The Search for di-Higgs Production in the $b\bar{b}b\bar{b}$ Final State with the ATLAS Detector

Yuwen Zhang

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics and Astronomy

University College London

February 5, 2025

I, Yuwen Zhang, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This thesis focuses on the search of non-resonant Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state at the LHC with the ATLAS detector.

The main published result uses 126 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV, targeting both the gluon-gluon fusion (ggF) and vector-boson fusion (VBF) production modes. No evidence of HH production is observed, with the observed (expected) upper limit on the cross-section being 5.4 (8.1) times the Standard Model(SM) prediction at 95% confidence level. The observed (expected) 2σ constraints on the HHH coupling modifier, κ_λ are [-3.9, 11.1]([-4.9, 10.8]), and that for $HHVV$ coupling modifier κ_{2V} are [-0.03,2.11]([-0.05,2.12]). Constraints on relevant coefficients are derived for the Standard Model effective field theory and the Higgs effective Higgs theory. A promising study of training a classifier on the background estimates versus simulated signal is also presented.

Several technical works about triggers are presented: improving track finding efficiency for b-jet signature for Inner Detector Triggers; and a brief investigation of triggers to be used in Run 3 analysis.

In addition to the ATLAS analysis, I also worked on a simulation study for the Circular Electron Positron Collider(CEPC). The study simulates the search for exotic decays of the Higgs boson into a pair of spin-zero particles considering the scenario of analysing 5000 fb^{-1} e^+e^- collision data at $\sqrt{s} = 240$ GeV from CEPC, showing promising potential for this future collider - the sensitivity can be significantly larger than that can be achieved at the LHC.

Impact Statement

The Large Hadron Collider has been providing fruitful insights into the fundamental particles that form the building blocks of our universe since it started operating in 2008. Multiple decades of research and development effort have been put into it, with hundreds of research groups, to engineer this giant machine and develop required data filtering, storage and analysis techniques. The ATLAS experiment plays an important role in completing and refining our understanding of the physical processes predicted by the Standard Model, as well as finding evidence of beyond Standard Model theories.

The search covered in this thesis is about the production of two Higgs bosons, a rare but very significant process that aims to shed light on one of the fundamental processes involving the Higgs boson. The use of advanced pattern recognition, and the data-driven background modelling technique, significantly improves the sensitivity compared to previous results. In addition, the novel technique of using a neural network-based classifier to discriminate signal from background explored in this thesis delivers promising potential. These methods can be adopted across future analyses.

The simulation study for the future Circular Electron Positron Collider (CEPC) demonstrates the potential for breakthroughs in experimental physics with its much higher projected luminosity and unprecedented sensitivity.

Beyond academia, the software engineering and machine learning techniques developed in ATLAS in order to process such a huge amount of collision data can benefit sectors outside physics, such as biology, technology and finance. On a social level, the ATLAS collaboration shows the power of international scientific coop-

eration. Its effort to advertise its findings through public engagement, educational outreach, and media will help stimulate interest in fundamental research, inspiring the next generation of scientists.

Acknowledgements

These four years being part of the UCL HEP group and ATLAS Collaboration has been a pleasant journey.

First and foremost, I would like to thank my supervisor, Nikos Konstantinidis, for his guidance and support throughout my PhD. His help has been invaluable, especially during the challenging times of starting my PhD in the middle of the pandemic. Without his support, along with that of the UCL staff, this period will be even more difficult.

Then, I would like to thank our analysis group at UCL - Liaoshan Shi, Thandi Madula, Noah Clarke Hall, Marin Mlinarevic, and Lucas Borgna (who has already graduated). I received a great deal of help from them, and it has truly been a pleasure working with them.

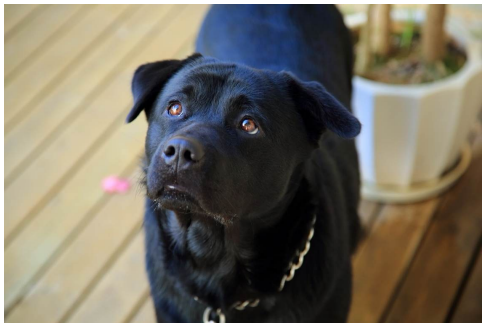
Thank you also to other ATLAS people I had the pleasure working with - especially Nicole Michelle Hartman, who provided invaluable help to my work and actively promoted the classifier study with the $HH4b$ group; and my technical supervisor Mark Sutton, who guided me throughout my qualification task.

Special thanks to the individuals with whom I collaborated on the classifier study. This project was initially started by Xudong Ke Lin during his summer internship with our group. I am also grateful to Raj Ramani, another summer student, for his significant contributions to the feature selection process. Finally, I would like to thank Thandi, Noah, Nicole, and Nikos for their active involvement and all the helpful insights they provided throughout the study.

Many thanks to all my friends, especially my bestie, Kaiwen Kang, for always encouraging me and providing emotional support. Wishing you the best of luck with

your career!

Finally, I would like to thank my family (and my lovely dogs Hu(1a) and Bei(1b)), and my loved one, Jimin Li, for their unwavering love and support.



(a) Hu (2012-2022). He sadly passed away on 28/07/2022. I missed him so much.

(b) Bei. My cutest little princess. She joined our family in early 2023.

Figure 1: My dogs

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The discovery of the 125 GeV Higgs boson [23, 24, 25, 26] and the measurements of its properties provide a strong indication that the mechanism of electroweak symmetry breaking (EWSB) is compatible with the prediction by Brout, Englert and Higgs (BEH) [27, 28, 29]. However, there are still many unprobed features, such as the form of the BEH potential. One of the parameters that determines the form of the BEH potential is the Higgs boson's trilinear self-coupling λ, which can be expressed in terms of its modifier κ_λ[1] and measured via Higgs-boson pair (HH) production. Another important parameter κ_{2V} represents the quartic coupling between two Higgs bosons and two vector bosons (W or Z). These processes can also provide insights into the search for new physics. It is a very challenging measurement since the Standard Model (SM) HH production cross-section is more than 1000 times lower than that of single Higgs production.

This thesis focuses on the $HH \to b\bar{b}b\bar{b}$ channel, which has the highest branching ratio ($\approx$33%) of all Higgs decay combinations and gives the largest potential signal yields. However, it has large QCD multijet background which needs to be suppressed. Prior to this thesis, the ATLAS Collaboration has published search results for non-resonant $HH \to b\bar{b}b\bar{b}$ production[2] using 27 fb^{-1} of early Run 2 data [30], as well as a dedicated search for vector-boson fusion HH production in 126 fb^{-1} of data

[1]The coupling modifier κ is defined as the ratio of the modified coupling to its SM value. For example, for the trilinear self-coupling λ, $\kappa_\lambda = \frac{\lambda}{\lambda^{\mathrm{SM}}}$. Then by definition, $\kappa = 1$ is the SM prediction. This will be detailed more in Chapter 4

[2]Non-resonant: the production of the two Higgs boson is directly from SM processes, as opposed to having a hypothetical heavy particle (resonance) decays into a pair of Higgs bosons.

collected between 2016 and 2018 [31]. Similarly, the CMS Collaboration has published results of a search for non-resonant $HH \to b\bar{b}b\bar{b}$ with their full Run 2 dataset [32], reporting an observed (expected) upper limit on the HH cross-section of 3.9 (7.8) times the SM prediction, and constraining the allowed interval for κ_λ to [-2.3, 9.4] ([-5.0, 12.0]), both at 95% confidence level (CL). A more recent CMS publication on $HH \to b\bar{b}b\bar{b}$ [33], which exploits the topologies arising from highly energetic Higgs boson decays into $b\bar{b}$, sets the observed (expected) upper limit at 9.9 (5.1) times the SM cross-section, and restricts the κ_{2V} interval to [0.62, 1.41] ([0.66, 1.37]), at 95% CL.

Other searches for non-resonant HH production were performed by ATLAS and CMS in the $b\bar{b}\tau^+\tau^-$ [34, 35] (moderately high branching ratio of ≈7.5% but has cleaner signature than $4b$), $b\bar{b}\gamma\gamma$ [36, 37] (very clean final state and low background contributions, but small branching ratio of ≈0.26%), $b\bar{b}l^+\nu l^-\nu$ [38, 39] (leptons provide a clean signature and are less affected by QCD backgrounds, but has low branching ratio (≈1.9%) and large $t\bar{t}$ background), multilepton [40, 41] (low background levels and sensitive to non-SM production modes, but very low branching ratio and requires complex event selection) decay channels, as well as by ATLAS in the $b\bar{b}qql\nu$ [42] (moderate branching ratio combining hadronic and leptonic Higgs decays, but with moderate to high backgrounds and complex final state), $WW^*\gamma\gamma$ [43] (clean final states but very small branching ratio), WW^*WW^* [44] (provides a unique probe of BSM Higgs coupling affecting W-boson interactions but very small branching ratio) and $b\bar{b}ll + E_T^{\text{miss}}$ [45] (combines the clean leptonic signature with the high branching ratio of $b\bar{b}$ to give a moderate branching ratio ≈1.6% and the missing energy allows suppression of QCD backgrounds, but there are other challenging backgrounds such as $t\bar{t}$) decay channels.

The most sensitive ATLAS result to date comes from a combination of results from $b\bar{b}b\bar{b}$ (combination of the latest ATLAS non-resonant $HH \to b\bar{b}b\bar{b}$ result [4] using full Run 2 data, which is the main focus of this thesis and will be detailed in Chapter 4; and the boosted vector-boson fusion result [46]), $b\bar{b}\tau^+\tau^-$, $b\bar{b}\gamma\gamma$, multilepton and $b\bar{b}ll + E_T^{\text{miss}}$ decay channels [47], probing more than half of the HH

decays, giving the observed (expected) 95% CL upper limit on the SM non-resonant HH cross-section at 2.9 (2.4) times the SM prediction and restricts the corresponding observed (expected) κ_λ interval to [-1.2, 7.2] ([-1.6, 7.2]), and κ_{2V} interval to [0.6, 1.5] ([0.4, 1.6]). The most sensitive CMS result to date comes from the combination of the $b\bar{b}ZZ$ [48], multilepton, $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ and $b\bar{b}b\bar{b}$ analyses [49], which set the observed (expected) 95% CL upper limit on the SM non-resonant HH cross-section at 3.4 (2.5) times the SM prediction and restricts the corresponding observed κ_λ interval to [-1.24, 6.49], and κ_{2V} interval to [0.67, 1.38].

The thesis is structured as follows. Chapter 2 provides an overview of the theoretical background, introducing the Standard Model and two relevant beyond Standard Model theories. Chapter 3 focuses on the ATLAS detector, its Trigger and Data Acquisition (TDAQ) system, and my ATLAS authorship qualification task - improving track finding efficiency for b-jet signature. The core of this thesis is presented in Chapter 4, which details our published results of the $HH \to b\bar{b}b\bar{b}$ analysis based on Run 2 dataset targeting both the gluon-gluon fusion and vector-boson fusion production modes. This analysis uses the full 2016-2018 data for both production channels, and also benefits from advancements in jet reconstruction and in the identification of jets arising from the hadronisation of b-quarks, made by the ATLAS Collaboration since the publication of [30].

The remaining chapters present promising future directions for $HH4b$ analysis, and high-energy experimental physics. Chapter 5 presents a study of training a Neural Network based classifier on the estimated background versus Monte-Carlo simulated signal. A study of triggers to be used in Run 3 is discussed in Appendix D, and a simulation study of an exotic Higgs decay process at the future Circular Electron Positron Collider (CEPC) is discussed in Appendix E.

Chapter 2

The Standard Model and Beyond

In this chapter, I will briefly outline important points of the Standard Model, focusing on the Higgs Boson, and the beyond standard model (BSM) theories relevant to the analysis.

2.1 The Standard Model

The Standard Model of Particle Physics [50, 51] describes the elementary particles and their fundamental interactions, with the exception of gravity. It is widely regarded as the most successful application of quantum field theory in terms of experimental verification.

The matter particles are elementary spin-$\frac{1}{2}$ fermions, divided into two fundamental types: quarks and leptons. Each group consists of six particles, organised in pairs and classified into three generations, as summarised in Table 2.1. The first generation includes the lightest and most stable particles, which form all stable matter in the universe. The heavier, less stable particles belong to the second and third generations and decay quickly into more stable forms. There are three generations of quarks, with each generation containing a pair of ”up-type” and ”down-type” quarks. Quarks carry colour charge and can only combine in ways that result in colourless objects. Together with gluons, quarks form atomic nuclei, and they participate in both electroweak and strong interactions. Particles composed of quarks are called hadrons, which are further classified as either baryons (made of three quarks, e.g. protons and neutrons) or mesons (a quark-antiquark pair). Each generation of leptons

consists of an electrically charged particle paired with a neutral "neutrino". Since leptons are colour-neutral, they do not participate in strong interactions. As a result, the electrically neutral neutrinos interact only via the weak force, making them particularly difficult to detect.

Fermions	Generation	I		II		III	
Quarks	mass[GeV]	u	0.00216	c	1.273	t	172.57
	charge	up	2/3	charm	2/3	top	2/3
	spin		1/2		1/2		1/2
	mass[GeV]	d	0.0047	s	0.0935	b	4.183
	charge	down	-1/3	strange	-1/3	bottom	-1/3
	spin		1/2		1/2		1/2
Leptons	mass[MeV]	e	0.511	μ	105.66	τ	1777
	charge	electron	-1	muon	-1	tau	-1
	spin		1/2		1/2		1/2
	mass[MeV]	ν_e	$< 2.2 \times 10^{-6}$	ν_μ	< 0.17	ν_τ	< 15.5
	charge	electron	0	muon	0	tau	0
	spin	nutrino	1/2	neutrino	1/2	neutrino	1/2

Table 2.1: The Standard Model fermions [19, 20]

The dynamics of the fermions are described by the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{2.1}$$

for γ the Dirac γ-matrices, and ψ the wave function. Writing $\phi(\vec{x},t) = u(E,\vec{p})e^{i(\vec{p}\cdot\vec{x}-Et)}$, for $u(E,\vec{p})$ a four-component Dirac spinor, we will find four independent solutions to the Dirac equation, with two of them having negative energies. The negative energy solutions are interpreted as positive energy particles propagating backwards in time - the antiparticles. As a result, every fermion has a corresponding antiparticle, which shares the same mass but has opposite electric charge and other quantum numbers. [52, 53]

The Standard Model describes the physics of three fundamental forces except gravity, each mediated by spin-1 gauge bosons. The electromagnetic force is described as interactions between electrically charged particles via photon exchange, as formulated in quantum electrodynamics (QED). Since photons are massless, electromagnetic interactions have an infinite range. The weak force, responsible for

certain types of radioactive decay, is mediated by the massive W and Z bosons, with their large mass limiting the range of the weak interaction and making it relatively weak. The weak charged-current interaction, which is responsible for nuclear β-decay and nuclear fusion, is mediated by the charged W^+ and W^- bosons. The weak neutral-current interaction closely related to the charged current is mediated by the electrically neutral Z^0 boson. The unification of the electromagnetic and weak forces is the electroweak theory [54, 55]. The strong force, which binds quarks into nucleons and nucleons into nuclei, is mediated by gluons and is described by quantum chromodynamics (QCD) [56]. Furthermore, the Brout-Englert-Higgs (BEH) mechanism [27, 28, 29, 57, 58] gives a mass to the W and Z bosons when they interact with the Higgs field, with the associated particle being the spin-0 Higgs boson. They are listed in Table 2.2.

Boson	mass[GeV]	charge	spin
g gluon	0	0	1
γ photon	0	0	1
Z^0 Z boson	91.188	0	1
$W^\pm$ $W^\pm$ boson	80.369	± 1	1
h Higgs boson	125.2	0	0

Table 2.2: The Standard Model bosons [19]

2.1.1 Quantum Electrodynamics (QED)

The Lagrangian for Maxwell's equation is

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.2}$$

For the field strength tensor $F_{\mu\nu}$ defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.3}$$

For A_μ the electromagnetic field. This Lagrangian is invariant under $A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x)$. We want to couple it to some matter field by adding a term representing the interaction between the matter and A_μ

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu \tag{2.4}$$

For j^μ the electromagnetic current density four-vector, satisfying $\partial_\mu j^\mu = 0$. It describes the sources of the electromagnetic field, with J^0 representing the charge density and $\mathbf{J}$ representing the current density. The term $-J^\mu A_\mu$ describes the interaction between the electromagnetic field and the sources.

To couple it to fermions, we have the Dirac Lagrangian

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

Which has an internal symmetry $\psi \to e^{i\alpha}\psi$ giving rise to the conserved current $j^\mu = \bar{\psi}\gamma^\mu \psi$. Hence we can write the Lagrangian of electromagnetism coupled to fermions with $U(1)$ gauge symmetry as

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{2.6}$$

where $D_\mu \psi = \partial_\mu \psi + ieA_\mu \psi$ is the covariant derivative. This Lagrangian is invariant under gauge transformation which acts as:

$$A_\mu \to A_\mu + \partial_\mu \lambda, \quad \psi \to e^{-ie\lambda}\psi \tag{2.7}$$

for any $\lambda(x)$. The resulting theory, known as quantum electrodynamics (QED) [59], is the first building block of the Standard Model.

2.1.2 Yang-Mills Theory

Yang-Mills theory is a generalisation of Maxwell theory in which $U(1)$ is replaced by Lie algebra G. For each element in the algebra, we introduce a gauge field A^A_μ,

referred to as gauge bosons, for $A = 1, \ldots, \dim(G)$, and the gauge potential

$$A_\mu = A_\mu^A T^A \tag{2.8}$$

For T^A the generators of the Lie algebra. The gauge fields are invariant under gauge transformations $\Omega(x) \in G$, $\Omega(x) = e^{i\theta(x)^A T^A}$

$$A_\mu \to \Omega A_\mu \Omega^{-1} + \frac{i}{g}\Omega \partial_\mu \Omega^{-1} \tag{2.9}$$

$$A_\mu^A \to A_\mu^A + \frac{1}{g}\partial_\mu \theta^A - f^{ABC}\theta^B A_\mu^C \tag{2.10}$$

for some coupling constant g. The covariant derivative satisfying this is

$$D_\mu = I\partial_\mu - igT^A A_\mu^A \tag{2.11}$$

and satisfies $[D_\mu, D_\nu] = -igT^A F_{\mu\nu}^A$

The field strength can be generally defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \tag{2.12}$$

Similarly to A_μ, the field strength can also be expanded as $F_{\mu\nu} = F_{\mu\nu}^A T^A$, and it transforms under gauge transformations as

$$F_{\mu\nu}^A \to F_{\mu\nu}^A - f^{ABC}\theta^B F_{\mu\nu}^C \tag{2.13}$$

and can be written as

$$F_{\mu\nu}^A = \partial_\mu A_\nu^A - \partial_\nu A_\mu^A + g f^{ABC} A_\mu^B A_\nu^C \tag{2.14}$$

The Lagrangian can therefore be written as

$$\mathcal{L} = \underbrace{-\frac{1}{4}F_{\mu\nu}^A F^{A,\mu\nu}}_{\mathcal{L}_{\text{YM}}} + \underbrace{\sum_f \bar{\psi}_f (i\gamma^\mu D_\mu - m_f)\psi_f}_{\mathcal{L}_{\text{matter}}} \tag{2.15}$$

for f the counts of the matters.

2.1.3 Weak Interactions and Electroweak Unification

Weak interaction has a more complex chiral $SU(2)_L$ gauge symmetry. The weak force only acts on left-handed fermions or right-handed antifermions. Under electroweak theory, the symmetry is $G = SU(2)_L \times U(1)_Y$, where $SU(2)_L$ weak isospin group is used to construct the weak current which only couples to left-handed fermions with generators T^A; and $U(1)_Y$ is the group of weak hypercharge Y related to the abelian QED $U(1)_Q$ by $Q = T_3 + \frac{1}{2}Y$, where Q is the electric charge and I_3 is the third component of weak isospin. The general group elements $\Omega \in G$ is $\Omega = e^{i\alpha^A T^A} e^{i\beta Y}$.

The gauge fields are

- W_μ^A for $SU(2)_L$ transforms as $\delta W_\mu^A = \frac{1}{g}\partial_\mu \alpha^A - \varepsilon^{ABC}\alpha^B W_\mu^C$ under infinitesimal gauge transformations, for g a gauge coupling. The field strength is

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - ig[W_\mu, W_\nu] \tag{2.16}$$

$$W_{\mu\nu}^A = \partial_\mu W_\nu - \partial_\nu W_\mu + g\varepsilon^{ABC} W_\mu^B W_\nu^C \tag{2.17}$$

- B_μ for $U(1)_Y$ transforms as $\delta B_\mu = \frac{1}{g'}\partial_\mu \beta$ for gauge coupling constant g'. The field strength is

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.18}$$

The fermions are grouped into left-handed weak isospin doublets and right-handed weak isospin singlets (projected into left and right-handed parts using $P_L, R = \frac{1}{2}(1 \mp \gamma^5)$)

$$Q_L^i = \left\{ \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right\}_{Y=1/6} \tag{2.19}$$

$$L_L^i = \left\{ \begin{pmatrix} \nu_L^e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_L^\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_L^\tau \\ \tau_L \end{pmatrix} \right\}_{Y=-1/2} \tag{2.20}$$

and

$$u_R^i(x) = \{u_R, c_R, t_R\}_{Y=2/3} \tag{2.21}$$

$$d_R^i(x) = \{d_R, s_R, b_R\}_{Y=-1/3} \tag{2.22}$$

$$e_R^i(x) = \{e_R, \mu_R, \tau_R\}_{Y=1} \tag{2.23}$$

$$\nu_R^i(x) = \{\nu_R^e, \nu_R^\mu, \nu_R^\tau\}_{Y=0} \tag{2.24}$$

for $i = 1,2,3$ the flavour index.

The gauge part of the Lagrangian is

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^A W^{A,\mu\nu} \tag{2.25}$$

The interaction term of the fermions is

$$\begin{aligned}\mathcal{L}_{\text{fermi}} = -i\sum_{i=1}^{3}(&\bar{Q}_L^i\gamma^\mu D_\mu Q_L^i + \bar{L}_L^i\gamma^\mu D_\mu L_L^i + \bar{u}_R^i\gamma^\mu D_\mu u_R^i \\ &+\bar{d}_R^i\gamma^\mu D_\mu d_R^i + \bar{e}_R^i\gamma^\mu D_\mu e_R^i + \bar{\nu}_R^i\gamma^\mu D_\mu \nu_R^i)\end{aligned} \tag{2.26}$$

Where D_μ is the covariant derivative

$$D_\mu = \partial_\mu - igW_\mu^A T^A - \frac{i}{2}g'B_\mu Y \tag{2.27}$$

2.1.4 Quantum Chromodynamics

The strong force affects the particles with colour change, similar to the electric charge, we have three colour charges RGB - red, green and blue. For strong interactions, the gauge symmetry group is $SU(3)$. The gauge field of the strong interactions is denoted as G_μ, a 3×3 Hermitian matrix. G_μ can be decomposed into gluon fields with $G_\mu = G_\mu^A T^A$ with $A = 1,\ldots,8$, where T^A are generators of $SU(3)$. A common choice of the generators are

$$T_{ij}^A = \frac{1}{2}\lambda_{ij}^A \tag{2.28}$$

for λ^A the Gell-Mann matrices [60]. The field strength is defined as

$$G^A_{\mu\nu} = \partial_\mu G^A_\nu - \partial_\nu G^A_\mu + g_s f^{ABC} G^B_\mu G^C_\nu \tag{2.29}$$

where g_s is the strong coupling constant. We also have the covariant derivative

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} - i g_s T^A_{ij} G^A_\mu \tag{2.30}$$

There is an extra Yang-Mills theta term in the Lagrangian

$$\frac{\theta g_s^2}{16\pi^2} G^A_{\mu\nu} \tilde{G}^{A\mu\nu} \tag{2.31}$$

where $\tilde{G}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma} G_{\rho\sigma}$, and θ a real parameter called theta angle, and taking a very small number $\theta < 10^{-10}$. Then the Lagrangian can be written as

$$\mathcal{L} = -\frac{1}{4} G^A_{\mu\nu} G^{A\mu\nu} + \sum_i \bar{q}_i (i\gamma^\mu D_\mu - m_f) q_i + \frac{\theta g_s^2}{16\pi^2} G^A_{\mu\nu} \tilde{G}^{A,\mu\nu} \tag{2.32}$$

Here q_i represented the quarks, for $i = 1, \ldots, 6$ the flavour indices.

The quarks are never observed freely, this is explained by the hypothesis of colour confinement [61], which states that coloured objects are always confined to colour singlet states and that no objects with non-zero colour charge can propagate as free particles. This places strong restrictions on the possible combinations of quarks that can form bound hadronic states. The observed colour singlets take the form

- $q_i \bar{q}_i$: mesons
- $\varepsilon_{ijk} q_i q_j q_k$: baryons

2.1.5 The Higgs Mechanism

For a gauge boson A_μ, a mass term in the Lagrangian would look like $\frac{1}{2} m^2 A_\mu A^\mu$, which is not invariant under the $SU(2)_L \times U(1)_Y$ gauge transformations. However, we know through experiments that gluon and photons are massless, but W and Z bosons are massive. This was solved in 1964 almost simultaneously by three

independent groups: by Robert Broug and Francois Englert [27]; by Peter Higgs [28]; and by Gerald Guralnik, C. R. Hagen, and Tom Kibble [29]. What we now call the Brout-Englert-Higgs (BEH) mechanism gives a mass to the W and Z bosons when they interact with the Higgs field through spontaneous symmetry breaking.

We first consider the spontaneous symmetry breaking of a complex scalar field with a potential

$$V(\phi) = \mu^2\phi^2 + \lambda\phi^4 \tag{2.33}$$

and we consider only the $U(1)$ local gauge symmetry for now. The combined Lagrangian is

$$\mathcal{L} = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - \mu^2\phi^2 - \lambda\phi^4 \tag{2.34}$$

Where D_μ is the covariant derivative $D_\mu = \partial_\mu + igB_\mu$ for the gauge field B_μ.

When $\lambda > 0$, the Lagrangian is invariant under $U(1)$ transformations. When $\mu^2 > 0$, the minimum is at 0. When $\mu^2 < 0$, the potential has an infinite set of minima at $|\phi| = v^2$ for $v = \sqrt{-\frac{\mu^2}{\lambda}}$. In this case, the lowest energy state is no longer at $\phi = 0$, the field is said to have a non-zero **vacuum expectation value** v. The physical vacuum state will point to a specific direction, breaking the global $U(1)$ symmetry - this is the process of **spontaneous symmetry breaking**. We can choose the vacuum state to be in the real direction, then $\phi = \frac{1}{\sqrt{2}}(\eta + v + i\xi)$. Substituting it into the Lagrangian leads to

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2}_{\text{massive } \eta} + \underbrace{\frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi)}_{\text{massless } \xi} \underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2v^2B_\mu B^\mu}_{\text{massive gauge field}}$$
$$- V_{int} + gvB_\mu(\partial^\mu\xi) \tag{2.35}$$

where $V_{int}(\eta,\xi,A)$ is the three and four-point interaction terms of the fields η,ξ and A. The breaking of the symmetry of the Lagrangian produces a massive scalar field η, and a massless Goldstone boson ξ. The gauge field B has also acquired a mass term $\frac{1}{2}g^2v^2B_\mu B^\mu$.

The goldstone boson can be eaten by the gauge boson with appropriate gauge transformations

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) + \frac{1}{gv}\partial_\mu \xi(x) \tag{2.36}$$

or correspondingly

$$\phi(x) \to \phi'(x) = e^{-i\xi(x)/v}\phi(x) \tag{2.37}$$

Then the Lagrangian becomes

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \lambda v^2 \eta^2}_{\text{massive } \eta} \underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2 v^2 B_\mu B^\mu}_{\text{massive gauge field}} + V_{int} \tag{2.38}$$

To the first order, $\phi(x) \approx \frac{1}{\sqrt{2}}(v+\eta(x))e^{i\xi(x)/v}$, then the gauge transformation gives $\phi'(x) = \frac{1}{\sqrt{2}}(v+\eta(x))$ - the Goldstone field ξ is eliminated. This is known as the **Unitary gauge**, corresponding to choosing the complex scaler field $\phi(x)$ to be entirely real

$$\phi(x) = \frac{1}{\sqrt{2}}(v+\eta(x)) \equiv \frac{1}{\sqrt{2}}(v+h(x)) \tag{2.39}$$

$\eta(x)$ has been written as the Higgs field h here. Now the Lagrangian becomes

$$\begin{aligned}\mathcal{L} = &\underbrace{\frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \lambda v^2 \eta^2}_{\text{massive } h \text{ scalar}} \underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2 v^2 B_\mu B^\mu}_{\text{massive gauge boson}} \\ &+ \underbrace{g^2 v B_\mu B^\mu h + \frac{1}{2}g^2 B_\mu B^\mu h^2}_{h,B \text{ interactions}} \underbrace{-\lambda v H^3 - \frac{1}{4}vh^4}_{h \text{ self-interactions}}\end{aligned} \tag{2.40}$$

This Lagrangian describes a massive scalar Higgs field h and a massive gauge boson A with the $U(1)$ local gauge symmetry.

In the Salam-Weinberg model [50, 51], the Higgs mechanism is embedded in the $SU(2)_L \times U(1)_Y$ gauge symmetry of the electroweak sector of the Standard Model. It consists of two complex scalar fields placed in a weak isospin doublet

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.41}$$

Here ϕ_0 is neutral, and ϕ^+ is charged with $(\phi^+)^* = \phi^-$. The potential is now

$$\mu^2 H^\dagger H + \lambda (H^\dagger H)^2 \tag{2.42}$$

After symmetry breaking, the neutral photon is required to remain massless, hence the minimum of the potential must correspond to a non-zero vacuum expectation value only of the neutral scalar field ϕ^0. We therefore pick the vacuum expectation value as

$$\langle H \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.43}$$

then the Higgs doublet can be written in unitary gauge as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.44}$$

Substituting H and the covariance derivative defined in 2.27 into the Lagrangian, we have a mass term $\lambda v^2 h^2$ for the **Higgs boson** h, experimentally measured to be $M_h = \sqrt{2\lambda v} = 125$ GeV.

Then the mass term for the gauge fields

$$\frac{1}{8}(v+h)^2 (g^2 (W^1_\mu W^{1,\mu} + W^2_\mu W^{2,\mu}) + (g W^3_\mu - g' B_\mu)^2) \tag{2.45}$$

giving masses for W^1 and W^2 to be $m_W = \frac{1}{2} g v$, experimentally measured to be approximately 80 GeV. The physical $W^\pm$ bosons are linear combinations

$$W^\pm = \frac{1}{\sqrt{2}} (W^1_\mu \mp i W^2_\mu) \tag{2.46}$$

to ensure $W^\pm$ has electric charge ± 1.

The remaining W^3 and B need to form the massless neutral boson A (photon)

and the massive neutral Z. This can be achieved by

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{2.47}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{2.48}$$

with the **weak mixing angle** θ_W defined as $\tan\theta_W = g'/g$. The Z boson mass is experimentally measured as $M_Z = \frac{v}{2}\sqrt{g^2+g'^2} \approx 91$ GeV.

The fermion masses are acquired through Yukawa couplings - the interactions between the Higgs field and the fermions. The Yukawa interaction for a single generation takes the form

$$\mathcal{L}_{\text{Yukawa}} = -y^d \bar{Q}_L H d_R - y^u \bar{Q}_L \tilde{H} u_R - y^e \bar{L}_L H e_R - y^\nu \bar{L}_L \tilde{H} \nu_R + \text{h.c.} \tag{2.49}$$

for ys the Yukawa coupling constants, $\tilde{H}$ the conjugated Higgs doublet defined as $\tilde{H}^a = \varepsilon^{ab} H_b^\dagger$ for $a,b = 1,2$. Expand Q and L as in Section 2.1.3, we have

$$\mathcal{L}_{\text{Yukawa}} = -\frac{(v+h)}{\sqrt{2}}(y^d \bar{d}_L d_R + y^u \bar{u}_L u_R + y^e \bar{e}_L e_R + y^\nu \bar{\nu}_L \nu_R) \tag{2.50}$$

which gives each of the fermions a mass $m^X = \frac{1}{\sqrt{2}} y^X v$, for $X = d, u, e, \nu$. It also gives the coupling of the fermions to the Higgs boson h with terms like $y^e h \bar{e} e = \sqrt{2} m^e / v$ - the coupling strength is proportional to the fermion masses.

2.1.6 The Standard Model Lagrangian

Now we have all the pieces, we can put together the Standard Model Lagrangian

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermi}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \tag{2.51}$$

where $\mathcal{L}_{\text{gauge}}$ is the gauge field kinetic terms including the $SU(3)$ gluon field strength $G_{\mu\nu}$, the $SU(2)$ gauge field strength $W_{\mu\nu}$, and the $U(1)$ hypercharge field stregnth $B_{\mu\nu}$

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{2}\text{Tr}\{W_{\mu\nu} W^{\mu\nu}\} - \frac{1}{2}\text{Tr}\{G_{\mu\nu} G^{\mu\nu}\} \tag{2.52}$$

and the kinetic terms of the fermions are defined in 2.26.

The Lagrangian of the Higgs term includes both its kinematic term and potential

$$\mathcal{L}_{\text{Higgs}} = D_\mu H^\dagger D^\mu H - \lambda(H^\dagger H - \frac{v^2}{2})^2 \tag{2.53}$$

with the covariance derivative $D_\mu H = \partial + \mu H - igW_\mu H - \frac{i}{2}g'B_\mu H$.

Finally the Yukawa terms $\mathcal{L}_{\text{Yukawa}}$ are defined in 2.49.

2.1.7 The Higgs Boson

Via the coupling term defined in 2.50, the Feynmann rule for the interaction vertex with a fermion of mass m_f with h is

$$-i\frac{m_f}{v} \equiv -i\frac{m_f}{2m_W}g_W \tag{2.54}$$

The Higgs boson can decay via $G \to f\bar{f}$ for all kinematically allowed decay modes with $m_H > 2m_f$. If it is sufficiently massive, it can also decay into W^+W^- or ZZ. However, since the coupling strength is proportional to the fermion mass, the largest branching ratios are to the more massive particles. For the SM Higgs boson of 125 GeV, the largest branching ratio is to bottom quarks via $b\bar{b}$, which is the focus of our analysis. The branching ratios of the main SM Higgs decay near $m_H = 125$ GeV are shown in Figure 2.1.

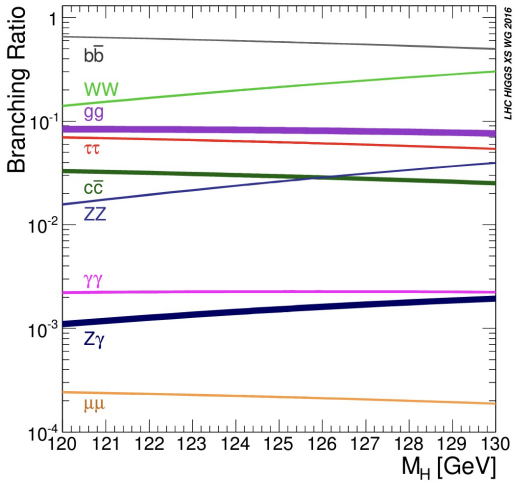
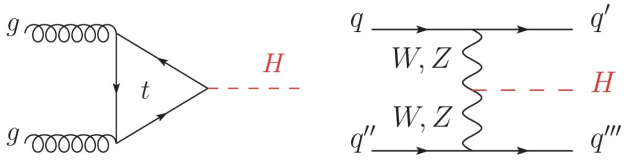


Figure 2.1: The branching ratios for the main decays of the SM Higgs boson near $m_H = 125$ GeV. The theoretical uncertainties are indicated as bands. [1]

At LHC, the main production modes of the Higgs boson are gluon-gluon fusion (ggF) and vector-boson fusion (VBF), as shown in figure 2.2. The ggF process is $gg \to H + X$, mediated by the exchange of a virtual, heavy top quark. In VBF, the process is $qq \to qqH$ - the scattering of two (anti-)quarks mediated by the exchange of a W or Z boson, with the Higgs boson radiated off the weak-boson propagator.



(a) ggF **(b)** VBF

Figure 2.2: Main leading order Feynmann digrams contributing to the single Higgs boson production in (a) gluon fusion, (b) Vector-boson-fusion. [1]

The ggF and VBF processes are also the main production modes of interest for our $HH \to b\bar{b}b\bar{b}$ analysis. This will be detailed more in Section 4.1.

2.2 Beyond the Standard Model

Despite the success of the Standard Model, several unresolved issues remain. The most prominent is the absence of gravity in the theory. Additionally, the Standard Model does not provide an explanation for Dark Matter, which constitutes approximately 85% of the universe's total matter. Another major issue is the "hierarchy problem", referring to the large discrepancy between aspects of the weak force and gravity - why the weak force is 10^{24} times stronger than gravity? These and other open questions suggest that the Standard Model is not the whole picture. Theories that extend the Standard Model in an effort to address these unresolved questions are referred to as "Beyond the Standard Model" (BSM) theories.

In the following, I will discuss the projection of the interactions of the known Standard Model states into an effective field theory (EFT) framework. In this setup, the experimental results can be projected into consistent field theories such as the SM, the SM supplemented with higher dimensional operators (the SMEFT), or an Electroweak chiral Lagrangian with a dominantly $J^P = 0^+$ scalar[1] (the HEFT). [62]

[1] J is the total angular momentum, and $J = 0$ means that the particle has no intrinsic spin, i.e.

2.2.1 Standard Model Effective Field Theory (SMEFT)

The Standard Model Effective Field Theory (SMEFT) [63, 64, 62] is based on the idea that the Standard Model is an effective field theory valid up to a certain energy scale, above which new physics might appear. This approach introduces higher-dimensional operators that encapsulate the effects of new physics at higher energies, while maintaining the symmetries of the Standard Model. These higher-dimensional operators have dimensions greater than 4, and are suppressed by the powers of the energy scale of new physics, Λ.

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \sum_k \frac{c_k^{(5)}}{\Lambda} \mathcal{O}_k^{(5)} + \sum_k \frac{c_k^{(6)}}{\Lambda^2} \mathcal{O}_k^{(6)} + \dots \tag{2.55}$$

where $\mathcal{O}_i^{(5)}$, $\mathcal{O}_i^{(6)}$, ... are higher-dimensional operators of dimensions 5, 6 and so on, and the $C_i^{(n)}$ are dimensionless Wilson coefficients determining the strength.

At dimension five, there is a single lepton-number-violating operator. It generates the neutrino masses and mixings after the electroweak symmetry breaking [21].

The dimension six operators are defined in Warsaw basis constructed by [21].

SMEFT allows model-independent search for new physics instead of specifying a particular BSM theory such as Supersymmetry. Any measured deviation from the SM prediction in experiments can be linked to the Wilson coefficients of the higher-dimensional operators, allowing for a systematic comparison of experimental data with theoretical predictions.

2.2.2 Higgs Effective Field Theory (HEFT)

The Higgs Effective Field Theory (HEFT) [65, 66] is an extension of the SMEFT, tailored to scenarios where the electroweak symmetry breaking sector may not behave in the way predicted by the Standard Model. It allows for more general deviations from the SM prediction, especially in cases where the Higgs boson might not be a fundamental $SU(2)$ doublet, or when new physics does not decouple cleanly

scalar. P is the parity, $P = +$ means the particle is parity-even. In HEFT, the Higgs boson is treated as a general scalar field without assuming it is part of an $SU(2)_L$ doublet. It is parametrised as a standalone $J^P = 0^+$ particle with flexible couplings.

at high energy scales.

The Higgs field can be decoupled from the $SU(2)_L \times U(1)_Y$ symmetry in a nonlinear representation, similar to how pions are treated in chiral perturbation theory [67]. The chiral perturbation theory describes the low-energy dynamics of pseudo-Goldstone bosons resulting from the spontaneous breaking of chiral symmetry[2] in QCD. The pion fields are written as

$$U(x) = \exp\left(\frac{2i}{f_\pi}\pi(x)\right), \quad \pi(x) = \pi^a(x)T^a \tag{2.56}$$

For T^a the generators of $SU(2)$ (or $SU(3)$, we will focus on $SU(2)$ for now), $\pi^a(x)$ the pions, and f_π the pion decay constant. The chiral Lagrangian can be written as

$$\mathcal{L}_{\text{pion}} = \frac{f_\pi^2}{4}\text{Tr}\Big\{\partial^\mu U^\dagger \partial_\mu U\Big\} \tag{2.57}$$

In HEFT, instead of assuming Higgs is embedded in a doublet, they are treated as a separate degree of freedom, which can have more general couplings to other Standard Model particles.

In analogy with the formalism in chiral perturbation theory for the pions, in HEFT the three Goldstone bosons π^a are embedded into a unitary matrix transforming as a bi-doublet of the global symmetry group $SU(2)_L \times SU(2)_R$

$$U(x) = \exp\left(i\frac{\pi^a(x)\sigma^a}{v}\right) \quad U \mapsto LUR^\dagger \tag{2.58}$$

The physical Higgs scalar is introduced as a gauge singlet h, then the SM Higgs doublet H can be written as a fixed combination of the fields and U according to

$$\begin{pmatrix}\tilde{H} & H\end{pmatrix} = \frac{v+h}{\sqrt{2}}U \tag{2.59}$$

[2]The formation of quark condensate $\langle \bar{q}_{Li} q_{Rj}\rangle = -\sigma\delta_{ij}$ leads to spontaneous breaking of $SU(2)_L \times SU(2)_R \to SU(2)_V$ if we consider two light up and down quarks, or $SU(3)_L \times SU(3)_R \to SU(3)_V$ if we further consider strange quark. This breaking fives rise to Goldstone bosons: three pions (π^+, π^-, π^0) with the breaking of $SU(2)_L \times SU(2)_R$, and with additional Kaons and eta $(\pi^+, \pi^-, \pi^0, K^+, K^-, K^0, \bar{K}^0, \eta)$ with the breaking of $SU(3)_L \times SU(3)_R$.

Being a singlet enables h to have arbitrary couplings

$$\mathcal{F}_i(h) = 1 + 2a_i\frac{h}{v} + b_i\frac{h^2}{v^2} + \dots \tag{2.60}$$

The HEFT Lagrangian is composed of $\mathcal{L}_{\text{gauge}}$ and $\mathcal{L}_{\text{fermi}}$, and the scalar fields U, h defined above. It takes the form

$$\mathcal{L}_{\text{HEFT}} = \mathcal{L}_0 + \Delta\mathcal{L} + \dots \tag{2.61}$$

where $\mathcal{L}_0$ contains the leading order terms and $\Delta\mathcal{L}$ includes the first order deviations. $\mathcal{L}_0$ can be written as

$$\begin{aligned}\mathcal{L}_0 = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermi}} + \frac{1}{2}\partial_\mu h\partial^\mu h - \frac{v^2}{4}\text{Tr}\{V_\mu V^\mu\}\mathcal{F}_C(h) - V(h) \\ -\frac{v}{\sqrt{2}}(\bar{Q}Uy_Q(h)Q_R + \text{h.c.}) - \frac{v}{\sqrt{2}}(\bar{L}Uy_Q(h)L_R + \text{h.c.})\end{aligned} \tag{2.62}$$

Where $V_\mu = (D_\mu U)U^\dagger$ for D_μ the corresponding covariance derivative. The Yukawa couplings include dependence on h as

$$y_Q(h) = \text{diag}(\sum_n y_u^{(n)}\frac{h^n}{v^n}, \sum_n y_d^{(n)}\frac{h^n}{v^n}), \quad y_L(h) = \text{diag}(0, \sum_n y_e^{(n)}\frac{h^n}{v^n}) \tag{2.63}$$

The first term in the sum ($n = 0$) generates fermion masses, while higher-order terms describe the interactions involving an arbitrary number of h insertions. This Lagrangian is equivalent to the SM Lagrangian, with the difference that it allows for an arbitrarily large number of Higgs and π insertions, and the Higgs couplings are parametrised by independent coefficients rather than being fixed by the doublet structure.

The Lagrangian $\Delta\mathcal{L}$ contains the leading deviations from $\mathcal{L}_0$. Unlike SMEFT framework, classifying the invariants in HEFT is not straightforward, as HEFT is a hybrid approach combining elements of chiral perturbation theory with SMEFT. Since these two frameworks use different counting rules, the structure of the mixed expansion becomes more complex.

Chapter 3

The ATLAS Experiment

The work presented in the thesis relies on the data recorded during the proton-proton (pp) collisions at the LHC with the ATLAS detector [68]. This chapter provides an overview of the LHC and ATLAS detector, with a detailed discussion about the trigger systems used in Run 2 [3], and my ATLAS authorship qualification task of improving track-finding efficiency for b-jet signature of the inner detector.

3.1 The Large Hadron Collider (LHC)

The LHC is the world's largest particle accelerator. Crossing the border of France and Switzerland, near Geneva, it lies 100m underground with 27 km circumference. There are four main detectors: ATLAS [68], CMS [69], LHCb [70] and ALICE [71].

To date, LHC has completed two operational runs: Run 1 (2009-2013) at $\sqrt{s} = 8$ TeV with around 25fb^{-1} of pp data collected [72, 73], and Run 2 (2015-2018) at $\sqrt{s} = 13$ TeV with 140.1fb^{-1} pp data collected [74]. Currently, it is on Run 3 (started in July 2023) at $\sqrt{s} = 13.6$ TeV. In Run 4 (scheduled to begin in 2027), the high-luminosity LHC (HL-LHC) which can increase the instantaneous luminosity by a factor of 10, will start to take data. The increase in the centre of mass energy allows access to higher-mass particles and processes, and the higher luminosity means more collision data can be collected. However, higher rates of collisions raise challenges to data selection and disentangle of the overlapping collisions.

3.2 The ATLAS detector

The ATLAS detector is a multi-purpose particle detector with a forward-backward symmetric cylindrical design and provides nearly 4π coverage in solid angle around the collision point. It operates using a right-handed coordinate system, with the origin at the nominal interaction point at the centre of the detector. The z-axis is aligned along the beam direction, the x-axis points from the interaction point toward the centre of the LHC ring, and the y-axis points upwards. In the transverse plane, cylindrical coordinates (r, ϕ) are used, where ϕ is the azimuthal angle around the z-axis. The pseudorapidity η is defined in terms of the polar angle θ, as $\eta = -\ln\tan(\theta/2)$ (visualised in Figure 3.1), and angular distance $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

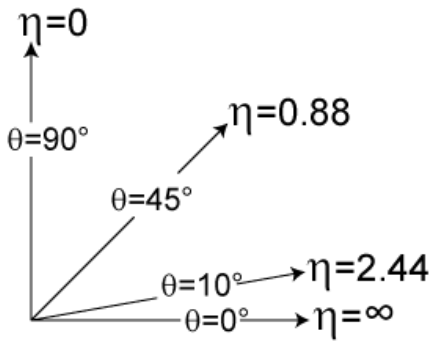


Figure 3.1: Illustration of pseudorapidity for different polar angles θ. [2]

It is composed of several key subsystems: an **inner tracking detector** surrounded by a thin superconducting solenoid providing a 2T axial magnetic field, **electromagnetic and hadronic calorimeters**, and a **muon spectrometer** equipped with large superconducting toroidal magnets. The inner detector is designed to measure the trajectories and properties (momenta, vertex positions, etc) of charged particles. The calorimeters are designed to measure the energy of electrons, photons and hadrons. The presence of neutrinos and other non-interacting particles can also be indicated by the missing transverse energy. Finally, the muon spectrometer is designed to measure the properties of muons, which penetrate other detector layers due to their minimal interaction with matter.

3.2.1 The Inner Detector

The inner detector is the first subsystem to encounter the decay products from collisions, enabling precise tracking of charged particles within the range $|\eta| < 2.5$. From the innermost to outermost layers, it consists of a Pixel Detector, Silicon Microstrip Tracker (SCT) and Transition Radiation Tracker (TRT) [75].

Positioned just 3.3cm from the LHC beam line, the high-granularity Pixel Detector is the first point of detection in the ATLAS experiment. As charged particles emerge from the collision point, they deposit small amounts of energy in the Pixel Detector. These signals are measured with a precision of nearly $10\mu m$. An insertable B-layer (IBL), placed closest to the interaction point, was installed before Run 2 [76, 77], spanning the entire pseudorapidity range of the inner detector. The pixel detector typically provides four space-point measurements per track.

The pixel detector is surrounded by the silicon microstrip tracker (SCT) [78, 79], which detects and reconstructs the tracks. It consists of over 4000 modules of 6 million ”micro-strips” of silicon sensors, arranged such that each particle crosses at least four layers of silicon. These layers provide axial hits, or azimuthal hits in the endcaps, as well as small-angle stereo hits that enhance spatial resolution in the longitudinal direction along the modules. The particle tracks can be measured with a precision of up to $25\mu m$. The SCT typically provides eight measurements per track.

The final layer is the Transition Radiation Tracker (TRT) [80, 81, 82], which consists of multiple layers of straw cylindrical drift tubes filled with a gas mixture. When charged particles pass through the straws, they ionise the gas, creating a detectable electric signal. This enables radially extended track reconstruction up to $|\eta| = 2.0$, and provides information on the particle type. It can make precision measurements of 0.17mm.

3.2.2 The Calorimeters

The calorimeter system measures the energy lost by particles as they pass through the detector [83], covering the pseudorapidity range $|\eta| < 4.9$. Electromagnetic calorimeters are responsible for measuring the energy of electrons and photons as they interact with matter, while Hadronic calorimeters measure the energy of hadrons

as they interact with atomic nuclei. They can stop most known particles, except for muons and neutrinos. The system is composed of an inner Liquid Argon (LAr) Calorimeter and an outer Tile Hadronic Calorimeter.

Within $|\eta| < 3.2$, electromagnetic calorimetry is provided by high-granularity barrel and endcap lead/liquid-argon (LAr) calorimeters. An additional LAr pre-sampler covering $|\eta| < 1.8$ corrects for energy loss in material upstream of the calorimeters.

Hadron calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadron endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic energy measurements respectively.

The energy resolution of the calorimeters quantifies their ability to precisely measure the energy of incoming particles. It is evaluated as $\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus c$, where E is the energy of the incident particle, a is sampling term coefficient representing fluctuations in the energy deposition process, and c is the constant term accounting for systematic effects such as calibration uncertainties. The LAr Calorimeter has an energy resolution of about $\frac{10\%}{\sqrt{E}} \oplus 0.2\%$ [84]. For the Tile Calorimeter, the energy resolutions for pions, kaons and protons are $\frac{47\%}{\sqrt{E}} \oplus 5\%$, $\frac{50\%}{\sqrt{E}} \oplus 2\%$ and $\frac{40.3\%}{\sqrt{E}} \oplus 5\%$ respectively [85].

3.2.3 The Muon Spectrometer

The muon spectrometer (MS) identifies and measures the momenta of muons [86]. It consists of separate trigger and high-precision tracking chambers, which measure the deflection of muons as they pass through a magnetic field generated by the superconducting air-core toroidal magnets. The toroids provide a magnetic field integral ranging from 2.0 to 6.0 $\mathrm{T} \cdot \mathrm{m}$ across most of the detector. Precision chambers cover the region $|\eta| < 2.7$, using three layers of monitored drift tubes, complemented by cathode-strip chambers in the forward region, where the background levels are highest. The muon trigger system covers the range $|\eta| < 2.4$, using resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

The spatial resolution is primarily achieved using the Monitored Drift Tube chambers, with an accuracy of approximately $82\mu m$ [87]. In terms of energy resolution, it is designed to measure the transverse momentum of muons with $p_T > 3$ GeV with a resolution of 4% up to p_T of 100 GeV, and increasing to 10% at 1 TeV [88].

3.2.4 The ATLAS Trigger and Data Acquisition (TDAQ) Systems

A trigger refers to the data selection system used to filter out irrelevant collision events. The ATLAS Trigger and Data Acquisition (TDAQ) system is responsible for selecting the most interesting events to be recorded from the billions of collisions that occur every second in the LHC. Without this system, it would be impossible to store or analyse the vast amount of data produced.

It has two primary modes in data processing: online and offline. The online system manages real-time data acquisition and initial event selection during data collection based on predefined criteria. It filters the vast number of collision events to a manageable rate for storage and further analysis. The offline system focuses on the detailed reconstruction, calibration and analysis of events that have been selected and recorded by the online system.

Interesting events are first selected for recording by the first-level trigger system (Level 1, L1) [89, 90], which is implemented in custom hardware. This is followed by selections made by algorithms implemented in software in High-Level Trigger (HLT) [89, 91]. In Run 2, the first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, while the high-level trigger reduces this rate to approximately 1.2 kHz for events to be recorded to disk [92].

In Run 1 [89, 93], the ATLAS TDAQ system consisted of a multi-level trigger system, including a pipelined L1 trigger and two large farms of commodity CPUs: one for the Level 2 (L2) software trigger, and the second for the third level, event filter (EF) processing. For Run 2, the centre-of-mass energy was increased from 8 to 13 TeV, with nominal bunch spacing reduced from 50 ns to 25 ns [94], and an increase in the beam intensity per bunch crossing. This led to a significant increase

in the mean pile-up[1] interaction multiplicity per bunch crossing, resulting in higher track and hit multiplicities. The trigger rates are therefore on average a factor of 2 to 2.5 times larger for the same luminosity and with the same trigger criteria [91]. To address this, the separate L2 and EF farms were integrated into a single, homogeneous HLT farm [95], allowing improved resource sharing and simplifying both the hardware and software components.

For the L1 trigger, a coarse granularity subset of the data from the calorimeters and the muon spectrometer is read out, processed, and sent to the off-detector hardware L1 processors. In parallel, data from the entire detector are stored in custom on-detector pipelined read-out buffers (ROBs) [96]. After a positive L1 trigger decision, data from the corresponding bunch crossing is read out from the ROBs to the off-detector read-out subsystem, ready for distribution to the HLT processing nodes.

The data volume in the SCT and pixel detectors are so high that they can only be read out following an L1 accept. Therefore, the track reconstruction in the silicon layers first occurs at HLT. The coarse granularity data readout of the L1 trigger from the muon spectrometer or calorimeter is used to reconstruct objects of interest, which are used to define regions of interest (RoI) in the detector that are worth reconstructing with the data read out at the full granularity.

3.3 The Inner Detector Trigger

The HLT is the first stage where information from the silicon detectors in the Inner Detector becomes available in the trigger processing. The track reconstruction in the HLT can be performed either within an RoI identified at L1 or for the full detector. It consists of a two-stage approach with a *fast tracking* reconstruction stage to reject the less interesting events; then for the remaining events, a slower precision reconstruction stage, known as *precision tracking*, is employed. The main steps are: the formation of space-points, track finding seeded by space-points, ambiguity

[1] pile-ups: Additional pp collisions that occur within the same or nearby bunch crossings as the primary collision of interest. These extra collisions introduce additional particles into the detector, which can overlap with the particles produced by the primary collision and complicate the analysis.

solving and TRT extension. The generic Run 2 tracking strategy for a single-stage is shown in Figure 3.2. Here single-stage refers to the single processing of a specific RoI. The tracking is split into steps separated by additional non-tracking-related algorithms to reconstruct extra features, and hypothesis algorithms that select objects meeting certain selection criteria. These are used to reduce the rate of RoI that need to be processed between the fast track finder and precision tracking steps.

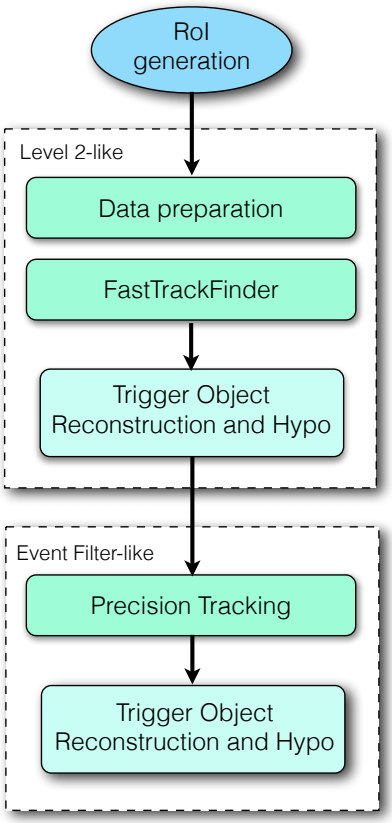


Figure 3.2: The single-stage tracking for a single RoI. [3]

3.3.0.1 Data preparation

The pixel and SCT data preparation involves decoding hits in the silicon modules of the Inner Detector from a binary stream format, grouping adjacent hits into clusters and forming space points from the clusters [97]. These space points represent locations in three-dimensional space and are used for pattern recognition, while the clusters themselves are used for the actual track fitting.

After L1 triggering, the data for each detector are read out to the ROB system.

In order to reduce the bandwidth, only the ROBs containing data corresponding to an RoI need to be processed for read-out to the HLT. The decoders iterate over ROB data fragments and the data words[2] within each fragment. While processing an individual fragment, the algorithm keeps track of the module being decoded, assigning raw data objects (RDOs) - in this case, decoded hits - to an appropriate in-memory container. Once the data words are decoded and RDOs are created and mapped to their respective modules, clustering algorithms are used to group adjacent hits within a module into clusters, referred to as reconstruction input objects. These pixel and SCT clusters are then converted to space points using a simple geometric transformation.

TRT data preparation is carried out only for precision tracking in the RoI, where the TRT track extension is included.

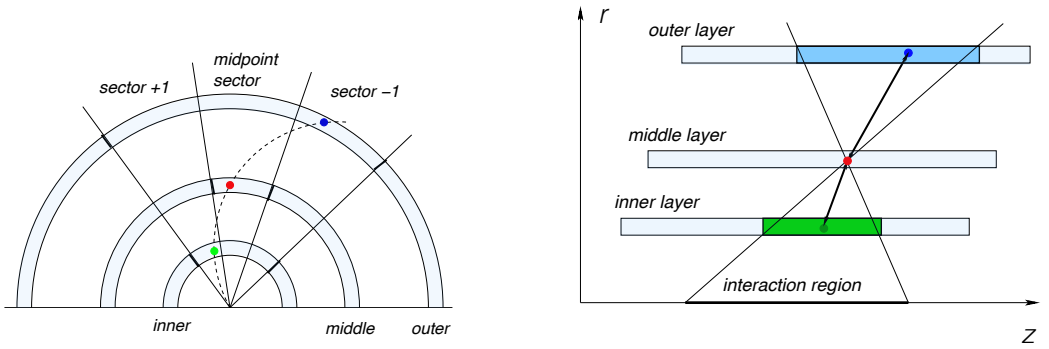
3.3.0.2 Fast tracking

Fast track finder (FTF) [98] was developed to provide track candidates for use early in the trigger, which are then used to seed the precision tracking stage. Given that the precision tracking supplies the track candidates used for the final objects selection in the trigger, the FTF design prioritises track finding efficiency over purity.

In the FTF pattern recognition process, a search for triplets of space-points (*track seeds*) is performed within bins of r and approximately 50 sectors in ϕ. Triplet formation begins with a $middle$ space-point and selects $outer$ and $inner$ space-points at larger and smaller radii. When tracking is performed within an RoI, this region can be replaced with a restricted z-region of the RoI along the beam line, provided there is information available about the z position of the interaction. This is illustrated in Figure 3.3.

Initial track candidates are then formed from the track seeds using a simple track finding algorithm which extends the track candidates into further layers to find additional hits, using the offline track-following algorithm [99] with a modified configuration for faster execution. The track-following algorithm extends track candidates into neighbouring layers to find additional hits, continuously refitting

[2] data word: the units of information, in binary representation



(a) Track seed formation in radial bins and azimuthal sectors, with the *midpoint sector* for a triplet being the sector that contains the midpoint. The innermost and outermost hits are allowed to originate in the same sector or from adjacent sectors

(b) Track seed formation in the $r-z$ plane. The extrapolated z position of the triplet at the bean line can be restricted to within a specific z region.

Figure 3.3: Track seed formation [3]

the track candidates as new hits are incorporated. To retain tracks of higher quality, an algorithm is applied to remove duplicate tracks that share track seeds. These preliminary tracks are then processed by a fast Kalman filter track fitter [100]. Track candidates with too large $|d_0|$ (d_0 is the traverse impact parameter, representing the distance of the closest approach of the particle's trajectory to the primary vertex in the transverse space) are rejected to maintain a manageable level of fake tracks.

3.3.0.3 Precision tracking

The precision tracking stage uses the FTF tracks as input and applies a version of the offline tracking algorithms [101, 102] configured to run online in the trigger [103, 104]. This includes a sophisticated algorithm for rejecting duplicate track candidates and clusters that were previously assigned to tracks, but were located too far from the track trajectory. Additionally, the track candidates are extended into the TRT to select TRT hits at larger radii to improve the track momentum resolution. The final ID track fit is performed using a more precise global χ^2 fitter algorithm [105] from code used for offline reconstruction. These processes are combined as components of the offline ambiguity solver [101] algorithm, which resolves ambiguity with respect to duplicated or falsely assigned hits, ranking the tracks and rejecting tracks with low qualities (detailed in Section 3.3.2).

More detailed handling of the detector conditions is feasible given the rate of event processing for precision tracking is much lower than fast tracking. The resulting precision tracks are much closer in performance to the offline tracks than the fast tracks. Since precision tracking relies on the tracks and clusters identified by fast tracking, precision tracking efficiency cannot exceed that of the fast tracking. Overall, the primary purpose of the precision tracking is to perform a higher quality fit, enhancing the purity and quality of the trigger tracks relative to those reconstructed offline.

3.3.0.4 Vertex Reconstruction

Two vertex algorithms are used online: a histogram-based algorithm, and the offline iterative vertex algorithm [106, 107]. Typically, trigger signatures employ only the offline vertex algorithm. However, for the *b*-jet trigger [108], both algorithms were used to maximise the vertex finding efficiency.

The histogramming algorithm begins by creating a histogram of the z_0 positions for the point of closest approach to the beam line of each track. It then calculates the vertex z position by calculating the mean of the bin centres, weighted by the number of tracks in each bin within the 1 mm sliding window that contains the most tracks. All tracks that meet some basic quality selection are included and are weighted equally.

The second algorithm uses the offline vertex finder, with minor modifications to make it work online. It iteratively clusters tracks to pinpoint the positions of vertex candidates, sorting them by the sum of the squared transverse momenta of the tracks assigned to each vertex.

Both algorithms operate on tracks reconstructed within the relevant RoI of the track finding process. For the *b*-jet trigger, both vertex algorithms are executed using tracks from a specific additional vertex tracking stage before conducting dedicated tracking in the *b*-jet RoI.

3.3.0.5 Multistage tracking

Running both fast and precision tracking sequentially in a single RoI is termed processing in a *single tracking stage*. When multiple tracking passes are intentionally

performed over similar but slightly different regions of the detector, where the second pass occurs in a different RoI that overlaps, extends, or updates the initial RoI, this is referred to as *multistage tracking*. Each set of steps within a specific RoI includes a single tracking stage. For example, the first stage may run fast tracking in a narrow RoI, while a subsequent stage runs the fast tracking again in a new, wider RoI along the same direction, as shown in Figure 3.5.

For the b-jet trigger [108], the initial tracking stage focuses on identifying the likely event vertex z-position to be used in the second stage. In the second stage, separate RoI are defined around each jet axis, with each RoI more tightly constrained along the bean line around the z-vertex position identified in the first stage. Note that each vertex or jet consists of multiple particles resulting from the hadronisation of a quark or gluon, the corresponding ROI can contain multiple tracks.

For the tracking in the initial vertex stage, all jets identified by the jet trigger with transverse energy $E_T > 30$ GeV are considered, and tracks are reconstructed using the fast tracking algorithm within a full width of 0.2 in η and ϕ around the jet axis for each jet, and a fully-extended z-range of $|z| < 225$ mm along the bean line.

To avoid redundant processing of overlapping detector regions, the RoI about each jet axis are first aggregated into a single "super RoI" for the event before running the fast tracking, as shown in Figure 3.4. This super RoI is then used during the data preparation for the first-stage vertex tracking to determine which detector elements - specifically individual silicon modules from the pixel and SCT detectors - need to be read out for further processing. This step occurs only once per events for the super RoI, which can include non-overlapping regions. With data from these selected detector elements, tracking proceeds as if it were processing the entire detector.

Following this stage, the tracks identified within the super RoI are used for primary vertex reconstruction [109]. This reconstructed vertex is then used to define broader RoI around each jet axis, with $|\Delta\eta| < 0.4$ and $|\Delta\phi| < 0.4$ relative to the jet axis, and $|\Delta z| < 10$mm relative to the primary vertex z-position. These RoI are used for the second-stage reconstruction, where fast tracking is performed again, but in the wider η and ϕ regions around the jets. This is followed by precision tracking,

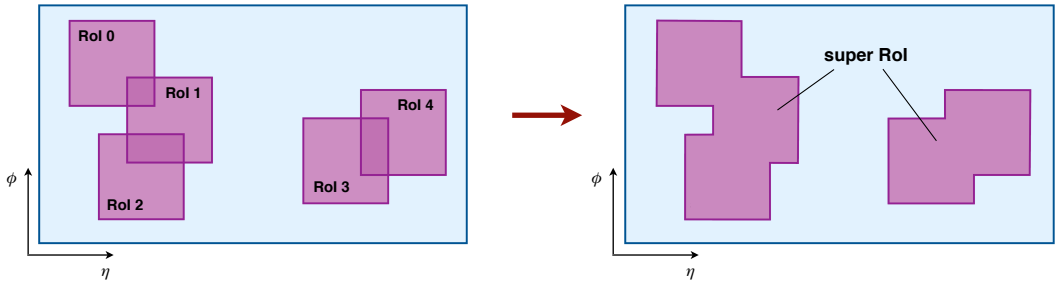


Figure 3.4: How the super ROI for the vertex tracking is created from all the trigger jets reconstructed. The RoI around each initial jet direction is fully extended in z along the bin line. [3]

secondary vertex reconstruction, and b-tagging algorithms.

3.3.1 High-level trigger tracking performance

This chapter outlines the measurement of the performance of the HLT, with a focus on the b-jet signature, which is what my qualification task is about.

3.3.1.1 Data selection

The performance analyses for muon and electron signatures use the full integrated luminosity from the 2016-2018 data-taking period. This is feasible because the processing for these signatures remained largely consistent across the years. In contrast, the tau and b-jet analyses, which employed two-stage tracking with modified second-stage RoI, have significant changes made to the reconstruction algorithms over the three years. As a result, only the 2018 data is analysed in full detail. For the tau signature, the complete integrated luminosity from each year is used. While for the b-jet analyses, only a subset of data from each year, sufficient to achieve a small statistical uncertainty is used. [3]

In all cases, events are used only if the data quality has been assessed as good for physics analyses, which means the data should be clean from any hardware or software-related issues that may compromise their integrity [110].

The trigger tracking efficiency relative to offline tracking is assessed using several support triggers. These support triggers, essentially identical to the physics triggers, function by reconstructing tracks in the usual manner but selecting objects based only on the muon spectrometer or calorimeter data, without using ID tracking

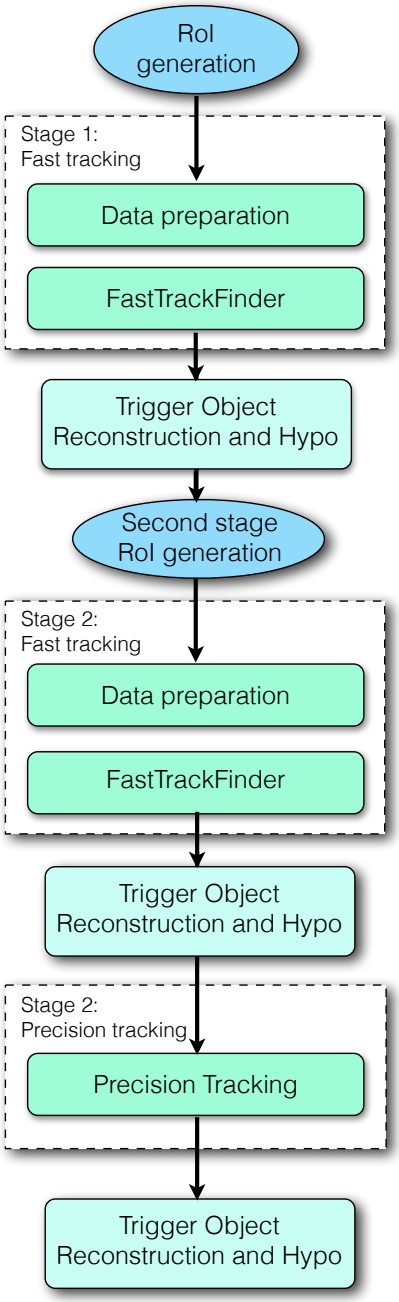


Figure 3.5: Illustration of the multistage tracking. Here multistage refers to the multiple passes of the track processing over specific parts of the detector, with an updated or modified RoI. [3]

information. This method allows the tracking efficiency to be estimated without bias from the ID track reconstruction. One limitation is that, because the trigger does not rely on inner detector track information for selection, the prescale on these triggers must be kept high to control the rate. Because the trigger objects used for the selection do not include tracking, the proportion of background selected by these triggers is significantly higher than that for fully selecting triggers. To mitigate such background in each analysis, a reference sample is selected using fully offline reconstructed objects, against which performance is evaluated. As a result, the sample of high quality offline objects, including the offline tracks used to monitor the trigger, may be statistically limited in regions of phase space.

3.3.1.2 Track Selection

To measure efficiency and resolution, the set of inner detector trigger tracks within the RoI are first matched to selected offline reference objects using a modified approach to the stable marriage problem [111] - only the closest matches are considered, taken from those tracks within a loose preselection cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.05$ around an offline reconstructed track. For each offline track, the closest matching trigger track is chosen as a match only if it is not more closely matched to any other offline track.

Efficiency is determined by calculating the ratio of offline reference objects with a matched trigger track to the total number of offline objects that meet the selection criteria.

To account for the long non-Gaussian tails often present in the residual distributions, the resolution is estimated by taking the root-mean-square (RMS) of the central 95% of the distribution. This value is then adjusted by scaling it with the inverse of the RMS for the central 95% of a Gaussian distribution with a unit standard deviation. This scaling ensures that the true width is accurately obtained for any Gaussian distribution.

Reference objects are selected based on full offline selection criteria, akin to those used in physics analyses, with additional selection. Offline candidate tracks must lie within the RoI used for trigger reconstruction, as tracks outisde this region

cannot be reconstructed. Furthermore, offline tracks are considered only if their transverse impact parameter at the point of closest approach to the bean line and the z-position relative to any relevant offline vertex are both under 10mm.

3.3.1.3 Tracking in the b-jet signature

For the b-jet trigger [91, 108], tracking is executed in two stages as discussed in 3.3.0.5. In the first stage, vertex tracking is performed within narrow $\eta - \phi$ RoI around the jet direction for jets with $p_T > 30$ GeV, but extending fully from $-255 < z < 225$ mm along the bean line. These are aggregated into a super RoI before track reconstruction. The resulting tracks are then used to reconstruct the vertex for the second stage of processing.

The primary goal of vertex tracking is to identify a sufficient number of high transverse momentum tracks to reconstruct the primary interaction vertex associated with the jets, thereby refining the RoI for tracking used for b-tagging.

In the second stage, both fast and precision tracking are conducted in wider RoIs centered on the jet axes, with $|\Delta\eta| < 0.4$ and $|\Delta\phi| < 0.4$ relative to the RoI direction, and centered at the primary vertex position identified in the first stage. These second stage RoIs have a narrower extend of $\Delta z < 10$ mm about this z position.

Only the tracks from the precision tracking in the second stage are used for b-jet selection. Additionally, these precision tracks are employed in the trigger for subsequent global jet energy calibration [112].

To select the reference tracks used for analysing the efficiency and resolution of the b-jet signatures, tight offline criteria that consider detector geometry and performance as a function of pseudorapidity are used. They are:

- Tracks with $\eta \leq 1.65$ must have more than eight silicon hits, while those with $\eta > 1.65$ must have 11 or more silicon hits.
- Tracks passing through active modules in the IBL and the next innermost pixel layer must have at least one hit in each of these layers.
- Tracks should have no missing pixel hits in any active modules in layers at a larger radius than the innermost pixel hit.

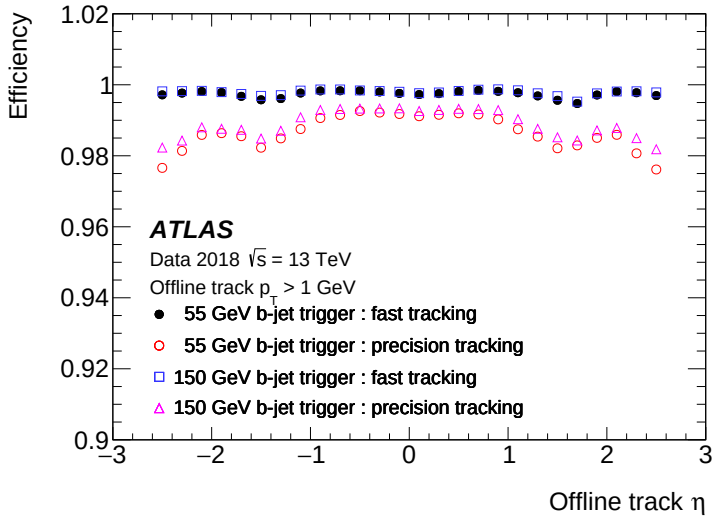
- If no hits are present in the IBL and the next innermost pixel layer, the tracks must have at least one pixel hit in any of the remaining layers.

The efficiency is measured using dedicated triggers that run as standard b-jet physics signatures but without selecting on tracking information or performing b-tagging selection.

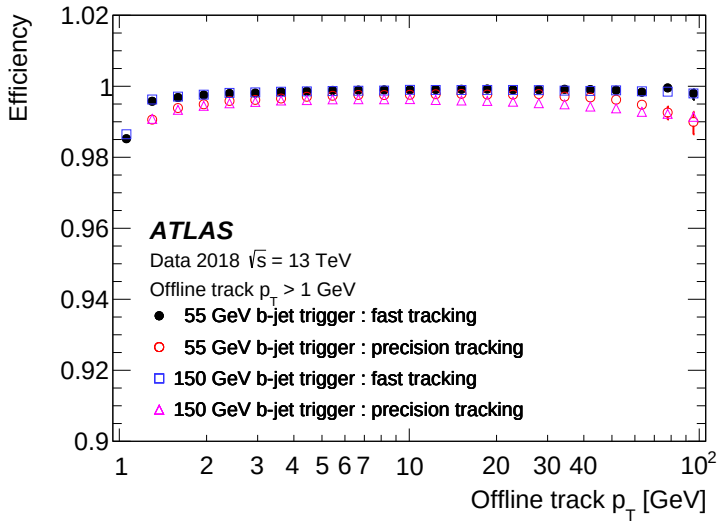
Figure 3.6 shows the precision and fast tracking efficiencies for jet tracking from the 55 GeV and 150 GeV threshold triggers. Because of the much larger radius of the outer SCT layer, the transition between barrel and endcap for tracks in the SCT begins much earlier in pseudorapidity, around $|\eta| = 1.2$. This results in dips in efficiency around $|\eta| = 1.2$ in Figure 3.6a. We can see that in Figure 3.6b, the fast tracking efficiency for tracks near 1 GeV exceeds 98%, whereas the precision tracking efficiency is very low - around 85%, and falls outside the range of the y-axis in the first bin. This discrepancy results from an explicit 1 GeV cut applied to the tracks from precision tracking in the b-jet signature to manage processing latency. The reconstruction for tracks near the threshold, where the trigger p_T is estimated to be lower than the 1 GeV p_T threshold, is terminated early. For offline tracks with $p_T > 10$ GeV, the precision tracking efficiency begins to decline slightly and is more significant for the higher threshold b-jet triggers. This is mainly due to the reduced pixel hit multiplicity observed on the precision tracks for the higher threshold b-jet triggers. While the hits are mostly present on the fast tracked tracks, the stricter hit selection criteria applied to the precision tracks lead to this reduction in efficiency. These affect of reduced b-tracking efficiencies will be studied in my ATLAS authorship qualification task, detailed in Section 3.3.2.

3.3.2 Improving Track Finding Efficiency for b-jet signature

The tracking efficiency for the Run 3 Trigger tracking was observed to be lower than Run 2, with inefficiencies of several per cent at high η, as well as the high p_T and low p_T regions as mentioned in 3.3.1.3. During my ATLAS authorship qualification task, I worked on recommissioning the Innder Detector Trigger for the start of Run 3. I investigated the causes of these losses by both examining where the candidates are being lost in the reconstruction, and studying how modifying the internal parameters



(a) Efficiencies wrt pseudorapidity



(b) Efficiencies wrt transverse momentum

Figure 3.6: Tracking efficiency in the b-jet signature for tracks with $p_T > 1$ GeV. [3]

of the reconstruction can increase these efficiencies.

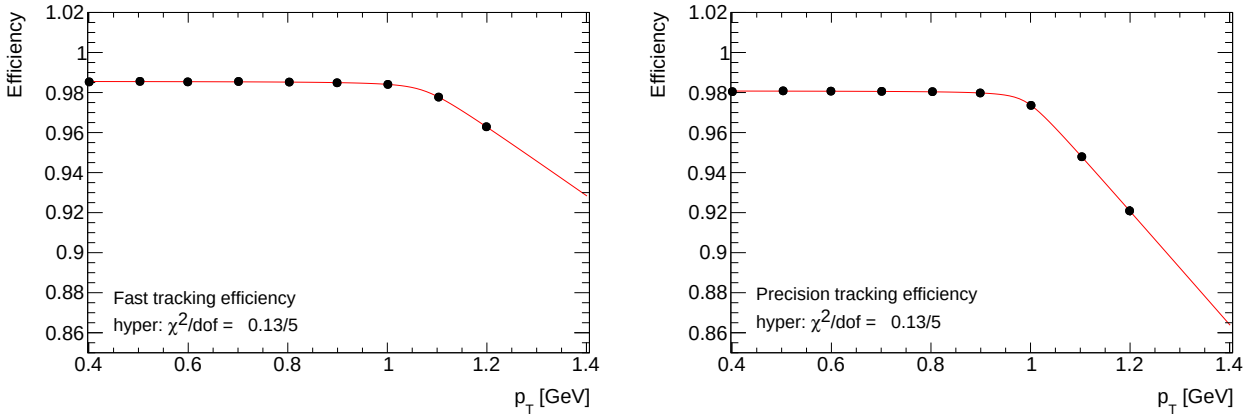
3.3.2.1 Lower the min p_T cut

I first investigated the minimum p_T cut applied during pattern recognition in order to fix the low efficiency for precision tracking for tracks near 1 GeV in Figure 3.6b, as discussed in Section 3.3.1.3.

Not that in Figure 3.6b, each bin represents the tracking efficiency within the corresponding track p_T range, and it is generated with the $p_T > 1$ GeV cut. By applying different overall p_T cuts to the tracks, the values of each data point, as

well as the overall integrated efficiency (the efficiency considering all tracks) would change.

Figure 3.7 illustrates a scan of the overall integrated efficiencies for p_T cuts ranging from 0.4 to 1.2 GeV, in increments of 0.1 GeV. The integrated efficiencies exhibit a decreasing trend with increasing p_T cuts. This decline becomes more pronounced at p_T cuts around 0.9-1.0 GeV, whereas for cuts below 0.8 GeV, the decrease is marginal.



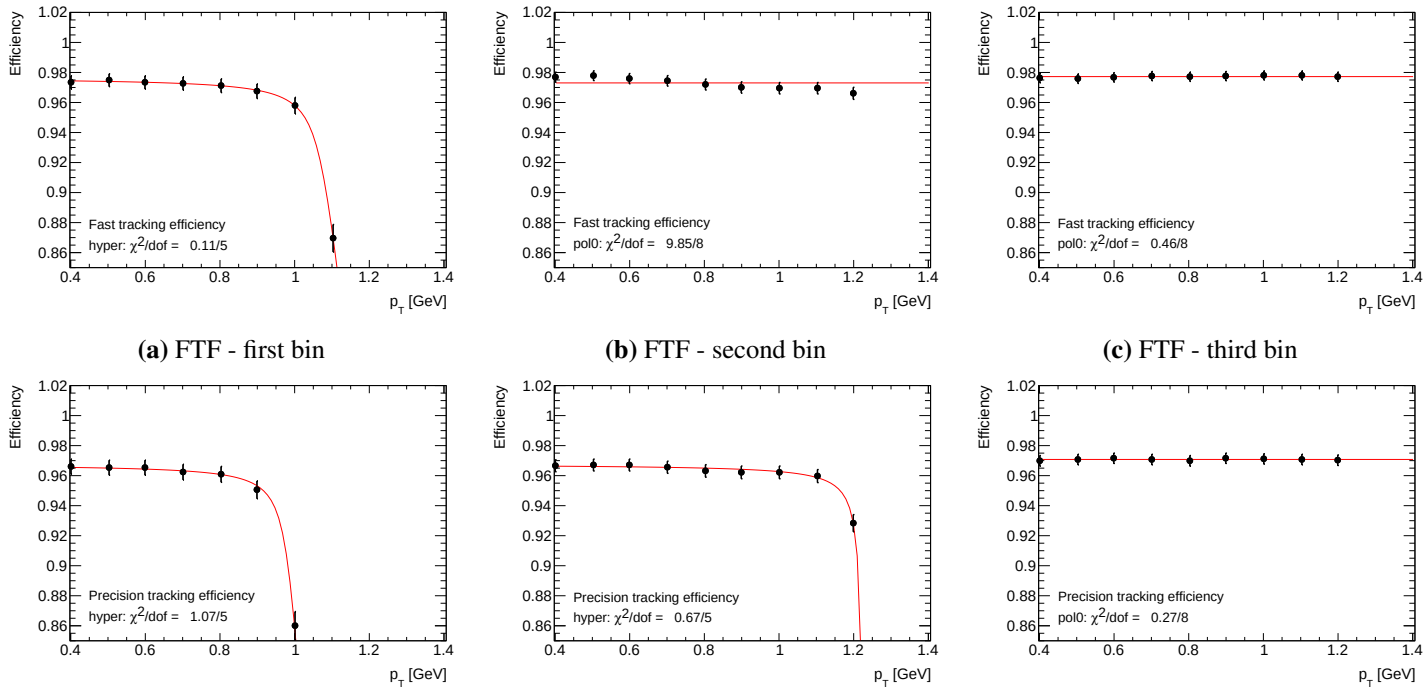
(a) FTF **(b)** Precision Tracking

Figure 3.7: Efficiencies vs minimum p_T cuts applied on the tracks, for (a) fast track finder (FTF) and (b) precision tracking.

Figure 3.8 shows how the efficiency changes for the first three bins in the efficiency vs track p_T distribution (like Figure 3.6b), when different p_T cuts are applied. As an example, Figure 3.8a shows how the fast tracking efficiency of the first bin (i.e. tracks with p_T of around 1.1 GeV) changes with varying p_T cuts. Each point in 3.8a corresponds to the efficiency of the leftmost data point in Figure 3.6b generated using the respective p_T cut.

We can observe that, reducing the p_T cut from 1 GeV results in a significant improvement in efficiency for the first bin, and a minor but noticeable improvement for the second bin. For the third and subsequent bins, the effect is negligible.

This observation leads to the proposal to lower the minimum p_T cut to 0.8 GeV in order to retrieve efficiencies for the first few bins. The increase in processing latency remains manageable, with this adjustment resulting in only about a 10%



(a) FTF - first bin **(b)** FTF - second bin **(c)** FTF - third bin

(d) Precision Tracking - first bin **(e)** Precision Tracking - second bin **(f)** Precision Tracking - third bin

Figure 3.8: Efficiencies vs minimum p_T cuts for the first three bins for FTF and precision tracking.

increase in execution time. This is considered an acceptable trade-off as it allows for the increase of approximately 1% of the total integrated efficiency and a notable 10% improvement in the first-bin efficiency in precision tracking. See Appendix C.1 for the supporting timing plots.

This proposal was accepted, and the change has been implemented for Run 3 data taking. The precision tracking efficiency for the first bin is now improved in later builds, as shown in Figure 3.9.

3.3.2.2 Relax Xi2max cut

The next stage of the project is to investigate any additional efficiencies of the Release 22 tracking compared to the Release 21 tracking[3].

In this study, the same RoI and reference tracks were used for R21 and R22 triggers. A clear reduction of efficiencies for both FTF and precision tracking from

[3]Normally abbreviated as R21 and R22, they are different versions of the Athena software framework used by ATLAS for various analysis tasks, including the Inner Detector trigger algorithm in this chapter. R21 was a major release used in Run 2, and R22 is the major release currently used for Run 3.

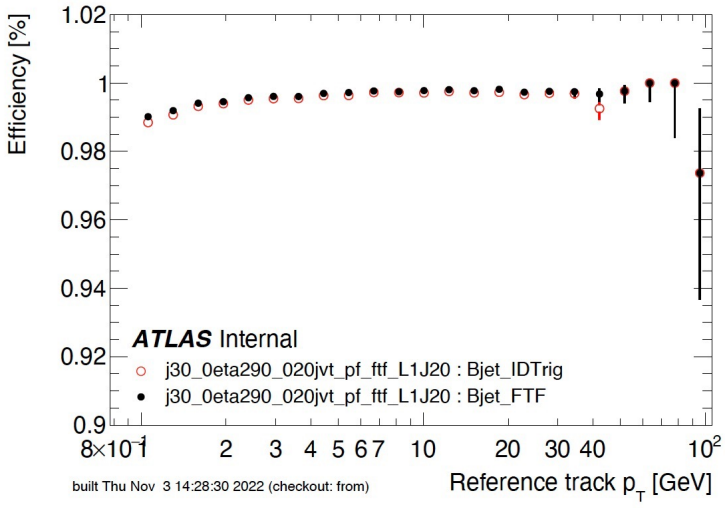


Figure 3.9: Efficiencies wrt transverse momentum with min p_T cut at 0.8 GeV

R22 to R21 is observed. We focus on the efficiency loss for high p_T tracks in precision tracking, as shown in Figure 3.10. A logical approach is to investigate the p_T bin where R22 exhibitis significantly lower efficiency than R21, which is for $p_T \in [37.8249, 46.4159]$.

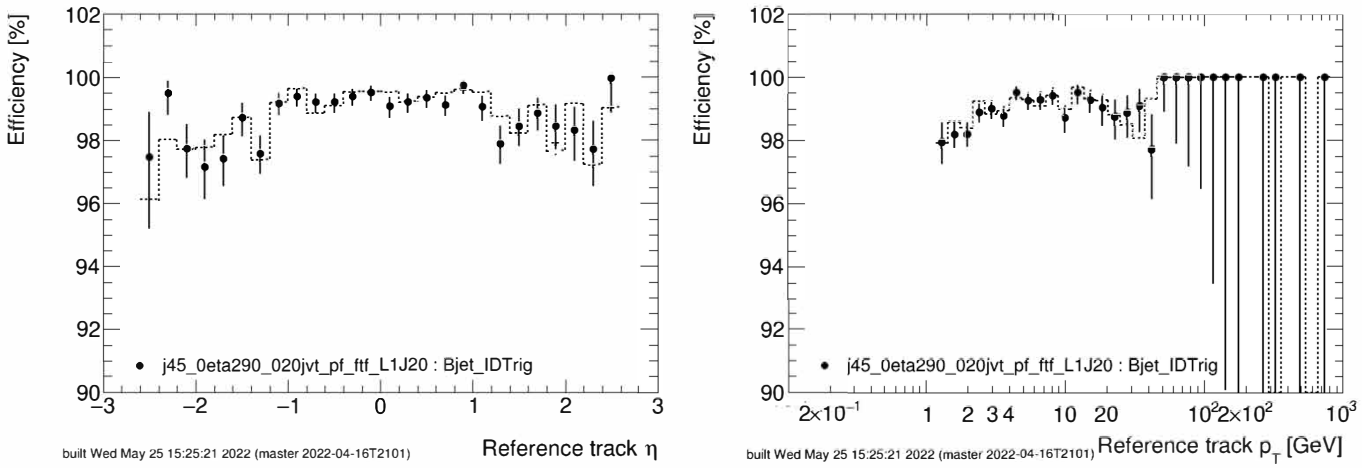


Figure 3.10: Efficiencies of R22 triggers vs R21 triggers. Using the same RoI, and same offline reference tracks. The dots represent efficiencies of R22 triggers, and the histograms represent efficiencies of R21 triggers.

Upon scanning through the events with track-p_T in this range, I identified an offline reference track in event 375 that had a matched track with the R21 trigger, but no matching track with the R22 trigger (listed in Table 3.1). This track has high transverse momentum and lies in the centre of RoI, thus it is a good quality track

which should be found.

	η	ϕ	z_0	p_T(GeV)	d_0	χ_2
R22						
offline	-1.1289	-1.84848	-11.6695	46.1243	0.206964	37.8259/37
FTF	-1.12866	-1.84688	-11.7473	29.9496	0.147066	100.665/7
IDTrig	**not found**					
R21						
offline	-1.12886	-1.84846	-11.6424	45.9495	0.206925	51.7086/38
FTF	-1.12788	-1.8454	-11.7714	21.8445	0.102233	115.858/9
IDTrig	-1.12899	-1.84869	-11.6188	49.1845	0.221677	1.7014/3

Table 3.1: The track matching information for the offline reference track in event 375 that has matching FTF and precision tracking tracks in R21, but missing the matching precision tracking track in R22. IDTrig represented precision tracking.

This discrepancy happens during the **SolveTrack** process, which is an ambiguity-solving process responsible for reconstructing the tracks from recorded hits. As described in Section 3.3.0.2, the pattern recognition process identifies track seeds, on which we can perform track fitting. However, multiple track seeds might share the same hits, or there might be multiple ways to interpret the hits and fit the tracks, resulting in ambiguities. The **SolveTrack** resolves these ambiguities by iteratively refining the track seeds and reconstructed tracks.

Initial track seeds are generated in the first iteration, and track fitting is applied to them to obtain initial tracks. Scores are assigned to tracks based on factors like number of hits, track momentum, etc. In each iteration, tracks with scores above a certain threshold and enough unique hits (not shared with other higher-scoring tracks) are accepted, and these tracks are removed from the loop. Lower-quality tracks are discarded or adjusted, and the shared hits may be reassigned correspondingly. Some tracks are refitted, normally with a more complicated algorithm, to improve their qualities. For tracks sharing a significant number of hits, or for tracks that are only partially good, subtracks can be created as fragments of the original tracks. New scores are assigned to these refitted tracks and subtracks, and the process continues until no further improvements can be made.

For R21, this process is successful for the track seeds associated with the offline reference track in event 375, resulting in a matching precision tracking track shown

in Table 3.1.

For R22, the track reconstructed from the raw data and is a close fit to the target track initially has $\eta = -1.12866, \phi = -1.84687, z_0 = -10.8928, p_T = 29.9495, d_0 = -0.197707$, which is kinematically similar to the offline reference track in Table 3.1, and has a reasonably high initial score. However, after several iterations of refitting and subtrack creation, it is finally rejected with a score of 0.

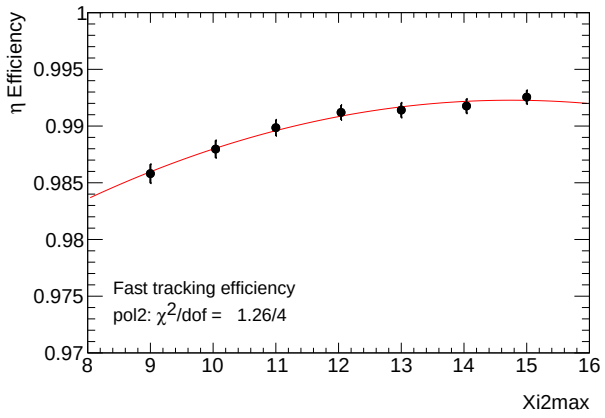
The rejection is due to the **InDetTrigTrackingCuts** configuration, which is used to define various cuts and criteria for track reconstruction, and tracks failing these cuts will be assigned a zero score and directly rejected. The **cutLevel**, which sets different levels of stringency for the cuts, was raised from 10 to 12 from R21 to R22. There are four additional criteria for Level 12:

- Xi2max, the maximum chi-squared (χ^2)[4] value allowed for a track fit, is lowered from 15 to 9. This change ensures only tracks with a better fit to the observed hits are accepted.
- Xi2maxNoAdd, the maximum chi-squared value for tracks without additional constraints, is lowered from 35 to 25.
- nHolesMax, the maximum number of holes (missing hits) allowed in the track, is reduced from 3 to 2. This change ensured that the tracks were more complete and reliable.
- nHolesGapMax, the maximum number of gaps (missing hits in consecutive tracker layers) allowed in the track, is reduced from 3 to 2. This change ensures that the tracks are more continuous and less likely to be spurious.

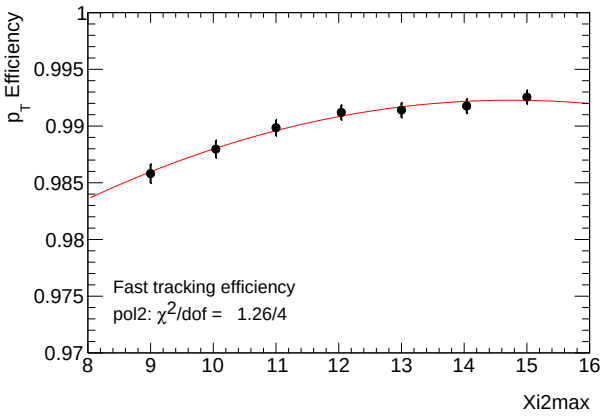
By tightening these cuts, the reconstruction algorithm can produce higher-quality tracks. However, this might come at a cost of lower efficiency since more tracks are rejected. In our case, the good-quality target track was rejected after this change. It can be recovered if we relax the Xi2max cut from 9 to 12.

[4]The χ^2 value is a measure of how well the reconstructed track fits the measure fits: $\chi^2 = \sum_i \frac{(x_i^{\text{measured}} - x_i^{\text{predicted}})^2}{\sigma_i^2}$, for x_i^{measured} the measured position of the hits, and $x_i^{\text{predicted}}$ the predicted position of the hits based on the track model, σ_i the uncertainties in the measurements, summing over the hits associated with the track.

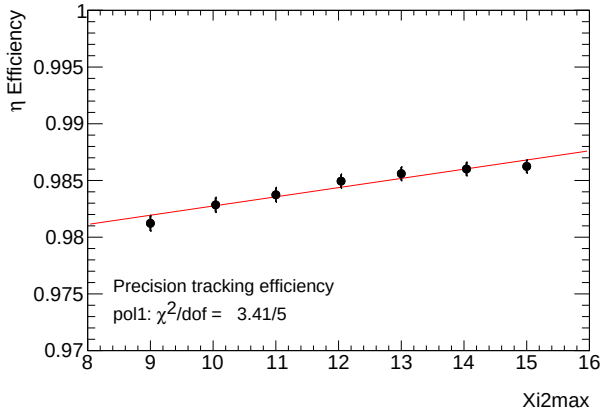
Figure 3.11 displays the scan of mean efficiencies with respect to η and p_T for Xi2max cuts ranging from 9 to 15 in increments of 1. Relaxing the Xi2max cut from 9 to 12 results in approximately a 1% increase in efficiency, beyond which the efficiency curve starts to plateau.



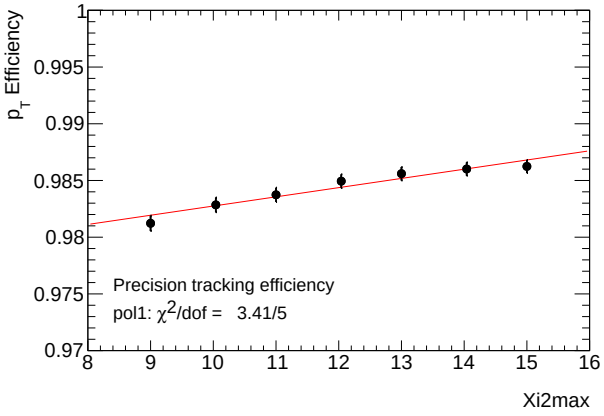
(a) η efficiencies for FTF



(b) p_T efficiencies for FTF



(c) η efficiencies for precision tracking



(d) p_T efficiencies for precision tracking

Figure 3.11: The efficiencies vs η and p_T for FTF and precision tracking, with various Xi2max cuts.

Figure 3.12 illustrates the number of input tracks and the number of accepted tracks in the precision tracking process for Xi2max cuts ranging from 9 to 15. We can see that while the number of input tracks continues to increase at an approximately constant rate with higher Xi2max cuts, the number of accepted tracks begins to plateau around a Xi2max cut of 12.

Therefore, I proposed to relax the Xi2max cut in **InDetTrigTrackingCuts** from 9 to 12. As shown in Figure 3.13, the efficiencies have been improved significantly

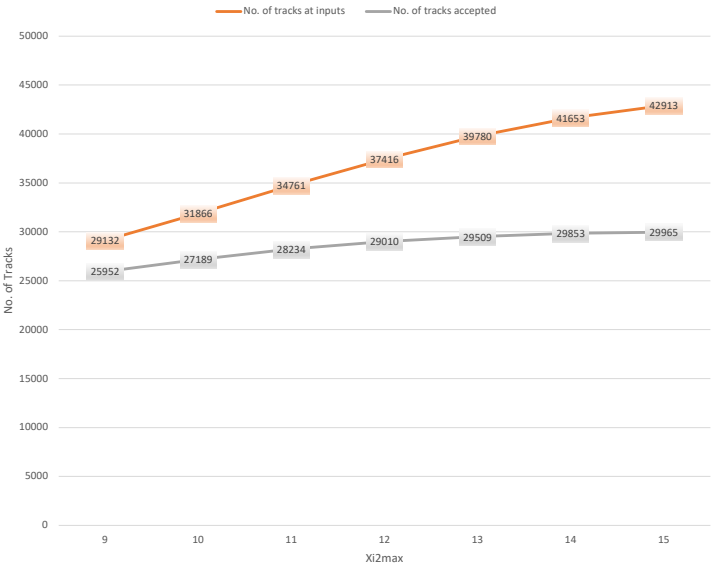


Figure 3.12: Track counts for the input and accepted tracks in the precision tracking process, using various Xi2max cut.

(around 1%) after relaxing the Xi2max cut from 9 to 12. Further supporting plots including timing, tight matching comparison to remove effects from fake tracks, and χ^2 values comparisons, are shown in Appendix C.2.

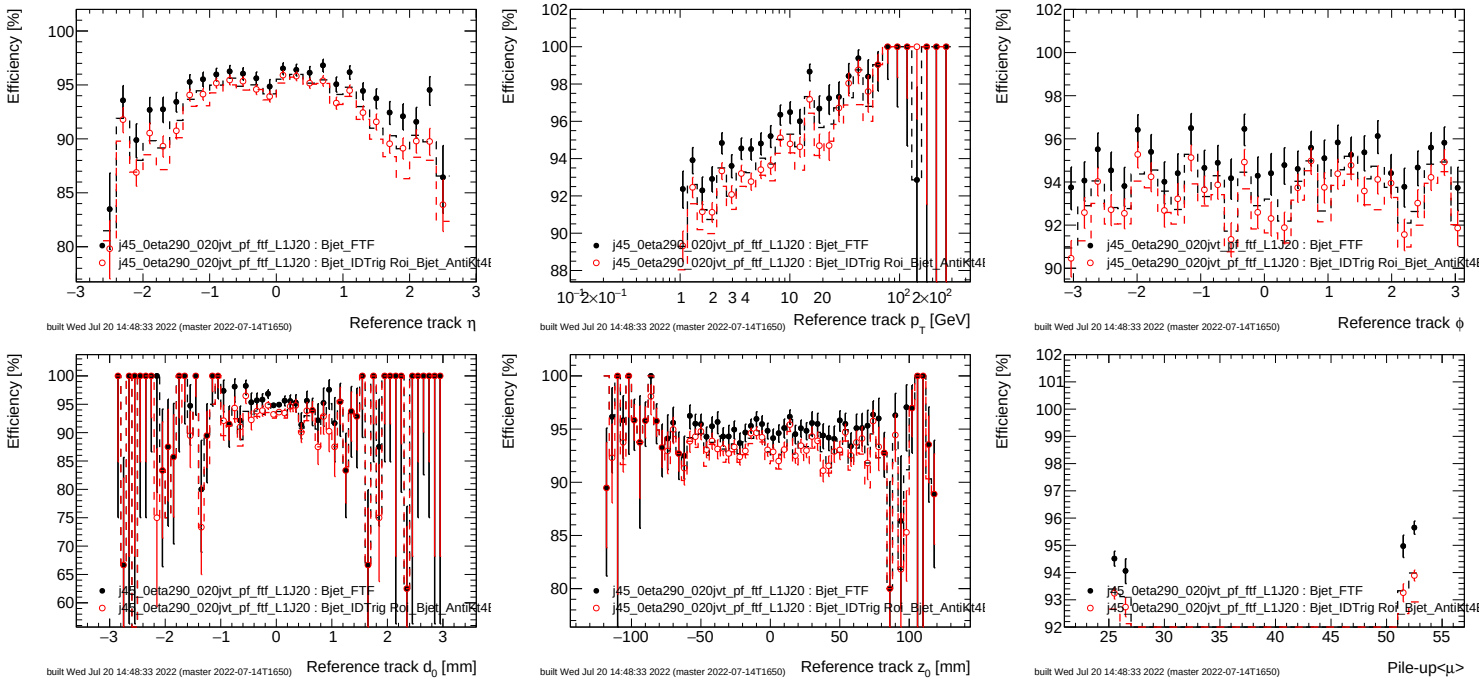


Figure 3.13: Efficiencies of *b*-jet signature of Xi2max cut at 12 vs 9, for FTF (red) and precision tracking (black). The dots represent efficiencies of Xi2max=12, and the histograms represent efficiencies of Xi2max=9. Clear improvement can be observed.

Chapter 4

Search for non-resonant HH4b production with Run 2 dataset

This chapter focuses on the published result [4] - the search for non-resonant $pp \to hh \to b\bar{b}b\bar{b}$ process with 126fb^{-1} of Run 2 pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector at the LHC.

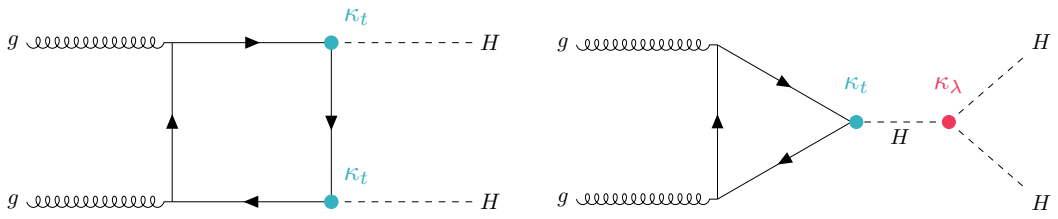
My main involvement in this analysis focuses on the background estimation part of the analysis, including investigating the analysis categorisation and binning, how different setups in background modelling impact background modelling uncertainties and sensitivities, and validation studies of the background analysis.

4.1 Introduction

The di-Higgs production process strongly depends on the Higgs self-coupling, which is a key ingredient of the electroweak symmetry breaking and a sensitive probe for physics beyond the Standard Model.

The dominant SM HH production process is the gluon-gluon fusion (ggF). For a Higgs boson mass $m_H = 125$ GeV, its cross-section calculated at next-to-next-to-leading order (NNLO) including finite top-quark-mass effects [113] is 31.05 fb at a centre-of-mass energy $\sqrt{s} = 13$TeV. The two main Feynman diagrams for this process are shown in Figure 4.1. Figure 4.1a is known as the *box* diagram, and Figure 4.1b is known as the *triangle* diagram. The red vertex in Figure 4.1b indicates the trilinear Higgs self-coupling λ, which can be expressed in terms of its modifier

κ_λ. Here the coupling modifier κ is defined as the ratio of the modified coupling to its SM value, i.e. $\kappa = c/c^{\mathrm{SM}}$, hence $\kappa = 1$ denotes the coupling at SM. In SM, these two diagrams interfere destructively, hence the HH production cross-section and kinematic properties have critical dependences on κ_λ. The blue vertex indicates the Higgs-top coupling with the corresponding coupling modifier κ_t. Though it is an important parameter in single Higgs production processes, the dependence of di-Higgs production on it is less direct comparing to κ_λ. Therefore it is not the focus of the analysis.



(a) The box diagram **(b)** The triangle diagram

Figure 4.1: The two leading order ggF di-Higgs production Feynman diagrams [4]

The production process with the second-highest cross-section is vector-boson fusion (VBF), with the value of 1.73 fb at next-to-next-to-next-to-leading order ($\mathrm{N^3LO}$) [114]. Figure 4.2 shows the Feynman diagrams for VBF di-Higgs production at LO. The coupling modifiers, κ_λ, κ_V and κ_{2V} are respectively shown at the HHH, HVV and $HHBB$ interaction vertices. Here V is the gauge vector bosons W or Z. In the SM, due to perturbative unitarity, the divergences in Figure 4.2a and 4.2b cancel out. In BSM scenarios, κ_V and κ_{2V} differ from their SM values, hence they no longer cancel out, which introduces a linear dependence of the cross-section on the effective centre-of-mass energy of the incoming vector bosons [115]. This increase in the energy of Higgs bosons with increasing deviation from the SM continues up to the scale of some new physics, which is necessary to unitarise the total amplitude.

This analysis targets the HH process with $b\bar{b}b\bar{b}$ final state in both ggF and VBF production modes. The data used for the analysis is collected by ATLAS from 2016-2018 during LHC Run 2. With the SM branching ratio of 58.2% for $HH \to b\bar{b}$ [116, 117], approximately one-third of the HH events end up with $b\bar{b}b\bar{b}$ final state. Though this is the most abundant final state, being fully hadronic challenges the

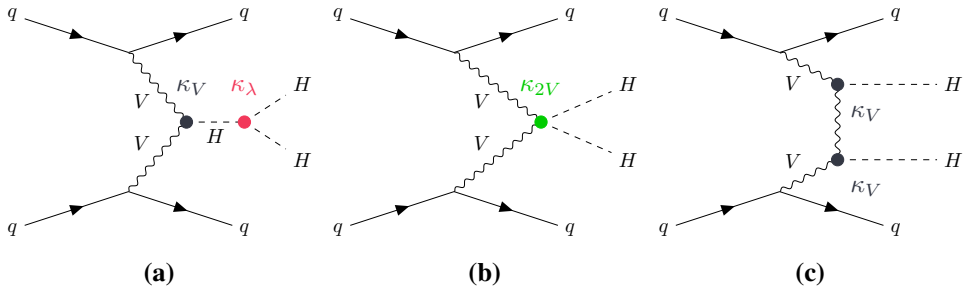


Figure 4.2: The tree-level VBF di-Higgs production Feynmann diagrams [4]

analysis by large backgrounds - mostly from non-resonant QCD production of multiple heavy quark jets and light-quark initiated jets misidentified as coming from heavy quarks.

The results are interpreted in terms of constraints on κ_λ and κ_{2V} with the assumption that $\kappa_V = 1$. The analysis also provides one and two-dimensional constraints on relevant couplings in the SM effective field theory (SMEFT) [63, 64, 62] and Higgs effective field theory (HEFT) [65, 66] frameworks.

As introduced in Section 2.2.1, in the SMEFT framework, the effects of new physics can be described with an effective Lagrangian:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_k c_k^{(6)} O_k^{(6)} \tag{4.1}$$

There is only one dimension-five operator which is ignored since it is only related to lepton and baryon number violation. The mass scale of new physics Λ is set to 1 TeV for this result. The five operators relevant to the HH process and their corresponding coefficients defined in Warsaw basis are listed in Table 4.1. The computation of amplitudes derived from the above Lagrangian involves three components: a pure SM term, a "quadratic" term of order $(1/\Lambda^4)$ including purely new physics, and a "linear" term of order $(1/\Lambda^2)$ captures the interference between SM and new physics. In this analysis, the SMEFT constraints account for both the linear and quadratic new physics terms.

As described in Section 2.2.2, in the HEFT framework, new physics in the electroweak sector is characterised by anomalous coupling involving the Higgs boson. Within this framework, anomalous single-Higgs-boson and HH couplings

Wilson Coefficient	Operator	Interpretation
c_H	$(H^\dagger H)^3$	Modifies the Higgs potential and directly affects the Higgs self-coupling, which can modify di-Higgs production.
$c_{H\Box}$	$(H^\dagger H)\Box(H^\dagger H)$	Modifies the kinetic term of the Higgs field and affects the propagation of the Higgs boson, potentially changing its production rates and decays.
c_{tH}	$(H^\dagger H)(\bar{Q}\tilde{H}t)$	Modifies the Yukawa coupling between the Higgs boson and the top quark, affecting processes like $t\bar{t}H$ production
c_{HG}	$H^\dagger H G^A_{\mu\nu} G_A^{\mu\nu}$	Modifies the coupling between the Higgs boson and the gluon field strength $G^A_{\mu\nu}$, which affects the Higgs-gluon interaction, impacting the ggF Higgs production.
c_{tG}	$(\bar{Q}\sigma^{\mu\nu}T^A t)\tilde{H}G^A_{\mu\nu}$	Introduces a chromomagnetic dipole moment interaction between the top quark and $G^A_{\mu\nu}$, which affects top quark production and Higgs production associated with top quarks.

Table 4.1: The five relevant SMEFT coefficients are their corresponding dimension-6 operators, as defined in Warsaw basis [21, 22]

are defined separately, allowing simplified interpretations of HH processes. The HEFT Lagrangian at LO describes ggF HH production using five key operators and their corresponding Wilson coefficients: $c_{HHH}, c_{t\bar{t}H}, c_{ggH}, c_{ggHH}$, and $c_{t\bar{t}HH}$. In this context, c_{HHH} is equivalent to κ_λ, while $c_{t\bar{t}HH}$ corresponds to the coupling modifier between the Higgs boson and the top quark, κ_t, as shown in fig 4.1a. Setting $c_{t\bar{t}H} = c_{HHH} = 1$ and $c_{ggH} = c_{ggHH} = c_{t\bar{t}HH} = 0$ recovers SM. At NLO, seven HEFT benchmark models (BM) [5] have been defined using cluster analysis [118] to probe a wide variety of characteristic shapes of the m_{HH} spectrum resulting from different BSM scenarios. The values of the coefficients used to define these scenarios are given in Table 4.2, and the corresponding m_{HH} distributions are shown in Figure 4.3. We can see the m_{HH} distributions assemble a few characterising features such as an

enhanced low-m_{HH} region, a double peak, a single peak or an enhanced tail. Some of these features can be easily attributed to a certain coefficient, such as large values of $|c_{HHH}|$ gives an enhanced low-m_{HH} region. Other features such as the double peak or SM-like distribution are harder to attribute to a certain configuration.

Beanchmark Model	c_{HHH}	$c_{t\bar{t}H}$	c_{ggH}	c_{ggHH}	$c_{t\bar{t}HH}$
SM	1	1	0	0	0
BM1	3.94	0.94	1/2	1/3	-1/3
BM2	6.84	0.61	0	-1/3	1/3
BM3	2.21	1.05	1/2	1/2	-1/3
BM4	2.79	0.61	-1/2	1/6	1/3
BM5	3.95	1.17	1/6	-1/2	-1/3
BM6	5.68	0.83	-1/2	1/3	1/3
BM7	-0.10	0.94	1/6	-1/6	1

Table 4.2: The values of the HEFT Wilson coefficients in the SM and in seven BSM benchmark models, as defined in [5]

4.2 Data and Monte Carlo Simulation

The analysis is conducted using pp collision data from the LHC at a centre of mass energy of $\sqrt{s} = 13$ TeV, collected between 2016 and 2018. The dataset, corresponding to an integrated luminosity of 126 fb^{-1}, includes only data collected under stable beam conditions with all relevant detector systems fully operational.

Monte Carlo (MC) simulations are used to model the signal events and to generate background process samples for cross-checks and validation studies. In these simulations, the Higgs boson mass is fixed at 125 GeV. All samples were processed using the ATLAS simulation framework [119], with the detector response simulated by GEANT4 [120].

The ggF signal process was simulated using the POWHEG Box v2 generator [121, 122, 122] at NLO, incorporating finite top-quark-mass effects, using the PDF4LHC15 [123] parton distribution function (PDF) set. Parton showers[1] and hadronization[2] were simulated with PYTHIA 8.244 [124] with the A14 set of tuned pa-

[1]Parton showers: the quarks and gluons produced cannot exist freely, and will begin to emit additional gluons and quark-antiquark pairs as they propagate through the detector.

[2]Hadronization: as partons lose energy through the showering process, they reach a point where they can no longer overcome the strong force and remain free. As a result, they start to combine to

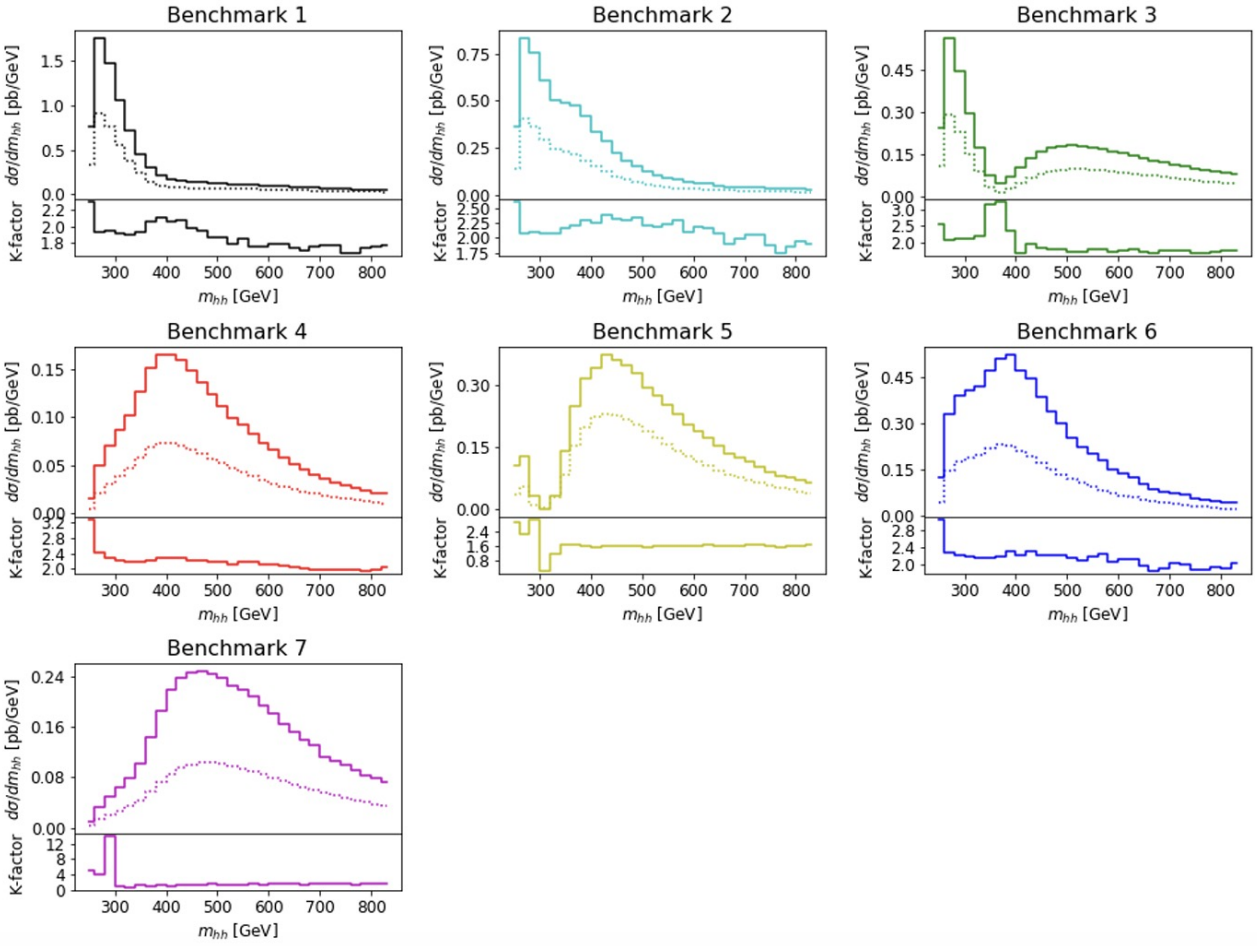


Figure 4.3: The m_{HH} distributions corresponding to the seven benchmark models listed in Table 4.2. The solid curves show the NLO result and the dotted curves show the LO results. The lower panels show the K-factor, which is $d\sigma_{\mathrm{NLO}}/d\sigma_{\mathrm{LO}}$. [5]

rameters [125] and the NNPDF2.3LOPDF set [126]. The SM ggF HH cross-section was taken as $\sigma_{\mathrm{ggF}} = 31.05$fb, calculated at NNLO including finite top-quark-mass effects [113]. Signal samples for the ggF process were explicitly generated for coupling modifier values of $\kappa_\lambda = 1$ and 10, with a reweighting method applied to obtain a ggF signal sample across different κ_λ values [127].

The VBF signal process was simulated using MADGRAPH 2.7.3 [128] at leading order with the NNPDF3.0NLO PDF set [129] , and interfaced with PYTHIA 8.244 for parton showering and hadronization. The VBF HH process cross-section, evaluated at N^{3}LO in QCD, is 1.73 fb in the SM. Signal samples for the VBF process were generated explicitly for $(\kappa_\lambda, \kappa_{2V}, \kappa_V)$=(1,1,1),(1,1.5,1),(2,1,1),(10,1,1),(1,1,0.5),(-

form hadrons.

5,1,0.5),(0,1,1),(1,0,1),(1,3,1). A linear combination of the first six of the listed samples is used to derive distributions with finer granularity for κ_{2V} values, following a technique used previously to generate κ_λ distributions [130].

To validate the background modelling procedure, simulations of top-quark pair production ($t\bar{t}$) and multijet background processes were carried out. The $t\bar{t}$ sample was simulated at NLO in α_s (the strong coupling constant, related to g_s defined in Equation 2.29 by $\alpha_s = \frac{g_s^2}{4\pi}$) using POWHEG Box v2 [131]. The multijet background samples were modeled using PYTHIA 8.235. This simulates pure QCD 2-to-2 interactions at LO in α_s.

Other background processes, such as SM Higgs boson, HH (in other final states) and electroweak diboson production, have been estimated to give negligible contributions and are therefore excluded from this analysis.

4.3 Triggers

The analysis focuses on events that satisfy one of two trigger signatures, each with different requirements on the number of jets and their b-tagging status [108]. The first trigger signature, used to select $b\bar{b}b\bar{b}$ events, requires two b-jet plus one additional jet (”2b1j”), while the second requires two b-jet plus two additional jets (”2b2j”). They are detailed in Table 4.3. See Appendix B for the interpretation of the trigger names. For the 2b2j trigger, all jets must have a minimum transverse energy (E_T) of 35 GeV. In the 2b1j trigger, the b-tagged jets are required to have $E_T > 55$ GeV, while the additional jet must have a minimum E_T between 100 and 150 GeV, depending on the specific year of data collection. The trigger efficiencies for various ggF and VBF HH signals versus m_{HH} are shown in Figure 4.4.

The choice of triggers with 2 b-tagged jets was constrained by the requirements of the background estimation method, which uses 2b events to predict 4b events (described in Section 4.6). Triggers with 3 b-tagged jets were deemed unsuitable for this method, as they disproportionally select 4b events without selecting more 2b events. This results in larger discrepancies between the source and target distributions during the background estimation process. Triggers were selected to maximise the

yield of SM HH events in the $4b$ signal region (defined in Section 4.5.1). The optimal configuration was found to be the combination of $2b1j$ and $2b2j$ triggers which yields the best results with non-resonant SM signal.

Year	Trigger Name	Trigger Type
2016	HLT_j100_2j55_bmv2c2060_split	2b1j
	HLT_2j35_bmv2c2060_split_2j35_L14J15.0ETA25	2b2j
2017	HLT_j110_gsc150_boffperf_split_2j35_gsc55_bmv2c1070_split_L1J85_3J30	2b1j
	HLT_2j15_gsc35_bmv2c1040_split_2j15_gsc35_boffperf_split_L14J15.0ETA25	2b2j
2018	HLT_j110_gsc150_boffperf_split_2j45_gsc55_bmv2c1070_split_L1J85_3J30	2b1j
	HLT_2j35_bmv2c1060_split_2j35_L14J15.0ETA25	2b2j

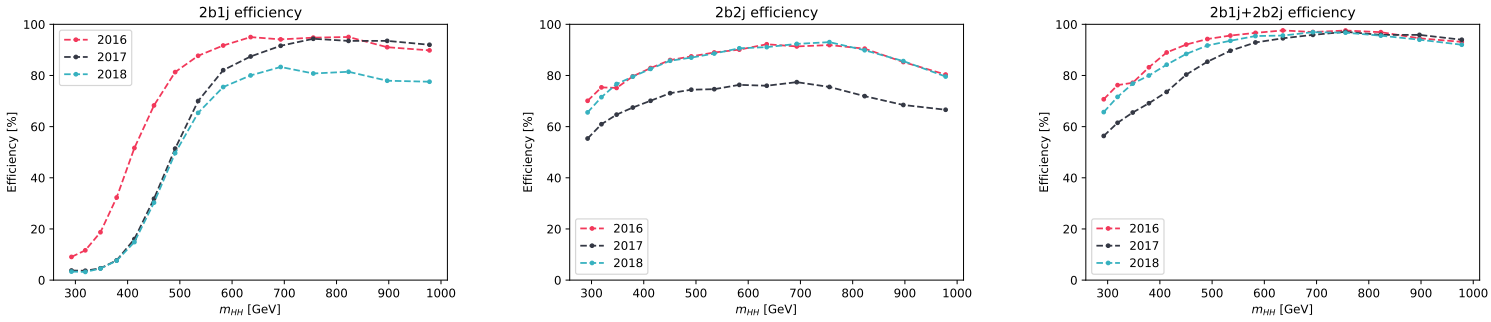
Table 4.3: Triggers used in Run 2 non-resonant searches

Several differences between simulation and data can lead to differences in trigger efficiency. MC scale factors (SFs) are the correction factors applied to MC data to ensure that they accurately represents the real experimental data.

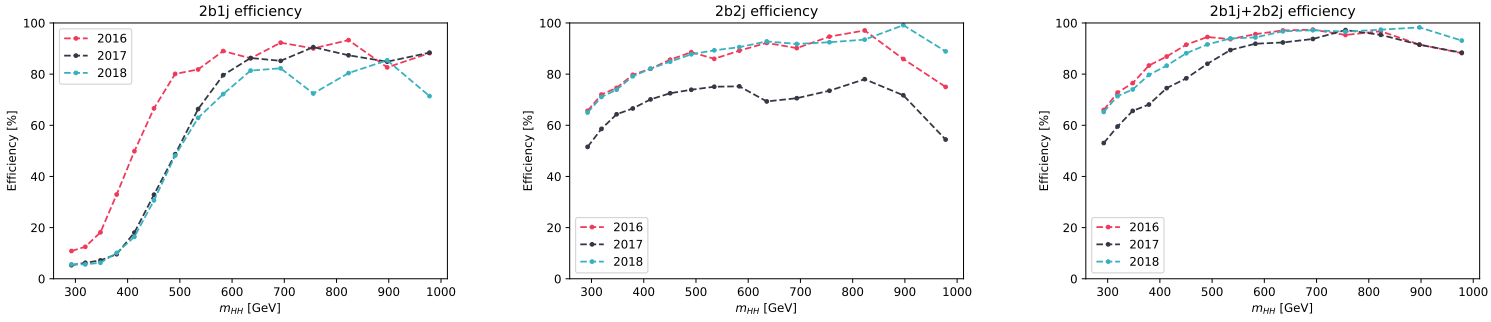
The trigger efficiencies are measured at the jet-level, while we need event-level SFs for our analysis. Hence we need to divide events to each trigger using the "trigger bucket" strategy (illustrated in Figure 4.5) - it splits the events into different categories by applying cuts using fully reconstructed offline variables, and only a single trigger is used to determine whether or not to keep an event within each category. These offline cuts are designed to enhance sensitivity to different regions of the m_{HH} spectrum across various trigger buckets. Additional selections based on offline kinematic quantities are introduced on top of the trigger bucket selection. Events are selected if they have a leading jet with $p_T > 170$ GeV, a third leading jet with $p_T > 70$ GeV, and pass the 2b1j trigger; alternatively, events that fail either of these jet-p_T requirements but pass the 2b2j trigger are also selected. This selection step keeps about 90% of signal efficiency, allowing for reliable calculation of simulation-to-data correction factors for estimating trigger efficiency in the remaining HH4b signal events, depending on which of the above two trigger classes they belong to.

The corresponding breakdown of the analysis signal yields by trigger bucket is illustrated in Figure 4.6. The 2b1j bucket targets the high m_{HH} spectrum, and the 2b2j bucket improves sensitivity to the low m_{HH} events.

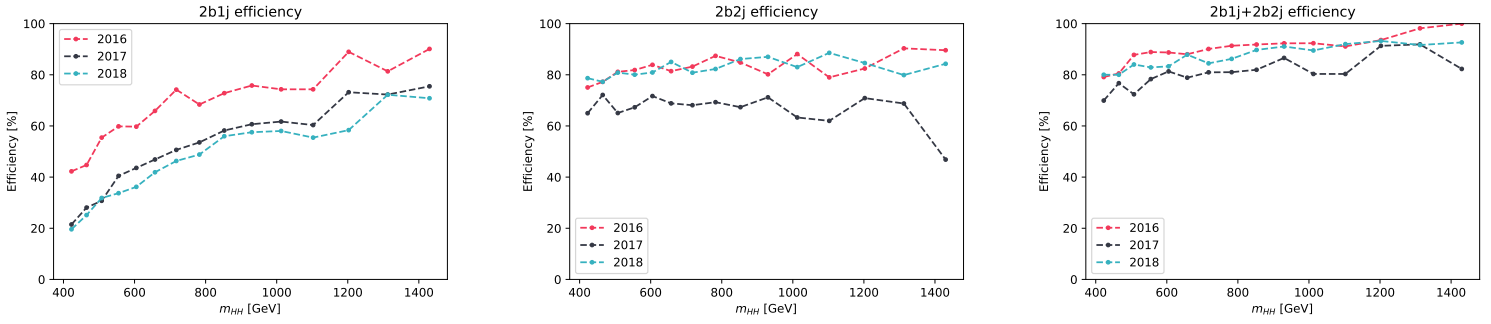
We derived the correction to both the online b-tagging efficiency and the correc-



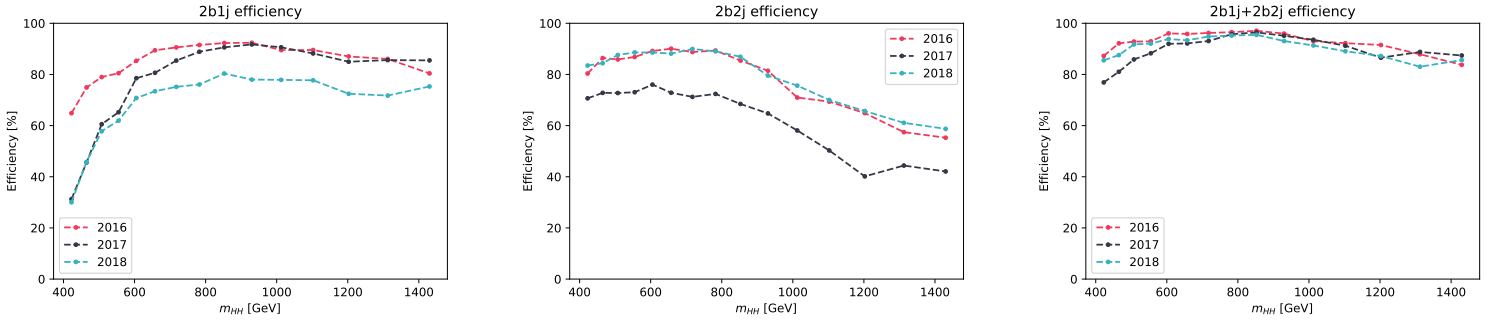
(a)SM ggF $\kappa_\lambda = 1$



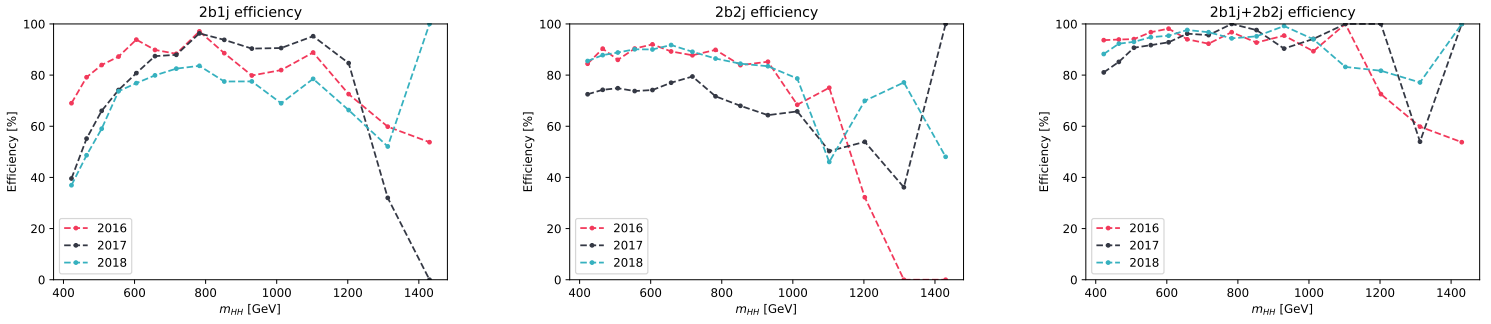
(b)SM ggF $\kappa_\lambda = 10$



(c)SM VBF $\kappa_\lambda = 1, \kappa_{2V} = 1$



(d)SM VBF $\kappa_\lambda = 1, \kappa_{2V} = 0$



(e)SM VBF $\kappa_\lambda = 10, \kappa_{2V} = 1$

Figure 4.4: Trigger efficiencies of the 2b1j, 2b2j and combined for the MC samples for years 2016-2018 for various HH signals as a function of m_{HH}. The bulk of the signal is around 500 GeV. A couple of bins in the high mass region are taking extreme values due to low statistics. The significantly lower efficiency for 2017 2b2j is due to tighter b-tagging requirements.[6]

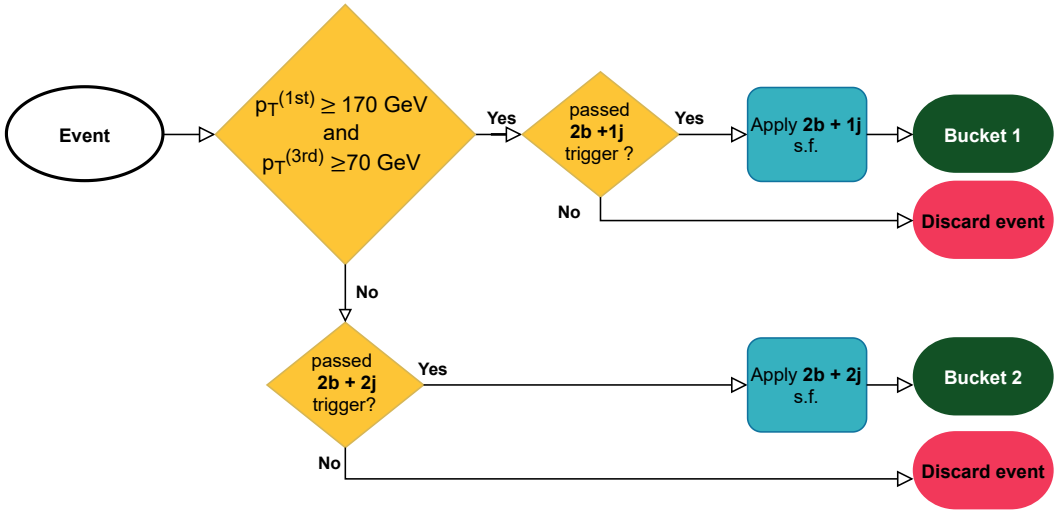
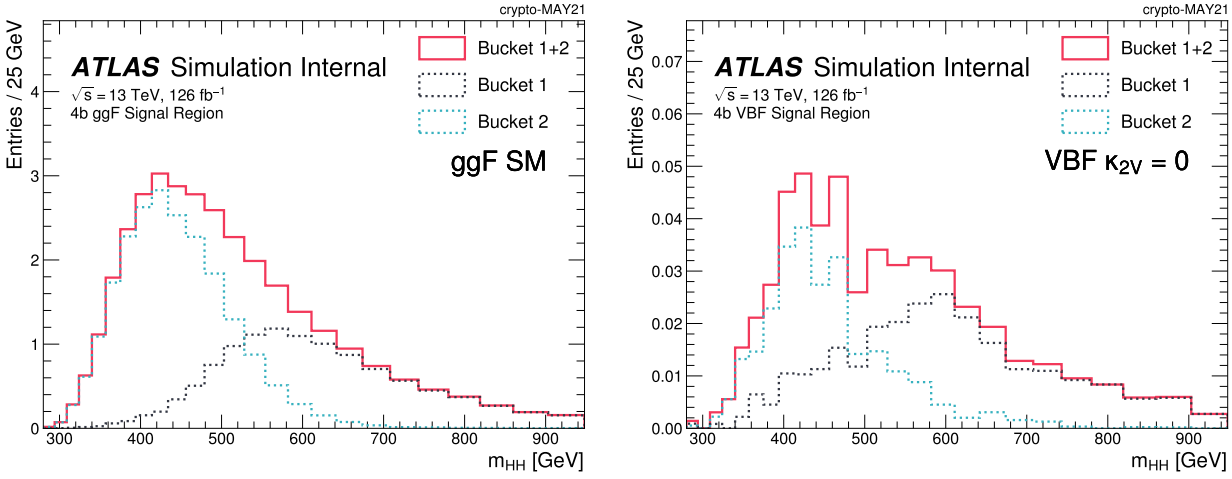


Figure 4.5: Trigger bucket strategy. [6]



(a) SM ggF **(b)** $\kappa_{2V} = 0$ VBF

Figure 4.6: The bucket comparison of m_{HH} for SM ggF and $\kappa_{2V} = 0$ VBF HH MC simulation in the 4b signal region. Bucket 1 corresponds to 2b1j trigger and Bucket 2 corresponds to the 2b1j trigger. [4]

tion to the offline flavour tagging efficiency. The two SFs are multiplied together on the signal MC events.

4.4 Object Reconstruction

Primary vertices from pp interactions are reconstructed [132] using a minimum of two charged-particle tracks with $p_T > 500$ MeV, measured by the Inner Detector. The vertex with the largest sum of squared track momenta ($\sum p_T^2$) is designated as

the hard-scatter primary vertex.

Hadronic jets are reconstructed using the anti-k_t algorithm [133, 134] with radius parameter $R = 0.4$. For this analysis, jets are clustered from particle-flow (PFlow) objects [135], which are charged-particle tracks matched to the hard-scatter vertex and calorimeter energy clusters, after applying an energy subtraction algorithm to remove calorimeter deposits associated with good-quality tracks from any vertex.

To suppress jets from pile-up processes, jets with $p_T < 60$ GeV and $|\eta| < 2.4$ are required to pass a Jet Vertex Tagger (JVT) [136] cut to distinguish jets coming from primary vertex to pileups. The tight working point is used with 96% efficiency for hard scatter jets.

Additionally, jets with $R = 0.4$ are reconstructed from topological clusters of energy deposits in the calorimeter [137] and are calibrated in the same manner as jets reconstructed from PFlow objects. These jets are only used to apply quality criteria to identify events that may be due to calorimeter noise or noncollision backgrounds [138]. Events containing at least one such jet with $p_T > 20$ GeV that meets the JVT requirements, but fails these quality checks are rejected.

b-jets are identified using the NN based DL1r algorithm [139], which is applied to all jets with $|\eta| < 2.5$. The DL1r working point selected for this analysis has a 77% efficiency for jets associated with true b-hadrons in simulated $t\bar{t}$ events. At this working point, the rejection rates for light-jet (charm-jet) measured in $t\bar{t}$ simulation is about a factor 130 (4.9). The DL1r algorithm is calibrated separately for each jet type [140, 141], with correction factors derived and applied to simulated samples to account for discrepancies between b-tagging efficiencies in data and simulation.

Muons are reconstructed by matching ID tracks with either Muon Spectrometer (MS) tracks or aligned individual hits in the MS, and performing a combined track fit [142]. They are required to have $p_T > 4$ GeV and $|\eta| < 2.5$, and must meet the ”Medium” identification criteria based on track-quality variables. In this analysis, Muons are only used to apply energy corrections to jets.

Additionally, a momentum correction is applied to b-jets to compensate for energy lost due to soft out-of-cone radiation, as well as to account for energy

carried away by muons and neutrinos in semileptonic b-hadron decays, following the procedures in [143].

4.5 Analysis Selection

The analysis uses a specific set of criteria to identify HH4b candidates, including requirements to separate events into orthogonal ggF and VBF regions.

We have "forward" and "central" jets classified as:

- central jets: $|\eta| < 2.5$ and $p_T > 40$ GeV. This $|\eta|$ region corresponds to the coverage of the Inner Detector which provides precise tracking information, allowing efficient b-jet identification.
- forward jets: $.25 < |\eta| < 4.5$ and $p_T > 30$ GeV. This $|\eta|$ region corresponds to the forward calorimeters, where tracking information is limited or absent, thus the b-tagging is less effective.

The event selection can be illustrated in Figure 4.7.

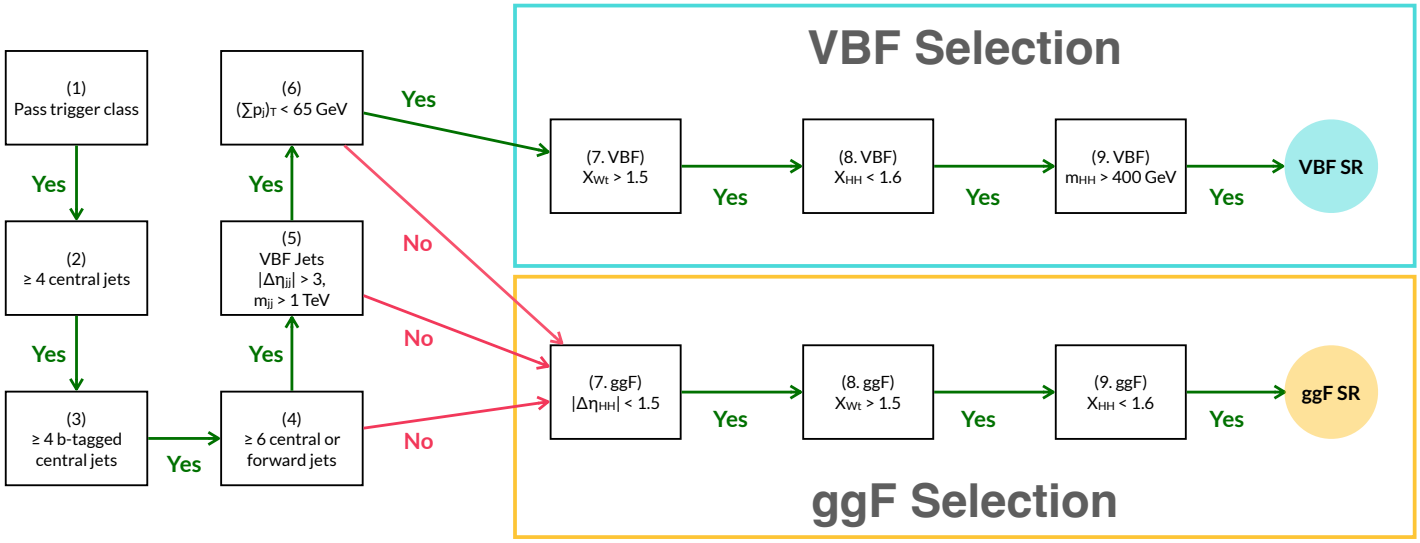


Figure 4.7: A flowchart summarizing the nine selection criteria used for the VBF and ggF analysis selections. Events must satisfy criteria 1-3 to be considered for each analysis signal region. Events failing to satisfy any of the selection criteria 4-6 are considered for inclusion in the ggF signal region, while those satisfying criteria 4-6 are considered for VBF signal region. [4]

The selections can be summarised as:

(1) Events must pass trigger bucket selection as described in 4.3.

(2) Events must contain **at least four central jets**.

(3) Among these central jets, we must have **at least 4 central jets with DL1r 77% Working Point (WP)** as b-tag selection.

The four highest-p_T b-tagged jets are chosen to reconstruct the decays of the two Higgs bosons. In approximately 75% of simulated signal events that reach this stage, these four jets can be each matched (within $\Delta R < 0.3$) to one of the four b-quarks from the HH decays. In cases where this matching fails, it is typically due to one of the b-quarks from the Higgs boson decays producing a jet that falls outside the analysis acceptance range. See Table 4.4 for the different b-tag categories used in this analysis.

Notation	Definition	Usage
2b	Exactly two central jets with DL1r 77% WP	Background Estimation
3b1f	Exactly three central jets tagged with DL1r 77% WP, and no central jets passing the 85% WP	Background estimate systematics
4b	At least four central jets tagged with DL1r 77% WP	Signal region

Table 4.4: Different b-tagged categories and their use

From the selected four b-tagged jets, three possible pairings can be made to form the two Higgs candidates, as shown in Figure 4.8. The analysis selects the pairing in which the jet pair with the highest p_T has the smallest ΔR separation. This pairing method leads to a relatively smooth distribution of the expected background in the plane of the invariant masses of the two Higgs boson candidates, which makes the background estimation described in Section 4.6 easier.

(4)-(6): These cuts are used to separate ggF and VBF events:

(4) Since VBF events characteristically produce two high-energy jets in addition to jets from Higgs decays, **events targeting VBF must contain at least two additional jets**, which can be either central or forward, excluding any b-tagged jets.

(5) The two jets with the highest invariant mass (m_{jj}) are chosen as "VBF" jets. These jets correspond to the t-channel exchange of a vector boson and typically have a significant rapidity gap and a high invariant mass. **A cut of $m_{jj} > 1$ TeV and**

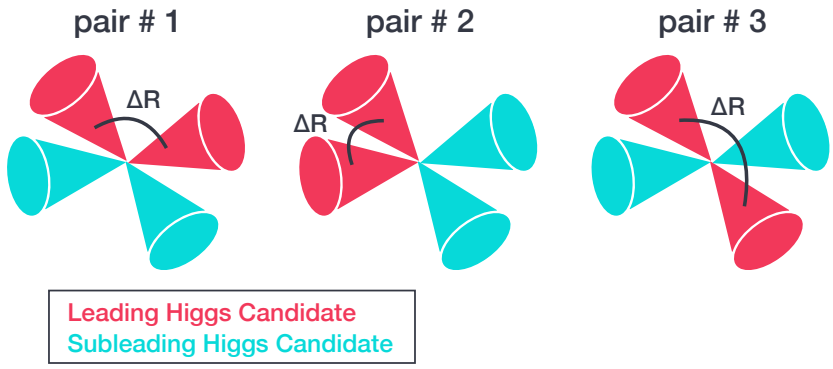


Figure 4.8: The three possible pairing permutations of the four HH jets into the two Higgs candidates. The opening angles between the jets in the leading Higgs Candidate are indicated, we select pair No.2. [4]

$|\Delta\eta_{jj}| > 3$ is applied to isolate VBF events.

(6) In VBF events, the Higgs bosons are produced centrally with minimal additional QCD radiation, leading to balanced transverse momentum distribution. Therefore for VBF events, **the vector sum of the transverse momentum of the two VBF jets and the four jets forming the Higgs candidates should be less than 65 GeV**.

(7) and (8) for ggF, and (7) for VBF: Background reduction and top veto:

To suppress the background contamination in the ggF signal region, events are required to have reconstructed Higgs bosons with a **pseudorapidity separation of** $|\Delta\eta_{HH}| < 1.5$. Figure 4.9 shows the $|\Delta\eta_{HH}|$ distribution for ggF di-Higgs signal and blinded data[3] in the ggF channel immediately after pairing. We can see the background events tend to have higher values of $|\Delta\eta_{HH}|$ than signals, thus this cut can help improve the signal purity. This selection is not applied for VBF SR, as SM VBF HH signal events already tend to have a larger $|\Delta\eta_{HH}|$.

Since top quark decays almost exclusively into a W boson and a b-quark ($t \to Wb$), and the W-boson can further decay into a pair of quarks leading to a dijet system, or into a lepton or neutrino. $t\bar{t}$ events can produce a final state with jets

[3]The events in the Signal Region are removed. The blinded $4b$ data is a good approximation of background.

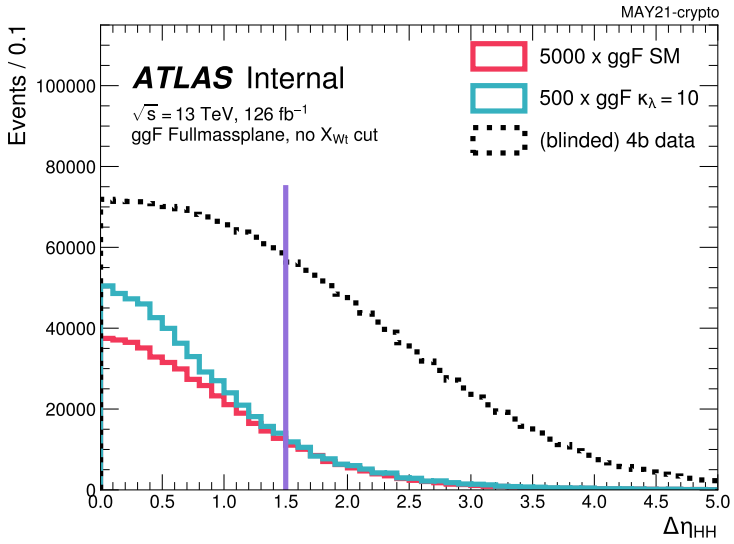
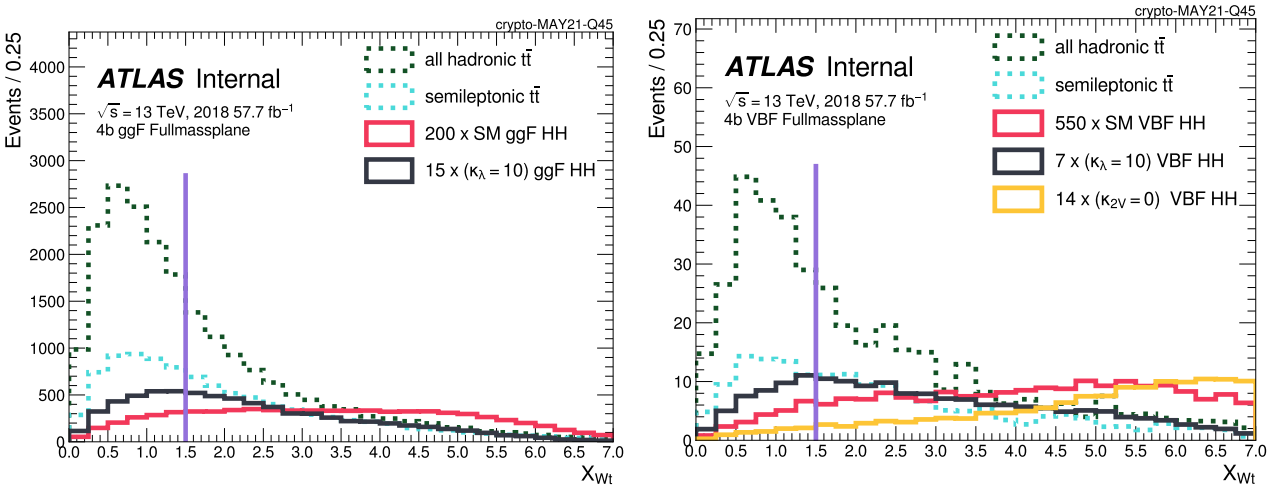


Figure 4.9: The $|\Delta\eta_{HH}|$ distributions for SM ggF HH MC simulation and blinded data in the ggF channel. The solid purple line indicates the $|\Delta\eta_{HH}| < 1.5$ cut applied to the ggF selection. Events to the right of the line are discarded. [6]

and b-jets similar to the signals, so it's important to suppress their contribution. A top-veto discriminant X_{W_t} is defined as:

$$X_{W_t} = \min\left[\sqrt{\left(\frac{m_{jj}-m_W}{0.1m_{jj}}\right)^2 + \left(\frac{m_{jjb}-m_t}{0.1m_{jjb}}\right)^2}\right] \tag{4.2}$$

where $m_W = 80.4$ GeV and $m_t = 172.5$GeV are the nominal W boson and top quark masses, and m_{jj} and m_{jjb} represent the invariant masses of W boson and top quark candidates formed from jet combinations in each event. The "minimum" refers to the minimum value from all possible jet combinations (of one b-tagged jet and two additional untagged jets) that would give a W boson candidate and a corresponding top candidate. The factor of 0.1 in the denominators is chosen to approximate the experimental di-jet mass resolution. The X_{W_t} discriminant quantifies the likelihood that an event contains a hadronic top quark decay. **Events with $X_{W_t} < 1.5$ are rejected**, reducing the $t\bar{t}$ background by approximately a factor of 2 in simulated events, with a small signal efficiency loss of around 15%, and a similar reduction in the non-$t\bar{t}$ multijet background. Figure 4.10 shows the distributions of X_{W_t} for $t\bar{t}$ backgrounds and signals in 2018 as an example.



(a) ggF **(b)** VBF

Figure 4.10: X_{W_t} distributions in 2018 dataset. The solid purple lines indicates the X_{W_t}¿1.5 cut applied to both ggF and VBF channels. Events to the left of the lines are rejected. [6]

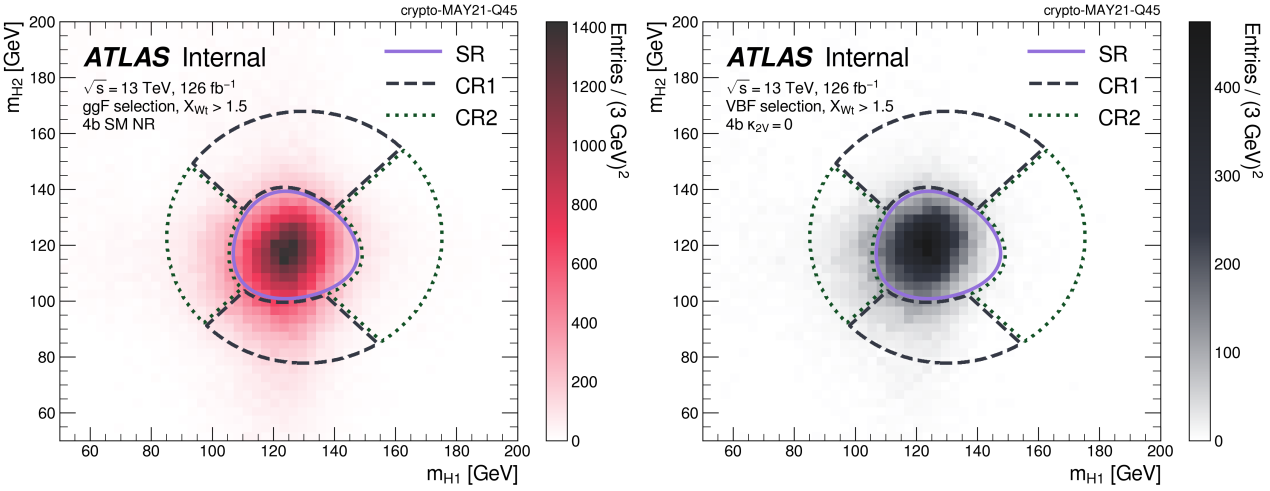
(9) ggF and (8) VBF: Signal Region (SR) definition:

We define

$$X_{HH} = \sqrt{\left(\frac{m_{H1} - 124\text{GeV}}{0.1 m_{H1}}\right)^2 + \left(\frac{m_{H2} - 117\text{GeV}}{0.1 m_{H1}}\right)^2} \tag{4.3}$$

where m_{H1} and m_{H2} are the masses of the leading and subleading reconstructed Higgs boson candidates respectively. The values of 124 and 117 GeV are chosen according to the centres of m_{H1} and m_{H2} distributions for correctly paired signal events from simulation. The reconstructed masses differ from the true mass of 125 GeV due to factors like detector effects (such as jet energy resolution and b-tagging efficiency), pileups and jet-pairing efficiencies. The leading Higgs candidate is formed from the jet pair with higher p_T while the sub-leading Higgs candidate comes from jets with lower p_T, the kinematic differences can influence the mass reconstruction process, resulting in different masses. **Events are required to have** $X_{HH} < 1.6$ **to be included in the signal region (SR) of the analysis**. Figure 4.11 shows the distributions of SM ggF signals, and κ_{2V} VBF signals as an example. We can see the signal events are nicely peaking inside the defined SR.

(10) VBF: **In VBF SR, only events with** $m_{HH} > 400$ **GeV are considered** since



(a) 4b ggF SM signal

(b) 4b VBF $\kappa_{2V} = 0$ signal

Figure 4.11: Selected Higgs Candidate signal mass planes. CR1 and CR2 are the control regions and will be defined in Section 4.5.1. [6]

the background in the lower m_{HH} region was found to be inadequately modelled by the data-driven method from validation studies.

4.5.1 Kinematic Region Definition

Besides the Signal Region defined above, we also define two control regions that are used to perform background modelling (detailed in Section 4.6): Control Region 1 (CR1) and Control Region 2 (CR2). The boundaries are outlined in Table 4.5. For the CR outer edge, there is shift of the centre by a factor of 1.05 relative to $X_{HH} = 0$. This shift provides an optimal trade-off between achieving a good number of events outside the SR and avoiding the low m_{H1}/m_{H2} regions, where the discrepancies between the $2b$ and $4b$ kinematic distributions are larger. CR1 is used to perform background modelling, and CR2 is used to derive a systematic uncertainty associated with the method. These control regions are situated within a band enclosed by the inner and outer CR edges, and are designed to be orthogonal to the SR.

The band is divided into quadrants, defined by 45° and 135° lines passing through the SR centre of (124, 117) GeV, creating four sectors of roughly equal area. The four quadrants are given labels based on compass directions: the upper quadrant Q_N, the lower Q_S, the left Q_W, and the right Q_E. CR1 and CR2 each consist of two quadrants ($Q_N + Q_S$ =CR1 and $Q_E + Q_W$ =CR2), with the quadrants paired such

that they are positioned on opposite sides of the band, as illustrated in Figure 4.12. We can see there is minimal 4b data in the control regions, and minimal 2b data in the signal region. We can also see two straight light-coloured bands centred around 80 GeV on both the x and y-axes, stretching horizontally and vertically across the plot. These bands result from the X_{W_t} cut applied in the selection, which functions as a W-mass veto for the reconstructed Higgs Candidates. This leads to a noticeable drop in the number of events where m_{H1} or m_{H2} are near 80GeV, corresponding to the mass of the W boson.

Figure 4.13 shows the mass planes for the unblinded reconstructed Higgs boson candidates. Figure 4.14 shows the distributions of the simulated ggF and VBF HH signals as a comparison.

SR	$X_{HH} < 1.6$
CR Inner Edge	$X_{HH} = 1.6$
CR Outer Edge	$\sqrt{(m_{H1} - 1.05 \cdot 125\text{GeV})^2 + (m_{H2} - 1.05 \cdot 117\text{GeV})^2} = 45\text{GeV}$

Table 4.5: Edges of CR and SR

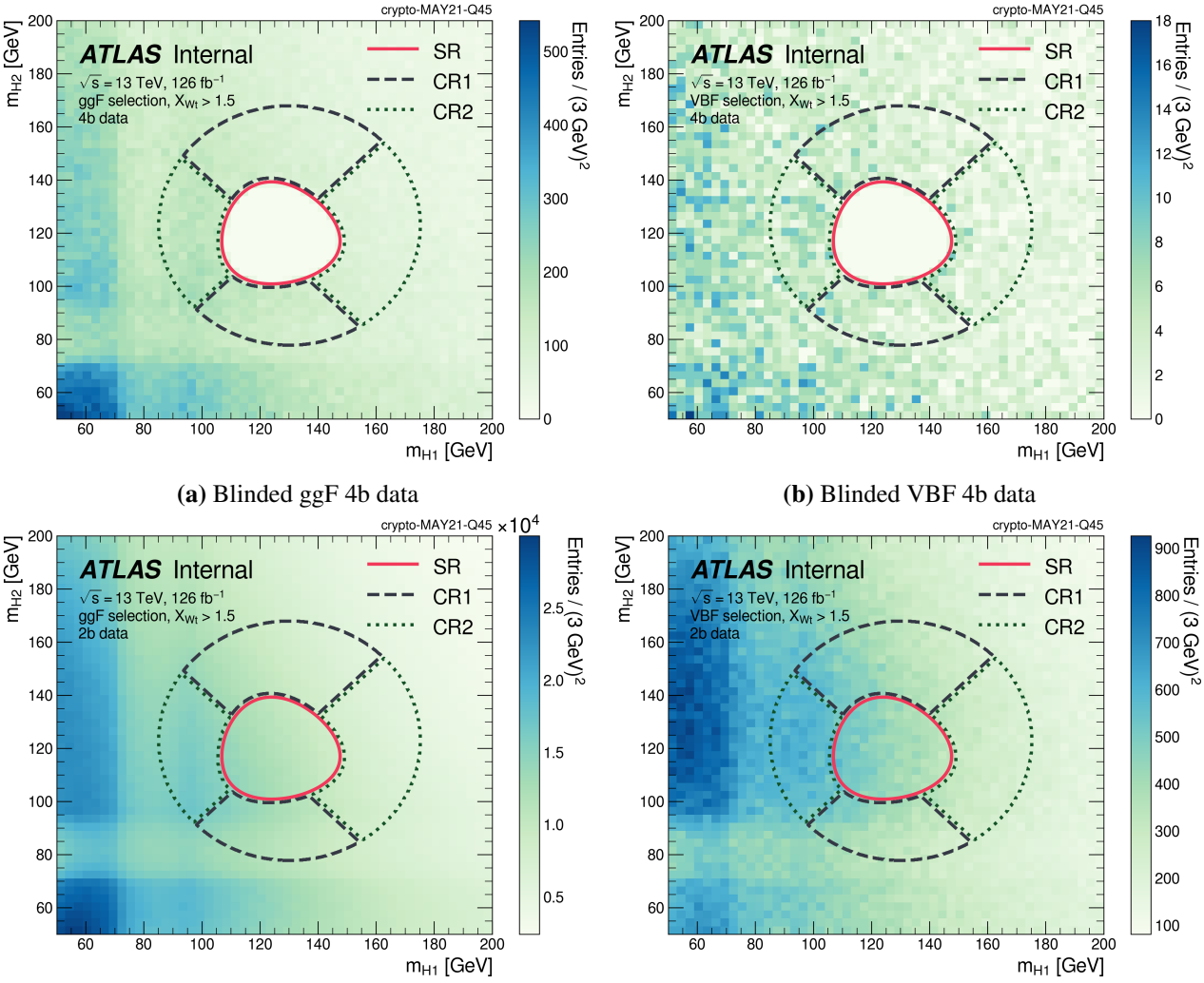
4.6 Background Estimation

After event selection, approximately 90% of the background events originate from multijet processes (excluding top quark production), with the remaining 10% predominantly consisting of $t\bar{t}$ events. A fully data-driven technique is employed to model the background.

4.6.1 Neural Network Background Modelling

As outlined in Table 4.4, events containing exactly two b-tagged jets (2b events) are used for background estimation. This sample is roughly two orders of magnitude larger than the 4b sample, meaning the presence of any HH4b signal in it is negligible. In 2b events, the jets used to reconstruct the two Higgs candidates are the two b-tagged jets and the two untagged central jets with the highest p_T (excluding the VBF jets in the VBF category).

The kinematic properties of the 2b and 4b events are not expected to be identical, due to different processes contributing to the two samples, as well as differences in



(a) Blinded ggF 4b data

(b) Blinded VBF 4b data

(c) ggF 2b data

(d) VBF 2b data

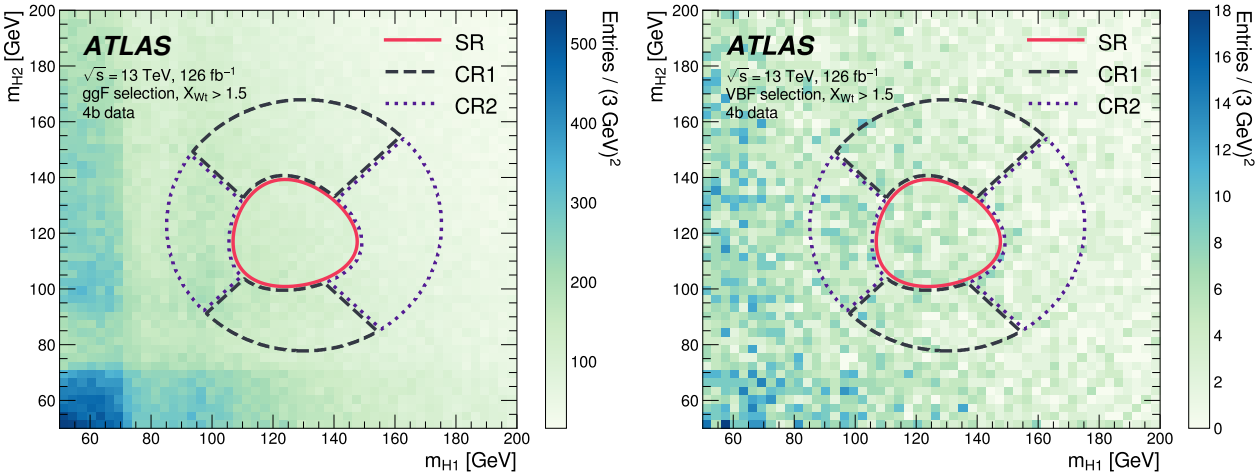
Figure 4.12: The blinded 4b and 2b data mass planes. [6]

trigger acceptance, and the variance in the probability of tagging a b-jet with respect to jet p_T and η. Therefore, a reweighting function is required. This function, when applied to the $2b$ events, maps their kinematic distributions onto the corresponding $4b$ events. The reweighting function is derived using the $2b$ and $4b$ events in Control Regions described in Section 4.5.1.

The reweighting function takes the form

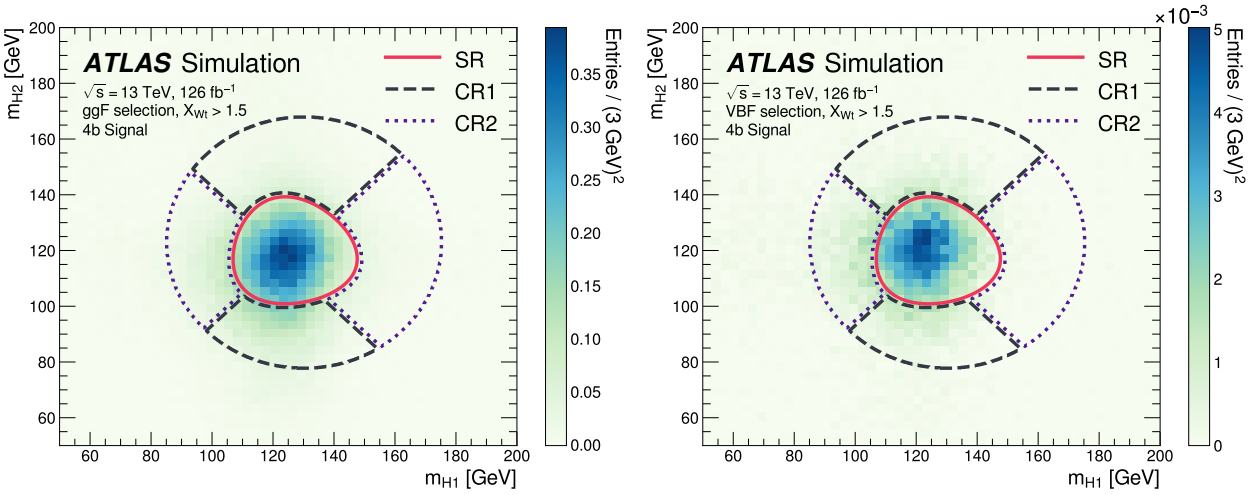
$$w(\vec{x}) = \frac{p_{4b}(\vec{x})}{p_{2b}(\vec{x})} \tag{4.4}$$

where $p_{4b}(\vec{x})$ and $p_{2b}(\vec{x})$ are probability density functions for $4b$ and $2b$ data, respec-



(a) ggF **(b)** VBF

Figure 4.13: The mass planes of the reconstructed Higgs candidates for (a) ggF and (b) VBF signal regions of the analysis, shown for the $4b$ data events. [4]



(a) ggF **(b)** VBF

Figure 4.14: The mass planes of the reconstructed Higgs boson candidates for (a) simulated ggF HH signal in the ggF SR and (b) simulated VBF HH signal in the VBF SR of the analysis. [4]

tively, over a set of kinematic variables $\vec{x}$. The task of computing $w(\vec{x})$ is a density ratio estimation problem. In this analysis, we employ a neutral network (NN) trained on data from the CR1 $2b$ and $4b$ events (and from CR2, but only for systematic uncertainties evaluation). The loss function for the training is

$$\mathcal{L}(w(\vec{x})) = \int d\vec{x} \left[\sqrt{w(\vec{x})} p_{2b}(\vec{x}) + \frac{1}{\sqrt{w(\vec{x})} p_{4b}(\vec{x})} \right] \tag{4.5}$$

We can show that

$$\arg\min_{w} \mathcal{L}(w(\vec{x})) = \frac{p_{4b}(\vec{x})}{p_{2b}(\vec{x})} \tag{4.6}$$

i.e. the reweighting function Eq 4.4 minimises the loss function in Eq4.5. The kinematic variables used: $\vec{x}$ are listed in Table 4.6. These variables are chosen based on the larger differences between the $2b$ and $4b$ events.

ggF	**VBF**
1. $\log(p_T)$ of the 2nd leading Higgs boson candidate jet	1. Maximum dijet mass from the possible pairing of the four Higgs boson candidate jets
2. $\log(p_T)$ of the 4th leading Higgs boson candidate jet	2. Minimum dijet mass from the possible pairings of the four Higgs boson candidate jets
3. $\log(\Delta R)$ between the closest two Higgs boson candidate jets	3. Energy of the leading Higgs boson candidate
4. $\log(\Delta R)$ between the order two Higgs boson candidate jets	4. Energy of the subleading Higgs boson candidate
5. Average absolute η value of the Higgs boson candidate jets	5. Second-smallest ΔR between the jets in the leading Higgs boson candidate (from the three possible pairings for the leading Higgs candidate)
6. $\log(p_T)$ of the di-Higgs system	6. Average absolute η value of the four Higgs boson candidate jets
7. ΔR between the two Higgs boson candidates	7. $\log(X_{W_t})$
8. $\Delta\phi$ between jets in the leading Higgs boson candidate	8. Trigger class index as one-hot encoder
9. $\Delta\phi$ between jets in the subleading Higgs boson candidate	9. Year index as one-hot encode (for years inclusive training)
10. $\log(X_{W_t})$	
11. Number of jets in the event	
12. Trigger class index as one-hot encoder	

Table 4.6: Input variables used for $2b$ to $4b$ reweighting in the ggF and VBG channels.

For the ggF SR, the NN has three fully connected hidden layers, each with 50 nodes and a rectified linear unit (ReLU) activation function [144], and single-linear output. A similar architecture is used for VBF NN, though with only 20 nodes per hidden layer due to the smaller sample sizes in this region - nearly two orders of magnitude smaller than ggF SR. This size difference is also the reason of using all data-taking years together for the VBF NN training, with the year encoded as a one-hot feature. For the ggF SR, separate reweightings are derived for each year, allowing the model to better account for year-to-year variations, especially those

arising from changes in trigger conditions.

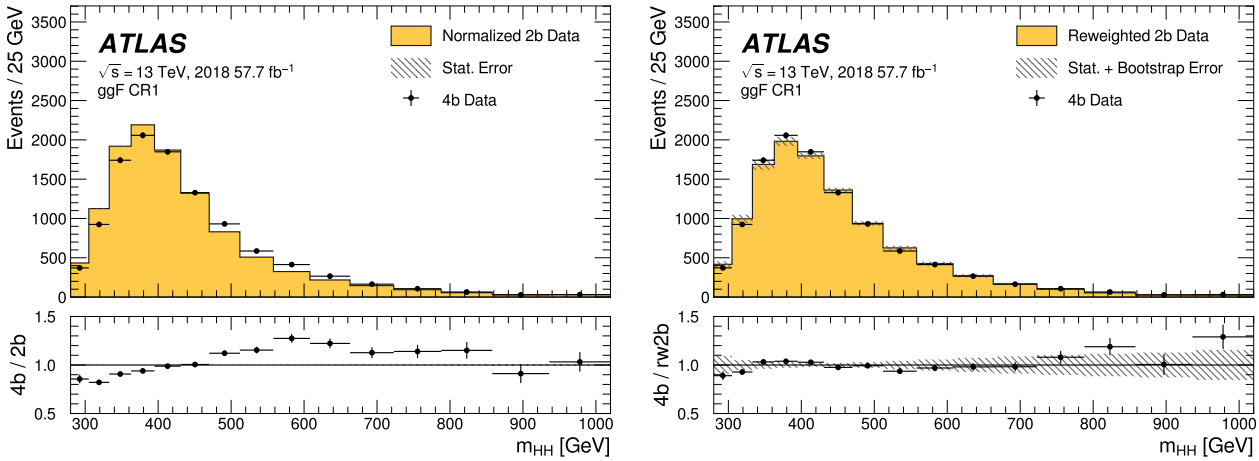
To account for the effects of varying initial conditions and the limited size of training samples in the NN, a deep ensemble technique [145] is employed in combination with bootstrap resampling [146] of the training data. This method involves constructing multiple training datasets by sampling with replacements from the original dataset. In this analysis, since the training set is large, the resampling procedure can be simplified through the relation $\lim_{n\to\infty} \text{Binominal}(n, 1/n) = \text{Poisson}(1)$, i.e. we can approximate the sampling with replacements by assigning different random training weights to each event, which follow a Poisson distribution with $\mu = 1$. The NN is trained independently on each dataset, using different initial conditions. This process results in an ensemble of reweighting functions; in our analysis, there are 100 of them.

Each reweighting function is further scaled by a normalisation factor so that the number of reweighted $2b$ events matches the number of $4b$ events in the region where the NN was trained. The nominal background estimate is produced using the mean of these weights for each event, and the variation across these predictions is used to calculate a systematic uncertainty, reflecting the stability of the NN training process.

The effect of the reweighting in CR1 is shown in Figure 4.15 for ggF m_{HH} distribution and Figure 4.16 for VBF X_{W_t} distribution. The reweighted $2b$ distributions show agreement with the corresponding $4b$ distributions within approximately 10% across most of the phase space, though some larger deviations are observed in bins near the distribution tails, where event counts are lower. Additionally, the agreement between the reweighted $2b$ and $4b$ distributions was examined for other kinematic variables for validation purposes. For all variables, the level of agreement, quantified using the χ^2 metric, either improves after the reweighting or, for variables where the $2b$ and $4b$ distributions are already similar, exhibited only minor changes.

4.6.2 Analysis Categories and binning

$|\Delta\eta_{HH}|$ is a strong discriminating variable: it is typically smaller for ggF events, and larger for VBF events due to additional forward jets characteristic of VBF processes;



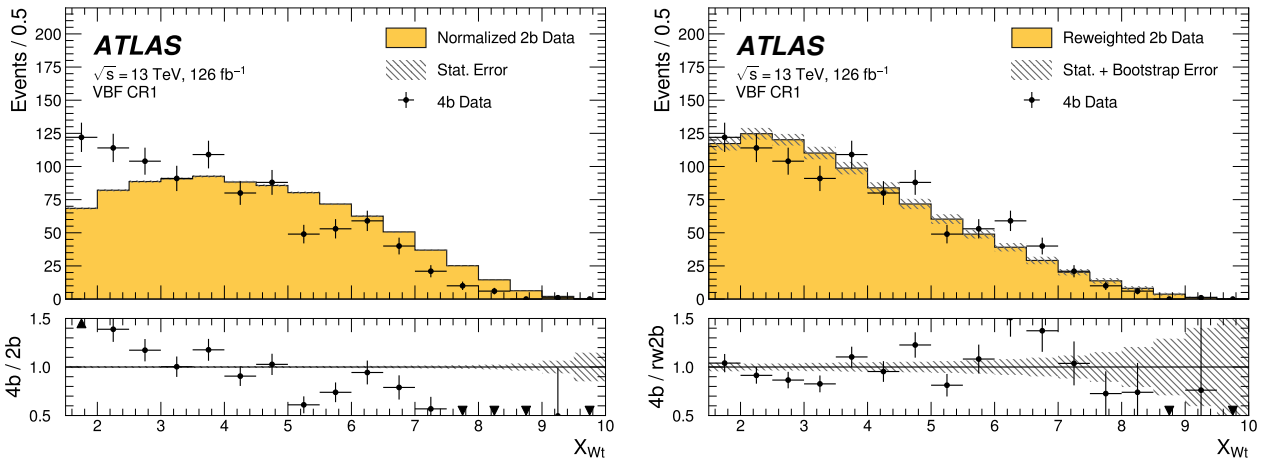
(a) Before reweighting **(b)** After reweighting

Figure 4.15: Comparison of the $2b$ and $4b$ m_{HH} distribution in CR1 of the ggF SR from 2018. The error bars indicate the statistical uncertainty of the $4b$ data, and the hatching indicates the statistical uncertainty of the $2b$ data. In (a), the statistical uncertainty is only the Poisson uncertainty, while in (b) it also includes bootstrap uncertainty. [4]

also, the background events normally have broader and flatter $\Delta\eta$ distribution, and signal events tend to have more central $|\Delta\eta_{HH}|$. X_{HH} is a discriminant variable to separate SR and CR. By using X_{HH} to categorise events, the analysis can focus on events with a higher likelihood of being signal with small X_{HH}. Events are divided into several categories based on $|\Delta\eta_{HH}|$ and X_{HH}, each with different signal-to-background ratios.

For ggF channels, both $|\Delta\eta_{HH}|$ and X_{HH} are used as categorisation variables. Specific cuts have already been applied: $|\Delta\eta_{HH}| < 1.5$ for QCD background rejection and $X_{HH} < 1.6$ for SR definition. The distributions of these variables of background predictions, 4b SM, and 4b $\kappa_\lambda = 10$ are shown in Figure 4.17. Given the flat distribution of $|\Delta\eta_{HH}|$ in the background, the events are divided into three equally spaced bins between 0 and 1.5. In addition, two X_{HH} bins are used, with a boundary at 0.95, chosen to equally split the signal events.

The VBF analysis only has a single categorisation based on $|\Delta\eta_{HH}|$. The boundary of 1.5 is selected to balance maximising significance and maintaining accurate background modelling within the categories. As shown in Figure 4.18, the non-SM coupling distribution peaks near $|\Delta\eta_{HH}| = 0$, while the SM prediction peaks



(a) Before reweighting **(b)** After reweighting

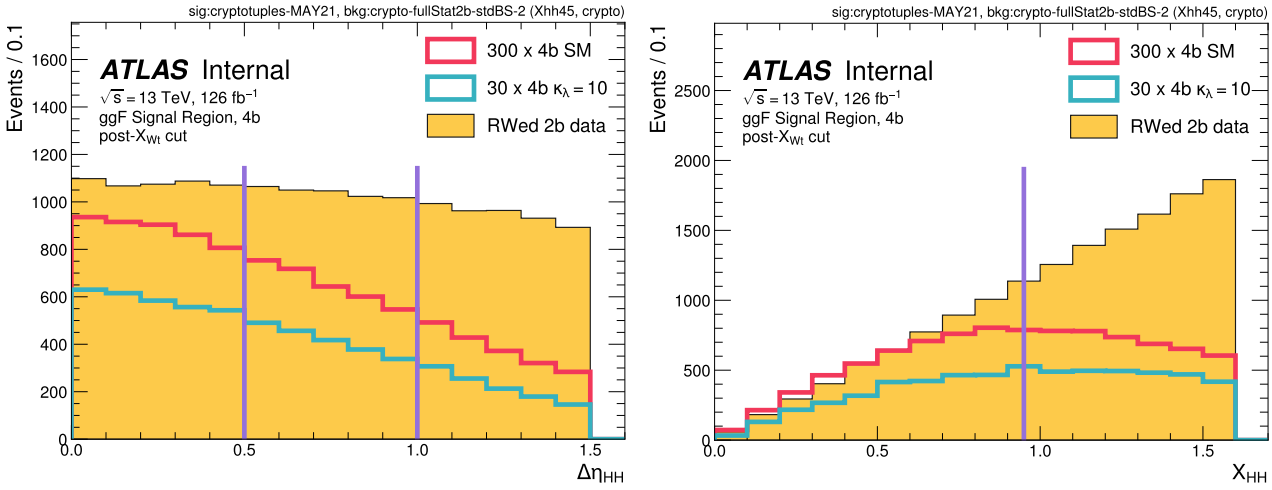
Figure 4.16: Comparison of the $2b$ and $4b$ X_{W_t} distribution in CR1 of the VBF SR. The error bars indicate the statistical uncertainty of the $4b$ data, and the hatching indicates the statistical uncertainty of the $2b$ data. In (a), the statistical uncertainty is only the Poisson uncertainty, while in (b) it also includes bootstrap uncertainty. [4]

around $|\Delta\eta_{HH}| = 2$. As a result, the $|\Delta\eta_{HH}| < 1.5$ category drives the sensitivity to non-SM couplings, whereas the $|\Delta\eta_{HH}| > 1.5$ category is more sensitive to the SM prediction.

To ensure a sufficient number of $4b$ events for training in both the ggF and VBF regions, the training is performed inclusively before categorizing events. Both $|\Delta\eta_{HH}|$ and X_{HH} are found to be insensitive to the kinematic reweighting, so inclusive training does not introduce any bias when events are later divided into the relevant categories.

Since the m_{HH} distribution is steeply falling, variable width histogram binning is employed, with narrower bins at low m_{HH} distribution and wider bins at higher values. This approach allows the analysis to benefit from high m_{HH} events while maintaining reasonable statistical uncertainties within these bins. After choosing the lowest bin edge, the second bin edge is set at (100+X%)$\times$ the lowest bin edge, where X is the specified percentage parameter. Subsequent bin edges are defined similarly, with bins of increasing width added until an upper threshold is reached.

For ggF, the lower and upper thresholds are 280 and 950, with X=9. Initially, for the lower $|\Delta\eta_{HH}|$ VBF category, the minimum bin edge is 280, the highest bin edge



(a) $4b$ ggF $|\Delta\eta_{HH}|$ **(b)** $4b$ ggF X_{HH}

Figure 4.17: Distributinos of the variables used for categorisation in the ggF channel, with all years merged. The solid purple lines indicate the boundaries of the categories. [6]

is 890, with X=10. For the higher $|\Delta\eta_{HH}|$ VBF category, the minimum bin edge, maximum bin edge, and X are 290, 1470 and 9 respectively. Later, due to the poor modelling of the background in lower m_{HH} region, only events with $m_{HH} > 400$ GeV are considered (the (10)th VBF selection criteria in Section 4.5). These parameters were optimised to ensure that the relative error on the quadrature sum of the bootstrap and $2b$ Poisson components of the background model remains below 30%, while avoiding overly wide bins that could reduce the sensitivity of the analysis.

4.6.3 Validation on Background Modelling

Since the reweighting is performed in CR1, which is kinematically different from SR, we need to assess how well CR-derived weights extrapolate into SR. Hence we define a number of alternative control and signal regions, away from the true CR and SR, to test the extrapolation ability.

4.6.3.1 Reverse $|\Delta\eta_{HH}|$

As described in Section 4.5, for the ggF channel, we impose the cut $|\Delta\eta_{HH}| < 1.5$ to suppress the QCD background. Hence inverting this cut, but with all other selection criteria applied, can give us a signal-depleted region to validate the background modelling. The distributions of the $4b$ events in 2018 in this region for the ggF

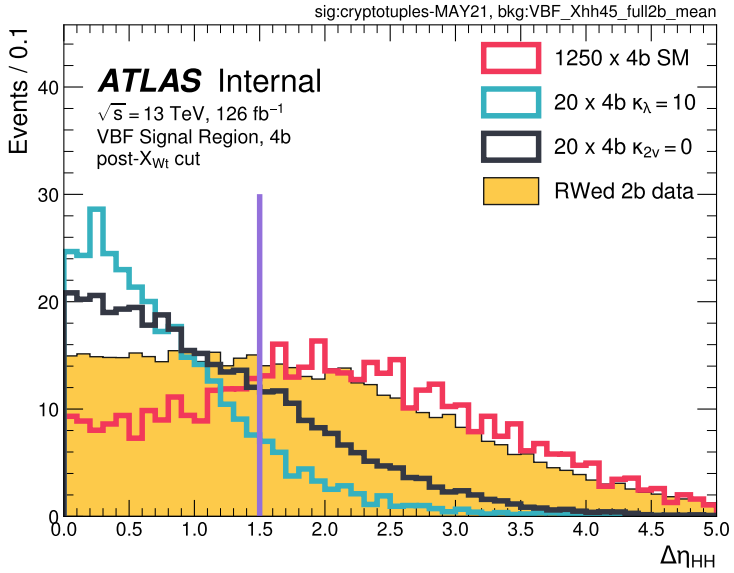


Figure 4.18: Distributions of $|\Delta\eta_{HH}|$ for signal MC with SM, $\kappa_\lambda = 10, \kappa_{2V} = 0$, and the background estimate in the VBF SR. The solid purple line indicates the boundary of the categories. [6]

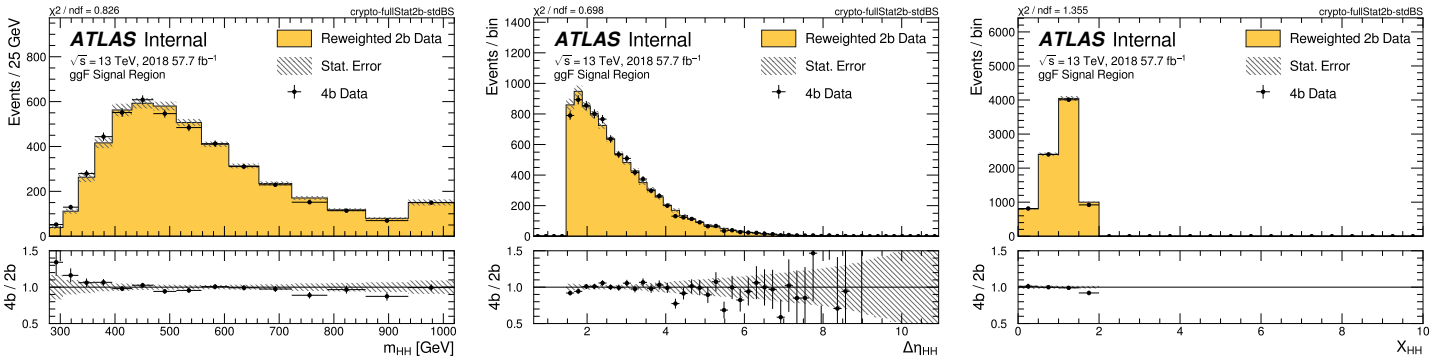
discriminating variables are shown in Figure 4.19. To validate the analysis with the ggF categories, we choose the $\Delta|\eta_{HH}|$ boundaries such that we have equal background statistics between each of the regions:

- $1.5 < |\Delta\eta_{HH}| < 2.5$
- $2.5 < |\Delta\eta_{HH}| < 3.6$
- $|\Delta\eta_{HH}| > 3.6$

The reweighted $2b$ background is consistent with $4b$ data in the kinematic signal region of this reversed $|\Delta\eta_{HH}|$ validation region, with any differences fully covered by uncertainties. The validation plot for 2018 with reweighted $2b$ in SR is shown in Figure 4.20 as an example.

4.6.3.2 Shifted Regions

We define five **shifted regions** identical to the true Signal and Control Regions but in different positions in the $m_{H1} - m_{H2}$ plane. The positions are chosen to be close



(a) m_{HH} **(b)** $|\Delta\eta_{HH}|$ **(c)** X_{HH}

Figure 4.19: Distributions for the discriminating variables ($m_{HH}, |\Delta\eta_{HH}|, X_{HH}$) for the ggF 4b 2018 reversed $|\Delta\eta_{HH}|$ samples. [6]

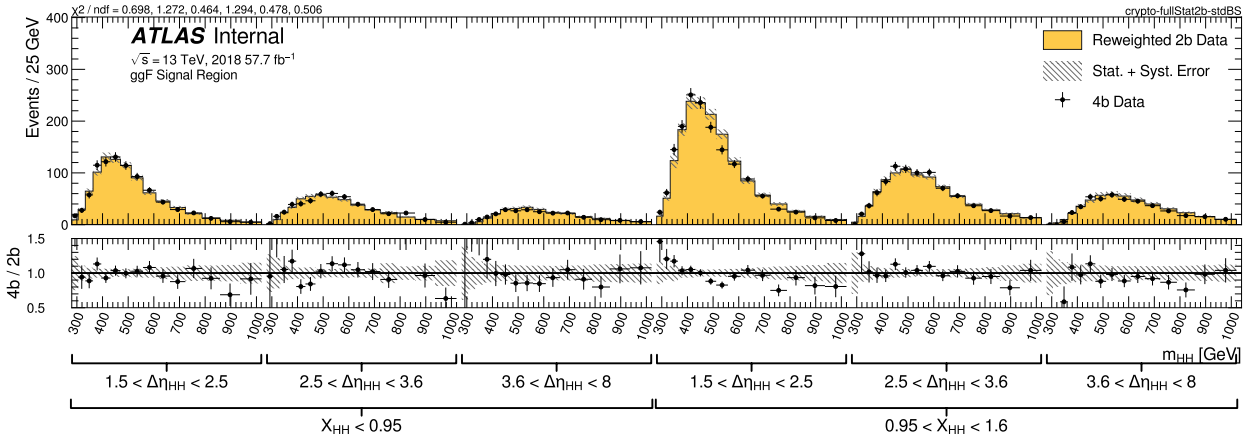


Figure 4.20: 2018 4b ggF reweighted $2b$ data and $4b$, reversed $|\Delta\eta_{HH}|$. The discrepancies between the reweighted $2b$ and $4b$ are larger in the lowest m_{HH} bins, causing some $4b$ data points to fall outside the plot range. However, this is primarily attributed to limited statistics, and the validation studies permit more lenient requirements. [6]

to the true SR but do not overlap with it. The SR of each shifted region is defined by

$$X_{HH,\text{shift}} = \sqrt{\left(\frac{m_{H1} - m_{H1,\text{centre}}}{\sigma_{m_{H1}} m_{H1}}\right)^2 + \left(\frac{m_{H2} - m_{H2,\text{centre}}}{\sigma_{m_{H2}} m_{H2}}\right)^2} \tag{4.7}$$

which is similar to the nominal SR, but the centre values and resolutions of m_{H1} and m_{H2} are adjusted to avoid overlapping with the nominal SR. The centre values, $m_{H1,\text{centre}}$ and $m_{H2,\text{centre}}$, are shown in Table 4.7. The resolutions are set to 1, except for upper centre $\sigma_{m_{H1}} = 0.1 \times \frac{124}{m_{H1,\text{centre}}}$ and centre right $\sigma_{m_{H2}} = 0.1 \times \frac{117}{m_{H2,\text{centre}}}$, where we make the regions a little smaller. The SR cut is set at $X_{HH,\text{shift}} < 1.6$, the same as

the nominal cut.

Shifted Region	$m_{H1,\text{centre}}$[GeV]	$m_{H2,\text{centre}}$[GeV]
Upper Left	78	166
Upper Centre	124	180
Upper Right	170	166
Centre Right	188	117
Lower Right	170	68

Table 4.7: m_{H1} and m_{H2} centre for the shifted regions, they are named according to their positions relative to the true SR in the $m_{H1} - m_{H2}$ plane.

The Control Regions for the shifted regions are defined similarly to the nominal Control Regions, but using centre values shown in Table 4.7. The boundaries are shown in Figure 4.21. For the Lower Right region, CR2 is used to derive background estimation instead of CR1 since CR1 extend below the acceptance threshold for m_{H2}, resulting in distinct kinematic differences, and is used to derive shape systematic uncertainty instead. Due to the acceptance threshold for m_{H1} and m_{H2}, the shape systematic uncertainties in the Lower Right and Upper Right regions are larger than others.

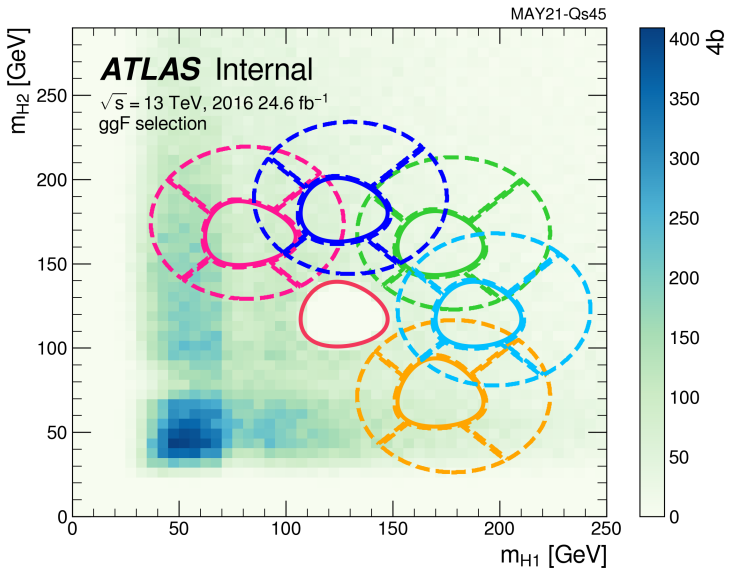
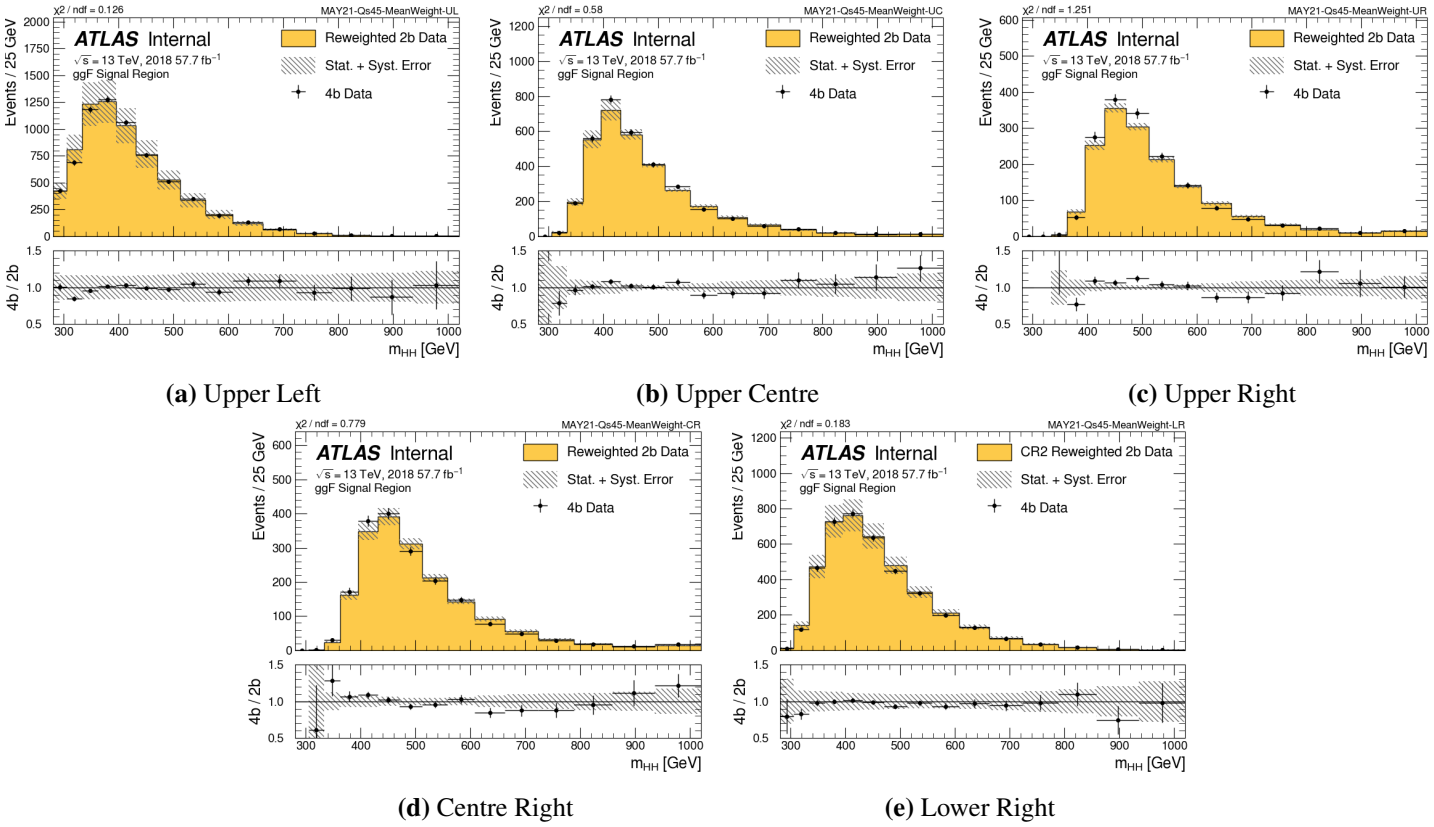


Figure 4.21: Shifted Regions. The red solid closed curve in the centre is the nominal SR. In clockwise order, we have: Upper Left (pink), Upper Centre (blue), Upper Right (green), Centre Right (light-blue), Lower Right (orange). [4]

The standard background estimation with the same training setup is conducted

on the CRs. Since the SRs of the shifted regions contain negligible signals and can be unblinded, the background estimate can be validated directly by comparing it with the data in SR. Figure 4.22 shows the 4*b* background estimation results in the shifted regions SR for ggF 2018 as an example. Good agreement is observed in both ggF and VBF channels, suggesting that the weights derived in the CR can be successfully extrapolated into SR.



(a) Upper Left **(b)** Upper Centre **(c)** Upper Right

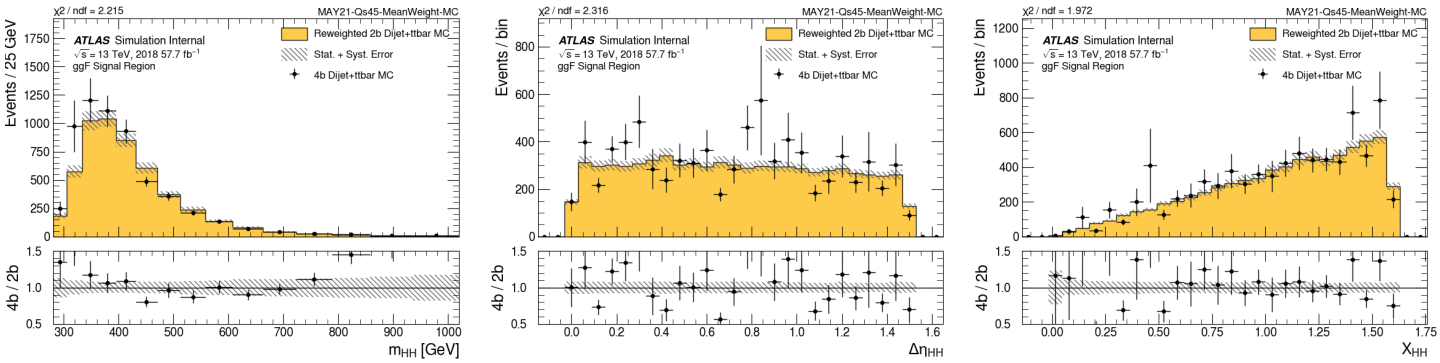
(d) Centre Right **(e)** Lower Right

Figure 4.22: m_{HH} distribution of reweighted 2*b* data and 4*b* data in SR in the shifted regions in 2018 ggF channel. The background error includes 2*b* Poisson statistic error, bootstrap error and shape systematics. [4]

4.6.3.3 QCD and $t\bar{t}$ MC

The two main sources of background in this analysis are QCD multijet processes and $t\bar{t}$ production. To verify the NN used to derive the data-driven background estimate also accurately models these background contributions, MC simulated QCD di-jet and $t\bar{t}$ samples were used in place of data when assessing the NN model. Note the NN itself was still trained on data.

Figure 4.23 shows the agreement of 4b events and reweighted 2b events in SR for 2018 as an example, both from MC. The MC QCD multijet and $t\bar{t}$ samples are combined by weighting them according to their respective cross-sections and the integrated luminosity of the dataset, and the resulting distribution is again normalised to match the total number of observed events in the data. Good agreement is observed across the years for m_{HH}, $|\Delta\eta_{HH}|$ and X_{HH}.



(a) m_{HH} **(b)** $\Delta\eta_{HH}$ **(c)** X_{HH}

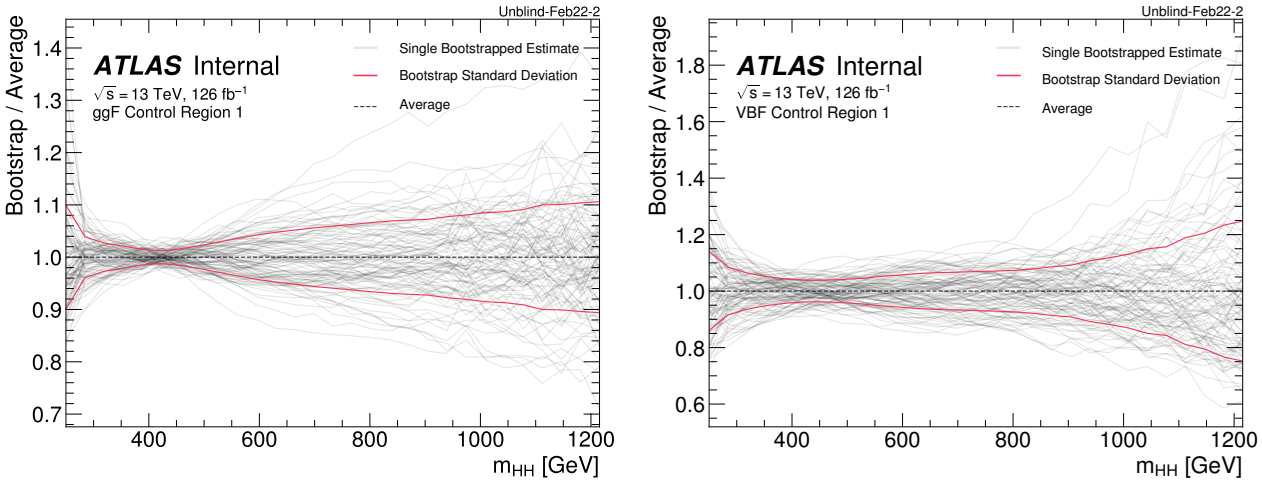
Figure 4.23: The m_{HH}, $|\Delta\eta_{HH}|$ and X_{HH} distributions in reweighted 2b and 4b events in QCD and $t\bar{t}$ MC simulation. The error in background estimation is 2b Poisson statistic error, bootstrap error and shape systematic error. [4]

4.7 Systematic Uncertainties

The uncertainties with the most significant impact on the analysis sensitivity come from the data-driven background estimate detailed in Section 4.6. These uncertainties primarily arise from two sources: the limited sample sizes in the CR and SR, and the physical differences between the CR, where the 2b reweighting function is derived, and the SR, where it is applied.

Bootstrap Uncertainty As described in Section 4.6, the ensemble of 100 reweighting functions generates 100 distinct background predictions. From each of these predictions, an m_{HH} histogram can be constructed, and the standard deviation of the predictions in each bin is taken as the bootstrap uncertainty. This uncertainty is considered uncorrelated across m_{HH} bins. Figure 4.24 demonstrates how the uncertainty envelope is calculated from this method, showing the variation in the m_{HH} histogram heights of the bootstraps vs nominal distributions in CR1, for ggF

and VBF channel and all years inclusive.



(a) ggF **(b)** VBF

Figure 4.24: Illustration of the bootstrap band procedure, shown as a ratio to the nominal estimate. Each grey line is from the m_{HH} prediction for a single bootstrap training, and the solid red line is the standard deviation. [4]

Poisson Statistical Uncertainty An additional source of statistical uncertainty arises from the limited sample size of the 2b SR dataset. A Poisson uncertainty is computed in the following way: for a given bin i, the background histogram value is $n_i = \sum_{j \in i} w_j$, where w_j represents the weight of event j that falls within bin i. The uncertainty is then calculated as $\delta n_i = \sqrt{\sum_{j \in i} w_j^2}$, which simplifies to the standard $\sqrt{N}$ Poisson error when all weights $w_j = 1$.

Background Shape Systematic Uncertainty To account for the kinematic differences between the SR and CR1 in the background estimate, an additional uncertainty component is evaluated using alternative predictions derived from the CR2 region. The SR is divided into four quadrants - each covering approximately equal area - with the alignment of the quadrants matching those that define the control regions. These quadrants are labeled according to their relative positions in the Higgs Candidate mass plane - Q_N(north), Q_S(south), Q_E(east) and Q_N(west), as shown in Figure 4.25. Four alternative background estimates are produced by applying CR1-derived weights to three of the SR quadrants, while CR2-derived weights are applied to the remaining SR quadrant. For instance, one alternative estimate is obtained by applying CR1-derived weights to Q_S, Q_E, and Q_W, while CR2-derived weights are

used for Q_N.

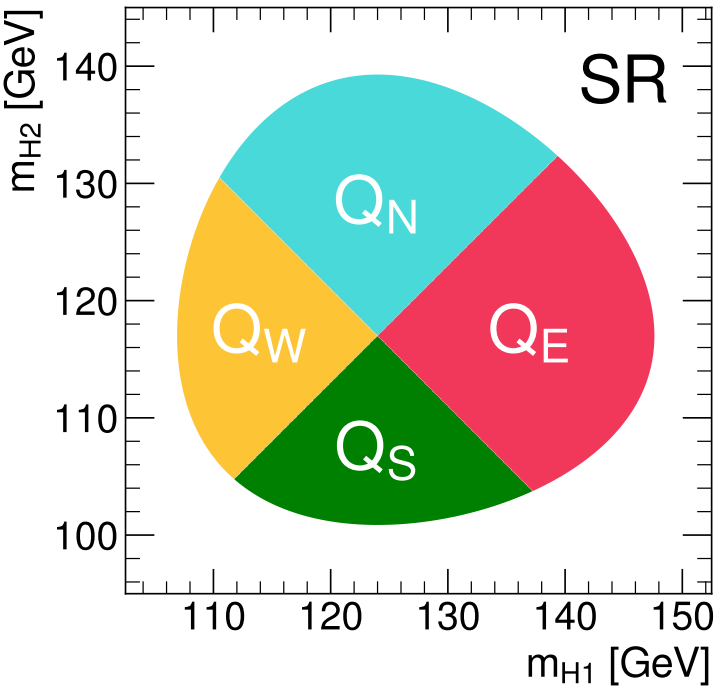


Figure 4.25: SR quadrants chosen to derive the four background variation nuisance parameters. [4]

This procedure minimises the constraints associated with a single nuisance parameter (NP) if the entire CR2 weight set is treated as a single uncertainty. Each of the four background predictions is then symmetrised around the nominal m_{HH} distribution to construct a two-sided uncertainty. Figure 4.26 illustrates the variation in Q_E for the 2018 ggF channel as an example.

Since the m_{HH} distribution varies across the four SR quadrants, replacing the CR1-based prediction with the CR2-based prediction in each quadrant separately and using a four-component uncertainty approach provides the fit model with greater flexibility to describe the m_{HH} variable with finer granularity. In the ggF SR, these uncertainties are treated as uncorrelated across three years. In both ggF and VBF SRs, the uncertainty is considered to be correlated across the different analysis categories (described in Section 4.6.2.)

3b1f Non-Closure Uncertainty An additional closure uncertainty is estimated by applying the full background modelling procedure to the $3b1f$ sample. The predicted

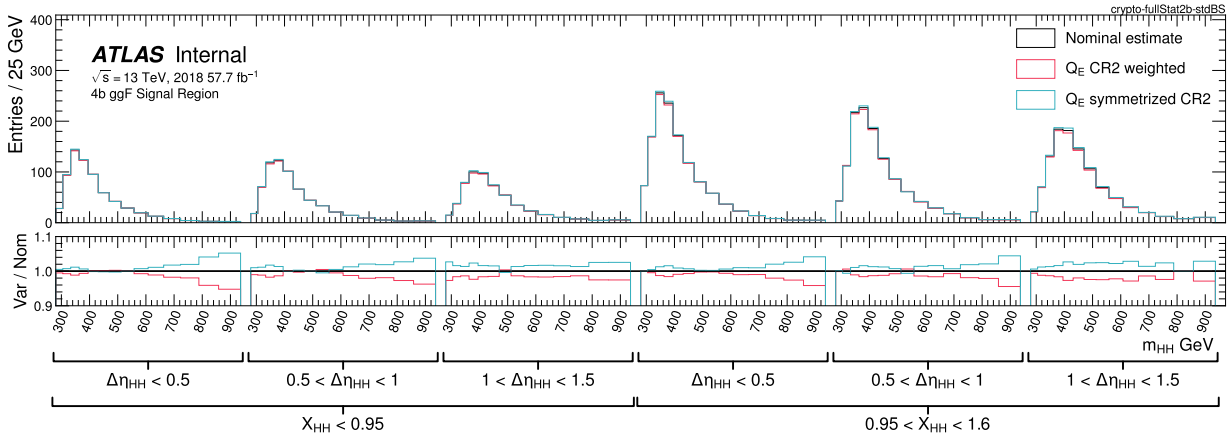
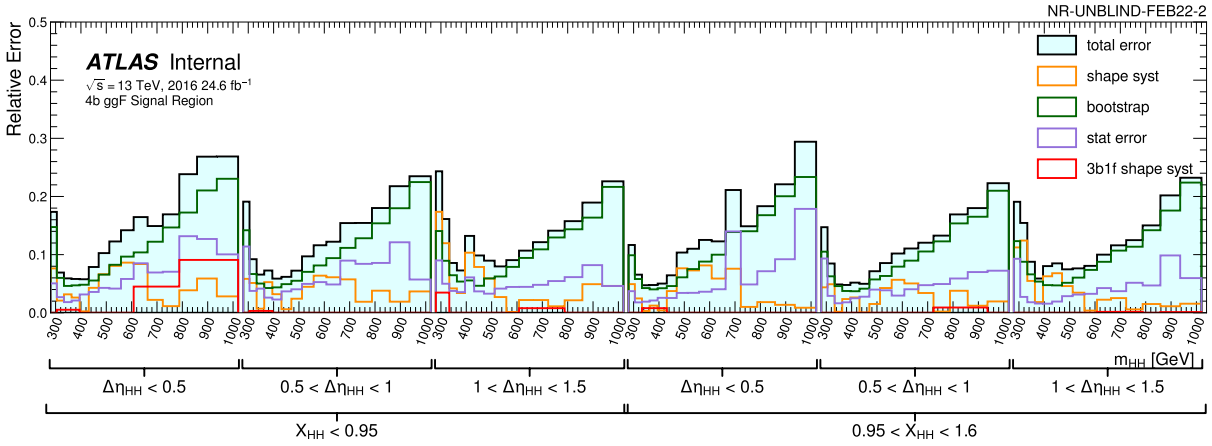


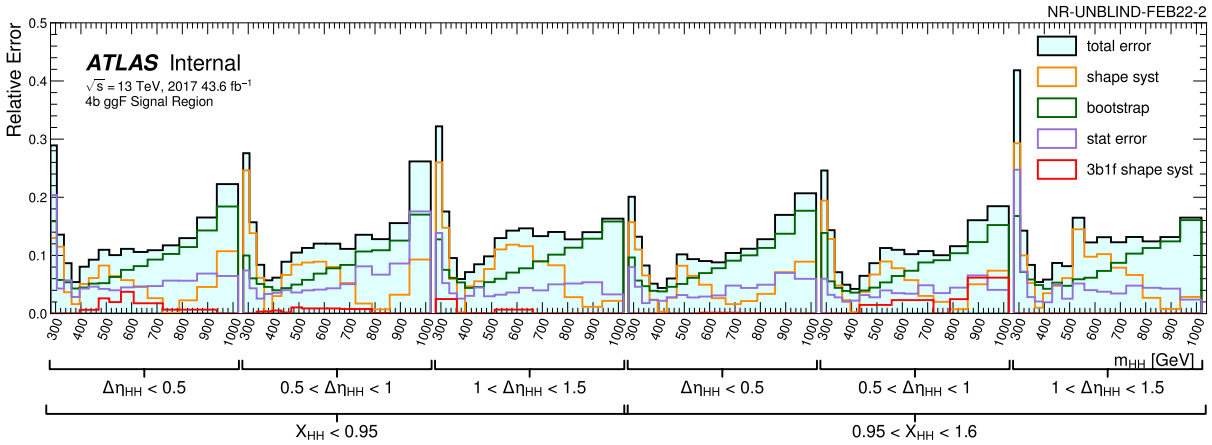
Figure 4.26: Example of variation in the SR NP quadrants for 2018 ggF discriminant. [4]

$3b1\text{f}$ m_{HH} distribution is then compared with the observed $3b1\text{f}$ data in the SR. For the VBF signal region, no statistically significant difference between the prediction and observation is found, and hence no additional uncertainty is applied. The uncertainty of ggF SR is derived as follows: In each analysis category, the ratio of the reweighted $2b$ m_{HH} distributions (predicted) and the $3b1\text{f}$ (observed) is computed and compared to unity bin-by-bin. If the difference is less than 1σ, where σ is obtained from all other background modelling uncertainties combined, no additional uncertainty is applied. For bins where the predicted and observed values differ by more than 1σ, the amount beyond 1σ is averaged with the corresponding amounts in the two adjacent bins, to limit the impact of statistical fluctuations, and is symmetrised around the nominal prediction to construct a two-sided uncertainty. This nonclosure uncertainty has a much smaller impact on the analysis sensitivity than the other sources of background modelling uncertainty.

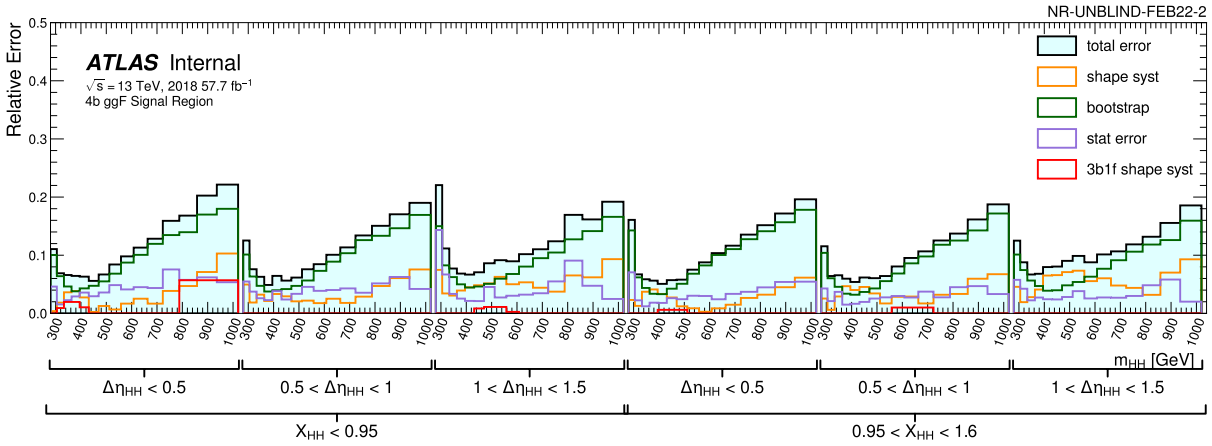
Illustrations of the magnitudes of the above background modelling uncertainties are shown in Figure 4.27 for ggF and Figure 4.28 for VBF. The statistical error dominates for high m_{HH} while the shape systematic contributes more in the moderate m_{HH} region that drives our analysis sensitivity. Although the statistical error is negligible relative to the bootstrap error in the bulk of the distribution, it becomes relevant in the high m_{HH} tail due to the reduced number of events, therefore the final statistical uncertainty used for the limit setting is the quadrature sum of them.



(a) 4b 2016 ggF



(b) 4b 2017 ggF



(c) 4b 2018 ggF

Figure 4.27: Relative error contributions of the background for the 4*b* ggF discriminant. [4]

Detector Modelling and Reconstruction Uncertainties The sources of experimental uncertainty in this analysis include contributions from the luminosity, jet energy

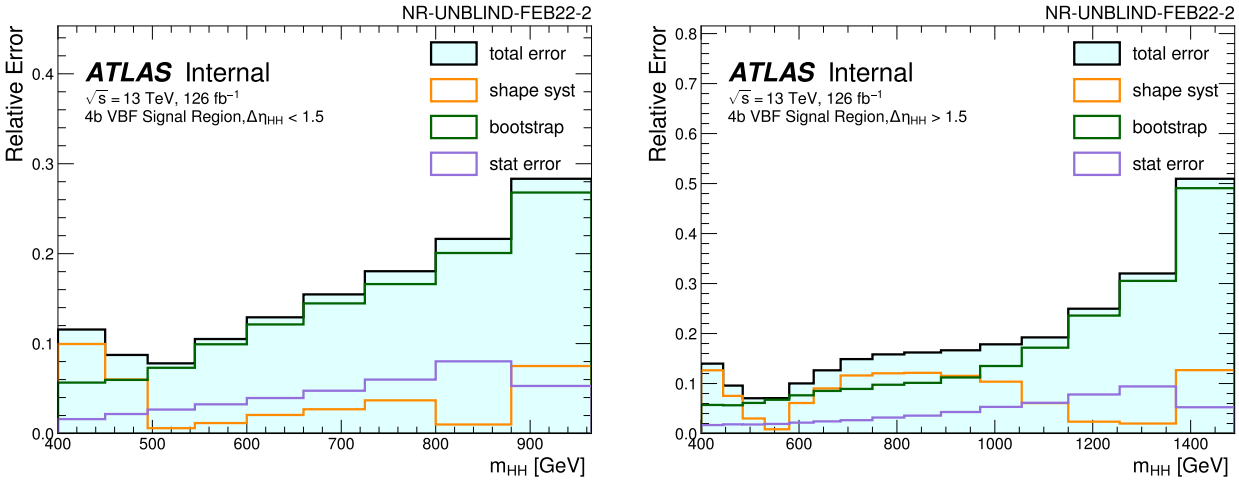


Figure 4.28: Relative error contributions of the background for the 4*b* VBF discriminant. [4]

scale and resolution (JES and JER), and *b*-tagging efficiency. These uncertainties affect only the signal modelling, as the background is derived entirely from data. Uncertainties related to JES and JER, as well as the jet vertex tagger (JVT), are treated according to the prescription in Refs [147, 136]. Additional uncertainties coming from the correction of the simulated pileup distribution are treated as per the method described in Ref [148]. The uncertainties in *b*-tagging efficiency are managed following the prescription provided in [149]. Trigger efficiency uncertainties are evaluated based on measurements of per-jet online efficiencies for both jet reconstruction and *b*-tagging, which are then used to calculate event-level uncertainties. These uncertainties are applied to the simulated events as overall weight variations. The uncertainty in the integrated luminosity for this analysis ranges from 2.0% to 2.4% for the individual years of data collection and 1.7% for the combined dataset [94]. These measurements were obtained using the LUCID-2 detector, which is the primary instrument used for luminosity determination [150].

Theoretical Uncertainties Uncertainties related to the modelling of the parton shower and underlying event are assessed by comparing the results obtained using two different generators: the nominal PYTHIA 8 and the alternative HERWIG 7. This comparison shows an impact of approximately 10% on the ggF and VBF signal acceptances, with a negligible effect on the shape of the m_{HH} distributions. Parton

showering uncertainties are derived separately for each analysis SR category, with uncertainties reaching up to 40% in certain categories where the acceptance for a given production mode is small. Uncertainties in the matrix element calculation are evaluated by varying the factorisation and renormalisation scales in the generator by a factor of two, both independently and simultaneously. This typically results in a 2% effect for both ggF and VBF, with a maximum effect of around 6% in certain analysis categories. PDF uncertainties are evaluated using the PDF4LHC_NLO_MC set [123], where the signal acceptance is calculated for each replica, and the standard deviation is taken as the uncertainty. This uncertainty is generally found to be below 1% for both ggF and VBF acceptances, with a maximum value of approximately 2%. Theoretical uncertainties in the $H \to b\bar{b}$ branching ratio [116] are also included, contributing an overall uncertainty of about 3.5% to the signal normalisation. The dependence of this uncertainty on κ_λ is neglected. Additionally, theoretical uncertainties in the ggF and VBF HH cross-sections, arising from uncertainties in the PDF and α_s, as well as the choice of renormalisation scheme and the scale of the top quark mass, are taken from Refs [113, 116, 7]. These cross-section uncertainties are incorporated into the derivation of the upper limits on the ggF, VBF, and combined HH signal strengths, as well as the likelihood-based constraints on the κ_λ and κ_{2V} modifiers, as discussed in Section 4.9.

An additional systematic uncertainty in signal modelling is considered for the SMEFT and HEFT measurements. The m_{HH} spectra of reweighted SMEFT/HEFT signal samples are compared against explicitly generated samples, with the average relative deviations across the m_{HH} bins in the ranges of 280 GeV$< m_{HH} <$938 GeV and $m_{HH} > 936$ GeV used to compute the uncertainty. This separation into low- and high-m_{HH} regions ensures that the better agreement in the densely populated low-m_{HH} region does not overly constrain the uncertainty in the more sparsely populated high-m_{HH} region.

4.8 Statistical Analysis

The signal extraction fit is performed using the m_{HH} variable, chosen for its strong discriminating power between signal and background, as well as between different signal hypotheses (e.g. different κ_λ signals). In the ggF SR, the fit is conducted simultaneously across the three data-taking years (2016, 2017 and 2018), whereas in the VBF SR, the fit is performed inclusively, using data from all years.

The number of categories, fits and all possible background nuisance parameters (NP) are listed in Table 4.8. For both the ggF and VBF channels, the shape and normalisation uncertainties for each category are determined by comparing the CR1 and CR2 models after each category selection process.

	ggF channel	VBF channel
Categories	3 years$\times 3|\Delta\eta_{HH}| \times 2X_{HH} = 18$	$2\ |\Delta\eta_{HH}|$
Background shape NP	3 years $\times(\mathrm{Q_N}, \mathrm{Q_E}, \mathrm{Q_S}, \mathrm{Q_W}) = 12$	$(\mathrm{Q_N}, \mathrm{Q_E}, \mathrm{Q_S}, \mathrm{Q_W}) = 4$
$3b1$f nonclosure NP	18 (1 per category)	0

Table 4.8: Summary of the categorisation strategy and background-related nuisance parameters (NP) in ggF and VBF channels

Uncertainties related to luminosity and signal modelling are treated as fully correlated across the analysis categories, as well as data-taking years for ggF. The components of the shape uncertainty derived from the quadrants are correlated across the data-taking years for ggF channel. These components are also correlated across analysis categories within the ggF and VBF SRs, though they are treated as uncorrelated between ggF and VBF. All other uncertainties in the background model are considered uncorrelated between different categories and across data-taking years.

The statistical model is implemented using RooFit [151].

The presence of the signal is tested using the profile likelihood ratio test with the standard asymptotic formulae [152]. Histograms in the discriminant variables with their systematics are used in the fit. Say that our measured variable x, in our case is

m_{HH}, constructs the histogram $\boldsymbol{n} = (n_1, \ldots, n_N)$. Then we have $\mathbb{E}[n_i] = \mu s_i + b_i$, for

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s) dx$$
$$b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b) dx$$

where μ is the parameter of interest (POI), for example, the signal strength. The signal strength(e.g. for combined ggF + VBF process) is defined as

$$\mu_{\text{ggF+VBF}} = \frac{\sigma_{\text{ggF+VBF}}}{\sigma^{\text{SM}}_{\text{ggF+VBF}}} \tag{4.8}$$

$\mu = 0$ corresponds to the background-only model and $\mu = 1$ corresponds to the background and predicted signal. $\boldsymbol{\theta}$s are the corresponding nuisance parameters and f is the probability density function. Event falling into the bin i can be regarded as a Poisson process with expectation $\lambda_i(\mu, \boldsymbol{\theta})$.

The likelihood function used to construct the test statistic has a standard form, consisting of a product of Poisson distributions of the yields in each bin, along with constraint functions for NPs related to systematic uncertainties. For uncertainties arising from limited sample size in either data or simulation, the constraint function is a Poisson distribution, whereas for all other systematic uncertainties, a Gaussian distribution is applied. Then we have the likelihood

$$L(\mu, \theta) = \prod_{c \in \text{categories}} \prod_i \text{Pois}(n_{c,i} | \lambda_{c,i}(\hat{\mu}, \hat{\theta})) \prod_{s \in \text{systematics}} \text{Gaus}(\hat{\theta}_s | \theta_s) \tag{4.9}$$

To test a hypothesized value of μ, we consider the profile likelihood ratio test statistics used in hypothesis testing $\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})}$, or more precisely

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\ \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(0, \hat{\theta}(0))} & \hat{\mu} < 0 \end{cases} \tag{4.10}$$

In the numerator, $\hat{\hat{\theta}}$ denotes the value of θ that maximizes L for the **given** μ. While the denominator represents the unconditional likelihood fit, where both $\hat{\mu}$

and $\hat{\theta}$ are maximum likelihood estimators (MLE) for L. During the fitting process, the algorithm maximises the likelihood function by adjusting the signal strength parameter μ and the NP θs. Since $\mu \geq 0$ in our case, whenever we find $\hat{\mu} < 0$, the best level of agreement between the data and any physical value of μ occurs for $\mu = 0$. Therefore when $\hat{\mu} < 0$, we take $\tilde{\lambda}(\mu) = \lambda(0)$.

The test statistic $\tilde{q}_\mu$ is defined as

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{0},\hat{\theta}(0))} & \hat{\mu} \leq 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} < \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{4.11}$$

The distributions of $\tilde{q}_\mu$ are given by the asymptotic formulae. We have the p-values:

$$p_\mu = \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu,\theta(\mu))d\tilde{q}_\mu \tag{4.12}$$

$$p_0 = 1 - \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|0,\theta(0))d\tilde{q}_\mu \tag{4.13}$$

The 95% confidence level CL exclusion limit, μ_{up} is the μ such that

$$p'_\mu = \frac{p_\mu}{1-p_0} \leq 5\% \tag{4.14}$$

The one and two standard deviation error bands on the limits are estimated using the following formula with some fine-tuning

$$\mu_{\text{up+N}} = \mu_{\text{up}}(\Phi^{-1}(1-0.05\Phi(\pm N)) \pm N) \tag{4.15}$$

where Φ^{-1} is the quantile function (i.e. Φ is the cumulative density function) of a standard Gaussian distribution.

4.9 Results

The observed m_{HH} distributions, along with the predicted background and example signal shapes, are shown in Figure 4.29 for each of the six ggF categories, with data from all years combined. The expected background distributions are obtained using the best-fit values of the NPs from the fit under the background-only hypothesis. The corresponding m_{HH} distributions for the two VBF categories are shown in 4.30.

In Figure 4.29a, the signal shape for $\kappa_{2V} = 0$ clearly illustrates the impact of the divergences in Figure 4.2b and 4.2c, which fail to cancel out, as discussed in Section 4.1. Although the deviations from the SM explored in this analysis do not reach levels that would violate unitarity, this behaviour makes the VBF topology particularly sensitive to κ_{2V}. A summary of the observed number of data events, predicted number of background events, and expected number of signal events for the SM ggF and VBF signals are summarised in Table 4.9.

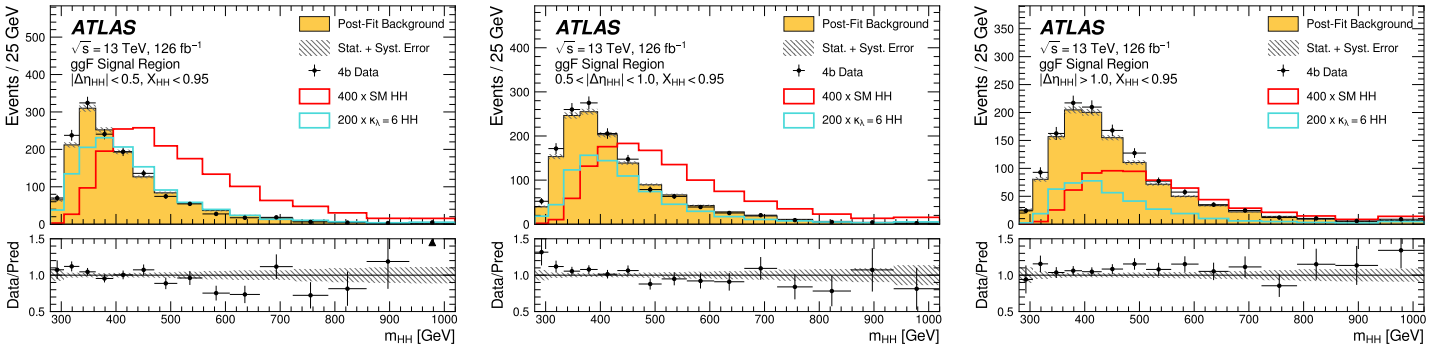
No evidence of signal is found in this analysis.

Category	Data	Expected Background	ggF Signal SM	VBF Signal SM
ggF signal region				
$\lvert\Delta\eta_{HH}\rvert < 0.5, X_{HH} < 0.95$	1941	1935 ± 25	7.0	0.038
$\lvert\Delta\eta_{HH}\rvert < 0.5, X_{HH} > 0.95$	3602	3618 ± 37	6.5	0.036
$0.5 < \lvert\Delta\eta_{HH}\rvert < 1.0, X_{HH} < 0.95$	1924	1874 ± 21	5.1	0.037
$0.5 < \lvert\Delta\eta_{HH}\rvert < 1.0, X_{HH} > 0.95$	3540	3492 ± 35	4.7	0.040
$\lvert\Delta\eta_{HH}\rvert > 1.0, X_{HH} < 0.95$	1880	1739 ± 22	2.9	0.043
$\lvert\Delta\eta_{HH}\rvert > 1.0, X_{HH} > 0.95$	3285	3212 ± 37	2.8	0.041
VBF signal region				
$\lvert\Delta\eta_{HH}\rvert < 1.5$	116	125.3 ± 4.4	0.37	0.090
$\lvert\Delta\eta_{HH}\rvert > 1.5$	241	230.6 ± 5.3	0.06	0.21

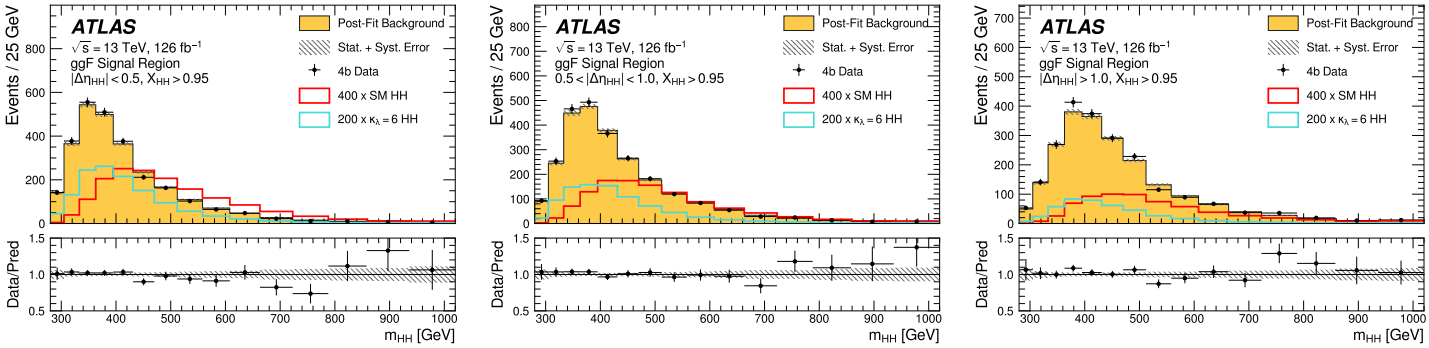
Table 4.9: The yields for each analysis category of the observed data, expected background, and expected SM ggF and VBF signals. The expected background yields are derived from a fit to the data with the background-only hypothesis, with the quoted uncertainties representing the quadrature sum of all per-bin systematic uncertainties. The expected signal yields are obtained from simulation.

4.9.1 Upper Limits on the cross-section

An upper limit on the combined ggF and VBF HH signal strength, $\mu_{\text{ggF+VBF}}$ is calculated using the asymptotic formula [152] and the CL_s method [153]. The



(a) $|\Delta\eta_{HH}| < 0.5, X_{HH} < 0.95$ **(b)** $0.5 < |\Delta\eta_{HH}| < 1.0$, $X_{HH} < 0.95$ **(c)** $|\Delta\eta_{HH}| > 1.0, X_{HH} < 0.95$

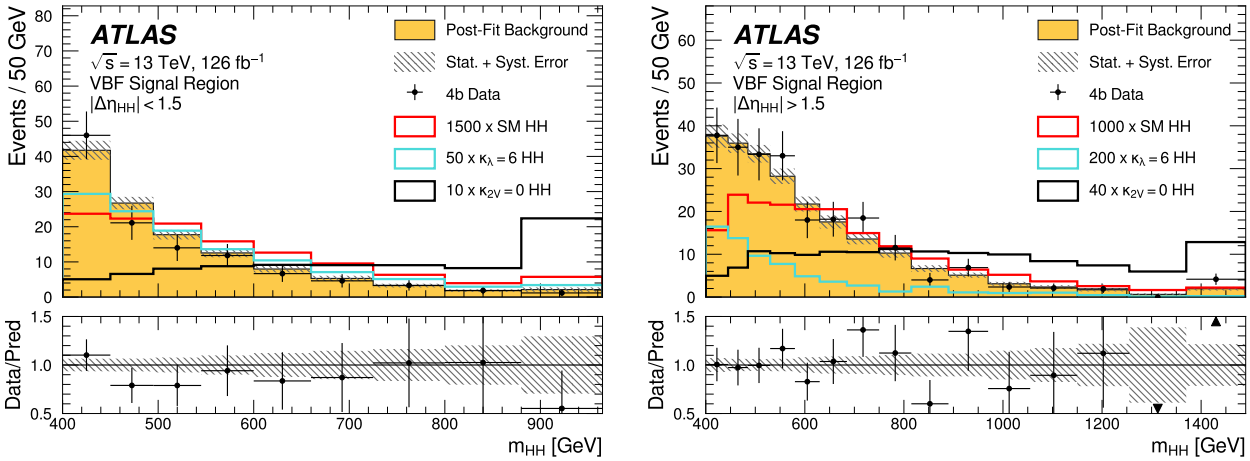


(d) $|\Delta\eta_{HH}| < 0.5, X_{HH} > 0.95$ **(e)** $0.5 < |\Delta\eta_{HH}| < 1.0$, $X_{HH} > 0.95$ **(f)** $|\Delta\eta_{HH}| > 1.0, X_{HH} > 0.95$

Figure 4.29: Distributions of the reconstructed m_{HH} in data (black points) and the estimated background (yellow histograms), in each of the six categories in the ggF SR, all years combined. The hatching shows the total uncertainty of the background estimate. The lower panels show the ratio of the observed data yield to the predicted background. [4]

observed limit is the actual upper limit derived directly from the data collected during the experiment. It is the maximum signal strength allowed by the observed data at given confidence level. The expected upper limit is what one would expect if there were no HH signal present, based on background-only hypothesis. This reflects the sensitivity of the analysis, since a lower expected limit indicates that the analysis is more capable of detecting smaller signals.

The observed (expected) 95% CL upper limit on $\mu_{\text{ggF+VBF}}$ is found to be 5.4 (8.1). Table 4.10 summarised the upper limit on the combined $\mu_{\text{ggF+VBF}}$, as well as the upper limit on the individual $\mu_{\text{ggF}} = \sigma_{\text{ggF}}/\sigma^{\text{SM}}_{\text{ggF}}$ and $\mu_{\text{VBF}} = \sigma_{\text{VBF}}/\sigma^{\text{SM}}_{\text{VBF}}$. For the individual μ_{ggF} and μ_{VBF} limits, the other production mode is treated as background



(a) $|\Delta\eta_{HH}| < 1.5$ **(b)** $|\Delta\eta_{HH}| > 1.5$

Figure 4.30: Distributions of the reconstructed m_{HH} in data (black points) and the estimated background (yellow histograms), in each of the two categories in the VBF SR. The hatching shows the total uncertainty of the background estimate. The lower panels show the ratio of the observed data yield to the predicted background. [4]

process, with its normalisation only loosely constrained during the fit.

	Observed Limit	-2σ	-1σ	**Expected Limit**	$+1\sigma$	$+2\sigma$
μ_{ggF}	5.5	4.4	5.9	8.2	12.4	19.6
μ_{VBF}	130	70	100	130	190	280
$\mu_{\text{ggF+VBF}}$	5.4	4.3	5.8	8.1	12.2	19.1

Table 4.10: The observed and expected upper limits on SM ggF HH production cross-section σ_{ggF}, SM VBF HH production cross-section σ_{VBF}, and combined SM ggF and VBF HH production cross-section $\sigma_{\text{ggF+VBF}}$ at the 95% CL, expressed as multiples of the corresponding SM cross-sections. The expected values are shown with corresponding one- and two-standard-deviation error bounds, and they are obtained using a background-only fit to the data. When extracting the limits on $\mu_{\text{ggF+VBF}}$, the relative contributions of ggF and VBF production to the total cross-section are fixed to the SM prediction.

Compared to the previous ATLAS measurement of ggF HH production in the $b\bar{b}b\bar{b}$ decay channel which uses 27fb^{-1} of early Run 2 data [30], the current upper limit on the ggF cross-section is more than 50% lower. Approximately 20% of this improvement is from advances in analysis techniques and object reconstruction. Similarly, when compared to the earlier ATLAS measurement of VBF HH production in the $b\bar{b}b\bar{b}$ decay channel, based on 126fb^{-1} of data collected between

2016 and 2018 [31], the upper limit on the VBF HH cross-section is reduced by 75%. This improvement is entirely due to advances in analysis techniques and object reconstruction.

The total uncertainty in the upper limit of the cross-section is primarily driven by uncertainties related to the background modelling procedure and theoretical predictions. When only the statistical uncertainties from the reweighted $2b$ data, observed $4b$ data, and simulated signal samples are included in the fit, the expected upper limit on $\mu_{\text{ggF+VBF}}$ is 6.0 times the SM prediction. However, when uncertainties from the background estimation - such as the bootstrap uncertainty and $3b1$f nonclosure uncertainty - are included, the expected upper limit on $\mu_{\text{ggF+VBF}}$ increases to 7.1 times the SM prediction. The further reduction in sensitivity, leading to 8.1 times the SM prediction as quoted in Table 4.10, is primarily driven by uncertainties arising from theoretical predictions. The relative impact of various systematic uncertainties on the expected upper limit on $\mu_{\text{ggF+VBF}}$ is summarised in Table 4.11.

Source of Uncertainty	$\Delta\mu/\mu$
Theory Uncertainties	
Theory uncertainty in signal cross-section	-9.0%
All other theory uncertainties	-1.4%
Backgroudn modeling uncertainties	
Bootstrap uncertainty	-7.1%
CR to SR extrapolation uncertainty	-7.5%
$3b1$f nonclosure uncertainty	-2.0%

Table 4.11: Breakdown of dominant systematic uncertainties. The impact of the uncertainties on the expected upper limit on $\mu_{\text{ggF+VBF}}$ when re-evaluating the profile likelihood ratio after fixing the NPs in question to their best-fit values, while all remaining NPs remain free to float. The impact is shown in %. Only the systematic uncertainties that have an impact of at least 1% are shown. The impact of each experimental source of systematic uncertainty described in the test, as well as of all of them together, is less than 1%.

4.9.2 Constraints on the coupling modifiers

Constraints on the κ_λ and κ_{2V} modifiers are derived using two different approaches: the ”95% CL” method and the ”profile likelihood ratio” method.

The ”95% CL” method sets the upper limit on the signal strength μ by determining the parameter values for which the observed data is compatible with the

background-only hypothesis at the 95% confidence level. The "profile likelihood ratio" method constructs a likelihood function of the parameter of interests (POI), the vector of coupling modifiers $\vec{\kappa} = (\kappa_\lambda, \kappa_{2V})$, while profiling over nuisance parameters to account for systematic uncertainties, and then uses this function to derive the constraints.

4.9.2.1 The 95% CL Method

The 95% CL method interprets the analysis as a traditional search for a signal. It uses a background-only hypothesis and sets limits on the maximum possible contribution of the signal without making specific assumptions about the signal's shape.

Constraints on κ_λ and κ_{2V} are obtained by determining the 95% CL upper limits on the cross-section as a function of the coupling modifiers - $\mu_{\text{ggF+VBF}}(\kappa_\lambda, \kappa_{2V})$. Values of the coupling modifiers $(\kappa_\lambda, \kappa_{2V})$ are excluded if the predicted cross-section for that configuration is excluded at the 95% CL. In the likelihood fit, the $H \to b\bar{b}$ branching ratio is fixed to its SM value, and any dependence on κ_λ is neglected. Upper limits on the HH signal strength as a function of κ_λ and κ_{2V} are shown in Figure 4.31, and the exclusion boundaries are summarised in Table 4.12. With the other modifiers (κ_V and either κ_{2V} or κ_λ) set to their SM values of 1, the analysis excludes values of κ_λ outside the range [-3.9, 11.1] and values of κ_{2V} outside [-0.03, 2.11].

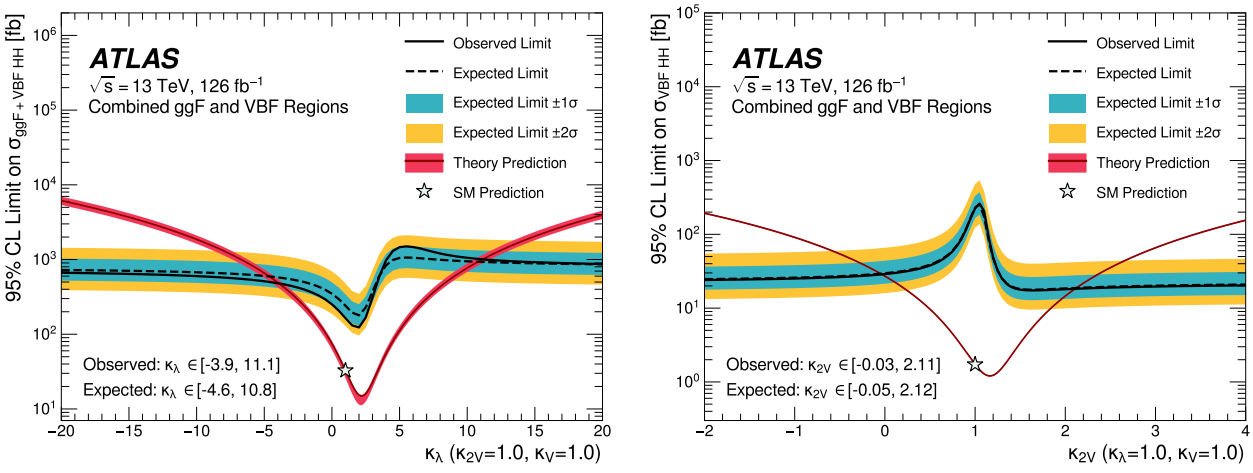
Parameter	**Expected Constraint**		**Observed Constraint**	
	Lower	Upper	Lower	Upper
κ_λ	-4.6	10.8	-3.9	11.1
κ_{2V}	-0.05	2.12	-0.03	2.11

Table 4.12: The observed and expected constraints on the κ_λ abd κ_{2V} coupling modifiers at 95% CL. For each modifier, the constraints were extracted with all other modifiers fixed to the SM prediction. [4]

Figure 4.32 shows the 95% CL exclusion limits in the two dimensional plane of the $\kappa_\lambda - \kappa_{2V}$ modifier space.

4.9.2.2 The Profile Likelihood Ratio Method

The profile likelihood ratio method uses a hypothesis consisting of the predicted background plus the SM HH signal. It tests whether the data are compatible with the



(a) **(b)**

Figure 4.31: The observed 95% CL exclusion limits as a function of (a) κ_λ (obtained using the signal strength $\mu_{\text{ggF+VBF}}$ as the POI) and (b) κ_{2V} (obtained using the signal strength μ_{VBF} as the POI) from the combined ggF and VBF SRs, as shown by the solid black line. In each case, the value of the other modifier is fixed to 1. The blue and yellow bands show respectively the 1σ and 2σ bands around the expected exclusion limits, which are shown by the dashed black line. The expected exclusion limits are obtained using a fit to the data with background-only hypothesis. The dark red line shows in (a) the predicted combined ggF and VBF HH cross-section as a function of κ_λ and (b) the predicted VBF HH cross-section as a function of κ_{2V}. The dark pink bands surrounding the predicted cross-section lines indicate the theoretical uncertainty of the cross-section, as taken from Ref [7]. [4]

specific cross-section and shape predictions of the κ framework. These alternative coupling modifier constraints are derived using the profile likelihood ratio method, where the coupling modifiers $\vec{\kappa} = (\kappa_\lambda, \kappa_{2V})$ are treated as the POIs instead of the signal strength μ:

$$\lambda(\vec{\kappa}) = \frac{L(\vec{\kappa}, \hat{\hat{\theta}}(\vec{\kappa}))}{L(\hat{\vec{\kappa}}, \hat{\theta})} \tag{4.16}$$

and then we are interested in

$$-2\Delta \ln L(\vec{\kappa}) = -2\ln\left(\frac{L(\vec{\kappa}, \hat{\hat{\theta}}(\vec{\kappa}))}{L(\hat{\vec{\kappa}}, \hat{\theta})}\right) \tag{4.17}$$

A scan of the profile likelihood ratio is performed as a function of the coupling modifiers at discrete points, producing the curve shown in Figure 4.33. At each point, the signal strength is now fixed to the prediction obtained for the specific coupling

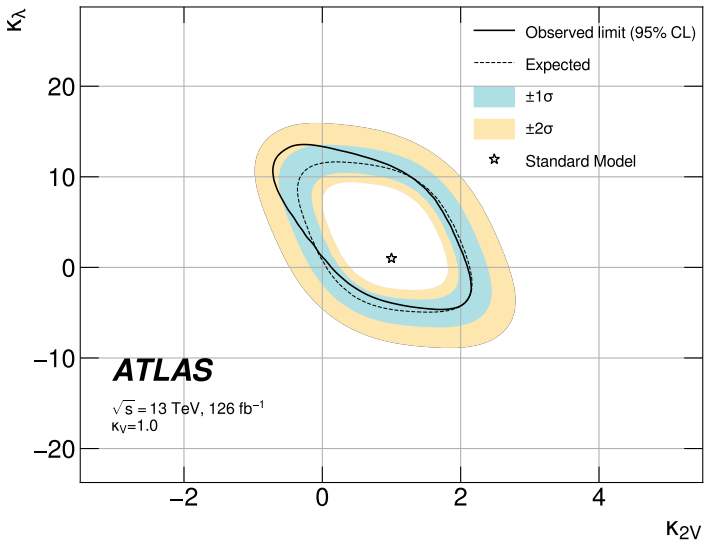
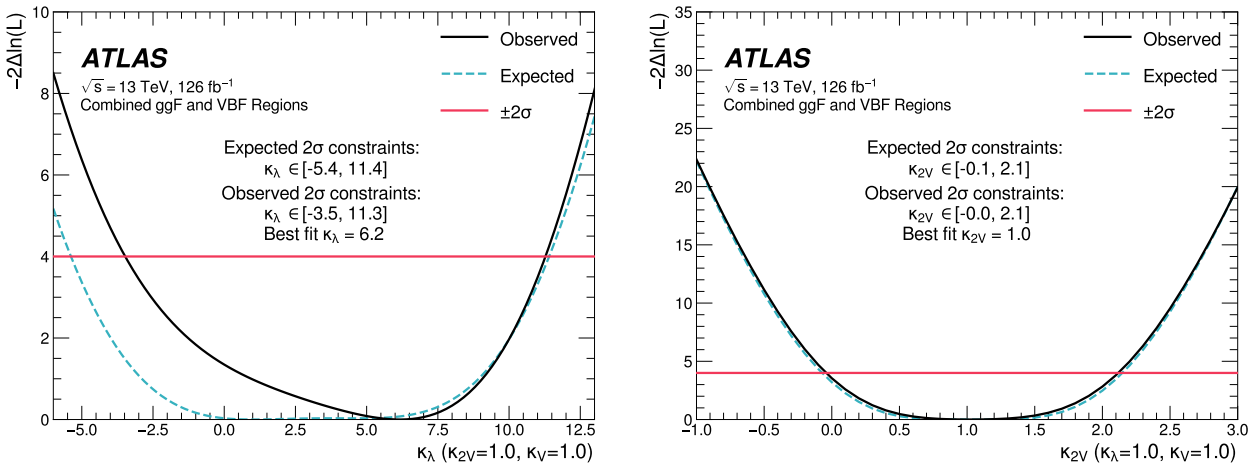


Figure 4.32: The observed 95% CL exclusion limit obtained using the signal strength $\mu_{\text{ggF+VBF}}$ as the POI in the two-dimensional κ_λ vs κ_{2V} space, obtained from the combined ggF and VBF signal model, as shown by the solid black line. The blue and yellow bands show respectively the 1σ and 2σ bands around the expected exclusion limits, which are shown by the dashed black line. The star denotes the SM prediction ($\kappa_\lambda = \kappa_{2V} = 1$). [4]

modifier configuration.

Though the physical assumption is different from the 95% CL method, which uses the background-only hypothesis, given the relatively small size of the SM HH signal compared to the background, this difference is not expected to significantly impact the results. From this scan, the best-fit value of κ_λ is found to be 6.2. The observed pull of the best-fit κ_λ value away from the SM prediction is attributed to a slight excess of observed data in the ggF signal region, particularly in the low-m_{HH} range. A signal model with κ_λ near 6 is favoured due to the balancing of two competing effects: as κ_λ increases from the SM value, the m_{HH} spectrum softens, while the cross-section also increases beyond the level of the observed excess when κ_λ increases much further. This slight excess also results in the observed limits in Figure 4.33 deviating from the expected limits by approximately 1σ. In contrast, no such excess is observed in the VBF SR, and the best-fit value of κ_{2V} from the likelihood scan is found to be 1.0.

The ranges of the coupling modifiers are obtained by 2σ-level constraints. With



(a) **(b)**

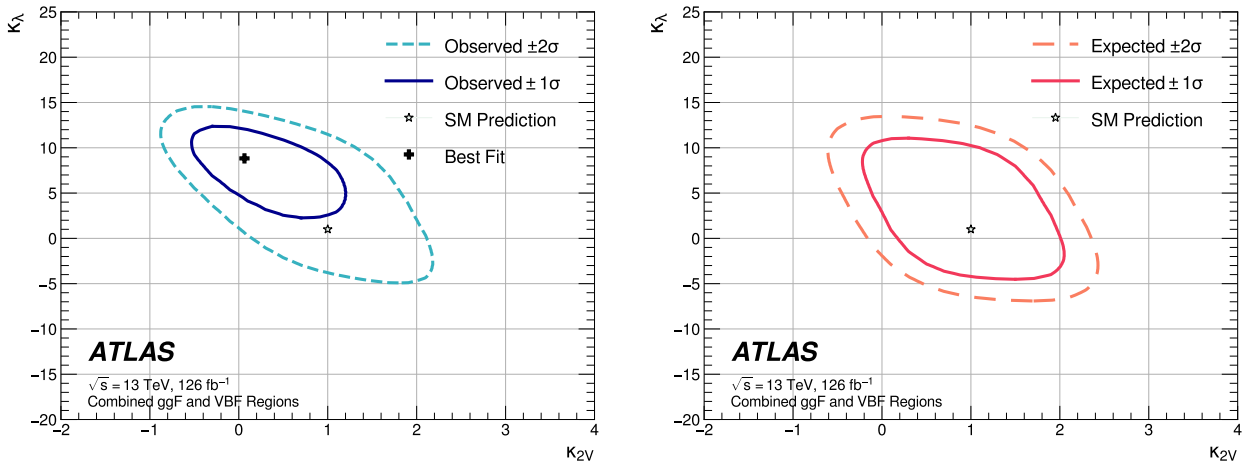
Figure 4.33: The observed profile likelihood ratio scans for (a) κ_λ and (b) κ_{2V}, shown by the solid black line using the coupling modifiers $\vec{\kappa}$ as the POIs. The dashed blue line shows expected profile likelihood ratio, as obtained using a fit to the data with the background-only hypothesis. Ths pink line indicates the 2σ exclusion boundary. [4]

the other coupling modifiers (κ_V and either κ_{2V} or κ_λ, respectively) fixed to their SM value of 1, the observed (expected) 2σ allowed range for κ_λ is found to be [-3.5, 11.3] ([-5.4, 11.4]) and the corresponding range for κ_{2V} is [-0.0, 2.1] ([-0.1, 2.1]).

The exclusion constraints derived using the profile likelihood ratio method are also presented in the two-dimensional $\kappa_\lambda - \kappa_{2V}$ coupling modifier space, similar to the 95% CL constraints described earlier in Section 4.9.2.1. The excluded regions are presented in Figure 4.34. When both modifiers are allowed to vary in the two-dimensional fit which combines the ggF and VBF SRs, the fit slightly shifts the coverage of κ_λ and κ_{2V} compared to the values found in the single-parameter fits where only one parameter is free at a time.

4.9.3 Constraints on coefficients of SMEFT and HEFT frameworks

In addition to constraints on the ggF and VBF HH cross-sections and the κ_λ and κ_{2V} coupling modifiers, constraints for relevant coefficients can be derived from the ggF selection of the analysis in the SMEFT and HEFT frameworks, as discussed in Section 4.1. The VBF HH process is ignored from both the SMEFT and HEFT



(a) **(b)**

Figure 4.34: (a) The observed profile likelihood ratio exclusion limits for the two-dimensional κ_λ vs κ_{2V} modifier space, shown by the solid line at the 1σ level and the dashed line at the 2σ level. The black cross denotes the best-fit values of $(\kappa_\lambda, \kappa_{2V})$. The expected exclusion limits are presented in (b), where the solid line denotes the 1σ-level exclusion and the dashed line denotes the 2σ-level exclusion. For both the expected and observed limit plots, the black star indicates the SM prediction of $\kappa_\lambda = \kappa_{2V} = 1$. [4]

results, since including it as a background was found to have a negligible impact on the extracted parameter limits. The slight dependence on the $H \to b\bar{b}$ branching fraction on the SMEFT and HEFT coefficients is also ignored, due to its minimal effect on the analysis sensitivity.

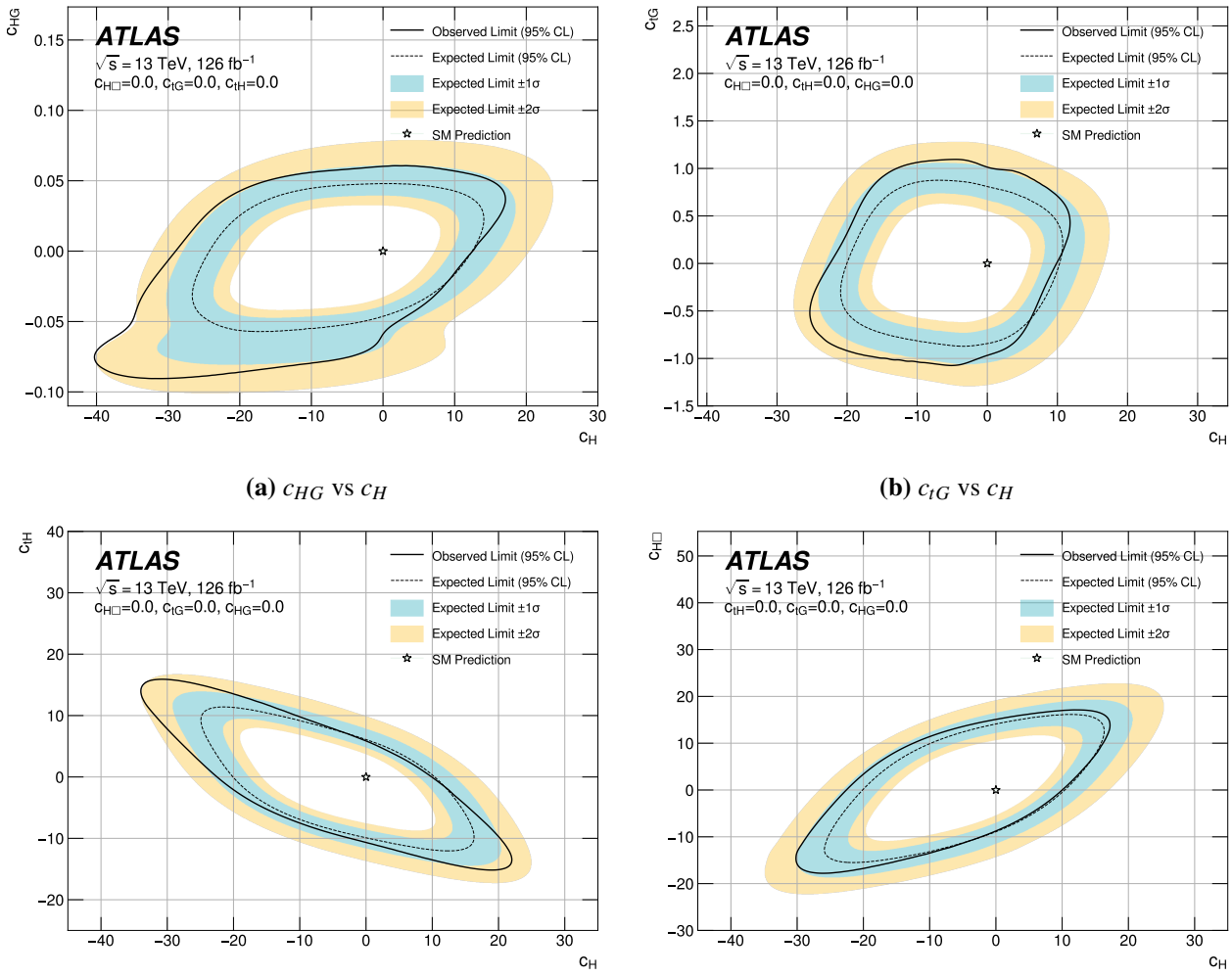
Constraints on the SMEFT coefficients are obtained by considering 95% CL exclusion limits for the cross-section as a function of each SMEFT parameter, similar to the κ_λ and κ_{2V} constraints. The extracted limits for individual parameters, assuming the others are fixed to 0, are summarised in Table 4.13. Caution is advised when interpreting limits approaching or exceeding $\pm 4\pi$, since there is potential influence from the missing higher-order model contributions. The exclusion limits for each coefficient as a function of the c_H coefficient (with the remaining three coefficients set to 0) are shown in Figure 4.35.

The upper limits on the HEFT ggF HH production cross-section for the seven benchmark models are displayed in Figure 4.36. The range of sensitivity across the seven benchmark models reflects the different signal kinematics and the correspond-

ing shapes of the m_{HH} distributions. The variation between observed and expected limits is linked to the slight excess observed in the low m_{HH} region, as discussed earlier in Section 4.9.2.2. Red crosses in Figure 4.36 represent the predicted HH cross-sections from the respective benchmark models. We can see that BM3, BM5 and BM7 are excluded with more than 95% confidence. Constraints are placed on the values of c_{ggHH} and $c_{t\bar{t}HH}$, with all other HEFT coefficients fixed to the SM values. The observed (expected) constraints on c_{ggHH} are found to be [-0.36, -0.78] ([-0.42,0.75]), while the observed (expected) constraints on $c_{t\bar{t}HH}$ are found to be [-0.55,0.51] ([-0.46,0.40]).

Parameter	**Expected Constraint**		**Observed Constraint**	
	Lower	Upper	Lower	Upper
c_H	-20	11	-22	11
c_{HG}	-0.056	0.049	-0.067	0.060
$c_{H\Box}$	-9.3	13.9	-8.9	14.5
c_{tH}	-10.0	6.4	-10.7	6.2
c_{tG}	-0.97	0.94	-1.12	1.15

Table 4.13: The extracted upper and lower limits on the SMEFT parameters to which the analysis is sensitive. The VBF HH process is ignored.



(a) c_{HG} vs c_H

(b) c_{tG} vs c_H

(c) c_{tH} vs c_H

(d) $c_{H\Box}$ vs c_H

Figure 4.35: The observed 95% CL exclusion limits on the SMEFT coefficients in the two-dimensional spaces, shown by the solid black lines. The dashed black line indicates the expected 95% CL exclusion limits. The shaded blue band indicates the $\pm 1\sigma$ uncertainty and the yellow band indicates the $\pm 2\sigma$ uncertainty. [4]

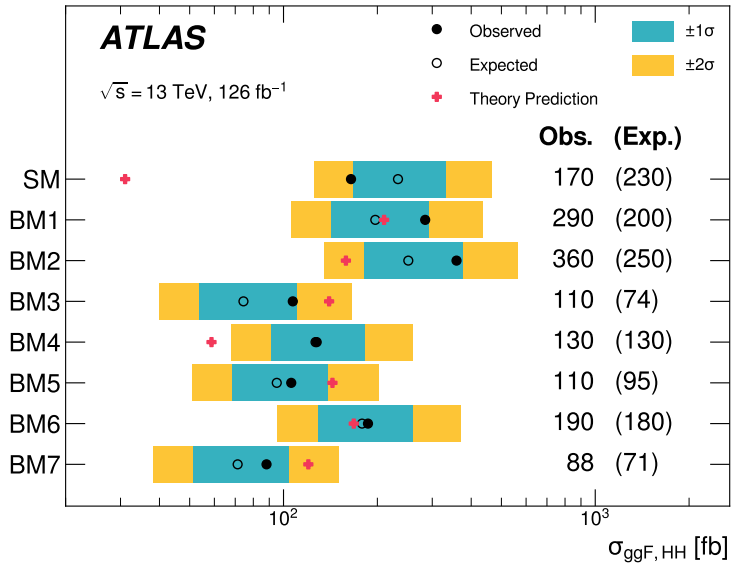


Figure 4.36: The observed 95% CL exclusion limits on the ggF HH production cross-section in the SM and each of the seven HEFT benchmark models, given by the solid black points. The blue and yellow bands show respectively the 1σ and 2σ bands around the expected upper limits, which are shown by the open circles. The predicted ggF HH production cross-section from each benchmark is indicated by a pink cross. Benchmarks where the theory cross-section is higher than the exclusion limit (i.e. to the right) are excluded. The VBF HH process is ignored for the result. [4]

Chapter 5

Classifier for the HH4B Analysis

As shown in Table 4.6, in the HH4b analysis, several variables are used as inputs for the background reweighting, while only m_{HH} is used to construct the histograms for the profile likelihood fit. Though m_{HH} has strong discriminating power, it is natural to investigate the feasibility of including more physical variables in constructing the discriminant.

In this Chapter, I will present my study of a Neural Network based classifier using the output probability as the discriminant. For exploration purposes, this study is only performed on the ggF channel.

5.1 Neural Networks for Classifier Training

The classifier is trained on Signal Region, with the background sample formed by the estimated background 4b events, and the signal sample formed by the MC simulated 4b signal events.

Two types of classifiers with different background estimation are trained:

- The **nominal classifier**: Uses the reweighted 2b events as background, the same as the background modelling method described in Section 4.6.

- The **flow classifier**: Instead of the background reweighting method, an alternative method of generating background 4b events using *normalising flows* is introduced (detailed in Section 5.1.2.1). The classifier is trained on these generated background 4b events.

5.1.1 Nominal Classifier

For the nominal classifier, the background samples are the $2b$ events in the Signal Region, reweighted using the nominal weight (the mean of 100 bootstrap weights), and the signal samples are the MC simulated $4b$ signal events in SR with MC Scale Factors (SF) applied (see Section 4.3). During model training, these reweighting factors are applied separately as the sample weight, adjusting each sample's contribution to the loss function by

$$L = \frac{1}{N}\sum_{i=1}^{N} w_i L_i \tag{5.1}$$

where N is the number of samples, w_i is the weight for the i-th sample, and L_i the loss for the i-th sample.

To balance the training datasets, the ratio of background to signal events is calculated as $R = N(\text{bkg})/N(\text{sig})$, where $N(\text{bkg})$ represents the sum of the background weights and $N(\text{sig})$ is the sum of MC SFs. During training, the sample weights for the signal events are scaled by this ratio R.

Due to the relatively small size of the dataset (approximately 150,000 background events and 100,000 signal events), complex architectures tend to overfit. After some parameter tuning, a simple NN model with 2 hidden layers of 200 nodes, each with a dropout of 0.2 and softmax activation is used, as shown in table 5.1. The train, validation and test data split is 70:15:15.

Layer(type)	Output Shape	Param #
dense(Dense)	(None, 200)	9200
dropout(Dropout)=0.2	(None, 200)	0
dense_1 (Dense)	(None, 200)	40200
dropout_1 (Dropout)=0.2	(None, 200)	0
dense_2 (Dense)	(None, 2)	204
Total params: 49802		

Table 5.1: The structure of the Neural Network used for Classifier Training

The loss used for training is categorial cross-entropy loss, for event i,

$$L_i = -y_i \log f(X_i) \tag{5.2}$$

where y_i is the class label, and $f(X_i)$ the model output when the input is X_i. In our case, the labels are one-hot encoded such that $y_i = \binom{1}{0}$ for background, and $y_i = \binom{0}{1}$ for signal. The model output after the final two hidden layers and softmax activation as $\binom{p_{i,b}}{p_{i,s}}$ for $p_{i,b}$, $p_{i,s}$ representing the probability of event i being background and signal respectively (will be detailed later). Then we have

$$L_i = -\log p_{i,b}, \quad \text{background}$$
$$L_i = -\log p_{i,s}, \quad \text{signal}$$

and minimising L is equivalent to maximise the probability predictions for the correct class. The optimisation method is Adam[154] optimisation.

It is worth noticing that some negative MC SFs are used to "scale down" certain events when estimating MC distributions, resulting in positive terms in the loss. The probability p_s will be suppressed instead, coherent with the purpose of scaling down the presence of the event in signal distribution. As shown in Figure 5.1, the fraction of signal events with negative sample weight is also small (around 8%), hence their effect on the convergence of the training process is not significant.

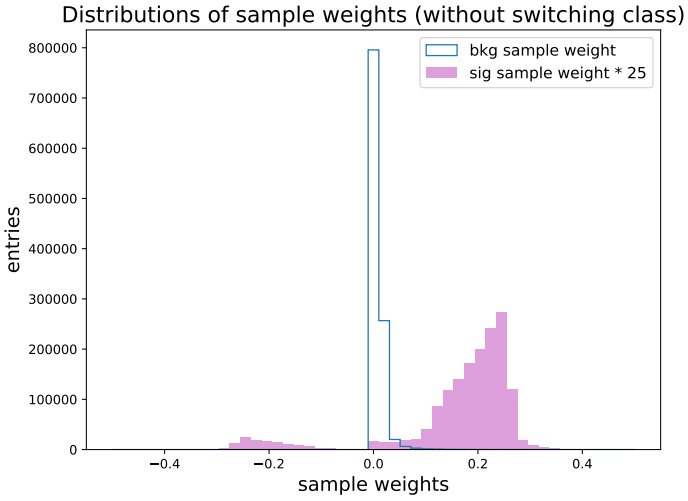


Figure 5.1: The distributions of background and signal sample weights. The signal histogram heights are multiplied by 25 for better visualisation.

Feature engineering is also important in this training. Including too many

physical variables will result in overfitting, and the curse of dimensionality means that high dimensional space makes the data sparse, which is detrimental since the size of the dataset is typically small. Apart from aiming to achieve a better upper limit, keeping good agreements of the model output for reweighted $2b$ events vs $4b$ events in the Control Region is also important. Since the reweighting achieves the same distribution of $2b$ and $4b$ in CR (see Figure 4.15), the model output of them should also agree. This is an important validation criterion which serves as a sanity check for the model, and also prevents overfitting. Starting with basic variables m_{HH}, X_{HH} and $\Delta\eta_{HH}$, I gradually added variables based on their impact (evaluated using SHAP analysis[155]) and their importance to the analysis (e.g. used in the background modelling, used for categorisations), to improve sensitivity while maintaining good CR agreement. Eventually, 11 variables are selected as inputs, as listed in table 5.2.

variables	Remarks
m_{HH}	The mass of the Higgs boson pair
m_{H_1}	The mass of the leading Higgs candidate
m_{H_2}	The mass of the second Higgs candidate
X_{HH}	As defined in 4.3
p_{THH}	The transverse momentum of the Higgs boson pair system
$\Delta\eta_{HH}$	The difference in pseudorapidity between the two Higgs bosons
njets	Number of jets in the event
Trigger class index	(one-hot encoded)
year	(one-hot encoded)
ΔR_{jj_1}	The distance of the two leading jets in the $\eta-\phi$ plane
ΔR_{jj_2}	The distance of the following pair of jets in the $\eta-\phi$ plane

Table 5.2: The set of input variables used for classifier training for ggF channel

The final layer of the model produces two outputs, called logits. The Softmax function exponentiates each logit and then normalising these values by the sum of all exponentiated logits, effectively converting the logits into probabilities that sum to 1

$$P(\text{class } i) = \frac{e^{z_i}}{e^{z_b} + e^{z_s}} \tag{5.3}$$

where z_i represents the logits for class i, with $i = b$ for background, and $i = s$ for

signal. $P(\text{signal})$, which is the probability of the event being signal (or later called classifier probability, or classifier score), is used as the discriminant. The limit fitting is performed on the binned histograms of this output probability. Note categorisation is not applied here since the variables $\Delta\eta_{HH}$ and X_{HH} are directly used in the training.

Similar to the m_{HH} distribution, the probability score is steeply falling at the high probability region. In addition, it experiences a rise in low-probability regions. A variable width histogram binning, similar to the method introduced in 4.6.2, is used to generate narrower bins in the middle and wider bins at both low and high-probability regions. The logarithmic binning scheme is again used. A default bin width, and the low and high score thresholds are set. The bins in between have a uniform width of the default bin width. After the high score threshold, each bin edge is set at (100+X%)$\times$ the previous edge, where X =default bin width/high score threshold. The edge for the lower probability region is calculated in a similar fashion - starting from the lower threshold, we set each bin edge at (100+Y%)$\times$ the previous edge for Y =default bin width/low score threshold, until we reach 2$\times$ the lower threshold; then we flip the edges to obtain the binning. The default bin width is set to be 0.025, and the low and high probability score threshold is set to 0.2 and 0.6 respectively. These parameters are optimised to keep the systematic uncertainties across the distribution small and achieve optimum sensitivity.

The distributions of the classifier probabilities for background (reweighted 2b data) and signal samples (MC simulated 4b data reweighted by MC Scale Factors as described in Section 4.3) in SR are shown in Figure 5.2. A good separation can be observed.

5.1.1.1 Systematic Uncertainties

The **Poisson statistical uncertainty** is calculated in the same way as in Section 4.7: for each bin i in the classifier probability histogram, the value is $s_i = \sum_{j\in i} w_j$, for w_j the weight for an event j with its probability score falls in bin i. The resulting statistical error is $\delta s_i = \sqrt{\sum_{j\in i} w_j^2}$.

The **Bootstrap Uncertainty** from background modelling is also calculated in the same fashion - the probability score of each 2b event is reweighted by the

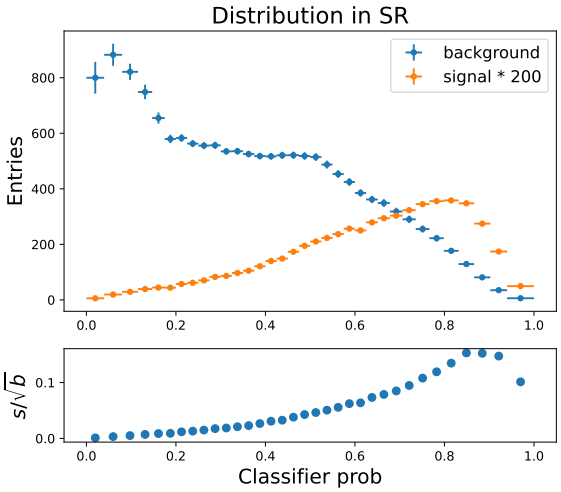


Figure 5.2: The distribution of the classifier for the reweighted $2b$ events (background), and MC simulated $4b$ events scaled by MC SFs (signal), in the Signal Region. The histogram heights for MC events are scaled up by 200 for better visualisation. Statistical uncertainty for signal, and statistical plus bootstrap uncertainty for background, are shown by the error bars.

100 bootstrap weights, and their standard deviation is recorded as the bootstrap uncertainty for the classifier. The **Background Shape Uncertainties** is also defined similarly as the difference between the scores reweighted by nominal weights (weights derived in CR1) and the alternative weights (weights derived in CR2).

For simplicity, the 3b1f-Non closure uncertainty, and other systematic and theoretical uncertainties are not considered.

The uncertainties are shown in fig 5.3. We can see that they are at manageable levels. The most dominant uncertainty is the bootstrap uncertainty, with the shape uncertainty dominating the low probability region, and statistical uncertainty becoming relevant in the high probability tail due to the reduced number of events.

The CR closure plots are shown in fig 5.4. Good agreement is achieved.

5.1.2 Flow Classifier

As shown in Table 4.11, the uncertainty from CRs extrapolating into SR degrades the upper limit by 7.5%, which is quite significant. An alternative background estimate training a generative model as a function of (m_{H1}, m_{H2}) is proposed[156].

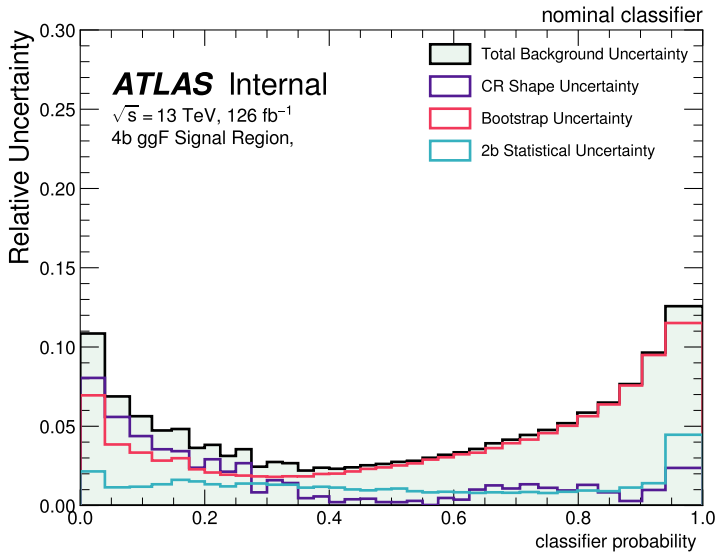


Figure 5.3: Systematic Uncertainties of the classifier probability output of the **nominal classifier**.

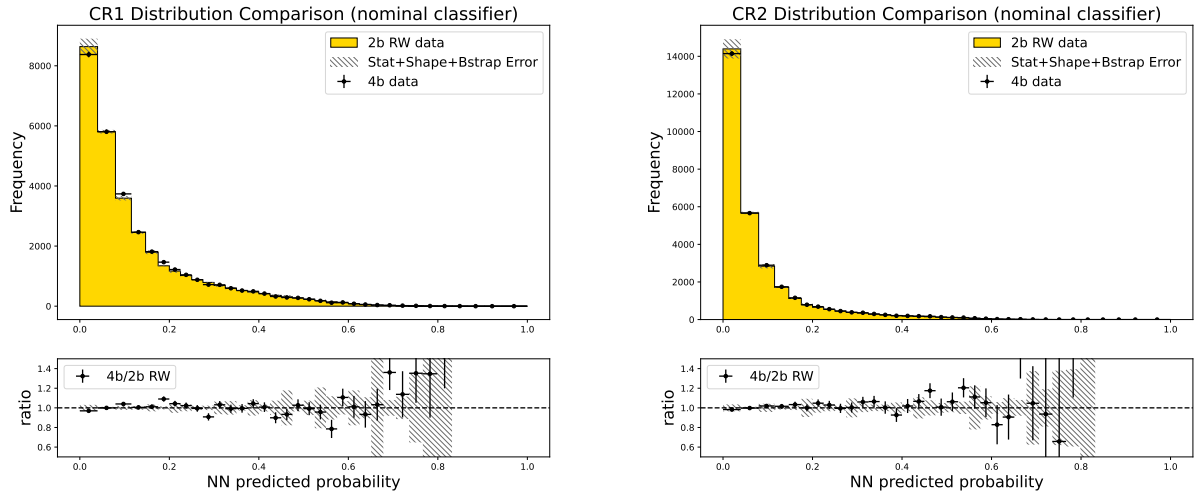
5.1.2.1 Generative background modelling

I will first give a brief introduction of the generative background modelling method. The full process and results are detailed in Nicole Hartman's (one of the main contributors of this method) thesis: [156].

The distributions of (m_{H1}, m_{H2}) in Signal Region are fitted with a Gaussian process(GP)[157], a non-parametric ML method. The input is a 2d histogram in the $m_{H1} - m_{H2}$ mass plane in a bounding box(edge length of 90 GeV, and 25 bins for both the m_{H1} and m_{H2} axes) surrounding the outer edge of the CR. The (m_{H1}, m_{H2}) bin centres are used to predict the number of events in the bin. During the training, the bins overlapping with SR are removed, and the inputs are standardised to have zero mean and unit variance. As usual, we denote the i-th training sample as $x^{(i)} = (m_{H1}^{(i)}, m_{H2}^{(i)}), y^{(i)} = f(x^{(i)}) + \varepsilon^{(i)}$ for some prior f, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, and $y^{(i)}$ is the number of events in the i-th bin.

The GP claims the priors follow a multivariate Gaussian probability distribution

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_m) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_m) \end{bmatrix}, \begin{bmatrix} K(x_1,x_1) & \dots & K(x_1,x_m) \\ \vdots & \ddots & \vdots \\ K(x_m,x_1) & \dots & K(x_m,x_m) \end{bmatrix} \right) \tag{5.4}$$



(a) CR1 **(b)** CR2

Figure 5.4: Control Region validation plots for (a) CR1 and (b) CR2.

where $x_1, \ldots, x_m$ are the input points, and the f is the prior over functions. To account for the difference between the years, priors for each year are modelled differently. The targets are standardised, hence the mean is set to zero and the variance is set to one. The covariance function used is the squared exponential kernel

$$K\left(\begin{bmatrix} m_{H1} \\ m_{H2} \end{bmatrix}, \begin{bmatrix} m'_{H1} \\ m'_{H2} \end{bmatrix}\right) = \exp\left(-\frac{(m_{H1}-m'_{H1})^2}{2l_1^2} - \frac{(m_{H2}-m'_{H2})^2}{2l_2^2}\right) \tag{5.5}$$

where ls are the length scales determine how strongly correlated the inputs are, and can be obtained by maximising the log marginal likelihood of the training inputs

$$\log p(\vec{y}|X,l) = \log \frac{1}{(2\pi)^{n/2}|K+\sigma^2 I|^{1/2}} \exp\left(-\frac{1}{2}\vec{y}^T(K+\sigma^2 I)y\right) \tag{5.6}$$

The joint marginal over the training points and the test (or interpolation) points X_* and $\vec{y}_*$(in our case X_* are the bins in SR, and y_* are the corresponding numbers of

events) also follows a multivariate Gaussian distribution

$$\begin{bmatrix} \vec{y} \\ \vec{y}_* \end{bmatrix} \Bigg| X, X_* = \begin{bmatrix} f \\ f_* \end{bmatrix} + \begin{bmatrix} \vec{\varepsilon} \\ \vec{\varepsilon}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X,X)+\sigma^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*)+\sigma^2 I \end{bmatrix}\right) \tag{5.7}$$

and the predictive posterior distribution of $\vec{y}_*$ follows a Gaussian distribution

$$\begin{aligned} \vec{y}_* \sim\ & (K(X_*,X)(K(X,X)+\sigma^2 I)^{-1}\vec{y}, \\ & K(X_*,X_*) - K(X_*,X)(K(X,X)+\sigma^2 I)^{-1}K(X,X_*)) \end{aligned} \tag{5.8}$$

The values (m_{H1}, m_{H2}) are used to condition the normalising flow model, which is a generative model that transforms a known base density into a target density by a sequence of invertible transformations f_i[158]. All three years are trained together, conditioned on the year. For the base distribution z (here is multivariate Gaussian), the forward mode of the flow $f = f_1 \circ \ldots \circ f_L$ samples from the Gaussian are transformed to give samples from the target distribution x. The reverse mode of the flow f^{-1} let us evaluate the probability of the samples

$$p_x(x) = p_z(z)\left|\frac{dz}{dx}\right| = p_z(f^{-1}(x))\left|\frac{df^{-1}}{dx}\right| \tag{5.9}$$

For $p_z(z)$ the Gaussian density, and x the features to model - m_{HH}, p_{TH1}, p_{TH2}, p_{THH}, η_{H1}, η_{H2}, X_{W_t}, $\Delta\phi_{HH}$, $\Delta\eta_{HH}$. The f_i used are splines with K bins and $K+1$ knots, with each bin parametrised by a rational quadratic transformation

$$f_{jk}(x_i) = \frac{a_{ijk}x_i^2 + b_{ijk}x_i + c_{ijk}}{d_{ijk}x_i^2 + e_{ijk}x_i + f_{ijk}} \tag{5.10}$$

where i is the dimension of the transforming variable, and j the corresponding flow layer, k is the bin. These parameters, together with the position of the knots, are optimised by maximising the likelihood of the training data points. NN is used to perform the optimisation, by using the negative log-likelihood of the training data $(-\sum_i \log p_x(x_i))$ as loss, and minimise it by stochastic gradient descent.

To summarise, 100,000 events are sampled from the interpolation box each

year to determine the (m_{H1}, m_{H2}) mass plane using GP (typically predicted around 15,000 events in the SR). These events are also used to determine the parameters of the flow model. The flow model is then applied to the SR massplane to generate the kinematic features of events in SR. These generated events are the estimation of the background $4b$ events in the Signal Region. This process is repeated 25 times using the same (m_{H1}, m_{H2}) to generate 25 flow samples. The final prediction is taken to be the mean of these 25 predictions.

The distribution of the flow-generated $4b$ events agrees with the real $4b$ events we sampled from in the Control Region.

5.1.2.2 Classifier Training using the Flow Generated Background

The 25 flow background estimates are concatenated together to form the background samples, and trained against the MC generated $4b$ signal samples. At inference time, the model will be applied separately on the 25 flow samples to obtain 25 classifier probability predictions, and the final discriminant output is taken to be the mean of them.

The optimal classifier structure is shown in Table 5.3, and has L2 regularisation of 1e-3, and batch normalisation to avoid overfitting. The optimised features to be included are m_{HH}, m_{H1}, m_{H2}, p_{TH1}, p_{TH2}, η_{H1}, η_{H2}, X_{W_t}, $\Delta\phi_{HH}$, and one-hot encoded year.

Layer(type)	Output Shape	Param #
dense(Dense)	(None, 256)	3584
batch normlisation	(None,256)	1024
dropout(Dropout)=0.2	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
batch normlisation	(None,128)	512
dropout_1 (Dropout)=0.2	(128)	0
dense_2 (Dense)	(None, 2)	258
Total params: 38274		

Table 5.3: The structure of the Neural Network used for flow Classifier Training

To obtain the final background probability estimation, the model is applied to the 25 flow-generated background samples, resulting in 25 probability estimations. Using the same binning as before, 25 histograms are constructed from these estima-

tions. The final histogram is then obtained by averaging them. i.e. for each bin b, the histogram height $h_b^{\text{final}} = h_b^{(i)}$ for $i = 1, \ldots, 25$. The signal probability is obtained by applying the model to the MC signal sample.

The distributions of the classifier probabilities for background and signal(reweighted by MC SFs) in SR are shown in Figure 5.5. A good separation can be observed.

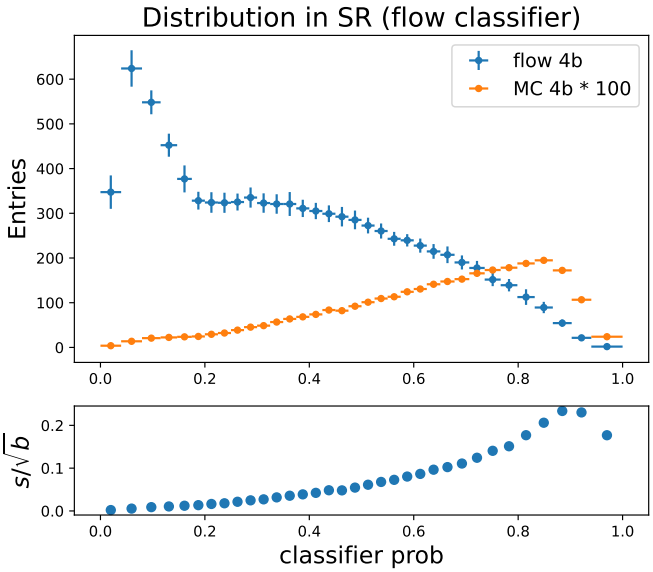


Figure 5.5: The distribution of the classifier probabilities estimations for the background and signal(reweighted by MC SFs), in the Signal Region. The histogram heights for MC events are scaled up by 100 for better visualisation. The uncertainties shown in the plots are the standard deviations of the 25 histograms.

To perform the validation, we similarly obtain the probability estimation of the real 4b events, and compare it with the probability estimation of the flow-generated 4b events, confined within the Control Region. As shown in Figure 5.6

5.1.2.3 Alternative systematic uncertainty estimation using eigenvalue decomposition

The current measure of the stability of the background modelling is the bootstrap uncertainty, which is the standard deviation of the ensemble of models (in the 2b reweighting method this is the standard deviation of the 100 bootstrap weights, and in the flow generated background estimates this is the standard deviation of the 25

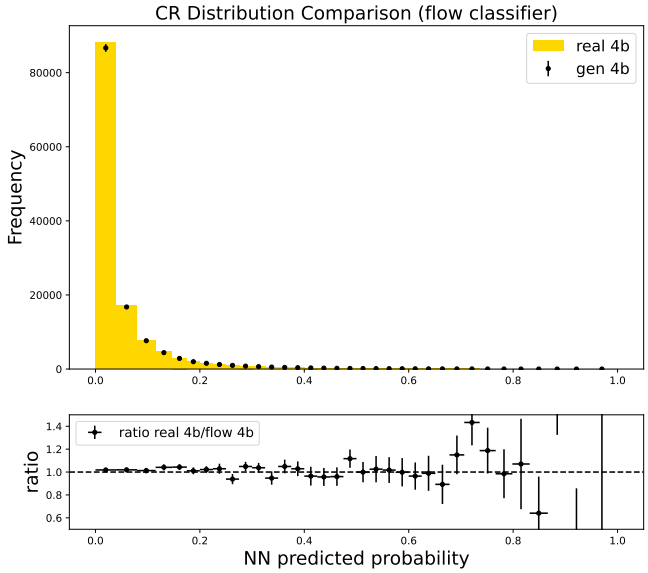


Figure 5.6: Control Region validation plots for **flow classifier**

models). This relies on the assumption that the variations are uncorrelated across bins. An alternative method of estimating the uncertainties using eigenvalue decomposition of the sample covariance matrix is introduced[159, 160]. This alternative systematic is used in the limit calculation for both the flow-based models.

Given N histograms (they are normally normalised during the calculation) over B bins, we denote the height of the j-th bin of the i-th histogram as b^i_j. i.e. the superscript i represents the numbering of the histograms and $i = 1,\ldots,N$, and the subscript j represents numbering of the bins and $j = 1,\ldots,B$. We build the $B \times B$ covariance matrices Q^i for each histogram, where

$$Q^i_{jk} = (b^i_j - \bar{b}_j)(b^i_k - \bar{b}_k) \tag{5.11}$$

where $\bar{b}_j = \frac{1}{N}\sum_{i=1}^{N} b^i_j$ is the average height of the j-th bin across the histograms. The **sample covariance matrix** Q is obtained as the average of Q^is

$$Q_{jk} = \frac{1}{N-1}\sum_{i=1}^{N} Q^i_{jk} \tag{5.12}$$

The diagonal entries of this sample covariance matrix are the variance of each bin across the histograms. In addition, the off-diagonal elements record the covariance between different bins.

We perform eigenvalue decomposition of Q to obtain eigenvalues λ_i corresponding eigenvectors v_i, for $i = 1, \ldots, B$. These eigenvalues and the eigenvectors represent magnitudes and the directions of the input histogram variations. Each $\sqrt{\lambda_i} v_i$ can be incorporated into the background shape Nuisance Parameter(NPs) for the limit calculation (see 4.8), allowing the inclusion of the full multivariate uncertainty structure of the histograms.

5.2 Results

The signal extraction fit is performed with the binned histograms of the probability scores, as introduced in 4.8.

Table 5.4 shows the expected limits obtained using the baseline background modelling method with m_{HH} as discriminant(m_{HH}) versus the **nominal classifier**; and the flow generated background estimates with m_{HH} as discriminant(flow m_{HH}) versus the **flow classifier**. For m_{HH} and the **nominal classifier**, only the background modeling related uncertainties discussed in Section 5.1.1.1 are considered. For the two flow-based models, the shape systematic uncertainty is inapplicable; and the variation of the modelling process is evaluated using the eigenvalue decomposition method discussed in Section 5.1.2.3.

Note that we should not directly compare the **nominal classifier** and **flow classifier**, since the differences lie not only within the classifier training, but also the background estimation methods they are trained on. The focus of this study is to compare the performances of the classifiers to their corresponding base models.

In general, good improvements in the sensitivity with the classifier can be observed, especially in the stats-only limit - 21% for the **nominal classifier** vs m_{HH}; and 22% for **flow classifier** vs flow m_{HH}.

This study shows a promising potential for using classifiers, or other multivariate analysis to derive probability estimates as the discriminant. It opens up an

model	stats-only limit	limit with systemics
m_{HH}(baseline)	6.07	6.83
nominal classifier	4.78	6.34
flow m_{HH}	4.23	4.44
flow classifier	3.44	4.39

Table 5.4: Limits for the baseline model fitted by m_{HH} distribution(m_{HH}), the **nominal classifier**, the flow-generated background estimation using m_{HH} as discriminant(flow m_{HH}), and the **flow classifier**.

encouraging exploration direction for the future of the 4b analysis or even other di-Higgs decay channels, for Run 3 and beyond.

Chapter 6

Conclusions

This thesis presents the search for HH in the $b\bar{b}b\bar{b}$ final state at the LHC with the ATLAS detector. The primary result is based on the analysis of the full Run 2 dataset, which uses 126 fb^{-1} of pp collision data at $\sqrt{13}$ TeV targeting both ggF and VBF production modes. No evidence for HH production is observed. The observed (expected) upper limit on the production cross-section is found to be 5.4 (8.1) times the Standard Model prediction at a 95% confidence level. The observed (expected) 2σ constraints on the HHH coupling modifier, κ_λ are [-3.5, 11.3] ([-5.4, 11.4]), while those for $HHVV$ coupling modifier κ_{2V} are [-0.0, 2.1] ([-0.1, 2.1]). The results are also used to derive constraints on the relevant coefficients for the beyond Standard Model SMEFT and HEFT theory frameworks. Significant improvements in sensitivity compared to previous analyses are observed, attributed to advances in analysis technique and object reconstruction.

Additionally, technical work aimed at improving track finding efficiency for b-jet signatures in the Inner Detector Trigger, undertaken as my ATLAS authorship qualification task, is presented. This work addressed the lower efficiency observed in Run 3 trigger tracking compared to Run 2 by implementing two key modifications: reducing the track p_T cut in precision tracking from 1 GeV to 0.8 GeV, which improved the integrated efficiency by approximately 1%, and the first-bin efficiency by a significant 10%; and replacing the χ^2 cut applied during track reconstruction and ambiguity solving process from 9 to 12, further improving the integrated efficiency by 1%. These modifications have been adopted and are currently in use for Run 3

data taking.

An extension of the $HH4b$ analysis is to train a classifier using several kinematic variables as input features, with the classifier's probability output serving as a discriminant, instead of only using m_{HH}. Besides the nominal data-driven background reweighting method, another classifier is trained on an alternative background estimation technique based on normalising flow. The classifier training showed significant improvements in the upper limit, particularly in the stats-only limit, with a 13% improvement for the nominal classifier, and a 22% improvement for the flow classifier. Although still in the exploratory phase, this approach holds great potential for enhancing analysis sensitivity and is encouraged to be further studied in future analysis.

An exploratory study on the use of analysis-level triggers for Run 3 is shown in the Appendix. This study suggests the appropriate trigger to use is a2b2j only, and suggests suitable transverse momentum cuts for the four leading jets.

Finally, a simulation study for exotic Higgs decay, $H \to ss \to 4b$, where ss is a pair of spin-zero particles, is presented in the Appendix. This study considers a scenario of 5000 fb^{-1} e^+e^- collision data at $\sqrt{s} = 240$ GeV from CEPC. The results indicate that CEPC could offer unprecedented sensitivity to this decay process, surpassing what can be achieved at the LHC.

Appendix A

Inner Detector Trigger ART

The ATLAS authorship qualification task to improve the track finding efficiency for *b*-jet signature described in Section3.3.2 is performed with the Inner Detector Trigger ART framework.

The ART (ATLAS Release Tester) system is used to validate performance changes in the nightly software releases. The nightly tests run on Athena tracks with the stages

1. **RDOtoRDOTrigger** - processes raw data objects (RDO) collected from the detector and applys trigger algorithms.

2. **RAWtoESD** - converts the selected raw data (RDOs) into Event Summary Data (ESD), which is a more processed form of the raw data and includes detailed offline reconstruction information.

3. **ESDtoAOD** - data reduction process to convert ESD into Analysis Object Data (AOD), a more compact format suitable for large-scale physical analysis.

We then execute a separate post-processing stage to read the AOD and write to our own flexible ntuple structure to facilitate subsequent performance analyses:

1. **TIDArdict** - standalone executable that performs track and event selection, track matching, and calculates resolutions and efficiencies, saving the results into histogram files.

2. **TIDAcomparitor** - plot the histograms generated by the **TIDArdict** stages

3. **TIDAcpucost** - plots the execution times from the athena expert timing histograms output by athena, or from the CostMonitoring.

Appendix B

Trigger Naming Conventions

The trigger names are composed of several components, each specifying a particular selection criterion. For Run 2 triggers (e.g. the triggers listed in Table 4.3), we have [161]:

- Trigger level identifier, such as **HLT** which stands for the High-Level Trigger.
- Jet requirements: **jX** means a jet with a p_T threshold of X GeV; **nJY** means n jets and each with a p_T threshold of Y GeV.
- *b*-tagging criteria: **bmv2c20Z** indicates the use of the MV2c20 *b*-tagging algorithm [139] with a working point of Z; **boffperf** means that the *b*-tagging performance optimised to match offline analysis criteria.
- Calibratoin and reconstruction: **gsc** means global sequential calibration, a method to improve jet energy measurements, e.g. j110_gsc150 means at least one jet with $p_T \geq 110$ GeV, calibrated using GSC to 150 GeV.; **split** indicates the use of a split processing mode.
- L1 trigger conditions: e.g. **L1JX** is L1 trigger requiring a jet with a p_T threshold of X GeV; **L1nJY** is L1 trigger requiring n jets and each with a p_T threshold of Y GeV; **L1nJY.0ETA25** is L1 trigger requiring n jets with $p_T \geq Y$ GeV and pseudorapidily $|\eta| < 2.5$.

For Run 3 triggers (e.g. the triggers listed in Table D.1), the naming convention has evolved and contain more detailed information [162]:

- The jet requirements are similar to Run 2. It can follow by ”c” to specify the jet is required to be within the central region of the detector, which is $|\eta| < 2.5$.
- *b*-tagging criteria: **bdl1dX** indicates the use of DL1d *b*-tagging algorithm [139] with working point X.
- **jvt**: Jet Vertex Tagger cut applied to supress pile-ups. Xjvt means applying the JVT cut with threshold X.
- **presel**: preselection cuts applied to reduce the cost of tracking.
- **pf_ftf**: indicates that the chain reconstructs particle flow jets which needs full scan fast track finding. If ”pf_ftf” is not specified then the chain runs topo-cluster-based jet reconstruction.

Appendix C

b-tagging Efficiency Study Supporting Plots

C.1 Supporting plots for lowering the min p_T cut

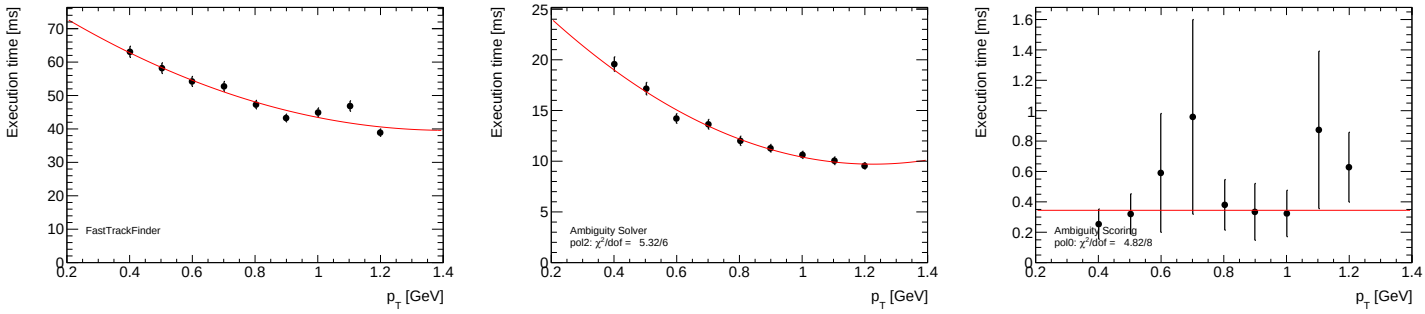
Figure C.1: the execution time per call for various p_T cuts, for the

- Fast Track Finder: explained in Section 3.3.0.2

- Ambiguity Solver: Resolves the track overlaps (which creates ambiguities in determining which hits in the detector correspond to which tracks) to ensure correct hits are assigned to the correct particle tracks. It collects scores from the Ambiguity Scorer, which assigns a score to each track; and then selects the best tracks by optimising the score.

- Ambiguity Scorer: assigns a score to each track based on how well it fits the detector hits and the expected particle trajectory

The Ambiguity Scoring time is negligible due to its rapid execution. As expected, the timing for FTF and Ambiguity Solver generally increases with lower p_T cuts. Reducing the minimum p_T cut from 1.0 GeV to 0.8 GeV results in an additional execution time of 5ms for FTF (from 40ms to 45ms), and 2ms out for Ambiguity Solver (from 10ms to 12ms), which is deemed acceptable.

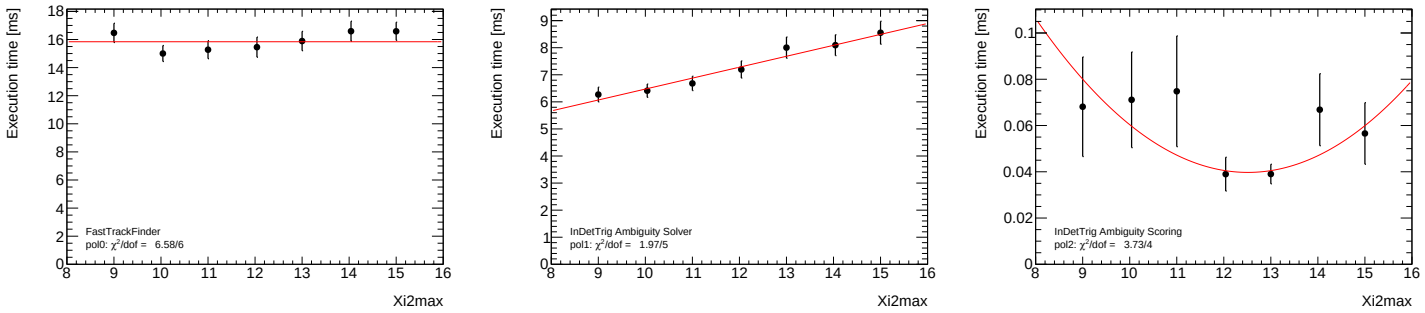


(a) FTF **(b)** Ambiguity Solver **(c)** Ambiguity Scoring

Figure C.1: Execution time per call vs minimum p_T cuts for (a) Fast Track Finder, (b) Ambiguity Solver and (c) Ambiguity Scoring.

C.2 Supporting plots for relax Xi2max cut

Figure C.2: the change in execution times for various Xi2max cuts. There is a slight increasing trend in the execution time for the Fast Track Finder. The execution time for the Ambiguity Solver increases by 1 ms when the Xi2max cut is raised from 9 (6 ms) to 12, and by an additional 2 mx if the cut is further increased to 15. The Ambiguity Scoring process remains very fast and negligible. Therefore, the increase in execution time when relaxing the Xi2max cut to 12 is acceptable.



(a) FTF **(b)** Ambiguity Solver **(c)** Ambiguity Scoring

Figure C.2: Execution time per call for (a) FTF, (b) Ambiguity Solver and (c) Ambiguity Scoring process with various Xi2max cuts.

In order to see if the increase in efficiency is due to fake tracks, we can use tighter track matching conditions to cut out tracks lying in the tails of residuals (see Figure C.3). As mentioned in 3.3.1.2, by default the tracks within a cone of size $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.05$ are considered, this is called the **Rmatcher** in the algorithm. Now we consider a stricter **DeltaRZSinThetamatcher** defined as

$\sqrt{\Delta RZ\sin\theta} = \sqrt{(\Delta R)^2 + (\Delta Z \cdot \sin\theta)^2}$. The matching conditions are also tightened, moving from $|\Delta\eta| < 0.02$, $|\Delta\phi| < 0.05$, $|\Delta Z\sin\theta| < 20$ to $|\Delta\eta| < 0.02$, $|\Delta\phi| < 0.01$ and $|\Delta Z\sin\theta| < 2$.

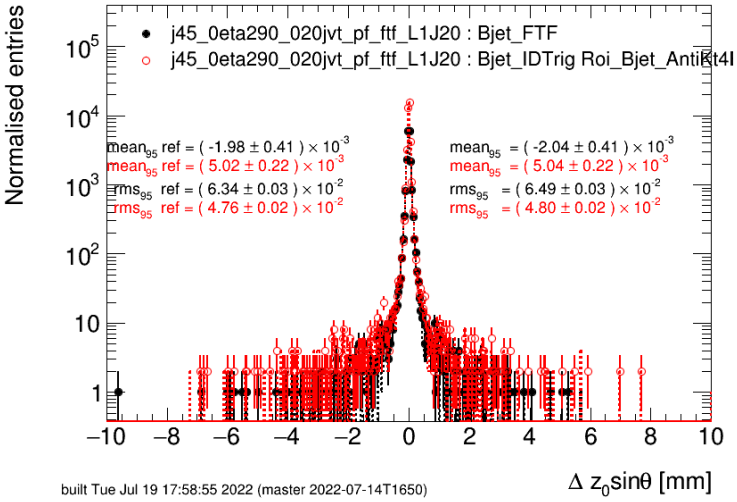


Figure C.3: The residuals of *b*-jet signature of Xi2max cut at 12 vs 9 wrt $\Delta z_0 \sin\theta$. The dots represent Xi2max=12, and the histograms represent Xi2max=9.

Figure C.4: the residuals of DeltaRZSinThetamatcher with tight matching conditions. We can see the tails of $\Delta\eta$, $\delta\phi$ and $\Delta z_0 \sin\theta$ are truncated, resulting in some genuine but less well-matched tracks being discarded.

Figure C.5: the improvement in efficiencies can still be observed with this tighter matcher, indicating this change in **cutLevel** does improve track efficiencies for high-quality central tracks without any significant increase in fake track multiplicities.

The final check is to look at the chi-squared values of the matched tracks. It is labeled as **Chi2dof** (chi-squared per degree of freedom) in the plots. Still using the tight DeltaRZSinThetamatcher, Figure C.6 shows the distributions of the Chi2dof for the matched tracks. Slightly higher multiplicities of matched tracks at higher Chi2dof regions are observed after relaxing the Xi2max cut to 12.

Figure C.7: the mean and standard deviation of $\Delta\chi^2/\text{dof} \equiv \chi^2/\text{dof}$(trigger track)-$\chi^2/\text{dof}$(reference track), for each pair of matched tracks. We can see that $\Delta\chi^2/\text{dof}$ does run out of statistics for high Chi2dof values. When Chi2dof is around 4 or 5, the standard deviation is nearly 2, indicating that the original cut at 9 would result in discarding some matched tracks.

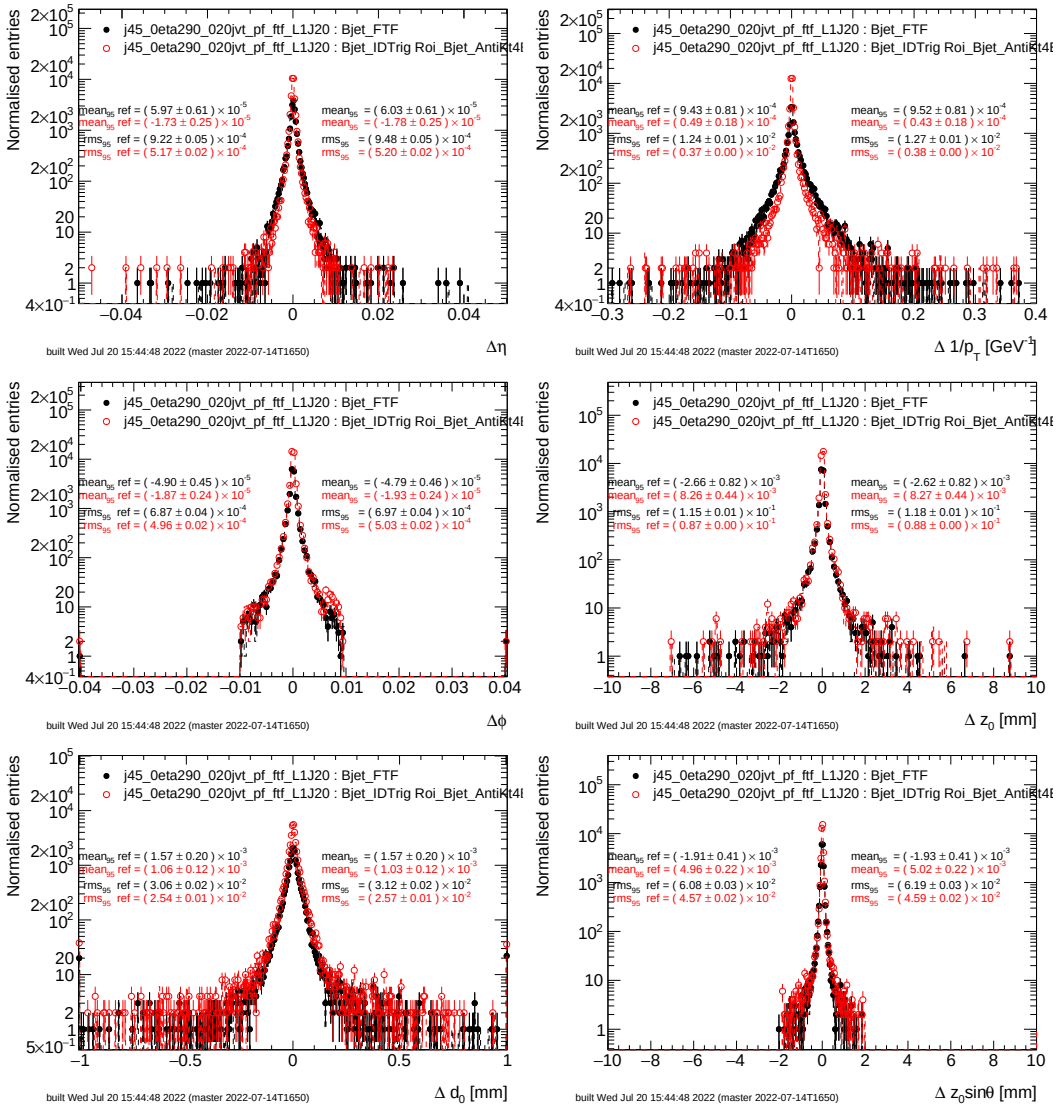


Figure C.4: The residuals of DeltaRZSinThetamatcher with tight matching conditions. The dots represent residuals of Xi2max=12, and the histograms represent efficiencies of Xi2max=9.

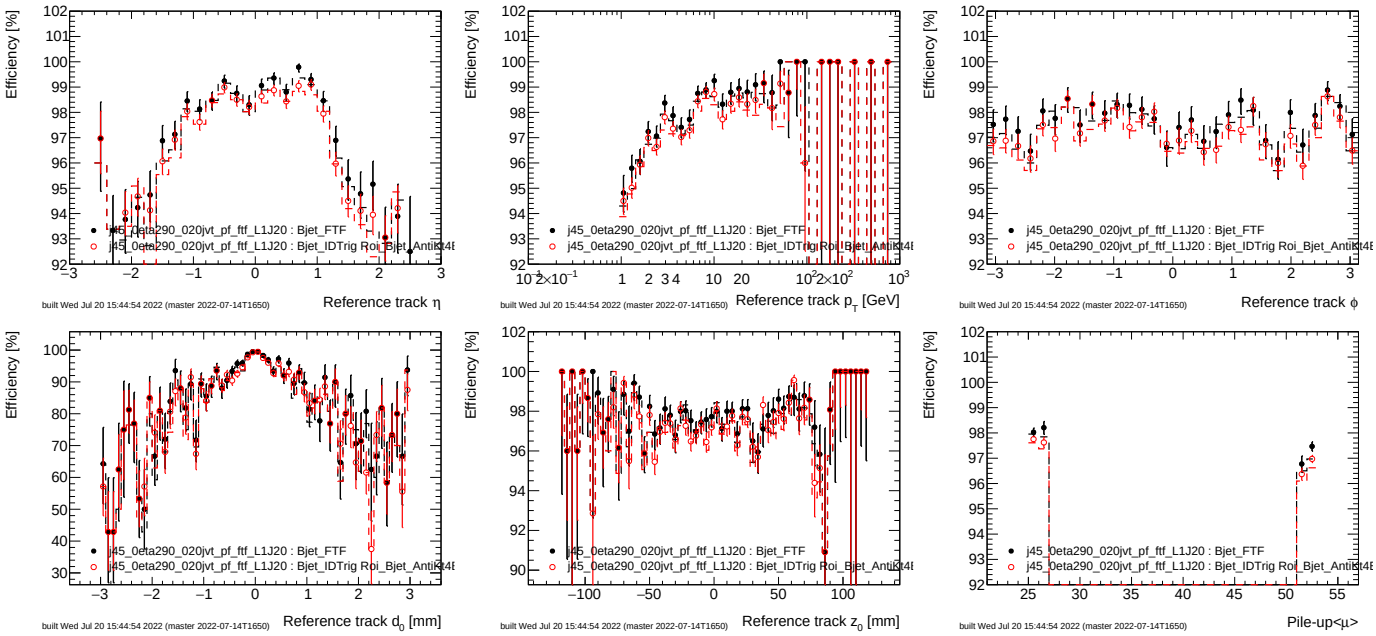
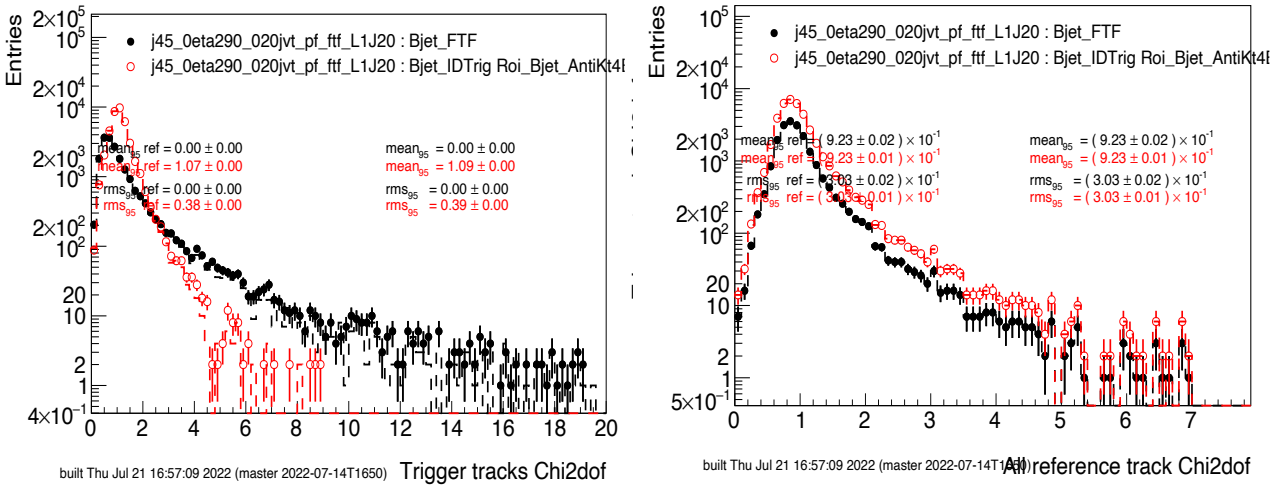
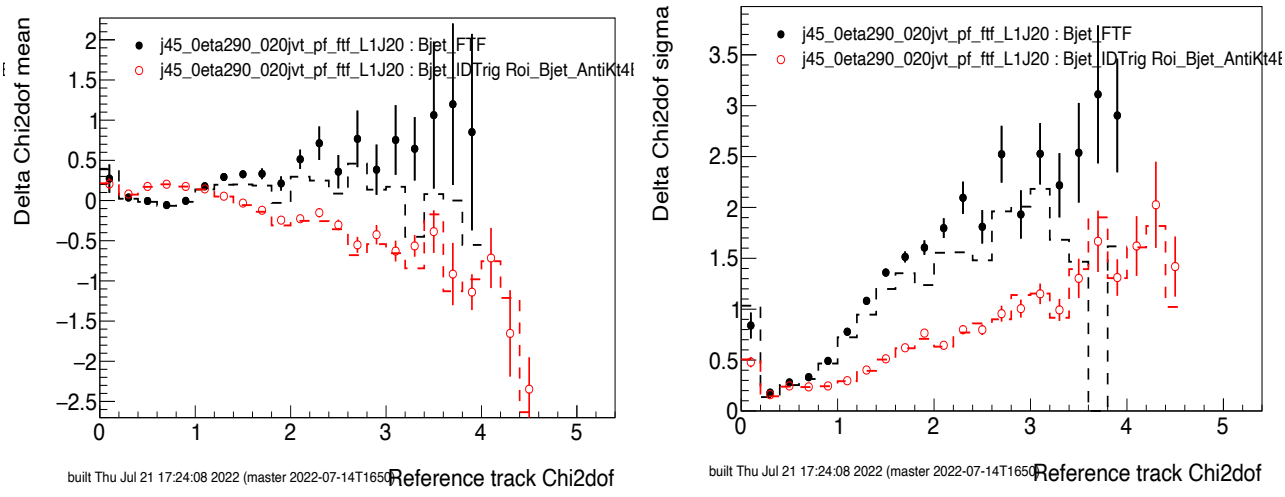


Figure C.5: The efficiencies of DeltaRZSinThetamatcher with tight matching conditions. The dots represent efficiencies of Xi2max=12, and the histograms represent efficiencies of Xi2max=9. Improvement can still be observed.



(a) Distributions of matched trigger tracks wrt Chi2dof **(b)** Distributions of matched reference tracks wrt Chi2dof

Figure C.6: Distributions of matched (a)trigger and (b)reference tracks with respect to Chi2dof. The dots represent Xi2max=12, and the histograms represent Xi2max=9.



(a) Mean **(b)** Standard Deviation

Figure C.7: (a)Mean and (b)standard deviation of ΔChi2dof for matched tracks, with respect to Chi2dof for the reference tracks. The dots represent Xi2max=12, and the histograms represent Xi2max=9.

Appendix D

Triggers for Run3

This chapter discusses a study investigating the triggers to use for Run3 ggF resolved analysis.

The multi-b-jet triggers to investigate are listed in Table D.1. The key trigger to consider is a2b2j, which is an asymmetric chain requiring four jets with p_T greater than 80, 55, 28 20 GeV. It also requires at least two jets with $p_T > 20$ GeV, b-tagged with the DL1d algorithm at the 77% working point, with preselection of jets with $p_T > 20$ GeV and b-tagging at 85% WP. a2b2jMU has the same requirement on the jet b-tagging and their p_T, with an additional requirement of a muon with $p_T > 8$ GeV along with jets. While having the same jet p_T requirement as the previous two triggers, a3b1j requires three b-tagged jets with $p_T > 20$ GeV, b-tagged at 82% WP. 2b2j is a symmetric chain, requires two jets with $p_T > 35$ GeV, and both are b-tagged at 60% WP. Its preselection requires two jets with $p_T > 25$ GeV and b-tagging at 85% WP. The final 2b1j trigger requires one high-energy jet with $p_T > 150$ GeV and two additional jets with $p_T > 55$ GeV, with two b-tagged jets at 70% WP. Its preselection requires a jet with $p_T > 80$ GeV and two jets with $p_T > 45$ GeV, along with b-tagging at 90% WP. See Appendix B for explanations of the naming conventions.

The data used is the mc21 sampleT data produced during the long shutdown. These are high-statistics samples that provide a good reference for performance. This study uses $\kappa_\lambda = 1$ and $\kappa_\lambda = 10$, with the selection of at least four jets with $p_T \geq 25$ GeV, and at least 2 of them are b-tagged with DL1dv01 77% working point (see

Section 4.4 and 4.5 for more details about the selection).

Trigger	Trigger Type	Trigger Name
Asymmetric 2b2j	a2b2jd	HLT_j80c_020jvt_j55c_020jvt_j28c_020jvt_j20c_020jvt_SHARED_2j20c_020jvt_bdl1d77_pf_ftf_presel2c20XX2c20b85_L1J45p0ETA21_3J15p0ETA25
Asymmetric 2b2j + muon	a2b2jdMU	HLT_j80c_020jvt_j55c_020jvt_j28c_020jvt_j20c_020jvt_SHARED_2j20c_020jvt_bdl1d77_pf_ftf_presel2c20XX2c20b85_L1MU8F_2J15_J20
Asymmetric 3b1j	a3b1jd82	HLT_j80c_020jvt_j55c_020jvt_j28c_020jvt_j20c_020jvt_SHARED_3j20c_020jvt_bdl1d82_pf_ftf_presel2c20XX2c20b85_L1J45p0ETA21_3J15p0ETA25
Symmetric 2b2j	2b2jd	HLT_2j35c_020jvt_bdl1d60_2j35c_020jvt_pf_ftf_presel2j25XX2j25b85_L14J15p0ETA25
Symmetric 2b1j	2b1jd	HLT_j150_2j55_0eta290_020jvt_bdl1d70_pf_ftf_preselj80XX2j45b90_L1J85_3J30

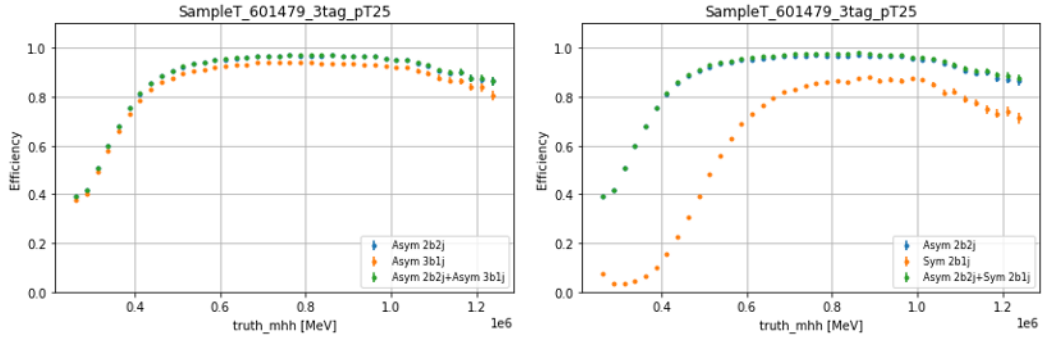
Table D.1: Triggers to study for the Run 3 ggF resolved analysis

Table D.2 shows the efficiencies of the triggers, calculated as the number of events passing the triggers divided by the total number of events, with baseline selections applied. Among the triggers, a2b2j has the highest efficiency, which makes it suitable to use as the baseline. We can use it as a reference point for pairwise comparisons with the efficiencies of the other triggers.

Triggers	$\kappa_\lambda = 1$ Efficiency	$\kappa_\lambda = 10$ Efficiency
a2b2j	0.828	0.604
a3b1j	0.803	0.584
a2b2jMU	0.316	0.217
2b2j	0.567	0.452
ab1j	0.337	0.146

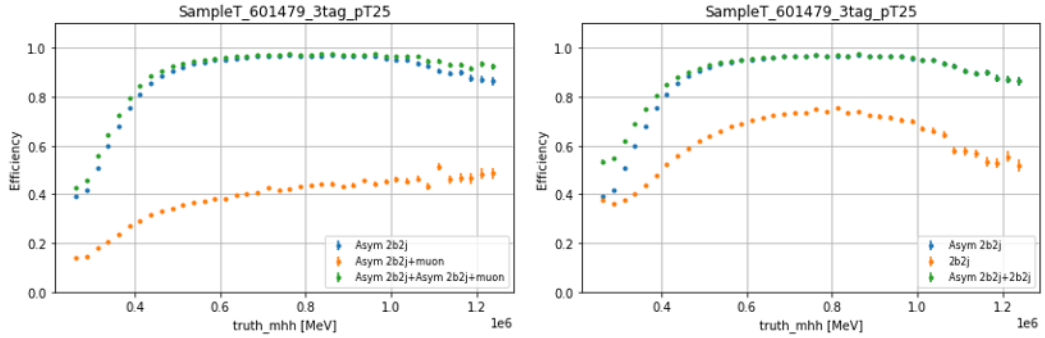
Table D.2: Efficiencies of the triggers at $\kappa_\lambda = 1$ and $\kappa_\lambda = 10$.

Figure D.1 and Table D.3 shows the pairwise efficiency comparisons of a2b2j vs other triggers for $\kappa_\lambda = 1$ samples as an example. We can see that in general most of the efficiencies come from the contribution of a2b2j, with the combined efficiency only slightly better than the efficiency of only a2b2j. a2b2jMU shows improvements in efficiency at higher m_{HH} region, and 2b2j shows improvements at lower m_{HH} regions, but the overall improvement is still minor. $\kappa_\lambda = 10$ samples give similar observations.



(a) a2b2j vs a3b1j

(b) a2b2j vs 2b1j



(c) a2b2j vs a2b2jMU

(d) a2b2j vs 2b2j

Figure D.1: Pairwise comparisons for the triggers of interests with a2b2j, for the $\kappa_\lambda = 1$ samples. The blue dots represent the bin-wise efficiencies of a2b2j trigger, the orange dots represent the efficiencies of the trigger to compare, and the green dots are their combined efficiencies.

Trigger of interest	Trigger Efficiency	Trigger + a2b2j Efficiency	Extra contribution
a2b2j (baseline)	0.828	0.828	0
a3b1j	0.803	0.829	0.001
ab1j	0.337	0.833	0.004
a2b2jMU	0.316	0.855	0.027
2b2j	0.567	0.861	0.032

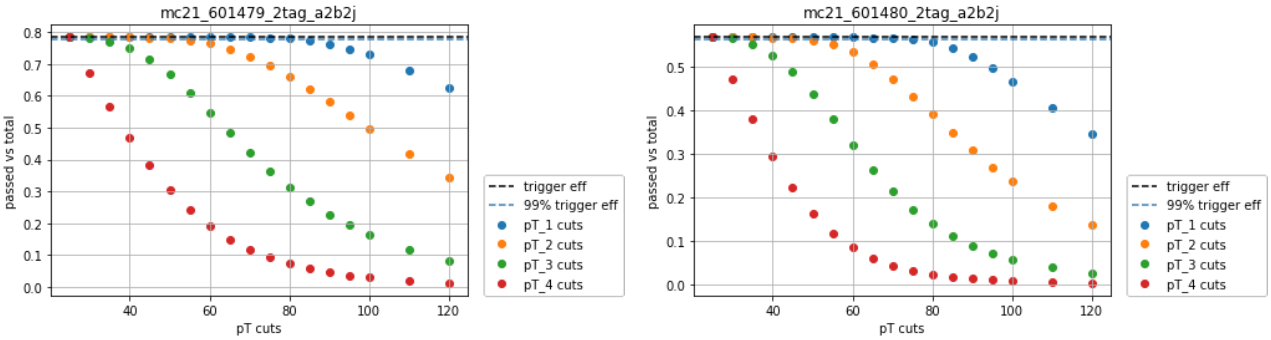
Table D.3: Pairwise efficiency comparisons of a2b2j vs other triggers, for $\kappa_\lambda = 1$ samples.

This motivates the proposal to adopt a single a2b2j trigger for Run 3, as the other triggers largely overlap with a2b2j. Additionally, a2b2j is part of the delayed stream, a dataset separate from the main stream where the other triggers are located. Combining triggers from two different datasets can introduce complexities.

Cuts are applied to the transverse momentum of the first four leading jets (denoted as p_{T1} to p_{T4}) to save processing time. It is important to investigate the impact of these cuts on trigger efficiencies, as the trigger efficiency increases sharply

for events with higher jet transverse momentum. The goal is to determine appropriate cuts without discarding too many high-quality events.

Figure D.2 shows a2b2j trigger efficiency as a function of the cuts applied to the transverse momentum of the four leading jets of $\kappa_\lambda = 1$ and $\kappa_\lambda = 10$ samples. All events must pass an additional selection requiting at least two b-tagged jets as events with fewer than two b-tagged jets are of less interest for this analysis. The efficiency is defined as the ratio of events passing the p_T cuts and the a2b2j trigger, to the total number of events (without any cuts or triggers applied). Table D.4 presents the overall trigger efficiency (the fraction of events passing the trigger out of the total number of events, both without applying p_T cuts) and the highest p_T cuts that can be applied while maintaining a loss of less than 1% efficiency.



(a) $\kappa_\lambda = 1$ **(b)** $\kappa_\lambda = 10$

Figure D.2: Efficiencies of a2b2j trigger with different jet p_T cuts for (a)$\kappa_\lambda = 1$ and (b)$\kappa_\lambda = 10$ samples. The black dashed line is the overall trigger efficiency, and the blue dashed line represents 99% of the overall trigger efficiency. The dots represent the ratio of events passing a2b2j trigger and the corresponding p_T cut, vs the total number of events. Base selection of having at least 2 b-tagged jets is applied to all events.

Sample	trigger efficiency	99% efficiency	p_{T1} cut	p_{T2} cut	p_{T3} cut	p_{T4} cut
$\kappa_\lambda = 1$	0.785	0.777	80	50	30	25
$\kappa_\lambda = 10$	0.569	0.563	75	45	30	25

Table D.4: The overall a2b2j trigger efficiency, 99% threshold of the overall trigger efficiency, and the maximum p_T cuts allowed to preserve at least 99% of the trigger efficiency, for (a)$\kappa_\lambda = 1$ and (b)$\kappa_\lambda = 10$ samples.

We concluded that suitable cuts can be applied is $p_{T1} > 75, p_{T2} > 45, p_{T3} > 30$ and $p_{T4} > 25$, resulting in trigger efficiencies of 0.778 for $\kappa_\lambda = 1$ samples, and 0.558

for $\kappa_\lambda = 2$ samples, both lose only approximately 1% of efficiency.

Appendix E

Study of Electroweak Phase Transition in Exotic Higgs Decays at the CEPC

This chapter focuses on the simulation study [8, 18] about the search for exotic decays of the Higgs boson into a pair of spin-zero particles, then these particles further decay into b-quarks, i.e. $h \to ss \to 4b$, considering a scenario of analysing 5000 fb^{-1} e^+e^- collisions data at $\sqrt{s} = 240$ GeV from the Circular Electron Positron Collider (CEPC) [163].

E.1 Introduction

Within the BSM scenarios, a strong first-order electroweak phase transition (SFOEWPT) is expected, which provides one of the necessary conditions for electroweak baryogenesis and gravitational wave background [164, 165]. Any new particles involved in such alternative thermal history can't have masses too heavy with respect to the electroweak temperature scale, and they can't interact too weakly with the SM Higgs boson [166]. This work considers a lighter BSM singlet particle s coupled to the SM Higgs h to catalyze a Higgs Boson, enabling exotic Higgs decay modes.

The future e^+e^- collider is projected to possess sufficient sensitivity to probe the SFOEWPT for new scalar masses down to approximately 10 GeV [9, 167]. This

study focuses on the strong first-order electroweak phase transition with the light real single, S, coupled with SM Higgs, H. The relevant scalar potential involving S and H is [9, 168]:

$$V = -\mu^2|H|^2 + \lambda|H|^4 + \frac{1}{2}a_2|H|^2S^2 + b_1S + \frac{1}{2}b_2S^2 + \frac{1}{3}b_3S^3 + \frac{1}{4}b_4S^4 \tag{E.1}$$

After electroweak symmetry breaking, the fields can be parametrized as

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix}, \quad S = v_s + s \tag{E.2}$$

where v_s is the vacuum expectation value(VEV) for the singlet, which is set to 0.

The mass eigenstates are the mixing of two scalar fields h and s:

$$h_1 = h\cos\theta + s\sin\theta, \tag{E.3}$$

$$h_2 = -h\sin\theta + s\cos\theta \tag{E.4}$$

Where h_1 is the singlet-like particle with mass m_1, and h_2 is the SM-like Higgs particle with $m_2 = 125$GeV. We consider the case in which h_2 decays visibly and promptly, which requires the mixing angle of the SM Higgs and the light scalar $\cos\theta \neq 0$. We can take $\cos\theta = 0.01$, sufficient for h_1 to decay promptly [9]. The trilinear scaler interactions in terms of mass eigenstates become

$$V = \frac{1}{6}\lambda_{111}h_1^3 + \frac{1}{2}\lambda_{211}h_2h_1^2 + \frac{1}{2}\lambda_{221}h_2^2h_1 + \frac{1}{6}\lambda_{222}h_2^3 \tag{E.5}$$

The cublic coupling λ_{211} is

$$\lambda_{211} = 2s^3cb_3 + \frac{a_1}{2}c(c^2-2s^2) + (2c^2-s^2)sva_2 - 6\lambda sc^2v \tag{E.6}$$

where $s \equiv \sin\theta$ and $c \equiv \cos\theta$. In this singlet extension model, only a_2, b_3, b_4 in the potential V are free parameters.

We consider m_1 in the range [5,60] GeV to ensure the occurrence of the exotic

Higgs decay and set $\cos\theta = 0.01$ to comply with the experimental constraints. The exotic Higgs decay provides a powerful probe for the small mixing angle $\cos\theta$, which can be detected up to $\mathcal{O}(0.01)$ at future colliders. For given values of m_1 and $\cos\theta$, we scan over $a_2, b_3/v \in [10^{-4}, 1]$ and $b_4 \in [10^{-5}, 1]$ in our numerical simulations.

At the center-of-mass energy of 240 GeV, the dominant Higgs production channel is the ZH process. In this exotic decay scenario, the SM Higgs boson is produced in association with a Z boson. Therefore the search focuses on the ZH process with $Z \to ll$ for $(l = e, \mu)$, and $H \to ss \to 4b$, as shown in fig E.1.

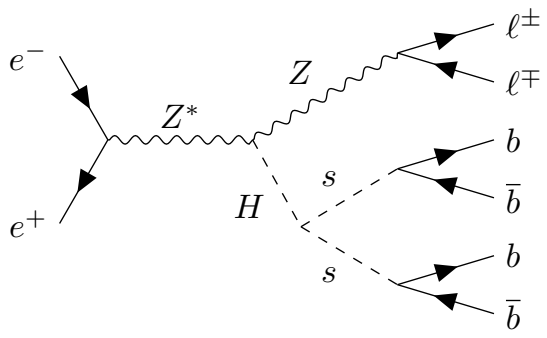


Figure E.1: Representative tree-level Feynman diagram for the ZH production process with subsequent decays $Z \to ll(l = e, \mu)$ and $H \to ss \to 4b$ [8]

.

E.2 CEPC Detector Concept

The CEPC detector is designed to fulfil the requirements of various physics programs, particularly for precise measurements. It has a low material tracking system, a high granularity calorimeter system and a flux return yoke embedded with a muon detector, from the innermost to the outermost. The tracking system comprises a silicon pixel vertex detector, a silicon inner tracker, a Time Projection Chamber and a silicon external tracker. This configuration provides a resolution of 3-5% for jet energies between 20 and 100GeV. It can also provide a 50 ps resolution for time measurement and can be used for particle identification.

E.3 Event Simulation

The study uses MADGRAPH5_aMC@NLO [128] and WHIZARD [169] Monte Carlo event generators for the signal and background events generation. The singlet-like exotics Higgs signal model is implemented and imported to MADGRAPH5_aMC@NLO using FEYN-RULES [170]. Signal events are generated with MADGRAPH5_aMC@NLO at leading order (LO), with Parton showering and hadronization modelling performed using PYTHIA8 [124]. Signal samples are generated for different singlet-like exotic Higgs masses starting from 15 GeV and 60 GeV in 5 GeV increments. The SM background events are generated using WHIZARD, with the parton showering and hadronization modelling also done using PYTHIA8. All background and signal samples are generated at non-polarized electron-position collision at centre of mass energy of $\sqrt{s} = 240$ GeV. The detector simulation is carried out using Mokka [171], a Geant4 [120] based detector simulation software. The simulated hits are digitized and reconstructed with ArborPFA [172]. All the backgrounds are modelled using the samples simulated by CEPC.

E.4 Object Reconstruction

The charged leptons are identified using the Lepton Identification in Calorimeter with High Granularity (LICH) algorithm8 [173], specifically designed for future detectors using high-granularity calorimeters. Electrons and muons are identified with efficiencies of 99.7% and 99.9%, respectively, while keeping misidentification rates smaller than 0.07%. Lepton isolation is applied by requiring $E^2_{\text{cone}} < 4E_l +$ 12.2 GeV, where E_l is the lepton energy and E_{cone} is the energy within a cone of $\cos\theta_{\text{cone}} > 0.98$ around the lepton. Jets in the events are reconstructed from the particle flow object (PFO) using the LCHIPlus Software package [174]. Leptons are removed from the ArborPFO before the jet reconstruction. Jets are reconstructed using the Durham algorithm [175], with exclusive clustering performed to reconstruct exactly four jets in the final states.

E.5 Event Selection

Events are selected targeting four b-jets coming from the singlet-like exotics Higgs and two same-flavour opposite-sign leptons coming from the Z-boson decay. The events are selected in two stages using a set of loose preselections, followed by a classifier using a Boosted Decision Tree (BDT).

The preselection requirements of the events are:

- have two well-isolated leptons with opposite charges and each having energy higher than 20 GeV.
- The polar angle between the two leptons satisfy $|\cos\theta_{e^+e^-}| < 0.71$ and $|\cos\theta_{\mu^+\mu^-}| < 0.81$
- The angle between two isolated tracks satisfies $\cos\theta > -0.72$.
- The invariant mass of the lepton system is within the Z-mass window of 77.5-104.5 GeV
- The recoil mass of the dilepton system is in the range of 124-140 GeV.

After selecting events with two leptons and four jets, the singlet-like exotic scalar is reconstructed by combining the jets.

The remaining particles in the event are used to reconstruct exactly four jets with polar angle θ_{jet} within the range of $|\cos\theta_{\text{jet}}| < 0.96$ and each jet containing at least 40 particles, each with an energy no less than 0.4 GeV.

Flavour tagging significantly enhances background suppression and improves the reconstruction of M_s. This is accomplished using the multi-variable-based flavour tagging toolkit in LCFIPLUS, which calculates the b-likeliness value for each jet, and a combined b-likeliness, or b-jet efficiency f_b. Selected events are required to have $\log(1-f_b) < -4.0$, and can be used to reconstruct the singlet scalar.

An alternative approach is to train a BDT to separate the singlet scalar Higgs decay signals from the backgrounds of SM processes. To maximize the sensitivity of signal samples at every mass point, separate BDTs are trained for different signal masses, thus we have 10 models. For each model, the hyperparameters are optimized

using Optuna, a black-box hyperparameter tuning framework. An area under the curve (AUC) of approximately 0.99 is achieved for every BDT.

In addition to the aforementioned variables, other variables with discriminating power, such as the opening angles between the two reconstructed exotic Higgs and the minimum distance between the particles used in the jet reconstruction, are used for BDT training. A total of 24 input variables are used in the BDT Training, as summarised in table E.1

Variable	**Description**
p_T^{l1}, p_T^{l2}	Lepton p_T
$M_{\ell\bar{\ell}}^{\text{recoil}}$	Recoil mass of the di-lepton system
$E^{j0}, E^{j1}, E^{j2}, E^{j3}$	Energy of each reconstructed jet
M^{4j}	Invariant mass of the four-jet system
M_{4j}^{recoil}	Recoil mass of the four-jet system
M_{bb}	Average mass of the two reconstructed s candidates
M_{diff}^{ss}	Difference between the two reconstructed s particles
N_{particle}^{4j}	Number of particles inside the four jets
$y_{12}, y_{23}, y_{34}, y_{45}, y_{56}, y_{67}$	Distance between jet constituents
$\cos\theta_{\text{boost}}^{s_1}, cos\theta_{\text{boost}}^{s_2}$	Boosted angle between two jets from each s candidate
$\cos\theta_s^{\text{open}}$	Opening angle between the reconstructed s particles
$\lvert\cos\theta_{j0}^{\text{Helicity}}\rvert, \lvert\cos\theta_{j2}^{\text{Helicity}}\rvert$	Helicity of the first and the third jets
b_{ineff}	b-jet inefficiency

Table E.1: Variables used for the BDT training. Helicity angle of the first jet and the third jet is used because the other two jets have the same value.

The background contributions are categorised into three groups: Higgs processes with bb final state ($llHbb$) (since the $H \to bb$ process is the dominant background), other Higgs processes(llH), and non-Higgs processes.

The main systematic uncertainty is the event yields of the fixed background. Event yields of primary $llHbb$ background are varied up and down by 5% while other background processes are varied by 100% to estimate the uncertainties from background modelling [173]. Contributions from luminosity [176] and lepton iden-

tification [173] are considered negligible and therefore ignored. Corresponding uncertainties with flavour tagging (following the method from [177]) and jet energy resolution(JER)(calculating the energy difference between a truth jet and a reconstructed jet using the $e^+e^-k_{t-}$ algorithm from the FastJet package [178]) are considered.

The JER uncertainty tends to be negligible after the statistical fit. The post-fit flavour tagging uncertainty is $\tilde{1}\%$. The background modelling uncertainty is around 4% post fit. The contribution from non-Higgs processes is between 2-5% above $M_S = 20$ GeV, and around 10% at $M_s = 15$ GeV. For the non-Higgs processes, it's between 10-30%.

E.6 Results

The output BDT score is used as the main discriminant. The bins of the BDT score distribution are used to define a test statistic based on the profile likelihood ratio. The fit is performed to calculate the expected upper limit on the cross-section at 95% confidence level using the CL_s method [152, 153, 179], as described in Section 4.8. All the systematic uncertainties discussed above and the statistical uncertainty are considered. The upper limits for the singlet signal for each m_s are shown in table E.2, and illustrated in fig E.2 with corresponding uncertainties.

This simulation study shows that CEPC can have an unprecedented sensitivity in the search for such scalar particles coming from Higgs decay compared to the HL-LHC at this low mass range.

Expected 95% CL upper limits	
M_s [GeV]	$\sigma_{ZH} \times \mathrm{B}(H \to ss)/\sigma_{\mathrm{SM}}$
15	4.07×10^{-4}
20	3.50×10^{-4}
25	5.45×10^{-4}
30	4.49×10^{-4}
35	5.34×10^{-4}
40	5.50×10^{-4}
45	6.32×10^{-4}
50	6.63×10^{-4}
55	6.09×10^{-4}
60	4.58×10^{-4}

Table E.2: The expected 95% CL upper limit for the product of the signal's cross-section times branching ratio relative to the Standard Model cross-section. The results are presented for the combined electron e^+e^-H and muon $\mu^+\mu^-H$ channels using the multivariate BDT approach.

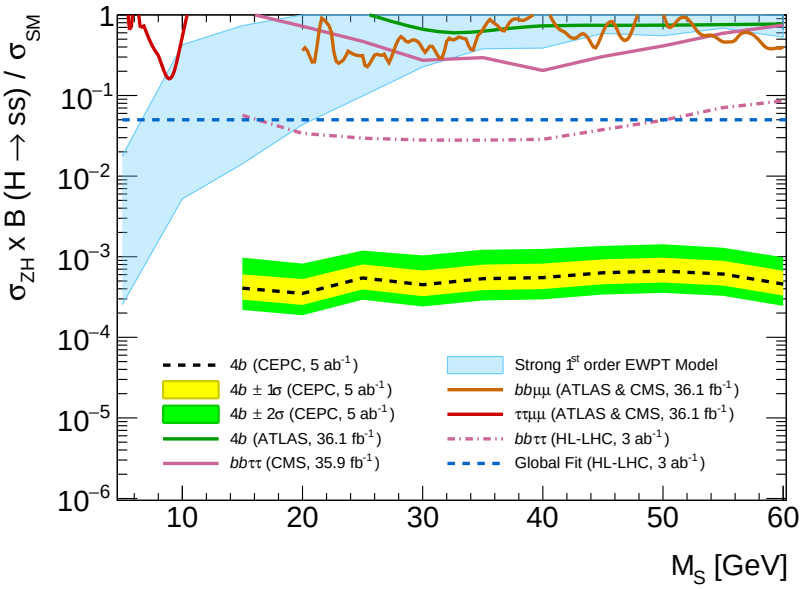


Figure E.2: The ratio of expected upper limits at 96% CL on the $\sigma_{\mathrm{ZH}} \times B(H \to ss)$ and SM cross section (σ_{SM}) vs singlet mass M_s. The expected limits obtained from the BDT analysis are represented by black dotted line, The blue-shaded region indicates points that predict a strong first-order electroweak phase transition with successful tunneling [9]. The current (solid) and projected (dashed) sensitivities(take the best limit between ATLAS and CMS for each mass point) of the LHC in the $\tau\tau\mu\mu$ [10, 11], $bb\mu\mu$ [12, 13], $bb\tau\tau$ [14, 15] and $4b$ [16] channels are also shown. The HL-LHC projection for the indirect limit on the total exotic branching fraction is displayed as the blue dashed line [17]. [18]

Bibliography

[1] Particle Data Group. Status of higgs boson physics, 2023.

[2] Wikipedia contributors. Pseudorapidity — Wikipedia, the free encyclopedia, 2024. [Online; accessed 11-October-2024].

[3] Georges Aad et al. The ATLAS inner detector trigger performance in pp collisions at 13 TeV during LHC Run 2. *Eur. Phys. J. C*, 82(3):206, 2022.

[4] Georges Aad et al. Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at s=13 TeV with the ATLAS detector. *Phys. Rev. D*, 108(5):052003, 2023.

[5] M. Capozi and G. Heinrich. Exploring anomalous couplings in higgs boson pair production through shape analysis. *Journal of High Energy Physics*, 2020(3), March 2020.

[6] Dale Abbott, William Keaton Balunas, Alessandra Betti, Raif Rafideen Bin Norisam, Lucas Santiago Borgna, Jay Chan, Arely Cortes Gonzalez, Marta Czurylo, Alex Emerman, James Frost, Sean Joseph Gasiorowski, Anna Goussiou, James Cameron Grundy, Chris Gubbels, Dan Guest, Michael Donald Hank, Nicole Michelle Hartman, Shota Hayashida, Yu Nakahama Higuchi, Todd Brian Huffman, Rachel Jordan Hyneman, Cigdem Issever, Michael Aaron Kagan, Teng Jian Khoo, Young-Kee Kim, Nikolaos Konstantinidis, Katharine Leney, Charles Lewis, Chris Milke, Marin Mlinarevic, Jana Schaarschmidt, Todd Michael Seiss, Stephen Jacob Sekula, Liaoshan Shi, Mel Shochet, Beojan Stanislaus, Maximilian J Swiatlowski, Rafael Teixeira

De Lima, Marco Valente, Iza Veliscek, Alex Zeng Wang, Stephane Willocq, Sau Lan Wu, Rui Zhang, Yuwen Zhang, Chen Zhou, Thandikire Madula, Tom Ingebretsen Carlson, and Joergen Sjoelin. Supporting Document: The Search for Non-Resonant ggF and VBF HH Production Decaying to the 4b Final State Using the Full Run-2 Data. Technical report, CERN, Geneva, 2021.

[7] J. Baglio, F. Campanario, S. Glaus, M. Mühlleitner, J. Ronca, and M. Spira. $gg \to HH$: Combined uncertainties. *Phys. Rev. D*, 103(5):056002, 2021.

[8] Zhen Wang, Xuliang Zhu, Elham E. Khoda, Shih-Chieh Hsu, Nikolaos Konstantinidis, Ke Li, Shu Li, Michael J. Ramsey-Musolf, Yanda Wu, and Yuwen E. Zhang. Study of Electroweak Phase Transition in Exotic Higgs Decays at the CEPC. In *Snowmass 2021*, 3 2022.

[9] Jonathan Kozaczuk, Michael J. Ramsey-Musolf, and Jessie Shelton. Exotic higgs boson decays and the electroweak phase transition. *Physical Review D*, 101(11), June 2020.

[10] Georges Aad et al. Search for Higgs bosons decaying to *aa* in the $\mu\mu\tau\tau$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment. *Phys. Rev. D*, 92(5):052002, 2015.

[11] Albert M Sirunyan et al. Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 800:135087, 2020.

[12] Morad Aaboud et al. Search for Higgs boson decays into a pair of light bosons in the $bb\mu\mu$ final state in pp collision at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Lett. B*, 790:1–21, 2019.

[13] Albert M Sirunyan et al. Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV. *Phys. Lett. B*, 795:398–423, 2019.

[14] Projection of searches for exotic Higgs boson decays to light pseudoscalars for the High-Luminosity LHC. 2019.

[15] Albert M Sirunyan et al. Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two b quarks and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 785:462, 2018.

[16] M. Aaboud et al. Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the $H \to aa \to 4b$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 10:031, 2018.

[17] M. Cepeda et al. Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC. *CERN Yellow Rep. Monogr.*, 7:221–584, 2019.

[18] Zhen Wang, Xuliang Zhu, Elham E. Khoda, Shih-Chieh Hsu, Nikolaos Konstantinidis, Ke Li, Shu Li, Michael J. Ramsey-Musolf, Yanda Wu, and Yuwen E. Zhang. Probing Electroweak Phase Transition at CEPC via Exotic Higgs Decays with 4b Final States. *LHEP*, 2023:436, 2023.

[19] K. A. Olive et al. Review of Particle Physics. *Chin. Phys. C*, 38:090001, 2014.

[20] Sarkamo J Peltoniemi J. Laboratory measurements and limits for neutrino properties.

[21] B. Grzadkowski, M. Iskrzyński, M. Misiak, and J. Rosiek. Dimension-six terms in the standard model lagrangian. *Journal of High Energy Physics*, 2010(10), October 2010.

[22] Combined effective field theory interpretation of Higgs boson and weak boson production and decay with ATLAS data and electroweak precision observables. 2022.

[23] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[24] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012.

[25] Morad Aaboud et al. Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector. *Phys. Lett. B*, 786:59–86, 2018.

[26] A. M. Sirunyan et al. Observation of Higgs boson decay to bottom quarks. *Phys. Rev. Lett.*, 121(12):121801, 2018.

[27] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[28] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[29] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[30] Morad Aaboud et al. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 01:030, 2019.

[31] Georges Aad et al. Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 07:108, 2020. [Erratum: JHEP 01, 145 (2021), Erratum: JHEP 05, 207 (2021)].

[32] Armen Tumasyan et al. Search for Higgs Boson Pair Production in the Four b Quark Final State in Proton-Proton Collisions at s=13 TeV. *Phys. Rev. Lett.*, 129(8):081802, 2022.

[33] Armen Tumasyan et al. Search for Nonresonant Pair Production of Highly Energetic Higgs Bosons Decaying to Bottom Quarks. *Phys. Rev. Lett.*, 131(4):041803, 2023.

[34] Georges Aad et al. Search for the nonresonant production of Higgs boson pairs via gluon fusion and vector-boson fusion in the bb¯τ+τ- final state in proton-proton collisions at s=13 TeV with the ATLAS detector. *Phys. Rev. D*, 110(3):032012, 2024.

[35] Albert M Sirunyan et al. Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s}$ =13TeV. *Phys. Lett. B*, 778:101–127, 2018.

[36] Georges Aad et al. Studies of new Higgs boson interactions through nonresonant HH production in the $b\bar{b}\gamma\gamma$ final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 01:066, 2024.

[37] Albert M Sirunyan et al. Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *JHEP*, 03:257, 2021.

[38] Georges Aad et al. Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 801:135145, 2020.

[39] Albert M Sirunyan et al. Search for resonant and nonresonant Higgs boson pair production in the $b\bar{b}\ell\nu\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 01:054, 2018.

[40] Georges Aad et al. Search for non-resonant Higgs boson pair production in final states with leptons, taus, and photons in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 08:164, 2024.

[41] Armen Tumasyan et al. Search for Higgs boson pairs decaying to WW*WW*, WW*$\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *JHEP*, 07:095, 2023.

[42] Morad Aaboud et al. Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 04:092, 2019.

[43] Morad Aaboud et al. Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 78(12):1007, 2018.

[44] Morad Aaboud et al. Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV. *JHEP*, 05:124, 2019.

[45] Georges Aad et al. Search for non-resonant Higgs boson pair production in the $2b + 2\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. *JHEP*, 02:037, 2024.

[46] Georges Aad et al. Search for pair production of boosted Higgs bosons via vector-boson fusion in the bb^{-}bb^{-} final state using pp collisions at s=13TeV with the ATLAS detector. *Phys. Lett. B*, 858:139007, 2024.

[47] Georges Aad et al. Combination of Searches for Higgs Boson Pair Production in pp Collisions at s=13 TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 133(10):101801, 2024.

[48] Armen Tumasyan et al. Search for nonresonant Higgs boson pair production in the four leptons plus twob jets final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *JHEP*, 06:130, 2023.

[49] Armen Tumasyan et al. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature*, 607(7917):60–68, 2022. [Erratum: Nature 623, (2023)].

[50] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968.

[51] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[52] Paul A. M. Dirac. The quantum theory of the electron. *Proc. Roy. Soc. Lond. A*, 117:610–624, 1928.

[53] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[54] Glashow-weinberg-salam theory of electroweak interactions and the neutral currents. *Physics Reports*, 90(2):73–157, 1982.

[55] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[56] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996.

[57] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.

[58] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967.

[59] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20:367–387, Apr 1948.

[60] Murray Gell-Mann. Symmetries of baryons and mesons. *Phys. Rev.*, 125:1067–1084, Feb 1962.

[61] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, 1974.

[62] Ilaria Brivio and Michael Trott. The standard model as an effective field theory. *Physics Reports*, 793:1–98, February 2019.

[63] W. Buchmuller and D. Wyler. Effective Lagrangian Analysis of New Interactions and Flavor Conservation. *Nucl. Phys. B*, 268:621–653, 1986.

[64] Chung Ngoc Leung, S. T. Love, and S. Rao. Low-Energy Manifestations of a New Interaction Scale: Operator Analysis. *Z. Phys. C*, 31:433, 1986.

[65] R. Alonso, M.B. Gavela, L. Merlo, S. Rigolin, and J. Yepes. The effective chiral lagrangian for a light dynamical "higgs particle". *Physics Letters B*, 722(4–5):330–335, May 2013.

[66] Gerhard Buchalla, Oscar Catà, and Claudius Krause. Complete electroweak chiral lagrangian with a light higgs at nlo. *Nuclear Physics B*, 880:552–573, March 2014.

[67] Gerhard Buchalla, Oscar Catà, and Claudius Krause. On the power counting in effective field theories. *Physics Letters B*, 731:80–86, April 2014.

[68] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[69] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[70] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[71] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[72] Georges Aad et al. Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 73(8):2518, 2013.

[73] Morad Aaboud et al. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 76(12):653, 2016.

[74] Georges Aad et al. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 83(10):982, 2023.

[75] ATLAS Collaboration. The inner detector. `https://atlas.cern/Discover/Detector/Inner-Detector`.

[76] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010.

[77] B. Abbott et al. Production and Integration of the ATLAS Insertable B-Layer. *JINST*, 13(05):T05008, 2018.

[78] G. Aad et al. ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008.

[79] A. Ahmad et al. The Silicon microstrip sensors of the ATLAS semiconductor tracker. *Nucl. Instrum. Meth. A*, 578:98–118, 2007.

[80] E. Abat et al. The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance. *JINST*, 3:P02013, 2008.

[81] Basic ATLAS TRT performance studies of Run 1. Technical report, CERN, Geneva, 2014. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-INDET-PUB-2014-001.

[82] Morad Aaboud et al. Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties. *JINST*, 12(05):P05002, 2017.

[83] ATLAS Collaboration. Calorimeter. `https://atlas.cern/Discover/Detector/Calorimeter`.

[84] M. Aharrouche, J. Colas, L. Di Ciaccio, M. El Kacimi, O. Gaumer, M. Gouanère, D. Goujdami, R. Lafaye, S. Laplace, C. Le Maner, L. Neukermans, P. Perrodo, L. Poggioli, D. Prieur, H. Przysiezniak, G. Sauvage,

F. Tarrade, I. Wingerter-Seez, R. Zitoun, F. Lanni, H. Ma, S. Rajagopalan, S. Rescia, H. Takai, A. Belymam, D. Benchekroun, M. Hakimi, A. Hoummada, E. Barberio, Y.S. Gao, L. Lu, R. Stroynowski, M. Aleksa, J.B. Hansen, T. Carli, I. Efthymiopoulos, P. Fassnacht, F. Follin, F. Gianotti, L. Hervas, W. Lampl, J. Collot, J.Y. Hostachy, F. Ledroit-Guillon, P. Martin, F. Ohlsson-Malek, S. Saboumazrag, M. Leltchouk, J.A. Parsons, M. Seman, S. Simion, D. Banfi, L. Carminati, D. Cavalli, G. Costa, M. Delmastro, M. Fanti, L. Mandelli, M. Mazzanti, G.F. Tartarelli, C. Bourdarios, L. Fayard, D. Fournier, G. Graziani, S. Hassani, L. Iconomidou-Fayard, M. Kado, M. Lechowski, M. Lelas, G. Parrour, P. Puzo, D. Rousseau, R. Sacco, L. Serin, G. Unal, D. Zerwas, A. Camard, D. Lacour, B. Laforge, I. Nikolic-Audit, Ph. Schwemling, H. Ghazlane, R.C. El Moursli, A.I. Fakhr-Eddine, M. Boonekamp, N. Kerschen, B. Mansoulié, P. Meyer, J. Schwindling, B. Lund-Jensen, and Y. Tayalati. Energy linearity and resolution of the atlas electromagnetic barrel calorimeter in an electron test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):601–623, December 2006.

[85] Jalal Abdallah et al. Study of energy response and resolution of the ATLAS Tile Calorimeter to hadrons of energies from 16 to 30 GeV. *Eur. Phys. J. C*, 81(6):549, 2021.

[86] ATLAS Collaboration. Muon spectrometer. `https://atlas.cern/Discover/Detector/Muon-Spectrometer`.

[87] Georges Aad et al. Resolution of the ATLAS muon spectrometer monitored drift tubes in LHC Run 2. *JINST*, 14(09):P09011, 2019.

[88] E. Diehl. Calibration and Performance of the ATLAS Muon Spectrometer. In *Meeting of the APS Division of Particles and Fields*, 9 2011.

[89] Georges Aad et al. Performance of the ATLAS Trigger System in 2010. *Eur. Phys. J. C*, 72:1849, 2012.

[90] *ATLAS level-1 trigger: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1998.

[91] Morad Aaboud et al. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J. C*, 77(5):317, 2017.

[92] Georges Aad et al. Operation of the ATLAS trigger system in Run 2. *JINST*, 15(10):P10004, 2020.

[93] M. Abolins et al. The ATLAS Data Acquisition and High Level Trigger system. *JINST*, 11(06):P06008, 2016.

[94] Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, Geneva, 2019. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021.

[95] N. Berger et al. The ATLAS high level trigger steering. *J. Phys. Conf. Ser.*, 119:022013, 2008.

[96] Peter Jenni, Marzio Nessi, Markus Nordberg, and Kenway Smith. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 2003.

[97] Georges Aad et al. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014.

[98] Studies for the development of the Inner Detector trigger algorithms at ATLAS. Technical report, CERN, Geneva, 2013. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2013-002.

[99] M. Aaboud et al. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2. *Eur. Phys. J. C*, 77(10):673, 2017.

[100] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987.

[101] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger. The new ATLAS track reconstruction (NEWT). *J. Phys. Conf. Ser.*, 119:032014, 2008.

[102] Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV. Technical report, CERN, Geneva, 2010. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2010-072.

[103] *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. CERN, Geneva, 1997.

[104] S Haywood, L Rossi, R Nickerson, and A Romaniouk. *ATLAS inner detector: Technical Design Report, 2*. Technical design report. ATLAS. CERN, Geneva, 1997.

[105] T G Cornelissen, M Elsing, I Gavrilenko, J F Laporte, W Liebig, M Limper, K Nikolopoulos, A Poppleton, and A Salzburger. The global ^{2} track fitter in ATLAS. *J. Phys.: Conf. Ser.*, 119:032013, 2008.

[106] Pamela Ferrari. Tracking and vertexing at atlas, 2007.

[107] Morad Aaboud et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC. *Eur. Phys. J. C*, 77(5):332, 2017.

[108] Georges Aad et al. Configuration and performance of the ATLAS *b*-jet triggers in Run 2. *Eur. Phys. J. C*, 81(12):1087, 2021.

[109] Federico Meloni. Primary vertex reconstruction with the ATLAS detector. *JINST*, 11(12):C12060, 2016.

[110] Georges Aad et al. ATLAS data quality operations and performance for 2015–2018 data-taking. *JINST*, 15(04):P04003, 2020.

[111] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[112] Trigger Menu in 2017. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2018-002.

[113] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzitelli. Higgs boson pair production at nnlo with top quark mass effects. *Journal of High Energy Physics*, 2018(5), May 2018.

[114] Frédéric A. Dreyer and Alexander Karlberg. Vector-Boson Fusion Higgs Pair Production at N^3LO. *Phys. Rev. D*, 98(11):114016, 2018.

[115] Fady Bishara, Roberto Contino, and Juan Rojo. Higgs pair production in vector-boson fusion at the lhc and beyond. *The European Physical Journal C*, 77(7), July 2017.

[116] CERN. Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector, 2017.

[117] A. Djouadi, J. Kalinowski, and M. Spira. Hdecay: a program for higgs boson decays in the standard model and its supersymmetric extension. *Computer Physics Communications*, 108(1):56–74, January 1998.

[118] Alexandra Carvalho, Martino Dall'Osso, Tommaso Dorigo, Florian Goertz, Carlo A. Gottardo, and Mia Tosi. Higgs pair production: choosing benchmarks with cluster analysis. *Journal of High Energy Physics*, 2016(4):1–28, April 2016.

[119] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823–874, 2010.

[120] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[121] Paolo Nason. A new method for combining nlo qcd with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040–040, November 2004.

[122] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *Journal of High Energy Physics*, 2007(11):070–070, November 2007.

[123] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, Ronan McNulty, Andreas Morsch, Pavel Nadolsky, Voica Radescu, Juan Rojo, and Robert Thorne. Pdf4lhc recommendations for lhc run ii. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, January 2016.

[124] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, June 2015.

[125] ATLAS Pythia 8 tunes to 7 TeV data. 11 2014.

[126] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions with lhc data. *Nuclear Physics B*, 867(2):244–289, February 2013.

[127] Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector. 2019.

[128] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-

level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), July 2014.

[129] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the lhc run ii. *Journal of High Energy Physics*, 2015(4), April 2015.

[130] Georges Aad et al. Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Lett. B*, 800:135103, 2020.

[131] John M. Campbell, R. Keith Ellis, Paolo Nason, and Emanuele Re. Top-pair production and decay at nlo matched with parton showers. *Journal of High Energy Physics*, 2015(4), April 2015.

[132] Vertex Reconstruction Performance of the ATLAS Detector at "$\sqrt{s}$ = 13 TeV". *Technical report, CERN, Geneva*, 2015. *All figures including auxiliary figures are availa*
//atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL − PHYS − PUB − 2015 − 026.

[133] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, April 2008.

[134] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. Fastjet user manual: (for version 3.0.2). *The European Physical Journal C*, 72(3), March 2012.

[135] Morad Aaboud et al. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77(7):466, 2017.

[136] Georges Aad et al. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C*, 76(11):581, 2016.

[137] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017.

[138] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical report, CERN, Geneva, 2015. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2015-029.

[139] Georges Aad et al. ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset. *Eur. Phys. J. C*, 83(7):681, 2023.

[140] Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-006.

[141] Georges Aad et al. Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector. *Eur. Phys. J. C*, 82(1):95, 2022.

[142] Georges Aad et al. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 81(7):578, 2021.

[143] M. Aaboud et al. Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector. *JHEP*, 12:024, 2017.

[144] Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 807–814, Madison, WI, USA, 2010. Omnipress.

[145] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles, 2017.

[146] B Efron. Bootstrap methods: another look at the jackknife annals of statistics 7: 1–26. *View Article PubMed/NCBI Google Scholar*, 24, 1979.

[147] Georges Aad et al. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021.

[148] M. Aaboud et al. Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.*, 117(18):182002, 2016.

[149] Georges Aad et al. 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(11):970, 2019.

[150] G Avoni, M Bruschi, G Cabras, D Caforio, N Dehghanian, A Floderus, B Giacobbe, F Giannuzzi, F Giorgi, P Grafström, V Hedberg, F Lasagni Manghi, S Meneghini, J Pinfold, E Richards, C Sbarra, N Semprini Cesari, A Sbrizzi, R Soluk, G Ucchielli, S Valentinetti, O Viazlo, M Villa, C Vittori, R Vuillermet, and A Zoccoli. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13(07):P07017, 2018.

[151] Wouter Verkerke and David Kirkby. The roofit toolkit for data modeling, 2003.

[152] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), February 2011.

[153] Alexander L. Read. Presentation of search results: The CL_s technique. *J. Phys. G*, 28:2693–2704, 2002.

[154] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[155] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc., 2017.

[156] Nicole Michelle Hartman. A Search for Non-Resonant $HH \to 4b$ at $\sqrt{s} = 13$ TeV with the ATLAS detector – or – $2b$, and then another $2b$... now that's the thesis question., 2022. Stanford U. PhD thesis presented 27 Oct 2022.

[157] Carl Edward Rasmussen. *Gaussian Processes in Machine Learning*, pages 63–71. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[158] Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows, 2016.

[159] Max Swiatlowski Thandi Madula, Nicole Hartman. Resolved background systematics, 2024. ATLAS hh4b meeting presentation.

[160] Rafael Teixeira de Lima Michael Kagan. Proposal on how to properly treat bootstraps uncertainty and their correlations. ATLAS hh4b meeting presentation.

[161] Trigger naming run2. https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TriggerNamingRun2.

[162] Trigger naming run3. https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TriggerNamingRun3.

[163] Waleed Abdallah et al. CEPC Technical Design Report: Accelerator. *Radiat. Detect. Technol. Methods*, 8(1):1–1105, 2024.

[164] David E Morrissey and Michael J Ramsey-Musolf. Electroweak baryogenesis. *New Journal of Physics*, 14(12):125003, December 2012.

[165] Chiara Caprini, Mark Hindmarsh, Stephan Huber, Thomas Konstandin, Jonathan Kozaczuk, Germano Nardini, Jose Miguel No, Antoine Petiteau,

Pedro Schwaller, Géraldine Servant, and David J. Weir. Science with the space-based interferometer elisa. ii: gravitational waves from cosmological phase transitions. *Journal of Cosmology and Astroparticle Physics*, 2016(04):001–001, April 2016.

[166] Michael J. Ramsey-Musolf. The electroweak phase transition: a collider target. *Journal of High Energy Physics*, 2020(9), September 2020.

[167] Marcela Carena, Zhen Liu, and Yikun Wang. Electroweak phase transition with spontaneous z2-breaking. *Journal of High Energy Physics*, 2020(8), August 2020.

[168] Stefano Profumo, Michael J Ramsey-Musolf, and Gabe Shaughnessy. Singlet higgs phenomenology and the electroweak phase transition. *Journal of High Energy Physics*, 2007(08):010–010, August 2007.

[169] Wolfgang Kilian, Thorsten Ohl, and Jürgen Reuter. Whizard—simulating multi-particle processes at lhc and ilc. *The European Physical Journal C*, 71(9), September 2011.

[170] Adam Alloul, Neil D. Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks. Feynrules 2.0 — a complete toolbox for tree-level phenomenology. *Computer Physics Communications*, 185(8):2250–2300, August 2014.

[171] P. Mora de Freitas and H. Videau. Detector simulation with MOKKA / GEANT4: Present and future. In *International Workshop on Linear Colliders (LCWS 2002)*, pages 623–627, 8 2002.

[172] Manqi Ruan, Hang Zhao, Gang Li, Chengdong Fu, Zhigang Wang, Xinchou Lou, Dan Yu, Vincent Boudry, Henri Videau, Vladislav Balagura, Jean-Claude Brient, Peizhu Lai, Chia-Ming Kuo, Bo Liu, Fenfen An, Chunhui Chen, Soeren Prell, Bo Li, and Imad Laketineh. Reconstruction of physics objects at the circular electron positron collider with arbor. *The European Physical Journal C*, 78(5), May 2018.

[173] Dan Yu, Manqi Ruan, Vincent Boudry, and Henri Videau. Lepton identification at particle flow oriented detector for the future e + e - higgs factories. *The European Physical Journal C*, 77(9), September 2017.

[174] Taikan Suehara and Tomohiko Tanabe. Lcfiplus: A framework for jet analysis in linear collider studies. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 808:109–116, February 2016.

[175] S. Catani, Yuri L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber. New clustering algorithm for multi - jet cross-sections in e+ e- annihilation. *Phys. Lett. B*, 269:432–438, 1991.

[176] Ivan Smiljanić, Ivanka Božović-Jelisavčić, Goran Kačarević, Nataša Vukašinović, Tatjana Agatonović-Jovin, Gordana Milutinović-Dumbelović, Jasna Stevanović, and Mirko Radulović. Integrated luminosity measurement at cepc, 2022.

[177] Yu Bai, Chunhui Chen, Yaquan Fang, Gang Li, Manqi Ruan, Jing-Yuan Shi, Bo Wang, Pan-Yu Kong, Bo-Yang Lan, and Zhan-Feng Liu. Measurements of decay branching fractions of $H \to b\bar{b}/c\bar{c}/gg$ in associated $(e^+e^-/\mu^+\mu^-)H$ production at the CEPC. *Chin. Phys. C*, 44(1):013001, 2020.

[178] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett. B*, 641:57–61, 2006.

[179] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(2–3):435–443, September 1999.