TÉCNICO LISBOA

Constraining Charge-Parity symmetry violation in the interactions of the Higgs and W bosons

Beatriz Rebelo Rosalino

Thesis to obtain the Master of Science Degree in

Engineering Physics

Supervisors: Dr. Maria Inês Abreu Julião Ochoa de Castro
Prof. Patricia Conde Muiño

Examination Committee

Chairperson: Prof. José Guilherme Teixeira de Almeida Milhano
Supervisor: Dr. Maria Inês Abreu Julião Ochoa de Castro
Members of the Committee: Prof. José Ricardo Morais Silva Gonçalo
Prof. Ruben Maurício da Silva Conceição

February 2025

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

Acknowledgments

Before anyone else, I would like to thank my supervisors for all their support throughout this year of work. To Dr. Inês Ochoa and Prof. Patricia Muíño, who helped me tremendously whenever I found myself struggling. Their guidance and their knowledge was invaluable to me.

My parents deserve my deepest thanks for always believing in me and never letting me forget that, especially when I doubted myself the most. I am extremely grateful them for their unwavering support and for encouraging my curiosity from the start. I also thank my father for all the ”pão de deus” and coffee he bought for me throughout this year.

I am very thankful for my friends, who kept me sane as I wrote this dissertation, from the companionship in the long days, to the much-needed breaks and get-togethers. Thank you for standing by me through everything.

I would also like to thank the ATLAS WH anomalous couplings group at LIP, for all their help and support in this work. I could not have done it without them.

Resumo

A violação da simetria de carga e paridade (CP) no universo primordial é necessária para explicar a discrepância entre a evidência de o universo ser presentemente dominado por matéria, e a hipótese de o mesmo número de partículas e anti-partículas ter sido produzido no Big Bang. O modelo padrão não explica esta discrepância. Este trabalho usa a teoria de campo efetiva, baseada no modelo padrão, para estudar acoplamentos anómalos na produção de um bosão de Higgs, que decai para dois quarks b, associado a um bosão W, que decai em leptões. O foco é num operador CP-ímpar, $c_{H\tilde{W}}$, que afeta este processo e é uma possível fonte de violação de CP. Dois outros operadores CP-par são também considerados: c_{HW} e $c_{Hq}^{(3)}$. Obtiveram-se distribuições de quatro variáveis para investigar a sua sensibilidade à presença destes operadores: o momento transverso do bosão W, a distância angular entre os jatos b e duas variáveis angulares. A combinação do momento transverso e uma das variáveis angulares é adequada para obter uma parametrização da secção eficaz do processo em função das contribuições anómalas, também feita para as variáveis separadas. A parametrização é finalmente usada num fit para impor limites nos valores de $c_{H\tilde{W}}$, e, portanto, na magnitude do acoplamento CP-ímpar. Este trabalho usa dados da experiência ATLAS de colisões de protões da Run 2 do LHC, com $\sqrt{s} = 13$ TeV, e é uma contribuição para um artigo em preparação sobre violação de CP em acoplamentos anómalos na produção associada de WH.

Palavras-chave: Bosão de Higgs, Violação de CP, Experiência ATLAS, Teoria de Campo Efetiva

Abstract

The violation of charge and parity symmetry (CP) must have occurred in the early universe to explain the discrepancy between the hypothesis that particles and anti-particles were produced in equal numbers in the Big Bang, and the observation that the universe today is matter-dominated. This is not explained by the Standard Model. This work uses the Standard Model effective field theory framework to study anomalous couplings of associated production of a Higgs boson, decaying into two b-quarks, and a W boson, decaying leptonically, focusing on $c_{H\tilde{W}}$, a CP-odd operator affecting this process, that can cause CP violation. Two other CP-even operators are also considered: c_{HW} and $c_{Hq}^{(3)}$. The distributions of four variables were studied to probe their sensitivity to these operators - the transverse momentum of the W boson, the angular distance between the b-jets, and two angular observables. The transverse momentum and one of the angular observables were chosen as the combination of observables with a suitable sensitivity to perform a binned parameterisation of the WH associated production cross section, in terms of the anomalous contributions. This parameterisation is performed independently for the two observables and in combination, and is later used in a likelihood fit to constrain the possible values of $c_{H\tilde{W}}$, and so, the magnitude of the CP-odd coupling. This work uses ATLAS Run 2 proton-proton collision data, with $\sqrt{s} = 13$ TeV, and is a contribution to a paper in preparation, studying CP violation in anomalous couplings in WH associated production.

Keywords: Higgs boson, CP violation, ATLAS experiment, Effective Field Theory

Contents

List of Tables

List of Figures

Glossary

2HDM	Two-Higgs Doublet Model
ATLAS	A Toroidal LHC Apparatus
BSM	Beyond the Standard Model
CERN	European Organization for Nuclear Research
CKM	Cabibbo–Kobayashi–Maskawa Matrix
CL	Confidence Level
CM	Center of Mass
CMS	Compact Muon Solenoid
CP	Charge-Parity Symmetry
CPV	Charge-Parity Symmetry Violation
EFT	Effective Field Theory
ggF	gluon-gluon Fusion
GSW	Glashow-Salam-Weinberg Model
LHC	Large Hadron Collider
MET	Missing Transverse Energy
NLO	Next-to-leading Order
PDF	Parton Distribution Function
SM	Standard Model of particle physics
SMEFT	Standard Model Effective Field Theory
STXS	Simplified Template Cross Sections
VBF	Vector Boson Fusion

Chapter 1

Introduction

One of the most important open questions in particle physics is related to explaining why there exists more matter than anti-matter in the universe. This observable difference goes against what is proposed by the theory, which suggests the same amount of both was produced in the Big Bang. To explain this imbalance when considering that anti-matter and matter were produced in the same numbers, there has to have existed a violation of charge-parity symmetry in the early universe.

All known elementary particles and the interactions between them, mediated by the fundamental forces, can be described by the Standard Model of particle physics (SM). Even though there are sources of CP violation in this theory, they do not account for so large an imbalance between matter and anti-matter as is observed. Thus the question remains.

There are, however, some theories Beyond the Standard Model (BSM) that suggest anomalous couplings with the Higgs boson (H) as a possible source of CP violation. This work focuses on the search for CP violation in interactions between the Higgs boson and W bosons via the HWW interaction vertex, using Effective Field Theory (EFT) to study and constrain a new additional CP-odd contribution to this coupling which may produce CP violation. The study here being reported is developed in the context of the ATLAS experiment at the LHC, as a part of an analysis probing the CP characteristics of leptonic WH associated production of a Higgs boson.

The goal of this work is to perform an interpretation of the differential cross section measurements of the WH associated production channel in terms of the strength of potential anomalous couplings, within the framework based on the SM fields. To this end, simulated events of the HWW interaction, with the H boson decaying into a pair of b-quarks and the W boson decaying leptonically ($pp \to WH \to \ell\nu b\bar{b}$) are used to study the distribution of a set of observables known to be sensitive to CP-odd contributions.

Following the standard of Higgs sector studies, the Simplified Template Cross Sections (STXS) framework [1] is used to classify the events in bins of specific kinematic and angular observables, to produce a parameterisation of the cross sections of the $pp \to WH \to \ell\nu b\bar{b}$ process in terms of BSM contributions that is sensitive to the CP-odd couplings. The resulting parameterisation coefficients can then be used in the fit of data from the ATLAS experiment, to constrain the possible magnitude of the CP-odd terms in the HWW vertex.

In Chapter 2, the theoretical basis of this work is explained and the latest results in this field are presented, including what has been achieved in CP-violation studies in the Higgs sector, and the measurement of CP-odd contributions. In Chapter 3, the LHC and the ATLAS detector are briefly described. Chapter 4 details the ATLAS study that provides a basis for this work. In Chapter 5, the methodology and results of this work are presented, including the study of the distributions of variables sensitive to anomalous couplings and the cross section parameterisation in terms of the STXS bins. The likelihood fit performed to constrain the limits of the CP-odd operator is also described, and its results are presented. Chapter 6 mentions the conclusions and future work of this analysis.

Chapter 2

The Standard Model of Particle Physics

2.1 Theoretical Framework

The Standard Model of particle physics (SM) took its present shape in the 1970s and describes the fundamental particles and the interactions between them. Although there are four fundamental forces, the model accounts for the effect of only three of them, as it cannot include the description of gravity. Nonetheless, the gravitational force between fundamental particles is so much weaker than the other interactions that it can be safely ignored.

In the Standard Model, the particles arise from quantum field excitations, so they are described using

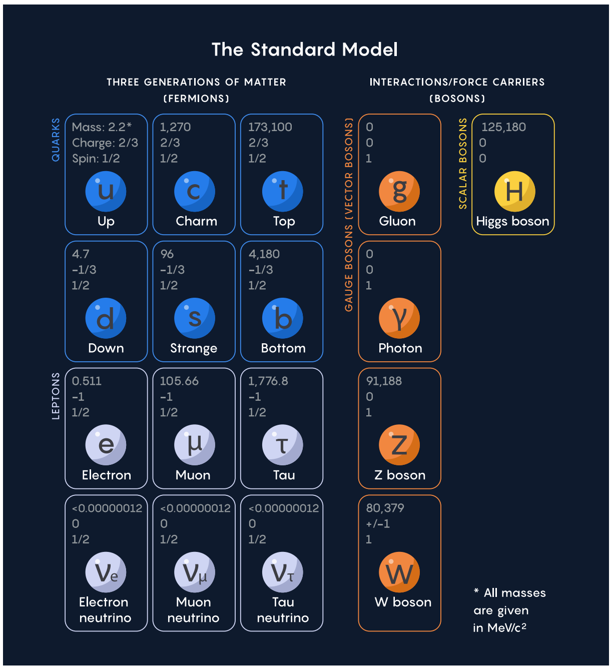


Figure 2.1: The Standard Model [2]

Quantum Field Theory Lagrangians, $\mathcal{L}$. Focusing on the electroweak sector, the dynamics of the fields are governed by the Glashow-Salam-Weinberg Model (GSW), which is a gauge theory requiring a U(1) and SU(2) local gauge symmetry [3]. To verify this symmetry, the gauge bosons, $W^{\pm}$ and Z, would have to be massless, in which case they could not mediate the weak processes they are known to be involved in. This means a mechanism needed to be introduced to solve this inconsistency [4], which is where the Higgs boson and field come in.
Despite making very accurate predictions confirmed by experiment, there are some observed phenomena that the SM does not explain, namely the fact that neutrinos have mass, nor does it incorporate gravity. Another phenomenon the SM does not explain is the asymmetry between matter and anti-matter in the universe. According to the theory, there was no imbalance between matter and anti-matter in the Big Bang, requiring the same amount of particles and anti-particles to have been produced. Observational evidence, however, shows that anti-matter has nearly fully disappeared, leaving a universe made up of matter [5].

2.1.1 The Higgs Field

The Higgs mechanism was first proposed in 1964 [6–8], allowing the SM to be consistent with massive gauge bosons while maintaining gauge symmetry. Besides this, the mechanism is also responsible for the other fundamental particles acquiring mass. Even though its existence was theorised in the 1960s, the Higgs boson was only discovered in 2012, by the ATLAS and CMS collaborations at CERN's LHC [9, 10].

The Higgs Mechanism

A generic Lagrangian includes two parts: a kinetic energy term, and a component that includes mass terms and interactions between the fields considered. In the SM, the Higgs field (ϕ) is a scalar field, which means that its Lagrangian takes the form of Eq. 2.1. The first term corresponds to the kinetic energy, where D is the covariant derivative, replacing the partial derivative, ∂. This guarantees the required SU(2) $\times$ U(1) local gauge symmetry. For SU(2), this symmetry arises from the middle term of the derivative (Eq. 2.2), which corresponds to the multiplication of $\boldsymbol{\sigma}$, the set of Pauli matrices and the symmetry generators, by $\mathbf{W}_\mu = \begin{bmatrix} W_\mu^{(1)} & W_\mu^{(2)} & W_\mu^{(3)} \end{bmatrix}$, which are three fields. As for the last term in Eq. 2.2, it provides U(1) symmetry, with B_μ corresponding to the final field being considered. g_W and g' are constants.

$$\mathcal{L}_H(\phi) = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{2.1}$$

$$D_\mu = \partial_\mu + \frac{1}{2}(ig_W\boldsymbol{\sigma}\cdot\mathbf{W}_\mu + ig'B_\mu) = \frac{1}{2}\begin{bmatrix} 2\partial_\mu + ig_W W_\mu^{(3)} + ig'B_\mu & ig_W(W_\mu^{(1)} - iW_\mu^{(2)}) \\ ig_W(W_\mu^{(1)} + iW_\mu^{(2)}) & 2\partial_\mu - ig_W W_\mu^{(3)} + ig'B_\mu \end{bmatrix} \tag{2.2}$$

The last two terms of Eq. 2.1 correspond to the potential for the ϕ field ($V(\phi)$), which we will focus on. A very relevant information about the Lagrangian potential is the position of its minimum values, corresponding to the lowest possible energy state. These are called the vacuum states, v, and depend on the two parameters of the potential: λ and μ^2. In order for a finite minimum to exist, λ has to be positive [3], but $V(\phi)$ has different possible behaviors, depending on the sign of μ^2. If $\mu^2 > 0$, the only possible minimum state require $\phi^\dagger\phi = 0$. If, on the other hand, $\mu^2 < 0$, then there is a set of degenerate minima.

In the SM, the Higgs field is a doublet of complex scalar fields. It can be written as Eq. 2.3, where ϕ is written in terms of ϕ^+ and ϕ^0, as there must be a charged, and a neutral field [3].

$$\phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{bmatrix} \tag{2.3}$$

The equation for the set of degenerate minima, corresponding to $\mu^2 < 0$, can be written in terms of these doublet components, as seen in Eq.2.4.

$$\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{2.4}$$

All the states that fulfill this requirement are possible vacuum states, but only one will be the physical state, which spontaneously breaks the symmetry of the potential. This is a key characteristic of the Higgs mechanism, represented in Figure 2.2, that allows other particles to acquire mass.

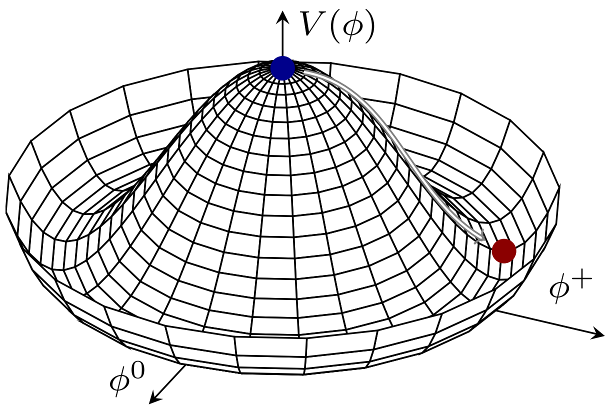


Figure 2.2: Standard Model Higgs potential. The red dot represents a broken symmetry. (Adapted from [11])

Because all states are equivalent, any can be chosen, without a loss of generality. This includes the one seen in the first part of Eq. 2.5. From here, ϕ can be described as perturbations about this minimum, as seen in the second part of the same equation, where $h(x)$ is the Higgs field. This, in turn, can be plugged into the Lagrangian of the Higgs field (Eq. 2.1), giving us the corresponding expanded Lagrangian shown in Eq. 2.6.

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v \end{bmatrix} \Rightarrow \phi(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v + h(x) \end{bmatrix} \tag{2.5}$$

$$\mathcal{L}_H = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{8}g_W^2(W_\mu^{(1)}W^{(1)\mu} + W_\mu^{(2)}W^{(2)\mu})(v+h)^2$$
$$+ \frac{1}{8}(g_W W_\mu^{(3)} - g' B_\mu)(g_W W^{(3)\mu} - g' B^\mu)(v+h)^2 - \frac{\mu^2}{2}(v+h)^2 - \frac{\lambda}{4}(v+h)^4 \quad (2.6)$$

The second term concerns the W^+ and W^- bosons. They relate to the $W^{(1)}$ and $W^{(2)}$ fields through Eq. 2.7. From this, the mass of these bosons can be found by equating this term to the mass term of the $W^{(i)}$ field, following Eq. 2.8, for $i = 1, 2$ [3].

$$W_\mu^\pm = \frac{W_\mu^{(1)} \mp iW_\mu^{(2)}}{\sqrt{2}} \quad (2.7)$$

$$\frac{1}{2}m_W^2 W_\mu^{(i)} W^{(i)\mu} = \frac{1}{8}g_W^2 v^2 W_\mu^{(i)} W^{(i)\mu} \Rightarrow \boldsymbol{m_W = \frac{1}{2}g_W v} \quad (2.8)$$

Repeating the same procedure for the third term, where the two neutral bosons - Z and the photon (represented by the field A) - are related to the $W^{(3)}$ and B fields through

$$Z_\mu = \frac{g_W W_\mu^{(3)} - g' B_\mu}{\sqrt{g_W^2 + g'^2}} \qquad A_\mu = \frac{g' W_\mu^{(3)} + g_W B_\mu}{\sqrt{g_W^2 + g'^2}}$$

From this, the masses of these two bosons are also acquired:

$$m_Z = \frac{v}{2}\sqrt{g_W^2 + g'^2} \qquad m_A = 0$$

The relation of Eq. 2.8 can be used, alongside the measurements of m_W and g_W to find the vacuum expectation value of the Higgs field, resulting in $v = 246$ GeV [3]. As to the mass of the Higgs boson, there was no prediction as to what its value should be. However, it has been measured by both the ATLAS and CMS collaborations, and simultaneously fitting the mass peaks obtained by the two experiments gives a Higgs boson mass value of 125.11 ± 0.11 GeV [12].

Adding to the Higgs $\mathcal{L}$ of Eq. 2.1 (I) the contributions related to the gauge boson fields (II), fermions (III) and fermion mass [13] (IV) gives the full SM Lagrangian of the Electroweak sector (Eq. 2.9). Here ψ correspond to the fermion fields.

$$\mathcal{L}_{\text{SM}} = \underbrace{(D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2}_{\text{(I)}} \underbrace{-\frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{\text{(II)}} + \underbrace{\overline{\psi}\not{D}\psi}_{\text{(III)}} - \underbrace{y_{ij}[\overline{\psi}_{iL}\phi\psi_{jR} + (\overline{\psi}_{iL}\phi\psi_{jR})^*]}_{\text{(IV)}} \quad (2.9)$$

The field strengths are given by:

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - ig_W[W_\mu, W_\nu], \qquad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

while, for γ_μ corresponding to the Dirac gamma matrices, and q to the fermion charge,

$$\not{D} = \gamma^\mu(i\partial_\mu - qA_\mu).$$

Finally, the Yukawa coupling, y_{ij}, characterises the interactions of fermions to the Higgs field from which their mass appears. The value of the coupling is not given by the Higgs mechanism, and is instead chosen to guarantee the mass of the fermions matches observational data, through $y_{ij} = \frac{m_f}{v}$ [3].

The Higgs Boson

Ever since the Higgs Boson was discovered, its properties have been continuously studied to achieve the most reliable characterisation of this particle. This includes its production and decay methods as well. There are several processes through which a Higgs boson is created in proton-proton collisions, seen in Figure 2.3. According to the SM [14], at a collision energy of $\sqrt{s} = 13$ TeV, the majority of events (87 %), correspond to the production of a Higgs boson via gluon-gluon fusion (ggF). Other relevant processes include vector boson fusion (VBF), accounting for 7% of produced Higgs bosons; associated production of a Higgs boson and a vector boson (VH), occurring 4% of the time; and generation of a Higgs boson together with a top or bottom quark, corresponding to 1% of events. This is in agreement with what is seen in the data [15, 16].

The Higgs boson has a very short lifetime, decaying after 1.8×10^{-22} s [17]. The most important Higgs

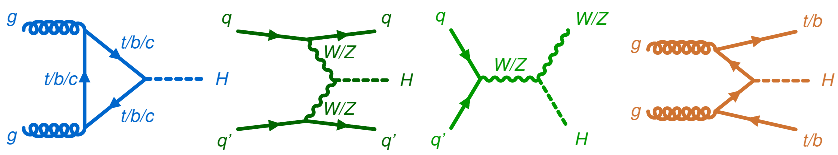


Figure 2.3: Feynman diagrams for the most relevant Higgs boson production processes. From left to right: gluon-gluon fusion, vector bosons fusion, VH associated production, and production of a Higgs boson with a top or bottom quark pair. (Adapted from [15])

decay processes can be seen in Figure 2.4, with their branching ratios mentioned in Table 2.1. There are other possible Higgs boson decay processes, which occur for under 1% of particles, so they are not mentioned here [14].

Table 2.1: Main Higgs boson decay processes and respective branching ratios [14]

Higgs Decay Process	Branching Ratio	Branching Ratio Theoretical Uncertainty (%)
$H \to b\bar{b}$	0.58	± 0.65
$H \to W^+W^-$	0.22	± 0.99
$H \to \tau^-\tau^+$	0.06	$+1.17 - 1.16$
$H \to ZZ$	0.03	± 0.99

2.2 Violation of Charge-Parity Symmetry

Except for their internal quantum numbers, like electric charge or lepton number, matter and anti-matter are indistinguishable, so the Big Bang would have produced the same amount of each. This is in

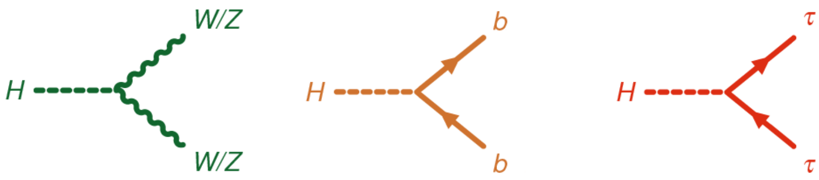


Figure 2.4: Feynman diagrams for the most relevant Higgs boson decay processes. From left to right, decay into a vector boson pair, into a pair of b-quarks, and into a pair of τ. (Adapted from [15])

opposition to what is seen in the present universe, which is nearly exclusively made up of matter. This asymmetry has to have arisen in the early universe, and to reconcile it with the theoretical predictions for the Big Bang, a set of conditions was proposed - the Sakharov conditions. In his 1966 paper [5], Sakharov enumerates these conditions: the need for interactions that violate baryonic number; which occur away from thermodynamic equilibrium; and that violate charge conjugation symmetry, as well as CP [18]. The LHC is well-suited in this search for CP violation, since, as it can access increasingly higher energies, it can reproduce the conditions in the early universe, and allow to better understand this baryonic asymmetry. Out of these aforementioned conditions, this work focuses on charge-parity symmetry violation (CPV).

The Standard Model has one source of CP violation, in the Cabibbo-Kobayashi-Maskawa matrix (CKM) that describes quarks interacting through the weak force. These processes, however, do not account for the magnitude of the asymmetry found in the universe, which means other sources are needed to explain this phenomenon [3].

This creates a need to look beyond the SM for possible sources of CPV that would explain the observed asymmetry. One of the simplest of BSM theories is the Two-Higgs Doublet Model (2HDM), in which the Higgs field is described by two doublets, instead of just the one of Eq. 2.3. This means that the Higgs mechanism described in Section 2.1.1 would produce five, instead of one SM Higgs boson. Among these, there would be one CP-odd boson, opening the door for CP violation in interactions with this particle [19]. A CP-odd particle involved in an interaction undergoing a charge conjugation and parity transformation produces a change of sign, as opposed to a CP-even one. Only a CP-odd behaviour can allow for CP violation.

2.3 Effective Field Theory

To study theories BSM in the search for new physics phenomena, it is useful to have a formalism that could be employed generally, independent of a specific model. In this pursuit, effective field theory (EFT) is particularly helpful.

EFT simplifies an analysis by separating the fields in a model by their energy scale, into light and heavy fields. An effective theory of the light fields is only valid for energy scales much lower than the masses of the heavy fields, so these have very little influence on the processes studied, and can therefore be integrated out. The New Physics scale, Λ, is the energy value that limits the validity of this approach, requiring that the energies this theory focuses on are much smaller than Λ. EFT makes the physics easier to handle since removing the heavy fields transforms some processes into an effective vertex, with much simpler calculations. This theory breaks down, however, when approaching the scale of the heavy fields, where their dynamics need to be taken into account to fully describe a model [20].

The particular framework used in this work is Standard Model effective field theory (SMEFT), which uses the SM fields and symmetries to build the framework. Considering only the electroweak sector, SMEFT adds to the Lagrangian of Eq. 2.9 a set of operators acting on the SM fields. These operators alter the physics of certain interactions and depend on the BSM model considered. To have a renormalisable theory, $\mathcal{L}$ has to have terms of mass-dimension 4 [20], which means that, as the order of operators increases, they are scaled by increasing powers of $\frac{1}{\Lambda}$. After adding these new terms, $\mathcal{L}$ becomes that of Eq. 2.10, which works for any BSM theory using the SM fields.

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda}\sum_i c_i^{(5)} O_i^{(5)} + \frac{1}{\Lambda^2}\sum_j c_j^{(6)} O_j^{(6)} + \cdots \tag{2.10}$$

In this equation, the sums run through every possible operator, $O^{(n)}$, of order n. A coefficient (c) affects each term - the Wilson coefficient - modifying the magnitude of each operator contribution.

There is only one mass-dimension 5 operator in SMEFT, the Weinberg operator, in which lepton number is not conserved [20]. Since the processes in this work conserve lepton number, this operator is not relevant, and can therefore be disregarded. The higher order operators have increasingly little influence on the processes, which means the terms above order 6 can be discarded, leaving only $O^{(6)}$ to be considered.

The cross section of a process is proportional to the square of the amplitude of the matrix element of the interaction, and thus to the square of the Lagrangian. This means the cross sections in this framework can be written as in Eq. 2.11, for a set of i operators and excluding quadratic cross terms between different operators. The first term corresponds to the cross section of all the processes encoded in the Standard Model. As it has no dependence on the operator, it is referred to as order 0. The last element arises from the Lagrangian portion featuring the operator. The operator contribution appears squared (and so does the Wilson coefficient), so this term is referred to as the order 2, or quadratic term. The middle term corresponds to the cross term, featuring the interference between the SM and operator contributions. It is linear in the operator, being therefore called order 1.

The quadratic term of the operator can only produce a CP-even term, but the order 1 term has a linear dependence on the operator, so it is here that a CP-odd contribution can appear. This makes it the only possible term where CP violation can be featured.

$$\sigma_{\text{SMEFT}} = \sigma_{\text{SM}} + \sum_i \left(\frac{c_i}{\Lambda^2}\sigma_{i,\text{int}} + \frac{c_i^2}{\Lambda^4}\sigma_{i,\text{BSM}} \right) \tag{2.11}$$

2.3.1 The HWW Vertex

This work focuses on studying CPV in interactions between the Higgs and the W bosons. The chosen coupling to perform this analysis is the HWW vertex in which a Higgs boson is produced by associated production with a W boson. The Feynman diagram for HW associated production, originating from proton collisions like those occurring at the LHC, can be seen in Fig. 2.5.

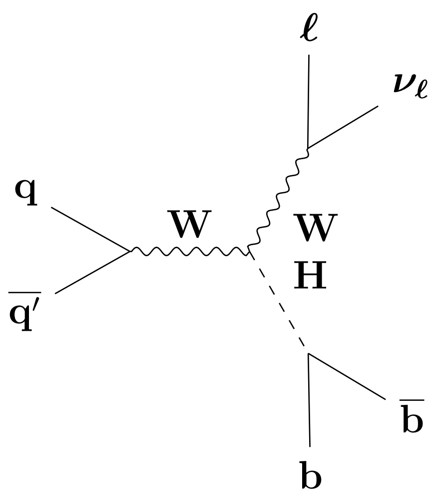
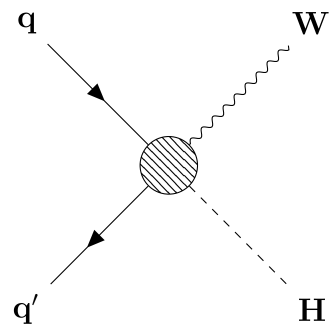
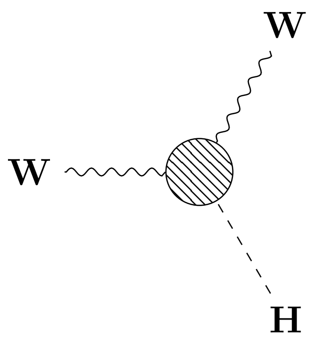


Figure 2.5: Feynman diagram of the HW associated production process

There are some advantages of studying interactions with the W through WH associated production. In VBF the only boson appearing in the final state is the Higgs boson, which makes the HWW and HZZ vertices indistinguishable. A Higgs boson decaying into a W^+W^- is also a HWW vertex, but its analysis may be less susceptible to BSM effects, due to the kinematic cuts that are implemented during selection and reconstruction of events [21], making this process, therefore, not adequate to study the relevant vertex.

There are only four CP-even mass-dimension 6 operators that affect the HWW vertex, referred from here on out by their respective Wilson coefficient. All can be seen in Table 2.2. Additionally, there is a CP-odd operator also considered, shown in the same table.

Table 2.2: Operators affecting WH associated production with their respective Wilson coefficients in the Warsaw basis and the vertices they affect. p is a quark flavour index, while I is an isospin index. $\tau = \frac{1}{2}\sigma$, σ being the Pauli matrices. $h^\dagger i\overleftrightarrow{D}^I_\mu h \equiv h^\dagger i\tau^I D_\mu h - (iD_\mu \tau^I h^\dagger)h$ and $\tilde{W}^I_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}W^{I\rho\sigma}$, where $\epsilon_{\mu\nu\rho\sigma}$ is the antisymmetric tensor. [22–25]

	Wilson coeff.	**Operator**	**Vertex**
CP-even	$c^{(3)}_{Hq}$	$(h^\dagger i\overleftrightarrow{D}^I_\mu h)(\overline{q}_p \tau^I \gamma^\mu q_r)$	
	$c_{H\Box}$	$(h^\dagger h)\Box(h^\dagger h)$	
	c_{HW}	$h^\dagger h W^I_{\mu\nu} W^{I\mu\nu}$	
	c_{HDD}	$(h^\dagger D^\mu h)^*(h^\dagger D_\mu h)$	
CP-odd	$c_{H\tilde{W}}$	$h^\dagger h \tilde{W}^I_{\mu\nu} W^{I\mu\nu}$	

2.4 State of the Art

2.4.1 CP violation in the Higgs sector

Studies of CP violation in the Higgs sector can be divided into two categories: couplings with fermions, and interactions with vector bosons, because of the different characteristics of the couplings.

Fermionic couplings

The SMEFT Lagrangian for the Yukawa coupling between the Higgs boson and fermions (f) is given by Eq. 2.12 [26]. k_f corresponds to the Yukawa coupling scaling factor, while α corresponds to the CP-mixing angle. This angle quantifies the CP-even and CP-odd contributions to a coupling, with $\sin\alpha = 1$ ($\alpha = 90°$ or $\alpha = 270°$) corresponding to a purely CP-odd interaction and $\alpha = 0$ or $180°$ ($\cos\alpha = 1$) leading to a pure CP-even one. If $\alpha = 0$ and $k_f = 1$, the SM coupling is recovered. This equation also shows that CP-odd contributions are added to the Lagrangian at tree level, without this term being suppressed by any power of Λ, unlike what happens for interactions between the Higgs and vector bosons.

$$\mathcal{L}_{\text{SMEFT, HFF}} = -\frac{m_f}{v}k_f(\cos(\alpha)\overline{\psi}\psi + \sin(\alpha)\overline{\psi}i\gamma_5\psi)h \tag{2.12}$$

The ATLAS experiment probed CP violation in ttH and tH production, with the Higgs boson decaying into two photons [27] or into a pair of b quarks [28]. Proton-proton collision data from Run 2 was used, with a luminosity [1] of 139 fb^{-1} and a center of mass (CM) beam energy of $\sqrt{s} = 13$ TeV, to constrain the mixing angle α, and, thus, the CP-odd contribution. A value of $|\alpha| > 43°$ was excluded at a 95% confidence level (CL), which rules out the possibility of having a purely CP-odd coupling [27]. A best-fit value was achieved in [28] of $\alpha = 11°^{+55°}_{-77°}$. The CMS experiment performed a similar study [29] on CP proprieties of ttH and tH couplings, including $H \to WW$, $H \to \tau\tau$, $H \to ZZ$ and $H \to \gamma\gamma$ decays, at a luminosity of 138 fb^{-1}. The possibility of a purely CP-odd coupling was also excluded, finding that $k_t \cos(\alpha) \in [0.86, 1.26]$ and $k_t \sin(\alpha) \in [-1.07, 1.07]$, at a 95% CL.

Bosonic couplings

When studying CP violation in interactions between the Higgs and vector bosons (HVV), the goal is often to constrain the values of the CP-odd Wilson coefficients, and thus the possibility for CPV. Several studies using data from the various channels have been performed, using observables sensitive to the various operators to constrain the coefficients. Some of these are mentioned in this section, culminating in a compilation of the most recent results.

The ATLAS experiment has published a series of studies, mostly focusing on VBF, as this is a production process with a large cross section. A group of these analyses focuses on a matrix element-based observable sensitive to CP violation, with different distributions for CP-even and CP-odd interactions, and which can be used for any Higgs decay [30]. Figure 2.6 shows the constraints on a set of CP-odd operators using this procedure for studies of VBF, where $H \to ZZ \to 4\ell$ and $H \to \gamma\gamma$ decays are considered [30, 31], again using data from proton-proton collisions with the proton beams having center of mass energy of $\sqrt{s} = 13$ TeV. The luminosity of the accelerator was 139 fb^{-1}. Besides $c_{H\tilde{W}}$, two other operators were considered, both of which affect only the HZZ vertex, $c_{H\tilde{B}}$ and $c_{H\tilde{W}B}$ [22], corresponding, respectively to the operators in Eq. 2.13 and 2.14 [25]. Together with $c_{H\tilde{W}}$, these are the only three CP-odd operators affecting the HVV vertices.

$$O_{H\tilde{B}} = h^\dagger h \tilde{B}_{\mu\nu} B^{\mu\nu} \qquad (2.13) \qquad\qquad O_{H\tilde{W}B} = h^\dagger \tau^I h \tilde{W}^I_{\mu\nu} B^{\mu\nu} \qquad (2.14)$$

VBF studies with the Higgs boson decaying into two W bosons are particularly interesting for this work since the HWW interaction can be present in both the production and decay vertices. One of these analyses was performed by the ATLAS collaboration, described in [32], studying the $H \to WW^* \to e\nu\mu\nu$ channel, using data in the same conditions as mentioned above. This analysis relies on an angular observable, $\Delta\phi_{\ell\ell}$, the difference between the azimuthal angle of the electron and muon, as this is a variable also sensitive to CP-odd couplings, with CP-odd operators creating an asymmetry in its distribution. The three operators were studied separately, by keeping the others at $c = 0$. The results in constraining the Wilson coefficients are also shown in Figure 2.6. As seen in this plot, all these results are consistent with the Standard Model predictions.

The most recent constraints on CP violation in the HWW vertex from the ATLAS experiment pertains

[1] This quantity refers to the integrated luminosity of the LHC and related to the number of events produced (see Chapter 3).

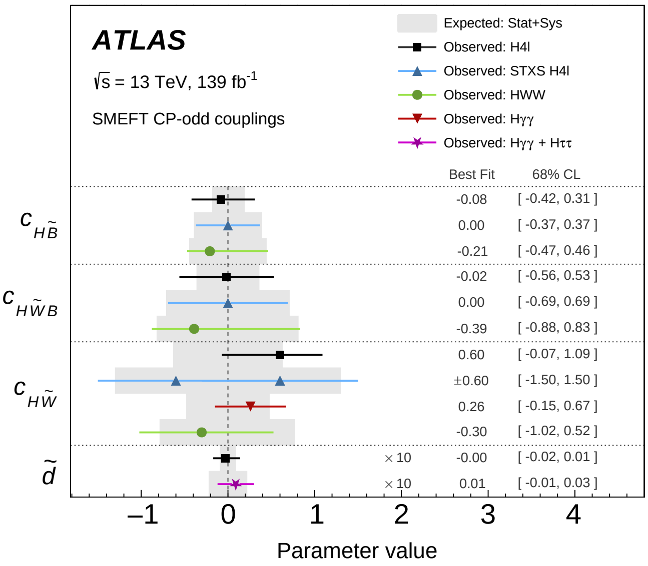


Figure 2.6: CP-odd Wilson coefficient results from VBF Higgs production and varying decay processes, considering both linear and quadratic terms for all channels but $H \to ZZ \to 4\ell$, where the quadratic contribution is negligible and only the linear term is included. The bars correspond to a 68% CL showing statistical and systematic uncertainties. $\tilde{d}$ corresponds to operators in a basis other than the Warsaw basis and is not relevant to this work. [26]

to the study of VBF production of the Higgs boson, with this boson decaying through $H \to \tau^+\tau^-$ [33]. Using data from Run 2 of the LHC, with a luminosity of 140 fb^{-1} and $\sqrt{s} = 13$ TeV, this study considered $c_{H\tilde{W}}$, $c_{H\tilde{B}}$ and $c_{H\tilde{W}B}$, and performed a binned analysis for two cases: considering only $\Delta\phi_{jj}$, the difference between the azimuthal angles of the two leading jets produced alongside the Higgs boson in VBF, and studying this variable alongside with the transverse momentum of the Higgs boson, p_T^H. There was a particular focus on $c_{H\tilde{W}}$, since $\Delta\phi_{jj}$ shows an asymmetric distribution for a non-zero $c_{H\tilde{W}}$, and a symmetric one when considering CP-even operators. Further dividing the data into bins of p_T^H increases the sensitivity to the CP-odd coupling.

The aforementioned analysis culminated in constraining the values of the coefficients of the CP-odd operators coefficients (Figure 2.7), as well as its CP-even counterparts (c_{HW}, c_{HB} and c_{HWB}), again considering all operators separately and setting the others at $c = 0$. Again these results are in agreement with the SM predictions, but this work was able to produce the strongest constraint on $c_{H\tilde{W}}$ up to date, expected to be within [-0.60,0.60] at a 95% confidence level.

The CMS collaboration, on the other hand, has performed a CPV study in Higgs associated production, considering both W and Z bosons decaying leptonically and the Higgs boson decaying through $H \to b\bar{b}$, as described in [34]. This analysis was, therefore, concerned with the same processes as is the focus of this thesis. Data from proton collisions at the LHC with a CM energy $\sqrt{s} = 13$ TeV and for a luminosity of 138 fb^{-1} is used, and the three CP-odd operators of HVV vertex, $c_{H\tilde{W}}$, $c_{H\tilde{B}}$ and $c_{H\tilde{W}B}$ are considered. In this study new variables are built from linear combinations of these Wilson coefficients, forming a new mass-eigenstate basis that makes the analysis depend on fewer parameters. Out of these, only one can be constrained in studies of VH associated production, so it will be the only variable mentioned. It can be found in Eq. 2.15, where θ_w is the weak mixing angle. A similar variable, g_2^{ZZ} can be built for the

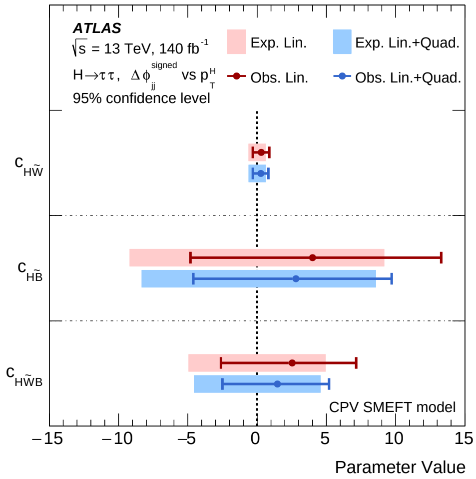


Figure 2.7: CP-odd Wilson coefficient results from VBF Higgs production and $H \rightarrow \tau^+\tau^-$, computed considering the distribution in bins of $\Delta\phi_{jj}$ and p_T^H simultaneously. [33]

CP-even operators affecting VH associated production, simply by replacing the CP-odd operators in the expression with their CP-even counterpart.

$$\tilde{g}_2^{ZZ} = -2\frac{v^2}{\Lambda^2}(\sin^2\theta_w c_{H\tilde{B}} + \cos^2\theta_w c_{H\tilde{W}} + \sin\theta_w \cos\theta_w c_{H\tilde{W}B}) \tag{2.15}$$

The values of these variables were then constrained by keeping the values of all coefficients as free parameters. Again, the results were in accordance with the SM predictions, with $\tilde{g}_2^{ZZ} \in [-0.366, 0.352]$ at a 68% CL, and $[-0.601, 0.615]$ at a 95% CL. As for the CP-even variable, $g_2^{ZZ} \in [-0.259, 0.409]$ at a 68% CL, and $[-0.573, 0.614]$ at a 95% CL.

Chapter 3

The LHC and the ATLAS experiment

This chapter briefly details the components of the LHC and the ATLAS detector, as well as some of their principal characteristics. It also describes some object reconstruction procedures relevant in the study of $q\bar{q}' \to WH \to \ell\nu b\bar{b}$.

3.1 The Large Hadron Collider

The LHC at CERN is a particle accelerator at the forefront of searches in high-energy physics, and the ATLAS experiment employs the largest general-purpose detector at the accelerator. The LHC has been running since 2009, having had three runs so far, separated by shutdown periods. Run 1 lasted from 2010 to 2012, while Run 2 spanned from 2015 to 2018. It is currently on its Run 3 which started in 2022 and will last until 2026.
This work uses simulated data of the ATLAS detector for Run 2, so this chapter focuses on the characteristics of the detector and the LHC during this period. Besides this, even though the accelerator can be used to collide other particles, like heavy lead ions, as this work is only concerned with proton-proton collisions, they are the only ones mentioned.
Two quantities characterise each run of the LHC, the first of which is the center of mass energy of the proton beams colliding, $\sqrt{s}$. During Run 2, this value was $\sqrt{s} = 13$ TeV, and the higher the energy of the proton beams, the higher the energy of the phenomena that can be probed in the collisions. The other one is the instantaneous luminosity of the accelerator. It gives the number of events that occur per unit area and time and can be integrated throughout the operation of the accelerator to find the integrated luminosity, following Eq. 3.1. $\mathcal{L}$ is the instantaneous luminosity, t is the operation time of the accelerator being considered, and N is the resulting number of events of a process with cross section σ.

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

The goal is to have an instantaneous luminosity as large as possible, so as to increase the statistics available for the various processes. With the improvement of technology, it has been increasing ever since the first LHC run. Figure 3.1 shows the integrated luminosity achieved for different years of Runs 1 and 2, but, the increasing slope of the plots for more recent data-taking, when comparing the same period of each year, shows the improvement in instantaneous luminosity in the LHC. As for the integrated luminosity, the ATLAS experiment measured it to be 147 fb^{-1} for Run 2. This is an improvement from Run 1, which delivered an integrated luminosity of 5 fb^{-1} for $\sqrt{s} = 7$ TeV and 21 fb^{-1} for $\sqrt{s} = 8$ TeV. As for Run 3, it is expected to reach a luminosity of 250 fb^{-1} over the four years of operation, for $\sqrt{s} = 13.6$ TeV [35].

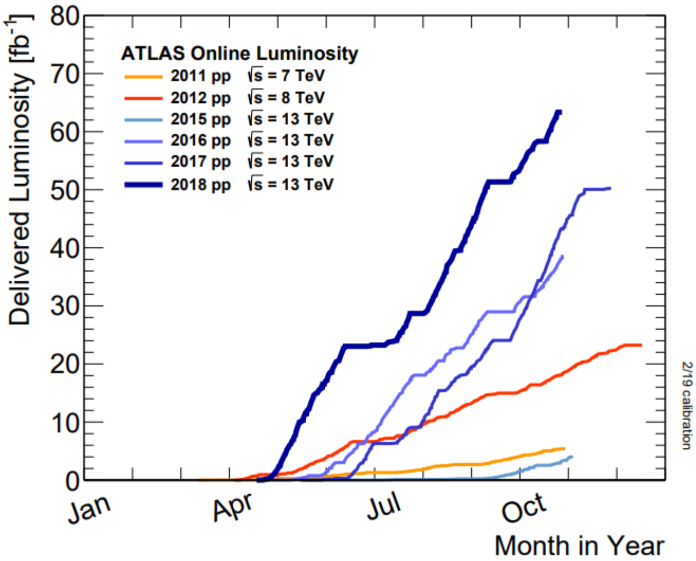


Figure 3.1: Integrated luminosity throughout the years delivered to the ATLAS detector with stable proton beams. [35]

3.2 The ATLAS detector

The ATLAS experiment is a general purpose detector that investigates a wide range of physics at the LHC. A schematic representation of the detector can be seen in Figure 3.2. The protons are made to collide at the center of the detector, and the different particles produced interact with the detector components, producing signatures that are then interpreted to identify them.

The ATLAS detector can cover nearly the full 4π solid angle around the collision point. Figure 3.3 shows the coordinate system used by the ATLAS experiment, where z corresponds to the beam axis, x is the direction from the collision point to the center of the LHC, and y points upwards. ϕ is the azimuthal angle about z, and θ is the polar angle, from the beam axis to the three-momentum of the particle, $\vec{p}$. The transverse momentum of a particle, $\vec{p_T}$, is the projection of $\vec{p}$ in the $x - y$ plane. Instead of using θ, however, it is the standard in the LHC experiments to use the pseudorapidity, η, which is related to θ through $\eta = -\ln\tan(\theta/2)$. This is done because in proton-proton collisions, the distribution of the proton momentum between the quarks and gluons it is composed of is not known. This means the momentum of the interacting particles is not known, and neither is their boost in the z direction, relative

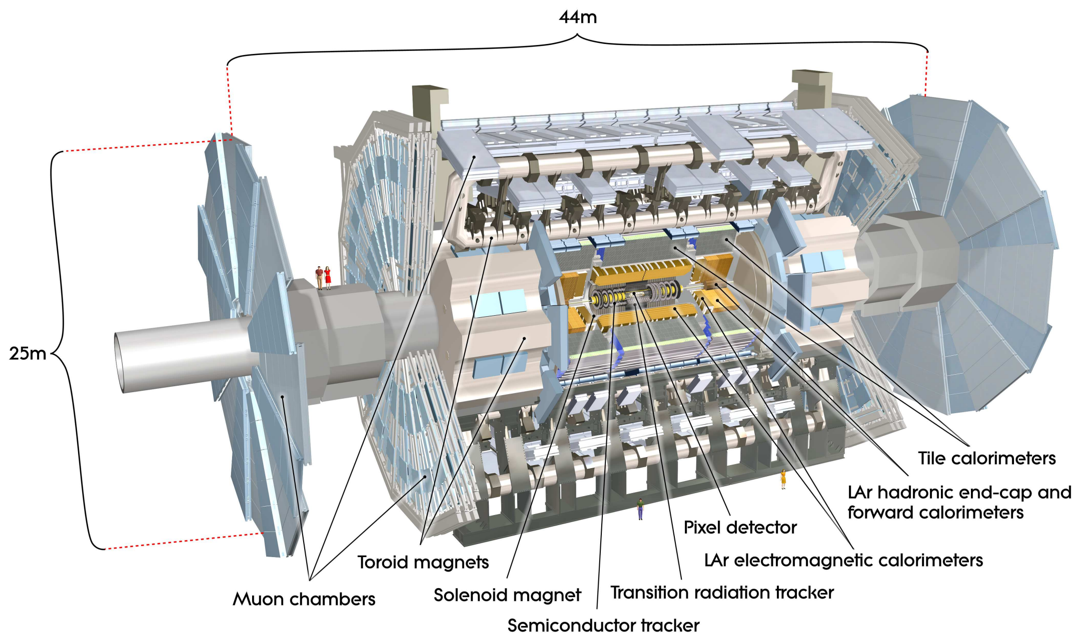


Figure 3.2: Schematic of the Run 2 ATLAS detector [36]

to the lab frame. The pseudorapidity then proves to be a useful variable, as the difference in η between two particles is independent of boosts in the beam direction [37]. Furthermore, the ATLAS detector can be divided into two regions depending on this variable: the central region, corresponding to $|\eta| < 2.5$, and the forward region, for $2.5 < |\eta| < 4.9$.

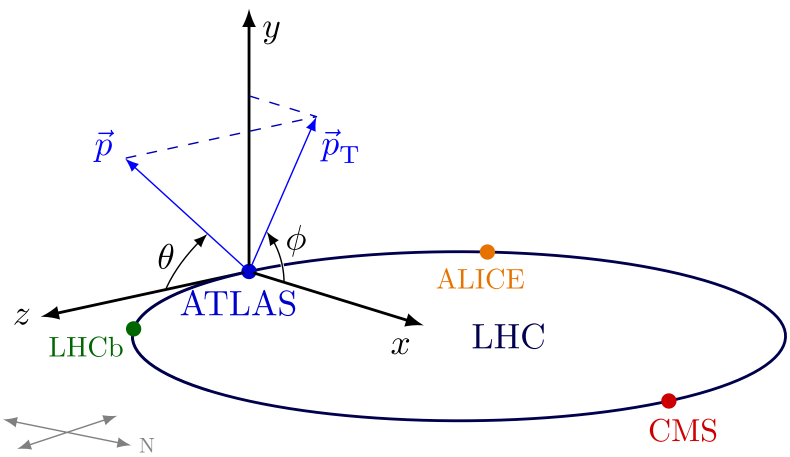


Figure 3.3: Coordinates system used by the ATLAS experiment (Adapted from `https://tikz.net/axis3d_cms/`)

3.2.1 ATLAS Sub-Detectors

Going from the collision point outwards, the particles resulting from the proton interactions first encounter the inner detectors (ID). This detector set is formed by a pixel detector, a semiconductor tracker, and the transition radiation tracker (TRT), covering a range of $|\eta| < 2.5$. All of these components contribute to information on the track of charged particles, with the TRT using the characteristic radiation

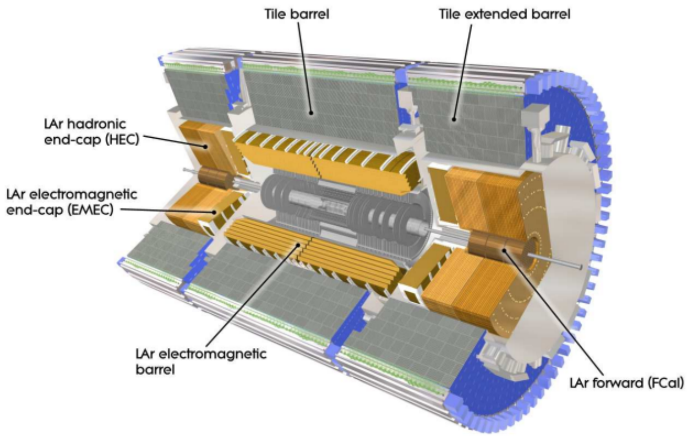


Figure 3.4: Schematics of the ATLAS detector calorimetry system [36]

that particles emit when passing through inhomogeneous media to allow for the better identification of electrons. A superconducting solenoid magnet around the ID provides a 2 T magnetic field in the direction of the beam axis, and is responsible for bending the trajectory of charged particles depending on their electric charge [35].

The next layer consists of a system of hadronic and electromagnetic calorimeters (Figure 3.4). This is where most particles are stopped, generating cascades of new particles and allowing for their energy to be measured. These calorimeters can be divided into two categories, depending on the type of particles they interact with. Lead and liquid argon electromagnetic calorimeters, both at the barrel and endcaps, are responsible for stopping photons and electrons, in the range of $|\eta| < 3.2$. Detecting hadrons requires another set of calorimeters. One made of scintillating tiles surrounds ID and can detect particles found at $|\eta| < 1.7$. There is also an endcap liquid argon hadronic calorimeter covering up to $|\eta| = 3.2$. To be able to account for particles traveling in nearly all directions, there are both electromagnetic and hadronic liquid argon forward calorimeters at the center of the endcaps that can measure particles in the forward region, up to $|\eta| < 4.9$ [35, 36].

There are, however, two kinds of particles, that do not end their trajectory at the calorimeters: the muons and the neutrinos. A muon spectrometer system surrounds the calorimeters, with detectors at the barrel and endcaps. This system has great precision in tracking muons, working together with toroidal magnets that deflect the muon trajectory to achieve high-resolution measurements of the momentum of these particles, up to a pseudorapidity of $\eta = 2.7$ [36]. The presence of neutrinos, however, has to be inferred from the presence of missing transverse momentum, a quantity that should sum up to zero by momentum conservation.

3.2.2 Trigger

The collisions at the LHC produce a very large amount of data every second and it is impossible to store the full volume at such a fast rate. If only a fraction of all collisions can be saved, one needs a system capable of choosing the events that can potentially harbour new physics. This is where the trigger system comes in.

The first step when selecting events is the Level 1 hardware trigger, which uses electronics to reduce the amount of data to be stored. It makes a decision within a $2.5\ \mu s$ latency and can reduce the data output rate from 40 MHz (at the detector output) to 100 kHz (after the trigger) [38]. This system makes very fast decisions, selecting events with high-momentum particles and with large missing transverse energy, based on information from the calorimeter and muon chambers.

The events that pass the selection of the Level 1 trigger are then directed to the High-Level Trigger (HLT), which is based on software. The HLT may use the regions of interest of the detector, flagged by the Level 1 trigger as containing objects that passed its selection criteria, or the full event information [36]. It does a more thorough evaluation of the event to verify if it indeed meets the criteria required, looking at the information not only from the calorimeters and muon chambers but also at the tracks from the inner detector, and using the full granularity of the detector. After passing through the HLT, the data is stored at a more feasible output rate of 1.2 kHz [38].

3.3 Object Reconstruction

Information from the different sub-detectors can be combined to reconstruct the particles that make up the final state of an event. Firstly, the tracks from the inner detector are matched to the vertices that originated the particles. The vertex with the largest squared transverse momentum sum is chosen as corresponding to the primary hard scattering vertex [39].

3.3.1 Electrons, Photons and Muons

Electrons are reconstructed by corresponding tracks from the inner detector to deposits of energy in the electromagnetic calorimeter. Naively, since the electrons are the only particles that leave both energy deposits in this calorimeter and a track in the ID (as the photons do not leave a track), this process should uniquely identify the electrons. There is, however, the possibility that a photon in an event produces an electron-positron pair in the ID (converted photon). There would be a cluster of deposited energy in the calorimeter and two tracks in the inner detector, and to identify this phenomenon, these signatures must be matched to a conversion vertex. Unconverted photons, have a signature made up of only an energy deposit in the calorimeters, which can also be reconstructed as converted photons [40]. The presence of these phenomena requires that likelihood-based evaluation criteria be used to evaluate electron candidates, based on information from the ID and electromagnetic calorimeter [41].

For an increasing number of rejected events, electron identification can meet loose, medium, or tight criteria, which correspond to 93%, 88%, and 80% classification efficiencies, respectively [40]. Similarly, to improve hadron background rejection further isolation criteria of electrons are used. Electrons are required to meet loose isolation and identification criteria.
Electrons candidates are also required to have a pseudorapidity $|\eta| < 2.47$, as this is the range of acceptance that guarantees efficient reconstruction of the electrons and background rejection in the ATLAS barrel, and the transverse momentum is required to be above 7 GeV, the minimum limit usually considered [41]. As to the ID tracks, there are restrictions on the quantities that characterise them as well. They are required to have $z_0 \sin\theta < 0.5$ mm and $|d_0|/\sigma_{d_0} < 5$ [39], where z_0 is the longitudinal impact parameter in relation to the primary vertex, while d_0 is the transverse one. Its uncertainty is given by σ_{d_0}, with their ratio corresponding to the significance of d_0 [40]. While the restriction on d_0 is related to its use in the likelihood, the restriction on z_0 is required as a parameter for isolating the electrons [40].
As for the muons, their reconstruction relies on matching the ID tracks to information from the muon spectrometer, instead of the calorimeter. Regarding muon acceptance, the muon spectrometer guarantees an efficient reconstruction for $|\eta| < 2.7$, and $|d_0|/\sigma_{d_0}<3$ is also required [42]. The limits on z_0 and transverse momentum are the same as for the electrons, and, like these particles, muon candidates are required to meet loose identification and isolation criteria [39].

3.3.2 Jets

The shower and hadronisation processes of quarks and gluons (and subsequent hadron decays) produce collimated sprays of hadrons called jets that can provide information about the partons that originated them [1]. These collimated sprays are clustered into jets via jet clustering algorithms. Jets act as proxies to the parton originating the spray.
The jets are created using information from the tracks in the inner detector, coupled with the energy deposits in both the electromagnetic and the hadronic calorimeters. They are defined using the anti-k_t jet clustering algorithm [43]. This algorithm groups the particles with the smallest distance between them, defined by their angular distance in the $\eta - \phi$ plane, multiplied by the inverse momentum of the particle with the largest transverse momentum. It starts with the most energetic particles, whose momentum and position are unlikely to be affected by other particles, and groups around them the less energetic ones, forming a cone. This algorithm is effective in the study of jets, clustering together particles likely to have the same origin [43] [44, Ch. 2]. One of the most important variables in this algorithm is the radius parameter of jets, set at $R = 0.4$ (small-R jets) or $R = 1.0$ (large-R jets).
If small-R jets are found in the central region ($|\eta|<2.5$), they are required to have a minimum transverse momentum of 20 GeV. In the forward region ($2.5<|\eta|<4.9$), outside of the acceptance of the inner detector, they are required to have a transverse momentum $p_T>30$ GeV [39]. This cut is performed to guarantee the efficiency and purity of jet reconstruction [36]. Furthermore, a likelihood-based jet-vertex-

[1] The final state includes not only high-energy particles originating from hard scattering, but also other low-energy ones, coming from additional partonic interactions in the collision (underlying event).

tagging procedure can also be applied to low-momentum jets with $p_T<60$ GeV in the central region and $p_T<120$ GeV in the forward one to better remove jets originating from pile-up interactions, which are more likely to have low energy [45].
As for large-R jets, they may include more energy deposits belonging to pile-up and underlying event, since these jets encompass a larger area of the detector. This means that a procedure is needed to remove unwanted energy contributions, namely using jet trimming. This method consists of subdividing the jet into smaller ones and then removing the sub-jets with the lowest transverse momenta, and keeping the rest [46]. It is also required that the large-R jets have $|\eta|<2.0$, which means they are found in a more interior region of inner detector, avoiding the outer regions of the ID where reconstruction is not as reliable [45].
One important feature of jet reconstruction is flavour tagging, where a jet can be classified as a b- or c-jet, depending on whether it contains a b- or a c-hadron. If a jet fits neither of these classifications, it is called a light jet. The B hadrons present have a long lifetime, which means they travel a few millimeters on average before decaying. This produces a signature of a secondary vertex emerging some millimeters away from the primary one, roughly along the same direction as the jet [3]. In addition, B-hadrons have a large mass, so the particles resulting from their decay tend to have large momenta, which does not happen when originating from light quarks. These two characteristics are used in b-jet identification [47].
The DL1r tagger developed by the ATLAS collaboration [48] uses a neural network to find the probability of a jet being a b- or c-jet. From this, the b-score of a jet can be calculated, with a high probability of having a b-hadron leading to a high b-score, and a high probability of having a c-hadron or no heavy hadrons leading to a low b-score. The same is done for c-jets, but as the lifetime of c-hadrons is shorter than that of b-hadrons, the efficiency of this process is lower [49]. If a jet meets the 70% operating point, which corresponds to a tagging efficiency of 70% in a specific test dataset, the jet is classified as a b-jet. If this does not happen, but the jet meets the 45% operating point for c-tagging, the jet is classified as a c-jet [39, 49].

3.3.3 Neutrinos

Neutrinos interact very little with matter. As such, these particles cannot be detected directly by the ATLAS detector. Their characteristics can, however, be inferred from the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) of an event. This variable corresponds to the symmetric of the sum of all detected momentum in the transverse direction, to guarantee the conservation of transverse momentum after the collision, which is zero before it. To calculate $E_{\mathrm{T}}^{\mathrm{miss}}$, the momenta of all reconstructed particles are included, along with an additional term corresponding to tracks in the inner detector left by charged particles that originate from the hard-scattering vertex of the event, but do not correspond to an object reconstructed in the detector. By requiring that the tracks be matched to the interaction vertex, the possible influence of pile-up is reduced [50].

In the case of there being a W boson that decays into one charged lepton and a neutrino, it is assumed that the neutrino momentum corresponds to $E_{\mathrm{T}}^{\mathrm{miss}}$, in the transverse direction. The longitudinal component of its momentum, however, is calculated from the kinematic constraint that the squared module of the neutrino and charged lepton four-momenta must yield the squared mass of the W-boson from which they originated [39].

Chapter 4

WH associated production measurements with the ATLAS detector

In this chapter, the ATLAS analysis providing the basis for this work is detailed, along with the modifications done in this thesis to be able to constrain the values of the CP-odd operator.
The legacy Run 2 analysis by the ATLAS experiment of VH associated production [39] is mainly concerned with measuring differential cross sections, to constrain the Yukawa coupling in Higgs bosons decaying into b or c quark pairs. Although the legacy analysis focuses on WH and ZH associated production, only the production of a Higgs and a W boson will be regarded, as this is the process that concerns this thesis. In [39], the analysis focuses on measurements of the cross sections of the WH(bb) process, which is also what is done for this work, so it provides a good guideline for this thesis. The legacy analysis uses 13 TeV pp data from Run 2 of the LHC, for signal and most backgrounds. Since the analysis does not take CP-odd couplings into account, modifications have to be made so it can be suitable for probing these interactions.

4.1 Event Generation

The focus of the analysis is the vertex in Figure 2.5, but the W and H bosons cannot be measured directly, since they decay too quickly to be detected. Only the stable products of this decay can be observed in the detector. As mentioned before, the leptonic decay of W bosons provides a cleaner signature for detection, since the majority of backgrounds are made up of hadrons [1]. As for the Higgs boson, the decay channel considered is into a $b\bar{b}$ quark pair, since this is the dominant decay.

[1] In this thesis, only W bosons decaying into electrons and muons (and their corresponding neutrinos) are considered. As for τ , its lifetime is 290×10^{-15} s [17], so it cannot be detected directly. It can, however, decay into muons or electrons, which are stable. If so, the W boson decaying into a τ (and its neutrino) is included, since electrons and muons are the charged leptons the thesis considers. If the tau decays hadronically, it is disregarded.

The first step in this analysis was the generation of the events, performed using Monte Carlo (MC) generation. Event generation is a complex endeavour, as thousands of particles are produced in each event. In MC event generation, the proton-proton collision is factorised into a series of processes, until the particles reach the detector, as illustrated in Figure 4.1. Briefly, this starts with the partons, the proton constituents, described through their distribution function (PDF). This function is based on perturbative theory and can be written as $f(x, Q^2)$, where x is the fraction of proton momentum carried by the parton and Q^2 is the factorization scale that determines the resolution at which the proton can be probed [51]. This is set to correspond to the energy transferred between protons in the collision [52]. One parton can branch into others spontaneously, which will result in the production of other partons, to guarantee momentum and flavour are conserved. These particles, which will not intervene in the collision, form the initial-state radiation, as they contribute to the final state, but are not related to the process being studied.

In the collision, two partons interact (hard scattering) and can produce high-energy particles. This process and the PDFs refer to distinct energy scales, separated by the factorization scale, which means these two processes can be treated separately. This results in a total cross section after hard scattering described by Eq. 4.1 [51, 52]. The first two integrals correspond to the probability of having a combination of partons q and $\overline{q}'$ in the event and the sum accounts for all possible quark flavours. The last integral corresponds to the probability of having the two partons interacting and producing the desired final state X, with $\mathrm{d}\Omega = \mathrm{d}\cos\theta \mathrm{d}\phi$ corresponding to the solid angle.

$$\sigma_{q\overline{q}' \to X} = \sum_{q,q'} \int \mathrm{d}x_1 \int \mathrm{d}x_2 f_q(x_1, Q^2) f_{q'}(x_2, Q^2) \int \mathrm{d}\Omega \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\Omega} \tag{4.1}$$

In WH associated production, hard scattering results in the production of the W and H bosons. If quarks and gluons are produced in the final state, they can again branch into other particles sequentially, forming a parton shower. These can again radiate other partons, akin to a Bremsstrahlung process, forming final-state radiation [51, 52]. In the process studied, a parton shower is included, since the Higgs boson decays into two b-quarks. Due to the effect of confinement, once their energy is low enough, partons join together, and form colourless hadrons (hadronisation) [3], which, if unstable, will decay. In the process studied, the b-quarks hadronize into B-hadrons, which then decay into stable particles.

The framework used to perform calculations for the Lagrangian in question and to produce the events is given by SMEFTSIM3 [54], which only considers leading order (LO) contributions. This package uses the SMEFT formalism detailed in Section 2.3, where the operators can be chosen, as well as the value of their coefficient. The specific model, U35, guarantees maximum flavour symmetry, $U(3)^5$, only allowing operators whose contributions do not distinguish between the quark flavours. As such, the flavour contribution of these operators can be removed, as it will be the same whichever particles are involved [54]. This greatly reduces the number of parameters of the model and is quite useful in an analysis such as this, where flavour is not relevant. This model is implemented using FEYNRULES [55] package, so it can link with the next steps. The calculation of the matrix elements and generation of the hard-scatter process is done with MADGRAPH_AMC@NLO [56]. This software uses the NNPDF23LO set [57] to de-

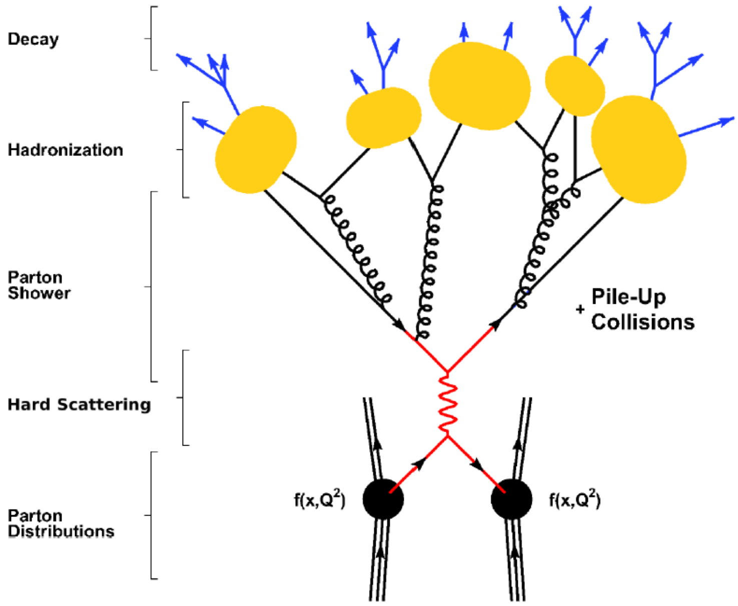


Figure 4.1: Illustration of event generation in a collision experiment [53]

fine the PDF, and it also models the parton level interactions, calculating their cross sections. After this, PYTHIA8 [58] is run, to simulate the parton shower and hadronisation of the particles and the unstable hadron decay is modeled by EVTGEN [59], giving the truth-level final state of the events.
The complete event generation also takes into account the simulation of the interactions of particles with the ATLAS detector, which is done using the Geant4 software [60].

4.2 Event Selection

The next step consists of selecting the events corresponding to the relevant signal, and efficiently rejecting background events. This requires implementing selection cuts on leptons, jets, and neutrinos to identify the signal.
In the final state considered, there is only one charged lepton - an electron or a muon. These events are filtered using the trigger system of the LHC. If the W boson decays into an electron, the event is selected by the single electron trigger. If it decays into a muon of p_T^W<150 GeV, it is similarly selected by the single muon trigger, but if the muon momentum is above this value, the E_T^{miss} trigger is used. The E_T^{miss} trigger is only efficient at high energies, due to its limited resolution, and uses information only from the calorimeters [61]. As $W \to \mu\nu_\mu$ results in a final state with a muon, which deposits little energy in the calorimeter, and a neutrino, which cannot be detected directly, the E_T^{miss} trigger labels the event as having a large missing transverse energy. As such, the E_T^{miss} trigger is efficient in selecting high-momentum muons [61].
When studying VH associated production, events are usually categorised depending on the V boson transverse momentum: $p_T^V < 400$ GeV corresponds to the resolved regime, and $p_T^V > 400$ GeV is called

the boosted regime. While in the resolved regime the two b-quarks can be resolved with two small-R jets [45], in the boosted regime, for a high-momentum Higgs boson, the b-quarks are emitted close together and have to be reconstructed from a single large-R jet. Because of this, jet event selection criteria are different depending on the regime.

In the resolved regime, an event must have two b-tagged jets to be considered in the signal region, with an invariant mass (m_{jj}) of at least 50 GeV, and with the jet of highest momentum fulfilling $p_T^{J_1}>45$ GeV [39]. Besides this, the angular distance between the two jets, ΔR_{jj}, defined in Eq. 5.1, must be lower than π. In the boosted regime, the large-R jet with the highest transverse momentum is considered, requiring that it has $p_T^j>250$ GeV. Within this large-R jet, smaller radius jets are built from track-only information (track-jets) and are considered for b-tagging. Akin to the resolved regime, the two leading track-jets must be b-tagged and the mass of the large-R jet (m_J) must be at least 50 GeV. In both regimes, the jets must also meet tight cleaning criteria, consisting of evaluating several variables that measure jet quality to effectively reject background signatures and select signal jets [62].

Additionally, when the final state includes an electron, the missing transverse energy of an event is required to be at least 30 GeV, in the resolved regime, or 50 GeV, in the boosted one. All the event selection cuts can be seen in Table 4.1.

After selecting these events, a series of boosted decision trees (BDT) are employed, based on the

Cuts	**Resolved Regime**	**Boosted Regime**
Number of leptons	1 electron or muon	
$\mathbf{p_T^W}$	< 400 GeV	> 400 GeV
Electron channel trigger	single electron trigger	
Muon channel trigger	$p_T^W<150$ GeV : single muon trigger $p_T^W>150$ GeV : E_T^{miss} trigger	E_T^{miss} trigger
Number of jets	≥ 2 small-R jets	≥ 1 large-R jet
b**-jets**	2 b-tagged small-R jets	2 b-tagged track-jets in the large-R jet
Jet invariant mass	$m_{jj} = 50$ GeV	$m_J = 50$ GeV
Leading jet $\mathbf{p_T^j}$	45 GeV	250 GeV
$\mathbf{\Delta R_{jj}}$	$< \pi$	-
$\mathbf{E_T^{miss}}$ **in electron channel**	30 GeV	50 GeV

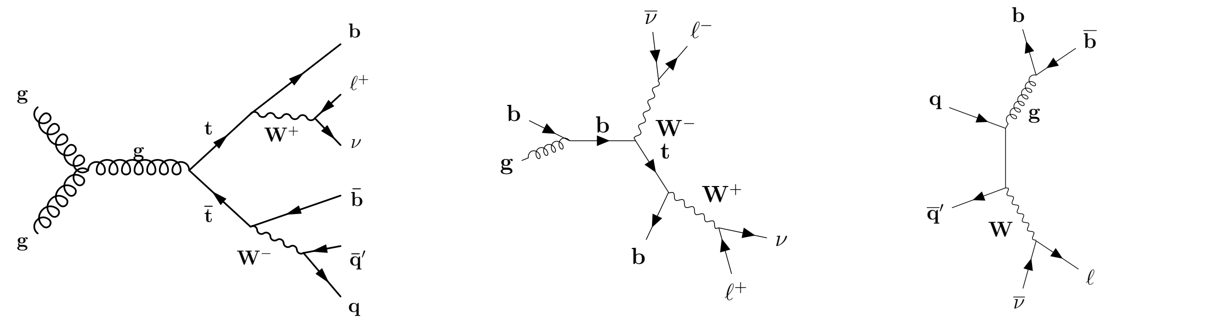
Table 4.1: Event selection cuts of the legacy analysis in the resolved and boosted regimes [39]

various observables in the analysis. The BDT outputs are used to find the yields of signal and background processes to use when fitting the data, and allow for a better differentiation between signal and background [39].

4.3 Background Events

Several processes produce similar final states to the signal event: $b\bar{b}\ell\nu_\ell$, which hinders the signal event identification. In this section, the most important of these background events are mentioned and shown in Figure 4.2.

One of the most significant backgrounds is $t\bar{t}$ production, where each top quark decays into a W boson and a b quark (which occurs for 96% of top quarks [17]), with one of the W bosons decaying leptonically



(a) (b) (c)

Figure 4.2: Feynman diagrams for the most important backgrounds in WH associated production with $WH \rightarrow b\bar{b}\ell\nu$: $t\bar{t}$ production (a), Wt production (b) and production of a W boson and b-jets (c).

and the other one hadronically (Figure 4.2a). Events involving the production of a top quark and a W boson make up another background, considering again that the top quark decays into a W boson and a b quark, with one W boson decaying hadronically and the other leptonically (Figure 4.2b).

The other main background process affecting this analysis is the production of a W boson along with jets. In the resolved regime, after event selection, about 85-95% of jets in this process are b-jets [39], in the region where signal is most likely to be found (signal region). This makes this process an irreducible background, when the W boson decays leptonically, as it has a final state identical to the signal events (Figure 4.2c).

Additionally, the associated production of a W and a Z boson (also called diboson) contributes to the set of backgrounds, since in the signal region after event selection, about 90% of these events correspond to a Z boson decaying into a $b\bar{b}$ quark pair [39], producing a final state identical to the signal events, when the W boson decays leptonically. The invariant mass of the b-jets in this process shows a peak at the mass of the Z boson that originates them (about 90 GeV), so it is useful for validation purposes. This behaviour can be seen in Figure 4.3, which shows the distribution of the mass of the two b-jets for signal and background, where a peak for the diboson background is evident around the mass of the Z boson.

A multi-jet background arises from hadronic jets being wrongly identified as leptons, and from semi-leptonic decays of b-hadrons. Multijet events, however, do not have a source of missing transverse energy, and the selection cuts described in the previous section are effective in suppressing this background in muon final states. In the electron channel, however, it may reach 2% of total backgrounds in electron final states [45]. This is, therefore, a reducible background.

4.3.1 Control Regions

The cuts presented in the previous section define the signal region of the analysis, but in addition to this, two control regions (CR) are also defined to contain a reduced number of signal events, but where the backgrounds still exist in similar number as in the signal region. This is done to better predict the background signatures, independently of the signal event so it can be more efficiently identified. The signal and control regions are defined so there is no overlap between them. In the legacy analysis, con-

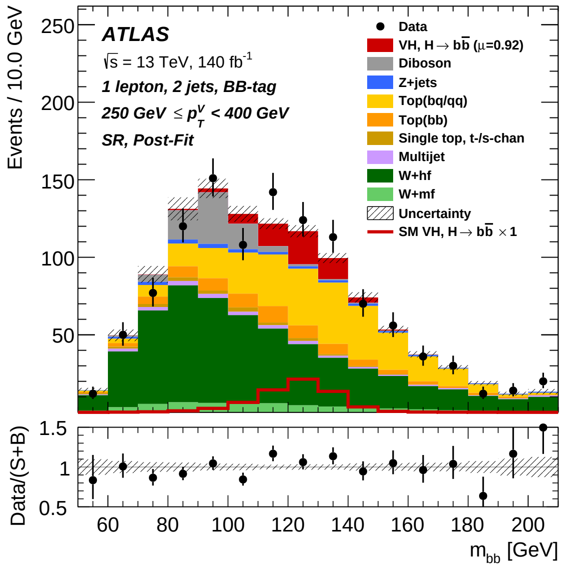


Figure 4.3: Distribution of the mass of two b-tagged jets for a one-lepton final state, considering signal and background events in the 250 GeV $< p_T^W < 450$ GeV [39]

trol regions are defined differently for the resolved and boosted regimes, targeting the most important background processes.

In the resolved regime, two control regions are defined, which differ from the signal region based solely on the angular distance between the two b-tagged jets, ΔR_{jj}. The High-ΔR region is made up of events with high values of ΔR_{jj}, above what is found for the signal region. This is used to study the $t\bar{t}$ production background. Similarly, the other CR, Low-ΔR, is characterised by events with ΔR_{jj} below what is seen in the signal region. This is a region with a higher number of events from the production of a W boson and jets than the other regions, making it suitable to make predictions on this background [39].

As for the boosted regime, only one CR region is defined, where events are required to have a b-tagged jet not belonging to the large-R jet, while keeping the remaining event selection criteria of the signal region. This requirement ensures this region includes a large number of $t\bar{t}$ production events, as is evident from the diagram of Figure 4.2a, where the decay products of the hadronically-decaying top-quark can be included in a large-R jet, while the other b-jet (from the semileptonic top decay) is produced far from this jet [39]. As the production of a W boson and jets is characterised by small angular distances between the final state objects, this background is not constrained in the boosted regime, where this is also a feature of the signal jets.

4.4 Simplified Template Cross Section measurements

A common framework used when studying cross sections in the Higgs sector is the simplified template cross sections (STXS) [1]. It provides a guideline of the regions in which to divide data when measuring couplings of the Higgs boson, based on truth-level quantities, to maximise the sensitivity to relevant information. The bin division is chosen so that, within each bin, there is only a small fluctuation

in signal acceptance, which corresponds to the fraction of events that pass the cuts applied. A small fluctuation of this quantity corresponds to a reduced dependence on theoretical predictions, which is an objective of the STXS framework [1].
For $q\bar{q}' \rightarrow WH$, data is divided into regions of p_T^W, and, within these, separated depending on the number of jets in the event, to further increase the sensitivity, and to separate events with different experimental uncertainties [14]. This division follows what is seen in Figure 4.4.

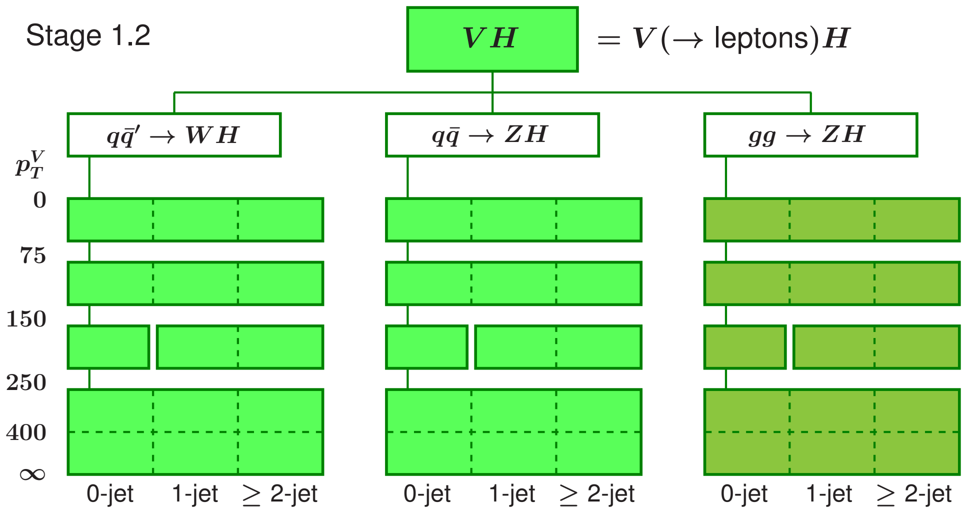


Figure 4.4: STXS bins for VH associated production [63]

4.4.1 Legacy Analysis Procedure

After performing the event selection in the ATLAS VH legacy analysis, the signal and background events are used to perform several likelihood fits, but only one is relevant to this thesis. Although the fit procedure is described in detail in Section 5.5, it is be briefly mentioned here. The events are then divided into the STXS bins and the number of signal and background events in each bin are input into the likelihood function. A fit of this data is then performed, as a function of the parameter of interest, which is, in this case, the cross section times branching ratio for the processes studied (WH or ZH associated production, with the Higgs boson decaying into b or c quarks). The goal is then to maximise the likelihood function to find the best-fit value of the signal cross section, for each of the STXS bins.

4.4.2 Systematic Uncertainties

In the legacy analysis, measurements are affected by systematic uncertainties that can originate from shortcomings in the theoretical model, or in the object reconstruction by the ATLAS detector.
The systematic uncertainties associated with simulating the signal events include acceptance uncertainties, which arise from the possibility of event migration between regions, and therefore depend on the event yields in the signal and control regions considered in the analysis. Besides this, systematic

uncertainties can arise from changes in the shape of the p_T^W distribution within a bin of this variable. Shape uncertainties are also included to account for electroweak and QCD scale corrections of the signal process [39].
As for background processes, their modeling also leads to systematic uncertainties that include normalisation uncertainties, which are independent of the binning, but affect each background process; as well as acceptance and shape uncertainties, which are analogous to what was described for signal events [39]. These uncertainties are calculated separately for each background.
As for experimental systematic uncertainties, the most relevant ones are related to jets. They include uncertainties due to the limited resolution and calibration of jet energy, as well as uncertainties from the flavour-tagging procedure [39].

4.4.3 Legacy Analysis Results

The legacy analysis used events from Run 2 of the LHC, corresponding to a luminosity of 140 fb^{-1} and a center-of-mass energy of the proton beams of $\sqrt{s} = 13$ TeV. It found that the WH cross section measurements presented large relative uncertainties, above 100% for some bins of W boson transverse momentum, and that there was a smaller number of events available than for ZH. This meant that the further division of events depending on the number of jets was only performed for the ZH analysis. The resulting cross section (times branching ratio) measurements can be seen in Figure 4.5, as well as the ratio to the SM values. All the results are consistent with the SM predictions.

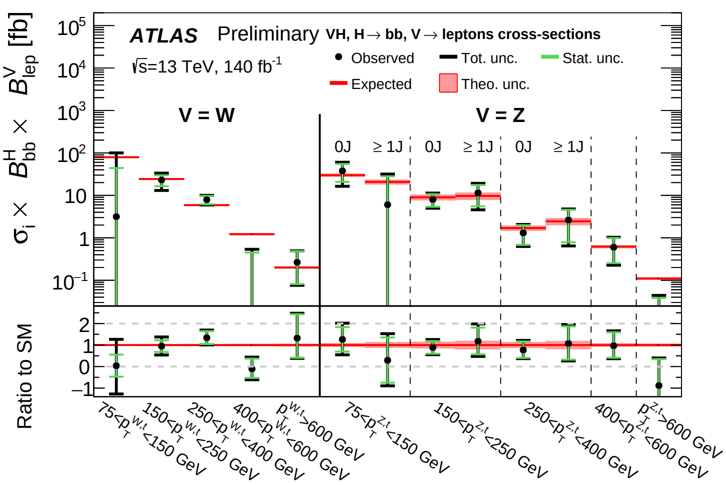


Figure 4.5: VH associated production cross section times the $H \to b\bar{b}$ and V leptonic decay branching ratios as a function of p_T^W and the number of jets. The lower panel shows the ratio of this quantity to the SM prediction. [39]

4.5 Probing CP-violating couplings in the HWW vertex

The search for CP-odd HWW interactions is based on the VH legacy analysis, with some changes designed to obtain the best sensitivity to these couplings. In this section, a series of angular observables that have proven to be sensitive to the HWW interaction are defined, along with the necessary changes to the STXS binning to include them in a study of this interaction.

4.5.1 Angular Observable Study

In the work of this thesis, to perform an analysis of CP-odd couplings in WH associated production, one needs suitable observables that are sensitive to the linear contribution of $c_{H\tilde{W}}$. A relevant set of angular variables to search for anomalous couplings is introduced in [21], which describes a CPV study for WH associated production with $WH \to \ell\nu b\bar{b}$. The variables are defined in Eq. 4.2, 4.3, and 4.4. In these equations, $\mathbf{p}_Y^{(X)}$ corresponds to the three-momentum of particle Y in the rest frame of X, and $\mathbf{p}_Y$ corresponds to the momentum of Y in the lab frame.

$$\cos\theta^* = \frac{\mathbf{p}_\ell^{(W)} \cdot \mathbf{p}_W}{|\mathbf{p}_\ell^{(W)}||\mathbf{p}_W|} \tag{4.2}$$

$$\cos\delta^+ = \frac{\mathbf{p}_\ell^{(W)} \cdot (\mathbf{p}_H \times \mathbf{p}_W)}{|\mathbf{p}_\ell^{(W)}||(\mathbf{p}_H \times \mathbf{p}_W)|} \tag{4.3}$$

$$\cos\delta^- = \frac{(\mathbf{p}_{\ell_1}^{(H-)} \times \mathbf{p}_{\ell_2}^{(H-)}) \cdot \mathbf{p}_W}{|(\mathbf{p}_{\ell_1}^{(H-)} \times \mathbf{p}_{\ell_2}^{(H-)})||\mathbf{p}_W|} \tag{4.4}$$

The angular variable $\cos\theta^*$ corresponds to the angle between the charged lepton produced in the W boson decay (in the W boson rest frame) and the trajectory of the W boson, giving information on the polarisation of this particle. This observable proved to have different distributions for the SM prediction and some BSM contributions. It could not, however, distinguish the CP-odd contribution from the CP-even ones.

To probe possible CP violation, other variables, such as $\cos\delta^+$ and $\cos\delta^-$, are built. For the latter, $\mathbf{p}_\ell^{(H-)}$ corresponds to the momentum of a lepton in the rest frame of a Higgs boson with symmetric momentum. These angular variables depend on the CP-odd linear term of the interaction. Their distributions show an asymmetry in the presence of the CP-odd coupling, dependent on the magnitude of the Wilson coefficient. This does not happen for the SM or CP-even operators. $\cos\delta^-$ seems to have a greater asymmetry than $\cos\delta^+$, but proved more difficult to differentiate between the SM and CP-even contributions [21]. Since this thesis also considers the CP-even operators, $\cos\delta^+$, in particular, proves to be a variable well-suited for the work.

As noted in [64], when considering the CP-odd operator, events with positive W bosons produce a $\cos\delta^+$ distribution symmetric to what is found for negative W bosons. This results in a reduced asymmetry, as $\cos\delta^+$ takes positive and negative values alternately, thus having a shape close to what is found for

CP-even or SM contributions. To improve on this, instead of considering $\cos\delta^{+}$, its value is multiplied by the charge of the lepton (and therefore of the W boson) involved: $Q_{\ell}\cos\delta^{+}$, to increase the desired asymmetry.

4.5.2 STXS-like bins

As $Q_{\ell}\cos\delta^{+}$ has proved to be suitable to study the contribution of the CP-odd operator $c_{H\tilde{W}}$, it is relevant to include it when establishing the bins to study in this work. As such, an STXS-like framework is used, keeping the p_T^W bins, but subdividing the data further into two $Q_{\ell}\cos\delta^{+}$ bins, as is sketched in Figure 4.6. To calculate $Q_{\ell}\cos\delta^{+}$, however, the full momentum of the W boson is needed, which requires the reconstruction of the longitudinal component of the neutrino momentum, as described in Section 3.3.3. The limited resolution of this reconstruction leads to a limited resolution in $Q_{\ell}\cos\delta^{+}$ as well. To avoid resolution effects that result in large bin migrations, only two bins are chosen for this observable: $Q_{\ell}\cos\delta^{+}<0$ and $Q_{\ell}\cos\delta^{+}>0$.

Furthermore, as the focus of this thesis is on constraining a BSM operator, instead of focusing on the

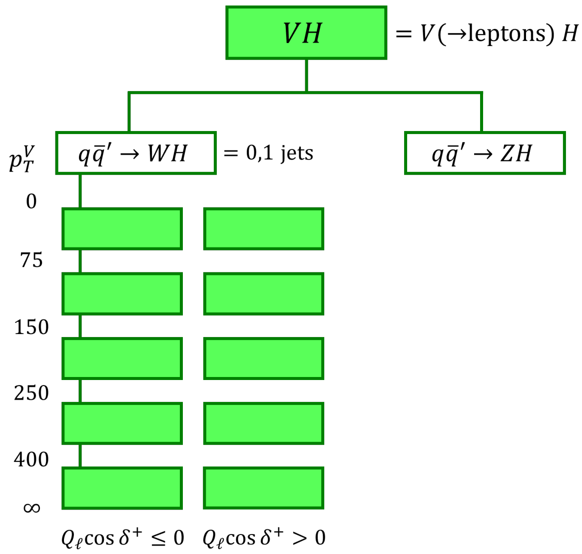


Figure 4.6: STXS-like bins used in the analysis

cross sections measurements (as the ATLAS legacy analysis does), it is useful to perform a likelihood fit using instead a parameterisation of the ratio of the BSM WH associated production cross section to the SM prediction. This procedure is described in detail in the following chapter.

In summary, the ATLAS $WH \rightarrow \ell\nu bb$ analysis is the basis for the methodology employed in this work.The project this thesis is a part of has, however, the novelty of exploring and adding an angular variable ($Q_{\ell}\cos\delta^{+}$) to the analysis, allowing, for the first time, to extract a measurement of the contribution of the $c_{H\tilde{W}}$ operator in the HWW vertex.

Chapter 5

Methodology and Results

This chapter details the methodology used in this thesis and the results achieved. The goal of this work was to produce a parameterisation of cross sections of WH associated production events, with $HW \to b\bar{b}\ell\nu$ making up the final state, in terms of different BSM contributions and with a special focus on the $c_{H\tilde{W}}$ operator. For this purpose, only signal, truth-level events were considered. To achieve this, the work was divided into five stages: the event generation for the SM and BSM hypothesis; the implementation of the event selection and analysis; the study and interpretation of several observables potentially sensitive to BSM effects; and the parameterisation of cross sections, dividing the events into bins of the transverse momentum of the W boson and of an angular variable. The parameterisation was then used as input for a fit to constrain the BSM contributions.

5.1 Event Generation

The parameterisation of the cross sections is done using truth-level events, without including the simulation of the final state particles interacting with the detector. As such, there is no need to include a software to provide particle reconstruction, as would happen in the detector. Only signal events are considered in this thesis, so $q\overline{q}' \to HW \to b\bar{b}\ell\nu$ is the only process for which events were generated. The backbone of this procedure was the EFTTOOLS repository [65], which employs a number of software packages to arrive at the final state particles.

The event generation is performed for a center-of-mass energy of $\sqrt{s} = 13$ TeV. The final state is specified requiring a $\nu_\ell \ell^+$ or a $\overline{\nu_\ell}\ell^-$ pair in the final state (for $\ell = e$ or μ), corresponding to the decay of a W^+ or a W^- boson, respectively; and a Higgs boson decaying into a $b\bar{b}$ pair using PYTHIA8, following the SM predictions. The charged leptons are required to have a minimum transverse momentum of 1 GeV and the transverse momenta of jets and b hadrons in the events have a minimum value of 20 GeV. These are usual cuts performed for the jets, so they can be calibrated and included in the event final state. The events are generated independently for the SM, and for the linear and quadratic terms of each operator.

5.1.1 Operator Choice

A study of the operators affecting the HWW vertex (seen in Table 2.2) was performed, to understand their influence on the cross section. This was done for the linear and quadratic terms (second and third terms in Eq. 2.11, respectively). The study was also done separately for $W^+H \to \nu_\ell \ell^+ H$ and $W^-H \to \overline{\nu_\ell}\ell^- H$, using MADGRAPH, since the two final states could only be generated separately. The resulting cross sections, found in Tables 5.1 and 5.2, respectively, for 10×10^3 events in each, demonstrate that the operators $c_{Hq}^{(3)}$, c_{HW} and $c_{H\tilde{W}}$ have the largest impact on the cross section. The remaining operators do not affect the kinematics of the process, only its cross section [66], and are therefore disregarded in what follows. In the case of $c_{H\tilde{W}}$, it is clear that the linear term of this operator (the CP-odd one) does not cause any change in the cross section of the processes. The coupling has to be studied, therefore, by looking at the changes in kinematics it can produce. Besides this, only the values for the CP-odd linear coupling show a difference between the two tables, which is motivated by statistical fluctuations, evident by the large uncertainties associated with each value. There is, however, an overlap between the uncertainty intervals of the two couplings.

Table 5.1: Relative cross section for each operator acting on the HWW vertex for a $\ell^+\nu_\ell H$ final state. All coefficients are set at $c = 1$, with $\Lambda = 1$ TeV and the SM cross section was found to be $\sigma_{SM} = 0.1257 \pm 0.0002$ pb.

Wilson coefficient	σ/σ_{SM}	
	Linear Term	Quadratic Term
$c_{Hq}^{(3)}$	1.976 ± 0.009	2.96 ± 0.02
c_{HW}	0.866 ± 0.004	0.360 ± 0.002
$c_{H\tilde{W}}$	$(3.7 \pm 13.5) \times 10^{-4}$	0.230 ± 0.001
$c_{H\Box}$	0.1161 ± 0.0005	$(3.67 \pm 0.01) \times 10^{-3}$
c_{HDD}	-0.02902 ± 0.00002	$(2.300 \pm 0.009) \times 10^{-4}$

Table 5.2: Relative cross section for each operator acting on the HWW vertex for a $\ell^-\overline{\nu}_\ell H$ final state. All coefficients are set at $c = 1$, with $\Lambda = 1$ TeV and the SM cross section was found to be $\sigma_{SM} = 0.0676 \pm 0.0002$ pb.

Wilson coefficient	σ/σ_{SM}	
	Linear Term	Quadratic Term
$c_{Hq}^{(3)}$	1.93 ± 0.01	2.52 ± 0.02
c_{HW}	0.830 ± 0.005	0.348 ± 0.002
$c_{H\tilde{W}}$	$(-1.8 \pm 1.3) \times 10^{-3}$	0.228 ± 0.001
$c_{H\Box}$	0.1154 ± 0.0006	$(3.68 \pm 0.02) \times 10^{-3}$
c_{HDD}	-0.02885 ± 0.00001	$(2.30 \pm 0.02) \times 10^{-4}$

After this study was performed, $c_{Hq}^{(3)}$, c_{HW} and $c_{H\tilde{W}}$ are considered to generate events at both the linear order (order 1) and the quadratic (order 2). For each of the operators, two symmetric values of

c were used, to study the differences between them, as can be seen in Table 5.3. The influence of c_{HW} and $c_{H\tilde{W}}$ on the cross section is much smaller than for $c_{Hq}^{(3)}$, which means that when comparing distributions proportionally to this quantity, it is helpful to use a larger value of c_{HW} and $c_{H\tilde{W}}$. Even though $c_{HW} = 1.0$ and $c_{H\tilde{W}} = \pm 1.0$ are excluded at 95 % CL based on the latest measurements [33], this choice helps in better seeing differences in the plots. This will not be relevant in later steps, since the parameterisation produced is independent of the values of the Wilson coefficients. To compare the CP-odd operator with its CP-even counterpart, $c_{H\tilde{W}}$ is studied for the same values as c_{HW}.

Table 5.3: Operator and their values used in the analysis.

Operator	Values
$c_{Hq}^{(3)}$	± 0.1
c_{HW}	± 1.0
$c_{H\tilde{W}}$	± 1.0

13 different sets of events were generated: one order 0 (SM), and four for each of the BSM operators (two values of c chosen, for order 1 and order 2 contributions). This was done by producing a total of 500×10^3 events, for each case.

5.2 Analysis of Simulated Data

The distributions of four observables were studied in the first part of this thesis. In this section, the variables and their definitions are presented, followed by a description of the analysis, along with its results.

5.2.1 Choice of Observables

Angular distance

The angular distance between two objects, i and j, ΔR_{ij}, is one of the observables considered. It is given by Eq. 5.1, where η is the pseudorapidity and ϕ is the azimuthal angle.

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{5.1}$$

The angular distance considered here is the distance between two b-jets in the event, resulting from the Higgs boson decay. This variable is sensitive to the boost of the Higgs boson, since the higher its momentum is, the closer together the b quarks from its decay are produced. This in turn leads to smaller values of ΔR. This is the variable that allows for the separation of signal and control regions.

W boson transverse momentum

The transverse momentum of the W boson, p_T^W, is also taken into account. This is the projection of the momentum in the plane perpendicular to the beam axis and can be calculated from the transverse momenta of the lepton and neutrino, following Eq. 5.2. If there is no other source of E_T^{miss}, as there is not in this case, and because the neutrino mass can be neglected, the transverse momentum of the neutrino will correspond to E_T^{miss}.

$$p_T^W = p_T^\ell + p_T^{\nu_\ell} = p_T^\ell + E_T^{\text{miss}} \tag{5.2}$$

$\cos\delta^+$

This angular variable follows Eq. 5.3, and its choice is motivated in Chapter 4.

$$\cos\delta^+ = \frac{\mathbf{p}_\ell^{(W)} \cdot (\mathbf{p}_H \times \mathbf{p}_W)}{|\mathbf{p}_\ell^{(W)}||(\mathbf{p}_H \times \mathbf{p}_W)|} \tag{5.3}$$

$\cos\theta^*$

This variable follows 5.4, with its choice again being motivated in Chapter 4.

$$\cos\theta^* = \frac{\mathbf{p}_\ell^{(W)} \cdot \mathbf{p}_W}{|\mathbf{p}_\ell^{(W)}||\mathbf{p}_W|} \tag{5.4}$$

5.2.2 Event selection and analysis

The analysis of the events was done using the RIVET3 analysis toolkit [67]. The routine was produced for this work, and can be found in the VHBB_CP repository [68], under NEWSPECTRAANALYSIS_WH.CC. It follows the same structure as the routines of EFTTOOLS, but is updated to work with a more recent version of Rivet, 3.1.6. The goal of this routine is to produce the distributions of the four variables mentioned above, for the different operators. In this section, the routine will be described.

This work focuses on the resolved regime, as the number of events in the boosted regime is small, which means that, firstly, events with two b-tagged jets are selected. It is also checked if the leptons originate from a W boson. In this routine, events are generated for W^+ and W^- simultaneously, in the same proportion of as was found in Tables 5.1 and 5.2. After these two conditions are met, the values of the four variables mentioned above are calculated using kinematic information of the event particles. Moreover, this analysis is performed by setting $\Lambda = 1$ TeV.

Jets and the H boson

In this routine only small-R jets are taken into account and a cut is applied to require the momentum of the jets is at least 20 GeV.

The only jets of relevance are those that include the two b-quarks from the Higgs boson decay. As such, there must be two jets in the event considered that are b-jets. This analysis is concerned with truth-level

data, so b-jet identification is done by ghost-association of B-hadrons to jets, instead of through experimental signatures like secondary vertices. In ghost-association, the B-hadron momentum is reduced to a negligible value and it is then added to the stable final state particle list, allowing the anti-k_t algorithm to choose to which jet to cluster each B-hadron [69].
After finding the two leading b-jets, which are considered to correspond to the Higgs boson decay, the angular distance between them is calculated, and the momentum of the Higgs boson is found by summing the momenta of the two jets. The Higgs boson momentum is also used to calculate the values of the angular observables.

Leptons and the W boson

Leptonically decaying W bosons can be reconstructed automatically in RIVET, using WFINDER. The momentum of the boson is found using the information from the leading charged lepton of the event, along with the missing four-momentum of the final state, which will correspond to that of the neutrino. The lepton is also required to have a pseudorapidity $|\eta| < 2.5$, and the transverse momentum larger than 15 GeV, over the minimum threshold usually considered for electron and muon selection in the ATLAS detector.
The momentum of the reconstructed W boson is used not only to produce a distribution of its transverse component but also to calculate the value of the angular variables for the event. Furthermore, the momentum of the lepton is also used to calculate the angular variables, after it is boosted to the W boson rest frame.

5.2.3 Observable Distributions

The routine described in the previous section was run for a sample of 50×10^3 events, for the linear and quadratic terms of the operators with the values of Table 5.3, as well as for the SM term. The results were then merged so that the distributions produced could combine the contributions from the order 0, 1, and 2.
Firstly, it is relevant to present the cross sections and sum of event weights of each anomalous contribution, as these are the two quantities that govern the histogram scaling, and give some information about the most influential contributions in the analysis. Each event has an associated weight, which is attributed to account for its corresponding differential cross section [70]. The event weights for the SM and the different BSM contributions are presented in Table 5.4.
The linear term corresponds to the interference between the contribution of an operator and the SM, which can be constructive or destructive. This means that there can be events with positive or negative weights. As was seen before in Tables 5.1 and 5.2, the linear term of the $c_{H\tilde{W}}$ contribution does not affect the cross section, which can also be seen in this table, including from the fact that the event weights sum to a value close to zero, orders of magnitude lower than for the other operators. As for the CP-even operators, there is no difference in the magnitude of the sum of weights for the two values studied, which

Table 5.4: Cross sections and sum of weights for 50×10^3 generated events, for a $\sigma_{\text{SM}} = 0.3929$ pb, with weights summing to 1.965×10^4. Here, σ_{int} corresponds to the linear contribution to the cross section, and σ_{BSM} corresponds to the quadratic term.

Operator	**c Value**	σ_{int} **(pb)**	**Sum of Weights Order 1**	σ_{BSM} **(pb)**	**Sum of Weights Order 2**
$\mathbf{c_{H\tilde{W}}}$	1.0	1.065×10^{-4}	5.755	0.08853	4.427×10^3
	-1.0	-5.487×10^{-4}	-27.58	0.08848	4.424×10^3
$\mathbf{c_{HW}}$	1.0	0.3480	1.740×10^4	0.1383	6.916×10^3
	-1.0	-0.3478	-1.740×10^4	0.1383	6.916×10^3
$\mathbf{c_{Hq}^{(3)}}$	0.1	0.07566	3.782×10^3	9.690×10^{-3}	484.2
	-0.1	-0.07574	-3.787×10^3	9.684×10^{-3}	484.2

suggests the weights have consistent values in both cases, although with opposite signs.

By looking at the sum of weights for order 1 in the table, it is evident that when considering negative values of the operators, there is a greater prevalence of negative event weights. This means that when the histograms for the three orders of an operator are merged, the negative coefficients will produce distributions with lower yields, when compared to their positive counterparts. This does not happen, however, for the quadratic term, as it is the same for the two symmetric values of each operator. For order 2, $c_{H\tilde{W}}$ and c_{HW} present similar event weights, but an order of magnitude above what is found for $c_{Hq}^{(3)}$. This is due to the fact that the former are studied for a value that is also one order of magnitude above the value chosen for $c_{Hq}^{(3)}$.

To further probe the relevance of negative weights in each contribution, Table 5.5 shows the number of events with negative weights for each coupling. As is expected, the number of events with negative weights for the SM and order 2 contributions is negligible. As for order 1, the CP-even operators with $c>0$ prove to include a low number of events with negative weights. When comparing this to the corresponding $c<0$ case, however, it is clear that the weight distributions are symmetric: the number of events with positive weights when $c>0$ is consistent with the number of events with negative weights for $c<0$. This, therefore, explains why the sum of weights in Table 5.4 is symmetric for order 1 CP-even contributions. As for the CP-odd operator, however, there is an equal number of events with positive and negative weights for both cases, $c = 1.0$ and $c = -1.0$, which again explains why the sum of weights shown in Table 5.4 is reduced for these contributions, as the weights cancel themselves out in the two cases. It is also of note that when dividing events onto the STXS bins, the fraction of negative events remained constant for increasing of p_T^W, so the influence of these event weights is the same for all the regions considered.

Figure 5.1 shows the distribution of ΔR between the two b-tagged jets, comparing the SM prediction to the three operators. It is noticeable that all plots have similar shapes, and that for $\Delta R \gtrsim 4.5$, not all bins are populated. It is clear that the yield for $c_{HW} = 1.0$ is larger than for all other cases (and smaller for $c_{HW} = -1.0$), but there is no marked difference in the distributions of the different operators (which show overlapping ratios to the SM), except for c_{HW}. This variable is important when establishing the

Table 5.5: Number of negative weights in 50×10^3 generated events of each operator value. The SM file contains 17 events with negative weights.

Operator	c Value	Order 1 Events	Order 2 Events
$\mathbf{c_{H\tilde{W}}}$	1.0	25008	31
	-1.0	25022	25
$\mathbf{c_{HW}}$	1.0	1628	24
	-1.0	48387	26
$\mathbf{c^{(3)}_{Hq}}$	0.1	31	94
	-0.1	49980	122

BDT and control regions mentioned above, but will not be used in the parameterisation that follows, as it does not clearly differentiate the CP-odd contribution from the CP-even ones.

When analyzing the distribution of $\cos\theta^*$ in Figure 5.2, looking at the ratio of BSM to SM contribution

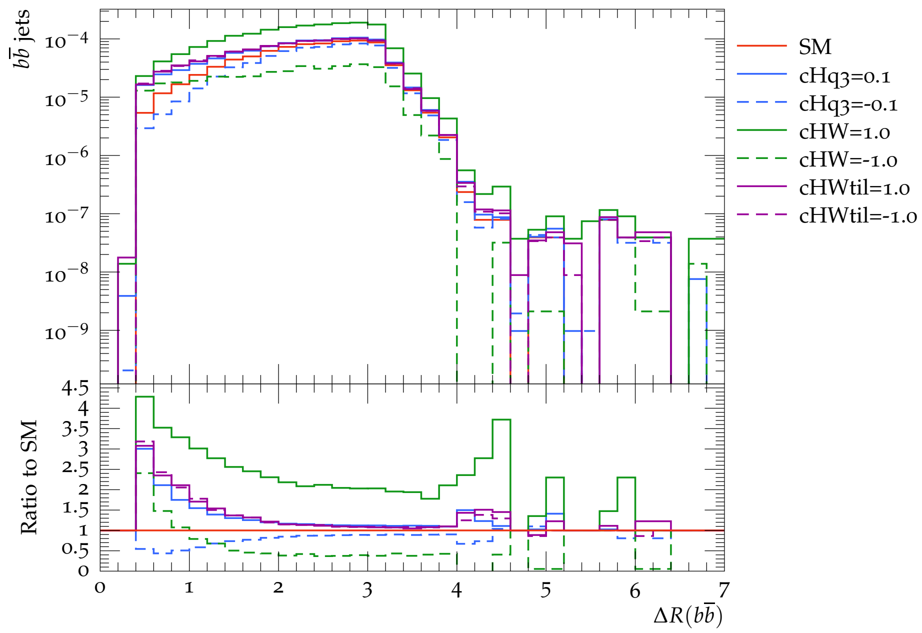


Figure 5.1: Distributions of ΔR, for the SM prediction and the sum of the SM, linear and quadratic contributions of three operators. A positive value (solid) and its symmetric (dashed), are represented. The ratio to the SM distribution is at the bottom. All plots are normalised to the cross-section.

provides easier-to-read information. This plot shows that, for $c^{(3)}_{Hq}$, the ratio to the SM appears nearly constant for all values of the observable, indicating the shape of the distribution is similar to the SM. There is, however, an asymmetry in the distribution ratios of both c_{HW} and $c_{H\tilde{W}}$. This variable can therefore provide a distinction between the SM and c_{HW} or $c_{H\tilde{W}}$, but, as $c_{H\tilde{W}} = \pm 1.0$ have overlapping distributions, it cannot differentiate between different contributions of this operator. As such, it is not suitable to study CP-odd couplings individually.

The study of p_T^W is shown separately for each of the operators in Figure 5.3. It is noticeable that most operators lead to an increase in the number of events with high-momentum W bosons, which suggests BSM contributions are likely to appear in this momentum region. This stems from the fact that BSM effects increase with the energy of a collision, so they will be more prevalent in high-momentum regions. Once again looking at the ratio of a contribution compared to the SM provides added information, with

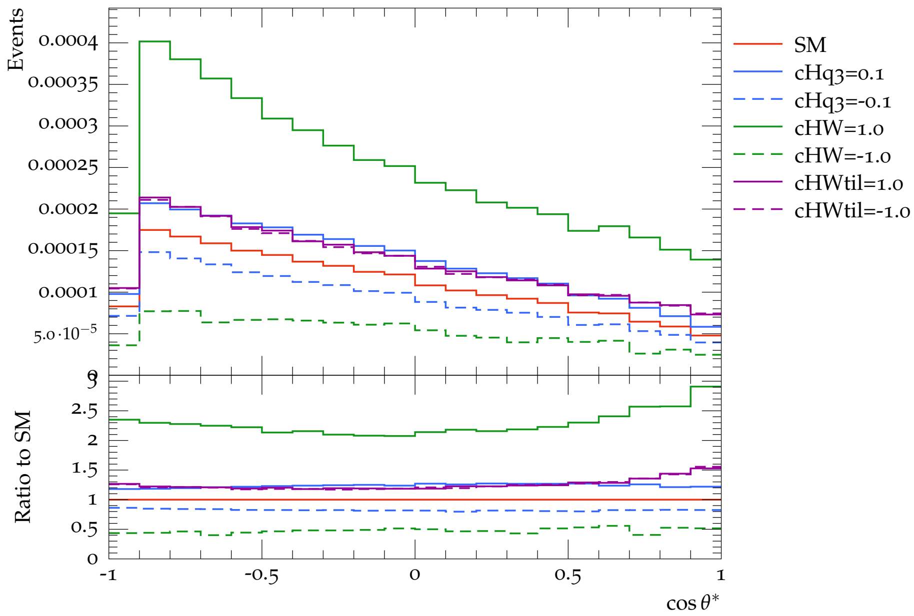


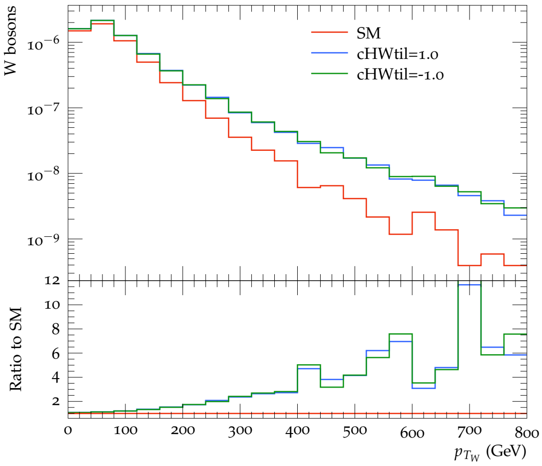
Figure 5.2: Distributions of $\cos\theta^*$, for the SM prediction and the sum of the SM, linear, and quadratic contributions of three operators. A positive value (solid) and its symmetric (dashed), are represented. The ratio to the SM distribution is at the bottom. All plots are normalised to the cross-section.

$c_{H\tilde{W}}$ showing the same behaviour independently of its value, which suggests that this variable is not sensitive to the linear term of this operator. The magnitude of BSM effects is larger in the quadratic contribution, so the increase in the number of high p_T^W events will be driven by this order 2 (and not order 1) terms. It is clear from these plots that p_T^W is a variable that can separate between SM and BSM couplings, and, following the standard of ATLAS analyses [1], it will be included in the following steps.

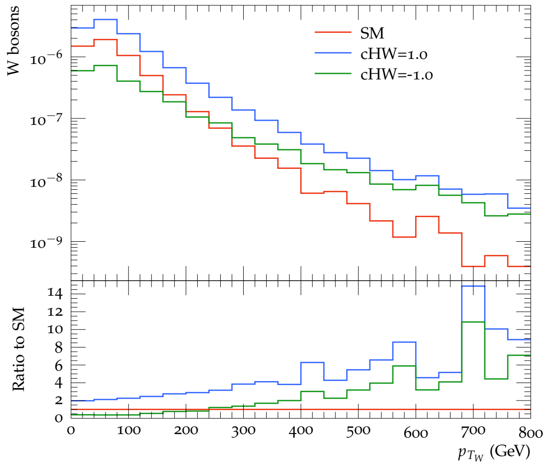
The distributions of $Q_\ell \cos\delta^+$ for the three operators considered can be seen in Figure 5.4. Both of the CP-even operators produce distributions similar in shape to the SM, but the CP-odd operator (Figure 5.4a), however, leads to an asymmetric distribution, distinct from all other cases. Moreover, changing $c_{H\tilde{W}}$ also produces alterations in the shape of $Q_\ell \cos\delta^+$, as the symmetric values of the Wilson coefficient produce symmetric distributions of this variable. This shows that this observable is mostly governed by the linear term of the operator, responsible for the CP-odd coupling, since the quadratic term is the same for both values of $c_{H\tilde{W}}$ [21]. $Q_\ell \cos\delta^+$ is thus suitable to be used in the next stages of this work.

5.3 Sorting events into the STXS-like bins

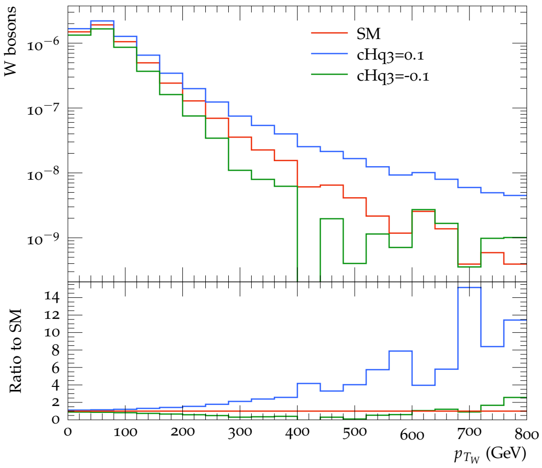
To obtain the distribution of the events into the STXS-like bins described in Section 4.5.2, a RIVET routine was developed, which can be found in STXSANALYSIS_WH.CC in the VHBB_CP repository [68]. It is used to create a set of histograms that are filled according to the number of events fitting each bin's criteria. Three sets of histograms are produced: a division of the events into the five p_T^W bins; a separation into the two $Q_\ell \cos\delta^+$ regions; and a combination of both into a 2D binning. This portion of



(a)



(b)



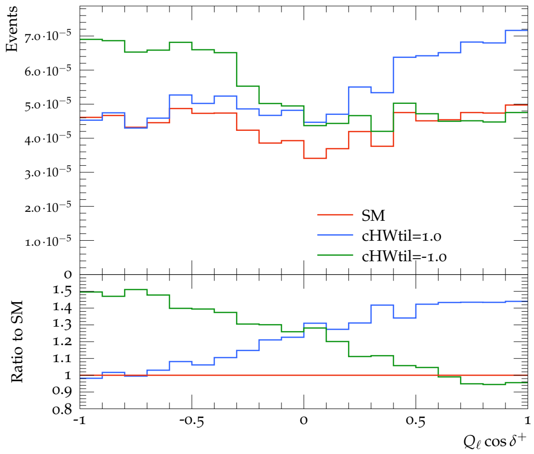
(c)

Figure 5.3: Distributions of p_T^W for the SM prediction and the sum of the SM, linear and quadratic contributions of operators $c_{H\tilde{W}}$ (a), c_{HW} (b) and $c_{Hq}^{(3)}$ (c). A positive value and its symmetric is represented in each and the ratio to the SM distribution is at the bottom of each figure. All plots are normalised to the cross-section.

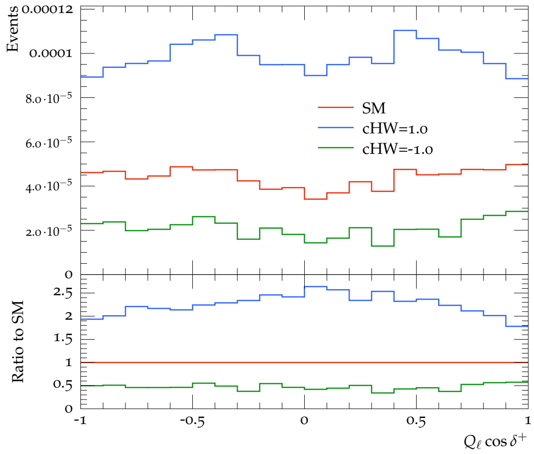
the work used a set of 500×10^3 events for each sample and, again, $\Lambda = 1.0$ TeV.

The basis of this routine is the same as what was described in 5.2.2. The main difference in this process is the way the Higgs is reconstructed. Instead of summing the jets momenta to get $\mathbf{p}_H$, the HEPMC event record [71] is accessed to recover the truth Higgs information, since the STXS framework requires the use of truth-level quantities. A cut is then applied, requiring that the Higgs boson is produced with $|\eta| < 2.5$, as this is the acceptance of ATLAS's inner detector, but no cut is applied on the leptons, as is recommended by the STXS framework [1]. The MC-truth momenta of the Higgs boson, the W boson and the charged lepton are accessed directly to calculate $Q_\ell \cos\delta^+$.

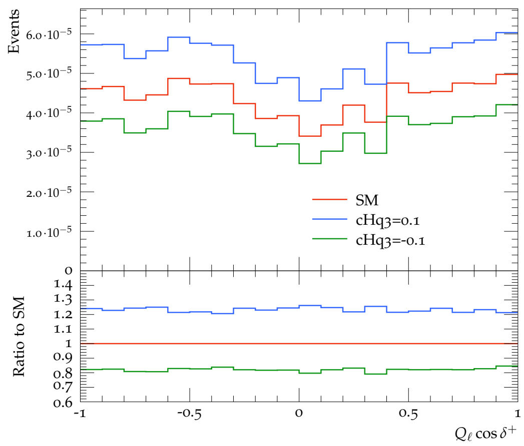
To understand the statistical significance of the results in each bin, the number of events in each of the p_T^W STXS bins was calculated, for the events in the SM sample. The number of events decreases with the increase of transverse momentum, and while in the first bin ($p_T^W < 75$ GeV) there were about 100×10^3



(a)



(b)



(c)

Figure 5.4: Distributions of $Q_\ell \cos\delta^+$ for the SM prediction and the sum of the SM, linear and quadratic contributions of operators $c_{H\tilde{W}}$ (a), c_{HW} (b) and $c^{(3)}_{Hq}$ (c). A positive value and its symmetric is represented in each and the ratio to the SM distribution is at the bottom of each figure. All plots are normalised to the cross-section.

events, the last bin ($p_T^W > 400$ GeV) had only about 1% of this number of events. The breakdown of how many events populate each bin can be found in Table 5.6. This shows that there will be a much higher statistical uncertainty in the results from the last bins.

Table 5.6: Number of events in each p_T^W STXS bin, for the SM signal sample

$\mathbf{p_T^W}$ **(GeV)**	[0, 75]	[75, 150]	[150, 250]	[250,400]	[400, ∞]
Number of events	117581	77217	24365	6112	1350

5.4 Cross section parameterisation

After the RIVET routine described has divided the events into the STXS-like bins, they are plugged into a parameterisation routine. This process follows what is described in [23], adding to the CP-even

operators used in the paper, the CP-odd one this work is concerned with.
The cross section of any process calculated within the EFT framework follows Eq. 2.11. This equation can be rewritten for an operator, i, and for each STXS region in consideration, p, as shown in the left side of Eq. 5.5. It is also useful to represent this equation in terms of the relative cross section, i.e., the ratio to the SM cross section, as shown on the right side of Eq. 5.5. Here, $A_i^{\sigma_p}$ and $B_i^{\sigma_p}$ are the parameterisation coefficients for the linear and quadratic operator terms, respectively. They are used to describe the dependence of the relative cross section on the Wilson coefficient.

$$\sigma_{p,\text{SMEFT}} = \sigma_{p,\text{SM}} + \frac{c_i}{\Lambda^2}\sigma_{i,\ p,\text{int}} + \frac{c_i^2}{\Lambda^4}\sigma_{i,\ p,\text{BSM}} \Leftrightarrow \frac{\sigma_{p,\text{SMEFT}}}{\sigma_{p,\text{SM}}} = 1 + A_i^{\sigma_p} c_i + B_i^{\sigma_p} c_i^2 \tag{5.5}$$

To implement the parameterisation, the histograms produced in the RIVET routine described in Section 5.3 are input into the parameterisation routine of the EFTTOOLS repository [65]. The parameterisation coefficients are found by solving a least squares problem for each term, $y = A_i^{\sigma_p} c_i$ for order 1 and $y = B_i^{\sigma_p} c_i^2$ for order 2, where y is the relative cross-section with respect to the SM. This procedure also includes the use of 200 bootstrap samples, the default value in the routine, to calculate the standard deviation of the parameterisation parameters for each bin. This is done by taking 200 random samples from the totality of events in a bin (with replacement) and finding the parameterisation coefficients for each one. The distribution of the parameterisation coefficients from each sample can then be built, with the standard deviation of this distribution being used as the value for uncertainty of the parameterisation coefficients for totality of events in a bin.

5.4.1 Parameterisation Results

The parameterisation coefficients A_i and B_i, along with their uncertainties are obtained by the aforementioned procedure, using 500×10^3 events for each sample. The tables presenting the parameterisation results for the $c_{H\tilde{W}}$ operator can be seen in this section (Tables 5.7-5.12), as this is the operator used in the subsequent steps of the analysis. The parameterisation results for the two other operators, c_{HW} and $c_{Hq}^{(3)}$ can be found in Appendix A.
To better visualise the A_i and B_i coefficients, however, a set of six plots is presented, with the linear and quadratic coefficients for a parameterisation of $Q_\ell \cos\delta^+$, p_T^W, and the combination of the two variables, considering the three operators. As the parameterisation of Eq. 5.5 describes, the values of A_i and B_i are independent of c. The parameterisation was, however, performed for the positive and negative values of each coefficient to guarantee this was fulfilled, even in the bins with high p_T^W where there is less statistics. The plots presented below use only the results for the positive values of Wilson coefficients, but are equivalent to what was found from the negative ones (as can be confirmed in the tables).

$Q_\ell \cos\delta^+$ parameterisation

The $Q_\ell \cos\delta^+$ parameterisation results for the CP-odd operator can be seen in Tables 5.7 and 5.8, while the CP-even operator coefficients are shown in Appendix A.1. Two plots are also presented, showing the parameterisation coefficients for the three operators in the same figure, for comparison. This can be seen in Figures 5.5 (considering order 1) and 5.6 (considering order 2).

Table 5.7: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{H\tilde{W}} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1, 0]$	-0.156 ± 0.001	0.280 ± 0.001
$[0, 1]$	0.157 ± 0.001	0.279 ± 0.001

Table 5.8: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{H\tilde{W}} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1, 0]$	-0.158 ± 0.001	0.279 ± 0.001
$[0, 1]$	0.159 ± 0.001	0.279 ± 0.001

As suggested by distributions of $Q_\ell \cos\delta^+$ (Figure 5.4), the parameterisation coefficients for CP-even contributions are the same for the two bins of this variable, within their uncertainties. As for the CP-odd operator (see Tables 5.7 and 5.8), its asymmetric distribution due to the linear order term leads to the two bins having A_i of opposite signs, but keeping the same magnitude. When considering order 2, however, there is no difference in B_i between the two bins (within the uncertainty intervals). This is to be expected, since, for the CP-odd operator, $Q_\ell \cos\delta^+$ is only sensitive to its linear contribution.

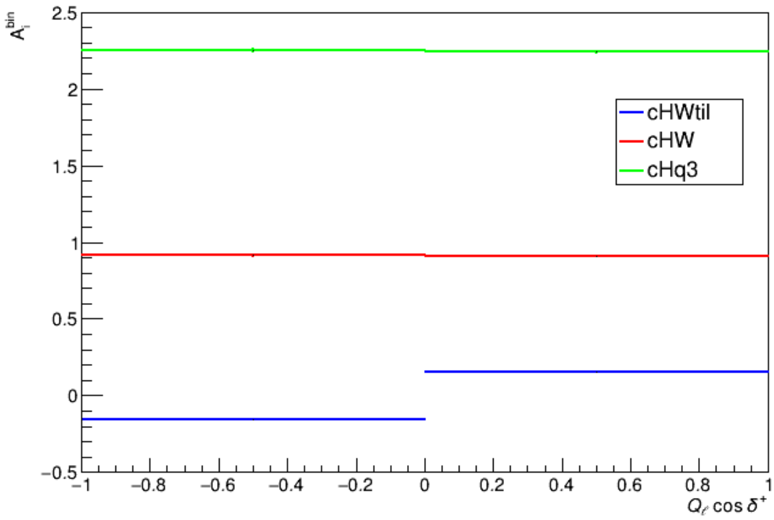


Figure 5.5: Linear parameterisation coefficient for each bin of $Q_\ell \cos\delta^+$, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$. The uncertainties of each bin are smaller than the width of the line.

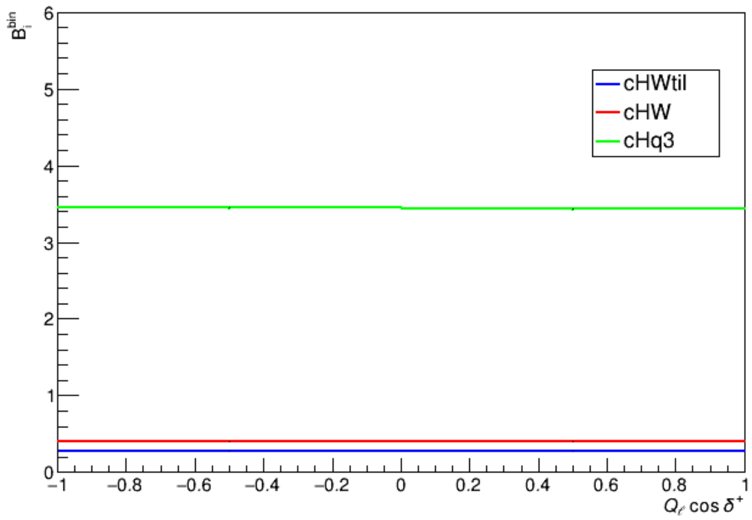


Figure 5.6: Quadratic parameterisation coefficient for each bin of $Q_\ell \cos\delta^+$, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$. The uncertainties of each bin are smaller than the width of the line.

Parameterisation on the transverse momentum of the W boson

As for the parameterisation of p_T^W, its results can be seen in Tables 5.9 and 5.10 for $c_{H\tilde{W}}$, and in Appendix A.2 for c_{HW} and $c_{Hq}^{(3)}$. Figures 5.7 and 5.8 show the plots with these results for A_i and B_i, respectively, for the three operators.

Table 5.9: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{H\tilde{W}} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	0.0000 ± 0.0009	0.1194 ± 0.0006
$[75, 150]$	-0.001 ± 0.002	0.257 ± 0.001
$[150, 250]$	0.002 ± 0.004	0.621 ± 0.005
$[250, 400]$	0.000 ± 0.009	1.45 ± 0.02
$[400, \infty]$	0.01 ± 0.02	4.0 ± 0.1

Table 5.10: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{H\tilde{W}} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	0.000 ± 0.001	0.1198 ± 0.0006
$[75, 150]$	0.000 ± 0.003	0.257 ± 0.001
$[150, 250]$	-0.003 ± 0.003	0.623 ± 0.004
$[250, 400]$	0.013 ± 0.009	1.46 ± 0.02
$[400, \infty]$	0.027 ± 0.03	3.9 ± 0.1

The linear order parameterisation coefficient for $c_{H\tilde{W}}$ in each bin is consistent with a null coefficient (see Tables 5.9 and 5.10). This is to be expected from the results of the previous sections which showed the order 1 term of this operator did not influence the cross section of the signal process. For c_{HW}, the linear term does modify the relative cross section, with the coefficient for each bin being within 10% of the value of the next bin. This means the interference term of this operator adds a nearly constant contribution to the WH production relative cross section. As for $c_{Hq}^{(3)}$, it is already clear just from order 1 that the higher the momentum of the W boson, the more a bin impacts the cross section value. As to the order 2 parameterisation, it is clear that for all operators, their contribution to the total WH production cross section increases with p_T^W. This argues in favor of using events in the boosted regime, as they are more sensitive to the BSM Wilson coefficients.

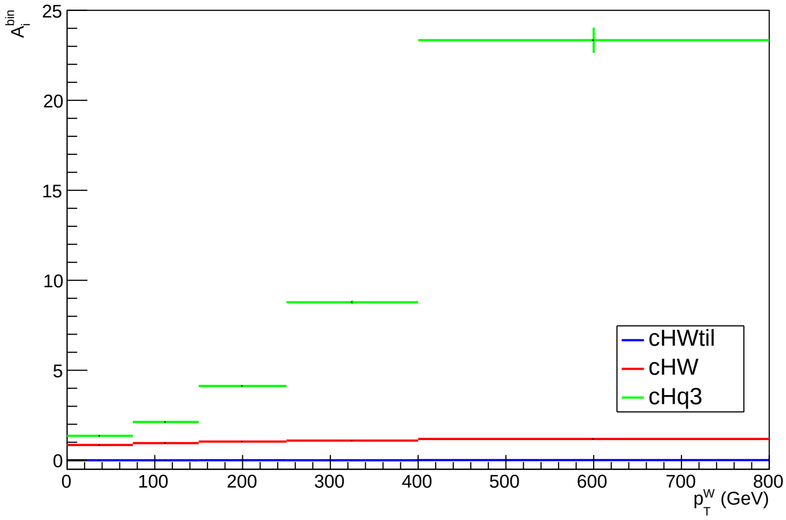


Figure 5.7: Linear parameterisation coefficient for each bin of p_T^W, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$.

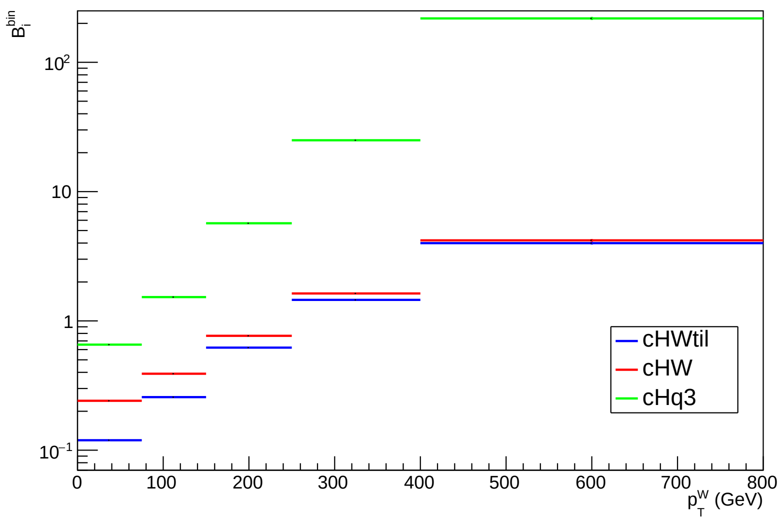


Figure 5.8: Quadratic parameterisation coefficient for each bin of p_T^W, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$. The uncertainties of each bin are smaller than the width of the line.

Two-dimensional parameterisation

The final parameterisation, from a simultaneous analysis of both $Q_\ell \cos\delta^+$ and p_T^W bins, can be found in Tables 5.11 and 5.12 for the $c_{H\tilde{W}}$ operator, and in Appendix A.3 for c_{HW} and $c_{Hq}^{(3)}$. Two plots combine the results for the positive values of the three operators, which can be seen in Figures 5.9 and 5.10 for orders 1 and 2, respectively.

Table 5.11: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{H\tilde{W}} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W **Bin**	$Q_\ell \cos\delta^+$ **Bin**	**Order 1**	**Order 2**
$[0, 75]$	$[-1, 0]$	-0.056 ± 0.001	0.1196 ± 0.0009
	$[0, 1]$	0.057 ± 0.001	0.1192 ± 0.0008
$[75, 150]$	$[-1, 0]$	-0.174 ± 0.002	0.257 ± 0.002
	$[0, 1]$	0.171 ± 0.002	0.257 ± 0.002
$[150, 250]$	$[-1, 0]$	-0.378 ± 0.006	0.615 ± 0.006
	$[0, 1]$	0.394 ± 0.007	0.628 ± 0.007
$[250, 400]$	$[-1, 0]$	-0.73 ± 0.02	1.50 ± 0.03
	$[0, 1]$	0.69 ± 0.02	1.41 ± 0.03
$[400, \infty]$	$[-1, 0]$	-1.16 ± 0.05	3.7 ± 0.1
	$[0, 1]$	1.36 ± 0.07	4.3 ± 0.2

Table 5.12: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{H\tilde{W}} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W **Bin**	$Q_\ell \cos\delta^+$ **Bin**	**Order 1**	**Order 2**
$[0, 75]$	$[-1, 0]$	-0.056 ± 0.001	0.1200 ± 0.0008
	$[0, 1]$	0.056 ± 0.001	0.1196 ± 0.0008
$[75, 150]$	$[-1, 0]$	-0.175 ± 0.002	0.257 ± 0.002
	$[0, 1]$	0.174 ± 0.002	0.257 ± 0.002
$[150, 250]$	$[-1, 0]$	-0.393 ± 0.006	0.617 ± 0.007
	$[0, 1]$	0.398 ± 0.006	0.630 ± 0.007
$[250, 400]$	$[-1, 0]$	-0.72 ± 0.02	1.50 ± 0.03
	$[0, 1]$	0.72 ± 0.02	1.43 ± 0.03
$[400, \infty]$	$[-1, 0]$	-1.23 ± 0.05	3.6 ± 0.1
	$[0, 1]$	1.48 ± 0.07	4.2 ± 0.2

As the events are divided into bins of p_T^W and $Q_\ell \cos\delta^+$, it provides a parameterisation most sensitive to the operator this work is concerned with. Both for order 1 and 2, the parameterisation coefficients are consistent within the uncertainty intervals between the two $Q_\ell \cos\delta^+$ bins with the same p_T^W, for the bins with p_T^W<150 GeV, which is to be expected from the parameterisation of $Q_\ell \cos\delta^+$. This is not true in some bins with higher transverse momenta (p_t^W>150 GeV), but here the results still agree within 2σ. The uncertainties quoted in the tables, in all bins, do not exceed 5%, which guarantees that the parameterisation will be accurate up to $c \lesssim \sqrt{\frac{1}{0.05}} \simeq 4.5$ [72], which is one order of magnitude above what is considered in EFT and above the experimental limits reported in the most recent studies mentioned in Chapter 2, since [33] found that $c_{H\tilde{W}} \in [-0.31, 0.88]$ at a 95% CL.

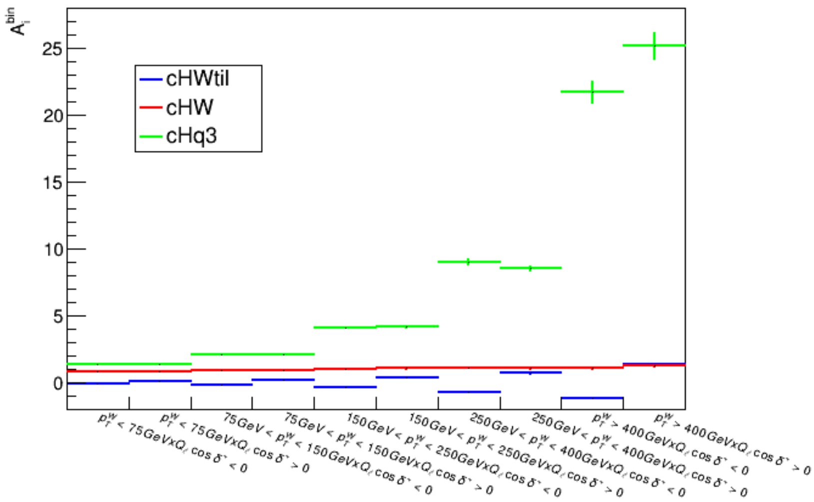


Figure 5.9: Linear parameterisation coefficient for each bin of $p_T^W \times Q_\ell \cos\delta^+$, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$.

5.5 Likelihood fit

The parameterisation results are used in a maximum-likelihood fit of signal and background data (now including object reconstruction) into the STXS bins, as was mentioned in the previous chapter. Although this portion of the analysis was developed outside of the scope of this thesis, a likelihood scan as a function of the value of $c_{H\tilde{W}}$ was performed, whose results are presented in this section. The likelihood fit is also briefly explained here, as the last step to reach a constraint on the operator values.

The inputs to binned maximum-likelihood fits are the BDT output distributions described in Chapter 4. Their bin contents are described by a Poisson distribution, as each bin is simply populated by the number of events that fulfill the bin's requirement. As such, the expected number of events in bin i with n_i total entries is given by $E[n_i] = \boldsymbol{\mu} s_i + b_i$, where s_i is the mean expected count of signal events (corresponding to the process described in this thesis) in that bin, while b_i is the equivalent for background events [73] (the backgrounds included are mentioned in Chapter 4). $\boldsymbol{\mu}$ is the signal strength [26], given by Eq. 5.6. This corresponds, therefore, to the cross section parameterisation described in the previous section.

$$\mu_i = \frac{\sigma_{i,\text{SMEFT}}}{\sigma_{i,\text{SM}}} \tag{5.6}$$

This leads to a likelihood following Eq. 5.7. Two other contributions are added, the first of which corresponds to the inclusion of nuisance parameters in the model, $\boldsymbol{\theta}$. These are quantities that change the distributions of the background and signal, but that the analysis is not concerned with. They include systematic uncertainties, from the detector, or the model itself (as mentioned in Section 4.4.2), that would affect both s_i and b_i. The last term models the limitation of having a reduced number of simulated events to study which will lead to a statistical uncertainty, especially for the more numerous background events. As such this term is added to account for this uncertainty, with γ_i corresponding to a scaling parameter for each i bin and $\beta_i = 1/\sigma_{\text{rel}}^2$, with σ_{rel} being the statistical uncertainty in a bin relative to the

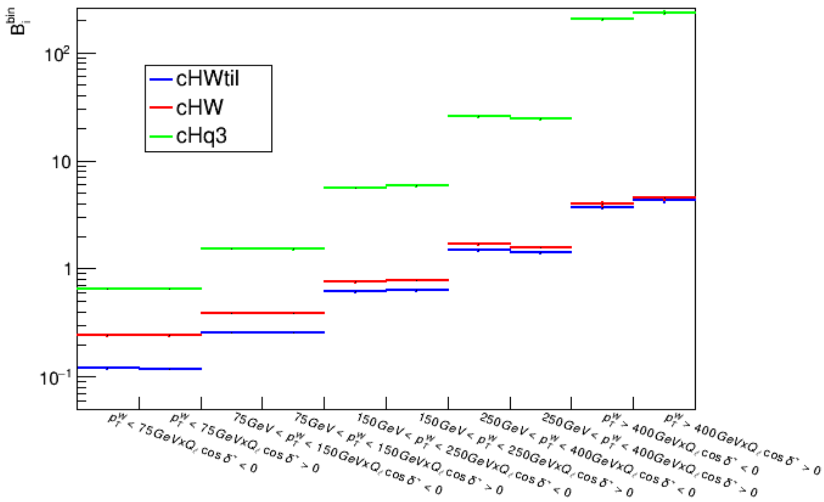


Figure 5.10: Quadatic parameterisation coefficient for each bin of $p_T^W \times Q_\ell \cos\delta^+$, for $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$.

total background sample. Lastly, $\boldsymbol{\tau}$ corresponds to background normalisation factors.

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\tau}) = \prod_{i \in bins} \text{Poisson}(n_i | \boldsymbol{\mu} s_i(\boldsymbol{\mu}) + b_i(\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\tau})) \prod_{\theta \in \boldsymbol{\theta}} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} \prod_{i \in bins} \text{Gauss}(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i}) \quad (5.7)$$

A fit is then performed to maximise the likelihood. After this is done, it is useful to calculate the log-likelihood ratio, q_μ (Eq. 5.8) [73]. The numerator denotes the maximum likelihood fixing a specific value of μ, with $\hat{\hat{\theta}}$ and $\hat{\hat{\gamma}}$ giving the parameters maximise the likelihood in this condition. The denominator is the maximum unconditional likelihood, where no parameters are restricted. The log-likelihood ratio then gives a measure of how good the agreement between a model with a certain μ, and the data is.

$$q_\mu = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}, \hat{\hat{\gamma}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}}, \hat{\gamma})} \quad (5.8)$$

It was shown that when considering only one parameter, q_μ could also be written in the form of Eq. 5.9 [73], where N is the sample size.

$$q_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \quad (5.9)$$

If the dataset considered is large so that the second term in Eq.5.9 can be disregarded, one can assume q_μ follows a non-central χ^2 distribution. Here is where it is useful to use an Asimov dataset, in which the value of each parameter matches its expectation value under the SM hypothesis and which allows for the use of a smaller data sample representative of the larger dataset, thus reducing the computational power required, while allowing for the non-central χ^2 distribution of q_μ [73].
One important metric when performing a maximum-likelihood fit to data is the p-value (p_μ), which measures the probability of finding the observed results under the null hypothesis. In this work, the null hypothesis corresponds to the fit of SM predictions and here $p_\mu = 2(1 - \Phi(\sqrt{q_\mu}))$, where Φ is the cumulative distribution function for a Gaussian. In particle physics, however, the significance, Z, is usually

preferred to the p-value. This quantity follows Eq. 5.10, setting the limit of Z above which a hypothesis can be validated.

$$Z_\mu = \Phi^{-1}(1 - p_\mu) = \sqrt{q_\mu} \tag{5.10}$$

A log-likelihood profile can also be obtained in terms of a range of Wilson coefficient values. From this profile, the best-fit value of the Wilson coefficients can therefore be found, corresponding to the log-likelihood minimum.

For a two-sided χ^2 distribution where only one parameter is considered, the threshold of the distribution marking the 68% and 95% confidence levels (CL) corresponds to a log-likelihood ratio of 0.99 and 3.84, respectively. This means confidence intervals of the Wilson coefficients correspond to the values that produce a log-likelihood ratio of ± 0.49 or ± 1.92, respectively.

The likelihood is calculated using WH associated production signal events, and background events corresponding to the processes described in the previous chapter. The event selection and analysis regions correspond to what was used in the legacy analysis, while the events are divided into bins of p_T^W and $Q_\ell \cos\delta^+$.

The profile likelihood for $p_T^W \times Q_\ell \cos\delta^+$ bins, considering the linear order parameterisation of the CP-odd operator, $c_{H\tilde{W}}$, can be seen in Figure 5.11, where to each point is subtracted the log-likelihood of the best-fit value. This shows the expected sensitivity of the analysis, that can constrain the Wilson coefficient to [-0.32, 0.33] at a 68% CL, and [-0.65,0.65] at 95% CL. These limits are consistent with the results of the analysis in [33], which found the expected contribution of the operator to be within [-0.60,0.60] at a 95% CL for a linear parameterisation. The results presented in this section are, therefore, competitive with the current ones of other analyses, with more symmetric limits. Furthermore, this is the first analysis that is sensitive only to the HWW vertex, as other studies consider VBF, which cannot unambiguously distinguish between HWW and HZZ.

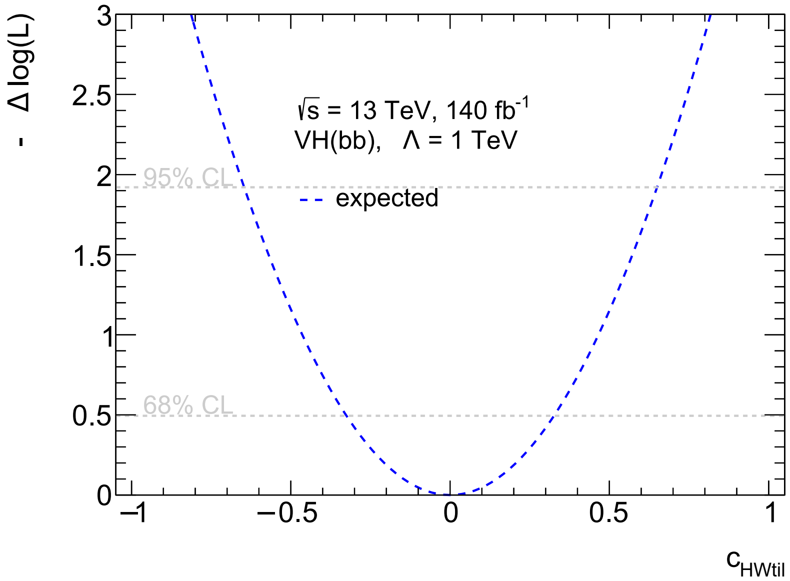


Figure 5.11: Profile likelihood including the linear parameterisation of $p_T^W \times Q_\ell \cos\delta^+$ as a function of the value of $c_{H\tilde{W}}$

Chapter 6

Conclusions

This thesis was developed as part of a study probing CP-violation in the HWW interaction. It used truth-level simulated data of LHC proton-proton collisions with $\sqrt{s} = 13$ TeV, integrated in the ATLAS experiment, and focusing on WH associated production with $WH \to \ell\nu b\bar{b}$. The basis of this analysis was the SMEFT framework, considering operators that modify the HWW vertex of the process. An initial study was performed to understand which operators influence the WH production cross section the most, resulting in three operators being considered: $c_{H\tilde{W}}$, c_{HW} and $c_{Hq}^{(3)}$.
The distributions of four variables sensitive to these operators were plotted, ΔR, p_T^W, $Q_\ell \cos\delta^+$, and $\cos\theta^*$. This was performed for different values of the Wilson coefficients considering both the linear and quadratic terms of an operator at a time.
From the study of their distributions, p_T^W, $Q_\ell \cos\delta^+$ and $\cos\theta^*$ are found to be sensitive to the BSM couplings. Although $\cos\theta^*$ showed an asymmetry for $c_{H\tilde{W}}$ and c_{HW}, being able to separate these contributions from the SM prediction, the behaviour of this variable was similar for these two operators, thus not being suitable in studies of the CP-odd coupling. $Q_\ell \cos\delta^+$, on the other hand, showed an asymmetric distribution when considering $c_{H\tilde{W}}$ only, thus being an observable sensitive mainly to this operator. The distribution of this variable was studied for $c_{H\tilde{W}} = \pm 1$, resulting in a behaviour of $Q_\ell \cos\delta^+$ for $c_{H\tilde{W}} = 1.0$ symmetric to what is found for $c_{H\tilde{W}} = -1.0$. This shows that this variable is sensitive to the linear term of the operator, the one responsible for the CP-odd contribution.
The STXS framework used in studies of the cross section of Higgs boson production processes was adapted to divide data into regions that would increase the sensitivity of the analysis to the CP-odd operator. In this work, the data was then divided into the STXS bins of p_T^W and two bins of $Q_\ell \cos\delta^+$: $Q_\ell \cos\delta^+ \leq 0$ and $Q_\ell \cos\delta^+ > 0$. The WH production cross section obtained from non-zero BSM Wilson coefficients (c_i) was parameterised in terms of $Q_\ell \cos\delta^+$, p_T^W and $p_T^W \times Q_\ell \cos\delta^+$, by solving a least squares problem: $\sigma_{\text{BSM}}/\sigma_{\text{SM}} = A_i c_i$, for the linear term, and $\sigma_{\text{BSM}}/\sigma_{\text{SM}} = B_i c_i^2$, for the quadratic one. Since A_i and B_i are independent of the value of c_i, the two values considered for each operator produced similar parameterisations. The linear coefficients had little sensitivity to the value of p_T^W, with A_i being constant for the whole range of momentum values, for both $c_{H\tilde{W}}$ and c_{HW}. It is the quadratic term that alters the momentum distribution, when compared to the SM, which can be seen in the increasing

values of B_i. As to the $Q_\ell \cos\delta^+$ linear parameterisation coefficients, for the CP-even operators there is no difference in the value of A_i in the two bins of this observable, in opposition to what is found for $c_{H\tilde{W}}$. For this operator, the two bins of $Q_\ell \cos\delta^+$ considered had symmetric linear parameterisation coefficients for the same value of the Wilson coefficient, thus showing the asymmetry found in the distribution of this variable for $c_{H\tilde{W}}$. With this parameterisation, the goals of this thesis were achieved.

This parameterisation was used as the signal strength in a likelihood function where WH associated production with $WH \to b\bar{b}\ell\nu$ is the signal process, and where events are divided into bins of p_T^W and $Q_\ell \cos\delta^+$. Although this was not developed in the scope of this thesis, a scan of the log-likelihood profile for a linear parameterisation was performed, finding the values of $c_{H\tilde{W}}$ to be in [-0.32,0.33] at a 68% CL and [-0.65,0.65] at a 95% CL, when considering the SM hypothesis.

The next steps in this analysis include adding the CP-even parameterisations to perform a likelihood fit, as these operators influence the distribution of p_T^W. Besides this, new machine learning methods can be employed to constrain the operator, such as simulation-based inference techniques [64]. This would improve the capability of this analysis of studying detector data.

Bibliography

[1] M. Delmastro, N. Berger, C. Bertella, M. Duehrssen-Debling, O. Kivernyk, J. M. Langford, P. Milenovic, C. E. Pandini, F. Tackmann, K. Tackmann, L. Viliani, M. Xiao, G. Marchiori, M. Calvetti, C. Li, T. P. Calvet, P. Francavilla, V. Dao, Y. Haddad, and E. J. T. Scott. Simplified Template Cross Sections – Stage 1.1. Technical Report LHCHXSWG-2019-003, DESY-19-070, CERN, Geneva, 2019.

[2] N. Wolchover, S. Velasco, and L. Reading-Ikkanda. A New Map of All the Particles and Forces. *Quanta Magazine*. URL `https://www.quantamagazine.org/a-new-map-of-the-standard-model-of-particle-physics-20201022/#example`.

[3] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.

[4] B. R. Martin and G. Shaw. *Particle physics*. Wiley, 3rd edition, 2008. ISBN 978-0-470-03294-7.

[5] A. D. Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967. doi: 10.1070/PU1991v034n05ABEH002497.

[6] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. doi: 10.1103/PhysRevLett.13.508.

[7] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi: 10.1103/PhysRevLett.13.321.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. doi: 10.1103/PhysRevLett.13.585.

[9] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(CERN-PH-EP-2012-218, CERN-PH-EP-2012-218):1–29, 2012. arXiv:1207.7214.

[10] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(CMS-HIG-12-028, CERN-PH-EP-2012-220, CMS-HIG-12-028, CERN-PH-EP-2012-220):30–61, 2012. arXiv:1207.7235.

[11] J. Riebesell and S. Bringuier. Collection of standalone TikZ images, 2020. 10.5281/zenodo.7486911 - https://github.com/janosh/tikz.

[12] The ATLAS Collaboration. Combined measurement of the higgs boson mass from the $h \to \gamma\gamma$ and $h \to zZ^* \to 4\ell$ decay channels with the atlas detector using $\sqrt{s} = 7$, 8, and 13 tev pp collision data. *Phys. Rev. Lett.*, 131:251802, Dec 2023. doi: 10.1103/PhysRevLett.131.251802.

[13] W. Buchmüller and C. Lüdeling. Field Theory and Standard Model. (hep-ph/0609174, DESY-06-151, DESY-2006-151), 2006. doi: 10.5170/CERN-2006-014.1.

[14] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. CERN, Geneva, 2017. doi: 10.23731/CYRM-2017-002.

[15] The ATLAS Collaboration. A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery. *Nature*, 607(7917):52–59, Jul 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04893-w.

[16] The CMS Collaboration. A portrait of the higgs boson by the cms experiment ten years after the discovery. *Nature*, 607(7917):60–68, Jul 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04892-x.

[17] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024. doi: 10.1103/PhysRevD.110.030001.

[18] D. H. Perkins. *Particle Astrophysics*. Oxford University Press, 1st edition, 2003. ISBN 978-0198509523.

[19] H. E. Haber and O. Stål. New LHC benchmarks for the CP-conserving two-Higgs-doublet model. *The European Physical Journal C*, 75(10):491, Oct 2015. ISSN 1434-6052. doi: 10.1140/epjc/s10052-015-3697-x.

[20] J. Rojo. The Standard Model Effective Theory: towards a pedagogical primer, 2020. SciPost Physics Lecture Notes.

[21] R. M. Godbole, D. J. Miller, K. A. Mohan, and C. D. White. Jet substructure and probes of CP violation in Vh production. *Journal of High Energy Physics*, 2015(4):103, Apr 2015. ISSN 1029-8479. doi: 10.1007/JHEP04(2015)103.

[22] B. Moser. The Beauty and the Boost: a Higgs Boson Tale - Measurements of Higgs Boson Production at High Energy in Decays to Bottom Quarks and Their Interpretations with the ATLAS Experiment at the LHC, 2021. Presented 04 Mar 2022.

[23] The ATLAS Collaboration. Methodology for EFT interpretation of Higgs boson Simplified Template Cross-section results in ATLAS. Technical Report ATL-PHYS-PUB-2019-042, CERN, Geneva, 2019.

[24] I. Brivio, Y. Jiang, and M. Trott. The SMEFTsim package, theory and tools. *Journal of High Energy Physics*, 2017(12):70, Dec 2017. ISSN 1029-8479. doi: 10.1007/JHEP12(2017)070.

[25] C. Degrande and J. Touchèque. A reduced basis for CP violation in SMEFT at colliders and its application to diboson production. *Journal of High Energy Physics*, 2022(4):32, Apr 2022. ISSN 1029-8479. doi: 10.1007/JHEP04(2022)032.

[26] G. Aad et al. Characterising the Higgs boson with ATLAS data from Run 2 of the LHC. (CERN-EP-2024-087), arXiv:2404.05498 [hep-ex].

[27] The ATLAS Collaboration. CP properties of higgs boson interactions with top quarks in the $t\bar{t}h$ and th processes using $h \to \gamma\gamma$ with the ATLAS detector. *Phys. Rev. Lett.*, 125:061802, Aug 2020. doi: 10.1103/PhysRevLett.125.061802.

[28] The ATLAS Collaboration. Probing the cp nature of the top–higgs yukawa coupling in tt¯h and th events with h→bb¯ decays using the atlas detector at the lhc. *Physics Letters B*, 849:138469, 2024. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2024.138469.

[29] The CMS Collaboration. Search for CP violation in t$\bar{t}$H and tH production in multilepton channels in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Journal of High Energy Physics*, 2023(7):92, Jul 2023. ISSN 1029-8479. doi: 10.1007/JHEP07(2023)092.

[30] The ATLAS Collaboration. Test of CP-invariance of the Higgs boson in vector-boson fusion production and in its decay into four leptons. *JHEP*, 2405(CERN-EP-2023-030):105, 2024. arXiv:2304.09612.

[31] The ATLAS Collaboration. Test of CP Invariance in Higgs Boson Vector-Boson-Fusion Production Using the $H \to \gamma\gamma$ Channel with the ATLAS Detector. *Phys. Rev. Lett.*, 131:061802, Aug 2023. doi: 10.1103/PhysRevLett.131.061802.

[32] The ATLAS Collaboration. Integrated and differential fiducial cross-section measurements for the vector boson fusion production of the Higgs boson in the $H \to WW^* \to e\nu\mu\nu$ decay channel at 13 TeV with the ATLAS detector. *Phys. Rev. D*, 108:072003, Oct 2023. doi: 10.1103/PhysRevD.108.072003.

[33] The ATLAS Collaboration. Differential cross-section measurements of Higgs boson production in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report CERN-EP-2024-198, CERN, Geneva, 2024.

[34] The CMS Collaboration. Constraints on standard model effective field theory effects with Higgs bosons produced in association with W or Z bosons in the H→b$\bar{\mathrm{b}}$ decay channel in proton-proton collisions at $\sqrt{s} = 13\ \mathrm{TeV}$. Technical Report CMS-PAS-HIG-23-016, CERN, Geneva, 2024.

[35] The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3. *Journal of Instrumentation*, 19(05):P05063, may 2024. doi: 10.1088/1748-0221/19/05/P05063.

[36] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, aug 2008. doi: 10.1088/1748-0221/3/08/S08003.

[37] V. A. Schegelsky, M. G. Ryskin, A. D. Martin, and V. A. Khoze. A note on rapidity distributions at the LHC. arXiv:1010.2051 [hep-ph].

[38] The ATLAS collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004, oct 2020. doi: 10.1088/1748-0221/15/10/P10004.

[39] The ATLAS Collaboration. Measurements of WH and ZH production with Higgs boson decays into bottom quarks and direct constraints on the charm Yukawa coupling in $13\,\mathrm{TeV}pp$ collisions with the ATLAS detector. arXiv:2410.19611 [hep-ex].

[40] The ATLAS collaboration. Electron and photon performance measurements with the atlas detector using the 2015–2017 lhc proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006, dec 2019. URL `https://dx.doi.org/10.1088/1748-0221/14/12/P12006`.

[41] The ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, Geneva, 2016.

[42] The ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$=13TeV. *The European Physical Journal C*, 76(5):292, May 2016. ISSN 1434-6052. doi: 10.1140/epjc/s10052-016-4120-y.

[43] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, apr 2008. doi: 10.1088/1126-6708/2008/04/063.

[44] A. Banfi. *Hadronic Jets*. 2053-2571. Morgan Claypool Publishers, 2016. ISBN 978-1-6817-4073-7. doi: 10.1088/978-1-6817-4073-7.

[45] The ATLAS Collaboration. Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett. B*, 816(CERN-EP-2020-093):136204, 2021. arXiv:2008.02508.

[46] D. Krohn, J. Thaler, and L.-T. Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2):84, Feb 2010. ISSN 1029-8479. doi: 10.1007/JHEP02(2010)084.

[47] L. Scodellaro. B tagging in ATLAS and CMS, arXiv:1709.01290 [hep-ex]. Proceeding for the Fifth Annual Large Hadron Collider Physics (LHCP2017) conference.

[48] The ATLAS Collaboration. Atlas flavour-tagging algorithms for the lhc run 2 pp collision dataset. *The European Physical Journal C*, 83(7):681, Jul 2023. ISSN 1434-6052. doi: 10.1140/epjc/s10052-023-11699-1.

[49] The ATLAS Collaboration. Atlas flavour-tagging algorithms for the lhc run 2 pp collision dataset. *The European Physical Journal C*, 83(7):681, Jul 2023. ISSN 1434-6052. doi: 10.1140/epjc/s10052-023-11699-1.

[50] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the atlas detector using proton–proton collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 78 (11):903, Nov 2018. ISSN 1434-6052. doi: 10.1140/epjc/s10052-018-6288-9.

[51] S. Höche. Introduction to parton-shower event generators, 2015. arXiv:1411.4085 [hep-ph].

[52] A. Papaefstathiou. How-to: write a parton-level monte carlo particle physics event generator. *The European Physical Journal Plus*, 135(6):497, Jun 2020. ISSN 2190-5444. doi: 10.1140/epjp/s13360-020-00499-1.

[53] J. Bauer. Prospects for the Observation of Electroweak Top Quark Production with the CMS Experiment. Perspektiven zur Beobachtung der elektroschwachen Produktion einzelner Top-Quarks mit dem CMS Experiment, 2010. URL `https://cds.cern.ch/record/1308713`. Presented on 7 May 2010.

[54] I. Brivio. SMEFTsim 3.0 — a practical guide. *Journal of High Energy Physics*, 2021(4):73, apr 2021.

[55] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks. FeynRules 2.0— A complete toolbox for tree-level phenomenology. *Computer Physics Communications*, 185(8):2250–2300, 2014. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2014.04.012.

[56] 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, Jul 2014. ISSN 1029-8479. doi: 10.1007/JHEP07(2014)079.

[57] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali. Parton distributions with lhc data. *Nuclear Physics B*, 867(2): 244–289, 2013. ISSN 0550-3213. doi: https://doi.org/10.1016/j.nuclphysb.2012.10.003.

[58] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, T. Sjöstrand, P. Skands, M. Utheim, and R. Verheyen. A comprehensive guide to the physics and usage of PYTHIA 8.3. *SciPost Phys. Codebases*, page 8, 2022. doi: 10.21468/SciPostPhysCodeb.8.

[59] D. 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. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(01)00089-4. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[60] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano,

A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. Wellisch, T. Wenaus, D. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(03)01368-8.

[61] The ATLAS Collaboration. Performance of the missing transverse momentum triggers for the atlas detector during run-2 data taking. *Journal of High Energy Physics*, 2020(8):80, Aug 2020. ISSN 1029-8479. doi: 10.1007/JHEP08(2020)080.

[62] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical report, CERN, Geneva, 2015.

[63] LHC Higgs Cross Section Working Group. Simplified template cross sections. URL `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGFiducialAndSTXS`.

[64] R. Barrué, P. C. Muíño, V. Dao, and R. Santos. Simulation-based inference in the search for CP violation in leptonic WH production. *Journal of High Energy Physics*, 2024(4):14, Apr 2024. ISSN 1029-8479. doi: 10.1007/JHEP04(2024)014.

[65] P. Windischhofer, R. Barrué, B. Moser, and H. Arnold. EFTTools Repository. URL `https://gitlab.cern.ch/atlas-physics/higgs/hbb/efttools/-/tree/dev-rcoelhob?ref_type=heads`.

[66] J. Brehmer, S. Dawson, S. Homiller, F. Kling, and T. Plehn. Benchmarking simplified template cross sections in w h production. *Journal of High Energy Physics*, 2019(11):34, Nov 2019. ISSN 1029-8479. doi: 10.1007/JHEP11(2019)034.

[67] C. Bierlich, A. Buckley, J. Butterworth, C. H. Christensen, L. Corpe, D. Grellscheid, J. F. Grosse-Oetringhaus, C. Gutschow, P. Karczmarczyk, J. Klein, L. Lonnblad, C. S. Pollard, P. Richardson, H. Schulz, and F. Siegert. Robust Independent Validation of Experiment and Theory: Rivet version 3. *SciPost Phys.*, 8:026, 2020. doi: 10.21468/SciPostPhys.8.2.026.

[68] B. Rosalino and I. Ochoa. VHbb_CP Repository. URL `https://gitlab.cern.ch/miochoa/vhbb_cp/-/tree/vhbb_cp_beatriz?ref_type=heads`.

[69] A. Buckley, D. Kar, and K. Nordström. Fast simulation of detector effects in Rivet. *SciPost Phys.*, 8: 025, 2020. doi: 10.21468/SciPostPhys.8.2.025.

[70] K. Danziger, S. Höche, and F. Siegert. Reducing negative weights in Monte Carlo event generation with Sherpa. arXiv:2110.15211 [hep-ph].

[71] M. Dobbs and J. B. Hansen. The HepMC C++ Monte Carlo Event Record for High Energy Physics. Technical Report ATL-SOFT-2000-001, CERN, Geneva, 2000.

[72] P. Windischhofer. Physics for particle detectors and particle detectors for physics: Timing performance of semiconductor detectors with internal gain and constraints on high-scale interactions of the Higgs boson, 2022. Presented 24 Jun 2022.

[73] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, Feb 2011. ISSN 1434-6052. doi: 10.1140/epjc/s10052-011-1554-0.

Appendix A

Parameterisation Tables

This appendix contains the results and uncertainties for three cross section parameterisations: considering $Q_\ell \cos\delta^+$; p_T^W; and both, presented for two values of the CP-even operators studied: c_{HW}, and $c_{Hq}^{(3)}$. The parameterisation for these operators was performed (even though it is not used in this thesis), since they affect the p_T^W bins used in this work, and can therefore be useful in future studies of the influence of these operators on the analysis.

A.1 $Q_\ell \cos\delta^+$ Parameterisation

A.1.1 c_{HW}

Table A.1: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{HW} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1,0]$	0.913 ± 0.004	0.410 ± 0.002
$[0,1]$	0.911 ± 0.004	0.409 ± 0.002

Table A.2: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{HW} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1,0]$	0.909 ± 0.004	0.409 ± 0.001
$[0,1]$	0.914 ± 0.004	0.410 ± 0.001

A.1.2 $c_{Hq}^{(3)}$

Table A.3: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{Hq}^{(3)} = 0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1,0]$	2.25 ± 0.01	3.45 ± 0.01
$[0,1]$	2.24 ± 0.01	3.44 ± 0.01

Table A.4: Parameterisation of cross section in terms of $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{Hq}^{(3)} = -0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[-1,0]$	2.24 ± 0.01	3.44 ± 0.01
$[0,1]$	2.25 ± 0.01	3.43 ± 0.01

A.2 p_T^W Parameterisation

A.2.1 c_{HW}

Table A.5: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{HW} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	0.848 ± 0.003	0.241 ± 0.001
$[75, 150]$	0.952 ± 0.004	0.390 ± 0.002
$[150, 250]$	1.04 ± 0.01	0.767 ± 0.006
$[250, 400]$	1.09 ± 0.02	1.63 ± 0.02
$[400, \infty]$	1.18 ± 0.05	4.2 ± 0.1

Table A.6: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{HW} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	0.850 ± 0.004	0.242 ± 0.001
$[75, 150]$	0.952 ± 0.005	0.392 ± 0.002
$[150, 250]$	1.02 ± 0.01	0.768 ± 0.006
$[250, 400]$	1.09 ± 0.02	1.60 ± 0.02
$[400, \infty]$	1.14 ± 0.06	4.1 ± 0.1

A.2.2 $c_{Hq}^{(3)}$

Table A.7: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{Hq}^{(3)} = 0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	1.358 ± 0.006	0.655 ± 0.005
$[75, 150]$	2.13 ± 0.01	1.528 ± 0.009
$[150, 250]$	4.13 ± 0.03	5.69 ± 0.04
$[250, 400]$	8.8 ± 0.1	25.0 ± 0.3
$[400, \infty]$	23.3 ± 0.7	218 ± 6

Table A.8: Parameterisation of cross section in terms of p_T^W (GeV) for $\boldsymbol{c_{Hq}^{(3)} = -0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

Bin	Order 1	Order 2
$[0, 75]$	1.354 ± 0.006	0.647 ± 0.004
$[75, 150]$	2.13 ± 0.01	1.546 ± 0.008
$[150, 250]$	4.15 ± 0.03	5.68 ± 0.04
$[250, 400]$	8.7 ± 0.1	25.3 ± 0.3
$[400, \infty]$	23.4 ± 0.6	216 ± 6

A.3 p_T^W and $Q_\ell \cos\delta^+$ Parameterisation

A.3.1 c_{HW}

Table A.9: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{HW} = 1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W Bin	$Q_\ell \cos\delta^+$ Bin	Order 1	Order 2
$[0, 75]$	$[-1, 0]$	0.851 ± 0.005	0.241 ± 0.002
	$[0, 1]$	0.845 ± 0.005	0.241 ± 0.002
$[75, 150]$	$[-1, 0]$	0.952 ± 0.007	0.393 ± 0.003
	$[0, 1]$	0.951 ± 0.007	0.388 ± 0.003
$[150, 250]$	$[-1, 0]$	1.02 ± 0.01	0.754 ± 0.008
	$[0, 1]$	1.05 ± 0.01	0.780 ± 0.009
$[250, 400]$	$[-1, 0]$	1.13 ± 0.03	1.68 ± 0.04
	$[0, 1]$	1.06 ± 0.03	1.59 ± 0.03
$[400, \infty]$	$[-1, 0]$	1.12 ± 0.07	3.9 ± 0.2
	$[0, 1]$	1.26 ± 0.07	4.5 ± 0.2

Table A.10: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{HW} = -1.0}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W **Bin**	$Q_\ell \cos\delta^+$ **Bin**	**Order 1**	**Order 2**
$[0, 75]$	$[-1, 0]$	0.848 ± 0.005	0.242 ± 0.002
	$[0, 1]$	0.852 ± 0.005	0.242 ± 0.001
$[75, 150]$	$[-1, 0]$	0.954 ± 0.007	0.392 ± 0.003
	$[0, 1]$	0.949 ± 0.007	0.392 ± 0.003
$[150, 250]$	$[-1, 0]$	1.01 ± 0.01	0.758 ± 0.009
	$[0, 1]$	1.04 ± 0.01	0.779 ± 0.009
$[250, 400]$	$[-1, 0]$	1.08 ± 0.03	1.62 ± 0.03
	$[0, 1]$	1.10 ± 0.03	1.57 ± 0.03
$[400, \infty]$	$[-1, 0]$	1.07 ± 0.07	3.9 ± 0.1
	$[0, 1]$	1.23 ± 0.08	4.4 ± 0.2

A.3.2 $c_{Hq}^{(3)}$

Table A.11: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{Hq}^{(3)} = 0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W **Bin**	$Q_\ell \cos\delta^+$ **Bin**	**Order 1**	**Order 2**
$[0, 75]$	$[-1, 0]$	1.367 ± 0.009	0.658 ± 0.006
	$[0, 1]$	1.349 ± 0.007	0.652 ± 0.006
$[75, 150]$	$[-1, 0]$	2.12 ± 0.02	1.54 ± 0.01
	$[0, 1]$	2.13 ± 0.02	1.51 ± 0.01
$[150, 250]$	$[-1, 0]$	4.10 ± 0.04	5.58 ± 0.06
	$[0, 1]$	4.15 ± 0.05	5.81 ± 0.06
$[250, 400]$	$[-1, 0]$	9.0 ± 0.2	25.6 ± 0.5
	$[0, 1]$	8.5 ± 0.2	24.3 ± 0.5
$[400, \infty]$	$[-1, 0]$	21.7 ± 0.8	204 ± 7
	$[0, 1]$	25 ± 1	235 ± 10

Table A.12: Parameterisation of cross section in terms of p_T^W (GeV) and $Q_\ell \cos\delta^+$ for $\boldsymbol{c_{Hq}^{(3)} = -0.1}$, excluding all other operators, considering the linear (order 1) or the quadratic term (order 2)

p_T^W **Bin**	$Q_\ell \cos\delta^+$ **Bin**	**Order 1**	**Order 2**
$[0, 75]$	$[-1, 0]$	1.354 ± 0.009	0.647 ± 0.006
	$[0, 1]$	1.353 ± 0.008	0.647 ± 0.006
$[75, 150]$	$[-1, 0]$	2.13 ± 0.02	1.56 ± 0.01
	$[0, 1]$	2.12 ± 0.02	1.54 ± 0.01
$[150, 250]$	$[-1, 0]$	4.09 ± 0.05	5.60 ± 0.06
	$[0, 1]$	4.21 ± 0.04	5.75 ± 0.06
$[250, 400]$	$[-1, 0]$	8.9 ± 0.2	26.0 ± 0.4
	$[0, 1]$	8.4 ± 0.2	24.6 ± 0.5
$[400, \infty]$	$[-1, 0]$	21.5 ± 0.9	202 ± 8
	$[0, 1]$	26 ± 1	232 ± 10