_T sum of all small-R jets and leptons. The main contributions for the missing transverse momentum of signal events are the mismeasurement of jets' and leptons' p_T. The variable estimates the significance of missing transverse energy since the resolution of E_T^{miss} is proportional to H_T. A cut is also applied to the mass of the Z candidates (m_{ll}). The two cuts suppress top-related backgrounds. The cuts are relaxed for high-m_{VH} events because of the low background and poor resolution of the high-mass signals.

The p_T of the Z boson candidate ($p_{\text{T},\ell\ell}$) is expected to increase as the signal mass increases. A m_{VH}-dependent cut on $p_{\text{T},\ell\ell}$ is applied to reduce the backgrounds.

Events are further split according to the number of b-tagged jets. The resolved events are categorized into regions with 1 b-tagged jet, 2 b-tagged jets and 3+ b-tagged jets. The merged events are categorized into regions with 1 b-tagged track jet without any b-tagged track jets not associated with the large-R jet, 2 b-tagged track jets without any b-tagged track jets not associated with the large-R jet, 1 b-tagged track jet with

one or more b-tagged track jets not associated with the large-R jet and 2 b-tagged track jets with one or more b-tagged track jets not associated with the large-R jet. The ggA and Z' analysis only uses 1 and 2 b-tagged resolved events and merged events without any additional b-tagged track jets not associated with the large-R jet. The 3+ b-tagged resolved events and merged events with additional b-tagged track jets are important for the bbA analysis due to the two additional b-jets in the final state.

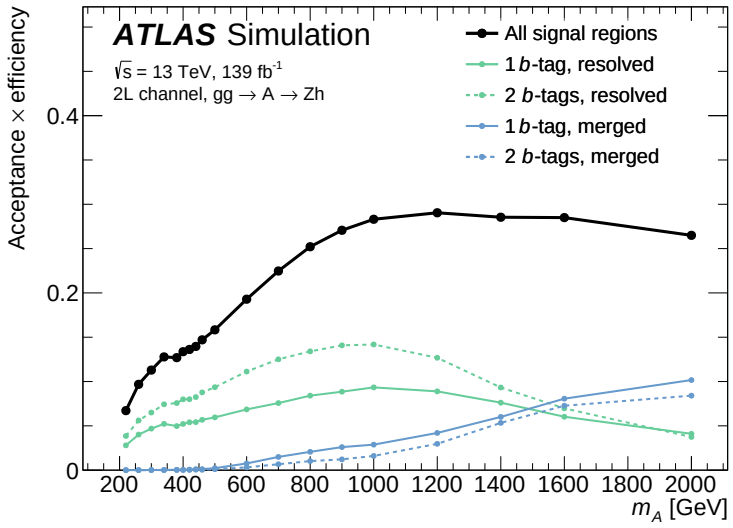


Figure 6.9: Event-selection acceptance × reconstruction efficiency for the ggA signals.

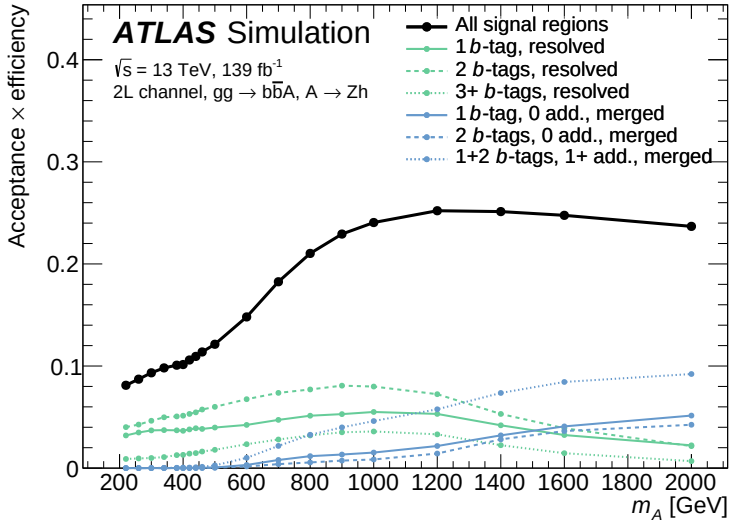


Figure 6.10: Event-selection acceptance × reconstruction efficiency for the bbA signals.

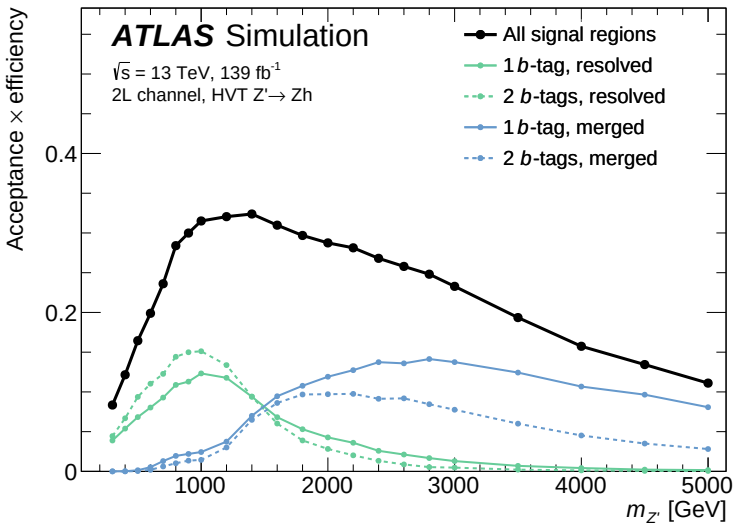


Figure 6.11: Event-selection acceptance × reconstruction efficiency for the Z' signals.

The event-selection acceptance × reconstruction efficiency for the ggA, bbA and Z' signals for different signal mass hypotheses are shown in Figure 6.9, Figure 6.10 and Figure 6.11 respectively. The plots show that the resolved regions have high efficiencies at low mass, while merged regions begin to have better efficiencies as the signal mass increases. As can be seen from Figure 6.11, the efficiencies in the merged regions begin to decrease after 2500 GeV. This is because the two boosted leptons get collimated and overlap with each other, which reduces the lepton-reconstruction efficiency.

Two types of control regions are defined using events which fail the signal-region event selections. Events failing the di-jet mass m_{bb} selection while still being within the range $50 < m_{bb} < 200$ GeV are assigned to the m_{bb} sideband control region which is dominated by the Z+jets and $t\bar{t}$ backgrounds. The resolved events with two leptons of different flavours and without the missing transverse energy requirement are assigned to the top control region, which helps to improve the modelling of the top backgrounds.

To improve the mass resolution of the signal, the four-momentum of the di-muon system is scaled by a factor of 91.2 GeV/m_{ll} after the event selection, which addresses the poor momentum measurement of the high-energy muons. The mass resolution is further improved by scaling the four-momentum of the di-jet system by a factor of 125GeV/m_{bb} in the resolved signal region. This is not applied to the merged signal region because it does not improve the sensitivity of the search.

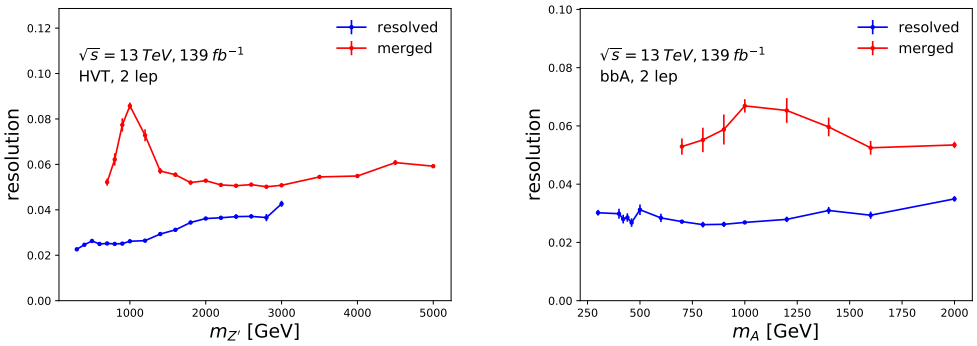


Figure 6.12: The signal mass resolution of the Z' signals (left) and the bbA signals (right) in the resolved and merged categories.

The signal mass resolution, shown in Figure 6.12, is determined by fitting a Bukin function [108] to the distributions of

$$\frac{m_{\text{recon}} - m_{\text{truth}}}{m_{\text{truth}}}, \tag{6.1}$$

where m_{recon} is the reconstructed mass and m_{truth} is the truth mass. The plots show the resolved category has a mass resolution of 2% to 4%, while the merged category has a mass resolution of about 5% to 6%.

6.3.1 Sub-leading Lepton p_T Requirement Study

In the previous analysis [7], the minimum p_T thresholds for the leading lepton and the second-leading lepton in the resolved category were 27 GeV and 7 GeV respectively. The loose requirement on the second-leading lepton p_T was based on the assumption that the non-prompt lepton contribution caused by the multijet background is negligible after the event selection. The multijet background and the W+jet background are poorly modelled and they may produce non-prompt leptons. To investigate the contribution of non-prompt leptons in the full Run 2 dataset, the event selection described in Table 6.2 is applied while requiring two muons of the same charge. The selection will enrich the non-prompt lepton contribution because the two main backgrounds (Z+jets and $t\bar{t}$) produce oppositely-charged prompt leptons . Figure 6.13 shows the distributions of the second-leading lepton p_T after the selection. The events are categorized into four regions according to the number of b-tagged jets. Events with 1 or 2 b-tagged jets are

sensitive to gluon-gluon fusion signal production, while events with 2 or more b-tagged jets are sensitive to b-associated signal production. The plots show the processes in the low-p_T region are underestimated by the simulation, which indicates the presence of non-prompt lepton contamination which cannot be ignored.

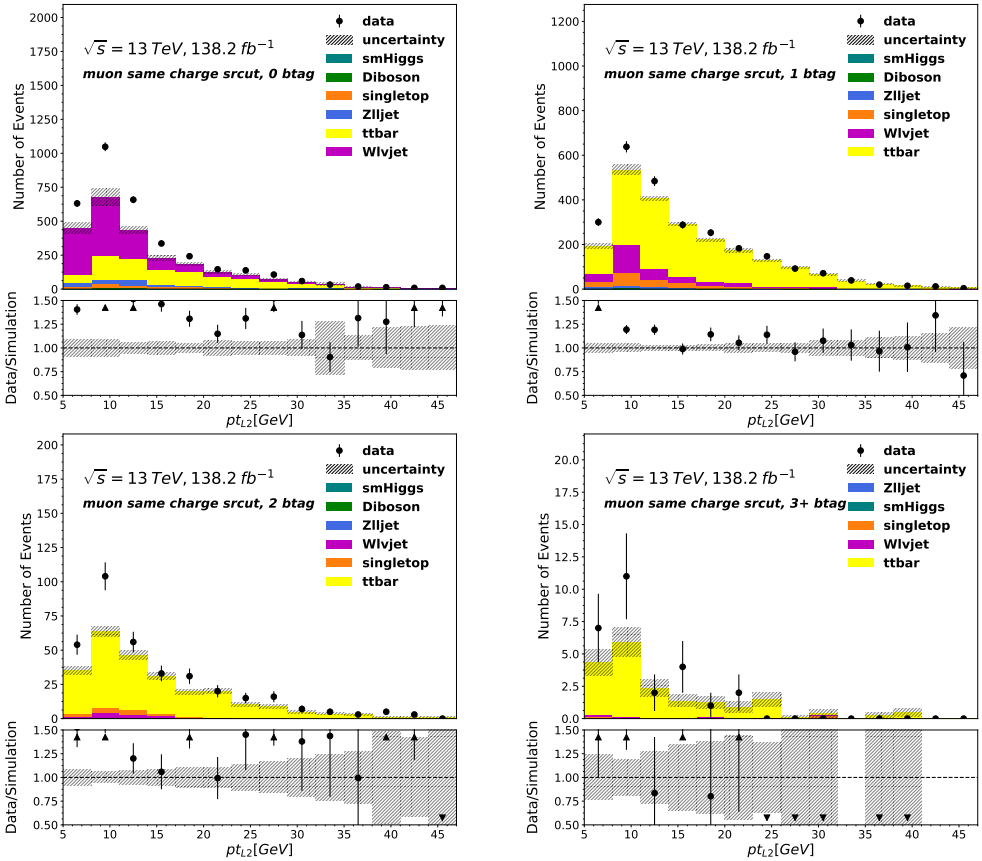


Figure 6.13: The distribution of the second lepton p_T with the resolved signal region cut while requiring two muons of the same charge in the 0 b-jet (top left), 1 b-jet (top right), 2 b-jets (bottom left) and 3+ b-jets (bottom right) regions. The uncertainties on data and MCs are statistical.

A tighter cut on second-leading lepton p_T will reduce the non-prompt lepton contribution but it may also affect the analysis sensitivity. The analysis sensitivity is estimated by the discovery significance given by the formula [109]

$$\sigma = \sqrt{\sum_i 2 \cdot [(S_i + B_i) \cdot ln(1 + \frac{S_i}{B_i}) - S_i]}, \tag{6.2}$$

where S_i and B_i are the numbers of expected signal and background events in the i^{th} bin. The significances of the 2HDM ggA search and the 2HDM bbA search with a mass

hypothesis of m_A=300, 700, 1000 GeV as a function of the second-leading lepton p_T are shown in Figure 6.14. The significances of m_A=700, 1000 GeV signals are stable as the cut threshold increases. For low mass signals (m_A=300 GeV), the ggA signal significance is stable as the cut value increases up to 20 GeV. There is a less than 10% of the bbA signal significance loss in the 2 or more b-tagged jets categories as the cut increases from 7 GeV to 20 GeV. Since there is no significant sensitivity loss, the second-leading lepton p_T cut is increased to 20 GeV to reduce the mismodelling effect in the full Run 2 analysis. It should be noted that the normalizations of the background distributions used in this study are from MC simulation, which generally deviates from the normalizations determined by fits to data. However the p_T cut optimisation should be robust with respect to the normalisation of the MC, as it was designed to remove the significant excess at low p_T corresponding to the mismodelled non-prompt contribution from W+jets and multi-jets, which persists after this is taken into account.

6.4 Systematic Uncertainty

This section describes the systematic uncertainties. Two types of systematic uncertainties are considered in the analysis: modelling uncertainties and experimental uncertainties. The modelling uncertainties are related to the modelling of the signal and background processes in the simulation, while the experimental uncertainties are related to the reconstruction and calibration of objects. The systematic uncertainties are taken into account by the means of nuisance parameters which will be discussed in Section 6.5.

6.4.1 Modelling Uncertainty

The modelling uncertainties affect the shape and the normalization of the distributions of kinematic variables. The modelling uncertainties arise due to the order in QCD calculation, the parton shower, the hadronization, the parton distribution functions and the renormalization and factorization scales.

The type of the quarks from gluon emissions has an effect on the shape and the normalization of the kinematic-variable distributions. To help model the backgrounds,

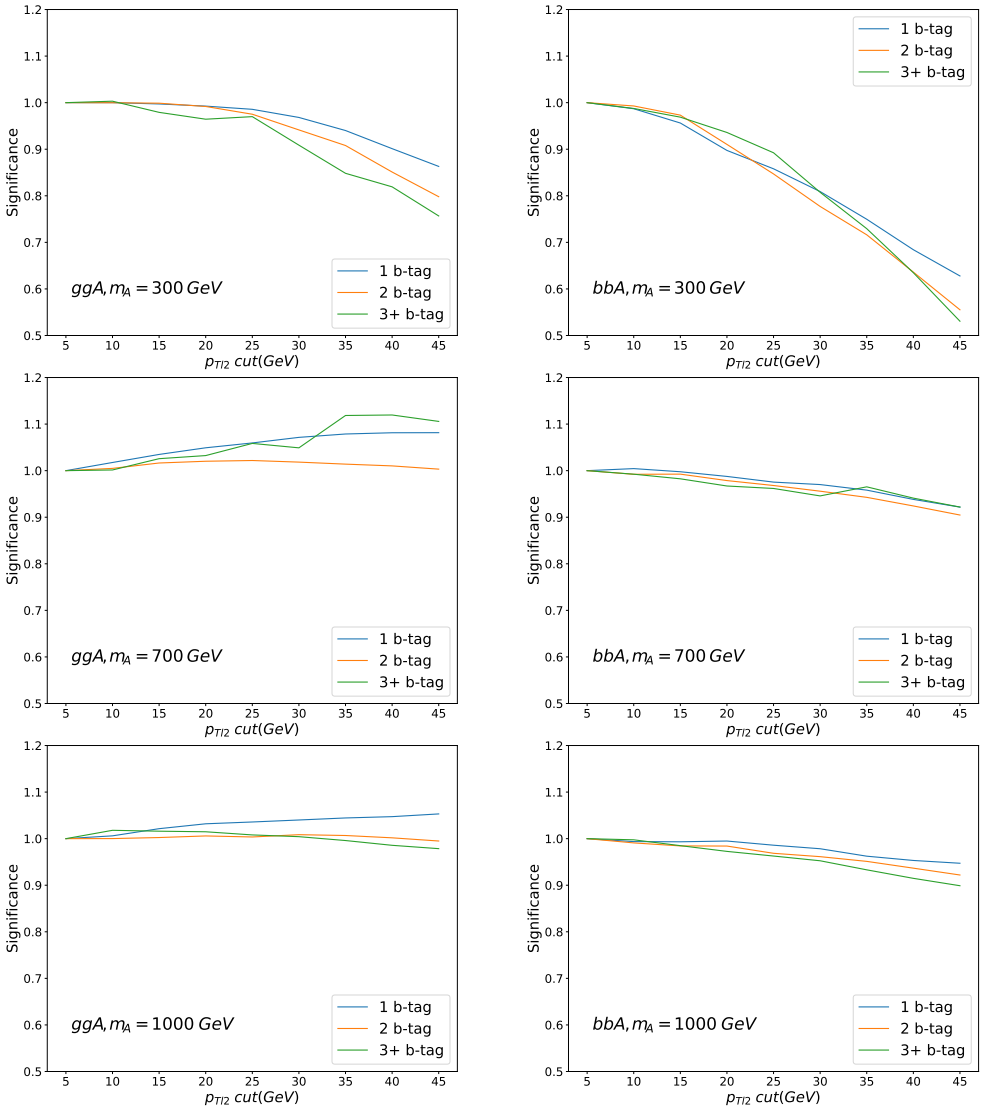


Figure 6.14: The discovery significance for ggA signals (left) and bbA signals (right) of $m_A = 300$ GeV (first row), $m_A = 700$ GeV (second row) and $m_A = 1000$ GeV (third row) in the resolved signal regions as a function of the second leading lepton p_T cut. The significance has been normalized to the expected value when the cut is at 7 GeV.

the V+jets ($V = Z, W$) background is split into $V + (bb, bc, cc)$, $V + (bl, cl)$ and $V + l$ components according to the truth flavour of the jets which are selected to reconstruct the SM Higgs boson, where l means up, down and strange quarks. For the bbA analysis, the $t\bar{t}$ background is split into $t\bar{t}$ with b-jets or c-jets ($t\bar{t} + HF$) and $t\bar{t}$ without b-jets or c-jets ($t\bar{t} + LF$) according to the flavour of the quarks from gluon emissions.

The normalizations of the dominating backgrounds, $t\bar{t}$, $Z + (bb, bc, cc)$ and $Z + (bl, cl)$, are obtained from data. The normalizations of other small backgrounds are

constrained to Standard-Model predictions with the prior normalization uncertainties shown in Table 6.3. Uncertainties on the prior normalizations of the diboson and Standard-Model Vh backgrounds are assigned to cover the effect on both the cross section and event-selection acceptance [110]. An uncertainty of 50% is assigned to the prior normalizations of the $t\bar{t}V$ and $t\bar{t}H$ backgrounds, which gives the fits freedom to adjust their normalizations. A 19% uncertainty is assigned to the normalization of the single top background to cover the effect on generator configurations [7]. The uncertainties on prior normalizations of the $W + (bb, bc, cc)$, $W + (cl, bl)$, $W + (l)$ and $Z + (l)$ backgrounds will be discussed later in this section.

Nuisance Parameter	Description	Regions/Backgrounds	Value	Effect
`XS_stop`	single top normalization	single top in all regions	19%	normalization
`XS_VV`	diboson normalization	diboson in all regions	20%	normalization
`XS_Whf`	$W + (bb, bc, cc)$ normalization	$W + (bb, bc, cc)$ in all regions	30%	normalization
`XS_Wclbl`	$W + (cl, bl)$ normalization	$W + (cl, bl)$ in all regions	30%	normalization
`XS_Wl`	$W + (l)$ normalization	$W + l$ in all regions	19%	normalization
`XS_VH125`	Vh normalization	Vh in all regions	32%	normalization
`XS_ttV`	$t\bar{t}V$ normalization	ttV in all regions	50%	normalization
`XS_ttH`	$t\bar{t}H$ normalization	ttH in all regions	50%	normalization
`XS_Zl`	$Z + (l)$ normalization	$Z + (l)$ in all regions	19%	normalization

Table 6.3: Summary of normalization nuisance parameters of small backgrounds. The forth column shows the values of the prior uncertainties.

The shape uncertainties of the three main backgrounds ($t\bar{t}$, $Z + (bb, bc, cc)$ and $Z + (bl, cl)$) are determined by a comparison between the nominal and alternative samples. The shape uncertainties of other small backgrounds are ignored because their contributions are negligible and these uncertainties have no impact on the result. The alternative distribution is calculated by either a different MC generator or the same generator with a different setting. During the statistical analysis, we assume that the nuisance parameters of those shape uncertainties follow a Gaussian distribution and the $\pm\sigma$ variation is the difference between the nominal distributions and the alternative distributions. The normalization effect of variations can be removed by rescaling the alternative distributions, so they have the same expected event yield as the nominal distribution.

The uncertainties related to the PDF measurement affect the distribution of the

kinematic variables. The analysis considers the PDF uncertainties of the $t\bar{t}$ and Z+jets backgrounds. The NNPDF internal uncertainty is the standard deviation of the 100 alternative NNPDF variations. The PDF modelling uncertainty is calculated by comparing the central value of the NNPDF with those of the CT14 [111] and MMHT2014 [112] and the envelope of all alternative distributions are taken as the systematic uncertainties. The α_s uncertainty is determined by scaling the α_s parameter of the NNPDF by a factor of 0.5 and 2 and the two alternative distributions are taken as $\pm\sigma$ variations.

The initial and final state radiation uncertainties of the $t\bar{t}$ background are calculated by varying the renormalization scales and factorization scales. The envelopes of alternative distributions are taken as the systematic uncertainties. The effect of varying the matrix element generator and parton shower matching on the $t\bar{t}$ distribution is estimated by comparing the nominal sample to the NLO $t\bar{t}$ samples produced by MadGraph5_aMC@NLO 2.2.2 + Pythia 8.212. The effect of the parton shower and the hadronization is estimated by comparing the nominal sample to the PowhegBox v2+Herwig 7.04 sample with the H7UE tuning [113]. The difference between the nominal distribution and the alternative distributions are taken as the uncertainty. The normalization effect of all the $t\bar{t}$ modelling uncertainties is removed because the normalization is determined from data.

The scale uncertainty of the Z/W+jets samples is calculated by halving/doubling the factorization scale and the renormalization scale, and the envelope of all alternative distributions are taken as the uncertainty. The combined effect of the parton shower, hadronization and matching scheme on the Z/W+jets distributions is estimated by the Z/W+jets samples produced by MadGraph5_aMC@NLO 2.2.2 + Pythia 8.186. The difference between the nominal distribution and the alternative distribution is taken as the systematic uncertainty. The normalization effects of the $Z+(bb, bc, cc)$ and $Z+(cl, bl)$ modelling uncertainties are removed because their normalizations are determined from data. The normalization effects of the $Z + l$ and W+jets modelling uncertainties are kept because they are constrained to the Standard-Model prediction.

The PDF uncertainties of the signals are derived by comparing the central values of the nominal PDF with those of the CT10 [114] PDF and the MSTW2008lo [115]

PDF. The signal uncertainties related to the parton shower are estimated by comparing with the HERWIG 7.04 samples. Signal uncertainties are also estimated by varying the renormalization scales and the factorization scales.

The uncertainties are symmetrized to make sure that the nominal distribution lies symmetrically between the $+\sigma$ variation and the $-\sigma$ variation. The distribution of kinematic variables are presented by histograms. To reduce statistical fluctuation during the uncertainty calculation, the binning of the histogram is chosen so that the statistical uncertainty in each bin is less than 20% of the total weight in the bin. Moreover, a Gaussian kernel smoothing [116] is applied to reduce the effect of the statistical fluctuation.

6.4.2 Experimental Uncertainty

The potential difference between experiment and detector simulation is addressed by experimental uncertainties. The relevant experimental uncertainties for this analysis are listed below:

***b*-tagging:** Scale factors accounting for the disagreement in the b-tagging efficiencies between data and simulation are calculated separately for b-jets, c-jets and light-flavour jets [117–119], introducing three, four and five systematic uncertainties respectively using eigenvector decomposition of error matrices. τ and high-p_T jets cannot be calibrated by the measurements, which introduces two extrapolation uncertainties. The number of b-tagging-related uncertainties are doubled because two different types of jets (small-R and large-R) are used for the reconstruct of h bosons.

Trigger: Uncertainties on the difference in scales factors which correct the trigger efficiency between measurements made using different processes introduce one electron trigger efficiency uncertainty [120] and one muon trigger efficiency uncertainty [121]. For the muon trigger efficiency correction, one statistical uncertainty and one systematic uncertainty due to sample dependence are also considered in the analysis.

Luminosity: The uncertainty on the integrated luminosity is 2.1% for the 2015-2016 dataset, 2.4% for the 2017 dataset and 2.0% for the 2019 dataset. The uncertainty on the integrated luminosity of the full Run 2 dataset is 1.7% [122].

Electron: Three systematic uncertainties are caused by the uncertainties on efficiency scale factors of electron reconstruction, identification and isolation accounting for the difference between data and simulation [120]. Two more electron systematic uncertainties are caused by the uncertainties on the energy resolution and energy scale [120].

Muon: Scale factors for muon reconstruction, isolation and track-vertex association efficiencies are applied to account for the difference between data and simulation [121]. Overall, there are four systematic uncertainties for the statistical and systematic uncertainties of the muon reconstruction efficiency measurement in the low-p_t and the high-p_T region, two systematic uncertainties for the statistical and systematic uncertainties of the track-vertex association efficiency measurement, and one for the isolation efficiency measurement. The detector energy resolution also contributes to the muon systematic uncertainties, which introduces four more systematic uncertainties related to muons.

Large-R jet: Five jet energy scale uncertainties related to the difference in jet p_T to track-jet p_T ratio between data and simulation, track measurement uncertainties, the difference between Pythia and Sherpa samples, statistical uncertainty and difference between the TCC jet and calo-jet p_T, and one jet energy resolution uncertainty are used in the analysis [123, 124].

Small-R jet: Thirty jet energy scale uncertainties related to in situ analysis which calculates the difference between data and simulation in different kinematic phase space, η intercalibration, pileup, flavour response and composition, punch through, high-p_T jet behaviours and b-jets response, eight jet energy resolution uncertainties and one jet vertex tagging uncertainty are used in the analysis [125, 126].

Missing Transverse Momentum: The missing transverse momentum is calculated using the reconstructed objects [127, 128]. As a result, all systematic uncertainties related to the reconstructed objects contribute to the missing transverse momentum systematic uncertainties. In addition, the longitudinal resolution, transverse resolution and energy scale of soft terms (unassociated tracks) introduce three systematic uncertainties. Furthermore, one additional systematic uncertainty is related to the energy scale of tracks in jets [129, 130].

All nuisance parameters related to experimental uncertainties are listed in Table 6.4

and Table 6.5.

Nuisance Parameter	Description
Event	
luminosity	uncertainty on integrated luminosity
PRW_DATASF	uncertainty on pileup reweighting
Track jet b-tagging	
FT_EFF_Eigen_Light_0-4_AntiKtVR30Rmax4Rmin02TrackJets	b-tagging efficiency uncertainties of light jets
FT_EFF_Eigen_B_0-2_AntiKtVR30Rmax4Rmin02TrackJets	b-tagging efficiency uncertainties of b-jets
FT_EFF_Eigen_C_0-3_AntiKtVR30Rmax4Rmin02TrackJets	b-tagging efficiency uncertainties of c-jets
FT_EFF_extrapolation_AntiKtVR30Rmax4Rmin02TrackJets	b-tagging efficiency uncertainties of high-p_T jets
FT_EFF_extrapolation_from_charm_AntiKtVR30Rmax4Rmin02TrackJets	b-tagging efficiency uncertainties of τ jets
Small-R jet b-tagging	
FT_EFF_Eigen_Light_0-4_AntiKt4EMTopoJets	b-tagging efficiency uncertainties of light jets
FT_EFF_Eigen_B_0-2_AntiKt4EMTopoJets	b-tagging efficiency uncertainties of b-jets
FT_EFF_Eigen_C_0-3_AntiKt4EMTopoJets	b-tagging efficiency uncertainties of c-jets
FT_EFF_extrapolation_AntiKt4EMTopoJets	b-tagging efficiency uncertainties of high-p_T jets
FT_EFF_extrapolation_from_charm_AntiKt4EMTopoJets	b-tagging efficiency uncertainties of τ jets
Electron	
EG_SCALE_ALL	energy scale uncertainty
EG_RESOLUTION_ALL	energy resolution uncertainty
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty
Muon	
MUON_MS	energy resolution uncertainty of muon system
MUON_ID	energy resolution uncertainty of inner detector
MUON_SCALE	energy scale uncertainty
MUON_SAGITTA_RESBIAS	charge-dependent energy scale uncertainty
MUON_SAGITTA_RHO	charge-dependent energy scale uncertainty
MUON_EFF_TTVA_SYS	track-to-vertex association efficiency uncertainty
MUON_EFF_TTVA_STAT	track-to-vertex association efficiency uncertainty
MUON_EFF_ISO_SYS	isolation efficiency uncertainty
MUON_EFF_ISO_STAT	isolation efficiency uncertainty
MUON_EFF_RECO_SYS_LOWPT	low-p_T reconstruction and ID efficiency uncertainty
MUON_EFF_RECO_STAT_LOWPT	low-p_T reconstruction and ID efficiency uncertainty
MUON_EFF_RECO_SYS	reconstruction and ID efficiency uncertainty
MUON_EFF_RECO_STAT	reconstruction and ID efficiency uncertainty
MUON_EFF_TrigSystUncertainty	trigger efficiency uncertainty
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty

Table 6.4: Summary of nuisance parameters of experimental uncertainties for events, track jets b-tagging, small-R jets b-tagging, electrons and muons.

Nuisance Parameter	Description
Missing transverse energy	
MET_SoftTrk_Scale	track-based soft term scale uncertainty
MET_JetTrk_Scale	scale uncertainty of track in jet
MET_SoftTrk_ResoPerp	track-based soft term transverse resolution uncertainty
MET_SoftTrk_ResoPara	track-based soft term longitudinal resolution uncertainty
METTrigTop	sample-dependence uncertainty on pileup reweighting
METTrigStat	trigger efficiency uncertainty due to statistics
Large-R jet	
FATJET_JER	energy resolution uncertainty
FATJET_Medium_JET_Rtrk_Modelling_pT	P_T scale uncertainty due to Pythia/Sherpa difference
FATJET_Medium_JET_Rtrk_TotalStat_pT	P_T scale uncertainty due to statistics
FATJET_Medium_JET_Rtrk_Tracking_pT	P_T scale uncertainty due to track measurement
FATJET_Medium_JET_Rtrk_Baseline_pT	P_T scale uncertainty due to data/simulation difference
FATJET_Medium_JET_Rtrk_Closure_pT	P_T scale uncertainty due to TCC/calo jets difference
Small-R jet	
JET_JvtEfficiency	vertex tagging efficiency uncertainty
JET_CR_JET_JER_EffectiveNP_1-7(restTerm)	energy resolution uncertainties
JET_CR_JET_JER_DataVsMC	energy resolution uncertainty
JET_CR_JET_BJES_Response	energy scale uncertainty for b-jets response
JET_CR_JET_SingleParticle_HighPt	energy scale uncertainty for high-p_T jets
JET_CR_JET_PunchThrough_MC16	energy scale uncertainty for punch through jets
JET_CR_JET_Pileup_RhoTopology	energy scale uncertainty due to pileup
JET_CR_JET_Pileup_PtTerm	energy scale uncertainty due to pileup
JET_CR_JET_Pileup_OffsetNPV	energy scale uncertainty due to pileup
JET_CR_JET_Pileup_OffsetMu	energy scale uncertainty due to pileup
JET_CR_JET_Flavor_Response	energy scale uncertainty related to flavour response
JET_CR_JET_Flavor_Composition	energy scale uncertainty related to flavour composition
JET_CR_JET_EtaIntercalibration_TotalStat	energy scale uncertainty of η intercalibration
JET_CR_JET_EtaIntercalibration_NonClosure_posEta	energy scale uncertainty of η intercalibration
JET_CR_JET_EtaIntercalibration_NonClosure_negEta	energy scale uncertainty of η intercalibration
JET_CR_JET_EtaIntercalibration_NonClosure_highE	energy scale uncertainty of η intercalibration
JET_CR_JET_EtaIntercalibration_Modelling	energy scale uncertainty of η intercalibration
JET_CR_JET_EffectiveNP_Statistical1-6	energy scale uncertainties of in situ analysis
JET_CR_JET_EffectiveNP_Modelling1-4	energy scale uncertainties of in situ analysis
JET_CR_JET_EffectiveNP_Mixed1-3	energy scale uncertainties of in situ analysis
JET_CR_JET_EffectiveNP_Detector1-2	energy scale uncertainties of in situ analysis

Table 6.5: Summary of nuisance parameters of experimental uncertainties for missing transverse energy and jets.

6.5 Background Modelling and Statistical Analysis

The analysis fits simulated distributions of kinematic variables to data for signal extraction and limit setting. The fit is performed using distributions of the reconstructed invariant mass of the VH object (m_{Vh}). In order to improve the data/simulation agreement, a data-driven method is developed to improve the modelling of the backgrounds by reweighting the simulated events.

This section will first introduce the statistical method of the fit. It will then discuss the reweighting of the MC events. The fit models and the results of the signal searches will be discussed afterwards.

6.5.1 Statistical Method

This section describes the statistical methods for signal extraction and limit setting. In the ggA, bbA and Z' searches, the possible existence of new physics will not have significant interference effects with the Standard-Model background processes, so the expected number of events n_i in the i^{th} bin of a histogram for a distribution can be written as

$$E[n_i] = b_{\text{total\_}i} + \mu s_i, \tag{6.3}$$

where s_i and $b_{\text{total\_}i}$ are the expected number of signal events and total background events in the i^{th} bin [131]. The expected number of signal events is scaled by a variable called signal strength μ. μ equals zero in the background-only hypothesis. $b_{\text{total\_}i}$ can be written as a sum of expected numbers of events from different background processes:

$$b_{\text{total\_}i} = \sum_j b_{ij}, \tag{6.4}$$

where b_{ij} is the expected number of events from the j^{th} background process in the i^{th} bin [131].

The statistical and systematic uncertainties also need to be taken into account when estimating the number of signal and background events in a bin. The effect of uncertainties is incorporated in the statistical analysis by the means of nuisance

parameters [132]. The nuisance parameter may shift the number of events from the nominal distribution. The estimated value of a nuisance parameter is not of interest in the search, but it may affect the result.

The expected numbers of signal and background events are estimated using a binned maximum-likelihood fit [133]. The likelihood of a possible simulated signal-plus-background distribution given a certain data distribution can be calculated by the product of Poisson probabilities:

$$L_{Poisson}(\mu, \vec{\alpha}, \vec{\tau}, \vec{\gamma}) = \prod_{i=0}^{n_{bins}} Poisson(N_i|\mu s_i(\vec{\alpha}) + b_{\text{total\_}i}(\vec{\alpha}, \vec{\tau}, \vec{\gamma})), \tag{6.5}$$

where $\vec{\alpha}, \vec{\tau}$ and $\vec{\gamma}$ are vectors of three different types of nuisance parameters, n_{bins} is the total number of bins and N_i is the number of data entries in the i^{th} bin.

Prior knowledge about nuisance parameters helps to constrain the fit if it is also included in the likelihood function. The systematic uncertainties with prior knowledge, such as the normalization uncertainties and shape uncertainties, are controlled by nuisance parameters in vector $\vec{\alpha}$. A nuisance parameter with a value of $\pm\alpha_i$ varies its corresponding uncertainty i up/down by $\alpha_i\sigma$, where σ is the prior uncertainty. These nuisance parameters follow the standard normal distribution, so the likelihood function which constrains $\vec{\alpha}$ is

$$L_{Alpha}(\vec{\alpha}) = \prod_{i=0}^{n_\alpha} Gauss(\alpha_i|0, 1), \tag{6.6}$$

where n_α us the total number of nuisance parameters in $\vec{\alpha}$.

The statistical uncertainty on the total background in each bin σ_{bi} is controlled by the nuisance parameters in vector $\vec{\gamma}$. A nuisance parameter which controls the statistical fluctuation of the i^{th} bin with a value of γ_i varies the bin height by $(\gamma - 1)\sigma_{bi}$. The nuisance parameters follow a Gamma distribution, so the likelihood function which constrains $\vec{\gamma}$ is

$$L_{Gamma}(\vec{\gamma}) = \prod_{i=0}^{n_{bins}} Gamma(\gamma_i|(\sigma_{bi}/b_{\text{total\_}i})^2, n_i - 1). \tag{6.7}$$

The nuisance parameters in vector $\vec{\tau}$ are unconstrained and they are allowed to float

freely during the fit. The value of these nuisance parameters are decided by the data completely, so there is no likelihood function which constrains them.

The total likelihood function is the product of the three likelihood functions described above:

$$L(\mu, \vec{\theta}) = L_{Poisson}(\mu, \vec{\alpha}, \vec{\tau}, \vec{\gamma}) L_{Alpha}(\vec{\alpha}) L_{Gamma}(\vec{\gamma}), \tag{6.8}$$

where $\vec{\theta}$ is a vector containing all nuisance parameters.

A fit to data maximizes the likelihood function by varying both the signal strength and the nuisance parameters. $L(\hat{\mu}, \hat{\theta})$ denotes a maximized likelihood function, where $\hat{\mu}$ and $\hat{\theta}$ are the signal strength and the nuisance parameter which maximize the likelihood function respectively. Since the maximum likelihood estimators of nuisance parameters are not of interest, a profile likelihood fit which maximizes the likelihood function by varying only the nuisance parameters at a fixed signal strength is used in the analysis. The signal strength is the parameter of interest in the profile likelihood fit. The maximized likelihood function at the signal strength μ is $L(\mu, \hat{\hat{\theta}})$, where $\hat{\hat{\theta}}$ is the maximum likelihood estimator of the nuisance parameters for a given value of μ .

A statistic [131] which increases as the incompatibility between the data and the signal strength increases is defined by

$$t_\mu = -2\ln\lambda(\mu), \tag{6.9}$$

where $\lambda(\mu)$ is

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{b}})}{L(\hat{\mu}, \hat{b})}. \tag{6.10}$$

The statistic t_μ approaches a chi-square distribution [132], which simplifies the computation. In the signal search, we assume that the signal strength should always be positive, so the analysis uses an alternative definition which is given by

$$t_\mu = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{b}})}{L(\hat{\mu},\hat{b})} & 0 \leq \hat{\mu}, \\ -2\ln\frac{L(\mu,\hat{\hat{b}})}{L(0,\hat{\hat{b}})} & 0 > \hat{\mu}. \end{cases} \tag{6.11}$$

To test if the distribution of data is compatible with the Standard-Model prediction or not,

the background-only hypothesis is tested against a signal-plus-background hypothesis using the statistic t_μ. The level of incompatibility between data and a hypothesis with signal strength μ is described by the p-value

$$p(\mu) = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu)dt_\mu, \tag{6.12}$$

where $t_{\mu,obs}$ is the statistic t_μ determined by the data and $f(t_\mu|\mu)$ is the conditional probability density function of t_μ under the condition that the signal strength is μ. The p-value gives the level of incompatibility with the background-only hypothesis when $\mu = 0$.

The significance of a signal Z can be determined by testing the data against the background-only hypothesis. The signal significance Z can be defined by the number of standard deviations of a point away from the zero point in the standard normal distribution. The relationship between the p-value and the signal significance is

$$p = 1 - \Phi(Z), \tag{6.13}$$

where Φ is the cumulative distribution function of the standard normal distribution. Discovery of a signal can be claimed when the significance excesses 5.

The search can set an upper limit on the signal strength in the case of no discovery. The 95% upper limit of the signal strength can be defined by the value of μ when the p-value $p(\mu)$ is 5%. However, the method may reject a signal if the probability density function of the background-only hypothesis is similar to that of the background-plus-signal hypothesis, hence the limit is determined by an alternative statistic CL_s [134] which is defined as

$$CL_s = \frac{p(\mu)}{1 - p(0)}. \tag{6.14}$$

The 95% upper limit is the signal strength when $CL_s = 5\%$. The value of the signal strength is determined by solving the equation numerically.

The post-fit values and uncertainties of nuisance parameters can help to understand the fit. The Cramér-Rao bound [135, 136] suggests that the covariance matrix of unbiased

estimators is the reverse of the Fisher information matrix. In the analysis, we assume that the bias is small, so the covariance matrix of the parameters for the fit is

$$\mathrm{cov}(\theta_i, \theta_j) = (-\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j}\bigg|_{\hat{\vec{\theta}}})^{-1}. \tag{6.15}$$

The pull of a constrained nuisance parameter is

$$\mathrm{pull} = \frac{\hat{\theta} - \theta}{\sigma_\theta}, \tag{6.16}$$

where θ is the prior value of the constrained nuisance parameter and σ_θ is its prior uncertainty. All prior values are zero in the analysis. The pull of an unconstrained nuisance parameter cannot be calculated because it does not have a prior uncertainty.

The expected limit of the search can be determined by replacing the observed data with an Asimov dataset which follows the same distribution as the simulation in a binned profile likelihood fit. All nuisance parameters are expected to have zero pulls in an Asimov fit. The Asimov fit estimates the sensitivity of a search. It also allows the study of possible correlations between nuisance parameters.

6.5.2 Background Simulation Reweighting

The dominating background in the analysis is the Z+jets process. However, the Z+jets background is poorly modelled by simulation, and disagreement between the data and the simulation have been observed in the m_{bb} sideband control regions. A data-driven method was developed to improve the modelling of the Z+jets background.

The reweighting of the Z+jets background in resolved regions

A slope in the data-to-MC ratio of the p_T^{BB} distributions was found in the m_{bb} sideband control regions. The main backgrounds in these regions are the Z+jets background and the $t\bar{t}$ background. The slope is not observed in the top control region, which indicates the mismodelling is related to the Z+jets background. The slope is corrected by reweighting the Z+jets background, where the reweighting function is determined using the p_T^{BB} distributions in the m_{bb} sideband control regions. To determine the

reweighting function, each constituent of the background is rescaled first using the normalization factors obtained from a fit to data in the m_{bb} sideband regions and the top control region using the m_{VH} distributions. The normalization factors are given in Table 6.6. The fit does not consider any experimental or theory uncertainties which affect the shape of the distributions. The distributions after the rescaling are shown in Figure 6.15. After the rescaling, the reweighting functions are determined by fits to data subtracting all backgrounds except Z+jets, which is shown in Figure 6.16. It should be noted that the Z+jets theory uncertainties are not shown on the figures and could account for some of the observed mis-modelling, leading to correlations between these uncertainties and this correction.

t, $t\bar{t}$	Z+(bl,cl,l)	Z+(bb,bc,cc)	VV	SM Higgs
1.017 ± 0.006	1.143 ± 0.005	1.286 ± 0.013	0.770 ± 0.169	0.790 ± 0.277

Table 6.6: Normalization factors applied to determine the reweighting function. The factors are determined from a m_{VH} fit in the 1 and 2 b-tagged resolved and merged sideband control regions and the 1+2 b-tagged resolved top control with normalization nuisance parameters only. The uncertainties on the normalization factors are statistical only.

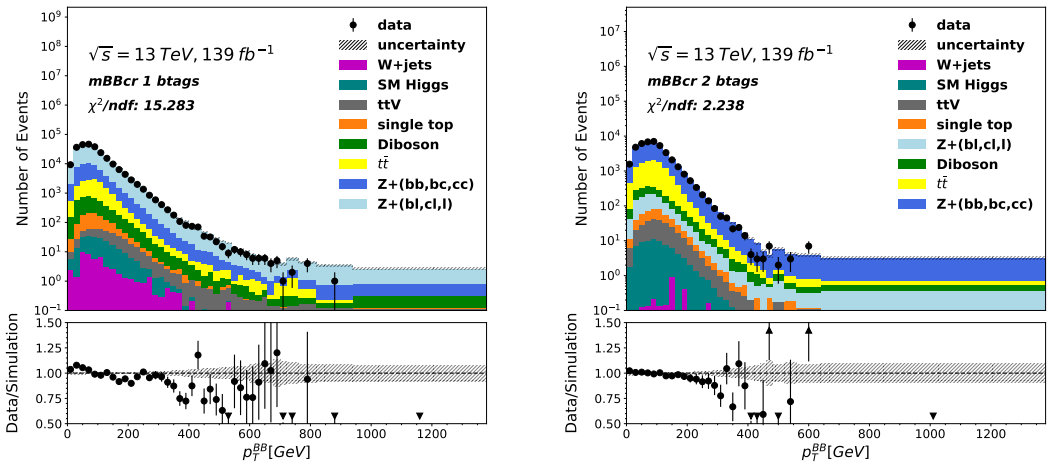


Figure 6.15: Distributions of p_T^{BB} in the 1 b-tagged (left) and 2 b-tagged (right) resolved m_{bb} sideband control regions after applying the normalization factors. The uncertainties on the simulated distributions are statistical only. The theory uncertainties of the Z+jets background modelling are not included, but they could account for some of the difference between data and simulation.

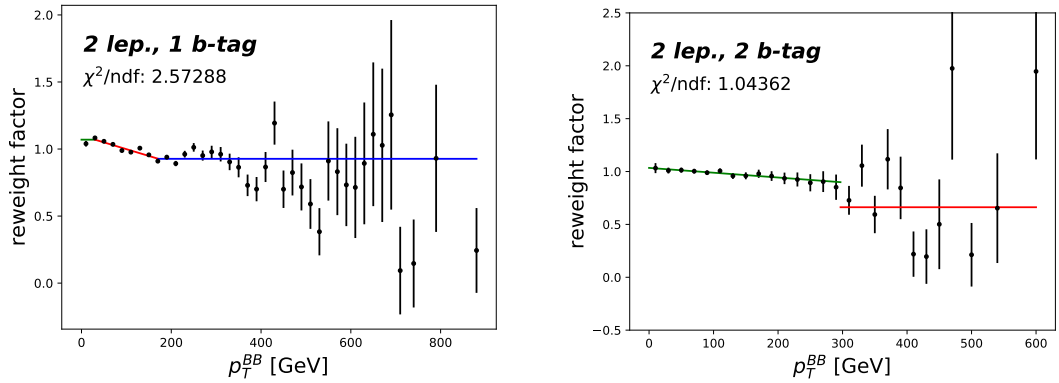


Figure 6.16: Reweighting functions for the Z+jets background in 1 b-tagged (left) and 2 b-tagged (right) resolved regions.

The reweighting functions are determined for 1 b-tagged events and 2 b-tagged events separately. The function in the 1 b-tagged region is

$$w = \begin{cases} p_0 & p_T^{BB} \leq p_1, \\ p_2 \times (p_T^{BB} - p_1) + p_0 & p_1 \leq p_T^{BB} \leq p_3, \\ p_2 \times (p_3 - p_1) + p_0 & \text{other}, \end{cases} \tag{6.17}$$

where p_0, p_1, p_2 and p_3 are free parameters in the fit. The function in the 2 b-tagged region is

$$w = \begin{cases} p_0 + p_1 \times p_T^{BB} & p_T^{BB} \leq p_2, \\ p_3 & \text{other}, \end{cases} \tag{6.18}$$

where p_0, p_1, p_2 and p_3 are free parameters in the fit. The fits use piecewise functions where the intervals are also determined by the fits. The functional form helps to account for the distinct mismodelling trends observed in the different p_T^{BB} regions of phase space and prevents unrealistic extrapolations at high p_T^{BB}. The parameters determined by the fit are shown in Table 6.7 and the distributions of the p_T^{BB} after the reweighting are shown in Figure 6.17. The plots show that the reweighting improves the agreement between data and simulation.

The reweighting function is determined in the m_{bb} sideband control region with $50\,\text{GeV} < m_{bb} < 100\,\text{GeV}$ or $145\,\text{GeV} < m_{bb} < 200$ GeV, but it will be applied to the signal regions with $100\,\text{GeV} < m_{bb} < 145\,\text{GeV}$. To check the m_{bb} dependence of

	1 b-tagged	2 b-tagged
p_0	1.070 ± 0.007	1.034 ± 0.017
p_1	32.00 ± 0.19	$-(4.546 \pm 1.538) \times 10^{-4}$
p_2	$-(1.011 \pm 0.093) \times 10^{-3}$	286.8 ± 10.8
p_3	171.3 ± 12.463	$(6.662 \pm 0.730) \times 10^{-1}$

Table 6.7: Parameters of the reweighting functions for the resolved events.

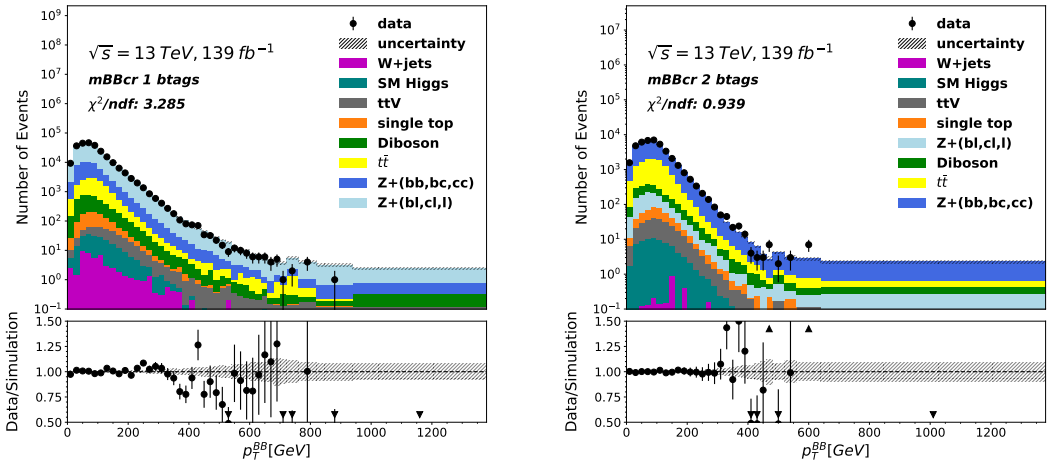


Figure 6.17: Distributions of p_T^{BB} the 1 b-tagged (left) and 2 b-tagged (right) m_{bb} sideband control regions after applying the normalization factors and the reweighting. The uncertainties on the simulated distributions are statistical only.

the reweighting, the reweighting functions are determined in the low-m_{bb} sideband control region and high-m_{bb} sideband control region separately. Figure 6.19 shows the distributions in the high-m_{bb} sideband control region before and after applying reweighting determined in the low-m_{bb} sideband control region. Figure 6.18 shows the distributions in the low-m_{bb} sideband control region before and after applying reweighting determined in the high-m_{bb} sideband control region. As can be seen from the plots, the slopes on the p_T^{BB} distributions are corrected and the χ^2 shows that the data/MC agreement is improved after applying reweighting functions determined using orthogonal events. However, the reweighting functions determined from the low-m_{bb} and high-m_{bb} sideband control regions are found to be different. To address the m_{bb} dependence of the reweighting, the distributions before the reweighting are used as systematic uncertainties, which allows the fit to determine the amount of slope correction needed.

The background distribution of the 3+ b-tagged merged m_{bb} region is shown in

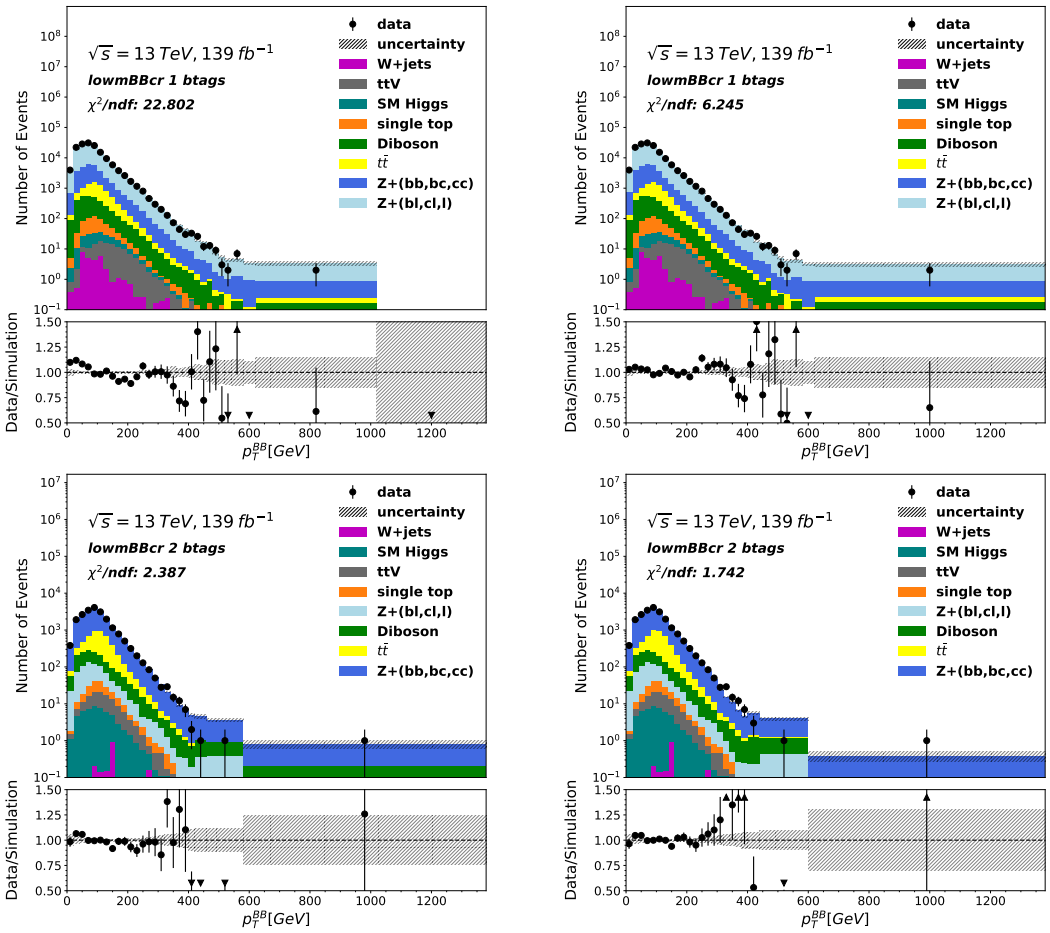


Figure 6.18: Distributions of p_T^{BB} in the 1 b-tagged (top) and 2 b-tagged (bottom) low-m_{bb} sideband control region before (left) and after (right) applying the reweighting determined using events in the high-m_{bb} sideband control regions. The uncertainties on the simulated distributions are statistical only.

Figure 6.20. The plot shows good data/MC agreement, and no reweighting is applied.

The reweighting of backgrounds in the merged regions

The reweighting is also applied to all background events in the 1 b-tagged merged region. The distribution of the transverse momentum of leading large-R jets (p_T^{FJ}) in the 1 b-tagged m_{bb} sideband control region is shown in Figure 6.21. It is clear that the slope in the resolved regions is related to the Z+jets background since the top control region has good data/MC agreement. As a result, the reweighting is only applied to the Z+jets background in the resolved regions. However, the reweighting and associated uncertainties are applied to all background events in the merged region because the

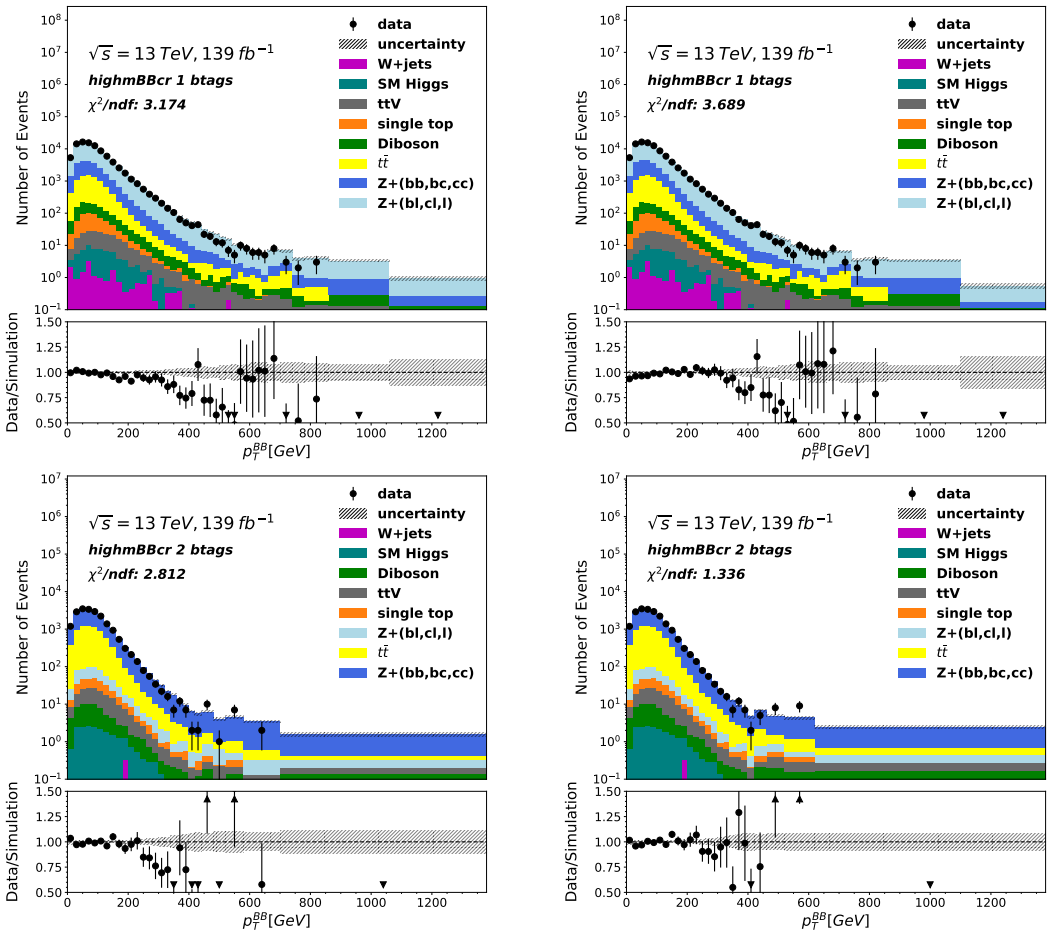


Figure 6.19: Distributions of p_T^{BB} in the 1 b-tagged (top) and 2 b-tagged (bottom) high-m_{bb} sideband control region before (left) and after (right) applying the reweighting determined using events in the low-m_{bb} sideband control regions. The uncertainties on the simulated distributions are statistical only.

cause of the mismodelling is less clear. The mismodelling is corrected by reweighting the p_T^{FJ} distribution, where the reweighting function is determined by fitting the distribution to data. Before the fit, the scale factors of each background process determined from a fit using the distributions of m_{VH}, shown in Table 6.8, are applied.

The reweighting function, shown in Figure 6.22, is $w = p_0 + p_1 \times p_T^{FJ} + p_2 \times (p_T^{FJ})^2$, where $p_0 = 1.219 \pm 0.211$, $p_1 = -(8.180 \pm 8.953) \times 10^{-4}$ and $p_2 = -(3.548 \pm 8.219) \times 10^{-7}$. The p_T^{FJ} and m_{VH} distributions before and after the reweighting are shown in Figure 6.21 and Figure 6.23, and the plots show the reweighting improves the modelling of backgrounds. The difference between the reweighed distribution and the original distribution is used as a systematic uncertainty during the fit. The reweighting is not

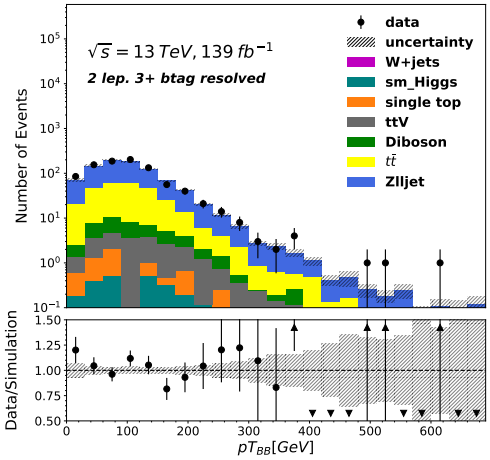


Figure 6.20: The distribution of p_T^{BB} in the 3+ b-tagged m_{bb} sideband control region after applying the normalization factors obtained from a fit. The normalization factors are: 1.301 for Z+jets, 1.316 for $t\bar{t}+HF$ and 0.904 for $t\bar{t}+LF$. The uncertainties on the simulated distributions are statistical only.

applied to the 2 b-tagged merged regions because of the large statistical uncertainties.

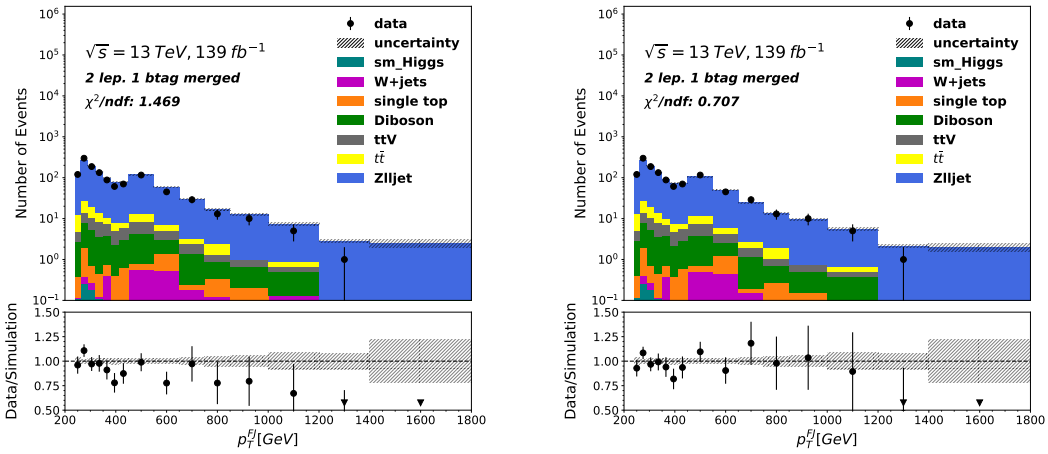


Figure 6.21: Distributions of p_T^{FJ} in the 1 b-tagged m_{bb} sideband control region before (left) and after (right) the reweighting. The uncertainties on the simulated distributions are statistical only.

t, $t\bar{t}$	Z+(bl,cl,l)	Z+(bb,bc,cc)	VV	SM Higgs
1.016 ± 0.177	0.968 ± 0.043	1.191 ± 0.108	0.770 ± 0.169	0.790 ± 0.277

Table 6.8: Normalization factors for the backgrounds in the merged regions. The factors are determined from a m_{VH} fit in the 1 and 2 b-tagged resolved and merged sideband control regions and the resolved top control. The uncertainties on the normalization factors are statistical only.

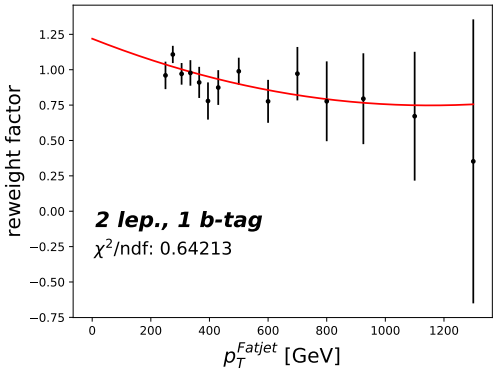


Figure 6.22: The reweighting function for events in the 1 b-tagged merged region.

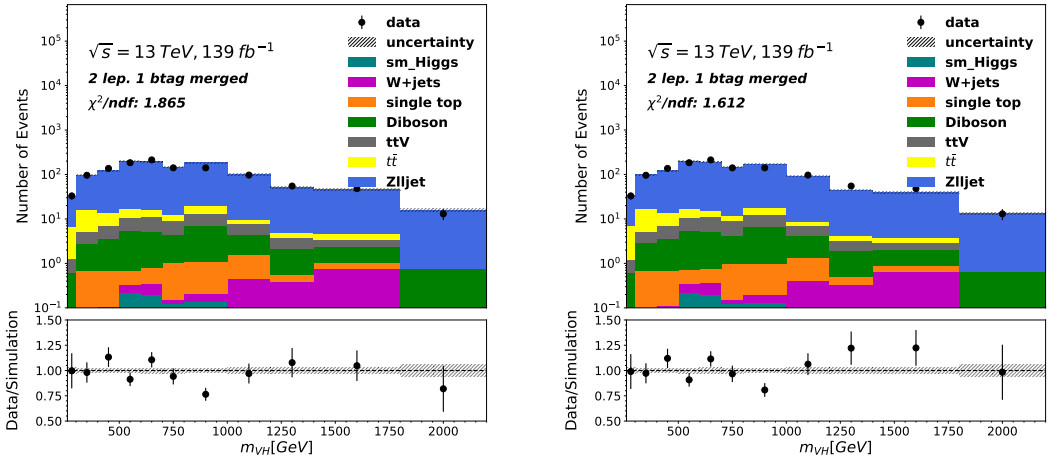


Figure 6.23: Distributions of the m_{VH} in the 1 b-tagged m_{bb} sideband control region before (left) and after (right) the reweighting. The uncertainties on the simulated distributions are statistical only.

6.5.3 Statistical Analysis

The analysis searches for three different types of signals. The ggA and the Z' searches share the same fit model because they have the same final states. A dedicated fit model is used for the bbA fit to model the additional b-jets in the final state. Before searching for resonance in the signal regions, the modelling of the background is first validated using events in the control regions.

The ggA and Z' fits

To extract the ggA and the Z' signals, the m_{VH} distributions of the simulated events are fitted to data in a binned profile likelihood fit as described in Section 6.5.1. The

parameter of interest is the normalization factor of the ggA or Z' signal distributions (signal strength), while the normalization factors of background distributions and the uncertainties on the background shapes are controlled by nuisance parameters. A mass range of 220 GeV to 5 TeV is probed in the search.

The 1 or 2 b-tagged resolved signal regions and the 1 or 2 b-tagged merged signal regions without additional b-tagged track jets not associated with the large-R jets are sensitive to the ggA and Z' signals, so those regions are used in the fit. In addition to the signal regions, the 1 or 2 b-tagged resolved top control regions are also included in the fit to constrain the $t\bar{t}$ background. To reduce the sensitivity to b-tagging-related uncertainties, the 1 and 2 b-tagged top control regions are combined.

Nuisance Parameter	Description	Regions/Backgrounds	Value	Effect
`norm_Zhf`	Z+(bb,bc,cc) normalization	Z+(bb,bc,cc) in all regions	Float	normalization
`norm_Zclbl`	Z+(bl,cl) normalization	Z+(cl,bl) in all regions	Float	normalization
`norm_top_L2`	top normalization	$t\bar{t}$, single top, ttV, ttZ in all regions	Float	normalization
`Zhf_MerResRatio`	merged / resolved ratio	Z+(bb,bc,cc) in all merged regions	19%	normalization
`Zclbl_MerResRatio`	merged / resolved ratio	Z+(cl, bl) in all merged regions	28%	normalization
`Zl_MerResRatio`	merged / resolved ratio	Z+(l) in all merged regions	23%	normalization
`top_MerResRatio_L2`	merged / resolved ratio	top in all merged regions	18%	normalization
`ttbar_topemuRatio_L2`	SR / topcr ratio	top in all resolved signal regions	1.2%	normalization
`Whf_MerResRatio`	merged / resolved ratio	W+(bb,bc,cc) in all merged regions	14%	normalization
`Wclbl_MerResRatio`	merged / resolved ratio	W+(cl, bl) in all merged regions	19%	normalization
`Wl_MerResRatio`	merged / resolved ratio	W+(l) in all merged regions	17%	normalization

Table 6.9: A summary of the normalization nuisance parameters in the ggA and Z' fit related to the Z+jets, W+jets and top backgrounds. The forth column shows the values of the prior uncertainties.

The normalizations of the largest backgrounds, Z+jets, are determined by data. The overall normalization factor of the Z+(bb,bc,cc) component and that of the Z+(bl,cl) component are allowed to float freely in the fit because there is enough data to constrain the normalization factors. The normalization of the Z+l component is constrained to its Standard-Model prediction. To address the mismodelling of the event acceptance difference between regions, resolved-to-merged normalization factors are applied in the merged regions. These normalization factors have a prior value of one, and their prior uncertainties are calculated by comparing the nominal samples with the alternative samples described in Section 6.4.1. The uncertainty on the acceptance ratio between

region A and region B is

$$\sigma = \sqrt{\sum_{i}^{M} \left(\frac{\left| \frac{n_A^{\mathrm{alt}^i}}{n_B^{\mathrm{alt}^i}} - \frac{n_A^{\mathrm{nom}}}{n_B^{\mathrm{nom}}} \right|}{\frac{n_A^{\mathrm{nom}}}{n_B^{\mathrm{nom}}}} \right)^2}, \tag{6.19}$$

where $n_A^{\mathrm{alt}^i}$ is the number of events of the alternative sample i in region A, $n_B^{\mathrm{alt}^i}$ is the number of events of the alternative sample i in region B, n_A^{nom} is the number of events of the nominal sample in region A, n_B^{nom} is the number of events of the nominal sample in region B, M is the number of alternative samples. A list of normalization factors related to the Z+jets backgrounds and their prior uncertainties is shown in Table 6.9.

The overall normalization factor for the $t\bar{t}$ background is a nuisance parameter which floats freely during the fit. Similar to the treatment for the Z+jets background, a resolved-to-merged normalization factor is applied in the merged regions to address the mismodelling of the event acceptance difference between the resolved and merged regions. Since a dedicated top control region is used to model the $t\bar{t}$ background, another normalization factor is applied to the signal regions to address the mismodelling of the event acceptance difference between the signal regions and the top control region. All normalization factors related to the $t\bar{t}$ background and their prior uncertainties are shown in Table 6.9.

The normalizations of other backgrounds are determined by the fit using constrained nuisance parameters whose prior uncertainties are shown in Table 6.3. A global normalization nuisance parameter accounting for the luminosity uncertainty is applied to both the signal and the background distributions in all regions.

Like the treatment for the Z+jets and $t\bar{t}$ backgrounds, resolved-to-merged normalization factors (shown in Table 6.9) are also applied to the single top background and the W+jets background. Those scale factors have trivial effects in the analysis because the single top and the W+jets are small backgrounds in the 2-lepton channel. However, the normalization factors are still used in the fit for the combination with the 0-lepton channel where the Z boson decays into two neutrinos.

The bin width of the histograms used in the fit also has an effect on the result. The current analysis follows the same binning strategy as the previous analysis [7]. The

signal-region bins are required to be wider than the mass resolution of the reconstructed signal. In addition to the mass resolution requirement, the statistical uncertainty of the simulation in each bin must be less than 20% of the total number of events of the simulated background distributions. In the top control regions, the bin width is fixed at 250 GeV. Since the distribution in the top control regions are not of interest, wider bins are used to reduce the fit's sensitivity to the shape of the distributions, and studies show that additional constraints from control regions do not significantly improve the limits.

Before the signal-region fit, the fit model was validated using events which are orthogonal to the signal regions. The study follows a similar analysis strategy as described above, while the signal region is replaced by the m_{bb} sideband control regions. The distributions of m_{VH} after the fit are shown in Figure 6.24. The plots show good agreement between data and simulation.

After the study in the m_{bb} sideband control regions, the fit was performed using the m_{VH} distribution in the signal regions. The pulls of nuisance parameters from the Asimov fit and the fit using data are shown in Figure 6.25. Overall, the nuisance parameters behave as expected in the fit because all pulls are within -1 to 1, and both the Asimov fit and the fit to data provide a similar constraint on nuisance parameters. The small uncertainties on the pulls of `norm_Zclbl`, `norm_zhl` and `norm_top_L2` show that the data provides a strong constraint on the normalizations of the $Z + (bl, cl)$, $Z + (bb, bc, cc)$ and $t\bar{t}$ backgrounds. The fit suggests that the simulation provides a good estimation for the normalization of the $t\bar{t}$ background, while the normalizations of the Z+jets backgrounds are underestimated. `RW_Zjets_PThReweight1T` and `RW_Zjets_PThReweight2T` are the nuisance parameter for the p_T^{BB} reweighting in the 1-tagged and 2-tagged resolved regions. The small pulls suggest that the reweighting improved the m_{VH} distributions in the signal regions since the fit did not reverse the reweighting.

The distributions of the m_{VH} after the fit in the signal regions and the top control region are shown in Figure 6.26. The data agrees with the Standard-Model prediction. A mild excess of data is observed at around 700 GeV in the resolved regions.

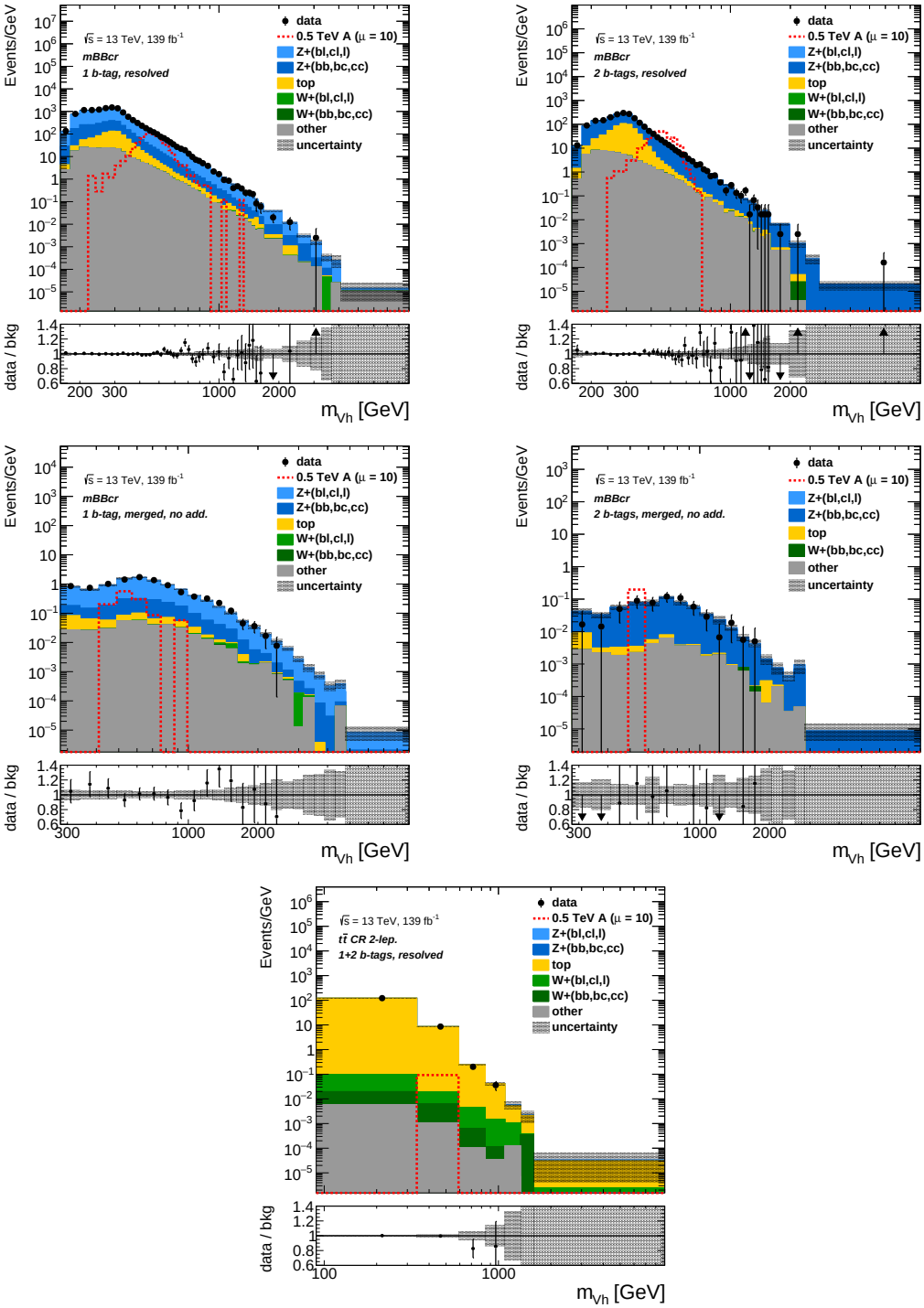


Figure 6.24: Post-fit distributions of the m_{VH} in the 1 b-tagged resolved m_{bb} sideband control region (top left), 2 b-tagged resolved m_{bb} sideband control region (top right), 1 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (middle left), 2 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (middle right) and 1+2 b-tagged resolved top control region (bottom) in the ggA/Z' fit.

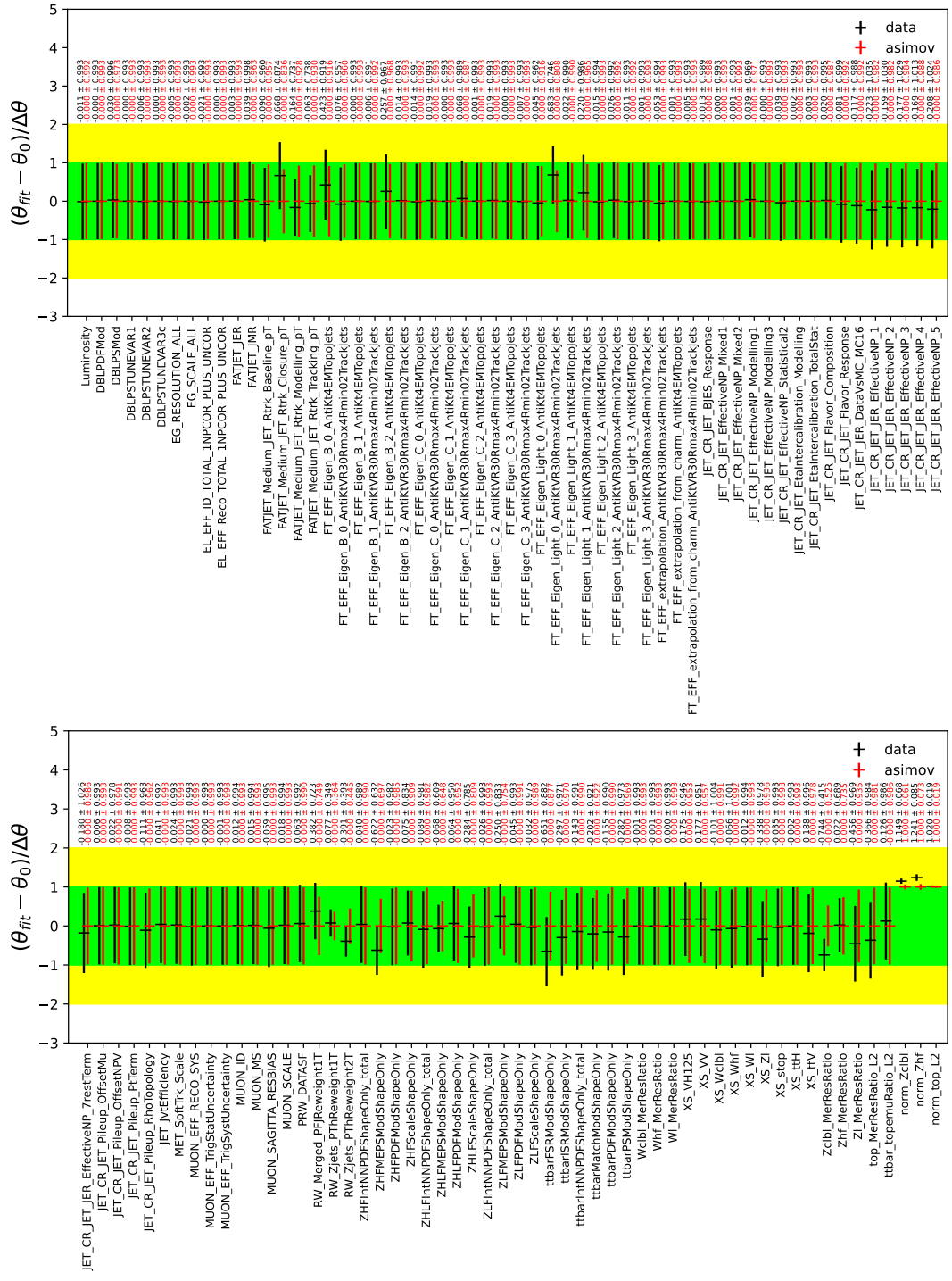


Figure 6.25: Pull plots for the 2-lepton ggA/Z' fits to the data (black) and the Asimov dataset (red).

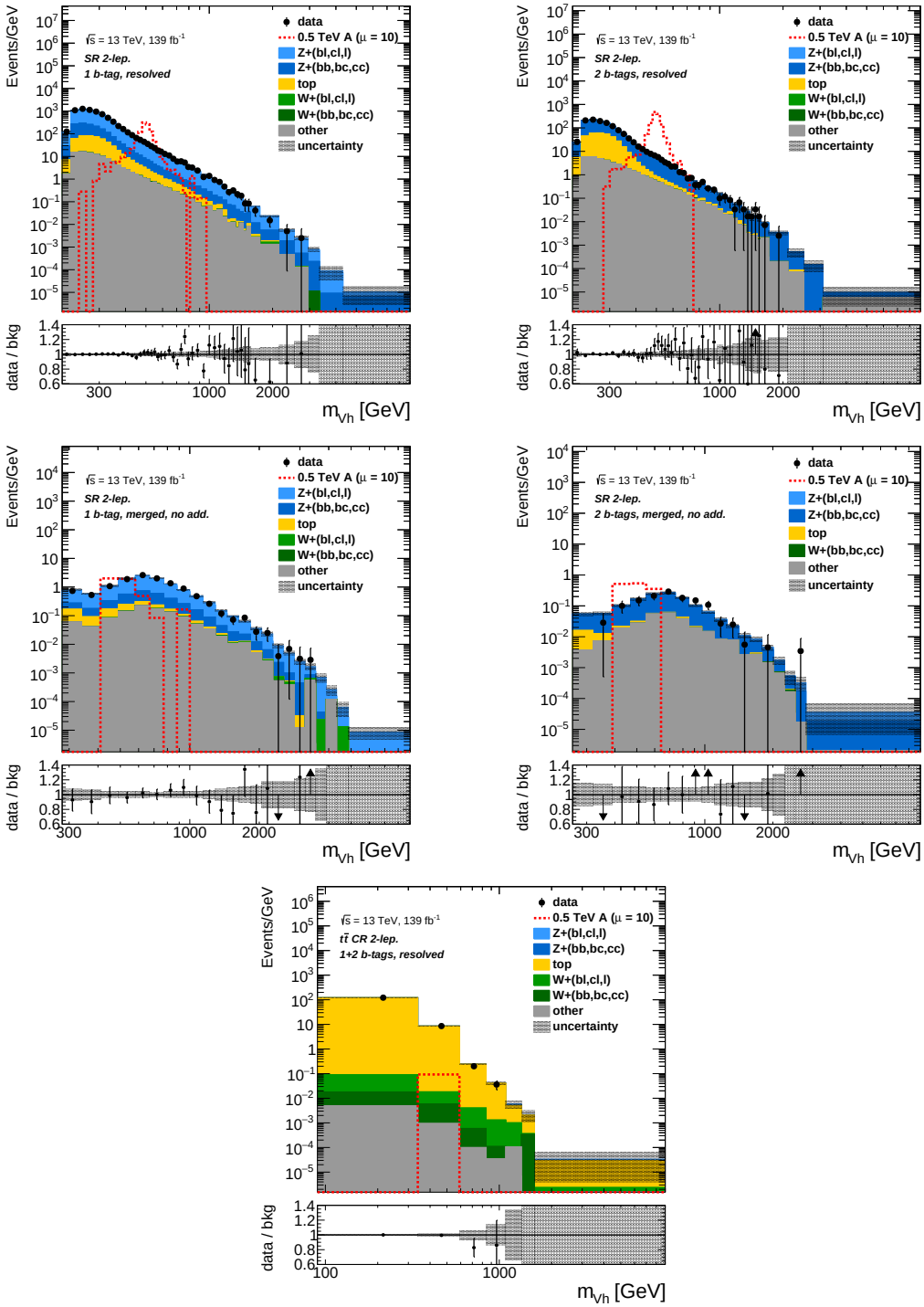


Figure 6.26: Post-fit distributions of the m_{VH} in the 1 b-tagged resolved signal region (top left), 2 b-tagged resolved signal region (top right), 1 b-tagged merged signal region with no additional b-tagged track jets (middle left), 2 b-tagged merged signal region with no additional b-tagged track jets (middle right) and 1+2 b-tagged resolved top control region (bottom) in the ggA/Z' fit.

The bbA Fit

The bbA signals are extracted by fitting the simulated m_{VH} distributions to data. The parameters of interest are the signal strengths of the bbA productions.

The bbA fit uses the 1, 2 and 3+ b-tagged events in the resolved signal regions, the 1, 2 and 3+ b-tagged resolved top control regions, the merged signal regions with no additional b-tagged track jets not associated with the large-R jet and the merged signal regions with at least one additional b-tagged track jet not associated with the large-R jet. The 1 b-tagged and 2 b-tagged top control regions are merged to reduce the sensitivity to b-tagging-related systematic uncertainties. The 1 b-tagged and 2 b-tagged merged signal regions with more than one additional b-tagged track jet not associated with the large-R jet are also combined.

Nuisance Parameter	Description	Regions/Backgrounds	Value	Effect
`norm_Zhf`	Z+(bb,bc,cc) normalization	Z+(bb,bc,cc) in all regions except 3+ b-tagged	Float	normalization
`norm_Zclbl`	Z+(bl,cl) normalization	Z+(cl,bl) in all regions except 3+ b-tagged	Float	normalization
`norm_top`	top normalization	$t\bar{t}+LF$, single top, ttV, tth in all regions	Float	normalization
`norm_ttbarHF`	top normalization	$t\bar{t}+HF$ in all regions	Float	normalization
`norm_Zjets_T3P`	Z+jets normalization	Z+jets in 3+ b-tagged resolved region	Float	normalization
`Zhf_MerResRatio_0AddTag`	merged / resolved ratio	Z+(bb,bc,cc) in 0 add b-tagged merged regions	19%	normalization
`Zhf_MerResRatio_1AddTag`	merged / resolved ratio	Z+(bb,bc,cc) in 1 add b-tagged merged regions	19%	normalization
`Zclbl_MerResRatio_0AddTag`	merged / resolved ratio	Z+(cl, bl) in 0 add b-tagged merged regions	28%	normalization
`Zclbl_MerResRatio_1AddTag`	merged / resolved ratio	Z+(cl, bl) in 1 add b-tagged merged regions	28%	normalization
`Zl_MerResRatio_0AddTag`	merged / resolved ratio	Z+(l) in 0 add b-tagged merged regions	28%	normalization
`Zl_MerResRatio_1AddTag`	merged / resolved ratio	Z+(l) in 1 add b-tagged merged regions	28%	normalization
`top_MerResRatio_L2`	merged / resolved ratio	$t\bar{t}+LF$, single top, ttV, tth in all merged regions	18%	normalization
`ttbarHF_MerResRatio`	merged / resolved ratio	$t\bar{t}+HF$, single top, ttV, tth in all merged regions	18%	normalization
`ttbar_topemuRatio_L2`	SR / topcr ratio	$t\bar{t}+LF$, single top, ttV, tth in all resolved signal regions	2.4%	normalization

Table 6.10: Summary of the normalization nuisance parameters in the bbA fit related to the Z+jets, W+jets and top backgrounds. The forth column shows the values of the prior uncertainties.

A list of nuisance parameters which control the normalization of background components is shown in Table 6.10. In the 1 and 2 b-tagged resolved regions and all merged regions, the normalization factors of the Z+(bb,bc,cc), Z+(bl,cl) , $t\bar{t}+HF$ and $t\bar{t}+LF$ backgrounds are allowed to float freely while the normalization factors of the Z+l, diboson and SM Higgs backgrounds are constrained. The normalization of 3+ b-tagged Z+jets backgrounds are decorrelated from the 1 or 2 b-tagged backgrounds, which means a different set of nuisance parameters which control the normalization factor and the modelling shape uncertainty are assigned to the 3+ b-tagged Z+(bb,bc,cc,bl,cl,l). The normalization factors for the 3+ b-tagged Z+(bb,bc,cc,bl,cl,l) backgrounds are allowed to

float freely in the fit. Other constrained normalization nuisance parameters are luminosity and the resolved/merged ratios of the top, Z+(bb,bc,cc) and Z+(bl,cl) backgrounds. The resolved/merged ratios of the Z+(bb,bc,cc) and Z+(bl,cl) backgrounds with no additional b-tagged track jets not associated with the large-R jets are decorrelated from the one with additional b-tagged track jets. The nuisance parameters which control the merged/resolved ratios are applied to the merged regime.

The bbA fit uses a similar binning strategy as the ggA and Z' fits. However, a coarser binning is used in the 1 b-tagged resolved region because the region is mainly used to model the backgrounds, and it has a smaller contribution to the signal discovery significance. The statistical uncertainty of the simulation in each bin is required to be less than 20% of the total weight of the simulated background distributions.

The fit model is first tested using events in the control region before applying it in the signal region. The control region fit is performed by replacing signal region distributions with the distributions in the m_{BB} sideband control region, and the post-fit distributions are shown in Figure 6.27 and Figure 6.28. The plots show overall good agreement between data and simulation.

The pulls of nuisance parameters from the Asimov fit and the fit using data are shown in Figure 6.29. As can be seen from the plots, most nuisance parameters behave as expected in the fit. It should be noticed that the nuisance parameter for the p_T^{BB} reweighting in the 1 b-tagged region (`RW_Zjets_PThReweight1T`) is pulled toward the positive direction, which means the fit tries to reverse the slope correction. The pull is caused by its correlation with other nuisance parameters which control the modelling of Z+jets background. Figure 6.30 shows that the nuisance parameter is strongly anti-correlated with a nuisance parameter for a Z+jets modelling uncertainty (`ZHLFMEPSModShapeOnly`). This correlation is caused by overlap between the Z+jets modelling uncertainties and the reweighting uncertainties mentioned in Section 6.5.2. Tests have shown that `RW_Zjets_PThReweight1T` is not pulled when it is the only nuisance parameter which affects the shape of the distributions in the fit. It is pulled if `ZHLFMEPSModShapeOnly` is also included in the fit. Other pulled nuisance parameters related to flavour tagging and jets are not of concern because their effect on signal

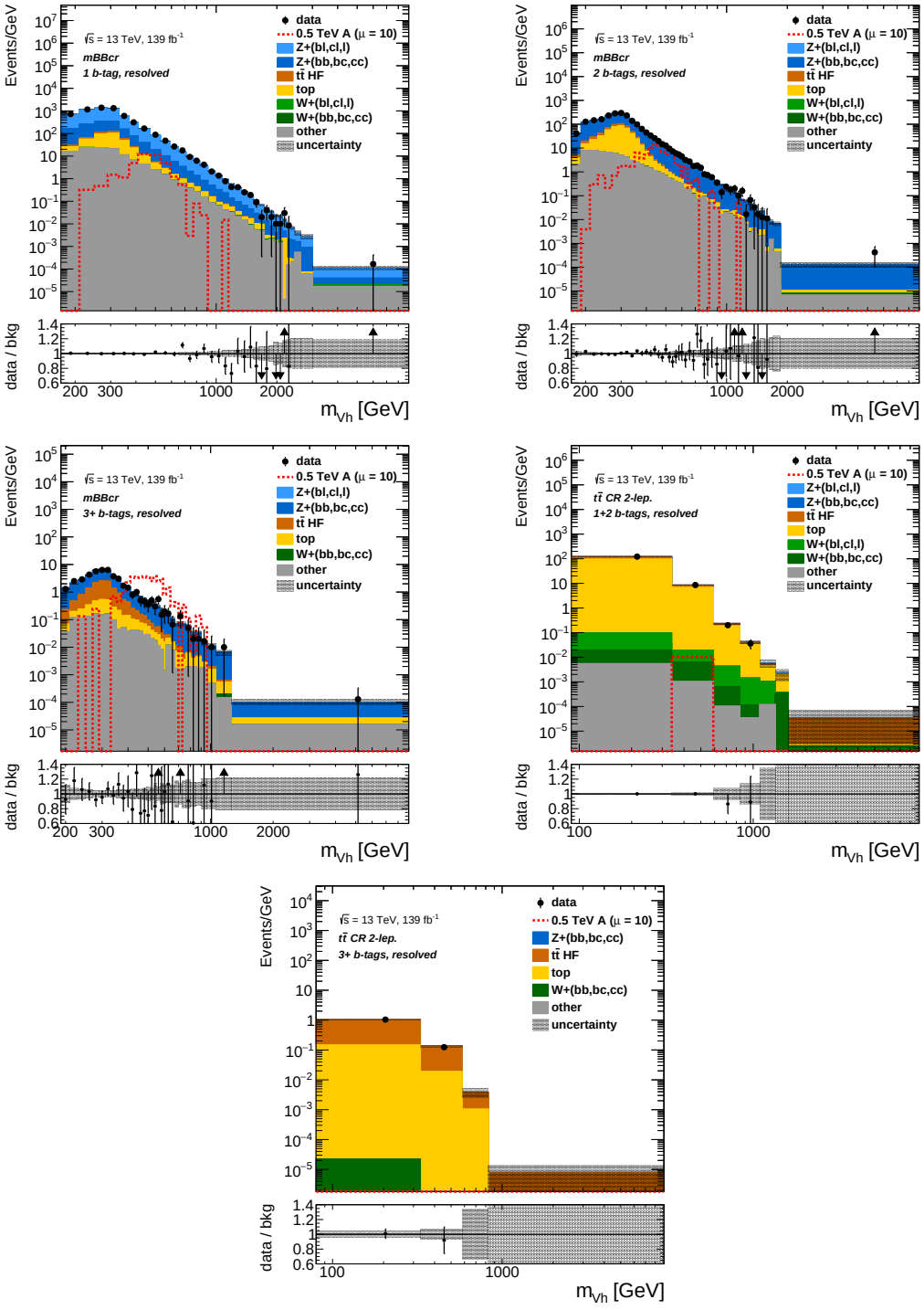


Figure 6.27: Post-fit distributions of the m_{VH} in the 1 b-tagged (top left), 2 b-tagged (top right) and 3+ b-tagged (middle left) resolved m_{bb} sideband control regions and 1+2 b-tagged (middle right) and 3+ b-tagged (bottom) resolved top control regions in the bbA fit.

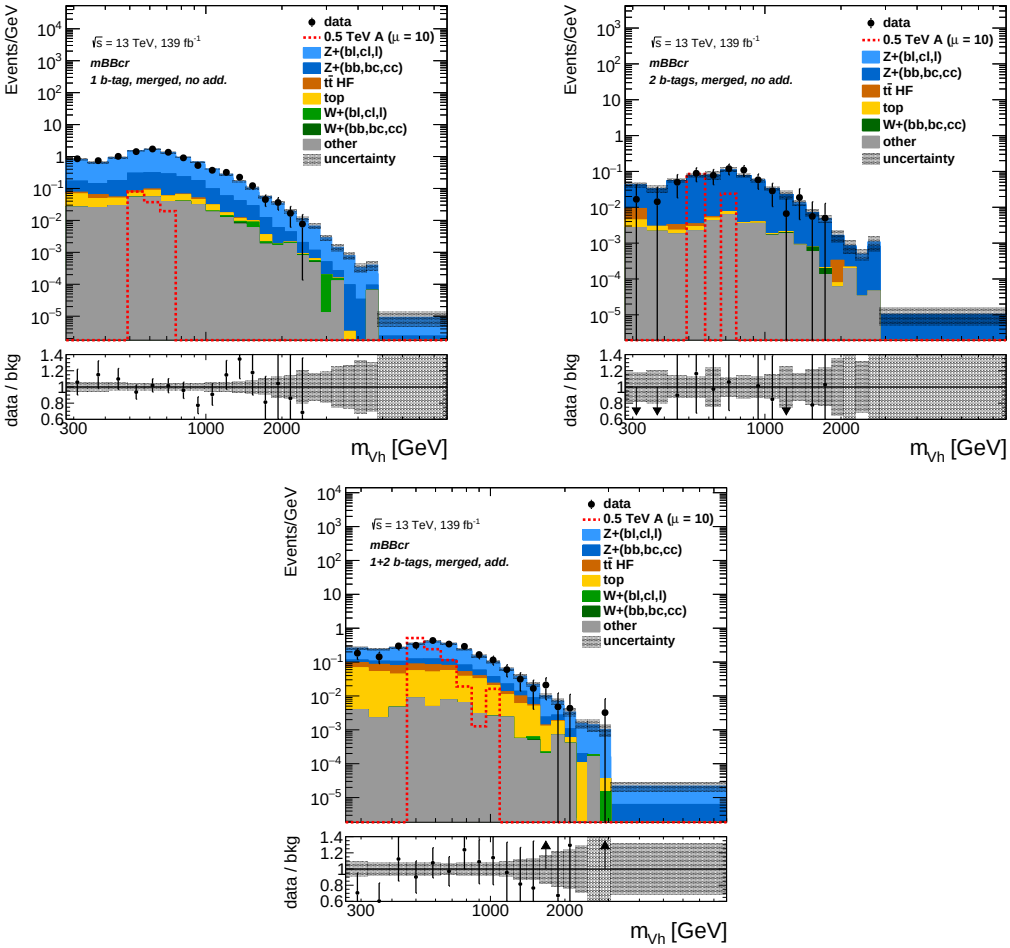


Figure 6.28: Post-fit distributions of the m_{VH} in the 1 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (top left), 2 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (top right), 1+2 b-tagged merged m_{bb} sideband control region with additional b-tagged track jets (bottom) in the bbA fit.

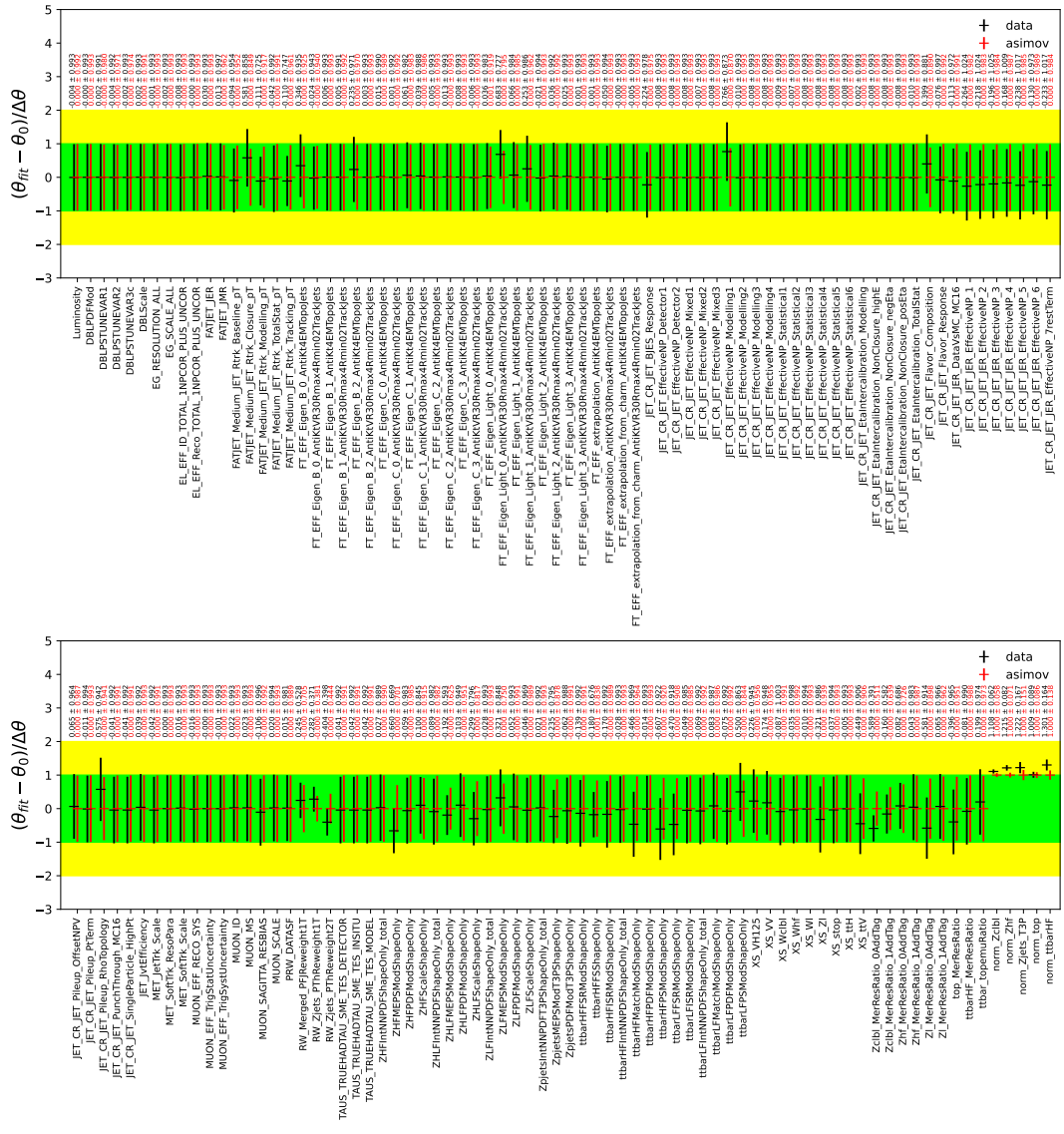


Figure 6.29: Pull plots for the 2-lepton bbA fits to the data (black) and the Asimov dataset (red).

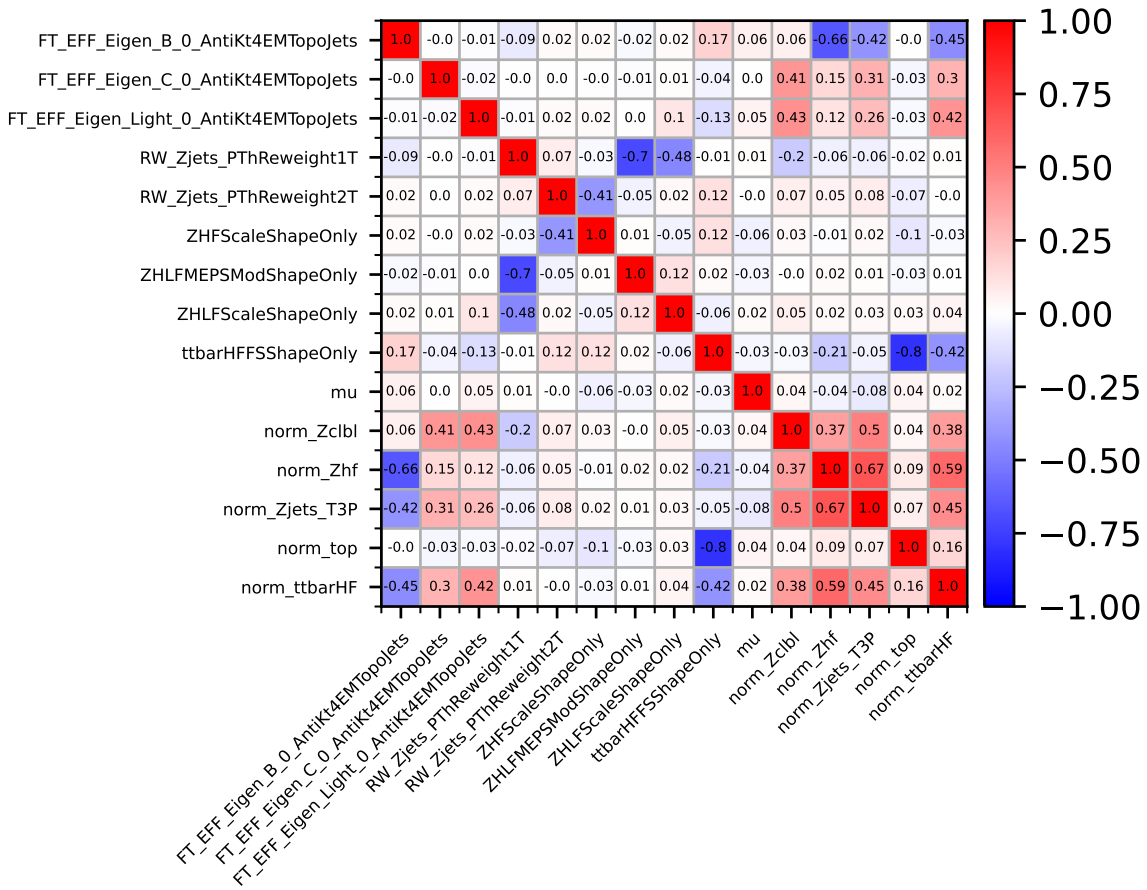


Figure 6.30: The correlation matrix of the bbA fit in the 2-lepton channel. The plot only shows rows or columns with elements whose absolute value is larger than 0.3.

strength is trivial, and the pulls have no impact on the result.

Figure 6.31 and Figure 6.32 show the post-fit distributions in the signal regions and the top control regions. The plots show that the data in general agrees with the Standard-Model prediction.

6.6 Results

The result of the search is compatible with the Standard-Model prediction, and upper limits on the production cross section of signals are calculated. The expected and observed limits on the production cross sections of the ggA, bbA and Z' signals and the p-values of the searches are shown in Figure 6.33. The ggA search has probed a mass range of 220 GeV to 2 TeV. The results show good agreement between the observed limits and the expected limits, and upper limits at the 95 % confidence level (CL) has been placed on $\sigma(ggA) \times \mathrm{BR}(A \to Zh) \times \mathrm{BR}(h \to b\bar{b})$ between 1×10^{-3} pb and 4×10^{-1} pb. A mild excess of about 2σ has been observed at 500 GeV. The bbA search has also probed a mass range of 220 GeV to 2 TeV with no excesses above 2σ observed. Upper limits at the 95 % CL has been placed on $\sigma(bbA) \times \mathrm{BR}(A \to Zh) \times \mathrm{BR}(h \to b\bar{b})$ between 1×10^{-3} pb to 2×10^{-1} pb. The search for the Z' boson has probed a mass range of 300 GeV to 5 TeV with a mild excess above 2σ observed at 500 GeV. The search placed upper limits at the 95 % CL on $\sigma \times \mathrm{BR}(Z' \to Zh) \times \mathrm{BR}(h \to b\bar{b}, b\bar{c})$ between 1×10^{-3} pb and 4×10^{-1} pb. The search in the $l^-l^+b\bar{b}$ final state has excluded the HVT Model A up to 2400 GeV and the HVT model B up to 2600 GeV [1].

[1] The two benchmark models were discussed in Section 2.3.3.

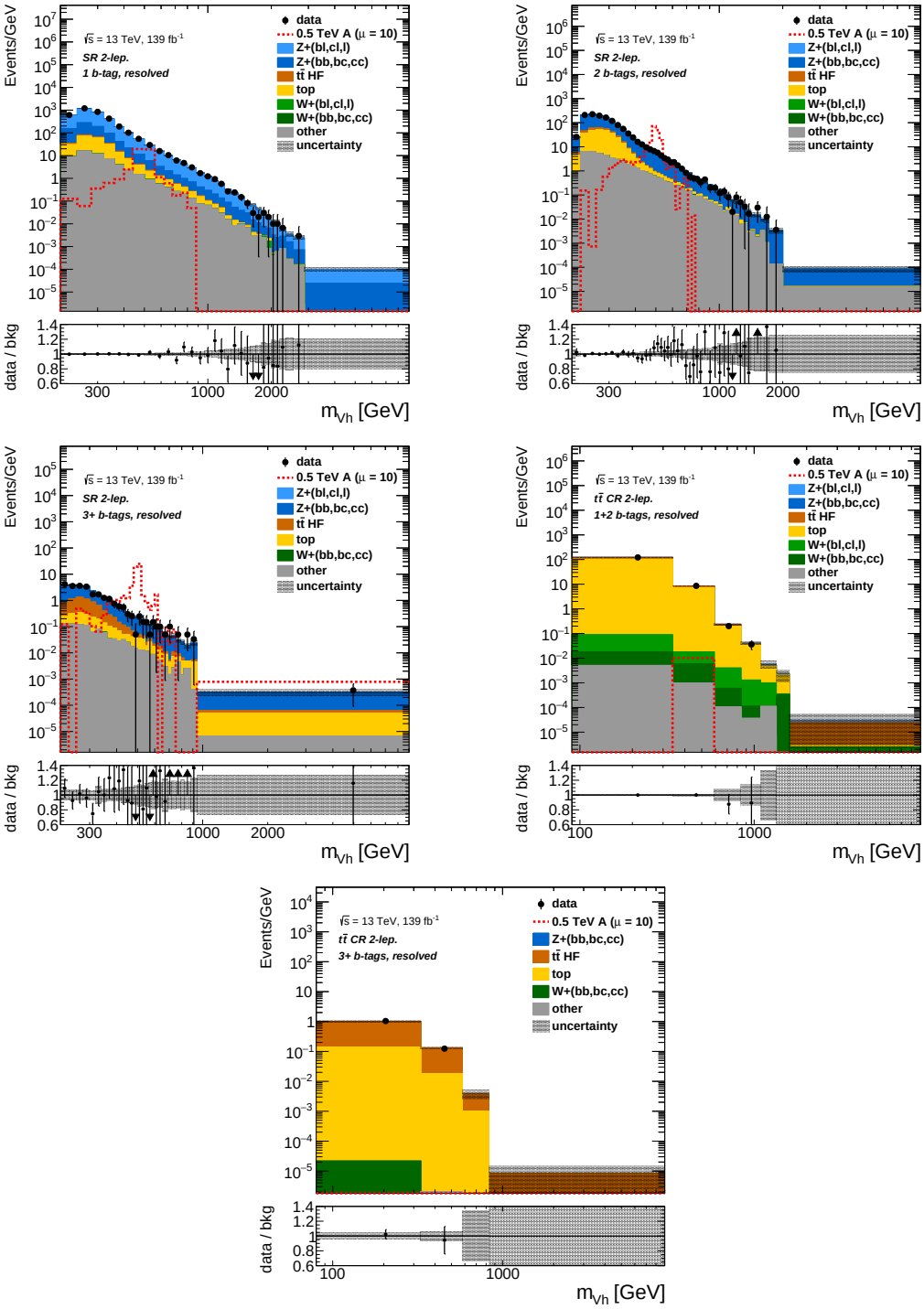


Figure 6.31: Post-fit distributions of the m_{VH} in the 1 b-tagged (top left), 2 b-tagged (top right) and 3+ b-tagged (middle left) resolved signal regions and 1+2 b-tagged (middle right) and 3+ b-tagged (bottom) resolved top control region in the bbA fit.

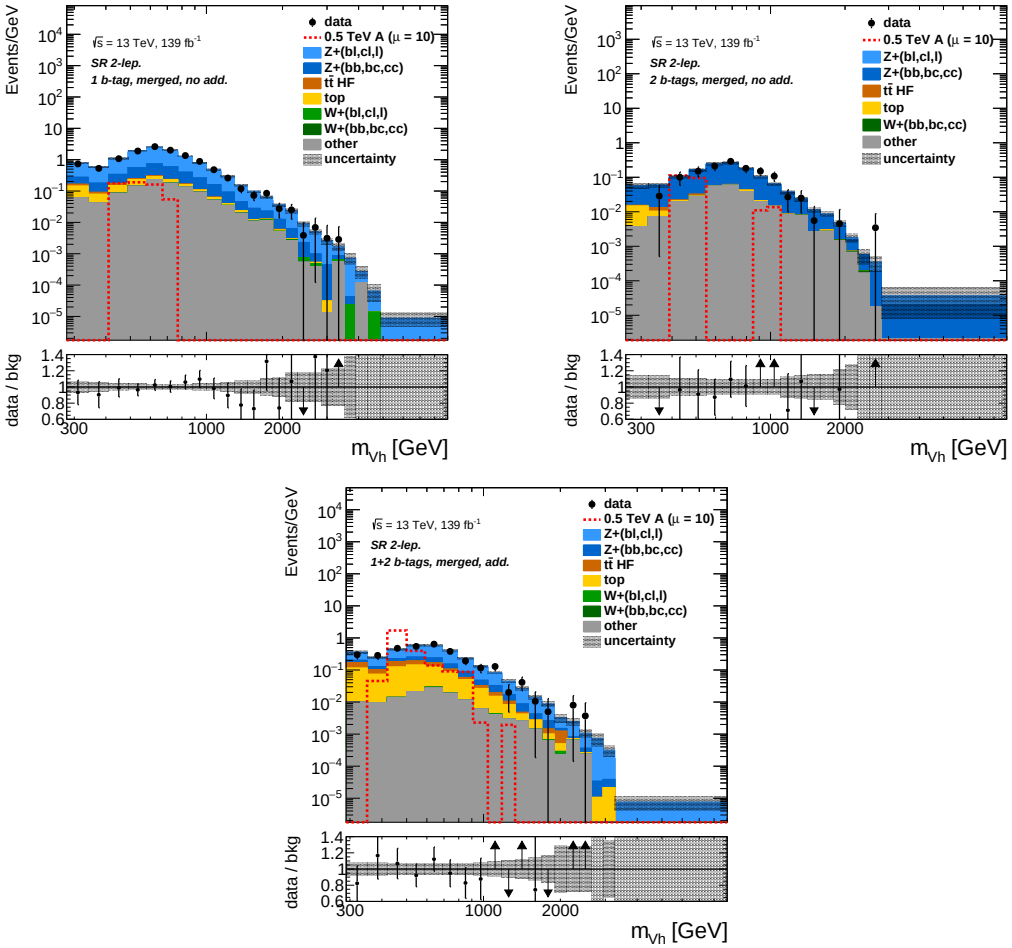


Figure 6.32: Post-fit distributions of the m_{VH} in the 1 b-tagged merged signal region with no additional b-tagged track jets (top left), 2 b-tagged merged signal region with no additional b-tagged track jets (top right), 1+2 b-tagged merged signal region with additional b-tagged track jets (bottom) in the bbA fit. The discontinuity in the signal distributions in the top right plot and the bottom plot is due to low statistics.

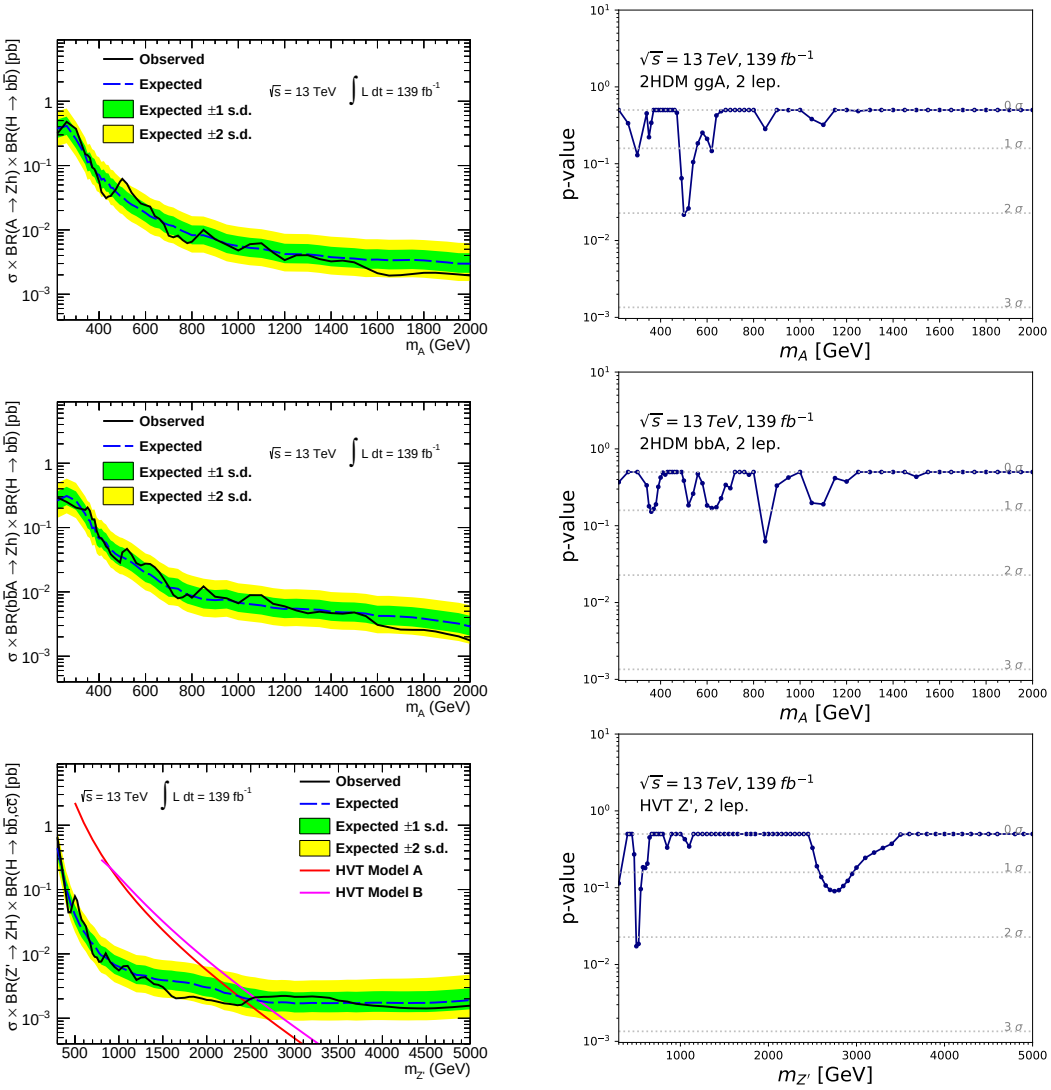


Figure 6.33: The expected and observed upper limits at the 95% CL on the production cross section of the ggA signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}$ (upper left), the production cross section of the bbA signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}$ (middle left) and the production cross section of the Z' signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}, c\bar{c}$ (bottom left). The p-values for the ggA (upper left), bbA (middle left) and Z' (bottom left).

Chapter 7

Combination of the $llbb$ and $\nu\nu bb$ Searches and 2HDM Interpretation

The searches in the $l^-l^+b\bar{b}$ final state are combined with the searches in the $\nu\bar{\nu}b\bar{b}$ final state where the Z boson decays into two neutrinos [1]. The combined results are then interpreted as exclusion limits in the 2HDM parameter space. In the $\nu\bar{\nu}b\bar{b}$ searches, the heavy resonance cannot be fully reconstructed due to the presence of two neutrinos in the final state. As a result, the final discriminator of the search is the transverse mass of the Zh system

$$m_{T,VH} = \sqrt{(E_T^{BB} + E_T^{\text{miss}})^2 - (\vec{p}_T^{BB} + \vec{E}_T^{\text{miss}})^2}, \tag{7.1}$$

where E_T^{BB} is the transverse energy of the h boson candidate, E_T^{miss} is the missing transverse energy, $\vec{p}_T^{BB}$ is the transverse momentum of the h boson candidate and $\vec{E}_T^{\text{miss}}$ is the missing transverse momentum [1].

The chapter mainly focuses on the author's contribution which is the pseudo-experiment limits, the signal smearing and the 2HDM interpretation. The 0-lepton channel analysis and the combination strategy of the fit are beyond the scope of this thesis [1]. However, they will still be briefly discussed in this chapter for completeness.

Variable	Resolved	Merged
Number of jets	≥2 signal jets	≥1 large-R jet ≥ 2 track jets (matched to leading $R = 1.0$ jet)
Leading jet p_T [GeV]	> 45	> 250
m_{bb} [GeV]	110–140	75–145
E_T^{miss} [GeV]	> 150	> 200
$\Delta\phi_{jj}$	$< 7\pi/9$	–
E_T^{miss} [GeV]	> 150	> 200
S_{T} [GeV]	> 150 (120^*)	–
$E_{\text{T,trk}}^{\text{miss}}$ [GeV]	> 60	
$\Delta\phi(\vec{E}_{\text{T}}^{\text{miss}}, \vec{E}_{\text{T,trk}}^{\text{miss}})$	$< \pi/2$	
$\Delta\phi(\vec{E}_{\text{T}}^{\text{miss}}, h)$	$> 2\pi/3$	
$\min\left[\Delta\phi(\vec{E}_{\text{T}}^{\text{miss}}, \text{small-}R\text{ jet})\right]$	$> \pi/9$ (2 or 3 jets), $> \pi/6$ (≥ 4 jets)	
E_T^{miss} significance $\mathcal{S}$	$\begin{cases} > 9 & \text{if } m_{Vh} < 240\,\text{GeV}, \\ > 6.6 + 0.01 \cdot m_{Vh} & \text{if } 240\,\text{GeV} \leq m_{Vh} < 700\,\text{GeV}, \\ > 13.6 & \text{if } m_{Vh} > 700\,\text{GeV}, \end{cases}$	

Table 7.1: The event selection for the $A/Z' \to Zh \to \nu\bar{\nu}b\bar{b}$ search. * applies when there are only two signal jets. The table is taken from [1].

7.1 Searches in the $\nu\nu bb$ Final State

The search for a heavy resonance decaying into a Z boson and a h boson in the $\nu\bar{\nu}b\bar{b}$ final state is described in [2] and the event selection is listed in Table 7.1. The 0-lepton channel uses the same object reconstruction methods as the 2-lepton channel. It also splits events into the resolved and the merged categories depending on the jets used to reconstruct the Higgs candidate. The 0-lepton channel uses the same cuts on the jets' p_T as the 2-lepton channel. It uses events with large missing transverse momentum due to the presence of two neutrinos in the final state, so the 0-lepton channel requires $E_T^{\text{miss}} > 150$ GeV for the resolved category and $E_T^{\text{miss}} > 200$ GeV for the merged category. To avoid double-counting of events with the 1-lepton channel[1] and the 2-lepton channel, the 0-lepton channel only uses events with no reconstructed leptons. To reduce the mismodelling on trigger efficiencies, the p_T sum of the three signal jets leading in p_T (S_T) is required to be greater than 150 GeV. The requirement is relaxed to 120 GeV if there are only two signal jets. The multijet background is mainly caused by the mismeasurement of jet energy. It can be reduced by a cut on the angular separation of

[1]The 1-lepton channel is used for the search for the W' boson. It is described in [137].

final-state products. The azimuthal angle between the two jets used for the h boson reconstruction $\Delta\phi_{jj}$ is required to be less than $7\pi/9$ in the resolved category. The cut is not needed in the merged category because the two track jets are required to be associated with the large-R jets. The azimuthal angle between the h boson candidate and the missing transverse momentum $\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, h)$ is required to be greater than $> 2\pi/3$. The minimal azimuthal angle between the missing transverse momentum and small-R jets $\min\left[\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \text{small-}R\text{ jet})\right]$ is required to be greater than $\pi/9$ when there are two or three small-R jets. It is required to be greater than $\pi/6$ when there are four or more small-R jets. The missing transverse momentum of tracks $E_{\mathrm{T,trk}}^{\mathrm{miss}}$ is less affected by pileup. The analysis selects events with $E_{\mathrm{T,trk}}^{\mathrm{miss}} > 60$ GeV, and the azimuthal angle between $E_{\mathrm{T,trk}}^{\mathrm{miss}}$ and the missing transverse momentum $\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{E}_{\mathrm{T,trk}}^{\mathrm{miss}})$ is required to be less than $\pi/2$. A m_{VH}-dependent cut on the object-based significance of the missing transverse momentum $\mathcal{S}$ [138] is applied to further suppress backgrounds. The 0-lepton channel selects resolved events with 110 GeV $< m_{bb} <$ 140 GeV. The cut on the m_{bb} is tighter than that of the 2-lepton channel because the 0-lepton channel has a higher background.

The 0-lepton channel analysis also assigns an event to the resolved category if it passes both resolved and merged event selections. Events failing the m_{bb} selection while still being within the range 50 GeV $< m_{bb} <$ 200 GeV are assigned to the m_{bb} sideband control region. The signal and control regions in the 0-lepton channel are mainly dominated by the Z+jets, W+jets and $t\bar{t}$ backgrounds.

The simulated events are reweighted to improve the agreement between the data and the simulation. The reweighting functions based on the transverse momentum of the h boson candidate and the missing transverse momentum are derived in the m_{bb} sideband control region and applied in the signal regions [1].

7.2 The Combined Fits and Results

In the 0-lepton channel, the bbA fit uses events in the 1, 2 and 3+ b-tagged resolved signal regions, the 1, 2 and 3+ b-tagged resolved top control regions, the merged signal

regions with no additional b-tagged track jets not associated with the large-R jet and the merged signal regions with at least one additional b-tagged track jet not associated with the large-R jet. The combined fit uses both the 0-lepton channel events and the 2-lepton channel events to improve the sensitivity of the search. The combined bbA fit uses the same nuisance parameters as the 2-lepton channel bbA fit. Most nuisance parameters are correlated between the two channels except for the normalization factors of the top background (`norm_top`) which are left uncorrelated due to the different regions of phase-space being probed in the two channels. Due to the large W+jets background in the 0-lepton channel, additional constrained normalization factors are applied to the W+jets background in the resolved regions to address the uncertainties on the acceptance ratios between resolved and merged regions. Additional constrained normalization factors are applied to the $t\bar{t}$ and Z+jets background to address the uncertainties on the acceptance ratios between the 0-lepton channel and the 2-lepton channel [1].

The 0-lepton ggA/Z' fit uses the 1, 2 b-tagged signal regions and m_{bb} sideband control regions in both the resolved and merged categories. The m_{bb} sideband control regions are also used in the fit because it helps to constrain the normalization factors of the backgrounds and thus improve the sensitivity of the search. The 0-lepton channel is combined with the 2-lepton channel using the same nuisance parameters described in Section 6.5.3. The normalization factor of the top backgrounds (`norm_top_L2`) is decoupled between the two channels because of the normalization difference. Additional constrained normalization factors are applied to the $t\bar{t}$, W+jets and Z+jets backgrounds to address the uncertainties on the acceptance ratios between the two channels [1].

The expected and observed upper limits of the combined fits at the 95% CL on the production cross sections of ggA, bbA and Z' signals time the branching ratio to Zh are shown in Figure 7.1. The results are compatible with the Standard-Model prediction. The largest excess is observed for the ggA and Z' signal at 500 GeV corresponding to a significance of 2.1σ. The plots also show the combined limits are dominated by the 2-lepton channel in the low-mass range below 1 TeV. The 0-lepton channel is important above 1 TeV because of the poor lepton-reconstruct efficiency in the high mass range

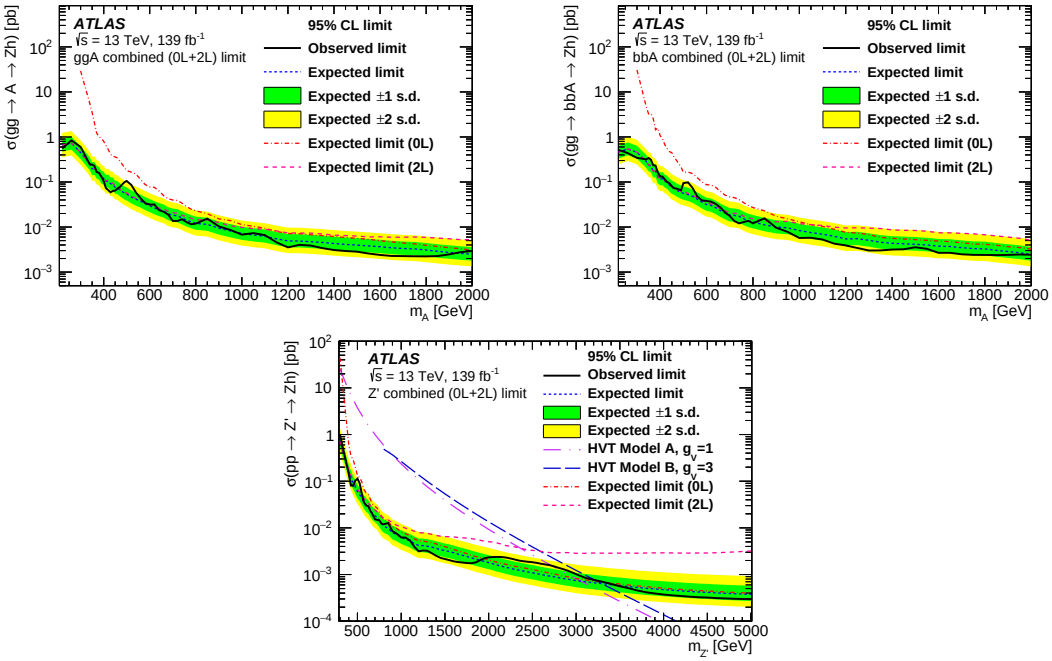


Figure 7.1: The expected and observed upper limits at the 95% CL on the production cross section of the ggA signal times the branching ratio to Zh (upper left), the production cross section of the bbA signal times the branching ratio to Zh (upper right) and the production cross section of the Z' signal times the branching ratio to Zh (bottom). The plots are taken from [1].

where the two leptons are collimated.

7.3 Pseudo-Experiment Limits

The limits shown in Figure 7.1 are calculated using the asymptotic approximation, where the conditional probability density function of the test statistic t_μ in Equation (6.12) is estimated by asymptotic formulas [131]. The approach is valid in the large-sample limit, but the asymptotic formulas begin to fail in the high mass range due to the low number of high-mass events. Alternatively, the t_μ distributions can also be estimated by pseudo-experiments. The pseudo-experiment provides a more accurate estimation than the asymptotic formulas, but it is also compute-intensive. It is not practical to calculate the pseudo-experiment limits for all signal hypotheses due to the limitation of computing resources. The Z' search covers the mass range up to 5 TeV, and pseudo-experiment limits of ten selected mass hypotheses are calculated to test the validity of the asymptotic approximation.

For each signal mass hypothesis, the probability density function of the null hypothesis ($f(t_\mu|0)$) and that of the signal (with strength μ) hypothesis ($f(t_\mu|\mu)$) are constructed from toy samples, while the value of the observed test statistic $t_{\mu,obs}$ is determined from data. The p-value at a given signal strength is

$$p(\mu) = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu)dt_\mu. \tag{7.2}$$

The p-values are calculated at multiple μ values, and the relationship between the p-value and μ is estimated by an error-function fit. The upper limit on μ at the 95% CL is the μ value when $p = 5\%$. The median limit and the $\pm1\sigma$ and $\pm2\sigma$ error bands are determined by testing the signal hypothesis against the hypothesis where t_μ is at the median, $\pm1\sigma$ and $\pm2\sigma$ of the null hypothesis distribution $f(t_\mu|0)$.

In the Z' search, the t_μ distributions are estimated by 10000 toy samples for each signal mass hypothesis. The upper limits at the 95% CL for signals of 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500 and 5000 GeV are shown in Figure 7.2. A comparison between the asymptotic limits and the pseudo-experiment limits is also shown in Figure 7.2. The figure shows that the asymptotic approximation is valid below 2000 GeV, and it begins to fail in the higher mass range. The comparison shows a 20% discrepancy in the observed limit between the asymptotic approximation and the pseudo-experiment at 5000 GeV. In the published results [1], the asymptotic formulas are used for the limit calculation because a 20% discrepancy is considered to be acceptable. However, a comparison between the asymptotic limits and the pseudo-experiment limits is also shown for the ten selected mass hypotheses to demonstrate the impact of this assumption.

7.4 Large-Width Signal Modelling

The 2HDM signal samples used in the analysis are generated under the narrow width approximation (NWA). The width of the A boson is sensitive to the parameters of the 2HDM. Large-width signal distributions with widths Γ_A ranging from 1% to 20% are estimated by smearing the distributions of the $m_{T,Vh}$ in the 0-lepton channel and the

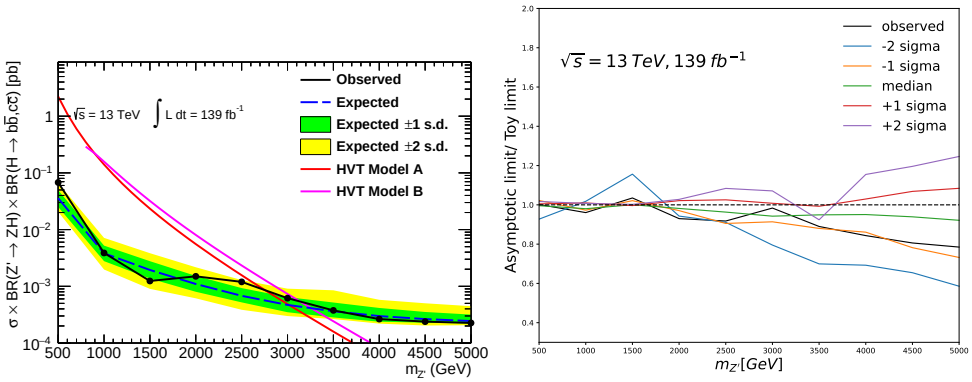


Figure 7.2: The expected and observed upper limits at the 95% CL on the production cross section of Z' signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}, b\bar{c}$ calculated by pseudo-experiments (left) and the ratio between the asymptotic limits and the pseudo-experiment limits (right).

distributions of the m_{Vh} in the 2-lepton channel.

The LO $ggA \to l^-l^+b\bar{b}/\nu\bar{\nu}bb$ samples with signal widths ranging from $\Gamma_A/m_A =$ 0% to 20% with a step size of 1% are generated using MadGraph5 2.7.3 and Pythia 8.244 [72, 76]. Cuts are applied to the truth-level leptons, jets and missing transverse momentum objects to mimic the event selection at the reconstruction level. The priority resolved strategy is applied to split the truth-level events into the resolved and merged categories.

The resonance peak follows the Breit-Wigner distribution. However, the reconstructed A boson mass distribution is distorted because of the event selection. Instead of using the Breit-Wigner function, the NWA distributions are smeared using a modified Breit-Wigner function [139] defined as

$$f(x) = \frac{e^{\frac{-(\ln(x-x_0)-\ln(m)^2)}{2\times\ln(s)^2}}}{\sqrt{2\pi}(x-x_0)\ln(s)} \times \frac{1}{x^2+0.25w^2}, \tag{7.3}$$

where w is the signal width , x_0, s and m are free parameters. The function is a product of a Breit-Wigner function and a log-normal function. The log-normal function makes the function asymmetric to improve the modelling of the reconstructed distributions. The three free parameters are determined by fitting the smeared NWA distributions to the generated large-width distributions.

The NWA distributions are smeared by the convolution between the NWA distributions and the modified Breit-Wigner function. The convolution is calculated using the RooFit package numerically [140]. The initial values of the free parameters in the fit are set by scanning the parameters space to avoid suboptimal solutions. The parameters are then optimized by the fit. Fits are performed for each set of initial values in the scan, and the best solutions are selected using the root mean square error. The fit is done for each signal mass of different widths in the 0/2 lepton channels and the resolved/merged regions separately. In the case where the modified Breit-Wigner function smearing gives poor modelling of the large-width distributions, the Breit-Wigner function is used for the smearing. The best function is determined by the root mean square error.

Figure 7.3 and Figure 7.4 compare the simulated large-width distributions with the distributions estimated by the smearing for the 700 GeV ggA signal in the 2-lepton channel. The plots show that the smearing produces the same line shape as the generated distributions.

The signal width also affects the event-selection efficiency, which is estimated using the ratios of total weight between the large-width distributions and the NWA distributions at the truth level. The ratios are shown in Figure 7.5. The plots show that the ratios are close to one in the 2-lepton channel. The signal width significantly affects the event-selection efficiencies of the low-mass signal in the 0-lepton channel because of the low event-selection efficiency of the low-mass signals. With larger signal widths, more events are migrated to the higher mass bins, and the event selection efficiency increases.

The smearing function and the ratios are applied to the reconstructed NWA distributions to estimate the large-width distributions. The statistical uncertainties of the large-width distributions are estimated by varying the nominal NWA distribution by $\pm\sigma$, and then applying the smearing function and the ratio. The halved difference between the two smeared $\pm\sigma$ distributions is taken as statistical uncertainties. Due to the lack of large-width bbA samples, the smearing function and the total weight ratios determined from the ggA sample are applied to the NWA bbA distributions to estimate the large-width bbA distributions.

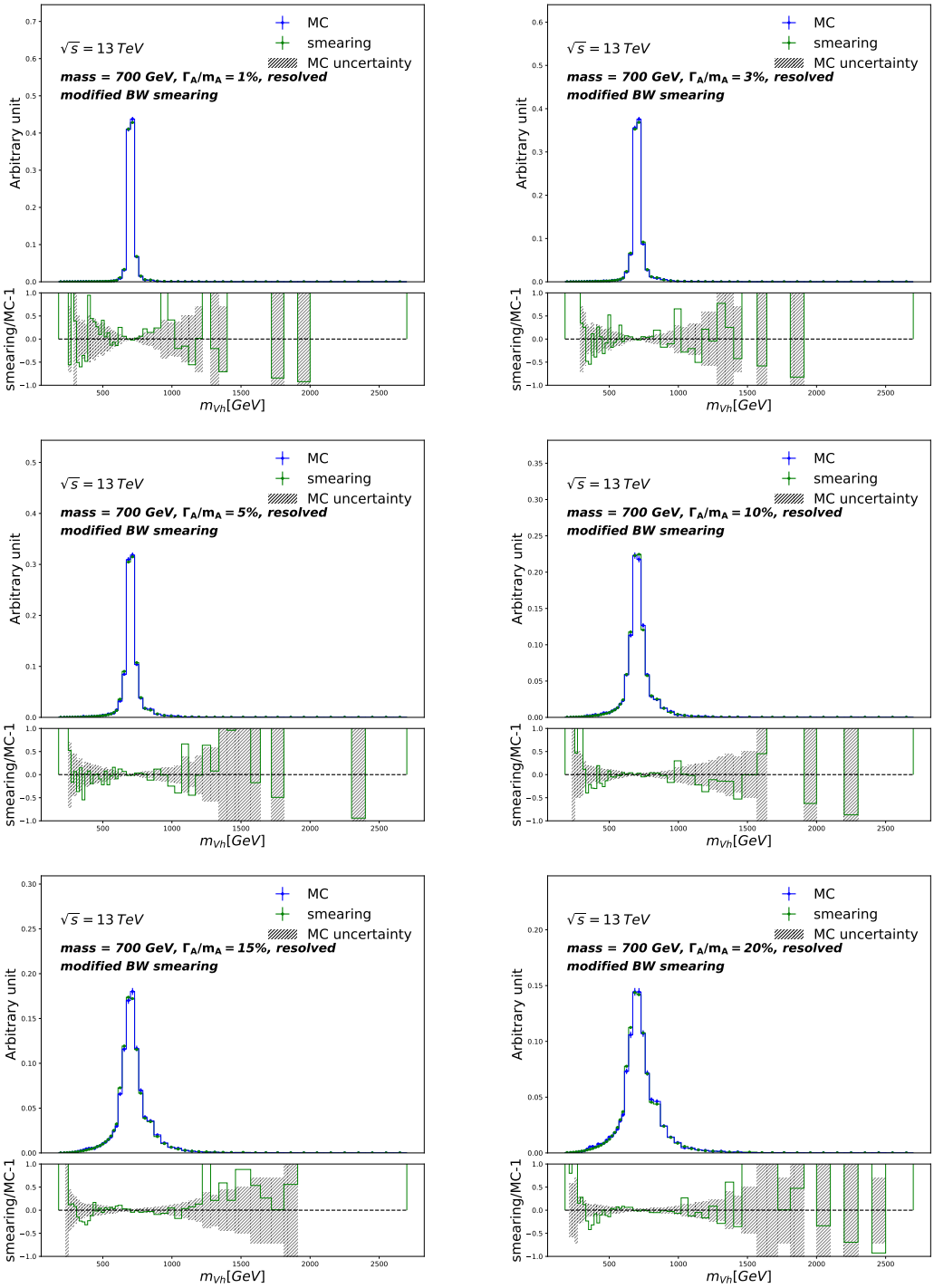


Figure 7.3: The distributions of m_{Vh} for $m_A = 700$ GeV ggA signals in the 2-lepton resolved regions with the widths of $\Gamma_A/m_A = 1\%$ (top left), 3% (top right), 5% (middle left), 10% (middle right), 15% (bottom left) and 20% (bottom right). The blue curve shows the simulated distributions while the green curve show the smeared NWA distribution. The lower panels show the ratios between the smeared NWA distributions and the simulated distributions. The MC uncertainty shown is statistical only.

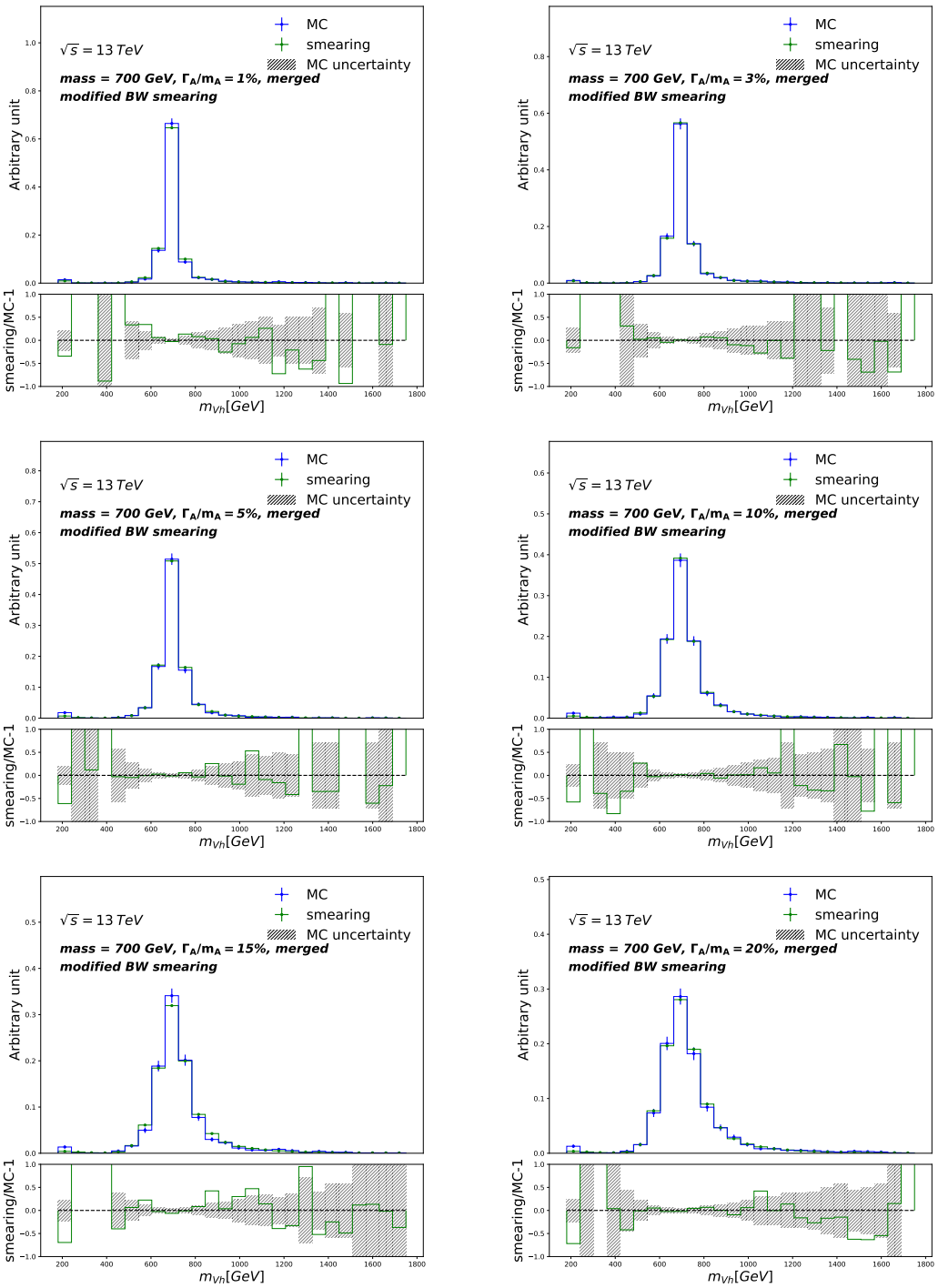


Figure 7.4: The distributions of m_{Vh} for $m_A = 700$ GeV ggA signals in the 2-lepton merged regions with the widths of $\Gamma_A/m_A = 1\%$ (top left), 3% (top right), 5% (middle left), 10% (middle right), 15% (bottom left) and 20% (bottom right). The blue curve shows the simulated distributions while the green curve show the smeared NWA distribution. The lower panels show the ratios between the smeared NWA distributions and the simulated distributions. The MC uncertainty shown is statistical only.

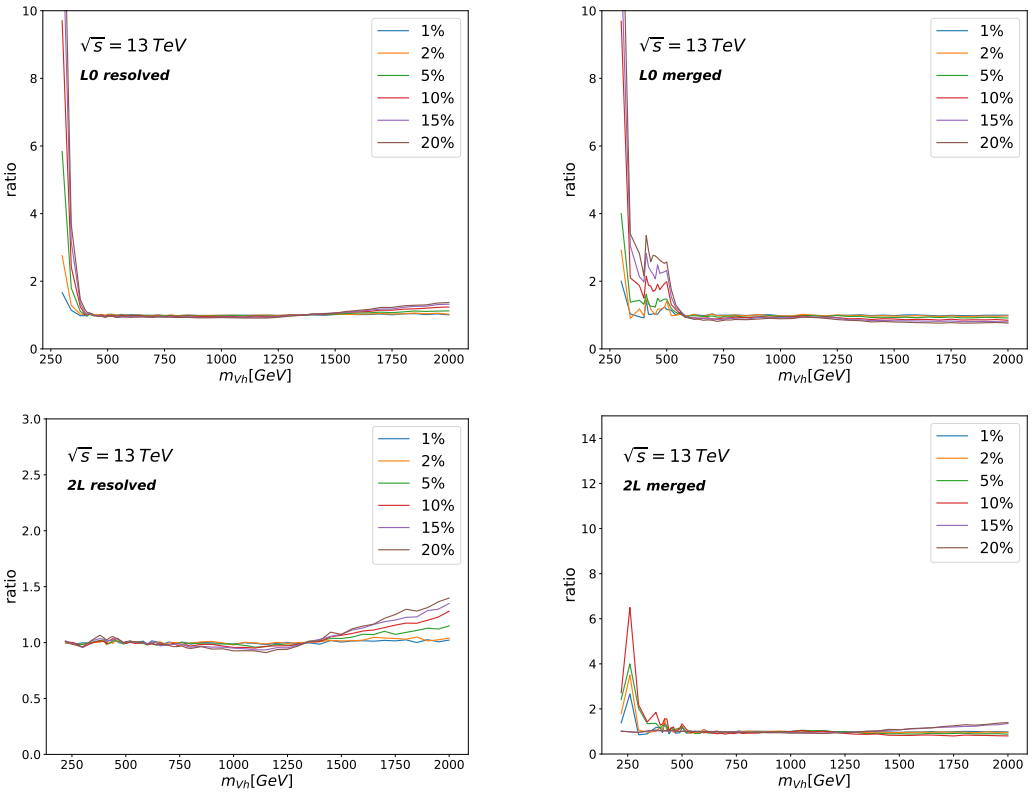


Figure 7.5: The total-weight ratios between the large-width distributions and the NWA distributions in the 0-lepton resolved (top left), 0-lepton merged (top right), 2-lepton resolved (bottom left) and 2-lepton merged (bottom right) regions.

7.5 The 2HDM Interpretation

The 2HDM parameters, discussed in Section 2.3.2, affect the fraction of ggA and bbA production modes and the natural width of the A boson. To obtain the upper limits at the 95% CL on the 2HDM parameters, the fraction of ggA and bbA production modes and the variation of the signal width are taken into account. The upper limits at the 95% CL on the cross section of 2HDM signals are calculated for various A boson natural widths up to $\Gamma_A/m_A = 20\%$. The bbA and the ggA signals are combined with different fractions ranging from 0 to 1, and the bbA fit model is used for the limit calculation. The expected and observed upper limit at the 95% CL as a function of bbA fraction and signal mass is in shown in Figure 7.6. The expected and observed upper limits at the 95% CL on $\sigma(pp \to b\bar{b}A \to Zh)$ and $\sigma(pp \to A \to Zh)$ as a function of signal width and signal mass are shown in Figure 7.7 and Figure 7.8 respectively. In general, better limits are expected as the signal width decreases. For high mass signals, the pure bbA/ggA signal has better limits than the mixed signals. However, the limits get worse as the fraction of bbA signal decreases for signals with mass below 300 GeV.

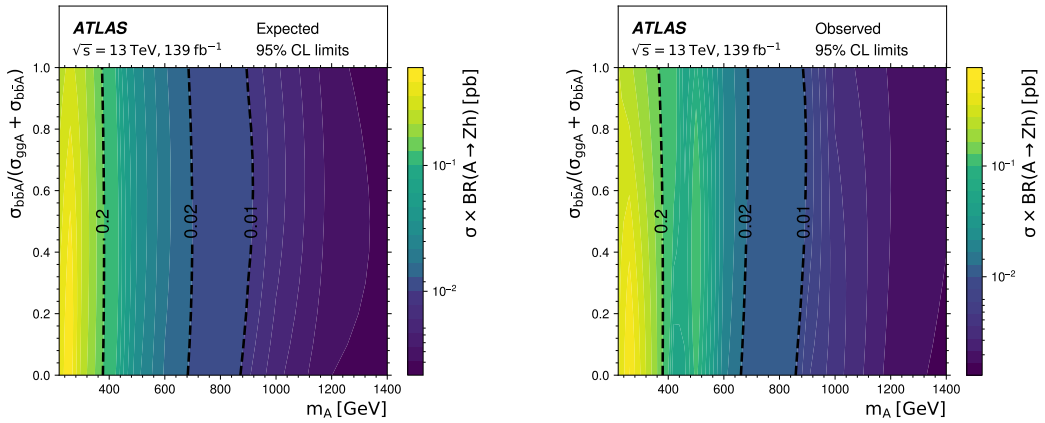


Figure 7.6: Distributions of expected (left) and observed (right) upper limits at the 95% CL on $\sigma(pp \to A/b\bar{b}A \to Zh)$ as a function of m_A and the fraction of bbA production. The dashed lines indicate where the upper limits reach values of 0.2 pb, 0.02 pb and 0.01 pb, respectively.

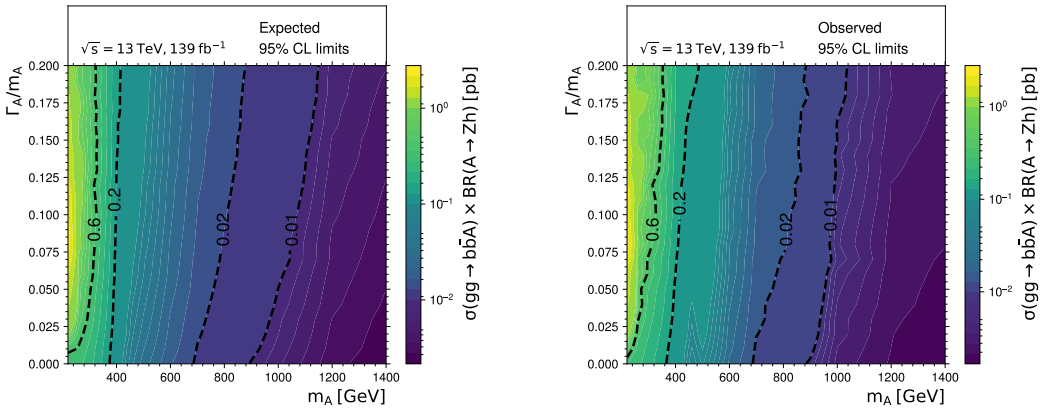


Figure 7.7: Distributions of expected (left) and observed (right) upper limits at the 95% CL on $\sigma(pp \to b\bar{b}A \to Zh)$ of signals with NWA as a function of m_A and the width. The dashed lines indicate where the upper limits reach values of 0.6 pb, 0.2 pb, 0.02 pb, and 0.01 pb, respectively.

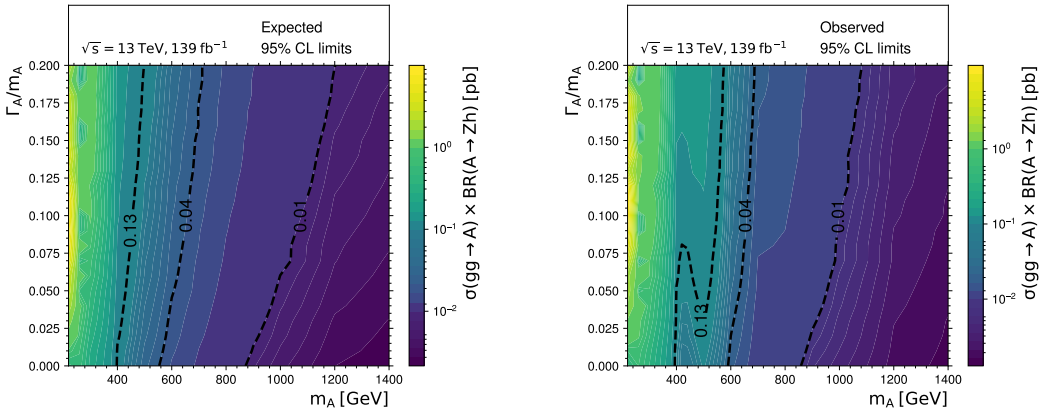


Figure 7.8: Distributions of expected (left) and observed (right) upper limits at the 95% CL on $\sigma(pp \to A \to Zh)$ of signals with NWA as a function of m_A and the width. The dashed lines indicate where the upper limits reach values of 0.13 pb, 0.04 pb, and 0.01 pb, respectively.

The results are interpreted in terms of exclusion limits at the 95% CL on the 2HDM parameters of the Type-I, Type-II, Lepton-specific and Flipped models. The theory cross section of the A boson production is calculated using SusHi [141], and the branching ratio is calculated using 2HDMC [142] under the assumption of $m_h = 125$ GeV and $m_A = m_H = m_{H^\pm}$.

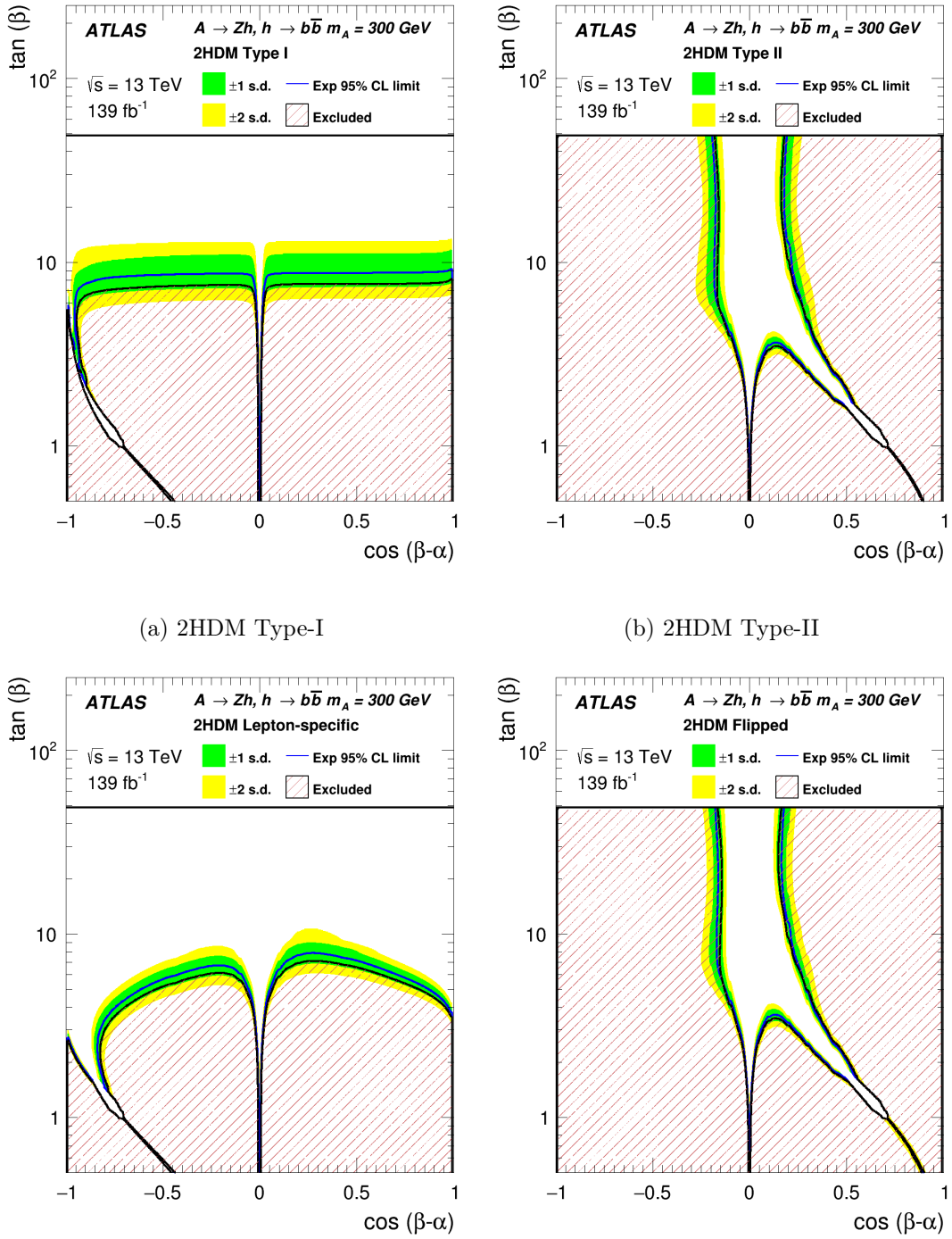
Figure 7.9 - 7.12 show the exclusion limits at the 95% CL on the 2HDM parameters for the Type-I, Type-II, Lepton-specific and Flipped 2HDMs in the $\tan(\beta)$ and $\cos(\beta-\alpha)$ space at $m_A = 300$ GeV, 500 GeV, 700 GeV and 1000 GeV respectively. The plots show that the search has no exclusion power at high $\tan(\beta)$ values for the Type-I and Lepton-specific models due to the reduction of the ggA production cross section as $\tan(\beta)$ increases. This is not the case for the Type-II and Flipped models due to the contribution of the bbA production cross section at high $\tan(\beta)$ values. The search has no sensitivity near the alignment limit ($\cos(\beta-\alpha) = 0$) because of the vanishing $BR(A \to Zh)$. At low $\tan(\beta)$ values, there are also unexcluded regions away from the alignment limit because of the vanishing $BR(h \to bb)$.

Figure 7.15 and Figure 7.16 show the 95% CL exclusion limits in the $\tan(\beta)$ and m_A space at $\cos(\beta-\alpha)$= 0.1 and 0.05 respectively. In the low $\tan(\beta)$ area, the sensitivity reduces at $m_A = 350$ GeV as a result of the $A \to t\bar{t}$ decay mode. At low $\tan(\beta)$ values, the dominating production mode is ggA. At $\cos(\beta-\alpha)$= 0.1, the exclusion area is available at high $\tan(\beta)$ values only for the Type-II and Flipped models because of the contribution from the bbA production. The exclusion area gets smaller as the $\cos(\beta-\alpha)$ value approaches the alignment limit. At $\cos(\beta-\alpha)$= 0.05, only a small area below the $t\bar{t}$ production threshold is excluded. Limits further away from the alignment limit at $\cos(\beta-\alpha)$= 0.2, 0.15 are shown in Figure 7.13, Figure 7.14 respectively.

In the previous analysis [7], limits in the $\tan(\beta)$ and $\cos(\beta-\alpha)$ space were set at $m_A = 300$ GeV. With the increased luminosity, more areas are expected to be excluded. However, the observed limits show that less region is excluded comparing with the previous analysis due to fluctuation at $m_A = 300$ GeV. The previous analysis did not calculate limits in the $\tan(\beta)$ and $\cos(\beta-\alpha)$ space above the $t\bar{t}$ production threshold. The current analysis excluded more areas in the m_A and $\cos(\beta-\alpha)$ space above the $t\bar{t}$

production threshold, which makes it possible to set limits on the $\tan(\beta)$ and $\cos(\beta - \alpha)$ parameters at $m_A = 500$ GeV, 700 GeV and 1000 GeV.

More regions are excluded compared with the previous analysis because of the new signal smearing method. The previous analysis smeared the signal using the Breit-Wigner function, and the method is only valid for signals with $\Gamma_A/m_A < 10\%$. The new signal smearing method is valid for signals with greater natural width, and limits on the parameters space are calculated up to $\Gamma_A/m_A < 20\%$.



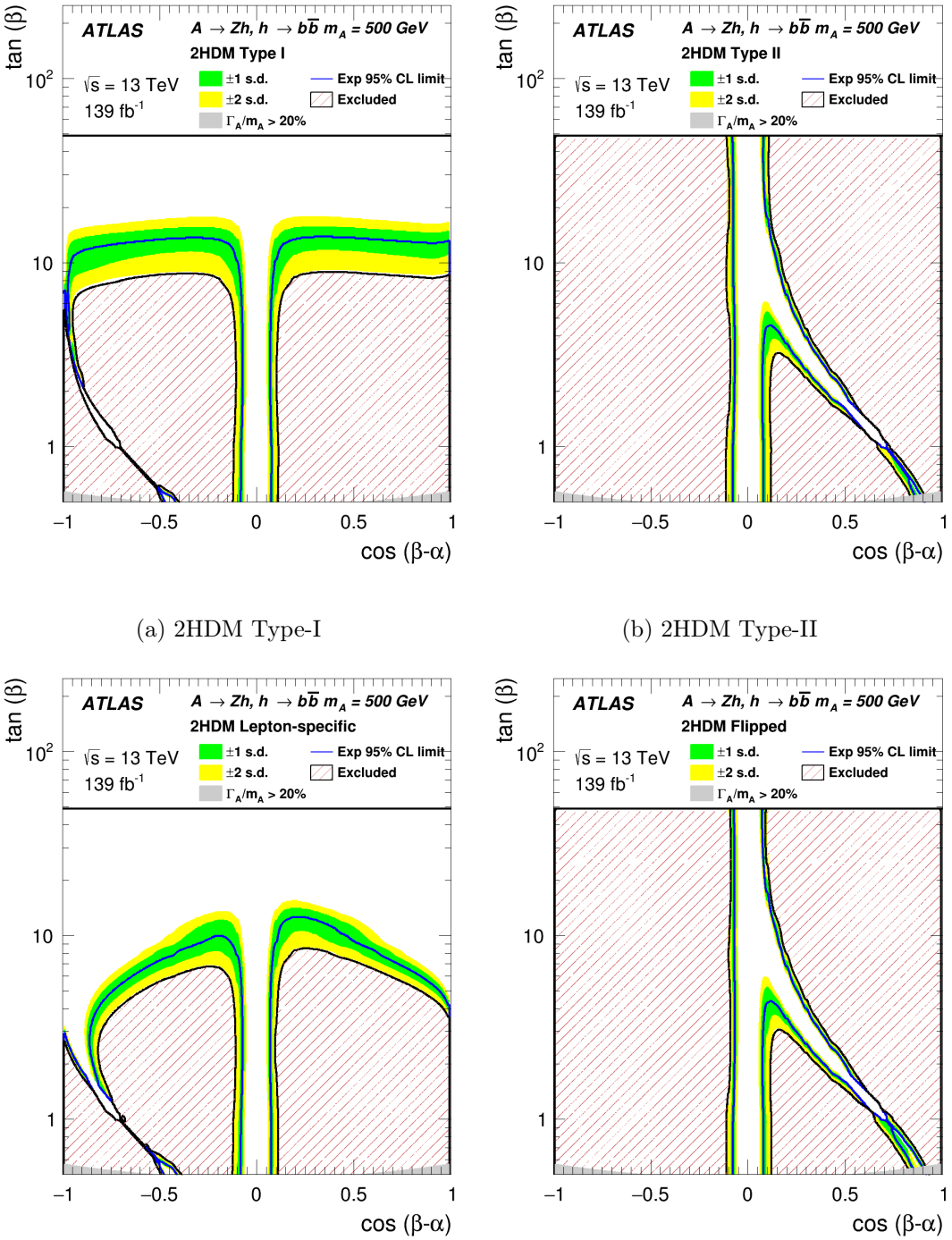
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.9: The exclusion limits at the 95% CL on tan(β) and $cos(\beta - \alpha)$ at m_A = 300 GeV for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



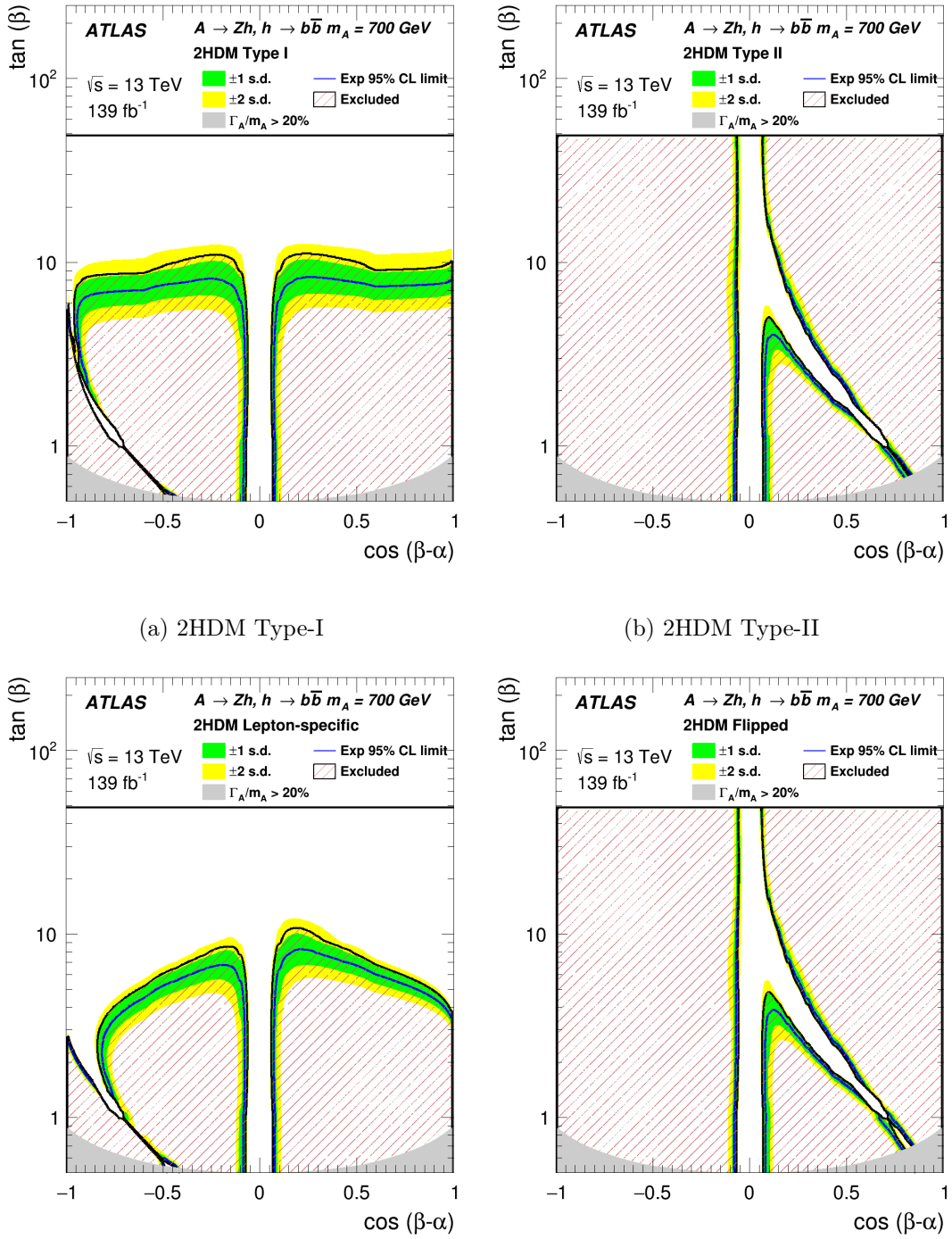
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.10: The exclusion limits at the 95% CL on $\tan(\beta)$ and $cos(\beta - \alpha)$ at $m_A = 500$ GeV for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



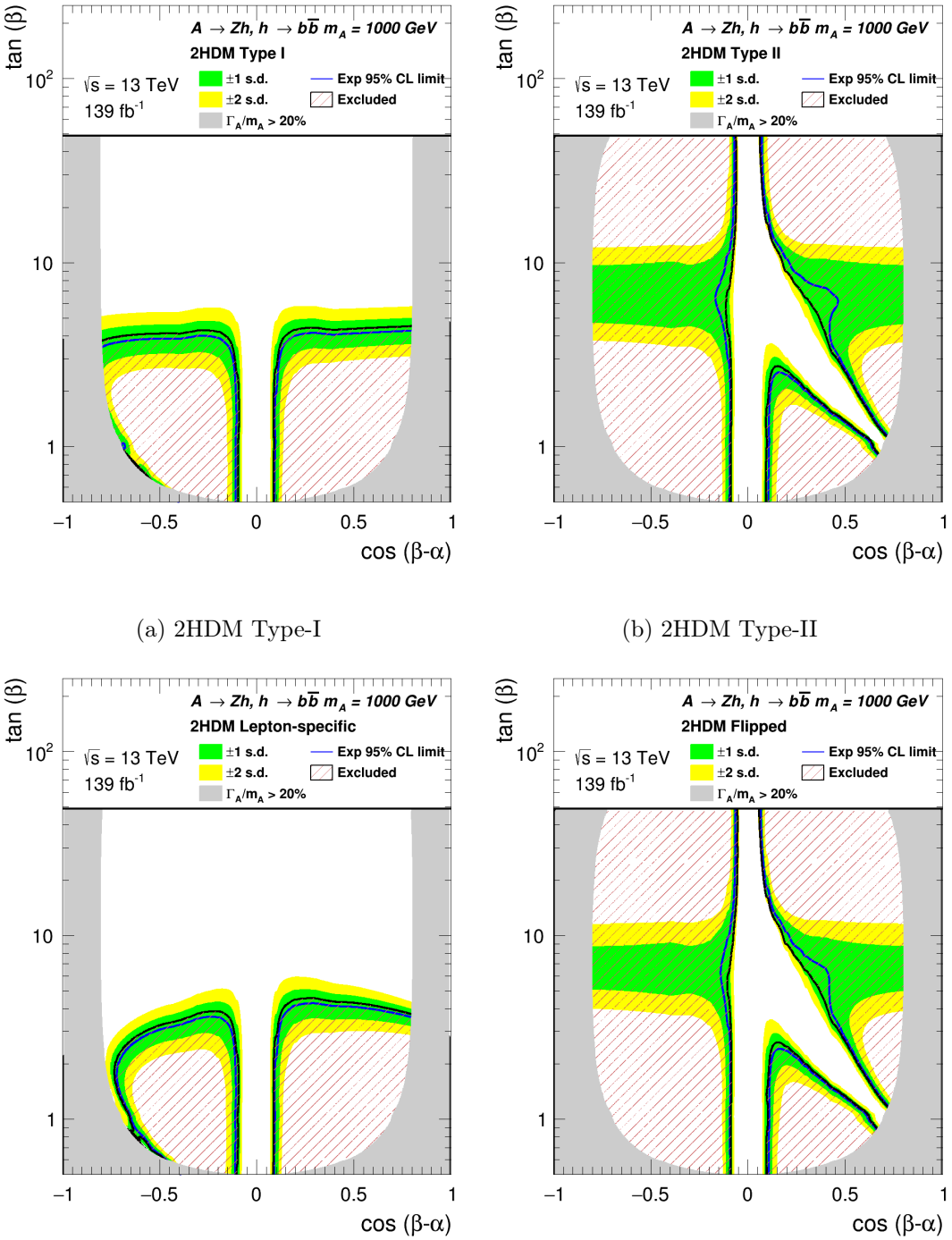
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.11: The exclusion limits at the 95% CL on $\tan(\beta)$ and $cos(\beta - \alpha)$ at $m_A = 700$ GeV for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



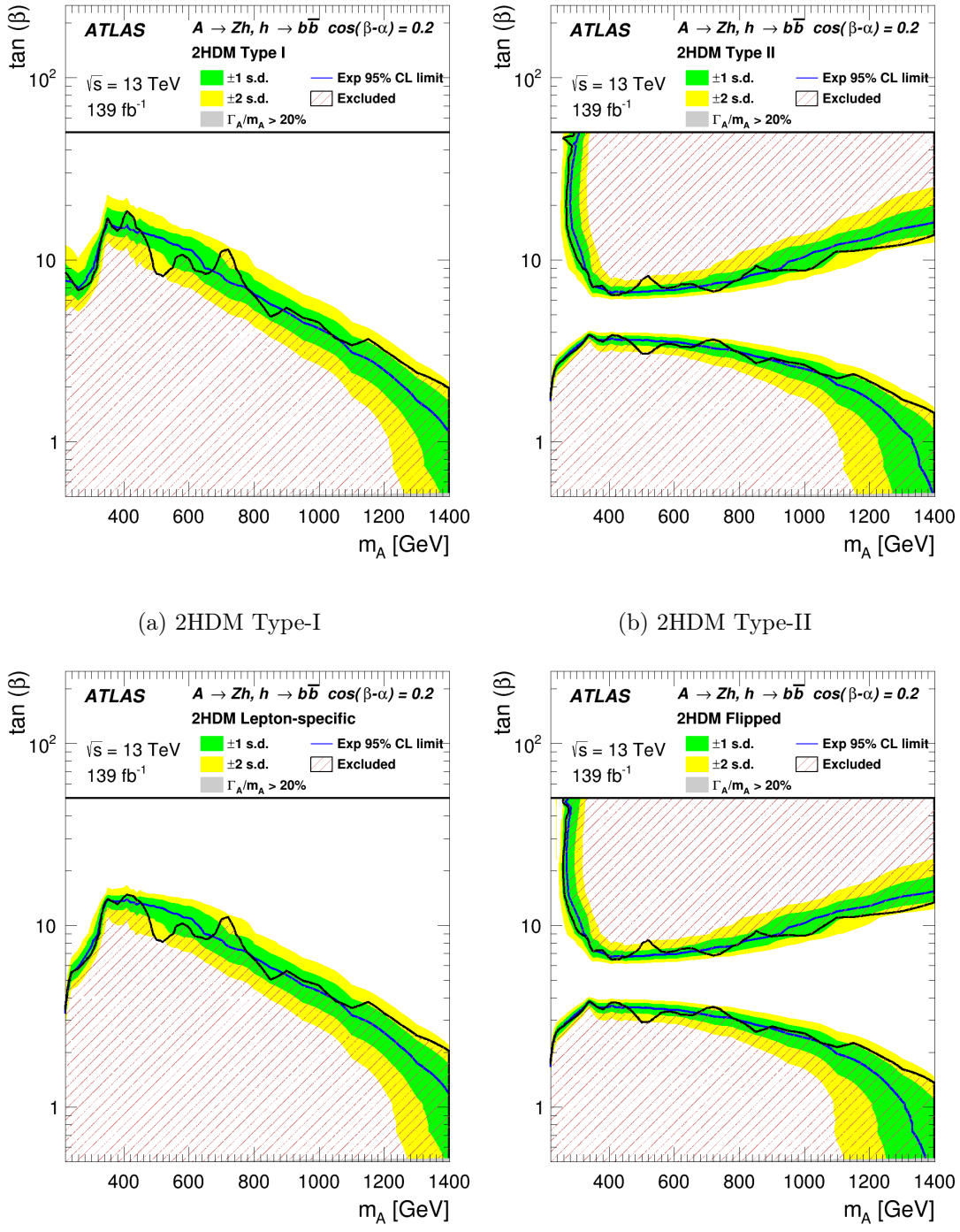
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.12: The exclusion limits at the 95% CL on $\tan(\beta)$ and $cos(\beta - \alpha)$ at $m_A = 1000$ GeV for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



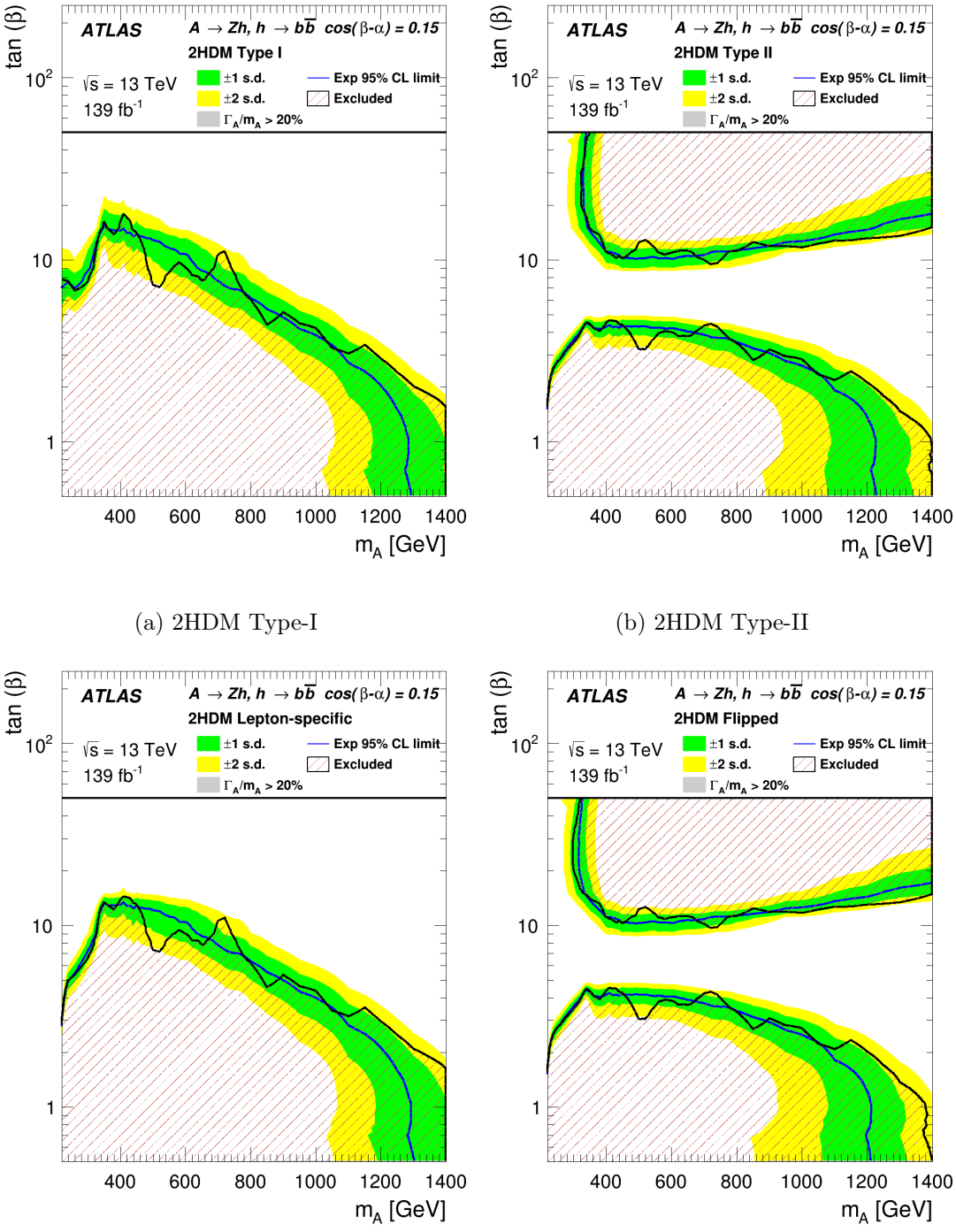
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.13: The exclusion limits at the 95% CL on tan(β) and m_A at $cos(\beta - \alpha) = 0.2$ for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



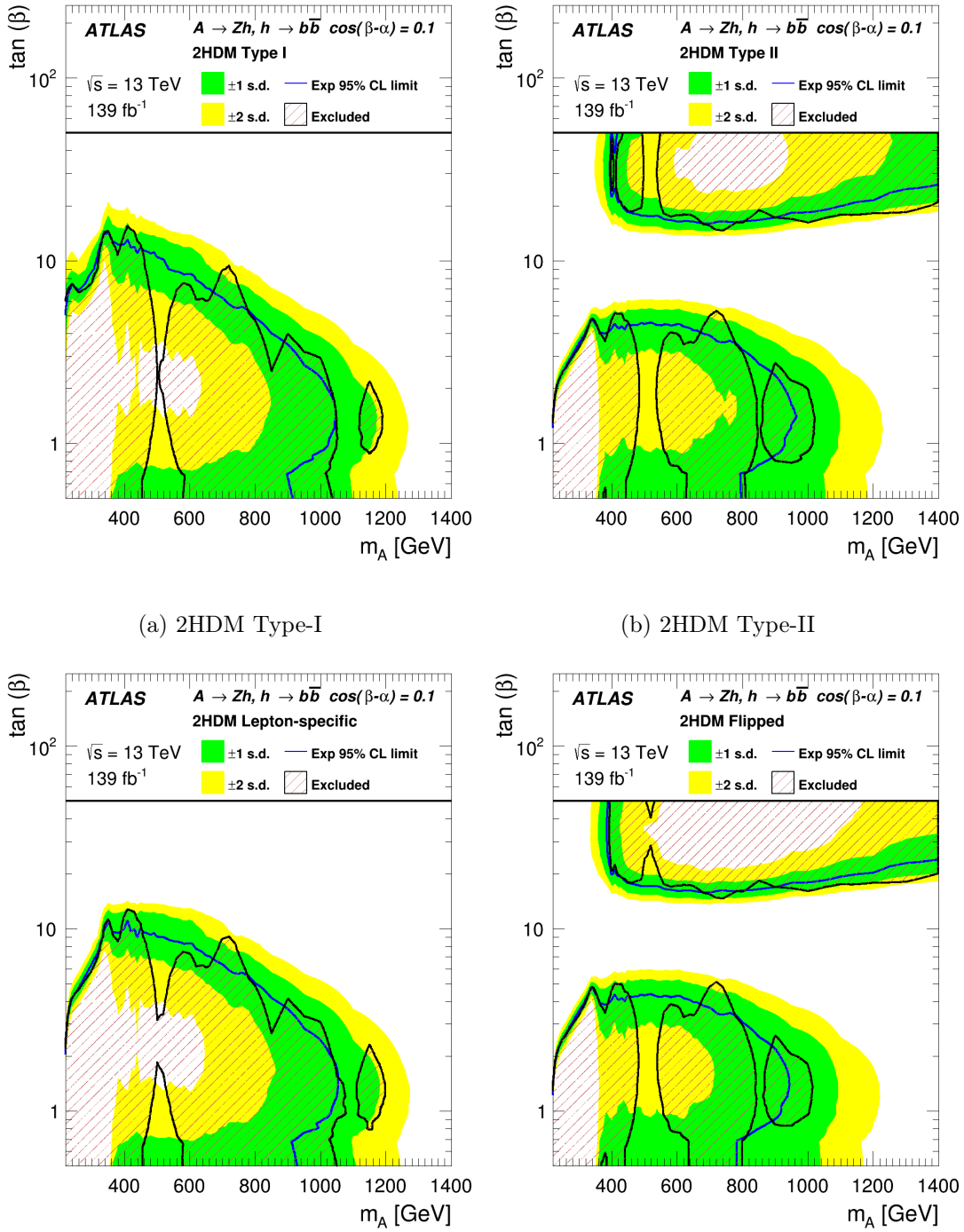
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.14: The exclusion limits at the 95% CL on tan(β) and m_A at $cos(\beta - \alpha) = 0.15$ for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



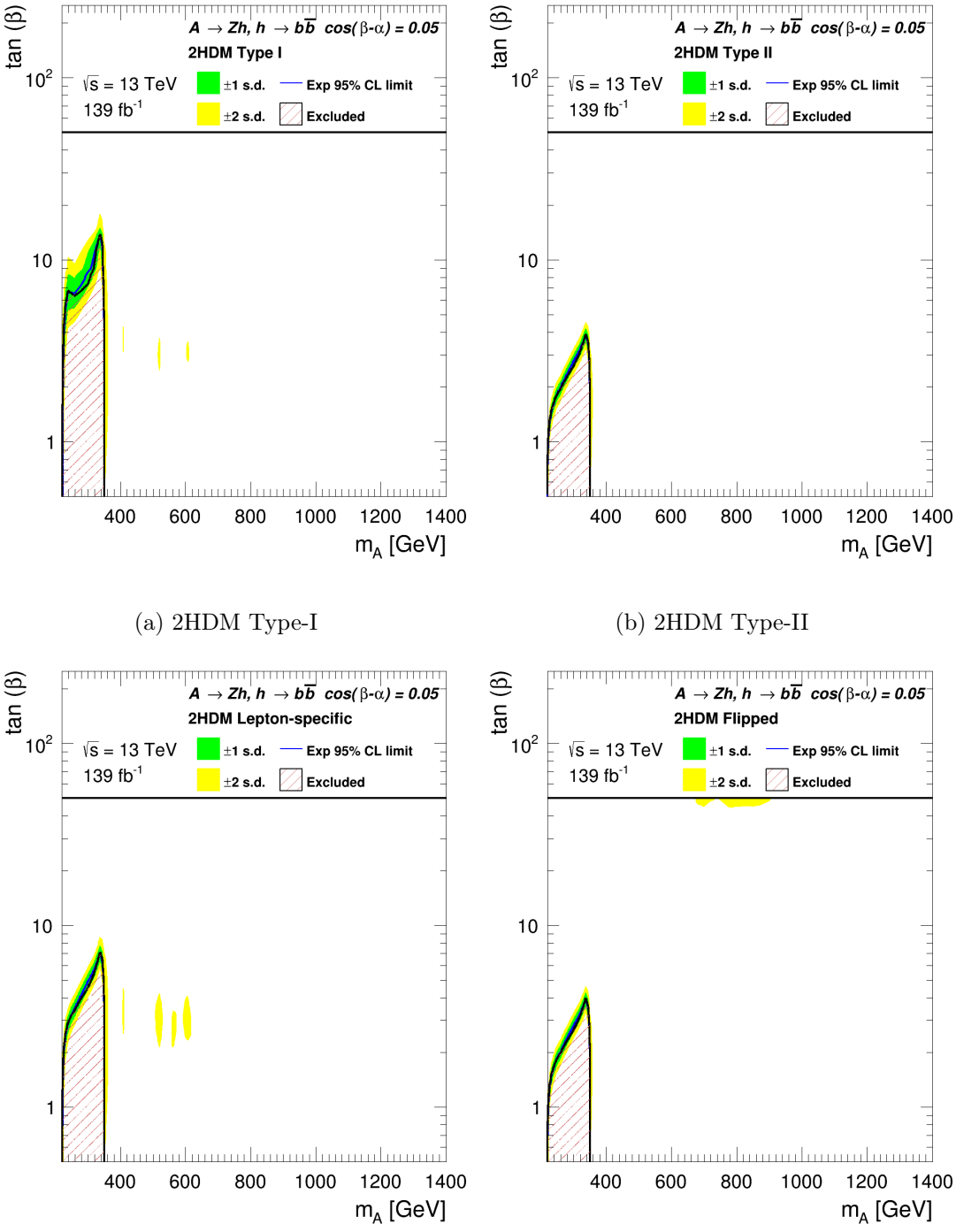
(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.15: The exclusion limits at the 95% CL on tan(β) and m_A at $cos(\beta - \alpha) = 0.1$ for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.



(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Lepton-specific

(d) 2HDM Flipped

Figure 7.16: The exclusion limits at the 95% CL on tan(β) and m_A at $cos(\beta - \alpha) = 0.05$ for the (a) Type-I , (b) Type-II, (c) Lepton-specific and (d) Flipped 2HDMs.

Chapter 8

The *llbb* Search with Neural Networks

Chapter 6 uses an analysis approach where the events are selected by cuts to improve the purity of the signals, and the m_{Vh} is used as the final discriminator. The approach is sensitive to the high-mass signals because of the low number of high-m_{Vh} background events. However, the approach is less sensitive to the low-mass signals because of the large number of low-m_{Vh} background events.

This chapter describes a new analysis approach based on machine learning, where the m_{Vh} is replaced by a variable constructed by deep neural networks as the final discriminator. The new analysis approach is tested using the ggA signal, and the expected limits on the ggA production cross section are calculated.

The chapter first gives a brief introduction to machine learning and deep neural networks. After that, the use of neural networks in the Vh resonance search will be discussed. The discussion of machine learning is mainly based on [143] and [144].

8.1 Machine Learning

Machine learning is a study which enables computers to learn from data. If the relationship between a quantity Y and a set of quantities $X = (X_1, X_2, X_3, ..., X_p)$ can be

written as

$$Y = f(X) + \epsilon, \tag{8.1}$$

where f is an unknown function and ϵ represents a random error, the goal of machine learning is to estimate the unknown function f [143]. The quantity Y is called the response (also called the label) and the quantities X are called features. The estimated response $\hat{Y}$ is

$$\hat{Y} = \hat{f}(X), \tag{8.2}$$

where $\hat{f}$ is an estimation of the unknown function f. The squared error of the estimation is

$$\mathrm{E}(Y - \hat{Y})^2 = [f(x) - \hat{f}(x)]^2 + \mathrm{Var}(\epsilon), \tag{8.3}$$

where the first part is the reducible error while the second part is the irreducible error [143]. A good estimation of f can reduce the reducible error, but the accuracy of machine learning is limited due to the presence of the irreducible error.

The unknown function f can be estimated by simple linear regression or complex nonlinear models such as boosted decision trees or neural networks. The linear models make a good estimation if the true relationship between X and Y is close to linear, but the model would have a high bias for nonlinear problems. The nonlinear models have a lower bias for nonlinear problems, but the variance increases as the model gets more complicated.

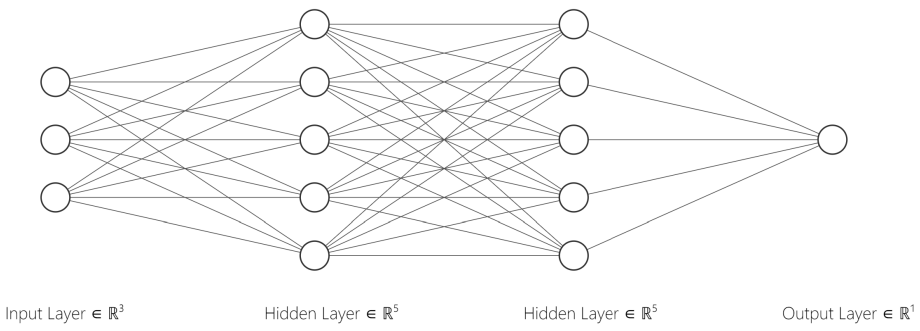


Figure 8.1: A neural network with two hidden layers. The image is created using NN-SVG [145].

8.1.1 Deep Neural Network

The thesis uses deep neural networks to estimate f. The architecture of a neural network is shown in Figure 8.1. The features are the input for the neural network, and they take the nodes in the input layer. The values in the input layer are fed into each node in the first hidden layer in the form of

$$g(w_0 + \sum_{i=1}^{p} w_i X_i), \tag{8.4}$$

where g is the activation function and ws are the weights which need to be determined [143]. The nodes in the second hidden layer accepts values from the first hidden layer in the same form

$$g(w_0 + \sum_{i=1}^{k} w_i X_i), \tag{8.5}$$

where k is the number of nodes in the previous hidden layer. The process continues until it reaches the output layer, and the output of the neural network is $h(v)$, where h is another activation function, and v is the value in the node of the output layer.

The difference between the true response and the estimated response is described by loss functions. In order to find the best estimation of f, the weights in the neural network are optimized by the stochastic gradient descent algorithm [146]. The algorithm minimizes the loss by selecting a small batch of the data randomly to calculate the gradient at each step of the descent, and the full dataset is processed multiple times. Each time the full dataset is processed is called an epoch.

In order to avoid overtraining, the full dataset is split into the training dataset, the validation dataset and the test dataset [144]. The gradient is calculated using the training dataset, while the loss is calculated using the validation dataset at the end of each epoch. The gradient descent is stopped when the loss evaluated using the validation dataset stops decreasing. After the training, the test dataset is used to evaluate the actual performance of the neural network.

In addition to the early stopping, a certain fraction of nodes in each layer is removed randomly when calculating the gradient at each step of the descent. The dropout improves the result of the neural networking training because it prevents the network

from overly relying on a few nodes and reduces overtraining [144].

8.1.2 Parameterized Neural Network

It is challenging to use machine learning in the Vh resonance search because the search covers different signal-mass hypotheses, while most features are correlated with the mass of the signals. Instead of training individual neural networks for each signal-mass hypothesis, the thesis uses parameterized neural networks [147] where one neural network is trained for all signal-mass hypotheses.

The parameterized neural network uses the same architecture as described in Section 8.1.1. It uses mass as an additional feature (X_{mass}). The neural network is trained using the signal events of all masses and the background events. During the training, the X_{mass} of the signal events is the pole mass of the heavy resonance, while the X_{mass} of the background events is chosen randomly according to the distributions of the X_{mass} of the signal events. When using the neural network on data for resonance searches, the X_{mass} of all events is the mass of the resonance it searches for.

Studies [147] show that the parameterized neural network has a better performance than the neural network trained on a mixture of signals of all mass hypotheses without the X_{mass} feature.

8.2 Training and Hyperparameters

The neural network is trained using the simulated events in the signal regions reconstructed and selected using the method described in Chapter 6. The activation function for the hidden layers is the ReLU function [148] which is defined as

$$\text{ReLU}(x) = \max(0, x). \tag{8.6}$$

The ReLU activation function introduces nonlinear behaviour to the neural network, and it is also computationally efficient. The Sigmoid activation function defined as

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{8.7}$$

is applied to the output layer of the neural network. The Sigmoid activation function makes sure that the output falls between 0 and 1. During training, the background events are labelled as 0, while the signal events are labelled as 1. The neural network uses the binary cross entropy function as the loss function which is defined as

$$L = \sum_{i=1}^{N} -w_i[y_i \log x_i + (1 - y_i)\log(1 - x_i)], \tag{8.8}$$

where i is the event number, N is the total number of events, w_i is the weight of event i, y_i is the true label of event i and x_i is the neural-network output of event i [144]. The binary cross entropy is expected to be low when the outputs are close to the true labels.

The neural networks were trained using the simulated Standard-Model backgrounds and the ggA signals with the PyTorch [149] deep learning framework. The full dataset is split into the training dataset, the validation dataset and the test dataset with a ratio of 3:1:1 using scikit-learn [150].

Regions	Number of hidden layers	Number of nodes per hidden layer	Dropout rate
1-tagged resolved	3	150	30%
2-tagged resolved	3	150	30%
1,2-tagged merged	2	100	30%

Table 8.1: Configurations of the neural networks.

Three neural networks were trained separately for the 1 b-tagged resolved events, 2 b-tagged resolved events and the 1 or 2 b-tagged merged events without additional b-tagged track jets not associated with the large-R jet. The configurations of the neural networks are shown in Table 8.1. The neural networks were optimized by starting from a small network and then increasing the number of nodes and hidden layers until the results obtained using the validation dataset stop improving. For each network trained during the optimization, different dropout rates are tested and the optimal one is selected. The neural network of the merged events has fewer nodes and layers compared with the neural networks for the resolved events because of the low number of merged events in the simulation.

The features of the neural networks in the resolved regions are the reconstructed mass of the the Z boson, the h boson and the A boson, the transverse momenta of the

three leading jets and the two leptons, the b-tagging scores of the three leading jets, the $E_T^{\text{miss}}/\sqrt{H_{\text{T}}}$, the angular separations between two leptons and the angular separations between the three leading jets. The neural network for the merged events uses the reconstructed mass of the the Z boson, the h boson and the A boson, the angular separation between the two leptons, the $E_T^{\text{miss}}/\sqrt{H_{\text{T}}}$ and the transverse momenta of the leptons and jets as features. The features are standardized by removing their means and scaling them to the same standard deviation using scikit-learn, which helps the gradient-descent algorithm converge.

The outputs of the three neural networks for the signal and background events in the test dataset are shown in Figure 8.2. The plots show that the neural networks correctly separate signals from the backgrounds because the background distributions peak at 0 and the signal distributions peak near 1. It should be noted that small peaks of signal distributions at low NN output values are observed, and future studies need to be conducted into this features.

Figure 8.3 compares the expected signal-discovery significance calculated using m_{Vh} with the one calculated using the neural-network outputs as discriminators with the same dataset in the 1 b-tagged resolved signal region, 2 b-tagged resolved signal region and the 1+2 b-tagged merged signal regions without additional b-tagged track jets not associated with the large-R jets, where the significance is defined in Equation (6.2) and the definition of the signal regions is described in Section 6.3. Comparing with the cut-based analysis, the neural network provides better performance for low-mass signals. However, the neural networks do not offer significantly better performance for high-mass signals. This is because the expected number of background events in the high-m_{Vh} range is low, and the discrimination power of the neural-network output is similar to that of the m_{Vh}.

8.3 Expected Limits

The performance of the neural networks is tested by a binned maximum-likelihood fit using the distributions of the neural-network output. The fits use the same fit model

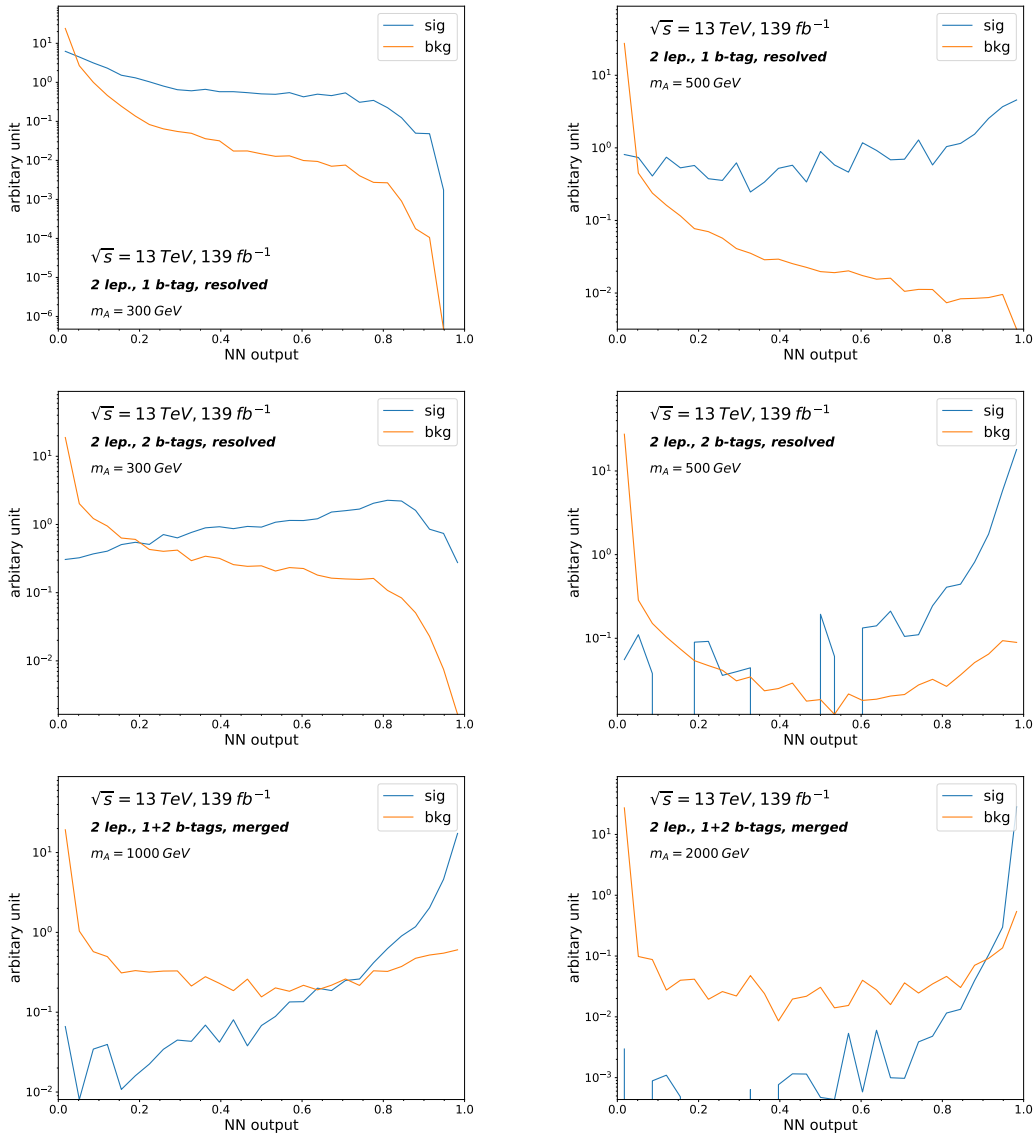


Figure 8.2: The outputs of the neural networks for the 1 b-tagged resolved events with $X_{\text{mass}} = 300$ GeV (top left), 1 b-tagged resolved events with $X_{\text{mass}} = 500$ GeV (top right), 2 b-tagged resolved events with $X_{\text{mass}} = 300$ GeV (middle left), 2 b-tagged resolved events with $X_{\text{mass}} = 500$ GeV (middle right), 1+2 b-tagged merged events with $X_{\text{mass}} = 1000$ GeV (bottom left) and 1+2 b-tagged merged events with $X_{\text{mass}} = 2000$ GeV (bottom right).

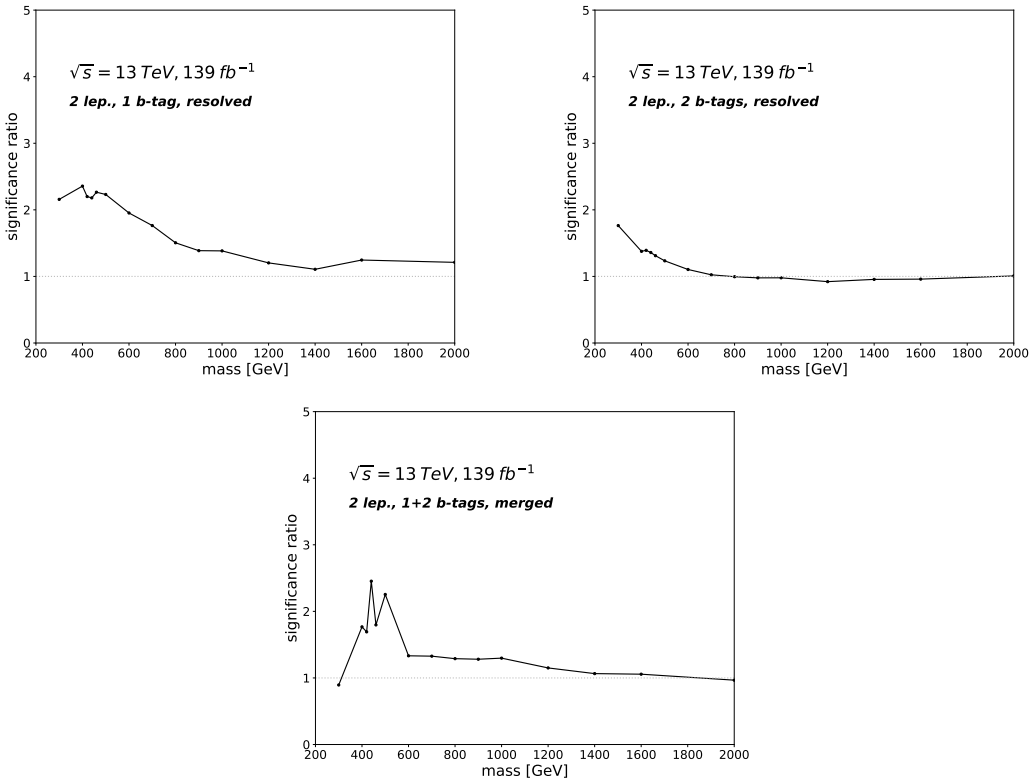


Figure 8.3: The ratios of the expected signal significance between the neural-network analysis and the cut-based analysis in the 1 b-tagged resolved region (top left), the 2 b-tagged resolved region (top left) and the 1+2 b-tagged merged regions without additional b-tagged track jets not associated with the large-R jets (bottom).

as described in Chapter 6. For this test, the fit only includes nuisance parameters that control the normalization of the individual background components, while the nuisance parameters that control the distributions' shape are ignored.

The modelling of the backgrounds is tested by fitting the simulated distribution of the neural network output to that of the data using the events in the m_{bb} sideband control regions and the top control region, and the post-fit distributions for neural networks with $X_{\text{mass}} = 500$ GeV are shown in Figure 8.4. The plots show overall good agreement between the data and the simulation. Mismodelling is observed in the 1 b-tagged regions where there is more data than expected backgrounds between 0 to 0.3.

The expected limit is calculated by an Asimov fit using the simulated events in the signal regions and the top control region, and the distributions of the signals and the backgrounds at $X_{\text{mass}} = 500$ GeV are shown in Figure 8.5. Figure 8.6 shows the upper

limits at the 95% CL on the production cross section of ggA signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}$. A comparison is made with expected limits calculated by the cut-based approach described in Chapter 6. For this comparison, the cut-based limits are also calculated without considering nuisance parameters which affect the shape of the distributions. The comparison shows that the parameterized neural network provides a 60% improvement for the limit of the signal at 300 GeV compared with the cut-based approach. However, the neural network is expected to have a similar performance as the cut-based analysis for signals with mass above 800 GeV. It should be noted that the training and validation datasets are also used in the limit calculation, and the effect of overtraining may affect the limits even if early stopping and dropout are applied. Moreover, the effect of shape uncertainties may affect the neural-network approach more than the cut-based approach. As a result, the neural network may offer less improvement than predicted in this study.

In order to understand the neural network, the importance of the features is calculated by estimating their SHAP values [151] which are the average contribution of the feature to the output in every possible feature subset. The most important feature for both the resolved and the merged events is the reconstructed mass of the A boson which is also the final discriminator of the cut-based analysis. For low-mass signals, the second most important feature is the reconstructed mass of the h boson since the events with the h mass close to 125 GeV are more likely to be from the $A \to Zh$ decay. The b-tagging scores and the angular separations between jets also contribute to the output, which suggests that the analysis strategy discussed in Chapter 6 often selects the wrong jets for the h boson reconstruction, and the neural network improves the sensitivity of the search by recovering information related to the third leading jets.

8.4 Outlook

The preliminary studies have shown that the parameterized neural network significantly improves the sensitivity of the search for the low-mass signals in the 2-lepton channel. The neural network does not improve the performance for the high-mass signals, but

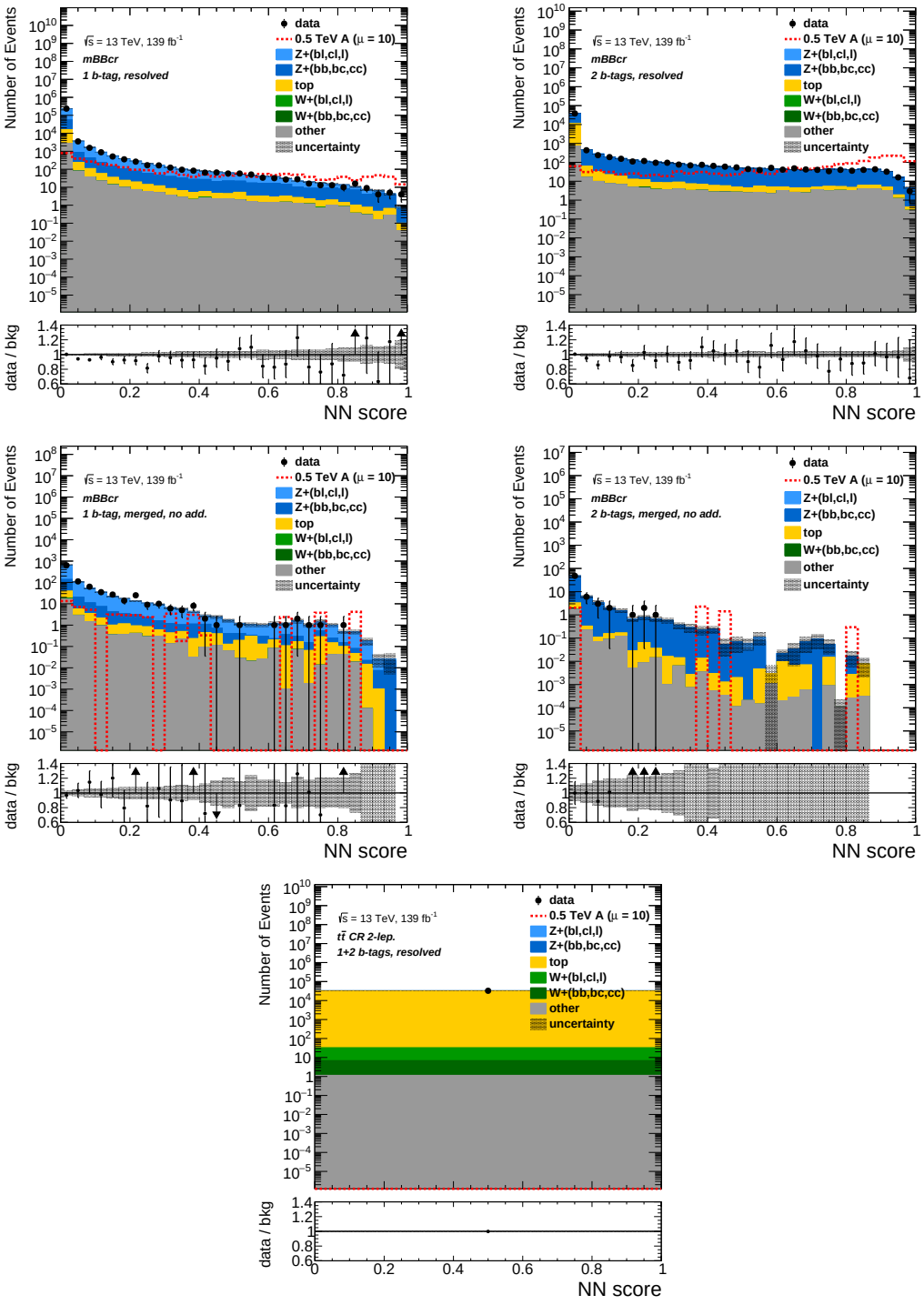


Figure 8.4: Post-fit distributions of neural-network output with $X_{\text{mass}} = 500$ GeV in the 1 b-tagged resolved m_{bb} sideband control region (top left), 2 b-tagged resolved m_{bb} sideband control region (top right), 1 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (middle left), 2 b-tagged merged m_{bb} sideband control region with no additional b-tagged track jets (middle right) and 1+2 b-tagged resolved top control region (bottom).

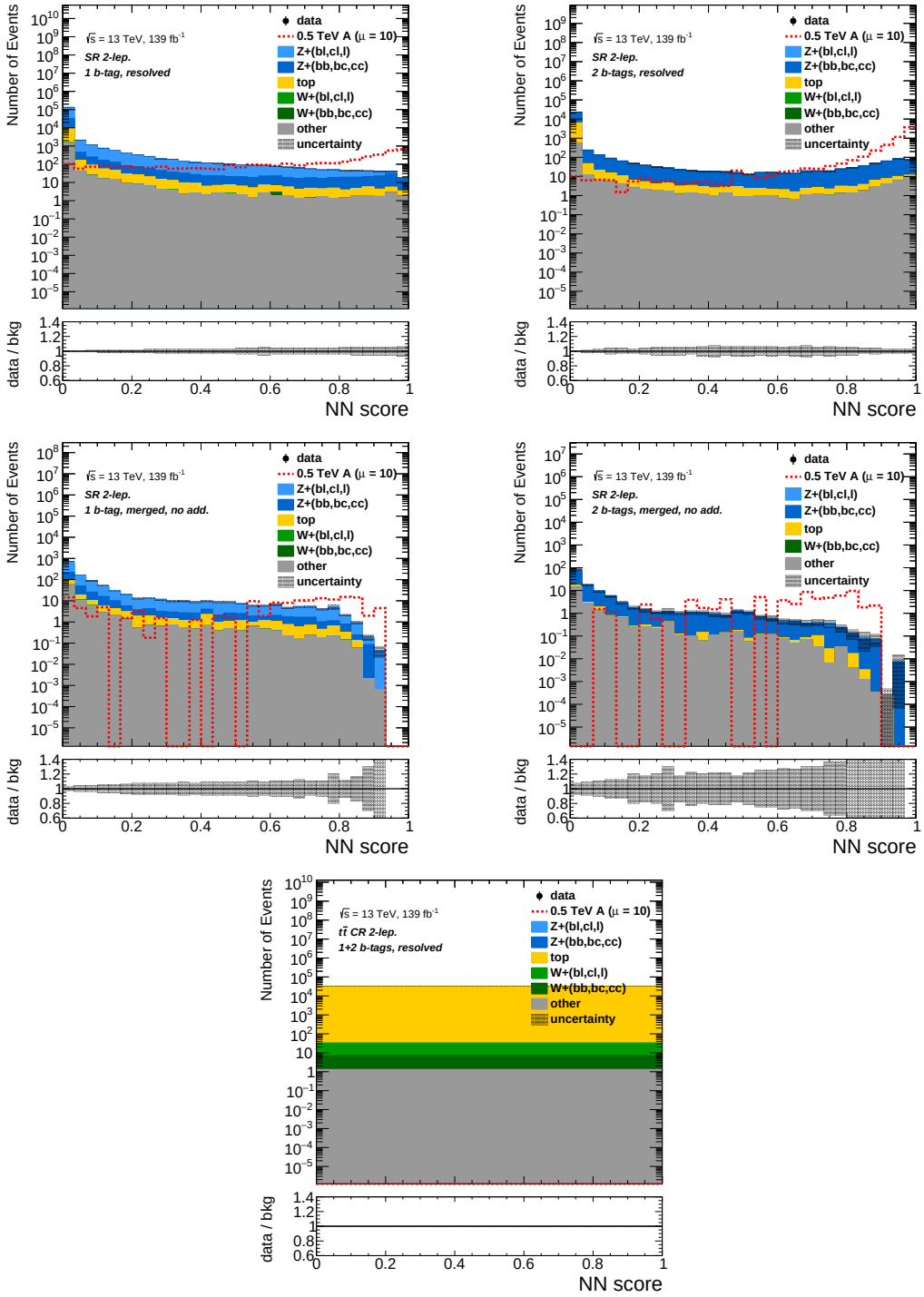


Figure 8.5: Pre-fit distributions of neural-network output with $X_{\text{mass}} = 500$ GeV in the 1 b-tagged resolved signal region (top left), 2 b-tagged resolved signal region (top right), 1 b-tagged merged signal region with no additional b-tagged track jets (middle left), 2 b-tagged merged signal region with no additional b-tagged track jets (middle right) and 1+2 b-tagged resolved top control region (bottom).

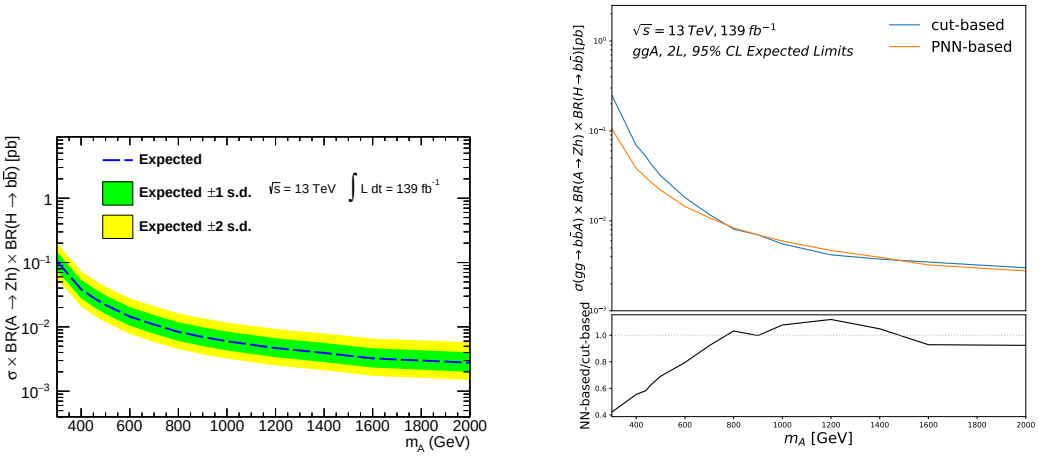


Figure 8.6: The expected upper limits at the 95% CL on the production cross section of ggA signal times the branching ratio to Zh and the branching ratio of $h \to b\bar{b}$ (left) and a comparison with the limits computed by the cut-based analysis (right).

the results are still promising because the 2-lepton channel is less important than the 0-lepton channel for the high-mass signals.

Although the neural network improves the sensitivity of the search, mismodelling of backgrounds has been observed. Studies are needed to understand the effect of the shape-related nuisance parameters on the modelling of backgrounds. Reweighting of the simulated events is required if the shape-related nuisance parameters do not improve the background modelling.

In this chapter, the neural networks are trained and tested on signals following the narrow-width approximation, but the signal width also affects the performance of the neural network. Studies are required to test the performance of the parameterized neural network on large-width signals.

The second round of the Vh resonance search is currently ongoing, and the parameterized neural network approach will be used to search for the A boson and the Z' boson using the full Run 2 ATLAS dataset.

Chapter 9

Conclusion

The Standard Model describes the existence of all known particles and their interactions, and the discovery of the Higgs boson in 2012 confirmed another of its predictions. However, the Standard Model has its limitations, and the observed Higgs boson could be part of an extended Higgs sector. The 2HDM predicts five Higgs-like bosons, and the pseudoscalar Higgs A can decay into an h boson and a Z boson. Alternatively, the HVT model predicts the existence of two new heavy vector bosons Z' and W', and they couple to the Standard-Model Z/W boson and the h boson.

The thesis described a search for a heavy resonance decaying into an h boson and a Z boson in the final states of $l^-l^+b\bar{b}/\nu\bar{\nu}b\bar{b}$. The search is split into the 0-lepton channel and the 2-lepton channel according to the number of the charged leptons in the final states. The low-p_T h boson candidates are reconstructed by two small-R jets, while the two b-jets from the high-p_T h boson candidates are collimated and the h boson is reconstructed by one large-R jet. The b-jets are identified using b-tagging algorithms, and the analysis selects events with at least one b-tagged jet. The events are selected by cutting on variables such as the p_T of electrons and jets, the reconstructed mass of the Z boson and h boson candidates, missing transverse momentum and angular separations between the reconstructed final-state objects. The m_{bb} sideband control regions are defined using events outside of the h boson mass window, and the top control region is defined by selecting events with one electron and one muon and without the requirement on the missing transverse energy. The shape of all backgrounds is estimated by the MC

simulation. The normalizations of the two largest backgrounds, the Z+jet background and the top background, are determined by data, while the normalizations of other small backgrounds are constrained to their Standard-Model prediction. The simulated events in the signal region are reweighted using reweighting functions derived from the m_{bb} sideband control regions to improve the modelling of the background.

The thesis first described a cut-based approach which used m_{Vh} as the final discriminator. The search probed the mass range from 220 GeV to 5 TeV, and its result is compatible with the Standard-Model prediction. Upper limits on $\sigma(pp \to Z' \to Zh)$ at the 95% CL are set from 0.9 pb at 300 GeV to 0.3 fb at 5 GeV. Upper limits on $\sigma(gg \to A \to Zh)$ at the 95% CL are set from 0.6 pb at 220 GeV to 3 fb at 2 GeV. Upper limits on $\sigma(gg \to bbA \to Zh)$ at the 95% CL are set from 0.5 pb at 220 GeV to 2.5 fb at 2 GeV. The limits are calculated using asymptotic formulas, and the difference between the asymptotic limits and the pseudo-experiment limits are within 20%. The distribution of the reconstructed A mass are smeared to estimate the distributions of the large-width signals. The exclusion contours in the parameter space of the Type-I, Type-II, Lepton-specific and Flipped 2HDMs are calculated using the upper limits on cross section as a function of the bbA/ggA fraction and the natural width of the A boson.

The thesis also explored the use of machine learning. Parametrized neural networks are trained using the simulated background events and the simulated ggA signal events in the 2-lepton channel. The neural networks use the ReLU activation function for the input and hidden layers, while the Sigmoid function is used for the output layer. The neural networks are trained using the binary cross entropy as the loss function, and early stopping and dropout are used to avoid overtraining. The Hyperparameters are optimized to improve the results. The analysis uses the output of the neural network as the final discriminator. The comparison between the cut-based analysis and the machine-learning-based analysis shows the neural network provides up to 60% improvement of the expected upper limits for low-mass signals, while its performance is similar to that of the cut-based analysis at high mass.

In addition to the resonance search, the thesis also described the simulation of

the L1Calo trigger for the ATLAS Phase-I upgrade. The author was involved in the development of the eFEX module which identifies e/γ and τ candidates using the energy deposit in the calorimeters. The development of the e/γ algorithm module, the mapping modules and the modules for the validation were discussed in the thesis.

The upcoming Run 3 and the High-Luminosity LHC upgrade will make it possible to measure the Standard-Model and beyond-the-Standard-Model processes to unprecedented precision. The proposed Future Circular Collider (FCC) [152] and Super Proton-Proton Collider (SppC) [153] will allow searches for new particles with the proton-proton data at a centre-of-mass energy beyond 100 TeV. In addition to collider physics, progress is also being made in the measurement of neutrino properties and the searches for dark matter. With new experiments and improved analysis techniques, particle-physics research will continue improving our understanding of the Universe.

Bibliography

1. ATLAS Collaboration. *Search for heavy resonances decaying into a Z or W boson and a Higgs boson in final states with leptons and b-jets in* 139 fb^{-1} *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector* tech. rep. submitted to JHEP. (CERN, Geneva, July 2022). arXiv: 2207.00230. https://cds.cern.ch/record/2814371.

2. ATLAS Collaboration. *Search for heavy resonances decaying into a Z boson and a Higgs boson in final states with leptons and b-jets in* 139 fb^{-1} *of pp collisions at* $\sqrt{s} = 13TeV$ *with the ATLAS detector* tech. rep. (CERN, Geneva, Aug. 2020). http://cds.cern.ch/record/2728053.

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,** 1–29. arXiv: 1207.7214 [hep-ex] (2012).

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,** 30–61. arXiv: 1207.7235 [hep-ex] (2012).

5. Branco, G. *et al.* Theory and phenomenology of two-Higgs-doublet models. *Physics Reports* **516.** Theory and phenomenology of two-Higgs-doublet models, 1–102. ISSN: 0370-1573. https://www.sciencedirect.com/science/article/pii/S0370157312000695 (2012).

6. Pappadopulo, D., Thamm, A., Torre, R. & Wulzer, A. Heavy Vector Triplets: Bridging Theory and Data. *JHEP* **09,** 060. arXiv: 1402.4431 [hep-ph] (2014).

7. ATLAS Collaboration. Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$

TeV *pp* collisions with the ATLAS detector. *JHEP* **03,** 174. eprint: 1712.06518. https://link.springer.com/article/10.1007/JHEP03(2018)174 (2018).

8. CMS Collaboration. Search for a heavy vector resonance decaying to a Z boson and a Higgs boson in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C* **81,** 688. arXiv: 2102.08198 [hep-ex] (2021).

9. Cranmer, K. & Yavin, I. RECAST: Extending the Impact of Existing Analyses. *JHEP* **04,** 038. arXiv: 1010.2506 [hep-ex] (2011).

10. Thomson, M. *Modern Particle Physics* (Cambridge University Press, 2013).

11. David, G. *Introduction to Elementary Particles, 2nd, Revised Edition* (Wiley, 2008).

12. Schwartz, M. D. *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2013).

13. Buckley, A., White, C. & White, M. *Practical Collider Physics* https://dx.doi.org/10.1088/978-0-7503-2444-1 (IOP Publishing, 2021).

14. Ellis, J. P. TikZ-Feynman: Feynman diagrams with TikZ. *Computer Physics Communications* **210,** 103–123. ISSN: 0010-4655. https://www.sciencedirect.com/science/article/pii/S0010465516302521 (2017).

15. Gross, D. J. & Wilczek, F. Asymptotically Free Gauge Theories. I. *Phys. Rev. D* **8,** 3633–3652. https://link.aps.org/doi/10.1103/PhysRevD.8.3633 (10 Nov. 1973).

16. Altarelli, G. & Parisi, G. Asymptotic freedom in parton language. *Nuclear Physics B* **126,** 298–318. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/0550321377903844 (1977).

17. Agostinelli, S. *et al.* Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506,** 250–303. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900203013688 (2003).

18. ATLAS Collaboration. The ATLAS Simulation Infrastructure. *The European Physical Journal C* **70,** 823–874. ISSN: 1434-6052. http://dx.doi.org/10.1140/epjc/s10052-010-1429-9 (Sept. 2010).

19. Yorikiyo, N. *Beyond the Standard Model of Elementary Particle Physics* (Wiley, 2014).

20. MARTIN, S. P. in *Perspectives on Supersymmetry* 1–98 (). eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789812839657_0001. https://www.worldscientific.com/doi/abs/10.1142/9789812839657_0001.

21. Ward, J. C. An Identity in Quantum Electrodynamics. *Phys. Rev.* **78,** 182–182. https://link.aps.org/doi/10.1103/PhysRev.78.182 (2 Apr. 1950).

22. Kim, J. E. Light pseudoscalars, particle physics and cosmology. *Physics Reports* **150,** 1–177. ISSN: 0370-1573. https://www.sciencedirect.com/science/article/pii/0370157387900172 (1987).

23. Paschos, E. A. Diagonal neutral currents. *Phys. Rev. D* **15,** 1966–1972. https://link.aps.org/doi/10.1103/PhysRevD.15.1966 (7 Apr. 1977).

24. Glashow, S. L. & Weinberg, S. Natural conservation laws for neutral currents. *Phys. Rev. D* **15,** 1958–1965. https://link.aps.org/doi/10.1103/PhysRevD.15.1958 (7 Apr. 1977).

25. Kominis, D. The phenomenology of the CP-odd scalar in two-doublet models. *Nuclear Physics B* **427,** 575–613. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/0550321394906416 (1994).

26. CMS Collaboration. Search for a heavy pseudoscalar boson decaying to a Z and a Higgs boson at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C* **79,** 564. arXiv: 1903.00941 [hep-ex] (2019).

27. Kling, F., Su, S. & Su, W. 2HDM neutral scalars under the LHC. *Journal of High Energy Physics* **2020.** https://doi.org/10.1007%2Fjhep06%282020%29163 (June 2020).

28. ATLAS Collaboration. Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. *Phys. Rev. D* **101,** 012002. https://link.aps.org/doi/10.1103/PhysRevD.101.012002 (1 Jan. 2020).

29. Sannino, F. & Tuominen, K. Orientifold theory dynamics and symmetry breaking. *Phys. Rev. D* **71,** 051901. https://link.aps.org/doi/10.1103/PhysRevD.71.051901 (5 Mar. 2005).

30. Belyaev, A. *et al.* Technicolor walks at the LHC. *Phys. Rev. D* **79,** 035006. https://link.aps.org/doi/10.1103/PhysRevD.79.035006 (3 Feb. 2009).

31. Schmaltz, M. & Tucker-Smith, D. LITTLE HIGGS THEORIES. *Annual Review of Nuclear and Particle Science* **55,** 229–270. eprint: https://doi.org/10.1146/annurev.nucl.55.090704.151502. https://doi.org/10.1146/annurev.nucl.55.090704.151502 (2005).

32. 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,** 1165. arXiv: 2004.14636 [hep-ex] (2020).

33. CMS Collaboration. Search for heavy resonances decaying to WW, WZ, or WH boson pairs in the lepton plus merged jet final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D* **105,** 032008. arXiv: 2109.06055 [hep-ex] (2022).

34. Evans, L. & Bryant, P. LHC Machine. *Journal of Instrumentation* **3,** S08001–S08001. https://doi.org/10.1088/1748-0221/3/08/s08001 (Aug. 2008).

35. ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation* **3,** S08003–S08003. https://doi.org/10.1088/1748-0221/3/08/s08003 (Aug. 2008).

36. CMS Collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software (2006).

37. LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation* **3,** S08005–S08005. https://doi.org/10.1088/1748-0221/3/08/s08005 (Aug. 2008).

38. ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation* **3,** S08002–S08002. https://doi.org/10.1088/1748-0221/3/08/s08002 (Aug. 2008).

39. ATLAS Collaboration. *Luminosity Public Results Run2* https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

40. Fartoukh, S. *et al. LHC Configuration and Operational Scenario for Run 3* tech. rep. (CERN, Geneva, Nov. 2021). https://cds.cern.ch/record/2790409.

41. Outreach, A. *ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC* 2010. https://cds.cern.ch/record/1457044.

42. ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *The European Physical Journal C* **70,** 787–821. https://doi.org/10.1140%2Fepjc%2Fs10052-010-1366-7 (2010).

43. La Rosa, A. The ATLAS Insertable B-Layer: from construction to operation. *JINST* **11** (eds Gemme, C. & Rossi, L.) C12036. arXiv: 1610.01994 [physics.ins-det] (2016).

44. Aharrouche, M. *et al.* Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **614,** 400–432. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900209023924 (2010).

45. Adragna, P. *et al.* Testbeam studies of production modules of the ATLAS Tile Calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **606,** 362–394. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S016890020900792X (2009).

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

47. Achenbach, R. *et al.* The ATLAS Level-1 Calorimeter Trigger. *Journal of Instrumentation* **3,** P03001–P03001. https://doi.org/10.1088/1748-0221/3/03/p03001 (Mar. 2008).

48. ATLAS Collaboration. *ATLAS level-1 trigger: Technical Design Report* http://cds.cern.ch/record/381429 (CERN, Geneva, 1998).

49. ATLAS Collaboration. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report* http://cds.cern.ch/record/616089 (CERN, Geneva, 2003).

50. ATLAS Collaboration. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J. C* **77,** 317. 76 p. arXiv: 1611.09661. https://cds.cern.ch/record/2235584 (Nov. 2016).

51. ATLAS Collaboration. *Trigger Menu in 2016* tech. rep. (CERN, Geneva, Jan. 2017). https://cds.cern.ch/record/2242069.

52. ATLAS Collaboration. *Trigger Menu in 2017* tech. rep. (CERN, Geneva, June 2018). https://cds.cern.ch/record/2625986.

53. ATLAS Collaboration. *Trigger menu in 2018* tech. rep. (CERN, Geneva, Oct. 2019). https://cds.cern.ch/record/2693402.

54. Stelzer, B. *The New Small Wheel Upgrade Project of the ATLAS Experiment* tech. rep. (CERN, Geneva, Oct. 2014). https://cds.cern.ch/record/1958265.

55. Kawamoto, T. *et al. New Small Wheel Technical Design Report* tech. rep. ATLAS New Small Wheel Technical Design Report (June 2013). https://cds.cern.ch/record/1552862.

56. Iguchi, R. The Phase-I Trigger Readout Electronics Upgrade of the ATLAS Liquid Argon Calorimeters. *PoS* **EPS-HEP2019,** 202. 5 p. https://cds.cern.ch/record/2760981 (2020).

57. ATLAS Collaboration. *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System* tech. rep. CERN-LHCC-2013-018. ATLAS-TDR-023. Final

version presented to December 2013 LHCC. (Sept. 2013). https://cds.cern.ch/record/1602235.

58. ATLAS Collaboration. *ATLAS Computing: technical design report* https://cds.cern.ch/record/837738 (CERN, Geneva, 2005).

59. Qian, W. Design and test performance of the ATLAS Feature Extractor trigger boards for the Phase-1 Upgrade. *Journal of Instrumentation* **12,** C01079–C01079. https://doi.org/10.1088/1748-0221/12/01/c01079 (Jan. 2017).

60. Schwienhorst, R. The Phase-1 upgrade of the ATLAS first level calorimeter trigger. *Journal of Instrumentation* **11,** C01018–C01018. https://doi.org/10.1088/1748-0221/11/01/c01018 (Jan. 2016).

61. Mohsen, E. in *White Paper* (2013). https://docs.xilinx.com/v/u/en-US/wp432-K7-Perf-Power-Cost.

62. Barrand, G. *et al.* GAUDI — A software architecture and framework for building HEP data processing applications. *Computer Physics Communications* **140.** CHEP2000, 45–55. ISSN: 0010-4655. https://www.sciencedirect.com/science/article/pii/S0010465501002545 (2001).

63. ATLAS Collaboration. *The ATLAS multithreaded offline framework* tech. rep. (CERN, Geneva, Nov. 2018). https://cds.cern.ch/record/2649080.

64. Calafiura, P., Leggett, C., Quarrie, D., Ma, H. & Rajagopalan, S. The StoreGate: A Data model for the Atlas software architecture. *eConf* **C0303241,** MOJT008. arXiv: cs/0306089 (2003).

65. Rumbaugh, J., Jacobson, I. & Booch, G. *Unified Modeling Language Reference Manual, The (2nd Edition)* ISBN: 0321245628 (Pearson Higher Education, 2004).

66. Barrand, G. *et al.* GAUDI - A software architecture and framework for building HEP data processing applications. *Comput. Phys. Commun.* **140,** 45–55 (2001).

67. Brun, R. & Rademakers, F. ROOT — An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389.** New Comput-

ing Techniques in Physics Research V, 81–86. ISSN: 0168-9002. http://www.sciencedirect.com/science/article/pii/S016890029700048X (1997).

68. Achenbach, R. *et al. Control, test and monitoring software framework for the ATLAS level-1 calorimeter trigger* in *Topical Workshop on Electronics for Particle physics* (Nov. 2008), 441–445.

69. Nason, P. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP* **11,** 040. arXiv: hep-ph/0409146 (2004).

70. Frixione, S., Nason, P. & Oleari, C. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP* **11,** 070. arXiv: 0709.2092 [hep-ph] (2007).

71. Alioli, S., Nason, P., Oleari, C. & Re, E. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP* **06,** 043. arXiv: 1002.2581 [hep-ph] (2010).

72. Sjöstrand, T. *et al.* An introduction to PYTHIA 8.2. *Comput. Phys. Commun.* **191,** 159–177. arXiv: 1410.3012 [hep-ph] (2015).

73. ATLAS Collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP* **09,** 145. arXiv: 1406.3660 [hep-ex] (2014).

74. Stump, D. *et al.* Inclusive jet production, parton distributions, and the search for new physics. *JHEP* **10,** 046. arXiv: hep-ph/0303013 (2003).

75. Ball, R. D. *et al.* Parton distributions with LHC data. *Nuclear Physics B* **867,** 244–289. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/S0550321312005500 (2013).

76. Alwall, J. *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,** 079. arXiv: 1405.0301 [hep-ph] (2014).

77. ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data* tech. rep. (Nov. 2014). https://cds.cern.ch/record/1966419.

78. Lange, D. J. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462.** BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, 152–155. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900201000894 (2001).

79. ATLAS Collaboration. *ATLAS Tracking Software Tutorial* https://atlassoftwaredocs.web.cern.ch/trackingTutorial/index.html.

80. ATLAS Collaboration. *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment* tech. rep. (CERN, Geneva, July 2010). https://cds.cern.ch/record/1281344.

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

82. 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,** 578. arXiv: 2012.00578 [hep-ex] (2021).

83. ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C* **77,** 490. arXiv: 1603.02934 [hep-ex] (2017).

84. CMS Collaboration. *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging* tech. rep. (CERN, Geneva, July 2009). https://cds.cern.ch/record/1194489.

85. ATLAS Collaboration. *Boosted Object Tagging with Variable-R Jets in the ATLAS Detector* tech. rep. (CERN, Geneva, July 2016). https://cds.cern.ch/record/2199360.

86. ATLAS Collaboration. *Variable Radius, Exclusive-*k_T*, and Center-of-Mass Subjet Reconstruction for Higgs(*$\rightarrow b\bar{b}$*) Tagging in ATLAS* tech. rep. (CERN, Geneva, June 2017). https://cds.cern.ch/record/2268678.

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

88. ATLAS Collaboration. *Improving jet substructure performance in ATLAS using Track-CaloClusters* tech. rep. (Geneva, July 2017). https://cds.cern.ch/record/2275636.

89. ATLAS Collaboration. Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton-proton collisions. *Eur. Phys. J. C* **81,** 334. 47 p. arXiv: 2009.04986. https://cds.cern.ch/record/2730141 (Sept. 2020).

90. ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector* tech. rep. (May 2014). https://cds.cern.ch/record/1700870.

91. Krohn, D., Thaler, J. & Wang, L.-T. Jet Trimming. *JHEP* **02,** 084. arXiv: 0912.1342 [hep-ph] (2010).

92. ATLAS Collaboration. Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *JHEP* **03.** [Erratum: JHEP 11, 051 (2018)], 174. arXiv: 1712.06518 [hep-ex] (2018).

93. ATLAS Collaboration. *Expected performance of the ATLAS b-tagging algorithms in Run-2* tech. rep. (CERN, Geneva, July 2015). https://cds.cern.ch/record/2037697.

94. Lai, H.-L. *et al.* New parton distributions for collider physics. *Phys. Rev. D* **82,** 074024. https://link.aps.org/doi/10.1103/PhysRevD.82.074024 (7 Oct. 2010).

95. Baak, M., Gadatsch, S., Harrington, R. & Verkerke, W. Interpolation between multi-dimensional histograms using a new non-linear moment morphing method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **771,** 39–48. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900214011814 (2015).

96. Honle, A. *Searches for new resonances decaying to a Higgs and a vector boson with the ATLAS detector at the LHC* Presented 22 Jul 2021 (June 2021). https://cds.cern.ch/record/2779497.

97. Bothmann, E. *et al.* Event Generation with Sherpa 2.2. *SciPost Phys.* **7,** 34. https://scipost.org/10.21468/SciPostPhys.7.3.034 (3 2019).

98. Ball, R. D. *et al.* Parton distributions for the LHC Run II. *JHEP* **04,** 040. arXiv: 1410.8849 [hep-ph] (2015).

99. Anastasiou, C., Dixon, L., Melnikov, K. & Petriello, F. High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order. *Phys. Rev. D* **69,** 094008. https://link.aps.org/doi/10.1103/PhysRevD.69.094008 (9 May 2004).

100. Beneke, M., Falgari, P., Klein, S. & Schwinn, C. Hadronic top-quark pair production with NNLL threshold resummation. *Nuclear Physics B* **855,** 695–741. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/S0550321311005803 (2012).

101. Cacciari, M., Czakon, M., Mangano, M., Mitov, A. & Nason, P. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Physics Letters B* **710,** 612–622. ISSN: 0370-2693. https://www.sciencedirect.com/science/article/pii/S0370269312002766 (2012).

102. Czakon, M. & Mitov, A. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications* **185,** 2930–2938. ISSN: 0010-4655. https://www.sciencedirect.com/science/article/pii/S0010465514002264 (2014).

103. Czakon, M., Fiedler, P. & Mitov, A. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$. *Phys. Rev. Lett.* **110,** 252004. https://link.aps.org/doi/10.1103/PhysRevLett.110.252004 (25 June 2013).

104. Czakon, M. & Mitov, A. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. *JHEP* **01,** 080. arXiv: 1210.6832 [hep-ph] (2013).

105. Bärnreuther, P., Czakon, M. & Mitov, A. Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t}+X$. *Phys. Rev. Lett.* **109,** 132001. https://link.aps.org/doi/10.1103/PhysRevLett.109.132001 (13 Sept. 2012).

106. Frixione, S., Laenen, E., Motylinski, P., Webber, B. R. & White, C. D. Single-top hadroproduction in association with a W boson. *JHEP* **07,** 029. arXiv: 0805.3067 [hep-ph] (2008).

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

108. Bukin, A. D. *Fitting function for asymmetric peaks* 2007. https://arxiv.org/abs/0711.4449.

109. Cowan, G. & Gross, E. *Discovery significance with statistical uncertainty in the background estimate* 2008. https://www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf.

110. Campbell, J. M. & Ellis, R. MCFM for the Tevatron and the LHC. *Nuclear Physics B - Proceedings Supplements* **205-206.** Loops and Legs in Quantum Field Theory, 10–15. ISSN: 0920-5632. https://www.sciencedirect.com/science/article/pii/S0920563210001945 (2010).

111. Dulat, S. *et al.* New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D* **93,** 033006. https://link.aps.org/doi/10.1103/PhysRevD.93.033006 (3 Feb. 2016).

112. Harland-Lang, L. A., Martin, A. D., Motylinski, P. & Thorne, R. S. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J. C* **75,** 204. arXiv: 1412.3989 [hep-ph] (2015).

113. Bellm, J. *et al.* Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C* **76,** 196. arXiv: 1512.01178 [hep-ph] (2016).

114. Gao, J. *et al.* CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev. D* **89,** 033009. arXiv: 1302.6246 [hep-ph] (2014).

115. Martin, A. D., Stirling, W. J., Thorne, R. S. & Watt, G. Parton distributions for the LHC. *Eur. Phys. J. C* **63,** 189–285. arXiv: 0901.0002 [hep-ph] (2009).

116. Cranmer, K. S. Kernel estimation in high-energy physics. *Comput. Phys. Commun.* **136,** 198–207. arXiv: hep-ex/0011057 (2001).

117. ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C* **79,** 970. arXiv: 1907.05120 [hep-ex] (2019).

118. ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV* tech. rep. (CERN, Geneva, 2018). https://cds.cern.ch/record/2314418.

119. ATLAS Collaboration. *Measurement of b-tagging Efficiency of c-jets in* $t\bar{t}$ *Events Using a Likelihood Approach with the ATLAS Detector* tech. rep. (CERN, Geneva, 2018). https://cds.cern.ch/record/2306649.

120. ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C* **79,** 639. arXiv: 1902.04655 [physics.ins-det] (2019).

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

122. ATLAS Collaboration. Luminosity determination in *pp* collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC (June 2019).

123. ATLAS Collaboration. *Identification of boosted, hadronically-decaying W and Z bosons in* $\sqrt{s} = 13$ *TeV Monte Carlo Simulations for ATLAS* tech. rep. (CERN, Geneva, Aug. 2015). https://cds.cern.ch/record/2041461.

124. ATLAS Collaboration. *Jet mass reconstruction with the ATLAS Detector in early Run 2 data* tech. rep. (CERN, Geneva, July 2016). https://cds.cern.ch/record/2200211.

125. 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,** 072002. arXiv: 1703.09665 [hep-ex] (2017).

126. ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *Eur. Phys. J. C* **73,** 2306. arXiv: 1210.6210 [hep-ex] (2013).

127. ATLAS Collaboration. E_T^{miss} *performance in the ATLAS detector using 2015-2016 LHC p-p collisions* tech. rep. (CERN, Geneva, June 2018). https://cds.cern.ch/record/2625233.

128. 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,** 903. arXiv: 1802.08168 [hep-ex] (2018).

129. ATLAS Collaboration. *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV* tech. rep. (CERN, Geneva, 2015). https://cds.cern.ch/record/2037700.

130. ATLAS Collaboration. *Jet energy measurement and systematic uncertainties using tracks for jets and for b-quark jets produced in proton-proton collisions at* $\sqrt{s} = 7$ *TeV in the ATLAS detector* tech. rep. (CERN, Geneva, 2013). http://cds.cern.ch/record/1504739.

131. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C* **71,** 1554. arXiv: 1007.1727 [physics.data-an] (2011).

132. Wilks, S. S. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.* **9,** 60–62 (1938).

133. Barlow, R. Extended maximum likelihood. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **297,** 496–506. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/0168900290913348 (1990).

134. Read, A. L. Presentation of search results: The CL(s) technique. *J. Phys. G* **28** (eds Whalley, M. R. & Lyons, L.) 2693–2704 (2002).

135. Cramér, H. in *Mathematical Methods of Statistics (PMS-9), Volume 9* (Princeton university press, 2016).

136. Rao, C. "Information and the accuracy attainable in the estimation of statistical parameters". Bulletin of the Calcutta Mathematical Society (1945).

137. ATLAS Collaboration. *Search for heavy resonances decaying into a W boson and a Higgs boson in final states with leptons and b-jets in* 139 fb^{-1} *of pp collisions at* $\sqrt{s} = 13$ TeV *with the ATLAS detector* tech. rep. (CERN, Geneva, June 2021). http://cds.cern.ch/record/2773302.

138. ATLAS Collaboration. *Object-based missing transverse momentum significance in the ATLAS detector* tech. rep. (CERN, Geneva, July 2018). https://cds.cern.ch/record/2630948.

139. ATLAS Collaboration. Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell bb$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett. B* **783,** 392–414. arXiv: 1804.01126 [hep-ex] (2018).

140. Verkerke, W. & Kirkby, D. P. The RooFit toolkit for data modeling. *eConf* **C0303241** (eds Lyons, L. & Karagoz, M.) MOLT007. arXiv: physics/0306116 (2003).

141. Harlander, R. V., Liebler, S. & Mantler, H. SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM. *Comput. Phys. Commun.* **184,** 1605–1617. arXiv: 1212.3249 [hep-ph] (2013).

142. Eriksson, D., Rathsman, J. & Stal, O. 2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual. *Comput. Phys. Commun.* **181,** 189–205. arXiv: 0902.0851 [hep-ph] (2010).

143. James Witten, H. & Tibshirani. *An Introduction to Statistical Learning with Applications in R Second Edition* (Springer, 2021).

144. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition* (O'Reilly, 2019).

145. LeNail, A. NN-SVG: Publication-Ready Neural Network Architecture Schematics. *Journal of Open Source Software* **4,** 747. https://doi.org/10.21105/joss.00747 (2019).

146. Robbins, H. E. A Stochastic Approximation Method. *Annals of Mathematical Statistics* **22,** 400–407 (2007).

147. Baldi, P., Cranmer, K., Faucett, T., Sadowski, P. & Whiteson, D. Parameterized neural networks for high-energy physics. *Eur. Phys. J. C* **76,** 235. arXiv: 1601.07913 [hep-ex] (2016).

148. Fukushima, K. Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics* **20,** 121–136 (1975).

149. Paszke, A. *et al.* Automatic differentiation in PyTorch (2017).

150. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).

151. Lundberg, S. M. & Lee, S.-I. *A Unified Approach to Interpreting Model Predictions* in *Advances in Neural Information Processing Systems* (eds Guyon, I. *et al.*) **30** (Curran Associates, Inc., 2017). https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

152. Mangano, M. *et al. FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. Future Circular Collider* tech. rep. (CERN, Geneva, Dec. 2018). https://cds.cern.ch/record/2651294.

153. Ahmad, M. *et al.* CEPC-SPPC Preliminary Conceptual Design Report. 1. Physics and Detector (Mar. 2015).

List of Figures

List of Tables