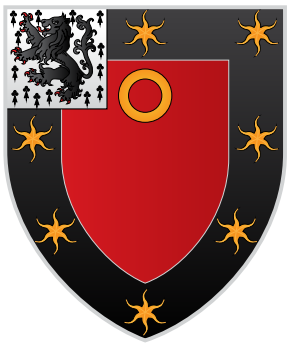
Measurements of Simplified Template Cross Sections in the $H \to WW^*$ decay and interpretations with Effective Field Theory

Gabija Zemaityte

St John's College

University of Oxford



A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Abstract

This thesis discusses the measurement of the Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel and the Effective Field Theory (EFT) interpretations of combined Higgs boson measurements. The results are obtained using 36 fb^{-1} proton-proton collision data produced at the Large Hadron Collider at a centre-of-mass energy of 13 TeV, and recorded by the ATLAS detector in 2015 and 2016. The measured $H \to WW^*$ cross-sections are defined within the Simplified Template Cross Sections (STXS) framework, where the gluon-gluon fusion (ggF) and vector boson fusion production modes are studied. In the case of the ggF production mode the cross-sections are measured as a function of transverse momentum and number of jets. With the $H \to WW^*$ measurements high transverse momentum can be accessed, which is sensitive to new physics. Seven cross-section measurements with the STXS framework are perfomed using the $H \to WW^*$ decay. This thesis also presents the first ATLAS EFT interpretation of combined Higgs boson data in $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels. Finally, the EFT interpretation is also performed including measurements from the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$ channels. The constraints are obtained on the EFT parameters sensitive to modifications of the Higgs boson couplings to strong and electroweak gauge bosons, and to the top quark.

To my sister Mantė.

Acknowledgements

The adversities during my DPhil were constantly followed by a beautiful combination of kindness and courage of other humans.

I would like to thank my supervisor Kathrin Becker who is the supervisor I always wanted to have and the most impressive physicist I met during my DPhil. There are no words to describe her excellence in physics research and pedagogy. I have a lot of gratitude to my supervisor Çiğdem İşsever without whom this thesis would not have been completed, and who has saved this DPhil. I would like to thank Çiğdem for giving a bright vision for my future. I would like to thank my former supervisor Chris Hays for initiating the projects presented in this thesis, for his patience, and a great contribution to my scientific development; the EFT project would be unimaginable without Dr Hays. I am very grateful and very proud of all three supervisors who supervised, mentored, inspired, protected, encouraged, supported me during the DPhil.

I would like to thank Tom Gillam for helping in so many ways before and during the DPhil and for the friendship. It was so much safer to know that if something fails at Oxford Tom Gillam will help.

I would like to thank Kristin Lohwasser for being a "shadow" EFT supervisor. I am very grateful for all Kristin's help and academic support.

I have been impressed and inspired by the following researchers and academics: Çiğdem İşsever, Kathrin Becker, Kristin Lohwasser, Chris Hays, Amanda Cooper-Sarkar, Daniela Bortoletto, Andy Parker, James Frost, Cecilia Tosciri, Claudia

Bertella, Ilaria Brivio, Tony Weidberg, Ian Shipsey, Chris Lester, Michael Trott, Hongtao Yang, and Chris Murphy.

I have a lot of gratitude to Dr Lester for trusting me with ideas, for setting high moral and quality standards in academia, and for invaluable encouragement even years after graduation. I would like to thank Dr Frost and Prof Parker for introducing me to ATLAS. I would like to thank "Fizikos Olimpas" and Prof Bogdanovičius where my physics stepped its first steps.

I would like to thank my physics advisors, mentors and dear friends Aurelijus Rinkevičius and Carsten Burgard. I would like to thank Eric Feng for recognising life problems beneath the physics questions and stepping in to help despite the risk. I would like to thank Koichi Nagai for being in ATLAS.

I would like to thank Amanda Cooper-Sarkar, Todd Huffman, Tony Weidberg, Philipp, Hayden, and Beojan for their help with this thesis. I would like to thank Sue Geddes and Kim Proudfoot for their help with administrative and other matters. I am grateful for the help I received from Ieva G.

I would like to thank St John's College and the Clarendon Fund.

Thanks to this field I met dear Hessam Kaveh, Bill Balunas, Nurfikri Bin Norjoharuddeen, and Marius Utheim.

It was a great gift and Oxford would not be Oxford without the dearest: Kelli Francis-Staite, Catherine Paverd, Henrique Laitenberger, Alex Calin, Yavor Novev, Domenico Giordani, and Andrea Loesel.

I would like to thank my dearest friends Gintautė Vitkauskaitė and Severina Aučinaitė for laughing and crying together, for countless years of friendship. You are like a family to me. Thank you to my dear Lukas and Kotryna Muzika Vaidžiulytė. My dearest Dainius Kilda. To my dear Alex Mortimer. Dear Siddharth Mishra-Sharma who I missed so much. To Tom Gillam because you are among my dearest friends. To my dear Petras Verbyla. To my dear Andi Bode.

I would like to thank my Grandmother Rima and Grandfather Antanas. Asta and Vytas. I would like to thank my Father's Godmother Elenutė for being a Godmother to us all. Finally, I would like to thank my Mother and Mantė, and I always have my Father in my heart.

Preface

The examination regulations of the University of Oxford for the Degree of Doctor of Philosophy state that the student is required to present "a significant and substantial piece of research" [1]. The primary reasons the research presented in this thesis was performed are that the author becomes as close to independent researcher as this study programme provides, acquires knowledge and skills meeting the quality expected by the University of Oxford, and produces original and intellectual contribution to science.

The $H \to WW^*$ measurement presented in chapters 3 and 4 was an excellent project to learn the contemporary experimental particle physics analysis. It was an honour to be supervised by and learn from Dr Kathrin Becker. The starting point of the analysis was the inclusive measurements performed by the ATLAS experiments. The dedicated analysis to measure transverse momentum of the Higgs boson for different jet multiplicities was performed as part of this thesis. Two original methods were developed by the author: the ranking method for a selection of observables for the training of multivariate classifiers (section 3.4.2) and the fit-driven binning method (section 4.2.1). The following work was also performed as part of this thesis: event selection for the $\geqslant$2-jet category and a study on 1-jet category, categorisation of signal regions of these jet categories, event classification using multivariate classifiers, fit optimisation using binning methods. Almost all theoretical uncertainties were calculated as part of this thesis (based on calculations performed in [2]): all gluon-gluon fusion signal uncertainties; WW, top quark, and

$Z \to \tau\tau$ background uncertainties in 1-jet and $\geqslant$2-jet categories. The expected and observed results were also obtained as part of this thesis. The project was initiated by Chris Hays and mainly supervised by Kathrin Becker; the author of this thesis was the only student working on this dedicated Simplified Template Cross Sections analysis.

Chapters 5 and 6 present the second project of this thesis about Effective Field Theory (EFT). Unlike the $H \to WW^*$ measurement case, the EFT analysis was produced almost from scratch. This thesis presents the first ATLAS interpretation of combined Higgs data, which was performed by the supervisor and initiator of the project Chris Hays and by the author of this thesis. The work is made public in [3]. The following work was performed as part of this thesis: dedicated analysis design which was later followed by ATLAS [4] and CMS analyses; generation of Monte Carlo (MC) samples and their translation into analytic EFT equations (made public in [5]), constraints on EFT parameters. The second EFT interpretation was performed outside ATLAS on a public measurement and made public in [5]. The third EFT interpretation using a different EFT model was performed almost without a supervision, and the final EFT interpretation, which includes the measurement presented in chapters 3 and 4, is produced independently by the author of this thesis. Two original methods were created independently by the author of this thesis: the function-of-function method (section 5.3) and the "Wilsons as functions" method (section 5.4). The function-of-function method improves the translation of MC samples into analytic EFT equations, while the "Wilsons as functions" method challenges the status quo of EFT analyses in this field. Most EFT analyses that constrain individual Wilson coefficients rather than their linear combinations typically set the un-fitted coefficients to zero. This induces model dependence, while EFT is designed to be model independent. The "Wilsons as functions" method proposes a systematic way to account for the un-fitted coefficients. The publication of this

method is in preparation.

This thesis does not present PHOTOS generator [6] validation, which was performed as ATLAS qualification task, and evaluation of ggF signal theoretical uncertainties for the $H \to WW^*$ in [2].

Unless stated otherwise, the figures presented in this thesis are produced by the author.

Contents

Introduction

"Magic Theatre
For Madmen Only
Price of Admission – Your Mind"
— Hermann Hesse

Scientific curiosity of the physical world is a powerful force which has shaped the evolution of human civilization. Although the past couple of millennia have been rich in discoveries, fundamental questions about our Universe still remain unanswered: what are the particles of which it consists, what is the form of the fundamental forces, and there are observations which cannot be explained by current theories.

The best theory describing nature at the particle level is called the Standard Model (SM) of particle physics. The last piece of it, the Higgs boson, was discovered in 2012 by the ATLAS [7] and CMS [8, 9] experiments at the Large Hadron Collider (LHC) at CERN. The Higgs boson is the integral part of spontaneous electroweak symmetry breaking [10] which explains the origin of elementary particle masses [11]. However, the SM is insufficient to explain certain phenomena in nature: dark matter, matter-antimatter asymmetry, the cosmological constant and the hierarchy problem. This indicates a need for Beyond Standard Model (BSM) physics.

There are two fundamentally different approaches to search for BSM physics: observation of specific new particles, or interpretation of any deviations observed in precise measurements of quantities predicted by the SM. So far there are no observations of new fundamental particles; this motivates a program of precise measurements of the SM, and Effective Field Theory (EFT) interpretations thereof. The

EFT framework expands the effects of new high-scale perturbative physics in orders of coupling/mass.

This thesis aims to contribute to the scientific knowledge with studies of the Higgs boson. The measurements of Higgs boson properties are a promising probe for new physics, since a new heavy particle that interacts with the Higgs boson would affect its production cross-section and decay width. Deviations from the SM predictions would be a clear indication of new physics. These measurements can be interpreted with the EFT framework in a model independent way.

The first half of this thesis is dedicated to the study of the $H \to WW^*$ decay channel, which provided the most precise Higgs boson cross-section measurements in LHC Run 1 [12]. This thesis presents the measurements in the $H \to WW^* \to e\mu\nu\nu$ channel, where the Higgs boson is produced by gluon-gluon fusion (ggF) and vector boson fusion (VBF) processes. The ggF process is the dominant Higgs boson production mode, and probes Higgs boson couplings to gluons and heavy quarks. VBF is the second-largest mode, and this probes Higgs boson couplings to the W and Z bosons. The $H \to WW^*$ decay has the second-largest branching fraction [13] and only leptonic W decays are considered due to the absence of a large QCD multijet background process. The measurements are performed in the context of the Simplified Template Cross Section (STXS) framework [14], which defines kinematic regions for the Higgs boson cross-section measurements. The STXS framework was designed to reduce theoretical dependencies in a systematic way and increase sensitivity to new high-scale physics. The STXS framework is convenient for BSM interpretations and allows the STXS measurements to be updated in the future by combining all Higgs boson decay channels.

The second half of this thesis is dedicated to the analysis design and the interpretation of the STXS measurements within the EFT framework. This thesis presents the first ATLAS EFT interpretation of the combined Higgs data for the $H \to \gamma\gamma$

and $H \to ZZ^* \to 4\ell$ channels. It also provides novel analysis methods that aim to contribute to the quality and robustness of future EFT analyses. The $H \to WW^*$ measurement is combined with the existing $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels, and is then interpreted in the context of an EFT. All measurements presented here are performed using 36.1 fb^{-1} proton-proton collision data at a centre-of-mass energy of 13 TeV, collected at the ATLAS detector in 2015 and 2016. The limits obtained within the EFT interpretation enrich our knowledge about the model-independent characteristics of new physics.

The theory regarding the SM, the Higgs boson, and EFT is explained in chapter 1. The description of the LHC and the ATLAS detector is presented in chapter 2. The $H \to WW^*$ signal extraction is introduced in chapter 3, which is followed by chapter 4 describing the STXS measurement in this channel. Chapter 5 is dedicated to EFT methodology, while the EFT application to the combination of STXS measurements in $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$ channels is presented in chapter 6. The thesis is closed by a discussion and conclusions in chapter 7.

Chapter 1

Theory

This chapter discusses the theoretical background of experimental particle physics, which covers particle physics theory (the Standard Model, Effective Field Theory), a theoretical framework for experimental analyses (Simplified Template Cross Sections framework) and the machinery used for simulations. First, the Standard Model (SM) is introduced in section 1.1. The Effective Field Theory (EFT) framework is discussed in section 1.2 and the Simplified Template Cross Sections (STXS) framework is introduced in section 1.3. Finally, Monte Carlo simulations are discussed in section 1.4.

1.1 Standard Model

The quantum field theory was developed to describe relativistic and quantum phenomena. One of the most successful quantum field theories is the Standard Model of particle physics, which well explains the experimental observations. It describes three fundamental forces: the electromagnetic, the weak and the strong interactions. The SM consists of force carrier particles, matter particles and the Higgs boson. Each force is mediated by gauge bosons. The electromagnetic interaction acts between charged particles via the exchange of photons. The weak interaction is mediated via the Z and $W^{\pm}$ bosons, where the mediation via the $W^{\pm}$ bosons is

responsible for the change of flavour. The strong force is mediated by gluons and is responsible for binding quarks into hadrons. The matter particles are fermions: neutrinos, charged leptons and quarks. Neutrinos are neutral leptons that interact via the weak force only. Charged leptons interact via the electromagnetic and weak forces, while quarks interact via all three fundamental forces. The Higgs boson generates mass for the $W^{\pm}$ and Z bosons via the Brout-Englert-Higgs (BEH) mechanism [15, 16]. The fundamental particles and the force carriers are illustrated in Figure 1.1.

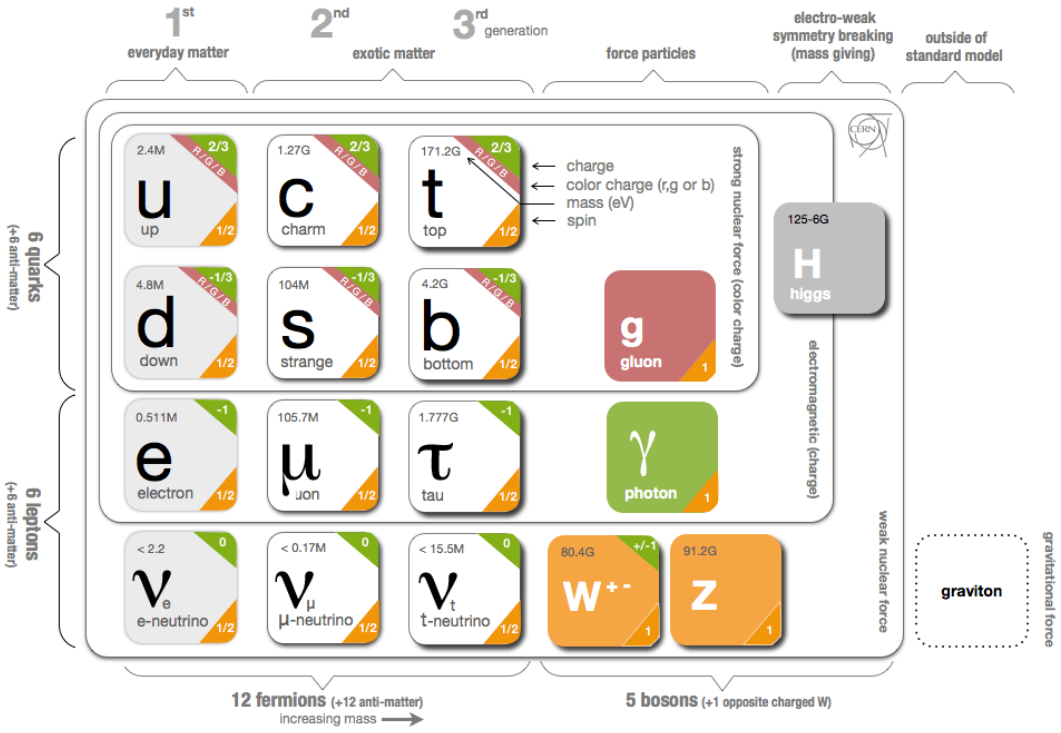


Figure 1.1: Fundamental particles in the Standard Model. Figure is produced by CERN.

The standard model gauge group is the direct product of $SU(3)_C \times SU(2)_L \times U(1)_Y$. The electroweak theory (EW) is described by $SU(2)_L \times U(1)_Y$ and the

quantum chromodynamics (QCD) is descried by $SU(3)_C$.

The BEH mechanism and the electroweak theory are described in section 1.1.1. It is followed by the Higgs boson characteristics discussed in section 1.1.2. The introduction to QCD is presented in section 1.1.3.

1.1.1 Brout-Englert-Higgs mechanism and the electroweak theory

This section explains the mathematics behind the Brout-Englert-Higgs (BEH) mechanism and the electroweak theory. It starts with the presentation of the idea of spontaneous symmetry breaking, which is followed by an explanation of the BEH mechanism. Finally, the electroweak theory is discussed.

Spontaneous symmetry breaking

Spontaneous symmetry breaking is a process through which a system in a symmetric state, obeyed by the Lagrangian, transforms into a state in which the system's perturbations around the vacuum do not have the same symmetry.

Let us consider a Lagrangian with an N-component real scalar field $\phi = (\phi_1, ..., \phi_N)^T$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi \cdot \partial^\mu\phi - V(\phi) \tag{1.1}$$

where the potential is:

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4; \quad \lambda > 0 \tag{1.2}$$

The Lagrangian is invariant under $O(N)$ transformation. Let us consider the case where $m^2 < 0$. Then the potential can be written as:

$$V(\phi) = \frac{\lambda}{4}(\phi^2 - v^2)^2 + \text{const}; \quad v^2 = -\frac{m^2}{\lambda} > 0 \tag{1.3}$$

where v is called the vacuum expectation value. The potential resembles a bottom of a wine bottle (an example of it is illustrated in Figure 1.2). When the potential

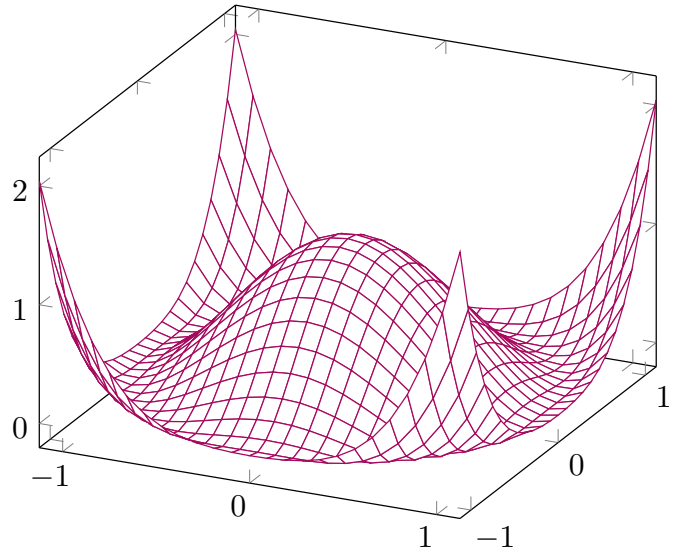


Figure 1.2: An example of $V(\phi) = \frac{\lambda}{4}(\phi^2 - v^2)^2$ potential in two dimensions.

is written in this way, there is a continuum of vacua at $\phi^2 = v^2$.

Without loss of generality let us choose the vacuum $\phi_0 = (0, ..., 0, v)^T$. Let $\pi_i(x)$ and $\sigma(x)$ be small fluctuations around this vacua. Then the scalar field is $\phi = (\pi_1(x), \pi_2(x), ..., \pi_{N-1}, v + \sigma(x))$. The Lagrangian then becomes:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \pi \cdot \partial^\mu \pi + \frac{1}{2}\partial_\mu \sigma \cdot \partial^\mu \sigma - V(\pi, \sigma) \tag{1.4}$$

where

$$V(\pi, \sigma) = \frac{1}{2}m_\sigma^2 \sigma^2 + \lambda v(\sigma^2 + \pi^2)\sigma + \frac{\lambda}{4}(\sigma^2 + \pi^2)^2 \tag{1.5}$$

where $m_\sigma = \sqrt{2\lambda v^2}$. Therefore, there is one massive field $\sigma(x)$ and N-1 massless fields $\pi(x)$. The $\sigma(x)$ corresponds to the radial excitation and thus has a potential barrier. The $\pi(x)$ fields are excitations in the equipotential space, hence do not have a potential barrier (in two dimensions the equipotentials are concentric rings).

Brout-Englert-Higgs mechanism

The Brout-Englert-Higgs mechanism explains how massive particles acquire mass. It is based on the idea of spontaneous symmetry breaking and the mathematical derivation is the following.

Let us look at the complex scalar field's $\phi(x)$ interactions with a photon field $A_\mu(x)$ and itself. The interactions are described with the following Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - V(\phi^*\phi) \tag{1.6}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is a field strength tensor and $D_\mu \equiv \partial_\mu + iqA_\mu$ is a covariant derivative. Let

$$V(\phi^*\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4, \text{ with } \lambda > 0 \tag{1.7}$$

If we consider the case $\mu^2 < 0$, the resulting potential is similar to the one that is described in the previous section. It has the vacua at:

$$|\phi_0|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{1.8}$$

which can be parameterised as:

$$\phi_0 = \frac{v}{\sqrt{2}}e^{i\zeta_0/v}, \text{ with } v > 0 \text{ and } \zeta_0 \in \mathbb{R} \tag{1.9}$$

Let $\zeta(x)$ and $\eta(x)$ be fluctuations around the vacuum configuration $\zeta_0 = 0$, then the field becomes:

$$\phi(x) = \frac{e^{i\zeta(x)/v}}{\sqrt{2}}(v + \eta(x)) \approx \frac{1}{\sqrt{2}}(v + \eta + i\zeta) \tag{1.10}$$

After the field is substituted into the Lagrangian, it shows that in this configuration η and A_μ are massive fields and ζ is a massless field.

When a unitary gauge transformation is made such that $\alpha(x) = -\frac{1}{v}\zeta(x)$, then the quadratic part of the Lagrangian becomes:

$$\mathcal{L}_{quad} = \frac{1}{2}(\partial_\mu\eta\partial^\mu\eta + 2\mu^2\eta^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{q^2v^2}{2}A_\mu A^\mu \tag{1.11}$$

Here ζ is not a massless field anymore since it is absorbed into A_μ and becomes the longitudinal polarisation of A_μ.

Electroweak theory

The previous sections lead to the discussion about the electroweak theory. In the SM, the electroweak symmetry breaking is responsible for generating masses for the $W^{\pm}$ and Z bosons. The electroweak Lagrangian is invariant under $SU(2) \times U(1)$ gauge symmetry and is constructed as:

$$\mathcal{L} = -\frac{1}{4}\mathrm{Tr}F^W_{\mu\nu}F^{W,\mu\nu} - \frac{1}{4}F^B_{\mu\nu}F^{B,\mu\nu} + (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2|\phi|^2 - \lambda|\phi|^4 \tag{1.12}$$

where $F^W_{\mu\nu}$, $F^B_{\mu\nu}$ are the field strength tensors and ϕ is a scalar field. The covariant derivative is defined as:

$$D_\mu\phi = (\partial_\mu + igW^a_\mu\tau^a + \frac{i}{2}g'B_\mu)\phi \tag{1.13}$$

where W^a_μ corresponds to three gauge bosons associated with $SU(2)_L$ gauge symmetry and B_μ corresponds to $U(1)_Y$ gauge boson. The τ^a is an $SU(2)$ generator $\tau^a = \frac{\sigma^a}{2}$, where σ^a is a Pauli matrix. The $SU(2)$ field strength tensor is:

$$F^{W,a}_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\varepsilon^{abc}W^b_\mu W^c_\nu \tag{1.14}$$

and the $U(1)$ field strength tensor is:

$$F^B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.15}$$

The scalar field ϕ, or the Higgs field, is defined in the doublet representation of $SU(2)_L$ with $U(1)_Y$ hypercharge of $\frac{1}{2}$. As a result, the unbroken theory has 3+1 massless gauge bosons. As seen before, the vacua of the potential with $\lambda > 0$ is:

$$|\phi_0|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{1.16}$$

Without loss of generality, let us choose the following vacuum doublet:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{1.17}$$

As a result, SM gauge symmetry $SU(2)_L \times U(1)_Y$ is broken into $U(1)_{em}$: the three out of four generators of the $SU(2)_L \times U(1)_Y$ gauge group are spontaneously broken. This results in three massless Goldstone bosons. The Higgs field couples to the W_μ and B_μ gauge fields through the following covariant derivative:

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2}\frac{v^2}{4}[g^2(W^1)^2 + g(W^2)^2 + (-gW^3 + g'B)^2] \tag{1.18}$$

In the unitary gauge the massless Goldstone bosons become the longitudinal components of the Z and $W^\pm$ gauge bosons. The fields W^i and B are redefined to represent the physical gauge bosons:

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \\ Z^0_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gW^3_\mu - g'B_\mu) \\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g'W^3_\mu + gB_\mu) \end{aligned} \tag{1.19}$$

The $W^\pm_\mu$ fields have mass $m_W = vg/2$ and correspond to $W^\pm$ bosons, the Z^0_μ field has mass $m_Z = v\sqrt{g^2+g'^2}/2$ and corresponds to the Z boson, while A_μ field is massless and corresponds to the photon. The massless A_μ is a result of the fourth unbroken generator associated with the $U(1)_{em}$ gauge symmetry. The Weinberg angle is defined to be

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}} \tag{1.20}$$

and the $W^\pm$ boson mass relates to the Z boson mass as $m_W = m_Z \cos\theta_W$.

1.1.2 Higgs boson

In 2012 the ATLAS and CMS experiments at the LHC discovered a particle compatible with the SM Higgs boson. The particle was discovered using the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ decay channels. The SM Higgs boson has a spin-parity of 0^+ and its coupling to the fundamental particles depends on their masses: it couples weakly to the light particles and strongly to the heavy particles, like the $W^\pm$ and Z bosons and

the top quark. The Higgs boson mass is measured to be 125.09 ± 0.21 (stat)$\pm$0.11 (syst) GeV [17].

The Higgs boson properties are accessed by measuring its production and decay modes, which are presented below.

Higgs boson production modes

The Higgs boson is produced via the main four production modes: the gluon-gluon fusion (ggF), the vector boson fusion (VBF), the associated production with vector bosons (VH) and the associated production with top quarks ($t\bar{t}$H) [18].

In the case of the ggF production mode, the Higgs boson is produced via a heavy quarks' (top or bottom) loop, which is generated by gluons. It is the largest Higgs boson production cross-section at the LHC. This production process probes Higgs boson couplings to heavy quarks. The ggF production is illustrated in Figure 1.3.

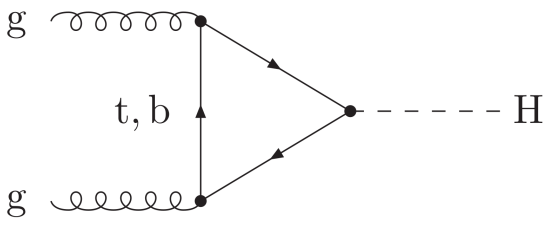


Figure 1.3: The ggF Higgs boson production mode. Figure is taken from [18].

In the case of the vector boson fusion mode, the Higgs boson is produced as a result of a fusion of two $W^{\pm}$ or Z bosons. Effectively, it is a scattering of two (anti)quarks that exchange $W^{\pm}$ or Z boson in t or u channel with the Higgs boson radiated off them. This production process probes Higgs boson couplings to $W^{\pm}$ or Z bosons. The VBF production is illustrated in Figure 1.4.

In the case of the associated Higgs production with vector bosons, the Higgs boson and Z or $W^{\pm}$ boson are final state particles. This production probes Higgs boson couplings to $W^{\pm}$ and Z bosons. The associated Higgs boson production with

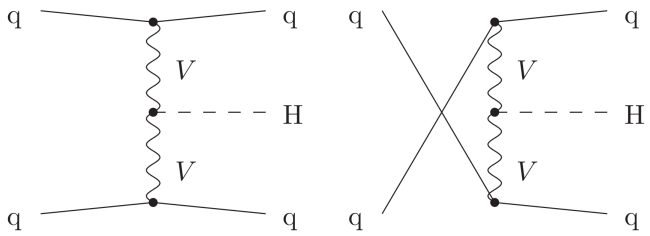


Figure 1.4: The VBF Higgs boson production mode. Figure is taken from [18].

vector bosons is illustrated in Figure 1.5.

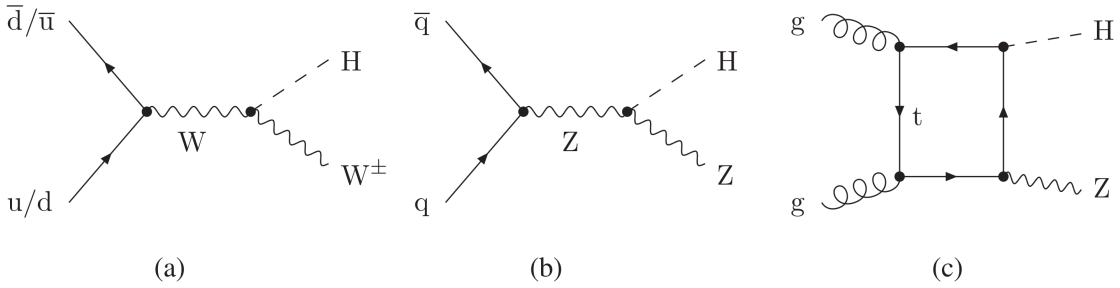


Figure 1.5: The VH Higgs boson production mode. Figure is taken from [18].

In the case of the associated production with top quarks, the final state contains the Higgs boson and a top anti-top pair. It provides a direct probe of the Higgs-top Yukawa coupling. The associated production with top quarks is illustrated in Figure 1.6.

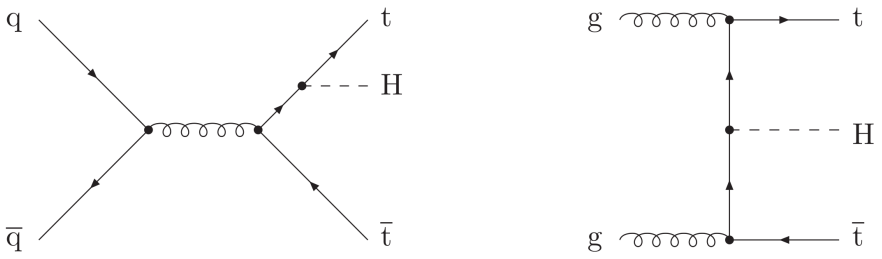


Figure 1.6: The $t\bar{t}$H Higgs boson production mode. Figure is taken from [18].

Higgs boson decay modes

The Higgs boson has a short life time (of order 10^{-22} s [14]), so it is measured via the particles it decays into. Since Higgs boson's electric charge is zero, the total charge of its decay products should be neutral. The predicted total width of the SM Higgs boson with mass of m_H =125 GeV is $\Gamma_H = 4.07 \times 10^{-3}$ GeV (with uncertainty $^{+4.0\%}_{-3.9\%}$) [14]. The leading Higgs boson decays are the decays into a pair of $b\bar{b}$ quarks with a branching ratio (BR) of 58% and $H \to WW^*$ with BR of 21% [14]. They are followed by $H \to gg$ with BR of 8.2%, $H \to \tau^+\tau^-$ with BR of 6.3%, $H \to c\bar{c}$ with BR of 2.9%, and $H \to ZZ^*$ with BR of 2.6% [14]. The even smaller decay rates are $H \to \gamma\gamma$ with BR of 0.23%, $H \to \gamma Z$ with 0.15% BR, and $H \to \mu^+\mu^-$ with 0.021% BR [14]. The BRs are given for the SM Higgs of mass 125 GeV. The Higgs boson BRs for the mass range around 125 GeV are illustrated in Figure 1.7.

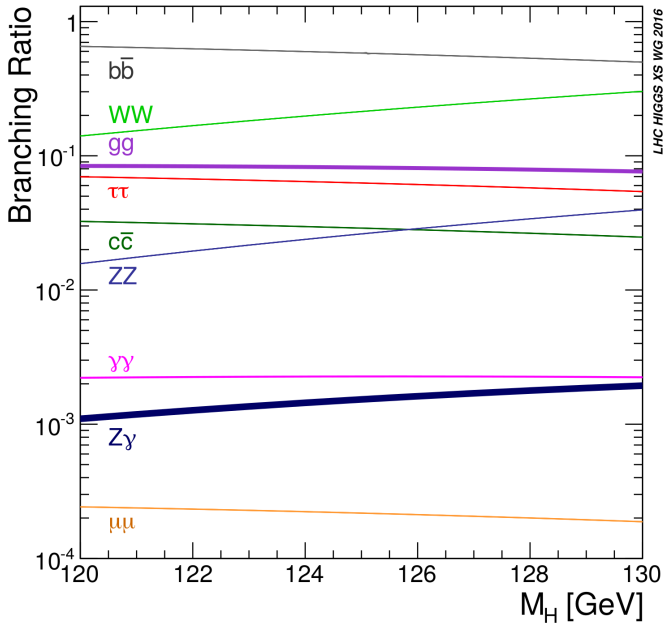


Figure 1.7: Higgs boson branching ratios and their uncertainties for the mass range around 125 GeV. Figure is taken from [14].

The sensitivity of a channel depends on the production cross-section, the decay

branching ratio, the reconstructed mass resolution, the selection efficiency and the background composition. In case of $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ decay modes, all final states can be measured precisely and has a very good Higgs mass resolution of 1-2% [13]. The $H \to WW^* \to \ell^+\nu\ell^-\bar{\nu}$ decay has a poor Higgs mass resolution of about 20% due to the neutrinos in the final state [13]. The $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ decay modes are the most promising channels to probe the Higgs boson couplings to quarks and leptons but these decay modes typically have large background contributions and the mass resolution of about 10% and 15%, respectively [13].

1.1.3 Quantum Chromodynamics

The quantum chromodynamics (QCD) is a locally invariant quantum field theory under the $SU(3)_C$ gauge group. The interaction is carried by the exchange of eight colour charge carriers - gluons. It is a non-abelian gauge theory meaning that self interactions of gluons are allowed [19].

The QCD experiences a colour confinement phenomenon which means that only colourless particles can be found in isolation and observed. The field transforms as:

$$\psi(x) \to \psi'(x) = e^{igt\cdot\theta(x)}\psi(x) \tag{1.21}$$

where g is the strong coupling constant, t^a is a group generator, and θ_a is a space-time phase function. The gluon field strength tensor is:

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f^{abc} A_b^\mu A_c^\nu \tag{1.22}$$

where f^{abc} is an $SU(3)$ structure constant. The Lagrangian density is:

$$\mathcal{L}_{QCD} = \sum_f \bar{\psi}_f^i (i\gamma_\mu D^\mu - m_f)_{ij}\psi_f^j - \frac{1}{4}F_a^{\mu\nu}F_{\mu\nu}^a \tag{1.23}$$

where m_f is a mass parameter and the covariant derivative is:

$$D_{ij}^\mu = \delta_{ij}\partial^\mu + ig(t^a)_{ij}A_a^\mu \tag{1.24}$$

1.2 Effective Field Theory

In the history of experimental particle physics, there are plenty of precision measurements consistent with the SM and the discoveries of particles are no longer frequent. The last three discovered particles are the top quark in 1995, the τ neutrino in 2000, and the Higgs boson in 2012 [20]. The direct searches for new physics, e.g. for dark matter, supersymmetry and other exotic particles, are fruitless to date. An effective field theory (EFT) is a model independent framework designed to interpret measurements as indirect searches for Beyond Standard Model (BSM) physics in a systematic way [20].

The main idea of EFT is the separation of IR (infrared or long distance) and UV (ultraviolet or short distance) physics. The aim of a well designed EFT is to capture the BSM effects on the IR regime since the full BSM theory is simplified by ignoring the short distance effects produced by the particles much heavier than the SM particles and the energy scale of the experiment [14]. The UV physics effects are systematically removed and therefore the underlying UV theory does not have to be known, which is an advantage of EFT. The choice of an EFT defines the UV suppression scale below which physics can be parameterised by the EFT. An EFT Lagrangian is constructed of dimension d operators that are suppressed by $d-4$ powers of the UV suppression scale. The EFT Lagrangian is a sum of operators O_i^{IR} that affect IR physics, and the scalar parameters c_i^{UV} that encapsulate UV physics are called the Wilson coefficients [20]:

$$\mathcal{L} = \sum_i c_i^{UV} O_i^{IR}. \tag{1.25}$$

Non-local interactions via the heavy particle exchange are replaced by the local interactions. This restricts the validity of the theory to below a certain energy scale [21]. One of the advantages of EFT is the possibility to re-interpret the EFT constraints on masses and couplings in different BSM theories [14].

The two EFT frameworks relevant to the research presented in this thesis are the Standard Model Effective Field Theory (SMEFT) and the Higgs Effective Field Theory (HEFT) [20]. The SMEFT is an EFT framework which describes SM particles' interactions and the Higgs field is introduced as the SM doublet. In the case of the HEFT, the Higgs doublet is not present [20]. The SMEFT is described in the following section in more detail.

1.2.1 The Standard Model Effective Field Theory

The SMEFT is a powerful and consistent EFT, which describes the lower energy limit of the BSM physics. The SMEFT consists of SM fields and the Higgs field modelled as the SM Higgs doublet. The higher dimensional operators are invariant under $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge group. The EFT Lagrangian is defined as:

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \mathcal{L}^{(7)} + ... \tag{1.26}$$

where higher order Lagrangian terms are:

$$\mathcal{L}^{(d)} = \sum_{i=1}^{N} \frac{c_i^{(d)}}{\Lambda^{d-4}} O_i^{(d)} \quad \text{for } d > 4 \tag{1.27}$$

where $O_i^{(d)}$ is a dimension d operator, $c_i^{(d)}$ is the corresponding Wilson coefficient, and Λ is the suppression scale [20]. A complete and non-redundant set of operators form a basis. Complete bases predict the same new physics effects [14].

All dimension-5 operators violate lepton number conservation [22], while all odd dimension operators violate baryon number minus lepton number [23]. Given current experimental constraints, the Wilson coefficients corresponding to these operators are not observable at the LHC [24].

In the context of the Brout-Englert-Higgs mechanism (section 1.1.1), the SMEFT brings a correction to the Higgs potential. In the SMEFT the EW symmetry is broken as in the SM, $SU(2)_L \times U(1)_Y \rightarrow U(1)_{em}$. In the Warsaw basis [25] with

dimension-6 operators the Higgs potential is modified by the operator $O_H^{(6)} \equiv (\phi^\dagger\phi)^3$ [20]. In this case the potential is:

$$V(\phi^\dagger\phi) = \lambda\Big(\phi^\dagger\phi - \frac{1}{2}v^2\Big)^2 - c_H^{(6)}(\phi^\dagger\phi)^3 \tag{1.28}$$

where ϕ is the Higgs doublet and $c_H^{(6)}$ is a Wilson coefficient corresponding to $O_H^{(6)}$ operator. The new minimum of the Higgs potential is:

$$\langle\phi^\dagger\phi\rangle = \frac{v^2}{2}\Big(1 + \frac{3c_H^{(6)}v^2}{4\lambda}\Big) \equiv \frac{1}{2}v_T^2 \tag{1.29}$$

1.3 Simplified Template Cross Sections

Simplified Template Cross Sections (STXS) [14] is a framework defined for Higgs boson production and decay measurements at particle level, where particle level refers to a Monte Carlo (MC) simulation, which includes all but detector simulation stages (section 1.4.2). The kinematic regions are designed such that the cross-sections have minimised dependence on theoretical uncertainties, maximised experimental sensitivity, and the framework includes dedicated regions potentially sensitive to BSM physics. The minimised theoretical uncertainties include both SM predictions and dependence on physics models. STXS is a convenient framework for a combination of multiple Higgs boson decay channels and for the BSM interpretations.

The STXS VH and VBF processes are defined differently than the production modes presented in section 1.1.2 because at higher orders VH production in association with hadronically decaying V and VBF definitions are ambiguous. As a result, the STXS VH mode is defined as a VH production in association with a leptonically decaying V, while the VH production, where V decays hadronically, is defined to be a part of the VBF production. The STXS VBF production is also called qqH production.

There are multiple stages of the STXS framework defined, which differ by the granularity of kinematic regions. The higher the stage, the finer the granularity of

the kinematic regions of the Higgs boson production modes. Stage 0 contains the inclusive ggF, $t\bar{t}$H, $b\bar{b}$H, and tH production cross-sections; the VBF and VH production modes have two and three kinematic regions, respectively. Stage 1 has a finer binning than stage 0. The Higgs boson production modes and their corresponding stage 1 kinematic regions are summarised in Figure 1.8.

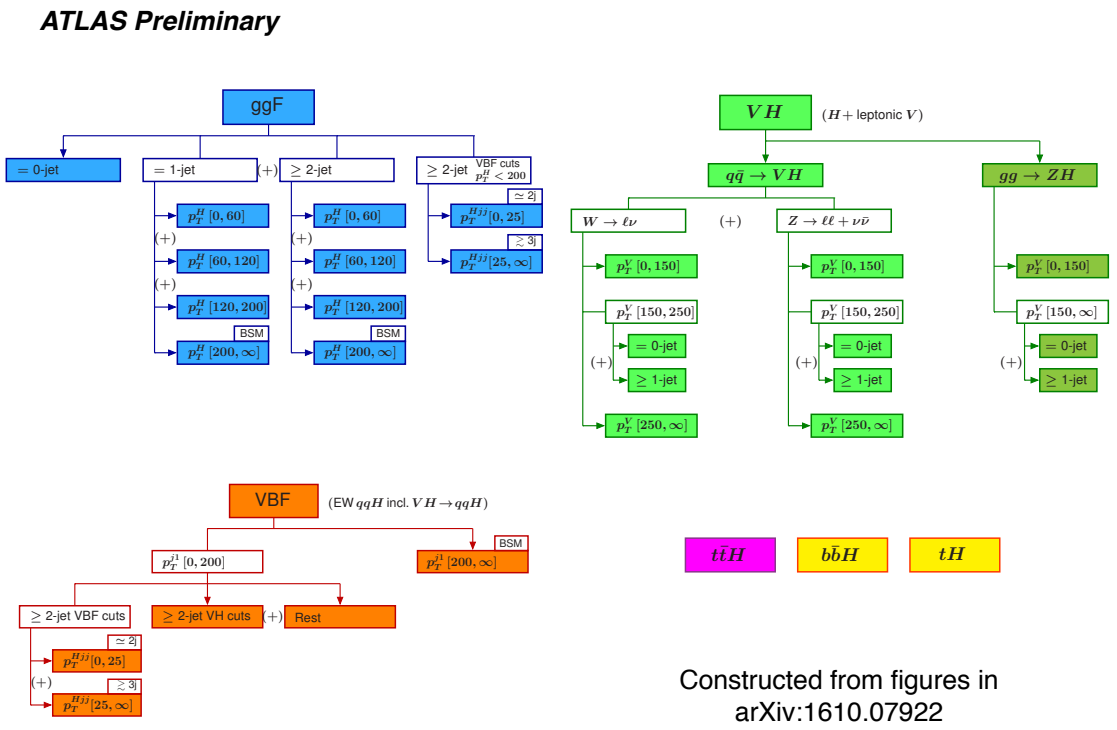


Figure 1.8: STXS stage 1 kinematic regions. The figure is constructed from figures in [14] and taken from [26].

The stage 1 ggF production mode is most relevant to this thesis. This is separated into jet multiplicity categories of zero, one, and two or higher. These are further split based on kinematic observables. There is one inclusive zero jet kinematic region; one jet category is split into four kinematic regions based on the Higgs boson transverse momentum p_T^H at 60 GeV, 120 GeV, and 200 GeV. The two or more jets category

is first split at $p_T^H = 200$ GeV, where only the $p_T^H < 200$ GeV region is split further. The $p_T^H < 200$ GeV region is first split into ggF-like and VBF-like regions, where VBF-like region is defined as $m_{jj} > 400$ GeV and $\Delta\eta_{jj} > 2.8$, where m_{jj} and $\Delta\eta_{jj}$ are the invariant mass and the difference in pseudorapidity of two leading jets, while the remaining phase space is the ggF-like region. The ggF-like region is further split into three regions at p_T^H equal to 60 GeV and 120 GeV. The VBF-like region is split into two regions at $p_T^{Hjj} = 25$ GeV, where p_T^{Hjj} is a transverse momentum of the system of the Higgs boson and two leading jets. The regions satisfying the $p_T^H > 200$ GeV criterion are expected to be sensitive to BSM physics. Each kinematic region has a cut on Higgs rapidity $|y_H| < 2.5$. The remaining region corresponds to the forward detector (like ATLAS or CMS) region $|y_H| > 2.5$ which is not included in the STXS definition. The ggF production kinematic regions are summarised in Figure 1.9.

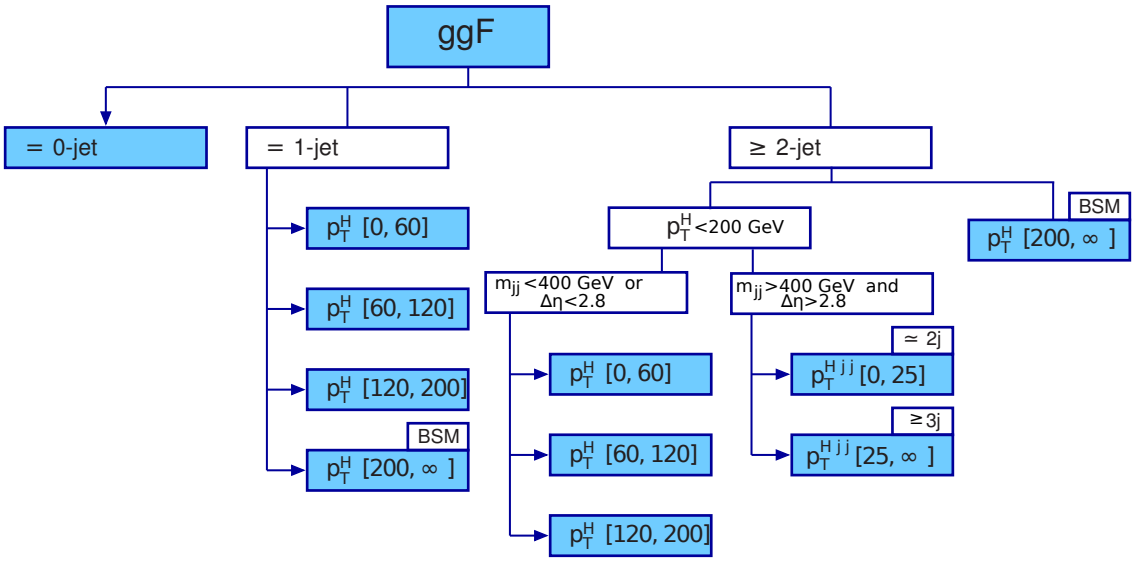


Figure 1.9: Summary of STXS stage 1 ggF kinematic regions. The figure is a modified version of the ggF part in Figure 1.8.

1.4 Monte Carlo simulations

The experimental analysis is based on the data comparison with the theory predictions. The predictions are produced using the Monte Carlo (MC) simulations. One particle collision is called an event. The generated events aim to imitate nature as closely as possible, where the MC generators provide all the necessary information for the experiment: the particle types, their momenta, and the absolute production rates [27].

First, section 1.4.1 discusses theoretical background for the MC modelling. It is followed by section 1.4.2, which describes different stages of the event generation. Finally, the MC generators are described in section 1.4.3.

1.4.1 Theoretical aspects of Monte Carlo simulation

This section discusses the relevant theoretical aspects used for MC simulations, including renormalisation, factorisation, and parton distribution functions.

Renormalisation

Renormalisation is a regulation which addresses the problem of divergences, which arise due to unrestricted integration over the momentum flow in the Feynman diagrams containing loops. The renormalisable theory is a theory which has a finite number of parameters which ensures finite results in all orders of a perturbation theory. The SM gauge theories are renormalisable [19].

In QCD quarks and gluons contribute to loop diagrams. The equation of QCD coupling running is:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2) b_0 \ln\left(\frac{Q^2}{\mu_R^2}\right)} \tag{1.30}$$

where Q is a scale parameter, $b_0 > 0$ is a constant, and μ_R is a scale at which the renormalised coupling is defined [19].

The asymptotic freedom is a condition when $\alpha_s(Q^2) \to 0$ as $Q^2 \to \infty$. It means that at large scale (or a short distance) strongly interacting particles act as free particles. One example is a hard process [19].

Factorisation

Factorisation is an approach where the cross-section is expressed as a convolution of a hard process and a parton density [19]. Due to factorisation the proton structure and the final state hadron creation are decoupled [27].

Parton distribution function

A parton distribution function (PDF) $f_i(x, \mu_F)$ is a number density of a parton i which has momentum fraction x with respect to its parent hadron and is probed at a factorisation scale μ_F. The PDFs are used in the simulation of hard processes, parton showers and multiple parton interactions (section 1.4.2). They affect both the cross-sections and the event kinematic properties. The PDFs are extracted from lepton-hadron and hadron-hadron collision data sets. The PDFs' evolution with momentum scale is described by the DGLAP equations [28].

1.4.2 Stages of event generation

The MC generation starts with the simulation of the hard process using perturbative QCD. The next stage is the parton shower associated with the incoming and outgoing coloured particles, which naturally leads to the hadron formation from the coloured participants, which is in a non-perturbative QCD regime. The secondary interactions to the hard scattering are modelled as underlying events.

All these production stages are described in more detail in this section. The section is finished by a discussion on the merging and matching schemes, which allow to combine the different stages of the event generation.

Parton level event generation

The parton level event generation considers the interactions of elementary constituents of protons: quarks and gluons. The fully differential cross-section of a hadron-hadron collision is:

$$\frac{d\sigma}{d\Omega dx_1 dx_2} = \frac{1}{\text{Flux}} \sum_{i,j,k,...n} f_i(x_1, \mu_F) f_j(x_2, \mu_F) \times \sum |M(i, j \to k, l, ...n)|^2 \quad (1.31)$$

where f_i is a PDF and M is a matrix element.

The hard scattering happens over very short timescales, therefore the details of the interactions within the proton and its evolution after it are irrelevant. The dynamics within the proton can be seen as static and therefore the hard process can be described as an interaction between freely moving particles. The hard scattering process is calculated perturbatively using expansion in powers of the strong coupling constant $\alpha_s(\mu_R)$, since it is small at high energy scales. Calculations are done up to a finite order in $\alpha_s(\mu_R)$ which leads to the intrinsic uncertainty. The uncertainty depends on the renormalisation scale μ_R of the coupling constant $\alpha_s(\mu_R)$ and the factorisation scale μ_F of the parton densities. In the uncertainty calculations μ_R is typically varied in the range of $\mu_0/2 < \mu_R < 2\mu_0$, where μ_0 is the nominal scale of the hard process [27].

Parton shower generation

Parton shower algorithms are used to model the internal structure of jets and accompanying particles. The parton shower generators model QCD radiation produced by accelerated coloured partons and which is calculated by perturbative QCD [27]. Gluons can radiate other gluons and quark-antiquark pairs, which can be modelled using step-wise Markov chain approach. The parton shower mainly consists of partons that are either soft or collinear with the outgoing partons [28]. The hard process is generated first and then the parton shower dresses the particles with additional

radiation. These algorithms evolve the momentum transfer from the high scales of the hard process to the low scales of about 1 GeV, associated to hadron formation [28].

Hadronisation

The hadron formation occurs at the end of parton shower: it is a transition from the partonic final state to a complete representation of a hadronic final state. As the distance between moving partons increases, the strong force grows and the confinement effect becomes important. As a result, the partons form colour neutral hadrons [27]. The hadronisation is in the non-perturbative regime: it cannot be computed from the first principles, so the hadronisation models should be used [28]. The two most common models of hadronisation are the string fragmentation and the cluster models [27]. The main difference is that the string model directly transforms the system into hadrons, while the cluster model first creates clusters that are further used to define hadrons [28]. The hadronisation is independent of the initial hard scattering process due to factorisation assumption. As a result, the tuning using one data set can be used for another data set and thus it can be later reused for other hard processes.

Jet production

A jet is a structure where most of the energy is carried localised in collinear bundles of partons. It originates from the final state of the parton shower which is hadronised, where the hadronisation procedure preserves the jet structure. The total momentum of constituent hadrons is described using perturbative calculations. Since jets are not fundamental objects, they are defined by jet algorithms instead of directly by theoretical models [28].

Multiple parton interactions

The multiple parton interactions (MPI) are several parton-parton interactions that occur during a single hadron-hadron collision. Most MPIs are relatively soft and they do not lead to easily identifiable jets [28].

Underlying event generation

The fragments of hadrons that do not participate in the hard collision construct the underlying event. Since these fragments are not colour neutral, they interact with the partons from the hard scattering [27].

Matching and merging

Double counting may occur when the same final state is generated by a parton level generator and a parton shower calculation. Since fixed order matrix elements describe separated hard partons well, while parton shower provides good modelling of soft and collinear partons, these two algorithms are combined for the best performance [28]. The merging scale is introduced to connect matrix element generation and parton shower. The partons that are produced above the merging scale are generated with a higher order matrix element, while the partons produced below this scale are generated by the parton shower [28]. The matching scheme is developed to avoid double counting and the dead regions where no events are generated [29].

The CKKW is a merging scheme where the second and higher parton shower emissions are corrected to match the tree level matrix element. Each jet multiplicity is simulated above the threshold cut with the corresponding matrix element and are dressed with parton shower [28].

In the MLM merging scheme the matrix element is generated and showered, the final state partons are combined into jets. The partons from the matrix element are required to match the final state jets, so the events are accepted only if all the jets

satisfy the matching criteria and the event contains no extra jets above the merging scale [30].

1.4.3 Monte Carlo generators

This section describes the relevant MC generators.

POWHEG

The POWHEG method produces positive weighted events, where the hardest radiation is generated using the exact next to leading order (NLO) matrix element. The POWHEG output can be interfaced with other parton shower generators [31].

Herwig++

Herwig++ is a general purpose MC event generator, which has an automatic generation of hard processes and decays with full spin correlations. It has built in matching of many hard processes at NLO with the POWHEG method [28]. The initial and final state QCD radiation is simulated with particular emphasis on radiation from heavy particles, and hadronisation is obtained using the cluster model [32]. Hard and soft multiple partonic interactions are modelled as well [28].

PYTHIA 8

PYTHIA 8 is a general purpose event generator. It is a standalone generator but can be interfaced with other generators [28]. PYTHIA 8 models hard and soft interactions, initial and final state parton showers. It has built in matching and merging methods to combine hard processes with parton shower. PYTHIA 8 simulates multiparton interactions, beam remnants, particle decays. It uses a string fragmentation approach to model hadronisation [33].

SHERPA

SHERPA is a general purpose event generator which simulates lepton-lepton, lepton-hadron, and hadron-hadron collisions. It has an automated generation of tree level matrix elements and uses advanced phase space integration methods. SHERPA is able to calculate cross-sections at NLO accuracy and has a built in multijet merging. Hadronisation is performed using the cluster model [28].

MadGraph5_aMC@NLO

MadGraph5_aMC@NLO is a framework that provides computations of tree level and NLO order cross-sections, their matching to parton shower simulations. It is used for generation of SM processes and BSM phenomenology [34].

1.4.4 Detector simulation

The detector simulation models the effects when the particles interact with the detector. It includes the geometry of the detector, particles' interaction with the material, the tracking of particles through materials and external magnetic fields and the response of the detector components [35].

GEANT4

GEANT4 is a toolkit to simulate the particles traversing the matter. It includes simulation of tracking, geometry, physics models and hits. It can be used for basic phenomena and geometries and for large scale complex simulations of LHC detectors [35].

Chapter 2

The ATLAS experiment

There is a long path from the production of protons to the objects that are interpreted and used in the analysis. First, protons are produced and accelerated in the Large Hadron Collider (LHC) accelerator complex, then collisions are induced at the interaction point in the ATLAS detector. The resulting particles leave traces in different subdetector systems. In order to be used in physics analysis, these traces are reconstructed and particles are identified. Only then a physics analysis can begin.

In this chapter first LHC is introduced in section 2.1. Collisions and detection are described in section 2.2. Section 2.3 explains the reconstruction and identification of physics objects.

2.1 LHC

The LHC is the largest and most powerful particle accelerator and collider in the world. LHC is built to provide data for the precise measurements of SM physics and the searches for new physics [36]. It is operated by the European Organisation for Nuclear Research (CERN) in Geneva, Switzerland.

LHC is built in a 26.7 km tunnel where there are two rings in which two high energy particle beams are accelerated in opposite directions. The beams are kept in a circular trajectory by twin bore superconducting magnets that consist of two

sets of coils. The beams collide at the four particle detectors in the accelerator ring: ATLAS, CMS, ALICE, and LHCb.

The protons are produced using hydrogen gas where electrons are removed by passing the gas through an electric field. Protons are provided via the injector chain of Linear accelerator 2 (Linac 2), Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS) [37]. At the Linac 2 protons are accelerated to 50 MeV energy. The PSB is built to increase the number of protons accepted by PS. In the PSB protons are accelerated to 1.4 GeV energy and then directed to PS. PS is a 628 m circumference accelerator which increases protons' energy to 25 GeV. Then protons enter the 7 km circumference SPS which increases protons' energy to 450 GeV. When protons reach this energy, they enter the LHC ring. LHC is designed to accelerate protons up to 7 TeV energy but it operated with up to 6.5 TeV before the shutdown in 2018. LHC accelerator complex is illustrated in Figure 2.1. At the LHC each proton beam has a nominal number of 2808 bunches with the spacing between them of 25 ns, where the spacing is set in the PS.

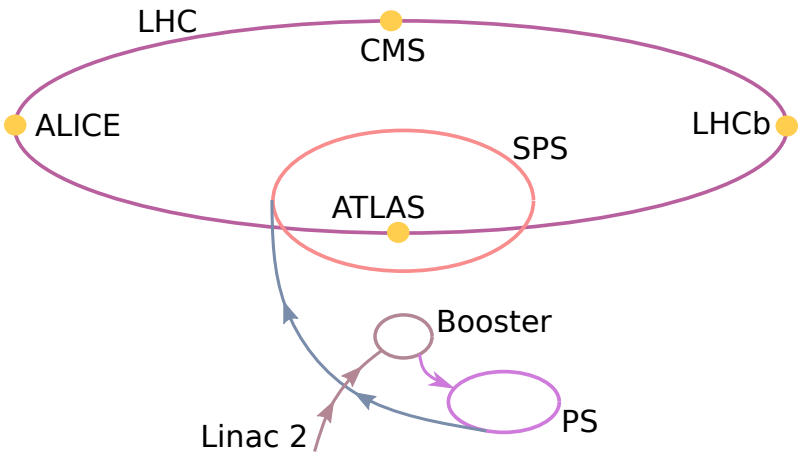


Figure 2.1: The LHC accelerator complex. Protons reach the LHC via the injector chain of Linear accelerator 2 (Linac 2), Proton Synchrotron Booster, Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). Figure is adapted from [38].

2.1.1 Luminosity

The rate of interactions, R, is expressed by:

$$R = L\sigma, \tag{2.1}$$

where L is the machine luminosity and σ is a cross-section of the process of interest. The luminosity depends on the beam parameters only and is given by the following formula:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\varepsilon_n \beta^*} F \tag{2.2}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, the revolution frequency is denoted by f_{rev} and the relativistic gamma factor by γ_r, ε_n is the normalised transverse beam emittance, β^* is the beta function at the collision point, and F is the geometric luminosity reduction factor [37].

LHC is designed to run with $10^{34} cm^{-2} s^{-1}$ luminosity. In 2016 luminosity exceeded the designed value and peaked at $1.38 \cdot 10^{34} cm^{-2} s^{-1}$; the peak luminosity per fill in 2016 is shown in Figure 2.2. The total delivered and recorded luminosity by ATLAS in 2015 and 2016 are presented in Figure 2.3.

LHC Runs

Since LHC activity started, there were two active phases when data were collected. The Run 1 is the first phase when LHC produced proton-proton collisions at a centre of mass energy of $\sqrt{s} = 7$ TeV and later at $\sqrt{s} = 8$ TeV. The Run 2 is the second phase when LHC produced collisions at a centre of mass energy of $\sqrt{s} = 13$ TeV.

2.2 ATLAS

The ATLAS (A Toroidal LHC Apparatus) detector is a general purpose particle detector installed at the LHC (Figure 2.4). The detector is about 44 m long and 25 m in diameter, and weights about 7000 t. ATLAS has a forward-backward symmetry

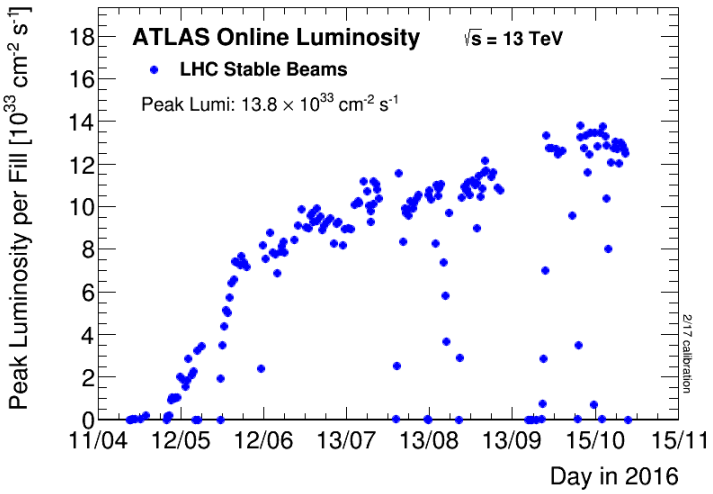
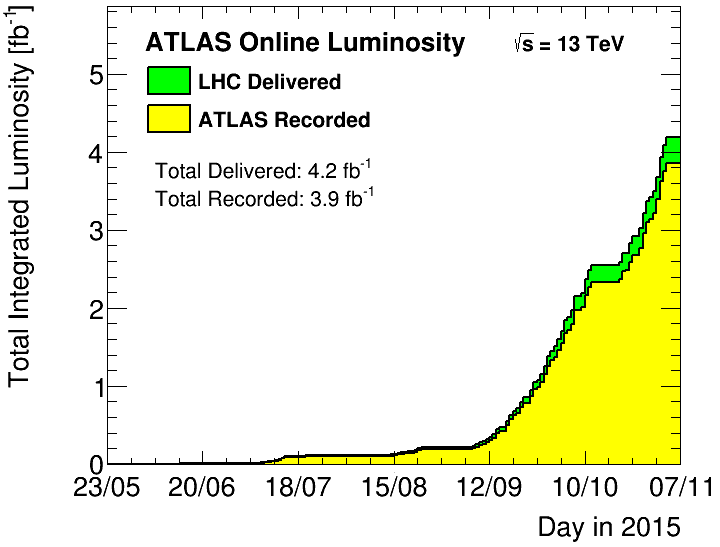


Figure 2.2: The peak luminosity per fill by ATLAS in 2016. Figure is taken from [39].

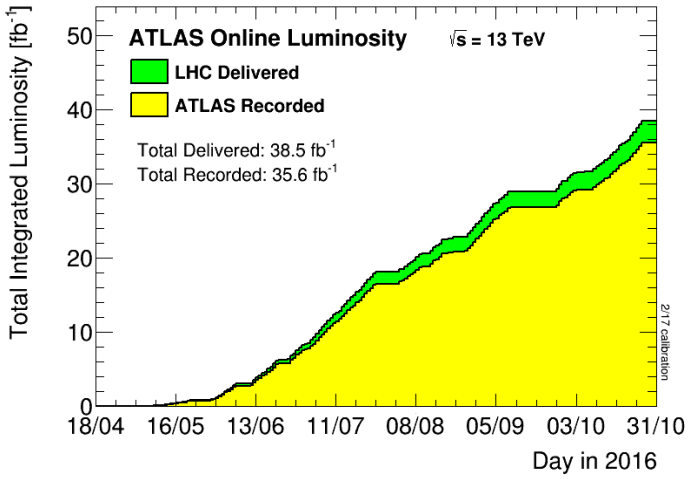
with respect to the interaction point. It consists of sub-detectors that are designed for specific particle detection. The detector consists of the inner detector that is responsible for charged particle tracking, the calorimeter system that is responsible for detection based on energy measurement, and the muon systems for muon identification. The inner detector is surrounded by a superconducting solenoid that produces 2 T magnetic field. The calorimeters are surrounded by three eight-fold superconducting toroid magnets with asimuthal symmetry: one in the barrel and one at each end-cap. [36]

Coordinate system in ATLAS

The origin of the coordinate system is the nominal interaction point. The ATLAS detector uses a right-handed coordinate system. The z-axis is defined in the direction of the beam, and $x-y$ plane is transverse to the beam direction. The x-axis points from the interaction point to the centre of the LHC ring, while the y-axis points upwards. The asimuthal angle ϕ is defined around the beam axis, while the polar angle θ is the angle measured from the beam axis. The pseudorapidity is defined as



(a)



(b)

Figure 2.3: The total delivered and recorded luminosity by ATLAS in (a) 2015 and (b) 2016. Figures are taken from [39].

$\eta = -\ln\tan(\theta/2)$, while for the massive objects (like jets) the rapidity measure is used $y = \frac{1}{2}\ln[(E+p_z)/(E-p_z)]$, where E and p_z are energy and momentum along the z-axis, respectively. The transverse quantities with subscript $_T$ are defined in the $x-y$ plane. The distance in the pseudorapidity-asimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ [36].

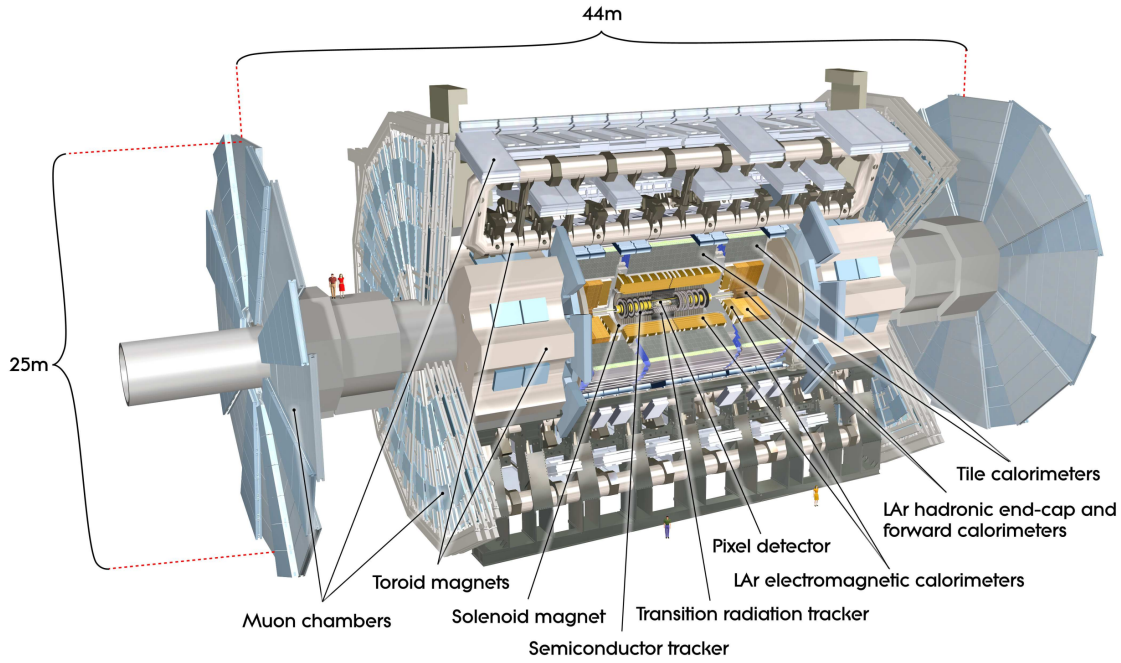


Figure 2.4: The cut-away view of the ATLAS detector. Figure is taken from [36].

2.2.1 The inner detector

The purpose of the inner detector (ID) is charged particle tracking (Figure 2.5). It is the closest sub-detector to the beam line. The detector is surrounded by a superconducting solenoid which provides 2 T axial magnetic field [40] and curves the trajectories of the charged particles. The inner detector has three subsystems: the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). These subsystems are described below.

The pixel detector

The pixel detector consists of four cylindrical layers in the central region and has three disks at each end-cap in the forward regions. The closest to the beam line lies the insertable B-layer (IBL) which was installed for the LHC Run 2.

The purpose of IBL is to improve impact parameter resolution and the vertex reconstruction [40], where the impact parameter is the closest distance from the

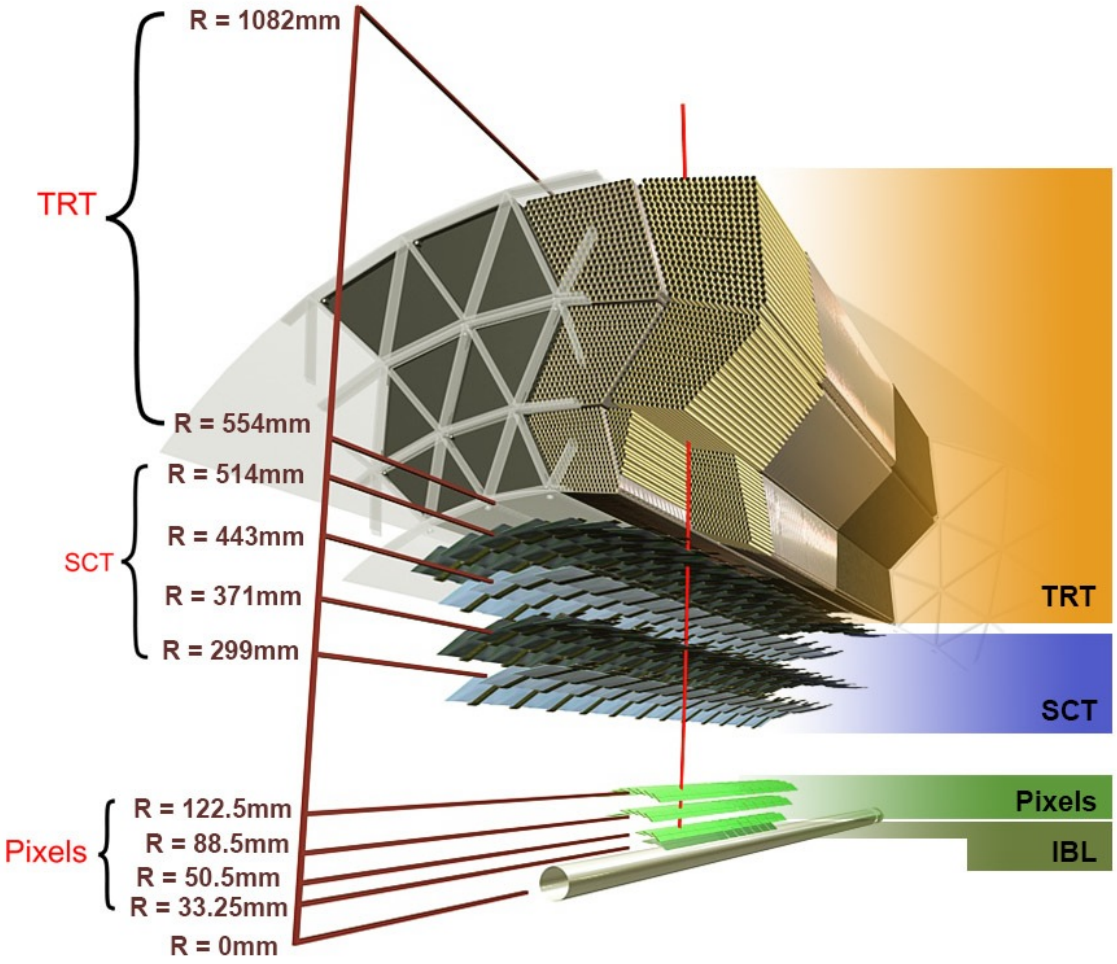


Figure 2.5: The cut-away view of the ATLAS inner detector. It shows the pixel detector, the SCT, and the TRT. Figure is taken from [36].

track to the vertex. It consists of 14 staves and covers the $|\eta| < 3.03$ region. The IBL contains about 12 million silicon pixels of size 50 μm $\times$ 250 μm in $r - \phi$ and z directions, respectively. The staves are placed between 29.0 mm and 42.5 mm distance from the beam line [41].

The remaining part of the pixel detector consists of three layers in the barrel and three disks in each of the two end-caps. The three barrel layers are located at 50.5 mm, 88.5 mm, and 122.5 mm distance from the beam line [41], and a typical pixel size is 50 μm $\times$ 400 μm in $r - \phi$ and z directions, respectively. The four disks

in the end-caps are designed such that most tracks hit at least three layers. The pixel detector contains about 80 million pixels in total. It is designed to tolerate high radiation doses, especially the IBL.

The operating principle of the pixel detector is based on semiconductor properties. The charged particle hits the silicon and creates electron-hole pairs. Under the electric field the generated electron-hole pairs move to the electrodes and the induced current is recorded.

The SCT

The silicon micro-strip semiconductor tracker (SCT) is located outside the pixel detector. The SCT consists of four cylindrical layers in the barrel region and nine end-cap disks. It covers the region of $|\eta| < 2.5$ [42]. SCT modules consist of two sensors of silicon micro-strips that are mounted on both sides of the module with a stereo angle of 40 mrad [43]. The strips are located longitudinally in the barrel region and radially in the end-cap region. There are nominally 8 hits per track (four space-points) [40]. The detector has about six million channels.

A space-point is determined from a combination of two strip measurements at a stereo angle [43]. The fine granularity of the SCT allows the reconstruction of particle trajectories, bent by the magnetic field. The SCT operates based on the same semiconductor mechanism as the pixel detector.

The transition radiation tracker

The transition radiation tracker's (TRT) purpose is to provide position information and particle identification, primarily the discrimination of charged particles with different mass, especially electrons and pions. The design is based on the fact that radiation depends on the Lorentz factor, which depends on particle mass.

The TRT covers the region of $|\eta| < 2.0$ [42] and is the outermost subsystem of the inner detector. It consists of 4 mm diameter straw drift tubes that are

located parallel to the beam axis in the barrel region and radially in the end-caps. The cathode is the inner part of the tube that is coated in aluminium, while the enclosed tungsten wire works as an anode. The tubes are filled with a mixture of gas composed of CO_2 (27%), O_2 (3%), and either Xe or Ar (70%). Xe is preferred due to its good X-ray absorption, while Ar is used in straws where large gas leaks are expected.

When a charged particle traverses a straw, ionised clusters are created, and the induced current is amplified and measured [44]. X-ray TR is emitted as a result of the change of the refractive index due to the inserted polypropylene material between the tubes, and is absorbed by Xe or Ar gas. The lighter particles produce more TR than heavier particles, so the TRT's threshold is set such that the electrons are more likely to exceed it than pions. For charged particles with transverse momentum $p_T > 0.5$ GeV, each track typically yields 35 hits [40].

2.2.2 Calorimetry

The ATLAS calorimeter system consists of two subsystems: the electromagnetic calorimeter which is closer to the beam line and the hadronic calorimeter which is in the outer part of the system (Figure 2.6). It covers the range of $|\eta| < 4.9$. The electromagnetic calorimeter is designed to measure energy deposits of electrons and photons, while the hadronic calorimeter is designed to measure hadronic jets.

The ATLAS calorimeters are sampling calorimeters, where energy deposition is sampled using alternating active and passive layers. Particle interaction with the material mainly in the passive layer creates particle showers; in the active layers a traversing charged particle ionises the active material and the current is recorded.

The additional purpose of the calorimeters is to limit high energy jets entering the muon chambers.

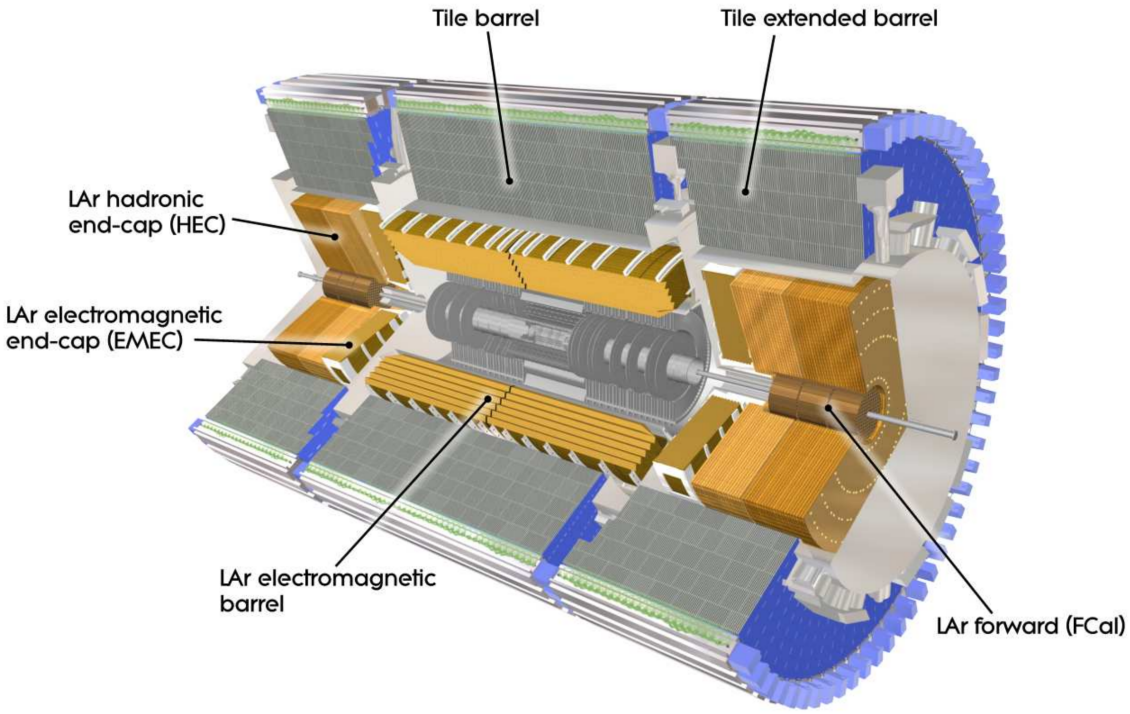


Figure 2.6: The cut-away view of the ATLAS calorimeter system. It shows the electromagnetic and hadronic calorimeters. Figure is taken from [36].

The electromagnetic calorimeter

The electromagnetic calorimeter (EM) consists of two half-barrels, which cover $-1.475 < \eta < 0$ and $0 < \eta < 1.475$ regions, and two end-caps covering $1.375 < |\eta| < 3.2$. There is a transition region filled with inactive material covering $1.37 < |\eta| < 1.52$.

The EM calorimeter has a presampler and three layers. The presampler is designed to account for sampling of the start of the particle shower. It is made of liquid argon (LAr) layer and covers $|\eta| < 1.52$ in the barrel area and $1.5 < |\eta| < 1.8$ in the end-caps region.

In the three layers the LAr is used as an active material and lead acts as a passive component. The calorimeter is designed in an accordion-like geometry symmetric around z axis, which ensures uniform ϕ coverage and fast extraction of the signal

[36].

The first layer of EM consists of strip cells, it has granularity of 0.0031×0.098 in $\eta - \phi$ plane, and its purpose is to discriminate photons and neutral pions. The second layer has granularity of 0.025×0.025 in $\eta - \phi$ plane. Its thickness is constructed to contain the greater part of the energy deposit. The third layer is used for the energy deposits of the tails of the showers [40]. It has a coarser resolution $\delta\eta \times \delta\phi = 0.05 \times 0.025$.

The hadronic calorimeter

The hadronic calorimeter consists of the tile central barrel, the tile extended barrel, the hadronic LAr end-caps, and a forward calorimeter.

The purpose of the tile calorimeters is to ensure identification and measurement of hadronic jets. The central barrel covers $|\eta| < 1.0$ and the extended barrels cover $0.8 < |\eta| < 1.7$ region. There the plastic scintillators act as an active material, while iron is used as a passive material. The light emitted by the scintillators is measured with photomultiplier tubes.

The LAr hadronic end-caps consist of two wheels per cap and cover $1.5 < |\eta| < 3.2$ region. In this subsystem copper is used as a passive material.

The forward calorimeter consists of three wheels on each side and covers $3.1 < |\eta| < 4.9$ region. The first layer uses copper as an absorbent material and the other two layers use tungsten. LAr is used as an active material.

2.2.3 Muon spectrometer

The muon spectrometer (MS) is designed to detect muons and measure their momentum (Figure 2.8). It is located in the outermost part of the ATLAS detector. The subdetector achieves 3% relative resolution over a wide p_T range. The MS consists of a barrel part and two end-cap parts, covering $|\eta| < 1.05$ and $1.05 < |\eta| < 2.7$ regions, respectively. The system is equiped with three toroidal magnets for each of

the previously discussed parts [45]. The toroid magnet system generates 0.5 T field in the barrel and 1 T in the end-caps.

The muon system consists of four types of chambers: the monitored drift tube chambers (MDT), the cathode strip chambers (CSC), the resistive plate chambers (RPC), and the thin gap chambers (TGC).

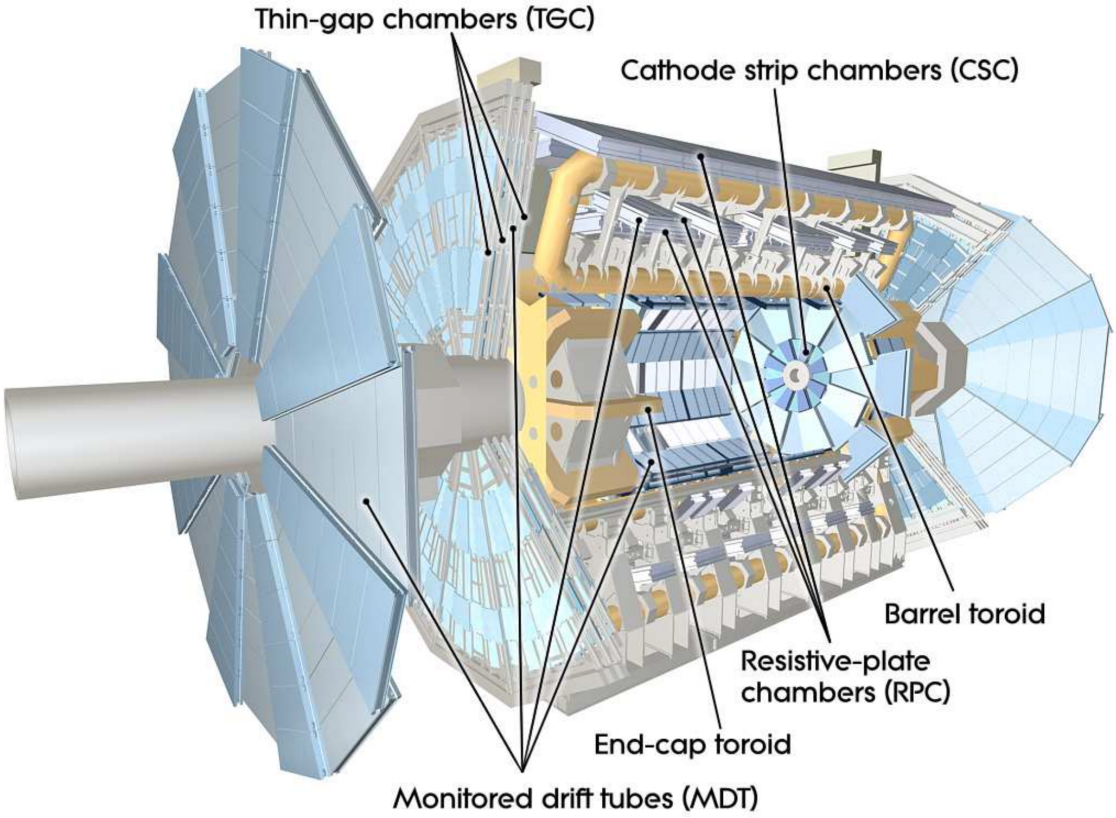


Figure 2.7: The cut-away view of the ATLAS muon system. Figure is taken from [36].

Monitored drift tube chambers

The monitored drift tube chambers (MDT) are designed to provide precise measurements in the $\eta - z$ plane, where the bending occurs. The MDTs cover $|\eta| < 2.7$ region and provide 6-8 η measurements per muon trajectory. The single hit resolution in the bending is a plane of 80 μm [45]. The MDTs consist of aluminium tubes

filled with Ar and CO_2 gas mixture.

Cathode strip chambers

The cathode strip chambers (CSC) are placed in the forward part of the detector. They also provide precise momentum measurements and are built to accommodate particle flux in $|\eta| > 2$ region which exceeds MDTs' threshold. The CSCs cover $2 < |\eta| < 2.7$ in the end-caps. The single hit resolution in the bending is a plane of 60 μm [45].

The cathodes are made of strips which are at right angles, while anodes are wires which are parallel to one of the strip directions. The electric signal arise from ionisation of Ar-CO_2 gas mixture.

Muon trigger chambers

Muon trigger chambers consist of the resistive plate chambers (RPC) and the thin gap chambers (TGC). The RPCs are located in the barrel region, while the TGCs are built in the forward region. The RPCs and TGCs cover $|\eta| < 1.05$ and $1.05 < |\eta| < 2.4$ regions, respectively. The purpose of this system is to provide the trigger with the fast readout speed: the RPCs serve the trigger with 10 ns readout speed. The trigger chambers also provide position measurements in $\eta - \phi$ plane with spacial resolution of 5-10 mm [45].

2.2.4 Trigger

The full information of collisions produced in ATLAS are both unfeasible to record due to technical restrictions and also many collisions are not of interest to physics analyses. The trigger system is build to aid the recording of interesting collisions. The trigger system consists of the two level triggers: Level 1 (L1) and Level 2 (L2), and an Event Filter (EF). The L2 and EF are referred to as the high level trigger

(HLT). The L1 trigger reduces readout rate from 40 MHz to 100 kHz. HLT reduces readout rate from about 100 kHz to 1 kHz.

The L1 trigger is implemented in hardware to increase the decision speed, which reaches 2.5 μs. This trigger takes information from reduced granularity calorimeter information and the muon system, including muon trigger chambers (section 2.2.3). It uses high p_T muons, electrons, photons, jets, and hadronically decaying τ-leptons.

The HLT is a software trigger. It also uses inner detector track information, matching of tracks and calorimeter energy deposits, and full event reconstructions.

2.3 Object reconstruction

The signals recorded by the subdetector systems are used to reconstruct physics objects which enter physics analyses. This section discusses physics object reconstruction.

2.3.1 Particle reconstruction in subdetectors

Charged particle reconstruction in the inner detector

Tracking detectors are used to reconstruct the trajectories and momenta of the charged particles [41]. Charged particle's reconstruction starts in the pixel detector and the SCT. Hits in the silicon detector layers are combined into clusters. Three dimensional space points are created from these clusters and are used to obtain track seeds. Then the pattern-recognition algorithm is applied and ambiguities are resolved.

Reconstruction in the electromagnetic calorimeter

First, the electromagnetic (EM) calorimeter is divided into cells in $\eta - \phi$ space of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, which corresponds to the granularity of the second EM calorimeter layer. The energy deposit is summed in each cell over presampler

and the three EM calorimeter layers. The candidates are selected by the sliding-window algorithm, where the window size is 3×5 cells. The selected candidates have the sum of transverse energy E_T higher than 2.5 GeV. If two candidates are found overlapping in the area of 5×9 cells, then the candidate with higher E_T is kept if it is at least 10% higher than the other candidate's E_T. Otherwise, the candidate with the central cell with highest E_T is kept. The reconstruction efficiency depends on $|\eta|$ and E_T [40].

2.3.2 Vertex reconstruction

The *primary vertices* are the inelastic proton-proton interaction positions, while the *secondary vertices* are defined as interactions resulting from in-flight decays, photon conversions, and hadronic interactions.

The primary vertices are reconstructed using the iterative vertex finding algorithm. A vertex must be formed of at least two charged particle tracks with p_T greater than 100 MeV. The primary vertices are categorised as the *hard-scatter* and the *pile-up* vertices based on the sum of the p_T^2 of the associated tracks. The highest sum defines the hard-scatter vertex while the remaining vertices are pile-up vertices. [41]. The secondary vertices are reconstructed using the inclusive secondary-vertex finding algorithm.

2.3.3 Electrons

Electron reconstruction

The electron reconstruction is based on three electron characterisation criteria: charged particle tracks in the ID, localised energy deposit clusters in the EM calorimeter, and the matching of the ID tracks and EM calorimeter clusters [40]. Charged particle tracks in the ID and clusters in the EM calorimeter are described in section 2.3.1.

When electron traverses material, it loses energy due to bremsstrahlung. The radiated photons may further convert into electron-positron pairs. Normally, these particles are emitted in a collimated fashion, therefore it is possible to match several tracks in the ID with the same EM cluster, associated with the primary electron. An electron candidate is considered to be a track with at least four hits in silicon layers and with no association with a vertex of photon conversion. Then, the extended EM calorimeter clusters are formed surrounding seed clusters. The best matching of the track and the seed cluster is used to calculate the energy of the electron candidate, which is computed using calibration and the extended clusters [40]. Electron reconstruction is illustrated in Figure 2.8.

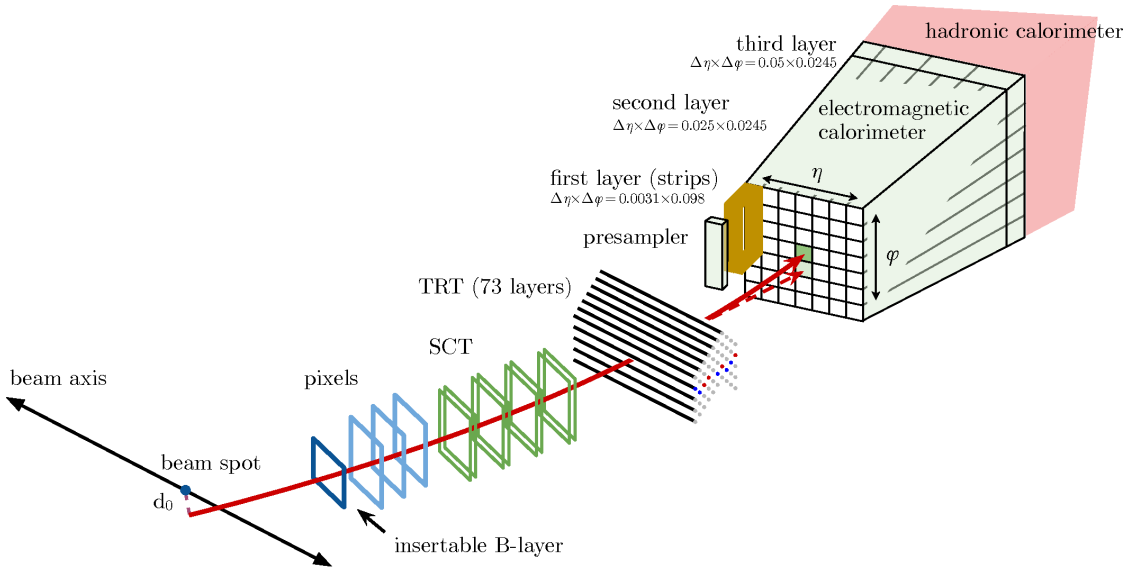


Figure 2.8: Electron reconstruction based on tracks in the inner detector and clusters in the electromagnetic calorimeter. Figure is taken from [40].

Electron identification, isolation, and reconstruction of charge

Electrons in the central detector region ($|\eta| < 2.47$) are identified using likelihood fit. There are three levels of electron identification: "Loose", "Medium", and "Tight" with corresponding identification efficiencies of 96%, 94%, and 88%, respectively.

The tighter the electron, the better background rejection. The discrimination is obtained using a likelihood fit with shower shapes, track properties and track and cluster matching as input parameters [40]. In the $H \to WW^*$ analysis in this thesis the electrons are required to have transverse energy (E_T) greater than 15 GeV. The electrons with $E_T < 25$ GeV are required to satisfy the "Tight" requirement, while the leptons satisfying $E_T > 25$ GeV should pass the "Medium" identification requirement. The pseudorapidity η is required to satisfy $|\eta| < 2.47$, excluding the region of $1.37 < |\eta| < 1.52$.

Electrons that originate from the hard-scattering vertex or are decay products of heavy bosons (Higgs, W, Z) have a characteristic signature of being isolated from other particles. In contrast, the electrons that are the products of semileptonic decays of heavy quarks or proton conversion tend to have other particle traces in the close vicinity. Isolation serves as a measure to identify the electrons interesting to physics analyses. The "Tight" and "Medium" electrons should satisfy the "Fix" and "Gradient" isolation requirements, respectively, which are described in [40].

The sign of the electric charge of the electrons is determined by the curvature of the tracks in the ID.

2.3.4 Muons

Muon reconstruction in the muon spectrometer

Muon reconstruction is based on the hit patterns in the muon spectrometer (MS) chambers. Hits in the monitored drift tube chambers are measured in the bending plane, while hits in the resistive plate chambers and the thin gap chambers are measured orthogonal to the bending plane. The segments in the monitored drift tube chambers are reconstructed using a straight line fit, while a combinatorial technique is applied in the cathode strip chambers case. Muon tracks are built by using the segments in the middle as seeds for extrapolation. Tracks are built from

at least two matching segments, except for the barrel-end-cap region where one segment is required. Overlap removal is applied if needed and a global χ^2 fit is used to select hits associated to each track [45].

Muon reconstruction

Muon reconstruction is performed independently in the ID (section 2.3.1) and the MS.

There are four types of muons based on how the ID and the MS reconstructions are combined. The four types are the combined (CB) muons, the segment-tagged (ST) muons, the calorimeter-tagged (CT) muons, and the extrapolated (ME) muons. The combined muons are defined by combining independently reconstructed ID and MS tracks. A global refit is performed where MS hits may be added or removed in favour of the fit. The segment-tagged muons' definition serves low p_T muons or muons in the low MS acceptance region. They are defined as ID tracks extrapolated to the MS with at least one track segment in the monitored drift tube chambers. The calorimeter-tagged muons are defined as ID tracks matched with an energy deposit in the calorimeter compatible with minimum ionising particle. The extrapolated muons are constructed using an MS track only and are compatible with the interaction point as the origin [45].

The relevant muons to this thesis are the combined muons.

Muon identification

The muon identification is based on applying conditions, which select muons with high efficiency and/or robust momentum measurement and suppressing backgrounds, especially pions and kaons. The quality requirements for a robust momentum measurement are the following: at least one hit in the pixel detector, at least five hits in the SCT, less than three track parts without hits in the pixel or the SCT subdetector, and at least 10% of TRT hits assigned to the track if $0.1 < |\eta| < 1.9$.

Muon identification has four categories: "Loose", "Medium", "Tight", and "High-p_T". "Medium" muons are defined such that they minimise the systematic uncertainties associated with muon calibration and reconstruction. "Loose" muons have good quality muon tracks and maximise the reconstruction efficiency. The combined and extrapolated muons in the "Medium" category are a subset of "Loose" muons. "Tight" category maximises the purity of muons. It is a subset of the combined "Medium" muons, which have at least two hits in the MS. "High-p_T" muons have transverse momentum above 100 GeV and maximised momentum resolution for tracks. They consist of the "Medium" combined muons with at least three hits in the MS [45].

In the $H \to WW^*$ analysis in this thesis the muons required to pass the "Tight" identification requirement and satisfy the p_T >15 GeV and $|\eta| < 2.5$ requirements.

Muon isolation

The muons which are interesting to physics analyses, i.e. originating from heavy particles like Higgs, W, Z bosons, tend to be isolated from other particles. In contrast, semileptonic decays produce muons that are part of jets. *Muon isolation* is a measure that is used to distinguish muons from different origins [45]. The track-based and the calorimeter-based isolation criteria are $p_T^{varcone30}/p_t < 0.06$ and $E_T^{topocone20}/p_T < 0.09$, respectively, where the $p_T^{varcone30}$ and $E_T^{topocone20}$ are isolation variables defined in [45].

2.3.5 Jets

Jets are objects widely used in physics analyses. They are made of collimated showers of hadrons. Jets are reconstructed using topological clusters (topo-clusters) [46]. The topo-clusters are formed to suppress the noise from the calorimeters. The noise is estimated from the measurements of calorimeter electronic noise and simulated pile-up noise. The algorithm starts by identification of the seed cell where

the signal (S) to noise (N) ratio is $S/N > 4$. Secondly, the neighbouring cells with $S/N > 2$ are added. Finally, one iteration of cells without S/N criterion are added. In 2015 the topo-cluster reconstruction algorithm was improved by forbidding topo-cluster seeds from the presampler layer. This avoids jet reconstruction from energy depositions that do not penetrate the calorimeters [46].

The jet reconstruction algorithm uses the following parameters: d_{ij} which is the distance between two objects i and j, and d_{iB} which is the distance between an object i and the beam.

$$d_{ij} = min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{R_{ij}^2}{R^2} \tag{2.3}$$

$$d_{iB} = p_{T,i}^{2n} \tag{2.4}$$

where $p_{T,i}$ is the transverse momentum of the object i, R is a fixed input parameter to the algorithm. $R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ where y_i and ϕ_i are the rapidity and the asimuthal angle of the object i. n is an exponent which depends on the algorithm choice: $n = 1$ corresponds to k_T algorithm [47], $n = -1$ corresponds to anti-k_T algorithm [48], and $n = 0$ is Cambridge-Aachen jet finding algorithm [49].

In this thesis jets are reconstructed using the anti-k_T algorithm with $R = 0.4$ distance parameter. This algorithm is favoured because it is infrared and collinear safe, i.e. is resilient with respect to soft and collinear radiation, and it prioritises to cluster hard particles first [50].

The pile-up jet contamination in jets is reduced using a jet vertex tagger (JVT) discriminant, which is constructed from two variables: the ratio of the sum of the transverse momentum of the tracks associated with the jet and originating from the primary vertex, and the sum of the transverse momentum of all tracks associated with the jet and corrected for the total number of primary vertices; and the ratio of the sum of the transverse momentum of the tracks associated with the primary vertex and the transverse momentum of the jet [51, 52]. The forward JVT (fJVT)

is a procedure to remove forward pile-up jets using multivariate techniques [53].

In the $H \to WW^*$ analysis in this thesis the jets are required to satisfy the $p_T > 30$ GeV and $|\eta| < 4.5$ requirements. The JVT variable is required to be larger than 0.59 for jets with $p_T < 60$ GeV and $|\eta| < 2.4$. For jets satisfying $|\eta| > 2.5$ the fJVT requirement is applied.

b-jets

b-jets (jets containing b-hadrons) with $p_T > 20$ GeV and $|\eta| < 2.5$ are identified based on a multivariate technique where track impact parameters and information from secondary vertices are used as inputs. The b-jet tagging efficiency is estimated using $t\bar{t}$ simulation and is about 85% in this analysis [2]. The rejection rate of the light quark or gluon jets is 34 and the rejection rate for the jets containing c-hadrons is 3.1.

Jet energy scale calibration

Jet energy scale (JES) calibration starts by recalculating a jet's four-momentum such that it points to the hard-scatter vertex instead of the centre of the detector, while energy is kept constant. Then, the pile-up correction is applied, which removes the excess energy from pile-up. The calorimeter, MS, and tracks' information is used to obtain further improvements. Finally, calibration is performed using a p_T balance to other well measured objects, e.g. photons, Z bosons, and other calibrated jets.

2.3.6 Missing transverse energy

The missing transverse energy (E_T^{miss}) is an important proxy to the particles invisible to the ATLAS detector. These particles may be SM neutrinos or undetectable BSM particles. The E_T^{miss} is defined as the negative sum of the observed objects. The reconstruction of the E_T^{miss} consists of two contributions: the hard event signals and the soft event signals.

The hard events consist of fully reconstructed and calibrated particles and jets. The particles are electrons, photons, τ-leptons, and muons. Electrons and τ-leptons are reconstructed using ID tracks and EM calorimeter signals. Photons and jets are obtained from calorimeter signals. Muons are reconstructed from ID and MS tracks. The soft events comprise particles reconstructed from charged particles' tracks associated with the hard scatter vertex excluding the hard events defined above [54].

The object selection is prioritised in this order: electrons, photons, hadronically decaying τ-leptons, and jets. The muon reconstruction overlaps insignificantly with the other objects [54]. The $\mathbf{E}_T^{\text{miss}}$ is defined as in [54]:

$$\mathbf{E}_T^{\text{miss}} = -\sum_{\substack{\text{selected}\\ \text{electrons}}} \mathbf{p}_T^{e} - \sum_{\substack{\text{accepted}\\ \text{photons}}} \mathbf{p}_T^{\gamma} - \sum_{\substack{\text{accepted}\\ \tau-\text{leptons}}} \mathbf{p}_T^{\tau_{\text{had}}} - \sum_{\substack{\text{selected}\\ \text{muons}}} \mathbf{p}_T^{\mu} - \sum_{\substack{\text{accepted}\\ \text{jets}}} \mathbf{p}_T^{\text{jet}} - \sum_{\substack{\text{unused}\\ \text{tracks}}} \mathbf{p}_T^{\text{track}} \tag{2.5}$$

Missing transverse energy from tracks

An alternative variable, the missing transverse energy from tracks ($E_T^{\text{miss,Track}}$) is reconstructed from ID tracks associated with the primary vertex only. Therefore, jets and reconstructed particles are ignored since the calorimeters and the MS are not included. As a result, $E_T^{\text{miss,Track}}$ is almost insensitive to pile-up.

2.3.7 Overlap removal

In ATLAS, reconstruction of electrons, muons, and jets is done independently, therefore an overlap of objects may occur. A distance parameter ΔR is used to remove overlapping objects in the following manner: if a reconstructed muon shares an ID track with a reconstructed electron, the electron is removed. The reconstructed jets are removed if they overlap with electrons or muons in a cone of radius $\Delta R = 0.2$. Electrons and muons are removed if they are with a jet within a cone of

$\Delta R = \min(0.4, 0.04 + 10 \text{ GeV}/p_T)$, where p_T is a transverse momentum of a lepton [2].

2.3.8 Pile-up

Pile-up is more than one recorded collisions per bunch crossing. There are two sources of pile-up collisions: the ones happening at the same time as the main collision, and the ones that are due to a long calorimetry read-out time since it is larger than the bunch spacing of 25ns.

Chapter 3

Extraction of the $H \to WW^* \to e\nu\mu\nu$ signal

This chapter presents the first part of the Simplified Template Cross Sections (STXS) measurement through the decay $H \to WW^* \to e\nu\mu\nu$ and the gluon-gluon fusion (ggF) and vector boson fusion (VBF) production processes. The results are obtained from a data sample corresponding to an integrated luminosity of 36.1 fb^{-1} from $\sqrt{s} = 13$ TeV proton-proton collisions recorded by the ATLAS detector at the LHC in 2015 and 2016. This Higgs boson decay channel has been studied by the ATLAS [2, 55] and CMS [56, 57] experiments. This chapter presents the extension of the inclusive cross-section measurement performed in [2] to the STXS measurement (section 1.3). In the case of the ggF production mode the cross-sections are measured as a function of transverse momentum and number of jets. The STXS framework is preferred since it allows a relatively straightforward combination with other Higgs decay channels and the fact that the cross-sections are measured at particle level makes the BSM interpretations easier. The analysis strategy is to start from the $H \to WW^*$ inclusive cross-section measurement presented in [2], change the one jet final state analysis and include the analysis of the two or more jets phase space. The new analysis design presented in this thesis includes the definition and optimisation of signal regions and the introduction of a new discriminant based on multivariate

analysis methods. Dedicated control regions are used for the evaluation of the background processes.

The chapter presents the extraction of the $H \to WW^* \to e\nu\mu\nu$ signal. The modelling of the signal and background processes is described in section 3.1 and the Monte Carlo samples used for the simulation of these processes are described in section 3.2. The event selection for the signal regions and splitting the signal regions into categories are presented in section 3.3. Finally, the development of the multivariate discriminants is described in section 3.4.

3.1 Characteristics of signal and background processes

The final state particles of the $H \to WW^* \to e\nu\mu\nu$ signal can be produced by other processes, denoted as background processes. This section identifies the signal and background processes and their characteristic features, where the latter aids the background reduction and signal extraction. The signal and background processes are discussed in sections 3.1.1 and 3.1.2, respectively.

3.1.1 Signal characteristics

The $H \to WW^*$ decay channel has the second largest branching fraction of the Higgs boson decays [2] and in the LHC Run 1 the cross-section of this channel was measured with the highest precision of all decay channel measurements [12]. The decay signature is the cleanest when W bosons decay leptonically $W \to \ell\nu$, where ℓ is either an electron, a muon, or a leptonically decaying τ lepton. In contrast, for hadronic W boson decays, there is a very large irreducible background from QCD jets. However, the presence of the neutrinos in the final state is a disadvantage which prevents the full reconstruction of the Higgs boson mass. The different lepton flavour $H \to WW^* \to e\nu\mu\nu$ signature is chosen due to the reduction of the Drell-

Yan background process in comparison to the same lepton flavour e^+e^- and $\mu^+\mu^-$ final states. As a result, the analysis selects the events whose final states contain neutrinos, one muon and one electron with opposite electric charge.

A characteristic aspect of the signal process is related to the spin values of the Higgs boson and its decay products. Since the Higgs boson is spin-0, it decays to two W bosons with opposite spins (Figure 3.1). The W boson is spin-1 and therefore decays into two leptons with aligned spins. Due to the spin correlations, the final state charged leptons are emitted at a small angle with respect to each other. As a result it leads to a low invariant mass $m_{\ell\ell}$ and a small asimuthal angle $\Delta\phi_{\ell\ell}$ of the two charged leptons.

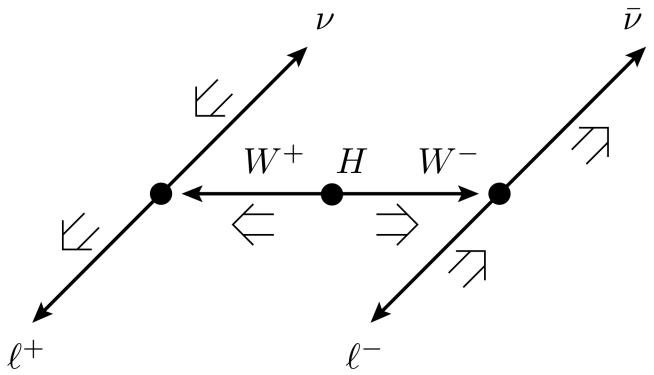


Figure 3.1: The illustration of the leptonic $H \to WW^*$ decay. The narrow and wide arrows indicate the direction of motion and the spin projections of the particles, respectively. The Higgs boson is a spin-0 particle which in its rest frame decays into two spin-1 W bosons, resulting in their spins pointing in opposite directions. In the W boson rest frame its decay products are spin-$\frac{1}{2}$ leptons in the configuration of aligned spins. The figure is taken from [55].

The ggF and VBF production processes are the dominant Higgs boson production modes considered in this analysis (section 1.1.2). The characteristic aspect of VBF production is two quarks scattered at a small angle leading to two highly energetic, well separated forward jets with a large invariant mass. Since the quarks radiate W or Z bosons (Figure 3.2), there is no colour exchange and so there is

no additional QCD activity between the two jets. Unlike the VBF case, the ggF production process has colour exchange, which results in additional QCD activity.

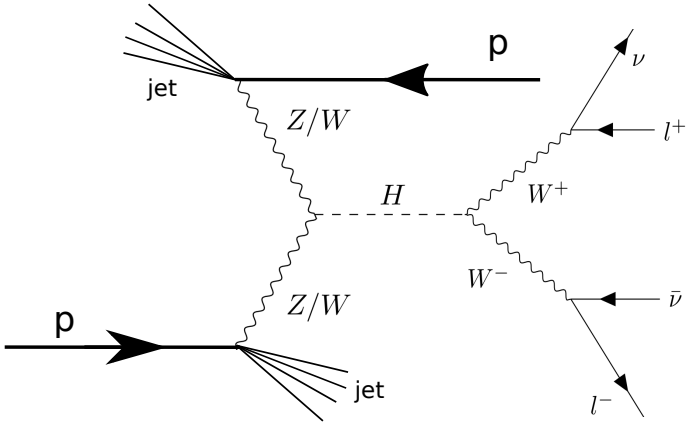


Figure 3.2: VBF production diagram for the $H \to WW^* \to e\nu\mu\nu$ decay.

In this analysis the VH production is fixed to the SM prediction, while the $t\bar{t}$H and $b\bar{b}$H production processes have a negligible contribution.

3.1.2 Background characteristics

The background processes are categorised in two types based on the final state particles: the reducible and irreducible background processes. The irreducible background processes have identical final state particles as the signal process, while in the case of a reducible background process a misidentification of at least one final state particle yields the process to have final state particles mimicking the signal final state. The main background processes in this analysis are WW bosons, other dibosons, top quarks, misidentified leptons, and the Drell-Yan (DY) process. The signal extraction is based on the background reduction by exploiting the properties of these backgrounds.

Nonresonant WW production

The nonresonant WW production is an irreducible background process, which consists of quark-quark ($q\bar{q} \to WW$) and gluon-gluon ($gg \to WW$) production mecha-

nisms, where W bosons decay leptonically. The dominant production is through the quark-antiquark t-channel scattering and s-channel annihilation including a triple-gauge boson vertex, illustrated in Figure 3.3 (a) and (b), respectively. A less significant contribution is via the gluon fusion through a quark loop, illustrated in Figure 3.3 (c). A characteristic feature of the WW background process is two well separated charged leptons in the final state and the background suppression is targeted based on the Higgs boson decay properties (section 3.1.1).

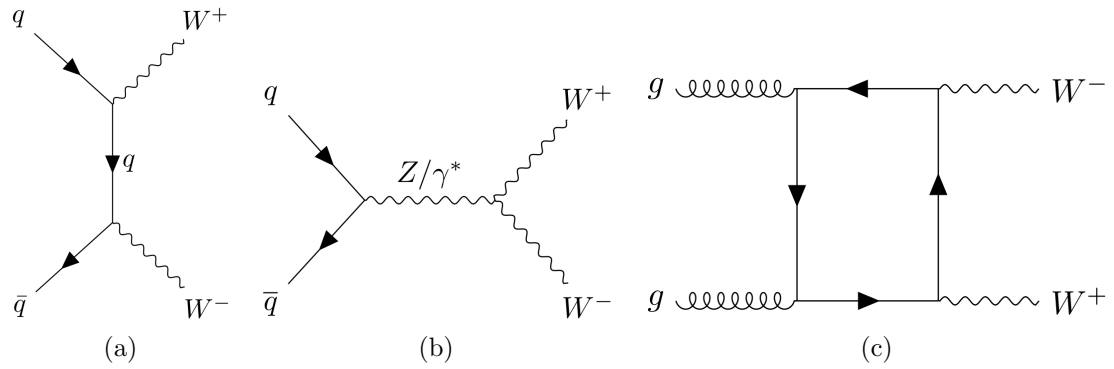
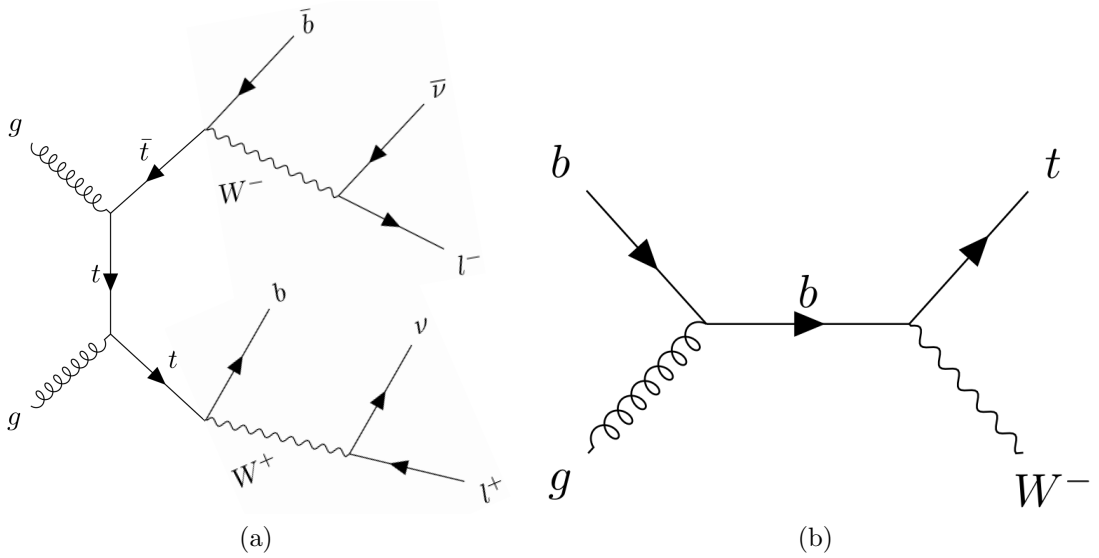


Figure 3.3: The SM tree level Feynman diagrams for WW background processes: (a) through the quark-antiquark t-channel scattering, (b) through the quark-antiquark s-channel annihilation including a triple-gauge boson vertex, and (c) gluon fusion through a quark loop.

Top quark production

The top quark production process is an irreducible background, which is characterised by top quark's decay almost exclusively into a W boson and a b quark. Only leptonic W boson decays are considered. The main contributing processes are top quark pair production ($t\bar{t}$) and a production of a top quark in association with a W boson (Wt), illustrated in Figure 3.4. Consequently, there are either one or two associated b-jets (originating from a b quark, section 2.3.5) in Wt and $t\bar{t}$ cases, respectively. This background process is characterised by high momentum jets and can be reduced by requiring no b-jets in the event selection, although the background

contribution still remains large due to inefficiencies in the b-jet identification algorithm. Due to associated jets, the top pair background process is more prominent at categories with higher jet multiplicity.



(a) (b)

Figure 3.4: The SM tree level Feynman diagrams for top quark background processes in (a) a top quark pair and (b) a top quark in association with a W boson cases.

Drell-Yan process

Drell-Yan (DY, also denoted as Z/γ^*) is a process when a virtual photon or a Z boson decays into a pair of charged leptons. Since the signal process has different flavour charged leptons in the final state, the dominant DY background contribution is from the fully leptonic decays of τ lepton pairs. τ leptons decay into W bosons and neutrinos, where W bosons further decay leptonically. This process is classified as an irreducible background process and an example of it is illustrated in Figure 3.5.

Other VV processes

The diboson background processes, which originate from two vector bosons other than WW, are denoted as VV. These processes include $W\gamma$, $W\gamma^*$, WZ, ZZ, $Z\gamma$, and $Z\gamma^*$, where the dominant background processes are $W\gamma$, $W\gamma^*$, and WZ.

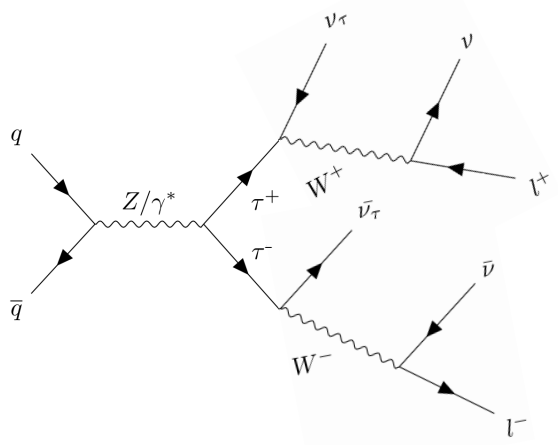


Figure 3.5: The SM tree level Feynman diagram for Drell-Yan background process, where a Z boson decays into a leptonically decaying pair of τ leptons.

The $W\gamma$ background process mimics the signal when the W boson decays leptonically and the photon converts into an e^+e^- or $\mu^+\mu^-$ pair in the detector material, where only one lepton satisfies electron selection criteria (section 2.3.3) (e.g. if the pair is very asymmetric in transverse momentum p_T). The missing transverse momentum originates from the leptonic W boson decay. The $W\gamma^*$ background process consists of a W boson decaying leptonically and a virtual photon producing to an e^+e^- or $\mu^+\mu^-$ pair, where only one lepton passes the lepton selection criteria. The WZ background process originates from the W boson decaying leptonically and the Z boson decaying into charged leptons where one lepton is not reconstructed. The $Z\gamma$ and $Z\gamma^*$ background processes occur when the Z boson decays into a pair of charged leptons where one of them is not reconstructed, and the photon decays into a different flavour pair of leptons where one is also not identified. These are classified as reducible background processes. The ZZ background process consists of one Z boson decaying into a pair of leptonically decaying τ leptons, and the other Z boson decaying into a pair of neutrinos.

Background process due to misidentified leptons

The dominant process of the misidentified lepton background occurs when a W boson is produced in association with one or more jets where one of the jets is misidentified as a lepton. The missing transverse momentum originates from the leptonic W boson decay. A less significant contribution arises from a multijet process when two jets are misidentified as leptons and the missing transverse momentum is recorded. The background process due to misidentified leptons is classified as a reducible background. The dominant processes are illustrated in Figure 3.6.

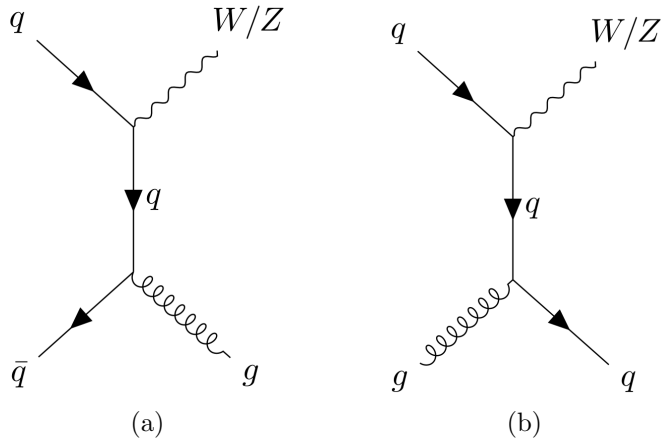


Figure 3.6: The SM tree level Feynman diagrams for W+jets background process, when the W boson is produced in association with (a) a gluon, and (b) a quark.

$H \to \tau\tau$ **decay**

The $H \to \tau\tau$ decay is considered as an irreducible background in this analysis. The signal process is mimicked when τ leptons decay into W bosons and neutrinos where W bosons decay leptonically. The characteristic aspect of the decay of the τ leptons is that the neutrinos are typically collinear with the charged leptons.

3.2 Monte Carlo samples

The Monte Carlo technique (section 1.4) is the main method used in this analysis for the signal and background modelling. Dedicated samples are generated to evaluate all but W+jets and multijet processes, which are estimated using a data driven techique (discussed in section 3.3.4). Sections 3.2.1 and 3.2.2 describe the production of signal and background samples, respectively.

The simulation tools used in the generation of signal and background processes, the PDF sets, and the prediction order in QCD are summarised in Table 3.1.

Table 3.1: Overview of simulation tools used for the generation of signal and background matrix element and parton level events, their PDF sets, and prediction order in QCD. The table is adapted from [2].

Process	Matrix element	PDF set	UEPS model	Prediction order for total cross-section
ggF H	POWHEG-Box v2 NNLOPS [59, 60, 61]	PDF4LHC15 NNLO [62]	PYTHIA 8 [58]	N^3LO QCD + NLO EW [63, 64, 65, 66, 67]
VBF H	POWHEG-Box v2 [59, 60, 61, 68]	PDF4LHC15 NLO	PYTHIA 8	NNLO QCD + NLO EW [63, 69, 70, 71]
VH	POWHEG-Box v2 [72]	PDF4LHC15 NLO	PYTHIA 8	NNLO QCD + NLO EW [73, 74, 75]
$qq \to WW$	SHERPA 2.2.2 [76, 77]	NNPDF3.0NNLO [79]	SHERPA 2.2.2 [80, 81]	NLO [78]
$gg \to WW$	SHERPA 2.1.1 [78, 82]	CT10 [83]	SHERPA 2.1	NLO [84]
$WZ/V\gamma^*/ZZ$	SHERPA 2.1	CT10	SHERPA 2.1	NLO [78]
$V\gamma$	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	NLO [78]
$t\bar{t}$	POWHEG-Box v2 [85]	NNPDF3.0NLO	PYTHIA 8	NNLO+NNLL [86]
Wt	POWHEG-Box v1 [87]	CT10 [83]	PYTHIA 6.428 [88]	NLO [87]
Z/γ^*	SHERPA 2.2.1	NNPDF3.0NNLO	SHERPA 2.2.1	NNLO [89, 90]

3.2.1 Signal modelling

The Higgs boson ggF production mode is generated at next-to-next-to-leading order (NNLO) QCD accuracy with POWHEG-Box v2 NNLOPS program [59, 60, 61] with

the PDF4LHC15 NNLO PDF set [62]. The parton shower, hadronisation, and the underlying event are generated using PYTHIA 8 [58] with AZNLO tune [91]. The Higgs boson rapidity spectrum in Hj-MiNLO [60] is reweighted to HNNLO [92] to achieve NNLO accuracy for inclusive $gg \to H$ observables [2]. Zero jet events are generated at NNLO accuracy in QCD, one jet events are generated at next to leading order (NLO) accuracy, and two jet events are generated at leading order (LO) accuracy. Samples are normalised to the cross-section equal to 48.58 pb obtained from the next-to-next-to-next-to-leading order (N3LO) prediction, which includes electroweak (EW) corrections [63, 64, 65, 66, 67].

The Higgs boson VBF production mode is generated at NLO accuracy in QCD with POWHEG-Box v2 [59, 60, 61, 68] with the PDF4LHC15 NLO PDF set [62]. PYTHIA 8 [58] is used to generate parton level events. The normalisation of the cross-section equal to 3.782 pb is obtained by taking into account full NLO QCD, full EW, and approximate NNLO QCD corrections [63, 69, 70, 71].

The Higgs boson VH production mode is generated with POWHEG-Box v2 [72] program with the PDF4LHC15 NLO PDF set. The parton level events are generated using PYTHIA 8 program. The samples are normalised to the cross-section obtained from NNLO QCD prediction, which includes NLO EW corrections [73, 74, 75]. The WH and ZH samples are normalised to 1.373 pb and 0.8839 pb cross-sections, respectively.

The Higgs boson mass is set to 125 GeV in all signal samples. The SM branching fraction of $H \to WW^*$ is calculated with HDecay v6.50 [93, 94] and is equal to 0.214 [14].

3.2.2 Modelling of background processes

WW background samples are generated separately for gluon-gluon ($gg \to WW$) and quark-quark ($qq \to WW$) production mechanisms. The $qq \to WW$ is modelled

with SHERPA 2.2.2. [76, 77] with NNPDF3.0 NNLO PDF set [79]. SHERPA 2.2.2. [80, 81] is also used to generate parton shower, hadronisation, and underlying event. Zero and one jet matrix elements are generated at NLO, while two and three jet matrix elements are generated at LO precision. $gg \to WW$ sample is simulated with SHERPA 2.1.1. with zero or one jet [78, 82] using CT10 PDF set [83]. The sample is normalised to the cross-section calculated at NLO in QCD [84].

The top quark background process consists mainly of top pair ($t\bar{t}$) and single top processes. The top quark pair samples are generated with POWHEG-Box v2 [85] using NNPDF3.0NLO PDF set. The parton shower is modelled with PYTHIA 8 with NNPDF2.3 PDF set [95] and A14 tune [96]. The $t\bar{t}$ samples are normalised using cross-sections derived at NNLO+NNLL (next-to-next-to-leading logarithm) [86]. The single top background consists mainly of Wt events. Samples are generated with POWHEG-Box v1 [87] with CT10 PDF set [83]. The parton level events are modelled with PYTHIA 6.428 [88] with Perugia2012 tune [97]. Bottom and charm hadron decays are modelled with EvtGen 1.2.0 [98].

Gluon-gluon initiated diboson other VV samples are generated with SHERPA 2.1.1. with zero or one jets. The $q\bar{q}/g \to Z^*Z^* \to \ell\nu\ell\nu$ process is generated with POWHEG-Box v2. The WZ samples are generated with SHERPA 2.1. with CT10 PDF set. The zero and one jet samples are generated at NLO, while the two and three jets samples are generated at LO. The EW processes of WW + 2 jets ($\ell\nu\ell\nu$), WZ + 2 jets ($\ell\nu\ell\ell$), ZZ + 2 jets ($\ell\ell\ell\ell$), and $q\bar{q} \to q\bar{q}$ with $Z \to \tau\tau$ are modelled with SHERPA 2.1.1. at LO. The $Z\gamma$ and $W\gamma$ processes are generated with SHERPA 2.2.2. The zero and one jet final states are modelled at NLO in QCD.

The DY samples are generated with SHERPA 2.2.1. with the NNPDF3.0NNLO PDF set. The normalisation of the samples is calculated at NNLO.

3.3 Event selection and categorisation

The purpose of the event selection is to reduce the background contributions without significantly affecting the signal acceptance. There is a standard selection developed for the former $H \to WW^*$ analyses [2, 55] based on the typical signal and background characteristics and detector geometry. These selection criteria are incorporated from the previous research, while a new selection is developed to aid the specifics of this Simplified Template Cross Sections' (STXS) analysis.

The basic selection applied to all events (denoted as preselection) is presented in section 3.3.1. The resulting set of events is further divided into categories based on jet multiplicity to target the characteristic signal and background aspects individually (section 3.3.2). The procedure of splitting the signal regions into categories enriched in STXS signal processes is presented in section 3.3.3. Finally, the estimation of the background processes is explained in section 3.3.4.

3.3.1 Preselection

The selection discussed in this section is a basic selection applied to all events entering the analysis. The selected event must have strictly two leptons, which are of different flavour and opposite sign. The higher p_T (leading) lepton ($p_T^{\ell_0}$) has to pass $p_T^{\ell_0} > 22$ GeV criterion and the lower p_T (subleading) lepton ($p_T^{\ell_1}$) has to have $p_T^{\ell_1} > 15$ GeV. These criteria are selected to reduce the background due to misidentified leptons. The $m_{\ell\ell} > 10$ GeV condition is applied to remove low mass meson resonances and to reduce the DY background process. The requirement on missing transverse energy reconstructed based on tracks (section 2.3.6) $E_T^{miss,Track} > 20$ GeV is applied to events which contain zero or one jet. This selection is adapted from [2].

The background composition strongly depends on the jet multiplicity as shown in Figure 3.7. Based on the background composition, there are three characteristic

regions: the zero jet category mainly contains the $Z \to \tau\tau$ and WW background events, one jet category is dominated by the WW and top quark background processes, while in two or more jets category the top quark process is dominant. This motivates to optimise background rejection for each of the three listed jet categories individually.

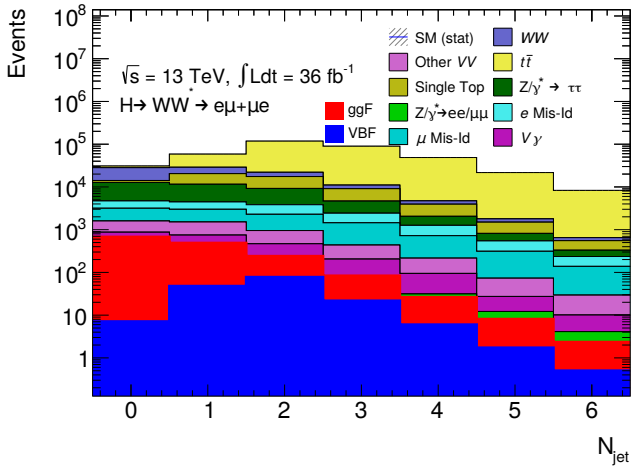


Figure 3.7: Jet multiplicity distribution obtained after the preselection criteria are applied.

3.3.2 Event selection for signal regions

This section presents the dedicated selection for the signal regions in each of the three jet categories. The zero jet category follows the measurement in [2], while in the remaining two categories the selection optimisation is investigated.

A b-veto requirement is applied for all jet categories, vetoing b-jets with transverse momentum $p_T > 20$ GeV, in order to reduce the top quark background contribution.

Zero jet category

The zero jet (0-jet) category corresponds to a set of events, which contain no reconstructed jets. Potentially mismeasured objects are removed using $\Delta\phi_{\ell\ell,E_T^{miss}}$ ob-

servable, which is an azimuthal angle between the dilepton system and the missing transverse energy (E_T^{miss}). Pathological events correspond to close to zero values of this observable, where E_T^{miss} points in the dilepton system's direction, and are reduced with the $\Delta\phi_{\ell\ell,E_T^{miss}} > \pi/2$ requirement. The $Z \to \tau\tau$ background is reduced by applying a requirement on the transverse momentum of the dilepton system $p_T^{\ell\ell} > 30$ GeV.

Charged leptons originating from the signal events are typically emitted at a small angle due to spin correlations (section 3.1.1), thus the dilepton observables corresponding to the dilepton invariant mass $m_{\ell\ell}$ and the azimuthal angle between two charged leptons $\Delta\phi_{\ell\ell}$ are useful to discriminate the signal. The values of $m_{\ell\ell} <$ 55 GeV and $\Delta\phi_{\ell\ell} < 1.8$ are applied in the selection. The selection is summarised in Table 3.2.

Table 3.2: Summary of the selection criteria applied on 0-jet, 1-jet, and $\geqslant$2-jet categories to construct the signal regions.

Observable \ Channel	0 jets	1 jet	$\geqslant$2 jets
Number of jets	0	1	>1
$E_T^{miss,Track}$/ GeV	> 20	> 20	-
b-tag veto	<1	<1	<1
$\Delta\phi_{\ell\ell,E_T^{miss}}$	$> \pi/2$	-	-
$p_T^{\ell\ell}$/ GeV	> 30	-	-
$m_{\ell\ell}$/ GeV	< 55	< 55	< 90
$\Delta\phi_{\ell\ell}$	< 1.8	< 1.8	-
m_T/ GeV	-	-	< 200
$m_{\tau\tau}$/ GeV	-	$<$ (m_Z -25)	< 66
max($m_T^{\ell_0}$,$m_T^{\ell_1}$)/ GeV	-	> 50	-
$\Delta\eta_{\ell\ell}$	-	-	< 2.6
$\Delta R_{\ell\ell}$	-	-	< 3.5

One jet category

The one jet (1-jet) category contains events with exactly one reconstructed jet. The selection is based on the characteristics of the signal and background processes.

The selection criteria based on the signal and background characteristics follow the analysis in [2]. The $Z \to \tau\tau$ and QCD background processes are reduced by applying the max($m_T^{\ell_0}$,$m_T^{\ell_1}$)> 50 GeV requirement, where max($m_T^{\ell_0}$,$m_T^{\ell_1}$) corresponds to the larger transverse mass of the two charged leptons, denoted as ℓ_0 and ℓ_1. The m_T^{ℓ} is defined as $m_T^{\ell} \equiv \sqrt{2p_T^{\ell}E_T^{miss}(1 - cos(\phi^{\ell} - \phi^{E_T^{miss}}))}$, p_T^{ℓ} and ϕ^{ℓ} are the transverse momentum and the azimuthal angle of the lepton and $\phi^{E_T^{miss}}$ is the azimuthal angle of the missing transverse energy. The processes with W bosons are affected less by this criterion since W bosons typically decay into charged leptons with large m_T^{ℓ} values. The $Z \to \tau\tau$ and $H \to \tau\tau$ processes are suppressed with a dedicated collinear mass of the di-τ system [99], $m_{\tau\tau}$. The selection criterion $m_{\tau\tau} < m_Z - 25$ GeV, where m_Z is the mass of the Z boson, is applied. The conditions on $m_{\ell\ell} < 55$ GeV and $\Delta\phi_{\ell\ell}$ <1.8 are applied based on the signal characteristics like in the 0-jet case.

The signal and background processes could be distinguished better using the fully reconstructed invariant mass of the final state particles, which is not possible in this case due to the presence of neutrinos in the final state. Therefore, a "transverse mass" m_T is introduced whose calculation does not require longitudinal neutrino momenta: $m_T = \sqrt{(E_T^{\ell\ell} + p_T^{\nu\nu})^2 - |\mathbf{p}_T^{\ell\ell} + \mathbf{p}_T^{\nu\nu}|^2}$, where $E_T^{\ell\ell} = \sqrt{(p_T^{\ell\ell})^2 + (m_{\ell\ell})^2}$, $\mathbf{p}_T^{\nu\nu}$ and $\mathbf{p}_T^{\ell\ell}$ are the vector sum of the neutrino and lepton transverse momenta, and $p_T^{\nu\nu}$ and $p_T^{\ell\ell}$ are their moduli, respectively. The m_T distribution has an upper bound at the Higgs boson mass, which can be used to separate the Higgs boson signal from the nonresonant WW and top quark pair background contributions.

The selection criteria in the 1-jet category are further optimised for this STXS analysis. The observables are selected based on the normalised signal and total

background distribution shapes, where the larger difference in shapes indicates a potential discriminating power between the signal and background processes. Consequently, the following observables are chosen: $m_{\tau\tau}$, $\Delta\phi_{\ell\ell}$, $\phi_{\ell\ell,E_T^{miss}}$, $p_T^{\ell\ell}$, m_T, E_T^{miss}, and the cone size of two charged leptons $\Delta R_{\ell\ell}$, where $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. The cut values are chosen based on the ratio of the total background and signal yields. The $\Delta R_{\ell\ell}$ observable's upper limit values are tested in the range [0.6, 3.8], for $m_{\tau\tau}$ in the range [20, 50] GeV, for $\Delta\phi_{\ell\ell}$ in the range [0.4, 3.0], for m_T in the range [110, 180] GeV, and E_T^{miss} in the range [90, 210] GeV. The lower limit is tested for observables $\phi_{\ell\ell,E_T^{miss}}$ and $p_T^{\ell\ell}$ in ranges [0.1, 2.6] and [30, 100] GeV, respectively. The performance of a selection criterion is evaluated by comparing the uncertainties of the STXS stage 1 ggF 1-jet signal strength parameters, which are defined as ratios of the observed and SM signal yields, corresponding to the regions defined in section 1.3. The uncertainties are obtained using a likelihood fit (section 4.1), where only statistical uncertainties are considered. The result of this optimisation is that no additional selection criteria improve the sensitivity of the measurement. The final selection is summarised in Table 3.2.

Two or more jets category

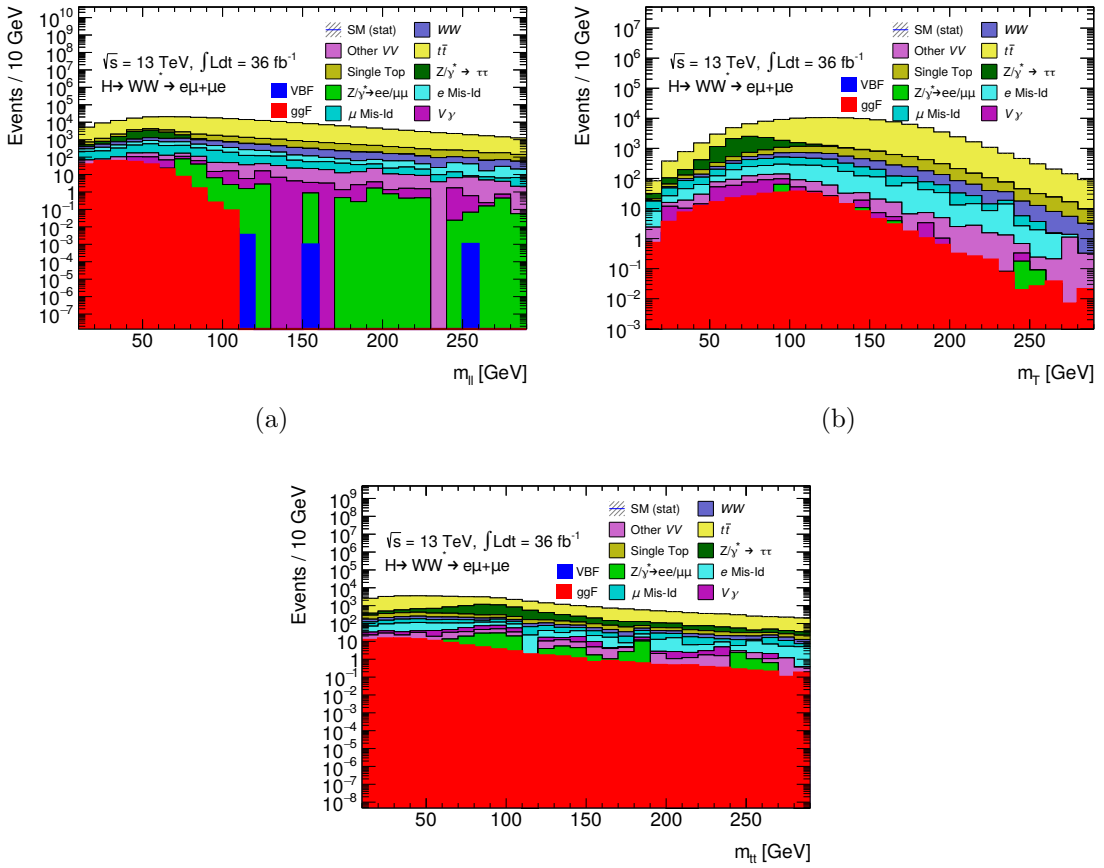
The two or more jets ($\geqslant$2-jet) category contains events with two or more reconstructed jets. The $\geqslant$2-jet analysis extends the 0-jet and 1-jet measurement in [2]. Moreover, this analysis uses a multivariate classifier (section 3.4.1), which is trained to separate signal and background events, instead of a potentially weaker single-variable classifier, m_T, used in the $H \to WW^*$ ggF analyses [2, 55]. The selection is primarily applied to reduce background contributions, so if a more powerful classifier is used, a larger signal region might be preferred due to the gain in signal yield.

The observables and the selection values are selected based on the combination of Run 1 analysis [55], argumentation in 0-jet and 1-jet cases, internal ATLAS group

work, and the shapes of normalised signal and total background distributions (like in 1-jet case). As a result, a wide range of values of the $m_{\ell\ell}$, m_T, $\Delta\phi_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, p_T^j, $p_T^{\ell\ell}$, and ΔR_{jj} observables are tested, where p_T^j is a transverse momentum of the highest p_T jet, and ΔR_{jj} is the difference in cone size of two leading jets (section 2.3.5). The selection criteria are selected based on the likelihood fit's output, as described in the 1-jet case. The requirement $m_{\ell\ell} < 90$ GeV mainly reduces WW, top quark, and W+jets events (Figure 3.8 (a)), while the $m_T < 200$ GeV requirement mainly reduces the top quark background (Figure 3.8 (b)). The requirement $m_{\tau\tau} < 66$ GeV is applied to reduce $Z \to \tau\tau$ events (Figure 3.8 (c)). The normalised total background and signal plots in Figure 3.9 show that the selection requirements on the difference of the pseudorapidities of two charged leptons $\Delta\eta_{\ell\ell} < 2.6$ and the difference in cone sizes of two charged leptons $\Delta R_{\ell\ell} < 3.5$ remove a fraction of background events without affecting the signal events since the signal distribution falls to zero sooner than the background distribution. The selection based on the observables $\Delta\phi_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, p_T^j, $p_T^{\ell\ell}$, and ΔR_{jj} is not effective. The selection is summarised in Table 3.2.

3.3.3 Categorisation of signal regions

The STXS analysis aims to measure cross-sections of the kinematic regions within one or more jets categories, which motivates to create dedicated signal subregions for each of the STXS signal categories (section 1.3). As a result, the one or more jets signal regions have to be divided into phase spaces enriched in individual STXS signal processes. Since the STXS framework is defined at particle level, the SR categorisation is not straightforward. Each of the three jet categories (section 3.3.2) are treated differently due to the granularity of the STXS regions and proportion of ggF and VBF production events. In the 1-jet SR case the Higgs boson signal is dominated by the ggF production, so the categorisation follows the STXS stage 1 ggF

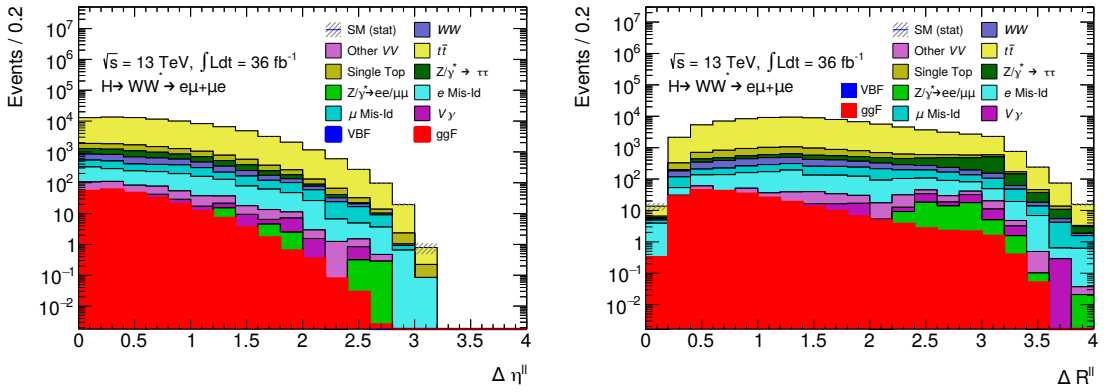


(a) (b)

(c)

Figure 3.8: Distributions of (a) $m_{\ell\ell}$, (b) m_T, and (c) $m_{\tau\tau}$ for the $\geqslant$2-jet category. The plot (a) is made just before the $m_{\ell\ell} < 90$ GeV, the (b) is made just before the $m_T < 200$ GeV, and (c) is made just before $m_{\tau\tau} < 66$ GeV requirements, listed in Table 3.2.

kinematic definitions. Finally, the $\geqslant$2-jet SR contains significant contributions from both the ggF and VBF Higgs boson production modes. It motivates the development of the categorisation targeted to create regions enriched in these processes. The categorisation is described below, where the categorisation of 1-jet and $\geqslant$2-jet SRs is developed as part of this thesis.



(a) (b)

Figure 3.9: Distributions of (a) $\Delta\eta_{\ell\ell}$ and (b) $\Delta R_{\ell\ell}$ observables. The plot in (a) is made just before the $\Delta\eta_{\ell\ell} < 2.6$, and (b) is made just before $\Delta R_{\ell\ell} < 3.5$ requirements, listed in Table 3.2.

Zero jet signal region categorisation

There is only one 0-jet STXS stage 1 ggF cross-section, so the division of the signal region is performed to increase the sensitivity of the measurement. The categorisation is performed based on the signal and background characteristics and follows the inclusive cross-section measurement in [2].

The $m_{\ell\ell}$ observable reflects $H \to WW^*$ spin correlation characteristic, which distinguishes the signal from the irreducible WW background process. As a result, the SR is divided into $m_{\ell\ell} \in [10, 30]$ GeV and $m_{\ell\ell} \in [30, 55]$ GeV regions based on different background distributions. These regions are further divided based on the subleading lepton p_T due to the composition of background processes. In the case of the background processes where W boson is produced in association with a jet or a photon (misreconstructed as a lepton) or an off-shell photon which decays into a low-mass lepton pair, the subleading lepton p_T distribution falls rapidly as p_T increases. Consequently, the regions are further divided into $[15, 20]$ GeV and $[20, \infty]$ GeV regions. Finally, the regions are further divided based on the leading lepton flavour since the jet and photon misidentification rates are different for electrons

and muons. The 0-jet SR categorisation is summarised in Figure 3.10.

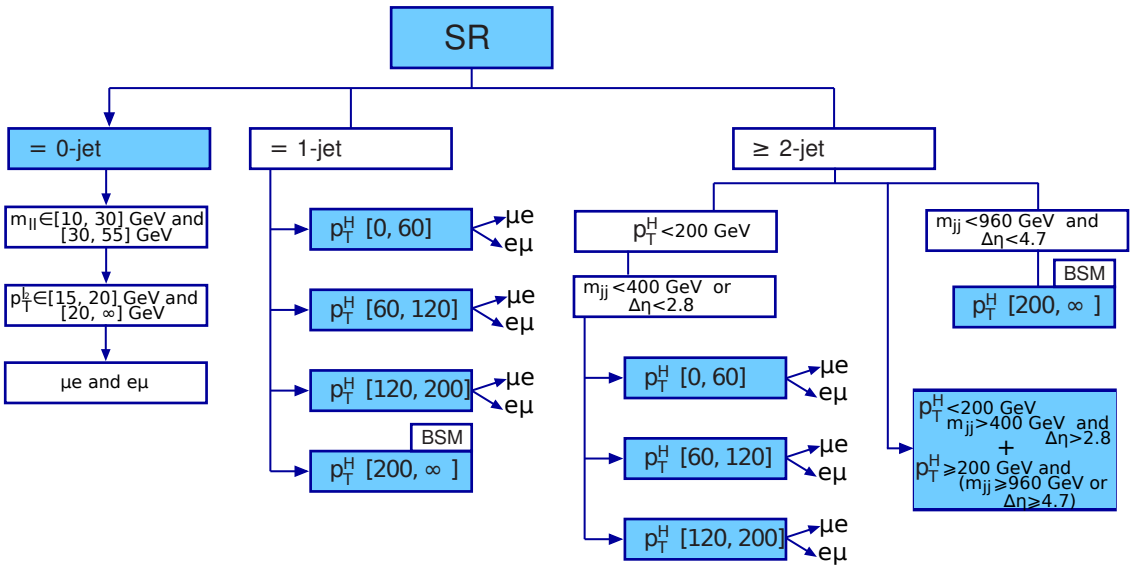


Figure 3.10: Overview of the SR's division into categories. The split of the region based on the leading lepton flavour is indicated by μe and $e\mu$ for the leading muon and electron, respectively. Here the notation p_T^H at the reconstructed level corresponds to the transverse momentum of the two charged leptons' and the missing energy system.

One jet signal region categorisation

In the 1-jet SR the Higgs boson signal is dominated by the ggF production, so the categorisation of the region is optimised only for the STXS stage 1 ggF 1-jet signal processes. The particle and reconstruction level events have similar kinematic properties, so it motivates to divide the SR into categories corresponding to the STXS kinematic definitions (section 1.3), where the notation p_T^H at the reconstructed level corresponds to the transverse momentum of the two charged leptons' and the missing energy system. The fractional contributions of the STXS signal processes in these reconstructed regions show that each region is enriched in the corresponding STXS signal process from 45% to 59% (Figure 3.11). As a result, the SR is split into four parts based on the p_T^H with the split boundaries at $p_T^H = 60$ GeV, $p_T^H = 120$

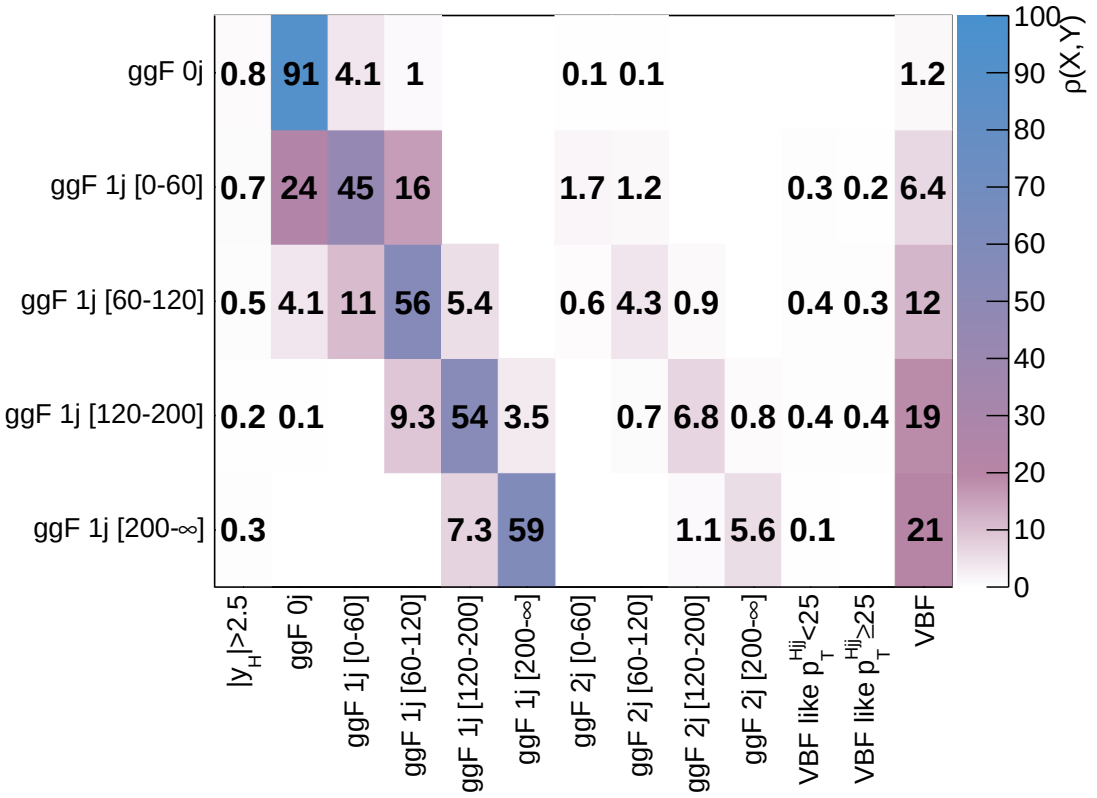


Figure 3.11: The relative contributions (in %) of the STXS ggF and VBF signals relative to the total Higgs boson signal, which includes VH and $H \to \tau\tau$ processes, in the reconstructed 1-jet signal regions. The STXS regions are indicated on the horizontal axis, and the reconstructed regions are indicated on the vertical axis.

GeV, and $p_T^H = 200$ GeV. The signal and background distributions in each of these categories are shown in Figure 3.12.

The regions satisfying $p_T^H < 200$ GeV are further split into leading lepton flavour due to different misidentification rates for electrons and muons, as described in 0-jet case. The region $p_T^H > 200$ GeV is not split based on the leading lepton flavour due to limited MC statistics. The categorisation of the 1-jet SR is summarised in Figure 3.10. Event yields of signal and background processes in the signal region categories before the split into leading lepton flavour are summarised in Table 3.3.

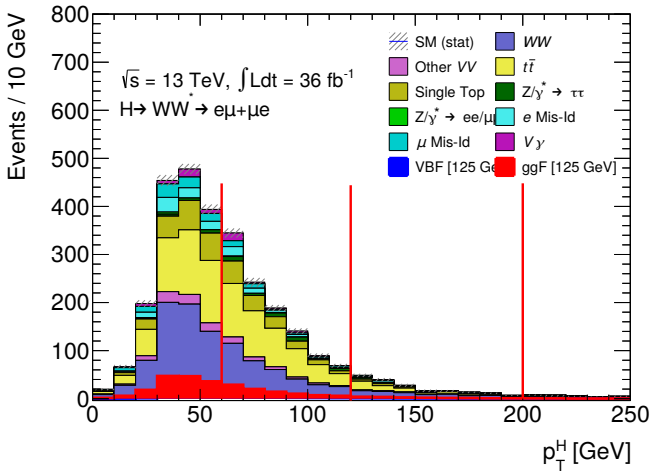


Figure 3.12: The signal and background distributions in the categories corresponding to the STXS stage 1 ggF kinematic regions, indicated by red bars, in 1-jet category. Here the notation p_T^H at the reconstructed level corresponds to the transverse momentum of the two charged leptons' and the missing energy system.

Table 3.3: Event yields of signal and background processes in 1-jet signal regions. The regions divided based on the leading lepton flavour are merged. WW, top quark and $Z \to \tau\tau$ yields are scaled with postfit normalisation factors listed in Table 3.6.

Process / Region	WW	Top quark	$Z \to \tau\tau$	Other $VV + V\gamma$	Mis-Id	ggF	VBF	Other Higgs
Inclusive SR	1019 ± 12	1242 ± 14	78 ± 9	193 ± 11	268 ± 13	264 ± 2	29.1 ± 0.2	2.69 ± 0.03
[0, 60] GeV	601 ± 10	633 ± 10	23 ± 5	115 ± 8	173 ± 10	155.9 ± 1.4	11.16 ± 0.13	1.52 ± 0.02
[60, 120] GeV	321 ± 6	532 ± 9	39 ± 6	66 ± 6	76 ± 8	87.1 ± 0.9	12.66 ± 0.14	0.89 ± 0.02
[120, 200] GeV	72 ± 3	64 ± 3	11 ± 2	9 ± 2	16 ± 4	16.6 ± 0.4	4.13 ± 0.08	0.22 ± 0.02
>200 GeV	25 ± 2	12 ± 2	4.9 ± 1.0	3 ± 2	4 ± 2	4.2 ± 0.2	1.19 ± 0.04	0.06 ± 0.01

Two or more jets signal region categorisation

The Higgs boson signal in $\geqslant$2-jet SR is enriched in both the ggF and VBF production processes, therefore the SR's division into categories corresponding to the STXS stage 1 ggF kinematic regions is not optimal. Moreover, the $\geqslant$2-jet STXS stage 1 ggF production category has two types: ggF ggF-like and ggF VBF-like (section 1.3). The current data set is not sensitive to the ggF VBF-like signal processes, so the analysis is designed to measure a combined VBF and ggF VBF-like cross-section.

The STXS stage 1 ggF region is already defined for the ggF-like and VBF-like parts, so the strategy is to split the SR based on the STXS ggF ggF-like $p_T^H < 200$ GeV kinematic regions and from the remaining phase space construct two regions enriched in the ggF ggF-like $p_T^H > 200$ GeV events, and the VBF and ggF VBF-like events. Consequently, the ggF ggF-like $p_T^H < 200$ GeV categories are defined in the phase space satisfying the selection of $p_T^H < 200$ GeV and $\{m_{jj} < 400$ GeV or $\Delta\eta_{jj} < 2.8\}$, where $\Delta\eta_{jj}$ and m_{jj} are the difference in pseudorapidity and the invariant mass of two leading jets, respectively. This region is further split at boundaries $p_T^H = 60$ GeV and $p_T^H = 120$ GeV into three parts. The signal and background distributions in each of the categories are shown in Figure 3.13. Finally, these regions are divided based on the leading lepton flavour due to different misidentification rates for electrons and muons, as in the 0-jet case. Fractional contributions of the STXS signal processes in these categories show that each region is enriched in the corresponding STXS type signal from 31% to 51% (Figure 3.14).

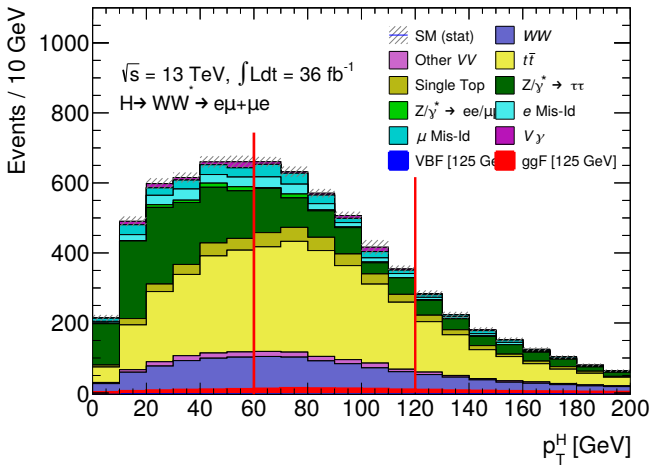


Figure 3.13: The signal and background distributions in the categories corresponding to the STXS stage 1 ggF kinematic regions, indicated by red bars, in ⩾2-jet category with $p_T^H < 200$ GeV and $\{m_{jj} < 400$ GeV or $\Delta\eta_{jj} < 2.8\}$ selection criteria. Here the notation p_T^H at the reconstructed level corresponds to the transverse momentum of the two charged leptons' and the missing energy system.

The remaining phase space is composed of the ggF ggF-like $p_T^H > 200$ GeV and

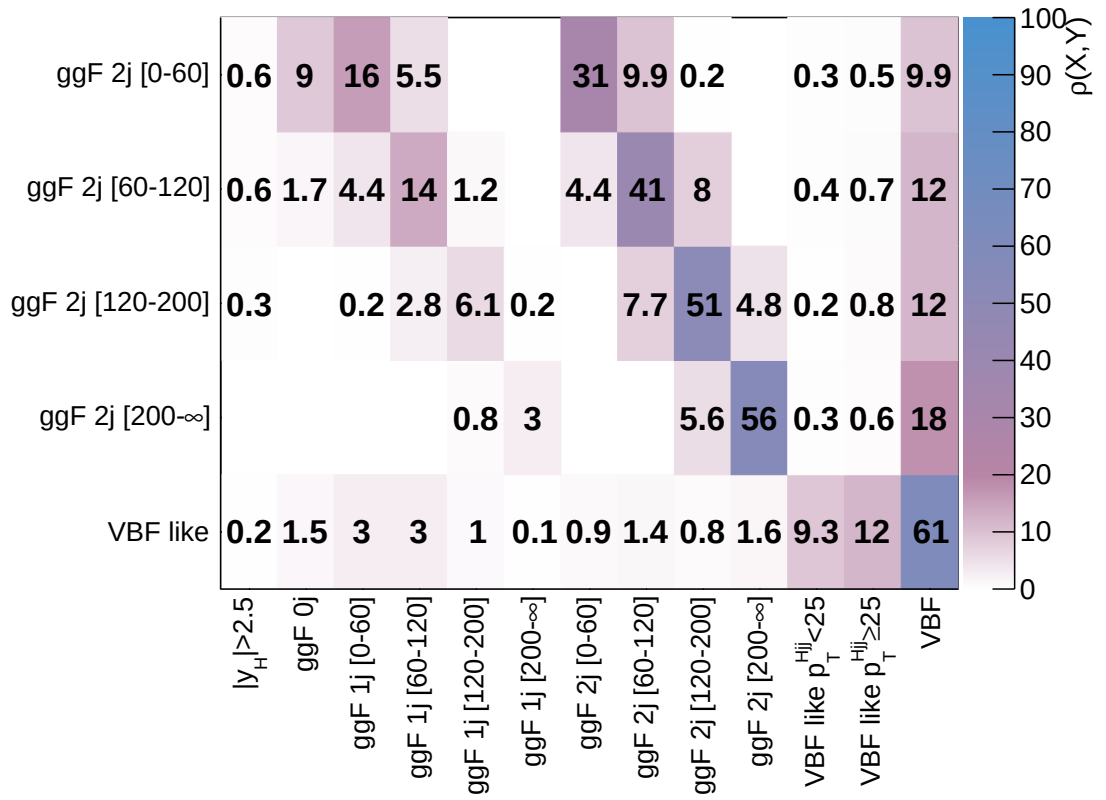
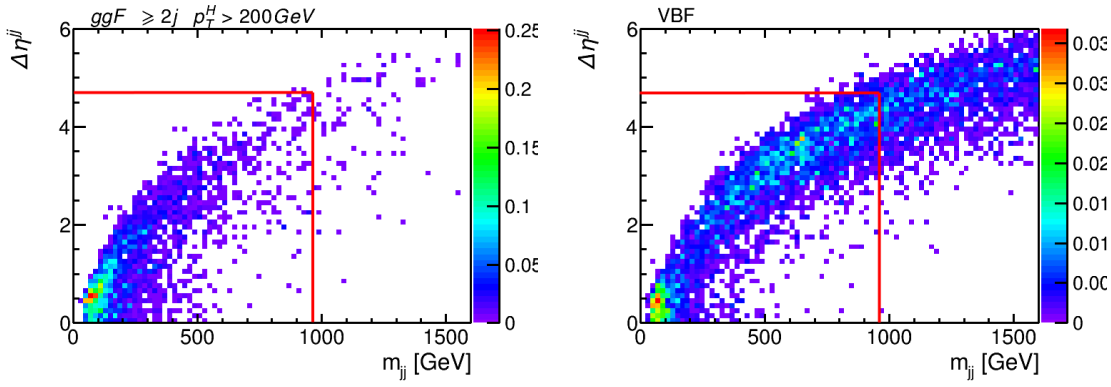


Figure 3.14: The relative contributions (in %) of the STXS ggF and VBF signals relative to the total Higgs boson signal, which includes VH and $H \to \tau\tau$ processes, in the reconstructed $\geqslant$2-jet signal regions. The STXS regions are indicated on the horizontal axis, and the reconstructed regions are indicated on the vertical axis.

ggF VBF-like regions. The former region has a ratio of the STXS ggF ggF-like $p_T^H > 200$ GeV and ggF VBF-like type events equal to 1.7, while in the case of the latter region this ratio is equal to $8 \cdot 10^{-3}$. Therefore, the strategy is to split the ggF ggF-like $p_T^H > 200$ GeV region into two parts enriched in the ggF ggF-like $p_T^H > 200$ GeV and the ggF VBF-like type events, respectively, and attach the latter region to the already defined ggF VBF-like region.

The ggF ggF-like $p_T^H > 200$ GeV region is split into two parts using $\Delta\eta_{jj}$ and m_{jj} observables since due to the typical two forward jets nature in the VBF case, the VBF signal tends to reside at higher $\Delta\eta_{jj}$ and m_{jj} values. A wide range of $\Delta\eta_{jj}$

and m_{jj} values are tested. The best $\Delta\eta_{jj}$ and m_{jj} values are determined using the ratio of the ggF ggF-like $p_T^H > 200$ GeV with the VBF and ggF VBF-like STXS signal yields, and the uncertainties on the corresponding signal processes obtained using a likelihood fit (section 4.1). The best performing criteria $\{m_{jj} \leqslant 960$ GeV and $\Delta\eta_{jj} \leqslant 4.7\}$ define a new region enriched in the STXS ggF ggF-like $p_T^H > 200$ GeV type events and is illustrated in Figure 3.15. This selection reduces the STXS ggF ggF-like $p_T^H > 200$ GeV signal events by 4.6% (from 16.3±0.4 to 15.6±0.4) while significantly reducing the VBF type events by 48% (from 9.3±0.1 to 4.9±0.1). The number of STXS ggF VBF-like type events is reduced negligibly.



(a) (b)

Figure 3.15: Distributions of (a) ggF $N_{\text{jet}} \geqslant 2$ with $p_T^H > 200$ GeV STXS type signal and (b) VBF signal. The region that satisfies $m_{jj} \leqslant 960$ GeV and $\Delta\eta_{jj} \leqslant 4.7$ selection is enriched in ggF $N_{\text{jet}} \geqslant 2$ with $p_T^H > 200$ GeV STXS type events and the remaining region is enriched in VBF type events.

The same strategy is also tested on the ggF ggF-like $p_T^H < 200$ GeV region. Various $\Delta\eta_{jj}$ and m_{jj} selection values are investigated. However, the nominal definition of $\{m_{jj} \leqslant 400$ GeV or $\Delta\eta_{jj} \leqslant 2.8\}$ is found to be optimal for this configuration. The ⩾2-jet SR categorisation is summarised in Figure 3.10. Event yields of signal and background processes in the signal region categories before the split into leading lepton flavour are summarised in Table 3.4.

Table 3.4: Event yields of signal and background processes in ⩾2-jet signal regions. The regions divided based on the leading lepton flavour are merged. Top quark and $Z \to \tau\tau$ yields are scaled with postfit normalisation factors listed in Table 3.6.

Process / Region	WW	Top quark	$Z \to \tau\tau$	Other $VV + V\gamma$	Mis-Id	ggF	VBF	Other Higgs
Inclusive SR	1545±9	4440±30	1720±30	362±16	660±20	198.1±1.4	73.7±0.3	11.0±0.3
[0, 60] GeV	457±6	1343±15	890±20	122±11	272±15	42.4±0.7	5.71±0.09	3.9±0.2
[60, 120] GeV	514±5	1840±20	380±20	115±7	245±13	70.7±0.9	10.93±0.13	2.53±0.12
[120, 200] GeV	256±3	649±11	169±9	44±5	71±8	40.4±0.6	6.47±0.10	1.27±0.08
>200 GeV	139±2	160±5	78±4	24±4	10±4	18.3±0.4	4.86±0.09	0.97±0.08
VBF-like	179±4	443±9	200±11	56±7	67±8	26.3±0.5	45.7±0.3	2.33±0.10

3.3.4 Background estimation

In this STXS analysis the background processes are estimated using three main techniques: templates and normalisations are derived from MC predictions, the MC templates are normalised using dedicated control regions (CR) in data, and templates and normalisations are derived with a fully data-driven method. The control region methodology is used to determine the normalisation factors for the WW (for the 0-jet and 1-jet categories only), top quark, and $Z \to \tau\tau$ background processes; W+jets background process is estimated using a data-driven technique, while the remaining background processes are evaluated using MC predictions. The following sections describe background estimation using the control region approach and the data-driven methodology.

Background measurement using control regions

The normalisation of the main background processes is determined using control regions (CR). CRs are constructed to be enriched in the corresponding background processes, the phase space of the CRs is orthogonal to all signal regions (SR) and kinematically similar to the signal regions on which the corresponding normalisation factors are applied. The main purpose of the CRs is to determine the normalisation

factor β, which is the ratio of the observed and the expected yields of the corresponding background process. The observed yield is determined by subtracting other background contributions from the data. The yield of the background process of interest N_{SR} is obtained from the following relation:

$$N_{SR} = B_{SR} \cdot N_{CR}/B_{CR} \tag{3.1}$$

where N_{CR} and B_{CR} are the observed and the MC yields in the CR, respectively. B_{SR} is the MC yield in the SR, and $\beta \equiv N_{CR}/B_{CR}$ [55].

The other advantage of using the CRs is the smaller theoretical uncertainties since they are estimated on the ratio B_{SR}/B_{CR} instead of on B_{SR} alone, and therefore the impact of the theoretical uncertainties is reduced as a result.

The normalisation factors for the WW (in 0-jet and 1-jet categories), top quark, and $Z \to \tau\tau$ processes are determined using control region methodology and the procedure follows [2]. The CRs are defined such that they do not overlap with SRs but are kinematically close. The WW CRs' selection has an $m_{\ell\ell}$ cut orthogonal to the SRs. In the top quark CRs case the requirement of at least one b-tagged jet ensures there is no overlap with the SRs. In the case of $Z \to \tau\tau$ process, the 0-jet regions do not overlap due to the $\Delta\phi_{\ell\ell}$ selection criterion, while the remaining CRs use an orthogonal to SRs $m_{\tau\tau}$ selection requirement. The selection criteria defining the CRs are summarised in Table 3.5.

The top quark CRs have a large fraction of top quark events: the 0-jet CR is 87% pure, the 1-jet CR is 96% pure, and the $\geqslant$2-jet CR is 95% pure. The $Z \to \tau\tau$ CRs are pure in Z/γ^* events to 88%, 69% and 65% in 0-jet, 1-jet, and $\geqslant$2-jet categories, respectively. The WW CR in 0-jet category is 70% pure in WW background events, while in the 1-jet CR there are only 39% WW events.

The normalisation factors for the corresponding background yields calculated using the methodology described in this section are summarised in Table 3.6. Both WW and top quark samples' normalisation factors are compatible with 1. In the

Table 3.5: Summary of the selection criteria for the WW, top quark, and $Z \to \tau\tau$ control regions. The selection starts with the preselection requirements defined in section 3.3.1. Here N_{jet} is a number of jets, which corresponds to the jet categories, and $N_{b\text{-jet}}$ is the number of b-tagged jets.

CR	$N_{\text{jet}} = 0$	$N_{\text{jet}} = 1$	$N_{\text{jet}} \geqslant 2$
WW	$\Delta\phi_{\ell\ell} < 2.6$ $E_T^{miss,Track} > 20$ GeV $\Delta\phi_{\ell\ell,E_T^{miss}} > \pi/2$ $p_T^{\ell\ell} > 30$ GeV 55 GeV $< m_{\ell\ell} <$ 110 GeV $N_{b\text{-jet}} = 0$	$E_T^{miss,Track} > 20$ GeV $N_{b\text{-jet}} = 0$ $\max(m_T^{\ell_0}, m_T^{\ell_1}) > 50$ GeV $\lvert m_{\tau\tau} - m_Z \rvert > 25$ GeV $m_{\ell\ell} > 80$ GeV	
Top quark	$m_{\ell\ell} > 10$ GeV $E_T^{miss,Track} > 20$ GeV $\Delta\phi_{\ell\ell,E_T^{miss}} > \pi/2$ $p_T^{\ell\ell} > 30$ GeV $\Delta\phi_{\ell\ell} < 2.8$ $N_{b\text{-jet}} > 0$	$m_{\tau\tau} < (m_Z$ -25 GeV) $E_T^{miss,Track} > 20$ GeV $N_{b\text{-jet}} = 1$ $\max(m_T^{\ell_0}, m_T^{\ell_1}) > 50$ GeV	$E_T^{miss,Track} > 20$ GeV $m_{\ell\ell} < 90$ GeV $\Delta R_{jj} > 1$ $m_T < 150$ GeV $N_{b\text{-jet}} = 1$ $p_T^{\ell\ell} > 20$ GeV $m_{\tau\tau} < 66$ GeV
$Z \to \tau\tau$	10 GeV $< m_{\ell\ell} <$ 80 GeV $N_{b\text{-jet}} = 0$ $\Delta\phi_{\ell\ell} > 2.8$	$N_{b\text{-jet}} = 0$ $\max(m_T^{\ell_0}, m_T^{\ell_1}) > 50$ GeV $m_{\tau\tau} > (m_Z$ -25 GeV) $m_{\ell\ell} < 80$ GeV	$E_T^{miss,Track} > 20$ GeV $m_{\ell\ell} < 90$ GeV $\Delta R_{jj} > 1$ $m_T < 150$ GeV $N_{b\text{-jet}} = 0$ 66 GeV $< m_{\tau\tau} <$ 116 GeV

case of the $Z \to \tau\tau$ CRs, the samples are scaled down in 0-jet and $\geqslant$2-jet categories.

Table 3.6: Postfit normalisation factors and their statistical and systematic uncertainties obtained using the dedicated control regions and the methodology described in this section. There is no WW CR defined in the $\geqslant$2-jet category.

Category	WW	Top quark	$Z \to \tau\tau$
$N_{\text{jet}} = 0$	1.04 ± 0.08	1.00 ± 0.15	0.81 ± 0.05
$N_{\text{jet}} = 1$	0.92 ± 0.15	0.98 ± 0.07	0.99 ± 0.10
$N_{\text{jet}} \geqslant 2$	-	1.02 ± 0.05	0.83 ± 0.08

The CRs discussed in this section were derived for the measurements in the

inclusive jet categories. This STXS analysis has dedicated signal subregions in each of the 1-jet and $\geqslant$2-jet signal regions (section 3.3.3). The modelling of the background processes has a potential to improve if dedicated CRs are derived for each of the signal subregions corresponding to the STXS stage 1 categories. It is only beneficial to create new CRs if they correspond to significantly different phase spaces and thus provide more accurate normalisation factors. Since the STXS ggF kinematic regions in 1-jet and $\geqslant$2-jet categories are defined based on the Higgs boson transverse momentum p_T^H, the new potential CRs are defined by dividing the existing CRs based on p_T^H at p_T^H values of 60 GeV, 120 GeV, and 200 GeV. This definition ensures that the new regions are kinematically close and orthogonal to the signal regions. The new CRs in $p_T^H > 200$ GeV category have limited data statistics. The remaining new CRs show similar background modelling with respect to the corresponding original CR. The modelling of the new CRs in the 1-jet category is illustrated in Figure 3.16. Thus, the definition of new CRs by splitting the existing CRs would not provide more accurate modelling of the background processes.

Background estimation using the data-driven technique

The lepton misidentification rate is low and difficult to model with MC, so the prediction of the background process due to misidentified leptons is obtained using a data-driven technique. The method is identical to the one used in [2].

The background estimation is obtained by extrapolating the number of events in a dedicated control sample to an analysis region. The control sample is constructed to be enriched in misidentified leptons. It follows the same kinematic selection as the analysis region, except it requires one of the leptons to be anti-id, which is defined as a lepton which does not meet the nominal identification and isolation criteria but satisfies looser identification requirements (section 2.3.3). The main source of misidentified leptons in control and analysis regions is W+jets. The estimated

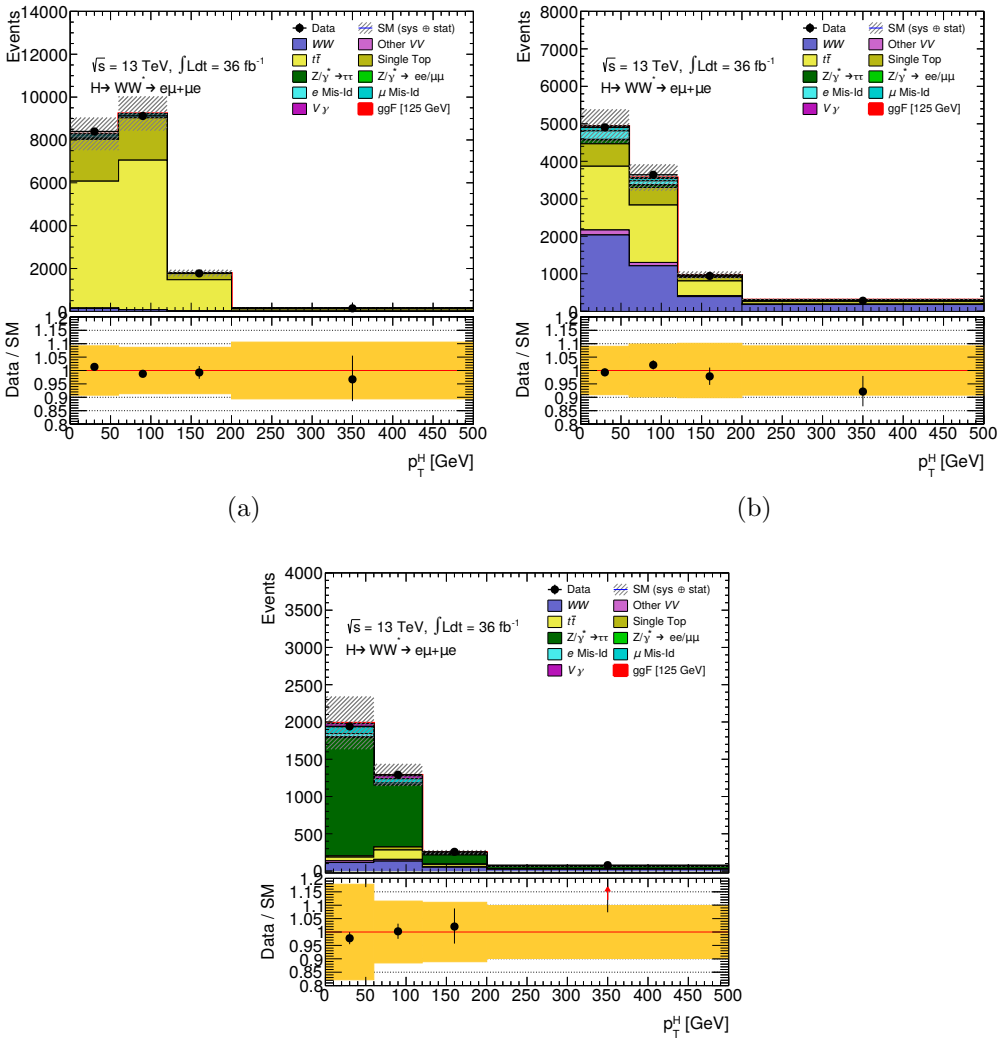


Figure 3.16: 1-jet control regions split based on the p_T^H values corresponding to the STXS definitions. (a) Top quark CR, (b) WW CR, and (c) $Z \to \tau\tau$ CR.

yield of the W+jets background process in an analysis region $N_{analysis}^{W+jets}$ to first order accuracy is:

$$N_{analysis}^{W+jets} = FF^{W+jets} \cdot N_{control}^{W+jets}, \tag{3.2}$$

where FF^{W+jets} is a fake factor, and $N_{control}^{W+jets}$ is the number of events in the control sample, which contains one anti-id lepton. The fake factor is determined as a

function of p_T and $|\eta|$ and is defined as:

$$FF(p_T, |\eta|) = \frac{N_{id}}{N_{\bar{id}}}(p_T, |\eta|), \tag{3.3}$$

where N_{id} is the number of leptons meeting the nominal identification and isolation criteria (also denoted as an id lepton) and $N_{\bar{id}}$ is the number of anti-id leptons. The fake factor is obtained for electrons and muons separately. Assuming the number of events with two anti-id leptons is negligible,

$$N_{control}^{W+jets} = N_{control}^{total} - N_{control}^{id,\bar{id},prompt} \tag{3.4}$$

where $N_{control}^{total}$ is the total number of data events in the control sample and $N_{control}^{id,\bar{id},prompt}$ is the number of events with prompt id and anti-id leptons derived using the MC samples. The FF^{W+jets} can be expressed as:

$$FF^{W+jets} = FF^{Z+jets}\frac{FF^{W+jets}}{FF^{Z+jets}} \tag{3.5}$$

where FF^{Z+jets} is a fake factor measured using a Z+jets data sample. The correction factor

$$CF = \frac{FF^{W+jets}}{FF^{Z+jets}} \tag{3.6}$$

is estimated using MC samples. Thus the W+jets background estimate is

$$N_{analysis}^{W+jets} = FF^{Z+jets} \cdot CF \cdot (N_{control}^{total} - N_{control}^{id,\bar{id},prompt}) \tag{3.7}$$

More details about the method can be found in [55, 100].

3.4 Event classification

The selection presented in the previous section defines phase space regions enriched in signal events. The signal events can be further discriminated from the background processes by redistributing events within signal regions using a classifier function. The classifier used in the previous $H \to WW^*$ ggF analyses [2, 55] is the

m_T distribution due to the different signal and background distributions' shapes. The availability of multivariate techniques allows one to improve the discrimination power of a classifier function by using multiple observables in its construction. This section discusses a construction of boosted decision tree (BDT) classifiers for the 1-jet and $\geqslant$2-jet categories.

The BDT technique is introduced in section 3.4.1. The methodology to select discriminating parameters for the BDT construction, including the ranking method created as part of this thesis, is presented in section 3.4.2. The construction of BDTs for the 1-jet and $\geqslant$2-jet categories is described in section 3.4.3.

3.4.1 Boosted decision tree theory and implementation

A decision tree is a multivariate binary classifier, which divides the phase space into a large number of hypercubes, each of which corresponds to one of the two categories enriched either in background or signal events. The decisions are made one variable at a time and an example decision tree is illustrated in Figure 3.17. The classification starts with the full data set. A single variable, which gives the best separation between the signal and background events, is selected and the optimal cut value is determined by scanning over the variable's range. This procedure is repeated on the resulting subsets of events until the preselected number of these procedures is reached. The resulting subsets are typically enriched in either background or signal events. A disadvantage of a decision tree is its instability due to the statistical fluctuations.

The instability problem is reduced by constructing a boosted decision tree (BDT), which is a multivariate classifier composed of multiple decision trees. The trees are trained on multiple training samples, constructed from the original sample. Finally, these trees are combined into a single classifier, which categorises events based on the response of the majority of the trees. More details can be found in [101].

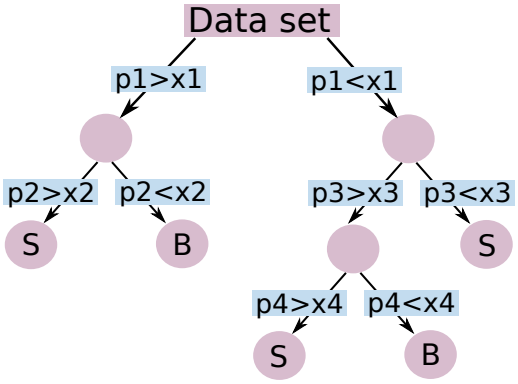


Figure 3.17: Example of a decision tree. The classification starts from the full data set. The best discriminating variable (p1) and cut value (x1) are selected to split the data set into two parts. The procedure is repeated on the resulting data sets until the stop criteria are met. Finally, the data set is divided into regions enriched in either signal or background events, here indicated as "S" and "B", respectively.

In this $H \to WW^*$ analysis BDT classifiers are constructed to discriminate the Higgs boson signal against the background processes. The BDTs are trained on the same MC and W+jets samples as are used for the main analysis (section 3.2). To avoid bias, the training data set is split into two statistically independent sets and the training of the BDTs is performed on each data set separately. In the main analysis the BDT trained on the first data set is used as a classifier on the second data set and vice versa. This way no BDT discriminant classifies events on the data set it was trained on. The data are also split into two sets and the BDT classifiers are applied in the same way as in the MC case.

3.4.2 Methodology for constructing boosted decision trees

A BDT's performance depends on the choice of the discriminating variables. A larger subset of discriminating variables typically produces a more powerful classifier, however, since the training is mainly performed on MC samples, each variable brings potential modelling uncertainty, which is difficult to control and validate. Therefore, an optimal subset of discriminating variables should be selected. In this

analysis different BDTs are trained for the 1-jet and $\geqslant$2-jet categories. 27 observables are identified in the 1-jet category as potential discriminating variables, and 45 observables are identified in the $\geqslant$2-jet case. Training BDTs and testing their performance on all combinatoric variations is computationally challenging, which leads to a need for a method to select discriminating observables.

This section discusses selection approaches applied in this thesis. The "observables' ranking method" is a systematic way to select the most promising candidates for the BDT training and is developed as part of this thesis. The method aims to rank observables such that the highest ranked observables are the best candidates for the BDT training. It is supplemented by less general approaches, which target the specific aspects of this $H \to WW^*$ analysis. These methods only propose potentially good combinations of discriminating parameters but the final set of discriminating variables should be determined by practically testing the performance of a BDT using a likelihood fit for the $H \to WW^*$ STXS signal cross-sections.

Three ranking methods

The idea of the BDT discriminating observables' ranking method (or the "ranking method") is based on the comparison of the shapes of normalised background and signal distributions of the discriminating observables of interest, illustrated in Figure 3.18. In this section three ways to quantify observable's discrimination power based on the shapes of background and signal distributions are presented. The methodology is created as part of this thesis.

The first approach is based on the idea of the centre of mass c of each normalised distribution $d(x)$ (Figure 3.19 (a)), which is defined as $c = \int d(x)xdx$, where x is an observable. Let D be the discrimination power:

$$D = \frac{2|c_{bkg} - c_{sig}|}{\Delta_{bkg} + \Delta_{sig}} \tag{3.8}$$

where c_{bkg} and c_{sig} are centres of mass of the background and signal distributions,

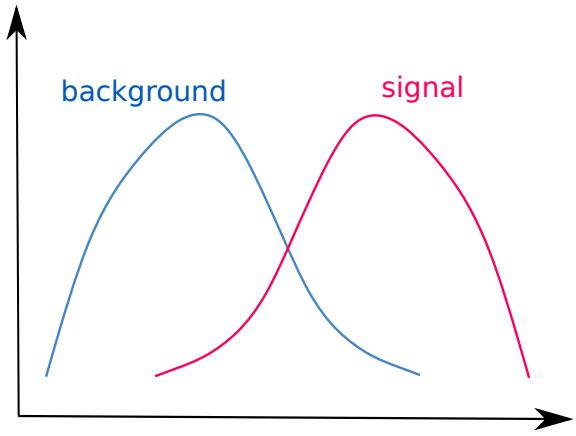


Figure 3.18: A schematic plot of normalised background and signal distributions.

while Δ_{bkg} and Δ_{sig} are equal to two square roots of variance of the background and signal distributions, respectively. The larger the D, the better the expected discrimination power of the observable.

The second approach to quantify the discrimination power is by comparing the area of the signal distribution that does not overlap with the background distribution, as illustrated in Figure 3.19 (b). The larger the area, the better the expected discrimination power.

The third discriminant is based on the χ^2 test idea, where background and signal distributions act as "theory" and "experiment" in a typical χ^2 test (Figure 3.19 (c)). First, distributions are binned with a predefined binning, then the χ^2 value is calculated as in equation 3.9, where N_{bins} is the total number of bins, b_i^{bkg} and b_i^{sig} are background and signal yields in i^{th} bin, respectively.

$$\chi^2 = \frac{1}{N_{bins}} \sum_{i=1}^{N_{bins}} (b_i^{bkg} - b_i^{sig})^2 \tag{3.9}$$

These methods are sensitive to the choice of binning of the histograms due to the finite MC/data samples' size. To ensure that deviations are physical rather than statistical, histograms are binned such that the bin content in each bin is not smaller than the MC statistical uncertainty on that bin and the background yield is uniform

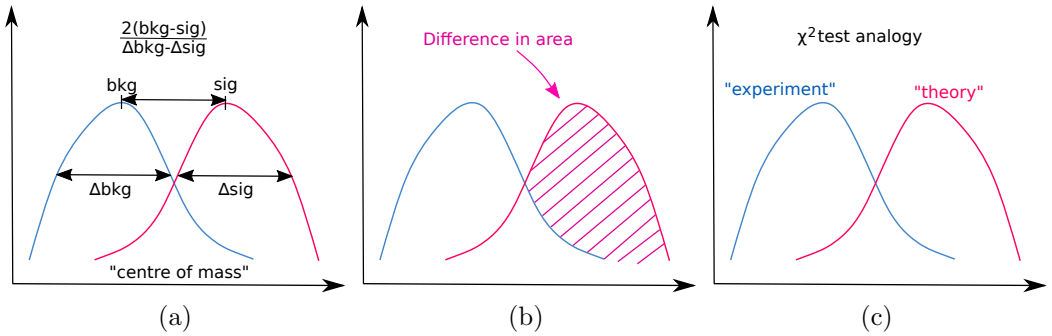


Figure 3.19: Three ranking methods: (a) based on the centre of mass idea, (b) based on the plot area that does not overlap with the other distribution, and (c) based on the χ^2 test analogy.

in the bins.

Comparison of three ranking methods

It is expected that the three ranking discriminants provide similar ranking of the observables. This is tested on the 1-jet inclusive signal region, defined in section 3.3.2. The comparison of the three discriminants is summarised in Table 3.7. All discriminants provide similar ranking: for the first 14 observables only one rank differs by three units, while the remaining differences are smaller. This is sufficient since the ranking is intended to be used as a guidance for further studies, thus the small differences in the results of the three ranking methods are not important.

Ranking method validation

The proposed ranking method is practically applicable if and only if the BDTs constructed from the higher ranked observables perform better than the BDTs constructed from other observable subsets. The validation strategy is the following: an STXS signal process and its corresponding signal subregion are selected; the ranking of observables is produced on the signal subregion and the higher ranked discriminating observables are used for the BDT training. Let us denote these BDTs as test BDTs. The performance of a BDT is evaluated based on the uncertainties on

Table 3.7: The comparison of three ranking methods. The first column indicates the rank of the observables, the next three columns list the ranking of the observables, while the last two columns indicate observable's rank with respect to the first discriminant. D_1 corresponds to the ranking method based on the plot area, D_2 is a discriminant derived based on the centre of mass, and D_3 discriminant is based on the χ^2 test analogy. The table is truncated after the first eight observables; the full list is in Table A.1.

Rank	D_1	D_2	D_3	Rank(D_1)-Rank(D_2)	Rank(D_1)-Rank(D_3)
1	m_T	m_T	m_T	0	0
2	E_T^{miss}	E_T^{miss}	E_T^{miss}	0	0
3	$\Delta R_{\ell_0 j_0}$	$\Delta R_{\ell_0 j_0}$	$\Delta R_{\ell_0 j_0}$	0	0
4	$m_{\ell\ell}$	$\Delta R_{\ell_1 j_0}$	$\Delta R_{\ell_1 j_0}$	1	1
5	$\Delta R_{\ell_1 j_0}$	$m_{\ell\ell}$	$m_{\ell\ell}$	-1	-1
6	$\Delta\eta_{\ell_0 j_0}$	$\Delta\phi_{\ell\ell,j}$	$m_{\ell_1 j}$	1	3
7	$\Delta\phi_{\ell\ell,j}$	$\Delta\eta_{\ell_0 j_0}$	$\Delta\phi_{\ell\ell,j}$	-1	0
8	$\Delta\eta_{\ell_1 j_0}$	$\Delta\eta_{\ell_1 j_0}$	$\Delta\eta_{\ell_0 j_0}$	0	-2

the STXS signal strength parameter, evaluated using a likelihood fit (section 4.1.1) with that BDT as a classifier. This BDT performance is compared with a control BDT set, which is independent of the ranking on this signal region (the set will be specified later). The observables' ranking method is concluded to be effective if and only if the STXS signal measurement improves when the BDT is obtained from the ranking method.

The validation is performed on all four 1-jet signal subregions corresponding to the STXS stage 1 ggF signal definitions, defined at Higgs boson transverse momentum p_T^H intervals of [0, 60] GeV, [60, 120] GeV, [120, 200] GeV, and $>$ 200 GeV. The validation is performed on each region independently of the other regions. The control BDT set is formed from the BDTs produced for the remaining signal regions, i.e. a BDT is a test BDT if it is used on the region on which it was ranked, and it is a control BDT if it is used on the other three regions. The control BDT set

is not random and is correlated with the test BDT set due to kinematic similarities between the regions, which means that the ranking method has to be accurate enough to account for these small kinematic differences in order to pass the validation. Table 3.7 showed that all three discriminants perform similarly, so any of the discriminants may be chosen. The ranking of observables is performed using the χ^2 approach and the ranking is listed in Table 3.8. The BDTs are constructed from seven discriminating observables ranked in the corresponding region, where the first six observables are the top six ranked observables and the seventh is taken from the set of remaining observables in Table 3.8 one per BDT. The seventh observable acts as a fluctuation and allows to have multiple BDTs with the dominant six highest ranked observables. It results in the construction of 21 BDTs per signal subregion. All BDTs are trained on the inclusive 1-jet SR (section 3.3.2) in order to train on a sufficiently large sample. The performance of a BDT is evaluated with the likelihood fit (section 4.1.1) and is compared with the nominal BDT, constructed from top six observables ranked on the inclusive 1-jet signal region (Table 3.8). It is expected that the BDTs obtained using the dedicated ranking for the signal subregion perform better than the nominal BDT, while the BDTs from the control sample (i.e. produced using the ranking on the remaining regions) perform worse than the nominal BDT.

The results are presented in Table 3.9. The diagonal values show the performance of the test BDT set, and off-diagonal terms show the performance of the control BDT set. Each value corresponds to the performance of 21 BDTs. The control set is split based on the regions where the BDTs were ranked to track the effect of potentially overlapping rankings in different signal subregions. The results demonstrate that there is a tendency that an STXS measurement improves if the ranking for the BDT is performed on the corresponding signal subregion, and does not improve otherwise. The reason an improvement is observed in some other neighbouring regions is due

Table 3.8: Observable ranking using the χ^2 type method on 1-jet reconstructed regions corresponding to the STXS stage 1 ggF definitions. E, m, p correspond to the energy, mass and momentum, respectively, while ϕ and η are the azimuthal angle and pseudorapidity. Subscript T indicates that the quantity is derived in the transverse direction. Subscripts ℓ_0 and ℓ_1 correspond to the leading and subleading charged leptons, respectively, while j corresponds to the leading jet. Subscript H indicates that the observable is calculated for the Higgs boson. Δ corresponds to the difference in the quantity followed by this sign.

Rank	Inclusive 1-jet region	1-jet $p_T^H \in$[0, 60] GeV	1-jet $p_T^H \in$[60, 120] GeV	1-jet $p_T^H \in$[120, 200] GeV	1-jet $p_T^H >$200 GeV
1	m_T	m_T	m_T	m_T	m_T
2	E_T^{miss}	E_T^{miss}	E_T^{miss}	$m_{\ell_1 j}$	E_T^{miss}
3	$\Delta R_{\ell_0 j}$	$p_T^{\ell_0}$	$\Delta R_{\ell_0 j}$	$\Delta\phi_{\ell\ell,j}$	$\Delta\phi_{\ell\ell,j}$
4	$\Delta R_{\ell_1 j}$	$\Delta R_{\ell_0 j}$	$\Delta R_{\ell_1 j}$	E_T^{miss}	$\Delta R_{\ell\ell}$
5	$m_{\ell\ell}$	$p_T^{\ell\ell}$	$\Delta\phi_{\ell\ell,j}$	$\Delta R_{\ell\ell}$	$p_T^{\ell_1}$
6	$m_{\ell_1 j}$	$m_{\ell\ell}$	$m_{\ell_1 j}$	$\Delta\phi_{\ell_1 j}$	$\Delta\phi_{\ell\ell,E_T^{miss}}$
7	$\Delta\phi_{\ell\ell,j}$	$\Delta R_{\ell_1 j}$	$\Delta\phi_{\ell_0 j}$	$\Delta R_{\ell_0 j}$	$\Delta\phi_{\ell\ell}$
8	$\Delta\eta_{\ell_0 j}$	$\Delta\phi_{\ell\ell,j}$	$m_{\ell_0 j}$	$\Delta R_{\ell_1 j}$	$\Delta\phi_{\ell_0 j}$
9	$\Delta\eta_{\ell_1 j}$	$\Delta\eta_{\ell_0 j}$	$m_{\ell\ell}$	$\Delta\phi_{\ell\ell,E_T^{miss}}$	$\Delta\phi_{\ell_1 j}$
10	$\Delta\phi_{\ell_0 j}$	$\Delta\eta_{\ell_1 j}$	$\Delta\phi_{\ell_1 j}$	$p_T^{\ell_1}$	$p_T^{\ell\ell}$
11	$\Delta\phi_{\ell_1 j}$	$m_{\ell_1 j}$	$\Delta\eta_{\ell_0 j}$	$\Delta\phi_{\ell_0 j}$	$\Delta R_{\ell_1 j}$
12	$m_{\ell_0 j}$	$\Delta\phi_{\ell_0 j}$	$\Delta\eta_{\ell_1 j}$	$\Delta\phi_{\ell\ell}$	$m_{\ell_1 j}$
13	$\Delta R_{\ell\ell}$	$\Delta\phi_{\ell_1 j}$	$\Delta R_{\ell\ell}$	$m_{\ell_0 j}$	$m_{\ell_0 j}$
14	$\Delta\phi_{j,E_T^{miss}}$	$\Delta\phi_{\ell\ell,E_T^{miss}}$	$\Delta\phi_{\ell\ell,E_T^{miss}}$	$p_T^{\ell_0}$	$p_T^{\ell_0}$
15	$p_T^{\ell_0}$	$m_{\ell_0 j}$	$\Delta\phi_{j,E_T^{miss}}$	$m_{\ell\ell}$	$\Delta\eta_{\ell\ell}$
16	$\Delta\eta_{\ell\ell}$	p_T^j	$\Delta\phi_{\ell\ell}$	$\Delta\eta_{\ell\ell}$	$\Delta R_{\ell_0 j}$
17	p_T^j	$\Delta\phi_{j,E_T^{miss}}$	$p_T^{\ell_0}$	$p_T^{\ell\ell}$	$m_{\ell\ell}$
18	η_{ℓ_0}	η_{ℓ_0}	$\Delta\eta_{\ell\ell}$	$\Delta\eta_{\ell_0 j}$	η_{ℓ_0}
19	η_{ℓ_1}	$p_T^{\ell_1}$	$p_T^{\ell\ell}$	$\Delta\eta_{\ell_1 j}$	$\Delta\phi_{j,E_T^{miss}}$
20	$\Delta\phi_{\ell\ell}$	$\Delta R_{\ell\ell}$	η_j	$\Delta\phi_{j,E_T^{miss}}$	$\Delta\eta_{\ell_0 j}$
21	η_j	η_{ℓ_1}	η_{ℓ_0}	η_{ℓ_1}	ϕ_{ℓ_0}
22	$p_T^{\ell\ell}$	η_j	η_{ℓ_1}	η_j	p_T^j
23	$p_T^{\ell_1}$	$\Delta\eta_{\ell\ell}$	$p_T^{\ell_1}$	p_T^j	p_T^H
24	p_T^H	$\Delta\phi_{\ell\ell}$	ϕ_{ℓ_0}	η_{ℓ_0}	ϕ_{ℓ_1}
25	$\Delta\phi_{\ell\ell,E_T^{miss}}$	p_T^H	ϕ_{ℓ_1}	ϕ_{ℓ_0}	η_j
26	ϕ_{ℓ_1}	ϕ_{ℓ_0}	p_T^j	ϕ_{ℓ_1}	$\Delta\eta_{\ell_1 j}$
27	ϕ_{ℓ_0}	ϕ_{ℓ_1}	p_T^H	p_T^H	η_{ℓ_1}

to similar ranking in neighbouring regions, e.g. m_T, E_T^{miss}. The reason the ranking on the $p_T^H \in$[60, 120] GeV region does not significantly improve the measurement of

the p_T^H ∈[60, 120] GeV signal strength parameter of interest (POI) with respect to the nominal BDT is because five out of top six ranked observables coincide in both regions (Table 3.8). It is not expected to have a 100% improvement with respect to the nominal BDT since, in the case the ranking method is effective, the nominal BDT should already provide a competitive performance. The remarkable outcome of this validation study is that the ranking method is accurate enough to account for the differences between kinematically similar regions.

Table 3.9: Results of the validation of the ranking method on 1-jet signal subregions. The BDTs are constructed based on the ranking of observables on the regions indicated in the columns' labels. The regions in which these BDTs classify the events are listed in the rows' labels. The percentage values indicate the fraction of the BDTs that perform better than the nominal BDT. 0% is replaced by "-" for readability. The diagonal values correspond to the performance of the ranking method, while the off-diagonal values correspond to the performance of the control BDT set.

Ranked on / Fit on region	p_T^H ∈[0, 60]GeV	p_T^H ∈[60, 120]GeV	p_T^H ∈[120, 200]GeV	p_T^H >200GeV
p_T^H ∈[0, 60] GeV	44%	6%	-	-
p_T^H ∈[60, 120] GeV	-	6%	-	-
p_T^H ∈[120, 200] GeV	-	-	50%	19%
p_T^H >200 GeV	-	-	6%	44%

It can be concluded that the observables' ranking method can be used as a starting point for the selection of the most promising observables for the BDT construction.

Other BDT construction aspects

The ranking method can be supplemented by a visual examination of the shapes of signal and background distributions to account not only for the difference between signal and background shapes for an individual observable but also for the difference

in shapes between different observables. Moreover, some observables describe similar kinematic aspects, so the ranking method is likely to assign them similar rank. The collective discimination power of these observables is likely to be similar to the power of a single observable since there is little additional kinematic information. This suggests to avoid having kinematically close observables in the BDT's input set.

In addition to the choice of discriminating observables, the three main aspects of the BDT discriminant's optimisation relevant to this $H \to WW^*$ analysis are the following: the phase space on which BDT is trained, the background processes used for the BDT training and for the ranking method, and the number of discriminating observables. Regarding the phase space of the training sample, it would be ideal to train a separate BDT for each signal subregion corresponding to the STXS POIs. The choice of a relatively small phase space results in limited MC statistics of a training sample. It may be beneficial to extend the training phase space to a larger kinematically close region, e.g. an inclusive SR of the corresponding jet category or even larger phace space by removing a part of selection criteria. The uncertainties due to specific background processes can be targeted by selecting only these processes in the set of training samples or in the ranking method. In particular, it applies to $Z \to \tau\tau$ and W+jets processes. In contrast, removing background processes which have small yields in this analysis may reduce unwanted statistical fluctuations in the training process. In the case of W+jets it would also significantly reduce computation time. A larger BDT discriminating observable set is expected to lead to a more powerful classifier, although each discriminating observable brings modelling uncertainty. It is only beneficial to use large sets of discriminating observables if the increase in BDT's discrimination power is significant enough. Sets with different number of differentiating variables can be tested to obtain the optimal number.

Finally, the p_T^H observable is used to define most of the STXS stage 1 ggF regions.

So, in the case where a single BDT is trained for the whole jet category, inclusion of the p_T^H observable in the set of discriminating observables might account for the different kinematic properties of the individual STXS categories.

3.4.3 Boosted decision tree construction for jet categories

This section discusses the construction of dedicated BDTs for each of the 1-jet and $\geqslant$2-jet categories. The BDTs are constructed using the "base" and "additional" subsets of observables, where the BDT's input set is constructed by taking all parameters from the base subset and adding an indicated number of observables (typically from one to three) from the additional subset. The performance of a BDT is evaluated using a likelihood fit for STXS signal strength parameters. The performance is quantified based on the size of uncertainties on the POIs corresponding to the SRs on which the BDT is applied. There are 27 and 45 discriminating observables defined for the 1-jet and $\geqslant$2-jet categories, respectively.

BDT construction for 1-jet category

This section is dedicated to the selection of the best performing BDT for the 1-jet category. The BDT candidates are constructed taking into account the aspects discussed in section 3.4.2, in particular: the ranking method, selection based on the shapes of signal and background distributions, the training on different phase spaces, inclusion of the p_T^H observable to the discriminating parameters' set, exclusion of the W+jets process from the training data set, and the impact of a large number of discriminating observables.

First, the selection of discriminating observables is performed by using the ranking method. The observables are ranked in both the inclusive 1-jet signal region and in the 1-jet signal subregions corresponding to the STXS kinematic regions using the χ^2 ranking approach (Table 3.8). The BDTs are constructed from the first five observables in each column plus one observable per BDT from the remaining set of

observables. A single BDT is used to classify events in all 1-jet signal subregions in the fit. The best performing BDT is obtained when the following subset is used: $\{m_T, E_T^{miss}, \Delta R_{\ell_0 j}, \Delta R_{\ell_1 j}, m_{\ell\ell}, m_{\ell_1 j}\}$, which corresponds to the first six observables ranked in the 1-jet inclusive signal region. Let us denote it as BDT-A.

The second BDT candidate is obtained based on the examination of the shapes of normalised signal and background distributions. The base and the additional subsets are created (Table A.2) and BDTs are constructed by adding one parameter per BDT to the base subset. BDTs are trained on the inclusive 1-jet signal region. The best performing BDT is constructed from the set of $\{m_{\ell\ell}, m_T, \Delta R_{\ell\ell}, \Delta\phi_{\ell\ell,E_T^{miss}}, \Delta\phi_{\ell\ell,j}, \Delta\eta_{\ell_0 j}\}$ observables, let us denote this BDT as BDT-B.

BDT-A and BDT-B are compared in Table 3.10. Both BDTs show similar performance. BDT-B is also compared with the fit configuration, where the m_T distribution acts as a classifier, in Table 3.11. The BDT discriminant performs significantly better than a single variable classifier m_T, especially for the POIs satisfying $p_T^H > 60$ GeV, which justifies the use of multivariate discriminants.

Table 3.10: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with BDT-A and BDT-B classifiers, which are obtained in the studies based on the ranking of discriminating observables, and the shapes of signal and background distributions, respectively. Red colour indicates that at least one of the uncertainties of this BDT is at least 0.01 units smaller than the other BDT. Only statistical uncertainties are considered.

BDT / STXS POI	BDT-B	BDT-A
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.56}_{-0.55}$	$1.00^{+0.56}_{-0.54}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.54}_{-0.53}$	$1.00^{+0.55}_{-0.53}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.02}_{-0.95}$	$1.00^{+1.04}_{-0.98}$
1-jet, $p_T^H >$200 GeV	$1.00^{+1.77}_{-1.50}$	$1.00^{+1.76}_{-1.48}$

Table 3.11: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in the 1-jet signal regions are classified with the m_T distribution and BDT-B classifier. Only statistical uncertainties are considered.

STXS POI \ Classifier	m_T	BDT-B
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.57}_{-0.56}$	$1.00^{+0.56}_{-0.55}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.63}_{-0.61}$	$1.00^{+0.54}_{-0.53}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.20}_{-1.15}$	$1.00^{+1.02}_{-0.95}$
1-jet, $p_T^H >$200 GeV	$1.00^{+2.15}_{-1.94}$	$1.00^{+1.77}_{-1.50}$

The impact of the size of the training samples is tested by training BDTs on smaller and larger phase spaces than the 1-jet SR. The new BDTs are trained using the parameters used for BDT-B training. In the case of the reduced phase space, the four BDTs are trained on individual signal regions corresponding to the STXS kinematic definitions (section 3.3.3). These BDTs and BDT-B are compared in Table 3.12. The measurement does not improve, the likely cause is the reduced MC statistics available for training. When a BDT is trained on a larger phase space in order to gain MC statistics, the inclusive 1-jet signal region is extended by removing $\max(m_T^{\ell_0}, m_T^{\ell_1})$, $m_{\tau\tau} - m_Z$, and $\Delta\phi_{\ell\ell}$ requirements (section 3.3.2). The BDT is constructed using the set of discriminating parameters used for BDT-A training. This BDT and BDT-A are compared in Table 3.13. There is a negligible improvement on the POIs satisfying $p_T^H < 200$ GeV condition, while the $p_T^H > 200$ GeV POI performs better with the BDT trained on the inclusive signal region. Since the $p_T^H > 200$ GeV POI is a POI sensitive to BSM physics, BDT trained on the inclusive signal region is preferred.

In the case a single BDT is used for all signal subregions, the p_T^H observable may account for the different characteristics of the signal subregions. A new BDT is

Table 3.12: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with the BDTs trained using the same set of discriminating parameters as in the BDT-B case and trained on the 1-jet inclusive signal region and on the individual signal regions corresponding to the STXS kinematic regions. Only statistical uncertainties are considered.

Classifier / STXS POI	BDT trained on inclusive SR	Individual BDTs trained for each SR corresponding to STXS definition
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.56}_{-0.55}$	$1.00^{+0.56}_{-0.55}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.55}_{-0.53}$	$1.00^{+0.55}_{-0.53}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.03}_{-0.96}$	$1.00^{+1.09}_{-1.03}$
1-jet, $p_T^H >$200 GeV	$1.00^{+1.67}_{-1.38}$	$1.00^{+1.98}_{-1.74}$

Table 3.13: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with the BDTs trained on the extended phase space and on the inclusive signal region, respectively. The BDTs are trained using the set of discriminating parameters as in the BDT-A case. Only statistical uncertainties are considered.

BDT trained on / STXS POI	Extended phase space	Inclusive signal region
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.55}_{-0.54}$	$1.00^{+0.55}_{-0.54}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.54}_{-0.52}$	$1.00^{+0.54}_{-0.53}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.04}_{-0.97}$	$1.00^{+1.04}_{-0.98}$
1-jet, $p_T^H >$200 GeV	$1.00^{+1.76}_{-1.49}$	$1.00^{+1.75}_{-1.48}$

trained using the parameters used for BDT-B training. The new BDT and BDT-B are compared in Table 3.14. There is a negligible improvement on the $p_T^H < 200$ GeV POIs, while the $p_T^H > 200$ GeV POI performs better with the BDT constructed without the p_T^H observable. Since the $p_T^H > 200$ GeV POI is more sensitive to BSM physics than the other POIs, the BDT constructed without the p_T^H observable is

preferred.

Table 3.14: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with BDT-B and the BDT trained using the set of parameters used for BDT-B training supplemented with the p_T^H observable. Only statistical uncertainties are considered.

STXS POI \ BDT type	BDT subset without p_T^H	BDT subset with p_T^H
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.56}_{-0.55}$	$1.00^{+0.56}_{-0.54}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.55}_{-0.53}$	$1.00^{+0.54}_{-0.52}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.03}_{-0.96}$	$1.00^{+1.03}_{-0.96}$
1-jet, $p_T^H >$200 GeV	$1.00^{+1.67}_{-1.38}$	$1.00^{+1.77}_{-1.5}$

The small fluctuations due to the W+jets sample are tested by training a BDT with a set of BDT-B's discriminating parameters on the training samples without W+jets process. This BDT and BDT-B are compared in Table 3.15. The BDT trained on the set without W+jets sample performs better on the $p_T^H < 200$ GeV POIs, while BDT-B performs significantly better on the $p_T^H > 200$ GeV POI. As a result, since the $p_T^H > 200$ GeV POI is more sensitive to BSM physics, the preference is to use the BDT trained on the samples including W+jets process.

The impact of the number of discriminating parameters is evaluated by composing BDTs using 11 discriminating parameters. The sets of discriminating parameters are constructed by taking a 10 observable base subset and adding one observable at a time from an additional subset (Table A.3). These subsets are selected based on the studies presented earlier in this section. One of the best performing BDTs is constructed when the $\Delta\eta_{\ell_1 j_1}$ observable is added to the base subset. This BDT is compared with BDT-B in Table 3.16. The improvement is 2.1-3.7%, where the highest impact is from the poorest measured $p_T^H > 200$ GeV POI. Therefore, almost doubling the number of observables does not provide a significant enough improve-

Table 3.15: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with the BDTs trained on samples including and omitting the W+jets process. In both cases a set of BDT-B's discriminating parameters is used. Only statistical uncertainties are considered.

BDT type / STXS POI	BDT with W+jets	BDT without W+jets
1-jet, p_T^H ∈[0, 60] GeV	$1.00^{+0.56}_{-0.55}$	$1.00^{+0.56}_{-0.55}$
1-jet, p_T^H ∈[60, 120] GeV	$1.00^{+0.55}_{-0.53}$	$1.00^{+0.54}_{-0.53}$
1-jet, p_T^H ∈[120, 200] GeV	$1.00^{+1.03}_{-0.96}$	$1.00^{+1.02}_{-0.95}$
1-jet, p_T^H >200 GeV	$1.00^{+1.67}_{-1.38}$	$1.00^{+1.77}_{-1.50}$

ment to motivate the increase of the number of discriminating observables.

Table 3.16: The expected best-fit values and uncertainties for the indicated STXS signal strength POIs when events in 1-jet signal regions are classified with the best performing BDT constructed using a set of 11 observables and BDT-B constructed using six observables. Only statistical uncertainties are considered.

BDT constructed from / STXS POI	6 observables	11 observables
1-jet, p_T^H ∈[0, 60] GeV	$1.00^{+0.60}_{-0.58}$	$1.00^{+0.58}_{-0.57}$
1-jet, p_T^H ∈[60, 120] GeV	$1.00^{+0.56}_{-0.55}$	$1.00^{+0.55}_{-0.54}$
1-jet, p_T^H ∈[120, 200] GeV	$1.00^{+1.06}_{-0.98}$	$1.00^{+1.06}_{-0.99}$
1-jet, p_T^H >200 GeV	$1.00^{+1.86}_{-1.63}$	$1.00^{+1.8}_{-1.57}$

The two best performing BDTs constructed using subsets of six observables are BDT-A and BDT-B. The two BDTs are compared in Table 3.10. The POIs with $p_T^H < 120$ GeV perform similarly in both cases, the p_T^H ∈[120, 200] GeV POI is measured better with BDT-B, while the $p_T^H > 200$ GeV POI is measured better with BDT-A. Since the p_T^H ∈[120, 200] GeV POI is measured much better, BDT-B

is chosen.

In summary, the best selected BDT is the BDT trained using $\{m_{\ell\ell}$, m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta\phi_{\ell\ell,j}$, $\Delta\eta_{\ell_0 j}\}$ discriminating observables, which are selected in the study based on the shapes of background and signal distributions. The BDT is trained on the inclusive 1-jet signal region on a data set, which includes all background processes. The BDT's performance on the training set and the validation set is shown in Figure 3.20. The normalised background and signal distributions of discriminating observables are presented in Figure 3.21.

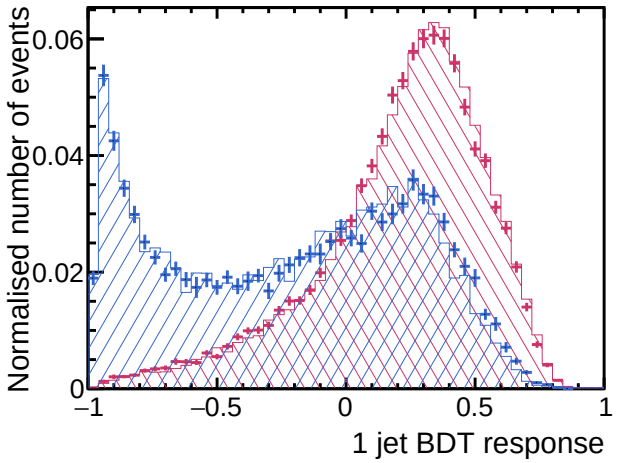


Figure 3.20: The selected BDT's for the 1-jet category performance on the training set (solid) and the validation set (dots with error bars). The total normalised background and signal distributions are in blue and red, respectively.

BDT construction for ⩾2-jet category

This section is dedicated to the BDT construction for the ⩾2-jet category. The BDT candidates are constructed using the following methods, discussed in section 3.4.2: the ranking method, evaluation of the optimal number of discriminating observables, and the effect of the correlations between the discriminating observables. All BDTs are trained on the inclusive ⩾2-jet signal region (3.3.2). Training on the signal

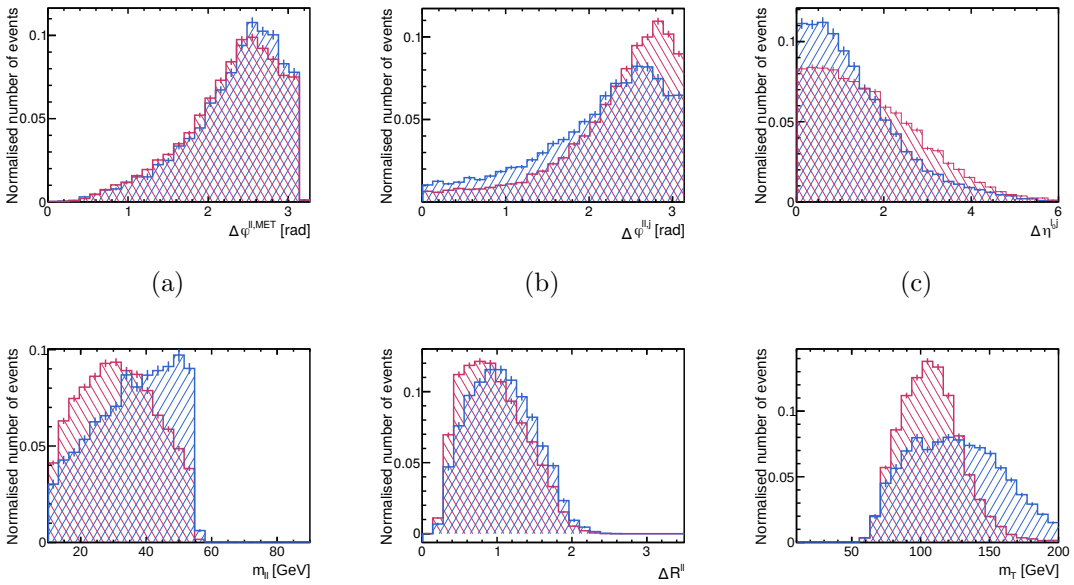


Figure 3.21: The normalised background (blue) and signal (red) distributions of the discriminating observables of the selected best BDT for the 1-jet category: (a) $\Delta\phi_{\ell\ell,E_T^{miss}}$, (b) $\Delta\phi_{\ell\ell,j}$, (c) $\Delta\eta_{\ell_0 j}$, (d) $m_{\ell\ell}$, (e) $\Delta R_{\ell\ell}$, (f) m_T.

subregions corresponding to the STXS kinematic regions is not performed based on the results in the 1-jet analysis showing that there is no significant improvement.

First, the BDTs are constructed using the ranking method with the χ^2 scheme. Five base sets of five observables are selected and the BDTs are constructed by adding the sixth observable from the additional subsets, summarised in Table A.4. The first two subsets consist of top five observables ranked in the inclusive $\geqslant 2$ jets SR using either all background processes or $Z \to \tau\tau$ process only; the last three subsets are constructed using normalised signal and background plots, the ranking in individual signal regions corresponding to the STXS kinematic definitions and taking potential correlations between the observables into account. The ranking obtained using only $Z \to \tau\tau$ sample is not promising. Based on the fit results, the best input subset is $\{\Delta R_{\ell\ell}, m_{\ell\ell}, \Delta\phi_{\ell\ell}, min\Delta R_{\ell_1 j_i}, m_T, E_T^{miss}\}$, let us denote this BDT as BDT-C.

Although the study in the 1-jet case showed that there is no significant improvement when more than six observables are used in the BDT training, the $\geqslant$2-jet case has more kinematic degrees of freedom. Thus more observables used in the BDT training may improve the results. This is tested by defining two base subsets of 16 and 23 observables, where the observables are selected based on the observable ranking method with the emphasis on the highest p_T^H region and avoiding highly correlated observables. BDT inputs are constructed by adding one observable per BDT from the additional subset to the base subset, where the subsets are summarised in Table A.5. There is no significant improvement when a higher number of observables is used. The impact of a larger number of discriminating observables is investigated by constructing BDTs by gradually adding parameters, starting with a six parameter set. The discriminating observables used for the best performing 16+1 BDT's (BDT^{best}_{16+1}, where "+1" corresponds to the $S\eta_{jj}$ observable, which is the sum of the pseudorapidities of two leading jets) training are used to construct subsets from 6 to 15 observables. First, this subset is reduced by two observables: $min\Delta R_{\ell_0 j_i}$ and Δy_{jj}, and then one at a time in the following order: p_T^H, $m_{\ell_0 j_0}$, $S\eta_{jj}$, p_T^{jj}, ΔR_{jj}, $\Delta\phi_{\ell\ell,E_T^{miss}}$, p_T^j, $m_{\ell_1 j_0}$, $p_T^{\ell_0}$. The performance of each BDT is presented in Table 3.17. There is a stable improvement for the $p_T^H \in$[0, 60] GeV POI starting with the set of 9 discriminating observables upwards, while uncertainties fluctuate up and down for the other POIs. A higher number of observables would increase the modelling uncertainties, therefore, an input set with six observables is concluded to be sufficient for the BDT construction.

The use of the highly correlated observables as inputs in the BDT training may contribute less to the discrimination power of the classifier. The observables with low mutual correlations are indentified by using a correlation matrix (Figure A.1) obtained by training BDT^{best}_{16+1}. The base and additional subsets are selected and listed in Table A.6. The BDTs are constructed by adding one or two observables per

Table 3.17: The expected best-fit uncertainties for the indicated STXS signal strength $\geqslant$ 2 jets POIs, when events in $\geqslant$2-jet signal regions are classified with the BDTs constructed using the subsets from 6 to 15 discriminating parameters. The starting subset of $\{m_{\ell\ell}, \Delta R_{\ell\ell}, m_T, min\Delta R_{\ell_1 j_i}, p_T^{\ell_1}, E_T^{miss}\}$ is supplemented by an observable one at a time in the following order: $p_T^{\ell_0}$, $m_{\ell_1 j_0}$, p_T^{j}, $\Delta\phi_{\ell\ell,E_T^{miss}}$, ΔR_{jj}, p_T^{jj}, $S\eta_{jj}$, $m_{\ell_0 j_0}$, p_T^{H}. The red (blue) colour indicates that this BDT has at least one uncertainty 0.01 (0.003) units smaller than the result obtained with the BDT constructed using six discriminating observables. Only statistical uncertainties are considered.

POI \ BDT type	6	7	8	9	10	11	12	13	14	15
[0, 60] GeV	$^{+2.23}_{-2.19}$	$^{+2.22}_{-2.18}$	$^{+2.23}_{-2.19}$	$^{+2.20}_{-2.16}$	$^{+2.21}_{-2.17}$	$^{+2.20}_{-2.16}$	$^{+2.20}_{-2.16}$	$^{+2.17}_{-2.13}$	$^{+2.16}_{-2.12}$	$^{+2.14}_{-2.10}$
[60, 120] GeV	$^{+1.14}_{-1.10}$	$^{+1.15}_{-1.11}$	$^{+1.16}_{-1.12}$	$^{+1.13}_{-1.10}$	$^{+1.13}_{-1.09}$	$^{+1.11}_{-1.08}$	$^{+1.12}_{-1.08}$	$^{+1.11}_{-1.07}$	$^{+1.17}_{-1.14}$	$^{+1.11}_{-1.08}$
[120, 200] GeV	$^{+0.95}_{-0.91}$	$^{+0.95}_{-0.91}$	$^{+0.93}_{-0.90}$	$^{+0.94}_{-0.90}$	$^{+0.93}_{-0.90}$	$^{+0.93}_{-0.90}$	$^{+0.93}_{-0.90}$	$^{+0.94}_{-0.90}$	$^{+0.93}_{-0.90}$	$^{+0.94}_{-0.90}$
>200 GeV	$^{+1.00}_{-0.94}$	$^{+1.00}_{-0.94}$	$^{+1.00}_{-0.95}$	$^{+0.98}_{-0.92}$	$^{+1.00}_{-0.95}$	$^{+1.01}_{-0.96}$	$^{+1.01}_{-0.96}$	$^{+1.00}_{-0.95}$	$^{+1.00}_{-0.95}$	$^{+1.01}_{-0.95}$

BDT. The best performing subsets are $\{S\eta_{jj}, min\Delta R_{\ell_1 j_i}, \Delta R_{\ell\ell}, \Delta R_{jj}, m_T, E_T^{miss}\}$ and $\{min\Delta R_{\ell_1 j_i}, \Delta R_{\ell\ell}, \Delta R_{jj}, m_T, E_T^{miss}\}$. More BDT variations are created by adding one observable per BDT from the additional subsets in Table A.7. The additional subsets are constructed based on the previous studies by observing the frequency with which the BDT performance improves. The results show that the subsets of observables with low correlations do not perform better than BDT-C.

Finally, the best BDT is constructed using the information from previous studies. The BDT is constructed by selecting $\{\Delta R_{\ell\ell}, m_T, E_T^{miss}\}$ as the base subset and adding three observables per BDT from the subset in Table A.8. The best performing BDT is constructed from $\{\Delta R_{\ell\ell}, m_T, E_T^{miss}, m_{\ell\ell}, min\Delta R_{\ell_1 j_i}, p_T^{\ell_1}\}$ observables. The BDT performance on the signal and background training and test samples is shown in Figure 3.22. The normalised background and signal distributions of the best performing BDT discriminating observables are presented in Figure 3.23.

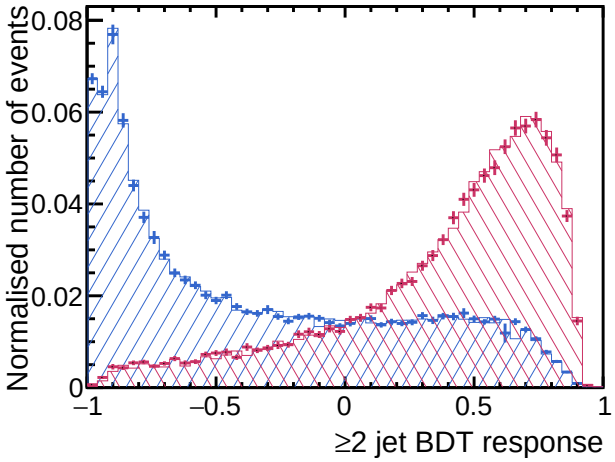
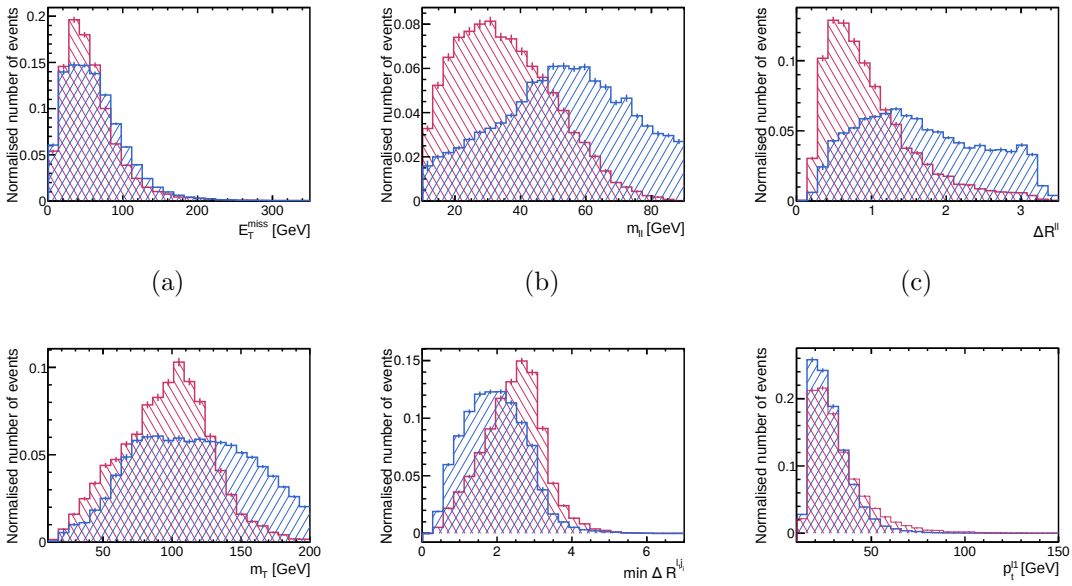


Figure 3.22: ⩾2-jet best BDT's performance on the training set (solid) and the validation set (dots with error bars). The total normalised background and signal distributions are in blue and red, respectively.



(a) (b) (c)

(d) (e) (f)

Figure 3.23: The normalised background (blue) and signal (red) distributions of the best performing BDT discriminating observables for the ⩾2-jet category: (a) E_T^{miss}, (b) $m_{\ell\ell}$, (c) $\Delta R_{\ell\ell}$, (d) m_T, (e) $min\Delta R_{\ell_1 j_i}$, (f) $p_T^{\ell_1}$.

Chapter 4

Measurement of Simplified Template Cross Sections in $H \to WW^*$

This chapter describes the measurement of the STXS cross-sections in the $H \to WW^*$ decay. It follows the previous chapter where the signal extraction is described. The following strategy is used: a binned maximum likelihood fit is applied to the discriminants introduced in chapter 3 to determine the signal cross-sections and background normalisations (section 4.1). The measurement includes estimation of experimental and theoretical uncertainties (section 4.3). The fit is optimisated based on the binning selection (section 4.2). The final result of the cross-section measurements is presented in section 4.4 and discussed in section 4.5.

4.1 Statistical analysis

The statistical analysis is based on a maximum likelihood fit. The likelihood function is constructed of probability density functions, which account for the uncertainties of all signal and background components. This analysis estimates data and MC statistical uncertainties, systematic uncertainties, and background normalisations obtained using the control regions. The likelihood function is maximised for the signal strength parameters μ_i, which are defined as a ratio of observed and SM

signal yields, where for the expected results with the SM prediction $\mu_i = 1$. There is one signal strength parameter per STXS cross-section.

The likelihood function is introduced in section 4.1.1 and the fit configuration is described in section 4.1.2.

4.1.1 The likelihood function

The likelihood function used in the fit is the following:

$$L = \Pi_i^{n_{\text{SR-bins}}} P(N_i|\lambda_i) \cdot \Pi_j^{n_{CR}} P(N_j|\lambda_j) \cdot \Pi_m^{syst} G(\tilde{\theta}_m|\theta_m, 1) \cdot \Pi_p^{n_{bins}} BBl_p(\eta_p|\Pi_k^{n_{bkg}} P(\tilde{\theta}_{pk}|\theta_{pk}\xi_{pk})) \tag{4.1}$$

The first term corresponds to data statistical uncertainties in the signal regions, the second term models the control regions, the third term models systematic uncertainties, while the last term accounts for the MC statistical uncertainties [55]. The first term in the likelihood function is iterated over all bins in all signal regions, while the second term is iterated over all control regions.

The data uncertainties are modelled with a Poisson distribution $P(N|\lambda) = \frac{\lambda^N e^{-\lambda}}{N!}$, where N and λ are the observed and the expected number of events, respectively. The expected number of events is estimated with the following equation:

$$\lambda = \Sigma_i^{STXS} \mu_i S_i + \Sigma_i^{\text{CR}} \beta_i B_i + \Sigma_i^{\text{other bkg}} B_i \tag{4.2}$$

where μ_i and S_i are the signal strength and the signal yield of the i^{th} STXS parameters of interest (POI); the second term is a sum of the background yields, which are modelled with dedicated control regions, from which the corresponding floating normalisation factors β_i are obtained. The third term is a sum of the remaining background yields.

The systematic uncertainties are modelled with a normalised Gaussian distribution $G(\tilde{\theta}|\theta, 1) = \frac{1}{\sqrt{2\pi}} \exp^{\frac{(\tilde{\theta}-\theta)^2}{2}}$, where $\tilde{\theta}$ corresponds to the central value of a measurement and θ is a nuisance parameter. The systematic uncertainties enter probability

density functions as multiplicative factors to signal and background yields, modelled with functions $\nu(\theta)$. There are two types of $\nu(\theta)$ functions: the first type models the scaling of the yields and the second type models the shape of classifier's distribution.

MC uncertainties are first modelled with Poisson functions and then processed with the Barlow-Beeston lite method [102]. The Poisson term $P(\tilde{\theta}_{nk}|\theta_{nk}\xi_{nk})$ models k^{th} background yield in n^{th} bin. The Barlow-Beeston lite method is an approximation applied to Poisson functions to reduce computational intensity. It is based on an assumption that there is only one floating parameter η per bin for all background processes, instead of assigning individual parameters for each process. The analytic function, which models the MC uncertainties, depends on fractional yield values of the background components. As a result, instead of $N \cdot K$ parameters θ_{nk} (where N is the number of bins and K is the number of background processes), there are N parameters η_n and the number of floating parameters is reduced K times.

4.1.2 Fit configuration for STXS analysis

The fit is maximised for the list of signal strength parameters. Most of the analysis optimisation studies are conducted by performing a fit for ten POIs. The first nine are the ggF ggF-like POIs, which correspond to the 0-jet, 1-jet and $\geqslant$2-jet STXS kinematic regions (section 1.3). The 10^{th} POI is a combination of VBF, ggF VBF-like POIs and the ggF POI corresponding to the Higgs boson signal in the forward detector region. The POIs are summarised in Figure 4.1.

The signal region categories and classifiers, which enter the likelihood fit, are described in section 3.3.3, and the control regions are defined in section 3.3.4.

4.2 Fit optimisation: binning

This analysis uses classifier functions for event redistribution within the SRs: the transverse mass m_T and BDT distributions. Consequently, the signal subregions are

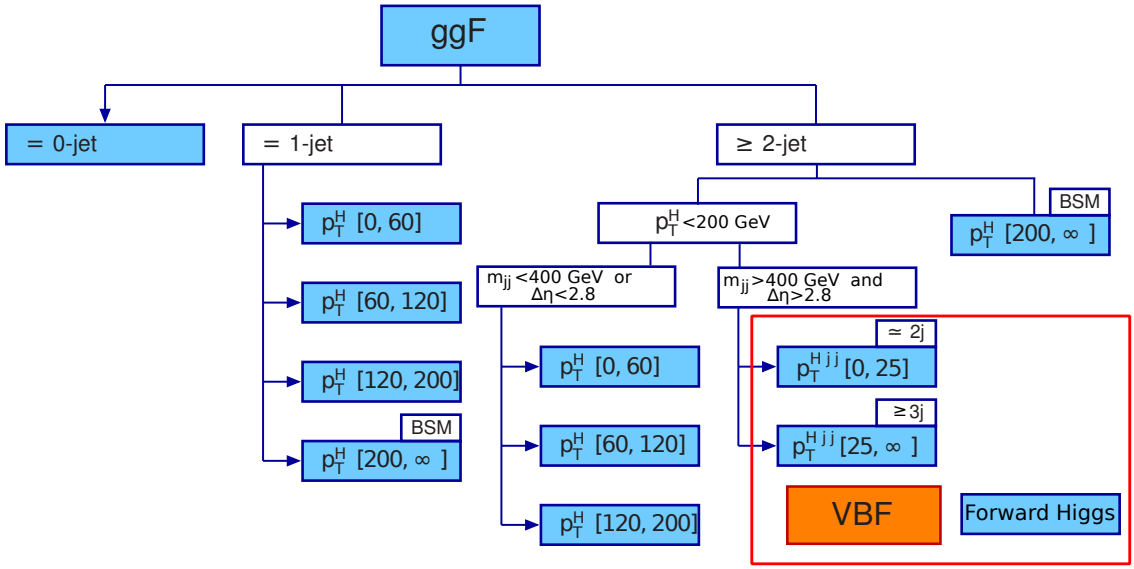


Figure 4.1: Parameters of interest measured using the likelihood fit. Red box indicates which STXS signals are merged to one POI.

further divided into smaller regions called bins, which is not a trivial task. Although a larger number of bins would access more features of a classifier, the number of bins is limited by the available data and MC statistics, and the estimation of systematic uncertainties. The following three binning methods are discussed: the flat signal binning, the telescope binning, and a method created as part of this thesis: the fit-driven binning (section 4.2.1). The methods are compared and the selected binnings for each of the signal regions are discussed in section 4.2.2.

4.2.1 Binning methods

This section discusses three binning methods. The choice of a binning method depends on a classifying discriminant. All three methods can be tuned for a preferred number of bins, given there are enough data and MC statistics. The procedure of number of bins and other specific aspects of each model, e.g. a choice of external constraints on the bins or internal parameters like step size, are analysis specific and are not discussed in this section.

Flat signal binning

The flat signal binning method divides signal regions into bins such that each bin has a uniform signal yield. This approach is applied in $H \to WW^*$ Run 1 [55] and Run 2 [2] analyses where the m_T distribution is fitted. However, it is not necessarily an optimal binning for BDT distributions, which is a motivation to explore other methods.

Telescope binning

The telescope binning is a method designed to bin BDT distributions [103]. The method defines an analytic function C:

$$C = \sqrt{2((s+b)\ln\frac{s+b}{b_0} - s - b + b_0)}\frac{|b-b_0|}{\Delta b} \tag{4.3}$$

where $b_0 \equiv 0.5b - \Delta b^2 + \sqrt{(b-\Delta b^2)^2 + 4(s+b)\Delta b^2}$, s and b are the total signal and total background yields in a bin, respectively, and Δb is a statistical uncertainty on the background yield.

The implementation of the telescope binning is the following: without loss of generality, let us consider a BDT distribution, defined in a range [-1, 1], where the signal distribution is located towards the positive values and the background events are concentrated towards the negative values. The binning procedure starts at 1 in the following manner: the right-most bin edge is fixed at BDT score equal to 1 and the left bin edge is moved towards the left step by step, where the step size is constant over the iterations and the choice depends on particular analysis. The C value is calculated at each step for the full region between the right and left edges. The process stops when C value decreases with respect to the previous step and the bin edge is placed where C value is largest. This edge becomes the new right-most edge and the process repeats. In addition, external constraints may be applied, e.g. a minimum signal yield in a bin. If the C value decreases but the external

constraints are not satisfied, the iteration continues until the external conditions are satisfied.

Fit-driven binning

If signal strength uncertainties could be analytically derived using a likelihood function, the best binning would be obtained by minimising these uncertainties. In case an analytic solution does not exist, the ambition to obtain the best binning is attempted by using the maximum likelihood fit itself. The idea of the method is that the fit will exploit the most important aspects of the statistical uncertainties and consequently the best binning will be obtained.

A standalone software is written for this purpose, which performs a simplified binned maximum likelihood fit for different binning configurations. The likelihood function used in the fit-driven technique is a simplified version of eq. 4.1 and has the following form:

$$\begin{aligned} L = & \Pi_i^{n_{bins}^{SR}} [P(N_i|\mu \cdot S_i \cdot \theta_i^s + (\Sigma_k^{\text{CRs}} \beta_k \cdot B_{ki} + \Sigma_k^{\text{other bkg}} B_{ki}) \cdot \theta_i^b) \\ & \Pi_l^{n^{CRs}} P(N_l|\mu \cdot S + \Sigma_k^{\text{CRs}} \beta_k \cdot B_{ki} + \Sigma_k^{\text{other bkg}} B_{ki}) \\ & \Pi_k^{n_{sig,bkg},n_{bins}} P(N_{ki}^{MC}|N_{ki}^{MC}\theta_i^k)] \end{aligned} \tag{4.4}$$

The first term models data statistical uncertainties, the second term models uncertainty associated with the profiled control regions, while the last term accounts for the MC statistical uncertainties. This likelihood function does not include the modelling of systematic uncertainties and does not use Barlow-Beeston lite method since the setup is not computationally intensive. The number of expected events is defined as:

$$\lambda = \mu \cdot S + \Sigma_k^{\text{CRs}} \beta_k \cdot B_{ki} + \Sigma_k^{\text{other bkg}} B_{ki} \tag{4.5}$$

where the first term consists of the SM total signal yield S scaled by a signal strength parameter μ, the second term is the sum of the background yields B_k which are scaled by normalisation factors β_k obtained from the control regions, and the last

term corresponds to the remaining background processes. n_{bins}^{SR} is the total number of bins in the signal region, n^{CRs} is the total number of control regions. N_i^{MC} is the number of signal or background MC events in i^{th} bin, $n_{sig,bkg}$ corresponds to the iteration through signal and background processes, and n_{bins} is the number of bins in the SR and CRs.

Ideally, the iteration should be done over all binning configurations, which effectively is a merging of bins. To merge a histogram with N internal bin edges (i.e. $N+1$ bins) into a histogram with B internal bin edges ($B+1$ bins) there are $\frac{N!}{(N-B)!B!}$ possibilities. To merge a histogram with 100 bins into 6 bins requires $\approx 7 \cdot 10^7$ variations, which is computationally lengthy. The computational intensity is significantly reduced when one bin edge is fixed in the middle of a histogram and each half is scanned separately, then the total number of scans is $2\frac{(\frac{1}{2}N-\frac{1}{2})!}{(\frac{1}{2}N-\frac{1}{2}-B)!B!}$. In this case to merge a histogram with 100 bins into 6 bins requires only $\approx 2 \cdot 10^3$ variations.

The fit-driven binning likelihood function, which includes statistical uncertainties only, can be extended to a likelihood function with systematic uncertainties included. Including all systematic uncertainties might make the technique computationally challenging but it could be simplified by including only the leading systematic uncertainties.

4.2.2 Binning on the signal regions

This section discusses the comparison of the three binning methods, which is followed by the choice of the binning for this STXS analysis.

Comparison of binning methods

The performance of three binning methods is evaluated on the 1-jet SR, classified with a BDT. The performance is quantified using a likelihood fit for ten STXS POIs,

as described in section 4.1.2. Lower uncertainties on the 1-jet POIs correspond to a better binning performance.

In the case of the telescope and fit-driven binning, the total signal and background yields in a bin are required to be larger than 1 event. Only statistical uncertainties are taken into account for this comparison. A comparison between the flat signal binning and the telescope binning is presented in Table 4.1. For all 1-jet but the p_T^H ∈[60, 120] GeV POI the use of the telescope binning improves the results by 0.9-5.3%. The flat signal binning and the fit-driven binning methods are

Table 4.1: Comparison of the best-fit values and uncertainties, where a BDT is binned with the flat signal binning and the telescope binning in 1-jet signal region. POIs which are measured better by at least 0.01 units with the telescope binning are in red. Only statistical uncertainties are considered.

POI \ Binning	BDT flat signal binning	BDT telescope binning
1-jet, p_T^H ∈[0, 60] GeV	$1.00^{+0.59}_{-0.57}$	$1.00^{+0.57}_{-0.55}$
1-jet, p_T^H ∈[60, 120] GeV	$1.00^{+0.56}_{-0.56}$	$1.00^{+0.58}_{-0.57}$
1-jet, p_T^H ∈[120, 200] GeV	$1.00^{+1.07}_{-1.00}$	$1.00^{+1.06}_{-1.01}$
1-jet, p_T^H >200 GeV	$1.00^{+1.88}_{-1.66}$	$1.00^{+1.78}_{-1.61}$

compared in Table 4.2. All POIs' uncertainties are reduced by 1.7-5.4%. Similar results are observed when the study is done on the ⩾2-jet signal region. When the fit-driven method is compared with the telescope binning method in Table 4.3, the 1-jet p_T^H ∈[0, 60] GeV performs better using the telescope binning. When the best binning from both methods is applied on 1-jet region, the results do not improve (Table 4.3). The likely reason is due to the correlations between the POIs.

Table 4.2: Comparison of the best-fit values and uncertainties, where a BDT is binned with the flat signal binning and the fit-driven binning. POIs which are measured better with the fit-driven binning method by at least 0.01 units are written in red. Only statistical uncertainties are considered.

Binning / POI	BDT flat signal binning	BDT fit-driven binning
1-jet, p_T^H ∈[0, 60] GeV	$1.00^{+0.59}_{-0.57}$	$1.00^{+0.58}_{-0.56}$
1-jet, p_T^H ∈[60, 120] GeV	$1.00^{+0.56}_{-0.56}$	$1.00^{+0.55}_{-0.54}$
1-jet, p_T^H ∈[120, 200] GeV	$1.00^{+1.07}_{-1.00}$	$1.00^{+1.03}_{-0.96}$
1-jet, p_T^H >200 GeV	$1.00^{+1.88}_{-1.66}$	$1.00^{+1.80}_{-1.57}$

Table 4.3: Comparison of the best-fit values and uncertainties, where a BDT is binned with the telescope binning, the fit-driven binning, and their best combination. Only statistical uncertainties are considered here.

Binning / POI	BDT telescope binning	BDT fit-driven binning	BDT mixed telescope and fit-driven binning
1-jet, p_T^H ∈[0, 60] GeV	$1.00^{+0.57}_{-0.55}$	$1.00^{+0.58}_{-0.56}$	$1.00^{+0.58}_{-0.56}$
1-jet, p_T^H ∈[60, 120] GeV	$1.00^{+0.58}_{-0.57}$	$1.00^{+0.55}_{-0.54}$	$1.00^{+0.55}_{-0.54}$
1-jet, p_T^H ∈[120, 200] GeV	$1.00^{+1.06}_{-1.01}$	$1.00^{+1.03}_{-0.96}$	$1.00^{+1.03}_{-0.96}$
1-jet, p_T^H >200 GeV	$1.00^{+1.78}_{-1.61}$	$1.00^{+1.80}_{-1.57}$	$1.00^{+1.78}_{-1.61}$

Binning for the STXS analysis

In the 0-jet signal regions events are classified using the m_T distribution. Based on [2], each 0-jet signal region, defined in section 3.3.3, is divided into eight bins using the flat signal binning method.

The discriminants in the 1-jet and ⩾2-jet SRs are BDT distributions and the study in section 4.2 shows that the fit-driven method provides the best binning. Thus the fit-driven method is applied. The middle point is chosen at several places

(around 10 cases in range [-0.4, 0.7]), and the iteration is run on the right hand side, as described in section 4.2.1. The output binning configurations are then tested with the full $H \to WW^*$ analysis setup. All binning variations are required to satisfy the criterion for the sum of yields in each bin to be greater than 1. The number of bins in each region is deduced based on MC statistics of the signal and background processes, e.g. it is preferred that W+jets background's yield is positive in each bin and other backgrounds' yields are not negative. Moreover, different numbers of bins in each signal region are tested with the fit and based on the results the optimal number of bins is deduced in each case. The final number of bins chosen for each signal subregion (section 3.3.3) is summarised in Table 4.4.

Table 4.4: Number of bins in each of the signal subregions defined in section 3.3.3. Subregions with the same number of bins are grouped into categories in the left column.

Category of signal subregions	Number of bin per signal subregion
0-jet	8
1-jet, p_T^H <120 GeV	6
1-jet, p_T^H ∈[120, 200] GeV	4
⩾2-jet, p_T^H <200 GeV	6
⩾1-jet, p_T^H >200 GeV	5
VBF-like	2

The VBF-like signal region is split into two bins. Since in this case there is only one internal bin edge, all variations are tested instead of using the fit-driven approach and the best performing configuration is chosen.

In order to check that the binning preserves the modelling of the background processes, the selected binning is applied to the corresponding control regions. A good data to MC agreement is observed, which indicates that these background processes are modelled well when the chosen binning is used.

4.3 Systematic uncertainties

Systematic uncertainties can be classified into two categories: experimental and theoretical. The former category concerns the sources related to the object reconstruction and identification, while the latter is mostly related to the simulation of the events. This section discusses the sources of theoretical uncertainties and their derivation in section 4.3.1, while experimental uncertainties are reviewed in section 4.3.2.

4.3.1 Theoretical uncertainties

This section discusses theoretical uncertainties on the ggF signal process as well as the dominant background processes, WW, top quark, and $Z \to \tau\tau$, which are evaluated specially for this STXS measurement. There are two types of uncertainties evaluated: on the yields and on the shapes of the classifiers' distributions. These uncertainties are estimated by comparing the nominal MC samples (section 3.2.1) to the "alternative" samples, where the "alternative" corresponds to a sample, which aims to represent identical physics but is produced with a different tool or configuration.

Uncertainties on yields are derived differently for signal and background processes and discussed in dedicated sections. The shape uncertainties are evaluated by comparing normalised histograms obtained from the alternative samples with the nominal normalised histogram. If only one alternative sample is used, both the uncertainty on yields and the shape uncertainty are symmetrised with respect to the nominal values in order to estimate the higher and lower uncertainties. In the cases where more than two alternative samples are considered (except for the PDF uncertainties), the highest and lowest deviations from the nominal sample are defined as the higher and the lower uncertainties, respectively.

ggF signal theoretical uncertainties

There are four theoretical uncertainties derived for the signal samples: QCD scale, PDF (section 1.4.1), parton shower (section 1.4.2), and generator (section 1.4.3) uncertainties. The uncertainties are derived for the 0-jet, 1-jet and $\geqslant$2-jet signal regions. The uncertainty on ggF signal yield corresponds to an acceptance uncertainty, where acceptance is defined as a fraction of events in a measured region with respect to the region this number is extrapolated to.

The uncertainty on acceptance is defined as a difference in acceptances from the nominal sample (except for the PDF uncertainty). The ggF signal sample is split into STXS categories, so the acceptance uncertainties are evaluated for each STXS category (section 1.3) and each signal region (section 3.3.2). The QCD scale uncertainties are determined by variations in the scale choice for the renormalisation and factorisation (section 1.4.1) in the nominal POWHEG generation and the NNLO calculation used for the NNLOPS sample. The alternative samples are generated with renormalisation and factorisation scale values set to 1/2, 1 (nominal), and 2. As a result, 26 alternative samples are used. The PDF acceptance uncertainties are evaluated by adding in quadrature the 68% C.L. PDF4LHC Hessian PDF [62] eigenvector deviations from the nominal sample. In addition, the nominal PDF set is compared to CT10 [83], MMHT14 [104], and NNPDF3.0 [79] PDF sets. An uncertainty is assigned on the choice of the ME generator by comparing the nominal POWHEG NNLOPS sample to an alternative MadGraph5_aMC@NLO [105, 106] sample. Only the NLO prediction by POWHEG is used to match the order in QCD of the alternative sample. Parton shower acceptance uncertainty is estimated by comparing the nominal POWHEG-Box v2 NNLOPS interfaced with PYTHIA 8 sample to POWHEG-Box v2 NNLOPS interfaced with Herwig 7 [107] using UEEE5 tune [108]. Parton shower uncertainty is evaluated on particle level only. Typical ggF signal theoretical uncertainties are illustrated on 1-jet $p_T^H \in$[0, 60] GeV and

$p_T^H > 200$ GeV regions in Table 4.5 and Table 4.6, respectively.

Table 4.5: ggF signal theoretical uncertainties evaluated for the STXS ggF 1-jet $p_T^H \in [0, 60]$ GeV signal category. 0% is replaced by "-" for readability. The uncertainties are statistically significant, except the generator uncertainty for the $p_T^H \in [0, 60]$ GeV $e\mu$ SR.

Uncertainty / Measured SR	QCD scale	PDF	Generator	Parton shower
1-jet, $p_T^H \in [0, 60]$ GeV μe	-0.4% / 0.6%	± 0.11%	± 4.7%	± 3.9%
1-jet, $p_T^H \in [0, 60]$ GeV $e\mu$	-0.54% / 0.98%	± 0.12%	± 1.4%	± 3.9%
1-jet, $p_T^H \in [60, 120]$ GeV μe	-4.5% / 2.6%	± 0.36%	± 14%	-
1-jet, $p_T^H \in [60, 120]$ GeV $e\mu$	-3.7% / 2.3%	± 0.45%	± 8.4%	-
1-jet, $p_T^H \in [120, 200]$ GeV μe	-6% / 4.7%	± 3.2%	-	-
1-jet, $p_T^H \in [120, 200]$ GeV $e\mu$	-	-	-	-
1-jet, $p_T^H > 200$ GeV	-	-	-	-

Table 4.6: ggF signal theoretical uncertainties evaluated for the STXS ggF 1-jet $p_T^H > 200$ GeV signal category. 0% is replaced by "-" for readability. The uncertainties are statistically significant, except the generator uncertainty for the $p_T^H > 200$ GeV SR.

Uncertainty / Measured SR	QCD scale	PDF	Generator	Parton shower
1-jet, $p_T^H \in [0, 60]$ GeV μe	-	-	-	-
1-jet, $p_T^H \in [0, 60]$ GeV $e\mu$	-	-	-	-
1-jet, $p_T^H \in [60, 120]$ GeV μe	-	-	-	-
1-jet, $p_T^H \in [60, 120]$ GeV $e\mu$	-	-	-	-
1-jet, $p_T^H \in [120, 200]$ GeV μe	-1.3% / 1.3%	± 1%	± 54%	-
1-jet, $p_T^H \in [120, 200]$ GeV $e\mu$	-1.4% / 1.4%	± 0.89%	± 20%	-
1-jet, $p_T^H > 200$ GeV	-0.28% / 0.39%	± 0.18%	± 2.7%	± 3.6%

Shape uncertainties are considered only in the dominant STXS category per

signal subregion. The parton shower shape uncertainties are evaluated at particle level and at particle level the SRs coincide with the corresponding STXS regions, so there is only one STXS category per SR. Moreover, the transverse mass m_T distribution is used as a classifier on the 0-jet SRs, which at particle level falls rapidly at $m_T = 125$ GeV (around Higgs boson mass). A more accurate modelling of 0-jet parton shower uncertainties is obtained by using a folding matrix from the particle level m_T to the reconstructed m_T. This matrix is obtained using the nominal POWHEG-Box v2 NNLOPS MC sample. Shape uncertainties due to PDF set variations were found to be negligible.

WW theoretical uncertainties

There are five sets of theoretical uncertainties derived for the $qq \to WW$ samples: QCD scale, PDF, parton shower, matrix element matching scale (CKKW) (section 1.4.2), and α_s (section 1.4.1) uncertainties. The uncertainties for the 0-jet category are adopted from [2]. The uncertainties on the 1-jet and $\geqslant$2-jet signal regions and the 1-jet WW control region are derived as part of this thesis. These $qq \to WW$ (denoted as WW) uncertainties are estimated at particle level.

The uncertainty on yields corresponds to the difference of selected number of events in a measured region using an alternative sample from the value obtained from the nominal sample. The QCD scale uncertainty is derived by varying renormalisation and factorisation scales by setting them to 1/2, 1 (nominal), and 2 values. The alternative 1-jet samples are generated using SHERPA 2.2.2 [76, 77], while $\geqslant$2-jet samples are generated with MadGraph5_aMC@NLO [105, 106]. For the evaluation of PDF uncertainties, one or more jets samples are generated with MadGraph5_aMC@NLO generator. The PDF uncertainties are estimated using NNPDF3.0 [79] and its variations. The uncertainty is derived by calculating the standard deviation of these variations. Parton shower uncertainty is estimated by

comparing the nominal samples generated with POWHEG-Box v2 interfaced with PYTHIA 8 with the samples obtained with POWHEG-Box v2 interfaced with Herwig++ [107] generator. The matrix element matching scale (CKKW) uncertainty is estimated by comparing the nominal sample where the matching scale is set to 20 GeV to two samples where the matching scale is set to 15 GeV and 30 GeV, respectively. The samples are generated using SHERPA 2.2.2 [76, 77]. The α_s uncertainty is estimated by comparing different PDF sets in SHERPA 2.2.2 with NNPDF3.0 NNLO [79]. The nominal sample has $\alpha_s(m_Z) = 0.118$ and is compared to two samples, which have $\alpha_s(m_Z)$ values set to 0.117 and 0.119, respectively. The WW theoretical uncertainties for the 1-jet category are summarised in Table 4.7, while $\geqslant$2-jet WW theoretical uncertainties are summarised in Table 4.8.

Table 4.7: WW theoretical uncertainties for the 1-jet category. The uncertainties are statistically significant.

Measured region \ Uncertainty	QCD scale	PDF	Parton shower	CKKW	α_s
1j SR, [0, 60] GeV $\mu e/e\mu$	-1.4% / 1.7%	± 0.43%	± 13%	-3.1% / 1.2%	-0.22% / 0.57%
1j SR, [60, 120] GeV $\mu e/e\mu$	-4.7% / 4.3%	± 0.5%	± 8.6%	-1.4% / 0%	-0.35% / 0.58%
1j SR, [120, 200] GeV $\mu e/e\mu$	-5.9% / 5.3%	± 0.56%	± 15%	-4.8% / 0%	-0.27% / 0.47%
1j SR, >200 GeV	-7.7% / 7.7%	± 0.9%	± 24%	-6.1% / 9.6%	-0.35% / 0.48%
1j CR, WW	-3.8% / 3.8%	± 0.48%	± 13%	-0.92% / 2.2%	-0.066% / 0.33%

The shape uncertainties are derived for all five sources of uncertainties. The $gg \to WW$ uncertainties are adopted from [2].

Top quark theoretical uncertainties

The following uncertainties are estimated on the $t\bar{t}$ and Wt background processes: parton shower, radiation, generator, and PDF. In addition, the interference between $t\bar{t}$ and Wt uncertainty is estimated on the Wt background process. The uncertainties for the 0-jet category are adopted from [2]. The estimation is performed on

Table 4.8: WW theoretical uncertainties for the $\geqslant$2-jet category. The uncertainties are statistically significant.

Uncertainty / Measured SR	QCD scale	PDF	Parton shower	CKKW	α_s
$\geqslant$2j [0, 60] GeV $\mu e/e\mu$	-13% / 13%	± 0.63%	± 11%	-4.2% / 5.1%	-1.1% / 1.3%
$\geqslant$2j [60, 120] GeV $\mu e/e\mu$	-7.4% / 6.9%	± 0.62%	± 12%	-1.6% / 3.6%	-1.2% / 1.4%
$\geqslant$2j [120, 200] GeV $\mu e/e\mu$	-7.6% / 7.3%	± 0.64%	± 5.4%	-2.2% / 3.3%	-1.2% / 1.3%
$\geqslant$2j >200 GeV	-7.1% / 7.3%	± 0.86%	± 4.8%	-1.7% / 2.7%	-1.2% / 1.2%
VBF-like	-11% / 10%	± 0.72%	± 9%	-7.9% / 14%	-1.2% / 1.2%

the 1-jet and $\geqslant$2-jet SRs and corresponding CRs. Due to limited MC statistics, the signal regions split based on the leading lepton flavour (section 3.3.3) are assigned uncertainties, which are evaluated on the region before the split. The shape uncertainties are not evaluated due to limited available MC statistics in the alternative samples.

The uncertainties on yields are performed in the same way as in the WW case. The $t\bar{t}$ parton shower uncertainties are evaluated by comparing the nominal POWHEG-Box v2 [85] interfaced with PYTHIA 8 samples with POWHEG-Box v2 interfaced with Herwig 7 [107] samples. The Wt parton shower uncertainties are evaluated by comparing POWHEG-Box v1 [87] interfaced with PYTHIA 6.428 samples with samples generated with POWHEG-Box v1 interfaced with Herwig++ [107]. The interference between $t\bar{t}$ and Wt uncertainty is evaluated on Wt background sample by comparing two samples where different techniques are used to remove the overlapping diagrams [109]. The $t\bar{t}$ radiation uncertainty is evaluated by comparing sample variations generated with POWHEG-Box v2 interfaced with PYTHIA 8 with different generator weights. The Wt radiation uncertainty is evaluated by comparing samples generated with POWHEG-Box v1 interfaced with PYTHIA 6 with different generator weights. The $t\bar{t}$ and Wt PDF uncertainties are defined as the standard deviation of variations in alternative samples from

the nominal sample (section 3.2.1). In the case of $t\bar{t}$, the PDF variations of the nominal NNPDF3.0NLO are used, while in the case of Wt, the nominal CT10 is compared to the NNPDF3.0. The generator uncertainty on the $t\bar{t}$ background process is evaluated by comparing the nominal samples generated with POWHEG-Box v2 interfaced with PYTHIA 8 with the alternative samples generated with SHERPA 2.2.1 [71]. The Wt generator uncertainty is derived by comparing samples generated with POWHEG-Box v1 interfaced with Herwig++ with samples generated with MadGraph5_aMC@NLO interfaced with Herwig++.

The 1-jet $p_T^H > 200$ GeV signal region has large $t\bar{t}$ parton shower and generator uncertainties. The issue is addressed by increasing MC statistics. The increase of MC statistics by removing selection requirements on $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ observables does not improve the estimation of uncertainties. The uncertainties are estimated on a merged $p_T^H > 120$ GeV signal region since the two merged regions have similar kinematic properties. Theoretical uncertainties on $t\bar{t}$ background process in 1-jet and $\geqslant$2-jet regions are summarised in Table 4.9 and Table 4.10, respectively. Table 4.11 and Table 4.12 present theoretical uncertainties on Wt background process in 1-jet and $\geqslant$2-jet regions, respectively.

$Z \to \tau\tau$ theoretical uncertainties

The uncertainties on $Z \to \tau\tau$ background process are determined for the 1-jet and $\geqslant$2-jet signal and $Z \to \tau\tau$ control regions. The uncertainties on the 0-jet category are adopted from [2]. The same procedure is used as in the WW and top quark cases. The $Z \to \tau\tau$ modelling is evaluated by estimating generator uncertainties. These theoretical uncertainties are determined by comparing the nominal SHERPA 2.2.1 sample with the sample generated with MadGraph5_aMC@NLO interfaced with PYTHIA 8. Due to limited MC statistics the uncertainties on the 1-jet and $\geqslant$2-jet signal subregions are evaluated on the corresponding inclusive SRs and are

Table 4.9: $t\bar{t}$ theoretical uncertainties in 1-jet category. The uncertainties are statistically significant, except the generator uncertainty for the [60, 120] GeV SR and the $Z \to \tau\tau$ CR.

Uncertainty / Measured region	Parton shower	Radiation	Generator	PDF
1j SR, [0, 60] GeV $\mu e/e\mu$	± 6.6%	± 4.4%	± 3.3%	± 0.65%
1j SR, [60, 120] GeV $\mu e/e\mu$	± 4.1%	± 2.3%	± 0.14%	± 0.54%
1j SR, [120, 200] GeV $\mu e/e\mu$	± 8.2%	± 11%	± 23%	± 0.43%
1j SR, >200 GeV	± 1.8%	± 11%	± 16%	± 0.68%
1j CR, top quark	± 0.7%	± 2.8%	± 3.1%	± 0.59%
1j CR, WW	± 7.5%	± 2.5%	± 2.2%	± 0.48%
1j CR, $Z \to \tau\tau$	± 5.5%	± 4%	± 1.2%	± 1.1%

Table 4.10: $t\bar{t}$ theoretical uncertainties in $\geqslant$2-jet category. The uncertainties are statistically significant, except the radiation uncertainty for the [0, 60] GeV SR and the top quark CR.

Uncertainty / Measured region	Parton shower	Radiation	Generator	PDF
$\geqslant$2j SR, [0, 60] GeV $\mu e/e\mu$	± 12%	± 0.13%	± 7%	± 1.1%
$\geqslant$2j SR, [60, 120] GeV $\mu e/e\mu$	± 8.9%	± 1.9%	± 2.1%	± 0.92%
$\geqslant$2j SR, [120, 200] GeV $\mu e/e\mu$	± 11%	± 3.2%	± 9.5%	± 0.7%
$\geqslant$2j SR, >200 GeV	± 16%	± 8%	± 4.7%	± 1%
VBF-like SR	± 12%	± 8.9%	± 16%	± 1.9%
$\geqslant$2j CR, $Z \to \tau\tau$	± 13%	± 6.2%	± 1.6%	± 1%
$\geqslant$2j CR, top quark	± 5.2%	± 0.62%	± 2.1%	± 0.36%

assigned to the corresponding signal subregions. The shape uncertainties are not significant compared to the MC statistical uncertainties of the $Z \to \tau\tau$ sample. The $Z \to \tau\tau$ theoretical uncertainties on 1-jet and $\geqslant$2-jet regions are summarised in

Table 4.11: Wt theoretical uncertainties in 1-jet category. The uncertainties are statistically significant, except the parton shower uncertainty for the [0, 60] GeV SR and the WW CR, the generator uncertainty for the $Z \to \tau\tau$ CR, and the interference uncertainty for the [0, 60] GeV SR.

Uncertainty / Measured region	Parton shower	Radiation	Generator	Interference	PDF
1j SR, [0, 60] GeV $\mu e/e\mu$	± 0.45%	± 2.6%	± 4.6%	± 0.13%	± 0.38%
1j SR, [60, 120] GeV $\mu e/e\mu$	± 2.9%	± 3.5%	± 9.5%	± 7.2%	± 0.38%
1j SR, [120, 200] GeV $\mu e/e\mu$	± 19%	± 6.2%	± 11%	± 12%	± 0.38%
1j SR, >200 GeV	± 11%	± 25%	± 37%	± 1.3%	± 0.3%
1j CR, top quark	± 2.8%	± 1.1%	± 1.8%	± 3.2%	± 0.74%
1j CR, WW	± 0.072%	± 3.9%	± 7.3%	± 2.2%	± 0.19%
1j CR, $Z \to \tau\tau$	± 9.4%	± 2.7%	± 0.1%	± 1.9%	± 0.24%

Table 4.12: Wt theoretical uncertainties in $\geqslant$2-jet category. The uncertainties are statistically significant, except the parton shower uncertainty for the VBF-like SR and the generator uncertainty for the [0, 60] GeV SR.

Uncertainty / Measured region	Parton shower	Radiation	Generator	Interference	PDF
$\geqslant$2j SR, [0, 60] GeV $\mu e/e\mu$	± 18%	± 2.7%	± 1.3%	± 9.9%	± 0.33%
$\geqslant$2j SR, [60, 120] GeV $\mu e/e\mu$	± 13%	± 4.3%	± 9.4%	± 11%	± 0.35%
$\geqslant$2j SR, [120, 200] GeV $\mu e/e\mu$	± 7.2%	± 3.3%	± 21%	± 29%	± 0.38%
$\geqslant$2j SR, >200 GeV	± 24%	± 7.3%	± 26%	± 2%	± 0.25%
VBF-like	± 0.73%	± 5.3%	± 15%	± 4.1%	± 0.38%
$\geqslant$2j CR, $Z \to \tau\tau$	± 5.5%	± 6%	± 24%	± 3%	± 0.16%
$\geqslant$2j CR, top quark	± 1.2%	± 6.1%	± 9.2%	± 2.4%	± 0.77%

Table 4.13.

4.3.2 Experimental uncertainties

Experimental uncertainties concern the reconstructed objects that enter the analysis rather than particular signal and background processes, as in the case of theoretical uncertainties. Since this analysis contains electrons, muons, the missing energy, untagged and b-tagged jets, the following sources are the main contributors to the

Table 4.13: $Z \to \tau\tau$ theoretical uncertainties in 1-jet and $\geqslant$2-jet categories. The uncertainties are statistically significant.

Region	Generator
1-jet SR	± 16%
1-jet CR, $Z \to \tau\tau$	± 1.4%
$\geqslant$2-jet SR	± 9.7%
$\geqslant$2-jet CR, $Z \to \tau\tau$	± 6.9%

experimental uncertainties: charged leptons' energy and momentum scale, resolution, identification and isolation; measurement of missing transverse momentum; jet energy scale, resolution, and b-tagging efficiency; modelling of pile-up and luminosity measurement.

The electron energy scale and resolution calibration and uncertainties are obtained from a comparison of data and simulation of $Z \to ee$ events [110]. The calibration and uncertainties on electron reconstruction, identification, trigger and isolation are obtained by adjusting the simulated predictions to the observed data using $Z \to ee$ and $J/\Psi \to ee$ decays [40]. Muon energy scale and resolution calibration and uncertainties are obtained from the comparison of the $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ data samples with the MC predictions [45]. $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ decays are also used to estimate muon reconstruction and isolation efficiencies and corresponding uncertainties [45].

The main uncertainties associated to jets arise from the jet energy scale (JES) and the jet energy resolution (JER). The JES uncertainties are measured as a function of jet transverse momentum (p_T) and a modulus of pseudorapidity ($|\eta|$) [46]. The JER uncertainties are obtained by varying the transverse momenta of the simulated jets based on their p_T and $|\eta|$. More details can be found in [111].

The calibration and uncertainties associated to the b-tagging performance are

calculated using simulation and observed data comparison and the resulting scale factors [112]. The missing transverse momentum uncertainty contains uncertainties due to smearing of energies and momenta of charged leptons and jets. More details about missing transverse momementum uncertainties can be found in [54]. The uncertainty on the integrated luminosity depends on the data set: for the data set collected in 2015 the uncertainty was measured to be $\pm 2.1\%$, while for the 2016 data set the uncertainty is $\pm 2.2\%$ [113].

4.4 Results

This section presents the results of the $H \to WW^*$ Simplified Template Cross Sections measurement. First, the results are presented for ten POIs, defined in section 4.1.2. A merging scheme is introduced to produce a smaller set of measurements than the set defined in the STXS framework. The expected results with the SM prediction are presented in section 4.4.1. The observed results are discussed in section 4.4.2.

4.4.1 Expected results

The expected results of the measurement of ten signal strength parameters with the SM prediction are summarised in Table 4.14. The table presents the results with statistical uncertainties only (data and MC statistical uncertainties and the uncertainties on CR normalisations) and the results with the total uncertainties (the statistical and systematic uncertainties). In this analysis a measurement is defined to be significant if the total uncertainty is not significantly higher than 100%. The 1-jet $p_T^H > 200$ GeV and $\geqslant$2-jet $p_T^H < 120$ GeV POIs have uncertainties significantly higher than 100%, therefore a merging scheme of POIs must be designed to obtain significant measurements.

Correlations between the POIs are given in Figure 4.6. The anti-correlations

Table 4.14: The expected best-fit values and uncertainties of a simultaneous fit for ten STXS signal strength POIs with the SM prediction. The middle column shows the results with statistical uncertainties only, while the results with the total uncertainties are presented in the last column.

POI	Statistical uncertainties only	Total uncertainties
0-jet	$1.00^{+0.16}_{-0.16}$	$1.00^{+0.23}_{-0.22}$
1-jet, p_T^H $\in$[0, 60] GeV	$1.00^{+0.58}_{-0.56}$	$1.00^{+0.72}_{-0.71}$
1-jet, p_T^H $\in$[60, 120] GeV	$1.00^{+0.55}_{-0.54}$	$1.00^{+0.62}_{-0.60}$
1-jet, p_T^H $\in$[120, 200] GeV	$1.00^{+1.03}_{-0.96}$	$1.00^{+1.12}_{-1.03}$
1-jet, p_T^H >200 GeV	$1.00^{+1.80}_{-1.57}$	$1.00^{+2.02}_{-1.80}$
$\geqslant$2-jet, p_T^H $\in$[0, 60] GeV	$1.00^{+2.13}_{-2.10}$	$1.00^{+2.76}_{-2.66}$
$\geqslant$2-jet, p_T^H $\in$[60, 120] GeV	$1.00^{+1.13}_{-1.09}$	$1.00^{+1.34}_{-1.28}$
$\geqslant$2-jet, p_T^H $\in$[120, 200] GeV	$1.00^{+0.93}_{-0.89}$	$1.00^{+1.04}_{-0.97}$
$\geqslant$2-jet, p_T^H >200 GeV	$1.00^{+0.97}_{-0.91}$	$1.00^{+1.10}_{-1.01}$
VBF-like	$1.00^{+0.41}_{-0.39}$	$1.00^{+0.49}_{-0.48}$

between the measured parameters are induced mainly due to migration between the signal regions. Figure 3.11 and Figure 3.14 show that the highest migration of the STXS signal processes happens from the neighbouring processes, where a neighbouring process is defined as a process, which shares the same boundary in phase space either in p_T^H or jet multiplicity, e.g. the 1-jet p_T^H $\in$[0, 60] GeV and p_T^H $\in$[60, 120] GeV regions share $p_T^H = 60$ GeV boundary, or 0-jet and 1-jet are neighbours, when all other parameters are the same. The VBF signal process has relatively significant contributions in all but the 0-jet signal region due to experimental difficulty to distinguish the ggF process with jets and the VBF process. Therefore, anti-correlations between the VBF-like POI and the ggF 1-jet and $\geqslant$2-jet POIs are expected. The correlation matrix in Figure 4.2 shows correlation coefficients between the POIs

when only statistical uncertainties are considered. The anti-correlations which are more negative than -4% are exclusively between neighbouring POIs or the VBF-like POI, as expected. In the case where both statistical and systematic uncertainties are considered (Figure 4.6), the moduli of correlation coefficients between neighbouring POIs in most cases are slightly reduced, while the moduli of correlation coefficients of non-neighbouring POIs tend to be slightly increased. This is an expected behaviour since most systematic uncertainties have a single nuisance parameter for all SRs and CRs, thus reduces the impact of contamination of other STXS signal processes in the dedicated signal regions.

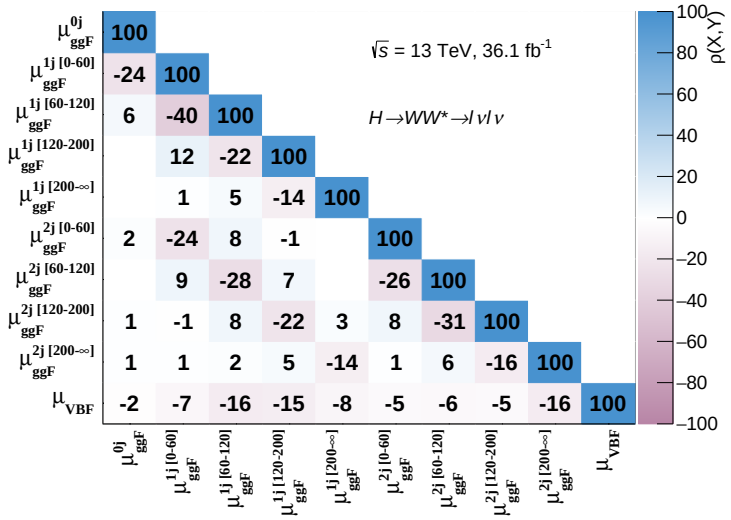


Figure 4.2: Correlation matrix, in %, for the expected values with the SM prediction of the POIs, shown in Table 4.14. Only statistical uncertainties are considered.

The measurements are relatively equally dominated by both systematic and data statistical uncertainties, as it is shown in Table 4.15. One of the largest single contributions to the total systematic uncertainty is limited MC statistics. This suggests that the measurement can be easily improved if more data is analysed (e.g. the full ATLAS Run 2 data set) and if larger MC samples are generated. Data statistical uncertainties are larger than systematic uncertainties for all but 0-jet and

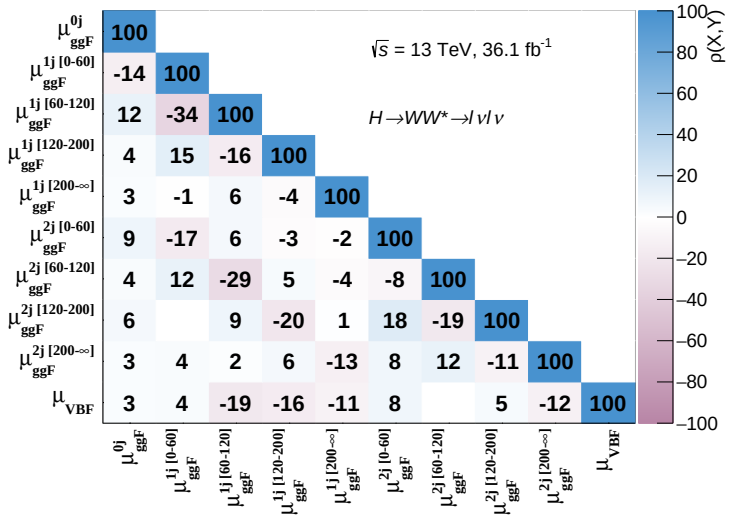


Figure 4.3: Correlation matrix, in %, for the expected values with the SM prediction of the POIs, shown in Table 4.14. Total uncertainties are considered.

one or more jets $p_T^H \in [0,\ 60]$ GeV POI measurements, while the MC statistical uncertainties are at least a half of all systematic uncertainties in all but 0-jet and $\geqslant$2-jet $p_T^H \in [60,\ 120]$ GeV measurements.

Merging scheme

The STXS results with the full stage 1 ggF production granularity (Table 4.14) show that a merging scheme should be designed in order to obtain significant measurements. The merging scheme proposed here combines ten POIs into seven. The 1-jet and $\geqslant$2-jet POIs with $p_T^H > 200$ GeV are designed to be sensitive to new particles via the gluon fusion loop, so they are combined together; the POIs in $\geqslant$2-jet category satisfying $p_T^H < 120$ GeV criterion have close kinematic properties and are poorly measured, so are merged into one POI; the rest of the POIs remain the same. The merging scheme is illustrated in Figure 4.4.

The expected results with the SM prediction are presented in Table 4.16. The POI sensitive to BSM physics (i.e. $\geqslant$1-jet $p_T^H > 200$ GeV POI) has an expected

Table 4.15: The expected absolute systematic uncertainties breakdown with the SM prediction for ten STXS signal strength POIs. The second column presents all but data statistical uncertainties, the third column shows data statistical uncertainties, and the last column lists MC statistical uncertainties.

STXS POI \ Uncertainty	All systematics	Data statistics	MC statistics
0-jet	0.19	0.12	0.07
1-jet, p_T^H ∈[0, 60] GeV	0.52	0.49	0.26
1-jet, p_T^H ∈[60, 120] GeV	0.37	0.49	0.23
1-jet, p_T^H ∈[120, 200] GeV	0.58	0.90	0.37
1-jet, p_T^H >200 GeV	1.04	1.60	0.60
⩾2-jet, p_T^H ∈[0, 60] GeV	1.96	1.87	1.07
⩾2-jet, p_T^H ∈[60, 120] GeV	0.73	1.08	0.18
⩾2-jet, p_T^H ∈[120, 200] GeV	0.58	0.82	0.39
⩾2-jet, p_T^H >200 GeV	0.58	0.88	0.33
VBF-like	0.33	0.35	0.20

uncertainties of $1.00^{+0.89}_{-0.83}$, while the ⩾2-jet $p_T^H < 200$ GeV POI is expected to be $1.00^{+0.72}_{-0.69}$. Therefore, this merging scheme is expected to produce significant results.

Since the uncertainties on the ⩾2-jet $p_T^H < 200$ GeV POI are reduced significantly below 100%, a new merging scheme is designed where this POI is split into p_T^H ∈[0,120] GeV and p_T^H ∈[120, 200] GeV POIs. The expected uncertainties on the first POI are $1.00^{+1.15}_{-1.11}$, and on the second the uncertainties are $1.00^{+1.00}_{-0.94}$. The uncertainties on the p_T^H ∈[0,120] GeV POI are significantly higher than 100% and therefore this merging scheme is not preferred.

The correlation matrix for the seven POIs shows similar pattern (Figure 4.5) as in the ten parameter case (Figure 4.6).

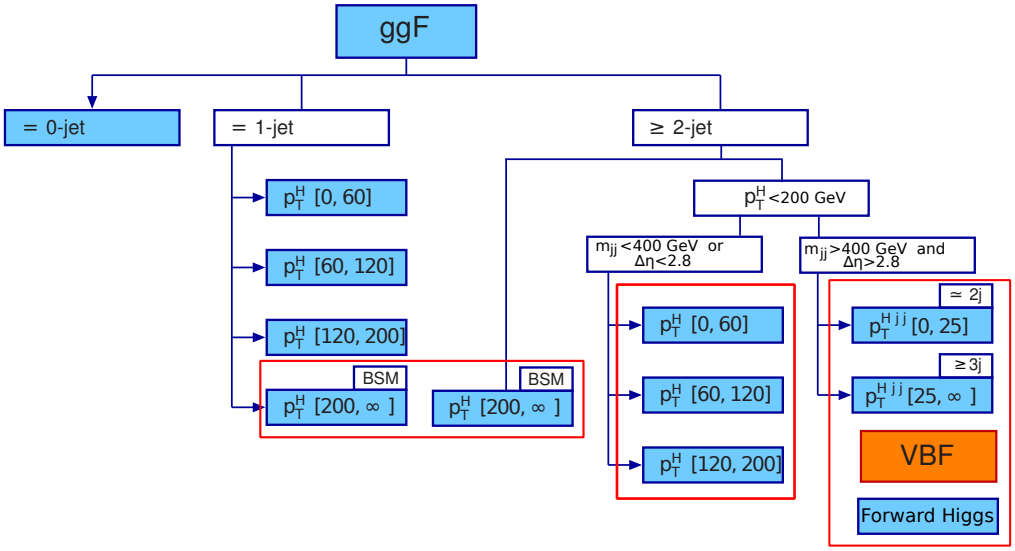


Figure 4.4: A merging scheme of the STXS stage 1 ggF and the inclusive VBF regions. Red boxes indicate which STXS processes are merged into new POIs.

Table 4.16: The expected best-fit values and uncertainties of a simultaneous fit for seven STXS signal strength POIs with the SM prediction, after a merging scheme of POIs is applied.

POI	Total uncertainties
0-jet	$1.00^{+0.23}_{-0.22}$
1-jet, $p_T^H \in$[0, 60] GeV	$1.00^{+0.70}_{-0.70}$
1-jet, $p_T^H \in$[60, 120] GeV	$1.00^{+0.59}_{-0.58}$
1-jet, $p_T^H \in$[120, 200] GeV	$1.00^{+1.11}_{-1.01}$
$\geqslant$2-jet, $p_T^H \in$[0, 200] GeV	$1.00^{+0.72}_{-0.69}$
$\geqslant$1-jet, $p_T^H >$200 GeV	$1.00^{+0.89}_{-0.83}$
VBF-like	$1.00^{+0.49}_{-0.48}$

The breakdown of the main contributions to the total uncertainties is presented in Table 4.17. As it was seen in the ten POIs case (Table 4.15), the uncertainties of the measurement have almost equal contributions from both data statistics and systematic uncertainties. The MC statistics are significant for all POIs. The theoretical

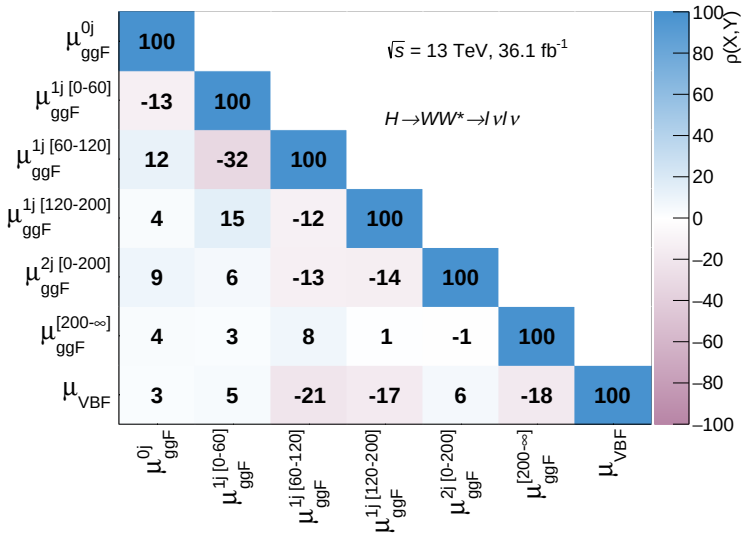


Figure 4.5: Correlation matrix, in %, for the expected values of the STXS signal strength parameters, presented in Table 4.16.

and experimental uncertainties have similar contributions. The largest theoretical uncertainties are on the ggF signal, WW and top quark background processes' estimation. The dominant experimental uncertainty is on jets. Large contributions arise also from uncertainties on b-tagging and the estimation of misidentified leptons.

4.4.2 Observed results

The measured STXS signal strength parameters and cross-sections times branching fraction, when a merging scheme is applied, are presented in Table 4.18. The SM predictions for cross-sections and branching fraction are obtained from [14]. The results show a good agreement with the SM prediction. The uncertainties are consistent with the expected values (Table 4.16), except the uncertainties on the 1-jet $p_T^H \in [0, 60]$ GeV POI are decreased and the uncertainties on the $\geqslant$1-jet $p_T^H >200$ GeV POI are increased, which can be accounted mainly by the ggF theoretical uncertainties. The ggF 0-jet region is the most precise, while one or more jets signal

Table 4.17: Relative breakdown of the main contributions to the total uncertainty (in %) for the expected STXS signal strength POIs with the SM prediction. The individual sources are listed with an indent. The b-tagging uncertainty corresponds to all uncertainties associated to the b-tagging algorithm; pile-up uncertainty considers only pile-up reweighting; and the combined jet uncertainty includes JES, JER, JVT, and fJVT uncertainties. The systematic uncertainties correspond to all but data statistical uncertainties.

Source	0-jet [%]	1-jet [0,60] [%]	1-jet [60,120] [%]	1-jet [120,200] [%]	$\geqslant$2-jet [0,200] [%]	$\geqslant$1-jet [200,∞] [%]	VBF-like [%]
Data statistics	54	68	80	83	69	83	73
CR statistics	32	36	24	15	32	10	22
MC statistics	32	36	38	34	31	31	41
Theoretical	52	43	31	35	29	37	31
ggF signal	34	15	16	15	15	14	5.3
WW	36	24	16	9.1	23	31	7.4
Top quark	23	25	17	26	12	15	26
Experimental	42	41	29	21	38	29	30
b-tagging	11	26	14	4.1	13	22	13
Pile-up	26	5.1	5.7	3.4	8.6	4.4	4.9
Jets	20	31	29	19	23	16	25
Lepton	19	2.7	4.2	2.3	3.6	4	3.8
E_T^{miss}	2	2.9	4.8	11	15	7	9.6
Misidentified leptons	23	30	5.4	6.9	25	14	21
Luminosity	11	4.3	3.4	2.8	4.4	5.1	5.6
Systematic	83	73	60	56	72	56	68

strength parameters have uncertainties between 0.56 and 1.09.

The anti-correlations (in range [-33, -9]) are observed between neighbouring POIs and between the ggF and VBF-like POIs only (Figure 4.6), which is consistent with the expected correlation matrix (Figure 4.5). The largest positive correlations are between POIs, where each of them has a large anti-correlation with the third POI, e.g. the correlation between the 1-jet p_T^H $\in$[120, 200] GeV and 1-jet p_T^H $\in$[0, 60] GeV POIs is due to their anti-correlation with 1-jet p_T^H $\in$[60, 120] GeV POI. The 0-jet and 1-jet p_T^H $\in$[60, 120] GeV POIs are both strongly anti-correlated with the

Table 4.18: The best-fit values and uncertainties of the STXS signal categories, when a merging scheme is applied. The second and third columns show observed signal strength parameters (μ_{STXS}) and cross-sections times branching fraction, respectively, while the last column shows the expected SM cross-sections times branching fraction [14].

STXS POI	μ_{STXS} *observed*	$\sigma \cdot \mathcal{B}_{H\to WW^*}$ [pb] *observed*	$\sigma \cdot \mathcal{B}_{H\to WW^*}$ [pb] *SM prediction*
0-jet	$1.20^{+0.23}_{-0.22}$	$7.0^{+1.3}_{-1.3}$	5.8 ± 0.4
1-jet, $p_T^H \in$[0, 60] GeV	$0.56^{+0.61}_{-0.62}$	$0.8^{+0.8}_{-0.9}$	1.39 ± 0.11
1-jet, $p_T^H \in$[60, 120] GeV	$1.07^{+0.59}_{-0.56}$	$1.0^{+0.6}_{-0.5}$	0.96 ± 0.07
1-jet, $p_T^H \in$[120, 200] GeV	$1.20^{+1.09}_{-1.01}$	$0.19^{+0.17}_{-0.16}$	0.158 ± 0.012
$\geqslant$2-jet, $p_T^H \in$[0, 200] GeV	$1.25^{+0.71}_{-0.66}$	$1.1^{+0.6}_{-0.6}$	0.87 ± 0.07
$\geqslant$1-jet, $p_T^H >$200 GeV	$1.47^{+0.95}_{-0.87}$	$0.18^{+0.12}_{-0.11}$	0.122 ± 0.009
VBF-like	$0.95^{+0.44}_{-0.43}$	$1.77^{+0.82}_{-0.80}$	1.86 ± 0.09

1-jet $p_T^H \in$[0, 60] GeV POI, which results in a positive correlation between these two POIs.

The main sources of uncertainties for this STXS measurement are shown in Table 4.19. Data statistics and systematic uncertainties are the dominant sources. The MC statistics are significant for all POIs. The theoretical and experimental uncertainties are comparable in size. The dominant theoretical uncertainties are associated to the modelling of the ggF signal, WW and top quark background processes; all of them are similar in size and the dominant source depends on the individual POI. The dominant source of experimental uncertainties is associated with jets, where the impact of JES uncertainty is slightly larger than the impact of JER. The other significant sources of experimental uncertainties are due to b-tagging and misidentified leptons. The breakdown of uncertainties is consistent with the expected breakdown (Table 4.17). The largest deviations are in ggF theoretical

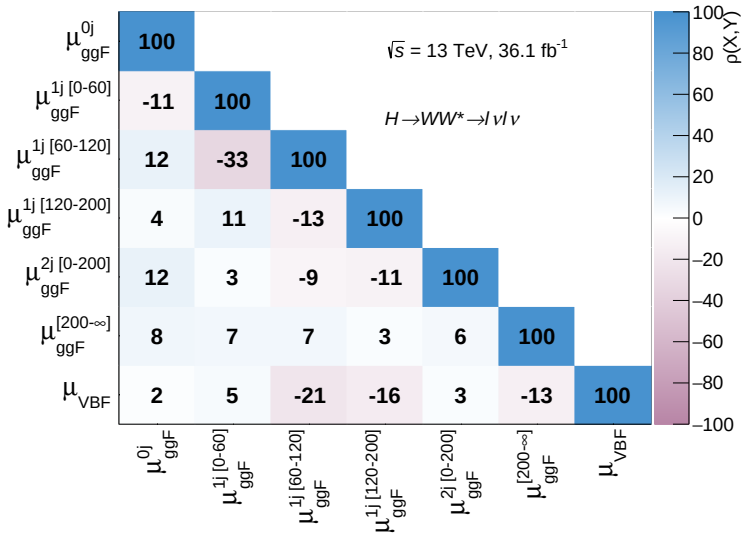


Figure 4.6: Correlation matrix for the measured values of the STXS signal strength parameters, shown in Table 4.18.

uncertainties, which is expected since the ggF uncertainties are scaled with the ggF yield. The observed higher/lower than the expected signal yield changes the fraction of ggF signal uncertainties accordingly as seen in Table 4.19. This explains the difference between the expected and observed uncertainties of the 1-jet $p_T^H \in [0, 60]$ GeV and $\geqslant$1-jet $p_T^H >$200 GeV POIs.

The results of the STXS signal strength parameters with data statistics and systematics breakdown are the following:

Table 4.19: Relative breakdown of the main contributions to the total uncertainty (in %) for the observed STXS signal strength POIs. The individual sources are listed with an indent. The b-tagging uncertainty corresponds to all uncertainties associated to the b-tagging algorithm; pile-up uncertainty considers only pile-up reweighting; and the combined jet uncertainty includes JES, JER, JVT, and fJVT uncertainties. The systematic uncertainties correspond to all but data statistical uncertainties.

Source	0-jet [%]	1-jet [0,60] [%]	1-jet [60,120] [%]	1-jet [120,200] [%]	$\geqslant$2-jet [0,200] [%]	$\geqslant$1-jet [200,∞] [%]	VBF-like [%]
Data statistics	51	69	78	83	70	81	74
CR statistics	31	37	25	16	25	9.4	18
MC statistics	32	38	39	36	29	32	39
Theoretical	57	43	32	31	36	42	31
ggF signal	40	8.9	18	13	24	21	10
WW	36	27	16	9.2	24	30	8.3
Top quark	23	29	18	26	12	19	27
Experimental	44	44	33	24	34	30	31
b-tagging	11	29	14	4.1	13	20	14
Pile-up	25	4.3	1.7	3.2	6.1	4.9	5.6
Jets	20	35	30	19	24	15	27
Lepton	19	3	4.2	2.3	3.8	3.8	4.3
E_T^{miss}	2	3.3	4.9	11	15	6.6	11
Misidentified leptons	20	24	3.8	3.2	21	11	18
Luminosity	13	2.6	4.5	2.1	5	3.7	5.3
Systematic	85	73	62	56	71	59	68

$$\mu^{\text{0-jet}} = 1.20 \pm 0.12(\text{stat.}) \pm 0.19(\text{syst.}) = 1.20^{+0.23}_{-0.22}$$

$$\mu^{\text{1-jet}}_{p_T^H \in [0,60]\text{GeV}} = 0.56 \pm 0.42(\text{stat.}) \pm 0.45(\text{syst.}) = 0.56^{+0.61}_{-0.62}$$

$$\mu^{\text{1-jet}}_{p_T^H \in [60,120]\text{GeV}} = 1.07 \pm 0.45(\text{stat.}) \pm 0.36(\text{syst.}) = 1.07^{+0.59}_{-0.56}$$

$$\mu^{\text{1-jet}}_{p_T^H \in [120,200]\text{GeV}} = 1.20 \pm 0.87(\text{stat.}) \pm 0.59(\text{syst.}) = 1.20^{+1.09}_{-1.01}$$

$$\mu^{\geqslant\text{2-jet}}_{p_T^H \in [0,200]\text{GeV}} = 1.25 \pm 0.48(\text{stat.}) \pm 0.49(\text{syst.}) = 1.25^{+0.71}_{-0.66}$$

$$\mu^{\geqslant\text{1-jet}}_{p_T^H > 200\text{GeV}} = 1.47 \pm 0.73(\text{stat.}) \pm 0.53(\text{syst.}) = 1.47^{+0.95}_{-0.87}$$

$$\mu_{\text{VBF-like}} = 0.95 \pm 0.32(\text{stat.}) \pm 0.29(\text{syst.}) = 0.95^{+0.44}_{-0.43}$$

Comparison with the Higgs combined measurement

The best to date ATLAS STXS measurement is obtained by a combined measurement of the Higgs boson decay modes $H \to \gamma\gamma$, ZZ^*, WW^*, $\tau\tau$, $b\bar{b}$, $\mu\mu$, searches for decays to invisible final states, and the measurements of the off-shell Higgs boson production [114]. The main channels which contribute to the Higgs ggF production are $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$, both measured using 79.8 fb^{-1} integrated luminosity, and $H \to WW^* \to e\nu\mu\nu$ and $H \to \tau\tau$ which are both measured using 36.1 fb^{-1} integrated luminosity. The $H \to WW^* \to e\nu\mu\nu$ is obtained using the same 0-jet and 1-jet data as in the analysis presented in this thesis but the combination analysis did not have a dedicated optimisation for the $H \to WW^*$ channel.

The results presented in this chapter are compared with the combined measurement in Table 4.20. The results presented in this chapter have higher uncertainties. However, the data set used to obtain this $H \to WW^*$ STXS measurement is significantly smaller than the ATLAS combined Higgs data set. It implies that the $H \to WW^*$ STXS measurement presented in this thesis is well optimised. Moreover, the $H \to WW^*$ contribution in the combination result [114] improved the measurements of the ggF 0-jet and 1-jet $p_T^H \in [0, 60]$ GeV POIs by 19% and 14%, respectively, while the remaining POIs were improved only by 1%-4% [115]. Therefore, this measurement should be included in future combinations to improve 1-jet and $\geqslant$2-jet combined measurements, especially at higher p_T^H.

4.5 Conclusions

Measurements of Simplified Template Cross Sections (STXS) via the gluon-gluon fusion (ggF) and vector boson fusion (VBF) modes in the $H \to WW^*$ decay channel have been performed using proton-proton collision data produced by the LHC at $\sqrt{s} = 13$ TeV and recorded by the ATLAS detector. Seven signal strength param-

Table 4.20: The comparison of the $H \to WW^*$ STXS measurement presented in this thesis (second column) with the best to date ATLAS combined measurement [114] (third column).

POI	$H \to WW^*$ 36.1 fb^{-1}	Combined 36.1 - 79.8 fb^{-1}
0-jet	$1.20^{+0.23}_{-0.22}$	$1.29^{+0.18}_{-0.17}$
1-jet [0, 60] GeV	$0.56^{+0.61}_{-0.62}$	$0.57^{+0.43}_{-0.41}$
1-jet [60, 120] GeV	$1.07^{+0.59}_{-0.56}$	$0.87^{+0.38}_{-0.34}$
1-jet [120, 200] GeV	$1.20^{+1.09}_{-1.01}$	$1.30^{+0.81}_{-0.72}$
$\geqslant$2-jet [0, 200] GeV	$1.25^{+0.71}_{-0.66}$	$1.11^{+0.56}_{-0.51}$
$\geqslant$1-jet >200 GeV	$1.47^{+0.95}_{-0.87}$	$2.05^{+0.84}_{-0.72}$
VBF-like	$0.95^{+0.44}_{-0.43}$	

eters and cross-sections times branching fraction corresponding to the STXS stage 1 regions are measured. The most precise measurement is of the ggF 0-jet parameter. The ggF $\geqslant$1-jet signal strength parameters are measured with uncertainties between 56% and 109%. The VBF-like signal strength parameter is measured with the average uncertainty of 44%. The results are consistent with the SM prediction.

The results can be improved by analysing a larger data set (e.g. the full ATLAS Run 2 data set) or using larger Monte Carlo samples. The results are comparable with the best to date ATLAS combined STXS measurement [114], which is performed on a significantly larger data set. The inclusion of the optimised $H \to WW^*$ STXS measurement should significantly improve the combined results. The measurement presented in this thesis can be used for Beyond Standard Model physics interpretations, for example using Effective Field Theory (EFT) framework. The EFT interpretation of these results in combination with $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels is presented in chapter 6.

Chapter 5

Effective Field Theory analysis methodology

Standard Model (SM) measurements, like the Simplified Template Cross Section (STXS) measurement presented in the previous chapter, provide information about the consistency with the SM and constraints on new physics. Constraints on new physics can be obtained in a model independent way using the Effective Field Theory (EFT) approach. The experimental analysis design for EFT interpretations was not well established when the analysis for this thesis started: the first ATLAS EFT interpretation [3] on combined Higgs data has only been performed as a part of this thesis. Consequently, the EFT analysis methodology had to be developed and this chapter is dedicated to present this work.

The main aspects of an EFT analysis are the following: the choice of an EFT model (it is typically an implementation in a Monte Carlo (MC) generator); the translation of MC samples into analytic EFT equations; a selection of a subset of Wilson coefficients in case the available measurements are insufficient to constrain a full set, which then requires an approach to account for the induced theory dependence; and fitting EFT equations against a set of measurements. The choice of an EFT model depends on the dimension of the EFT operators (section 1.2), the order in QCD, and the basis (section 1.2.1). The MC samples have to be translated

into analytic expressions since the domain of the Wilson coefficients is a continuous subset of the real numbers rather than a discrete set, which would lead to an infinite number of MC samples. Typically, only a subset of Wilson coefficients can be constrained due to a limited set of available measurements. This induces a model dependence, so a rigorous procedure must be developed to motivate the choice of the subset of Wilson coefficients. This model dependence can be addressed by taking into account the known constraints from other EFT fits or by the "Wilsons as functions" method, proposed in this thesis. Finally, a fit is performed by maximisation of a likelihood function.

The EFT models relevant to this thesis are introduced in 5.1. The procedure to translate MC generator output into analytic expressions is discussed in two parts: a translation based on couplings is described in section 5.2, while an original methodology to account for the particle widths' dependence on Wilson coefficients is explained in section 5.3. It is followed by another original contribution in section 5.4, where the "Wilsons as functions" method is proposed to account for the Wilson coefficients removed from the fit. Finally, a general discussion on EFT analysis design is presented in section 5.5. Discussions regarding the selection of theoretical constraints from other EFT fits and construction of likelihood functions are deferred to the next chapter since the methodology is both relatively straightforward and analysis dependent.

5.1 EFT models

Two EFT models are considered in this thesis: Higgs Effective Lagrangian (HEL) [116], and SMEFTsim [117]. The models are implemented in the MadGraph_aMC@NLO (MadGraph) generator (section 1.4.3). Both models simulate the SMEFT (section 1.2.1) at leading order (LO) in QCD with dimension-6 operators and represent the same physics. The main difference between these models is that the HEL is a par-

tial implementation of the strongly-interacting light Higgs (SILH) basis [118], and the SMEFTsim is written in the Warsaw basis [25] (section 1.2.1). The complete renormalisation program was performed for dimension-6 operators in the Warsaw basis only [14], which led to the development of the SMEFT at next-to-leading order (NLO) in QCD in the Warsaw basis [119]. At the start of this analysis the HEL model was the best available EFT implementation. Although EFT interpretations in both the HEL and SMEFTsim models carry equivalent value, it is anticipated that future EFT interpretations are likely to be performed in the Warsaw basis. Therefore, this thesis presents interpretations using both the HEL and SMEFTsim models. The models are introduced in this section.

5.1.1 Higgs Effective Lagrangian model

The Higgs Effective Lagrangian (HEL) model is a partial implementation of the SILH basis at leading order in perturbation theory and includes dimension-6 operators [116]. The model contains 39 operators related to Higgs physics but only 37 are linearly independent. The four-fermion operators are neglected in this implementation.

In the HEL model the Z boson mass m_Z is also dependent on EFT parameters in the following way:

$$m_Z^2 = m_{Z,SM}^2 \Big[1 - \mathtt{cT} + \frac{\mathtt{cA} \cdot 8\sin^4(\theta_W) + \mathtt{cWW} \cdot 2\cos^2(\theta_W) + \mathtt{cB} \cdot \sin^2(\theta_W)}{\cos^2(\theta_W)} \Big] \quad (5.1)$$

where $m_{Z,SM}$ is the SM Z boson mass, θ_W is Weinberg angle and `cT`, `cA`, `cWW` and `cB` are EFT parameters. This dependence is important when the model is translated into analytic expressions.

The HEL model uses the m_W, α_{EM}, and G_F input scheme, where m_W is the W boson mass, α_{EM} is the electromagnetic coupling constant, and G_F is the Fermi constant, which are the fixed inputs to the model.

5.1.2 SMEFTsim model

The SMEFTsim model is an implementation of the Warsaw basis at leading order in perturbation theory and includes dimension-6 operators [117]. The model option that is relevant to this thesis has a flavour symmetric $U(3)^5$ assumption including non-SM CP violating phases. The $U(3)^5$ symmetry corresponds to the limit of unbroken global flavour symmetry in the SM Lagrangian when Yukawa matrices $\mathbf{Y}_{u,d,e} \to \mathbf{0}$[1]. All Wilson coefficients are scalar parameters [117].

The version of the SMEFTsim model used in this thesis employs the m_W, m_Z, and G_F input parameter scheme, where m_W is the W boson mass, m_Z is the Z boson mass and G_F is the Fermi constant [117].

5.1.3 Generation of the samples

The parton level events are generated from the Matrix Elements using the MadGraph generator with the EFT model implementation. The parton shower is generated using the SM PYTHIA 8. The BSM dependence enters the process only via the MadGraph generator. The BSM operators in parton shower would have a subleading effect and would mainly affect vertices that do not contain the Higgs boson.

5.2 Creation of analytic expressions from MC samples

The continuous nature of the domain of the Wilson coefficients leads to the necessity to translate the generator output into analytic expressions. The EFT Lagrangian is expressed as a series of Wilson coefficients (section 1.2), which allows one to express cross-sections and decay widths as a series of Wilson coefficients. The derivation of EFT equations is described in section 5.2.1 and its validation is presented in section

[1]Here each of the five $U(3)$s correspond to a symmetry for each of the five fermion representations: l_L, e_R, q_L, d_R, u_R [120].

5.2.2.

5.2.1 Derivation of EFT equations

This section presents the mathematics and the theory behind the extraction of EFT expressions. It is followed by a discussion of the treatment of the ratios of these expressions. Finally, practical aspects of obtaining the expressions are briefly reviewed.

Analytic expressions for cross-sections and decay widths

Let us consider the effective Lagrangian including dimension-6 operators O_i, where the dependence on Wilson coefficients enters only via the couplings (Einstein's summation convention is used in this chapter[2]):

$$L = L_{SM} + c_i O_i \Lambda^{-2} \tag{5.2}$$

Since the effective Lagrangian is a sum of SM and BSM Lagrangian parts ($L_{BSM} \equiv c_i O_i \Lambda^{-2}$), the cross-section can be expressed as a sum proportional to pure SM, an interference of SM and BSM, and a pure BSM matrix elements squared:

$$\sigma \propto |M_{SM}|^2 + 2Re\{M_{SM} M_{BSM}^{\dagger}\} + |M_{BSM}|^2 \tag{5.3}$$

where M_{SM} and M_{BSM} are the SM and BSM matrix elements, respectively. The $|M_{BSM}|^2$ term has a suppression of Λ^{-4} but it is the leading order term that is not dependent on the SM amplitude and is the only term carrying the dependence on CP odd operators which motivates keeping it in the equation.

Wilson coefficients, entering the effective Lagrangian in this form, are scaling factors of BSM vertices of the Feynman diagrams. If up to N BSM vertices per Feynman diagram are considered, then the matrix element has the following form:

$$M = M_{SM} + D_{i_1} c_{i_1} + D_{i_1 i_2} c_{i_1} c_{i_2} + ... + D_{i_1 ... i_N} c_{i_1} ... c_{i_N} \tag{5.4}$$

[2]Under Einstein's summation convention the summation sign $\sum$ is omitted; the summation is indicated by an index appearing twice in a single term. More information can be found in [121]

where the constant prefactors D are complex numbers. The analysis in this thesis considers only the leading contribution, when at most one BSM vertex per diagram is allowed. Then the matrix element is linear in Wilson coefficients and equation 5.4 becomes:

$$M = M_{SM} + D_i c_i \tag{5.5}$$

In a basis where the Wilson coefficients are real, a quadratic equation is derived by substituting eq. 5.5 into eq. 5.3:

$$\sigma = \sigma_{SM} + A_i c_i + B_{ij} c_i c_j \tag{5.6}$$

where constant prefactors A_i, B_{ij} are real numbers. The prefactors A_i and B_{ij} are typically extracted from MC samples. The $|M_{BSM}|^2$ term can be dropped by neglecting the B_{ij} coefficients.

A decay width has the same form as cross-sections:

$$\Gamma = \Gamma_{SM} + A_i c_i + B_{ij} c_i c_j \tag{5.7}$$

The equations [5.6] and [5.7] are the mapping from the EFT to any cross-section or decay width, assuming that the Wilson coefficient dependence enters the Lagrangian via the couplings only.

Ratios of branching fractions

The Higgs boson decay modes are usually measured as a ratio of branching fractions. The EFT parameterisation for a ratio can be expressed either as a ratio of EFT equations or a polynomial using the Taylor expansion, where the latter ensures that there are no zero values in the denominator. Here an example of the ratio of branching fractions of $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ is given.

The ratio of branching fractions of two processes is equal to the ratio of their decay widths:

$$\frac{\mathrm{BF}(H \to \gamma\gamma)}{\mathrm{BF}(H \to ZZ^* \to 4\ell)} = \frac{\Gamma(H \to \gamma\gamma)}{\Gamma(H \to ZZ^* \to 4\ell)} \tag{5.8}$$

Using equation 5.7, let

$$\Gamma(H \to \gamma\gamma) = \Gamma^{\gamma\gamma}_{SM} + A^{\gamma\gamma}_i c_i + B^{\gamma\gamma}_{ij} c_i c_j \tag{5.9}$$

and

$$\Gamma(H \to ZZ^* \to 4\ell) = \Gamma^{4\ell}_{SM} + A^{4\ell}_i c_i + B^{4\ell}_{ij} c_i c_j \tag{5.10}$$

Then substituting eq. 5.9 and eq. 5.10 into eq. 5.8 yields:

$$\frac{\mathrm{BF}(H \to \gamma\gamma)}{\mathrm{BF}(H \to ZZ^* \to 4\ell)} = \frac{\Gamma^{\gamma\gamma}_{SM} + A^{\gamma\gamma}_i c_i + B^{\gamma\gamma}_{ij} c_i c_j}{\Gamma^{4\ell}_{SM} + A^{4\ell}_i c_i + B^{4\ell}_{ij} c_i c_j} \tag{5.11}$$

The equation 5.11 is expanded using the Taylor expansion as follows:

$$\begin{aligned}\frac{\mathrm{BF}(H \to \gamma\gamma)}{\mathrm{BF}(H \to ZZ^* \to 4\ell)} \approx \frac{\Gamma^{\gamma\gamma}_{SM}}{\Gamma^{4\ell}_{SM}}\Bigg(1 &+ \frac{A^{\gamma\gamma}_i c_i}{\Gamma^{\gamma\gamma}_{SM}} + \frac{B^{\gamma\gamma}_{ij} c_i c_j}{\Gamma^{\gamma\gamma}_{SM}} - \frac{A^{4\ell}_i c_i}{\Gamma^{4\ell}_{SM}} - \frac{B^{4\ell}_{ij} c_i c_j}{\Gamma^{4\ell}_{SM}} \\ &+ \frac{A^{\gamma\gamma}_i A^{4\ell}_j c_i c_j}{\Gamma^{\gamma\gamma}_{SM}\Gamma^{4\ell}_{SM}} + \frac{A^{4\ell}_i A^{4\ell}_j c_i c_j}{(\Gamma^{4\ell}_{SM})^2}\Bigg)\end{aligned} \tag{5.12}$$

Implementation of EFT equation derivation

The EFT equations are derived by solving a set of linear equations for the variables σ_{SM}, A_i, and B_{ij}. The Wilson coefficients are chosen freely and the cross-section or decay width values are obtained from an MC generator. $(N+1)(1+N/2)$ is the minimum number of samples needed to produce one quadratic EFT equation with N Wilson coefficients: one sample is needed for the SM value extraction, N samples for A_i, and $N(N+1)/2$ samples for B_{ij} extraction. Since it is a linear problem, there are several classical methods to obtain the solution, which correspond to different MC generator configurations and potentially different precision given the same number of generated events.

One approach is to solve the system using a matrix inversion method, called Morphing [122]. In this case the samples are generated by setting a subset of Wilson coefficients to non-zero values. It may require high statistics samples to separate A_i and B_{ij} coefficients that are mixed in one MC sample.

An alternative method, which is used in this thesis, is based on the idea of maximally decoupling variables in the EFT equations. In the context of the previously discussed Morphing method, it is equivalent to constructing a close to diagonal matrix that does not need to be inverted. This is achieved by producing pure SM, SM-BSM interference, and pure BSM terms separately. As a result, σ_{SM}, A_i, and B_{ij}, where $i = j$, can be retrieved directly, which reduces statistical uncertainties since there is no mixing. The term B_{ij}, where $i \neq j$, is derived using equation 5.13, where σ_{BSM^2} is an output from MC generator, and B_{11}, B_{22} are obtained by the procedure described above.

$$\sigma_{BSM^2} = B_{11}c_1^2 + B_{12}c_1c_2 + B_{22}c_2^2 \tag{5.13}$$

5.2.2 Validation of the equations

Translation of the MC samples into the EFT analytic equations relies on the assumption that the equations are capable of predicting the cross-sections (or decay widths) with the preferred precision. The validation is performed by comparing a cross-section obtained directly with an MC generator with the one calculated using the EFT equations. The MadGraph input propagator widths are set to the SM values for both the production of the equations and the validation. The validation procedure is described in this section.

Validation procedure

The validation is performed using the HEL model. The validation is performed on the Simplified Template Cross Sections (STXS) stage 1 cross-sections (section 1.3) and H$\to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and the total Higgs boson decay widths. Without loss of generality, the following subset of seven Wilson coefficients is considered: {`cA`, `cG`, `cu`, `cWW`, `cHW`, `cB`, `cHB`}, with a constraint `cWW`+`cB`=0. The `cWW`+`cB`=0 constraint is motivated by the constraint on the Peskin-Takeuchi S parameter [123, 124]. In

the HEL model the Z boson mass m_Z has a dependence on Wilson coefficients (equation 5.1), and the EFT equations are derived such that they are consistent with the SM m_Z measurements. As a result, the processes including Z bosons are treated differently by choosing Wilson coefficient values such that m_Z is consistent with the SM uncertainties. Sets of randomly chosen Wilson coefficients' values are used to obtain cross-section and decay width values from both the generated MC samples and the EFT equations.

The validation is separated in three versions based on the external constraints and the required precision. The production processes are split into the kinematic regions corresponding to the extended stage 1 version of the STXS framework [26]. Version 1 contains production and decay processes independent of m_Z (ggF, WH where W decays into leptons, WH where W decays into hadrons, and $t\bar{t}$H productions and H$\rightarrow \gamma\gamma$ decay) and the total width of the Higgs boson since there the m_Z contribution is small. For each process 10 samples are generated using different Wilson coefficient sets, where each sample contains $5 \cdot 10^5$ events. The versions 2 and 3 contain processes that include Z bosons. Version 2 consists of ZH where Z decays into leptons, ZH where Z decays into hadrons, and VBF productions. 10 different samples are generated for each process. Version 3 consists of $H \rightarrow ZZ^* \rightarrow 4\ell$ decay since its deviations from the SM are very small and a large number of 1000 samples is needed to verify the deviations and avoid significant statistical fluctuations. Each sample in versions 2 and 3 has $5 \cdot 10^4$ events.

The dominant uncertainty is the MadGraph statistical uncertainty of $< 1\%$, and the difference between the predicted and generated cross-sections is found to be Gaussian-distributed with respect to this uncertainty (Figure 5.1). The equations for each STXS bin are determined to a statistical precision of $< 0.01\%$, and are therefore negligible. Several example cross-sections are shown in Table 5.1.

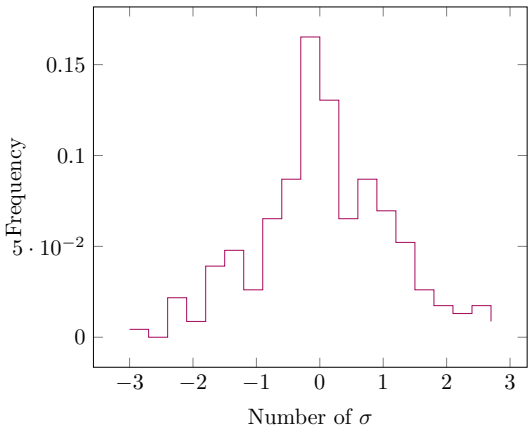


Figure 5.1: The plot of the frequency of the difference between the STXS cross-sections calculated with the MadGraph generator and those calculated with the EFT equations that is divided by the leading uncertainty - the statistical uncertainty on the MadGraph calculations (σ). Results are shown for ggF, WH where W decays into leptons, WH where W decays into hadrons, and $t\bar{t}$H production processes and the $H \to \gamma\gamma$ decay process. The MadGraph generator calculations have a statistical uncertainty $< 1\%$.

Table 5.1: Example results of the validation of the EFT equations in the listed STXS regions. The cross-section generated with MadGraph (σ_{MG}) is compared to the cross-section obtained using the EFT equation (σ_{eq}). The two are consistent within the statistical uncertainty on the MadGraph calculation (δ^{stat}_{MG}).

STXS region	σ_{MG}/σ_{SM}	σ_{eq}/σ_{SM}	δ^{stat}_{MG}
$gg \to H$ (≥ 2 jets, $p_T^H \geq 200$ GeV)	0.807	0.799	0.006
$gg \to H$ (≥ 2 jets, $p_T^H \geq 200$ GeV)	1.169	1.174	0.008
$q\bar{q} \to Hl\nu$ ($p_T^V \geq 250$ GeV)	1.001	1.001	0.001
$q\bar{q} \to Hl\nu$ ($p_T^V \geq 250$ GeV)	0.944	0.944	0.002
$q\bar{q} \to Hl\nu$ ($p_T^V \geq 250$ GeV)	1.392	1.390	0.002
$gg/q\bar{q} \to t\bar{t}H$	0.876	0.876	0.001
$gg/q\bar{q} \to t\bar{t}H$	1.532	1.533	0.002

5.3 Dependence on widths

In the previous section the derivation of equations was discussed in the case where the input decay widths in the generator configuration are set to constant values.

This assumption works well for small Wilson coefficient values. In this section the cross-section dependence on the input decay widths is explored. The SMEFTsim model is used in this case.

The SMEFTsim model in MadGraph has a dependence on Wilson coefficients via the scaling factors of the vertices, input decay widths, and the other input parameters, e.g. m_Z, G_F, α_{EW}, etc. For large Wilson coefficient values a significant contribution comes from the change of the input decay widths. Empirical studies of the Higgs STXS cross-sections have shown that in almost all cases the other input parameters' dependence on Wilson coefficients is insignificant. Figure 5.2 shows how the linear parameterisation deviates from the MadGraph output of the SM and SM-BSM interference parts, which are linear in case the input widths are set to constant SM values and non linear otherwise.

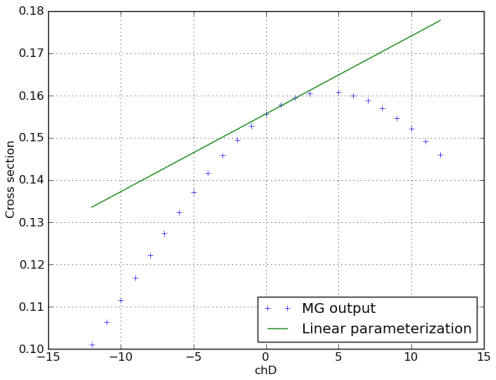


Figure 5.2: ZH production, where Z boson decays into leptons, cross-section dependence on chD Wilson coefficient for a sample containing the SM and SM-BSM interference terms. The solid green line shows the output of a linear parameterisation, where the input widths are set to the SM values, while the blue crosses correspond to cross-sections directly generated with MadGraph, where the input widths are estimated for the input Wilson coefficient values.

There are two main ways to derive the EFT parameterisation including the dependence on widths: to fit each distribution to a Kth order polynomial (e.g. using the Professor tool [125]), or use a function-of-function method, which is developed

for this analysis. The function-of-function method is introduced in section 5.3.1.

5.3.1 Function-of-function method

Empirically, in the case of Higgs production cross-sections, the strongest dependence on Wilson coefficients comes from two sources: via the scaling factors of the vertices and via the decay widths of the intermediate particles. The idea of the method is first to express a cross-section (σ) as a polynomial of Wilson coefficients (c_i) assuming the dependence on them comes only via the scaling factors of the vertices (i.e. the case of constant propagator decay widths):

$$\sigma = \sigma_{SM} + A_i c_i + B_{ij} c_i c_j \tag{5.14}$$

where σ_{SM} is the SM cross-section, and A_i, B_{ij} are constant prefactors. Then each prefactor in this polynomial is expressed as a function of propagator widths (w_k):

$$A_i \to A_i(w_k), \qquad B_{ij} \to B_{ij}(w_k) \tag{5.15}$$

Following that, the decay widths are expressed as polynomials of Wilson coefficients:

$$w_p \to w_p(c_i) \tag{5.16}$$

Decay widths are mainly affected by scaling factors of the vertices and therefore are polynomials (eq. 5.7). Finally, decay width equations 5.16 are substituted into functions of decay widths 5.15, while the functions of decay widths are substituted into the original equation of the cross-section 5.14.

In this study the SMEFTsim model is used. The SMEFTsim model has only four decay widths as inputs - the Higgs, Z, W bosons, and the top quark. So, for the last set of functions only four functions have to be calculated once and can be used for any production.

The first function is the expression of a cross-section as a polynomial, where Wilson coefficients come only via the scaling factors of the vertices, while other

dependence is encapsulated in the prefactors:

$$\sigma = \sigma_{SM}(\Delta w_p) + A_i(\Delta w_p)c_i + B_{ij}(\Delta w_p)c_i c_j \tag{5.17}$$

where $A_i(\Delta w_p)$, $B_{ij}(\Delta w_p)$ are prefactors derived from MadGraph as functions of the difference of particle widths for a given set of Wilson coefficients and the SM value of the propagator (Δw_p), and c_i is a Wilson coefficient.

Without loss of generality, let us focus on the interference term of the SM and BSM: $A_i(w_p)c_i$. The prefactor $A_i(w_p)$ is not restricted to be a polynomial but empirically (from Higgs STXS studies) it is sufficiently well modeled by a quadratic polynomial. Let

$$A(\Delta w) = A(0) + a_p \Delta w_p + b_{pq}\Delta w_p \Delta w_q \tag{5.18}$$

where $A(0)$, a_p and b_{pq} are constants determined using MadGraph generator, and p, q indices indicate the iteration over the all four particle widths (Higgs, W, Z bosons' and top quark's). The decay width is not affected by itself, so it has a quadratic form as in equation 5.7:

$$w_p = w_{SM,p} + k_{pi}c_i + t_{pij}c_i c_j \tag{5.19}$$

where $w_{SM,p}$ is an SM decay width, k_{pi} and t_{pij} are constant prefactors determined using MadGraph generator. Then from eq. 5.19:

$$\Delta w_p \equiv w_p - w_{SM,p} = k_{pi}c_i + t_{pij}c_i c_j \tag{5.20}$$

The only missing expression is $\sigma_{SM}(\Delta w)$. Let:

$$\sigma_{SM}(\Delta w) = \sigma_{SM}(0) + r_p \Delta w_p + s_{pq}\Delta w_p \Delta w_q \tag{5.21}$$

where r_p, s_{qp} are constant prefactors. The cross-section equation linear in Wilson coefficients is obtained by substituting equations 5.18, 5.19 and 5.21 into equation

5.17. Then the linear part in Wilson coefficients of the equation 5.17 is:

$$\begin{aligned}\sigma &= \sigma_{SM}(0)\Big(1 + \frac{r_p}{\sigma_{SM}(0)}\Delta w_p\Big) + A_i(0)c_i \\ &= \sigma_{SM}(0)\Big(1 + \frac{r_p}{\sigma_{SM}(0)}k_{pi}c_i\Big) + A_i(0)c_i \\ &= \sigma_{SM}(0) + r_p k_{pi}c_i + A_i(0)c_i\end{aligned} \tag{5.22}$$

The new term that accounts for the propagator decay widths is $r_p k_{pi} c_i$. This component is computationally simple: the index p is summed over the number of input decay widths, which in the SMEFTsim or the HEL case is four, while k_{pi} corresponds to prefactors in the equation 5.20 of width's dependence on Wilson coefficients, which is calculated once for EFT model and can be applied for all equations.

Validation of the method

The method is validated by investigating ZH production, where Z boson decays into leptons. The SMEFTsim EFT model with $U(3)^5$ flavour symmetry assumption and m_W input scheme is used to produce the SM and SM-BSM interference cross-section. Only the Wilson coefficient chD is considered. The ZH production cross-section depends only on the Z boson width since the Higgs boson is a final state particle.

The equations 5.17, 5.18, 5.19 now become:

$$\sigma = \sigma_{SM}(\Delta w) + A(\Delta w) \cdot chD \tag{5.23}$$

$$\sigma_{SM}(\Delta w) = \sigma_{SM}(0) + a_\sigma \cdot \Delta w + b_\sigma \cdot (\Delta w)^2 \tag{5.24}$$

$$A(\Delta w) = A(0) + a \cdot \Delta w + b \cdot (\Delta w)^2 \tag{5.25}$$

$$w = w_{SM} + k \cdot chD + t \cdot chD^2 \tag{5.26}$$

The chD in the eq. 5.26 is varied in range [-12, 12]. As a result, w in eq. 5.25 is varied in range [2.4, 3.3] GeV based on y axis range in Figure 5.3.

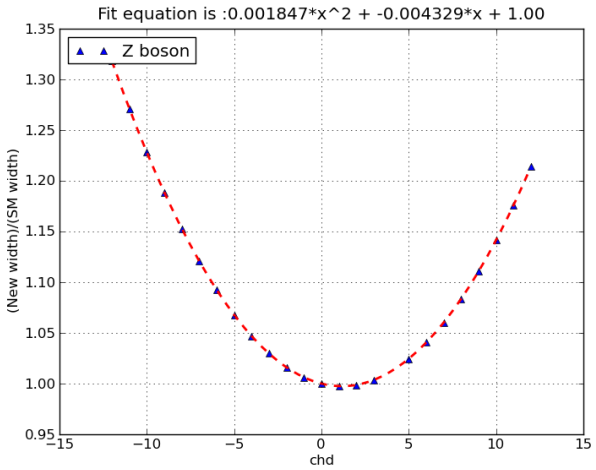


Figure 5.3: The Z boson width dependence on chD calculated using MadGraph. The equation extracted from a fit to quadratic polynomial is printed on top of the plot.

In this case both $\sigma_{SM}(\Delta w)$ and $A(\Delta w)$ have the same dependence on the width (deduced from MadGraph output), i.e. $a_\sigma/\sigma_{SM}(0) = a/A(0)$ and $b_\sigma/\sigma_{SM}(0) = b/A(0)$, so for the final result the dependence derived for $\sigma_{SM}(\Delta w)$ is used in both cases. Figure 5.4 shows that modeling $\sigma_{SM}(\Delta w)$ as a quadratic polynomial in Δw is sufficient.

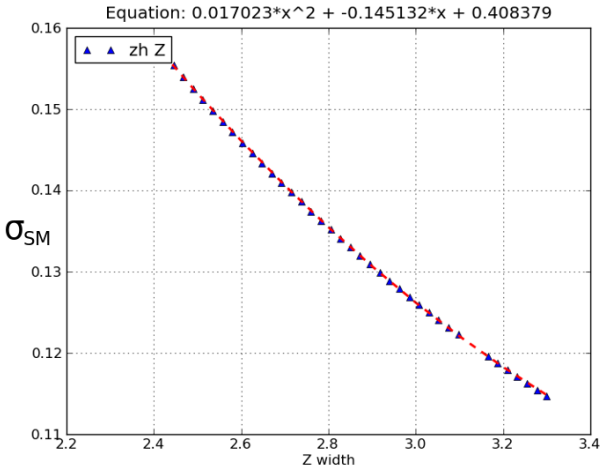


Figure 5.4: The $\sigma_{SM}(w)$ fitted to a quadratic polynomial. The equation of the fitted polynomial is printed on top of the plot.

Using this information and substituting equations 5.24 and 5.25 into 5.23, the cross-section is:

$$\sigma = (\sigma_{SM}(0) + A(0) \cdot chD)(\sigma_{SM}(0) + a_\sigma \cdot \Delta w + b_\sigma \cdot (\Delta w)^2)/\sigma_{SM}(0) \qquad (5.27)$$

Substitution of equation 5.26 into 5.27 yields a fifth order polynomial in chD. Figure 5.5 shows that this polynomial agrees well with the direct MadGraph output in the whole range. It also shows that assuming σ_{SM} is linear in widths (i.e. $b_\sigma = 0$), the resulting cubic polynomial does not model the cross-section well for large chD values, as expected.

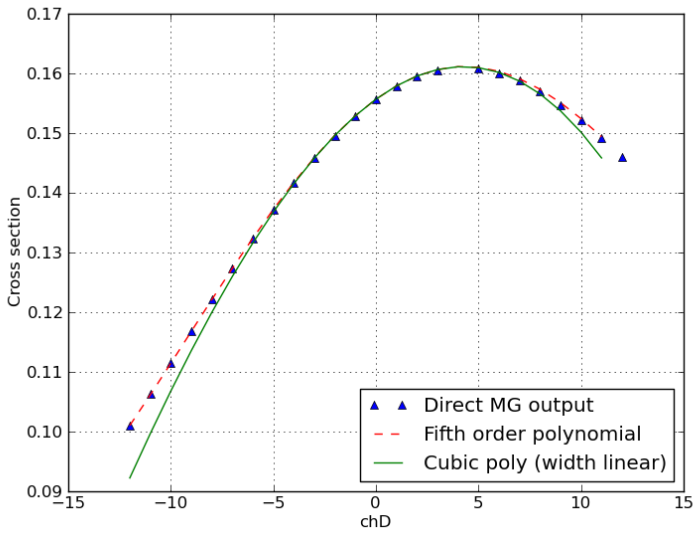


Figure 5.5: The dependence of the ZH production cross-section on the chD coefficient. Fifth order polynomial (red dashed) derived using function-of-function method is compared with the direct MadGraph output (propagator widths are calculated by MadGraph) (blue triangles). Green solid line corresponds to a cubic equation derived assuming σ_{SM} is linear in propagator widths.

In the validation of this method only the SM-BSM interference terms are tested. The resulting functions are higher order polynomials of chD than the neglected pure BSM terms. The same procedure is applicable for any other prefactor, e.g. the prefactors for pure BSM terms. In the final EFT equation, the terms that exceed

the order of the highest suppression term, should be truncated. In the case of the linear fit, the linearised version as in equation 5.22 is sufficient.

Comparison of function-of-function method with directly fitted polynomials

If the generator output is parameterised by directly fitting it to polynomials, then in the case of N Wilson coefficients one sample is needed for estimating the SM cross-section, N samples to evaluate linear terms, $N(N+1)/2$ samples for quadratic prefactors, and cubic terms contribute $(N^3+3N^2+2N)/6$ additional prefactors in a polynomial. For example, the derivation of one cubic equation with 10 Wilson coefficients would require 286 MC samples.

In case the function-of-function method is used, the dependence on propagator widths would be calculated only once for four particles (Higgs, Z, W bosons and top quark). In the case of the Higgs boson production, only VBF, WH and ZH production modes depend on these widths: VBF - on Z and W bosons' widths, while WH only on W boson's width, in the ZH case - only on Z boson's width. The equation of one width has $1+N(N+3)/2$ prefactors (when modelled as a quadratic polynomial). In order to estimate a prefactor's dependence on propagator widths, two additional MadGraph samples are needed when the dependence is on one width (one for a_p and one for b_{qp} in eq. 5.18), and five are needed when the dependence is on two widths (two for a_p and three for b_{qp}). The first EFT equation dependent on two widths would need $(1+N)(1+5)+2(1+N(N+3)/2)$ samples, while any other equation would only need $(1+N)(1+5)$ samples since the width equations have to be calculated once. In case pure BSM terms are included, $(1+5)N(N+1)/2$ additional samples would be needed. In the case of a cross-section consisting of SM and SM-BSM interference terms with 10 Wilson coefficients, 193 prefactors would be needed for the first EFT equation. Each additional equation would need only 11 samples in case it does not depend on propagator widths (e.g. ggF), 33 samples if it

depends on one propagator width (WH, ZH, $t\bar{t}$H), or 66 samples in case it depends on two particle widths (VBF).

In the discussed example, the function-of-function method is slightly more accurate (Figure 5.6) than a direct fit to a polynomial and for larger number of equations requires significantly smaller number of samples. This method could also be beneficial for estimating uncertainties due to truncation of higher order terms.

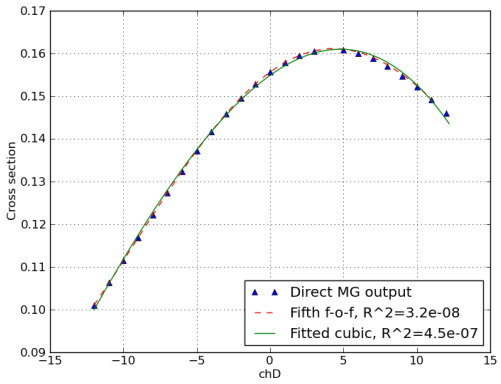


Figure 5.6: The dependence of the ZH production cross-section on the chD coefficient. Fifth order polynomial (red dashed line) derived using the function-of-function method agrees with the direct MadGraph output (blue triangles) slightly better than the fitted cubic polynomial (green solid line).

5.4 The "Wilsons as functions" method

One of the main objectives of the EFT is model independence. Satisfying this condition becomes problematic when a limited set of measurements is available. Consequently, the full set of Wilson coefficients entering EFT equations cannot be constrained by this set of measurements. A general procedure often applied is to select of a subset of Wilson coefficients, which can be constrained by the available set of measurements, while the rest of the coefficients are set to zero. This pragmatic choice induces model dependence unless the removed set of parameters is motivated by other measurements, where these Wilson coefficients are highly constrained. In

the scope of this thesis an alternative approach is developed, which is introduced in this section.

The idea of the method is to include the other Wilson coefficients, which are otherwise set to zero, by defining them as random variables independent of the fitted coefficients. The final result is presented as a function of these EFT coefficients and can be reused in the future when other measurements and constraints on these Wilson coefficients are available. In addition, this method can also be used to estimate the uncertainties on the analysis approach when parameters are set to zero. The method is derived in the following section.

5.4.1 Mathematical derivation of the method

The EFT analysis can be performed either using complex likelihood functions containing the full information on individual uncertainties or using finalised measurements when only the total uncertainty is available. In the latter case the EFT likelihood function is constructed from a multivariate Gaussian distribution [126]:

$$G(\boldsymbol{x}) = \frac{\exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{d})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{d}))}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}} \tag{5.28}$$

where $\boldsymbol{x}$ is a k dimensional vector, $\boldsymbol{d}$ is a vector of its mean values, and $\boldsymbol{\Sigma}$ is a covariance matrix with $|\boldsymbol{\Sigma}|$ as its determinant. In the EFT case $\boldsymbol{d}$ is a vector of a measurement's values' deviations from the SM values, where k is the number of measurements, and $\boldsymbol{x}$ is a vector of EFT equations.

The negative log likelihood is:

$$-2\log L = (\boldsymbol{x}-\boldsymbol{d})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{d}) + \text{const} \tag{5.29}$$

Let us denote $\boldsymbol{M} \equiv \boldsymbol{\Sigma}^{-1}$, and let us consider the case where linear EFT parameterisation is used[3], so we express $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{c}$, where $\boldsymbol{A}$ is a linear mapping from Wilson

[3]By the construction of the EFT, linear EFT parameterisation provides the leading EFT contribution.

coefficient space to the measurement space, and $\boldsymbol{c}$ is a Wilson coefficient vector. $\boldsymbol{A}$ is a matrix, which is not necessarily a square matrix, therefore not invertable in a general case. Consequently, the likelihood becomes (ignoring the constant terms):

$$-2\log L = (\boldsymbol{Ac} - \boldsymbol{d})^T \boldsymbol{M}(\boldsymbol{Ac} - \boldsymbol{d}) \tag{5.30}$$

The most likely Wilson coefficient values are obtained by minimising the negative log likelihood in eq. 6.2. The likelihood in eq. 5.30 can be differentiated with respect to $\boldsymbol{c}$:

$$\frac{\partial(-\log L)}{\partial c_i} = A^T_{ij} M_{jn}(A_{nm}c_m - d_n) \tag{5.31}$$

which leads to:

$$\Rightarrow \boldsymbol{A}^T\boldsymbol{MAc} - \boldsymbol{A}^T\boldsymbol{Md} = 0 \tag{5.32}$$

$$\Rightarrow \boldsymbol{A}^T\boldsymbol{MAc} = \boldsymbol{A}^T\boldsymbol{Md} \tag{5.33}$$

The solution for the Wilson coefficient vector is:

$$\boldsymbol{c} = (\boldsymbol{A}^T\boldsymbol{MA})^{-1}\boldsymbol{A}^T\boldsymbol{Md} \tag{5.34}$$

Note: $(\boldsymbol{A}^T\boldsymbol{MA})^{-1} \neq \boldsymbol{A}^{-1}\boldsymbol{M}^{-1}(\boldsymbol{A}^T)^{-1}$ since $\boldsymbol{A}$ is not a square matrix in general case and therefore the inverse is not defined.

Uncertainties are retrieved from the correlation matrix $(\boldsymbol{A}^T\boldsymbol{MA})^{-1}$, which is an inverse of the Hessian.

The equation for the un-fitted Wilson coefficients

In this section the result of the likelihood fit will be expressed as a function of the un-fitted Wilson coefficients.

First, let the EFT vector $\boldsymbol{x}$ be expressed as:

$$\boldsymbol{x} = \boldsymbol{Ac} + \boldsymbol{A'c'} \tag{5.35}$$

where $\boldsymbol{c}$ is a vector of Wilson coefficients which are floating and fitted in the fit, $\boldsymbol{c'}$ is a vector of Wilson coefficients which are assumed to be unknown random variables (which would be set to zero otherwise), and $\boldsymbol{A}$, $\boldsymbol{A'}$ are their corresponding mappings from Wilson coefficient space to the measurement space. This method targets the evaluation of model dependence rather than a combination with other measurements, so to ensure that external measurements of $\boldsymbol{c}$ are not included, the following condition:

$$\frac{dc'_i}{dc_j} = 0 \tag{5.36}$$

is applied. Let us also assume that $P(\boldsymbol{c'})$ is a probability function for $\boldsymbol{c'}$s. The likelihood function is constructed from the distribution function of the data $\boldsymbol{x}$ and the external information about the un-fitted Wilson coefficients:

$$L = P(\boldsymbol{x})P(\boldsymbol{c'}) \tag{5.37}$$

Assuming the data is distributed as a multivariate Gaussian distribution (eq. 5.28), we substitute eq. 5.35 into eq. 5.28 and it yields:

$$L = \frac{\exp(-\frac{1}{2}(\boldsymbol{Ac} + \boldsymbol{A'c'} - \boldsymbol{d})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Ac} + \boldsymbol{A'c'} - \boldsymbol{d}))}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} P(\boldsymbol{c'}) \tag{5.38}$$

Since $\frac{dc'_i}{dc_j} = 0$, then $\frac{dP(\boldsymbol{c'})}{dc_i} = 0$. As a result, the negative log likelihood function becomes (ignoring constant terms and the terms which do not depend on $\boldsymbol{c}$):

$$-2\log L = (\boldsymbol{Ac} + \boldsymbol{A'c'} - \boldsymbol{d})^T \boldsymbol{M} (\boldsymbol{Ac} + \boldsymbol{A'c'} - \boldsymbol{d}) \tag{5.39}$$

which is independent of external conditions on $\boldsymbol{c'}$.

The equivalent interpretation of the equation 5.39 is that the floating Wilson coefficients $\boldsymbol{c}$ are fitted to the data, which is corrected by the un-fitted Wilson coefficients instead of directly to data $\mathbf{d}$:

$$\boldsymbol{d} \longrightarrow \boldsymbol{d} - \boldsymbol{A'c'} \tag{5.40}$$

Substituting eq. 5.40 to eq. 5.34[4] yields that the fitted Wilson coefficient vector is:

$$\begin{aligned} \boldsymbol{c} &= (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{M}(\boldsymbol{d} - \boldsymbol{A'}\boldsymbol{c'}) \\ &= (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{d} - (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A'}\boldsymbol{c'} \end{aligned} \tag{5.41}$$

which is a linear function of $\boldsymbol{c'}$. Let us denote the prefactors of the un-fitted Wilson coefficients as $\boldsymbol{K}$:

$$\boldsymbol{K} = -(\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A'} \tag{5.42}$$

There is no correction for the Hessian uncertainty on $\boldsymbol{c}$ since the Hessian does not change because of $\frac{dc'_i}{dc_j} = 0$:

$$\frac{\partial^2 \log L(\boldsymbol{c}, \boldsymbol{c'})}{\partial c_j \partial c_i} = \frac{\partial^2 \log L(\boldsymbol{c})}{\partial c_j \partial c_i} \tag{5.43}$$

The eq. 5.41 can also be used to evaluate the uncertainties on the assumption when un-fitted Wilson coefficients are set to zero.

Practical application of the "Wilsons as functions" method

For a single coefficient c_0 eq. 5.41 is:

$$c_0 = \omega_0 + \sum_i k_i c'_i \tag{5.44}$$

where ω_0 is its central value and $k_i \equiv K_{0i}$. It can be seen either as a linear equation of multiple variables or as a hyperplane.

The evaluation of uncertainties of c'_i depends on the distributions of c'_i. If c'_i are uniformly distributed, e.g. within $\pm b_i$ where b_i is a real number, then the correction is $\max(\sum_i k_i c'_i)$:

$$\max\Big(\sum_i k_i c'_i\Big) = \sum_i |k_i b_i| \tag{5.45}$$

In practice this would be the case when no additional information is known about the c'_i and its maximum expected range (typically $\sim$10 or 4π) is used.

[4]Minimisation with respect to a subset of random variables is allowed as described in [127].

In case the c_i' are Gaussian distributed, then $\sum_i k_i c_i'$ is treated as a sum of Gaussian variables and added in quadrature[5]. In practice the Gaussian constraints are likely to come from an external global fit.

In case there is a mixture of uniform and Gaussian distributed variables, the new uncertainty is:

$$\sqrt{\sigma_0^2 + \sum_i^m (k_i\sigma_i)^2} + \sum_i^{N-m} |k_i b_i| \tag{5.46}$$

where σ_0^2 is uncertainty of c_0 obtained from the Hessian matrix, m is the number of un-fitted Gaussian distributed coefficients and σ_i are their standard deviations, N is the total number of un-fitted coefficients, b_i are the upper/lower limit of un-fitted uniformly distributed coefficients.

The shift of the central value is obtained by substituting central values to equation 5.44.

5.5 Discussion

The main four aspects of an experimental EFT analysis are the choice of measurements, the modelling of an EFT, accounting for the coefficients which cannot be constrained by the given set of measurements, and a construction of a statistical model. The measurements in this thesis are seen as a model independent source of information about Higgs physics and are not optimised for a specific EFT model[6]. Here EFT is used as a framework to interpret the measurements rather than tune the measurements for a specific EFT model. There are two main reasons for that: the main reason is that EFT is model independent and can be used to test specific theoretical models, and the secondary that the current tools are available at LO only

[5]For correlated $\boldsymbol{c'}$ the addition in quadrature would contain a correlation matrix.

[6]Here an EFT model refers to a version of EFT where the versions may differ by the basis, order in QCD, dimensionality of operators, etc. While model independence in the context of EFT refers to no preference to a specific physics model.

and before there is a concrete reason to prefer certain parameters (like cH is section 1.2.1), it is preferred to measure cross-sections independently of EFT models.

Conceptually the modelling of an EFT is quite well defined since the aim is for it to be as general as possible. The aspect which is discussed in this thesis is that of obtaining analytic equations from the MC samples. This chapter discussed how analytic EFT equations are obtained when EFT dependence enters both via the couplings and the propagator widths, where the latter is developed as part of this thesis. This aspect requires mathematical rigour rather than particular assumptions about physics.

The remaining two aspects: accounting for the un-fitted EFT parameters and the construction of a statistical model depend on assumptions about physics. After the development of the "Wilsons as functions" method (as part of this thesis), the discussion about model dependence became rather straightforward. In case the un-fitted Wilson coefficients are set to zero, there might be a strong model dependence induced and the results correspond to only a subset of physics models. More accurate constraints can be obtained using the "Wilsons as functions" method, where the results can be used for any set of theory interpretations, only limited by the EFT model.

The likelihood function is defined based on the statistical modelling of the measurements. The deviations from the SM are parameterised with Wilson coefficients. The additional external knowledge about the Wilson coefficients would typically enter the likelihood function as a multiplicative factor. For example, propagator widths depend on the Wilson coefficients while the analytic EFT equations (as well as the EFT models used to derive these equations) do not contain a priori information about the known constraints on the propagator widths (in the case of the models used in this thesis: the Higgs, W, Z bosons' and the top quark's widths). Typically EFT fits do not include this information about propagator widths, therefore

the resulting constraints on Wilson coefficients include a subset of physics models, which are not compatible with the known constraints on propagator widths. This motivates a careful construction of the likelihood function. The two choices which have to be made are whether the dependence on propagator widths in the EFT equations is included and whether the known constraints on propagator widths and their dependence on Wilson coefficients are added to the likelihood function as external constraints. This well illustrates how the construction of the likelihood function corresponds to different physics models and that the external constraints should be consistent with the EFT modelling.

In case both the dependence on propagator widths in the EFT equations and the external constraints on the widths are not included, the fit results would correspond to a set of models where propagator widths are close to the SM values. In case the dependence on propagator widths is included in the EFT equations without an additional information about the existing constraints on the propagator widths, the fit results would include the models which are inconsistent with the known measurements of the Higgs, W, Z bosons' and the top quark's widths. If the external information about the widths is included in addition, this restricts results to the models which are consistent with the measurements of particle widths. Finally, in case the external constraints on propagator widths are included but the EFT equations are produced assuming the SM widths, theoretically this is an inconsistent modelling (practically it depends on the uncertainties on propagator widths).

In the case of a global fit, the MC samples should be translated into analytic expressions including full dependence on Wilson coefficients via couplings, propagator widths, other parameters, and all available (and not-overlapping) measurements should be included. While in the case of local fits, the construction of a likelihood function should reflect the ultimate aim of the interpretation.

Chapter 6

Effective Field Theory interpretations

The Simplified Template Cross Section (STXS) measurements, like the one presented in chapter 4, are a source of information about the limits on new physics, especially via the Higgs couplings. Moreover, the STXS framework includes kinematic regions sensitive to Beyond Standard Model (BSM) physics, and the cross-sections are measured at particle level (section 1.4.2), so BSM Monte Carlo (MC) predictions do not require computationally intensive detector simulation. The Effective Field Theory (EFT) framework offers an interpretation of these measurements in a model independent way. This chapter presents EFT interpretations of the combined Higgs boson measurements in the STXS framework using the methodology presented in chapter 5.

This EFT analysis targets the modifications of the Higgs boson couplings to strong and electroweak gauge bosons, and to the top quark. The constraints on EFT coefficients are obtained using the SMEFT (section 1.2) EFT with dimension-6 operators, written in the strongly-interacting light Higgs (SILH) and Warsaw bases (section 5.1). The corresponding models in these bases are the Higgs Effective Lagrangian (HEL) and SMEFTsim (section 5.1), respectively. The interpretation is performed on the combined $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels and the

combination of $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ [26], and $H \to WW^*$ (chapter 4) STXS measurements. In addition, an EFT analysis performed on public measurements is compared with a fit to event yields, which contains full experimental information, available to the ATLAS Collaboration only.

This chapter presents the first ATLAS EFT interpretation of the combined Higgs measurements. Constraints on an effective Lagrangian from the combined $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels are obtained using the HEL model. This work is made public in an ATLAS note [3] and presented in section 6.1. In order to evaluate if any useful information is lost when an EFT interpretation is performed on a public measurement instead of using the full ATLAS information, the EFT interpretation of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurements [26] with the HEL model is made public in a LHC Higgs Cross Section Working Group public note [5] and described in section 6.2. The availability of the new SMEFTsim EFT model led to the interpretation of the same STXS measurements, which is presented in section 6.3. Finally, the EFT interpretation is performed using the SMEFTsim EFT model on the combination of the $H \to WW^*$ STXS measurement with the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurement. This interpretation is described in section 6.4. The interpretations are summarised in Table 6.1. All EFT interpretations presented in this chapter are performed as part of this thesis. A general discussion about EFT fits is presented in section 6.5.

6.1 EFT interpretation of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ combined channels using the HEL model

This section presents the first ATLAS EFT fit to the combined $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels. A dedicated analysis design is created for this study, which was later used in other studies beyond this thesis. This includes a selection of a

Table 6.1: Summary of the EFT interpretations presented in this chapter.

Channel	Fit to	EFT model	Sections	Made public
$H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$	Full ATLAS information	HEL	6.1	[3]
	Public result [26]	HEL	6.2	[5]
	Public result [26]	SMEFTsim	6.3	-
$H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$	Public result [26] and the measurement performed for this thesis	SMEFTsim	6.4	-

subset of Wilson coefficients, derivation of dedicated EFT equations, and multiple validation tests. The results are made public in [3].

The presentation of the analysis starts with the data set used for the interpretation, which is described in section 6.1.1. The production of EFT equations is presented in section 6.1.2. Section 6.1.3 describes how a full set of Wilson coefficients is reduced to a subset that is later used in the analysis. Finally, the fitting procedure is described in section 6.1.4, and the results of the fit are discussed in section 6.1.5.

6.1.1 $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ combination

The combination of $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ ($\ell = e$ or μ) decay channels is obtained using 36.1 fb^{-1} of proton-proton collision data recorded at the ATLAS experiment at the LHC at a centre of mass energy of $\sqrt{s} = 13$ TeV during the period 2015-2016. The analysis is performed for ggF, VBF, VH, and $t\bar{t}$H production processes [26].

The $H \to \gamma\gamma$ analysis consists of 31 exclusive event categories, and the $H \to ZZ^* \to 4\ell$ analysis consists of 9 event categories. The categories are defined such

that they are enriched in the Higgs boson production modes and the kinematic regions corresponding to STXS categories. More details of the experimental analyses are given in [128] and [129].

The event yields are classified into extended STXS stage 1 categories, where the stage 1 kinematic regions in the $qq \to Hqq$ (VBF + VH where V decays hadronically) category are split into pure VBF production, $W(\to q\bar{q}')H$, and $Z(\to q\bar{q}')H$ productions [3].

6.1.2 Production of EFT equations

The EFT processes are simulated at leading order (LO) QCD accuracy using the MadGraph5_aMC@NLO program (section 1.4.3) with the HEL model implementation and the NNPDF2.3 PDF set [95]. The parton shower is simulated using PYTHIA 8 (section 1.4.3) with CTEQ6l.LHpdf PDF set [130]. The ggF production is generated with up to two jets in MadGraph5_aMC@NLO, and matching is performed using the MLM matching scheme (section 1.4.2). A dedicated quadratic EFT parameterisation is produced for the extended STXS stage 1 categorisation based on the principles described in section 5.2.1. The parameterisation is made public in [5].

6.1.3 Parameter reduction in the HEL model

This analysis considers only dimension-6 operators since dimension-5 operators violate lepton and baryon number conservation, while higher than dimension-6 operators are suppressed by higher powers of suppression scale Λ. The data set considered in this analysis is still insufficient to constrain the full dimension-6 operator set and therefore the operator set has to be reduced.

Each production and decay process is affected by a subset of EFT operators, which depends on the vertices in that particular process. In consequence, if all the

vertices associated with a certain Wilson coefficient are not present in the process, these Wilson coefficients do not enter the EFT equations. These Wilson coefficients are removed from the full Wilson coefficients' set first.

In the HEL model, at LO, 15 dimension-6 operators reduce to SM Higgs interactions through $H \to vev$ (vacuum expectation value) and are not directly constrained by precision electroweak data. Four CP-odd operators ($\tilde{\mathcal{O}}_g = |H|^2 G^A_{\mu\nu}\tilde{G}^{A\mu\nu}$, $\tilde{\mathcal{O}}_\gamma = |H|^2 B_{\mu\nu}\tilde{B}^{\mu\nu}$, $\tilde{\mathcal{O}}_{HW} = i\,(D^\mu H)^\dagger\,\sigma^a (D^\nu H)\tilde{W}^a_{\mu\nu}$, $\tilde{\mathcal{O}}_{HB} = i\,(D^\mu H)^\dagger\,(D^\nu H)\tilde{B}_{\mu\nu}$) are neglected since they do not have any SM and BSM interference contribution to STXS. The operators $\mathcal{O}_\ell = y_\ell |H|^2 \bar{L}_L H \ell_R$ and $\mathcal{O}_d = y_d |H|^2 \bar{Q}_L H d_R$ correspond to lepton and down-type Yukawas, respectively, and they are not directly constrained by the $H \to ZZ^* \to 4\ell$ and H$\to \gamma\gamma$ processes. The operator $\mathcal{O}_6 = \left(H^\dagger H\right)^3$ corresponds to di-Higgs production and has a negligible contribution to STXS. Operator $\mathcal{O}_H = \frac{1}{2}\left(\partial^\mu |H|^2\right)^2$ normalises the Higgs field and produces a global change in rates, which is not well constrained with the current precision of the data. These operators are not included in the fit [5].

At LO, the Peskin-Takeuchi S parameter [123] is proportional to `cWW+cB`, where the corresponding operators are $\mathcal{O}_W = \frac{i}{2}\left(H^\dagger \sigma^a D^\mu H\right) D^\nu W^a_{\mu\nu}$ and $\mathcal{O}_B = \frac{i}{2}\left(H^\dagger D^\mu H\right)\partial^\nu B_{\mu\nu}$ [124]. The S parameter is constrained by the precision electroweak data, so the linear combination `cWW+cB` is set to 0 and a linear combination `cWW−cB` is used in the fit.

The remaining operators are $\{\mathcal{O}_g,\ \mathcal{O}_\gamma,\ \mathcal{O}_u,\ \mathcal{O}_{HW},\ \mathcal{O}_{HB},\ \mathcal{O}_W,\ \mathcal{O}_B\}$ and their expressions are listed in Table 6.2. The HEL model implementation redefines the Wilson coefficients into the dimensionless parameters as indicated in Table 6.2 with monospace typing. This EFT analysis aims to constrain the following set of HEL parameters corresponding to the EFT operators: {`cG`, `cA`, `cu`, `cHW`, `cHB`, `cWW−cB`}.

Table 6.2: The HEL model operators constrained by the fit for EFT parameters and their relevant vertices.

Operator	Expression	HEL coefficient	Vertices
$\mathcal{O}_g$	$\lvert H\rvert^2 G^A_{\mu\nu} G^{A\mu\nu}$	$\texttt{cG} = \frac{m_W^2}{g_s^2}\bar{c}_g$	Hgg
$\mathcal{O}_\gamma$	$\lvert H\rvert^2 B_{\mu\nu} B^{\mu\nu}$	$\texttt{cA} = \frac{m_W^2}{g'^2}\bar{c}_\gamma$	$H\gamma\gamma, HZZ$
$\mathcal{O}_u$	$y_u \lvert H\rvert^2 \bar{u}_l H u_R +$ h.c.	$\texttt{cu} = v^2 \bar{c}_u$	$Ht\bar{t}$
$\mathcal{O}_{HW}$	$i\left(D^\mu H\right)^\dagger \sigma^a \left(D^\nu H\right) W^a_{\mu\nu}$	$\texttt{cHW} = \frac{m_W^2}{g}\bar{c}_{HW}$	HWW, HZZ
$\mathcal{O}_{HB}$	$i\left(D^\mu H\right)^\dagger \left(D^\nu H\right) B_{\mu\nu}$	$\texttt{cHB} = \frac{m_W^2}{g'}\bar{c}_{HB}$	HZZ
$\mathcal{O}_W$	$i\left(H^\dagger \sigma^a D^\mu H\right) D^\nu W^a_{\mu\nu}$	$\texttt{cWW} = \frac{m_W^2}{g}\bar{c}_W$	HWW, HZZ
$\mathcal{O}_B$	$i\left(H^\dagger D^\mu H\right) \partial^\nu B_{\mu\nu}$	$\texttt{cB} = \frac{m_W^2}{g'}\bar{c}_B$	HZZ

6.1.4 Statistical modelling

The set of six EFT parameters {cG, cA, cu, cHW, cHB, and cWW−cB} is fitted to event yields in the extended STXS stage 1 categories and using the data set described in section 6.1.1. In this section the fit configuration and its validation are explained, and the evaluation of uncertainties is described.

Fit configuration

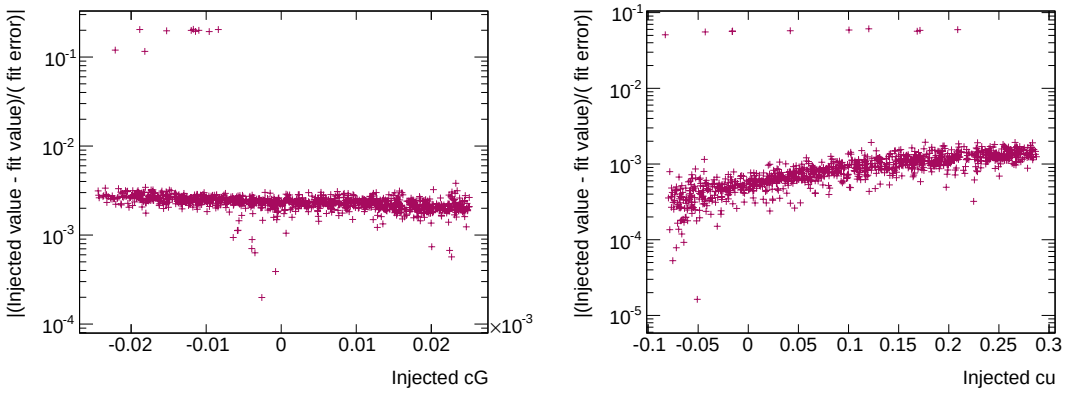
The statistical procedure for the EFT analysis is similar to the statistical procedure described in section 4.1.1. In the $H \to ZZ^* \to 4\ell$ case the number of events in each bin is modelled with a Poisson distribution. The $H \to \gamma\gamma$ categories are modelled either as Poisson variables or are described with the unbinned likelihood function based on the number of data events in the bin. The likelihood function includes the modelling of background processes and the full information about individual theoretical and systematic uncertainties. The fit is performed to obtain the most likely values for the EFT parameters.

Validation of the fit: the fit bias test

The fitting procedure relies on the minimisation algorithm's ability to find the correct minimum with the uncertainties much smaller than the total uncertainties on the parameters of interest. The uncertainties are tested by injecting artificially produced data sets, which are generated using the EFT equations and randomly generated sets of EFT parameters. Data sets generated in this way are not contaminated with the fluctuations induced by the mapping of MC samples to EFT equations, and directly tests the fit's ability to find the correct minimum. The goodness of the fit machinery is determined based on the absolute value of the difference between the fitted and the injected values divided by the uncertainty of the fitted value; let us denote this ratio as f. The fit's tolerance is set to 10^{-7}, where the tolerance indicates when the fit should stop iterating. The tolerance is defined to be the maximum expected distance to the minimum in the parameter space. The results are presented in Figure 6.1, where 1000 artificially produced data sets are used. A reliable fit machinery should produce low f values, preferably smaller than 10^{-2}. Less than 1% of fits to these data sets do not satisfy this criterion since the fit failed to reach the requested tolerance. The results based on the remaining 99% of the data sets show that f value is typically between 10^{-2} and 10^{-3}. The fits presented later in this section are required to satisfy the tolerance bound to ensure there is no significant fit bias.

Derivation of uncertainties from negative log-likelihood scans

The uncertainties are estimated by performing a negative log-likelihood L scan, where $-2 \ln L = 1$ corresponds to 1σ deviation. Event yields are modelled as Poisson variables, which at large enough yield values can be approximated as Gaussian. Therefore, when a quadratic EFT parameterisation is used, a stable negative log-likelihood scan is expected to be a quartic polynomial. In the case of double minima,



(a) (b)

Figure 6.1: Fit results to STXS values corresponding to non-zero input EFT parameters using the HEL model. Shown is the base-10 logarithm of the absolute value of the difference between the measured and injected values, divided by the measurement uncertainty, for the EFT parameters (a) `cG`, (b) `cu`.

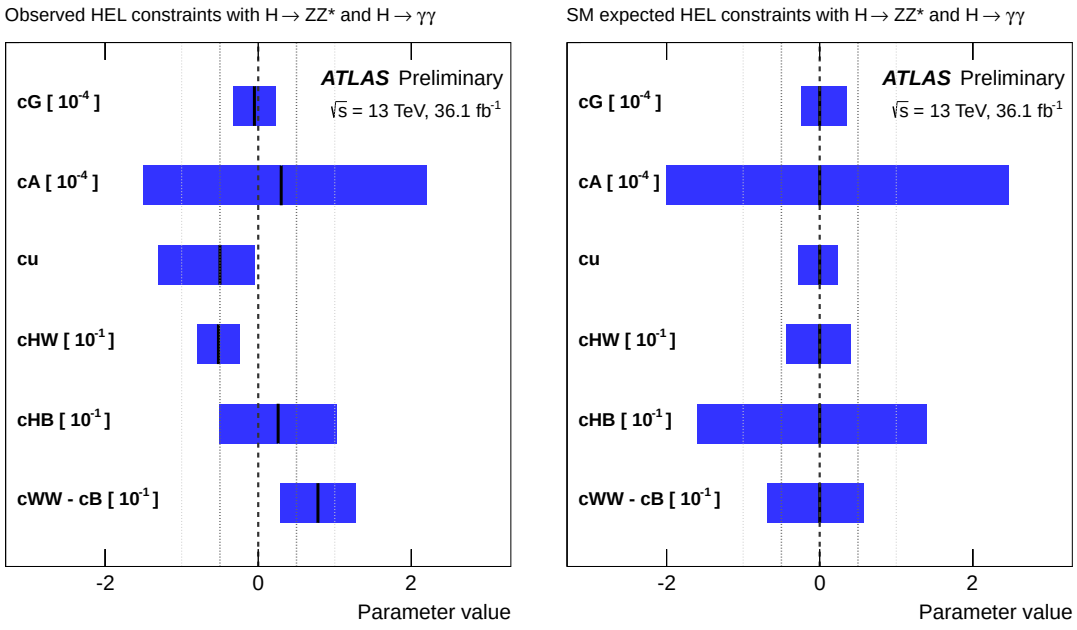
the solution that is closer to the SM value is chosen.

6.1.5 Fit results

The most likely EFT parameters' values and their uncertainties are presented in Figure 6.2 and Table 6.3. The uncertainties on EFT parameters are estimated using negative log-likelihood scans in Figure 6.3. The correlation matrix of the EFT parameters is presented in Figure 6.4.

The central values of the EFT parameters are consistent with the patterns observed in the STXS measurements in [26]. The coefficient `cG` is measured close to the SM value since the inclusive ggF production cross-section is measured close to the SM value [26]. The ratio of branching fractions $BR(H \to \gamma\gamma)/BR(H \to 4\ell)$ is measured below the SM expected value, leading to a positive `cA` since $BR(H \to \gamma\gamma)/BR(H \to 4\ell) - 1 \propto -\texttt{cA}$ to first order in the coefficients; the measured VBF cross-section is above the SM expected value, leading to a positive `cWW` and a negative `cHW`, since $\sigma(VBF)/\sigma_{SM}(VBF) - 1 \approx \texttt{cWW} - \texttt{cHW}$; and the measured $t\bar{t}H$ cross-section is below

the SM expected value, leading to a negative cu since $\sigma(t\bar{t}H)/\sigma_{SM}(t\bar{t}H) - 1 \approx$ cu to first order. Also, the coefficient cu has asymmetric uncertainty since in Figure 6.3 (c) the negative log-likelihood scan has a local maximum below 1σ line, and the solution closer to SM value is quoted.



(a) (b)

Figure 6.2: The (a) observed and (b) SM predicted best-fit values and 68% C.L. intervals for the six-parameter fit to the event yields. The presented results are produced by the author of the thesis and are made public in [3].

Compared to the SM expectation, the sensitivity to the dimension-6 HVV (Higgs and vector bosons) couplings (cHW, cHB, cWW-cB) is higher because the measured VBF cross-section is higher than (though still compatible with) the SM prediction [26]. For the $t\bar{t}H$ coupling cu the observed sensitivity is lower than expected since the low observed yield [26] effectively removes the degeneracy and gives a flatter negative log-likelihood profile.

Figure 6.5 demonstrates EFT parameters' individual impacts on the STXS regions. The value of each parameter is set to the expected $\approx +1\sigma$ deviations assuming

Table 6.3: The fit values and 68% C.L. uncertainties from the fit for six EFT parameters using the STXS stage 1 framework to relate the data to the parameters. The observed and SM expected results are shown. The results are produced by the author of this thesis and the table is made public in [3].

Operator	Fit result (observed)	Fit result (SM expected)
$\mathcal{O}_g$	$\texttt{cG} = -0.05^{+0.27}_{-0.28} \times 10^{-4}$	$\texttt{cG} = 0.00^{+0.38}_{-0.26} \times 10^{-4}$
$\mathcal{O}_\gamma$	$\texttt{cA} = 0.3^{+1.9}_{-1.8} \times 10^{-4}$	$\texttt{cA} = 0.0^{+2.8}_{-2.2} \times 10^{-4}$
$\mathcal{O}_u$	$\texttt{cu} = -0.50^{+0.45}_{-0.81}$	$\texttt{cu} = 0.00^{+0.24}_{-0.28}$
$\mathcal{O}_{HW}$	$\texttt{cHW} = -0.052 \pm 0.028$	$\texttt{cHW} = 0.000^{+0.041}_{-0.043}$
$\mathcal{O}_{HB}$	$\texttt{cHB} = 0.026 \pm 0.077$	$\texttt{cHB} = 0.00^{+0.14}_{-0.16}$
$\mathcal{O}_W$, $\mathcal{O}_B$	$\texttt{cWW} - \texttt{cB} = 0.078 \pm 0.049$	$\texttt{cWW} - \texttt{cB} = 0.000^{+0.057}_{-0.074}$

the SM. The similar shape and amplitude impact profile of the coefficients cHW and cWW − cB results in strong anticorrelation of these coefficients (Figure 6.4).

The EFT parameters {cHW, cHB, cWW-cB} can be further improved by including other VBF Higgs measurements and $H \to WW^*$ decays. Tighter constraints on cu can be obtained by including measurements of $t\bar{t}H$ production in other Higgs boson decay channels.

6.2 Fit to STXS measurements using the HEL model

This section presents the fit to the ATLAS STXS measurement [26] using the HEL model. This ATLAS STXS measurement consists of 11 measurements. Unlike in the previous section, in this case only combined information about the uncertainties is known. The EFT is parameterised in quadratic equations as before. This study investigates whether there are advantages to constraining the EFT parameters using full ATLAS information versus using the already available public results derived from the same data set. The results are made public in [5].

The measurement used in this analysis is described in section 6.2.1. It is followed

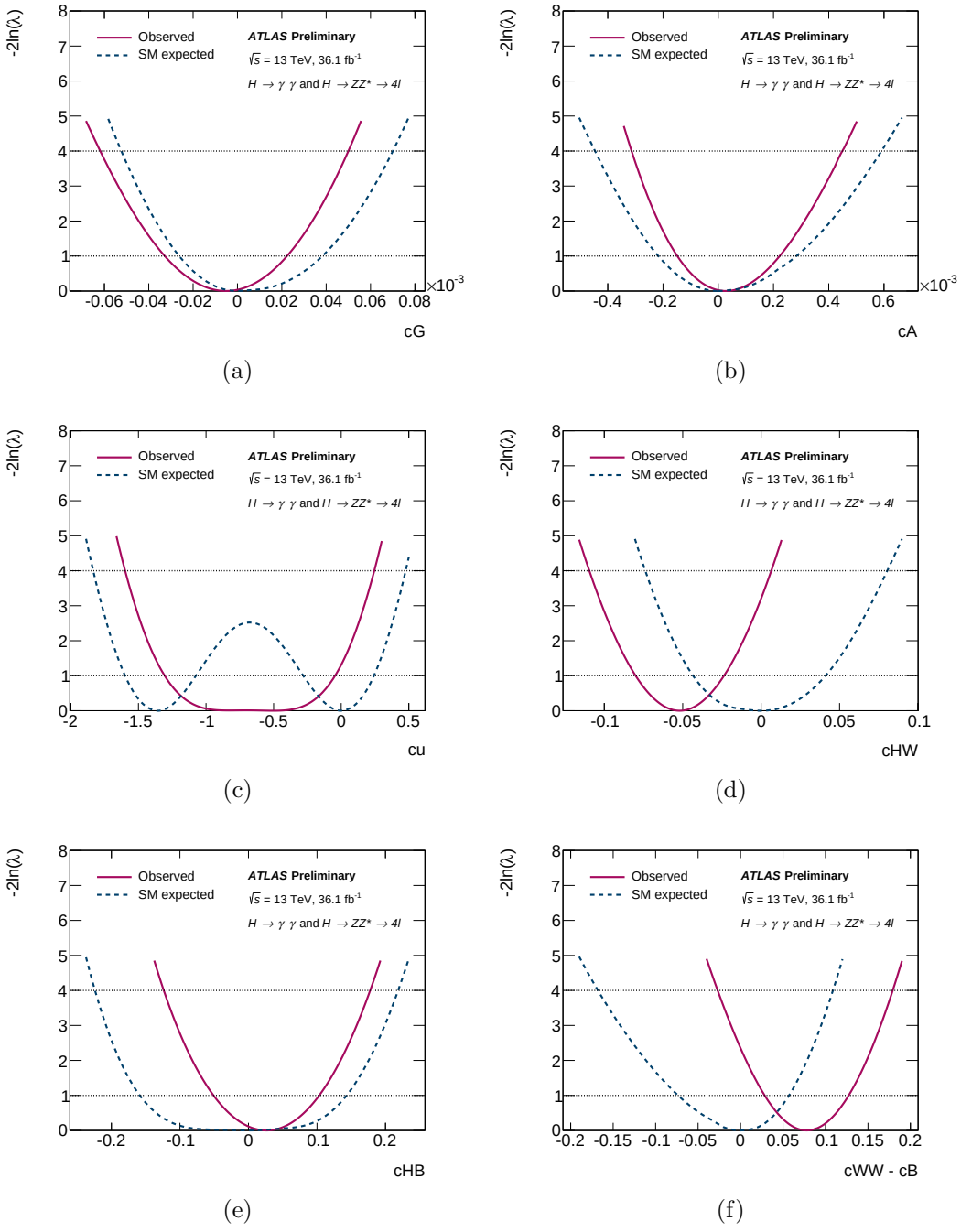


Figure 6.3: The observed (solid) and SM expected (dashed) profiled negative log-likelihood scans for the six-parameter fit to the event yields. The parameters are (a) `cG`, (b) `cA`, (c) `cu`, (d) `cHW`, (e) `cHB`, and (f) `cWW` − `cB`. The plots are produced by the author of this thesis and are made public in [3].

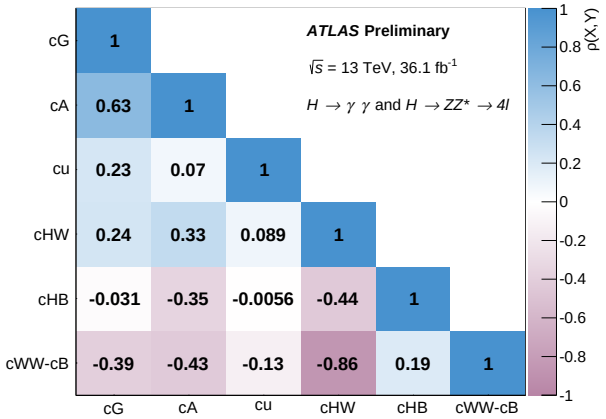


Figure 6.4: The observed correlation matrix from six-parameter fit using HEL model. The plot is produced by the author of this thesis and is made public in [3].

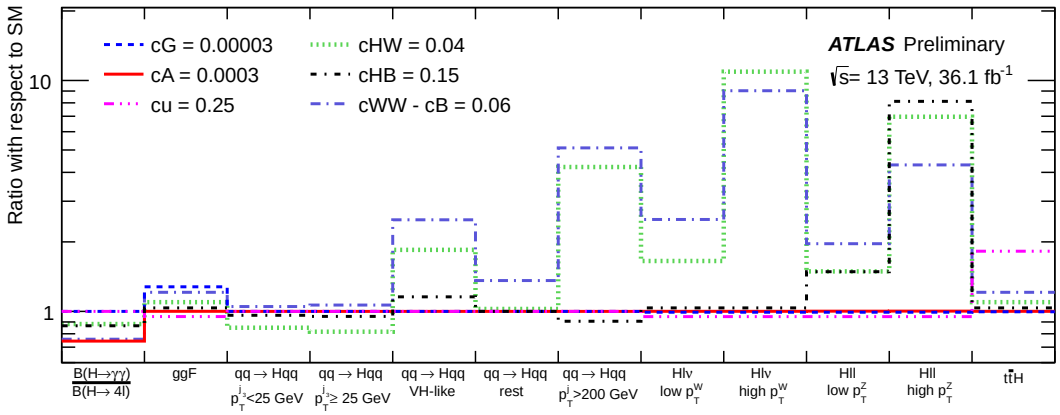


Figure 6.5: The values of modified STXS regions, relative to the SM, for $\approx +1\sigma$ expected SM values of `cG` (3×10^{-5}), `cA` (3×10^{-4}), `cu` (0.25), `cWW` $-$ `cB` (0.06), `cHW` (0.04), and `cHB` (0.15). The STXS regions have been modified by merging all ggF regions, which are affected in the same way by all parameters, and by merging VH categories with $p_T^V > 150$ GeV ("high p_T^W" and "high p_T^Z"). The plot is made public in [3].

by statistical modelling in section 6.2.2, and finally the results are discussed in section 6.2.3.

6.2.1 The combined STXS measurement

The combined ATLAS STXS measurement [26] is performed on the data set described in section 6.1.1. This data set is not sensitive to the full STXS stage 1 categorisation, as a result a merging scheme was constructed. The merging scheme is designed to increase sensitivity to SM production and minimise correlations. The ggF two or more jets regions with $p_T^H < 200$ GeV are merged together. The ggF one or more jets regions with $p_T^H \geqslant 200$ GeV and VBF regions with the leading jet $p_T \geqslant 200$ GeV are used to define two quantities: one corresponds to the sum of the three regions, while the other is defined as the difference of ggF regions and the VBF region. VBF regions with the leading jet $p_T < 200$ GeV are merged. All VH categories are merged into one; also $t\bar{t}$H and tH productions are combined together. The merging scheme is summarised in Figure 6.6.

Eleven STXS measurements are produced and the results are summarised in Figure 6.7. The 11th measurement corresponding to the difference between ggF $p_T^H \geqslant 200$ GeV and VBF leading jet $p_T \geqslant 200$ GeV is $1.7^{+1.7}_{-1.5}$ pb and is not presented in the plot due to the low experimental sensitivity.

6.2.2 Statistical modelling

The likelihood fit is constructed from a multivariate Gaussian distribution:

$$G(\boldsymbol{x}) = \frac{\exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{d})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{d}))}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}} \tag{6.1}$$

where $\boldsymbol{x}$ is a k dimensional vector corresponding to EFT equations, $\boldsymbol{d}$ is the vector of measured STXS values, and $\boldsymbol{\Sigma}$ is a covariance matrix of the STXS measurement and $|\boldsymbol{\Sigma}|$ is its determinant. The negative logarithm of the likelihood function in equation 6.2, where $L(\boldsymbol{x}) = G(\boldsymbol{x})$:

$$-2\log L(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{d})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{d}) + \text{const} \tag{6.2}$$

is minimised using the RooFit data analysis framework [131].

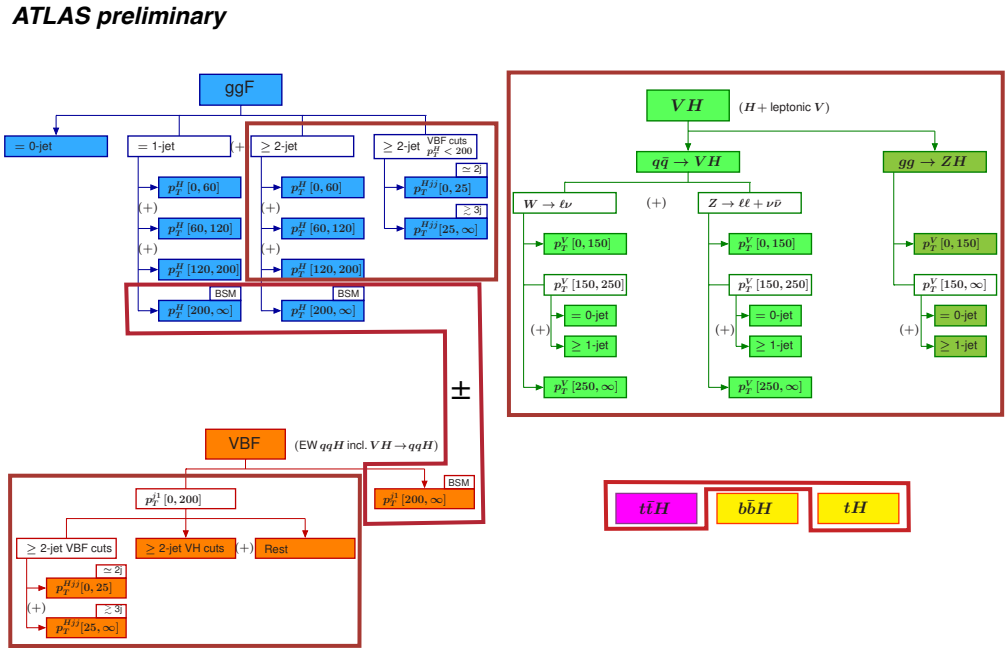
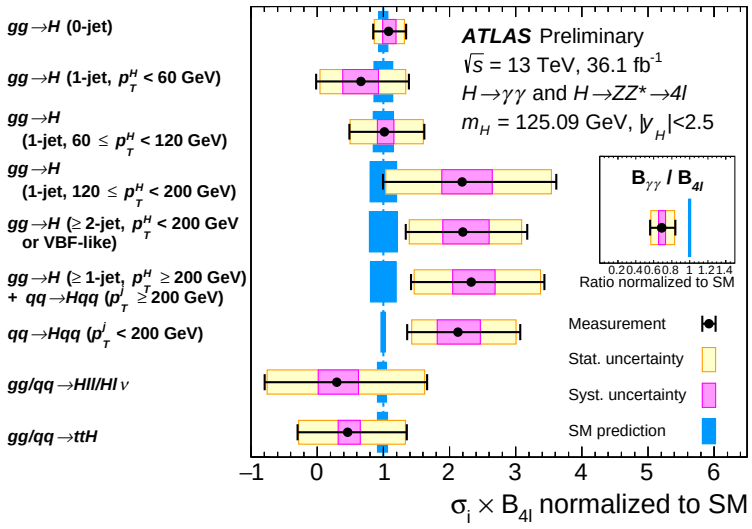


Figure 6.6: The merging scheme of the STXS stage 1 regions applied for the Higgs combined measurement presented in [26]. The regions enclosed by the red boxes are merged, unless there is a ± sign which indicates the sum and the difference of enclosed regions. The plot is taken from [26].

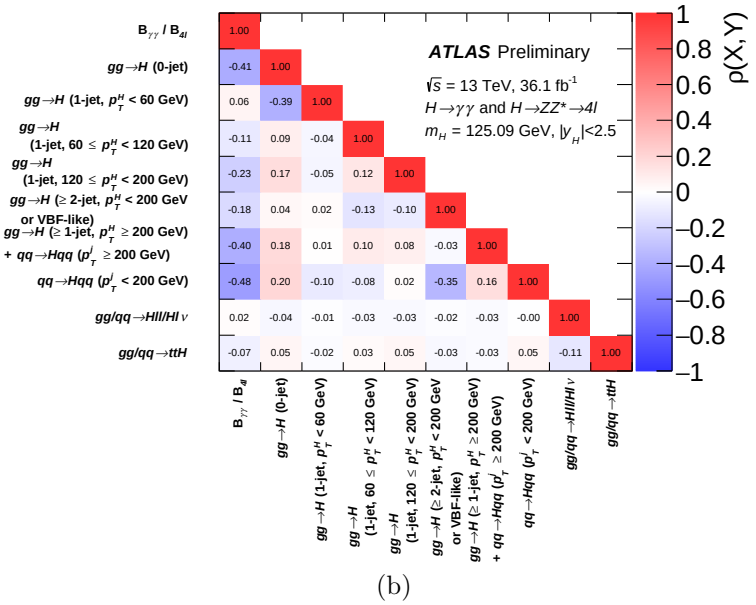
6.2.3 Fit results

In the fit to these STXS measurements, only five parameters are constrained (Figure 6.8) since the sensitivity to one of the parameters `cHB`, `cHW` or `cWW-cB` is lost due to the merging of WH and ZH processes. The parameter `cHB` is chosen to be removed from the final result since it is least constrained.

The sensitivity to `cHB` is recovered when the VH kinematic region is split into WH and ZH regions, which can be seen in the negative log-likelihood scan in Figure 6.9. The likelihood inputs are parameterised with six EFT parameters and a profiled scan is performed for the fit to the extended STXS stage 1 (section 6.1.1), the merged case (section 6.2.1), and the VH-split case, defined here. The log-likelihood scan in



(a)



(b)

Figure 6.7: The Higgs combined $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS stage 1 measurement using a merging scheme. (a) The central values and uncertainties of the POIs and (b) the correlation matrix of these POIs. The plots are taken from [26].

the VH-split case is a half of a close-to-quartic polynomial and passes $-2\ln(\Lambda) = 1$ value, so the uncertainties are well defined (section 6.1.4). In contrast, the log-likelihood scan of the original merged version has an irregular shape and the fit fails

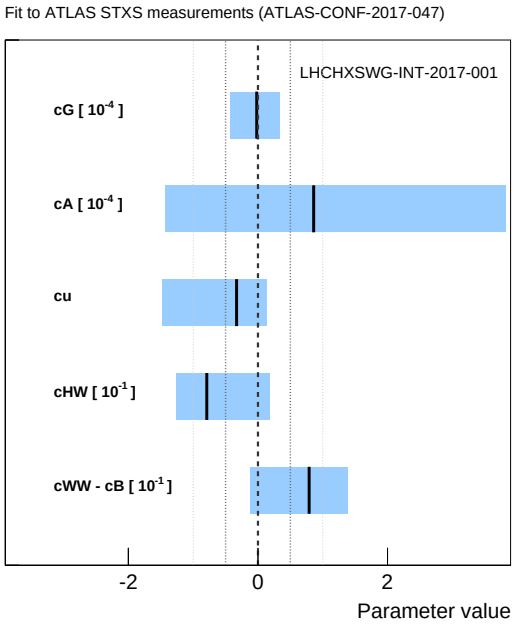


Figure 6.8: The best-fit values and 68% C.L. intervals for the five-parameter fit to the STXS measurements. The plot is produced by the author of this thesis and is made public in [5].

to converge before reaching $-2\ln(\Lambda) = 1$: the `cHB` coefficient cannot be constrained at 68% C.L. Similar behaviour is observed in the log-likelihood scans of the remaining five coefficients. As mentioned in the previous paragraph, the problematic aspect is not the `cHB` parameter but the fit to six rather than five coefficients. This study implies that it is valuable to measure WH and ZH productions separately in which case the EFT results produced using full ATLAS information and only the final measurements would be very similar.

6.3 Fit to STXS measurements using the SMEFTsim model

The availability of the SMEFTsim model motivated the reinterpretation of the ATLAS combined Higgs $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurement [26] presented in section 6.2. The advantage of the SMEFTsim is that it is a full imple-

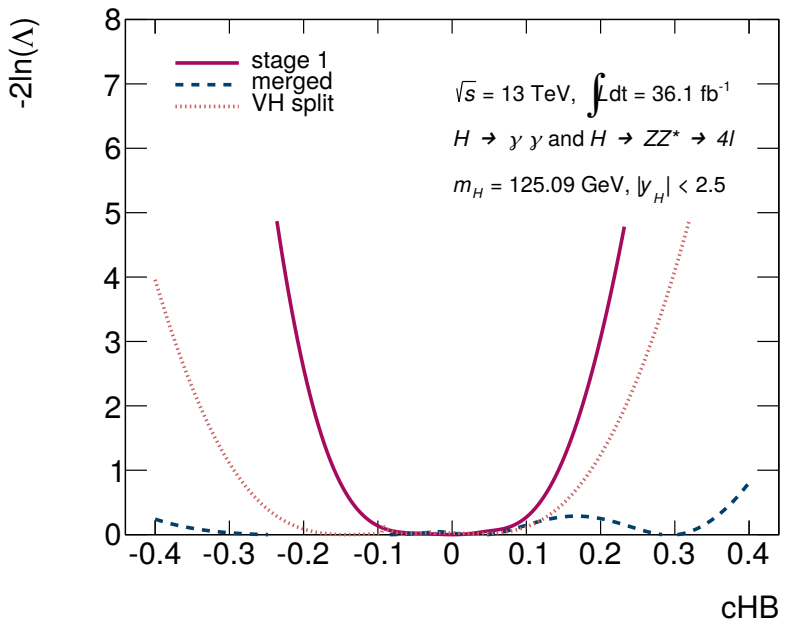


Figure 6.9: Profiled negative log-likelihood scan for modified STXS stage 1 binning (solid), ATLAS merged public measurement (dashed), and a modified STXS stage 1 merging where VH production is split into WH and ZH bins (dotted). The likelihood inputs are parameterised with six EFT coefficients. When VH is split into WH and ZH, the sensitivity to the EFT parameter `cHB` is recovered.

mentation of a renormalisable basis. The dedicated EFT equations are derived as described in section 6.1.2, except that the SMEFTsim model is used. The EFT is modelled as linear equations.

This section starts with the selection of a set of Wilson coefficients, which is described in section 6.3.1. The fit procedure is presented in 6.3.2, and the results are discussed in 6.3.3.

6.3.1 The selection of SMEFTsim Wilson coefficients

The STXS framework defines Higgs boson production processes, thus the measurements mainly constrain the set of Wilson coefficients corresponding to the vertices which include the Higgs boson. The relevant operators and Higgs boson production processes and decays are listed in Table 6.4.

The Wilson coefficients cuH and cHG are mainly constrained by the $t\bar{t}$H and the ggF productions, while the remaining coefficients from Table 6.4 simultaneously affect VBF and VH productions, and $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ decay widths. VH production split into WH and ZH would increase the distinction of cHW from cHB and $cHWB$, as can be seen in Figure 6.10. Since in the STXS measurement presented in section 6.2.1 VH kinematic regions are merged into one region, VBF regions with leading jet $p_T < 200$ GeV are merged, and the remaining VBF phase space is attached to ggF kinematic region, the ability to constrain one, two or all three of these Wilson coefficients depends on the uncertainties of these measurements. In case the measurement is only able to constrain two out of three Wilson coefficients, then any subset of two will lead to a convergent fit.

Table 6.4: The SMEFTsim model operators aimed to be constrained by the fit for Wilson coefficients and their relevant vertices.

Operator	Expression	SMEFT coefficient	Production/Decay
$\mathcal{O}_{HG}$	$H^\dagger H G^A_{\mu\nu} G^{A\mu\nu}$	cHG	ggF
$\mathcal{O}_{HW}$	$H^\dagger H W^I_{\mu\nu} W^{I\mu\nu}$	cHW	$H\gamma\gamma$, HZZ, HWW, WH, ZH, VBF
$\mathcal{O}_{HB}$	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$	cHB	$H\gamma\gamma$, HZZ, ZH, VBF
$\mathcal{O}_{HWB}$	$H^\dagger H \tau^I W^I_{\mu\nu} B^{\mu\nu}$	$cHWB$	$H\gamma\gamma$, HZZ, ZH, VBF
$\mathcal{O}_{uH}$	$H^\dagger H \bar{q}_p u_r \widetilde{H}$	cuH	$t\bar{t}H$, ggF

6.3.2 Fit procedure

The likelihood fit is constructed from a multivariate Gaussian function and the description is equivalent to that in section 6.2.2. The vector $\boldsymbol{x}$ corresponds to linear EFT equations, so it can be expressed as a linear map $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{c}$ as in section 5.4.1. Therefore, the analytic solution exists as in equation 5.34: $\boldsymbol{c} = (\boldsymbol{A}^T \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{M} \boldsymbol{d}$, where $\boldsymbol{M}$ is the inverse of a covariance matrix of the STXS measurements. The

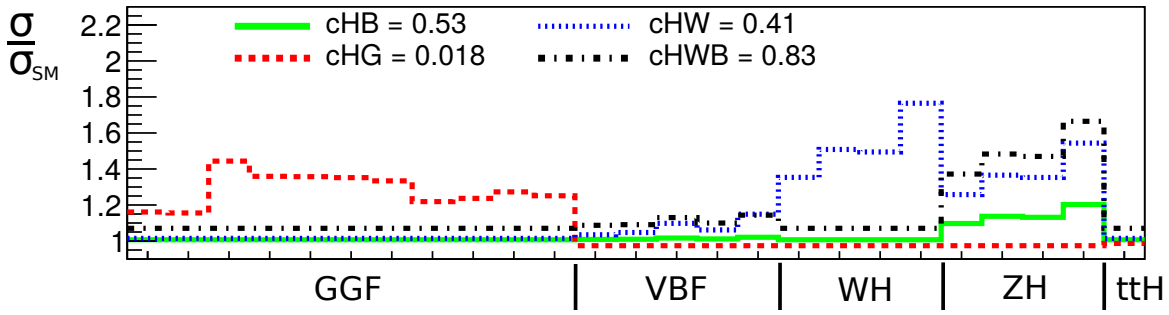


Figure 6.10: The values of STXS stage 1 regions, relative to the SM, for cHB, cHG, cHW, and $cHWB$ Wilson coefficients.

EFT results, presented in this section, are obtained using this analytic procedure instead of using numerical minimisation of the likelihood function.

6.3.3 Fit results

This STXS measurement is able to constrain only four Wilson coefficients. Any two Wilson coefficients from $\{cHB, cHW, cHWB\}$ can be chosen for a convergent fit. The Peskin-Takeuchi S parameter is proportional to $cHWB$ [132], and as in the case of the analysis using the HEL model (section 6.1.3), $cHWB$ can be removed from the fit. The fit result is shown in Figure 6.11, where the subset $\{cHB, cHG, cHW, cuH\}$ is fitted. The cHG coefficient is well constrained due to the large number of well-measured ggF production measurements. The cuH coefficient has a relatively large uncertainty. This is consistent with the constraint on cu in the HEL model case (Figure 6.8) since both coefficients are mainly constrained by the $t\bar{t}$H production. All but cuH Wilson coefficients agree well with the SM, while cuH has only a 1σ agreement with the SM.

The case where the parameter cHB is replaced by $cHWB$ in $\{cHB, cHG, cHW, cuH\}$ is compared to the previously obtained result and presented in Figure 6.12. cHG and cuH have almost not changed, as expected, while the uncertainty on cHW is reduced at the expense of increased uncertainty on $cHWB$ with respect to cHB.

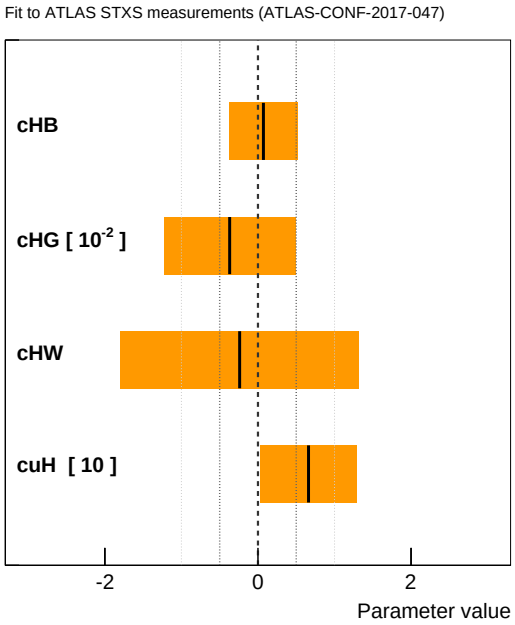


Figure 6.11: The best-fit values and 68% C.L. intervals for the four-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurements from [26].

This disagreement between the ranges of confidence limits is a manifestation of the induced model dependence. These two EFT scenarios intersect only at $cHB = cHWB = 0$. The effect of fitting a subset rather than a full set of Wilson coefficients is evaluated later in this section using "Wilsons as functions" method.

The effect of propagator widths

The effect of propagator widths' parameterisation (section 5.3) is tested by fitting a subset $\{cHWB, cHG, cHW, cuH\}$. From this subset of Wilson coefficients, only $cHWB$ affects the propagator widths, Z boson's width in particular. The two cases are compared in Figure 6.13. Due to the correlations that originate from multiple Wilson coefficients appearing in the same EFT equation, the values of cHB, cHG and cHW are slightly shifted closer to the SM value in the case where propagator widths are fixed to SM values. These three Wilson coefficients are consistent with

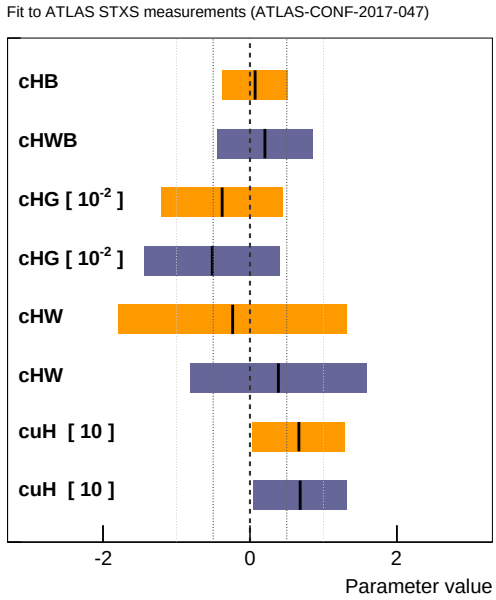


Figure 6.12: The comparison of the best-fit values and 68% C.L. intervals of the four-parameter $\{cHB, cHG, cHW, cuH\}$ (orange) and $\{cHWB, cHG, cHW, cuH\}$ (violet) fits using SMEFTsim model to the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurements.

the SM in both cases. The uncertainties of all parameters remained similar. The implementation of dependence on propagator widths for this set of measurements does not make a significant impact.

Interpretation using "Wilsons as functions" method

This set of measurements is insufficient to constrain the full set of Wilson coefficients. The induced model dependence is addressed with the "Wilsons as functions" method (section 5.4). This method provides the corrections to central values of the fitted Wilson coefficients, where the corrections are expressed as functions of un-fitted Wilson coefficients. The constraints on the un-fitted Wilson coefficients may be taken from other EFT fits, e.g. in [132] a global fit was performed on 20 Wilson coefficients.

The model dependence is evaluated on a four parameter fit of $\{cHB, cHG, cHW,$

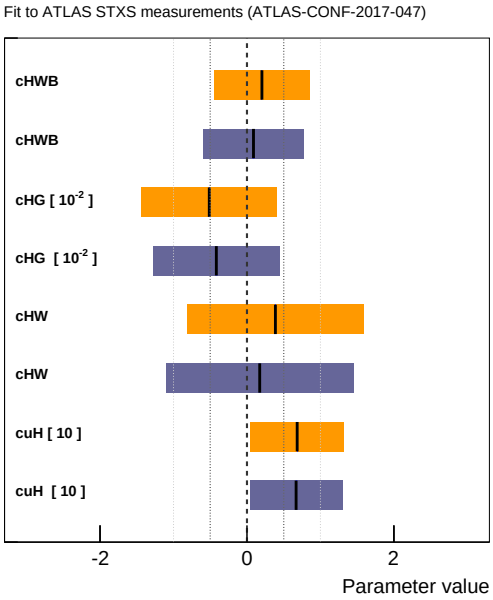


Figure 6.13: The comparison of the best-fit values and 68% C.L. intervals of the four-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurements when the propagator widths' dependence on Wilson coefficients is taken into account (orange) versus when widths are fixed to SM values (violet).

$cuH\}$, and an example fit output for cHW reads like this:

$$\begin{aligned} cHW = &- 0.2 - 6.4 \cdot 10^{-8} \cdot cH - 0.19 \cdot cH\Box - 0.16 \cdot cHDD - 0.51 \cdot cHWB \\ &- 3.4 \cdot 10^{-9} \cdot ceH + 3.6 \cdot 10^{-4} \cdot ceW + \cdot 10^{-6} \cdot ceB - 0.15 \cdot cHl1 \\ &+ 5.1 \cdot cHl3 + 5.1 \cdot 10^{-2} \cdot cHe - 0.22 \cdot cll - 4 \cdot cll1 + 5.6 \cdot 10^{-4} \cdot cuW \\ &- 9 \cdot 10^{-6} \cdot cuB + 2.9 \cdot 10^{-4} \cdot cdW + 1.5 \cdot 10^{-4} \cdot cdB - 3.6 \cdot 10^{-2} \cdot cHq1 \\ &- 7 \cdot cHq3 - 0.26 \cdot cHu + 0.12 \cdot cHd - 8.8 \cdot 10^{-8} \cdot cdH + 8.5 \cdot 10^{-5} \cdot cuG \\ &+ 10^{-8} \cdot cdG + 2.2 \cdot 10^{-12} \cdot cHud - 7.2 \cdot 10^{-3} \cdot cG - 1.2 \cdot 10^{-16} \cdot cquqd1 \\ &+ 4 \cdot 10^{-17} \cdot cquqd8 \end{aligned} \tag{6.3}$$

Equivalently, there is an equation for each of the remaining three coefficients. The fit results expressed in this form are technically model independent.

The correction to the central value of the fitted Wilson coefficients is estimated by substituting the best-fit values of 16 Wilson coefficients from [132] to the "Wilsons

as functions" equations. This correction encapsulates the information about the measurements' used in the global fit central values' deviations from the SM values. The correction due to the 1σ uncertainties on these external Wilson coefficients from the global fit is evaluated by adding the terms in eq. 6.3 in quadrature[1] since the uncertainties from the global fit are extracted from a likelihood fit, which is constructed as a multivariate Gaussian distribution.

The original uncertainties on the four fitted Wilson coefficients are independent of the un-fitted Wilson coefficients since the original uncertainties are derived from a Hessian matrix, which is independent of deviations from the SM. The correction to the original confidence interval is eq. 5.46. The un-fitted coefficients whose constraints are not available from the global fit are treated as uniformly distributed random variables within ± 10. The results are listed in Table 6.5 and compared with the original confidence intervals in Figure 6.14. When only the Gaussian constraints are considered, the confidence intervals for cHW and cHB have increased 1.6 times, while for the cHG it has increased 3.6 times. This is a reasonable increase since here an impact of 16 un-fitted coefficients is considered, and similar corrections are seen when a four parameter fit is extended to only a five parameter fit (section 6.4.1). The correction to cuH shows that the new confidence interval is larger than the expected Wilson coefficients' range of [-10, 10], so in practice cuH is unconstrained. It is consistent with the result obtained in [132], where $cuH = -79 \pm 43$.

These corrections correspond to accounted model dependence and are more accurate representation of actual constraints on Wilson coefficients. However, these constraints still cannot compete with the constraints obtained in a global fit. This method would be more interesting for a subset of measurements that is able to constrain a Wilson coefficient, which was not well constrained before. Then it would be

[1]This is true for independent Gaussian distributed variables; for a correlated set of these variables ($\boldsymbol{c'}$) the estimation would be $\boldsymbol{c'^T M c'}$, where $\boldsymbol{M}$ is a correlation matrix of the Wilson coefficients. $\boldsymbol{M}$ is not provided in [132].

Table 6.5: The best-fit values and 68% C.L. intervals of the four-parameter fit using SMEFTsim model to the combinated STXS measurement in $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels are presented in the middle column. The correction due to the induced model dependence, which is estimated using the "Wilsons as functions" method, is shown in the last column.

Wilson coefficient	Original fit result	"Wilsons as functions" fit result
cHB	0.07 ± 0.45	0.22 ± 0.73
cHG	$(-0.37 \pm 0.86) \times 10^{-2}$	$(-1.1 \pm 3.1) \times 10^{-2}$
cHW	-0.2 ± 1.6	-0.7 ± 2.5
cuH	6.6 ± 6.3	-69 ± 48

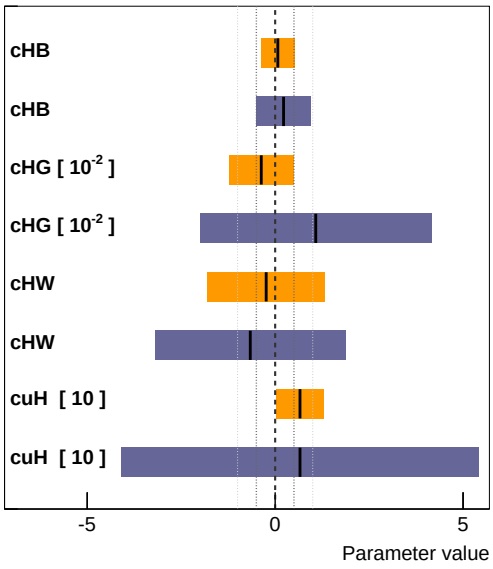


Figure 6.14: The comparison of the best-fit values and 68% C.L. intervals of the four-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ STXS measurement (orange), and the corrections obtained using the "Wilsons as functions" method (section 5.4) using the constraints on the un-fitted parameter values from [132] (violet). In the cuH case, the correction due to the shift of central value is not indicated in the plot since the new limit is already outside the expected Wilson coefficient range of [-10, 10] and hence cuH is practically unconstrained.

sufficient to perform a local fit (i.e. where only a subset of coefficients is fitted) and the uncertainties on the non fitted coefficients could be evaluated with "Wilsons as

functions" approach. This could apply for a cuH in the near future, or cH which corresponds to double-Higgs production. In the case of cH, the estimated cH could be directly used for an estimation of a corrected vacuum expectation value (section 1.2.1) without a need of a full global fit.

6.4 EFT interpretation of $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ combination

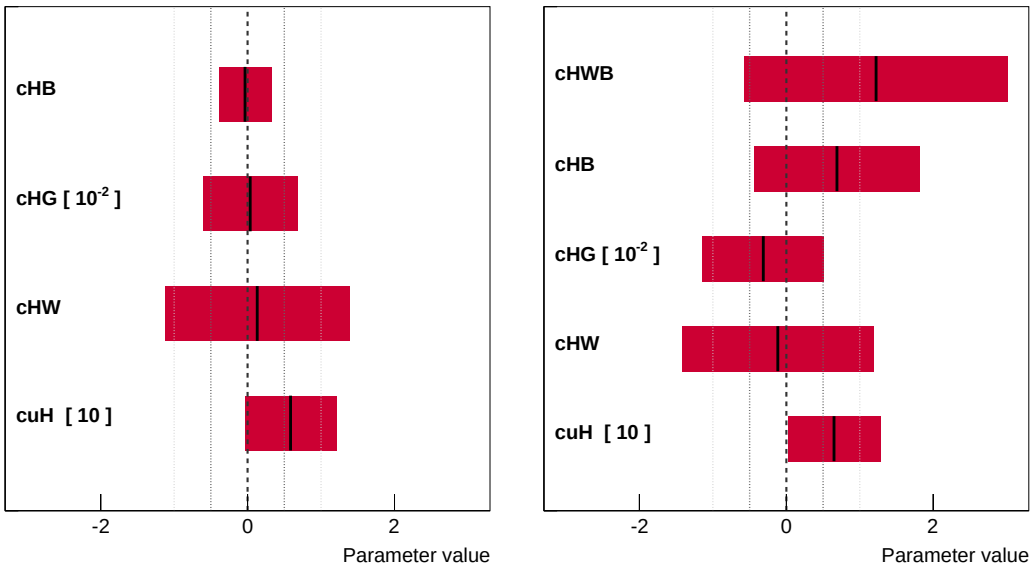
The $H \to WW^*$ STXS measurement, presented in chapter 4, is a valuable source of information about Higgs boson properties. The $H \to WW^*$ STXS measurement using a merging scheme (section 4.4.1) is combined with the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ combined measurement from [26] with the assumption that the correlations between the two sets of measurements are equal to zero. The SMEFTsim model is used in the EFT interpretation, where the parameterisation is expressed as linear equations. The fit procedure is identical to the strategy described in section 6.3.2.

The larger set of measurements is expected to improve the four parameter fit in Figure 6.11 and increase the chances of constraining all five Wilson coefficients listed in Table 6.4. The EFT fit results are presented in the following section.

6.4.1 Fit results

The four parameter subset $\{cHB, cHG, cHW, cuH\}$ is constrained using the combined STXS measurement and is presented in Figure 6.15 (a) and Table 6.6. The fit results using the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ combination are compared to the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ combination in Figure 6.16.

The cHB and cHW Wilson coefficient uncertainties are reduced by 19% and cHG uncertainty is reduced by 25%. The cuH parameter's uncertainty is reduced only slightly (1%) since the added $H \to WW^*$ measurement consists of ggF, VBF



(a) (b)

Figure 6.15: The best-fit values and 68% C.L. intervals for the (a) four-parameter fit and (b) five-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ STXS measurements. The observed results are shown.

Table 6.6: The best-fit values and 68% C.L. intervals of the four-parameter and five-parameter fits using SMEFTsim model to the combination of $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ STXS measurements. The observed results are presented.

Wilson coefficient	Four parameter fit result	Five parameter fit result
cHB	-0.03 ± 0.36	0.7 ± 1.1
cHG	$(0.4 \pm 6.4) \times 10^{-3}$	$(-3.1 \pm 8.2) \times 10^{-3}$
cHW	0.1 ± 1.3	-0.1 ± 1.3
cuH	5.9 ± 6.2	6.5 ± 6.3
$cHWB$	-	1.2 ± 1.8

productions, and $H \to WW^*$ decay, which have negligible cuH dependence. This shows the importance of $H \to WW^*$ measurements for the EFT fits to combined Higgs boson measurements.

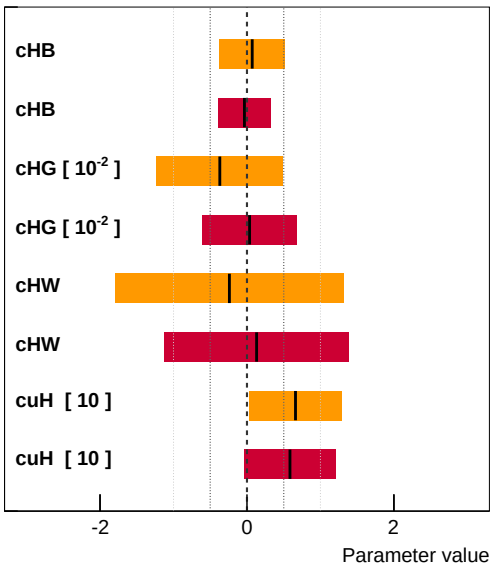


Figure 6.16: The comparison of the best-fit values and 68% C.L. intervals of the four-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ (orange) versus $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ STXS measurements (red). The observed results are shown.

This set of STXS measurements is also able to constrain all five Wilson coefficients listed in Table 6.4. The fit result is presented in Figure 6.15 (b) and Table 6.6. The four and five parameter fits using the set of measurements containing $H \to WW^*$ are compared in Figure 6.17. Since one Wilson coefficient is added using the same set of measurements, the uncertainties of the parameters used in the four parameter fit increased. The most affected parameter is cHB whose uncertainty increased three times. The cHG parameter's uncertainty increased by 28%, while the cHW and cuH coefficients' uncertainties increased only by 4% and 1%, respectively. The increase in uncertainties can be interpreted as an uncertainty on induced model dependence when $cHWB$ is set to zero. The $cHWB$ is constrained mainly on account of cHB and cHG, which is a result of correlations between measurements and the mixing of Wilson coefficients within EFT equations.

The "Wilsons as functions" method (section 5.4) is used to account for the

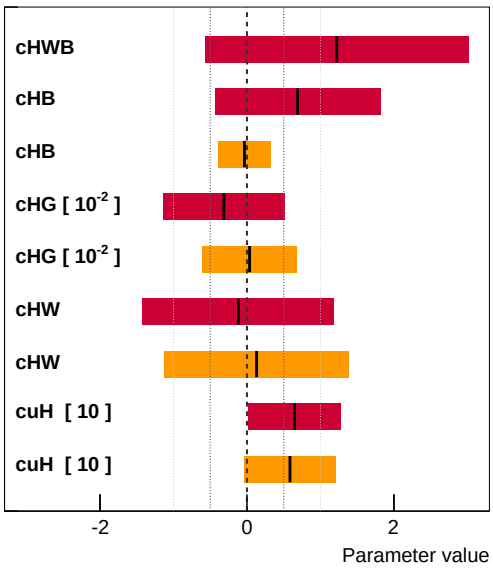


Figure 6.17: The comparison of the best-fit values and 68% C.L. intervals of the four-parameter (orange) and five-parameter (red) fits using SMEFTsim model to the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ STXS measurements. The observed results are shown.

un-fitted Wilson coefficients. The Gaussian constraints are taken from [132], the remaining coefficients are treated as uniformly distributed random variables within ± 10. The five parameter set with cuH is not constrained. Both cuH and cuG (corresponding to the operator $\mathcal{O}_{uG} = (\bar{q}_p \sigma^{\mu\nu} T^A u_r)\tilde{H} G^A_{\mu\nu}$) coefficients enter the fit via the $t\bar{t}$H measurement only. The cuG coefficient enters the equation with a larger prefactor, therefore the cuH coefficient is replaced by cuG. When only external Gaussian constraints are included, the uncertainties of all but cHB coefficient increased 1.3-3.1 times (Table 6.7), while when the remaining Wilson coefficients are accounted for, the uncertainties increased 3.4-4.7 times with respect to the fit without the "Wilsons as funcions" method. The increase due to uniformly distributed variables is almost exclusively from the cll coefficient (the corresponding operator is $\mathcal{O}_{ll} = (\bar{l}_p \gamma_\mu l_r)(\bar{l}_s \gamma^\mu l_t)$). The results when all un-fitted coefficients are included either as Gaussian or uniformly distributed random variables are presented in Figure 6.18

and the correlation matrix is presented in Figure 6.19. All five Wilson coefficients are constrained and this result represents constraints on Wilson coefficients with removed model dependence due to un-fitted coefficients and thus provides reliable and consistent results for the interpretations using specific theoretical models.

Table 6.7: The best-fit values and uncertainties of the five-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^*$ STXS measurements. The correction due to the induced model dependence, which is estimated using the "Wilsons as functions" method, is shown in the last two columns: the first column includes only external Gaussian constraints, while the results in the second column accounts for the full set of Wilson coefficients.

Wilson coefficient	Original fit result	"Wilsons as functions"	
		Gaussian only	Full set
cHB	0.7 ± 1.1	0.6 ± 1.1	0.6 ± 4.8
cHG	$(-0.33 \pm 0.83) \times 10^{-2}$	$(1.5 \pm 2.6) \times 10^{-2}$	$(1.5 \pm 2.9) \times 10^{-2}$
cHW	-0.1 ± 1.3	-0.7 ± 2.9	-0.7 ± 4.2
cuG	0.91 ± 0.89	1.8 ± 1.2	1.8 ± 3.0
$cHWB$	1.2 ± 1.8	0.7 ± 2.3	0.7 ± 8.5

6.5 Interpretation of local EFT fits

The EFT fits which are presented in this chapter are local EFT fits since only a subset of Wilson coefficients is constrained. The use of a subset of Wilson coefficients induces model dependence. Nevertheless, Higgs STXS measurements are expected to be a dominant contributor in constraining these Wilson coefficients in a global fit, so the HEL and SMEFTsim results in Figure 6.2(a) and Figure 6.15(b) are good indicators of how the constraints on these Wilson coefficients are expected to look.

The techniques developed while designing and performing this EFT analysis have already contributed to other scientific work. The result presented in section 6.1 is the first ATLAS EFT fit to combined Higgs data constraining as many as six parameters. This analysis also inspired and was followed closely by the other

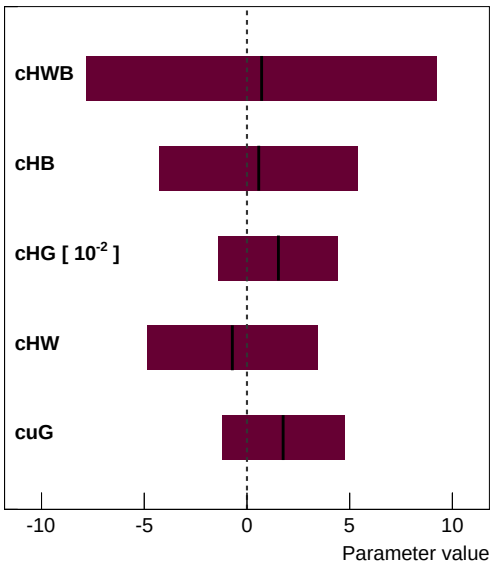


Figure 6.18: The best-fit values and uncertainties of the five-parameter fit using SMEFTsim model to the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$ STXS measurements. The uncertainties are evaluated using the "Wilsons as functions" method.

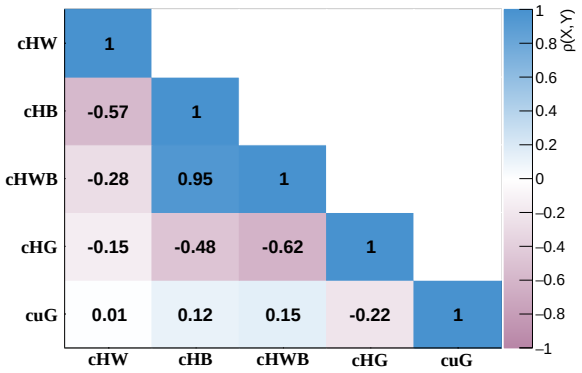


Figure 6.19: The observed correlation matrix from five-parameter fit using SMEFTsim model.

ATLAS EFT analysis in [4]. The HEL model equations made public in [5] were used in the global EFT analysis in [132].

The analysis design expertise and insights in EFT constraints in the Higgs sector

can and will contribute to the ongoing effort for the first ATLAS global EFT fit.

Chapter 7

Conclusions

This thesis has presented the measurement of the $H \to WW^*$ decay channel in the context of the Simplified Template Cross Section (STXS) framework [14], where the Higgs boson is produced via gluon-gluon fusion (ggF) and vector boson fusion (VBF) production modes, and the Effective Field Theory (EFT) interpretations on combined Higgs boson data in $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$ channels. These results are produced using proton-proton collision data recorded by the ATLAS detector at the LHC at a centre-of-mass energy of $\sqrt{s}$ =13 TeV that correspond to an integrated luminosity of 36.1 fb^{-1}.

The STXS $H \to WW^*$ measurement provides both precision measurements, which are sensitive to dimension-6 EFT operators, and measurements of the Higgs boson transverse momentum distribution, which directly probe new physics effects. The most precise ggF measurement is of an inclusive zero jet category and has the upper/lower uncertainties of 23%/22%. The remaining ggF measurements have uncertainties between 56% and 109%, while the VBF-like measurement is measured with the upper/lower uncertainties of 44%/43%. The results are consistent with Standard Model (SM) predictions. These results can be combined with STXS measurements in other Higgs boson decay channels and used for Beyond Standard Model (BSM) interpretations. The comparison with the best-to-date ATLAS STXS measurement of combined Higgs data shows that the results obtained in this thesis are

competitive with the ATLAS measurement and demonstrates that the dedicated $H \to WW^*$ analysis is performed successfully. The two main factors that would improve the measurement are the availability of a larger data set (e.g. the full LHC Run 2 data set), and the availability of larger Monte Carlo samples for theoretical modelling.

The second half of this thesis presented the first ATLAS EFT interpretation of combined Higgs data in $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels using dimension-6 operators in the strongly-interacting light Higgs (SILH) basis, which corresponds to the Higgs Effective Lagrangian (HEL) model. The results are made public in [3]. Six EFT parameters are constrained corresponding to the Higgs boson couplings to strong and electroweak gauge bosons, and to the top quark. The most stringent constraints are on the effective couplings to photons and to gluons; these constraints are stronger than those from a prior fit to a similar parameter set [124]. In order to evaluate the benefits of an EFT analysis performed by the ATLAS experiment, an EFT interpretation is performed using the same EFT parameterisation on the public ATLAS STXS measurements [26]. This interpretation is made public in [5]. The ATLAS STXS measurements are also interpreted using dimension-6 operators written in the Warsaw basis, which corresponds to the SMEFTsim model. Four Wilson coefficients, corresponding to the Higgs boson couplings to weak vector bosons, photons, and gluons, are constrained.

The $H \to WW^*$ results presented in this thesis have enabled an EFT interpretation on a combination of measurements in the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, and $H \to WW^*$ channels. In the case of a four EFT parameter fit described in the previous paragraph, the constraints on three Wilson coefficients corresponding to the Higgs boson couplings to strong and electroweak gauge bosons improved by 19-25%. This combination of STXS measurements is also able to constrain five Wilson coefficients. These results demonstrate the significance of the $H \to WW^*$

measurement.

The EFT analysis would greatly benefit from a set of measurements from other Higgs boson decay channels as well as outside the Higgs sector. This would expand the fits to larger operator sets. The results would also improve if more accurate theoretical predictions were available, for example by using an implementation at NLO in QCD, or by including dimension-8 operators. The results obtained in the EFT analysis can also be used to test specific BSM models [132].

The EFT analysis design presented in this thesis has inspired and been adopted by other ATLAS analyses, e.g. [4], CMS analysis, and the EFT parameterisation provided herein has been used in a global EFT fit interpretation [132]. The ATLAS collaboration is preparing for the first global EFT fit, so the experience presented in this thesis can hopefully contribute both analysis design expertise and insights into forming EFT constraints in the Higgs sector.

Appendix A

Boosted decision tree analysis

A.1 Boosted decision tree ranking method table

Table A.1: The comparison of three BDT ranking methods. First three columns list the ranking of the observables, while the last two columns indicate observable's rank with respect to the first discriminant. D_1 corresponds to the ranking method based on the plot area, D_2 is discriminant derived based on the centre of mass, and D_3 discriminant is based on the χ^2 test analogy. Red colour indicates that the difference in rank is larger than two units.

D_1	D_2	D_3	D_2 with respect to D_1	D_3 with respect to D_1
m_T	m_T	m_T	0	0
E_T^{miss}	E_T^{miss}	E_T^{miss}	0	0
$\Delta R_{\ell_0 j_0}$	$\Delta R_{\ell_0 j_0}$	$\Delta R_{\ell_0 j_0}$	0	0
$m_{\ell\ell}$	$\Delta R_{\ell_1 j_0}$	$\Delta R_{\ell_1 j_0}$	1	1
$\Delta R_{\ell_1 j_0}$	$m_{\ell\ell}$	$m_{\ell\ell}$	-1	-1
$\Delta\eta_{\ell_0 j_0}$	$\Delta\phi_{\ell\ell,j}$	$m_{\ell_1 j}$	1	3
$\Delta\phi_{\ell\ell,j}$	$\Delta\eta_{\ell_0 j_0}$	$\Delta\phi_{\ell\ell,j}$	-1	0
$\Delta\eta_{\ell_1 j_0}$	$\Delta\eta_{\ell_1 j_0}$	$\Delta\eta_{\ell_0 j_0}$	0	-2
$m_{\ell_1 j}$	$\Delta\phi_{\ell_0 j_0}$	$\Delta\eta_{\ell_1 j_0}$	1	-1
$\Delta\phi_{\ell_0 j_0}$	$\Delta\phi_{\ell_1 j_0}$	$\Delta\phi_{\ell_0 j_0}$	1	0
$\Delta\phi_{\ell_1 j_0}$	$m_{\ell_1 j}$	$\Delta\phi_{\ell_1 j_0}$	-2	0
$\Delta R_{\ell\ell}$	$\Delta R_{\ell\ell}$	$m_{\ell_0 j}$	0	1
$m_{\ell_0 j}$	$\Delta\phi_{j,E_T^{miss}}$	$\Delta R_{\ell\ell}$	1	-1
$\Delta\phi_{j,E_T^{miss}}$	$m_{\ell_0 j}$	$\Delta\phi_{j,E_T^{miss}}$	-1	0
$\Delta\eta_{\ell\ell}$	$\Delta\eta_{\ell\ell}$	$p_T^{\ell_0}$	0	4
p_T^j	$p_T^{\ell_0}$	$\Delta\eta_{\ell\ell}$	3	-1
η_{ℓ_0}	$p_T^{\ell\ell}$	η_{ℓ_0}	4	0
$p_T^{\ell_0}$	p_T^j	subη_{ℓ_0}	-3	-2
$\Delta\phi_{\ell\ell}$	$\Delta\phi_{\ell\ell,E_T^{miss}}$	$\Delta\phi_{\ell\ell}$	4	0
η_j	p_T^H	η_j	2	0
$p_T^{\ell\ell}$	$p_T^{\ell_1}$	$p_T^{\ell\ell}$	3	0
p_T^H	η_j	$p_T^{\ell_1}$	-2	2
$\Delta\phi_{\ell\ell,E_T^{miss}}$	ϕ_{ℓ_0}	p_T^H	3	-1

A.2 Boosted decision tree construction material

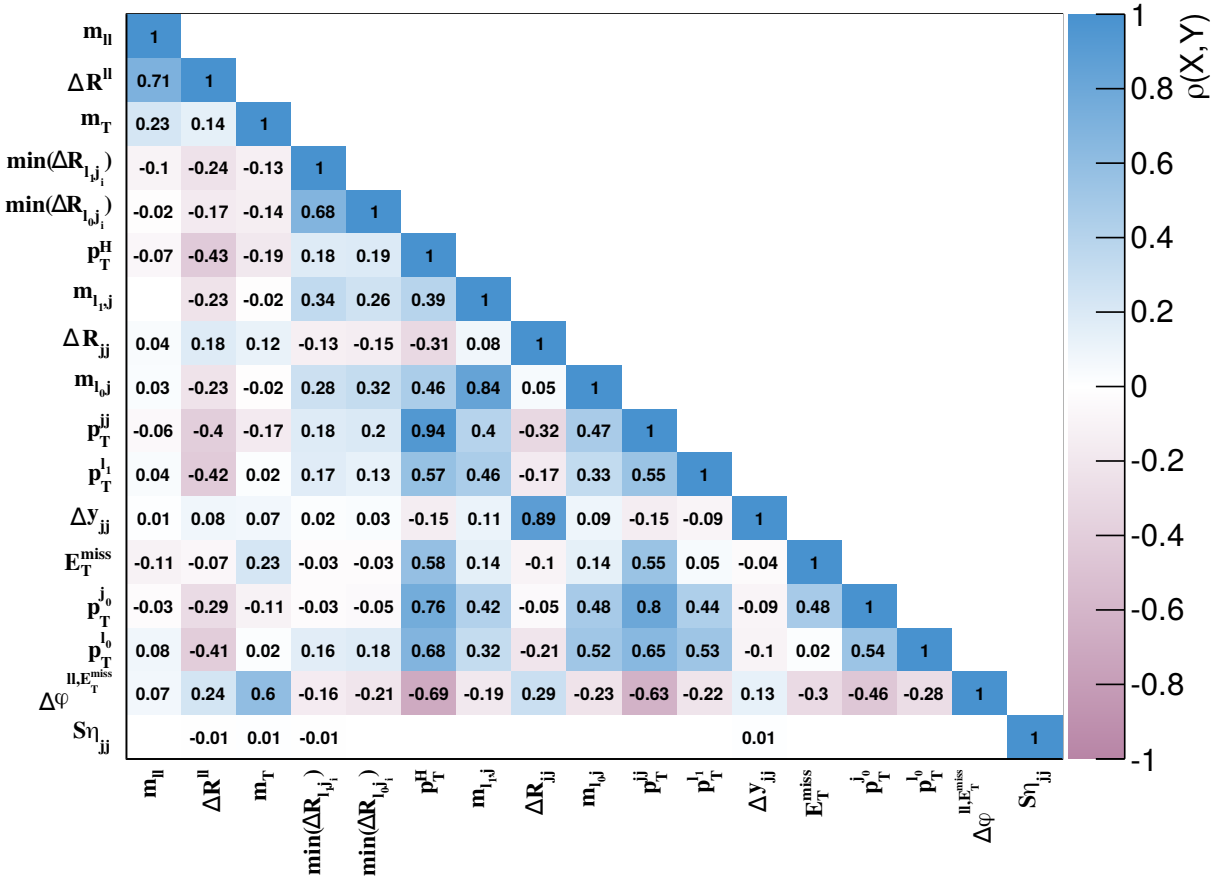


Figure A.1: Correlation matrix of the BDT discriminating parameters obtained in the training of the best performing BDT in $\geqslant$ 2-jet category with 17 discriminating parameters.

Table A.2: The base and the additional subsets used to construct the subsets of discriminating observables for the BDT training for 1-jet signal region and selected based on the shapes of the signal and background distributions. A BDT input subset is constructed by taking the full base subset and adding one observable at a time from the additional subset.

Base subset	+1 observable from this subset
$m_{\ell\ell}$, m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta\phi_{\ell\ell,j}$	$\Delta\eta_{\ell_1 j}$, $\Delta\eta_{\ell\ell}$, $\Delta\phi_{j,E_T^{miss}}$, $\Delta\phi_{\ell_0 j}$, $\Delta\phi_{\ell_1 j}$, $\Delta\phi_{\ell\ell}$, $\Delta R_{\ell_0 j}$, $\Delta R_{\ell_1 j}$, p_T^j, $p_T^{\ell_0}$, $\max(m_T^{\ell_0}, m_T^{\ell_1})$, E_T^{miss}, $m_{\ell_0 j}$, $m_{\ell_1 j}$, $m_{\tau\tau}$, $p_T^{\ell\ell}$, $p_T^{\ell_1}$, $E_T^{miss,Track}$, $\Delta\eta_{\ell_0 j}$
m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta\phi_{\ell\ell,j}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $p_T^{\ell_1}$
$m_{\ell\ell}$, m_T	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta\eta_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $\Delta\phi_{\ell\ell,j}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta R_{\ell\ell}$, $\max(m_T^{\ell_0}, m_T^{\ell_1})$, $min\Delta R_{\ell_0 j_i}$, $min\Delta R_{\ell_1 j_i}$, $p_T^{\ell\ell}$, $E_T^{miss,Track}$
$m_{\ell\ell}$, m_T, $\Delta\phi_{\ell\ell}$, $p_T^{\ell\ell}$	$min\Delta R_{\ell_0 j_i}$, $min\Delta R_{\ell_1 j_i}$
m_T, $\Delta R_{\ell\ell}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta\phi_{\ell\ell,j}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta R_{\ell_1 j}$, $m_{\ell\ell}$, $p_T^{\ell_1}$
m_T, $\Delta R_{\ell\ell}$, $m_{\ell\ell}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $p_T^{\ell_1}$
m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $m_{\ell\ell}$, $p_T^{\ell_1}$
m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $m_{\ell\ell}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $p_T^{\ell_1}$
m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,j}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $m_{\ell\ell}$, $p_T^{\ell_1}$
m_T, $\Delta R_{\ell\ell}$, $\Delta\phi_{\ell\ell,j}$, $m_{\ell\ell}$	$\Delta\eta_{\ell_0 j}$, $\Delta\eta_{\ell_1 j}$, $\Delta R_{\ell_1 j}$, $p_T^{\ell_1}$

Table A.3: The base and the additional subsets used to construct the subsets of discriminating observables for the BDT training for 1-jet signal region to construct BDTs composed of 11 discriminating observables. In $min\Delta R_{\ell_0 j_i}$ and $min\Delta R_{\ell_1 j_i}$ notation the j_i is chosen from $\{j_0, j_1\}$.

Base subset of 10 observables	+1 observable from this subset
m_T, E_T^{miss}, $m_{\ell\ell}$, $p_T^{\ell_0}$, $\Delta\phi_{\ell\ell,j}$, $p_T^{\ell_1}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $\Delta R_{\ell_0 j_0}$, $\Delta R_{\ell_1 j_0}$, $\Delta R_{\ell\ell}$	$\Delta\phi_{\ell\ell}$, $\Delta\phi_{\ell_0 j_1}$, $\Delta\eta_{\ell_0 j_1}$, $\Delta\phi_{\ell_1 j_1}$, $\Delta\eta_{\ell_1 j_1}$, $\Delta R_{\ell_0 j_1}$, $\Delta R_{\ell_1 j_1}$, $min\Delta R_{\ell_0 j_i}$, $m_{\ell_0 j_0}$, $m_{\ell_0 j_1}$, $m_{\ell_1 j_0}$, $m_{\ell_1 j_1}$, $\Delta\eta_{\ell\ell}$, p_T^j, $m_{\tau\tau}$, $\Delta\phi_{\ell_0 j_0}$, $\Delta\eta_{\ell_0 j_0}$, $\Delta\phi_{\ell_1 j_0}$, $\Delta\eta_{\ell_1 j_0}$, $\Delta\phi_{j,E_T^{miss}}$, p_T^H, $p_T^{\ell\ell}$, $min\Delta R_{\ell_1 j_i}$

Table A.4: The base and the additional subsets used to construct the subsets of discriminating observables for the BDT training for $\geqslant$ 2-jet signal region. The origin of the subsets is indicated in the first column. A BDT input subset is constructed by taking the full base subset and adding one observable at a time from the additional subset. excluding the already included observables.

Origin	Base subset	+1 observable from this subset
Top five observables ranked in the inclusive $\geqslant$ 2-jet SR using all background processes in the procedure.	$\Delta R_{\ell\ell}$, $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $min\Delta R_{\ell_1 j_i}$, m_T	$\Delta R_{\ell\ell}$, m_T, $\Delta\phi_{\ell\ell}$, $min\Delta R_{\ell_1 j_i}$, $p_T^{\ell\ell}$, $\Delta R_{\ell_1 j_0}$, $min\Delta R_{\ell_0 j_i}$, $\Delta\phi_{\ell_1 j_0}$, $\Delta\eta_{\ell\ell}$, $\Delta R_{\ell_0 j_0}$, $\Delta\phi_{\ell\ell,j}$, p_T^H, $m_{\ell_1 j_0}$, $\Delta\phi_{\ell_0 j_0}$, ΔR_{jj}, $m_{\ell_0 j_0}$, $\Delta R_{\ell_0 j_1}$, $\Delta R_{\ell_1 j_1}$, p_T^{jj}, $\Delta\phi_{\ell_1 j_1}$, $p_T^{\ell_1}$, Δy_{jj}, $S\eta_{jj}$, $\Delta\eta_{jj}$, m_{jj}, $\Delta\phi_{\ell_0 j_1}$, $\Delta\eta_{\ell_0 j_1}$, $\Delta\eta_{\ell_1 j_1}$, $m_{\ell_0 j_1}$, $m_{\ell_1 j_1}$, E_T^{miss}, p_T^j, $p_T^{\ell_0}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $m_{\tau\tau}$, $\Delta\eta_{\ell_0 j_0}$, $\Delta\eta_{\ell_1 j_0}$, $\Delta\phi_{j,E_T^{miss}}$
Top five observables ranked in the inclusive $\geqslant$ 2-jet SR using only $Z \to \tau\tau$ process in the procedure.	$\Delta R_{\ell\ell}$, $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell_1}$, $p_T^{\ell\ell}$	
The subsets are selected using the plots of normalised signal and background distributions, the ranking in individual $\geqslant$ 2-jet signal regions corresponding to STXS kinematic definitions and taking potential correlations between the observables into account.	$m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell_1}$, m_T, $\Delta\phi_{jj}$	
	$m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell_1}$, p_T^{jj}, $\Delta\phi_{jj}$	
	$\Delta R_{\ell\ell}$, $m_{\ell\ell}$, $\Delta\phi_{jj}$, $min\Delta R_{\ell_1 j_i}$, m_T	

Table A.5: The base and the additional subsets used to construct the subsets of discriminating observables for the BDT training for $\geqslant$ 2-jet signal region for constructing BDTs with large discriminating observable sets, consisting of 17 or 24 observables.

Base subset	+1 observable from this subset	Size of base subset
$m_{\ell\ell}$, $\Delta R_{\ell\ell}$, m_T, $min\Delta R_{\ell_1 j_i}$, $min\Delta R_{\ell_0 j_i}$, p_T^H, $m_{\ell_1 j_0}$, ΔR_{jj}, $m_{\ell_0 j_0}$, p_T^{jj}, $p_T^{\ell_1}$, Δy_{jj}, E_T^{miss}, p_T^j, $p_T^{\ell_0}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$	$S\eta_{jj}$, $\Delta\eta_{jj}$, m_{jj}, $\Delta\phi_{\ell_0 j_1}$, $\Delta\eta_{\ell_0 j_1}$, $\Delta\eta_{\ell_1 j_1}$, $m_{\ell_0 j_1}$, $m_{\ell_1 j_1}$, $m_{\tau\tau}$, $\Delta\eta_{\ell_0 j_0}$, $\Delta\eta_{\ell_1 j_0}$, $\Delta\phi_{j,E_T^{miss}}$, $\Delta\phi_{\ell\ell}$, $p_T^{\ell\ell}$, $\Delta R_{\ell_1 j_0}$, $\Delta\phi_{\ell_1 j_0}$, $\Delta\eta_{\ell\ell}$, $\Delta R_{\ell_0 j_0}$, $\Delta\phi_{jj}$, $\Delta\phi_{\ell\ell,j}$, $\Delta\phi_{\ell_0 j_0}$, $\Delta R_{\ell_0 j_1}$, $\Delta R_{\ell_1 j_1}$, $\Delta\phi_{\ell_1 j_1}$	16
$m_{\ell\ell}$, $\Delta R_{\ell\ell}$, m_T, $\Delta\phi_{\ell\ell}$, $min\Delta R_{\ell_1 j_i}$, $p_T^{\ell\ell}$, $\Delta R_{\ell_1 j_0}$, $min\Delta R_{\ell_0 j_i}$, $\Delta\phi_{\ell_1 j_0}$, $\Delta\eta_{\ell\ell}$, $\Delta R_{\ell_0 j_0}$, $\Delta\phi_{jj}$, $\Delta\phi_{\ell\ell,j}$, p_T^H, $m_{\ell_1 j_0}$, $\Delta\phi_{\ell_0 j_0}$, ΔR_{jj}, $m_{\ell_0 j_0}$, $\Delta R_{\ell_0 j_1}$, $\Delta R_{\ell_1 j_1}$, p_T^{jj}, $\Delta\phi_{\ell_1 j_1}$, $p_T^{\ell_1}$	Δy_{jj}, $S\eta_{jj}$, $\Delta\eta_{jj}$, m_{jj}, $\Delta\phi_{\ell_0 j_1}$, $\Delta\eta_{\ell_0 j_1}$, $\Delta\eta_{\ell_1 j_1}$, $m_{\ell_0 j_1}$, $m_{\ell_1 j_1}$, E_T^{miss}, p_T^j, $p_T^{\ell_0}$, $\Delta\phi_{\ell\ell,E_T^{miss}}$, $m_{\tau\tau}$, $\Delta\eta_{\ell_0 j_0}$, $\Delta\eta_{\ell_1 j_0}$, $\Delta\phi_{j,E_T^{miss}}$	23

Table A.6: Summary of the BDT construction for the $\geqslant$ 2-jet category based on the low correlations between the discriminating observables. BDTs are constructed by taking the full set from the first column and either adding one observable from the second column or adding two observables from the third column.

Base subset	+1 observable from this subset	+2 observable from this subset
$S\eta_{jj}$, $min\Delta R_{\ell_1 j_i}$, $m_{\ell\ell}$, ΔR_{jj}, $m_{\ell_1 j_0}$	m_T, E_T^{miss}, p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j	p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$
$min\Delta R_{\ell_1 j_i}$, $m_{\ell\ell}$, ΔR_{jj}, $m_{\ell_1 j_0}$, m_T, E_T^{miss}	p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j	
$S\eta_{jj}$, $min\Delta R_{\ell_1 j_i}$, $m_{\ell\ell}$, ΔR_{jj}, $m_{\ell_1 j_0}$, m_T, E_T^{miss}	p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j	
$min\Delta R_{\ell_1 j_i}$, $m_{\ell\ell}$, ΔR_{jj}, $m_{\ell_1 j_0}$	m_T, E_T^{miss}, p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j	

Table A.7: The base and the additional subsets used to construct the subsets of discriminating observables for the BDT training for $\geqslant$ 2-jet signal region selected based on the results of the correlation study.

Base subset	+1 observable from this subset
$S\eta_{jj}$, $min\Delta R_{\ell_1 j_i}$, $\Delta R_{\ell\ell}$, ΔR_{jj}, m_T, E_T^{miss}	p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j
$min\Delta R_{\ell_1 j_i}$, $\Delta R_{\ell\ell}$, ΔR_{jj}, m_T, E_T^{miss}	p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $p_T^{\ell_1}$, p_T^j

Table A.8: The base and the additional subsets used to construct the subsets of discriminating observables for the selection of the best BDT discriminator. The subsets are selected based on all studies in the $\geqslant$ 2-jet case. A BDT input subset is constructed by taking the full base subset and adding three observables at a time from the additional subset.

Base subset	+3 observables from this subset
$\Delta R_{\ell\ell}$, m_T, E_T^{miss}	$m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, $min\Delta R_{\ell_1 j_i}$, $\Delta\phi_{\ell_1 j_0}$, $\Delta\phi_{\ell_1 j_1}$, $p_T^{\ell_1}$, p_T^H, p_T^{jj}, $p_T^{\ell_0}$, $m_{\ell_0 j_1}$, $\Delta R_{\ell_0 j_1}$, p_T^j

Bibliography

[1] Postgraduate Research Student Handbook. `https://www2.physics.ox.ac.uk/sites/default/files/page/2011/09/05/handbook-2016-2017-41329.pdf`.

[2] ATLAS Collaboration. Measurements of gluon–gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 789:508 – 529, 2019.

[3] ATLAS Collaboration. Constraints on an effective Lagrangian from the combined $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data collected with the ATLAS detector. 2017. ATL-PHYS-PUB-2017-018.

[4] ATLAS Collaboration. Measurement of VH, $H \to b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector. 2019. CERN-EP-2019-019.

[5] Chris Hays, Veronica Sanz, and Gabija Zemaityte. Constraining EFT parameters using Simplified Template Cross Sections. *LHCHXSWG-INT-2017-001*, 2017.

[6] Piotr Golonka and Zbigniew Was. PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays. *Eur. Phys. J.*, C45:97–107, 2006.

[7] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012.

[8] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.

[9] CMS collaboration. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *Journal of High Energy Physics*, 2013(6):81, 2013.

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

[11] D0 Collaboration. Measurement of the W boson mass with the D0 detector. *Phys. Rev. D*, 89:012005, 2014.

[12] The ATLAS and CMS collaborations. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. *Journal of High Energy Physics*, 2016(8):45, 2016.

[13] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98:030001, 2018.

[14] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.

[15] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

[16] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[17] 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, May 2015.

[18] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. 2011.

[19] Robin Devenish and Amanda Cooper-Sarkar. *Deep Inelastic Scattering*. Oxford University Press, 2004.

[20] Ilaria Brivio and Michael Trott. The Standard Model as an Effective Field Theory. *Phys. Rept.*, 793:1–98, 2019.

[21] H. Georgi. Effective field theory. *Ann. Rev. Nucl. Part. Sci.*, 43:209–252, 1993.

[22] Weinberg, Steven. Baryon and Lepton Nonconserving Processes. *Phys. Rev. Lett.*, 43:1566–1570, 1979.

[23] André de Gouvêa, Juan Herrero-García, and Andrew Kobach. Neutrino masses, grand unification, and baryon number violation. *Phys. Rev. D*, 90:016011, 2014.

[24] Frank F. Deppisch, Julia Harz, Wei-Chih Huang, Martin Hirsch, and Heinrich Päs. Falsifying high-scale baryogenesis with neutrinoless double beta decay and lepton flavor violation. *Phys. Rev. D*, 92:036005, 2015.

[25] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek. Dimension-Six Terms in the Standard Model Lagrangian. *JHEP*, 10:085, 2010.

[26] ATLAS Collaboration. Combined measurements of Higgs boson production and decay in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels using $\sqrt{s} = 13$ TeV

pp collision data collected with the ATLAS experiment. 2017. ATLAS-CONF-2017-047.

[27] M. L. Mangano and T. J. Stelzer. Tools for the simulation of hard hadronic collisions. *Ann. Rev. Nucl. Part. Sci.*, 55:555–588, 2005.

[28] Andy Buckley et al. General-purpose event generators for LHC physics. *Phys. Rept.*, 504:145–233, 2011.

[29] Stefan Hoeche, Frank Krauss, Nils Lavesson, Leif Lonnblad, Michelangelo Mangano, Andreas Schalicke, and Steffen Schumann. Matching parton showers and matrix elements. In *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*, pages 288–289, 2005.

[30] Nils Lavesson and Leif Lonnblad. Merging parton showers and matrix elements: Back to basics. *JHEP*, 04:085, 2008.

[31] Carlo Oleari. The POWHEG-BOX. *Nucl. Phys. Proc. Suppl.*, 205-206:36–41, 2010.

[32] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[33] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[34] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[35] S. Agostinelli et al. GEANT4 - A simulation toolkit. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. Cited By :11493.

[36] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003–S08003, 2008.

[37] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001–S08001, 2008.

[38] Esma Mobs. The CERN accelerator complex. 2016. `https://cds.cern.ch/record/2197559`.

[39] ATLAS experiment public results picture gallery. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[40] ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *Submitted to: Eur. Phys. J.*, 2019.

[41] ATLAS Collaboration. Study of the material of the ATLAS inner detector for Run 2 of the LHC. *Journal of Instrumentation*, 12(12):P12009–P12009, 2017.

[42] ATLAS Collaboration. Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeVs. 2015. ATL-PHYS-PUB-2015-026.

[43] ATLAS Collaboration. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014.

[44] ATLAS Collaboration. Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties. *JINST*, 12(05):P05002, 2017.

[45] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$=13 TeV. *The European Physical Journal C*, 76(5):292, 2016.

[46] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$=13TeV with the ATLAS detector. *Phys. Rev. D*, 96:072002, 2017.

[47] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993.

[48] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[49] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. *JHEP*, 08:001, 1997.

[50] Ryan Atkin. Review of jet reconstruction algorithms. *J. Phys. Conf. Ser.*, 645(1):012008, 2015.

[51] 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(11):581, Oct 2016.

[52] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. May 2014. ATLAS-CONF-2014-018.

[53] ATLAS Collaboration. Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector. *Eur. Phys. J.*, C77(9):580, 2017. [Erratum: Eur. Phys. J.C77,no.10,712(2017)].

[54] ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$=13TeV. *The European Physical Journal C*, 78(11):903, 2018.

[55] ATLAS Collaboration. Observation and measurement of Higgs boson decays to WW* with the ATLAS detector. *Phys. Rev.*, D92(1):012006, 2015.

[56] CMS Collaboration. Observation of a New Boson with Mass Near 125 GeV in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06:081, 2013.

[57] CMS Collaboration. Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 122(12):121803, 2019.

[58] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[59] Keith Hamilton, Paolo Nason, Emanuele Re, and Giulia Zanderighi. NNLOPS simulation of Higgs boson production. *JHEP*, 10:222, 2013.

[60] John M. Campbell, R. Keith Ellis, Rikkert Frederix, Paolo Nason, Carlo Oleari, and Ciaran Williams. NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM. *JHEP*, 07:092, 2012.

[61] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *Journal of High Energy Physics*, 2010(2):37, 2010.

[62] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.

[63] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013. CERN-2013-004.

[64] Charalampos Anastasiou, Claude Duhr, Falko Dulat, Franz Herzog, and Bernhard Mistlberger. Higgs Boson Gluon-Fusion Production in QCD at Three Loops. *Phys. Rev. Lett.*, 114:212001, 2015.

[65] Charalampos Anastasiou, Claude Duhr, Falko Dulat, Elisabetta Furlan, Thomas Gehrmann, Franz Herzog, Achilleas Lazopoulos, and Bernhard Mistlberger. High precision determination of the gluon fusion Higgs boson cross-section at the LHC. *JHEP*, 05:058, 2016.

[66] Stefano Actis, Giampiero Passarino, Christian Sturm, and Sandro Uccirati. NLO electroweak corrections to Higgs boson production at hadron colliders. *Physics Letters B*, 670(1):12 – 17, 2008.

[67] Charalampos Anastasiou, Radja Boughezal, and Frank Petriello. Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion. *Journal of High Energy Physics*, 2009(04):003–003, apr 2009.

[68] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator. *JHEP*, 05:140, 2015.

[69] M. Ciccolini, A. Denner, and S. Dittmaier. Strong and Electroweak Corrections to the Production of a Higgs boson +2 jets via Weak Interactions at the Large Hadron Collider. *Phys. Rev. Lett.*, 99:161803, 2007.

[70] M. Ciccolini, Denner A., and Dittmaier S. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC. *Phys. Rev. D*, 77:013002, 2008.

[71] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro. Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD. *Phys. Rev. Lett.*, 105:011801, 2010.

[72] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. HW/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013.

[73] T. Han and S. Willenbrock. QCD correction to the pp $\rightarrow$ WH and ZH total cross sections. *Physics Letters B*, 273(1):167 – 172, 1991.

[74] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579:149–156, 2004.

[75] M. L. Ciccolini, S. Dittmaier, and M. Krämer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev. D*, 68:073003, 2003.

[76] T Gleisberg, S Höche, F Krauss, M Schönherr, S Schumann, F Siegert, and J Winter. Event generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007–007, feb 2009.

[77] Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *Journal of High Energy Physics*, 2008(12):039–039, dec 2008.

[78] F. Cascioli, S. Höche, F. Krauss, P. Maierhöfer, S. Pozzorini, and F. Siegert. Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production. *JHEP*, 01:046, 2014.

[79] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[80] S Schumann and F Krauss. A parton shower algorithm based on Catani-Seymour dipole factorisation. *Journal of High Energy Physics*, 2008(03):038–038, mar 2008.

[81] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers. The NLO case. *Journal of High Energy Physics*, 2013(4):27, 2013.

[82] Sherpa Team. Sherpa 2.1.1 Manual. `https://sherpa.hepforge.org/doc/SHERPA-MC-2.1.1.html`.

[83] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, 2010.

[84] Fabrizio Caola, Kirill Melnikov, Raoul Röntsch, and Lorenzo Tancredi. QCD corrections to W^+W^- production through gluon fusion. *Physics Letters B*, 754:275 – 280, 2016.

[85] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *Journal of High Energy Physics*, 2007(09):126–126, sep 2007.

[86] Michal Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$. *Phys. Rev. Lett.*, 110:252004, 2013.

[87] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J.*, C71:1547, 2011.

[88] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, may 2006.

[89] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$. *Phys. Rev. D*, 74:114017, 2006.

[90] Charalampos Anastasiou, Lance Dixon, Kirill Melnikov, and Frank Petriello. High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order. *Phys. Rev. D*, 69:094008, 2004.

[91] ATLAS collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2014(9):145, 2014.

[92] Stefano Catani and Massimiliano Grazzini. Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-Boson Production at the Large Hadron Collider. *Phys. Rev. Lett.*, 98:222002, 2007.

[93] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension. *Computer Physics Communications*, 108(1):56 – 74, 1998.

[94] Abdelhak Djouadi, Jan Kalinowski, Margarete Muehlleitner, and Michael Spira. HDECAY: Twenty++ years after. *Comput. Phys. Commun.*, 238:214–231, 2019.

[95] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013.

[96] Andrew Buckley. ATLAS Pythia 8 tunes to 7 TeV data. In *Proceedings of the Sixth International Workshop on Multiple Partonic Interactions at the Large Hadron Collider*, page 29, Geneva, 2014. CERN. ATL-PHYS-PROC-2014-273.

[97] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev.*, D82:074018, 2010.

[98] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152 – 155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[99] T. Plehn, David L. Rainwater, and D. Zeppenfeld. A Method for identifying $H \to \tau^+\tau^- \to e^\pm\mu^\mp p_T$ at the CERN LHC. *Phys. Rev.*, D61:093005, 2000.

[100] Edvin Sidebo. *Measurements of the Standard Model Higgs boson cross sections in the WW^* decay mode with the ATLAS experiment.* PhD thesis, 2018.

[101] A. Hoecker et al. TMVA - Toolkit for Multivariate Data Analysis. *arXiv e-prints*, page physics/0703039, 2007.

[102] J. S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pages 115–120, 2011.

[103] The method was first developed by Claudia Bertella.

[104] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J.*, C75(5):204, 2015.

[105] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of

tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):79, 2014.

[106] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.

[107] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016.

[108] Stefan Gieseke, Christian Röhr, and Andrzej Siódmok. Colour reconnections in Herwig++. *The European Physical Journal C*, 72(11):2225, 2012.

[109] Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R Webber. Single-top hadroproduction in association with a W boson. *Journal of High Energy Physics*, 2008(07):029–029, jul 2008.

[110] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data. *JINST*, 14(03):P03017, 2019.

[111] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. 2015. ATL-PHYS-PUB-2015-015.

[112] ATLAS Collaboration. Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. 2016. ATL-PHYS-PUB-2016-012.

[113] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C76(12):653, 2016.

[114] ATLAS Collaboration. Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. 2019. ATLAS-CONF-2019-005.

[115] Communication with the ATLAS Higgs Combination analysis team.

[116] Adam Alloul, Benjamin Fuks, and Verónica Sanz. Phenomenology of the Higgs effective Lagrangian via FeynRules. *Journal of High Energy Physics*, 2014(4):110, 2014.

[117] Ilaria Brivio, Yun Jiang, and Michael Trott. The SMEFTsim package, theory and tools. *JHEP*, 12:070, 2017.

[118] Roberto Contino, Margherita Ghezzi, Christophe Grojean, Margarete Muhlleitner, and Michael Spira. Effective Lagrangian for a light Higgs-like scalar. *JHEP*, 07:035, 2013.

[119] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, C. Zhang. Standard Model Effective Theory at Next-to-Leading-Order in QCD. Technical report. `https://feynrules.irmp.ucl.ac.be/wiki/SMEFTatNLO`.

[120] D.B. Kaplan. Strongly coupled supersymmetry as the possible origin of flavor. In *'97 electroweak interactions and unified theories. Proceedings, 32nd Rencontres de Moriond, Les Arcs, France, March 15-22, 1997*, pages 219–226, 1997.

[121] S.J. Cowley. Lecture notes for the Mathematical Tripos: IA Vectors & Matrices. 2010. `http://www.damtp.cam.ac.uk/user/sjc1/teaching/VandM/notes.pdf`.

[122] ATLAS Collaboration. A morphing technique for signal modelling in a multidimensional space of coupling parameters. 2015. ATL-PHYS-PUB-2015-047.

[123] Michael E. Peskin and Tatsu Takeuchi. New constraint on a strongly interacting Higgs sector. *Phys. Rev. Lett.*, 65:964–967, 1990.

[124] John Ellis, Veronica Sanz, and Tevong You. The Effective Standard Model after LHC Run I. *JHEP*, 03:157, 2015.

[125] Andy Buckley, Hendrik Hoeth, Heiko Lacker, Holger Schulz, and Jan Eike von Seggern. Systematic event generator tuning for the LHC. *Eur. Phys. J.*, C65:331–357, 2010.

[126] Glen Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.

[127] Lecture notes. `https://www.stat.tamu.edu/~suhasini/teaching613/chapter3.pdf`.

[128] ATLAS Collaboration. Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev.*, D98:052005, 2018.

[129] ATLAS Collaboration. Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 10:132, 2017.

[130] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[131] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [186(2003)].

[132] John Ellis, Christopher W. Murphy, Verónica Sanz, and Tevong You. Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data. *JHEP*, 06:146, 2018.