GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Untersuchung der CP-Natur des Higgs-Bosons in VBF-$H \to \tau\tau$-Ereignissen bei $\sqrt{s} = 13$ TeV

Investigation of the CP Nature of the Higgs Boson in VBF $H \to \tau\tau$ Events at $\sqrt{s} = 13$ TeV

prepared by

Serhat Ördek

from Celle

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-MSc-2017/01
Thesis period:	22nd February 2016 until 26th January 2017
First referee:	Prof. Dr. Stanley Lai
Second referee:	Prof. Dr. Ariane Frey

Contents

1 Introduction

In 2012, ATLAS and CMS, the two largest experiments located at CERN in Geneva, independently reported the discovery of a new particle [1, 2]. This particle is assumed to be the Higgs boson, whose existence was predicted almost 50 years before [3–5]. All experimental evidence collected so far supports this conjecture, but some of its properties are not yet measured very precisely, leaving room for speculation. This thesis will describe studies of the "Optimal Observable" method used to measure the CP nature of the new particle based on its production properties in vector boson fusion.

The optimal observable method can distinguish between the cases of a CP-even, CP-odd or CP-mixed coupling of the new particle to weak vector bosons. An even coupling would be in agreement with the Standard Model, meaning that the particle could indeed be the predicted Higgs boson. A pure CP-odd coupling is not likely at this point, as it was ruled out at more than 99.9% confidence level by experimental observations [6, 7]. The CP-mixed case, however, is not as strongly constrained by observations, and it might be the most interesting one because it would mean a violation of CP symmetry. The existence of such processes is one of the Sakharov conditions [8], so in this scenario, the Higgs boson's couplings could be a reason for the observed baryon asymmetry in our universe.

In Chapter 2, the theoretical fundamentals of this analysis are described, such as the Standard Model of particle physics, vector boson fusion (VBF) Higgs production at the Large Hadron Collider (LHC) and a motivation of the optimal observable method. Chapter 3 covers the experimental setup of the ATLAS detector at the LHC. Moreover, the methods of data acquisition used in the ATLAS collaboration, concepts of hadron collider phenomenology and analysis techniques for simulated collision data are discussed. Chapter 4 is the first part of the analysis done in context of this thesis. It includes studies on the event sample that is used for a large part of this work and the validation of a method to reweight Monte Carlo data in order to simulate different CP mixing scenarios at leading order. In chapter 5, modifications to make the reweighting method applicable to next-to-leading order (NLO) events are presented. Using the best performing modification, the reweighting is then validated at NLO. Chapter 6 describes an attempt to improve the accuracy of the optimal observable calculation by applying a scaling factor to

the measured momenta of the jets. In chapter 7, the results of the analysis are summarised and an outlook on the future use of the tested methods is given.

2 Theoretical Overview

2.1 The Standard Model of Particle Physics

2.1.1 The Particle Content of the Standard Model

The Standard Model (SM) [9–11] is the most widely accepted theory in particle physics and also the theory tested in this analysis. It is a quantum field theory describing interactions between elementary particles. The SM particles are shown in figure 2.1, and with these particles the SM describes strong and weak nuclear force as well as electromagnetism, but not gravitation. These forces are incorporated in the SM by requiring local gauge invariance with respect to certain symmetry groups. These groups are $SU(3)$ for the strong force, and $SU(2) \times U(1)$ for weak and electromagnetic force.
The resulting symmetry group is often denoted as $SU(3)_C \times SU(2)_L \times U(1)_Y$, and it has 12 generators, corresponding to 12 gauge bosons in the theory. These are the eight gluons that couple to particles with colour charge C and mediate the strong force, the three

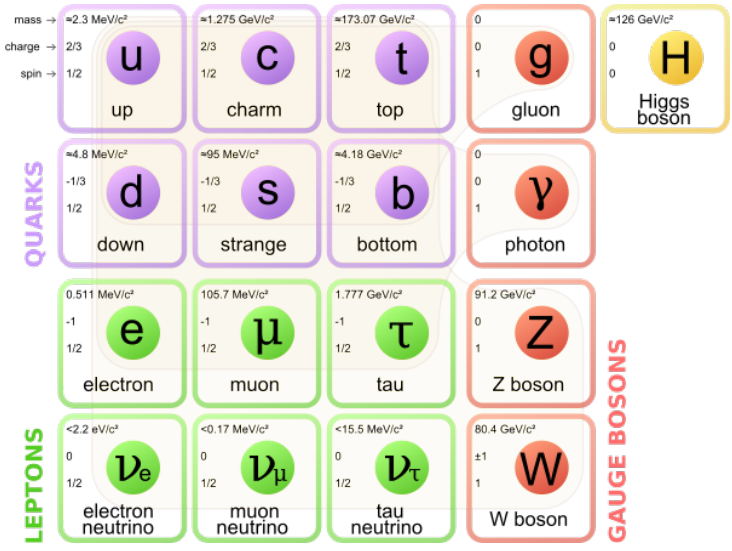

Figure 2.1: Particle content of the SM. The fermions on the left, divided into quarks and leptons, build up matter, while the bosons on the right carry the forces and allow interactions.

mediators of the weak force W^+, W^- and Z^0 and the photon γ responsible for electromagnetism, which couples to electric charge Q. The index L means that only left-handed particles and right-handed antiparticles are subject to the $SU(2)$ symmetry.
The charges associated to the gauge groups $SU(2)$ and $U(1)$ are the weak isospin T and the weak hypercharge Y, but these charges do not directly correspond to the charges to which the gauge bosons W, Z and γ couple. Instead, they are linear combinations of each other, such that the electric charge can be expressed as $Q = T_3 - \frac{Y}{2}$, with T_3 being the third component of the weak isospin. Consequently, weak and electromagnetic interactions cannot be viewed separately at this fundamental level, which is why they are often referred to in combination as electroweak interactions.
The $U(1)$ and $SU(2)$ gauge symmetries are part of the reason why the Higgs boson was first introduced. When imposing a local gauge symmetry on the SM Lagrangian, the bosons mediating the force must be massless. The reason is that a term representing the mass of the gauge bosons is not gauge invariant, meaning that it would break the symmetry imposed. Therefore, the measurement of non-zero W and Z boson masses [12, 13] is not compatible with the required gauge symmetry.

2.1.2 Electroweak Symmetry Breaking

To solve this apparent contradiction, a self-interacting complex scalar doublet field Φ can be introduced to the theory. The doublet and its potential V can be written as

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}$$
$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.1}$$

with the parameter μ of mass dimension 1 and the dimensionless self-coupling strength λ. For a realistic theory, the vacuum state of the theory needs to be stable, meaning that the potential must be bounded from below. Consequently, the self-coupling parameter λ needs to be positive. For the sign of μ^2, we consider two cases. For $\mu^2 \leq 0$, the potential has a single minimum at 0. If $\mu^2 > 0$, however, the potential takes on the shape depicted in figure 2.2. In this case, there is a circle of minima with radius $\sqrt{(\Phi^\dagger \Phi)_{\text{min}}} = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$ around the origin, corresponding to infinitely many equivalent vacuum states. The parameter v is called the vacuum expectation value of the field.
Such a potential is symmetric with respect to rotations around the origin, since $V(\mathrm{e}^{i\theta}\Phi) = V(\Phi)$ for any $\theta \in \mathbb{R}$, as can be seen in equation 2.1. The important feature for solving the problem of the gauge boson masses is that it allows for "spontaneous symmetry breaking",

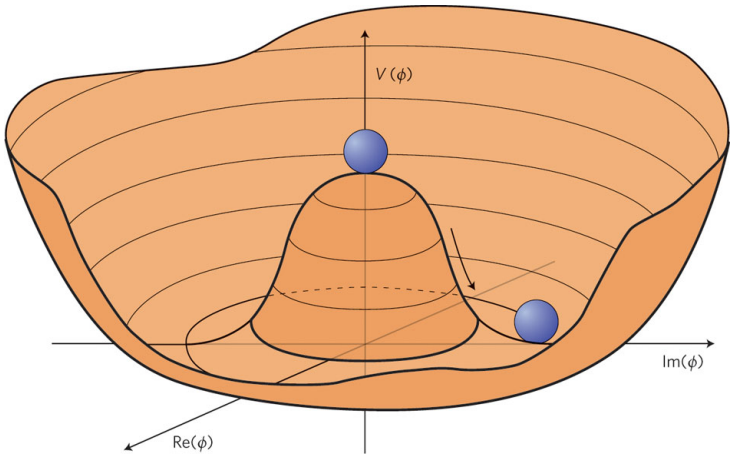

Figure 2.2: Image of the Higgs potential. It takes on its lowest value on a circle around the origin.

meaning that regardless of the phase space region in which one starts, the system will eventually end up in one of the vacuum states in the circle of minima of its potential, about which the potential is not rotationally invariant.
Let this scalar doublet now be coupled to the gauge fields W^i, $i \in \{1,2,3\}$, and B of the $SU(2)$ and $U(1)$ symmetries. Then the Lagrangian of this scalar doublet and the covariant derivative D_μ contained in it take the form

$$D_\mu = \partial_\mu + \mathrm{i}\frac{g}{2}\vec{\tau}\cdot\vec{W}_\mu + \mathrm{i}\frac{g'}{2}B_\mu$$
$$\mathcal{L} = (D_\mu\Phi^\dagger)(D^\mu\Phi) - V(\Phi)\ .$$

The new parameters appearing in the covariant derivative are the dimensionless coupling parameters g and g' and the vector $\vec{\tau}$ of Pauli matrices. The latter are defined as follows, using the Kronecker δ_{ij} factor.

$$\tau_j = \begin{pmatrix} \delta_{j3} & \delta_{j1} - \mathrm{i}\delta_{j2} \\ \delta_{j1} + \mathrm{i}\delta_{j2} & \delta_{j3} \end{pmatrix}$$

To determine this theory's particle spectrum, one can consider small excitations around the vacuum state. One can choose the state $\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ out of the infinitely many possible

vacua, so the field Φ can be rewritten as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + \mathrm{i}\phi_2 \\ \phi_3 + \mathrm{i}\phi_4 \end{pmatrix}$$

with fields ϕ_1, ϕ_2 and ϕ_4 that have vanishing vacuum expectation values. With the gauge freedom one has, one can choose $\phi_1 = \phi_2 = \phi_4 = 0$, the so-called unitary gauge [14]. Thus, one ends up with $\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \phi_3 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}$ with a field h that has a vanishing vacuum expectation value.

The goal of this calculation is to find mass terms in the kinetic term of the doublet Φ, defined as the sum of its Lagrangian and potential. If it contains mass terms of fields, they must be quadratic in these fields and not contain any derivatives. Therefore, all terms that do not meet these criteria will be omitted in the following. Examining the remaining terms makes it apparent that one of the mass eigenstates of the theory is defined by $Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W$, where the weak mixing angle θ_W is defined by $\tan(\theta_W) = \frac{g'}{g}$. The orthogonal mixed state $A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W$ does not have a mass term.

Moreover, one finds that the physical states of this theory are not W^1 and W^2 but W^+ and W^- with $W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp W^2)$. With these definitions, the remainder of the kinetic term is equal to

$$\frac{v^2}{8} \left(g^2 (W^{+\mu} W_\mu^+ + W^{-\mu} W_\mu^-) + (g^2 + g'^2) Z^\mu Z_\mu \right) .$$

From this result, it is apparent that, starting from the four massless gauge fields W^1, W^2, W^3 and B, by coupling them to a massive scalar doublet, one arrives at a situation with three massive vector bosons W^+, W^-, Z and one massless vector boson A.

The masses of these bosons are $m_W = \frac{v}{2} g$ for the W bosons and $m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}$ for the Z boson. This electroweak symmetry breaking explains how the W and Z bosons, despite being gauge bosons, can have mass terms that do not break the $SU(2)$ gauge invariance needed in the theory. In a theory where this mechanism is realised, however, there must also be a massive scalar particle, the remaining degree of freedom h of the complex doublet Φ. This scalar particle is the Higgs boson, and after the non-zero masses of W and Z bosons were confirmed in 1983, its existence was already widely accepted, although it had not been observed.

A similar theoretical problem arises when introducing fermion masses. The mass term for a fermionic field ψ with the mass m_ψ, when separated into the different helicity states,

can be written as

$$-m_\psi\overline{\psi}\psi = -m_\psi(\overline{\psi}_R + \overline{\psi}_L)(\psi_R + \psi_L) = -m_\psi(\overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R) \ .$$

The $SU(2)$ symmetry imposed on the SM Lagrangian only applies to left-handed and not to right-handed fields. Thus, the expressions $\overline{\psi}_R\psi_L$ and $\overline{\psi}_L\psi_R$ are not invariant under gauge transformations, which forbids the presence of these mass terms in the SM Lagrangian. By allowing Yukawa couplings of fermions to the doublet Φ, in the unitary gauge and expanded around $\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ as before, spontaneous symmetry breaking again results in mass generation for fermions. For a charged lepton ℓ and the corresponding neutrino ν_ℓ, the coupling term is

$$-\frac{\lambda_\ell}{\sqrt{2}}\left(\begin{pmatrix}\overline{\nu}_\ell & \overline{\ell}\end{pmatrix}_L \begin{pmatrix} 0 \\ v+h \end{pmatrix} \ell_R + \overline{\ell}_R \begin{pmatrix} 0 & v+h \end{pmatrix} \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L\right) = -\frac{\lambda_\ell v}{\sqrt{2}}\overline{\ell}\ell - \frac{\lambda_\ell}{\sqrt{2}}h\overline{\ell}\ell$$

with the dimensionless coupling parameter λ_ℓ. The lepton ℓ therefore gains the mass $\frac{\lambda_\ell v}{\sqrt{2}}$, while the neutrino gains no mass and does not couple to the Higgs boson. In the case of quark masses, the calculation is similar but more complicated, and will not be covered in this report.

2.1.3 The Higgs Boson Discovery

The discovery of what is now assumed to be the Higgs boson was announced in July 2012 by the two LHC experiments ATLAS and CMS. These experiments took proton-proton collision data at a center-of-mass energy of $\sqrt{s} = 7$ TeV until 2011 and at 8 TeV in 2012. Since the SM prediction for the lifetime of the Higgs boson is very short, it cannot be directly observed. Therefore, the goal of the experiments is to observe the products of its decay and to reconstruct the Higgs boson from its decay products. But not every allowed decay mode of the Higgs boson in the SM can be effectively used to find evidence of its existence, since the experimental signatures in the detectors can be hard to distinguish from background events.

The decay channels that contributed to the Higgs boson discovery are $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$. Their branching ratios along with that of $H \to \tau\tau$ are listed in table 2.1. Since the strengths of couplings between the Higgs field and other particles is proportional to their mass or their squared mass, there is no direct coupling of the Higgs to the massless photons γ. Nevertheless, there is a non-zero branching ratio for $H \to \gamma\gamma$. The reason is that these decays can still occur indirectly by inserting inter-

Higgs boson decay products	$\gamma\gamma$	$\ell\ell\ell\ell$	$e\nu\mu\nu$	$\tau\tau$
Branching ratio	$2.27 \cdot 10^{-3}$	$1.24 \cdot 10^{-4}$	$5.04 \cdot 10^{-3}$	$6.27 \cdot 10^{-2}$

Table 2.1: Selection of branching ratios of a 125 GeV SM Higgs boson. The listed decay channels are the most relevant for this analysis and the Higgs boson discovery. The symbol ℓ stands for electrons or muons, but not for τ leptons. Source: [15].

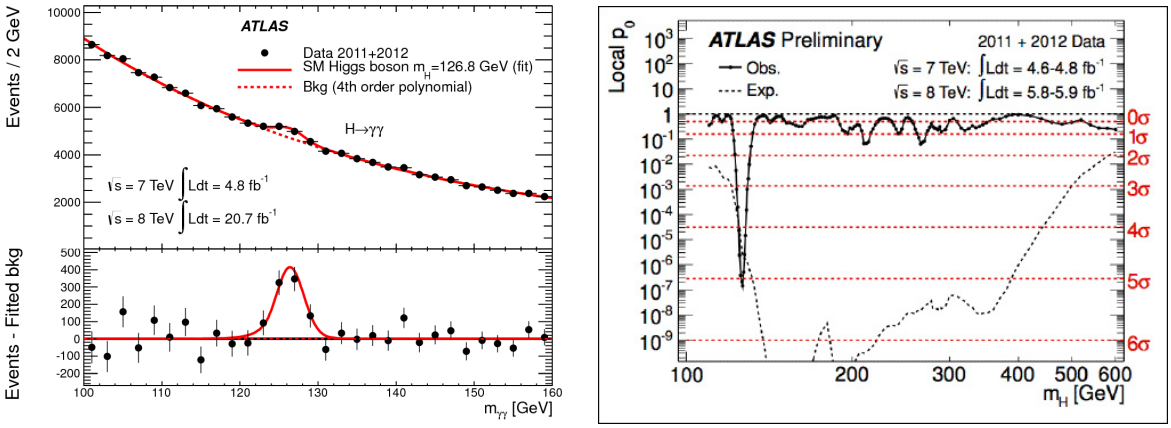

Figure 2.3: Left: Diphoton invariant mass distribution measured at ATLAS. At about 126 GeV, a peak above the background prediction can be seen, indicating the production of a new particle. Right: p-values for the background-only hypothesis with respect to the hypothesis of background + the signal of a SM-like Higgs with different masses. The minimum is around 126 GeV. Source: [1].

mediate loops containing heavy particles such as top quarks or W bosons that themselves couple to photons.

The general strategy was to identify the events in which the searched process might have taken place and detect the final state particles. By measuring their energy and momenta, one can calculate their combined invariant mass. The invariant mass distribution measured by the ATLAS collaboration in the $H \to \gamma\gamma$ channel, which was the most sensitive one at the time of the discovery, can be seen in figure 2.3(left). The peak at around 126 GeV implies that there is resonant Higgs boson production in this mass region.

The right part of figure 2.3 shows the significance of the result obtained by ATLAS for the combined search over all channels. The local p-value shown is the probability for only statistical fluctuations of non-Higgs events to produce invariant mass spectra of the tested final states that look more like the spectra caused by production of a SM Higgs boson of the mass given on the abscissa than the data.

For a 126 GeV SM Higgs boson, this is particularly unlikely, and therefore the published result was that a new particle around this mass was discovered. Whether or not this is the

Hypothesis	0_h^+	0^-	2^+	2^+ quark-enhanced	2^+ gluon-enhanced
$10^5 \cdot p$-value	7.1	< 3.1	4.3	7.3	4.3

Table 2.2: Alternative spin/parity hypotheses tested by the ATLAS collaboration with their p-values tested against the SM prediction of a 0^+ Higgs particle. The "basis" of the hypothesis description gives the spin, while the "exponent" is the eigenvalue of the state with respect to the parity operator. The hypothesis "0_h^+" describes non-SM CP-even couplings.

Higgs boson with properties as predicted by the SM is still an ongoing research question. Open questions include: whether the couplings of the new particle to other particles are proportional to their masses as predicted for the Higgs boson, and whether it truly is a spin-0 particle with positive parity and CP eigenvalues. The focus of this report will be on a method to test whether the particle is a CP eigenstate.

2.1.4 Status of Higgs Boson Spin and CP Measurements

Measurements for the particle's spin and CP properties have been conducted in the channels $H \to \gamma\gamma$, $H \to ZZ^* \to \ell\ell\ell\ell$, $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\ell\nu$ with $\ell \in \{e, \mu\}$ at ATLAS and CMS [6, 7]. Different approaches are used by the collaborations to analyse their data. While ATLAS uses an effective field theory (EFT) approach to characterise the Higgs couplings, CMS uses an anomalous couplings approach. The EFT approach will be presented in more detail in section 2.2.2. The goal of these analyses is to draw inferences on the spin and CP properties of the Higgs boson from the spin correlations of its decay products. This is done by considering different spin/CP scenarios for the particle and comparing kinematic distributions found in data to the predictions of the SM prediction and the alternative hypothesis.
Some of the alternative hypotheses are listed in table 2.2, along with the resulting p-values of the test against the SM, which are the probabilities for the non-SM hypotheses to produce kinematic distributions that look more SM-like than the data. For the spin-2 scenarios, "enhanced" means that hypotheses are tested in which the new particle couples more strongly to gluons than to quarks ("gluon-enhanced") or vice versa ("quark-enhanced"). All the tested scenarios are excluded at more than 99.9% confidence level in favour of the SM spin-0 Higgs boson hypothesis with positive parity. Spin-1 hypotheses are not listed because of the Landau-Yang theorem [16, 17] which states that it is impossible for massive spin-1 particles of neutral electric charge to decay into two photons. Since this was one of the discovery decay channels of the new particle, spin-1 scenarios can be excluded. The CMS collaboration used the same decay channels for their spin/parity analysis [7].

In addition to the models listed in table 2.2, hypotheses with the new particle being a superposition of different parity-states with the same spin were also tested at both ATLAS and CMS, with different mixing angles. These hypotheses imply CP violation, and testing their compatibility with the data is the applied strategy to check for CP invariance of the Higgs couplings to vector bosons. Listing and describing all of the slightly different tested hypotheses is beyond the scope of this section. The conclusion of these analyses is that all tested spin-1 and spin-2 scenarios are excluded at 99% confidence level or higher, and that if the new particle is a C and CP eigenstate, it is of spin 0 and has parity and charge conjugation eigenvalue +1, like the SM Higgs boson.
However, at this point neither ATLAS nor CMS can rule out the possibility that the new resonance is a CP-violating state, which is the case if it is a superposition of two spin zero states with different CP eigenvalues. To test whether the couplings of the new particle violate CP conservation, ATLAS and CMS use different approaches. The approach of CMS is described in the previous paragraph. At ATLAS, there are separate analyses probing CP invariance in the production and decay of the Higgs boson.

In the Higgs CP decay analysis, which was conducted in the same decay channels as

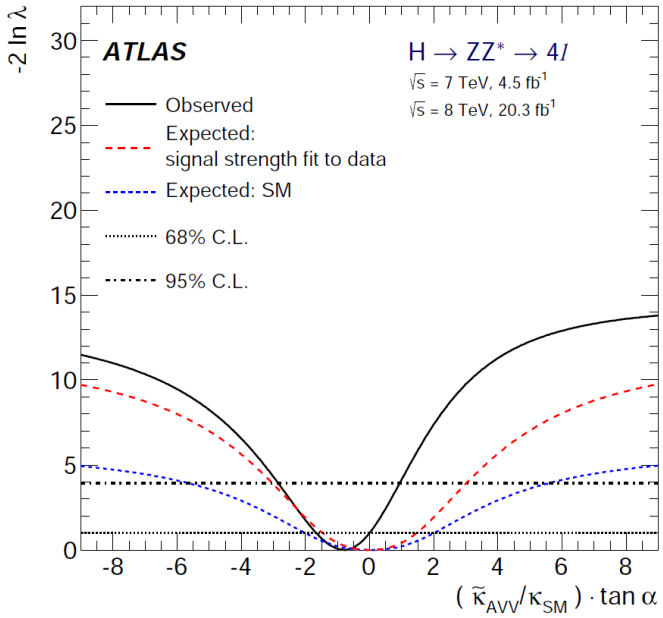

Figure 2.4: Result of the maximum likelihood fit for the CP-mixing parameter $(\tilde{\kappa}_{AVV}/\kappa_{\mathrm{SM}}) \cdot \tan(\alpha)$. The result is consistent with the SM prediction of zero.

the spin/parity analysis, the angular distributions of its decay products are compared to theoretical predictions depending on the mixing angle between the two Higgs states. The best-fitting mixing parameters are then determined with a maximum likelihood fit.

The result is expressed in terms of the CP-mixing parameter $(\tilde{\kappa}_{AVV}/\kappa_{\mathrm{SM}}) \cdot \tan(\alpha)$ with α and $\tilde{\kappa}_{AVV} = \frac{v}{4\Lambda}\kappa_{AVV}$ as defined in section 2.2.3. This parameter is proportional to the parameter $\tilde{d}$ described in section 2.2.2 that is used in the Higgs production analysis. The result was that the regions $(-\infty, -2.18]$ and $[0.83, \infty)$ could be excluded for the parameter $(\tilde{\kappa}_{AVV}/\kappa_{\mathrm{SM}}) \cdot \tan(\alpha)$ at 95% confidence level, as can be seen in figure 2.4. Since the observed value at zero for $-2\ln(\lambda)$ with the likelihood function λ is well below the 95% line, this result is consistent with the SM prediction of 0, meaning no CP-mixing and an even CP state.
The ATLAS Higgs CP analysis probing the production mechanism is, at the moment, restricted to the vector boson fusion (VBF) Higgs production mechanism and the decay channel $H \to \tau\tau$. Since the objective of this thesis is to verify and improve this analysis, its foundations will be described in more detail in the following.

2.2 Vector Boson Fusion

This section focuses on the Higgs boson production mechanism probed in this analysis, namely vector boson fusion. The reasons why this production mechanism is chosen, the theoretical models applied to describe the process, and the test of the CP nature of the Higgs boson are described.

2.2.1 General Vertex Structure

When trying to produce the SM Higgs boson in a proton-proton accelerator, there are several different production processes. Their Feynman diagrams and cross sections as a function of the center-of-mass energy $\sqrt{s}$ of the proton-proton system are shown in figure 2.5, and one can see that the gluon-gluon fusion is the dominant production process. Nevertheless, this work will focus on the VBF production so the coupling of a Higgs boson to vector bosons can be probed. Despite expecting less VBF events than for gluon fusion, these leave a characteristic signature in the detector which allows for better distinction from the background than in the case of gluon fusion. The two jets produced along with the Higgs boson, called VBF jets in the following, are typically detected in the forward regions of the detector and deposit their energy in opposite hemispheres, which happens more seldomly for non-VBF events.
Considering a scalar boson coupling to two weak vector bosons, the most general Lorentz

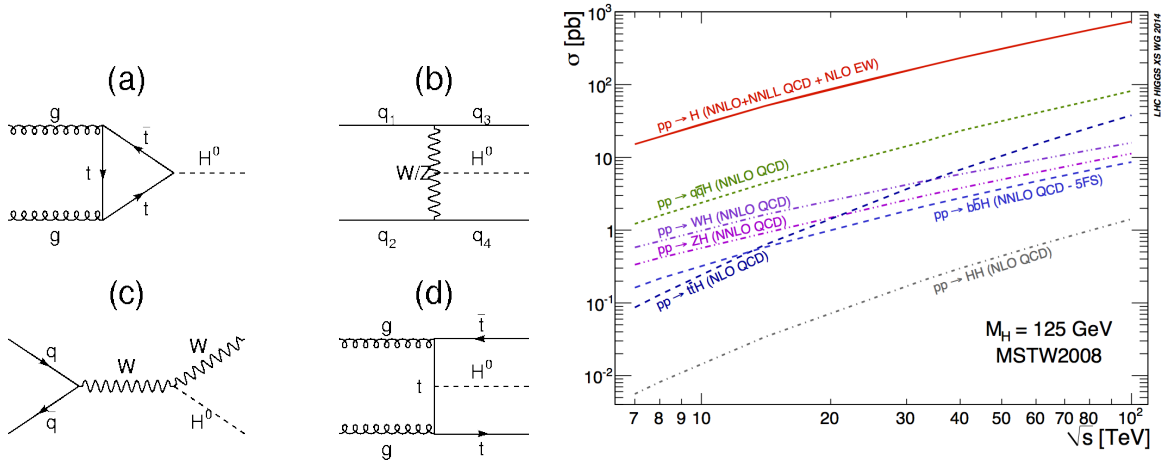

Figure 2.5: Higgs production processes and their cross sections depending on the Higgs mass. The gluon-gluon fusion, depicted in diagram (a), has the highest cross section. For the mass relevant in this work, $m_H = 125$ GeV, it is an order of magnitude higher than the cross section for VBF, shown in diagram (b). The associative productions with a vector boson (c) or top quarks (d) have even lower rates.

invariant form [18] for the vertex structure is

$$\begin{aligned} T^{\mu\nu}(q_1, q_2) = {} & a_1(q_1, q_2) g^{\mu\nu} \\ & + a_2(q_1, q_2) \left(q_1 q_2 g^{\mu\nu} - q_2^\mu q_1^\nu\right) \\ & + a_3(q_1, q_2) \epsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma} \ , \end{aligned} \tag{2.2}$$

where q_1 and q_2 are the four-momenta of the gauge bosons, $g^{\mu\nu}$ is the Minkowski metric, ϵ is the totally antisymmetric tensor in four dimensions and a_i for $i \in \{1, 2, 3\}$ are scalar functions of the momenta of the gauge bosons. The terms proportional to a_1 and a_2 are both CP-even, while the third term is CP-odd. In the SM, a_2 and a_3 are equal to zero, resulting in a CP-even coupling. Any deviation from this SM coupling is called an anomalous coupling. In the following, it is assumed that a_1 has its non-zero SM value and that $a_2 = 0$.[1]

The goal of this analysis is to verify and improve a method to test whether a_3, and thus the contribution from the third line in equation 2.2, is actually zero for vector bosons coupling to the recently discovered particle (called "HVV couplings" in the following). If it is not (and neither is a_1), there is an even and an odd CP contribution to the coupling. Such a mixed coupling would imply CP violation, meaning that the VBF process would

[1] The hypothesis $a_1 = a_3 = 0$ and $a_2 \neq 0$ corresponds to the non-SM CP-even case 0_h^+ listed in table 2.2.

distinguish matter from antimatter.
In this scenario, the Higgs boson's couplings could be a reason for the observed baryon asymmetry, since C and CP symmetry are broken [8].

2.2.2 Anomalous Couplings

There are several ways of parametrising the Lagrangian of the anomalous HVV couplings. The one that is mainly used in this analysis employs operators of mass dimension 6. The effective Lagrangian [18] for the Higgs sector can be written

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \frac{f_{\tilde{B}B}}{\Lambda^2} O_{\tilde{B}B} + \frac{f_{\tilde{W}W}}{\Lambda^2} O_{\tilde{W}W} \ ,$$

where the operators $O_{\tilde{B}B}$ and $O_{\tilde{W}W}$ are defined as

$$O_{\tilde{B}B} = \Phi^\dagger \hat{\tilde{B}}_{\mu\nu} \hat{B}^{\mu\nu} \Phi \qquad\qquad O_{\tilde{W}W} = \Phi^\dagger \hat{\tilde{W}}_{\mu\nu} \hat{W}^{\mu\nu} \Phi \ . \tag{2.3}$$

Here, Λ is the energy scale of new physics, Φ is the Higgs doublet, $V^{\mu\nu}$ is the field strength tensor of the vector field V and $\tilde{V}_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$ is the dual field strength tensor. The tensors denoted with a hat are defined via $\hat{B}_{\mu\nu} = \mathrm{i}\frac{g'}{2} B_{\mu\nu}$ and $\hat{W}_{\mu\nu} = \mathrm{i}\frac{g}{2}\sigma^a W^a_{\mu\nu}$. After electroweak symmetry breaking in the unitary gauge, this Lagrangian can be written as

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \tilde{g}_{HAA} H \tilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{HAZ} H \tilde{A}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HZZ} H \tilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} H \tilde{W}^+_{\mu\nu} W^{-\mu\nu} \tag{2.4}$$

with four coupling parameters $\tilde{g}$. When calculating their values, it appears that only two of them are independent. Therefore, they can be expressed with two dimensionless parameters $\tilde{d}$ and $\tilde{d}_B$.

$$\begin{aligned} \tilde{g}_{HAA} &= \frac{g}{2m_W} \left(\tilde{d} \sin^2(\theta_W) + \tilde{d}_B \cos^2(\theta_W) \right) \quad & \tilde{g}_{HAZ} &= \frac{g}{2m_W} \sin(2\theta_W)(\tilde{d} - \tilde{d}_B) \\ \tilde{g}_{HZZ} &= \frac{g}{2m_W} \left(\tilde{d} \cos^2(\theta_W) + \tilde{d}_B \sin^2(\theta_W) \right) \quad & \tilde{g}_{HWW} &= \frac{g}{m_W} \tilde{d} \end{aligned} \tag{2.5}$$

Calculating these coupling parameters by comparing the coefficients of the different operators in equations 2.3 and 2.4 leads to

$$\tilde{d} = -\frac{m_W^2}{\Lambda^2} f_{\tilde{W}W} \quad \tilde{d}_B = -\frac{m_W^2}{\Lambda^2} \tan^2(\theta_W) f_{\tilde{B}B} \ .$$

As one cannot experimentally distinguish the contributions to VBF Higgs production from the different gauge boson species, one can arbitrarily choose $\tilde{d} = \tilde{d}_B$ for simplicity.

Thus, one can specify the strength of the anomalous CP-odd coupling with only the parameter $\tilde{d}$. With the Lagrangian above and the choice made for $\tilde{d}_B$, one can express the Lorentz structure of the VBF Higgs production vertex as

$$T^{\mu\nu} = \sum_{V=W,Z} \frac{2m_V^2}{v} g^{\mu\nu} + \sum_{V=W,Z,\gamma} \frac{2g}{m_W} \tilde{d}\, \epsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma} \,.$$

When comparing this to equation 2.2, one obtains

$$a_1^{HVV} = \frac{2m_V^2}{v} \qquad a_2 = 0 \qquad a_3 = \frac{2g}{m_W}\tilde{d}\,,$$

where a_3 is independent of the species of vector bosons that is involved. Consequently, the matrix element $\mathcal{M}$ for this process can be written as

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \tilde{d}\, \mathcal{M}_{CP\text{-odd}}$$

with a CP-even standard model term and the contribution from CP-odd anomalous couplings. This equation shows that the SM prediction is equivalent to $\tilde{d} = 0$, leading to a purely CP-even coupling, whereas for $|\tilde{d}| \to \infty$, the SM part becomes negligible and the coupling purely CP-odd. Other values of $\tilde{d}$ are tantamount to mixed VBF coupling and CP violation.
Squaring the matrix element leads to a differential cross section proportional to

$$|\mathcal{M}|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2\tilde{d}\, \Re\left(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{CP\text{-odd}}\right) + \tilde{d}^2 \;\; |\mathcal{M}_{CP\text{-odd}}|^2 \;. \tag{2.6}$$

The first term is just the SM prediction for the cross section. The second part is the interference between odd and even coupling and causes CP violation. However, it does not increase the total cross section of the process. The third part is the square of the CP-odd matrix element and it gives a CP-even additional contribution to the cross section which is proportional to $\tilde{d}^2$.

2.2.3 Alternative Parametrisation

Apart from this parametrisation in terms of $\tilde{d}$, there is an alternative parametrisation that is also used to describe VBF Higgs couplings. Since these parameters are needed as well, their definitions [19] and how to convert them to $\tilde{d}$ are shown in this section.
When using the Higgs characterisation [20] (HC) UFO [21] model in the Monte Carlo generator "MadGraph5_aMC@NLO" [22] to generate VBF events with different $\tilde{d}$ values,

one cannot specify the model by simply choosing a value for $\tilde{d}$. Instead, the values of the parameters of the following effective Lagrangian for interactions between the model's Higgs boson and any gauge bosons can be changed.

$$\begin{aligned}
\mathcal{L}^{M}_{eff} = \Big\{ & c_\alpha \kappa_{\mathrm{SM}} [\frac{1}{2} g_{HZZ} Z_\mu Z^\mu + g_{HWW} W^+_\mu W^{-\mu}] \\
& - \frac{1}{4} [c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma} A_{\mu\nu} \tilde{A}^{\mu\nu}] \\
& - \frac{1}{2} [c_\alpha \kappa_{HZ\gamma} g_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} g_{AZ\gamma} Z_{\mu\nu} \tilde{A}^{\mu\nu}] \\
& - \frac{1}{4} [c_\alpha \kappa_{Hgg} g_{Hgg} G^a_{\mu\nu} G^{a,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg} G^a_{\mu\nu} \tilde{A}^{a,\mu\nu}] \\
& - \frac{1}{4\Lambda} [c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} Z_{\mu\nu} \tilde{Z}^{\mu\nu}] \\
& - \frac{1}{2\Lambda} [c_\alpha \kappa_{HWW} W^+_{\mu\nu} W^{-\mu\nu} + s_\alpha \kappa_{AWW} W^+_{\mu\nu} \tilde{W}^{-\mu\nu}] \\
& - \frac{1}{\Lambda} c_\alpha [\kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial_\mu Z^{\mu\nu} + (\kappa_{H\partial W} W^+_\nu \partial_\mu W^{-\mu\nu} + h.c.)] \Big\} H
\end{aligned} \tag{2.7}$$

In this Lagrangian, c_α and s_α are the sine and cosine of α, the mixing angle between the even and odd Higgs states. Therefore, c_α is always a prefactor of the CP-even couplings. The CP-even Higgs state is usually denoted as H, while a hypothetical CP-odd Higgs state is denoted as A. Thus, parameters κ and g of the strength of CP-odd couplings are denoted with an A in the index where the even couplings have an H. The parameters κ are dimensionless, and in MadGraph they are used to specify the couplings of a model. The g parameters are usually not dimensionless and have fixed values from table 2 of ref. [19] in MadGraph.

Since the models investigated in this report reproduce the SM couplings and have an additional CP-odd coupling, the operators of the SM couplings should have a prefactor of 1. These couplings can be seen in line 1 of equation 2.7, and the prefactor is $c_\alpha \kappa_{\mathrm{SM}}$. These would both be one in the SM case, but in order to allow for CP-mixed coupling, the choice $\kappa_{SM} = \frac{5}{3}$ and $c_\alpha = 0.6$ is made. This implies $s_\alpha = 0.8$ and that the strengths of the SM Yukawa couplings of the Higgs to fermions need to be multiplied by $\frac{5}{3}$.

The question remaining is how a set $\tilde{d}$ value relates to the coupling parameters κ of this parametrisation. From equation 2.5, one infers that $\tilde{d}$ only affects the CP-odd couplings of the Higgs to two Z, two W or two γ bosons. The $HZ\gamma$ coupling vanishes for $\tilde{d} = \tilde{d}_B$. Thus, κ_{AZZ}, κ_{AWW} and $\kappa_{A\gamma\gamma}$ are the only parameters affected by a variation of $\tilde{d}$. The conversion formulas can be derived by comparing the coefficients of the operators corresponding to

these parameters in the equations 2.4 and 2.7. This leads to

$$\kappa_{AZZ} = \kappa_{AWW} = -\frac{2g\Lambda}{m_W s_\alpha}\tilde{d}$$
$$\kappa_{A\gamma\gamma} = -\frac{3\pi v g}{2m_W s_\alpha \alpha_{\mathrm{EM}}}\tilde{d}\,. \tag{2.8}$$

The other non-SM couplings do not contribute to the VBF models discussed in this report. Therefore, many of them, for instance the coupling to gluons, are set to zero.

2.3 The Optimal Observable Method

For a simple test of CP invariance, one can choose a CP-odd observable $\mathcal{O}$ of the investigated process and check whether its distribution is centred around zero. This follows from

$$CP[\mathcal{O}] = -[\mathcal{O}]\,,$$

which, within uncertainties, has to hold exactly for the mean $\langle\mathcal{O}\rangle$ of the distribution. $\langle\mathcal{O}\rangle = 0$ is consistent with CP conservation, thus $\langle\mathcal{O}\rangle \neq 0$ implies CP violation.
In order to decide which observable of a process to use for this type of test, one option is to consider its differential cross section [23]. It depends on the phase space Ω and a CP violation inducing parameter α. An expansion in α yields

$$\frac{\mathrm{d}\sigma(\Omega,\alpha)}{\mathrm{d}\Omega} = S^{(0)}(\Omega) + \alpha\, S^{(1)}(\Omega) + \alpha^2\, S^{(2)}(\Omega) \tag{2.9}$$

with scalar functions $S^{(0)}$, $S^{(1)}$ and $S^{(2)}$ of Ω. Since the cross section determines the probability of a process, the likelihood for N events can be written as

$$L(\alpha) = \prod_i^N \left(S^{(0)}(\Omega) + \alpha\, S^{(1)}(\Omega) + \alpha^2\, S^{(2)}(\Omega)\right)\,.$$

To estimate α, one has to maximise L.

$$\left.\frac{\mathrm{d}\log(L)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_0} = 0 = \sum_i^N \frac{S^{(1)}(\Omega) + 2\alpha_0\, S^{(2)}(\Omega)}{S^{(0)}(\Omega) + \alpha_0\, S^{(1)}(\Omega) + \alpha_0^2\, S^{(2)}(\Omega)}$$
$$= \sum_i^N \frac{O^{(1)}(\Omega) + 2\alpha_0\, O^{(2)}(\Omega)}{1 + \alpha_0\, O^{(1)}(\Omega) + \alpha_0^2\, O^{(2)}(\Omega)}$$

The variables $O^{(1)}$ and $O^{(2)}$ are defined as $\frac{S^{(1)}}{S^{(0)}}$ and $\frac{S^{(2)}}{S^{(0)}}$, respectively. Expressed this way, only the values of these two so-called "optimal observables" O enter the likelihood function that actually depends on all the available phase space variables. Therefore, a maximum likelihood fit to the high-dimensional phase space is equivalent to a two-dimensional fit to only $O^{(1)}$ and $O^{(2)}$.
Since this calculation started with an expansion of the cross section, this is only valid for small values of the parameter α. For very small values, even $O^{(2)}$, the second order optimal observable, can be neglected.
Transferred to this analysis, the CP violation inducing parameter is $\tilde{d}$ and the expansion in equation 2.9 corresponds exactly to equation 2.6. Following the procedure above, one finds that the optimal observables $\mathcal{O}_1$ and $\mathcal{O}_2$ for testing CP invariance in vector boson Higgs fusion are

$$\mathcal{O}_1 = \frac{S^{(1)}}{S^{(0)}} = \frac{2\Re\left(\mathcal{M}^*_{\mathrm{SM}}\mathcal{M}_{CP\text{-odd}}\right)}{\left|\mathcal{M}_{\mathrm{SM}}\right|^2} \qquad \mathcal{O}_2 = \frac{S^{(2)}}{S^{(0)}} = \frac{\left|\mathcal{M}_{CP\text{-odd}}\right|^2}{\left|\mathcal{M}_{\mathrm{SM}}\right|^2} \,.$$

Since the data are consistent with the SM predictions and the purely CP-odd Higgs state has already been excluded, assuming $|\tilde{d}| < 1$ is reasonable. Therefore, using the optimal observable method in the way shown above, which is only valid for a small parameter, is justified.

3 Setup of Experiment and Analysis

This chapter will briefly describe the LHC and the ATLAS detector which provide the data investigated in these studies. The different subdetectors as well as their uses and limitations will be covered. The method of data acquisition and the various data types used in the ATLAS experiment will also be described, along with basic hadron collider phenomenology.

3.1 The ATLAS Experiment at the LHC

3.1.1 The LHC Accelerator Complex

The Large Hadron Collider (LHC) is a hadron accelerator and collider located at CERN near Geneva [24]. It is installed in the 26.7 km tunnel that was constructed between 1984 and 1989 for the LEP electron-positron collider after LEP was removed in 2000. The tunnel system is 50 to 175 m deep underground, and there are two transfer tunnels of approximately 2.5 km length connecting the LHC to the older CERN accelerators that now act as injectors. An image of the entire complex can be seen in figure 3.1. The LHC accelerates protons in order to make them collide at a center-of-mass energy of $\sqrt{s} = 13$ TeV. To keep the protons inside the beampipe during acceleration, magnets with field strengths of up to 8 T are used. The highest instantaneous luminosities reached are of the order of 10^{34} m^{-2} s^{-1}.
In figure 3.1, one can also see the locations of the four largest experiments at CERN: ALICE, ATLAS, CMS and LHCb. The ALICE experiment uses heavy ion collisions to produce and examine quark-gluon plasma, a state of matter that the universe was in for a few milliseconds directly after the big bang. LHCb specialises in investigating decays of hadrons containing b or c quarks to perform precision measurements of CP violation and rare decays. ATLAS and CMS are multi-purpose detectors that can test a wide variety of models for new physics in addition to measuring the properties of the SM Higgs boson.

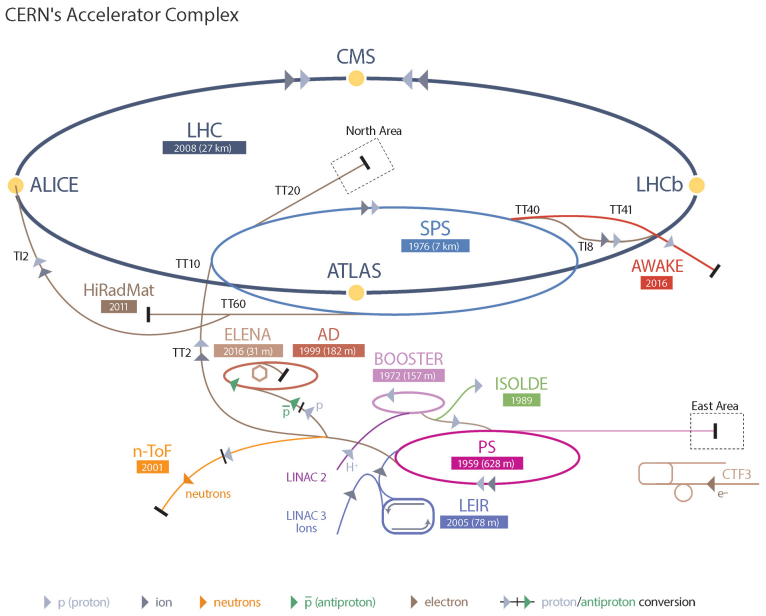

Figure 3.1: The accelerators at CERN. PS and SPS are used to pre-accelerate protons before injecting them in the LHC. Also, the locations of the four largest experiments using the LHC are shown.

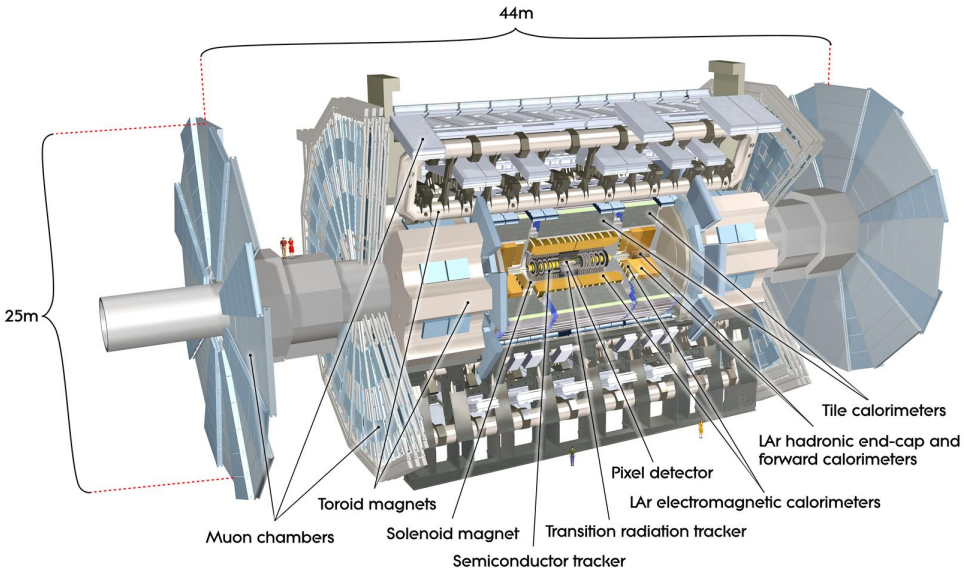

Figure 3.2: The ATLAS detector with all its subcomponents in cut-away view.

3.1.2 Structure of the ATLAS Detector

This section will provide a short description of the ATLAS detector and its components, which can be seen in figure 3.2. It is 25 metres high and 44 metres long, and it weighs approximately 7000 tons [25]. It consists of a barrel located in the centre and two end-caps surrounding it.

A cylindric coordinate system is used. The z-axis is at the beam pipe, ϕ is the azimuthal angle measured around the beam axis and θ is the polar angle from the beam axis. In most cases, instead of θ, the pseudorapidity $\eta = -\ln(\tan(\frac{\theta}{2}))$ will be used to specify the direction of an object in the detector. The transverse momentum of an object is the projection of its momentum orthogonal to the z-axis. As a measure of the angular distance between two objects i and j, the quantity $\Delta R = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$ is employed.

The inner solenoid magnet produces a magnetic field of strength 2 T, which is necessary for momentum measurements, measuring the ratio $\frac{q}{p}$ of charge over momentum of tracks, electron identification and determining where an interaction took place, the so-called vertex reconstruction. All of this takes place in the inner detector (ID) in the centre of the ATLAS detector, which consists of three parts: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). The pixel detector consists of three layers in the barrel and three disk layers in each end-cap [26]. These layers are mainly composed of silicon and are the closest to the interaction point where the proton beams meet, and they detect charged tracks, which is in particular necessary for τ lepton reconstruction, in the region $|\eta| < 2.5$. The SCT consists of four layers of silicon microstrip detectors in the barrel and up to three rings in the forward regions of the detector. These detectors can measure momenta of charged tracks and the interaction points of the event. The TRT consists of straw-tubes and increases the accuracy of the track measurement as well as electron identification in the region $|\eta| < 2.0$. Moreover, it provides additional discrimination between electrons and hadrons. The relative momentum resolution for tracks traversing the inner detector is $\frac{\sigma_p}{p} = (4.83 \pm 0.16) \cdot 10^{-4}(\text{GeV})^{-1} \cdot p_T$ [27].

Around the inner detector, the calorimeters are installed. The liquid argon (LAr) electromagnetic (EM) calorimeter is used to identify photons and electrons. It is divided in a barrel part and two end-caps covering the pseudorapidity ranges $|\eta| < 1.475$ and $1.375 < |\eta| < 3.2$, respectively. The depth of the EM calorimeter is > 22 radiation lengths X_0 in the barrel and > 24 X_0 in the end-caps. One radiation length is $\frac{7}{9}$ of the mean free path for pair-production by a highly energetic photon and also the mean distance an electron traverses until its energy is reduced by $\frac{1}{\mathrm{e}}$. The energy resolution of the EM barrel and end-cap calorimeters is $\frac{\sigma_E}{E} = \left(0.7 + 10\sqrt{\frac{\text{GeV}}{E}}\right)\%$ [28].

The hadronic tile calorimeter is used to reconstruct hadronic jets and stop hadronic

showers before reaching the muon spectrometer. The barrel parts measure jets with $|\eta| < 1.7$ and have a relative energy resolution of $\frac{\sigma_E}{E} = \left(3 + 50\sqrt{\frac{\text{GeV}}{E}}\right)\%$. The depth of the hadronic calorimeter is typically given in nuclear interaction lengths λ defined as the mean path distance relativistic particles can pass until their energy is reduced by $\frac{1}{\text{e}}$. Since the hadronic calorimeter in the barrel consists of several layers located in different regions, its depth is dependent on η. In the central region around $\eta = 0$, it is 9.7 λ. The end-cap hadronic LAr calorimeters cover $1.5 < |\eta| < 3.2$ and have a relative energy resolution of $\frac{\sigma_E}{E} = \left(3 + 50\sqrt{\frac{\text{GeV}}{E}}\right)\%$.
In addition to these calorimeters, there are forward LAr calorimeters close to the beam pipe covering the region $3.1 < |\eta| < 4.9$. These calorimeters can reconstruct hadronic and electromagnetic showers, but their resolution of $\frac{\sigma_E}{E} = \left(10 + 100\sqrt{\frac{\text{GeV}}{E}}\right)\%$ is not as good as that of the other subdetectors.
All the calorimeters are surrounded by toroid magnets producing a field of 0.5 T in the barrel and 1 T in the end-caps. The outmost part, the muon spectrometers, can detect muons with pseudorapidities $|\eta| < 2.7$. Its relative transverse momentum resolution varies between $\frac{\sigma_{p_T}}{p_T} = 10\%$ at 1 TeV and 2% at 50 GeV.
Since there are more proton collisions than can be saved, a two-level trigger system [29] is used to select the most relevant events for further use. The hardware-based Level-1 trigger uses a subset of detector information to reduce the event rate of 30 MHz to 100 kHz. The decision time for the Level-1 trigger is 2.5 μs. The rate of accepted events is then reduced to about 1 kHz by the software-based "high-level trigger" that uses full detector information and needs about 0.2 s for its selection algorithms.

3.1.3 Data Acquisition and Processing at ATLAS

This section describes the most important steps of the data processing procedure applied by the ATLAS collaboration. The first step is for events measured in the detector to fulfill certain requirements in order to fulfill the trigger conditions and be saved for further analysis. Events that do not fit these criteria are discarded. The remaining events undergo an object reconstruction where different particles and jets are identified, and are then saved in large files called xAODs.
Each analysis team now applies a different so-called "derivation" to this xAOD that deletes unwanted information in order to keep mostly events and event information relevant for its purposes. This can include kinematic cuts and conditions on the number of jets in the events. Within the ATLAS τ working group that works on processes involving τ leptons, such as the VBF $H \rightarrow \tau\tau$ process examined in this thesis, the xAODs are processed according to figure 3.3. The figure shows that after the steps that are independent of the

event signature and the more specific derivation, the resulting DxAOD is converted to a so-called n-tuple using the "xTauFramework", a code package developed within ATLAS. This file contains all the important information and kinematic variables, and it is the file that is mainly used for the analysis.

3.2 Hadron Collider Phenomenology

In order to properly interpret the data measured by the detector, understanding the origin of the observed objects is essential. Therefore, the following section describes the processes that lead to the production of hadronic jets and how these processes are simulated with Monte Carlo methods. Additionally, the algorithm used in the analysis to define the jets is outlined.

3.2.1 Hadronization in Nature and Simulation

During proton-proton collisions at the LHC, the constituents of the protons scatter, which, in many events, leads to some of them being separated from the rest of the proton. Nonetheless, one never observes single quarks or gluons in a detector. Instead, the detector finds jets that consist of a large number of hadrons. The reason for this phenomenon is called confinement, and it means that particles carrying a colour charge can only exist in bound states that are colourless. The mechanism is not calculable using perturbation theory, but the main reasons for this phenomenon to occur can be given: The running coupling strength α_s of strong interactions is close to 1 at low energies, leading particles with colour charge to emit quarks and gluons as their energy decreases.

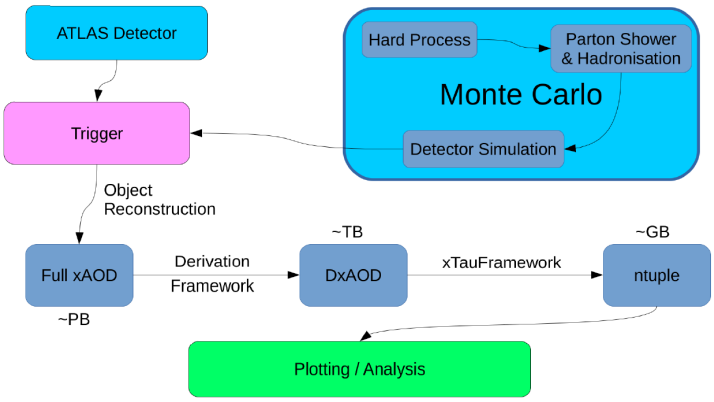

Figure 3.3: Schematic data flow in ATLAS for processes involving τ leptons.

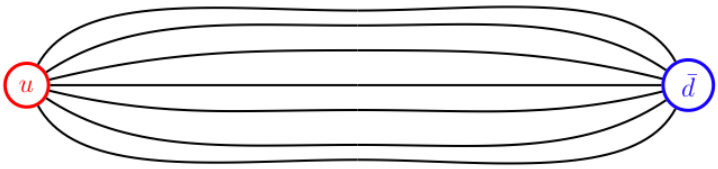

Figure 3.4: Pictorial representation of the attractive force field between two isolated quarks.

The other cause of confinement is that the gluons, the mediator particles of the strong force, carry colour charge themselves. This separates the strong force from other forces, such as electromagnetism, and it leads to a force field between two quarks as sketched in figure 3.4 whose energy increases with the distance between the quarks. Therefore, when quarks out of a bound state such as a proton are separated during a collision at the LHC, it is energetically more favourable to create a quark-antiquark pair between the initial quarks at a certain distance to reduce the size of the gluon field. This pair production reduces the kinetic energy of the involved quarks, and it happens until there are no single quarks remaining and all the final state particles are grouped into colour neutral hadrons. Simulations of such events are essential for the analysis of LHC data, but at the moment there is no exact way to calculate hadronization processes. Instead, the event generation is split into the following steps: Hard interaction, parton showering and detector simulation. These make simplifying assumptions to reach feasible calculation times while maintaining a satisfactory accuracy.

To start the calculation, the parton distribution functions are used to simulate which constituents of the two protons are involved in the interaction. Then the hard interaction, the interaction with the highest momentum transfer, is simulated by a generator such as MadGraph 5 [22]. The generators can also calculate the properties of the interaction products as well as their decay products. Most generators leave unstable particles such as τ leptons and also single quarks in their final state.

After that, a parton showering program such as PYTHIA 8 [30] is used. It simulates the further evolution of the event from where the event generator stopped. The simulated hadronization processes are usually based on simple models that force the gluons exchanged between single partons, meaning quarks or gluons, in the process to split into quark-antiquark pairs when reaching a specified energy. This process is repeated multiple times, also for hadronic decays of other particles, until a final state that is comparable to those obtained in experiments is reached.

The last step is to simulate the interactions between the particles in the simulated process and the detector with which the real collisions are to be reconstructed, in this case the

ATLAS detector. The detector simulator that is mostly used by the ATLAS collaboration is GEANT 4 [31].

3.2.2 Jet Algorithms

Jet clustering algorithms are used to combine large numbers of collimated hadrons produced in events to jets. Therefore, they are an important part of almost every analysis in ATLAS. The most widely used algorithms in ATLAS are of the so-called "k_t" family, and they work very similarly [32]. One example is the "anti-k_t" algorithm with distance parameter $R = 0.4$, the main algorithm used for this analysis.
The first step is defining a measure of distance d_{ij} between two objects i and j. For the k_t-family of algorithms it is commonly defined as

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \ , \tag{3.1}$$

where $p_{T,i}$ is the transverse momentum of object i, the exponent p is an algorithm dependent parameter, ΔR_{ij} is the ΔR value between objects i and j and R is the distance parameter of the algorithm which is set to 0.4 in this case. Additionally, the reference distance $d_{iB} = p_{T,i}^{2p}$ is defined.
The algorithm starts by choosing the object i with the highest d_{iB} value and calculating $d_{\min} = \min(d_{ij}, d_{iB})$ for all objects $j \neq i$. If $d_{\min} = d_{ik}$, then i and k are combined and form a new object, replacing the old i. This is repeated until $d_{\min} = d_{iB}$. Then object i is considered a successfully clustered jet, and the procedure is restarted with the remaining object l with the highest d_{lB} value.
For anti-k_t algorithms, the exponent p is set to -1. With this choice, infrared and collinear safe jets are produced, and their shape is not influenced by soft radiation.

3.3 Analysis of Simulated Events

3.3.1 Levels of Monte Carlo Data

In this analysis, different types of simulated Monte Carlo data are used. This section is meant to clarify the differences in generation and applicability of these types.
The simplest type are generator level data. These are the direct output of Monte Carlo generators. Consequently, they have unstable particles in their final states, which makes these events unphysical. In parts, however, they can still be used for analysis. If one is only interested in initial state information, the missing hadronization is not a problem.

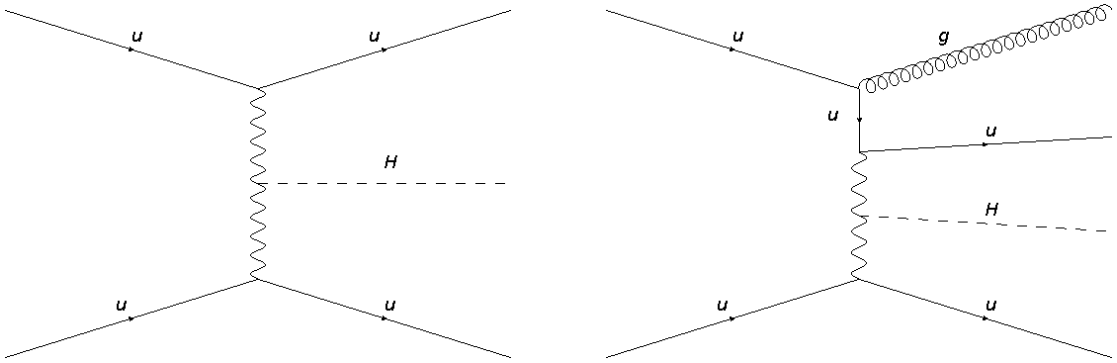

Figure 3.5: LO (left) and NLO (right) diagrams for the same VBF process. The NLO diagram contains one QCD vertex more than the LO calculation, leading to radiation of a gluon.

These data can still be used to calculate the VBF production matrix element and distributions of the optimal observables free from detector and showering effects. The kinematic distributions of the final state partons, however, should not be directly compared to the corresponding distributions of jets in data.
Another type of simulated data are the truth level data. These are the result of applying a parton shower to generator level data, and consequently, the final state is more realistic. Truth level data are used to find problems occurring during object reconstruction and determine the detector resolution. Also, they can be used to test the accuracy of different analysis methods used on data.
The last type of simulated data are reconstruction level data which are truth level data that have gone through a detector simulation. Hence, reconstruction level data are what a detector, in this analysis the ATLAS detector, would measure if the corresponding truth level events took place inside the detector. The resulting events are comparable to real collision data, and can be used to devise, test and improve analysis strategies.
In the event generation step, there are two possibilities used in this analysis: leading order (LO) or next-to-leading order (NLO) in QCD calculation. The difference is in the complexity of the calculation, and is illustrated in figure 3.5. LO calculations calculate processes to get from a given initial state to a given final state with as few couplings as possible. For a LO cross section calculation and event simulation, only these processes with minimal number of couplings are taken into account. NLO calculations additionally take diagrams with one extra QCD vertex into account. This can lead to additional final state particles, as in figure 3.5.

3.3.2 HAWK and the Event Reweighting Method

The main goal of the optimal observable method and this analysis is to determine which CP-mixing scenario expressed through the parameter $\tilde{d}$ fits experimental data the best. To that end, one can compare different simulated $\mathcal{O}$ distributions to the experimental data and calculate which yields the best agreement. But within a given range of possible values, one cannot cover all possibilities by simply simulating events for $\tilde{d}$ values at very small intervals. This would be too computationally expensive. Instead, one can simulate only SM events and calculate a weight for each of them depending on the $\tilde{d}$ that is of interest, so that the reweighted SM events are equivalent to Beyond Standard Model (BSM) events with the value of $\tilde{d}$ chosen beforehand.
These weights, as well as the optimal observables of first and second order, were calculated with HAWK [33] for this analysis, a Monte Carlo (MC) generator that specialises in simulating VBF and VH Higgs production in hadron colliders. As kinematic input, it uses the centre-of-mass energy which is 13 TeV for this analysis, the Bjorken x values of the initial state partons, the four-vectors of the two VBF jets, and the four-vector of the Higgs boson. The Bjorken x values x_1 and x_2 can be calculated from

$$x_{1/2} = \frac{M_{\text{final}}}{\sqrt{s}} \mathrm{e}^{\pm y_{\text{final}}} \ .$$

The index "final" means it is the value of the whole final state obtained by adding the four-vectors of all particles and then calculating the mass M or the rapidity $y = \frac{1}{2} \ln\left(\frac{E+p_z}{E-p_z}\right)$. Instead of using the Higgs four-vector it is given, HAWK uses x_1 and x_2 to calculate the motion of the center-of-mass frame and sets the Higgs four-momentum such that its sum with the jet four-vectors yields exactly the center-of-mass momentum. This is done in order to ensure energy-momentum conservation which is needed for calculating non-vanishing matrix elements. In the case of the Higgs decaying to τ leptons, this method is more accurate than reconstructing it from the Higgs decay products, because subsequent τ decays produce neutrinos that cannot be detected, leading to missing momentum and large uncertainties.
The weights w chosen for the reweighting method are

$$w = \frac{|\mathcal{M}_{\text{BSM}}(\tilde{d})|^2}{|\mathcal{M}_{\text{SM}}|^2} \ ,$$

since the probability for an event to occur is proportional to its squared matrix element. In this expression, $\mathcal{M}_{\text{BSM}}$ is the matrix element for non-zero $\tilde{d}$ according to equation 2.6. Therefore, events that are more likely in a SM scenario receive smaller weights than those

with higher $|\mathcal{M}_{\mathrm{BSM}}|^2$. The reweighting method uses the same kinematic variables as the calculations for $\mathcal{O}$, but takes additional information on the model to which it reweights the SM events.
Using this method, one can create $\mathcal{O}$ distributions for arbitrary $\tilde{d}$ values without need of different simulated samples.

4 Kinematic Tests and Validation of the Reweighting on MC Samples at $\sqrt{s} = 13$ TeV

This chapter gives an overview over the Monte Carlo samples used for the analysis and the computing framework. It also describes the production of new MC samples for validating the event reweighting and shows validation results for the reweighting algorithm.

4.1 Datasets and Event Selection

The analysis used a simulated VBF $H \to \tau_{\text{lep}}\tau_{\text{had}}$ event sample[1] to begin with, for which the kinematic distributions of the jets and the Higgs boson decay products were to be tested. The notation $\tau_{\text{lep}}\tau_{\text{had}}$ indicates the decays of the τ leptons: One of them decays hadronically, one leptonically into a muon or an electron. This sample was simulated using the event generator Powheg [34] and the parton shower Pythia 8 [30] using the AZNLO tune [35] and the PDF sets CT10 and CTEQ6L1 [36].
The sample contains 130119 events at $\sqrt{s} = 13$ TeV. These are events that passed the derivation "HIGG4D2", which introduces cuts on reconstruction level. In this derivation, $p_T > 12$ GeV is required for muons, as well as a loose isolation as defined in ref. [37]. For electrons to pass the derivation, they need to fulfill $p_T > 15$ and medium identification quality [38]. For jets to be treated as candidates for the product of a hadronic tau decay, they must fulfill $p_T > 18$. Moreover, they must contain either one or three tracks and the absolute value of the sum of the electric charges of the tracks should be equal to 1. Also, they are overlap removed against electrons and muons, and must not be located in the crack region $1.37 < |\eta| < 1.52$ of the detector. Overlap removal means that, if the τ candidate is within $\Delta R < 0.2$ of an electron or muon, this candidate is discarded as a possible hadronic τ decay. Another requirement for them is $|\eta| < 2.5$, since they must

[1] File name: mc15_13TeV.341156.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_VBFH125 _tautaulh.merge.DAOD_HIGG4D2.e3888_s2608_s2183_r6765_r6282_p2419

have associated tracks and these can only be reconstructed by the inner detector which is limited to this region.
Using the xTauFramework, as described in section 3.1.3, a ROOT [39] n-tuple file was produced. In this step, an additional cut on the jet vertex tagger (JVT) score [40] is introduced. It is a measure of the fraction of the p_T of the charged tracks associated to a jet that can be traced back to the hardest interaction that took place while the considered event was recorded over the p_T of all associated tracks. The cut is introduced to suppress event contamination from the products of secondary proton collisions. These unwanted objects in the detector are called pile-up. Jets that fulfill $p_T < 60$ GeV, $|\eta| < 2.4$ and JVT < 0.59 are discarded. All jets are constructed using the anti-k_t algorithm [32] with distance parameter $R = 0.4$ using all stable hadrons produced in hadronic jets, as described in section 3.2.2, and the energy measurement was calibrated using the "local hadronic calibration" scheme presented in ref. [41]. An overlap removal with prompt leptons and hadrons coming from tau-lepton decays within a radius of 0.4 is performed.

4.2 Tests of Kinematic Distributions

4.2.1 Comparison of truth and reconstruction level distributions

With the n-tuple file, kinematic distributions such as the transverse momentum p_T or the pseudorapidity η of the leading and sub-leading jets, which are necessary to calculate $\mathcal{O}_1$ and $\mathcal{O}_2$, can be investigated. The leading jet is the jet with the highest p_T in the event, the sub-leading jet is the one with the second highest p_T. These distributions are shown in figures 4.1 to 4.2 at both truth and reconstruction levels. In case there are less than two reconstructed jets, all the properties of the missing jets are set to zero, which means that the bin containg zero for each distribution is filled in addition to the usual filling whenever a jet is missing. In order to avoid distorted distributions from this effect, the condition $p_T > 1$ GeV for both the reconstruction and truth level sub-leading jets is demanded in the creation of the plots 4.1 to 4.2 in this section. The total number of events, without cuts, is 24461.
The truth and reconstruction level p_T distributions in figure 4.1 are similar, but for low p_T values they deviate, making the mean reconstructed p_T value lower than the mean truth value. One reason for this deviation is that truth jets are occasionally mislabelled as hadronic tau decay products. These jets are then deleted from the list of reconstructed jets, leading to the remaining jets with lower p_T to move one spot up in the p_T-ordered jet list, and when no jet with lower p_T is reconstructed, the event is treated as containing

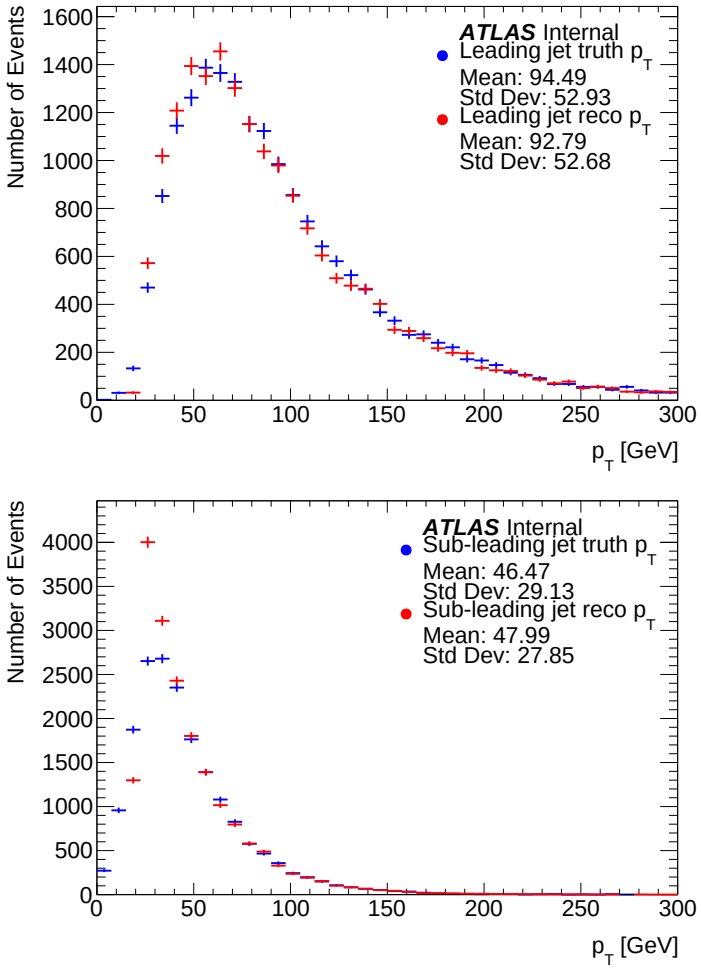

Figure 4.1: Transverse momentum p_T distributions for the leading and sub-leading jets in the used VBF $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ sample with only the cut $p_T > 1$ GeV applied at truth and reconstruction level applied after the n-tuple production. In events where no jet was found, the bin containing $p_T = 0$ is filled.

a jet with $p_T = 0$. Since this occurs more often than the case where hadronic tau decay products are misidentified as hadronic jets, the reconstructed jet p_T is lower than the truth jet p_T on average.

The shape of the η distributions has with a minimum at zero sets the VBF Higgs boson production apart from other production mechanisms and allows for better distinction from background processes.

The Bjorken x distributions in figure 4.3 on truth and reconstruction level look very similar, but the $\mathcal{O}$ distributions in figure 4.4 do not. On truth level, they have significantly lower standard deviations and higher peaks than the corresponding reconstruction level distributions. This can be attributed to measuring uncertainties that lead to the observed

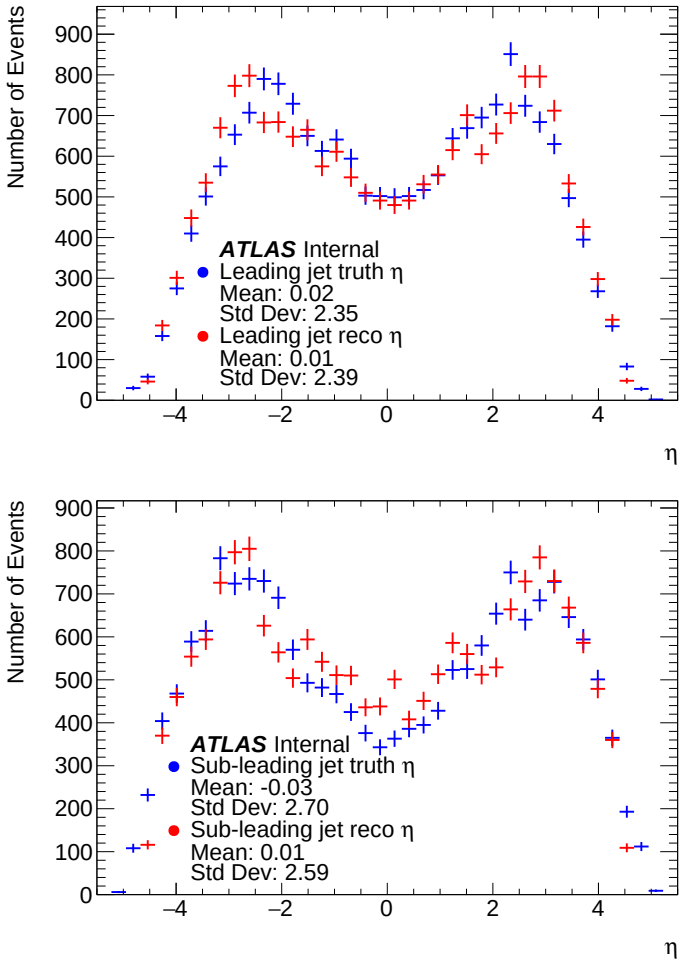

Figure 4.2: Pseudorapidity η distributions for the leading and sub-leading jets in the used VBF $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample with only the cut $p_T > 1$ GeV applied at truth and reconstruction level after the n-tuple production. When no jet is found, the bin containing $\eta = 0$ is filled. The two peaks at $|\eta| \approx 2.5$ are typical of VBF events.

distribution being equal to the truth distribution convolved with a Gaussian distribution.

To select VBF events in data, kinematic cuts based on the distributions in figures 4.1 to 4.4 are applied on reconstruction level. The cuts used in the following are $p_{T,\text{lead}}$, $p_{T,\text{sub-lead}} > 30$ GeV, $\eta_{\text{lead}} \cdot \eta_{\text{sub-lead}} < 0$ and $|\eta_{\text{lead}} - \eta_{\text{sub-lead}}| > 3$. After introducing these cuts, the resulting distributions are shown in figures 4.5, 4.6, 4.7 and 4.8. The biggest differences to the distributions before introducing the cuts are in the η distributions, where the minimum at 0 is now more distinct. Moreover, the $\mathcal{O}$ distributions now have significantly smaller standard deviations since the events with low p_T jets, which are the most susceptible to reconstruction errors and pile-up, are no longer used.

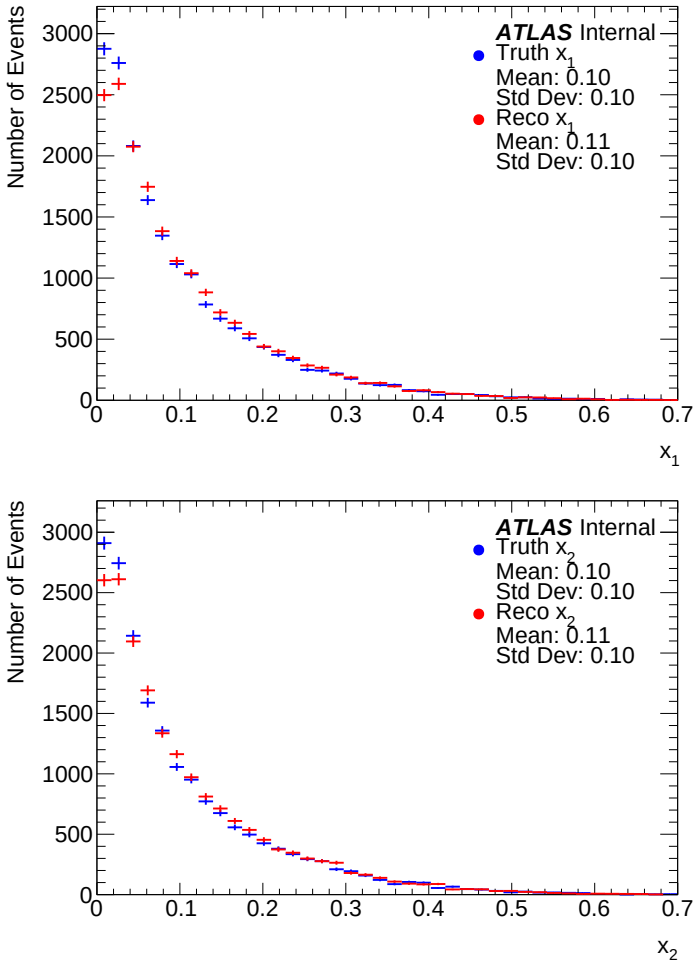

Figure 4.3: Bjorken x distributions in the VBF $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample with only the cut $p_T > 1$ GeV applied at truth and reconstruction level after the n-tuple production.

4.2.2 Mislabelling of truth level jets

This section will cover the reasons for a jet to be mislabelled. Understanding these is important for correctly interpreting results on reconstruction level or data.

As described in section 4.2.1, possible mislabelling of jets can lead to discrepancies between truth and reconstruction level distributions of the same observables. In figure 4.9, one can identify the main causes of mislabelling. It shows the p_T and η values of the truth jets that could not be matched to reconstructed jets. For this study, a reconstructed and a truth jet are considered matched if their angular distance fulfills $\Delta R < 0.4$. It appears that there are three types of unmatched truth jets: jets with $|\eta| < 2.5$, jets with $|\eta| > 4$, and jets with $p_T < 30$ GeV, regardless of their η value.

As described in section 3.1.2, the ATLAS detector can only reconstruct τ leptons in the

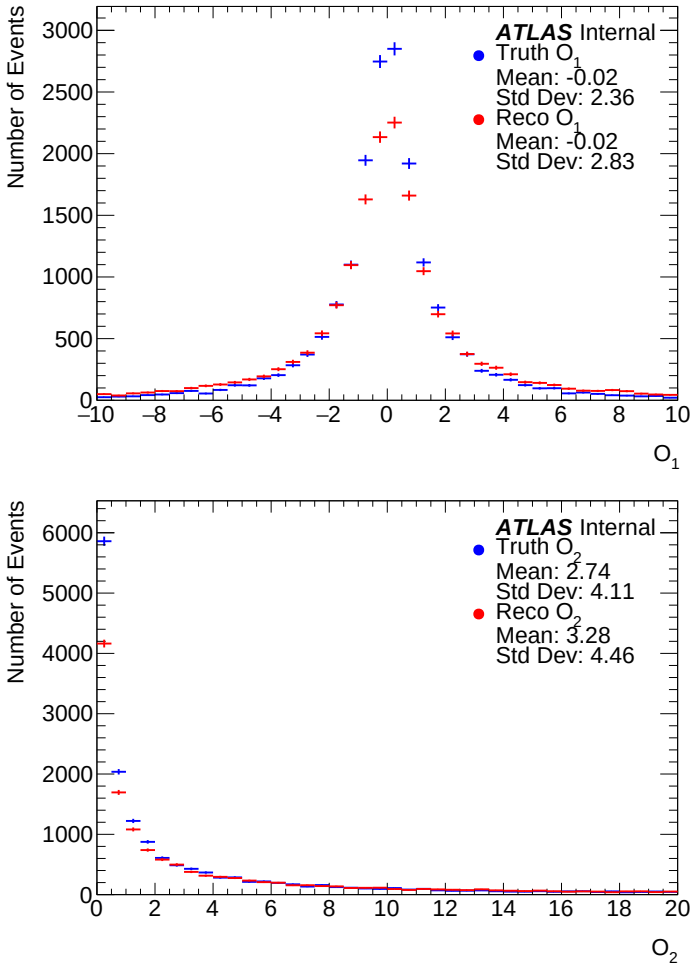

Figure 4.4: Optimal observable distributions in the VBF $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample with only the cut $p_T > 1$ GeV applied at truth and reconstruction level after the n-tuple production.

region $|\eta| < 2.5$ because track information is needed. This η region is the same as for the first group of truth jets without matched reconstructed jets. Therefore, it seems natural to assume that these are jets mistakenly identified as hadronically decaying τ leptons. This was tested by calculating the following ratio:

$$\frac{\#\text{truthjets(no reco jet, leading reco tau)}}{\#\text{truthjets(no reco jet)}} .$$

The denominator is the number of truth jets with $|\eta| < 2.5$ that have no matching reconstructed jet, the numerator is the number of the same jets that additionally are matched to the leading reconstructed hadronically decaying τ lepton candidate. The result for the leading jets is 28.6%, for the sub-leading jets the ratio is 12.8%. Consequently, in

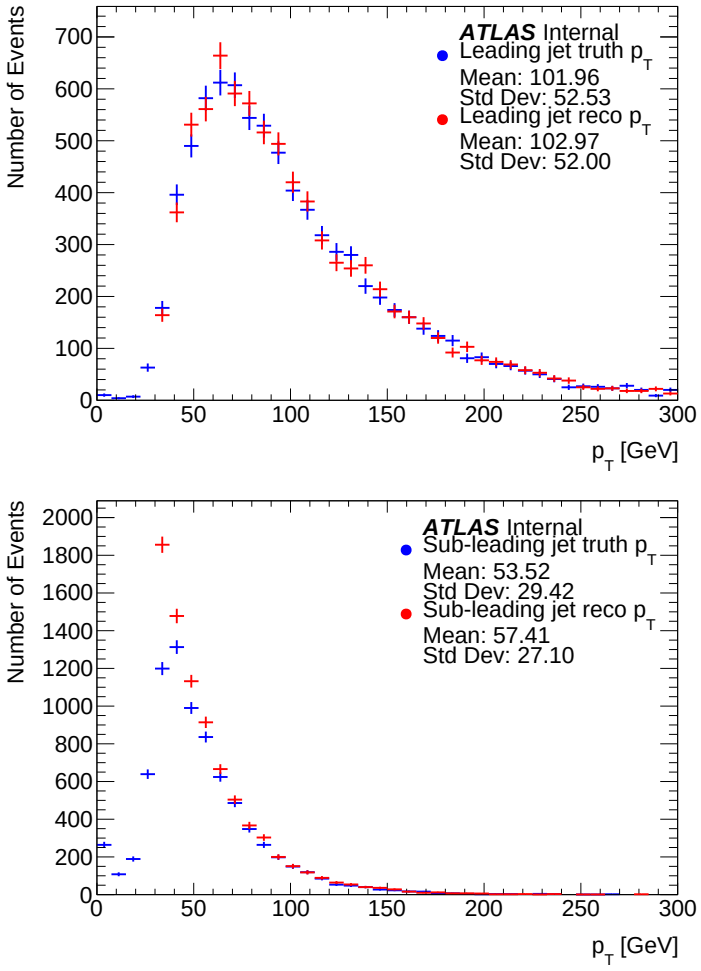

Figure 4.5: Transverse momenta of the two leading jets on truth and reconstruction level after introducing additional kinematic cuts.

more than 40% of the events containing truth jets with $|\eta| < 2.5$ that cannot be matched to reconstructed jets, one of them can be matched to the leading reconstruction level hadronically decaying τ lepton. This proves that a large part of the considered truth jets is mistakenly labelled as hadronically decaying tau leptons.

The second group of unmatched jets is very close to the limit of detector acceptance of $|\eta| \approx 4.5$. These jets were most likely not correctly reconstructed due to limited detector resolution which causes the reconstructed jet to be outside of the accepted range $|\eta| < 4.5$ while the original parton was not. In that case, the truth jet is not reconstructed. Hence, there is no reconstructed jet that can be matched to it.

In the third group, the jets are close to the p_T cut of 30 GeV applied to the reconstructed jets. These events are not matched to reconstructed jets because their p_T was slightly

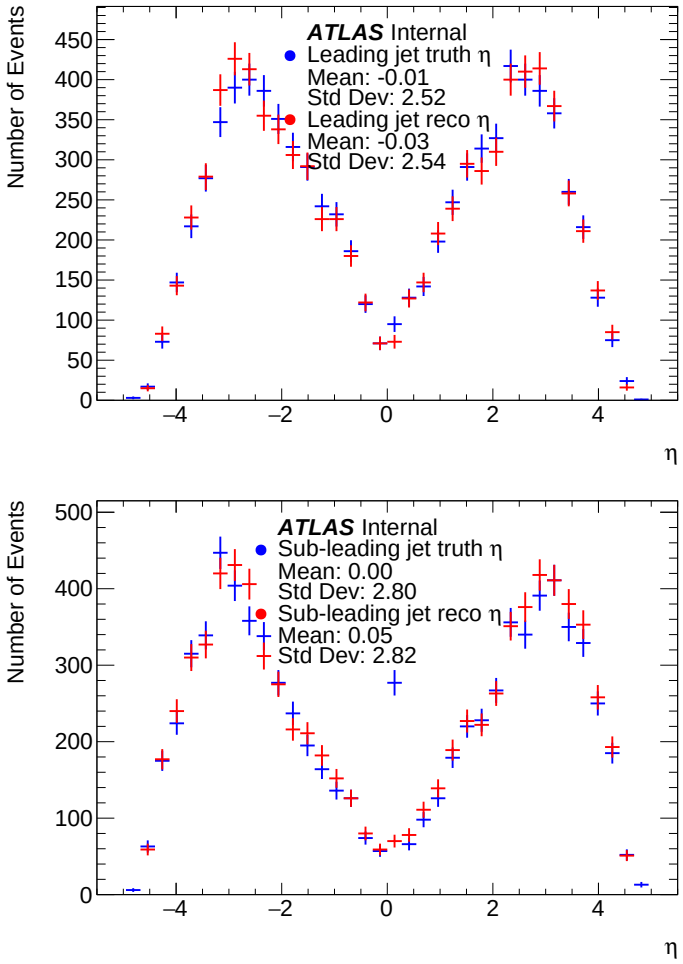

Figure 4.6: Pseudorapidities of the two leading jets on truth and reconstruction level after introducing additional kinematic cuts.

mismeasured, so that it falls below the 30 GeV threshold.
To conclude, there is still room for improvement in jet reconstruction. But since this is not the main focus of this work, the achieved level of precision, together with understanding the remaining problems, was considered sufficient.

4.2.3 Detector Resolution Effects

Comparing truth and reconstruction level data can also help estimate the effect of the finite detector resolution on the distributions of different measured observables. In figures 4.10 to 4.13, the resolution distributions of the jet momenta and the optimal observables are shown. They show the relative resolution $\frac{p_T^{\text{reco}} - p_T^{\text{truth}}}{p_T^{\text{truth}}}$ for the transverse momentum and the absolute resolution for the other quantities. To avoid influence from the misidentifi-

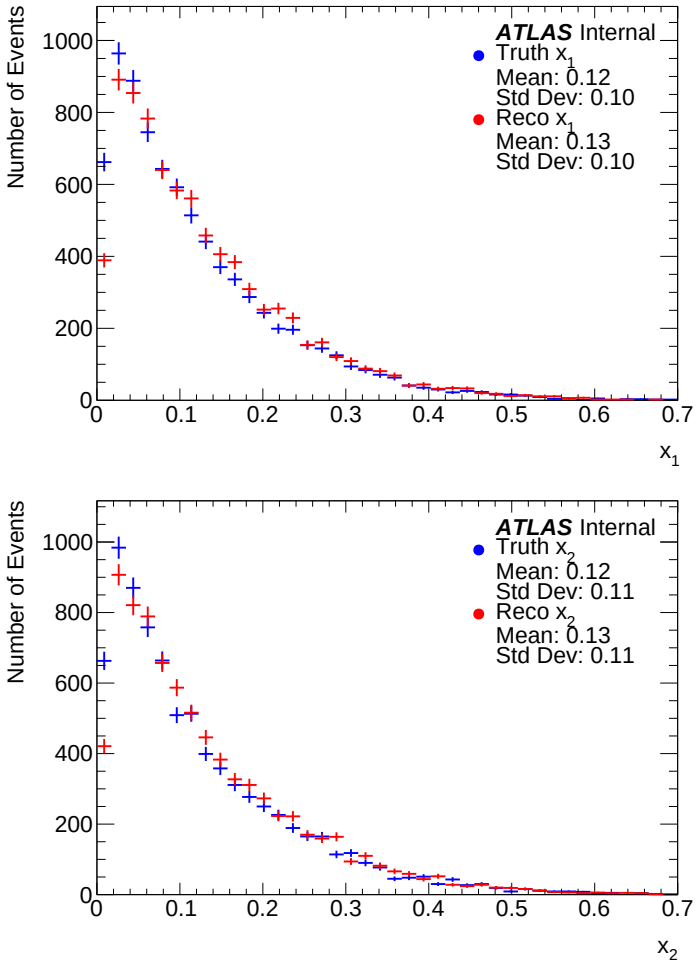

Figure 4.7: Bjorken x distributions for truth and reconstruction level events after introducing additional kinematic cuts.

cation of jets, a jet matching is required for this study. As in section 4.2.2, a truth and a reconstruction level jet are considered matched if they fulfill $\Delta R < 0.4$.
The p_T resolution in figure 4.10 shows that the resolution for the leading jet is better than for the sub-leading jet. This becomes apparent when comparing the standard deviations of the two distributions, and it shows that for higher jet p_T, the relative uncertainty of the measurement becomes smaller. Moreover, an asymmetry in the momentum measurement can be observed, which shifts the mean reconstructed p_T to higher values. This bias is introduced by applying a p_T cut on the reconstructed jets. If a truth jet with $p_T < 30$ GeV is reconstructed as a jet with p_T slightly above 30 GeV, the event is accepted. But in the reverse case, the event is rejected, which leads the p_T resolution distribution to have a longer tail in the positive direction. Since mostly jets with low p_T are affected by this,

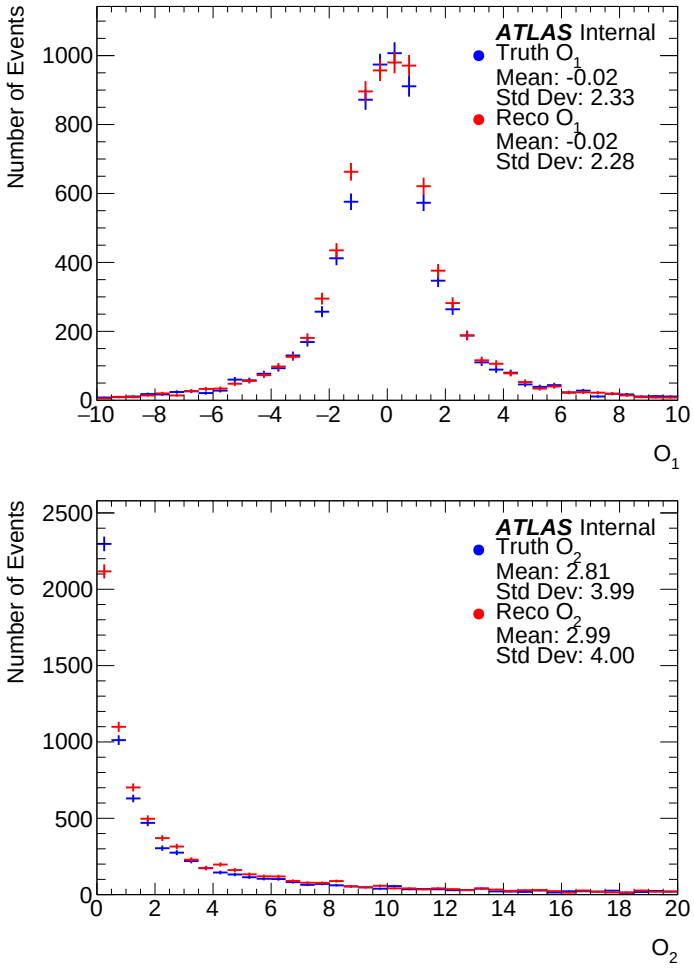

Figure 4.8: Optimal observable distributions for truth and reconstruction level events after introducing additional kinematic cuts.

the asymmetry is more apparent in the resolution distribution of the sub-leading jet.
The angular resolutions shown in figures 4.11 and 4.12 show no strong p_T dependence, as they are very similar for the leading and sub-leading jets. Moreover, they are symmetric, so the reconstruction of the angles is unbiased.
For the $\mathcal{O}$ resolution distributions in figure 4.13, the reconstruction level $\mathcal{O}$ distributions are calculated by using the four-vectors of the two jets that were matched to the two leading truth jets. They show that the reconstruction uncertainty is large, especially for $\mathcal{O}_2$, and that the $\mathcal{O}_2$ distribution is biased such that the reconstructed values are higher. The reason is suspected to be that the asymmetric shift in p_T between leading and sub-leading jets observed in figure 4.10 leads to less SM-like events. Since these asymmetries, however, do not favour one sign of $\mathcal{O}_1$, this effect only becomes apparent in the mean of

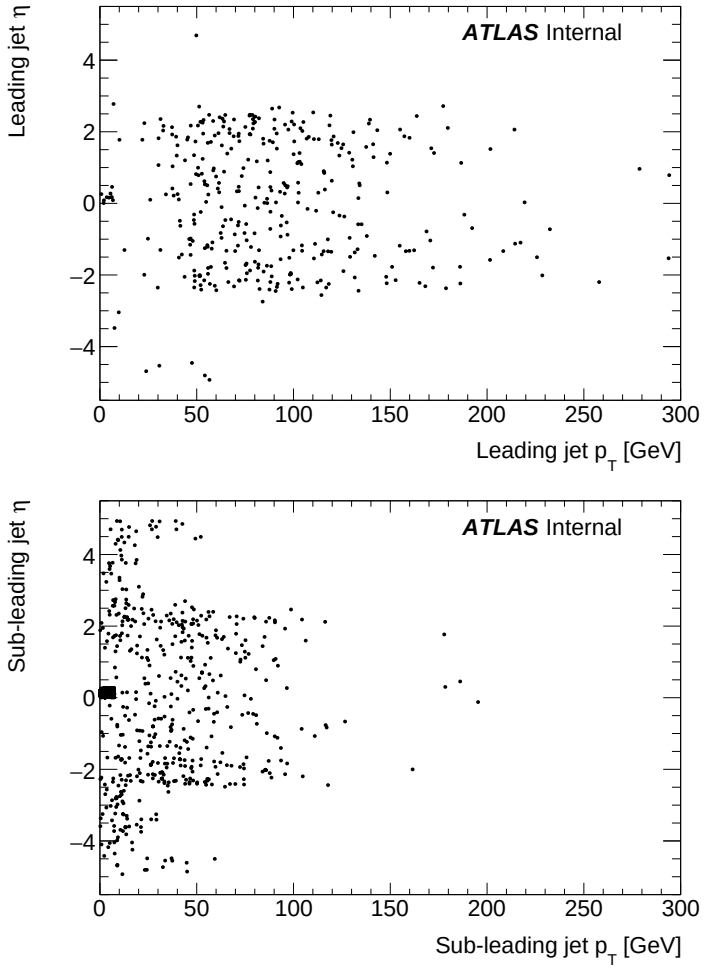

Figure 4.9: Scatter plot η vs p_T for leading and sub-leading truth jets, only for the events where no matching reconstructed jet was found. One can see that most of the unmatched jets have an absolute η value smaller than 2.5 or between 4 and 4.5, and mostly low p_T values.

the $\mathcal{O}_2$ resolution distribution. It increases the standard deviation of the $\mathcal{O}_1$ resolution distribution, but not the mean.
This concludes the comparison of truth and reconstruction level distributions. The simulated $H \to \tau_{\text{lep}}\tau_{\text{had}}$ kinematics are as expected of a SM sample, and discrepancies between truth and reconstruction level distributions can all be explained. Therefore, it qualifies for further use.

4.3 Reweighting Simulated VBF Events

As described in section 3.3.2, the strategy to test CP invariance of the HVV couplings is to find the CP-mixing scenario whose predictions best describe the experimentally observed

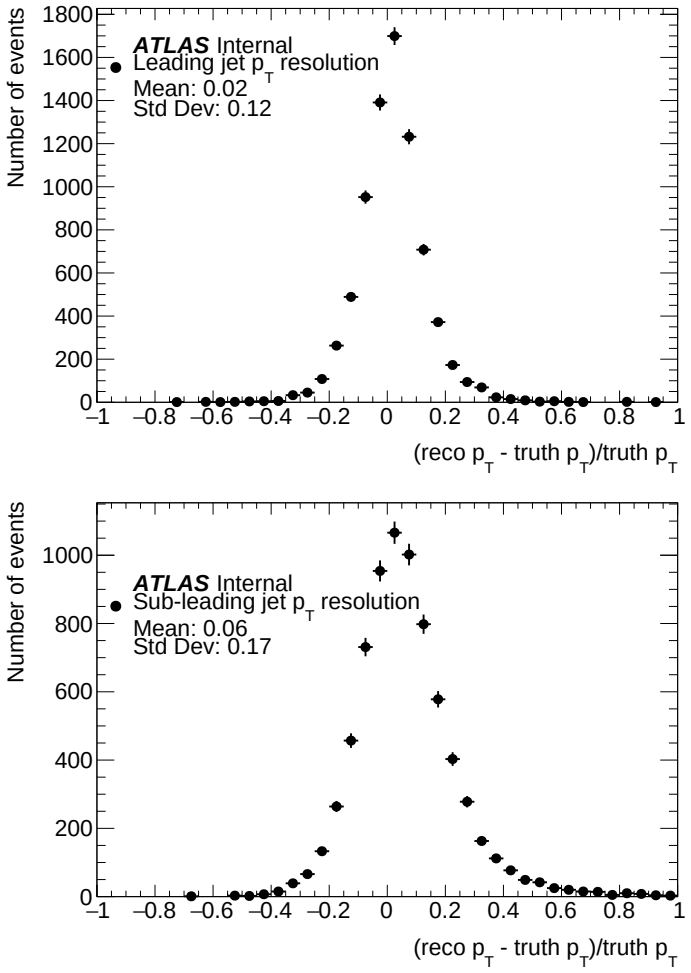

Figure 4.10: Relative resolution distributions of the transverse momentum of the leading and the sub-leading jet.

$\mathcal{O}_1$ and $\mathcal{O}_2$ distributions. Since computation power is limited, an event reweighting method is applied to simulated data in order to produce the $\mathcal{O}$ distributions predicted by scenarios with different $\tilde{d}$ values. This reweighting method has already been used for the ATLAS run 1 analysis on events with $\sqrt{s} = 8$ TeV [18], and in order to reutilise this method for the run 2 analysis, its suitability for use on events with $\sqrt{s} = 13$ TeV is tested in this section.

To that end, VBF event samples with non-SM $\tilde{d}$ values were generated. A SM sample is then reweighted to the same $\tilde{d}$ value by computing a matrix element ratio as described in section 3.3.2 and the agreement of the $\mathcal{O}$ distributions of the reweighted SM and the non-SM sample is tested.

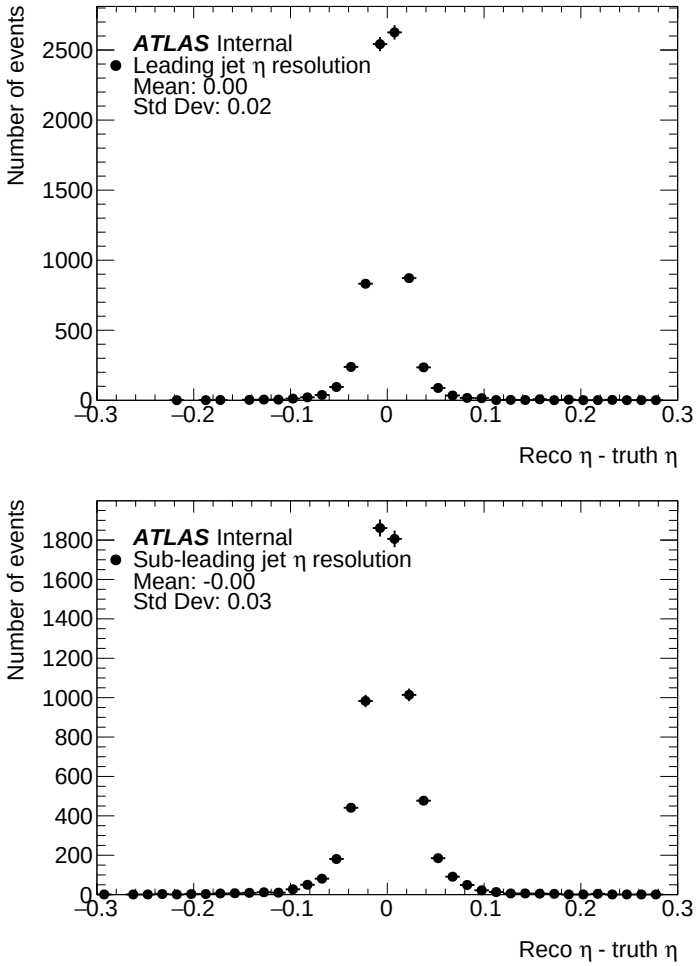

Figure 4.11: Resolution distributions of the pseudorapidity of the leading and the sub-leading jet.

4.3.1 Generation of VBF Events

The SM sample and the BSM events for this test were produced at LO using version 2.4.0 of the Monte Carlo generator "MadGraph5_aMC@NLO" [22]. One sample with SM couplings and $\tilde{d} = 0$ and three BSM sample with the $\tilde{d}$ values -0.2, 0.6 and 1.0 were produced. To simulate events with BSM HVV couplings, the Higgs characterisation (HC) model [20] was used, which uses the effective Lagrangian in equation 2.7 to parametrise the couplings. Hence, to specify the the $\tilde{d}$ parameter, the HC model parameters were calculated according to equation 2.8. Additionally, the couplings of the CP-odd Higgs state to photons was set to 10^{-4} for this test in order to be insensitive to the coupling of the CP-odd Higgs to top quarks which occurs in the effective coupling of the considered Higgs boson states to photons.

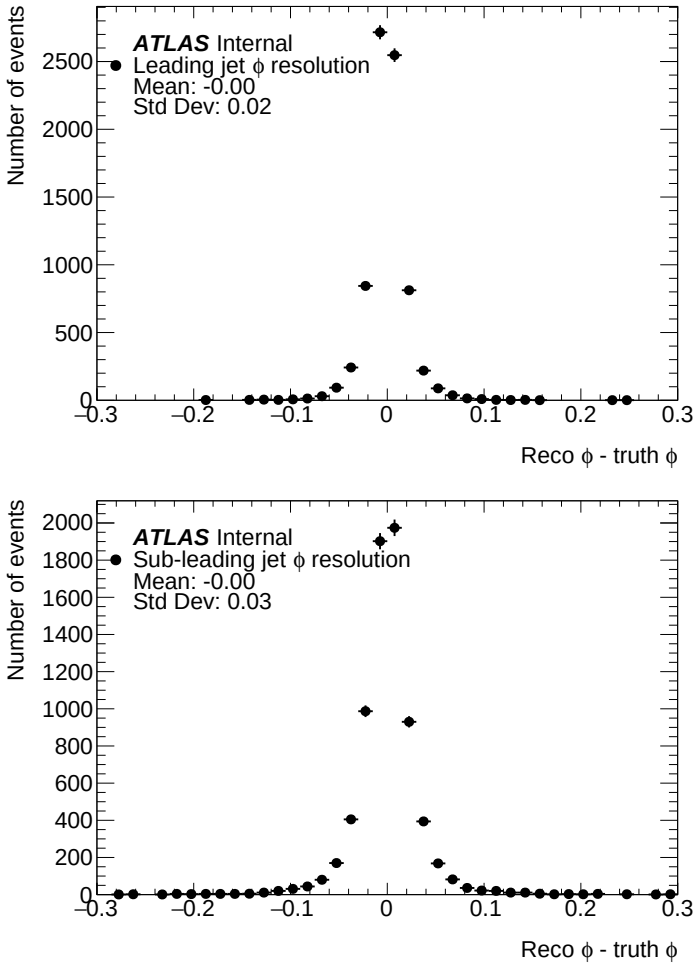

Figure 4.12: Resolution distributions of the azimuthal angle of the leading and the sub-leading jet.

In MadGraph, files called "parameter cards" contain all model-specific parameters used to generate events. Other parameters, such as the number of events to simulate, the centre-of-mass energy or kinematic cuts, can be set in "run cards". For the four samples with different CP-mixing scenarios used to test the reweighting, these parameter cards and the run card, which was almost the same for all samples, are shown in appendix A. The only difference in the run cards is that the SM sample contains 500000 events, while the BSM samples only contain 50000 events.
The samples produced this way contain LO generator level data, since no parton shower or detector simulation programme was applied.

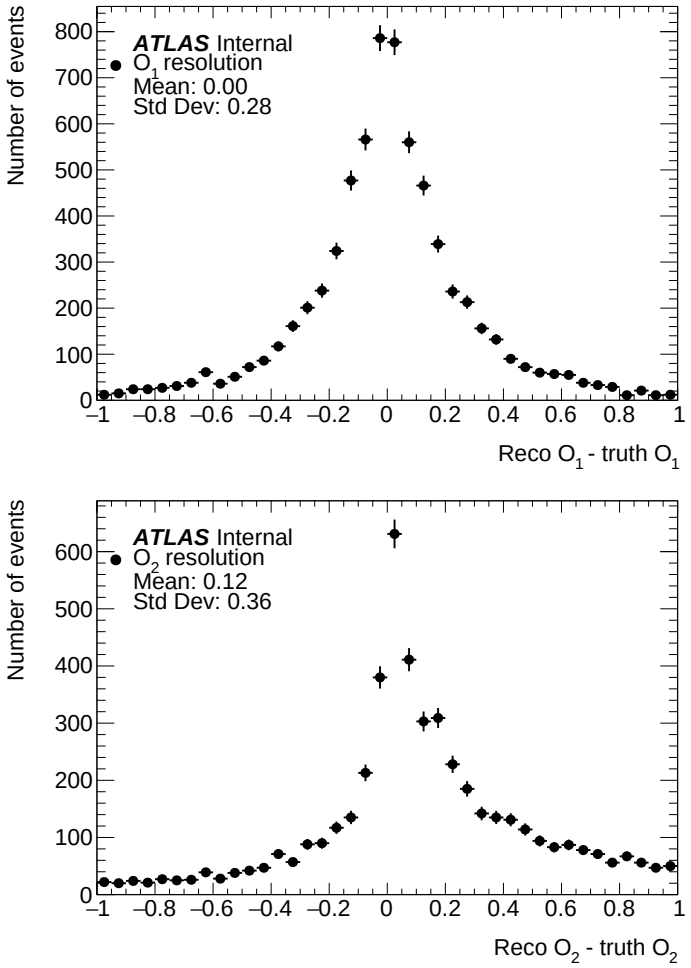

Figure 4.13: Resolution distributions of the two optimal observables.

4.3.2 Performance of the Reweighting Algorithm

The $\mathcal{O}$ distributions of the events described in the previous section are shown in the figures 4.14, 4.15 and 4.16, which show tests of the reweighting for the $\tilde{d}$ values -0.2, 0.6 and 1.0. The green distributions in these figures are SM distributions and the red ones are the same SM distributions reweighted in order to reproduce the blue BSM distributions. For a correct reweighting procedure, the red and blue distributions should be identical for all $\tilde{d}$ values. For a better comparison, below the $\mathcal{O}$ distribution plots, each bin of the red and blue distributions is divided by the content of the respective bin in the red distribution to show the ratio between the two distributions.

The reweighted distributions show good agreement with the BSM distributions, but to quantify the agreement of the distributions, Kolmogorov-Smirnov (KS) tests and χ^2 tests were conducted, and the test results are shown in the figures 4.14, 4.15 and 4.16. The

KS test statistic is a measure of the probability for both histograms to originate from the same probability distribution, which would mean that the differences between the reweighted and BSM distributions would only occur due to statistical fluctuations. The output of the ROOT KS test is uniformly distributed between 0 and 1, if both histograms originate from the same probability distribution, and is smaller for strongly deviating distributions. However, the expectation value of the KS test is only 0.5 when applied to continuous distributions. For histograms, it is biased towards higher values, so the expectation value of the KS test in this study is 0.56.
The χ^2 test is another test for comparing two histograms. It compares the histograms bin by bin, and from the deviations in these bins it calculates a test statistic χ^2 that has an expectation value equal to the number of degrees of freedom (d. o. f.) of the histograms, in this case the number of bins minus 1, which is 39. Deviations that are larger than expected lead to a χ^2 value that is higher than 39. The resulting χ^2 is also converted into a χ^2 probability. This is the probability that, under the assumption that both histograms are independently produced from the same probability distributions, replacing one of the histograms with a third histogram produced by this probability distribution would lead to a higher χ^2 value. Consequently, higher χ^2 probabilities mean better agreement.
It appears that the reweighting yields better results applied to $\mathcal{O}_1$ than to $\mathcal{O}_2$. Especially in the $\mathcal{O}_2$ distribution for $\tilde{d} = 0.6$, the test scores are low, which is, however, not a problem since individual low scores are expected. The average test results are $\chi^2 = 42.7$, 38 % for the χ^2 probability and a KS test output of 0.52. These results are all slightly worse than the expected values, which are $\chi^2 = 39$ and 0.56 for the KS test, but still consistent with them. In conclusion, this test shows that the reweighting algorithm can be used to simulate LO BSM Monte Carlo data at $\sqrt{s} = 13$ TeV.

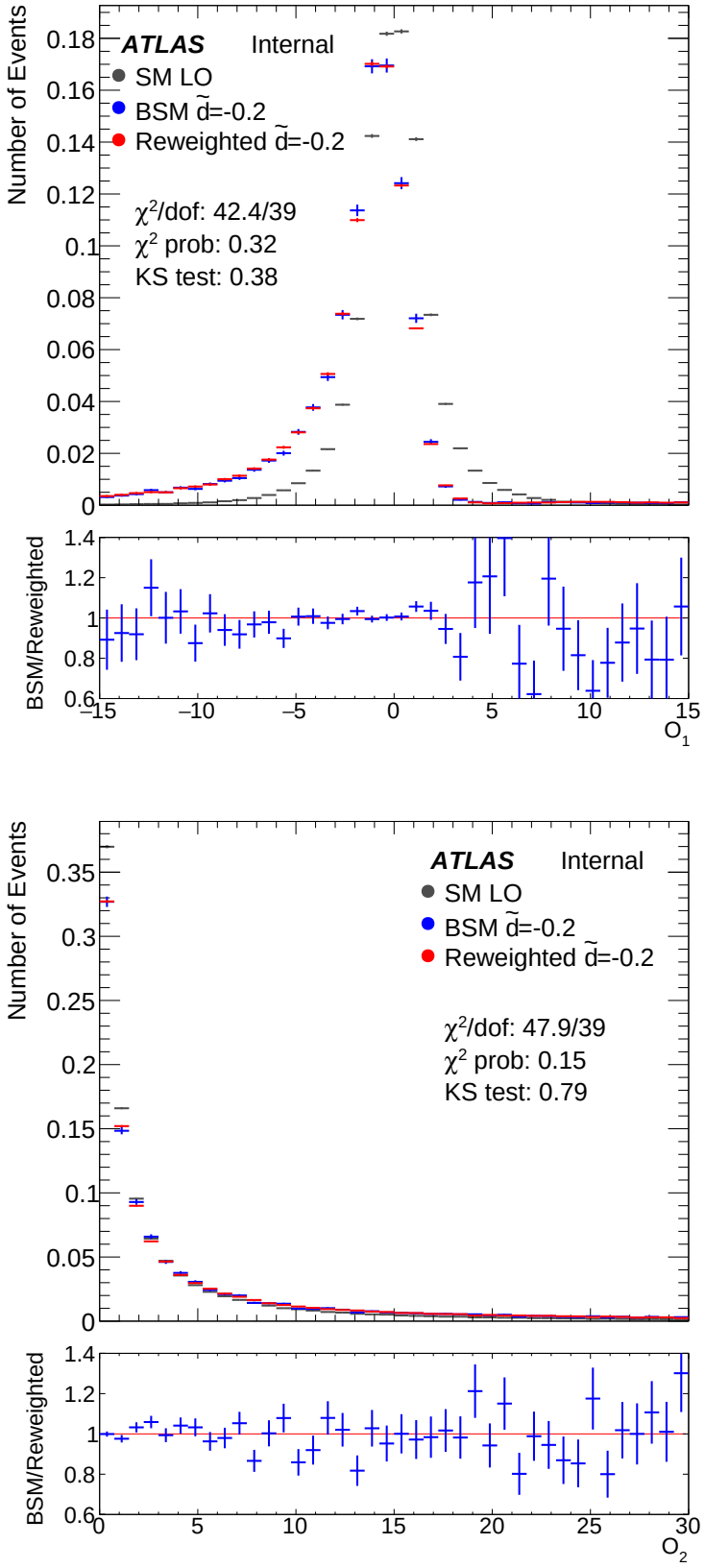

Figure 4.14: Test of the HAWK reweighting against a LO $\tilde{d} = -0.2$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = -0.2$.

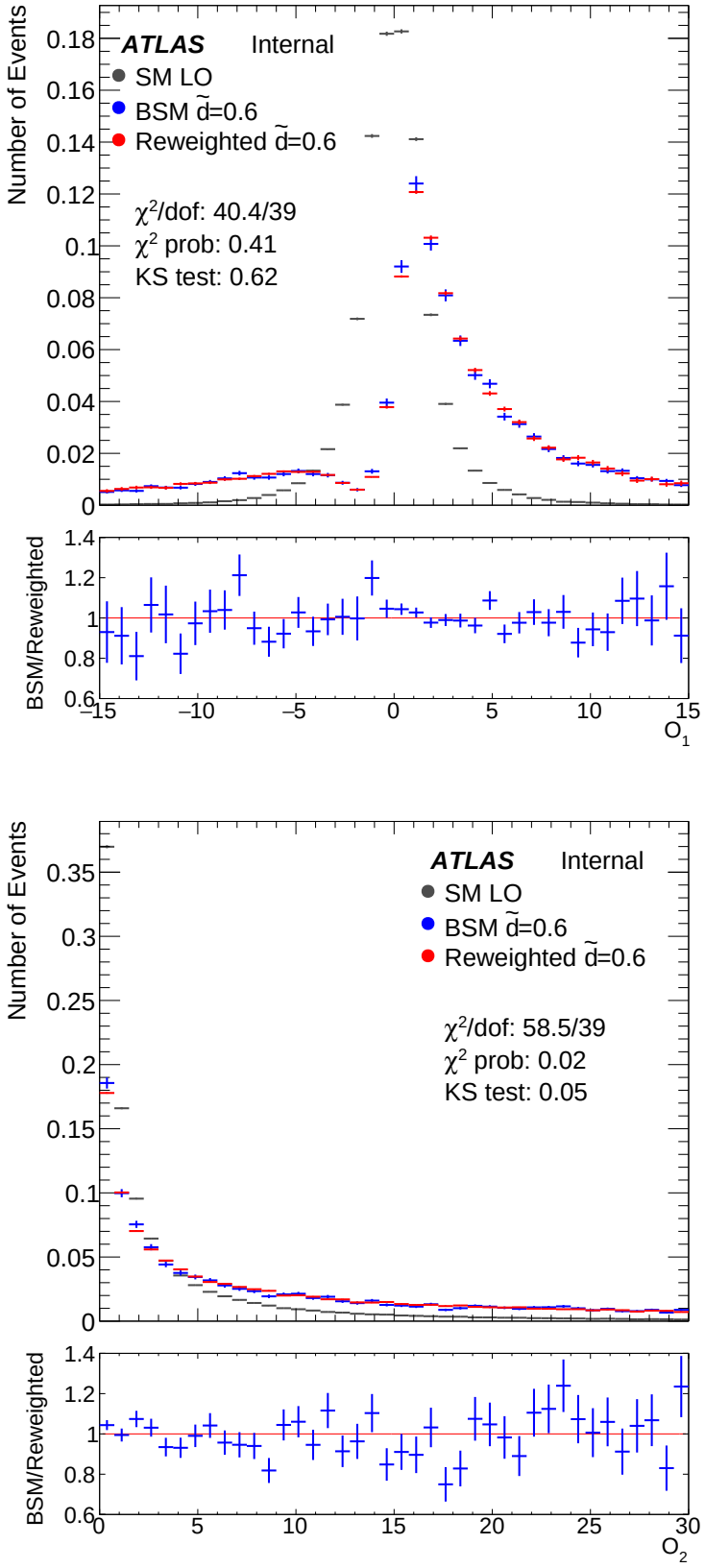

Figure 4.15: Test of the HAWK reweighting against a LO $\tilde{d} = 0.6$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = 0.6$.

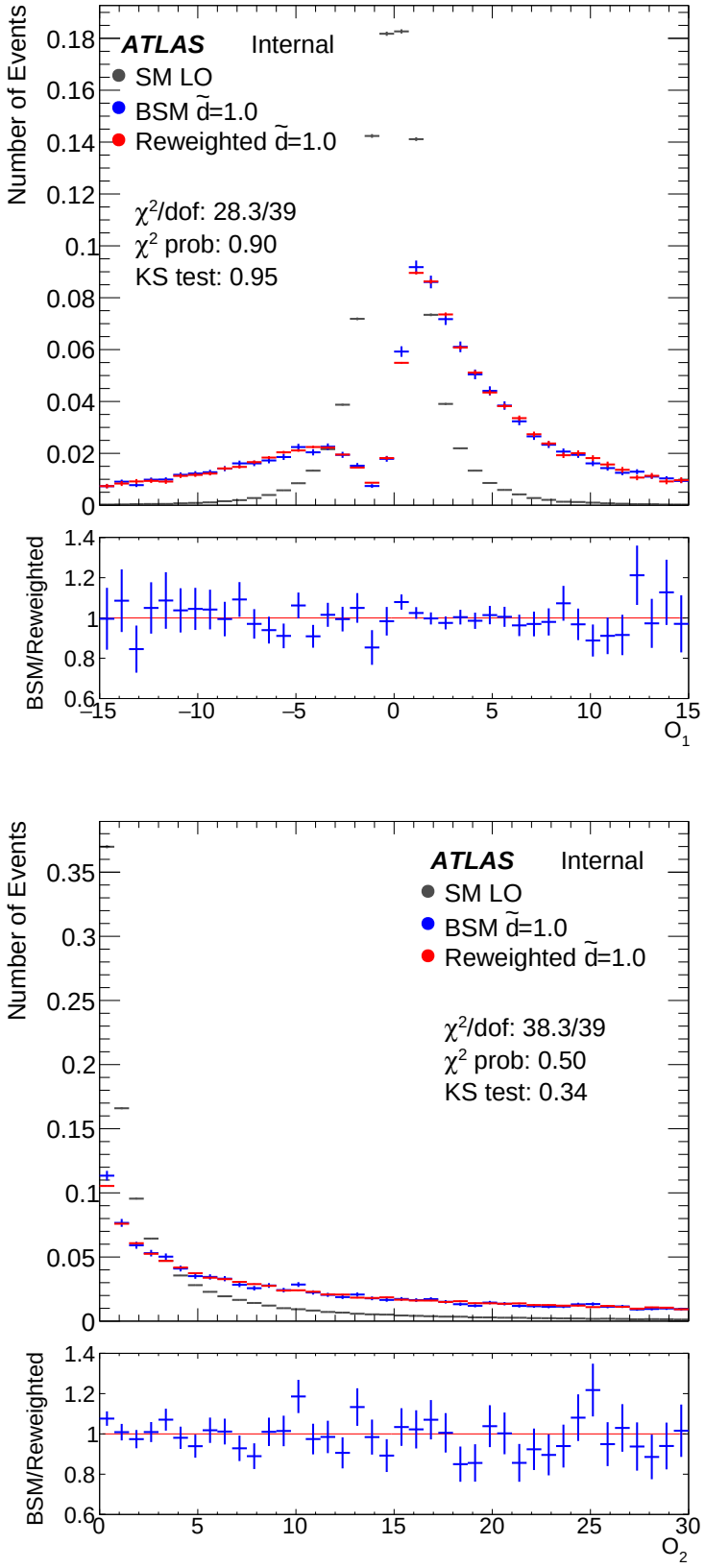

Figure 4.16: Test of the HAWK reweighting against a LO $\tilde{d} = 1.0$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = 1.0$.

5 Modified Reweighting on NLO Events

The test to validate the reweighting that is described in the previous chapter was conducted on LO events. To simulate a more realistic signal process, and since the signal samples in ATLAS analyses are generated at NLO, the same reweighting routine was also tested on NLO events, albeit with necessary modifications.
This chapter will describe why modifications are necessary and possible approaches for these modifications. These approaches are then compared, and using the most performant approach, the reweighting method is validated at NLO.

5.1 Interpreting the MadGraph Event Record

To calculate the matrix elements needed for the optimal observables and event weights, HAWK must be given the four-vectors of two VBF partons. When using generator level LO events, each event only has two VBF partons, and their four-momenta can be passed to the HAWK routine to calculate the event weights, as described in section 4.3.2. At NLO, however, there can be three partons, and one can only pass two of these to HAWK. In this case, one must decide what information to use for the calculation and what to neglect. To make that decision, the MadGraph event record was examined. Figure 5.1 shows two Feynman diagrams of processes that describe events that are indistinguishable only based on information in the event. The reason for that is that the event record only contains information on the initial state and final state particles, but not on the intermediate particles. However, this distinction is necessary in order to find out which quarks emitted the vector bosons that fused to produce the Higgs boson. In figure 5.1(left), these were two down quarks, and the two VBF partons are the up and anti-up quarks, while the final state down quark did not contribute to the Higgs boson production. In the right diagram, however, the anti-up quark did not contribute to the Higgs boson production. This ambiguity is not limited to the two specific processes shown in figure 5.1. In fact, no VBF NLO event can only be associated to one diagram since no distinction is made

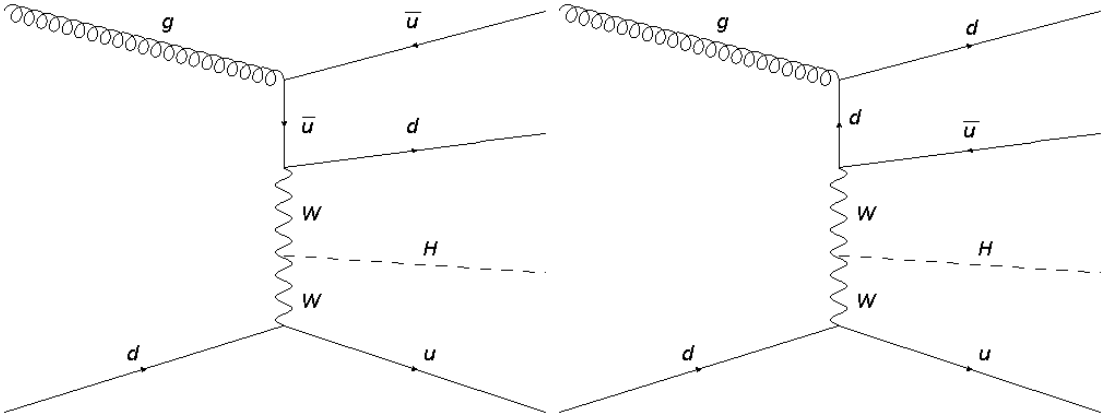

Figure 5.1: Feynman diagrams of two different VBF processes. Each of these processes has the same initial and final state particles. Therefore, they cannot be distinguished. This means that one cannot unambiguously determine the two final state quarks that coupled to the vector bosons and participated in the VBF Higgs production.

between initial and final state gluon radiation.
Consequently, simply choosing two final state partons in NLO events and passing them to the HAWK routine in order to calculate the weights for NLO events is not necessarily correct. Therefore, one needs to decide on a method to combine the information of the three final state partons into two four-vectors.

5.2 Parton Combination Methods

The MadGraph event record contains the information on the colour charge of each particle. Using that, one can determine the final state parton that did not couple to gluons in this event. In figure 5.1, this is the final state up quark for both diagrams. This particle always participates in the VBF Higgs production. Therefore, its four-vector should be one of the four-vectors to be used for the $\mathcal{O}$ and event weight calculations. This method is unambiguous, since the different colour charges in the event records are expressed through numbers. Even if two final state partons carry the same colour charge, these charges are expressed through different numbers.
Regarding the information of the other two partons, there are several ways to combine the information on them to construct the second four-vector that should be used. The following three methods were tested:

- "sum" method: Add the two four-vectors of the remaining partons and use the sum.
- "avg" method: Use only one of the two four-vectors at a time to calculate the

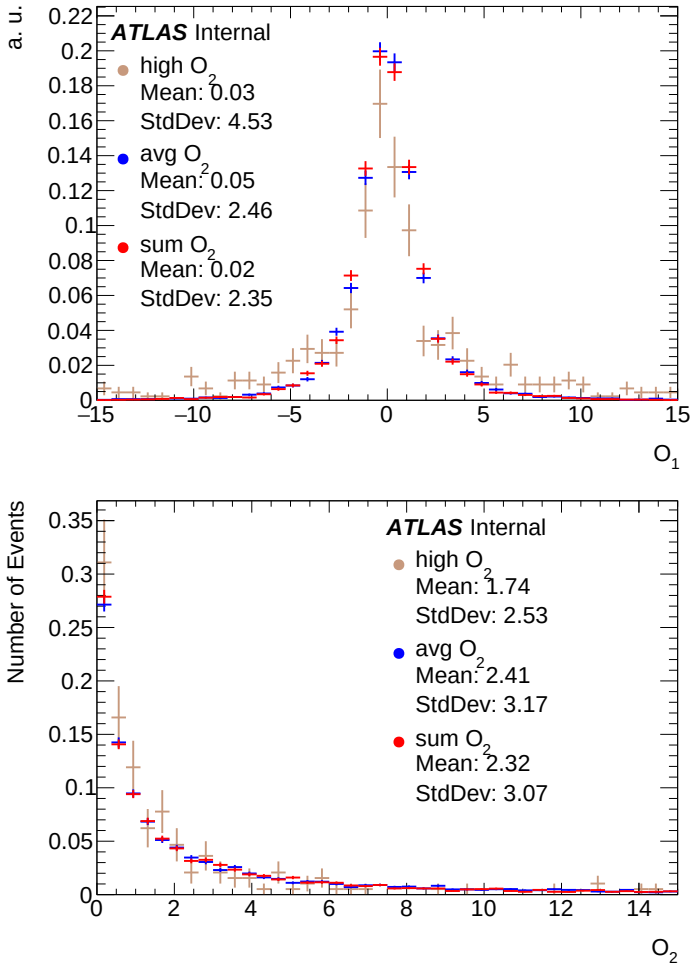

Figure 5.2: $\mathcal{O}$ distributions of NLO events using the different parton combination methods. The distributions look similar for the "avg" and "sum" methods, while the "high" method produces larger uncertainties and a flatter distribution.

optimal observables, use the average of these values as the final result.

- "high" method: Use only the four-vector of the parton with higher p_T, neglect the other one.

To compare these methods, an NLO SM generator level VBF sample was generated with MadGraph using the same parameters as for the SM LO sample used for the tests in section 4.3.2. The $\mathcal{O}$ distributions obtained by applying these methods on the SM NLO sample are shown in figure 5.2. This comparison disfavours the "high" method, since it produces a distribution with a higher standard deviation than the other two methods, which makes the uncertainty on the calculation larger. This would mean a decreased sensitivity for measuring $\tilde{d}$ and the test of CP invariance. The other two methods look

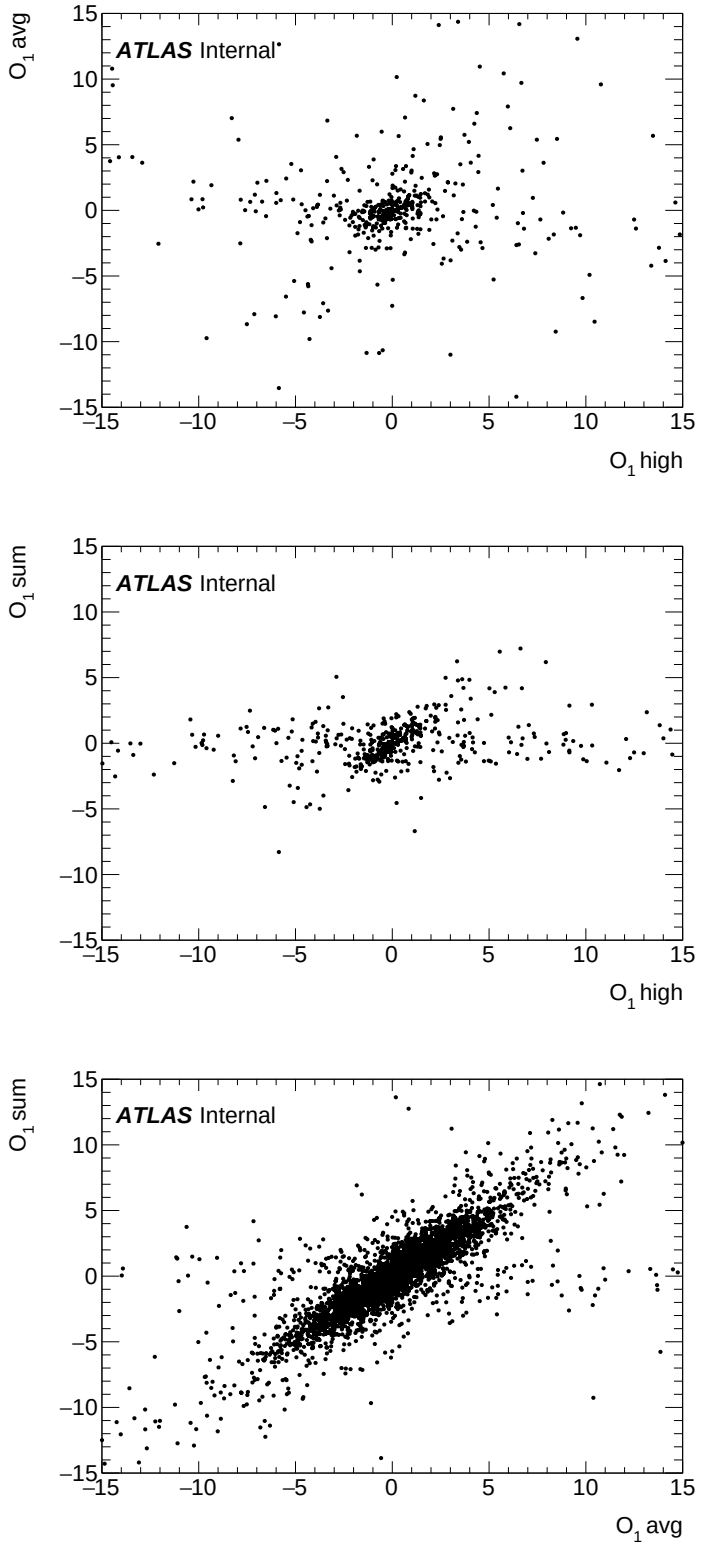

Figure 5.3: Scatter plots comparing the $\mathcal{O}_1$ distributions of the different methods. While the "avg" vs "high" scatter plot does not show a distinct pattern, the other two plots feature a horizontal line where the "sum" method yields $\mathcal{O}_1 \approx 0$ and the values from the other methods vary.

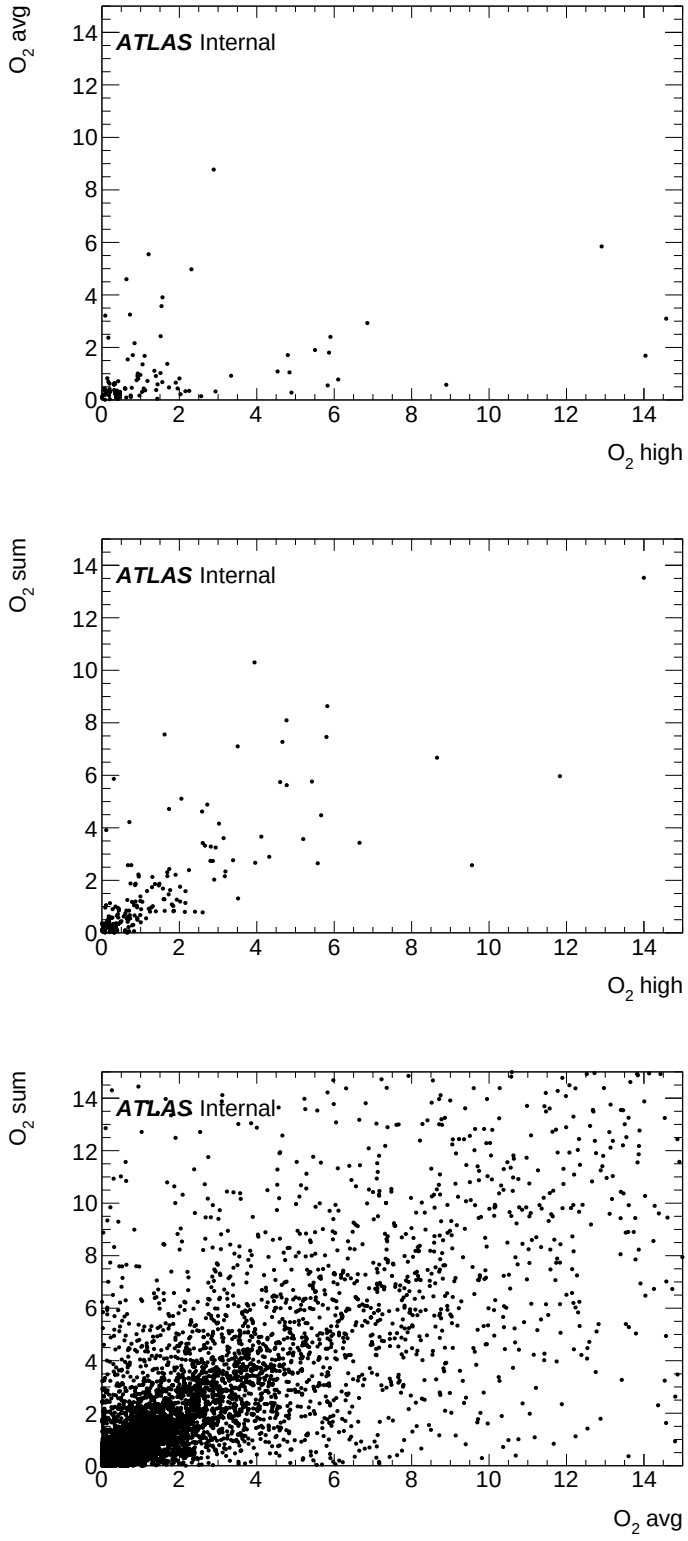

Figure 5.4: Scatter plots comparing the $\mathcal{O}_2$ distributions of the different methods.

similarly well-suited.
Therefore, the three methods were additionally compared in 2D scatter plots, which is shown in figures 5.3 and 5.4, where the $\mathcal{O}$ distributions from different methods are compared. Figure 5.3 does not show a strong correlation between the $\mathcal{O}_1$ values of the "high" and the "avg" method. The plot comparing the "high" and the "sum" method, however, shows that the values produced by the "high" method vary strongly in the shown range from -15 to 15, while the values from the "sum" method are mostly close to zero. The values are expected to be around zero, since the test is conducted on a SM sample. Thus, the "sum" method produces more realistic $\mathcal{O}_1$ values than the "high" method. The last plot in figure 5.3 shows that, for many events, the "sum" and "avg" methods yield similar results, since most entries are on the diagonal of the plot. The entries that are not on the diagonal, however, almost all have $\mathcal{O}_1 \approx 0$ for the sum method, while the values for the "avg" method vary. Consequently, the "avg" method produces unexpectedly high absolute values for $\mathcal{O}_1$ in some events. This is also reflected in the mean value of the distributions in figure 5.2 where the "sum" method has a slightly lower value than the other methods.
The scatter plots of the $\mathcal{O}_2$ values are shown in figure 5.4. With the same reasoning, one can argue that the methods producing lower $\mathcal{O}_2$ values perform better, since these are more SM-like. However, none of the plots in figure 5.4 shows a strong tendency for any method to produce lower $\mathcal{O}_2$ values than the other ones, since none of the plots shows significantly more or less entries above the main diagonal than below it. Therefore, the $\mathcal{O}_2$ scatter plots do not favour any of the methods.
The results seen in figures 5.3 and 5.4 can be explained with the fact that the "high" and "avg" methods ignore parts of the final state, which leads the partons used for the HAWK routine to no longer balance the momentum of the Higgs boson. The SM expectations for the mean of $\mathcal{O}_1$ and $\mathcal{O}_2$ are low values, while these methods with the non-SM kinematic distributions they use then produce higher $\mathcal{O}$ values. Conversely, the "sum" method takes the whole final state into account, which leads to lower $\mathcal{O}$ values overall. Another argument in favour of the "sum" method is that, for events with final state radiation, applying the "sum" method incorporates this radiation, yielding the final state that would be produced if the same process was generated at LO. This is the ideal case, since the reweighting algorithm that is to be tested was devised for LO events. For these reasons, the "sum" method will be used in the following.

5.3 Performance of the Reweighting Algorithm at NLO

The test of the reweighting at NLO is conducted analogously to the LO test described in section 4.3.2. Again, one SM sample and three BSM samples with coupling parameters corresponding to the $\tilde{d}$ values -0.2, 0.6 and 1.0 are produced on generator level, with the same parameter cards as the LO samples. These cards can be found in appendix A. The only difference is that the samples for this test are generated at NLO, allowing for an additional gluon vertex. In events with three final state partons, one of them will be added to one of the other two using the "sum" method introduced in section 5.2.
Figures 5.5, 5.6 and 5.7 compare the reweighted $\mathcal{O}_1$ and $\mathcal{O}_2$ distributions of the SM sample with and without reweighting to the $\mathcal{O}$ distributions in the BSM samples for the values -0.2, 0.6 and 1.0 in $\tilde{d}$. As in the LO case, the agreement of the BSM and the reweighted SM distributions was tested with KS and χ^2 tests, the results of which one also finds in the figures. As described in section 4.3.2, the expectation value for χ^2 is the number of degrees of freedom, for the χ^2 probability it is slightly below 0.5 depending on the number of degrees of freedom, and for the KS test it is 0.56. For χ^2, lower values indicate better agreement between the reweighted SM and the BSM distributions, while for the χ^2 probability and the KS test, the values increase with better agreement.
The worst results were achieved for the $\mathcal{O}_1$ distribution with $\tilde{d} = 1$, where the χ^2 probability is only 14%. Apart from that distribution, however, the distributions are modelled very well. The average result is $\chi^2 = 39.7$ with a probability of 46 %.This is slightly worse than the expectation value of 39, but still consistent with it. The average KS test result is 0.86, which is better than expected. This validates the reweighting method at NLO at $\sqrt{s} = 13$ TeV, and it also shows that using the "sum" method to make the reweighting algorithm applicable to NLO events was a good choice.

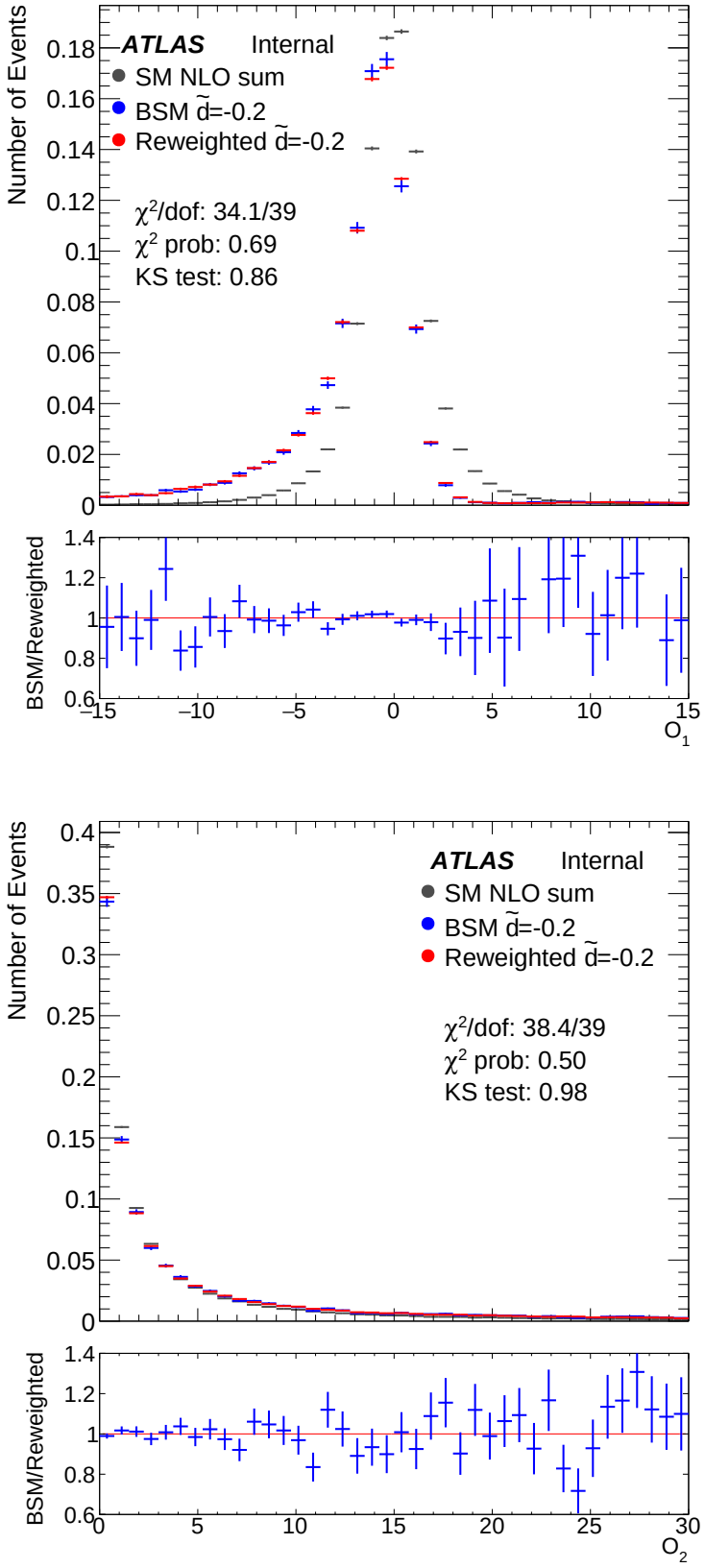

Figure 5.5: Test of the HAWK reweighting against an NLO $\tilde{d} = -0.2$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = -0.2$.

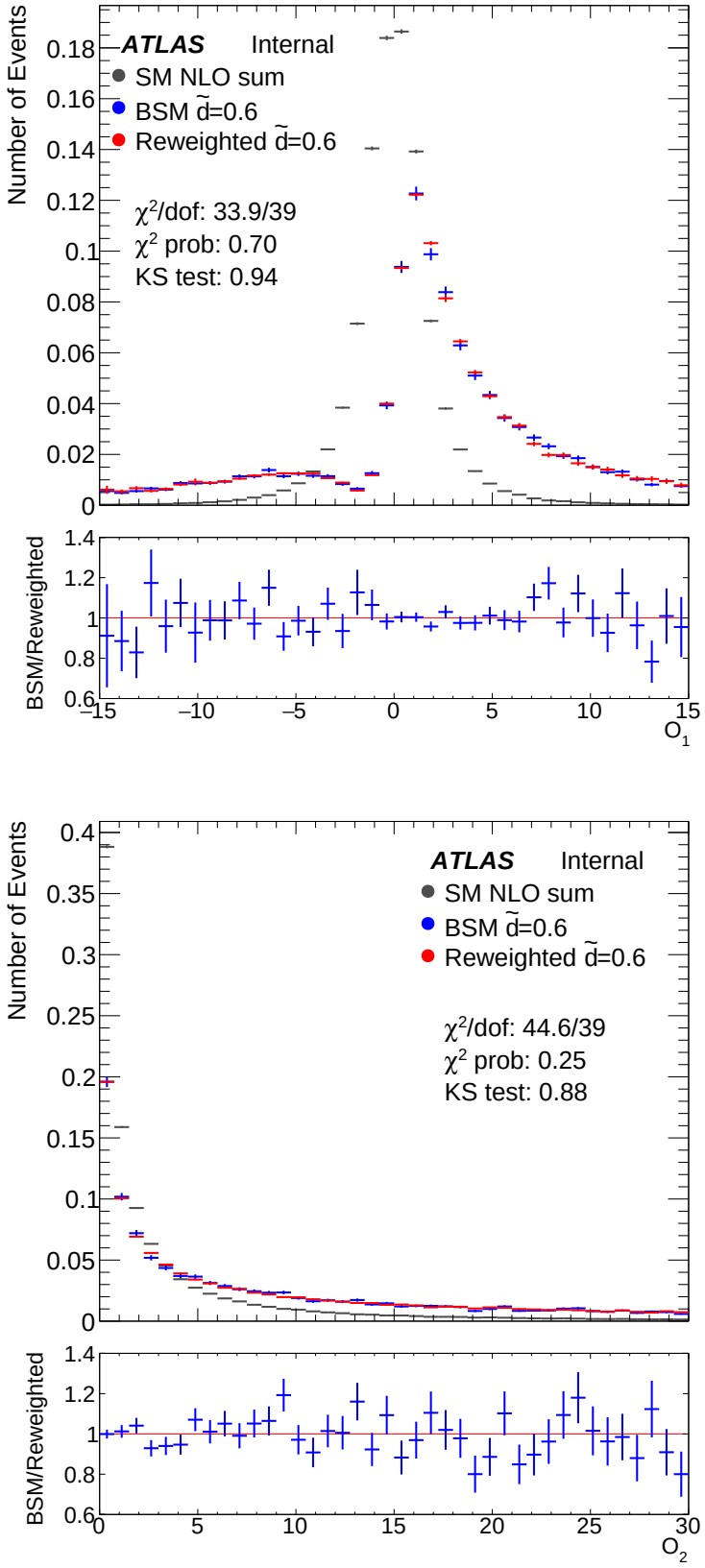

Figure 5.6: Test of the HAWK reweighting against an NLO $\tilde{d} = 0.6$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = 0.6$.

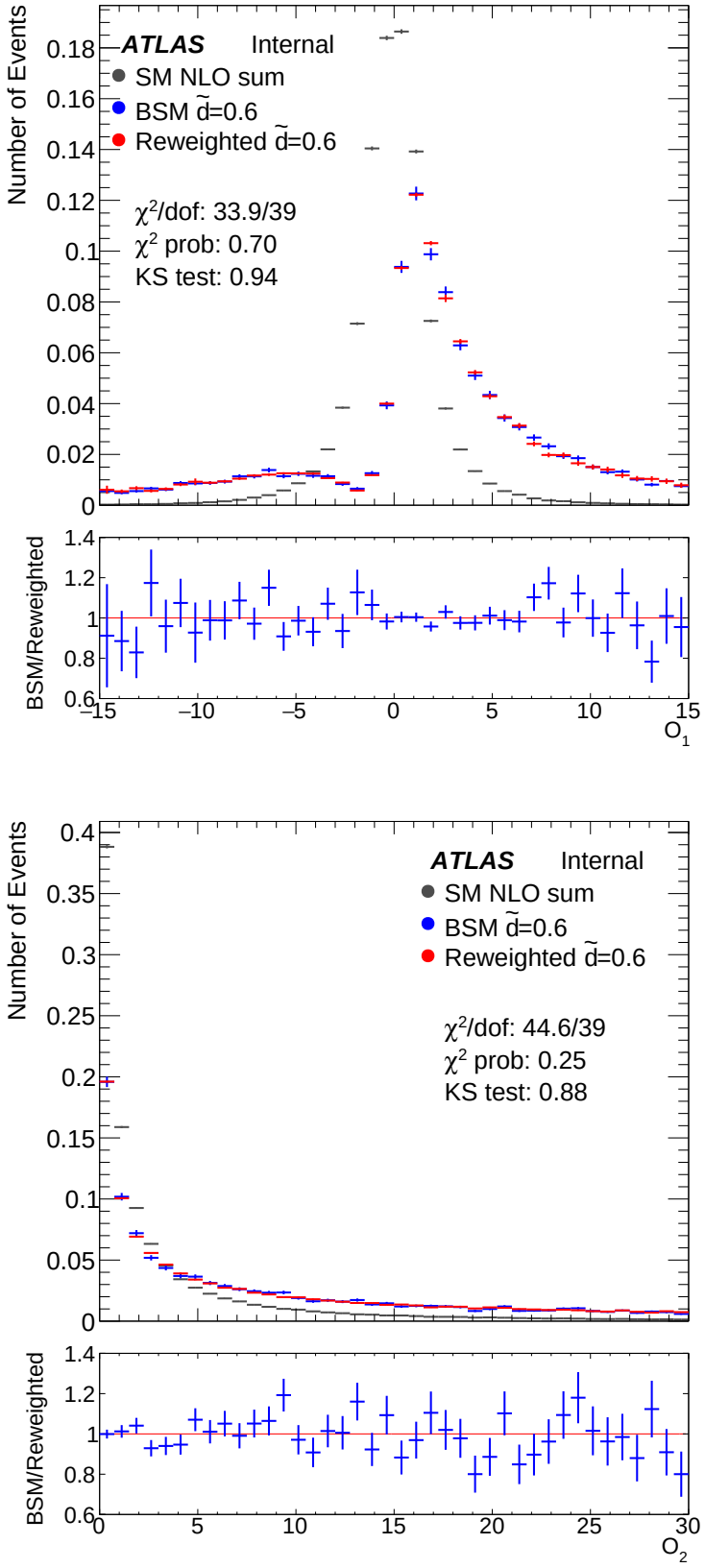

Figure 5.7: Test of the HAWK reweighting against an NLO $\tilde{d} = 1.0$ sample. The green and red distributions are obtained from a SM sample, while the blue distributions are from a non-SM sample generated with $\tilde{d} = 1.0$.

6 Jet Scaling Studies for the Optimal Observable

As described in section 3.3.2, the method used by HAWK to process information on the Higgs four-vector is expected to yield a more exact calculation of $\mathcal{O}_1$ and $\mathcal{O}_2$ in Higgs decay channels that cannot be reconstructed with high precision such as $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ or $H \rightarrow \tau\tau$. The reason is that, in order to achieve a p_T balance between the jets and the Higgs boson, HAWK uses the jet momenta to calculate an expression that replaces the Higgs boson momentum. This procedure might, however, lead to a loss in precision for events in which the Higgs boson decays into particles such as photons that can be reconstructed with a better energy resolution than the jets. This chapter describes different methods to scale the jets according to the reconstructed Higgs boson momentum in order to achieve more precise results in these cases.

6.1 Simple Jet Scaling Algorithm

The goal of these scaling algorithms is to scale the transverse momentum of the two leading jets using information from the p_T of the Higgs boson. If the uncertainty on the Higgs p_T is smaller than that on the jet p_T, such an algorithm should increase the accuracy of the jet p_T, which should produce narrower $\mathcal{O}$ distributions. The VBF $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ sample described in section 4.1 is used to test the performance of such algorithms. For this test, a gauge curve is constructed for each algorithm to compare their sensitivity.

6.1.1 Gauge Curves and Sensitivity Estimate

Gauge curves (GC) are a mapping between the CP violation inducing parameter $\tilde{d}$ and the mean of the $\mathcal{O}$ distributions of the BSM model that has Higgs boson couplings as defined by the chosen $\tilde{d}$ value. The gauge curves used in this test are constructed in the following way on a SM MC sample. At first, the SM $\mathcal{O}_1$ and $\mathcal{O}_2$ distributions are determined. Then a parameter range for considered $\tilde{d}$ values and a step width in $\tilde{d}$ are

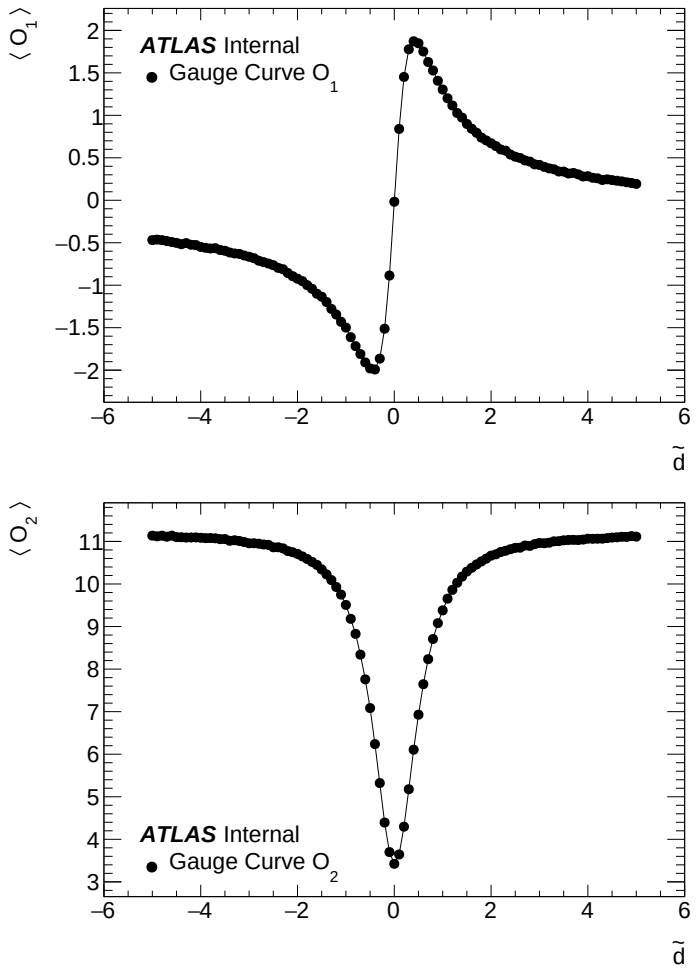

Figure 6.1: Simple gauge curves for the optimal observables with no uncertainty estimation. They show the dependence of $\langle \mathcal{O}_1 \rangle$ and $\langle \mathcal{O}_2 \rangle$ from the CP violation inducing parameter $\tilde{d}$ in the range $|\tilde{d}| \leq 5$. The curves were produced using the MadGraph LO SM sample described in section 4.3.1.

chosen. The SM $\mathcal{O}$ distributions are reweighted to each of these $\tilde{d}$ values, and the mean values of these distributions are plotted against the $\tilde{d}$ that they were reweighted to. The results for the means of the $\mathcal{O}_1$ and $\mathcal{O}_2$ distributions can be seen in figure 6.1, where the range $|\tilde{d}| \leq 5$ is chosen with a step width of 0.1. These curves were produced using the MadGraph LO SM sample described in section 4.3.1. In general, the $\mathcal{O}_1$ gauge curve is point-symmetric around the origin, and the $\mathcal{O}_2$ curve is axisymmetric with respect to the $\tilde{d} = 0$ axis apart from statistical fluctuations. The cuts applied are the same as in section 4.2.1: $p_{T,\text{lead}}$, $p_{T,\text{sub-lead}} > 30$ GeV, $\eta_{\text{lead}} \cdot \eta_{\text{sub-lead}} < 0$ and $|\eta_{\text{lead}} - \eta_{\text{sub-lead}}| > 3$. These are also the cuts that will be used for producing the following gauge curves in this chapter. Since the $\mathcal{O}_1$ and $\mathcal{O}_2$ gauge curves are almost flat for high values of $|\tilde{d}|$, it is apparent that

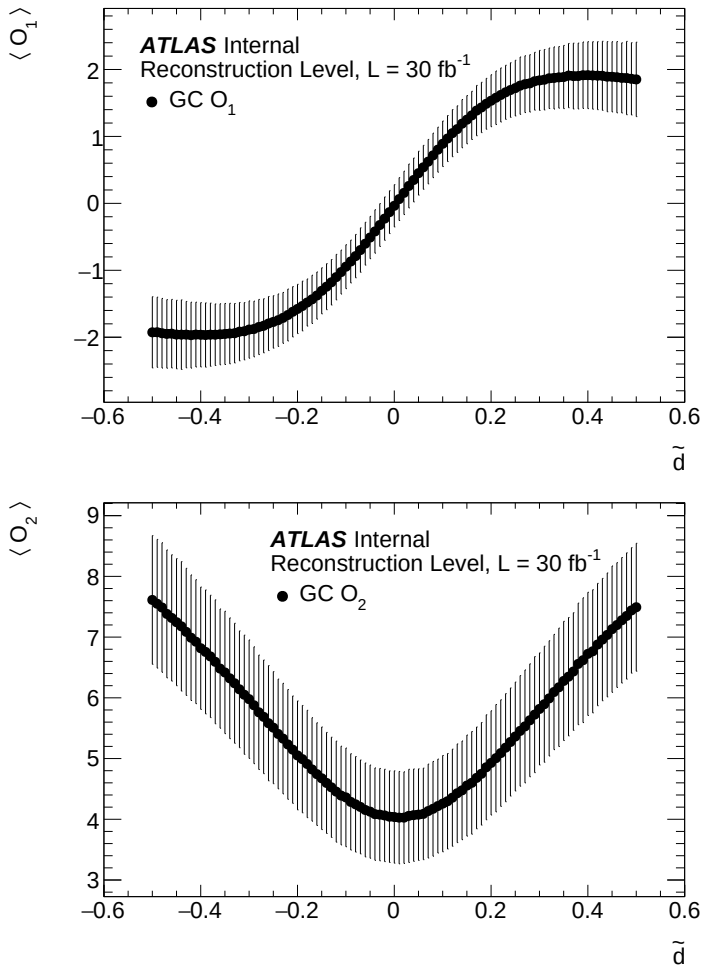

Figure 6.2: Gauge curves for the reconstruction level optimal observables on the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample scaled to an integrated luminosity of 30 fb^{-1}. The error bars are the smallest around $\tilde{d} = 0$.

the optimal observables are most effective at distinguishing CP mixing scenarios with small $|\tilde{d}|$ values. However, this is not a problem since the data collected so far are consistent with the SM Higgs boson [6, 7], and consequently a small value for $\tilde{d}$ is expected. Therefore, the gauge curves used in the following will only cover the range $|\tilde{d}| < 0.5$.
To find uncertainties of the points making up a gauge curve, the Neyman construction [42] is applied. The first step is choosing a luminosity to which one wishes to scale the uncertainties. Then the expected number of VBF events to be measured using the ATLAS detector with the chosen luminosity is calculated, and the reweighted $\mathcal{O}_1$ and $\mathcal{O}_2$ distributions are scaled such that their integral equals this number of events. The statistical uncertainties on each bin are calculated as Gaussian uncertainties, not Poisson. This is not a problem, since the mean number of scaled and generated events is in the regime of large

numbers. In the next step, so-called "pseudodata" are created by using the reweighted $\mathcal{O}$ distributions to generate a large number of $\mathcal{O}$ distributions with the same number of entries and the same $\tilde{d}$ values. For each $\tilde{d}$, the value inserted into the gauge curve is now the mean $\langle\mathcal{O}\rangle$ of all the pseudo-distributions generated with the corresponding reweighted $\mathcal{O}$ distribution. The uncertainty on this mean value is half the length of the smallest interval symmetric around $\langle\mathcal{O}\rangle$ that contains 68% of the means of all pseudo-distributions. For the default HAWK calculation and the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample used thus far, the resulting reconstruction level gauge curves scaled to the expected number of signal events in a luminosity of 30 fb^{-1} can be seen in figure 6.2, where the so-called "Neyman confidence belt" surrounds the gauge curve. The expected number of events in the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ channel is 74.1. This number is the product of the VBF Higgs production cross section of 3.75 pb, the branching ratio of 0.063 of a 125 GeV SM Higgs boson to two tau leptons, the branching ratios for one tau lepton to decay hadronically and one leptonically times two, which is 0.46, the assumed luminosity of 30 fb^{-1} and the simulated reconstruction efficiency from the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample examined in section 4.1. The efficiency is calculated as the ratio of the number of events in the sample that fulfill all the derivation and event selection cuts and the initial sum of Monte Carlo events that were generated to produce the sample file, which equals approximately 0.023. Note that the uncertainties given throughout this chapter contain only a statistical component, and distributions from signal-only event samples are considered.

These gauge curves can be used to determine the 68% confidence interval for $\tilde{d}$ that follows from the observation of a certain $\mathcal{O}_{\text{obs}}$ value. The boundaries of the confidence interval are the $\tilde{d}$ values where the line $\mathcal{O} = \mathcal{O}_{\text{obs}}$ intersects with the confidence belt. In case there are only two intersections, this method leads to exactly one confidence interval. If there are, however, four intersections, which is more likely for the $\mathcal{O}_2$ gauge curve since it has a minimum, the 68% region for $\tilde{d}$ would consist of two disjoint intervals.

For figure 6.2, measuring exactly the SM prediction of $\mathcal{O}_1 = \langle\mathcal{O}_1\rangle_{\tilde{d}=0}$, meaning the value for $\langle\mathcal{O}_1\rangle$ in the gauge curve at $\tilde{d} = 0$, would lead to the 68% confidence level $\tilde{d} \in [-0.030, 0.038]$. The analogous calculation for $\mathcal{O}_2$ leads to $\tilde{d} \in [-0.181, 0.196]$. The size of this interval is a measure for the sensitivity of the analysis towards CP violation. The intervals are slightly asymmetric due to finite MC statistics, which leads to statistical fluctuations in the $\mathcal{O}_1$ and $\mathcal{O}_2$ distributions.

To test the accuracy of a jet scaling method presented in the following compared to the default calculation, the lengths of the confidence intervals resulting from the gauge curves of the different methods are compared.

6.1.2 Performance of the Jet Scaling

This section will discuss a method to make the p_T of the two leading jets equal to the p_T of the Higgs boson by scaling the transverse momenta of the jets. The scaling is applied before passing the jet four-momenta to HAWK. The tested method consists of the following steps:

- Multiply the leading and sub-leading jet momenta by $\frac{p_T(H)}{p_T(j_0+j_1)}$.
- Adjust the energy of the jets such that their respective invariant masses are not affected by the scaling.
- Calculate new Bjorken x values using the scaled jets.

Here, $p_T(H)$ is the transverse momentum of the Higgs boson and $p_T(j_0 + j_1)$ is the transverse momentum of the sum of the four-vectors of the two leading jets.
After applying these steps, the two leading jets balance the Higgs boson in p_T, but their η and ϕ values are unchanged. Hence, if the measuring precision for the Higgs boson p_T is higher than for the jets, this should improve the precision of the jet p_T measurement, leading to more precise $\mathcal{O}$ values.
Figure 6.3 shows a comparison of the gauge curves between the default calculation and this method. The curves were produced on reconstruction level with the $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample and the uncertainties were scaled to 30 fb^{-1}. It shows that the $\mathcal{O}_1$ gauge curve becomes steeper after scaling the jets, meaning that different $\tilde{d}$ scenarios are more distinguishable. However, the sensitivity of the measurement does not increase, since the uncertainties of the gauge curve points also become larger. The reason for the larger error bars can be seen in figures 6.4 and 6.5. They show the underlying distributions used to determine the gauge curve points at $\tilde{d} = 0$ and $\tilde{d} = -0.2$. The distributions obtained with rescaled jets have larger standard deviations, leading the means of the pseudodata distributions to fluctuate more, resulting in larger error bars in the gauge curves. The confidence intervals around $\tilde{d} = 0$ from the $\mathcal{O}_1$ curves are $[-0.030, 0.038]$ for the default calculation and $[-0.33, 0.41]$ for the rescaled version. The $\mathcal{O}_2$ curves yield [-0.181, 0.196] and [-0.204, 0.213] for the default and rescaled calculations. Consequently, applying this method does not improve the accuracy of the calculation, but worsens it instead. However, since this is a $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample and the Higgs boson momentum cannot be reconstructed very precisely in the di-tau final state, an improvement was not expected. Since there are neutrinos in the final state of these events, the Higgs momentum is reconstructed using a so-called “missing mass calculator” [43]. This algorithm analyses the mutual orientation of the tau lepton decay products to calculate the most likely mass for the mother particle

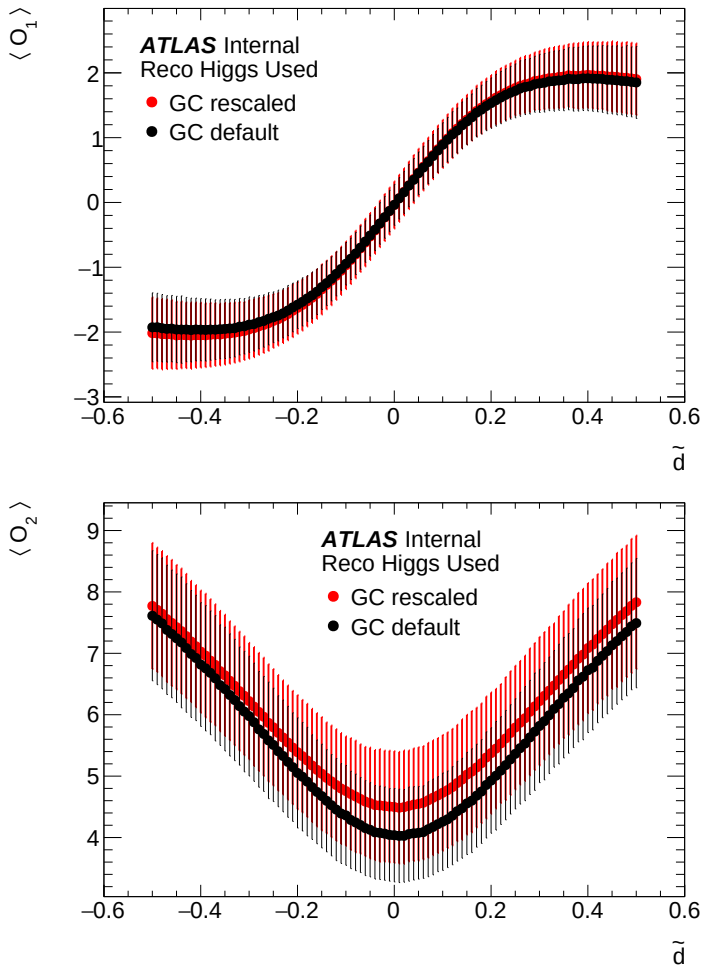

Figure 6.3: Comparison of the reconstruction level gauge curves produced by the calculation with jet scaling and the default $\mathcal{O}$ calculation.

of the tau pair.

To test whether this method could increase the accuracy in other decay channels, the calculation is repeated with the truth level Higgs boson four-vector replacing the reconstruction level Higgs boson. The truth level Higgs momentum can be viewed as the measured momentum of a perfectly reconstructed Higgs boson. Thus, this calculation can test whether the scaling method can help in any scenario. The confidence intervals being larger after rescaling than before in this test would mean that there is no decay channel in which the proposed scaling method improves the accuracy.

Figure 6.6 shows that, again, the $\mathcal{O}_1$ gauge curve using rescaled jets is steeper, but the width of the confidence belt is so much larger that the sensitivity still drops after jet scaling. The $\mathcal{O}_2$ gauge curve after jet rescaling has both larger error bars and is less steep

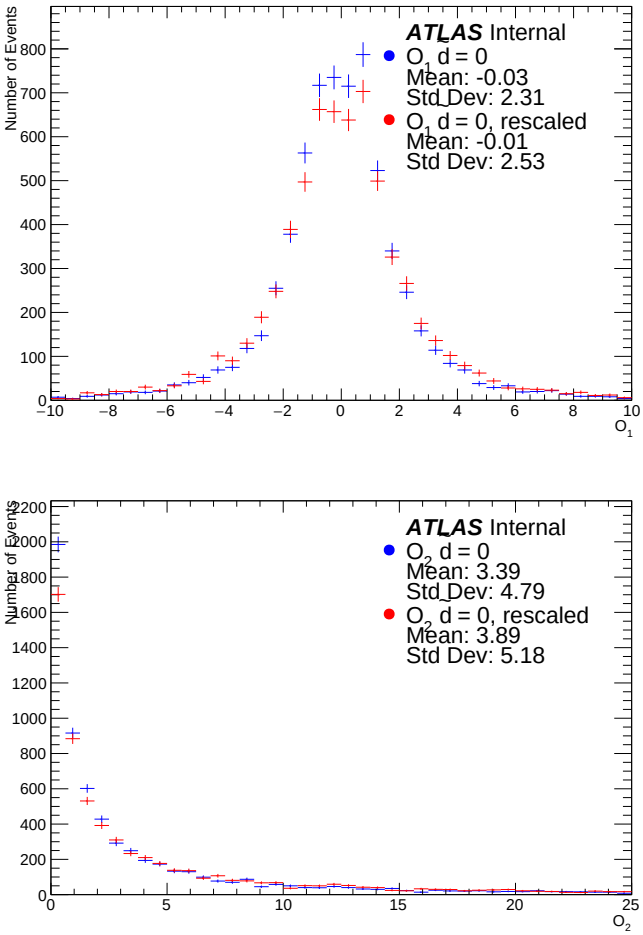

Figure 6.4: Comparison of the SM reconstruction level $\mathcal{O}$ distributions with and without jet scaling.

than before, making its performance worse than the default calculation. The confidence intervals at $\tilde{d} = 0$ from $\mathcal{O}_1$ are $[-0.029, 0.036]$ before rescaling and $[-0.033, 0.039]$ after. $\mathcal{O}_2$ yields $[-0.177, 0.187]$ before and $[-0.191, 0.202]$ after rescaling. The result is that the tested rescaling algorithm should not be used in any analysis.

6.2 Effect of the Rescaling on Event Kinematics

In this section, the origins of the drop in sensitivity after rescaling are investigated. For that, the p_T of the jets before and after rescaling is compared in figure 6.7. It shows that the mean p_T for both jets increases. This would not be the case if the scaling only affected the precision of the measurement. Consequently, scaling the jets as it was done in section

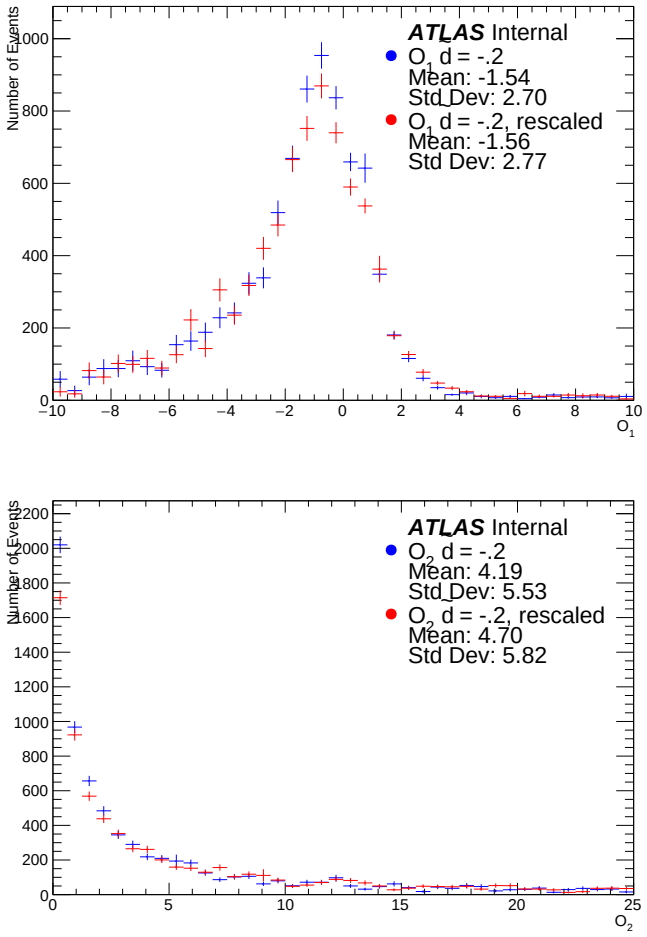

Figure 6.5: Comparison of the reconstruction level $\mathcal{O}$ distributions at $\tilde{d} = -0.2$ with and without jet scaling.

6.1.2 changes the jet kinematics in a way it was not supposed to. The reason for this shift in p_T was thought to be that a possible third jet was neglected, since the Higgs boson should balance the p_T of all VBF jets and the used sample was generated at NLO.
To test this idea, figure 6.8(above) shows the leading jet p_T only for events where p_T of the third truth jet is larger than half the p_T of the second truth jet, while 6.8(below) shows only the other events. One can see that the shift in p_T is significantly higher in events where the p_T of the third jet is large than in events with a low p_T third jet. This proves that a neglected third jet, if existent, indeed influences the jet scaling and has to be accounted for.

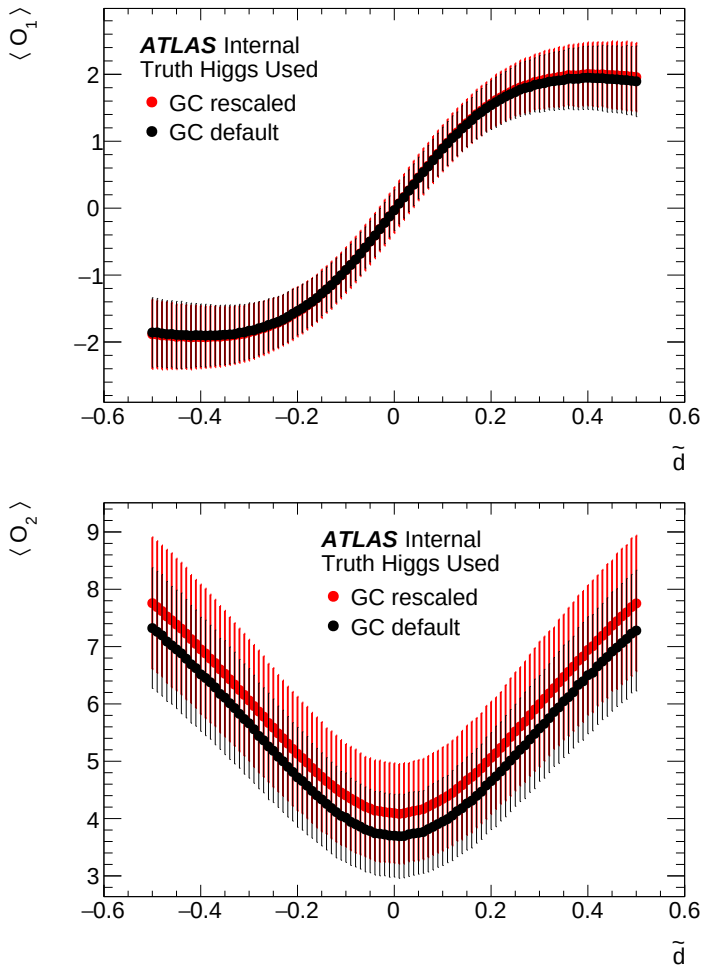

Figure 6.6: Comparison of the reconstruction level gauge curves using the truth level Higgs boson produced by the calculation with jet scaling and the default $\mathcal{O}$ calculation.

6.3 Scaling with Three Jets

In the following test, the sub-leading jet four-vector is replaced by the sum of the sub-leading and the sub-sub-leading jet four-vector in the calculation of the scaling factor $\frac{p_T(H)}{p_T(j_0+j_1)}$ used in section 6.1.2 and the optimal observables. However, in order to avoid being influenced by pile-up jets,which typically have low p_T values, and to have a cut that only uses reconstruction level information, this is only done in events where the sub-sub-leading jet fulfills $p_T > 25$ GeV. This is done in addition to the JVT condition described in section 4.1. In the figures in this section, the curves obtained from calculations that neglect the presence of a third jet are denoted with "2jet", while the ones taking a third jet into account are denoted with "3jet".

Figure 6.9 probes only the effects of adding a third jet with $p_T > 25$ GeV on the gauge

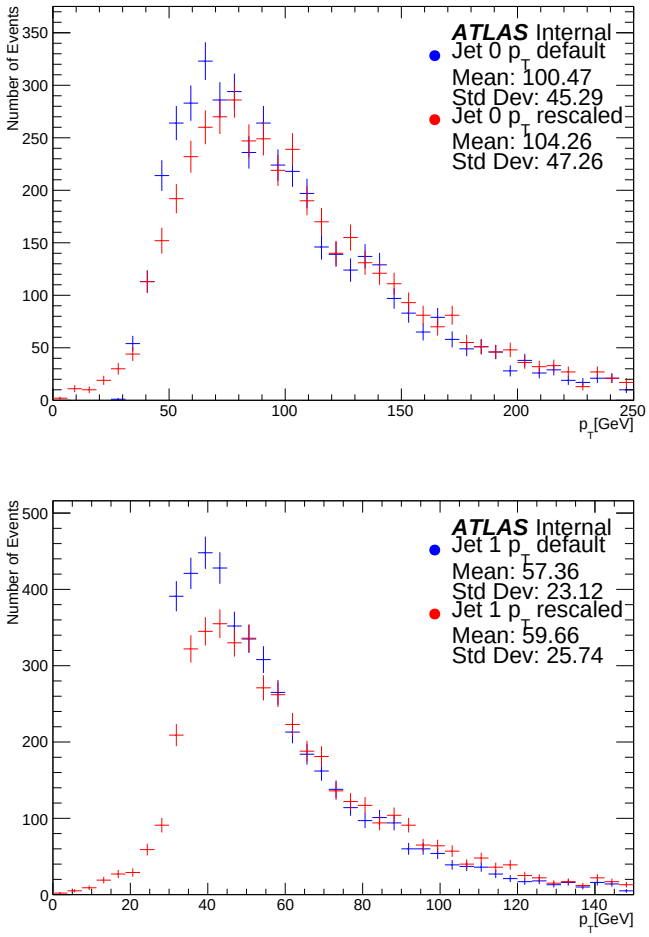

Figure 6.7: Leading and sub-leading jet p_T before and after rescaling with respect to the truth level Higgs boson. The mean values increase.

curves, with no rescaling applied. While the size of the error bars does not change significantly, the curves become flatter. This leads to confidence level intervals around $\tilde{d} = 0$ of [-0.029, 0.036] without and [-0.031. 0.039] with the third jet for $\mathcal{O}_1$. For $\mathcal{O}_2$, the intervals change from [-0.177, 0.187] to [-0.208, 0.192] when including the third jet in the calculation in the described way.

This test shows that the gauge curves become flatter after including a possible third jet in the scaling calculation. When neglecting the third jet, the constant scaling factor applied to the two leading jets leads to a bigger difference in p_T than before, which causes the event to be less SM-like and produce larger $\mathcal{O}$ values. Therefore, correctly including a possible third jet in an event is expected to produce flatter gauge curves. However, the method tested in this section seems to force events to appear SM-like even for $\tilde{d} \neq 0$,

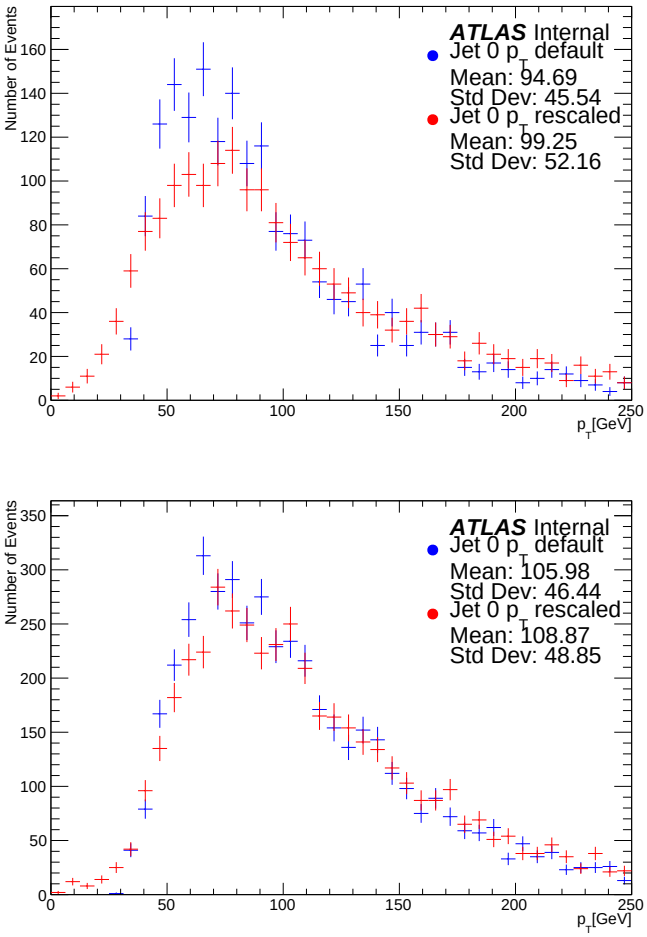

Figure 6.8: Leading jet p_T before and after rescaling with respect to the truth level Higgs boson. The upper distribution only shows events where the p_T value of the truth level sub-sub-leading jet is higher than half the p_T of the truth sub-leading jet, while the lower distribution only shows the other events.

overcompensating for the asymmetry introduced by completely neglecting the possibility of a third jet as in section 6.1.2.

In figure 6.10, the calculation with scaling of up to three jets with respect to the truth level Higgs boson momentum is compared to the default calculation. It shows that the sensitivity does not improve. The error bars are now about as large as for the default calculation, but the gauge curves obtained with the new calculation are flatter than the curves from the default calculation, as expected from the previous test. Therefore, the confidence intervals become larger. For $\mathcal{O}_1$, the confidence intervals on $\tilde{d}$ around $\tilde{d} = 0$ change from $[-0.029, 0.36]$ for the default calculation to $[-0.033, 0.041]$, and $\mathcal{O}_2$ yields $[-0.177, 0.187]$ before and $[-0.191, 0.202]$ after rescaling. Therefore, this method also

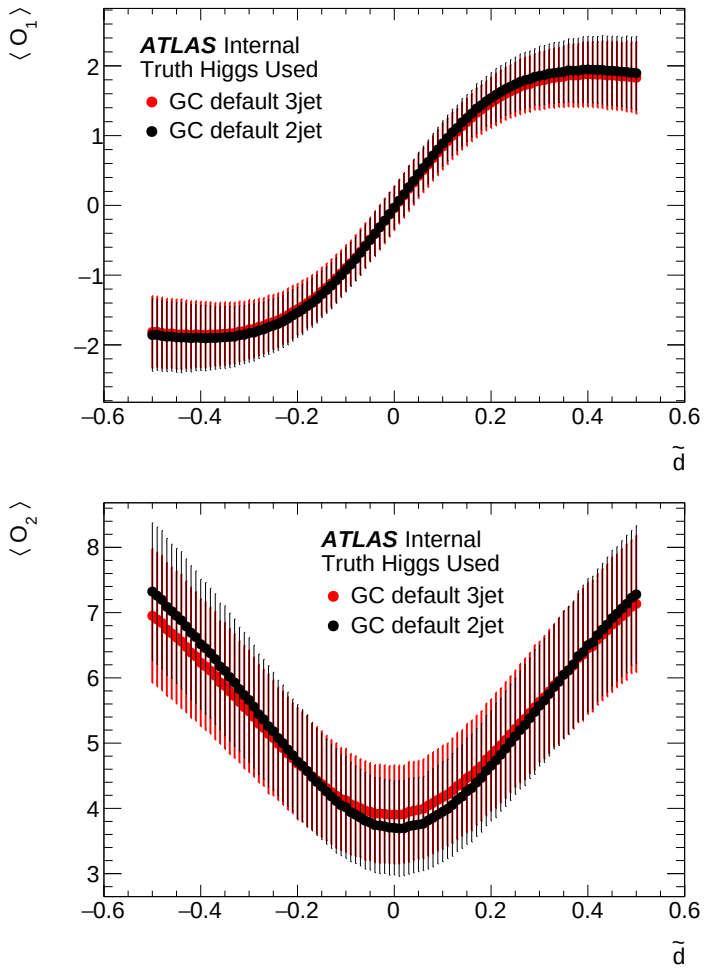

Figure 6.9: Comparison of the default calculation on reconstruction level with the calculation that takes the third jet into account, but does not scale jets.

does not improve the sensitivity of the $\mathcal{O}$ calculation.

To conclude, there seems to be not much sensitivity to be gained from the application of jet scaling methods of the kind presented in this chapter, if at all. The studies show that no improvement of the $\mathcal{O}$ calculation accuracy can be expected without taking NLO effects into consideration. The nature of the corrections that need to be applied for the presented algorithms to yield any improvement, however, is still unclear.

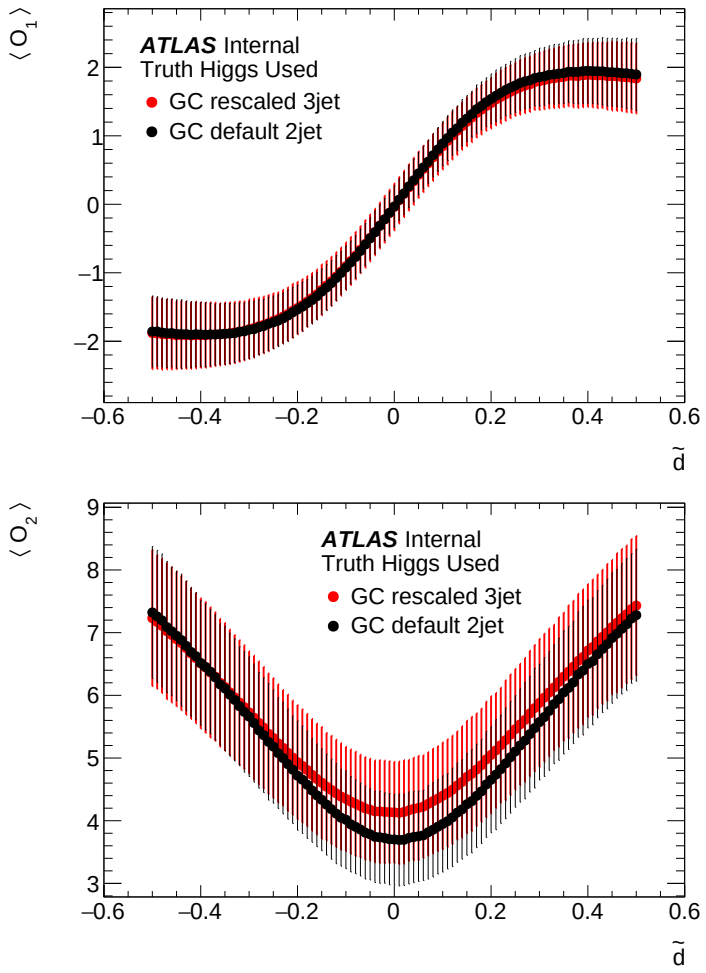

Figure 6.10: Comparison of gauge curves for the scaling using three jets and the default calculation. The jets were rescaled with respect to the truth level Higgs momentum.

7 Conclusion and Outlook

Chapter 2 gave an overview over the theoretical fundamentals of anomalous Higgs couplings and the optimal observable method. Chapter 3 described the LHC, the ATLAS detector and analysis techniques used within the ATLAS collaboration.
In chapter 4, the reweighting procedure that was used for the LHC run 1 analysis was tested. Before the test was started, the VBF $H \to \tau_{\text{lep}}\tau_{\text{had}}$ sample most of the analysis was conducted on was analysed. The differences between truth and reconstruction level distributions were determined and the causes were found. Then the test of the reweighting algorithm at leading order was conducted, and it showed that the run 1 HAWK reweighting algorithm is capable of reweighting 13 TeV MC samples.
This was, however, a test at leading order, and the standard for Monte Carlo data used within the ATLAS collaboration for run 2 is at least NLO. Therefore, it is important to also test the reweighting at NLO, which was done in chapter 5. But since the reweighting algorithm was devised for application on LO data, the event weight and $\mathcal{O}$ calculation had to be modified. After testing different alternative modifications, the best performing one was to add the momenta of two partons so that a situation comparable to that in LO events is reached. With this modification, testing the validity of the reweighting yielded a slightly better result than at LO. This result shows that the modification method was chosen correctly, and it is a stronger argument in favour of the usage of the tested algorithm in run 2 than the test at LO described in chapter 4.
The studies in chapter 6 were conducted to see whether the accuracy of the $\mathcal{O}$ calculation with HAWK could be improved. The code was suspected to perform suboptimally in Higgs decay channels where the Higgs momentum can be measured with higher accuracy than the jet momentum, since the Higgs momentum is replaced in order to balance the momentum of the two leading jets at an early stage of the calculation. Therefore, the leading jets were scaled so that they balance the momentum of the Higgs boson before the HAWK calculation was started. Tests of the performance of this modified calculation show, however, that this approach does not lead to an improved accuracy of the $\mathcal{O}$ calculation, regardless of the decay channel of the Higgs boson, since not even a scaling with respect to the truth level Higgs boson momentum improved the accuracy. A more

sophisticated approach seems to be needed in order to gain sensitivity to $\tilde{d}$, but this approach is not yet found. Moreover, from the test results collected in this study, it seems unlikely that another scaling algorithm can yield an improvement that is significant when compared to other errors on the calculation. For now, it might be better to test other methods in order to improve the $\mathcal{O}$ calculation efficiency, if other methods are available. Regarding the tested reweighting algorithm, it is worth noting that there are more complex BSM Higgs coupling scenarios than the ones tested in this thesis, for example with both a_2 and a_3 in equation 2.2 being non-zero, meaning that there are both CP-even and CP-odd non-SM contributions to the HVV couplings at the same time. In these scenarios, a single parameter like $\tilde{d}$ used in this analysis is not sufficient to specify one of the considered HVV coupling scenarios. In these cases, the reweighting algorithm would have to be validated again, and as the number of parameters needed to describe the models increases, reweighting techniques are known to become very computationally expensive. It will be interesting to see whether new techniques such as morphing might be a better choice in more complex studies in the future, since their computational cost does not increase as quickly with the number of model parameters as for reweighting methods. The run 2 $H \rightarrow \tau\tau$ CP analysis has mainly been focusing on probing the decay vertex of the Higgs boson, and therefore there has not yet been much progress in the optimal observable analysis of the production vertex. However, this analysis is planned to be started in early 2017, so the optimal observable method will be applied for 13 TeV very soon. Since in run 1 only around 26 fb^{-1} of data were recorded at ATLAS and in 2016 alone, 35.6 fb^{-1} were recorded , the constraints on the CP mixing in HVV couplings after run 2 can be expected to be significantly more rigorous than the current ones.

A Simulated Events

```
######################################################################

##  Process generated with '-no_b_mass' model restrictions    ##
##  5F scheme (MB=0)                                          ##
##  no gluon fusion                                           ##
##                                                            ##
##  NB: Please check that before generating this process,     ##
##  you have correctly set the content of the multiparticles  ##
##  'p' (proton) and 'j' (jet) inside MG5_aMC,                ##
##  accordingly to the massless flavour scheme                ##
##  5F: p, j = g d d~ u u~ s s~ c c~ b b~                     ##
######################################################################


####################################
## INFORMATION FOR FRBLOCK
####################################
Block frblock
    1 1.000000e+03 # Lambda
    2 1.000000e+00 # cosa
    3 1.000000e+00 # kSM
    4 1.000000e+00 # kHtt
    5 1.000000e+00 # kAtt
    8 1.000000e+00 # kHll
    9 1.000000e+00 # kAll
   10 1.000000e+00 # kHaa
   11 1.000000e+00 # kAaa
   12 1.000000e+00 # kHza
   13 1.000000e+00 # kAza
```

```
   16 0.000000e+00 # kHzz
   17 0.000000e+00 # kAzz
   18 0.000000e+00 # kHww
   19 0.000000e+00 # kAww
   20 0.000000e+00 # kHda
   21 0.000000e+00 # kHdz
   22 0.000000e+00 # kHdwR (real part of kHdw)
   23 0.000000e+00 # kHdwI (imaginary part of kHdw)

###################################
### INFORMATION FOR LOOP
###################################
Block loop
  1 9.118800e+01 # MU_R

###################################
## INFORMATION FOR SMINPUTS
###################################
Block SMINPUTS
    1 1.325070e+02 # aEWM1
    2 1.166390e-05 # Gf
    3 1.180000e-01 # aS

###################################
## INFORMATION FOR MASS
###################################
Block MASS
    6 1.730000e+02 # MT
   15 1.777000e+00 # MTA
   23 9.118800e+01 # MZ
   25 1.250000e+02 # MX0
## Dependent parameters, given by model restrictions.
## Those values should be edited following the
## analytical expression. MG5 ignores those values
## but they are important for interfacing the output of MG5
## to external program such as Pythia.
```

```
  1 0.000000 # d : 0.0
  2 0.000000 # u : 0.0
  3 0.000000 # s : 0.0
  4 0.000000 # c : 0.0
  5 0.000000 # b : 0.0
  11 0.000000 # e- : 0.0
  12 0.000000 # ve : 0.0
  13 0.000000 # mu- : 0.0
  14 0.000000 # vm : 0.0
  16 0.000000 # vt : 0.0
  21 0.000000 # g : 0.0
  22 0.000000 # a : 0.0
  24 80.419002 # w+ : cmath.sqrt(MZ__exp__2/2. + cmath.sqrt(
     MZ__exp__4/4. - (aEW*cmath.pi*MZ__exp__2)/(Gf*sqrt__2)))
  82 0.000000 # gh : 0.0

###################################
## INFORMATION FOR YUKAWA
###################################
Block YUKAWA
    6 1.730000e+02 # ymt
   15 1.777000e+00 # ymtau

###################################
## INFORMATION FOR DECAY
###################################
DECAY  6 1.491500e+00  # top width
DECAY 23 2.441404e+00 # Z width
DECAY 24 2.047600e+00  # W width
DECAY 25 4.070000e-03 # X0 width
## Dependent parameters, given by model restrictions.
## Those values should be edited following the
## analytical expression. MG5 ignores those values
## but they are important for interfacing the output of MG5
## to external program such as Pythia.
DECAY 1 0.000000 # d : 0.0
```

```
DECAY 2 0.000000 # u : 0.0
DECAY 3 0.000000 # s : 0.0
DECAY 4 0.000000 # c : 0.0
DECAY 5 0.000000 # b : 0.0
DECAY 11 0.000000 # e- : 0.0
DECAY 12 0.000000 # ve : 0.0
DECAY 13 0.000000 # mu- : 0.0
DECAY 14 0.000000 # vm : 0.0
DECAY 15 0.000000 # ta- : 0.0
DECAY 16 0.000000 # vt : 0.0
DECAY 21 0.000000 # g : 0.0
DECAY 22 0.000000 # a : 0.0
DECAY 82 0.000000 # gh : 0.0
#===========================================================
# QUANTUM NUMBERS OF NEW STATE(S) (NON SM PDG CODE)
#===========================================================

Block QNUMBERS 82 # gh
       1 0  # 3 times electric charge
       2 1  # number of spin states (2S+1)
       3 8  # colour rep (1: singlet, 3: triplet, 8: octet)
       4 1  # Particle/Antiparticle distinction (0=own anti)
```

File A.1: Default MG parameter card for the HC model. The Higgs boson couplings are listed in the "FRBLOCK" starting in line 16. These parameters are the same as the κ parameters in equation 2.7. Setting them as shown here reproduces the SM prediction for the Higgs couplings.

```
#*********************************************************************

#                      MadGraph5_aMC@NLO                            *
#                                                                   *
#                    run_card.dat MadEvent                          *
#                                                                   *
#  This file is used to set the parameters of the run.              *
#                                                                   *
#  Some notation/conventions:                                       *
#                                                                   *
```

```
#  Lines starting with a '# ' are info or comments           *
#                                                            *
#  mind the format: value   = variable    ! comment          *
#************************************************************

#
#*******************
# Running parameters
#*******************
#
#************************************************************

# Tag name for the run (one word)                            *
#************************************************************

  tag_1    = run_tag ! name of the run
#************************************************************

# Run to generate the grid pack                              *
#************************************************************

  False    = gridpack !True = setting up the grid pack
#************************************************************

# Number of events and rnd seed                              *
# Warning: Do not generate more than 1M events in a single run *
# If you want to run Pythia, avoid more than 50k events in a run.
    *
#************************************************************

  50000 = nevents ! Number of unweighted events requested
  0   = iseed  ! rnd seed (0=assigned automatically=default))
#************************************************************

# Collider type and energy                                   *
# lpp: 0=No PDF, 1=proton, -1=antiproton, 2=photon from proton, *
```

```
#                                       3=photon from electron  *
#*********************************************************************

     1       = lpp1   ! beam 1 type
     1       = lpp2   ! beam 2 type
     6500.0    = ebeam1 ! beam 1 total energy in GeV
     6500.0    = ebeam2 ! beam 2 total energy in GeV
#*********************************************************************

# Beam polarization from -100 (left-handed) to 100 (right-handed)
    *
#*********************************************************************

     0.0    = polbeam1 ! beam polarization for beam 1
     0.0    = polbeam2 ! beam polarization for beam 2
#*********************************************************************

# PDF CHOICE: this automatically fixes also alpha_s and its evol.
    *
#*********************************************************************

     lhapdf   = pdlabel    ! PDF set
     10800    = lhaid    ! if pdlabel=lhapdf, this is the lhapdf
        number
# lhaid 21000 = MSTW2008lo68cl
#*********************************************************************

# Renormalization and factorization scales                      *
#*********************************************************************

 True = fixed_ren_scale ! if .true. use fixed ren scale
 True       = fixed_fac_scale ! if .true. use fixed fac scale
 80.385 = scale          ! fixed ren scale
 80.385 = dsqrt_q2fact1  ! fixed fact scale for pdf1
 80.385 = dsqrt_q2fact2  ! fixed fact scale for pdf2
```

```
 -1 = dynamical_scale_choice ! Choose one of the preselected
    dynamical choices
 1.0  = scalefact      ! scale factor for event-by-event scales
#*********************************************************************

# Time of flight information. (-1 means not run)
#*********************************************************************

 -1.0 = time_of_flight ! threshold below which info is not
    written
#*********************************************************************

# Matching - Warning! ickkw > 1 is still beta
#*********************************************************************

 0 = ickkw          ! 0 no matching, 1 MLM, 2 CKKW matching
 1 = highestmult    ! for ickkw=2, highest mult group
 1 = ktscheme       ! for ickkw=1, 1 Durham kT, 2 Pythia pTE
 1.0 = alpsfact       ! scale factor for QCD emission vx
 False = chcluster      ! cluster only according to channel diag
 True = pdfwgt         ! for ickkw=1, perform pdf reweighting
 5 = asrwgtflavor   ! highest quark flavor for a_s reweight
 True = clusinfo       ! include clustering tag in output
 3.0 = lhe_version     ! Change the way clustering information
    pass to shower.
#*********************************************************************

#*******************************************************
#
#*******************************************************
# Automatic ptj and mjj cuts if xqcut > 0
# (turn off for VBF and single top processes)
#*******************************************************
   True  = auto_ptj_mjj ! Automatic setting of ptj and mjj
#*******************************************************
#
```

```
#*********************************
# BW cutoff (M+/-bwcutoff*Gamma)
#*********************************
  15.0 = bwcutoff     ! (M+/-bwcutoff*Gamma)
#***********************************************************
# Apply pt/E/eta/dr/mij/kt_durham cuts on decay products or not
# (note that etmiss/ptll/ptheavy/ht/sorted cuts always apply)
#***********************************************************
   False = cut_decays   ! Cut decay products
#***********************************************************
# Number of helicities to sum per event (0 = all helicities)
# 0 gives more stable result, but longer run time (needed for
# long decay chains e.g.).
# Use >=2 if most helicities contribute, e.g. pure QCD.
#***********************************************************
   0  = nhel         ! Number of helicities used per event
#*******************
# Standard Cuts
#*******************
#
#*****************************************************************************

# Minimum and maximum pt's (for max, -1 means no cut)         *
#*****************************************************************************

 20.0  = ptj      ! minimum pt for the jets
 0.0  = ptb      ! minimum pt for the b
 10.0  = pta      ! minimum pt for the photons
 10.0  = ptl      ! minimum pt for the charged leptons
 0.0  = misset   ! minimum missing Et (sum of neutrino's momenta)
 0.0  = ptheavy  ! minimum pt for one heavy final state
 -1.0  = ptjmax   ! maximum pt for the jets
 -1.0  = ptbmax   ! maximum pt for the b
 -1.0  = ptamax   ! maximum pt for the photons
 -1.0  = ptlmax   ! maximum pt for the charged leptons
```

```
 -1.0  = missetmax ! maximum missing Et (sum of neutrino's momenta
    )
#*********************************************************************

# Minimum and maximum E's (in the center of mass frame)       *
#*********************************************************************

  0.0  = ej    ! minimum E for the jets
  0.0  = eb    ! minimum E for the b
  0.0  = ea    ! minimum E for the photons
  0.0  = el    ! minimum E for the charged leptons
  -1.0   = ejmax ! maximum E for the jets
 -1.0   = ebmax ! maximum E for the b
 -1.0   = eamax ! maximum E for the photons
 -1.0   = elmax ! maximum E for the charged leptons
#*********************************************************************

# Maximum and minimum absolute rapidity (for max, -1 means no cut
   ) *
#*********************************************************************

  5.0 = etaj   ! max rap for the jets
  -1.0  = etab   ! max rap for the b
 2.5  = etaa   ! max rap for the photons
 2.5  = etal   ! max rap for the charged leptons
 0.0  = etajmin ! min rap for the jets
 0.0  = etabmin ! min rap for the b
 0.0  = etaamin ! min rap for the photons
 0.0  = etalmin ! main rap for the charged leptons
#*********************************************************************

# Minimum and maximum DeltaR distance                         *
#*********************************************************************

 0.4 = drjj   ! min distance between jets
 0.0   = drbb   ! min distance between b's
```

```
 0.4 = drll   ! min distance between leptons
 0.4 = draa   ! min distance between gammas
 0.0  = drbj   ! min distance between b and jet
 0.4 = draj   ! min distance between gamma and jet
 0.4 = drjl   ! min distance between jet and lepton
 0.0  = drab   ! min distance between gamma and b
 0.0  = drbl   ! min distance between b and lepton
 0.4 = dral   ! min distance between gamma and lepton
 -1.0 = drjjmax ! max distance between jets
 -1.0 = drbbmax ! max distance between b's
 -1.0 = drllmax ! max distance between leptons
 -1.0 = draamax ! max distance between gammas
 -1.0 = drbjmax ! max distance between b and jet
 -1.0 = drajmax ! max distance between gamma and jet
 -1.0 = drjlmax ! max distance between jet and lepton
 -1.0 = drabmax ! max distance between gamma and b
 -1.0 = drblmax ! max distance between b and lepton
 -1.0 = dralmax ! maxdistance between gamma and lepton
#*********************************************************************

# Minimum and maximum invariant mass for pairs                *
# WARNING: for four lepton final state mmll cut require to have *
#         different lepton masses for each flavor!            *
#*********************************************************************

 0.0  = mmjj   ! min invariant mass of a jet pair
 0.0  = mmbb   ! min invariant mass of a b pair
 0.0  = mmaa   ! min invariant mass of gamma gamma pair
 0.0  = mmll   ! min invariant mass of l+l- (same flavour) lepton
     pair
 -1.0 = mmjjmax ! max invariant mass of a jet pair
 -1.0 = mmbbmax ! max invariant mass of a b pair
 -1.0 = mmaamax ! max invariant mass of gamma gamma pair
 -1.0 = mmllmax ! max invariant mass of l+l- (same flavour)
    lepton pair
```

```
#*********************************************************************

# Minimum and maximum invariant mass for all letpons          *
#*********************************************************************

 0.0  = mmnl   ! min invariant mass for all letpons (l+- and vl)
 -1.0 = mmnlmax ! max invariant mass for all letpons (l+- and vl)
#*********************************************************************

# Minimum and maximum pt for 4-momenta sum of leptons         *
#*********************************************************************

 0.0  = ptllmin ! Minimum pt for 4-momenta sum of leptons(l and
    vl)
 -1.0 = ptllmax ! Maximum pt for 4-momenta sum of leptons(l and
    vl)
#*********************************************************************

# Inclusive cuts                                              *
#*********************************************************************

 0.0 = xptj ! minimum pt for at least one jet
 0.0 = xptb ! minimum pt for at least one b
 0.0 = xpta ! minimum pt for at least one photon
 0.0 = xptl ! minimum pt for at least one charged lepton
#*********************************************************************

# Control the pt's of the jets sorted by pt                   *
#*********************************************************************

 0.0  = ptj1min ! minimum pt for the leading jet in pt
 0.0  = ptj2min ! minimum pt for the second jet in pt
 0.0  = ptj3min ! minimum pt for the third jet in pt
 0.0  = ptj4min ! minimum pt for the fourth jet in pt
 -1.0 = ptj1max ! maximum pt for the leading jet in pt
 -1.0 = ptj2max ! maximum pt for the second jet in pt
```

```
 -1.0  = ptj3max ! maximum pt for the third jet in pt
 -1.0  = ptj4max ! maximum pt for the fourth jet in pt
 0   = cutuse ! reject event if fails any (0) / all (1) jet pt
    cuts
#*********************************************************************

# Control the pt's of leptons sorted by pt                    *
#*********************************************************************

 0.0   = ptl1min ! minimum pt for the leading lepton in pt
 0.0   = ptl2min ! minimum pt for the second lepton in pt
 0.0   = ptl3min ! minimum pt for the third lepton in pt
 0.0   = ptl4min ! minimum pt for the fourth lepton in pt
 -1.0  = ptl1max ! maximum pt for the leading lepton in pt
 -1.0  = ptl2max ! maximum pt for the second lepton in pt
 -1.0  = ptl3max ! maximum pt for the third lepton in pt
 -1.0  = ptl4max ! maximum pt for the fourth lepton in pt
#*********************************************************************

# Control the Ht(k)=Sum of k leading jets                     *
#*********************************************************************

 0.0   = htjmin ! minimum jet HT=Sum(jet pt)
 -1.0  = htjmax ! maximum jet HT=Sum(jet pt)
 0.0   = ihtmin !inclusive Ht for all partons (including b)
 -1.0  = ihtmax !inclusive Ht for all partons (including b)
 0.0   = ht2min ! minimum Ht for the two leading jets
 0.0   = ht3min ! minimum Ht for the three leading jets
 0.0   = ht4min ! minimum Ht for the four leading jets
 -1.0  = ht2max ! maximum Ht for the two leading jets
 -1.0  = ht3max ! maximum Ht for the three leading jets
 -1.0  = ht4max ! maximum Ht for the four leading jets
#*********************************************************************

# Photon-isolation cuts, according to hep-ph/9801442           *
# When ptgmin=0, all the other parameters are ignored          *
```

```
# When ptgmin>0, pta and draj are not going to be used          *
#*********************************************************************

 0.0 = ptgmin ! Min photon transverse momentum
 0.4 = R0gamma ! Radius of isolation code
 1.0 = xn ! n parameter of eq.(3.4) in hep-ph/9801442
 1.0 = epsgamma ! epsilon_gamma parameter of eq.(3.4) in hep-ph
    /9801442
 True = isoEM ! isolate photons from EM energy (photons and
    leptons)
#*********************************************************************

# WBF cuts                                                       *
#*********************************************************************

 0.0   = xetamin ! minimum rapidity for two jets in the WBF case
 0.0   = deltaeta ! minimum rapidity for two jets in the WBF case
#*********************************************************************

# KT DURHAM CUT                                                  *
#*********************************************************************

 -1.0   =  ktdurham
 0.4  =  dparameter
#*********************************************************************

# maximal pdg code for quark to be considered as a light jet  *
# (otherwise b cuts are applied)                              *
#*********************************************************************

 5 = maxjetflavor  ! Maximum jet pdg code
#*********************************************************************

# Jet measure cuts                                               *
#*********************************************************************
```

```
 0.0   = xqcut   ! minimum kt jet measure between partons
#*********************************************************************


#
#*********************************************************************

# Store info for systematics studies                          *
# WARNING: If use_syst is T, matched Pythia output is         *
#         meaningful ONLY if plotted taking matchscale        *
#         reweighting into account!                           *
#*********************************************************************

   True  = use_syst     ! Enable systematics studies
#
#*************************************
# Parameter of the systematics study
#  will be used by SysCalc (if installed)
#*************************************
#
0.5 1 2 = sys_scalefact # factorization/renormalization scale
   factor
None = sys_alpsfact # \alpha_s emission scale factors
30 50 = sys_matchscale # variation of merging scale
# PDF sets and number of members (0 or none for all members).
NNPDF23_lo_as_0130_qed = sys_pdf # matching scales
# MSTW2008nlo68cl.LHgrid 1 = sys_pdf
```

File A.2: MG run card for event generation. This file sets parameters such as the number of events to be simulated or the centre-of-mass energy.

```
#*********************************************************************

##  Process generated with '-no_b_mass' model restrictions    ##
##  5F scheme (MB=0)                                          ##
##  no gluon fusion                                           ##
##                                                            ##
##  NB: Please check that before generating this process,     ##
##  you have correctly set the content of the multiparticles  ##
```

```
##  'p' (proton) and 'j' (jet) inside MG5_aMC,                    ##
##  accordingly to the massless flavour scheme                    ##
##  5F: p, j = g d d~ u u~ s s~ c c~ b b~                         ##
#####################################################################



###################################
## INFORMATION FOR FRBLOCK
###################################
Block frblock
    1 1.000000e+03 # Lambda
    2 0.600000e+00 # cosa
    3 1.666667e+00 # kSM
    4 1.666667e+00 # kHtt
    5 1.000000e+00 # kAtt
    8 1.666667e+00 # kHll
    9 1.000000e+00 # kAll
   10 1.000000e-04 # kHaa
   11 3.030619e+02 # kAaa
   12 1.000000e-04 # kHza
   13 1.000000e-04 # kAza
   16 0.000000e+00 # kHzz
   17 3.955050e+00 # kAzz
   18 0.000000e+00 # kHww
   19 3.955050e+00 # kAww
   20 0.000000e+00 # kHda
   21 0.000000e+00 # kHdz
   22 0.000000e+00 # kHdwR (real part of kHdw)
   23 0.000000e+00 # kHdwI (imaginary part of kHdw)
```

File A.3: Part of the MG parameter card used to produce events with $\tilde{d} = -0.2$. The part of the file following the "FRBLOCK" is identical to the SM card A.1. Therefore, it is not shown here. The coupling strengths of the CP-even Higgs state are multiplied by $\frac{5}{3}$ to compensate for the smaller value of $\cos(\alpha)$. In addition, the couplings of the CP-odd state to vector bosons have been recalculated with equation 2.8.

```
#####################################################################
```

```
##  Process generated with '-no_b_mass' model restrictions    ##
##  5F scheme (MB=0)                                          ##
##  no gluon fusion                                           ##
##                                                            ##
##  NB: Please check that before generating this process,     ##
##  you have correctly set the content of the multiparticles  ##
##  'p' (proton) and 'j' (jet) inside MG5_aMC,                ##
##  accordingly to the massless flavour scheme                ##
##  5F: p, j = g d d~ u u~ s s~ c c~ b b~                     ##
################################################################################



###################################
## INFORMATION FOR FRBLOCK
###################################
Block frblock
    1 1.000000e+03 # Lambda
    2 0.600000e+00 # cosa
    3 1.666667e+00 # kSM
    4 1.666667e+00 # kHtt
    5 1.000000e+00 # kAtt
    8 1.666667e+00 # kHll
    9 1.000000e+00 # kAll
   10 1.000000e-04 # kHaa
   11 -9.091857e+02 # kAaa
   12 1.000000e-04 # kHza
   13 1.000000e-04 # kAza
   16 0.000000e+00 # kHzz
   17 -1.186515e+01 # kAzz
   18 0.000000e+00 # kHww
   19 -1.186515e+01 # kAww
   20 0.000000e+00 # kHda
   21 0.000000e+00 # kHdz
   22 0.000000e+00 # kHdwR (real part of kHdw)
   23 0.000000e+00 # kHdwI (imaginary part of kHdw)
```

File A.4: Part of the MG parameter card used to produce events with $\tilde{d} = 0.6$.

```
######################################################################

##  Process generated with '-no_b_mass' model restrictions    ##
##  5F scheme (MB=0)                                          ##
##  no gluon fusion                                           ##
##                                                            ##
##  NB: Please check that before generating this process,     ##
##  you have correctly set the content of the multiparticles  ##
##  'p' (proton) and 'j' (jet) inside MG5_aMC,                ##
##  accordingly to the massless flavour scheme                ##
##  5F: p, j = g d d~ u u~ s s~ c c~ b b~                     ##
######################################################################


####################################
## INFORMATION FOR FRBLOCK
####################################
Block frblock
    1 1.000000e+03 # Lambda
    2 0.600000e+00 # cosa
    3 1.666667e+00 # kSM
    4 1.666667e+00 # kHtt
    5 1.000000e+00 # kAtt
    8 1.666667e+00 # kHll
    9 1.000000e+00 # kAll
   10 1.000000e-04 # kHaa
   11 -1.51531e+03 # kAaa
   12 1.000000e-04 # kHza
   13 1.000000e-04 # kAza
   16 0.000000e+00 # kHzz
   17 -1.977525e+01 # kAzz
   18 0.000000e+00 # kHww
   19 -1.977525e+01 # kAww
   20 0.000000e+00 # kHda
   21 0.000000e+00 # kHdz
   22 0.000000e+00 # kHdwR (real part of kHdw)
```

```
35   23 0.000000e+00 # kHdwI (imaginary part of kHdw)
```

File A.5: Part of the MG parameter card used to produce events with $\tilde{d} = 1.0$.

Bibliography

[1] 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. B716 1-29 (2012)

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 30 (2012)

[3] F. Englert, R. Brout, *Broken Symmetries and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. 13 (9): 321-323 (1964)

[4] P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. 13 (16): 508-509 (1964)

[5] G. Guralnik, C. Hagen, T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. 13 (20): 585-587 (1964)

[6] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. C75 476 (2015)

[7] CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, Phys. Rev. D 92, 012004 (2015)

[8] A. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. 5, 32 (1967)

[9] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22**, 579 (1961)

[10] A. Salam, *Elementary Particle Physics: Relativistic Groups and Analyticity*, in N. Svartholm, editor, *8-th Nobel Symposium*, page 367, Almquvist and Wiksell, Stockholm (1968)

[11] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967)

[12] UA1 Collaboration, *Experimental observation of isolated large transverse energy electrons with associated missing energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B122 103-116 (1983)

[13] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys.Lett. B122 476-485 (1983)

[14] S. Weinberg, *Physical Processes in a Convergent Theory of the Weak and Electromagnetic Interactions*, Phys. Rev. Lett. 27 (24): 1688-1691 (1971)

[15] K. Olive, et al., *Review of Particle Physics*, Chin. Phys. C, 38, 090001 (2014)

[16] L. Landau, *The moment of a 2-photon system*, Dokl. Akad. Nauk. 60: 207-209 (1948)

[17] C. Yang, *Selection Rules for the Dematerialization of a Particle into Two Photons*, Phys. Rev. 77, 242 (1950)

[18] ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, Eur. Phys. J. C76 658 (2016)

[19] P. Artoisenet, et al., *A framework for Higgs Characterisation*, arXiv:1306.6464 (2014)

[20] F. Malton, K. Mawatari, M. Zaro, *Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects*, arXiv:1311.1829 [hep-ph] (2014)

[21] C. Degrande, et al., *UFO - The Universal FeynRules Output*, Comput.Phys.Commun. 183 1201-1214 (2012)

[22] J. Alwall, et al., *MadGraph 5 : Going Beyond*, arXiv:1106.0522 (2011)

[23] D. Atwood, A. Soni, *Analysis for magnetic and electric dipole moment form factors of the top quark via* $e^+e^- \to t\bar{t}$, Phys.Rev. D45 2405 (1992)

[24] P. Bryant, L. Evans, *LHC Machine*, JINST 3 S08001 (2008)

[25] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 S08003 (2008)

[26] ATLAS Collaboration, *ATLAS Inner Detector*, ATLAS TDR 4, CERN/LHCC 97-16 (1997)

[27] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. C70: 787-821 (2010)

[28] ATLAS LARG Unit, *ATLAS Liquid Argon Calorimeter*, Technical report, ATLAS Collaboration (1996), CERN/LHCC 96-41

[29] A. Ruiz MartÃnez, A. Collaboration, *The Run-2 ATLAS Trigger System* (2016), ATL-DAQ-PROC-2016-003

[30] T. Sjöstrand, S. Mrenna, P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. 178: 852-867 (2007)

[31] S. Agostinelli, et al. (Geant4), *GEANT4: A Simulation Toolkit*, Nucl. Instrum. Methods A **506**, 250 (2003)

[32] M. Cacciari, G. Salam, *The anti-k_t jet clustering algorithm*, JHEP 0804:063 (2008)

[33] A. Denner, et al., *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, arXiv:1412.5390 (2014)

[34] P. Nason, C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP 1002 037 (2010)

[35] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP 09 145 (2014)

[36] J. Pumplin, et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP 07, 012 (2002)

[37] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C76 292 (2016)

[38] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. C74:2941 (2014)

[39] R. Brun, F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A389, 81 (1997)

[40] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C76:581 (2016)

[41] ATLAS Collaboration, *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001 (2009)

[42] J. Neyman, *Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability*, Royal Society (1937)

[43] A. Elagin, et al., *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Phys. A 654 (1) 481-489 (2010)

Erklärung nach §17(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 22. Mai 2017

(Serhat Ördek)