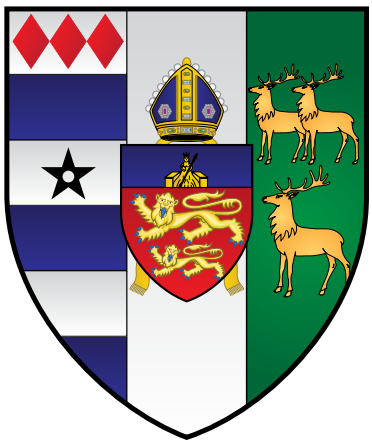
The search for pair production of Higgs Bosons at the ATLAS experiment



Iza Veliscek
Lincoln College, Oxford

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Summer, 2023

Abstract

The thesis presents the search for $HH \rightarrow b\bar{b}b\bar{b}$ that exploits the $126\,\mathrm{fb}^{-1}$ of proton-proton data collected by the ATLAS detector during the Large Hadron Collider Run 2 at $\sqrt{s} = 13\,\mathrm{TeV}$. Done for the first time in an ATLAS diHiggs analysis, the independent optimisation of the gluon-gluon fusion (ggF) and vector boson fusion (VBF) production channels not only improves the sensitivity to anomalies in the trilinear Higgs self-coupling (κ_λ) but also to anomalies in the barely explored HHVV coupling (κ_{2V}). The combination of the ggF and VBF channels allows the exploration of the $\kappa_\lambda - \kappa_{2V}$ parameter space and extracting the $\pm 2\sigma$ confidence region with the $-2\ln\Delta L$ test statistic. The observed upper limit on the Standard Model (SM) HH signal strength is found to be 5.4 times the SM value. Lastly, the ggF channel is interpreted using Higgs and SM Effective Field Theory, the latter is the first such diHiggs ATLAS result. Taking a step further the combination of the three most sensitive diHiggs channels in the $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^-\tau^+$ final states is performed, which is further joined with the single Higgs combination. The thesis mainly focuses on the technical adjustments needed in the non-resonant $HH \rightarrow b\bar{b}b\bar{b}$ analysis to allow for its inclusion in the single Higgs and diHiggs combination. The upper bound on the diHiggs SM signal strength extracted from the diHiggs combination is 2.4 times the SM prediction. The power of the single Higgs and diHiggs combination is that it allows the consideration of a generic κ_λ model, where in addition to the trilinear Higgs self-coupling, the couplings of the Higgs to the top quark, bottom quark, tau lepton and vector bosons can vary. The observed (expected) 95% confidence level is measured to be $-1.3 < \kappa_\lambda < 6.1$ $(-2.1 < \kappa_\lambda < 7.6)$, when the generic κ_λ model is fitted. The thesis concludes with the in-situ calibration of the mis-tag rate for the multijet background in the boosted $X \rightarrow b\bar{b}$ tagger. The calibration uses simulated dijet events enriched with $g \rightarrow b\bar{b}$ and the full Run 2 data to derive preliminary scale factors that quantify the difference in the mis-tag efficiency between data and simulations.

Dedicated to my parents and sister.

Acknowledgements

I would like to thank my supervisors, Prof. Todd B. Huffman and Dr. James A. Frost, for their mentoring and guidance throughout my DPhil journey. Special thank you goes to Dr William Balunas for helping me unravel the problems at hand and trace down all the bugs along the way. The analyses presented would not be possible without the whole $HH \to b\bar{b}b\bar{b}$ team led by Maximilian Swiatlowski and Rafael Teixeira de Lima. Shout out to Chris and Shota for pushing through the final steps. I would like to thank the Science and Technology Facilities Council (STFC) for provding the STFC studenship and Lincoln College Oxford - University of Oxford for supproting me through the Briggs/Myers Scholarship.

The completion of this thesis would not be possible without the unconditional support of my friends and family. I would also like to thank my beloved dog Medo, who sadly passed away this May before this thesis was finished, but was my dear companion for fourteen long years. Lastly, I would like to thank my furry friends Taci and Meda for the cuddles and distractions while writing up.

Preface

The thesis would not be possible without the *ATLAS Collaboration*, with 2659 active authors and 263 institutes participating [1]. The work presented in this thesis has been done in collaboration with institutes from all over the world. The author contributed to the non-resonant $HH \rightarrow b\bar{b}b\bar{b}$ analysis presented in Chapter 6, the Single Higgs and diHiggs combination presented in Chapter 7 and the Calibration of the boosted $X \rightarrow b\bar{b}$ tagger presented in Chapter 8. The exact contributions of the author are stated in the introduction to each chapter. Unless stated otherwise all figures have been made by the author of the thesis.

Publications

Search for nonresonant 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 (ATLAS Collaboration, 2023, CERN-EP-2022-235) [2], *Chapter 6*

The author of this thesis was one of the main analysis members in charge of the statistical analysis. The contributions included the implementation and validation of the $-2\ln\Lambda$ test statistic method for setting limits, helping to implement a setup that significantly speeded up the limit-setting procedure, and studying the impact of the systematics. The author was among the first analysis members who explored the two-dimensional parameter space of $\kappa_\lambda - \kappa_{2V}$ and her work enabled the publication of the limits in the $\kappa_\lambda - \kappa_{2V}$ plane. A significant contribution to the study and final definition of the systematic uncertainty on the acceptance and reweighting of the ef-

fective field theory signal was made by the author. Minor contributions of the author include work on the software and studies of the impact of changing the definition of the cut that reduces the $t\bar{t}$ background. Finally, a crucial role was played by the author in the final stages of the analysis which involved studying and validating the unblinded results, finalising the interpretations, evaluating the impact of systematics and producing several figures showing the final results of the analysis that were published in the above-cited paper. The contributions to the analysis from the author not discussed in the thesis include studies on alternative kinematic region definitions and background estimation.

Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV (ATLAS Collaboration, 2022, CERN-EP-2022-149) [3], *Chapter 7*

The author of the thesis prepared the nonresonant $HH \to b\bar{b}b\bar{b}$ workspace that was used in the final combination. In addition, the author performed a minor study to assess the impact of a different Higgs mass used in the $HH \to b\bar{b}b\bar{b}$ and $HH \to b\bar{b}b\bar{b}$ and $HH \to bb\tau^-\tau^+$ channels.

Efficiency corrections for a tagger for boosted $H \to b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector (ATLAS Collaboration, 2021, ATL-PHYS-PUB-2021-035) [4], *Chapter 8*

The author was one of the main analysers that contributed to the PUB note, contributing to the studies on the mis-tag rate of the multijet background. However, the thesis focuses on the *in-situ* calibration of the mis-tag rate of the multijet background using the secondary vertex invariant mass, which has been introduced fairly recently by the author. The author implemented the fitting method and several of the systematics uncertainties. She helped to maintain the framework and was responsible for leading the production of the datasets used for the calibration. In addition, she studied the behaviour of the fit and explored several validation methods. The thesis does not present studies with the mean signed transverse impact parameter significance, which was found to contain a too large a mis-modelling to be used.

Contents

1 The Standard Model of Particle Physics 1

1.1 The Lagrangian . . . 2

1.2 Fermionic Matter Fields . . . 2

1.2.1 The Dirac Equation . . . 3

1.3 The Bosonic Gauge Fields . . . 4

1.4 Interactions . . . 7

1.5 Strong Force . . . 8

1.6 Electroweak Unification . . . 9

1.6.1 Electroweak Symmetry Breaking . . . 9

1.6.2 Electroweak Force . . . 13

1.6.3 The Yukawa Terms . . . 14

1.7 The Standard Model Lagrangian . . . 17

1.8 The Standard Model - An Incomplete Picture . . . 19

2 Higgs self-coupling 22

2.1 DiHiggs Production . . . 23

2.2 Effective Field Theory . . . 26

2.3 Single Higgs Production . . . 30

2.4 The Higgs potential . . . 32

2.5 New Physics . . . 34

3 The ATLAS Experiment **37**

3.1 The Large Hadron Collider . . . 37

3.1.1 Synchrotrons . . . 39

3.1.2 Run 2 Performance . . . 40

3.2 The ATLAS Detector . . . 42

3.2.1 The Coordinate System . . . 43

3.2.2 The Inner Detector . . . 44

3.2.3 The Calorimeter System . . . 47

3.2.4 The Muon Spectrometer . . . 49

3.3 The Trigger System . . . 51

3.3.1 Level-1 Trigger . . . 51

3.3.2 HLT . . . 52

4 Simulation **53**

4.1 Event Generation . . . 54

4.2 Detector Simulation . . . 59

5 Physics Object Reconstruction in ATLAS **60**

5.1 Building Blocks - Tracks, Vertices, Topo-clusters . . . 60

5.2 Muons . . . 63

5.3 Jets . . . 64

5.3.1 Jet Constituents . . . 64

5.3.2 Jet Clustering . . . 65

5.3.3 Jet Calibration . . . 68

5.3.4 Jet Types . . . 72

5.4 Identification of b-quark Jets (*b-tagging*) . . . 74

6 Non-resonant $HH \to b\bar{b}b\bar{b}$ Analysis 79

6.1 Overview . . . 79

6.2 Data and Simulations . . . 82

6.3 Sample Reweighting . . . 84

6.4 Event Selection . . . 88

6.4.1 Object Specifications . . . 88

6.4.2 Triggers . . . 89

6.4.3 Reconstructing the diHiggs System . . . 90

6.4.4 Reducing the Background . . . 94

6.4.5 Kinematic Regions . . . 96

6.4.6 Categorisation . . . 99

6.5 Background Estimation . . . 99

6.5.1 Kinematic Reweighting . . . 101

6.5.2 Validation . . . 103

6.6 Systematic Uncertainties . . . 104

6.6.1 Background Estimation Uncertainties . . . 107

6.6.2 Experimental Uncertainties . . . 109

6.6.3 Theoretical Uncertainties in diHiggs Signal . . . 111

6.6.4 Summary of Standard Model HH uncertainties . . . 112

6.6.5 The Acceptance Reweighting EFT Uncertainty . . . 113

6.7 Statistical Analysis . . . 120

6.7.1 The Likelihood Function . . . 120

6.7.2 Test Statistic and p-value . . . 124

6.7.3 The Profile Likelihood Ratio Test Statistic . . . 125

6.7.4 The $-2\ln\Lambda$ Test . . . 129

6.7.5 CLs Technique . . . 130

6.7.6 Combining Measurements from Different Channels . . . 133

6.7.7 Asimov Data . . . 134
6.8 Results . . . 134
6.8.1 $\kappa-$ framework Interpretations . . . 136
6.8.2 Effective Field Theory Interpretations . . . 141
6.9 Conclusion . . . 148

7 Combining single Higgs and diHiggs Measurements 150
7.1 Input Channels . . . 151
7.2 Combination Strategy and Statistical Analysis . . . 151
7.3 Tweaking the $HH \to b\bar{b}b\bar{b}$ Analysis . . . 153
7.4 Impact of Higgs Mass Assumptions . . . 155
7.5 Results . . . 156
7.6 Conclusion . . . 159

8 Calibrating the boosted $X \to b\bar{b}$ tagger 160
8.1 The Tagger . . . 161
8.1.1 Data and Simulations . . . 162
8.1.2 Object Definition and Event Selection . . . 162
8.1.3 Double-b Tagging Performance . . . 163
8.2 Calibrating the mis-tag rate of the multijet background . . . 165
8.2.1 Data and Monte Carlo Simulations . . . 166
8.2.2 Event Selection . . . 167
8.2.3 Flavour Discriminating Variables . . . 170
8.2.4 Flavour Fraction Correction . . . 174
8.2.5 Pre-fit Distributions . . . 177
8.2.6 Systematic Uncertainties . . . 178
8.2.7 Results . . . 186
8.2.8 Conclusions . . . 197

9 Summary **199**

9.1 Outlook 203

A CR12 Non-Closure Uncertainty **206**

B Ranking Plots **207**

C $< s_{d0} >$ Distribution **208**

Bibliography **212**

Chapter 1

The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a quantum field theory that successfully describes the observed matter in the Universe and its interactions through three out of the four known forces. Given the observed interactions of quarks and leptons through the weak, electromagnetic, and strong force the most general Lagrangian density, from now on referred to simply as the Lagrangian, that correctly describes the system is constructed in this chapter. The Higgs mechanism is briefly discussed, and a more detailed discussion of the Higgs boson with an emphasis on the Higgs self-coupling, diHiggs production, and effective field theories involving modifications of the Higgs potential can be found in the next chapter. The chapter concludes by discussing the shortcomings of the SM. Natural units are used throughout the thesis and basic knowledge of quantum field theory, group theory, and special relativity is assumed. Further details can be found in Symmetry and the standard model: mathematics and particle physics by Matthew B. Robinson [5] and the lecture notes on Unification by Arttu Rajantie [6], which are the general references for this chapter.

1.1 The Lagrangian

The principle of least action implies that the dynamics of any system can be captured by constructing a Lagrangian that is both mathematically and physically sound. Thus, when considering the behaviour of observed matter at high energies, around the electroweak scale, the following conditions must be met by the Lagrangian:

I is real,
II has dimensions of GeV^4,
III is a local function of the fields and their derivatives,
IV is invariant under symmetries to reflect the physics observed [6].

These four requirements are the guiding principles in the construction of the SM Lagrangian, with especially IV placing very stringent constraints on the allowed terms. The SM must be invariant under the Lorentz and $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry transformations. However, there is no guarantee that such a Lagrangian can explain all observed phenomena at every energy scale.

1.2 Fermionic Matter Fields

Fermions are elementary spin-half particles that obey Fermi–Dirac statistics and are the building blocks of all known matter. There are two types of fermions: leptons and quarks, a defining property of the latter is that they carry a colour charge. As can be seen in Figure 1.1, quarks (leptons) come in 6 flavours, each flavour carrying a different set of quantum charges[1] and are grouped into three generations or families. Within a generation, the pairing is done so that the two particles differ by one unit of electric charge, while the mass of particles increases with each

[1] The quantum charges are: mass, colour, isospin, electric charge

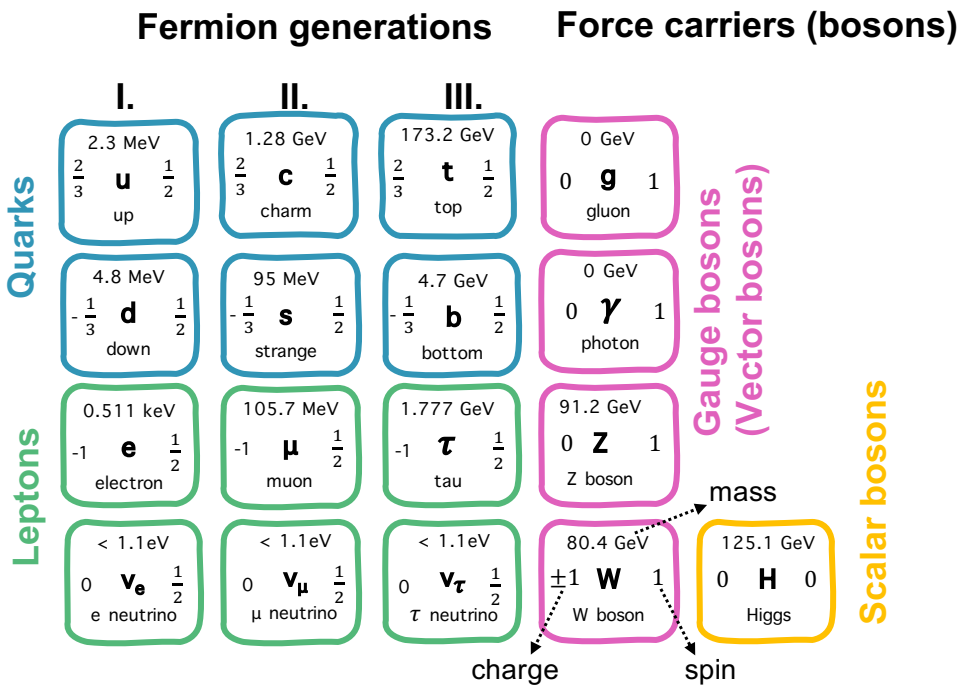


Figure 1.1: The particle content of the Standard Model is shown, each particle has its mass, electric charge, and spin indicated. The fermions are ordered into three generations, the only difference between generations is in the particle's mass which increases per generation.

generation. For example, let's consider the first quark generation, $\begin{pmatrix} u \\ d \end{pmatrix}$, and the second quark generation, $\begin{pmatrix} c \\ s \end{pmatrix}$. The up and down quarks are the lightest quarks, carrying $\frac{2}{3}$e and $-\frac{1}{3}$e respectively. The charm quark is a heavier version of the up quark, it has all the same quantum numbers except for its mass, and the same can be said for the strange quark with respect to the down quark.

1.2.1 The Dirac Equation

The behaviour of fermions in terms of relativistic quantum mechanics was first described by Dirac in [7], who sought an alternative formula to the Klein-Gordan equation. The solution to having a relativistic quantum description of fermions that

has only first-order derivatives is the Dirac equations,

$$0 = (i\partial\!\!\!/ - m)\Psi, \tag{1.1}$$

where $\partial\!\!\!/ = \gamma_\mu \partial^\mu$, γ_μ are the Dirac matrices and Ψ is a four-component Dirac spinor (after quantisation spinor field). The properties of Ψ, such as spin and chirality, are completely determined by the algebra of the γ-matrices. The Dirac equation also predicts the existence of anti-particles, which are interpreted as the negative energy solution to the Dirac equation.

Requiring Lorentz invariance the Lagrangian of a freely propagating fermionic field is

$$\mathcal{L} = \bar{\Psi}(i\partial\!\!\!/ - m)\Psi, \tag{1.2}$$

where $\bar{\Psi} = \Psi^\dagger \gamma^0$ and m is the mass of the fermion. It can be easily shown that it is invariant under both, charge conjugation (C) and parity (P) transformation.

It is useful to note that Equation 1.2 can be written as

$$\mathcal{L} = \bar{\Psi_L} i\partial\!\!\!/ \Psi_L + \bar{\Psi_R} i\partial\!\!\!/ \Psi_R - m(\bar{\Psi_L}\Psi_R + \bar{\Psi_R}\Psi_L), \tag{1.3}$$

where the Dirac spinors are decomposed into right- and left-handed fields. The above equation implies that for massless fermions the left- and right-handed fields are allowed to exist on their own, allowing parity violation.

1.3 The Bosonic Gauge Fields

Bosons, on the other hand, mediate the interactions between fermions and are represented as quantum vector fields with integer spin. There are four spin-1 bosons: gluon, $W^\pm$, Z, and photon, which mediate the strong, weak, and

electromagnetic forces respectively. These bosonic fields are often called gauge vector fields, as under a local gauge symmetry[2] the transformations of the field can be expressed in terms of a Lie group.

Let's first consider an Abelian gauge vector field, B_μ, which under an Abelian gauge transforms as $B_\mu \to B_\mu - \frac{1}{g}\partial_\mu\theta$, where g is some constant and θ is an element of the Lie algebra. Defining the field strength tensor as $F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and enforcing that the Lagrangian must be invariant under Lorentz transformation and the Abelian gauge symmetry the most general Lagrangian is simply

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{1.4}$$

If the symmetry chosen is $U(1)$ the Lagrangian above describes the free propagation of the electromagnetic field in the absence of any sources. The picture gets slightly more complicated when the gauge symmetry is non-Abelian, as is the case for the strong and weak forces. According to pure gauge theory, also called Yang-Mills theory [8], the non-Ablian gauge transformation of vector fields, A_μ, is $A_\mu \to A_\mu - \frac{1}{g}\partial_\mu + i[\theta, A_\mu]$, where $[\theta, A_\mu]$ is the Lie bracket[3]. Consequently, the only allowed Lorentz and gauge invariant term of the Lagrangian is

$$\mathcal{L} = -\frac{1}{2}Tr\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}, \tag{1.5}$$

where the field strength tensor takes the form $\mathcal{G}_{\mu\nu} = \partial\mu A_\nu - \partial\nu A_\mu + ig[A_\mu, A_\nu]$ and as before g is a constant introduced for convenience[4].

To get a better grasp of the physics it is useful to note that if A_μ, and consequently $\mathcal{G}_{\mu\nu}$, is an element of the Lie algebra of the symmetry group, $A_\mu = A^a_\mu t^a$, where t^a are

[2] Gauging the symmetry simply means that the symmetry that $\theta = \theta(\boldsymbol{x})$, i.e. the transformation is not global but local.

[3] The Lie bracket defines the product of the Lie algebra and must satisfy $[x, x] = 0$ and the Jacobi identity [9].

[4] Maybe it is a bit of a stretch to say by convenience, as it comes about naturally when writing down a general transformation of A_μ under the gauge transformation.

the generators of the algebra of the Lie group. Substituting this into 1.5 one gets

$$\mathcal{L} = -\frac{1}{4}(\partial\mu A^a_\nu - \partial\nu A^a_\mu)(\partial^\mu A^{\nu a} - \partial^\nu A_{\mu a}) \tag{1.6}$$

$$+\frac{g}{2} f^{abc}(\partial\mu A^a_\nu - \partial_\nu A^a_\mu) A^{\mu b} A^{\nu c} \tag{1.7}$$

$$-\frac{g^2}{4} f^{abc} f^{ade} A^b_\mu A^c_\nu A^{\mu d} A^{\nu e}, \tag{1.8}$$

where f^{abc} is the structure constant from the commutation relation $[t^a, t^b] = i f^{abc} t^c$, which determines the structure of the Lie group. This then describes the dynamics of D massless non-Abelian gauge fields in the absence of matter, where D is the dimensionality of the gauge symmetry group [6]. The first line of the equation represents the free propagation of a field, while the second and fourth lines are the triple and quartic self-interactions of those fields.

Taking a step back, the question to ask is whether Equation 1.6 correctly describes the interaction of the strong and weak force. Let's say the strong force is generated by the $SU(3)$ group, the dimensionality of the group is 8, which means that there should be 8 massless fields in agreement with the 8 gluons observed. Similarly, choosing $SU(2)$ to generate the weak force, should give the 3 bosons involved in the weak interaction. So far so good, however, both forces are short-ranged, which at first glance seems in contradiction with the predictions of the Yang-Mills theory. For the strong force, it turns out that g is $g_s \sim 1$, is sufficiently large to lead to confinement, i.e. the self-interacting terms are so strong that they lead to the formation of only neutral (colourless) states. On the other hand, g of the weak force is not large enough to cause confinement, and a different solution to generate a short-range force is needed. The weak gauge bosons, opposite to the predictions of the Yang-Mills theory, turn out to have non-zero masses. The mechanism of generating the masses of weak bosons is outside the scope of the pure gauge theory and will be discussed in more detail in Section 1.6.

1.4 Interactions

Putting together the results from solving Dirac's equation and pure gauge theory the Lagrangian naively would be

$$\mathcal{L} = \sum_{f=1}^{6} \bar{\Psi_L} i\partial\!\!\!/ \Psi_L + \bar{\Psi_R} i\partial\!\!\!/ \Psi_R - m(\bar{\Psi_L}\Psi_R + \bar{\Psi_R}\Psi_L) \tag{1.9}$$

$$-\frac{1}{2} Tr\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu} - \frac{1}{2} TrW_{\mu\nu}W^{\mu\nu} - \frac{1}{4} F_{\mu\nu}F^{\mu\nu}, \tag{1.10}$$

where $\mathcal{G}$, W, F are the strength tensors of the strong, weak and electromagnetic force respectively. This is a rather incomplete picture as, besides the self-interactions of the non-Abelian vector fields, there are no interactions between the various fields.

The interaction terms can be introduced by requiring that the total Lagrangian is invariant under the $SU(3) \times SU(2) \times U(1)$ symmetry. It follows that spinor fields gauge transformation is

$$\Psi \to e^{ig_s\alpha^b(x)T^b + ig_w\theta^a t^a + iY}\Psi, \tag{1.11}$$

where $T^b_{SU(3)}$ are the generators of $SU(3)$ and the generators of $SU(2)$ are $t^a_{SU(2)}$, Y is the hypercharge generator of $U(1)$, g_s is the strong coupling constant, g_W is the weak coupling constant. There are two things to note here, firstly $U(1)$ is not the symmetry of electromagnetic force but $U(1)_Y$, the choice becomes clear later. Secondly, the representations of $T^b_{SU(3)}$ and $t^a_{SU(2)}$ are chosen to match the observations. For example, leptons do not carry a colour charge, which means that they transform trivially under SU(3) and Equation 1.11 can be simplified to

$$l \to e^{ig_w\theta^a t^a + iY} l, \tag{1.12}$$

where l is the lepton field.

Given 1.11 the covariant derivative

$$D_\mu = \partial_\mu + ig_s G^b_\mu T^b_{SU(3)} + ig_W W^a_\mu t^a_{SU(2)} + ig_1 Y B_\mu \tag{1.13}$$

has to replace ∂_μ in 1.9 to ensures that the total Lagrangian is invariant under the $SU(3) \times SU(2) \times U(1)$, where g_1 is the coupling constant to the $U(1)$ field. The covariant derivative gives rise to cross terms between the vector and spinor terms fields, which represent the interactions. Clearly, the strength of the interaction is mediated by the coupling constants, which need to be measured experimentally.

1.5 Strong Force

Quarks carry colour charge and therefore transform non-trivially under $SU(3)$. The strong interaction they experience is described by the interaction term

$$\mathcal{L}_{IntQCD} = \sum_q i g_s \bar{q}_i \gamma^\mu (G^b_\mu T^b_{SU(3)})_{ij} q_j \tag{1.14}$$

$$= \sum_q i \frac{g_s}{2} \bar{q}_i \gamma^\mu (G^b_\mu \lambda_b)_{ij} q_j, \tag{1.15}$$

where q runs over all quarks, i is the colour state of the quark and T^b are in the fundamental representation given by the 3×3 Gell-Mann matrices ($T^b = \frac{1}{2}\lambda_b$). Hence, quarks come in three possible (colour) states,

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{1.16}$$

red (r) or green (g) or blue (b) and corresponding anti-colour states.

From examining the structure of the Gell-Mann matrices it is clear that strong force changes the colour states while conserving colour at the interaction vertex. The 8

colours states of the gluons can be deducted from the Gell-Mann matrices to be: $r\bar{g}$,$r\bar{b}$,$g\bar{r}$, $g\bar{b}$, $b\bar{r}$, $g\bar{g}$, $\frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b})$ and $\frac{1}{\sqrt{6}}(r\bar{r} + g\bar{g} - 2b\bar{b})$.

1.6 Electroweak Unification

The short-ranged weak force, contrary to the strong and electromagnetic force, can change the flavour of the fermions and is not invariant neither under the P or CP transformation. Naively, a gauge field invariant under $SU(2)_L$, would be expected to describe the weak force, where L indicates that the bosons only couple to left-handed spinors. While there are indeed three mediators of the weak force and P is maximally violated in the beta-decay of cobalt-60 [10], several observed phenomena fail to be explained. The predicted force mediators are massless and only couple to the left-handed fields. However, the observed bosons are massive and the observation of CP violation in the Kaon sector [11] implies the existence of a neutral current boson that, while able to distinguish between them, couples to both, the left- and right-spinor fields. Both problems are resolved through electroweak unification and spontaneous symmetry breaking (SSB).

1.6.1 Electroweak Symmetry Breaking

The weak and electromagnetic forces are low energy manifestations of the electroweak force generated by the $SU(2) \times U(1)$ symmetry group. At low energy this symmetry is spontaneously broken, giving the vector fields mass without actually breaking the gauge symmetry. In the SM this is known as the Higgs mechanism and, based on works [12, 13, 14, 15, 16], is brought about by the existence of a scalar boson field with a non-zero vacuum expectation value (vev).

Let's first introduce a scalar field, ϕ, and call it the Higgs field. The simplest form the

potential with a non-zero vev ϕ can take is

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2, \tag{1.17}$$

where $\mu^2 < 0$ and $\lambda > 0$. The shape of the potential is illustrated in Figure 1.2, which clearly shows the circle of minima located at $|\phi| = v = \frac{|\mu|}{\sqrt{\lambda}}$. Postulating that the Higgs

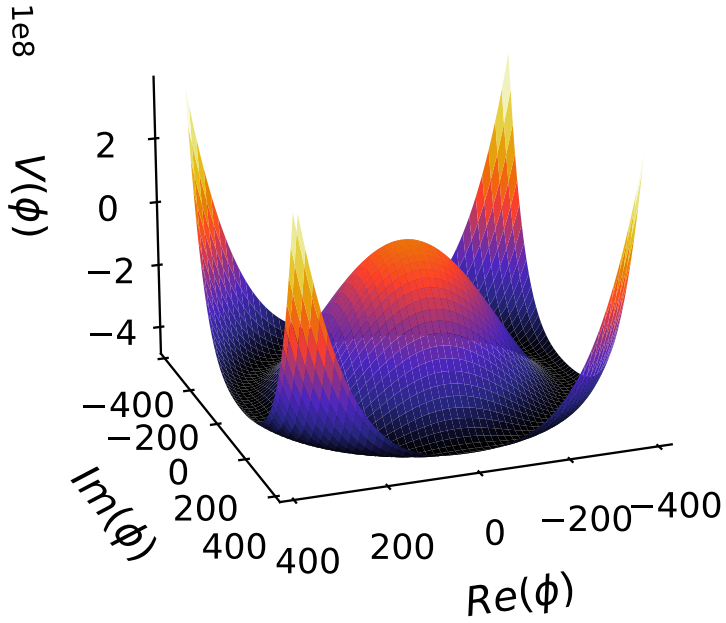


Figure 1.2: The Higgs potential, $V(\phi) = \mu^2\phi + \lambda(\phi^\dagger\phi)^2$, where $\mu^2 < 0$.

field transforms under $SU(2) \times U(1)$ as a complex scalar doublet the vev state can be chosen to be

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.18}$$

thanks to the freedom of choice coming from gauge invariance.

The electroweak part of the Higgs Lagrangian then is

$$\mathcal{L}_{Higgs} = (\mathcal{D}_\mu \phi)^\dagger \mathcal{D}^\mu \phi - V(\phi), \tag{1.19}$$

with the covariant derivative

$$D_\mu = \partial_\mu + + i g_2 W_\mu^a t_{SU(2)}^a + i g_1 B_\mu \tag{1.20}$$

$$= \partial_\mu + + i g_2 A_\mu^a T^a. \tag{1.21}$$

In the second line, for convenience, A_μ is defined as $A_\mu = \{W_\mu^1, W_\mu^2, W_\mu^3, B_\mu\}$. The generators of $SU(2)_L$ are $\{T^1, T^2, T^3\}$ and the $U(1)_Y$ field is redefined as $T^4 = \frac{g_1}{2g_2}\mathbb{1}$. The coupling strengths of the $SU(2)_L$ and $U(1)$ fields to the Higgs field are g_2 and g_1 respectively.

Skipping over the mathematical derivation[5] expanding over the vev results in the covariant derivative part of $\mathcal{L}_{Higgs}$ taking the form

$$\mathcal{L}_{Higgs} \supset \frac{1}{2} g_2^2 A_\mu^a A^{\mu b} \phi_0^\dagger \{T^a, T^b\} \phi_0 \tag{1.22}$$

$$= \frac{1}{8} A_\mu g_2^2 v^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{g_1}{g_2} \\ 0 & 0 & -\frac{g_1}{g_2} & \frac{g_1^2}{g_2} \end{pmatrix} A^\mu. \tag{1.23}$$

The equation can be simplified by finding the eigenvectors and eigenvalues of the above matrix. The result is the physical, i.e. mass, vector boson states,

$$\mathcal{A}_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3; m_\gamma = 0, \tag{1.24}$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2); m_{W^\pm} = \frac{g_2 v}{2}, \tag{1.25}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu; m_Z = \sqrt{g_1^2 + g_2^2}\frac{v}{2}, \tag{1.26}$$

[5] See among others [5, 17, 6] for the detailed derivation.

where the Weinberg angle, θ_W, is defined as

$$\cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} \qquad \sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}}. \tag{1.27}$$

It is clear from Equations 1.24, 1.25, 1.26 that the electroweak symmetry breaking results in three vector fields, $W^{\pm}$ and Z, acquiring mass while $\mathcal{A}_\mu$ remains massless. The massless photon implies that there is a residual symmetry, $U(1)_{EM}$ of the electroweak interaction, and the symmetry breaking pattern is $SU(2) \times U(1)_Y \to U(1)_{EM}$. Note that $U(1)_Y$ and $U(1)_{EM}$ are not the same as $\mathcal{A}_\mu$ mixes B_μ of $U(1)_Y$ and A^3_μ of $SU(2)_L$. Z is also a mix of B_μ and A^3_μ, while $W^{\pm}_\mu$ are composed of solely $SU(2)_L$ fields. This explains why P is maximally violated in weak flavour changing processes, as $W^{\pm}_\mu$ purely couple to left-handed fields. On the other hand, neutral weak interactions violate CP as Z couples to both, right- and left-handed fermions, but with a different strength.

The predicted particle spectrum in the broken state matches what is experimentally observed and correctly predicts the outcomes of electroweak interactions to a high precision. It is important to note that the couplings, and so indirectly the masses, of the massive gauge bosons are not predicted by SSB. However, their relative strength determines the mass relationship of the $W^{\pm}$ and Z,

$$\frac{m_W}{m_Z} = \cos\theta_W. \tag{1.28}$$

If this relationship is found not to withstand measurements it implies the assumptions made to describe the Higgs mechanism are not quite correct. The Higgs was chosen to be a complex doublet as this is the simplest choice that gives the correct particle spectrum. It is not yet ruled out that the Higgs field is more complex, for instance, composed of two doublets, as is the case in the two-Higgs-doublet model (2HDM) extension of the SM [18], in which case the relationship in Equation 1.28 does not hold.

Similarly, by expanding the Higgs potential around the vev as

$$\mathcal{L}_{Higgs} \supset -V_0 - \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4, \tag{1.29}$$

the Higgs field (h) acquires mass, $m_H = \sqrt{2\lambda}v = \sqrt{2}|\mu|$, and is expected to self-interact. The coupling strength of the trilinear self-interaction,

$$\lambda_{hhh} = \lambda v = \frac{m_H}{2v} \tag{1.30}$$

and quartic self-interaction

$$\lambda_{hhhh} = \lambda = \frac{m_H}{2v^2}, \tag{1.31}$$

are in the SM determined by m_H and the vev, which are free parameters of the theory.

The Higgs boson with a mass of $125\,\mathrm{GeV}$ was discovered by the CMS and ATLAS experiments at the LHC in 2012 [19, 20] and the vev can be determined from the Fermi constant to be $v = \frac{G_F}{\sqrt{2}} = 246\,\mathrm{GeV}$, which in the SM completely determines the self-couplings of the Higgs that are yet to be measured precisely.

1.6.2 Electroweak Force

Postulating an unbroken symmetry $SU(2)_L \times U(1)_Y$ and introducing a scalar doublet field with non-zero vev correctly predicts the particle spectrum of the electroweak sector. Rewriting the covariant derivative in 1.20 in terms of the physical fields results in

$$D_\mu = \partial_\mu + i\frac{g_1 g_2}{\sqrt{g_1^2+g_2^2}}(T^3+Y)\mathcal{A}_\mu + i\frac{g_2^2 T^3 - g_1^2}{\sqrt{g_1^2+g_2^2}}Z_\mu \tag{1.32}$$

$$+\frac{ig_2}{\sqrt{2}}(W_\mu^+(T^1+iT^2)+W_\mu^-(T^1-iT^2)), \tag{1.33}$$

where Y was introduced for generality, and applying it to the kinematic term of the fermionic Lagrangian correctly generates the electromagnetic and weak interactions[6]. However, requiring $SU(2)_L$ only to couple to left-handed fields introduces a new problem.

Left-handed fermion fields transform under $SU(2)_L \times U(1)_Y$ as a doublet

$$f_L \to e^{i\theta^a t^2 + iY} f_L, \tag{1.34}$$

while right-handed fermion fields transform as a singlet

$$f_R \to e^{iY} f_R. \tag{1.35}$$

This implies, in contradiction with observations, that fermions must be massless as the mass term in 1.3 is not invariant under $SU(2)_L \times U(1)_Y$ and therefore not allowed. Luckily, it is trivial to generate fermion masses through their interaction with the Higgs field.

1.6.3 The Yukawa Terms

Given the Higgs field is in the double representation of $SU(2)_L \times U(1)_Y$ an additional term, which combines the left- and right-handed fermionic fields, is allowed to be

[6] It is well known from experiment that e is the coupling strength of the electromagnetic interaction and can be identified as $e = \frac{g_1 g_2}{\sqrt{g_1^2+g_2^2}}$, while the electric charge of the fermion in question is $q = (T^3+Y)$

added to the Lagrangian, the Yukawa term,

$$\mathcal{L}_{=}... + y_f(\bar{f}_L \phi f_R + \bar{f}_R \phi^\dagger f_L), \tag{1.36}$$

where y_f is a coupling constant. For the first generation of the lepton (l), the term after SSB is

$$\mathcal{L}_{=}... - \frac{1}{\sqrt{2}} y_e v(\bar{l}_L l_R + \bar{l}_R l_L) - \frac{1}{\sqrt{2}} y_l(\bar{l}_L l_R + \bar{l}_R l_L)h, \tag{1.37}$$

generating the Dirac mass term, with $m_l = \frac{y_l v}{\sqrt{2}}$. Therefore, the most general gauge invariant Yukawa term for leptons after diagonalisation is

$$\mathcal{L}_{=}... - \sum_{f=1}^{3} y_l^f \left(\bar{l}_L^f \phi l_R^f + \bar{l}_R^f \phi^\dagger l_L^f \right), \tag{1.38}$$

where f runs over the three generations and y_l^f is the Yukawa couplings determining the mass of the charged leptons f. The three Yukawa couplings determining the masses of the three charged leptons are free parameters. It is worth noting that as the diagonalisation redefines the left- and right-handed fields in the same way the kinematic term is unchanged, which leaves the three generations of leptons decoupled.

The down quark masses can be generated in the same way as the lepton masses. However, to generate the up quark masses the vev state is

$$\tilde{\phi} = i\sigma_2 \phi^* = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix}, \tag{1.39}$$

where σ_2 is a Pauli matrix. The most general Yukawa term for the quarks is

$$\mathcal{L}_{=}... - \sum_{f,g=1}^{3} \left(y_d^{fg} \bar{q}_L^f \phi d_R^g + (y_d^{fg})^* \bar{d}_R^g \phi^\dagger q_L^f + y_u^{fg} \bar{q}_L^f \tilde{\phi} u_R^g + (y_u^{fg})^* \bar{u}_R^g \tilde{\phi}^\dagger q_L^f \right), \tag{1.40}$$

where f, g runs over the three generations, y_d^{fg} and y_u^{fg} are elements of the 3×3 complex matrix, Y_d and Y_u respectively. In the above equation, q_L is the left-handed quark field, and u_R (d_R) is the right-handed up-like (down-like) quark field. Each of the matrices can be decomposed into a diagonal matrix by redefining the quarks fields as $d_L^f \to V_{dL}^{fg} d_L^g$ and $u_L^f \to V_{uL}^{fg} u_L^g$. The Yukawa term for the quarks then can then be simplified as

$$\mathcal{L} = \ldots - \sum_{f=1}^{3} y_d^f \left(\bar{q}_L^f \phi d_R^f + \bar{d}_R^f \phi^\dagger q_L^f \right) \tag{1.41}$$

$$- \sum_{f=1}^{3} y_u^f \left(\bar{q}_L^f \tilde{\phi} u_R^f + \bar{u}_R^f \tilde{\phi}^\dagger q_L^f \right), \tag{1.42}$$

where the nine Yukawa couplings define the quark masses and are as in the lepton case not fixed by the theory. It is important to note that the up and down parts of the left-handed field are redefined (”rotated”) differently, leading to the kinematic term changing as

$$\mathcal{L}_= \ldots - \frac{g_2}{\sqrt{2}} \left(\bar{u}_L^f (V_{uL}^\dagger V_{dL})^{fg} \gamma^\mu W_\mu^+ d_L^g + h.c. \right), \tag{1.43}$$

where $V_{uL}^\dagger V_{dL}$ define the Cabibbo-Kobayashi- Maskawa (CKM) matrix. The non-zero off-diagonal terms of the CKM matrix are manifested as the change of quark generations in the interaction with W bosons. Again the CKM matrix has 4 free parameters that need to be determined experimentally.

1.7 The Standard Model Lagrangian

Piecing it all together the most general Lagrangian that is renormalisable and invariant under the $SU(3) \times SU(2) \times U(1)$ gauge transformation is

$$\mathcal{L} = \underbrace{-\frac{1}{2}Tr\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu} - \frac{1}{2}Tr\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu} - \frac{1}{4}\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}}_{\text{The dynamics terms of the gauge groups.}} \tag{1.44}$$

$$\underbrace{+(\mathcal{D}_\mu\phi)^\dagger\mathcal{D}^\mu\phi + \mu^2\phi^\dagger\phi - \frac{1}{2}\lambda(\phi^\dagger\phi)^2}_{\text{The dynamics and mass terms of the Higgs}} \tag{1.45}$$

$$\underbrace{+\sum_{f=1}^{3}\left(\bar{l}_L^f i\not{\mathcal{D}} l_L^f + \bar{l}_R^f i\not{\mathcal{D}} l_R^f + \bar{q}_L^f i\not{\mathcal{D}} q_L^f + \bar{d}_R^f i\not{\mathcal{D}} d_R^f + \bar{u}_R^f i\not{\mathcal{D}} u_R^f\right)}_{\text{The dynamics terms of leptons and quarks}} \tag{1.46}$$

$$\underbrace{-\sum_{f=1}^{3} y_l^f\left(\bar{l}_L^f\phi l_R^f + \bar{l}_R^f\phi^\dagger l_L^f\right)}_{\text{Yukawa term for the leptons}} \tag{1.47}$$

$$\underbrace{-\sum_{f,g=1}^{3}\left(y_d^{fg}\bar{q}_L^f\phi d_R^g + (y_d^{fg})^*\bar{d}_R^g\phi^\dagger q_L^f + y_u^{fg}\bar{q}_L^f\tilde{\phi}u_R^g + (y_d^{fg})^*\bar{u}_R^g\tilde{\phi}^\dagger q_L^f\right)}_{\text{Yukawa term for the quarks}}. \tag{1.48}$$

The theory is fully determined by fixing the 19 free parameters; 3 Yukawa couplings y_l^f in the lepton sector, 10 parameters of Y_d and Y_u in the quark sector, Higgs self-coupling λ and Higgs mass term μ^2 from the Higgs sector, 3 gauge couplings g_1, g_2, g_3 of the vector fields and θ_{QCD} (not discussed in the thesis), which all need to be determined experimentally.

The tremendous success of the SM can be seen in Figure 1.3 which shows the total and fiducial production cross-section measurements of various SM processes at different center-of-mass energies, which all are in good agreement with its theoretical predictions. It is only fair to state that the SM holds up to a high degree of accuracy when it comes to predicting the outcomes of particle collisions about the electroweak scale. Yet, from failing to describe the force of gravity to not including dark matter, the SM is an incomplete description of the observed Universe.

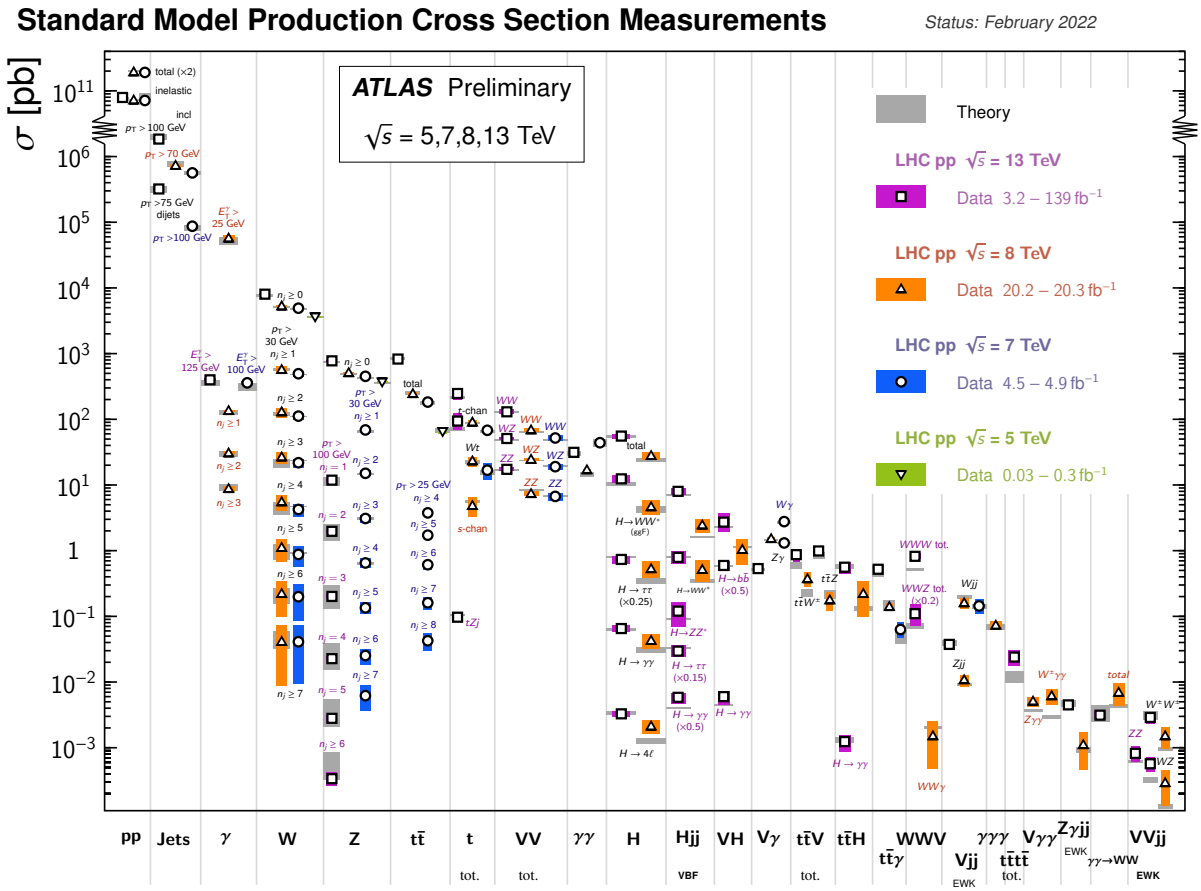


Figure 1.3: The total and fiducial production cross-section measurements of various SM processes at different center-of-mass energies are summarised. The measurements are based on the data collected by the ATLAS experiment and compared to the theoretical predictions of the SM, which are shown in grey. The figure is taken from [21].

1.8 The Standard Model - An Incomplete Picture

The SM is a rather incomplete picture of the observed Universe. On the experimental side, the SM fails to predict or explain the following phenomena;

1. Neutrino masses

   The experimental discovery of neutrino oscillations by Super-Kamiokande and SNO+ established that neutrinos have mass, which is not predicted by the SM [22, 23]. Currently, only the differences between the neutrino masses can be established from the neutrino oscillation experiment, while cosmological data implies that the masses are tiny, $\sim 0.05\,\mathrm{eV}$ [24]. It might seem trivial to extend the SM and include the neutrino masses through the mechanism used for the up-like quarks to generate a Dirac mass term. However, this is not done as the Yukawa coupling would need to be orders of magnitudes smaller than the fermionic counterparts and it is not the only possible extension of the SM.

   Alternatively, provided the right-handed neutrino exists the Majorana mass term, which couples the right-handed neutrino and left-handed antineutrino, is gauge invariant and can be added to the Lagrangian. If neutrino mass is generated through the Dirac and Majorana term, the seesaw mechanism produces a heavy and light neutrino [17, 25].

2. Matter-antimatter symmetry

   The Big Bang predicts an equal amount of matter and anti-matter to be created at the beginning of the Universe and given that in the SM matter and anti-matter behave nearly identically, there are small sources of CP violation, the prevalence of matter cannot be explained by the SM. According to Sakharov [26] the three conditions that need to be satisfied to explain the observed matter-antimatter symmetry in the Universe are

   - baryon number violation,

- sufficient C- and CP-violation,
- an interaction that takes place out of thermal equilibrium.

While there are sources of baryon number as well as C- and CP-violation in the SM, though they might not be sufficient to generate matter-antimatter asymmetry, there is no mechanism to bring the Universe out of equilibrium, and a beyond SM explanation is needed.

3. Gravity
   While the electroweak and strong force can be described by means of quantum field theory and are therefore part of the SM, the quantisation of gravity is non-trivial. As of now, there is no quantum description of gravity but only a classical explanation, General Relativity. Loop quantum gravity [27] and string theory, where the latter predicts supersymmetry (SUSY) [28], are some theoretical attempts at quantizing gravity, yet neither has been experimentally confirmed.

4. Dark matter
   Several astronomical and cosmological observations indicate that there is ”missing” matter, called dark matter (DM), in the Universe, not accounted for by the SM. To name a few, observed rotation curves of a typical spiral galaxy require the mass distributions of the galaxy to be approximately $M(r) \sim r$, which is not the case if only luminous matter is taken into account [29, 30]. The observed gravitational lensing and the dynamics of the Bullet Cluster, also cannot be explained by general relativity without the inclusion of DM [31]. To explain the above phenomena while taking into account the constraints from the electroweak measurements, DM has to be massive and only weakly or not at all interact through the weak force. No such particle is predicted within the SM.

The SM, as the name itself indicated, is just a model that can be fine-tuned to match

the observations at the electroweak scale through the 19 free parameters and is expected to break if not sooner at the Planck scale. This hints that it might be merely a low-energy manifestation of a more general theory. Additionally, the need for new physics, specifically quantum gravity, at the Planck scale brings about the problem of naturalness, known as the hierarchy problem. Given quantum gravity corrections to the Higgs propagator, the Higgs mass becomes enormous and huge cancellations are needed to get the Higgs mass down to its experimentally observed value [28]. In other words, given some new heavy particles and the cutoff scale about the Planck scale the Higgs mass would be expected of the order of the Planck scale and not the electroweak scale as observed. The source of the cancellations of the quantum gravity corrections is a mystery. SUSY again is a good candidate to solve this problem, as the contribution perfectly cancels out. Another possible resolution is that the Higgs is in fact composed and not a singular scalar field. Possible BSM theories that modify the Higgs sector are discussed in the next chapter.

Chapter 2

Higgs self-coupling

The scalar boson discovered in 2012 at the LHC by the ATLAS and CMS experiments [19, 20] is a good candidate for the predicted Higgs boson needed in electroweak symmetry breaking (EWSB). Since then one of the main goals of current and future colliders is to pin down all properties of the discovered scalar and conclusively answer the question of whether it is indeed the SM Higgs boson.

The trilinear and quartic Higgs self-couplings are such yet to be measured quantities. At the centre-of-mass energies reached by current hadron colliders, only the measurement of the trilinear self-coupling (λ_{HHH}) is possible. Any observed deviations from the SM prediction would shrink the set of allowed BSM theories and have important implications for the evolution of the Universe.

This chapter outlines how diHiggs and single Higgs productions can be used to probe the Higgs trilinear coupling. Examples of BSM theories that modify the self-coupling are given and their role in cosmology is discussed. Two approaches to quantify anomalies in λ_{HHH} are outlined, the $\kappa-$ framework and effective field theory.

2.1 DiHiggs Production

The leading diHiggs production channel is the gluon-gluon fusion (ggF), for which the SM cross-section at $\sqrt{s} = 13\,\mathrm{TeV}$ in proton-proton collisions is $\sigma = 31.05\,\mathrm{fb}$. The calculation is done at next-to-next-to-leading order (NNLO) assuming $m_H = 125\,\mathrm{GeV}$ and using the large top mass approximation to compute the QCD NNLO corrections including effects of finite top mass [32]. At LO the box and triangle diagrams, which interfere destructively, produce the HH pair through a top loop[1] as shown in Figure 2.1. The matrix element of the box diagrams depends on c_t^2, the coupling of Higgs to the top quark, while the contribution of the triangle diagram is proportional to λ_{HHH}. As seen in Figure 2.2, these contributions produce HH pairs with different kinematics. The triangle diagram dominates at low invariant mass of the diHiggs system (m_{HH}), while the box diagram peaks around $m_{HH} \sim 400\,\mathrm{GeV}$. Therefore, any anomaly in λ_{HHH} is reflected in the cross-section and kinematics of the final state as can be seen in Figure 2.3.

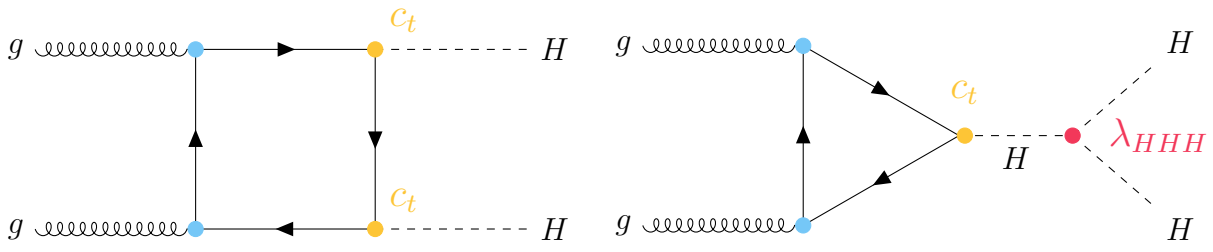


Figure 2.1: The two tree-level Feynman diagrams that contribute to the DiHiggs production via gluon-gluon fusion.

The $\kappa-$ framework is a simple way to quantify deviations from the SM, through the coupling modifiers $\kappa_m = \frac{c_m}{c_m^{SM}}$, where c_m is the measured coupling between the Higgs and particle m and c_m^{SM} its SM prediction. The main point of research in this thesis is the anomalous trilinear Higgs self-coupling (κ_λ). Unless stated otherwise the other coupling modifiers are fixed at their SM values ($\kappa_m = 1$). The maximum destructive

[1] The bottom loop has a small contribution as well but is not taken into account in the calculation.

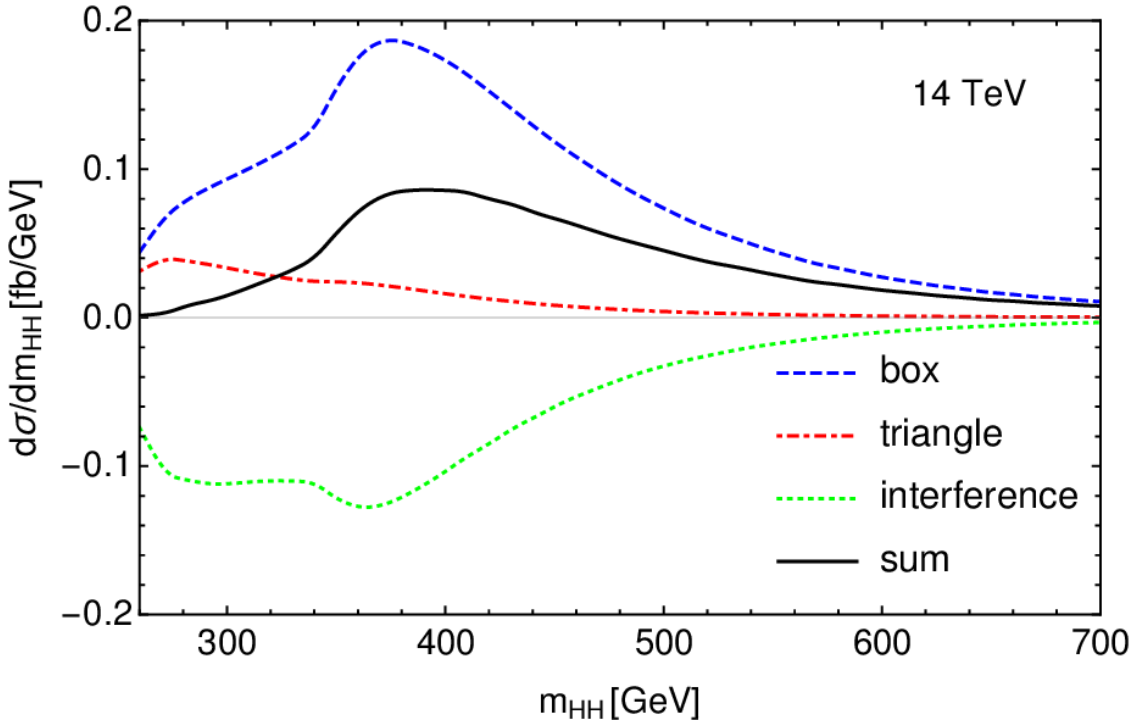


Figure 2.2: The expected m_{HH} spectrum in the ggF SM diHiggs production channel at $\sqrt{s} = 14$ TeV in a proton-proton collider is shown as the solid black line. The dashed lines show the different contributing terms, in blue is the box diagram, in green is the interference, and in red is the triangle term. The figure is taken from [33].

interference between the triangle and box diagrams is at $\kappa_\lambda = 2.4$ characterized by a dip around $350\,\mathrm{GeV}$ and is shown in Figure 2.3. For $\kappa_\lambda < -1$ and $\kappa_\lambda > 5$ the triangle diagram contributes more and more, which is reflected in the m_{HH} distribution peaking at lower masses and an increased cross-section. The simulated m_{HH} distribution for various κ_λ variations in the ggF production channel after applying an approximate event selection resembling the $HH \to b\bar{b}b\bar{b}$ analysis and no detector effects are shown in Figure 2.3.

The VBF channel is the second leading diHiggs production mode and about an order of magnitude smaller than the ggF channel. Its cross-section of $\sigma = 1.73\,\mathrm{fb}$ is calculated at $\sqrt{s} = 13$ TeV in proton-proton collisions to N^{3}LO [34] for $m_H = 125\,\mathrm{GeV}$ and the QCD corrections are computed by neglecting any interactions between the two external quarks. At LO the three contributing diagrams are shown in Figure 2.4 and the trilinear Higgs self-coupling is accessible directly through the diagram in

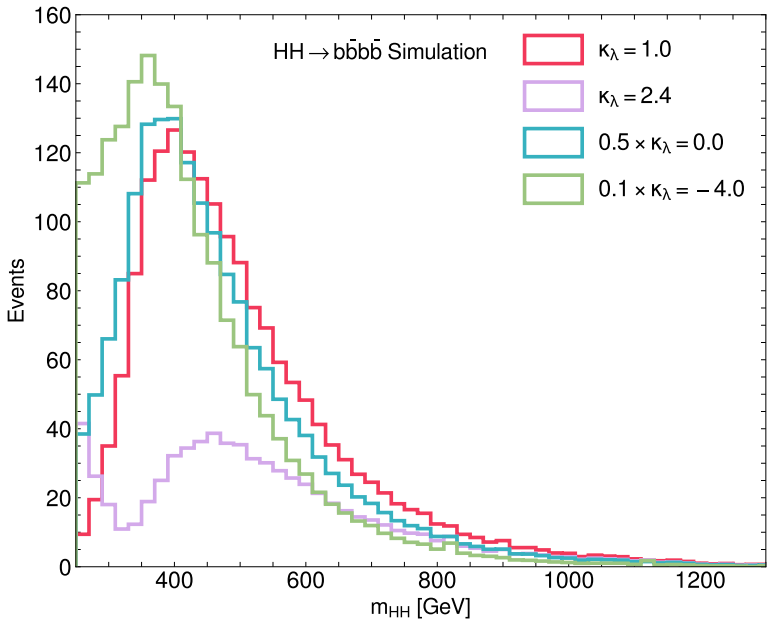


Figure 2.3: The reconstructed m_{HH} distribution for different κ_λ values is shown. Selection and reconstruction criteria resembling the one used in the $HH \to b\bar{b}b\bar{b}$ analysis presented in this thesis are applied to the SM sample. The sample is showered with Pythia 8 and no detector simulation is used. To obtain the various κ_λ variations the SM sample is reweighted according to the strategy described in Chapter 6. The characteristic dip at $\kappa_\lambda = 2.4$ seen on the plot is wiped out by detector resolution effects. The $\kappa_\lambda = 0.0$ signal is scaled by 0.5, while the $\kappa_\lambda = -4.0$ signal is scaled by 0.1.

Figure 2.4a. The VBF production also has a direct dependency on the HVV (c_v coupling) and barely explored HHVV (c_{2v} coupling) vertex, where V stands for either the W or Z boson.

The small cross-section of the VBF production is the result of the destructive interference between the diagrams in Figure 2.4b and 2.4c. As a result of perturbative unitarity the two contributions cancel each other out exactly [35]. However, in some BSM scenarios, c_{2v} and c_v are allowed to deviate from their SM values, and their contribution grows with the centre-of-mass energy roughly as $\mathcal{A} \simeq \frac{s}{v^2}(c_{2v} - c_v^2)$[35]. Thus, any deviation of c_{2v} and c_v away from the $c_{2v} - c_v^2 = 0$ parabola results in an enhancement in the m_{HH} tails, making boosted HH pair production very likely.

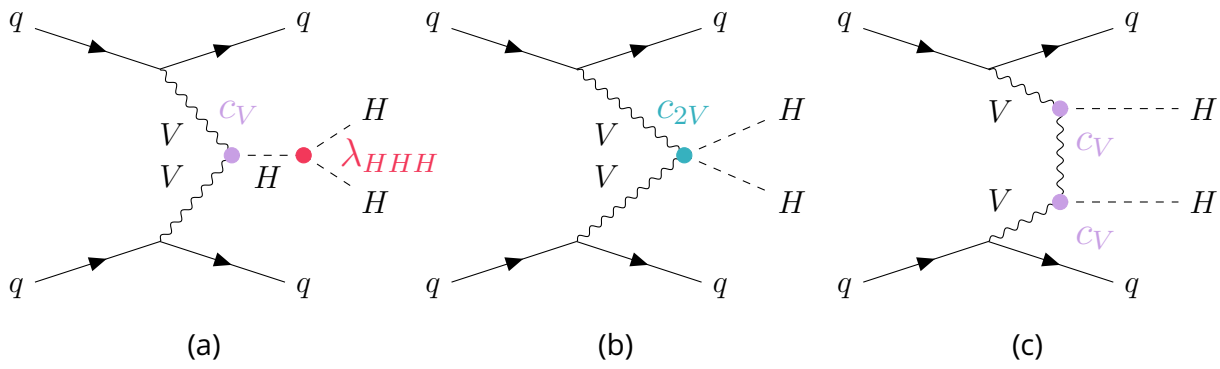


Figure 2.4: The three tree-level Feynman diagrams that contribute to the diHiggs production via Vector Boson Fusion.

2.2 Effective Field Theory

New physics might appear at an energy scale Λ well above the EW scale (vev $\sim 246\,\mathrm{GeV}$) making the direct detection of its dynamics impossible with current collider experiments. However, an indirect model-independent way to quantise deviation from the SM due to these effects is the construction of an Effective Field Theory (EFT). The EFT framework gives a rigorous mechanism to predict finite amplitudes while integrating out the details of new physics above Λ making assumptions only about locality, gauge invariance, and power counting techniques[36]. The requirement of renormalizability is relaxed through the appropriate choice of the power counting scheme and scale, as the Lagrangian is required to be renormalizable only order by order. The Standard Model Effective Field Theory (SMEFT) and Higgs Effective Field Theory (HEFT) are used for reinterpreting the ggF diHiggs production channel of the non-resonant $HH \to b\bar{b}b\bar{b}$ results and presented in this section.

Analogous to Fermi theory, decoupling the two energy scales is possible by introducing Wilson coefficients (c_i) that act as coupling constants between SM fields. It follows that the EFT Lagrangian can be written as an order-by-order expansion of

the SM Lagragian,

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(5)}}{\Lambda} Q_i^{(5)} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} Q_i^6 + \mathcal{O}(\frac{1}{\Lambda^3}) + ..., \tag{2.1}$$

where c_i^n are the Wilson coefficients and $Q_i^{(n)}$ are operators of dimensionality n greater than four[2] [37]. The two scales considered are connected as the physics at Λ determines c_i, while fundamental or composed SM fields make up Q_i [38, 39].

SMEFT can be built from Equation 2.1 by imposing $SU(3) \times SU_L(2) \times U_Y(1)$ gauge symmetry, assuming that the Higgs field (H) is a complex doublet in the $SU_L(2) \times U(1)$ representation as in the SM and setting $\Lambda >> m_Z$. Higher-dimensional operators are suppressed by Λ^{4-n} but it turns out that dimension 5 and 7 operators introduce lepton or baryon number violation, and therefore are not included in the SMEFT Lagrangian [33]. Only dimension 6 operators are relevant for current measurements, as operators of dimensionality ≥ 8 impact the HH production by $< 10^{-4}$ %. Their modification of the SM Lagrangian in the Warsaw basis[3] [39] is

$$\begin{aligned}
\Delta\mathcal{L}^6 \sim & c_H (H^\dagger H)^3 && (2.2)\\
& + c_{H\Box} (H^\dagger H) \partial^{mu} \partial_{mu} (H^\dagger H) && (2.3)\\
& + c_{tH} (H^\dagger H) \bar{q_L} H t_R + h.c. && (2.4)\\
& + c_{HG} H^\dagger H G^A \mu\nu G^{A\mu\nu} && (2.5)\\
& + c_{tG} \bar{Q}_L \sigma^{\mu\nu} T^A \tilde{H} t_R G^A_{\mu\nu} + h.c. + ..., && (2.6)
\end{aligned}$$

where H is the Higgs field, G is the gluon field strength tensor and Q is the (b, t) doublet. The Wilson coefficients involved in the SM-like diHiggs production are c_H and c_{tH}, which change the λ_{HHH} and c_t vertex shown in Figure 2.1 respectively. The

[2]Only $n > 4$ operators of SM fields are considered as the SM is already the most general possible effective field theory up to and including dimension four operators after imposing the SM gauge and Lorentz invariance.

[3]The Warsaw basis gives all allowed independent dimension 6 operators assuming the SM gauge symmetry and that the operators can be made only out of composed or fundamental SM fields. The significance of the Warsaw basis is the reduction of operators, which is a result of eliminating linear combinations that make the SM equations of motion vanish [39].

remaining Wilson coefficients introduce BSM diagrams, shown in Figure 2.5, otherwise forbidden in the SM. The c_{tG} affects both the SM and BSM-like contributions and is marked by a turquoise dot in Figure 2.1 and 2.5. Deviations of c_i from the SM point at zero can impact HH measurements at a few percent level.

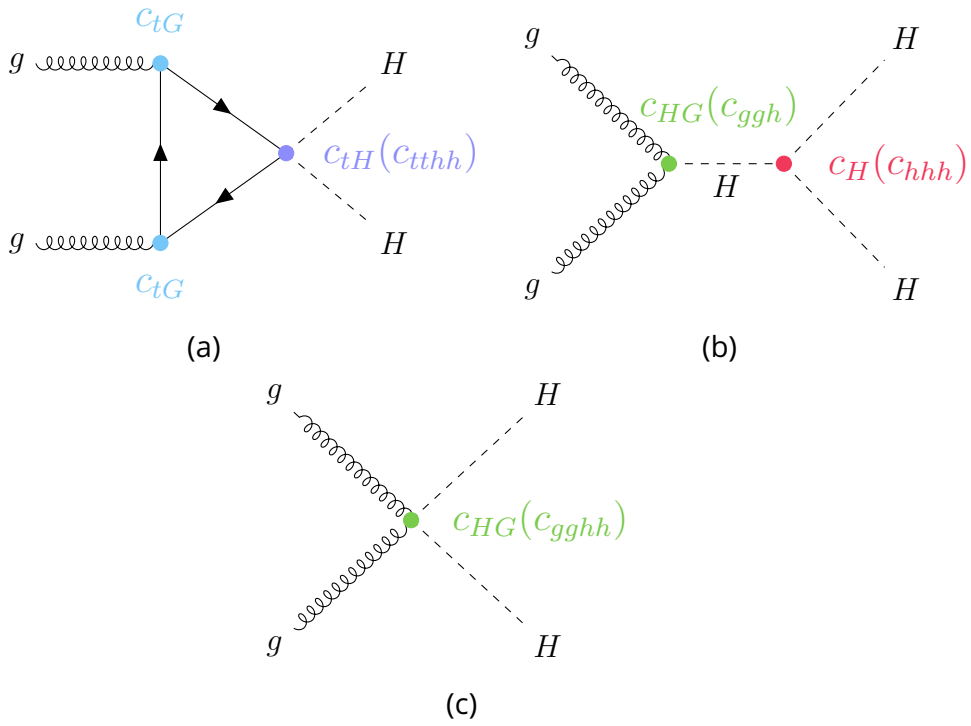


Figure 2.5: The BSM contributing diagrams from SMEFT dimension 6 and HEFT operators that can have a significant impact on the diHiggs production, the later are shown in brackets. The main difference is the Higgs coupling to the gluons, the ggHH and ggH vertex are moderated by the same operator in SMEFT but not in HEFT.

On the other hand, HEFT assumes a non-linear manifestation of electroweak symmetry breaking (EWSB) where the global symmetry $SU(2) \times SU(2)$ is broken into $SU(2)$ [33]. The Higgs field is redefined and transformed as a $SU(2) \times U(1)$ singlet (h), which leads to the decoupling of the single Higgs and diHiggs interactions. The power-counting is done per loop order expanded by $\frac{f^2}{\Lambda^2} = \frac{1}{16\pi^2}$, where f is the energy scale at which the EFT is valid[4] [36, 37]. The five coefficients relevant for ggF diHiggs production and considered in this thesis are κ_λ, κ_t, c_{ttHH}, c_{ggH} and c_{ggHH}, which

[4]This is analogous to the chiral perturbative theory used for describing pions[36, 40].

modify the SM Lagrangian as

$$\Delta\mathcal{L}_{HEFT} = -m_t\left(\kappa_t\frac{h}{v} + c_{tthh}\frac{h^2}{v^2}\right)t\bar{t} - \kappa_\lambda\frac{m_H^2}{v^2}h^3 \tag{2.7}$$

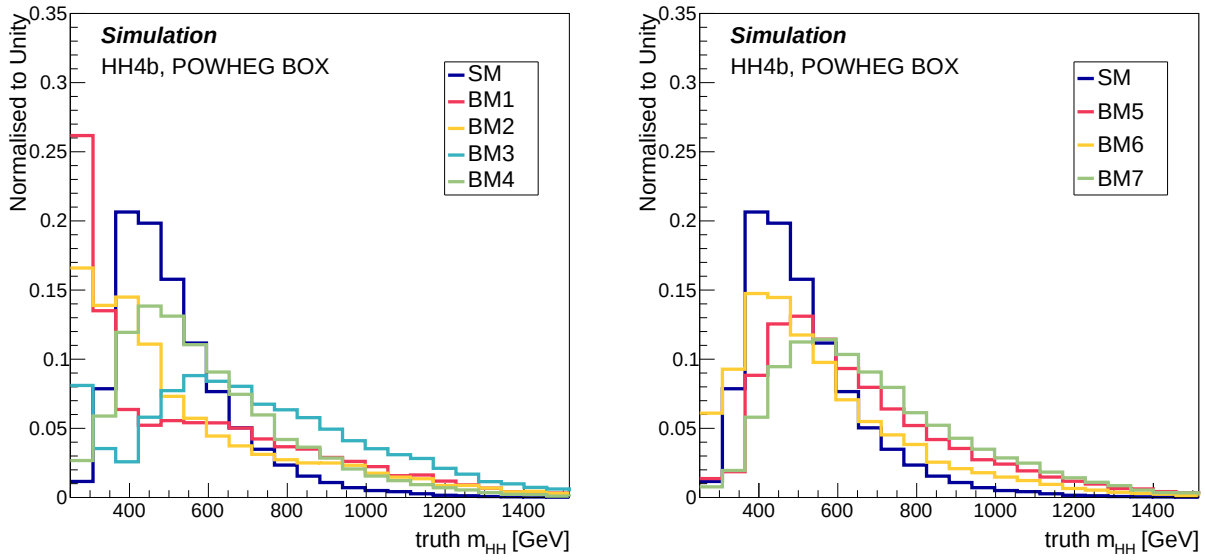
$$+\frac{\alpha_S}{8\pi}\left(c_{ggh}\frac{h}{v} + c_{gghh}\frac{h^2}{v^2}\right)G^a_{\mu\nu}G^{a,\mu\nu} \tag{2.8}$$

[33]. The so-constructed EFT assumes that modifications of Higgs couplings are the main sources of new physics in the EW sector and provides a basis for the $\kappa-$ framework [33]. The BSM diagrams that come about from non-zero c_{ttHH}, c_{ggH} and c_{ggHH} contributions are shown in Figure 2.5. Benchmark models shown in Table 2.1 are defined to assess the allowed parameter space taken by HEFT parameters as they cannot be constrained from single Higgs measurements [33, 41]. As seen in Figure 2.6 the seven benchmarks cover a range of different final state kinematics manifested in the m_{HH} spectrum.

The SMEFT and HEFT interpretations include both the linear and quadratic terms, where the linear contributions are the interference effects between SM and new physics, while the quadratic terms are pure BSM dynamics interactions.

Benchmark Model	κ_λ	κ_t	c_{ggH}	c_{ggHH}	c_{ttHH}
SM	1	1	0	0	0
BM1	3.94	0.94	$\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$
BM2	6.84	0.61	0	$-\frac{1}{3}$	$\frac{1}{3}$
BM3	2.21	1.05	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{3}$
BM4	2.79	0.61	$-\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$
BM5	3.95	1.17	$\frac{1}{6}$	$-\frac{1}{2}$	$-\frac{1}{3}$
BM6	5.68	0.83	$-\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$
BM7	-0.10	0.94	$\frac{1}{6}$	$\frac{1}{6}$	1

Table 2.1: The Benchmark models as defined in [41] are defined in terms of the five most relevant Willson coefficients in the ggF HH production within HEFT. Note that κ_λ and κ_t are 1 in the SM, while the other coefficients are zero.



(a) (b)

Figure 2.6: The truth m_{HH} distribution for the HEFT BMs considered as compared to the SM HH production are shown. PowhegBox has been used to produce the HH events.

2.3 Single Higgs Production

The EW NLO corrections in single Higgs production modes offer an indirect way of measuring the Higgs self-coupling. Examples of how anomalies in the Higgs self-coupling can affect single Higgs production are shown in Figure 2.7. On the one-loop level, the Higgs self-energy has a quadratic dependence on κ_λ, while the production mode and decay channel each depend on κ_λ linearly. The general reference for this section is [33].

The impact of κ_λ on the differential cross-section is parametrised in the $\kappa-$framework, taking into account κ_λ at NLO and κ_τ, κ_t, κ_V and κ_b at LO. According to [42, 43, 44], the deviation from the SM can be quantified as the multiplicative correction

$$\mu_{fi} = \mu_i \times \mu_f = \frac{\sigma_i^{BSM}}{\sigma_i^{SM}} \times \frac{BR_f^{BSM}}{BR_f^{SM}}, \tag{2.9}$$

where μ_i modifies the signal strength of the production mode $i \to H$ and μ_f modifies the branching fraction of the decay $H \to f$.

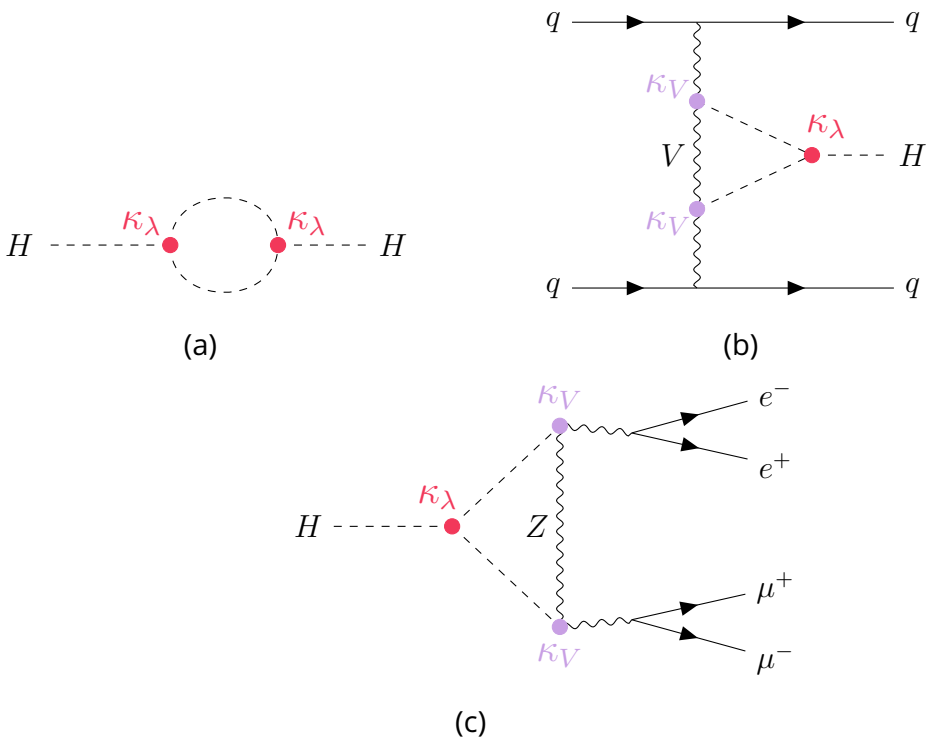


Figure 2.7: The impact of the trilinear Higgs self-coupling on single Higgs measurements via NLO corrections is shown. The quadratic dependence on κ_λ of the Higgs self-energy is depicted in (a). The diagram in (b) shows the dependence of the VBF production channel on κ_λ, while in (c) its impact on the Higgs decaying to two Z bosons can be seen.

The effect of anomalous Higgs couplings on the signal strength is

$$\mu^i(\kappa_\lambda, k_i) = Z_H^{BSM}(\kappa_\lambda)\left[k_i^2 + \frac{(\kappa_\lambda - 1)C_1^i}{K_{EW}^i}\right], \tag{2.10}$$

where

- i is the production mode ($i \to H$).
- C_1^i quantifies the linear dependence of i on κ_λ, which is process and kinematics specific. Details on its calculation are in [44].
- $k_i^2 = \frac{\sigma_{LO,i}^{BSM}(\kappa_\lambda=1)}{\sigma_{LO,i}^{SM}(\kappa_\lambda=1)}$, takes into account changes due to the other κ modifiers

considered, when $\kappa_\lambda = 1$.

- $K^i_{EW} = \frac{\sigma^{SM}_{NLO,i}}{\sigma^{SM}_{LO,i}}$ is the NLO k-factor.
- $Z_H^{BSM}(\kappa_\lambda) = \frac{1}{1-(\kappa_\lambda^2-1)\delta Z_H}$ with $\delta Z_H = -1.536 \times 10^{-3}$, renormalizes the Higgs wave function [42]. This is needed, due to the modifications in the self-energy of the Higgs shown in 2.8a.

The branching fraction changes as

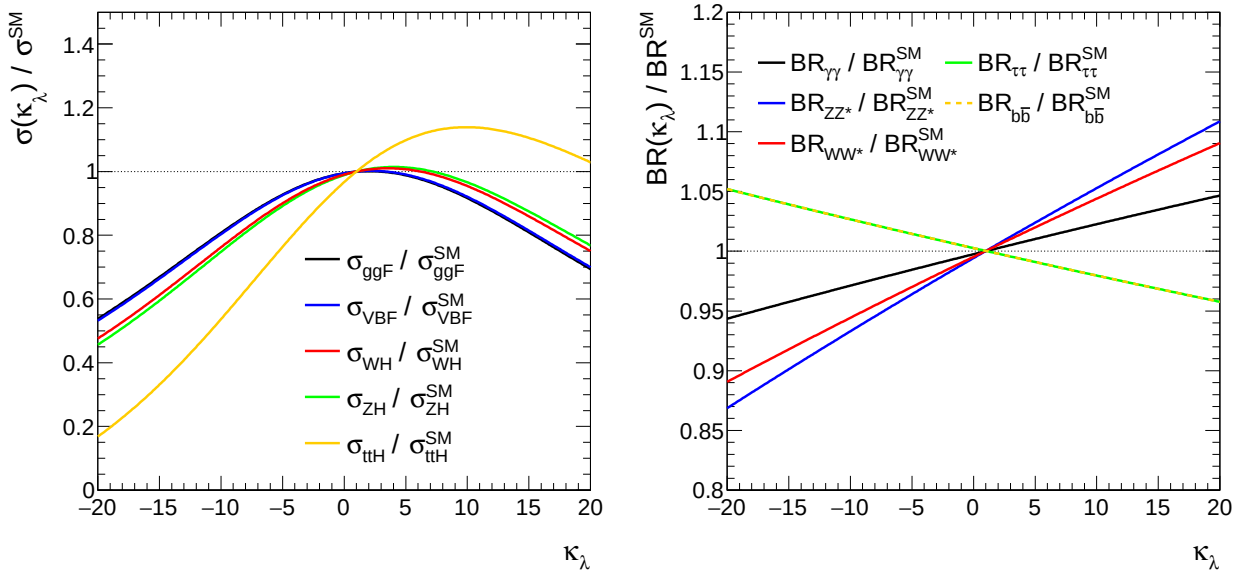
$$\mu^f(\kappa_\lambda, \kappa_f) = \frac{BR_f^{BSM}}{BR_f^{SM}} = \frac{\kappa_f^2 + (\kappa_\lambda - 1)C_1^f}{\sum_j BR^{SM}(j)[\kappa_j^2 + (\kappa_\lambda - 1)C_1^j]}, \tag{2.11}$$

where f in the decay mode of the Higgs, j in $\sum_j$ runs over all Higgs decay channels and BR_j is the SM decay rate of j. There is no modification coming from the Higgs self-energy as it cancels in the Higgs decay [44].

The multiplicative correction factors as a function of κ_λ are shown in Figure 2.8. Despite the smaller significance of κ_λ on the single H measurements as compared to the diHiggs production, the method still proves to be competitive as precision measurements of various single H channels and their combination are possible [33].

2.4 The Higgs potential

The determination of the Higgs potential is not only a probe of the electroweak theory but also plays a crucial role in cosmology. In the SM it is entirely predicted by the Higgs self-couplings, which are known from the vacuum expectation value (vev) and Higgs mass (m_H). However, given there exist yet undiscovered particles that could couple to the Higgs, the self-couplings would be modified with possible significant implications on the evolution of the early Universe as well as its ultimate fate.



(a) (b)

Figure 2.8: The modification of the SM production signal strength, μ^i, is shown in (a), while the branching fraction correction, μ^f, is shown in (b). The figure is taken from [42], which calculates μ^i and μ^f according to Equation 2.10 and 2.11 using the values for C_1^j, K_{EW}^j and k_j^2 as obtained by [43, 44].

In the early Universe, the temperature dependence of the Higgs potential determined the nature of the electroweak phase transition (EWPT)[5]. When the Universe expanded and cooled down, a second minima developed and co-existed at a critical temperature (T_C) with the minima at zero [45, 46]. The EWPT is of second-order if the transition from the symmetric to the broken EW phase is smooth and at no point is the Universe in a false minimum. On the other hand, the EWPT undergoes a first-order phase transition if at $T < T_C$ there is a local minimum at zero and the true minima with a non-zero vev. In this case, the Universe is out-of-equilibrium, and parts of it tunnel to the true minima initiating bubble nucleation[6]. The state inside the bubbles has broken EW symmetry, while outside the EW symmetry is unbroken. The bubbles expand, collide, and merge until the whole Universe is in the true EW minima. For electroweak baryogenesis to take place the first-order transition needs to be strong, as explained in [47]. In the SM the

[5]The transition of the Universe from the symmetric $SU_L(2) \times U(1)$ state with a zero vev to the broken symmetry state with the remnant $U(1)_{EM}$ symmetry and a non-zero vev.

[6]Some additional conditions need to be met for this to happen, see [45, 46] for details.

EWPT of second order fails to generate the out-of-equilibrium phase required for EW baryogenesis. However, the effective Higgs potential depends on the trilinear Higgs self-coupling, and experimental constraints still allow for a first-order phase transition [47].

The shape of the Higgs potential also determines the stability of the EW vacuum, which must be stable or metastable. Given the current measurements of the Higgs and top mass, it is metastable with a 95% confidence level within the SM [48]. Similarly, as in the EW baryogenesis discussion, metastability implies that the Universe could tunnel to the true minima at any point and the known Universe would cease to exist. The vacuum (meta)stability requirement sets bounds on the allowed anomalies in the Higgs self-coupling. However, as the modifications of the Higgs self-coupling could arise from various BSM sources which themselves can impact the evolution of the Higgs potential they are model dependent [49, 47].

2.5 New Physics

Triggering EWSB by a single $SU_L(2)$ doublet is the simplest but still arbitrary choice made within the SM to generate gauge boson masses. This simplistic nature of the Higgs sector might not be the whole story, as suggested by the outstanding problems of the SM, described in Section 1.8, and the rather complex structure of the fermion and gauge sectors [50]. Several proposed BSM theories, such as composite Higgs models [50], SUSY [28] or the two-Higgs-doublet model [18], depart from the single $SU(2)$ doublet assumption and extend the Higgs sector. Independent of the exact nature of the BSM Higgs sector, Higgs self-couplings as well as other properties of the Higgs would deviate from the SM predictions. The Higgs couplings can also be modified through Higgs portal models [51, 52], which postulate the

existence of a hidden sector that is coupled to the SM only through the Higgs[7].

In composite Higgs models the Higgs boson is a bound state rather than an elementary particle. For instance, in the Strongly Interacting Light Higgs model [53] the strongly-coupled sector produces a SM Higgs-like pseudo-Goldstone boson. In this scenario, the couplings would be close to the SM prediction, but even small deviations could be observed as enhancements in the boosted diHiggs production. This is because in VBF production the Higgs couplings to the gauge boson could deviate from the $c_V^2 - c_{2V} = 0$ parabola, leading the longitudinal vector boson scattering amplitude to grow with centre-of-mass energy [50].

Current LHC measurements, see [54] and [55], however, disfavour a composite Higgs and suggest that the observed scalar boson is an elementary particle in the $SU_L(2)$ doublet representation. Besides the SM, two-Higgs-doublet models (2HDM) [18], which assume two scalar $SU_L(2)$ doublet fields, are possible theories consistent with LHC measurement within a parameter space. The particle spectrum of the Higgs sector in the 2HDM consists of two charged particles, one neutral pseudo-scalar and two neutral scalars, one of which is the observed SM-like Higgs. The 2HDM is popular and well-motivated as it can provide a DM candidate and is required by SUSY. Under certain assumptions, it can be an additional source of CP violations and modifications of the Higgs potential needed in EW baryogenesis. One of the main challenges in constructing a plausible 2HDM is the suppression of tree-level flavour violating neutral currents while taking into account EW precision measurements [56]. Different strategies can be employed to solve the problem and as a result, there are several types of the 2HDMs. Hence, it is not possible to place an upper bound on the size of the Higgs self-coupling as it depends on the 2HDM type and free parameters within the chosen theory itself [18].

As tweaking the Higgs sector could solve several shortcomings of the SM, from

[7]Compostie Higgs models, as well as some types of two-Higgs-doublet model, can also be Higgs portal models.

baryogenesis to DM candidates, it is likely for new physics to manifest through the Higgs sector. Regardless if nature is manifested as a type of 2HDM or any other Higgs portal model, given there are new states that couple to the Higgs modifications of the Higgs self-couplings are expected. Extracting limits on the HHH vertex strength from observations complements direct searches for new particles of the extended Higgs sector and is crucial to narrow down the vast space of allowed BSM extension on the Higgs sector. Expanding the measurements to the HHVV vertex strength or considering alternative interpretations of the data through EFTs offer an additional more general insight into what is possible.

Chapter 3

The ATLAS Experiment

This thesis is based on the proton-proton collision data collected by the ATLAS detector between 2015-2018 at a centre-of-mass energy of $13\,\mathrm{TeV}$ during The Large Hadron Collider (LHC) Run 2. The operation and performance of the LHC are outlined in Section 3.1. The experimental apparatus, i.e. the ATLAS detector, during the Run 2 period is described in the 3.2. The general reference for the LHC Section is [57], while the general reference for the ATLAS Experiment is [58].

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator and collider in the world with a circumference of $26.7\,\mathrm{km}$ [57]. Located approximately $100\,\mathrm{m}$ underground on the French-Swiss border close to Geneva in the tunnel that once hosted the Large Electron-Positron Collider (LEP) it is designed to collide two oppositely circulating beams of protons, with a per beam energy of $7\,\mathrm{TeV}$ as well as lead ions with a $2.8\,\mathrm{TeV}$ of energy per nucleon. Due to the difficulty of training the superconducting dipole magnets, the centre-of-mass energy of $14\,\mathrm{TeV}$ has never been reached.

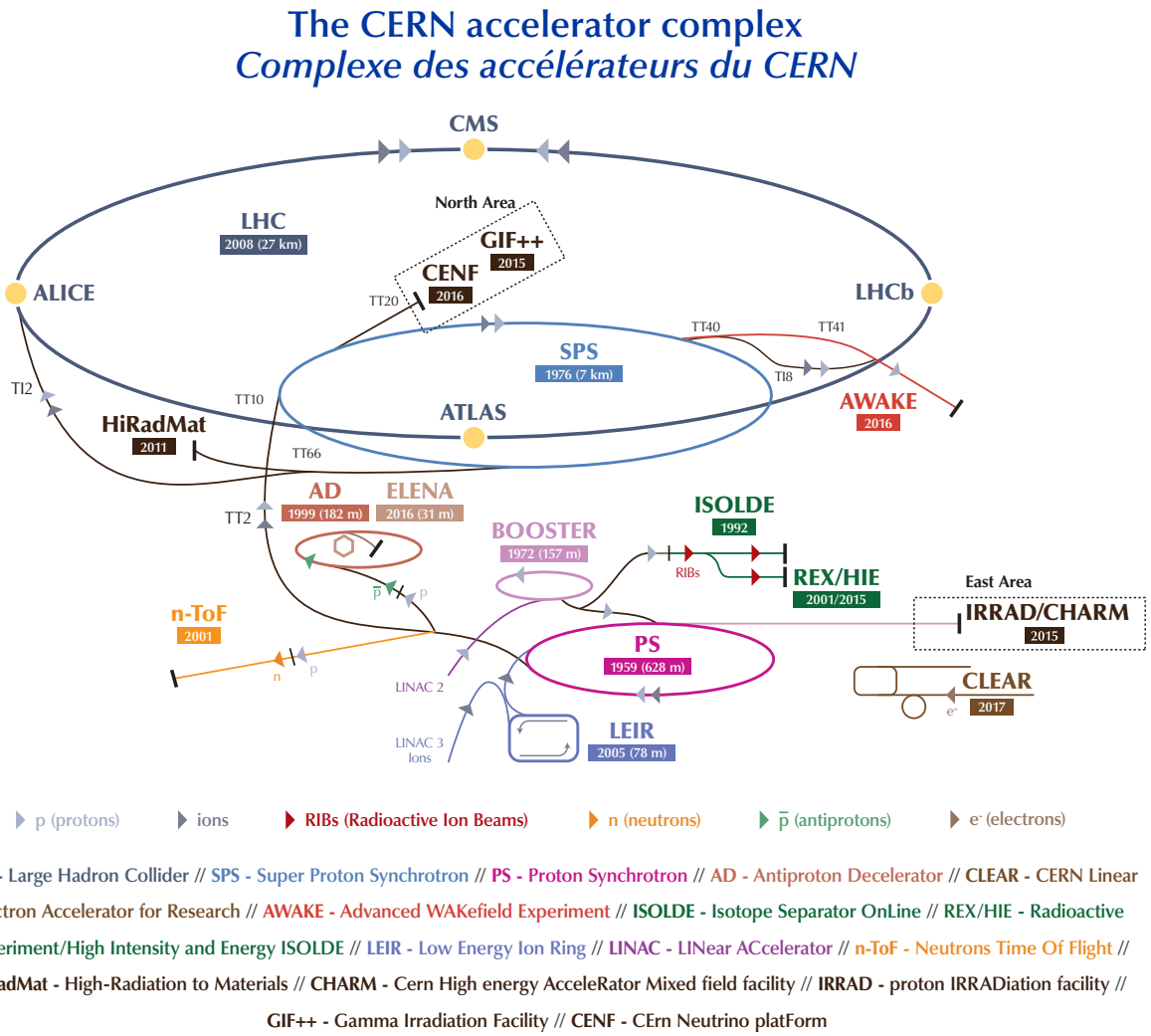


Figure 3.1: The LHC complex during Run2 is shown. The acceleration chain of the LHC consists of the Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). The accelerated proton beams at an energy of $450\,\mathrm{GeV}$ are injected into the LHC from the SPS (line T18). The proton-proton beams are crossed to collide at four interaction points marked on the plot as yellow circles. At each interaction point, one of the four LHC experiments is placed, ATLAS, CMS, LHCb, or ALICE. The Figure is taken from [59].

As can be seen in Figure 3.1 the acceleration of the protons is done in four stages after which the beam is finally injected into the LHC and further accelerated before the two beams are crossed to collide. The LHC injection chain consists of the Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS).

The acceleration chain starts by injecting protons obtained from a hydrogen atom which electrons are stripped away by an electric field into the Linac2[1]. The protons are accelerated to $50\,\mathrm{MeV}$ before being injected into the PSB, where they are further pushed to $1.4\,\mathrm{GeV}$. The proton energy is increased to $25\,\mathrm{GeV}$ in the PS. Finally, the SPS gives a final push and the beam energy is raised to $450\,\mathrm{GeV}$ before injecting the beam into the two pipes of the LHC in bunches of approximately 10^{11} protons separated by $25\,\mathrm{ns}$. The two proton beams are counter-circulating and accelerate to their maximum per beam energy of $6.5\,\mathrm{TeV}$ in less than half an hour. The beams are focused, crossed and made to collide at four different interaction points of the LHC, each hosting an experiment: the two general purpose detectors ATLAS and CMS [60], LHCb [61], built to precisely measure the charge-parity violation in nature and ALICE [62] whose primary design focuses on collecting lead-lead collision data.

3.1.1 Synchrotrons

Collisions at centre-of-mass energy on a couple of a TeV scale are possible thanks to the innovation of synchrotrons. As opposed to their predecessors, the cyclotrons, synchrotrons can accelerate the particles without changing the radius of curvature by changing the magnetic field. In the case of the LHC, the energy of the protons is increased by eight radiofrequency (RF) cavities each operating between $4\,\mathrm{MeV}$ and $14\,\mathrm{MeV}$ [63]. The changing electromagnetic wave acts like a self-regulator, the protons that arrive too late (early) will experience a different acceleration according to their position in the EM wave and therefore arrive earlier (later) at the next RF

[1] Linac2 has been replaced with Linac4 in the currently ongoing Run 3.

cavity focusing the bunch. The protons are kept on a circular path by 1232 superconducting dipole magnets made out of niobium-titanium cooled to $1.9\,\mathrm{K}$. In order to maximise the number of collisions the beam needs to be focused, which is achieved by 858 quadrupole magnets that, depending on the positioning of the poles either focus the width or height of the beam.

3.1.2 Run 2 Performance

The physics that can be explored by the LHC is dictated by the centre-of-mass energy and the integrated luminosity delivered by the machine. The number of events expected to occur for a given process is:

$$N_{event} = L\sigma_{event} = \int dt\mathcal{L}\sigma_{event}, \tag{3.1}$$

where L is the integrated luminosity, in other words, the instantaneous luminosity ($\mathcal{L}$) integrated over the sensitive time t. The cross-section of the process is σ_{event} and depends on the centre-of-mass energy. The centre-of-mass energy is limited by the maximum magnetic field that can be carried by dipole magnets, while the instantaneous luminosity according to [64, 65] is completely determined by beam parameters:

$$\mathcal{L} = \frac{N_1 N_2 n_b f_{rev}}{4\pi\sigma_x\sigma_y} \cdot F. \tag{3.2}$$

The fraction in the above formula represents the head-on collision of two Gaussian beams, where N_1 and N_2 are the per bunch proton populations, n_b is the number of bunches per beam, σ_x and σ_y are the Gaussian widths of the beam and f_{rev} is the frequency of revolution. However, the beams at the LHC are made to collide at a small crossing angle, reducing the luminosity by the geometric luminosity reduction factor, F. The crossing angle is needed as before colliding several bunches from the two beams share a vacuum pipe for more than $120\,\mathrm{m}$ and the introduction of a

crossing angle mitigates beam-beam effects.

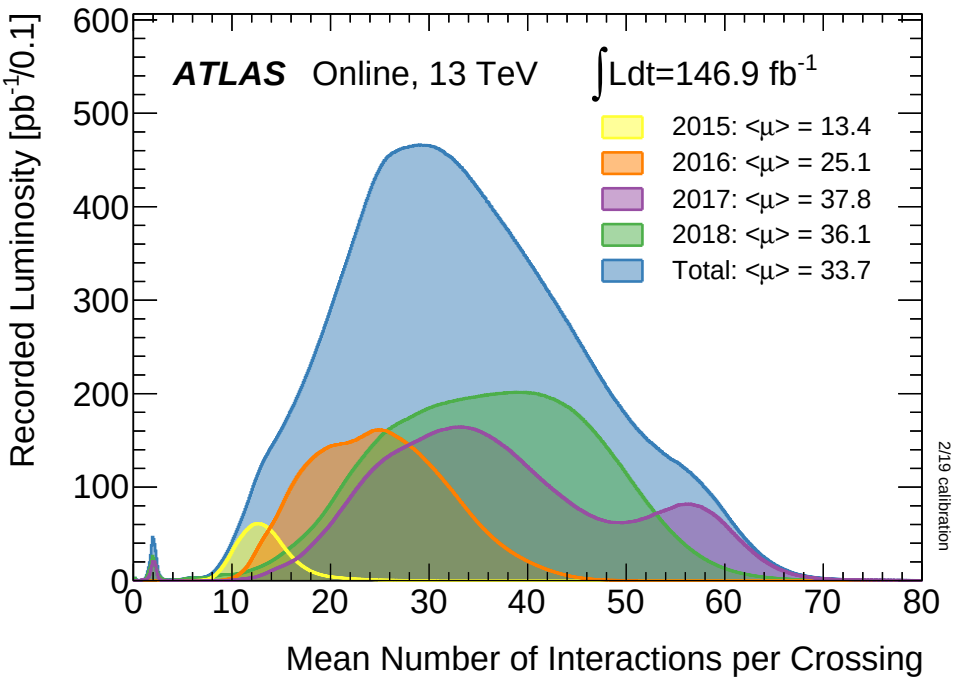


Figure 3.2: The mean number of interactions per crossing as a function of the recorded luminosity. The means of the number of interactions per year of data taking and the total integrated luminosity collected are stated on the plot taking into account all data collected during stable beam conditions in Run 2. The Figure is taken from [66].

During Run 2 of the LHC operation, the centre-of-mass energy was $13\,\mathrm{GeV}$, while exceeding expectations during the last two years of operation when the maximum instantaneous luminosity recorded was $2.07 \times 10^{34}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ [67]. Figure 3.2 shows that the mean number of interactions was 33.7 recorded by the ATLAS experiment during Run 2. It follows that each collision of interest will have a background made of tens of other collisions, referred to as pile-up. In-time pile-up refers to background collisions that occur in the same bunch crossing as the event of interest, while out-of-time pile-up refers to background collisions from a previous or later bunch crossing that signature in the detector is still present during the collision of interest.

3.2 The ATLAS Detector

With a diameter of $25\,\mathrm{m}$ and length of $44\,\mathrm{m}$, the ATLAS detector is one of the biggest detectors ever built for particle physics and is described in detail in [58], which is the general reference for this section. As it is a general-purpose detector, its design allows it to target (nearly) everything; from dark matter searches, and SM precision measurements in proton-proton collisions to the study of the phase transitions in the quark-gluon plasma of heavy ion collisions. The detector reconstructs the energy, momentum, and tracks of almost all visible particles emerging from the proton-proton collisions at the TeV scale. This excludes neutrinos whose signature in the detector can be missing energy. The physics objects are reconstructed by combining the measurements of several sub-systems; trackers, calorimeters, and muon chambers, all immersed in a magnetic field covering nearly the whole solid angle around the collision points, as shown in Figure 3.3.

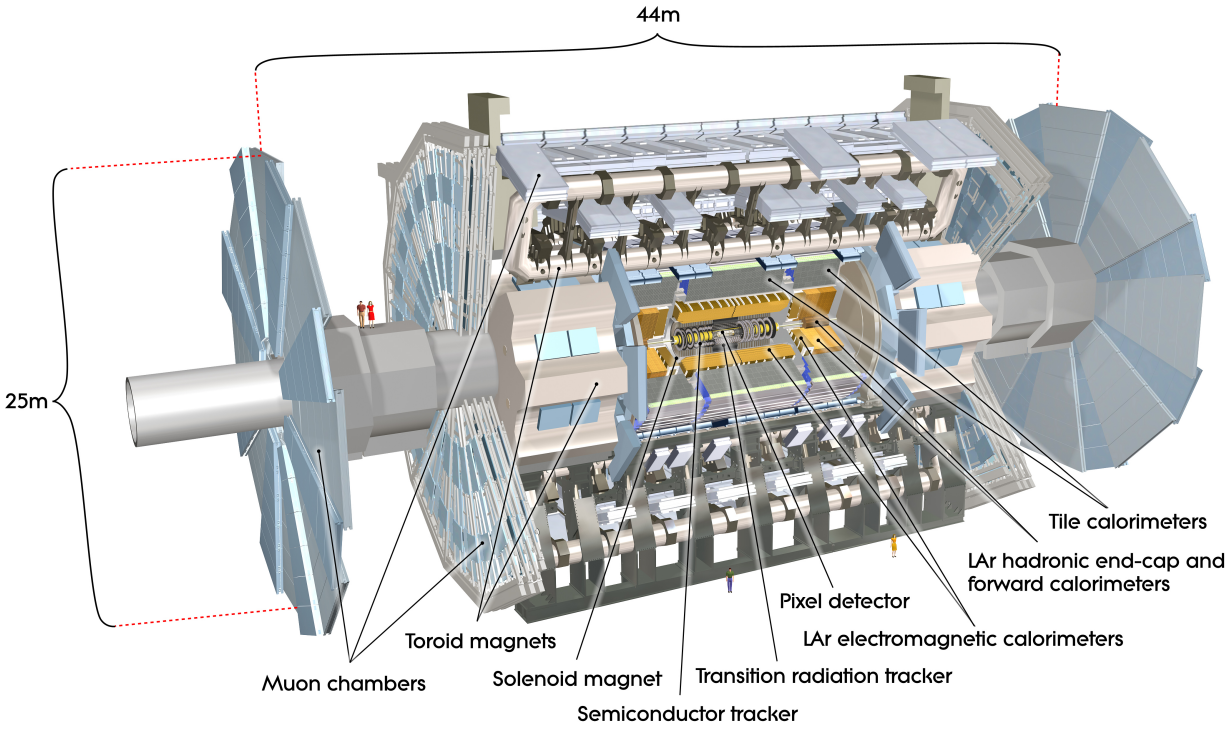
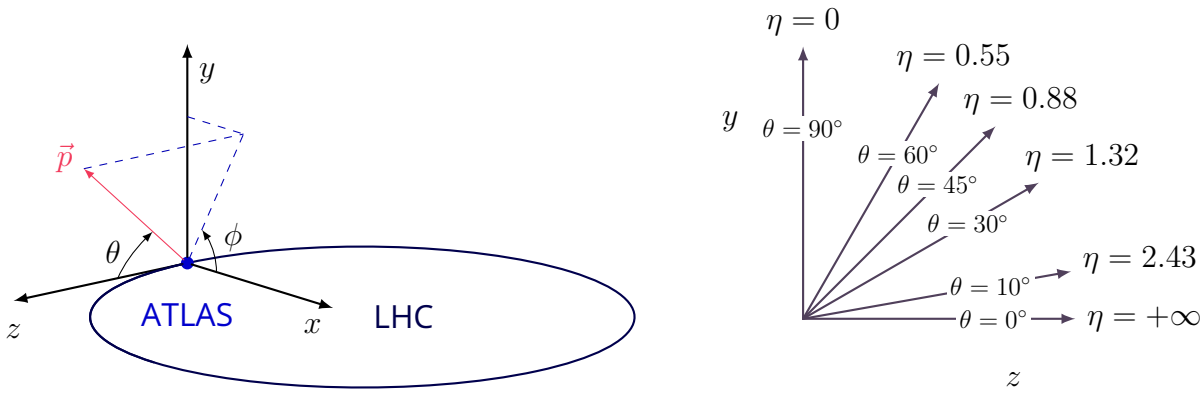


Figure 3.3: The ATLAS detector at the LHC as operated during Run 2 is depicted. The subdetector systems are labeled on the plot that has been taken from [68].

3.2.1 The Coordinate System

The right-handed coordinate system used by the ATLAS Experiment is depicted in Figure 3.4a. The origin is set at the proton-proton interaction point and the z-axis is defined along the beam axis, where the part of the detector on the positive z axis is labelled as the A-side while the negative z axis is the C-side. The $x-y$ plane is perpendicular to the beam direction and used to define the transverse momentum (p_T), transverse energy (E_T), and missing transverse energy (E_T^{miss}). The angles defined to the beam axis are the around-pointing azimuthal angle (ϕ) and the away-pointing polar angle (θ). It is the convention to measure the angular separation of a particle pointing away from the beam axis with pseudorapidity[2], $\eta = -\ln\tan(\frac{\theta}{2})$, instead of θ. Figure 3.4b illustrates the relation between θ and η in the $x-y$ plane. The angular geometrical distance ΔR is commonly used to measure the separation between physical objects, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.



(a) (b)

Figure 3.4: An illustration of the coordinate system used by ATLAS to describe the position of particles emerging from a proton-proton collision. The coordinate system as defined with respect to the LHC ring (black circle) for the ATLAS detector (blue dot) is depicted in (a). The momentum of a particle emerging from the proton-proton collision is shown as $\vec{p}$ (dark pink vector). The relationship between η and θ in the $y-z$ plane is shown in (b). The graphics in (a) is adapted from [69] and the graphics in (b) from [70].

[2]The rapidity of a massless particle is the same as the pseudorapidity (η).

3.2.2 The Inner Detector

The inner detector (ID) is positioned closest to the interaction point and is responsible for tracking in the $|\eta| < 2.5$ region. It consists of three sub-detectors, as shown in Figure 3.5; the pixel and silicon microstrip (SCT) detector in the $|\eta| < 2.5$ region and Transition Radiation Tracker (TRT) in the $|\eta| < 2$ region all immersed in a $2\,\mathrm{T}$ magnetic field produced by a solenoid magnet. The Pixel and SCT provide high-precision tracking, while the TRT compensates for its lower granularity by recording a large number of tracking points per particle (generally > 30) [71]. Each sub-detector has a barrel component and two end-caps to ensure a high resolution in the ϕ and z directions. The precise coverage of each ID component is shown in Figure 3.6, while the intrinsic spatial resolution in the $r - \phi$ and z direction is shown in Table 3.1. The high resolution of the tracker is crucial for jet reconstruction and jet flavour identification (flavour tagging) as discussed in Chapter 5.

Sub-detector	**Intrinsic resolution in $r - \phi$ [μm]**	**Intrinsic resolution in z [μm]**
IBL [72]	10	60
Pixel [58]	10	115
SCT [58]	17	-
TRT [58]	130	-

Table 3.1: The intrinsic resolution of each ID sub-detector in the $r-\phi$ and z directions.

Pixel Detector

The sensors that make up the Pixel detector have to tolerate high radiation levels and maintain a good charge-collection efficiency after possible type inversion due to radiation effects. Therefore, the semiconductor sensors are based on a double-side processing technique where $p+$ and $n+$ regions are implemented on both sides of the wafer while the n-type bulk material is enriched with oxygen impurities [75, 58]. The size of approximately 90% of the sensors is $50 \times 400\,\mu\mathrm{m}^2$, while the dimensions of the remaining are $50 \times 600\,\mu\mathrm{m}^2$.

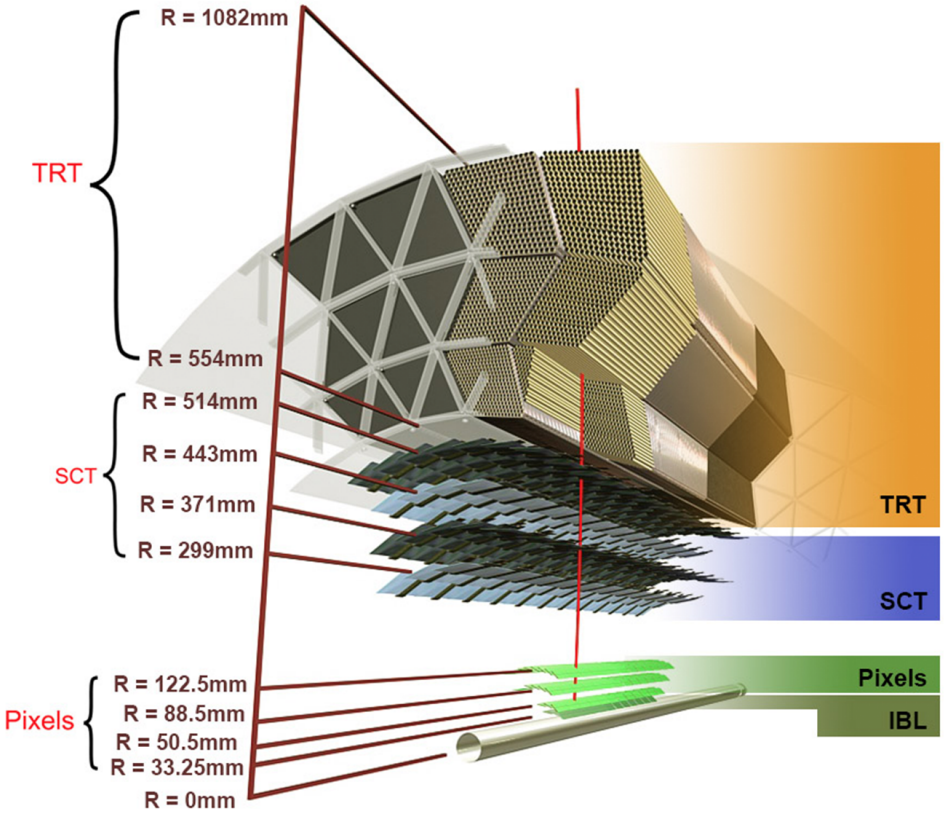


Figure 3.5: The cutout of the Inner Detector, each subdetector and its distance from the beam axis is indicated. The figure is taken from [73].

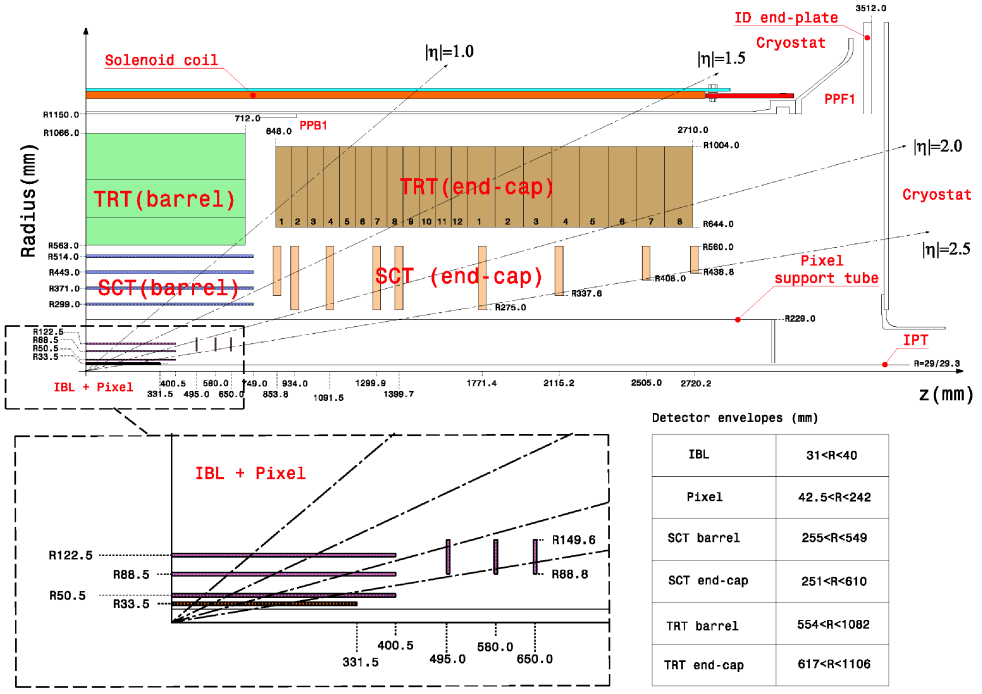


Figure 3.6: The precise layout of the Inner Detector also shows the inner positioning tube (IPT), mechanical support structures, and solenoid coils. The graphic is taken from [74].

The layering of the pixel detector ensures that a track leaves a signature in at least three layers. There are four layers in the barrel with the sensors parallel to the beam pipe, as the Insertable B-Layer (IBL) [76, 74] was installed before the start of Run 2. The IBL is located between the beam pipe and the previous innermost layer of the pixel and significantly improves b-tagging through better reconstruction of the B-hadron decay vertex (secondary vertex). On the other hand, the side-A and side-C end-caps consist of three disk layers with the sensors positioned perpendicular to the beam pipe. The layering enables the accuracy of up to $10\,\mu\mathrm{m}$ in $R-\phi$ and $115\,\mu\mathrm{m}$ in z directions in the barrel region.

Semiconductor Tracker

The SCT sensors are based on semiconductor detectors as the pixel sensors but use p-in-n single-sided sensors instead of the double-sided technology as the latter would be too costly [77]. The geometrical design is such that each track crosses at least four double layers of the SCT, so similar to the pixel detector there are four layers of co-axial sensors in the barrel region and nine disks in each end-caps. The strip detectors are $80 \times 64\,\mu\mathrm{m}^2$ in the barrel region, enabling an accuracy of $17\,\mu\mathrm{m}$ in $R-\phi$.

Transition Radiation Tracker

The outermost part of the ID covering the $|\eta| < 2.0$ range is the TRT made of 120 Polyimide drift (straw) tubes of $4\,\mathrm{mm}$ diameter. Its resolution is $130\,\mu\mathrm{m}$ per straw in the $R-\phi$ plane. It provides good particle identification, due to the differences in the transition radiation produced in the Xenon based gas mixture part. Electrons and pions are distinguished by the number of hits produced that exceed a high threshold[3], while charged particles, in general, are identified through $\frac{dE}{dx}$ [78]. Due to

[3] A lower threshold on the pulse is used to identify hits used for track reconstruction. Electrons produce significantly more hits high threshold than pions.

leakage issues part of the detector is filled with an Argon-based gas mixture as it is much cheaper [79].

3.2.3 The Calorimeter System

Sampling calorimeters are used to measure the energy of all particles while preserving the directional information through interchanging layers of absorption and active material. The calorimeter system consists of the Electromagnetic Calorimeter (ECal) and the Hadron Calorimeter (HCal). The ECal is desinged to measure the energy of photons and electrons precisley, while the HCal measures the energy of hadrons. The absorption material causes the incoming particle to shower as it is made of a material with a large radiation lengths X_0 (nuclear interaction length λ), usually a metal, in the ECal (HCal). On the other hand, the active layer has a shorter $X_0(\lambda)$ and samples the energy of the particles produced in the shower. Figure 3.7 shows the loss of energy of a particle through electromagnetic showering in the ECal, where an electron radiates particles approximately every X_0 with an energy loss of $\frac{E}{e}$. Hadrons start showering in the ECal as well, but the hadron component of the shower dominates only in the HCal, where the majority of their energy is lost. The energy resolution of the calorimeter system is characterised by the equation

$$\frac{\sigma(E_T)}{E_T} = \frac{N}{E_T} \oplus \frac{S}{\sqrt{E_T}} \oplus C, \tag{3.3}$$

where N is the noise term, S is the stochastic term, C is the constant term, and E_T is measured in GeV [80]. The constant term captures resolution effects that do not depend on p_T, such as energy deposition outside the active material and the non-uniform calorimeter response. The terms extracted from charged pion measurements in the central barrel region ($\eta = 0.3$) are $N = 1.0\%$, $S = 50\%$, $C = 3.4\%$, while the resolution deteriorates in the extended barrel region ($\eta = 1.3$) to $N = 1.5\%$,

$S = 68\%$, $C = 3.0\%$ [81].

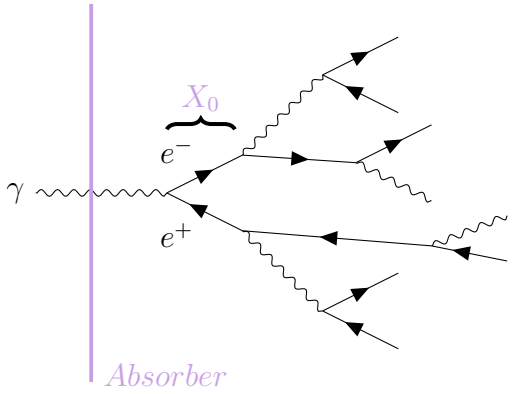


Figure 3.7: A simplified illustration of an electromagnetic shower initiated by a photon entering an absorption material with a radiation length X_0.

Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECal) consists of lead absorption plates and liquid argon (LAr) layers folded together with the read-out electronics in an accordion-like shape. Particles entering the active material ionize LAr, the so-produced ionization is linearly proportional to the energy of the ionizing particle and can be measured by the electronics. Similarly, as the ID, the ECal consists of the barrel covering $|\eta| < 1.475$ and the A- and C-side end-caps covering the $1.375 < |\eta| < 3.2$ with a higher granularity in the region bordering the ID. The depth of the ECal is $X_0 > 22$ in the barrel and $X_0 > 24$ in the end-cap region, which is enough to contain all electromagnetic showers. The granularity of the cells changes with the depth of the ECal. For instance, in the central region of the barrel ($\eta = 0$) the $(\eta \times \phi)$ cell size is 0.0031×0.098, 0.025×0.0245 and 0.05×0.0245 in layer 1, 2 and 3 respectively.

Hadron Calorimeter

The tile calorimeter is located in the barrel ($|\eta| < 1.0$) and extended-barrel ($0.6 < |\eta| < 1.7$) region. Steel is used as the absorption material and the active

material is made of scintillating tiles. The tiles emit light when hit by charged particles, which is read out by wavelength-shifting fibres connected to photomultiplier tubes as the signal. The choice of active material and geometry was partially driven by the budget. The end-cap of the hadronic calorimeter covers the $1.5 < |\eta| < 3.2$ region with LAr used as the active material and copper as the absorption material. The forward hadron calorimeter covers the $3.1 < |\eta| < 4.9$ and as the end-cap partially uses LAr/Copper technology. Further away in the z direction in the forward hadron calorimeter copper is replaced with tungsten, which is better for energy measurements of hadrons. The depth of the HCal is between 9.7 λ in the barrel and 11 λ at $\eta = 0$. The HCal is expected to stop all particles from punching through with the exception of muons and multi-TeV jets.

In the ECal and HCal, the cells are combined into trigger towers with the size $(\Delta\eta \times \Delta\phi) = 0.1 \times 0.1$ (slightly larger at higher η), which serve as an input to the L1 hardware trigger.

3.2.4 The Muon Spectrometer

The muon spectrometer (MS) is immersed in a magnetic field and measures the deflected tracks of muons, at least in the majority of cases, that punch through the calorimeter. It covers the pseudorapidity range $2.7 < |\eta|$ over which four subdetectors are arranged. The subdetectors responsible for precision measurements are the monitored drift tubes (MDT) and cathode strip chambers (CSC), while the triggering is done by the barrel resistive plate chambers (RPC) and thin gap chambers (TGC). The magnetic field, provided by the barrel and smaller end-cap air-core toroids, points perpendicular to the direction of the incident muons and has a bending power from $1\,\mathrm{T}$ to $7.5\,\mathrm{T}$. The general reference for this section is [82].

Monitored Drift Tubes

Monitored drift tubes cover a large portion of the pseudorapidity region, both in the barrel and end-cap and provide precision tracking information in the direction of the bending of the magnetic field. The tubes have a diameter of $30\,\mathrm{mm}$ and $400\,\mu\mathrm{m}$ wall thickness and are of varying length, between $70\,\mathrm{cm}$ and $630\,\mathrm{cm}$. The single wire resolution of an aluminium tube is $80\,\mu\mathrm{m}$.

Cathode Strip Chambers

Located close to the beam pipe at $2.0 < |\eta| < 2.7$ the cathode strip chambers (CSC) are of finer granularity and greater radiational hardness to withstand the higher backgrounds. Essentially, the CSC is a multiwire proportional chamber with a resolution of about $60\,\mu\mathrm{m}$, achieved by the segmentation of the readout.

Resistive Plate Chambers

Triggering in the barrel and an additional measurement in the direction parallel to the magnetic field are provided by the resistive plate chambers (RPC). The gap between two resistive plates is filled with $C_2H_2F_4$, which incident particles ionize. A high electric field between the plates causes the ionized particles to trigger an avalanche of charged particles that can be detected. The achieved $1\,\mathrm{cm} \times 1\,\mathrm{ns}$ resolution satisfies the triggering requirements needed for bunch crossing identification and p_T resolution.

Thin Gap Chambers

The thin gap chambers are the RPC used for triggering and direction measurements in the ϕ plane located in the end cap. The technology is very similar to the CSC with

the distinction that the distances between the anode-cathode, and anode-wire are not the same. A gas mixture of 55% CO_2 and 45% $n - C_5H_{12}$ is used.

3.3 The Trigger System

It is not possible to record all events, as the rate is about $40\,\mathrm{MHz}$. Therefore, a triggering system is needed that makes a decision whether or not an event is of interest for offline analysis. In Run 2 a two-stage system of trigger architecture was used, the hardware-based Level-1 (L1) Trigger and software-based High-Level Trigger (HLT) [83]. Details on the performance and operation of the Run 2 trigger system are documented in [83], which serves as the main reference for the summary provided in this section.

3.3.1 Level-1 Trigger

Limited information is taken from the calorimeter and muon spectrometer measurements to form a decision in approximately $2.5\,\mu\mathrm{s}$ if an event should be kept for further consideration. The L1 Calorimeter (L1Calo) trigger calculates the total and missing energy while identifying jets and other physics objects of interest, such as electrons and photons, that surpass the predefined thresholds [84]. The L1 Muon (L1Muon) trigger compares the hit pattern with that of a muon with infinite energy while placing additional requirements on coincidental hits between different parts of the muon spectrometer [85]. The information from L1Calo and L1Muon are combined to make a decision, if an event is kept for further processing based on topological requirements or physical objects passing thresholds [83]. Topologies of interest are identified by the L1 Topological Processor (L1Topo), which combines physical objects from L1Calo and L1Muon [86]. In this step, the rate is reduced to $100\,\mathrm{kHz}$, and the selected events together with found Regions-of-Interest (ROIs) in

the $\eta - \phi$ plane are passed to the HLT.

3.3.2 HLT

The HLT is a farm made of approximately 40,000 Central Processing Units, that make the final decision if an event is kept for offline analysis in less than $230\,\mu\mathrm{s}$. The software-based trigger first checks for early rejection by executing fast algorithms. Next more CPU-intensive algorithms are used that can require the full event information to make a decision, which is the case for E_T^{miss} reconstruction. Data is saved for online analysis at an average rate of about $1.1\,\mathrm{kHz}$.

In summary, the full detector is read out and stored for offline analysis, if an event contains a signature of interest, such as jets or E_T^{miss}. Each signature of interest has its trigger chain, which defines the required thresholds reconstructed physics objects need to meet to fire the trigger. The trigger chains used can vary from run to run and depend on the beam conditions and physics requirements. To mitigate the rate prescaled triggers are used sometimes, where only every n-th event that satisfied the trigger chain in question is saved for offline analysis.

Chapter 4

Simulation

Particles produced in a proton-proton collision usually cannot be directly detected by the experiment, as any unstable final states immediately further decay and partons produce a cascade of collimated particles (jets). More fundamentally, what is measured in the detector are hits and energy deposits of the stable particles produced produced. Hence, comparing what an experiment observes to $\mathcal{L}_{\mathcal{SM}}$ or $\mathcal{L}_{\mathcal{BSM}}$ is not as simple as using Equation 3.1 to calculate the expected number of events. Complex simulation, called Monte Carlo (MC) samples in ATLAS, are needed to model an event from the bunch crossing to the electrical signals observed by the detector.

It is not feasible to generate every signal process that might have occurred in a proton-proton collision. MC samples are generated only for processes of interest and all non-negligible backgrounds that could have produced the same signature in the detector. Markov-Chain Monte Carlo sampling techniques are employed to sample representative ensembles of events from underlying probability distributions of a process of interest. Key aspects of proton-proton collisions and their simulation, including the detector simulation, are described in this chapter. The general references for this section are [87, 88].

4.1 Event Generation

The first step in the simulation is event generation, from the scatterings in a bunch crossing to the formation of hadrons. A simplified schematic representation of a $pp \to t\bar{t}$ event, as generated by Pythia, is shown in Figure 4.1. Everything shown takes place before interactions with the detector.

The differential cross-section of a process factorises into largely independent contributions; the hard scatter (HS), parton shower (PS), hadronisation, and multiple interactions (MPI). This allows for more precise theoretical predictions. For example, the HS will determine the initial conditions of the parton shower, but its evolution is universal.

Hard-scatter interaction

The MC is constructed around a particular hard scatter (HS), which is the parton-parton interaction with the largest momentum transfer in a proton-proton collision. This is true in the large majority of cases, as the experiments at the LHC have been primarily designed to study HS events. The inclusive cross-section of an HS process $ab \to X$ can be written as;

$$\sigma = \sum_{a,b} \int d\boldsymbol{x} \int PDF(x_a, \mu_R) PDF(x_b, \mu_R) d\hat{\sigma}_{(ab\to X)}(\mu_F, \mu_R), \tag{4.1}$$

where

- $PDF(x, \mu)$ is the parton distribution function, which models the momentum carried by partons in the hadron.
- $d\hat{\sigma}$ is the differential cross-section of the HS. It depends on the matrix element, which can be calculated perturbatively and the final state phase space.

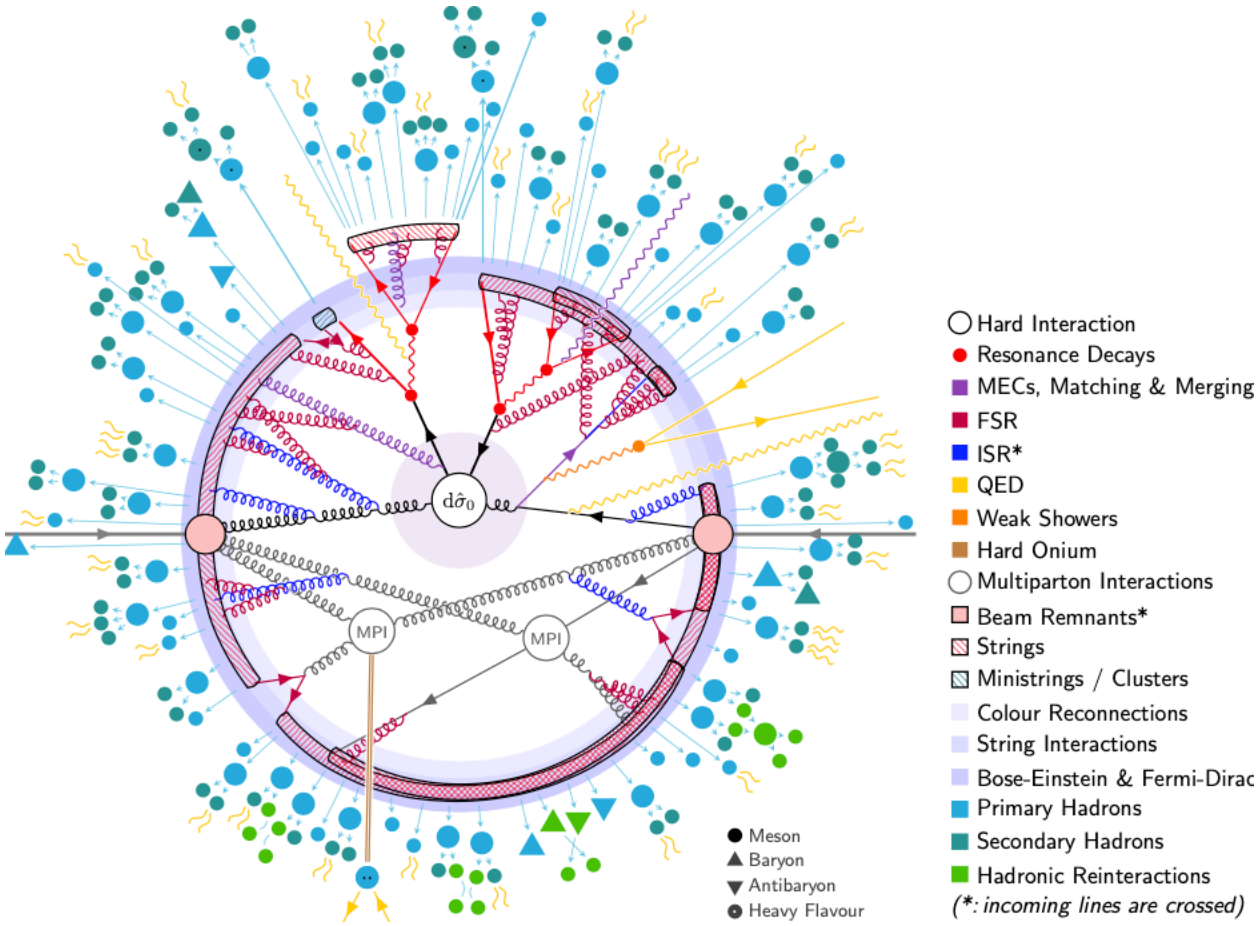


Figure 4.1: A simplified sketch of the structure of a $pp \to t\bar{t}$ event, as simulated by Pythia. The colliding protons, light pink circles, contribute partons to the hard scatter (HS), the white circle labelled $d\sigma_0$, and underlying events (UE), white circles labelled MPI. Before the HS or UE, the partons entering the interactions can emit initial state radiation, represented by dark blue lines. Similarly, the partons that emerge from the HS or UE can radiate, the final state radiation is indicated by purple lines. The $t\bar{t}$ final state and its subsequent decays are represented by the red circles. The purple and hatched areas represent hadronisation, the formation of colour singlets. Primary hadrons are shown as blue-coloured shapes, while green-coloured shapes are the hadrons from subsequent decays, usually what is observed by the detector. The simplifications made are; there are fewer final-state hadrons, the recoil is depicted incorrectly, light-flavour hadrons are shown as stable and the incoming momentum is not accurately depicted [89]. The illustration is taken from [89].

- μ_F, μ_R are the factorisation and renormalization scale.
- x is the fraction of momentum from the colliding hadron carried by the interacting parton.

Assumptions need to be made to make the cross-section calculation feasible. The matrix element calculation is perturbative but is usually terminated at leading or next-to-leading order, while the μ_F, μ_R are unphysical and determined from experience. The PDFs represent a non-perturbative regime applied at the perturbative scale[1]. They can only be determined empirically from deep inelastic scattering experiments and therefore carry an uncertainty. Lastly, the integral related to the phase space is rather complex, solved by employing Monte Carlo sampling techniques by event generators. Examples of generators that evaluate the matrix element of an HS and sample its phase space generating representative events are MadGraph5_aMC@NLO [90] or Powheg Box [91, 92, 93].

Parton shower

Incoming and outgoing partons at high momentum radiate gluons and quarks, which can themselves radiate, in a process called parton showering. The cascade of mainly soft and collinear gluons and quarks produced before the HS is classified as initial state radiation (ISR), while the one after the HS is final state radiation (FSR). The cascade of emissions evolves the momentum scale from the perturbative to the non-perturbative regime (about $1\,\mathrm{GeV}$) [88]. Thanks to Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations it is possible to model this p_T evolution, as they vary PDFs with changing momentum scale [94].

Simulated parton showers are not perfect representations of the underlying theory, as several assumptions go into the event generator. Soft and collinear

[1] This generalisation is called factorisation.

approximations for estimating the soft gluon emission term are needed. The total cross-section of the parton shower is calculated only to LO or NLO. A cut-off value is chosen to determine the point at which the parton showering process terminates and non-perturbative effects take over and as before PDFs and DGLAP equations have an uncertainty as they are derived empirically.

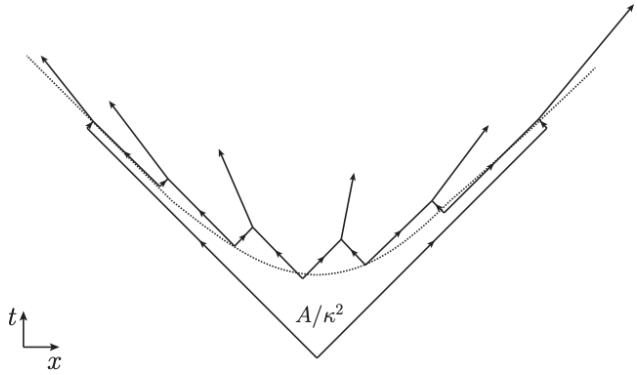


Figure 4.2: Illustration of the idea behind the Lund string taken from [89]. The light-like trajectory is the originally produced $q\bar{q}$ parent. The breaking of the string and consequential hadron production is indicated by the trajectories along the hyperbola.

Hadronisation

Once the produced partons lose enough energy through parton showering hadronisation, the process of forming colourless states (hadrons) begins. The formation of hadrons is only possible at low momentum transfer, where α_S is large and QCD becomes non-perturbative. Therefore, the formation of hadrons is described by phenomenological models. The Lund string model [95, 96], used by Pythia, and the cluster model [97], used by Herwig are employed most widely.

Hadrons are created collectively from colour confinement states, which are described differently in the two models. For example, the essence of the Lund string model is that the strong QCD field between partons is a string with a kink, shown in Figure 4.2. The kink represents the gluon at the interaction vertex, which created the colour connected $q\bar{q}$ pair. As the quarks separate the strength of the QCD field

between them grows until it becomes energetically feasible for $q\bar{q}$ pairs to tunnel out of the vacuum. The string starts breaking from the kink outwards until only colourless states are left. If the so-created hadrons are unstable, they future decay into lighter particles.

On the other hand, the cluster hadronisation model (used by Herwig) builds on the colour preconfinement of the parton shower discovered and described in [98]. In other words, the colour structure of the shower is independent of the energy of the HS and only determined by the colour composition of the partons involved in the HS [99]. Once the non-perturbative regime begins, colour singlet clusters are formed from the present partons, whose mass distribution is universal and depends on the parton shower energy scale [88]. A cluster undergoes fissions if its mass and available phase space allows it. Otherwise, colour reconnection with neighbouring clusters is needed before hadrons can be formed.

Multiple parton interactions (MPI)

The HS is not the only interaction in a proton-proton collision, as other protons in the bunch can scatter and produce additional final state particles. The majority of additional activity is soft and has a non-negligible effect on the number of final state partons. The final state partons can increase the number of the created hadrons through colour exchange with the soft scatters [88]. Additionally, underlying events, the cascade of particle interactions and decays initiated by the remnants of the protons involved in the HS, can be relatively hard and are possible sources of non-HS jets. MPIs are challenging to simulate as they involve both perturbative and non-perturbative interactions. In addition, the PDFs of the remnant partons from the HS are tricky to estimate in underlying events.

In this thesis MadGraph5_aMC@NLO [90] or Powheg Box [91, 92, 93] are used for matrix element calculation and event sampling. The event generators are combined

with either Pythia 8.2 [100] or Herwig 7 [101] to simulate the parton shower and hadronise the produced partons [88]. In this thesis, Pythia and consequently the Lund model has been used to model the hadronisation process. Alternative simulations are produced with Herwig to assess the systematic uncertainty on the parton shower and hadronisation model used by Pythia.

4.2 Detector Simulation

All generated particles that remain after hadronisation are propagated through a simulation of the ATLAS detector [87]. The geometry of the detector and the response of the active material with the incoming particles is modelled by the Geant4 simulation toolkit [102]. Interactions of particles with the active material of the detector are saved as hits, i.e. the energy deposition at a specific location and time. Hits are converted into the expected electrical signal produced if a certain threshold is reached and are digitized to match the output format of the collected raw collision data. In the digitization step, additional information about the varying responses of channels, different noise levels, or dead modules is taken into account and the hardware trigger decision is recorded. These conditions can be run-dependent. In addition to the raw data, each event's truth information is saved. After the simulated raw data is passed through reconstruction, discussed in the next chapter, the truth information can be used to check the performance of reconstruction algorithms and for their calibration.

Chapter 5

Physics Object Reconstruction in ATLAS

Reconstruction is the process of turning the electrical signals collected by the detector into physical objects, such as electrons and missing transverse energy. Firstly low-level algorithms are run to define tracks and calorimeter clusters (topo-clusters). These serve as inputs to higher-level algorithms that create physical objects as four-vectors. The higher-level algorithms use the characteristic patterns left by different SM particle species in the detector, as shown in Figure 5.1.

Jets originating from b-quarks, and muons are the pinnacle of this thesis. This chapter discusses their reconstruction from raw data measured by the ATLAS detector.

5.1 Building Blocks - Tracks, Vertices, Topo-clusters

Trajectories of charged particles, called tracks, are reconstructed mainly from hits in the ATLAS Inner Detector (ID) (see Section 3.2.2) and used throughout the analysis. They provide precision momentum measurements and charge information thanks to the bending from the $2\,\mathrm{T}$ axial magnetic field. As all charged particles travel outwards following a helical trajectory, pattern recognition forms the basis of their

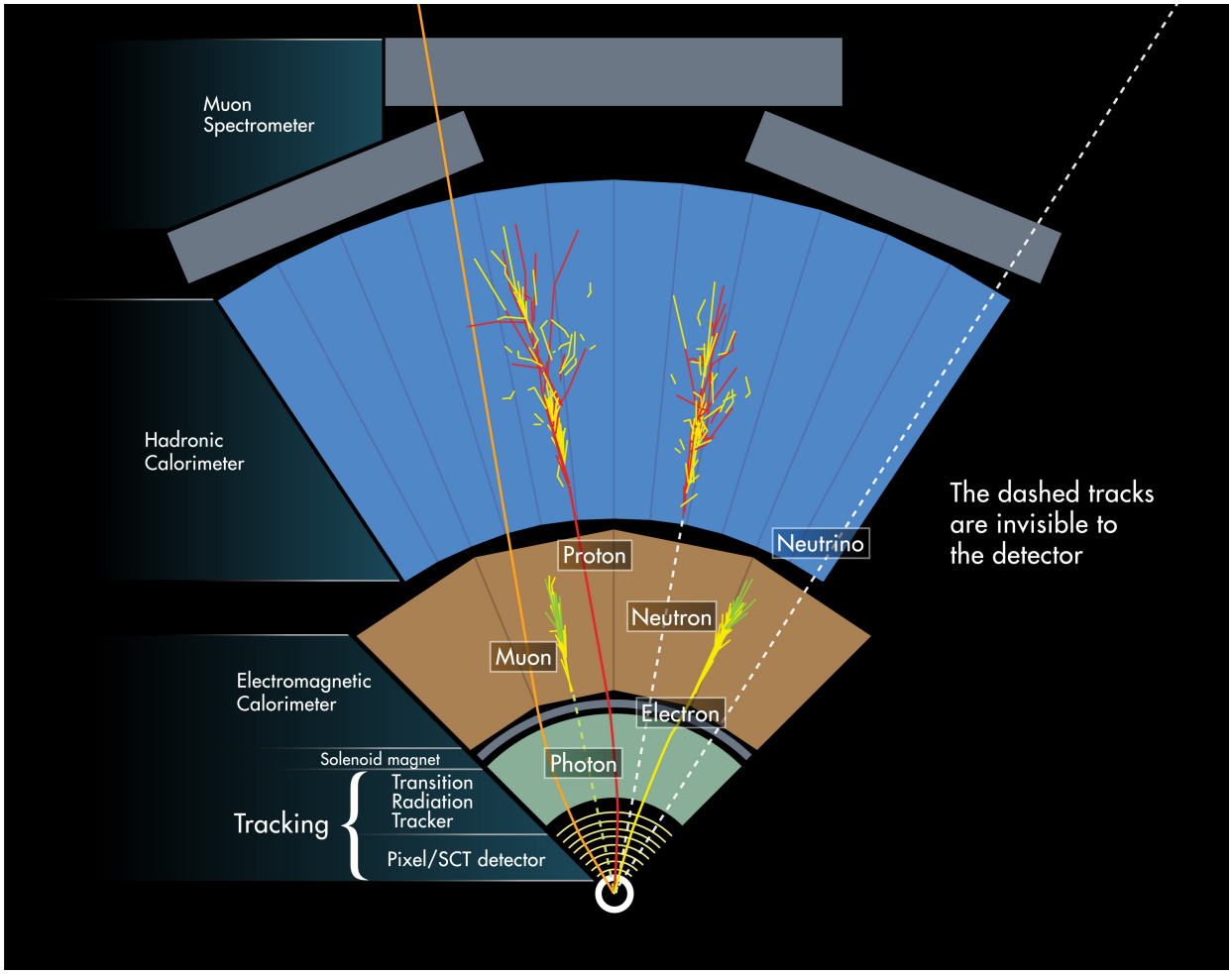


Figure 5.1: The illustration, taken from [103], shows the signatures left by various SM particles in the ATLAS detector.

trajectory reconstruction. Of course, there is an ambiguity in defining a track from many hits. When multiple tracks are associated with the same cluster of hits, this ambiguity is resolved with an artificial neural net as described in [104]. The combinatorics from extra hits from pile-up and eliminating fakes make track finding the most CPU-intensive step in object reconstruction.

Primary vertices (PV) are points in space where protons collide and are reconstructed from tracks with $p_T > 500\,\mathrm{MeV}$. The PV with the highest $\sum p_T^2$ of tracks is identified as the hard scatter vertex [105].

The signal of a single particle left in the calorimeter is gathered into a topo-cluster, a three-dimensional set of topologically connected calorimeter cells [106]. Cells below the noise threshold and far enough from cells with sufficiently high energy are removed and not considered by the clustering algorithm. The calibration of a topo-cluster is done on a cell per-cell level and corrects for the non-compensating calorimeter response, the out-of-cluster lost cells, and dead material interactions. The corrections are derived independently for hadronic (HAD) and electromagnetic (EM) energy deposition from MC simulations of single pions generated at energies from $200\,\mathrm{MeV}$ to $2\,\mathrm{TeV}$ covering various η directions. Charged pions are used to derive the HAD calibrations, while neutral pions are used for the EM calibrations. In the first step of the calibration procedure, each topo-cluster is assigned a probability of having the energy deposited from an EM shower ($\mathcal{P}_{clus}^{EM}$). Once $\mathcal{P}_{clus}^{EM}$ is assigned to the cluster the difference in the calorimeter response between hadrons and electrons/photons is evaluated. The calorimeter response is defined as the ratio between the deposited energy and the detected signal on the EM scale. Hence, no correction is needed for an EM-induced shower and the correction weight is $w_{cell}^{em-cal} = 1$. On the other hand, hadrons with the same energy deposition as electrons generate a smaller signal, and a per cell correction ($w_{cell}^{had-cal}$) is applied to obtain an on-average same response between pions and electrons. Secondly, an algorithm is applied to identify the so-called lost cells, which are cells with energy

deposition from the true signal that failed to be assigned to a topo-cluster due to noise suppression. The out-of-cluster correction (w_{cell}^{ooc}) employs an algorithm that searches the neighbouring area, which size is defined by η[1], for lost cells. Lastly, the energy lost outside the active material is calculated from the energy deposited and the direction of flight of the particles that ignited the topo-clusters. Eight regions of dead material in the $\phi - \eta$ calorimeter plain are defined. The region used to calculate the dead material correction ($w_{cell}^{EM/HAD-like}$) depends on the direction of flight of the particle matched to the topo-cluster. The different dead material regions are mapped to the topo clusters in a similar way as the lost clusters. The size of all the desired corrections is derived separately for EM and HAD showers and combined in a weighted fashion according to $\mathcal{P}_{clus}^{EM}$ as

$$w_{cell}^{calib} = \prod_{m \in \{calib,ooc,dm\}} \mathcal{P}_{clus}^{EM} \cdot w_{cell}^{EM-calib} + (1 - \mathcal{P}_{clus}^{EM}) \cdot w_{cell}^{HAD-calib}, \tag{5.1}$$

where w_{cell}^{calib} is the full calibration of a topo-cluster's cell. The described procedure is summarised from [106].

5.2 Muons

The main identifying criteria for muons is the presence of tracks in the Muon Spectrometer (MS) that can be matched to those in the ID [107]. A muon is then defined by fitting the matched MS and ID components. In this step, the energy loss in the calorimeter is accounted for. Several working points (WPs) are defined to ensure a high quality of the reconstructed muons tuned to the needs of various analyses. Their definition varies in the required track fit properties, and the number of hits in the ID and MS stations as detailed in [107]. The purity and efficiency of reconstructed muons are used as the figure of merit in defining the WPs.

[1] Defining the size of the search area by η ensures that the granularity of the calorimeter is reflected.

Only muons found in jets passing the medium WP are considered in this thesis. In the $HH \to b\bar{b}b\bar{b}$ they are used for the b-jet calibration, while in the *in-situ* $g \to b\bar{b}$ calibration they are used to increase the purity of $b\bar{b}$ jets in the final event selection.

5.3 Jets

Many analyses target final states with quarks, especially b-quarks. As they shower and hadronise they can be identified only indirectly through a spray of highly collimated particles forming a jet. The signatures left in the detector by jets are dense hits in the ID and large energy depositions over several cells in the calorimeter. The reconstruction aims to identify jets in such a way that their four-vectors can be approximated with that of the initiating parton. In general, there are three stages involved in reconstructing a jet; defining its constituent parts, clustering them into jets, and finally grooming and calibrating the jets.

5.3.1 Jet Constituents

Individual particles that make up a jet are represented by tracks, topo-clusters, or their combined information through PFlow objects. The PFlow algorithm [108] is widely employed by ATLAS, due to the energy resolution improvements it brings. It matches ID tracks and topo-clusters to form PFlow objects as follows:

- ID tracks fulfilling a set of quality requirements are p_T ordered.
- The track with the highest p_T is considered first and matched to a topo-cluster forming a topo-track system.
- The probability that the particle initiating the track deposited energy in more than one topo-cluster is calculated, if exceeding a threshold more top-clusters are matched to the track-topo system.

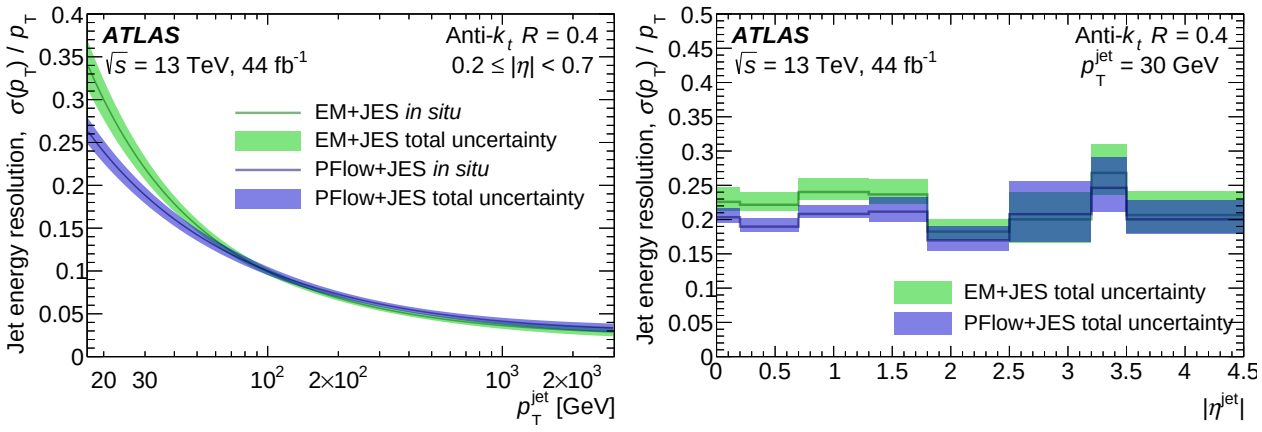
- Once the topo-track system is complete, the track energy is removed from the topo-cluster resulting in a PFlow object.

Once all the tracks and topo-clusters are considered the list of PFlow objects is complete.

Combining the tracking and calorimeter information improves the energy and angular resolution because the ID and calorimeter provide complementary measurements. ID tracks provide precision measurements for low-energy charged particles with $p_T > 500\,\mathrm{MeV}$ and can account for particles whose energy does not exceed the calorimeter cell threshold. The calorimeter has better energy resolution for high-energy particles. Additionally, only the calorimeter can provide measurements of the energy coming from neutral particles and the forward region, where only topo-clusters form PFlow objects. Figure 5.2 shows the improvement in jet energy resolution when jets are clustered from PFlow objects (PFlow jets) instead of topo-clusters (Topo jets) only. The jet energy resolution is shown as a function of p_T in Figure 5.2a and as a function of η in Figure 5.2b. In both cases, the jets reconstructed from PFlow objects have a notably better resolution at low values of p_T and η.

5.3.2 Jet Clustering

The main challenge in constructing jets is to achieve collinear and infrared safety (IR) while maintaining a well-defined jet boundary. IR safety requires the clustering to be insensitive to soft radiation, meaning that adding infinitely many soft particles does not change the number of reconstructed hard jets [109, 110]. Similarly, a clustering algorithm is collinear safe, if the reconstructed jets are independent of the collinear splitting of a parton. The anti-kt algortihm [111] is such a jet clustering method. It is a distance-based sequential algorithm that clusters hard objects first and forms conical jets, both of which ensure infrared and collier safety.



(a) (b)

Figure 5.2: The relative jet energy resolution is compared between PFlow (blue line) and EMTopo (green line) jets. All jets have been clustered with the anti-k_T algorithm setting $R = 0.4$ and are fully calibrated. The energy resolution is plotted as a function of p_T in (a) and η in (b). The figure is taken from [80].

The anti-kt algorithm takes as the figure of merit d_{ij} and d_{iB}, the distance between particle (or pseudo-jet[2]) i and particle (or pseudo-jet) j and the distance between particle i and the beam (B).

The algorithm works as follows:

- d_{ij} is calculated for all possible particle and pseudo-jet combinations.
- The smallest d_{ij} is taken and compared to d_{iB}, where i is the harder particle.
  - Provided d_{ij} is smaller, d_{iB} i and j are clustered forming a pseudo-jet.
  - If d_{iB} is smaller, d_{ij} i is declared a jet and removed from the list of particles.

This is repeated until no particles or pseudo-jets are left to be recombined.

[2]A pseudo-jet is the four-vector of already recombined particles by the algorithm.

The distance parameters are defined as

$$d_{ij} = min(p_{Ti}^{m2}, p_{Tj}^{m2}) \frac{R_{ij}^2}{R^2}, \tag{5.2}$$

$$d_{iB} = p_{Ti}^{m2}, \tag{5.3}$$

$$\tag{5.4}$$

where $R_{ij} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The momentum, rapidity, and azimuth angle of particle i are p_{Ti}, y_i and ϕ_i respectively and R is the radius parameter of the jet. The definition of d_{ij} accounts for geometrical and kinematic properties and m determines their respective impact. The success of the anti-kt algorithm comes from setting $m = -1$, which ensures that the hard particles are clustered first.

The kt algorithm [112] and Cambridge-Aachen [113] have the same clustering procedure and only differ in the m parameter. The kt algorithm sets $m = 1$, which results in firstly clustered soft components. This property is often put to use in pile-up removal procedures. The Cambridge-Aachen algorithm has $m = 0$, which recombined the closest particles first.

Matching Objects to Jets

Particles, which can be tracks, truth hadrons, or other jets, are matched to a jet by either cone or ghost association. Cone association relies on the angular separation between the jet axis and the four-momentum of the particle in question. A particle is matched to a jet, given it is within $\Delta R < 0.3$ of the jet axis. Ghost association relies on defining ghosts, infinitely soft particles with the same η and ϕ as the particle to be matched. The so-defined ghosts are added to the list of particles to be clustered into jets [114]. A match is made, if the ghost is clustered into a jet. The constituents of the latter would remain unchanged as the anti-kt algorithm is IR-safe. Both matching techniques are used in this thesis, which one is used is pointed out when needed.

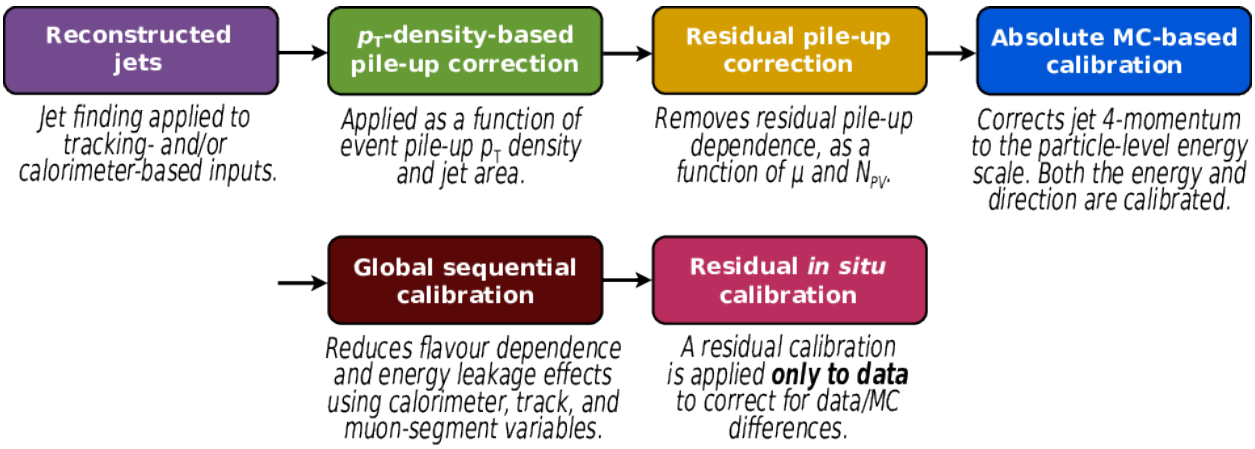


Figure 5.3: The JES calibration chain that modifies the four-momentum of a jet. The Figure is taken from [80].

5.3.3 Jet Calibration

The jet energy scale (JES) calibration applies a chain of corrections to the energy, p_T, mass, and direction of a jet. It aims to adjust the four-vector of the jet to match that of the truth level particle by removing effects of pile-up, and underlying events besides other detector-induced differences between data and MC jets [80]. In addition, the jet energy resolution (JER) correction smears the simulated jets so that the resolution matches that observed in the data. This section outlines the main steps taken in calibrating small-R jets ($R = 0.4$) and is heavily based on [80], which is the general reference of this section.

The JES calibration can be split into two parts, the simulation-based jet calibration and the *in situ* calibration which removes any remaining differences between data and MC. The first is applied to both data and MC, while the latter only corrects the jets reconstructed from the collected data. The exact sequence of steps taken by the JES calibration is shown in Figure 5.3.

In the simulation-based JES calibration first, the contribution to the jet p_T from pile-up is removed as a function of the jet p_T-density and jet area, while any residual dependence of the jet p_T on the pile-up activity is removed as a function of the

number of primary vertices (N_{PV}) and the mean number of interactions per bunch (μ). Next, the absolute jet energy scale and direction (η) are calibrated in bins of truth jet p_T and detected η from truth jets which have been cone matched to reconstructed jets in dijet simulations. Lastly, the global sequential calibration corrects for the jet flavour-dependent differences in the detector response. The four-momentum correction is defined in six flavour-dependant observables independently and applied sequentially. The order of the applied corrections is found to have no impact on the final improvement of the jet resolution, which is used as the figure of merit. Similarly, as in the previous step, the corrections are derived from reconstructed jets matched to truth jets from the dijet simulation and mainly take care of the difference between gluon and quark initiated jets.

Any remaining discrepancy between data and MC is accounted for by the *in-situ calibration* applied to data. The ratio between the jet response in simulations and data is used to derive the jet response correction factor in three independent measurements. The jet response correction factor is computed from balancing a jet against a well-calibrated physics object in the η intercalibration, Z/γ+ jet, and multijet balance analyses. Each analysis is performed independently and the corrections are applied sequentially but the sequence itself does not impact the final result. The η intercalibration takes advantage of the dijet balance and corrects the jet energy scale of forward jets to match that of central jets. Next, the hadronic response of the jet is corrected by using the Z or photon as a probe, as the electromagnetic scale is well-calibrated already. The p_T of the probed Z or photon is balanced against the recoiling jet in the Z/γ+ jet analysis. Lastly, the multijet calibration extracts the corrections for the calibration of high p_T jets. A high-p_T jet is balanced against recoiling low-p_T jets, where the Z/γ+ jet correction is applied to the recoiling jets. The main sources of uncertainties are the mis-modelling from simulations, event topology, and measurement error of the reference objects, which all are propagated through the analysis in the same sequence as the applied corrections [80].

After the full set of JES calibration is applied the jet energy resolution (JER) in simulations is smeared to match that observed in the data. The resolution depends on the jet p_T and is derived from the expected calorimeter resolution in Equation 3.2.3. It is determined from asymmetries in the dijet balance measurements, as dijet events are abundant as well as the balance requirement allows for precise p_T measurements. The noise term, which is used to put constraints on the dijet measurement, is estimated from MC and data collected by unbiased triggers by randomly forming cones corresponding with the jet area but not overlapping with the reconstructed jets.

The simulated jets are smeared until on average the data and MC resolutions match and only if the data resolution is larger. A JER component of the systematic uncertainty is then derived by sampling from a Gaussian with width σ_{smear}

$$\sigma^2_{smear} = (\sigma_{data} + \Delta\sigma_i)^2 - \sigma^2_{data}, \tag{5.5}$$

where σ_{data} is the resolution in data the nominal MC is corrected to and $\Delta\sigma_i$ is the one-standard deviation on the uncertainty component i. If the data resolution is smaller than the one measured in MC no smearing is applied and only a systematic uncertainty is defined as

$$\sigma = \sigma_{data} - \sigma_{MC}, \tag{5.6}$$

where σ_{data} and σ_{MC} are the jet energy resolutions of data and simulations respectively. Each component of the systematic uncertainty is evaluated by propagating the jets with the alternative JER value throughout the whole analysis.

The relative jet energy resolution of PFlow and Topo jets clustered with the anti-k_T algorithm setting $R = 0.4$ after the JER and JES calibrations are applied is shown as a function of p_T in Figure 5.2a and as a function of η in Figure 5.2b.

The JES and JER calibrations set the stage for the definition of the detector,

statistical, and modelling systematics uncertainties associated with jets. As each calibration step in Figure 5.3 comes with its own set of uncertainties, they are often evaluated as a set of nuisance parameters (NPs) within the calibration itself and propagated through to the next stage. The result is more than 100 NPs that cannot be simply added up into a single JES uncertainty, due to the underlying correlations. However, the majority of the NPs are from the in-situ calibration step and only have a p_T dependence. Thus, thanks to eigenvalue decomposition a highly reduced set of orthogonal NPs can be constructed. Depending on the analysis need, the number of NPs can be reduced to around 30 or 60. Similarly, JER systematics are derived from each resolution measurement stage and propagated through the whole analysis as a Gaussian smearing function on the MC [80].

b-jet Calibration

JES and JER calibrations are derived to correct light and gluon jets, which lead to the energy of b-jets being systematically underestimated. The sources of the underestimations are the semi-leptonic decays of b-hadrons and differences in the nature of the parton shower.

The muon-in-jet strategy is employed to correct for the energy missing from the semi-leptonic decays, $W \to \mu\nu_\mu$, due to the muon punching through the calorimeter. The missing muon energy is accounted for by adding the four-vector of a muon found within the jet and subtracting the muon's calorimeter energy deposition. The muon is said to be part of the jet, if $\Delta R(jet, \mu) < min(0.4, 0.04 + \frac{10\,\mathrm{GeV}}{p_T^\mu})$ and in the case of multiple muons in the jet, only the one closest to the jet axis is considered. The missing neutrino energy as well as the out-of-cone effects due to the wider parton shower of b-jets are captured by the p_T-reco correction. A separate correction factor is derived from $t\bar{t}$ events that have semi-leptonically decaying B hadrons, i.e. contain a muon, and non-semi-leptonically decaying B hadrons in $log(p_T)$ bins. The calibration strategy has been first described in the search for $H \to b\bar{b}$ decays using

partial Run 2 data in [115].

5.3.4 Jet Types

Jet collections are defined by their constituents, the algorithm used and the value of R. For example, jets clustered from PFlow objects with $R = 0.4$ are Small-R PFlow jets. Three different jet collections are used in this thesis, fixed radius PFlow jets with $R = 0.4$, large radius calorimeter (large-R) jets with $R = 1.0$, and variable radius track (VR) jets.

Fixed Radius Jets, R=0.4

Small R-jets have $R = 0.4$ and are clustered with the anti-kt algorithm. They are widely used in ATLAS analyses that search for final states with well-separated jets. The non-resonant $HH \to b\bar{b}b\bar{b}$ analysis uses PFlow jets, while topo-cluster jets are used for identifying the trigger jets in the $g \to b\bar{b}$ Calibration. In addition to applying the JES and JER calibration, jets with $p_T < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ have to pass the tight Jet Vertex Tagger cut defined in [116] to limit the pile-up contamination. Bad jets are identified in a process called jet cleaning, which requires all jets in an event to pass a set of quality cuts [117]. If any jet fails those quality criteria it is labelled as a bad jet and the whole event is vetoed. This ensures that the jet kinematics are not distorted by jets coming from non-collision background processes, such as cosmic rays.

Large Radius Calorimeter Jets, R=1.0

The angular separation between decay products is approximately $\frac{2m}{p_T}$, which means that in the boosted regime[3] jets containing the decay products are not well resolved.

[3] When a particle with a large transverse momentum decays (significantly larger than its rest mass), its decay products are highly collimated in the direction of the Lorentz boost.

It becomes more feasible to reconstruct decay products from a heavy resonance as a single jet. The usual procedure is to define large radius (large-R) jets with $R = 1.0$ clustered from topo-clusters. A jet trimming algorithm [118] is then applied to remove the effects of pile-up, initial state radiation, and underlying events. It uses the kt algorithm to recombine the constituents of the large-R jet into small-R jets with $R = 0.2$. If the small-R jet's p_T is smaller than 5% of the large-R jet p_T it is removed from it. After trimming only the JES calibration is applied, as JER uncertainties are sufficient to cover the data-MC discrepancy in resolution [119, 120].

Variable Radius Jets

Adjusting the R of highly boosted jets in the clustering step improves the performance, as jets with higher p_T tend to be more collimated. Therefore, variable radius (VR) jets are constructed by letting R vary as $R = \frac{\rho}{p_T}$, where ρ is a constant without dimensions that determines how fast the jet radius alters [121]. A lower and upper bound on R is set to prevent too large jets or jets smaller than the detector resolution to be clustered. The optimal parameters for analyses targeting boosted $H \to b\bar{b}$ have been found to be $\rho = 30$, $R_{max} = 0.4$ and $R_{min} = 0.02$ [122]. The optimisation is based on truth double b-tagging efficiency as VR jets are often used to capture the sub-structure of large-R jets and used in b-tagging. The convention adapted by analyses that target a boosted heavy resonance decaying as $X \to b\bar{b}$ is to reconstruct a large-R topo jet to which VR track jets are ghost associated. The large-R jets captures the kinematics of the whole decay, while the VR jets encode the substructure needed to extract flavour information.

5.4 Identification of b-quark Jets (*b-tagging*)

The identification of jets initiated by b-quarks (b-jets) is done with b-tagging algorithms, simply referred to as b-taggers from now on. The taggers exploit the distinct properties of the B hadrons; large mass, high decay multiplicity, long lifetime, and subsequently a secondary vertex displaced from the hard scatter. The algorithms were optimised to distinguish between three classes of jets, b-jets from c-jets and l-jets, which in MC are defined as follows:

- ΔR is defined as the angular separation between the four-momentum of a hadron and the jet-axis of the jet considered.
- The jet is a b-jet if a B-hadron that satisfies $\Delta R < 0.3$ is found.
- If no B-hadron is found, next all C-hadrons are considered.
- The jet is a c-jet if there is a C-hadron within $\Delta R < 0.3$.
- The jet is tau jet if there is a tau within $\Delta R < 0.3$.
- The jet is an l-jet[4], if neither B- nor C- hadrons are found within $\Delta R < 0.3$ of the jet.

The presence of more than one B- or C- hadrons within $\Delta R < 0.3$ can be saved as additional information. For example, if two B-hadrons are found within $\Delta R < 0.3$ the jet is initially classified as a b-jet and the number of B-hadrons is checked only if required.

To evaluate the performance of the b-tagging algorithm it is useful to define the b-tagging efficiency (ϵ_b) as the fraction of tagged b-jets, and the mis-tag rate is the fraction of the tagged jets that are not true b-jets (background). The latter is commonly defined for c- jets (ϵ_c) and l-jets separately (ϵ_l). It is conventional to plot

[4]The l-jets, where l stands for light jets are initiated by a u-, d-, s-quark or a gluon.

the background rejection ($\frac{1}{\epsilon_{c/l}}$) as a function of ϵ_b when assessing the performance of the tagger.

During Run 2 ATLAS adopted a two-stage approach to b-tagging; low-level and high-level algorithms with several optimisations as compared to Run 1 [123]. Low-level algorithms associate tracks to jets of interest by requiring the tracks to be within ΔR of the jet axis, where ΔR is p_T-dependent[5].

The transverse impact parameter (d_0) of tracks is exploited by the IP2D tagger, while the IP3D tagger uses the transverse and longitudinal impact parameter (z_0), where d_0 (z_0) is the distance between the PV and the point of closest approach of the track in the $r-\phi$ (longitudinal) plane [123]. Both taggers construct flavour-dependent probability density functions from impact parameter-based discrimination variables that are combined into log-likelihood ratios used for tagging [124]. The recurrent neural network called RNNIP [125] is also based on the measurement of the impact parameter. The main difference between the baseline IP taggers and the RNNIP is that the latter learns the dependence between tracks and can exploit correlations inherent in the decays of a b-hadron [125]. For tagging purposes, it combines the probabilities of a jet being a b-, c- or l-jet according to Equation 5.7.

Vertices originating from b-hadron and c-hadron decays, called Secondary Vertices (SVs), are crucial for the flavour identification of jets. The Single Secondary Vertex Finder (SVF) [126] is applied on tracks associated with reconstructed jets to find an SV. High purity of SVs originating from b-hadron and c-hadron decays is achieved by applying a cleaning that removes track pairs from light long-lived particles, photon conversion and material interactions[6]. Additionally, only SVs with an invariant mass below 6 GeV are kept, as higher masses are not expected given the B hadron invariant mass range. The properties from the SV are exploited by the

[5] ΔR gets narrower with higher track p_T, i.e. $\Delta R = 0.45$ for $20\,\mathrm{GeV}$ and $\Delta R = 0.26$ for $150\,\mathrm{GeV}$ [123].

[6] Light long-lived particles are vetted based on the reconstructed invariant mass of track pairs. Fake vertices originating from material interactions are removed by comparing the radial position of the track pairs in comparison to the location of the pixel detector and beam pipe.

secondary-vertex-tagging algorithm, SV1. Discriminating variables, such as SV invariant mass and SV track multiplicity, are constructed and used as inputs to higher-level algorithms. The algorithm is optimised for low p_T environments and a higher number of fake vertices is expected in dense track environments, as is the case in the $g \to b\bar{b}$ calibration [124]. More details on the SV1 algorithm and performance can be found in [126]. .

JetFitter [127] reconstructs the full decay chain of the b-hadron through the topology of b- and c-hadron decays. Similarly to SV1 it reconstructs discriminating variables that serve as inputs to higher-level algorithms and its performance is checked by constructing a simple neural net tagger.

The outputs from the low-level taggers in addition to the jet p_T and η serve as inputs to high-level tagging algorithms used in analyses. A deep feed-forward neural network called DL1r is an example of a high-level tagger and has been the default tagger for Run 2 analyses, including the $HH \to b\bar{b}b\bar{b}$ analysis [124, 128]. It outputs the probabilities of a jet to be a b-, c- or l-jet (p_b, p_c, and p_l). The probabilities are combined into a discriminant based on the log-likelihood as

$$D_{DL1r} = ln\frac{p_b}{f_c p_c - (1 - f_c)p_{light}}, \tag{5.7}$$

where f_c is the expected c-jet fraction. A jet is said to be b-tagged, if its D_{DL1r} score exceeds a pre-defined threshold. Several thresholds, called working points (WPs), are defined as benchmarks for the b-tagging efficiency (ϵ_b) measured in $t\bar{t}$ samples [129, 123].

The performance of low-level taggers is compared to the high-level DL1r tagger in Figure 5.4 by showing the dependence of the background rejection on the b-tagging efficiency. From the Figure, it is clear that for the same number of b-tagged jets, DL1r has a significantly higher c-jet and l-jet rejection than any of the low-level algorithms it is compared to.

A calibration is needed to correct for differences in b-tagging efficiency between MC and data. A flavour dependent scale factor (SF), $SF = \frac{\epsilon_b^{Data}}{\epsilon_b^{MC}}$, is applied to MC. Separate calibrations for the b-, c- and l-jets are presented in [129, 130, 131]. The b-jet calibration uses semi-leptonically decaying $t\bar{t}$ events, where one of the top quarks decays as $t \to Wb$. The c-jet mis-tag calibration uses semi-leptonically decaying $t\bar{t}$ events as well only that it targets the $W \to cs$ decays that have a large enough and known branching fraction. The l-jet mis-tag calibration targets Z bosons produced in association with jets and uses a modified b-tagging algorithm to estimate the mis-tag rate due to track resolution affects, whilst minimising the impact from b and c-jets.

As boosted regimes become accessible with the increase in statistics there is a demand for double-b taggers. As the name suggests they are used for identifying heavy resonance decaying into a $b\bar{b}$ pair. The first double-b tagger used in ATLAS and its calibration is described in Chapter 8.

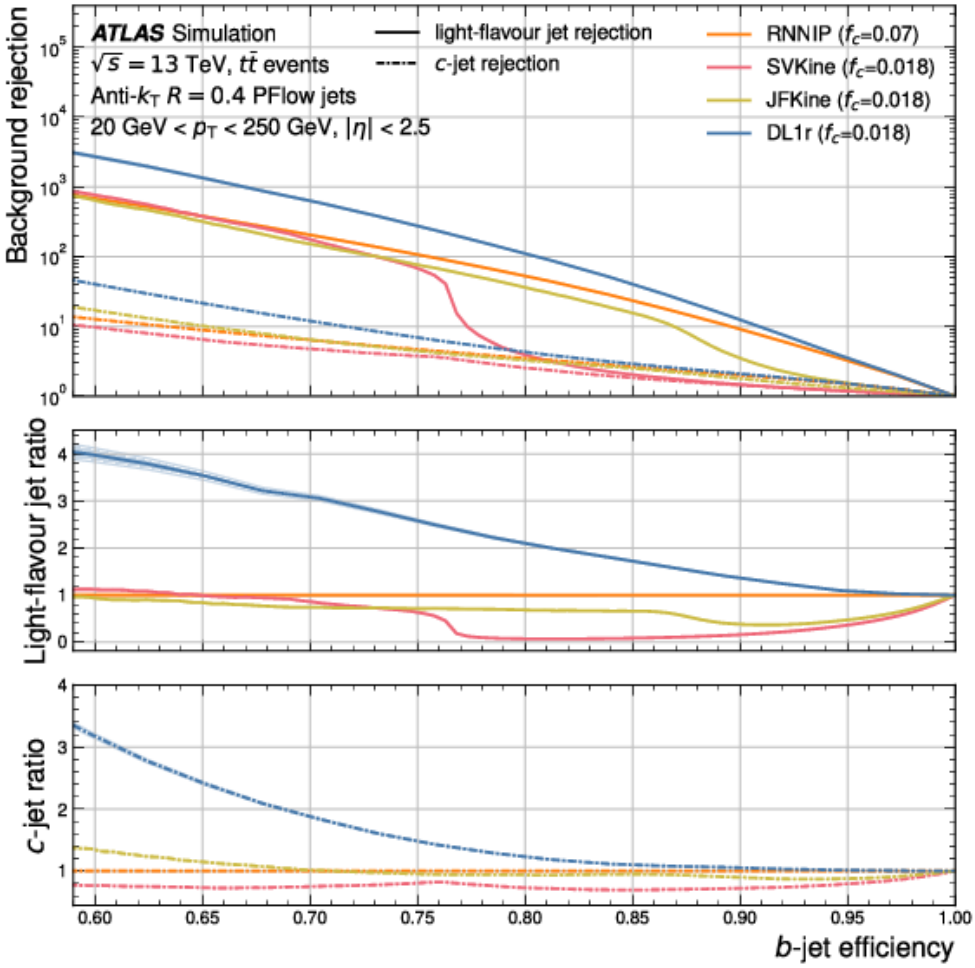


Figure 5.4: The background rejection is defined as the inverse of the mis-tag rate, evaluated independently for c- and l-jets. The upper panel shows the c-jet (solid line) and l-jet (dashed line) rejection as a function of ϵ_b. The lower two panels show the ratio of the c- and l-jet rejection as compared to the RNNIP tagger. JVFKine is the tagger constructed from JetFitter variables and SVKine is the tagger constructed from the SV variable. The figure is taken from [124].

Chapter 6

Non-resonant $HH \to b\bar{b}b\bar{b}$ Analysis

6.1 Overview

The diHiggs production at $\sqrt{s} = 13\,\mathrm{TeV}$ in proton-proton collisions has a relatively low cross-section of $\sigma_{HH} \sim 32.7\,\mathrm{fb}$ in the SM and is not expected to be observed in Run 2 searches at the LHC [32, 34]. Nevertheless, in the collected dataset BSM scenarios are in reach and the trilinear Higgs self-couplings can be constrained. The ATLAS and CMS experiments have performed resonant and non-resonant searches for diHiggs in several decay channels. The three most sensitive decay channels considered by the ATLAS experiment are $b\bar{b}b\bar{b}$ [2], $b\bar{b}\tau^+\tau^-$ [132, 133], and $b\bar{b}\gamma\gamma$ [134, 135]. In addition, analyses also have tackled the following final states, $b\bar{b}WW^*$ [136], WW^*WW^* [137], $WW^*\gamma\gamma$ [138] and $b\bar{b}l\nu l\nu$ [139]. In the last one the Higgs boson decays into $H \to b\bar{b}$ while the other as $H \to WW^*/ZZ^*/\tau^+\tau^-$.

This chapter focuses on the $b\bar{b}b\bar{b}$ decay channel, which has a branching fraction (BR) of 34% as can be read in Figure 6.1. However, it is not the most sensitive channel as the large QCD background has to be modelled with a data-driven technique, which has a large uncertainty, and the inefficiencies in triggering and b-tagging lead to a reduced signal acceptance, to name a few of the challenges. First, the dataset and

Monte Carlo simulations used are described in Section 6.2. The number of signal sample simulations is significantly reduced by applying a reweighting technique, as outlined in Section 6.3. The event selection and data-driven background estimation are described in Section 6.4 and 6.5. Section 6.6 discusses the systematic uncertainties applied to the background estimation and the signal samples. The statistical analysis used to interpret the results is outlined in Section 6.7, while the results, including the κ-framweork and EFT interpretations, are presented in Section 6.8. The conclusions that can be drawn from the search are presented in Section 6.9. The work presented is based on the non-resonant $HH \to b\bar{b}b\bar{b}$ paper [2].

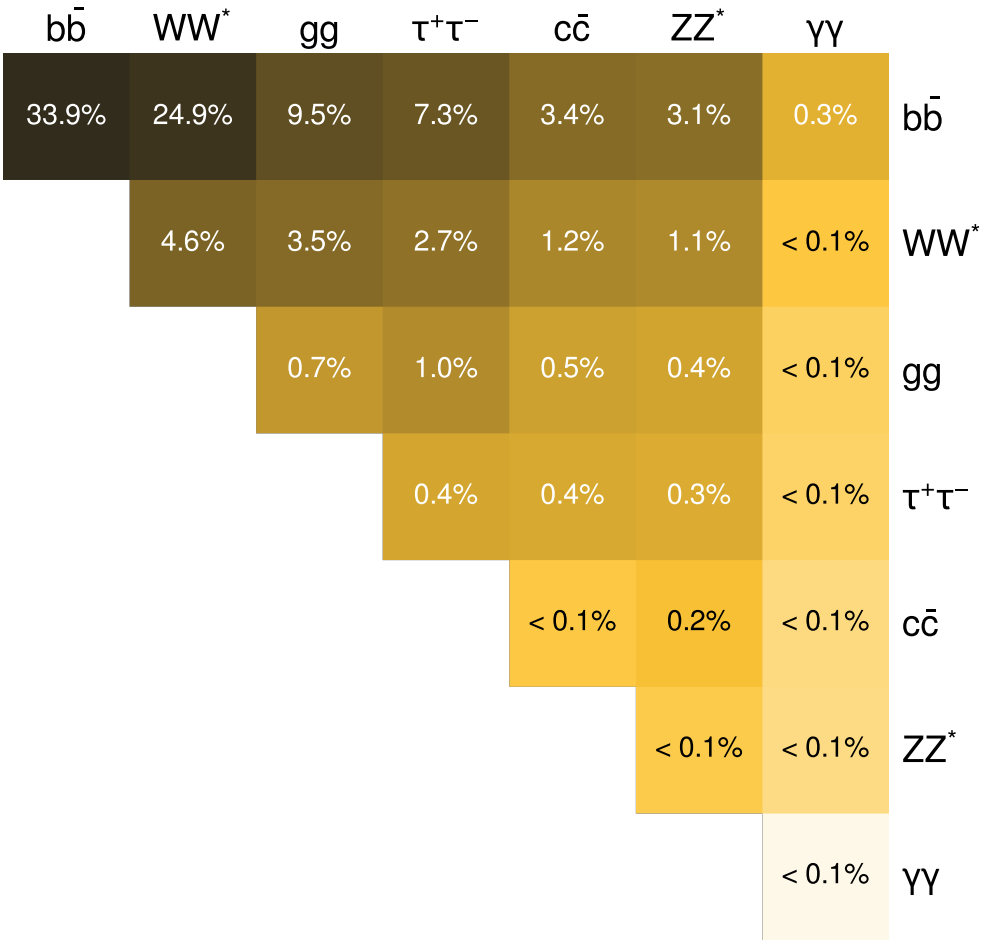


Figure 6.1: The expected diHiggs branching fractions as reported by [140]. The $b\bar{b}b\bar{b}$ channel has by far the largest branching fraction of 33.9%.

Acknowledgements and Author's Contributions

The non-resonant $HH \to b\bar{b}b\bar{b}$ analysis has been possible thanks to the collaboration of scientists from fourteen institutes from all over the world. To ensure a complete overview of the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis the thesis includes contributions from other analysers as well. Their contributions are documented in the internal documentation [141], which is the general reference for this chapter in addition to the published paper [2].

The author of this thesis made significant contributions to the software framework of the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis. Their work involved implementing scale factors for the Monte Carlo (MC) signals, specifically focusing on the total b-tagging scale factors. Two different treatments of these scale factors were explored, allowing for the use of two different b-tagging requirements in the final signal regions of the analysis. However, the latter treatment was ultimately discarded and is not discussed in the thesis. Additionally, the author played a major role in implementing and validating the statistical analysis, especially in Section 6.7. They were particularly involved in establishing the $-2\ln\Delta L$ test outlined in Sections 6.7.3 and 6.7.4. The author's contributions also extended to establishing and validating the correlation scheme of systematics, as outlined in Section 6.6. A significant contribution of the author was to define the uncertainty related to the acceptance of the effective field theory signal and the associated reweighting uncertainty, described in Section 6.6.5. The author was instrumental in studying the unblinded results presented in Section 6.8 and contributed significantly to verifying the fit result on the unblinded data which is critical for this type of analysis. The likelihood scans depicted in Figures 6.22 and 6.23, which were later published in [2], are the author's work. The author advocated for and developed a broadened interpretation able to demonstrate the full power of the analysis, by introducing the 2D $\kappa_\lambda - \kappa_{2V}$ scans shown in Figure 6.23, which extends naturally to EFT models. Furthermore, the author played a vital role in optimising the statistical analysis of the effective

field theory, significantly reducing the result turnaround time. This improvement facilitated the exploration of wider parameter space in both the HEFT and SMEFT interpretations, as evidenced by Figures 6.24, 6.25, 6.26, and 6.27. In addition to their major contributions, the author conducted a study to assess the impact of the VBF channel when modifying the cut definition to reduce $t\bar{t}$, which is mentioned in Section 6.4.4. Lastly, the author investigated an alternative background estimation procedure not presented in the thesis.

6.2 Data and Simulations

The non-resonant $HH \to b\bar{b}b\bar{b}$ analysis uses data from proton-proton collisions collected by the ATLAS Experiment at $\sqrt{s} = 13\,\mathrm{TeV}$ in the years 2016, 2017, and 2018 during Run 2 of the LHC operation. The data collected during 2015 has an issue associated with the jet triggers, which is fixable but because the b-jet triggers used are inefficient and the dataset represents only 2.5% of the total integrated luminosity collected during Run 2 it is discarded. The integrated luminosity collected per year of data taking is $24.6\,\mathrm{fb}^{-1}$ in 2016[1], $43.7\,\mathrm{fb}^{-1}$ in 2017, and $57.7\,\mathrm{fb}^{-1}$ in 2018, which results in the total integrated luminosity of $126\,\mathrm{fb}^{-1}$. The dataset described is used to derive the background estimation of the predominantly QCD multi-jet background.

Monte Carlo simulations are used only to model the various signal processes considered, from κ variations of the HH signal to EFT modifications. The software used to compute the matrix element and sample events from it varies between the processes considered. However, the event generator is always combined with Pythia 8.244 [100], which uses the NNPDF2.3lo PDF set [142] and the A14 set for the tuned parameters [143], to obtain the nominal samples. The uncertainty on the parton shower and hadronisation is evaluated by using Herwig 7.1.6 instead of Pythia 8.244

[1] $8.3\,\mathrm{fb}^{-1}$ of data is not used as there was an inefficiency in the PV reconstruction.

[144]. It uses a different hadronisation model, PDF set (MMHT2014lo PDF set [145]) and tune (default set of tuned parameters from Herwig [146]). The Higgs mass is assumed to be $125\,\mathrm{GeV}$ for all produced simulations. Lastly, the interactions with the ATLAS detector are simulated with Geant4. A sample is said to be fully simulated if, after the event generation, each event is parton showered, hadronised, and propagated through the simulation of the ATLAS detector, as detailed in Section 4.2.

MadGraph 2.7.3 [90] at LO with the NNPDF3.0nlo PDF set [147] interfaced with Pythia 8.244 is used to generate the VBF signals, all $\boldsymbol{\kappa}$ variations that are fully simulated and used in the analysis are shown in Table 6.1. A reweighting technique, described in detail in the next section, is used to produce a finer grid of $\boldsymbol{\kappa}$ variations. The SM sample is normalised to $\sigma_{VBF} \times BR(4b) = 1.73\,\mathrm{fb} \times 0.3392 = 0.59\,\mathrm{fb}$, where σ_{VBF} is the N^{3}LO cross-section [34], see Section 2.1 for details, and $BR(4b)$ is the branching fraction of diHiggs to two pairs of b-quarks.

Process	κ_λ	κ_{2V}	κ_V
ggF	1	N/A	N/A
ggF	10	N/A	N/A
VBF	1	1	1
VBF	1	1.5	1
VBF	2	1	1
VBF	10	1	1
VBF	1	1	0.5
VBF	-5	1	0.5
VBF	0	1	1
VBF	1	0	1
VBF	1	3	1

Table 6.1: The fully simulated signal samples used on the reconstruction level. The generated samples have been parton showered, hadronised, and passed through the ATLAS detector simulation. The ggF signal is not sensitive to κ_V and κ_{2V}.

The ggF events for various κ_λ values are simulated with the Powheg Box v2 generator [148, 149, 150] at NLO including the full NLO correction assuming finite top mass, and using the PDF4LHC15 PDF set [151]. At the event generator level the samples for $\kappa_\lambda = 0.0$, $\kappa_\lambda = 1.0$, $\kappa_\lambda = 10.0$ and $\kappa_\lambda = 20.0$ HH signals are produced. Only the $\kappa_\lambda = 1.0$ and $\kappa_\lambda = 10.0$ samples are fully simulated. All other κ_λ variations are

derived at the reconstruction level from reweighting the SM sample. The coefficients needed for the reweighting are derived from the truth m_{HH} distributions, as described in Section 6.3. As discussed in Section 2.1, the SM cross-section is $\sigma_{ggF} = 31.05\,\text{fb}$, while the branching ratio is $BR(4b) = 0.3392$ resulting in a normalisation $\sigma_{ggF} \times BR(4B) = 10.53\,\text{fb}$ to be applied to the SM sample.

SMEFT and HEFT variations are not fully simulated, as information from the truth m_{HH} is sufficient to reweight the SM sample to any EFT signal, see next section for details. MadGraph5_aMC@NLO [152, 153, 154] with the SMEFT@NLO model [155] is used for the SMEFT event generator level variations produced, while HEFT variations are simulated with Powheg Box v2.

6.3 Sample Reweighting

It is not feasible to generate a signal sample for all possible $\boldsymbol{\kappa}$ values nor for all EFT Wilson coefficients. Only considering κ_λ, the true value is expected to be in the range $\kappa_\lambda \in [-6.0, 10.0]$ at a 95% confidence level according to the white paper considering early Run-2 searches [33]. It is possible to perform a weighted linear combination, provided a few carefully selected Monte Carlo samples are available to obtain a signal for any coupling considered. The reweighting technique is outlined and validated in [156] and is summarised in the following section.

Starting with the ggF production it is clear from Figure 2.1 that the amplitude of the process is

$$A(\kappa_\lambda, \kappa_t) = \kappa_\lambda \kappa_t A_1 + \kappa_t^2 A_2, \tag{6.1}$$

where A_1 and A_2 are the amplitudes of the box and triangle diagrams in Figure 2.1,

respectively. Therefore, the approximate differential cross-section at LO is

$$\begin{aligned} \frac{d\sigma(\kappa_\lambda, \kappa_t)}{dm_{HH}} \approx |A(\kappa_\lambda, \kappa_t)|^2 = \\ \kappa_t^4[\frac{\kappa_\lambda^2}{\kappa_t^2}|A_1|^2 + \frac{\kappa_\lambda}{\kappa_t}|A_1 A_2 + A_2 A_1| + |A_2|^2], \end{aligned} \tag{6.2}$$

which is a 2nd order polynomial in $\frac{\kappa_\lambda}{\kappa_t}$ scaled by k_t^4. Ignoring the normalisation factor the differential cross-section can be treated as a system of linear equations with three unknowns, $|A_1|^2$, $|A_2|^2$ and $|A_1 A_2 + A_2 A_1|$. To solve it $\frac{d\sigma(\kappa_\lambda,\kappa_t)}{dm_{HH}}$ is simulated for the $(\kappa_\lambda = 1, \kappa_t = 1)$, $(\kappa_\lambda = 0, \kappa_t = 1)$ and $(\kappa_\lambda = 20, \kappa_t = 1)$ variations, which form the orthogonal ggF production basis samples. It follows that the solution is

$$\begin{aligned} \frac{d\sigma(\kappa_\lambda, \kappa_t)}{dm_{HH}} \approx & \left(\kappa_t^2 + \frac{1}{20}\kappa_\lambda^2 - \frac{399}{380}\kappa_\lambda\kappa_t\right)\left|\frac{d\sigma}{dm_{HH}}(0,1)\right|^2 \\ & + \left(\frac{40}{38}\kappa_\lambda\kappa_t - \frac{2}{38}\kappa_\lambda^2\right)\left|\frac{d\sigma}{dm_{HH}}(1,1)\right|^2 \\ & + \left(\frac{\kappa_\lambda^2 - \kappa_\lambda\kappa_t}{380}\right)\left|\frac{d\sigma}{dm_{HH}}(20,1)\right|^2. \end{aligned} \tag{6.3}$$

The equation dictates how to reweight $\frac{d\sigma}{dm_{HH}}(0,1)$, $\frac{d\sigma}{dm_{HH}}(1,1)$ and $\frac{d\sigma}{dm_{HH}}(20,1)$ to get $\frac{d\sigma}{dm_{HH}}(\kappa_\lambda, \kappa_t)$, where $(\kappa_\lambda, \kappa_t)$ can be any variation. The reweighting can be done at the truth or reconstruction level.

At the truth level, the ggF production basis samples are binned in truth m_{HH}. The histograms are then combined to produce the truth m_{HH} distributions for κ_λ variations between -30 and 30 in steps of 0.2 according to Equation 6.3. Each distribution is additionally divided bin per bin by the SM distribution and the yields are saved as event-by-event weights to be applied to the fully simulated SM sample after reconstruction. This approach saves computation power, as only the SM sample has to be fully simulated. The reweighting is done at LO but the cross-section is known at NNLO, therefore only the shape of the m_{HH} is obtained from the reweighting, while the cross-section is corrected to

$\sigma(\kappa_\lambda) = 70.3874 - 50.4111 \times \kappa_\lambda + 11.0595 \times \kappa_\lambda^2$ [157]. Similarly, the amplitude of the VBF productions at LO can be read from Figure 2.4 as

$$A(\kappa_{2V}, \kappa_\lambda, \kappa_V) = \kappa_{2V} A_1 + \kappa_V^2 A_2 + \kappa_V \kappa_\lambda A_3, \tag{6.4}$$

where A_1, A_2, and A_3 are the amplitudes of the diagrams in Figure 2.4b, 2.4c, 2.4a, respectively.

It follows that the differential cross-section is

$$\begin{aligned} \frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}} \approx |A(\kappa_{2V}, \kappa_\lambda, \kappa_{2V})|^2 = \\ \kappa_{2V}^2 |A_1|^2 + \kappa_V^4 |A_2|^2 + \kappa_V^2 \kappa_\lambda^2 |A_3|^2 \\ + \kappa_{2V}\kappa_V^2 |A_1 A_2 + A_2 A_1| + \kappa_{2V}\kappa_\lambda\kappa_V |A_1 A_3 + A_3 A_1| + \kappa_\lambda \kappa_V^3 |A_2 A_3 + A_3 A_3| , \end{aligned} \tag{6.5}$$

which again is a linear equation with six unknowns and can be solved given $\frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}}$ is known for six variations of $(\kappa_{2V}, \kappa_\lambda, \kappa_V)$ that form orthogonal vectors. Choosing $(\kappa_{2V} = 1, \kappa_\lambda = 1, \kappa_V = 1)$, $(\kappa_{2V} = \frac{3}{2}, \kappa_\lambda = 1, \kappa_V = 1)$, $(\kappa_{2V} = 1, \kappa_\lambda = -5, \kappa_V =, \frac{1}{2})$, $(\kappa_{2V} = 1, \kappa_\lambda = 2, \kappa_V = 1)$, $(\kappa_{2V} = 1, \kappa_\lambda = 10, \kappa_V = 1)$ and $(\kappa_{2V} = 1, \kappa_\lambda = 1, \kappa_V = \frac{1}{2})$ as the VBF basis sample set results in

$$\begin{aligned} \frac{d\sigma(\kappa_{2V}, \kappa_\lambda, \kappa_V)}{dm_{HH}} \approx \\ \left(\frac{68\kappa_{2V}^2}{135} - 4\kappa_{2V}\kappa_V^2 + \frac{20\kappa_{2V}\kappa_V\kappa_\lambda}{27} + \frac{772\kappa_V^4}{135} - \frac{56\kappa_V^3\kappa_\lambda}{27} + \frac{\kappa_V^2\kappa_\lambda^2}{9}\right) \times \left|\frac{d\sigma}{dm_{HH}}(1,1,1)\right|^2 \\ + \left(-\frac{4\kappa_{2V}^2}{5} + 4\kappa_{2V}\kappa_V^2 - \frac{16\kappa_V^4}{5}\right) \times \left|\frac{d\sigma}{dm_{HH}}\left(\frac{3}{2},1,1\right)\right|^2 \\ + \left(\frac{11\kappa_{2V}^2}{60} + \frac{\kappa_{2V}\kappa_V^2}{3} - \frac{19\kappa_{2V}\kappa_V\kappa_\lambda}{24} - \frac{53\kappa_V^4}{30} + \frac{13\kappa_V^3\kappa_\lambda}{6} - \frac{\kappa_V^2\kappa_\lambda^2}{8}\right) \times \left|\frac{d\sigma}{dm_{HH}}(1,2,1)\right|^2 \\ + \left(-\frac{11\kappa_{2V}^2}{540} + \frac{11\kappa_{2V}\kappa_V\kappa_\lambda}{216} + \frac{13\kappa_V^4}{270} - \frac{5\kappa_V^3\kappa_\lambda}{54} + \frac{\kappa_V^2\kappa_\lambda^2}{72}\right) \times \left|\frac{d\sigma}{dm_{HH}}(1,10,1)\right|^2 \\ + \left(\frac{88\kappa_{2V}^2}{45} - \frac{16\kappa_{2V}\kappa_V^2}{3} + \frac{4\kappa_{2V}\kappa_V\kappa_\lambda}{9} + \frac{152\kappa_V^4}{45} - \frac{4\kappa_V^3\kappa_\lambda}{9}\right) \times \left|\frac{d\sigma}{dm_{HH}}\left(1,1,\frac{1}{2}\right)\right|^2 \\ + \left(\frac{8\kappa_{2V}^2}{45} - \frac{4\kappa_{2V}\kappa_V\kappa_\lambda}{9} - \frac{8\kappa_V^4}{45} + \frac{4\kappa_V^3\kappa_\lambda}{9}\right) \times \left|\frac{d\sigma}{dm_{HH}}\left(1,-5,\frac{1}{2}\right)\right|^2 , \end{aligned} \tag{6.6}$$

which only depends on $\kappa_{2V}, \kappa_\lambda, \kappa_V$. The six VBF basis samples are fully simulated and the weighted combination is performed after reconstruction and event selection. In summary, the m_{HH} spectrum for any $\boldsymbol{\kappa}$ variation[2] is constructed by performing the weighted combination of the differential cross-sections with Equations 6.3 for the ggF signal or 6.6 for the VBF signal.

The EFT interpretations require several different values of the Wilson coefficients to be considered, which modify the kinematics and cross-section of the HH ggF signal[3]. To avoid the computationally intensive and time-consuming full simulation of all possible signals a similar reweighting method is taken as for the κ_λ variations in the ggF production. The weights to be applied to the SM sample at the reconstruction level are calculated from the parameterised differential cross-section in bins of truth m_{HH}. Events that have truth $m_{HH} > 1040\,\mathrm{GeV}$ are weighted with the highest mass bin weight.

The differential cross-section for HEFT at NLO as derived by [158] is

$$\begin{aligned}\frac{d\sigma(c_{hhh}, c_{tth}, c_{tthh}, c_{ggh}, c_{gghh})}{dm_{HH}} = \\ A_1 c_{tth}^4 + A_2 c_{tthh}^2 + (A_3 c_{tth}^2 + A_4 c_{ggh}^2)\cdot c_{hhh}^2 + A_5 c_{gghh}^2 + (A_6 c_{tthh} + A_7 c_{tth} c_{hhh})\cdot c_{hhh}^2 \\ +(A_8 c_{tth} c_{hhh} + A_9 c_{ghh} c_{hhh})\cdot c_{tthh} + A_{10} c_{tthh} c_{gghh} + (A_{11} c_{ggh} c_{hhh} + A_{12} c_{gghh})\cdot c_{tth}^2 \\ +(A_{13} c_{hhh} c_{ggh} + A_{14} c_{gghh})\cdot c_{tth} c_{hhh} + A_{15} c_{ggh} c_{gghh} c_{hhh} + A_{16} c_{tth}^3 c_{ggh} \\ +A_{17} c_{tth} c_{tthh} c_{ggh} + A_{18} c_{tth} c_{ggh}^2 c_{hhh} + A_{19} c_{tth} c_{ggh} c_{gghh} + A_{20} c_{tth}^2 c_{ggh}^2 \\ +A_{21} c_{tthh} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{13} c_{ggh}^2 c_{gghh},\end{aligned} \tag{6.7}$$

where A_i are given in truth m_{HH} bins in the range $240\,\mathrm{GeV} < m_{HH} < 1040\,\mathrm{GeV}$ with a $20\,\mathrm{GeV}$ width.

On the other hand, SMEFT weights are obtained using LO truth parton level samples generated with MadGraph_aMC@NLO, SMEFT@NLO model [155] from the

[2] $\boldsymbol{\kappa} = (\kappa_\lambda, \kappa_t)$ for ggF and $\boldsymbol{\kappa} = (\kappa_{2V}, \kappa_\lambda, \kappa_V)$ for VBF production.
[3] The EFT interpretation is not done for the VBF production.

differential cross-section

$$\begin{aligned}\frac{d\sigma(c_H, c_{H\Box}, c_{tH}, c_{tG}, c_{HG})}{dm_{HH}} = \\ 1 + A_1 c_H + A_2 c_{H\Box} + A_3 c_{tH} + A_4 c_{tG} + A_5 c_{HG} \\ + A_6 c_H^2 + A_7 c_{tH}^2 + A_8 c_{H\Box}^2 + A_9 c_{tG}^2 + A_{10} c_{HG}^2 \\ + A_{11} c_H c_{H\Box} + A_{12} c_H c_{tH} + A_{13} c_H c_{tG} + A_{14} c_H c_{HG} \\ + A_{15} c_{H\Box} c_{tH} + A_{16} c_{tG} c_{H\Box} + A_{17} c_{H\Box} c_{HG} + A_{18} c_{tH} c_{tG} \\ + A_{19} c_{tH} c_{HG} + A_{20} c_{tG} c_{HG}.\end{aligned} \tag{6.8}$$

In some cases, it is possible to relate HEFT coefficients to SMEFT coefficients, the translations can be found in [41]. Therefore, a k-factor defined as the ratio of the HEFT NLO cross-section to the HEFT LO cross-section, the latter is derived in [32], for the corresponding signal is used to normalise the SMEFT signal to a proxy NLO cross-section. In the set of coefficients considered only c_{tG} cannot be related to any HEFT variation and therefore higher-order effects are not accounted for and the LO cross-section is taken.

6.4 Event Selection

6.4.1 Object Specifications

In the event selection PFlow jets with $R = 0.4$, as defined and described in Section 5.3, are split into central and forward jets. Central jets are required to have $|\eta| < 2.5$ and $p_T \geq 40\,\mathrm{GeV}$. Provided b-tagging requirements are met they make up the Higgs candidates and are used to compute the trigger SFs. A jet is b-tagged, if it passes the 77% WP of the DL1r tagger. Forward jets must have $|\eta| \geq 2.5$ and $p_T \geq 30\,\mathrm{GeV}$ and

help capture the topology of VBF events. Muons are only needed to correct the energy of b-tagged jets with the procedure described in Section 5.3, i.e. for the event to be a signal candidate a muon is not required.

6.4.2 Triggers

A combination of b-jet triggers is used to select events that are most likely to record the $HH \to b\bar{b}b\bar{b}$ signal and cover the whole m_{HH} spectrum[4]. The triggers used are the 2b1j and 2b2j, which are fired if the event contains two b-tagged jets and 1 (2) extra jet(s). The b-tagging WP and additional p_T requirements on the jets are summed up in Table 6.2 for each trigger. The 2b2j trigger targets the low m_{HH} spectrum, while the 2b1j trigger covers higher m_{HH} masses.

The challenge of collecting events with multiple trigger chains is to determine the event-level trigger scale factor (SF), as the difference between data and simulation due to trigger inefficiencies is evaluated per object that fired the trigger. Hence, to obtain the total event-level SF all possible jet combinations and correlations between the triggers would need to be accounted for. To simplify the calculation events are put into "buckets" that consider only a single trigger decision. Inevitably some events are lost and to minimise the loss, the bucket categorisation is done by matching the kinematic cuts applied on offline fully reconstructed objects close to the trigger thresholds. The first bucket contains events in which the leading jet[5] and third jet have $p_T^{1st} > 170\,\mathrm{Gev}$ and $p_T^{3rd} > 70\,\mathrm{Gev}$ and pass the 2b+1j trigger. The second bucket contains events that do not meet the kinematic requirement and pass the 2b+2j trigger.

The online b-tagging requirements and online jet kinematic cuts are the two sources of the trigger decision efficiency difference between data and simulation. These need to be corrected through the online per jet-derived SFs. As the online and

[4] A general overview of triggers in ATLAS can be found in Section 3.3.
[5] The jets are sorted in p_T.

offline b-tagging efficiencies are correlated the event level SF has to be calculated jointly. First, the online jets need to be matched to offline jets by requiring their separation to be $\Delta R < 0.4$. The whole event is disregarded if the online jet does not have a match. On the other hand, if the offline jet has no trigger object matched to it, only the offline SF ($SF_i(off_{DL1r_{77}})$) is considered. For matched jets, the conditional online b-tagging SF, ($SF_i(\text{on}|\text{off}_{DL1r_{77}})$), is derived with respect to the chosen offline b-tagging WP, which is the 77% WP of the DL1r tagger in the analysis. The total event level b-tagging SF is computed according to

$$SF_{b-tag} = \prod_i SF_i(on|off_{DL1r_{77}})SF_i(off_{DL1r_{77}}), \tag{6.9}$$

where i runs over all offline jets in the event. Similarly, the per-jet trigger kinematics SF is derived with respect to its matched jets. The two event-level SFs are then simply multiplied together.

Year	Trigger Name	Trigger Type	Conditions
2016	HLT_j100_2j55_bmv2c2060_split	2b+1j	2 b-tagged jets at the 60% WP with $p_T > 55\,\text{Gev}$ in addition to 1 jet with $p_T > 100\,\text{Gev}$
	HLT_2j35_bmv2c2060_split_2j35_L14J15.0ETA25	2b+2j	2 b-tagged jets at the 60% WP with $p_T > 25\,\text{Gev}$ in addition to 2 jets with $p_T > 35\,\text{Gev}$
2017	HLT_j110_gsc150_boffperf_split_2j35_gsc55_bmv2c1070_split_L1J85_3J30	2b+1j	2 b-tagged jets at the 70% WP with $p_T > 55\,\text{Gev}$ in addition to 1 jet with $p_T > 110\,\text{Gev}$
	HLT_2j15_gsc35_bmv2c1040_split_2j15_gsc35_boffperf_split_L14J15.0ETA25	2b+2j	2 b-tagged jets at the 60% WP with $p_T > 35\,\text{Gev}$ in addition to 2 jet with $p_T > 35\,\text{Gev}$
2018	HLT_j110_gsc150_boffperf_split_2j45_gsc55_bmv2c1070_split_L1J85_3J30	2b+1j	2 b-tagged jets at the 70% WP with $p_T > 55\,\text{Gev}$ in addition to 1 jet with $p_T > 110\,\text{Gev}$
	HLT_2j35_bmv2c1060_split_2j35_L14J15.0ETA25	2b+2j	2 b-tagged jets at the 60 % WP with $p_T > 35\,\text{Gev}$ in addition to 2 jet with $p_T > 35\,\text{Gev}$

Table 6.2: The trigger used in the $HH \rightarrow b\bar{b}b\bar{b}$ as defined per year of data taking. The Trigger Type is the shortened name, while the Trigger Name contains the exact conditions that need to be met to trigger an event.

6.4.3 Reconstructing the diHiggs System

Events are required to have at least four central jets, which are b-tagged by the DL1r tagger at the 77% WP. The four b-tagged central jets leading in p_T are paired to form the two Higgs candidates (HC). The pairing strategy chosen, called minΔR, considers all three possible pairings. The b-jet pair with the highest p_T four-vector sum is the leading HC, and the second pair is the subleading HC. The angular separation (ΔR)

between jets in the $b\bar{b}$ pair of the leading HC is computed for each pairing. The chosen pairing is the one with the smallest ΔR out of the three, as shown in Figure 6.2. The motivation behind the strategy is that the HC has some initial momentum and its decay products are collimated to some extent. The minΔR pairing strategy is rather simplistic compared to alternative machine learning methods, which can account for several event properties and are more efficient at low diHiggs masses. The rapid fall in the efficiency of the minΔR pairing strategy below $m_{HH} < 400\,\mathrm{GeV}$ when pairing ggF SM signal events is shown in Figure 6.3. The main advantage of the chosen pairing strategy is the smooth background with little sculpting, which is beneficial for the data-driven background estimation technique.

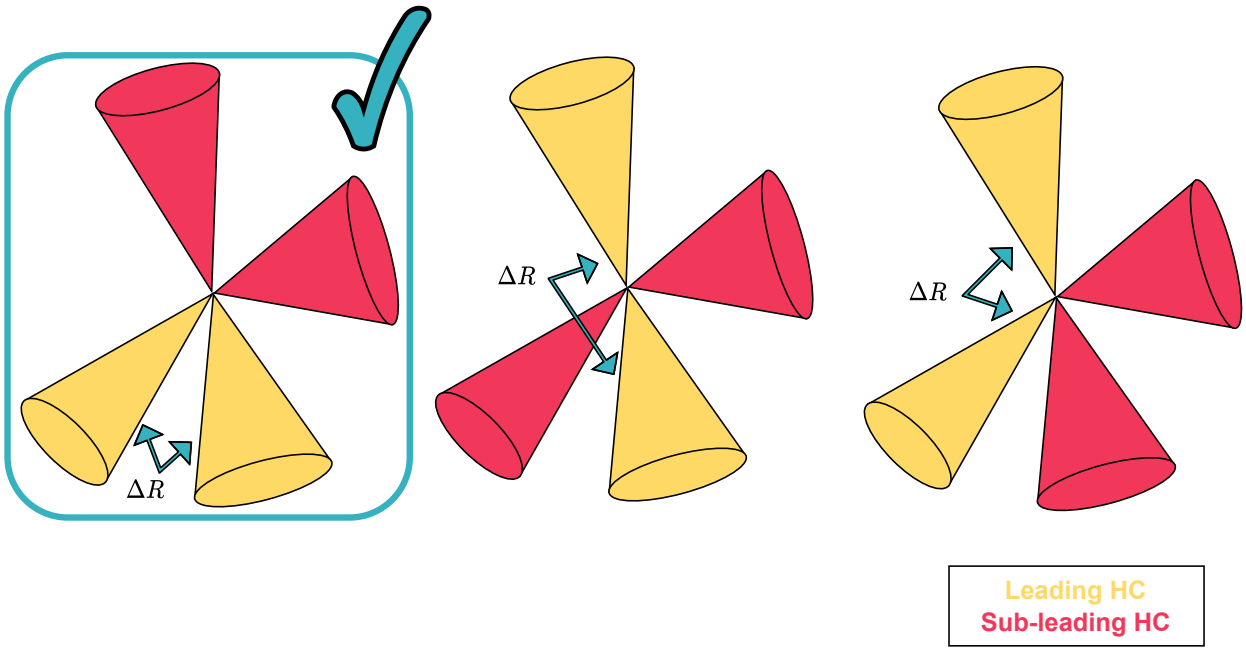


Figure 6.2: The sketched minΔR pairing strategy. The correct pairing is circled and indicated by a tick. The leading HC jets are shown in yellow and the sub-leading HC jets are shown in pink. Only the ΔR between the jets in the leading HC is considered in the pairing strategy.

Only events with four or more central jets that pass the b-tagging, referred to as 4b events, are considered signal candidates. However, events with exactly two b-tagged central jets, 2b events, come into play in the background estimation strategy. Besides the number of required b-jets, the selection of 2b and 4b events is the

same. In the HC pairing of 2b events the two b-tagged jets and two more jets leading in p_T but failing the b-tag requirement are considered.

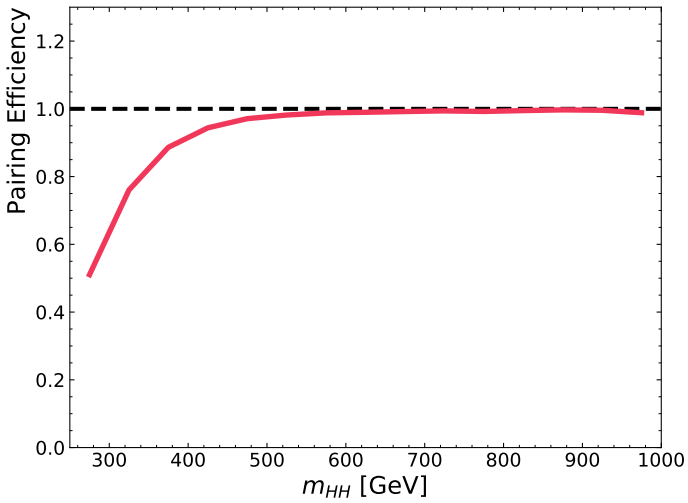


Figure 6.3: The efficiency of minΔR as the mass of the diHiggs system for the ggF SM signal. The efficiency is evaluated on events that have two truth Higgs bosons and at least four central jets, no other event selection is applied.

Once the Higgs candidates are reconstructed, the VBF-specific selection is applied. It is based on the characteristic VBF topology, the presence of two well-separated energetic VBF jets. Forward, and central jets not used to construct the HCs, are used for the VBF jets, which are defined as two extra jets with a combined mass greater than $1\,\mathrm{TeV}$ and separated by $\Delta\eta_{jj} \geq 3$. In addition to the VBF jets, the total vector sum of the p_T of HC and VBF jets is required to be less than $65\,\mathrm{GeV}$ for the event to pass the VBF-specific selection. To keep the ggF and VBF selections completely orthogonal only events that do not pass the VBF selection are further considered in the ggF selection, as shown in the event selection flow chart in Figure 6.4. The VBF selection ensures that only a small amount of the ggF signal leaks into the VBF channel. In the case of the ggF SM signal, only about 1.3% of the events pass the VBF selection, as can be seen from the cut flow shown in Table 6.3.

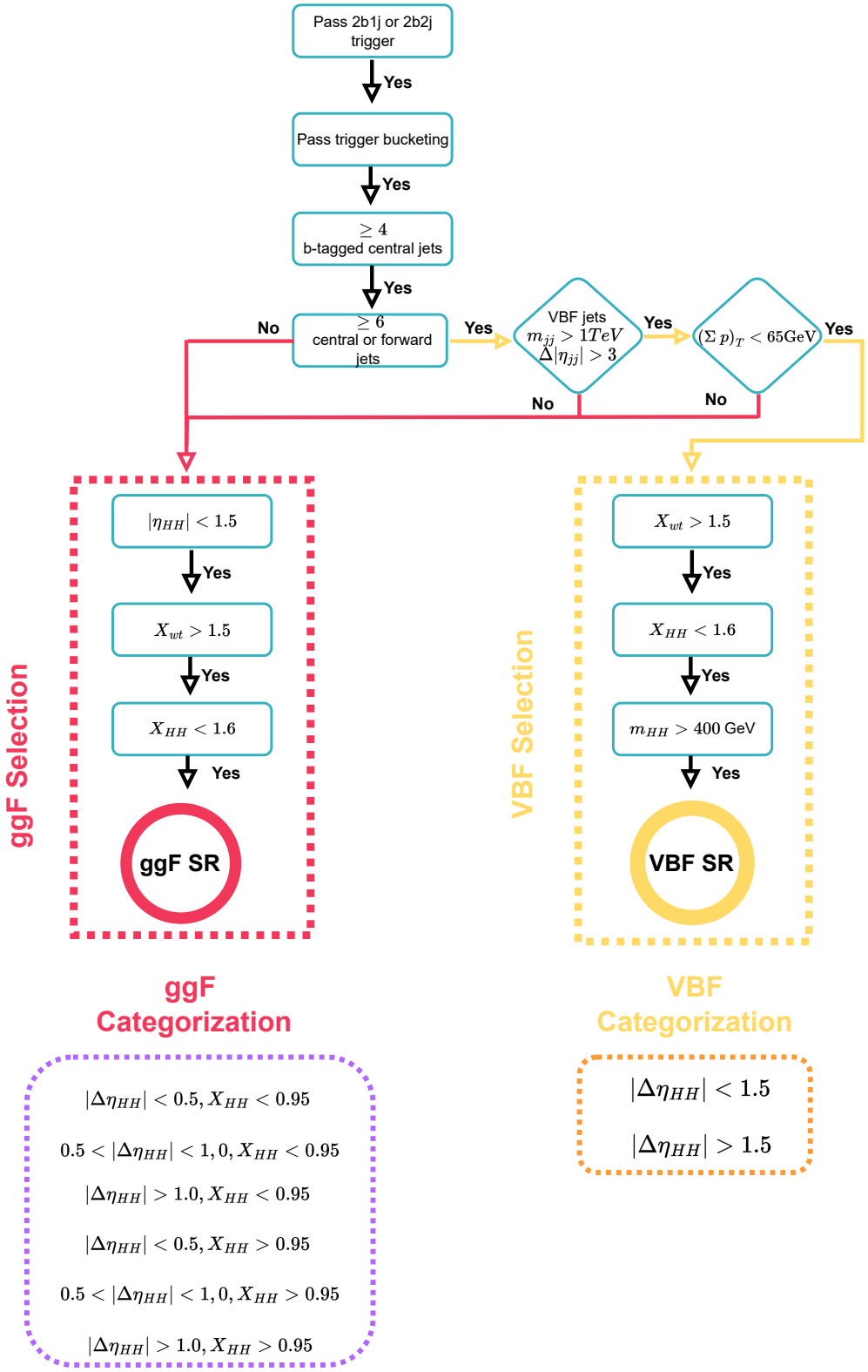


Figure 6.4: The summary of the event selection. The final categorisation in the ggF and VBF channels that increases the sensitivity to the SM signal is shown in the two dotted boxes.

6.4.4 Reducing the Background

The $t\bar{t}$ background makes up approximately 18% (12%) of the total background in the ggF (VBF) channel, which is reduced by applying a top veto that picks up events that are kinematically consistent with a hadronically decaying top quark. These events are identified by first constructing all possible W candidates from all possible pairings of central jets. Each W candidate is used to capture a top decay by again considering all possible combinations with a b-tag jet out of the remaining central and forward jets not used to form the W. The discriminant X_{Wt}, defined as

$$X_{Wt} = \sqrt{\Big(\frac{m_W - 80.4\,\mathrm{GeV}}{0.1 m_W}\Big)^2 + \Big(\frac{m_t - 172.5\,\mathrm{GeV}}{0.1 m_t}\Big)^2}, \tag{6.10}$$

is calculated for all possible combinations that could form the W and top. Only the minimum X_{Wt} in an event is considered and the event is discarded, if it is smaller than 1.5. The cut reduces the $t\bar{t}$ background component in the ggF (VBF) channel to approximately 10% (4%) of the total background, depending on the year of data taking and channel, based on $t\bar{t}$ simulations. The b-tagged requirement of the jet paired with the W candidate increases the accuracy of the cut as more than 95% of all top decays are $t \to Wb$ [24]. An improvement in the VBF cross-section limits on the SM, $\kappa_{2V} = 0$ and $\kappa_\lambda = 10$ between 9%-18% is observed as compared to applying no such requirement. It can be seen from Figure 6.5 that the X_{Wt} distributions for fully hadronically and semi-hadronically decaying $t\bar{t}$ events peaks at low values ($X_{Wt} < 1.5$), while the majority of the non-resonant SM, as well as an example BSM signal, is located at $X_{Wt} > 1.5$ in both the ggF and VBF channels.

The total background is further reduced in the ggF channel by requiring the two Higgs candidates not to be separated by a pseudo-rapidity larger than 1.5. Figure 6.6 shows that the bulk of the ggF SM signal events indeed have $\Delta\eta < 1.5$, while the blinded[6] 4b data has a larger spread.

[6]The observed data in the signal region is not considered in order to avoid introducing any bias when optimising the analysis.

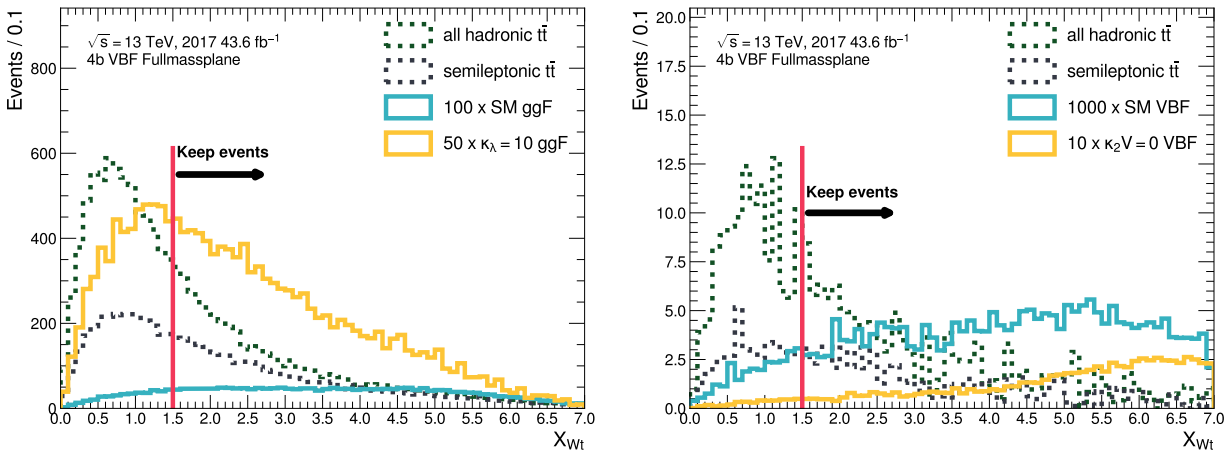


Figure 6.5: The X_{Wt} distribution for the fully hadronically (dotted green line) and semi-leptonically (dotted grey line) decaying $t\bar{t}$ background in the ggF and VBF channels are shown along the SM (solid turquoise line) and a BSM (solid yellow line) non-resonant signal. The normalisation of the $t\bar{t}$ sample is taken from the simulation, while the HH cross-section is the theoretically calcualted value. All the histograms shown are scaled to the integrated luminosity collected during 2017. The SM ggF and VBF signals are scaled by a factor of 100, while the BSM ggF $\kappa_\lambda = 10$ (VBF $\kappa_{2V} = 0$) signal is scaled by a factor of 50 (10) for clarity.

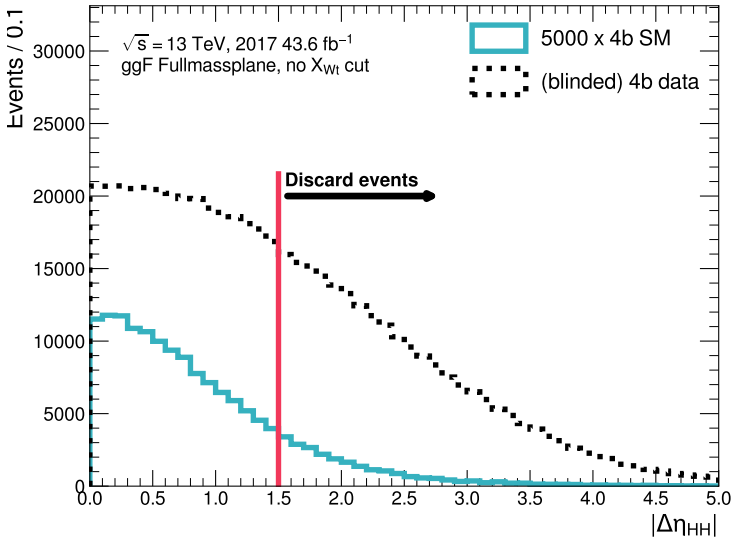


Figure 6.6: The distribution of the blinded 4b data (dotted line) and the ggF SM 4b signal (solid line) is shown. It can be seen that placing a cut $\Delta\eta < 1.5$ reduces the background, while the majority of the signal is still kept. The data shown was collected during the year 2017 of data taking. The SM HH signal is normalised to the SM cross-section obtained from theory, and scaled to the integrated luminosity collected during 2017. In addition, the SM HH signal is scaled by a factor of 5000 for easier visualisation.

	Data	ggF SM Signal	ggF $\kappa_\lambda = 10$	VBF SM Signal	VBF $\kappa_{2V} = 0$
4 jets with $p_T > 40\,\mathrm{GeV}$, of which ≥ 2 b-tagged	5.70×10^8	526.6	7337.7	22.3	626.1
Trigger Buckets	2.49×10^8	381.8	5279.1	16.1	405.2
ggF Channel					
Fail VBF Selection	2.46×10^8	376.6	5198.0	13.9	334.4
≥ 4 b-tagged central jets	1.89×10^6	86.0	1001.7	1.9	65.2
$\lvert\Delta\eta_{HH}\rvert < 1.5$	1.03×10^6	71.9	850.6	0.9	46.4
$X_{Wt} > 1.5$	7.51×10^5	60.4	569.0	0.7	43.1
$X_{HH} > 1.6$	1.62×10^4	29.1	182.7	0.2	23.0
VBF Channel					
Pass VBF Selection	3.30×10^6	5.2	81.1	2.2	70.7
≥ 4 b-tagged central jets	2.71×10^4	1.1	15.3	0.7	27.6
$X_{Wt} > 1.5$	2.18×10^4	1.0	11.2	0.7	26.5
$X_{HH} > 1.6$	5.02×10^2	0.5	3.1	0.3	17.3
$m_{HH} > 400\,\mathrm{GeV}$	3.57×10^2	0.4	1.8	0.3	16.4

Table 6.3: The cut flow for the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis as presented in [2] is shown. The yield after each selection step is shown for data, the ggF SM signal, the ggF $\kappa_\lambda = 10$ signal, the VBF SM signal and the VBF $\kappa_{2V} = 0$ signal. To obtain the HH signal yields the corresponding simulations are normalised to the theoretically calculated cross-section and luminosity of $126\,\mathrm{fb}^{-1}$. After and including the trigger bucket step, which includes the whole trigger selection, SFs are applied to the MC samples.

6.4.5 Kinematic Regions

The kinematic regions are defined in the (m_{H1}, m_{H2}) mass plane shown in Figure 6.7, where m_{H1} is the mass of the leading Higgs candidate and m_{H2} is the mass of the sub-leading Higgs candidate. The signal region (SR) is defined as $X_{HH} < 1.6$, where

$$X_{HH} = \sqrt{\left(\frac{m_{H1} - 124\,\mathrm{GeV}}{0.1 m_{H1}}\right)^2 + \left(\frac{m_{H2} - 117\,\mathrm{GeV}}{0.1 m_{H2}}\right)^2}. \tag{6.11}$$

The resultant shape of a deformed ellipse is well-motivated. In a diHiggs signal event, each of the $b\bar{b}$ pairs is expected to have an invariant mass close to $125\,\mathrm{GeV}$, whereas the background has no such resonant structure. Therefore, the centre of the SR is at $m_{H1} = 124\,\mathrm{GeV}$ and $m_{H2} = 117\,\mathrm{GeV}$, slightly moved from $125\,\mathrm{GeV}$ to

account for energy losses in the ATLAS detector. The elliptical shape is distorted by the HC-dependent radius, to ensure that the resolution effects are taken into account. Figure 6.8 shows that such an SR definition indeed captures the resonance produced by the ggF and VBF SM signals with only little leakage into the CRs. The

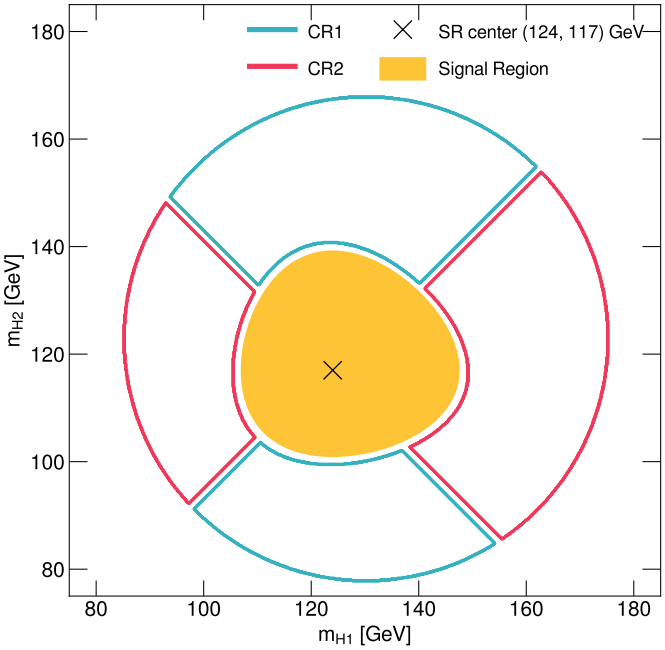


Figure 6.7: The mass plane shows the kinematic regions used in the analysis, signal region (yellow area), control region CR1 (bound by the turquoise counter), and control region CR2 (bound by the pink counter). The m_{H1} is the mass of the reconstructed Higgs candidate leading in p_T, while m_{H2} is the mass of the reconstructed Higgs candidate sub-leading in p_T.

control region (CR) surrounds the SR, the inner boundary is defined as $X_{HH} = 1.6$, while the outside boundary is defined as the counter $CR_{HH} = 45\,\mathrm{GeV}$, where

$$CR_{HH} = \sqrt{(m_{H1} - 1.05 \cdot 124\,\mathrm{GeV})^2 + (m_{H2} - 1.05 \cdot 117\,\mathrm{GeV})^2}. \tag{6.12}$$

As shown in Figure 6.7 the CR is further split into CR1, the north and south quadrant, and CR2, the east and west quadrants. The nominal background estimation is derived in CR1, while CR2 is used to evaluate its uncertainty. The choice of defining the quadrants by a line going through the SR centre of the mass plane at a $45°$ and

$135°$ is arbitrary, varying it does not make a significant difference. The distribution of 4b data events passing the ggF and VBF event selection up to the X_{HH} cut in the (m_{H1}, m_{H2}) mass plane is shown in Figure 6.8c and 6.8d respectively.

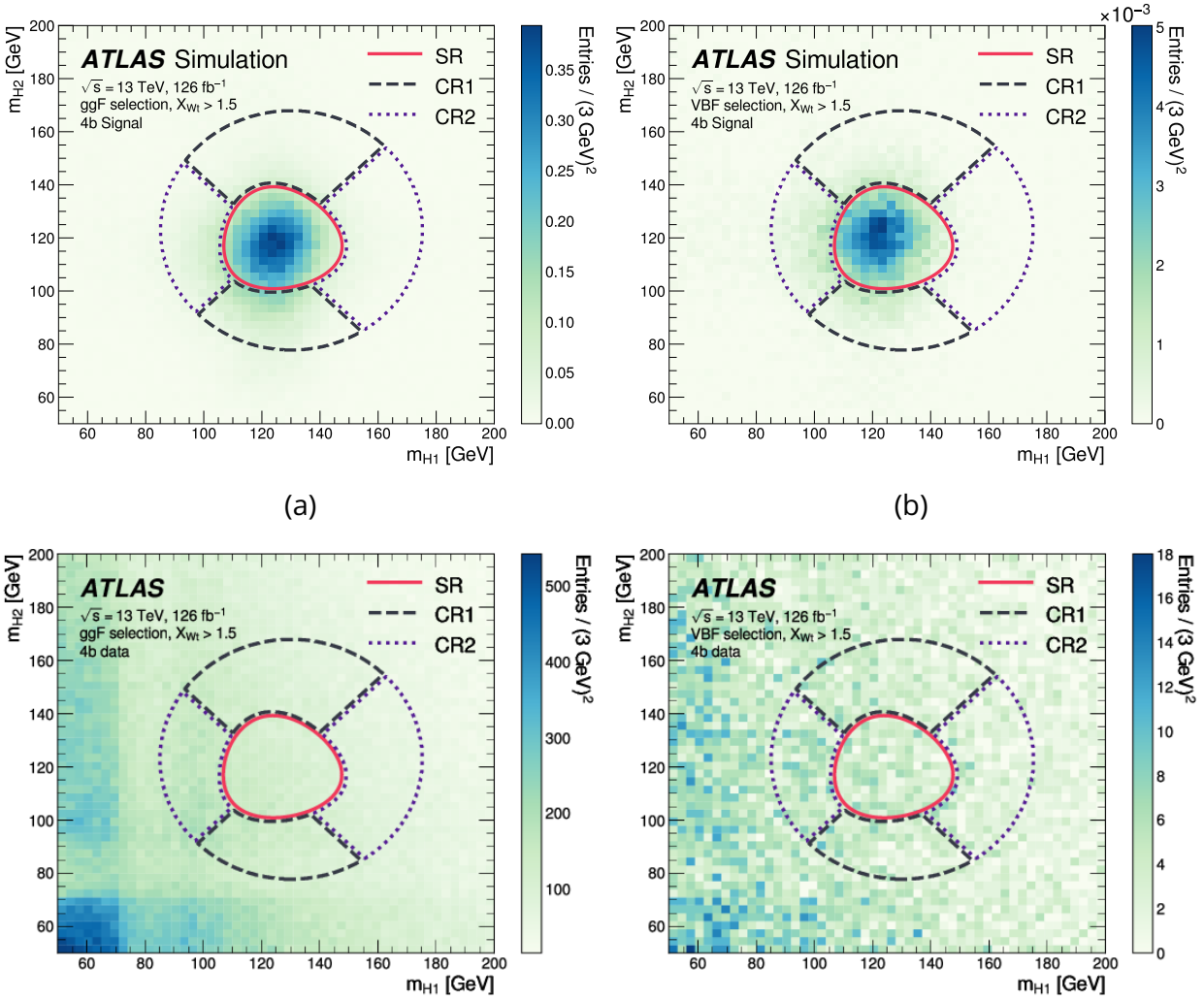


Figure 6.8: The upper subfigures show the non-resonant ggF (left) and VBF (right) SM signal with a clear peak in the SR and only a small leakage into the CR. The lower subfigures show the 4b data in the ggF (left) and VBF (right) channels. The effect of the X_{Wt} cut can be seen on the ggF 4b data, as stripes around $m_{H1/H2} \sim 80\,\mathrm{GeV}$. The plots have been published in the non-resonant $HH \to b\bar{b}b\bar{b}$ paper [2].

6.4.6 Categorisation

The full event selection chain including the final categorisation used in the ggF and VBF channel is shown in Figure 6.4. The figure of merit used in choosing the final categorisation is the upper limit on the SM signal strength. There are six categories in the ggF SR, where events are split into the inner SR, $X_{HH} < 0.95$, and outer SR, $X_{HH} > 0.95$, which are additionally categorised by $|\Delta\eta|$ between the Higgs candidates. The $|\Delta\eta|$ cuts are, $|\Delta\eta| < 0.5$, $0.5 < |\Delta\eta| < 1.0$, $|\Delta\eta| > 1.0$. The VBF SR is only split into two categories, $|\Delta\eta| > 1.5$ and $|\Delta\eta| < 1.5$, due to limited statistics. The distribution of 4b data events passing the ggF and VBF event selection up to the X_{HH} cut in the (m_{H1}, m_{H2}) mass plane is shown in Figure 6.8c and 6.8d respectively. The yields in data and some simulated HH signals, including the ggF and VBF non-resonant SM signals, after each event selection step but excluding the final categorisation are shown in Table 6.3.

6.5 Background Estimation

After applying the event selection the background is still significant. It is largely composed of QCD multi-jet processes (about 90%), while the remaining 10% are misidentified $t\bar{t}$ decays. The single H and ZZ decay contribution is negligible and not accounted for explicitly by the analysis.

The QCD and $t\bar{t}$ backgrounds are estimated simultaneously using a data-driven approach inspired by the ABCD method of which a classic example is presented in [159]. The primary motivation for a data-driven technique is that simulated multi-jet events carry large uncertainties making their modelling of physical quantities unreliable. In addition, generating a reliable high-statistics background model is very CPU intensive. The $t\bar{t}$ decays can be well-modelled with available simulations, which are used for checks in analysis (see Figure 6.5). However, they are not used in the

background estimation as the data-driven method models the $t\bar{t}$ contribution well and the process is sub-dominant.

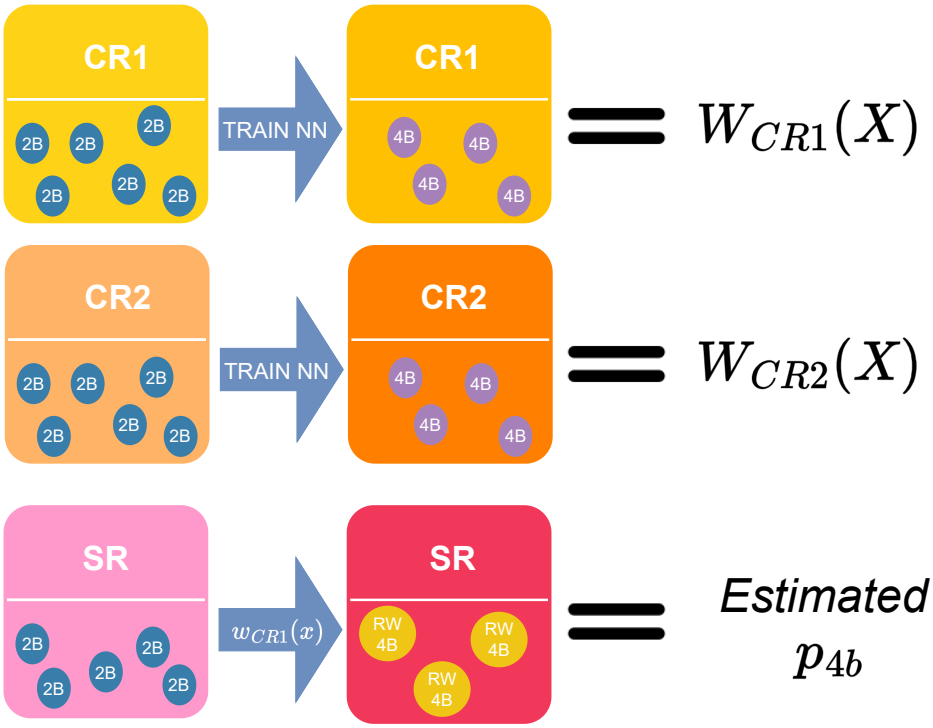


Figure 6.9: The background estimation technique inspired by the ABCD method is depicted. It shows that two independent sets of weights ($w(\vec{x})$) are derived from CR1 and CR2. The NN is trained on a set of kinematics properties of 2b and 4b events in each control region (CR1 or CR2) to learn the mapping. The analysis is blinded, meaning that the 4b kinematics in the SR are unknown. To estimate its background components the 2b data probability density function ($p_{2b}(\vec{x})$) in the SR is mapped with the learned $w(\vec{x})$ to the corresponding 4b data probability density function ($p_{4b}(\vec{x})$) in the SR. To obtain the nominal background estimation only weights derived from CR1 are applied. The learned mapping in CR1 is evaluated by comparing the nominal background estimation to alternative background estimations obtained by applying a combination of $w(\vec{x})_{CR1}$ and $w(\vec{x})_{CR2}$ to the 2b events in the SR as described in detail in the next section.

The data-driven method is based on the assumption that there exists a mapping between 2b and 4b events. The mapping is not as trivial as normalising 2b to 4b events, as different physics processes contribute to the 2b and 4b events that lead to notable differences in kinematics. A reweighting function that corrects the kinematics of a 2b event to match those of a 4b event can be obtained by training an

artificial neural net (NN). The learning step is done in CR1 and CR2 independently on 2b and 4b events, while the learned mapping is applied to 2b events in the signal region (SR) to estimate the 4b background as depicted in Figure 6.9. The kinematic regions used for training are described in Section 6.4.5 and depicted in Figure 6.7. The mapping extracted from CR1 is used to predict the background in the SR, while the reweighting function derived in CR2 is used to define an uncertainty on the nominal background estimation, as described in Section 6.6. The kinematic reweighting has already been introduced in the resolved resonant $HH \to b\bar{b}b\bar{b}$ analysis [160] but with differently defined kinematic regions.

6.5.1 Kinematic Reweighting

The task at hand is to learn $w(\vec{x})$, which maps the 2b data probability density function $p_{2b}(\vec{x})$ into the 4b data probability density function $p_{4b}(\vec{x})$ according to

$$p_{2b}(\vec{x}) \cdot w(\vec{x}) = p_{4b}(\vec{x}), \tag{6.13}$$

where $\vec{x}$ are carefully chosen inputs that capture the kinematic difference between 2b and 4b events. Table 6.4 shows the $\vec{x}$ inputs, which differ between the ggF and VBF channels. The technique to solve for $w(\vec{x})$ is adapted from [161, 162] and

$$\mathcal{L}(w(\vec{x})) = \int d\vec{x} \Big[\sqrt{w(\vec{x}) p_{2b}(\vec{x})} + \frac{1}{w(\vec{x})} p_{4b} \Big] \tag{6.14}$$

is the loss function to be minimised by the NN. The NN used to learn the reweighting of data passing the ggF selection consists of three connected hidden layers, each with 50 nodes, while only 20 nodes per layer are used for the VBF selection.

The training is done for each year separately in the ggF channel due to expected kinematic differences coming from the trigger conditions. All years are trained inclusively in the VBF channel as the accuracy of the NN training would suffer

ggF	**VBF**
Average $\|\eta\|$ of the Higgs boson candidate jets	
Trigger bucket index as one-hot encoder	
$\log(X_{Wt})$	
$\log(p_T)$ of the 2^{nd} leading HC jet	Of all possible four HC pairing, the maximum dijet mass
$\log(p_T)$ of the 4^{th} leading HC jet	Of all possible four HC pairing, the minimum dijet mass
$\log(p_T)$ of the diHiggs system	Energy of leading HC
$\log(\Delta R)$ of the diHiggs system between the closest two HC jets	Energy of sub-leading HC
$\log(\Delta R)$ of the diHiggs system between the remaining two HC jets	Of all possible four HC pairing, the 2^{nd}-smallest ΔR between the leading HC
ΔR between the two HCs	Year index as one-hot encoder
$\Delta\Phi$ of the jets in the leading HC	
$\Delta\Phi$ of the jets in the sub-leading HC	
Number of jets in the event	

Table 6.4: The inputs ($\boldsymbol{x}$) used by the NN to learn the kinematic mapping between 2b and 4b events in the ggF and VBF channels. One-hot encoding assigns each bucket (category) a unique binary array of length k, where k is the number of classes, with exactly one element set to 1 and all the others to 0 [163].

significantly due to low per-year statistics. The X_{Wt} cut is applied to the training dataset in the ggF channel but not in the VBF channel. The background in the final SR categorisation is not expected to have large enough kinematic differences to require the NN training to be performed per category.

It follows that the reweighting function needs to be extracted in each of the eight orthogonal ensembles of data; the per year ggF selection and the VBF selection per control region. The bootstrap resampling technique [164] with replacement is used to generate hundred training datasets per ensemble. Each dataset is used for training independently. The mean $w(\vec{x})$ calculated from the hundred datasets trained is taken to reweight 2b events in the SR per training region, where $w(\vec{x})$ is the per-event weight.

6.5.2 Validation

The performance of reweighting the 2b data is checked for the kinematic variables used and not used in the NN training. All checked distributions are found to have the same or improved modelling of the 4b data as compared to the simple normalisation correction of the 2b data. The initial checks are performed in CR1, and the 4b data and reweighted 2b data are found to agree within 10%. Figure 6.10 shows the improvement in modelling the 4b data with reweighted 2b data as compared to simply correcting the normalisation in the ΔR_{HH} distribution (upper row) used for training and $p_{\mathrm{T\,h1}}$ (middle row) and the number of jets (bottom row) which are not used in training. The modelling of the m_{HH} distribution in CR1 (CR2) is shown in Figure 6.11, wherein $w_{CR1}(\boldsymbol{x})$ ($w_{CR2}(\boldsymbol{x})$) reweight the 2b data. It can be seen that the reweighting derived in CR1 and CR2 have good closure. However, when comparing the reweighted 2b data in the SR using $w_{CR1}(\boldsymbol{x})$ or $w_{CR2}(\boldsymbol{x})$ the differences are not covered by the statistical error. Hence, a non-closure systematic is assigned to the background estimation. The systematic is defined in Section 6.6.1.

Additionally, the NN reweighting is validated in several regions orthogonal to the 4b SR data, where the $HH \to 4b$ signal contamination is negligible. Good closure is seen when;

- The shape of the kinematic region is shifted in the mass plane and the background procedure is evaluated in the shifted regions. Four alternative kinematic regions, which do not overlap with the SR at any point, are used to validate the background estimation.
- The events in the ggF channel that do not pass the $|\eta| < 1.5$ cut (reverse η region) are used to test the mapping of 2b to 4b events.

Lastly, the reweighting is checked by learning to map the 2b events to the distribution of 3b1f events in the SR. The 3b1f events have exactly 3 b-tagged events, while the jets that fail the 77% WP must fail the 85% WP as well to ensure that the signal contamination is negligible. There is good agreement within statistical uncertainties observed in the VBF channel, while non-closure is seen in the ggF channel. Therefore, an additional systematic uncertainty is defined from the discrepancy between the reweighted and observed 3b1f events in the SR in the ggF regions with the procedure described in 6.6.1.

6.6 Systematic Uncertainties

This section outlines the dominant sources of uncertainty, the samples they affect, and the correlation scheme applied. The dominant source of uncertainty is on the background estimation, see Section 6.6.1 for details. As the background estimation is entirely data-driven, only the statistical uncertainty and non-closure systematics, are considered. The experimental and theoretical uncertainties, outlined in Section 6.6.2 and 6.6.3, only affect the signal samples. Finally, Section 6.6.5 describes the extraction of the uncertainty on the EFT signal reweighting.

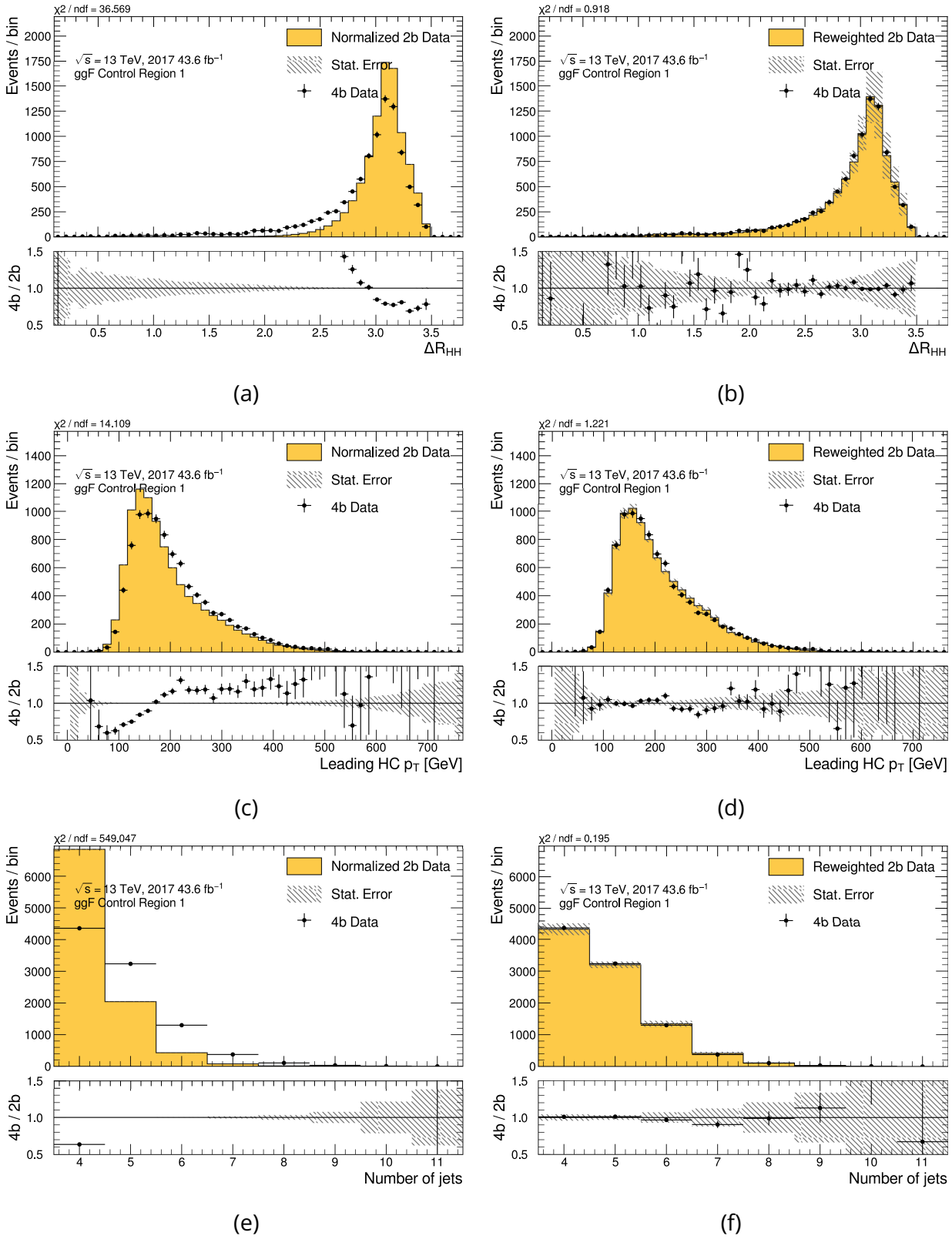


Figure 6.10: The comparison of the 2b normalised (left) and reweighted (right) events to the 4b distribution of ΔR_{HH} (upper row), $p_{\mathrm{T\,h1}}$ (middle row) and number of jets (bottom row) in the CR1 is shown. Only the ΔR_{HH} variable is used in the training. The reweighted 4b distribution is obtained by applying per-event weights derived in CR1 from the NNs trained. The statistical uncertainty on the reweighted 2b distributions includes the Possion error from 2b data and the bootstrapping uncertainty defined in Section 6.6.1.

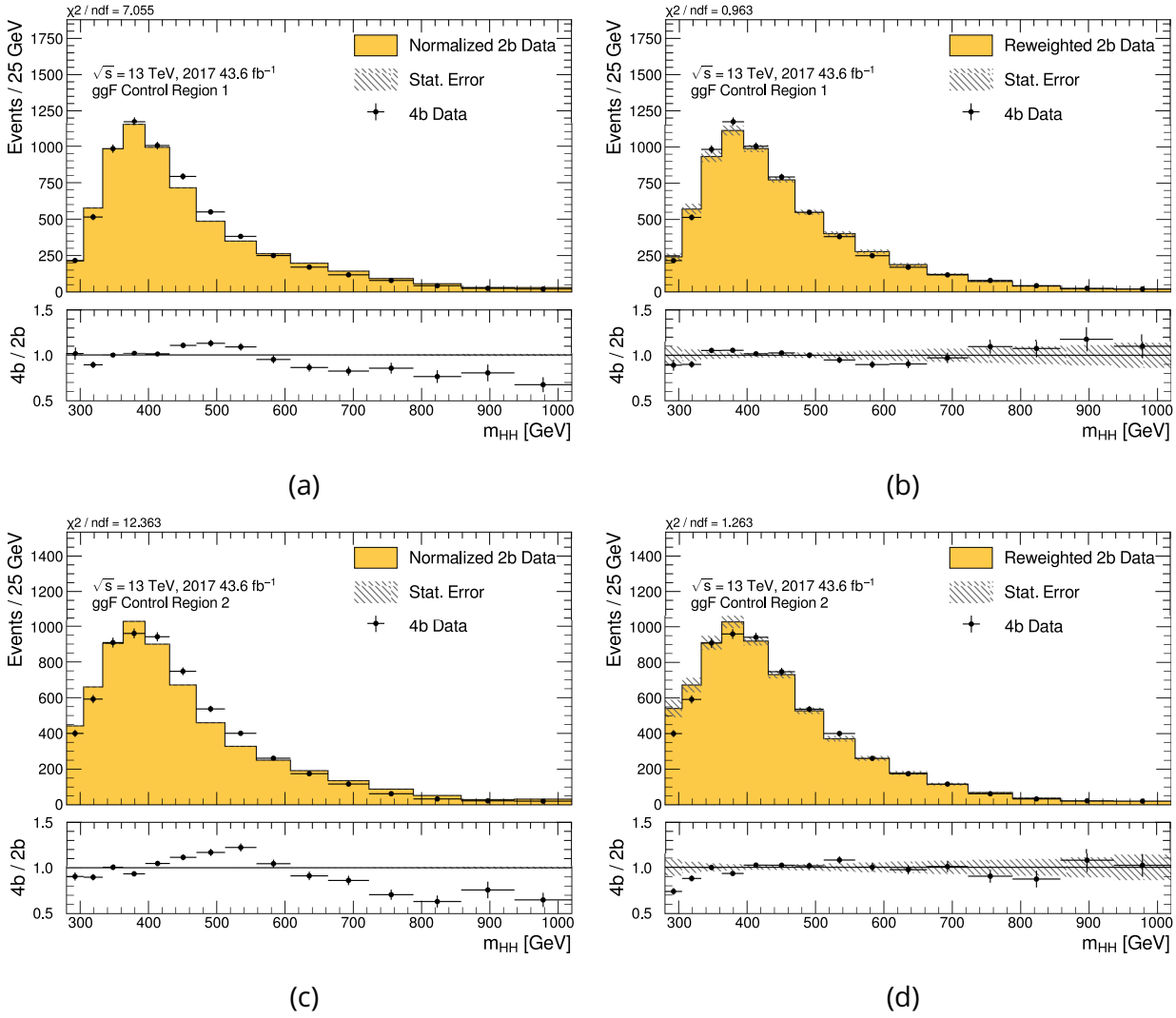


Figure 6.11: The 2b data normalised (left) and reweighted (right) plot compared to the 4b data. Clearly the reweighting outperforms the normalisation, despite m_{HH} not being used in the training. The upper two plots show CR1, while the lower two plots show CR2. The 2b distributions in (b) and (d) have weights derived from CR1 and CR2 respectively. The Poisson error from 2b data and the bootstrapping uncertainty are included in the statistical error on reweighted 2b data.

6.6.1 Background Estimation Uncertainties

The statistical uncertainty has two components, the Poisson error on the reweighted 2b SR data and the uncertainty on the NN-derived weights, due to the finite training dataset. The uncertainties are combined in quadrature, and treated as a single statistical uncertainty on the background estimation and are uncorrelated across bins. The per bin uncertainty from the statistical fluctuations of the reweighted 2b SR data is $\delta N_i = \sqrt{\sum_{k \in i} w_k^2}$, where N_i is the number of events in bin i and w_k is the weight of event k which falls into bin i. The bootstrap component of the statistical uncertainty is derived by generating 100 background estimations from each NN trained for the 4b SR m_{HH} distribution. The per bin mean is taken as the nominal background estimation, while the standard deviation provides the statistical uncertainty on the training. The dominant source of the bootstrap uncertainty is the 4b data used for training.

The background estimation method assumes that the same physics processes contribute to the QCD and $t\bar{t}$ backgrounds in the CR and SR. However, this is only approximately true, some differences in the kinematics are expected, which are covered by the non-closure systematic estimated from the alternative background estimation derived in CR2. A four-point systematic is constructed by comparing the nominal background estimation to the alternative background estimations obtained from applying w_{CR2} to a part of the SR, while w_{CR1} to the remaining 2b events in the SR. To be precise the SR is split into quadrants, labelled by the cardinal directions as Q_N,Q_S,Q_W,Q_E as shown Figure 6.12. The nuisance parameter (NP) associated with Q_W, labelled as CR12_shape_W or CR12_shape_W_{year of data taking} in the ggF channel, for example, is defined as follows;

- Apply weight derived from CR2 to 2b data in Q_W
- Apply weight derived from CR1 to 2b data in the renaming three quadrants, Q_N, Q_S, Q_E

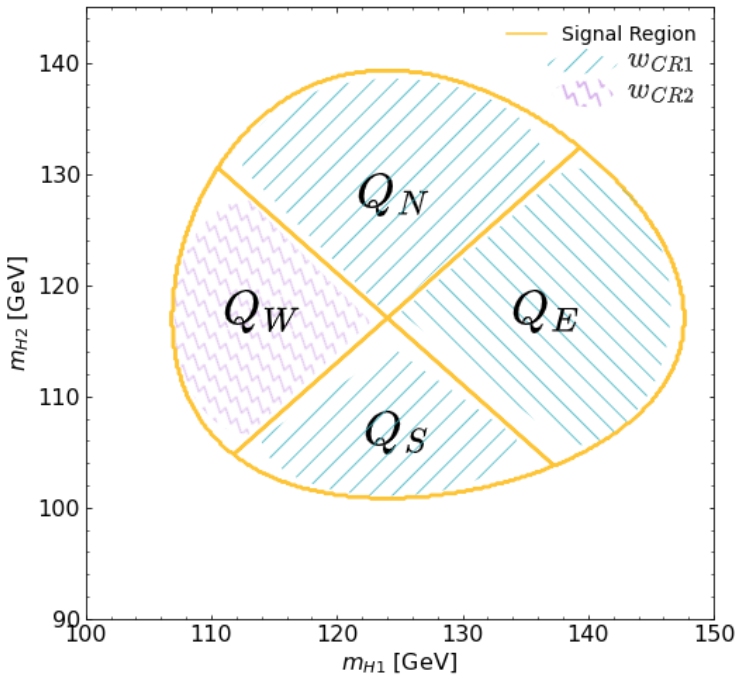


Figure 6.12: The SR split into quadrants is used to define the systematics of the CR12 non-closure from deriving an alternative background using weights derived in CR2. The example shown is how the CR12_shape_W NP is derived, the $w_{CR2}(\boldsymbol{x})$ are applied only on 2b events in the SR quadrant Q_W (purple zigzag region), while in all other quadrants, the 2b events are reweighted with $w_{CR1}(\boldsymbol{x})$ (turquoise hatched region).

- Symmetrise the alternative m_{HH} distribution over the nominal background estimation in each category

- Define the NP as the up and down variation as compared to the nominal m_{HH} background estimation.

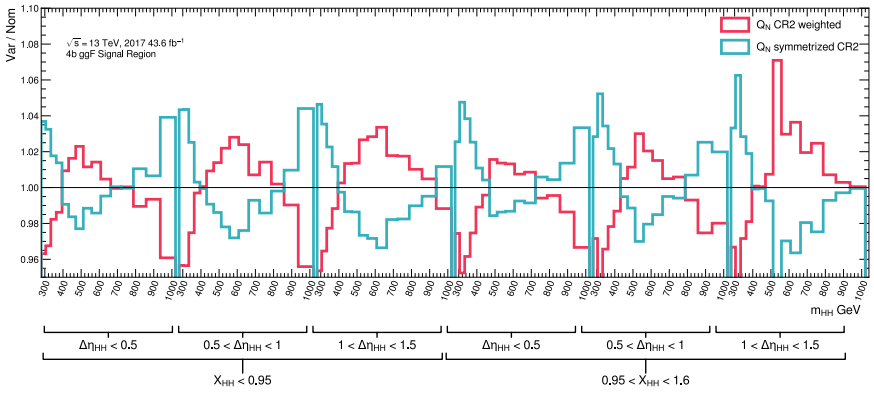
The above procedure is repeated in each quadrant and the given example is illustrated in Figure 6.12. The four nuisance parameters are derived per analysis category. An example of the four-point non-closure CR12 systematic is shown in Figure 6.13 for the ggF categories in the data taking year 2017. The strategy taken differs from the one used in the resonant analysis, which pioneered the background estimation technique. The motivation for the four-point rather than the simpler two-point systematic uncertainty is to give the fit more freedom and not wash out possible kinematic differences arising in the quadrants.

In the ggF channel, each NP is correlated across the categories but not years, while in the VBF channel, they are correlated across the two categories. The motivation to correlate the individual NPs across categories comes from a closer examination of Figure 6.13. For instance, the CR12_shape_W NP shown in Figure 6.13b exhibit the same shape across the ggF categories in the 2017 dataset. This hints that the CR12_shape_W NP captures the same systematic effect and the NP should be correlated accordingly. The plots of the CR12_shape_{N/W/S/E} NPs for the 2016 and 2018 data taking years in the ggF categories are shown in Appendix A and found to exhibit similar behaviour per NP.

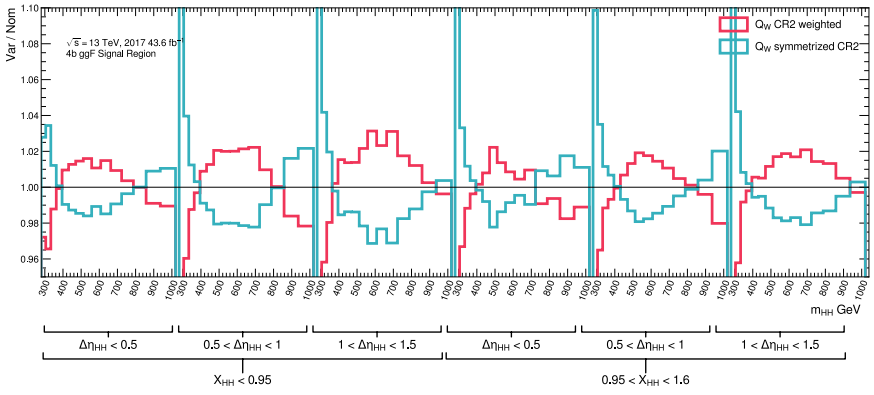
An additional systematic is assigned to the ggF channel, which comes from the non-closure in the 3b1f validation region described in Section 6.5.2. An uncertainty is assigned only to bins that have $> 1\sigma$ difference between the nominal and alternative background estimation taking into account all other uncertainties on the nominal background. To mitigate statistical fluctuations the systematics is defined by averaging over the $a\sigma - 1\sigma$, where $a\sigma$ is the variation in the bin with $> 1\sigma$ and its adjacent bins. The VBF regions are not assigned a non-closure 3b1f systematic as the agreement in the 3b1f validation VBF categories is observed to be good.

6.6.2 Experimental Uncertainties

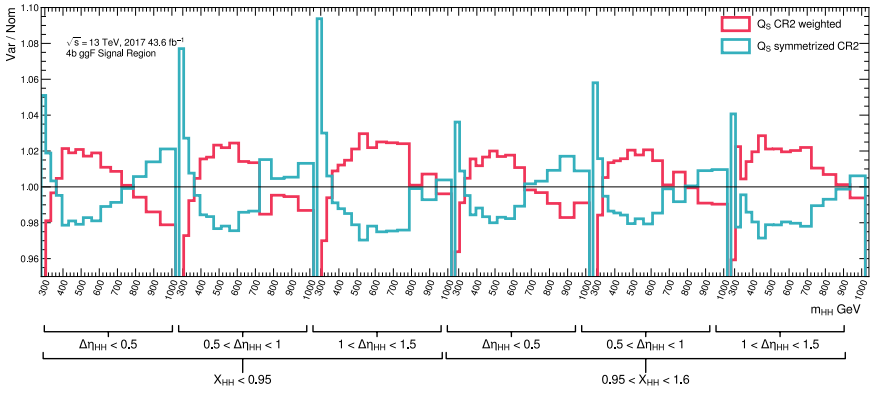
Detector and reconstruction uncertainties are taken into account by shifting parameters used in the jet reconstruction by their $\pm\sigma$ uncertainty obtained from auxiliary measurements. The whole analysis chain is then repeated with the physical objects defined by the shifted value. The up/down systematic uncertainty is taken as the envelope of the difference between the distribution obtained by the $\pm\sigma$ shift and the nominal. Jet energy scale and resolution as well as the uncertanties associated with the jet vertex tagger cut are varied as detailed in [80, 165], and briefly summarised in Chapter 5 as well. The uncertainty associated with the b-tagging



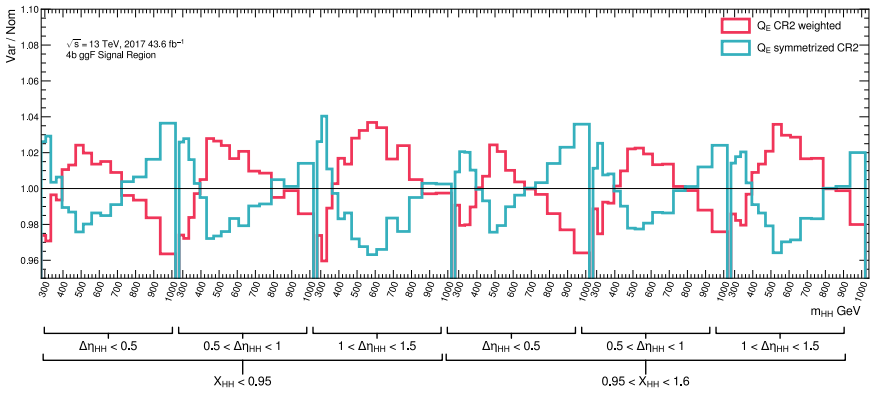
(a)



(b)



(c)



(d)

Figure 6.13: Non-closure systematics are derived from the alternative background estimation in each SR quadrant. The data taken in the year 2017 in the ggF channel is shown.

efficiency is calculated from event weight variations, covering the high p_T as well as the b-, c- and l-jet calibration uncertainties as several separate NPs, see [129] for the methodology. Similarly, trigger efficiency uncertainties are computed from event weight variations and accounted for as the uncertainty in the trigger SF calculation. The uncertainty on the simulated pile-up is accounted for as prescribed in [166]. Lastly, the preliminary uncertainty on the luminosity has a size of 1.7% for the three data taking years considered, where the primary luminosity determination for this dataset is done with the LUCID-2 detector [167, 168]. The experimental uncertainties are only applied to the signal samples as the background estimation is completely data-driven. They are correlated across categories as well as production modes and are subdominant, as compared to statistical error and background systematics.

6.6.3 Theoretical Uncertainties in diHiggs Signal

The uncertainties on the HH cross-section and $H \to b\bar{b}$ branching fraction are taken from theoretical calculations presented in [32, 140, 169] and are summarised in [170]. Notably, the uncertainty on the ggF production cross-section is quite large, the scale choice and finite top mass scheme alone result in an uncertainty of $^{+6\%}_{-23\%}$.

Theoretical Uncertainty	**Size [ggF Signal]**	**Size [VBF Signal]**
HH production cross-section	$^{+6\%}_{-23\%}$(scale+top mass) $+ \pm 3\%$(PDF+α_S)	$\pm 2.1\%$ & $^{+0.03\%}_{-0.04\%}$ (scale) $\pm 2.1\%$ (PDF+α_S)
Parton Shower	$10\% - 40\%$	
PDF and α_S	$\sim 1\% - 2\%$	
Initial state & final state radiation scale	$2\% - 6\%$	
$H \to b\bar{b}$ branching fraction	3.5%	

Table 6.5: The relative sizes of theoretical uncertainties on the ggF and VBF signals. The parton shower uncertainty is around 40% only in certain categories, where the contribution of the signal from the considered production mode is very small and therefore it only has a minor impact upon the final result.

The generators used to simulate the HH signal make a series of assumptions, from the hadronisation model to the PDF set, each with its systematic uncertainty. The

uncertainty on the parton shower and hadronisation model is estimated by repeating the full analysis chain for signal samples generated with Herwig instead of Pythia. The choice of the PDF set and α_S scale affects the matrix element calculation and its uncertainty is evaluated by varying each independently. The alternative PDF set is PDF4LHC_NLO_MC [143], while different α_S values are evaluated using the nominal PDF set. The two variations are combined to define the PDF and α_S uncertainty. Lastly, the renormalisation and factorisation scales used to evaluate PDFs are not unique. The uncertainty on the initial and final state radiation scale is evaluated by considering a 7-point variation, where the two scales are changed jointly or independently. The size of each theoretical uncertainty is stated in Table 6.5.

6.6.4 Summary of Standard Model HH uncertainties

Source of Uncertainty	Impact	Correlation
Background Uncertainties	**Applied the Background Estimation**	
Bootstrapping	Dominant	Treated as a statistical uncertainty
Non-closure CR12	Dominant	Across all categories but not between the ggF & VBF channel.
Non-closure 3b1f	Significant	
Theoretical Uncertainties	**Applied to the ggF & VBF HH Signals**	
HH production cross-section	Dominant	Across all categories and channels but not between the two production modes.
Parton Shower	Significant	
PDF and α_S	Significant	
Initial state & final state radiation (scale)	Significant	
$H \to b\bar{b}$ branching fraction	Significant	Across all categories, channels and the two production modes.
Experimental Uncertainties	**Applied to the ggF & VBF HH Signals**	
Luminosity	Subdominant	Across categories, channels, and production modes.
Jet energy scale and resolution		
Flavour tagging efficiency		
Trigger efficiency		
Pile-up		

Table 6.6: The sources of uncertainty applied when setting upper limits on the SM HH signal strength and constraining κ couplings. For each, its importance and correlation scheme are listed. Dominant impact implies the change in the observed signal strength of about 7-10%, significant of a few % and subdominant $\leq 1\%$.

All sources of uncertainty, their importance and correlation scheme that are accounted for when deriving the upper limits on the SM HH signal strength and constraining κ couplings are summarised in Table 6.6.

6.6.5 The Acceptance Reweighting EFT Uncertainty

The signal reweighting method performs well in the $\kappa -$ framework, good closure is observed across ggF and VBF signals when comparing the reweighted distributions to the simulated samples. Thus, no systematic uncertainty on the reweighting is applied. On the other hand, when comparing the simulated EFT signals to the reweighted signals non-closure is observed in the SR m_{HH} distributions. The source of the systematic uncertainty is the finite statistics used to derive the weights and the difference in acceptance between the SM and non-SM signals. In addition, the SMEFT reweighting is derived only at LO. The acceptance reweighting EFT uncertainty is derived at the truth level using parton showered signal samples and the SM sample, where the latter is reweighted to the considered signals. No detector simulation is run on the samples to spare CPU resources. To account for the difference in acceptance between the SM and non-SM signals, events must pass approximate ggF selection, described in Section 6.4, at the truth level. The SR is only a proxy as the lack of detector simulations means that approximate cuts are needed to mimic the trigger decision and truth information is used instead of the b-tagging algorithms.

For the HEFT interpretations, each of the HEFT benchmarks (BM) is simulated. As the c_{gghh} and c_{tthh} coefficients in HEFT, and the c_H, c_{tG}, $c_{H\Box}$, c_{tH} and c_{HG} coefficients in SMEFT are scanned independently at least two possible values of each coefficient are considered. In SMEFT the two-dimensional parameter space of c_H and one additional Wilson coefficient is explored. Therefore, samples, where c_H and another coefficient take non-SM values are generated as well. To be exact, for the HEFT c_{gghh}

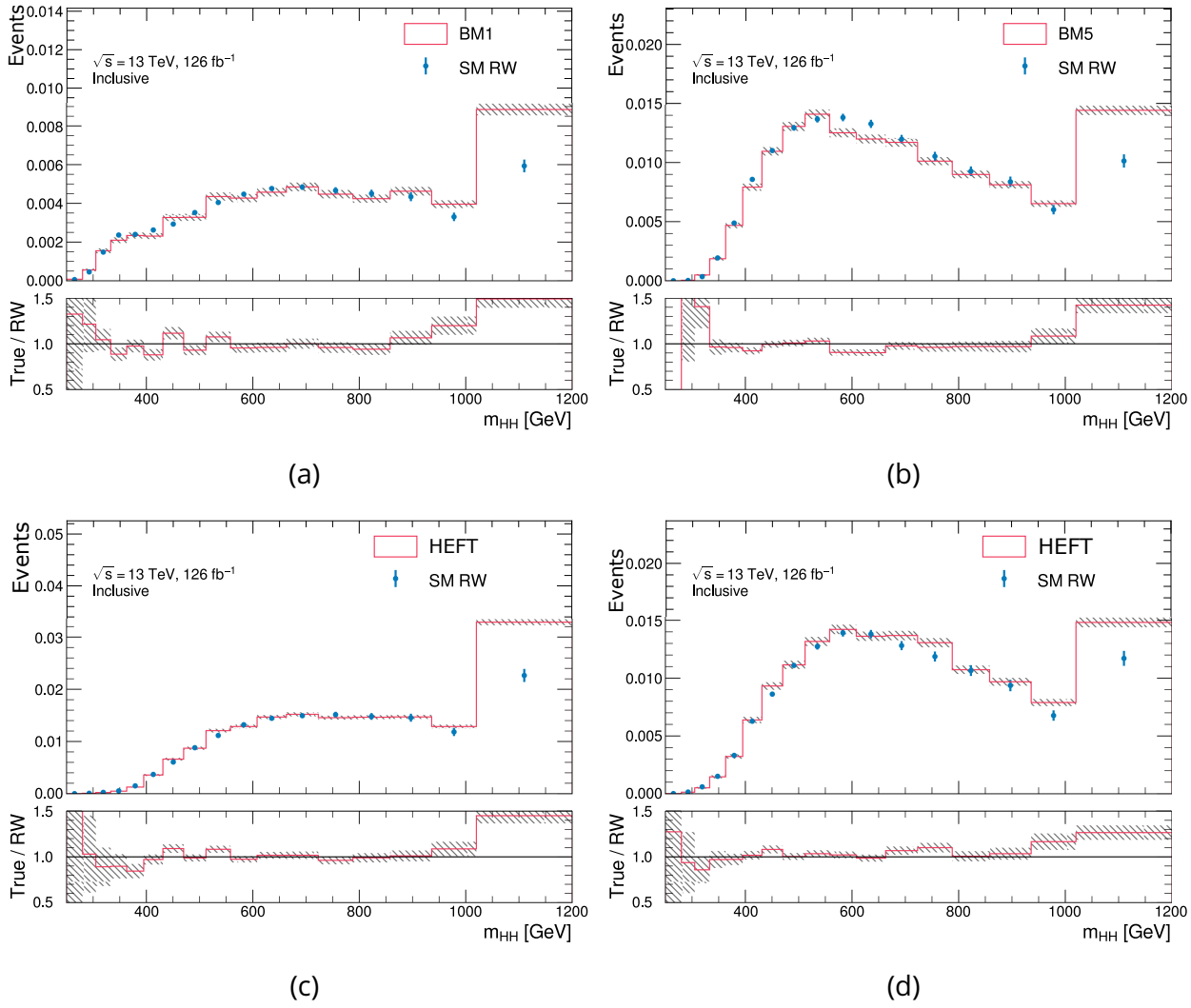


Figure 6.14: The simulated HEFT signal (red line) is compared to the SM sample reweighted to the examined variation (blue dots) in bins of m_{HH}. The HEFT variation shown are BM1 in (a), BM5 in (b), $c_{gghh} = 0.5$ in (c), and $c_{tthh} = 0.5$ in (d). Only the statistical error is shown. The overflow (last bin) and underflow (first bin) bins are shown for visualisation purposes (in the actual analysis they are merged into the highest m_{HH} bin) and an explanation for the discrepancy observed is offered in the text of the thesis.

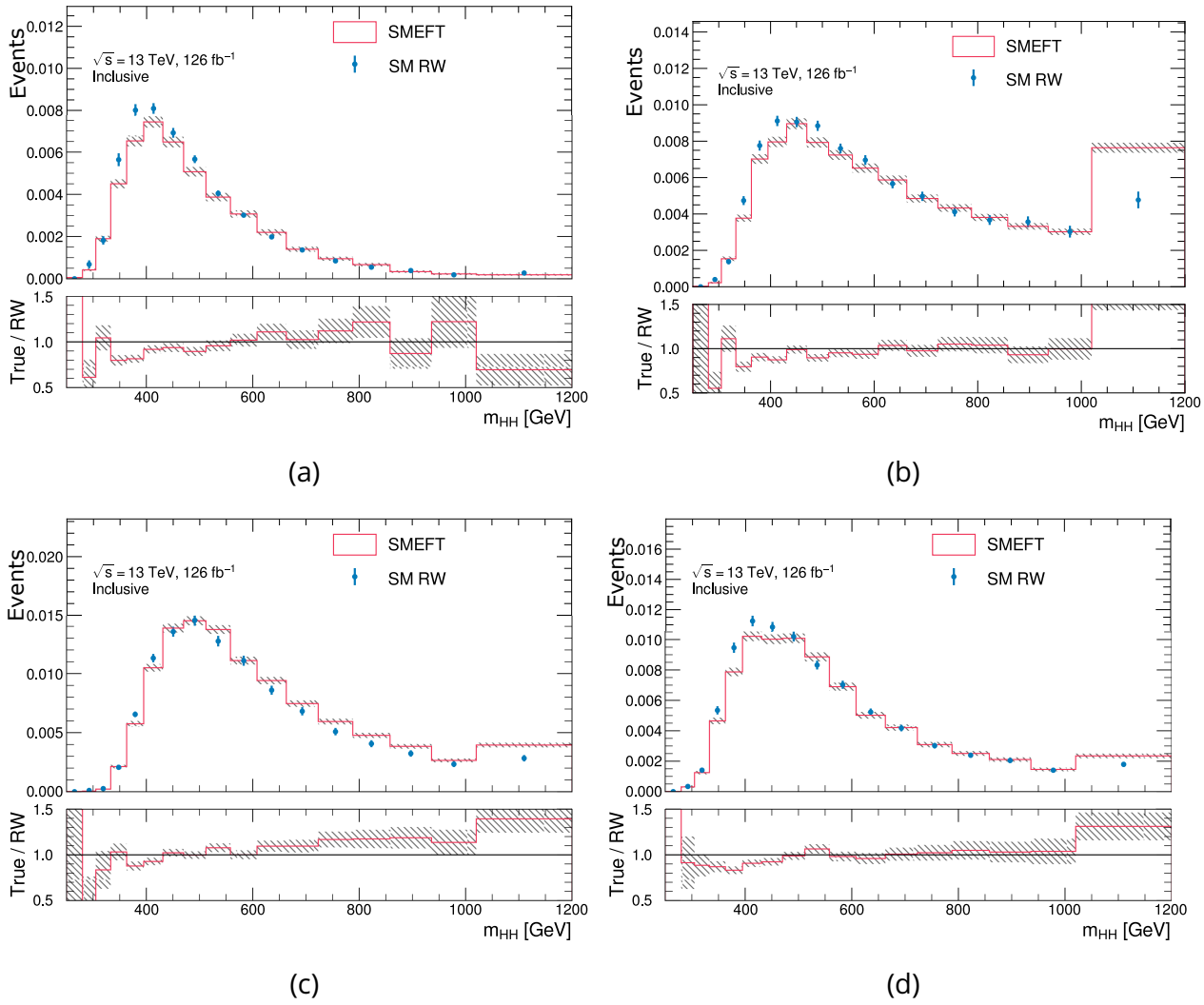


Figure 6.15: The simulated SMEFT signal (red line) is compared to the SM sample reweighted to the examined variation (blue dots) in bins of m_{HH}. The SMEFT variation shown are $c_H = 14.93$ in (a), $c_{tG} = -0.9$ in (b), $c_H = -2, c_{tH} = 1.5$ in (c), and $c_H = -2.5, c_{tG} = 0.15$ in (d). Only the statistical error is shown. As is the case for the HEFT signals as well the overflow (last bin) and underflow (first bin) exhibit the largest discrepancy.

scan the variations $c_{gghh} = 0.5$, $c_{gghh} = -0.5$ and $c_{gghh} = 1.0$ (similarly $c_{tthh} = 0.5$, $c_{tthh} = -0.5$ and $c_{tthh} = 1.0$ for the c_{tthh} scan) are evaluated. The simulated SMEFT signals are stated in Table 6.7 and cover both negative and positive values of the coefficients scanned. Approximately 100 000 events are simulated for each EFT signal, while 300 000 events are generated for the SM signal. The events that pass the proxy ggF selection are then used to derive the acceptance reweighting EFT uncertainty. The final SR categorisation is found to have no impact on the size of the mis-modelling, as the differences observed are due to statistical fluctuations. Therefore, the systematic uncertainty is derived in the inclusive SR. The reweighted SM distribution is compared to the corresponding simulated EFT signal in bins of m_{HH}, as shown in Figures 6.14 and 6.15.

The uncertainty up to the highest m_{HH} bin is found to be flat within the statistical uncertainty, while the highest m_{HH} has a significantly larger uncertainty in most EFT signals checked. Taking a closer look at Figures 6.14 and 6.15, it can be seen that the large discrepancy in the yield of the highest m_{HH} bin comes largely from the overflow bin and to a lesser extent from the underflow bin, which is shown. The reason for this is that the weights are derived only up to $1040\,\mathrm{GeV}$, if truth m_{HH} of a reweighted event is larger than that, the weight from the $1040\,\mathrm{GeV}$ bin is applied. Hence, a ”two-step” shape normalisation uncertainty is defined for the acceptance reweighting uncertainty on the EFT reweighting. A flat uncertainty which is the difference in the yield of all but the last m_{HH} bin is applied on the low m_{HH} distribution, while the highest m_{HH} bin has a dedicated uncertainty. The final uncertainties applied are defined in Tables 6.8 for the HEFT variations and 6.9 for the SMEFT variations. The differences between the simulated and reweighted distributions come about, due to the difference in acceptance, the range in which the weights for the SM sample are derived, and in the SMEFT case, the missing higher orders. The latter is reflected as a slightly higher uncertainty across the m_{HH} distribution. The acceptance reweighting EFT uncertainty has an impact of a few percent on the upper cross-section exclusion limit of an EFT signal.

Applicable SMEFT Scan	c_H	c_{tG}	$c_{H\Box}$	c_{tH}	c_{HG}	Used for final systematic?
c_H	-19.2	0	0	0	0	No
c_H	14.93	0	0	0	0	Yes
c_{tG}	0	-0.9	0	0	0	Yes
c_{tG}	0	0.9	0	0	0	No
$c_{H\Box}$	0	0	-8	0	0	No
$c_{H\Box}$	0	0	14	0	0	Yes
c_{tH}	0	0	0	-11	0	Yes
c_{tH}	0	0	0	5	0	No
c_{HG}	0	0	0	0	0.5	Yes
c_{HG}	0	0	0	0	-0.5	No
$c_H - c_{H\Box}$	-1	0	-1	0	0	No
$c_H - c_{H\Box}$	1	0	-1	0	0	Yes
$c_H - c_{tH}$	-2	0	0	-1.5	0	Yes
$c_H - c_{tH}$	2	0	0	-1.5	0	No
$c_H - c_{tG}$	-2.5	0.15	0	0	0	No
$c_H - c_{tG}$	2.5	0.15	0	0	0	Yes
$c_H - c_{HG}$	14.93	0	0	0	-0.5	No
$c_H - c_{HG}$	14.93	0	0	0	0.5	Yes

Table 6.7: The SMEFT variations considered to evaluate the reweighting acceptance uncertainty. Two variations per Wilson coefficient scan performed are considered, and the systematic is extracted from the variation with the larger difference.

HEFT variation	low m_{HH} bins	highest m_{HH} bin
BM1	0.99	0.72
BM2	0.99	0.62
BM3	0.99	0.74
BM4	0.98	0.9
BM5	0.97	0.77
BM6	0.97	0.8
BM7	0.99	0.81
$\mathbf{c_{gghh} = 0.5}$	1.0	0.75
$\mathbf{c_{tthh} = 0.5}$	0.97	0.81

Table 6.8: The differences between the reweighted and simulated HEFT signals, when compared in bins of m_{HH}. The quoted values are symmetrised to obtain the upper variation used in the fit. The quoted high m_{HH} difference is applied to the highest m_{HH} bin in the fitted distribution, which contains the overflow and underflow bins as well. Only the differences observed in $c_{tthh} = 0.5$ and $c_{tthh} = 0.5$ signals are quoted and used by the analysis, as the other variations checked yield a smaller uncertainty. In the case of $c_{gghh} = 0.5$ ($c_{tthh} = 0.5$), all other coefficients are fixed to their SM expected values.

SMEFT variation	**low m_{HH} bins**	**highest m_{HH} bin**
$\mathbf{c_H = 14.93}$	0.91	0.90
$\mathbf{c_{tG} = -0.9}$	0.94	0.73
$\mathbf{c_{H\Box} = 14.00}$	0.90	0.74
$\mathbf{c_{tH} = -11}$	0.91	0.77
$\mathbf{c_{HG} = 0.5}$	0.97	0.72
$\mathbf{c_H = -1, c_{H\Box} = -1}$	0.92	0.97
$\mathbf{c_H = -2, c_{tH} = 1.5}$	0.94	0.80
$\mathbf{c_H = -2.5, c_{tG} = 0.15}$	0.95	0.84
$\mathbf{c_H = 14.93, c_{HG} = 0.05}$	0.91	0.72

Table 6.9: The differences between the reweighted and simulated SMEFT signals, when compared in bins of m_{HH}. The quoted high m_{HH} difference is applied to the highest m_{HH} bin in the fitted distribution, which contains the overflow and underflow bins. Two variations per coefficient scans are used to validate the reweighting strategy, only the difference extracted from the largest variation is quoted. All but the quoted coefficients are fixed to their SM expected values.

6.7 Statistical Analysis

The final step of any analysis is to take the (processed) collected data, compare it to a theory and draw conclusions from the results. Usually, the aim is to either search for a new phenomenon or measure a physical quantity. The non-resonant $HH \to b\bar{b}b\bar{b}$ analysis is interested in both; searching for the SM HH production and constraining couplings involved in the processes.

The likelihood, which quantifies the agreement of data with a theory and provides the means to estimate unknown parameters of the theory itself is the key element of the statistical analysis independently of the final goal [171]. The likelihood function and the estimation of its parameters are outlined in Section 6.7.1. The building blocks of any hypothesis testing, the test statistic, and p-value, are introduced in Section 6.7.2. The profile likelihood ratio test statistic and profile likelihood ratio test used to set confidence intervals on the coupling modifiers ($\boldsymbol{\kappa}$) are outlined in Sections 6.7.3 and 6.7.4. The CLs method, which sets upper limits on the signal strength, is described in Section 6.7.5. The chapter concludes by outlining how to perform a combination of orthogonal channels and the description of Asimov data in Sections 6.7.6 and 6.7.7 respectively.

6.7.1 The Likelihood Function

To define the likelihood ($\mathcal{L}$), it is assumed that the observed data is fixed, the events are independent and the underlying theory can be described by a probability density function (PDF) that is parameterised by some unknown parameters. The unknown parameters are commonly denoted as the vector $\boldsymbol{\theta}$ and can be free or constrained by some auxiliary measurements. Additionally, it is usually the case that events are split into categories and presented as histograms of a discriminant variable that provides good background and signal separation. For example, the

$HH \to 4b$ uses the diHiggs invariant mass (m_{HH}) as the discriminant variable and the inputs for $\mathcal{L}$ are the observed data, background estimation, and signal histograms of each of the ggF and VBF categories. Hence, $\mathcal{L}$ can be written as:

$$\mathcal{L}(\boldsymbol{\Psi}, \boldsymbol{\theta}) = \prod_{c \in \text{categories}} \prod_{b \in \text{bins}} \mathcal{P}\langle\int(n_{c,b}|\lambda_{c,b}(\boldsymbol{\Psi}, \boldsymbol{\theta})) \cdot \prod_{\chi \in \boldsymbol{\chi}} \mathcal{C}_{\chi}(a_{\chi}|\chi), \tag{6.15}$$

where:

1. $\boldsymbol{\Psi}$ is the parameter of interest (POI). This can be the signal strength (μ) when the limits on the cross-section are set or the $\boldsymbol{\kappa}$ coupling modifier that changes the signal hypothesis in the fit. One or more coupling modifiers can be set as the POI.
2. $\mathcal{P}\langle\int$ is the joint probability function of the Poisson distribution. The Poisson distribution is an obvious choice when one wants to model the event rates per bin in an experiment.
3. $\boldsymbol{\theta}$ are the nuisance parameters (NPs), that encode the experimental, theoretical, and statistical uncertainties and any other unknown in the theory modelled on the expected rate.
4. $n_{c,b}$ is the total number of events measured in the observed dataset in bin b of category c.
5. $\lambda_{c,b}(\boldsymbol{\Psi}, \boldsymbol{\theta})$ is the expected total number of events in bin b of category c that depends on both the POI(s) and NPs. More precisely the rate is $\lambda_{c,b}(\boldsymbol{\Psi}, \boldsymbol{\theta}) = \prod_{s \in \text{sample}} \lambda_{s,c,b}(\boldsymbol{\Psi}, \boldsymbol{\theta})$, where $\lambda_{s,c,b}(\boldsymbol{\Psi}, \boldsymbol{\theta})$ is the rate of sample s in bin b of category c.
6. $\boldsymbol{\chi}$ are a subset of NPs that are constrained, i.e. their impact on the expected total number of events is limited in some way.
7. $\boldsymbol{a}$ is the auxiliary data, the global observable dataset which comes from independent measurements from another experiment or analysis. It is needed

to constrain $\boldsymbol{\chi}$ correctly. Together with the observed data it forms the measurements.

8. $\mathcal{C}_\chi(a_\chi|\chi)$ is the constraint term for the constrained NP χ and depends on the auxiliary data a_χ. The constraint term takes the form of a probability function of the Poisson distribution for statistical uncertainties, for all other systematic variations it is a Gaussian probability distribution.

Method of Maximum Likelihood

The best estimators of $\boldsymbol{\theta}$ can be obtained by maximising the likelihood, which in the large sample limit provides estimators that are efficient, constant, and unbiased [172]. For convenience, it is easier to take the derivatives of $\ln \mathcal{L}$ and solve the set of equations $\frac{\partial \ln \mathcal{L}}{\partial \theta_i}$, which give the same estimators of $\boldsymbol{\theta}$ as maximising $\mathcal{L}$. This procedure is often referred to as fitting, when one or more unknown parameters are fixed the fit is referred to as the profile likelihood fit. From now on $\hat{\square}$ represents the best estimators of a parameter in an unconditional maximum likelihood fit, while $\hat{\hat{\square}}$ denotes the best estimators of a parameter in a conditional (profiled) maximum likelihood fit.

Nuisance parameters

Unfortunately, it is not possible to have perfect knowledge of either the background or signal hypothesis. Thus, the predicted sample rate ($\lambda_{s,c,b}$) might not be the true rate. To capture the lack of knowledge, $\lambda_{s,c,b}$ in Equation 6.15 is parameterised in terms of nuisance parameters (NP), $\boldsymbol{\Psi}$ and $\boldsymbol{\theta}$. Note that the only difference between $\boldsymbol{\Psi}$ and $\boldsymbol{\theta}$ is that the post-fit value of $\boldsymbol{\Psi}$ is of special interest. The NPs are essentially the systematic uncertainties that could modify the expected rate. The NPs consist of free parameters and constrained parameters, the latter introduces an additional term to the likelihood, the constraint term for auxiliary measurements. This term

requires consistency with the auxiliary measurements by penalising variations in tension with auxiliary measurements and so limits the extent to which a systematic uncertainty can change the sample rate. Auxiliary measurements refer to some external measurements that inform us about the systematic uncertainties, for example, the b-tagging uncertainty has been determined in calibration, completely independent of the analysis considered [173].

The majority of systematic uncertainties applicable to the analysis are either normalisation uncertainties or shape uncertainties. A normalisation uncertainty is a multiplicative modification of the sample rate that takes the same value across all bins. It can be written as

$$\lambda_{c,b}(\alpha) = \lambda_{c,b} \times \begin{cases} \sigma_{+}^{|\alpha|} & \alpha > 0, \\ \sigma_{-}^{|\alpha|} & \alpha < 0, \\ 1 & \alpha = 0 \end{cases} \tag{6.16}$$

where α is the NP, nominally set to zero, and $\sigma_{+/-}(b)$ are the $\pm 1\sigma$ normalisation uncertainties [174]. The constrained term is the probability function of the Gaussian distribution centred at zero with a mean equal to α and one as the standard deviation. For instance, the uncertainty on the $H \to b\bar{b}$ branching fraction and the uncertainty on the HH production cross-section defined in the previous section are normalisation only uncertainties.

On the other hand, shape uncertainties have an additive effect on $\lambda_{c,b}$ and can be quantified as

$$\lambda_{c,b}(\alpha) = \lambda_{c,b} + |\alpha| \times \begin{cases} \sigma_{+}(b) + \lambda_{c,b} & \alpha \geq 0, \\ \sigma_{-}(b) - \lambda_{c,b} & \alpha < 0, \end{cases} \tag{6.17}$$

where the NP α again takes zero as the nominal value [174]. Here $\sigma_{+/-}(b)$ are the bin per bin $\pm 1\sigma$ systematic uncertainties implemented as (absolute) up/down

histograms of the variation. The constrained term is the same as for the normalisation uncertainty. An example, of a shape uncertainty, is the non-closure CR12 extrapolation uncertainty described in Section 6.6.1.

6.7.2 Test Statistic and p-value

At this point, it is useful to introduce the p-value and test statistic, which form the core of hypothesis testing. The test-statistic ($Q(\boldsymbol{x})$) is a scalar function of the observable $\boldsymbol{x}$, whose distribution, $f(Q(\boldsymbol{x})|H)$, is fixed by a hypothesis distribution, H, followed by $\boldsymbol{x}$ [175]. In the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis $\boldsymbol{x}$ is the m_{HH} distribution observed across all SR categories. From this point on $\boldsymbol{x}$ is omitted from the equations for clarity.

The incompatibility of data with a given H can be quantified with the p-value,

$$p_H = \int_{Q_H,obs}^{\infty} f(Q|H)dQ_H, \tag{6.18}$$

which defines how likely is it to observe a value of Q at least as extreme as the observed Q_{obs}. It is common to reject a hypothesis at a confidence level $\geq 1-\alpha$, provided $p_H \leq \alpha$, where α is a value of choice.

The distribution of Q is defined by H, which in the analysis discussed is the signal hypothesis. To be precise the HH signal parameterised by the signal strength (μ) or the coupling modifiers ($\boldsymbol{\kappa}$) on top of the background estimation is referred to as the signal hypothesis. For example, the $\mu = 0$ hypothesis is simply the background-only hypothesis, while the $\kappa_\lambda = 0$ hypothesis assumes the existence of the $\kappa_\lambda = 0$ HH signal on top of the background model. The signal strength (μ) or the coupling modifiers ($\boldsymbol{\kappa}$) are the parameter to be measured by the observed dataset.

6.7.3 The Profile Likelihood Ratio Test Statistic

The most powerful statistical test according to the Neyman-Pearson lemma [176] is the profile likelihood ratio

$$\Lambda = \frac{\mathcal{L}(\Psi, \hat{\hat{\theta}}(\Psi))}{\mathcal{L}(\hat{\Psi}, \hat{\theta})}, \tag{6.19}$$

where the numerator is the conditional maximum likelihood and the dominator is the unconditional maximum likelihood. Hence, $\hat{\hat{\theta}}$ are the best-fit values given a fixed Ψ, while $\hat{\Psi}$ and $\hat{\theta}$ are the best maximum likelihood estimators.

It is the convention to take the logarithm of 6.19 and multiply it by minus two to form

$$-2\ln\Lambda = -2\ln\frac{\mathcal{L}(\Psi, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\Psi}, \hat{\theta})}, \tag{6.20}$$

which is the basis of the test statistic for setting upper limits on the HH cross-section and finding the confidence intervals on $\boldsymbol{\kappa}$.

Example with Toy Experiments

An important question to answer when designing an analysis is given the observed dataset follows H_0 is it possible to exclude an alternative hypothesis H_1.

Let's say that H_0 is the SM, so $\kappa_\lambda = 1.0$ HH signal and H_1 is the $\kappa_\lambda = 6.0$ signal hypothesis. As discussed in Section 6.7.2, the p-value is a good measure to assess the compatibility between a dataset that follows H_0 and H_1. In this example, $H(\kappa_\lambda = -6.0)$ is rejected if $p_H \leq 0.05$, so a confidence level of at least 95% is required. The results shown are based on a preliminary version of the ggF channel.

Choosing two times the negative log-likelihood ratio as the test statistic, Equation

6.20 in this example becomes

$$Q = -2\ln\Lambda = \frac{\mathcal{L}(\kappa_\lambda = -6.0, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\kappa_\lambda}, \hat{\theta})}, \tag{6.21}$$

where $\hat{\kappa}_\lambda$ is the signal hypothesis ($H(\kappa_\lambda)$) that best fits the observed dataset. As in the $HH \to b\bar{b}b\bar{b}$ analysis the final observable is the binned m_{HH} distribution, the likelihood function in Equation 6.15 and therefore $-2\ln\Lambda$ are functions of the m_{HH} distribution.

To compute the p-value both, the distribution of $-2\ln\Lambda$ denoted as $f(Q|H_1)$ and the observed test statistic, need to be determined. This can be always done by running toy experiments, as follows;

1. The PDF for H_1 is constructed, i.e. the assessed model is the m_{HH} distribution of the $\kappa_\lambda = -6.0$ signal on top of the background distribution as expected in the ggF channel. In this simplified example only the statistical uncertainties and the non-closure CR12 extrapolation uncertainty are accounted for.

2. A toy dataset of the m_{HH} distribution is drawn from the PDF.

3. $-2\ln\Lambda$ is calculated given the toy m_{HH} distribution.

Figure 6.16a shows $f(Q|H_1)$ generated by repeating steps 2-3 about 850 times as the turquoise histogram. In an analysis usually, thousands of toy datasets are generated, as the more toy datasets are drawn from the PDF the closer the produced Q distribution resembles the true $f(Q|H_1)$.

The expected observed test statistic (Q_{obs}) is defined as the median of the observed test statistic distribution ($f(Q|H_0)$) [175]. The later is generated by the same procedure as described above, except that the toy datasets are sampled from a PDF that represents H_0. The purple histogram in Figure 6.16a shows $f(Q|H_0)$ obtained from about 850 toy datasets, while the Q_{obs} is plotted as the dashed dark pink line.

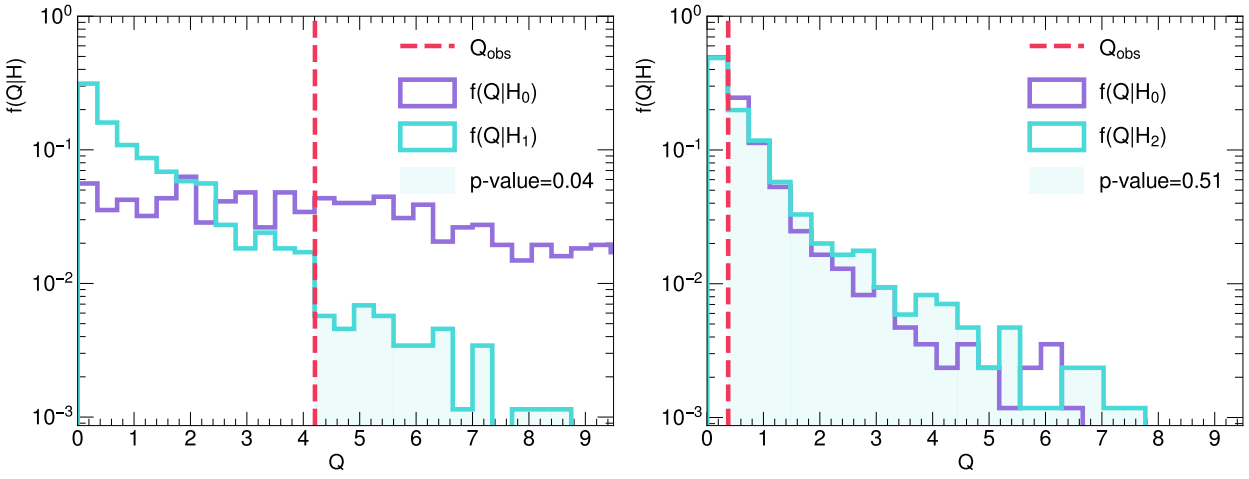
It is clear from the small overlap between $f(Q|H_0)$ and $f(Q|H_1)$ in Figure 6.16a that the two hypotheses are not compatible. To be exact, the p-value is 0.04, illustrated as the shaded turquoise area in the plot. Hence, the H_1 is expected to be excluded at a confidence level greater than 95% given that the observed dataset follows H_0. In other words, it is possible to separate between the m_{HH} distribution that contains the $\kappa_\lambda = 1$ HH signal and one that contains the $\kappa_\lambda = -6$ HH signal on the top of the background in the ggF $HH \to b\bar{b}b\bar{b}$ analysis. On the other hand, the ggF channel studied is not sensitive to a dataset that contains the $\kappa_\lambda = 0$ hypothesis, provided the expected observed dataset follows the SM hypothesis. It can be seen in Figure 6.16b that the distributions of the two test statistics are nearly identical with a p-value of 0.51.

By repeating the above-described procedure for several κ_λ hypotheses the 95% confidence interval on κ_λ can be deduced. The interval is bounded by the κ_λ values at which the p-value is 0.05. The optimisation of the analysis is driven by narrowing down the 95% confidence interval so that κ_λ is constrained as tightly as possible.

In the example above only the expected observed test statistic is shown, as at the time of the study, the $HH \to b\bar{b}b\bar{b}$ analysis was blinded[7]. Once the analysis is unblinded the observed test statistic can be calculated simply by computing $-2\ln\Lambda$ in Equation 6.21 given the observed m_{HH} distribution. If the expected and observed 95% confidence interval on κ_λ are found to be notably different, it may indicate that the observed data does not follow the SM hypothesis or potential issues with the background modelling, among other possible factors.

The final results presented in this chapter have not been produced with running toy experiments as the asymptotic approximation (discussed next) was found to be valid. However, toys were used to identify an issue with the definition of the uncertainty on the reweighted HH signals in the statistical analysis.

[7] In collider physics the observed dataset is not used to optimise and validate the analysis strategy to avoid introducing any bias. Instead expected (Asimov) datasets are used.



(a) Testing the $\kappa_\lambda = -6.0$ hypothesis.

(b) Testing the $\kappa_\lambda = 0.0$ hypothesis.

Figure 6.16: The computation of the p-value, assuming the $\kappa_\lambda = -6$ hypothesis (H_1) in (a) and the $\kappa_\lambda = 0$ hypothesis (H_2) in (b) given the expected dataset follows the SM hypothesis. The turquoise histogram is the distribution of the test statistic $f(Q|H_1)$ ($f(Q|H_2)$), while the purple histogram is the distribution of the expected observed test statistic $f(Q|H_0)$. It should be noted that for simplicity the the test statistic is denoted by Q in both plots but in fact $Q = -2\ln(\Lambda(\kappa_\lambda = -6))$ in (a) and $Q = -2\ln(\Lambda(\kappa_\lambda = 0))$ in (b). The expected observed test statistic is defined as the median of the sampled test statistic and marked as the pink dashed line on the plots. The shaded turquoise area is the p-value given the expected data generated assuming the SM case and testing the alternative hypothesis. Each of the plotted distributions is generated by sampling approximately 850 toy datasets (m_{HH} distributions) from the PDF of the hypothesis considered. The background estimation and the κ_λ signal model as defined in the ggF channel of the $HH \to b\bar{b}b\bar{b}$ analysis are used to define the PDF of a hypothesis. Only the statistical uncertainties and dominant uncertainties on the background estimation are accounted for. The same set of toy datasets is used to compute $f(Q|H_0)$ in (a) and (b), while Q in $f(Q|H_0)$ is distinct.

The Asymptotic Approximation

The computation of the test statistic by generating toys, demonstrated in Figure 6.16, is computationally expensive and time-consuming. In this thesis, the asymptotic approximation can be used, which enables the analytic expression for the test statistic to be determined. The approximation is based on results from Wilk [177] and Wald [178], which show that in the limit of large statistics the test statistic $-2\ln\Lambda$ follows a non-central χ^2 distribution with the number of free parameters equal to the number of POIs. Another property that follows from the asymptotic approximation is that given a region can be identified, where the probability to make a measurement is less or equal to an arbitrary value α it is bound by a surface of constant $-2\ln\Lambda$ [172]. The latter is used to extract confidence regions on $\boldsymbol{\kappa}$.

6.7.4 The $-2\ln\Lambda$ Test

The coupling modifiers are measured through the unconditional maximum likelihood fit, where $\boldsymbol{\kappa}$ are set as the POIs. Hence, the measured $\hat{\boldsymbol{\kappa}}$ given the observables are the best estimators found to maximise the likelihood. All that is left to do is determine the uncertainty on $\hat{\boldsymbol{\kappa}}$. It is the convention to refer to the uncertainty as the coverage probability, which is extracted from the profile likelihood ratio test statistic ($-2\ln\Lambda(\boldsymbol{\kappa})$) with $\boldsymbol{\kappa}$ as the POIs.

As the large number limit can be taken, $\boldsymbol{\kappa}$ are expected to follow a multivariant Gaussian distribution, and $-2\ln\Lambda(\boldsymbol{\kappa})$ a non-central χ^2 distribution [171, 172]. Hence, the coverage probability can be defined by a constant value of the test statistic

$$C(1-\alpha) = -2\ln\Lambda(\boldsymbol{\kappa}), \tag{6.22}$$

where α is the probability to measure a value at least as extreme as $\boldsymbol{\kappa}$ [172, 175]. To be exact the interval (a region in case of multiple POIs) bound by $C(1-\alpha)$ contains

the true value of $\boldsymbol{\kappa}$ in $(1-\alpha)$% of measurements performed. In the large sample limit, this holds for non-Gaussian likelihoods and is a good approximation even for finite samples given the coverage probability only weakly depends on the true values of the POIs [172].

The confidence region with a 2σ coverage probability is determined by $C(1-\alpha) = -2\Delta \ln L \leq 4.0$ and $C(1-\alpha) = -2\Delta \ln L \leq 6.18$ for one and two POIs respectively [172]. The bounds of the confidence interval (region) are usually found by brute force. Several values of the POI(s) are scanned to determine the shape of the $-2\ln\Lambda(\boldsymbol{\kappa}) = C(1-\alpha)$ by interpolation. This can result in asymmetric confidence intervals (regions) around the best estimator(s). Figure 6.17 shows the expected confidence region with a 2σ coverage probability that bounds the $\kappa_\lambda - \kappa_{2V}$ parameter space as extracted from the $HH \to b\bar{b}b\bar{b}$ ggF, VBF and combined ggF and VBF channels. Each contour captures $\boldsymbol{\kappa} = (\kappa_\lambda, \kappa_{2V})$ values that result in $-2\Delta \ln L = 6.18$.

6.7.5 CLs Technique

The CLs technique is widely used in HEP to establish the discovery of a new signal as well as set an exclusion interval on the cross-sections of a process. For example, this diHiggs analysis is used to set upper limits on the SM diHiggs signal strength.

Let's first define the null hypothesis (H_0) as the background prediction only case assuming no SM HH production at all. Given some observables the confidence level of H_0 (CL_b) is given as

$$CL_b = P_b(Q \leq Q_{obs}) = \int_{-\infty}^{Q_{obs}} f(Q|H_0)dQ, \tag{6.23}$$

where Q is the test statistic, Q_{obs} is the observed test statistic and $f(Q|H_0)$ is the probability density function of the test statistic assuming H_0.

The alternative hypothesis (H_1) can then be defined as the background prediction

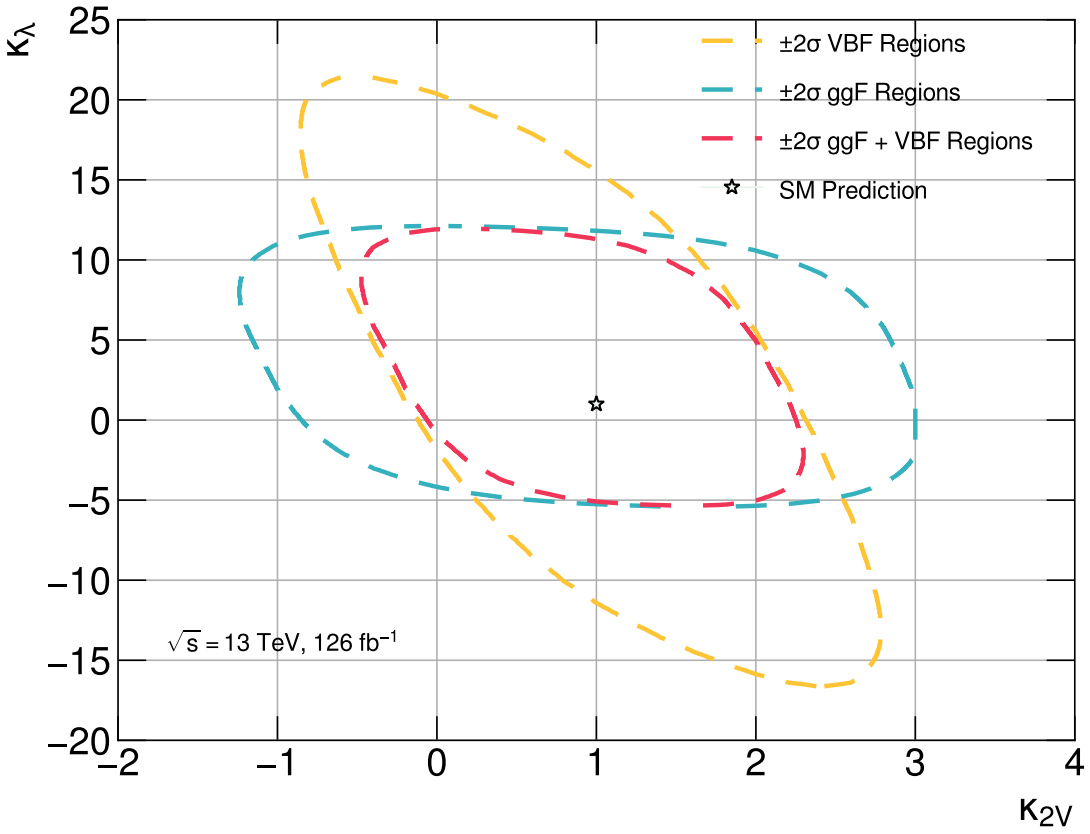


Figure 6.17: The expected 2σ confidence regions deducted from the $HH \to b\bar{b}b\bar{b}$ ggF (turquoise), VBF (yellow), and combined ggF and VBF (dark pink) channels. The contour represents the $\boldsymbol{\kappa} = (\kappa_\lambda, \kappa_{2V})$ values of constant $-2\ln\Lambda$, where $-2\Delta\ln L = 6.18$ as the bounded region has a 2σ coverage probability to capture the true $(\kappa_\lambda, \kappa_{2V})$ values. The sensitivity to κ_{2V} in the ggF channel comes from the VBF signal that leaks into the region, whose contribution can be significant as κ_{2V} varies (the $\kappa_{2V} = 0$ signal is expected to have a contribution of 23 events to the ggF signal region as evident from Table 6.3). The 2σ confidence regions are derived from the expected Asimov dataset and only the statistical and background uncertainties are considered.

and SM HH production. Similarly, as for the H_0 the confidence of H_1 given the observables (CL_{s+b}) can be written as

$$CL_{s+b} = P_{s+b}(Q \leq Q_{obs}) = \int_{-\infty}^{Q_{obs}} f(Q|H_1)dQ, \tag{6.24}$$

where $f(Q|H_1)$ is the probability density function of Q given H_1.

Q is chosen to be a monotonically increasing function so that values of CL_b close to one mean that the observed data is not background-like and the alternative hypothesis is favoured, while small values of CL_{s+b} disfavour the alternative hypothesis, indicating the data is more compatible with the null hypothesis [172].

In principle, CL_{s+b} should be enough to either discover or exclude a new signal, yet when the distributions of the test statistic for H_0 and H_1 are not well separated this method can lead to a large false exclusion rate. This is the case in diHiggs searches, as a small HH signal in comparison to a large background is expected. To avoid this the CLs technique is introduced, which essentially re-normalises the CL_{s+b} as

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_\mu}{1-p_b} = \frac{P_{s+b}(Q \leq Q_{obs})}{P_b(Q \leq Q_{obs})}, \tag{6.25}$$

where p_μ is the p-value of H_1 given a signal strength μ and p_b is the p-value of the H_0.

While this leads to more conservative limits, it mitigates the false exclusion rate and prevents unphysical results from a downward fluctuation of the background in those cases where there is only little sensitivity to the signal in the data.

The Q based on the likelihood ratio

$$q_\mu = \begin{cases} -2\ln\left(\frac{\mathcal{L}(\mu,\hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu},\hat{\theta})}\right) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu, \end{cases} \tag{6.26}$$

is commonly used for setting upper limits and is utilised in this thesis [175].

Equipped with the test statistic q_μ and CLs as the hypothesis test, the upper limit on the signal strength at 95% confidence level can be found by calculating the CLs for several μ and subsequently using interpolation to the value of μ for which $CLs = 0.05$[173]. It is the convention to report the upper limit given at 95% confidence level, but one might as well choose a different value to reject the signal strength. The CLs can also be obtained using toy experiments, however, as the analysis is in the large sample limit the asymptotic formula can be used to approximate the distribution of the q_μ, see [175] for details.

6.7.6 Combining Measurements from Different Channels

The statistical combination of different channels is done by multiplying the likelihood functions by each channel to be included. The combination can be viewed as adding more bins to Equation 6.15 [172]. The resultant combined likelihood function captures the observed data, and NPs from each analysis and can be used to perform a simultaneous fit of all the channels. The only assumptions made for the described procedure to be valid are that the channels entering the combination need to be orthogonal or the number of shared events negligible and the POIs have a common definition [179]. Special care must be taken in the treatment of NPs as well, for example, correlations of the uncertainties that are common between the analyses need to be considered [180].

The power of performing a combination is illustrated in Figure 6.17, which compares the expected $\kappa_\lambda - \kappa_{2V}$ 2σ confidence regions extracted from fitting the ggF and VBF channels independently and after their combination in the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis. The ggF-only channel (turquoise line) places stringent constraints on κ_λ, while κ_{2V} is only roughly bound compared to the VBF-only channel (yellow line). As the ggF and VBF channels are orthogonal by design and POIs (κ_λ and κ_{2V}) have the same definition, the likelihood functions of the two channels are multiplied to

construct the combined likelihood. Simultaneously fitting ggF and VBF channels result in a more tightly constrained 2σ confidence regions (dark pink line) across the $\kappa_\lambda - \kappa_{2V}$ parameter space as compared to the individual channels.

6.7.7 Asimov Data

Given a hypothesis, it is useful to know what are the confidence intervals of a measurement. The expected confidence intervals can be calculated using expected data, also referred to as the Asimov dataset. They are useful for optimising the analysis by pushing for the highest possible expected significance that can be reached and also serve as a cross-check for the unblinded measurements.

The Asimov dataset is constructed in such a way that the true parameters can be correctly estimated by running a profiled likelihood fit [175]. Depending on the task at hand it can take different definitions, which is the case here. The background-only model is taken as the Asimov data for when cross-section limits are set within the SM HH production. However, when finding the confidence intervals on $\boldsymbol{\kappa}$, the Asimov data is taken to be the background estimation with the SM signal on top.

6.8 Results

The analysis strategy is developed with the data in all signal regions (SR) blinded. The results presented in this section are the final limits derived from the unblinded dataset and have been published in [2].

A slight excess in the observed data as compared to the pre-fit background model is found at low m_{HH} in each ggF SR across all categories. However, when the background-only hypothesis is fitted there is a good post-fit agreement between the background estimation and the observed data, as seen in Figure 6.18, in the

absence of any HH signal. It follows that the observed excess is largely covered by the non-closure CR12 background estimation uncertainties (CR12_shape_{N/W/S/E}_{year}) and partially by the 3b1f non-closure (NC_3b1f) uncertainties, which is evident from the good post-fit agreement seen Figure 6.18 that results from pulling those systematic by about $\gtrsim 1\sigma$. The excess is not consistent with any of the non-resonant signals considered, as none of the HH signals considered on top of the pre-fit background estimation alone can explain the excess observed. Even when allowing the normalisation of the HH signal to freely float in the fit the pull on the non-closure background systematics lessens only slightly, which can be seen in the ranking plot, Figure 6.20 showing the forty NPs with the biggest impact in the fit. In addition, the signal injection test disfavours the presence of an actual signal in the data as well.

Given that the pull on the background systematics is of similar size independently of the signal hypothesis considered and there being good post-fit agreement once the systematics are accounted for, the slight excess is likely caused by a mis-modelling of the turn-on peak around $m_{HH} \sim 350\,\mathrm{GeV}$ in the background estimation. That the modelling of the turn-on peak in the ggF categories, where the excess is observed, comes with a large systematic uncertainty is evident from the CR12 non-closure background estimation uncertainties shown in Figure 6.13. Each NP (CR12_shape_{N/W/S/E}_{year}) captures a slightly different uncertainty of the shape of the peak, which gives the fit enough freedom to correct the mis-modelling of the background estimation. The VBF channel was evaluated to not have sufficient statistics to accurately model the turn-on peak before unblinding and therefore only $m_{HH} > 400\,\mathrm{GeV}$ is considered in the final fit. In the VBF SR categories, good pre-fit and post-fit agreement between the observed data and background estimation is observed. The post-fit m_{HH} distribution in the $|\eta| < 1.5$ VBF category is shown in Figure 6.19a and the $|\eta| > 1.5$ VBF category in Figure 6.19b. There is no significant pull on any of the systematic uncertainties applied in the VBF channels, as is evident from Figure 6.20. The ranking plot only shows the forty systematics with the highest

impact, as derived from fitting the ggF and VBF channels with the HH signal strength as the parameter of interest (POI). Appendix B shows the pulls and impacts of the subdominant systematics.

The upper limits on the SM signal strength ($\mu = \frac{\sigma^{measured}}{\sigma^{SM}}$) are extracted for the ggF, VBF, and ggF+VBF HH production with the CLs method using the asymptotic formula. The VBF (ggF) signal is treated as a background process, when extracting μ_{ggF} (μ_{VBF}), and is constrained in the fit only by the systematic uncertainties. The expected signal strength of the ggF+VBF HH production is extracted by fixing the relative strength between the ggF and VBF signals to the SM ratio. The observed (expected) upper limit on the SM HH signals strength is $\mu_{ggF+VBF} = 5.4(8.1)$ when all systematic uncertainties are accounted for including the uncertainties on the SM cross-section. The theoretical uncertainty on the cross-section and the systematics on the background estimation have the largest impact on the final limit, as seen in the ranking plot in Figure 6.20. The expected limit on $\mu_{ggF+VBF}$, when only the statistical error on the 4b data and reweighted 2b data is included, it is 6.0. The limits deteriorate by about 18% to $\mu_{ggF+VBF} = 7.1$ when accounting for all systematic uncertainties on the background estimation. Additionally, including all systematic uncertainties on the SM signals the sensitivity drops to $\mu_{ggF+VBF} = 8.1$ due to the significant size of the theoretical uncertainties. The theoretical uncertainty on the cross-section from the finite top mass assumption of the ggF production has the largest impact on the final limit. The observed and expected upper limits on the ggF, VBF, and ggF+VBF signal strengths are summarised in Figure 6.21, the VBF result is scaled down by a factor of ten for easier visualisation.

6.8.1 $\kappa-$ framework Interpretations

The κ_λ only interpretation is examined, which assumes that new physics in the electroweak sector has a detectable impact on the Higgs self-coupling but not the

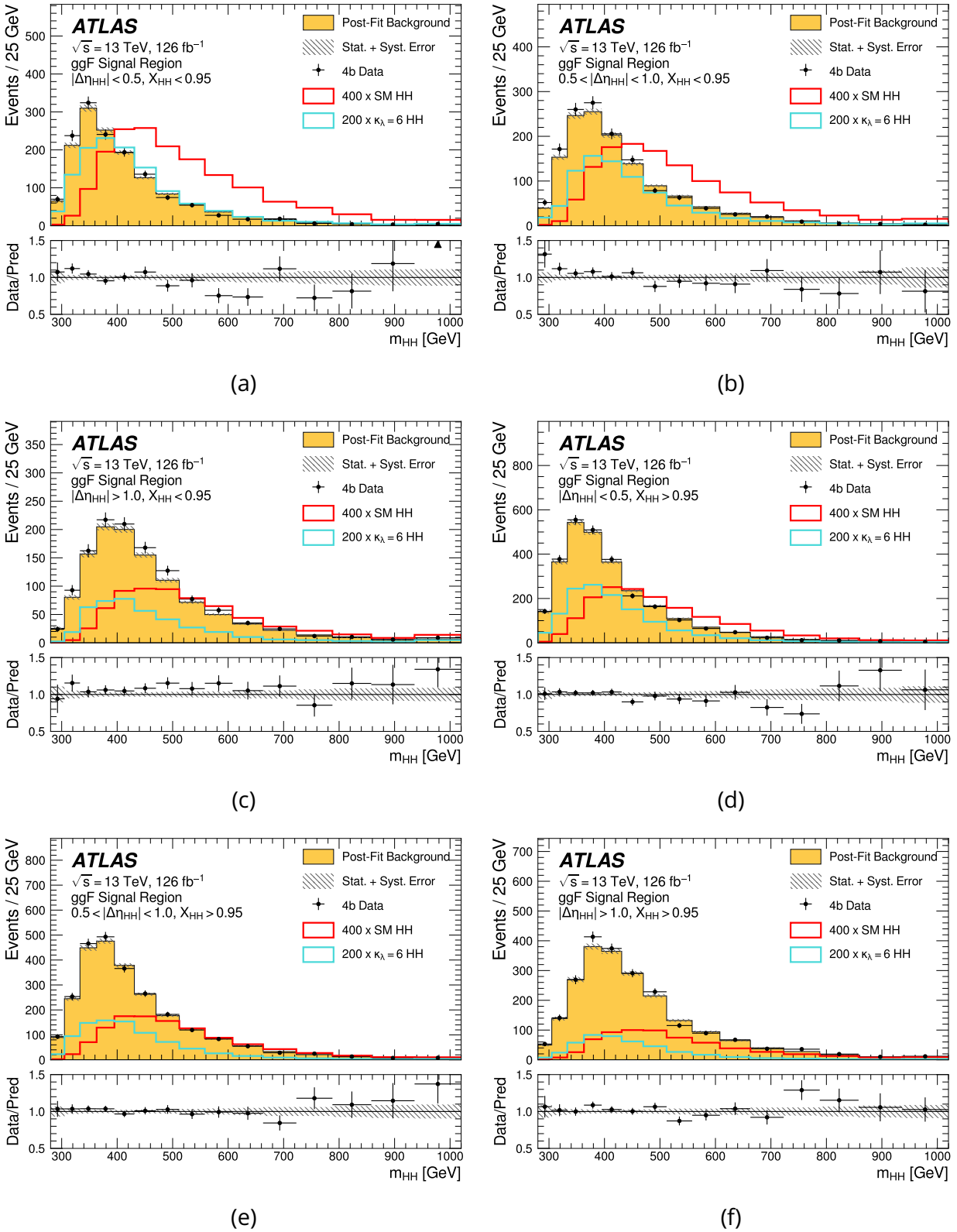
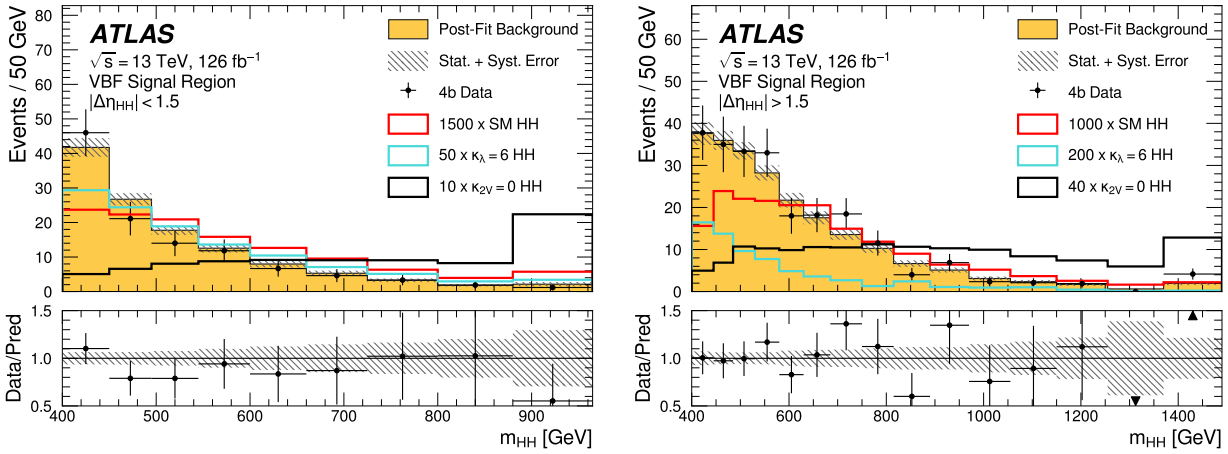


Figure 6.18: The data and post-fit background estimation, with the background-only hypothesis fitted, agree in all considered categories. Each year is considered to be a separate category in the fit, but for visualisation purposes, all the years are combined in the plots per kinematic category in $|\eta_{HH}|$ and $|X_{HH}|$. The plot is overlaid with the scaled SM and $\kappa_\lambda = 6$ HH signals for greater clarity. The plots are published in [2].



(a) (b)

Figure 6.19: Good agreement between data and the post-fit background estimation is observed in both VBF categories, the $|\eta_{HH}| < 1.5$ category shown in (a) and $|\eta_{HH}| > 1.5$ category in (b). The post-fit background estimation is shown for the profile likelihood fit assuming the background-only hypothesis. The expected SM, $\kappa_\lambda = 6$, and $\kappa_{2V} = 0$ HH signals are overlaid on the plot and scaled for better visualisation. The two plots have been published in [2].

other SM couplings. The $-2\Delta\ln(L)$ test statistic is used to extract the allowed $\pm 2\sigma$ confidence range on κ_λ, as shown in Figure 6.22a. The expected measurement assumes the existence of SM HH production. The maximum likelihood fit is performed on the combined ggF and VBF channels and has all couplings fixed to their SM expectations except for κ_λ. The observed (expected) measurement of κ_λ constrains it to the interval $-3.5 < \kappa_\lambda < 11.3$ ($-5.4 < \kappa_\lambda < 11.4$) with a $\pm 2\sigma$ confidence level. The best-fit value is moved away from unity to $\kappa_\lambda = 6.2$ (Figures 6.18 and 6.19 show the $\kappa_\lambda = 6.0$ signal against the post-fit background). The non-SM best-fit is driven by the observed data in the ggF channels, which has a small excess at low m_{HH} about where the $\kappa_\lambda = 6.2$ signal peaks. The excess has been extensively checked and found to be due to the mis-modelling of the turn-on in the peak in the background estimation.

Similarly, the κ_{2V} only interpretation is considered, where instead of κ_λ new physics is expected to be manifested only through modifications of the HHVV vertex. As deduced from Figure 6.22b the observed (expected) $\pm 2\sigma$ confidence level interval is

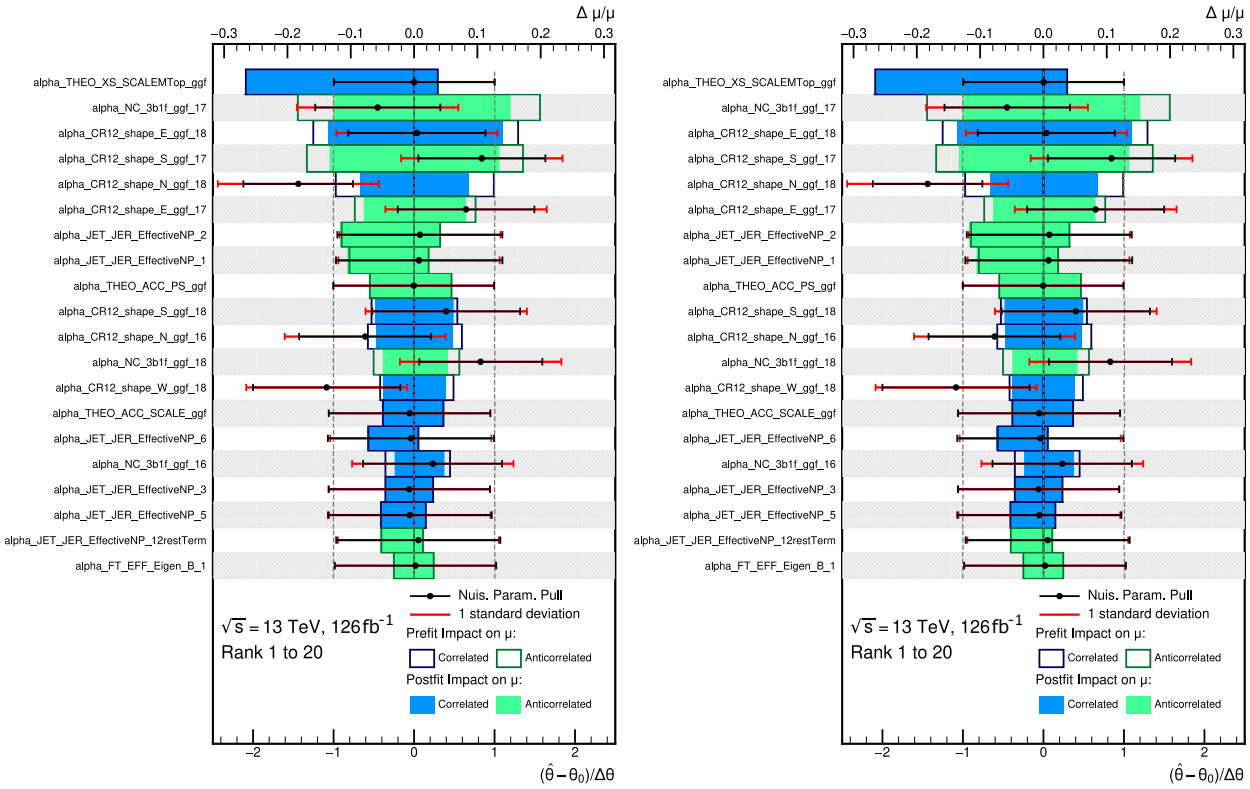


Figure 6.20: The pulls on the nuisance parameters (NPs) and impacts as extracted from the fit with $\mu_{ggF+VBF}$, as the POI. To pre-fit impact (green or blue rectangle) of a nuisance parameter (θ) is defined as the change in the POI when the systematic under question is fixed to $\theta \pm \sigma_\theta$ in the likelihood fit. Similarly, the post-fit impact (blue or green shaded area) is computed, only that θ is fixed to $\hat{\theta} \pm \hat{\sigma}$, where the hat indicates the values that maximise the likelihood in the unconstrained fit. The procedure is repeated for each nuisance parameter in the fit. The plot shown is called a ranking plot as it shows the nuisance parameters ranked by the size of their relative post-fit impact on the POI. The upper x-axis indicated the size of the relative impact $\frac{\Delta\mu}{\mu}$, where $\mu = \mu_{ggF+VBF}$ and $\Delta\mu$ is the change in $\mu_{ggF+VBF}$ when the NP examined is fixed in the fit. The lower x-axis shows the size of the pull on an NP ($\frac{\hat{\theta}-\theta_0}{\Delta\theta}$), $\hat{\theta}$ is the best-fit value taken by the NP, θ_0 is the pre-fit value and $\Delta\theta$ is the uncertainty on θ_0. The red line indicates the unconstraint (1σ) NP. Only the 40 highest-ranked nuisance parameters are shown. The NPs representing the non-closure systematic estimated from the alternative background estimation are labelled as alpha_CR12_shape_{quadrant}_{channel}_{year}. The alpha_NC_3B1f_ggF_{year} is the non-closure systematic derived in the 3b1f validation region. The NPs encoding the systematic uncertainties in the jet reconstruction due to energy resolution, scale, and pileup modelling to name a few have the prefix alpha_JET_{NP name}, while the flavour tagging NPs have the prefix alpha_FT_{NP name} and alpha_SF_PRW is the uncertainty in the pileup reweighting. The theoretical uncertainties have the prefix alpha_THEO_{NP name}_{signal}, for example, alpha_THEO_XS_SCALEMTop_ggf is uncertainty in the ggF production cross-section, due to the scale choice and finite top mass scheme.

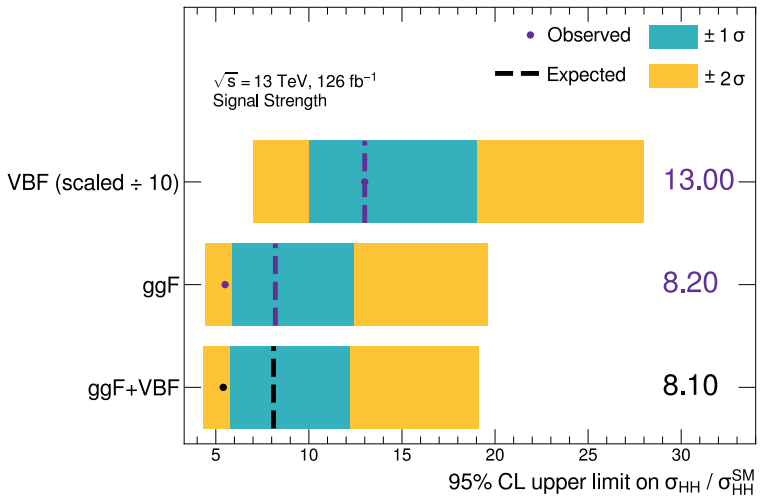
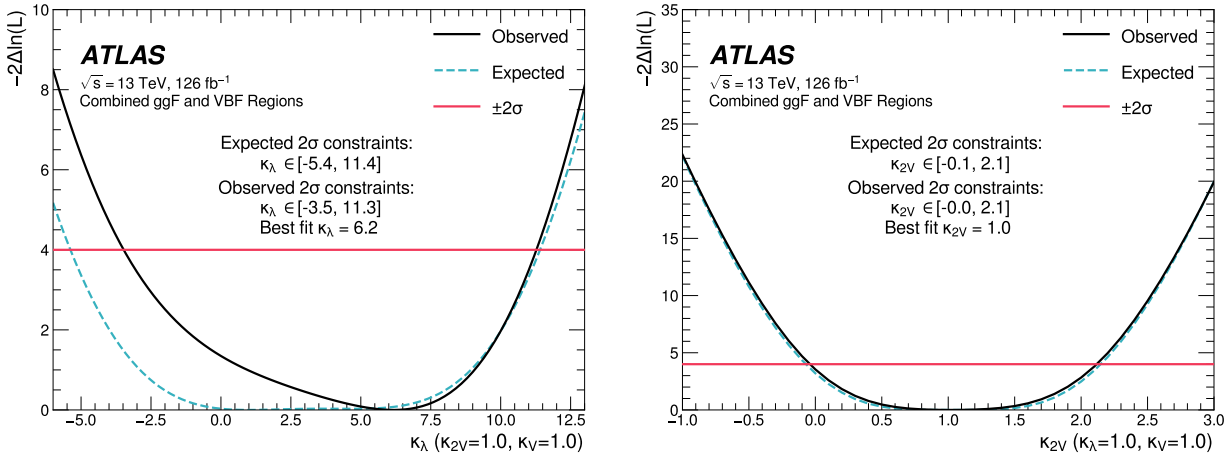


Figure 6.21: The upper limit on the signal strength of the SM VBF, ggF, and combined ggF+VBF HH production is plotted. The limits on the VBF signal are scaled down by a factor of 10 for better visualisation. The observed limits are shown as a circle. The expected limits are shown as dashed lines and the 1σ (2σ) error bounds are plotted as turquoise (yellow) bands.

κ_{2V} $-0.0 < \kappa_{2V} < 2.1$ ($-0.1 < \kappa_{2V} < 2.1$) and the observed best-fit of κ_{2V} is the SM point. The confidence intervals are extracted from the combined ggF and VBF channels, as the contribution from the VBF signal that leaks into the ggF channels becomes significant at certain non-SM κ_{2V} values and improves the sensitivity of the measurements. A more general approach to set constraints on κ_λ and κ_{2V} is to relax the assumption that only one modifier is allowed to deviate from unity significantly. Allowing both κ_λ and κ_{2V} to float in the fit while keeping all other couplings fixed to their SM values results in the $\pm1\sigma$ and $\pm2\sigma$ observed (expected) confidence contours in the $\kappa_\lambda - \kappa_{2V}$ plane shown in Figure 6.23a (6.23b). The SM point is within the $\pm2\sigma$ confidence contours but the best-fit takes a non-SM value in the fit to observed data, due to the data excess at low m_{HH} in the ggF channel. The constraints on κ_λ and κ_{2V} are slightly looser compared to the extracted values from the 1D $-2\Delta\ln(L)$ scans, which is expected due to similar signal yield and the shift in the non-SM best-fit value of $(\kappa_\lambda = 8.8, \kappa_{2V} = 0.1)$. The $\boldsymbol{\kappa}$ HH signal that minimises the log-likelihood in both, the 1D and 2D scans, is a BSM signal that peaks at low m_{HH} masses. The observed data prefers the non-SM $\boldsymbol{\kappa}$ best-fit value, as the increase of expected events at low m_{HH} mass slightly lessens the pulls on the NPs that



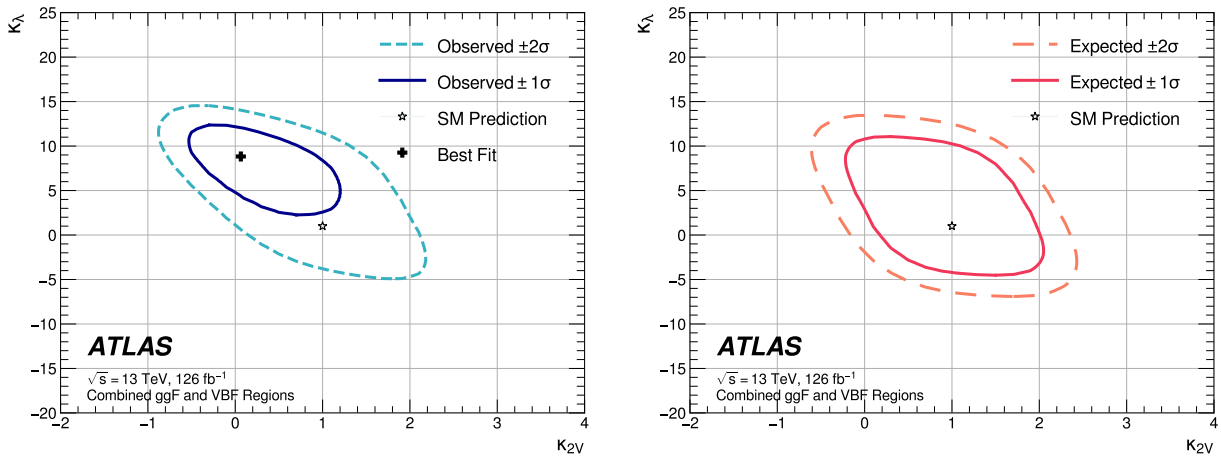
(a) (b)

Figure 6.22: The $-2\Delta\ln(L)$ with κ_λ in (a) and κ_{2V} in (b) as a POI are shown. The expected $-2\Delta\ln(L)$ scans derived from an Asimov dataset (dashed line) assume the existence of an SM HH signal. In the fit, the ggF and VBF channels are combined and all $\boldsymbol{\kappa}$ but the POI are fixed to unity. The VBF HH signal is included in the ggF channels and vice versa. The signal strengths are fixed to their theoretical predictions and can be modified only through their systematic uncertainties. The observed (expected) $\pm2\sigma$ confidence levels are $-3.5 < \kappa_\lambda < 11.3$ ($-5.4 < \kappa_\lambda < 11.4$) and $-0.0 < \kappa_{2V} < 2.1$ ($-0.1 < \kappa_{2V} < 2.1$) as derived from the $-2\Delta\ln(L)$ scans shown in (a) and (b) respectively. The best-fit value in the κ_λ $-2\Delta\ln(L)$ scan is shifted away from the SM to $\kappa_\lambda = 6.2$, as the data distribution at low m_{HH} has a slightly more prominent peak than the pre-fit background estimation. The results have been published in [2].

represent the background uncertainties. Consequently, the penalty from the constraint terms is reduced, which minimises the log of the likelihood.

6.8.2 Effective Field Theory Interpretations

The Higgs Effective Field Theory (HEFT) and Standard Model Effective Field Theory (SMEFT) interpretations are considered as an alternative to the $\kappa-$ framework based measurements. The limits on the EFT coefficients and benchmarks are set by only considering the ggF channels, as the theoretical EFT framework for the VBF HH production has not yet been developed. In addition, the SM VBF HH signal in the ggF categories has a negligible impact on the extracted limits and is therefore neglected



(a) (b)

Figure 6.23: The observed in (a) and expected in (b) $\pm 1\sigma$ (solid line) and $\pm 2\sigma$ (dashed line) confidence contours in the $\kappa_\lambda - \kappa_{2V}$ parameter space as derived from the $-2\Delta \ln(L)$ scans are shown. The star indicates the SM point, while the best-fit $(\kappa_\lambda, \kappa_{2V})$ point in the observed data is shown as the cross. All other $\boldsymbol{\kappa}$ are fixed to their SM values. The figure has been published in [2].

in the interpretations. Non-SM values of the EFT coefficients can affect the branching fraction of $H \to b\bar{b}$. However, the effects are not accounted for as the analysis has little sensitivity to the expected $\Delta Br(H \to b\bar{b})$. All 95% confidence levels are extracted by using the asymptotic formula and CLs technique described in Section 6.7.5. The expected limits and corresponding error bands are extracted from the background-only hypothesis.

Upper limits on the HH production cross-section for seven different HEFT benchmarks (BM), which are representative of various possible shapes that the m_{HH} distribution can take, are shown in Figure 6.24. From comparing the theoretically predicted and the observed upper limit of the HH production cross-section it can be deduced that the BM3, BM5, and BM7 are excluded at more than 95% confidence level. The discrepancy of about 1σ between the expected and observed upper limits in some BMs is the result of the slight excess at low m_{HH}. In addition, limits are set on the allowed values of $c_{t\bar{t}HH}$ and c_{ggHH}, which introduce non-SM like Feynman diagrams, see Figure 2.5a and 2.5c, that contribute to the HH production. The

allowed range of $c_{t\bar{t}HH}$ (c_{ggHH}) is determined by setting the upper limit on the cross-section at 95% confidence level for several $c_{t\bar{t}HH}$ (c_{ggHH}) points. If at a given $c_{t\bar{t}HH}$ (c_{ggHH}) value the observed upper cross-section limit is below the theoretically predicted cross-section the point is excluded with more than 95% confidence level. The results of the $c_{t\bar{t}HH}$ and c_{ggHH} scans, where each but the examined coefficients are set to their SM values, are shown in Figure 6.25. Using the strategy described above it is observed that $c_{t\bar{t}HH} \in [-0.36, 0.78]$ and $c_{ggHH} \in [-0.55, 0.51]$, the values of the coefficients outside the stated intervals are excluded at $\geq 95\%$ confidence level. The observed constraints on $c_{t\bar{t}HH}$ and c_{ggHH} have to be taken with reservation as the assumption that new physics would have an observable impact only on $c_{t\bar{t}HH}$ or c_{ggHH} does not hold for all BSM models.

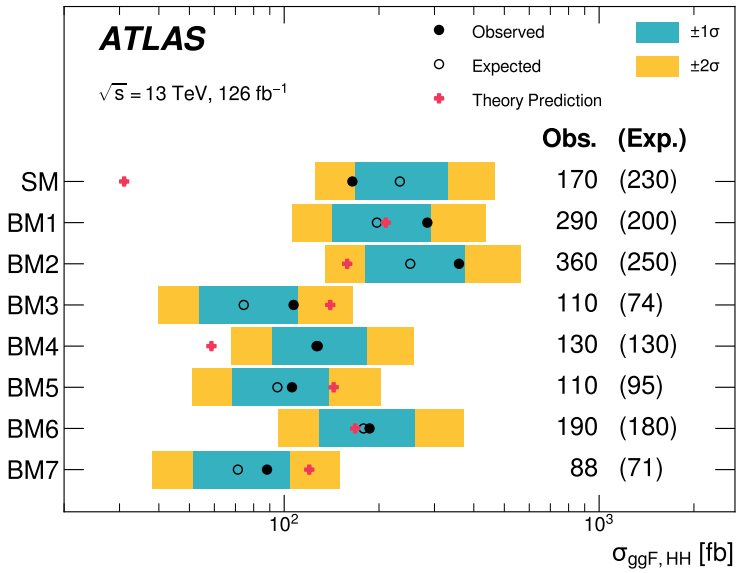
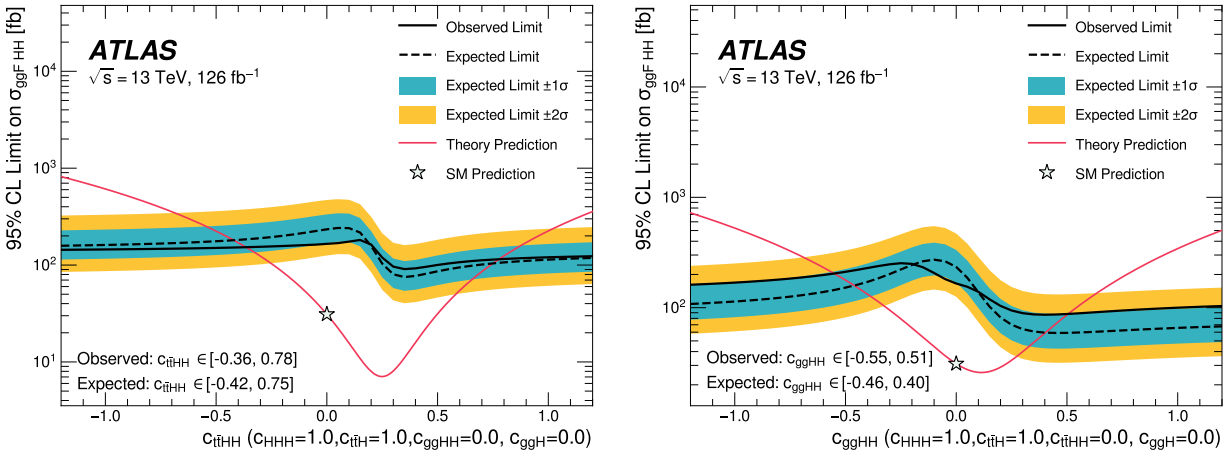


Figure 6.24: The upper limits set at a 95% confidence level on HEFT benchmarks and the SM point, defined in Table 2.1. The hollow black circle represents the expected limit and its corresponding 1σ (2σ) error bounds as the turquoise (yellow) band. The observed upper limits are marked as a filled-in black circle. The pink crosses mark the theoretically predicted cross-section for each BM. The plot is shown as published in [2].

Finally, in the SMEFT framework constraints are placed on the five parameters to which the HH production is sensitive. Figure 6.26 shows the upper exclusion cross-section limits derived with a 95% confidence level on various c_H, c_{HG}, $c_{H\Box}$, c_{tH},



(a) (b)

Figure 6.25: The observed 95% confidence level upper exclusion limits on the $c_{t\bar{t}HH}$ in (a) and c_{ggHH} in (b) in the range [-1.2,1.2] is shown by the black line. The red line indicates the theoretically predicted cross-section, if the observed limit is below the theoretical prediction the point is excluded with a confidence level of more than 95%. The star indicates the SM point, which is zero for both coefficients considered. Expected upper exclusion limits are shown as the dashed line with the 1σ (2σ) error bound in turquoise (yellow). In each of the scans, only the considered coefficient is varied, while all other HEFT coefficients are fixed at their SM value. From these, the extracted observed (expected) upper exclusion limits are $c_{t\bar{t}HH} \in [-0.36, 0.78]$ ($c_{t\bar{t}HH} \in [-0.42, 0.75]$) and $c_{ggHH} \in [-0.55, 0.51]$ ($c_{ggHH} \in [-0.48, 0.40]$). The figures have been published in the auxiliary material of [2].

and c_{tG} points. All but the scanned coefficients are fixed to their SM values. For example, in the c_H scan c_{HG}, $c_{H\Box}$, c_{tH}, and c_{tG} are fixed to zero while the different signals are obtained by varying c_H. As in the HEFT 1D scans, a c_H point is excluded at more than 95% confidence level if its upper cross-section limit is smaller than the theoretically predicted value. The 95% confidence level intervals for both observed and expected values of each scanned SMEFT parameter are extracted from Figure 6.26. The c_{HG} parameter is tightly constrained to $c_{HG} \in [-0.067, 0.60]$, deviations outside the interval are excluded with an above 95% confidence level. The tight constraints are expected, as non-zero c_{HG} values introduce the two additional BSM Feynman diagrams shown in Figures 2.5b and 2.5c which significantly amplify the HH production. However, it is unlikely that new physics at high energy modifies only one of the SMEFT coefficients considered. Two-dimensional parameter scans are performed of c_H with respect to the remaining four SMEFT parameters. The allowed contours in the 2D parameters space are shown in Figure 6.27. The $c_H - c_x$, where $x \in [HG, tG, tH, H\Box]$ contours are derived using the same method as for the 1D 95% confidence intervals described above. All results, which allow the coefficients of size $\pm 4\pi$ should be taken with a grain of salt, as the SMEFT interpretation can break down. If one or more parameters are of the size $\pm 4\pi$, the power-cutting scheme might not be valid and the missing higher-orders can contribute significantly.

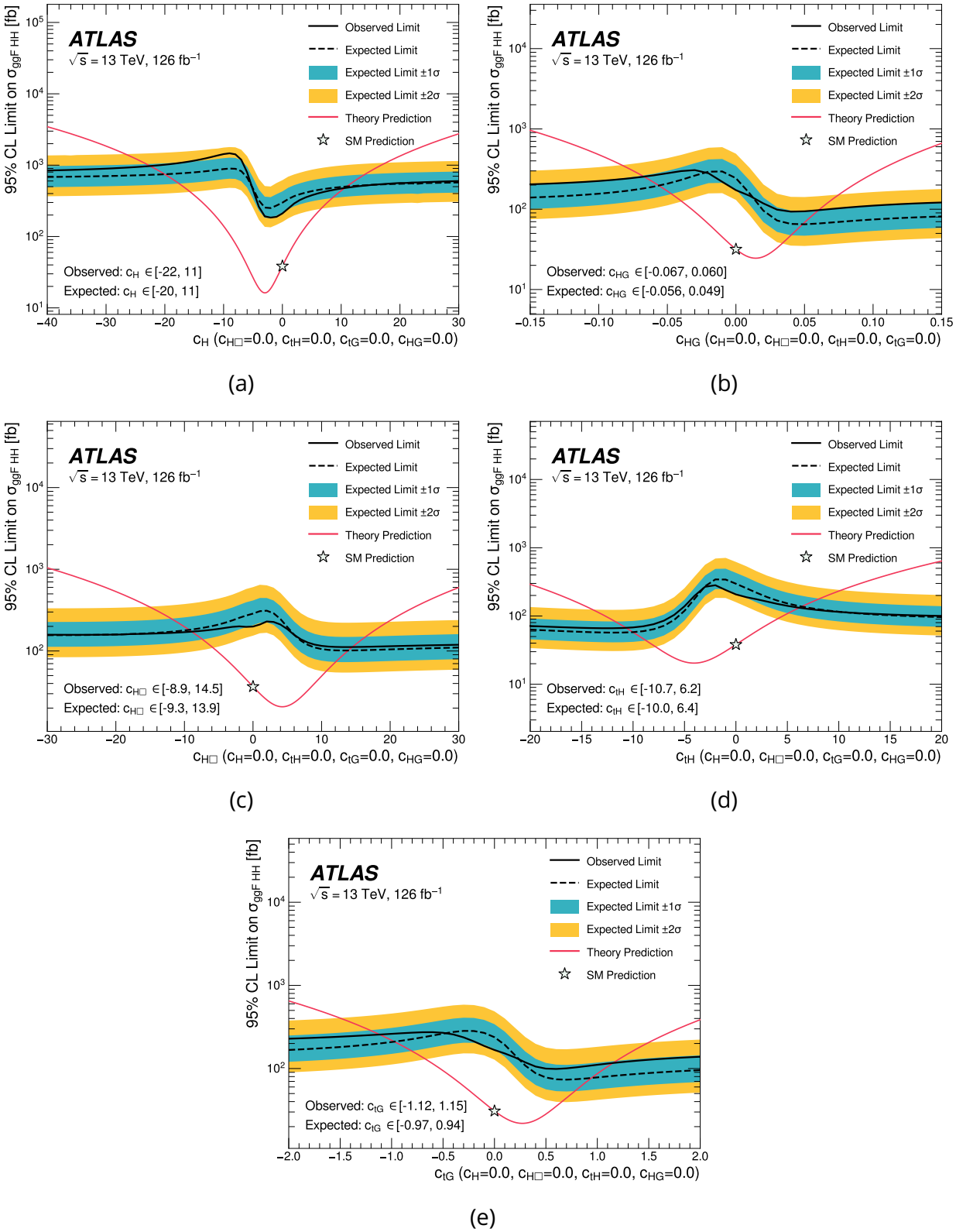


Figure 6.26: The 95% upper exclusion limits on c_H in (a), c_{HG} in (b), $c_{H\Box}$ in (c), c_{tH} in (d), and c_{tG} in (e) are shown. The solid black line represents the observed limits, while the pink line illustrates the theoretical cross-section corresponding to each considered variation. The dashed black line shows the expected limits and the turquoise (yellow) bands are the corresponding 1σ (2σ) error bounds. The SM point is located at zero. The figures have been published in [2].

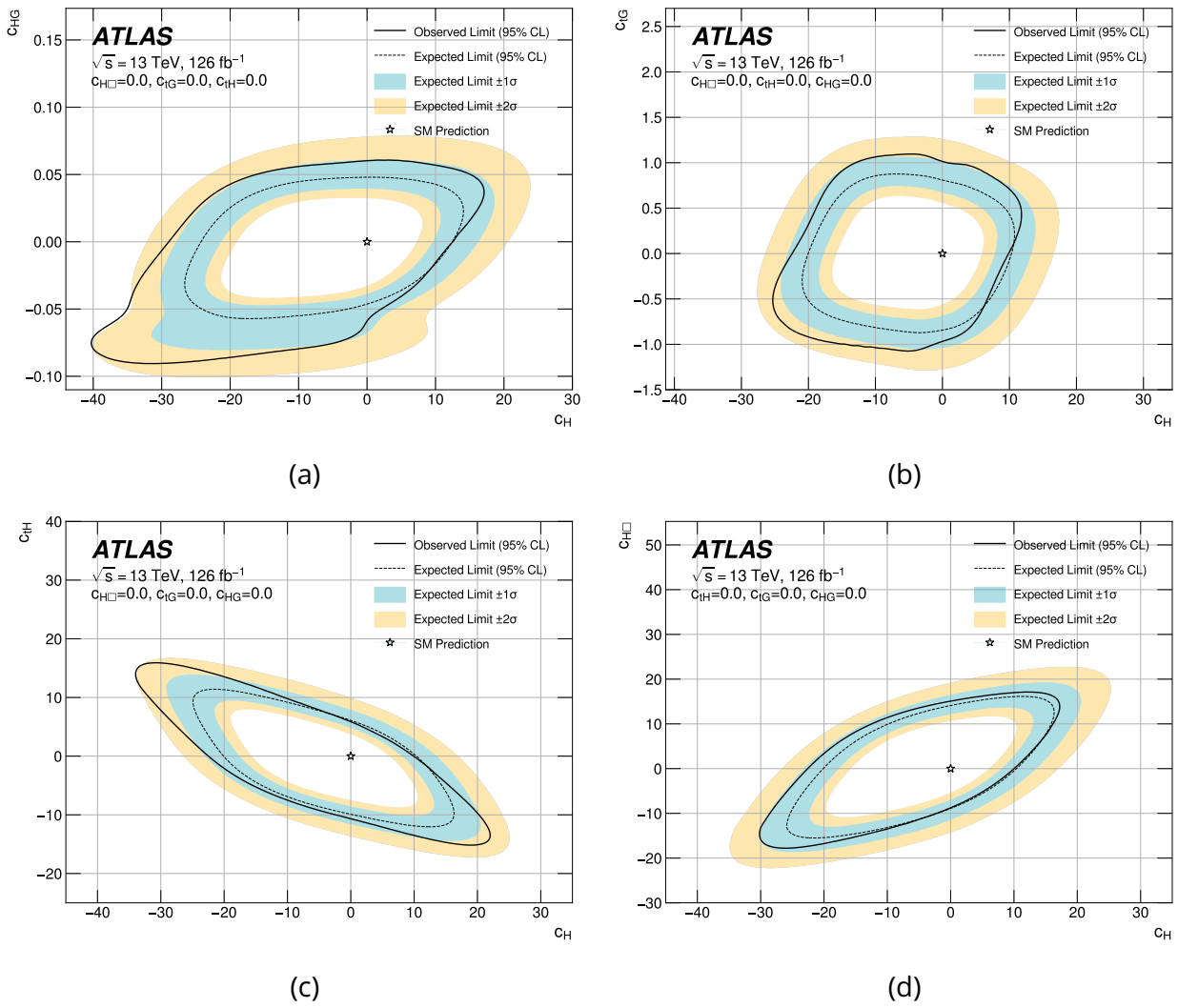


Figure 6.27: The observed (expected) contours that bound the allowed two-dimensional parameter space $c_H - c_{HG}$ in (a), $c_H - c_{tG}$ in (b), $c_H - c_{tH}$ in (c), and $c_H - c_{H\Box}$ in (d) at 95% confidence level are shown as a black solid (dashed) line. The 1σ (2σ) error bounds on the expected contours are illustrated as turquoise (yellow) bands. The SM point is within the allowed parameter space in all cases considered. The two-dimensional contours have in published in [2].

6.9 Conclusion

The search for non-resonant diHiggs production in the $b\bar{b}b\bar{b}$ final state is presented. The upper limits on the ggF, VBF, and combined signal strength are computed. The improvement as compared to the Run 2 analysis done with $36.1\,\mathrm{fb}^{-1}$ ($126\,\mathrm{fb}^{-1}$) of proton-proton collision data collected published in [181] ([182]) is more than 50% (70%) on the ggF (VBF) signal strength limit. The gain in sensitivity is a result of the larger luminosity (only in the ggF channels), and advancements in analysis strategies including better b-tagging [2]. The observed (expected) upper limit on the HH production computed by combining the ggF and VBF channels is 5.4 (8.1). As no statistically significant excess has been found the analysis interprets the results within the κ and EFT frameworks independently.

The extracted $\pm 2\sigma$ confidence levels from the observed data are $-3.5 < \kappa_\lambda < 11.3$ on κ_λ and $-0.0 < \kappa_{2V} < 2.1$ on κ_{2V}. In addition, a more general κ model is considered, which extracts the two-dimensional $\pm 1(2)\sigma$ bounds on κ_λ and κ_{2V}. Three of the seven HEFT benchmark models are excluded at more than 95% confidence level. The allowed ranges for c_{ttHH} and c_{ggHH} is determined. The five parameters considered within the SMEFT interpretation are c_H, c_{HG}, $c_{H\Box}$, c_{tH}, and c_{tG}. One-dimensional scans in addition to two-dimensional scans are performed, in the later c_H is varied in addition to one other coefficient. The results should be interpreted with caution, as allowing two of the five SMEFT coefficients to deviate from zero results in a significant loss of sensitivity. Constraints of the size $\sim 4\pi$ might be outside the validity regime assumed in SMEFT.

The ultimate goal is to provide the most powerful and comprehensive set of constraints on the Higgs self-coupling. To achieve this, the next step has been to combine the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis with other diHiggs and single Higgs channels. The combination provides the most stringent constraints on the Higgs self-coupling up to date and is presented in Chapter 7. Future plans involve the

combination of all the HH Run 2 analyses in the various final states considered and a possible EFT combination with single Higgs channels.

Chapter 7

Combining single Higgs and diHiggs Measurements

Measurements of the Higgs self-coupling are accessible directly through the diHiggs (HH) and indirectly through single-Higgs (H) production modes, as described in Sections 2.1 and 2.3. The combination of the diHiggs channels enables the extraction of the most stringent upper limits on the SM diHiggs signal strength and measurements of the Higgs self-coupling. However, to constrain the Higgs self-coupling modifier (κ_λ) without making too many assumptions on the values of other Higgs vertices, the combination of single-Higgs and diHiggs (H+HH) channels is necessary.

This chapter describes the most up-to-date combination of H and HH channels published by the ATLAS Collaboration in [3], which is the general reference for this chapter. The input analysis and the datasets used are described in Section 7.1, while the combination strategy is outlined in Section 7.2. Adjustments needed in the $HH \to b\bar{b}b\bar{b}$ channel are described in Section 7.3, while additional checks on the difference in the Higgs mass definitions are in Section 7.4. The results of the H+HH combination are presented in Section 7.5.

Acknowledgements and Author's Contributions

The author of this thesis made significant contributions to the preparation of the $HH \to b\bar{b}b\bar{b}$ channel, presented in Section 7.3. In addition, the author helped performed minor checks on the diHiggs combination, one of which is presented in Section 7.4. The internal documentation [183] states the other ATLAS authors who contributed significantly to the combination.

7.1 Input Channels

The inputs to the HH combination are the three most sensitive final states, $HH \to b\bar{b}\gamma\gamma$ [135], $HH \to bb\tau^-\tau^+$ [133] and $HH \to b\bar{b}b\bar{b}$ [2]. The analyses account for the two leading production modes, the ggF and VBF channels, as the others are negligible. The H combination consists of the $H \to \gamma\gamma$ [184], $H \to \tau^-\tau^+$ [185] and $H \to ZZ^* \to 4l$ [186] decay channels, the $H \to b\bar{b}$ channel produced via the VH, VBF and $t\bar{t}H$ modes [187, 188, 189] and the $H \to WW^* \to e\nu\mu\nu$ [190] channel produced via ggF and VBF modes. The input analyses use the full Run 2 proton-proton collision dataset collected between 2015-2018 at $\sqrt{s} = 13\,\mathrm{TeV}$ by the ATLAS Experiment. The $HH \to b\bar{b}b\bar{b}$ and $H \to b\bar{b}$ (VBF) channels use $126\,\mathrm{fb}^{-1}$ of integrated luminosity due to the trigger requirements. All other analyses use $139\,\mathrm{fb}^{-1}$ of integrated luminosity. The mass of the Higgs (m_H) is assumed to be $125.09\,\mathrm{GeV}$ [191] in single Higgs inputs and the $b\bar{b}\gamma\gamma$ diHiggs channel, while the diHiggs $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^-\tau^+$ final states set m_H to $125.00\,\mathrm{GeV}$.

7.2 Combination Strategy and Statistical Analysis

The single Higgs channels are sensitive to κ_λ via next-to-leading-order (NLO) electroweak (EW) corrections that modify the kinematics and abundance of the final

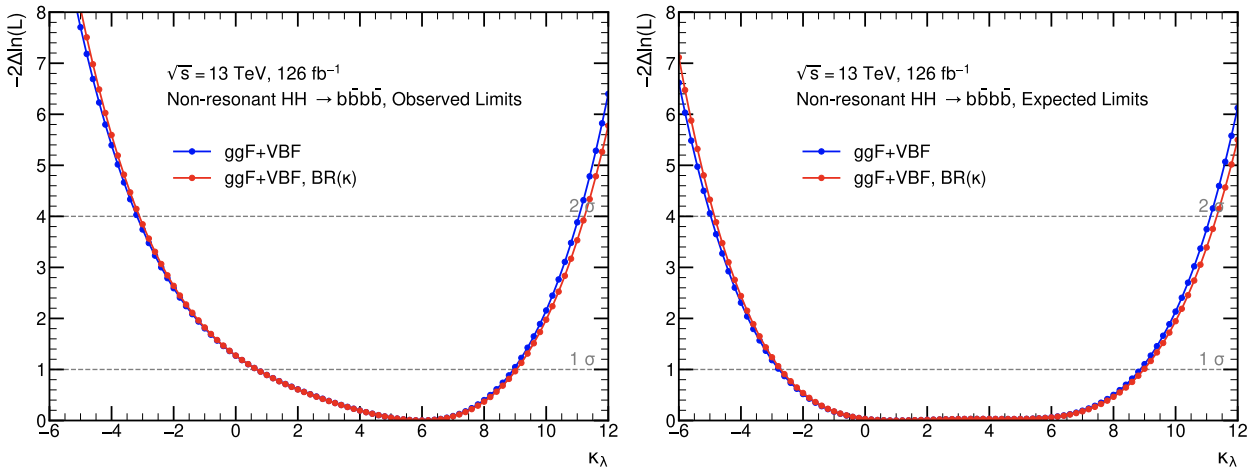
states. The impact of the NLO EW corrections is quantified via the model based on the $\kappa-$ framework and described in Section 2.3. In the model the differential cross-section and branching fraction vary independently, provided the Higgs couplings deviate from their SM expectation. The H and HH signals are parametrised with respect to the model, where the simplified template cross-sections (STXS) [192] are used in H channels.

The single Higgs (diHiggs) channels are combined by extracting the total likelihood function by multiplying together the likelihood functions obtained in each analysis ($\mathcal{L}_{total} = \prod_{c\in\text{channel}} \mathcal{L}_c(\boldsymbol{\kappa}, \boldsymbol{\theta})$, $\mathcal{L}_c(\boldsymbol{\kappa}, \boldsymbol{\theta})$ defined in Equation 6.15). This simple combination is possible as the event overlap between channels is negligible. Similarly, the H+HH combination is assembled by joining the H and HH likelihoods. The $t\bar{t}H$ category in the $H \to \tau^-\tau^+$ channel is removed from the H combination as there is a few percent overlap with the $HH \to b\bar{b}\tau^-\tau^+$ analysis. The systematic uncertainties on the data-taking conditions are correlated between all the channels, while the systematic uncertainties on physics objects are correlated only if the same object definitions, working points, and calibration steps have been used [3]. Systematics that come from theoretical uncertainties in the signal or background simulations are in principle correlated (per production modes in the case of the signal). The theoretical uncertainty on the branching fraction is correlated across the channels and production modes. The choice to treat systematics constrained by the data fit as correlated or uncorrelated is made after testing both schemes and choosing the more conservative one.

The upper limits on the HH production signal strength are set using the CLs technique described in Section 6.7.5. The 95% confidence intervals on κ_λ measurement are extracted from the log-likelihood ratio test with two times the negative log-likelihood ratio ($-2\ln\Lambda$) as the test statistic as outlined in Section 6.7.3.

7.3 Tweaking the $HH \to b\bar{b}b\bar{b}$ Analysis

To be consistent with the single H analyses the HH channels have to account for the NLO EW corrections, which are not taken into account by default due to their small effect on κ_λ in HH searches. The NLO EW dependence of the diHiggs production in the $b\bar{b}b\bar{b}$ final state is accounted for through modifying the branching fraction of the Higgs decaying to $b\bar{b}$ using Equation 2.11. The $BR(H \to b\bar{b})$ only depends on κ_λ through the denominator in Equation 2.11, as $C_1^b \approx 0$. This leads to a maximum of 5% increase (decrease) in the $BR(H \to b\bar{b})$ in the allowed negative (positive) κ_λ range as can be seen in Figure 2.8b, when only κ_λ is varied. The change in the κ_λ log-likelihood scan, shown in Figure 7.1, is as expected. The 2σ confidence interval has a stronger lower bound, while the upper bound on κ_λ becomes weaker in both the expected and observed limits. The change is smaller than the granularity of the scan ($\Delta\kappa_\lambda < 0.2$) and therefore can be neglected in the HH analyses.

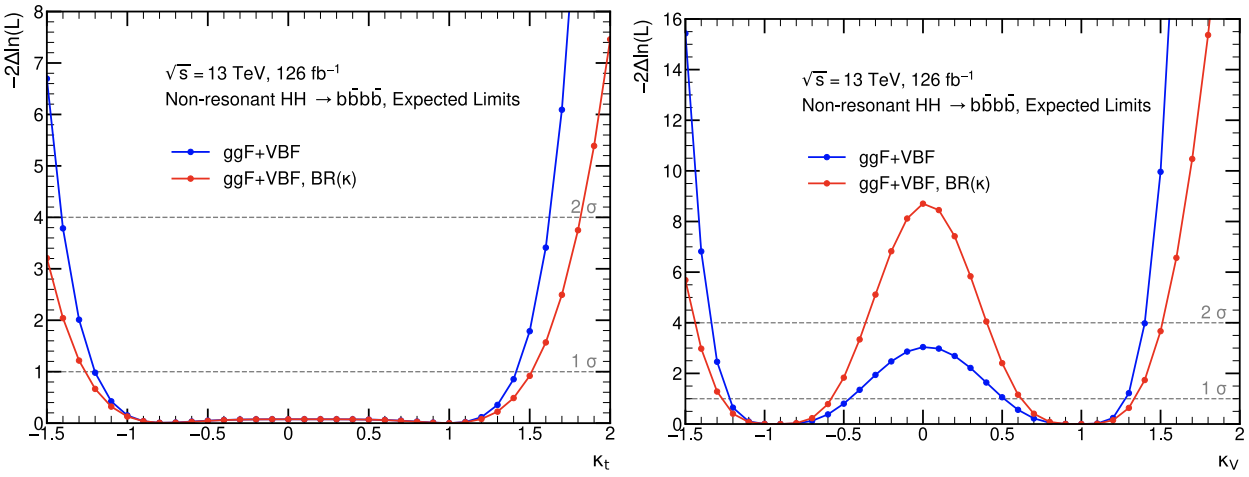


(a) (b)

Figure 7.1: The observed (left) and expected (right) $-2\ln\Lambda$ scans show the scenario that accounts for the NLO EW corrections (red line) and the one that ignores the effects (blue line) in $HH \to b\bar{b}b\bar{b}$. The POI is κ_λ, all other κ_m are fixed to unity.

Single Higgs precision measurements put tight constraints on several couplings between the Higgs and SM particles, which enables the relaxation of some assumptions in the H+HH combined fit. In addition to κ_λ one or more κ_m are floated,

where m is the coupling of Higgs to the SM particle m ($m \in t, V, \tau, b$). All HH channels are directly sensitive to κ_λ, κ_t, κ_V and κ_{2V} and only indirectly through $BR(\kappa_\lambda, \kappa_m)$ to κ_τ and κ_b. In the HH ggF production the dependence on κ_t is clear from Equation 6.3, which is valid to all orders as no new ttH or HHH vertices are introduced through additional loops [3]. The HH production depends on κ_V and κ_{2V} at LO through the VBF mode according to Equation 6.6. The sensitivity to κ_t and κ_V in the $HH \to b\bar{b}b\bar{b}$ channel can be seen in Figure 7.2, which shows the $-2\ln\Lambda$ scans to Asimov data. NLO EW corrections have a larger impact on the κ_t and κ_V as compared to κ_λ when the respective couplings are varied with all other modifiers fixed to unity. The κ_{2V} coupling is not varied as there is no parametrisation available in the single Higgs channels and it is tightly constrained by κ_V.



(a) (b)

Figure 7.2: The change in the $-2\ln\Lambda$ scan when the NLO EW corrections are (red line) and are not (blue line) taken into account as a function of κ_t (left) and κ_V (right). All other κ_m are fixed to unity.

It can be concluded that, while NLO EW corrections can be neglected in individual HH analyses, such as the $HH \to b\bar{b}b\bar{b}$ analysis presented in Chapter 6, they cannot be ignored in the H+HH combination. This is because the H+HH combination allows several κ modifiers, including κ_t and κ_V, to vary. As evident from Figure 7.2, NLO EW corrections when κ_t or κ_V deviate from unity even slightly have a significant impact on the HH channels. Therefore, for the sake of consistency in the H+HH

combination, the HH channels should account for NLO EW corrections associated with all κ modifiers that impact the HH signal.

7.4 Impact of Higgs Mass Assumptions

The single H and the HH $b\bar{b}\gamma\gamma$ analyses included in the combination all assume $m_H = 125.09\,\mathrm{GeV}$, while the HH $bb\tau^-\tau^+$ and $b\bar{b}b\bar{b}$ channels assume $m_H = 125.00\,\mathrm{GeV}$. The impact of the different Higgs mass assumptions on the expected limits is checked by scaling the cross-section and branching fraction of the $bb\tau^-\tau^+$ and $b\bar{b}b\bar{b}$ signals as well as single H background in the $bb\tau^-\tau^+$ channel[1] by a scale factor (SF). For example in the $b\bar{b}b\bar{b}$ channel the modification is $SF = (\frac{BR(m_H=125.09)}{BR(m_H=125.00)})^2 \times \frac{\sigma_{ggF(VBF)}(m_H=125.09)}{\sigma_{ggF(VBF)}(m_H=125.00)} = 0.994(0.993)$, the values are obtained from [170]. To account for the impact on the m_{HH} shape the signal samples and weights derived for the κ reweighting would need to be regenerated. As the process is computationally expensive, time-consuming and the expected difference is small this is not pursued. The impact of scaling m_H to 125.09 GeV in the $b\bar{b}b\bar{b}$ and $bb\tau^-\tau^+$ HH channels have a maximum change on the lower 2σ constraint in the likelihood scan of 1.2%, while the changes on the signal strength and the upper 2σ constraint from the likelihood scan as well as the best-fit value are below 0.5%. The changes coming from the different values of m_H used are found to have a negligible impact on the HH results and the $b\bar{b}b\bar{b}$ and $bb\tau^-\tau^+$ HH analyses can be used in the combination without any further changes. In addition, to there being no sensitivity to the yield difference also no effects on the reconstruction of physics objects are expected as the mass resolution is worse than $90\,\mathrm{MeV}$.

[1] The single Higgs background in the $b\bar{b}b\bar{b}$ channel is negligible.

7.5 Results

The results of the single Higgs and diHiggs combination have been published in [3], the highlights are summarised in this section.

In preparation for the final combination, the upper limits on the SM HH signal strength are extracted from each channel entering the HH combination and the combination itself. The HH ggF and HH VBF signals are normalised to their SM prediction, which results in $\sigma_{HH} \sim 32.7\,\mathrm{fb}$. The other production modes are not considered, as their contribution is negligible. In the combination the HH signal strength (μ_{HH}) floated in the fit is common between the ggF and VBF signals and the combined decay channels. The results are summarised in Figure 7.3, which shows the upper limit on μ_{HH} in the individual channels and their combination. The observed (expected) upper limit on μ_{HH} derived from the HH combination is 2.4 (2.9) times the SM HH signal strength. The H and HH channels are combined with the goal to extract the allowed κ_λ range given different assumptions. The considered models are;

- κ_λ **only model** in which κ_λ is the POI and other coupling modifiers are fixed to their SM values.
- κ_λ **generic model**, where κ_t, κ_b, κ_V and κ_τ are free NPs in the fit in addition to κ_λ.
- $\kappa_\lambda - \kappa_\mathbf{t}$ **model**, all coupling modifiers except for κ_λ and κ_t are fixed to unity.

The κ_λ only model is the simplest interpretation, as it assumes that new physics has a detectable impact only on the trilinear Higgs self-coupling. The $-2\ln\Lambda$ scans for the observed and expected data when fitting the H, HH, and H+HH combinations are shown in Figure 7.4. The observed (expected) 95% confidance interval is $-0.6 < \kappa_\lambda < 6.6\ (-2.1 < \kappa_\lambda < 7.8)$ given the HH only combination, while the combination of H+HH channels gives $-0.4 < \kappa_\lambda < 6.3\ (-1.9 < \kappa_\lambda < 7.5)$. Little is

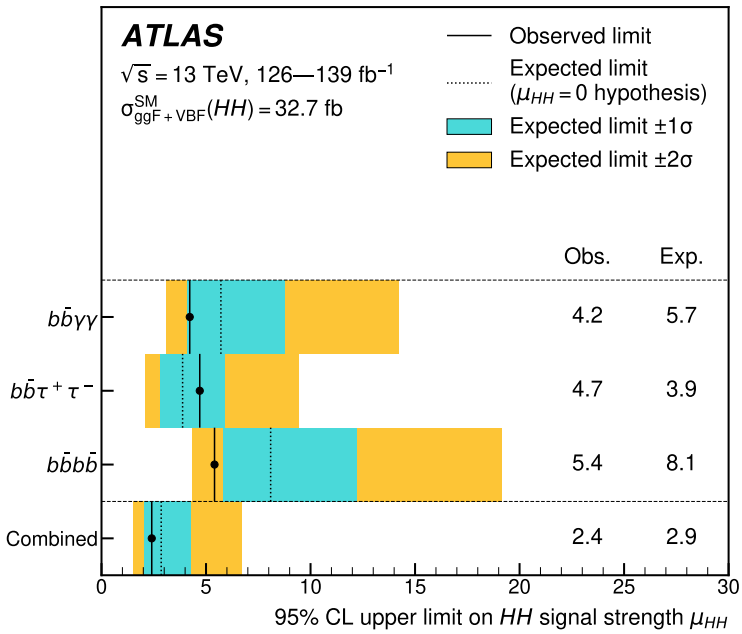
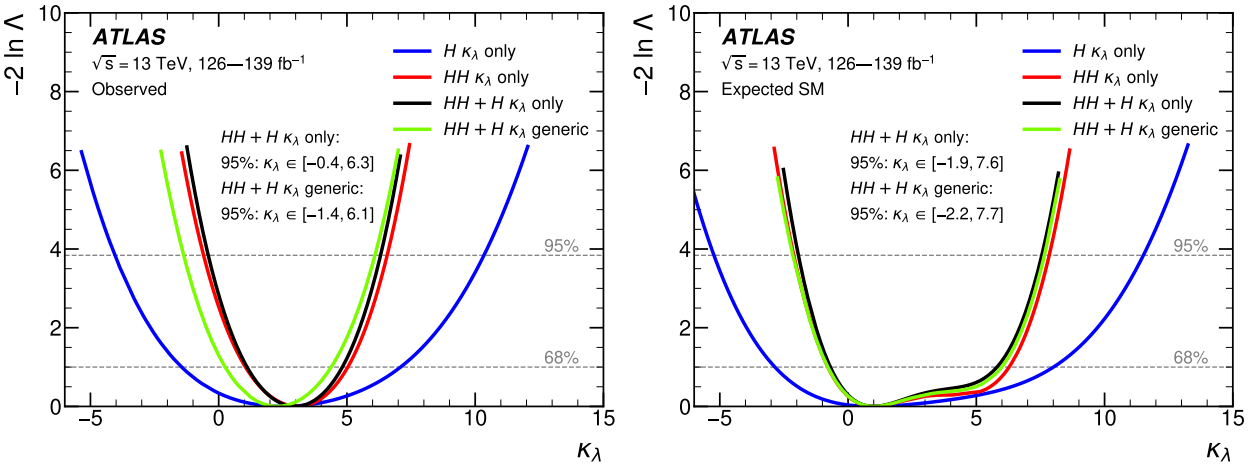


Figure 7.3: The diHiggs only combination includes the three diHiggs channels leading in sensitivity, $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^{-}\tau^{+}$, and $b\bar{b}b\bar{b}$. The HH combination provides the latest 95% CL upper limits on the SM signal strength published by the ATLAS Collaboration, shown in the plot in addition to the upper limits in each channel. The extracted expected limits and corresponding error bands assume no HH production and take the NP values from the profiled fit to data. The results are published in [3].

gained by combing the single Higgs and diHiggs analyses, as κ_λ is dominantly constrained by the HH searches. The power of including the single Higgs precision measurements is that it is possible to measure κ_λ by loosening up the constraints on other Higgs couplings. For example, HH measurements alone do not have enough sensitivity to the ttH coupling for its modification (κ_t) to be a free parameter in the fit with $POI = \kappa_\lambda$ (see Figure 7.5). However, when considering both constraints from single Higgs and diHiggs measurements it is possible to use the κ_λ generic model, float several Higgs coupling modifiers at once, and still extract a competitive confidence interval on κ_λ as compared to the more restrictive κ_λ only model. In the κ_λ generic model the observed (expected) 95% confidance interval is $-1.3 < \kappa_\lambda < 6.1 (-2.1 < \kappa_\lambda < 7.6)$, which differs from the κ_λ only model. The difference in the confidence interval between the two models arises from the correlations between the floated $\boldsymbol{\kappa}$. The observed combined dataset prefers κ_t, κ_b,

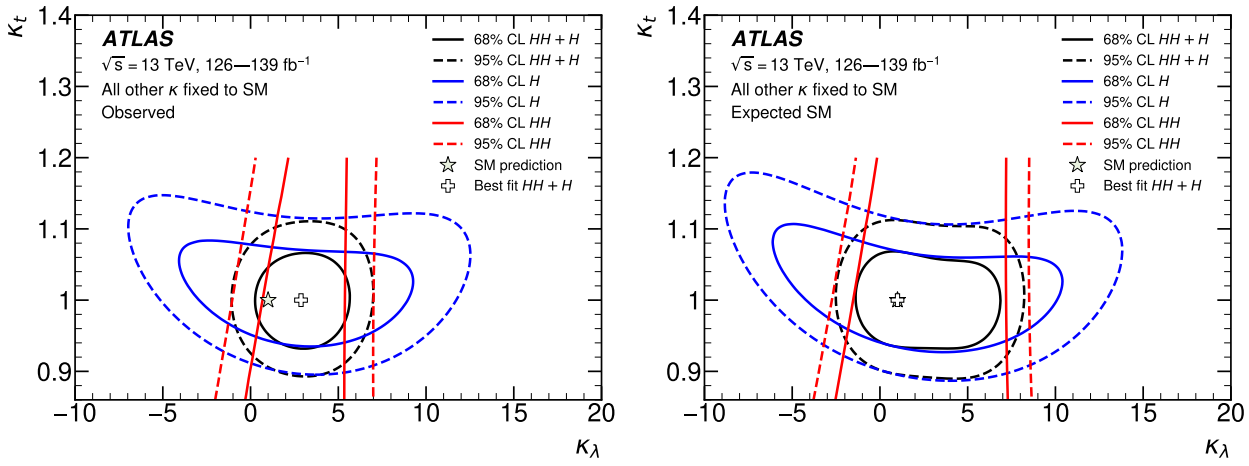
and κ_τ values below unity which shifts the best-fit values for κ_λ and consequently the confidence intervals. It should be noted that the post-fit measurements of κ_t, κ_b, and κ_τ are nevertheless consistent with unity within uncertainties.



(a) (b)

Figure 7.4: The $-2\ln\Lambda$ test statistic with κ_λ as a POI is shown for the κ_λ only model (black line) and the κ_λ generic model (green line). Besidies κ_λ the generic model allows κ_t, κ_b, κ_V and κ_τ to float in the fit as well. The intersection of the test statistic with the 68% (95%) dashed lines sets the 68% (95%) confidence levels on κ_λ. The expected data assumes the SM HH signal exists. The publication containing the results is [3].

In the $\kappa_\lambda - \kappa_t$ model it is possible to put tight constraints on the two-dimensional parameter space when combing H and HH measurements. Figure 7.5 shows the 68% and 95% confidence contours for the H (blue), HH (red), and H+HH (black) combination. It can be seen that the HH combination only cannot constrain both coupling modifiers. Yet, the H+HH combination shrinks the allowed $\kappa_\lambda - \kappa_t$ parameter space as compared to the H-only combination. Comparing the different assumptions in the H+HH combined fit, it can be noted that the $\kappa_\lambda - \kappa_t$ and κ_λ only models set the same 95% confidence interval on κ_λ, as the best-fit value of k_t is approximately unity.



(a) (b)

Figure 7.5: The solid (dashed) lines show the 68% (95%) CL contours in the $\kappa_\lambda - \kappa_t$ plane, the other Higgs couplings are fixed to their SM value. The observed constraints are shown in 7.5a, while the expected limits are shown in 7.5b and have been published in [3]. In expected limits fit it is assumed the SM HH signal exists.

7.6 Conclusion

To tightly constrain the Higgs self-coupling while relaxing assumptions usually placed on other Higgs couplings the combination of single Higgs and diHiggs channels has been performed. Modifications of the $HH \to b\bar{b}b\bar{b}$ diHiggs analysis to take into account the dependence on κ_λ, κ_b, κ_V and κ_τ through electroweak next-to-leading order corrections are described and the impact of different values assumed for the Higgs mass between the channels is checked. The observed (expected) 95% confidance level is measured to be $-1.3 < \kappa_\lambda < 6.1$ $(-2.1 < \kappa_\lambda < 7.6)$ in the generic κ_λ model. Hence, it can be concluded that without simplifying the assumption there is enough sensitivity in the combined Higgs measurements to constrain the nearly full coupling space. Additionally, upper bounds are set on the diHiggs SM signal strength and extracted to be 2.4 (2.9) times the SM HH signal strength in the HH combination.

Chapter 8

Calibrating the boosted $X \to b\bar{b}$ tagger

A signature in the detector targeted by several Run 2 BSM searches at the LHC is a Higgs boson decaying at high p_T to a $b\bar{b}$ pair. An example when the boosted regime becomes highly relevant is diHiggs production, as even tiny modifications of the couplings involved in the process or resonantly decaying new particles can lead to a heavier expected mass spectrum. A notable example is the modification of the HHVV coupling in the diHiggs production through the VBF channel that can lead to an enhancement in the number of Higgs bosons decaying at high p_T. One of the main challenges in such analyses is to efficiently identify the boosted $b\bar{b}$ final state. The convention in ATLAS is to reconstruct the whole decay within a single large-R jet to which VR track jets are ghost associated. Initially, only VR subjets have been used for b-tagging, with single taggers such as the DL1r (described in Section 5.4). A large-R jet is double-b tagged, provided the two VR subjets leading in p_T satisfy a DL1r score requirement to be b-tagged. However, this proved to have a poor signal efficiency and background rejection above approximately 500 GeV, as mainly low p_T jets have been used for optimising the low-level taggers and any information in the large-R jet system is neglected.

To recover some efficiency lost while keeping the background rejection high, a specialised tagger, *The Boosted $X \to b\bar{b}$ Tagger*, that aims to identify the boosted

Higgs boson decaying into a pair of b-quarks has been developed by the ATLAS Experiment [193]. The first section of the chapter briefly outlines the double-b tagger and its performance. The second half of the chapter is dedicated to the tagger's calibration, focusing on deriving scale factors for the mis-tag rate of the multijet background.

Acknowledgements and Author's Contributions

The author of this thesis made significant contributions to the *In-situ* calibration of the mis-tag rate of the multijet background. Only a fraction of the contributions are presented in Section 8.2. The author's contributions include the setup of the flavour correction fit outlined in Section 8.2.4, introducing the secondary vertex invariant mass in Section 8.2.3 and studying the impact of the flavour mapping on the templates. The author's helped to define the scale and parton shower uncertainties, implemented the systematics into the fit and study their impact, parts are outlined in Sections 8.2.6 and 8.2.7. Lastly, the author has studied the switch to the secondary vertex invariant mass as the discriminant variable to be used to derive the scale factors, which are shown in Section 8.2.7. The author also performed several other studies not present in this thesis, which include studying the flavour correction fit in bins of signed transverse impact parameter significance and validating the calibration method by deriving the scale factors with the DL1r tagger among others.

8.1 The Tagger

The boosted $X \to b\bar{b}$ tagger is a feed-forward neural network (NN) that combines the flavour information of up to three VR subjets with the kinematics of the large-R jet to determine the origin of the latter. The main backgrounds to the $H \to b\bar{b}$ signal are top and multijet events. Consequently, the NN is trained to classify the large-R jet as

a signal, $t\bar{t}$, or QCD jet. To be precise a signal (top) jet is defined as a large-R jet passing a set of kinematic requirements, see Section 8.1.2, with one ghost-associated Higgs (top quark), while a QCD jet is the large-R jet leading in p_T and passing the kinematic requirements from the multijet sample. The general reference for this section is [193].

8.1.1 Data and Simulations

The signal, $t\bar{t}$, and multijet samples used for training and validation of the NN are all generated with the ATLAS A14 tune and NNPDF2.3 LO PDF set. For showering and hadronisation, Pythia 8.186 is used. The signal events are simulated through the decay of the Randall-Sundrum (RS) graviton [194] to a pair of Higgs bosons, of which each decays into a $b\bar{b}$ final state with MadGraph5 at leading order. The parton shower and hadronisation are generated with Pythia 8.186 [195] with the ATLAS A14 tune [143] and the NNPDF2.3 LO PDF [142]. Graviton masses between $300\,\mathrm{GeV}$ and $6\,\mathrm{TeV}$ are generated. An event in the signal sample is used in the training only if a Higgs boson can be matched to the reconstructed large-R jet. The $Z' \to t\bar{t}$ production, with Z' invariant masses between $750\,\mathrm{GeV}$ and $5\,\mathrm{TeV}$, is used to simulate the top background. On the other hand, multijet samples are generated in different p_T regimes at the particle level to ensure sufficient statistics in the high p_T range. Pythia 8.186 with the ATLAS A14 tune and the NNPDF2.3 LO PDF is used to simulate the background processes.

8.1.2 Object Definition and Event Selection

The large-R jet is reconstructed from topo-clusters and required to have $250\,\mathrm{GeV} < p_T < 3000\,\mathrm{GeV}$ and $|\eta| < 2$. The VR subjets are clustered from tracks passing a set of criteria specified in [193] and need to consist of at least two tracks with $p_T > 7\,\mathrm{GeV}$ and $|\eta| < 2.5$. Up to three VR jets ghost-associated with the large-R jet are

considered for training the tagger. If more than three subjets are found, the ones leading in p_T are picked. Details on the reconstruction and calibration of the large-R jet and VR subjets can be found in Section 5.3.

The probabilities for the VR subjet to be a b-, c-, and l-jet are extracted from the DL1r algorithm. This flavour information in addition to the p_T and η of the large-R jet is used for training the $X \to b\bar{b}$ algorithm. To ensure no bias, the large-R jet p_T and η distributions must have the same shape in all samples, which is achieved through downsampling. A detailed description of the NN's architectures can be found in [193].

8.1.3 Double-b Tagging Performance

The outputs of the NN are probabilities of the large-R jet to be a signal (Higgs) jet (p_{Higgs}), top jet (p_{top}), and QCD jet (p_{QCD}). Inspired by the log-likelihood ratio used to define the DL1r discriminant for single b-tagging, see Equation 5.7, the three probabilities are combined to form the double-b tagging discriminant

$$D_{Xbb} = ln \frac{p_{Higgs}}{f_{top} p_{top} - (1 - f_{top}) p_{QCD}}, \tag{8.1}$$

where f_{top} is the expected fraction of top events. The D_{Xbb} distribution for each of the processes trained on is shown in Figure 8.1 and shows good discrimination. A loose, medium, and tight working point (WP) are defined, which specify the D_{Xbb} threshold for double-b tagging to achieve a predefined signal tagging efficiency. The signal efficiency is calculated as the ratio of the number of double-b tagged Higgs jets divided by all selected Higgs jets, given the above-described selection with an additional requirement on the invariant Higgs mass to be between $75\,\mathrm{GeV}$ and $145\,\mathrm{GeV}$ [193]. For use in analyses $f_{top} = 0.25$ and working points at 50%, 60%, and 70% tagging efficiencies are defined, see Table 8.1.

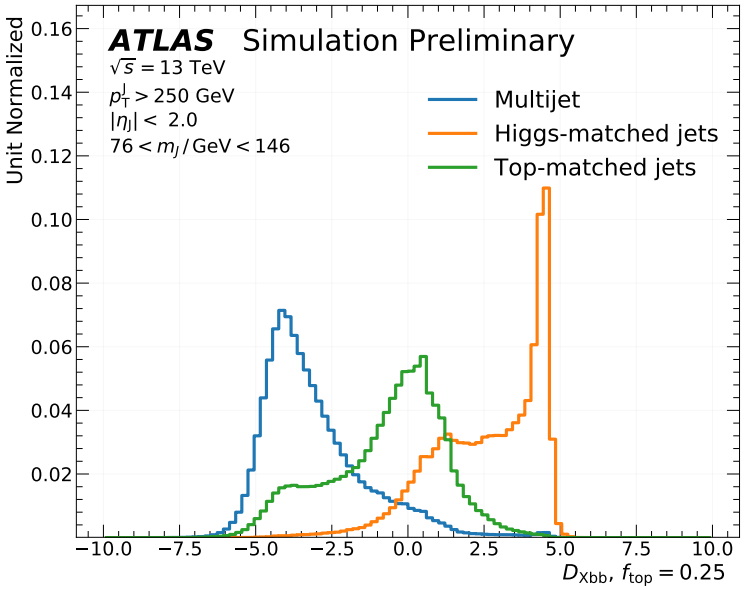


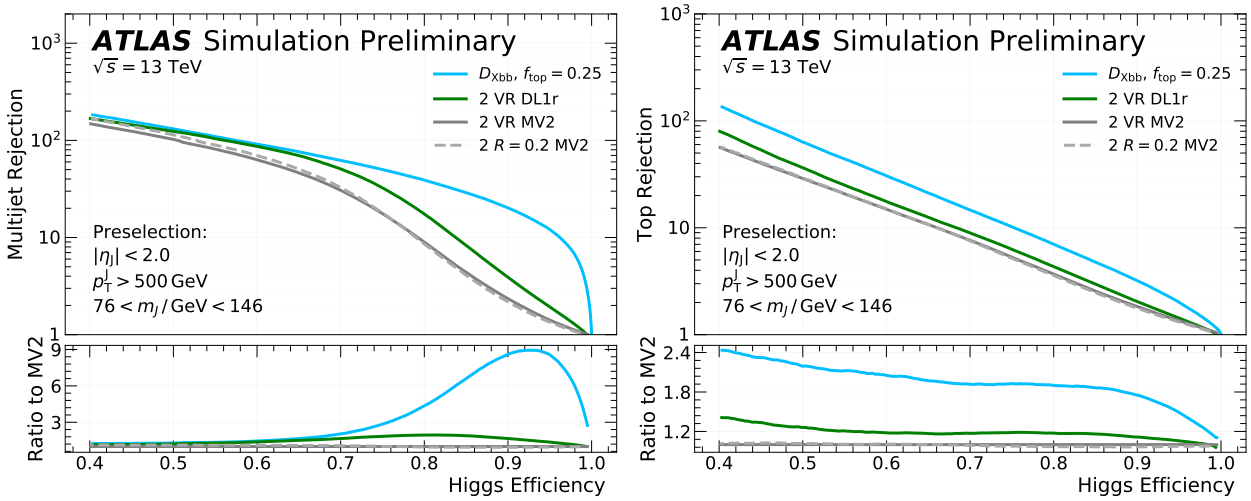
Figure 8.1: The distributions of D_{Xbb} with $f_{top} = 0.25$ for the Higgs, multijet, and top jets. The multijet and Higgs jets are well separated by D_{Xbb}, while the Higgs and top jets D_{Xbb} distributions overlap more. The graphic is taken from [193].

Working Point	50%	60%	70%
D_{Xbb} Threshold	3.13	2.55	1.92

Table 8.1: The working points and corresponding threshold values on the D_{Xbb} score are defined at $f_{top} = 0.25$.

A quantitative representation of the relationship between the ability of the tagger to reject the background while maintaining a high double-b tagging efficiency is shown in Figure 8.2. The performance of the double-b tagging using D_{Xbb} compared to the standard (single) taggers used during Run 2, MV2 and DL1r[1] [123]. In the range of the considered WPs the multijet rejection is similar for the boosted $X \to b\bar{b}$ tagger and DL1r, while the top rejection improves with the D_{Xbb} tagger.

[1] Only the comparison with DL1r will be discussed in the text, as by the time of writing MV2 is obsolete.



(a) (b)

Figure 8.2: The multijet rejection in (a) and top rejection in (b) are shown as functions of the Higgs identification efficiency. All jets are required to have a $p_T > 500\,\mathrm{GeV}$ to capture the performance in the boosted regime better. The figure is taken from [193].

8.2 Calibrating the mis-tag rate of the multijet background

A calibration is needed to correct for the difference in the taggers' performance on data and MC by deriving scale factors (SF) in bins of large-R jet p_T. The response of the boosted $X \to b\bar{b}$ tagger on data is unknown, as the algorithm is trained solely on MC simulations. The SFs for the signal efficiency are derived using $Z(\to b\bar{b}) +$ jets and $Z(\to b\bar{b})\gamma$ samples, while the SFs for the mis-tag rate of the two background processes are derived from $t\bar{t}$ and multijet samples independently. The preliminary calibrations of the signal and $t\bar{t}$ background are described in [4], while the rest of the chapter focuses on the rather complicated calibration of the multijet background (the event selection closely follows the one already presented in [4]).

The mis-tag rate of the QCD background is extracted from dijet MC samples enriched in $g \to b\bar{b}$ splittings that are primarily misidentified as signal events. The mis-tag efficiency, in data and MC, is

$$\epsilon_{data} = \frac{N_{data}^{BB,tag}}{N_{data}^{BB}} \tag{8.2}$$

$$\epsilon_{MC} = \frac{N_{MC}^{BB,tag}}{N_{MC}^{BB}} \tag{8.3}$$

respectively. The notation $N_{<data/mc>}^{BB,<tag/"">}$ is read as the number of true $g \to b\bar{b}$ events in data/MC that are double b-tagged/before any tagging. The main challenge of the calibration is to determine the number of true BB events in data. The problem seems trivial at first, as the number of BB events pre-tag and post-tag is known in MC. However, it turns out the flavour fractions are not correctly modelled by the dijet simulation and a per flavour ($flav$) normalisation correction factor (NF_{flav}) has to be extracted from a fit to data. It follows that

$$N_{data}^{BB} = N_{BB}^{MC} \times NF_{BB}, \tag{8.4}$$

$$N_{data}^{BB,tag} = N^{\text{data,tag}} - \sum_{flav \in \text{non-BB flav.}} N_{flav}^{\text{MC,tag}} \times NF_{flav}. \tag{8.5}$$

The SF then becomes

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}} = \frac{N^{\text{data,tag}} - \sum_{flav \in \text{non-BB flav.}} N_{flav}^{\text{MC,tag}} \times NF_{flav}}{N_{BB}^{\text{MC,tag}} \times NF_{BB}}. \tag{8.6}$$

8.2.1 Data and Monte Carlo Simulations

The full Run 2 dataset collected from proton-proton collisions by the ATLAS Experiment at $\sqrt{s} = 13\,\text{TeV}$ in the years 2015-2018 is used, which results in the total integrated luminosity $139\,\text{fb}^{-1}$.

Pythia 8.235 with NNPDF2.3LO PDF set is used to generate dijet events which are passed through the GEANT4 full ATLAS detector simulation. To enrich the sample in $g \to b\bar{b}$ an additional sample is generated with a filter that requires a muon with $p_T > 3\,\text{GeV}$ and $|\eta| < 2.8$ at the event generation level. This sample is called the muon-filtered sample, as it has a substantial number of semi-muonically decaying B-hadrons from the $g \to b\bar{b}$ splitting. To allow the combination of the two samples

HLT Trigger	Offline Trigger Jet p_T [GeV] Threshold			
	2015	2016	2017	2018
HLT_j420	-	-	463.5	454.5
HLT_j380	-	436.0	-	-
HLT_j360	416.5	416.5	395.5	388.5
HLT_j260	297.5	297.5	282.5	282.5
HLT_j175	205.5	205.5	-	-

Table 8.2: The triggers used per year and the corresponding offline trigger jet p_T thresholds. The HLT_{j420} is unprescaled and every other trigger is prescaled.

and increase the MC statistics, events that have a truth muon meeting the requirements placed by the filter in the non-filtered samples are removed on the calibration level. To derive the systematics on the parton shower and hadronisation model an alternative sample is produced with Herwig 7.1.6 at NLO.

8.2.2 Event Selection

The object definition and event selection closely follow the one used for training the tagger, which is briefly outlined in section 8.1.2. The events are initially selected by applying the logical OR on a set of small-R jet triggers, listed in Table 8.2. The table also states the offline p_T thresholds derived per year of data taking, which needs to be passed by an offline calo small-R jet matched to the online trigger jet. This additional requirement is in place to ensure the full efficiency of the fired triggers. All but HLT_j420 are prescaled, meaning that only every p$^{\text{th}}$ triggered event, where p is the prescale, is recorded. Therefore the data needs to be weighted by the combined effect of the prescaled triggers fired. The calibration-specific cuts are that the large-R jet with $250 < p_T < 1000\,\mathrm{GeV}$, $m > 50\,\mathrm{GeV}$, and $\eta > 2.0$ have to be separated from the trigger jet by at least $\Delta R > 1.5$. Mimicking the training dataset, the large-R jet is required to be ghost-associated to at least two VR track jets with $p_T > 7\,\mathrm{GeV}$ and $\eta < 2.5$. The whole event is vetoed if overlapping VR jets are found. At least two

tracks need to be associated with the VR jet, which have $p_T > 0.5\,\mathrm{GeV}$, $|d_0| < 5.0\,\mathrm{mm}$, and $|z_0 sin\theta| < 3.0\,\mathrm{mm}$. The additional tracking requirements are in place to mitigate pile-up while keeping large IP tracks originating from B-hadron decays. The VR jets are synonymous with subjets, and the two terms are used interchangeably.

To increase the fraction of b-jets in the simulated sample without biasing the output of impact parameter based taggers a non-isolated muon with $p_T > 10\,\mathrm{GeV}$, $\eta < 2.4$, $|d_0| < 2.0\,\mathrm{mm}$ and $|z_0 sin\theta| < 2.0\,\mathrm{mm}$ passing the medium muon WP is required to be matched to one of the subjets. The muon is matched to the VR jet, provided it is within its radius. The VR jet with the associated muon is referred to as the muon-jet (j_μ), while among the remaining subjets, the one leading in p_T is picked as the non-muon jet ($j_{non-\mu}$). If more than one VR jet has a muon matched to it, the leading p_T subjet is j_μ and the sub-leading subjet is $j_{non-\mu}$. A truth label based on cone truth matching is given to the j_μ, $j_{non-\mu}$ and if present the third subjet leading in p_T (j_3) in MC as described in Section 5.4, which classifies the jet as a b-, c- or l-jet[2].

It follows that there are 36 possible flavour compositions of the large-R jet. A flavour merging scheme is applied to reduce the number of flavour categories to only five (di-flavour) categories, as several of the flavour compositions have a contribution smaller than 1%. Each simulated large-R jet is categorized as either a BB, BL, CC, CL, or LL jet based on the number of b-, c-, and l- subjets in the large-R jet. For instance, in the BB category, the event has a large-R jet with two or more b-subjets. Table 8.3 breaks down the flavour mapping scheme, and Table 8.4 shows the flavour composition of the MC after applying the flavour merging. The MC simulation after the event selection contains a sufficient fraction of the BB component, around 16% inclusively, to perform the calibration.

The benefit of basing the flavour mapping on the number of b-, c-, and l-jets in the large-R jet is the similarity of the D_{Xbb} distributions between the merged three flavour compositions. The downside of the method is that as only j_μ and $j_{non-\mu}$ are

[2]If one of the VR jets is a tau jet the event is discarded.

Flavor Category	Flavour Content
BB	BBB,BBC,BBL,BBx,BCB,BLB,CBB,LBB
BL	BCC,BCL,BCx,CBC,CBL,CBx,CCB,BLC,BLL,BLx,LBC,LBL,LBx,CLB,LCB,LLB
CC	CCC,CCL,CCx,CLC,LCC
CL	CLL,CLx,LCL,LCx,LLC
LL	LLL,LLx

Table 8.3: The flavour mapping used for merging the flavour into the five categories in the right column. The right column lists the flavour content of the large-R jet that is merged into the same category. The three letters indicate the truth flavour of $(j_\mu, j_{non-\mu}, j_3)$, where j_3 is the third jet, respectively. The absence of the third sub-jet is indicated by x.

Flavor Category	Flavor Fraction (%)
BB	16
BL	22
CC	11
CL	23
LL	28

Table 8.4: The flavour composition in the MC after applying the flavour merging.

used to correct the flavour composition, and there is degeneration in the flavour purity of each of the jets considered. It follows that there is less discrimination power between the flavour categories as compared to alternative flavour merging schemes. Notably, the $j_{non-\mu}$ of the BL category contains about 15% b- and c- jets while the remainder are l-jets. Similarly, the $j_{non-\mu}$ in the BB and CC categories is not purely made of b-jets and c-jets respectively, the l-jet component is about 22%.

Figure 8.3 shows the D_{Xbb} distributions of the five flavour categories after the event selection. The three vertical dotted lines indicate the double-b tagging cuts applied for three WPs, events on the right of the line fall into the pass region (are double-b tagged). It can be seen that the pass region is very pure in BB events, with a small fraction of BL and CC events while barely any LL events pass the double b-tagging requirement.

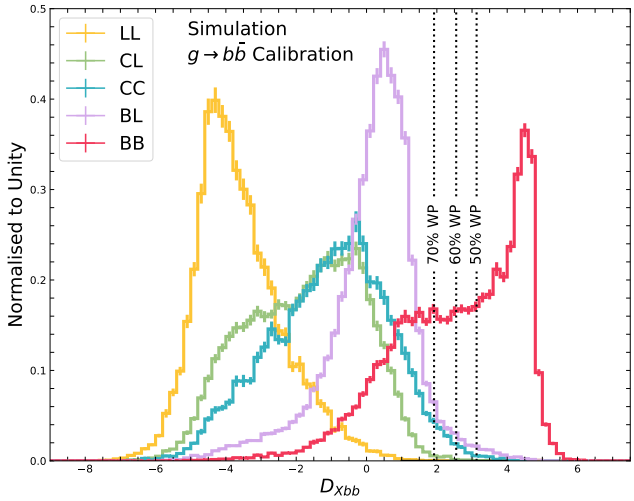


Figure 8.3: The D_{Xbb} distributions in dijet MC of the five flavour categories. The three dotted vertical lines indicate the cuts applied when the 70%, 60% and 50% WP are chosen for double-b tagging. The events on the right of the vertical line are said to be double-b tagged at the 70%, 60% or 50% WP (depending on the cut considered).

8.2.3 Flavour Discriminating Variables

Signed Transverse Impact Parameter Significance

The signed transverse impact parameter significance of a track is defined as

$$s_{d_0} = s_j \frac{d_0}{\sigma_{d0}}, \tag{8.7}$$

where d_0 is the transverse impact parameter relative to the primary vertex (PV), σ_{d0} denotes its uncertainty, and s_j is its sign. As depicted in Figure 8.4, tracks crossing the jet axis behind the PV are assigned a value of $s_j = -1$, whereas tracks crossing the jet axis in front of the PV have $s_j = 1$. B- and C-hadron decays are more likely to have a positive s_j, due to the hadrons carrying a large fraction of the jet's p_T and their long lifetime [4, 196]. On the other hand, the s_j of l-jets depends on the resolution of the measured tracks, which leads to a more or less symmetric s_{d0} distribution. To mitigate the impact of long-lived light hadrons, such as K_s, that can

skew the l-jet s_{d0} distribution towards large positive values the calibration uses mean s_{d0} ($< s_{d0} >$) constructed from three tracks leading in p_T. Only tracks ghost-associated with the jet are assigned s_{d_0}.

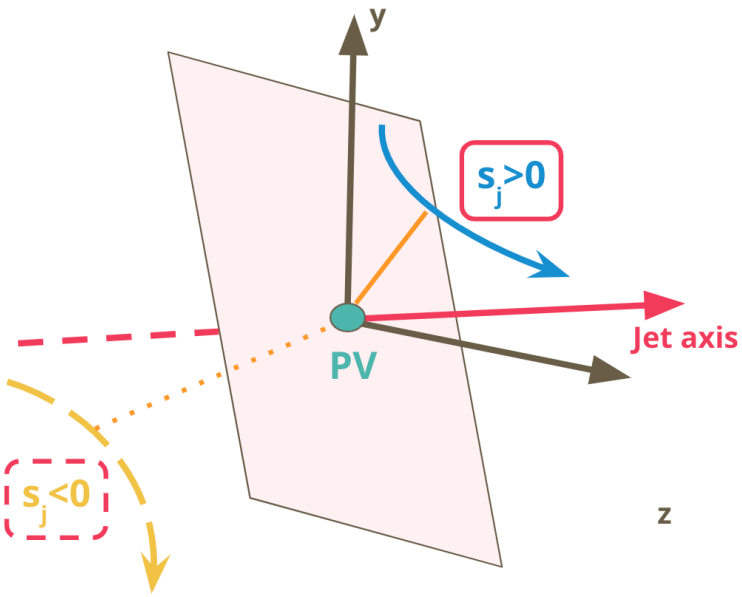
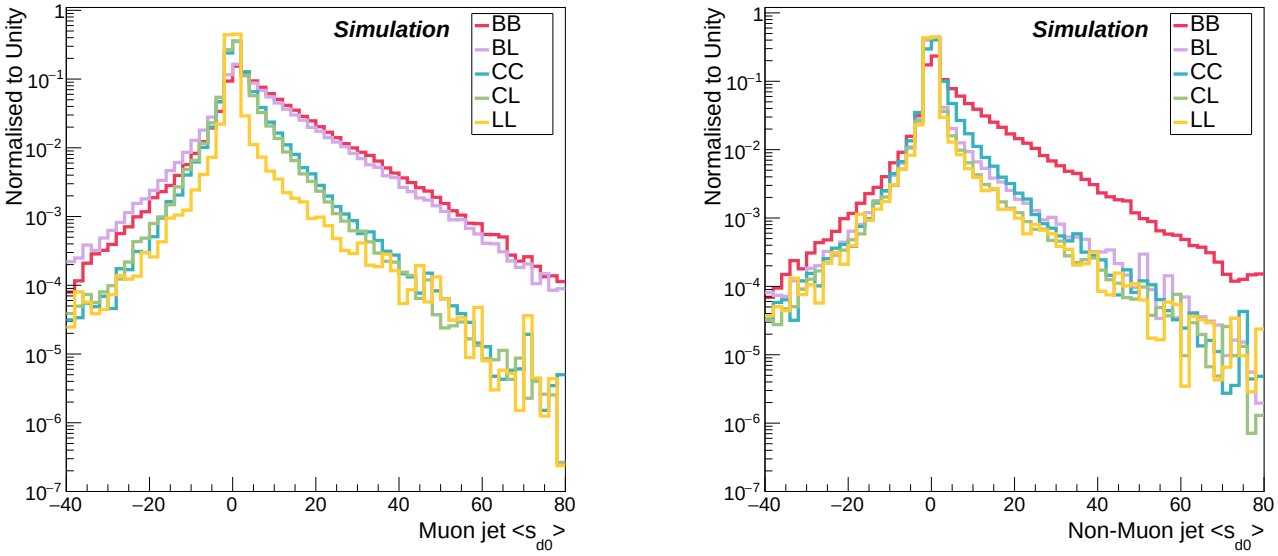


Figure 8.4: The method for determining the sign of the transverse impact parameter, denoted as s_j, for tracks that are ghost-associated with a jet. If the angle between the jet axis and the line connecting the closest approach point to the PV is less than $90°$, then s_j is positive. In the sketch, the blue track has $s_j > 0$, while the yellow track has $s_j < 0$. The shown plane is perpendicular to the jet axis.

The different masses and consequently lifetimes of the hadrons lead to excellent flavour discrimination power between b-, c- and l-jets. The $< s_{d0} >$ templates for each flavour category in j_μ and $j_{non-\mu}$ can be seen in Figure 8.5. The flavour discrimination power of $< s_{d0} >$ is evident from the exponentially falling tails, which can differ by orders of magnitude depending on the flavour category. Taking a closer look at the muon jet the BB and BL templates are characteristic for b-jets, the CC and CL templates for c-jets, and the LL template for l-jets. Thus, to distinguish between the five large-R jet flavours categories the combined information from the muon and non-muon jet is needed.

As pointed out before, the definition of the flavour categories diminishes the flavour discrimination power, which is evident from the somewhat surprising behaviour in



(a) Muon jet.

(b) Non-muon jet.

Figure 8.5: The templates of the simulated $< s_{d0} >$ distributions for each flavour category in the inclusive large-R jet p_T bin.

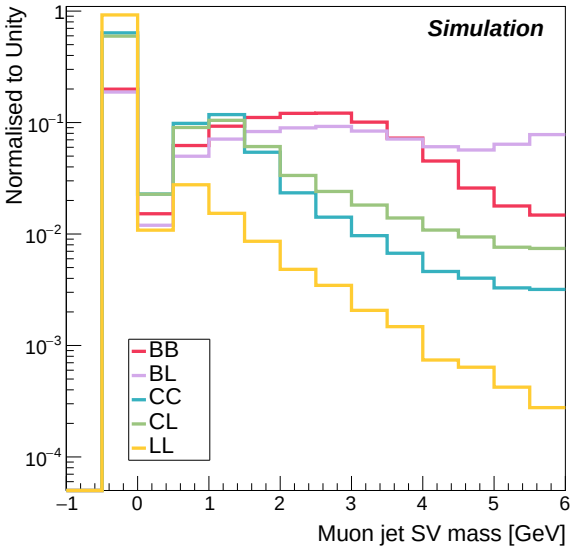
the $j_{non-\mu}$. Compared to the j_μ BB template, the peak of the BB template in the $j_{non-\mu}$ is more pointed, and the BL template has heavier tails than expected. The reason for this is that the $j_{non-\mu}$ BB and BL templates do not exhibit the typical characteristics of b- and l-jets purely, respectively, which can be explained by examining the flavour mapping outlined in Table 8.3 more closely. The flavour categorisation of the large-R jet is based on the total number of B and C subjets, which means that the subject templates are not purely one flavour. In particular, the $j_{non-\mu}$ flavour categories represent different compositions of b-, c-, and l-jets. For example, the muon-jet BB category is very pure in b-jets as only the BLB composition with the second B in the third subject has a non-negligible contribution. On the other hand, the BLB component explains the less prominent peak and less hard positive tail in the $j_{non-\mu}$. Similarly, the shape of the BL template in $j_{non-\mu}$ can be understood by the LBx, LBL, BCx, and BCL contributions. The distinction between CC and CL becomes especially challenging, the $j_{non-\mu}$ CC templates include a non-negligible fraction of CLC events, and the CL template includes LCL events, making the templates more similar. In addition, the two templates are nearly identical in j_μ.

Secondary Vertex Mass

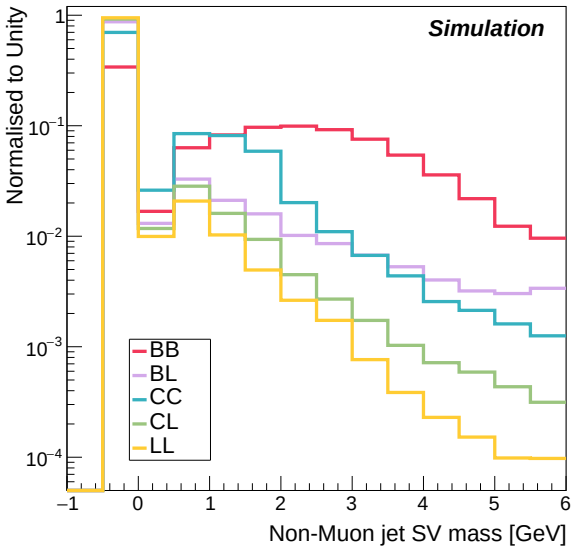
Another way to identify B-hadrons is through the mass of their decay products coming from their secondary vertex (SV), whose reconstruction is outlined in Section 5.4. A higher track multiplicity and a heavier mass spectrum are expected from b-jets, while l-jets primarily do not have a reconstructable SV[3] or are mistakenly associated with fake vertices. In case case no SV is found a negative value is assigned to the jet. The c-jets are an in-between case, as the C-hadrons' lifetime varies. Its behaviour can be either l-jet or b-jet like. The SV invariant mass, reconstructed and exploited by the low-level SV1 tagger, is calculated from all the tracks associated with the found SV. Its flavour discrimination power has been used for fit validation. However, more recently, it has also been tested as the fitted variable in the flavour correction fit described in the next section. .

Figure 8.6 shows the SV mass flavour templates in the muon and non-muon jet after the flavour merging. The BB, CC, and LL templates in both subjets agree well with the expected properties for b-,c-, and l-jets respectively. Similarly, the CL template agrees with the standard c-jet template in j_{μ}, while being very alike to the l-jet template in $j_{non-\mu}$. On the other hand, the BL template has features not observed in the standard b-, c- or l-jet templates. In j_{μ}, the number of BL events peaks around $3\,\mathrm{GeV}$ and $6\,\mathrm{GeV}$, contrary to the single peak observed in the characteristic b-jet template. The increasing number of events at masses above $4.5\,\mathrm{GeV}$ is the result of highly collimated $g \to b\bar{b}$ decays. The two b-jets fail to be reconstructed as resolved VR subjets, so tracks from both B-hadron decays are matched to the same SV resulting in a higher than expected SV mass. In $j_{non-\mu}$, similarly, as in the corresponding $< s_{d0} >$ distribution, a heavier tail is observed than expected. The significant enhancement in the tails might seem surprising, as the composition of

[3]The Single Secondary Vertex Finder (SVF) [126] applies a cleaning that removes track pairs originating from light long-lived particles, photon conversion, and material interactions. Light long-lived particles are vetted based on the reconstructed invariant mass of track pairs. Fake vertices originating from material interactions are removed by comparing the radial position of the track pairs in comparison to the location of the pixel detector and beam pipe.



(a) Muon jet.



(b) Non-muon jet.

Figure 8.6: The simulated templates of the SV mass distribution for each flavour category in the inclusive large-R jet p_T bin. The enhanced tail at masses above $3.5\,\mathrm{GeV}$ in the muon jet, see the plot in (a), is observed only in the BL category. The source of the enhancement is the reconstruction of $g \to b\bar{b}$ splittings as a single subject.

the BL template in $j_{non-\mu}$ is about 85% l-jets, 8% b-jets, and 7% c-jets. Although the number of b-jets is small, their signature is prominent, as they make a significant contribution to the high-mass region. It is clear from Figure 8.7 that the l-jets barely contribute events above $3\,\mathrm{GeV}$, while the b-jets have an enriched tail from the large fraction of double-B VR jets. As is the case for the $< s_{d0} >$ discriminant, combining the information from the muon and non-muon jet breaks the degeneracy between the flavour templates, and all five flavour categories can be distinguished.

8.2.4 Flavour Fraction Correction

The simulated flavour composition is corrected by extracting the flavour normalisation factors (NF) for each flavour category from a binned profiled likelihood fit, described in Section 6.7. The flavour templates are fitted to data in four regions; the pass muon jet, the fail muon jet, the pass non-muon jet, and the

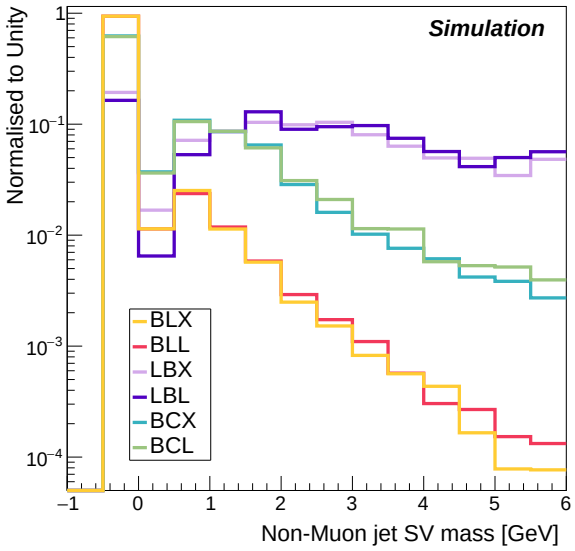


Figure 8.7: A selection of templates that are merged into the non-muon jet BL template. The second letter on each label indicates the truth flavour of the non-muon jet. Only a few large-R jet flavour compositions that are merged into the BL category are plotted for clarity.

fail non-muon jet region independently in a number of large-R jet p_T bins. The pass regions contain only events that are double-b tagged, while the fail regions contain the events that fail the double-b tagging. This additional categorisation of events enhances the flavour discrimination power of the templates, as the pass region is very pure in BB events, and enables the direct extraction of the SF. The distributions in the pass (fail) j_μ and $j_{non-\mu}$ regions share the same events but as they are found to be uncorrelated this is not an issue for the fit setup. Thus, the total likelihood is obtained by multiplying the likelihood function of each region, see Equation 6.15. The correlation of the pass and fail regions is taken care by applying the SF,

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}}, \tag{8.8}$$

in the pass regions and anti-SF,

$$anti-SF = \frac{1-\epsilon_{data}}{1-\epsilon_{MC}} = \frac{1-\epsilon_{MC}SF}{1-\epsilon_{MC}}, \tag{8.9}$$

in the fail regions to the BB MC templates, where ϵ_{data} (ϵ_{MC}) is defined in Equation 8.2 (8.3). It follows that the expected rate of events in bin k belonging to either the j_μ or $j_{non-\mu}$ pass region is

$$\lambda_k^{pass} = SF \times NF_{BB} \times \lambda_k^{\text{BB}} + \sum_{flav \in \{BL,CC,CL,LL\}} NF^{flav} \times \lambda_k^{flav}, \tag{8.10}$$

and in bin i in the fail region is

$$\lambda_i^{\text{fail}} = anti-SF \times NF_{BB} \times \lambda_i^{\text{BB}} + \sum_{\text{flav} \in \text{BL,CC,CL,LL}} NF^{\text{flav}} \times \lambda_i^{\text{flav}} \tag{8.11}$$

$$= \frac{1 - SF \times \epsilon_{MC}}{1 - \epsilon_{MC}} \times NF_{BB} \times \lambda_i^{\text{BB}} + \sum_{\text{flav} \in \text{BL,CC,CL,LL}} NF^{\text{flav}} \times \lambda_i^{\text{flav}}. \tag{8.12}$$

The rate per flavour category in a bin is λ_{bin}^{flav}, where $flav \in \{BB, BL, CC, CL, LL\}$. The index for the region and dependence on other nuisance parameters are omitted for clarity, i.e. $\lambda_i^{fail} = \lambda_i^{\text{fail region R}}(\boldsymbol{\theta})$ and $\lambda_k^{pass} = \lambda_k^{\text{pass region R}}(\boldsymbol{\theta})$ where R $\in \{j_{non-\mu}, j_\mu\}$.

It follows that the total likelihood function to be maximized is

$$\mathcal{L}(\boldsymbol{\Psi}, \boldsymbol{\theta}) = \prod_{p \in \{j_\mu^{pass}, j_{non-\mu}^{pass}\}} \prod_{k \in \text{bins}} \mathcal{P}\wr\int(n_{p,k} | \lambda_{p,k}^{pass}(SF, \boldsymbol{\theta})) \tag{8.13}$$

$$\times \prod_{f \in \{j_\mu^{fail}, j_{non-\mu}^{fail}\}} \prod_{i \in \text{bins}} \mathcal{P}\wr\int(n_{f,i} | \lambda_{f,i}^{fail}(SF, \boldsymbol{\theta})) \tag{8.14}$$

where the $\boldsymbol{\theta}$ are the nuisance parameter (NPs) that can modify the expected rate and include the NF_{flav}. The constrained term is omitted for clarity. The NPs consist of the systematic uncertainties (described in the next Section) and the per di-flavour category NFs whose impact on the rate is shown in Equations 8.10 and 8.11. Varying the NFs leads to changes in the shape of the total MC distribution, as changing the contributions of the flavour templates can modify the shape of the total MC distribution. The NF values that maximise the likelihood function in a profiled likelihood fit give the MC shape that best agrees with the data. Assuming that there is enough discrimination in the fitted region between the five di-flavour templates

the post-fit flavour composition is unique and corresponds to the data. If the post-fit NF is unity, the contribution of the simulated di-flavour category is correct and kept to its pre-fit value.

The observed data used to correct the flavour fractions in the MC is weighted, due to the use of prescaled triggers. Therefore, it is not guaranteed that the data is Poisson distributed as assumed in the profiled likelihood fit. For Equation 8.13 to hold the bin of each contribution is scaled by $\frac{n}{\sigma_n^2}$, where n is the number of data events in a bin and σ_n the uncertainty on n [197]. Given that there are at least ten data events per bin this scaling ensures that the uncertainty on the data per bin is $\sqrt{n}$.

MINUIT is used to minimise the negative log-likelihood of 8.13 in a profiled maximum likelihood fit and the errors on the NPs are estimated by the MINOS technique [198]. The NPs are correlated across regions and in general across flavour categories. Each large-R jet bin is fitted independently with TRExFitter, which is built on HistFactory [173].

SFs are extracted in independent large-R jet regions defined by the large-R jet p_T with the bin ages: $280\,\mathrm{GeV}$, $310\,\mathrm{GeV}$, $340\,\mathrm{GeV}$, $380\,\mathrm{GeV}$, $420\,\mathrm{GeV}$, $460\,\mathrm{GeV}$, $500\,\mathrm{GeV}$, $550\,\mathrm{GeV}$, $600\,\mathrm{GeV}$, $750\,\mathrm{GeV}$, $1000\,\mathrm{GeV}$. Each of the p_T bins below $p_T < 420\,\mathrm{GeV}$ are populated by a single trigger.

8.2.5 Pre-fit Distributions

The agreement between the observed data and MC SV invariant mass distribution is fairly reasonable in each region used by the calibration, as can be seen in Figures 8.8-8.11. Across the large-R jet p_T bins a similar degree of mis-modeling is observed in all pass j_μ, fail j_μ, and pass $j_{non-\mu}$ regions. The fail $j_{non-\mu}$ regions display a greater degree of mis-modelling in large-R jet p_T bins with p_T below $600\,\mathrm{GeV}$. The source of the observed discrepancy in the tails of SV mass distributions is likely a p_T dependent mis-modelling of the number of $g \to b\bar{b}$ events reconstructed as a single

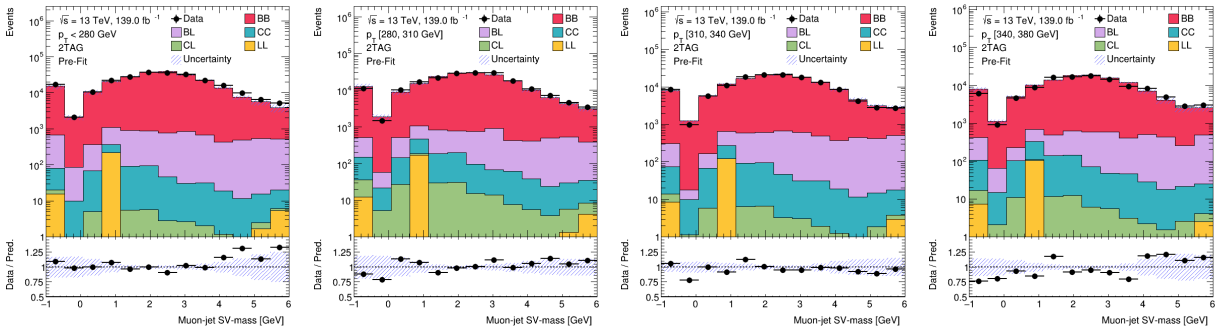
VR jet in $j_{non-\mu}$. As p_T increases, the fraction of b-jets in $j_{non-\mu}$ in the BL category increases while the number of l-jets decreases, resulting in an overall 5% change. Upon revisiting Figure 8.7, it becomes evident that even these small changes can have a significant impact on the number of events above $3\,\mathrm{GeV}$ as can be seen in Figure 8.13. The change of large-R jet flavour composition is reflected as a less prominent tail at lower p_T in the BL non-muon jet template. The mis-modelling increases as the same p_T dependent behaviour is not observed in data, where the shape of the distribution is similar across all large-R jet bins considered.

The pre-fit data-MC disagreement in $< s_{d0} >$ is quite large as compared to the SV mass, especially in the pass $j_{non-\mu}$ region shown in Figure 8.12. After several studies, it has been determined that the mis-modelling of $< s_{d0} >$ is predominantly the result of a lack of calibration for σ_{d0} rather than the incorrect flavour composition of the MC.

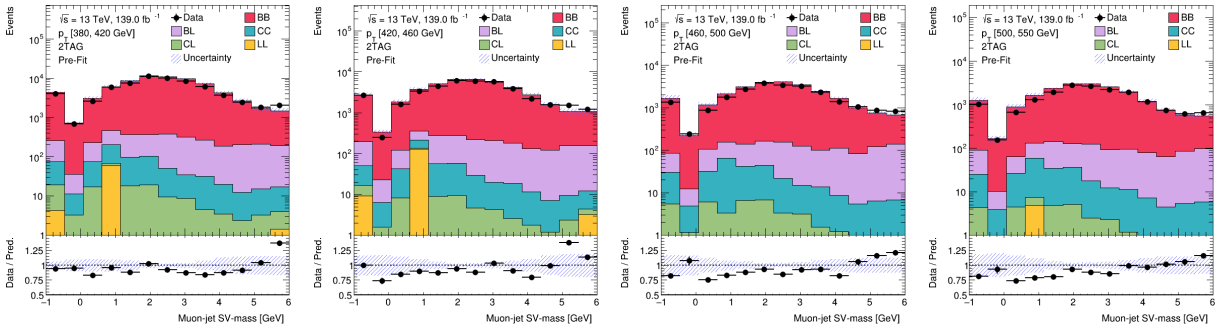
The degree of $< s_{d0} >$ mis-modelling renders it unsuitable for use in the calibration process. As a result, SV mass is undergoing testing as an alternative and the remainder of this chapter details the flavour correction fit utilising SV mass as the discriminating variable. The data-MC plots in the other $< s_{d0} >$ regions across the large-R jet p_T bins are shown in Appendix C.

8.2.6 Systematic Uncertainties

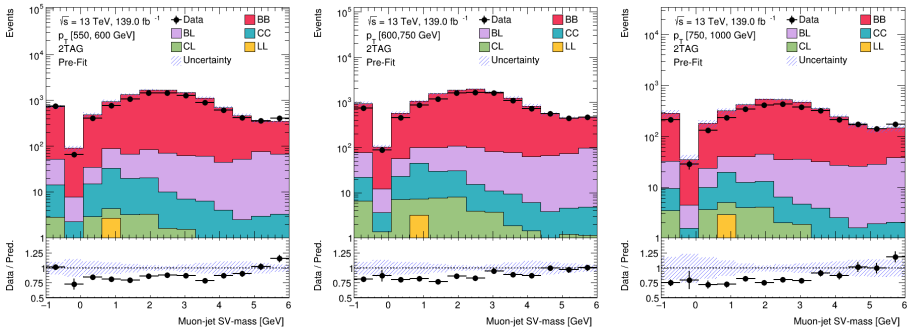
The systematic uncertainties that affect the flavour templates are described in this section. The strategy of defining and applying the majority of the experimental uncertainty on the luminosity, physics object reconstruction, parton shower, and renormalisation scale is the same or similar to the one used in the non-resonant $HH \to b\bar{b}b\bar{b}$ analysis described in Section 6.6. To minimise the noise levels in the templates the systematic uncertainties are smoothed with the ”353QH, twice” algorithm described in [199]. Pruning removes the normalisation or shape of a



(a) $280\,\text{GeV} < p_T$ (b) $280 < p_T < 310\,\text{GeV}$ (c) $310 < p_T < 340\,\text{GeV}$ (d) $340 < p_T < 380\,\text{GeV}$

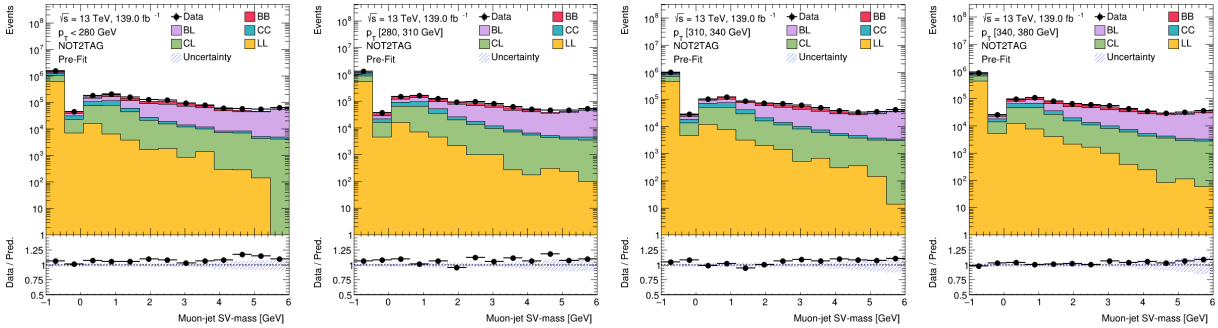


(e) $380 < p_T < 420\,\text{GeV}$ (f) $420 < p_T < 460\,\text{GeV}$ (g) $460 < p_T < 500\,\text{GeV}$ (h) $500 < p_T < 550\,\text{GeV}$

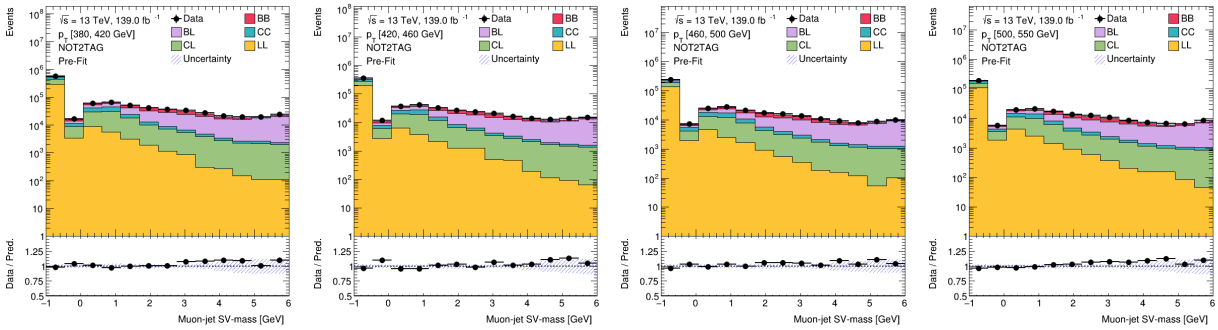


(i) $550 < p_T < 600\,\text{GeV}$ (j) $600 < p_T < 750\,\text{GeV}$ (k) $750 < p_T < 1000\,\text{GeV}$

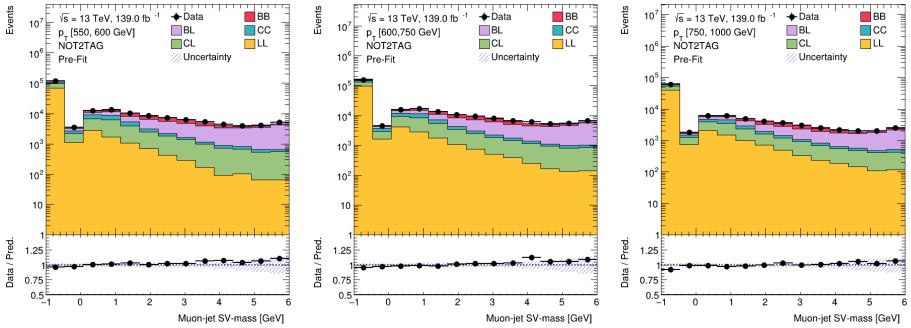
Figure 8.8: The pre-fit SV mass distribution in the pass region of the muon jet in different large-R jet p_T bins for the 50% WP.



(a) $280\,\text{GeV} < p_T$ (b) $280 < p_T < 310\,\text{GeV}$ (c) $310 < p_T < 340\,\text{GeV}$ (d) $340 < p_T < 380\,\text{GeV}$

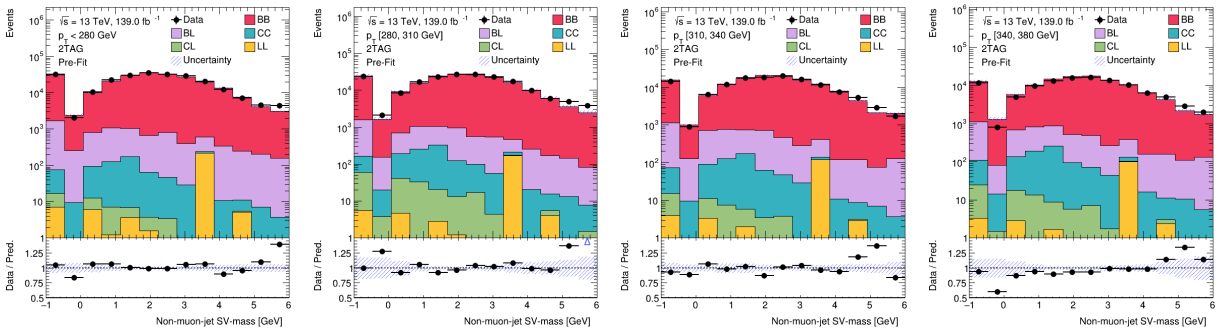


(e) $380 < p_T < 420\,\text{GeV}$ (f) $420 < p_T < 460\,\text{GeV}$ (g) $460 < p_T < 500\,\text{GeV}$ (h) $500 < p_T < 550\,\text{GeV}$

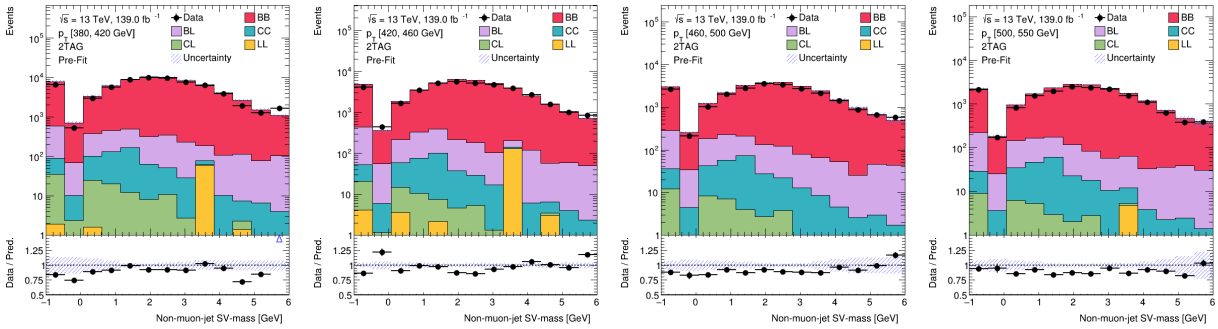


(i) $550 < p_T < 600\,\text{GeV}$ (j) $600 < p_T < 750\,\text{GeV}$ (k) $750 < p_T < 1000\,\text{GeV}$

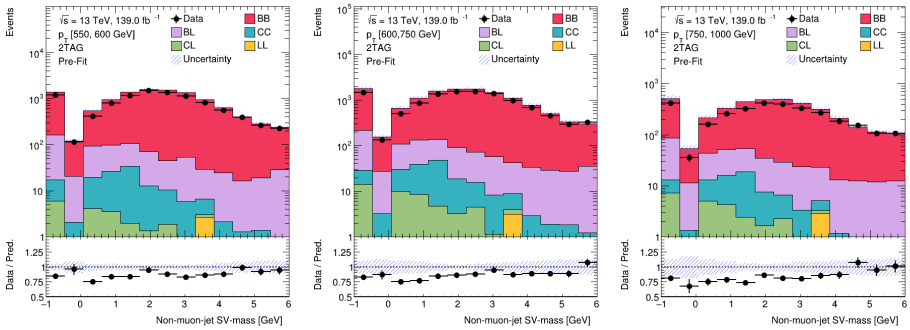
Figure 8.9: The pre-fit SV mass distribution in the fail region of the muon jet in different large-R jet p_T bins for the 50% WP.



(a) $280\,\text{GeV} < p_T$ (b) $280 < p_T < 310\,\text{GeV}$ (c) $310 < p_T < 340\,\text{GeV}$ (d) $340 < p_T < 380\,\text{GeV}$

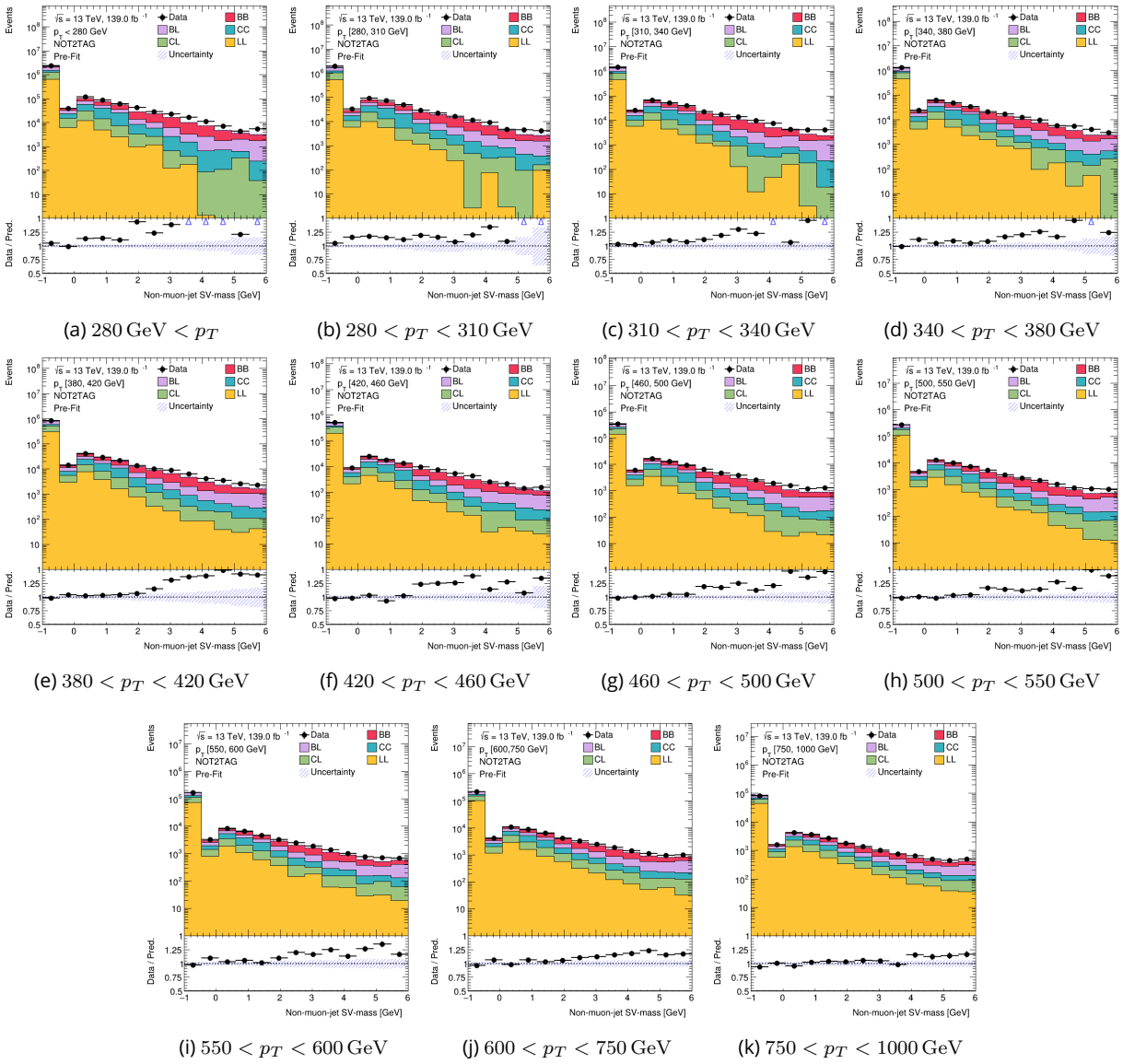


(e) $380 < p_T < 420\,\text{GeV}$ (f) $420 < p_T < 460\,\text{GeV}$ (g) $460 < p_T < 500\,\text{GeV}$ (h) $500 < p_T < 550\,\text{GeV}$



(i) $550 < p_T < 600\,\text{GeV}$ (j) $600 < p_T < 750\,\text{GeV}$ (k) $750 < p_T < 1000\,\text{GeV}$

Figure 8.10: The pre-fit SV mass distribution in the non-muon jet pass region in different large-R jet p_T bins for the 50% WP.

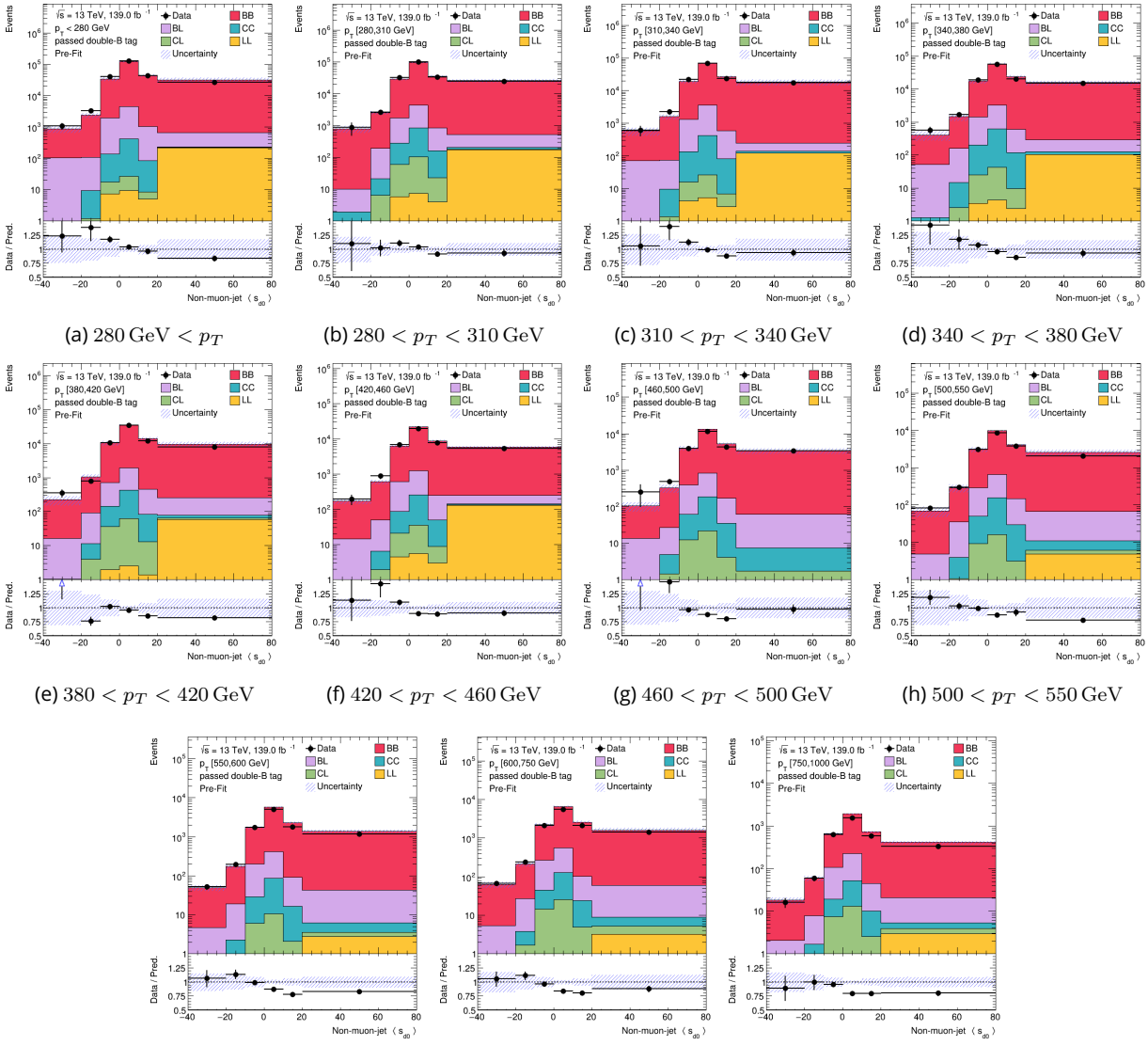


(a) $280\,\text{GeV} < p_T$ (b) $280 < p_T < 310\,\text{GeV}$ (c) $310 < p_T < 340\,\text{GeV}$ (d) $340 < p_T < 380\,\text{GeV}$

(e) $380 < p_T < 420\,\text{GeV}$ (f) $420 < p_T < 460\,\text{GeV}$ (g) $460 < p_T < 500\,\text{GeV}$ (h) $500 < p_T < 550\,\text{GeV}$

(i) $550 < p_T < 600\,\text{GeV}$ (j) $600 < p_T < 750\,\text{GeV}$ (k) $750 < p_T < 1000\,\text{GeV}$

Figure 8.11: The pre-fit SV mass distribution in the non-muon jet fail region in different large-R jet p_T bins for the 50% WP. Mis-modelling in the tail of the SV mass distribution can be observed across the large-R jet p_T bins.



(a) $280\,\text{GeV} < p_T$ (b) $280 < p_T < 310\,\text{GeV}$ (c) $310 < p_T < 340\,\text{GeV}$ (d) $340 < p_T < 380\,\text{GeV}$

(e) $380 < p_T < 420\,\text{GeV}$ (f) $420 < p_T < 460\,\text{GeV}$ (g) $460 < p_T < 500\,\text{GeV}$ (h) $500 < p_T < 550\,\text{GeV}$

(i) $550 < p_T < 600\,\text{GeV}$ (j) $600 < p_T < 750\,\text{GeV}$ (k) $750 < p_T < 1000\,\text{GeV}$

Figure 8.12: The $< s_{d0} >$ distribution in the non-muon jet pass region in different large-R jet p_T bins for the 50% WP.

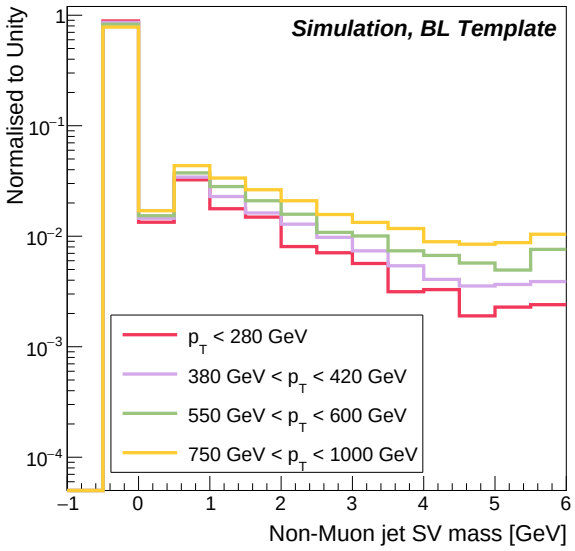


Figure 8.13: The large-R jet p_T dependence of the BL template shape. Only a few p_T bins are plotted for clarity.

systematic uncertainty per region if the variation is smaller than 0.5%.

Experimental Uncertainties

The physical objects used in the calibration are large-R jets, VR track jets, and muons, which come with several NPs to capture the detector and reconstruction uncertainties, as described in more detail in Chapter 5 and Section 6.6. The calibrations that have defined the systematics of the object reconstruction used a fairly inclusive phase space. Yet, it is possible that the boosted regime is not that well captured by it leading to a underestimation of the systematic variations. The uncertainties are subdominant and only a few NPs survive the pruning, which removed the normalisation or shape smaller than 0.5%.

B- and C-hadron decays are not the only source of large s_{d0} and SV mass measurements, which can result from photon conversion, long-lived light particles or hadronic material interactions with the tracker as well. The simulated rate of each of these processes comes with an uncertainty of $\pm\sigma_{rate} = 10\%$, as derived in the

mis-tag rate calibration presented in the Run 1 paper on *Performance of b-Jet Identification in the ATLAS Experiment* [200]. A systematic uncertainty on the template shape is assigned per process, by increasing the weight of the event by $\pm\sigma_{rate}$ if the event has a photon conversion, long-lived light particle, or hadronic interactions on the truth level. After smoothing there is only little shape dependence and the systematic variation is about 2-4%.

Theoretical Uncertainties

The two-point systematic uncertainty on the parton shower and hadronisation model choice is defined as the shape difference between the distributions in the flavour discrimination variable between Herwig and Pythia. The systematic is de-correlated across flavour categories, as both the parton shower and hadronisation have a direct impact on the formation of hadrons and the jet kinematics. It is the largest source of uncertainty in the calibration. Figures 8.14 and 8.15 show an example of the variation that affects only the BB and BL template in the fitted regions. The parton shower and hadronisation uncertainty is pruned in the pass region of the non-B templates, while the uncertainty is significant in the tails of the fail region distributions shown in Figure 8.16. After dropping the normalisation and smoothing the templates the difference between the nominal template and the template generated with Herwig is as large as 20%.

The uncertainty on the renormalisation scale and fragmentation scale (μ) used to evaluate PDFs is evaluated by varying the final-state radiation (FSR) and initial-state radiation (ISR) scales independently. For FSR the variations considered are $\{\mu_R = 0.625, \mu_R = 2.\}$ and for ISR $\{\mu_R = 0.5, \mu_R = 2.\}$, where factor μ_R is a multiplicative correction to μ. For FSR $\mu_R = 0.625$ is picked, as for $\mu_R = 0.5$ the phase space may be ill-defined. The variations can change both the shape and normalisation of an observable and the up (down) systematic is defined by taking

the difference between the nominal and varied flavour discrimination variable distribution in each region fitted. However, only the shape difference is taken into account as differences in the normalisation are captured by the NFs.

The lifetime and branching fractions of a B-hadron are dependent on its species. Therefore, it is not surprising that the SV mass and s_{d0} distribution may vary for each type of B-hadron. The fractions of the B-hadron species in the simulation differ from those derived by the Heavy Flavour Averaging Group from Z decay measurements in e^+e^- collisions [201]. The differences between the two are utilised as the $\pm\sigma$ uncertainties to re-weight the simulation to match the B-hadron fractions with theoretical predictions. The systematic error is expected to be small, with the most significant difference of approximately 5%, as it is only applied to the BB and BL category.

8.2.7 Results

The SV mass distributions in the pass muon-jet, the fail muon-jet, the pass non-muon jet, and the fail non-muon jet regions are fitted to derive preliminary SFs for the 50%, 60%, and 70% WPs in independent bins of large-R jet p_T. However, the fit fails to converge in large-R jet bins with $p_T < 380\,\mathrm{GeV}$, so only the regions where the fit converges are shown. Figure 8.17 shows the SFs for the 50%, 60%, and 70% WP, while the side-to-side comparison of the SFs and NFs between the three WPs in Figure 8.18 illustrates consistent behaviour between the WPs. A similar smooth transition of SFs as a function of the large-R jet p_T is observed for all three considered WPs, which adds confidence in the results.

To understand the behaviour of the flavour correction fit the large-R jet bin with $600 < p_T < 750\,\mathrm{GeV}$ is examined closely. The pre-fit and post-fit distributions in the pass regions are shown in Figure 8.19, while the fail regions are shown in Figure 8.20. Good agreement between data and MC is found in the post-fit SV mass

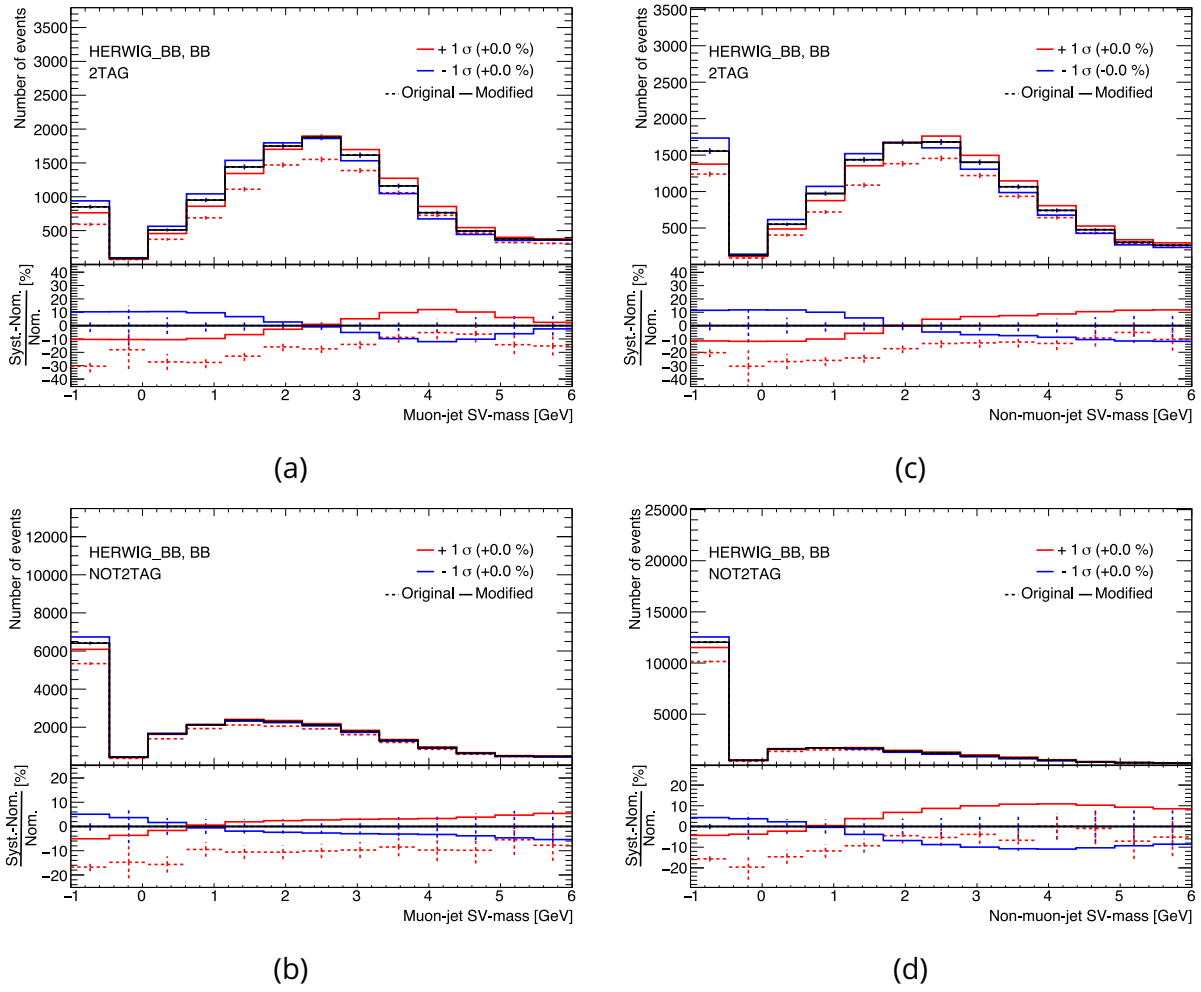


Figure 8.14: The parton shower and hadronisation uncertainty on the BB template before (dotted cross) and after smoothing (solid line) in the fitted SV mass distributions at the 50% WP in the bin with large-R jet p_T: $600 < p_T < 750\,\mathrm{GeV}$. The $+1\sigma$ (-1σ) up (down) variation is shown by the red (blue) solid line, with the normalisation removed.

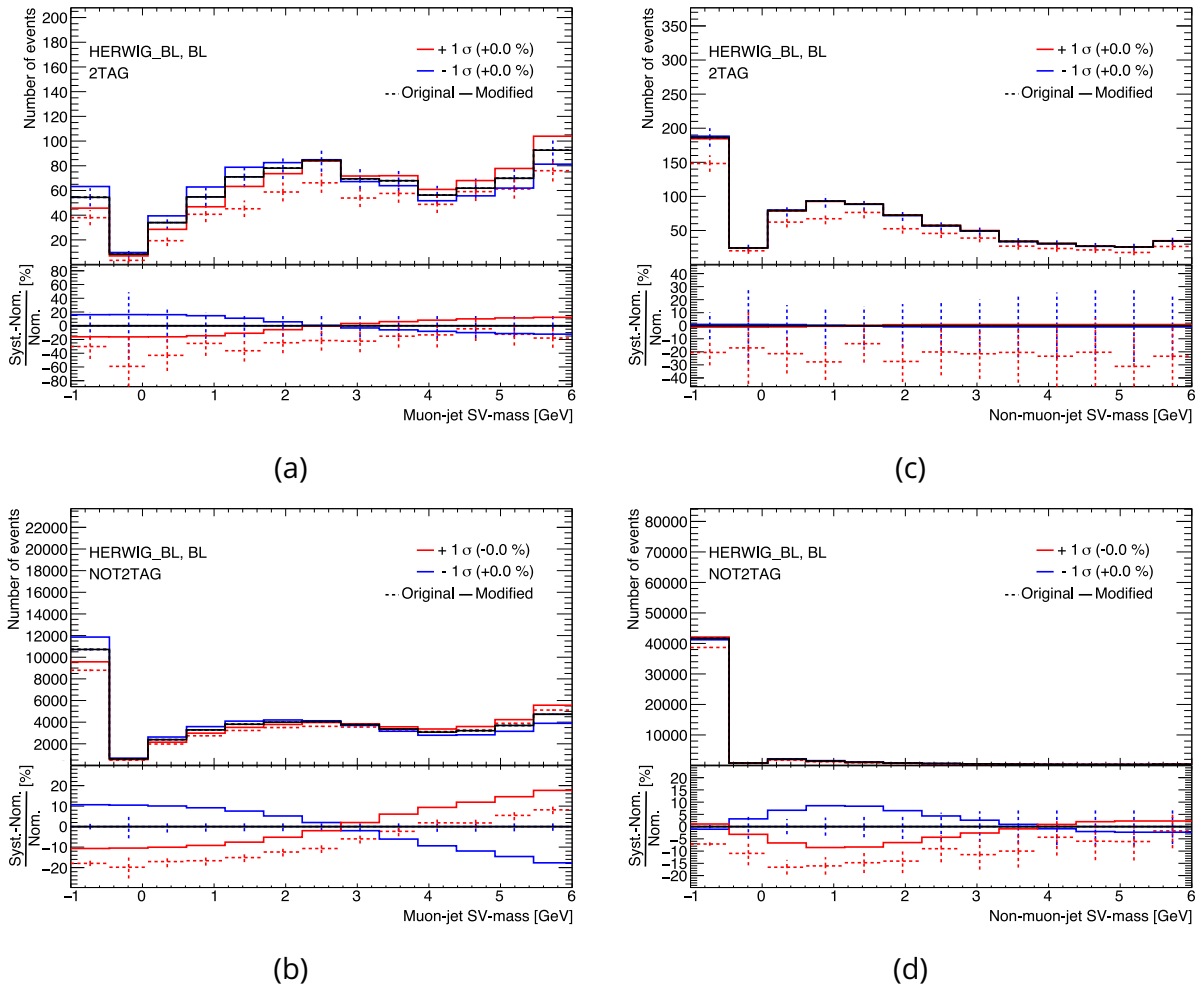


Figure 8.15: The parton shower and hadronisation uncertainty on the BL template before (dotted cross) and after smoothing (solid line) in the fitted SV mass distributions at the 50% WP in the bin with large-R jet p_T: $600 < p_T < 750\,\mathrm{GeV}$. The $+1\sigma$ (-1σ) up (down) variation is shown by the red (blue) solid line, with the normalisation removed.

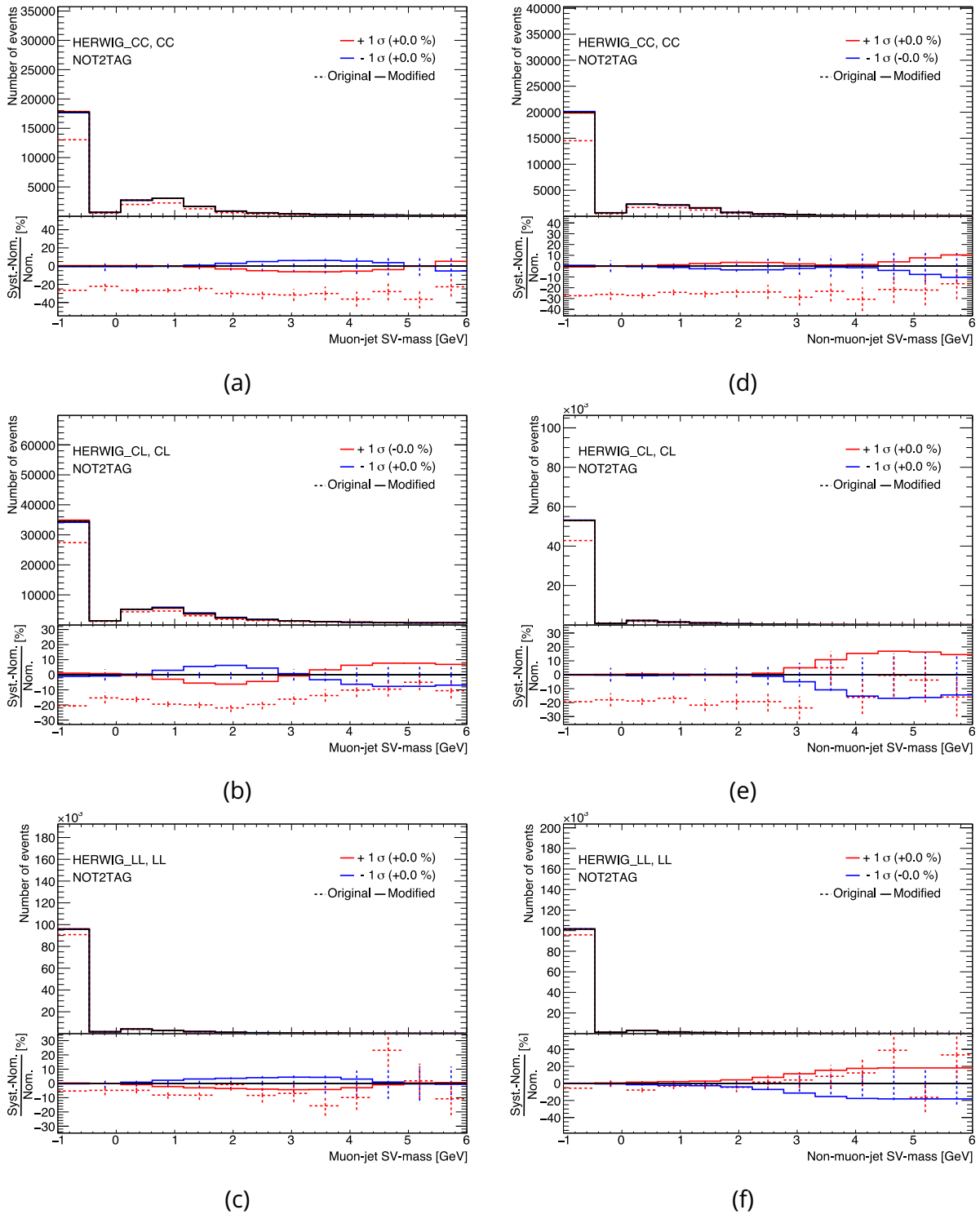
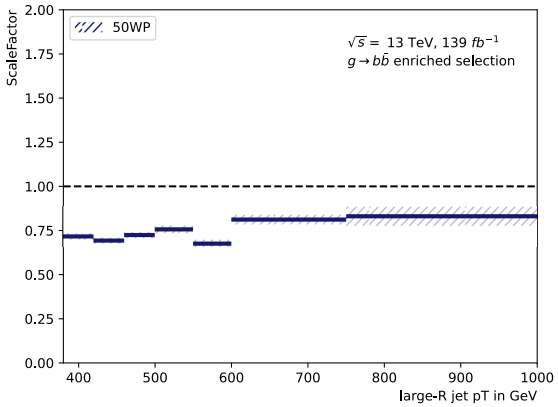
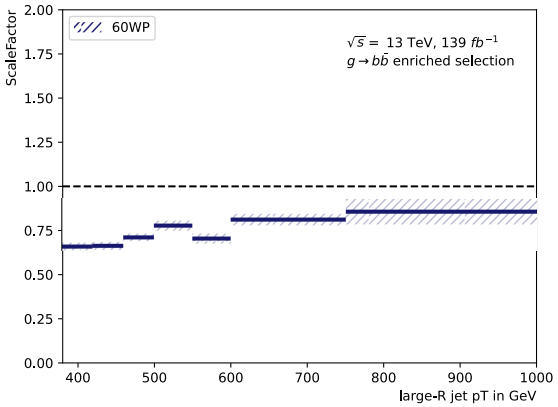


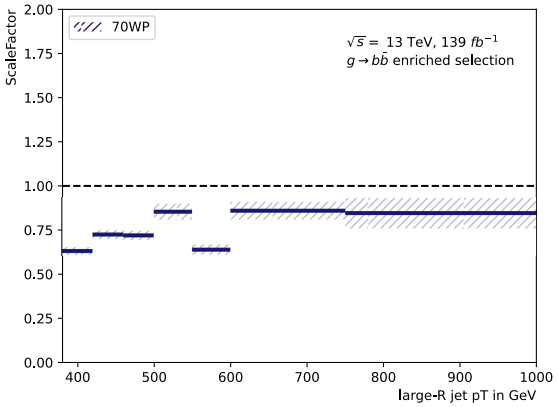
Figure 8.16: The parton shower and hadronisation uncertainty on the CC, CL, and LL templates before (dotted cross) and after smoothing (solid line) in the fitted SV mass distributions at the 50% WP in the bin with large-R jet p_T: $600 < p_T < 750\,\mathrm{GeV}$. The $+1\sigma$ (-1σ) up (down) variation is shown by the red (blue) solid line, with the normalisation removed. The systematics in the pass regions are dropped by pruning.



(a) 50% WP



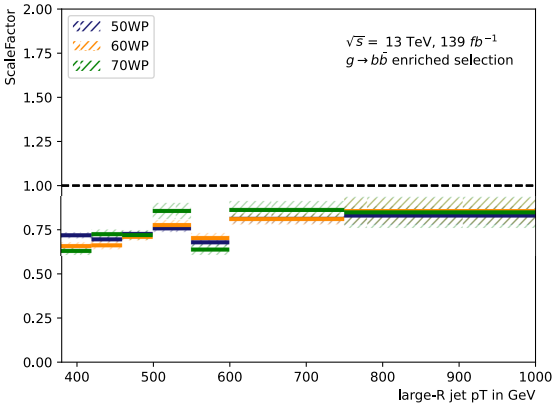
(b) 60% WP



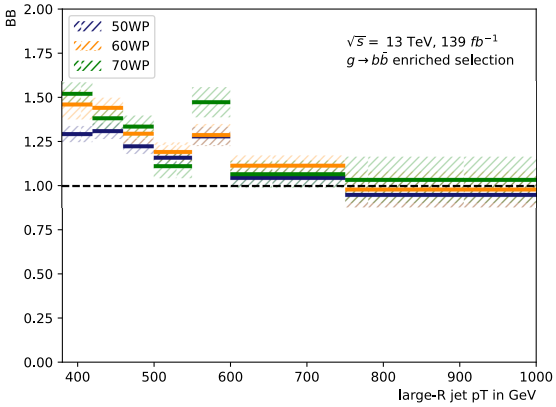
(c) 70% WP

Figure 8.17: The Scale Factors derived in bins of large-R jet p_T for the 50%, 60%, and 70% WPs. The flavour fractions are corrected by fitting SV mass flavour templates in the fails and pass regions of the muon and non-muon jets. The hatched region represents the total uncertainty on the derived SFs.

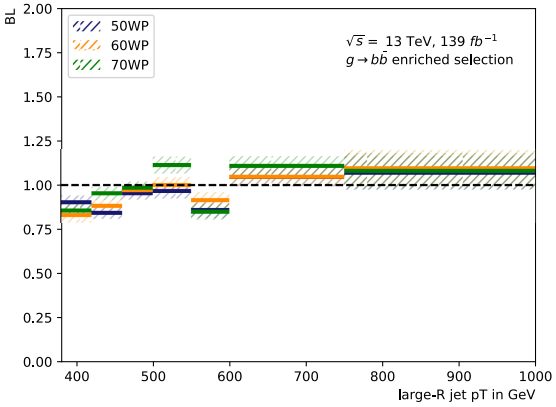
distributions. The fit behaviour is found to be mainly driven by the excess in the tail of the fail $j_{non-\mu}$ distribution, as it is statistically more significant than the mis-modelling in the pass regions. The tail of the fail $j_{non-\mu}$ SV mass distribution consists nearly entirely of BB and BL events, the fraction of each is increased by a factor of 1.05 post-fit in an attempt to improve the modelling. The good agreement at low SV masses already observed in the pre-fit is maintained as while the contribution of the heavy flavour increases the fraction of CL and LL components, which barely contribute to high SV mass, decreases (the normalisation factors are



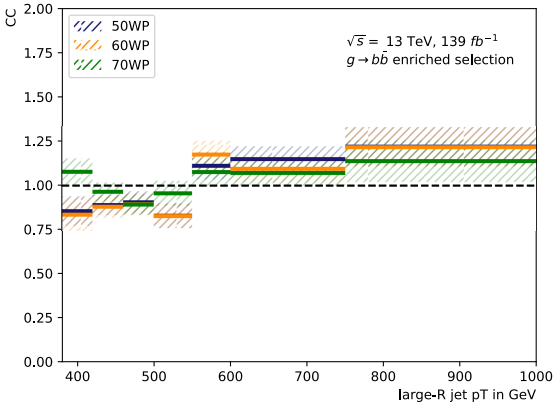
(a) SF



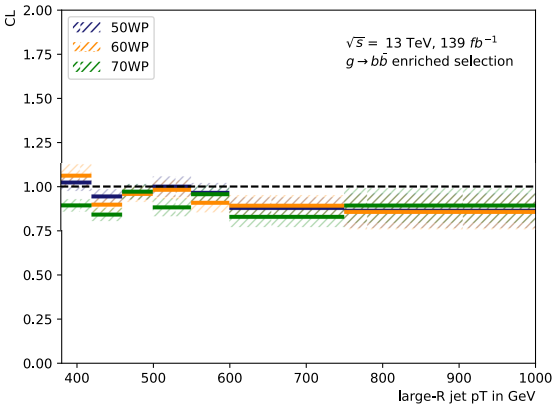
(b) NF_{BB}



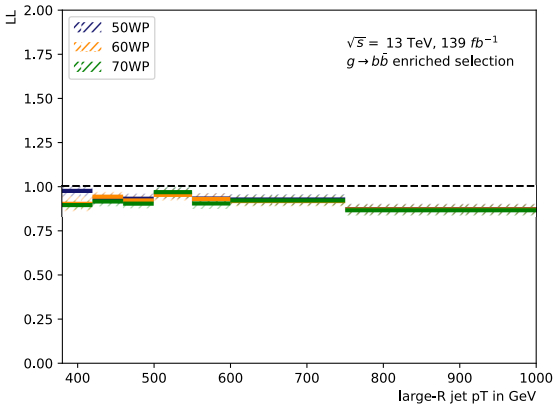
(c) NF_{BL}



(d) NF_{CC}



(e) NF_{CL}



(f) NF_{LL}

Figure 8.18: The comparison of the derived SFs and NF between the 50%, 60%, and 70% WP in bins of large-R jet p_T. The hatched region represents the total uncertainty on the derived SFs and NFs.

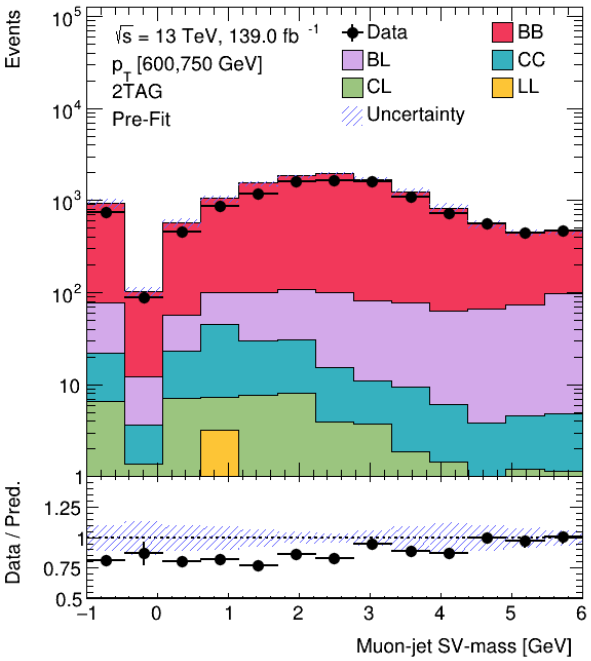
below unity). The NPs associated with the FSR scale and parton shower systematic uncertainties that capture the mis-modelling in the fail $j_{non-\mu}$ distribution are pulled ($\gtrsim 0.5\sigma$) and have a significant impact on the final SF as shown in Figure 8.21. To be exact, the parton shower uncertainties on the CC, CL, and LL temples, shown in 8.16, and the FSR scale uncertainty have a similar effect in the fail regions[4]. The parton shower uncertainties on the CL template and the FSR scale uncertainty have a significant pull, nearly 2σ, as it corrects the discrepancy in the tails of the fail regions. The NPs on the CC and LL templates are pulled less or, in the CC case, not at all as their contribution to the tails is much smaller. The BB and BL NPs representing the parton shower uncertainties cover the discrepancy between the data and MC in the fail regions but to a lesser extent in the pass regions where if pulled they can worsen the mis-modelling, as evident from Figures 8.14 and 8.15. Therefore, there is a compromise on how much the BB and BL NPs representing the parton shower uncertainties are pulled. On the other hand, the large majority of the experimental uncertainties are dropped or have either the normalisation or shape pruned. The remaining difference in normalisation is taken care of by the SF, which is derived to be 0.81 ± 0.03. This is expected as the MC overshoots the data in the pass regions in the pre-fit and the difference increases post-fit as the BB and BL fractions are increased to fix the mis-modelling in the fail $j_{non-\mu}$ SV mass distribution.

After closer inspection, it has been found that the post-fit MC and data discrepancy in fitted regions does not improve or worsen in bins with large-R jet $p_T < 600\,\mathrm{GeV}$. The flavour correction fit is unreliable in the large-R jet $p_T < 600\,\mathrm{GeV}$ and the SFs should be interpreted with caution. The poor post-fit agreement is likely the result of the worse mis-modelling in the fail $j_{non-\mu}$ region, as can be seen in Figure 8.11, which nearly entirely drives the behaviour of the fit. The result is an increase in the NF_{BB}, with decreasing large-R jet p_T, see Figure 8.18b. However, this improves only the MC-data agreement in the fail regions, while the pass regions have the same or

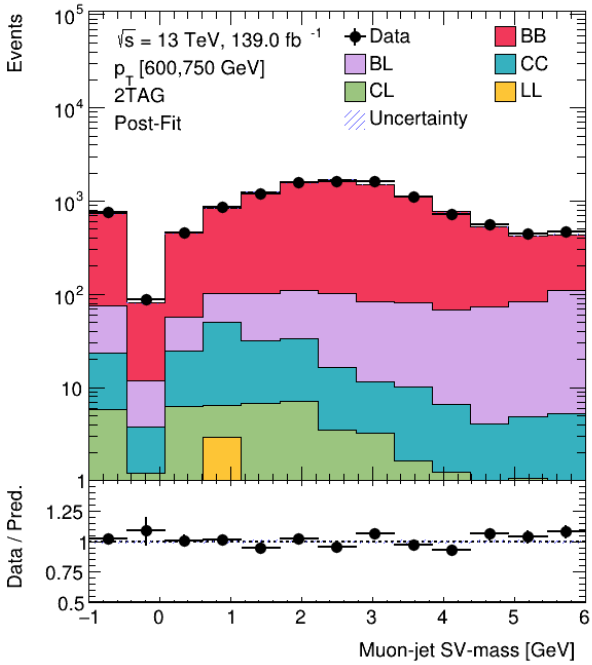
[4]No systematic is applied in the pass regions as the variation is found to be smaller than 0.5% for Herwig_CC, Herwig_CL,Herwig_LL.

worse MC-data agreement compared to the pre-fit distributions. The reason that the mis-modelling in the fail regions determines the behaviour of the fit is the significantly larger constraints from statistics compared to the pass regions. No improvement in the post-fit MC-data agreement in the pass regions and the failure of the fit to converge in some bins indicated that the flavour correction fit might be missing some flexibility.

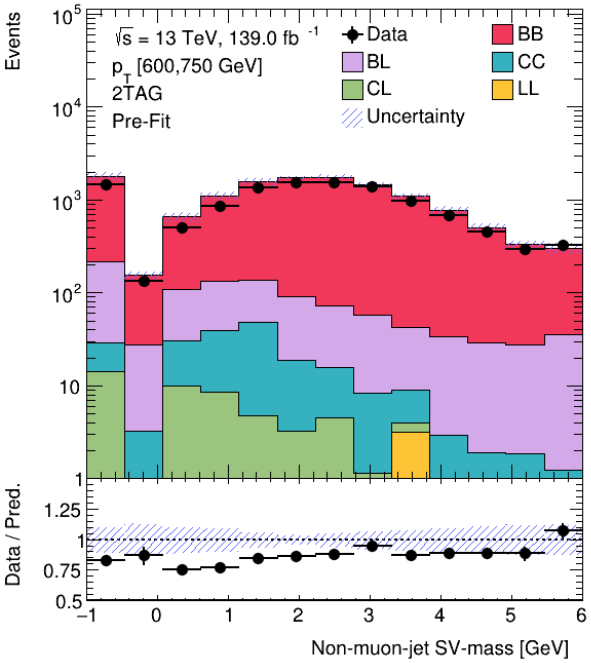
The parton shower and the variation of the FSR scale are the only substantive NPs that can correct the discrepancy between data and MC and have a notable impact on the SF, as deduced from the ranking plot in Figure 8.21. However, these NPs fail to capture the p_T dependent migration of double-B subjets between the j_μ and $j_{non-\mu}$ observed in the BL category but not in data. Thus, the mis-modelling of the double-B subjet migration in the MC cannot be corrected by the fit, as the ratio of the BL sub-components is constant when incorporated into a single template. An NP that changes the fraction of double-B jets in BL templates by varying it within some uncertainty is needed. The NP would introduce a small variation in the tail of the BL templates. The systematic is well justified as the fraction of double-B hadrons reconstructed within a single VR jet is known to come with a non-negligible uncertainty in the MC.



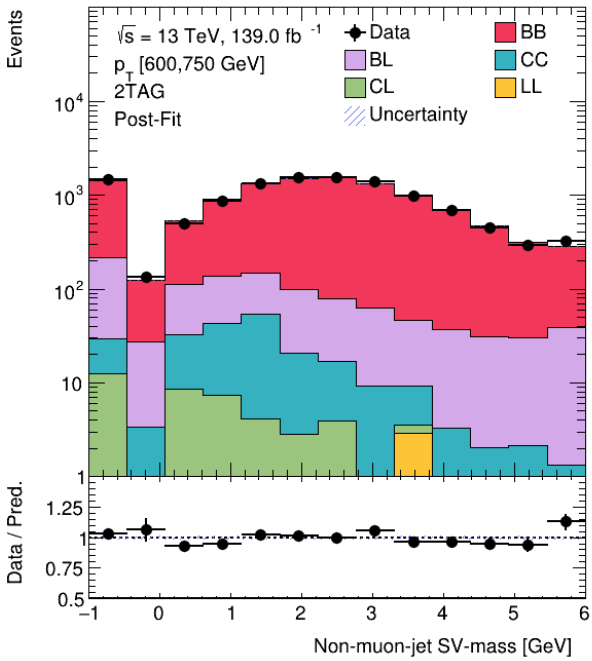
(a) Pre-fit in j_μ pass region



(c) Post-fit in j_μ pass region

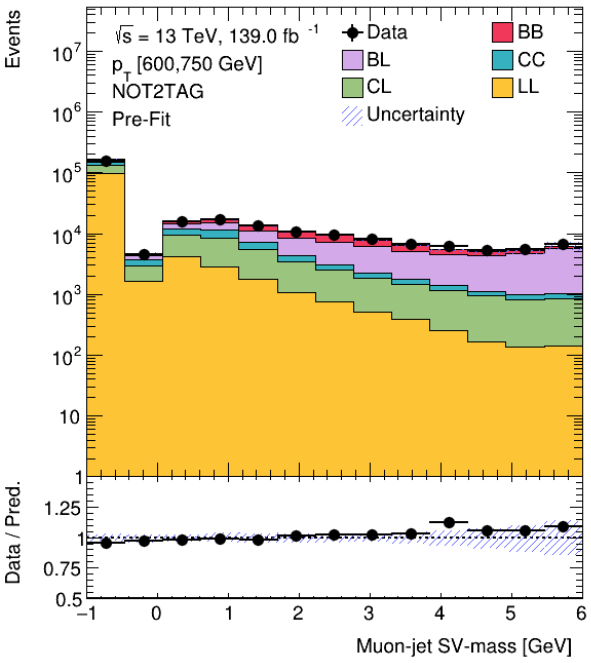


(b) Pre-fit in $j_{non-\mu}$ pass region

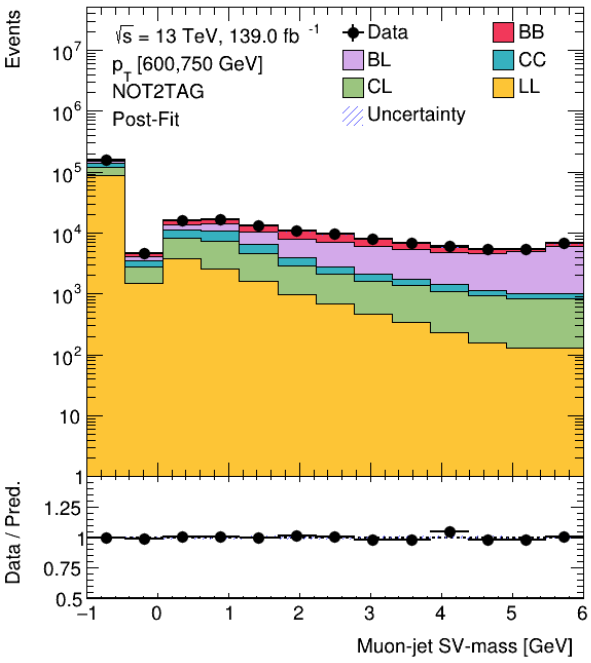


(d) Post-fit in $j_{non-\mu}$ pass region

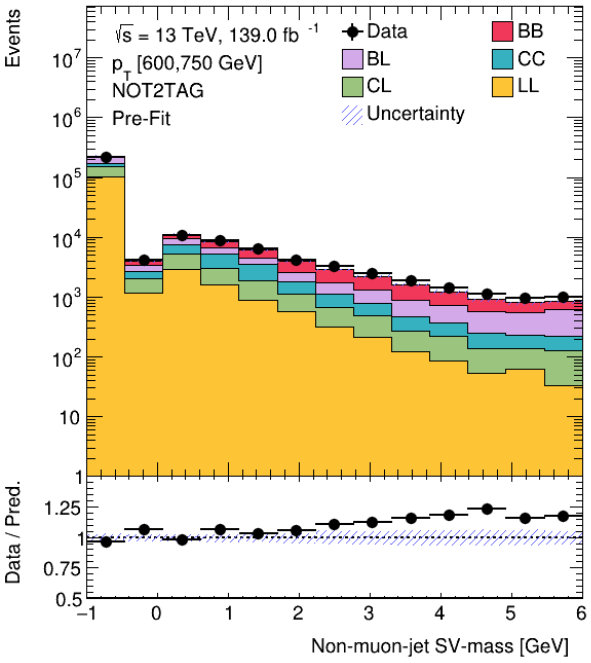
Figure 8.19: Pre- and post-fit SV mass distributions in the pass region at the 50 % WP in the bin with large-R jet p_T: $600 < p_T < 750\,\mathrm{GeV}$. The pre-fit distributions are on the left-hand side, the muon jet in (a) and the non-muon jet in (b). The post-fit distributions are on the right-hand side, the muon jet in (c) and the non-muon jet in (d).



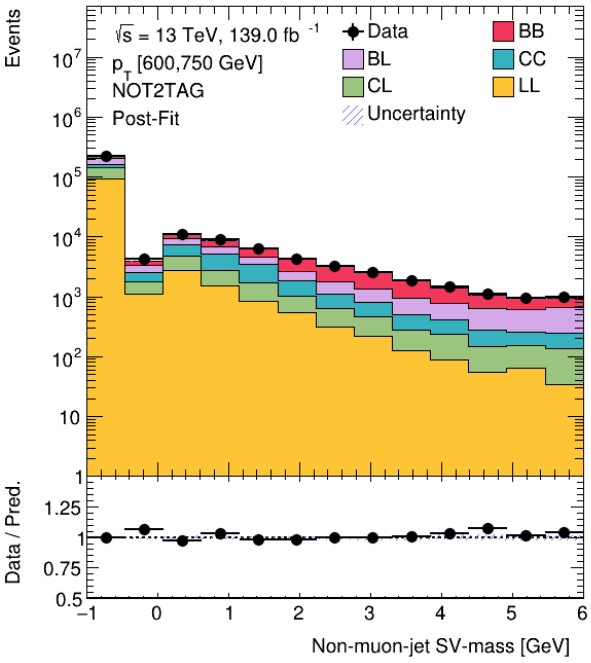
(a) Pre-fit in j_μ fail region



(c) Post-fit in j_μ fail region



(b) Pre-fit in $j_{non-\mu}$ fail region



(d) Post-fit in $j_{non-\mu}$ fail region

Figure 8.20: Pre- and post-fit SV mass distributions in the fail region at the 50% WP in the bin with large-R jet p_T: $600 < p_T < 750\,\mathrm{GeV}$. The pre-fit distributions are on the left-hand side, the muon jet in (a) and the non-muon jet in (b). The post-fit distributions are on the right-hand side, the muon jet in (c) and the non-muon jet in (d).

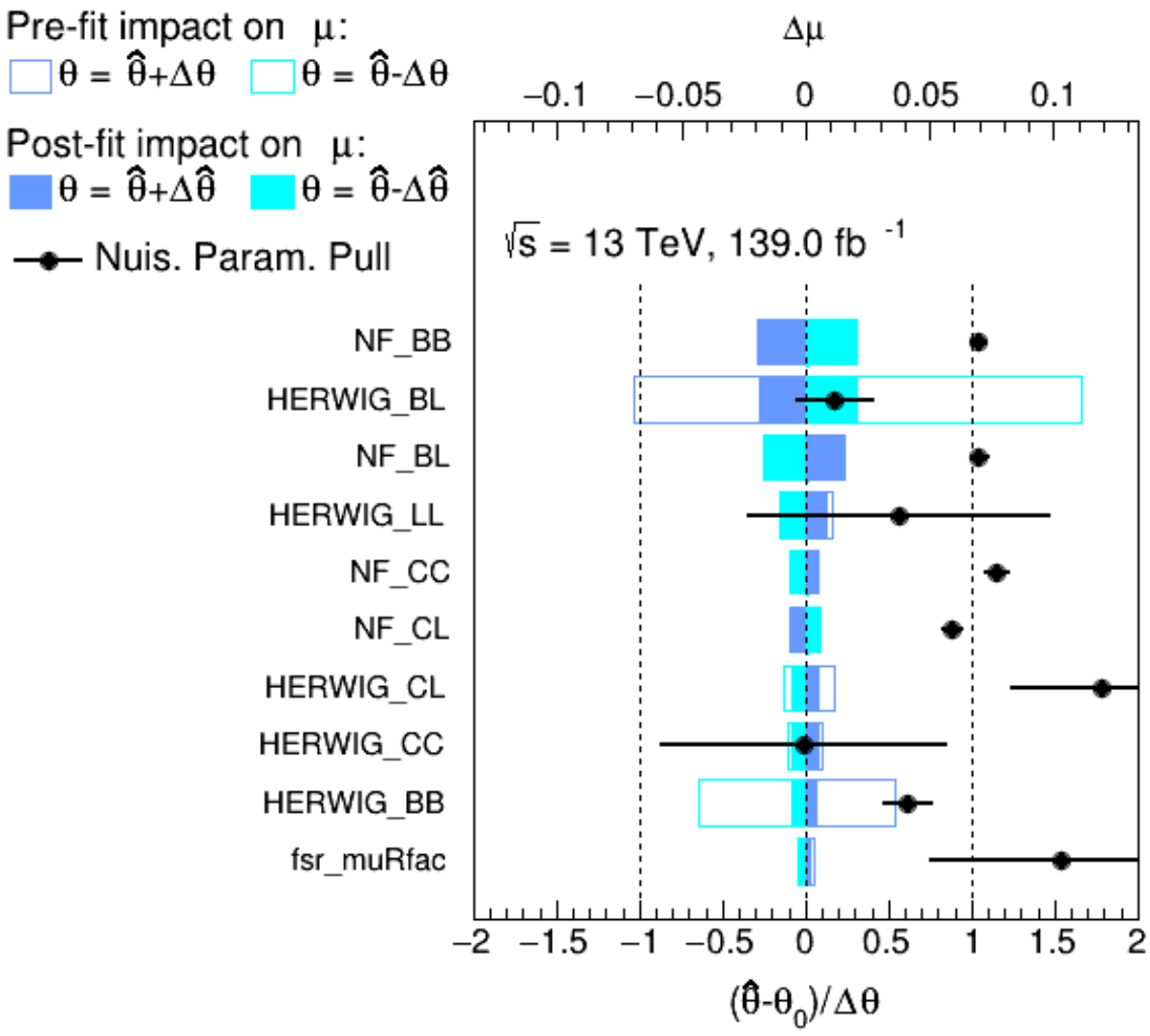


Figure 8.21: The ten highest-ranked nuisance parameters (NPs) found in the large-R jet bin $600 < p_T < 750\,\mathrm{GeV}$ at 50% WP. The pulls (scale on the bottom x-axis) and the impacts on the SF (scale on the top x-axis) are shown. The normalisation factors (NF_{flav}) pre-fit value is one, while the other NPs have a pre-fit value of zero. The constraints observed on the parton shower and hadronisation uncertainty on the BB and BL templates (HERWIG_BB, HERWIG_BL) are expected from the Asimov fit. The source of the constraints is the pass region data, as the pass region consists of BB and to a lesser BL, while the non-B flavours contribute a negligible amount. Therefore, nearly all of the data statistics are used to constrain the BB and BL parton shower systematic uncertainty. As discussed in the text of this thesis some NPs, especially the final state factorisation scale (fsr_mRfac) and parton shower and hadronisation uncertainty on the CL template (HERWIG_CL), are pulled significantly as they correct the mis-modelling in the tails of the fail regions.

8.2.8 Conclusions

In summary, the calibration of the multijet mis-tag efficiency is presented and preliminary scale factors are derived in bins of large-R jet p_T, where $p_T > 380\,\mathrm{GeV}$. The scale factors are extracted from a simultaneous fit in four regions in bins of secondary vertex invariant mass. The use of a flavour-sensitive variable in the fit and the split of the event into regions that fail and pass the double-b tagging enables the fit to correct for the mis-modelling of the flavour composition of the large-R jet in MC and to also compute the scale factor. The flavour correction fit struggles in the large-R jet p_T with $p_T < 600\,\mathrm{GeV}$ and does not converge when $p_T < 380\,\mathrm{GeV}$ as the mis-modelling in the fail non-muon jet region gets worse and there are no appropriate NPs to cover the discrepancy. The source of the mis-modelling in the MC is likely the incorrect fraction of $g \to b\bar{b}$ events reconstructed as a single subject.

The boosted $X \to b\bar{b}$ tagger is currently being retrained to be used as a benchmark to which to compare new taggers developed for Run 3. Although the tagger will be obsolete in ATLAS analysis based on Run 3 data thanks to the constant advances in machine learning techniques, the new state-of-art tagger is likely to be process-based as well. For instance, the current trend in flavour tagging in the LHC community is towards the adoption of graph neural networks. These networks enable the classification of jets by flavour, production process, and vertex origin of the input tracks. An example of such an algorithm used in ATLAS can be found in [202]. Therefore, the need for calibration of the mis-tag rate of the multijet background will remain crucial. The presented *in-situ* calibration of mis-tagged multijet background with dijet Monte Carlo samples enriched in $g \to b\bar{b}$ splittings serves as a proof of concept and with further studies should make an important contribution to any future calibration of the mis-tag rate of the multijet background. Areas for further studies include the addition of extra systematic uncertainties (for instance to cover the modelling of $g \to b\bar{b}$ reconstruction), robustness of the fit model, correlations between entries in the muon and non-muon jet distributions

and the impact of scale-factors for other flavour combinations. Advancements in simulating multijet events and the increase in their rejection by the next-generation taggers are expected to simplify the calibration task at hand.

Chapter 9

Summary

This thesis presents the search for non-resonant diHiggs in the $b\bar{b}b\bar{b}$ final state using $126\,\mathrm{fb}^{-1}$ of the Run 2 dataset collected in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ by the ATLAS detector during Run 2 of the Large Hadron Collider. In addition, the combination of the $HH \to b\bar{b}b\bar{b}$ search with several diHiggs and single Higgs channels that allows for a more general measurement of the Higgs self-coupling and is discussed. Lastly, the thesis concludes with the process based *in-situ* calibration of the mis-tag efficiency of the multijet background of the boosted $X \to b\bar{b}$ tagger.

The first five chapters lay out the theory, experimental apparatus, and reconstruction techniques needed to carry out each of the analyses presented. The Standard Model of Particle Physics (SM) and its shortcomings are presented in Chapter 1, while the Higgs self-coupling and diHiggs production are discussed at length in Chapter 2. Chapter 3 briefly described the Large Hadron Collider and the ATLAS detector. The simulations that tie together the theoretical predictions and experimental observation are outlined in Chapter 4. The reconstruction of physical objects, such as jets, is outlined in Chapter 5.

The search for the non-resonant diHiggs production in the most abundant final state, where each Higgs decays into a pair of $b\bar{b}$ jets, is presented in Chapter 6.

Proton-proton collision data collected at $\sqrt{s} = 13\,\mathrm{TeV}$ by the ATLAS detector during Run 2 of the Large Hadron Collider is used in the analysis. The search is split into two orthogonal channels, each optimised for one of the two leading diHiggs production modes: gluon-gluon fusion (ggF) and vector boson fusion (VBF). The ggF and VBF channels are combined to extract the final results on the SM HH signal strength and constraints on modifications of the trilinear Higgs self-coupling (κ_λ) and HHVV coupling (κ_{2V}).

The CLs method is used to derive the observed (expected) upper limit on the SM HH signal strength, which is found to be 5.4 (8.1) times the SM signal strength, where only the ggF and VBF production modes are considered. The $-2\ln\Delta L$ test statistic is used to measure κ_λ and κ_{2V}. Figure 6.22a (6.22b) shows the observed and expected $-2\ln\Delta L$ scans with κ_λ (κ_{2V}) as the parameter of interest and all other couplings fixed to their expected SM values. The extracted observed (expected) $\pm 2\sigma$ confidence intervals are $\kappa_\lambda \in [-3.5, 11.3]$ ($\kappa_\lambda \in [-5.4, 11.4]$) and $\kappa_{2V} \in [-0.0, 2.1]$ ($\kappa_{2V} \in [-0.1, 2.1]$). Furthermore, the combination of the independently optimised yet orthogonal ggF and VBF production channels provides enough sensitivity to allow both κ_λ and κ_{2V} to vary in the binned likelihood fit performed to extract the $-2\ln\Delta L$ test statistic. The observed (expected) $\pm 1(2)\sigma$ confidence regions in the $\kappa_\lambda - \kappa_{2V}$ plane as derived from $-2\ln\Delta L$ are presented in Figure 6.23. Compared to the first round of Run 2 non-resonant analysis presented in [181] and [182], the legacy Run 2 non-resonant $HH \to b\bar{b}b\bar{b}$ not only improves the upper limits on the diHiggs SM signal strength by more than 50% but also puts more stringent constraints on κ_λ and κ_{2V} measurements. The introduction of the dedicated channels that target the ggF and VBF signals independently unlocks the $\kappa_\lambda - \kappa_{2V}$ phase space. In other words, a more generic κ_λ measurement can be performed where κ_{2V} is constrained by the observed data.

The Higgs Effective Field Theory (HEFT) and Standard Model Effective Field Theory (SMEFT) are used to reinterpret the ggF only channel. The upper cross-section limits

on various HEFT and SMEFT signals considered are derived with the CLs method. Figure 6.24 shows the upper limits on the cross-section derived at 95% confidence level of seven different HEFT benchmarks (BM) [41], which capture a range of possible m_{HH} distribution shapes. It is found that the BM3, BM5, and BM7 are excluded at more than 95% confidence level. Additionally, upper limits on the cross-sections are extracted for the $c_{t\bar{t}HH}$ Wilson coefficient, shown in Figure 6.25a, and the c_{ggHH} Wilson coefficient, shown in Figure 6.25b, in the range [-1.2,1.2]. Scanning several $c_{t\bar{t}HH}$ (c_{ggHH}) allows extraction of the 95% confidence intervals on the Wilson coefficient by comparing the upper limits to the theoretically calculated cross-section. Values of $c_{t\bar{t}HH}$ outside the interval $[-0.36, 0.78]$ are excluded at $\geq 95\%$ confidence level, while the corresponding interval for c_{ggHH} is $[-0.55, 0.51]$.

Similarly, bounds are set on the SMEFT Wilson coefficients considered, c_H, c_{HG}, $c_{H\Box}$, c_{tH}, and c_{tG}, independently. The observed and expected upper limits on the cross-section of the scanned coefficients are shown in Figure 6.25. The observed 95% confidence intervals are, $c_H \in [-22, 11.]$, $c_{HG} \in [-0.067, 0.060]$, $c_{H\Box} \in [-8.9, 14.5]$, $c_{HG} \in [-8.9, 14.5]$, $c_{tH} \in [-10.7, 6.2]$ and $c_{tG} \in [-1.12, 1.15]$ extracted per Wilson coefficients with all other coefficients fixed to their SM expectations (zero). A more general setup is considered as well, where c_H is varied in addition to one other coefficient to extract the limits in the two-dimensional parameter space. Figure 6.27 shows the extracted confidence regions at 95% confidence level in the two-dimensional parameter spaces. However, constraints on the Wilson coefficients of the size $\sim 4\pi$ should be interpreted with caution as the validity of the SMEFT framework is not guaranteed. It can be concluded that ggF $HH \to b\bar{b}b\bar{b}$ search has very limited sensitivity to Wilson coefficients in the SMEFT framework. The interpretation is valuable as it sets up the analysis to be included in the global EFT fit where several of the Wilson coefficients can be constrained by precision single Higgs measurements.

The $b\bar{b}b\bar{b}$ final state is not the only diHiggs decay channel that the Run 2 dataset

collected by the ATLAS Experiment is sensitive to. It is natural to combine all the decay channels considered to extract the most stringent upper limits on the SM HH production signal strength. To pin down the Higgs self-coupling a step further can be taken by accounting for single Higgs electroweak precision measurements, that are sensitive to anomalies in Higgs self-coupling through electroweak loop corrections. The latest diHiggs, as well as single Higgs and diHiggs (H+HH) combinations published by the ATLAS Collaboration, are described in Chapter 7. The chapter emphasises the steps taken to harmonise the $HH \to b\bar{b}b\bar{b}$ channel with the other inputs to the combination. The intermediate step of the H+HH combination is the combination of only the diHiggs channels. The $HH \to b\bar{b}b\bar{b}$, $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau^{-}\tau^{+}$ decays channels are combined, resulting in an observed (expected) upper limit on the HH SM signal strength to be 2.4 (2.9) times the SM prediction. The upper limits on the SM HH signal strength extracted from the individual as well as the combined diHiggs searches considered are shown in Figure 7.3. The significance of adding the single Higgs channels to the combination is that, while their sensitivity to the anomalous Higgs self-coupling is limited, they can place stringent constraints on several other Higgs couplings. Thus, the $-2ln\Delta L$ test statistic used to measure κ_λ treats the Higgs couplings to the top, bottom, tau and vector bosons as free parameters left to be constrained by the observed dataset. Figure 7.4 shows that indeed in the H+HH combination only little sensitivity to κ_λ is lost when loosening the assumptions on the other Higgs couplings besides κ_λ. To be precise, when no assumptions are placed on the Higgs couplings to the top, bottom, tau and vector bosons the observed (expected) 95% confidence levels are $\kappa_\lambda \in [-1.3, 6.1]$ $(\kappa_\lambda \in [-2.1, 7.6])$ in the H+HH combination. On the other hand, only varying κ_λ in $-2ln\Delta L$ leads to the observed (expected) 95% confidance interval $\kappa_\lambda \in [-0.4, 6.3]$ $(\kappa_\lambda \in [-1.9, 7.5])$.

All three diHiggs channels that enter the H+HH combination have at least one of the Higgs bosons decaying to a $b\bar{b}$ pair. The sensitivity of these searches would benefit from improvements in b-tagging efficiency at high transverse momentum, as

anomalous HHVV couplings enhance the number of produced Higgs bosons with a large boost. Therefore, a specialized tagger, the boosted $X \to b\bar{b}$ tagger, has been developed to target the Higgs decays to a pair of b quarks at high transverse momentum. Chapter 8 described the boosted $X \to b\bar{b}$ tagger and the processed-based *in-situ* calibration of the mis-tag of the multijet background. Figure 8.17 shows the preliminary scale factors to be applied to the simulation and derived for the 50%, 60% and 70% working points. The correction of the flavour components in the dijet simulation as well as the scale factors on the mis-tag rate are derived on the fly from a binned likelihood fit in four regions of secondary vertex invariant mass.

The measured constraints on the Higgs self-coupling, while consistent with the SM scenario, still allow for a BSM Higgs sector. For instance, several values of κ_λ within the measured 95% confidence interval allow the Higgs potential to undergo the first-order phase transition needed for the electroweak baryogenesis to occur in the early Universe [47]. Therefore, the investigation of the Higgs self-coupling that will ultimately determine the shape of the Higgs potential must continue, as knowing it will not only help to understand the origin of electroweak symmetry breaking but alose the evolution of the Universe.

9.1 Outlook

To measure the Higgs self-coupling remains one of the main goals of current and future particle colliders. The Run 2 searches for the diHiggs VBF production are statistically limited, while the ggF production mode, especially the $b\bar{b}b\bar{b}$ final state, suffered from a low trigger acceptance rate of the $HH \to b\bar{b}b\bar{b}$ signal and a large multijet irreducible background that is challenging to model.

In the already ongoing Run 3 of the LHC, the search for the diHiggs VBF production

mode is expected to benefit from the increase in luminosity, which is likely to enable searchers for the $t\bar{t}H$ production mode as well. The low acceptance of the $HH \to b\bar{b}b\bar{b}$ signal has already been tackled by the Run 3 trigger, as outlined in [203]. The use of triggers that identify multi-b-hadron signatures increases the acceptance of the $HH \to b\bar{b}b\bar{b}$ SM signal by more than 50%, especially at low m_{HH}. Higher efficiency in b-tagging in hand with a higher background rejection is also expected to come from the switch to graph neural network based b-tagging algorithms [202]. Furthermore, advancements in process based tagging algorithms are likely to reduce the now irreducible $g \to b\bar{b}$ background.

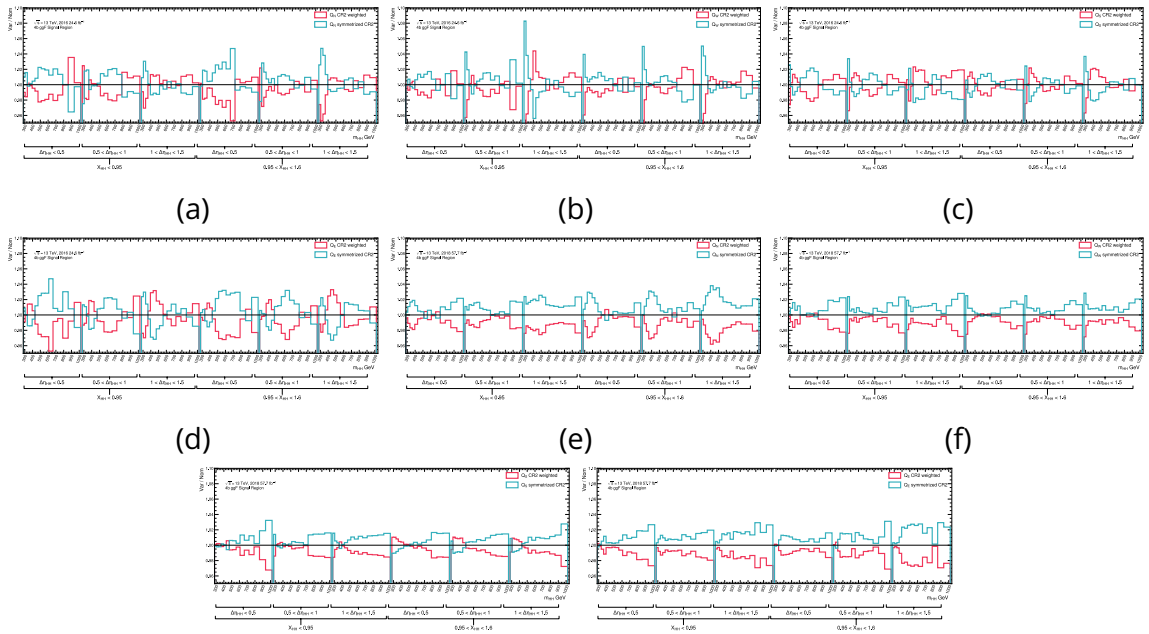
The next step will be the high luminosity phase of the LHC (HL-LHC) [204], which is expected to start in 2029 and in its ten years of running deliver a total of $3000\,\mathrm{fb}^{-1}$ integrated luminosity at $\sqrt{s} = 14\,\mathrm{TeV}$. The latest HL-LHC projections, as published in [205], predict the discovery significance of the SM diHiggs production of 3.4σ (4.9σ) with (without) systematic uncertainties. Given a BSM Higgs-self coupling of $\kappa_\lambda < 1.7$ or $\kappa_\lambda > 7.8$ the discovery significance for the diHiggs signal is greater than 3σ at HL-LHC. The projections are obtained by extrapolating and combining the three Run 2 diHiggs channels leading in sensitivity, $HH \to b\bar{b}b\bar{b}$, $HH \to b\bar{b}b\gamma\gamma$ and $HH \to b\bar{b}\tau^-\tau^+$ searches. The projections are rather pessimistic and the obtainable discovery significance at HL-LHC is likely much higher. The reason is that while the project scales down the statistical uncertainties and the majority of the experimental and theoretical uncertainties the data-driven uncertainties on the background estimation techniques or spurious signal are fixed to their Run 2 values. For instance, the uncertainty on the spurious signal in $b\bar{b}b\gamma\gamma$ decay channel and the uncertainty on the background estimation in $b\bar{b}b\bar{b}$ decay channel are assumed not to or only partially decrease. However, this is highly unlikely as advancements in analysis strategies are inevitable to come from lessons from the past or to be diHiggs searches and the systematical uncertainty on the background estimation is expected to shrink with more data collected. Especially in the $b\bar{b}b\bar{b}$ final state, it is crucial to push for a better understanding of the irreducible multijet background

and develop novel ways not only to reduce it but also model it with a high degree of certainty. The Higgs self-coupling can be precisely measured at the percent level only by combining the observations of the (possible) next-generation particle colliders, the FCC-ee and its successor FCC-hh [206, 207, 208, 209]. The FCC-ee alone is not enough to discover the SM diHiggs production, as at the planned centre-of-mass energy the Higgs self-coupling is accessible only though NLO electroweak corrections in the single Higgs production [208].

Appendix A

CR12 Non-Closure Uncertainty

The appendix shows the CR12 non-closure uncertainty, which definition is discussed in Section 6.6.1, in the 2016 and 2018 ggF categories.



(a) (b) (c)

(d) (e) (f)

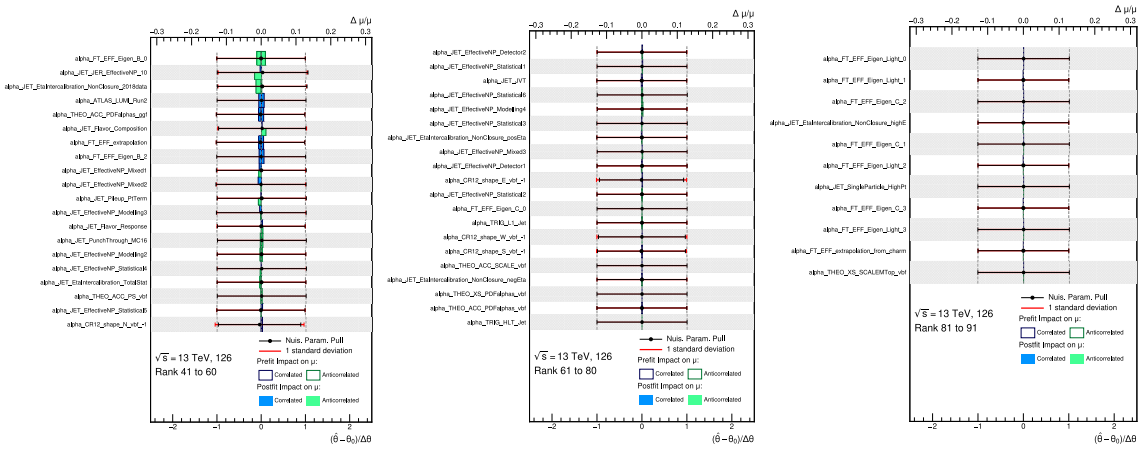
(g) (h)

Figure A.1: Non-closure systematics derived from the alternative background estimation in SR quadrants in the ggF channel in the data taking the year 2016 and 2018.

Appendix B

Ranking Plots

The pulls on the nuisance parameters (NPs) and impacts as extracted from the fit with $\mu_{ggF+VBF}$, as the POI. The top forty pulls ranked by impact are shown in Figure 6.21 in Section 6.8.



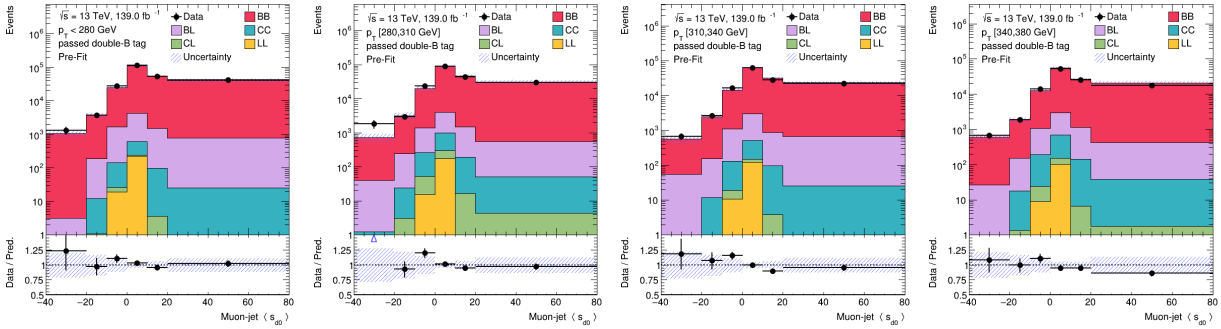
(a) (b) (c)

Figure B.1: The pulls on the nuisance parameters and impacts on the fit to extract the upper limit on $\mu_{ggF+VBF}$. The nuisance parameters ranked by impact from the 41st to the 91st place are shown.

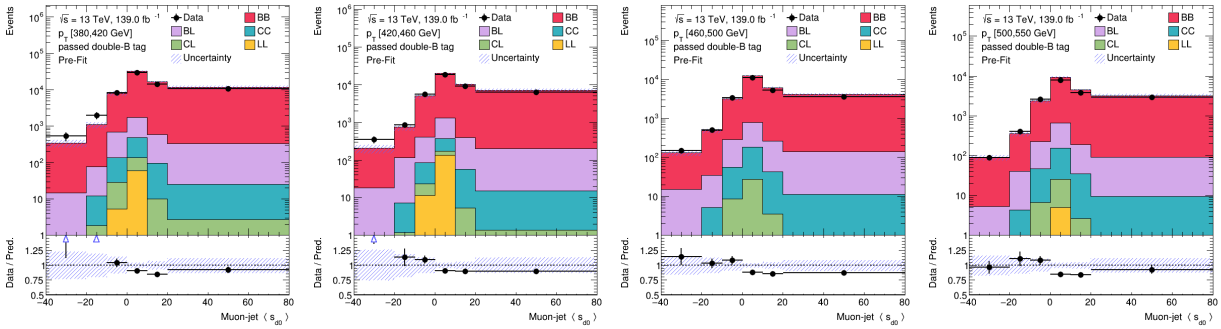
Appendix C

$< s_{d0} >$ Distribution

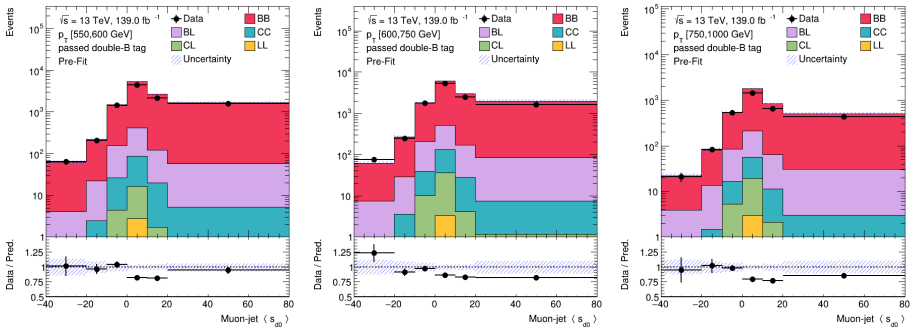
The appendix shows the pre-fit $< s_{d0} >$ distributions in the pass muon jet (Figure C.1), fail muon jet (Figure C.2) and fail non-muon jet (Figure C.3) regions in bins of large-R jet p_T.



(a) $280 < p_T$ GeV (b) $280 < p_T < 310$ GeV (c) $310 < p_T < 340$ GeV (d) $340 < p_T < 380$ GeV

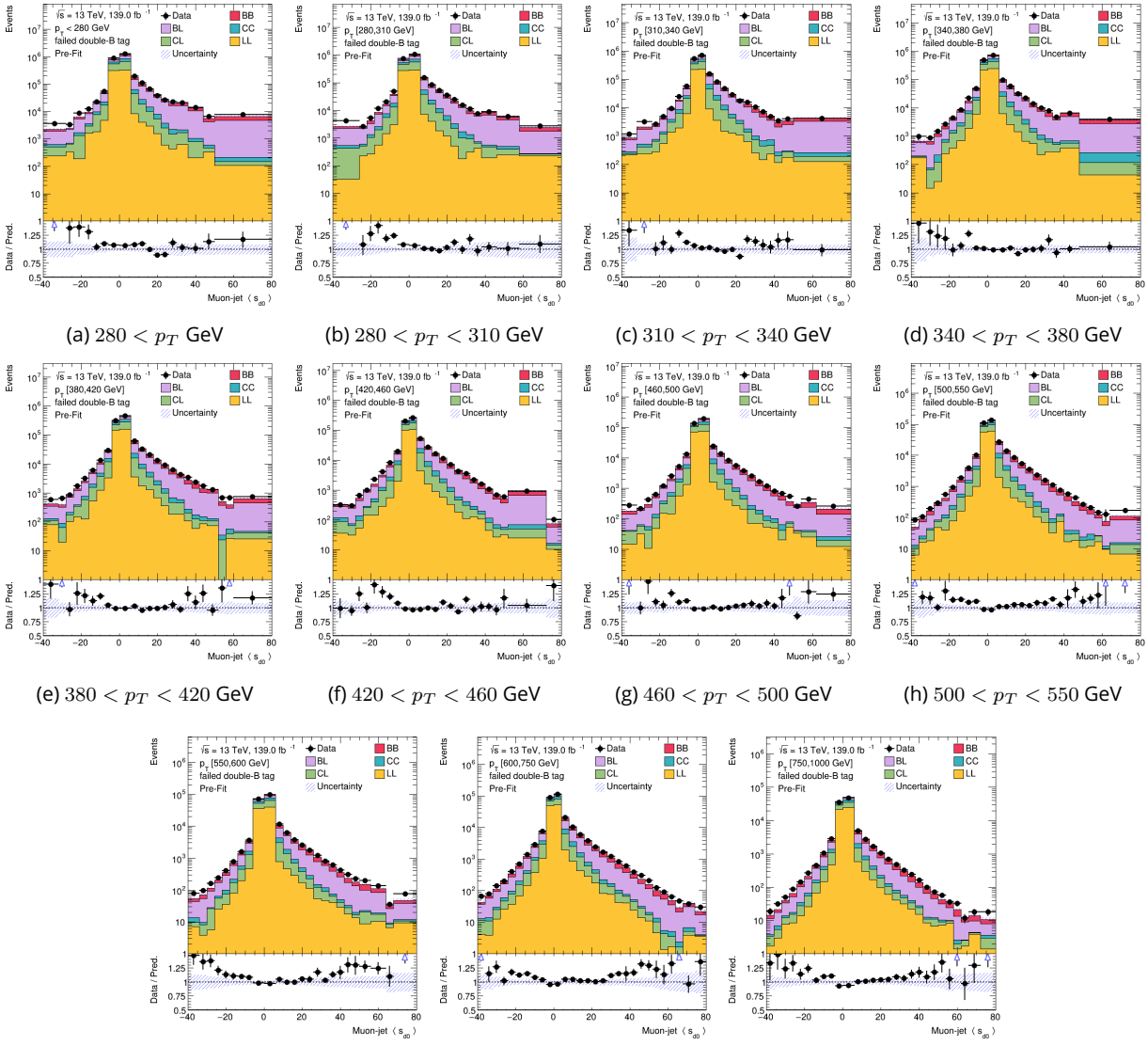


(e) $380 < p_T < 420$ GeV (f) $420 < p_T < 460$ GeV (g) $460 < p_T < 500$ GeV (h) $500 < p_T < 550$ GeV



(i) $550 < p_T < 600$ GeV (j) $600 < p_T < 750$ GeV (k) $750 < p_T < 1000$ GeV

Figure C.1: The $< s_{d0} >$ distribution in the pass region of the muon jet in in different large-R jet p_T bins for the 50% WP.

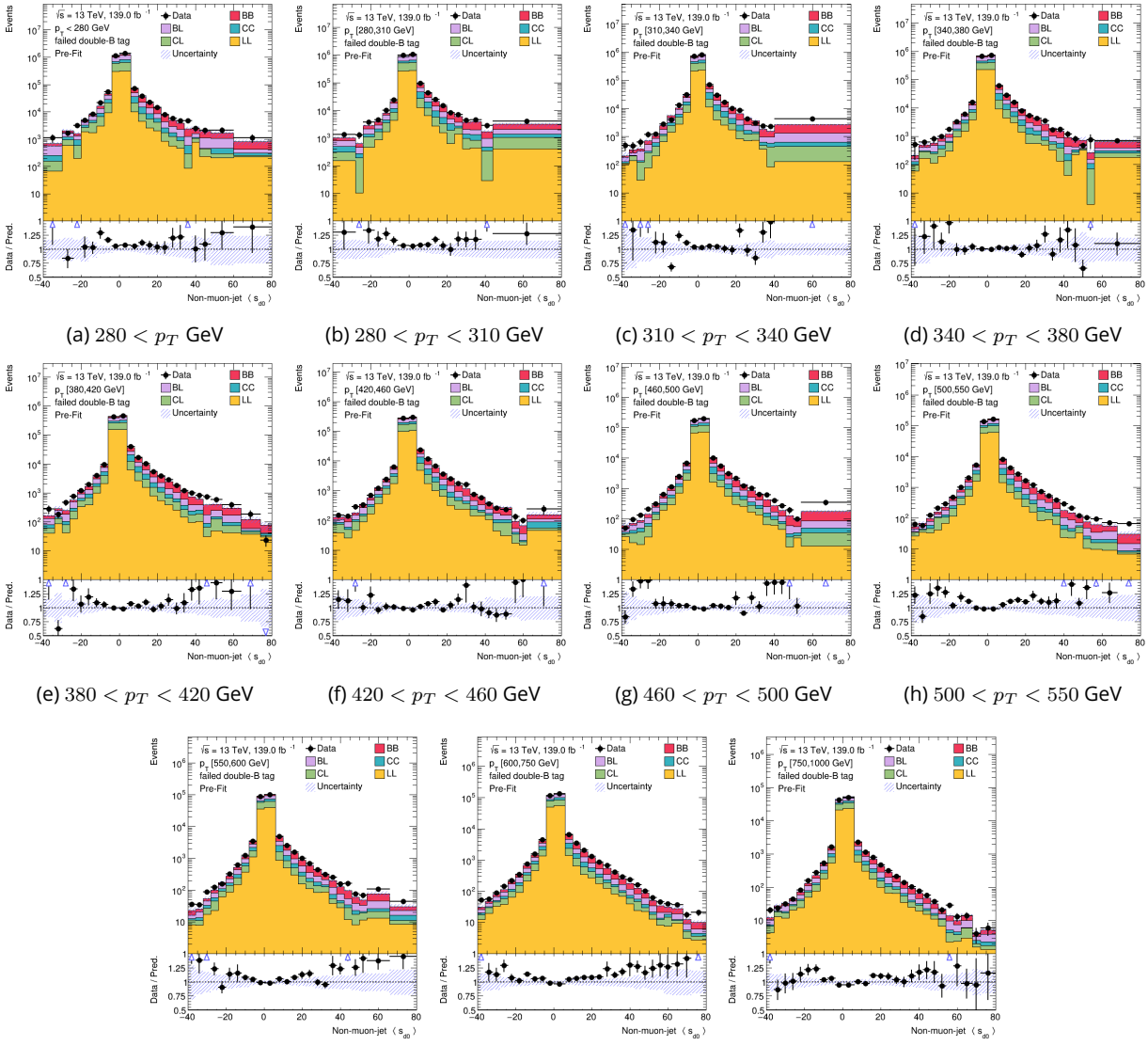


(a) $280 < p_T$ GeV (b) $280 < p_T < 310$ GeV (c) $310 < p_T < 340$ GeV (d) $340 < p_T < 380$ GeV

(e) $380 < p_T < 420$ GeV (f) $420 < p_T < 460$ GeV (g) $460 < p_T < 500$ GeV (h) $500 < p_T < 550$ GeV

(i) $550 < p_T < 600$ GeV (j) $600 < p_T < 750$ GeV (k) $750 < p_T < 1000$ GeV

Figure C.2: The $< s_{d0} >$ distribution in the fail region of the muon jet in in different large-R jet p_T bins for the 50% WP.



(a) $280 < p_T$ GeV (b) $280 < p_T < 310$ GeV (c) $310 < p_T < 340$ GeV (d) $340 < p_T < 380$ GeV

(e) $380 < p_T < 420$ GeV (f) $420 < p_T < 460$ GeV (g) $460 < p_T < 500$ GeV (h) $500 < p_T < 550$ GeV

(i) $550 < p_T < 600$ GeV (j) $600 < p_T < 750$ GeV (k) $750 < p_T < 1000$ GeV

Figure C.3: The $< s_{d0} >$ distribution in the fail region of the non-muoj jet in different large-R jet p_T bins for the 50% WP.

Bibliography

1. CERN. *The CERN Experimental Programme - ATLAS* https://greybook.cern.ch/experiment/detail?id=ATLAS (2023).
2. ATLAS Collaboration. *Search for nonresonant 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*. 2023. arXiv: 2301.03212 [hep-ex].
3. 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*. 2022. arXiv: 2211.01216 [hep-ex].
4. ATLAS Collaboration. *Efficiency corrections for a tagger for boosted $H \to b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. (CERN, Geneva, 2021). http://cds.cern.ch/record/2777811.
5. Robinson, M. B. *Symmetry and the standard model : mathematics and particle physics [electronic resource]* eng. ISBN: 9781441982674 (electronic bk.) (Springer, New York, 2011).
6. Rajantie, A. *Unification, Autumn Term 2018*. Oct. 2018.
7. Dirac, P. A. M. *The quantum theory of the electron*. *Proc. Roy. Soc. Lond. A* **117,** 610–624 (1928).
8. Yang, C. N. & Mills, R. L. *Conservation of Isotopic Spin and Isotopic Gauge Invariance*. *Phys. Rev.* **96,** 191–195. https://link.aps.org/doi/10.1103/PhysRev.96.191 (1 Oct. 1954).

9. Burde, D. *Lie algebras and representation theory*. 2021. https://homepage.univie.ac.at/dietrich.burde/papers/burde_70_lie_algebra_course.pdf.

10. Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. *Experimental Test of Parity Conservation in Beta Decay*. *Phys. Rev.* **105,** 1413–1415. https://link.aps.org/doi/10.1103/PhysRev.105.1413 (4 Feb. 1957).

11. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. *Evidence for the 2π Decay of the K_2^0 Meson*. *Phys. Rev. Lett.* **13,** 138–140. https://link.aps.org/doi/10.1103/PhysRevLett.13.138 (4 July 1964).

12. Salam, A. *Weak and Electromagnetic Interactions*. *Conf. Proc. C* **680519,** 367–377 (1968).

13. Weinberg, S. *A Model of Leptons*. *Phys. Rev. Lett.* **19,** 1264–1266. https://link.aps.org/doi/10.1103/PhysRevLett.19.1264 (21 Nov. 1967).

14. Glashow, S. L. *Partial-symmetries of weak interactions*. *Nuclear Physics* **22,** 579–588. ISSN: 0029-5582. https://www.sciencedirect.com/science/article/pii/0029558261904692 (1961).

15. Higgs, P. W. *Broken Symmetries and the Masses of Gauge Bosons*. *Phys. Rev. Lett.* **13,** 508–509. https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (16 Oct. 1964).

16. Englert, F. & Brout, R. *Broken Symmetry and the Mass of Gauge Vector Mesons*. *Phys. Rev. Lett.* **13,** 321–323. https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (9 Aug. 1964).

17. Thomson, M. *Modern particle physics* eng. ISBN: 9781107034266 (Cambridge, 2013).

18. Branco, G., Ferreira, P., Lavoura, L., Rebelo, M., Sher, M. & Silva, J. P. *Theory and phenomenology of two-Higgs-doublet models*. *Physics Reports* **516,** 1–102. https://doi.org/10.1016%2Fj.physrep.2012.02.002 (July 2012).

19. 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,** 1–29. https://doi.org/10.1016%2Fj.physletb.2012.08.020 (Sept. 2012).

20. CMS Collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. *Physics Letters B* **716,** 30–61. ISSN: 0370-2693. https://www.sciencedirect.com/science/article/pii/S0370269312008581 (2012).

21. ATLAS Collaboration. *Standard Model Summary Plots February 2022*. Tech. rep. (CERN, Geneva, 2022). https://cds.cern.ch/record/2804061.

22. Ahmad, Q. R. *et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*. *Physical Review Letters* **89.** https://doi.org/10.1103%2Fphysrevlett.89.011301 (June 2002).

23. Fukuda, Y. *et al. Evidence for Oscillation of Atmospheric Neutrinos*. *Physical Review Letters* **81,** 1562–1567. https://doi.org/10.1103%2Fphysrevlett.81.1562 (Aug. 1998).

24. Workman, R. L. *et al. Review of Particle Physics*. *PTEP* **2022,** 083C01 (2022).

25. Bilenky, S. in *Introduction to the Physics of Massive and Mixed Neutrinos* 79–88 (Springer International Publishing, Cham, 2018). ISBN: 978-3-319-74802-3. https://doi.org/10.1007/978-3-319-74802-3_5.

26. Sakharov, A. D. *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*. *Pisma Zh. Eksp. Teor. Fiz.* **5,** 32–35 (1967).

27. Ashtekar, A. & Bianchi, E. *A short review of loop quantum gravity*. *Reports on Progress in Physics* **84,** 042001. https://doi.org/10.1088%2F1361-6633%2Fabed91 (Mar. 2021).

28. Martin, S. P. in *Perspectives on Supersymmetry* 1–98 (WORLD SCIENTIFIC, July 1998). https://doi.org/10.1142%2F9789812839657_0001.

29. Babcock, H. W. *The rotation of the Andromeda Nebula*. *Lick Observatory Bulletin* **498,** 41–51. https://ui.adsabs.harvard.edu/abs/1939LicOB..19...41B (Jan. 1939).

30. Oort, J. H. *Some Problems Concerning the Structure and Dynamics of the Galactic System and the Elliptical Nebulae NGC 3115 and 4494. Astrophysical Journal* **91,** 273. https://ui.adsabs.harvard.edu/abs/1940ApJ....91..273O (Apr. 1940).

31. Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C. & Zaritsky, D. *A Direct Empirical Proof of the Existence of Dark Matter*. The Astrophysical Journal* **648,** L109. https://dx.doi.org/10.1086/508162 (Aug. 2006).

32. Grazzini, M., Heinrich, G., Jones, S., Kallweit, S., Kerner, M., Lindert, J. M. & Mazzitelli, J. *Higgs boson pair production at NNLO with top quark mass effects*. *Journal of High Energy Physics* **2018.** https://doi.org/10.1007%2Fjhep05%282018%29059 (May 2018).

33. Di Micco, B. *et al. Higgs boson pair production at colliders: status and perspectives. Higgs boson potential at colliders: Status and perspectives*. *Rev. Phys.* **5** (ed Di Micco, B.) 100045. arXiv: 1910.00012. http://cds.cern.ch/record/2692014 (2020).

34. Dreyer, F. A. & Karlberg, A. *Vector-boson fusion Higgs pair production at* $\mathrm{N^3LO}$. *Phys. Rev. D* **98,** 114016. https://link.aps.org/doi/10.1103/PhysRevD.98.114016 (11 Dec. 2018).

35. Bishara, F., Contino, R. & Rojo, J. *Higgs pair production in vector-boson fusion at the LHC and beyond*. *The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-5037-9 (July 2017).

36. Manohar, A. V. *Introduction to Effective Field Theories*. 2018. arXiv: 1804.05863 [hep-ph].

37. Heinrich, G., Lang, J. & Scyboz, L. *SMEFT predictions for $gg \to hh$ at full NLO QCD and truncation uncertainties*. *Journal of High Energy Physics* **2022.** https://doi.org/10.1007%2Fjhep08%282022%29079 (Aug. 2022).

38. Brivio, I. & Trott, M. *The standard model as an effective field theory*. *Physics Reports* **793,** 1–98. https://doi.org/10.1016%2Fj.physrep.2018.11.002 (Feb. 2019).

39. Grzadkowski, B., Iskrzyński, M., Misiak, M. & Rosiek, J. *Dimension-six terms in the Standard Model Lagrangian*. *Journal of High Energy Physics* **2010.** https://doi.org/10.1007%2Fjhep10%282010%29085 (Oct. 2010).

40. Gasser, J. & Leutwyler, H. *Chiral perturbation theory: Expansions in the mass of the strange quark*. *Nuclear Physics B* **250,** 465–516. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/0550321385904924 (1985).

41. Capozi, M. & Heinrich, G. *Exploring anomalous couplings in Higgs boson pair production through shape analysis*. *Journal of High Energy Physics* **2020.** https://doi.org/10.1007%2Fjhep03%282020%29091 (Mar. 2020).

42. ATLAS Collaboration. *Constraint of the Higgs boson self-coupling from Higgs boson differential production and decay measurements*. Tech. rep. (CERN, Geneva, 2019). https://cds.cern.ch/record/2667570.

43. Degrassi, G., Giardino, P., Maltoni, F. & Pagani, D. *Probing the Higgs self coupling via single Higgs production at the LHC*. *Journal of High Energy Physics* **2016.** https://doi.org/10.1007%2Fjhep12%282016%29080 (Dec. 2016).

44. Maltoni, F., Pagani, D., Shivaji, A. & Zhao, X. *Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC*. *The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-5410-8 (Dec. 2017).

45. Morrissey, D. E. & Ramsey-Musolf, M. J. *Electroweak baryogenesis*. *New Journal of Physics* **14,** 125003. https://doi.org/10.1088%2F1367-2630%2F14%2F12%2F125003 (Dec. 2012).

46. Hindmarsh, M., Lüben, M., Lumma, J. & Pauly, M. *Phase transitions in the early universe*. https://doi.org/10.21468%2Fscipostphyslectnotes.24 (Feb. 2021).

47. Reichert, M., Eichhorn, A., Gies, H., Pawlowski, J. M., Plehn, T. & Scherer, M. M. *Probing baryogenesis through the Higgs boson self-coupling*. *Phys. Rev. D* **97,** 075008. https://link.aps.org/doi/10.1103/PhysRevD.97.075008 (7 Apr. 2018).

48. Degrassi, G., Di Vita, S., Elias-Miró, J., Espinosa, J., Giudice, G. F., Isidori, G. & Strumia, A. *Higgs mass and vacuum stability in the Standard Model at NNLO*. *Journal of High Energy Physics* **2012,** 98. https://doi.org/10.1007/JHEP08(2012)098 (2012).

49. Di Luzio, L., Gröber, R. & Spannowsky, M. *Maxi-sizing the trilinear Higgs self-coupling: how large could it be? The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-5361-0 (Nov. 2017).

50. Gröber, R. & Mühlleitner, M. *Composite Higgs boson pair production at the LHC*. *Journal of High Energy Physics* **2011.** https://doi.org/10.1007%2Fjhep06%282011%29020 (June 2011).

51. Assamagan, K. *et al. The Higgs Portal and Cosmology* in (Apr. 2016). arXiv: 1604.05324 [hep-ph].

52. Arcadi, G., Djouadi, A. & Raidal, M. *Dark Matter through the Higgs portal*. *Physics Reports* **842,** 1–180. https://doi.org/10.1016%2Fj.physrep.2019.11.003 (Feb. 2020).

53. Giudice, G. F., Grojean, C., Pomarol, A. & Rattazzi, R. *The strongly-interacting light Higgs*. *Journal of High Energy Physics* **2007,** 045–045. https://doi.org/10.1088%2F1126-6708%2F2007%2F06%2F045 (June 2007).

54. ATLAS Collaboration. *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*. *Nature* **607,** 52–59. https://doi.org/10.1038/s41586-022-04893-w (2022).

55. CMS Collaboration. *A portrait of the Higgs boson by the CMS experiment ten years after the discovery*. *Nature* **607,** 60–68. https://doi.org/10.1038/s41586-022-04892-x (2022).

56. Wang, L., Yang, J. M. & Zhang, Y. *Two-Higgs-doublet models in light of current experiments: a brief review*. *Communications in Theoretical Physics* **74,** 097202. https://dx.doi.org/10.1088/1572-9494/ac7fe9 (Aug. 2022).

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

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

59. Mobs, E. *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*. General Photo. https://cds.cern.ch/record/2636343 (2018).

60. CMS Collaboration. *The CMS experiment at the CERN LHC*. *Journal of Instrumentation* **3,** S08004. https://dx.doi.org/10.1088/1748-0221/3/08/S08004 (Aug. 2008).

61. LHCb Collaboration. *The LHCb Detector at the LHC*. *JINST* **3.** Also published by CERN Geneva in 2010, S08005. https://cds.cern.ch/record/1129809 (2008).

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

63. Timko, H. *et al. LHC Longitudinal Beam Dynamics during Run 2*, 245–251. https://cds.cern.ch/record/2750299 (2019).

64. Herr, W. & Muratori, B. *Concept of luminosity*. https://cds.cern.ch/record/941318 (2006).

65. ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. 2022. arXiv: 2212.09379 [hep-ex].

66. ATLAS Collaboaration. *Public ATLAS Luminosity Results for Run-2 of the LHC* https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (2022).

67. Steerenberg, R. *et al. Operation and performance of the CERN Large Hadron Collider during proton Run 2*, MOPMP031. https://cds.cern.ch/record/2696126 (2019).

68. Pequenao, J. *Computer generated image of the whole ATLAS detector* 2008. https://cds.cern.ch/record/1095924.

69. Neutelings, I. *CMS coordinate system*. Website. 2021. https://tikz.net/axis3d_cms/.

70. Neutelings, I. *Pseudorapidity*. Website. 2021. https://tikz.net/axis2d_pseudorapidity/.

71. ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 1* https://cds.cern.ch/record/331063 (CERN, Geneva, 1997).

72. ATLAS Collaboration. *Alignment of the ATLAS Inner Detector in Run 2*. *The European Physical Journal C* **80,** 1194. https://doi.org/10.1140/epjc/s10052-020-08700-6 (2020).

73. ATLAS Collaboration. *Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment* General Photo. 2020. https://cds.cern.ch/record/2723878.

74. ATLAS Collaboration. *Journal of Instrumentation* **13,** T05008. https://dx.doi.org/10.1088/1748-0221/13/05/T05008 (May 2018).

75. ATLAS Collaboration. *ATLAS pixel detector electronics and sensors*. *Journal of Instrumentation* **3,** P07007. https://dx.doi.org/10.1088/1748-0221/3/07/P07007 (July 2008).

76. Capeans, M., Darbo, G., Einsweiller, K., Elsing, M., Flick, T., Garcia-Sciveres, M., Gemme, C., Pernegger, H., Rohne, O. & Vuillermet, R. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. (2010). https://cds.cern.ch/record/1291633.

77. Ahmad, A. *et al. The silicon microstrip sensors of the ATLAS semiconductor tracker*. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **578,** 98–118. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900207007644 (2007).

78. Bingül, A. *The ATLAS TRT and its Performance at LHC*. *Journal of Physics: Conference Series* **347,** 012025. https://dx.doi.org/10.1088/1742-6596/347/1/012025 (Feb. 2012).

79. Mindur, B. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **845.** Proceedings of the Vienna Conference on Instrumentation 2016, 257–261. ISSN: 0168-9002. https://www.sciencedirect.com/science/article/pii/S0168900216301905 (2017).

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

81. ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1* https://cds.cern.ch/record/391176 (CERN, Geneva, 1999).

82. ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report* https://cds.cern.ch/record/331068 (CERN, Geneva, 1997).

83. ATLAS Collaboration. *Operation of the ATLAS trigger system in Run 2*. *Journal of Instrumentation* **15,** P10004–P10004. https://doi.org/10.1088%2F1748-0221%2F15%2F10%2Fp10004 (Oct. 2020).

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

85. ATLAS Collaboration. *Performance of the ATLAS muon triggers in Run 2*. *Journal of Instrumentation* **15,** P09015–P09015. https://doi.org/10.1088%2F1748-0221%2F15%2F09%2Fp09015 (Sept. 2020).

86. Simoniello, R. *The ATLAS Level-1 Topological Processor: from design to routine usage in Run-2*. Tech. rep. (CERN, Geneva, 2019). https://cds.cern.ch/record/2649959.

87. ATLAS Collaboration. *The ATLAS Simulation Infrastructure*. *The European Physical Journal C* **70,** 823–874. https://doi.org/10.1140%2Fepjc%2Fs10052-010-1429-9 (Sept. 2010).

88. Buckley, A. *et al. General-purpose event generators for LHC physics*. *Physics Reports* **504,** 145–233. https://doi.org/10.1016%2Fj.physrep.2011.03.005 (July 2011).

89. Bierlich, C. *et al. A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph].

90. Alwall, J., Frederix, R., Frixione, S., Hirschi, V., Maltoni, F., Mattelaer, O., Shao, H.-S., Stelzer, T., Torrielli, P. & Zaro, M. *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.** https://doi.org/10.1007%2Fjhep07%282014%29079 (July 2014).

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

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

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

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

95. Sjöstrand, T. *Jet fragmentation of multiparton configurations in a string framework*. *Nuclear Physics B* **248,** 469–502. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/0550321384906072 (1984).

96. Andersson, B., Gustafson, G. & Söderberg, B. *A probability measure on parton and string states*. *Nuclear Physics B* **264,** 29–59. ISSN: 0550-3213. https://www.sciencedirect.com/science/article/pii/0550321386904712 (1986).

97. Marchesini, G., Webber, B., Abbiendi, G., Knowles, I., Seymour, M. & Stanco, L. *HERWIG 5.1 - a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons*. *Computer Physics Communications* **67,** 465–508. ISSN: 0010-4655. https://www.sciencedirect.com/science/article/pii/0010465592900554 (1992).

98. Amati, D. & Veneziano, G. *Preconfinement as a property of perturbative QCD*. *Physics Letters B* **83,** 87–92. ISSN: 0370-2693. https://www.sciencedirect.com/science/article/pii/0370269379908967 (1979).

99. Winter, J.-C., Krauss, F. & Soff, G. *A modified cluster-hadronisation model*. *The European Physical Journal C* **36,** 381–395. https://doi.org/10.1140%2Fepjc%2Fs2004-01960-8 (Aug. 2004).

100. Sjöstrand, T., Ask, S., Christiansen, J. R., Corke, R., Desai, N., Ilten, P., Mrenna, S., Prestel, S., Rasmussen, C. O. & Skands, P. Z. *An introduction to PYTHIA 8.2*. *Comput. Phys. Commun.* **191,** 159–177. arXiv: 1410.3012 [hep-ph] (2015).

101. Bellm, J. *et al. Herwig 7.0/Herwig++ 3.0 release note*. *The European Physical Journal C* **76.** https://doi.org/10.1140%2Fepjc%2Fs10052-016-4018-8 (Apr. 2016).

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

103. Pequenao, J. & Schaffner, P. *How ATLAS detects particles: diagram of particle paths in the detector* 2013. https://cds.cern.ch/record/1505342.

104. ATLAS Collaboration. *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*. *The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-5225-7 (Oct. 2017).

105. ATLAS Collaboration. *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*. *The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-4887-5 (May 2017).

106. ATLAS Collaboration. *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*. *The European Physical Journal C* **77.** https://doi.org/10.1140%2Fepjc%2Fs10052-017-5004-5 (July 2017).

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

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

109. Blazey, G. C. *et al. Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*. 2000. arXiv: hep-ex/0005012 [hep-ex].

110. Salam, G. P. & Soyez, G. *A practical seedless infrared-safe cone jet algorithm*. *Journal of High Energy Physics* **2007,** 086–086. https://doi.org/10.1088%2F1126-6708%2F2007%2F05%2F086 (May 2007).

111. Cacciari, M., Salam, G. P. & Soyez, G. *The anti-k_t jet clustering algorithm*. *JHEP* **04,** 063. arXiv: 0802.1189 [hep-ph] (2008).

112. Ellis, S. D. & Soper, D. E. *Successive combination jet algorithm for hadron collisions*. *Physical Review D* **48,** 3160–3166. https://doi.org/10.1103%2Fphysrevd.48.3160 (Oct. 1993).

113. Dokshitzer, Y., Leder, G., Moretti, S. & Webber, B. *Better jet clustering algorithms*. *Journal of High Energy Physics* **1997,** 001–001. https://doi.org/10.1088%2F1126-6708%2F1997%2F08%2F001 (Aug. 1997).

114. Cacciari, M. & Salam, G. P. *Pileup subtraction using jet areas*. *Physics Letters B* **659,** 119–126. ISSN: 0370-2693. https://www.sciencedirect.com/science/article/pii/S0370269307011094 (2008).

115. ATLAS Collaboration. *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*. *Journal of High Energy Physics* **2017,** 24. https://doi.org/10.1007/JHEP12(2017)024 (2017).

116. ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. (CERN, Geneva, 2014). https://cds.cern.ch/record/1700870.

117. ATLAS Collaboration. *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*. Tech. rep. (CERN, Geneva, 2015). https://cds.cern.ch/record/2037702.

118. Krohn, D., Thaler, J. & Wang, L.-T. *Jet trimming*. *Journal of High Energy Physics* **2010.** https://doi.org/10.1007%2Fjhep02%282010%29084 (Feb. 2010).

119. ATLAS Collaboration. *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*. *Eur. Phys. J. C* **79,** 135. arXiv: 1807.09477 [hep-ex] (2019).

120. ATLAS Collaboration. *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data*. Tech. rep. (CERN, Geneva, 2020). https://cds.cern.ch/record/2724442.

121. Krohn, D., Thaler, J. & Wang, L.-T. *Jets with variable R*. *Journal of High Energy Physics* **2009,** 059–059. https://doi.org/10.1088%2F1126-6708%2F2009%2F06%2F059 (June 2009).

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

123. ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. (CERN, Geneva, 2017). https://cds.cern.ch/record/2273281.

124. ATLAS Collaboration. *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*. 2022. arXiv: 2211.16345 [physics.data-an]. https://arxiv.org/abs/2211.16345.

125. ATLAS Collaboration. *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. (CERN, Geneva, 2017). https://cds.cern.ch/record/2255226.

126. ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. (CERN, Geneva, 2017). https://cds.cern.ch/record/2270366.

127. ATLAS Collaboration. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. (CERN, Geneva, 2018). https://cds.cern.ch/record/2645405.

128. ATLAS Collaboration. *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. Tech. rep. (CERN, Geneva, 2018). https://cds.cern.ch/record/2306649.

129. 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,** 970. https://doi.org/10.1140/epjc/s10052-019-7450-8 (2019).

130. ATLAS Collaboration. *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*. *The European Physical Journal C* **82,** 95. https://doi.org/10.1140/epjc/s10052-021-09843-w (2022).

131. ATLAS Collaboration. *Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 fb^{-1} of ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. 2023. arXiv: 2301.06319 [hep-ex].

132. ATLAS Collaboration. *Search for Resonant and Nonresonant Higgs Boson Pair Production in the $b\bar{b}\tau^+\tau^-$ Decay Channel in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*. *Phys. Rev. Lett.* **121,** 191801. https://link.aps.org/doi/10.1103/PhysRevLett.121.191801 (19 Nov. 2018).

133. ATLAS Collaboration. *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*. *Journal of High Energy Physics* **2023.** https://doi.org/10.1007%2Fjhep07%282023%29040 (July 2023).

134. ATLAS Collaboration. *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment*. *Journal of High Energy Physics* **2018,** 40. https://doi.org/10.1007/JHEP11(2018)040 (2018).

135. ATLAS Collaboration. *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. *Phys. Rev. D* **106,** 052001. https://link.aps.org/doi/10.1103/PhysRevD.106.052001 (5 Sept. 2022).

136. ATLAS Collaboration. *Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector*. *Journal of High Energy Physics* **2019.** https://doi.org/10.1007%2Fjhep04%282019%29092 (Apr. 2019).

137. ATLAS Collaboration. *Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV*. *Journal of High Energy Physics* **2019.** https://doi.org/10.1007%2Fjhep05%282019%29124 (May 2019).

138. ATLAS Collaboration. *Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector*. *The European Physical Journal C* **78,** 1007. https://doi.org/10.1140/epjc/s10052-018-6457-x (2018).

139. ATLAS Collaboration. *Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*. *Physics Letters B* **801,** 135145. https://doi.org/10.1016%2Fj.physletb.2019.135145 (Feb. 2020).

140. CERN. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*. en. 2017. https://e-publishing.cern.ch/index.php/CYRM/issue/view/32.

141. 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*. Tech. rep. (CERN, Geneva, 2021). https://cds.cern.ch/record/2780536.

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

143. ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. (CERN, Geneva, 2014). https://cds.cern.ch/record/1966419.

144. Bähr, M. *et al. Herwig++ physics and manual*. *The European Physical Journal C* **58,** 639–707. https://doi.org/10.1140%2Fepjc%2Fs10052-008-0798-9 (Nov. 2008).

145. *Parton distributions in the LHC era: MMHT 2014 PDFs*. *The European Physical Journal C* **75,** 204. https://doi.org/10.1140/epjc/s10052-015-3397-6 (2015).

146. Gieseke, S., Röhr, C. & Siódmok, A. *Colour reconnections in Herwig++*. *The European Physical Journal C* **72,** 2225. https://doi.org/10.1140/epjc/s10052-012-2225-5 (2012).

147. Ball, R. D. *et al. Parton distributions for the LHC run II*. *Journal of High Energy Physics* **2015.** https://doi.org/10.1007%2Fjhep04%282015%29040 (Apr. 2015).

148. Alioli, S., Nason, P., Oleari, C. & Re, E. *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*. *Journal of High Energy Physics* **2010.** https://doi.org/10.1007%2Fjhep06%282010%29043 (June 2010).

149. Frixione, S., Nason, P. & Oleari, C. *Matching NLO QCD computations with parton shower simulations: the POWHEG method*. *Journal of High Energy Physics* **2007,** 070. https://dx.doi.org/10.1088/1126-6708/2007/11/070 (Nov. 2007).

150. Nason, P. *A new method for combining NLO QCD with shower Monte Carlo algorithms*. *Journal of High Energy Physics* **2004,** 040. https://dx.doi.org/10.1088/1126-6708/2004/11/040 (Dec. 2004).

151. Butterworth, J. *et al. PDF4LHC recommendations for LHC Run II*. *Journal of Physics G: Nuclear and Particle Physics* **43,** 023001. https://doi.org/10.1088%2F0954-3899%2F43%2F2%2F023001 (Jan. 2016).

152. ATLAS Collaboration. *A study of optimal parameter setting for* $MADGRAPH5_AMC@NLO$ *+ PYTHIA 8 matched setup*. Tech. rep. (CERN, Geneva, 2015). https://cds.cern.ch/record/2103221.

153. Alwall, J., Frederix, R., Frixione, S., Hirschi, V., Maltoni, F., Mattelaer, O., Shao, H. .-., Stelzer, T., Torrielli, P. & Zaro, M. *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,** 79. https://doi.org/10.1007/JHEP07(2014)079 (2014).

154. Alwall, J., Herquet, M., Maltoni, F., Mattelaer, O. & Stelzer, T. *MadGraph 5: going beyond*. *Journal of High Energy Physics* **2011,** 128. https://doi.org/10.1007/JHEP06(2011)128 (2011).

155. Degrande, C., Durieux, G., Maltoni, F., Mimasu, K., Vryonidou, E. & Zhang, C. *Automated one-loop computations in the standard model effective field theory*. *Phys. Rev. D* **103,** 096024. https://link.aps.org/doi/10.1103/PhysRevD.103.096024 (9 May 2021).

156. ATLAS Collaboration. *Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector*. Tech. rep. (CERN, Geneva, Feb. 2019). https://cds.cern.ch/record/2665057.

157. Amoroso, S. *et al. Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report*. 2020. arXiv: 2003.01700 [hep-ph]. https://arxiv.org/abs/2003.01700.

158. Buchalla, G., Capozi, M., Celis, A., Heinrich, G. & Scyboz, L. *Higgs boson pair production in non-linear Effective Field Theory with full mt-dependence at NLO QCD*. *Journal of High Energy Physics* **2018.** https://doi.org/10.1007%2Fjhep09%282018%29057 (Sept. 2018).

159. Abe, F. *et al. Measurement of* $\sigma B(W \to e\nu)$ *and* $\sigma B(Z^0 \to e^+e^-)$ *in* $\overline{p}p$ *collisions at* $\sqrt{s} = 1800$ *GeV*. *Phys. Rev. D* **44,** 29–52. https://link.aps.org/doi/10.1103/PhysRevD.44.29 (1 July 1991).

160. 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*. *Phys. Rev. D* **105,** 092002. https://link.aps.org/doi/10.1103/PhysRevD.105.092002 (9 May 2022).

161. Moustakides, G. V. & Basioti, K. *Training Neural Networks for Likelihood/Density Ratio Estimation*. 2019. arXiv: 1911.00405 [eess.SP].

162. Kanamori, T., Hido, S. & Sugiyama, M. *A Least-squares Approach to Direct Importance Estimation*. *Journal of Machine Learning Research* **10,** 1391–1445. http://jmlr.org/papers/v10/kanamori09a.html (2009).

163. Pedregosa, F. *et al. Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).

164. Efron, B. *Bootstrap Methods: Another Look at the Jackknife*. *The Annals of Statistics* **7,** 1–26. https://doi.org/10.1214/aos/1176344552 (1979).

165. ATLAS Collaboration. *Performance of pile-up mitigation techniques for jets in* pp *collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*. *The European Physical Journal C* **76,** 581. https://doi.org/10.1140/epjc/s10052-016-4395-z (2016).

166. ATLAS Collaboration. *Measurement of the Inelastic Proton-Proton Cross Section at* $\sqrt{s} = 13$ TeV *with the ATLAS Detector at the LHC*. *Phys. Rev. Lett.* **117,** 182002. https://link.aps.org/doi/10.1103/PhysRevLett.117.182002 (18 Oct. 2016).

167. ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. (CERN, Geneva, 2019). https://cds.cern.ch/record/2677054.

168. Avoni, G. *et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*. *Journal of Instrumentation* **13,** P07017. https://dx.doi.org/10.1088/1748-0221/13/07/P07017 (July 2018).

169. Baglio, J., Campanario, F., Glaus, S., Mühlleitner, M., Ronca, J. & Spira, M. *$gg \to HH$: Combined uncertainties*. *Phys. Rev. D* **103,** 056002. https://link.aps.org/doi/10.1103/PhysRevD.103.056002 (5 Mar. 2021).

170. CERN. *LHC Cross-section Working Group* https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGHH?rev=70 (2023).

171. Dauncey, P. *Second Year Statistics of Measurements*. Oct. 2016.

172. Tanabashi *et al.*, M. *Review of Particle Physics*. *Phys. Rev. D* **98,** 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001 (3 Aug. 2018).

173. Cranmer, K., Lewis, G., Moneta, L., Shibata, A. & Verkerke, W. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. (New York U., New York, Jan. 2012). https://cds.cern.ch/record/1456844.

174. Stanislaus, B. *Searching for Beyond the Standard Model Resonances in the $HH \to b\bar{b}b\bar{b}$ Final State Using the ATLAS Detector*. Presented 30 Sep 2020 (2020). https://cds.cern.ch/record/2743178.

175. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. *Asymptotic formulae for likelihood-based tests of new physics*. *The European Physical Journal C* **71.** https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0 (Feb. 2011).

176. Neyman, J. & Pearson, E. S. *IX. On the problem of the most efficient tests of statistical hypotheses*. *Phil. Trans. R. Soc. Lond. A* **231,** 289–337 (1933).

177. Wilks, S. S. *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*. *The Annals of Mathematical Statistics* **9,** 60–62. https://doi.org/10.1214/aoms/1177732360 (1938).

178. Wald, A. *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*. eng. *Transactions of the American Mathematical Society* **54,** 426–482. ISSN: 0002-9947 (1943).

179. Cowan, G. *Statistical data analysis* eng. ISBN: 9780198501565 (Clarendon Press, Oxford, 1998).

180. ATLAS Collaboration. *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector supporting note*. Tech. rep. (CERN, Geneva, July 2021). https://cds.cern.ch/record/2775165.

181. 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,** 30. https://doi.org/10.1007/JHEP01(2019)030 (2019).

182. ATLAS Collaboration. *Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. *JHEP* **07.** [Erratum: JHEP 01, 145 (2021), Erratum: JHEP 05, 207 (2021)], 108. arXiv: 2001.05178 [hep-ex] (2020).

183. Bernardi, G. *et al. DiHiggs and single-Higgs combination using full Run2 dataset*. Tech. rep. (CERN, Geneva, 2022). https://cds.cern.ch/record/2807720.

184. ATLAS Collaboration. *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*. 2022. arXiv: 2207.00348 [hep-ex]. https://doi.org/10.48550/arXiv.2207.00348.

185. ATLAS Collaboration. *Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS*

detector. *Journal of High Energy Physics* **2022,** 175. https://doi.org/10.1007/JHEP08(2022)175 (2022).

186. ATLAS Collaboration. *Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s}$ =13 TeV with the ATLAS detector*. *Eur. Phys. J. C* **80.** [Erratum: Eur.Phys.J.C 81, 29 (2021), Erratum: Eur.Phys.J.C 81, 398 (2021)], 957. arXiv: 2004.03447 [hep-ex] (2020).

187. ATLAS Collaboration. *Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*. *Eur. Phys. J. C* **81,** 178. arXiv: 2007.02873 [hep-ex] (2021).

188. ATLAS Collaboration. *Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s} = 13$ TeV*. *Eur. Phys. J. C* **81,** 537. arXiv: 2011.08280 [hep-ex] (2021).

189. ATLAS Collaboration. *Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. *Journal of High Energy Physics* **2022,** 97. https://doi.org/10.1007/JHEP06(2022)097 (2022).

190. ATLAS Collaboration. *Measurements of Higgs boson production by gluon−gluon fusion and vector-boson fusion using $H \to WW^* \to e\nu\mu\nu$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. July 2022. arXiv: 2207.00338 [hep-ex]. https://arxiv.org/abs/2207.00338.

191. ATLAS Collaboration and CMS Collaboration. *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*. *Phys. Rev. Lett.* **114,** 191803. https://link.aps.org/doi/10.1103/PhysRevLett.114.191803 (19 May 2015).

192. Badger, S. *et al. Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*. 2016. arXiv: 1605.04692 [hep-ph].

193. ATLAS Collaboration. *Identification of Boosted Higgs Bosons Decaying Into $b\bar{b}$ With Neural Networks and Variable Radius Subjets in ATLAS*. Tech. rep. (CERN, Geneva, 2020). https://cds.cern.ch/record/2724739.

194. Randall, L. & Sundrum, R. *Large Mass Hierarchy from a Small Extra Dimension*. *Physical Review Letters* **83,** 3370–3373. https://doi.org/10.1103%2Fphysrevlett.83.3370 (Oct. 1999).

195. Sjöstrand, T., Mrenna, S. & Skands, P. *A brief introduction to PYTHIA 8.1*. *Computer Physics Communications* **178,** 852–867. https://doi.org/10.1016%2Fj.cpc.2008.01.036 (June 2008).

196. ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. Tech. rep. (CERN, Geneva, 2018). https://cds.cern.ch/record/2314418.

197. Pinamonti, M., Valery, L., Held, A., Dado, T. & Aly, M. *Documentation for TRExFitter*. Website. 2023. https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

198. James, F. *MINUIT: Function Minimization and Error Analysis Reference Manual*. CERN Program Library Long Writeups. https://cds.cern.ch/record/2296388 (1998).

199. Sumner, F. H., Guiboud-Ribaud, S. H., Mazare, G., Verkerk, C., Monrad-Krohn, L., Friedman, J. H., Jank, W., Buxton, N. J., Böck, R. K. & Dahl, O. J. *Proceedings of the 1974 CERN School of Computing: Godøysund, Norway 11 - 24 Aug 1974. 3rd CERN School of Computing* in (CERN, Geneva, 1974). https://cds.cern.ch/record/186223.

200. ATLAS Collaboration. *Performance of b-Jet Identification in the ATLAS Experiment*. *Journal of Instrumentation* **11,** P04008–P04008. https://doi.org/10.1088%2F1748-0221%2F11%2F04%2Fp04008 (Apr. 2016).

201. Amhis, Y. *et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*. *The European Physical Journal C* **81,** 226. https://doi.org/10.1140/epjc/s10052-020-8156-7 (2021).

202. ATLAS Collaboration. *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*. Tech. rep. (CERN, Geneva, 2022). https://cds.cern.ch/record/2811135.

203. ATLAS Collaboration. *Fast b-tagging at the high-level trigger of the ATLAS experiment in LHC Run 3*. Tech. rep. (2023). arXiv: 2306.09738. https://cds.cern.ch/record/2862499.

204. Aberle, O. *et al. High-Luminosity Large Hadron Collider (HL-LHC): Technical design report* https://cds.cern.ch/record/2749422 (CERN, Geneva, 2020).

205. ATLAS Collaboration. *HL-LHC prospects for the measurement of Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state and combination with the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states at the ATLAS experiment*. Tech. rep. (CERN, Geneva, 2022). https://cds.cern.ch/record/2841244.

206. Benedikt, M. *et al. FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2. Future Circular Collider*. Tech. rep. 2 (CERN, Geneva, 2019). https://cds.cern.ch/record/2651299.

207. Mangano, M. L., Ortona, G. & Selvaggi, M. *Measuring the Higgs self-coupling via Higgs-pair production at a 100 TeV p–p collider*. *The European Physical Journal C* **80,** 1030. https://doi.org/10.1140/epjc/s10052-020-08595-3 (2020).

208. Blondel, A. & Janot, P. *Future strategies for the discovery and the precise measurement of the Higgs self coupling*. 2019. arXiv: 1809.10041 [hep-ph]. https://arxiv.org/abs/1809.10041.

209. Abada, A. *et al. FCC Physics Opportunities*. *The European Physical Journal C* **79,** 474. https://doi.org/10.1140/epjc/s10052-019-6904-3 (2019).