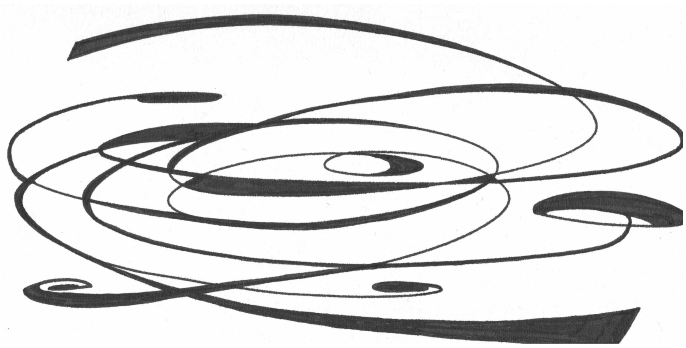
CP-violation in the GGF HWW channel



Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Physics

Author :	Liya Charlaganova
Student ID :	s2548852
Supervisor :	Prof.dr. I.B. van Vulpen (LION[1])
Second Supervisor :	dr. L. Brenner (Nikhef[2])
Daily Supervisor :	dr. O. Rieger (Nikhef[2])

Leiden, The Netherlands, March 2, 2025

CP-violation in the GGF HWW channel

Liya Charlaganova

[1]Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

[2]Nikhef, Amsterdam Science Park
P.O. Box 41882, 1009 DB Amsterdam, The Netherlands

March 2, 2025

Abstract

The Standard Model (SM) does not account for enough CP-violating processes to explain the matter-antimatter asymmetry seen in the universe. One promising area to look for additional CP-violation is the Higgs boson, which still houses unexplored areas due to its recent discovery. In this thesis, the Standard Model Effective Field Theory (SMEFT) approach was used to extend the Higgs sector beyond the SM. The sensitivity of the ATLAS detector at CERN to the CP-odd SMEFT operator $\hat{O}_{H\tilde{G}}$ in the ggF$\rightarrow$HWW channel was studied. Different analyses on simulated data were compared in order to find an optimized strategy. It can be concluded that applying phase-space cuts on the leading di-jet mass m_{jj} leads to a significant increase in sensitivity, especially when compared to cuts on p_T^H. The $m_{jj} > 200$ GeV is deemed to be the most effective.

Contents

Chapter 1

Introduction

For centuries, humanity has tried to unravel the underlying nature of reality. As evidence emerged, old theories were replaced by new ones, and the search for a Theory of Everything continues. Currently, one of the fundamental pillars of modern physics is the Theory of the Standard Model (SM). And even though its predictions have been thoroughly verified by experiment, it cannot be fully correct. Still, no new theory has been able to replace SM just yet.

One unanswered question of the SM is its apparent lack of CP-violation. The existence of matter in the universe suggests that the necessary CP-violation can only be explained by theories which lie Beyond the Standard Model (BSM). In order to test as many BSM theories as possible, a theory-agnostic approach is used: Standard Model Effective Field Theory (SMEFT). This approach does not presume some specific theory to describe observed phenomena, rather, it extends the SM beyond its limits with higher-order terms. Most importantly, the vast amount of data generated by the Large Hadron Collider allows us to test the existence of such terms.

This thesis presents a sensitivity study of the ATLAS detector at CERN to CP-violating behaviour in the Higgs sector using simulated data. More specifically, the effects of the CP-odd SMEFT operator $\hat{O}_{H\tilde{G}}$ are studied in the ggF$\rightarrow$HWW channel.

The Higgs boson is the last SM particle to be discovered, and no CP-odd operators have been studied in the Higgs sector yet.

Chapter 2 will provide an overview of the Standard Model and introduce the CP-violation question that is prevalent in SM. Afterwards it will introduce

the SMEFT approach that is used to study the Higgs boson beyond the SM. The ATLAS detector is described in chapter 3. Production and decay of the Higgs boson is discussed in Chapter 4 along with relevant background processes. The analysis is described in Chapter 5, and the simulations in Chapter 6. Lastly, results are presented in Chapter 7 and further discussed in Chapter 8.

Chapter 2

Theory: the Standard Model and Beyond

The Standard Model (SM) of particle physics and the theory of General Relativity are the two best tested theories in physics. So far no experiment has been able to definitively contradict the predictions made by these two models. SM describes all known fundamental particles and their interactions through the weak, strong and electromagnetic forces. Together with gravity, which is described by General Relativity, they make up the four fundamental forces. Although both theories have been thoroughly tested, the models are known to be incomplete. Therefore new ways of extending the SM are being proposed. One of these Beyond the Standard Model (BSM) theories is the Standard Model Effective Field Theory (SMEFT). It treats SM as a low-energy approximation of a grander theory. This chapter provides an overview of the relevant aspects SM and an introduction to SMEFT.

2.1 The Standard Model

SM is a quantum field theory which describes the fundamental particles of matter (fermions) and interactions between particles through three forces (mediated by gauge bosons). An overview of the particles and their characteristics can be seen in Figure 2.2 and the interactions are visualised in Figure 2.3.

Fermions
Fermions are particles that have spin 1/2 and make up all known matter in

the universe. These are split into two groups: quarks and leptons. There are three generations of fermions, each next generation being heavier than the previous. Only first generation quarks are stable and make up protons and neutrons, while higher generations will quickly decay to either the up or down quark. Quarks are the only particles to interact with all three forces. They carry colour charge, facilitating interaction through the strong force. Colour confinement states that isolated coloured particles cannot exist for prolonged periods of time. Therefore, quarks clump together into hadrons. Two quarks form mesons (pions, kaons) while three quarks form baryons (protons, neutrons). When two quarks get torn apart, quark-antiquark pairs are produces from the binding energy of the two original quarks. The further two quarks are pulled apart, the more pairs are produced in a process called hadronization. For the top quark, hadronization time is longer than its lifetime, which means the top-quark decays into a bottom-quark via the weak force. It is the only quark that can decay, all others will undergo hadronization. Leptons also come in three generations, but do not have colour charge and therefore do not interact with strong force. Each generation consists of a charged lepton (electron: e, muon: μ, tau:τ) and an associated neutral neutrino (ν_e, ν_μ, ν_τ). Neutrinos are very light and only interact through the weak force making them difficult to detect. Finally, all fermions have an antiparticle version of themselves. Antiparticles have the same spin and mass as their counterpart, but a flipped charge.

Gauge bosons
Gauge bosons are particles with spin-1 and mediate three fundamental forces through their interaction with fermions. The strong force is mediated by the gluon: a massless particle that carries the colour charge and able to interact with other colour-charged particles (quarks) as well as itself. The photon mediates the electromagnetic force and is also massless. It interacts with all electrically charged particles, but lacks an electric charge itself so there is no self-interaction. The $\mathrm{W}^{\pm}$ and Z^0 bosons are massive particles that together mediate the weak force and interact with all fermions. Additionally the $W^{\pm}$-bosons are electrically charged and are therefore coupled to photons as well.

Higgs boson
The last particle to be found predicted by the Standard Model was confirmed to exist by ATLAS and CMS experiments in 2012[3][2]. It is a spin-0 particle with a mass of approximately 125 GeV[8] and leads to the masses of all particles through the Higgs mechanism.

[Higgs mechanism] A Higgs field Φ with four degrees of freedom (d.o.f.) is

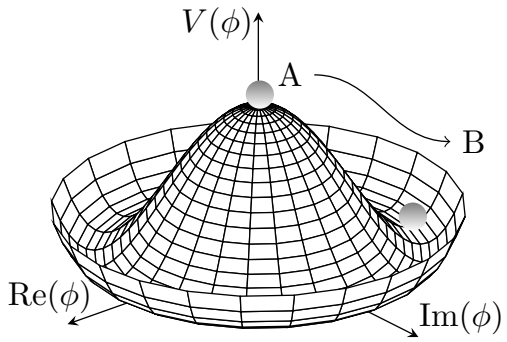


Figure 2.1: Visual representation of the Higgs potential [18]. The expectation value is non-zero (point A), and symmetry is broken as the Higgs field transitions to the ground state (point B).

introduced to the SM Lagrangian. It is subject to a scalar potential with a non-zero expectation value (see Figure 2.1) which has the form

$$V(\phi) = m^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$$

As the Higgs field transitions to the ground state, spontaneous symmetry breaking occurs. This eliminates three degrees of freedom which are absorbed by the $W^\pm$- and Z^0-boson fields, giving them mass. The unbroken d.o.f. becomes the massive Higgs boson. Lastly, fermions acquire their mass through Yukawa interactions between the Higgs field and the fermionic fields.

2.1.1 Mathematical description

The SM can be famously written down in a single mathematical equation using the Lagrangian formalism, which takes up a whole page. In simplified terms, it describes the kinetic and potential energies of a system and allows for easier calculations of the equations of motions than Newtonian physics. Fermions and bosons are quantum fields which evolve through time and can interact with one another at vertices. The evolution of an isolated field is described by the kinetic part of the Lagrangian: a mass term of the form $-\bar{\psi}m_\psi^2\psi$, with m the mass of the particle, and a purely kinetic terms of the form $\partial^\mu\bar{\psi}\partial_\mu\psi$. If the mass of a particle coincides with the resonance peak m, it is said to be on-shell. However, it is possible to create off-shell particles; particles that have a mass not equal to the resonance mass, but these will be less stable and are more unlikely to form in general. An example of an interaction term is $-g\bar{\psi}\gamma^\mu\psi A_\mu$, which would describe the interaction between a fermionic field ψ and a bosonic field A_μ with a strength g.

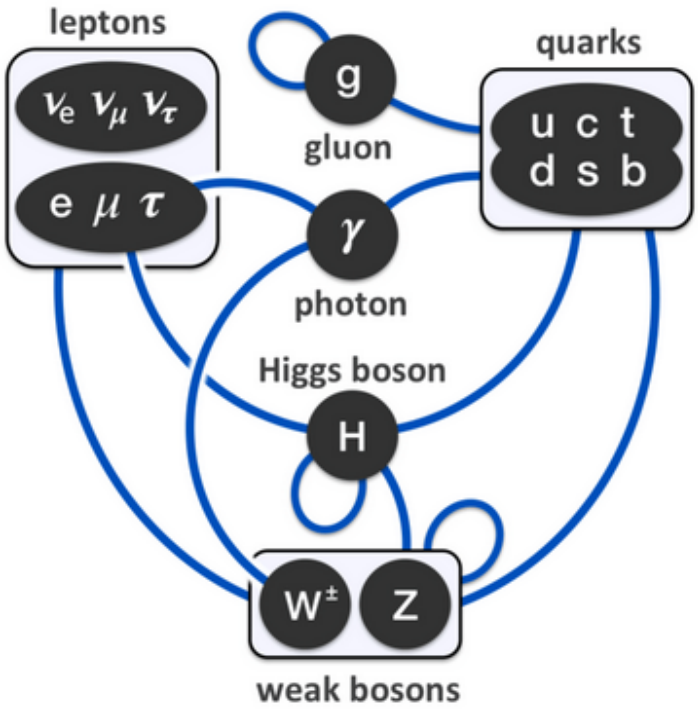


Figure 2.2: A visualisation of the interactions between the particles of the SM. Two particles connected by a solid line directly couple to each other in the SM. A particle with a loop couples to itself. Figure taken from [15].

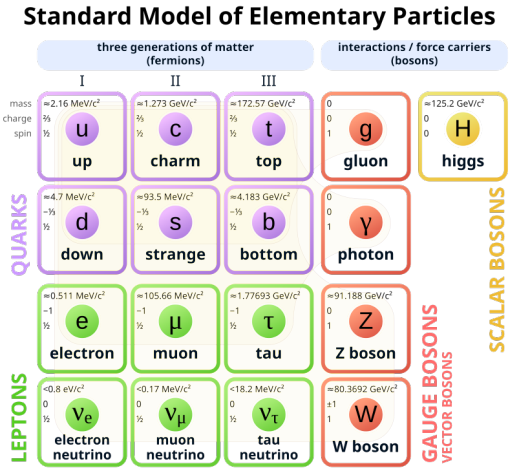


Figure 2.3: An overview of the particles of the SM and their properties[11].

In Quantum Field Theories (QFT), the probability of a certain process happening can be calculated. This is then compared to experimental results and can be used to verify a theory, or to determine values for its free parameters. This probability is expressed as a cross-section σ, and is proportional to the amplitude A of a process. A higher cross-section corresponds to a higher probability of a specific process occurring. Richard Feynman developed a method[13] to calculate cross-sections by using Feynman diagrams which visually represent the interaction in an intuitive way. Though no such calculations will be performed in this thesis, Feynman diagrams will be used to visualise processes. The reader will likely be able to grasp these even without a formal course on QFT.

2.1.2 CP-violation and the SM

Even though all predictions of the SM have been thoroughly verified, we know that it is incomplete. One observation that the SM cannot explain is that the universe is made up of matter and virtually no antimatter. From SM there is no reason for this asymmetry to exist to the extend that it does. This problem is referred do as the baryon asymmetry problem, and in 1967 Andrei Sakharov proposed three conditions that need to be met in order to produce more matter than antimatter[19]:

- Baryon number violation
- Reactions out of thermal equilibrium
- CP-violating processes

This thesis will focus on the last condition: CP-violating processes. CP is the combination of Charge and Parity transformations: charge transformation amounts to taking the antiparticle, and parity transformation reverses the three spacial axes. A process is said to be CP-symmetric if it is the same before and after this CP-transformation, and CP-violating if it is not. An experiment by Wu and collaborators in 1956[22] proved that the combined CP-symmetry is violated by the weak force. Since then, other CP-violating processes predicted by the SM have been observed, but so far none are sufficient to explain baryon asymmetry. The search has been extended to BSM theories, but the Higgs sector is so far relatively unexplored due to its recent discovery.

2.2 Standard Model Effective Field Theory

Standard Model Effective Field Theory (SMEFT) is a very general approach to probe BSM physics without assuming a specific theory. In SMEFT, SM is seen as a low energy approximation of a higher order theory in which BSM physics come into play at high energies ($\gg$TeV) and are suppressed at lower energies. The SM Lagrangian is expanded by adding all possible higher-order operators $\hat{O^{(d)}}$, with $d > 4$*, that comply to certain symmetry arguments[12]:

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_{d}\sum_{i} \frac{c_i^{(d)}}{\Lambda^{(d-4)}} \hat{O}_i^{(d)}$$

The operators $\hat{O^{(d)}}$ describe SMEFT interactions, which are modelled as unresolved blobs: physics within this blob remain unknown, but its effects on outgoing particles can be measured. These BSM interactions are suppressed by Λ (=1 TeV), which represents the energy scale of new physics. Additionally, this reduces the dimensionality of the term back to 4 in order to match $\mathcal{L}_{SM}$. The strengths of the operators are scaled by Wilson coefficients $c_i^{(d)}$. These coefficients cannot be predicted by theory, instead they are free parameters that must be measured experimentally. If a Wilson coefficient is constrained to equal 0, this means the corresponding operator does not play a role. Even though we cannot probe the high energy scale of SMEFT, small effects could be visible at the limits of the accessible energy range in the form of deviations from SM prediction, which can be used to constrain Wilson coefficients (see [6] for an example).

These analyses have been performed in the Higgs sector, but no CP-violating operators have been considered yet. This makes studying CP-violating SMEFT operators in the Higgs sector especially promising as it is an unexplored area of research and it could lead to an explanation of the baryon asymmetry problem. For this thesis, a single CP-odd dimension-six† operator is considered: $\hat{O}_{H\tilde{G}}$. It describes the direct interaction between the Higgs boson and gluons, an interaction which is impossible in SM as the gluon is massless, but can be facilitated through a quark loop.

*Operators with $d \leq 4$ are considered in $\mathcal{L}_{SM}$

†Dimension-five operators violate lepton number, meaning dimension-six are lowest order lepton-number-conserving operators. The higher the order, the more the operator is suppressed.

Version of March 2, 2025– Created March 2, 2025 - 15:18

2.2.1 Parametrization

Ultimately, particle physics is limited to measuring cross-sections of specific processes. The BSM contribution, referred to as EFT, can be added to the observed cross-section as follows:

$$\begin{aligned}\sigma_{SMEFT} &\propto |A_{SMEFT}|^2 \\ &= |A_{SM} + A_{EFT}|^2 \\ &= A_{SM}^2 + A_{int} + A_{EFT}^2 \\ &\propto \sigma_{SM} + \sigma_{int} + \sigma_{BSM}\end{aligned}$$

This thesis focuses only on the SM term (σ_{SM}) and the interaction term (σ_{int}), the quadratic EFT term σ_{BSM} is not considered. We can rewrite σ_{SMEFT} as

$$\sigma_{SMEFT} = \sigma_{SM} \times \left(1 + \frac{\sigma_{int}}{\sigma_{SM}}\right)$$

which let's us parametrize the equation in a simple manner:

$$\sigma_{SMEFT} = \sigma_{SM} \times (1 + \sum_i c_i A_i)$$

Here c_i is the Wilson coefficient that we ultimately want to determine after performing a global SMEFT fit. The parametrization coefficient A_i is determined from simulations that calculate the cross-section for a known c_i value, usually $c_i = 1$. The goal is to model the cross-section as follows:

- For $c_i = 0$, SM result is obtained
- For $c_i = 1$, the SM + EFT result is obtained.

This is a direct fit for A_i, which can be compared between simulations. This parametrization scheme ensures that the expected values for c_i are on the order of 1. If the best-fit coefficients deviate significantly from 1, it can be concluded that we are outside of the regime where our assumptions hold.

Version of March 2, 2025– Created March 2, 2025 - 15:18

Chapter 3

The ATLAS experiment

In order to test SM and probe its limits, we need to observe its behaviour in highly-energetic environments. Particle accelerators can achieve these regimes in a controlled fashion, and are the most widely used type of experiment for high energy particle physics. The characteristics of accelerators and detectors have an impact on the data analysis, and therefore differ for each detector. The simulations and analyses of this research are based on the ATLAS detector at the Large Hadron Collider at CERN.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest particle accelerator capable of achieving proton-proton collision energies of 13.6 TeV. It was built by the European Organisation for Nuclear Physics (CERN) at the border between France and Switzerland. The circular 27 km-circumference tunnel, which currently houses the LHC, was originally built for the Large Electron-Positron Collider (LEP). The LEP collider used leptons instead of protons, which greatly reduced the collision energy due to synchrotron radiation.

Particles are accelerated in multiple stages before being injected into the LHC. Here the protons travel in opposite directions in bunches, each containing on average 10^{11} protons. The diameter of the proton beam is only tens of microns[1], which means that head-on collisions where the full centre-of-mass energy of 13 TeV is available, hardly happen. Finding suitable collisions in data obtained at the LHC is a big challenge for all analyses and requires storing and processing vast numbers of events. If the process under to study

is rare, there might not be enough collisions of the required type in order to achieve a significant result. This can be mitigated by collecting more data over a longer period, often expressed in terms of integrated luminosity, measured in m^2. Currently, the LHC is on its third run which started in 2022 and is expected to last until 2026. The collision energy increased to 13.6 TeV and a total luminosity of 300 fb^{-1} can be expected.

The LHC houses 4 big experiments: ATLAS, CMS, ALICE and LHCb. ATLAS and CMS are all-purpose detectors working separately and verifying each other's results. ALICE's main focus is studying the quark-gluon plasma that would have filled the universe shortly after the Big Bang. Lastly, the LHCb is specialised in studying b-hadrons and CP-violation in this sector.

3.2 ATLAS experiment

The ATLAS detector is the largest detector at the LHC. It is designed as a general purpose detector, meaning it can study a broad range of physics process. Particles of interest often have short lifetimes and decay quickly, making direct detection impossible. Instead, properties of the original particle need to be reconstructed from its decay products.

3.2.1 Structure of detector

The whole ATLAS experiment is built as a layered cylindrical detector around the beamline along which the particles travel. Each layer is made up of different detectors whose measurements are combined in order to reconstruct and identify particles produced in the collision. A schematic overview of the detector can be seen in Figure 3.1. At the centre is the Inner Detector which is made up of three sensors systems immersed in a magnetic field. Together they provide information on the origin, type and and momentum of charged particles, which are bent by a magnetic field. Next follows the Electromagnetic Calorimeter and the Hadronic Calorimeter which measure the total energy of particles. When a particle enters the calorimeter, it interacts with it and forms a shower of lower energy particles. The energy that is radiated away in the interactions is measured and recombined to obtain the total energy of the original particle. The calorimeters are able to fully stop most hadrons, electrons and protons. Muons, which are significantly heavier than electrons, pass through these calorimeters and are measured by the Muon Spectrometer: a collection of five types of detectors able to measure the mo-

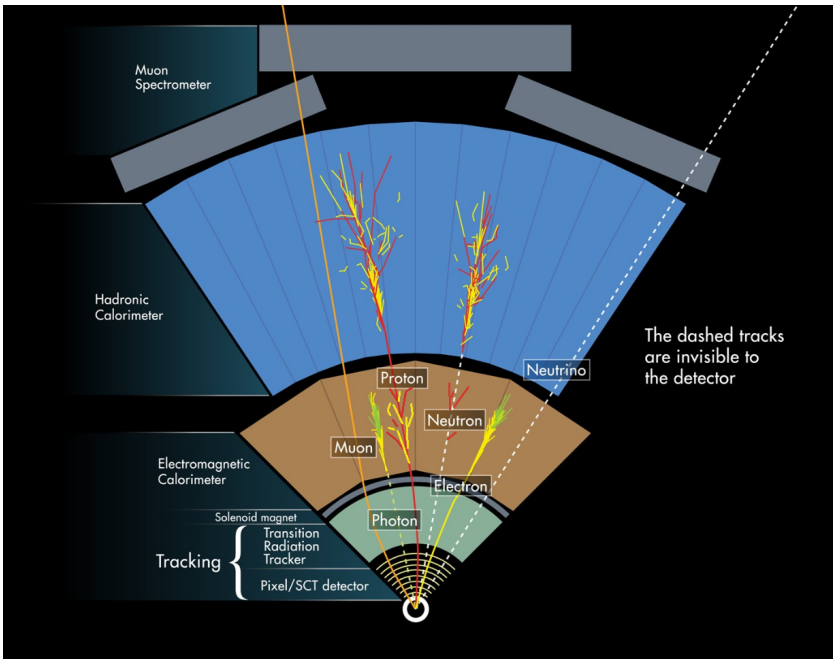


Figure 3.1: Schematic representation of the ATLAS detector with particle tracks. Figure taken from [14]

menta of muons. Still, the weakly-interacting neutrino passes through the whole detector unimpeded, carrying away momentum. Although neutrinos cannot be directly measured, some properties can be reconstructed from the Missing Transverse Energy (MET).

3.2.2 Coordinates

A cylindrical coordinate system is used with the origin at the centre of the detector, shown in Figure 3.2. The z-axis is defined along the travel path of the proton bunches: the beamline. The xy-plane, perpendicular to the beamline, is the transverse plane. Directions in the transverse plane are labelled by their angle around the beamline axis ϕ (azimuthal angle). The angle away from the beamline can considered θ, however, usually pseudo-rapidity (η) is used instead:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$$

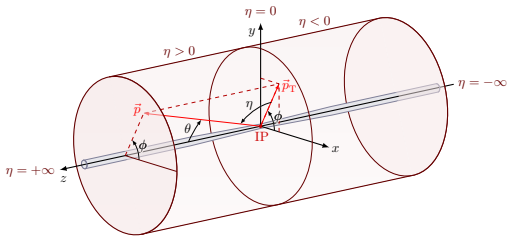


Figure 3.2: Schematic representation of the ATLAS detector. The z-axis points along the beamline, the x-axis towards the centre of the main LHC ring. Adapted from [17].

3.2.3 Event and object reconstruction

Colliding protons at high energies brings with it numerous challenges. Firstly, the collision of protons happens in bunches, which results in multiple collisions along the beamline in the detector. The additional (not interesting) collisions are referred to as pile-up. Furthermore, most collisions that happen at the LHC do not produce any particles of interest. A lot of systems are in place in order to filter out likely candidates. These two aspects are not considered in this thesis, but they serve as a reminder of how complicated a full-scale analysis truly is. In this thesis, two other aspects of a detector are considered: jets and MET.

Jets

Every single (hard-scatter) collision vertex produces a multitude of different particles of which some may decay further depending on their lifetime. As mentioned before, quarks will undergo hadronization and form hadrons which all travel in a similar direction. When this hadron enters a detector, it will produce a large number of new particles in which results in a cone structure. This cone is referred to as a 'jet'. Distinguishing different jets can be difficult due to overlap leading to loss of the total number of jets.

B-jets

B-jets are jets that originate from a bottom quark. It is possible to reconstruct the path of the jet backwards in order to find its origin: the point where a quark began hadronization. Since hadronization is not an instantaneous process, a quark can travel some distance from the primary impact

vertex of the two protons before turning into a jet. How fast a quark will undergo hadronization depends on its mass: the more massive the particle, the farther it can travel in the detector. The origin point of the jet is a secondary vertex. By measuring the distance between the two, it is possible to distinguish, or 'tag', different types of quarks. Currently, only the b-quark can be tagged reliably as it is the heaviest particle that undergoes hadronization (the top-quark decays to a b-quark). These are the b-jets, all other jets are light-jets.

Missing Transverse Energy

Since the particles travel in a straight line, the transverse (perpendicular) momentum is zero both before and after the collision. If the reconstructed transverse energy is non-zero, it indicates the presence of undetected particles. Their total transverse momentum can be inferred and is known as Missing Transverse Energy (MET). This quantity is used to infer neutrino transverse momentum, but could also be caused by new weakly interacting unknown particles which are not considered in this thesis.

Chapter 4

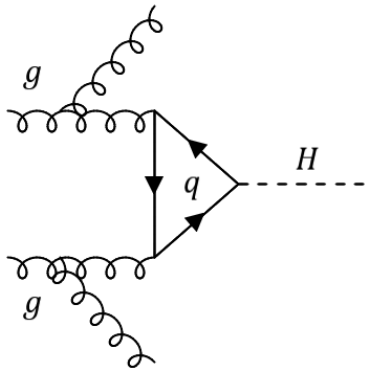
The ggF→HWW channel

The Higgs boson is an unstable particle with a short lifetime of 2.06×10^{-22} seconds[4]. It is produced during proton-proton collisions, but then quickly decays, making it impossible to study it directly. There are multiple ways to produce and decay a Higgs boson. In this thesis, Higgs production through the gluon-gluon fusion (ggF) channel is considered along with decay through the HWW channel. The operator $\hat{O}_{H\tilde{G}}$ modifies the ggF production mode.

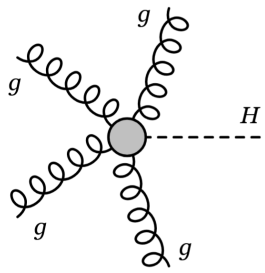
4.1 Higgs production through ggF

In the ggF production mode, two gluons, originating from the energetic protons that travel in opposite direction, interact through a quark-loop to produce a Higgs boson. Additional jets from gluons or quarks can be produced as well. A diagram of ggF with two additional jets is shown in Figure 4.1a. The loop can be made up of any of the six quarks, but since the coupling of the top quark to the Higgs is strongest, it is the most efficient and likely mediator.

In SMEFT, the $\hat{O}_{H\tilde{G}}$ reduces the whole ggF process to a single interaction with two incoming gluons and one outgoing Higgs boson and, optionally, additional outgoing jets. Figure 4.1b shows a the diagram for ggF with two additional jets as seen in SMEFT. The observable most sensitive to $\hat{O}_{H\tilde{G}}$ is the angle between the two additional outgoing jets. So the CP-sensitive signal (process of interest) is narrowed down to $\geq$2-jet ggF.



(a) ggF production mode in the SM.



(b) ggF interaction vertex in SMEFT, which becomes unresolved. The vertex is modified by the $\hat{O}_{H\tilde{G}}$ operator. Image courtesy of [20].

Figure 4.1: Gluon-gluon fusion (ggF) production mode of the Higgs boson in SM and SMEFT. Here, the two additional jets come from gluons, but quark-jets are also possible.

4.2 Higgs Decay $H \to WW^* \to l\nu l\nu$

The Higgs boson has a decay width of 4.07×10^{-3} GeV, corresponding to a lifetime of 1.6×10^{-22} seconds, so it does not live long enough to be directly observed by the detector. After it's formation, a Higgs boson quickly decays into a pair of fermions or a pair of bosons. The leading contribution is a decay into a $b\bar{b}$ pair at 53%[16]. Decay into other quark-antiquark pairs and lepton anti-lepton pairs are possible but have only small contributions. The second most likely decay mode is decay into a pair of W-bosons, denoted as HWW decay, at 26%. Since we need at least two jets produced in ggF, it is preferable to have fewer jets present from the decay mode. Any quark produced during decay will constitute another jet. For this reason, the analysis does not consider $b\bar{b}$ decay and is restricted to purely leptonic decays of the W-bosons: $WW* \to l^+\nu l^-\nu$ (see Figure 4.2). The two highly energetic oppositely leptons are required tags for the HWW channel.

Even though two W-bosons have a higher rest mass (161 GeV) than a Higgs boson, this decay is still possible since one (or both) W-bosons are be off-shell (denoted as W^*) for a short period of time before decaying further. The W-boson has a decay width of 2.085 GeV which translates to a lifetime of

3.2×10^{-25}s through

$$\tau_x^{[½]} = \hbar/\Gamma_x$$

So a W-boson quickly decays further into a meson (one quark and one anti-quark with a total charge of ± 1) or a lepton-neutrino pair. The lifetime of the τ is 2.903×10^{-13}s, which is too short to be detected. Further decay of the τ is considered as well. It can also decay into a lighter lepton-neutrino pair or into a pion. The total branching ratio for a W-boson decaying into a an electron or a muon, with an associated neutrino, is 25%. The total branching ratio for the Higgs to two leptons is therefore 6.4%. It is important to note that this calculation is based on on-shell W decays and that due to one of the W-bosons being off-shell the ratio might deviate slightly.

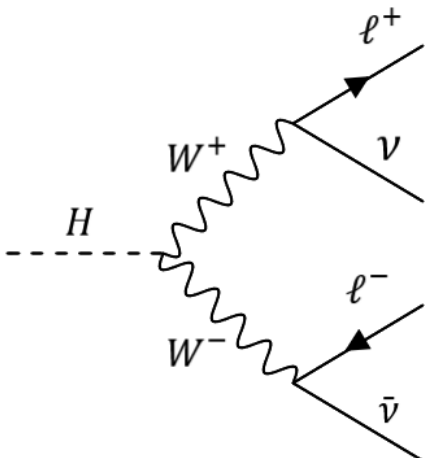


Figure 4.2: Feynman diagram of HWW decay channel. Final state consists of two oppositely charged leptons and two neutrinos.

The properties of the Higgs boson are reconstructed from its decay products. In order to take into account the effects this reconstruction has on the result, such a reconstruction is also applied to the simulated data.* The reconstruction only takes into account the two leptons which the Higgs decays into through the W-bosons, and ignores the neutrinos.

$$\vec{p}_T^{\,H} = \vec{p}_T^{\,l1} + \vec{p}_T^{\,l2} + \vec{E}_T^{\,miss}$$

*Also, virtual Higgs boson is not saved by default in the Monte Carlo generator.

4.3 Process Backgrounds

In order to increase the sensitivity of our analysis, background processes need to be removed. These are processes that look similar to our channel (ggF→HWW) in the detector, but are not useful for our purposes. The top-quark background will be modelled alongside the signal channel.

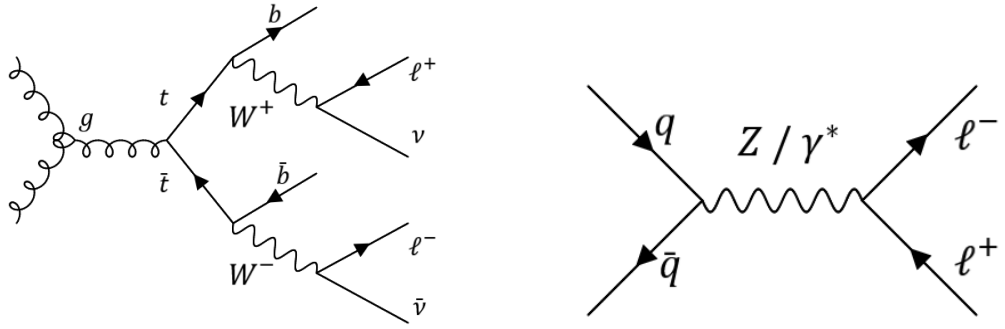
4.3.1 Top-quark background

The top-quark background ($t\bar{t}$ background) is the most prominent background in the HWW decay channel and is shown in Figure 4.3a. Two partons (quark-antiquark pair or two gluons) from the two initial protons interact form a single high energy gluon, which can then decay into a $t\bar{t}$ pair. Top quarks decay to bottom quarks before undergoing hadronization, producing W-bosons in the process. Due to the neutral charge of the intermediate gluon, the two W-bosons have opposite charge and will decay into a lepton-antilepton pair (and neutrinos). This final state is the same as the HWW final state, with the only difference being that $t\bar{t}$-background has additional b-quarks. It is possible to distinguish b-jets formed from these quarks, but with limited efficiency. Furthermore, the jets in 2-jet ggF can be b-jets, meaning that some signal is rejected as well.

4.3.2 Same flavour lepton backgrounds

Numerous processes exist that result in a two-lepton end state. One example is the Drell-Yan process where a quark and an anti-quark from colliding protons annihilate and form a neutral boson (Z^0/γ), which then decays into a lepton/anti-lepton pair (see Figure 4.3b). Such a decay is not isolated to bosons directly produced in the collision, but is also possible from neutral bosons that appear elsewhere. This type of background always has same flavour leptons because they appear from a single particle. Even though this type of background is not simulated in this research, an additional requirement of two different-flavour leptons is added to the analysis.

Version of March 2, 2025– Created March 2, 2025 - 15:18



(a) $t\bar{t}$ background with a dilepton final state.

(b) Example of Drell-Yan process where a lepton-antilepton pair is created.

Figure 4.3: Two background processes present in the HWW decay channel

4.4 $\Delta\phi_{jj}$ observable

The observable sensitive to the CP-odd operator $\hat{O}_{H\tilde{G}}$ is $\Delta\phi_{jj}$ in $\geq$2-jet ggF. More specifically, it is the η-ordered signed $\Delta\phi_{jj}$. $\Delta\phi_{jj}$ designates the the azimuthal angle between the two jets with the highest transverse momentum (p_T). "Signed" means that this angle lies between $-\pi$ and π, covering the whole circle. This angle depends on the order of subtraction, with the two possibilities differing by a minus sign. The η-ordered determines this order: if the jet with the higher η is $j1$, and the other jet is $j2$ (so $\eta_{j1} > \eta_{j2}$), then the angle is $\Delta\phi_{jj} = \phi_{j1} - \phi_{j2}$. Figure 4.4 shows how the $\Delta\phi_{jj}$ distribution changes when the $c\_H\tilde{G}$ (EFT) contribution is added.

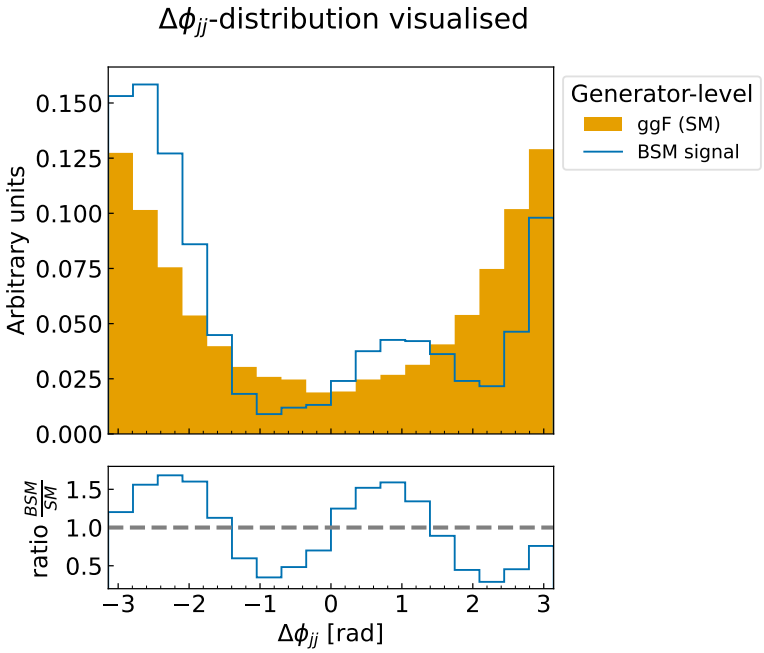


Figure 4.4: $\Delta\phi_{jj}$ distributions of a ggF SM sample (orange bars), and ggF BSM (SM + EFT) (blue line). The ratio between the two is shown at the bottom.

Chapter 5

Analysis strategy

In the search for BSM physics, the goal is to find some deviation from the SM prediction in the data. For this thesis, the goal is to study the sensitivity of ATLAS to the CP-odd SMEFT operator $\hat{O}_{H\tilde{G}}$, an operator which has not been properly studied before. If the Wilson coefficient $c_{H\tilde{G}}$ is non-zero, we expect the $\Delta\phi_{jj}$-distribution to deviate from the SM prediction. Simply calculating this angle distribution for all data would not be effective since there is a lot of noise: not all proton-proton collisions produce a ggF Higgs boson. In fact, most collisions do not even produce a Higgs boson. Furthermore, the strength of $\hat{O}_{H\tilde{G}}$ can depend on the characteristics of the event. The analysis strategy should isolate events that are modified by $\hat{O}_{H\tilde{G}}$ while simultaneously discarding events that contain background processes or are not sensitive to $\hat{O}_{H\tilde{G}}$. The better an analysis does this, the higher the signal-to-noise ratio will be. This is done by making cuts in the measurement phase space: only events with certain parameter values (e.g. energy above some threshold) are accepted. Multiple analysis strategies will be compared in order to determine how to best increase sensitivity to CP-odd effects.

5.1 The CP-odd ggF→HWW analysis

The analysis for this thesis can be split into two parts: selection cuts and kinematic cuts. The goal of the selection cuts is to reject background processes and remove unreliable data but keep the signal (ggF H→ $WW^* \to l\nu l\nu$ + 2-jets) events. Afterwards, the goal is to increase the sensitivity of the analysis by selecting events that most affected by $\hat{O}_{H\tilde{G}}$.

5.1.1 Selection cuts

The selection cuts for the ggF→HWW channel were adapted from [10]. The cuts are as follows:

1. The two highest-p_T leptons are required to have opposite sign and different flavours, so: $e^-\mu^+$ or $e^+\mu^-$ (OSDF leptons)
2. The event must contain at least 2 reconstructed jets. ($n_{jets} \geq 2$ jets:)
3. Events that contain b-jets are discarded (b-jet veto)
4. The combined mass of the two highest-p_T leptons must be below 55 GeV. ($m_{ll} < 55$ GeV)
5. The absolute azimuthal angle between the two leptons must be above 1.8 radians. ($\Delta\phi_{ll} > 1.8$)

We only look at OSDF leptons, as this cut is commonly applied in order to reduce certain backgrounds in real data. Even though these backgrounds are not simulated for this thesis, this selection is implemented in order to be more representative of a real analysis in this channel. The $m_{ll} < 55$ GeV, $\Delta\phi_{ll} > 1.8$ and b-jet veto cuts are applied to reduce the $t\bar{t}$ background. Lastly, the event must contain at least 2 jets so that $\Delta\phi_{jj}$ can be calculated.

5.1.2 Kinematic cuts

After the selection cuts, there are few constrains on the properties of the Higgs boson. This is the full, or inclusive cross-section. By applying additional cuts on event properties, sensitivity can be increased and more precise comparisons between theory and experiment can be made. However, this decreases the number of events that are accepted in the final binning which decreases statistical significance. In recent years, it has become possible to study Higgs-events in this way due to the increase of the total integrated luminosity of the LHC. A standardised method of applying phase space cuts in Higgs boson events is the Simplified Template Cross Section (STXS) scheme[23]. The latest version of the STXS scheme for the ggF channel is shown in Figure 5.1. Each event is categorised in a single bin based on its properties: p_T^H (transverse momentum of the Higgs boson), m_{jj} (total mass of the two leading-p_T) and the number of jets. This version does not contain any binning in the $\Delta\phi_{jj}$ variable, so it is impossible to determine any

Version of March 2, 2025– Created March 2, 2025 - 15:18

CP-odd properties of the Higgs using these cuts. Future versions of STXS will introduce CP-sensitivity.

STXS is taken as inspiration for this analysis. After event selection, cuts on the variables p_T^H and m_{jj} are applied. Afterwards, events are binned in $\Delta\phi_{jj}$, making the analysis sensitive to CP-odd behaviour. Different types of p_T^H and m_{jj} cuts are compared:

Binary cuts A single cut is applied to a single parameter (e.g. $p_T^H > 200$ GeV). This method has the advantage of being very simple (and therefore more easily compatible with other analyses) and keeping a large number of events. Cuts above and below a value are considered (e.g. > 200 GeV and < 200 GeV) for both variables.
Multi-bin cuts The phase space of a single variable is split into multiple bins. This is the closest analogy to the STXS scheme where all final bins are simultaneously fitted to the data, which significantly increases sensitivity. Such a fit is not performed here, instead the final sensitivity is approximated from the sensitivity of individual bins.
Combination cuts The best cuts on each individual variable are combined into a single cut (e.g. $m_{jj} > 200$ GeV + $p_T^H > 200$ GeV

Events that have passed the phase space cuts are then binned according to their $\Delta\phi_{jj}$, which ranges from $-\pi$ to $+\pi$. This interval has been split into 4 equally sized regions, after careful investigation of the sensitivities for different number of bins.

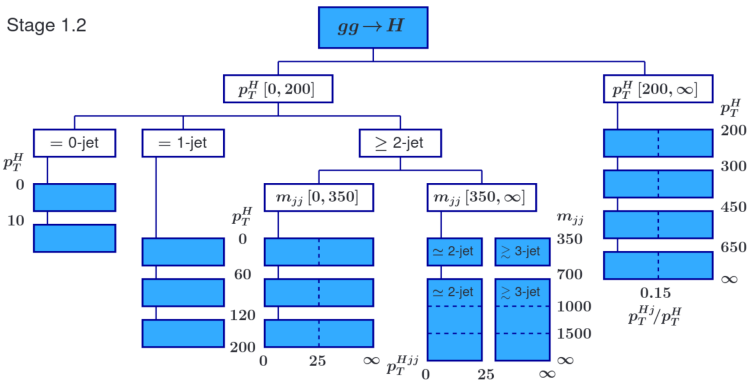


Figure 5.1: STXS stage 1.2 scheme for the ggF Higgs production channel. Blue boxes indicate the final categories that an event is placed in. Dashed lines indicate additional cuts that can be applied which result in the STXS stage 1.2 'fine binning' scheme.

5.2 Sensitivity

The sensitivity of an analysis is determined by performing a profile likelihood scan and determining the 1-σ interval*. The likelihood is calculated for the following scenario: how likely are we to observe the SM $\Delta\phi_{jj}$ distribution if the true distribution is S_c. Here, the SM distribution has $c_{H\tilde{G}} = 0$, and the S_c distribution can have any $c_{H\tilde{G}}$-value. The latter can be obtained through parametrisation as discussed previously. Both SM and S_c contain ggF and $t\bar{t}$ contributions. For a each c-value considered, $-2\ln\mathcal{L}(S_c)$ is calculated.

For each $\Delta\phi_{jj}$ bin, the probability of seeing the SM amount of events k_i, when expecting S_c amount of events m_i is calculated according to the Poisson distribution: $P(k_i, m_i)$. The total log likelihood is then:

$$\begin{aligned}
-2\ln\mathcal{L}(S_c) &= -2\ln\left(\prod_i P(k_i, m_i)\right) \\
&= -2\sum_i \ln\left(P(k_i, m_i)\right) \\
&= -2\sum_i \ln\left(\frac{m_i^{k_i} e^{-m_i}}{k_i!}\right) \\
&= -2\sum_i \left[k_i \ln(m_i) - m_i - \ln(k_i!)\right]
\end{aligned}$$

This can be further simplified by using the Stirling approximation

$$\ln(k!) = k\ln(k) - k + \frac{1}{2}\ln(2\pi k) + O(\frac{1}{k})$$

Often the first two terms are enough for a good approximation, but in this case the third term is added in order to avoid log likelihoods of 0 when $k_i = m_i$.

$$\begin{aligned}
-2\ln\mathcal{L}_m &= -2\sum_i \left[k_i \ln(m_i) - m_i - k_i \ln(k_i) + k_i - \frac{1}{2}\ln(2\pi k_i)\right] \\
&= -2\sum_i \left[k_i \ln\left(\frac{m_i}{k_i}\right) - m_i + k_i - \frac{1}{2}\ln(2\pi k_i)\right]
\end{aligned}$$

The 1σ interval is then determined by finding the values for which the following holds: $-2\ln\mathcal{L}(S_c) = \min(-2\ln\mathcal{L}(S_c)) + 1$, and is referred to

*Wilks' theorem is assumed to hold[21]

Version of March 2, 2025– Created March 2, 2025 - 15:18

as $\Delta\lambda$ from now on. The narrower this interval, the better the sensitivity. Additionally, the 2σ interval can be determined with $-2\ln\mathcal{L}(S_c) = \min(-2\ln\mathcal{L}(S_c)) + 4$. The minimum $-2\ln\mathcal{L}$ value is always at $c = 0$ because we are fitting to the SM distribution.

A simultaneous fit to multiple m_{jj} or p_T^H bins can be approximated with

$$\Delta\lambda_{simul} = \frac{1}{N}\sqrt{\sum_i^N \Delta\lambda_i^2}$$

Chapter 6

Simulations

Three types of samples were created for this analysis: SM ggF→HWW, EFT ggF→HWW and SM $t\bar{t}$ background. The EFT samples only contain the linear interaction term and were generated with $c_{H\tilde{G}} = 1$. This term does not make physical sense on its own, as it can have negative cross-sections, but must be combined with the SM ggF samples to obtain physical results. All samples were created at generator level and additionally passed through a detector simulation in order to obtain detector-level samples. All simulations were performed at the Run 3 centre-of-mass energy of 13.6 TeV.

6.1 Generator level

The ggF →HWW events were generated using `MadGraph5\_aMC@NLO` version 3.5.4 [5] and the `SMEFTsim` model version 3.0.2. Showering was added with `Pythia8` version 8.312. 130.000 SM samples were generated at NLO (next to leading order) precision. To ensure that additional jets, needed for $\Delta\phi_{jj}$ are generated, they need to be generated explicitly:

```
generate p p > h > l+ vl l- vl~ /z QED=4 QCD=2 SMHLOOP=1
add process p p > h > l+ vl l- vl~ jb /z  QED=4 QCD=3   SMHLOOP=1
add process p p > h > l+ vl l- vl~ jb jb /z QED=4 QCD=4   SMHLOOP=1
```

Here `jb` is defined as a gluon or a quark (excluding the top quark). There is no NLO level available yet for SMEFT operators, so these are at LO precision. 500.000 EFT ($\hat{O}_{H\tilde{G}}$) samples were generated for the linear interference term with $c_{H\tilde{G}} = 1$.

```
generate p p > h > l+ vl l- vl~ /z QED=4 QCD=2 NP<=1 \
    NP^2==1 SMHLOOP<=1 NPcHGtil<=1 NPcHGtil^2==1 SMHLOOP^2==1
add process p p > h > l+ vl l- vl~ jb /z  QED=4 QCD=3 \
    NP<=1 NP^2==1 SMHLOOP<=1 NPcHGtil<=1 NPcHGtil^2==1 SMHLOOP^2==1
add process p p > h > l+ vl l- vl~ jb jb /z QED=4 QCD=4 \
    NP<=1 NP^2==1 SMHLOOP<=1 NPcHGtil<=1 NPcHGtil^2==1 SMHLOOP^2==1
```

Lastly, 200.000 $t\bar{t}$ samples have been generated

```
generate p p > t t~ > b l+ vl b~ l- vl~
```

6.2 Detector level

The characteristics of the ATLAS detector were simulated using `Delphes` version 3.5.0, which models the effects of detector efficiency and resolution on observed data. For SM and EFT, the generator-level samples were used as Delphes input. For $t\bar{t}$, 10.000.000 new samples were generated with `POWHEG` and passed through Delphes in order to improve statistical significance. The default ATLAS configuration card, included in the Delphes distribution, which specifies these effects is representative of the ATLAS detector during Run 1. In order to better match the current state of the detector, three changes were made to the configuration card:

1. The b-Tagging efficiency was updated
2. The radius parameter R was updated
3. The number of saved branches was greatly reduced

b-Tagging efficiency: One of the simulated effects is the efficiency of correctly tagging a b-jet. This is the defining characteristic of the $t\bar{t}$ background and therefore greatly affects the background rejection efficiency. Figure 6.1 shows the efficiency function used in the standard ATLAS card which is dependent on the p_T of the jet. The inclusive (average) efficiency of the tagger is 70%, calculated over the range of 20 GeV to 250 GeV. This is below the 77%-efficiency working point of Run 2 `DLr1` b-tagger[9], shown in Figure 6.2. Furthermore, current b-taggers run with efficiencies of 85% percent.

For this reason, the b-tagging efficiency has been updated by matching the shape to the `DLr1` tagger and adding an overall scaling to reach an inclusive efficiency of 85%.

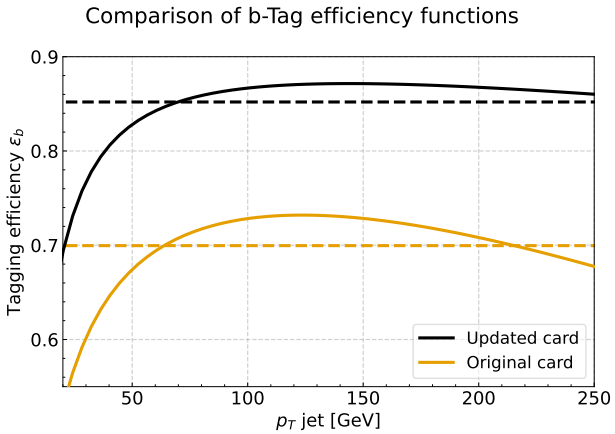


Figure 6.1: Comparison of the b-Tagging efficiency of the original ATLAS configuration card included in Delphes v. 3.5.0 and the updated card. Inclusive efficiency is calculated over the range 20 GeV to 250 GeV.

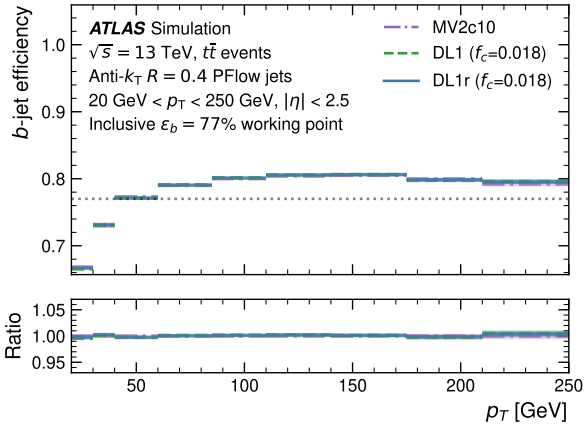


Figure 6.2: The 77%-efficiency working point of the *DL1r* b-tagger. Taken from [9], Figure 11a.

ParameterR: This is a parameter that is used in the anti-kt algorithm[7] which reconstructs the jets from detector data. Old value was given as `{Parameter}R=0.6` and has been updated to `ParameterR=0.4`.

Branches: By default, Delphes would save almost all information of an event. This resulted in big Delphes files (750 Mb for 10.000 events). If only the necessary branches are kept (`Jet`, `GenJet`, `Electron`, `Muon`, `Photon`, `MissingET`), the size reduces by a factor of 50 (15 Mb for 10.000 events). This allows us to generate the 10M samples for detector-level.

6.3 Normalisation and additional scaling

Samples were normalised to the expected Run 3 luminosity of $\mathcal{L} = 300 fb^{-1}$ in order to obtain the expected number of observed events:

$$N = \mathcal{L}\sigma_{eff} = \mathcal{L}\sum_i |\sigma_i| \frac{w_i}{\sum_i |w_j|}$$

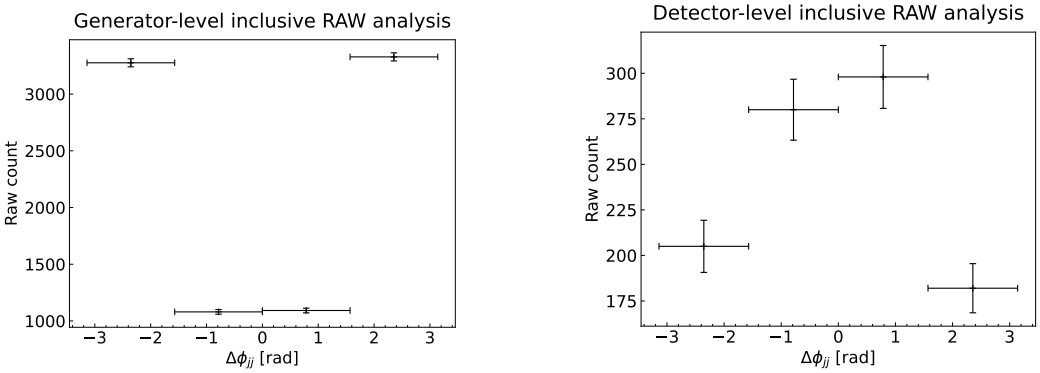
The cross-section σ_i and weight w_i are determined per event by the generation frameworks. At generator-level, σ_{eff} is calculated automatically in Rivet and results in a final cross-section that matches the generator values. At generator-level, this normalisation is done manually. w_i and σ_i are strictly positive for SM and $t\bar{t}$, but can be negative for EFT. In order to preserve the correct CP-odd shape of EFT, the absolute value of σ_i is used.
The $t\bar{t}$ σ_{eff} differs between generator- and detector-level since samples were generated by two different methods. It was chosen to scale the detector-level σ_{eff} to match generator-level, as this agrees better with [10]. Even though the same EFT samples are used for generator- and detector-level, their cross-sections differ. The reason for this change is so far not known. For consistency, EFT σ_{eff} is also scaled to match generator-level. An overview of σ_{eff} is provided in Table 6.1.

Process	Gen σ_{eff} [pb]	Det σ_{eff} [pb]
ggF→HWW SM	0.387	0.387
ggf→HWW EFT	1.19×10^{-3}	1.92×10^{-3}
$t\bar{t}$	23.4	91.6

Table 6.1: Effective cross-sections σ_{eff} for detector-level (Det) and generator-level (Gen).

Due to the low number of samples of generator-level $t\bar{t}$ and detector-level

SM*, their $\Delta\phi_{jj}$ distributions lose their symmetrical shape after passing event selection. Figure 6.3b shows the deterioration of the shape due to low sample count after passing through detector simulation. The results of the same analysis on the generator level are shown in Figure 6.3a. Here, no shape deterioration is seen. For this reason, the $\Delta\phi_{jj}$ distributions of these two samples were artificially created by scaling the $\Delta\phi_{jj}$-shape before event selection to the number of events after event selection and additional kinematic cuts.



(a) Generator level

(b) Generator level

Figure 6.3: Comparison of RAW event count of the inclusive analysis for generator and detector level. Both analyses were performed on the same generator-level samples.

*The event selection at detector-level for the ggF→HWW process is much more strict than at generator-level.

Chapter 7

Results

Analyses were performed using two different frameworks: the Rivet framework in C++ was used to analyse at generator-level, and `pyROOT` was used for detector-level. This was necessary due to the slightly different formats of the two types of samples.

7.1 Event selection

The normalised number of events left after each cut are summarised in Tables 7.1 and 7.2. This normalisation includes the additional rescaling for cross-section matching.

	ggF SM	ggF EFT	$t\bar{t}$
None	116039	356	7013478
OSDF leptons	57883	100	3509823
$\geq$ 2-jets	45501	118	3494706
b-veto	43275	123	48363
m_{ll} <55 GeV	25188	115	10731
$\Delta\phi_{ll} < 1.8$ rad	20575	44	9013

Table 7.1: Generator-level cutflow, $\sqrt{s} = 13.6$ TeV $\mathcal{L} = 136$ fb^{-1}

Both the SM and $t\bar{t}$ counts are a factor of 2 larger than the final event count of [10], but agree well on the SM/$t\bar{t}$ sample ratio. Comparing the SM-cutflows at generator and detector level reveals that both the OSDF leptons and $\Delta\phi_{ll} < 1.8$ cuts are much stricter at detector level. The reason for this

	ggF SM	ggF EFT	$t\bar{t}$
None	115984	356	7422822
OSDF leptons	12514	42	623019
$\geq$ 2-jets	12514	42	558491
b-veto	11462	61	42714
m_{ll} <55 GeV	7439	-44	33908
$\Delta\phi_{ll} < 1.8$ rad	871	-35	12183

Table 7.2: Detector-level cutflow, $\sqrt{s} = 13.6$ TeV $\mathcal{L} = 136$ fb^{-1}

is most likely a detector-simulation effect, but it has not been attempted to discover which setting is responsible. The EFT signal is the pure linear $c_{H\tilde{G}}$ contribution. The cross-sections of linear terms of SMEFT operators may be negative, which results in a negative event count for the EFT signal

7.2 Kinematic cuts

7.2.1 Inclusive analysis

The inclusive analysis does not include any additional cuts after event selection and is used as a control. Other analyses should improve upon the inclusive sensitivity in order to be useful. The resulting $\Delta\phi_{jj}$ distributions are shown in Figure 7.1 (generator level) and Figure 7.2 (detector level). As stated previously, fewer ggF events are left after event selection for detector-level, leading to a significant decrease in sensitivity. The $\Delta\lambda$ interval increases by a factor of 10; from $\Delta\lambda = 0.062$ at generator level, to $\Delta\lambda = 0.668$ at detector level.

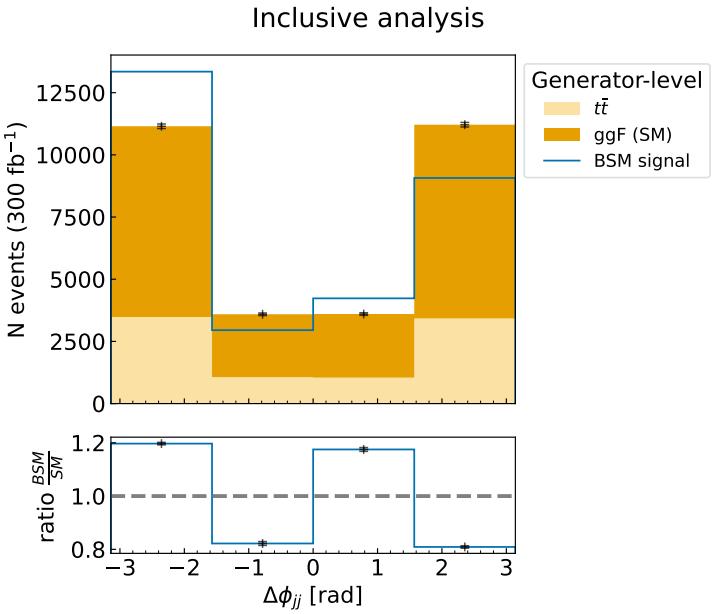


Figure 7.1: Data visualisation of the inclusive analysis at generator level. The upper plot shows the makeup of the final distribution. The ratio between the BSM distribution (ggF SM + ggf EFT + $t\bar{t}$ SM) and the SM distribution (ggF SM + $t\bar{t}$ SM) is shown in the bottom plot.

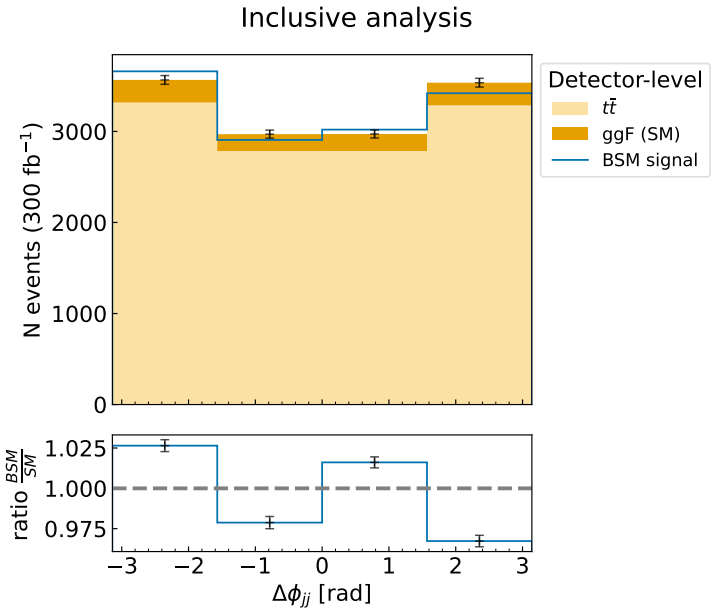
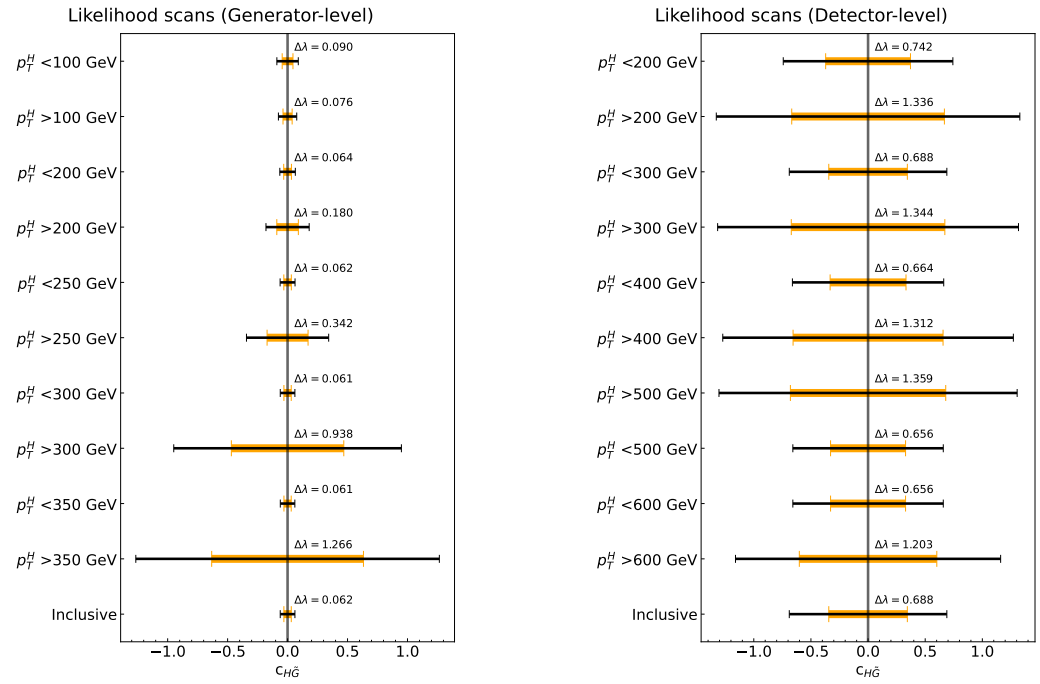


Figure 7.2: Sample distributions of the inclusive analysis at detector level.

Version of March 2, 2025– Created March 2, 2025 - 15:18

7.2.2 Binary cuts

The simplest kinematic cut added is the binary cut. All events are split into two categories: either below or above a certain threshold. Figures 7.3a and 7.3b summarise the results of these cuts. For both generator and detector level, removing events with high p_T^H is beneficial. At detector level, the greatest improvement, with respect to the inclusive analysis, is achieved by $p_T^H < 500$ and $p_T^H < 600$. However. $\Delta\lambda$ is only decreased by 4.7%, so the change is not significant. Cuts on m_{jj} are much more effective. The best result (for detector level) is achieved by the $m_{jj} > 200$ GeV cut which decreases $\Delta\lambda$ by 27% to $\Delta\lambda = 0.504$. Again, a summary is shown in Figures 7.5a and 7.5b. Figure 7.4 visualises the difference between a good cut ($p_T^H < 500$ GeV) and a bad cut ($p_T^H > 500$ GeV): there are far fewer events left after the bad cut, and the shape of the ratio-plot has deteriorated.



(a) Generator level (b) Detector level

Figure 7.3: Overview of the sensitivities of different p_T^H binary cuts for detector and generator level. The orange and black intervals correspond to 1σ ($\Delta\lambda$) and 2σ intervals, respectively.

Version of March 2, 2025– Created March 2, 2025 - 15:18

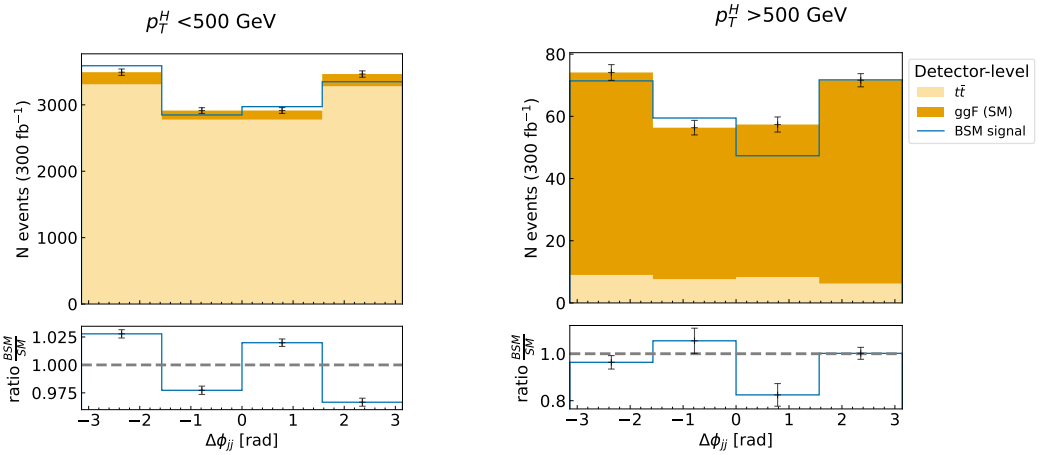
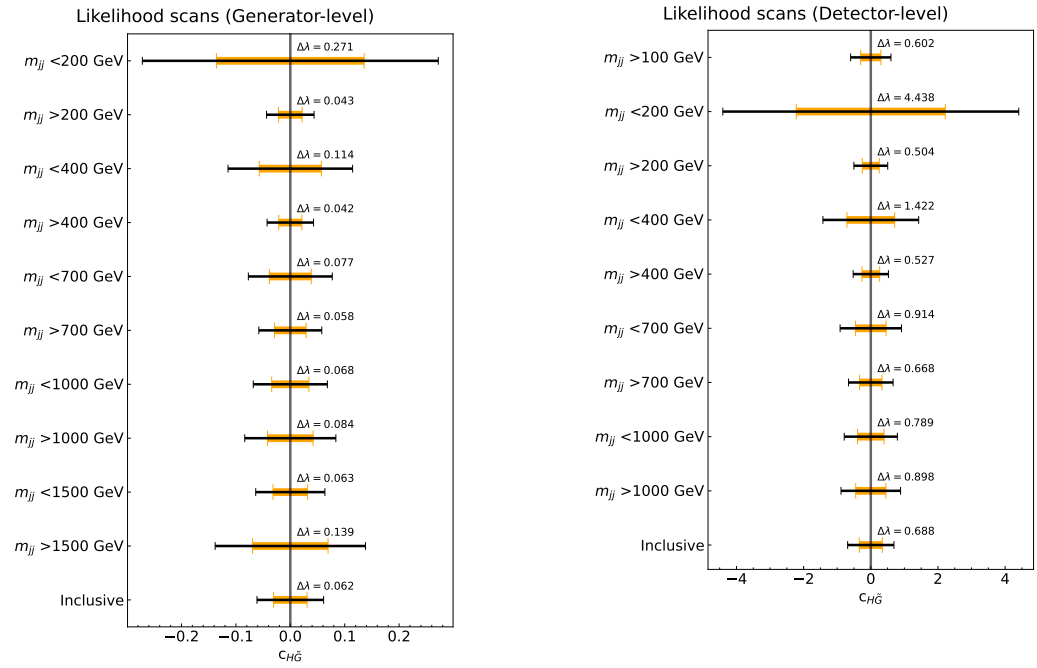


Figure 7.4: Sample distributions of two binary p_T^H cuts at detector level.



(a) Generator level (b) Detector level

Figure 7.5: Overview of the sensitivities of different m_{jj} binary cuts for detector and generator level.

7.2.3 Combination cuts

By combining the two best binary cuts ($p_T^H < 500$ GeV, $m_{jj} > 200$ GeV), it is expected that sensitivity will increase further. Small variations were added

in order to probe a slightly larger set of cuts, however this had little effect. Instead, all combination cuts around the optimal values result in practically the same sensitivity, as seen in Figure 7.6. In order to keep maximise the number of events that pass the selection, the $p_T^H < 600$ GeV + $m_{jj} > 200$ GeV combination is taken to be optimal with a $\Delta\lambda$ of 0.469, a decrease of 32% with respect to inclusive $\Delta\lambda$.

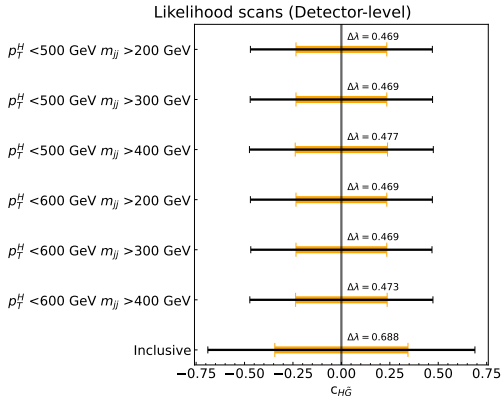


Figure 7.6: Sensitivities of the combination cuts analysis.

7.2.4 Multi-bin

Splitting the domain of a single parameter into multiple bins does not increase the sensitivity in one single bin compared to inclusive, see Figure 7.7. But, approximating a simultaneous fit (as described in Section 5.2) does improve the result: the three p_T^H bins yield a $\Delta\lambda_{simul}$ of 0.757, and the three best m_{jj} bins (ignoring $m_{jj} < 200$ GeV) yield $\Delta\lambda_{simul}$ 0.502.

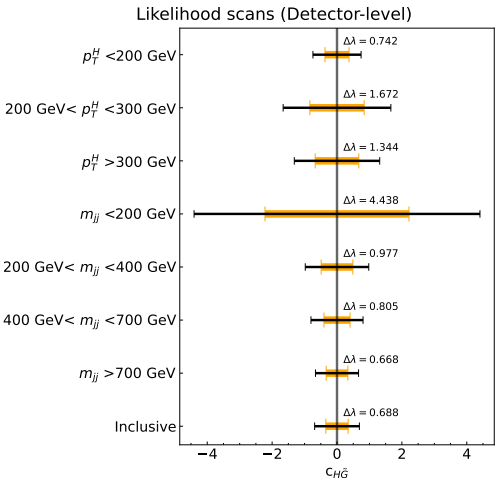


Figure 7.7: Sensitivities of the multi-bin analysis at detector level.

7.3 Detector effects

A small study was done on the loss of sensitivity as a result of detector simulation. Due to the finite accuracy of the detector, events will sometimes land in bins adjacent to the truth-bin. This leads to smearing in parameters, which degrades the overall resolution. The degree to which the jet-parameters are change after passing through the ATLAS detector simulation is shown in Figure 7.8. The smearing of parameters was calculated by comparing detector-level jets with the closest (minimum ΔR) generator-level jet: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The resulting distribution looks like a normal distribution, so the standard error can be calculated from its Full Width at Half Max (FWHM): $\sigma = FWHM/2.355$. These are shown in Table 7.3. For the primary observable, $\Delta\phi_{jj}$, ϕ_{jet}-smearing is most relevant. From the small relative σ of 2.0% we can conclude that the ATLAS simulation does not significantly affect the reconstruction of $\Delta\phi_{jj}$.
Another effect of detector simulation is the decrease of observed leptons. From the cutflow comparison of generator- and detector-level, we can see that the OSDF lepton selection becomes much stricter in at detector-level.

Parameter	p_T	ϕ	η
σ [%]	6.9	2.0	1.4

Table 7.3: Detector resolutions of three jet-parameters.

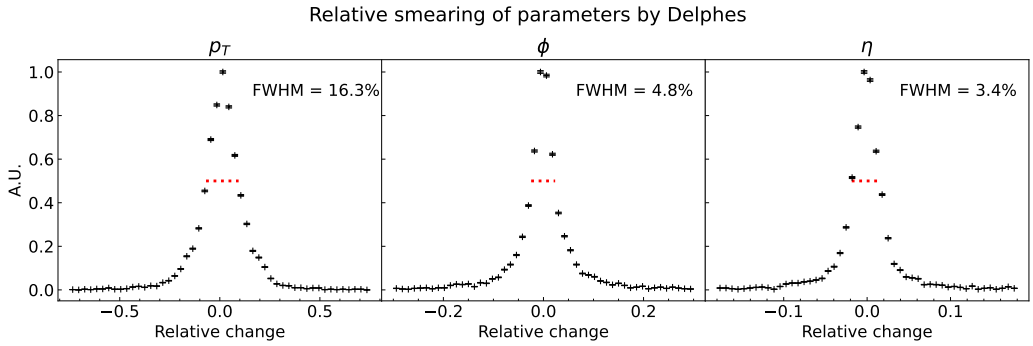


Figure 7.8: Smearing of three jet-parameters by Delphes detector simulation of ATLAS.

Chapter 8

Discussion

Comparing the three approaches of kinematic cutting in phase space (binary, combination and multi-bin), it can be concluded that the sensitivity to CP-odd effects of $\hat{O}_{H\tilde{G}}$ is best increased by cutting on m_{jj}. In particular, variations on the criterion $m_{jj} > 200$ GeV seem to be best. An additional cut ($p_T^H < 600$ GeV) to remove high-p_T^H events increases the sensitivity further, but only by 7%. Only cutting on p_T^H does not result in notable better sensitivity when compared to the inclusive sensitivity.

This thesis is a small-scale study which is not fully representative of a full analysis, but has the advantage of being fast. Still, there are small improvements that can be made in order to obtain more accurate results. Firstly, the effects of detector simulation on leptons should be better understood. It has been noted that the configuration card, which was used to model the detector, lead to a substantial decrease in detected leptons. The degree to which this effect is representative of the current state of the ATLAS detector is unknown, as is the exact parameter responsible for this decrease. Furthermore, the research would benefit from a larger sample of SM ggF events. It was necessary to artificially create the SM ggF distribution for detector-level, which likely improved the sensitivity, but decreased reliability.

At the writing of this thesis, an effort is underway to measure Higgs boson production cross-sections in the ggF$\rightarrow$HWW channel. This research will also contain a study on CP-odd SMEFT operators, among which $\hat{O}_{H\tilde{G}}$. Currently, $\hat{O}_{H\tilde{G}}$-effects are being isolated by performing p_T^H cuts, which this thesis has shown to be less efficient than using m_{jj} cuts. In the future, it would be interesting to explore the feasibility of using m_{jj} as the primary parameter for cuts in a full-scale analysis.

Acknowledgements

The months I've spent at the Nikhef ATLAS corridor were immensely vivid and fulfilling. Not only did I get to work on a cool project and learn a lot about a field I had little previous knowledge about, I got to do this with a wonderful group of people. And it's the people who I want to thank here.

First of all, thank you Lydia for giving the girl from Leiden the chance to work with you. I was immediately struck by your kindness and enthusiasm which made me so happy and excited. You are a true inspiration for me as a physics, supervisor and role-model. Ivo, thank you for making the project possible and always managing to make me laugh with your jokes.

Of course I have to thank you Oliver for your incredible help. I am so grateful for the effort you have put into helping me improve my work. From talking me through whatever physics thing I couldn't understand to the detailed feedback on the thesis and presentation, I would always come away having learned something new and encouraged to continue improving. Too bad you're too good at ping-pong.

Meetings on Monday mornings were never dull in the EFT group. Even though the problems were often niche and complicated, everyone still took time to explain what was going on so that everyone could follow along. Wouter, thank you for the mini-lectures on statistics and the fun stories. Andrea, thanks for always helping me with the slightly random problems throughout the project that I couldn't solve myself. And Robin, thank you for taking time to help me to understand STXS (& more) even though you were probably swamped with work.

Every day at Nikhef would have me laughing at some point. Whether it's ping-pong, Dutch coffee or the borrels on Friday, the atmosphere is always incredibly welcoming and fun. Alejandro, Ali, Ambre, Andrea & Andrea, Arturo, Charlotte, Duncan, Dylan, Karel, Jurjan, Lisa, Liza, Luca, Maria, Marten, Oliver, Osama, Petja, Pranati, Robin, Slava, Zef and Zhouran, thanks for a great time. Though my stay was not long, I'm grateful to have gotten to know you all and thinking back makes me smile.

Danke Lisa for putting me in that German group chat. I feel like you took me under your wing and supported me mentally throughout the whole project. I am so happy that I've gotten to know you and I have no doubts that you'll do an amazing job with your PhD and beyond.

Лиза, спасибо тебе за твою теплую дружбу! Мне всегда было очень приятно когда ты проходя мимо заглядывала ко мне поболтать, и я

надеюсь ещё увидимся или на квизе, или просто так.

My fellow masters, Alejandro, Arturo, Charlotte, Duncan, Jurjan and Slava, it has been great fun to get to spend time with you all. I wish you the best with your future endeavours, wherever they may lead you. Jurjan, bedankt voor de leuke natuurkundige gesprekken tijdens de lunch. Ze hebben me geholpen mijn eigen blik te verbreden. Char, ik ben blij dat we de os pab praktijken hebben kunnen voortzetten. Altijd wanneer je naar Amsterdam kwam was het voor mij een leuke dag. Niet alleen inspireer je mij met je doorzetvermogen en je onredelijke smashes, je bent ook een hele goede vriendin.

A big source of inspiration and knowledge were the two PhD theses permanently open on my desk. Thank you Federica Pasquali and Rahul Balasubramanian for writing great theses.

Als laatste wil ik een klein bedankje geven aan mijn dierbare PSW die ondertussen een beetje verspreid is over Nederland en Europa. Jullie zijn onvervangbaar.

Bibliography

[1] Beam. https://lhc-machine-outreach.web.cern.ch/beam.htm. Accessed: 2025-01-29.

[2] Observation of a new boson with a mass near 125 GeV. Technical report, CERN, Geneva, 2012.

[3] Observation of an Excess of Events in the Search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Technical report, CERN, Geneva, 2012. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-093.

[4] Measurement of the higgs boson width and evidence of its off-shell contributions to zz production. Nature physics, 18(11):1329–1334, 2022.

[5] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. Journal of High Energy Physics, 2014(7), July 2014.

[6] Rahul Balasubramanian et al. Effectively going beyond the standard model: Combined interpretations of Higgs boson and electroweak measurements with the ATLAS experiment. 2024.

[7] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. Journal of High Energy Physics, 2008(04):063–063, April 2008.

[8] ATLAS collaboration et al. 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. arXiv preprint arXiv:2308.04775, 2023.

[9] 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.

[10] The ATLAS Collaboration. Measurements of higgs boson production by gluon-gluon fusion and vector-boson fusion using $h \to wW^* \to e\nu\mu\nu$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the atlas detector. Phys. Rev. D, 108:032005, Aug 2023.

[11] Cush. Standard model of elementary particles. Wikimedia Commons, Public Domain, 2021. Accessed: 2025-01-29.

[12] Adam Falkowski. Lectures on smeft. The European Physical Journal C, 83(7):1–41, 2023.

[13] Richard Phillips Feynman. Space-time approach to quantum electrodynamics. In Quantum Electrodynamics, pages 178–198. CRC Press, 2018.

[14] Sascha Mehlhase. ATLAS detector slice (and particle visualisations). 2021.

[15] Theodoros Nakas. Searching for Localized Black-Hole solutions in Brane-World models. PhD thesis, 07 2017.

[16] S. Navas et al. Review of particle physics. Phys. Rev. D, 110(3):030001, 2024.

[17] I. Neutelings. Cms coordinate system. `https://tikz.net/axis3d_cms/`.

[18] J. Riebesell. Higgs potential. `https://tikz.net/higgs-potential/`.

[19] Andrei D Sakharov. Violation of cp-invariance, c-asymmetry, and baryon asymmetry of the universe. In In The Intermissions... Collected Works on Research into the Essentials of Theoretical Physics in Russian Federal Nuclear Center, Arzamas-16, pages 84–87. World Scientific.

[20] Stephen Nicholas Swatman and Rahul Balasubramanian. Visualising eft feynman vertices. `https://rahulb.web.cern.ch/SMEFTviz.html`. Accessed: 2025-02-23.

[21] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. The Annals of Mathematical Statistics, 9(1):60 – 62, 1938.

[22] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. Phys. Rev., 105:1413–1415, Feb 1957.

[23] Hongtao Yang. Status of STXS (stage 1.2) and differential measurements. 2020.