

Dissertation
Combined Faculty of the Natural Sciences and Mathematics
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Marta Maja Czuryło
Born in: Wrocław, Poland
Oral examination: February 14, 2023

Search for the non-resonant Vector Boson
Fusion production of the Higgs boson pairs
decaying to $b\bar{b}b\bar{b}$ final state using the
ATLAS detector

Referees: Prof. Dr. Andre Schöning
Prof. Dr. Monica Dunford

Kurzreferat

Unter Verwendung von 126 fb^{-1} an $\sqrt{s} = 13 \text{ TeV}$ Proton-Proton-Kollisionen, welche vom ATLAS-Detektor in Run 2 gesammelt wurden, wird nach nicht-resonanten Higgs-Boson-Paaren gesucht, die durch Vektorboson-Fusion (VBF) erzeugt wurden und die in den $b\bar{b}b\bar{b}$ Endzustand zerfallen. Die vom Standardmodell vorhergesagte Produktion von Higgs-Boson-Paaren ist bisher nicht experimentell belegt. Die beobachtete (erwartete) Ausschlussgrenze auf den Produktions-Wirkungsquerschnitt des Standardmodells entspricht dem 132,3 (132,8) -fachen der Vorhersage des Standardmodell-Wirkungsquerschnitts; mit einem Konfidenzniveau von 95%. Durch Optimierung der Analysestrategie wird die bisherige Grenze auf den Standardmodell $HH \rightarrow b\bar{b}b\bar{b}$ VBF-Wirkungsquerschnitt um einen Faktor sechs verbessert. Darüber hinaus werden nicht mit dem Standardmodell konforme Kopplungen untersucht. Dadurch wird κ_{2V} auf das beobachtete Intervall $[0,09, 1,99]$ eingegrenzt. Die vollständig datengesteuerte QCD-Untergrundabschätzung mittels neuronaler Netze hat einen großen Einfluss auf das verbesserte Ergebnis und wird in dieser Arbeit ausführlich diskutiert. Zusätzlich werden Projektionsstudien für die zukünftigen experimentellen Szenarien, HL-LHC und FCC-hh, durchgeführt. Hierfür wird ein vereinfachter Ansatz mittels Monte-Carlo-generierten Signal- und Untergrunddaten gewählt. Das Ergebnis der Studie ist eine erwartete Ausschlussgrenze auf den Standardmodell-Wirkungsquerschnitt, die im Falle des HL-LHC das 248-fache der Vorhersage des Standardmodells und für das FCC-Szenario das 18-fache des Standardmodells betragen; ebenfalls mit einem Konfidenzniveau von 95%.

Abstract

Non-resonant Higgs boson pairs produced via Vector Boson Fusion (VBF) and decaying to $b\bar{b}b\bar{b}$ final state are searched for using 126 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ proton-proton collisions data collected by the ATLAS detector in Run 2. There is no evidence for the Standard Model production of Higgs boson pairs. The observed (expected) 95% CL exclusion limit on the Standard Model production cross-section is set at 132.3 (132.8) times the Standard Model cross-section prediction. The analysis strategy is optimised so that the previous limit set on the Standard Model $HH \rightarrow b\bar{b}b\bar{b}$ VBF cross-section is improved by a factor of 6. Non Standard Model couplings are tested as well, leading to the observed κ_{2V} constrained interval of $[0.09, 1.99]$. The fully data-driven neural network QCD background estimation has a considerable impact on the improved result and is discussed in detail in this thesis. Additionally, projection studies for the future experimental scenarios: the HL-LHC and the FCC-hh, are performed using a simplified approach with Monte Carlo generated signal and background samples. The study results in a 95% CL expected limit on the Standard Model cross-section reaching 248 times the Standard Model cross-section prediction for for the HL-LHC and 18 times the Standard Model cross-section prediction for the FCC scenario.

Contents

1	Introduction	1
2	Theory	5
2.1	The Standard Model	5
2.1.1	Structure of the Standard Model	5
2.1.2	QCD	7
2.1.3	Electroweak Unification	8
2.1.4	The Higgs Mechanism	10
2.1.5	Phenomena Unexplained by the SM	13
2.2	Higgs Pair Production	14
2.2.1	Shape of Higgs Potential	14
2.2.2	Electroweak Baryogenesis	15
2.2.3	gluon-gluon Fusion	16
2.2.4	Vector Boson Fusion	18
2.2.5	$b\bar{b}b\bar{b}$ Decay Channel	19
2.2.6	Monte Carlo Modelling of Signal and Background	22
2.2.7	Di-Higgs Searches Status	23
3	LHC and ATLAS detector	25
3.1	The Large Hadron Collider	25
3.1.1	Luminosity	27
3.2	The ATLAS experiment	28
3.2.1	Coordinate System of the Detector	29
3.2.2	Inner Detector	29
3.2.3	Calorimeter	32
3.2.4	Muon Spectrometer	34
3.2.5	Triggering and Data Acquisition	34
4	Machine Learning	37
4.1	Supervised Learning	37
4.2	Neural Networks	38
5	Analysis Relevant Objects	45
5.1	Jets	45
5.1.1	The <i>anti</i> - k_t Algorithm	45
5.1.2	Jets in ATLAS	46

5.1.3	Jet Calibrations and Corrections	48
5.2	<i>b</i> -tagging	49
5.2.1	Algorithms in ATLAS	50
5.2.2	Working Points	50
5.2.3	Offline Flavour Tagging Calibrations	51
6	Event Selection	55
6.1	Data Samples	55
6.2	Monte Carlo Samples	56
6.2.1	κ_{2V} Reweighting	57
6.3	Triggers	58
6.3.1	<i>b</i> -jet Trigger Efficiency Correction	59
6.3.2	Trigger Bucket Algorithm	62
6.4	VBF Event Selection	64
6.4.1	Nomenclature of the Analysis Jets	64
6.4.2	Steps of the Event Selection	64
6.5	Higgs Candidate Pairing	65
6.5.1	Overview of Considered Algorithms	65
6.6	Top veto for the Background Reduction	67
6.7	Kinematic Regions Definitions	69
6.8	Cut-flows	70
6.9	Acceptance times Efficiency	71
6.10	Discriminating Variable	71
7	Background Estimation	75
7.1	Method of Event Reweighting with Bins	75
7.1.1	2b to 4b Event Reweighting	76
7.2	2b to 4b Neural Network Reweighting	77
7.3	Neural Network Reweighting in Practice	78
7.3.1	Network Input Variables	78
7.3.2	Network Architecture	83
7.4	Background Estimate Uncertainties	83
7.4.1	Statistical Uncertainty and Bootstrapping	83
7.4.2	Shape Systematic Uncertainty	85
7.5	Performance of the Reweighting in the Control Regions	89
7.6	Validation of the Background Estimate	96
7.6.1	3b1f Region Validation	96
7.6.2	Shifted Kinematic Regions Validation	98
7.7	Background Prediction in the 4b Signal Region	101
8	Statistical Analysis	105
8.1	Mathematical Methods	105
8.1.1	Hypotheses of the Experiment	105
8.1.2	Likelihood	106

8.1.3	Profile Likelihood Method	107
8.2	Implementation in the PYHF Framework	108
8.3	Uncertainties	109
8.3.1	Summary of Background Uncertainties	109
8.3.2	Summary of Signal Uncertainties	109
8.4	Results	111
8.4.1	Post-fit Background Distribution	113
8.4.2	95% CL Upper Limits	115
9	HL-LHC and FCC Projection Studies	119
9.1	Overview of the Considered Scenarios	119
9.2	Methods	120
9.2.1	MC Production Prescription	120
9.2.2	Calorimeter Emulation, Jet Clustering and b-tagging	121
9.2.3	Simplified VBF Analysis Selection	124
9.2.4	Cut-flows	127
9.2.5	Scaling Factors	131
9.3	Results	133
9.3.1	Signal Significance Studies	133
9.3.2	m_{HH} Distributions	137
9.3.3	95% CL Expected Upper Limit	139
9.3.4	Discussion of Results	140
10	Conclusion	143
	Bibliography	147

Chapter 1

Introduction

The missing piece of the Standard Model, a particle responsible for the masses of other fundamental particles and its own, was experimentally confirmed in 2012 with the discovery of the Higgs boson by the ATLAS [1] and the CMS [2] collaborations. Since then, the properties of the new particle have been extensively studied, and many have been measured. Nevertheless, the exact shape of the Higgs potential at high energies is not yet known, while it hides important insights, for example, on the matter of the universe stability. The Higgs potential is studied through Higgs boson pair production. The measurement of the Higgs couplings: the self-coupling (from the gluon-gluon di-Higgs production) or the quartic coupling between two vector bosons and two Higgs bosons (from the Vector Boson Fusion di-Higgs production) is possible. The non-SM values of these couplings would hint at the existence of the new physics. This thesis explores the VBF di-Higgs production using Run 2 proton-proton collision data collected by the ATLAS detector with the total integrated luminosity of 126 fb^{-1} . The Higgs bosons decay to various final states. The final state with the highest branching fraction, namely $b\bar{b}b\bar{b}$, is studied here.

The analysis presented here is not the first ATLAS attempt to search for the VBF di-Higgs production. However, compared to the previous results [3], the methods are greatly improved, both in the reconstruction phase, the new and better performing Particle Flow jet objects and the DL1r flavour tagging algorithm are used, as well as in the analysis techniques. A new Higgs pairing algorithm is proposed, and the fully data-driven neural network background estimation is introduced. The background estimation is based on the assumption of the kinematic similarity of the 2 b-tagged events to the 4 b-tagged events and the possibility of removing any kinematic differences by neural network reweighting in the control regions. The derived weights are applied to the signal region 2-tagged events constituting the predicted 4b background. Additionally, for the amelioration of the overall signal significance, event categorization in $|\Delta\eta_{HH}|$ variable is applied to the signal region events.

Given the limited amount of statistics that are available in ATLAS and the lack of possibility to precisely test Higgs self and quartic couplings, the future experimental scenarios, the HL-LHC [4], and the FCC-hh [5], are considered in the complementary

projection studies. The Standard Model VBF Higgs pair production is studied using the Monte Carlo generated signal and background samples. The study is simplified, but it stretches beyond the generator-level study by introducing the calorimeter emulation.

The document's structure is the following: the theory of the Standard Model and Higgs pair production are explained in Chapter 2. Chapter 3 describes the basics of the LHC and the ATLAS detector operation. Machine Learning concepts, which are applied in the analysis, especially the neural networks, are discussed in Chapter 4. Chapter 5 introduces the reconstruction of the b-tagged jets in ATLAS. The newest b-tagging algorithm is introduced, including an evaluation of the flavour tagging scale factors. Chapter 6 presents the datasets used in the analysis and continues with the analysis strategy and optimizations: triggering and trigger calibrations, selection of the Higgs Candidates and VBF events, the Higgs Candidates pairing, definitions of the top quark veto variable, and the analysis signal and control regions. The data-driven neural network background estimation is derived in Chapter 7. The motivation for this sophisticated method, as well as its practical implementation in the analysis, is presented. Background estimation performance, together with the derivation of the statistical and systematic uncertainties, and the background estimate validation, are described in detail. The statistical analysis, the signal modelling uncertainties and the results are presented in Chapter 8. Chapter 9 is dedicated to the projection studies for future experiments. It presents the methodology, including the sample generation, the calorimeter emulation and the event selection. The chapter continues with the results from the signal significance studies and presents the expected 95% CL upper limits on the SM di-Higgs cross-section for the HL-LHC and the FCC-hh scenarios. The conclusion is presented in Chapter 10.

Author's Contribution

In the $HH \rightarrow b\bar{b}b\bar{b}$ ATLAS analysis team, I was one of the main contributors in the VBF production mode studies. My main responsibility was the implementation and the validation of the VBF background estimation. I implemented neural network reweighting to the VBF channel, selected datasets for training and validation, and studied the optimal set of input variables and the network architecture. I integrated the procedure with the ggF part of the analysis and studied different methods of evaluating background shape systematics for the VBF. I evaluated the background estimate performance in the control regions and validated the VBF background estimate using the 3b1f method with downsampling. I also contributed to the VBF channel selection steps optimization and the Higgs Candidates pairing selection. I documented the VBF background estimate sections in the Internal Note [6]. The analysis Conference Proceeding for the combined ggF and VBF non-resonant search is available in [7].

I also performed the projection studies outside of the ATLAS analysis team. I generated signal and background samples, implemented calorimeter emulation and b-tagging, and performed the VBF analysis.

As part of my PhD, I spent one year at CERN designing and implementing the Central Online Monitoring Application for the FastTrack ATLAS detector subsystem, which constituted my ATLAS qualification task. The online monitoring allowed for the direct communication of the FTK hardware components, replacing an older, less efficient tool. The publication summarizing the FTK system, including the Online Monitoring, can be found in [8].

Chapter 2

Theory

The chapter introduces the theoretical concepts explored experimentally in this thesis. The Standard Model (SM) is introduced first, followed by the Higgs pair production process both in the context of verifying the SM predictions and new physics searches. The challenges related to the estimation of the underlying background are also discussed.

2.1 The Standard Model

The Standard Model is one of the two greatest existing theories that describe the fundamental building blocks of nature¹. In the following, the constituent pieces of the SM are described, with emphasis on the Higgs Mechanism. This chapter closely follows [9].

2.1.1 Structure of the Standard Model

Figure 2.1 lists the constituents of the Standard Model. The SM consists of two groups of fundamental particles: *fermions*, which follow the Fermi-Dirac statistics and therefore have the half-integer spin, and *bosons*, which obey the Bose-Einstein statistics instead, hence have an integer spin 0 or 1.

The first group are the particles (and anti-particles) that make up matter; quarks - stable matter, and leptons - unstable matter, in three generations of increasing mass. Additionally, each lepton has an accompanying neutrino partner. In the second group, bosons with spin-1 are responsible for carrying the fundamental forces that govern the interactions between particles. W and Z bosons carry the weak force, photon, γ , carries the electromagnetic force while gluon carries the strong force. Some of these particles are massive thanks to the spin-0 Higgs boson. Note that the last fundamental force, gravity, is not included in the SM. Among a few theoretical

¹The second being the General Relativity and its description of gravity.

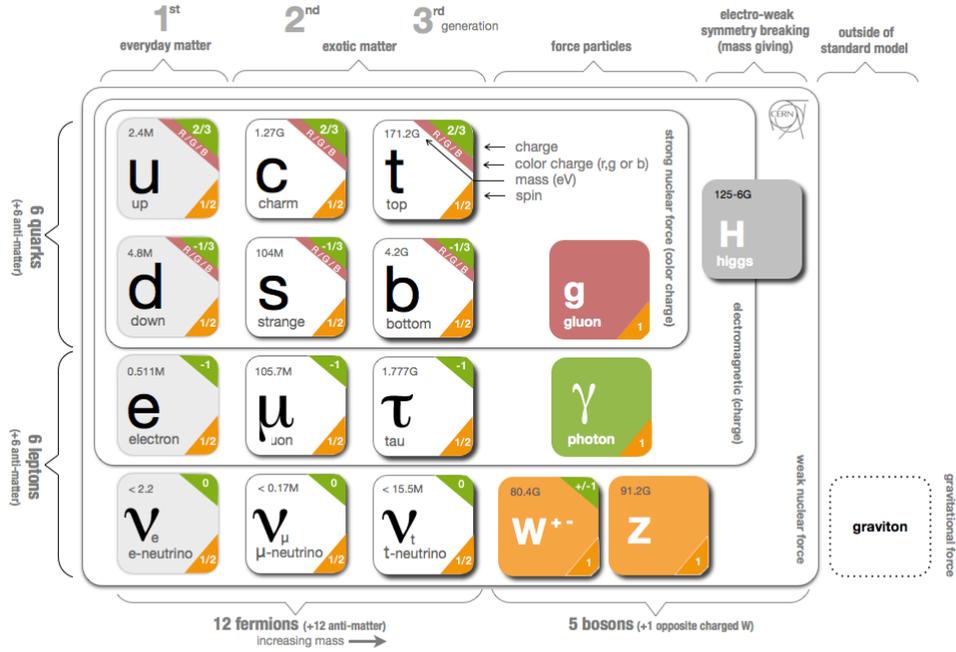


Figure 2.1: Particles constituting the Standard Model. Both fermions (quarks and leptons), as well as bosons, are shown. The description follows in the text. Figure taken from [10].

reasons, gravity’s very different operational scale makes it impossible to directly include it in this framework.

Formally, the SM belongs to the quantum gauge field theories. It means that particles are excitations of the quantum field and are defined by sets of quantum numbers, while the field itself is described by a Lagrangian density, \mathcal{L} . The field must be invariant under the local transformations of the constituent gauge groups: $SU(3) \times SU(2) \times U(1)$. The invariance condition is required for the theory renormalizability which is needed to make meaningful predictions at higher energies. The constituent $SU(3)$ is responsible for the strong force binding the atomic nuclei together, more details of QCD are discussed in Section 2.1.2. The $SU(2) \times U(1)$ is responsible for the unified electroweak force, with more details of the unification mechanism discussed in Section 2.1.3. To preserve electroweak invariance, all gauge bosons should be massless. However, W and Z are not massless. They have to be massive as weak force is a short-ranged interaction. In 1964, the publications by Brout and Englert [11] and, separately, Higgs [12], explained how this could be possible. They introduced the mechanism of spontaneous electroweak symmetry breaking that results in W and Z bosons acquiring mass. This is the famous Higgs mechanism discussed in Section 2.1.4.

2.1.2 QCD

The strong interactions are described by the Quantum Chromodynamics (QCD) gauge theory. In experiments, quarks never appear alone, but only as *hadrons*, the bound states of quarks. These can either have two constituents: quark and anti-quark, *mesons*, or three: combined quarks and antiquarks, *baryons*. In order to preserve the Pauli exclusion principle in the baryonic case, the colour charge was introduced. Each quark has one of the RGB (red, green, blue) colours, or anti-colours, while the observable states must be colourless. This implies the lack of existence of free quarks: the *quark confinement*. However, at very high energies quarks may exist on their own, leading to the phenomenon of *asymptotic freedom*.

The theory is associated with the $SU(3)_C$ symmetry group where the conserved quantity under transformations is the colour charge C . The compact form of the QCD Lagrangian reads:

$$\mathcal{L} = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} + \sum_f \bar{q}_f (i\gamma_\mu D^\mu - m_f) q_f, \quad (2.1)$$

where G_a^μ is the gluon field with $a = 1, \dots, 8$ as there are 8 possible gluons, $q_f = q_f^T = (q_f^1, q_f^2, q_f^3)$ is the quark triplet, with one of the six flavours (u, d, c, s, b, t) and each having one of the three colours. m_f is mass of the quark and D^μ is the covariant derivative:

$$D^\mu = \partial^\mu - ig_s \frac{\lambda_a}{2} G_a^\mu(x) = \partial^\mu - ig_s \frac{\lambda_a}{2} G_a^\mu(x). \quad (2.2)$$

g_s is the strong coupling, λ_a are generators associated with the symmetry group: the Gell-Mann 3×3 matrices.

The unitary transformation of the group is defined as:

$$U = \exp\left(ig_s \frac{\lambda_a}{2} \theta^a\right) \quad (2.3)$$

where θ^a is a parameter. When U is applied to quarks it results in:

$$q_f \rightarrow (q_f)' = U q_f, \quad (2.4)$$

to gluons:

$$G_a^\mu(x) \rightarrow (G_a^\mu(x))' = U G_a^\mu(x) U^\dagger + \frac{i}{g_s} (\partial^\mu U) U^\dagger, \quad (2.5)$$

and D_μ transforms as:

$$(D_\mu)' = U D_\mu U^\dagger. \quad (2.6)$$

Evaluating Lagrangian with definition of D^μ , we get:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} + \sum_f (\bar{q}_f i\gamma_\mu \partial^\mu q_f + g_s \bar{q}_f \gamma_\mu G^\mu q_f - m_f \bar{q}_f q_f), \quad (2.7)$$

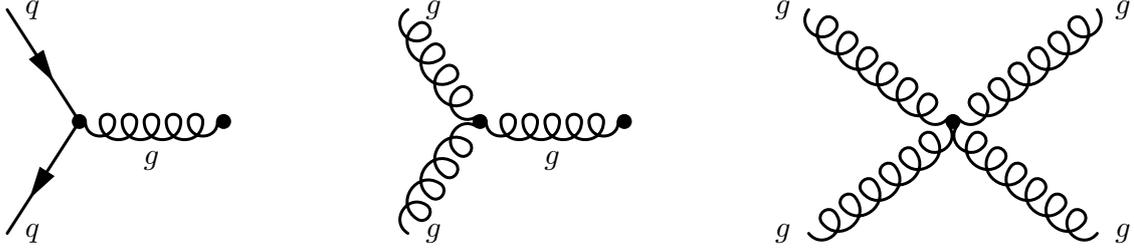


Figure 2.2: QCD vertices.

with gluon fields $G_{\mu\nu}^a$ expanding as:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - f^{abc} G_\mu^b G_\nu^c, \quad (2.8)$$

where f^{abc} are the symmetry group structure constants. When Eq. 2.8 is plugged into Eq. 2.7, the resulting Lagrangian gives the three vertices of QCD, shown in Figure 2.2. These are the gluon-quark-quark (quarks are of the same flavour) interaction and the trilinear and quartic gluon self-interactions.

2.1.3 Electroweak Unification

The observed charged currents are only possible with W bosons carrying the electric charge, meaning W 's must also interact electromagnetically with photons. Furthermore, the cross-section of the $e^-e^+ \rightarrow W^+W^-$ process was theoretically predicted as divergent. The Glashow-Weinberg-Salam unification mechanism introduced the fourth boson, the Z boson, in order to make this cross-section finite.

The basis of this mechanism is that under transformations of $SU(2)$, only the left-handed isospin, I , is conserved. It acts on the left-handed doublets of quark and lepton fields, which for the first generation of both are:

$$\begin{aligned} \psi_1 = l_L &= \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \\ \psi_1 = q_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix}. \end{aligned} \quad (2.9)$$

The right-handed fields exist but are singlets, for quarks:

$$\begin{aligned} \psi_2 = q_{R,1} &= u_R, \\ \psi_3 = q_{R,2} &= d_R. \end{aligned} \quad (2.10)$$

The symmetry group associated with the electromagnetic interaction, $U(1)$, conserves the weak hypercharge quantity under the transformation. The hypercharge, Y , is defined as:

$$Y = 2(Q - I_3), \quad (2.11)$$

where Q is the electric charge and I_3 is the third component of the isospin I .

Using the quark doublet and two quark singlets², the Lagrangian is:

$$\mathcal{L} = -\frac{1}{4}W_{\mu\nu}^k W_k^{\mu\nu} + \sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu D_\mu\psi_j(x), \quad (2.12)$$

where D_μ is the covariant derivative defined as:

$$D_\mu = \partial_\mu + ig\frac{\sigma_i}{2}W_\mu^i + ig'\frac{Y}{2}B_\mu, \quad (2.13)$$

in which σ_i are the three Pauli spin matrices, g and g' are the coupling constants. The four gauge bosons are introduced: three W_μ^i associated with the weak isospin I and one B_μ associated with hypercharge Y .

The unitary transformation that leaves the Lagrangian invariant acts on the left-handed doublet only. It reads:

$$U_L = \exp\left(i\frac{\sigma_i}{2}\alpha^i\right). \quad (2.14)$$

The resulting global transformations of the fields become:

$$\begin{aligned} \psi_1(x) &\rightarrow (\psi_1(x))' = \exp(iy_1\beta) U_L\psi_1(x), \\ \psi_2(x) &\rightarrow (\psi_2(x))' = \exp(iy_2\beta) \psi_2(x), \\ \psi_3(x) &\rightarrow (\psi_3(x))' = \exp(iy_3\beta) \psi_3(x). \end{aligned} \quad (2.15)$$

Finally, the physical, observed gauge bosons W_μ^+ , W_μ^- , Z_μ and the photon field A_μ are defined as:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2) \quad (2.16)$$

and

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \quad (2.17)$$

where θ_W is the weak mixing angle. The introduced gauge bosons solve the $e^-e^+ \rightarrow W^+W^-$ cross-section issue, with the Z^0 present, the cross-section is finite for a range of centre of mass energies.

However, the unification theory introduces all four bosons as massless, while we know it is not the case. The Higgs mechanism is needed to explain how W and Z bosons acquire mass.

²Same could be done with lepton doublet and singlets.

2.1.4 The Higgs Mechanism

In order for W and Z bosons to be massive, it is required that the local symmetry is spontaneously broken. For this purpose, the additional massive scalar boson, the Higgs boson, is introduced.

The local symmetry to be broken is the $SU(2)_L \times U(1)_Y$ symmetry. We first write the SM electroweak Lagrangian for the gauge bosons in terms of bosons W^i and B , which are assumed massless:

$$\mathcal{L} = -\frac{1}{4}W^{i\mu\nu}W_{\mu\nu}^i - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}. \quad (2.18)$$

To give mass to those bosons, an introduction of the two complex scalar fields is needed; a doublet invariant under $SU(2)_L$ reads:

$$\Phi(x) = \begin{pmatrix} \phi_+(x) \\ \phi_0(x) \end{pmatrix}, \quad (2.19)$$

where $\phi_+(x)$ represents a positively charged particle and $\phi_0(x)$ represents a neutral particle.

The potential of this doublet, the *Higgs potential* is given as:

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \quad (2.20)$$

where μ^2 and λ are, for now, just constants.

The potential is then incorporated into the electroweak Lagrangian from Eq. 2.18 resulting in:

$$\mathcal{L}_H = (D_\mu\Phi)^\dagger D_\mu\Phi - V(\Phi), \quad (2.21)$$

where D^μ is a covariant derivative as:

$$D_\mu = \left(\partial^\mu - ig\frac{\sigma_i}{2}W_i^\mu - ig'\frac{Y}{2}B^\mu \right), \quad (2.22)$$

with g and g' being the electroweak coupling constants and σ_i the 2×2 Pauli spin matrices.

Let's now look at the illustration of the Higgs potential in Figure 2.3. In both cases λ is positive, but μ^2 is either positive (on the left), or negative (on the right). In the negative case, the potential has an infinite number of non-zero minima which are given by:

$$|\Phi|^2 = \Phi^\dagger\Phi = \sqrt{-\frac{\mu^2}{2\lambda}}. \quad (2.23)$$

When the field is at the position of $|\Phi|^2 = 0$, everything is symmetric. The core of the spontaneous symmetry breaking is the transition of the field from $|\Phi|^2 = 0$ to

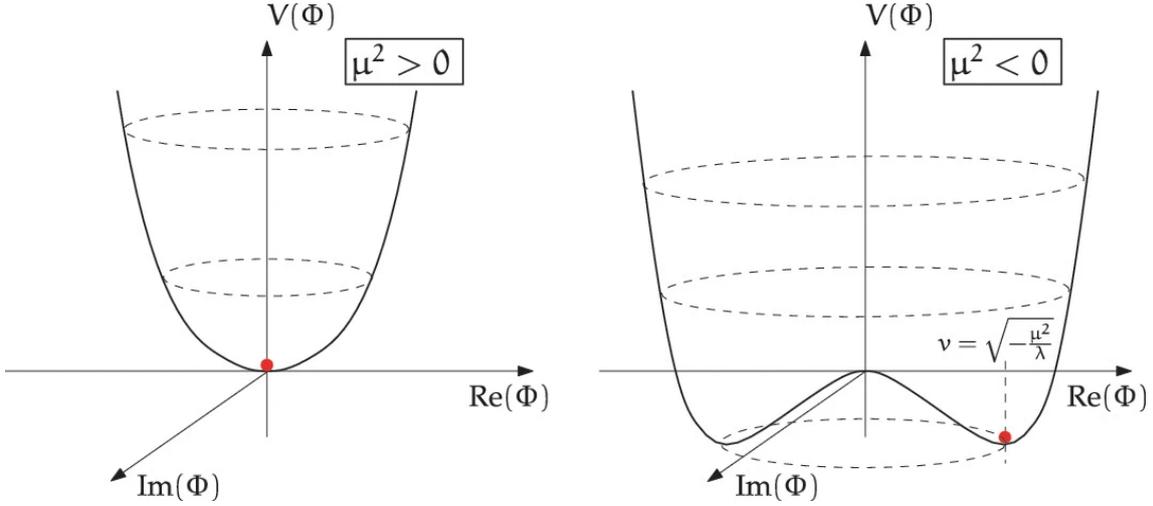


Figure 2.3: Schematic of the Higgs potential for positive λ and two cases of μ^2 . Notice the red dot, on the left, the field is at the equilibrium position, while on the right, at the minimum; the field is at the vacuum expectation value and the electroweak symmetry is broken. Figure taken from [13].

one of the minima states at $|\Phi|^2 = \sqrt{-\frac{\mu^2}{2\lambda}}$. As a result of this transition, some of the particles become massive. The reached minimum is called the *vacuum expectation value*, v or *VEV*.

In order to show that W and Z are massive, while the photon remains massless, we choose the particular field Φ , written as:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H(x) + v \end{pmatrix}, \quad (2.24)$$

where $H(x)$ is the scalar field associated with the Higgs boson and $v = \sqrt{-\frac{\mu^2}{\lambda}}$.

We now plug Eq. 2.24 into the kinetic part of the Higgs Lagrangian (Eq. 2.21):

$$(D_\mu \Phi)^\dagger D_\mu \Phi = \frac{1}{2} \partial^\mu H \partial_\mu H + (v+H)^2 \left(\frac{g^2}{8} (W_\mu^{(1)} + iW_\mu^{(2)})(W^{(1)\mu} - iW^{(2)\mu}) + \frac{1}{8} (gW_\mu^{(3)} - g'B_\mu)(gW^{(3)\mu} + g'B^\mu) \right). \quad (2.25)$$

The W boson fields, $W^{(1)}$ and $W^{(2)}$ are included in the equation, while the Z boson field, together with the massless photon field, are mixtures of the $W^{(3)}$ and B fields,

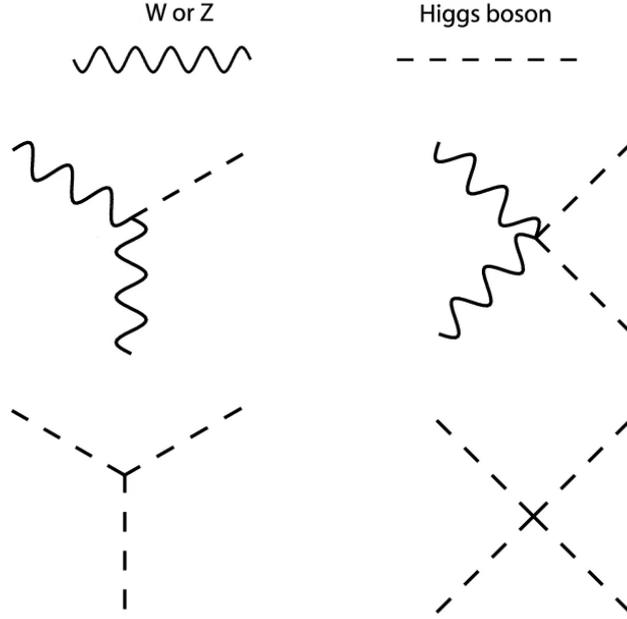


Figure 2.4: Bosonic interactions of the Higgs boson.

see Eq. 2.17. Now, skipping the photon field and using the potential term with the field Φ defined in Eq. 2.24, we finally arrive at the full Lagrangian as:

$$\mathcal{L}_\Phi = \frac{1}{2} \partial^\mu H \partial_\mu H + (v + H)^2 \left(\frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2(\theta_W)} Z_\mu Z^\mu \right) - \mu^2 H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4. \quad (2.26)$$

As can be inferred from the equation, a number of interactions are possible between W , Z and H bosons, namely, the trilinear and quartic Higgs self-interactions, and the VVH and VVHH interactions between the gauge bosons and Higgs. Figure 2.4 illustrates these four Higgs vertices. After recognizing the mass terms in the Lagrangian, we finally obtain:

$$M_W = \frac{vg}{2}, \quad (2.27)$$

$$M_Z = \frac{vg}{2 \cos \theta_W}, \quad (2.28)$$

$$M_H = \sqrt{-2\mu^2} = \sqrt{2\lambda v}, \quad (2.29)$$

substituting μ^2 from the definition of v .

All these values, $M_W = 80.379 \pm 0.012$ GeV, $M_Z = 91.1876 \pm 0.0021$ GeV and $M_H = 125.10 \pm 0.14$ GeV, [14], have been measured in experiments. It is now possible to explore the value of the constant λ .

The last three terms of the Lagrangian in Eq. 2.26 are the Higgs potential. The term proportional to H^2 is the Higgs mass term, H^3 is responsible for the Higgs

trilinear self-interaction and H^4 is the quartic self-interaction. For the Standard Model, we end up with the following equality:

$$\begin{aligned} V(H) &= -\lambda v^2 2H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4 \\ &= -\frac{1}{2} m_H^2 H^2 - \lambda_{HHH}^{SM} v H^3 - \lambda_{HHHH}^{SM} H^4. \end{aligned} \quad (2.30)$$

Therefore, $\lambda = \lambda_{HHH}^{SM}$ is the Higgs trilinear self-coupling. Using the H^2 term and the value of $v = 246$ GeV in the SM, we get:

$$\lambda_{HHH}^{SM} = \frac{m_H^2}{2v^2} \approx 0.13. \quad (2.31)$$

This value has not been measured yet, as this can only be done through the processes with the multi-Higgs final state; starting from the Higgs pair production, which is already an extremely rare process. Assuming the SM at $\sqrt{s} = 13$ TeV, di-Higgs production is 1500 times rarer than the single Higgs production. Taking into account detector inefficiencies, only around 100 Higgs pairs are expected in the full Run 2 data, yet the search for such pairs is the topic of this thesis. If the measured Higgs self-coupling value is different from the calculated SM value, it will indicate the existence of physics Beyond the Standard Model (BSM).

2.1.5 Phenomena Unexplained by the SM

Despite the Standard Model's successes in predicting various processes, there are observed phenomena that are not included in the model, a list of chosen ones is presented here. Firstly, gravity, the fourth fundamental force, is not included in the SM. Secondly, also from the cosmological observations, it is known that only 5% of the universe is visible matter, while *dark matter* and *dark energy* constitute the rest. Including the Dark Matter particle in the SM has been attempted for some time. No signs of any dark matter particles have been seen, while many models were tested, for example, the Weakly Interacting Massive Particle (WIMP) models [15]. Some of the decay rates of B^0 - mesons [16] have hinted at the new physics where the lepton flavour universality would not be required. Lastly, the prevalence of matter over anti-matter and the question of how this is possible is also not considered in the SM.

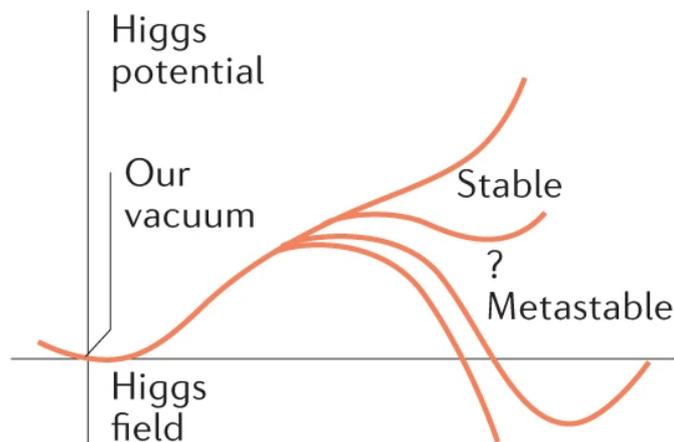


Figure 2.5: Possible scenarios for the shape of the Higgs potential at higher energies, while it is the known Higgs potential at the lower energies. The first possibility is that the potential keeps increasing, making our vacuum the global minimum and leaving the universe in a stable state. The second possibility is that, at higher energies, the potential reaches another minimum, making our vacuum only a local minimum. In that case, the universe would be metastable [18].

2.2 Higgs Pair Production

In this section the motivation to search for Higgs boson pairs is given first, followed by the details of the pair production mechanism in the two main production modes: the gluon-gluon fusion and the vector boson fusion, and the decay channel of interest, $b\bar{b}b\bar{b}$.

2.2.1 Shape of Higgs Potential

The Higgs potential plays a crucial role in the Higgs mechanism. However, what we really know about this potential is its shape at low energies, near the minimum, but not at high energies. Figure 2.5 illustrates different full potential shapes that originate from the same potential shape at low energies. Depending on the full shape of the potential, our VEV could be a global or local minimum, which leads to the question of vacuum stability. Additionally, exploring the Higgs boson self-coupling can lead to understanding electroweak baryogenesis [17].

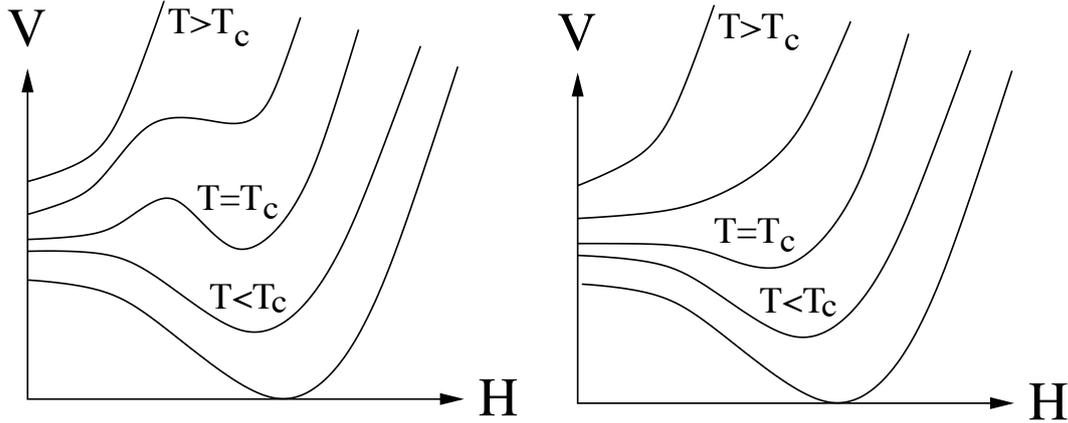


Figure 2.6: Illustration of the difference in the temperature-dependent phase transitions: the non-SM first-order phase transition that can lead to *bubble nucleation* on the left and the cross-over transition of the second-order, the SM-like, on the right. Figure taken from [17].

2.2.2 Electroweak Baryogenesis

The three Sakharov conditions point out what is necessary in order to explain the matter-antimatter asymmetry. These are; the baryon number violation, C and CP violation, and the CPT violation [5]. Even though the CP violation is present in the SM CKM matrix, it is not strong enough to fulfil the second Sakharov condition, while the CPT violation is not included in the SM at all. The electroweak baryogenesis model solves this problem by requiring the electroweak symmetry breaking to proceed via the strong, first-order phase transition instead of the smooth, cross-over second-order transition, which is the case for the SM Higgs mechanism. Such a transition could happen if a second singlet scalar Higgs field is coupled with the SM Higgs field [19]. The measured Higgs self-coupling different from the SM value could be a sign of the Higgs composite.

Phase Transitions

Figure 2.6 shows the two phase transitions. The crucial point is at the critical temperature T_c . For the cross-over, second-order transition (right), this temperature is defined as the temperature at which the second derivative of the potential is zero; the sign of the Higgs potential curvature changes smoothly. In the case of the first-order transition (left), T_c is defined at the point where the two minima are degenerate. There is a bump in between the symmetric and broken phases, which is overcome through the bubble nucleation transition [17].

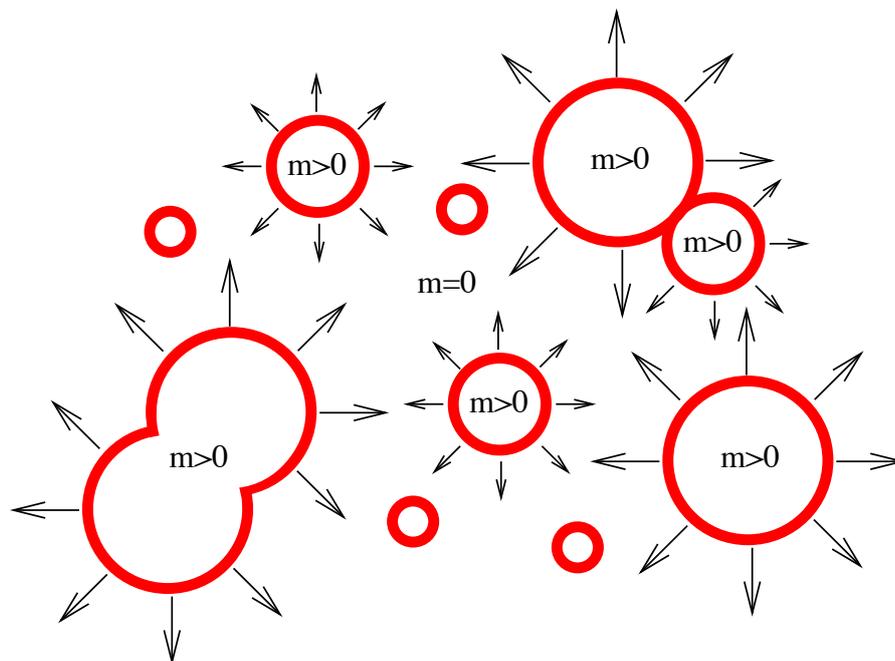


Figure 2.7: Illustration of the bubble nucleation for the electroweak baryogenesis. Inside the bubbles, VEV is zero, only while the bubbles expand, the antimatter is lost at the bubble walls. Figure taken from [17].

Bubble Nucleation

The schematic of the bubble nucleation is shown in Figure 2.7. Inside the bubbles the VEV of Higgs is non-zero, and so are the masses of the particles, the symmetry is broken. Outside of the bubbles VEV is zero, the electroweak symmetry is not broken and all particles are massless. At the point of thermal non-equilibrium, the bubbles expand, eventually collide and occupy the whole universe, hence the symmetry is broken everywhere. The baryon asymmetry arises at the bubble walls. Inside the bubbles, the baryon number is violated, while outside, the baryon-antibaryon equilibrium holds. The transition happens through the walls and it is not smooth, leading to the antimatter being lost at the walls. The sphalerons explain how the baryon asymmetry propagates to leptons, thus creating the matter-antimatter asymmetry [17].

2.2.3 gluon-gluon Fusion

Although this thesis is focused on the vector boson fusion production of Higgs pairs, for completeness the gluon-gluon fusion production is discussed first.

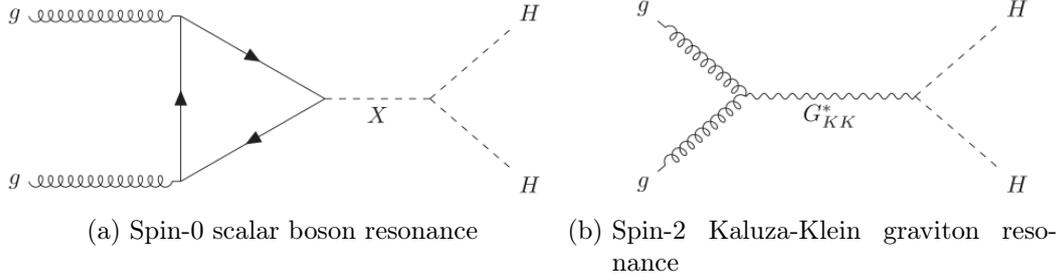


Figure 2.8: Feynman diagrams illustrating the resonant ggF production of a Higgs pair.

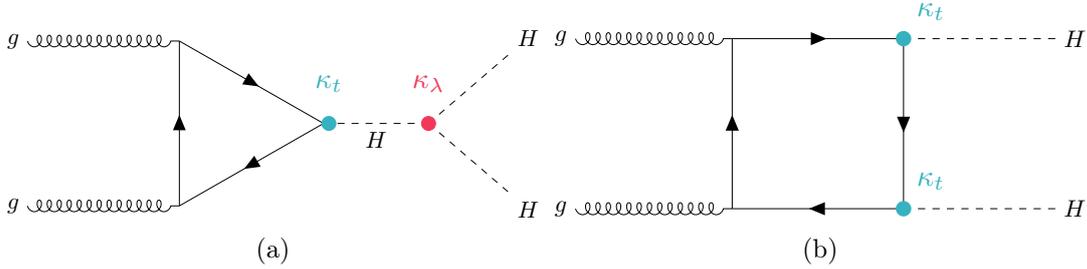


Figure 2.9: Feynman diagrams of the non-resonant gluon-gluon fusion production of two Higgs bosons in the final state. (a) shows the production via a triangular loop resulting in access to the Higgs self-coupling κ_λ , (b) shows the box diagram, where each Higgs couples directly to the top quark from the loop.

The gluon-gluon fusion (ggF) is the dominant production mode of Higgs pairs. The di-Higgs production proceeds either via a massive resonance or is a non-resonant production. The resonant scenario involves, for example, a spin-0 scalar boson X in the two Higgs doublet model [20] or spin-2 boson, for example, in the Kaluza-Klein theory [21]. Figure 2.8 shows the two resonant Feynman diagrams.

In this thesis, only the non-resonant channel is explored. Figure 2.9 shows the non-resonant ggF Feynman diagrams. Diagram in (a) is of the main interest as it allows direct access to the Higgs self-coupling, introduced in the SM as λ_{HHH}^{SM} in Section 2.1.4. Since its value is to be explored experimentally, for convenience, the parameter called the coupling strength modifier κ_λ , is defined:

$$\kappa_\lambda = \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}}, \quad (2.32)$$

for the SM $\kappa_\lambda = 1$. The same notation is applied to all other couplings; values of κ 's are always 1 for the SM. Formally, these are *coupling strength modifiers* but for simplicity, they will be referred to as *couplings*.

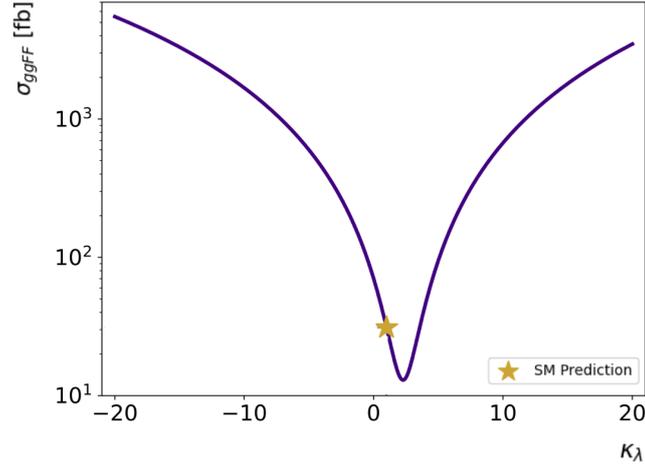


Figure 2.10: The theoretical cross-section values for different κ_λ values in gluon-gluon fusion. The prediction for the SM is indicated with a star. Cross-section values taken from [22].

The cross-section of the ggF production in the SM at $\sqrt{s} = 13$ TeV is 31.05 fb, evaluated at the next-to-next-to-leading order (NNLO) with the *FTApprox* top mass correction [22]. This low cross-section value comes from the destructive interference of the two diagrams shown in Figure 2.9. However, for the modified values of κ_λ , the cross-section increases, facilitating observation of Higgs pair production, see Figure 2.10.

2.2.4 Vector Boson Fusion

The second most dominant di-Higgs channel is the vector boson fusion production. Similarly to ggF, it can proceed via resonance or not. Figure 2.11 shows the resonant VBF production. The resonance is a scalar boson, for example, a heavy scalar from the 2HDM Type II model [23].

The non-resonant Feynman diagrams at the leading order are shown in Figure 2.12. A number of couplings are accessible through this channel; the self-coupling κ_λ , the coupling of Higgs to a single vector boson, κ_V , and the coupling of two vector bosons to two Higgs bosons, the quartic coupling κ_{2V} . The last coupling is the most interesting because it is only accessible via this particular process. Experimental measurement of $\kappa_{2V} \neq 1$ would be another indication of the BSM physics existence.

In the SM, the VBF production is around 20 times less likely than the ggF production, and the VBF cross-section at $\sqrt{s} = 13$ TeV, evaluated at N3LO, is 1.73 fb [24]. This is because the exact divergences are cancelled between the diagrams 2.12b and 2.12c in the SM due to the perturbative unitarity. Once the non-SM scenarios are

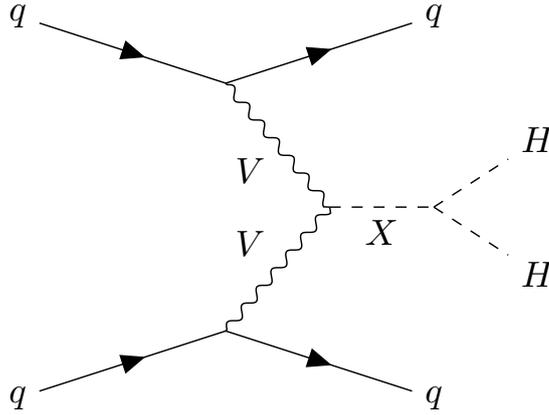


Figure 2.11: Feynman diagram illustrating the resonant VBF production of a Higgs pair.

considered, and κ_V and κ_{2V} are away from the SM value, the cross-section increases [25], illustrated in Figure 2.13.

2.2.5 $b\bar{b}b\bar{b}$ Decay Channel

Higgs boson can decay to many different final states, with the highest branching ratio of 58.2% for $H \rightarrow b\bar{b}$ in the SM. Figure 2.14 shows the SM branching ratios of all possible decays of the Higgs boson pairs. The final state of $b\bar{b}b\bar{b}$ is dominant, with the BR of 33%.

Although the $4b$ final state has the highest branching ratio, the signal process is accompanied by a difficult-to-estimate QCD background. Therefore, for the full exploration of the di-Higgs production, other decay channels are considered. To the most sensitive decay channels belong $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$. The two channels have much lower BRs, 7.3% and 0.3% respectively, but in the case of $b\bar{b}\tau^+\tau^-$, taus effectively help to reject the multi-jet QCD background, and in the case of $b\bar{b}\gamma\gamma$, great di-photon mass resolution provides remarkable signal purity.

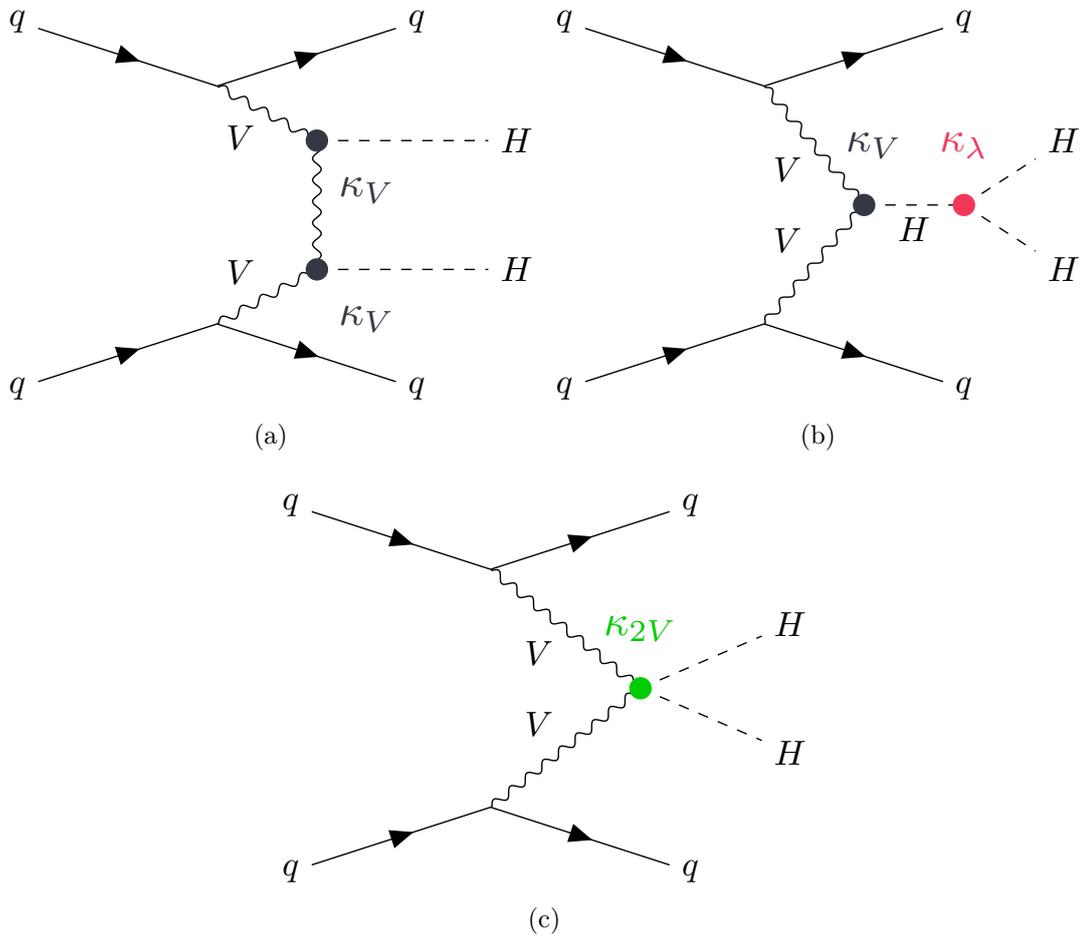


Figure 2.12: Feynman diagrams of the non-resonant vector boson fusion production of two Higgs bosons. (a) shows production in which each Higgs couples to one vector boson, (b) shows the production through Higgs self-interaction, while in (c) the quartic coupling κ_{2V} between two vector bosons and two Higgs bosons is accessed.

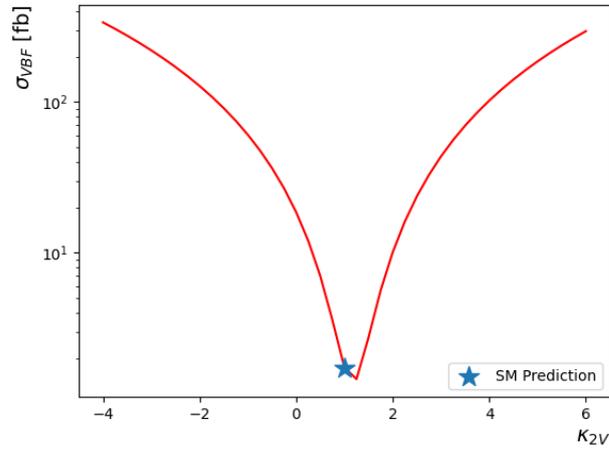


Figure 2.13: The theoretical cross-section values for different κ_{2V} values in vector boson fusion. The prediction for the SM is indicated with a star. Cross-section values taken from [3].

$b\bar{b}$	WW^*	gg	$\tau^+\tau^-$	$c\bar{c}$	ZZ^*	$\gamma\gamma$	
33.9%	24.9%	9.5%	7.3%	3.4%	3.1%	0.3%	$b\bar{b}$
	4.6%	3.5%	2.7%	1.2%	1.1%	< 0.1%	WW^*
		0.7%	1.0%	0.5%	0.4%	< 0.1%	gg
			0.4%	0.4%	0.3%	< 0.1%	$\tau^+\tau^-$
				< 0.1%	0.2%	< 0.1%	$c\bar{c}$
					< 0.1%	< 0.1%	ZZ^*
						< 0.1%	$\gamma\gamma$

Figure 2.14: The SM branching ratios of all final states of the Higgs pair production.

2.2.6 Monte Carlo Modelling of Signal and Background

Hadrons consist of quarks, anti-quarks and gluons. These constituents, *partons*, rather than the whole hadrons, participate in hadron collisions. For the process of interest, $pp \rightarrow F$, where p is the proton and F is the final state, the goal is to measure the cross-section, σ :

$$\sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{ij} f_i(x_1, \mu_F) f_j(x_2, \mu_F) \hat{\sigma}_{ij}(sx_1x_2, \mu_R), \quad (2.33)$$

where i and j can be any of the flavours of any of the QCD constituents, while x_1 and x_2 are the fractions of the total energy of the proton, carried by a given parton involved in the collision. f_i and f_j are the *parton distribution functions* (PDFs) which give the probability of finding the parton with a given momentum fraction. They depend on μ_F which is the factorization scale. $\hat{\sigma}_{ij}$ is the cross-section of the hard process which depends on the renormalization scale μ_R and s , the centre of mass energy squared.

To model what happens in the hadron-hadron collision, Monte Carlo simulations of the processes are used. The computation of the cross-section from Eq. 2.33 is based on the perturbation theory. Each cross-section is evaluated by the calculation of the process' amplitude, \mathcal{M} , as $\sigma \sim |\mathcal{M}|^2$. The perturbative expansion of the amplitude reads:

$$\mathcal{M} = \alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \alpha^3\mathcal{M}_2 + \mathcal{O}(\alpha^4), \quad (2.34)$$

where α , representing the coupling strength, must be sufficiently small for the perturbative expansion to work. The electroweak coupling at the electroweak scale is $\alpha_{\text{QED}} = \frac{1}{129}$, which makes the higher-order corrections very small and hence negligible. For QCD, the strength of the coupling varies depending on the energy range. For the LHC energy range $|q| > 100$ GeV (where q^2 is the amount of momentum transferred), resulting in $\alpha_s \approx 0.1$. Such a small value of α allows for the use of the perturbation theory [9] but it is not small enough for the corrections to be negligible. Therefore, calculations beyond the leading order are usually required for the QCD processes at the LHC.

The leading order term is the $\alpha\mathcal{M}_0$, which gives the cross-section at the leading order as:

$$\sigma_0 \sim \alpha^2 |\mathcal{M}_0|^2 \quad (2.35)$$

For the more precise determination of the cross-section, with smaller theoretical uncertainties, higher-order corrections to the amplitude are required. At the next-to-leading order, the cross-section reads:

$$\sigma_1 \approx |\alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \mathcal{O}(\alpha^3)|^2. \quad (2.36)$$

which, when expanded, leads to the inclusion of the loop diagrams and the interference term between \mathcal{M}_0 and \mathcal{M}_1 . That is why the calculations are very difficult.

The HH VBF production involves a large momentum transfer, which means that the perturbation theory can be used for the cross-section calculation. The SM cross-section has been computed up to next-to-next-to-next-to-leading order³ (N³LO) [24].

The *multi-jet* QCD background consists of jets either initiated by heavy quarks: bottom or top, or by light quarks incorrectly tagged as b-jets. The complexity of the multi-jet QCD processes, for example, the presence of the non-perturbative effects, causes a computational complication, especially for the computation of higher order diagrams which also include the loop diagrams. Moreover, the QCD background cross-section is orders of magnitude larger compared to the VBF $HH \rightarrow b\bar{b}b\bar{b}$ signal cross-section. Since the background cross-section is large, the amount of computational expense to produce sufficient background event statistics for the analysis is enormous. Therefore, to avoid the complication and expensive computation, data-driven methods are chosen for the background estimation in the analysis.

2.2.7 Di-Higgs Searches Status

The VBF di-Higgs production decaying to $b\bar{b}b\bar{b}$ channel was explored by the ATLAS collaboration in the non-resonant, resolved⁴ analysis with the full Run 2 data [3], [26]. It finds the observed interval for κ_{2V} values as [-0.55, 2.72], while other κ_{2V} values are excluded at 95% CL. The most recent, resolved non-resonant CMS result finds this interval to be [-0.1, 2.2] [27]. The boosted $b\bar{b}b\bar{b}$ non-resonant search was also performed by CMS [28], the κ_{2V} observed interval is then [0.62, 1.41], for the first time excluding the $\kappa_{2V} = 0$ scenario in which the HHVV vertex would vanish.

The ggF $b\bar{b}b\bar{b}$ results are the ATLAS partial Run 2 (data collected in 2015 and 2016) non-resonant publication [29], the ATLAS full Run 2 resonant analysis [30], and both CMS publications: resolved non-resonant and boosted non-resonant [27], [28].

A number of Higgs pair searches were performed both by ATLAS and CMS in other decay channels. The best up-to-date non-resonant result for the VBF comes from the combination of the $b\bar{b}b\bar{b}$ with the $b\bar{b}\tau^+\tau^-$ and the $b\bar{b}\gamma\gamma$ decay channels resulting in the κ_{2V} interval of [0.1, 2.0] [31], while from the CMS combination of the multilepton, the $b\bar{b}\gamma\gamma$, the $b\bar{b}\tau^+\tau^-$ and the $b\bar{b}b\bar{b}$ decay channels it is [0.67, 1.38] [32].

³In the analysis the MC signal samples are generated at the leading order but the theoretical N³LO result for the SM cross-section allows to rescale the LO cross-sections to obtain more precise values.

⁴Resolved means four b-jets are identified separately in the final state. The complementary category is *boosted*, in which not all b-jets are resolved.

Chapter 3

LHC and ATLAS detector

The focus of this chapter is the experimental set-up used to collect the data: the Large Hadron Collider and the ATLAS detector.

3.1 The Large Hadron Collider

In order to obtain the highly energetic particles, in this analysis the two Higgs bosons with a mass of at least 125 GeV each, a huge amount of energy must be involved, given by the famous $E = mc^2$ equation. Such energy can be produced by first accelerating and then colliding beams of particles. In order to accelerate particles to the desired energies efficiently, a circular rather than a linear collider is used, as the higher the required acceleration, the larger the distance the particles must travel in the electric field. The boost from the electric field is felt by the particles multiple times when they continue moving around the circle. To achieve the same energy levels in a linear collider, a much larger dimension of an accelerator would be needed.

The Large Hadron Collider (LHC) is the biggest of such circular accelerators that has ever existed. It was built around 100 meters below the ground at the French-Swiss border near Geneva and has a circumference of 27 km [33]. The colliding particles are hadrons, protons or heavy ions, which are circulating in two beam tubes of opposite directions and are forced to collide at four points. These four experimental sites are ATLAS, CMS, LHCb and ALICE. The first two are the general purpose detectors, the third concentrates on the flavour physics, while the last focuses on heavy ion collisions.

The data analysed in this thesis come from the beams of protons colliding at the centre of mass energy of 13 TeV. There are several stages between the injection of the hydrogen atoms until the collision itself. Figure 3.1 shows the accelerating complex. Its functions are described below based on [33] and [34].

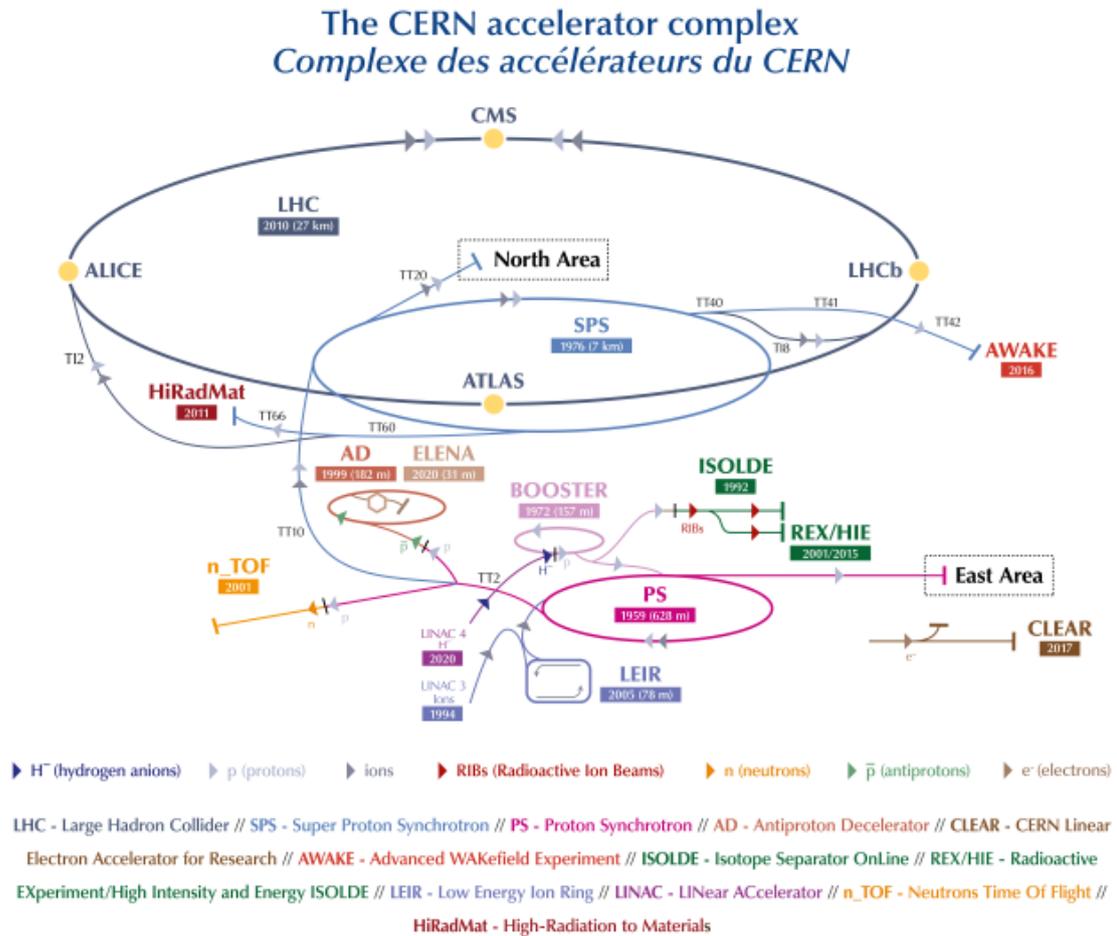


Figure 3.1: Schematic of the CERN accelerator complex. The most important stages needed to obtain the proton-proton collisions are described in the text [34].

First, in the electric field, electrons are removed from the hydrogen resulting in protons. Such protons are injected into a linear accelerator (*LINAC 2*) to be initially accelerated to 50 MeV. Protons are then processed by the *Proton Synchrotron Booster* which increases the beam energy to 1.4 GeV. Next, the *Proton Synchrotron* pushes the beam energy further to 25 GeV, after which the *Super Proton Synchrotron* allows to reach the energy of 450 GeV. At this point, the LHC itself takes over and makes each beam travel with the final energy of 6.5 TeV.

The key hardware components needed to accelerate particles in a circular collider are magnets and radiofrequency (RF) cavities. Magnets take care of the transverse acceleration, i.e. perpendicular to the velocity of the charged particles in the beam, keeping the circular paths of the beams. The RF cavities produce longitudinal acceleration through oscillating electric fields.

The main magnets are the state-of-the-art superconducting niobium-titanium dipoles that can provide a magnetic field of 8.22 T. For the purpose of superconductivity, they are cooled down to 1.9 K using liquid helium. Since it is not efficient to collide just one single proton with another single proton at a time, protons travel in bunches. Such bunches contain tightly clustered protons, which all have a positive charge. Quadrupole magnets are used to keep the protons together and to focalize the beams.

The RF cavities make the electric field change direction in the specific time intervals so that each time the bunch of protons from the beam passes through, it is accelerated. The frequency of the electric field oscillation corresponds to the proton bunch spacing of the LHC, which is 25 ns.

3.1.1 Luminosity

In order to quantify the size of the dataset that can be obtained from the experiment, namely the number of collisions produced per second per cm^2 , the integrated luminosity is quoted. The integrated luminosity is defined as:

$$L^{int} = \int \mathcal{L} dx \quad (3.1)$$

where \mathcal{L} is the luminosity. With colliding proton beams approximated by the Gaussian distributions and while the LHC collisions are head-on, the luminosity reads:

$$\mathcal{L} = f \frac{N_b n_1 n_2}{4\pi\sigma_x\sigma_y}, \quad (3.2)$$

where f is the frequency of one revolution of a bunch, N_b is the total number of bunches in the collision, n_1 and n_2 are the numbers of protons in either bunch and σ_x, σ_y are the RMS of the Gaussian profiles in horizontal and vertical directions respectively.

Eventually, the number of events that are produced with a corresponding quantum-mechanical cross-section σ is:

$$N = L^{int} \sigma, \quad (3.3)$$

meaning the larger the total integrated luminosity, the more events are accessible. Figure 3.2 shows the total integrated luminosity delivered by the LHC in Run 2 (operation of the LHC between 2015 and 2018). The total luminosity was 156 fb^{-1} , out of which 139 fb^{-1} [35] is usable for the physics analysis in the ATLAS experiment.

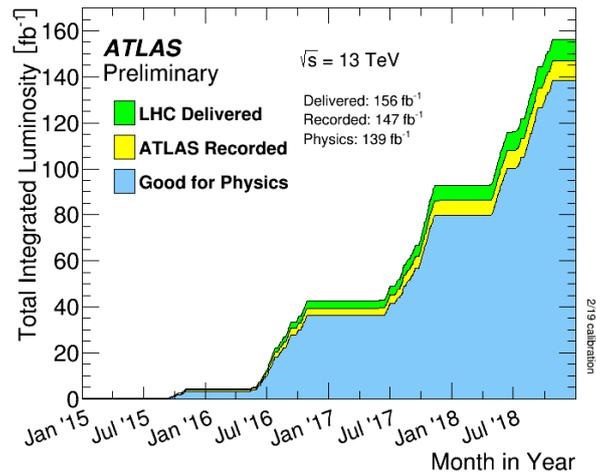


Figure 3.2: The total integrated luminosity in Run 2. The total luminosity recorded by ATLAS and the total luminosity good for physics analyses is indicated as well [35].

3.2 The ATLAS experiment

ATLAS is one of the two general-purpose experiments on the LHC ring. Its scientific aims are broad and include both precision measurements and searches for new physics using highly energetic proton-proton collisions. The detector weighs about 7 000 tonnes and is of a symmetrical cylindrical shape with a length of 45 meters and a diameter of 25 meters. It covers the range of almost 4π in solid angle and is split into three parts: the central barrel region, the end cap and the forward detector. Figure 3.3 presents the schematic of the detector with all sub-detectors and magnets indicated. These are:

- Inner Detector placed closest to the collision axis whose aim is to provide tracks of charged particles,
- Solenoid built from a thin superconducting material that allows the charged particles to bend in the Inner Detector (magnetic field of 2 T),
- Calorimeters (both electromagnetic and hadronic) that measure the energy deposited by charged and neutral particles interacting in the detector,
- Toroidal magnets needed for the charged muons to bend after they pass through the calorimeters,
- Muon spectrometer that measures the muon leptons created at the collision point.

Each part will be briefly discussed in this chapter, while the full description can be found in [36].

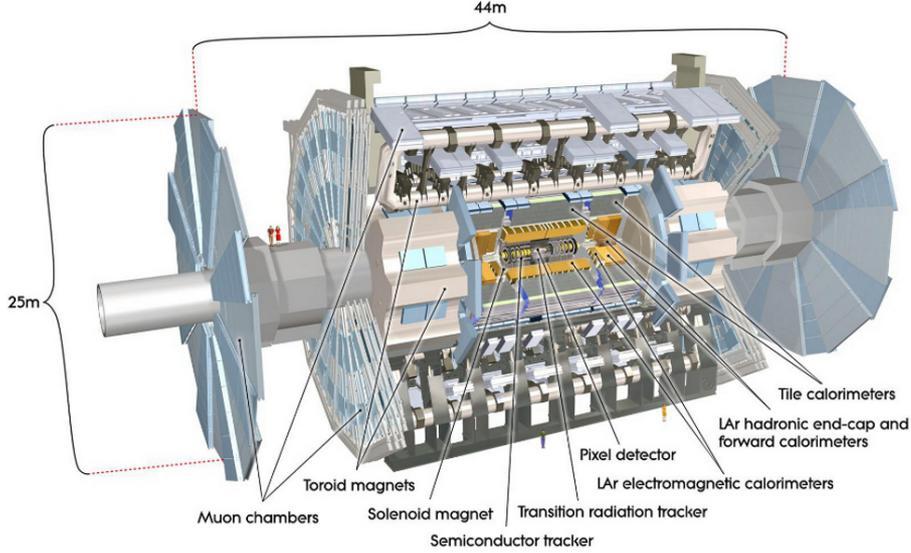


Figure 3.3: Schematic presenting the ATLAS detector and its constituents described in the text [37].

3.2.1 Coordinate System of the Detector

For clearer descriptions to follow, the ATLAS coordinate system is introduced first. Figure 3.4 schematically illustrates the coordinate system. The coordinate system is right-handed with the origin at the Collision Point. The z -axis is placed along the Beam Line, the x -axis points towards the centre of the LHC and the y -axis points upwards. To define the transverse plane, the cylindrical coordinates are used: (r, ϕ) , with ϕ being the azimuthal angle around the z -axis. The polar angle θ is also given and is used to define the two variables; pseudorapidity:

$$\eta = -\ln \tan(\theta/2), \quad (3.4)$$

and angular distance:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (3.5)$$

3.2.2 Inner Detector

The Inner Detector (ID) is a tracking device. Figure 3.5 shows the components of the Inner Detector: the Insertable B layer, Pixels, SemiConductor Tracker and Transition Radiation Tracker. The components are built from three types of sensors placed in a number of layers. Each layer receives an electrical signal when the charged particle passes through and ionises the active detector volume. By connecting these signals, *hits*, the particle's trajectory can be inferred.

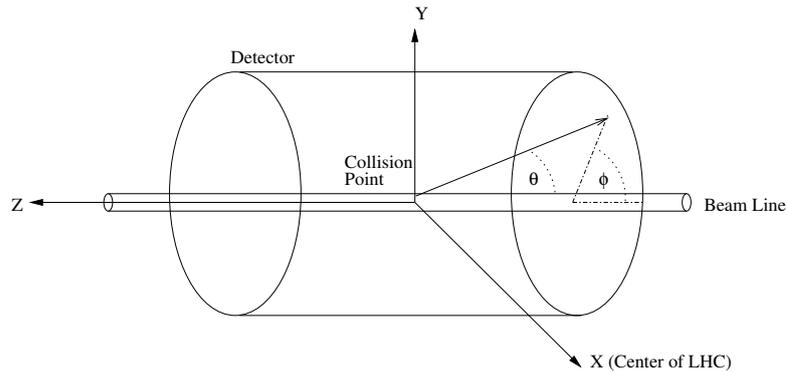


Figure 3.4: Schematic of the coordinates of the ATLAS detector with the centre at the Collision Point. Figure taken from [38].

Silicon Detectors

The first two layers closest to the beam pipe are the silicon detectors. Figure 3.6 shows a sketch of a silicon sensor. The silicon sensor chip is bump-bonded to its readout chip. When a charged particle goes through the silicon, it creates electron-hole pairs along the way. Electrons and holes naturally move via diffusion, but since the high voltage is applied to the sensors, the strong electric field causes the drift motion. Due to the charge carriers' movement, a signal in the readout chip is induced, as the chip is electrically coupled to the solder bumps. The readout chip pixel, closest to where the free charge carriers were generated, gets the most signal. By connecting all the readouts, the path of the particle can be inferred.

Closest to the beam line lies the high-granularity pixel detector, providing four layers in total. Each pixel is of size $50 \mu\text{m} \times 400 \mu\text{m}$, making it a total of around 80 million pieces, individual readouts in the detector. The innermost layer of the pixel detector is the *Insertable B layer*, IBL, which was added to the detector shortly before Run 2 in order to enhance the b-tagging efficiency [41].

The next ID component, the *SemiConductor Tracker* (SCT), consists of the silicon strip detectors, which work similarly to the pixels, but are not as small ($80 \mu\text{m} \times 12 \text{cm}$), hence their granularity is lower. There are four double layers of these strips. Pixel and SCT together cover the entire 2π region of the detector with $|\eta| < 2.5$.

Transition Radiation Tracker

On top of the two silicon detectors, at the range of $|\eta| \leq 2$, a technology-wise different Transition Radiation Tracker (TRT) detector is placed. The *drift tubes* filled with the Xenon gas are placed in the material instead of the pixel chips.

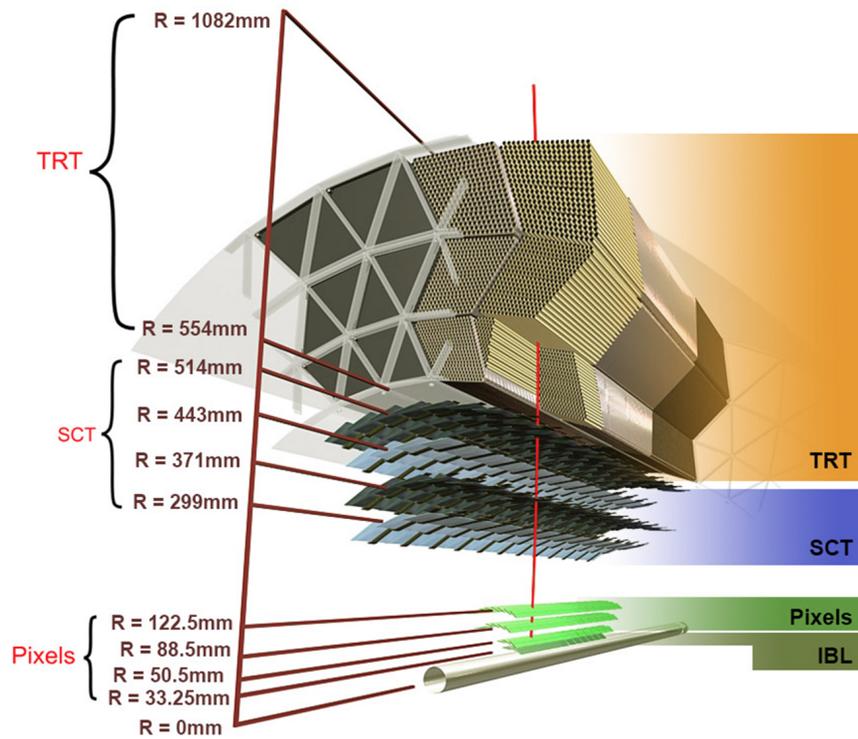


Figure 3.5: The structure of the Inner Detector components with the indication of the radial distances from the centre. Figure taken from [39].

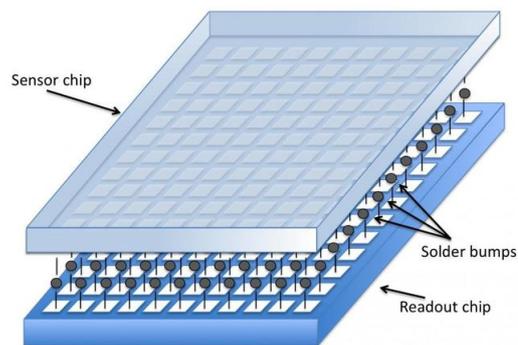


Figure 3.6: Sketch of a pixel detector with sensor and readout chips. Figure taken from [40].

When a charged particle passes through the material in between the tubes, it loses energy and photons are produced. The particle enters the drift tube and ionises the gas inside, thereby liberating electrons. The particle is accompanied by the radiated photons from the material surrounding the tubes, which also interact with the gas. The radiated photons free even more electrons which are then transferred to the gold wire, where the electrical signal is measured. Importantly, depending on the particle of interest, for example, whether it is an electron or a charged pion, the number of radiated photons differs. It comes from the fact that the loss of energy of the particle coming through the non-homogenous medium is proportional to the Lorentz factor; $\gamma = E/m$. Hence, the lighter the particle, the more energy it deposits and the more photons are produced causing a larger electrical signal in the gold wires. On average, around 36 hits are found in TRT for each collision.

The inner tracker detectors provide the positional information on the particle's track. In order to assign the momentum to such tracks, the ID is surrounded by the solenoid which produces the magnetic field of 2 T. In this magnetic field the relation

$$r = \frac{p_T}{qB}, \quad (3.6)$$

holds and with the arc radius measured from the hits, the transverse momentum is determined.

3.2.3 Calorimeter

To measure the particle's energy a system of calorimeters is used. In contrast to the tracker, both charged and neutral particles are recognized here. In the electromagnetic calorimeter, these are electrons, positrons and muons, and in the hadronic calorimeter, these are particles produced via a strong force, for example, pions, neutrons or protons.

The type of calorimeter that best serves the purpose of both stopping the particle and precisely measuring its energy, is the sampling calorimeter which consists of two material layers. The two types of materials are layered on top of each other; the heavy absorbers and the active materials. Most generally, particles are stopped in the absorbers while their energy is measured in the active material.

In the ATLAS detector, looking from its middle axis, calorimeters are placed after the Inner Detector covering the range $|\eta| < 4.9$. Figure 3.7 shows the ATLAS calorimeters. Both calorimeters are divided into those covering the central, barrel region, and those covering the end-cap and the forward regions.

The electromagnetic calorimeter is placed closer to the beam and it covers $|\eta| < 3.2$. The absorber is made of lead while the active material is liquid argon with copper electrodes inserted inside in order to collect the electric charge from the showered particles.

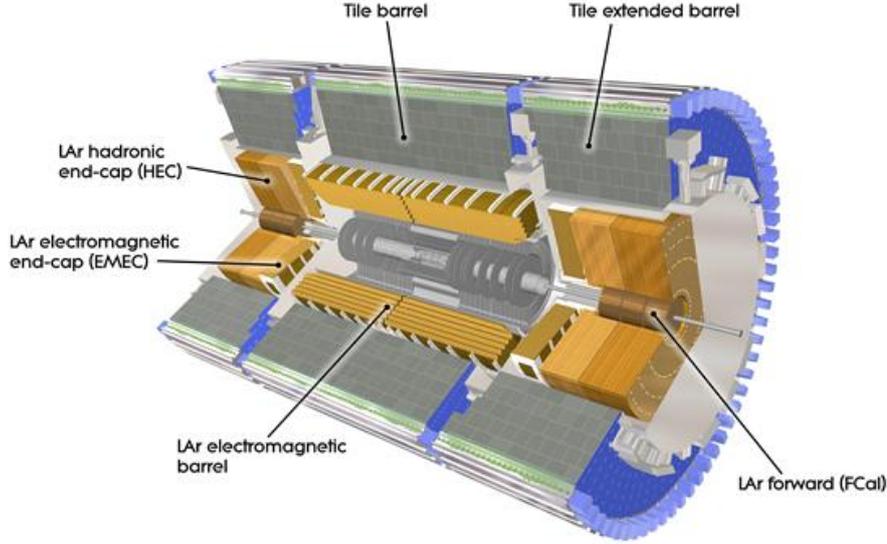


Figure 3.7: The calorimeter system in ATLAS [42].

When a high energy electron enters the absorber material, it is deflected by the electromagnetic field of the heavy nucleus and therefore it radiates a photon (bremsstrahlung occurs). The radiated photon produces a positron-electron pair and the new electron and positron behave in the same way as the first electron and radiate a new photon. Hence, the particle shower is produced in the absorber material and the original particle gradually loses all its energy until it is completely stopped. In the meantime, the low-energy showered particles ionize the active material placed in between the absorber layers, facilitating the measurement of the energy deposited by the original particle.

The hadronic calorimeter is placed on top of the electromagnetic calorimeter. Since hadrons are heavier, the amount of absorber needed in order to stop them completely is also greater. The central, barrel region ($|\eta| < 1.7$), is made of a steel absorber and the plastic scintillator tiles serve as an active material. When showered particles pass through the scintillator, a proportional amount of scintillating light is emitted and it is collected by the wavelength-shifting optical fibres connected to the photomultipliers in order to measure the deposited energy. At the end-cap ($1.5 < |\eta| < 3.2$), the active material is liquid argon while the absorber is copper. In the front calorimeter ($3.1 < |\eta| < 4.9$), the absorber is made of copper-tungsten layers.

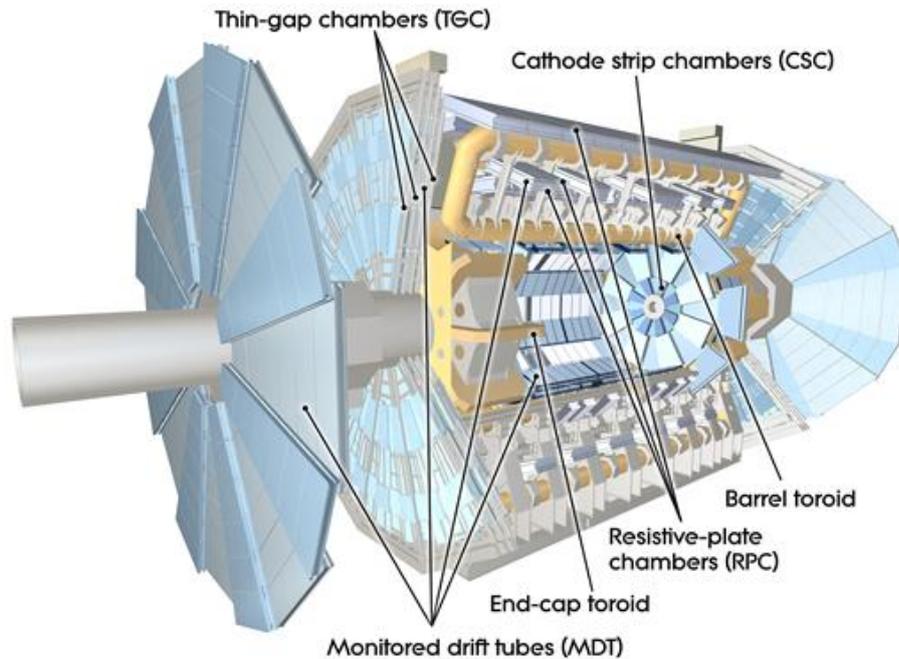


Figure 3.8: Schematic of the muon sub-system [43].

3.2.4 Muon Spectrometer

Since muons are much heavier than electrons ($m_\mu = 106$ MeV compared to $m_e = 0.511$ MeV), they can not be stopped by the electromagnetic calorimeters. Instead, they are detected in the *Muon Spectrometer* which is the last, outermost part of the ATLAS detector. Figure 3.8 shows the schematic of the muon spectrometer. Similarly to the TRT design in the Inner Detector, the muon chambers consist of drift tubes which allow for measuring the tracks of muons. The magnetic field (now produced by the massive toroidal magnets) allows for measuring the momentum of the tracks. The tracking chambers are placed in three layers at the region of $|\eta| < 2.7$ and in the forward region, where the cathode-strip chambers are used in addition to the drift tubes. The complementary chambers are used for the fast-track triggering. They cover the $|\eta| < 2.4$ region and are made of the resistive-plate chambers in the barrel, and the thin-gap chambers in the end-caps.

3.2.5 Triggering and Data Acquisition

The number of collisions that happen in the ATLAS detector reaches 40 million each second. Even with cutting-edge technology, it is impossible to save all this information, and it is also not necessary for the ATLAS physics goals. The job of choosing and saving only the potentially interesting information is done by the

triggering system. Figure 3.9 schematically shows the ATLAS triggering system in Run 2, while more details on the TDAQ Run 2 performance can be found in [44].

Firstly, events must pass a hardware-based, Level 1 Trigger requirement. Level 1 Trigger takes the raw data from the calorimeters, the muon spectrometers and the inner detector. With a simple and quick object reconstruction, L1 reduces the number of events to be processed per second from 40 million to 100 000. Secondly, the selected events, from different Regions of Interest of the detector, are passed to the software-based High Level Trigger (HLT), where information from the full detector is available. At the same time, if the Level 1 Trigger accepts an event, the Readout Drivers (detector’s front-end electronics) send the event data to the Readout System, which makes them available to the HLT so that different trigger algorithms can be run on the sampled data while the rest of the data is buffered. After the acceptance by the HLT, the data, around 1500 events per second, are moved to the permanent storage via the Data Logger. They can be accessed for object reconstruction first, and finally for physics analysis.

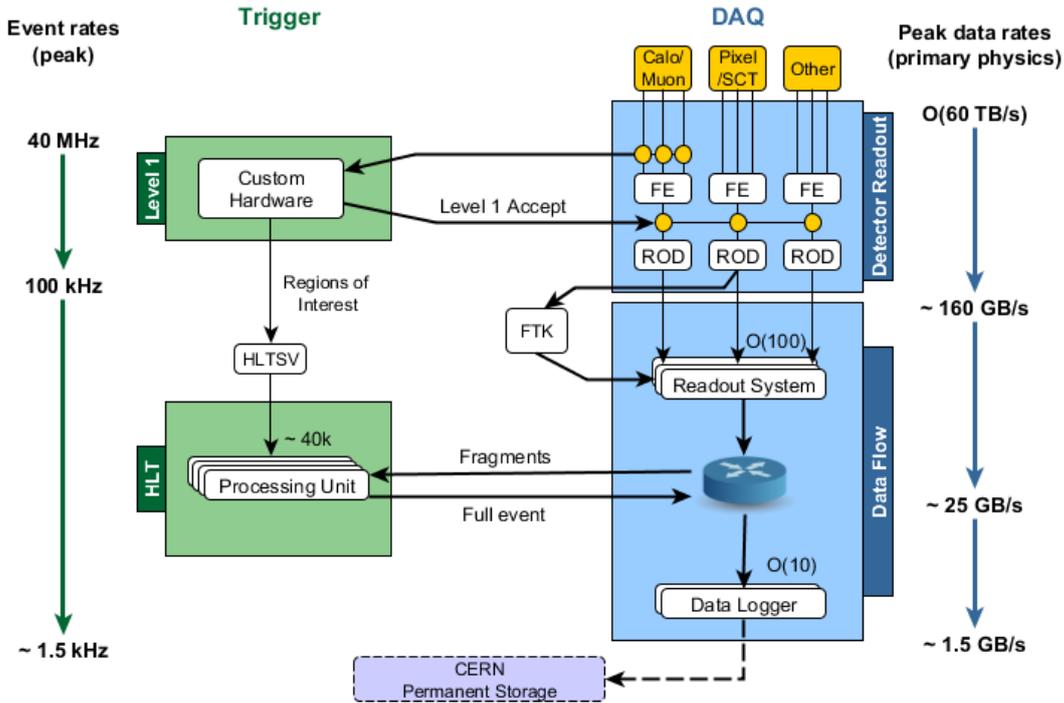


Figure 3.9: The TDAQ system of the ATLAS detector in Run 2. The most important components are briefly described in the text. Figure taken from [45].

Chapter 4

Machine Learning

In the past few years, the popularity of Machine Learning (ML) in various fields of science has sky-rocketed, including Particle Physics in general, and the analysis presented in this thesis in particular. In this chapter, a short overview of ML, with a focus on neural networks, is presented in order for the reader to fully grasp the ideas developed in the next chapters.

4.1 Supervised Learning

First of all, what Machine Learning really is? In a usual, non-ML computing system, a programmer gives the program some inputs and writes an algorithm, so that the program can process these inputs and give the outputs. In ML, the computers have the ability to learn the inputs and sometimes outputs (in supervised learning) all by themselves, without needing an explicit programming step. They can become very good at their task, which can be assessed using various quantifiers, explained in this chapter.

Supervised learning is what is used in the analysis, and generally in ATLAS, hence solely this part of the field will be discussed here, following [46] and [47].

In the first step of the supervised learning, the program is given n known sets of data, which include the inputs $X = \{x_1, x_2, \dots, x_n\}$ and corresponding outputs $Y = \{y_1, y_2, \dots, y_n\}$ ¹, so that the program can learn the functional dependence between the two, $Y = f(X)$. Once $f(X)$ is known, it can be applied to the new set of X for which Y is unknown.

Two developments are possible, depending on whether X and Y are continuous or discrete. For the continuous case, regression modelling is performed with the classic equation being:

$$y_i = \alpha x_i + \beta, \tag{4.1}$$

¹Note, both x_i and y_i can be single data points, but they are usually tensors.

where y_i is a vector of outputs, x_i of inputs, α is a vector of parameters and β is the scalar offset. In order to find values of α and β , an optimization procedure is necessary. In the simplest case, this involves minimizing the least-squared error (LSE) which reads:

$$\sum_{i=1}^n (\alpha x_i + \beta - y_i)^2. \quad (4.2)$$

If instead, X and Y are discrete, a classification approach is taken. As an outcome of learning, the program answers the *yes* or *no* question and assigns inputs to the pre-defined categories. For example, during an event selection, the signal from background distinction is made. For such a task, the model called *logistic regression* is used in most cases. Instead of optimization of the LSE, a logistic function is optimized:

$$f(x_i) = \frac{1}{1 + e^{-(\alpha x_i + \beta)}}. \quad (4.3)$$

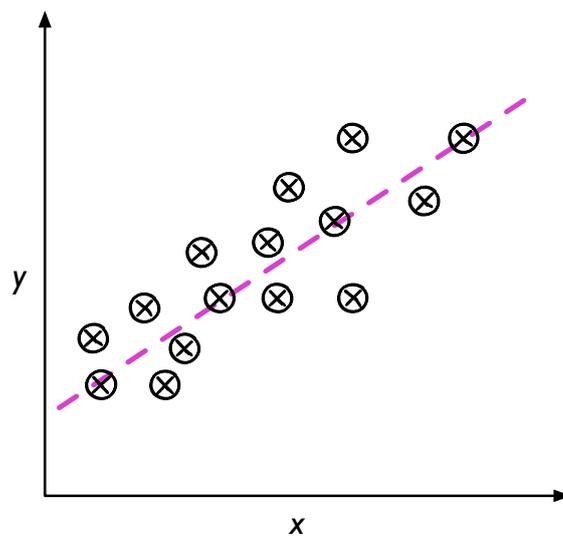
In particular, the α and β parameters have to be maximized via the maximum likelihood method. The values of $f(x_i)$ range between 0 and 1. Hence, the value obtained for a given input x_i can be interpreted as a probability of its classification in one of the categories. Figure 4.1 illustrates linear regression and classification approaches.

4.2 Neural Networks

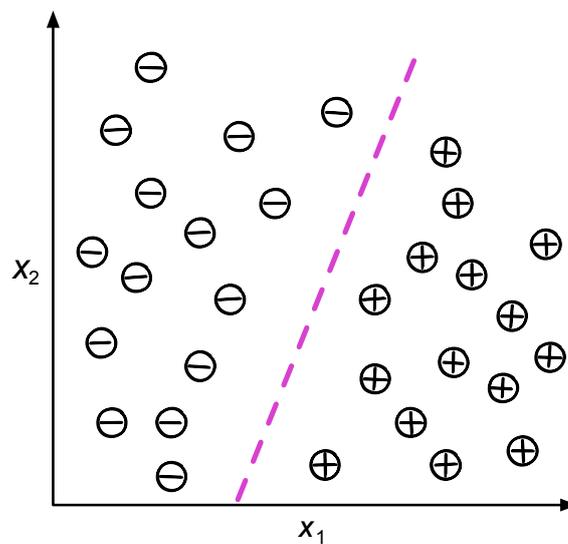
The goal is for the model to learn the connections between inputs and outputs and to predict unknown outputs as a function of inputs. The simplest, although very powerful, functional relations are explained above. However, not all dependencies between variables are as simple. The dependencies which are not linear, as in the case of many particle physics variables, need more advanced models. Inspired by the biological network of neurons and an appreciation of brain functionality, neural networks have been developed. In simplest words, they are networks of layers of nodes interconnected via weighted neurons, as schematically presented in Figure 4.2.

We focus on the classic feed-forward networks. Therefore, following the schematic, we start from the left and continue to the right. Firstly, the network receives the data via the input layer. Each node is a different x_i connected by neurons to all nodes in the second layer. Each of these neurons is randomly assigned a weight (this is the *random initialization*). The question arises, what values do the nodes have in the second layer?

Let's assume there are m nodes in the input layer. Each node in the second layer is connected to all nodes of the input layer. There are m connections per node, each with a random weight, w_{ji} , here connecting $j = 2$ layer with $j = 1$ input layer.



(a) Linear regression



(b) Classification

Figure 4.1: Illustration of the two types of modelling done in Machine Learning: regression and classification. In (a) the pink line indicates the final regression output while in (b) the pink line is a linear decision boundary in order to classify data into either "-" or "+" category. Figure taken from [47].

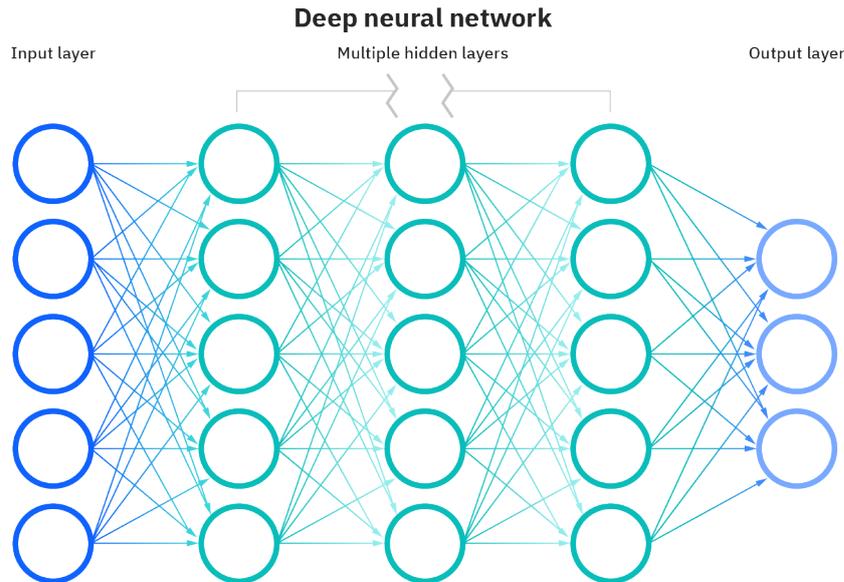


Figure 4.2: Schematic of a deep neural network with nodes indicated as circles and connected to other nodes by neurons, the first layer is the input layer, while the last is the output layer [48].

Moreover, it has to be noted that each non-input node has an associated *bias* or *threshold* term, b_j and an *activation function*, $f_{\text{activation}, j}$. The value of the threshold decides whether a given node is activated or if it is discarded. The value of a node in the second layer, x_j with $j = 2$, is given by:

$$x_j = \left(\sum_{i=1}^m w_{ji} x_i + b_j \right) \times f_{\text{activation}, j}. \quad (4.4)$$

The second layer is fed the values from the first layer and therefore the method is called *feed-forward*. For the third layer, or any further layer k , it continues as:

$$\begin{aligned} x_k &= \left(\sum_{j=1}^{m_2} w_{kj} x_j + b_k \right) \times f_{\text{activation}, k} \\ &= \left(\sum_{j=1}^{m_2} w_{kj} \left(\sum_{i=1}^m w_{ji} x_i + b_j \times f_{\text{activation}, j} \right) + b_k \right) \times f_{\text{activation}, k}, \end{aligned} \quad (4.5)$$

with m_2 being a number of nodes in layer 2 and w_{kj} the weights of the neurons connecting layer two with the layer three. The direct inheritance from Eq. 4.4 onto Eq. 4.5 can be noticed. The procedure continues until the output layer is reached. All layers in between the input and output layers are the *hidden layers*. Networks

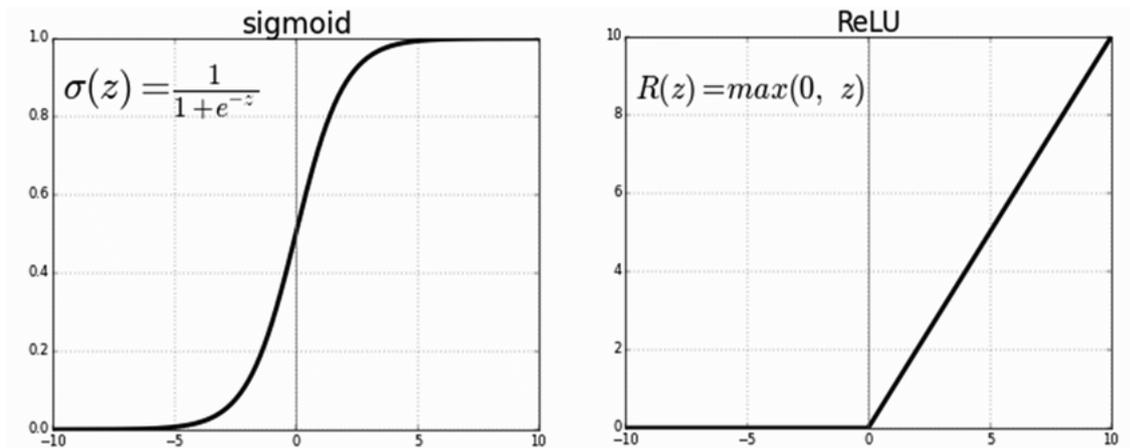


Figure 4.3: Example graphs of sigmoid and ReLU functions which are the most common activation functions used by Neural Networks [49].

with more than one hidden layer are referred to as *Deep Neural Networks*. The number of such layers as well as the number of nodes in each layer are the network *hyperparameters*. Depending on the complexity of the problem and the number of training data available, both of these have to be tuned.

Defining the activation function is an important part of the neural network method. Without an activation function, the problem would narrow down to linear regression and the non-linear relations would not be explored. There is a number of activation functions commonly used. For instance, in the classification task, the output is the probability of the data being assigned to a given category. That is why, in the output layer, the used activation functions are functions with values smoothly rising between 0 and 1. The most classical example of such a function is the sigmoid function, which reads:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (4.6)$$

For the hidden layers, the common choice is the *Rectified Linear Unit* function, or for short ReLU. For the inputs that are less than 0, it gives the output of 0, while it is a linear function for the positive inputs, as:

$$R(x) = \max(0, x). \quad (4.7)$$

The illustration of both functions is shown in Figure 4.3.

The general idea of how we reach the output layer is shown. To assess whether the network predicted the output correctly and whether the set of weights is optimal (since a large part of the modelling is done based on randomly initialized weights), the so-called *loss function* is defined. The goal for the network is to iteratively continue choosing weights such that the loss function is minimized. Many such

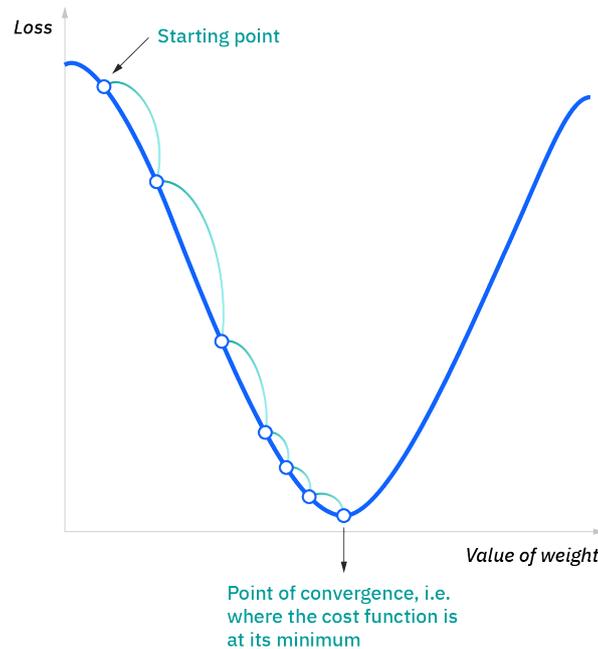


Figure 4.4: Illustration of gradient descent method. As weights are optimised, the Loss function reaches its minimum [48].

functions are available. A common, simplest choice, is the *Mean Square Error, MSE*:

$$\text{Loss function} = \frac{1}{2n} \sum_{i=1}^n (y_i - f(x_i))^2, \quad (4.8)$$

where n is the number of outputs, y_i are the known outputs from the dataset corresponding to given x_i , while $f(x_i)$ are the outputs predicted by the network.

The iterative technique which allows updating of the weights and minimises MSE is called the *gradient descent*. The gradient of the loss function with respect to the value of the weight has to be computed. It is computed for all the weights in the network starting from the last layer and going backwards to the beginning (hence this is the *back-propagation* method). The weights are updated by the fixed value, called the *learning rate*, and the same procedure is repeated until the minimum is reached. Each repetition is an *epoch* of the training. An illustration of how gradient descent works is shown in Figure 4.4.

Two issues can appear during the network training. If the network does not reach the minimum of the loss function, *underfitting* occurs. The other case is that the network learns the detailed connections between X and Y of the training set, but it is unable to make correct predictions on the unknown data, leading to an *overfitting*.

The overfitting is usually monitored on an additional set of data, the validation data. A comparison of the training and the validation loss is performed. In the case of overfitting, for the training data, the loss decreases with the increased model complexity, while for the validation data, the loss increases with the increased complexity. The methods such as the regularization of penalizing large weights, or dropout, where some connections are randomly omitted during each epoch, have to be employed to avoid overtraining [47].

Chapter 5

Analysis Relevant Objects

The method of obtaining data from the particle collisions was introduced in Chapter 3. This chapter describes the process of *reconstruction* of the physical objects for the analysis. Since four b-quarks constitute the final state, the reconstructed objects of interest are the b-tagged jets. This chapter specifically focuses on jets and the b-tagging procedure.

5.1 Jets

Observing pure quarks in the detector is impossible because of the quark confinement. Instead, the high-energy hadrons built from the collimated quarks and gluons, *jets*, are identified in the detector. For a systematic identification, an algorithm-based definition of a jet is required. A number of such algorithms exist, but the focus is given to the algorithm used in ATLAS¹ and thus in this analysis: the sequential-recombination algorithm called the *anti* – k_t [50].

5.1.1 The *anti* – k_t Algorithm

The sequential-recombination algorithm works based on a reversed approach with respect to the jet creation process, in which one parton successively branches out into more particles. In the *anti* – k_t algorithm, the final particles, as seen in the detector, are the starting point². They are grouped one by one based on either the inter-particle distance $d_{i,j}$ or the jet-beam distance, d_{iB} . The two are defined as:

$$d_{ij} = \min \left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad (5.1)$$

¹However, the algorithm is used not only in ATLAS.

²Same applies for the k_t and Cambridge-Aachen algorithms with the difference of the d_{iB} definition in Eq. 5.2.

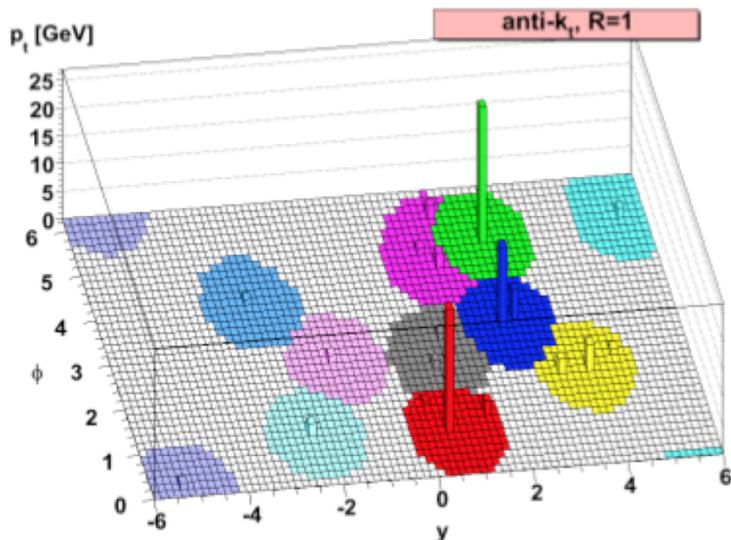


Figure 5.1: The result of the jet clustering with $anti - k_t$, here for jet radius of 1. Figure taken from [52].

where $k_{t,i}$ and $k_{t,j}$ are the momenta of the two objects i and j w.r.t. the beam axis, ΔR_{ij} is the geometric distance in the rapidity - azimuthal angle ($y - \phi$) plane and R is a jet radius (a free parameter of choice), while

$$d_{iB} = k_{t,i}^n, \quad (5.2)$$

where $n = -2$ in case of the $anti - k_t$ algorithm.

An object/particle is chosen, and it is associated with the next object/particle that lies closest in the $y - \phi$ plane. The one-by-one clustering is repeated as long as $d_{ij} < d_{iB}$. When $\Delta R_{ij} > R$, $d_{ij} > k_{t,i}^{-2} = d_{iB}$, the object i (the harder from objects i and j), becomes the jet and no more objects are attached to it.

In Eq. 5.2, $n = -2$ is the way of weighting the impact of the momentum as $\min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right)$. Therefore, the hard particles are clustered first, and the soft particles are attached. This way, the soft radiation is avoided, and the shape of the jets becomes circular, making the jet calibration easier in the experimental method [51]. Figure 5.1 illustrates the result of jet clustering with $anti - k_t$ algorithm, with the jet radius of 1.

5.1.2 Jets in ATLAS

In ATLAS Run 1, only the topologically clustered calorimeter cells, *topo-clusters* [53], were used to reconstruct jets. In Run 2, a new algorithm, called *Particle Flow*, is proposed, in which both tracking information and the topo-clusters are used [54].

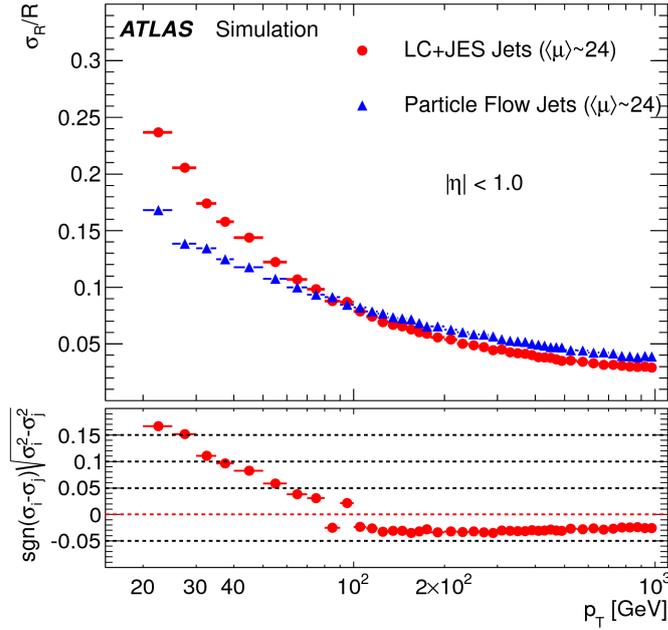


Figure 5.2: Comparison of the p_T resolution of the particle flow jets to the topological clusters jets (LC+JES) as a function of the jet's p_T . The bottom panel serves as the performance comparison of the two algorithms as the square root of the difference of the squares of the resolution of the p-flow and topo-clusters jets [54].

Particle Flow Jets

The tracks of the charged particles are first matched to the energies of the particles deposited in the calorimeter. The deposited energy is removed from the calorimeter in order to avoid double counting. This results in the *particle flow objects*, which are either the tracks matched to the charged hadrons³ or the remaining calorimeter energy deposits for neutral hadrons (those that had no tracks associated with them). The jets are clustered from both types of objects with the *anti* - k_t algorithm.

Figure 5.2 shows the comparison of the p_T resolution of the particle flow jets to the p_T resolution of the topological clusters. The overall improvement in energy resolution performance is significant, especially at low p_T [54]. This can be explained by the topo-clusters suffering from the high levels of stochastic noise at low energies, which is avoided when tracking information is used instead, as momenta of tracks are well-known. The improved performance motivates the use of this algorithm, despite the necessity of the complicated track-to-calorimeter signal matching to avoid the energy double counting.

³Those have the energy deposits removed.

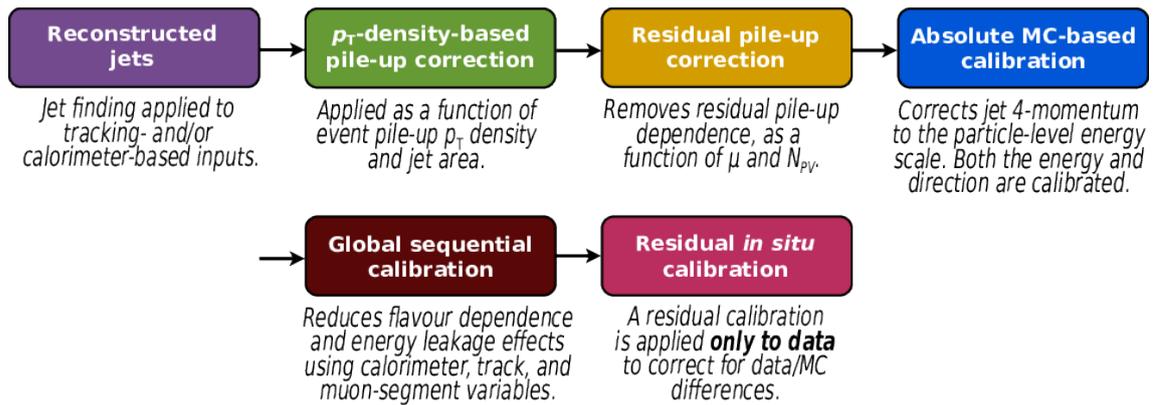


Figure 5.3: Illustration of jet calibration steps in order to match the detector reconstructed jets to the MC jets. Figure taken from [55].

5.1.3 Jet Calibrations and Corrections

The detector output is biased due to a number of associated effects, such as pile-up and calorimeter inefficiencies. Figure 5.3 shows the steps of jet calibration in ATLAS. The reconstructed jet’s energy, mass and p_T are matched to the corresponding, unbiased values for the Monte Carlo jets.

Figure 5.3 presents the summary of the calibration steps, while the full description can be found in [55]. The first step is the removal of any energy remaining in the calorimeter which is left from the previous bunch crossing, this is the so-called *out-of-time-pileup* correction. The second step is the removal of the *in-time-pileup* based on the average number of particles per bunch crossing (μ) and the number of primary vertices (N_{PV}). Next, the MC jet p_T four-vector is used to correct the p_T of the detector reconstructed jet. Subsequently, the correction factors for the jet energy resolution are applied as six multiplicative correction factors denoted as *Global Sequential Calibration*. Finally, the in-situ calibrations are used to correct for any remaining discrepancies between the data and MC based on precisely studied reference objects, for instance, Z bosons or photons.

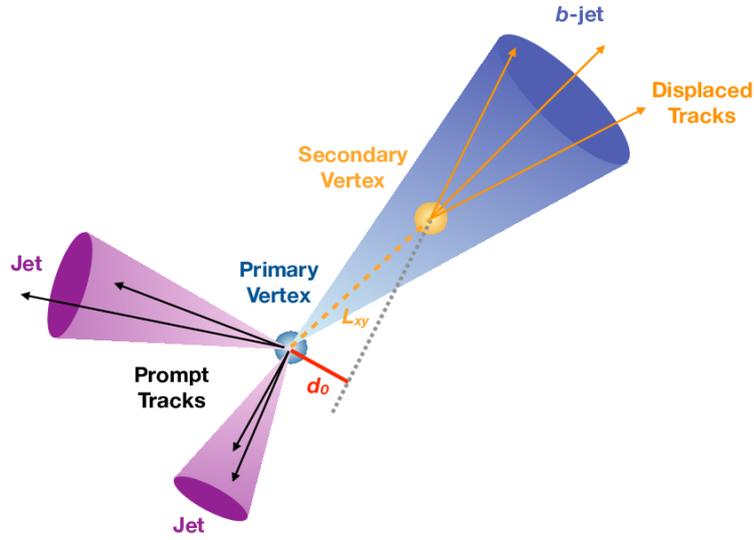


Figure 5.4: Schematic representing jet tagging in ATLAS. Attention should be drawn to the b -jet and its secondary vertex, which is distance L_{xy} away from the primary vertex, associated with the B-hadron’s longer lifetime. As a result, displaced tracks are formed. The existence of tracks in the detector can be inferred by inspecting the transverse impact parameter d_0 . B -jets are compared to the prompt jets, decaying at the primary vertex directly. Figure taken from [56].

5.2 b -tagging

In order to know which type of quark initiated the given jet, the flavour tagging procedure is performed. Jets are divided into light-flavoured: coming from light quarks such as d , u or s , or gluons, or heavy-flavoured: either c -tagged or b -tagged. The last is our tagging of interest.

The b -quarks (B hadrons resulting from the confinement) can be easily distinguished in the detector due to their fundamental properties: long lifetime (~ 1.5 ps) and large mass. Because of the long lifetime, the path that a B hadron travels before the decay is $c\tau \approx 0.45\text{mm}$, resulting in a displaced secondary vertex in the detector. In practice, this displacement is recognised based on the transverse and longitudinal impact parameters of the tracks. Large values of the impact parameters indicate the displaced vertex. A schematic illustrating the b -jet’s displaced vertex is shown in Figure 5.4.

5.2.1 Algorithms in ATLAS

The impact parameters and information on the secondary vertices of the B hadrons are used in ATLAS flavour tagging algorithms. The methods are distinguished as the *low-level* and the *high-level* taggers. The low-level taggers directly use track and vertex information in order to initialise the tagging procedure. These algorithms are the IP2D (3D) algorithm which uses the 2D (3D) impact parameters, SV1 which uses the secondary vertex information and the JETFITTER, which fits the displaced vertices. The RNNIP is a novel neural network-based tool to identify b-jets using the track information only. More information on the first three algorithms can be found in [57], while on RNNIP in [58].

The outputs of the low-level taggers are used by the high-level taggers which are complex multivariate algorithms. Such an approach is proven to enhance the b-tagging performance. The final jet flavour assignment comes from the high-level taggers' discrimination power. The newest of such taggers is a deep neural network tool, called *DL1r* [59], which is used in the analysis. It replaces its Boosted Decision Tree predecessor, the *MV2* algorithm [57]. The new algorithm takes all of the above low-level taggers as inputs, including the RNNIP. The deep neural network DL1 algorithm also exists, but it does not use the RNNIP output as its input.

Figure 5.5 shows the performance of the DL1r tagger compared to the MV2 and DL1 taggers. The y -axis represents either the light-flavour rejection efficiency (in a) or the c -jets rejection efficiency (in b), while the x -axis is the b-tagging efficiency. The DL1r algorithm outperforms the other two algorithms. The comparison with the MV2 is especially important as the MV2 was used in the previous VBF analysis [3]. The use of the new b-tagging algorithm contributes to the improved result of the analysis.

5.2.2 Working Points

The high-level DL1r algorithm discriminates whether the jet is, or is not, b-tagged. By setting a threshold on the discriminant value such that the desired efficiency of the b-jet identification is achieved, an operating point of the algorithm is defined [61]. The efficiency is determined by studying the Monte Carlo simulated data sample consisting of the $t\bar{t}$ events. In this analysis, the single-cut operating point at 77% b-jet identification efficiency is chosen as the nominal Working Point.

For validation purposes of the background estimation in Chapter 7, b-jets tagged at the Working Point of 85% are considered. The higher score, quite conversely, means that the selection/cut on the tagger's discriminant is in fact looser (as 100% efficiency would indicate the complete removal of such cut), but the b-jet efficiency is higher.

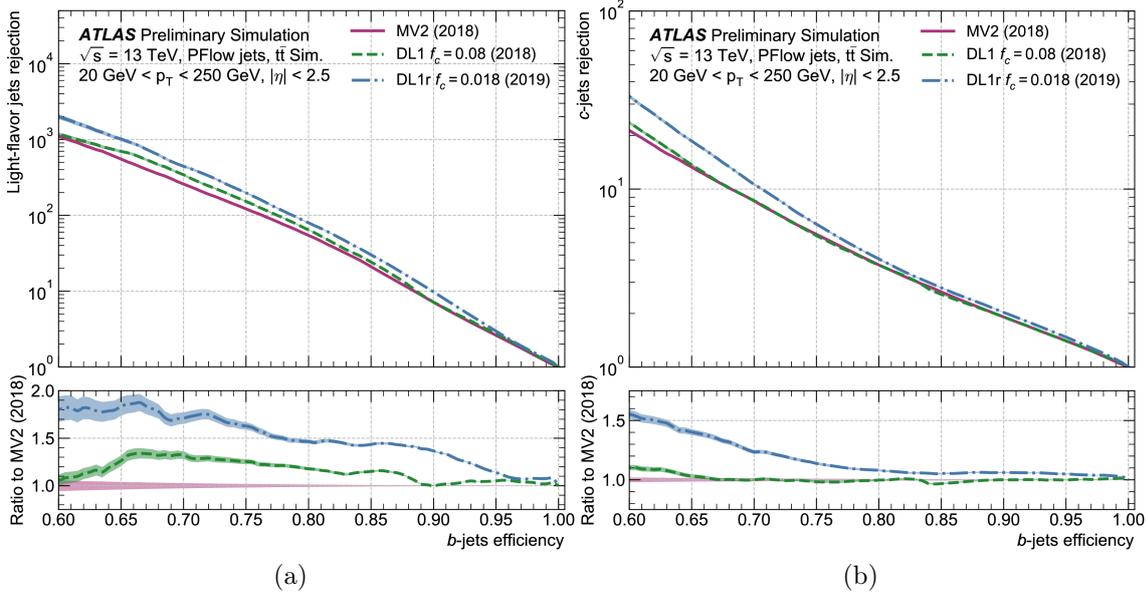


Figure 5.5: Comparison of light-flavour (in a) or c-jets (in b) rejection performance of three high-level taggers: MVA, DL1 and DL1r, as a function of b-jets efficiency. The lower panel shows the light-flavour or c-jets rejection as a ratio to the MV2 rejection [60].

5.2.3 Offline Flavour Tagging Calibrations

The efficiency is not a global value but depends on the kinematic properties of jets such as p_T or η . Since these can differ between the MC samples and the real data, calibration is needed. As a result, simulation-to-data scale factors are derived [57] for different working points, and are applied to the MC samples to correct for the efficiency in data. The general flavour tagging scale factor evaluated per-jet, is defined:

$$SF_j = \frac{\epsilon_j^{data}}{\epsilon_j^{MC}}, \quad (5.3)$$

where ϵ is the efficiency of tagging the flavour j in data or MC, and the efficiencies are measured as functions of p_T . For jets that are not flavour tagged, a tagging *inefficiency* scale factor is calculated as:

$$\overline{SF}_j = \frac{1 - \epsilon_j^{data}}{1 - \epsilon_j^{MC}} = \frac{1 - SF_j \cdot \epsilon_j^{MC}}{1 - \epsilon_j^{MC}}. \quad (5.4)$$

In order to evaluate the per-event scale factor, the product of all per-jet efficiency and inefficiency scale factors is calculated.

Figure 5.6 shows the scale factors for five different pseudo-continuous operating points. The values of the scale factors are close to 1. The associated total statistical

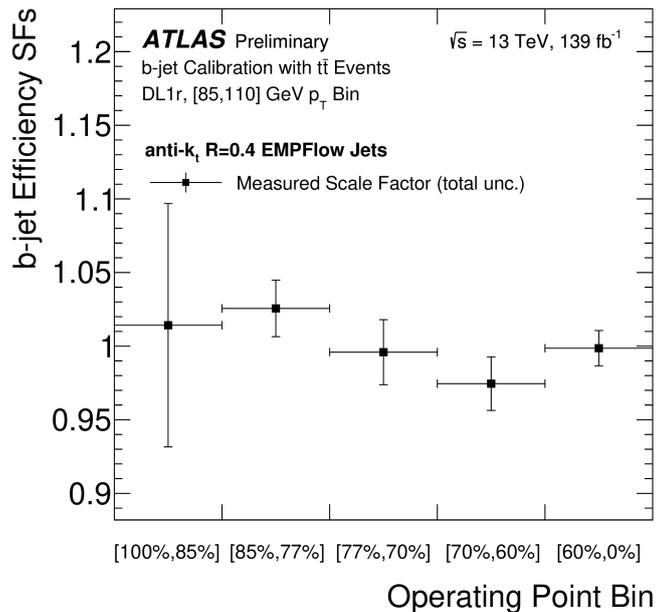


Figure 5.6: Scale factors as b-jet efficiency ratios in data over MC for different pseudo-continuous operating points [62].

and systematic uncertainty varies. For the 77% working point, the uncertainty is around 2% when measured with the b-jets in the [85, 110] GeV p_T bin. Figure 5.7 shows the offline b-tagging scale factors as a function of the particle flow jet p_T for the 70% working point⁴. The SF is close to 1 for most of the p_T spectrum, except for the $p_T > 250$ GeV bin, where it is 0.97. The corresponding total statistical and systematic uncertainties are no larger than 2%.

⁴As no SF vs p_T plot is available for the 77% working point.

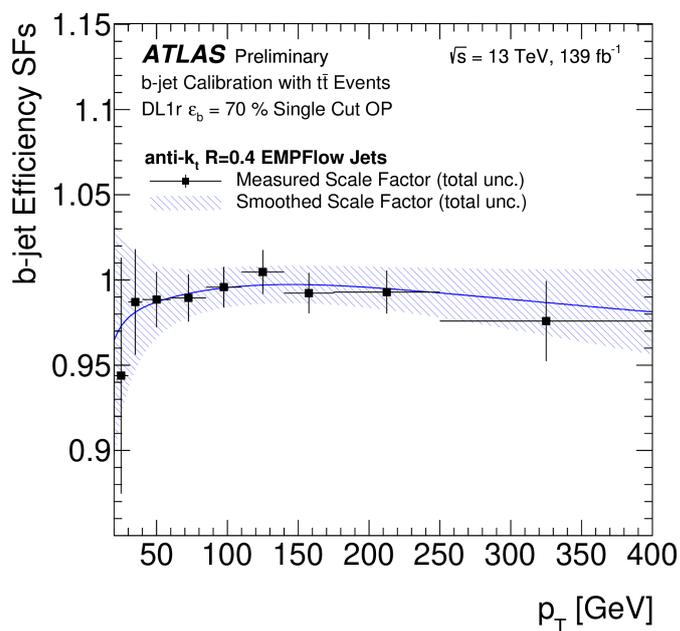


Figure 5.7: Scale factors as b-jet efficiency ratios in data over MC as a function of p-flow jet p_T for the 70% working point [62].

Chapter 6

Event Selection

Previous chapters introduced both the theoretical and experimental context needed to perform the ATLAS data analysis with the aim of finding Higgs boson pairs decaying to the four b-quarks final state. Starting from this chapter, the focus is put on the analysis and the results.

The scope of the non-resonant Run 2 $HH \rightarrow b\bar{b}b\bar{b}$ analysis performed by the ATLAS $HH4b$ analysis team, includes the search for the non-resonant ggF and VBF productions. This thesis work focuses on the VBF channel and details of the ggF analysis will not be discussed unless the choice of certain selections is motivated by the specifics of the ggF channel in particular.

The previous $HH \rightarrow 4b$ analyses performed by the ATLAS Collaboration are mentioned in Section 2.2.7. Given such a legacy, some analysis choices were directly motivated by the studies performed previously. In the course of this analysis, various improvements have been applied, leading to the final result being ameliorated by a factor of 6 compared to the previous non-resonant Run 2 VBF analysis [3], in which the same total integrated luminosity was used. Most importantly, regarding the analysis itself, the data-driven background estimation is changed from the histogram-based event reweighting technique to the neural network reweighting. Similarly, other important pieces of analysis such as the triggering approach and the Higgs Candidates pairing algorithm are optimised.

6.1 Data Samples

The data considered in this analysis are the ATLAS Run 2 $\sqrt{s} = 13$ TeV proton-proton collision data collected between 2016 and 2018. Run 2 also includes data from 2015, however, due to the lack of b-jet trigger scale factors for these data, they are not used in this analysis. The total used integrated luminosity is 126 fb^{-1} split as follows: 24.60 fb^{-1} coming from 2016, 43.65 fb^{-1} from 2017 and 57.70 fb^{-1} from 2018.

Table 6.1: All Monte Carlo generated signal points for the non-resonant VBF analysis.

DSID	κ_λ	κ_{2V}	κ_V
502970	1	1	1
502971	1	0	1
502972	1	0.5	1
502973	1	1.5	1
502974	1	2	1
502975	1	3	1
502976	0	1	1
502977	2	1	1
502978	10	1	1
502979	1	1	0.5
502980	1	1	1.5
502981	0	0	1

6.2 Monte Carlo Samples

Monte Carlo simulations are used for signal events only. The main background is comprised of various QCD *multi-jet* processes, which are not well modelled by Monte Carlo simulations as explained in Section 2.2.6. The background is estimated using data-driven techniques.

Signal samples, including the SM couplings point and the points with varied κ_λ and κ_{2V} couplings, were produced using MADGRAPH, version 2.7.3 [63] for the event generation at the leading order (LO), with the PDF set: NNPDF3.0NLO [64]. PYTHIA8.244 [65], was used for parton showering and hadronization with NNPDF2.3LO PDF set with A14 set of tuned parameters [64]. In order to simulate the detector response GEANT4 [66] was used. The generated signal samples are listed in Table 6.1.

Since samples are produced at LO, but the N3LO result for the $\kappa_{2V} = 1$ is available, all cross-sections are rescaled to the N3LO value as:

$$\sigma_{\text{N3LO}}^{\kappa_{2V}=x} = \sigma_{\text{N3LO}}^{\kappa_{2V}=1} \frac{\sigma_{\text{LO}}^{\kappa_{2V}=x}}{\sigma_{\text{LO}}^{\kappa_{2V}=1}}, \quad (6.1)$$

with $\sigma_{\text{N3LO}}^{c_{2v}=1} = 1.723 \text{ fb}$ [24] and $\sigma_{\text{LO}}^{c_{2v}=1} = 1.394 \text{ fb}$ [67].

6.2.1 κ_{2V} Reweighting

Even though twelve samples are produced, they are not enough for the overall study of the di-Higgs VBF production cross-section, for which a continuous range of the κ_{2V} values is needed. Producing more samples would be one possibility, however, it is computationally very expensive. The combination of a handful of samples is used instead. The combination has its mathematical foundations in the expansion of the differential cross-section. For VBF with its three contributing Feynman diagrams at the tree level, it reads:

$$\frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}} = |A(\kappa_{2V}, \kappa_\lambda, \kappa_V)|^2 = |\kappa_V \kappa_\lambda M_s(m_{HH}) + \kappa_V^2 M_t(m_{HH}) + \kappa_{2V} M_x(m_{HH})|^2. \quad (6.2)$$

After the expansion, the expression has six terms meaning six coefficients, a_i , dependent on the m_{HH} as:

$$\begin{aligned} \frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}} &= \kappa_V^2 \kappa_\lambda^2 a_1(m_{HH}) + \kappa_V^4 a_2(m_{HH}) + \kappa_{2V}^2 a_3(m_{HH}) \\ &+ \kappa_V^3 \kappa_\lambda a_4(m_{HH}) + \kappa_V \kappa_\lambda \kappa_{2V} a_5(m_{HH}) + \kappa_V^2 \kappa_{2V} a_6(m_{HH}). \end{aligned} \quad (6.3)$$

Because of this dependence, the analytical evaluation of the coefficients is very difficult. Instead, the six basis samples of the three couplings (κ_{2V} , κ_λ , κ_V) are chosen. Each is run through the analysis reconstruction and selection to finally give the m_{HH} distribution. A cross-section for each basis and for each bin is then evaluated.

The six samples are chosen using the optimization technique designed specifically for this analysis. It is based on the minimization of the unphysical, negative bins while producing m_{HH} distributions for a continuous combination of κ_λ and κ_{2V} . The result of this optimization gives the following six orthogonal basis samples: (κ_{2V} , κ_λ , κ_V): (1,1,1), (1.5,1,1), (1,2,1), (1,10,1), (1,1,0.5), (0,-5,0.5).

The total differential cross-section equation reads:

$$\begin{aligned}
& \frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}} = \\
& \left(\frac{68\kappa_{2V}^2}{135} - 4\kappa_{2V}\kappa_V^2 + \frac{20\kappa_{2V}\kappa_V\kappa_\lambda}{27} + \frac{772\kappa_V^4}{135} - \frac{56\kappa_V^3\kappa_\lambda}{27} + \frac{\kappa_V^2\kappa_\lambda^2}{9} \right) \times \frac{d\sigma}{dm_{HH}}(1, 1, 1) \\
& \quad + \left(-\frac{4\kappa_{2V}^2}{5} + 4\kappa_{2V}\kappa_V^2 - \frac{16\kappa_V^4}{5} \right) \times \frac{d\sigma}{dm_{HH}}\left(\frac{3}{2}, 1, 1\right) \\
& + \left(\frac{11\kappa_{2V}^2}{60} + \frac{\kappa_{2V}\kappa_V^2}{3} - \frac{19\kappa_{2V}\kappa_V\kappa_\lambda}{24} - \frac{53\kappa_V^4}{30} + \frac{13\kappa_V^3\kappa_\lambda}{6} - \frac{\kappa_V^2\kappa_\lambda^2}{8} \right) \times \frac{d\sigma}{dm_{HH}}(1, 2, 1) \\
& \quad + \left(-\frac{11\kappa_{2V}^2}{540} + \frac{11\kappa_{2V}\kappa_V\kappa_\lambda}{216} + \frac{13\kappa_V^4}{270} - \frac{5\kappa_V^3\kappa_\lambda}{54} + \frac{\kappa_V^2\kappa_\lambda^2}{72} \right) \times \frac{d\sigma}{dm_{HH}}(1, 10, 1) \\
& + \left(\frac{88\kappa_{2V}^2}{45} - \frac{16\kappa_{2V}\kappa_V^2}{3} + \frac{4\kappa_{2V}\kappa_V\kappa_\lambda}{9} + \frac{152\kappa_V^4}{45} - \frac{4\kappa_V^3\kappa_\lambda}{9} \right) \times \frac{d\sigma}{dm_{HH}}\left(1, 1, \frac{1}{2}\right) \\
& \quad + \left(\frac{8\kappa_{2V}^2}{45} - \frac{4\kappa_{2V}\kappa_V\kappa_\lambda}{9} - \frac{8\kappa_V^4}{45} + \frac{4\kappa_V^3\kappa_\lambda}{9} \right) \times \frac{d\sigma}{dm_{HH}}\left(1, -5, \frac{1}{2}\right),
\end{aligned} \tag{6.4}$$

giving the prescription for how to obtain any specific value of the κ_{2V} coupling required for the final coupling scan.

6.3 Triggers

Triggers are used in order to select interesting events for the analysis. Because of the b-jet only final state, the multi-b-jet triggers are chosen. A simple approach would be to only use one trigger, but using multiple triggers, with each working better in a particular phase-space, enhances the event acceptance. Such enhancement is especially important in the case of a statistically deficient VBF channel.

The data-driven background estimation is based on the assumption that the 2b data is kinematically similar to the 4b data and the 2b data is used for the 4b background estimation in the signal region. Therefore, the chosen triggers require a maximum of two b-jets. With additional requirements on the number of other jets, two categories are proposed, namely, $2b1j$ and $2b2j$. Table 6.2 presents the $2b2j$ and $2b1j$ triggers for 2016, 2017 and 2018 data taking years. Jets are required to pass the minimum transverse energy, E_T , thresholds. For the $2b1j$ the threshold is 100 or 150 GeV (varies for different years of data collection) required on the non-b-tagged jet of the $2b1j$ trigger, while for the $2b2j$ trigger, the thresholds are 35 GeV for all four jets, including the b-jets.

Table 6.2: Names and types of triggers used in each year of data taking.

Year	Name	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

6.3.1 *b*-jet Trigger Efficiency Correction

Online to Offline b-tagging correction

The *b*-jet triggers work based on the online *b*-tagging applied at the High Level Trigger. The online *b*-tagging algorithms are the same as the offline *b*-tagging algorithms, described in Section 5.2.1, but they work with the lower-quality tracks and jets.

Only the reconstructed jets that are matched to the online trigger-level jets are considered in the analysis¹. These jets have to be both online and offline *b*-tagged. The two tagging efficiencies vary and the online to offline *b*-tagging correction is calculated. Furthermore, the performance of both online and offline *b*-tagging varies in the MC and the real data due to the detector discrepancies not accounted for in the MC. Therefore, the overall correction takes into account both the online-to-offline correction and the MC-to-data correction. The correction is applied to the MC efficiency as a scale factor.

For each jet, the online trigger efficiency conditioned on the offline efficiency for the D11r *b*-tagging at the 77% working point is evaluated and denoted as $\epsilon_b^{\text{Online|Offline}}$.

If the jet is both online and offline *b*-tagged, the combined flavour tagging efficiency is:

$$\epsilon_b^{\text{Online}\cap\text{Offline}} = \epsilon_b^{\text{Online|Offline}} \epsilon_b^{\text{Offline}}. \quad (6.5)$$

If the jet is not both online and offline *b*-tagged, the efficiencies are: for passing offline but failing online:

$$\epsilon_b^{\text{Online}\cap\text{Offline}} = (1 - \epsilon_b^{\text{Online|Offline}}) \epsilon_b^{\text{Offline}}, \quad (6.6)$$

for passing online and failing offline:

$$\epsilon_b^{\text{Online}\cap\text{Offline}} = \epsilon_b^{\text{Online}} - \epsilon_b^{\text{Online|Offline}} \cdot \epsilon_b^{\text{Offline}}, \quad (6.7)$$

¹For details of the jet matching procedure see Chapter 4 in [6].

for failing both online and offline:

$$\epsilon_b^{\text{Online} \cap \text{Offline}} = 1 - \epsilon_b^{\text{Offline}} - \epsilon_b^{\text{Online}} + \epsilon_b^{\text{Online}|\text{Offline}} \cdot \epsilon_b^{\text{Offline}}. \quad (6.8)$$

The efficiencies are evaluated as functions of p_T . Values of $\epsilon_b^{\text{Offline}}$, $\epsilon_b^{\text{Online}}$ and $\epsilon_b^{\text{Online}|\text{Offline}}$ for the MC jets and corresponding per-jet MC-to-data scale factors are provided by the ATLAS b-jet Trigger Flavour Tagging Group as CDI files [56].

As shown in Eq. 6.7 and 6.8, the online jet kinematics and flavour tagging efficiencies cannot be evaluated separately, the combined MC-to-data scale factors for the online and the offline b-tagging are calculated as:

$$\text{SF}_{\text{FTAG}}^{\text{per-jet}} = \frac{\epsilon_{\text{Data}}}{\epsilon_{\text{MC}}}, \quad (6.9)$$

where ϵ_{MC} are evaluated by plugging in the efficiencies values provided in the CDI files into one of the eq. 6.5 - 6.8 and ϵ_{Data} are evaluated by multiplying the CDI efficiencies values by the CDI SF values. For events with more than one jet, the product of per-jet scale factors is evaluated to give the per-event scale factor.

Online Jet Kinematic Correction

Online jet kinematic conditions also differ between MC and data. A scale factor for each reconstructed jet is derived specifically for this analysis to obtain the MC-to-data correction as a function of the matched offline jet p_T .

Firstly, each of the multi-b-jet triggers is split into its constituent L1 and HLT trigger jets. For example, the 2017 trigger of the type $2b1j$ is:

HLT_j110_gsc150_boffperf_split_2j35_gsc55_bmv2c1070_split_L1J85_3J30,

which means it consists of three HLT jets:

- HLT_j110_gsc150_boffperf_split_L1J85
- HLT_j110_gsc150_j35_gsc55_boffperf_J30
- HLT_j35_gsc55_boffperf_J30,

for simplicity L1 jets are not considered here.

Figure 6.1 shows the scale factors as the ratio of efficiency in data to the efficiency in MC for each of the above HLT jets as a function of the matched offline jet p_T for the 2017 $2b1j$ trigger. For the leading jet with $p_T > 170$ GeV, the second leading jet with $p_T > 60$ and the third leading jet with $p_T > 70$ GeV, the scale factors are close to 1 and the corresponding total uncertainty is not larger than 5%.

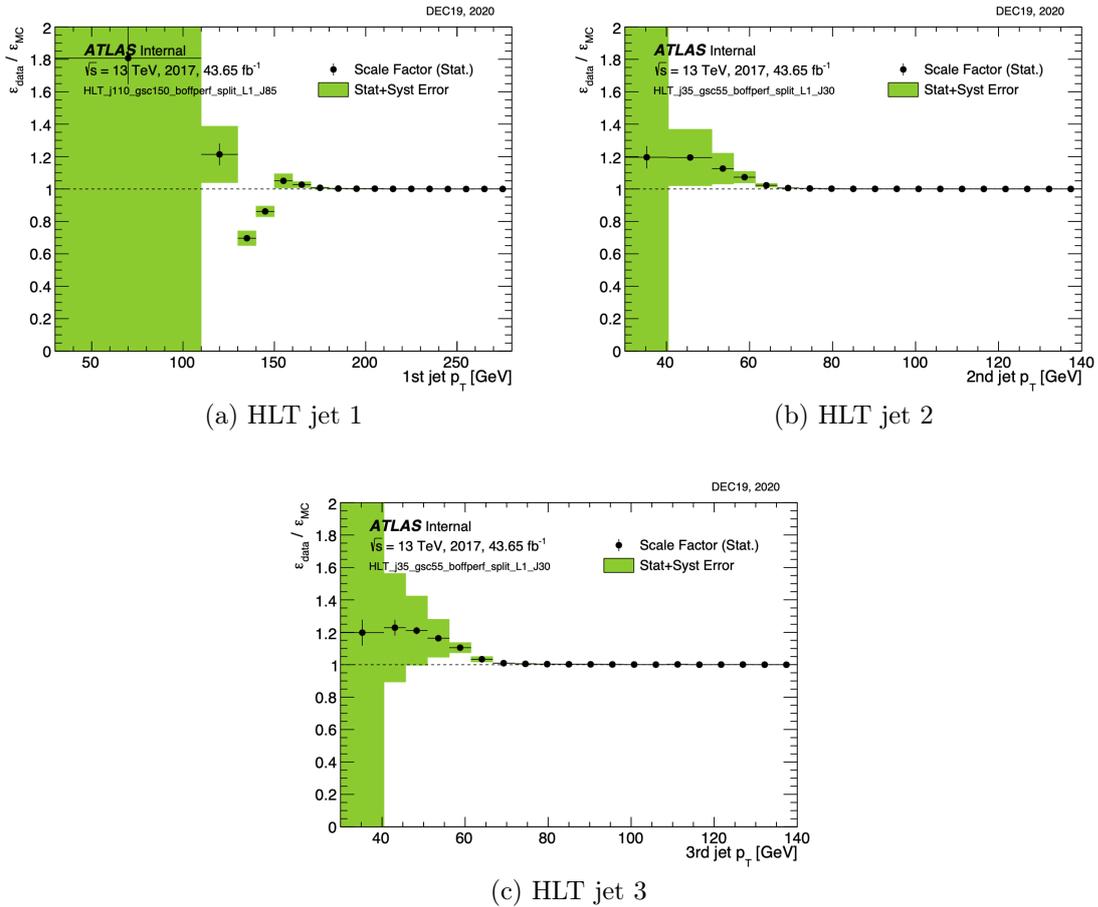


Figure 6.1: 2017 $2b1j$ trigger scale factors for the three leading online HLT jets measured as a function of the matched offline jet p_T . Statistical uncertainties are shown as the black bars, while the total statistical and systematic uncertainty is shown as the green bands.

Since scale factors are derived as functions of offline jet p_T , they can be combined with the total flavour tagging scale factors discussed in Section 6.3.1. As a result, a single trigger scale factor is applied to the MC events. More details on the method can be found in Appendix H of [6].

6.3.2 Trigger Bucket Algorithm

If an event passes multiple triggers, a combination of per-jet scale factors into a single per-event scale factor becomes very complicated. In order to avoid the complication, a technique called *trigger buckets* is used so that every event is assigned to exactly one trigger.

In Figure 6.2 the flowchart of the trigger buckets algorithm is shown. The algorithm is the offline and hierarchical algorithm which provides the selection steps in order to split events into two categories. Each category directly matches one of the two online triggers: $2b1j$ or $2b2j$. Only if an event is discarded from the first category, the $2b1j$, it can be tried for the $2b2j$ category. After events are placed in the buckets, it is checked whether they also pass the nominal, online trigger selection of the same type as the bucket they are in. Only if this is a positive decision, events are selected.

The first category mimics the $2b1j$ trigger. Its cuts are defined for the first leading jet ($p_T > 170$ GeV) and for the third leading jet ($p_T > 70$ GeV). The values of the cuts match the minimum p_T values for which the online jet kinematic scale factors are close to 1, as shown in Figure 6.1. If an event does not pass these requirements, it is placed in the second bucket which collects $2b2j$ -like events.

Figure 6.3 shows the trigger buckets composition for the $\kappa_{2V} = 0$ signal as a function of the di-Higgs mass. Bucket 1 represents the $2b1j$ trigger, which is dominant for high m_{HH} events, and Bucket 2 represents the $2b2j$ trigger, dominant for the low m_{HH} events. 38.9% of events are placed in $2b1j$ bucket, 54.5% in $2b2j$, the remaining events are discarded. More details on the trigger bucket algorithm studies can be found in Chapter 9 of [68].

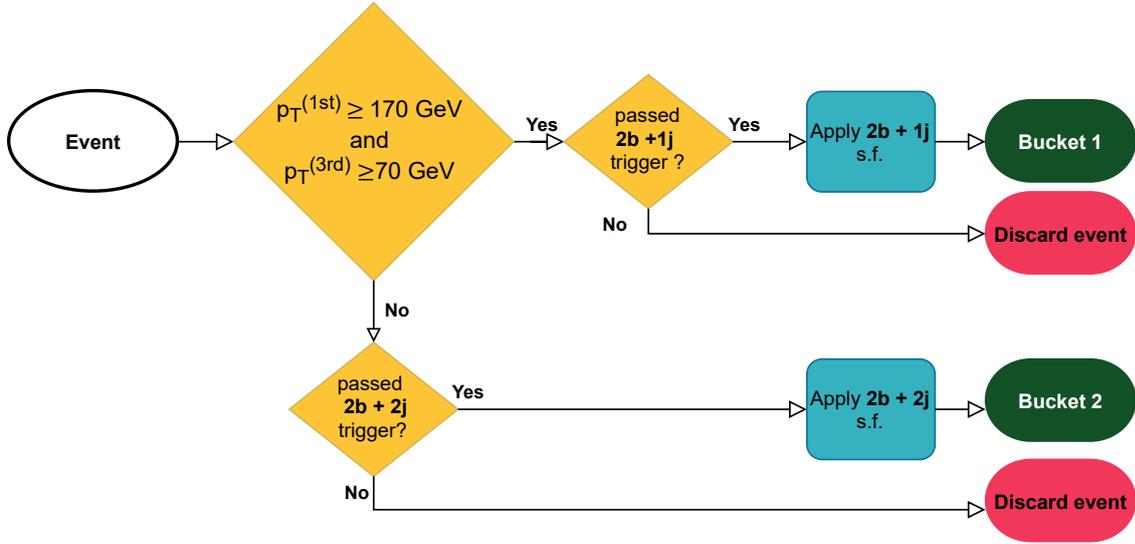


Figure 6.2: Flowchart illustrating the trigger bucket algorithm used in this analysis. Only if an event fails the first bucket, it can be tried for the second bucket. Each event needs to pass the corresponding online trigger requirements as well. The corresponding $2b2j$ or $2b1j$ event-level scale factors are applied to each event. Figure taken from [68].

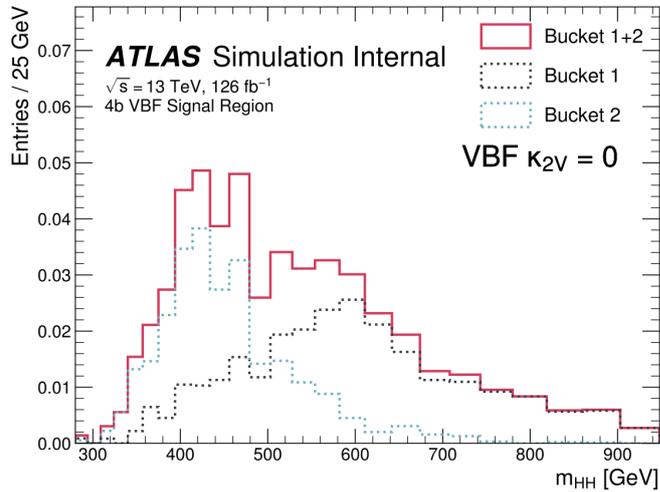


Figure 6.3: Buckets composition in the 4b signal region for the $\kappa_{2V} = 0$ signal as function of m_{HH} . Bucket 1 corresponds to $2b1j$ trigger and bucket 2 to $2b2j$ trigger [68].

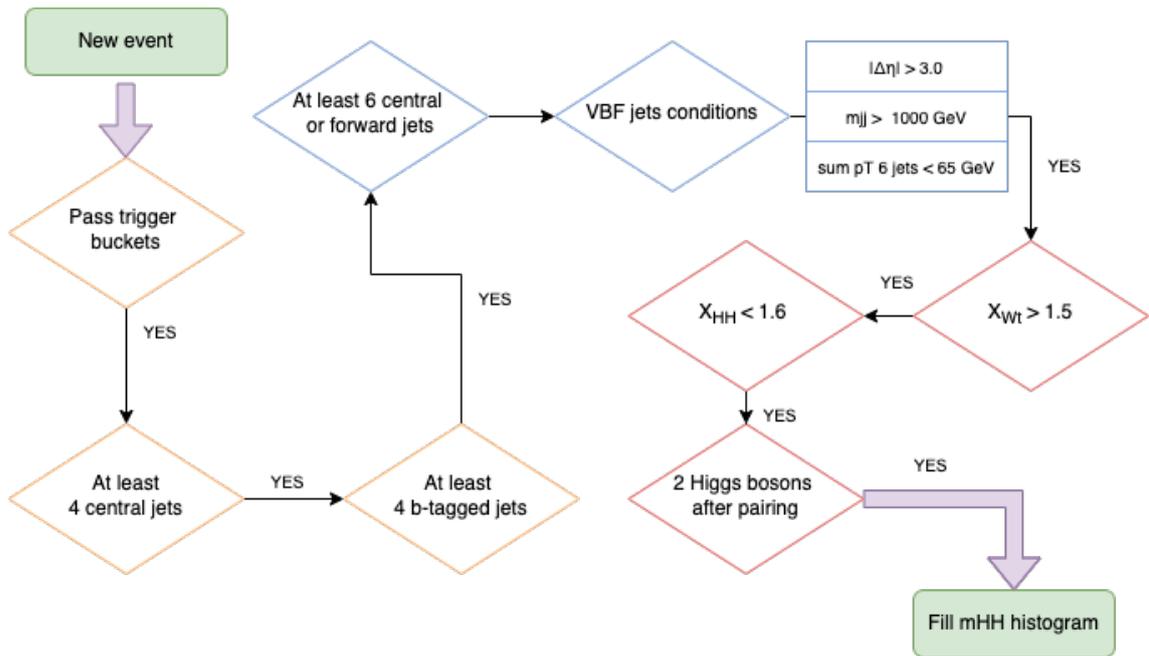


Figure 6.4: Flow-chart of all selections that need to be passed by an event to be considered for the VBF analysis.

6.4 VBF Event Selection

6.4.1 Nomenclature of the Analysis Jets

First of all, central and forward jets are defined. Kinematically, central jets must fulfil $p_T > 40 \text{ GeV}$ and $|\eta| < 2.5$ requirements. Central jets, if b-tagged, are used for reconstructing the two Higgs Candidates in the event. The η requirement is motivated by the b-tagging algorithm, which only works in the region covered by the tracker, as described in Chapter 3. The choice of the p_T threshold is motivated by the selection of the b-jet trigger. Secondly, the forward jets are defined as those with $p_T > 30 \text{ GeV}$ and $|\eta| > 2.5$. The presence of the two well-separated, forward jets is the crucial aspect of the VBF topology, as described in Section 2.2.4.

6.4.2 Steps of the Event Selection

All events considered here are the events that pass the triggering requirements and the trigger buckets algorithm. For a better overview of all selections, a flowchart showing VBF analysis steps is presented in Figure 6.4, while all selections are described in the text below.

The nominal signal events need to have at least four b-tagged central jets from which the Higgs Candidates can be built. This would be a sufficient selection for

the gluon-gluon fusion, however, since the VBF production is the focus here, the three additional VBF-specific conditions must be fulfilled. Namely, events must have at least six central or forward jets. Since at least four of those are b-tagged, the VBF jets are defined as the pair of the non-b-tagged jets with the highest m_{jj} . The events that pass the next selections must have the VBF jets with $m_{jj} > 1000$ GeV and $|\Delta\eta_{jj}| > 3$. The VBF events are also characterized by a significantly suppressed hadronic activity in the central region. In order to quantify such a suppression, the vector sum of six jets p_T (four Higgs Candidate jets and two VBF jets) is evaluated with the requirement of the total $p_T < 65$ GeV.

2 b-tagged Region

Additionally, for the purpose of the data-driven background estimation, the 2b category is defined. Events are selected with the same selection steps as the 4b events, with the only difference being the requirement that exactly two jets are b-tagged instead of at least four. The two remaining jets needed for Higgs Candidates construction are two, non-tagged highest p_T central jets.

6.5 Higgs Candidate Pairing

After the event selection, the four central b-tagged jets are paired into Higgs Candidates. Figure 6.5 illustrates the three possibilities of such pairings. Choosing the correct pairing is crucial because the further analysis cuts are dependent on the Higgs Candidates' properties.

6.5.1 Overview of Considered Algorithms

The ways Higgs Candidates can be paired are numerous. To assess if the pairing in data is correct, the Monte Carlo samples are used for comparison with the reconstructed pairs. The pairing algorithm which results in a smooth di-Higgs massplane is preferred for the background estimation reasons. A smooth massplane means that no events with a specific mass of leading or sub-leading Higgs Candidates are preferred. Pairing is performed on all events that pass the selection to this point, including the 2b events.

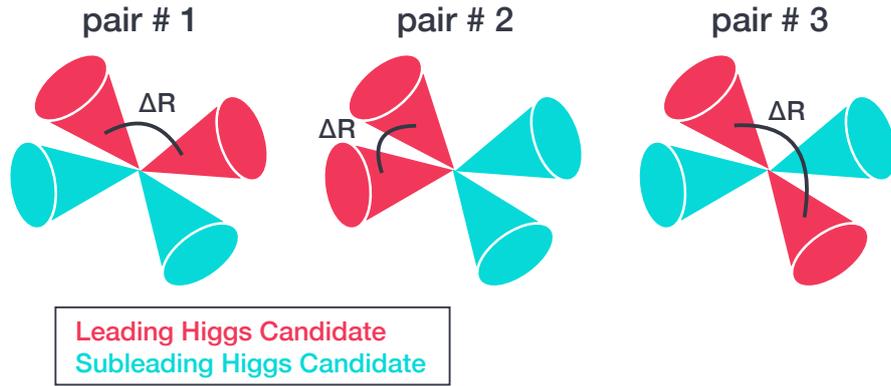


Figure 6.5: Three possibilities of pairing four jets into two Higgs candidates. The angular separations between jets of the leading Higgs candidates are indicated. In this schematic, the second pair is the one with the smallest ΔR hence it would be chosen as the Higgs pair for this event based on the minimum ΔR pairing algorithm.

D_{HH} Pairing

The algorithm that was used in the previous VBF analysis [3] is aimed at minimizing the quantity D_{HH} given by equation:

$$D_{HH} = \frac{|m_{H1} - \frac{120}{110}m_{H2}|}{\sqrt{1 + (\frac{120}{110})^2}} \quad (6.10)$$

where the centre of the signal region is at $m_{H1} = 120$ GeV, $m_{H2} = 110$ GeV. The D_{HH} measures the distance between the reconstructed Higgs masses of each of the three possible Higgs pairings, and the line crossing the origin and the centre of the m_{H1} and m_{H2} massplane. Even though the efficiency of such pairing compared to the true Higgs bosons is high (83% of the signal events were correctly paired in the previous VBF analysis [3]), the algorithm leads to the undesired shaping of the massplane; pairs laying nearby the central diagonal line are preferred. Therefore, a new algorithm, in which massplane sculpting is avoided, is proposed.

Minimum ΔR Pairing

The new pairing strategy is based on choosing the leading p_T Higgs Candidate in each of the three possible pairs. The rapidity separation, $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, is

calculated for the three leading Higgs Candidates. The pairing in which the leading HC has the smallest ΔR is chosen.

Figure 6.6 shows the resulting massplanes for the different samples: the 4b data, the 2b data and the $\kappa_{2V} = 0$ signal Monte Carlo. No peaks in the Higgs Candidates' massplanes are observed. Figure 6.7 shows the pairing accuracy of the truth di-Higgs mass for the VBF SM signal, the $\kappa_{2V} = 0$ signal and the $\kappa_\lambda = 10$ signal. For $m_{HH} > 450$ GeV, the accuracy is at least 70% for the SM signal, while it is at least 80% for the non-SM signals. The harder p_T spectrum of the non-SM signals at high m_{HH} results in pairs of jets that are more collimated, leading to the higher pairing accuracy. At low m_{HH} masses and for the SM point, the p_T spectrum is softer. The degradation of the accuracy at low m_{HH} is acceptable because the analysis is aimed at exploring the high m_{HH} events, where the QCD background is reduced.

Because of the smooth massplane and sufficiently high pairing accuracy, the minimum ΔR pairing method is chosen for the analysis.

6.6 Top veto for the Background Reduction

The multijet QCD processes are the dominant background of this analysis, while events in which pairs of top quarks, $t\bar{t}$, decay hadronically as $t \rightarrow b(W \rightarrow q_1\bar{q}_2)$ constitute the second most dominant background. The method of reducing the number of $t\bar{t}$ events is presented.

The possible W boson candidates, to which t quarks decay, are constructed as the combination of any two central jets that are neither Higgs Candidates nor VBF jets. From combinatorics, given there are n jets, there will be $\binom{n}{2}$ W candidates. Then, to create the hadronic t candidates, each of the possible W candidates has to be paired with a b-jet. There are exactly four b-jets in the event, hence there are $4\binom{n}{2}$ possible t candidates².

Subsequently, we introduce the top veto variable X_{Wt} :

$$X_{Wt} = \sqrt{\left(\frac{10(m_W - 80.4 \text{ GeV})}{m_W}\right)^2 + \left(\frac{10(m_t - 172.5 \text{ GeV})}{m_t}\right)^2}, \quad (6.11)$$

where m_W and m_t are the invariant masses of the W and t candidates respectively.

For each event, the X_{Wt} for $4\binom{n}{2}$ combinations is evaluated and the event is vetoed if the minimum X_{Wt} from among all combinations is < 1.5 .

²The Higgs Candidate b-jets and the jets constructing W boson must be distinct.

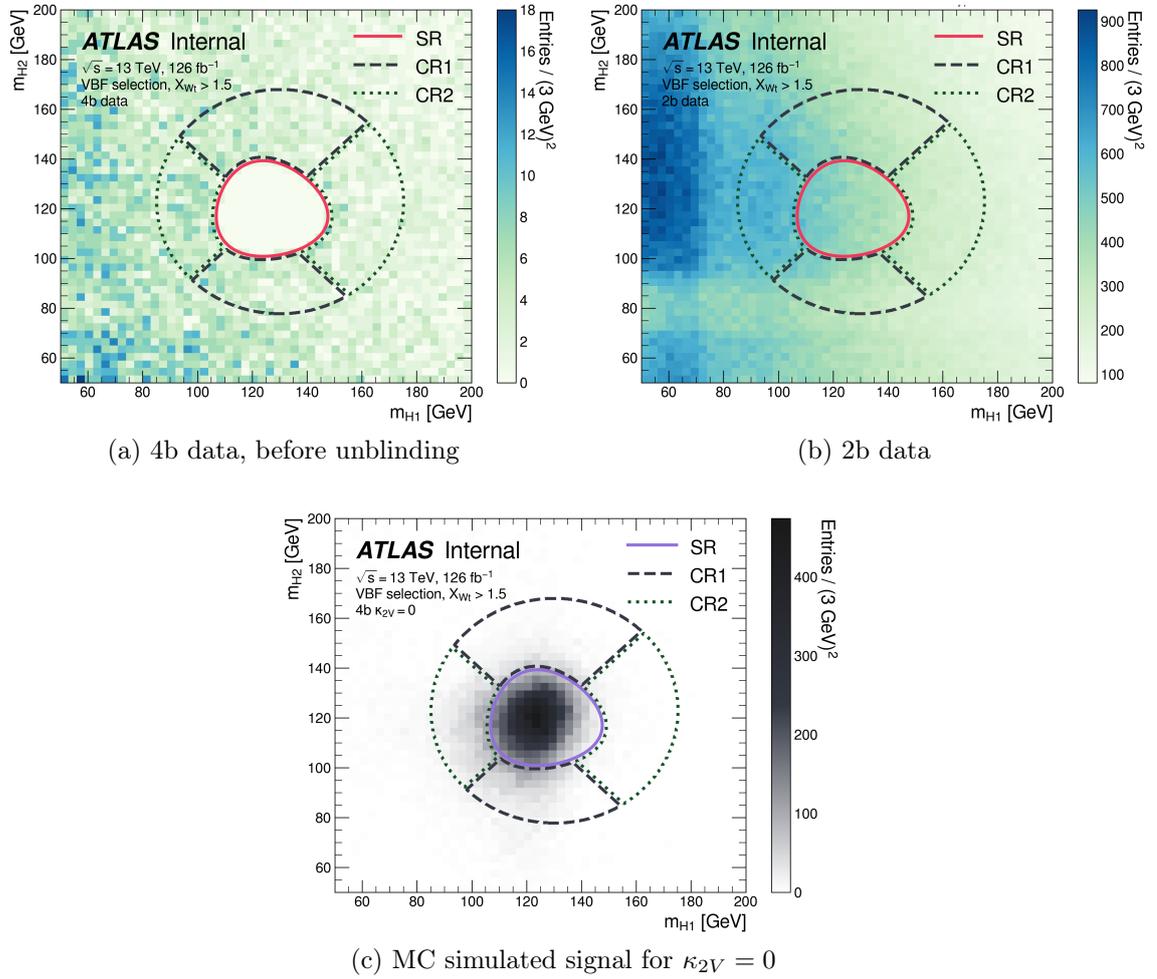


Figure 6.6: The m_{H_1} and m_{H_2} massplanes for the VBF selection for the 4b data (pre-unblinding), the 2b data used for the background estimation and the simulated signal for the $\kappa_{2V} = 0$. The kinematic regions are defined in Section 6.7. Signal and two control regions CR1 and CR2, defined in Section 6.7, are overlaid on each massplane.

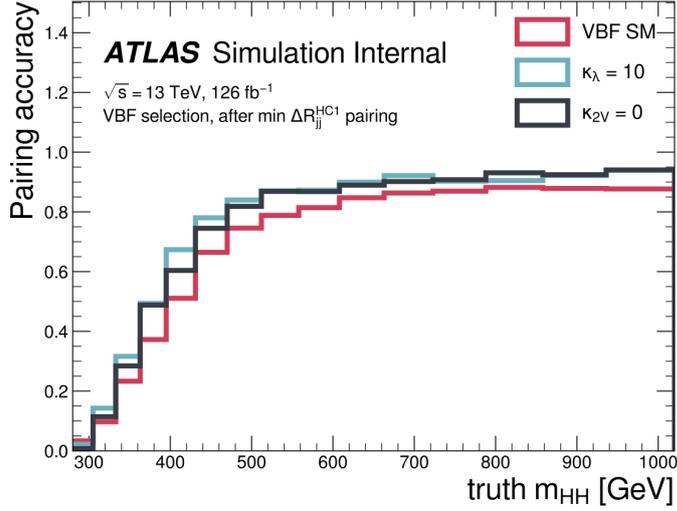


Figure 6.7: ΔR pairing accuracy as a function of the truth m_{HH} distribution for three VBF signals: VBF SM, $\kappa_{2V} = 0$ and $\kappa_{\lambda} = 10$.

6.7 Kinematic Regions Definitions

The kinematic signal (SR) and control (CR) regions are defined in the di-Higgs massplane. The signal region is defined by requiring $X_{HH} < 1.6$, where:

$$X_{HH} = \sqrt{\left(\frac{10(m(H_1) - 124 \text{ GeV})}{m(H_1)}\right)^2 + \left(\frac{10(m(H_2) - 117 \text{ GeV})}{m(H_2)}\right)^2}, \quad (6.12)$$

with $m(H_1)$ and $m(H_2)$ representing masses of leading and subleading Higgs Candidates respectively. The centre of the massplane is at $m(H_1) = 124 \text{ GeV}$, $m(H_2) = 117 \text{ GeV}$. During the Higgs Candidate pairing, there is an unavoidable energy loss due to the semi-leptonic B-decays resulting in the two Higgs candidates of unequal masses. 124 GeV and 117 GeV are chosen for consistency with the ggF analysis for which this is an optimal signal region centre [6].

Control regions are defined for the background estimation procedure, which is based on deriving scaling factors in the control regions and applying them to the events in the signal region. To ensure a smooth extrapolation of the scaling factors, the control region is bordering the signal region to preserve the two regions' kinematic resemblance. The outer boundary is selected such that there are enough statistics to perform background estimation. The outer boundary equation reads:

$$\sqrt{((m(H_1) - 1.05 \times 124 \text{ GeV})^2 + ((m(H_2) - 1.05 \times 117 \text{ GeV})^2 < 45 \text{ GeV}. \quad (6.13)$$

The centre of the control region is shifted by 5% with respect to the centre of the signal region in order to ensure that the event statistics are similar all around the

SR while the very low Higgs mass events are avoided. For the low mass events, the differences between the 2b and the 4b events are large, hence undesired for the background estimation procedure.

Control Region Splitting

In order to better account for the systematic uncertainty for the background estimation, the control region is split into two parts. Two separate background models are then derived, and the difference between them constitutes the background shape systematic uncertainty. For illustration, we refer to Figure 6.6b, which shows the 2b di-Higgs massplane, while the boundaries of the signal and two control regions, CR1 and CR2, are overlaid. Each control region consists of two quadrants lying on opposite sides of the signal region.

The lines splitting the control region are chosen such that they cross the centre of the signal region at the angles 45° and 135° measured from the positive x-axis. In order to refer to each quadrant a compass-like naming scheme is applied: the top quadrant is *N*-orth and following in the clockwise direction these are *E*-ast, *S*-outh and *W*-est.

To motivate such a choice of control region splitting, note that in Figure 6.6b at $70 < m_{H_2} < 90$ GeV, a horizontal lighter coloured strip is visible, similarly, a vertical strip for the $70 < m_{H_1} < 90$ GeV. This is the result of introducing the top veto variable, which in practice removes Higgs Candidates of approximately the *W* boson mass. To account for these horizontal and vertical reductions in a number of events with Higgs Candidate masses of around 80 GeV, the two control regions are defined as CR1: vertically, (N, S), and CR2: horizontally, (W, E), around the signal region. This way, the deficit in the number of events is the same for CR1 and CR2.

6.8 Cut-flows

Table 6.3 shows the number of the 2b and the 4b events passing each of the selection steps. The numbers are quoted for all years of data taking inclusively (2016, 2017 and 2018). The only requirement that differs for both samples is the number of the b-tagged jets. For the 4b events, ≥ 4 b-tags are required, while for the 2b events exactly 2 b-tags are required. It can be observed that out of the initially available 5.70×10^8 events, only 502 4b events and 58 700 2b events are selected in the signal region.

Table 6.3: Cut-flow table indicating the number of 4b and 2b events passing each of the described selection steps. 4b signal region data are unblinded.

	4b data	2b data
Initial Yield	5.70×10^8	
Trigger	2.81×10^8	
Trigger Buckets	2.49×10^8	
VBF Selections	3.30×10^6	
≥ 4 central jets, ≥ 2 b-tags	3.16×10^6	
≥ 4 b-tags (4b) or exactly 2 btags (2b)	2.71×10^4	2.76×10^6
Top Veto	2.18×10^4	2.47×10^6
Control Region 1	9.47×10^2	1.11×10^5
Control Region 2	9.06×10^2	1.10×10^5
Signal Region (post-unblinding)	5.02×10^2	5.87×10^4

6.9 Acceptance times Efficiency

Figure 6.8 shows the cumulative acceptance times efficiency³ as a function of κ_{2V} for the MC simulated samples for the selection steps described in this chapter. The final acceptance times efficiency is the worst for the SM point. At the SM point cancellation due to the perturbative unitarity between the Feynman diagrams, shown in Figures 2.12a and 2.12c in Chapter 2, is observed [25]. For the points away from the SM, the non-cancellation effects result in a harder m_{HH} spectrum. This leads to an increased signal acceptance times efficiency at the non-SM points, while the prediction around the SM region remains unimproved.

6.10 Discriminating Variable

Following [69], the full characterization of the di-Higgs system is provided by the di-Higgs mass variable, m_{HH} . The m_{HH} is the discriminating variable of the analysis and it is complemented with the angular variable between two Higgs bosons, $\Delta\eta_{HH}$, for categorization.

Categorization

It is beneficial to split data into categories based on kinematic variables such that the signal to background significance, $\frac{S}{\sqrt{B}}$, is maximized in each category, resulting

³Number of events passing each selection divided by the total number of generated events

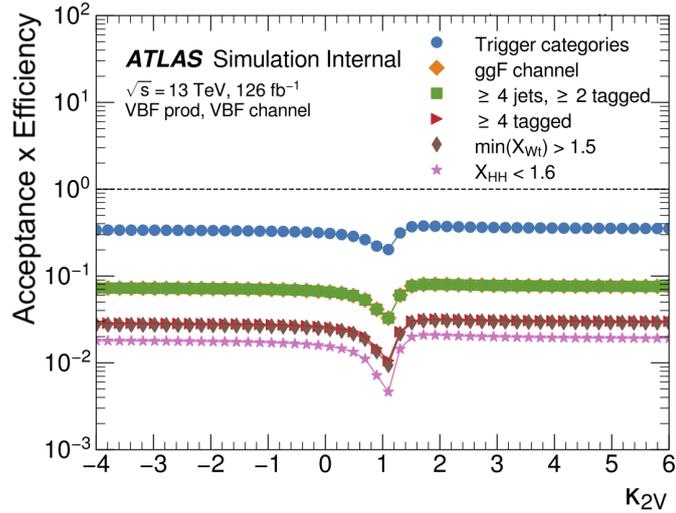


Figure 6.8: Cumulative acceptance times efficiency as a function of κ_{2V} for the MC simulated samples for the selection steps described in the chapter.

in a better overall signal significance. $|\Delta\eta_{HH}|$ is used for the categorization of events in the signal region.

Figure 6.9 shows the $|\Delta\eta_{HH}|$ distribution. The SM signal and the non-SM signals for $\kappa_{2V} = 0$ and $\kappa_{\lambda} = 10$, as well as the 2b reweighted data are shown. Signals are scaled by integers for visibility. The category split is chosen at $|\Delta\eta_{HH}| = 1.5$. $|\Delta\eta_{HH}| \leq 1.5$ is sensitive to the non-SM signals, while $|\Delta\eta_{HH}| > 1.5$ is sensitive to the SM signal.

Binning

Signal region events with $|\Delta\eta_{HH}| \leq 1.5$ and $|\Delta\eta_{HH}| > 1.5$ are separately binned in the m_{HH} histograms. Variable binning with bins gradually increasing in width is proposed, as the m_{HH} is a steeply falling distribution. The prescription for the consecutive bin edges reads:

$$bin_{low}^{i+1} = bin_{low}^i + (100 + x\%) \times bin_{low}^i, \quad (6.14)$$

where the first bin edge is chosen at 400 GeV. The $x\%$ also varies for $|\Delta\eta_{HH}|$ categories and it is 10% and 9% respectively for the low and high $|\Delta\eta_{HH}|$ categories. The binning algorithm iterates until the upper data limit is reached but it does not stop exactly at this limit; it calculates the last bin edge as programmed. The maximum data limits for the two categories are 890 GeV (low) and 1470 GeV (high). The choice of the specific percentage x is justified by posing the cap on the maximum statistical error at 30%. For the asymptotic approximation to work, 10 events are the minimum number of events per bin, so the Poisson error is $\sqrt{N} = \sqrt{10} \approx 3$ events, making it $\frac{3}{10} \approx 30\%$ the maximum relative error.

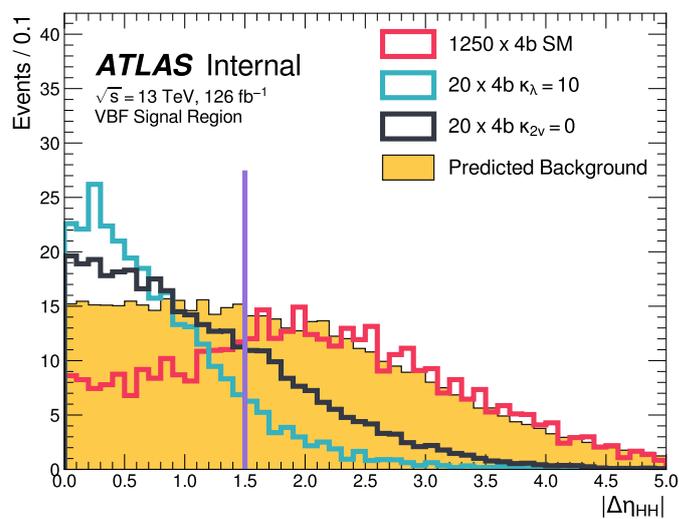


Figure 6.9: The $|\Delta\eta_{HH}|$ distribution in the 4b signal region of the 2b reweighted data and the signal shapes scaled by integers (for visibility). The cut at $|\Delta\eta_{HH}| = 1.5$ is shown to indicate the proposed category splitting.

Chapter 7

Background Estimation

Background estimation is the key component of the analysis. Because of the fully hadronic final state of the 4b channel and the multijet QCD background, the analysis uses a data-driven method instead of a Monte Carlo simulation. The details of the neural network reweighting technique are discussed in this chapter.

7.1 Method of Event Reweighting with Bins

In many cases, the background is estimated using the Monte Carlo method. Classic event reweighting [70] is a correction method for the imperfect MC background prediction obtained for a process when the simulated sample does not agree well with the real data or the real data are unknown. In order to reduce the disagreement, a new, but similar process, for which the MC and the real data are known, is introduced. The goal is to find a multiplying factor by which the imperfect MC distribution should be multiplied so that it is calibrated with the real data distribution. This calibration is the nominal reweighting.

Two distributions (MC and data) of a well-understood process are estimated using binned histograms of the kinematic variables describing the process. The multiplying factor m is calculated in the bin-by-bin fashion as:

$$m_{\text{bin}} = \frac{w_{\text{bin, data, NP}}}{w_{\text{bin, MC, NP}}}, \quad (7.1)$$

where $w_{\text{bin, data, NP}}$ and $w_{\text{bin, MC, NP}}$ are the total event weights in a given bin of either the real data or the MC kinematic variable distribution of the new process. Subsequently, factor m is applied, bin by bin, to the original, imperfect MC event weight to obtain the corrected MC event weight:

$$w_{\text{bin, MC, corrected}} = m_{\text{bin}} \times w_{\text{bin, MC, imperfect}}. \quad (7.2)$$

Usually, more than one variable is needed to calibrate the MC process and to perform the desired reweighting. The distributions of several variables are multiplied, which

leads to the problem of data scarcity. The *curse of dimensionality* occurs; the phase space increases while the available data remains unchanged, much scarcer in the enlarged space.

7.1.1 2b to 4b Event Reweighting

The idea of the event reweighting was modified and applied to the data-driven background estimation in the 4b ggF analysis in the 2015/2016 dataset analysis [29].

The 2b events, events with exactly two b-tagged jets, are selected on top of the 4b events, assuring no difference in the event selection other than the number of the b-tagged jets. The main assumption, verified with various kinematic distributions, is that the background events with exactly two b-jets are kinematically similar to the four b-jets events.

The desired background distribution is the 4b background in the signal region, but only the 2b background distribution in the signal region is known, the 2-tag events are unblinded while the 4-tag events are blinded. This is analogous to the imperfect MC from the general event reweighting method. The new processes, from which the scaling factor m is derived, are the 2b and 4b distributions in the control region, where both 4-tagged and 2-tagged events are unblinded. The control regions contain only the background events, the signal contamination is assumed to be negligible. In Figures 6.6a and 6.6b in Chapter 6, the 4b and 2b massplanes are shown with the overlaid signal and control regions boundaries¹. The multiplying factor m is obtained as:

$$m_{\text{bin}} = \frac{4\text{b distribution}_{\text{bin}}}{2\text{b distribution}_{\text{bin}}}. \quad (7.3)$$

The multiplier m is applied to the 2-tag events in the signal region resulting in the prediction of the 4b signal region background events.

Practically, to account for the kinematic differences between 2b and 4b events, ratios of binned histograms of various kinematic variables are used for reweighting, which leads to the mentioned *curse of dimensionality* problem. More details on this method can be found in [29] and [3]. In this analysis, the reweighting method for background estimation is retained, but a truly multivariate neural network approach is proposed.

¹Note that in the original analysis where the event reweighting method was used, different definitions of the kinematic control and validation regions were used but the fundamentals of the method do not differ. The original regions definitions can be found in [29].

7.2 2b to 4b Neural Network Reweighting

In order to obtain the multiplying factor, from now on referred to as a *weight*, using neural networks, the density ratio estimation of the two probability distributions is required. For the 4 and 2-tag data, the probability density functions: $p_{4b}(x)$ and $p_{2b}(x)$ across some input variables x are given. The aim of the neural network is to learn a reweighting function $w(x)$ between the 2 and 4-tag data such that:

$$p_{2b}(x) \cdot w(x) = p_{4b}(x) \quad (7.4)$$

from which it follows that:

$$w(x) = \frac{p_{4b}(x)}{p_{2b}(x)}. \quad (7.5)$$

This density ratio estimation can be treated using various multivariate methods. The method adapted for this analysis is based on [71]. Starting from the following loss function:

$$\mathcal{L}(R(x)) = \mathbb{E}_{x \sim p_{2b}} \left[\sqrt{R(x)} \right] + \mathbb{E}_{x \sim p_{4b}} \left[\frac{1}{\sqrt{R(x)}} \right]. \quad (7.6)$$

where, again, x represents some input variables in a vector form in \mathbb{R}^n , R is the estimator, while p_{4b} and p_{2b} are arbitrary, but fixed, probability density functions.

Using the definition of expectation value, we get:

$$\begin{aligned} \mathcal{L}(R) &= \int \sqrt{R(x)} p_{2b}(x) dx + \int \frac{1}{\sqrt{R(x)}} p_{4b}(x) dx \\ &= \int dx \left[\sqrt{R(x)} p_{2b}(x) + \frac{1}{\sqrt{R(x)}} p_{4b}(x) \right]. \end{aligned} \quad (7.7)$$

In order to minimise this loss, the estimator, R , has to be found. This results in a standard problem solvable by the calculus of variations. We start from:

$$\mathcal{I}(x_1, \dots, x_n, R, R_1, \dots, R_n) = \sqrt{R(x)} p_{2b}(x) + \frac{1}{\sqrt{R(x)}} p_{4b}(x), \quad (7.8)$$

with $R_i = \frac{\partial R}{\partial x_i}$. From the Euler-Lagrange equations, \mathcal{L} has an extreme value (minimum or maximum) only if R satisfies the partial differential equation:

$$\frac{\partial \mathcal{I}}{\partial R} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial \mathcal{I}}{\partial R_i} \right) = 0. \quad (7.9)$$

As \mathcal{I} does not depend on R_i for any $i \in \{1, \dots, n\}$, the second term vanishes. The extremizing estimator R is found after solving the following equation:

$$\begin{aligned} 0 = \frac{\partial \mathcal{I}}{\partial R} &= \frac{1}{2} R^{-\frac{1}{2}} p_{2b} - \frac{1}{2} R^{-\frac{3}{2}} p_{4b} \\ &= \frac{1}{2} R^{-\frac{1}{2}} (p_{2b} - p_{4b} R^{-1}), \end{aligned} \quad (7.10)$$

which is:

$$\arg \min_R \mathcal{L}(R) = \frac{p_{4b}(x)}{p_{2b}(x)}, \quad (7.11)$$

as desired. Note that the above only guarantees that R is an extremizing estimator, not necessarily a minimizing one. However, showing the latter property is outside of the scope of this thesis.

In practice, to avoid imposing explicit positivity constraints, the substitution $Q(x) \equiv \log R(x)$ is made. The form of the loss function then becomes

$$\mathcal{L}(R(x)) = \mathbb{E}_{x \sim p_{2b}} \left[\sqrt{e^{Q(x)}} \right] + \mathbb{E}_{x \sim p_{4b}} \left[\frac{1}{\sqrt{e^{Q(x)}}} \right], \quad (7.12)$$

with the solution

$$\arg \min_Q \mathcal{L}(Q(x)) = \log \frac{p_{4b}(x)}{p_{2b}(x)}. \quad (7.13)$$

Taking the exponent of the Eq. 7.13 results in the desired reweighting function.

7.3 Neural Network Reweighting in Practice

Figure 7.1 illustrates the neural network background estimation in practice, based on the 2b and 4b massplanes introduced with Figure 6.6. The top row illustrates the training stage in which the 2b data are mapped to the 4b data in the control regions and the network learns weights so that the differences between the kinematic distributions are minimized. The bottom row illustrates the background estimation stage in which the learned weights are applied to the 2b signal region events. As a result, the 4b background is predicted for the blinded 4b signal region.

7.3.1 Network Input Variables

An optimal set of input variables on which the network can efficiently learn the differences between the 2-tag and 4-tag distributions is required. The discriminating variable in this analysis is the di-Higgs mass distribution, the m_{HH} , for which background modelling is crucial. In order to avoid bias, the network is not directly

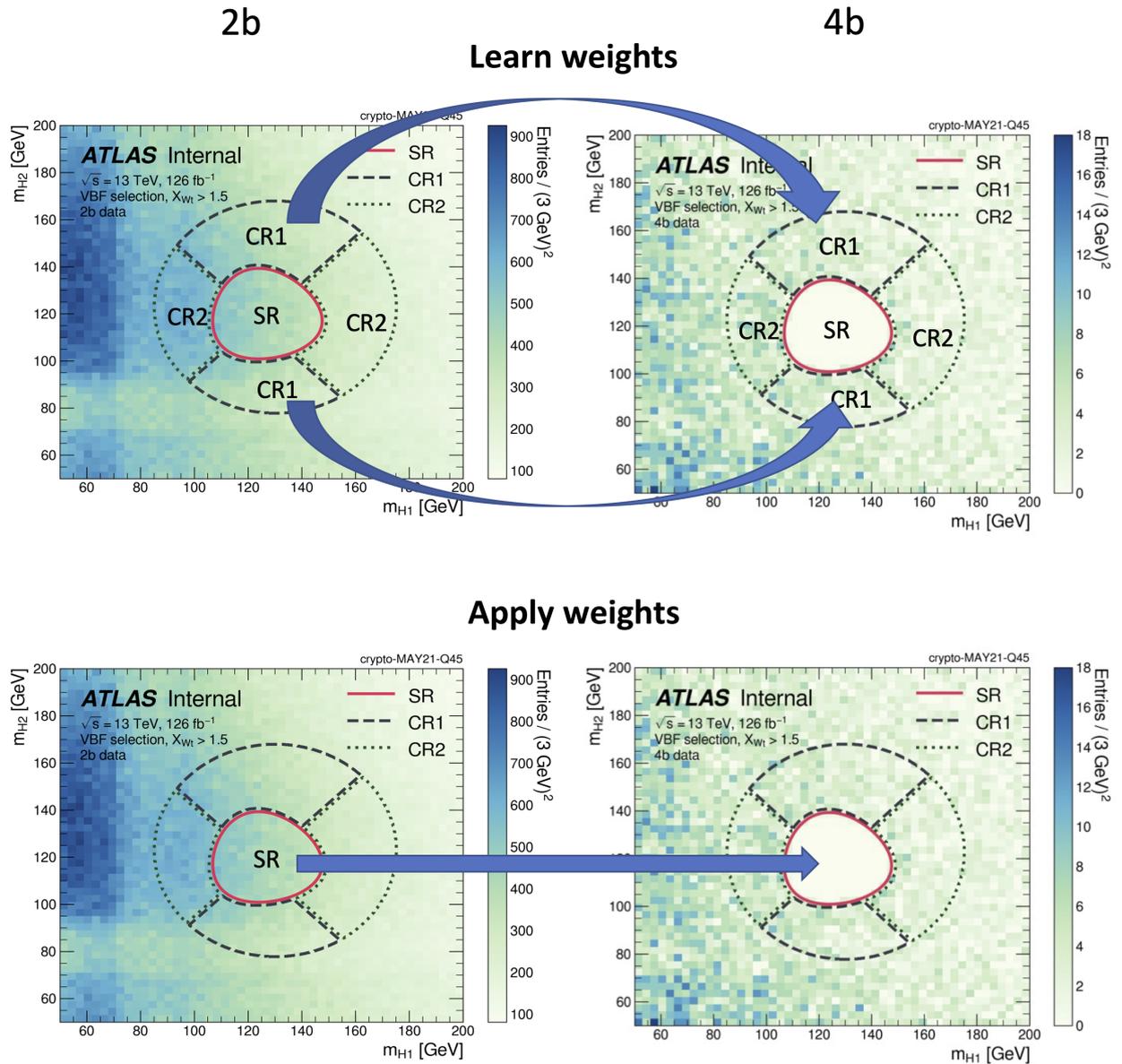


Figure 7.1: Illustration of the neural network reweighting approach based on the 2b and 4b di-Higgs massplanes. In the top row, the training stage of the procedure is shown; the network learns weights of the CR1 2b and 4b distributions (same is done in CR2). In the bottom row, the learned weights, are applied to the 2b events in the SR to obtain an estimation of the signal region 4b background events.

trained on the m_{HH} variable, but instead on a set of variables chosen such that they describe the m_{HH} distribution well by being closely correlated to the m_{HH} itself.

The correlations are checked using the Pearson Correlation Coefficients method [72], defined in the following equation:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (7.14)$$

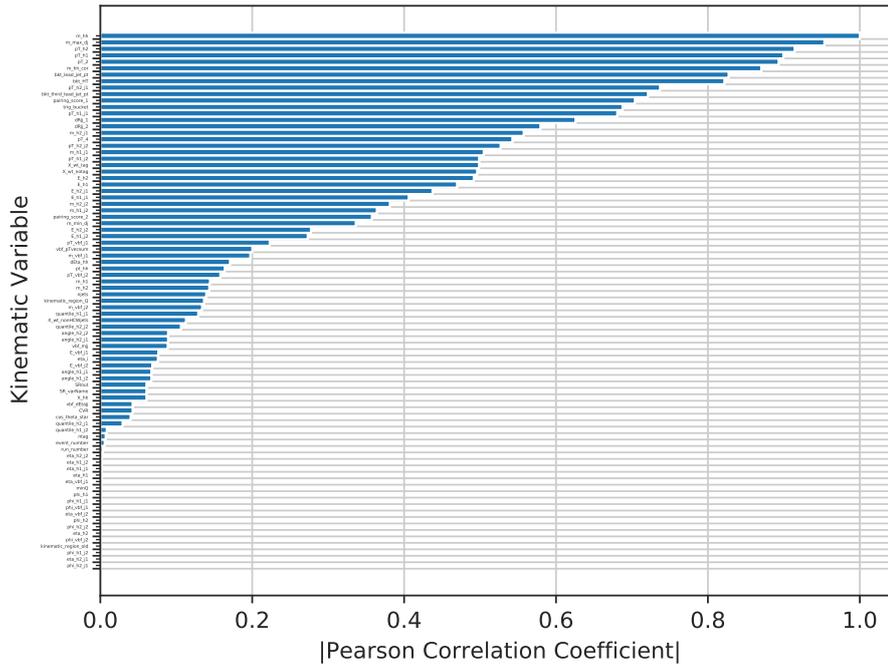
where r is the correlation coefficient, x_i and y_i are the values of the x and y variable respectively, while \bar{x} and \bar{y} are their means. In the analysis, each of the considered variables, x , is compared to the discriminating variable m_{HH} , y .

The coefficients are calculated for the full 2b control region dataset split into two categories in the $|\Delta\eta_{HH}|$ variable to assure background modelling is correct for both categories. The results of the coefficients calculation are presented in Figure 7.2.

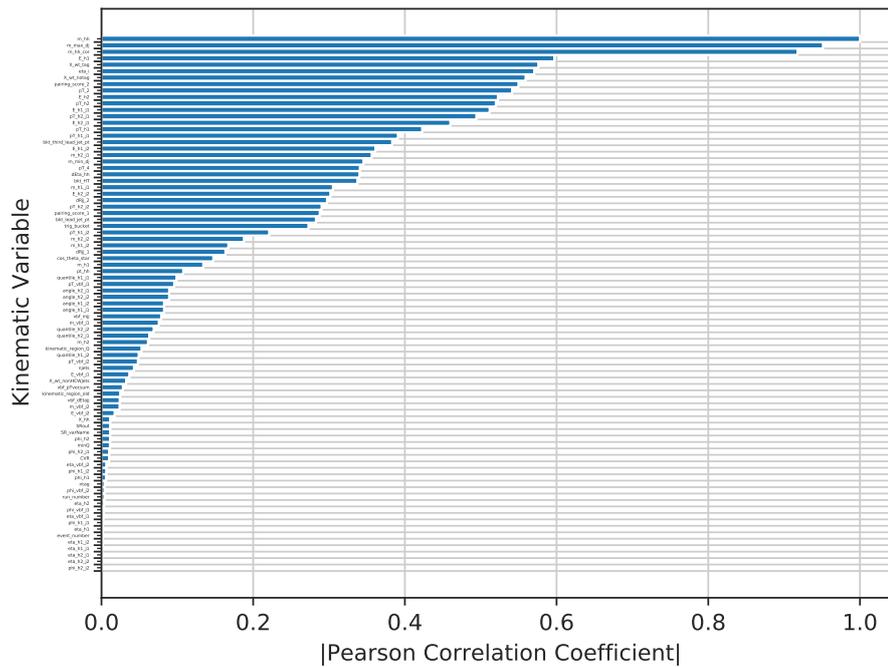
The variables, for which the correlation coefficient is high (above 0.5) in both categories, are chosen. The second test is then performed in order to reduce the set of input variables to the order of 10. The method is again the Pearson Correlation Coefficient method, but now the correlations among variables from the set are measured. Figure 7.3 shows the resulting 2D correlation matrix. Seven variables that are least correlated with each other are chosen. This way the redundant information that is carried by the closely correlated variables is avoided.

The list of the network input variables is presented below. The variable names displayed in the Figures are given in brackets. Additionally, two *one hot encoder* variables are added. These are related to the encoded trigger bucket information and the year in which the given event was recorded, even though the training is based on the entire dataset, where information from mixed years of data collection and all available triggers is included. The list is the following:

1. Maximum di-jet mass out of the possible pairings of the four Higgs candidate jets (m_max_dj),
2. Minimum di-jet mass out of the possible pairings of the four Higgs candidate jets (m_min_dj),
3. Energy of the leading Higgs candidate (E_h1),
4. Energy of the subleading Higgs candidate (E_h2),
5. Second smallest ΔR between the jets in the leading Higgs candidate (out of the three possible pairings for the leading Higgs candidate) (pairing_score_2),
6. $\log(X_{Wt})$, where X_{Wt} is the variable used for the top veto (X_wt_tag),
7. Average absolute value of Higgs candidate jet η (eta_i),
8. Trigger bucket index (for the trigger buckets inclusive training) as one hot encoder,
9. Year index (for the years inclusive training) as one hot encoder.



(a) $|\Delta\eta_{HH}| \leq 1.5$



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

Figure 7.2: The display of the Pearson Correlation Coefficients for the kinematic variables of the 2b Control Region 1 data correlated with the m_{HH} and evaluated in two $|\Delta\eta_{HH}|$ categories. The variables are shown in the descending order of the coefficient value r with the m_{HH} being correlated to itself at the top, with the r value of 1.

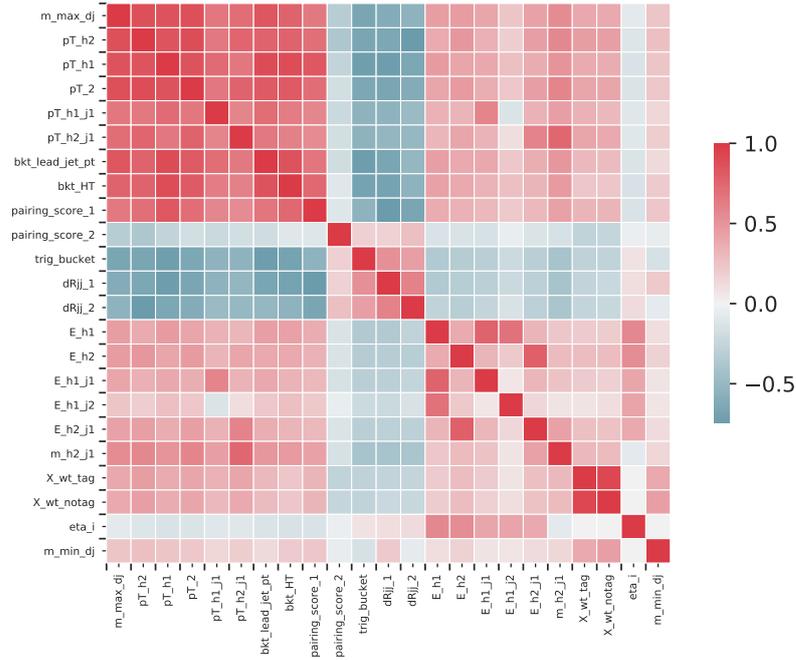
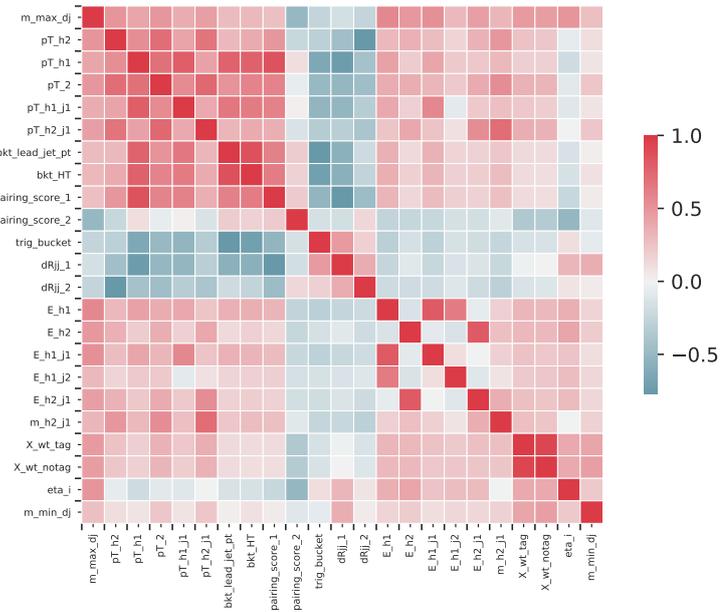

 (a) $|\Delta\eta_{HH}| \leq 1.5$

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

Figure 7.3: The matrices of the Pearson Correlation Coefficients for the kinematic variables of the 2b Control Region 1 data evaluated in two $|\Delta\eta_{HH}|$ categories. The darker the displayed colour (either red or blue), the stronger the correlation between the variables with the strongest (100%) correlation at the diagonal as the variables are correlated with themselves.

7.3.2 Network Architecture

As already outlined in Chapter 4, network architecture also impacts the network performance. Given the amount of available training data (limited by around 900 4b events in each of the control regions, see Table 6.3), the number of input variables and the computational expense, networks of 3 to 5 hidden layers with a varying number of nodes (between 10 to 50 in the first hidden layer), were tried. The tests involved establishing numbers of nodes, layers and network *shape* (i.e. the *bottle neck* shape of consecutive layers having fewer nodes each gradually, e.g. 10:5:2). The optimal network was chosen as a three-layered network with 20 nodes in each layer. Each node is activated with the ReLU activation function and the output is a linear, single node.

7.4 Background Estimate Uncertainties

7.4.1 Statistical Uncertainty and Bootstrapping

The statistical uncertainty on the background modelling is applied. It originates from the nature of the counting experiment and the standard Poisson error is evaluated. The value of the background histogram for any bin i is :

$$n_i = \sum_{j \in i} w_j \quad (7.15)$$

when an event j has the weight w_j . The standard error for such problem is then:

$$\delta n_i = \sqrt{\sum_{j \in i} w_j^2}. \quad (7.16)$$

If all weights are equal to 1, this summation results in the usually quoted Poisson error: \sqrt{N} [73].

Eq. 7.16 does not consider uncertainties on weights, w_j . In order to predict the uncertainty of weights, the neural network is retrained a number of times using the technique known as *bootstrap resampling with replacement* [74]. A number of statistically equivalent subsets are sampled with replacement from the original training set and the neural network is trained separately on each of these subsets. This results in a number of statistically equivalent sets of weights leading to statistically equivalent background estimates. Each time the network is retrained, it is initialised randomly. This aspect is implicitly included in the bootstrapping error.

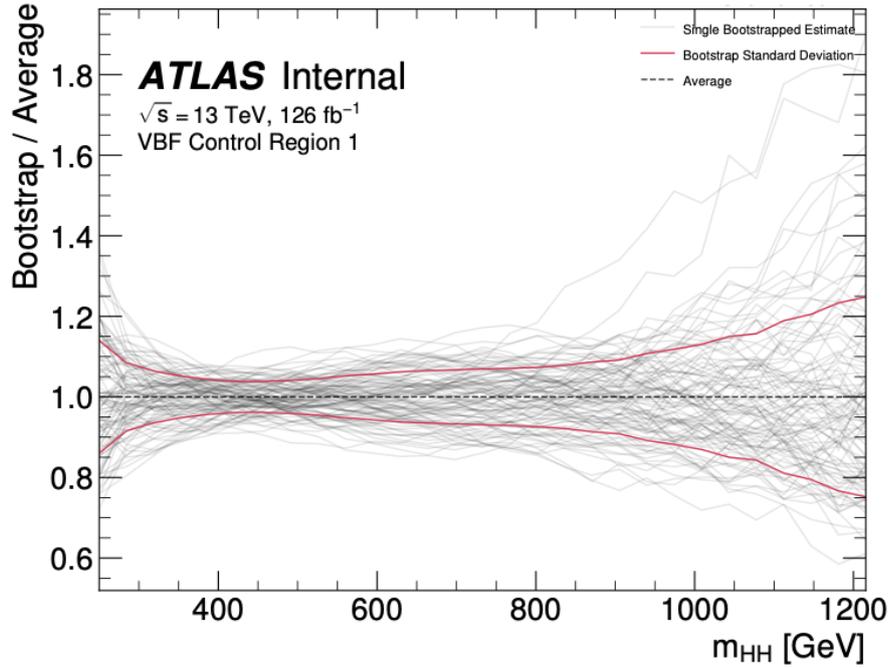


Figure 7.4: Illustration of 100 individual bootstrapped estimates of weights for the m_{HH} spectrum in control region 1, normalized to the average bootstrap weight. The red band indicates the bootstrap standard deviation around the average weight.

Derivation of Bootstrap Error Band

A set of 100 bootstraps is used in the background estimation procedure. This means that for each kinematic variable considered, 100 statistically equivalent histograms are created, each with a set of the bin-level weights predicted by a given network. The nominal background estimate for a given variable is the mean value in each bin. The associated uncertainty is evaluated as a standard deviation about that mean.

Figure 7.4 shows 100 bootstrap bin-level predictions of weights (light grey) normalized to the average weight for each bin as a function of m_{HH} . The red band shows the bootstrap uncertainty which is the standard deviation from the average weight in each bin.

7.4.2 Shape Systematic Uncertainty

The method of bootstrapping accounts for an unstable neural network training, but it does not account for the extrapolation uncertainty related to deriving the background model in one kinematic region, CR1, and applying it to another, the signal region, SR. In order to tackle this problem an alternative background model is derived in CR2. The differences between the CR1 and CR2 predictions constitute the background model’s systematic uncertainty.

A simple approach would be to take the difference between CR1 and CR2 predictions and assume this is a single uncertainty. However, that would result in a single, heavily constrained shape systematic that would impact the fitting procedure described in Chapter 8. Therefore, a method is proposed in which four separate alternative background predictions, corresponding to the splitting of the control regions into four quadrants, are evaluated. As a result, four less constrained shape systematics are used in the final fit in Chapter 8.

Figure 7.5 illustrates the splitting of the massplane for the systematic uncertainty evaluation. The control regions are split into quadrants as outlined in Section 6.7. Additionally, the signal region is split into four parts as shown in Figure 7.5a and the compass naming scheme is used for the SR as well. An orange triangle in Figure 7.5b indicates the N -orth part of the signal region for illustration of the N -orth shape systematic derivation.

For any event in SR, two background predictions are available: CR1 prediction (weights derived in CR1) and CR2 prediction (weights derived in CR2). The CR1 prediction is applied to all SR 2b events giving the nominal 4b background prediction. For the alternative background prediction, here the background prediction N , the CR2 prediction is applied to the N -orth part of the signal region, while CR1 prediction is applied to the remaining signal region parts: E , S and W (indicated with the green triangles). The bin-by-bin difference between the nominal prediction and the alternative prediction N is then evaluated for the signal region m_{HH} distribution.

Figures 7.6a and 7.6b illustrate the difference between the CR1 and CR2-North predictions for the m_{HH} signal region distribution evaluated for each bin in the two $|\Delta\eta_{HH}|$ categories. For an easier overview of the overall shape systematic, the calculated difference between the two predictions is symmetrised about the nominal estimate (from the CR1). The procedure is repeated for shape systematics E , S and W . Figure 7.6 shows shape systematic variations N and E for both $|\Delta\eta_{HH}|$ categories, while 7.7 shows variations S and W .

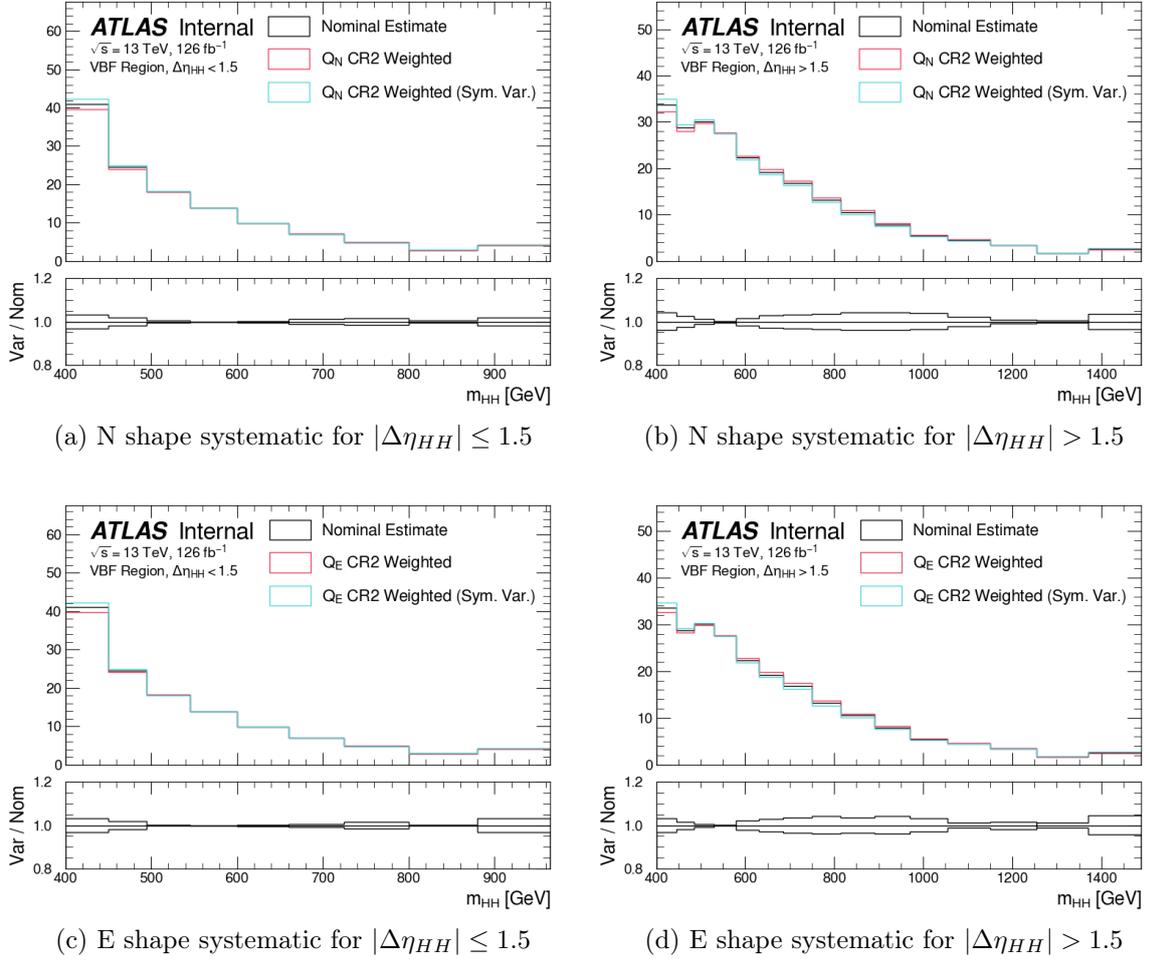


Figure 7.6: Systematic uncertainties evaluated for both regions split in $|\Delta\eta_{HH}|$. Plots in the left column are for $|\Delta\eta_{HH}| \leq 1.5$ category, plots in the right column for $|\Delta\eta_{HH}| > 1.5$. The first row shows differences between CR1 and CR2: N , and the second: E .

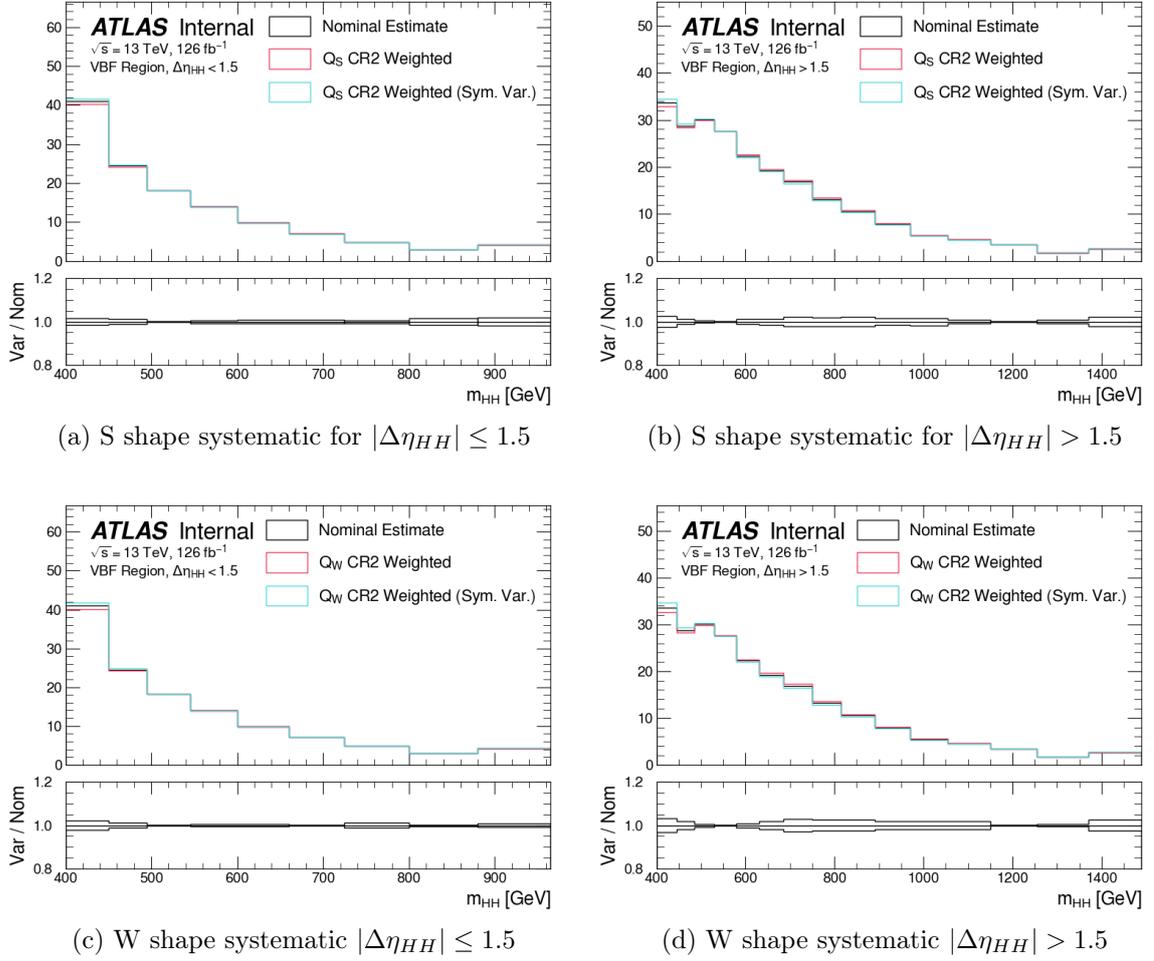


Figure 7.7: Systematic uncertainties evaluated for both regions split in $|\Delta\eta_{HH}|$ categories. Plots in the left column are for $|\Delta\eta_{HH}| \leq 1.5$ category, plots in the right column for $|\Delta\eta_{HH}| > 1.5$. The first row shows differences between CR1 and CR2: S , and the second: W .

7.5 Performance of the Reweighting in the Control Regions

A closure check in the control regions in which the network was trained is performed (in CR1 and CR2 for the networks trained in these respective regions), to assess the performance of the neural network technique. The level of agreement between the reweighted 2b data and the 4b data is checked. The agreement is quantified by the calculation of χ^2/NDF , where NDF is the number of degrees of freedom.

Figures 7.8 - 7.13 show closure checks performed on a set of chosen distributions; Figures 7.8 and 7.9 for the variables on which the network was trained, while Figures 7.10 - 7.13 show kinematic distributions that were not involved in the training, including the discriminant distribution, the m_{HH} , in Figure 7.13. In the left columns, the plots show the pre-reweighting 2-tagged data, normalized to match the number of the 4b events in the control region. In the right columns, the 2-tagged data for which the reweighting was applied are presented. In each figure, the top row shows the CR1 training, while the bottom row presents the CR2 (alternative) training. In all plots, the 4b data points are indicated together with associated 4b Poisson error. The bottom panel shows the ratio of the normalized, or the reweighted, 2-tagged data to the 4-tagged data. The background estimate statistical error, including the total of the 2b Poisson and bootstrap errors, is marked as a grey hatched area.

Generally, the good agreement between the reweighted 2-tagged data and the 4-tagged data is observed; the pre-reweighting 2b to 4b agreement is improved after the reweighting. The χ^2/NDF values are quoted in each plot. Since this is a closure check, the lower the χ^2/NDF value, the better the closure. χ^2/NDF values obtained after reweighting are in most cases no larger than 2, meaning the closure was achieved. For the m_{HH} distribution in the CR1, $\chi^2/\text{NDF} = 0.362$ post reweighting, compared to 1.568 pre-reweighting.

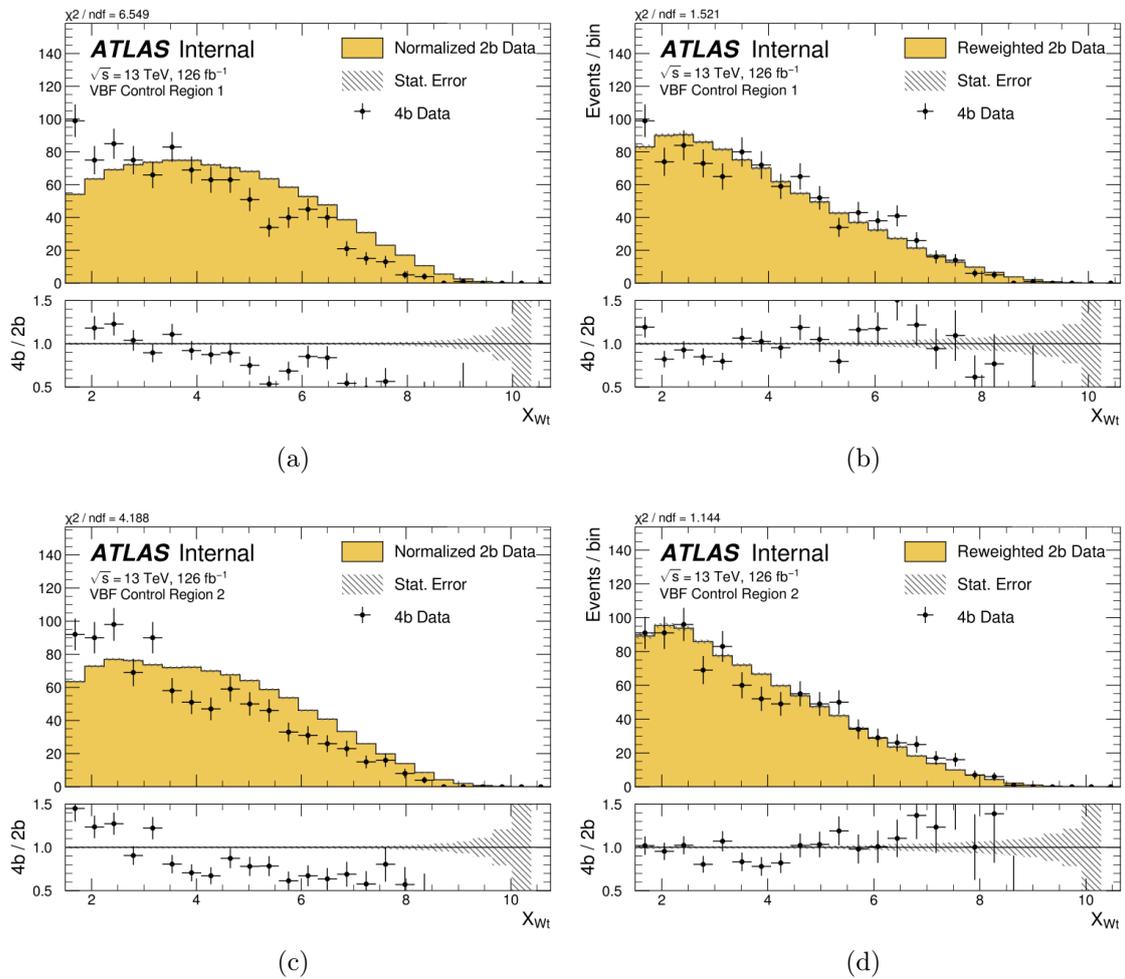


Figure 7.8: Distributions of the training variable: X_{Wt} in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

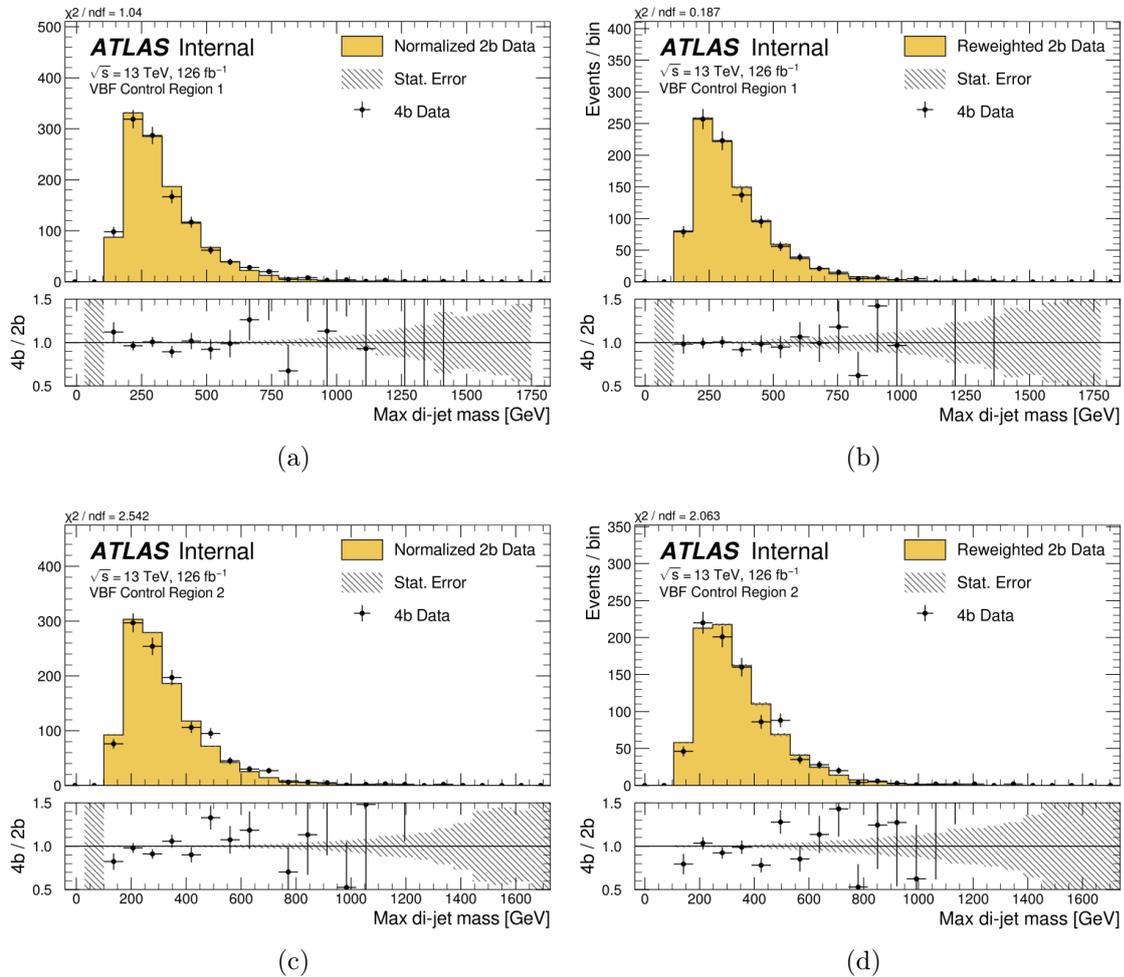


Figure 7.9: Distributions of the training variable: $m_{\text{max_dj}}$ in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

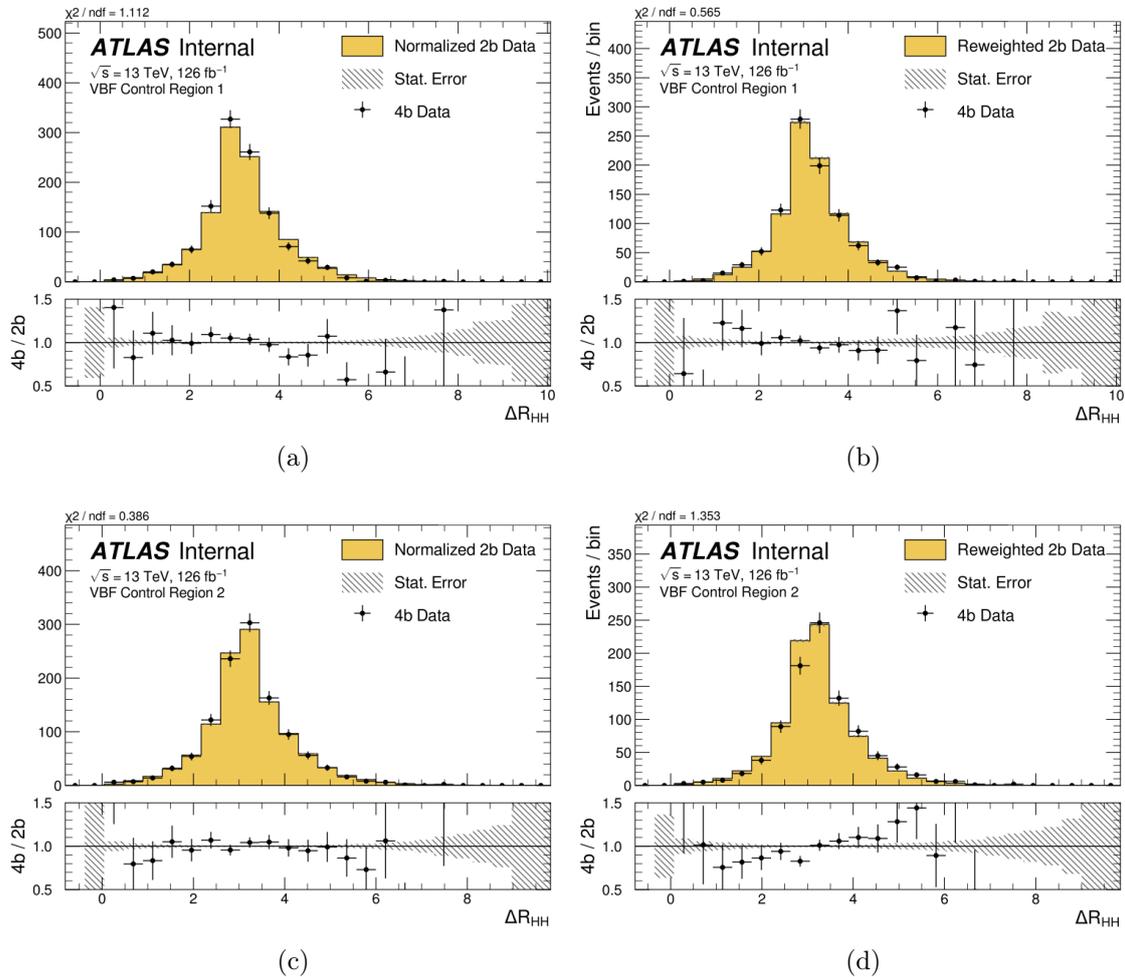


Figure 7.10: Distributions of the non-training variable: ΔR_{hh} in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

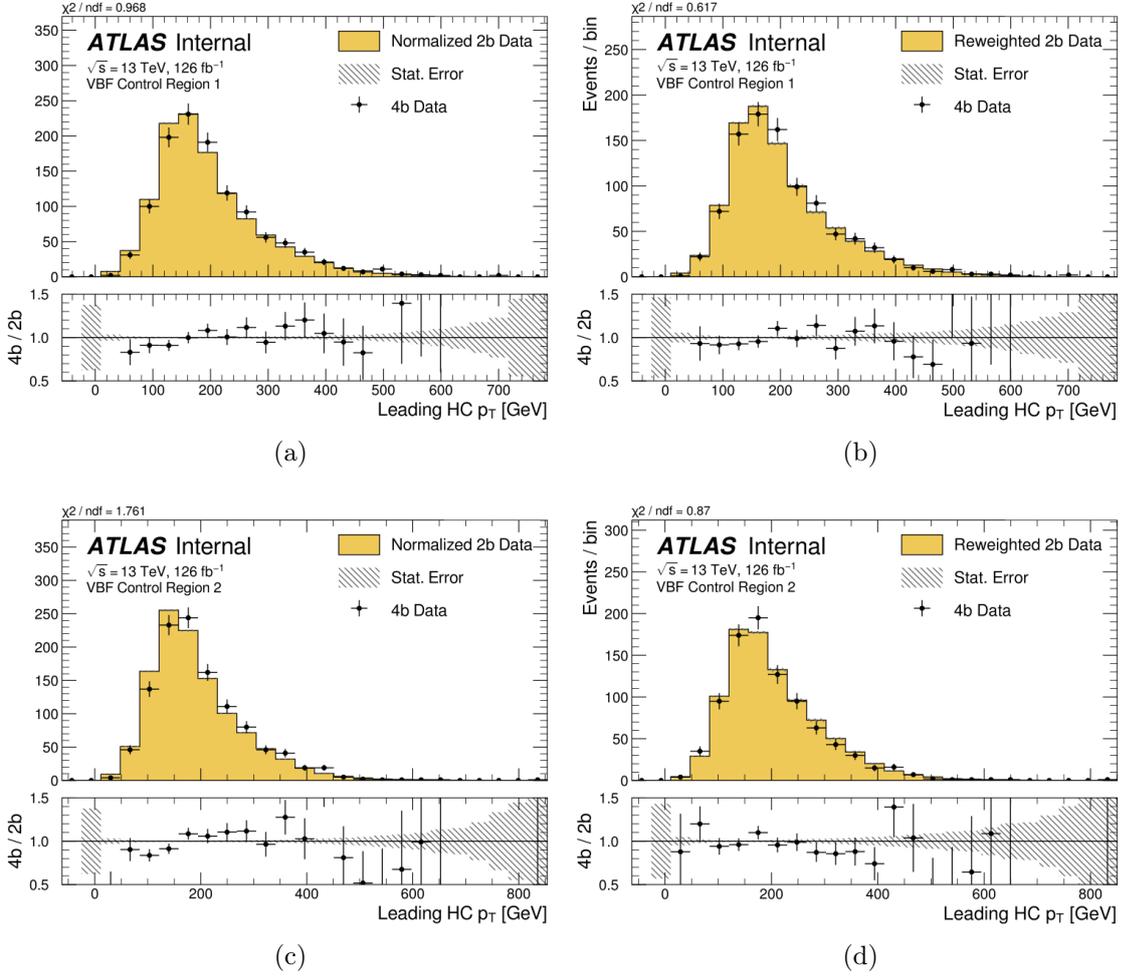


Figure 7.11: Distributions of the non-training variable: p_T of the leading Higgs Candidate in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

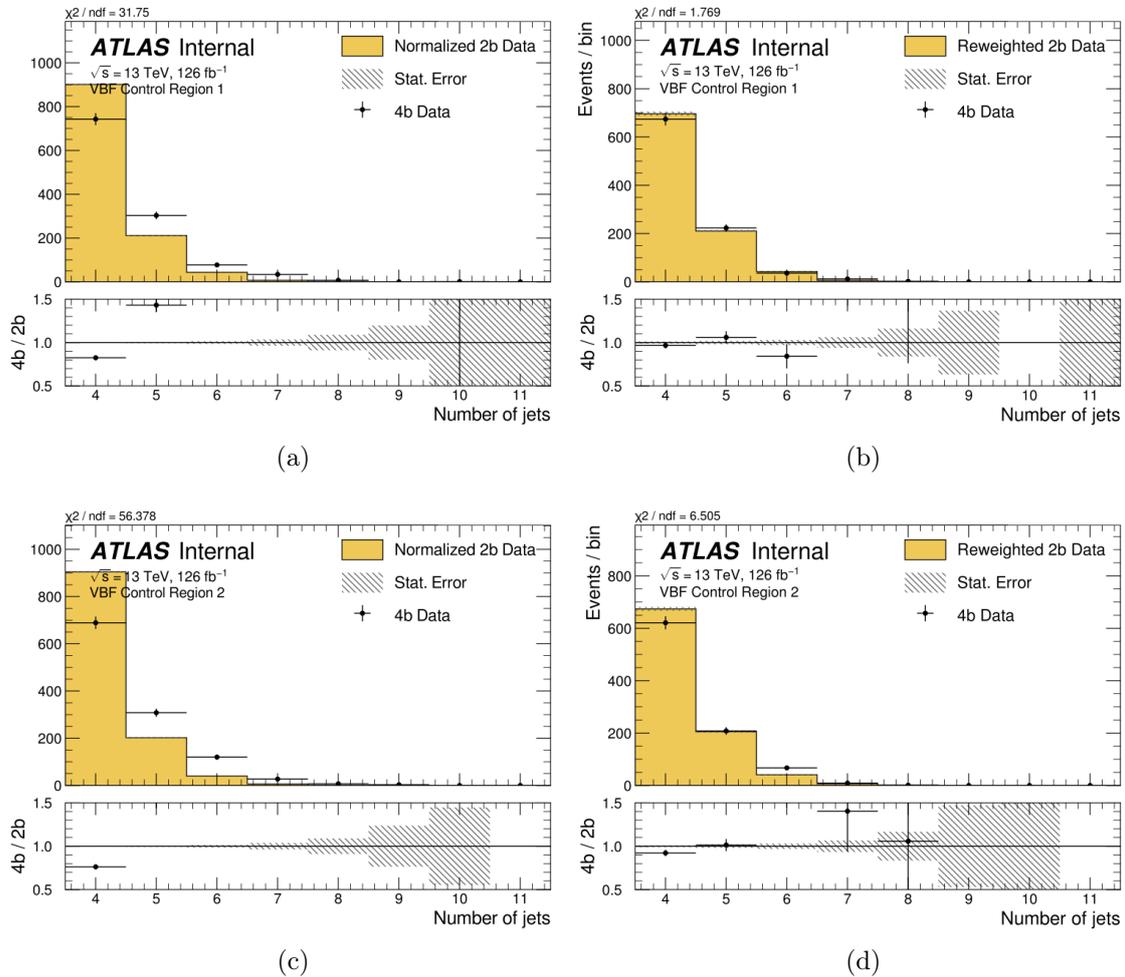


Figure 7.12: Distributions of the non-training variable: number of jets in the event in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

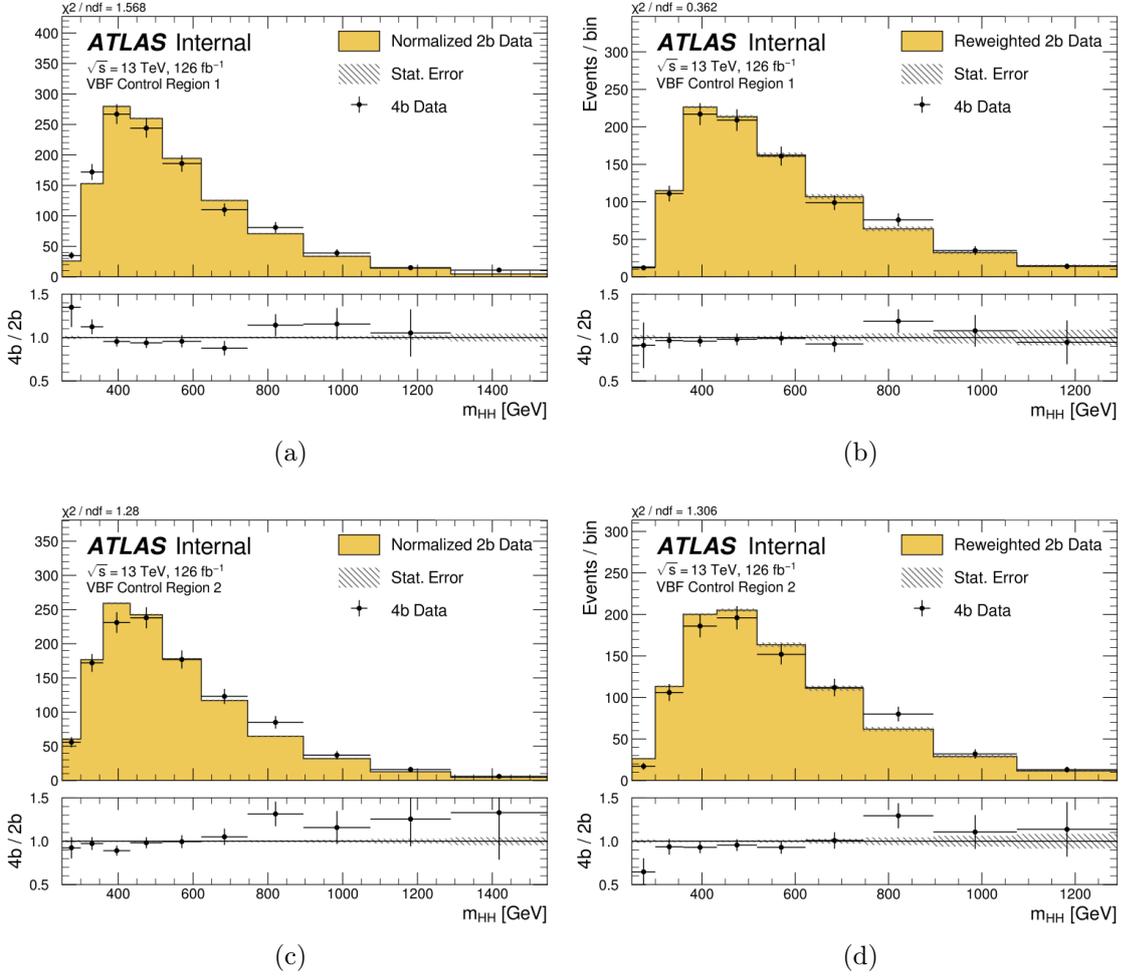


Figure 7.13: Distributions of the non-training variable, discriminating variable of the analysis: m_{HH} in Control Region 1 (top row) and Control Region 2 (bottom row). The errors indicated in grey are the total background estimate statistical uncertainty: 2b Poisson and bootstrap. The 4b data points are shown with the Poisson uncertainty.

7.6 Validation of the Background Estimate

The background estimation method is validated in two different ways. The idea of the validation is to use the exact same method of neural network reweighting as on the analysis data, but for reweighting of the 2b data to the data excluded by the analysis selections. This way, an unbiased confirmation that the method is working, without unblinding the 4b data of the signal region, is possible.

7.6.1 3b1f Region Validation

Any b-tagging referred to in the analysis, is the b-tagging at the 77% working point. A new analysis region for the background validation is proposed in which the first three leading p_T central jets are b-tagged at the 77% working point and no central jet passes the looser 85% working point². This validation region is referred to as $3b+1f$, where f stands for *fail*³.

Analogically to the main background estimation, reweighting starts with the 2-tag data. This time instead of aiming to learn the 4-tag distributions in the control regions, the network learns the $3b+1f$ distributions. Since $3b+1f$ is not our analysis signal region, it is unblinded. The 3b1f data has much more statistics, a factor of 10 more, than the original 4b data. Such a boost in the statistics untruly improves the network’s performance. This is not desired as we want to validate the original network performance which learns on a scarcer dataset. Therefore, at the start of the validation procedure, the data is first *downsampled*. It is split into 10 sets and the whole procedure is repeated 10 times. However, this must not be confused with the earlier described bootstrap repetitions; in each of the 10 sets here, 100 bootstraps are still performed to estimate the mean weight. Figure 7.14 shows the m_{HH} distribution of the reweighted 2b data and the 3b1f data in the signal region for the 10 downsampled trainings. Most importantly, it is checked that the ratio of the 2b to 3b1f data (bottom panel) is constant for all the downsampled trainings. It is therefore validated that the background estimation works well in the 3b1f signal region.

²More details on the b-tagging working points can be found in Section 5.2.2.

³The initial motivation for the use of a loose 85% working point on the fourth leading b-jet was considered as an additional signal category. This category was eventually dropped but its complementary category, 3b1f, is kept as the validation region.

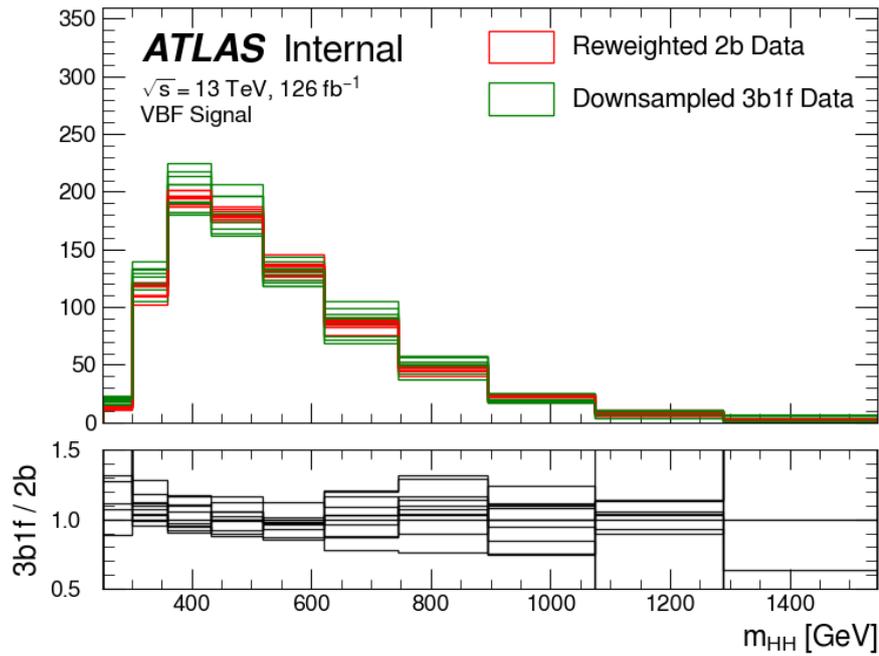


Figure 7.14: Background validation performed in the signal region of the 3b1f data sample shown as the 3b1f SR m_{HH} distribution. There are 10 separate green predictions for 10 downsampled sets of 3b1f data. Analogously, the 10 separate 2b reweighted data, the background estimates, are shown in red. The bottom panel shows the ratio of the 3b1f data to the reweighted 2b data for 10 separate samples.

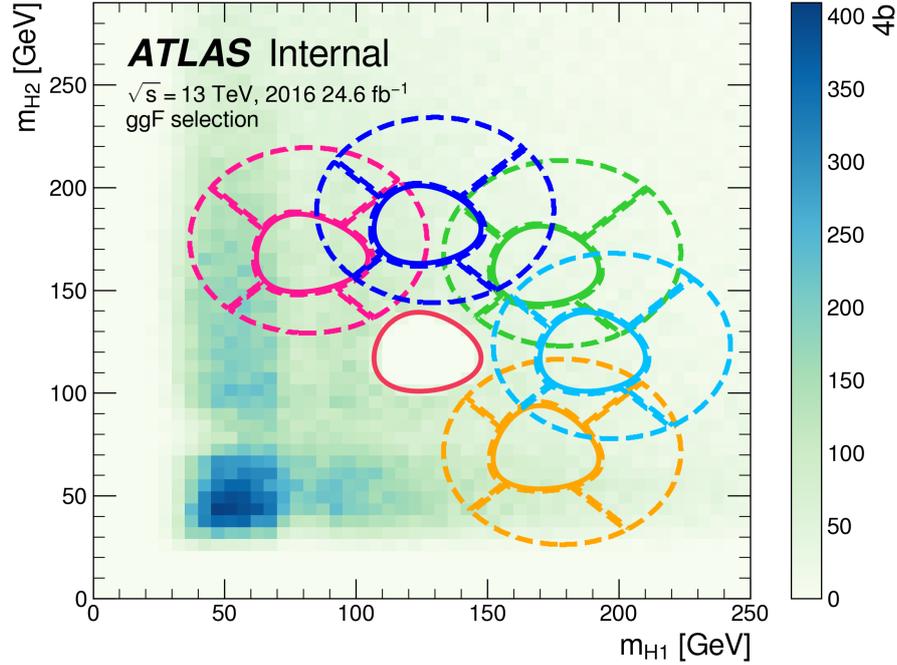


Figure 7.15: Illustration of the shifted kinematic regions, here overlaid on the ggF massplane (Figure taken from the same non-resonant analysis internal note [6]), the same is done on the VBF massplane. The new regions are named based on their relative position on the massplane: upper left is pink, upper centre is dark blue, upper right is green, centre right is light blue and lower right is orange.

7.6.2 Shifted Kinematic Regions Validation

The second validation method is based on redefining five full signal and control regions areas, as shown in Figure 7.15. The regions overlap neither the original signal region nor themselves. They are in proximity to the original SR so that the validation is reliable. The naming scheme of the five regions is given in the figure’s caption.

Regions are defined with a similar equation to the original SR equation (Eq. 6.12):

$$X_{HH,\text{shift}} = \sqrt{\left(\frac{m_{H1} - m_{H1,\text{centre}}}{\sigma_{m_{H1}} m_{H1}}\right)^2 + \left(\frac{m_{H2} - m_{H2,\text{centre}}}{\sigma_{m_{H2}} m_{H2}}\right)^2} < 1.6, \quad (7.17)$$

where the $m_{H1,\text{centre}}$ and $m_{H2,\text{centre}}$ are given in Table 7.1. The resolutions: $\sigma_{m_{H1}}$ and $\sigma_{m_{H2}}$ are set to 1 for *upper left*, *upper right* and *lower right* regions, while for the regions in between, *upper centre* and *lower right*, one of these values is slightly

Table 7.1: Centers of the shifted regions.

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

decreased in order to avoid overlap. For *upper centre*:

$$\sigma_{m_{H1}} = 0.1 \times \frac{124}{m_{H1,\text{centre}}}, \quad (7.18)$$

and for *centre right*:

$$\sigma_{m_{H2}} = 0.1 \times \frac{117}{m_{H2,\text{centre}}}, \quad (7.19)$$

with 124 and 117 being the centre of the original signal region. As can be seen in Eq. 7.17, the new signal regions are again required to be contained within the X_{HH} value of 1.6, as for the original SR. The control regions surrounding each of the shifted signal regions are defined analogously to the original CR1 and CR2.

The background estimation procedure is performed separately for each shifted region. The comparison of the results, the 4b event yields with the Poisson uncertainties, and the 2b reweighted event yields in the shifted signal regions, is shown in Table 7.2. The total uncertainties for the background prediction, added in quadrature 2b Poisson, bootstrap and systematic, are quoted as well. The predictions agree well (within uncertainties) with the 4b background yields in the shifted signal regions. The highest total relative error of 21% is observed in the lower right region due to its S quadrant's position at very low m_{H2} . The second highest total relative error of 18% is observed for the upper left region. In that case, quadrant W lies in the low m_{H1} region. At low m_{H1} and m_{H2} , bigger differences between the 2b and the 4b kinematics are present, causing an increase in the shape systematic error.

Figure 7.16, shows the m_{HH} distributions in the shifted signal regions. The good agreement is seen throughout, validating that extrapolating weights from the control regions into the signal region works well.

Based on these two background estimate validation methods, the neural network background reweighting technique is validated. For the non-resonant ggF analysis, two more validations are performed. For the ggF, an additional selection cut is performed on the $\Delta\eta_{HH}$ variable. Events with $\Delta\eta_{HH} > 1.5$ are discarded to reduce the QCD background [6]. The 2b to 4b reweighting is performed with the discarded events and the signal region closure is checked. The $\Delta\eta_{HH}$ selection is not used in the VBF analysis, as many signal events would be lost. The validation with the

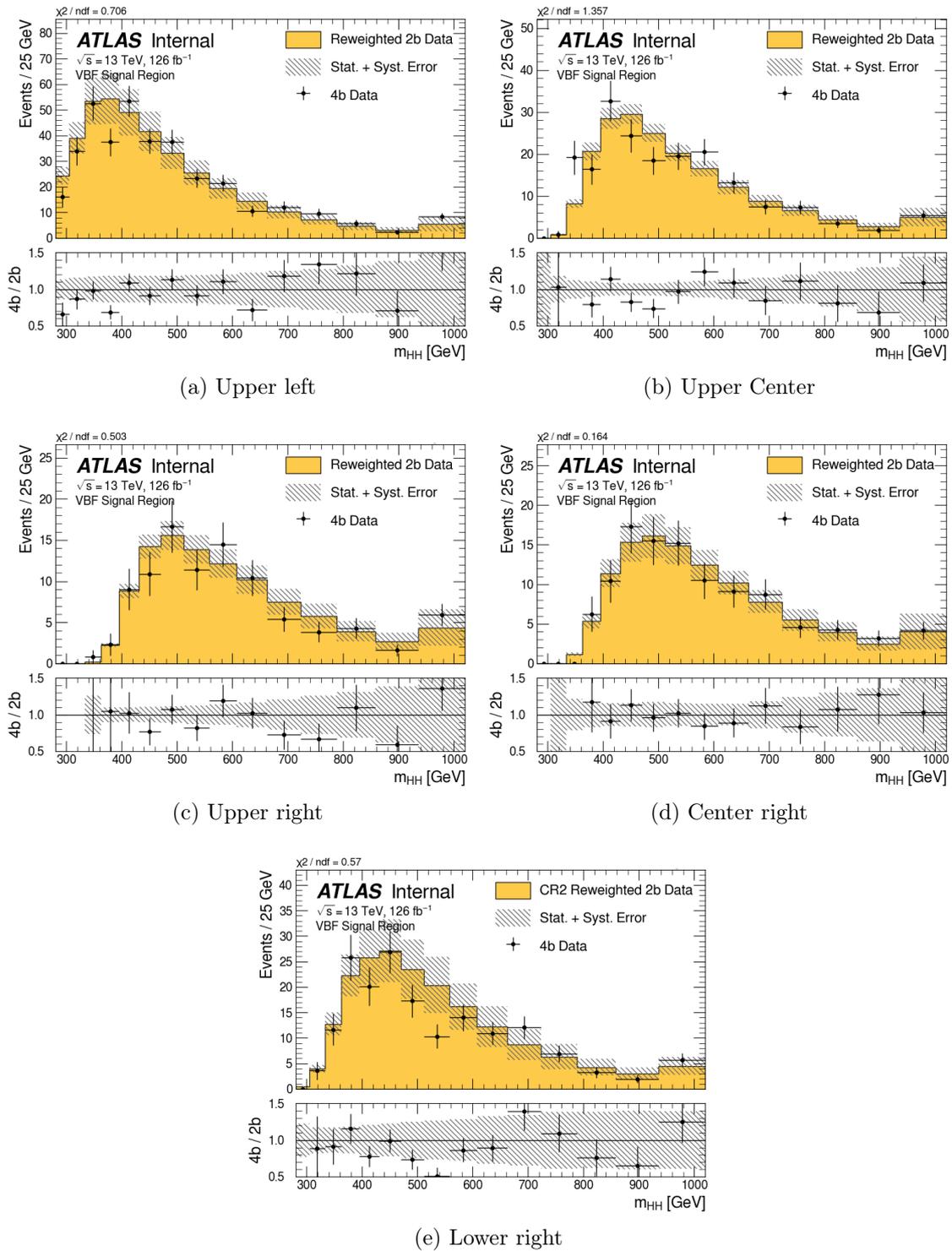


Figure 7.16: The m_{HH} distributions in the five shifted signal regions showing the reweighted 2b data in yellow and the 4b signal region data with the 4b Poisson errors in black. The grey band shows the total statistical and systematic uncertainty for the background prediction.

Table 7.2: Comparison of the 4b yields (with Poisson error) and the reweighted 2b (background prediction) yields in the shifted signal regions. The quoted Background Error includes the Poisson error, the bootstrap error and the shape systematic error.

Shifted Region	SR 4b Yield	Background Prediction	Background Error [%]
Upper Left	578 ± 24	595 ± 110	18
Upper Center	338 ± 18	339 ± 14	4
Upper Right	195 ± 14	203 ± 10	5
Center Right	214 ± 15	216 ± 14	6
Lower Right	303 ± 17	338 ± 71	21

Monte Carlo samples of the $t\bar{t}$ and the di-jet events was performed in the ggF as well. It was decided that the conclusion from the ggF, namely, that the closure is observed, is sufficient and the study was not re-performed for the VBF.

7.7 Background Prediction in the 4b Signal Region

Finally, the background prediction in the 4b Signal Region is evaluated. Background distributions in the two $|\Delta\eta_{HH}|$ categories together with the unblinded 4b data in the signal region are shown in Figure 7.17. The corresponding total background uncertainty: statistical and systematic, is indicated as well. The background prediction agrees well with the unblinded 4b data, the closure is observed with χ^2/NDF values of 1.00 and 1.25 for the two categories respectively.

Figure 7.18 shows the relative background uncertainties as a function of the m_{HH} in two $|\Delta\eta_{HH}|$ categories in the 4b signal region. In both categories, for the low values of m_{HH} , the shape systematic error is dominant. Starting at $m_{HH} = 500$ GeV for $|\Delta\eta_{HH}| < 1.5$ category and at $m_{HH} = 1000$ GeV for $|\Delta\eta_{HH}| > 1.5$ category, the bootstrap error dominates.

Table 7.3 presents the total magnitudes of the statistical and systematic background uncertainty components in the 4b signal region for both categories inclusively to assess the total background error. The background uncertainties are evaluated as the event counts and as a relative percentage error. The predicted background yield in the 4b signal region is 354.6. The total statistical error quoted in the table is the result of the addition of Poisson and bootstrap errors in quadrature. The total systematic error is the result of adding four individual shape uncertainties in quadrature. The total relative statistical error is 9.1%, while the total relative background shape uncertainty is 0.45%. Both, added in quadrature, result in the total background uncertainty of 9.1%.

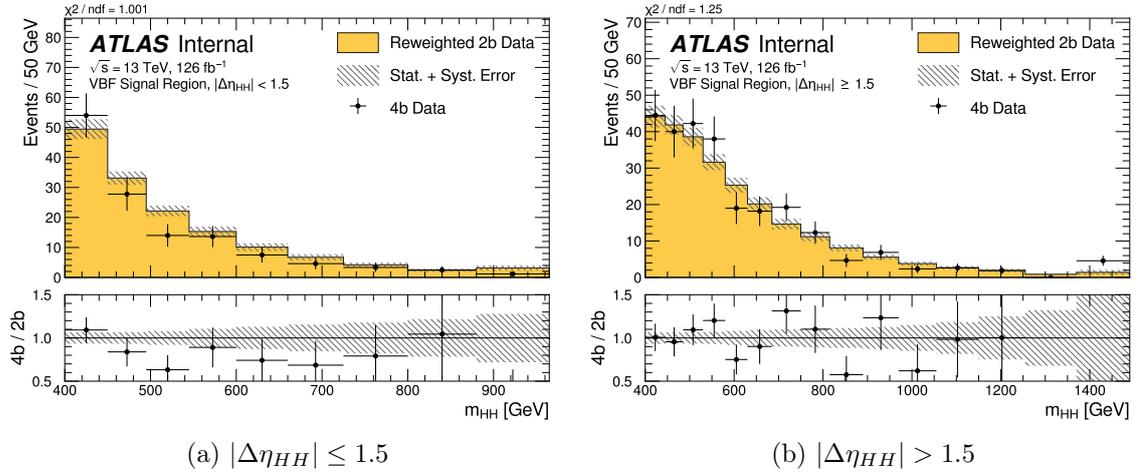


Figure 7.17: The m_{HH} distributions in $|\Delta\eta_{HH}|$ categories with the 2b reweighted data, the background prediction, marked in yellow, and the unblinded 4b data with associated Poisson error marked in black. Total statistical and systematic error for the background estimate is shown as hatched area. The bottom panel presents the 4b-to-reweighted-2b ratio.

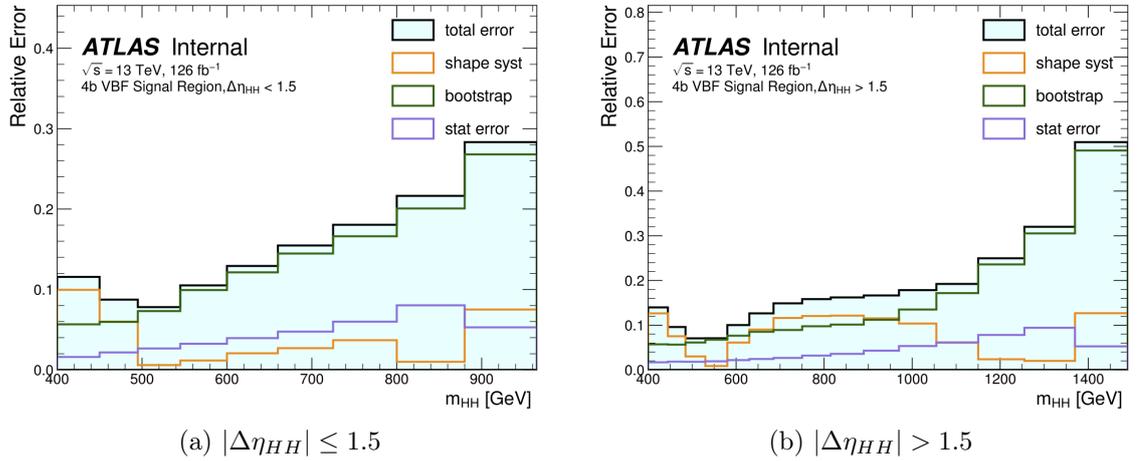


Figure 7.18: Composition of the relative background error as a function of the m_{HH} in both $|\Delta\eta_{HH}|$ categories.

Table 7.3: Magnitudes of the statistical and systematic background uncertainty components in the 4b signal region. The Poisson and bootstrap errors are added in quadrature to obtain the total statistical error.

	Counts	Relative Error %
Total Background Prediction	354.6	
Total Statistical Error	32.3	9.1
Shape N	1.2	0.34
Shape E	0.3	0.10
Shape S	0.5	0.15
Shape W	0.8	0.23
Total Systematic Error	1.6	0.45
Total Error	32.3	9.1

Chapter 8

Statistical Analysis

8.1 Mathematical Methods

8.1.1 Hypotheses of the Experiment

The main goal of the statistical inference is to confirm or reject a hypothesis. In the case of our experiment, there are two kinds of hypotheses. Namely, the *null hypothesis* (or the *background only hypothesis*), which assumes that the data describes only the SM background, and the *alternative hypothesis* in which the data describes both background and signal: $s + b$.

In practice, for this analysis, the real data are plotted in the form of histograms of the m_{HH} variable. It is then possible to check how well the data agree with the background-only hypothesis. To test the alternative hypotheses, similar histograms are prepared with the Monte Carlo simulated data for different κ_{2V} signal shapes. If there is no excess in data, suggesting the presence of the signal, an upper 95% confidence level limit on the signal cross-section is evaluated; we want to find the amount of signal which is not rejected by the data according to the chosen confidence level, here 95%. The CL_s technique is used for such limit evaluation [75]. It ensures that the signal events which are not well separated from the background events are not prematurely rejected. It is given by:

$$CL_s = \frac{CL_{s+b}}{CL_b}, \quad (8.1)$$

where CL_{s+b} is the confidence level for the alternative hypothesis and CL_b for the null hypothesis. Many alternative hypotheses are possible, but instead of testing them separately, a continuous, parametrized alternative hypothesis reads:

$$H_{\text{alternative}} = s(\mu) + b = \mu \cdot s_0 + b, \quad (8.2)$$

where μ is the signal strength parameter and s_0 is the reference signal strength taken as 1 fb here. In other words, μ can be treated as a signal cross-section relative to the SM signal cross-section.

8.1.2 Likelihood

The CL_s method is often based on the profile likelihood ratio. In order to first introduce the likelihood, we take a step back and first assume a hypothesis H . In the Frequentist regime, the probability for the frequency of data x given H is considered; if an experiment is repeated a number of times, the probability for the frequency of data is $P(x|H)$.

The experimental outcomes and their frequencies are already known. These data are used to extract the parameters of the distribution, the so-called parameters of interest (POIs). $P(x|H)$ is now treated as a function of the hypothesis H given the data x . Instead of a probability for data, we define the Likelihood, L , of the hypothesis;

$$L(H) = L(H|x) = P(x|H). \quad (8.3)$$

Since our data are coming from a counting experiment, this likelihood is taken as a Poisson distribution of n repeated experiments. Our hypothesis is then expressed by a probability distribution function $f(m_{HH})$ that depends on the signal strength μ , and the systematic uncertainties.

The general equation for the unbinned likelihood reads:

$$L(H) = L(f(m_{HH})|x) = \text{Pois}(n|\lambda = n_{exp}) = e^{-\lambda} \frac{\lambda^n}{n!}, \quad (8.4)$$

where *exp* is short for *expected*.

Since our signal and background events are categorized in $|\Delta\eta_{HH}|$ and binned in the m_{HH} histogram, Eq. 8.4 becomes:

$$L(f(m_{HH})|x) = \prod_{c \in \text{categories}} \prod_{b \in \text{bins}} \text{Pois}(n_b|\lambda = n_{b,exp}) \quad (8.5)$$

where n_b is the number of events in bin b and $n_{b,exp}$ is the expected number of events in that bin.

The uncertainty associated with the Poisson distribution, the statistical uncertainty, is defined as one standard deviation from the expected number of events, n_{exp} ,

$$\sigma_{n_{exp}} = \sqrt{\text{Var}(n_{exp})}, \quad (8.6)$$

which with a relatively large sample, defined as $n_{exp} \gtrsim 30$ [76], simplifies to:

$$\sigma_{n_{exp}} = \sqrt{n_{exp}}. \quad (8.7)$$

On top of the statistical fluctuations, there is a number of systematic uncertainties related to the methods of data collection and analysis at its various stages. These

systematic uncertainties constitute the Nuisance Parameters (NPs) of the model. These are the parameters of no interest in terms of physics interpretations but are still present in the model description. They are denoted by θ . Each NP has a constraint term, $p(\theta'|\theta)$ in which θ' is an auxiliary measurement through which the information about a given θ is received. The used constraining function is usually Gaussian, but sometimes Poisson constraints are used instead [77]. All these elements are summarized in the following likelihood equation:

$$\begin{aligned}
 L(f(m_{HH})|x) &= \prod_{c \in \text{categories}} \prod_{b \in \text{bins}} \text{Pois}(n_{c,b}|\lambda = n_{c,b,exp}) \prod_{s \in \text{syst}} p(\theta'|\theta) = \\
 L(\mu, \theta) &= \prod_{c \in \text{categories}} \prod_{b \in \text{bins}} \text{Pois}(n_{c,b}|\lambda(\hat{\mu}, \hat{\theta})) \prod_{s \in \text{syst}} p(\hat{\theta}|\theta). \quad (8.8)
 \end{aligned}$$

It is important to note that the *hatted* parameters are the best fit values of the likelihood. Since they were estimated from the data, we say they were *profiled* and the likelihood from Eq. 8.8 is called the profile likelihood.

8.1.3 Profile Likelihood Method

Having defined the Profile Likelihood function, we proceed with setting the limits using the CL_s method as shown in Eq. 8.1. In order to test our hypothesis, we define a test statistic q , simply a measure of compatibility of the data with the signal and background model. When q is small, the level of compatibility is large and if the $CL_s \leq 0.05$, a given signal strength μ is excluded; we obtain a limit on the cross-section. The CL_s at 95% CL is defined as:

$$CL_s(\mu) = CL_{s(\mu)} = \frac{CL_{s(\mu)+b}}{CL_b} = \frac{p(q_\mu \geq q_{\mu, \text{obs}}|s(\mu) + b)}{p(q_\mu \geq q_{\mu, \text{obs}}|b)}, \quad (8.9)$$

with both p-values calculated in the asymptotic approximation, which is possible due to a large number of points under consideration [76], as the background is large. Limits are evaluated using the test statistic defined by the following one-sided profile likelihood ratio:

$$q_\mu = \begin{cases} -2 \ln \left(\frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right) & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \quad (8.10)$$

where μ is the POI, the signal strength, and θ is the set of nuisance parameters. $\mathcal{L}(\mu, \hat{\theta}(\mu))$ is the conditional maximum likelihood fit with the NPs being set to their profiled values represented by $\hat{\theta}$. For $\hat{\theta}$, the likelihood is maximized while the value

of μ is fixed. $\mathcal{L}(\hat{\mu}, \hat{\theta})$ is the unconditional likelihood fit, both μ and θ can vary, while the overall likelihood is maximized for a pair $\hat{\mu}$ and $\hat{\theta}$. For $\hat{\mu} > \mu$, the test statistic is zero, because the method of setting the upper limits is developed to test signal strength μ against the alternative hypotheses which predict lower values of μ .

To recapitulate, the statistical procedure of setting the upper limits on the signal strength can be done in the following steps:

- Define a range of the signal strengths μ over which the limit setting scan will be performed.
- At each point of this range, calculate the test statistic q_μ given by Eq. 8.10. Since this is done using the observed data, the final results are $q_{\mu, \text{observed}}$.
- Now the $CL_{s(\mu)+b}$ and CL_b can be calculated using Eq. 8.9 and their ratio can be taken to obtain $CL_{s(\mu)}$.
- Finally, exclude all signal strengths for which the calculated $CL_{s(\mu)} < 0.05$. Namely, observing this data given the particular μ is the true signal strength, has the probability of less than 5%: μ is excluded at 95% confidence level.

8.2 Implementation in the pyhf Framework

The profile likelihood method was implemented in the PYHF framework [78], [79]. The framework is the pure Python-based implementation of the HISTFACTORY p.d.f. template package¹. The data, MC and NP histograms are first read by the framework, which then creates a workspace for each specific combination of the coupling values $(\kappa_\lambda, \kappa_{2V}, \kappa_V)$, as described in Section 6.2.1, also including the NP variations. Finally, the fitting procedure is performed in each workspace.

The final discriminating variable, the m_{HH} , is represented by various histograms, subsequently read by the framework. The types of histograms used are:

- background histogram - the reweighted 2 b-tagged data,
- data histogram - the 4 b-tagged data in the signal region,
- signal histogram - the MC generated signal for a given coupling values combination,
- NP variation histograms - to account for systematics.

All histograms are evaluated in the two $|\Delta\eta_{HH}|$ categories and using the binning method introduced in Section 6.10.

¹Histfactory is usually used with ROOT and hence frameworks such as RooFit or RooStats.

8.3 Uncertainties

The uncertainties are divided into the background uncertainties for the fully data-driven background estimate, and the signal uncertainties for the Monte Carlo generated signal samples. The signal uncertainties include detector modelling, luminosity and theory uncertainties. All uncertainties are evaluated as up and down variations of the Nuisance Parameters m_{HH} histograms. For the uncertainties evaluated specifically in the course of the analysis, as well as the luminosity and the theory cross-section uncertainties, the total relative errors are quoted. The per-event trigger scale factors and the related uncertainties as well as those derived by the ATLAS *Combined Performance* groups (Jet Energy Scale, Jet Energy Resolution and flavour tagging), are not quoted as they depend on the p_T and η distributions. The $\pm 1\sigma$ uncertainty variations histograms of the scale factors for each event are evaluated during the fitting procedure instead.

8.3.1 Summary of Background Uncertainties

The background uncertainties are discussed in Chapter 7. The total statistical uncertainty consists of the 2b events uncertainty, the Poisson uncertainty, and the statistical uncertainty of the weights, the bootstrap uncertainty. The total relative statistical uncertainty is 9.1%, while the total relative background shape uncertainty is 0.45%. The two uncertainties added in quadrature give the total background uncertainty of 9.1%.

8.3.2 Summary of Signal Uncertainties

Luminosity

The LUCID-2 detector [80] in the LHC complex is responsible for the luminosity studies. Based on the luminosity scale calibration, the uncertainty applied to the total integrated luminosity is 1.7% [81].

Theory Uncertainties

The first theory uncertainty addresses the uncertainty of the underlying event and parton shower modelling. The customary procedure, used in many ATLAS analyses and recommended by [82] and [83], consists of a comparison of results produced by two event generators, in which modelling of the underlying event and of parton shower differ. The results from PYTHIA8 [65], used in the analysis, are compared to the results from HERWIG7 [84],[85]. The result of the comparison is a 10% impact

on the VBF signal acceptance and none on the m_{HH} shape. Detailed studies on the parton shower uncertainties for this analysis are described in [6].

The factorization and renormalization scale uncertainty is evaluated by varying both scales up and down by a factor of two. For this analysis, the uncertainty varies between around 2% and up to 6%, more detailed studies are presented in [6]. The PDF and α_s uncertainties are assigned to evaluate the standard deviation of the signal acceptance for the PDFs from the PDF4LHC_NLO_MC set [86]. The result is 1 - 2 % relative error.

As each Higgs decays to a pair of bottom quarks, the $H \rightarrow b\bar{b}$ branching fraction uncertainty of 1.7%, evaluated in [87], is doubled to obtain the total branching ratio uncertainty of 3.4% for the HH . The uncertainty on the cross-section is related to the PDF used and the α_s variations, which impact the cross-section by 2.1% [87], while the scale of the top-quark mass impacts the cross-section by 0.04% [87]. The summary of the theory uncertainties can be found in Table 8.1.

Jet Energy Scale and Resolution (JES and JER)

The performance of the MC-simulated jets is compared to the performance recorded in the data. The evaluation of the jet uncertainties is based on the prescription given in [55]. To evaluate the uncertainty, calibrations are performed using the in-situ methods: dijet η -intercalibration, Z+jet, γ +jet and the multi-jet balance. The single particle uncertainties are used for the high p_T regime. The calibration plots for the MC jets to account for the performance of the PFlow jets can be found in [55], for all the listed methods. The methods to evaluate the Jet Energy Resolution uncertainties for the MC consist of the calibrations with the PFlow jets using noise term evaluation from random cones in zero bias data and the dijet p_T balance asymmetry [55]. The analysis uses 24 JES NPs and 18 JER NPs, as recommended by the ATLAS JES and JER recommendations internal TWiki [88]. All jet-related NPs are listed in Table 8.2.

Flavour Tagging

The b-tagging performance in the MC is calibrated with the b-tagging performance in data resulting in an application of the scale factors to the MC weights provided by the ATLAS Flavour Tagging group [89], [61]. For the used b-tagging working point, the scale factors are close to 1, as discussed in more detail in Section 5.2.3. All b-tagging, and c and light-flavour mistagging uncertainties used in the analysis (13 NPs in total) are listed in Table 8.3.

Table 8.1: List of all background and theory uncertainties used in the analysis. Types and names of all uncertainties as well as total relative percentage values are presented.

Type of Uncertainty	Name of NP	Total relative value [%]
Background Uncertainties	shape_N	0.34
	shape_E	0.10
	shape_S	0.15
	shape_W	0.23
	Total Statistical Uncertainty	9.1
Luminosity Uncertainty	ATLAS_LUMI_Run2	1.7
Theory Uncertainties	THEO_ACC_PS	10
	THEO_ACC_SCALE	2 - 6
	THEO_ACC_PDFalphas	1 - 2
	THEO_BR_Hbb	3.4
	THEO_XS_PDFalphas	2.1
	THEO_XS_SCALEMTop	0.04

Trigger

The trigger scale factors correct the online (both L1 and HLT) efficiencies of reconstructed jet kinematics as well as the online jet b-tagging as discussed in Section 6.3.1. Two Trigger NPs are listed in Table 8.3.

8.4 Results

The maximum likelihood fits are performed using the prescription from Eq. 8.8. The Poisson distributions govern the event yields in each bin of each category, while NPs corresponding to the systematic uncertainties are represented with the Gaussian constraint functions. The majority of uncertainties: those related to the signal modelling (JES, JER, b-tagging and trigger SFs, theory) and the background shape systematic, are correlated across the two $|\Delta\eta_{HH}|$ categories, which means that only one NP of each type is used in the fit. The bootstrap uncertainty is used as an uncorrelated NP, which means that for each $|\Delta\eta_{HH}|$ category, independent NPs are introduced. Since the bootstrap uncertainty describes the statistics-related uncertainty, it is constrained using the Poisson constraint instead of the Gaussian constraint.

Table 8.2: List of all systematics NPs used in the fit. Since these NPs depend on the kinematic variables such as p_T or η no single values are given and the uncertainties are evaluated by generating two NP histograms with $+1\sigma$ and -1σ .

Type of Uncertainty	Name of NP	
JES NPs	JET_EffectiveNP_Detector1	
	JET_EffectiveNP_Detector2	
	JET_EffectiveNP_Mixed1	
	JET_EffectiveNP_Mixed2	
	JET_EffectiveNP_Mixed3	
	JET_EffectiveNP_Modelling1	
	JET_EffectiveNP_Modelling2	
	JET_EffectiveNP_Modelling3	
	JET_EffectiveNP_Modelling4	
	ET_EffectiveNP_Statistical1	
	ET_EffectiveNP_Statistical2	
	ET_EffectiveNP_Statistical3	
	ET_EffectiveNP_Statistical4	
	ET_EffectiveNP_Statistical5	
	ET_EffectiveNP_Statistical6	
	JET_EtaIntercalibration_Modelling	
	JET_EtaIntercalibration_NonClosure_2018data	
	JET_EtaIntercalibration_NonClosure_highE	
	JET_EtaIntercalibration_NonClosure_negEta	
	JET_EtaIntercalibration_NonClosure_posEta	
	JET_EtaIntercalibration_TotalStat	
	JET_BJES_Response	
	JET_Flavor_Composition	
	JET_Flavor_Response	
	JER NPs	JET_JER_EffectiveNP_1
		JET_JER_EffectiveNP_2
JET_JER_EffectiveNP_3		
JET_JER_EffectiveNP_4		
JET_JER_EffectiveNP_5		
JET_JER_EffectiveNP_6		
JET_JER_EffectiveNP_7		
JET_JER_EffectiveNP_8		
JET_JER_EffectiveNP_9		
JET_JER_EffectiveNP_10		
JET_JER_EffectiveNP_11		
JET_JER_EffectiveNP_12restTerm		
JET_Pileup_OffsetMu		
JET_Pileup_OffsetNPV		
JET_Pileup_PtTerm		
JET_Pileup_RhoTopology		
JET_PunchThrough_MC16		
JET_SingleParticle_HighPt		

Table 8.3: Table 8.2 continued.

Type of Uncertainty	Name of NP
Flavour Tagging NPs	FT_EFF_Eigen_B_0
	FT_EFF_Eigen_B_1
	FT_EFF_Eigen_B_2
	FT_EFF_Eigen_C_0
	FT_EFF_Eigen_C_1
	FT_EFF_Eigen_C_2
	FT_EFF_Eigen_C_3
	FT_EFF_Eigen_Light_0
	FT_EFF_Eigen_Light_1
	FT_EFF_Eigen_Light_2
	FT_EFF_Eigen_Light_3
	FT_EFF_extrapolation
	FT_EFF_extrapolation_from_charm
	Trigger SF NPs
TRIG_HLT_Jet	

8.4.1 Post-fit Background Distribution

Figure 8.1 shows the m_{HH} distributions in the signal region in the two $|\Delta\eta_{HH}|$ categories after the background-only profile likelihood fit is performed. The expected background is evaluated using the best-fit values of NPs from the background-only fit to the data. The HH signal shapes for two coupling points, the SM point and $\kappa_{2V} = 0$ (with κ_λ and κ_V set to 1), are overlaid. Each signal shape is rescaled by an integer for a better comparison with the data. The predicted background from the 2b to 4b reweighting is shown in yellow, while the 4b data points with the Poisson uncertainty are shown in black. The bottom panel shows the ratio of data to the prediction with the total background statistical and systematic error indicated as a grey hatched area. The data agree well with the background-only hypothesis and no excess is seen throughout in either category. The HH signal for the $\kappa_{2V} = 0$ point can be compared to the SM signal. As discussed in Section 2.2.4, for the SM point, the cancellation effects between the two Feynman diagrams influence the very low number of the expected signal events. The cancellations do not occur for the $\kappa_{2V} = 0$ signal and therefore the signal yields obtained from simulations are greater for $\kappa_{2V} = 0$ signal as compared to the SM signal. As a result, this analysis is more sensitive to the non-SM values of κ_{2V} . The summary of the number of the observed 4b events and the number of the expected background and the SM signal (from simulation) events is shown in Table 8.4.

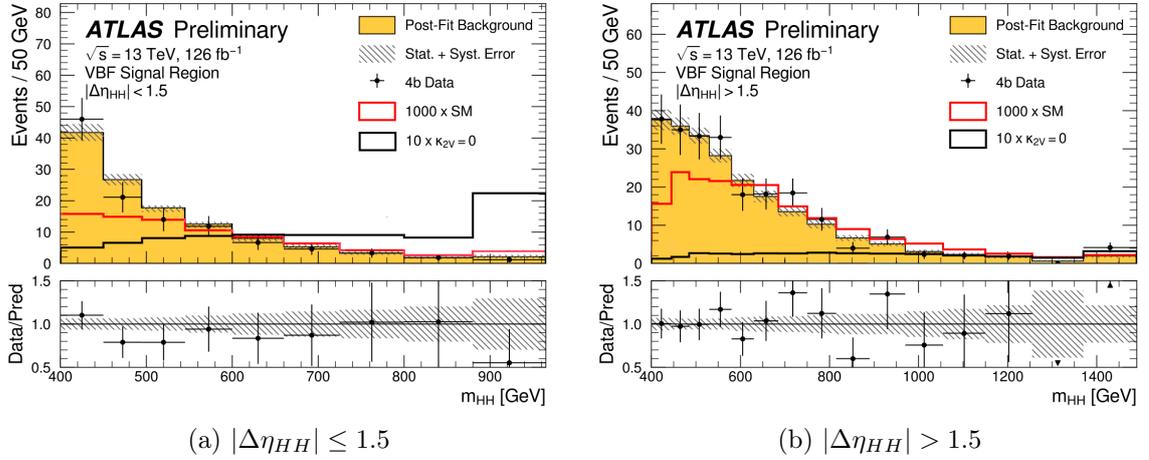


Figure 8.1: m_{HH} distributions in both $|\Delta\eta_{HH}|$ categories for data (black points) and expected background (yellow histogram) after background-only profile likelihood fit is performed. Total statistical and systematic background uncertainties are marked as grey hatched area, and statistical uncertainties in 4b data are shown as black error bars. Best-fit values of the nuisance parameters in the background-only fit to the data are used to obtain the expected background distribution. The Standard Model and the $\kappa_{2V} = 0$ signal shapes are overlaid and scaled by integers for the visibility on the plots. The lower panel shows the ratio of the 4b data to the reweighted 2b background prediction.

Table 8.4: Counts of the observed and the expected events in the two $|\Delta\eta_{HH}|$ categories of the 4b signal region.

Category	Data	Expected Background	Expected VBF SM signal
$ \Delta\eta_{HH} \leq 1.5$	116	125 ± 12	0.09
$ \Delta\eta_{HH} > 1.5$	241	231 ± 20	0.21

Table 8.5: Observed and expected 95% confidence level upper limit on the signal cross-section for the SM point in units of the theoretical Standard Model cross-section. The 1 and 2 standard deviation error bands are quoted as well. The result includes all background and signal uncertainties.

Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
132.3	71.3	95.7	132.8	192.0	277.9

8.4.2 95% CL Upper Limits

Figure 8.3 shows the observed and expected upper 95% CL_s limits on the signal strength μ as a function of the κ_{2V} values. The red line represents the theoretical prediction, which together with the observed limit is used to establish the excluded κ_{2V} values. If the observed (expected) limit on the cross-section is below the theoretical prediction line, the corresponding κ_{2V} value is excluded at 95% CL. The interval of the observed values of the κ_{2V} which are not excluded is [0.09, 1.99], while the expected interval is [-0.08, 2.16].

Table 8.5 presents observed and expected upper 95% CL_s limits on the signal strength $\mu = \sigma_{\text{VBF}}/\sigma_{\text{VBF}}^{\text{SM}}$. The measured SM cross-section is factor 132.3 higher than the theoretical prediction. The discrepancy between the prediction and the measured limit can be explained by the worse acceptance times efficiency value at the SM point compared to the non-SM coupling points shown in Section 6.9.

The results include all background and signal modelling uncertainties. The background uncertainties have the largest impact on the final result which is around 10%, with the exact value depending on the κ_{2V} point considered, based on the studies detailed in [6]. The normalized pulls of all systematic uncertainties are calculated as a ratio of the difference of the pre-fit and the post-fit NP value over the pre-fit NP value itself. The pull equal to 0 means that the pre-fit value of the systematic is correct. Figure 8.2 shows pulls of all NPs, in alphabetical order, after the fit is performed. All pulls are of value 0, meaning that their pre-fit values were correctly estimated.

Compared to the previous non-resonant 95% CL observed (expected) upper limits set on the VBF SM cross-section of 840 (550) times the SM cross-section prediction [3], the 6-fold improvement is observed coming purely from the analysis methods as the total integrated luminosity of 126 fb^{-1} was used for both analyses.

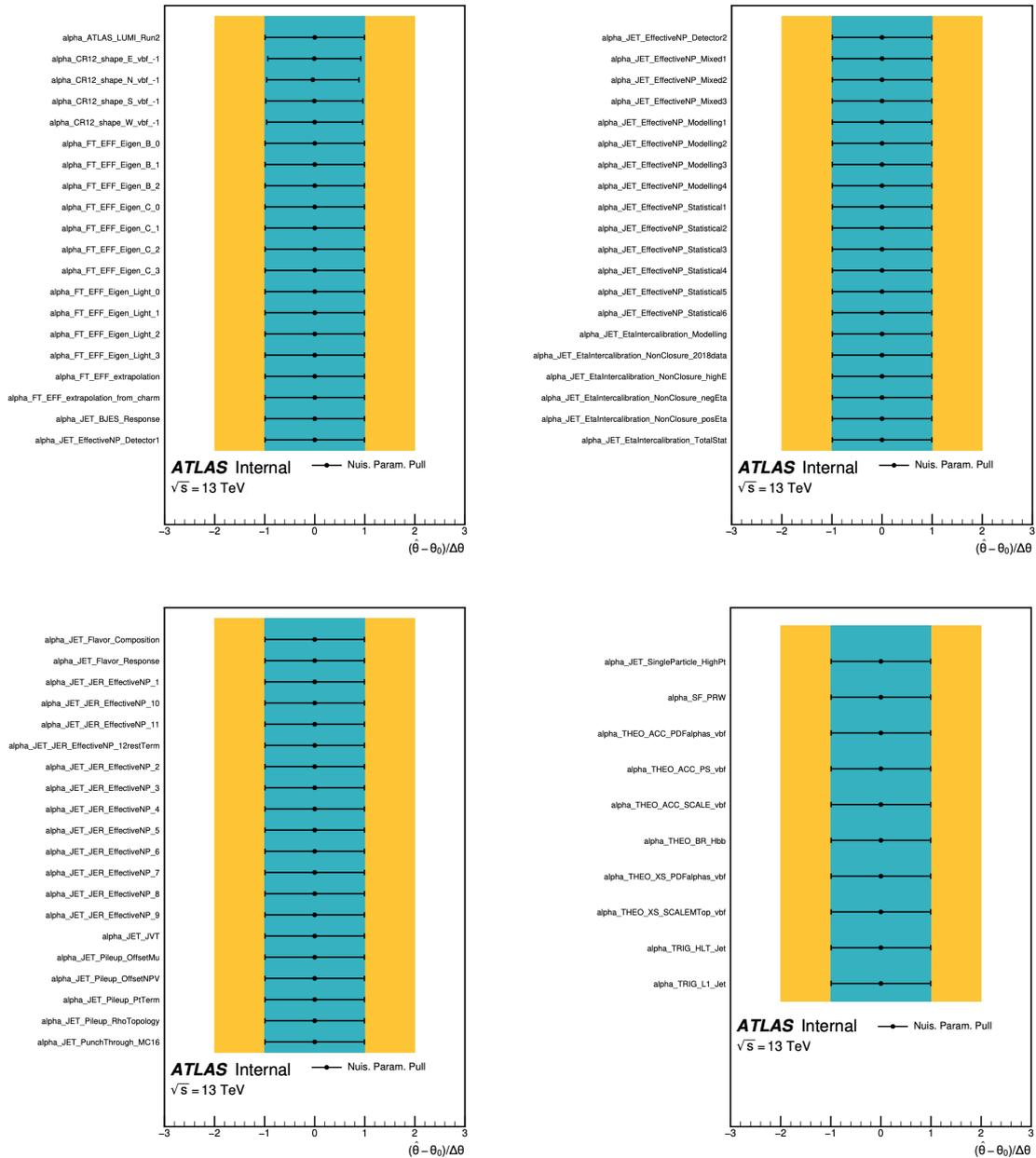


Figure 8.2: Pull plots for all systematic uncertainties NPs used in the analysis. The NPs are correlated across categories therefore there are no separate NPs for the two categories. The pulls order is alphabetical.

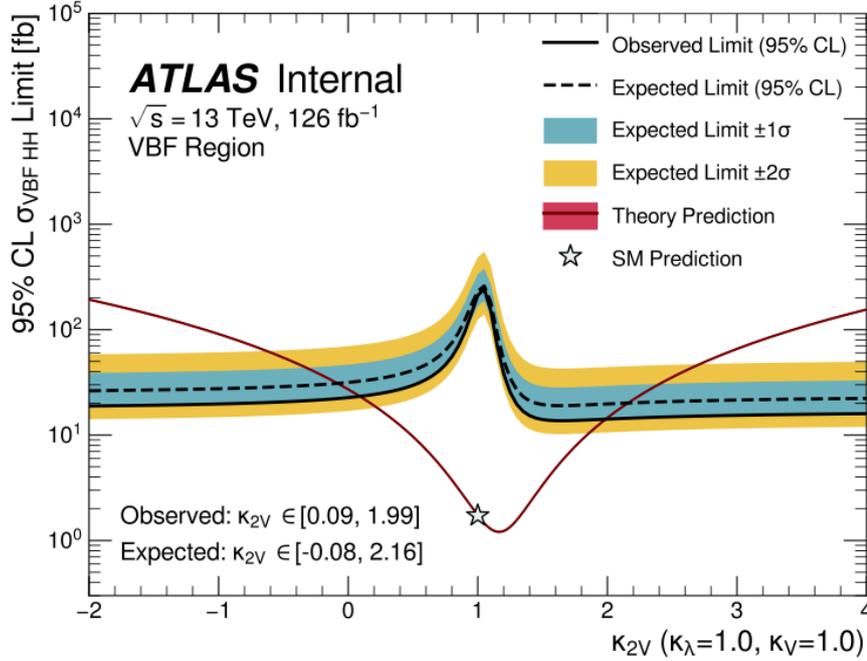


Figure 8.3: Observed and expected 95% CL upper limits on the VBF cross-section as a function of κ_{2V} . The observed limit is indicated with the solid black line, while the expected limit is indicated with the dashed black line. Blue and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty on the expected limit. The theory cross-section prediction is shown in red with the cross-section uncertainty included, but its width is smaller than the width of the plotted line. The SM point is indicated with a star. The observed (expected) 95% CL upper limit on the SM cross-section is 228.9 (232.7) fb which corresponds to 132.3 (132.8) times the SM theory prediction. The observed (expected) non-excluded κ_{2V} interval is $[0.09, 1.99]$, $[-0.08, 2.16]$ when only the VBF signal is included.

Chapter 9

HL-LHC and FCC Projection Studies

This chapter presents a set of simplified projection studies. The simulated datasets, produced assuming the future collider experimental conditions of the HL-LHC and the FCC-hh, are explored in order to understand what could be achieved with the larger centre of mass energy and luminosity, and the improved detector efficiencies compared to the ATLAS Run 2 conditions.

9.1 Overview of the Considered Scenarios

The LHC Run 3 is scheduled to end in 2025. Therefore, a lot of effort has already been put into what is coming next. The current LHC research programme will continue with the High Luminosity LHC (HL-LHC) which is scheduled to run between 2029 and 2040 at 14 TeV centre of mass energy and delivering the total integrated luminosity of 3000 fb^{-1} [4]. Furthermore, given the complexity of any accelerator experiment, the post-HL-LHC era plans have already started. One of the options is the Future Circular Collider (FCC-hh) [5]. Its basic properties, most importantly the centre of mass energy of 100 TeV and the total integrated luminosity of $30\,000 \text{ fb}^{-1}$, constitute the conditions used in the second futuristic scenario considered in this chapter. Additionally, ATLAS Run 2 - like scenario is used as a test and validation of the simplified method compared to the main analysis described earlier in the thesis. The list of important, especially in the context of this study, properties for all scenarios is shown in Table 9.1.

Table 9.1: Overview of the beam and detector properties, b-tagging efficiency and the pre-selection cut for the three scenarios considered in this chapter.

	ATLAS Run 2	HL-LHC	FCC-hh
\sqrt{s} [TeV]	13	14	100
Luminosity [fb^{-1}]	139	3000	30000
Calorimeter Granularity ($\eta \times \phi$)	0.1×0.1	0.025×0.1	0.025×0.025
Tracking Radius [m]	1.2	1.2	2
B-field [T]	2	2	4
η range	4.5	4.5	6
b-tagging efficiency [%]	77	80	80
Jet p_T cut for the MC BG sample [GeV]	10	10	20

9.2 Methods

9.2.1 MC Production Prescription

For each of the three scenarios, both signal and background samples, containing 0.5 and 10 million events respectively, are generated. The matrix element generation at the leading order (LO) is done with MADGRAPH [63] after which events are showered using PYTHIA8 [90]. The scaling factors to account for the lack of matrix element generation at higher orders are discussed in Section 9.2.5.

Signal Samples

The signal samples are produced following the ATLAS recommendations [91], as discussed in Section 6.2. The consistency between the main ATLAS analysis and the analysis presented here is thereby maximized.

Only the SM signal is studied here, meaning that κ_λ , κ_{2V} and κ_V are set to 1. The process is generated in such a way that the contamination from the VHH process is avoided by explicitly excluding diagrams with a photon exchange and an on-shell vector boson¹. For the $\sqrt{s} = 100$ TeV sample, the vertex displacement acceptance is altered to 10 m from the 1 m set by default for 13 and 14 TeV cases, as with the much higher centre of mass energy the b-jet can travel a longer distance from the primary vertex before decaying. No cuts are set on any kinematic variables.

¹The used MADGRAPH syntax: "generate p p > h h j j \$\$ z w+ w- /a j QED =4".

Background Samples

Since background used in the main ATLAS analysis is fully data-driven, it can not serve as the basis for the projection studies in this chapter. Instead, the simulated datasets are also used for the background, despite the limitations discussed in Section 2.2.6.

Generator-level di-Higgs VBF studies have been performed, in which a number of Monte Carlo QCD background channels are considered, for example, in [25]. It has to be highlighted that, at the time being, the QCD process in which six jets are generated directly is not computationally beneficial, although in principle possible, for example with the ALPGEN generator [92], which allows for the generation of up to five jets [93]. Therefore, to study the six-jet QCD process, the four-jet final state is first generated and the remaining jets are the product of a parton showering.

From the number of background channels discussed in [25], the $4b$ and $2b2j$ background samples contribute most significantly. Initially, both samples are produced for this study, but the $2b2j$ sample does not imitate the di-Higgs background well, meaning that no di-Higgs-like events are reconstructed. Therefore, only the $4b$ QCD background sample is considered.

MADGRAPH version 2.6.7 [63] is used for the event generation. In order to set the PDF set, the LHAPDF6 [94] is interfaced with MADGRAPH. The chosen PDF set is NNPDF3.0NLO [95] with $\alpha_s = 0.118$ and the flavour scheme of 4. The re-normalization and factorization scales are set to $\frac{H_T}{2}$, where H_T is the total sum of the transverse energies of all final state partons. The generated Les Houches event (lhe) file is showered using PYTHIA8, version 8.307 [90].

The pre-selection cut on the quark p_T is applied in order to efficiently generate interesting events for the analysis. The p_T pre-selection is set to 10 GeV for the ATLAS Run 2 and HL-LHC scenarios, and to 20 GeV for the FCC scenario. The cut is higher for the FCC scenario as with the increased centre of mass energy, the p_T of the interesting events is also increased.

9.2.2 Calorimeter Emulation, Jet Clustering and b-tagging

From the generator level consideration, we move to the detector level studies, where objects such as jets, and b-jets in particular, have to be reconstructed. The calorimeter emulation is used to emulate the energy deposition instead of the full detector simulations that exist for the current experiments, such as GEANT4 [66]. The energy depositions are clustered into jets using the anti- k_T FASTJET jet clustering algorithm [96] and are b-tagged using a simplified b-tagging approach.

Calorimeter Emulation

The calorimeter towers can be represented on a two dimensional plane as a grid of $\eta \times \phi$ of a certain dimension. Following the design reports, the calorimetry starts at a radius of 1.2 m in ATLAS [36] and HL-LHC [4], and of 2 m for the FCC [5]. The surrounding magnetic field is 2 T for the first two scenarios, and it is 4 T for the FCC. The η range is set to 4.5 for ATLAS and HL-LHC, while it is extended to 6 for the FCC. The η distributions for the four b-quarks are shown in Figure 9.1 for the signal and background samples, for both HL-LHC and FCC scenarios. Importantly, with the $\eta > 6$ cut for the FCC scenario, not much signal is lost. In the case of the HL-LHC scenario, only a minimal amount of signal, for the fourth leading b-quark, is lost with the $\eta > 4.5$ cut.

Depending on the considered calorimeter system, $\eta \times \phi$ dimensions, the granularity of the detector, can differ resulting in a better or a worse resolution of the energy deposition. Generally, with technological developments, more granular calorimeters will be possible in the future. The granularity of 0.025×0.1 is used for the HL-LHC and 0.025×0.025 for the FCC, while the nominal ATLAS granularity of 0.1×0.1 is used for ATLAS Run 2 - like scenario. Additionally, energy smearing of 50% is applied to the jet energies in each of the $\eta - \phi$ cells, to account for detector resolution effects. The same calorimeter emulation algorithm is used for all three scenarios. The summary of the discussed parameter values can be found in Table 9.1.

Jet Clustering

Using the energy depositions obtained from the calorimeter emulation, jets are clustered with the FASTJET package [96]. The anti- k_t algorithm is employed with the jet radius of 0.4 following the resolved ATLAS analysis.

b-tagging and *c*-mistagging

The last part of the object reconstruction method is the flavour tagging of the clustered jets. As described in Chapter 5, the present-day flavour tagging algorithms are extremely efficient, as they hugely benefit from machine learning techniques. In this simplified study, a simple b-tagging algorithm is used. It is defined such that if a truth b quark of $p_T > 15$ GeV is within the 0.4 radius of the reconstructed jet, then this jet is b-tagged. In order to follow the efficiencies of the current algorithms, the probability of such tagging is set to 77% for the ATLAS Run 2 - like scenario and 80% for the other two cases. It means that from all identified b-tagged jets, 23% (20%) are randomly rejected. The c-mistagging works in a very similar, but reverse, way. Firstly, the c-tagging is performed, if a c-quark is within the 0.4 radius of the jet, the jet is c-tagged. The c-mistagging means that some c-tagged jets are identified as the b-tagged jets instead of the c-jets they really are. In the studies

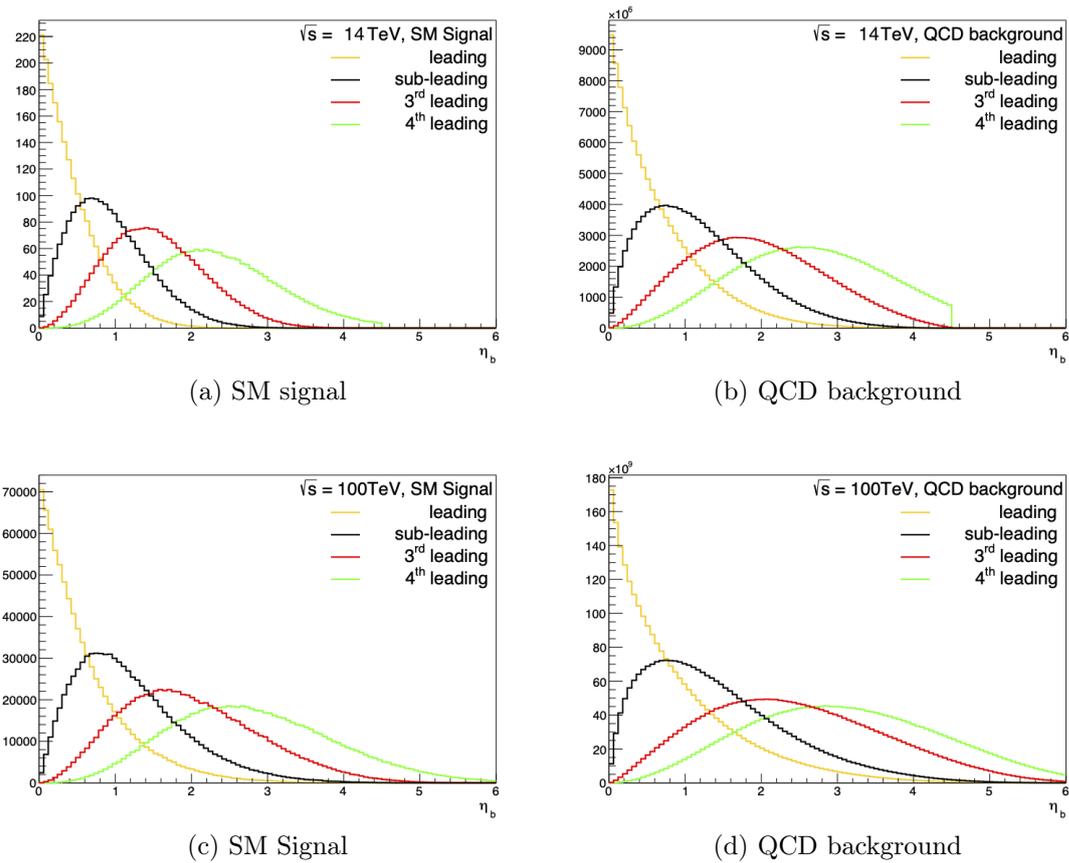


Figure 9.1: η distributions for both signal and background samples for HL-LHC $\sqrt{s} = 14 \text{ TeV}$ scenario (top row) and for the FCC scenario, $\sqrt{s} = 100 \text{ TeV}$ (bottom row).

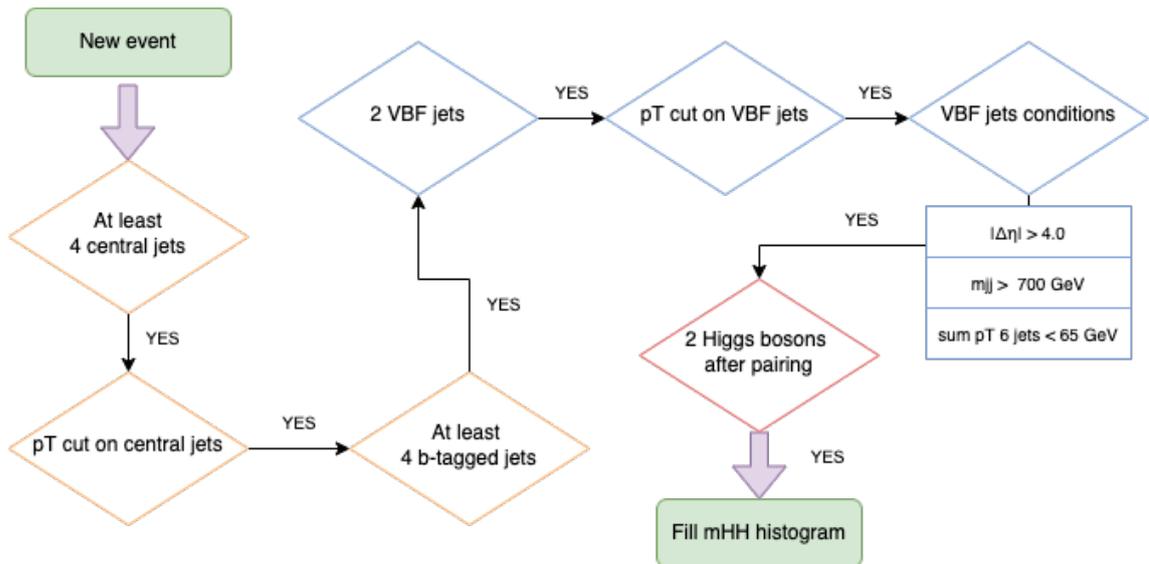


Figure 9.2: Flowchart of the analysis selections for all signal and background events for all scenarios.

performed here, the c -mistag rate is considered as 10%. In the same way, but to a significantly lesser extent, the light jet mistag rate is accounted for as 1%.

9.2.3 Simplified VBF Analysis Selection

After the object reconstruction, the analysis selections are performed. For a clearer overview, the flowchart of the analysis selections is shown in Figure 9.2. The steps are described below.

Object Selection

The analysis objects are chosen first. These are the central and the forward jets. The first requirement is that events must have at least four central ($|\eta| < 2.5$) b -tagged jets. Since the background samples are produced with the p_T pre-selection cut of 10 GeV (20 GeV for the FCC scenario), the same pre-selection cut is applied to the signal events. The central jet p_T distributions for the signal and background samples for the FCC scenario after the pre-selection cut are shown in Figure 9.3. The p_T distributions of the signal jets are much harder compared to the background jets' p_T distributions.

The second requirement is that events must have at least two, non- b -tagged forward jets, with $|\eta| > 2$. The same p_T pre-selection is applied to the forward jets. Out of all forward jets, the two highest p_T jets are chosen as the initial scattered jets.

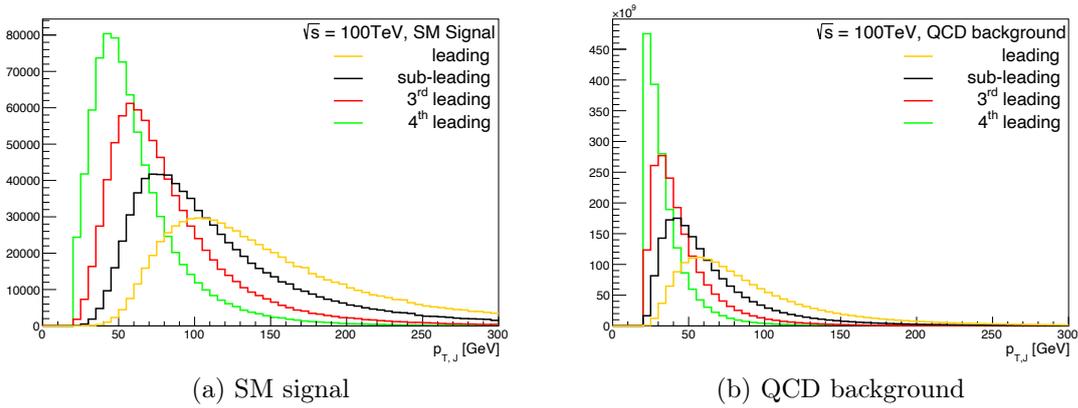


Figure 9.3: Central jets p_T distributions after the pre-selection p_T cut is applied to all jets for the FCC scenario.

Event Selection

Not all events with four central b-jets and two forward jets are the di-Higgs VBF events of interest. The event selection proceeds by requiring the VBF-specific conditions. Namely, to ensure that the forward jets are in the opposite direction, $\Delta\eta_{jj}$ is set as greater than 4, while the total invariant mass of the two VBF jets is greater than 700 GeV.

In the main analysis, $\Delta\eta_{jj}$ is set as greater than 3. As discussed in Section 6.4, such separation requirement allows to choose the VBF jets from both the non-b-tagged central and forward jets; the selection of the VBF jets pair in which one jet is central and one jet is forward is possible. The total number of the selected VBF events is thereby increased, as studied in more detail in [6]. The approach is altered here for simplicity.

In the main analysis, the invariant mass of the VBF jets condition, $m_{jj} > 1000$ GeV, is much tighter compared to the m_{jj} selection here. The looser cut allows to efficiently study the MC generated signal and background samples. If too tight cuts are set, very few background events are selected and no conclusions can be drawn due to very low sample statistics.

The third VBF condition, the central jet veto, is introduced to reject events with the significant kinematic activity in the central η region as this is not a characteristic of the VBF process. The requirement is that the vector sum of p_T of four b-tagged jets and the two VBF jets is less than 65 GeV. The condition directly follows the condition used in the main analysis.

The distributions of the three variables: m_{jj} , $\Delta\eta_{jj}$ and the vector sum of six jets p_T are shown in Figure 9.4. The vertical lines indicate the position of the cut and the arrows indicate which events are kept. Note that the background events are scaled

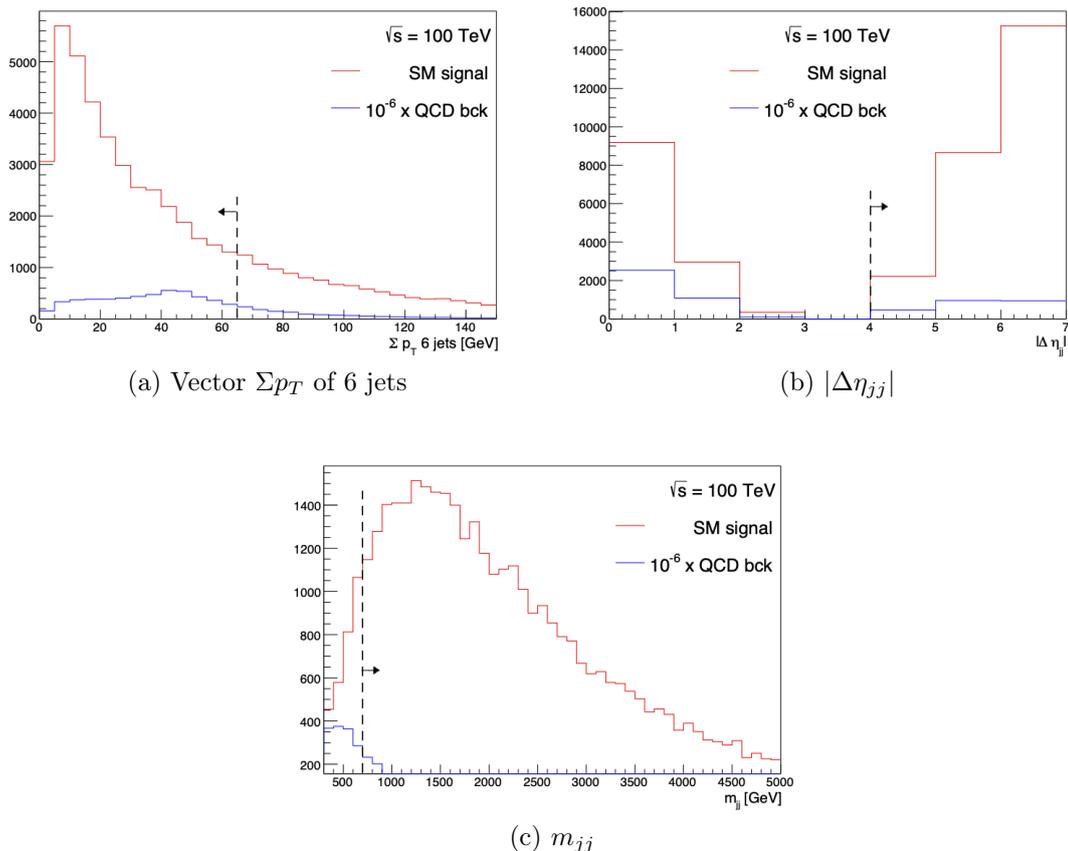


Figure 9.4: The VBF variables distributions for the FCC scenario for both signal and background samples. Cuts, as shown by the vertical lines with the arrows indicating which events are kept, are performed on these variables in order to select the VBF events.

by a factor of 10^{-6} in order to compare their shapes to the shapes of the signal samples. Especially for the m_{jj} cut, it can be clearly seen that a higher cut would remove almost all of the background events, making it impossible to perform the study.

Higgs Candidate Pairing

As was discussed in Section 6.5, many pairing algorithms are possible. A simplified pairing method is chosen here.

Three different pairings of Higgs boson candidates are possible from four b-jets. An event is accepted if at least one of these pairings results in the two Higgs Candidates whose mass differs by no more than 30 GeV from the SM Higgs boson mass, taken

here as 125 GeV. The higher p_T reconstructed Higgs boson is the leading Higgs boson, H_1 , while the other is the sub-leading Higgs boson, H_2 .

All three combinations are represented on the two-dimensional Higgs boson mass-planes, shown for the FCC scenario in Figure 9.5. It can be observed that the third pairing is chosen by the algorithm most often, i.e. the pairing in which the leading Higgs boson is built from the first and the fourth leading b-jets, while the sub-leading Higgs boson is built from the second and the third b-jets.

A circular area centred around 125 GeV indicates the presence of the most Higgs boson pairs in the signal sample plots (left column), while the background samples have a much more random distribution of the events. If more than one pairing combination is possible for a given event, the combination in which both Higgs candidates are closest to the SM Higgs mass is chosen.

9.2.4 Cut-flows

The impacts of the individual selection steps are presented in cut-flows. The bare numbers of the remaining events after each selection, without accounting for any normalization factors, are shown in Tables 9.2 - 9.5. Tables 9.2 and 9.3, present selections for the category with the pre-selection p_T cuts only. Tables 9.4 and 9.5 include tighter cuts on the VBF jets, namely, 30 GeV for the first two scenarios and 50 GeV for the FCC, the p_T threshold for the b-jets remains unchanged. The number of the bare background events which are left after all selection cuts is 7 for the HL-LHC scenario and 10 for the other two cases. Tightening the cuts further would result in a category with very few background events, 1 or 2, which means too much statistical uncertainty fluctuation for a reliable result.

Looking at the signal cut flows comparison from Table 9.2, one can note that the actual number of the bare events, after the selection of four b-jets, is lower for the FCC scenario than for the ATLAS Run 2 and HL-LHC cases. This can be explained by the increased boosted regime activity at 100 TeV, which is illustrated in Figure 9.6. Peaks at around 125 GeV are clearly visible for the first two leading jets. In that case, not one but two b-quarks lie within the required $R = 0.4$ cone making it a *doubly* b-tagged jet, which is not recognised in the process of this resolved analysis. The *doubly* b-tagged jet is still counted as a singly b-tagged jet, resulting in a loss of events which do not have the total of the required 4 b-tagged jets. As expected, this does not apply to the background events. It is an indication that performing a boosted analysis may be crucial for using the full potential of the FCC data, however, this is beyond the scope of this thesis.

In order to compare the general trend of the events reduction between signal and background samples, we compare Tables 9.2 and 9.3 (or 9.4 and 9.5). For signal samples, around 30% of events make it through the four b-jets selection, while for the background samples, only between 2 – 5% of events pass this selection. For

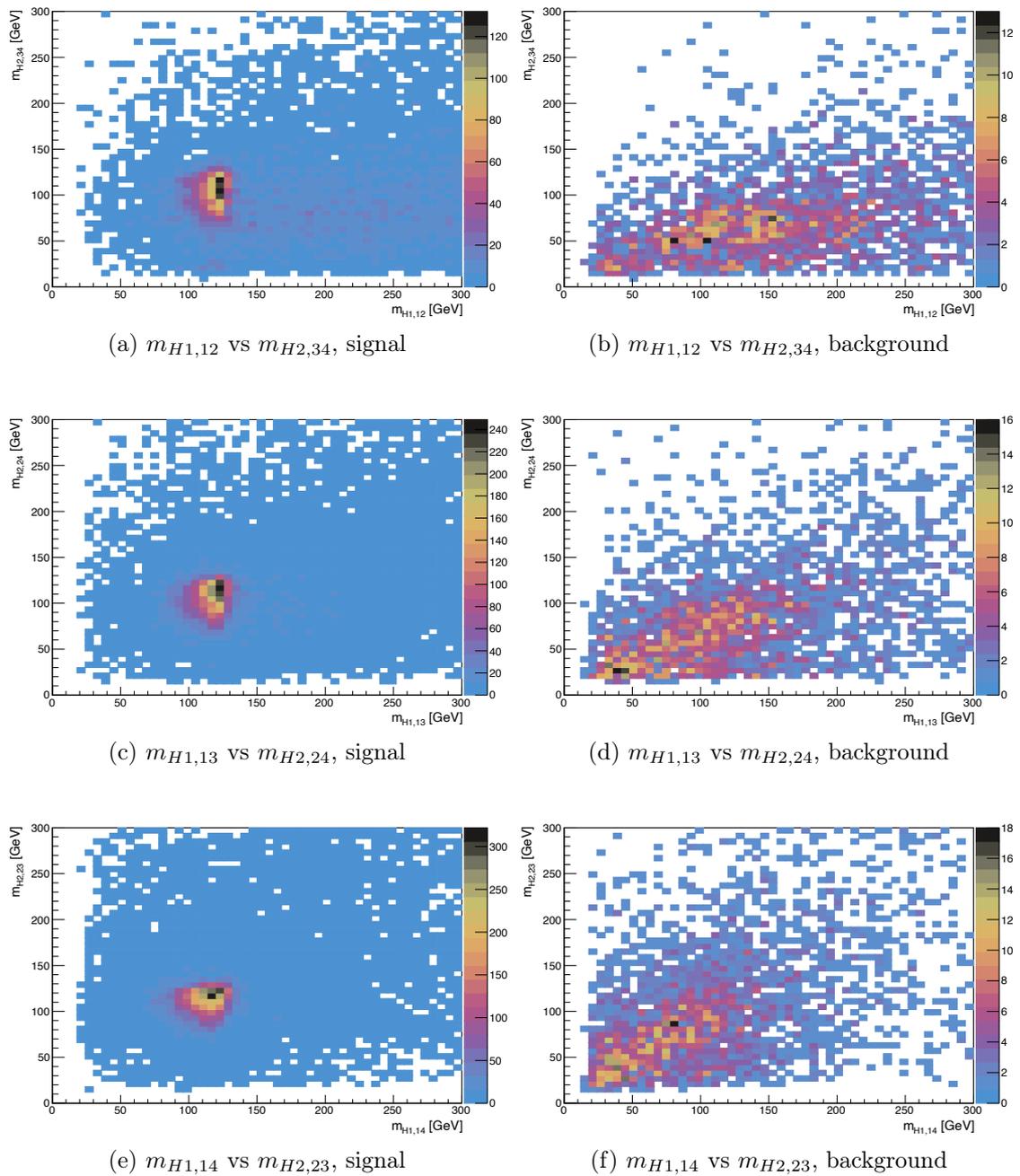


Figure 9.5: Massplanes of the leading and sub-leading Higgs Candidates for the three possible pairing combinations for the FCC scenario. Left column shows results for the signal sample, while the right for the background.

the signal samples in Tables 9.2 and 9.4, after the VBF selections, around 4 – 8% of events are left, while for the background in Tables 9.3 and 9.5, 0.01 – 0.05% of events are left. Out of the selected VBF events, in the case of the signal samples, around 56 – 68% of events pass the Higgs pairing algorithm, while for the background sample, only 11 – 18% of events pass the Higgs Candidate pairing. Thus it can be stated that the analysis selection method is at its limit in terms of selecting VBF-like events in the background sample.

 Table 9.2: Cut-flows, pre-selection p_T cuts: 10 (20) GeV for all jets, signal samples.

	ATLAS Run-2 like	HL-LHC	FCC-hh
4 jets in event	499 685	499 628	489 461
4 b-jets, no p_T , no η requirement	134 418	136 661	122 826
4 b-jets with p_T requirement	80 343	90 263	49 873
2 VBF jets	56 969	64 642	32 421
η requirement VBF jets	46 349	52 399	24 946
m_{jj} VBF >700 GeV	36 873	42 360	23 324
Vector $\sum p_T$ of 6 jets <65 GeV	33 812	38 674	18 684
Two Higgs bosons per event	19 033	22 056	12 803

 Table 9.3: Cut-flows, pre-selection p_T cuts: 10 (20) GeV for all jets, background samples.

	ATLAS Run-2 like	HL-LHC	FCC-hh
4 jets in event	6 413 703	6 590 928	6 946 482
4 b-jets, no p_T , no η requirement	167 804	204 804	299 268
4 b-jets with p_T requirement	102 240	123 454	136 009
2 VBF jets	17 111	22 004	25 559
η requirement VBF jets	7 358	9 512	11 034
m_{jj} VBF >700 GeV	685	832	4 856
Vector $\sum p_T$ of 6 jets <65 GeV	579	691	3 227
Two Higgs bosons per event	66	81	567

Table 9.4: Cut-flows, best category p_T cuts: 10 (20) GeV for b-jets, 30 (50) GeV for VBF jets, signal samples.

	ATLAS Run-2 like	HL-LHC	FCC-hh
4 jets in event	499 685	499 628	489 461
4 b-jets, no p_T , no η requirement	134 418	136 661	122 826
4 b-jets with p_T requirement	80 343	90 263	49 873
2 VBF jets	30 267	34 643	13 784
η requirement VBF jets	27 223	31 133	11 633
m_{jj} VBF >700 GeV	25 146	28 994	11 587
Vector $\sum p_T$ of 6 jets <65 GeV	23 641	27 177	9 484
Two Higgs bosons per event	13 299	15 550	6 536

 Table 9.5: Cut-flows, best category p_T cuts: 10 (20) GeV for b-jets, 30 (50) GeV for VBF jets, background samples.

	ATLAS Run-2 like	HL-LHC	FCC-hh
4 jets in event	6 413 703	6 590 928	6 946 482
4 b-jets, no p_T , no η requirement	167 804	204 804	299 268
4 b-jets with p_T requirement 10	102 240	123 454	136 009
2 VBF jets	575	712	1640
η requirement VBF jets	193	219	534
m_{jj} VBF >700 GeV	109	125	487
Vector $\sum p_T$ of 6 jets <65 GeV	52	54	93
Two Higgs bosons per event	10	7	10

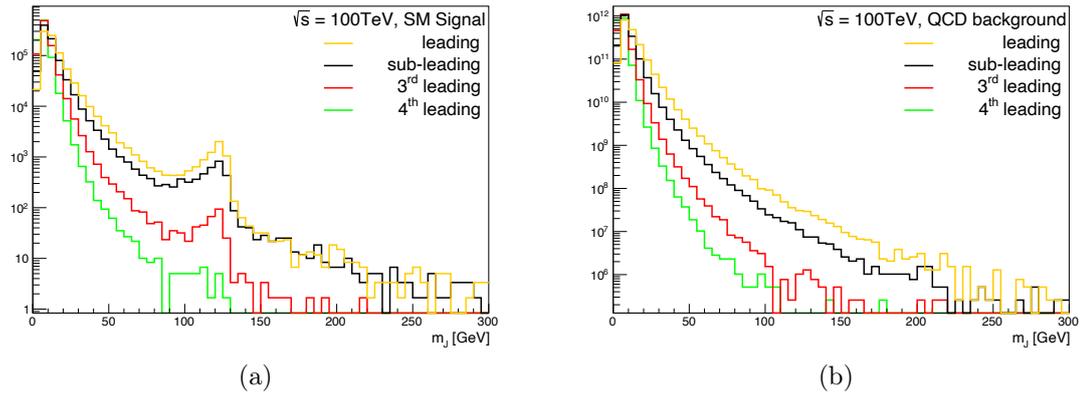


Figure 9.6: First four leading jets mass distribution for both (a) SM signal and (b) QCD background samples.

Table 9.6: Cut-flows comparison of the signal sample at 13 TeV for the calorimeter emulation and simplified analysis vs full simulation and full analysis.

	Calo emulation, simplified analysis [%]	ATLAS full simulation, full analysis [%]
Initial count	100.00	100.00
VBF selections	27.94	18.72
4 b-jets, $p_T > 10$ GeV	4.73	6.11
Higgs pairs	2.66	2.48

Brief Comparison to the Main ATLAS Analysis

The cut-flow analysis serves one more purpose, namely, assessing the agreement of the calorimeter emulation with the full simulation. Table 9.6 shows a cut-flow comparing the efficiency of cuts (percentage of total yield passing a given cut) performed on the signal events treated with the calorimeter emulation and simplified analysis, to the events treated with the full detector simulation, as well as the full ATLAS analysis. The signal sample for the full simulation was produced using the main analysis framework with the same pre-selection p_T jet cut of 10 GeV.

It can be stated that the calorimeter emulation method agrees to an acceptable extent with the full simulation. The biggest difference is for the VBF selection. However, as detailed in Section 9.2.3, the m_{jj} and $\Delta\eta_{jj}$ cuts differ between the two scenarios. The tighter m_{jj} cut removes a significant number of VBF events in the case of the main analysis. The final yield of events that contain Higgs pairs agrees within 10%.

9.2.5 Scaling Factors

Before evaluation of the final results, the scaling factors, the NLO-to-LO k -factors and cross-section normalization, are applied to the event counts².

k-factor Studies for the Background Samples

The k -factor studies are performed using the MG5@NLO framework [63] together with PYTHIA8 parton shower matching technique [90].

²Note that in all distributions presented earlier in this chapter, normalizations and scaling factors have already been applied for plotting, however, up until now they played no major role in the discussion.

The nominal calculation of the k -factor is:

$$k = \frac{\sigma_{NLO}}{\sigma_{LO}}, \quad (9.1)$$

where both σ 's are the total cross-sections. For both LO and NLO samples, 100 000 events are generated³, using the run parameters summarized in Table 9.1. The global k -factor of 2.7 is applied for all scenarios. The simplified global k -factor is chosen as no dependence on the kinematic variables such as jets p_T was established.

Normalization Calculation

The total cross-section is computed at the leading order in the generation process, as explained in Section 9.2.1. The normalization factor is applied to all background samples as:

$$norm = \frac{L \times \sigma_{LO} \times k}{N_{tot}}, \quad (9.2)$$

where L is the integrated luminosity, σ_{LO} is the leading order cross-section, k is a scaling factor between LO and NLO cross-sections and N_{tot} is the total number of the MC events used.

As discussed in Chapter 2, for the VBF signal the total N3LO cross-section is available [24]. Since only the 4b final state is considered, the total cross-section is multiplied by the 4b branching fraction, in the same way as in the main analysis, giving:

$$\sigma_{\text{signal}} = \sigma_{3\text{NLO}} \times BR = \sigma_{3\text{NLO}} \times 0.58^2. \quad (9.3)$$

The summary of the cross-sections pre- and post-scaling factor application is presented in Table 9.7.

³Since the NLO generation is very computationally expensive, the factor of 10 reduction in the study samples size is introduced.

Table 9.7: Cross-section values for all three scenarios considered in the projection studies. Calculation of the values used in the analysis for both signal and background events is outlined in the text.

	ATLAS Run-2 like	HL-LHC	FCC-hh
N_{sig}	500 000	500 000	500 000
Total σ_{sig} , N3LO [24] [fb]	1.726	2.055	82.84
σ_{sig} used in analysis [fb]	0.58	0.69	27.87
Norm factor (signal)	1.614×10^{-4}	4.148×10^{-3}	1.672
N_{BG}	9 990 000	10 000 000	9 971 039
σ_{BG} , LO [fb]	1.22×10^7	1.42×10^7	3.16×10^7
k-factor (background)	2.7	2.7	2.7
σ_{BG} used in analysis [fb]	3.29×10^7	3.83×10^7	8.53×10^7
Norm factor (background)	456	11 499	256 643

9.3 Results

9.3.1 Signal Significance Studies

Signal significance is a measure of how well a signal can be separated from a background. We have N total observed events, B expected background events and $S = N - B$ signal events. The signal significance z relates to the probability of S , given the null hypothesis:

$$p = \frac{1 - \text{erf}(z/\sqrt{2})}{2}, \quad (9.4)$$

where p is the p-value and $\text{erf}(\alpha)$ is the standard error function given by:

$$\text{erf}(\alpha) = \frac{2}{\sqrt{\pi}} \int_0^\alpha e^{-\beta^2} d\beta. \quad (9.5)$$

As was discussed in Chapter 8, the case here is a counting experiment, described by the Poisson distribution. Therefore, assuming a null hypothesis with only B events expected, z becomes:

$$z = \frac{N - B}{\sqrt{N}} = \frac{S}{\sqrt{N}} \approx \frac{S}{\sqrt{B}}, \quad (9.6)$$

with $\frac{S}{B} \ll 1$ and the known B . From the above equations, it can be inferred that the greater the significance, the smaller the p-value, meaning the probability of S being only a statistical fluctuation of a null hypothesis is very small for a large z .

As discussed in Section 6.10, event categorization in a kinematic variable allows for an improved total significance. The cumulative significance is computed as:

$$Z = \sqrt{\sum z_i^2}, \quad (9.7)$$

with z_i computed using Eq. 9.6.

Categorization

One of the main selections is the p_T cut on both central b-jets and the VBF jets. It was shown that depending on the p_T cut of the VBF jets, the number of background events can significantly decrease. Therefore, the categorization is performed in the p_T of the VBF jets primarily. Additionally, where statistically possible the categorization in the central jets p_T is performed as well. *Statistically possible* means that each category must have at least 5 bare events, the limiting factor being the statistics of the background events. The categories are exclusive and independent. For the proper computation of significances, the bare signal and background events, as from the cut-flow tables, are first multiplied by the scaling factors from Section 9.2.5. Tables 9.8 - 9.10 show the computed significances per p_T category.

The significances in each table are summed using Eq. 9.7. The results, compared with the significances computed without any p_T categorization, are shown in Table 9.11.

The highest, 4-fold, gain from the p_T categorization is obtained for the FCC scenario. For the ATLAS Run 2 - like scenario, it is a gain of factor 1.8, while for the HL-LHC, it is a factor of 2.4.

Table 9.8: Significance in categories for the ATLAS Run 2 - like scenario.

		VBF jet p_T [GeV]			
		10 – 20	20 – 30	>30	
b-jet p_T [GeV]	10 – 20	0.0005	0.0015	0.0094	0.0318
	20 – 30	0.001			
	>30	0.0019			

Table 9.9: Significance in categories for the HL-LHC scenario.

		VBF jet p_T [GeV]			
		10 – 20	20 – 30	>30	
b-jet p_T [GeV]	10 – 20	0.0022	0.0471	0.2273	
	20 – 30	0.0051			
	30 – 40	0.0062			
	>40	0.0044			

Table 9.10: Significance in categories for the FCC-hh scenario.

		VBF jet p_T bins [GeV]				
		20 – 30	30 – 40	40 – 50	>50	
b-jet p_T [GeV]	20 – 30	0.1498	0.2875	0.6074	6.8217	
	30 – 40	0.1673	0.5106	0.5424		
	40 – 50	0.1682	0.6906			
	>50	0.2609				

Table 9.11: Total significance for all scenarios. Comparison is made for the cases without and with the categorization (as shown in Tables 9.8 - 9.10 above). In the latter case, the total significance is calculated using Eq. 9.7.

	ATLAS Run-2 like	HL-LHC	FCC-hh
Z (no categorization)	1.77×10^{-2}	9.48×10^{-2}	1.77
Z (with categorization)	3.32×10^{-2}	2.32×10^{-1}	6.94

Evaluation of the Analysis Gain Factors

The analysis gain factors are evaluated for the higher energetic scenarios compared to the lower energetic scenarios. Namely, for the HL-LHC (FCC), the gain factor compared to ATLAS is evaluated as:

$$\text{GF c.f. ATLAS} = \frac{Z_{\text{HL-LHC (FCC)}}}{Z_{\text{ATLAS}}}, \quad (9.8)$$

and analogously, for the FCC compared to the HL-LHC,

$$\text{GF c.f. HL-LHC} = \frac{Z_{\text{FCC}}}{Z_{\text{HL-LHC}}}. \quad (9.9)$$

The results are presented in Table 9.12. For simplicity, the cross-sections are normalized to the luminosity. The left column shows the expected gain factors before the analysis is performed, based on the cross-sections from Table 9.7, row: $\sigma_{sig/BG}$ used in analysis. Based on the cross-sections only (before the analysis), the FCC result is 30 times better than the ATLAS Run 2 - like result. To assess the impact of the analysis, the post-analysis gain factors are evaluated in the right column. Post analysis, the FCC result is only 14 times better than the ATLAS result. This means that the analysis worsens the FCC performance by a factor of 2.

Performing the same comparison for the HL-LHC, before the analysis (noting no inclusion of luminosity), the gain factor compared to ATLAS is 1, meaning no gain, no loss, while after the analysis the gain factor is 1.5, meaning the analysis improves the HL-LHC performance by 50%.

In summary, the simplified analysis is optimised for the ATLAS Run 2 scenario in which the resolved channel is dominant. The analysis improves the significance of the HL-LHC scenario, but an ameliorated analysis technique for the FCC scenario is required. For instance, the boosted analysis should be explored.

Table 9.12: The significance gain factors (GF) for HL-LHC and FCC scenarios: before (left) and after (right) the analysis.

	Before Analysis			After Analysis		
	ATLAS	HL-LHC	FCC	ATLAS	HL-LHC	FCC
Z, $L = 1 \text{ fb}^{-1}$	1×10^{-4}	1×10^{-4}	3×10^{-3}	3×10^{-3}	4×10^{-3}	4×10^{-2}
GF c.f. ATLAS		1	30		1.5	14
GF c.f. HL-LHC			27			10

9.3.2 m_{HH} Distributions

No Categorization

Figure 9.7 shows the m_{HH} distributions for all three scenarios for the signal and background events when no categorization is performed. The width of bins is chosen as 20 GeV. The indicated error bars are the statistical Poisson errors, \sqrt{N} .

The peaks of the signal distributions are at around 300 GeV for the low energetic scenarios and at around 350 GeV for the FCC, as with the higher centre of mass energy, the constituent jets are also more energetic. All distributions demonstrate high background concentration in the low m_{HH} region.

Best Category

For each scenario, the total significance is driven by the highest p_T category, referred to as the *best category*. The further study focuses on the events from the best category alone.

Given only 7 - 10 bare background events pass selections in the best categories of each scenario, a proposed m_{HH} binning for the first two scenarios is 4 bins in the range of 0 - 1400 GeV. For the FCC scenario, the m_{HH} range of 0 - 1600 GeV is split into 5 bins. The m_{HH} distributions with coarse binning for all scenarios are shown in Figure 9.8.

There are no background events in the two highest m_{HH} bins of the ATLAS Run 2 - like and the HL-LHC m_{HH} distributions. In both cases, the second bin is extended to the maximum m_{HH} range to include the remaining signal events. For the FCC scenario, the three highest m_{HH} bins are added together instead. The final binning scheme is summarized in Table 9.13.

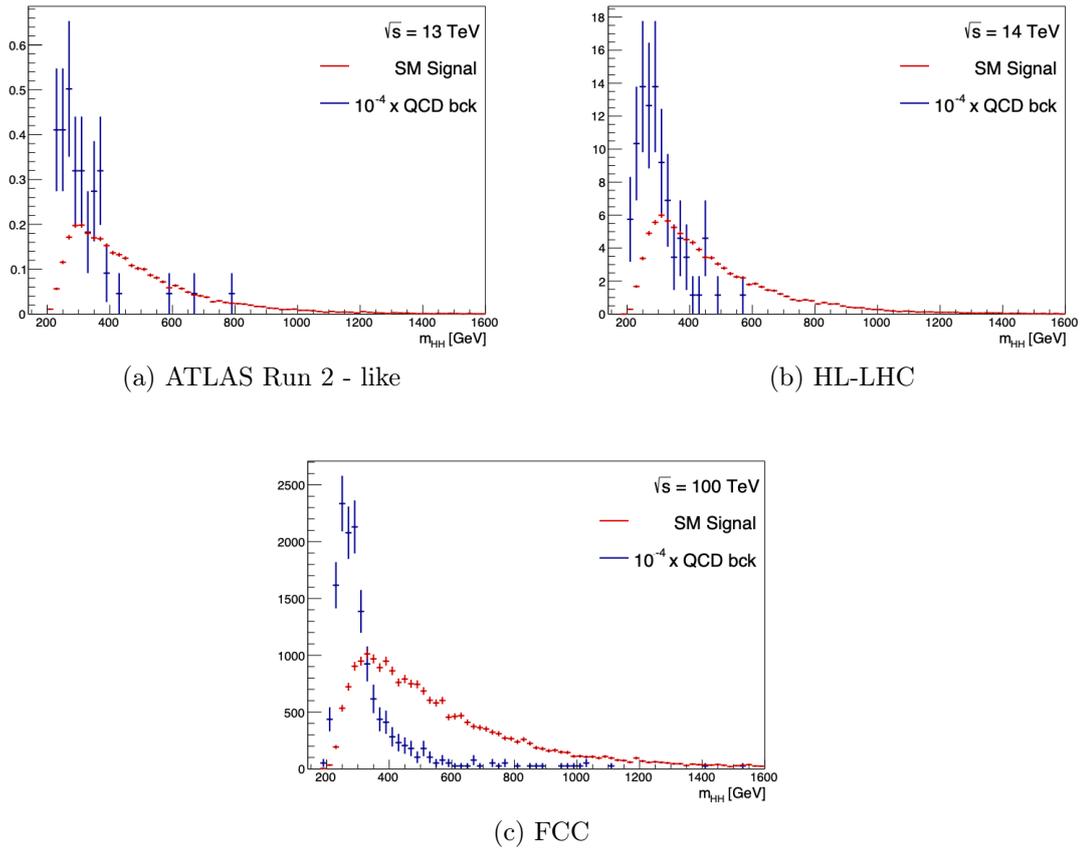


Figure 9.7: Normalized di-Higgs invariant mass distributions for ATLAS Run 2 - like, HL-LHC and FCC. Background distributions are downsampled by a factor of 10^{-4} for a better comparison.

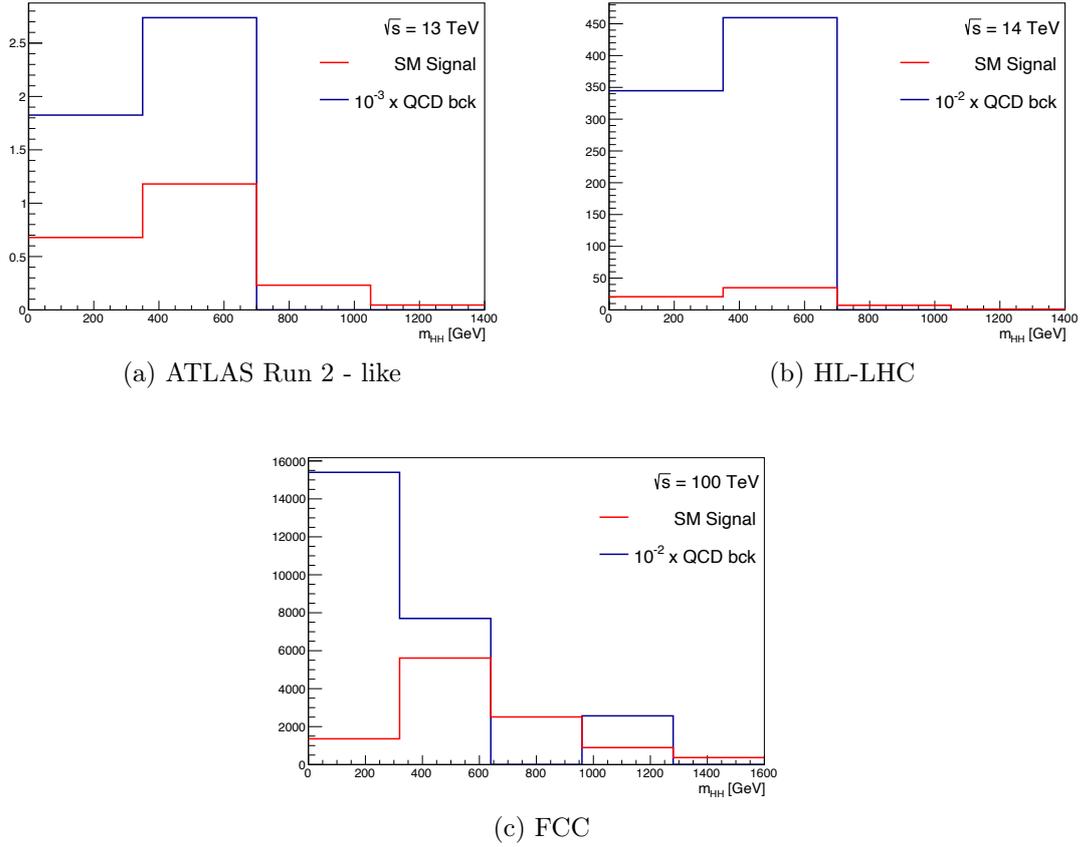


Figure 9.8: Normalized di-Higgs invariant mass distributions for all scenarios in the best p_T category (VBF jets cut at 30 (50) GeV) with coarse binning. Background samples are downsampled by a factor of 1000 (ATLAS Run 2 - like) or 100 (HL-LHC and FCC).

9.3.3 95% CL Expected Upper Limit

The limit-setting procedure is performed with the signal and background events from the best p_T category. In the main analysis, the hypothesis was tested based on the observed events against the expected counts of the signal and background events. Here, no real observed events are available and the analysis is based on the expected MC events only.

A systematic uncertainty is assigned to both signal and background events. Following [25], an approach is proposed in which the relative signal uncertainty is 10% and the relative background uncertainty is 15%. Both uncertainties are uncorrelated across the bins. It is a conservative treatment compared to the main ATLAS Run 2 analysis as the main systematic in the main analysis, the background modelling, is around 9%.

Table 9.13: Final binning scheme of the m_{HH} distribution. From the 4 (5) bins shown in Figure 9.8, only 2 (3) bins are populated with the background events. The unpopulated bins are added to the populated bins.

	Bin 1 [GeV]	Bin 2 [GeV]	Bin 3 [GeV]
ATLAS Run-2 like	0 - 350	350 - 1400	–
HL-LHC	0 - 350	350 - 1400	–
FCC-hh	0 - 320	320 - 640	640 - 1600

As in the main analysis, the PYHF framework [78] is used to set the expected 95% upper limits on the signal strength μ . Table 9.14 shows the expected limits for the three scenarios, evaluated in the best category. For the ATLAS Run 2 - like scenario and the HL-LHC, this is the category with the b-jets with $p_T > 10$ GeV and VBF jets with $p_T > 30$ GeV, while for the FCC the two cuts are 20 and 50 GeV respectively. For the first two scenarios events are split into two bins of the m_{HH} spectrum, while for the FCC scenario, events populate three bins.

Table 9.14: Limits on the signal strength μ . The comparison of the three scenarios considered in the projection studies to the final limit of the ATLAS Run 2 analysis is presented.

	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
ATLAS Run 2 Analysis	72	96	133	192	278
ATLAS Run-2 like	237	314	430	589	778
HL-LHC	137	182	248	340	448
FCC-hh	10	13	18	26	34

9.3.4 Discussion of Results

As can be seen in Table 9.14, the expected result coming from the full Run 2 ATLAS analysis is three times better compared to the result from the simplified analysis. However, given the number of simplifications throughout the analysis presented in this chapter, and also throughout the statistical treatment, this is understandable. The quoted ATLAS limit, as explained in Chapter 8, is based on two categories in $|\Delta\eta_{HH}|$, while here only the best category is taken. Moreover, for the main analysis, the number of bins in the m_{HH} spectrum which are populated is greater than in the case of this simplified method. The more granular binning allows for better signal to background discrimination and hence the improved limits. Lastly, the conservative treatment of systematics also has a non-negligible impact on the final result.

The more general limitations of the analysis method contain all the simplifications made throughout. Firstly, to a large extent, the use of the MC background simulation for the six jets in the final state process with the production of the only four-jet final state and the other jets coming from the showering procedure, and consideration of only one type of the QCD background, constitutes the weakest point of the procedure. As was shown throughout Section 9.2.3, the selections section, the tighter cuts on the kinematic variables were not possible due to the limiting number of background events. With the real experimental data available, this method will either be replaced by a data-driven technique, similar to what was done in the main Run 2 ATLAS analysis when data became available, or the more accurate MC calculations will be available for the multi-jet QCD background processes. Secondly, as explained earlier, the used reconstruction algorithms, namely the calorimeter emulation and the b-tagging, were simplified for the projection studies. Lastly, the boosted and semi-boosted regimes are not considered here, while they become much more important with the increasing centre of mass energy of the experiment.

Chapter 10

Conclusion

In this thesis, the non-resonant Vector Boson Fusion production of the Higgs boson pairs decaying to the $b\bar{b}b\bar{b}$ final state is studied. The analysis of the Run 2 data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV with the total integrated luminosity of 126 fb^{-1} results in the observed (expected) 95% CL upper limit on the SM cross-section of 132.3 (132.8) times the SM cross-section. This is a 6-fold improvement compared to the limit of 840 (550) times the SM cross-section set by the previous $HH \rightarrow b\bar{b}b\bar{b}$ VBF analysis [3], which used the same total integrated luminosity of 126 fb^{-1} . The κ_{2V} interval is constrained at [0.09, 1.99], compared to the interval of [-0.55, 2.72] from the previous publication [3], [26].

The search for the non-resonant Higgs pair production is motivated by the aspiration to know the shape of Higgs potential, leading to a statement on vacuum stability and, possibly, an insight into the electroweak baryogenesis. The main goal of studying the VBF channel is to gain information on the quartic coupling between two vector bosons and two Higgs bosons represented with a coupling strength modifier parameter κ_{2V} . Measurement of $\kappa_{2V} = 1$ would confirm the SM prediction. The $b\bar{b}b\bar{b}$ final state has the highest branching ratio among all di-Higgs decays. However, since it has a fully hadronic final state, the estimation of the dominant QCD background is non-trivial. A fully data-driven neural network-based method is employed for the background estimation in the analysis.

The analysis is improved compared to the previous VBF $HH \rightarrow b\bar{b}b\bar{b}$ analysis [3] in many aspects. Particle flow objects and a new, improved b-tagging algorithm, the DL1r algorithm [59], are used for the first time. The analysis events are triggered by the b-jet triggers of types $2b2j$ and $2b1j$. The Higgs Candidates pairing strategy is based on the ΔR separation between jets of the leading Higgs Candidate. In addition to selecting events with the four b-tagged jets, events with exactly two b-tagged jets are selected for the data-driven background estimation purpose. The selection steps, except for the number of the b-tagged jets, remain the same for the 4b and the 2b events. The 2b and the 4b events are assumed to be kinematically similar, and the differences are accounted for by applying event weights to the 2b events. The weights are derived in the control region using neural network reweighting and are applied to the 2b data in the signal region, resulting in the 4b background prediction

in the signal region. The reliability of the background estimate is confirmed by the good agreement between the 2b and the 4b distributions for a number of kinematic variables in the control regions. Additionally, the method is validated by applying background reweighting to the events vetoed by the analysis selections.

The limitations of the ATLAS Run 2, regarding both the centre of mass energy and the luminosity, motivate to improve the sensitivity of Higgs pair production in future experiments, namely, the HL-LHC at $\sqrt{s} = 14$ TeV and the FCC-hh at $\sqrt{s} = 100$ TeV. Even with a factor of 10 higher luminosity at the HL-LHC, the SM couplings will not be precisely tested, motivating the need for the FCC-hh. A simplified analysis is performed on the Monte Carlo simulated data to study the Standard Model di-Higgs VBF production in the HL-LHC and the FCC-hh. The treatment is simplified in many aspects; a calorimeter emulation is used instead of a full detector simulation, b-tagging is simplified and background is estimated using the simulation of the 4b QCD processes instead of the data-driven technique.

A benchmark study of the ATLAS Run 2 - like scenario is performed. The same analysis selections are applied to the signal samples from the calorimeter emulation and the full ATLAS simulation. The number of selected di-Higgs events agrees within 10%. The final result of the simplified analysis is the expected 95% CL upper limit of 430 times the SM cross-section prediction. This is three times worse compared to the main analysis result, indicating the negative impact of the employed background prediction method in the simplified study.

The expected 95% CL upper limit on the VBF di-Higgs SM cross-section for the HL-LHC scenario is set at 248 times the SM cross-section using the simplified analysis. For the FCC-hh the expected 95% CL upper limit is set at 18 times the SM cross-section. The VBF $HH \rightarrow b\bar{b}b\bar{b}$ generator-level projection study in [25] results in the expected 95% CL upper limit of 49 times the SM for the HL-LHC scenario and 12 times the SM for the FCC-hh scenario. These results are factors of 5 and 1.5 better compared to the results obtained in the simplified analysis here. The two better results can be explained by the use of the background samples with higher event statistics (hence with the reduced statistical uncertainties), the consideration of the semi-boosted and boosted regimes in addition to the resolved analysis, and lastly, the lack of consideration of the detector effects.

It is difficult to predict the results of the HL-LHC or the FCC-hh real-data analyses. Even though with more available data the data-driven techniques' performance should improve, the scaling of the systematic uncertainties of the data-driven methods in the future can only be estimated. The HL-LHC projection studies were performed based on rescaling of the early Run 2 ggF analysis results [97]. Focusing on the background uncertainties only, the best-case scenario assumes that the background uncertainty could scale down to $\sqrt{\frac{L_0}{L}}$, where L_0 is the total integrated luminosity of the ATLAS Run 2 (the reference luminosity) and L is the total integrated luminosity of the new scenario, as studied and prescribed in [98]. This means,

for the HL-LHC with $L = 3000 \text{ fb}^{-1}$, the background uncertainty could be as low as 20% of the current background uncertainty, while the worst case scenario would be to retain the current uncertainty values. The data-driven methods are used for the background estimation because the Monte Carlo estimation of the QCD background is not reliable at the moment. However, work on the improvements in calculations has already started [99]. The ideal case would be the one in which data-driven methods for the QCD background modelling are reliably and independently cross-checked with the theoretical predictions, as for now only the data-driven validation and the consistency checks are possible.

To summarize, the ATLAS analysis presented in this thesis improves the previous result 6-fold, because of the improved background modelling, the event categorization and the reconstruction algorithms ameliorations. The simplified study shows the potential of exploring the VBF production mode of Higgs pairs decaying to the $4b$ final state, when the higher total luminosities (3000 and $30\,000 \text{ fb}^{-1}$) and the centre of mass energies (14 and 100 TeV) are available. Although it is not possible to make a precise projection of the uncertainties of the HL-LHC upgrade experiment or of the proposed FCC-hh collider experiment, the increased luminosities and the collision energies will help to significantly reduce the statistical and systematic uncertainties in the $HH \rightarrow b\bar{b}b\bar{b}$ analysis and thereby better constrain the di-Higgs couplings.

Bibliography

- [1] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, Sep 2012. ISSN 0370-2693. doi:10.1016/j.physletb.2012.08.020. URL <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- [2] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, Sep 2012. ISSN 0370-2693. doi:10.1016/j.physletb.2012.08.021. URL <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [3] The ATLAS Collaboration. Search for the $HH \rightarrow b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2020(7), Jul 2020. ISSN 1029-8479. doi:10.1007/jhep07(2020)108. URL [http://dx.doi.org/10.1007/JHEP07\(2020\)108](http://dx.doi.org/10.1007/JHEP07(2020)108).
- [4] Aberle O. et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020. doi:10.23731/CYRM-2020-0010. URL <https://cds.cern.ch/record/2749422>.
- [5] Mangano M. et al. *Physics at the FCC-hh, a 100 TeV pp collider*. CERN Yellow Reports: Monographs. CERN, Geneva, Jun 2017. doi:10.23731/CYRM-2017-003. URL <https://cds.cern.ch/record/2270978>.
- [6] Abbott D. et al. 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. URL <https://cds.cern.ch/record/2780536>.
- [7] The ATLAS Collaboration. Search for non-resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical report, CERN, Geneva, 2022. URL <https://cds.cern.ch/record/2811390>.
- [8] The ATLAS Collaboration. The ATLAS Fast TracKer system. *Journal of Instrumentation*, 16(07):P07006, jul 2021. doi:10.1088/1748-0221/16/07/P07006. URL <https://dx.doi.org/10.1088/1748-0221/16/07/P07006>.

- [9] Thomson M. *Modern particle physics*. Cambridge University Press, New York, 2013. ISBN 978-1-107-03426-6.
- [10] Purcell A. Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules. page 10, 2012. URL <https://cds.cern.ch/record/1473657>.
- [11] Englert F. and Brout R. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi:10.1103/PhysRevLett.13.321. URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [12] Higgs P. W. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. doi:10.1103/PhysRevLett.13.508. URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [13] Valente M. *The Standard Model of Particle Physics*. Springer International Publishing, Cham, 2022. ISBN 978-3-030-94047-8. doi:10.1007/978-3-030-94047-8. URL <https://doi.org/10.1007/978-3-030-94047-8>.
- [14] Tanabashi M. et al. Review of Particle Physics. *Phys. Rev. D*, 98:030001, Aug 2018. doi:10.1103/PhysRevD.98.030001. URL <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [15] Kanemura S. et al. Can WIMP dark matter overcome the nightmare scenario? *Phys. Rev. D*, 82:055026, Sep 2010. doi:10.1103/PhysRevD.82.055026. URL <https://link.aps.org/doi/10.1103/PhysRevD.82.055026>.
- [16] The LHCb Collaboration. Test of lepton universality in beauty-quark decays. *Nature Physics*, 18(3):277–282, mar 2022. doi:10.1038/s41567-021-01478-8. URL <https://doi.org/10.1038/s41567-021-01478-8>.
- [17] Cline J. M. Baryogenesis, 2006. URL <https://arxiv.org/abs/hep-ph/0609145>.
- [18] Bass S. D. et al. The Higgs boson implications and prospects for future discoveries. *Nature Reviews Physics*, 3(9):608–624, Sep 2021. ISSN 2522-5820. doi:10.1038/s42254-021-00341-2. URL <https://doi.org/10.1038/s42254-021-00341-2>.
- [19] Servant G. Electroweak baryogenesis. 2022. URL <https://cerncourier.com/a/electroweak-baryogenesis/>.
- [20] Branco G. C. et al. Theory and phenomenology of two-Higgs-doublet models. *Physics Reports*, 516(1):1–102, 2012. ISSN 0370-1573. doi:<https://doi.org/10.1016/j.physrep.2012.02.002>. URL <https://www.sciencedirect.com/science/article/pii/S0370157312000695>.

- [21] Randall L. and Sundrum R. Large Mass Hierarchy from a Small Extra Dimension. *Phys. Rev. Lett.*, 83:3370–3373, Oct 1999. doi:10.1103/PhysRevLett.83.3370. URL <https://link.aps.org/doi/10.1103/PhysRevLett.83.3370>.
- [22] Grazzini M. et al. Higgs boson pair production at NNLO with top quark mass effects, journal=Journal of High Energy Physics. 2018(5):59, May 2018. ISSN 1029-8479. doi:10.1007/JHEP05(2018)059. URL [https://doi.org/10.1007/JHEP05\(2018\)059](https://doi.org/10.1007/JHEP05(2018)059).
- [23] Coleppa B. et al. Constraining type II 2HDM in light of LHC Higgs searches. *Journal of High Energy Physics*, 2014(1), jan 2014. doi:10.1007/jhep01(2014)161. URL <https://doi.org/10.1007%2Fjhep01%282014%29161>.
- [24] Dreyer F. et al. Vector-boson fusion Higgs pair production at N3LO. *Physical Review D*, 98(11), Dec 2018. ISSN 2470-0029. doi:10.1103/physrevd.98.114016. URL <http://dx.doi.org/10.1103/PhysRevD.98.114016>.
- [25] Bishara F. et al. Higgs pair production in vector-boson fusion at the LHC and beyond. *The European Physical Journal C*, 77, 11 2016. doi:10.1140/epjc/s10052-017-5037-9.
- [26] The ATLAS Collaboration. Erratum to: Search for the $HH \rightarrow b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2021(5):207, May 2021. ISSN 1029-8479. doi:10.1007/JHEP05(2021)207. URL [https://doi.org/10.1007/JHEP05\(2021\)207](https://doi.org/10.1007/JHEP05(2021)207).
- [27] The CMS Collaboration. Search for higgs boson pair production in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review Letters*, 129(8), aug 2022. doi:10.1103/physrevlett.129.081802. URL <https://doi.org/10.1103%2Fphysrevlett.129.081802>.
- [28] The CMS Collaboration. Search for nonresonant pair production of highly energetic higgs bosons decaying to bottom quarks, 2022. URL <https://arxiv.org/abs/2205.06667>.
- [29] The ATLAS Collaboration. 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. *Journal of High Energy Physics*, 2019(1), Jan 2019. ISSN 1029-8479. doi:10.1007/jhep01(2019)030. URL [http://dx.doi.org/10.1007/JHEP01\(2019\)030](http://dx.doi.org/10.1007/JHEP01(2019)030).
- [30] The ATLAS Collaboration. Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physical Review D*, 105(9), may 2022. doi:10.1103/physrevd.105.092002. URL <https://doi.org/10.1103%2Fphysrevd.105.092002>.

- [31] The ATLAS Collaboration. Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2022. URL <https://cds.cern.ch/record/2816332>.
- [32] The CMS Collaboration. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature*, 607(7917):60–68, Jul 2022. ISSN 1476-4687. doi:10.1038/s41586-022-04892-x. URL <https://doi.org/10.1038/s41586-022-04892-x>.
- [33] CERN. CERN’s accelerator complex. <https://home.cern/science/accelerators/accelerator-complex>. Accessed: 2022-09-26.
- [34] Mobs E. The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019. Jul 2019. URL <https://cds.cern.ch/record/2684277>.
- [35] The ATLAS Collaboration. Total Integrated Luminosity and Data Quality in 2015-2018. <https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2018/DataSummary/figs/intlumivstimeRun2DQall.png>, . Accessed: 2022-09-26.
- [36] The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008. doi:10.1088/1748-0221/3/08/s08003. URL <https://doi.org/10.1088/1748-0221/3/08/s08003>.
- [37] Pequeno J. Computer generated image of the whole ATLAS detector. 2008. URL <https://cds.cern.ch/record/1095924>.
- [38] Schott M. et al. Review of single vector boson production in pp collisions at $\sqrt{7}$ TeV. *The European Physical Journal C*, 74(7), jul 2014. doi:10.1140/epjc/s10052-014-2916-1. URL <https://doi.org/10.1140/2Fepjc%2Fs10052-014-2916-1>.
- [39] The ATLAS Collaboration. Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment. 2020. URL <https://cds.cern.ch/record/2723878>.
- [40] Vigani L. et al. Study of prototypes of LFoundry active and monolithic CMOS pixels sensors for the ATLAS detector. *Journal of Instrumentation*, 13, 10 2017. doi:10.1088/1748-0221/13/02/C02021.
- [41] Capeans M et al. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010. URL <https://cds.cern.ch/record/1291633>.
- [42] Pequeno J. Computer Generated image of the ATLAS calorimeter. Mar 2008. URL <https://cds.cern.ch/record/1095927>.
- [43] Pequeno J. Computer generated image of the ATLAS Muons subsystem. Mar 2008. URL <https://cds.cern.ch/record/1095929>.

- [44] The ATLAS Collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004–P10004, oct 2020. doi:10.1088/1748-0221/15/10/p10004. URL <https://doi.org/10.1088/2F1748-0221%2F15%2F10%2Fp10004>.
- [45] The ATLAS Collaboration. The ATLAS Data Acquisition system in LHC Run 2. Technical report, CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2244345>.
- [46] Raschka S. Introduction to Deep Learning Lecture Notes, 2021. URL <https://sebastianraschka.com/blog/2021/dl-course.html>.
- [47] Raschka S. et al. *Python Machine Learning - Third Edition*. Packt Publishing, Dec 2019.
- [48] IBM Cloud Education. Neural Networks. 2020. URL <https://www.ibm.com/cloud/learn/neural-networks>.
- [49] Ersoy S. et al. Automatic decision making system with environmental and traffic data. *Mathematical Models in Engineering*, 7, 06 2021. doi:10.21595/mme.2021.22020.
- [50] Cacciari M. et al. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008. ISSN 1029-8479. doi:10.1088/1126-6708/2008/04/063. URL <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [51] Marzani S. et al. Looking Inside Jets. *Lecture Notes in Physics*, 2019. ISSN 1616-6361. doi:10.1007/978-3-030-15709-8. URL <http://dx.doi.org/10.1007/978-3-030-15709-8>.
- [52] Salam G. P. Towards jetography. *The European Physical Journal C*, 67(3-4): 637–686, may 2010. doi:10.1140/epjc/s10052-010-1314-6. URL <https://doi.org/10.1140%2Fepjc%2Fs10052-010-1314-6>.
- [53] The ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *The European Physical Journal C*, 77(7), jul 2017. doi:10.1140/epjc/s10052-017-5004-5. URL <https://doi.org/10.1140%2Fepjc%2Fs10052-017-5004-5>.
- [54] The ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *The European Physical Journal C*, 77(7), jul 2017. doi:10.1140/epjc/s10052-017-5031-2. URL <https://doi.org/10.1140%2Fepjc%2Fs10052-017-5031-2>.
- [55] The ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 81(8):689, Aug 2021. ISSN 1434-6052.

doi:10.1140/epjc/s10052-021-09402-3. URL <https://doi.org/10.1140/epjc/s10052-021-09402-3>.

- [56] The ATLAS Collaboration. Configuration and performance of the ATLAS b-jet triggers in Run 2. *The European Physical Journal C*, 81(12):1087, Dec 2021. ISSN 1434-6052. doi:10.1140/epjc/s10052-021-09775-5. URL <https://doi.org/10.1140/epjc/s10052-021-09775-5>.
- [57] The ATLAS Colaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ tev. 79(11), nov 2019. doi:10.1140/epjc/s10052-019-7450-8. URL <https://doi.org/10.1140/2Fepjc%2Fs10052-019-7450-8>.
- [58] The ATLAS Collaboration. Identification of Jets Containing b -Hadrons with Recurrent Neural Networks at the ATLAS Experiment. Technical report, CERN, Geneva, 2017. URL <https://cds.cern.ch/record/2255226>.
- [59] The ATLAS Collaboration. Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration. Technical report, CERN, Geneva, 2020. URL <https://cds.cern.ch/record/2718610>.
- [60] The ATLAS Collaboration. Expected performance of the 2019 ATLAS b-taggers, 2019. URL <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/>. Accessed: 2022-11-20.
- [61] The ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 79(11), Nov 2019. ISSN 1434-6052. doi:10.1140/epjc/s10052-019-7450-8. URL <http://dx.doi.org/10.1140/epjc/s10052-019-7450-8>.
- [62] The ATLAS Collaboration. Flavour-tagging efficiency corrections for the 2019 ATLAS PFlow jet and VR track jets b-taggers with the full LHC Run II dataset, . URL <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2021-001.png>.
- [63] Alwall J. et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014. doi:10.1007/jhep07(2014)079. URL <https://doi.org/10.1007%2Fjhep07%282014%29079>.
- [64] Ball R.D. et al. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4), apr 2015. doi:10.1007/jhep04(2015)040. URL <https://doi.org/10.1007%2Fjhep04%282015%29040>.
- [65] Sjöstrand T. et al. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, jun 2015. doi:10.1016/j.cpc.2015.01.024. URL <https://doi.org/10.1016%2Fj.cpc.2015.01.024>.

- [66] Agostinelli S. et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi:[https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [67] Albrand S. et al. The AMI Database Project: Atlas Data Challenge Bookkeeping, and the Tag Collector, a new tool for Release Management. *eConf*, 0303241 (hep-ex/0304029):MONT003. 10 p, Apr 2003. URL <https://cds.cern.ch/record/613910>.
- [68] L. S. Borgna. Search for pair production of Higgs bosons decaying to four bottom quarks with data collected by the ATLAS detector, 2021. URL <https://cds.cern.ch/record/2812193>.
- [69] Carvalho A. et al. Higgs pair production: choosing benchmarks with cluster analysis. *Journal of High Energy Physics*, 2016(4):1–28, apr 2016. doi:10.1007/jhep04(2016)126. URL <https://doi.org/10.1007%2Fjhep04%282016%29126>.
- [70] Rogozhnikov A. Reweighting with Boosted Decision Trees. *Journal of Physics: Conference Series*, 762:012036, oct 2016. doi:10.1088/1742-6596/762/1/012036. URL <https://doi.org/10.1088/1742-6596/762/1/012036>.
- [71] Kanamori T. et al. A Least-Squares Approach to Direct Importance Estimation. *J. Mach. Learn. Res.*, 10:1391–1445, December 2009. ISSN 1532-4435.
- [72] Freedman D. et al. *Statistics (international student edition)*. 2007.
- [73] Cowan G. *Statistical data analysis*. Oxford University Press, USA, 1998.
- [74] Efron B. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 01 1979. doi:10.1214/aos/1176344552. URL <https://doi.org/10.1214/aos/1176344552>.
- [75] Read A. L. Presentation of search results: the CLs technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693–2704, 2002. ISSN 0954-3899. doi:10.1088/0954-3899/28/10/313.
- [76] Cowan G. et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), 2011. ISSN 1434-6052. doi:10.1140/epjc/s10052-011-1554-0.
- [77] Cranmer K. et al. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical report, New York U., New York, 2012. URL <https://cds.cern.ch/record/1456844>.
- [78] Heinrich L. et al. pyhf: v0.6.3. URL <https://doi.org/10.5281/zenodo.1169739>. <https://github.com/scikit-hep/pyhf/releases/tag/v0.6.3>.

- [79] Heinrich L. et al. pyhf: pure-Python implementation of HistFactory statistical models. *Journal of Open Source Software*, 6(58):2823, 2021. doi:10.21105/joss.02823. URL <https://doi.org/10.21105/joss.02823>.
- [80] Avoni G. et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *Journal of Instrumentation*, 13(07):P07017–P07017, jul 2018. doi:10.1088/1748-0221/13/07/p07017. URL <https://doi.org/10.1088/1748-0221/13/07/p07017>.
- [81] The ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, Geneva, 2019. URL <https://cds.cern.ch/record/2677054>.
- [82] Dao V. Parton Shower Uncertainties in Higgs Measurements, ATLAS treatment, 2019. URL <https://indico.cern.ch/event/827858/#3-atlas-treatment>. Accessed: 2022-11-25.
- [83] The ATLAS Collaboration. Theory systematics, 2019. URL <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TheorySystematics>. Accessed: 2022-11-25.
- [84] Bähr M. et al. Herwig++ physics and manual. *The European Physical Journal C*, 58(4):639–707, nov 2008. doi:10.1140/epjc/s10052-008-0798-9. URL <https://doi.org/10.1140/epjc/s10052-008-0798-9>.
- [85] Bellm J. et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76(4):196, 2016. doi:10.1140/epjc/s10052-016-4018-8.
- [86] Butterworth J. et al. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, jan 2016. doi:10.1088/0954-3899/43/2/023001. URL <https://doi.org/10.1088/0954-3899/43/2/023001>.
- [87] CERN. CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector, 2017. URL <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>.
- [88] The ATLAS Collaboration. Jet Uncertainties Package, 2018. URL <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertaintiesRel21Summer2018SmallR?rev=11>.
- [89] The ATLAS Collaboration. Tagger Calibrations Recommendation, 2021. URL https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalibrationRecommendationsRelease21?rev=60#Recommendation_April_2021.
- [90] Bierlich C. et al. A comprehensive guide to the physics and usage of PYTHIA 8.3, 2022. URL <https://arxiv.org/abs/2203.11601>.

- [91] The ATLAS Collaboration. Internal Twiki: VBF Production Prescription. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToVBF>, . Accessed: 2022-09-20.
- [92] Mangano M. et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *Journal of High Energy Physics*, 2003(07):001–001, jul 2003. doi:[10.1088/1126-6708/2003/07/001](https://doi.org/10.1088/1126-6708/2003/07/001). URL <https://doi.org/10.1088/1126-6708/2003/07/001>.
- [93] Alwall J. et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014. doi:[10.1007/jhep07\(2014\)079](https://doi.org/10.1007/jhep07(2014)079). URL <https://doi.org/10.1007%2Fjhep07%282014%29079>.
- [94] Buckley A. et al. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J. C*, 75(132), 2015. doi:<https://doi.org/10.1140/epjc/s10052-015-3318-8>.
- [95] Ball R. D. et al. Parton distributions with LHC data. *Nuclear Physics B*, 867(2):244–289, 2013. ISSN 0550-3213. doi:<https://doi.org/10.1016/j.nuclphysb.2012.10.003>. URL <https://www.sciencedirect.com/science/article/pii/S0550321312005500>.
- [96] Matteo C. et al. FastJet user manual. *The European Physical Journal C*, 72(3), mar 2012. doi:[10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2). URL <https://doi.org/10.1140%2Fepjc%2Fs10052-012-1896-2>.
- [97] The ATLAS Collaboration. Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC. Technical report, CERN, Geneva, 2018. URL <https://cds.cern.ch/record/2652727>.
- [98] The ATLAS Collaboration. Expected performance of the ATLAS detector at the High-Luminosity LHC. Technical report, CERN, Geneva, 2019. URL <http://cds.cern.ch/record/2655304>.
- [99] Buckley A. et al. Monte Carlo event generators for high energy particle physics event simulation, 2019. URL <https://arxiv.org/abs/1902.01674>.

Acknowledgments

I would like to thank professor André Schöning for his guidance throughout my PhD, and his patience, especially in the last few weeks of consultations and corrections. I was inspired and motivated to keep improving as a physicist, researcher and, in the last few months, a more concise and less vague writer. I hope this was achieved, at least to some extent.

I would also like to thank professor Monica Dunford for agreeing to be the second referee of this thesis. I am very grateful to the Research Training Group in Heidelberg, *Physics Beyond the Standard Model*, for funding my research and allowing me to meet a number of amazing fellow PhD students.

It was a pleasure to work with all the Heidelberg colleagues, both from ATLAS and Mu3e. I will always remember you as a cool bunch of smart and quirky people, who know lots about physics, but who also know how to enjoy life (#Bosseln). Cheers! Big kudos to Tamasi for the massive amount of constructive support in this whole process and an even greater amount of thesis corrections up until the very end. I would also like to thank Luigi for being the Bellini coffee card correspondent, those *lungos* were very much needed, especially in the last few weeks. Kudos to Christof for many fruitful analysis-oriented discussions and to Joachim and Sebastian for many on-point analysis clarification questions in our weekly meetings and proofreading chapters of this thesis.

I would also like to thank the whole 4b analysis team, especially, the analysis contacts, Max and Rafael, and the *VBF Push* team. Although to a large extent online, throughout the pandemic, it was great to work on this challenging analysis with you.

My PhD would have not been the same without the pre-pandemic time I spent at CERN. I would like to thank my postdoc, Louis, for all the help and advice very much needed, especially at the beginning, and the whole FTK team I worked with while in Geneva. I also made a bunch of amazing friends. Special thanks go to Amar and Pedja on whom I can always count, no matter what or when, you guys are the best! And to my fellow physicist in training (from the 2nd best experiment), Rhys, for all the inter-experimental support.

Thank you (so much) Annie and Tamasi, you have been much more than colleagues. You guided me through all the intricacies of living in Germany at the beginning and became my *Heidelberg family* soon after. This family would not be complete without my German *Mama* Anna, and our private photographer Arturo (I really hope to see all those pictures one day...). Thank you guys for simply being there!

I want to thank my closest friends based all over the world for their support at all stages of my PhD and for keeping in touch when I continued moving to new countries. Last but not least, none of this would have been possible without my Family.

Pragnę podziękować Rodzicom, Dziadkom, Bratu i całej Rodzinie za od zawsze otrzymywane wsparcie. Dodatkowo dziękuję Dziadkowi Dzikowi i Ś.P. Wujowi Rafałowi za inspiracje do podążania drogą naukową, oraz Tymonowi za wiele dyskusji o tym, że fizyka w swojej szerokiej perspektywie jest naprawdę cool.