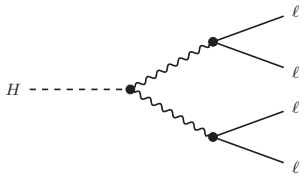
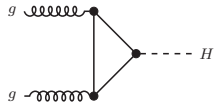
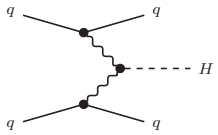
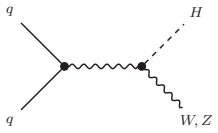
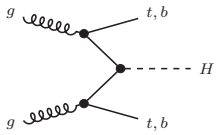
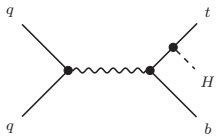


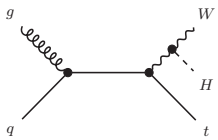


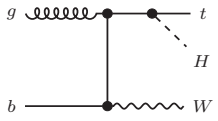












Technische Universität München
Fakultät für Physik

Dissertation

Measurement of the Higgs Boson Coupling Structure in Decays to Four Leptons with the ATLAS Detector

von
Verena Maria Walbrecht
Januar 2020

Fakultät für Physik

TECHNISCHE UNIVERSITÄT MÜNCHEN

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Measurement of the Higgs Boson Coupling Structure in Decays to Four Leptons with the ATLAS Detector

Verena Maria Walbrecht

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	apl. Prof. Dr. Norbert Kaiser
Prüfer der Dissertation:	
1.	apl. Prof. Dr. Hubert Kroha
2.	Prof. Dr. Stefan Schönert

Die Dissertation wurde am 10.01.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 31.01.2020 angenommen.

Für meine Mutter,
die mich mein Leben lang unterstützt hat, mir immer geholfen hat meine Träume zu verwirklichen und mich zum dem gemacht hat der ich heute bin.

Abstract

The properties of the Higgs boson couplings have been studied in decays into four leptons, $H \to ZZ^* \to 4\ell$ ($\ell = e, \mu$), using proton-proton collision data recorded with the ATLAS detector at the Large Hadron Collider (LHC) in the years 2015 to 2018 at a centre-of-mass energy of 13 TeV (LHC Run 2). Measurements have been performed using different data sets corresponding to integrated luminosities of 36.1 fb^{-1}, 79.8 fb^{-1} and 139 fb^{-1}.
The cross sections for the four main Higgs boson production modes have been measured inclusively in the 4ℓ final states as well as in exclusive phase space regions probing the Higgs boson couplings. The inclusive Higgs boson production cross section times branching ratio into 4ℓ decays in the rapidity range $|\eta| < 2.5$ of 1.34 ± 0.12 pb measured with the full set of Run 2 data set is in very good agreement with the Standard Model (SM) prediction of 1.33 ± 0.08 pb. All measurements in exclusive phase space regions are compatible with the SM predictions as well. The sensitivity of the measurements was projected to the ultimate integrated luminosity of the high-luminosity LHC (HL-LHC) of 3000 fb^{-1}. With the HL-LHC data, the precision of the cross section measurements is expected to improve by a factor four for gluon fusion production and by a factor of six for weak gauge boson fusion production and for the associated production with a weak gauge boson. The results provided input for the long-term planning of particle physics infrastructure in the 2020 European Strategy process.
The tensor structure of the Higgs boson couplings to massive vector bosons and to gluons has been studied in two effective field theory approaches to probe for small beyond-SM (BSM) CP-even and CP-odd admixtures to the CP-even SM coupling. In the first analysis performed with 36.1 fb^{-1} of Run 2 data, the BSM couplings are defined within the so-called Higgs Characterisation framework. From the measurements of the rates of the four main Higgs boson production modes, constraints on the CP-odd BSM coupling to gluons (κ_{Agg}) and on the CP-even and CP-odd BSM couplings to heavy vector bosons (κ_{HVV} and κ_{AVV}) have been obtained: $-0.68 < \sin\alpha \cdot \kappa_{Agg} < 0.68$, $-0.6 < \cos\alpha \cdot \kappa_{HVV} < 4.2$ and $-4.4 < \sin\alpha \cdot \kappa_{AVV} < 4.4$ at 95% confidence level (CL), where the mixing angle α describes the relative contributions of the CP-even or CP-odd terms. These are the first limits on BSM Higgs boson couplings in an EFT framework based on the measurement of the Higgs boson production rates providing substantially sensitivity.
The second analysis is based on the production cross section measurements in exclusive phase space regions using the full Run 2 data set (139 fb^{-1}) interpreted in the so-called Standard Model Effective Field Theory (SMEFT). Constraints have been derived on the CP-even and CP-odd BSM coupling parameters to gluons (c_{HG} and $c_{H\widetilde{G}}$), heavy vector bosons (c_{HW}, c_{HB}, c_{HWB} and $c_{H\widetilde{W}}, c_{H\widetilde{B}}, c_{H\widetilde{W}B}$) and top quarks ($c_{uH}$ and $c_{\tilde{u}H}$). The most

stringent limits at 95% CL are $-0.0074 < c_{HG} < 0.0080$ on the CP-even BSM coupling to gluons, as well as $-2.4 < c_{H\widetilde{W}} < 2.4$, $-0.56 < c_{H\widetilde{B}} < 0.56$ and $-1.03 < c_{H\widetilde{W}B} < 1.03$ on the CP-odd BSM coupling parameters to weak gauge bosons. The constraints on the CP-even BSM couplings to weak gauge bosons are weaker. These are the first constraints on SMEFT coupling parameters using the $H \to ZZ^* \to 4\ell$ decay channel, which will be combined with the results in other Higgs boson decay channels.
Finally, the sensitivity of kinematic variables of the final state particles for CP-odd contributions to the Higgs boson coupling in vector boson fusion production has been evaluated in the Higgs Characterisation framework. Limits of $-2.56 < \kappa_{AVV} < 2.64$ at 68% CL are expected for the full Run 2 data set. The measurement will provide the first limits on CP-violation in the $H \to ZZ^* \to 4\ell$ decays.

Contents

1 Introduction

The fundamental structure of the known matter in the universe consists of a few elementary particles, interacting with four fundamental forces. The relation between these particles and three of the forces, i.e. the electromagnetic, the weak and the strong interaction, is described by a quantum field theory, the Standard Model of particle physics. The theory was completed in the early 1970's and has demonstrated ever since enormous success in describing experimental observations. The masses of the fundamental particles are generated by the Higgs mechanism which involves the breaking of the electroweak gauge symmetry. The Higgs mechanism predicts a massive scalar particle, the Higgs boson, which until recently was the last missing particle predicted by the Standard Model.

The discovery in 2012 of a new particle with a mass of 125.09 ± 0.24 GeV and with spin zero, even CP (charge conjugation and parity) quantum numbers and coupling strengths compatible with the Standard Model predictions by the ATLAS and the CMS experiments at the Large Hadron Collider (LHC) completed the particle spectrum of the Standard Model. It also opened a new window for precision tests of the Standard Model and for the search for physics beyond it.

However, despite of the good agreement between Standard Model predictions and measurements with high precision, there are many questions left open, such as the unification of the three forces with gravity, the origin of dark matter, the mass values of the fermions and the asymmetry between matter and antimatter in the universe. Physics beyond the Standard Model can, for example, manifest itself in deviations of Higgs boson production cross sections and decay rates from the Standard Model predictions. Many extensions of the Standard Model predict a modified Higgs sector with additional Higgs singlets or doublets. In such theories, the discovered boson may not be a CP-even eigenstate, but rather contain CP-odd admixtures modifying the tensor coupling structure of the Higgs boson. CP-odd contributions lead to CP-violation in the Higgs boson production and decays which may help to explain the baryon and anti-baryon asymmetry in the universe.

So far, all measurements of Higgs boson production cross sections and decay rates and of quantum numbers agree with the Standard Model predictions. However, the uncertainties in the measurements are still large compared to expectations from physics beyond the

Standard Model (BSM). Therefore, small CP-even or CP-odd admixtures to the SM CP-even state are not excluded. Such BSM contributions can be probed by the measurements of the Higgs boson production and decay rates and of the kinematic properties of the final state particles. While the production and decay rates are sensitive to both CP-even and CP-odd BSM contributions, tests for CP-violation can only be performed using final state kinematic observables which are not sensitive to new CP-even contributions.

In this thesis, the Higgs boson discovery channel into four leptons, $H \to ZZ^* \to 4\ell$ ($\ell = e, \mu$), was analysed. Despite its small branching ratio of $1.250 \cdot 10^{-2}\%$, it is one of the most sensitive decay channels for the search for deviations from the Standard Model due to the very clear signature and high signal-to-background ratio and the precise reconstruction of the final state particles. The data analysed in this thesis were recorded during Run 2 of the LHC in the years from 2015 to 2018 at a centre-of-mass energy of 13 TeV. Three data sets of proton-proton collisions with integrated luminosities of 36.1 fb^{-1}, 79.8 fb^{-1} and 139 fb^{-1} have been analysed.

The Standard Model of particle physics is introduced in the next chapter, with particular emphasis on the Higgs sector. The kinematic properties of the Higgs boson production and decay allowing for BSM physics tests are discussed and frameworks for the description of deviations of the Higgs boson couplings parameters from the Standard Model predictions are introduced. Finally, an overview of the current measurements of Higgs boson properties is given.
The Large Hadron Collider and the ATLAS detector together with the particle reconstruction performed with it are described in Chapter 3, followed by the discussion of the signal and background processes in the $H \to ZZ^* \to 4\ell$ decay channel and the description of the Higgs boson candidate selection in Chapter 4. The inclusive analysis serves as input for the further Higgs boson property measurements in this thesis.

The measurement of the Higgs boson production cross sections in exclusive phase space regions performed with an integrated luminosity of 79.8 fb^{-1}, is presented in Chapter 5. Multivariate methods have been employed to separate signal and background as well as the different production modes.

The ultimate precision of the Higgs boson property measurements at the LHC will be reached after completion of the full LHC program, including the upgrade to the high-luminosity LHC (HL-LHC) which is expected to deliver 3000 fb^{-1} of luminosity, about 20 times more than recorded up to now. Projections of the precision of the current Higgs boson cross section measurements to the full HL-LHC luminosity are described in Chapter 6.

The tensor structure of the Higgs boson couplings to weak gauge bosons and gluons has been studied using two different approaches described in Chapter 7. In the first analysis,

performed with 36.1 fb^{-1} of data, the event yields in event categories targeting the different Higgs boson production modes have been measured and interpreted in terms of CP-even and CP-odd BSM contributions. The second analysis was performed with the full Run 2 data set corresponding to 139 fb^{-1}. The results of the production cross section measurements in exclusive phase space regions were interpreted in terms of BSM coupling parameters within an effective field theory extension of the Standard Model. For this purpose, particle level parametrisations of the cross section, of the 4ℓ decay branching ratio and of the detector acceptance have been developed.

In Chapter 8, CP-violation in vector boson fusion production of the Higgs boson is introduced. To probe for CP-violation, kinematic properties of the final state particles sensitive to the CP quantum numbers of the Higgs boson have been studied, namely two CP-odd observables, the azimuthal angle between the two tagging jets and the so-called first-order Optimal Observable based on an event-by-event matrix element calculation of the process. The measurement was performed for the 36.1 fb^{-1} data set and then extrapolated to the sensitivity expected with the full Run 2 data set.

The results of the thesis are summarised in Chapter 9.

2

The Standard Model and Beyond

In this chapter, the theoretical framework for the performed studies is introduced. First, an introduction to the Standard Model (SM) of particle physics is given with particular emphasis on the Higgs mechanism and Higgs boson phenomenology. Secondly, extensions to the SM in the framework of effective field theories are described. Finally, an overview of the current status of Higgs boson property measurements is given.

The SM of particle physics [1–6] is a relativistic quantum field theory classifying all known elementary particles and describing three of the four fundamental forces, the strong, the electromagnetic and the weak interaction, leaving out only gravity which is described by the classical field theory of General Relativity [7]. Up to the now, all SM predictions are in agreement with the measurements with high precision.

2.1 Particle Content of the Standard Model

The SM is a relativistic local gauge field theory. The SM Lagrangian is invariant under local phase transformations of the gauge symmetry group,

$$\mathcal{G}_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{2.1}$$

where the strong interaction acting on all particles with colour quantum numbers is governed by the $SU(3)_C$ colour group and the unified electroweak interactions by the $SU(2)_L \otimes U(1)_Y$ group. The $SU(2)_L$ interaction only acts on left-handed particle states accounting for the parity violation by the weak interaction.

The SM contains three types of particles, spin-$1/2$ fermions as matter constituents, spin-1 gauge bosons mediating the interactions, and the scalar Higgs boson as a consequence of the Higgs mechanism. (see Figure 2.1).

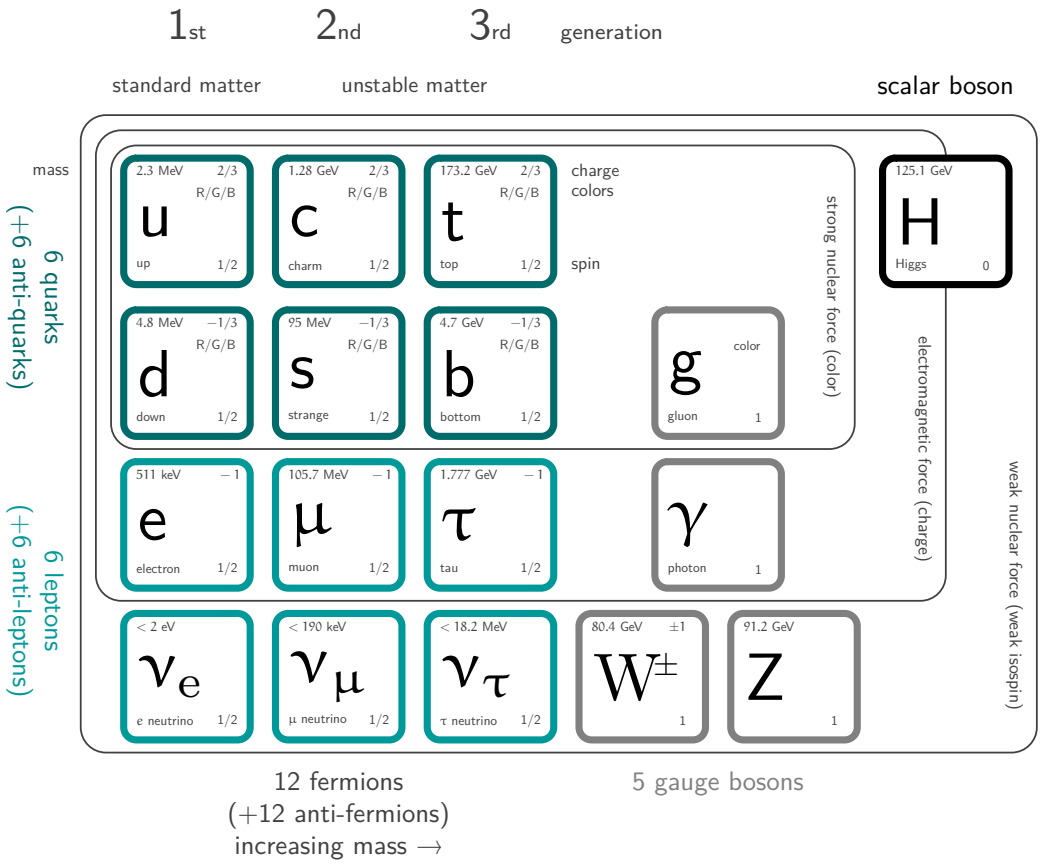


Figure 2.1: Particle content of the Standard Model (SM).

The fermions appear in three generations with increasing masses. Each generation consists of a lepton pair and a pair of up- and down-type quarks. The lepton pair consists of a charged lepton and an associated electrically neutral neutrino. The charged leptons participate in the electromagnetic and the weak interaction, while the neutrinos are interacting only weakly. Up-type quarks have electromagnetic charge $+2/3$, down-type quarks $-1/3$. Since quarks also carry colour quantum numbers r, g, b, they are participating the strong interaction in addition to the weak and electromagnetic interactions. Each particle has a corresponding anti-particle with the same mass and opposite electric charge sign.

The interactions are mediated by the spin-1 gauge bosons, which correspond to the generators of the gauge symmetries. Eight gluons (g) mediate the strong force carrying combinations of colour quantum numbers. The mediator of the electromagnetic interaction is the neutral photon (γ) while the gauge bosons of the weak interaction are the two electrically charged $W^{\pm}$ bosons and the neutral Z boson. While the gluons and the photon are massless, the gauge bosons of the weak interaction are massive.

The only fundamental scalar (spin-0) particle in the SM is the Higgs boson. This colourless, electrically neutral massive particle is predicted by the Higgs mechanism which is responsible for electroweak symmetry breaking [8, 9]. It was discovered last, in 2012 by the ATLAS and the CMS experiments at the Large Hadron Collider with a mass of $m_H = 125.09$ GeV [10, 11]. The electroweak gauge symmetry breaking introduces masses of the fermions and weak gauge bosons and will be discussed in Section 2.3.

2.2 The Electroweak Interaction

The electroweak interaction is described by the Glashow-Salam-Weinberg (GSW) theory, combining the electromagnetic and weak interactions [12–14]. Left-handed chirality eigenstates $\psi(x)_L = 1/2\,(1-\gamma^5)\,\psi(x)$ of the fermions form $SU(2)_L$ doublets with the third component of the weak isospin $I_3 = \pm 1/2$ while the right-handed states $\psi_R(x) = 1/2\,(1+\gamma^5)\,\psi(x)$ form singlets with $I_3 = 0$. The conserved charge of the $U(1)_Y$ symmetry is the weak hypercharge Y_W which is a combination of the third component of the weak isospin I_3 and of the electric charge Q: $Y_W = 2\,(Q - I_3)$. Table 2.1 summarises the electroweak quantum numbers of the fermions.

Table 2.1: Quantum numbers of the electroweak multiplets, the electric charge Q, the third component of the weak isospin I_3 and the weak hypercharge Y_W

Fermion $SU(2)_L$ eigenstates	Quantum Number		
	Q	I_3	Y_W
$\boldsymbol{L}_L = \left[\begin{pmatrix}\nu_{eL}\\ e_L\end{pmatrix}, \begin{pmatrix}\nu_{\mu L}\\ \mu_L\end{pmatrix}, \begin{pmatrix}\nu_{\tau L}\\ \tau_L\end{pmatrix}\right]$	0 -1	$+1/2$ $-1/2$	-1 -1
$\boldsymbol{\ell}_R = [e_R, \mu_R, \tau_R]$	-1	0	-2
$\boldsymbol{Q}_L = \left[\begin{pmatrix}u_L\\ d_L\end{pmatrix}, \begin{pmatrix}c_L\\ s_L\end{pmatrix}, \begin{pmatrix}t_L\\ b_L\end{pmatrix}\right]$	$+2/3$ $-1/3$	$+1/2$ $-1/2$	$+1/3$ $+1/3$
$\boldsymbol{u}_R = [u_R, c_R, t_R]$	$+2/3$	0	$+4/3$
$\boldsymbol{d}_R = [d_R, s_R, b_R]$	$-1/3$	0	$-2/3$

The SM Lagrangian can be divided into four parts:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge kinetic}} + \mathcal{L}_{\text{fermion kinetic}} + \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.2}$$

The first term corresponds to the kinetic energy of the gauge bosons including self-interactions of the non-Abelian gauge bosons. The kinetic terms of the fermions include the gauge invariant interaction terms between fermions and gauge bosons. The Higgs field and its self-interaction are described by from $\mathcal{L}_{\text{scalar}}$ and the Yukawa interactions of the fermions with the Higgs field by $\mathcal{L}_{\text{Yukawa}}$.

Gauge Kinetic Terms

The gauge invariant kinetic term for the gauge bosons is given by

$$\mathcal{L}_{\text{gauge kinetic}} = -\frac{1}{4}G^{\mu\nu}_A(x)G^A_{\mu\nu}(x) - \frac{1}{4}W^{\mu\nu}_I(x)W^I_{\mu\nu}(x) - \frac{1}{4}B^{\mu\nu}(x)B_{\mu\nu}(x), \tag{2.3}$$

where $G^A_{\mu\nu}(x)$, $A = 1,\dots 8$, are the field strength tensors for $SU(3)_C$, $W^I_{\mu\nu}(x)$, $I = 1,2,3$, for $SU(2)_L$ and $B_{\mu\nu}(x)$ the Abelian gauge group $U(1)_Y$ given by

$$G^A_{\mu\nu}(x) = \partial_\mu G^A_\nu(x) - \partial_\nu G^A_\mu(x) - g_s f^{ABC} G^B_\mu(x) G^C_\nu(x), \tag{2.4}$$

$$B_{\mu\nu}(x) = \partial_\mu B_\nu(x) - \partial_\nu B_\mu(x) \quad \text{and} \tag{2.5}$$

$$W^I_{\mu\nu}(x) = \partial_\mu W^I_\nu(x) - \partial_\nu W^I_\mu(x) - g_2 \epsilon^{IJK} W\mu^J(x) W^K_\nu(x), \tag{2.6}$$

where g_s is the strong gauge coupling constant, g_1 the hypercharge coupling and g_2 the weak isospin coupling. The tensor elements f^{ABC} are the structure constants of the $SU(3)_C$ Lie algebra defined by

$$\left[t^A, t^B\right] = i f^{ABC} t^C, \tag{2.7}$$

while elements of the the totally antisymmetric tensor ϵ^{IJK} are the structure constants of the $SU(2)_L$ Lie algebra. The strong and weak gauge fields transform non-trivially under the non-Abelian gauge groups $SU(3)_C$ and $SU(2)_L$, respectively, i.e. introducing three- and four-point self-interactions.

Fermion Kinetic Terms

The fermion kinetic terms of the electroweak Lagrangian are given in each generation by

$$\begin{aligned}\mathcal{L}_{\text{fermion kinetic}} =& \bar{\boldsymbol{Q}}_L(x)(i\not{D}_L)\boldsymbol{Q}_L(x) + \bar{\boldsymbol{u}}_R(x)(i\not{D}_R)\boldsymbol{u}_R(x) + \bar{\boldsymbol{d}}_R(x)(i\not{D}_R)\boldsymbol{d}_R(x) \\ &+ \bar{\boldsymbol{L}}_L(x)(i\not{D}_L)\boldsymbol{L}_L(x) + \bar{\boldsymbol{\ell}}_R(x)(i\not{D}_L)\boldsymbol{\ell}_R(x),\end{aligned} \tag{2.8}$$

with $\not{D} = \gamma^\mu D_\mu$ and the adjusted spinors $\bar{\boldsymbol{Q}} = \gamma^0 \boldsymbol{Q}^\dagger$ etc. Since right-handed fermion fields do not couple to the weak isospin, the covariant derivatives $D_{\mu_{L,R}}$ are defined as

$$\begin{aligned} D_{\mu_R}\psi_R(x) &= \left[\partial_\mu + ig_1 \frac{Y_W}{2} B_\mu(x)\right] \psi_R(x) \quad \text{and} \\ D_{\mu_L}\psi_L(x) &= \left[\mathbb{1}\left(\partial_\mu + ig_1 \frac{Y_W}{2} B_\mu(x)\right) + ig_2 \frac{1}{2}\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu(x)\right] \psi_L(x), \end{aligned} \tag{2.9}$$

where the elements of the weak isospin vector $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ are the two-dimensional Pauli matrices and $W_\mu = \left(W_\mu^1, W_\mu^2, W_\mu^3\right)$ is the weak isospin vector of the $SU(2)_L$ gauge fields.

Scalar Sector

In contradiction to the experimental observation, the weak gauge symmetry predicts massless weak gauge bosons and, in combination with parity violation, massless fermions as in Equation 2.3 and 2.8. Explicit mass terms $-\frac{m^2}{2} A_\mu A^\mu$ for gauge bosons would violate the local gauge symmetry and Dirac mass terms

$$-m\bar{\psi}_D\psi_D = -m\bar{\psi}_D\left(P_L^2 + P_R^2\right)\psi_D = m\left(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L\right), \tag{2.10}$$

for the fermions the global $SU(2)_L$ invariance. This conflict is solved by the Brout-Englert-Higgs mechanism of electroweak symmetry breaking [8, 9, 15–19], which will be discussed in detail in Section 2.3. The minimal field content needed to give masses to the three weak gauge bosons is a complex scalar $SU(2)_L$ doublet

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.11}$$

with hypercharge $Y_W = 1$ which is also a colour singlet . The minimal scalar Lagrangian invariant under $SU(2)_L \otimes U(1)_Y$ transformations preserving renormalisability of the theory [20–22] and allowing for the electroweak symmetry breaking is given by

$$\mathcal{L}_{\text{scalar}} = (D_{\mu_L}\Phi)^\dagger (D_L^\mu \Phi) - V(\Phi), \tag{2.12}$$

with the electroweak covariant derivative as in Equation 2.9 and the Higgs potential

$$V(\Phi) = -\mu^2 \Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \tag{2.13}$$

where $\lambda > 0$ is the dimensionless self-coupling and $-\mu^2$ a mass parameter.

Yukawa Interactions

With the same scalar doublet field finite masses of the fermions can be obtained after the electroweak symmetry breaking via Yukawa interactions with the Higgs field [23] of the form

$$\mathcal{L}_{\text{Yukawa}} = -\left(\bar{L}_L^i Y_{ij}^\ell \Phi \ell_R^j + \bar{Q}_L^i Y_{ij}^u \widetilde{\Phi} u_R^j + \bar{Q}_L^i Y_{ij}^d \Phi d_R^j + \text{h.c.}\right), \tag{2.14}$$

where Y^{ℓ}_{ij}, Y^{u}_{ij} and Y^{d}_{ij} are the Yukawa coupling matrices with the generation indices $i, j = 1, 2, 3$ and $\widetilde{\Phi} = i\tau_2 \Phi^*$.

2.3 The Higgs Mechanism

The potential of the Higgs field (Equation 2.13) after the electroweak symmetry breaking with $-\mu^2 > 0$ (illustrated in Figure 2.2) has an infinite set of degenerate ground states where the Higgs potential has a minimum at the vacuum expectation value v of the Higgs field:

$$\langle \Phi^\dagger \Phi \rangle = \frac{v^2}{2} \quad \text{and} \quad v \equiv \sqrt{\frac{-\mu^2}{\lambda}}. \tag{2.15}$$

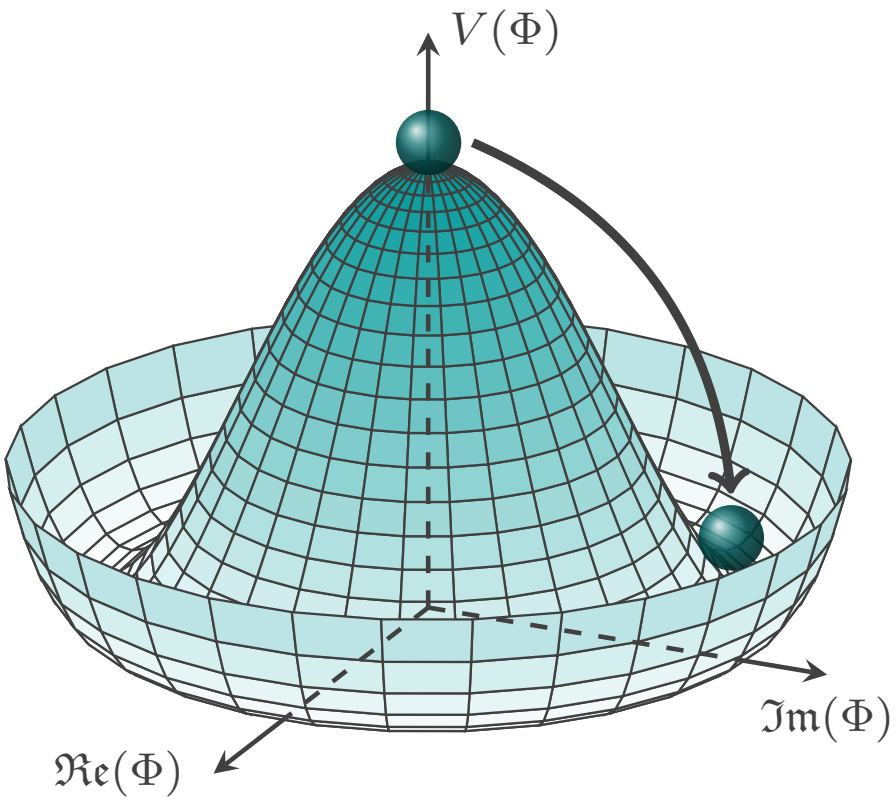


Figure 2.2: Illustration of the potential of a complex scalar field Φ.

Due to electric charge conservation, only the neutral component of the scalar doublet can acquire a finite vacuum expectation value. Therefore,

$$\langle\Phi\rangle_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}, \tag{2.16}$$

can be chosen as the ground state without restrictions. The choice of a particular vacuum state breaks the $SU(2)_L \otimes U(1)_Y$ gauge symmetry spontaneously. Invariance under the electromagnetic $U(1)_Q$ gauge symmetry is still preserved:

$$Q\langle\Phi\rangle_0 = \left(I_3 + \frac{Y_W}{2}\right)\langle\Phi\rangle_0 \equiv 0. \tag{2.17}$$

Three massless tangential excitations of the Higgs field are absorbed in the longitudinal polarisation of the weak gauge bosons, which acquire their masses in this way. In the unitary gauge, the scalar field $\Phi(x)$ can be expanded around the ground state as

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+h(x)\end{pmatrix}, \tag{2.18}$$

where the massive radial excitation $h(x)$ is the physical Higgs field. By inserting Equation 2.18 into $\mathcal{L}_{\text{scalar}}$ one obtains

$$\begin{aligned}\mathcal{L}_{\text{scalar}} = {} & \frac{1}{2}\partial_\mu h(x)\partial^\mu h(x) - \lambda v^2 h^2(x) - \lambda v h^3(x) - \frac{\lambda}{4}h^4(x) \\ & + \frac{g_2^2}{4}\left(2vh(x) + h^2(x)\right)\left(W_\mu^+(x)W^{-\mu}(x) + \frac{1}{2\cos^2\theta_W}Z_\mu(x)Z^\mu(x)\right) \\ & + \frac{g_2^2}{8}v^2\left(W_\mu^+(x)W^{+\mu}(x) + W_\mu^-(x)W^{-\mu}(x)\right) + \frac{g_2^2 v^2}{4\cos^2\theta_W}Z_\mu(x)Z^\mu(x).\end{aligned} \tag{2.19}$$

The first and the second term in the scalar Lagrangian are the kinetic and the mass term of the Higgs field. The triple and quadratic Higgs self-couplings are described by the two following terms. Interactions between the Higgs and the weak vector bosons appear in the second line of Equation 2.19 while the third line contains the mass terms of the weak gauge bosons. The charged gauge fields $W_\mu^\pm$ are linear combinations of the W_μ^1 and W_μ^2 isotriplet gauge fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right). \tag{2.20}$$

The mass eigenstates Z_μ and A_μ of the neutral electroweak gauge fields are obtained by rotation

$$Z_\mu = \frac{-g_1 B_\mu + g_2 W_\mu^3}{\sqrt{g_1^2+g_2^2}} \quad \text{and} \quad A_\mu = \frac{g_1 B_\mu + g_2 W_\mu^3}{\sqrt{g_1^2+g_2^2}} \tag{2.21}$$

around the Weinberg angle θ_W defined by

$$\sin\theta_W = \frac{g_1}{\sqrt{g_1^2+g_2^2}} = \frac{e}{g_2} \quad \text{and} \quad \cos\theta_W = \frac{g_2}{\sqrt{g_1^2+g_2^2}}, \tag{2.22}$$

where e is the elementary charge, the coupling constant of the electromagnetic interaction.

According to the third line of Equation 2.19, the masses of the three massive weak gauge bosons and of the Higgs boson are given to lowest order of perturbation theory by

$$\begin{aligned} m_{W^\pm} &= \frac{g_2 v}{2}, \\ m_Z &= \frac{g_2 v}{2\cos\theta_W} = \frac{m_{W^\pm}}{\cos\theta_W} \quad \text{and} \\ m_h &= v\sqrt{2\lambda}. \end{aligned} \tag{2.23}$$

The measured values [24] are $m_{W^\pm} = 80.379 \pm 0.012$ GeV, $m_Z = 91.1876 \pm 0.0021$ GeV and $m_h = 125.09 \pm 0.24$ GeV.
According to Equation 2.14, in the basis of the fermion mass eigenstates f_i the masses of the fermions proportional to the Yukawa couplings eigenvalues Y_i

$$\mathcal{L}_{\text{Yukawa}} = -\frac{v + h(x)}{\sqrt{2}} Y_i \left(\bar{f}_{i_L} f_{i_R} + \bar{f}_{i_R} f_{i_L} \right), \tag{2.24}$$

are

$$m_i = Y_i \frac{v}{\sqrt{2}}, \tag{2.25}$$

where i runs over all fermion flavours.

The coupling factors for the Higgs boson self-interactions in amplitude calculations [25] are given by

$$g_{h^3} \propto (3!)\,\lambda v = 3\frac{m_h^2}{v} \quad \text{and} \quad g_{h^4} \propto (4!)\,\frac{\lambda}{4} = 3\frac{m_h^2}{v^2}, \tag{2.26}$$

where the factors $n!$ account for the number n of identical Higgs particles in the interaction vertex (see Figure 2.3). The coupling factors for Higgs boson interactions with gauge bosons $V = W^\pm, Z$ and for fermions are

$$g_{hVV} \propto 2\frac{m_V^2}{v}, \quad g_{hhVV} \propto 2\frac{m_V^2}{v^2} \quad \text{and} \quad g_{hff} \propto \frac{m_f}{v}, \tag{2.27}$$

(see Figure 2.3). These coupling parameters of the Higgs boson to the SM particles are proportional to the fermion masses and to the squared of the weak gauge boson masses with $1/v$ as proportionality factor. The Higgs boson mass itself and the self-coupling constant λ are free parameter of the SM. Once they have been measured, all Higgs boson production and decay rates can be predicted by the SM.

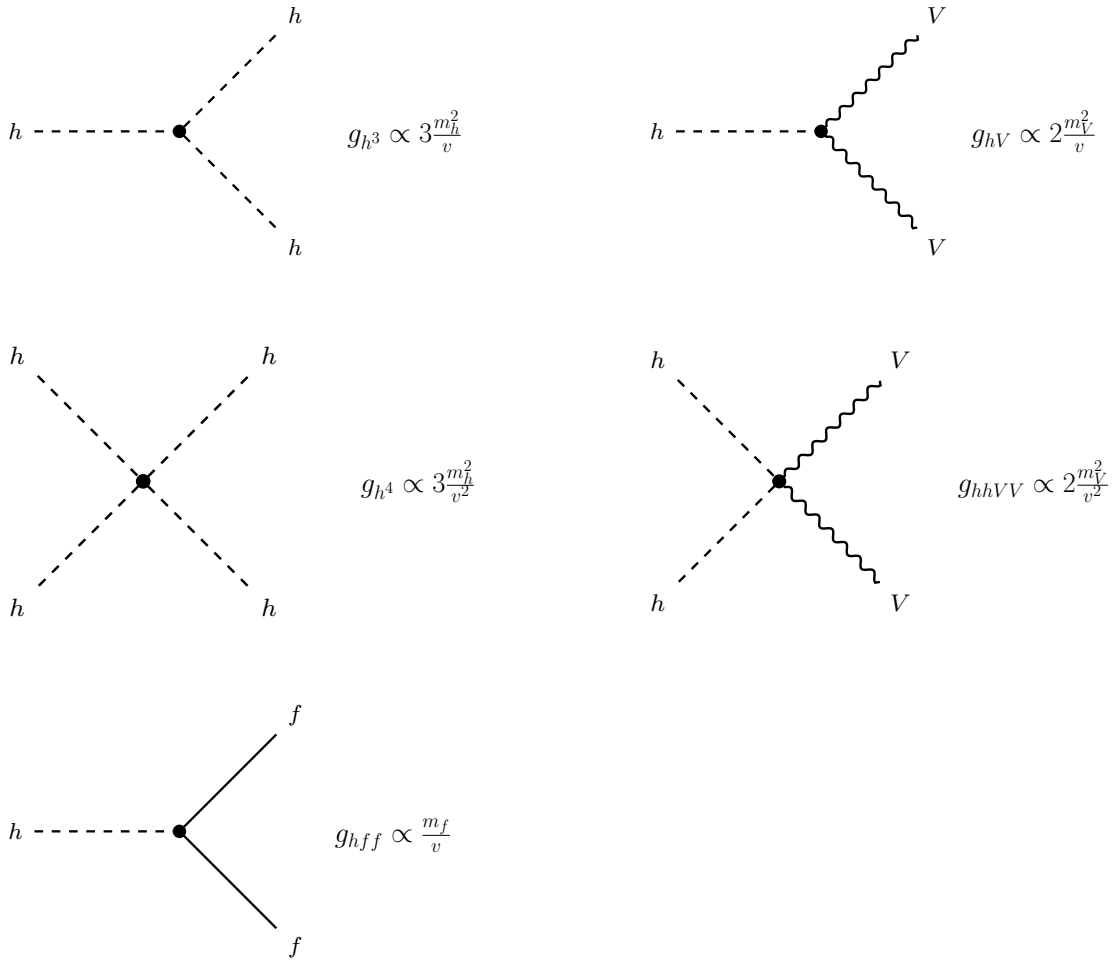


Figure 2.3: Couplings of the Higgs boson in the Standard Model.

2.4 Phenomenology of the Standard Model Higgs Boson

In the following, the spin and CP properties of the SM Higgs boson are introduced and its production in proton-proton collisions as well as its decay modes.

2.4.1 Spin and CP Properties

The CP quantum numbers of a particle describe the behaviour under charge (C) conjugation and parity (P) transformations. CP eigenstates with eigenvalue +1 are CP-even and do not change the sign under CP transformations while CP-odd states with eigenvalue −1 change the sign. According to their spin J and parity P bosonic particles are classified as scalars, pseudo-scalars, vectors and pseudo-vectors (see Table 2.2). The Higgs boson in the SM is a CP-even spin-0, i.e. scalar particle. For the determination of the Higgs boson coupling properties, its quantum numbers has to be identified. After the discovery of the Higgs-like boson by the ATLAS and CMS experiments [10, 11], the quantum numbers have been probed [26–28] (see Section 2.6) and are in accordance with the SM predictions.

Table 2.2: Spin and CP quantum numbers of bosonic particles in the SM

Quantum number	scalar	pseudo-scalar	vector	pseudo-vector
Spin: J	0	0	1	1
Charge conjugation: C	+1	+1	+1	+1
Parity: P	+1	−1	−1	+1
J^{CP}	0^{++}	0^{+-}	1^{+-}	1^{++}
Example	SM Higgs boson	π, K, η, η'	Z, W, γ, g	f_1, a_1

Small CP-even and CP-odd admixtures to the SM CP-even state of the Higgs field are predicted by theories beyond the Standard Model (BSM) [29]. In the case of CP-odd admixtures, the observed Higgs boson would not be a CP-eigenstate and CP would be violated in Higgs boson production and decays.

By measuring the tensor structure of the Higgs boson couplings to SM particles, its CP-eigenstate composition can be investigated. The tensor structure can be studied using distributions of kinematic variables of the Higgs boson decay and associated production products as well as in relative rates of different production modes.

2.4.2 Production in Proton-Proton Collisions

The couplings of the Higgs boson to SM particles are proportional to the particle masses. Therefore, the production of the Higgs boson in hadron collisions (Section 3.2) is dominated by processes involving heavy particles, e.g. the top and bottom quark and the massive weak bosons [30]. The main production mechanisms are gluon fusion (ggF), vector boson fusion (VBF), production in association with a W or Z boson (WH and ZH) and the associated production with top (ttH) or bottom (bbH) quark pairs or with single top quarks (tH). Figure 2.4 shows the cross sections of the dominant production modes as a function of the proton-proton centre-of-mass energy.

The dominant production mode with a relative rate of 88% is ggF shown in Figure 2.5(a). Since the gluons do not couple directly to the Higgs boson, it is produced via an intermediate quark loop. Because the Higgs boson couplings are proportional to the mass, the largest contribution is the top quark loop. With about 7%, the next frequent production mode is VBF (Figure 2.5(b)) where the Higgs boson is produced in the scattering of two weak gauge bosons emitted from the incoming quarks, which produce two energetic jets in the forward regions of the detector. This signature provides a powerful discrimination between the Higgs boson signal and the QCD background. The production mode in association with vector bosons (VH), also called Higgs strahlung, contributes with 4%.

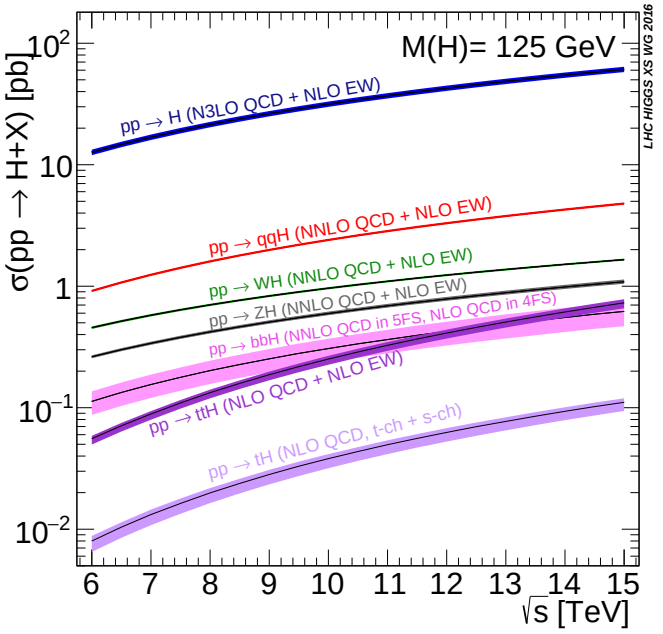
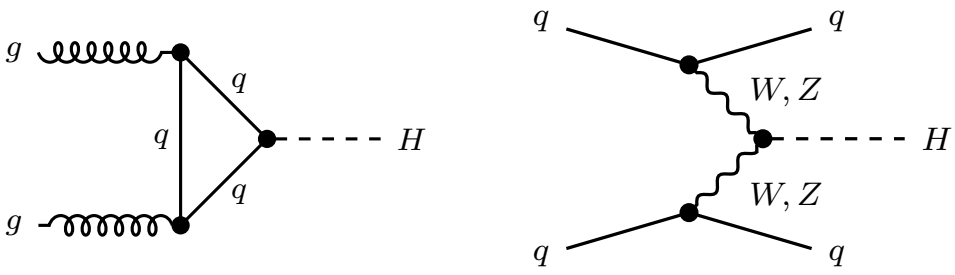


Figure 2.4: Cross sections of the dominant production modes of the Standard Model Higgs boson with a mass of $m_H = 125$ GeV in proton-proton collisions at the Large Hadron Collider as a function of the centre-of-mass energy $\sqrt{s}$.

A W or Z boson created in quark-antiquark annihilation or gluon fusion emits the Higgs boson (see Figure 2.6). The smallest contributions of about 1% come from the production in association with heavy quark pairs, ttH and bbH, as well as tH production with additional W bosons or quark jets shown in Figure 2.7 and 2.8, respectively. Table 2.3 summarises the predictions for the main SM Higgs boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV for $m_H = 125$ GeV.



(a) (b)

Figure 2.5: Tree-level Feynman diagrams for the Standard Model Higgs boson production in proton-proton collisions via (a) gluon fusion and (b) vector boson fusion.

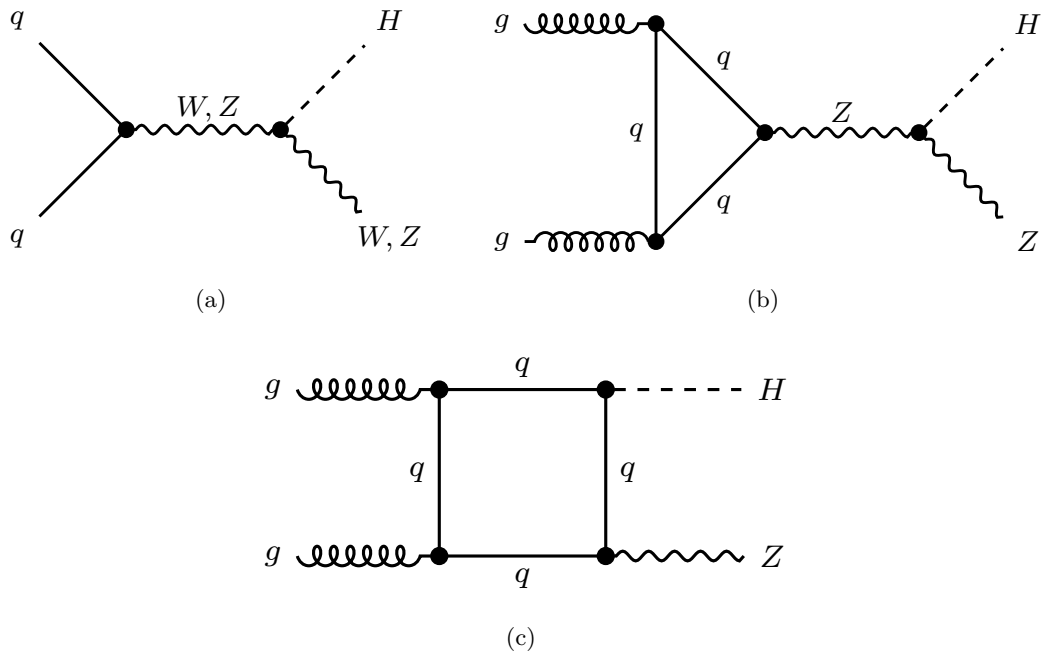


Figure 2.6: Feynman diagrams for the Standard Model Higgs boson production in proton-proton collisions (a) at tree-level in association with weak gauge bosons and (b) and (c) with loop processes via $gg \to ZH$.

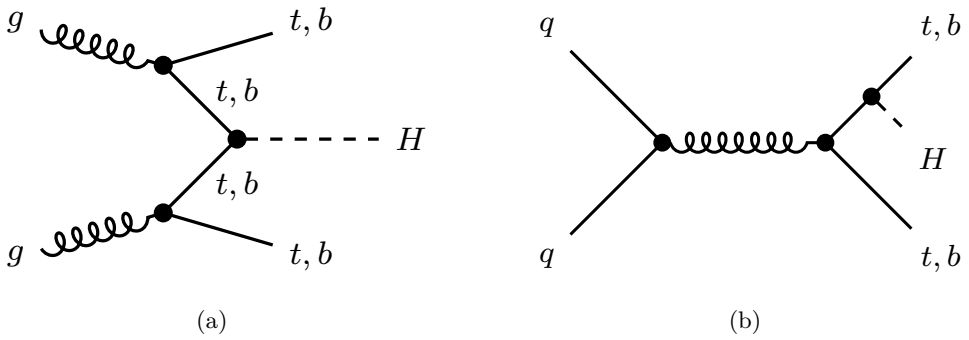


Figure 2.7: Tree-level Feynman diagrams for the Standard Model Higgs boson production in proton-proton collisions in association with heavy quark pairs.

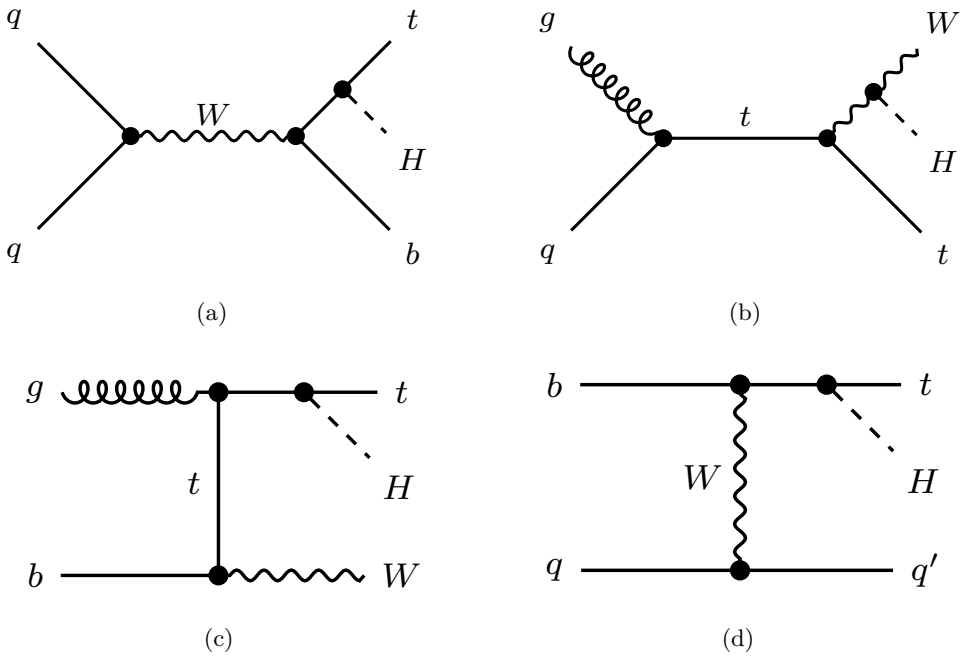


Figure 2.8: Tree-level s- and t-channel Feynman diagrams for the Standard Model Higgs boson production in proton-proton collisions in association with a top quark and (a) a bottom quark, (b) and (c) a W boson and (d) a light or heavy quark.

Table 2.3: Production cross sections for the dominant production modes of the SM Higgs boson with a mass of $m_H = 125$ GeV in proton-proton collisions at the Large Hadron Collider at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The quoted uncertainties correspond to the total theoretical uncertainties calculated by adding in quadrature the QCD scale, PDF and α_s uncertainties [31].

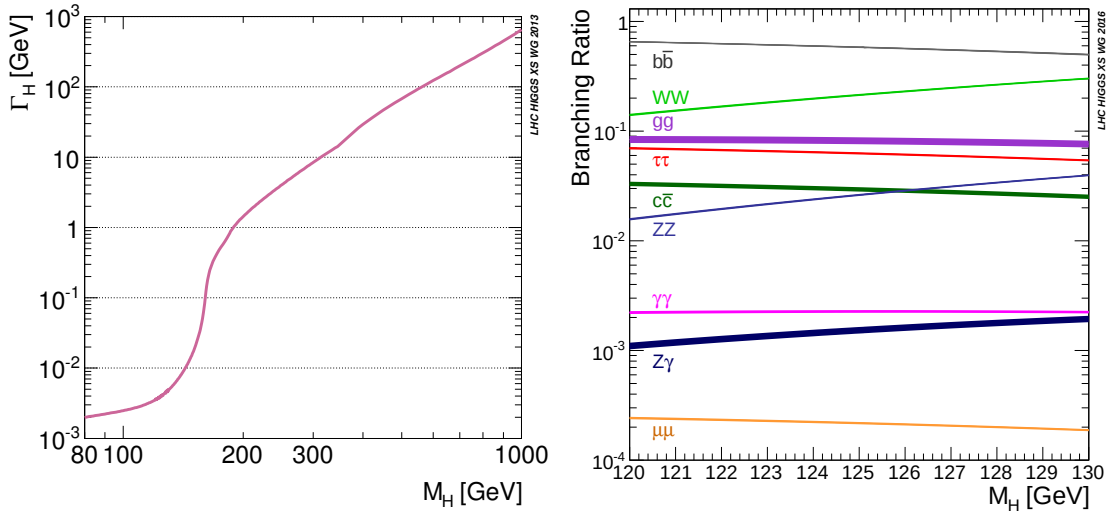
Production mode	Production mechanism	Cross section [pb]
ggF	$gg \to H$	$48.6\ ^{+\ 5.62\%}_{-\ 7.43\%}$
VBF	$qq \to H$	$3.78\ ^{+\ 2.20\%}_{-\ 2.18\%}$
WH	$qq \to WH$	$1.37\ ^{+\ 1.99\%}_{-\ 2.05\%}$
ZH	$gg/qq \to ZH$	$0.884\ ^{+\ 4.12\%}_{-\ 3.48\%}$
ttH	$gg/qq \to t\bar{t}H$	$0.507\ ^{+\ 6.83\%}_{-\ 9.88\%}$
bbH	$gg/qq \to b\bar{b}H$	$0.488\ ^{+20.20\%}_{-23.90\%}$
tH	$qq \to tHq$ (t-ch.)	$0.0743\ ^{+\ 7.48\%}_{-15.35\%}$
	$qq \to tHb$ (s-ch.)	$0.00288^{+\ 3.26\%}_{-\ 2.85\%}$
	$gb \to tHW$ (W-ass.)	$0.0152\ ^{+\ 7.97\%}_{-\ 9.18\%}$

2.4.3 Decay Modes

The Higgs boson decay branching ratios ($\mathcal{B}$) are defined as the ratios of particular partial decay widths and the total decay width:

$$\mathcal{B}(H \to X) = \frac{\Gamma(H \to X)}{\sum_Y (H \to Y)}. \tag{2.28}$$

The decay widths depend only on the square of the Higgs coupling to the decay products which is directly proportional to their mass and on phase space factors including the Higgs boson mass. The Higgs boson decays preferentially into the heaviest particles allowed by energy conservation. The total Higgs boson decay width and the branching ratios of the SM Higgs boson decays as a function of the Higgs boson mass m_H are shown in Figure 2.9(a) and 2.9(b), respectively.

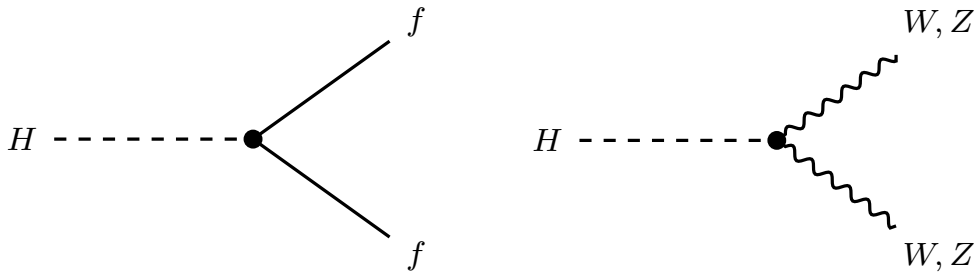


(a) (b)

Figure 2.9: (a) Total Higgs boson decay width and (b) branching ratios of the SM Higgs boson decays as a function of the Higgs boson mass m_H [31].

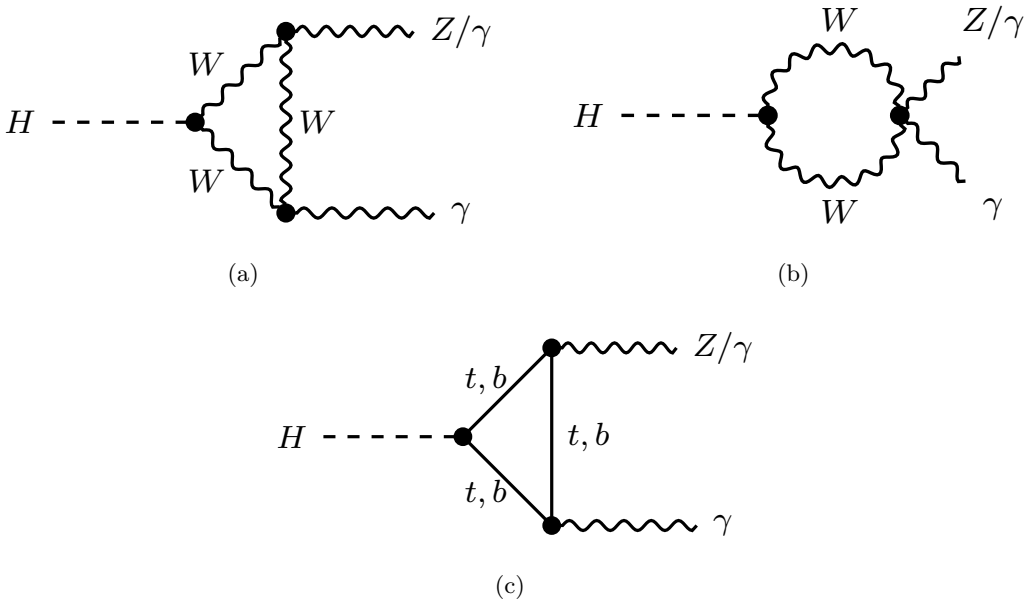
The most frequent decay mode of the SM Higgs boson with a mass of $m_H = 125$ GeV is the decay into pairs of bottom quarks, followed by the decay into W bosons. The decay channels with the next highest branching ratio are the decays into gluons, τ leptons and charm quark and Z boson pairs. Decays into pairs of photons, $Z\gamma$ and into pairs of muons have very small branching ratios at the per mill level. Since photons and gluons do not couple directly to the Higgs boson, the decays into these particles is only possible via loop processes involving heavy particles. The Feynman diagrams for the decays into fermions and weak gauge bosons are shown in Figure 2.10 and for decays into photon pairs and $Z\gamma$ in Figure 2.11. The branching fractions are summarised in Table 2.4.

Due to the large QCD background of hadron colliders, fully hadronic final states are experimentally difficult, e.g. the decays into bottom quark pairs and, even more, into gluons are light quarks. The largest sensitivity is achieved for final states with leptons or photons which can be easily identified. Therefore, Higgs boson decay into weak gauge bosons pairs which decay into leptons and into photon pairs are central for the study of the properties of the Higgs boson.



(a) (b)

Figure 2.10: Tree-level Feynman diagrams for the Standard Model Higgs boson decays into (a) fermions and (b) weak gauge bosons.



(a) (b) (c)

Figure 2.11: Tree-level Feynman diagrams for the Standard Model Higgs boson decays into two photons or $Z\gamma$ mediated by (a) and (b) W bosons and (c) heavy quarks.

Table 2.4: Predictions of the SM Higgs boson decay branching ratios for $m_H = 125$ GeV. The quoted theoretical uncertainties are due to missing higher order corrections and input parameter uncertainties in the calculation and have been added in quadrature [31].

Decay mode	Decay process	Branching ratio [%]
bb	$H \to b\bar{b}$	$58.24\ ^{+1.24\%}_{-1.27\%}$
WW	$H \to WW$	$21.37\ ^{+1.55\%}_{-1.53\%}$
gg	$H \to gg$	$8.187\ ^{+5.14\%}_{-5.09\%}$
$\tau\tau$	$H \to \tau\tau$	$6.272\ ^{+1.65\%}_{-1.65\%}$
cc	$H \to c\bar{c}$	$2.891\ ^{+5.54\%}_{-1.99\%}$
ZZ	$H \to ZZ$	$2.619\ ^{+1.55\%}_{-1.53\%}$
$\gamma\gamma$	$H \to \gamma\gamma$	$0.2270\ ^{+2.06\%}_{-2.08\%}$
$Z\gamma$	$H \to Z\gamma$	$0.1533\ ^{+5.82\%}_{-5.83\%}$
$\mu\mu$	$H \to \mu\mu$	$0.02176^{+1.67\%}_{-1.70\%}$

2.5 Beyond the Standard Model Higgs Boson Analysis Frameworks

The measurement of the production cross sections in exclusive phase space regions is sensitive to the tensor structure of the Higgs boson couplings. Any deviation of the measurements from the SM predictions would indicate BSM physics. Since significant deviations from the SM have not been observed so far (see Section 2.6), only small BSM effects are still allowed. Deviations are parametrised independently of the specific extensions of the SM by means of the so-called κ-framework or effective field theories approach, introduced in the following.

2.5.1 The Coupling Modification Framework

The measured Higgs boson cross sections can be interpreted to leading-order in the so-called κ-framework [29] in which multiplicative coupling modifiers $\boldsymbol{\kappa}$ are introduced to parametrise deviations of the Higgs boson couplings from the SM predictions under the assumption of a single CP-even Higgs boson state and of the SM tensor coupling structure. Only the coupling strengths are allowed to be modified by BSM physics. The narrow-width approximation for the Higgs boson is assumed, such that production and the decay can be factorised,

$$\sigma \cdot \mathcal{B}\ (i \to H \to f) = \sigma_i(\boldsymbol{\kappa}) \cdot \frac{\Gamma_f(\boldsymbol{\kappa})}{\Gamma_H}, \tag{2.29}$$

where σ_i is the production cross section from the initial state i, $\mathcal{B}$ and Γ_f are the branching ratio and partial decay width for the decay into the final state f, respectively, and Γ_H is the total width of the Higgs boson. For Higgs boson production and decay process via couplings i and f, respectively, coupling strength modifiers

$$\kappa_i^2 = \frac{\sigma_i}{\sigma_i^{\mathrm{SM}}} \quad \text{and} \quad \kappa_f^2 = \frac{\Gamma_f}{\Gamma_k^{\mathrm{SM}}}, \tag{2.30}$$

are defined.

2.5.2 Effective Field Theory Frameworks

An Effective Field Theory (EFT) parametrises new physics occurring at an energy scale Λ much larger than the electroweak scale v. Interactions of new particles are integrated out and absorbed into operators consisting of the SM fields with dimensions larger than the operators of the SM Lagrangian to which they are added.

In the following, two EFT versions are introduced. In the Standard Model Effective Field Theory (SMEFT) [32] a complete set of dimension-six operators invariant under the SM gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ is constructed from the SM fields. The Higgs Characterisation (HC) model [33] is formulated in terms of the mass eigenstates of the SM fields after the electroweak symmetry breaking and the effective Higgs boson couplings to the SM gauge bosons and fermions are described by in total 20 parameters. Since the HC model does not include all possible extensions of the SM Lagrangian predicted by dimension-six operators, it is less general than the SMEFT.

2.5.2.1 The Standard Model Effective Field Theory

In the SMEFT, the SM Lagrangian is extended by higher-dimension operators which have the same field content and the same, linearly realised, $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ local gauge symmetry as the SM. The general effective Lagrangian takes the form

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \mathcal{L}^{(7)} + \ldots, \quad \text{with} \quad \mathcal{L}^{(D)} = \sum_i^{n_D} \frac{C_i^{(D)}}{\Lambda^{D-4}} O_i^{(D)} \text{ for } D > 4, \tag{2.31}$$

where Λ is the energy scale at which new physics is assumed to appear, $O_i^{(D)}$ are the operators of the dimension-D invariant under the SM gauge group and $C_i^{(D)}$ are the corresponding dimensionless coupling constants, the so-called Wilson coefficients. In the present studies, Λ is set to 1 TeV, since current experimental results show no evidence for new physics up to this scale.

Complete sets of operators are known for dimensions up to eight [34–41]. Dimension-five operators violates lepton number (L conservation), while dimension-seven operators violate the observed $B-L$ symmetry, where B is the baryon number. Previous experiments showed [42] that the L and $B-L$ violating terms are suppressed such that they are not

observable at LHC. They are, therefore, neglected here. Thus, the leading contributions of new physics are from dimension-six operators and the SMEFT Lagrangian becomes

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{i}^{n_6} c_i O_i, \tag{2.32}$$

where the $1/\Lambda^2$ factor is absorbed into the Wilson coefficients, $c_i = C_i/\Lambda^2$.

There are several possible complete sets of these dimension-six operators. The SMEFT used here is formulated in the Warsaw basis [34], which contains 59 independent operators assuming lepton and baryon number conservation. These operators are divided into eight classes according to the field content and the number of covariant derivatives, class 1: X^3, class 2: Φ^6, class 3: $\Phi^4 D^2$, class 4: $X^2\Phi^2$, class 5: $\psi^2\Phi^3$, class 6: $\psi^2 X\Phi$, class 7: $\psi^2\Phi^2 D$ and class 8: ψ^4, where $X = G^A_{\mu\nu}, W^I_{\mu\nu}, B_{\mu\nu}$ are the gauge field strength tensors, Φ the scalar doublet Higgs field, ψ the fermion spinor of $SU(2)_L$ eigenstates and D_μ the covariant derivative. The operators of class 1 to 7 are shown in Table 2.5, the class 8 operators in Table 2.5.

Table 2.5: Class 1 to 7 of dimension-six operators built from Standard Model gauge fields $X = G^A_{\mu\nu}, W^I_{\mu\nu}, B_{\mu\nu}$, the scalar Higgs doublet Φ and the fermion fields ψ with flavour indices p, r, s, t conserving lepton and baryon number in the so-called Warsaw basis defined in [34]

Class 1: X^3	
O_G	$f^{ABC} G^{A\nu}_{\mu} G^{B\rho}_{\nu} G^{C\mu}_{\rho}$
$O_{\widetilde{G}}$	$f^{ABC} \widetilde{G}^{A\nu}_{\mu} G^{B\rho}_{\nu} G^{C\mu}_{\rho}$
O_W	$\epsilon^{IJK} W^{I\nu}_{\mu} W^{J\rho}_{\nu} W^{K\mu}_{\rho}$
$O_{\widetilde{W}}$	$\epsilon^{IJK} \widetilde{W}^{I\nu}_{\mu} W^{J\rho}_{\nu} W^{K\mu}_{\rho}$

Class 2 and 3: Φ^6 and $\Phi^4 D^2$	
O_Φ	$(\Phi^\dagger\Phi)^3$
$O_{\Phi\Box}$	$(\Phi^\dagger\Phi)\Box(\Phi^\dagger\Phi)$
$O_{\Phi D}$	$(\Phi^\dagger D^\mu\Phi)^*(\Phi^\dagger D_\mu\Phi)$

Class 5: $\psi^2\Phi^3 + \text{h.c.}$	
$O_{\ell\Phi}$	$(\Phi^\dagger\Phi)(\bar{L}_p \ell_r \Phi)$
$O_{u\Phi}$	$(\Phi^\dagger\Phi)(\bar{Q}_p u_r \widetilde{\Phi})$
$O_{d\Phi}$	$(\Phi^\dagger\Phi)(\bar{Q}_p d_r \Phi)$

Class 4: $X^2\Phi^2$	
$O_{\Phi G}$	$\Phi^\dagger\Phi G^A_{\mu\nu} G^{A\mu\nu}$
$O_{\Phi\widetilde{G}}$	$\Phi^\dagger\Phi \widetilde{G}^A_{\mu\nu} G^{A\mu\nu}$
$O_{\Phi W}$	$\Phi^\dagger\Phi W^I_{\mu\nu} W^{I\mu\nu}$
$O_{\Phi\widetilde{W}}$	$\Phi^\dagger\Phi \widetilde{W}^I_{\mu\nu} W^{I\mu\nu}$
$O_{\Phi B}$	$\Phi^\dagger\Phi B_{\mu\nu} B^{\mu\nu}$
$O_{\Phi\widetilde{B}}$	$\Phi^\dagger\Phi \widetilde{B}_{\mu\nu} B^{\mu\nu}$
$O_{\Phi WB}$	$\Phi^\dagger\tau^I\Phi W^I_{\mu\nu} B^{\mu\nu}$
$O_{\Phi\widetilde{W}B}$	$\Phi^\dagger\tau^I\Phi \widetilde{W}^I_{\mu\nu} B^{\mu\nu}$

Class 6: $\psi^2 X\Phi + \text{h.c.}$	
$O_{\ell W}$	$(\bar{L}_p \sigma^{\mu\nu} \ell_r)\tau^I \Phi W^I_{\mu\nu}$
$O_{\ell B}$	$(\bar{L}_p \sigma^{\mu\nu} \ell_r)\Phi B_{\mu\nu}$
O_{uG}	$(\bar{Q}_p \sigma^{\mu\nu} t^A u_r)\widetilde{\Phi} G^A_{\mu\nu}$
O_{uW}	$(\bar{Q}_p \sigma^{\mu\nu} u_r)\tau^I \widetilde{\Phi} W^I_{\mu\nu}$
O_{uB}	$(\bar{Q}_p \sigma^{\mu\nu} u_r)\widetilde{\Phi} B_{\mu\nu}$
O_{dG}	$(\bar{Q}_p \sigma^{\mu\nu} t^A d_r)\Phi G^A_{\mu\nu}$
O_{dW}	$(\bar{Q}_p \sigma^{\mu\nu} d_r)\tau^I \Phi W^I_{\mu\nu}$
O_{dB}	$(\bar{Q}_p \sigma^{\mu\nu} d_r)\Phi B_{\mu\nu}$

Class 7: $\psi^2\Phi^2 D$	
$O^{(1)}_{\Phi L}$	$(\Phi^\dagger i \overleftrightarrow{D_\mu} \Phi)(\bar{L}_p \gamma^\mu L_r)$
$O^{(3)}_{\Phi L}$	$(\Phi^\dagger i \overleftrightarrow{D^I_\mu} \Phi)(\bar{L}_p \tau^i \gamma^\mu L_r)$
$O_{\Phi\ell}$	$(\Phi^\dagger i \overleftrightarrow{D_\mu} \Phi)(\bar{\ell}_p \gamma^\mu \ell_r)$
$O^{(1)}_{\Phi Q}$	$(\Phi^\dagger i \overleftrightarrow{D_\mu} \Phi)(\bar{Q}_p \gamma^\mu Q_r)$
$O^{(3)}_{\Phi Q}$	$(\Phi^\dagger i \overleftrightarrow{D^i_\mu} \Phi)(\bar{Q}_p \tau^i \gamma^\mu Q_r)$
$O_{\Phi u}$	$(\Phi^\dagger i \overleftrightarrow{D_\mu} \Phi)(\bar{u}_p \gamma^\mu u_r)$
$O_{\Phi d}$	$(\Phi^\dagger i \overleftrightarrow{D_\mu} \Phi)(\bar{d}_p \gamma^\mu d_r)$
$O_{\Phi ud} + \text{h.c.}$	$i(\widetilde{\Phi}^\dagger D_\mu \Phi)(\bar{u}_p \gamma^\mu d_r)$

Table 2.6: Class 8 of dimension-six operators built from Standard Model fermion fields ψ with flavour indices p, r, s, t conserving lepton and baryon number in the so-called Warsaw basis defined in [34]

Class 8: $(\bar{\psi}_L\psi_L)(\bar{\psi}_L\psi_L)$		Class 8: $(\bar{\psi}_R\psi_R)(\bar{\psi}_R\psi_R)$		Class 8: $(\bar{\psi}_L\psi_L)(\bar{\psi}_R\psi_R)$	
O_{LL}	$(\bar{L}_p\gamma_\mu L_r)(\bar{L}_s\gamma^\mu L_t)$	$O_{\ell\ell}$	$(\bar{\ell}_p\gamma_\mu \ell_r)(\bar{\ell}_s\gamma^\mu \ell_t)$	$O_{L\ell}$	$(\bar{L}_p\gamma_\mu L_r)(\bar{\ell}_s\gamma^\mu \ell_t)$
$O_{QQ}^{(1)}$	$(\bar{Q}_p\gamma_\mu Q_r)(\bar{Q}_s\gamma^\mu Q_t)$	O_{uu}	$(\bar{u}_p\gamma_\mu u_r)(\bar{u}_s\gamma^\mu u_t)$	O_{Lu}	$(\bar{L}_p\gamma_\mu L_r)(\bar{u}_s\gamma^\mu u_t)$
$O_{QQ}^{(3)}$	$(\bar{Q}_p\gamma_\mu\tau^I Q_r)(\bar{Q}_s\gamma^\mu\tau^I Q_t)$	O_{dd}	$(\bar{d}_p\gamma_\mu d_r)(\bar{d}_s\gamma^\mu d_t)$	O_{Ld}	$(\bar{L}_p\gamma_\mu L_r)(\bar{d}_s\gamma^\mu d_t)$
$O_{LQ}^{(1)}$	$(\bar{L}_p\gamma_\mu L_r)(\bar{Q}_s\gamma^\mu Q_t)$	$O_{\ell u}$	$(\bar{\ell}_p\gamma_\mu \ell_r)(\bar{u}_s\gamma^\mu u_t)$	$O_{Q\ell}$	$(\bar{Q}_p\gamma_\mu Q_r)(\bar{\ell}_s\gamma^\mu \ell_t)$
$O_{LQ}^{(3)}$	$(\bar{L}_p\gamma_\mu\tau^I L_r)(\bar{Q}_s\gamma^\mu\tau^I Q_t)$	$O_{\ell d}$	$(\bar{\ell}_p\gamma_\mu \ell_r)(\bar{d}_s\gamma^\mu d_t)$	$O_{Qu}^{(1)}$	$(\bar{Q}_p\gamma_\mu Q_r)(\bar{u}_s\gamma^\mu u_t)$
		$O_{ud}^{(1)}$	$(\bar{u}_p\gamma_\mu u_r)(\bar{d}_s\gamma^\mu d_t)$	$O_{Qu}^{(8)}$	$(\bar{Q}_p\gamma_\mu T^A Q_r)(\bar{u}_s\gamma^\mu T^A u_t)$
		$O_{ud}^{(8)}$	$(\bar{u}_p\gamma_\mu T^A u_r)(\bar{d}_s\gamma^\mu T^A d_t)$	$O_{Qd}^{(1)}$	$(\bar{Q}_p\gamma_\mu Q_r)(\bar{d}_s\gamma^\mu d_t)$
				$O_{Qd}^{(8)}$	$(\bar{Q}_p\gamma_\mu T^A Q_r)(\bar{d}_s\gamma^\mu T^A d_t)$

Class 8: $(\bar{\psi}_L\psi_R)(\bar{\psi}_R\psi_L)+\text{h.c.}$		Class 8: $(\bar{\psi}_L\psi_R)(\bar{\psi}_L\psi_R)+\text{h.c.}$	
$O_{L\ell Qq}$	$(\bar{L}_p^j\ell_r)(\bar{d}_s Q_{tj})$	$O_{QuQd}^{(1)}$	$(\bar{Q}_p^j u_r)\epsilon_{jk}(\bar{Q}_s^k d_t)$
		$O_{QuQd}^{(8)}$	$(\bar{Q}_p^j T^A u_r)\epsilon_{jk}(\bar{Q}_s^k T^A d_t)$
		$O_{L\ell Qu}^{(1)}$	$(\bar{L}_p^j \ell_r)\epsilon_{jk}(\bar{Q}_s^k u_t)$
		$O_{L\ell Qu}^{(3)}$	$(\bar{L}_p^j \sigma_{\mu\nu}\ell_r)\epsilon_{jk}(\bar{Q}_s^k \sigma^{\mu\nu} u_t)$

In perturbation theory, the measured input parameters have to be properly chosen to determine the couplings at a given scale with the best precision. In the following, parameters carrying a hat superscript are the measured input parameters and the related physical observables derived from it. In this thesis $\hat{m}_Z, \hat{G}_F, \hat{m}_W$ are used as electroweak input parameters. The parameters of the SM Lagrangian are, therefore, given by

$$\hat{e} = 2\sqrt[4]{2}\hat{m}_W\sqrt{\hat{G}_F}\sin\hat{\theta}_W, \quad \hat{g}_1 = \sqrt[4]{2}\hat{m}_Z\sqrt{\hat{G}_F}\sin\hat{\theta}_W, \quad \hat{g}_2 = 2\sqrt[4]{2}\hat{m}_W\sqrt{\hat{G}_F},$$
$$\sin^2\hat{\theta}_W = 1 - \frac{\hat{m}_W^2}{\hat{m}_Z^2}, \qquad \hat{v} = \frac{1}{\sqrt[4]{2}\sqrt{\hat{G}_F}}, \tag{2.33}$$

A detailed discussion of the two different input parameter schemes can be found in [32].

Due to the three generations and Yukawa coupling matrices, the 59 independent dimension-six operators in the Warsaw basis lead to 2499 hermitian operators and real coupling parameters of which 1350 are CP-even and 1149 CP-odd. Experimentally, such a large number of parameters cannot be independently determined. In order

to reduce the number of parameters, flavour symmetry assumptions are used for the dimension-six operators, which are motivated by previous measurements.
One assumption used here is $U(3)^5$ global flavour symmetry, described in detail in [32], which requires all Yukawa couplings except the ones for top and bottom quarks to vanish. The operators of classes 1, 2, 3 and 4 remain unchanged in the $U(3)^5$ symmetry limit and allowing for complex Wilson coefficients and thus for SM and BSM CP-violation. The Yukawa matrices for classes 5, 6 and 8 become diagonal with two non-zero entries. With this assumption, the number of parameters reduces to 52 CP-even and 17 CP-odd ones.

In the $U(3)^5$ flavour symmetric limit, 19 operators contribute to Higgs boson interactions, the operator O_Φ of class 2, the two operators $O_{\Phi\Box}$ and $O_{\Phi D}$ of class 3, the eight operators of class 4, the three operators $O_{\ell\Phi}$, $O_{u\Phi}$ and $O_{d\Phi}$ of class 5 and the five operators $O_{\Phi L}$, $O_{\Phi\ell}$, $O_{\Phi Q}$, $O_{\Phi u}$ and $O_{\Phi d}$ of the class 7 operators. The Ψ^6 operator O_Φ is neglected because it can only be measured in double-Higgs production. In addition, operators which do not directly affect the $H \to ZZ^*$ measurement, i.e. the Higgs boson self-couplings and the Yukawa couplings of the Higgs boson to down-type quarks and leptons, are not taken into account. The remaining couplings are the top Yukawa coupling and the Higgs couplings to the weak gauge bosons (class 4 operators). The former can be measured in the ttH production, while the latter can be measured in the ggF, VBF and VH production as well as in the four lepton decay (HZZ).

Five CP-even and five CP-odd parameter remain, which are listed in Table 2.7. For the Wilson coefficients the SMEFT naming convention [32] is used. The corresponding Feynman diagrams for the CP-even and CP-odd effective couplings are shown in Figure 2.12 and Figure 2.13, respectively.

Table 2.7: SMEFT CP-even and CP-odd dimension-six operators in the Warsaw basis relevant for the measurements in the $H \to ZZ^* \to 4\ell$ channel

CP-even			CP-odd			Impact on	
Operator	Structure	Coeff.	Operator	Structure	Coeff.	production	decay
$O_{\Phi G}$	$\Phi^\dagger\Phi G^A_{\mu\nu}G^{A\mu\nu}$	c_{HG}	$O_{\Phi\widetilde{G}}$	$\Phi^\dagger\Phi\widetilde{G}^A_{\mu\nu}G^{A\mu\nu}$	$c_{H\widetilde{G}}$	ggF	Yes
$O_{u\Phi}$	$(\Phi^\dagger\Phi)(\bar{Q}_p u_r\widetilde{\Phi})$	c_{uH}	$O_{\widetilde{u}\Phi}$	$(\Phi^\dagger\Phi)(\bar{Q}_p u_r\widetilde{\Phi})$	$c_{\tilde{u}H}$	ttH	-
$O_{\Phi W}$	$\Phi^\dagger\Phi W^I_{\mu\nu}W^{I\mu\nu}$	c_{HW}	$O_{\Phi\widetilde{W}}$	$\Phi^\dagger\Phi\widetilde{W}^I_{\mu\nu}W^{I\mu\nu}$	$c_{H\widetilde{W}}$	VBF, VH	Yes
$O_{\Phi B}$	$\Phi^\dagger\Phi B_{\mu\nu}B^{\mu\nu}$	c_{HB}	$O_{\Phi\widetilde{B}}$	$\Phi^\dagger\Phi\widetilde{B}_{\mu\nu}B^{\mu\nu}$	$c_{H\widetilde{B}}$	VBF, VH	Yes
$O_{\Phi WB}$	$\Phi^\dagger\tau^I\Phi W^I_{\mu\nu}B^{\mu\nu}$	c_{HWB}	$O_{\Phi\widetilde{W}B}$	$\Phi^\dagger\tau^I\Phi\widetilde{W}^I_{\mu\nu}B^{\mu\nu}$	$c_{H\widetilde{W}B}$	VBF, VH	Yes

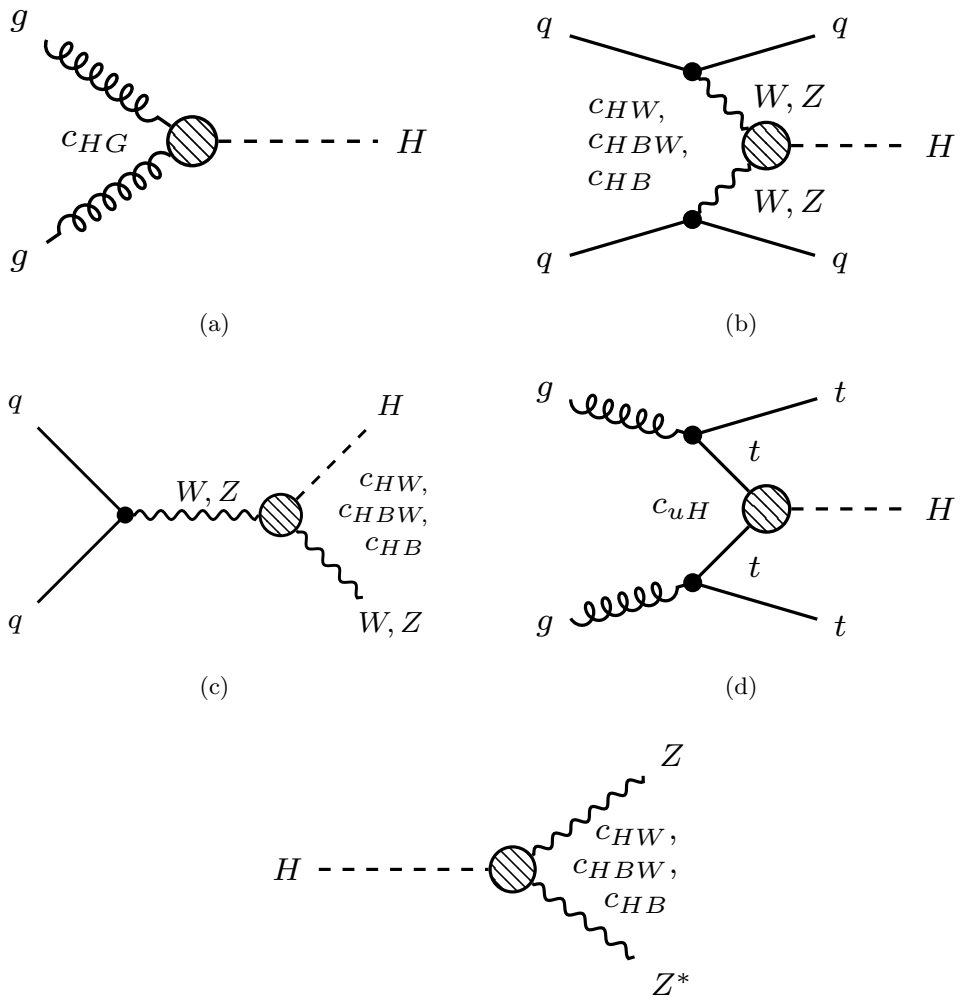


Figure 2.12: Tree-level Feynman diagrams for (a)ggF, (b) VBF, (c) VH and (d) ttH production and (e) for HZZ^* decay with CP-even effective couplings in the SMEFT.

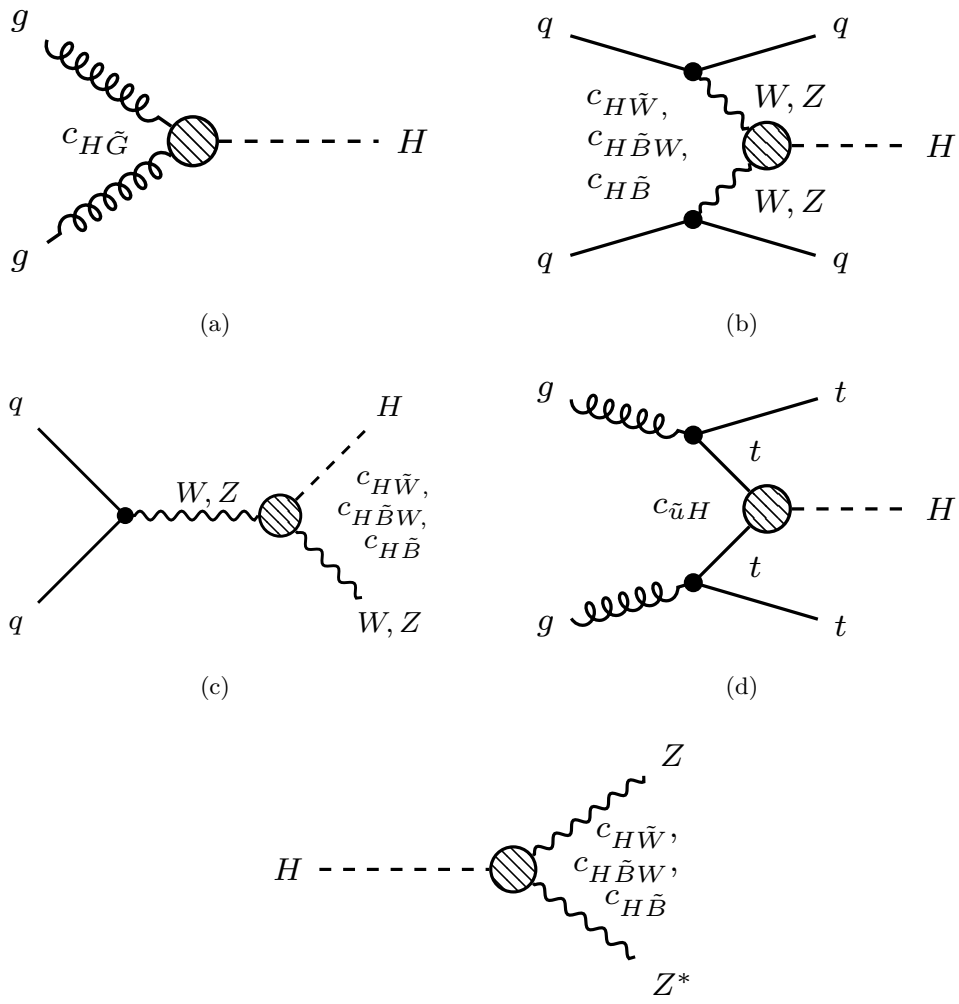


Figure 2.13: Tree-level Feynman diagrams for (a)ggF, (b) VBF, (c) VH and (d) ttH production and (e) for HZZ^* decay with CP-odd effective couplings in the SMEFT.

2.5.2.2 The Higgs Characterisation Model

In the HC model the effective Higgs boson couplings to the SM gauge bosons and fermions from BSM physics are characterised by 20 new parameters. The effective Lagrangian is constructed under the assumptions that the Higgs boson resonance at 125 GeV corresponds to a bosonic state and does not couple to any other new state below the cutoff scale $\Lambda = 1$ TeV. It is assumed, that new physics is described by the lowest dimensional additional operators, i.e. of dimension-six. In principle, all interactions generated by these gauge invariant operators above the electroweak scale can be included, but for practical reasons only operators modifying the three-particle Higgs boson (spin J= 0) interactions are considered. CP conservation is not required. The effective HC interaction Lagrangian

$$\mathcal{L}_{\mathrm{HC}} = \mathcal{L}_{\mathrm{SM-H}} + \mathcal{L}_{\mathrm{J=0}}, \tag{2.34}$$

is formulated in terms of the fermion and gauge boson mass eigenstates, where $\mathcal{L}_{\mathrm{SM-H}}$ is the SM interaction Lagrangian without terms involving the Higgs boson after electroweak symmetry breaking. The modified Higgs boson interactions including dimension-six contributions are described by

$$\mathcal{L}_{\mathrm{J=0}} = -\sum_{t,b,\tau} \bar{\psi}_f \left(c_\alpha \kappa_{Hff} g_{Hff} + s_\alpha \kappa_{Aff} g_{Aff} \gamma_5\right) \psi_f h \tag{2.35}$$

$$\begin{aligned}
&+ \Big\{ c_\alpha \kappa_{SM} \left[\frac{1}{2} g_{HZZ} Z_\mu Z^\mu + g_{HWW} W^+_\mu W^{-\mu}\right] \\
&- \frac{1}{4}\left[c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma} A_{\mu\nu} \widetilde{A}^{\mu\nu}\right] \\
&- \frac{1}{2}\left[c_\alpha \kappa_{HZ\gamma} g_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} g_{AZ\gamma} Z_{\mu\nu} \widetilde{A}^{\mu\nu}\right] \\
&- \frac{1}{4}\left[c_\alpha \kappa_{Hgg} g_{Hgg} G^A_{\mu\nu} G^{A,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg} G^A_{\mu\nu} \widetilde{G}^{A,\mu\nu}\right] \\
&- \frac{1}{4}\frac{1}{\Lambda}\left[c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} Z_{\mu\nu} \widetilde{Z}^{\mu\nu}\right] \\
&- \frac{1}{2}\frac{1}{\Lambda}\left[c_\alpha \kappa_{HWW} W^+_{\mu\nu} W^{-\mu\nu} + s_\alpha \kappa_{AWW} W^+_{\mu\nu} \widetilde{W}^{-\mu\nu}\right] \\
&- \frac{1}{\Lambda} c_\alpha \left[\kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial_\mu Z^{\mu\nu} + \left(\kappa_{H\partial W} W^+_\nu \partial_\mu W^{-\mu\nu} + \mathrm{h.c.}\right)\right] \Big\} h,
\end{aligned} \tag{2.36}$$

with the definition $c_\alpha = \cos\alpha$ and $s_\alpha = \sin\alpha$. The angle α parametrises the mixing between the two CP-eigenstates and implies CP-violation if $\alpha \neq 0, \pi/2$. The first line describes the effective Yukawa interactions with fermions and the remaining lines the effective couplings to the vector bosons. The effective HC coupling modifiers κ_{SM}, κ_{HXX} and κ_{AXX}, respectively, describe the SM and BSM CP-even and CP-odd interactions of the fermions and weak vector bosons with the Higgs field h. For the Lagrangian to be hermitian, the coupling parameters have to be real (except for $\kappa_{H\partial W}$). The SM coupling strengths of the Higgs boson to fermions and bosons are given by g_{Hff} and g_{HXX}, while g_{Aff} and g_{AXX} are the coupling strengths of a pseudo-scalar state A for instance in a

two Higgs-doublet model (2HDM), which is one of the simplest extensions of the SM [43]. A second complex spin-0 field leads to five mass eigenstates after electroweak symmetry breaking, two CP-even states (h, H), one CP-odd state (A) and two charged Higgs bosons ($H^{\pm}$). The SM can is recovered by setting $c_{\alpha} = 1$, $\kappa_{SM} = \kappa_{Hff} = \kappa_{Hgg} = 1$ and the remaining HC coupling parameters κ_i to zero.

Since with the current data no sensitivity is expected to operators describing BSM $\gamma\gamma$ and $Z\gamma$ as well as Yukawa interactions, these are neglected in the presented analysis. In addition it is assumed, that the CP-even and the CP-odd BSM coupling parameters to the Z bosons, κ_{HZZ} and κ_{AZZ}, are identical to the couplings to W bosons, κ_{HWW} and κ_{AWW} and are denoted by $\kappa_{XVV} = \kappa_{XZZ} = \kappa_{XWW}$ in the following.
In the presented studies, the HC coupling parameters considered are the ones to Z and W bosons and to gluons given in Table 2.8. The corresponding Feynman diagrams with effective coupling are shown in Figure 2.14.

Table 2.8: The most relevant Higgs Characterisation (HC) model coupling parameters for the measurements in the $H \to ZZ^* \to 4\ell$ decay channel

Coupling term		Production and decay mode	HC coupling parameter	
$G^A_{\mu\nu}G^{A,\mu\nu}h$		ggF	κ_{Hgg}	SM
$G^A_{\mu\nu}\widetilde{G}^{A,\mu\nu}h$		ggF	κ_{Agg}	BSM CP-odd
$Z_\mu Z^\mu h$,	$W^+_\mu W^{-\mu}h$	VBF, VH, HZZ	κ_{SM}	SM
$Z_{\mu\nu}Z^{\mu\nu}h$,	$W^+_{\mu\nu}W^{-\mu\nu}h$	VBF, VH, HZZ	κ_{HVV}	BSM CP-even
$Z_{\mu\nu}\widetilde{Z}h$,	$W^+_{\mu\nu}\widetilde{W}^{-\mu\nu}h$	VBF, VH, HZZ	κ_{AVV}	BSM CP-odd

2.5.3 Comparability between the Analysis Frameworks

The two EFT analysis frameworks described above use different effective field theories to introduce the BSM contributions to the Higgs boson interaction. The SMEFT framework (Section 2.5.2.1) is formulated before the electroweak symmetry breaking, while the HC framework (Section 2.5.2.2) is formulated in mass eigenstates after the electroweak symmetry breaking. A direct comparison, i.e. a translation between the SMEFT and the HC model, is only possible under certain assumptions. In the SMEFT model BSM contributions to the XVV vertex are described by three coupling parameters (c_{HW},c_{HB},c_{HWB} or $c_{H\widetilde{W}}$,$c_{H\widetilde{B}}$,$c_{H\widetilde{W}B}$), while these correspond to one in the HC model (κ_{HVV} or κ_{AVV}). Therefore, a comparison is only possible if all BSM coupling parameters in the SMEFT model are measured at the same time which was not possible in the presented study due to less amount of data (see Section 7.3.6). In addition, with the assumption in the HC

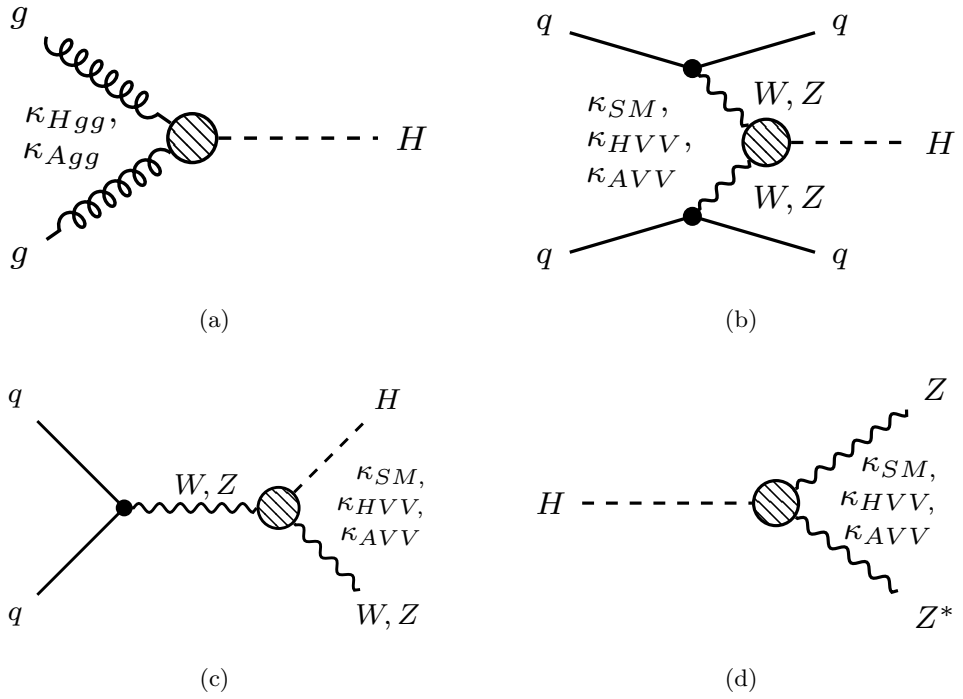


Figure 2.14: Tree-level Feynman diagram for (a) ggF, (b), VBF, (c) VH production mode and (d) for HZZ^* decay with effective couplings in the HC model.

model that the BSM coupling parameters to the Z bosons have the same correlation to corresponding couplings to W bosons as in the SM ($\kappa_{XWW} = \kappa_{XVV}$) a translation from HC to SMEFT is not possible. Thus, the results of the HC analysis and the interpretation within the SMEFT are not compared to each other.

The coupling modification framework (Section 2.5.1) parametrises deviations from the Higgs boson couplings under the assumption of a single CP-even state with SM tensor coupling. Therefore, these coupling strength modifiers can be related to the SM-like coupling parameter of the HC framework, $c_\alpha \kappa_{SM}$.

2.6 Status of Higgs Boson Property Measurements

In 2012, the ATLAS and CMS collaborations announced independently the discovery of a new particle with a mass of 125 GeV and with properties consistent with those of the SM Higgs boson [10, 11]. Further studies of its spin and CP properties, as well as the measurement of its couplings confirmed predictions of the SM [26–28, 44].

In this section the current status of the measurements by the ATLAS and CMS collaborations of the Higgs boson properties are summarised based on data taken in the years

2011 and 2012 at centre-of-mass energies of $\sqrt{s} = 7$ and 8 TeV, respectively, and in 2015 to 2018 at $\sqrt{s} = 13$ TeV.

2.6.1 Measurement of the Higgs Boson Mass

The Higgs boson mass m_H is not predicted by the SM. Once it is measured, the Higgs boson production cross sections as well as the decay rates and the total decay width are fully determined. Other SM predictions of precision electroweak observables in higher-orders perturbation theory depend on m_H. The highest mass resolution can be achieved in final states which can be fully reconstructed with high precision, namely in the $H \to 4\ell$ and $H \to \gamma\gamma$ decay channels. The mass resolution of both experiments in the four-lepton and diphoton final states is about 1-2% [45], while in other channels like $H \to WW^* \to \ell\nu\ell\nu$ only a resolution of about 20% can be obtained due to the missing neutrinos.

Table 2.9 summarises all measurements of the Higgs boson mass, the individual and combined Run 1 measurements of both experiments [46] and the ATLAS and CMS Run 2 measurement in four-lepton and diphoton final states [47, 48].

2.6.2 Measurement of the Higgs Boson Width

The width of the Higgs boson is precisely predicted once the Higgs boson mass is known. For $m_H = 125$ GeV, the width is only 4 MeV. Since the Higgs boson mass resolution of ATLAS and CMS is only about 1-2 GeV, direct measurements of the width of the resonance are much larger than the predicted natural width. In addition to the direct measurements, also indirect measurements have been performed from both experiments [45]. So far no deviations from the SM prediction have been observed.

Direct constraints on the width are determined by measuring the shape of the Higgs boson resonance in final states with high mass resolution, the four-lepton and diphoton final states [47, 50]. Another direct constraint has been derived by CMS from the lifetime of the Higgs boson obtained from the measurement of the displacement of the four-lepton vertex in the $H \to ZZ^* \to 4\ell$ decays from the proton-proton interaction point [51]. The limits on the natural width of the Higgs boson from direct measurements are summarised in Table 2.10.

Indirect constraints on the width can be obtained from the resonance shift in the diphoton mass spectrum. The interference between $gg \to H \to \gamma\gamma$ signal and the irreducible $gg \to \gamma\gamma$ background leads to a visible resonance mass shift, which depends on the total width of the Higgs boson [52–54]. The mass shift has been measured by ATLAS to be 35 ± 9 MeV [55]. At the high-luminosity LHC (HL-LHC) an upper limit on the total width

Table 2.9: Higgs boson mass measurements performed in the four-lepton and diphoton final states by the ATLAS and CMS experiments with Run 1 and Run 2 data [46–49]

Experiment	Channel	Measured mass	Errors		
			Tot.	Stat.	Syst.
		Run 1			
ATLAS	$H \to \gamma\gamma$	126.02	±0.51	(±0.43	±0.27) GeV
CMS	$H \to \gamma\gamma$	124.70	±0.34	(±0.31	±0.15) GeV
ATLAS	$H \to 4\ell$	124.51	±0.52	(±0.52	±0.04) GeV
CMS	$H \to 4\ell$	125.59	±0.45	(±0.45	±0.17) GeV
ATLAS+CMS	$H \to \gamma\gamma$	125.07	±0.29	(±0.25	±0.14) GeV
ATLAS+CMS	$H \to 4\ell$	125.15	±0.40	(±0.37	±0.15) GeV
ATLAS+CMS	$H \to \gamma\gamma + 4\ell$	125.09	±0.24	(±0.21	±0.11) GeV
		Run 2			
ATLAS	$H \to \gamma\gamma$	125.32	±0.35	(±0.19	±0.29) GeV
ATLAS	$H \to 4\ell$	124.71	±0.30	(±0.30	±0.05) GeV
ATLAS	$H \to \gamma\gamma + 4\ell$	124.86	±0.27	(±0.18	±0.20) GeV
CMS	$H \to \gamma\gamma$	125.78	±0.26	(±0.18	±0.19) GeV
CMS	$H \to 4\ell$	125.26	±0.21	(±0.20	±0.08) GeV
CMS	$H \to \gamma\gamma + 4\ell$	125.46	±0.17	(±0.13	±0.11) GeV

Table 2.10: Observed (and expected) 95% confidence level (CL) limits on the natural width of the Higgs boson with mass of $m_H = 125$ GeV from fits of the four-lepton and diphoton mass spectra and from the four-lepton vertex displacement

Experiment	$\gamma\gamma$ mass spectrum	4ℓ mass spectrum	4ℓ vertex lifetime
ATLAS	$< 5.0(6.2)$ GeV	$< 2.6(6.2)$ GeV	–
CMS	$< 2.4(3.1)$ GeV	$< 1.1(1.6)$ GeV	$> 3.5 \cdot 10^{-12}$ GeV

of approximately 200 MeV at 95% confidence level (CL) is expected for an integrated luminosity of 3 ab^{-1} and 14 TeV centre-of-mass energy.

In addition, the interference signal and background in the diphoton channel causes a change of the off-resonance Higgs boson production cross section. The size of this change also depends on the total width of the Higgs boson, leading to an upper bound of 800 MeV on the Higgs boson total width [56]. The limit can be improved to 60 MeV with 3 ab^{-1} at 14 TeV at the HL-LHC [57].

Finally, the Higgs boson width can be determined indirectly from the on- and off-resonance signal strengths in the VV decay channels [58–61]. CMS has set the most accurate constraints on the Higgs boson width from measurements the on- and off-resonance Higgs boson production in the four-lepton final state with the data collected in 2011 to 2017. The Higgs boson width is measured to be $\Gamma_H \Gamma_H^{SM} = 3.2^{+2.8}_{-2.2}$ MeV, while $4.1^{+5.0}_{-4.0}$ MeV is expected from simulation [62]. The upper limit obtained with 36.1 fb^{-1}of Run 2 data by ATLAS is weaker, corresponding to $\Gamma_H / \Gamma_H^{SM} < 3.5$, while 3.7 is expected from the SM [63]. Prospective studies by ATLAS and CMS using only the four-lepton final states show that the width can be measured with this method with a precision of $\Gamma_H = 4.1^{+0.7}_{-0.8}$ MeV with 3 ab^{-1} of 14 TeV data [64].

2.6.3 Measurements of the Higgs Boson Couplings

A crucial test of the SM is the measurement of the Higgs boson couplings to SM particles. At the LHC, the Higgs boson couplings are determined from the signal strength μ, the product of production cross sections σ_i for a given process i and decay branching ratios into a final state f, $\sigma_i \cdot \mathcal{B}_f$, relative to the SM prediction [65, 66]. The latest combination of measurements by ATLAS and CMS of the production cross-section times the branching ratio, are shown in Figure 2.15 and Figure 2.16, respectively. The results are all compatible with the SM prediction $\mu = 1$. In ATLAS, only the combination VH of WH and ZH production has been measured. Due to the small sensitivity for the

$H \to bb$ decay in ggF production, the SM prediction is used for the combination. The same has been done for the $H \to WW^*$ and $H \to \tau\tau$ decays in VH production and the for the $H \to WW^*$ and $H \to ZZ^*$ decays (VV) in $ttH + tH$ production. CMS on the other hand fixed the signal strength for $H \to \tau\tau$ decays in WH and ZH production, $H \to bb$ decays in VBF production and $H \to \mu\mu$ in WH, ZH and ttH production to the SM predictions. For $H \to ZZ^*$ decays in WH, ZH, and ttH production and $H \to \gamma\gamma$ decays in ZH production, negative signal strength results due to negative event yield after background subtraction have been set to zero.

The signal strength μ_i^f is related to the couplings of the Higgs boson to the fermions or vector bosons in the initial and final states via the multiplicative coupling modifiers of the κ-framework introduced in Section 2.5.1 and defined in Equation 2.30. The Higgs boson couplings are probed in addition by measuring the so-called reduced coupling strength scale factor defined as

$$y_F = \kappa_F \frac{g_F}{\sqrt{2}} = \kappa_F \frac{m_F}{v} \quad \text{and} \quad y_V = \sqrt{\kappa_V \frac{g_V}{2v}} = \sqrt{\kappa_V} \frac{m_v}{v}. \tag{2.37}$$

where m_F (m_V) corresponds to the mass of the fermions (weak gauge boson) and g_V the absolute values of Higgs boson coupling strength. Figure 2.17(a) and 2.17(b) illustrate this linear dependence over a wide range of particle masses in the SM with the latest measurement by ATLAS and CMS, respectively, in excellent agreement with the SM prediction [65, 66].

2.6.4 Measurements of the Higgs Boson Self-Coupling

With the measurement of the Higgs boson self coupling the Higgs potential and the mechanism of the electroweak symmetry breaking can be probed. This important and very challenging measurement is an outstanding long term goal at the LHC, at the HL-LHC and at future colliders.

ATLAS and CMS studied the sensitivity to the trilinear self-coupling g_{h^3} (see Figure 2.3) in the full data set of 3 ab^{-1} of the HL-LHC [67–69] in the measurements of the Higgs boson pair production in the $HH \to bb\gamma\gamma$, $HH \to bb\tau\tau$ and $HH \to bbWW$ decays. The most sensitive channel is $HH \to bb\gamma\gamma$, where signal significances of 1.05 σ and 1.6 σ are expected by ATLAS [67] and CMS [68], respectively.

2.6.5 Measurements of the Higgs Boson Spin and Parity

In the SM, the Higgs boson is a CP-even scalar particle with $J^P = 0^{++}$. Its spin and parity properties have been measured by ATLAS and CMS [27, 28] using decays $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to \ell\nu\ell\nu$. The SM $J^P = 0^+$ hypothesis has been compared

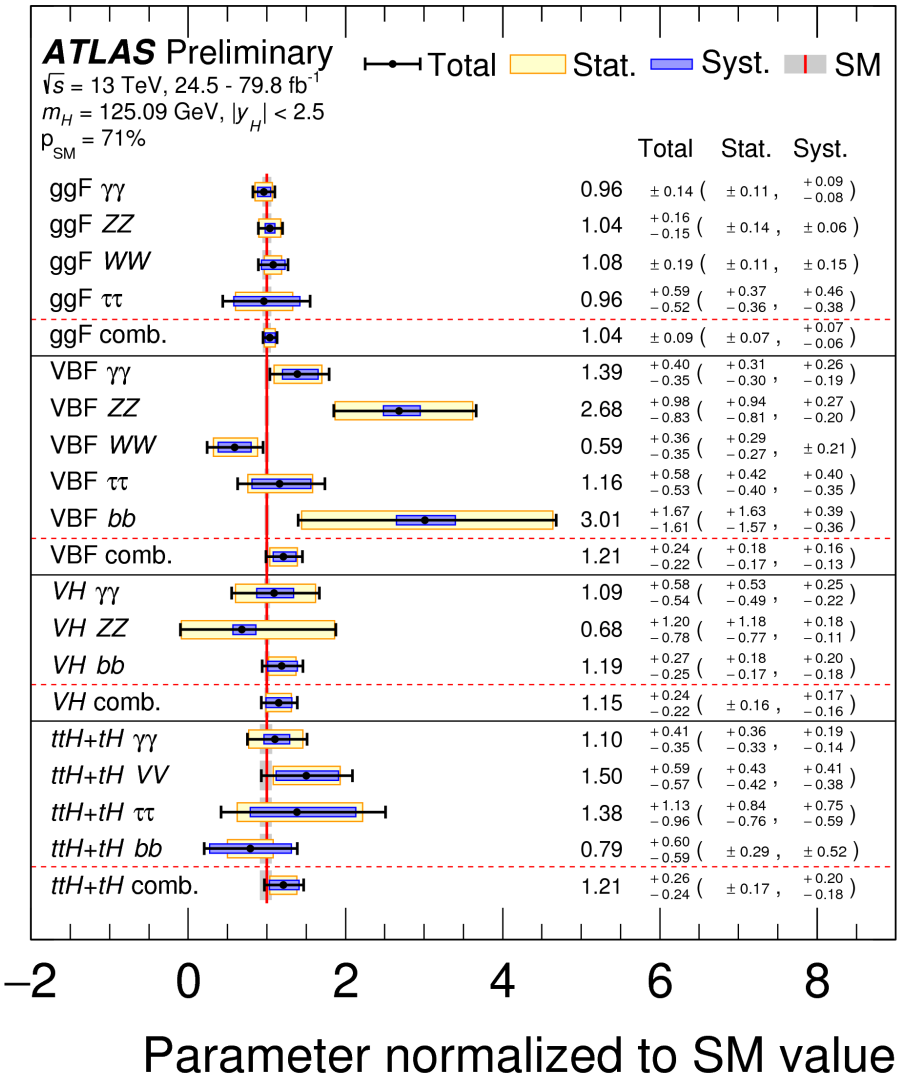


Figure 2.15: ATLAS combined measurements of the signal strength μ_i^f for the different Higgs boson productions (i) and decay modes (f), the product of the production cross section and the branching ratios, $\sigma_i \cdot \mathcal{B}_f$, normalised to the SM prediction. The results are based on the Run 2 data set with an integrated luminosity of 79.8 fb^{-1} at a centre-of-mass energy of 13 TeV [65].

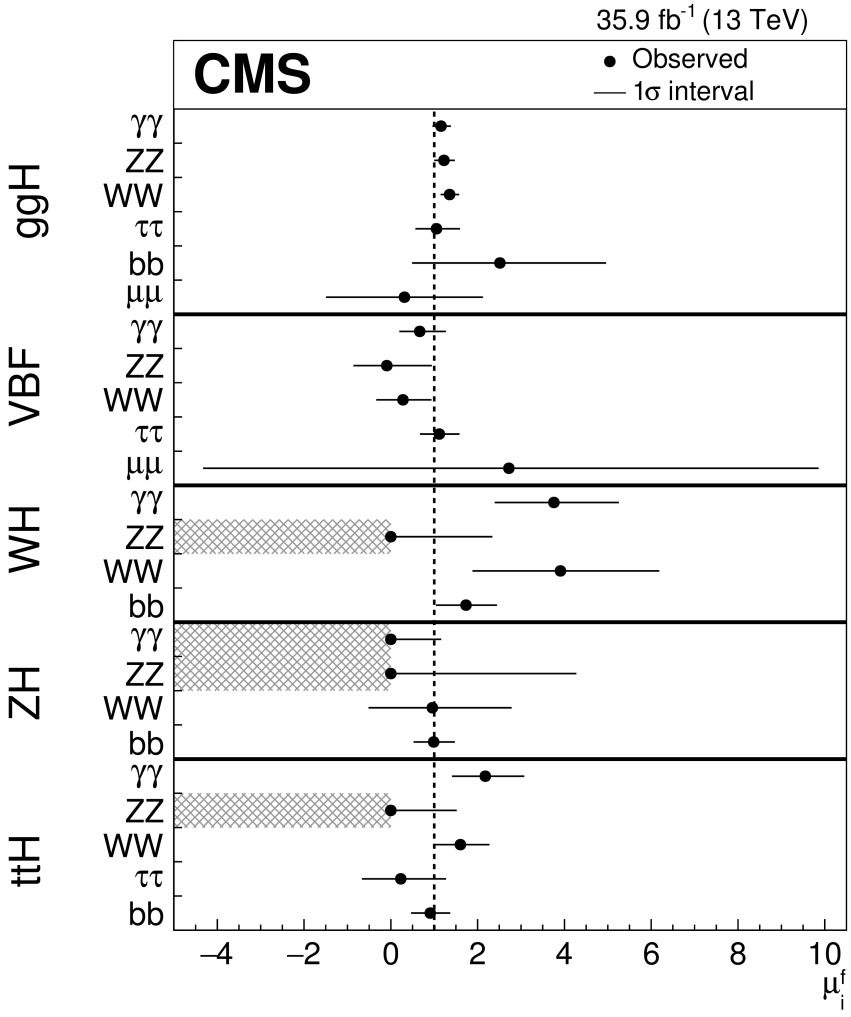
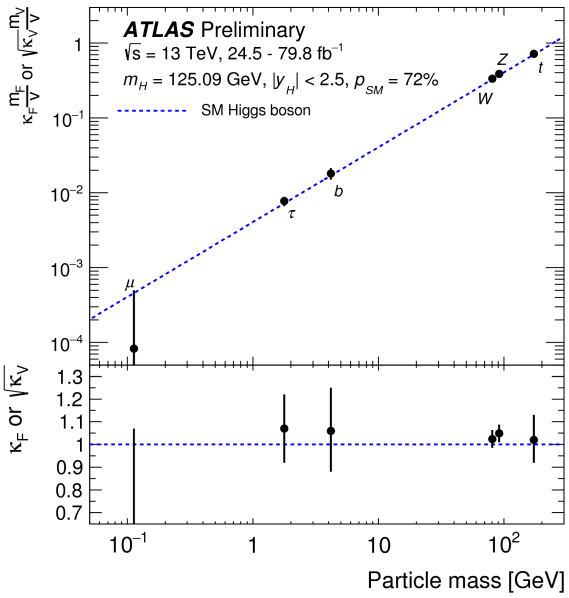
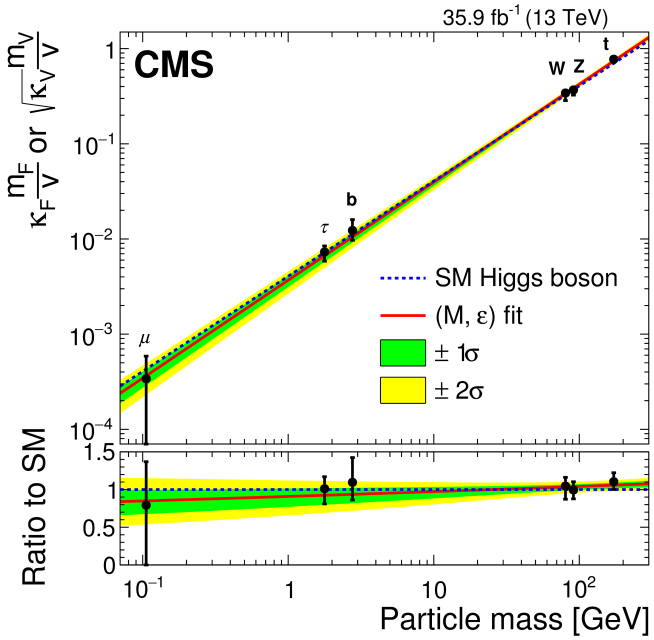


Figure 2.16: CMS combined measurements of the signal strength μ_i^f for the different Higgs boson productions (i) and decay modes (f), the product of the production cross section and the branching ratios, $\sigma_i \cdot \mathcal{B}_f$, normalised to the SM prediction. The results are based on the Run 2 data set with an integrated luminosity of 35.9 fb^{-1} at a centre-of-mass energy of 13 TeV [66].



(a)



(b)

Figure 2.17: Measurements of the modified coupling parameters $\kappa_F \frac{m_F}{v}$ and $\sqrt{\kappa_V}\frac{m_v}{v}$ to the SM fermions F and weak gauge bosons V by (a) ATLAS and (b) and CMS as a function of the particle mass. The SM prediction is shown as dashed line. The lower panels show the ratios of the measured values to the SM predictions, i.e. κ_F and $\sqrt{\kappa_V}$. The results are based on Run 2data with an integrated luminosities of 79.8 fb^{-1}(ATLAS) and 35.9 fb^{-1}(CMS) at a centre-of-mass energy of 13 TeV [65, 66].

to several alternative hypotheses, namely $J^P = 0^-, 1^+, 1^-$ and 2^+. The spin-1 hypothesis is in principle excluded due to the observation of $H \to \gamma\gamma$ decays and the Landau-Yang theorem, which forbids the decay of a spin-1 particle into two massless vector bosons [70, 71]. However, the theorem applies only in the narrow-width approximation and the two reconstructed photons could actually be a pair of two collinear photons which cannot be separated experimentally. Hence, the spin-1 hypotheses has also been probed.

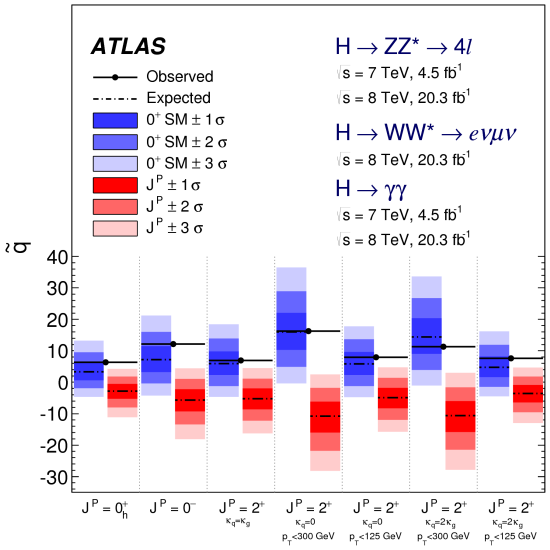
Figure 2.18(a) and Figure 2.18(b) show the expected and observed distributions of the test statistics of the SM hypothesis against all alternative spin-parity hypotheses for ATLAS and CMS, respectively. In all cases the quantum numbers predicted by the SM, $J^P = 0^+$, are favoured by the data. Higher-order scalar or pseudo-scalar contributions have been excluded. However, it is still possible to have small CP-even or CP-odd admixtures to the predominately CP-even SM Higgs boson state, as predicted by many theories beyond the SM. These admixtures can be tested by measuring the tensor structure of the Higgs boson couplings. This measurement has also been performed in $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to \ell\nu\ell\nu$ decays [28]. The ratio of the HC coupling parameters (Section 2.5.2.2),

$$\frac{\widetilde{\kappa}_{HVV}}{\kappa_{SM}} = \frac{v}{4 \cdot \Lambda} \cdot \frac{\kappa_{HVV}}{\kappa_{SM}} \quad \text{and} \quad \frac{\widetilde{\kappa}_{AVV}}{\kappa_{SM}} \cdot \tan\alpha = \frac{v}{4 \cdot \Lambda} \cdot \frac{\kappa_{AVV}}{\kappa_{SM}} \cdot \tan\alpha, \tag{2.38}$$

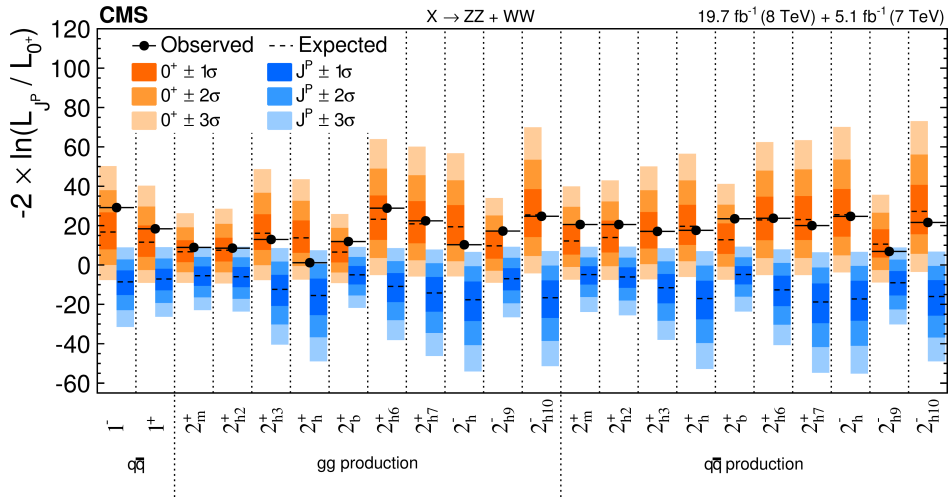
has been measured. The measurement was based on the shape of the CP-sensitive kinematic distributions of the Higgs boson decay products. The expected and observed values of coupling parameters excluded at a 95% CL after combination of both decay channels are given in Table 2.11. For the CP-odd coupling parameter the best fit value is $(\tilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan\alpha = -0.68$ and the region outside $-2.18 < (\tilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan\alpha < 0.83$ is excluded at 95% CL. For the CP-even coupling parameter the best fit value is $\tilde{\kappa}_{HVV}/\kappa_{SM} = -0.48$ and the region outside $-0.73 < \tilde{\kappa}_{HVV}/\kappa_{SM} < 0.63$ is excluded at 95% CL. All results are consistent with the SM prediction.

Table 2.11: Expected and observed best-fit values and exclusion regions at 95% confidence level (CL) for the higher-order CP-odd and CP-even BSM coupling parameters, $(\tilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan(\alpha)$ and $\tilde{\kappa}_{HVV}/\kappa_{SM}$, respectively. Results are obtained from the combined analysis of the $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to \ell\nu\ell\nu$ decay channels [28]

Coupling ratio	Best-fit value	95% CL Exclusion regions	
combined	Observed	Expected	Observed
$(\tilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan\alpha$	-0.68	$(-\infty, -2.33] \bigcup [2.30, \infty)$	$(-\infty, -2.18] \bigcup [0.83, \infty)$
$\tilde{\kappa}_{HVV}/\kappa_{SM}$	-0.48	$(-\infty, -0.55] \bigcup [4.80, \infty)$	$(-\infty, -0.73] \bigcup [0.63, \infty)$



(a)



(b)

Figure 2.18: Distributions of the test statistics of the SM hypothesis $J^P = 0^+$ against the alternative spin-parity hypotheses $J^P = 0^-, 1^+, 1^-, 2^+$). For (a) ATLAS [28]. SM hypothesis is shown in blue, alternative hypotheses in red and (b) CMS [27] the SM hypothesis in orange, alternative hypotheses in blue. The results are based on the Run 1 data at centre-of-mass energies of 7 TeV and 8 TeV.

3

The ATLAS Detector at the LHC

In this chapter the ATLAS experiment at the Large Hadron Collider (LHC) is described. A short introduction of the CERN accelerator complex is given, followed by the characterisation of proton-proton collisions at the LHC. Subsequently, the ATLAS detector with its sub-detectors is described. Finally, the particle reconstruction and identification with the ATLAS detector is outlined.

3.1 The Large Hadron Collider

The Large Hadron Collider [72, 73] is a circular particle accelerator and collider situated at the European Centre for Particle Physics (CERN) in Switzerland. The LHC tunnel with 26.5 km circumference is located 45 m to 170 m below ground. Two particle beams circulate in the tunnel in opposite directions in two separate vacuum pipes. Superconducting dipole-magnets of 8.33 T field strength bend the two beams on the accelerator ring. The LHC is designed to collide protons up to a centre-of-mass energy of $\sqrt{s} = 14$ TeV and heavy ions, lead nuclei, up to $\sqrt{s} = 5.02$ TeV.

Before the protons or heavy ions are injected into the LHC, they traverse a chain of pre-accelerators shown in Figure 3.1. First the protons and lead ions are accelerated in a linear accelerator (LINAC2) up to energies of 50 MeV and 4.2 MeV, respectively. Afterwards, the protons are injected into the Proton Synchrotron Booster (PSB), which increases their energy up to 1.4 GeV, while the lead ions are injected into the Low Energy Ion Ring (LEIR) and accelerated to 72 MeV. Protons and lead ions are then accelerated in the Proton Synchrotron (PS) up to energies of 25 GeV and 6 GeV, respectively. The last pre-acceleration is performed in the Super Proton Synchrotron (SPS) which accelerates protons up to 450 GeV and heavy ions up to 117 GeV. After that they are piped to the LHC, where they are further accelerated to the collision energy.

Up to now two successful periods of data taking, referred to as Run 1 and Run 2, took place at the LHC. During the Run 1 period in the years 2010 to 2012, proton-proton collision data were taken at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The Run 2 data taking period took place in the years 2015 to 2018 with proton-proton

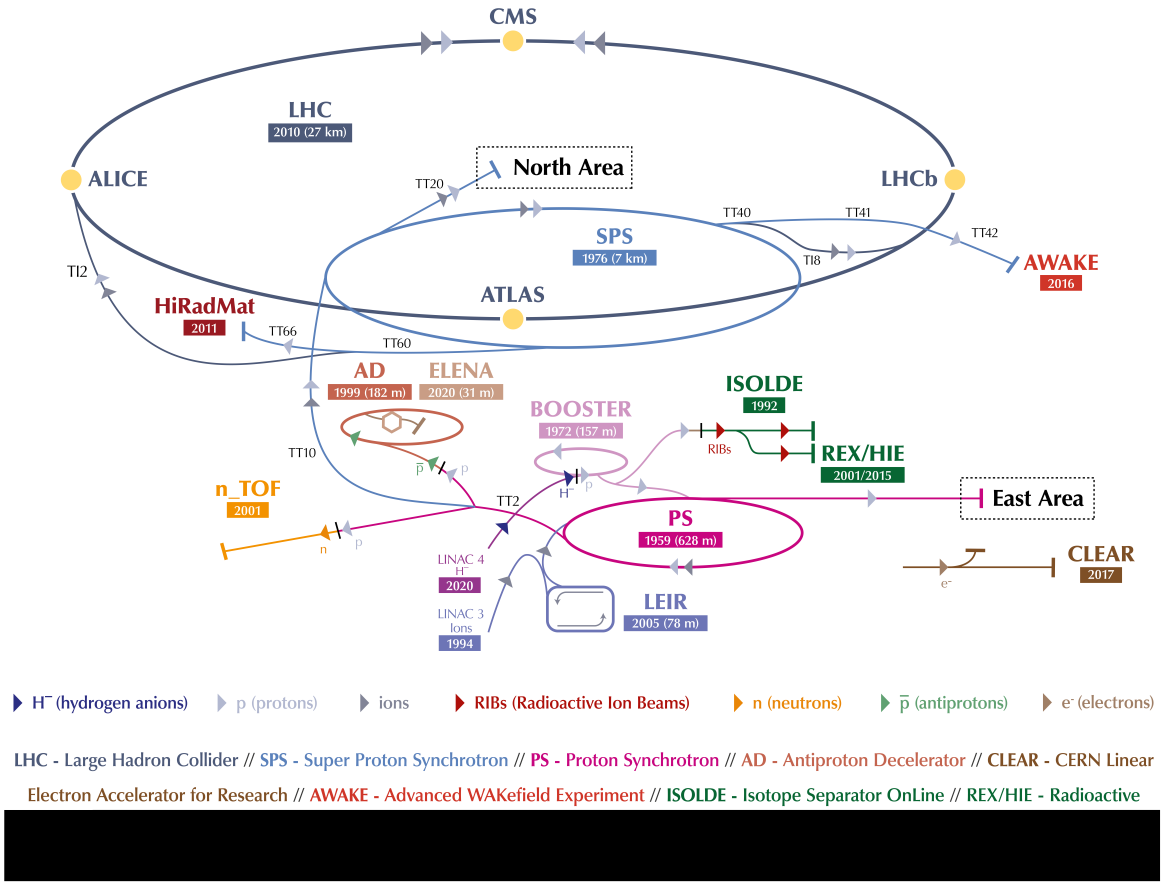


Figure 3.1: The CERN accelerator complex [74]. The illustration shows the injection chain traversed by protons colliding in the LHC.

collisions at a centre-of-mass energy of 13 TeV. The design collision energies of 14 TeV for protons and 5.6 TeV for heavy ions are expected to be reached during the Run 3 data taking which will start after the second long LHC shutdown in 2021.

The four main experiments at the LHC are ATLAS [75], CMS [76], ALICE [77] and LHCb [78]. The two general purpose experiments, ATLAS and CMS, are designed to search for the Higgs boson and a broad range of new physics phenomena, such as dark matter, supersymmetry or extra dimensions. In addition, the detector allows for high-precision measurements of SM processes and parameters. The heavy ion detector ALICE studies the physics of strongly interacting matter at extreme energy densities, the so-called quark-gluon plasma. In contrast to the other detectors, the LHCb experiment is designed to detect particles mainly in one forward direction. In order to investigate the difference between matter and antimatter in the decays products of B mesons. The studies presented in this thesis are based on the proton-proton collision data recorded with the ATLAS detector in Run 2.

3.2 Proton-Proton Collisions at the Large Hadron Collider

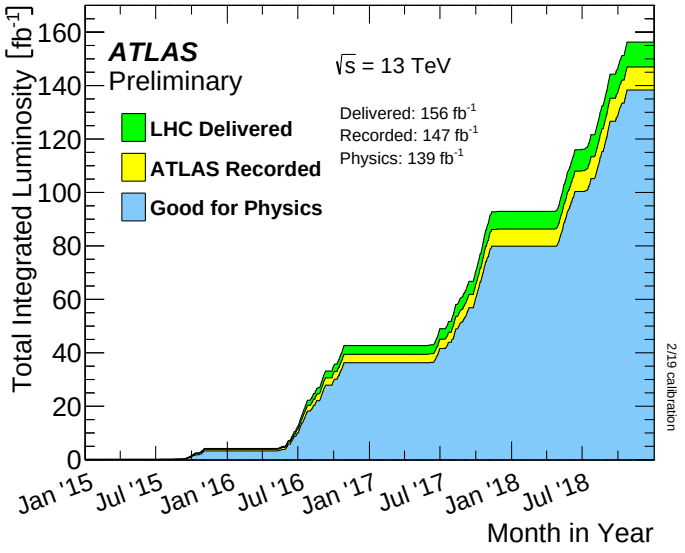
For any process occurring in the collisions of particles, the number of collision events N over a given time period is given by

$$N = L \cdot \sigma = \int L_{\text{inst}} \cdot \sigma \, \mathrm{d}t, \tag{3.1}$$

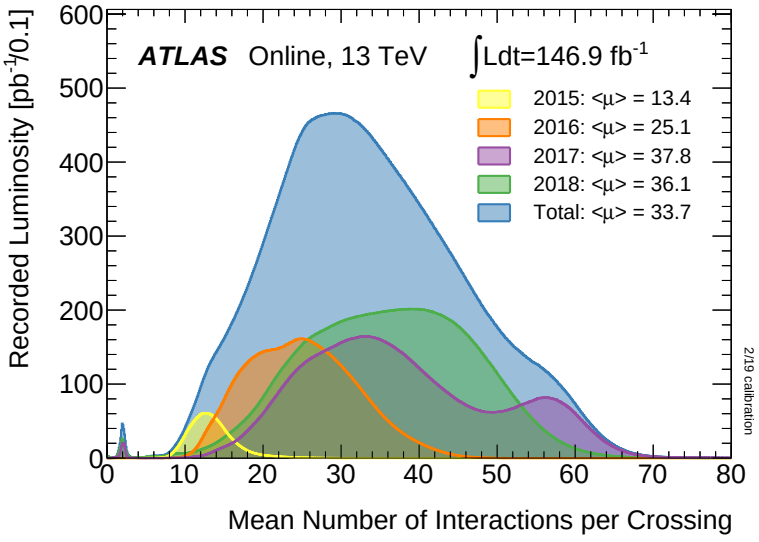
where σ is the cross section of the collision process, and L and L_{inst} are the integrated and instantaneous luminosities, respectively. The cross section depends on the centre-of-mass energy. L_{inst} is determined by the number of particle bunches per beam n_b, the number of particles N_B per bunch $B = 1, 2$, the bunch circulation frequency f and the widths Σ_x and Σ_y of the horizontal and vertical beam profiles at the collision point [79]:

$$L_{\text{inst}} = \frac{f n_b N_1 N_2}{\Sigma_x \Sigma_y}. \tag{3.2}$$

The LHC is designed to deliver proton-proton collisions at an instantaneous luminosity of $L_{\text{inst}} = 10^{34}$ cm^{-2}s^{-1}, with $N_B = 1.15 \cdot 10^{11}$ protons per bunch in $n_b = 2808$ bunches per beam and 25 ns bunch spacing [72]. These values have been reached in 2016, except for the number of bunches per beam. Figure 3.2(a) shows the integrated luminosity L recorded by ATLAS in Run 2. In 2018, the largest instantaneous of $L_{\text{inst}} = 21.0 \cdot 10^{33}$ cm^{-2}s^{-1} was reached. It was $L_{\text{inst}} = 20.9 \cdot 10^{33}$ cm^{-2}s^{-1} in 2017, $L_{\text{inst}} = 13.8 \cdot 10^{33}$ cm^{-2}s^{-1} in 2016 and $L_{\text{inst}} = 5.0 \cdot 10^{33}$ cm^{-2}s^{-1} in 2015. In this thesis, the analyses are based on several Run 2 proton-proton collision data sets described in more detail in Section 4.2.1, corresponding to integrated luminosities of 36.1 fb^{-1} (2015-2016), 79.8 fb^{-1} (2015-2017) and 139 fb^{-1} (2015-2018).



(a)



(b)

Figure 3.2: The total integrated luminosity delivered (green), recorded by the ATLAS detector (yellow) and good for physics (blue) in the years 2015 to 2018 is shown in (a), while the luminosity weighted distribution of the mean number of proton-proton interactions per bunch crossing (pile-up) for the full Run 2 data set (blue histogram) is shown in (b). Individual distributions for the data taking years are also shown.

Protons are no elementary particles, but are bound states consisting of strongly interacting partons, the valence quarks (uud), sea quarks ($\bar{q}q$) and gluons which carry fractions x of the total momentum of the proton.

The scattering processes in proton-proton collisions involve soft and hard parton interactions. Since hard interactions, allowing for the production of heavy resonances, are characterised by high momentum transfer, Q^2, they can be described by perturbation theory. Soft interactions of the participating partons show low momentum transfer and lead to additional quarks and gluons in the final-state (the so-called underlying event). They cannot cannot be described by perturbation theory. The quarks and gluons in the final-state event and further gluons and quarks in the parton showers eventually fragment into hadrons. An illustration of a hadronic interaction, including both hard and soft parton scattering interactions is shown in Figure 3.3.

The production cross section in proton-proton collisions can be described by [81]

$$\sigma_{\text{tot}} = \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \sum_{ij} f_i(x_1, Q^2) f_j(x_2, Q^2) \hat{\sigma}(x_1 x_2 s). \tag{3.3}$$

where $\hat{\sigma}(x_1 x_2 s)$ is the perturbatively calculable parton hard scattering cross section at energy fraction $\sqrt{x_1 x_2 s}$ of the LHC centre-of-mass energy $\sqrt{s}$ and $f_i(x_1, Q^2)$ and $f_j(x_2, Q^2)$ are the parton distribution functions (PDFs) of the colliding partons i and j. The PDFs are the probability density functions for finding a parton $i(j)$ with momentum fraction $x_1(x_2)$ in the colliding protons 1 and 2 at a given momentum transfer Q^2. The PDFs cannot be calculated perturbatively. They have been measured in dedicated experiments e.g. deep-inelastic electron/neutrino-proton scattering, and then extrapolated to the LHC energies using the DGLAP equation [82, 83].

Several PDF sets have been determined, like CT [84], MMHT [85] and NNPDF [86, 87]. For the LHC measurements, a combination of the available PDF sets is used, called PDF4LHC [88]. Figure 3.4 shows the proton PDF for momentum transfers of $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2. Gluons carry the main fraction of the total momentum, followed by the sea and valence quarks. The PDFs of the gluons and the sea quarks are large at small x values and become small above $x = 0.1$, where the valence quarks dominate. At large Q^2, the contribution of the sea quarks becomes larger and also bottom quarks carry a non-negligible momentum fraction. The most relevant momentum fraction of the colliding partons is between $x = 10^{-3}$ and $x = 10^{-1}$. Therefore, at LHC energies, the production processes are often dominated by the gluon fusion.

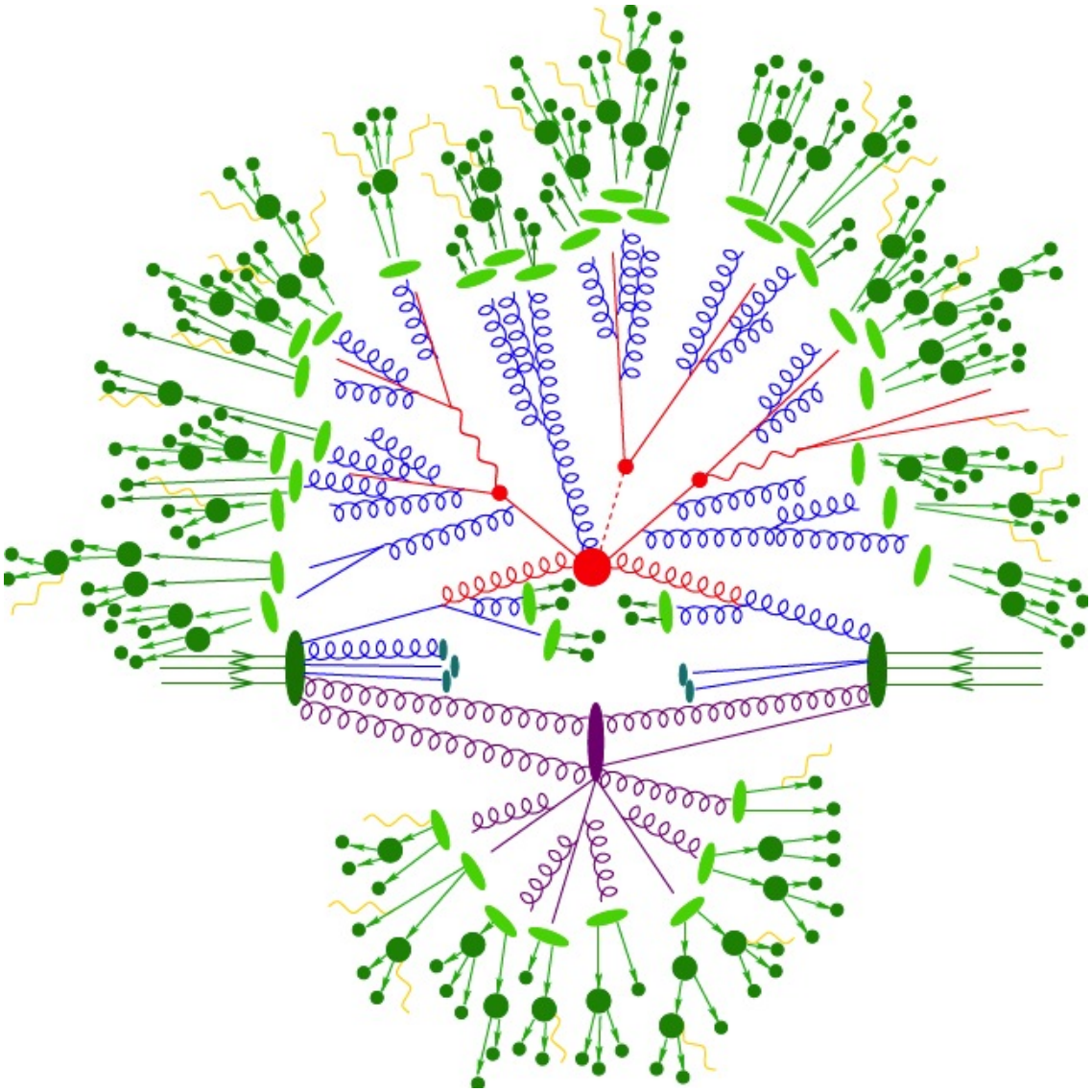
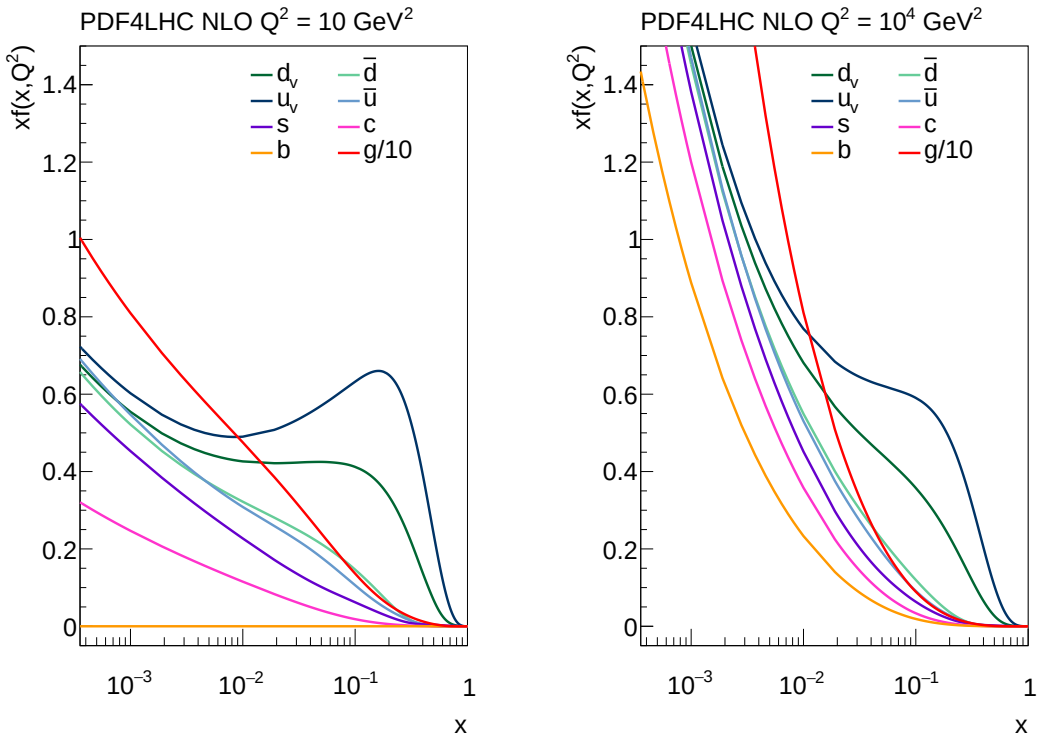


Figure 3.3: Illustration of a proton-proton collision at high energies [80]. The hard scattering of partons is shown in red, Bremsstrahlung processes occurring in parton showers in blue, the hadronising partons in green and secondary soft interactions in purple. Hadron decays are shown in dark green and QED Bremsstrahlung in yellow.



(a) (b)

Figure 3.4: The PDF4LHC PDF sets at next-to-leading-order (NLO) QCD at centre-of-mass-energies (a) $Q^2 = 10$ GeV^2 and (b) $Q^2 10^4$ GeV^2.

In addition to the collision process with the hard interaction multiple proton-proton collisions per bunch crossing can occur, the so-called pile-up events. The number of pile-up events depends on the run conditions and increases with the instantaneous luminosity as shown Figure 3.2(b). The performance of the detector is affected by two types of pile-up the in-time pile-up of additional proton-proton interactions in the same bunch crossing as the hard scattering process, and the out-of-time pile-up due to the particle interactions from neighbouring bunch crossings recorded during the detector read out.
The hard-scattering processes of signal and background events are simulated by Monte Carlo event generators. Parton shower programs, which are interfaced to the event generators, simulate the hadronisation of the final-state. The parton shower and the fragmentation simulation are tuned to the measurements. The generated events are then passed through the ATLAS detector simulation [89] which is based on GEANT4 program [90]. The hits produced by the detector simulation are converted into electronic in the digitalisation step. Afterwards, the simulates data are reconstructed in the same way as the recorded data. Multiple proton-proton collisions (pile-up) are taken into account by su-

perimposing minimum bias events onto the simulated signal events. In order to match the mean number of interactions per bunch crossing to the data, the simulated pile-up profile is reweighted.

3.3 The ATLAS Detector

The ATLAS experiment [75] is one of the two multi-purpose experiments at the LHC. Its physics programme is based on proton-proton collisions covers precision measurements of SM processes and the search for physics beyond the SM. In addition, the ATLAS experiment explores the strongly interacting matter at extreme energy densities in heavy ion collisions. The first main goal of the ATLAS experiment, the discovery of the Higgs boson, was reached in 2012.

The LHC beam conditions, i.e. the high luminosity and the centre-of-mass energy and the aim to cover a broad spectrum of physics lead to the design requirements for the ATLAS detector: fast, highly-granular and radiation hard tracking detectors providing excellent momentum and vertex resolution for charged particles, highly-granular calorimeter covering almost the entire solid angle, a muon spectrometer with high momentum resolution up to highest muon energies and an selective and efficient trigger system.

In this section, the coordinate system and conventions of the ATLAS experiment and the different detector components are described.

3.3.1 Coordinate System and Conventions of the ATLAS Experiment

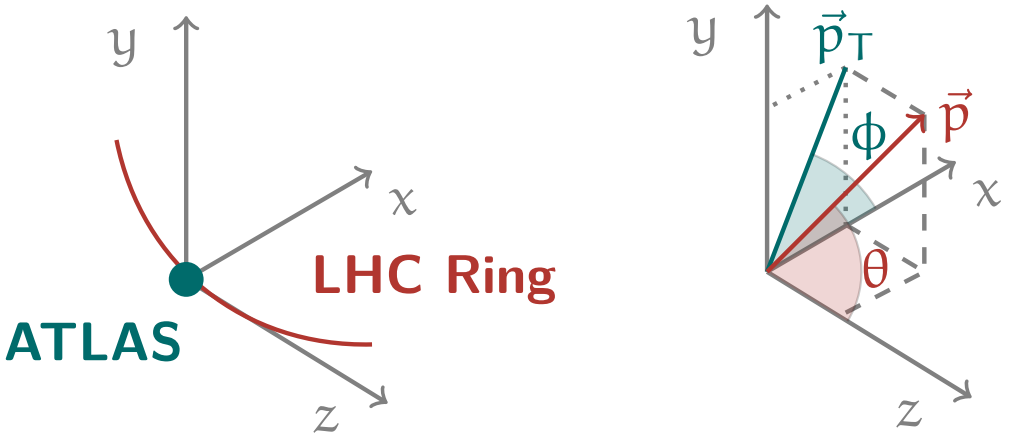
The origin of the right-handed coordinate system of the ATLAS detector (see Figure 3.5(a)) is set to the nominal interaction point. The x-axis points to the centre of the LHC ring, the y-axis upwards and the z-axis along the beam direction. In cylindrical coordinates (see Figure 3.5(b)), the azimuthal angle $\phi \in [-\pi, \pi]$ is measured in the x-y plane (transverse plane) relative to the x-axis. The polar angle $\theta \in [0, \pi]$ is measured from the positive z-axis.

The pseudorapidity η of a particle, illustrated in Figure 3.6, is a measure of the particles direction invariant boosts in z-direction:

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \tag{3.4}$$

Large pseudorapidities correspond to small polar angles, while values around $\eta = 0$ are close to the transverse plane. The rapidity

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.5}$$



(a) (b)

Figure 3.5: Illustration of the ATLAS coordinate system in (a) Cartesian and (b) cylindrical coordinates.

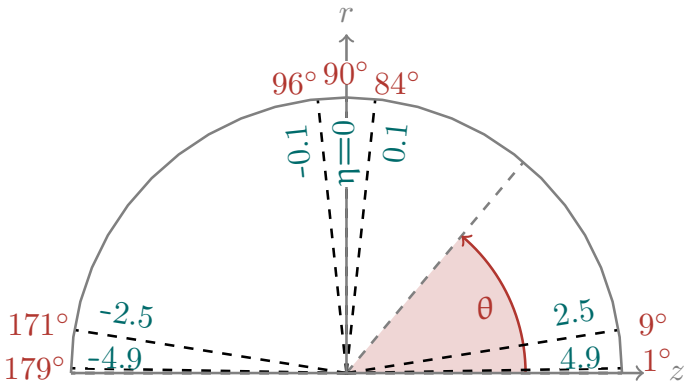


Figure 3.6: Illustration of the correspondence between pseudorapidity η (in green) and polar angles θ (in red).

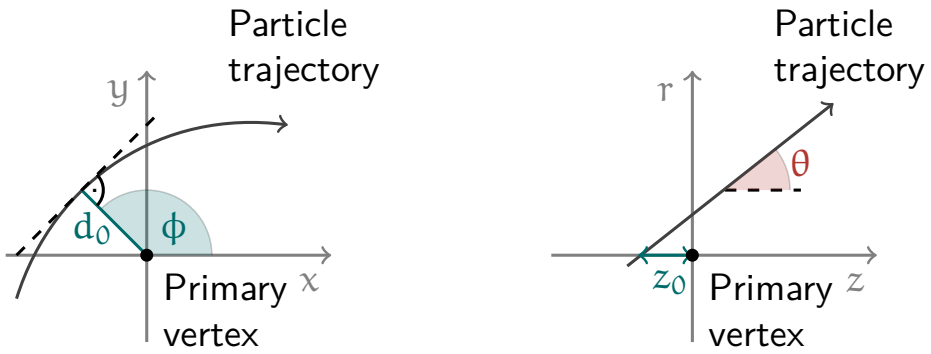
with p_z the momentum component along the z-axis, is alternatively used for particles whose rest mass is comparable to their energy E. For massless particles rapidity and pseudorapidity are identical.

The angular separation ΔR of two particles is given accordingly by

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} = \sqrt{\Delta\eta^2 + \Delta\phi^2}, \tag{3.6}$$

where $\eta_{1(2)}$ and $\phi_{1(2)}$ are the pseudorapidities and azimuthal angles of the two particles.

The distance of closest approach between the trajectory of a particle and the reconstructed primary vertex in the transverse plane is the transverse impact parameter d_0 illustrated in Figure 3.7(a). The longitudinal impact parameter z_0, illustrated in Figure 3.7(b), is defined as the distance of closest approach to the primary vertex in longitudinal z-direction.



(a) (b)

Figure 3.7: Illustration of (a) the transverse impact parameter d_0 and (b) the longitudinal impact parameter z_0.

The transverse momentum $\boldsymbol{p}_T$ of a particle, indicated in Figure 3.5(b), is the momentum component orthogonal to the beam axis

$$\boldsymbol{p}_T = \boldsymbol{p} \cdot \sin\theta. \tag{3.7}$$

As the colliding protons are composite particles, the momenta of the incoming partons are unknown. In addition, fragmentation products of partons which do not participate in

the hard scatter process leave the detector undetected in the direction of the beam pipe. Hence, the total longitudinal momentum of collision products is unknown. On the other hand, the colliding partons have negligible transverse momenta, such that the transverse momentum sum of all collision products vanishes,

$$\sum_i \boldsymbol{p}_{T,i} = 0, \tag{3.8}$$

due to momentum conservation. This constraint helps for the reconstruction of neutrinos or new weakly interacting particles, which are not recorded in the detector. Their transverse energies are detected as missing transverse energy defined as transverse momentum imbalance:

$$E_T^{miss} = -\left|\sum_i \boldsymbol{p}_{T,i}\right|. \tag{3.9}$$

3.3.2 The Detector Components

A schematic view of the ATLAS detector is shown in Figure 3.8. The detector has a length of about 44 m along the beam axis and a diameter of about 25 m. It consists of three regions, the central one with cylindrical shape (the barrel region) and two disk-shaped regions at both ends of the barrel (the endcap regions).

In each of these regions, several detector sub-system technologies are placed. Closest to the beam axis is the inner detector measuring the momenta and vertices of charged particle tracks in a 2 T solenoidal magnetic field provided by a thin superconducting coil. It is surrounded by the electromagnetic and hadronic calorimeters used for the identification of photons, electrons and hadron jets and for the measurement of their energies. The outermost part of the detector is the muon spectrometer used for the identification and the momentum measurement of muons. The magnetic field in the muon spectrometer is provided by three large superconducting air-core toroidal magnets with a field integral of 2-6 Tm.

In the following, the detector sub-system are described in detail.

3.3.2.1 The Trigger System

The detector is operated at a high occupancy with a bunch crossing frequency of 40 MHz. With the available bandwidth and computing resources it is not possible to record every collision event. Instead, a dedicated trigger system is used to select the interesting hard-interaction events. In Run 2, a two-level trigger system is used, consisting of a hardware-based level-1 trigger and a software-based high-level trigger [91, 92].

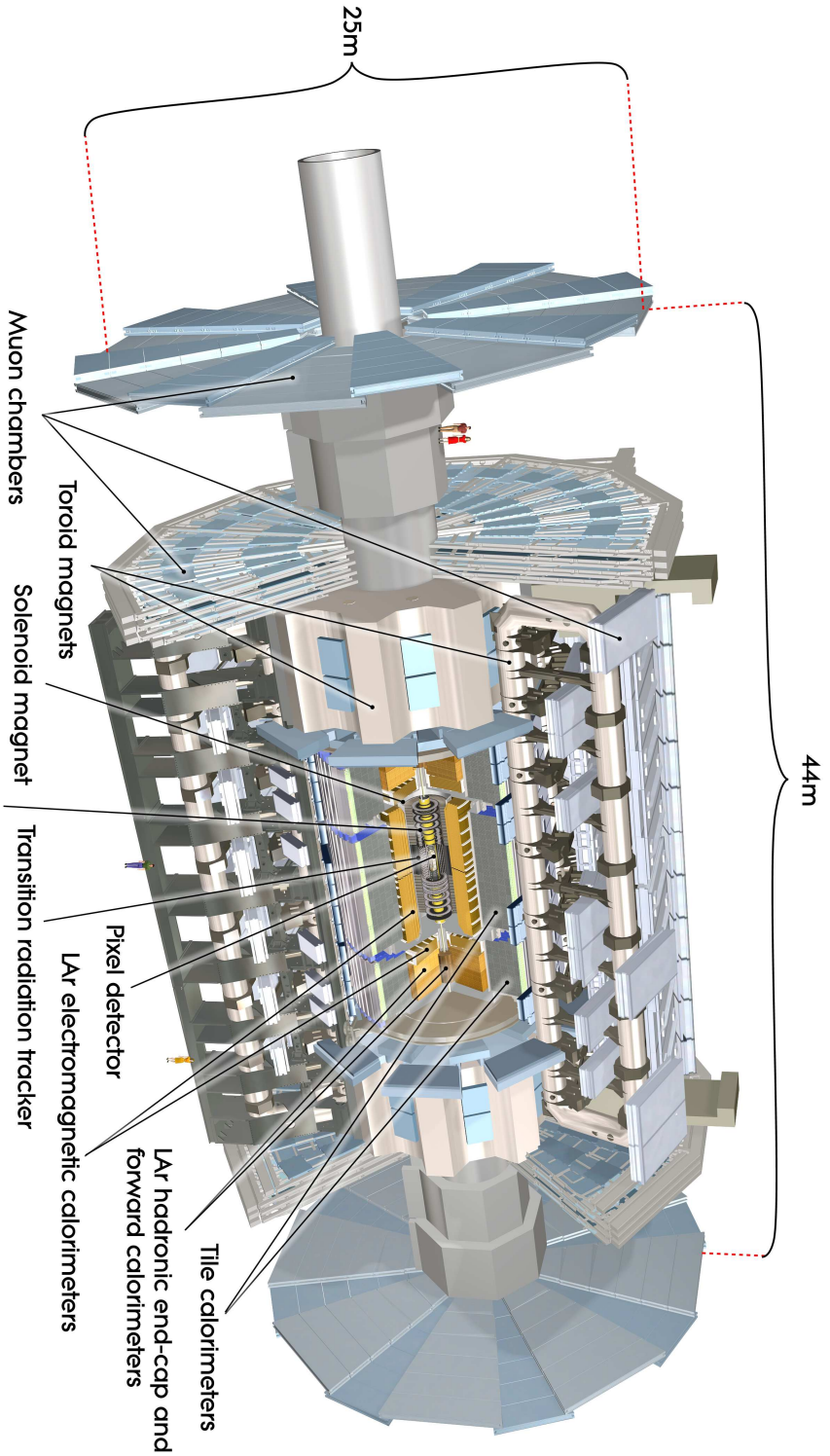


Figure 3.8: Cut-away view of the ATLAS detector [75] with its detector sub-systems.

The initial event selection is performed by the level-1 trigger within 2 µs after the bunch crossing, reducing the event rate from 40 MHz to 100 kHz. Until the trigger decision, the detector signals are stored in pipelines of the detector front-end electronics. In addition to accepting or rejecting an event, the level-1 trigger defines for the accepted events the regions of interests in the detector, which are analysed in more detail at the next trigger level.

The software-based high-level trigger processes the information within the regions of interest using fast algorithms, based on the offline reconstruction software. The trigger-level reconstruction consists of a fast tracking and a precision tracking part. The fast tracking is seeded by the information of the level-1 trigger and decides within 40 ms whether the event should be further processed. Afterwards, a detailed tracking is performed seeded by the fast tracking results. The decision of the high-level trigger is made within 4 s and reduces the output event rate down to 1 kHz. The triggered events are recorded and fully reconstructed by the ATLAS offline reconstruction [90].

3.3.2.2 The Inner Detector

The inner detector shown in Figure 3.9 is immersed in a 2 T solenoidal magnetic field which is oriented parallel to the z-axis and generated by a cylindrical superconducting coil. It measures the trajectories of charged particles and provides a precise determination of their momenta. In order to identify the primary collision vertex as well as the secondary decay vertices of particles, the reconstructed tracks are extrapolated to a common point of origin. The secondary vertex reconstruction is important for the identification of particles with relatively long lifetime, such as b-hadrons and τ-leptons.

The inner tracking detector consists of three sub-detectors: the silicon pixel detector, the semiconductor central tracker and the transition radiation tracker [93, 94]. The pixel detector is arranged in three concentric cylindrical layers starting at 110 mm distance from the beam axis in the barrel region and three discs in each of the two e endcap regions. During the first long shutdown of the LHC between Run 1 and Run 2, an additional inner pixel layer (the insertable b-layer) was installed at a distance of 33 mm to the beam axis to improve the vertex resolution and the b-tagging performance [95].

The second high-resolution tracking detector, the semiconductor central tracker, is arranged in four concentric layers of silicon micro-strip sensors in the barrel and nine discs in each of the two in the endcap regions. In order to measure the second coordinate along the z-axis in the barrel and radial in the endcaps, the strips in the sensor pairs glued back to back to each other at a small stereo angle are placed in each barrel layer parallel to the beam axis. The outermost tracking component, the transition radiation tracker, is composed of straw drift tubes with a diameter of 4 mm, which are filled with a $Xe/CO_2/O_2$ (70%/27%/3%) gas mixture.

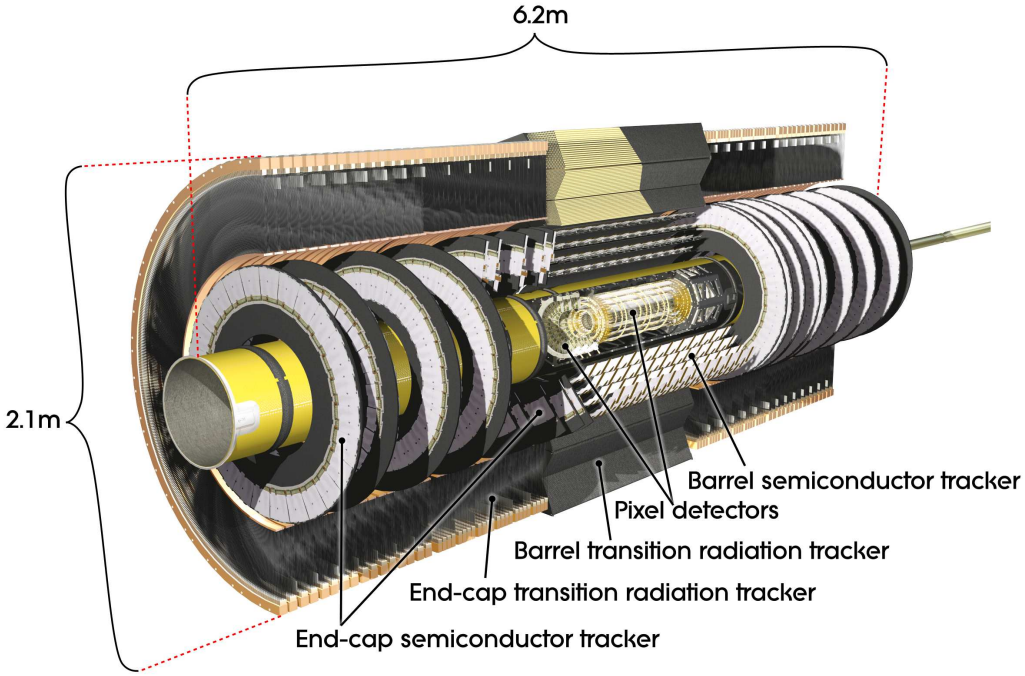


Figure 3.9: Cut-away view of the ATLAS inner detector including the pixel detector, the semiconductor tracker and the transition radiation tracker [75].

While the insertable b-layer covers a pseudorapidity range of $|\eta| < 2.9$, the semiconductor central tracker and the transition radiation tracker provide precise track measurement up to $|\eta| = 2.5$ and $|\eta| = 2.0$, respectively. The transverse momenta of charged particles are measured by the inner tracker with a resolution of $\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$ within $|\eta| < 2.5$.

3.3.2.3 The Calorimeter System

The energy deposits of electrons, photons, hadrons and jets are measured by the calorimeter system of the ATLAS detector shown in Figure 3.10. The particles are absorbed in high-density absorber plates and the deposited energy in the form of secondary scattered particles is measured in active material in between. The calorimeter system is divided into the electromagnetic and the hadronic calorimeter. The former is employed to measure the energies and directions of electromagnetically interacting particles, electrons and photons. The strongly interacting hadrons are measured in the latter. The entire calorimeter system covers the solid angle up to $|\eta| = 4.9$ [96, 97], allowing for a precise measurement of the missing transverse energy.

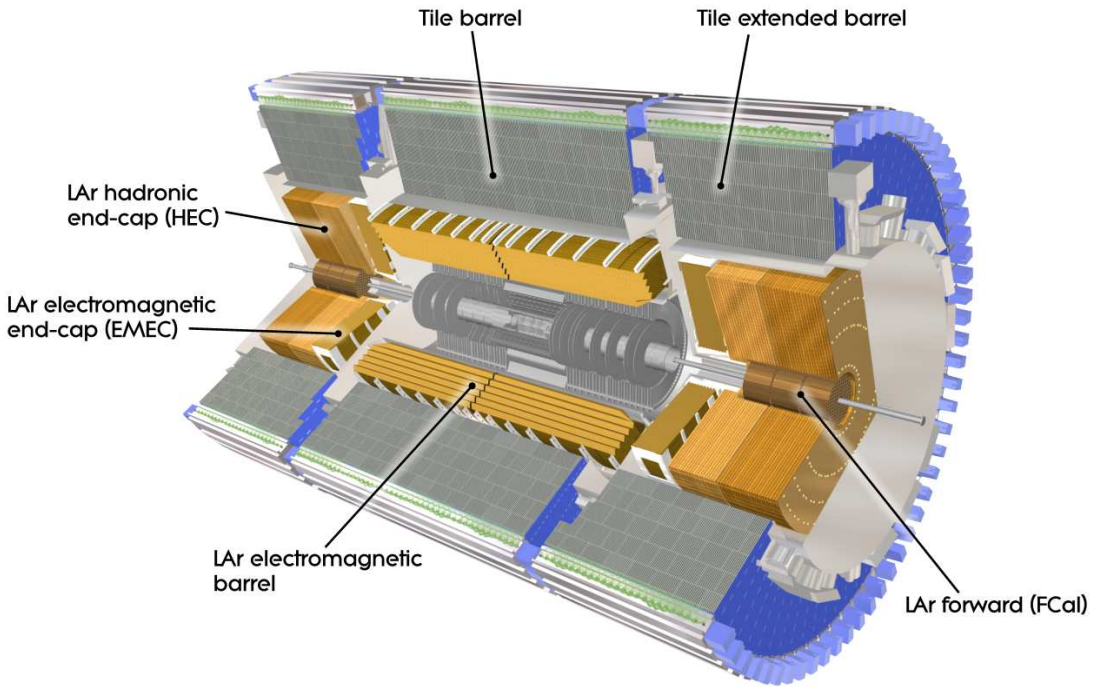


Figure 3.10: Cut-away view of the ATLAS calorimeter system with the electromagnetic and the hadronic calorimeters [75].

The electromagnetic calorimeter surrounding the inner detector covers a pseudorapidity range of $|\eta| < 3.2$. It consists of three parts, the barrel region within $|\eta| < 1.475$ and two endcap regions covering $1.375 < |\eta| < 3.2$. The active material of these highly-granular calorimeters is liquid argon (LAr) linking the gaps between accordion-shaped lead absorber plates. The energy resolution of the electromagnetic calorimeter is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$. To trigger on electrons and photons at level-1, a reduced granularity information from the electromagnetic calorimeters is used.

The hadronic calorimeters surround the electromagnetic calorimeter. In the central region covering $|\eta| < 1.7$, steel is used as absorber material and scintillating tiles as active medium. The endcap calorimeters cover the region $1.5 < |\eta| < 3.2$ using copper absorbers and liquid argon as active material. In order to increase the pseudorapidity coverage, hadronic forward calorimeters are installed in the region $3.1 < |\eta| < 4.9$ which use also liquid argon as active material but with different absorber materials. The first layer of the forward calorimeter consists of a copper absorber matrix with LAr filled bores optimised for electromagnetic calorimetry. The second and third layers use a tungsten absorber matrix measure hadronic energy deposits. The energy resolution of the hadronic calorimeter is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ in the barrel region and $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ in the endcap calorimeters.

3.3.2.4 The Muon Spectrometer

Muons are minimum ionising particles and the only charged particles which traverse the calorimeter system with only small energy loss. They are detected with high efficiency in the outermost and largest part of the ATLAS detector, the muon spectrometer (see Figure 3.11). The spectrometer consists of three parts, one barrel region within $|\eta| < 1.05$ and two endcap regions within $1.05 < |\eta| < 2.7$. The muon momenta are measured with high precision in the magnetic field of three large superconducting toroidal magnets, one in the barrel region ($|\eta| < 1.7$) and two in the endcap regions, with field strength ranging from 0.5 to 1 T [98].

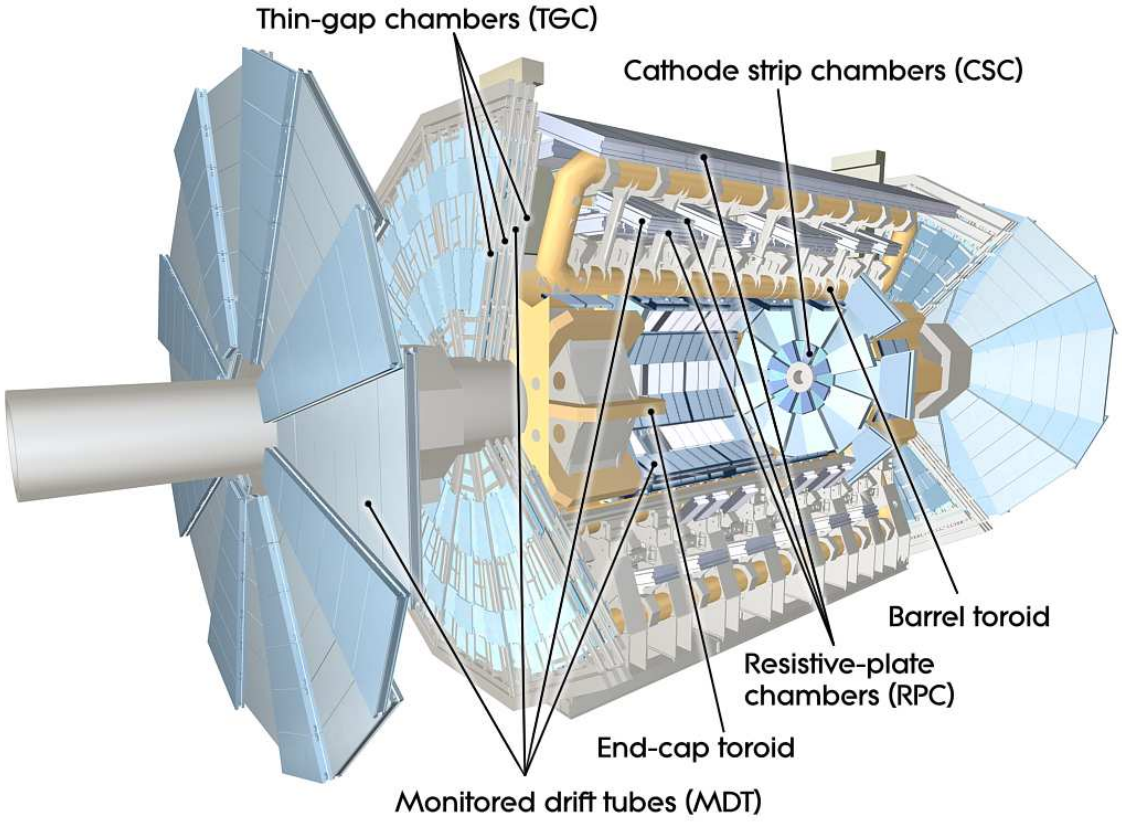


Figure 3.11: Cut-away view of the ATLAS muon spectrometer with the trigger chambers (RPC, TGC) and the precision tracking chambers (CSC, MDT) [75].

The precision muon tracking is in general performed in three layers of Monitored Drift Tube (MDT) chambers. In order to cope with the high background rates from neutrons and γ-rays emerging from the interactions in the calorimeters in shielding and filling the detector cavern, the innermost endcap layer uses Cathode Strip Chambers (CSC) covering $|\eta| > 2.0$. In addition, small-diameter Muon Drift Tube (sMDT) chambers have been installed inside the detector feet in the barrel region during the winter shutdown 2016/2017 [99]. The relative muon momentum resolution of the muon spectrometer ranges from about 3% at $p_T \approx 100$ GeV to about 10% at $p_T \approx 1$ TeV.

There are two types of fast trigger chambers: Resistive Plate Chambers (RPC) in the barrel region within $|\eta| < 1.05$ and Thin Gap Chambers (TGC) in the endcap regions within $1.05 < |\eta| < 2.7$. The level-1 muon trigger covers the pseudorapidity $|\eta| < 2.4$ with additional second-coordinate position measurement with reduced resolution along the drift tubes.

3.4 Particle Reconstruction and Identification

The different particles produced in proton-proton collisions interact in specific ways in one or more of the ATLAS detector sub-systems as illustrated in Figure 3.12, allowing for their reconstruction and identification.

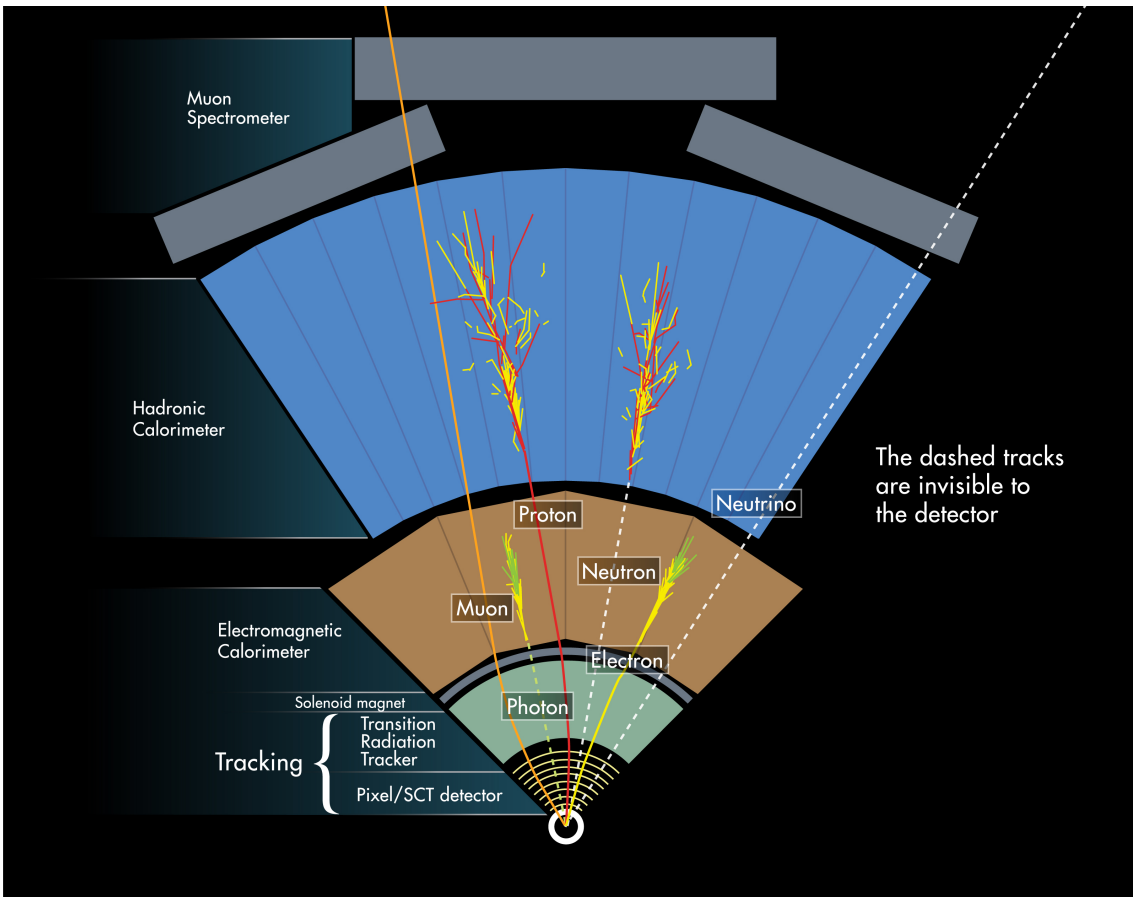


Figure 3.12: Illustration of the signatures of the interacting of different particle types in the ATLAS detector produced in proton-proton collisions [100].

Charged particles, such as electrons, muons or charged hadrons, are first detected in the inner detector. The direction of the curvature of the reconstructed particle trajectory allows for distinguishing between positively and negatively charged particles. All electromagnetically interacting particles such, as electrons and photons, deposit most of their energy in the electromagnetic calorimeter. The electrons can therefore be identified as electromagnetic showers matched to tracks in the inner detector, while photons as neutral particles, leave energy deposits in the electromagnetic calorimeter but no associated track in the inner detector.

Strongly interacting particles traverse the electromagnetic calorimeter with relatively little energy deposit and are stopped in the hadronic calorimeter producing hadronic showers. Charged hadrons (such as protons) are identified by matching hadronic showers to corresponding tracks in the inner detector and to energy deposits in the electromagnetic calorimeter. Neutral hadrons (such as neutrons), only produce hadronic showers. The hadrons produced in the hadronisation of highly energetic quarks and gluons and their decay products form hadron jets.

Muons as minimum ionising particles leave only small amounts of their energy in the calorimeter system and are subsequently identified in the muon detector. The muon trajectories reconstructed in the inner detector are combined with the muon tracks in the muon spectrometer to improve the muon momentum resolution and reconstruction efficiency and to measure the production vertex. The associated small energy deposition in the calorimeters are used to extend the muon identification coverage where there are gaps in the muon spectrometer. Weekly interacting particles (such as neutrinos), pass the detector undetected and can only be identified via missing transverse energy [101].

The subject of this thesis is the study of the Higgs boson properties in the $H \to ZZ^* \to 4\ell$ decay channel. The final-state particles are muons and electrons, as well as quark and gluon jets from associated Higgs boson production. The reconstruction and identification of these particles is described in more detail in the following.

3.4.1 Electron Reconstruction and Identification

Electrons are reconstructed from tracks in the inner detector and energy deposits in the electromagnetic calorimeter [102–104]. As first step, the energy deposits in the electromagnetic calorimeter are reconstructed using a dedicated clustering algorithm [105]. For the studies performed with the two larger data sets with 79.8 fb^{-1} and 139 fb^{-1}, the cluster algorithm has been improved to include low-energy Bremsstrahlung photons radiated by electrons in the inner detector [106]. The reconstructed clusters have to fulfil requirements on the shower shape, before they are matched to inner detector tracks. Radiative energy losses in the inner detector are taken into account by track refitting using a Gaussian-Sum Filter [107].

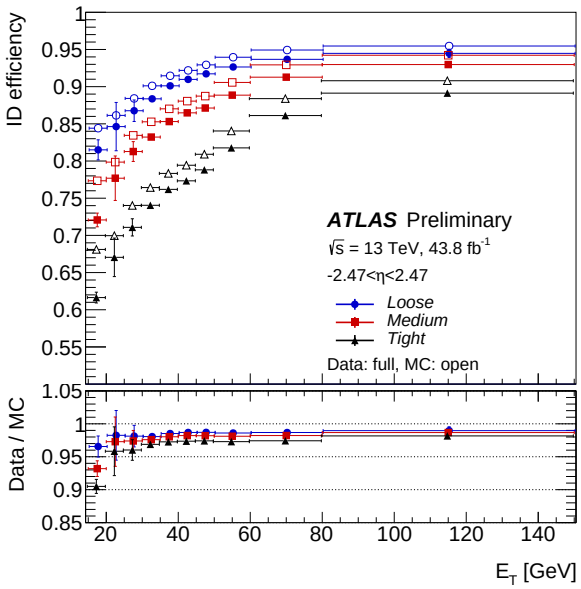
Prompt electrons are distinguished from background electrons due to misidentification of hadron jets, photon conversions or Dalitz decays of pions using a likelihood-based multivariate discriminant [102–104]. Discriminating variables are e.g. a hit in the insertable b-layer, the longitudinal and transverse electromagnetic showers shape, track-cluster matching accuracy and signals from the transition radiation detector. There are three sets of electron identification criteria with increasing background rejection: the so-called *loose, medium* and *tight* working points. For the $H \to ZZ^* \to 4\ell$ analysis, the *loose* working point with an electron efficiency of about 93% is chosen. The average electron efficiencies for the *medium* and *tight* working points are about 88% and 80%, respectively.

The electron identification efficiency of each of the three working points is measured using the tag-and-probe method in $Z \to ee$ and $J/\psi \to ee$ events [102–104]. The latter process allows for the measurement of electron transverse energies E_T in the energy range from 4.5 GeV to 20 GeV, while the former covers the energy range above 15 GeV. Electron candidates pairs are required to have an invariant mass consistent with the Z boson mass, 76.2 GeV $< m_{ee} <$ 106.2 GeV or with the J/ψ mass, 1.8 GeV $< m_{ee} <$ 4.6 GeV, respectively. The electrons are reconstructed within $|\eta| < 2.47$ and are required to have $E_T > 24$ GeV ($E_T > 4.5$ GeV) for Z boson (J/ψ) events.

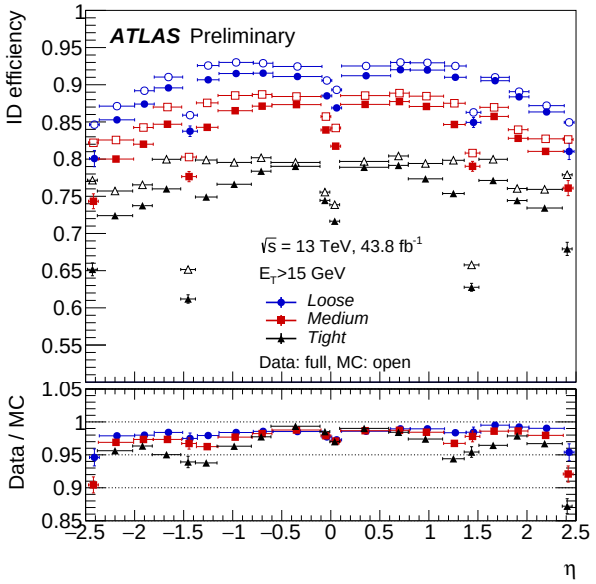
The tag electron of each pair has to fulfil the *tight* identification criteria in addition to the requirements on the transverse energy and is required not to be reconstructed in the calorimeter transition region $1.37 < |\eta| < 1.52$. The second electron candidate of the pair is called probe electron on which the *loose, medium* or *tight* identification criteria are applied. The identification efficiency is defined as the ratio of the number of probe electrons identified at a given working point and the total number of probe electrons.

The electron identification efficiency measurements for 2017 data has been performed with an integrated luminosity of 43.8 fb^{-1} and are shown in Figure 3.13 as a function of E_T and η of the probe electron.

The identification efficiency at all three working points increases with increasing E_T since the background rates are lower at higher transverse energies. The dependence on η is determined by the material distribution in the inner detector and the calorimeter. The transition region between the barrel and the endcap calorimeters around $|\eta| = 1.5$ is not instrumented and the efficiency therefore decreases in this region and in the calorimeter gap at $|\eta| = 0$. Due to known incomplete modelling of the calorimeter shower shapes by the detector simulation, the measured electron identification efficiencies differ from those predicted by the simulation. Therefore, a scale factor, defined as the ratio of the measured and the simulated electron efficiency, is applied to the simulated electron events depending on electron E_T and η to correct for the discrepancy. Measured uncertainties on the scale factors are taken into account as systematic uncertainty in the data analyses.



(a)



(b)

Figure 3.13: Electron identification efficiency determined from $Z \to ee$ and $J/\psi \to ee$ events using the tag-and-probe method for *loose, medium* and *tight* identification criteria as a function of (a) the transverse energy E_T and (b) the pseudorapidity η of the probe electron (see text). The measurements performed with data (full markers) are compared to the Monte Carlo simulation (open markers) [108].

The electron energy scale and resolution are calibrated using $Z \to ee$ decays [109]. The differences in energy scale (E) and energy resolution (σ_E/E) between data and Monte Carlo simulation are described by a factor α_i and a constant term c_i, respectively, depending on the $|\eta|$ region i:

$$E^{\text{data}} = E^{\text{MC}}(1+\alpha_i) \quad \text{and} \quad \left(\frac{\sigma_E}{E}\right)^{\text{data}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c_i \tag{3.10}$$

Both corrections are applied to the data and determined by comparing the shapes of the ee invariant mass distributions in $Z \to ee$ data and simulation.

The energy scale and resolution corrections are measured in intervals of η in the region $|\eta| < 2.47$. The accuracy of the energy scale measurement varies from 0.03% to 0.2%. The additional constant term in the energy resolution is smaller than 1% in the calorimeter central region and between 1% and 2% in the endcaps. In the transition region between barrel and endcaps, values of about 3% are reached.

3.4.2 Muon Reconstruction and Identification

The muon reconstruction is performed with tracks reconstructed in the inner detector and the muon spectrometer. In addition, the information from the calorimeters is employed in the central region of the detector ($|\eta| < 0.1$), where the routing read out cables of the inner detector require a gap in the instrumentation of the muon spectrometer. Depending on the combination of sub-systems used, four types of muon reconstruction can be distinguished [110, 111]:

Combined muon reconstruction

In the pseudorapidity range covered by both the inner detector and the muon spectrometer, muons are reconstructed by means of a combined fit of the hits in the inner detector and in the muon spectrometer. It is the most frequent and most precise type of muon reconstruction.

Segment-tagged muon reconstruction

For very low-p_T muons or in the regions of limited acceptance of the muon spectrometer, no full track reconstruction is possible in the muon spectrometer. In that case, a inner detector track is combined with track segments in one or more layers of the muon spectrometer. The muon momentum is then only measured in the inner detector.

Stand-alone muon reconstruction

Outside of the coverage of the inner detector in the forward region $2.5 < |\eta| < 2.7$ of the muon spectrometer, muon tracks are reconstructed in the muon spectrometer alone, where the track impact parameter is determined by extrapolation to the interaction point. The measured muon momentum is corrected for the expected energy loss in the calorimeters.

Calorimeter-tagged muon reconstruction

In the central detector region ($|\eta| < 0.1$) with a gap in the muon spectrometer instrumentation, tracks in the inner detector are identified as muons if they can be matched to an energy deposit in the calorimeter compatible with a minimum ionising particle.

If a muon candidate is reconstructed by more than one reconstruction method, the type with the highest priority is chosen. The highest priority is given to the combined muon reconstruction, followed by the segmented-tagged and finally calorimeter-tagged reconstruction. If a muon is reconstructed stand-alone in the muon spectrometer, in addition to of the other reconstruction methods, the reconstruction with the best fit quality and with the larger number of hits is chosen.

Muons are identified by applying quality requirements suppressing the background originating mainly from muons from decays of pions and kaons. The misidentified inner detector muon tracks from in-flight decays of charged hadrons, are characterised by kinks resulting in reduced quality of the combined track fit and incompatibility of the momenta measured in the inner detector and in the muon spectrometer. In addition to these criteria, a minimum number of hits on track in the inner detector and in the muon spectrometer is required in order to ensure a robust momentum measurement.

As for the electron identification, *loose*, *medium* and *tight* muon identification working points are defined. The *medium* muon identification working point is the default in ATLAS minimising the systematic uncertainties due to muon reconstruction and calibration. Only combined and stand-alone muon reconstruction is employed at this working point. For the *tight* muon identification working point, only combined muon reconstruction is used. This working point maximises the muon purity at the cost of some efficiency. For the $H \to ZZ^* \to 4\ell$ analysis, the *loose* working point is used for all muon candidates, maximising the reconstruction efficiency while keeping sufficient quality of the muon tracks. For this working point all four types of muon reconstruction are employed.

Within the acceptance of the inner detector ($|\eta| < 2.5$), a tag-and-probe method is used to determine the muon identification efficiency in $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events. The dimuon invariant mass system is required to be within 10 GeV of the Z boson mass for $Z \to \mu\mu$ events and within $2.7 - 3.5$ GeV for $J/\psi \to \mu\mu$ events. The tag muon is required to pass

the *medium* identification criteria and the single muon trigger requirements. The second muon (the probe muon), can be either an inner detector track, a muon spectrometer track or a calorimeter-tagged muon. The ratio of the number of probe muons identified as *loose*, *medium* or *tight* to the number of all probe muons is defined as the identification efficiency.

The measurement of the *medium* and *tight* identification efficiencies is performed in two stages. First, the efficiency $\epsilon(\mathrm{X}|\mathrm{CT})$ to identify a inner detector track with the *medium* or *tight* criteria (X=*medium/tight*) is determined using calorimeter-tagged muons (CT) as probes. Secondly, this efficiency is corrected to account for the inner detector (ID) track reconstruction efficiency $\epsilon(\mathrm{ID}|\mathrm{MS})$ using muon spectrometer (MS) tracks as probes. The final identification efficiency $\epsilon(\mathrm{X})$ is given by

$$\epsilon(\mathrm{X}) = \epsilon(\mathrm{X}|\mathrm{CT}) \cdot \epsilon(\mathrm{ID}|\mathrm{MS}). \tag{3.11}$$

The *loose* muon identification efficiency is measured separately for calorimeter-tagged muons within $|\eta| < 0.1$ and all other muon reconstruction methods. For the calorimeter-tagged muon identification efficiency, muon spectrometer tracks are used as probes, while for the remaining muon types the same approach as for the *medium* and *tight* working points is used.

For $Z \to \mu\mu$ events the expected background contribution from Monte Carlo simulation is subtracted. For $J/\psi \to \mu\mu$ events efficiency and background contribution are determined simultaneously in a maximum-likelihood fit of the $\mu\mu$ invariant mass distribution. To account for differences between the identification efficiencies measured in data and in simulation, an efficiency scale factor (SF) is applied to the muons in simulated events:

$$\mathrm{SF} = \frac{\epsilon^{\mathrm{Data}}(\eta, \phi)}{\epsilon^{\mathrm{MC}}(\eta, \phi)}. \tag{3.12}$$

The identification efficiency for muons with $|\eta| > 2.5$, i.e. outside of the inner detector acceptance, is determined using muon spectrometer tracks of *loose* and *medium* muons. The number of $Z \to \mu\mu$ events with a muon in the high-η region ($2.5 < |\eta| < 2.7$) is normalised to the number of $Z \to \mu\mu$ events observed in the region $2.2 < |\eta| < 2.5$. Muons in the central region are corrected with the SF as defined in Equation 3.12. This ratio is calculated in both, data and simulation and the double-ratio is applied as scale factor[110].

Figure 3.14 shows the muon identification efficiency as a function of muon p_T for the *medium* working point and $|\eta| > 0.1$ and as a function of η for the *loose*, *medium* and *tight* working points and muons $p_T > 10$ GeV. The measurement has been performed using the 2017 data corresponding to an integrated luminosity of 15.4 fb^{-1}. In the region $|\eta| > 0.1$, the efficiency increases from 82% for low-p_T muons to about 99% for $p_T > 10$ GeV. As a function of η the efficiency for muons with $p_T > 10$ GeV is almost constant and near to 99% for *loose* and *medium* working points and near to 97% for the *tight* working point

except for the region where there is an acceptance gap in the muon spectrometer for inner detector services to pass through. The efficiency near $\eta = 0$ can be fully recovered by using the *loose* identification which accepts the segment-tagged and calorimeter-tagged muons in addition to combined and stand-alone muons. The ratio of efficiencies measured in data and simulation is applied as a correction factor to the simulated muons events.

The muon momentum scale and resolution are measured from the invariant mass distribution of the muon pairs in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events [111]. In the central region of the detector, the momentum scale is determined with a precision of about 0.05%, increasing to 0.3% at $|\eta| \approx 2.5$. The relative p_T resolution $\sigma(p_T)/p_T$ is 1.7% (2.3%) in the central detector region and 2.3% (2.9%) in the endcaps for muons from Z boson (J/ψ) decays.

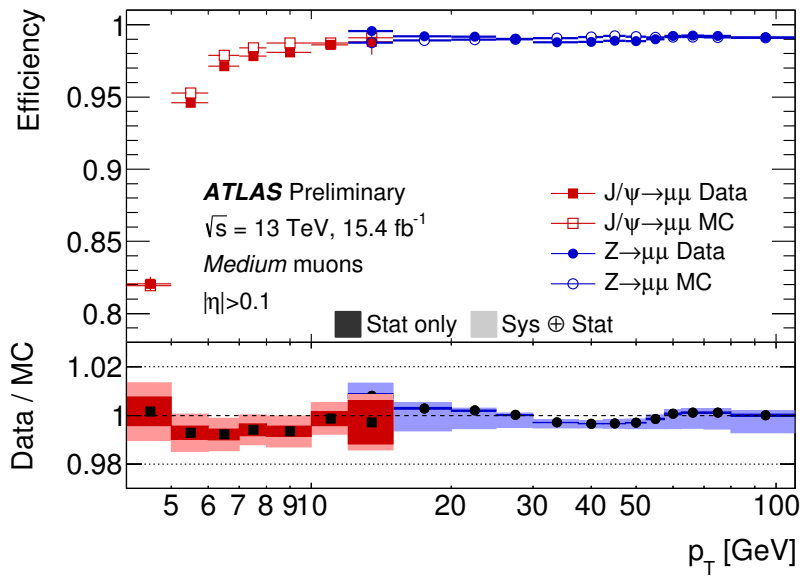
3.4.3 Jet Reconstruction and Identification

A jet is a narrow cone of hadrons and other particles produced in the hadronisation of highly energetic quarks and gluons. The hadrons produce showers in the calorimeters and charged particle tracks in the inner detector. Jets are reconstructed from the energy clusters deposited in the calorimeter using the anti-k_T jet finding algorithm [113, 114] which includes also the soft QCD radiation in the hadronisation process. The input to the jet finding algorithm are three-dimensional clusters of topologically connected energy deposits in calorimeter cells. The algorithm calculates the distances d_{ij} between pairs of topological clusters i and j and the distances d_{iB} of the topological clusters i to the beam direction B,

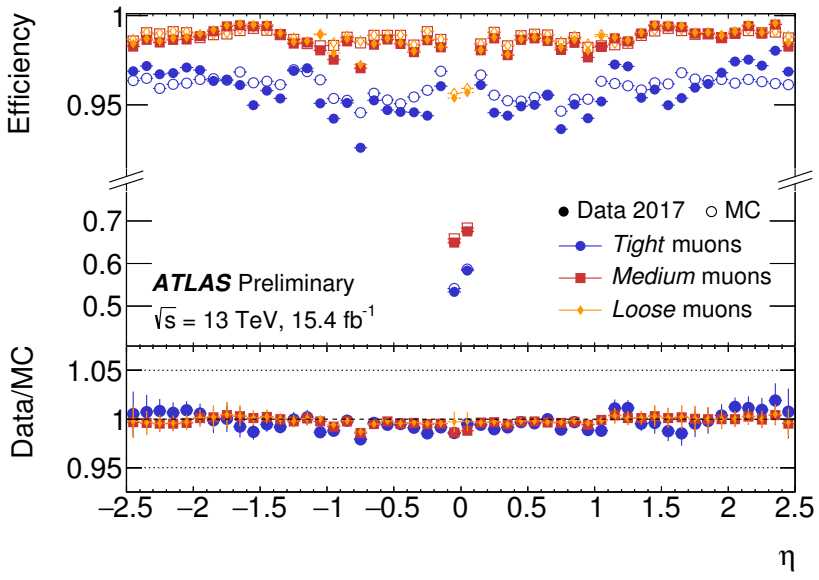
$$d_{ij} = \min\left(k_{Ti}^{-2}, k_{Tj}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2} \quad \text{and} \quad d_{iB} = k_{Ti}^{-2}, \tag{3.13}$$

where k_{Ti} is the transverse momentum of the topological cluster i, ΔR_{ij}^2 is the angular distance in the $\eta - \phi$ space between the clusters and R the radius parameter defining to the jet cone size. The jets in the $H \to ZZ^* \to 4\ell$ analysis are reconstructed using a radius parameter of $R = 0.4$. If the distance d_{ij} is smaller than d_{iB} and d_{jB}, the two clusters are combined. If d_{iB} is smaller than any d_{ij} value, the cluster i is defined to form a jet and is removed from the list of clusters. The distances are then recalculated and the process is repeated until no clusters are left.

The jet energy scale is calibrated in several stages, accounting for the features of the detector, for the reconstruction algorithm, for the fragmentation process and for corrections for pile-up interactions [115]. The correction factors are obtained from simulation and control data sets with dijet, multijet, Z+jets and $Z + \gamma$ events. In most of the calibration stages, the full four-momentum of the jet is corrected. The first step of the energy scale calibration is the *origin correction*, which improves the η resolution of the jets. The four-momentum of the reconstructed jet is recalculated to point to the primary vertex instead of to the centre of the detector.



(a)



(b)

Figure 3.14: Muon identification efficiency determined from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events for (a) the *medium* identification working point as a function of p_T and $|\eta| > 0.1$ for (b) the *loose*, *medium* and *tight* identification working points as a function of η and muon $p_T > 10$ GeV. The efficiencies have been measured using 15.4 fb^{-1} of data collected in 2017 at $\sqrt{s} = 13$ TeV [112].

In the next step, the pile-up induced energy from the same or nearby colliding proton bunches is subtracted in an *area-based* and *residual corrections*. In the *area-based correction*, the median of the distributions of the product of the jet p_T density and the jet area in the $\eta - \phi$ plane is subtracted from the jet four momentum. The jet p_T density is the ratio of the jet transverse momentum and the jet area estimated from dijet Monte Carlo events. It is calculated in the central region of the detector and does not fully describe the pile-up in the forward calorimeter region. Therefore, the *residual correction* is applied, given by the difference between the reconstructed and the true particle-level jet p_T.

After the pile-up correction, the reconstructed jet four-momentum is corrected to the particle-level energy scale by the *absolute jet energy scale* and η calibration taking into account the bias in the jet η reconstruction. The correction factor is derived from simulated dijet events by matching geometrically the reconstructed jets to the particle-level jets and is given by the inverse of the average energy response, which is defined as the mean of a Gaussian function fitted to the distribution of the ratio of reconstructed to true jet energy. The jet response is determined as a function of η and of the true jet energy.

The reconstructed jet energy and the related uncertainties are further improved by using the calorimeter, muon spectrometer and inner detector information in the *global sequential calibration*. Using several observables, like the fraction of jet energy measured in the first layer of the hadronic tile calorimeter or the number of tracks with $p_T > 1$ GeV associated with the jet, an independent correction of the jet four-momentum is derived as a function of the true p_T and $|\eta|$ values by inverting the reconstructed jet response in Monte Carlo events. The last step of the calibration is the *residual in situ calibration* which is performed using well-measured photons, Z bosons and calibrated jets in data as reference.

In the central region of the detector, the jet energy scale is calibrated with an accuracy of better than 1% for 100 GeV$< p_T <$ 500 GeV and to better than 4.5% for lower p_T jets. For jets in the forward region, i.e. $|\eta| > 0.8$, an additional uncertainty of 2% is added.

In the calibration steps described above, the average transverse energy from pile-up interactions is subtracted from the energy of the reconstructed jet. However, due to local fluctuation in the pile-up activity, additional pile-up jets can still occur. It is essential for data analyses to reject such jets. Within the $|\eta|$ coverage of the inner detector ($|\eta| < 2.4$) this can be done using the jet-vertex tagger method [116, 117] which makes use of inner detector track information. The method uses two discriminating variables related to the number of tracks associated to the jet and originating from the primary vertex. The jet-vertex tagger gives small output values for pile-up jets, while they tend to be large for signal jets.

In order to select Higgs boson candidates from bbH or ttH production, the identification of b-jets (b-tagging) is important. b-jets are distinguished from light quark and gluon jets based on the long lifetime of the b hadrons and reconstruction of their decay vertices. b-jet identification is only possible within the coverage of the inner detector $|\eta| < 2.5$. Secondary vertices are searched for or the primary vertex association is tested using three algorithms. Their outputs are combined in a multivariate discriminant by the so-called MV2c10b-tagging algorithm [118, 119]. The b hadrons are identified at different working points with efficiencies of 60%, 70%, 77% and 85%.

4

Detection of Higgs Boson Decays to Four Leptons at 13 TeV

The studies of the Higgs boson coupling properties presented in this thesis, including the measurements of the production cross sections in exclusive phase space regions and CP-violating observables, are performed in the $H \to ZZ^* \to 4\ell$ decay channel. In this chapter the signal and background processes and the signal selection for the inclusive reconstruction are described.

After the introduction of the signal and background processes, the recorded collision data sets and the corresponding Monte Carlo samples are described, followed by a discussion of the inclusive event selection criteria and the background estimation procedure and presentation of the measurement results. The measurements in this thesis are performed with data from different Run 2 data taking periods. The inclusive analysis is similar for all data sets. Deviations from the descriptions in this chapter for individual measurements are indicated in the respective chapters.

4.1 Signal and Background Processes

The $H \to ZZ^* \to 4\ell$ signal is characterised by final states with four prompt leptons originating from the same primary vertex and with low energy deposit around each lepton. There are also other background processes, with similar final state topology. For the measurement of the Higgs boson properties a good understanding and good separation of the signal and background processes is needed.

4.1.1 The $H \to ZZ^* \to 4\ell$ Decay Channel

The SM Higgs boson with a measured mass of $m_H = 125.09 \pm 0.24$ GeV [46] decays into a pair of Z bosons with a branching ratio of about 3%. Since the Higgs boson mass is below the threshold for the decay into two real Z bosons, one of the two Z bosons is produced off-shell (Z^*). The Z bosons decay either hadronically or leptonically [24].

The dominant Z boson decay mode is the hadronic decay with a branching ratio of about 70%. However, it is rather difficult to distinguish such decays from the large QCD multijet background. The next to dominant Z boson decay mode with a branching ratio of 20% is the decay into neutrinos, which escape the detector and can only be identified via the emitting total missing transverse energy. About 3% of the Z bosons decay into τ leptons pairs. The reconstruction of the hadronically decaying τ leptons is accompanied with a higher misidentification rate and worse energy resolution than decays of the Z bosons into electrons or muons. Leptonic τ decays have a small branching ratio and also worse energy resolution due to the two neutrinos in the final state. The Z boson decays into light charged leptons (where lepton is electron or muon; $\ell = e, \mu$) have a branching ratio of about 7% and the clearest signature in the detector with the best signal-to-background ratio. Therefore, the experimentally most promising $H \to ZZ^*$ decay mode, with a branching ratio of $(1.250 \pm 0.027) \cdot 10^{-2}$ % [31], is when each of the two Z bosons decays into a pair of opposite charged electrons or muons, $H \to ZZ^* \to 4\ell$. The respective tree-level Feynman diagram is shown in Figure 4.1.

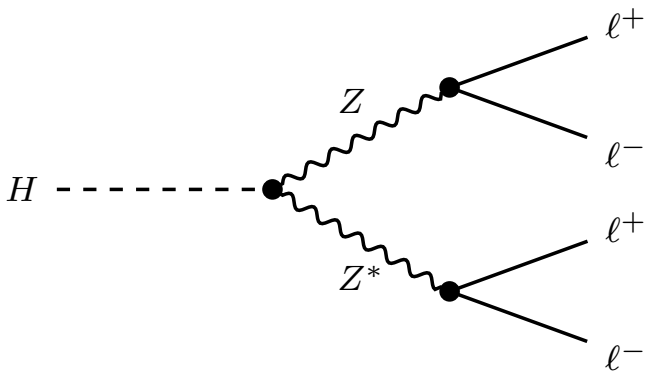


Figure 4.1: Tree-level Feynman diagram of the $H \to ZZ^* \to 4\ell$ decay (with $\ell = e, \mu$).

Since the Higgs and the Z bosons have a very short lifetime, the four leptons originate from the same primary vertex. The electrons and muons can be efficiently reconstructed in the ATLAS detector, allowing also for the full reconstruction of the Z and Higgs boson four-momenta. In addition, the high lepton energy and momentum resolution allows for a very good four-lepton invariant mass resolution of about 1-2%. Since the branching ratio for the $H \to ZZ^* \to 4\ell$ decay is very small, the measurement of this process can only be achieved with high muon and electron reconstruction efficiency.

In case of additional non-SM Higgs boson couplings, the kinematic properties vary with respect to the SM prediction. Such processes are referred to as effects beyond the SM (BSM).

As opposed to non-prompt leptons from hadronic decays, the leptons from Z boson decays have only small amounts of energy deposits around them. This allows for the discrimination against the reducible background processes, described in the following.

4.1.2 Background Processes

The background processes in the $H \to ZZ^* \to 4\ell$ decay channel can be divided into the irreducible background with four prompt leptons in the final state as in the signal, and the reducible background which contains non-prompt or fake leptons, which can be distinguished from the signal.

4.1.2.1 Irreducible Background

Non-resonant SM ZZ^* pair production with subsequent decays into four leptons is the main background process in the $H \to ZZ^* \to 4\ell$ decay channel. The dominant production mode for this process is quark-antiquark annihilation, $q\bar{q} \to ZZ^*$. If the four-lepton invariant mass is near the Z boson resonance, the dominant process is $q\bar{q} \to Z/\gamma^* \to 4\ell$. The Feynman diagrams for these processes are shown in Figure 4.2(a) and Figure 4.2(b). A much smaller contribution to the ZZ^* background originates from gluon induced ZZ^* production, $gg \to ZZ^*$. Since the Z bosons do not couple directly to the gluons, this process is only possible via a quark loop, as shown in Figure 4.2(c). An even smaller contribution to the final state with four prompt leptons is due to triboson ($VVV = WWZ, WZZ, ZZZ$) and $t\bar{t}Z$ production. The corresponding Feynman diagrams are shown in Figure 4.3.

Even though the above processes have the same number of prompt leptons in the final state as the signal, discrimination of this irreducible background from the signal is nevertheless possible to a certain degree by exploiting kinematic properties of the four-lepton final state and its invariant mass. The contribution of the irreducible background is estimated from simulation as described later on.

4.1.2.2 Reducible Background

Smaller background contributions originate from Z+jets, $t\bar{t}$ and WZ production (see Figure 4.4), where the dominant contribution is from Z+jets events. These processes are referred to as reducible background, because there are less than four prompt leptons in the final state and at least one of the leptons originates from a gluon or quark jet. Hence, lepton isolation criteria allow for a good discrimination from the signal process.

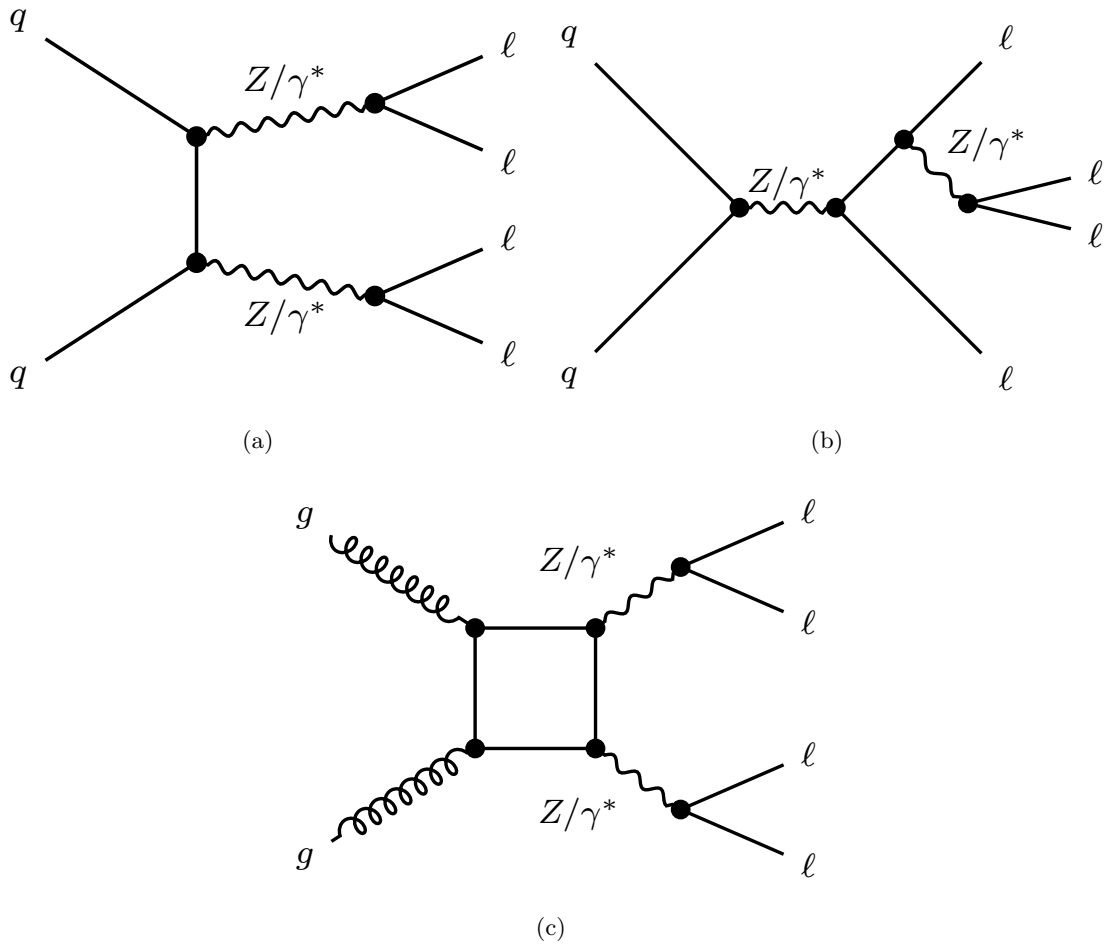


Figure 4.2: Tree-level Feynman diagrams for non-resonant $ZZ^* \to 4\ell$ production via (a), (b) quark-antiquark annihilation and (c) gluon fusion.

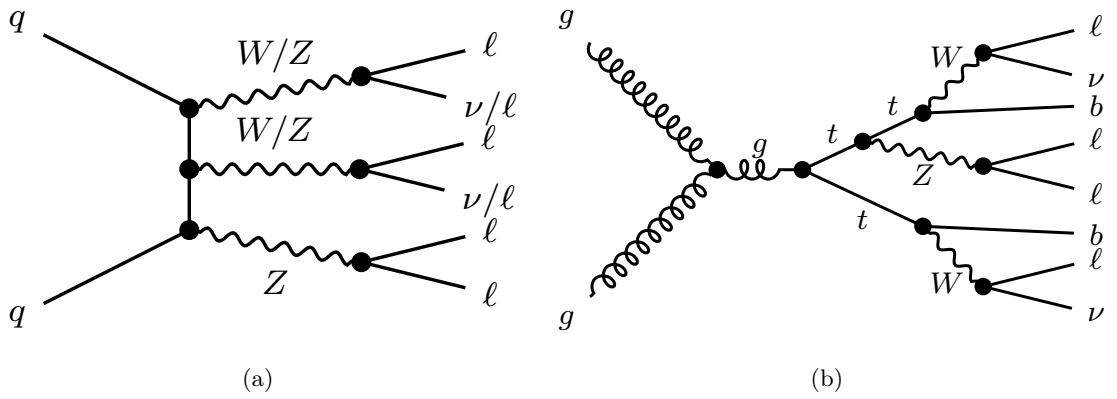


Figure 4.3: Tree-level Feynman diagrams for the minor irreducible background processes, (a) VVV and (b) $t\bar{t}Z$.

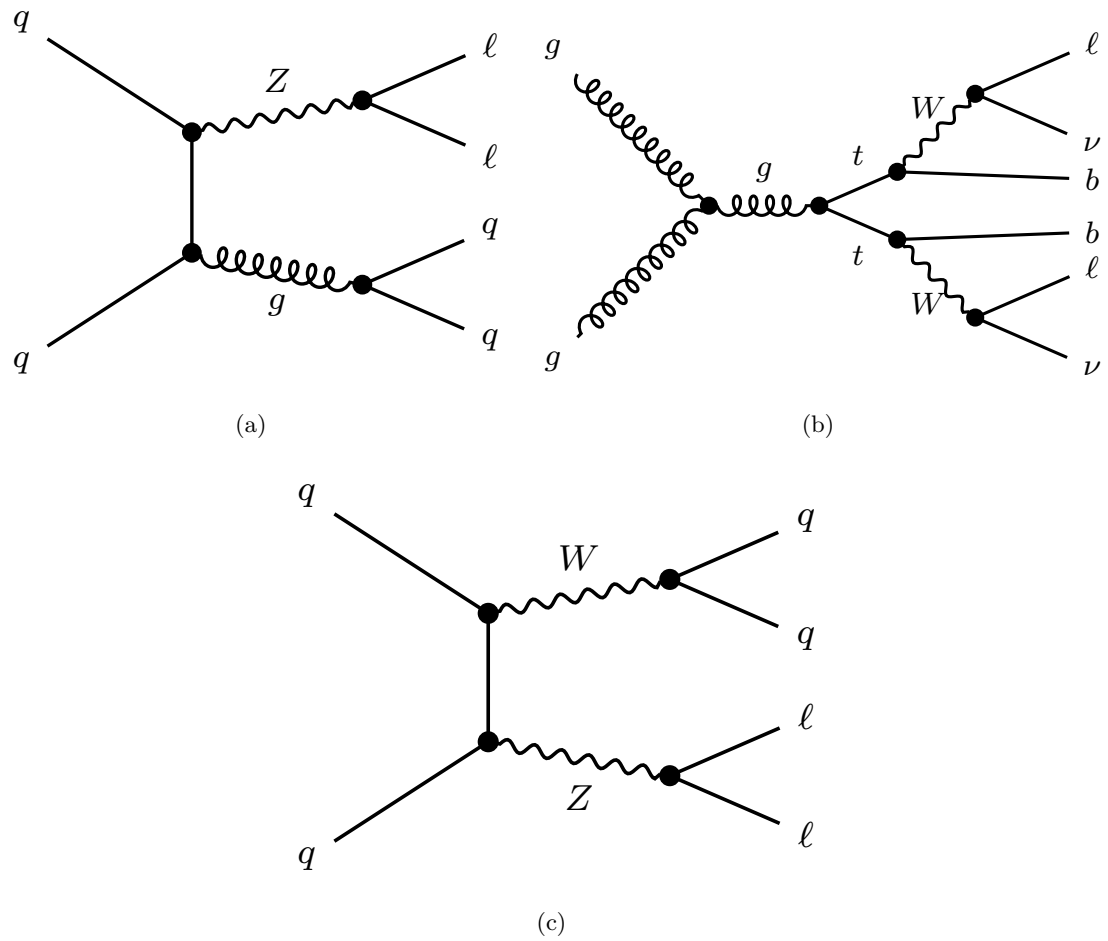


Figure 4.4: Tree-level Feynman diagrams for the main reducible background processes, (a) Z+jets, (b) $t\bar{t}$ and (c) WZ.

The reducible background events are classified according to the origin of the non-prompt leptons. The same-flavour and opposite-charge lepton pair closest to the Z boson mass in such a process is assigned to a Z boson decay (prompt lepton pair $\ell\ell$). Depending on the flavour of the second lepton pair, the background events are then divided into $\ell\ell+\mu\mu$ and $\ell\ell+ee$ final states. The non-prompt muons in the $\ell\ell+\mu\mu$ final state originate mainly from semi-leptonic decays of heavy-flavour hadrons. Non-isolated muons in Z+jets and $t\bar{t}$ production contribute to the reducible background, while WZ production in this final state is very rare. The non-prompt electrons in $\ell\ell+ee$ final states originate mainly from photon conversions or light-flavour jets misidentified as electrons, with a small contribution from semi-leptonic decays of heavy-flavour hadrons. The contribution of the Z+jets, $t\bar{t}$ and WZ backgrounds is estimated using dedicated control data depleted in signal.

4.2 Event Samples

Various measurements in this thesis are performed with three different data sets of proton-proton collisions with increasing integrated luminosities, as described in Section 4.2.1. All data sets are recorded during the Run 2 data taking period from 2015 to 2018 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The corresponding simulated Monte Carlo samples of all signal and background processes are introduced in Section 4.2.2.

4.2.1 Data Sets

The integrated luminosity of the proton-proton collisions delivered in each year of the Run 2 data taking is shown in Table 4.1, together with the luminosity recorded by the ATLAS detector. The data used for the analysis is subject to several quality requirements, i.e. recorded events are rejected if any relevant detector components was not fully functional. Therefore, the integrated luminosity of the analysed data is smaller than the recorded one, as summarised in Table 4.1.

Table 4.1: The integrated luminosity for each year of the Run 2 data taking, as delivered by the LHC, recorded by the ATLAS detector and analysed in this thesis. The corresponding average amount of pile-up interactions is also shown.

Data taking	Integrated luminosity [fb^{-1}]			Average
period	Delivered	Recorded	Analysed	pile-up $\langle\mu\rangle$
2015	4.2	3.9	3.2	13.4
2016	38.5	35.6	32.9	25.1
2017	50.2	46.9	43.7	37.8
2018	63.3	60.6	59.2	36.1

For the measurement of the Higgs boson production cross sections in exclusive phase space regions and their projection to high luminosity, discussed in Chapters 5 and 6, the proton-proton collision data from 2015-2017 corresponding to a total integrated luminosity of 79.8 fb^{-1} have been analysed. The average number of pile-up interactions in these three years is 31.9 with a peak value of 80.

The first of the two analyses within an effective field theory framework described in Section 7.3, as well as the study of the CP-invariance in Chapter 8, are performed on the data set taken during 2015 and 2016. The analysed integrated luminosity of this data set is 36.1 fb^{-1} with an average and a maximum number of pile-up interactions being 22.9 and 50, respectively.

The second analysis within an effective field theory framework uses a slightly different approach compared to the first one and is performed with the Run 2 data set (Section 8). This study uses as input the measurement of the Higgs boson production cross sections in exclusive phase space regions which has also been updated with the full Run 2 data set. The analysed integrated luminosity of this data set is 139 fb^{-1} with an average number of pile-up interactions of 33.7 and the peak value of 80.

4.2.2 Event Simulation

The hard-scattering process in each signal and background processes is simulated using Monte Carlo event generators. Parton shower programs, which are interfaced to the event generators are used to simulate the hadronisation of the final state and initial and final state radiations. The parameters of the parton shower and the fragmentation modes are tuned to agree with the observed data.

The generated events are then fed through the detailed, full simulation of the the ATLAS detector response [89] within the GEANT4 framework [90]. The hits produced by the simulation in the detector elements are converted into electronic signals and digitised. Afterwards, they are reconstructed in the same way as the data.

Multiple proton-proton collisions (pile-up) are taken into account by superimposing minimum bias events on the simulated signal events. In order to match the mean number of interactions per bunch crossing to data, the simulated pile-up profile is reweighted to the corresponding observed distribution.

4.2.2.1 Standard Model Signal Processes

The production of the SM Higgs boson via the ggF, VBF, VH and ttH production modes is in general simulated with the POWHEG-BOX v2 event generator [120–124]. The bottom and charm hadron decays are simulated with EVTGEN [125] which is interfaced to the event generator. The parton distribution functions are in general modelled with the next-to-leading-order (NLO) PDF set PDF4LHCnlo [88]. This set is replaced by a more precise PDF set at next-to-next-to-leading-order (NNLO), the PDF4LHCnnlo, for the ggF production in the 79.8 fb^{-1} and 139 fb^{-1} data sets. For ggF production with additional jets in the final state additionally the POWHEG [122] method for merging the NLO cross section calculation with the parton shower and the MiNLO method [124] is used to achieve NLO accuracy for the inclusive cross section. Finally, to obtain an NNLO accuracy in QCD corrections for ggF production, a reweighting procedure (NNLOPS) is applied using the HNNLO program [126, 127].

The matrix elements for VBF, $q\bar{q} \to VH$ and ttH production are calculated to NLO QCD. For VH production, the MiNLO method is used to merge 0-jet and 1-jet event samples. The $g\bar{g} \to ZH$ contribution is modelled at leading-order (LO) QCD with the PDF4LHC set [88]. For the modelling of the ttH production for the 36.1 fb^{-1} analysis, the Powheg-Box v2 event generator is replaced by MadGraph5_aMC@NLO [128] using the CT10nlo PDF set [129] at NLO QCD.

The Higgs boson production in association with bottom quarks, bbH, is simulated with MadGraph5_aMC@NLO using the NNPDF23 PDF set [130]. The process is modelled at NLO in QCD. For the analyses performed with the 79.8 and 139 fb^{-1} data sets, also the Higgs boson production in association with a single top quark (tH) has been taken into account. These processes is modelled with MadGraph5_aMC@NLO using the NNPDF30 PDF set at NLO QCD. For the analysis with the 79.8 fb^{-1} data set, the tHq process has been simulated with MadGraph5. For this data set, the CT10nlo PDF set has been used for all tH processes.

Except for the case of ttH production, the $H \to ZZ^* \to 4\ell$ decays as well as the parton shower modelling have been simulated with the Pythia8 [131] generator using the AZNLO tuned parameter set [132]. For the Powheg ttH sample the parton showering has also been performed with Pythia8, but with the A14 tuned parameter set [133]. The Herwig++ [134] parton shower program was employed for the MadGraph5_aMC@NLO ttH sample.

All SM Higgs boson signals are simulated for a Higgs boson mass of $m_H = 125$ GeV. The Higgs boson production cross sections are scaled to the highest-accuracy prediction available at the time of the analysis [29]. The branching ratio to the four-lepton final states is predicted by Prophecy4f [135, 136]. This program includes the complete NLO QCD+EW corrections, the interference effects between identical final state fermions, and the leading two-loop heavy Higgs boson corrections to the four-fermion width. The cross section prediction used for the normalisation of the Monte Carlo samples, as well as the accuracy of the calculations are summarised in Table 4.2.

4.2.2.2 Beyond the Standard Model Signal Processes

Additional signal samples with different values of non-SM coupling parameters (introduced in Section 2.5) are simulated for the measurements of the tensor structure of the Higgs boson coupling within effective field theory frameworks. Separate sets of BSM signal samples are produced for the two measurements with 36.1 fb^{-1} and with 139 fb^{-1} of data.

Table 4.2: The predicted SM Higgs boson production cross sections (σ) for ggF, VBF and five associated production modes in proton-proton collisions for $m_H = 125$ GeV at $\sqrt{s} =$ 13 TeV [29]. The quoted uncertainties correspond to the total theoretical systematic uncertainties calculated by adding in quadrature the uncertainties due to missing higher-order corrections and PDF+α_s. In addition, the accuracy of the calculations in QCD and EW is shown [137–162]. The decay branching ratios ($\mathcal{B}$) with the associated uncertainty for $H \to ZZ^*$ and $H \to ZZ^* \to 4\ell$, with $\ell = e, \mu$, are given.

Production process		σ [pb] used for		Accuracy of the calculation
		36.1 fb^{-1}	$>$ 79.8 fb^{-1}	
ggF	$gg \to H$	48.5 ± 2.400	48.6 ± 2.400	N^{3}LO QCD, NLO EW corrections
VBF	$qq \to H$	3.78 ± 0.080		NLO QCD, NLO EW corrections with approximate NNLO QCD corrections
WH	$qq \to WH$	1.369 ± 0.028	1.373 ± 0.028	NNLO QCD, NLO EW corrections
ZH	$gg/qq \to ZH$	0.88 ± 0.040		NNLO QCD, NLO EW corrections
ttH	$gg/qq \to t\bar{t}H$	0.51 ± 0.050		NLO QCD, no NLO EW correction
bbH	$gg/qq \to b\bar{b}H$	0.49 ± 0.120		NNLO QCD (5FS), NLO QCD (4FS), no EW correction
tH	$qq/gg \to tH$	–	0.09 ± 0.010	NLO QCD (5FS), no NLO EW correction
Decay process		$\mathcal{B}$ [$\cdot 10^{-4}$]		
$H \to ZZ^*$		262.00 ± 6.000		NLO QCD, NLO EW corrections
$H \to ZZ^* \to 4\ell$		1.250 ± 0.027	1.240 ± 0.027	NLO QCD, NLO EW corrections

For the measurement with the 36.1 fb^{-1} data set, the ggF, VBF and VH signal samples with different values of the BSM coupling parameters are simulated with MadGraph5_aMC@NLO. The samples correspond to the effective Lagrangian of the Higgs Characterisation framework described in Section 2.5.2.2, implemented into the event generator via the HC_NLO_X0_UFO [33, 163, 164] model. The model incorporates the new dimension-six operators in the standard Universal FeynRules Output (UFO) format created using the FeynRules framework [165, 166]. The ggF production is modelled at NLO QCD using the NNPDF30nlo PDF set. The merging of additional jets in the final state to parton shower simulations is done via the FxFx merging scheme [167]. The VBF process is simulated together with the VH process with hadronic V boson decays (VH-Had). Additional samples are produced for the VH production with leptonic V boson decays. Both the VBF+VH-Had and the VH-Lep are modelled at LO QCD

using the NNPDF23lo PDF set. In addition, a smaller number of equivalent VBF and VH samples is generated at NLO QCD accuracy using the NNPDF30nlo PDF set to estimate the relative uncertainties of higher-order QCD effects as a function of the BSM coupling parameters.

For the interpretation within effective field theories, which is performed with the 139 fb^{-1} data set, the ggF, VBF+VH-Had, VH-Lep and ttH samples are generated at LO QCD with MadGraph5_aMC@NLO and with the NNPDF23lo PDF set. The CKKW-L method [168] is used for the jet merging in the ggF process. For these samples the SMEFT model, described in Section 2.5.2.1, is used instead of the Higgs Characterisation model. It is implemented in the event generator via the SMEFTsim_A_U35_MwScheme_UFO_v2.1 UFO [32, 169] model.

Similarly as for the SM Higgs boson signal samples, the bottom and charm hadron decays are simulated with EvtGen. The parton showering in all BSM processes is modelled with Pythia8 and the A14 set of tuned parameters [133] is used. All BSM signal processes are simulated with the Higgs boson mass of $m_H = 125$ GeV. To account for higher-order QCD effects comparable to those in the SM scenario, the cross sections of the BSM signal samples are scaled to the best prediction according to the k-factors from the SM, as described in Section 7.2.2.

4.2.2.3 Background Processes

The SM ZZ^* background production via quark-antiquark annihilation is simulated with the Sherpa2.2.1 event generator for the analysis of the 36.1 fb^{-1} data set and with Sherpa2.2.2 for the two others [170–172]. Both provide a matrix element calculation accurate to NLO QCD for 0-jet and 1-jet final states. Final states with more than two jets are modelled in LO QCD. The parton showering is modelled with Sherpa [173] using the ME+PS@NLO prescription [174] for the jet merging. In addition, NLO EW corrections are applied as a function of the invariant mass of the two Z bosons [175, 176].

The $gg \to ZZ^*$ process is generated with gg2VV [177] at LO QCD. Bottom and charm hadron decays are simulated with EvtGen and the parton shower is obtained with Pythia8. Higher-order QCD corrections for the gluon induced ZZ^* production are available for massless quark loops [178, 179] in the heavy top-quark approximation [180], which include $gg \to H^* \to ZZ$ processes. According to these, the LO cross section from the simulation is scaled by a k-factor of 1.7 ± 1.0 to account for higher-order effects.

The Z+jets process is modelled with the Sherpa2.2.2 for the 36.1 fb^{-1} data set and Sherpa2.2.1 event generator for the 79.8 and 139 fb^{-1} data sets. The matrix elements are calculated at NLO accuracy for up to two partons and at LO accuracy for three

and four partons with COMIX [171] and OPENLOOPS [172]. The parton showering is performed with SHERPA using the ME+PS@NLO prescription for the jet merging.

The $t\bar{t}$ background is generated with POWHEG-BOX v2 at NLO QCD. The showering of the partons and the hadronisation is performed with PYTHIA8. Heavy hadron decays are simulated with EVTGEN. For the 36.1 fb^{-1} data set, PYTHIA6 is used instead.

The WZ background is generated at NLO QCD using POWHEG-BOX v2 interfaced to PYTHIA8 and EVTGEN. The triboson background VVV is modelled at LO QCD using SHERPA2.2.1 (36.1 fb^{-1}) and SHERPA2.2.2 (79.8 and 139 fb^{-1}). The small contributions from $t\bar{t}Z$ events with both top quarks decaying semi-leptonically and the Z boson decaying leptonically are modelled with MADGRAPH5_AMC@NLO interfaced to PYTHIA8 and EVTGEN at LO QCD. Higher-order corrections are taken into account by normalising the cross section to the prediction which includes the two dominant terms at LO and NLO in a mixed perturbative expansion in the QCD and EW coupling. For the 36.1 fb^{-1} data set only the $t\bar{t}Z$ with leptonically decaying Z boson is taken into account. The process is generated with MADGRAPH5, interfaced to PYTHIA8 and EVTGEN. The analysis done on the 139 fb^{-1} data set also includes the estimation of smaller background contributions from tWZ, $t\bar{t}W$, $t\bar{t}t$, $t\bar{t}t\bar{t}$ and tZ processes. They are simulated with MADGRAPH5_AMC@NLO interfaced to PYTHIA8.

4.3 $H \to ZZ^* \to 4\ell$ Event Selection

Following the properties of the signal and background topologies, a set of event selection criteria are defined, starting from the general data taking quality criteria, over the trigger selection and particle reconstruction, to the selection of Higgs boson candidates events. The full set of selection requirements is summarised in Table 4.3.

Data Quality

Prior to every data analysis, each recorded collision event has to pass certain data quality requirements. First, events are rejected if not all detector components were operating correctly when they were recorded. In addition, the events are required to have at least one collision vertex with at least two associated tracks. The collision vertices are reconstructed from all inner detector tracks with a transverse momentum larger than 400 MeV. The primary vertex is defined as the vertex with the highest sum of squared track transverse momenta, $\sum p_T^2$.

Trigger Selection

Four-lepton events are selected with single-lepton, di-lepton and tri-lepton triggers including the mixed electron-muon triggers in the two latter one. The trigger accepts an event if certain lepton identification and p_T-threshold criteria are satisfied. The p_T-thresholds are increasing during the data taking periods, due to an increasing peak luminosity. A summary of the trigger requirements for the $H \to ZZ^* \to 4\ell$ analysis for the different data taking periods is given in Appendix A. The trigger efficiency for the Higgs boson candidates after the final state selection is 98%.

Selection of Reconstructed Final State Particles

The general description of the reconstruction, identification and calibration of the final state particles is given in Section 3.4. The exact criteria employed for the certain selection of electrons, muons, jets and b-jets are summarised in the following. The latter two are needed for the production cross section measurements in exclusive phase space regions, as performed in the subsequent chapters.

The electrons are required to satisfy the *loose* identification criteria allowing for the highest electron reconstruction efficiency and therefore providing the highest four-lepton selection efficiency. Electrons are required to have a pseudorapidity within the inner detector coverage of $|\eta| < 2.47$ and to have a transverse energy larger than 7 GeV. As signal electrons are expected to emerge from the primary vertex, small longitudinal impact parameters of $|z_0 \cdot \sin(\theta)| < 0.5$ mm are required .

Similarly as electrons, muons are identified with the *loose* identification criteria. Only muons with $|\eta| < 2.47$ and a minimum transverse momentum of 5 GeV are selected. Due to a lower purity, calorimeter-tagged muons are required to have $p_T > 15$ GeV. In order to suppress the cosmic muon background a cut on the absolute value of the muon transverse impact parameter is applied, $|d_0| < 1$ mm. The background from decays of hadronised bottom quarks can be reduced by the requirement $|z_0 \cdot \sin(\theta)| < 0.5$ mm on the longitudinal impact parameter, since these muon tracks are displaced from the primary vertex. For the analysis of the 79.8 fb^{-1} and 139 fb^{-1} data sets an additional criteria is applied. In order to minimise the reducible background contribution, at most one calorimeter-tagged or stand-alone muon is allowed per event.

Jets are reconstructed with the anti-k_T algorithm with a radius parameter of $R = 0.4$ and are required to have $p_T > 30$ GeV and $|\eta| < 4.5$. The contribution from pile-up jets is reduced by applying a cut on the jet-vertex tagger (JVT) discriminant described in Section 3.4.3. Jets within $p_T < 60$ GeV and $\eta < 2.4$ are rejected if the corresponding jet-vertex tagger discriminant has a value JVT< 0.59 which leads to a selection efficiency

of 92% for hard-scattered jets. In addition, for the analysis of 79.8 fb^{-1} and 139 fb^{-1} data sets, forward pile-up jets ($p_T < 50$ GeV and $|\eta| > 2.5$) are rejected using the jet-vertex tagger working point with 90% selection efficiency for hard scattered jets.

Jets within $|\eta| < 2.5$ containing b hadrons are identified with the MV2c10b-tagging algorithm using the working points with a 70% selection efficiency for true b hadron jets. For the 139 fb^{-1} data set, a pseudo-continuous b-tagging weight is assigned to each jet, which combines the 60%, 70%, 77% and 85% efficiency working points.

If the same detector information is used for the reconstruction of more than one particle type, the ambiguities are solved with the so-called overlap removal. In case that the reconstructed electron and muon share the same inner detector track and the muon is obtained from calorimeter-tagged reconstruction, the muon is removed from the event. For all other muon types, the electron is removed. In addition, electrons or muons that geometrically overlap with a reconstructed jet within a cone of radius $\Delta R = 0.2$ are also removed. For the analysis of the 139 fb^{-1} data set, the cone radius for the overlap between muons and jets is reduced to $\Delta R = 0.1$. In addition, an electron-electron overlap removal is applied. If the calorimeter clusters of two electrons overlap, the one with the higher E_T is kept.

Selection of Lepton Quadruplets

In the next selection step, each event is required to have at least four leptons which can be combined into two same-flavour and opposite-charged lepton pairs, building a so-called lepton quadruplet. In signal events with four leptons, there are two such quadruplet combinations possible and additional leptons introduce further combinations. The following selection criteria are applied on each quadruplet combination.

The three leptons with highest transverse momenta in each quadruplet are required to have $p_T > 20$ GeV, $p_T > 15$ GeV and $p_T > 10$ GeV, respectively. The same-flavour and opposite charged lepton pair with an invariant mass m_{12} closest to the Z boson mass is defined to be the leading lepton pair, while the second lepton pair corresponding to the off-shell Z boson candidate with an invariant mass m_{34} is referred to as the subleading one. Each quadruplet is classified as $4e$, 4μ, $2e2\mu$ and $2\mu2e$ quadruplet according to the lepton flavour of the leading and subleading lepton pairs, with the first two leptons represent the leading and the second the subleading lepton pair.

The leading lepton pair associated to the on-shell Z boson in each quadruplet is required to have an invariant mass of 50 GeV $< m_{12} <$ 106 GeV, while the invariant mass of the subleading lepton pair has to satisfy $m_{min} < m_{34} <$ 115 GeV. The threshold m_{min} depends

Table 4.3: Event selection criteria applied in the inclusive $H \to ZZ^* \to 4\ell$ analysis

Selection of Reconstructed Final State Particles	
Data quality:	All detector components have be fully operational
Vertex:	At least one collision vertex with two associated tracks
Trigger:	Single-lepton, di-lepton and tri-lepton trigger
Physical Object Selection	
Electron:	*Loose* identification criteria, $\|\eta\| < 2.47$, $E_T > 7$ GeV, $\|z_0 \cdot \sin(\theta)\| < 0.5$ mm
Muon:	*Loose* identification criteria, $d_0 < 1$ mm, $\|z_0 \cdot \sin(\theta)\| < 0.5$ mm
	Combined-muons: $\|\eta\| < 2.5$, $p_T > 5$ GeV
	Calorimeter-tagged muons: $\|\eta\| < 0.1$, $p_T > 15$ GeV
	Segmented-tagged muons: $\|\eta\| < 0.1$, $p_T > 5$ GeV
	Stand-alone muons: $2.5 < \|\eta\| < 2.7$, $p_T > 5$ GeV
	At most one calorimeter-tagged or stand-alone muon (*79.8 fb^{-1} and 139 fb^{-1}*)
Jets:	Anti-k_T algorithm (R=0.4), $\|\eta\| < 4.5$, $p_T > 30$ GeV
	Pile-up suppression: JVT> 0.59 ($p_T < 60$ GeV, $\|\eta\| < 2.4$)
	Forward pile-up suppression: 90% JVT working point ($p_T < 50$ GeV and $\|\eta\| > 2.5$) (*79.8 fb^{-1} and 139 fb^{-1}*)
	b-jets: MV2c10b-tagging algorithm
Overlap removal:	Ambiguities are resolved between overlapping lepton-lepton or lepton-jet pairs
Selection of Lepton Quadruplets	
Quadruplet selection:	Each combination of two same-flavour and opposite-charged lepton pairs, with
	$p_T > 20, 15, 10$ GeV for the three leptons with highest p_T
Quadruplet	Quadruplet types: $4e$, 4μ, $2e2\mu$, $2\mu2e$
classification:	based on the lepton flavour in the leading and subleading pair
	Leading lepton pair: invariant mass m_{12} closest to the Z boson mass (on-shell Z boson candidate)
	Subleading lepton pair: remaining pair (off-shell Z boson candidate with invariant mass m_{34})
Kinematic	50 GeV $< m_{12} <$ 106 GeV
requirements:	$m_{min} < m_{34} <$ 115 GeV
	$\Delta R(\ell, \ell') > 0.1(0.2)$ for same-flavour (different-flavour) lepton pairs (*36.1 fb^{-1} and 79.8 fb^{-1}*)
	$\Delta R(\ell, \ell') > 0.1$ for all lepton pairs (*139 fb^{-1}*)
	$m_{\ell\ell} > 5$ GeV alternative for same-flavour opposite-charge lepton pairs $m_{\ell\ell} = m_{14}, m_{23}$
Impact parameter	Electrons: $\|d_0/\sigma_{d_0}\| < 5$
significance:	Muons: $\|d_0/\sigma_{d_0}\| < 3$
Lepton isolation:	Electron track-based isolation ($\Delta R = 0.2$): $I^{\mu}_{track} < 15\%$ (*36.1 fb^{-1} and 79.8 fb^{-1}*)
	Muon track-based isolation ($\Delta R = 0.3$): $I^{\mu}_{track} < 15\%$ (*36.1 fb^{-1} and 79.8 fb^{-1}*)
	Electron calorimeter-based isolation ($\Delta R = 0.2$): $I^{e}_{track} < 20\%$ (*36.1 fb^{-1} and 79.8 fb^{-1}*)
	Muon calorimeter-based isolation ($\Delta R = 0.2$): $I^{e}_{track} < 30\%$ (*36.1 fb^{-1} and 79.8 fb^{-1}*)
	All leptons: $I_{track} + 0.4 \cdot I_{calo} < 16\%$ (*139 fb^{-1}*)
Common 4ℓ-vertex:	$\chi^2/N_{dof} < 6$ for 4μ quadruplet candidates
	$\chi^2/N_{dof} < 9$ for $4e$, $2e2\mu$ and $2\mu2e$ quadruplet candidates
Selection of the Higgs Boson Candidate	
Final candidate:	Event with exactly four leptons: quadruplet with m_{12} closest to m_Z
	Event with additional leptons ($p_T > 12$ GeV): quadruplet with highest matrix-element value
$m_{4\ell}$:	Final state radiation correction, Z boson mass constraint (*36.1 fb^{-1}*)
	118 GeV $< m_{4\ell} <$ 129 GeV (*36.1 fb^{-1}*)
	115 GeV $< m_{4\ell} <$ 130 GeV (*79.8 fb^{-1} and 139 fb^{-1}*)

on the four-lepton invariant mass $m_{4\ell}$, starting at 12 GeV for $m_{4\ell} < 140$ GeV and rising linearly to a maximum value of 50 GeV for $m_{4\ell} \geq 140$ GeV.

The leptons in each lepton pair should be well separated in space, with the angular distance $\Delta R(\ell, \ell')$ between the leptons in the same-flavour (different-flavour) lepton pair required to be larger than 0.1 (0.2). For the analysis of the 139 fb^{-1} data set, the same angular separation is required for all lepton pairs, $\Delta R(\ell, \ell') > 0.1$.

Leptons from J/ψ decays are rejected by the requirement that the invariant mass of two alternative opposite-charge lepton pairs within a $4e$ or 4μ quadruplet be $m_{\ell\ell} > 5$ GeV.

Due to the short life time of the Higgs boson the signal leptons are expected to originate from the primary vertex. As opposed to that, leptons from heavy-flavour hadron decays originate from secondary vertices leading to an impact parameter offset with respect to the primary vertex. Therefore, such background processes are suppressed by the requirement on the transverse impact parameter significance $|d_0/\sigma_{d_0}|$, where σ_{d_0} is the uncertainty of the d_0 measurement. All muons (electrons) are required to have $|d_0/\sigma_{d_0}| < 3(5)$.

Signal leptons are well isolated from other particles in the final state. In contrast to that, non-prompt leptons are often produced in jets, resulting in a large activity in the detector material surrounding the lepton trajectory. The contribution from Z+jets and $t\bar{t}$ production can be therefore reduced by the requirements of the track-based and calorimeter-based lepton isolation.

The track-based muon (electron) isolation is defined as the scalar sum of the p_T of all tracks within a cone of $\Delta R = 0.3(0.2)$ around the muon (electron) track, $I_{track} = \sum p_T / p_T$. Both muons and electrons have to satisfy $I_{track} < 15\%$. The calorimeter-based isolation I_{calo} is defined in a similar way. The sum of all calorimeter energy deposits within a cone size of $\Delta R = 0.2$ around the lepton is required to be smaller than 30% (20%) of the muons (electrons) transverse momenta. The described isolation criteria have been optimised for the analysis of the 139 fb^{-1} data set. For leptons with $p_T > 33$ GeV the cone size of the track-based isolation decreases linearly with p_T from $\Delta R = 0.3$ to a minimum cone size of 0.2 at 50 GeV. The cone size for the calorimeter-based isolation remains fixed at 0.2. A lepton is retained if $I_{track} + 0.4 \cdot I_{calo} < 16\%$. The described optimisation of the isolation requirement improves the signal efficiency by about 5% compared to the previous analysis for the same background rejection.

For a further suppression of the reducible background, the four quadruplet leptons are required to originate from the common 4ℓ vertex point. The corresponding four inner detector tracks are fitted under the assumption that the leptons emerge from the same vertex point. The quality of the fit, i.e. the ratio of the χ^2 and the number of degrees of freedom,

N_{dof}, is expected to be high for signal quadruplets, as opposed to the background processes. The 4μ (other 4ℓ) quadruplet candidates are accepted if the fit quality, χ^2/N_{dof} is less than 6 (9), corresponding to a signal efficiency of 99.5% for this selection requirement.

Selection of Higgs Boson Candidates

If there is more than one quadruplet per event satisfying the quadruplet selection criteria, the final Higgs boson candidate is selected according to criteria which depend on the number of leptons in the final state. In events with exactly four leptons in the final state, the quadruplet containing the leading lepton pair with an invariant mass closest to the Z boson mass is chosen as Higgs boson candidate, separately for $4e$, 4μ, $2e2\mu$ and $2\mu2e$ quadruplet classes. If there is more than one class with such a Higgs boson candidate, the quadruplet from the class with highest efficiency (i.e. ordered as 4μ, $2e2\mu$, $2\mu2e$, $4e$) is selected as the final Higgs boson candidate in the event.

For the VH and ttH production, there may be additional leptons present in the final state. Therefore, it is possible that one or more leptons are selected in the quadruplet do not descend from the Higgs boson decay, but from a V boson or top quark decay. In order to improve the lepton pairing in such cases, a matrix element based lepton pairing method is used instead of the above selection in all events with at least one additional lepton which has $p_T > 12$ GeV and satisfy the same identification and isolation criteria as the four quadruplet leptons. A matrix element of the Higgs boson decay is computed at LO with MadGraph5_aMC@NLO generator [181] for each possible quadruplet combination passing the previously described quadruplet selection criteria. The quadruplet with the largest matrix element value is chosen as the final Higgs boson candidate.

The resolution of the invariant mass of the four-lepton final state is improved by taking into account lepton energy losses due to final state radiation (FSR). The energy of a photon with $p_T > 1$ GeV within a cone of $\Delta R < 0.15$ around one of the two leading muons is added to the energy of the muon. In addition, isolated photons with $p_T > 10$ GeV are associated with the closest lepton. The overall improvement of the invariant mass resolution of this correction is about 1%.

For the analysis with the 36.1 fb^{-1} data set an additional correction on the invariant mass of the four-lepton final state is applied. The Z boson mass constraint is applied to the leptons from the leading lepton pair. A kinematic fit is performed with the knowledge of the Z boson line shape and the momentum resolution of the leptons, correcting the lepton transverse momentum and therefore improving the four-lepton invariant mass resolution by about 15% [49].

In the analysis performed with the 36.1 fb^{-1} data set, the Higgs boson candidates are required to have the invariant mass within a mass window of 118 GeV $< m_{4\ell} <$ 129 GeV, a wider mass range of 115 GeV $< m_{4\ell} <$ 130 GeV is used for the two other analyses (79.8 fb^{-1} and 139 fb^{-1}).

4.4 Background Estimation

The above selection criteria define the signal region (SR) in which the background contributions have to be estimated. The irreducible background from processes with four prompt and isolated leptons, ZZ^*, $t\bar{t}Z$ and VVV, can be reliably estimated from Monte Carlo simulation. This is not the case for the reducible Z+jets, $t\bar{t}$ and WZ background processes with non-prompt or misidentified leptons.

The simulation of these processes is subject to substantial modelling uncertainties. In addition, the rate of misidentified leptons is very low, such that the number of simulated events would have to be very large. The contribution of reducible backgrounds is therefore estimated from data-driven methods, i.e. from the measurements in background-enriched while signal-depleted control regions (CR) of data.

4.4.1 Irreducible Background Estimation

The contribution of SM ZZ^* production, which is the main background in the $H \to ZZ^* \to 4\ell$ decay channel, is estimated from simulation as described in Section 4.2.2.3. The NNLO QCD and NLO EW corrections are applied on the cross section from simulation. The normalisation and the shape of the four-lepton invariant mass distribution are validated in the signal depleted sideband region with 105 GeV $< m_{4\ell} <$ 115 GeV and 130 GeV $< m_{4\ell} <$ 350 GeV. The minor backgrounds with prompt leptons such as decays VVV and $t\bar{t}Z$ are also estimated from simulation.

For the analysis of the 139 fb^{-1} data set the statistics is large enough to obtain the normalisation for some or all of these processes, in particular the dominant ZZ^* production, using a data-driven approach. This method involves adding a normalisation factor for the estimation of the ZZ^* processes to the fit. The sideband region with 105 GeV $< m_{4\ell} <$ 115 GeV and 130 GeV $< m_{4\ell} <$ 160 GeV is simultaneously fitted with the signal region (115 GeV $< m_{4\ell} <$ 130 GeV) within N_j-bins (0, 1 and ≥ 2 jets) to obtain the normalisation from data. The advantage of this method is that theoretical systematic uncertainties on the normalisation of the ZZ^* background, as well as luminosity systematic uncertainty on the background are removed. Also the tVV background is estimated with this method using a wider sideband region of 105 GeV $< m_{4\ell} <$ 115 GeV and 130 GeV $< m_{4\ell} <$ 350 GeV. Other triboson backgrounds (ZZZ,WZZ, and WWZ) are taken directly from Monte Carlo simulation.

4.4.2 Data-Driven Reducible Background Estimation

The contribution of reducible backgrounds is estimated by means of a data-driven method in control regions (CR) of data with enhanced background and suppressed signal contributions. The control regions are constructed by relaxing or inverting the lepton selection criteria for one or two of the four-final state leptons. For each of the background components a dedicated control region is built. The higher amount of data in the control regions compared to the signal region allows for a good precision of the background estimate.

The expected yields of the different background components in the control regions are obtained from a simultaneous fit of the background normalisation in all control regions. The resulting yields are then extrapolated to the signal region by means of so-called transfer factors which are calculated by dividing the signal lepton efficiency by the efficiency of the relaxed or inverted selection requirements. The transfer factors are determined both in data and Monte Carlo simulation. The difference of the two measurements is accounted for as a systematic uncertainty on the transfer factors.

Non-prompt or misidentified leptons have usually a low transverse momenta. Thus, they are mostly assigned to the subleading lepton pair representing the off-shell Z boson candidate. The sources of such leptons are different for electrons and muons. Therefore, the estimation of the background contribution is performed separately for events with two different flavour of the subleading lepton pair, i.e. $\ell\ell + ee$ and $\ell\ell + \mu\mu$ events. The $\ell\ell + \mu\mu$ background events originate form decays of heavy-flavour hadrons, which are present in Z+jets and $t\bar{t}$ production. The muons from such decays are surrounded by a hadron jet. The $\ell\ell + ee$ background mainly originates from light-flavour jets (u, d, s) and gluon jets misidentified as electrons. The background estimation methods for the $\ell\ell + ee$ and $\ell\ell + \mu\mu$ classes of reducible background are described in the following.

4.4.2.1 Reducible $\ell\ell + ee$ Background

The electrons in the $\ell\ell + ee$ background from Z+jets, $t\bar{t}$ and WZ processes can be classified into three different types according to the source of the background electron. The largest fraction are misidentified electrons from light-flavour jets with energy deposits in the calorimeter, referred as fake electrons (f). Smaller contributions come from photon conversions (γ) and electrons from semi-leptonic decays of heavy-flavour quarks (q). The background contribution of the first two types of electrons is estimated with the method described below, while the latter is taken from simulation.

The $\ell\ell + ee$ background contribution is estimated in the $3\ell + e$ *control region*. The leading three leptons have to pass the lepton selection criteria of the full inclusive analysis described in Section 4.3. The selection and identification criteria are relaxed only for the lowest E_T electron from the subleading lepton pair. Instead of the full identification

criteria only the basic track quality criteria are applied for that electron, i.e. requirements on the number total number of silicon and pixel hits. In addition, no isolation requirements are applied. All selected quadruplets from the control region have to satisfy the common vertex criteria. To suppress the contribution from heavy-flavour hadron decays in this control region the d_0 significance requirement is used. The contribution from the ZZ^* production is suppressed by requiring that the electrons in the subleading lepton pair have a same-sign charge instead of the opposite-sign. The contribution of the ZZ^* background after this requirement is still 5-10% and has to be subtracted from the total yield in the control region before extrapolating to the signal region. The subtraction is based on simulated ZZ^* events. The criteria defining the $3\ell + e$ control region are summarised in Table 4.4. The resulting control region after applying the above requirements, is enriched in f and γ background components which can therefore be constrained by a fit to data. The small contribution of q is estimated from simulation.

Table 4.4: Selection criteria defining the control region for the estimation of the reducible $\ell\ell + ee$ background in the $H \to ZZ^* \to 4\ell$ analysis. SF SS denotes a same-flavour same-sign lepton pair, the check mark (✓) indicates that a given lepton selection requirement is applied and the cross (✗) that this requirement is not applied.

Control region	Vertex selection	m_{34}	Requirements on the electron e		
			Identification	d_0 sig. cut	Isolation
$3\ell + e$	✓	SF SS	relaxed	✓	✗
$2\ell + e$	✗	-	relaxed	✓	✗

The background yields in the $3\ell + e$ control region are constrained by data using a template fit based on the distribution of the number of hits $n_{InnerPix}$ in the insertable b-layer of the pixel detector. Beyond the coverage of the insertable b-layer the number of hits in the next-to-innermost pixel layer is fitted instead. This observable discriminates between the $f + q$ and γ background components, since photons leave no hits in the insertable b-layer, while $f + q$ electron candidates leave at least one hit. The ZZ^* contribution from simulation is taken into account in the fit and finally subtracted from the fitted yield.

The templates for the fit are taken from Monte Carlos simulation in a separate $2\ell + e$ *control region* (see Table 4.4) with a larger amount of data. In this region, only one electron is required in addition to the leading lepton pair. The additional electron candidate has to satisfy the same criteria as in the $3\ell + e$ control region. Only the common vertex requirement is not imposed, since there are less than four leptons in the final state.

A separate template from the $2\ell + e$ control region is used for each of the background components (f, γ and q). Since the $n_{InnerPix}$ distribution is similar for electrons from q and ZZ^* background a combined template is used whose normalisation is determined from simulation and set constant in the fit. The contribution of each of the background sources (f and γ) is obtained by fitting the $n_{InnerPix}$ distribution to the data in bins of the electron p_T and number of jets in the final state. In this way, the fit assigns to each a probability of being an f or γ background component. The final yield for each background component in the $3\ell + e$ control region is obtained by summing up these probabilities for all events.

The transfer factor to extrapolate the yields of the f and γ background component from the $3\ell + e$ control region to the signal region corresponds to the efficiency that the f or γ electron pass the nominal selection criteria (full electron identification and isolation criteria), relative to the efficiency of passing relaxed criteria form Table 4.4. These efficiencies are calculated from simulation in the $2\ell + e$ control region separately for f and γ components as a function of p_T and the number of jets in the final state. An additional correction factor is applied on these measurements to take into account the differences to the efficiency measurement in data. The correction factors are calculated in data in f enriched ($n_{InnerPix} > 0$) and γ enriched ($n_{InnerPix} = 0$) samples only in bins of electron p_T since the data show only a little dependence on the number of jets in the final state. The number of background events in the signal region for the f and γ background is therefore given as

$$N_{SR} = \sum_i c_i \sum_j \epsilon_{ij} w_{ij}, \tag{4.1}$$

where the index i runs over the p_T-bins and j over the N_j-bins. c corresponds to the correction factor estimated in data, ϵ is the electron efficiency measured in simulation and w is the probability being f or γ background obtained from the fit.

The results of the $\ell\ell + ee$ background estimation for all analysed data sets are shown in Table 4.5. The yields in the control and signal regions are shown together with the estimated ZZ^* background and the transfer factors. The transfer factors are in the same order for all analysed data sets.

4.4.2.2 Reducible $\ell\ell + \mu\mu$ Background

The reducible $\ell\ell + \mu\mu$ background with non-prompt muons originating from the Z boson production with additional muons from semi-leptonic decays of heavy-flavour hadrons (Z+HF) and from the Z boson production accompanied by muons from in-flight decays of π/K in light jets (Z+LF). Additional contributions come from the $t\bar{t}$ and WZ production. The Z+HF, Z+LF and the $t\bar{t}$ background are estimated from control data, while the WZ contribution is predicted by Monte Carlo simulation.

Table 4.5: Results of the reducible $\ell\ell + ee$ background estimation for the inclusive $H \to ZZ^* \to 4\ell$ analyses with the different analysed data sets taken from 2015-2016 (36.1 fb^{-1}), 2015-2017 (79.8 fb^{-1}) and 2015-2018 (139 fb^{-1}). The quoted uncertainty corresponds to the combined statistical and systematic uncertainty.

Background type	Total event yield in $3\ell + e$ control region	ZZ^*	Transfer factor	$\ell\ell + ee$ event yield in the signal region
		2015-2016 data set (36.1 fb^{-1})		
f	3075 ± 56	280 ± 6	0.0020 ± 0.0004	5.68 ± 1.24
γ	208 ± 17	19.4 ± 0.5	0.0071 ± 0.0014	1.34 ± 0.44
q	Monte Carlo based estimation			6.34 ± 1.93
		2015-2017 data set (79.8 fb^{-1})		
f	6620 ± 83	567 ± 9	0.0021 ± 0.0005	12.32 ± 1.84
γ	480 ± 27	49.9 ± 1.2	0.0070 ± 0.0025	3.02 ± 0.75
q	Monte Carlo based estimation			14.09 ± 3.03
		2015-2018 data set (139 fb^{-1})		
f	10352 ± 104	1455 ± 7	0.0017 ± 0.0004	15.6 ± 3.55
γ	753 ± 34	122 ± 0.84	0.0068 ± 0.0014	4.31 ± 1.18
q	Monte Carlo based estimation			12.15 ± 3.65

There are four dedicated control regions (CR) enhanced in different types of $\ell\ell + \mu\mu$ background and depleted in signal are used for the estimation. Table 4.6 summarises the requirements defining different control regions.

1. Inverted d_0 CR:
This control region is enhanced in Z+HF and in $t\bar{t}$ processes. The leading lepton pair is required to pass the full signal lepton selection criteria, while no common vertex requirement is applied. Leptons from heavy-flavour hadron decays in Z+HF and $t\bar{t}$ processes are characterised by a large d_0 significance. This requirement is therefore inverted for at least one of the leptons in the subleading lepton pair. In addition, no isolation criteria are applied on the leptons in the subleading pair.

2. $e\mu + \mu\mu$ CR:
The *$e\mu + \mu\mu$ CR* is enriched in $t\bar{t}$ production. In order to ensure that the leading lepton pair does not originate from a Z boson decay, the leading lepton pair is required to have

two opposite-charge different-flavour leptons ($e\mu$) passing the full signal lepton selection of the inclusive analysis. The common vertex cut is not applied. The d_0 significance cut and the isolation criteria are not applied on the two leptons from the subleading lepton pair ($\mu\mu$). For the latter, both same-charge and opposite-charge lepton pairs are selected.

3. Inverted isolation CR:
The third control region is employed to constrain the Z+LF component of the reducible $\ell\ell+\mu\mu$ background. The full lepton selection criteria are applied on the leptons from the leading lepton pair including the common vertex requirement on the lepton quadruplet to reduce the contributions from heavy-flavour decays. At least one of the leptons in the subleading pair is not allowed to pass the isolation criteria, while the d_0 significance requirement has to be satisfied.

4. Same-sign (SS) CR:
The same-sign (SS) control region is depleted in signal, while enhanced in all rather than a specific background component. The standard four-lepton event selection is applied on leptons from the leading lepton pair. The leptons of the second pair in the quadruplet are required to have the same-sign charge, while no requirements are applied on the d_0 significance and the isolation criteria.

Table 4.6: Requirements defining the control (CR) and validation (VR) regions for the estimation of the reducible $\ell\ell+\mu\mu$ background in the $H \to ZZ^* \to 4\ell$ analysis. The check mark (✓) and crosses (✗) indicate whether a specific requirement is applied or not. SF OS, SF SS and OF OS, stand for the lepton pairs with same-flavour leptons with opposite-sign charge, same-flavour leptons with same-sign charge and opposite-flavour leptons with opposite-sign charge, respectively.

Control region	Vertex selection	m_{12}	XX requirements		
			m_{34}	d_0 sig. cut	Isolation
Inverted d_0 CR	✗	–	–	inverted	✗
$e\mu+\mu\mu$ CR	✗	OF OS	SF OS + SF SS	✗	✗
Inverted isolation CR	✓	–	–	✓	inverted
SS CR	✗	–	SF SS	✗	✗
Relaxed VR	✗	–	–	✗	✗

In order to verify the background normalisation obtained by a global fit to the above four control regions, a fifth validation region with larger amount of data (*Relaxed VR*) is used (see Table 4.6). This validation region is not orthogonal to the other control regions nor to the signal region and contains all three types of backgrounds. The quadruplet is required to satisfy the signal event selection, except of the common vertex cut. The isolation and d_0 significance criteria are not applied on the leptons in the subleading lepton pair.

The discriminating variable used for the fit in all orthogonal control regions is the distribution of the invariant mass m_{12} of the leading lepton pair. It provides a good separation between the Z+jets and $t\bar{t}$ production. In the *Inverted isolation CR* and *SS CR* an additional discriminant is introduced in order to enhance the Z+LF component to improve its estimate. As stated above, the Z+LF contribution originates from Z boson production accompanied by muons from in-flight decays of π/K in light jets. Muons from in-flight decays are detected only in the muon spectrometer, while the original π or K is detected in the inner detector. Therefore, the corresponding muon momentum balance defined as

$$\frac{\Delta p_T}{p_T} = \frac{p_T^{ID} - p_T^{MS}}{p_T^{ID}}, \tag{4.2}$$

where $p_T^{ID(MS)}$ is the transverse momentum measured in the inner detector (muon spectrometer), is expected to be on average larger than zero, because the π or K is decaying only after the inner detector. In order to enhance the Z+LF background at least one of the muons from the subleading lepton pair has to satisfy $\Delta p_T/p_T > 0.2$.

In the combined fit, the m_{12} distribution for each of the background components is modelled with an analytical function. The non-resonant $t\bar{t}$ contribution is described with a second order Chebyshev polynomial in all control regions. Except for the $e\mu + \mu\mu$ *CR*, the Z+jets background is described with a Breit-Wigner function convoluted with a Crystal Ball function in order to model the resonant shape. In the $e\mu + \mu\mu$ *CR* the leading lepton pair does not originate from the Z decay, therefore a first order polynomial is used instead. For the diboson (WZ and ZZ^*) and Higgs boson processes the shape of the m_{12} distribution is assumed to be the same as for Z+jets in every control region, while the normalisation is taken from simulation. For the 139 fb^{-1} analysis also the tail of the diboson contribution is modelled, while it is not taken into account in the analyses performed with the 36.1 fb^{-1} and 79.8 fb^{-1} data sets. The analysis done on the 79.8 fb^{-1} and 139 fb^{-1} data sets take also into account the ttV and VVV contributions, which are combined with the diboson sample with the prediction of the Monte Carlo simulation.

For the 79.8 fb^{-1} and 139 fb^{-1} analyses, the expected yields are obtained by a simultaneous fit of all four orthogonal control regions. In the early Run 2 analysis difficulties have

occurred in simulating the Z+LF production. A generator level filter was applied on the Z+jets sample, which lead to a large Z+HF contribution and less Z+LF. Therefore, a slightly different approach is used for the 36.1 fb^{-1}analysis. The Z+HF and $t\bar{t}$ contributions are first estimated from the global fit only in the *Inverted d_0 CR* and $e\mu + \mu\mu$ *CR*. The Z+LF yield is retrieved from the fit in the *Inverted isolation CR* by fixing the Z+HF and $t\bar{t}$ contribution to the values from the global fit. The requirement on the muon momentum balance is not applied.

The normalisation obtained from the global fit is tested with data in the *Relaxed VR* and extrapolated to the signal region using simulation-based transfer factors. The transfer factors are defined for each background component as the ratio of the corresponding yields in the signal region and *Relaxed VR*. Due to a small number of simulated Z+LF events that satisfy the selection criteria the transfer factors for the Z+LF component are taken from the Z+HF sample in case of the 36.1 fb^{-1}and 79.8 fb^{-1} analyses. This is valid, because both components have similar isolation cut efficiencies. To take into account differences observed between data and simulation for the isolation and d_0 significance distributions in the 79.8 fb^{-1} analysis, the transfer factor is doubled for that analysis. For the analysis with the 139 fb^{-1} data set, the transfer factor for the Z+LF contribution is delivered from data in the same way as for the simulation.

In addition to the statistical uncertainty of transfer factors due to limited size of the Monte Carlo samples, the differences between data and simulation are assigned as a systematic uncertainty. The latter is delivered in a separate control region with only one additional muon in the final state instead of two, $\ell\ell + \mu$, where the leading lepton pair has to pass the standard four-lepton event selection. This control region is also used to derive uncertainties for the Z+HF transfer factor. The Z+HF and Z+LF component are separated by a requirement on the muon momentum balance.

The results of the $\ell\ell + \mu\mu$ background estimation for all analysed data sets are shown in Table 4.7. The yields in the control and signal regions are shown for the different background contributions together with the corresponding transfer factors. The transfer factors for the $t\bar{t}$ production are similar for all data sets. The transfer factor for Z+HF is the same for 79.8 fb^{-1}and 139 fb^{-1} data sets, while it is larger in the 36.1 fb^{-1} data set. The transfer factor for Z+LF is different in every data sets, because of the changes in the simulation for each data set.

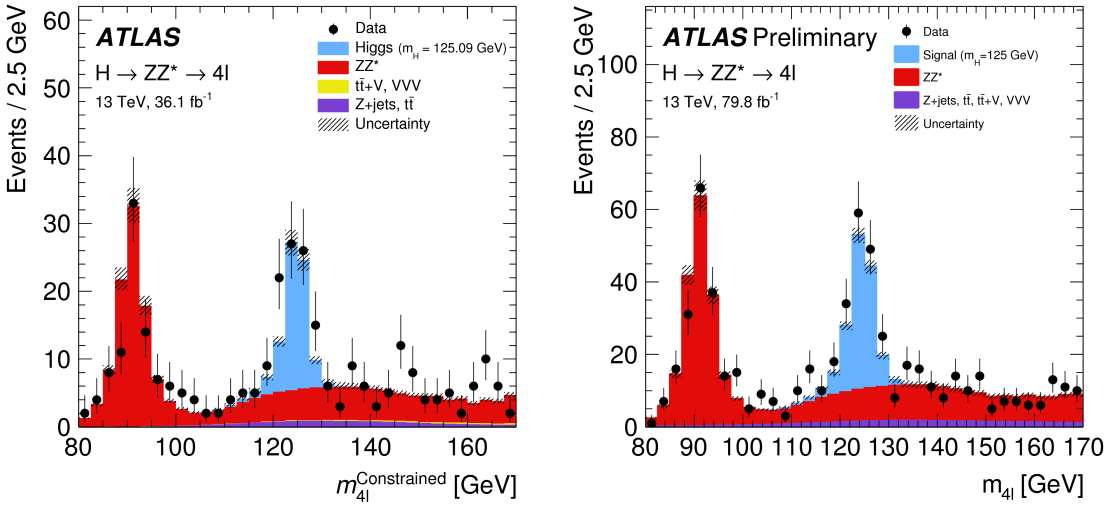
Table 4.7: Result of the reducible $\ell\ell + \mu\mu$ background estimation for the inclusive $H \to ZZ^* \to 4\ell$ analyses. The different analysed data sets are shown: 2015-2016 (36.1 fb^{-1}), 2015-2017 (79.8 fb^{-1}) and 2015-2018 (139 fb^{-1}). The total uncertainty corresponds to the combined statistical and systematic uncertainty.

Background type	Event yield in *Relaxed VR*	Transfer factor	Event yield in signal region
	2015-2016 data set (36.1 fb^{-1})		
$t\bar{t}$	917.57 ± 22.92	0.0025 ± 0.0003	2.29 ± 0.26
Z+HF	908.38 ± 52.37	0.0075 ± 0.0009	6.82 ± 0.90
Z+LF	50.10 ± 21.31	0.0075 ± 0.0075	0.38 ± 0.41
WZ	Monte Carlo based estimation		0.91 ± 0.50
	2015-2017 data set (79.8 fb^{-1})		
$t\bar{t}$	1878.9 ± 34.5	0.0027 ± 0.0004	5.00 ± 0.65
Z+HF	1624.6 ± 75.4	0.0043 ± 0.0005	7.01 ± 0.62
Z+LF	149.4 ± 40.2	0.0086 ± 0.0044	1.31 ± 0.58
WZ	Monte Carlo based estimation		2.14 ± 0.20
	2015-2018 data set (139 fb^{-1})		
$t\bar{t}$	3118 ± 47	0.0025 ± 0.0003	7.65 ± 0.85
Z+HF	2905 ± 113	0.0043 ± 0.0004	12.47 ± 1.27
Z+LF	299 ± 68	0.0100 ± 0.0010	3.10 ± 0.75
WZ	Monte Carlo based estimation		4.52 ± 0.32

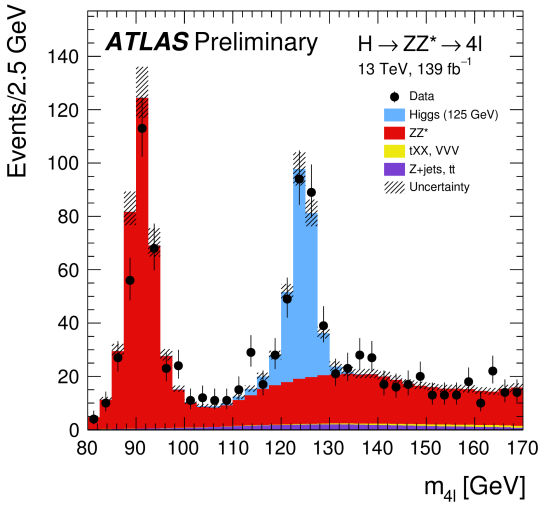
4.5 Results

The measured and expected four-lepton invariant mass ($m_{4\ell}$) distributions after the inclusive event selection are shown in Figure 4.5 for the three data sets corresponding to 36.1 fb^{-1}, 79.8 fb^{-1} and 139 fb^{-1}. The Higgs boson resonance is clearly visible at a mass of about 125 GeV. The peak at the Z boson resonance mass of 91 GeV corresponds to resonant $q\bar{q} \to Z/\gamma^* \to 4\ell$ production where the additional lepton pair is radiated from one of the two Z decay leptons as shown in Figure 4.2(b).

The measured and expected number of events are summarised in Table 4.8. For the 36.1 fb^{-1} data set a mass window of 118 GeV $< m_{4\ell} <$ 129 GeV and for the other two sets a wider window of 115 GeV $< m_{4\ell} <$ 130 GeV has been used. Therefore, the signal-to-background ratio is 2.3 for the former and 1.9 for the latter.



(a) (b)



(c)

Figure 4.5: The expected and observed four-lepton invariant mass distribution (a) for the 36.1 fb^{-1}, (b) 79.8 fb^{-1} and (c) 139 fb^{-1} data sets after applying the $H \to ZZ^* \to 4\ell$ event selection criteria of the inclusive analysis [182–184].

In the 36.1 fb^{-1} and 79.8 fb^{-1}data sets a slight excess of Higgs boson candidates with respected to the SM prediction is observed. This excess disappears for the full Run 2 data set which is in very good agreement with the SM prediction. The measured signal strength parameters μ, defined as the measured Higgs boson production rate in the 4ℓ decay channel relative to the SM prediction are $1.28^{+0.21}_{-0.19}$ [182] for the 36.1 fb^{-1} data set, $1.19^{+0.16}_{-0.15}$ [183] for the 79.8 fb^{-1} data set and $1.04^{+0.12}_{-0.10}$ [184] for the full Run 2 data set corresponding to 139 fb^{-1}. All measurements are in agreement with the the SM prediction.

Table 4.8: Expected and observed number of events after the inclusive $H \to ZZ^* \to 4\ell$ selection for the data sets corresponding to integrated luminosities of 36.1 fb^{-1}, 79.8 fb^{-1}and 139 fb^{-1}in the 4ℓ invariant mass ranges of 118 GeV $< m_{4\ell} <$ 129 GeV (36.1 fb^{-1}) and 115 GeV $< m_{4\ell} <$ 130 GeV (79.8 and 139 fb^{-1}) [182–184]. Statistical and systematic uncertainties have been added in quadrature.

	SM Higgs boson signal	ZZ^*	Other backgrounds	Total expected	Observed
2015-2016 data set (36.1 fb^{-1}), 118 GeV $< m_H <$ 129 GeV					
4μ	19.7 ± 1.6	7.5 ± 0.6	1.00 ± 0.21	28.1 ± 1.7	32
$2e2\mu$	13.5 ± 1.0	5.4 ± 0.4	0.78 ± 0.17	19.7 ± 1.1	30
$2\mu2e$	10.4 ± 1.0	3.57 ± 0.35	1.09 ± 0.19	15.1 ± 1.0	18
$4e$	9.9 ± 1.0	3.35 ± 0.32	1.01 ± 0.17	14.3 ± 1.0	15
Total	54 ± 4	19.7 ± 1.5	3.9 ± 0.5	77 ± 4	95
2015-2017 data set (79.8 fb^{-1}), 115 GeV $< m_H <$ 130 GeV					
4μ	40.5 ± 1.7	19.0 ± 1.1	1.71 ± 0.10	61.2 ± 2.0	64
$2e2\mu$	28.2 ± 1.2	13.3 ± 0.8	1.38 ± 0.10	42.8 ± 1.4	64
$2\mu2e$	22.1 ± 1.4	9.2 ± 0.9	2.99 ± 0.09	34.3 ± 1.7	39
$4e$	21.1 ± 1.4	8.6 ± 0.8	2.90 ± 0.09	32.5 ± 1.6	28
Total	112 ± 5	50 ± 4	8.96 ± 0.12	171 ± 6	195
2015-2018 data set (139 fb^{-1}), 115 GeV $< m_H <$ 130 GeV					
4μ	78 ± 5	38.1 ± 2.2	2.87 ± 0.18	119 ± 5	118
$2e2\mu$	52.8 ± 3.1	26.1 ± 1.4	3.01 ± 0.19	81.9 ± 3.4	98
$2\mu2e$	40.0 ± 2.9	17.4 ± 1.3	3.5 ± 0.5	60.9 ± 3.2	57
$4e$	35.3 ± 2.6	15.1 ± 1.5	2.9 ± 0.4	53.3 ± 3.1	43
Total	206 ± 13	97 ± 6	12.3 ± 0.9	315 ± 14	316

5

Measurement of Higgs Boson Production Cross Sections

In this chapter, the measurements of the Higgs boson production cross section in several exclusive phase space regions is described. The measurements have been performed for all data sets corresponding to integrated luminosities of 36.1 fb^{-1}, 79.8 fb^{-1} and 139 fb^{-1} as specified in Section 4.2.1. The input for the analyses are the Higgs boson candidates selected by the inclusive $H \to ZZ^* \to 4\ell$ analysis as described in Section 4.5.

The focus of this thesis was especially the evaluation of the theoretical uncertainties for the main Higgs boson production modes. For the 139 fb^{-1} data set, also the parton shower uncertainties in the $qqZZ$ background estimation were evaluated. In addition, dedicated multivariate discriminants have been designed for the event categorisation in the analysis of the 79.8 fb^{-1} data set.

The measured cross sections in this chapter provide input to the measurements presented in the next chapters. The extrapolation of the Higgs boson production cross section measurements to predictions for high-luminosity LHC, described in detail in Chapter 6, is based on the analysis of the 79.8 fb^{-1} data set, while the two EFT interpretations in Chapter 7 are based on results of the 36.1 fb^{-1} and 139 fb^{-1} analyses. The analyses of the three data sets differ slightly, but only the 79.8 fb^{-1} analysis is described in detail. The differences for other data sets are described in the relevant subsequent chapters.

This chapter is structured as follows. First the exclusive phase space regions for the production cross section measurement are described, followed by the criteria for the categorisation of the reconstructed Higgs boson candidates and the introduction of discriminants used in different categories to distinguish between the different signal production modes and between signal and background processes. The categories and discriminants are chosen such as to provide the optimal sensitivity based on the production phase space bins defined at particle level in Section 5.1 in the simplified template cross sections (STXS). The statistical model and the results of the measurement are outlined after describing the evaluation of the systematic uncertainties.

5.1 Dedicated Regions for the Measurement of Production Cross Sections

In the Run 1 analysis, the Higgs boson couplings to SM particles have been determined from measurements of the signal strength $\mu_i^f = (\sigma_i \cdot \mathcal{B}_f)/(\sigma_i \cdot \mathcal{B}_f)_{\mathrm{SM}}$ in the various production and decay modes by using multiplicative coupling modifiers κ [44]. In order to reduce the dependence on uncertainties in the SM predictions and to allow for a combination of the different production and decay channels, the so-called Simplified Template Cross Section (STXS) framework has been introduced which is described below. In the STXS framework cross section measurements in dedicated regions of the phase space are implemented for the coupling measurements. In addition it simplifies the use of analysis techniques like event categorisation and multivariate selection methods which improve the sensitivity.

The dedicated phase space regions in which the cross sections are measured are called "production bins" are defined at particle level for each production mode. The definition of the bins is guided by several aspects. The first is to minimise the theory dependence on uncertainties by avoiding extrapolation of the measurement from a certain phase space region to the full phase space. The bins are chosen such that there are only small variations of the experimental acceptance within a bin. Another aspect is the optimisation of the sensitivity of the coupling measurement by combining all decay channels and by using dedicated event categories. A final goal is the isolation of BSM effects in certain phase space bins where they are relatively large. At the same time the number of bins must not become to large to limit the complexity.

The simplest implementation, called STXS Stage-0, the bins only separate the different production modes of the Higgs boson. At the next stage, STXS Stage-1, the production bins are further divided into exclusive phase space regions. In practice not all analyses are able to set constraints on all defined particle level production bins at STXS Stage-1 due to statistical limitations, as it is the case in the $H \to ZZ^* \to 4\ell$ decay channel. In these cases, several particle level production bins are merged in order to achieve the necessary statistics. The merged scheme is called STXS Reduced-Stage-1.

To allow for comparisons between the measured cross section in the particle level production bins and theoretical predictions from analytic calculations or Monte Carlo simulations, a well-defined definition of the particle level final state is necessary. The final state particles are leptons, jets and the Higgs boson. Electrons and muons are defined as dressed leptons, i.e. the definition includes photons around charged leptons. The τ leptons are defined as the sum of all decay products. No kinematic constraints and no thresholds on the phase space are employed. Particle level jets are defined as described

in Section 3.4.3 and the transverse momentum p_T^j is required to be larger than 30 GeV. The Higgs boson is defined as off-shell particle and required to be within $|y_H| < 2.5$.

5.1.1 Stage-0 Regions

The particle level production bins of STXS Stage-0 (see Figure 5.1) separate the different production modes of the Higgs boson. The VBF and the VH production, with hadronically vector boson decays (VH-Had) are denoted as electroweak qqH (EW qqH) production. The VH production with leptonically decaying vector boson (VH-Lep) is split into three bins distinguishing the production mechanism and the final state vector boson, $qq \to WH$, $qq \to ZH$ and $gg \to ZH$.

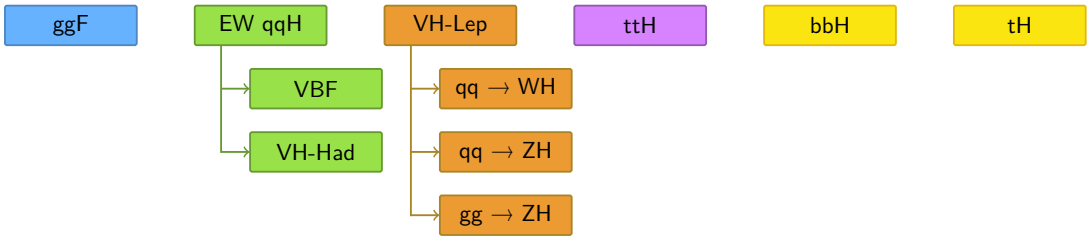


Figure 5.1: Particle level production bins of STXS Stage-0.

5.1.2 Stage-1 Regions

Figure 5.2 shows the particle level production bins of STXS Stage-1 for ggF production. The first classification is according the number of jets N_j in the final state: $N_j = 0$, $N_j = 1$, $N_j \geq 2$ and $N_j \geq 2$ with VBF cuts, motivated by the N_j splitting in experimental analyses. The $N_j = 0$ bin is not divided further, while the $N_j = 1$ and $N_j \geq 2$ bins are split according to the transverse momentum of the Higgs boson p_T^H into a low-p_T^H bin ($p_T^H < 60$ GeV), two medium bins (60 GeV $< p_T^H <$ 120 GeV and 60 GeV $< p_T^H <$ 120 GeV) and a BSM sensitive bin ($p_T^H > 200$ GeV). The $N_j \geq 2$ bin with VBF cuts is used to estimate the gluon fusion contamination in the VBF selection. The cuts are chosen to target the VBF signature (two high-energetic jets in the forward region). A large invariant mass ($m_{jj} > 400$ GeV) and a good separation ($\Delta\eta_{jj} > 2.8$) of the two leading jets is required. This category is further divided according to p_T^{Hjj} with a threshold of 25 GeV.

Figure 5.3 shows the particle level production bins of STXS Stage-1 for EW qqH (VBF and VH-Had) production. BSM effects are separated with cut on the transverse momentum of the leading jet, $p_T^{j1} > 200$ GeV. Events with VBF topology ($N_j \geq 2$, $m_{jj} > 400$ GeV and $\Delta\eta_{jj} > 2.8$) are divided according to p_T^{Hjj} with a threshold of 25 GeV, providing a good separation between ggF and VBF production. VH-Had production is targeted by requiring at least two jets in the final state with an invariant mass of 60 GeV $< m_{jj} <$ 120 GeV. Remaining events are assigned to the so-called "Rest" bin.

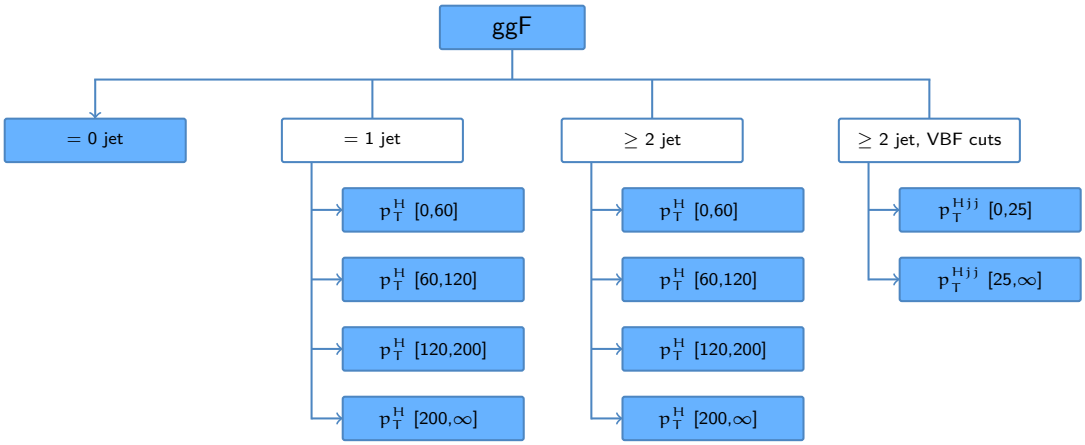


Figure 5.2: Particle level production bins of STXS Stage-1 for ggF production.

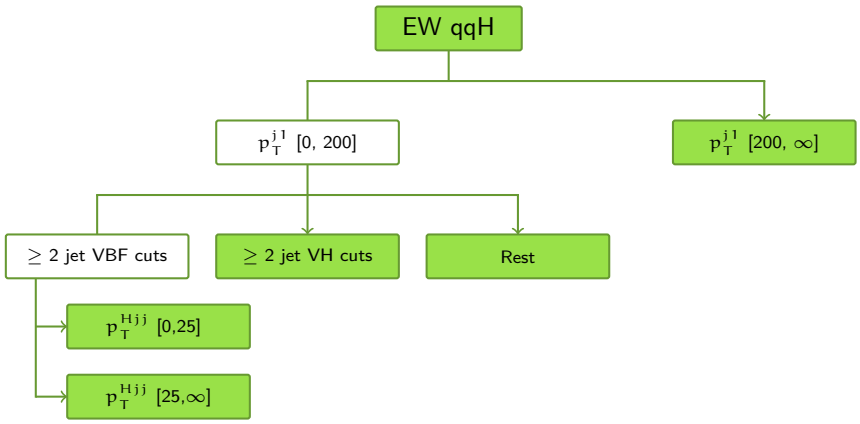


Figure 5.3: Particle level production bins of STXS Stage-1 for EW qqH production.

The STXS Stage-1 particle level production bins for VH-Lep production are shown in Figure 5.4. At first qq and gg initial states are separated. $qq \to VH$ processes are split according to the weak gauge boson in the final state, $W \to \ell\nu$ or $Z \to \ell\ell + \nu\nu$. These production bins are divided with thresholds on the transverse momentum of the weak boson p_T^V into a low ($p_T^V < 150$ GeV), a medium (150 GeV $< p_T^H <$ 20 GeV) and a high ($p_T^V > 250$ GeV) bin. The medium bin is split further into final states with $N_j = 0$ and $N_j \geq 1$. For $gg \to ZH$ processes the same splitting is used, except of the p_T^V-high bin which is merge into the medium bin.

The remaining production modes, ttH, bbH and tH are defined as in STXS Stage-0.

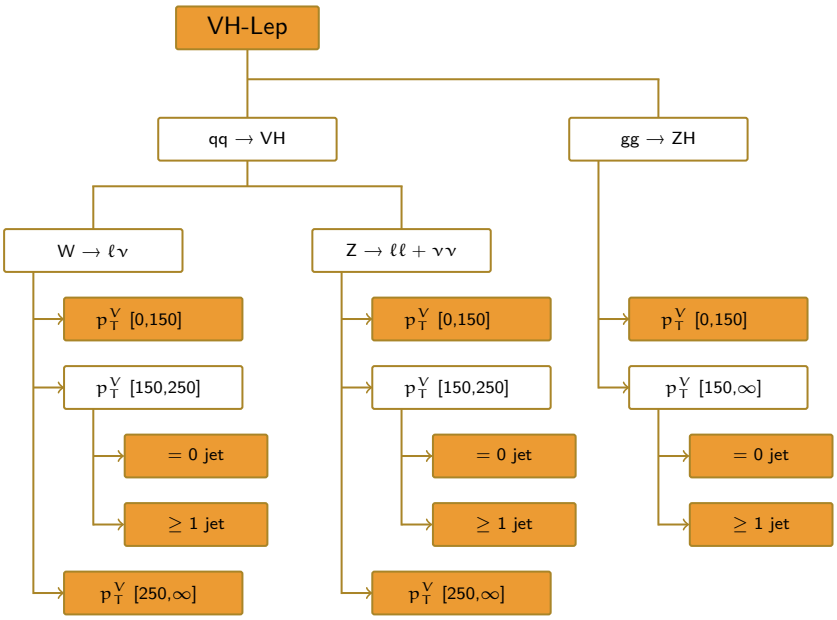


Figure 5.4: Particle level production bins of STXS Stage-1 for VH-Lep production.

5.1.3 Reduced-Stage-1 Regions

In the $H \to ZZ^* \to 4\ell$ decay channel the available amount of data does not allow for a measurement in all production bins defined at STXS Stage-1. Production bins are merged till a sufficient statistics is reached. The production bins are chosen to maximise the measurement precision and to isolate possible BSM contributions. The studies presented in this thesis use two different types of merged schemes. The production cross section measurement described in this chapter is performed within the so-called STXS Reduced-Stage-1 shown in Figure 5.5 and the EFT interpretation (Section 7.3) within the STXS Reduced-Stage-1.1 shown in Figure 5.6.

For the STXS Stage-0 production cross section measurement in $H \to ZZ^* \to 4\ell$ decays only four particle level production bins are used. The bbH production bin is not accessible with current statistics. Its contributions is measured together with the ggF production since both have a similar acceptance. Similarly, the tH production is considered together with the ttH production, referred as ttH in the following. In contrast to the STXS Stage-0 scheme introduced in Section 5.1.1 the VH-Had production is not combined with the VBF production. At Stage-0 it is merged with the VH-Lep production in a combined VH production bin.

In the STXS Reduced-Stage-1 five production bins for the ggF production are defined. The first split is according to the number of jets, $N_j = 0$, $N_j = 1$ and $N_j \geq 2$. The zero jet bin is denoted as ggF-0j. The one jet bin is divided further in bins of p_T^H, one low with $p_T^H < 60$ GeV (ggF-1j-p_T^H-Low), a medium with 60 GeV $< p_T^H <$ 120 GeV (ggF-1j-p_T^H-Med) and a high bin with $p_T^H > 120$ GeV (ggF-1j-p_T^H-High). The two jet bin is not split further and is denoted as ggF-2j. For the VBF production only the first split of STXS Stage-1 is used, the cut on the transverse momentum of the leading jet p_T^{j1}, separating SM and potential BSM contributions. The two bins are denoted as VBF-p_T^j-Low and VBF-p_T^j-High. The VH production is separated according to the hadronically or leptonically decays of the associated vector boson, denoted as VH-Had and VH-Lep, respectively. The ttH production mode remains the same as for STXS Stage-0.

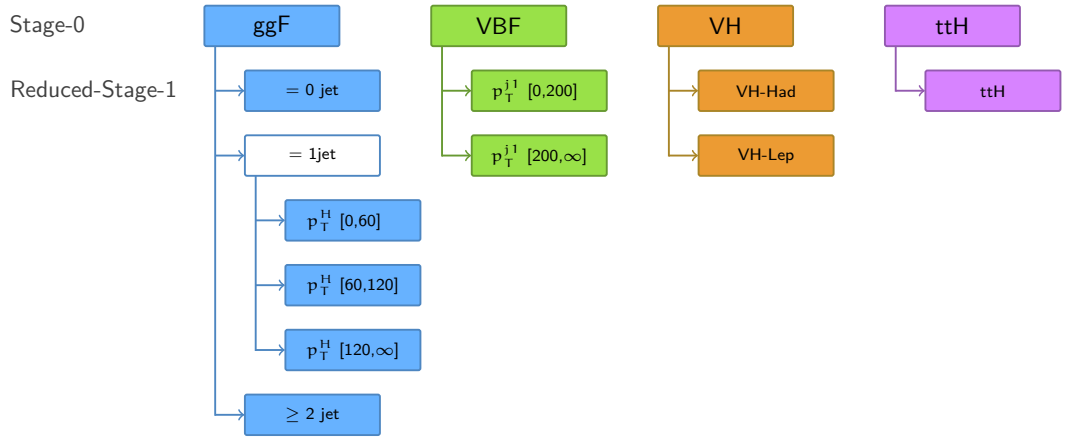


Figure 5.5: Particle level production bins of STXS Reduced-Stage-1.

The main difference of the STXS Reduced-Stage-1.1 compared to STXS Reduced-Stage-1 is that more production bins are defined due to the larger available amount of data. The $gg \to Z(2j) + H$ and ggF production are combined into the gg2H production bin, which is further divided into seven exclusive bins. First a high p_T^H bin is split targeting BSM physics with $p_T^H > 200$ GeV denoted as gg2H-p_T^H-High production bin. The remaining events are

split in a similar way as in the STXS Reduced-Stage-1, according to the number of jets ($N_j = 0$, $N_j = 1$ and $N_j \geq 2$) and the transverse momenta of the Higgs boson. The zero jet bin has an additional split at $p_T^H = 10$ GeV defining the gg2H-0j-p_T^H-Low and gg2H-0j-p_T^H-High production bin. One jet events are split with the same cuts on p_T^H into the gg2H-1j-p_T^H-Low, gg2H-1j-p_T^H-Med and gg2H-1j-p_T^H-High production bins. Final states with at least two jets are collected into the gg2H-2j production bin. In order to have a definition of productions bins closer to the Stage-1 described in Section 5.1.2 the VBF production mode is considered together with the VH-Had production. First the VH-Had production is separated with a cut on the invariant mass of the two leading jet, 60 GeV $< m_{jj} <$ 120 GeV defining the qq2Hqq-VH-Like bin. The production bin qq2Hqq-BSM sensitive to BSM contributions is defined as $m_{jj} > 350$ GeV and $p_T^H > 200$ GeV. The remaining events are collected in the qq2Hqq-VBF production bin. The VH-Lep (qq/gg2HLep) and the ttH production are not divided further.

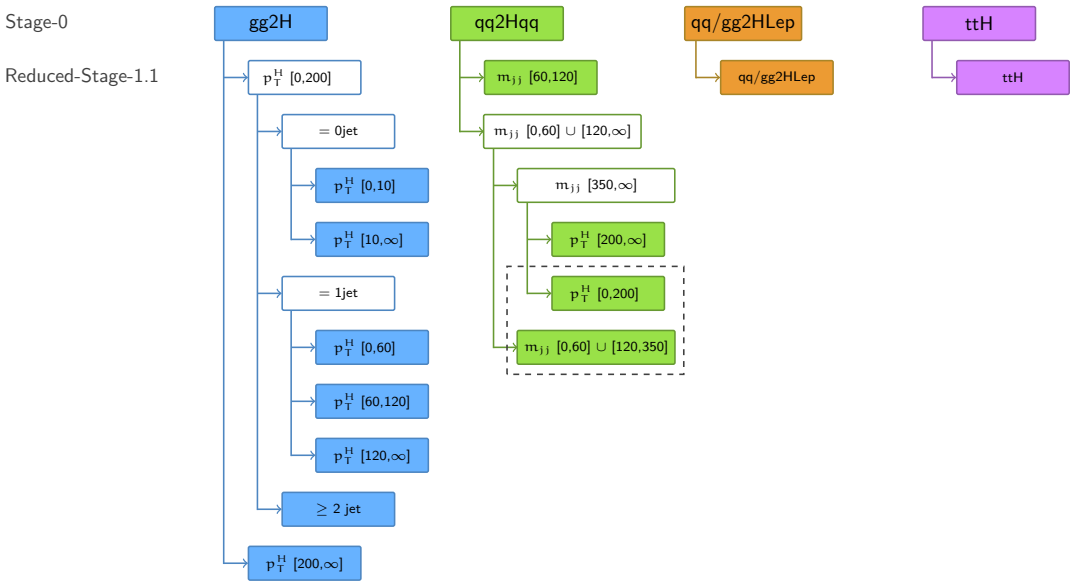


Figure 5.6: Particle level production bins of STXS Reduced-Stage-1.1.

5.2 Event Categorisation

The Higgs boson candidates selected by the inclusive $H \to ZZ^* \to 4\ell$ analysis in a mass range of 115 GeV $< m_{4\ell} <$ 130 GeV, as described in Section 4.3, are classified into nine categories in order to provide sensitivity to different Higgs boson production modes and more specifically to the more exclusive particle level production bins defined within the STXS framework. The production cross sections for each production mode are measured in the particle level production bins of STXS Stage-0. Additional exclusive regions of the production phase space are explored with the STXS Reduced-Stage-1 particle level production bins.

The ggF, VBF, VH and ttH production modes of the Higgs boson are disentangled according to their specific final state topology. The bbH production is measured together with the ggF production in the ggF production bin, since it is difficult to distinguish between the two.

The event classification starts with the selection of the ttH event candidates. The top quark almost always decays via the electroweak interaction into a W boson and a bottom quark. Therefore, at least one b-tagged jet is required in the final state. The selected event candidates are then distinguished according to the decay mode of the W boson. Events with fully hadronic decays of the both W bosons are collected in the *ttH-Had-enriched* category, requiring at least four additional jets in the final state. Events with at least one leptonic W boson decay are identified by requiring at least one additional lepton with $p_T > 12$ GeV and at least two jets in the final state and are collected in the *ttH-Lep-enriched* event category.

The VH production with leptonically decaying vector boson ($V \to \ell\nu/\ell\ell$) is also characterised by at least one additional lepton in the final state. Events with additional leptons in the final state which do not fulfil the jet requirements of the ttH selection are assigned to the *VH-Lep-enriched* category.

Events which do not fall into one of the categories described above are further sub-divided according to the number of jets in the final state. The VBF and the VH production with hadronically decaying vector boson (VH-Had) are targeted by requiring at least two jets in the final state. Even through there are no jets emitted for the tree-level ggF production, this production mode still significantly contaminates the 2-jet final state due to initial state radiation. The VBF and VH-Had production are separated by the requirement on the invariant mass m_{jj} of the two leading jets. The m_{jj} distribution from the VH-Had production peaks around the mass of the decaying vector boson, while for the VBF production larger values are expected due to the characteristic signature of two high-energy jets. Events with $m_{jj} < 120$ GeV are therefore collected in the *VH-Had-enriched* category. The events with $m_{jj} > 120$ GeV are further divided into two categories, *VBF-enriched-p_T^j-Low* and *VBF-enriched-p_T^j-High*, separated by the 200 GeV threshold on the transverse momentum p_T^{j1} of the highest p_T (leading) jet.

The ggF production mode is expected to be dominant in final states with 0-jet or 1-jet. In the 1-jet category, there is also a non-negligible contribution from the VBF production, processes in which one of the two jets is not reconstructed. Motivated by the STXS particle level production bins, the 1-jet events are split into three categories. Events with the four-lepton transverse momentum $p_T^{4\ell}$ smaller than 60 GeV are assigned to the *1j-$p_T^{4\ell}$-Low* category, while events with 60 GeV $< p_T^{4\ell} < 120$ GeV and $p_T^{4\ell} > 120$ GeV are collected in the *1j-$p_T^{4\ell}$-Med* and *1j-$p_T^{4\ell}$-High* categories, respectively.

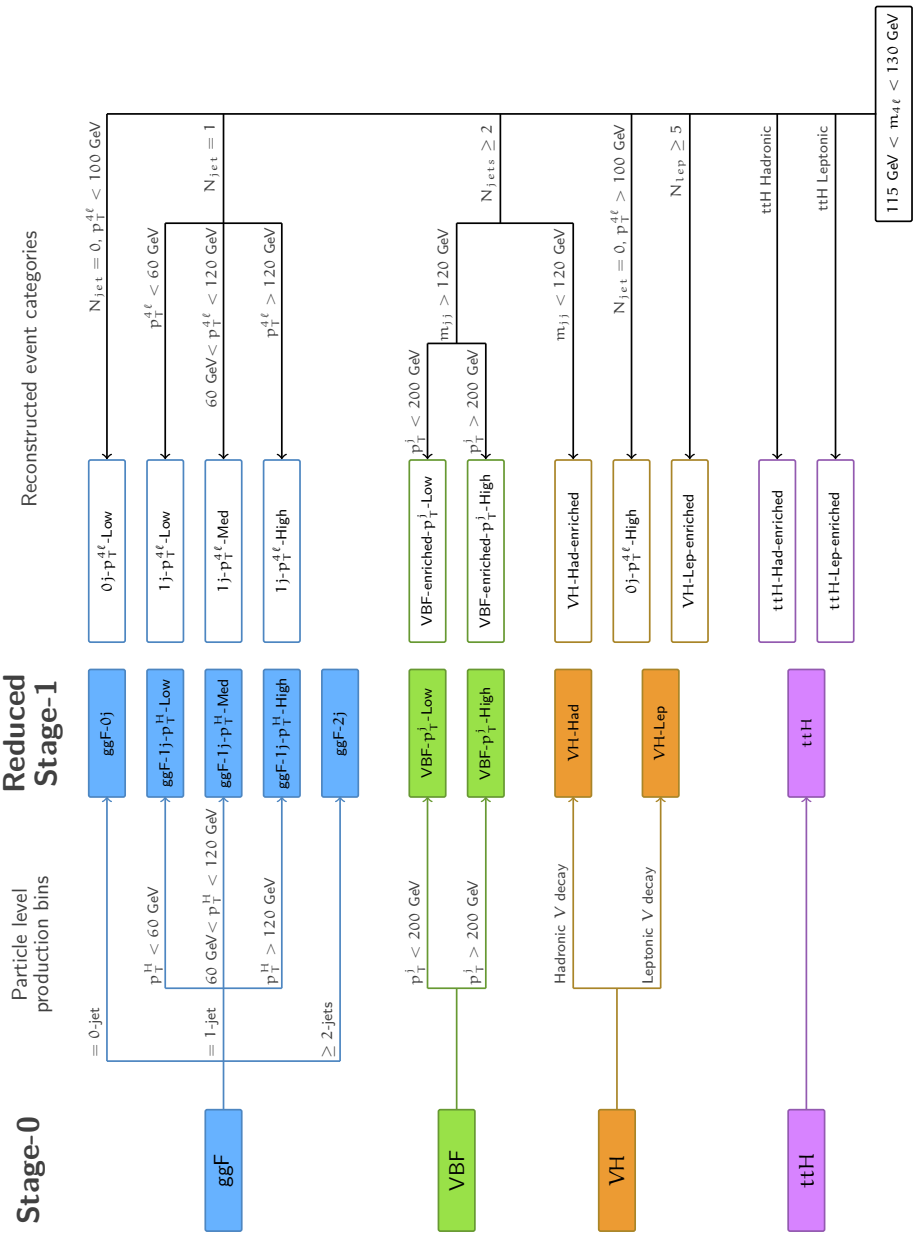


Figure 5.7: The exclusive regions of the Higgs boson production phase space (production bins) and the corresponding reconstructed event categories. The regions for the measurement of the Higgs boson production cross sections are defined at the particle level in two stages, STXS Stage-0 and Reduced-Stage-1.

The remaining event with no reconstructed jets in the final state are collected in the *0j-$p_T^{4\ell}$-Low* and *0j-$p_T^{4\ell}$-High* categories. A threshold on the $p_T^{4\ell}$ value is applied at 100 GeV to discriminate between the ggF and VH production with leptonic V boson decays.

The event categorisation scheme with the corresponding particle level STXS production bins is shown in Figure 5.7.

The expected number of events from the key production modes of the SM Higgs boson are shown in Table 5.1 for each of the reconstructed event categories. The expected yields for the ggF and bbH signal are shown separately, while in the fit of the Stage-0 prediction to data, the two contributions are combined into the same production bin (ggF). Jets produced in bbH processes tend to be more forward than the one produced in ttH, therefore escaping the acceptance of the ttH selection and contributing mostly to the $0j$-$p_T^{4\ell}$-Low category. The expectations for the WH and ZH contributions, as well as ttH and tH processes are also shown separately, while they are merged due to low statistics into a single VH and ttH production bin, respectively, when performing the Stage-0 measurement. The table includes the expected statistical and systematic uncertainties which are added in quadrature. The sources and the evaluation of the systematic uncertainties are discussed in detail in Section 5.4.

Table 5.1: The expected number of events from the production of the SM Higgs boson with a mass of $m_H = 125$ GeV via different production modes in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV and an integrated luminosity of 79.8 fb^{-1} separately for each of the reconstructed event categories. Statistical and systematic uncertainties are added in quadrature [183].

Reconstructed	SM Higgs boson production mode											
event category	ggF		VBF		WH		ZH		ttH		bbH		tH	
$0j$-$p_T^{4\ell}$-Low	54	±5	0.64	±0.12	0.213	±0.032	0.199	±0.030	–		0.56	±0.28	–	
$1j$-$p_T^{4\ell}$-Low	16.1	±2.2	1.05	±0.06	0.291	±0.035	0.173	±0.021	0.0017	±0.0010	0.23	±0.12	0.00140	±0.00030
$1j$-$p_T^{4\ell}$-Med	9.6	±1.5	1.38	±0.15	0.292	±0.033	0.194	±0.022	0.0018	±0.0011	0.049	±0.025	0.0021	±0.0004
$1j$-$p_T^{4\ell}$-High	2.4	±0.5	0.60	±0.07	0.115	±0.014	0.106	±0.013	0.0018	±0.0006	0.009	±0.004	0.0017	±0.0004
VBF-enriched-p_T^j-Low	7.8	±1.6	4.1	±0.4	0.35	±0.05	0.29	±0.04	0.124	±0.013	0.10	±0.05	0.055	±0.007
VBF-enriched-p_T^j-High	5.5	±1.1	0.43	±0.04	0.68	±0.07	0.52	±0.05	0.051	±0.008	0.053	±0.027	0.0169	±0.0022
VH-Had-enriched	0.70	±0.20	0.38	±0.04	0.062	±0.010	0.050	±0.008	0.038	±0.005	0.0014	±0.0007	0.0119	±0.0013
VH-Lep-enriched	0.030	±0.004	0.0084	±0.0004	0.44	±0.04	0.116	±0.011	0.083	±0.011	0.0028	±0.0014	0.0172	±0.0018
$0j$-$p_T^{4\ell}$-High	0.059	±0.022	0.0096	±0.0017	0.030	±0.004	0.085	±0.010	–		–		–	
ttH-Had-enriched	0.09	±0.09	0.020	±0.004	0.0130	±0.0027	0.028	±0.006	0.38	±0.04	0.012	±0.006	0.054	±0.006
ttH-Lep-enriched	–		–		0.0026	±0.0006	0.0018	±0.0004	0.212	±0.025	–		0.0204	±0.0022
Total	97	±8	8.6	±0.4	2.49	±0.25	1.76	±0.17	0.90	±0.09	1.0	±0.5	0.181	±0.020

5.3 Multivariate Production Mode Discriminants

The expected relative fractions of SM Higgs boson signal contributions from each production bin of the Reduced-Stage-1 are shown in Figure 5.8 separately for each reconstructed event category. Reconstructed event categories with high purity in a single production mode are the $0j$ and ttH-Lep-enriched, with more than 90% of the respective signal component. Also in the VH-Lep-enriched event category there is a large fraction of about 85% from the targeted VH-Lep production mode. In other categories, the fraction of the respective targeted signal mode is visibly smaller. For example, in the VBF-enriched reconstructed event categories with at least two jets in the final state, the fraction of the VBF component is only about 40%, with large contamination from ggF-$2j$ production. Therefore, in a number of categories there are further multivariate discriminants introduced to improve the separation of the different signal components and in case of the $0j$ categories, to distinguish the ZZ^* background from the ggF signal.

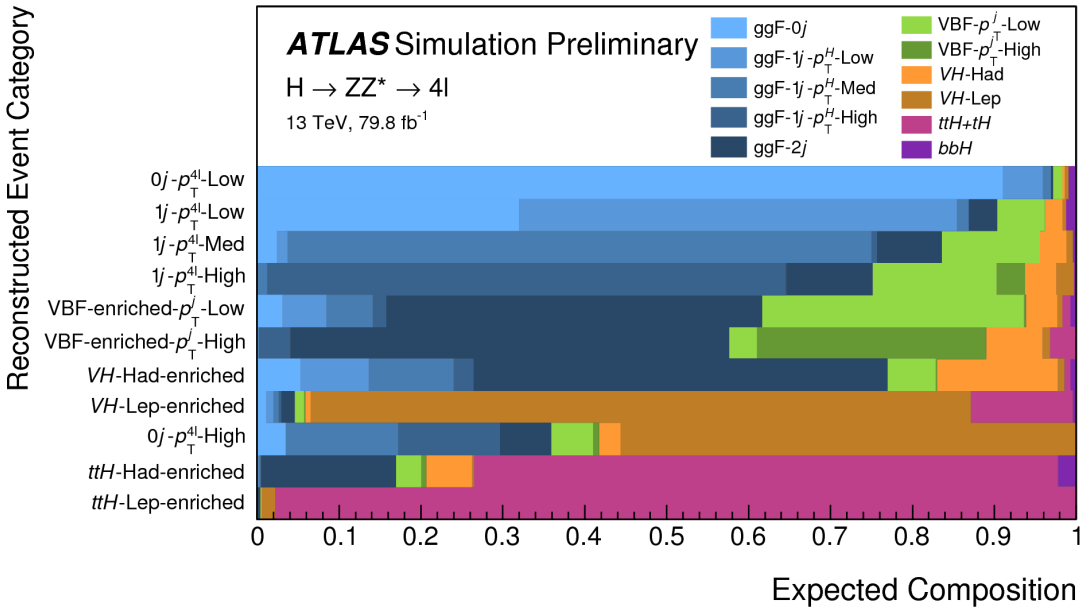


Figure 5.8: Expected signal composition of the SM Higgs boson signal for the Reduced-Stage-1 particle level production bins in each of the reconstructed event categories. The contribution of the ggF and bbH production are shown separately, while they both contribute to the same production bin (ggF) [183].

The discriminating observables based on boosted decision trees (BDT) [185] are developed with the Toolkit for Multivariate Data Analysis [186] with the ROOT framework [187]. Boosted decision trees are a machine learning method for classification of different processes. The classification is based on Monte Carlo events. The simulated samples are

split into two half, one for training the algorithm (training sample) and the other for testing its performance (test sample). The trained algorithm is then applied to the data.

A decision tree performs the classification by means of a tree-structured decision making. Starting from the topmost decision node (root node), the data set is split into two smaller one based on criteria imposed on an input variable. Subsequently, the two subsets are further split based on further input variables. The procedure is repeated until the algorithm reaches the final node (leave node) with which the trained events are classified into signal or background processes. In order to enhance the classification performance of such single trees, so-called boosting is employed by building several trees which form a forest. All trees are derived from the same training samples by reweighting events. Finally, the output of each tree is combined into a single classifier. The training procedure is described in more detail in the following.

The training of a decision tree starts with the root node at which an initial splitting criterion i.e. the input variables and the corresponding threshold values for the full training sample is determined. This split leads to two subsets of training samples. Each of these subsets is processed by the same algorithm to determine the next iteration of the splitting criteria. There are several types of splitting criteria that can be employed to provide the best separation between signal and background processes. For this analysis the default separation quantifier of the TMVA framework was used, the *Gini Index* defined as

$$G = p \cdot (1 - p), \tag{5.1}$$

where p corresponds to the purity of the training sample defined as the number of signal events in the samples divided by the total number of events. The *Gini Index* reaches a maximum value of $p = 0.5$ for a fully mixed sample with both signal and background contributions and is zero in case of only signal or only background contributions. The splitting criterion i.e. a threshold on a single input variable, is defined as the input variable that maximises the difference between the separation G of the parent node and the sum of the separations of the two daughter nodes weighted by their relative event fraction. The optimal threshold value is obtained by scanning over the input variable values with a specific scan granularity defined by the TMVA parameter `nCuts`. The performance of the classifier can in some cases be enhanced by using randomised trees (TMVA parameter `UseRandomisedTrees`), i.e. by considering only a random subset of all input variables for the splitting. The number of such random variables can be set with the `UseNvars` options. The iterative process of the node splitting ends once a minimum number of events is reached, defined by the parameter `nEventsMin`. The corresponding leaf node is then classified as signal or background depending on which type of events dominates. It is also possible to terminate the splitting process by setting the limit on the fraction of all training events that reach the leaf node (`MinNodeSize`) or by fixing the depth of the decision tree, i.e. the length of the longest path from the tree root to a leaf

(`MaxDepth`). The number of trees used for the training can be adjusted with the `NTrees` parameter.

The procedure of combining the trees into a single classifier is called the boosting method. There are two types of boosting that can be applied in the TMVA framework, the gradient boost and the adaptive boost. For the production cross section measurement in this thesis the former one is used as described in the following.

Each trained decision tree m is assigned to a parametrised base function (so-called weak learner) $f_m(x)$, where x correspond to the input variables. A weighted sum of the weak learners defines the function

$$F(x) = \sum_{m=0}^{M} \beta_m f_m(x), \tag{5.2}$$

where M corresponds to the number of trees and β_m to the weight of each tree. The gradient boost method tries to find in accordance with the empirical risk minimisation principle an approximation of $F(x)$ that minimises the averaged value of the loss function $L(F(x), y)$, which is defined as the deviation between the model response and the output value y. In case of the gradient boosting the loss function is a binomial log-likelihood function

$$L(F(x), y) = \ln\left(1 + e^{-2F(x)y}\right). \tag{5.3}$$

The minimum of this loss function is obtained with a steepest-descent approach. The current gradient of the loss function is calculated and a new growing regression tree $h(x)$ is fitted to the gradient values with the condition

$$h(x_i) = -\frac{\partial L\left(F(x_i), y\right)}{\partial F(x_i)}. \tag{5.4}$$

The new tree $h(x)$ is added to the function $F(x)$ multiplied by a step-size which controls the learning rate of the algorithm. This procedure is repeated till the minimum of the loss function is reached. The learning rate is steered with the `Shrinkage` parameter, which controls the weight of the individual trees. Small parameters (0.1-0.3) allow for more trees to be grown enhancing the robustness of the training procedure. In some cases the algorithm benefits from the introduction of a resampling procedure which uses random subsamples of the training sample in order to build the tree (stochastic gradient boosting, enabled with the `UseBaggedGrad` option). The fraction of the training sample used for the random subsamples can be adjusted with the `BaggingSampleFraction` parameter.

The final classifier obtained by the above minimisation procedure is the BDT response for each event with values inside an interval from -1 to 1. The larger the value of the BDT response, the more likely the event is a signal event.

One of the most important requirements on the machine learning procedure is to avoid overtraining. For example, it is in principle possible to generate a tree which provides a full separation between the signal and the background. However, such a tree could in general be strongly overtrained. Overtraining refers to the adaptation of the training procedure to statistical variations in the trained sample, which thus no longer corresponds to a representation of the underlying characteristics. Small trees (smaller depth) are less prone to overtraining, but due to the small size the classification performance is smaller and vice versa. Since gradient boosting provides the best performance with several small decision trees, this algorithm is less prone to overtraining compared to a single decision tree. Overtraining can be detected by comparing the BDT response from the trained sample with the response obtained with a test sample. A boosted decision tree is overtrained if the shapes of the BDT distributions for signal or background in the training sample differ from those in the test sample.

The different options which can be used to train the boosted decision trees and their short description are summarised in Table 5.2. If the option is not set in the setting, the default value is used.

The boosted decision tree observables are build in six reconstructed event categories: $0j$-$p_T^{4\ell}$-Low, $1j$-$p_T^{4\ell}$-Low, $1j$-$p_T^{4\ell}$-Med, VH-Had-enriched, VBF-enriched-p_T^j-Low and ttH-enriched. Except for the latter category, the training is based on the previous Run 2 analysis performed on the 36.1 fb^{-1} data set [182]. The boosted decision tree observable for the ttH-enriched category has been introduced only for data sets with 79.8 fb^{-1} or more, after sufficient amount of expected events was reached for this category. There are no boosted decision tree observables employed in other reconstructed event categories due to a very small corresponding number of expected events.

The setting for the training for the different categories is summarised in Table 5.3, while the training samples, the input variables (see Table 5.4) used in the training, the final BDT classifier and the performance of the algorithm are discussed for each reconstructed event category in the following.

The performance of the boosted decision tree classification can be estimated with the so-called Receiver Operating Characteristic (ROC) curve, which is obtained on the test samples. It corresponds to the background rejection in dependence on the signal efficiency for different values of BDT thresholds. In general, an algorithm performs better than another one if its associated ROC curve reaches higher background rejection for the same signal efficiency. Therefore, the area under the ROC curve is suitable to compare various configurations. The larger the area under the ROC curve, the better the BDT performance.

Table 5.2: Configuration options for the training of boosted decision trees within the TMVA framework [186]

Option	Default	Description
	Options for growing the decision tree	
`nCuts`	20	Number of grid points in the range of the variable used to find the optimal cut for the node splitting
`UseRandomisedTrees`	False	Determine at each node splitting the cut variable only as the best out of a random subset of variables
`UseNvars`	2	Size of the subset of variables used with `RandomisedTree` option
`nEventsMin`	0	Minimum events in the leaf node
`MinNodeSize`	5%	Minimum percentage of training events required in a leaf node
`MaxDepth`	3	Maximum allowed depth of the decision tree
	Options for gradient boosting	
`NTrees`	800	Number of trees in the forest
`Shrinkage`	1	Learning rate of the gradient boost algorithm
`UseBaggedGrad`	False	Use only a random subsample of all events for growing the trees in each iteration
`BaggingSampleFraction`	0.6	Relative size of bagged event sample to original size of the data sample (only if stochastic gradient boosting is used)

Table 5.3: Settings to train the of boosted decision trees within the TMVA framework [186] for the different reconstructed event categories

Option	$0j$-$p_T^{4\ell}$-Low	$1j$-$p_T^{4\ell}$-Low	$1j$-$p_T^{4\ell}$-Med	VH-Had-enriched	VBF-enriched-p_T^j-Low	ttH-enriched
`nCuts`	20	20	20	20	20	20
`UseRandomisedTrees`	False	False	False	False	False	True
`UseNvars`	–	–	–	–	–	9
`nEventsMin`	–	–	–	–	1000	–
`MinNodeSize`	2.5%	8.0%	5.0%	5.0%	–	2.5%
`MaxDepth`	2	3	3	5	5	3
`NTrees`	1000	400	150	100	50	1300
`Shrinkage`	0.10	0.03	0.06	0.275	0.275	0.03
`UseBaggedGrad`	True	True	True	True	True	True
`BaggingSampleFraction`	0.5	0.5	0.5	0.6	0.6	0.5

Table 5.4: The BDT discriminants and their corresponding input variables used for the measurement of the cross sections per particle level production bin

Reconstructed event category	Input variables	Processes	BDT discriminant
$0j$-$p_T^{4\ell}$-Low	$p_T^{4\ell}, \eta_{4\ell}, \mathcal{KD}\,(ZZ^*)$	ggF,ZZ^*	$\mathrm{BDT}_{\mathrm{ggF}}$
$1j$-$p_T^{4\ell}$-Low	$p_T^j, \eta_j, \Delta R(j, 4\ell)$	ggF,VBF	$\mathrm{BDT}_{\mathrm{VBF}}^{1j-p_T^{4\ell}-\mathrm{Low}}$
$1j$-$p_T^{4\ell}$-Med	$p_T^j, \eta_j, \Delta R(j, 4\ell)$	ggF,VBF	$\mathrm{BDT}_{\mathrm{VBF}}^{1j-p_T^{4\ell}-\mathrm{Med}}$
VH-Had-enriched	$m_{jj}, \Delta\eta_{jj}, p_T^{j1}, p_T^{j2}, \eta_{j1}, \Delta R_{jZ}^{min}, \eta_{4\ell}^*$	ggF,VBF,VH	$\mathrm{BDT}_{\mathrm{VH-Had}}$
VBF-enriched-p_T^j-Low	$m_{jj}, \Delta\eta_{jj}, p_T^{j1}, p_T^{j2}, \eta_{j1}, \Delta R_{jZ}^{min}, p_T^{4\ell jj}$	ggF,VBF	$\mathrm{BDT}_{\mathrm{VBF}}$
ttH-enriched	$m_{jj}, \Delta\eta_{jj}, \Delta R(j, 4\ell), \Delta R_{jZ}^{min}, \eta_{4\ell}^*, E_T^{miss}, p_T^{jj}, N_j, N_{b-j}, m_{sig}, H_T$	ggF,VBF,VH,ttV	$\mathrm{BDT}_{\mathrm{ttH-Had}}$

5.3.1 The 0-Jet Categories

The dominant Higgs boson signal production mode in the $0j$ categories is ggF. Since, similarly as the LO ggF production, the irreducible ZZ^* background production also leaves no additional jets in the final state it has a relatively large contribution of about 35% of all events in this category. In order to distinguish the ggF signal from the ZZ^* background a boosted decision tree is trained based on kinematic properties of the four-lepton events. Due to the slight differences of these properties for different types of lepton quadruplets, the training is performed separately for $4\mu/4e$ and $2e2\mu/2\mu2e$ events. Since only a very small fraction of events falls into the $0j$-$p_T^{4\ell}$-High category, the training is performed with events in both $0j$ categories combined.

The input variables which provide a good discrimination between the signal and the background are the transverse momentum and the pseudorapidity of the four lepton final state, $p_T^{4\ell}$ and $\eta_{4\ell}$, respectively. In addition a matrix element based observable is used, defined as

$$\mathcal{KD}\left(ZZ^*\right) = \log\left(\left|\mathcal{M}_{HZZ^*}\right|^2 / \left|\mathcal{M}_{ZZ^*}\right|^2\right), \tag{5.5}$$

where $\left|\mathcal{M}_{HZZ^*}\right|^2$ and $\left|\mathcal{M}_{ZZ^*}\right|^2$ correspond to the matrix elements calculated at LO with the MadGraph5_aMC@NLO generator [181] for ggF and ZZ^* production, respectively. The matrix element calculation is based on the four-momenta of four leptons in the final state. The expected signal and background distributions of the discriminating variables used for training are shown in Figure 5.9 separately for $4\mu/4e$ and $2e2\mu/2\mu2e$ final states. For both final states, the $p_T^{4\ell}$ distribution is harder and $\eta_{4\ell}$ more central in case of the ggF production. The matrix element based observable $\mathcal{KD}\left(ZZ^*\right)$ tends to smaller values for the ZZ^* continuum.

Figure 5.10(a) and Figure 5.10(b) show the classifier output distributions (BDT response) of the training and testing sample for signal (ggF) and background (ZZ^*) separately in the $4\mu/4e$ and $2e2\mu/2\mu2e$ final states. A good separation between the signal and the background is obtained. The shapes of the BDT response distributions agree rather well for testing and training sample, indicates that no significant overtraining is expected. The corresponding ROC curves are shown in Figure 5.10(c) and Figure 5.10(d). The BDT selection for the $2e2\mu/2\mu2e$ states with an integral of the ROC curve of $I_{\mathrm{ROC}} = 0.7820$ has a slightly better performance than for the $4\mu/4e$ final states with $I_{\mathrm{ROC}} = 0.7647$.

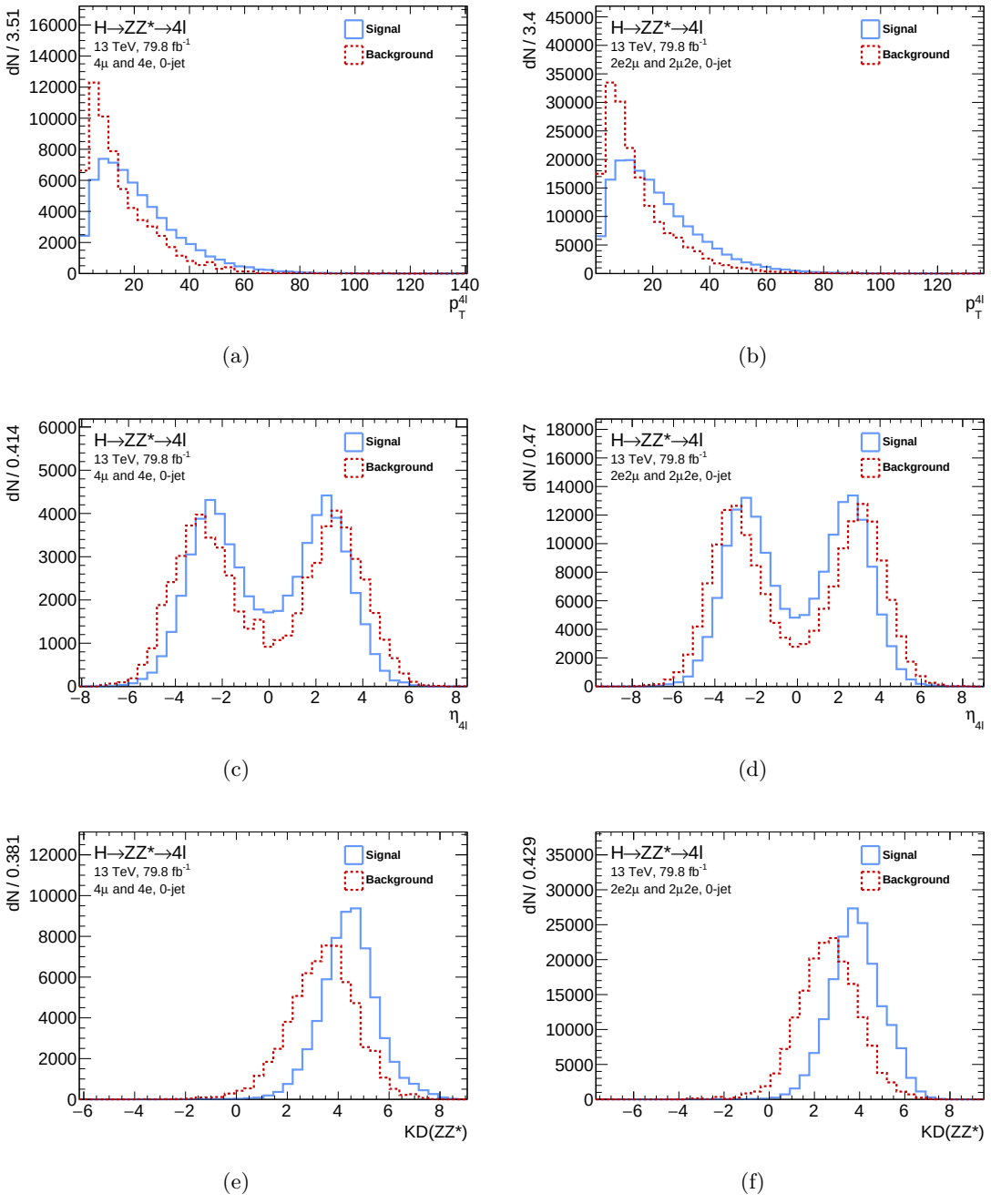


Figure 5.9: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) and (b) $p_T^{4\ell}$, (c) and (d) $\eta_{4\ell}$, (e) and (f) $\mathcal{KD}(ZZ^*)$ in the $0j$ reconstructed event category. The signal (ggF) distributions are shown in blue and background (ZZ^*) in red. The training is performed with samples in the combined $0j$ category, separately for $4\mu/4e$ and $2e2\mu/2\mu2e$ final states.

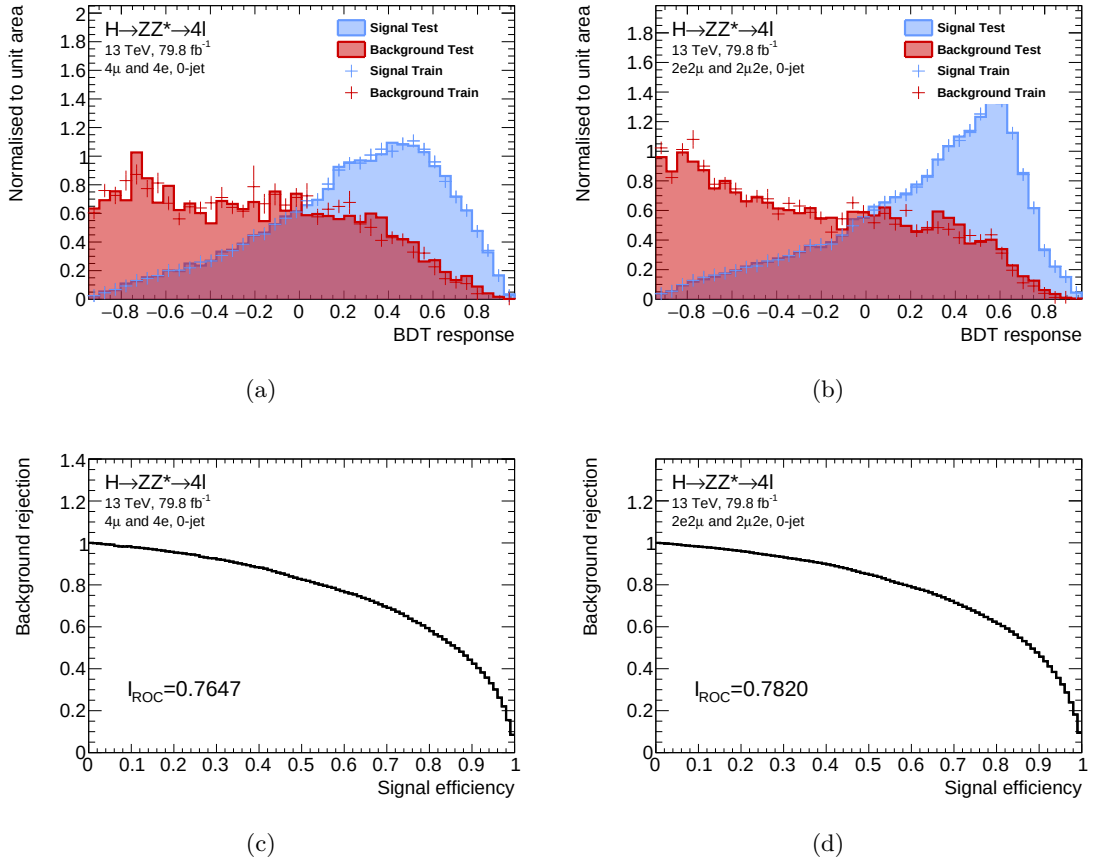


Figure 5.10: Classifier output distributions (BDT response) of the training and testing sample for signal (ggF) and background (ZZ^*) processes in the $0j$ category for (a) $4\mu/4e$ and (b) $2e2\mu/2\mu2e$ final states and (c) and (d) the corresponding ROC curves.

5.3.2 The 1-Jet Categories

The VBF production has a characteristic signature of two high-energy forward jets in the final state. However, it is possible that one of these jets is not reconstructed in the detector. Hence, about 10% of the VBF events fall into one of the $1j$ categories contaminating the dominant ggF contribution. To separate the Higgs boson events produced via the VBF from the ggF production, a boosted decision tree is trained separately for events in the $1j$-$p_T^{4\ell}$-Low and $1j$-$p_T^{4\ell}$-Med category. Since the same input variables are used in both categories, the same description of the procedure applies for both. In the $1j$-$p_T^{4\ell}$-High category no boosted decision tree discriminant is used due to a small expected number of events.

Three input variables are used: the transverse momentum of the leading jet (p_T^{j1}), its pseudorapidity (η_{j1}) and the angular separation between the reconstructed jet and the

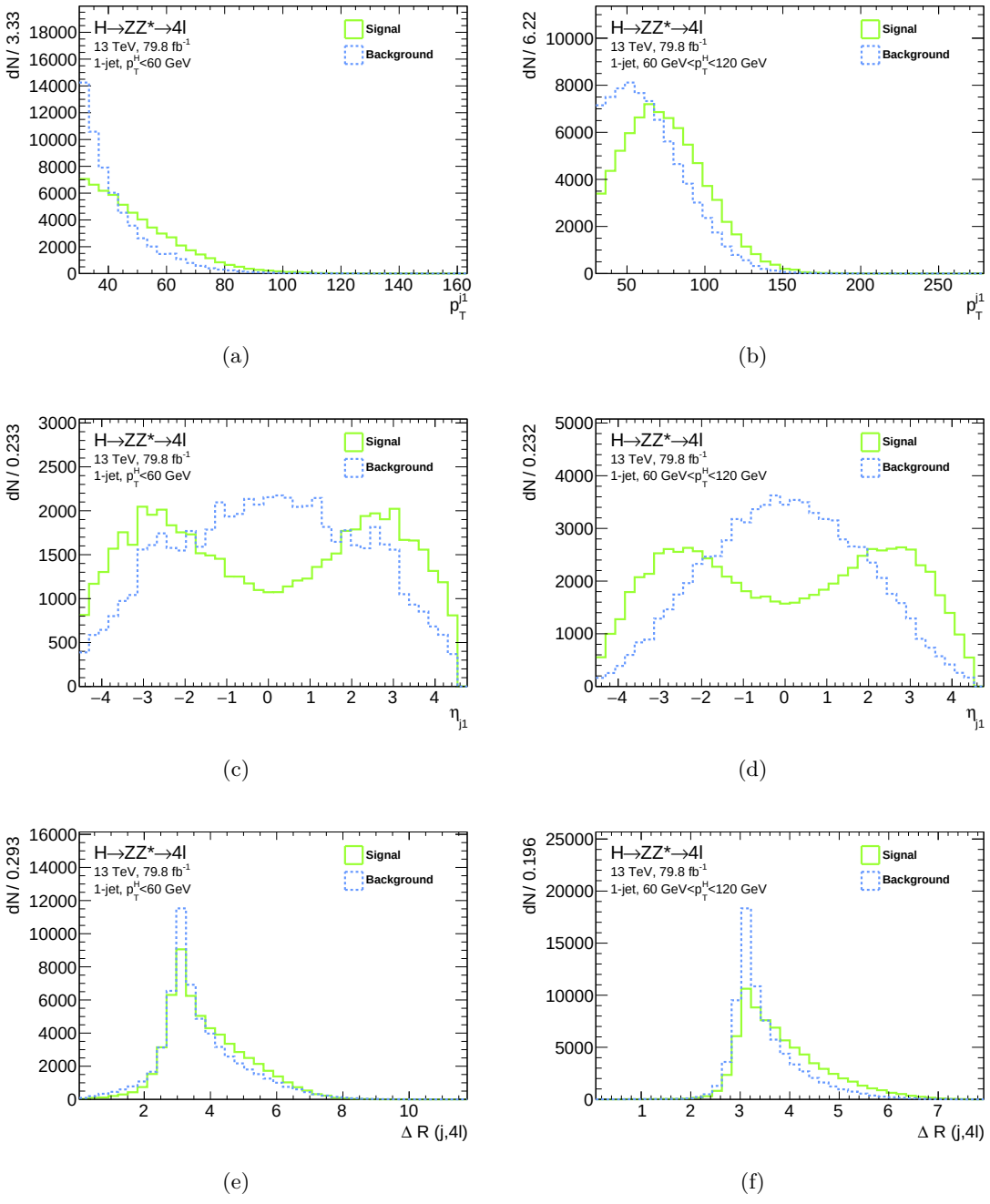


Figure 5.11: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) and (b) p_T^{j1}, (c) and (d) η_{j1}, (d) and (e) $\Delta R(j, 4\ell)$ in the $1j$-$p_T^{4\ell}$-Low and $1j$-$p_T^{4\ell}$-Med reconstructed event category. The signal (VBF) distributions are shown in green and background (ggF) in blue. The training is done separately in each of the categories.

4ℓ system ($\Delta R(j, 4\ell)$). The expected signal (VBF) and background (ggF) distributions of discriminating input variables are shown in Figure 5.11 separately for the $1j$-$p_T^{4\ell}$-Low and $1j$-$p_T^{4\ell}$-Med category. The jet from the VBF process tend to have a larger p_T compared to the jet from the initial state radiation in the case of the ggF production. The VBF jets are emitted in the forward region, which is clearly visible in the pseudorapidity distribution peaking at around $|\eta_{j1}| = 3$. The ggF jets, on the other hand, are produced more centrally. The η_{j1} distribution from ggF has a slight double peak structure similar to VBF. This is due to the contamination of pile-up jets. At least, the jets from the VBF production have a larger angular separation from the 4ℓ system compared to ggF jets.

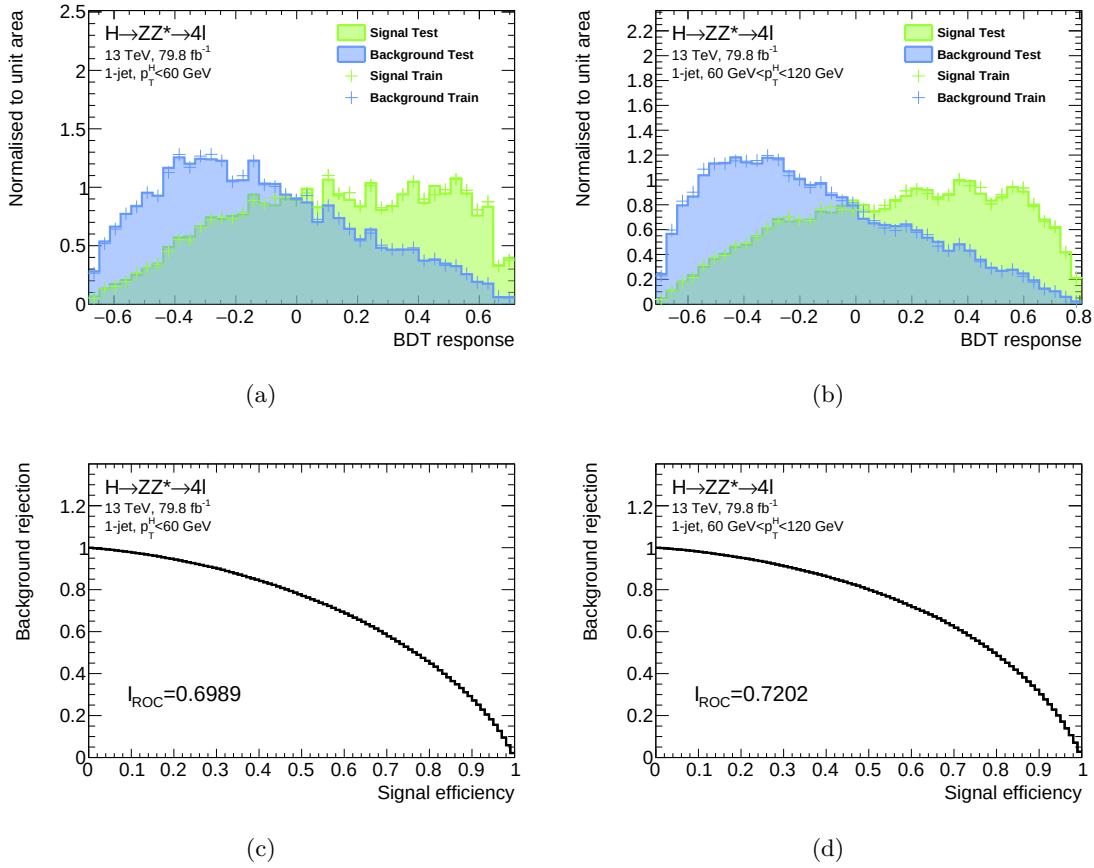


Figure 5.12: Classifier output distributions (BDT response) of the training and testing sample for signal (VBF) and background (ggF) processes in the (a) $1j$-$p_T^{4\ell}$-Low and (b) $1j$-$p_T^{4\ell}$-Med category and (c) and (d) the corresponding ROC curves.

The boosted decision tree output distributions for the testing and training samples, as well as the corresponding ROC curves are shown in Figure 5.12 separately for each event category. In both cases, the boosted decision tree output distributions show a good sep-

aration between the VBF signal and the ggF background. The shapes of the distributions from testing and training samples agree very well, therefore no overtraining is expected. The integral of the ROC curve for the $1j$-$p_T^{4\ell}$-Low category is $I_{\mathrm{ROC}} = 0.6989$, indicating a slightly worse performance than for the $1j$-$p_T^{4\ell}$-Med category with $I_{\mathrm{ROC}} = 0.7202$.

5.3.3 The VH-Hadronic Enriched Category

The VH-Had-enriched category is designed to provide sensitivity to VH production with hadronic V boson decays, but is highly contaminated with ggF events comprising 70% of the Higgs signal in this category. Also events from the VBF production contribute but with a smaller fraction of about 6%. In order to increase the sensitivity to the VH-Had production mode in this category, a boosted decision tree is used to separate the signal (VH-Had) from the contaminating background production modes (ggF and VBF).

Seven observables are employed to provide the separation between signal and background processes: the invariant mass of the dijet system (m_{jj}), its pseudorapidity difference ($\Delta\eta_{jj}$), the transverse momenta of the two leading jets (p_T^{j1} and p_T^{j2}), the pseudorapidity of the leading jet (η_{j1}), the minimal angular separation between the jets and two Z boson candidates (ΔR_{jZ}^{min}) and a variable describing the pseudorapidity difference between the 4ℓ and dijet system, $\eta_{4\ell}^*$ defined as $\eta_{4\ell}^* = \eta_{4\ell} - \langle\eta_{j1}, \eta_{j2}\rangle$. The expected distributions of discriminating observables used for the training are shown in Figure 5.13 and Figure 5.14. A very good separation is obtained with the m_{jj} distribution. The distribution from the VH-Had production peaks at the mass of the vector boson resonance, while ggF and VBF events produce a flat distribution in the considered mass range. The difference in pseudorapidity of the dijet system $\Delta\eta_{jj}$ has a rather small separation power, with background processes tending to slightly larger values. The spectra of p_T^{j1} and p_T^{j2} are harder for the VH-Had process, since the main background originates from initial state radiation jets in the ggF production. The leading jet from the decaying vector boson is produced more centrally as in the case of the ggF and VBF production. The minimal angular between the jets and Z bosons tends to be smaller for signal process. A very good discrimination is provided by the $\eta_{4\ell}^*$ observable. In case of the targeted VH-Had signal process the four leptons and the two jets are produced close to each other leading to smaller $\eta_{4\ell}^*$ values compared to the ggF and VBF production.

The BDT response distributions for signal (VH-Had) and background (ggF and VBF), as well the corresponding ROC curve are shown in Figure 5.15. In the VH-Had-enriched category a separation is achieved with no overtraining expected. The ROC integral of $I_{\mathrm{ROC}} = 0.8164$ indicates a better boosted decision tree performance compared to the $0j$ and $1j$ categories.

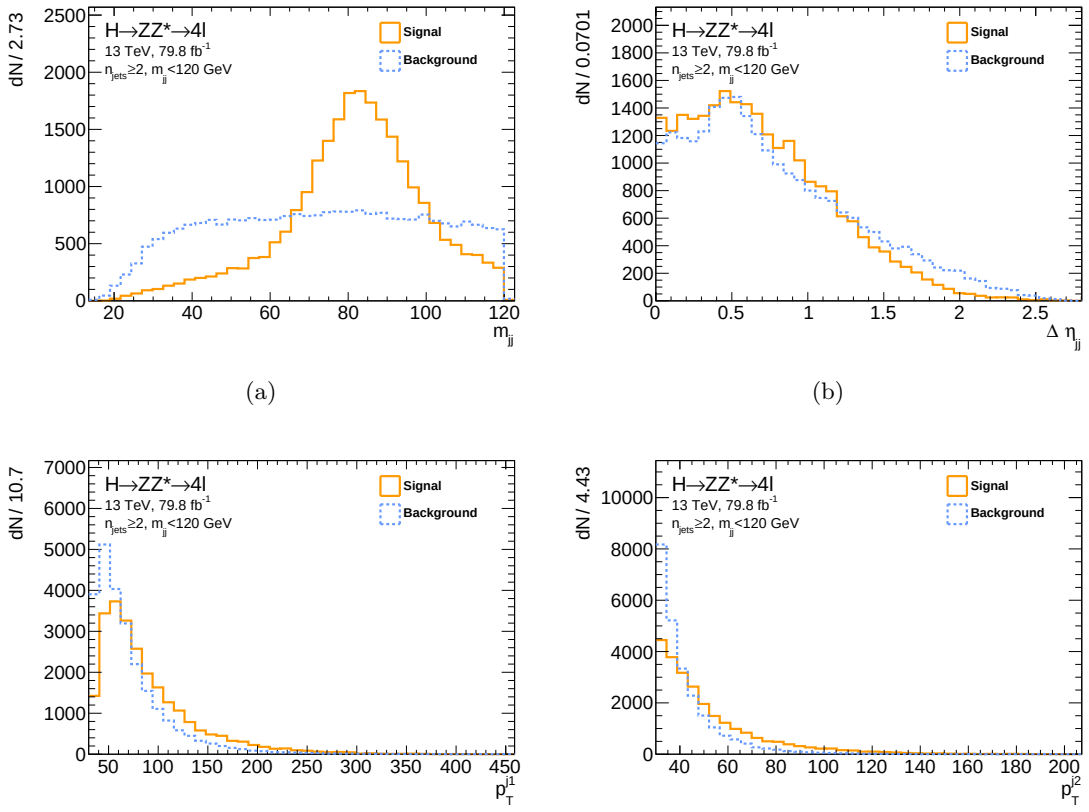


Figure 5.13: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) m_{jj}, (b) $\Delta\eta_{jj}$, (c) p_T^{j1}, (c) p_T^{j2} in the VH-Had-enriched reconstructed event category. The signal (VH-Had) distributions are shown in orange and background (ggF and VBF) in blue.

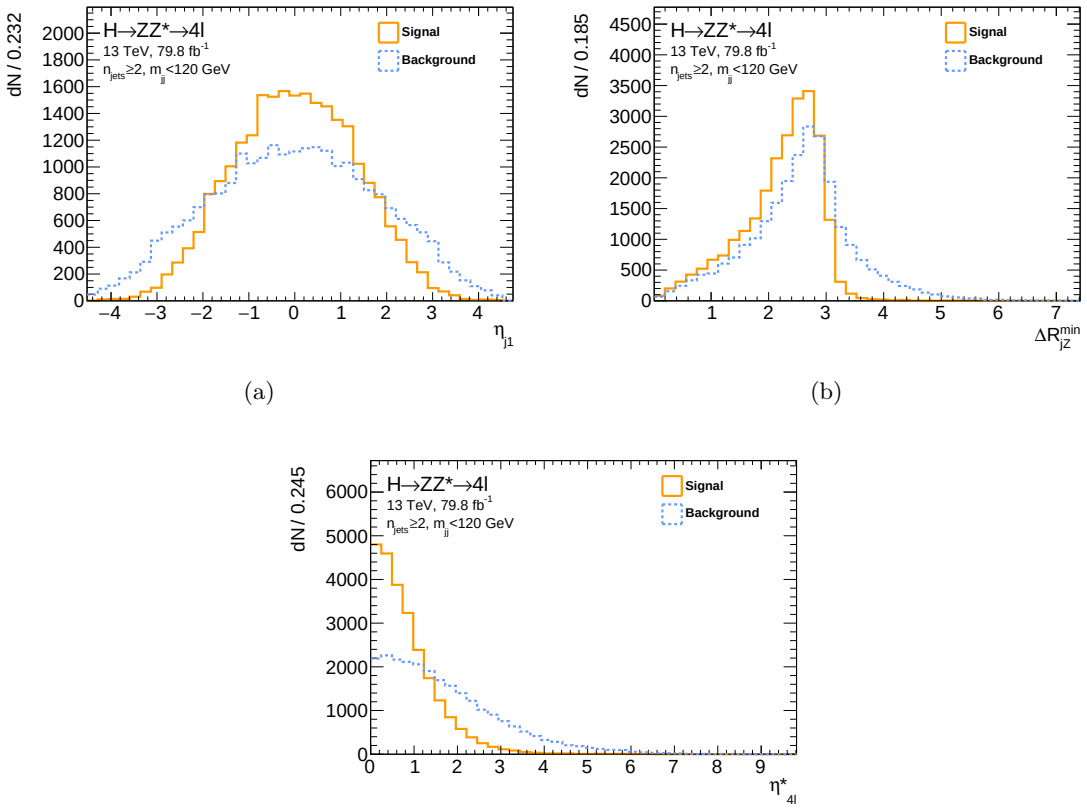
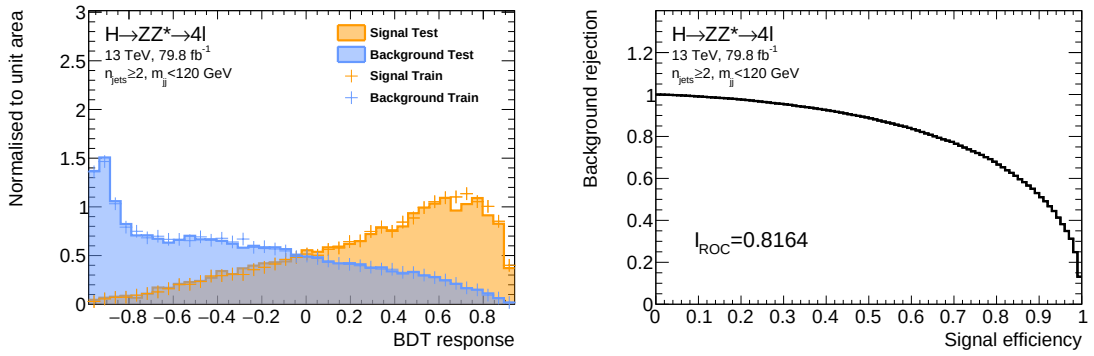


Figure 5.14: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) η_{j1}, (b) ΔR_{jZ}^{min}, (c) $\eta_{4\ell}^{*}$ in the VH-Had-enriched reconstructed event category. The signal (VH-Had) distributions are shown in orange and background (ggF and VBF) in blue.



(a) (b)

Figure 5.15: (a) Classifier output distributions (BDT response) of the training and testing sample for signal (VH-Had) and background (ggF and VBF) processes in the VH-Had-enriched category and (b) the corresponding ROC curve.

5.3.4 The VBF Enriched Category

Similar as in the case of the VH-Had-enriched category, the VBF-enriched category is also contaminated by the ggF production with a fraction of up to 50%. Therefore, the VBF production is further discriminated from ggF by means of a boosted decision tree discriminant. In order to enhance the number of events for the training, the two VBF-enriched categories are used together, i.e. no threshold is imposed on p_T^{j1}.

Input observables which provide a good separation between the VBF and ggF production are the invariant mass of the dijet system (m_{jj}), the difference of jet pseudorapidities ($\Delta\eta_{jj}$), the transverse momenta of the two leading jets (p_T^{j1} and p_T^{j2}), the minimal angular separation between the two jets and Z boson candidates (ΔR_{jZ}^{min}) and the distance in the pseudorapidity between the four lepton and dijet system $\eta^*_{4\ell}$. In addition, the transverse momentum of the Higgs-dijet system $p_T^{'4\ell jj}$ is used. To avoid large theoretical uncertainties for small values of the Higgs-dijet system transverse momentum, the observable is defined as

$$p_T^{'4\ell jj} = \begin{cases} p_T^{4\ell jj} & \text{if } p_T^{4\ell jj} > 50 \text{ GeV} \\ 50 \text{ GeV} & \text{if } p_T^{4\ell jj} < 50 \text{ GeV} \end{cases}. \tag{5.6}$$

The expected distributions of input variables used for boosted decision tree training are shown in Figure 5.16 and 5.17. Due to the characteristic VBF final state of two forward high-energy jets, the m_{jj} distribution is harder and the two jets are more separated in pseudorapidity as in the case of the ggF production. Also the transverse momenta of the two leading jets tend to be larger for the targeted VBF signal. A good separation is also provided by the $\eta^*_{4\ell}$ observable, whose values are smaller for the VBF production than for

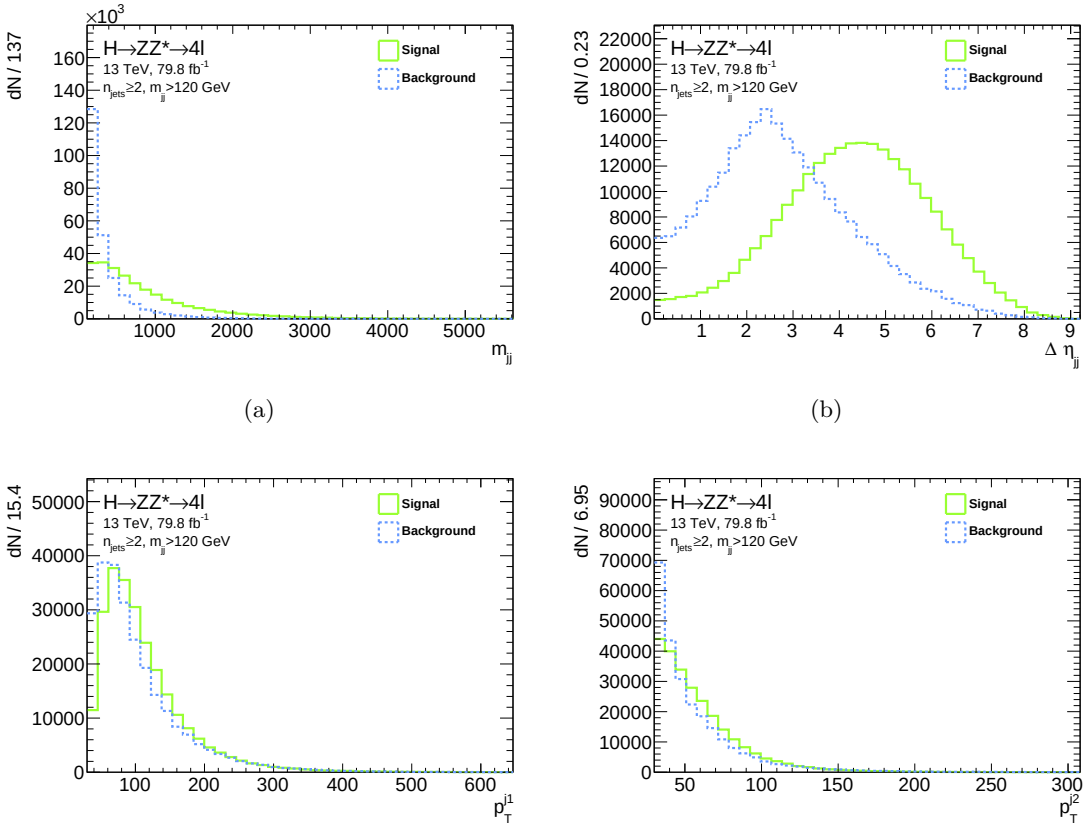


Figure 5.16: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) η_{j1}, (b) $\Delta\eta_{jj}$, (c) p_T^{j1}, (d) p_T^{j2}, in the VBF-enriched reconstructed event category. The signal (VBF) distributions are shown in green and background (ggF) in blue.

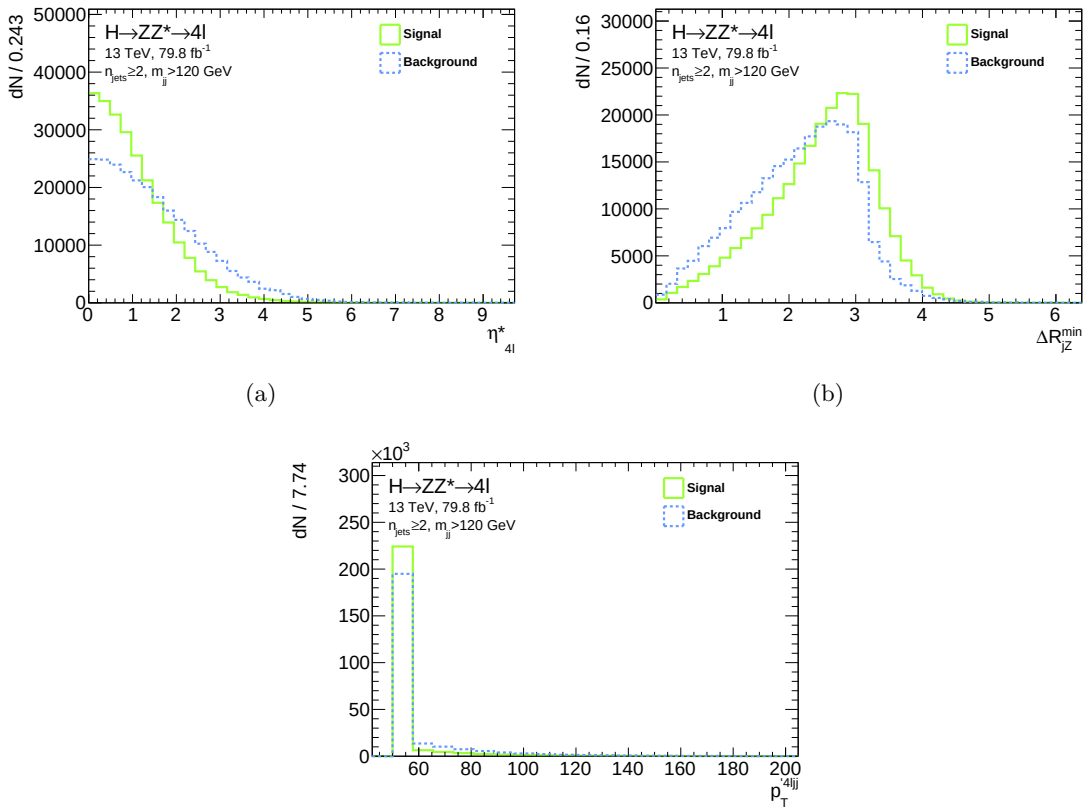
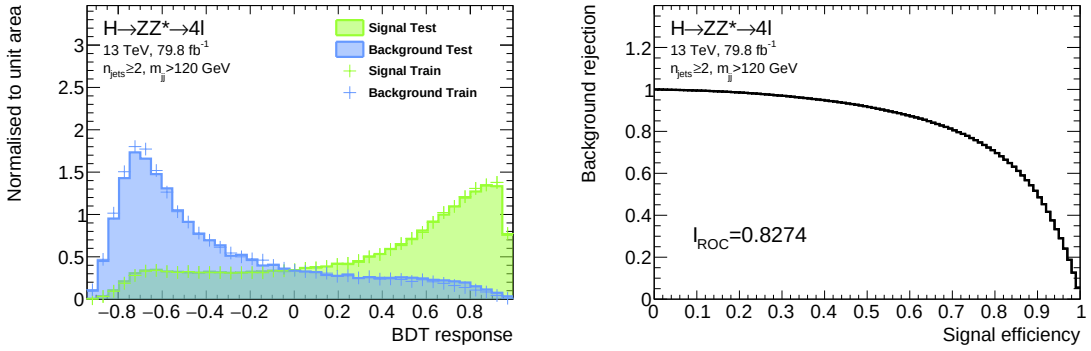


Figure 5.17: Distributions of the discriminating variables used as an input for the boosted decision tree training: (a) $\eta^*_{4\ell}$, (b) ΔR^{min}_{jZ}, (c) $p_T^{'4\ell jj}$ in the VBF-enriched reconstructed event category. The signal (VBF) distributions are shown in green and background (ggF) in blue.

ggF. In addition, the minimal angular separation ΔR_{jZ}^{min} is larger and the transverse momentum of the Higgs-dijet system production is smaller for the considered VBF signal process compared to the ggF background.

The BDT response distribution for signal (VBF) and background (ggF), as well as the ROC curve are shown in Figure 5.18. A good separation of the two production modes is expected with no significant overtraining. The integral of the ROC curve $I_{\mathrm{ROC}} = 0.8274$, showing the best performance compared to previously introduced boosted decision tree discriminants in other reconstructed event categories.



(a) (b)

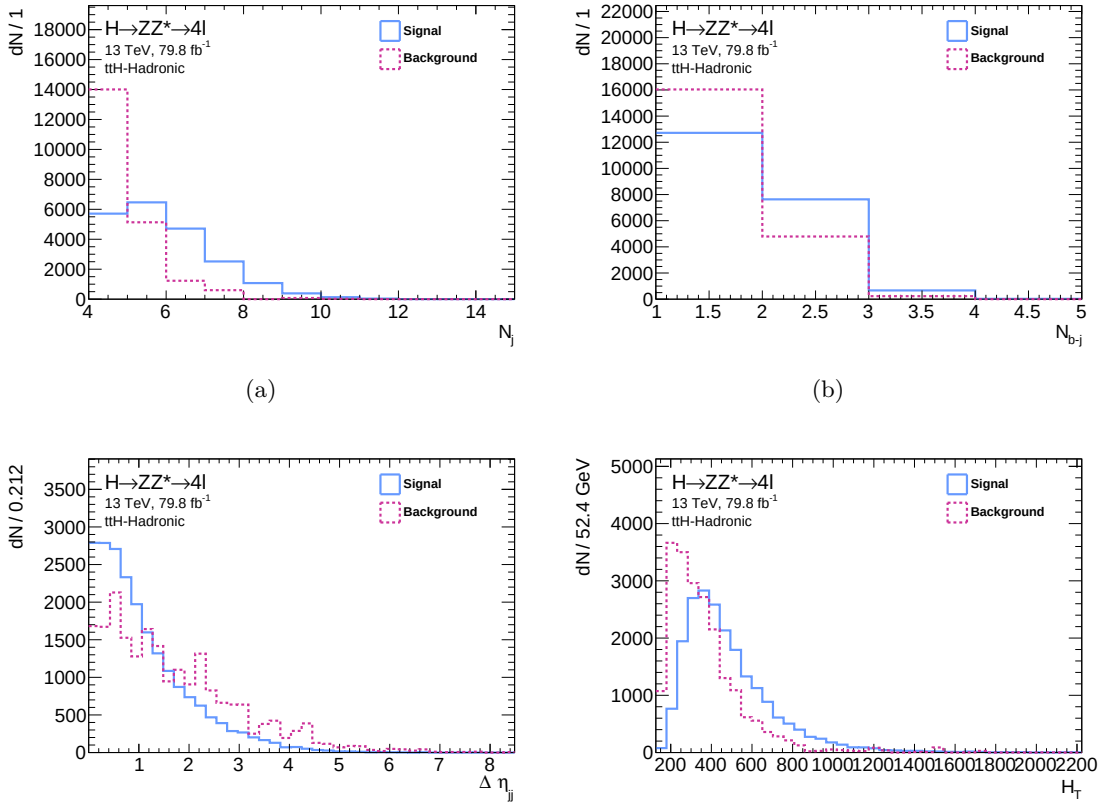
Figure 5.18: (a) Classifier output distributions (BDT response) of the training and testing sample for signal (VBF) and background (ggF) processes in the VBF-enriched category and (b) the corresponding ROC curve.

5.3.5 The ttH-Hadronic Enriched Category

Unlike the very pure ttH-Lep-enriched category, the contamination from other Higgs boson signal processes as well as from ttV production is about 30% in the ttH-Had-enriched category. To distinguish the ttH signal from ggF, VBF, VH and ttV processes a boosted decision tree is trained, with the ttH process treated as background and the sum of ggF, VBF, VH and ttV as signal.

As it is difficult to distinguish the ttH signal from the other processes contributing in the ttH-Had-enriched category, a large set of discriminating input variables is chosen: the invariant mass of the two leading jets (m_{jj}), the difference in jet pseudorapidity ($\Delta\eta_{jj}$), the dijet transverse momentum (p_T^{jj}), the number of jets (N_j) and b-tagged jets (N_{b-j}), the minimal angular separation between the two jets and Z boson candidates (ΔR_{jZ}^{min}), the angular separation between the reconstructed leading jet and the 4ℓ system

($\Delta R(j, 4\ell)$), as well as $\eta^*_{4\ell}$. Further input variables are the scalar sum of transverse momenta of the four leptons and the jets (H_T), the computed matrix element for the lepton pairing ($\mathcal{M}_{sig}$, see Section 4.3) and the missing transverse energy (E_T^{miss}). The expected distributions of discriminating observables with largest separation power used for the training are shown in Figure 5.19. A good separation is provided by the number of jets and b-tagged jets in the final state, which tend to be larger in case of the ttH production compared to the others. The difference in jet pseudorapidity is larger in case of the targeted ttH production, while the scalar sum the of transverse momenta of the four leptons and the jets is smaller.

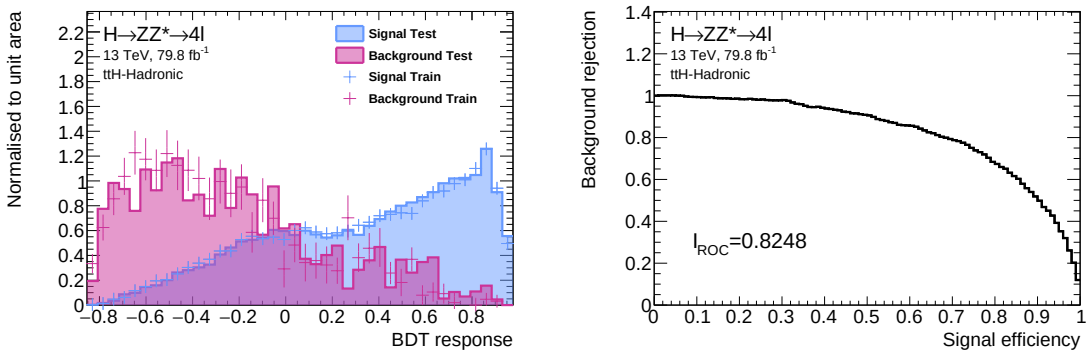


(a) (b)

(c) (d)

Figure 5.19: Distributions of the discriminating variables with largest separation power used as an input for the boosted decision tree training: (c) $\Delta\eta_{jj}$, (a) N_j, (b) N_{b-j}, (d) H_T in the ttH-Had-enriched reconstructed event category. The signal (ggF, VBF, VH and ttV) distributions are shown in blue and background (ttH) in violet.

The BDT response distribution for signal (ggF, VBF, VH and ttV) and background (ttH), as well as the corresponding ROC curve shown in Figure 5.20. A good separation of the ttH production from ggF, VBF, VH and ttV processes is expected with a little amount of overtraining. The integral of the ROC curve $I_{\mathrm{ROC}} = 0.8248$ showing a similar performance as the boosted decision tree discriminant in the VBF-enriched-p_T^j-Low category.



(a) (b)

Figure 5.20: (a) Classifier output distributions (BDT response) of the training and testing sample for signal (ggF, VBF, VH and ttV) and background (ttH) processes in the ttH-Had-enriched category and (b) the corresponding ROC curve.

5.4 Systematic Uncertainties

In addition to the statistical uncertainty of the data and simulated samples, systematic uncertainties associated to limitations of the experimental equipment or theory modelling are taken into account in the analysis. The systematic uncertainties can have an impact on the expected signal and background yields as well as the signal acceptance and reconstruction efficiency. The distributions of multivariate discriminants can also be affected.

The experimental uncertainties originate from uncertainties on the reconstruction of leptons and jets, their identification, isolation and trigger efficiencies, as well as on their energy scale and resolution measurement. Uncertainty on the measurement of the total integrated luminosity and uncertainty on the data-driven background estimates also contribute to the total experimental uncertainty.

Theory uncertainties account for the uncertainty on theoretical modelling of signal and background processes, such as the choice of renormalisation and factorisation scales (QCD scale), migration of events between different jet and p_T bins, PDF and α_s, as well as parton shower uncertainties.

The experimental and theoretical uncertainties are described in the following. Since, the main contribution for this thesis was the evaluation of the impact of theoretical uncertainties on the signal, this is described in more detail in Section 5.4.2. For the final statistical interpretation of the data, only systematic uncertainties which have an impact larger than 1% are taken into account. The impact of the dominant systematic uncertainties on the cross section measurement in the Stage-0 production bins (see Section 5.6) is summarised in Table 5.5.

Table 5.5: Impact of the dominant systematic uncertainties on the cross section measurement in Stage-0 production bins for 79.8 fb^{-1} of data at $\sqrt{s} = 13$ TeV. Systematic uncertainties from similar sources are grouped together. Luminosity, electron and muon reconstruction and identification efficiencies and pile-up (e, μ, pile-up), jet energy scale, energy resolution and b-tagging efficiencies (jets, flavour tagging), uncertainties on the reducible background (reducible background), theoretical uncertainties on the ZZ^* background (ZZ^* background), and theoretical uncertainties on the signal due to parton density function (PDF), QCD scale and showering algorithm (Parton shower) uncertainties [183].

	Experimental uncertainties [%]				Theoretical uncertainties [%]			
	Luminosity	e, μ pile-up	Jets, flavour tagging	Reducible background	ZZ^* background	PDF	QCD scale	Signal parton shower
ggF	2.9	3.9	1.3	0.7	2.3	0.4	2.1	0.7
VBF	1.7	1.5	10.5	0.5	2.3	2.3	9.5	5.1
VH	2.0	1.7	7.8	1.8	5.6	2.1	14.9	3.1
ttH	2.5	1.9	3.9	1.5	1.9	0.3	8.8	9.6

5.4.1 Experimental Uncertainties

The integrated luminosity is measured with a precision of 2%. This uncertainty affects the predicted number of simulated signal and background events resulting in cross section uncertainties.

The uncertainty on pile-up modelling introduce an uncertainty on simulated event yields in the range from 2% and 5%.

The size of uncertainties on the lepton identification, isolation and reconstruction efficiencies, as well as on the energy and momentum scale and resolution depends on the detector region. The uncertainty on muon reconstruction efficiencies are in the order of $1-2\%$, while they are in the range of $1-4\%$ for electrons. The lepton identification uncertainty is 2%. Uncertainties on the lepton momentum scale and resolution are smaller than 1% and are, therefore, not considered for the final result.

The uncertainties on the jet energy scale and resolution affect the predicted event yields by $3-7\%$ and $2-4\%$, respectively. The uncertainty on the efficiency on the b-tagging algorithm is relevant for the ttH-enriched category, impacting the predicted yields by approximately 5%. The uncertainty related to the E_T^{miss} reconstruction has a negligible impact on the ttH cross section measurement.

The uncertainties affecting the data-driven estimation of the reducible background can be classified into three sources. The smallest uncertainty originates from the statistical uncertainty of the fit in the control data which affects the background normalisation by about 4%. As described in Section 4.4, the identification, isolation and impact parameter efficiencies defining the transfer factors are evaluated in data and simulation in $Z+X$ control regions. The difference between data and simulation is assigned as an uncertainty on the transfer factor extrapolation from the control to the signal region. The uncertainty is about 6%. The largest source of uncertainty is the determination of the fractions of the reducible background in each experimental category, which is limited by the statistical precision of the simulated samples. This uncertainty ranges from $8-70\%$ depending on the reconstructed event category.

5.4.2 Theoretical Uncertainties

The dominant sources of theoretical uncertainties are the choice of the renormalisation and factorisation scales (QCD scale), the PDF and α_s uncertainties, as well as the modelling of the parton showers and underlying events.

In general, these uncertainties are evaluated with the help of event generators. The QCD scale uncertainties are evaluated with generated events which are reweighted by varying values of renormalisation and factorisation scales. The PDF and α_s uncertainties are evaluated by considering all internal variations (eigenvectors) in the employed PDF set [188]. The parton shower uncertainties are obtained by varying the nominal values of the parton shower generator tune and by using additional generated samples with a different choice of the parton shower generator. In all cases the sum of event weights related to a given variation w_i^{var} is than compared to the sum of nominal event weights w_i^{nom} to compute the corresponding resulting uncertainty δ,

$$\delta = \frac{\sum_i w_i^{var}}{\sum_i w_i^{nom}}. \tag{5.7}$$

The signal uncertainty defined in Equation 5.7 takes into account both the uncertainty on the cross section σ for a given signal process, as well as the acceptance ($\mathcal{A} \times \epsilon$) uncertainty introduced by the event selection and categorisation criteria. Both have to be taken into account for the measurement of the signal strength, i.e. of the ratios of the measured cross sections and their SM predictions. In case that the absolute value of the signal cross section is fitted to data, rather than the ratio with respect to the SM, only the uncertainty

on the signal acceptance needs to be taken into account. Theory uncertainties, therefore, have to be evaluated such that the cross section variations cancel out. For a given particle level bin k in each of the reconstructed event categories j the uncertainty on the acceptance is given as

$$\frac{(\mathcal{A}\times\epsilon)_{kj}^{var}}{(\mathcal{A}\times\epsilon)_{kj}^{nom}} = \frac{\sigma^{k,nom}}{\sigma^{k,var}}\left(\frac{\sum_i w_i^{kj,var}}{\sum_i w_i^{kj,nom}}\right) \approx \frac{\sum_n w_n^{k,nom}}{\sum_n w_n^{k,var}}\delta, \tag{5.8}$$

where index i runs over all events of a particle level bin in a given reconstructed category and n runs over all events in the same particle level bin in all reconstructed categories. The factor in front of the variation δ cancels the contributions of the cross section variation. It is calculated in particle level samples, prior to the event selection, separately for each of the particle level bin.

The determination of theoretical uncertainties on the Higgs boson signal is described in more detail in the following. The impact of theoretical uncertainties on the shape of the boosted decision tree discriminants has been also evaluated as a part of this thesis. Since the same methodology is used as for the Higgs boson signal yields, this is not described in detail. The impact of the QCD scale and PDF+α_s uncertainties on the shapes of the boosted decision tree discriminants is approximately $1-2\%$. In the $1j$-$p_T^{4\ell}$-Low, $1j$-$p_T^{4\ell}$-Med, VH-Had-enriched and VBF-enriched-p_T^j-Low reconstructed event categories, the statistical uncertainties are of the same size as the calculated parton shower uncertainties. Therefore, no additional shape uncertainty is included. The uncertainty on the shape of the boosted decision tree response distribution in the ttH-Had-enriched category due to parton shower variations is about 25% for the bin with the largest boosted decision tree score and only 5% for the other one.

The uncertainties on the background yields are also derived in a similar way. The QCD scale uncertainty affects the expected ZZ^* background yield by about 4% for the inclusive analysis, increasing to up to 30% in reconstructed categories with additional jets. For the irreducible background the impact of PDF and α_s uncertainties is approximately $1-2\%$ and the parton shower uncertainties are estimated to contribute with approximately $1-5\%$. The impact of QCD scale and PDF uncertainties on ttV and VVV processes is estimated to be 10% and 4%, respectively.

5.4.2.1 QCD Scale and Jet Reconstruction

One of the dominant systematic uncertainties is the uncertainty on the ggF production cross section due to missing higher-order calculation terms in QCD. In general, these uncertainties are evaluated by varying the QCD renormalisation (μ_R) and the factorisation (μ_F) scales. However, in case of the ggF production mode this procedure would lead to an underestimated uncertainty. In the reconstructed 1-jet and 2-jet event categories, there is a large contribution from the ggF production mode. This splitting in jet multiplicity (N_j)

may induce additional uncertainties on the fixed order cross section predictions for the ggF production which contains no jets at tree-level [189]. More precisely, predictions of the QCD scale variation uncertainties in NNLO or NLO fixed order can be unrealistic for the ggF production mode due to the presence of unresummed Sudakov logarithms in the perturbative expansion. These terms occur by including soft, virtual and collinear gluon effects for the cross section calculation by performing a threshold next-to-next-to-next-to-leading logarithmic (N^3LL) resummation. In general, the presence of higher-order real emission terms offsets the unresummed Sudakov logarithms terms. However, by applying an exclusive cut on the number of jets effectively cuts out these corrective emission terms from the perturbative expansion, which leads to an unrealistically small error when μ_R and μ_F are varied.

Therefore, when performing an analysis with exclusive jet multiplicity categories, the choice of the resummation scale and the migrations between the N_j bins, i.e. in this case migrations between 0-jet and 1-jet bins and between 1-jet and ≥ 2-jets bins, have to be taken into account simultaneously in addition to the choice of the factorisation and renormalisation scale.

In order to estimate the scale variation and the migration effects the uncertainties from Boughezal-Liu-Petriello-Tackmann-Walsh (BLPTW) approach are used [189]. They are propagated through the SM ggF sample by randomly shifting the event weights with different smearing factors depending on the particle level jet multiplicity.

Similarly, the classification of events into exclusive regions of Higgs boson transverse momentum p_T^H also induces non-trivial migration effects affecting the cross section prediction. Therefore, additional uncertainties on the yields are taken into account to reflect the migration of events between the $1j$-$p_T^{4\ell}$-Low and $1j$-$p_T^{4\ell}$-Med, as well as between the $1j$-$p_T^{4\ell}$-Med and $1j$-$p_T^{4\ell}$-High categories. These are accounted for by introducing an additional uncertainty on the shape of the p_T^H distribution, which is derived using internal QCD scale variations in a ggF POWHEG sample generated with NNLOPS [126, 127].

An additional uncertainty is assigned for the ggF production mode to account for the heavy top quark approximation used in cross section calculation. The ggF Higgs boson production mainly takes place through top and bottom quark loops. The exact calculation of this process, accounting the loops with finite quark masses, are computationally intensive. The calculation can be simplified by assuming that the masses of top and bottom quark are infinitely large, $m_t, m_b \to \infty$, due to which the loop is replaced by an effective Higgs gluon coupling. The two approaches are in good agreement for the majority of the p_T^H spectra of the Higgs boson. However, for highly boosted Higgs bosons significant differences are observed. Therefore, an additional systematic uncertainty is assigned to account for the differences with respect to the exact calculation. The uncer-

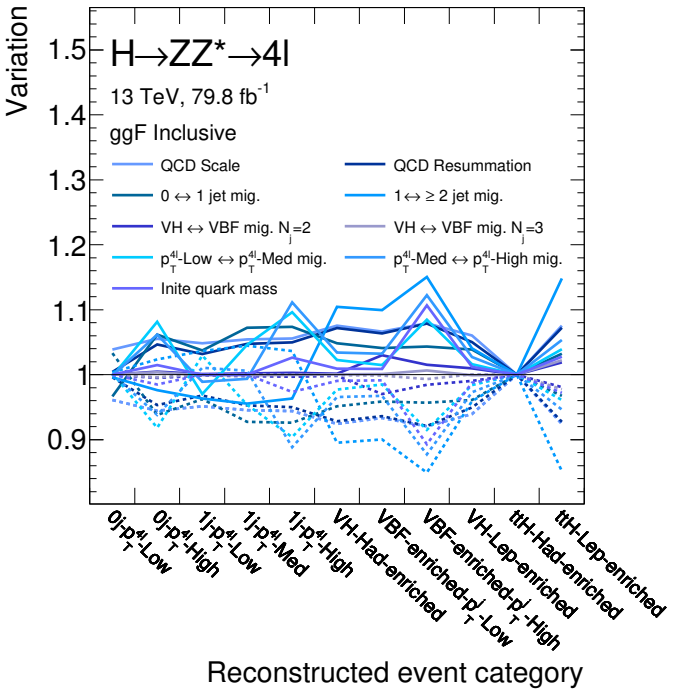
tainty is derived in a similar way as in the case of the p_T^H bin migration uncertainties by comparing the NNLOPS samples with and without the finite quark mass effects.

Finally, an additional uncertainty is assigned to account for the uncertainties in the migration of ggF events from the 0-jet into the VH-Had-enriched or VBF-enriched categories with 2-jet or 3-jets due to missing higher order in QCD. The uncertainties are estimated by variations of the renormalisation and factorisation scales using fixed-order calculations with MCFM [190].

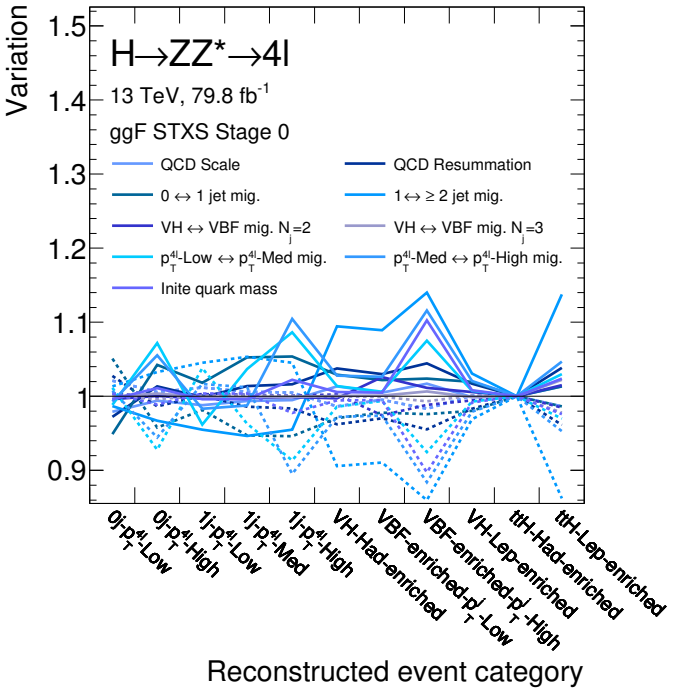
The uncertainties on the signal yields for the inclusive signal strength measurement are calculated in each reconstructed event category using Equation 5.7. The uncertainties on the signal acceptance in cross section measurements are determined for each particle level bin of Stage-0 and Reduced-Stage-1 using Equation 5.8. The resulting ggF uncertainties from QCD scale variations are shown in Figure 5.21(a) and Figure 5.21(b) for the inclusive signal strength measurement and the Stage-0 cross section measurement, respectively. The impact of all scale variations is symmetric around the nominal value of 1. The *up* variations of signal predictions are obtained by $1+\delta$ (solid lines), while *down* variation are given by $1-\delta$ (dashed lines). In general, the uncertainties for the cross section measurement in the particle level bins are smaller than for the signal strength measurement, since the signal yield uncertainty cancel out. For the signal strength measurement the size of uncertainties is up to 15%. The largest contribution come from the N_j and p_T^H bin migration in the corresponding reconstructed event categories. The VH and VBF migration uncertainties have a small impact of less than 2%. Due to the small number of events, no systematic uncertainties are assigned in the ttH-Had-enriched category.

In case of the production cross section measurement in particle level bins, the largest uncertainties also originate from N_j and p_T^H bin migration uncertainties with contributions of up to 14%. Minor impact have the QCD scale and VH and VBF migration uncertainties. As for the signal strength measurement no systematic uncertainty is assigned in the ttH-Had-enriched category.

For the other production modes, i.e. VBF, VH and ttH production, the uncertainties due to missing higher-order calculations can be simply evaluated by varying the QCD renormalisation and factorisation scales. To calculate the uncertainty, the nominal choice of the scale $\mu_{0,R}$ and $\mu_{0,F}$ is compared to different pairs of QCD renormalisation and factorisation variations of $\mu = \mu_0/2$ or $\mu = 2 \cdot \mu_0$. Eight combinations are considered in

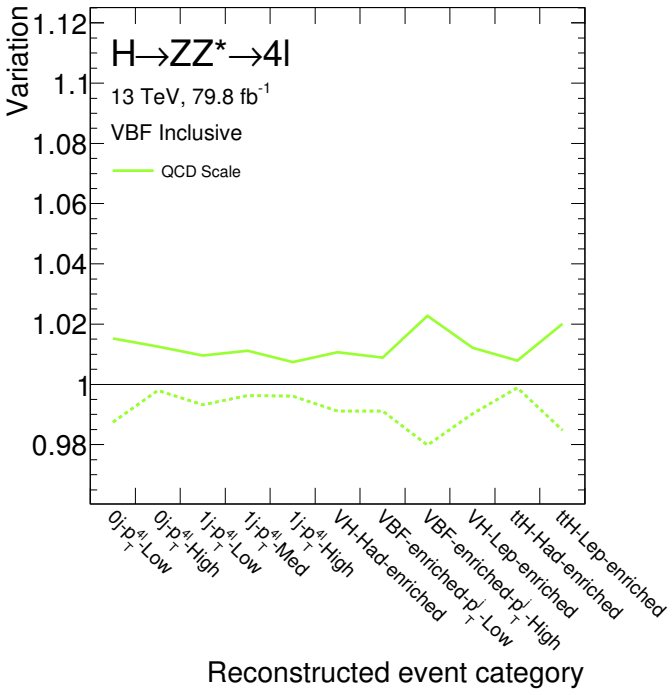


(a)

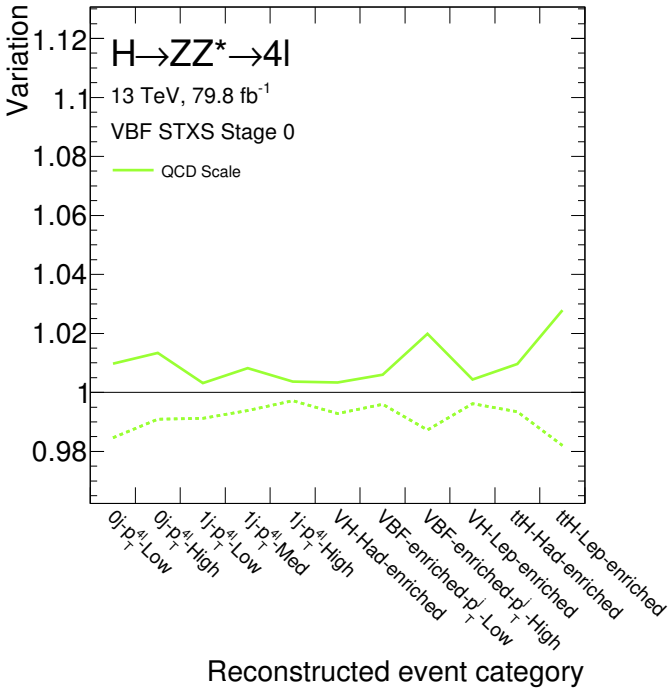


(b)

Figure 5.21: Systematic uncertainties from QCD scale variations in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for ggF production. *Up* (*down*) variations are shown as solid (dashed) lines.

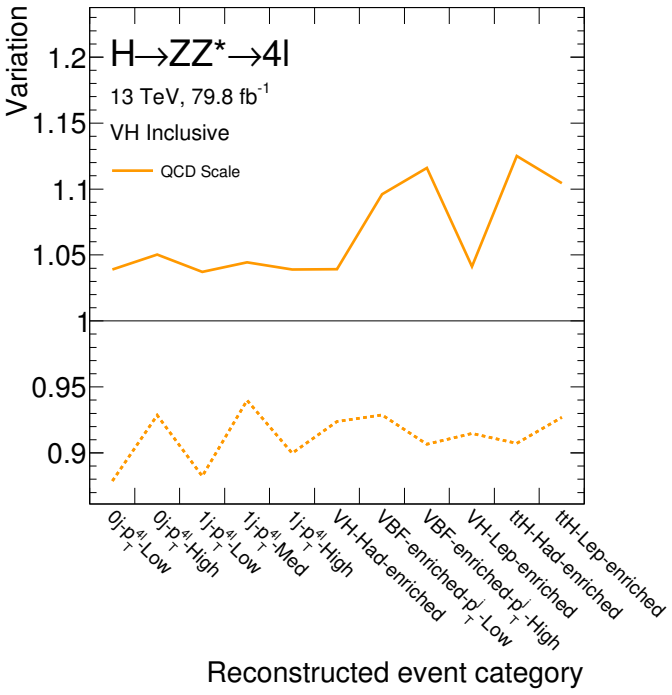


(a)

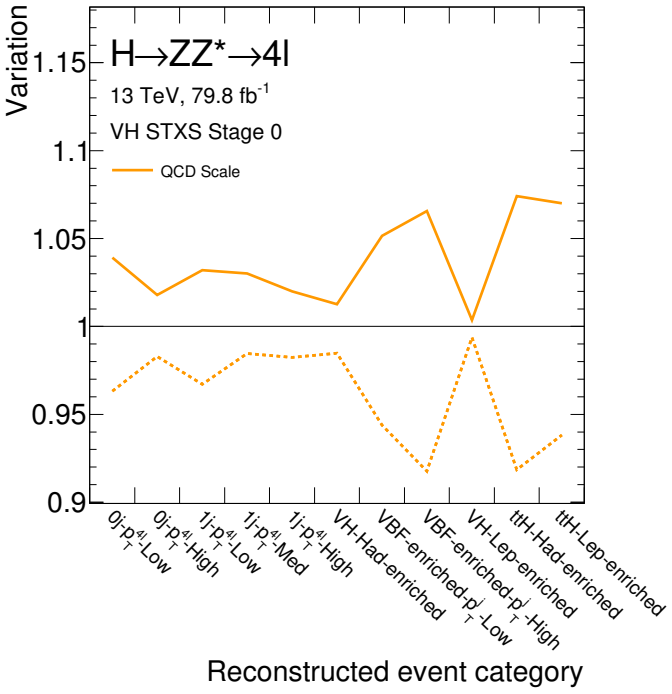


(b)

Figure 5.22: Systematic uncertainties from QCD scale variations in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for VBF production. *Up* (*down*) variations are shown as solid (dashed) lines.

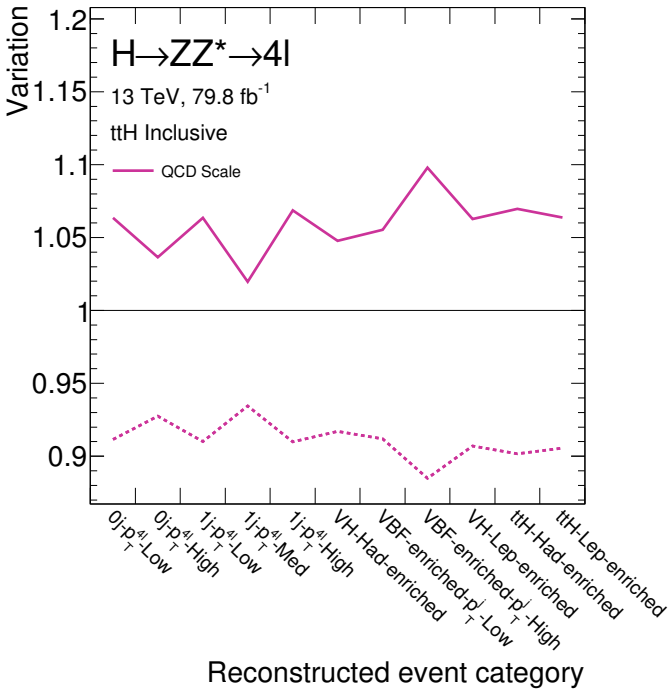


(a)

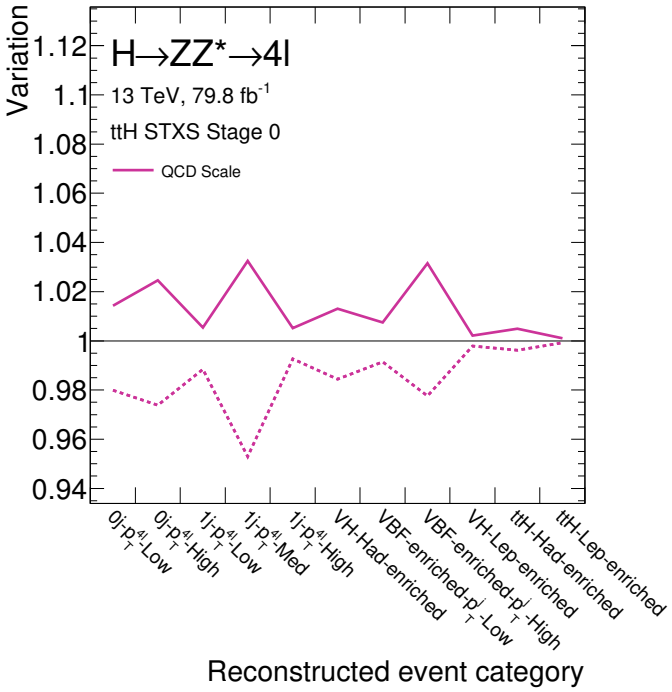


(b)

Figure 5.23: Systematic uncertainties from QCD scale variations in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for VH production. *Up* (*down*) variations are shown as solid (dashed) lines.



(a)



(b)

Figure 5.24: Systematic uncertainties from QCD scale variations in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for ttH production. *Up* (*down*) variations are shown as solid (dashed) lines.

total:

$$
\begin{aligned}
&(0.5 \cdot \mu_{0,R}, 0.5 \cdot \mu_{0,F}) \qquad && (2.0 \cdot \mu_{0,R}, 0.5 \cdot \mu_{0,F}) \\
&(0.5 \cdot \mu_{0,R}, 2.0 \cdot \mu_{0,F}) \qquad && (2.0 \cdot \mu_{0,R}, 2.0 \cdot \mu_{0,F}) \\
&(1.0 \cdot \mu_{0,R}, 0.5 \cdot \mu_{0,F}) \qquad && (0.5 \cdot \mu_{0,R}, 1.0 \cdot \mu_{0,F}) \\
&(1.0 \cdot \mu_{0,R}, 2.0 \cdot \mu_{0,F}) \qquad && (2.0 \cdot \mu_{0,R}, 1.0 \cdot \mu_{0,F})
\end{aligned}
$$

The final assigned uncertainty in each of the reconstructed event categories corresponds to the envelope of the uncertainties from all variations, i.e. the contribution with the largest uncertainty of all *up* and *down* variations. By definition such uncertainties are in general not symmetric.

The resulting VBF uncertainties from QCD scale variations are shown in Figure 5.22(a) and Figure 5.22(b) for the signal strength and the Stage-0 cross section measurement, respectively. Corresponding results for the VH production are shown in Figure 5.23(a) and Figure 5.23(b) and for ttH production in Figure 5.24(a) and Figure 5.24(b). The QCD scale uncertainty in the VBF production mode is in the order of 2% and is dominated by the yield uncertainties. The uncertainty on the VH production corresponding to the *up* variations are about 5% (3%) for the signal strength (production cross section) measurement in the categories with larger amount of data, while they reach up to 10% (7%) in categories with fewer events. The *down* variations have a stronger impact on the signal strength measurement and are in the order of 10% even in high-statistics 0-jet and 1-jet categories, as well as in the VH-Had-enriched category. The impact of QCD scale uncertainties on the ttH signal yield in the signal strength measurement is up to 10%, while the acceptance uncertainties for the Stage-0 cross section measurement are significantly smaller with values up to 2%.

5.4.2.2 PDF and Strong Coupling Constant

The choice of the PDF set has non-trivial effects on the calculations of cross sections and the kinematic properties of the final state produced in proton-proton collisions.

There are two approaches to evaluate the impact of the systematic uncertainties related to the choice of the PDF set: using internal variations (eigenvectors) of a given PDF set, or event generation with different choices of PDFs which is then compared to the sample with the nominal PDF set. The PDF uncertainties on the cross sections for the inclusive signal production from the latter approach are reported in Ref. [189]. The first approach is used in this analysis due to the splitting of events in exclusive jet and p_T^H event categories. The uncertainties are evaluated from the set of PDF eigenvectors within the nominal PDF set, which are orthogonal to each other and spam the PDF parameter space [188]. To allow for a combination with other Higgs boson decay channels, which are

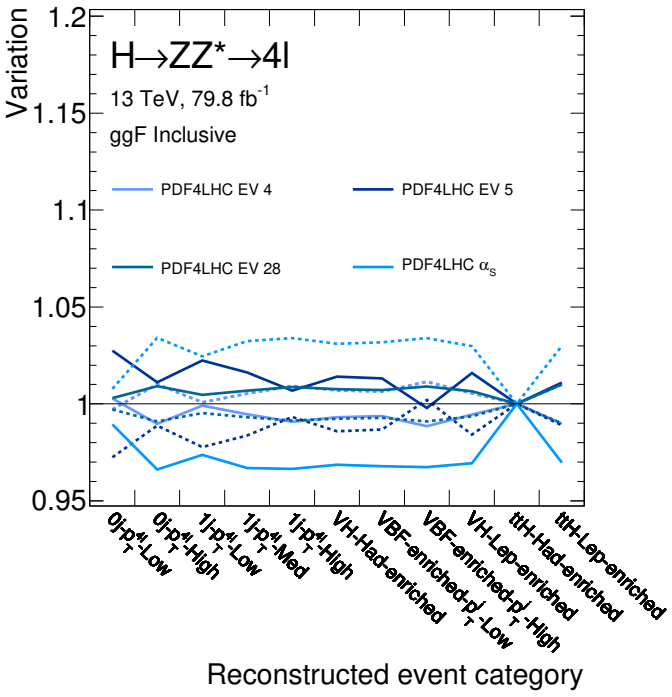
more sensitive to the effects of PDF variations, a consistent set of nuisance parameters (see Section 5.5) is needed. Therefore, the individual uncertainties from each eigenvector are propagated as a separate systematic uncertainty, rather than using an envelope of all individual uncertainties. The uncertainties are evaluated with the PDF4LHC_nlo_30 PDF set [88]. In total 30 eigenvector variations are available. The PDF set additionally contains two variations of the strong coupling constant, $\alpha_s = 0.1180 \pm 0.0015$, which are also taken into account.

The PDF uncertainties from individual eigenvectors, as well as the uncertainties from the variation of the strong coupling constant are shown in Figure 5.25,5.26,5.27 and 5.28 separately for the signal strength and the Stage-0 production cross section measurement. Since only systematic uncertainties larger than 1% are propagated through the analysis, the variations having smaller effects are not displayed. The variations of the individual PDF eigenvectors are treated as symmetric uncertainties, i.e. *up* variations are defined as $1+\delta$ (solid lines) and *down* variations as $1-\delta$ (dashed lines). The variation of the strong coupling constant $\alpha_s = 0.1180 + 0.0015$ corresponds to the *up* and $\alpha_s = 0.1180 - 0.0015$ to the *down* variation.

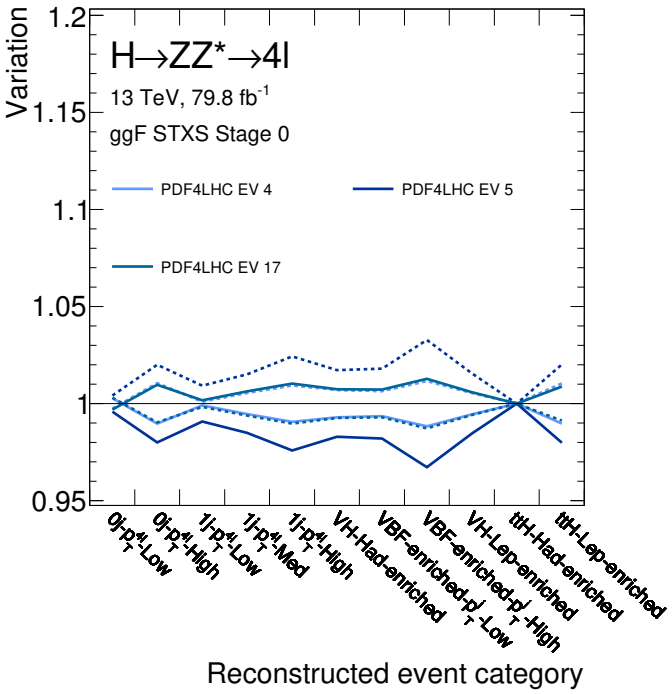
The PDF+α_s uncertainties in the ggF production mode amount up to 4% (3%) for the signal strength (production cross section) measurement. The variations with the largest impact on the signal strength measurement is the α_s uncertainty (except in the $0j$-$p_T^{4\ell}$-Low category), while for the cross section measurement the major uncertainty originates from one of the individual PDF eigenvector variations. No uncertainty is assigned for the ttH-Had-enriched category due to a small number of expected ggF events in this category.

The impact of the PDF+α_s uncertainty on the VBF production mode is smaller compared to ggF. For the signal strength measurement the largest uncertainty is 3% in the reconstructed $0j$-$p_T^{4\ell}$-High category. In some of the categories the uncertainties are smaller than 1% and therefore not propagated through the analysis. For the cross section measurement only the uncertainties in the $0j$-$p_T^{4\ell}$-High, ttH-Had-enriched and ttH-Lep-enriched which are in the order of 2% are taken into account. In the other categories the variations are smaller than 1%.

The size of the PDF+α_s uncertainties in the VH and ttH production mode is about 2% and 1% for the signal strength and cross section measurement, respectively.

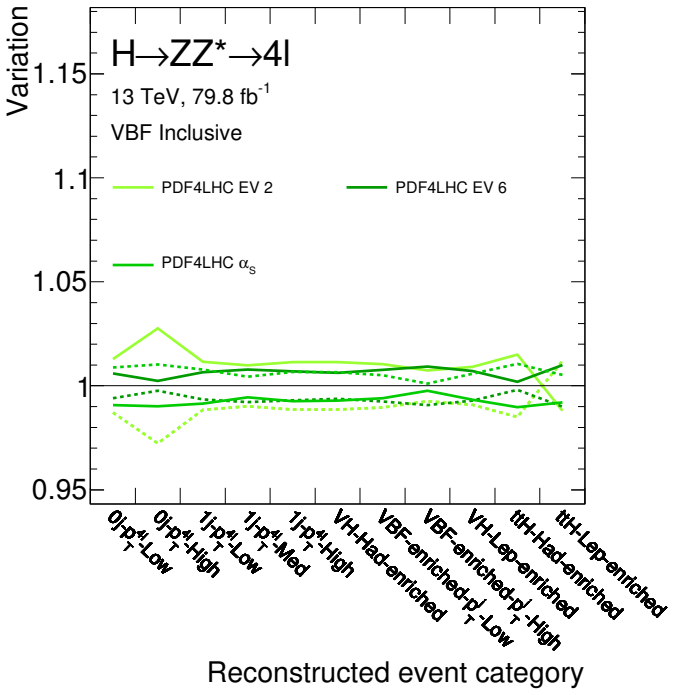


(a)

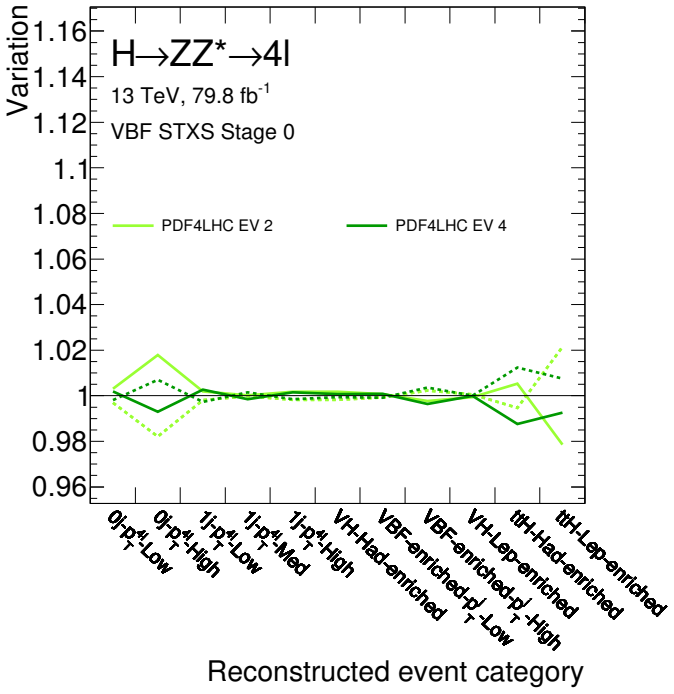


(b)

Figure 5.25: Systematic uncertainties from PDF+α_s variations ($> 1\%$) in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for ggF production. *Up* (*down*) variations are shown as solid (dashed) lines.

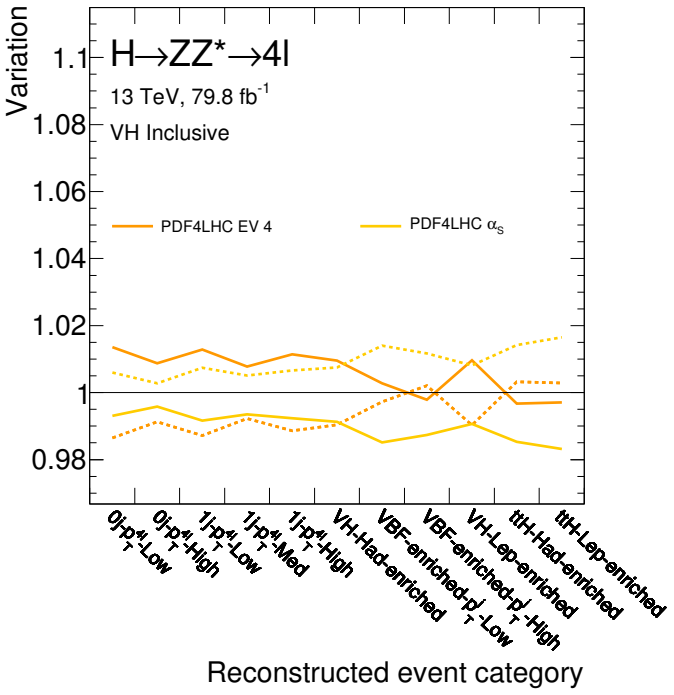


(a)

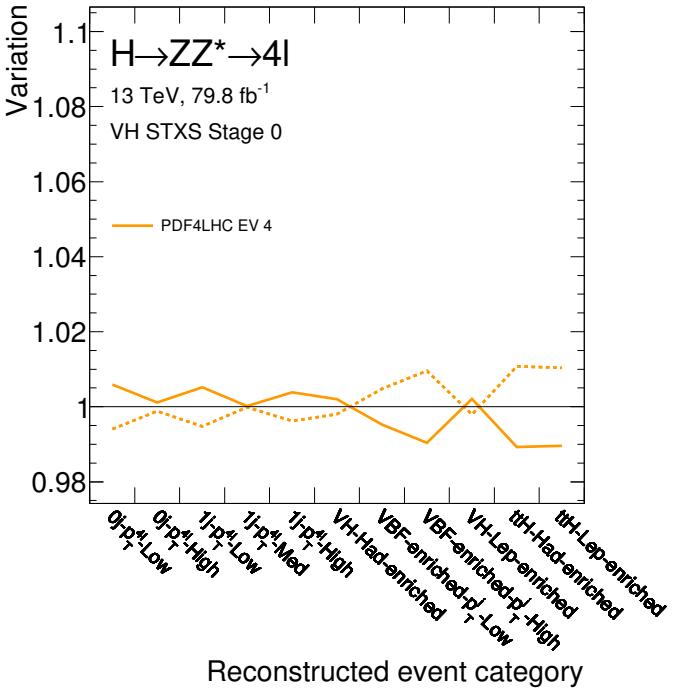


(b)

Figure 5.26: Systematic uncertainties from PDF+α_s variations ($> 1\%$) in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for VBF production. *Up* (*down*) variations are shown as solid (dashed) lines.

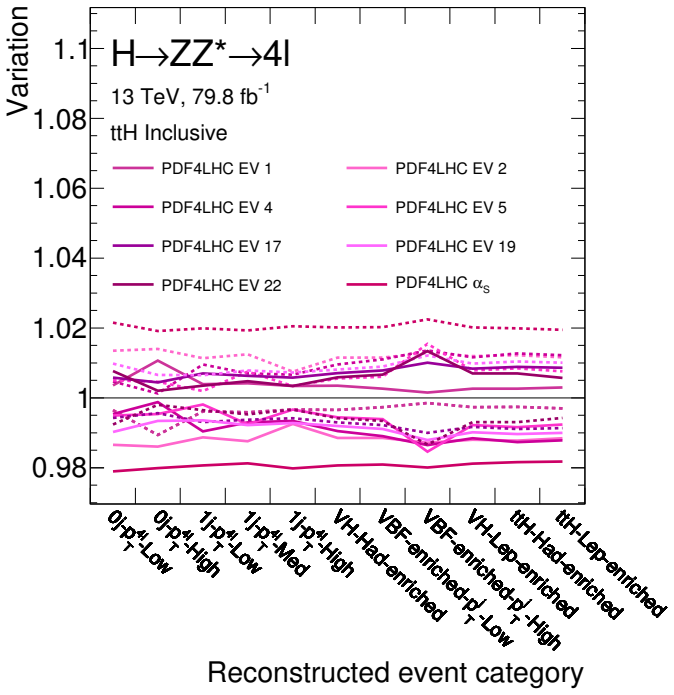


(a)

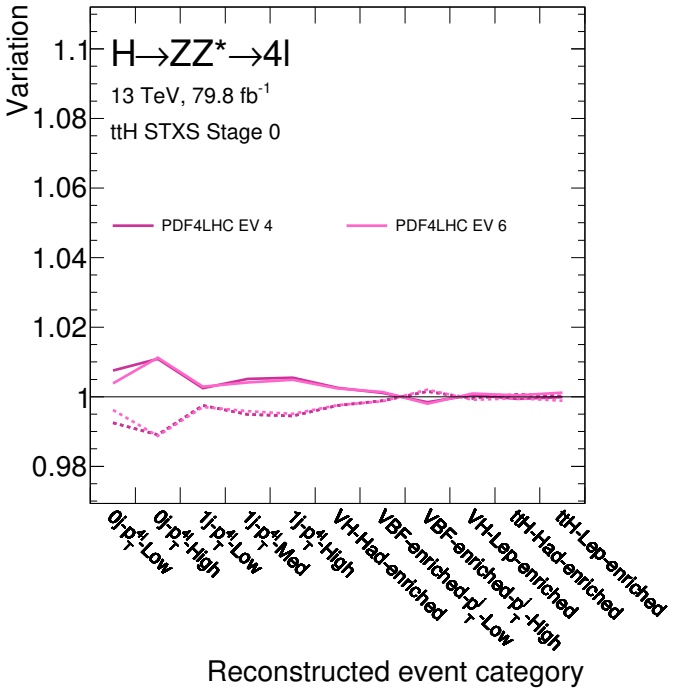


(b)

Figure 5.27: Systematic uncertainties from PDF+α_s variations ($> 1\%$) in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for VH production. *Up* (*down*) variations are shown as solid (dashed) lines.



(a)



(b)

Figure 5.28: Systematic uncertainties from PDF+α_s variations ($> 1\%$) in each of the reconstructed event categories for the (a) signal strength and (b) Stage-0 cross section measurement for ttH production. *Up* (*down*) variations are shown as solid (dashed) lines.

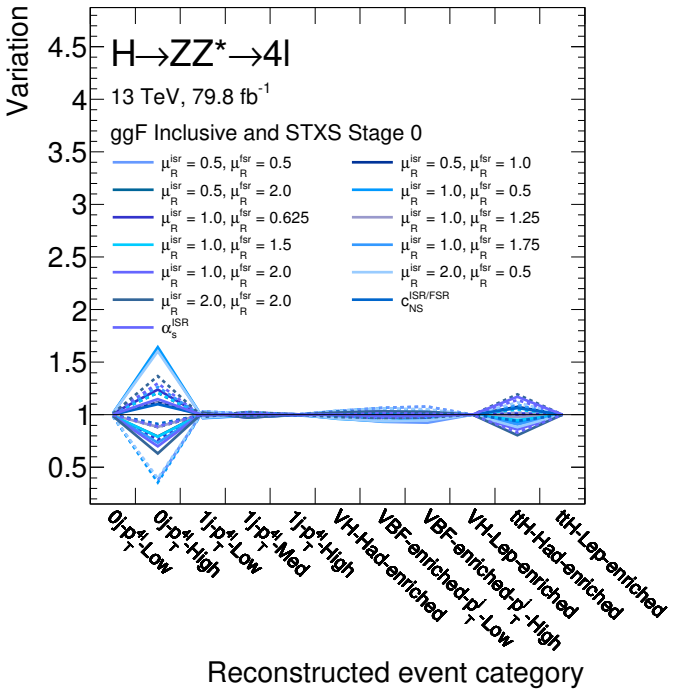
5.4.2.3 Parton Shower Simulation

The impact of parton shower uncertainties can be evaluated with two possible approaches, using internal parameter variations within the nominal parton shower generator tune, or using an alternative parton shower generator. Since it is not expected that a different generator configuration or the choice of the generator systematically shift the total cross section of the process, the per-category uncertainties are assumed to emerge entirely due to acceptance effects for both the signal strength and the cross section measurement. Since tune variations were not available in the reconstructed signal samples, particle level samples are used instead for the evaluation.

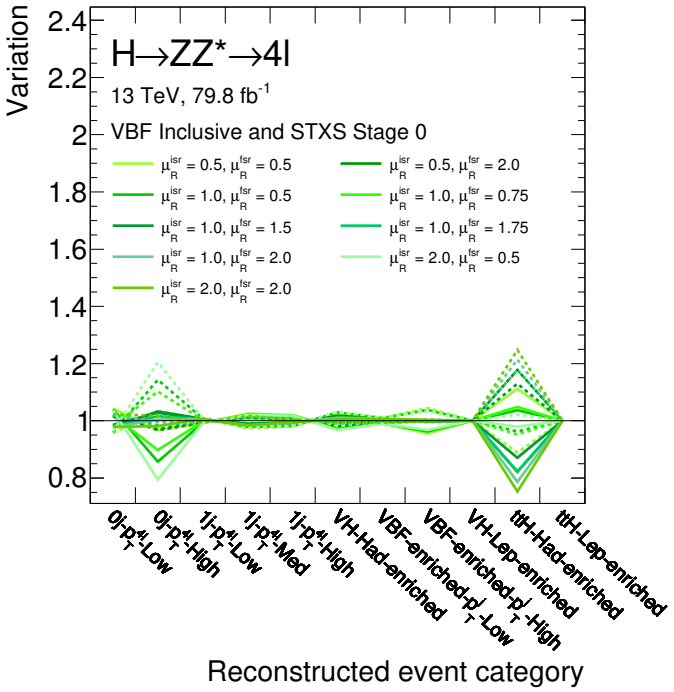
The uncertainties from the parton shower tune in the nominal generator are estimated using the automated shower variations in Pythia8 [191]. These include the variation of the renormalisation scales μ_R^{FSR} and μ_R^{ISR} for QCD emissions in final and initial state radiations, respectively. In total 20 combinations of the renormalisation scales $\left(f^{FSR}\cdot\mu_R^{FSR}, f^{ISR}\cdot\mu_R^{ISR}\right)$ are considered. The multiplicative factors f^{FSR} and f^{ISR} are varied in the range between 0.5 and 2.0. In addition variations of non-singular terms for QCD emissions in final and initial state radiations $\left(c_{NS}^{FSR/ISR}\right)$ are taken into account. The *up* variation of the nominal parameters is given by the combination $\left(2\cdot c_{NS}^{FSR}, 2\cdot c_{NS}^{ISR}\right)$, the corresponding *down* variation is $\left(-2\cdot c_{NS}^{FSR}, -2\cdot c_{NS}^{ISR}\right)$. Furthermore, the variations accounting for the $t\bar{t}$ gap, dijet decorrelations and the Z boson p_T are taken into account by using, $\alpha_s^{ISR} = 0.155$ for the *up* variation and $\alpha_s^{ISR} = 0.140$ for the *down* variation. Each of the variations is treated as an individual uncertainty. The uncertainties of the renormalisation scale are implemented as symmetric variations.

The parton shower uncertainties, which affect only the acceptance in the reconstructed event categories are shown in Figure 5.29 and Figure 5.30 separately for the ggF, VBF, VH and ttH production mode. The uncertainties are relatively large compared to the QCD scale and PDF uncertainties. Therefore, for presentation properties only uncertainties with an impact larger than 10% are shown. For the ggF production mode, the uncertainties reach up to 70% in the $0j$-$p_T^{4\ell}$-High category, while they are in the order of 10% for the remaining categories. For the VBF, VH and ttH production mode, the impact is smaller, reaching $5-10\%$ except for the categories with low number of events.

The parton shower uncertainties have also been evaluated with an alternative approach using a different parton shower generator. The same generated parton-level events are showered using two different parton shower generators. As described in Section 4.2, the default SM signal samples are showered with Pythia8 [131]. For the ggF, VBF and VH production modes the alternative generator used for the comparison is Herwig7 [192]. For the ttH production mode Herwig++ [134] is used.

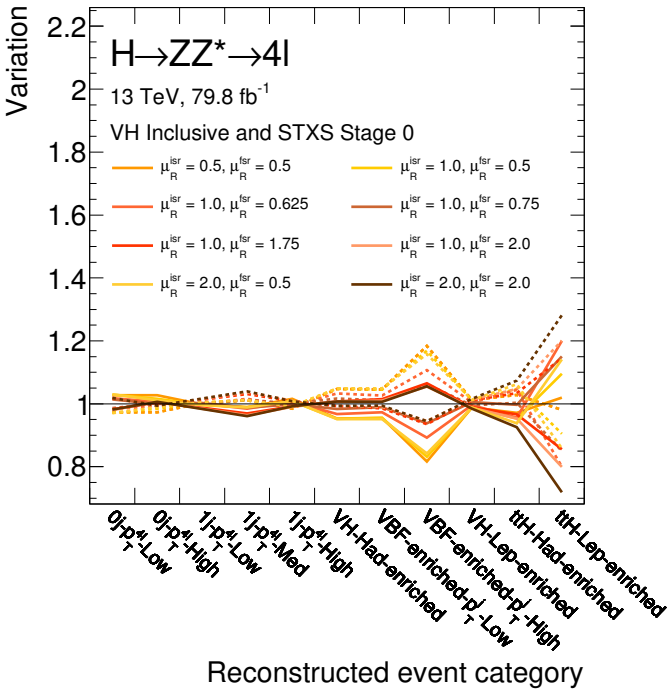


(a)

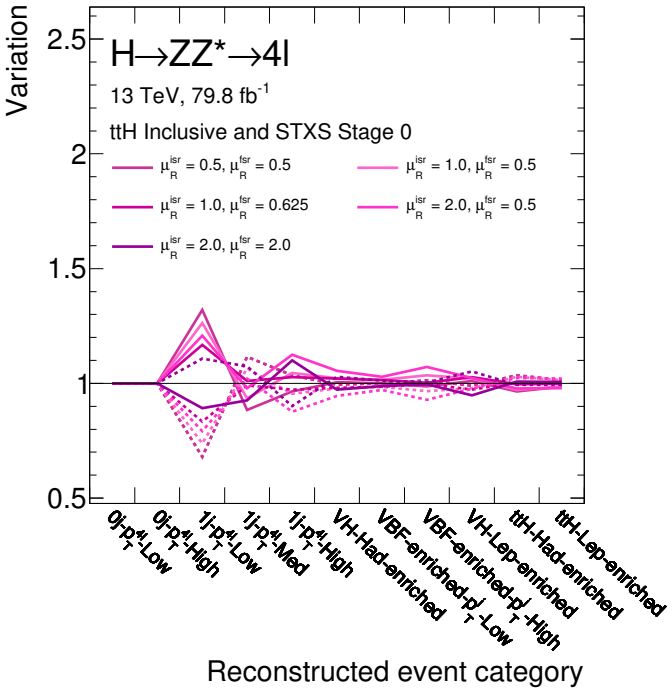


(b)

Figure 5.29: Systematic uncertainties from parton shower variations ($> 10\%$) in each of the reconstructed event categories for the (a) ggF and (b) VBF production. *Up* (*down*) variations are shown as solid (dashed) lines.

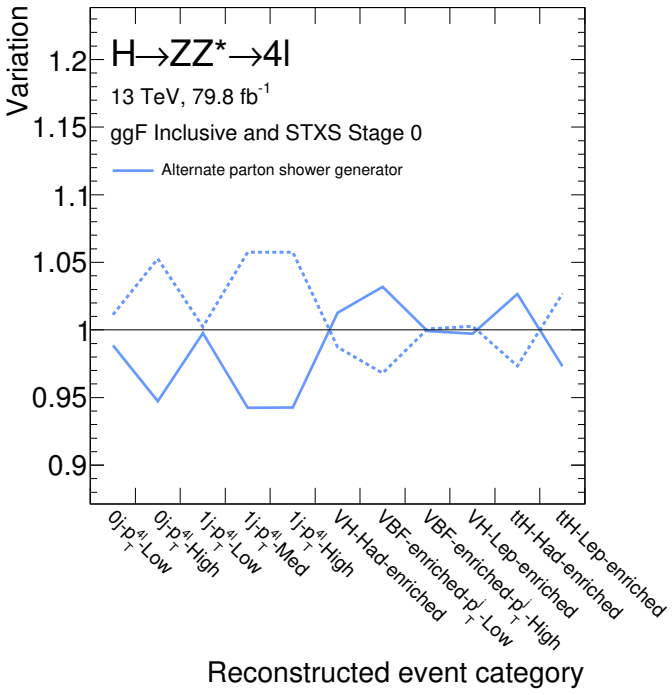


(a)

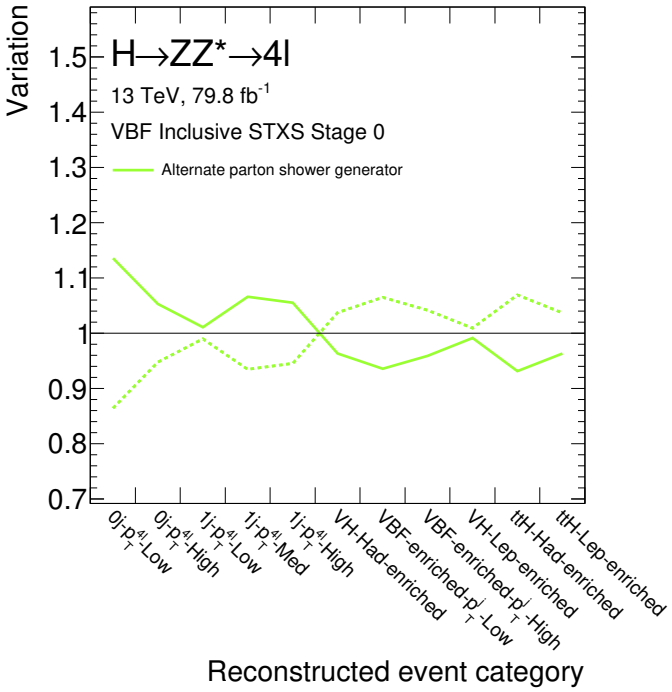


(b)

Figure 5.30: Systematic uncertainties from parton shower variations ($> 10\%$) in each of the reconstructed event categories for the (a) VH and (b) ttH production. *Up* (*down*) variations are shown as solid (dashed) lines.

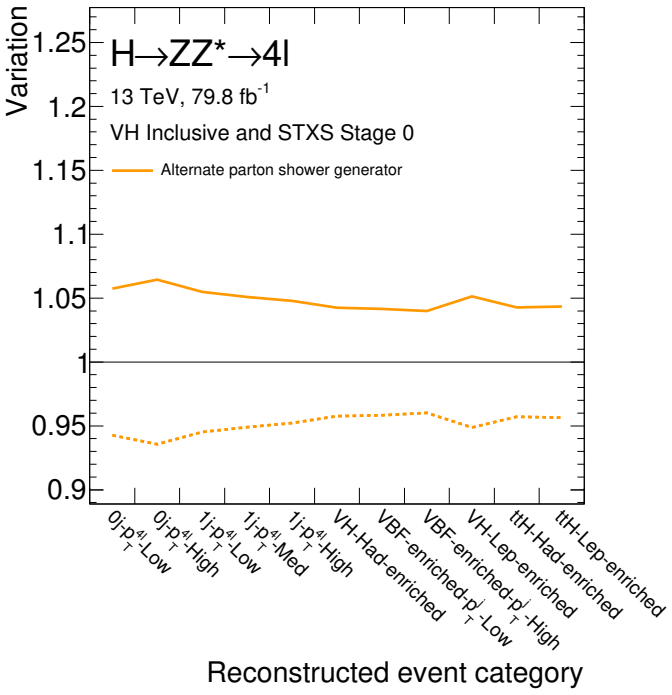


(a)

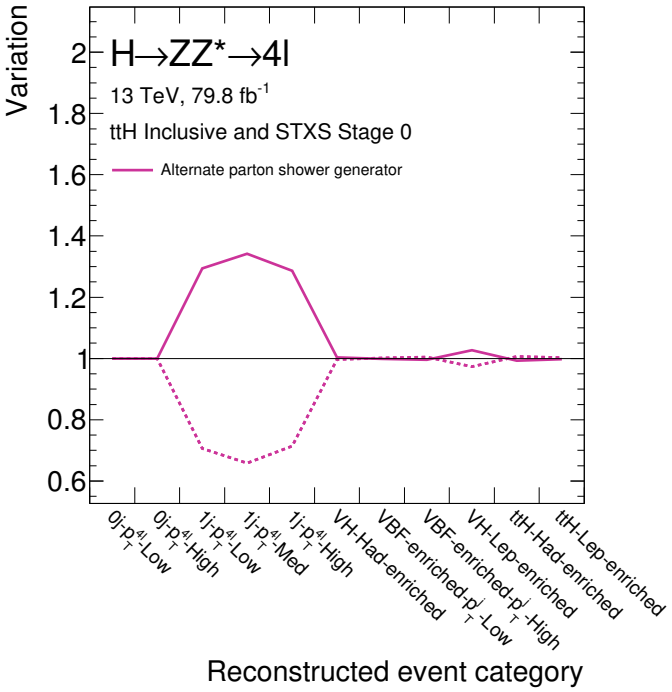


(b)

Figure 5.31: Systematic uncertainties from alternate parton shower variations in each of the reconstructed event categories for the (a) ggF and (b) VBF production. *Up* (*down*) variations are shown as solid (dashed) lines.



(a)



(b)

Figure 5.32: Systematic uncertainties from alternate parton shower variations in each of the reconstructed event categories for the (a) VH and (b) ttH production. *Up* (*down*) variations are shown as solid (dashed) lines.

In this approach of comparing different generators one difficulty arises. There is a possibility that one and the same event will enter a different particle level or reconstructed bin after it is showered by two different generators. Therefore, the change of the generator can cause events to migrate between different particle level or reconstruction categories, which can lead to uncertainties of up to ±100% in depopulated bins, when the uncertainty for a given category is defined as the ratio of the number of events from HERWIG and PYTHIA8 generators in each of the particle level bins. In order to avoid such large uncertainties due to statistical fluctuations and to derive a more reliable estimate of the alternate parton shower generator uncertainty, a particle level reweighting of PYTHIA8 events to the HERWIG prediction is used. The reweighting is performed with two-dimensional histograms $\left(\eta_{4\ell}, p_T^{4\ell}\right)$ in particle level bins of the Reduced-Stage-1 derived separately for each of the two parton shower generators. Scale factors are derived in each histogram bin to reweight the particle level Higgs boson p_T and $\eta_{4\ell}$ distributions from PYTHIA8 to match the ones from HERWIG. The scale factors are then applied to events in the corresponding nominal SM reconstruction-level samples. In this way, the impact of the described statistical fluctuations is mitigated.

The uncertainties obtained from alternate parton shower generator are Figure 5.31 and Figure 5.32 for both the signal strength and Stage-0 production cross section measurement. On average, the uncertainties amount to about 5% for all production modes. For the VBF and ttH processes these uncertainties reach up to 30% in categories with a small number of events from these process.

5.5 Statistical Analysis Model

The measurement of production cross sections in particle level bins, as well as the measurement of the corresponding signal strengths is performed via a binned maximum-likelihood fit to the data using the profiled likelihood ratio method [193].

The likelihood function is a product of Poisson distributions P corresponding to the observed number of events $n_{i,j}$ in each bin i of the discriminant observable in a given reconstructed category j, for a given signal $s_{i,j}(\boldsymbol{\theta}) = \boldsymbol{\sigma} \cdot \mathcal{B} \cdot \mathcal{A}_{i,j}(\boldsymbol{\theta})$ and background $b_{i,j}(\boldsymbol{\theta})$ expectation,

$$\mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\theta}) = \prod_{j}^{N_{\text{Categories}}} \prod_{i}^{N_{\text{Bins}}} P\left(n_{i,j} \mid L \cdot \boldsymbol{\sigma} \cdot \mathcal{B} \cdot \mathcal{A}_{i,j}(\boldsymbol{\theta}) + b_{i,j}(\boldsymbol{\theta})\right) \times \prod_{m}^{N_{\text{Nuisance}}} C_m(\boldsymbol{\theta}). \quad (5.9)$$

The vector $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ represents the Higgs boson production cross sections in different particle level bins, which are treated as independent parameters of interest and correlated among the different reconstructed event categories. $\mathcal{B}$ is the branching ratio

and the vector $\mathcal{A}_{i,j}$ corresponds to the signal acceptance in each reconstructed event category j and particle level production bin i. The vector $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_N\}$ represents the nuisance parameters corresponding to the systematic uncertainties described in Section 5.4. The nuisance parameters are constrained by a Gaussian constraint term $C_m(\boldsymbol{\theta})$ which represents an external measurement of the corresponding parameter.

The test statistic q used for the comparison of different fitted hypotheses is the ratio of profiled likelihoods,

$$q = -2\ln\frac{\mathcal{L}\left(\boldsymbol{\sigma}, \hat{\hat{\boldsymbol{\theta}}}(\boldsymbol{\sigma})\right)}{\mathcal{L}\left(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}\right)} = -2\ln(\lambda). \tag{5.10}$$

$\mathcal{L}(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}})$ is the maximum-likelihood estimator of an unconditional fit in which both $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$ parameters are free parameters of the fit. $\mathcal{L}(\boldsymbol{\sigma}, \hat{\hat{\boldsymbol{\theta}}}(\boldsymbol{\sigma}))$ is the likelihood estimator of a conditional fit with profiled parameters of interest $\boldsymbol{\sigma}$, i.e. with parameters $\boldsymbol{\sigma}$ fixed to a given value in the fit. The nuisance parameters $\hat{\hat{\boldsymbol{\theta}}}(\boldsymbol{\sigma})$ denote the free-floating parameters in such a conditional fit. The best-fit values of parameters $\boldsymbol{\sigma}$ are obtained from the scan of the test statistic q over a range of these parameters of interest, assuming that the asymptotic approximation [193] is valid. The smaller the value q, the better the agreement of the data with the tested hypotheses of parameter values $\boldsymbol{\sigma}$ is expected. The best-fit result corresponds to the minimal value of the test statistic q.

Alternatively, the parameter of interest in each particle level production bin i, σ_i, is replaced by $\mu_i \sigma_{i,\mathrm{SM}}(\boldsymbol{\theta})$, allowing for the measurement of the signal strength μ_i relative to the SM prediction $\sigma_{i,\mathrm{SM}}(\boldsymbol{\theta})$. In this case, the measurement includes the uncertainties on the signal yield, while in the former case only the uncertainty on the signal acceptance is relevant.

5.6 Results

The observed and expected numbers of events after the full $H \to ZZ^* \to 4\ell$ event selection (Section 4.3) and classification into the reconstructed event categories are shown in Table 7.10. As discussed in Section 4.5, a slight excess of Higgs boson candidates is observed compared to the SM prediction. The largest excess is observed in the two VBF-enriched categories. No events are observed in the $0j$-$p_T^{4\ell}$-High and in the two ttH-enriched categories.

Table 5.6: Expected and observed number of events per reconstructed $H \to ZZ^* \to 4\ell$ event category for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV after the inclusive event selection in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV. Statistical and systematic uncertainties are added in quadrature [183].

Reconstructed event category	SM Higgs boson signal	ZZ^*	Other backgrounds	Total expected	Observed
$0j$-$p_T^{4\ell}$-Low	56 $\pm$5	35.2 $\pm$2.5	4.1 $\pm$0.4	95 $\pm$ 6	112
$1j$-$p_T^{4\ell}$-Low	17.9 $\pm$2.2	7.6 $\pm$1.1	1.36 $\pm$0.14	26.9 $\pm$ 2.5	25
$1j$-$p_T^{4\ell}$-Med	11.5 $\pm$1.5	2.25 $\pm$0.32	0.72 $\pm$0.08	14.5 $\pm$ 1.5	14
$1j$-$p_T^{4\ell}$-High	3.3 $\pm$0.5	0.31 $\pm$0.05	0.198 $\pm$0.028	3.8 $\pm$ 0.5	4
VBF-enriched-p_T^j-Low	12.8 $\pm$1.6	2.8 $\pm$0.7	1.04 $\pm$0.10	16.6 $\pm$ 0.8	24
VBF-enriched-p_T^j-High	1.25 $\pm$0.20	0.28 $\pm$0.09	0.155 $\pm$0.025	1.68 $\pm$ 0.23	4
VH-Had-enriched	7.2 $\pm$1.1	1.6 $\pm$0.4	0.59 $\pm$0.07	9.4 $\pm$ 1.2	11
VH-Lep-enriched	0.70 $\pm$0.05	0.068 $\pm$0.013	0.035 $\pm$0.008	0.80 $\pm$ 0.05	1
$0j$-$p_T^{4\ell}$-High	0.183 $\pm$0.025	0.0082 $\pm$0.0027	0.164 $\pm$0.023	0.355 $\pm$ 0.034	0
ttH-Had-enriched	0.60 $\pm$0.05	0.035 $\pm$0.016	0.194 $\pm$0.030	0.83 $\pm$ 0.06	0
ttH-Lep-enriched	0.238 $\pm$0.026	0.0005 $\pm$0.0005	0.0144 $\pm$0.0034	0.253 $\pm$ 0.026	0
Total	112 $\pm$5	50 $\pm$4	8.96 $\pm$0.12	171 $\pm$61	195

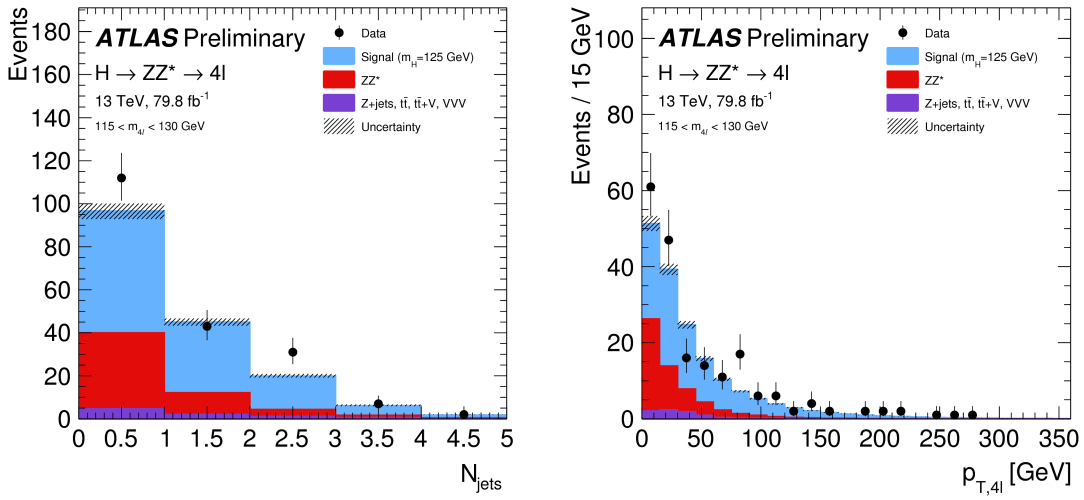
The expected and observed distributions of the number of jets in the final state (N_j) and the transverse momentum of the four leptons in the final state ($p_T^{4\ell}$) used for the classification in reconstructed categories are shown in Figure 5.33. Both are in good agreement with the data.

The expected and observed distributions of the BDT are shown in Figure 5.34 and Figure 5.35. The small excess of observed events in the VBF-enriched-p_T^j-Low categories is observed at larger values of the BDT response in which the highest contribution of VBF production is expected. All other distributions are in good agreement with the data.

The inclusive $H \to ZZ^*$ production cross section within the Higgs boson rapidity $|y_H| < 2.5$ is measured under the assumption that the relative signal fractions in each particle level bin are given by the SM prediction, yielding

$$\sigma \cdot \mathcal{B} = 1.57 \pm 0.15 \text{ (stat.) } 0.08 \text{ (exp.) } \pm 0.04 \text{ (th.) pb} = 1.57 \pm 0.18 \text{ pb.} \tag{5.11}$$

This is in good agreement with the SM prediction $(\sigma \cdot \mathcal{B})_{\text{SM}} = 1.33 \pm 0.09$ pb.



(a) (b)

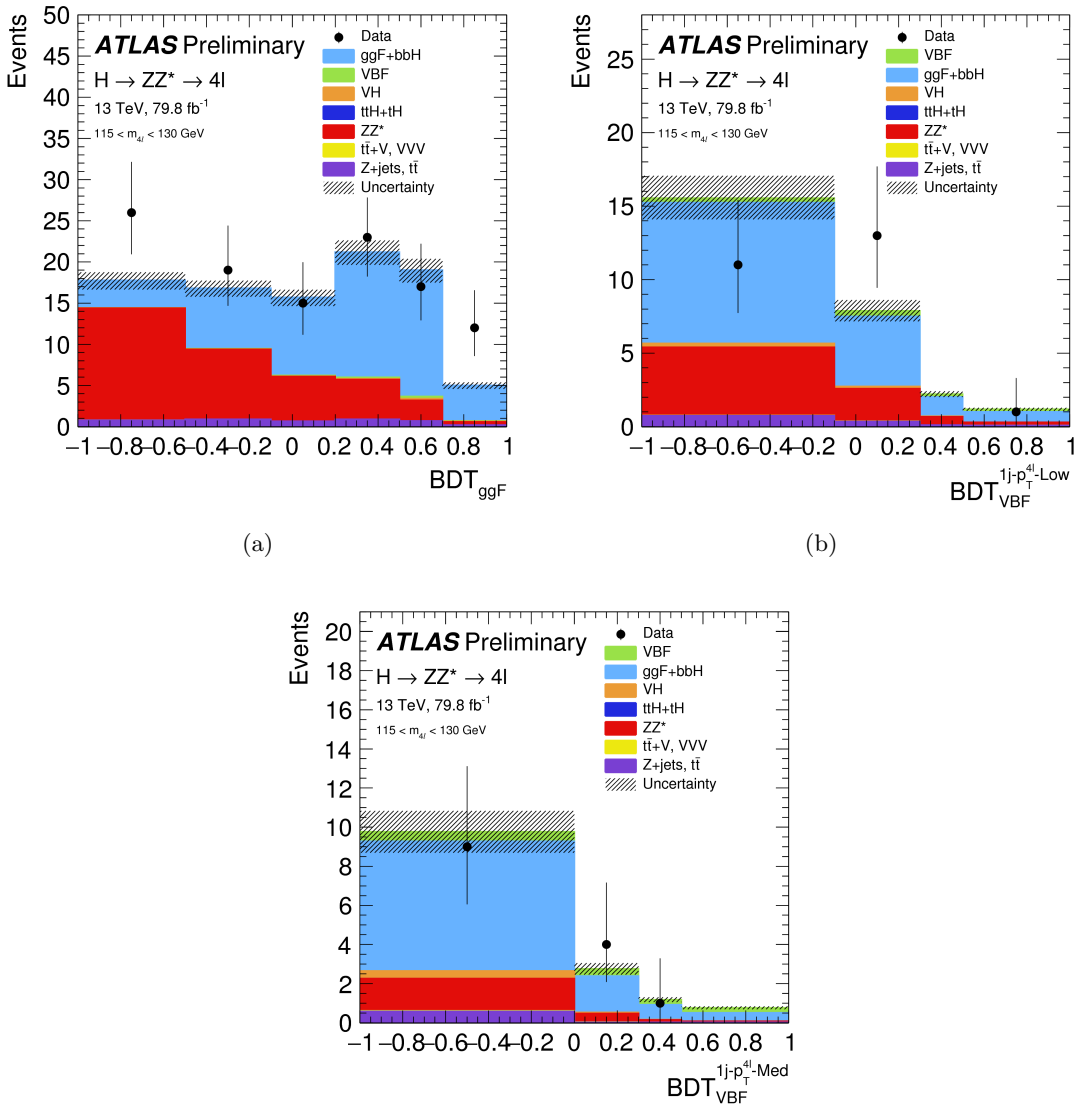
Figure 5.33: The observed and expected distribution of the (a) number of jets in the final state and (b) $p_T^{4\ell}$ for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV assuming a Higgs boson signal with a mass of $m_H = 125$ GeV [183].

In a similar way, the global signal strength is measured to be

$$\mu = 1.19 \pm 0.12\ (\text{stat.})\ \pm 0.06\ (\text{exp.})\ ^{+0.08}_{-0.07}\ (\text{th.}) = 1.19^{+0.16}_{-0.15}. \tag{5.12}$$

The measured signal strength and the inclusive production cross section agree within the SM prediction at a level of 1.4 and 1.5 standard deviations, respectively.

The corresponding observed and expected distributions of the test statistics (likelihood scans) for the signal strength and cross section measurement are shown in Figure 5.36(a) and Figure 5.36(b), respectively. The likelihood scans are shown before and after including systematic uncertainties. The signal strength measurement is equally affected by the experimental and theoretical systematic uncertainties, while for the cross section measurement the experimental uncertainties are dominant, since the uncertainty is on the signal yield.

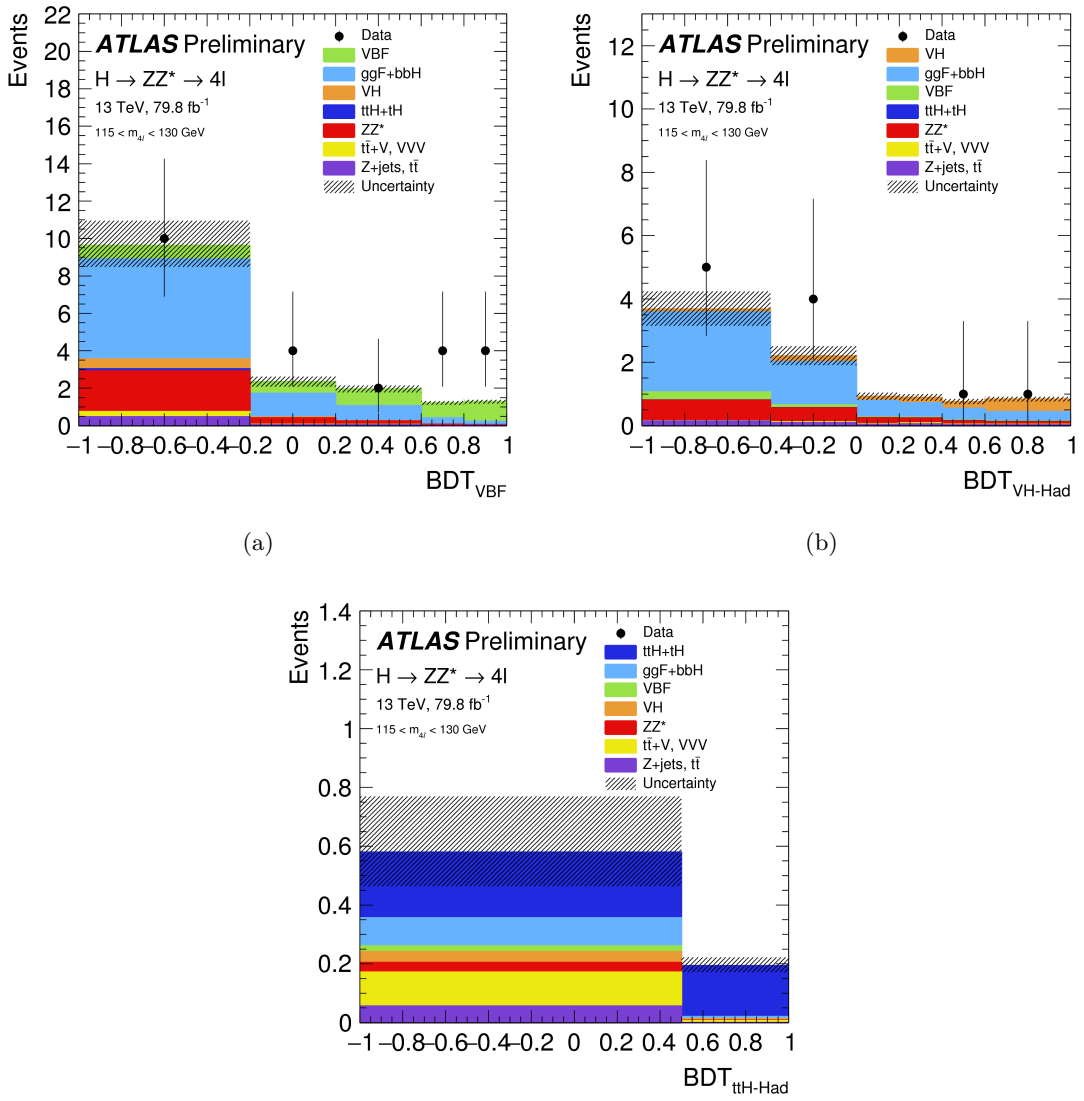


(a)

(b)

(c)

Figure 5.34: The observed and expected BDT output distributions in the (a) $0j$-$p_T^{4\ell}$-Low, (b) $1j$-$p_T^{4\ell}$-Low and (c) $1j$-$p_T^{4\ell}$-Med reconstructed event category for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV assuming a Higgs boson signal with a mass of $m_H = 125$ GeV [183].

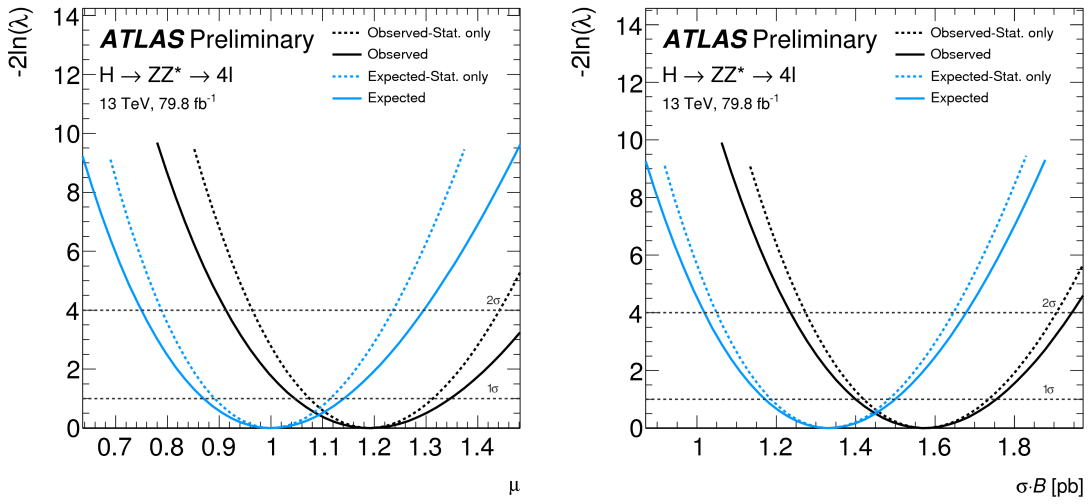


(a)

(b)

(c)

Figure 5.35: The observed and expected BDT output distributions in the (a) VBF-enriched-p_T^j-Low, (b) VH-Had-enriched and (c) ttH-Had-enriched reconstructed event category for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV assuming a Higgs boson signal with a mass of $m_H =$ 125 GeV [183].



(a) (b)

Figure 5.36: The expected (blue) and observed (black) test statistic $q = -2\ln(\lambda)$ for the (a) inclusive signal strength μ and (b) the $\sigma \cdot \mathcal{B}(H \to ZZ^*)$ measurement. The scans are shown both with (solid line) and without (dashed line) systematic uncertainties [183].

The expected and observed cross sections $\sigma \cdot \mathcal{B}$, as well as their ratio to the SM prediction are shown in Table 7.10 and in Figure 5.37 for the inclusive, Stage-0 and in Figure 5.38 for Reduced-Stage-1 particle level production bins. The ratio with respect to the SM prediction is calculated without the uncertainties on the SM prediction. Due to an excess of observed events compared to the expectation in the reconstructed event categories sensitive to the VBF production mode, a slight deviation of more than two standard deviations is observed for the VBF production cross section in the Stage-0 and Reduced-Stage-1 particle level production bins. All other production cross sections agree with the SM prediction within two standard deviations. Since no events are observed in the ttH-enriched categories, only upper limits are set on the cross section for the ttH production. To avoid negative total event yields in the ttH-enriched categories after the fit, the parameter of interest is constrained to be positive. The impact of this constraint on the final fit results is negligible.

Table 5.7: The expected SM cross section $(\sigma \cdot \mathcal{B})_{\mathrm{SM}}$, the observed cross section $(\sigma \cdot \mathcal{B})$ and their ratio $(\sigma \cdot \mathcal{B}) / (\sigma \cdot \mathcal{B})_{\mathrm{SM}}$ for the inclusive production, for the Stage-0 and Reduced-Stage-1 production bin for the $H \to ZZ^*$ decay for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV. The upper limits correspond to the 95% CL. The uncertainties are given as (stat.)+(exp.)+(th.) for the inclusive cross section and Stage-0, while they are given as (stat.)+(syst.) for Reduced-Stage-1, since the actual impact of theory uncertainties are smaller than the rounding.

Production bin	Cross section $(\sigma \cdot \mathcal{B})$ [pb]		$(\sigma \cdot \mathcal{B}) / (\sigma \cdot \mathcal{B})_{\mathrm{SM}}$
	SM expected	Observed	Observed
	Inclusive production, $\lvert y_H \rvert < 2.5$		
Inclusive	1.33 ± 0.09	$1.57 \pm 0.16 \pm 0.07 \pm 0.04$	$1.18 \pm 0.12 \pm 0.05 \pm 0.03$
	Stage-0 production bins, $\lvert y_H \rvert < 2.5$		
ggF	1.17 ± 0.08	$1.22 \pm 0.17 \pm 0.07 \pm 0.04$	$1.04 \pm 0.14 \pm 0.06 \pm 0.03$
VBF	0.0917 ± 0.0028	$0.25 \pm 0.08 \pm 0.02 \pm 0.01$	$2.8 \pm 0.9 \pm 0.2 \pm 0.2$
VH	$0.0524^{+0.0026}_{-0.0047}$	$0.05 \pm 0.05 \pm 0.01 \pm 0.01$	$0.9 \pm 1.0 \pm 0.1 \pm 0.1$
ttH	$0.0154^{+0.0011}_{-0.0016}$	< 0.07	< 4.04
	Reduced-Stage-1 production bins, $\lvert y_H \rvert < 2.5$		
ggF-0j	0.72 ± 0.05	$0.87 \pm 0.14 \pm 0.08$	$1.20 \pm 0.20 \pm 0.10$
ggF-1j-p_T^H-Low	0.170 ± 0.020	$0.10 \pm 0.10 \pm 0.05$	$0.59 \pm 0.55 \pm 0.27$
ggF-1j-p_T^H-Med	0.120 ± 0.020	$0.08 \pm 0.06 \pm 0.02$	$0.70 \pm 0.48 \pm 0.13$
ggF-1j-p_T^H-High	0.024 ± 0.005	$0.007 \pm 0.026 \pm 0.006$	$0.3 \pm 1.1 \pm 0.3$
ggF-2j	0.140 ± 0.030	$0.16 \pm 0.10 \pm 0.03$	$1.19 \pm 0.76 \pm 0.26$
VBF-p_T^j-Low	0.0872 ± 0.0027	$0.24 \pm 0.09 \pm 0.02$	$2.8 \pm 1.0 \pm 0.2$
VBF-p_T^j-High	$0.0041^{+0.0004}_{-0.0002}$	$0.03 \pm 0.02 \pm 0.01$	$7.5 \pm 6.0 \pm 0.7$
VH-Had	$0.0359^{+0.0019}_{-0.0033}$	$0.02 \pm 0.10 \pm 0.01$	$0.6 \pm 2.9 \pm 0.3$
VH-Lep	$0.0165^{+0.0008}_{-0.0014}$	$0.02 \pm 0.03 \pm 0.01$	$1.2 \pm 1.5 \pm 0.1$
ttH	$0.0154^{+0.0011}_{-0.0016}$	< 0.06	< 4.02

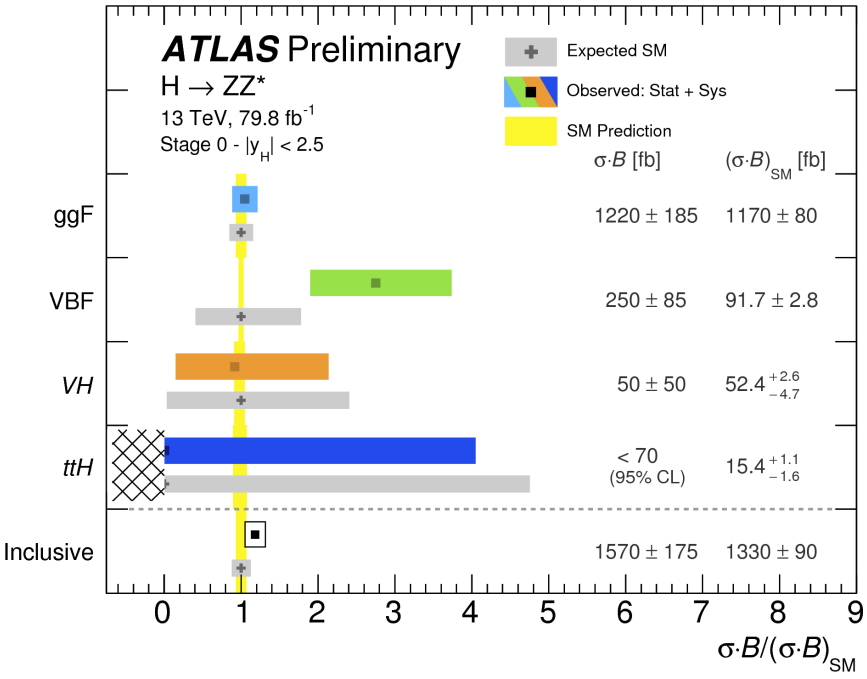


Figure 5.37: The observed and expected SM values of the cross section rations $\sigma \cdot \mathcal{B}$ normalised to the SM prediction $(\sigma \cdot \mathcal{B})_{\text{SM}}$ for the inclusive production and Stage-0 production bins for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV. The different colours of the observed results indicate different Higgs boson production modes. Hatched area indicates that the ttH parameter of interest is constrained to positive values. For visualisation purposes, the VBF-p_T^j-High value is divided by a factor of three. Yellow vertical bands represent theory uncertainty in the signal prediction, while horizontal grey bands represent the expected measurement uncertainty [183].

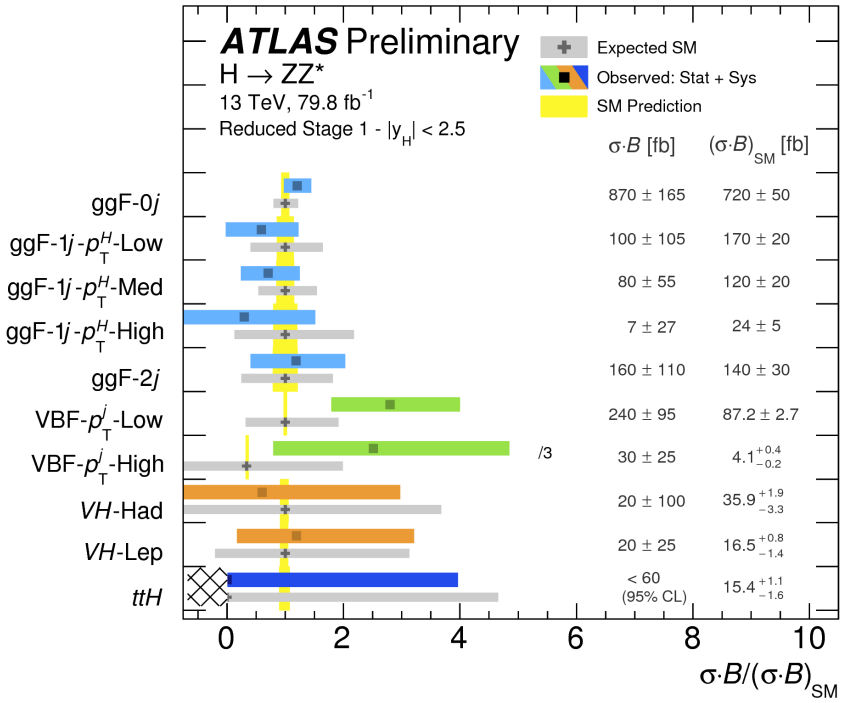


Figure 5.38: The observed and expected SM values of the cross section rations $\sigma \cdot \mathcal{B}$ normalised to the SM prediction $(\sigma \cdot \mathcal{B})_{\mathrm{SM}}$ for the Reduced-Stage-1 production bins for an integrated luminosity of 79.8 fb^{-1} at $\sqrt{s} = 13$ TeV. The different colours of the observed results indicate different Higgs boson production modes. Hatched area indicates that the ttH parameter of interest is constrained to positive values. For visualisation purposes, the VBF-p_T^j-High value is divided by a factor of three. Yellow vertical bands represent theory uncertainty in the signal prediction, while horizontal grey bands represent the expected measurement uncertainty [183].

6

Projection of Higgs Cross Section Measurement to HL-LHC

One of the main goals of the high-luminosity upgrade of the LHC (HL-LHC) is the precision measurement of the Higgs boson properties. In order to estimate the expected precision of the measurement of the Higgs boson production cross sections and couplings, a comprehensive projection study has been performed by the ATLAS Collaboration in 2014 based on the extrapolation of the Run 1 results to an integrated luminosity of 3000 fb^{-1} expected at the HL-LHC [57]. Since, several analyses have been improved with Run 2 data, the extrapolation is updated based on most recent analyses using Run 2 data sets of 36.1 fb^{-1} or 79.8 fb^{-1} collected at $\sqrt{s} = 13$ TeV. In addition, to the increase in integrated luminosity, the extrapolation accounts for the increase in the total cross section due to an increase of the centre-of-mass energy from 13 to 14 TeV [194]. The corresponding results serve as an input for the long-term planning of the research strategy in particle physics being prepared by the European Strategy [195].

The extrapolation for the $H \to ZZ^* \to 4\ell$ decay channel was performed as a part of this thesis and is presented in this chapter. A detailed description of the extrapolation is given followed by the discussion of the results. The projected results are obtained based on the results of the Stage-0 Higgs boson cross section measurements with 79.8 fb^{-1} of data as described in Chapter 5.

6.1 Extrapolation Procedure

The high-luminosity LHC machine is expected to start in 2026 and operate at a centre-of-mass energy of $\sqrt{s} = 14$ TeV with a baseline instantaneous luminosity of $5 \cdot 10^{34}$ cm^{-2}s^{-1}. It should deliver a total luminosity of about 3000 fb^{-1} over about 10 years of operation.

In order to account for the higher centre-of-mass energy, the signal expected yields from the Run 2 analysis (Chapter 5) are scaled according to the ratio of Higgs boson production cross sections at 14 and 13 TeV [189], as summarised in Table 6.1. The increase of the background yields is also taken into account. Processes which are predominantly quark

pair initiated are scaled by a factor of 1.08, while gluon pair initiated processes are scaled by 1.12 [29].

Table 6.1: Cross sections for the production of the SM Higgs boson with $m_H = 125$ GeV at a centre-of-mass energy of 13 TeV and 14 TeV [189], as well as their ratio corresponding to the scale factor applied on the expected signal yields for the projection from Run 2 to HL-LHC

	Higgs boson production mode								
	ggF	VBF	WH	$qq \to ZH$	$gg \to ZH$	ttH	bbH	tHq/b	tHW
σ (14 TeV) [pb]	54.67	4.278	1.513	0.8418	0.1443	0.6137	0.5529	0.0935	0.0186
σ (13 TeV) [pb]	48.58	3.782	1.373	0.7620	0.1227	0.5071	0.4880	0.0771	0.0151
Scale factor	1.125	1.131	1.102	1.105	1.176	1.121	1.133	1.210	1.224

It is expected that the systematic uncertainties described in Section 5.4 can be reduced for the HL-LHC program as described later on. In order to show the impact of this reduction of systematic uncertainties two extrapolation scenarios are considered. In the first scenario (HL-LHC S1), systematic uncertainties from the current Run 2 analysis are kept, while the event yields are extrapolated to higher luminosity and cross sections. The second scenario (HL-LHC S2) takes into account the reduction of the systematic uncertainties. It is assumed that the improved performance of the ATLAS detector will compensate for the performance degradation due to higher pile-up. Therefore, to simplify the extrapolation, the reconstruction efficiency, resolution and fake rate for some of the final state particles are assumed to be the same as in the baseline Run 2 analysis used for the projection. In addition, it is assumed that uncertainties related to the limited size of the simulated samples will become negligible. The systematic uncertainties from Section 5.4 are separated into the individual components for the extrapolation. Their values are reduced such to reflect the improvements reachable at the end of the HL-LHC area [196]. The sources of uncertainties with the corresponding reduction scale factors are summarised in Table 6.2.

The expected uncertainty on the integrated luminosity for the full HL-LHC data set is reduced by a factor of two down to 1% due to an increased statistical precision of the measurement [196]. All sources of uncertainties that are measured with calibrated data and thus driven by statistical precision are assumed to scale by a factor of $\sqrt{L}$. In an optimistic case the uncertainty on the scale factor applied in the pile-up reweighting would scale with $\sqrt{L}$ since it is measured in data. However, for the extrapolation a more conservative case is assumed. It is assumed that this uncertainty reduces by a factor of two. Uncertainties on the energy scale and resolution, as well as reconstruction and

identification efficiency uncertainties of electrons, muons and photons are kept unchanged for the extrapolation. This is also the case for the jet energy scale uncertainties, while uncertainties related to the jet energy resolution are reduced by a factor of two. The uncertainties on the light flavour, b- and c-tagging efficiencies are scaled by a factor of $1/3$.

The theoretical systematic uncertainties for the signal and background processes are in general reduced by a factor of two, except for the PDF+α_s uncertainties. In the latter case, the reduction factor depends on the quark/gluon initial states as well as the energy scale at which the collision takes place [197]. Since the initial state cannot be distinguished for the extrapolation an averaged scale factor of 0.45 is applied.

Table 6.2: Reduction scale factor applied to the systematic uncertainties in the $H \to ZZ^* \to 4\ell$ analysis of 79.8 fb^{-1} Run 2 data (see Section 5.4) to extrapolate to 3000 fb^{-1} at HL-LHC [196, 197]

Source of systematic uncertainty	Reduction scale factor
Integrated luminosity	0.5
Sources with data-based uncertainty measurement	$\sqrt{L}$
Scale factor applied in the pile-up reweighting	0.5
Lepton energy scale and resolution	1
Lepton reconstruction and identification efficiency	1
Jet energy scale	1
Jet energy resolution	0.5
Light flavour, b- and c-tagging efficiencies	0.33
Theory predictions for signal and background (except of PDF+α_s)	0.5
Theory predictions for signal and background (PDF+α_s)	0.45

6.2 Results

The expected results are obtained using the same statistical framework with binned maximum-likelihood fits as described in Section 5.5. The HL-LHC S1 (S2) projection of the cross section measurement for SM Higgs boson production in the Stage-0 gives [194]

$$
\begin{aligned}
\sigma_{ggF}\cdot\mathcal{B} &= 1.305^{+0.073}_{-0.072}\,\left(^{+0.057}_{-0.056}\right)\ \text{pb}\\
&= 1.305^{+0.026}_{-0.026}\ (\text{stat})\ ^{+0.055}_{-0.056}\,\left(^{+0.046}_{-0.046}\right)\ (\text{exp})\ ^{+0.034}_{-0.031}\,\left(^{+0.021}_{-0.020}\right)\ (\text{sig})\ \ ^{+0.009}_{-0.009}\,\left(^{+0.008}_{-0.008}\right)\ (\text{bkg})\ \text{pb},\\
\sigma_{VBF}\cdot\mathcal{B} &= 0.104^{+0.015}_{-0.014}\,\left(^{+0.013}_{-0.012}\right)\ \text{pb}\\
&= 0.104^{+0.010}_{-0.010}\ (\text{stat})\ ^{+0.006}_{-0.006}\,\left(^{+0.006}_{-0.005}\right)\ (\text{exp})\ ^{+0.009}_{-0.008}\,\left(^{+0.005}_{-0.005}\right)\ (\text{sig})\ \ ^{+0.001}_{-0.001}\,\left(^{+0.001}_{-0.001}\right)\ (\text{bkg})\ \text{pb},\\
\sigma_{VH}\cdot\mathcal{B} &= 0.058^{+0.012}_{-0.011}\,\left(^{+0.011}_{-0.010}\right)\ \text{pb}\\
&= 0.058^{+0.010}_{-0.010}\ (\text{stat})\ ^{+0.003}_{-0.002}\,\left(^{+0.002}_{-0.002}\right)\ (\text{exp})\ ^{+0.005}_{-0.004}\,\left(^{+0.004}_{-0.003}\right)\ (\text{sig})\ ^{+0.001}_{-0.001}\,\left(^{<0.001}_{<0.001}\right)\ (\text{bkg})\ \text{pb},\\
\sigma_{ttH}\cdot\mathcal{B} &= 0.016^{+0.004}_{-0.003}\,\left(^{+0.004}_{-0.003}\right)\ \text{pb}\\
&= 0.016^{+0.003}_{-0.003}\ (\text{stat})\ ^{+0.001}_{-0.001}\,\left(^{+0.001}_{-0.001}\right)\ (\text{exp})\ ^{+0.002}_{-0.001}\,\left(^{+0.001}_{-0.001}\right)\ (\text{sig})\ ^{+0.001}_{-0.001}\,\left(^{<0.001}_{<0.001}\right)\ (\text{bkg})\ \text{pb}.
\end{aligned}
$$

The corresponding scans of the expected test statistic per production mode are shown in Figure 6.1. For comparison, the results of the Run 2 analysis are shown together with both extrapolation results of the HL-LHC S1 and HL-LHC S2 scenarios in Table 6.3. Their cross section uncertainties are shown per particle level production bin relative to the corresponding SM cross section.

For all production modes the statistical uncertainties Δ_{stat} approximately scale with the factor luminosity, i.e. by a factor $(3000\ \text{fb}^{-1}/80\ \text{fb}^{-1})^{1/2} \approx 6.3$. Small differences are due to the expected higher centre-of-mass energy. The reduction of the systematic uncertainties in scenario HL-LHC S2 improves the precision of the measurement (Δ_{tot}) by a factor of about 1.2 for all production modes compared to HL-LHC S1. The Stage-0 ggF production bin is dominated by the systematic uncertainties for both the HL-LHC S1 and HL-LHC S2 scenario, whereas the measurements from the 79.8 fb^{-1} data set are still statistically dominated. The systematic uncertainties with the largest impact are the experimental ones. The measurement of the cross sections of other Stage-0 production modes (VBF, VH and ttH) is statistically dominated for all considered scenarios. Among the systematic uncertainties the experimental uncertainties dominate in the current Run 2 measurement of the VBF production cross section. In scenario HL-LHC S1 the theoretical uncertainties on the signal are the most important, while the experimental ones are again dominant in scenario HL-LHC S2. The measurement of the production cross section of the VH and ttH production mode is strongly limited by statistical uncertainties, while the largest contribution of systematic uncertainties in all scenarios are the experimental uncertainties.

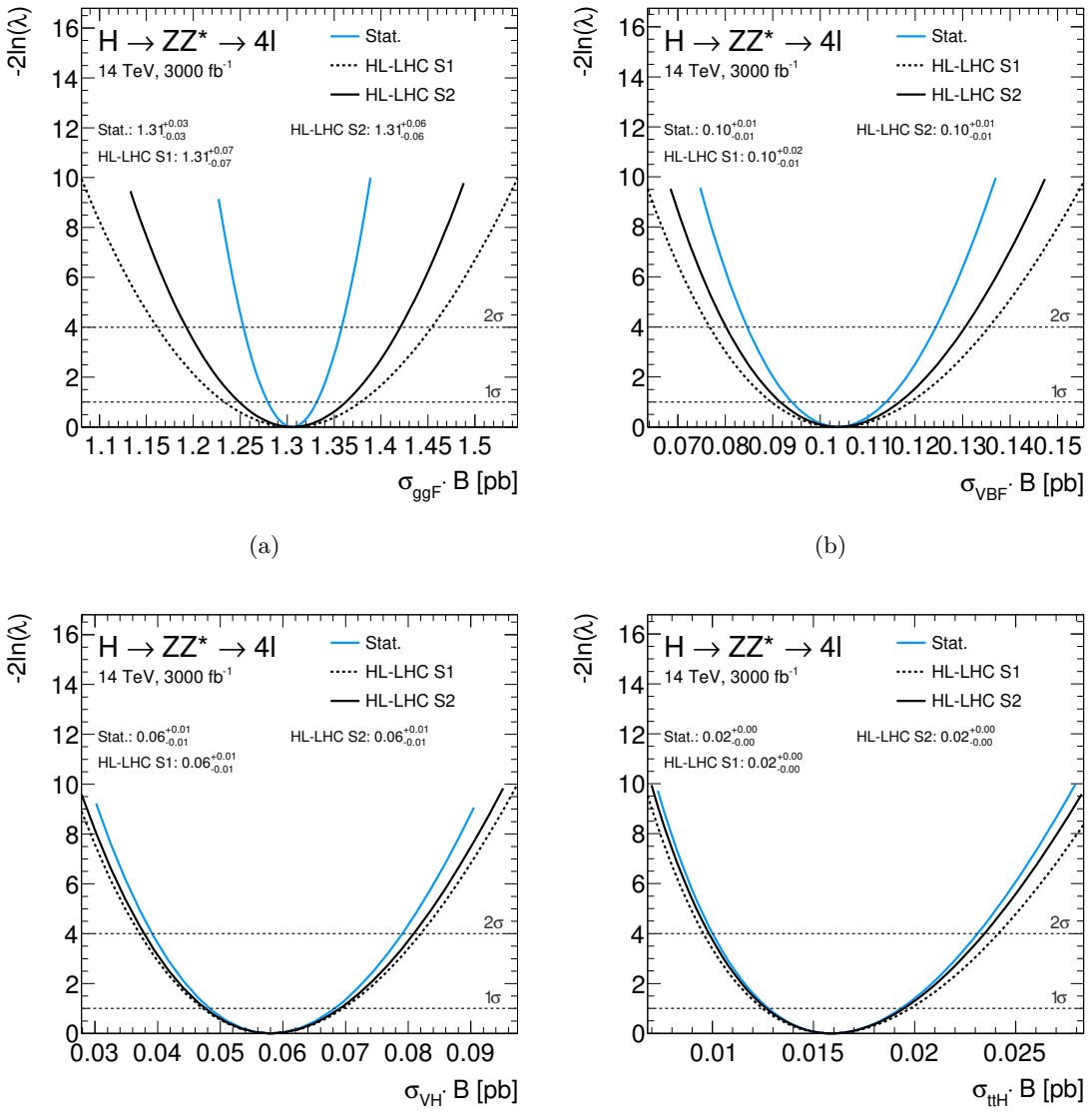


Figure 6.1: The expected test statistic $q = -2\ln(\lambda)$ at an integrated luminosity of 3000 fb^{-1} and a centre-of-mass energy of 14 TeV for the cross section measurement for the (a) ggF, (b) VBF, (c) VH and (d) ttH production mode. The scenario "stat" without systematic uncertainties (blue line) is shown, as well as the scenario HL-LHC S1 with systematic uncertainties as applied in the Run 2 analysis (black dashed line) and the scenario HL-LHC S2 with reduced systematic uncertainties (black solid line).

Table 6.3: Expected results of the Stage-0 production cross section measurements in the $H \to ZZ^* \to 4\ell$ decay channel with 79.8 fb^{-1} of Run 2 data and at HL-LHC. Uncertainties (Δ) are reported relative to the SM cross section σ_{SM} at the corresponding centre-of-mass energy. Two HL-LHC scenarios, S1 and S2, with different assumptions on the systematic uncertainties are shown for the HL-LHC extrapolation. The upper limit on the cross section for ttH production in the Run 2 analysis is given at 95% CL [194].

Analysis	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\text{sig}}/\sigma_{\text{SM}}$	$\Delta_{\text{bkg}}/\sigma_{\text{SM}}$
			Stage-0 ggF production bin		
Run 2 79.8 fb^{-1}	$^{+0.160}_{-0.152}$	$^{+0.143}_{-0.136}$	$^{+0.053}_{-0.052}$	$^{+0.043}_{-0.036}$	$^{+0.011}_{-0.014}$
HL-LHC S1	$^{+0.056}_{-0.055}$	$^{+0.020}_{-0.020}$	$^{+0.042}_{-0.043}$	$^{+0.026}_{-0.024}$	$^{+0.007}_{-0.007}$
HL-LHC S2	$^{+0.043}_{-0.043}$	$^{+0.020}_{-0.020}$	$^{+0.035}_{-0.035}$	$^{+0.016}_{-0.015}$	$^{+0.006}_{-0.006}$
			Stage-0 VBF production bin		
Run 2 79.8 fb^{-1}	$^{+0.782}_{-0.598}$	$^{+0.753}_{-0.583}$	$^{+0.157}_{-0.095}$	$^{+0.136}_{-0.074}$	$^{+0.014}_{-0.029}$
HL-LHC S1	$^{+0.147}_{-0.135}$	$^{+0.097}_{-0.094}$	$^{+0.059}_{-0.054}$	$^{+0.088}_{-0.078}$	$^{+0.007}_{-0.008}$
HL-LHC S2	$^{+0.125}_{-0.117}$	$^{+0.097}_{-0.094}$	$^{+0.057}_{-0.052}$	$^{+0.051}_{-0.047}$	$^{+0.007}_{-0.006}$
			Stage-0 VH production bin		
Run 2 79.8 fb^{-1}	$^{+1.410}_{-0.959}$	$^{+1.381}_{-0.946}$	$^{+0.155}_{-0.075}$	$^{+0.228}_{-0.137}$	$^{+0.012}_{-0.008}$
HL-LHC S1	$^{+0.200}_{-0.185}$	$^{+0.176}_{-0.167}$	$^{+0.051}_{-0.042}$	$^{+0.082}_{-0.070}$	$^{+0.002}_{-0.001}$
HL-LHC S2	$^{+0.190}_{-0.178}$	$^{+0.176}_{-0.167}$	$^{+0.043}_{-0.033}$	$^{+0.064}_{-0.056}$	$^{<0.001}_{<0.001}$
			Stage-0 ttH production bin		
Run 2 79.8 fb^{-1}	< 5.75		–		
HL-LHC S1	$^{+0.246}_{-0.213}$	$^{+0.217}_{-0.195}$	$^{+0.056}_{-0.042}$	$^{+0.100}_{-0.074}$	$^{+0.020}_{-0.026}$
HL-LHC S2	$^{+0.226}_{-0.202}$	$^{+0.217}_{-0.195}$	$^{+0.042}_{-0.032}$	$^{+0.047}_{-0.037}$	$^{+0.010}_{-0.015}$

The expected evolution of different sources of systematic uncertainties with integrated luminosity is shown in Figure 6.2 for all four production modes assuming the scenario HL-LHC S2 at a centre-of-mass energy of 14 TeV. It is assumed, that the systematic uncertainties scale with $\sqrt{L}$ until they reach the value from the HL-LHC S2 scenario at 3000 fb^{-1} (compare Table 6.2).

For the ggF production, the statistical uncertainties are dominant up to a luminosity of 1000 fb^{-1}, after which the experimental uncertainties become dominant. At 3000 fb^{-1} the statistical uncertainties are of the similar size as the theoretical systematic uncertainties. For other Stage-0 production modes, the statistical uncertainties are clearly dominant. Much smaller impact have the experimental systematic uncertainties, followed by the theoretical uncertainties on the signal and background.

In order to understand in more detail the impact of each individual source of systematic uncertainty, a so-called ranking of nuisance parameters is performed. The impact of each source of systematic uncertainty on the cross section measurement is computed by fitting the statistical model described in Section 5.5 on the pseudo data [1] corresponding to the SM hypothesis.

As a first step, the variations of each of the nuisance parameters corresponding to one standard deviations ($\pm 1\ \sigma$) are determined from an unconditional likelihood fit. Subsequently, the value of a given nuisance parameter is fixed (pulled) to $\pm 1\ \sigma$ away from the nominal value and the minimisation is performed again with this fixed value. This is done for each source of systematic uncertainties. The impact of a given source corresponds to the difference between the value of the fitted cross section from the unconditional fit and from the fit with the nuisance parameters pulled to $\pm 1\ \sigma$.

The respective nuisance parameter ranking for the ten dominant sources of systematic uncertainties is shown in Figure 6.3 and Figure 6.4 for the cross sections of the Stage-0 ggF, VBF, VH and ttH production modes in the HL-LHC S2.

In the Run 2 cross section measurement, the dominant systematic uncertainties for the ggF production are the electron and muon reconstruction as well as identification efficiency and pile-up modelling uncertainties (Table 5.5). For the extrapolation to HL-LHC, these still have a large impact. In addition, the uncertainty on the luminosity measurement and theory uncertainties related to the irreducible ZZ^* background become important, as well as the uncertainty on ggF production from alternative parton shower generators.

[1] Pseudo data sets are defined in such way that, if they are used to evaluate the estimators for all parameters, one obtains the true parameter values [193].

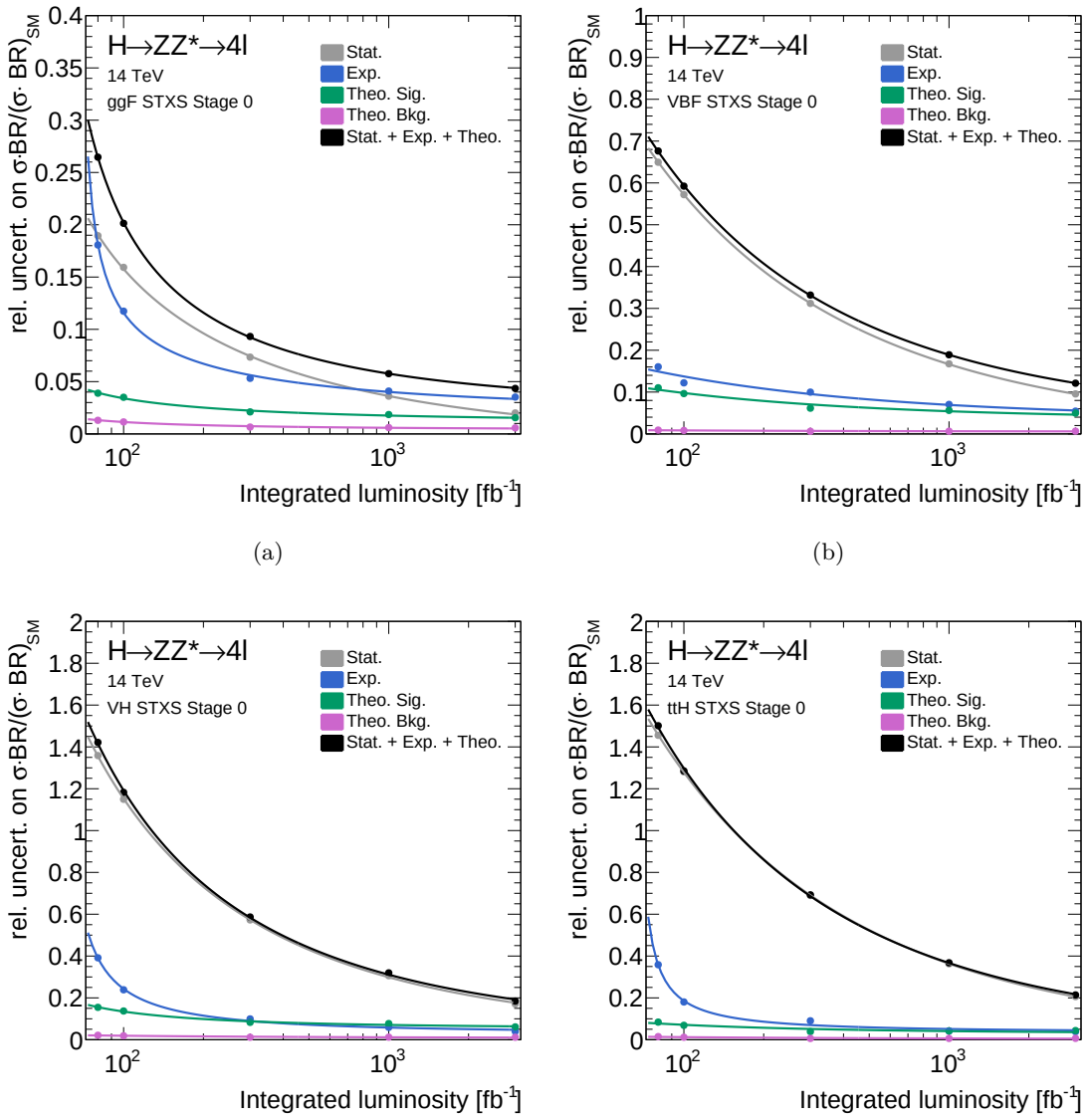
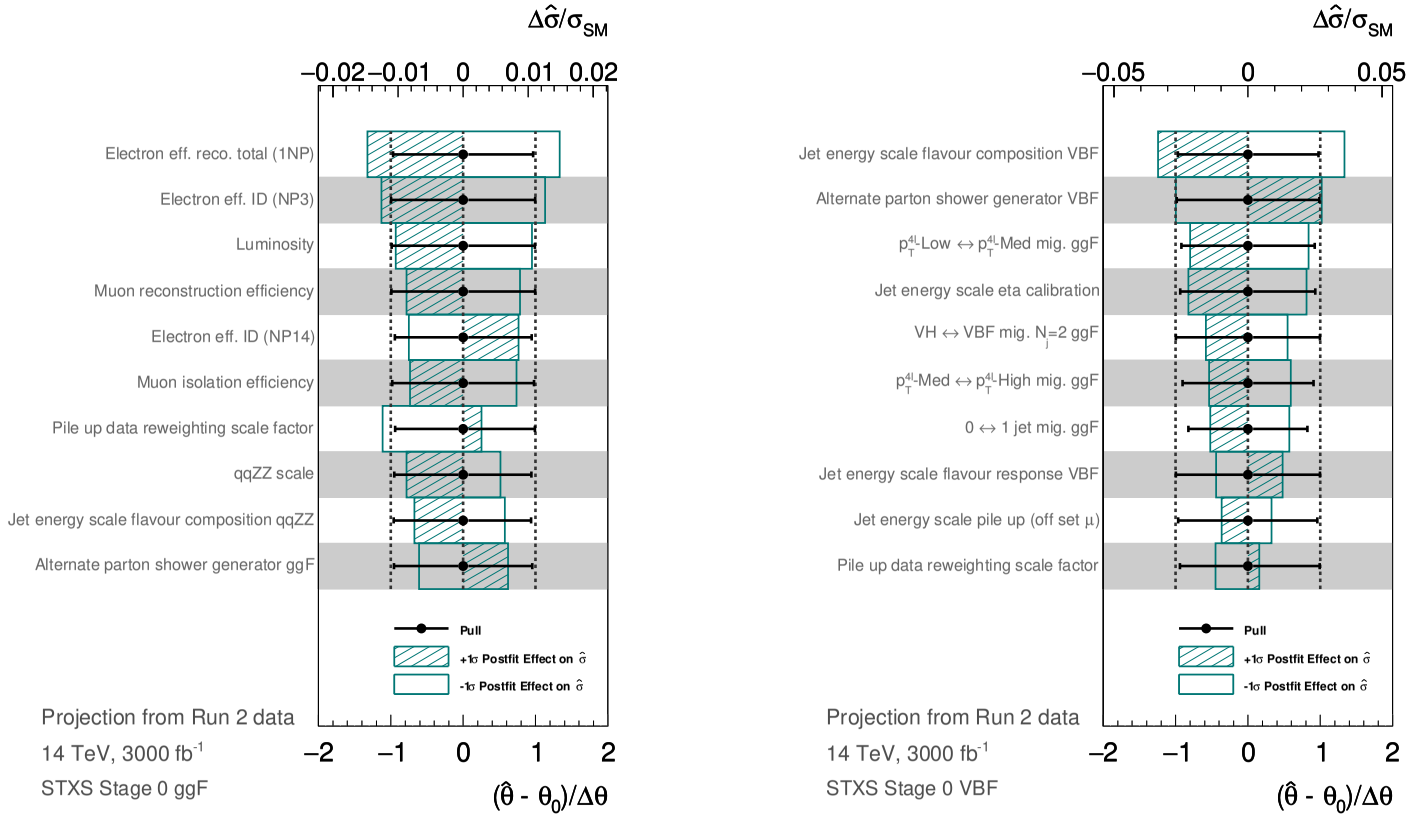


Figure 6.2: Expected statistical (grey line), experimental (blue line) and theoretical signal (green line) and background (pink line) uncertainties on the production cross section measurement in the HL-LHC S2 scenario as a function of the integrated luminosity at a centre-of-mass energy of 14 TeV for (a) ggF, (b) VBF, (c) VH and (d) ttH production mode. The total uncertainty, given by the squared sum of all sources of uncertainties, is shown as black line. The systematic uncertainties scale with $\sqrt{L}$ until the values from Table 6.2 are reached.

Major sources of systematic uncertainties for the VBF production mode in the Run 2 analysis are the jet energy scale and resolution, as well as QCD scale uncertainties related to jet and p_T^H bin migrations. For the HL-LHC S2 scenario these uncertainties are still dominant. In addition the alternative parton shower generator uncertainty is one of the ten dominant sources.

Dominant systematic uncertainties for the VH production mode in the Run 2 analysis are the same as for the VBF production. In the HL-LHC S2 scenario some of them are still dominant ones, but the uncertainty on the luminosity measurement, parton shower uncertainties and on electron reconstruction and identification become important.

In case of the ttH production mode, the largest uncertainty in the Run 2 analysis originates from the theory uncertainties related to parton shower and to the heavy flavour quark production modelling for the ggF background contribution. These are still dominant in the HL-LHC S2 scenario. Additional, systematic uncertainties with large impact are the uncertainties on electron and muon reconstruction efficiencies as well as on jet energy scale.



(a) (b)

Figure 6.3: Nuisance parameter ranking for the ten dominant sources of systematic uncertainties on the Stage-0 production cross section measurement for (a) ggF and (b) VBF production modes, as obtained for the HL-LHC S2 scenario in the $H \to ZZ^* \to 4\ell$ decay channel. The upper axis belongs to the impact on the cross section with fixed nuisance to ± 1 σ, while the lower axis belongs to the pull, i.e. value of a given nuisance parameter fixed to ± 1 σ away from the nominal value.

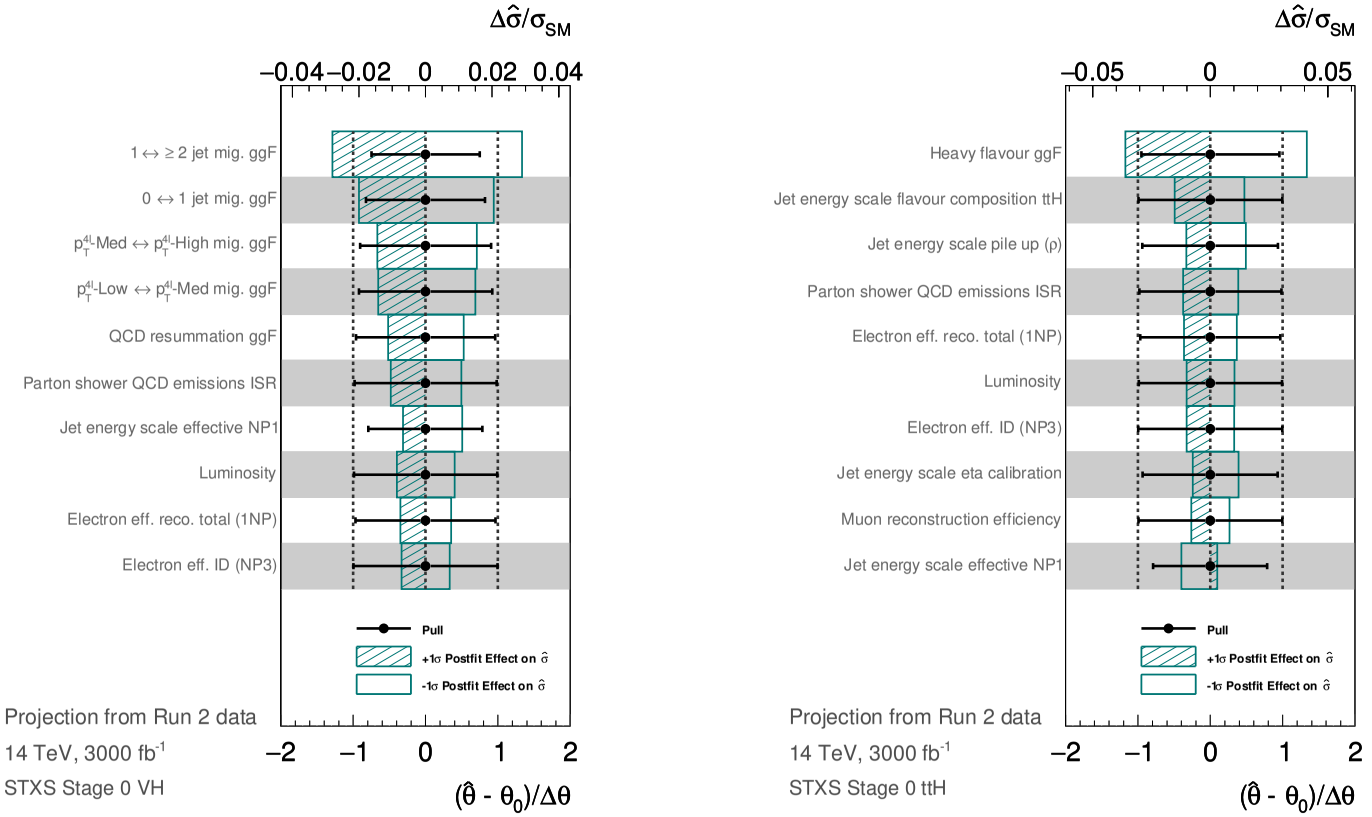


Figure 6.4: Nuisance parameter ranking for the ten dominant sources of systematic uncertainties on the Stage-0 production cross section measurement for (a) VH and (b) ttH production modes, as obtained for the HL-LHC S2 scenario in the $H \to ZZ^* \to 4\ell$ decay channel. The upper axis belongs to the impact on the cross section with fixed nuisance to ± 1 σ, while the lower axis belongs to the pull, i.e. value of a given nuisance parameter fixed to ± 1 σ away from the nominal value.

7

Measurement of the Tensor Structure of Higgs Boson Couplings

The studies presented in this chapter probe the tensor structure of the Higgs boson couplings to weak vector bosons, gluons and top quarks in search for new physics phenomena beyond the SM (BSM). In general, there are two types of observables sensitive to the tensor structure: the total production and decay rates in different Higgs boson production modes in exclusive phase space regions and the shape of the distributions related to the kinematic properties of final state particles.

The latter case is discussed in Chapter 8, while the studies presented in this chapter are based on the former one. The production and decay rates are in most cases more strongly affected by the BSM effects as compared to the kinematic properties. However, these observables cannot directly probe the CP-violation in the Higgs sector, i.e. they cannot distinguish between CP-even and CP-odd BSM contributions. CP-violating effects can only be measured through interference effects in SM and BSM interactions, which manifest themselves in modified distributions of the kinematic properties of final state particles.

Two separate analyses were performed to probe the tensor structure of the Higgs boson couplings by means of measured production and decay rates in the $H \to ZZ^* \to 4\ell$ decay channel. Both analyses have been introduced and performed almost entirely as a part of this thesis work.

The first analysis is performed with the first 36.1 fb^{-1} of Run 2 data and starts with the Higgs boson candidates selected by the full reconstruction chain of the inclusive analysis described in Section 4.5. These events are classified into categories of reconstructed events similar to those in Chapter 5 and the observed number of events in all categories is interpreted in terms of additional BSM Higgs boson couplings to vector bosons (HVV interaction vertex) and to gluons (Hgg interaction vertex). Both, CP-even and CP-odd BSM contributions are considered and are modelled by the Higgs Characterisation (HC) framework introduced in Section 2.5.2.2. In the following, this study will be referred to as the HC analysis and is described in the first part of this chapter (Section 7.2).

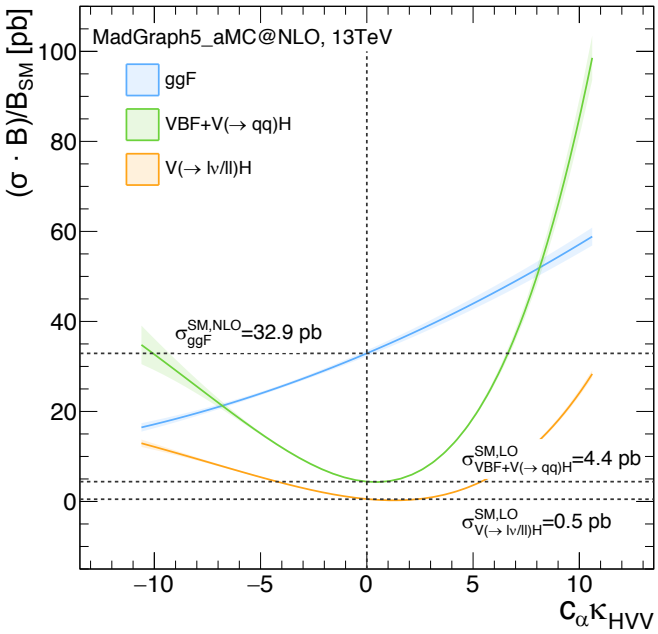
The second analysis, the so-called EFT interpretation is performed with 139 fb^{-1} of data. Rather than interpreting directly the event yields of fully reconstructed events, constraints on CP-even and CP-odd BSM coupling parameter are set by interpreting the cross section measurements in the Reduced-Stage-1.1 scheme (see Section 5.1.3) in terms of BSM couplings with the Standard Model Effective Field Theory (SMEFT) introduced in Section 2.5.2.1. The cross sections expected in each particle level production bin are parametrised as functions of BSM coupling parameters. In addition to the HVV and Hgg interaction vertex, also the top Yukawa interaction vertex ttH is probed. The analysis of the full Run 2 data set is quite similar to the one described in Chapter 5 with small changes introduced to further improve the sensitivity. Theses changes and the results of the EFT interpretation are described in detail in Section 7.3.

7.1 Observables Sensitive to BSM Contributions

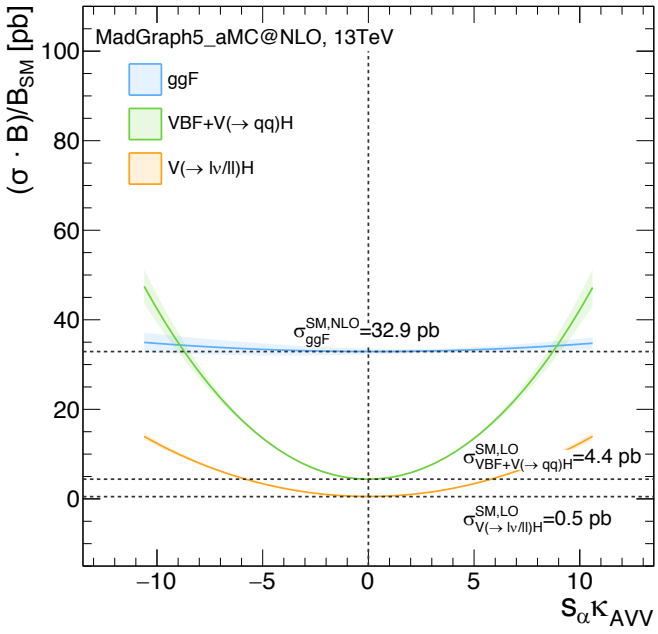
The presence of BSM contributions in Higgs boson interactions with SM particles can significantly enhance the cross sections of different Higgs boson production modes and change the relative production mode contributions compared to the SM prediction.

The expected cross section dependence in the $H \to ZZ^* \to 4\ell$ decay channel on different BSM couplings of the HC framework is shown in Figure 7.1 and Figure 7.2. The impact of CP-even ($c_\alpha\kappa_{HVV}$) and CP-odd ($s_\alpha\kappa_{AVV}$) BSM contributions in the HVV vertex is shown in Figure 7.1(a) and Figure 7.1(b), respectively, separately for the ggF, VBF+VH-Had and VH-Lep production. As described in Section 4.2.2.2, VBF and VH-Had processes both have the same initial and final state particles, $qq \to Hqq$ and are generated simultaneously in order to take possible interference effects into account.

The BSM coupling parameter κ_{XVV}, where $X = H, A$, affect the production rates of all production modes. VBF and VH production rates are changing more steeply ($\propto \kappa^4_{XVV}$) than the ggF rate ($\propto \kappa^2_{XVV}$), since in the former case the BSM coupling contributes to both production and decay interaction vertex, while it affects only the decay vertex in the latter case. The dependence of production rates on the CP-odd BSM coupling parameter is symmetric with respect to the SM point at $s_\alpha\kappa_{AVV} = 0$ providing no sensitivity to the sign of the coupling parameter. In contrast, the sign of the CP-even BSM coupling parameter $c_\alpha\kappa_{HVV}$ can be probed due to interference with the CP-even coupling term of the SM interaction. Due to the interference effects, the minimal Higgs boson production cross section is expected at a non-zero value of the BSM coupling parameter. For the VBF and ggF (VH) production modes this minima are located at negative (positive) values of BSM coupling parameters. The combination of the VBF process with VH-Had contribution in Figure 7.1(a) shifts the minimum to the positive values.



(a)



(b)

Figure 7.1: The expected Higgs boson production cross section times branching ratio normalised to the SM prediction of the branching ratio at a centre-of-mass energy of 13 TeV for ggF (blue), VBF+VH-Had (green) and VH-Lep (orange) Higgs boson production in dependence on the BSM coupling parameters (a) $c_\alpha\kappa_{HVV}$ and (b) $s_\alpha\kappa_{AVV}$ [198]. The cross sections are calculated at next-to-leading-order for ggF and leading-order for VBF and VH using MadGraph5_aMC@NLO. Horizontal dashed lines indicate the value of the predicted SM production cross section for a given production mode.

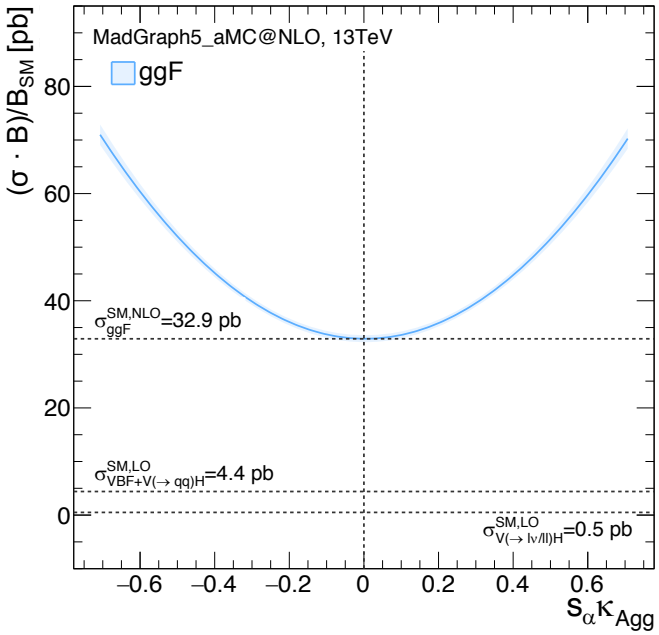


Figure 7.2: The expected Higgs boson production cross section times branching ratio normalised to the SM prediction of the branching ratio at a centre-of-mass energy of 13 TeV for ggF Higgs boson production in dependence on the BSM coupling parameter $s_\alpha \kappa_{Agg}$ [198]. The cross sections are calculated at next-to-leading-order using MadGraph5_aMC@NLO. Horizontal dashed lines indicate the value of the predicted SM production cross section for a given production mode.

Figure 7.2 shows the expected ggF production cross section in dependence on the CP-odd BSM coupling parameter ($s_\alpha \kappa_{Agg}$) related to the couplings of the Higgs boson to gluons. The dependence is quadratic ($\propto \kappa^2_{Agg}$), since the BSM coupling contributes only in the production vertex. In general, a similar impact on the expected cross section is also seen for the BSM coupling parameters of the SMEFT model.

As stated previously the distributions of kinematic properties of the final state particles are also sensitive to CP-even and CP-odd BSM contributions. An example of such observables which are related to the Higgs boson decay system are illustrated in Figure 7.3. In addition, to the decay observables, also the distributions of the transverse momentum of the final state jets, the transverse momentum of the Higgs boson and the mass of subleading lepton pair m_{34} can be modified by the presence of BSM contributions. As an example, theses distributions are shown for the VBF+VH-Had production mode for the SM and selected BSM Higgs boson signals in Figure 7.5 and Figure 7.4.

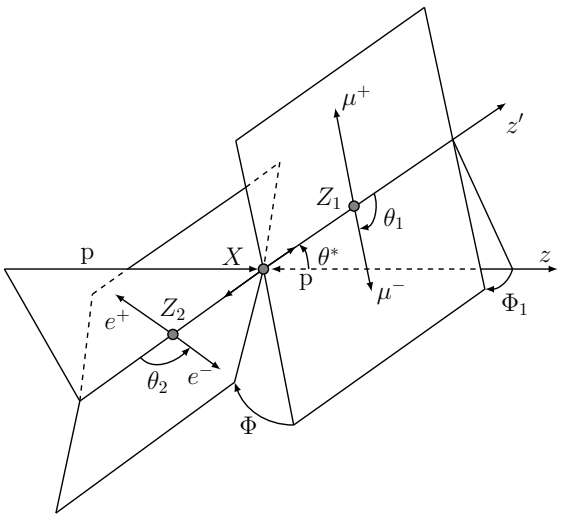


Figure 7.3: Definition of the decay angles in the four-lepton final state from $H \to ZZ^* \to 4\ell$ decays shown in the Higgs boson (X) rest frame [28]. θ_1 (θ_2) is the angle between the negatively charged lepton from the decay of the on-shell (off-shell) Z boson in the direction of flight of the respective Z boson, Φ is the angle between the two decay planes spanned by the di-lepton pairs in the four-lepton rest frame, Φ_1 is the angle between the direction of the momentum of the on-shell Z boson and the decay plane spanned by its decay products and θ^* is the production angle of the on-shell Z boson.

The Higgs boson candidates on which the HC analysis is performed are classified into several reconstructed categories targeting different production modes. The categorisation is based on the kinematic properties of final state particles. Therefore, both the production rates as well as the shape of distributions of kinematic properties are employed for the measurement of the tensor coupling structure. Similarly, the EFT interpretation is based on cross sections in exclusive phase space regions of particle level production bins (Section 7.3.2). Thus, also here the kinematic properties of the final state products are employed in addition to the total production rate.

The BSM contributions can also be probed by a complementary approach with inclusive differential distributions of kinematic properties without classification into event categories or particle level production bins. This approach is not covered by this thesis.

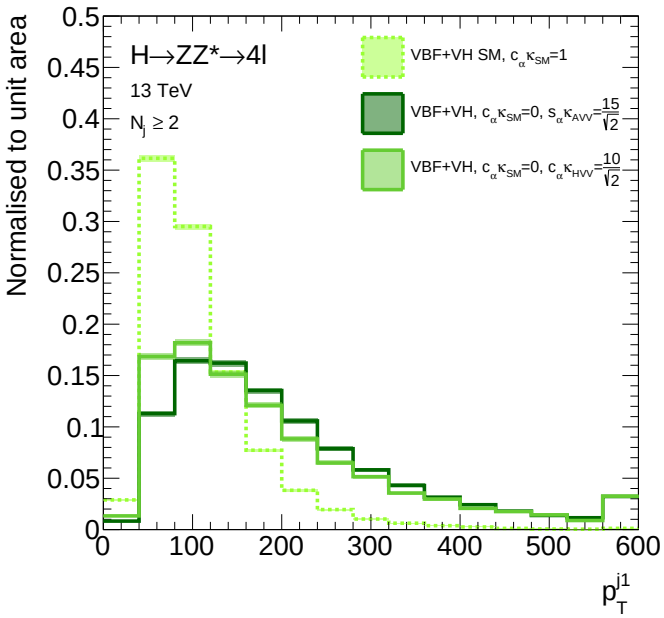
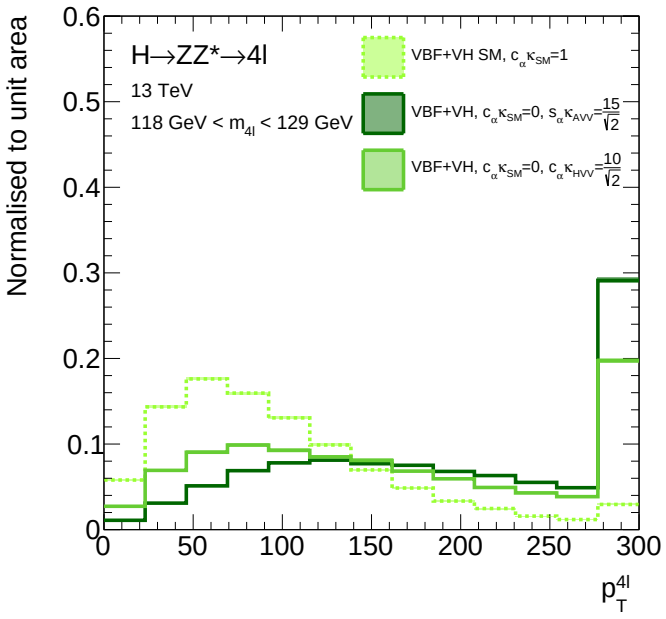
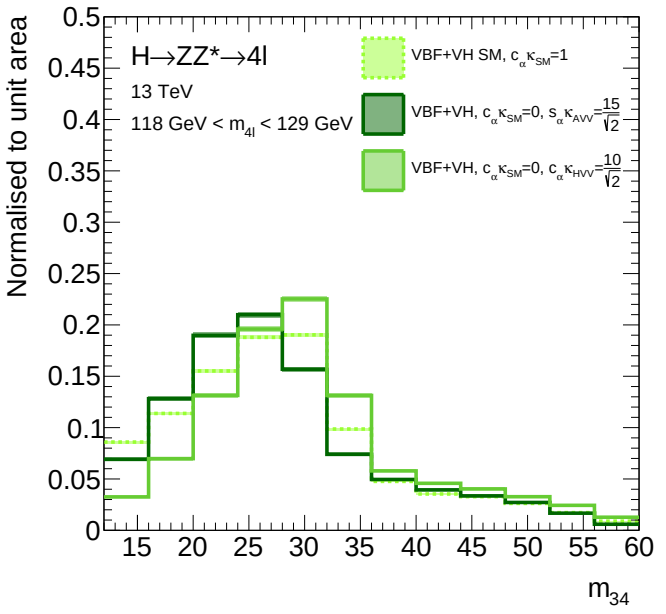


Figure 7.4: The distribution of p_T^{j1} for $N_j \geq 2$ in the mass range of 118 GeV $< m_{4l} <$ 129 GeV for SM and selected BSM Higgs boson signals in the $H \to ZZ^* \to 4\ell$ decay channel. The distributions are normalised to unit area. The last bin is an overflow bin collecting all events beyond the plotted x-axis range.



(a)



(b)

Figure 7.5: The distribution of (a) $p_T^{4\ell}$ and (b) m_{jj} in the mass range of 118 GeV $< m_{4l} <$ 129 GeV for SM and selected BSM Higgs boson signals in the $H \to ZZ^* \to 4\ell$ decay channel. The distributions are normalised to unit area. The last bin is an overflow bin collecting all events beyond the plotted x-axis range.

7.2 Tensor Structure Measurement in the Higgs Characterisation Framework

The CP-even and CP-odd BSM contributions to the Higgs boson interactions introduced within the Higgs Characterisation (HC) model, are probed with the first Run 2 data of 36.1 fb^{-1}.

The contributions to Higgs boson couplings to weak vector bosons in the XVV effective interaction vertex and to gluons in the Xgg effective interaction vertex are probed separately. It is assumed that the CP-even and CP-odd BSM coupling parameters to the Z bosons are directly correlated to the couplings to the W bosons, $\kappa_{XVV} = \kappa_{XZZ} = \kappa_{XWW}$. The cross section times branching ratio dependence on different BSM couplings is summarised in Table 7.1 separately for the XVV and Xgg interaction vertices.

Table 7.1: Dependence of the production cross section times branching ratio $\sigma \cdot \mathcal{B}(H \to ZZ^*)$ in different Higgs boson production modes on the BSM coupling contributions to the effective XVV and Xgg interaction vertices. The BSM couplings are defined in the Higgs Characterisation framework.

Interaction	Dependence on the BSM coupling				
vertex	ggF	VBF	VH	ttH	bbH
HVV	$\propto (c_\alpha \kappa_{HVV})^2$	$\propto (c_\alpha \kappa_{HVV})^4$	$\propto (c_\alpha \kappa_{HVV})^4$	$\propto (c_\alpha \kappa_{HVV})^2$	$\propto (c_\alpha \kappa_{HVV})^2$
AVV	$\propto (s_\alpha \kappa_{AVV})^2$	$\propto (s_\alpha \kappa_{AVV})^4$	$\propto (s_\alpha \kappa_{AVV})^4$	$\propto (s_\alpha \kappa_{AVV})^2$	$\propto (s_\alpha \kappa_{AVV})^2$
Agg	$\propto (s_\alpha \kappa_{Agg})^2$	const.	const.	const.	const.

In the XVV interaction vertex, the parameters $c_\alpha \kappa_{HVV}$ and $s_\alpha \kappa_{AVV}$ of the CP-even and CP-odd BSM couplings to vectors bosons, respectively, are probed. In the VBF and VH production modes they are contributing both to the production as well as to the decay vertex such that the production rates are proportional to the corresponding BSM coupling parameter to the power of four. In case of ggF, ttH and bbH production modes, the coupling parameters appear only in the decay leading to a quadratic dependence of production rates. The measurement of the VBF and VH production process therefore provides the highest sensitivity to BSM contributions in the XVV vertex.

In the effective Xgg interaction vertex, the parameter $s_\alpha \kappa_{Agg}$ of the CP-odd BSM coupling to gluons is probed. This coupling affects only the ggF production vertex resulting in a quadratic dependence of the ggF production rate on the BSM coupling. The analysis is not sensitive to the parameter $c_\alpha \kappa_{Hgg}$ of the CP-even Higgs boson coupling. This parameter is therefore assumed to be equal to the SM value of $c_\alpha \kappa_{Hgg} = 1$.

7.2.1 Event Categorisation

In order to measure the signal production rates in different production modes, the reconstructed Higgs boson candidates selected by the inclusive $H \to ZZ^* \to 4\ell$ analysis (Section 4.5) in a mass window of 118 GeV $< m_{4\ell} <$ 129 GeV in the 36.1 fb^{-1} data set are classified into several categories based on the properties of final state products related to the production vertex. In addition, some of the categories are subdivided according to the transverse momentum of the four-lepton final state or the leading jet p_T to provide an additional sensitivity to BSM couplings.

To be consistent with the cross section measurement (see Chapter 5), the categorisation scheme is closely following the one described in Section 5.2. There are small differences since the categorisation in Chapter 5 is optimised for the larger data set of 79.8 fb^{-1}. One additional category is added to increase the sensitivity to BSM couplings. The differences and the expected BSM sensitivity in the relevant categories is discussed in the following.

The schematic overview of the categorisation scheme for the HC analysis is summarised in Figure 7.6. As for the scheme used in the cross section measurement (Section 5.2) the ttH-enriched events are selected first, by requiring at least one b-tagged jet tag and either more than four additional reconstructed jets or at least two jets and at least one additional lepton in the final state. The former requirement targets the fully hadronic decay of the top quark pair, while the latter one selects the semi-leptonic tt decays. Due to a very small number of events in this category for 36.1 fb^{-1}, the ttH-enriched category is not further subdivided according to hadronic and semi-leptonic tt final states as it was the case for the scheme in Section 5.2.

Events with additional leptons in the final state ($N_{lep} \geq 5$) which do not enter the ttH-enriched category are assigned to the VH-Lep-enriched category targeting the Higgs boson production in association with leptonically decaying vector bosons. This category provides sensitivity to the measurement of BSM couplings in the XVV interaction vertex.

The remaining production modes are again selected according to the number of jets in the final state. VBF and VH-Had production are separated from the ggF production by requiring at least two jets in the final state, $N_j \geq 2$. The former two production modes are further disentangled by the threshold on the invariant mass m_{jj} of the two leading jets resulting in the VBF-enriched ($m_{jj} > 120$ GeV) and the VH-Had-enriched ($m_{jj} < 120$ GeV) categories.

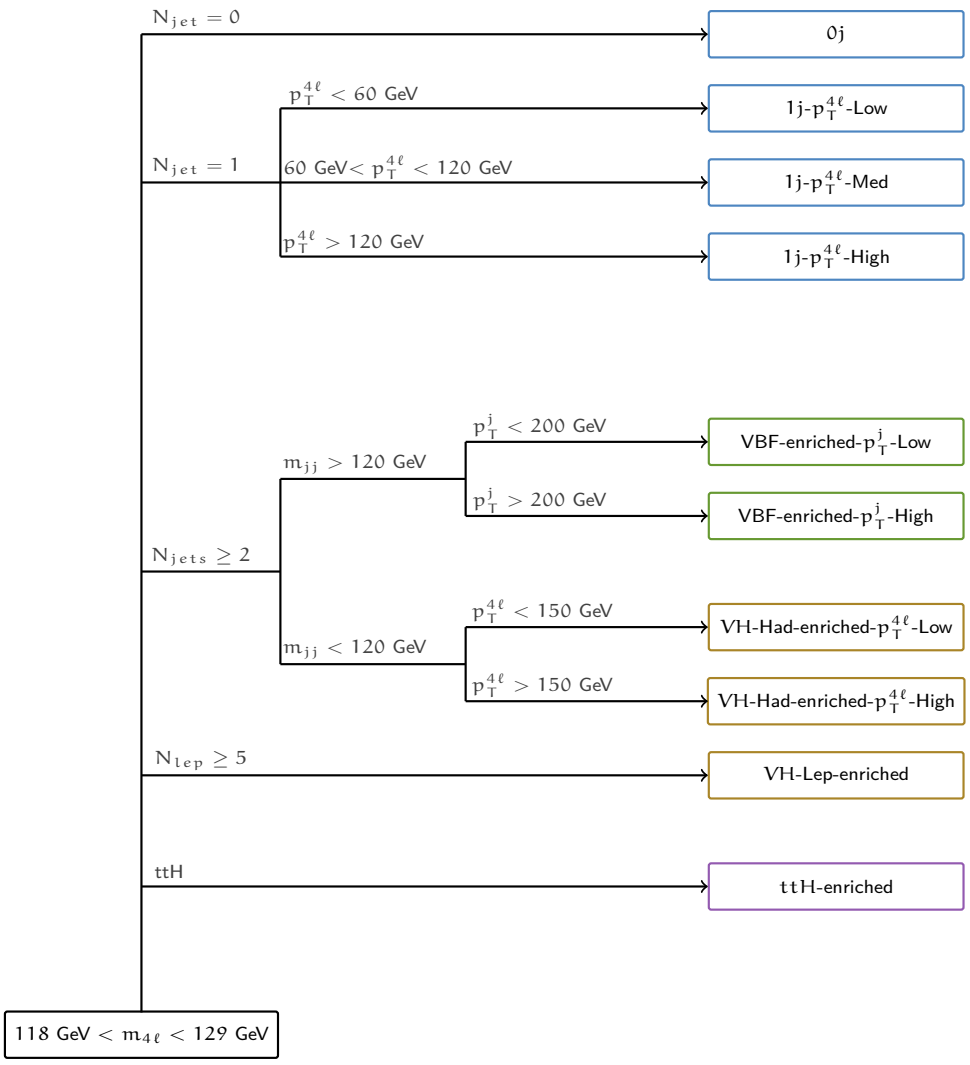


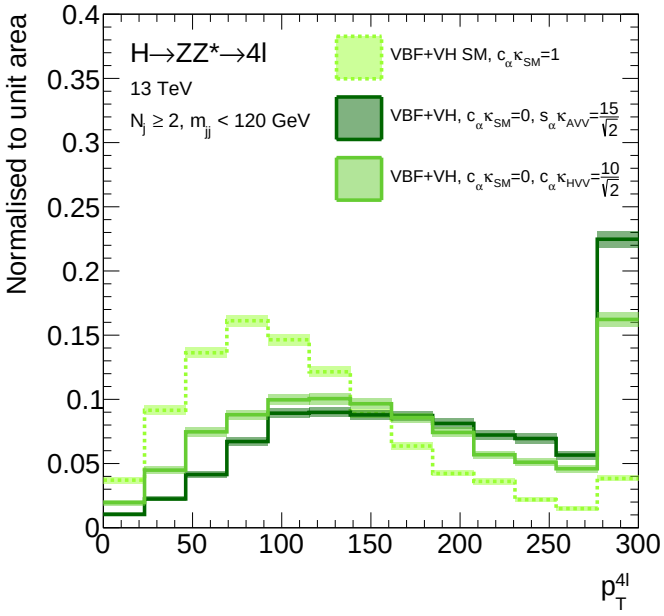
Figure 7.6: Overview of the event categories employed for the HC analysis of the Higgs boson couplings tensor structure.

The presence of non-vanishing BSM couplings in the XVV vertex is expected to lead to a harder distribution of the Higgs boson and leading jet transverse momenta, $p_T^{4\ell}$ and p_T^{j1}, respectively. To further increase the sensitivity to BSM contributions, the VH-Had-enriched category is therefore split according to the $p_T^{4\ell}$ into the VH-Had-enriched-$p_T^{4\ell}$-Low ($p_T^{4\ell} < 150$ GeV) and the VH-Had-enriched-$p_T^{4\ell}$-High ($p_T^{4\ell} > 150$ GeV) category. The expected distributions in the events with $N_j \geq 2$ and $m_{jj} < 120$ GeV are shown in Figure 7.7(a) for SM and several BSM Higgs boson signals. BSM contributions to the XVV vertex lead to higher population of bins with higher $p_T^{4\ell}$ values.

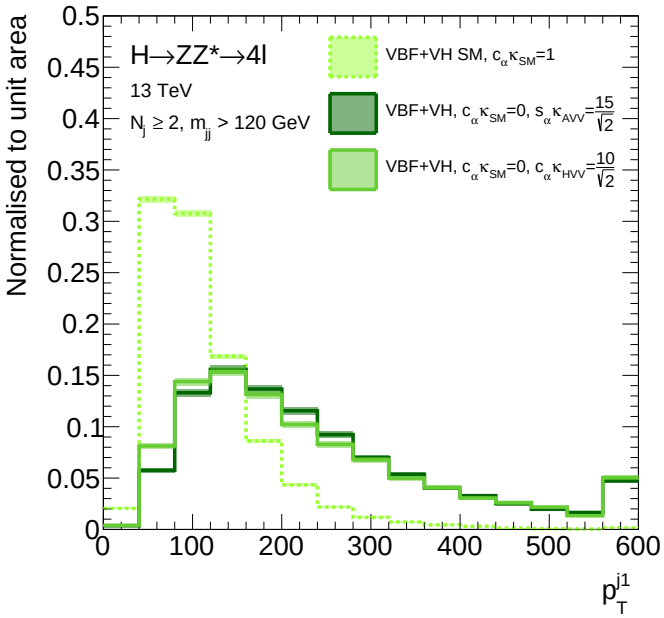
Similarly, to gain additional BSM sensitivity in the VBF-enriched category a threshold on the transverse momentum of the leading jet is applied at $p_T^{j1} = 200$ GeV. In the case of the SM hypotheses the Higgs boson candidates tend to populate the VBF-enriched-p_T^j-Low category ($p_T^{j1} < 200$ GeV), while in the case of BSM contributions they mainly contribute to the VBF-enriched-p_T^j-High category ($p_T^{j1} > 200$ GeV). The expected p_T^{j1} distribution in the VBF-enriched category is shown in Figure 7.7(b).

The ggF, VBF and VH-Had production processes contribute to the category with one additional jet in the final state. In order to distinguish the different production modes and to isolate BSM effects, this category is split according to the four-lepton transverse momentum as in the case of cross section measurements. Events with $p_T^{4\ell}$ smaller than 60 GeV are assigned to the $1j$-$p_T^{4\ell}$-Low category, while events with 60 GeV $< p_T^{4\ell} < 120$ GeV and $p_T^{4\ell} > 120$ GeV are collected in the $1j$-$p_T^{4\ell}$-Med and $1j$-$p_T^{4\ell}$-High categories, respectively. The expected $p_T^{4\ell}$ distribution in the combined $1j$ category is shown is Figure 7.8, separately for ggF and VBF+VH-Had production modes. While no differences between SM and BSM hypotheses are visible for the ggF production, a harder spectrum is expected for the VBF+VH-Had production in case of BSM contributions. In addition, the VBF+VH-Had spectra are harder than the ggF spectra independent of BSM contributions. Therefore, the categorisation according to $p_T^{4\ell}$ in the $1j$ category provides at the same time discrimination between the two production modes and the sensitivity to BSM contributions in the XVV vertex in the $1j$-$p_T^{4\ell}$-High category.

The $0j$ category is almost pure in ggF contribution and therefore provides the largest sensitivity to BSM contributions in the Xgg interaction vertex.



(a)



(b)

Figure 7.7: The distribution of (a) $p_T^{4\ell}$ in the VH-Had-enriched category with $N_j \geq 2$ and $m_{jj} < 120$ GeV and (b) p_T^{j1} in the VBF-enriched category with $N_j \geq 2$ and $m_{jj} > 120$ GeV for SM and selected BSM Higgs boson signals in the $H \to ZZ^* \to 4\ell$ decay channel. The distributions are normalised to unit area. The last bin is an overflow bin collecting all events beyond the plotted x-axis range.

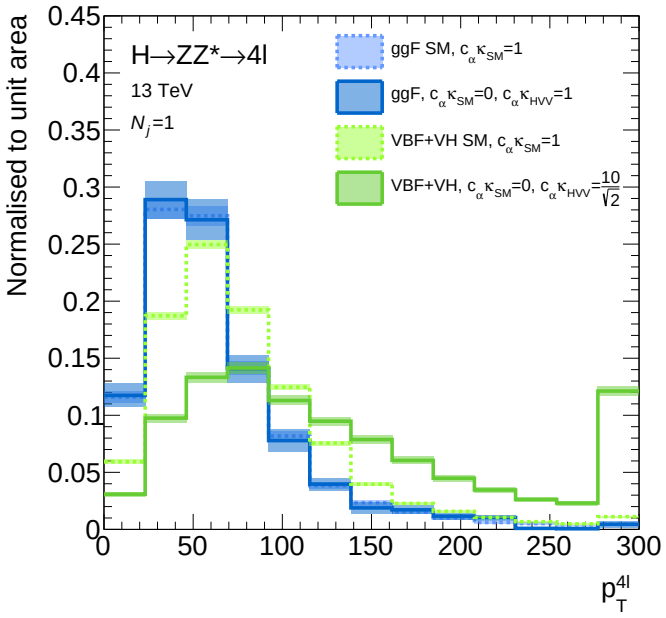


Figure 7.8: The distribution of $p_T^{4\ell}$ in the $1j$ category ($N_j = 1$) for SM and selected BSM Higgs boson signals in the $H \to ZZ^* \to 4\ell$ decay channel. The distributions are normalised to unit area. The last bin is an overflow bin collecting all events beyond the plotted x-axis range.

The expected number of events in each of the reconstructed event categories are shown in Table 7.2 for the different Higgs boson production modes assuming the SM hypotheses. The corresponding expected signal composition in each category is shown in Figure 7.9. The VH-Lep-enriched and ttH-Had-enriched reconstructed event categories are very pure with about 80% contribution of the respective production mode. However, the ttH-Had-enriched category plays only a minor role in the measurement for the tensor structure of Higgs boson couplings, providing no sensitivity to the Xgg production vertex.

Table 7.2: The expected number of SM Higgs boson events with a mass of $m_H = 125$ GeV in the mass range of 118 GeV $< m_{4\ell} <$ 129 GeV for an integrated luminosity of 36.1 fb^{-1} at a centre-of-mass energy of 13 TeV in each of the reconstructed event category, shown separately for the different Higgs boson production modes. The ggF and bbH contributions are shown separately but they are merged into a single production bin (ggF) for the final result. Statistical and systematic uncertainties are added in quadrature [182].

Reconstructed	SM Higgs boson production mode									
event category	ggF		VBF		VH		ttH		bbH	
$0j$	25.9	$\pm$2.5	0.29	$\pm$0.09	0.253	$\pm$0.025	0.00025	$\pm$0.00019	0.29	$\pm$0.14
$1j$-$p_T^{4\ell}$-Low	8.0	$\pm$1.1	0.514	$\pm$0.034	0.230	$\pm$0.018	0.0007	$\pm$0.0005	0.09	$\pm$0.05
$1j$-$p_T^{4\ell}$-Med	4.5	$\pm$0.7	0.64	$\pm$0.09	0.227	$\pm$0.019	0.0010	$\pm$0.0005	0.026	$\pm$0.013
$1j$-$p_T^{4\ell}$-High	1.10	$\pm$0.24	0.27	$\pm$0.04	0.095	$\pm$0.007	0.00080	$\pm$0.00024	0.0036	$\pm$0.0018
VBF-enriched-p_T^j-Low	3.9	$\pm$0.8	2.03	$\pm$0.19	0.285	$\pm$0.024	0.065	$\pm$0.009	0.045	$\pm$0.023
VBF-enriched-p_T^j-High	0.33	$\pm$0.09	0.185	$\pm$0.024	0.050	$\pm$0.004	0.0159	$\pm$0.0027	0.00058	$\pm$0.00029
VH-Had-enriched-$p_T^{4\ell}$-Low	2.3	$\pm$0.5	0.169	$\pm$0.014	0.418	$\pm$0.023	0.022	$\pm$0.004	0.025	$\pm$0.013
VH-Had-enriched-$p_T^{4\ell}$-High	0.42	$\pm$0.09	0.048	$\pm$0.008	0.162	$\pm$0.005	0.0090	$\pm$0.0015	< 0.0001	
VH-Lep-enriched	0.0129	$\pm$0.0018	0.00310	$\pm$0.00021	0.263	$\pm$0.018	0.038	$\pm$0.005	0.0009	$\pm$0.0005
ttH-enriched	0.050	$\pm$0.016	0.010	$\pm$0.006	0.0196	$\pm$0.0031	0.301	$\pm$0.032	0.0064	$\pm$0.0035
Total	47	$\pm$4	4.16	$\pm$0.23	2.00	$\pm$0.11	0.48	$\pm$0.05	0.48	$\pm$0.24

The ggF production contributes with a larger fraction in almost all categories. Its largest contribution with a purity of 97% is in the $0j$ category. With increasing $p_T^{4\ell}$ the ggF fraction in $1j$ categories decreases from 90% to 75%. In the $2j$ categories, the relative ggF contribution is still quite high with about 60% purity.

The VBF production mode is targeted in the VBF-enriched-p_T^j-Low and VBF-enriched-p_T^j-High categories with purities of 58% and 75%, respectively. There is also a substantial VBF contribution in the three $1j$ categories, with purities between 5% and 18%.

The VH-Had production mode is targeted by the VH-Had-enriched-$p_T^{4\ell}$-Low and VH-Had-enriched-$p_T^{4\ell}$-High categories with corresponding signal purities of 14% and 25%, respectively. Smaller VH-Had contributions are also expected in the VBF-enriched and $1j$ categories.

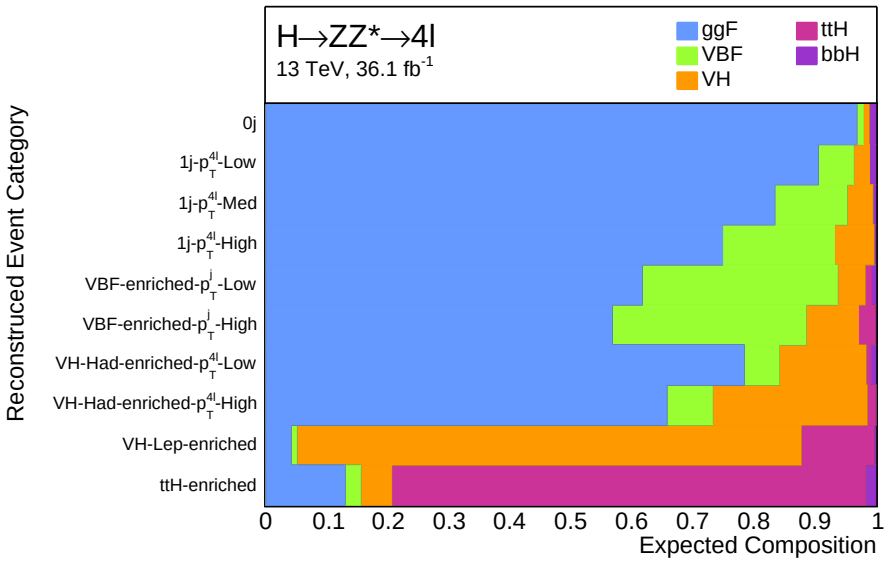
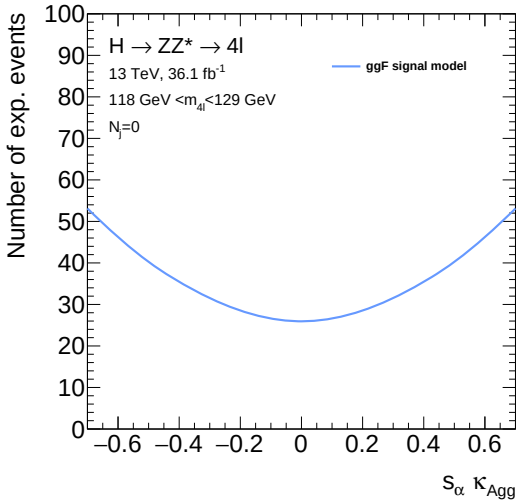


Figure 7.9: Expected signal composition of the SM Higgs boson signal in each of the reconstructed event category.

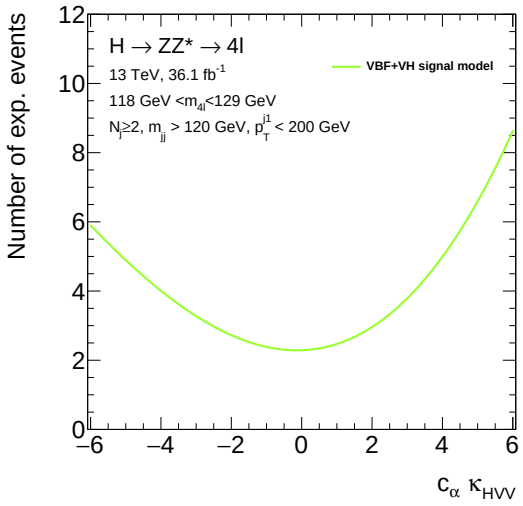
7.2.2 Signal Modelling

To study the tensor structure of the Higgs boson couplings to vector bosons and gluons, a dedicated signal model is needed which predicts the event yields from the different production modes in each reconstructed event category as a function of the BSM coupling parameters. The signal model is based on the so-called morphing method [199] which allows for a signal prediction in an arbitrary point of the multi-dimensional BSM parameter space by means of interpolation between the cross section predictions at discrete points of this space, obtained from dedicated samples. In order to take into account higher-order cross section calculations, which are not implemented in generated signal samples, a further correction is applied on the predicted signal yields taking into account highest-order prediction of the inclusive cross sections as listed in Table 4.2. As described in Section 4.2.2.2, the BSM signal samples are simulated with MadGraph5_aMC@NLO. The presence of BSM coupling contributions affects also the total width of the Higgs boson decay, which is not taken during the simulation. Therefore, the signal model is additionally corrected to take these effects into account. Each of the three signal modelling steps are described in more detail in the following subsections.

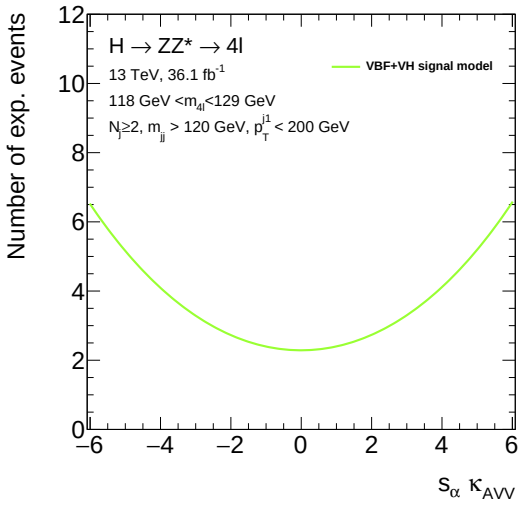
Examples of the resulting signal model are shown in Figure 7.10 for the ggF production in the $0j$ category and for VBF+VH production in the VBF-enriched-p_T^j-Low category. The former is sensitive to the Higgs boson couplings to gluons, the latter to couplings to vector bosons.



(a)



(b)



(c)

Figure 7.10: The signal model describing the expected event yield as a function of the BSM coupling parameter (a) $s_\alpha \kappa_{Agg}$ for the ggF production in the $0j$ event category, (b) $c_\alpha \kappa_{HVV}$ and (c) $s_\alpha \kappa_{AVV}$ for the VBF+VH production in the VBF-enriched-p_T^j-Low category. The signal models include the signal morphing via simulated samples, the scaling to the best-prediction cross sections and the scaling corrections accounting for BSM effects on the total Higgs boson decay width.

7.2.2.1 The Morphing Method

The tensor structure of the Higgs boson couplings is probed separately for the effective Xgg and XVV interaction vertices, using separate signal models. Each signal production mode is modelled on its own, using the BSM signal samples described in Section 4.2.2.2. The samples of ggF, VBF+VH-Had and VH-Lep events are simulated with different values of BSM coupling parameters, $s_\alpha\kappa_{Agg}$, $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$. No BSM samples are generated for the ttH and bbH processes, since there is no significant BSM sensitivity expected from the corresponding production vertices. The coupling parameters κ_{Hgg} and κ_{Agg} include the Higgs boson coupling to top or bottom quarks in the effective coupling to gluons, but the sensitivity to ttH production in the $H \to ZZ^* \to 4\ell$ decay is very small and there is no sensitivity to the bbH production mode. The production vertex of the ttH and bbH processes is therefore assumed to be described by the SM. The BSM contributions in the decay vertex (XVV) are taken into account by assuming that the production rates scale by the same amount as for the ggF production with no BSM contributions in the Xgg vertex.

Three signal models with different sets of coupling parameters are constructed for the ggF, VBF+VH-Had and VH-Lep production using the morphing method. The first is the Xgg model employed for the study of the tensor structure of Higgs boson couplings to gluons. Since the presence of BSM contributions in the Xgg vertex only affects the ggF production and not the decays, the other production modes are modelled by the SM prediction. The CP-even ggF production and decay contributions with SM-like tensor structure are fixed to the SM prediction with $c_\alpha\kappa_{Hgg} = 1$ and $c_\alpha\kappa_{SM} = 1$. The expected number of events in each reconstructed category is therefore modelled only in dependence on the CP-odd coupling parameter,

$$N_{exp}^{\text{ggF}} = s_{\text{ggF}}^{Xgg}\left(s_\alpha\kappa_{Agg}\right). \tag{7.1}$$

For the study of the XVV interaction vertex a three-dimensional (3D XVV) signal model is used, describing CP-even and CP-odd BSM coupling contributions simultaneously. In addition, the parameter $c_\alpha\kappa_{SM}$ of the CP-even coupling contribution with the SM-like structure in the XVV vertex is taken into account, while the $c_\alpha\kappa_{Hgg}$ contributions in the ggF production vertex is fixed to the SM prediction, $c_\alpha\kappa_{Hgg} = 1$. The expected signal yield is modelled separately for each of the production modes,

$$\begin{aligned} N_{exp}^{\text{ggF}} &= s_{\text{ggF}}^{\text{3D }XVV}\left(c_\alpha\kappa_{SM}, c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV}\right), \\ N_{exp}^{\text{VBF+}VH\text{-Had}} &= s_{\text{VBF+}VH\text{-Had}}^{\text{3D }XVV}\left(c_\alpha\kappa_{SM}, c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV}\right) \quad \text{and} \\ N_{exp}^{VH\text{-Lep}} &= s_{VH\text{-Lep}}^{\text{3D }XVV}\left(c_\alpha\kappa_{SM}, c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV}\right). \end{aligned} \tag{7.2}$$

The introduced three-dimensional signal model is employed for the simultaneous measurement of the BSM parameters $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ from CP-even and CP-odd couplings and for the measurement of the BSM CP-even coupling parameter $c_\alpha\kappa_{HVV}$ with the CP-odd coupling $s_\alpha\kappa_{AVV}$ set to zero.

In general, this model can also be employed for the measurement of the CP-odd BSM coupling parameter $s_\alpha\kappa_{AVV}$ with $c_\alpha\kappa_{HVV} = 0$. However, statistical fluctuations in this model cause asymmetric predictions with respect to the sign, i.e. introduce a small artificial sensitivity to the sign of the CP-odd BSM coupling parameter. To avoid this the CP-odd BSM coupling contribution to the XVV vertex is modelled with a two-dimensional (2D XVV) signal model, with $c_\alpha\kappa_{HVV}$ set to zero:

$$\begin{aligned} N_{exp}^{\text{ggF}} &= s_{\text{ggF}}^{\text{2D } XVV}\left(c_\alpha\kappa_{SM}, s_\alpha\kappa_{AVV}\right), \\ N_{exp}^{\text{VBF+}VH\text{−Had}} &= s_{\text{VBF+}VH\text{−Had}}^{\text{2D } XVV}\left(c_\alpha\kappa_{SM}, s_\alpha\kappa_{AVV}\right) \quad \text{and} \\ N_{exp}^{VH\text{−Lep}} &= s_{VH\text{−Lep}}^{\text{2D } XVV}\left(c_\alpha\kappa_{SM}, s_\alpha\kappa_{AVV}\right). \end{aligned} \tag{7.3}$$

The SM-like coupling contributing in the Xgg vertex is set to the SM prediction, $c_\alpha\kappa_{Hgg} = 1$ and the corresponding CP-odd BSM contribution $s_\alpha\kappa_{Agg}$ is set to zero.

To avoid artificial sensitivity to the sign of the CP-odd BSM coupling parameter in the signal models from Equation 7.3, the input samples with opposite signs of the coupling parameters $s_\alpha\kappa_{AVV}$ are combined together and averaged. Similar procedure is applied to the signal model of Equation 7.1.

The input samples for the three described signal models obtained via the morphing technique are summarised in Tables 7.3, 7.4 and 7.5. The values of BSM parameters are chosen such to cover most of the BSM parameter space to which the measurement is sensitive.

The signal models described above are obtained with the morphing technique [199]. The model provides a continuous multi-dimensional description of the BSM parameter space, by describing simultaneously the production rates and the distributions of the kinematical variables in dependence on the BSM coupling parameters.

The target value T_t of the physical quantity T, which is related to the cross section whether inclusive or in a given bin of a differential distribution, at an arbitrary point t in the coupling parameter space $\boldsymbol{g}_t$ can be described as a linear combination of a fixed number N of the corresponding input quantities (cross sections) $T_i(\boldsymbol{g}_i)$ obtained from the simulation for a given set of parameters $\boldsymbol{g}_i$,

$$T_t\left(\boldsymbol{g}_t\right) = \sum_{i=1}^{N} w_i\left(\boldsymbol{g}, \boldsymbol{g}_i\right) T_i\left(\boldsymbol{g}_i\right). \tag{7.4}$$

Table 7.3: List of simulated input samples employed for the one-dimensional Xgg signal model of the ggF production and the corresponding production cross section (σ) times the branching ratio $\mathcal{B}(H \to ZZ^* \to 4\ell)$

Input sample	BSM parameters				$\sigma \cdot \mathcal{B}$ [fb]
	c_α	κ_{Hgg}	κ_{SM}	κ_{Agg}	
ggF SM	1	1	1	0	4.1
ggF Mixture_P	$1/\sqrt{2}$	1	1	0.5	1.6
ggF Mixture_N	$1/\sqrt{2}$	1	1	-0.5	1.6

Table 7.4: List of simulated input samples employed for 2D XVV signal models of the ggF and VBF+VH productions and the corresponding production cross section (σ) times the branching ratio $\mathcal{B}(H \to ZZ^* \to 4\ell)$

Input sample	BSM parameters					$\sigma \cdot \mathcal{B}$ [fb]
	c_α	κ_{Hgg}	κ_{SM}	κ_{HVV}	κ_{AVV}	
ggF production						
ggF SM	1	1	1	0	0	4.1
ggF Mixture_P	$1/\sqrt{2}$	1	1	0	6	1.1
ggF Mixture_N	$1/\sqrt{2}$	1	1	0	-6	1.1
VBF+VH production						
VBF+VH SM	1	–	1	0	0	0.62
VBF+VH Mixture_P1	$1/\sqrt{2}$	–	1	0	-2.5	0.25
VBF+VH Mixture_N1	$1/\sqrt{2}$	–	1	0	2.5	0.25
VBF+VH Mixture_P2	$1/\sqrt{2}$	–	1	0	5	0.53
VBF+VH Mixture_N2	$1/\sqrt{2}$	–	1	0	-5	0.53

Table 7.5: List of simulated input samples employed for 3D XVV signal models of the ggF and VBF+VH productions and the corresponding production cross section (σ) times the branching ratio $\mathcal{B}(H \to ZZ^* \to 4\ell)$

Input sample	BSM parameters					$\sigma \cdot \mathcal{B}$ [fb]
	c_α	κ_{Hgg}	κ_{SM}	κ_{HVV}	κ_{AVV}	
ggF production						
ggF SM	1	1	1	0	0	4.1
ggF Pure BSM (H)	1	1	0	1	0	$5.4 \cdot 10^{-3}$
ggF Pure BSM (A)	$1/\sqrt{2}$	1	0	0	1	$0.5 \cdot 10^{-3}$
ggF Mixture_P (A/H)	1	1	1	−6	0	2.8
ggF Mixture_1 (A/H)	$1/\sqrt{2}$	1	1	6	6	1.5
ggF Mixture_2 (A/H)	$1/\sqrt{2}$	1	1	−6	−6	0.7
VBF+VH production						
VBF+VH SM	1	–	1	0	0	0.62
VBF+VH Pure BSM (H)	1	–	0	10	0	1.12
VBF+VH Pure BSM (A)	$1/\sqrt{2}$	–	0	0	15	1.29
VBF+VH Mixture_P1 (A)	$1/\sqrt{2}$	–	1	0	5	0.53
VBF+VH Mixture_N1 (A)	$1/\sqrt{2}$	–	1	0	−5	0.53
VBF+VH Mixture_P2 (A)	$1/\sqrt{2}$	–	1	0	2.5	0.25
VBF+VH Mixture_P1 (H)	1	–	1	5	0	2.76
VBF+VH Mixture_N1 (H)	1	–	1	−5	0	2.57
VBF+VH Mixture_P2 (H)	1	–	1	2.5	0	0.92
VBF+VH Mixture_1 (A/H)	$1/\sqrt{2}$	–	1	2.5	−5	0.66
VBF+VH Mixture_2 (A/H)	$1/\sqrt{2}$	–	1	5	5	1.19
VBF+VH Mixture_3 (A/H)	$1/\sqrt{2}$	–	1	5	6	1.42
VBF+VH Mixture_4 (A/H)	$1/\sqrt{2}$	–	1	−2.5	−5	0.65
VBF+VH Mixture_5 (A/H)	$1/\sqrt{2}$	–	1	−5	5	0.93
VBF+VH Mixture_6 (A/H)	$1/\sqrt{2}$	–	1	−5	−6	1.05

The weight function $w_i(\boldsymbol{g}, \boldsymbol{g}_i)$ depends on the targeted and input values of the coupling parameters, $\boldsymbol{g}_t$ and $\boldsymbol{g}_i$, respectively, determining relative contribution of each sample. This holds under the assumption that the physical quantity T is proportional to the squared matrix element of the underlying process

$$T(\boldsymbol{g}) \propto |\mathcal{M}(\boldsymbol{g})|^2. \tag{7.5}$$

The matrix element $\mathcal{M}(\boldsymbol{g})$ can be written as a linear sum of products of the coupling parameters g_α and the corresponding operator $\mathcal{O}(g_\alpha)$,

$$|\mathcal{M}(\boldsymbol{g})|^2 = \left(\sum_{\alpha \in p,s} g_\alpha \mathcal{O}(g_\alpha)\right)^2 \cdot \left(\sum_{\alpha \in d,s} g_\alpha \mathcal{O}(g_\alpha)\right)^2, \tag{7.6}$$

with p, d and s running over couplings appear only in production, only in decay and in both production and decay vertices, respectively. The weights w_i are functions of the targeted coupling parameters $(g_\alpha, g_\beta, g_\gamma, g_\delta)$, defined as

$$w_i = \sum_{\alpha,\beta,\gamma,\delta=1}^{N} C_{\alpha,\beta,\gamma,\delta} g_\alpha g_\beta g_\gamma g_\delta, \tag{7.7}$$

where the values of the coefficients $C_{\alpha,\beta,\gamma,\delta}$ are obtained by requiring that for the target points which are equal to one of the input points T_i the corresponding weight w_i must satisfy

$$w_i(\boldsymbol{g}_t = \boldsymbol{g}_i, \boldsymbol{g}_i) = 1, \tag{7.8}$$

while the weights of other input samples T_j for the same target point $T_{i,i\neq j}$ must be zero,

$$w_{i\neq j}\left(\boldsymbol{g}_t = \boldsymbol{g}_i, \boldsymbol{g}_j\right) = 0. \tag{7.9}$$

The resulting linear system of equations, is called a morphing matrix and the coefficients $C_{\alpha,\beta,\gamma,\delta}$ are calculated by inverting the morphing matrix.

The minimal number of input samples $N(n_s, n_p, n_s)$ needed for the morphing method procedure is dependent on the number of coupling parameters in the production and decay vertices [199],

$$\begin{aligned} N = \quad & \frac{n_p(n_p+1)}{2} \cdot \frac{n_d(n_d+1)}{2} + \binom{4+n_s-1}{4} \\ & + \left(n_p \cdot n_s + \frac{n_s(n_s+1)}{2}\right) \cdot \frac{n_d(n_d+1)}{2} \\ & + \left(n_d \cdot n_s + \frac{n_s(n_s+1)}{2}\right) \cdot \frac{n_p(n_p+1)}{2} \\ & + \frac{n_s(n_s+1)}{2} \cdot n_p \cdot n_d + (n_p+n_d)\binom{3+n_s-1}{3}, \end{aligned} \tag{7.10}$$

where n_p corresponds to the number of coupling parameters appearing only in the production vertex, n_d only in the decay and n_s in both the production and decay vertex. The number of input samples needed to construct the signal models for the measurement of the effective Xgg and XVV couplings is shown in Table 7.6.

Table 7.6: Number of input samples needed to construct the signal model in a given BSM coupling parameter space, shown separately for the ggF and VBF+VH Higgs boson production modes

Model	Production couplings	Decay couplings	n_p	n_d	n_s	N
		ggF production				
1D Xgg	$c_\alpha\kappa_{Hgg}$, $s_\alpha\kappa_{Agg}$	$c_\alpha\kappa_{SM}$	2	1	0	3
2D XVV	$c_\alpha\kappa_{Hgg}$	$c_\alpha\kappa_{SM}$, $s_\alpha\kappa_{AVV}$	1	2	0	3
3D XVV	$c_\alpha\kappa_{Hgg}$	$c_\alpha\kappa_{SM}$, $c_\alpha\kappa_{HVV}$, $s_\alpha\kappa_{AVV}$	1	3	0	6
		VBF+VH production				
2D XVV	$c_\alpha\kappa_{SM}$, $s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM}$, $s_\alpha\kappa_{AVV}$	0	0	2	5
3D XVV	$c_\alpha\kappa_{SM}$, $c_\alpha\kappa_{HVV}$,$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM}$, $c_\alpha\kappa_{HVV}$,$s_\alpha\kappa_{AVV}$	0	0	3	16

The complete set of N input samples can in general be chosen arbitrarily, as long as it provides an independent set of linear equations defining the morphing matrix. However, the statistical accuracy of the target quantity depends on position of the target couplings in the coupling parameter space covered by the input samples. The statistical error increases if the target coupling is outside of the spanned parameter space. Therefore, the values of the BSM parameters input samples are chosen such to span the range which can be probed by data and provide sufficiently small errors in the studied parameter range.

7.2.2.2 Best-Prediction-Scaling of the Signal Model

The ggF and VBF+VH input samples for the signal modelling are generated at NLO and LO, respectively, with the MadGraph5_aMC@NLO generator. As outlined in Section 4.2.2, the corresponding nominal SM signal samples are simulated with Powheg which includes the higher-order effects. Furthermore, the inclusive SM cross section in each production mode is normalised by the factor k to the highest-order inclusive SM cross section calculation available.

In order to take these best-predictions of the cross sections into account, the expected number of events ($N^{exp,i}$) predicted by the signal model in each of the categories is scaled by a scale factor (K) defined as the ratio of the highest-order SM prediction and the

SM prediction from MadGraph5_aMC@NLO. Since the SM signal samples with the highest-order corrections do not distinguish between hadronically and leptonically decays of the vector boson in VH production, the signal model for the VBF+VH-Had and VH-Lep production is combined to apply best-prediction-scaling. This is possible, because the input samples are generated with the same BSM coupling values. It is assumed, that the same scale factor is valid for all BSM samples, i.e. that the impact of the higher-order corrections under a SM hypothesis is not affected by the presence of BSM coupling parameters.

The scale factors are calculated separately for the ggF and VBF+VH processes. For each reconstructed event category i, the scale factor for the ggF production is given as

$$K_i(\text{ggF}) = \frac{k_{\text{ggF}} N^{exp,i}_{\text{Powheg,SM}}(\text{ggF})}{N^{exp,i}_{\text{MG5,SM}}(\text{ggF})}, \tag{7.11}$$

and for the combined VBF+VH process

$$\text{SF}_\text{i}(\text{VBF}+VH) = \frac{k_{\text{VBF}} N^{exp,i}_{\text{Powheg,SM}}(\text{VBF}) + k_{WH} N^{exp,i}_{\text{Powheg,SM}}(WH) + k_{ZH} N^{exp,i}_{\text{Powheg,SM}}(ZH)}{N^{exp,i}_{\text{MG5,SM}}(\text{VBF}+VH\text{-Had}) + N^{exp,i}_{\text{MG5,SM}}(VH\text{-Lep})}. \tag{7.12}$$

The calculated scale factors for each reconstructed event category are shown in Table 7.7.

Table 7.7: Best-prediction scale factors K in each reconstructed event category, shown separately for the ggF and VBF+VH production modes

Event category	Best-prediction scale factor ggF	VBF+VH
$0j$	1.6	1.4
$1j$-$p_T^{4\ell}$-Low	1.4	1.3
$1j$-$p_T^{4\ell}$-Med	1.5	1.4
$1j$-$p_T^{4\ell}$-High	1.2	1.5
VBF-enriched-p_T^j-Low	1.2	1.0
VBF-enriched-p_T^j-High	1.0	1.0
VH-Had-enriched-$p_T^{4\ell}$-Low	1.4	1.3
VH-Had-enriched-$p_T^{4\ell}$-High	0.9	1.4
VH-Lep-enriched	1.2	1.3
ttH-enriched	0.9	1.9

7.2.2.3 Scaling of the Total Decay Width

The branching ratio ($\mathcal{B}$) is defined as the ratio of the partial and the total decay width. Both of these change in the presence of BSM couplings in the decay vertex. While the change of the partial decay width is taken into account in MadGraph5_aMC@NLO, the total decay width is kept at its SM value. Therefore, the event yield predicted from the simulated samples has to be corrected to take into account the dependence of the total decay width on BSM contributions.

The event yield is proportional to the inverse of total Higgs boson decay width Γ_{SM}. Accordingly, the correction function $f_\Gamma(g)$ applied on the event yield for a given set of BSM couplings g is

$$f_\Gamma(g) = \frac{\Gamma_{\text{SM}}}{\Gamma_{\text{BSM}}} = \frac{\Gamma_{\text{SM}}}{\sum_i \Gamma^i_{\text{BSM}}} = \frac{1}{\sum_i \mathcal{B}^i_{\text{SM}} \cdot f^i(g)}, \tag{7.13}$$

where i corresponds to a particular Higgs boson decay channel, $\mathcal{B}^i_{\text{SM}} = \Gamma^i_{\text{SM}}/\Gamma_{\text{SM}}$ is the SM branching ratio and $f^i(g)$ are polynomials of BSM coupling parameters in the decay vertex, which are derived from the matrix element morphing method of the corresponding process. In the case of the SM, the correction factor for the total decay width is $f_\Gamma(SM) = 1$.

The modification of the total decay width due to BSM effects are taken into account for the measurement of the tensor structure of the Higgs boson couplings in $H \to WW$, $H \to ZZ$ and $H \to gg$ decays. Although loop-induced $H \to Z\gamma$ and $H \to \gamma\gamma$ decays are also affected by the BSM contributions through the effective XWW vertex, the corresponding branching ratios are small compared to other decays and their values are kept at the SM prediction. The correction function applied on the total decay width is, therefore, for the presented analysis given as

$$f_\Gamma(g) = \left(\mathcal{B}^{gg}_{\text{SM}} \cdot f^{gg}(g) + \mathcal{B}^{WW}_{\text{SM}} \cdot f^{WW}(g) + \mathcal{B}^{ZZ}_{\text{SM}} \cdot f^{ZZ}(g) + \mathcal{B}^{\text{Other}}_{\text{SM}}\right)^{-1}, \tag{7.14}$$

with the values of the SM branching ratios summarised in Table 2.4. The branching ratio $\mathcal{B}^{\text{Other}}_{\text{SM}} = 0.67$ corresponds to the sum of the branching ratios of all other decay modes.

The correction functions are f^{gg} for the $H \to gg$ decays,

$$\begin{aligned} f^{gg}(c_\alpha \kappa_{Hgg}, s_\alpha \kappa_{Agg}) &= \\ &= 1 + c_1^{gg} \cdot (s_\alpha \kappa_{Agg})^2 + c_2^{gg} \cdot c_\alpha s_\alpha \kappa_{Hgg} \kappa_{Agg}, \end{aligned} \tag{7.15}$$

and $f^{WW/ZZ}$ for the $H \to WW$ and $H \to ZZ$ decays

$$\begin{aligned} f^{WW/ZZ}(c_\alpha\kappa_{SM}, c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV}) = \\ = 1 + c_1^{WW/ZZ} \cdot (c_\alpha\kappa_{HVV})^2 + c_2^{WW/ZZ} \cdot (s_\alpha\kappa_{AVV})^2 \\ + c_3^{WW/ZZ} \cdot c_\alpha^2\kappa_{SM}\kappa_{HVV} \\ + c_4^{WW/ZZ} \cdot c_\alpha s_\alpha \kappa_{SM}\kappa_{AVV} \\ + c_5^{WW/ZZ} \cdot c_\alpha s_\alpha \kappa_{HVV}\kappa_{AVV}, \end{aligned} \tag{7.16}$$

where c_i^{gg}, c_i^{WW} and c_i^{ZZ} are constant coefficients, which are given relative to the respective SM terms, i.e. the terms $(c_\alpha\kappa_{Hgg})^{-2}$ and $(c_\alpha\kappa_{SM})^{-2}$ are absorbed into the factors c_i^{gg} and $c_i^{WW/ZZ}$, respectively, and set to one. Therefore, if $c_\alpha\kappa_{Hgg}$ or $c_\alpha\kappa_{SM}$ are free parameters in the fit, the branching ratio has to be corrected with $(c_\alpha\kappa_{Hgg})^2$ and $(c_\alpha\kappa_{SM})^2$, respectively. The coefficients are derived using the morphing method, similarly as deriving the signal model. To evaluate the f^{gg} correction function, the cross section times the branching ratio is calculated for the ggF production with different values for κ_{Agg} and the cross section dependence on κ_{Agg} is modelled to obtain the dependence of Γ^{gg}_{BSM} with the assumption that it has the same dependence as the cross section since the respective coupling does not contribute to the $H \to ZZ^*$ decay. For the $f^{WW/ZZ}$ correction functions ggF samples with $H \to ZZ^* \to 4\ell$ ($\ell = e, \mu$) and $H \to WW^* \to \ell\nu\ell\nu$ ($\ell = e, \mu, \tau$) decays are produced with different values of κ_{HVV} and κ_{AVV}. Since these couplings are contributing only in the decay, the dependence of $\Gamma^{WW/ZZ}_{\text{BSM}}$ is obtained directly without any assumption. The input samples used for the derivation of the constant coefficients c_i are summarised in Table 7.8. As for the signal modelling with the morphing method, the inputs for the calculation of the f^{gg} correction function are symmetrised. Three input samples are used for the f^{gg} correction function and six for $f^{WW/ZZ}$. The obtained values of the coefficients c_i^{gg}, c_i^{WW} and c_i^{ZZ} are summarised in Table 7.9.

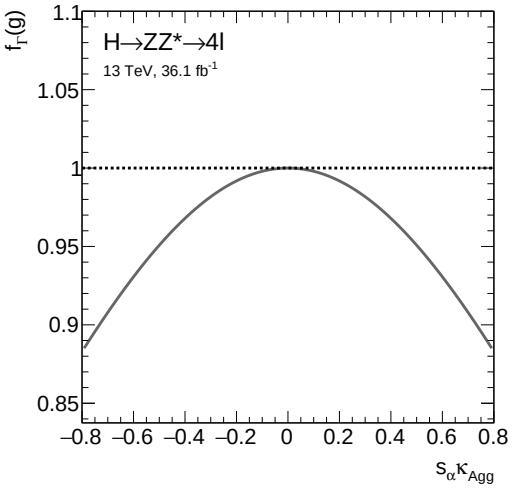
The resulting final correction f_Γ of the total decay width is shown in Figure 7.11 in dependence on the BSM coupling parameters $s_\alpha\kappa_{Agg}$, $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$. In the presence of non-vanishing CP-odd BSM coupling parameter the correction of the total decay width amounts up to -9% for the sensitive range $s_\alpha\kappa_{Agg} = \pm 0.7$ and up to -1% for $s_\alpha\kappa_{AVV} = \pm 6$. Due to the quadratic dependence on the CP-odd BSM contributions, the correction functions are symmetric around the SM value. Since the CP-even BSM interactions interfere with the SM ones the corresponding correction function is not symmetric around the SM point. The expected event yield is scaled up by $+11\%$ for $c_\alpha\kappa_{HVV} = -6$ and down by -13% for $c_\alpha\kappa_{HVV} = +6$ after applying the described corrections.

Table 7.8: List of input samples and their respective cross section times branching ratio employed for the calculation correction functions f^{gg}, f^{WW} and f^{ZZ} with the morphing method

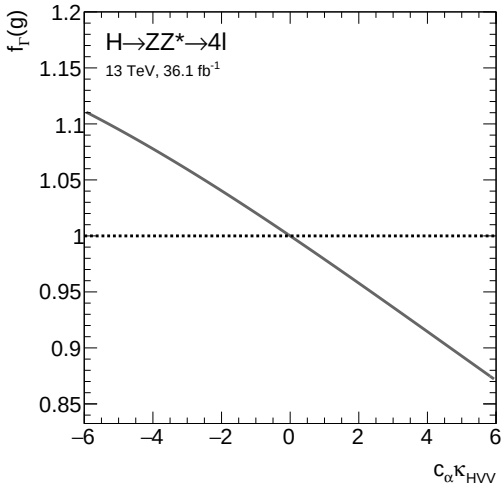
Input sample	BSM parameters						$\sigma \cdot \mathcal{B}$ [fb]	
	c_α	κ_{Hgg}	κ_{SM}	κ_{Agg}	κ_{HVV}	κ_{AVV}	$H \to ZZ^*$	$H \to WW^*$
f^{gg} correction function								
SM	1	1	1	0	0	0	4.1	–
Mixture	$1/\sqrt{2}$	1	1	0.5	0	0	1.6	–
Mixture	$1/\sqrt{2}$	1	1	−0.5	0	0	1.6	–
f^{WW} and f^{ZZ} correction function								
SM	1	1	1	0	0	0	4.1	745
BSM CP-even	1	1	0	0	1	0	$5.3 \cdot 10^{-3}$	2
Mixture	1	1	1	0	−6	0	2.8	418
BSM CP-odd	$1/\sqrt{2}$	1	0	0	0	1	$0.5 \cdot 10^{-3}$	0.2
Mixture	$1/\sqrt{2}$	1	1	0	6	6	1.5	317
Mixture	$1/\sqrt{2}$	1	1	0	−6	−6	0.7	120

Table 7.9: Values of the coefficients in the functions f^{XX} ($XX = g, W, Z$) correcting the decay width of the Higgs boson in presence of BSM contributions

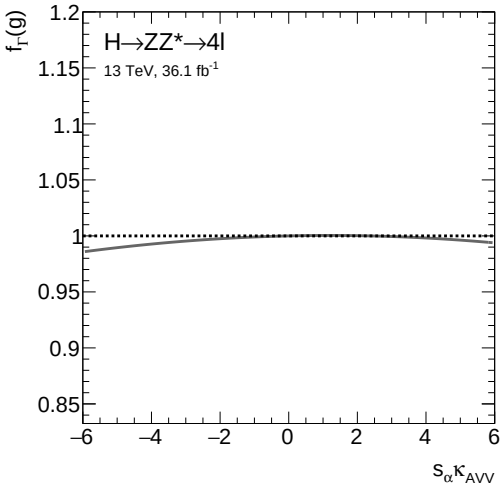
Correction function	Coefficients				
	c_1^{XX}	c_2^{XX}	c_3^{XX}	c_4^{XX}	c_5^{XX}
f^{gg}	2.3	0	–	–	–
f^{WW}	$3.0 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$	0.09	$-3.2 \cdot 10^{-3}$	$0.5 \cdot 10^{-3}$
f^{ZZ}	$1.2 \cdot 10^{-3}$	$0.5 \cdot 10^{-3}$	0.06	$-0.4 \cdot 10^{-3}$	$0.02 \cdot 10^{-3}$



(a)



(b)



(c)

Figure 7.11: Total width correction f_Γ as a function of (a) $s_\alpha\kappa_{Agg}$, (b) $c_\alpha\kappa_{HVV}$ and (c) $s_\alpha\kappa_{AVV}$.

7.2.3 Systematic Uncertainties

The HC analysis of the coupling tensor structure is affected by the same sources of experimental and theoretical systematic uncertainties as in the case of the cross section measurements described in Section 5.4. Since this analysis is performed on a smaller data set, the size of the systematic uncertainties is in general slightly larger. However, the analysis of the 36.1 fb^{-1} is dominated by the statistical error, such that the systematic uncertainties have a small impact on the results.

The uncertainty on the integrated luminosity is measured with a precision of 3%. A 2% uncertainty is assigned on the event yields due to the uncertainties on the modelling of the pile-up. Uncertainties on the lepton identification efficiencies are from $0.5\% - 1.0\%$ for muons and from $1\% - 1.3\%$ for electrons. The uncertainty on the lepton isolation efficiency is about 2%, while uncertainties on lepton energy scale and resolution are small and have a negligible impact on the final state result. Jet energy scale and resolution uncertainties are in the range from $3\% - 7\%$ and $2\% - 4\%$, respectively. Flavour tagging uncertainties due to the efficiency of the b-tagging algorithm are in the order of a few percent over most of the jet p_T range. Three sources of uncertainties associated with the data-driven estimation of the reducible background are taken into account. The first is the uncertainty of about 6% on the inclusive background estimation from the determination of the event selection efficiencies related to the lepton identification, isolation and impact parameter significance. The second source of uncertainty with the size of about 4% is also assigned to the inclusive background estimation and corresponds to the statistical uncertainty of the control data sample. It is in the order of 4%. The last and the largest uncertainty of up to 13% is due to the modelling and statistical uncertainties after the classification into the reconstructed event categories.

The theoretical uncertainties are derived using the same methods described in Section 5.4.2. As in the case of the signal strength measurement the impact of both the uncertainties on the acceptance and on the prediction of the SM cross sections is taken into account. QCD scale uncertainties affect the ggF signal prediction by up to 18% and VBF and VH production modes by up to 10%. For all production modes the PDF uncertainties are in the order of 2%. Uncertainties related to the parton shower amount to about 5%.

In addition, further uncertainties are assigned specifically for the HC analysis of the coupling tensor structure. An uncertainty on the VBF+VH signal is assigned to account for the uncertainty of the assumption of the best-prediction-scaling that higher-order scale factors K are the same for the SM and the BSM signal hypotheses. In order to qualify this uncertainty, additional reconstructed BSM VBF+VH samples are generated at NLO QCD accuracy with MadGraph5_aMC@NLO for the same values of BSM

coupling parameters as for the nominal LO BSM samples. The expected number of events predicted by these LO and NLO samples is compared in each reconstructed category and the relative difference is assigned as uncertainty. Since the best-prediction scale factors are derived at the SM point, no such uncertainty is assigned on the SM hypothesis. Figure 7.12 shows the best-prediction uncertainties in each reconstructed event category for the five BSM samples with largest uncertainty. Depending on the reconstructed event category and the point in the BSM parameter space, the systematic uncertainties on the best-prediction-scaling range from 10% to 40%. The total impact of this uncertainty on the final measurement is less than 1%.

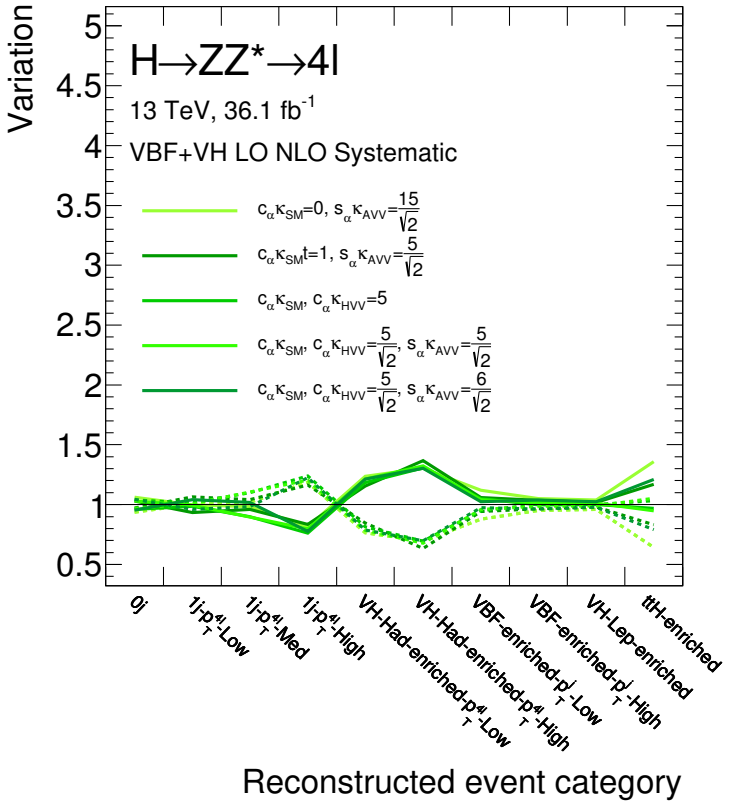


Figure 7.12: Systematic uncertainty on the best-prediction-scaling for the VBF+VH BSM signal model in each reconstructed event category.

The experimental and best-prediction systematic uncertainties on the VBF+VH signal yield are evaluated for the SM and all BSM signal samples and interpolated to an arbitrary point in the BSM parameter space using the morphing method. The theoretical signal uncertainties are only determined with the SM sample. It is assumed that they are of the same size also for all BSM signals.

In addition to the signal uncertainties, also the background uncertainties are taken into account. The largest uncertainty impacting the irreducible ZZ^* background is the QCD scale uncertainty. It affects the total event yield by 4% and the yields in the reconstructed event categories with additional jets by up to 30%. PDF and parton shower uncertainties affect the ZZ^* event yields by $1\%-2\%$ and $1\%-5\%$, respectively.

7.2.4 Statistical Analysis Model

As for the production cross section measurement (Section 5.5), the constraints on the BSM couplings within the HC framework are obtained from binned maximum likelihood fits to the data using the profiled likelihood ratio method.

The likelihood function is constructed as a product of Poisson distributions P corresponding to the observed number of events n_i in each of the reconstructed categories i for a given signal $s_i(\boldsymbol{\kappa},\boldsymbol{\theta})$ and background $b_i(\boldsymbol{\theta})$ expectation,

$$\mathcal{L}(\boldsymbol{\kappa},\boldsymbol{\theta}) = \prod_{i}^{N_{\text{Categories}}} P(n_i \mid s_i(\boldsymbol{\kappa},\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})) \times \prod_{m}^{N_{\text{Nuisance}}} C_m(\boldsymbol{\theta}), \tag{7.17}$$

with the vector of coupling parameters $\boldsymbol{\kappa} = \{c_\alpha\kappa_{Hgg}, s_\alpha\kappa_{Agg}, c_\alpha\kappa_{SM}, c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV}\}$. The vector $\boldsymbol{\theta} = \{\theta_1,\theta_2,\ldots,\theta_N\}$ represents the set nuisance parameters corresponding to the systematic uncertainties. The nuisance parameters are constraint by a Gaussian constrained term $C_m(\boldsymbol{\theta})$ which reflects the constraints from external measurement.

The constraints on the BSM coupling parameters are set by using a test statistic q similar as the one in Equation 5.10,

$$q = -2\ln\frac{\mathcal{L}\left(\boldsymbol{\kappa},\hat{\hat{\boldsymbol{\theta}}}\right)}{\mathcal{L}\left(\hat{\boldsymbol{\kappa}},\hat{\boldsymbol{\theta}}\right)} = -2\ln(\lambda), \tag{7.18}$$

based on a ratio of profiled likelihoods [193] of the conditional and the unconditional maximum-likelihood estimators in the numerator and the denominator, respectively, which is tested of the cross section $\boldsymbol{\sigma}$ in Equation 5.10, the BSM parameters $\boldsymbol{\kappa}$ are fitted as parameter of interest instead.

7.2.5 Results

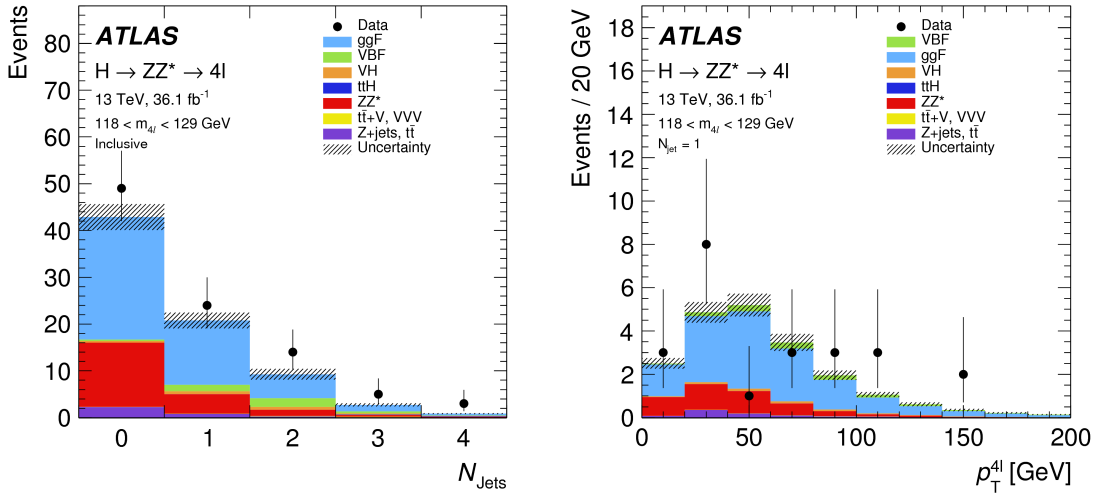
The expected number of events from the SM signal and the background as well as the observed number of events are shown in Table 7.10, separately in each reconstructed category resulting from the HC Analysis of the data set with an integrated luminosity of 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV. The data and the SM prediction are in a reasonable agreement. A slight excess of events above the SM prediction is observed. The largest deviation

from the SM prediction is observed in the VBF-enriched-p_T^j-Low and VBF-enriched-p_T^j-High categories, with respectively about two and four times more observed events than expected. The expected and the observed distributions of observables used for the classification of events into reconstructed event categories are shown in Figure 7.13 and Figure 7.14: the number of jets N_j in the final state, the invariant mass m_{jj} of the two leading jets, the transverse momentum p_T^{j1} of the leading jet and the transverse momentum $p_T^{4\ell}$ of the four-lepton final state. An excess of events is observed for the dijet invariant mass distribution in events with $N_j \geq 2$ and for the jet p_T^{j1} distribution in the subset of events with $m_{jj} > 120$ GeV (see Figure 7.13(c) and Figure 7.14(a)). All other distributions are in a good agreement with the SM expectation. The measurement of the Higgs boson tensor coupling structure is performed based on these results.

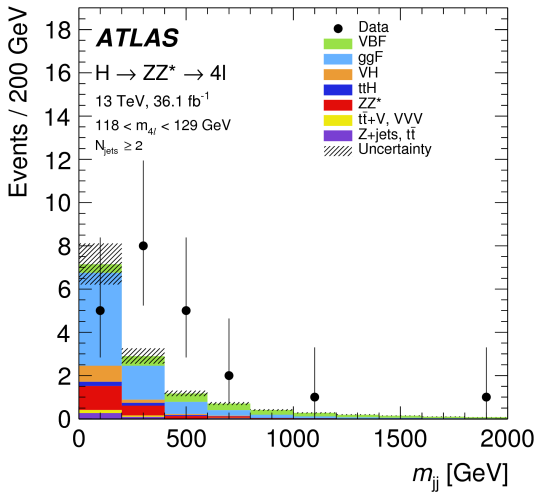
Table 7.10: Expected and observed number of events at an integrated luminosity of 36.1 fb^{-1} after the full event selection in the mass range of 118 GeV $< m_{4\ell} <$ 129 GeV at $\sqrt{s} = 13$ TeV in each reconstructed event category assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV. Statistical and systematic uncertainties are added in quadrature [182].

Reconstructed event category	SM Higgs boson signal	ZZ^*	Other backgrounds	Total expected	Observed
ggF-0j	26.8 $\pm$2.5	13.7 $\pm$1.0	2.23 $\pm$0.31	42.7 $\pm$2.7	49
1j-$p_T^{4\ell}$-Low	8.8 $\pm$1.1	3.1 $\pm$0.4	0.53 $\pm$0.07	12.5 $\pm$1.2	12
1j-$p_T^{4\ell}$-Med	5.4 $\pm$0.7	0.88 $\pm$0.12	0.38 $\pm$0.05	6.7 $\pm$0.7	9
1j-$p_T^{4\ell}$-High	1.47 $\pm$0.24	0.139 $\pm$0.022	0.045 $\pm$0.007	1.65 $\pm$0.24	3
VBF-enriched-p_T^j-Low	6.3 $\pm$0.8	1.08 $\pm$0.32	0.40 $\pm$0.04	7.7 $\pm$0.9	16
VBF-enriched-p_T^j-High	0.58 $\pm$0.10	0.093 $\pm$0.032	0.054 $\pm$0.006	0.72 $\pm$0.10	3
VH-Had-enriched-$p_T^{4\ell}$-Low	2.9 $\pm$0.5	0.63 $\pm$0.16	0.169 $\pm$0.021	3.7 $\pm$0.5	3
VH-Had-enriched-$p_T^{4\ell}$-High	0.64 $\pm$0.09	0.029 $\pm$0.008	0.0182 $\pm$0.0022	0.69 $\pm$0.09	0
VH-Lep-enriched	0.318 $\pm$0.019	0.049 $\pm$0.008	0.0137 $\pm$0.0019	0.380 $\pm$0.020	0
ttH-enriched	0.39 $\pm$0.04	0.014 $\pm$0.006	0.07 $\pm$0.04	0.47 $\pm$0.05	0
Total	54 $\pm$4	19.7 $\pm$1.5	3.9 $\pm$0.5	77 $\pm$4	95

Constraints on BSM coupling parameters are obtained with different set of assumptions. Firstly each coupling parameter $s_\alpha\kappa_{Agg}$, $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ is fitted separately assuming that all other BSM parameters are equal to zero. The parameters related to the SM-like tensor structure are set to the SM predictions, $c_\alpha\kappa_{Hgg} = 1$ and $c_\alpha\kappa_{SM} = 1$. In addition, the two BSM coupling parameters contributing to the XVV interaction vertex, $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ are probed also with a free-floating value of the SM-like coupling parameter $c_\alpha\kappa_{SM}$. In this way, it can be studied which fraction of the observed deviations from the SM can be absorbed in the SM-like coupling contribution related to the coupling strength modifiers of the κ framework.

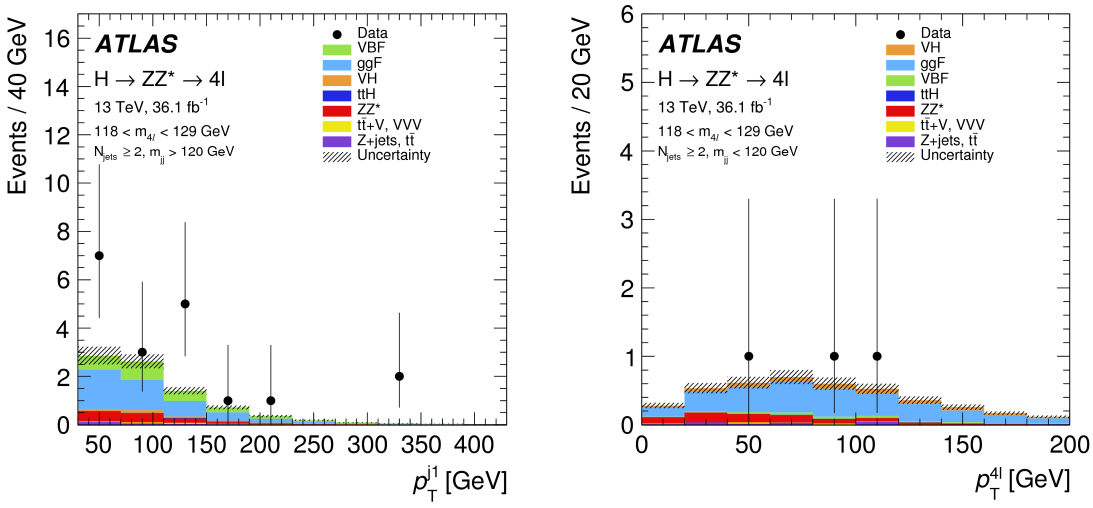


(a) (b)



(c)

Figure 7.13: The observed and expected distributions of (a) N_j after the inclusive selection, (b) $p_T^{4\ell}$ in the $1j$ categories and (c) m_{jj} in the $2j$ categories for an integrated luminosity of 36.1 fb^{-1} in the mass range of 118 GeV $< m_{4\ell} <$ 129 GeV at a centre-of-mass energy of $\sqrt{s} = 13$ TeV assuming a Higgs boson signal with a mass the SM $m_H = 125$ GeV [182].



(a) (b)

Figure 7.14: The observed and expected distributions of (a) p_T^{j1} in the VBF-enriched categories and (b) $p_T^{4\ell}$ in the VH-Had-enriched categories for an integrated luminosity of 36.1 fb^{-1} in the mass range of 118 GeV $< m_{4\ell} <$ 129 GeV at a centre-of-mass energy of $\sqrt{s} = 13$ TeV assuming a Higgs boson signal with a mass the SM $m_H = 125$ GeV [182].

Table 7.11 summarises the resulting constraints on BSM parameters obtained from the one-dimensional likelihood scans. The observed and expected values of the test statistics from the scan over the $s_\alpha\kappa_{Agg}$ parameter is shown in Figure 7.15. The excess observed in data in the $0j$ and VBF-enriched categories results in a non-zero best-fit value of this parameter $s_\alpha\widehat{\kappa}_{Agg} = \pm 0.43$, which corresponds to a small deviation from the SM expectation at the level of 1.8 σ. Since the size of the CP-odd BSM coupling contributions are not sensitive to the sign of the coupling parameter, the presented likelihood scan is symmetric around the SM value of $s_\alpha\kappa_{Agg} = 0$.

Figure 7.16 shows the observed and expected values of the test statistics from the scans of BSM coupling parameters $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ with fixed parameter $c_\alpha\kappa_{SM}$ of the SM-like coupling. The observed excess of data in the VBF-enriched categories leads to a non-zero best-fit value of $c_\alpha\kappa_{HVV} = 2.9$ and $s_\alpha\kappa_{AVV} = \pm 2.9$. By allowing the parameter $c_\alpha\kappa_{SM}$ of the SM-like coupling to be free-floating in the fit (Figure 7.17), the excess observed in data is partially absorbed leading to a best-fit value of $c_\alpha\widehat{\kappa}_{SM} = 1.2$, while the best-fit values of BSM coupling parameters move closer to the SM expectation. The corresponding deviation from the SM prediction also decreases from 2.3 σ (1.4 σ) to 1.7 σ (0.5 σ) for $c_\alpha\kappa_{HVV}$ ($s_\alpha\kappa_{AVV}$) parameters. Due to the interference of the BSM CP-even and the SM-like interactions, the data is sensitive to the sign of the $c_\alpha\kappa_{HVV}$ parameter. The fit to the data prefers the positive value. The CP-odd BSM coupling parameter in

the XVV vertex has no sign sensitivity. The deviations from the SM prediction for all one-dimensional scans of BSM coupling parameters are at most 2.3 σ, thus indicates no evidence of new couplings.

Table 7.11: The expected and observed 95% confidence level (CL) intervals from the one-dimensional likelihood scans over the BSM coupling parameters $s_\alpha\kappa_{Agg}$, $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$, together with the corresponding best-fit values and the deviation from the SM expectation. The results are obtained from the HC analysis using 36.1 fb^{-1} of data at a centre-of-mass energy of 13 TeV. The parameter related to the SM-like coupling to the gluons is fixed to the SM value, $c_\alpha\kappa_{Hgg} = 1$ [182].

BSM coupling	Fit	95% CL interval		Observed best-fit value		Deviation
κ_{BSM}	configuration	Expected	Observed	$\widehat{\kappa}_{\text{BSM}}$	$c_\alpha\widehat{\kappa}_{SM}$	from SM
$s_\alpha\kappa_{Agg}$	$c_\alpha\kappa_{SM} = 1$	$[-0.47, 0.47]$	$[-0.68, 0.68]$	± 0.43	–	1.8 σ
$c_\alpha\kappa_{HVV}$	$c_\alpha\kappa_{SM} = 1$	$[-2.9, 3.2]$	$[-0.8, 4.5]$	2.9	–	2.3 σ
$c_\alpha\kappa_{HVV}$	$c_\alpha\kappa_{SM}$ free	$[-3.1, 4.0]$	$[-0.6, 4.2]$	2.2	1.2	1.7 σ
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM} = 1$	$[-3.5, 3.5]$	$[-5.2, 5.2]$	± 2.9	–	1.4 σ
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM}$ free	$[-4.0, 4.0]$	$[-4.4, 4.4]$	± 1.5	1.2	0.5 σ

The CP-even (CP-odd) BSM coupling parameter $c_\alpha\kappa_{HVV}$ ($s_\alpha\kappa_{AVV}$) is also probed in a two-dimensional fit together with the parameter $c_\alpha\kappa_{SM}$ of the SM-like coupling. One BSM coupling is probed at a time with the other BSM coupling parameter set to zero. The corresponding best-fit values and the deviations from the SM expectation are summarised in Table 7.12. The corresponding observed and expected two-dimensional contours of the test statistics at 95% CL for the likelihood scans of the $(c_\alpha\kappa_{HVV}, c_\alpha\kappa_{SM})$ and $(c_\alpha\kappa_{HVV}, c_\alpha\kappa_{SM})$ parameter space are shown in Figure 7.18. Values of BSM coupling parameters inside the contours surrounding the SM point are allowed at 95% CL, while values outside are excluded. The best-fit values for the BSM coupling parameters, $c_\alpha\widehat{\kappa}_{HVV} = 2.1$ and $s_\alpha\widehat{\kappa}_{AVV} = \pm 1.5$ from the two-dimensional scans are very similar to those obtained from one-dimensional scans with a free-floating SM-like coupling parameter (Table 7.11). The data prefer positive values of the SM-like coupling parameter, again $c_\alpha\widehat{\kappa}_{SM} = 1.2$ for both scans. As in the case of one-dimensional scans, there is no sign sensitivity to the BSM CP-odd coupling. The deviation from the SM expectation are at most 2.1 σ, providing no evidence for new BSM couplings contributions.

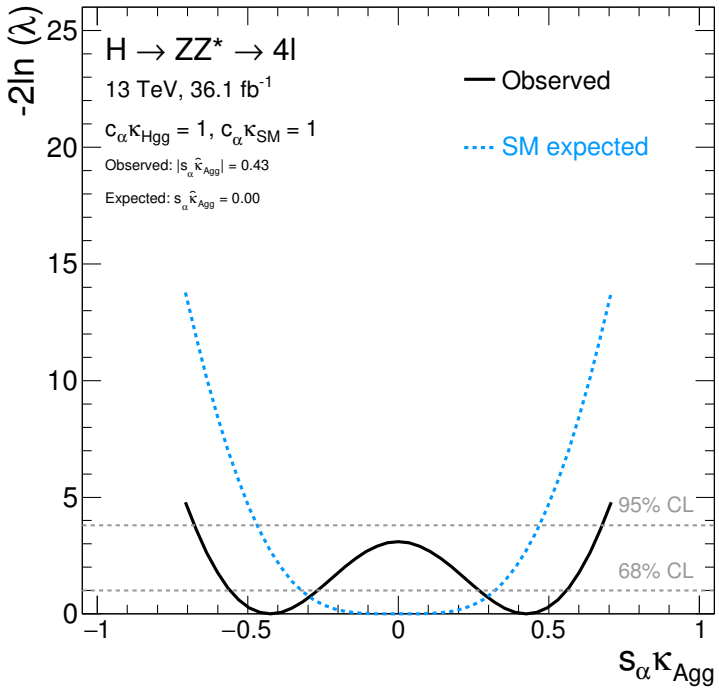
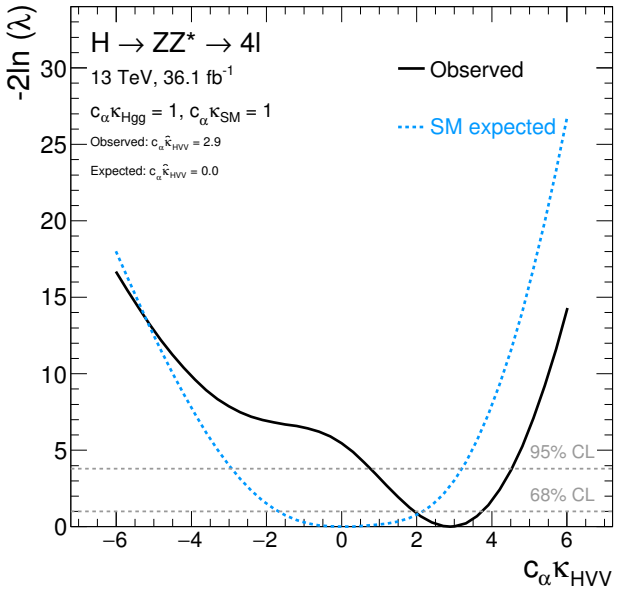
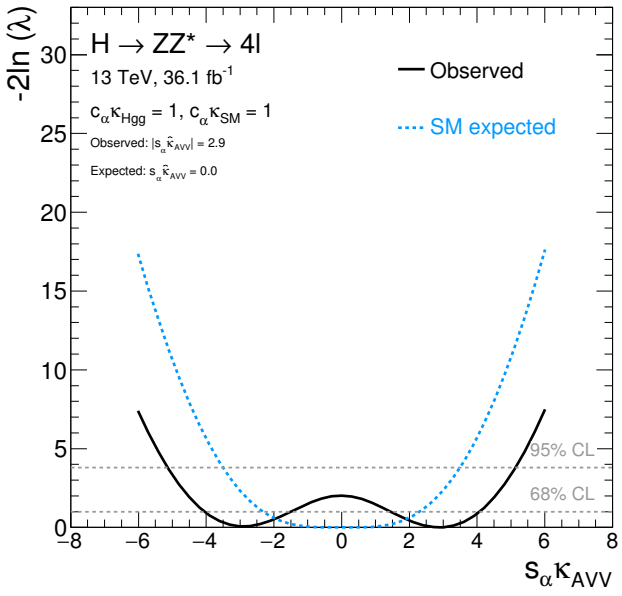


Figure 7.15: The observed (solid black line) and SM expected (dashed blue line) test statistic $q = -2\ln(\lambda)$ from the scan of the BSM coupling parameter $s_\alpha \kappa_{Agg}$ using 36.1 fb^{-1} of data at a centre-of-mass energy of 13 TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.

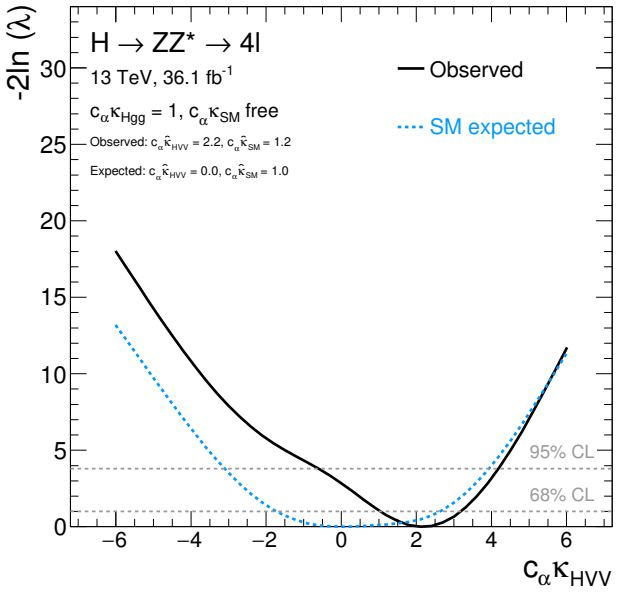


(a)

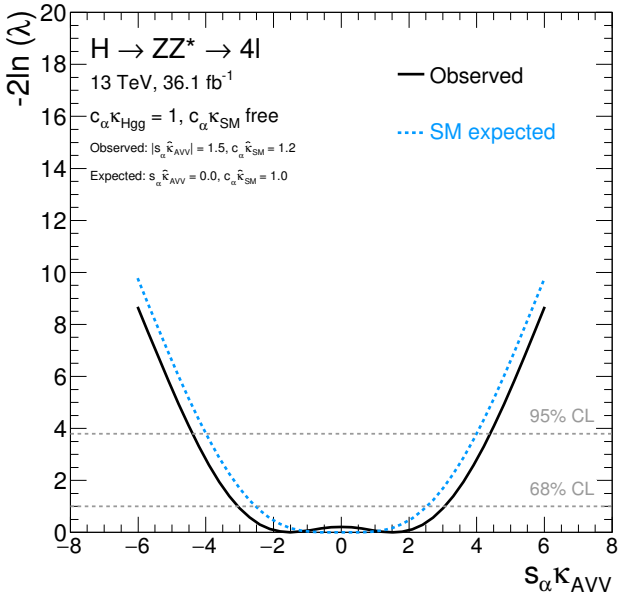


(b)

Figure 7.16: The observed (solid black line) and SM expected (dashed blue line) test statistic $q = -2\ln(\lambda)$ from scans of the BSM coupling parameters (a) $c_\alpha\kappa_{HVV}$ and (b) $s_\alpha\kappa_{AVV}$ with parameters of the SM-like couplings fixed to $c_\alpha\kappa_{SM} = 1$. Results are obtained with 36.1 fb^{-1} data at a centre-of-mass energy of 13 TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.

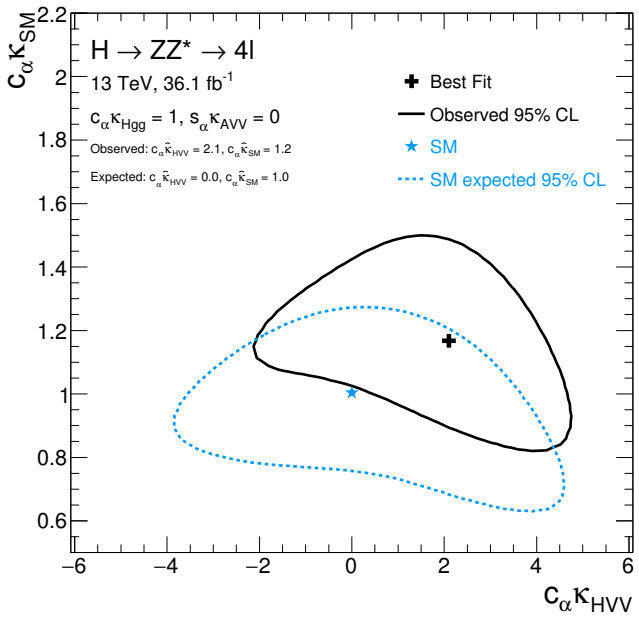


(a)

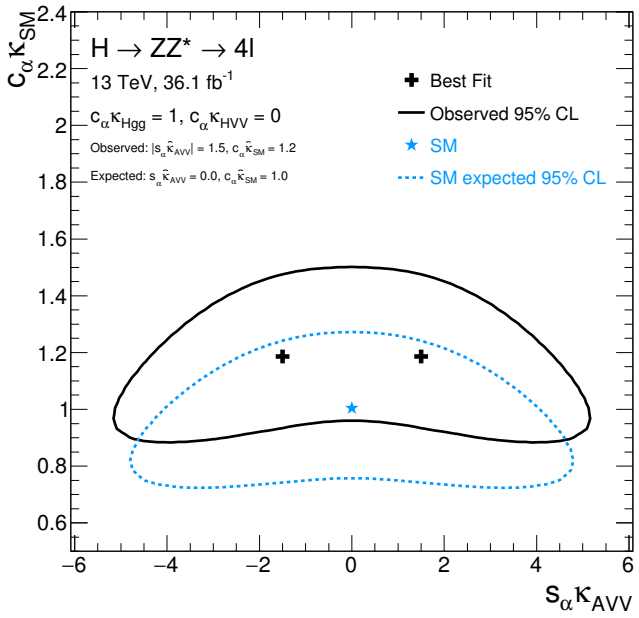


(b)

Figure 7.17: The observed (solid black line) and SM expected (dashed blue line) test statistic $q = -2\ln(\lambda)$ from scans of the BSM coupling parameters (a) $c_\alpha\kappa_{HVV}$ and (b) $s_\alpha\kappa_{AVV}$ with parameter of the SM-like coupling $c_\alpha\kappa_{SM}$ as free parameter of the fit. Results are obtained with 36.1 fb^{-1} data at a centre-of-mass energy of 13 TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.



(a)



(b)

Figure 7.18: The observed (black solid line) and SM expected (dashed blue line) contours of the two-dimensional negative log-likelihood at 95% CL for the (a) $c_\alpha\kappa_{SM}$ and $c_\alpha\kappa_{HVV}$ and (b) $c_\alpha\kappa_{SM}$ and $s_\alpha\kappa_{AVV}$ BSM coupling parameters using 36.1 fb^{-1} data at a centre-of-mass energy of 13 TeV.

Table 7.12: The best-fit values and the corresponding deviation from the SM prediction obtained from the two-dimensional likelihood scans of the $(\kappa_{XVV}, \kappa_{SM})$ parameter space performed with 36.1 fb^{-1}data at a centre-of-mass energy of 13 TeV. The parameter of the SM-like coupling gluons is kept at the SM value, $c_\alpha\kappa_{Hgg} = 1$ in the fits.

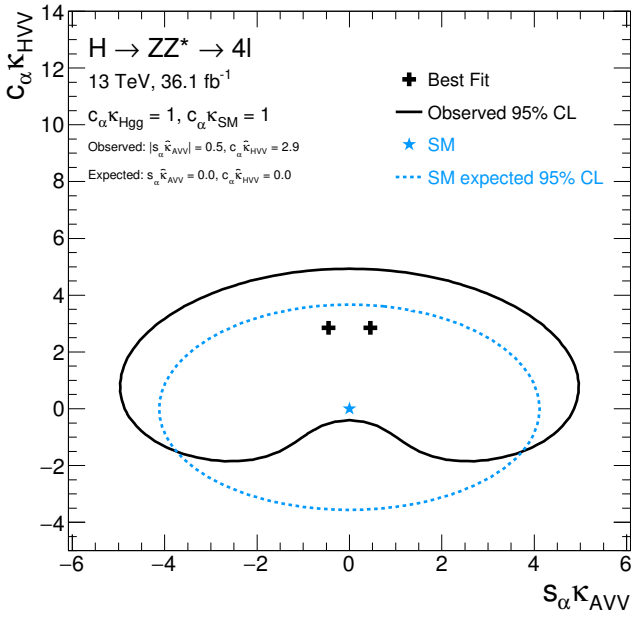
Fit	Observed best-fit value			Deviation
configuration	$c_\alpha\widehat{\kappa}_{HVV}$	$s_\alpha\widehat{\kappa}_{AVV}$	$c_\alpha\widehat{\kappa}_{SM}$	from SM
$(c_\alpha\kappa_{HVV}, c_\alpha\kappa_{SM})$, $s_\alpha\kappa_{AVV} = 0$	2.1	–	1.2	2.1 σ
$(c_\alpha\kappa_{HVV}, c_\alpha\kappa_{SM})$, $c_\alpha\kappa_{AVV} = 0$	–	$\pm$1.5	1.2	1.5 σ

Finally, the constraints on the BSM contributions in the XVV vertex are set simultaneously on the $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ coupling parameters, assuming either $c_\alpha\kappa_{SM} = 1$ or letting $c_\alpha\kappa_{SM}$ as free-floating in the fit. The corresponding two-dimensional likelihoods are symmetrised with respect to the sign of the CP-odd BSM coupling parameter to avoid nonphysical results due statistical fluctuations. A more direct approach of symmetrising the input samples for the two-dimensional signal modelling was not possible with the available simulated input samples. The best-fit values of fitted parameters and the deviation from the SM prediction are summarised in Table 7.12. The respective observed and expected two-dimensional contours of the test statistics at 95% CL for the likelihood scans of the $(c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV})$ parameter space are shown in Figure 7.19. The best-fit value $c_\alpha\widehat{\kappa}_{HVV} = 2.9$ in the scan with a fixed SM-like coupling parameter is similar to the one obtained from the one-dimensional scan, while value $s_\alpha\widehat{\kappa}_{AVV} = \pm 0.5$ is closer to the SM point. The respective deviation from the SM prediction is 1.9 σ.

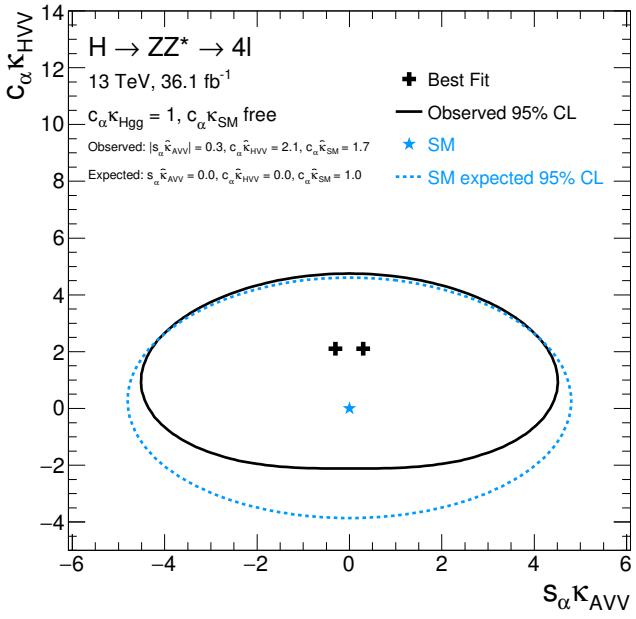
Table 7.13: The best-fit parameter values and the corresponding deviation from the SM prediction obtained from the two-dimensional likelihood scans in the $(c_\alpha\kappa_{HVV}, s_\alpha\kappa_{AVV})$ parameter space or a free-floating value of the SM-like coupling parameter performed with 36.1 fb^{-1} of data at a centre-of-mass energy of 13 TeV. The parameter of the SM-like coupling to gluons is fixed to the SM value, $c_\alpha\kappa_{Hgg} = 1$ [182].

Fit	Observed best-fit			Deviation
configuration	$c_\alpha\widehat{\kappa}_{HVV}$	$s_\alpha\widehat{\kappa}_{AVV}$	$c_\alpha\widehat{\kappa}_{SM}$	from SM
$c_\alpha\kappa_{SM} = 1$	2.9	$\pm$0.5	–	1.9 σ
$c_\alpha\kappa_{SM}$ free	2.1	$\pm$0.3	1.7	1.2 σ

By letting $c_\alpha \kappa_{SM}$ be free-floating parameter in the fit, the coupling parameter phase space is expected to become less constrained. However, this is not fully the case for the observed test statistics. The excess in data is partially absorbed by the SM-like coupling parameter ($c_\alpha \widehat{\kappa}_{SM} = 1.7$), reducing also the deviation from the SM prediction to 1.2 σ. Therefore, a larger region of the phase space can be excluded for the BSM CP-odd coupling parameter compared to the fit with $c_\alpha \kappa_{SM} = 1$. For the CP-even BSM parameter the excluded area is similar to the one with fixed value $c_\alpha \kappa_{SM} = 1$, except for negative values close to the SM point. In this region, the interference with the SM CP-even coupling is the strongest, corresponding to a minimum value of the predicted event yield. The observed excess of data can in this region only be compensated by the free-floating of the SM-like coupling in the fit. The different behaviour of the observed test statistics compared to the expected one is caused by the fact, that the observed excess in data is more consistent with a modification of the SM-like coupling, than with a non-zero value of the BSM CP-odd coupling.



(a)



(b)

Figure 7.19: The observed (black solid line) and SM expected (dashed blue line) contours of the two-dimensional negative log-likelihood at 95% CL for the BSM coupling parameter space $(s_\alpha\kappa_{AVV}, c_\alpha\kappa_{HVV})$ with the SM coupling parameter $c_\alpha\kappa_{SM}$ (a) fixed to the SM value of one and (b) free floating in the fit, obtained with 36.1 fb^{-1} of data at a centre-of-mass energy of 13 TeV.

7.3 Interpretation within the Standard Model Effective Field Theory

In the second approach that probes the tensor structure of the Higgs boson couplings to weak vector bosons, gluons and top quarks, in this thesis, the results of the cross section measurements in particle level production bins are interpreted in terms of the parameters of the SMEFT model (see Section 2.5.2.1). The main difference of this so-called EFT interpretation compared to the study presented in Section 7.2 are the observables used in the fit to data. While in the HC analysis (Section 7.2), the BSM predictions are fitted to event yields in each reconstructed event category, the study presented in this section interprets the measured particle level cross sections in each bin of the STXS framework in terms of EFT parameters. In the former approach, the BSM contributions are fully parametrised up to the reconstruction level, including production cross sections, branching ratio, detector acceptance as well as the signal reconstruction efficiency. In the latter approach, the cross sections, branching ratio and acceptance are parametrised, while the reconstruction efficiency is assumed not to be affected by BSM contributions. This assumption should be valid for the studied decay channel as presented later on. It allows for an EFT interpretation at particle level, instead of relying on the resources consuming full detector simulation for the BSM signal samples.

The interpretation is based on the results of the SM production cross section measurement obtained with the full Run 2 data set, which is similar to one performed with the 79.8 fb^{-1} data as presented in Chapter 5. Differences between the cross section measurements with the two data sets are described in the next subsection. The production cross sections, the branching ratio as well as the acceptance in each particle level production bin are parametrised as functions of BSM coupling parameters (Wilson coefficients) using particle level event generation. These predictions are then fitted to the corresponding measurement.

With the described EFT interpretation, constraints are set on ten CP-even and CP-odd BSM coupling parameters of the SMEFT model (summarised in Table 2.7). The entire analysis was developed as a part of this thesis, together with the derivation and implementation of the signal parametrisation for the CP-even BSM coupling parameters. The parametrisation for the CP-odd BSM coupling parameters has been performed elsewhere following the same approach. Therefore, the evaluation of the parametrisation and the validation with fully reconstructed signal samples are discussed in detail for the CP-even BSM coupling parameters in the following. The corresponding studies for the CP-odd BSM coupling parameters are very similar and summarised in Appendix C.

In general, if the Wilson coefficients are small enough, the so-called linear approximation holds, in which only terms with linear dependence on Wilson coefficients are significant, while the quadratic terms can be neglected. Taking into account only the dimension-six terms and neglecting higher dimensions, these interference terms between the SM and the introduced BSM physics are suppressed by a factor $1/\Lambda^2$. However, in the $H \to ZZ^* \to 4\ell$ decay channel, the linear approximation is not valid for the VBF+VH-Had, VH-Lep and ttH+tH production modes. As discussed later, the sensitivity to the corresponding BSM parameters is in the range of parameter values for which also quadratic terms have a significant impact. Therefore, the dimension-six quadratic terms suppressed by a factor $1/\Lambda^4$ are also taken into account for the presented study. The linear terms from the CP-odd operators do not contribute to the total production cross section. The sensitivity to CP-odd BSM parameters is only reached when taking into account the dimension-six quadratic terms suppressed by a factor $1/\Lambda^4$. Linear terms from dimension-eight operators are also suppressed by a factor $1/\Lambda^4$ and can therefore in general contribute by a similar amount as the quadratic dimension-six terms. These dimension-eight terms are currently not available in the SMEFT model and are neglected in the following.

The following Wilson coefficients are probed (see Section 7.3.6): c_{HW}, c_{HB} and c_{HWB} related to the Higgs boson couplings to vector bosons, c_{HG} for the Higgs boson coupling to gluons and c_{uH} for the top Yukawa coupling. With the current amount of data it is only possible to measure one or two BSM couplings at a time, while others have to be fixed to a constant value (zero for the SM). Some of the SMEFT Wilson coefficients are already constrained using the LEP and LHC data [200]. Therefore, contributions from these BSM coupling parameters are neglected in the presented analysis, including their contribution to the total Higgs boson decay width. No constraints are evaluated for the mixtures of CP-even and CP-odd BSM contributions, since it is difficult to model the BSM dependence of the detector acceptance in multi-dimensional phase space with such a high number of free parameters as discussed later.

7.3.1 Higgs Boson Cross Section Measurements for Full Run 2

For the measurement of the Higgs boson cross section in particle level production bins using the full Run 2 data set [201] with an integrated luminosity of 139 fb^{-1}, the Higgs boson candidates are reconstructed by the inclusive analysis described in Section 4.3. Only the candidates in the mass window of 115 GeV $< m_{4\ell} <$ 130 GeV are considered.

The analysis follows closely the previous one with 79.8 fb^{-1} of data (see Chapter 5). However, the definition of reconstructed event categories and the discriminating variables used to improve the background rejection and to distinguish between different Higgs boson production modes have been optimised for the larger data set to improve the

measurement sensitivity. In addition, the larger amount of data allows to constrain the normalisation of the ZZ^* and tXX background with dedicated samples of control data from side-bands of the $m_{4\ell}$ distribution. Systematic uncertainties are calculated in a similar way as described in Section 5.4. A detailed description is given in [201].

The definition of particle level bins for the production cross section measurement in exclusive phase space regions has also been updated (see Figure 7.20) compared to the 79.8 fb^{-1} analysis (Figure 5.7). The larger amount of data allows for a finer granularity of the production bins. In addition, the definition of several production bins is adjusted to better fit the full STXS Stage-1 scheme (see Section 5.1.2). A detailed description of the updated scheme, the so-called Reduced-Stage-1.1, is given in Section 5.1.3. The $gg \to Z(\to jj) + H$ production process is now considered as a part of the gluon fusion production, rather than of the VH production. The corresponding production bin is therefore renamed from ggF to gg2H. Instead of the five ggF production bins, there are new seven gg2H bins. Before any classification according to the number of jets, a high-p_T^H bin ($p_T^H > 200$ GeV) is split targeting the BSM contributions. The remaining events are then split into categories with 0, 1 and ≥ 2 jets. A threshold of $p_T^H = 10$ GeV is introduced in the $0j$ gg2H bin, defining the two gg2H-$0j$ categories. The remaining gg2H production bins remain the same. Compared to the previous STXS scheme, the VBF and VH-Had production modes are combined into a single qq2Hqq production bin, which is further divided into three exclusive production bins. The production bins for the VH-Lep and ttH production mode are unchanged. To be consistent with the qq2Hqq production bin, the VH-Lep bin is referred as qq/gg2HLep production bin the following.

The redefinition of the particle level production bins calls also for an adjusted definition of the reconstructed event categories. An overview of the reconstructed event categories as well as the new introduced side-band categories is shown in Figure 7.20, together with production bins of the Reduced-Stage-1.1 scheme.

The ttH production mode is targeted by means of requirements on the number of b-jets with a certain b-jet selection efficiency $\eta\%$ ($N_b^{\eta\%}$). Events are collected in the ttH-Lep-enriched category, requiring at least one additional lepton in the final state and one of the following requirements: $N_b^{85\%} \geq 2$, $N_b^{85\%} \geq 1$ and $N_j \geq 5$ or $N_b^{60\%} \geq 1$. Remaining events satisfying $N_j \geq 5$ and $N_b^{85\%} \geq 2$ or $N_j \geq 4$ and $N_b^{85\%} \geq 2$ are assigned to the ttH-Had-enriched event category. Events with one additional lepton which do not fulfil the ttH requirements are classified into the VH-Lep-enriched event category.

The remaining events are categorised according to the jet multiplicity. Higgs boson candidates produced via VBF or VH-Had production are targeted by requiring at least two jets in the final state. Events with a dijet invariant mass larger than 120 GeV and

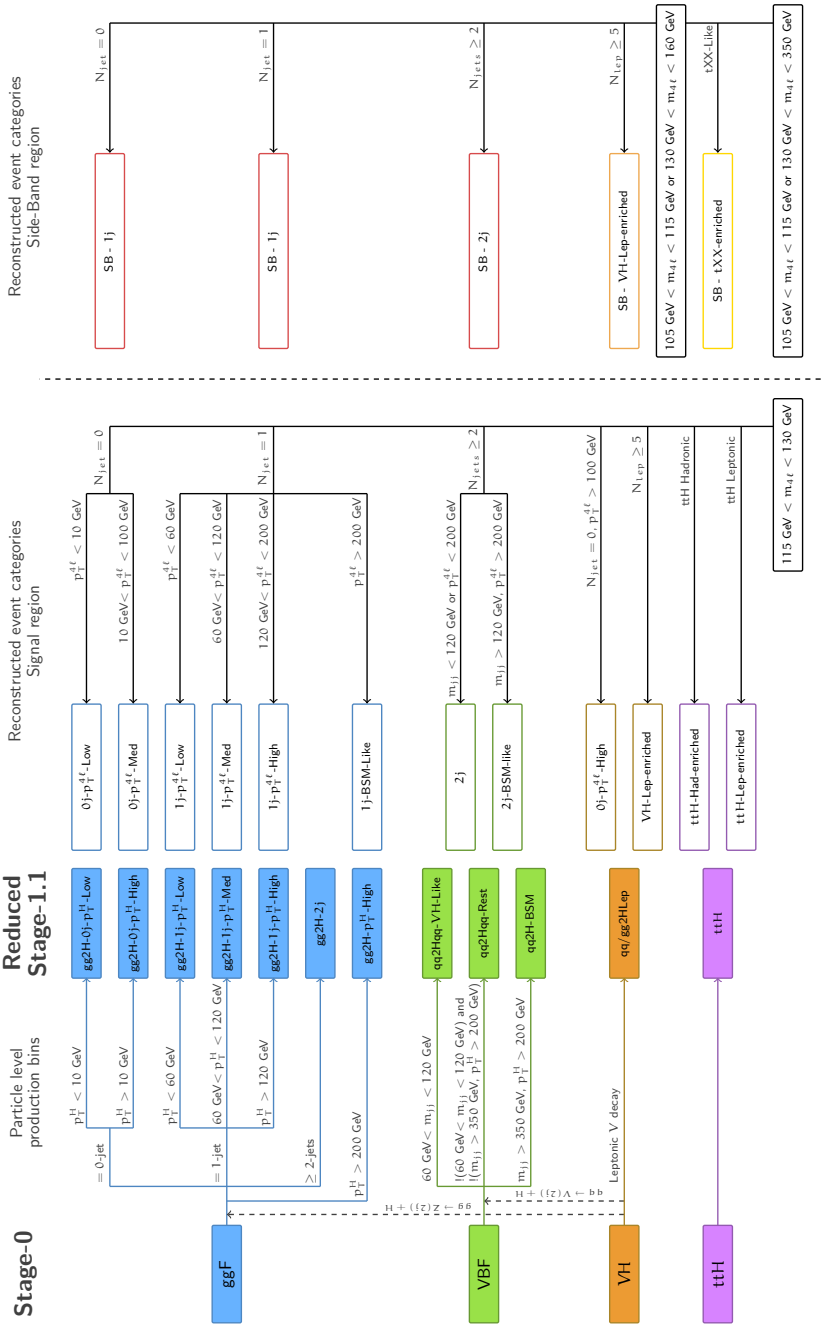


Figure 7.20: The exclusive regions of the Higgs boson production phase space (production bins) and the corresponding reconstructed event categories. The regions for the measurement of the Higgs boson production cross sections are defined at the particle level in two stages, STXS Stage-0 and Reduced-Stage-1.1.

$p_T^{4\ell} > 200$ GeV are assigned to the $2j$-BSM-Like reconstructed category which is expected to be sensitive to BSM contributions. Events failing these requirements are collected in the $2j$ category.

Events with one additional jet in the final state are enriched in both ggF and VBF contributions. Motivated by the particle level production bins they are divided into four categories with different requirements on the four-lepton transverse momentum: smaller than 60 GeV ($1j$-$p_T^{4\ell}$-Low); between 60 and 120 GeV ($1j$-$p_T^{4\ell}$-Med); between 120 and 200 GeV ($1j$-$p_T^{4\ell}$-High); and larger than 200 GeV ($1j$-$p_T^{4\ell}$-BSM).

Finally, the three categories with no jets in the final state are the $0j$-$p_T^{4\ell}$-Low ($p_T^{4\ell} <$ 10 GeV), $0j$-$p_T^{4\ell}$-Med (10 GeV$< p_T^{4\ell} < 100$ GeV) and $0j$-$p_T^{4\ell}$-High ($p_T^{4\ell} > 100$ GeV), motivated by the corresponding particle level production bins. The requirement $p_T^{4\ell} > 100$ GeV enhances the contribution of the $V(\to \ell\nu/\ell\ell)H$ process.

Additional side-band categories are introduced to constrain the background contributions, selecting events in the mass windows of 105 GeV $< m_{4\ell} <$ 115 GeV and 130 GeV $< m_{4\ell} <$ 350 GeV. Side-band events with at least two jets in the final state, including at least one b-tagged jet with 60% b-tagging efficiency and $E_T^{miss} > 100$ GeV are assigned to the SB-tXX-enriched category. The larger side-band mass range introduced above is only used in this category, to improve the statistical precision. For the remaining events, the upper mass window is reduced to 130 GeV $< m_{4\ell} <$ 160 GeV. Events in this mass window with an additional lepton in the final state are assigned to the SB-VH-Lep-enriched category. The remaining events are split according to the number of jets into the SB-$2j$, SB-$1j$ and SB-$0j$ categories.

The numbers of expected and observed events in each reconstructed event category are shown in Table 7.14 together with the number of events in the side-band regions. The expected event yields are in a reasonable agreement with the observed ones. A small excess of about two times more observed than expected events is observed in the $0j$-$p_T^{4\ell}$-High and in the $2j$-BSM-Like reconstructed event categories. A deficit of events is observed in the $1j$-$p_T^{4\ell}$-Low category. The event yields in the side-band categories constraining the ZZ^* contributions are in good agreement with the SM prediction, while a small excess is observed in the SB-VH-Lep-enriched and SB-tXX-enriched categories.

As for the previous analysis of the 79.8 fb^{-1} data set, the separation of contributions from different production modes in each reconstructed event category is improved by means of multivariate discriminating observables. Neural network (NN) discriminants [202, 203] are used, instead of the BDT discriminants (Section 5.3), improving the measurement sensitivity by up to 20% in several particle level production bins.

Table 7.14: Expected and observed number of events after the full event selection at an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV in each reconstructed event category assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV. Statistical and systematic uncertainties are added in quadrature [201].

Reconstructed event category	SM Higgs boson signal	ZZ^* background	tXX background	Other backgrounds	Total expected	Observed
Signal region: 115 GeV < m_H < 130 GeV						
$0j$-$p_T^{4\ell}$-Low	24.2 ± 3.5	30 ± 4	–	0.93 ± 0.13	55 ± 5	56
$0j$-$p_T^{4\ell}$-Med	76 ± 8	37 ± 4	–	6.5 ± 0.6	120 ± 9	117
$0j$-$p_T^{4\ell}$-High	0.355 ± 0.031	0.020 ± 0.012	0.0094 ± 0.0027	0.30 ± 0.05	0.69 ± 0.06	1
$1j$-$p_T^{4\ell}$-Low	34 ± 4	15.5 ± 2.7	–	1.91 ± 0.29	52 ± 5	41
$1j$-$p_T^{4\ell}$-Med	20.8 ± 2.8	4.0 ± 0.7	0.114 ± 0.013	1.02 ± 0.19	26.0 ± 2.9	31
$1j$-$p_T^{4\ell}$-High	4.7 ± 0.8	0.48 ± 0.10	0.043 ± 0.008	0.27 ± 0.04	5.5 ± 0.8	4
$1j$-$p_T^{4\ell}$-BSM	1.23 ± 0.23	0.069 ± 0.031	0.0067 ± 0.0031	0.062 ± 0.012	1.37 ± 0.23	2
$2j$	38 ± 5	9.1 ± 2.7	0.95 ± 0.08	2.13 ± 0.31	50 ± 6	48
$2j$-BSM-Like	3.3 ± 0.6	0.18 ± 0.06	0.032 ± 0.005	0.091 ± 0.017	3.6 ± 0.6	6
VH-Lep-enriched	1.29 ± 0.07	0.156 ± 0.025	0.039 ± 0.009	0.0194 ± 0.0032	1.50 ± 0.08	1
ttH-Had-enriched	1.02 ± 0.18	0.058 ± 0.025	0.252 ± 0.032	0.119 ± 0.033	1.45 ± 0.18	2
ttH-Lep-enriched	0.42 ± 0.04	0.002 ± 0.005	0.0157 ± 0.0023	0.0028 ± 0.0029	0.44 ± 0.04	1
Side-band region: 105 GeV < m_H < 115 GeV or 130 GeV < m_H < 160 GeV						
SB-$0j$	4.5 ± 0.5	150 ± 13	–	16.2 ± 2.2	171 ± 13	183
SB-$1j$	2.80 ± 0.30	51 ± 7	1.29 ± 0.16	8.4 ± 1.2	63 ± 7	64
SB-$2j$	2.02 ± 0.27	25 ± 7	4.4 ± 0.5	6.0 ± 0.9	38 ± 7	41
SB-VH-Lep-enriched	0.273 ± 0.015	0.48 ± 0.06	0.125 ± 0.018	0.126 ± 0.019	1.00 ± 0.07	3
Side-band region: 105 GeV < m_H < 115 GeV or 130 GeV < m_H < 350 GeV						
SB-tXX-enriched	0.071 ± 0.012	0.32 ± 0.12	12.1 ± 1.3	0.84 ± 0.33	13.3 ± 1.4	19

As in the case of BDT discriminants the NN discriminants are also trained in each reconstructed event category using several discriminating input distributions in simulated SM Higgs boson signal and background samples. Two types of neural networks are combined to define the final discriminant: one multilayer perceptron (MLP); and two recurrent neural networks (rNN), one for the four-momenta of the four leptons and one the four-momenta for up to three jets. These three components are chained into another MLP to obtain the final NN discriminants. In each reconstructed category, one or two NN discriminants are used to distinguish between two or three signals (ggF, VBF, VH or ttH production) or between signal and background (ZZ^* or tXX production) processes. The input variables employed for the MLP and the two rNN are shown in Table 7.15 together with the targeted processes. The definition of the input variables is given in Section 5.3.

Depending on the reconstructed event category and the number of targeted signal and background processes the neural network processing results in two or three output quantities, corresponding to probabilities for the event to originate from the given process. For example, in the $0j$ category two probabilities are given, $\mathrm{NN_{ggF}}$ for the ggF production mode and $\mathrm{NN_{VBF}}$ for the VBF production. Since these values are linearly correlated to each other, only one of them ($\mathrm{NN_{ggF}}$) is used as discriminant. In case of three output values, two of them are independent. In this case, sub-categories of events are defined by the threshold on one of the two independent NN discriminants and in each sub-category only one of the two output discriminants is used. The final NN discriminants used in the respective reconstructed event categories are also shown in Table 7.15, together with the definition of the related sub-categories. The expected and observed distributions of the final NN discriminants are shown in Figures 7.21,7.22,7.23 and 7.24. Figure 7.24 also shows the expected and observed yields in the remaining reconstructed categories in which no NN discriminant is used, including the side-band regions. All distributions are in good agreement with the data.

Table 7.15: The input variables for the multilayer perceptron (MLP) and lepton and jet recurrent neural networks (lepton rNN, jet rNN) chained into the final neural network discriminant (NN) for each reconstructed event category. The corresponding targeted processes, the final NN discriminants and the definition of the NN sub-categories are also shown. The transverse momentum (pseudorapidity) of each of the four leptons and each of the up to three jets is denoted as p_T^{ℓ} (η_ℓ) and p_T^{j} (η_j) [201]. In addition, the final discriminant used in the respective reconstructed event category as well as the definition of the sub-category is shown.

Reconstructed event category	MLP	Lepton rNN	Jet rNN	Processes	Final NN Discriminant	Sub-category
$0j$-$p_T^{4\ell}$-Low	$p_T^{4\ell}, \mathcal{KD}(ZZ^*), m_{12}, m_{34}, \cos\theta^*, \cos\theta_1, \phi_{ZZ}$	p_T^{ℓ}, η_ℓ	–	ggF,ZZ^*	$\mathrm{NN_{ggF}}$	–
$1j$-$p_T^{4\ell}$-Low	$p_T^{4\ell}, p_T^{j}, \eta_j, \Delta R(j,4\ell), \mathcal{KD}(ZZ^*)$	p_T^{ℓ}, η_ℓ	–	ggF,VBF,ZZ^*	$\mathrm{NN_{VBF}}$ $\mathrm{NN_{ZZ}}$	$\mathrm{NN_{ZZ}} < 0.25$ $\mathrm{NN_{ZZ}} > 0.25$
$1j$-$p_T^{4\ell}$-Med	$p_T^{4\ell}, p_T^{j}, \eta_j, \Delta R(j,4\ell), E_T^{miss}, \mathcal{KD}(ZZ^*), \eta_{4\ell}$	p_T^{ℓ}, η_ℓ	–	ggF,VBF,ZZ^*	$\mathrm{NN_{VBF}}$ $\mathrm{NN_{ZZ}}$	$\mathrm{NN_{ZZ}} < 0.25$ $\mathrm{NN_{ZZ}} > 0.25$
$1j$-$p_T^{4\ell}$-High	$p_T^{4\ell}, p_T^{j}, \eta_j, \Delta R(j,4\ell), E_T^{miss}, \eta_{4\ell}$	p_T^{ℓ}	–	ggF,VBF	$\mathrm{NN_{VBF}}$	–
$2j$	$m_{jj}, p_T^{4\ell jj}$	p_T^{ℓ}, η_ℓ	p_T^{j}, η_j	ggF,VBF,VH	$\mathrm{NN_{VBF}}$ $\mathrm{NN_{VH}}$	$\mathrm{NN_{VH}} < 0.2$ $\mathrm{NN_{VH}} > 0.2$
$2j$-BSM-Like	$\eta^*_{4\ell}, p_T^{4\ell jj}$	p_T^{ℓ}, η_ℓ	p_T^{j}, η_j	ggF,VBF	$\mathrm{NN_{VBF}}$	–
VH-Lep-enriched	$N_j, N_b^{70\%}\ E_T^{miss}, H_T$	p_T^{ℓ}	–	VH,ttH	$\mathrm{NN_{ttH}}$	–
ttH-Had-enriched	$p_T^{4\ell}, m_{jj}, \Delta R_{jZ}^{min}, N_b^{70\%}$	p_T^{ℓ}, η_ℓ	p_T^{j}, η_j	ggF,ttH, tXX	$\mathrm{NN_{ttH}}$ $\mathrm{NN_{tXX}}$	$\mathrm{NN_{tXX}} < 0.4$ $\mathrm{NN_{tXX}} > 0.4$

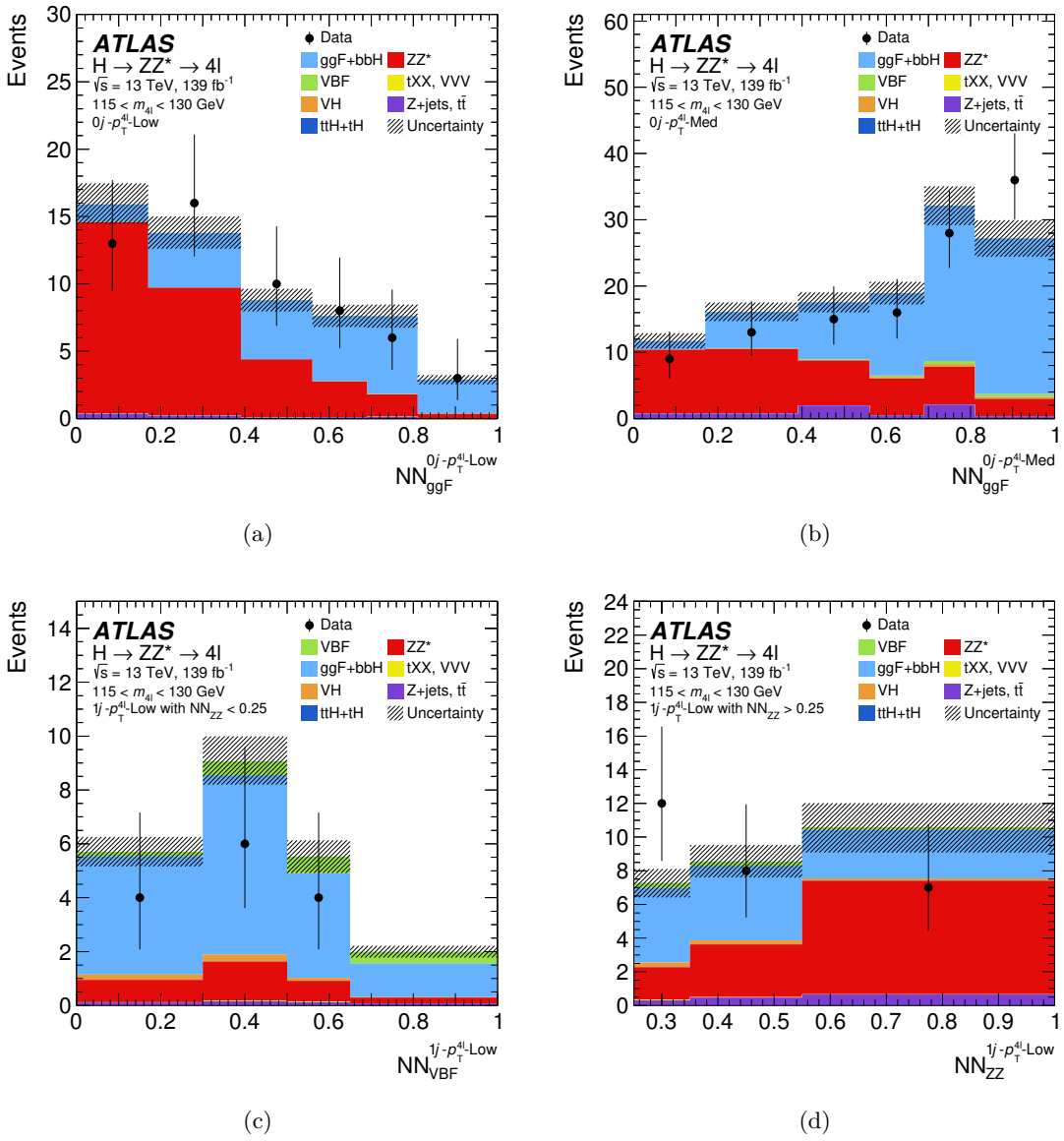


Figure 7.21: The observed and expected neural network (NN) output distributions for an integrated luminosity of 139 fb^{-1} and at $\sqrt{s} = 13$ TeV in the different reconstructed event categories, (a) $\mathrm{NN_{ggF}}$ in $0j$-$p_T^{4\ell}$-Low, (b) $\mathrm{NN_{ggF}}$ in $0j$-$p_T^{4\ell}$-Med, (c) $\mathrm{NN_{VBF}}$ in $1j$-$p_T^{4\ell}$-Low with $\mathrm{NN_{ZZ}}$ <0.25 and (d) $\mathrm{NN_{ZZ}}$ in $1j$-$p_T^{4\ell}$-Low with $\mathrm{NN_{ZZ}}$ >0.25. The boundaries of the bins are chosen to maximise the significance of the targeted signal in each category. The hatched band shows the combined statistical and theoretical uncertainty [201].

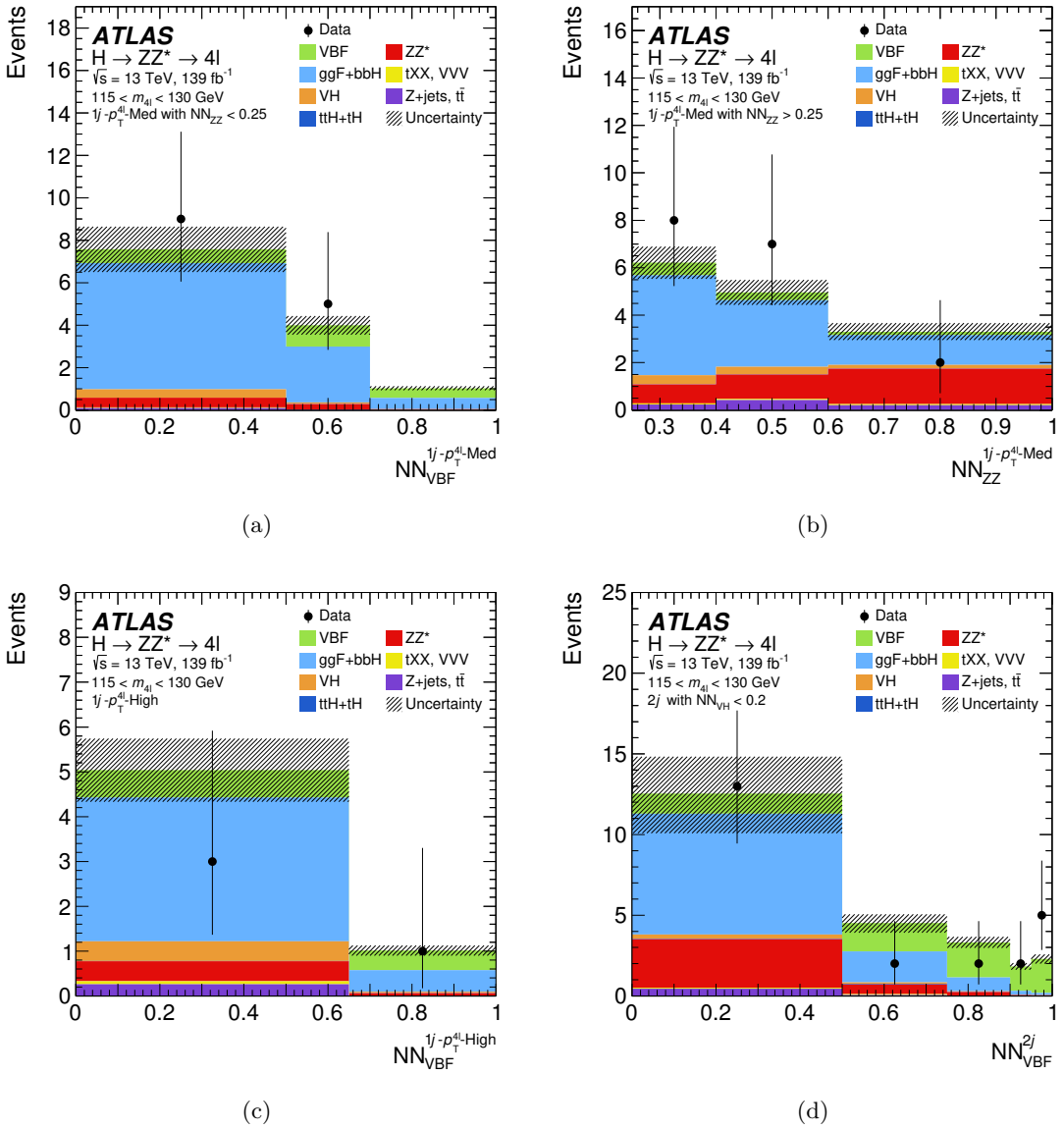


Figure 7.22: The observed and expected neural network (NN) output distributions for an integrated luminosity of 139 fb^{-1} and at $\sqrt{s} = 13$ TeV in the different reconstructed event categories, (a) $\mathrm{NN_{VBF}}$ in $1j$-$p_T^{4\ell}$-Med with $\mathrm{NN_{ZZ}}$ <0.25, (b) $\mathrm{NN_{ZZ}}$ in $1j$-$p_T^{4\ell}$-Med with $\mathrm{NN_{ZZ}}$ >0.25, (c) $\mathrm{NN_{VBF}}$ in $1j$-$p_T^{4\ell}$-High and (d) $\mathrm{NN_{VBF}}$ in $2j$ with $\mathrm{NN_{VH}}$ <0.2. The boundaries of the bins are chosen to maximise the significance of the targeted signal in each category. The hatched band shows the combined statistical and theoretical uncertainty [201].

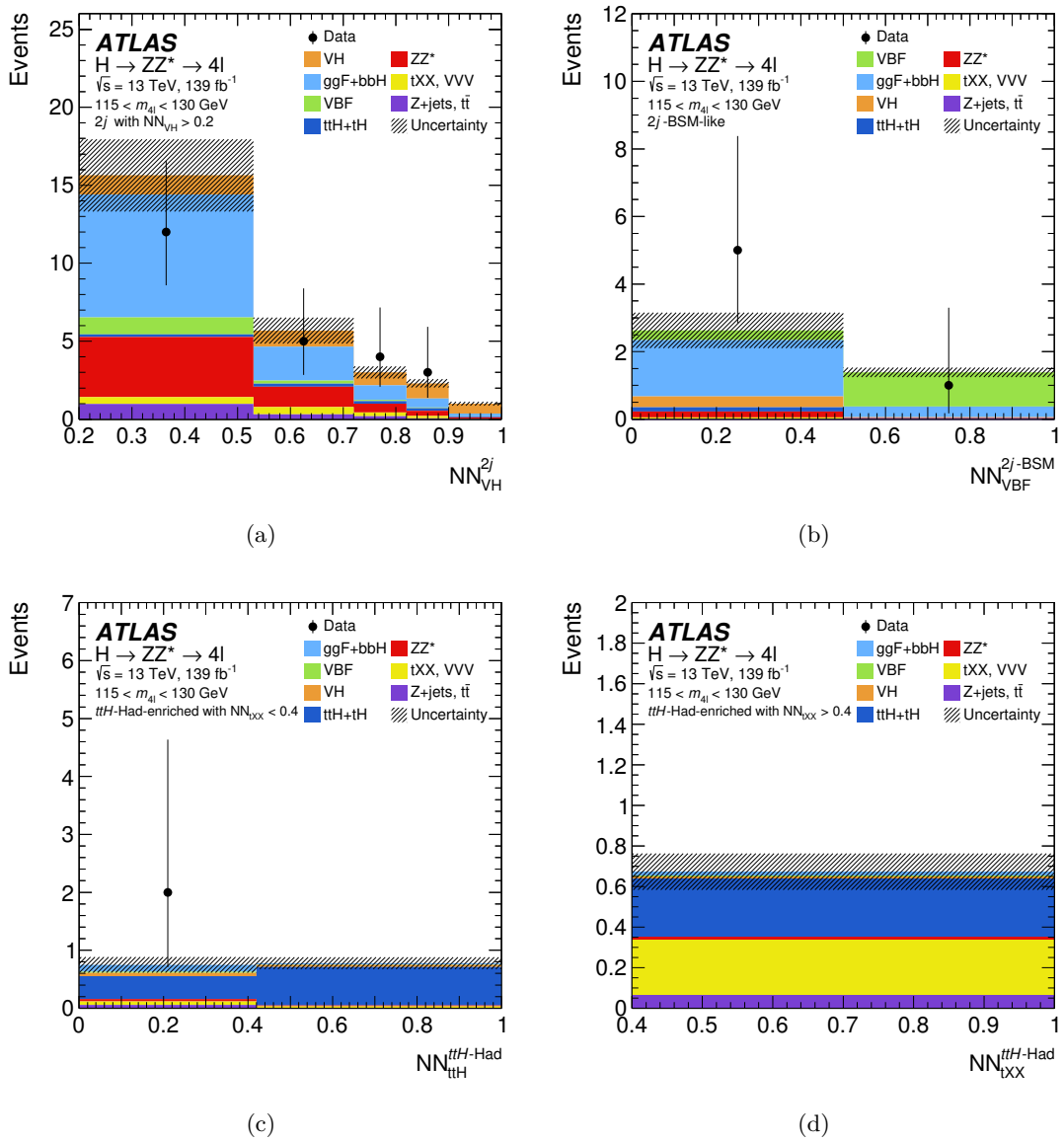


Figure 7.23: The observed and expected neural network (NN) output distributions for an integrated luminosity of 139 fb^{-1} and at $\sqrt{s} = 13$ TeV in the different reconstructed event categories, (a) $\mathrm{NN_{VH}}$ in $2j$ with $\mathrm{NN_{VH}}$ >0.2 (b) $\mathrm{NN_{VBF}}$ in $2j$-BSM-Like, (c) $\mathrm{NN_{ttH}}$ in ttH-enriched with $\mathrm{NN_{tXX}}$ <0.4 and (d) $\mathrm{NN_{tXX}}$ in ttH-enriched with $\mathrm{NN_{tXX}}$ >0.4. The boundaries of the bins are chosen to maximise the significance of the targeted signal in each category. The hatched band shows the combined statistical and theoretical uncertainty [201].

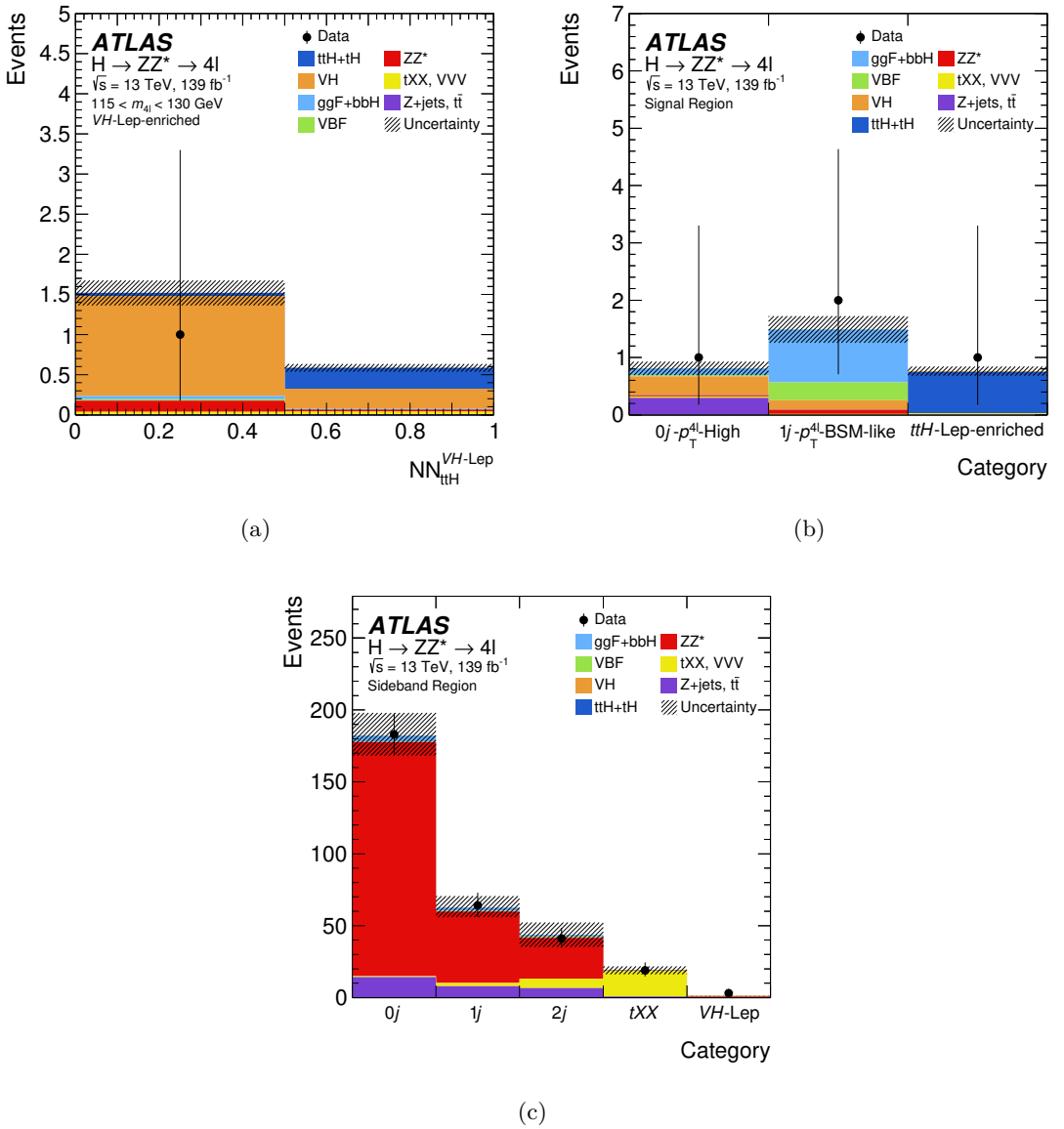


Figure 7.24: The observed and expected neural network (NN) output distribution for an integrated luminosity of 139 fb^{-1} and at $\sqrt{s} = 13$ TeV in the (a) VH-Lep-enriched reconstructed event category $\mathrm{NN}_{\mathrm{ttH}}$. The expected and observed number of events in the reconstructed event categories and in the side-band regions are shown in (b) and (c), respectively. The boundaries of the bins are chosen to maximise the significance of the targeted signal in each category. The hatched band shows the combined statistical and theoretical uncertainty [201].

The constraints on the production cross section measurement in exclusive regions of the phase space as defined in Reduced-Stage-1.1 are obtained from a binned maximum likelihood fits to the data using the profiled likelihood ratio method as described in Section 5.5. The normalisation of the irreducible ZZ^* and tXX background are included as additional nuisance parameters in the fit. The respective likelihood function is given as

$$\mathcal{L}(\boldsymbol{\sigma},\boldsymbol{\theta}) = \prod_{j}^{N_{\text{Categories}}} \prod_{i}^{N_{\text{Bins}}} P\left(n_{i,j} \mid L\cdot\boldsymbol{\sigma}\cdot\mathcal{B}\cdot\mathcal{A}_{i,j}(\boldsymbol{\theta}) + b_{i,j}(\boldsymbol{\theta})\right) \times \prod_{m}^{N_{\text{Nuisance}}} C_m(\boldsymbol{\theta}). \quad (7.19)$$

The constraints on $\boldsymbol{\sigma}\cdot\mathcal{B}$ are again obtained using a test statistic based on the ratio of profiled likelihoods as given in Equation 5.10.

The expected and the observed cross sections $\sigma\cdot\mathcal{B}$, as well as their ratio to the SM prediction are shown in Table 7.16 for each particle level bin of the Reduced-Stage-1.1 scheme. The results are summarised in Figure 7.25. The uncertainties on the SM prediction are not included in the uncertainties on the ratio $(\sigma\cdot\mathcal{B})/(\sigma\cdot\mathcal{B})_{\text{SM}}$. The results agree with the SM prediction within two standard deviations.

Table 7.16: The expected SM cross section $(\sigma\cdot\mathcal{B})_{\text{SM}}$, the observed cross section $(\sigma\cdot\mathcal{B})$ and their ratio $(\sigma\cdot\mathcal{B})/(\sigma\cdot\mathcal{B})_{\text{SM}}$ measured in the Reduced-Stage-1.1 production bins with the $H\to ZZ^*$ decay for an integrated luminosity of 139 fb^{-1} at $\sqrt{s}=13$ TeV. The uncertainties on the observed values are given as (stat.)+(syst.). The impact of the theory uncertainties is smaller than the rounding of the total uncertainty [201].

Production bin	Cross section $(\sigma\cdot\mathcal{B})$ [pb]		$(\sigma\cdot\mathcal{B})/(\sigma\cdot\mathcal{B})_{\text{SM}}$
	SM expected	Observed	Observed
gg2H-0j-p_T^H-Low	0.176 ± 0.025	$0.17 \pm 0.05 \pm 0.02$	$0.96 \pm 0.30 \pm 0.09$
gg2H-0j-p_T^H-High	0.55 ± 0.04	$0.63 \pm 0.09 \pm 0.06$	$1.15 \pm 0.17 \pm 0.11$
gg2H-1j-p_T^H-Low	0.172 ± 0.025	$0.05 \pm 0.07\ ^{+0.04}_{-0.06}$	$0.30 \pm 0.40\ ^{+0.2}_{-0.3}$
gg2H-1j-p_T^H-Med	0.119 ± 0.018	$0.17 \pm 0.05\ ^{+0.02}_{-0.01}$	$1.4 \pm 0.4 \pm 0.1$
gg2H-1j-p_T^H-High	0.020 ± 0.004	$0.009^{+0.016}_{-0.011} \pm 0.002$	$0.5^{+0.8}_{-0.6} \pm 0.1$
gg2H-2j	0.127 ± 0.027	$0.04 \pm 0.07 \pm 0.04$	$0.3 \pm 0.5 \pm 0.3$
gg2H-p_T^H-High	0.015 ± 0.004	$0.038^{+0.021\ +0.003}_{-0.016\ -0.002}$	$2.5^{+1.3\ +0.2}_{-1.0\ -0.1}$
qq2Hqq-VH-Like	$0.0138^{+0.0004}_{-0.0006}$	$0.021^{+0.037\ +0.009}_{-0.029\ -0.006}$	$1.5^{+2.7\ +0.6}_{-2.1\ -0.4}$
qq2Hqq-VBF	$0.1076^{+0.0024}_{-0.0035}$	$0.15 \pm 0.05\ ^{+0.02}_{-0.01}$	$1.4 \pm 0.5\ ^{+0.2}_{-0.1}$
qq2Hqq-BSM	0.00420 ± 0.00018	$0.0005^{+0.0079}_{-0.0047} \pm 0.008$	$0.1^{+1.9}_{-1.1} \pm 0.2$
qq/gg2HLep	0.0164 ± 0.0004	$0.022^{+0.028\ +0.003}_{-0.018\ -0.001}$	$1.3^{+1.7\ +0.2}_{-1.1\ -0.1}$
ttH	$0.0154^{+0.0010}_{-0.0013}$	$0.025^{+0.026\ +0.005}_{-0.017\ -0.003}$	$1.6^{+1.7\ +0.3}_{-1.1\ -0.2}$

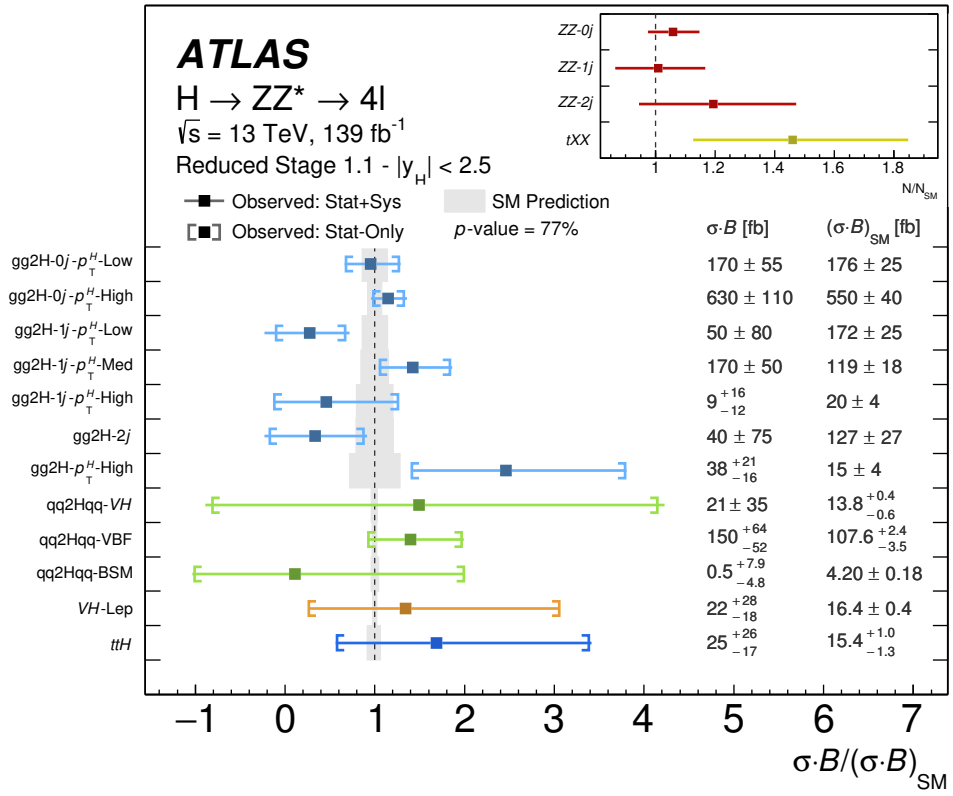


Figure 7.25: The observed and expected values of the cross sections $\sigma \cdot \mathcal{B}$ Reduced-Stage-1.1 production bins, normalised to the SM prediction $(\sigma \cdot \mathcal{B})_{\text{SM}}$ with the $H \to ZZ^*$ decays for an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV. The fitted normalisation factors for the ZZ^* and tXX background are shown in the inserts. Different colours indicate different Higgs boson production modes or background sources. The grey vertical band represents the theory uncertainty on the signal prediction. [201].

The presented results of the cross section measurements in production bins of the Reduced-Stage-1.1 scheme are interpreted in the framework of the SM Effective Field Theory (SMEFT) as described in the following.

7.3.2 BSM Signal Modelling

The observables in each particle level production bins of the Reduced-Stage-1.1 scheme are parametrised in terms of BSM parameters (SMEFT Wilson coefficients) within an effective field theory. The parametrisation is determined separately for the production cross sections, the branching ratio and the particle level acceptance of the $H \to ZZ^* \to 4\ell$ Higgs boson candidates. Each of these parametrisations is described in more detail in the following.

7.3.2.1 Cross Section Parametrisation

First, the production cross section can be parametrised as a function of BSM coupling parameters by employing the linear correlation of a production cross section to the squared matrix element of the corresponding process. The matrix element including new physics via Wilson coefficients C_i is given as

$$\mathcal{M}_{\text{Mix}} = \mathcal{M}_{\text{SM}} + \sum_i \frac{C_i}{\Lambda^2} \cdot \mathcal{M}_i, \tag{7.20}$$

where Λ corresponds to the energy scale at which new physics is assumed to appear. The squared matrix element is then

$$|\mathcal{M}_{\text{Mix}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \sum_i \frac{C_i}{\Lambda^2} \cdot 2\mathfrak{Re}\left(\mathcal{M}_{\text{SM}}^* \mathcal{M}_i\right) + \sum_{ij} \frac{C_i C_j}{\Lambda^4} \cdot 2\mathfrak{Re}\left(\mathcal{M}_i^* \mathcal{M}_j\right). \tag{7.21}$$

The first term corresponds to the cross section of the given SM process (SM term), the second one to the interference between the SM and the introduced BSM interactions (INT term) which is suppressed by a factor of $1/\Lambda^2$ and the last one describes the pure BSM contributions (BSM term) suppressed by a factor of $1/\Lambda^4$. The production cross section can, therefore, be separated into the corresponding three terms,

$$\sigma = \sigma_{\text{SM}} + \sigma_{\text{INT}} + \sigma_{\text{BSM}}. \tag{7.22}$$

The dependence of the Higgs boson production cross section $\sigma^p(\boldsymbol{c})$ in a given particle level production bin p on a set of Wilson coefficient $\boldsymbol{c} = \boldsymbol{C}/\Lambda^2$ can then be expressed as

$$\frac{\sigma^p(\boldsymbol{c})}{\sigma^p_{\text{SM}}} = 1 + \sum_i A_i^p c_i + \sum_{ij} B_{ij}^p c_i c_j, \tag{7.23}$$

where A_i^p $\left(B_{ij}^p\right)$ are the pre-factors of the linear interference (quadratic) terms that can be obtained from simulation.

The considered BSM coupling parameters c_i are summarised in Table 2.7. The pre-factors A_i^p and B_{ij}^p are computed using the SMEFTsim_A_U35_MwScheme_UFO_v2.1 UFO model [32, 169] within the MadGraph5_aMC@NLO event generator. The currently available implementation of the SMEFT model only provides the computation at LO in QCD and SM EW processes with additional corrections introducing the LO effective couplings of the Higgs boson to gluons and photons. It is assumed that the higher-order corrections are the same for both the SM and the BSM LO predictions [204], such that the parametrisation in Equation 7.23 can be used as a relative correction to the SM prediction calculated at NLO or NNLO,

$$\sigma^p(\boldsymbol{c}) = \sigma^{p,(\text{N})\text{NLO}}_{\text{SM}} \left(1 + \sum_i A_i^p c_i + \sum_{ij} B_{ij}^p c_i c_j\right). \tag{7.24}$$

The pre-factors A_i^p and B_{ij}^p are computed with MadGraph5_aMC@NLO generator by calculating separately the SM, the interference and the BSM part of the total production cross section for processes involving BSM physics. The input syntax used in the generator to obtain these fractions (SM term, INT term and BSM term) is shown in Table 7.17.

For example, in the case of two BSM coupling parameters entering into the interaction vertex, six Monte Carlo samples are required to calculate the parametrisation: one SM sample; two BSM samples which include only the interference terms, with one BSM coupling parameters c_i set to one and the other to zero; two BSM samples with the same configuration but including only the pure quadratic BSM terms $|\mathcal{M}_i|^2$; and one mixed BSM sample including the BSM terms with both BSM coupling parameters set to one. The configuration of the Monte Carlo samples used for the pre-factor calculation is summarised in Table 7.18, together with the resulting six production cross sections σ_{X}, where X indicates the type of the contribution (SM, INT or BSM). Each of these Monte Carlo samples is analysed on particle level to obtain the fraction f_p of events that fall into a given particle level production bin p. Accordingly, the production cross section in the particle level production bin corresponds to

$$\sigma_{\mathrm{X}}^p = f_p \cdot \sigma_{\mathrm{X}}. \tag{7.25}$$

The pre-factors corresponding to the linear and quadratic term can be directly obtained from the calculated production cross sections in the particle level production bin as follows

$$A_1^p = \frac{\sigma_{\mathrm{A_1}}^p}{\sigma_{\mathrm{SM}}^p}, \quad A_2^p = \frac{\sigma_{\mathrm{A_2}}^p}{\sigma_{\mathrm{SM}}^p}, \quad B_{11}^p = \frac{\sigma_{\mathrm{B_{11}}}^p}{\sigma_{\mathrm{SM}}^p} \quad \text{and} \quad B_{22}^p = \frac{\sigma_{\mathrm{B_{22}}}^p}{\sigma_{\mathrm{SM}}^p}. \tag{7.26}$$

Since the production cross sections $\sigma_{\mathrm{B_{11}}}^p$ and $\sigma_{\mathrm{B_{22}}}^p$ are included in $\sigma_{\mathrm{B_{12}}}^p$, they are subtracted to determine the pre-factor B_{12}^p related to the mixed term $c_1 c_2$

$$B_{12}^{\sigma_p} = \frac{\sigma_{\mathrm{B_{12}}}^p - \sigma_{\mathrm{B_{11}}}^p - \sigma_{\mathrm{B_{22}}}^p}{\sigma_{\mathrm{SM}}}. \tag{7.27}$$

Table 7.17: MadGraph5_aMC@NLO syntax to obtain the three different contributions (SM, INT and BSM) of the production cross section with non-vanishing values of BSM coupling parameters

	Syntax	Cross section	BSM dependence
SM term:	`NP^2==0`	σ_{SM}	–
INT term:	`NP^2==1`	σ_{INT}	linear
BSM term:	`NP^2==2`	σ_{BSM}	quadratic

Table 7.18: Configuration of the Monte Carlo samples used for the cross section parametrisation in case two arbitrary BSM coupling parameters c_1 and c_2

Sample Type	Value c_1	Value c_2	Syntax	Cross section
SM	0	0	`NP^2==0`	σ_{SM}
INT term c_1	1	0	`NP^2==1`	σ_{A_1}
BSM term c_1^2	1	0	`NP^2==2`	$\sigma_{\mathrm{B}_{11}}$
INT term c_2	0	1	`NP^2==1`	σ_{A_2}
BSM term c_1^2	0	1	`NP^2==2`	$\sigma_{\mathrm{B}_{22}}$
BSM term $c_1 c_2$	1	1	`NP^2==2`	$\sigma_{\mathrm{B}_{12}}$

To obtain the BSM parametrisation, several Monte Carlo samples are generated, with 100 000 events in each sample. Without the loss of generality, a massless version of the SMEFT model is used, i.e. the mass of the light quarks and the mass of e, μ and τ leptons is set to zero. A Higgs boson mass of $m_H = 125$ GeV is assumed. The generated events are showered with Pythia8 using the CKKW-L matching scheme to match matrix element and parton shower computations with different jet multiplicities [131]. The generator level cuts $p_T^j > 10$ GeV and $m_{jj} > 3$ GeV are applied. All processes are simulated within the four-flavour scheme (4FS) apart from the tHW production for which the five-flavour scheme (5FS) is used [167].

The MadGraph5_aMC@NLO input syntax for the simulation of a given Higgs boson production process, the required number of simulated samples and the detailed configuration of each sample is shown in Appendix B.1. Since the bbH production mode is treated as a part of the gg2H production bin, the ggF and bbH process are simulated in the same sample (gg2H production). For each of the processes, the SM cross section, the interference (linear) and pure BSM (quadratic) terms with the respective BSM coupling parameters set to one is calculated.

In case of the gg2H (WH-Lep) production, only one of the considered BSM coupling parameters enters into the interaction vertex, c_{HG} (c_{HW}). Therefore, the corresponding parametrisation is obtained with three Monte Carlo samples. The VBF+VH-Had and ZH-Lep production modes depend on three BSM coupling parameters (c_{HW}, c_{HB} and c_{HWB}), requiring ten Monte Carlo samples for the parametrisation. The ttH, $tHjb$ and tHW processes are affected by the c_{uH} and c_{HG} BSM coupling parameters corresponding to six Monte Carlo samples. Contributions from the $gg \to Z(\to \ell\ell)H$ process are assumed to have the same parametrisation like the $qq \to Z(\to \ell\ell)H$ process such that no

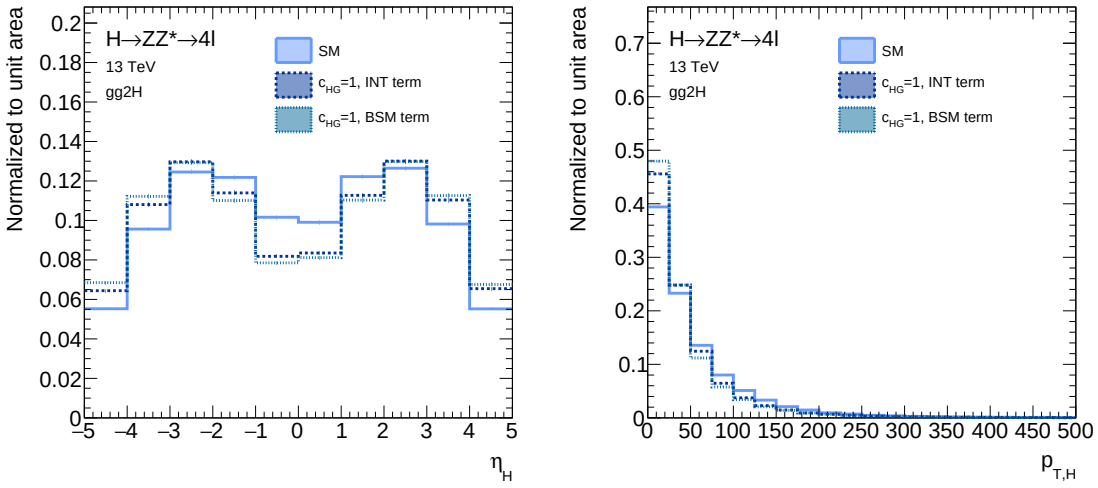
corresponding BSM Monte Carlo samples are produced. The impact of BSM coupling parameters on the background is assumed to be negligible.

The resulting total production cross sections for each generated SM and BSM process are summarised in Appendix B.2. Similar parametrisation procedure is also used for the CP-odd BSM coupling parameters, assuming that CP-even BSM parameters are not contributing. Since the CP-odd interference term does not contribute to the total cross section, a smaller number of samples is required in this case. The configuration of the corresponding generated Monte Carlo samples and the resulting total inclusive production cross sections can be found in Appendix C.1 and C.2, respectively.

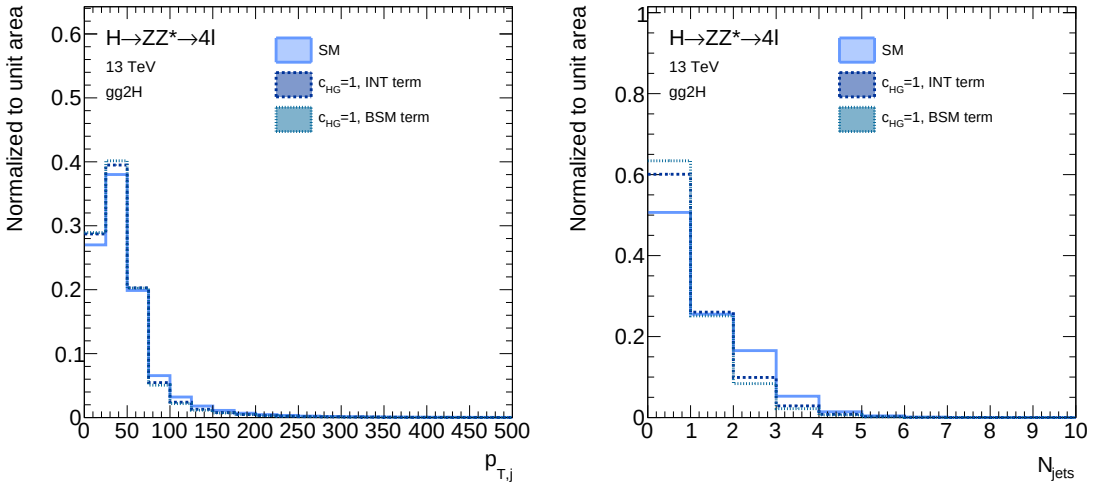
As mentioned above, all simulated samples are analysed on particle level to obtain the fraction of events falling into a particular particle level production bin. The gg2H production bins are defined based on the pseudorapidity η_H and the transverse momentum p_T^H of the Higgs boson, the p_T^j of the final state jets as well as the number of jets with $p_T^j > 30$ GeV in the final state N_j. The respective distributions from the SM, the interference and the BSM terms with $c_{HG} = 1$ are shown in Figure 7.26. In the SM case, the Higgs boson is produced more centrally in the detector and with a larger transverse momentum. The jets in the final state tend to have larger transverse momenta and a higher jet multiplicity as expected for the SM.

The relevant observables for the production bins of the VBF+VH-Had production (referred as qq2Hqq production in the following), are η_H, p_T^H and p_T^j. In addition, the invariant mass of the two leading jets m_{jj} is used for the classification. As an example, the kinematic distributions of these variables are shown for $c_{HW} = 1$ in Figure 7.27. The equivalent distributions with $c_{HB} = 1$ and $c_{HWB} = 1$ are shown in Appendix B.3. Compared to the SM, the η_H distribution from the interference term is wider while it is more central in the case of the BSM term. The Higgs boson has a larger transverse momentum in case of the interference and BSM terms compared to the SM term. The jet transverse momenta tend to have smaller values for the interference term, while the spectra is harder for the BSM term. The invariant mass of the two leading jets tends to have slightly higher values in case of the SM Higgs boson.

For the ZH-Lep, WH-Lep and ttH+tH production only the pseudorapidity of the Higgs boson and the jet transverse momenta is used for the definition of the respective particle level production bin. The corresponding distributions are shown in Figure 7.28 for the WH-Lep ($tHjb$) production for $c_{HW} = 1$ ($c_{uH} = 1$). The remaining distributions are shown in Appendix B.3. The interference term $c_{HW} = 1$ in the WH-Lep production gives similar distributions as the SM, while the Higgs boson is produced more central and the jet transverse momenta are larger in case of the BSM term. In case of $tHjb$ production the Higgs boson and the jets tend to have smaller values of η_H and p_T^j as compared to the SM term.



(a) (b)



(c) (d)

Figure 7.26: The expected distributions of (a) η_H, (b) p_T^H (c) p_T^j and (d) N_j with $p_T^j >$ 30 GeV in the gg2H production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HG} = 1$.

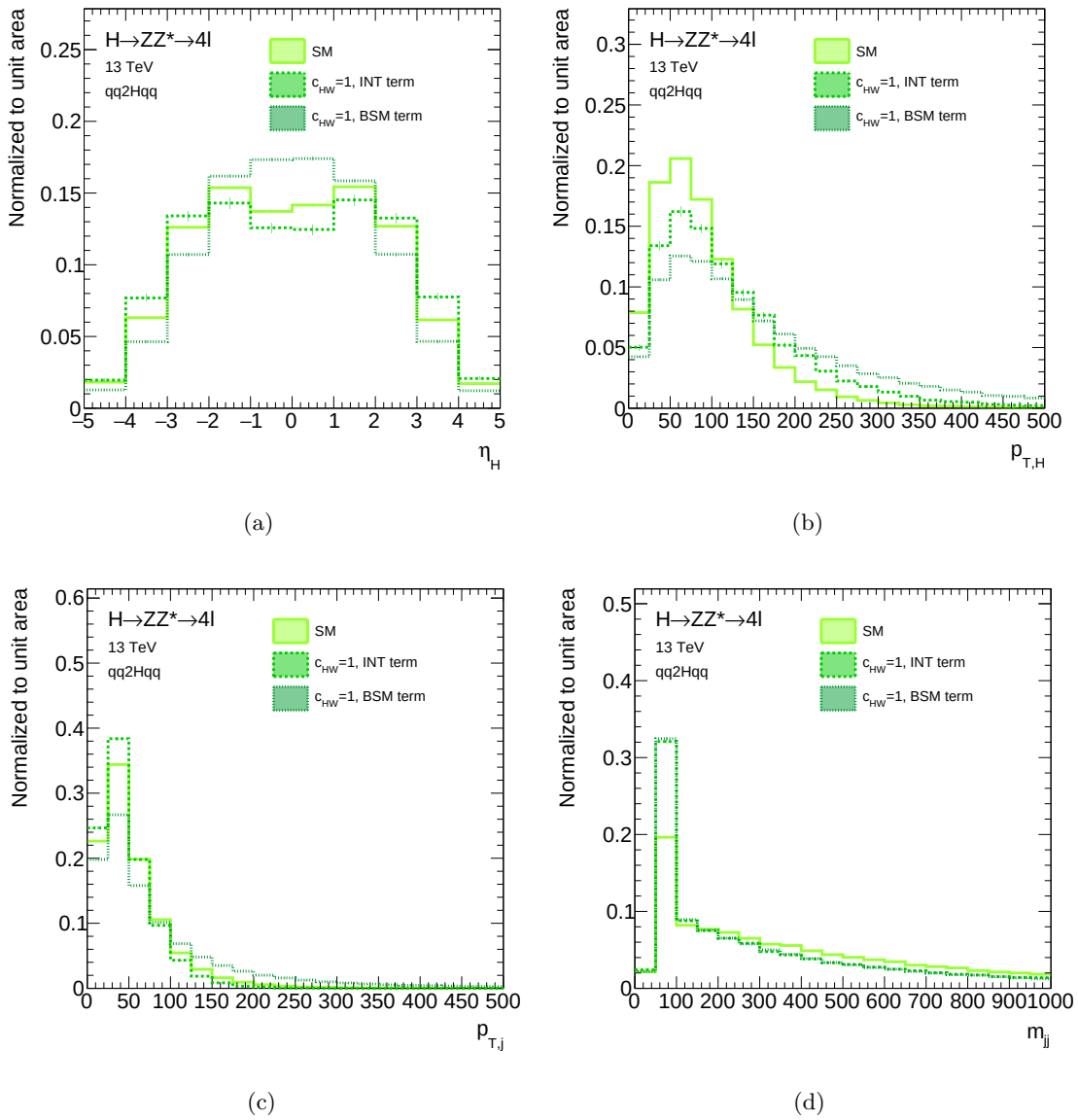
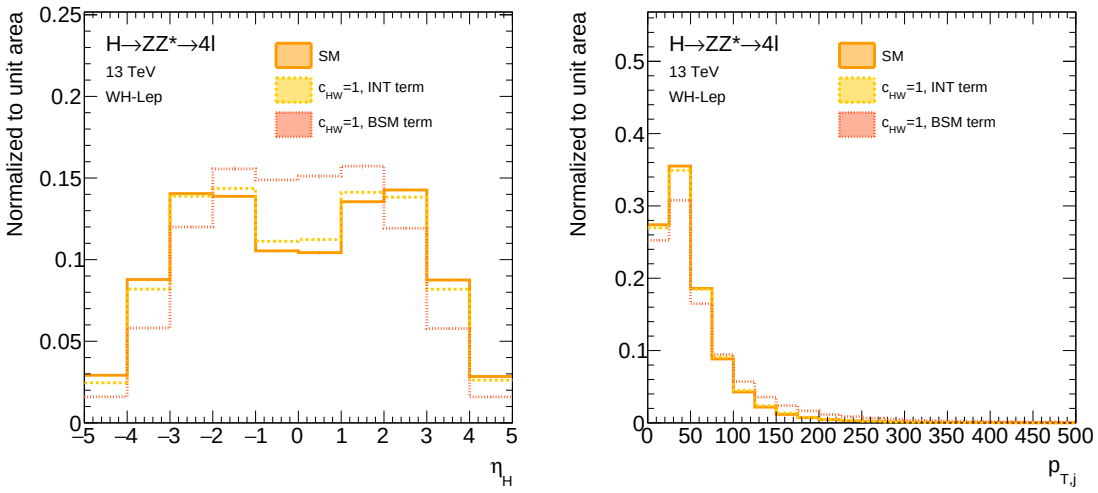
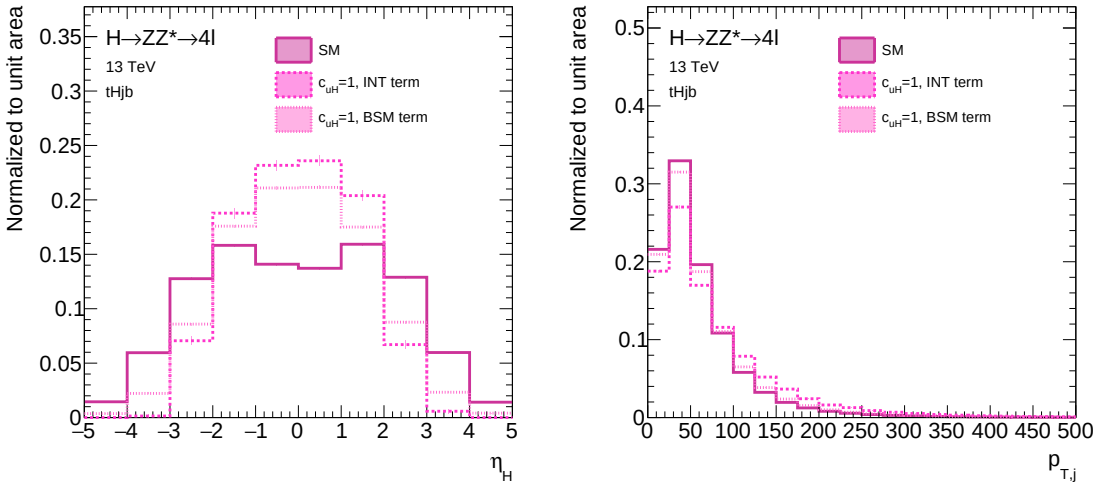


Figure 7.27: The expected distributions of (a) η_H, (b) p_T^H, (c) p_T^j and (d) m_{jj} in the qq2Hqq production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HW} = 1$.



(a) (b)



(c) (d)

Figure 7.28: The expected distributions of η_H and N_j shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HW} = 1$ in (a) and (b) WH-Lep and $c_{uH} = 1$ in (c) and (d) $tHjb$ production.

The fractions of events in each particle level production bins of the Reduced-Stage-1.1 scheme are summarised in Appendix B.4 and C.3 for all considered Higgs boson production modes with CP-even and CP-odd BSM coupling parameters, respectively. For the EFT parametrisation of the qq/gg2HLep (ttH) production bin, the calculated cross section in the production bin is summed up for ZH-Lep and WH-Lep (ttH, $tHjb$ and tHW) production. The resulting EFT parametrisation is summarised in Table 7.19. The EFT parametrisation for CP-odd BSM coupling parameters is given in Appendix C.4.

In order to validate the derived EFT parametrisation, dedicated validation Monte Carlo samples are generated for each of the production modes including all three terms in a single sample (SM, INT and BSM). The samples are produced for different values of BSM coupling parameters. The production cross section is calculated in each particle level production bin and is compared to the expected value from the EFT parametrisation. The validation results are shown in Figure 7.29. Within the statistical errors all calculated production cross sections agree with the predictions of the EFT parametrisation.

7.3.2.2 Branching Ratio Parametrisation

In addition to the cross section also the branching ratio of the $H \to ZZ^*$ decay is affected by the studied BSM coupling parameters. The parametrisation of the branching ratio is obtained from the parametrisation of partial and total decay width, in a similar way as for the cross sections. The partial and the total decay width can also be divided into the SM, interference and BSM terms

$$\Gamma^{4\ell/\text{tot}} = \Gamma^{4\ell/\text{tot}}_{\text{SM}} + \Gamma^{4\ell/\text{tot}}_{\text{INT}} + \Gamma^{4\ell/\text{tot}}_{\text{BSM}}, \tag{7.28}$$

where $\Gamma^{4\ell}$ is the partial width for the $H \to ZZ^* \to 4\ell$ decay and $\Gamma^{\text{tot}} = \sum_f \Gamma^f$ is the total Higgs boson decay width, defined as the sum of all partial decay widths from all decays. Therefore, the dependence of the branching ratio on the BSM coupling parameters can be expressed as

$$\mathcal{B}^{4\ell}(\boldsymbol{c}) = \frac{\Gamma^{4\ell}(\boldsymbol{c})}{\Gamma^{\text{tot}}(\boldsymbol{c})} = \frac{\Gamma^{4\ell}_{\text{SM}}}{\Gamma^{\text{tot}}_{\text{SM}}} \cdot \frac{1 + \sum_i A^{4\ell}_i c_i + \sum_{ij} B^{4\ell}_{ij} c_i c_j}{1 + \sum_f \left(\sum_i A^f_i c_i + \sum_{ij} B^f_{ij} c_i c_j \right)}. \tag{7.29}$$

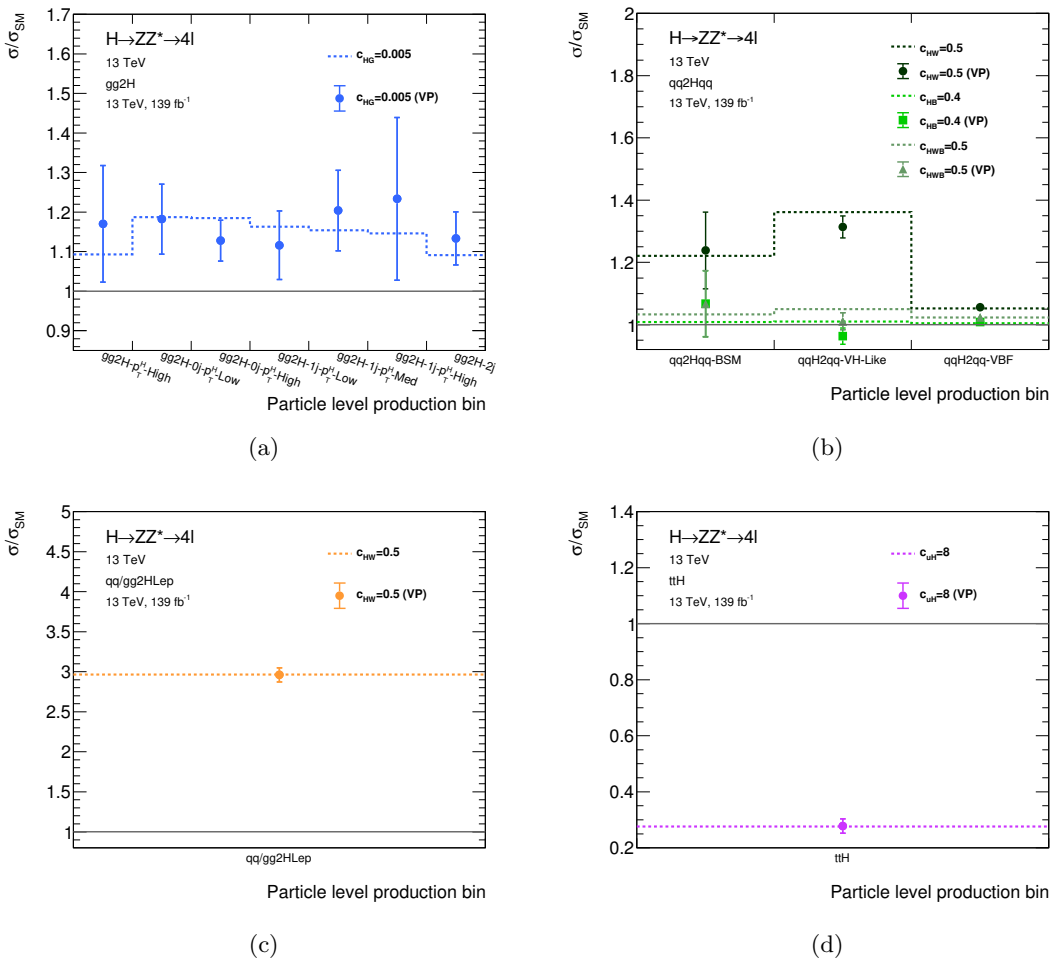


Figure 7.29: Comparison of the relative cross section $\sigma/\sigma_{\mathrm{SM}}$ per particle level production bin as predicted by the EFT parametrisation for different BSM parameters (dotted lines) and the production cross section of the validation points (VP) obtained from the Monte Carlo samples (points) for (a) gg2H, (b) qq2Hqq, (c) VH-Lep and (d) ttH production.

The pre-factors $A_i^{4\ell}, B_{ij}^{4\ell}, A_i^f$ and B_{ij}^f are computed at LO using MadGraph5_aMC@NLO samples with the Higgs boson decay alone. The branching ratio is expressed as a relative correction to the SM prediction calculated at NLO

$$\mathcal{B}^{4\ell}(\boldsymbol{c}) = \mathcal{B}_{\mathrm{SM}}^{4\ell,\mathrm{NLO}} \cdot \frac{1+\sum_i A_i^{4\ell} c_i + \sum_{ij} B_{ij}^{4\ell} c_i c_j}{1+\sum_f \left(\sum_i A_i^f c_i + \sum_{ij} B_{ij}^f c_i c_j\right)}, \tag{7.30}$$

assuming that the higher-order corrections to the branching ratio are the same for the SM and for the BSM samples.

Each Monte Carlo sample comprises 100 000 events generated with the massless version of the SMEFT model. A cut is applied on the Breit-Wigner distribution of the resonance, defining the Higgs boson to be on-shell if its invariant mass is within the mass window of $m_H \pm \Gamma^{\text{tot}} \cdot 15$. In addition, an angular separation between leptons of $\Delta R(\ell, \ell') > 0.05$ is required. For the generated Monte Carlo samples the obtained total Higgs boson decay width at LO, $\Gamma^{\text{tot,LO}}_{\text{SM}} = 4.995 \times 10^{-3}$ GeV, is used. Only the dominant decay modes are considered for the calculation.

Table 7.19: EFT parametrisation of the cross section ratio $\sigma/\sigma_{\text{SM}}$ for each particle level production bin of the Reduced-Stage-1.1 scheme and of the ratio of decay widths $\Gamma/\Gamma_{\text{SM}}$ in dependence on the CP-even BSM coupling parameters

Production bin	Cross section parametrisation, $\sigma/\sigma_{\text{SM}}$
gg2H-0j-p_T^H-Low	$1 + 35.80c_{HG} + 326.23c_{HG}^2$
gg2H-0j-p_T^H-High	$1 + 35.33c_{HG} + 319.05c_{HG}^2$
gg2H-1j-p_T^H-Low	$1 + 31.27c_{HG} + 264.44c_{HG}^2$
gg2H-1j-p_T^H-Med	$1 + 29.55c_{HG} + 236.09c_{HG}^2$
gg2H-1j-p_T^H-High	$1 + 28.00c_{HG} + 225.53c_{HG}^2$
gg2H-2j	$1 + 17.62c_{HG} + 118.92c_{HG}^2$
gg2H-p_T^H-High	$1 + 17.77c_{HG} + 162.65c_{HG}^2$
qq2Hqq-VH-Like	$1 + 0.593c_{HW} + 0.258c_{HW}^2 + 0.019c_{HB} + 0.014c_{HB}^2 + 0.088c_{HWB} + 0.023c_{HWB}^2 + 0.005c_{HW}c_{HB} + 0.034c_{HW}c_{HWB} + 0.005c_{HB}c_{HWB}$
qq2Hqq-BSM	$1 + 0.190c_{HW} + 0.506c_{HW}^2 - 0.002c_{HB} + 0.057c_{HB}^2 + 0.042c_{HWB} + 0.048c_{HWB}^2 + 0.012c_{HW}c_{HB} - 0.059c_{HW}c_{HWB} - 0.028c_{HB}c_{HWB}$
qq2Hqq-VBF	$1 + 0.059c_{HW} + 0.092c_{HW}^2 + 0.002c_{HB} + 0.027c_{HB}^2 + 0.037c_{HWB} + 0.018c_{HWB}^2 + 0.015c_{HW}c_{HB} - 0.016c_{HW}c_{HWB} - 0.024c_{HB}c_{HWB}$
qq/gg2HLep	$1 + 0.828c_{HW} + 0.321c_{HW}^2 + 0.035c_{HB} + 0.013c_{HB}^2 + 0.127c_{HWB} + 0.026c_{HWB}^2 - 0.218c_{HW}c_{HB} - 0.155c_{HW}c_{HWB} + 0.020c_{HB}c_{HWB}$
ttH/tH	$1 + 0.483c_{HG} + 0.590c_{HG}^2 - 0.108c_{uH} + 0.009c_{uH}^2 - 0.015c_{HG}c_{uH}$
Decay process	**Decay width parametrisation, $\Gamma/\Gamma_{\text{SM}}$**
$\Gamma^{4\ell}/\Gamma^{4\ell}_{\text{SM}}$	$1 - 0.199c_{HW} + 0.753c_{HW}^2 - 0.112c_{HB} + 2.665c_{HB}^2 + 0.181c_{HWB} + 0.760c_{HWB}^2 - 0.043c_{HW}c_{HB} - 1.288c_{HW}c_{HWB} - 1.397c_{HB}c_{HWB}$
$\Gamma^{\text{tot}}/\Gamma^{\text{tot}}_{\text{SM}}$	$1 - 0.054c_{HW} + 0.162c_{HW}^2 - 0.076c_{HB} + 1.209c_{HB}^2 + 0.062c_{HWB} + 0.357c_{HWB}^2 + 1.519c_{HG} + 14.922c_{HG}^2 + 0.518c_{HW}c_{HB} - 0.441c_{HW}c_{HWB} - 1.132c_{HB}c_{HWB}$

The MadGraph5_aMC@NLO syntax for all considered decays is shown in Appendix B.1. The total decay width depends on four BSM coupling parameters: c_{HG}, c_{HW}, c_{HB} and c_{HWB}. Since c_{HG} does not contribute to the decay vertex, twelve simulated samples are required. The HVV interaction vertex depends on three BSM coupling parameters (c_{HW}, c_{HB} and c_{HWB}) leading to ten simulated Monte Carlo samples. The values of the partial and total decay width for each of the simulated samples are summarised in Appendix B.2 and C.2 for CP-even and CP-odd BSM coupling parameters, respectively. The resulting EFT parametrisation for CP-even (CP-odd) BSM coupling parameters is shown in Table 7.19 (Appendix C.3).

7.3.2.3 Detector Acceptance Parametrisation

The selection criteria of the Higgs boson candidates in the $H \to ZZ^* \to 4\ell$ decay channel (see Section 4.3) introduce an additional dependence on the BSM coupling parameters. The largest impact of BSM contributions is expected on the invariant mass m_{34} of the subleading lepton pair as shown in Figure 7.30 for the gg2H production. The larger the absolute value of the BSM coupling parameter (c_{HW}, c_{HB} and c_{HWB}) the smaller the invariant mass of the subleading lepton pair. Therefore, the four-lepton selection criteria reject more events in case of BSM hypotheses as compared to the SM. In order to take this effect into account, the signal acceptance is parametrised in dependence on the Wilson coefficients c_{HW}, c_{HB} and c_{HWB} assuming vanishing CP-odd BSM coupling parameters values.

A common acceptance parametrisation is used, since the modifications introduced by the different production bins are shown to be very small. The parametrisation is derived using the dominant gg2H production mode, while the same parametrisation is expected also for other production modes.

The acceptance correction relative to the SM prediction is described by a three-dimensional Lorentzian function with free acceptance parameters α_0,α_1,α_2, β_i, δ_i, $\delta_{i,j}$ and $\delta_{i,j,k}$,

$$\frac{\mathcal{A}(\boldsymbol{c})}{\mathcal{A}_{\mathrm{SM}}} = \alpha_0 + (\alpha_1)^2 \cdot \left[\alpha_2 + \sum_i \delta_i \cdot (c_i + \beta_i)^2 + \sum_{\substack{ij \\ i \neq j}} \delta_{i,j} \cdot c_i c_j + \sum_{i \neq j \neq k} \delta_{i,j,k} \cdot c_i c_j c_k \right]^{-1} , \quad (7.31)$$

where indices i, j and k run over HW, HB and HWB. In addition, a quadratic modelling of the detector acceptance has been implemented. However, in the expected sensitivity range this approximation is not valid. In order to determine the values of the acceptance parameters in Equation 7.31 , a three-dimensional fit of the signal acceptance ratio $\mathcal{A}/\mathcal{A}_{\mathrm{SM}}$ is performed. For this purpose, 117 Monte Carlo samples with different values of BSM coupling values and including all contributing terms (SM, INT and BSM) are

generated for the Higgs boson production via gg2Hand the $H \to ZZ^* \to 4\ell$ decay. The simulated events are fed through the event selection of the inclusive analysis and the signal acceptance defined as,

$$\mathcal{A} = \frac{N_{\text{after } 4\ell \text{ selection}}}{N_{\text{total}}}, \tag{7.32}$$

is calculated for each of the samples. The ratio to the corresponding SM prediction $\mathcal{A}_{\text{SM}}$ is used as input for the fit. The configuration of the generated Monte Carlo samples, as well as the calculated ratio of acceptances is shown in Appendix B.5.

The fit result, i.e. the values of the acceptance parameters are summarised in Table 7.20. The EFT parametrisation of the signal acceptance is visualised by means of one-dimensional projections for c_{HW}, c_{HB} and c_{HWB} shown Figure 7.31. Each projections agrees well with a one-dimensional Lorentzian function. The signal acceptance parametrisation for the CP-odd BSM coupling parameters is derived in the same way, by assuming vanishing CP-even BSM coupling parameters values. The fit result is shown in Appendix C.5.

For the modelling of the detector acceptance with mixtures of CP-even and CP-odd BSM contributions a description with a six-dimensional Lorentzian function is required. Since for the three-dimensional case already a very large number of generated Monte Carlo samples is required and the number of free acceptance parameters in the fit is large, the detector acceptance is not modelled for both CP-even and CP-odd BSM contributions at the same time.

Table 7.20: Values of acceptance parameters obtained from the fit of the three-dimensional Lorentzian function (Equation 7.31)

Acceptance parameter	Fit result	Acceptance parameter	Fit result
α_0	0.153 ± 0.003	δ_{HW}	0.614 ± 0.027
α_1	0.874 ± 0.010	δ_{HB}	2.294 ± 0.033
α_2	0.881 ± 0.019	δ_{HWB}	0.703 ± 0.029
β_{HW}	-0.133 ± 0.012	$\delta_{HW,HWB}$	-1.210 ± 0.040
β_{HB}	0.005 ± 0.005	$\delta_{HB,HWB}$	-1.220 ± 0.060
β_{HWB}	0.120 ± 0.011	$\delta_{HW,HB}$	0.080 ± 0.070
		$\delta_{HW,HB,HWB}$	0.050 ± 0.060

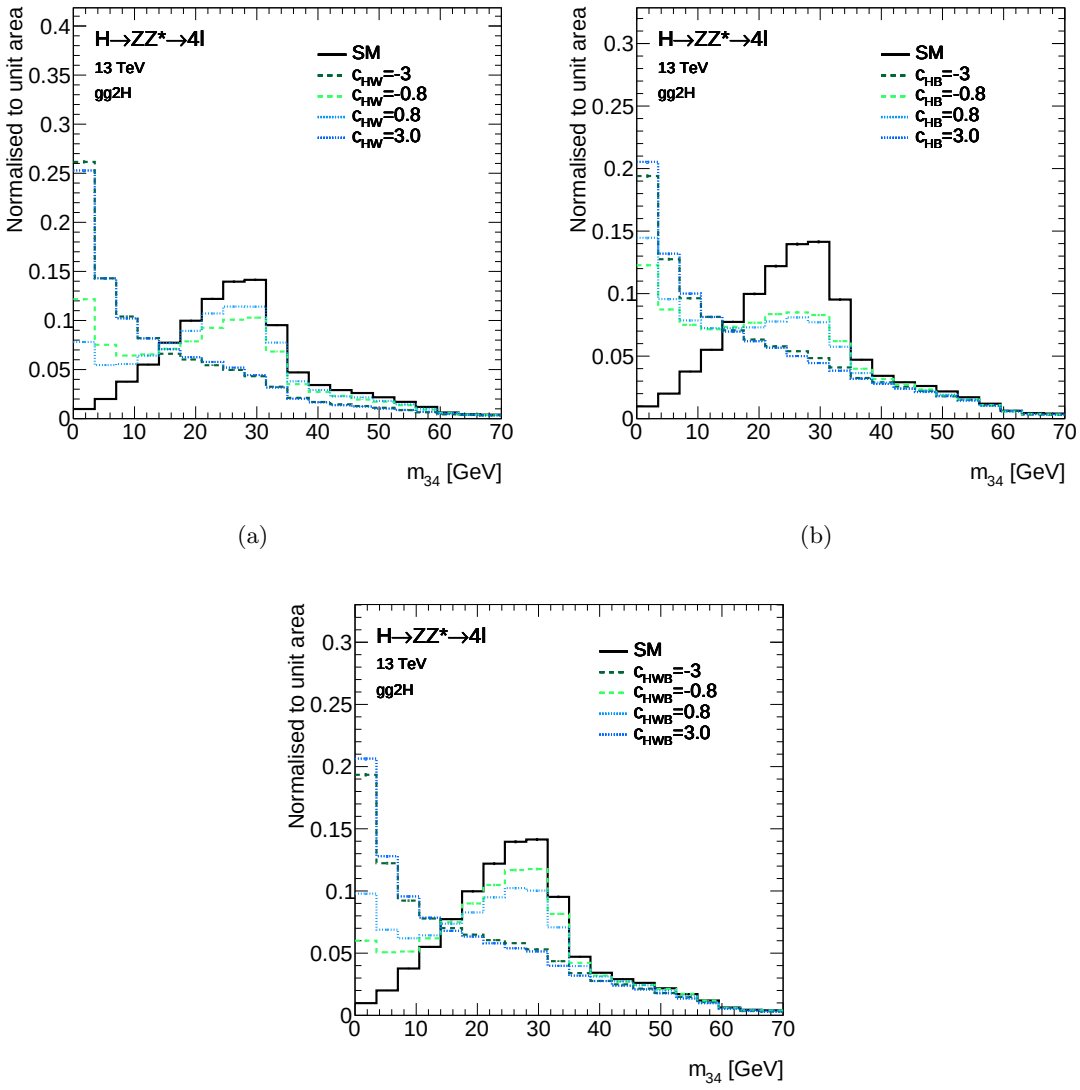


Figure 7.30: Particle level distribution of the invariant mass m_{34} of the subleading lepton pair for different values of BSM coupling parameters (a) c_{HW}, (b) c_{HB} and (c) c_{HWB}.

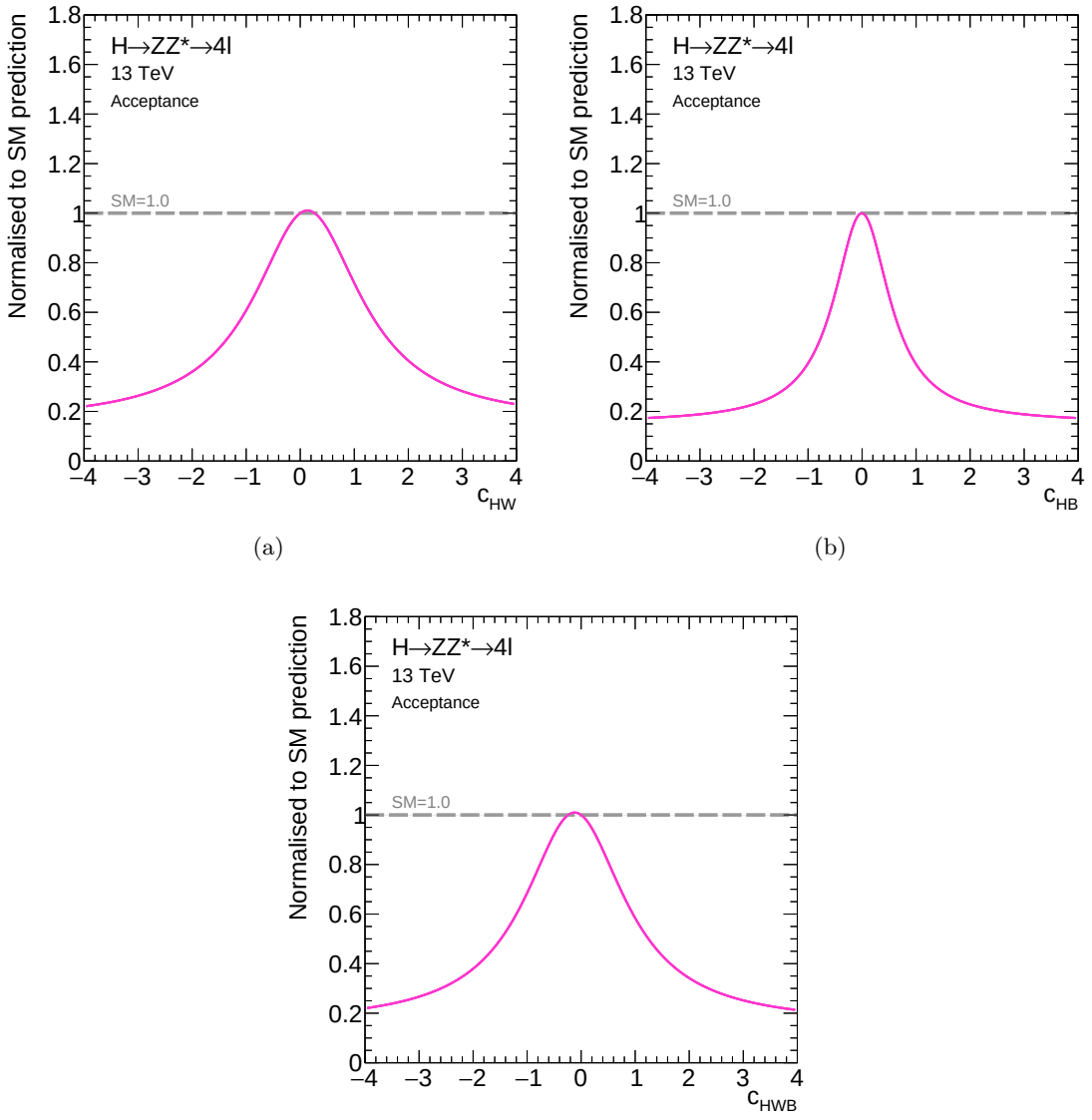


Figure 7.31: One-dimensional projections of the BSM parametrisation of the acceptance ratio $\mathcal{A}/\mathcal{A}_{SM}$ for the (a) c_{HW}, (b) c_{HB} and (c) c_{HWB} Wilson coefficients.

7.3.2.4 Full Signal Model

The expected number of events in each particle level production bin p of Reduced-Stage-1.1 scheme is parametrised as a product of the cross section, branching ratio and acceptance. The final signal parametrisation is, therefore, given as

$$\begin{aligned}
&\sigma^{p}(\boldsymbol{c}) \cdot \mathcal{B}^{4\ell}(\boldsymbol{c}) \cdot \mathcal{A}^{p}(\boldsymbol{c}) = \\
&\sigma^{p,(\mathrm{N})\mathrm{NLO}}_{\mathrm{SM}} \cdot \mathcal{B}^{4\ell,\mathrm{NLO}}_{\mathrm{SM}} \cdot \mathcal{A}^{p}_{\mathrm{SM}} \\
&\cdot \left(1 + \sum_i A^p_i c_i + \sum_{ij} B^p_{ij} c_i c_j\right) \\
&\cdot \left(\frac{1 + \sum_i A^{4\ell}_i c_i + \sum_{ij} B^{4\ell}_{ij} c_i c_j}{1 + \sum_f \left(\sum_i A^f_i c_i + \sum_{ij} B^f_{ij} c_i c_j\right)}\right) \\
&\cdot \left(\alpha_0 + (\alpha_1)^2 \cdot \left[\alpha_2 + \sum_i \delta_i \cdot (c_i + \beta_i)^2 + \sum_{\substack{ij \\ i \neq j}} \delta_{i,j} \cdot c_i c_j + \sum_{i \neq j \neq k} \delta_{i,j,k} \cdot c_i c_j c_k\right]^{-1}\right).
\end{aligned} \tag{7.33}$$

The expected event yields relative to the SM prediction are shown in Figure 7.34 for a given Wilson coefficient after setting all other coefficients to zero. Only particle level production bins with the highest sensitivity to the corresponding BSM coupling parameters are shown, in addition as an example the production bin with the lowest sensitivity (gg2H-p_T^H-High) is shown for the gg2H production. The BSM dependence of expected event yields in all remaining production bins is given in Appendix B.6. For comparison, also the predictions without the acceptance ($\sigma \cdot \mathcal{B}$), and without both acceptance and branching ratio corrections (σ) are shown. In addition, the grey bands indicate the prediction of the SM production cross section measurement at 1 σ level, closely related to the measurement sensitivity in the corresponding production bin.

The BSM coupling parameter c_{HG} contributes only to the effective Higgs boson coupling to gluons. Therefore, the largest sensitivity is expected in the gg2H production bins. In the most sensitive production bin (gg2H-0j-p_T^H-High) a linear EFT approximation is valid, while in production bins with ten times smaller sensitivity also the quadratic term contributes. In the expected sensitivity, range the impact of non-vanishing values of c_{HG} on the total Higgs boson decay width is small, and there is no impact on the partial decay width. The BSM coupling parameters c_{HW},c_{HWB} and c_{HB} contribute to the qq2Hqq production and to the $H \to ZZ^* \to 4\ell$ decay. The particle level production bin with the largest sensitivity to VBF or VH production is qq2Hqq-VBF. The parametrisation of the production cross section has a strong quadratic dependence in the expected sensitivity range, such that the contributions from the BSM term (suppressed by $1/\Lambda^4$) of the EFT para-

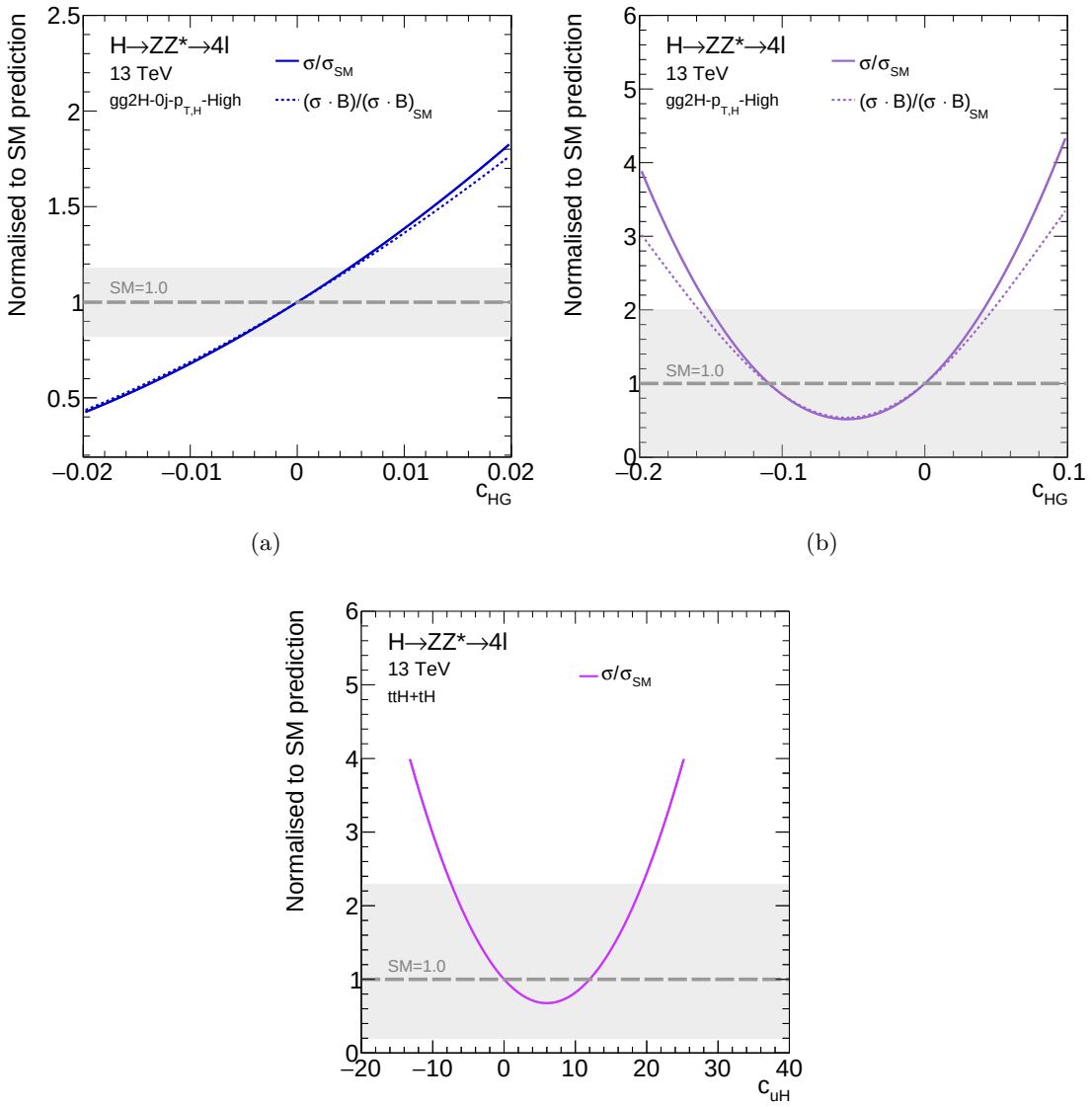
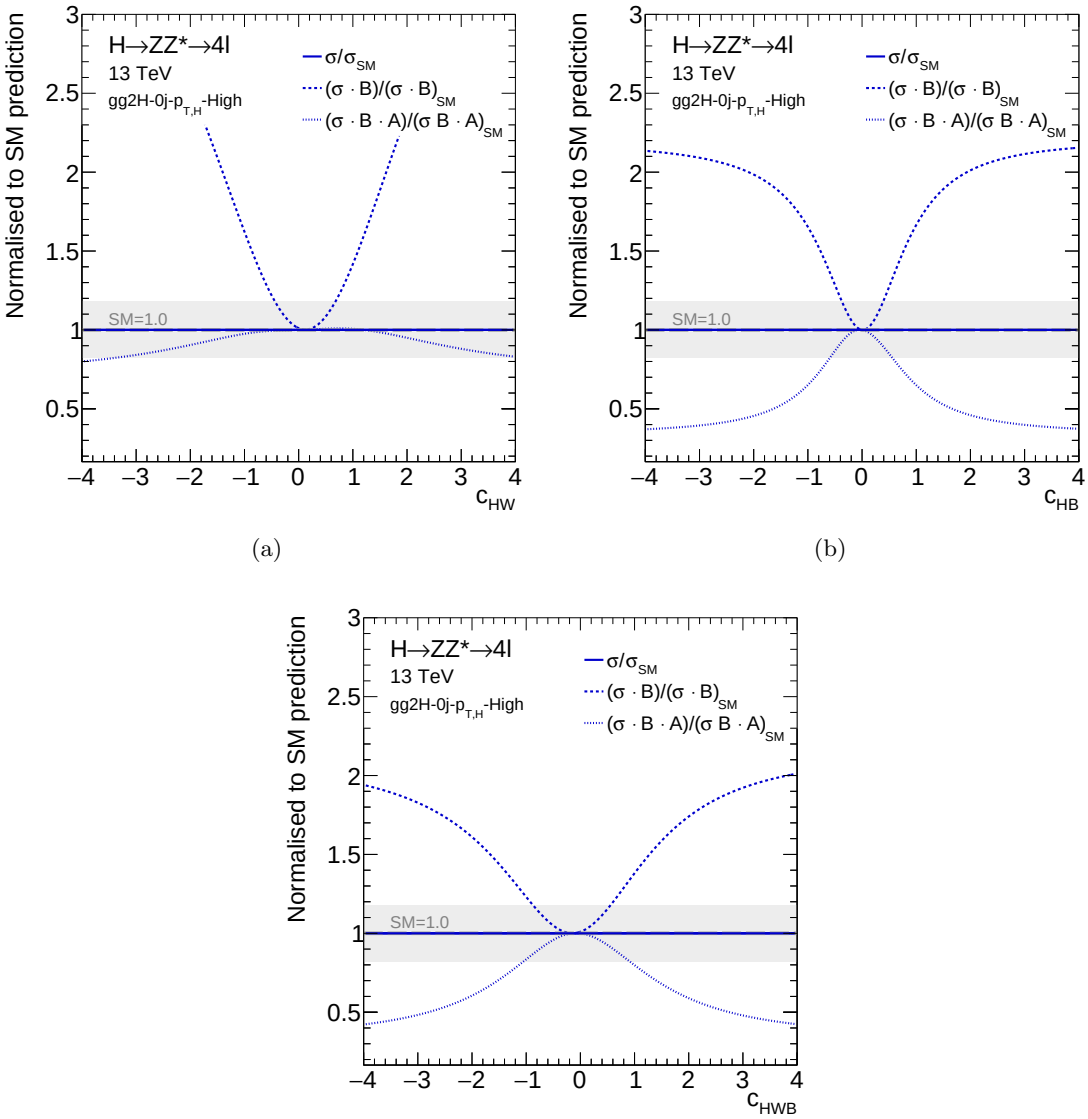


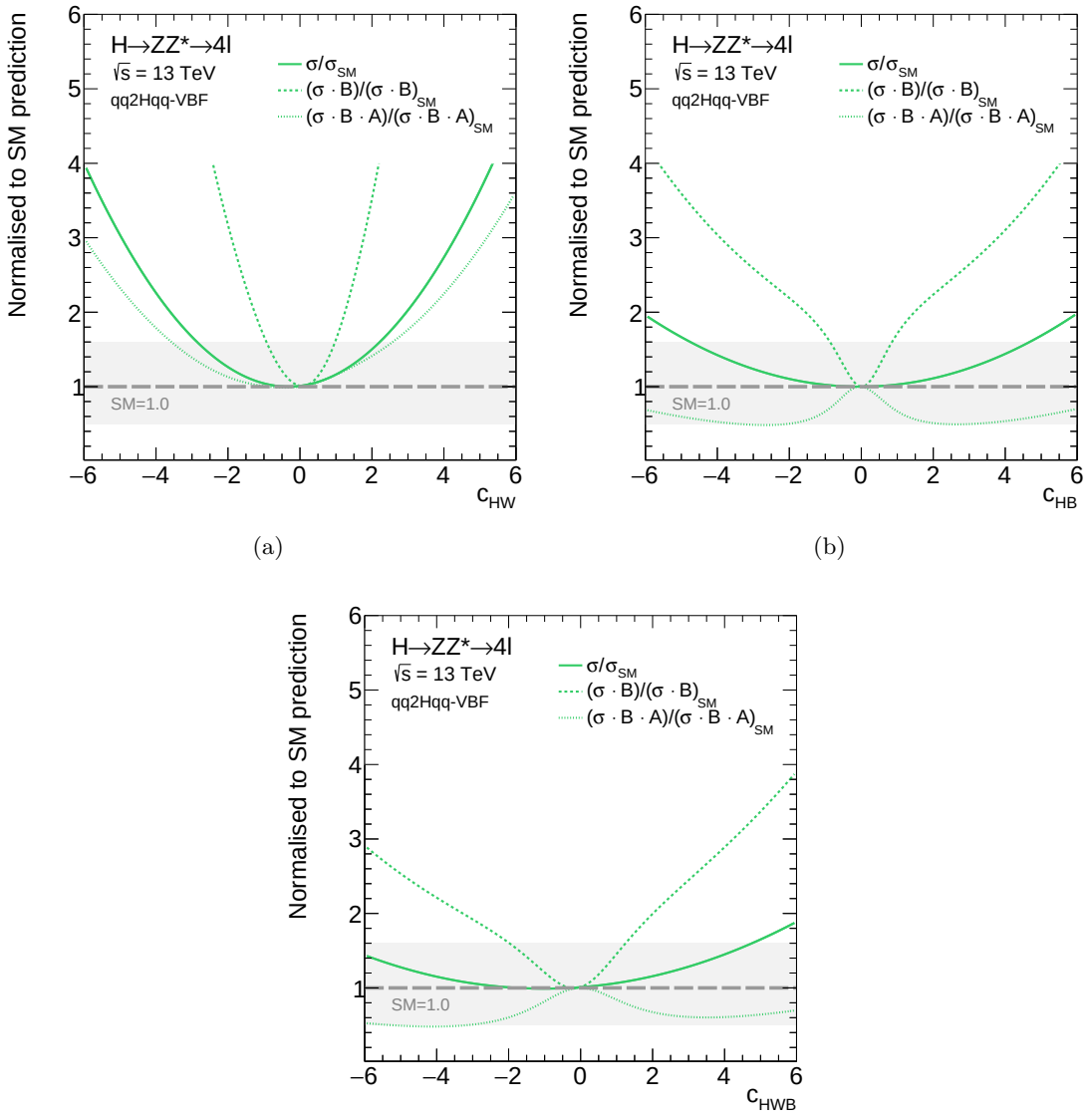
Figure 7.32: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter c_{HG} in (a) gg2H-0j-p_T^H-High and (b) gg2H-p_T^H-High production bin and (c) in dependence on c_{uH} in the ttH production bin. The grey band indicates the expected sensitivity at 1 σ level from the SM production cross section measurement.



(a) (b)

(c)

Figure 7.33: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter (a) c_{HW}, (b) c_{HB} and (c) c_{HWB} in the gg2H-0j-p_T^H-High production bin. The grey band indicates the expected sensitivity at 1 σ level from the SM production cross section measurement.



(a) (b)

(c)

Figure 7.34: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter (a) c_{HW}, (b) c_{HB} and (c) c_{HWB} in the qq2Hqq-VBF production bin. The grey band indicates the expected sensitivity at 1 σ level from the SM production cross section measurement.

metrisation cannot be neglected. Adding the parametrisation of the branching ratio the sensitivity is improved. However, this improvement is lost for the BSM coupling parameter c_{HW} after accounting for the signal acceptance. On the other hand, the expected sensitivity remains approximately the same for c_{HWB} and c_{HB}, because the parametrisation of the signal acceptance for these two parameters flops the sign compared to the one for the c_{HW} parameter. Since the BSM coupling parameters c_{HW}, c_{HWB} and c_{HB} enter in the decay vertex, they are measured in all Higgs boson production modes, including the dominant gg2H production. The largest sensitivity for these BSM coupling parameters comes from the gg2H production bins due to the largest expected number of events in that bins. There is also a non-negligible sensitivity in qq2Hqq bins since the Wilson coefficient appear in the corresponding production vertex. The effective top Yukawa interaction is measured in the ttH production bin. Also here the quadratic terms cannot be neglected in the expected sensitivity range.

7.3.3 Validation with Simulated Data

In order to validate the EFT parametrisation several fully reconstructed BSM signal samples (see Section 4.2.2.2) are used. The configuration of the samples and the corresponding cross sections times branching ratios are summarised in Table 7.21. The values of the BSM coupling parameters are chosen such to cover the expected sensitivity range at the 1 σ level.

As an example the comparison of the $p_T^{4\ell}$ distribution obtained with the SM (N)NLO validation samples and with the SMEFT LO Monte Carlo samples are shown Figure 7.35 for the different Higgs boson production modes. Further distributions are shown in Appendix B.7. Since the VBF and VH-Had production mode are simulated as one sample in case of the BSM signal process, the corresponding EFT prediction is combined with the one for the VH-Lep production in order to compare with the best-prediction from the SM VBF and VH signal samples. In case of the gg2H production mode, the shape of the $p_T^{4\ell}$ distribution of the SMEFT LO sample differs from the SM higher-order predication (NNLO)sample. For the gg2H production the higher-order corrections have a larger impact on the cross section and the expected kinematic distributions compared to the other production modes. However, the distribution matches the LO expectation. Other distributions agree well with the NLO prediction.

Similarly, the $p_T^{4\ell}$ distributions for the reconstructed validation samples with the BSM signal are compared to the respective SM sample in Figure 7.36 separately for the gg2H, qq2Hqq, VH-Lep and ttH production mode. Further distributions are shown in Appendix B.7. Non-vanishing BSM coupling values lead to a slightly harder spectra of the four-lepton transverse momentum.

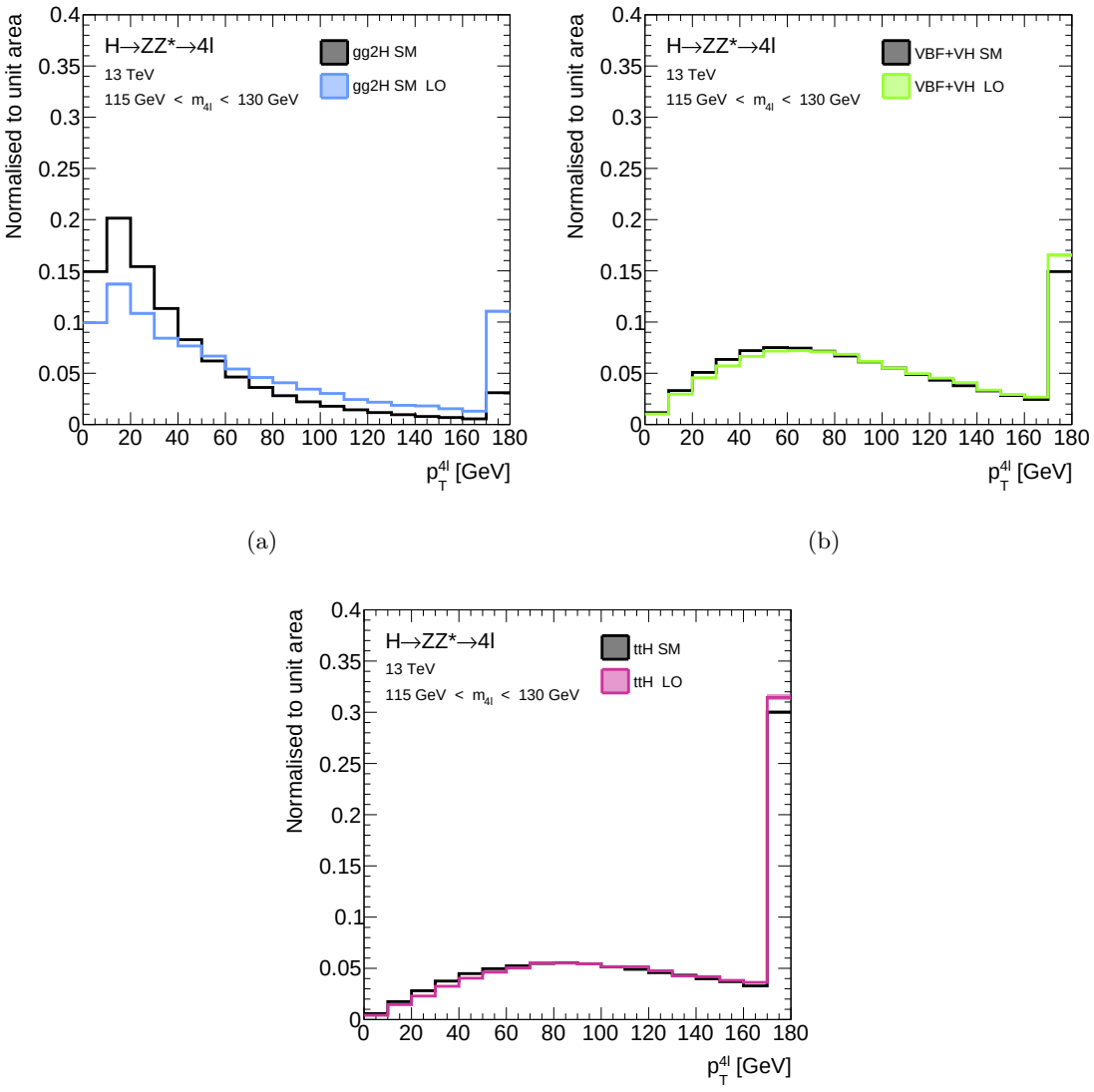
Table 7.21: Configuration of the fully simulated and reconstructed BSM signal samples used for the validation of the EFT parametrisation at reconstructed level. The corresponding cross sections times branching ratio ($\sigma \cdot \mathcal{B}$) are shown in addition.

BSM coupling parameter	$\sigma \cdot \mathcal{B}$ [fb]	BSM coupling parameter	$\sigma \cdot \mathcal{B}$ [fb]
gg2H production		*qq2Hqq production*	
SM	2.44	SM	0.41
$c_{HG} = 0.005$	2.82	$c_{HW} = 0.5$	0.48
$c_{HW} = 0.50$	2.64	$c_{HB} = 0.5$	0.55
$c_{HB} = 0.5$	3.32	$c_{HWB} = 0.5$	0.53
$c_{HWB} = 0.5$	3.12		
VH-Lep production		*ttH production*	
SM	0.038	SM	0.040
$c_{HW} = 0.5$	0.061	$c_{uH} = 0.5$	0.011

The EFT parametrisation of the production bins is performed on particle level. However, as described in Section 7.3.1, the measurement of the production cross sections is performed simultaneously in several event categories designed to target the different Higgs boson production bins. Since some of the observables used for the categorisation are in general sensitive to BSM physics, non-vanishing BSM coupling parameters could lead to the BSM dependence of the reconstruction efficiency in each category. In this case, the EFT parametrisation could only be done at reconstructed level, i.e. with a large set of fully simulated samples. To investigate whether such a BSM dependence exists, the selection efficiency ϵ in each of the reconstructed event categories i, defined as

$$\epsilon_i = \frac{N_i}{N_{\text{after } 4\ell \text{ selection}}}, \tag{7.34}$$

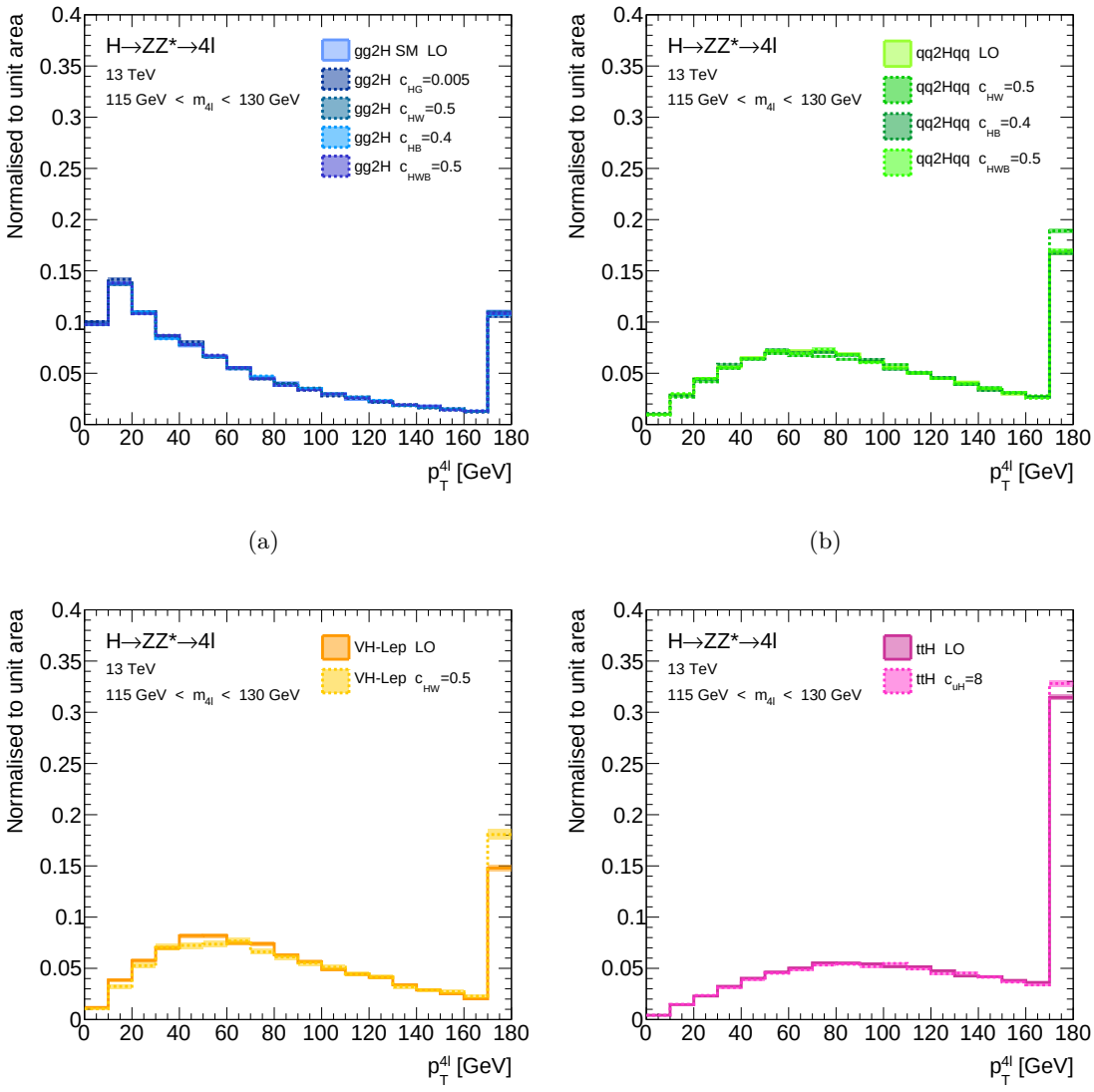
is calculated for each of the BSM signal samples, see Figure 7.37. N_i is the total number of events in a given reconstructed category and $N_{\text{after } 4\ell \text{ selection}}$ is the total number of events after the inclusive 4ℓ selection. Within the statistical uncertainties a good agreement is observed between the SM and various BSM hypotheses. Therefore, no additional EFT parametrisation of the event reconstructed efficiency is needed.



(a) (b)

(c)

Figure 7.35: Expected distribution of $p_T^{4\ell}$ for the SM (N)NLO and SM SMEFT LO sample for (a) gg2H, (b) VBF+VH and (c) ttH production.



(a) (b)

(c) (d)

Figure 7.36: Expected distribution of $p_T^{4\ell}$ for SM and selected BSM Higgs boson signals for (a) gg2H, (b) qq2Hqq, (c) VH-Lep and (c) ttH production.

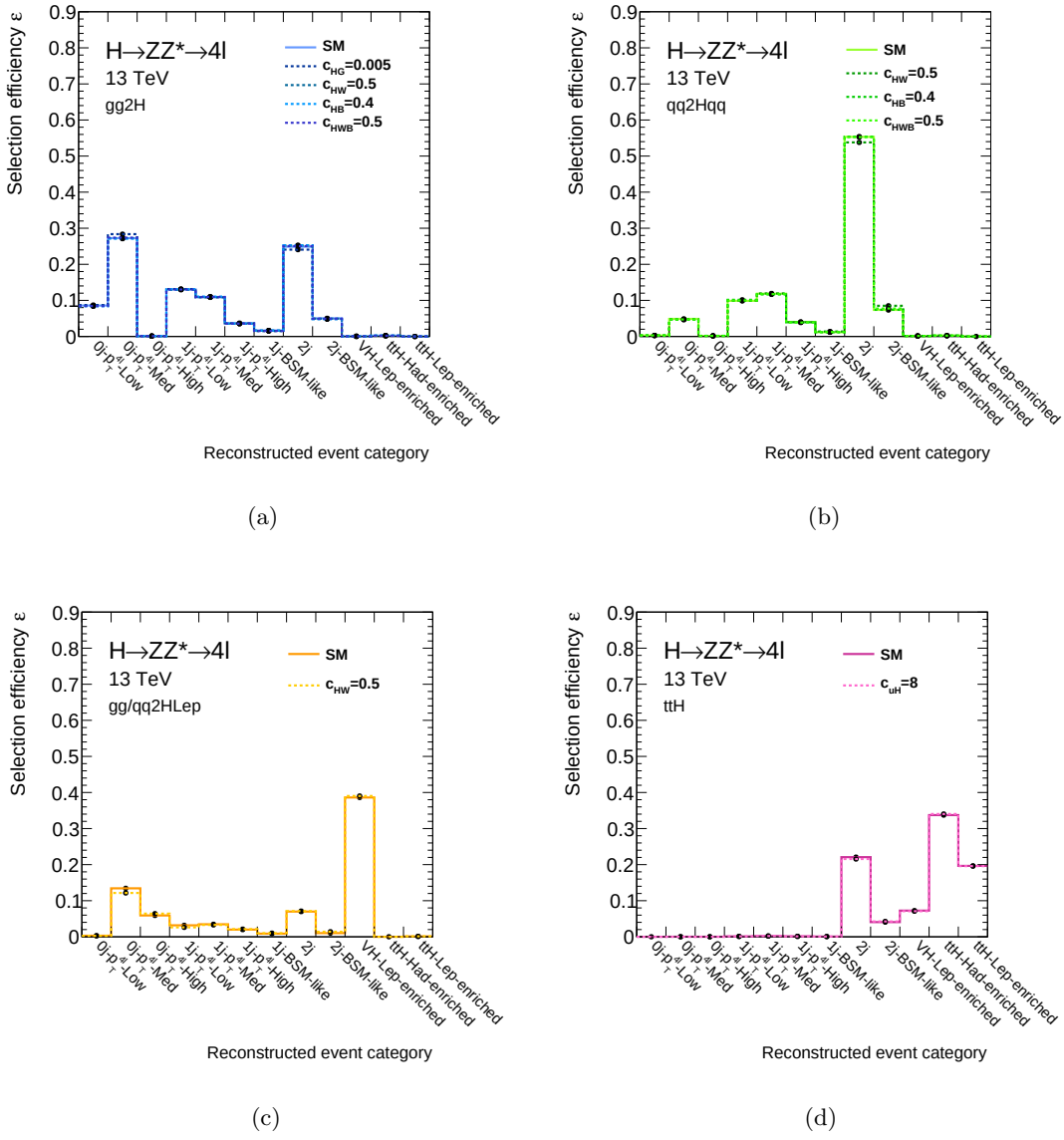
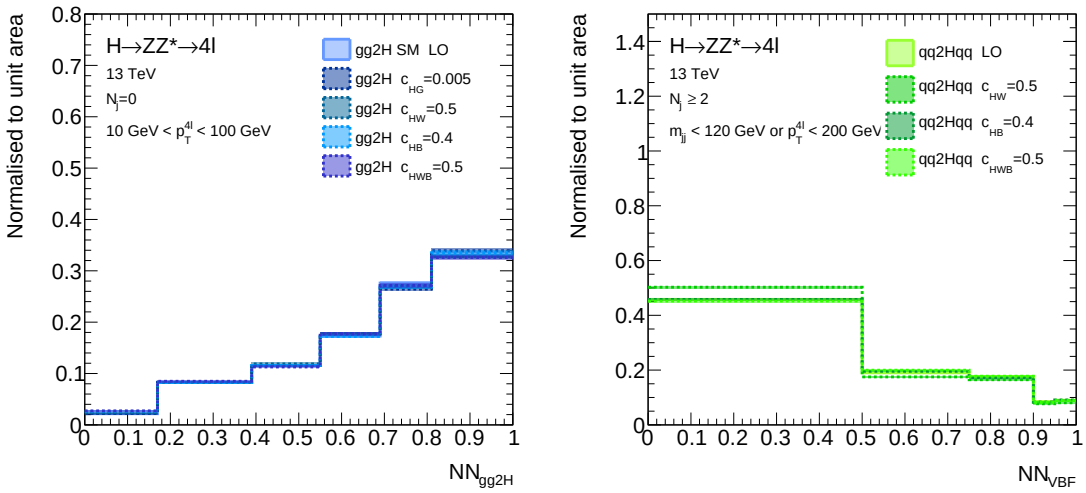
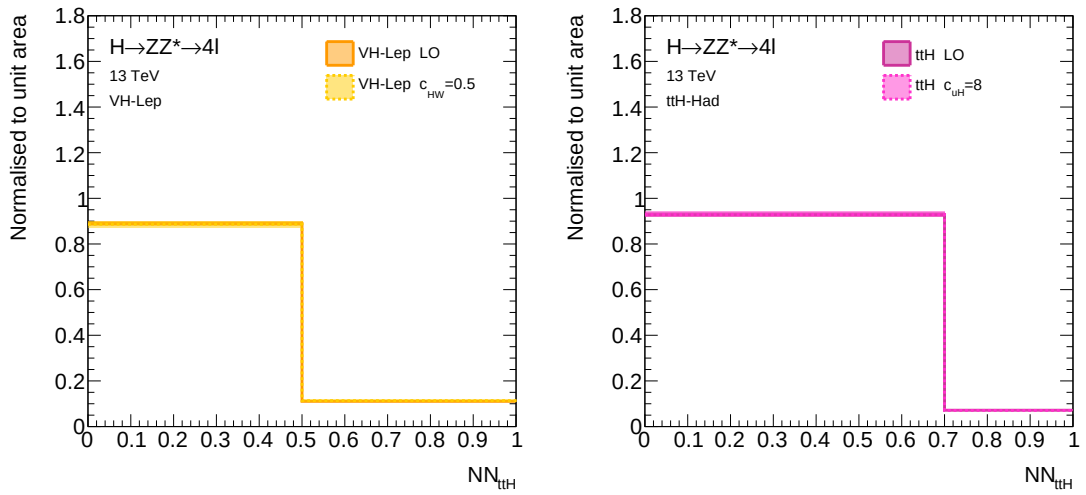


Figure 7.37: Event reconstruction efficiency in each event category, shown for different values of BSM coupling parameters and for the SM point, separately for the (a) gg2H, (b), qq2Hqq, (c) VH-Lep and (d) ttH production.

In order to increase the sensitivity of the cross section measurements, NN discriminants (Section 7.3.1) are employed in the reconstructed event categories. However, the BSM contributions can modify the distribution of a given variable which is used for the neutral network. If this BSM dependence is not taken into account for the signal parametrisation, the measurement results may be biased. The distributions of the NN discriminants in reconstructed event categories with large sensitivity to a given production bin are shown in Figure 7.38. For all NN discriminants, a small BSM dependence is observed, which can be neglected with the current statistical precision of the measurement.



(a) (b)



(c) (d)

Figure 7.38: Expected distributions of NN discriminants (a) $\mathrm{NN}_{\mathrm{ggF}}$ in the $0j$-$p_T^{4\ell}$-Med, (b) $\mathrm{NN}_{\mathrm{VBF}}$ in the $2j$, (c) $\mathrm{NN}_{\mathrm{ttH}}$ in the $0j$-$p_T^{4\ell}$-High and (d) $\mathrm{NN}_{\mathrm{ttH}}$ in the ttH-Had-enriched category.

7.3.4 Systematic Uncertainties

The impact of the QCD scale and PDF variations is computed using the MadGraph5_aMC@NLO generator and evaluated in a similar way as described in Section 5.4.2. The QCD scale uncertainties are evaluated by varying the renormalisation and factorisation scale and by comparing the nominal choice ($\mu_{0,R}$ and $\mu_{0,F}$) to different pairs of QCD renormalisation and factorisation variations of $\mu = \mu_0/2$ or $\mu = 2 \cdot \mu_0$. The impact of the PDF uncertainties is calculated using the internal variations of the NNPDF23lo PDF set.

As an example, the evaluated uncertainties on the event yield for the gg2H production in the production bin with largest sensitivity (gg2H-0j-p_T^H-High) for each of the reconstructed samples are shown in Figure 7.39. The impact of the QCD scale and PDF uncertainties is very large (in the order of 10-20%) and the size of the variations is approximately the same for different values of Wilson coefficients. This holds also for the other particle level production bins.

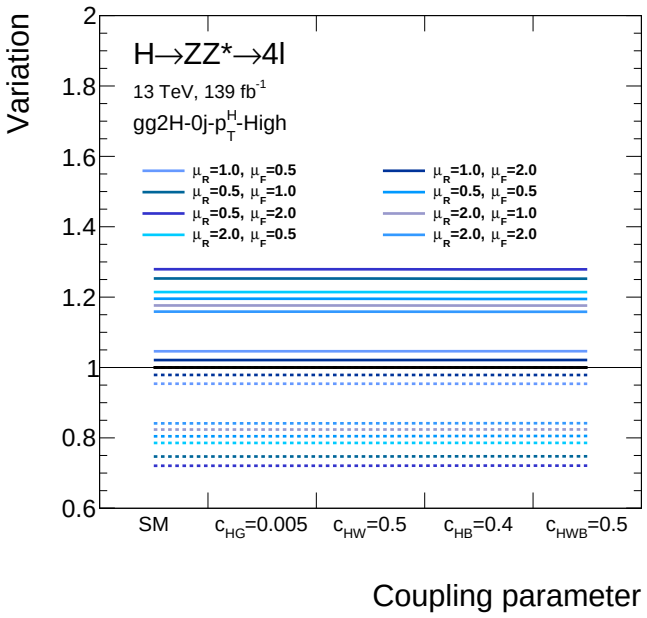
The size of the calculated systematic uncertainties is large, because the Monte Carlo samples are generated at LO accuracy. Since there is only a negligible dependence on the value of the BSM coupling parameter and the BSM parametrisation is introduced as a relative correction to the best-prediction of the SM, the theoretical uncertainties assigned on the SM best-prediction are used. In addition, it is assumed that there is no BSM dependence of the experimental systematics. They are also taken from the production cross section measurement.

7.3.5 Statistical Analysis Model

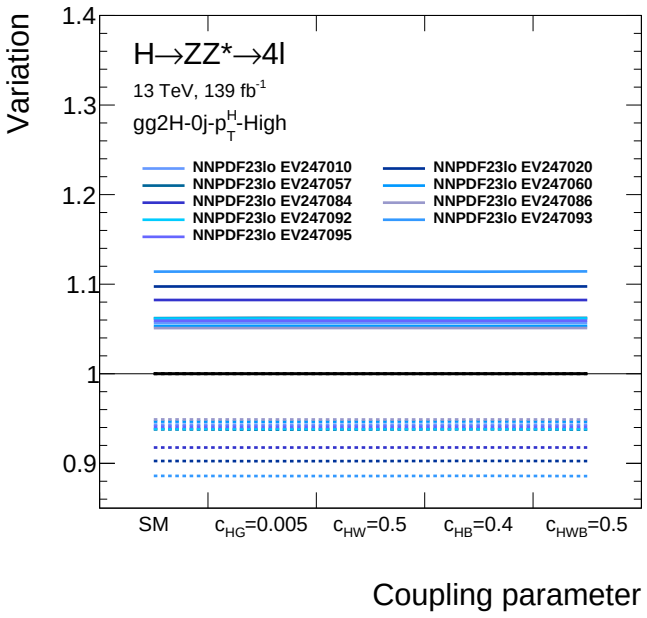
The signal EFT parametrisation is implemented in the likelihood function of Equation 7.19 by means of the BSM dependent signal strength parameters $\mu^p(\boldsymbol{c})$ for each given production bin p,

$$\mu^p(\boldsymbol{c}) = \frac{\sigma(\boldsymbol{c})}{\sigma_{\mathrm{SM}}} \cdot \frac{\mathcal{B}^{4\ell}(\boldsymbol{c})}{\mathcal{B}^{4\ell}_{\mathrm{SM}}} \cdot \frac{\mathcal{A}(\boldsymbol{c})}{\mathcal{A}_{\mathrm{SM}}}. \tag{7.35}$$

This is then fitted to the observed event yields. Default SM predictions at the highest available order are employed for the cross sections and the branching ratio multiplying the signal strengths in the likelihood function from Equation 7.19. Constraints on the Wilson coefficients are obtained using a test statistic based on the ratio of profiled likelihoods as given in Equation 5.10.



(a)



(b)

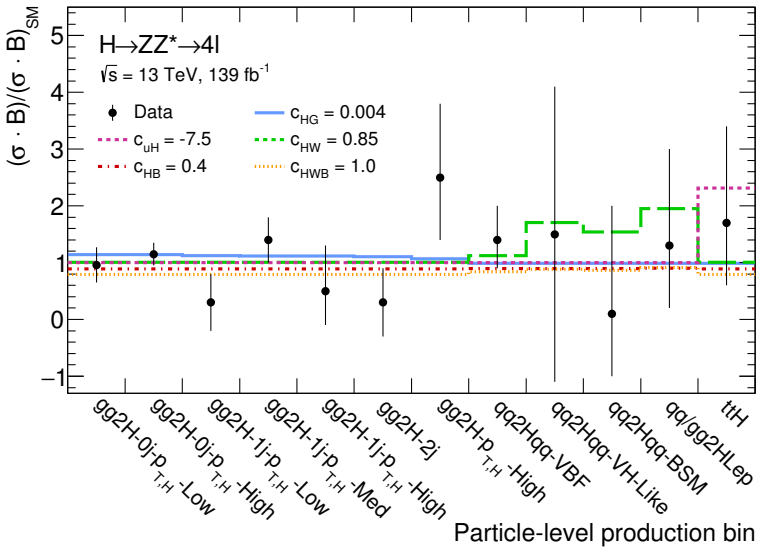
Figure 7.39: The impact of (a) QCD scale and (b) PDF variations on the event yield in the gg2H-0j-p_T^H-High production bin for gg2H production with different BSM coupling values. For visualisation, only a subset of PDF uncertainties is shown. *Up* (*down*) variations are shown as solid (dashed) lines.

7.3.6 Results

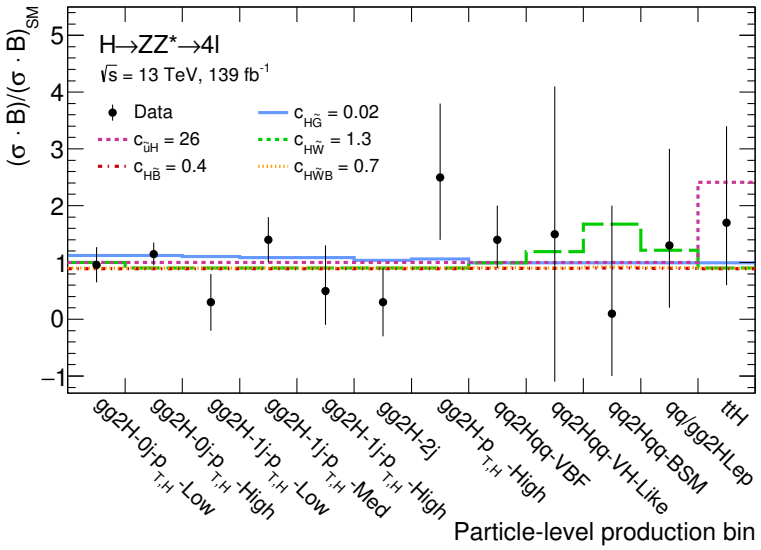
The observed cross section times branching ratio normalised to the SM prediction $(\sigma \cdot \mathcal{B})/(\sigma \cdot \mathcal{B})_{\mathrm{SM}}$ measured in the Reduced-Stage-1.1 production bins is shown in Figure 7.40. In all particle level production bins the measured value is in a good agreement with the SM prediction of one. A small excess is observed in the gg2H-p_T^H-High production bin which originates from the about two times more events observed than expected in the $1j$-$p_T^{4\ell}$-BSM reconstructed event category. A small deficit is measured in the gg2H-$1j$-p_T^H-Low production bin due to less events observed than expected in the corresponding $1j$-$p_T^{4\ell}$-Low event category.

The observed $(\sigma \cdot \mathcal{B})/(\sigma \cdot \mathcal{B})_{\mathrm{SM}}$ ratio is compared to the prediction of the signal model in each production bin. The values of the Wilson coefficients are chosen to be within the quoted 68% CL of the cross section measurement (Section 7.3.1). All Wilson coefficients, except c_{HW}, c_{uH} and $c_{\tilde{u}H}$, affect the cross section times branching ratio by $10-20\%$ in the gg2H production bins. Although the effective coupling to the gluons (c_{HG} and $c_{H\widetilde{G}}$) contributes in the gg2H and ttH production bins, but it is mainly measured in the gg2H bins due to a larger number of expected events. Deviations from the SM prediction in the qq2Hqq production bins can be mostly described by the Wilson coefficient c_{HW} or $c_{H\widetilde{W}}$. The sensitivity for these Wilson coefficients originates mainly from BSM contributions to the production vertex, since modifications in the decay vertex are compensated by considering the BSM dependence of the detector acceptance. The BSM couplings c_{uH} and $c_{\tilde{u}H}$ can only be constraint in the ttH production bin. The sensitivity to $c_{HB}, c_{H\widetilde{B}}, c_{HWB}$ and $c_{H\widetilde{W}B}$ is mainly obtained from the $H \to ZZ^* \to 4\ell$ decay predominantly measured in the gg2H production bins. The prediction of the BSM parametrisation in each of the particle level production bins does not match the quoted error of the cross section measurement, since all production bins are fitted simultaneously.

Due to the limited number of observed events, at most two BSM coupling parameter are fitted at a time while the others are set to zero. Firstly, each individual BSM coupling parameter is fitted by assuming that all other BSM coupling parameters are zero. The respective fit results are summarised in Figure 7.41. The observed and expected values of the test statistics from scans over the CP-even and CP-odd Wilson coefficients are shown in Figure 7.42, 7.43, 7.44, 7.45 and 7.46, respectively. Table 7.22 shows the corresponding observed and expected 95% CL interval, the best-fit parameter values and the corresponding deviation from the SM expectation. The strongest constraints are obtained on the c_{HB} and $c_{H\widetilde{W}B}$ BSM coupling parameters which are mostly affect the $H \to ZZ^*$ decays. The corresponding 95% CL intervals are $[-0.61, 0.58]$ and $[-1.03, 1.03]$, respectively. The constraints obtained on c_{HW} and $c_{H\widetilde{W}}$ are weaker ($[-3.40, 2.12]$ and $[-2.32, 2.32]$ at 95% CL), since they are mainly determined from the VBF and VH production vertices. However, the largest deviation from the SM is observed for these Wilson coefficients corresponding to 0.4 σ and 0.2 σ deviation from the SM, respectively.

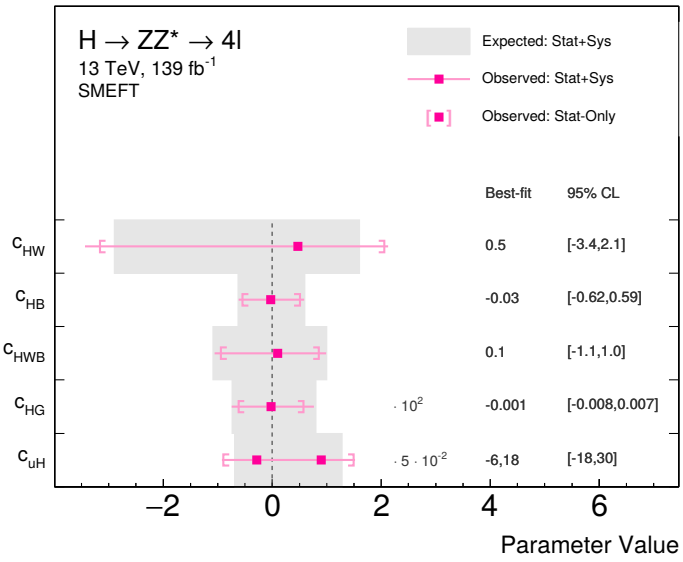


(a)

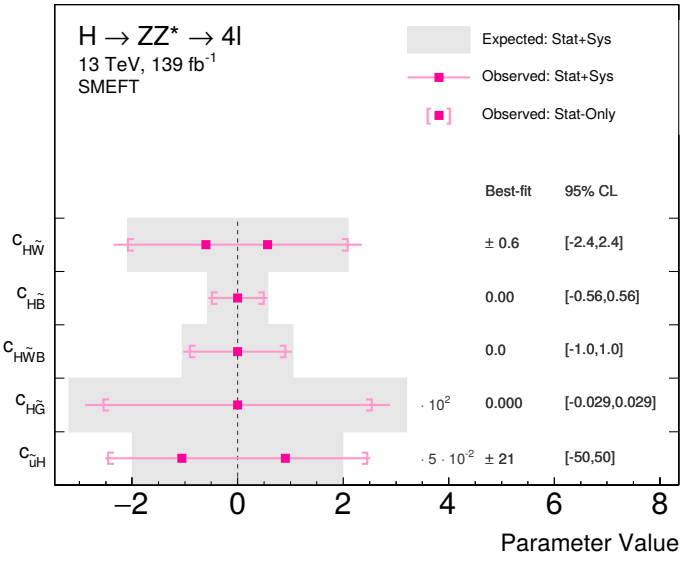


(b)

Figure 7.40: Observed cross section times branching ratio normalised to the SM prediction $(\sigma \cdot \mathcal{B})/(\sigma \cdot \mathcal{B})_{\mathrm{SM}}$ measured in the Reduced-Stage-1.1 production bins compared to the prediction in each production bin derived from the (a) CP-even and (b) CP-odd BSM parametrisation. The error on the cross section measurement is given at 68% CL interval.



(a)



(b)

Figure 7.41: The observed and expected values of BSM coupling parameters from (a) CP-even and (b) CP-odd operators obtained for an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV. Only one BSM coupling parameter is fitted at a time while all others are set to zero.

Table 7.22: The expected and observed confidence intervals at 95% CL on the BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV. The limits are computed using the confidence level interval method. Only one BSM coupling parameter is fitted at a time with all others set to zero.

BSM coupling parameter	95% CL interval		Observed best-fit	Deviation from SM
	Expected	Observed		
	CP-even BSM coupling parameter			
c_{HG}	$[-0.007, 0.008]$	$[-0.008, 0.007]$	-0.001	0.3 σ
c_{uH}	$[-14, 26]$	$[-18, 30]$	$-6, 18$	0.6 σ
c_{HW}	$[-2.9, 1.6]$	$[-3.4, 2.1]$	0.5	0.4 σ
c_{HB}	$[-0.62, 0.60]$	$[-0.62, 0.59]$	-0.03	0.0 σ
c_{HWB}	$[-1.09, 0.99]$	$[-1.06, 0.99]$	0.1	0.1 σ
	CP-odd BSM coupling parameter			
$c_{H\widetilde{G}}$	$[-0.031, 0.031]$	$[-0.029, 0.029]$	0.000	0.0 σ
$c_{\tilde{u}H}$	$[-40, 40]$	$[-50, 50]$	± 21	0.7 σ
$c_{H\widetilde{W}}$	$[-2.1, 2.1]$	$[-2.4, 2.4]$	± 0.6	0.2 σ
$c_{H\widetilde{B}}$	$[-0.57, 0.57]$	$[-0.56, 0.56]$	0.00	0.0 σ
$c_{H\widetilde{W}B}$	$[-1.05, 1.05]$	$[-1.03, 1.03]$	0.0	0.0 σ

The observed 95% CL intervals for the top Yukawa coupling are $[-18, 30]$ ($[-50, 50]$) for the CP-even (CP-odd) BSM coupling parameter c_{uH} ($c_{\tilde{u}H}$). The best-fit values are in agreement with the SM prediction within 0.6 σ (0.7 σ) deviation from the SM.

For the Wilson coefficients where the linear approximation is not valid, i.e. the quadratic terms are not negligible, the constraints of the CP-even Wilson coefficients are in the same order as for the corresponding CP-odd Wilson coefficients. This is not the case for the effective Higgs boson coupling to the gluons. More stringent constraints can be set on the CP-even BSM coupling parameter c_{HG} corresponding to $[-0.0074, 0.0080]$ at 95% CL compared to $[-0.029, 0.029]$ for $c_{H\widetilde{G}}$. The reason is, that for c_{HG} the linear approximation is valid and the linear terms from the CP- odd operators do not contribute to the total production cross section. Therefore, the slightly smaller measured cross section times branching ratio in the gg2H-1j-p_T^H-Low, gg2H-1j-p_T^H-High and gg2H-2j is assigned to the CP-even BSM coupling parameter with a best-fit value of $\hat{c}_{HG} = -0.0002$ corresponding

to a deviation from the SM prediction of 0.3 σ. The deviations from the SM prediction for all one-dimensional scans of BSM coupling parameter are at most 0.7σ, thus indicates no evidence for new couplings.

In addition, two-dimensional scans of several combination of the BSM coupling parameters are performed. All other BSM coupling parameters are set to zero in the fit, except the two fitted ones. No mixtures of CP-even and CP-odd BSM contributions are probed. The observed and expected two-dimensional contours of the test statistics at 95% CL intervals are shown in Figure 7.47 and 7.48 for the CP-even and in Figure 7.49 and 7.50 for the CP-odd BSM coupling parameters. The values inside the shown contours surrounding the SM point are allowed at 95% CL interval, while the values outside are excluded. Table 7.23 and 7.24 summarises the best-fit values of the two-dimensional scans and the corresponding deviations from the SM. All such possible combinations are in good agreement with the SM predictions. The observed best-fit values agree well with the one obtained in the one-dimensional scans.

Table 7.23: The best-fit values and the corresponding deviation from the SM prediction obtained from the two-dimensional likelihood scans of the CP-even BSM coupling parameters performed with 139 fb^{-1} data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The limits are computed using the confidence level interval method. Except for the two fitted BSM coupling parameters, all others are set to zero.

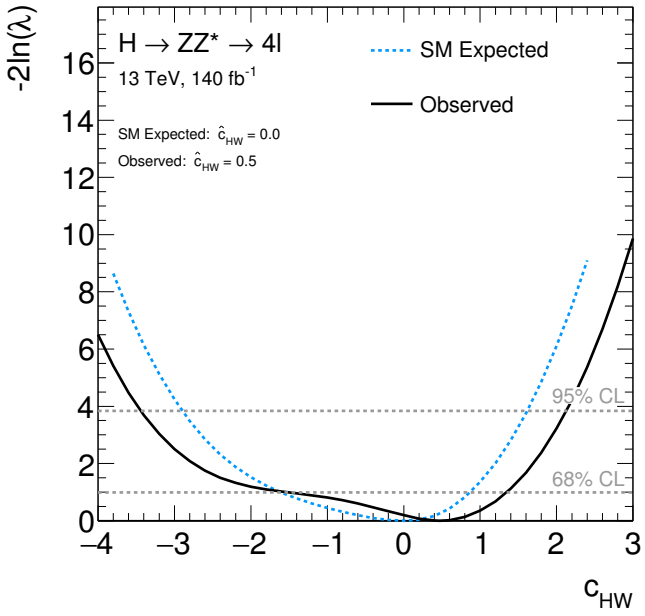
BSM coupling parameter	Observed best-fit		Deviation from SM
c_{HW}, c_{HB}	$\hat{c}_{HW} = 0.57$	$\hat{c}_{HB} = 0.05$	0.15 σ
c_{HW}, c_{HWB}	$\hat{c}_{HW} = 0.59$	$\hat{c}_{HWB} = -0.14$	0.15 σ
c_{HB}, c_{HWB}	$\hat{c}_{HB} = 1.74$	$\hat{c}_{HWB} = 2.24$	0.23 σ
c_{HG}, c_{HW}	$\hat{c}_{HG} = -0.001$	$\hat{c}_{HW} = 0.54$	0.20 σ
c_{HG}, c_{HB}	$\hat{c}_{HG} = -0.001$	$\hat{c}_{HB} = -0.04$	0.51 σ
c_{HG}, c_{HWB}	$\hat{c}_{HG} = -0.001$	$\hat{c}_{HWB} = -0.02$	0.00 σ
c_{HG}, c_{uH}	$\hat{c}_{HG} = -0.001$	$\hat{c}_{uH} = \pm 5.7$	0.29 σ

The shape of the contours of the two-dimensional scans originates from taking into account the detector acceptance. Around the SM point the parametrisation of the cross sections times branching ratio is quite flat which leads by constraining the parameter space to an elliptical contour. However, by introducing the BSM dependence of the de-

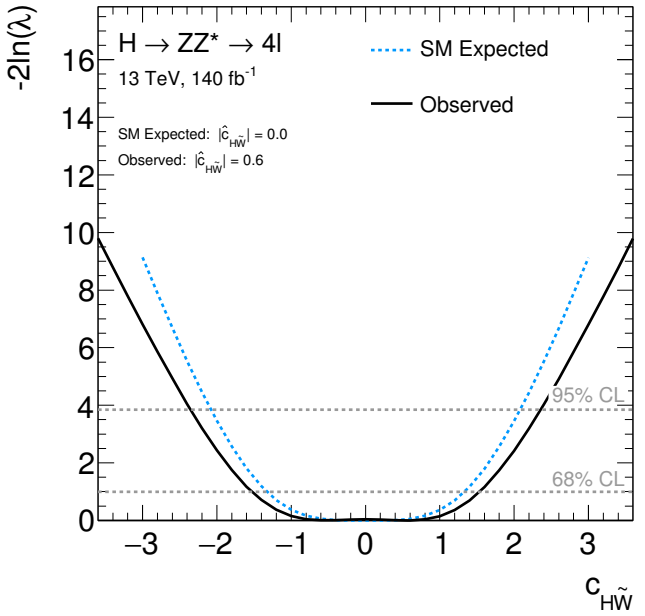
Table 7.24: The best-fit values and the corresponding deviation from the SM prediction obtained from the two-dimensional likelihood scans of the CP-odd BSM coupling parameters performed with 139 fb^{-1} data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The limits are computed using the confidence level interval method. Except for the two fitted BSM coupling parameters, all others are set to zero.

BSM coupling parameter	Observed best-fit		Deviation from SM
$c_{H\widetilde{W}}, c_{H\widetilde{B}}$	$\hat{c}_{H\widetilde{W}} = \pm 1.12$	$\hat{c}_{H\widetilde{B}} = \mp 0.21$	0.11 σ
$c_{H\widetilde{W}}, c_{H\widetilde{W}B}$	$\hat{c}_{H\widetilde{W}} = \pm 1.13$	$\hat{c}_{H\widetilde{W}B} = \pm 0.39$	0.11 σ
$c_{H\widetilde{B}}, c_{H\widetilde{W}B}$	$\hat{c}_{H\widetilde{B}} = 0.00$	$\hat{c}_{H\widetilde{W}B} = 0.00$	0.00 σ
$c_{H\widetilde{G}}, c_{H\widetilde{W}}$	$\hat{c}_{H\widetilde{G}} = 0.000$	$\hat{c}_{H\widetilde{W}} = \pm 0.56$	0.02 σ
$c_{H\widetilde{G}}, c_{H\widetilde{B}}$	$\hat{c}_{H\widetilde{G}} = 0.000$	$\hat{c}_{H\widetilde{B}} = 0.00$	0.00 σ
$c_{H\widetilde{G}}, c_{H\widetilde{W}B}$	$\hat{c}_{H\widetilde{G}} = 0.000$	$\hat{c}_{H\widetilde{W}B} = 0.00$	0.00 σ
$c_{H\widetilde{G}}, c_{\tilde{u}H}$	$\hat{c}_{H\widetilde{G}} = 0.000$	$\hat{c}_{\tilde{u}H} = \pm 21$	0.27 σ

tector acceptance a larger BSM dependence appears around the SM point and in addition the sensitivity decreases. Therefore, by constraining the parameter space the mixed terms between the two Wilson coefficients gets more important, which introduce the shape of the contours of the two-dimensional scans. The largest deviations from the SM are observed for c_{HG} vs. c_{HB} and $c_{H\widetilde{G}}$ vs. $c_{\tilde{u}H}$ corresponding to 0.51 σ and 0.3 σ. All results are in a very good agreement with the SM predictions, therefore no evidence for new BSM coupling contributions.

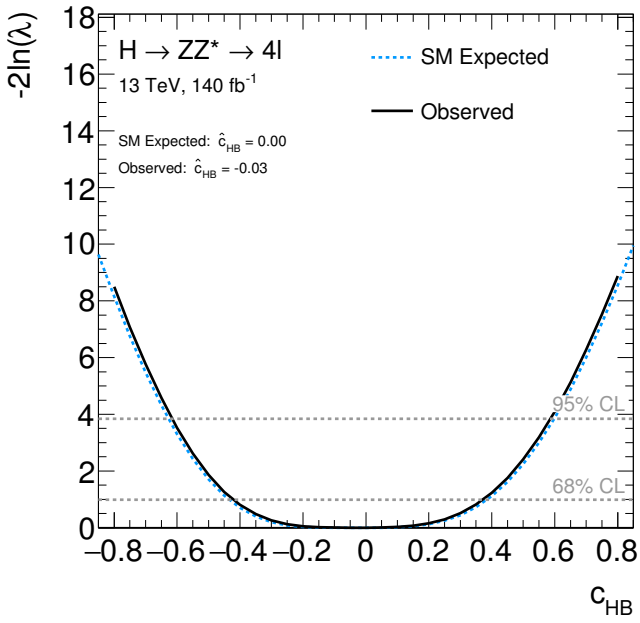


(a)

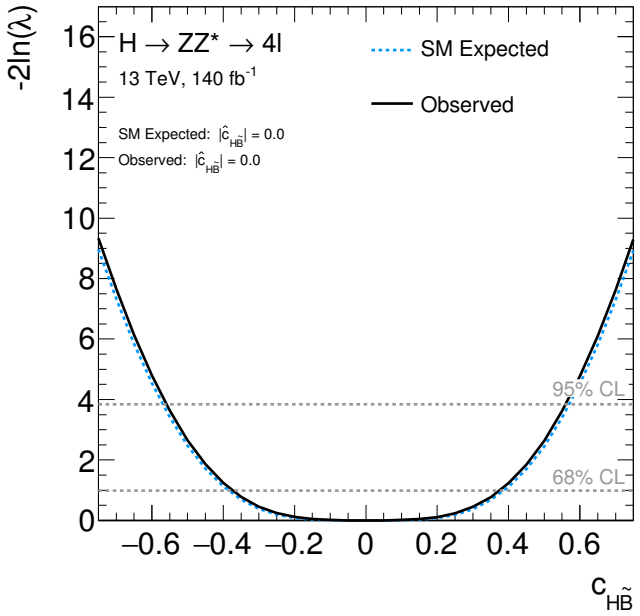


(b)

Figure 7.42: The observed (solid black line) and SM expected (dashed blue line) test statistic from scans of the BSM coupling parameters (a) c_{HW} and (b) $c_{H\widetilde{W}}$. Results are obtained with 139 fb^{-1} data at $\sqrt{s} = 13$ TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic. Except for the fitted BSM coupling parameter, all others are set to zero.

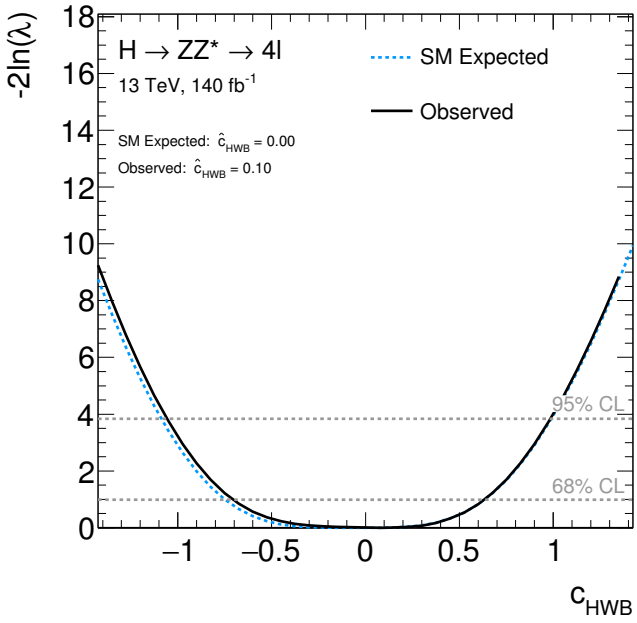


(a)

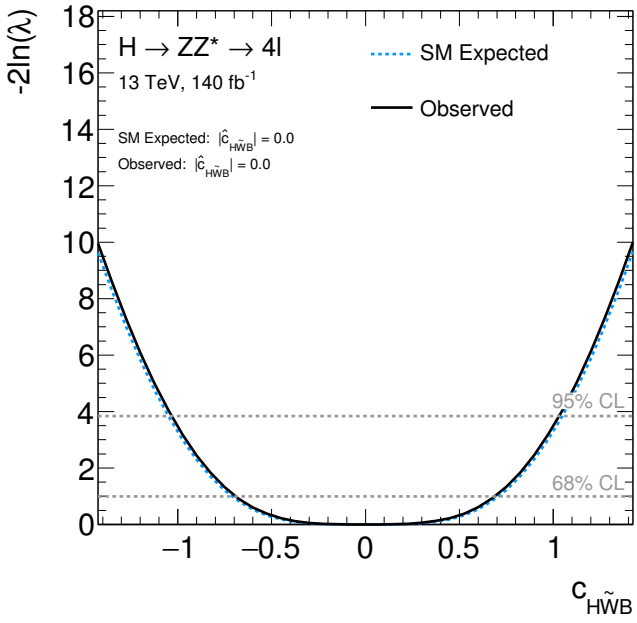


(b)

Figure 7.43: The observed (solid black line) and SM expected (dashed blue line) test statistic from scans of the BSM coupling parameters (a) c_{HB} and (b) $c_{H\widetilde{B}}$. Results are obtained with 139 fb^{-1} data at $\sqrt{s} = 13$ TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic. Except for the fitted BSM coupling parameter, all others are set to zero.

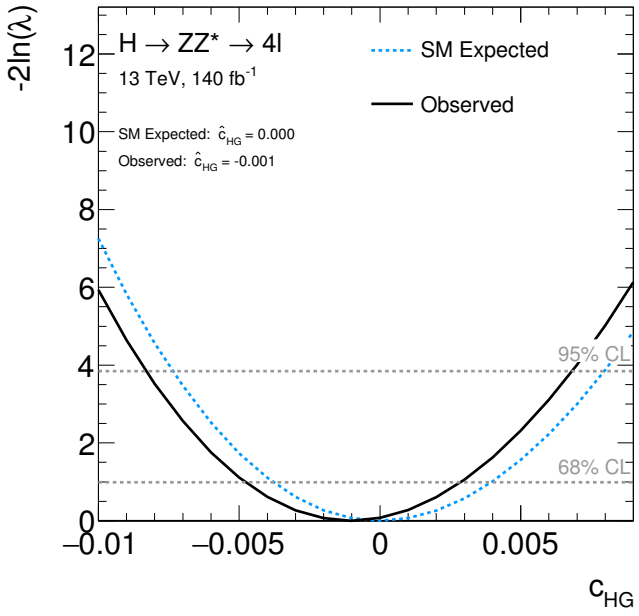


(a)

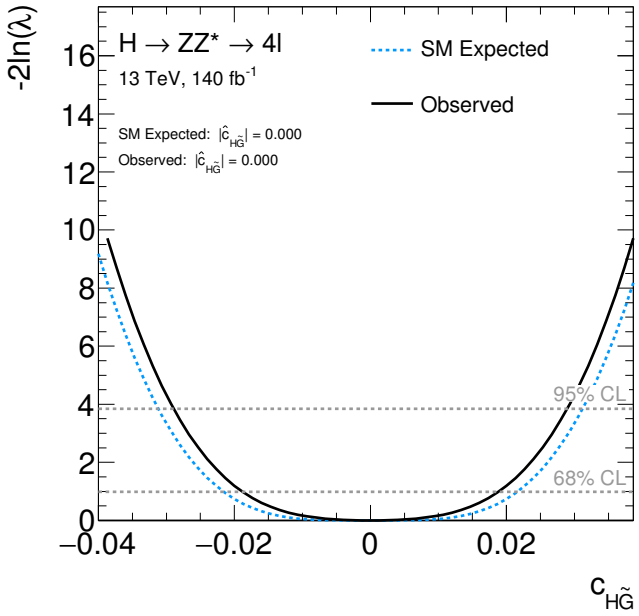


(b)

Figure 7.44: The observed (solid black line) and SM expected (dashed blue line) test statistic from scans of the BSM coupling parameters (a) c_{HWB} and (b) $c_{H\widetilde{W}B}$. Results are obtained with 139 fb^{-1} data at $\sqrt{s} = 13$ TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic. Except for the fitted BSM coupling parameter, all others are set to zero.

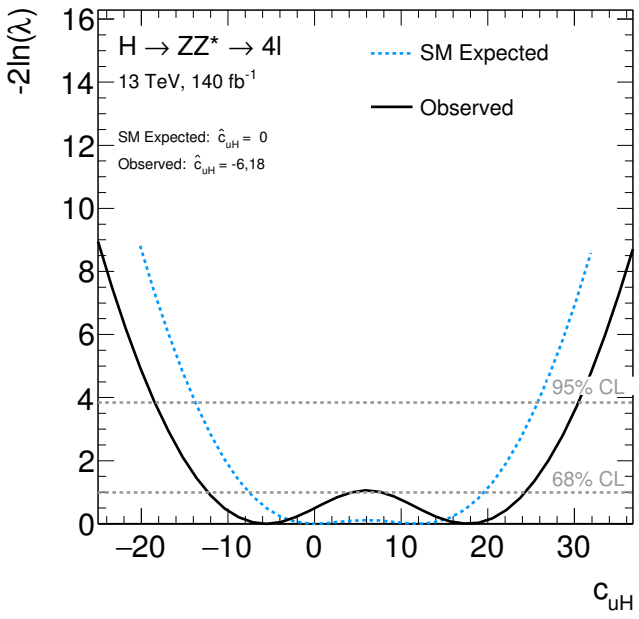


(a)

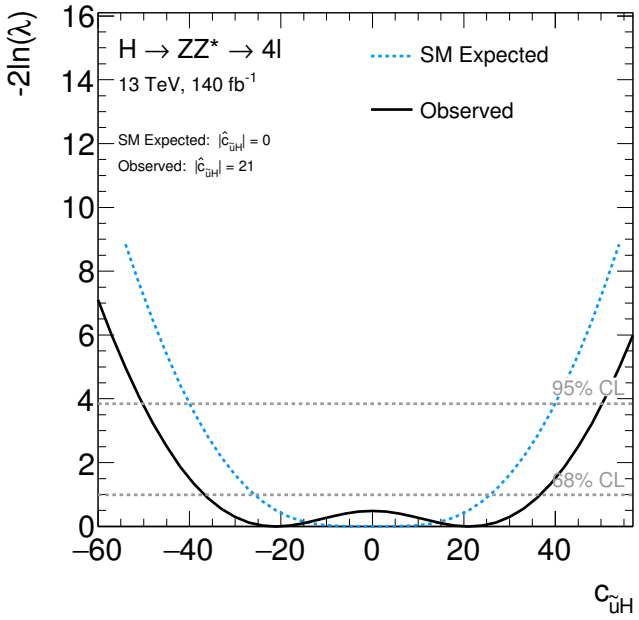


(b)

Figure 7.45: The observed (solid black line) and SM expected (dashed blue line) test statistic from scans of the BSM coupling parameters (a) c_{HG} and (b) $c_{H\widetilde{G}}$. Results are obtained with 139 fb^{-1} data at $\sqrt{s} = 13$ TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic. Except for the fitted BSM coupling parameter, all others are set to zero.



(a)



(b)

Figure 7.46: The observed (solid black line) and SM expected (dashed blue line) test statistic from scans of the BSM coupling parameters (a) c_{uH} and (b) $c_{\tilde{u}H}$. Results are obtained with 139 fb^{-1} data at $\sqrt{s} = 13$ TeV. The horizontal lines indicate the value of the test statistic corresponding to the 68% CL and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic. Except for the fitted BSM coupling parameter, all others are set to zero.

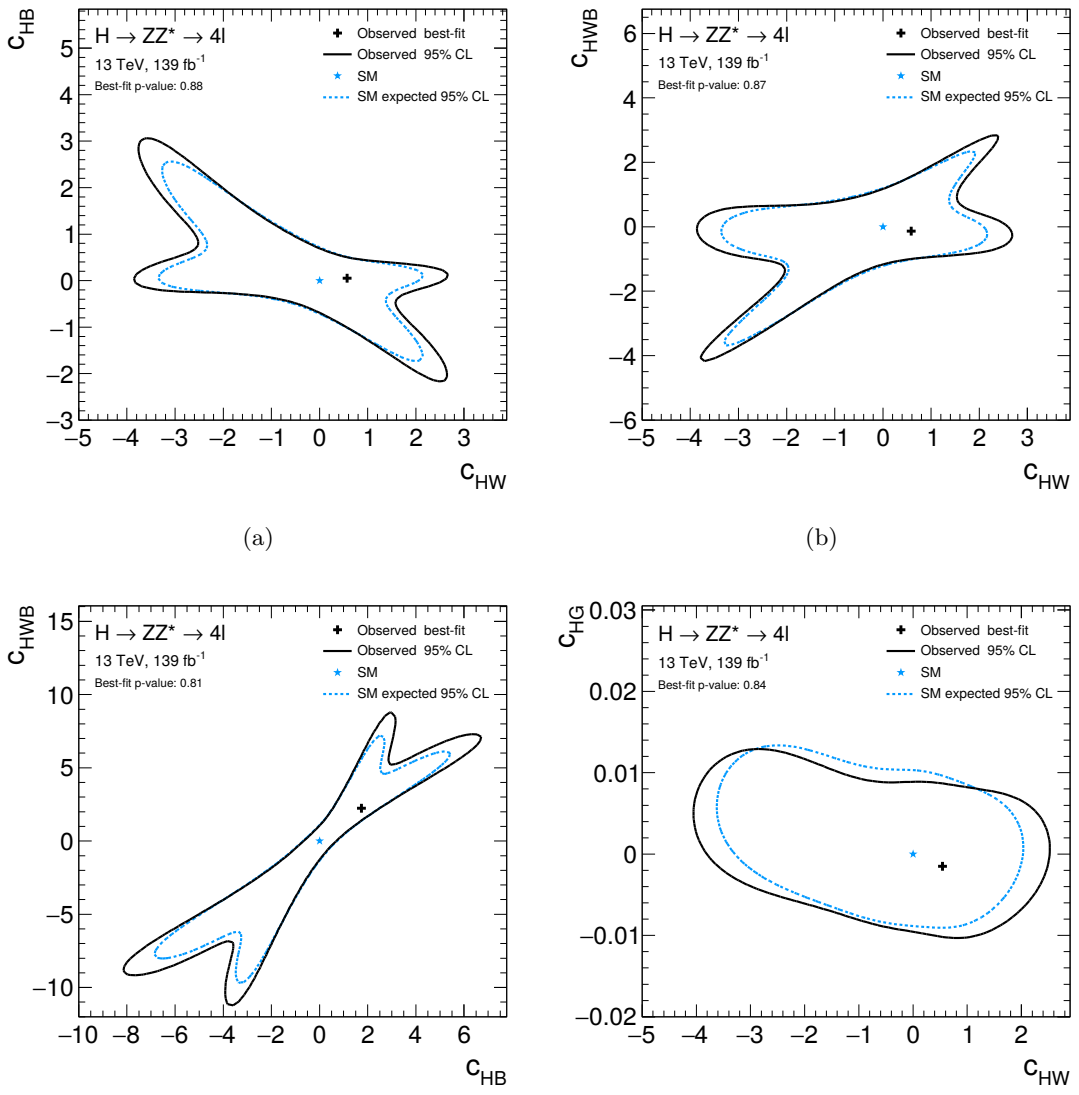
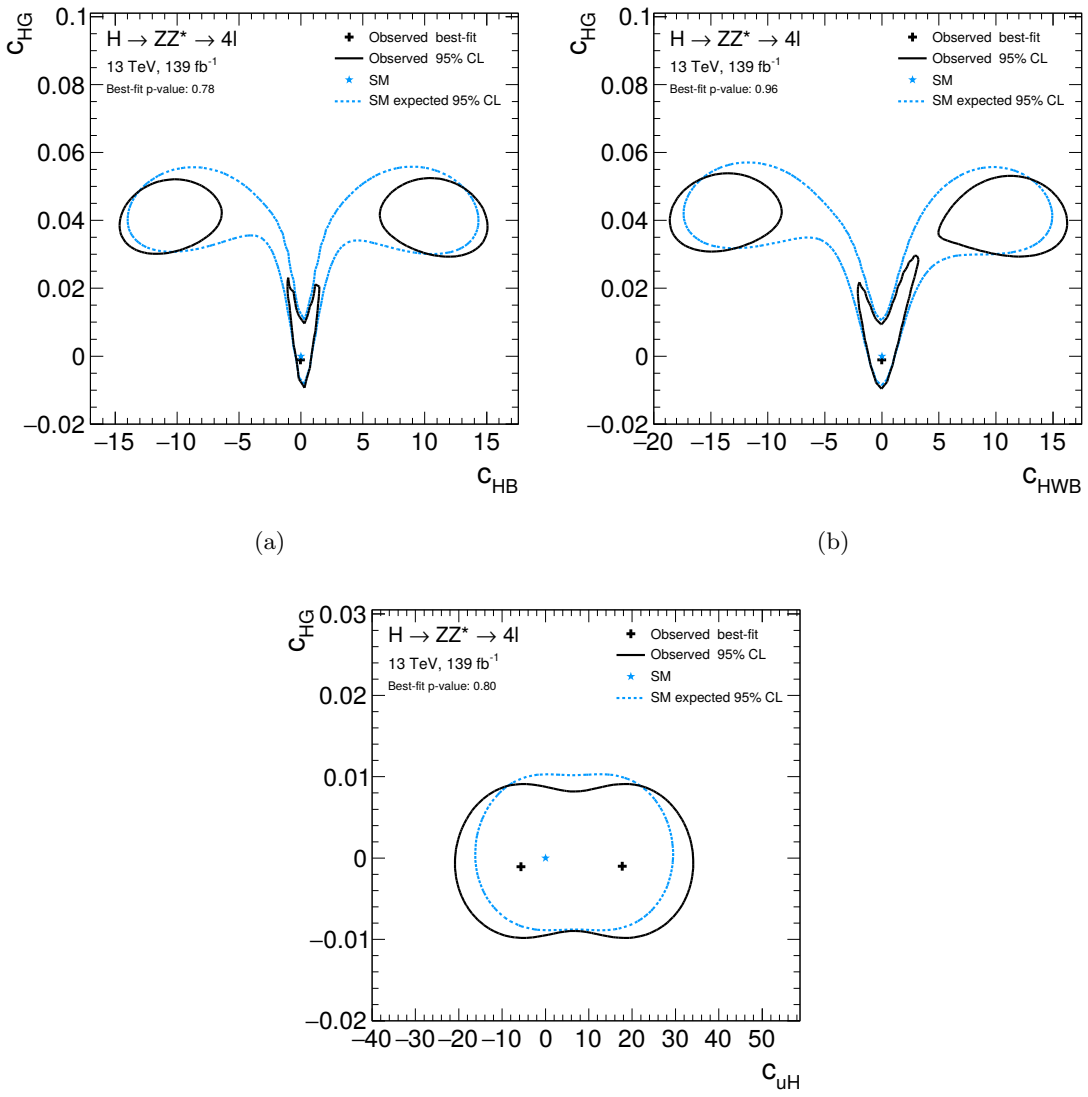


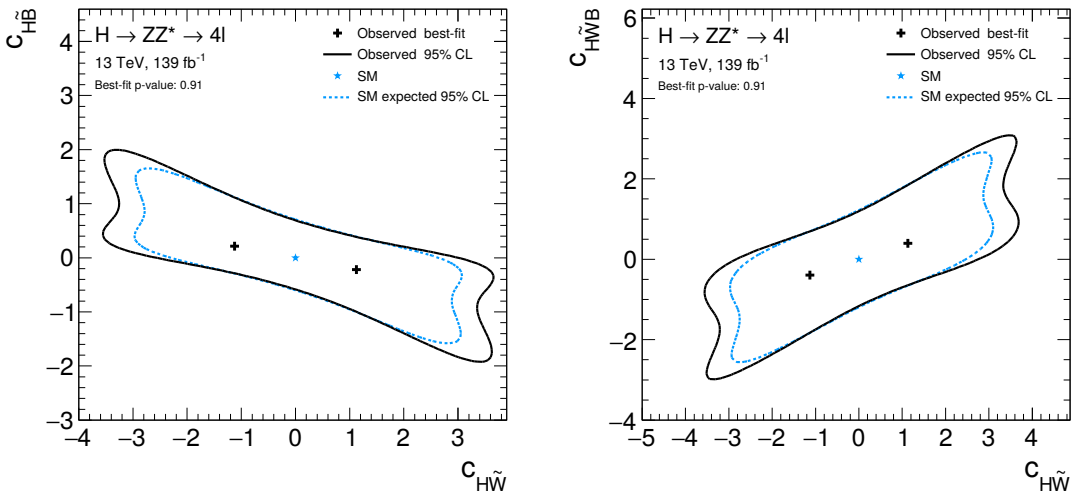
Figure 7.47: Expected and observed likelihood contours for the two-dimensional fit of the CP-even BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV. The values of the Wilson coefficients, which are not among the two parameters of interest, are set to zero.



(a) (b)

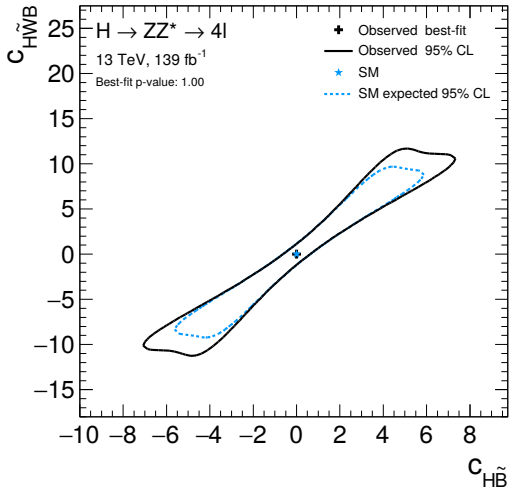
(c)

Figure 7.48: Expected and observed likelihood contours for the two-dimensional fit of the CP-even BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV. The values of the Wilson coefficients, which are not among the two parameters of interest, are set to zero.

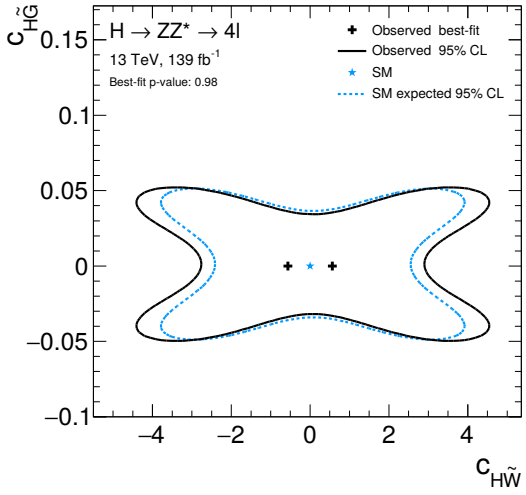


(a)

(b)

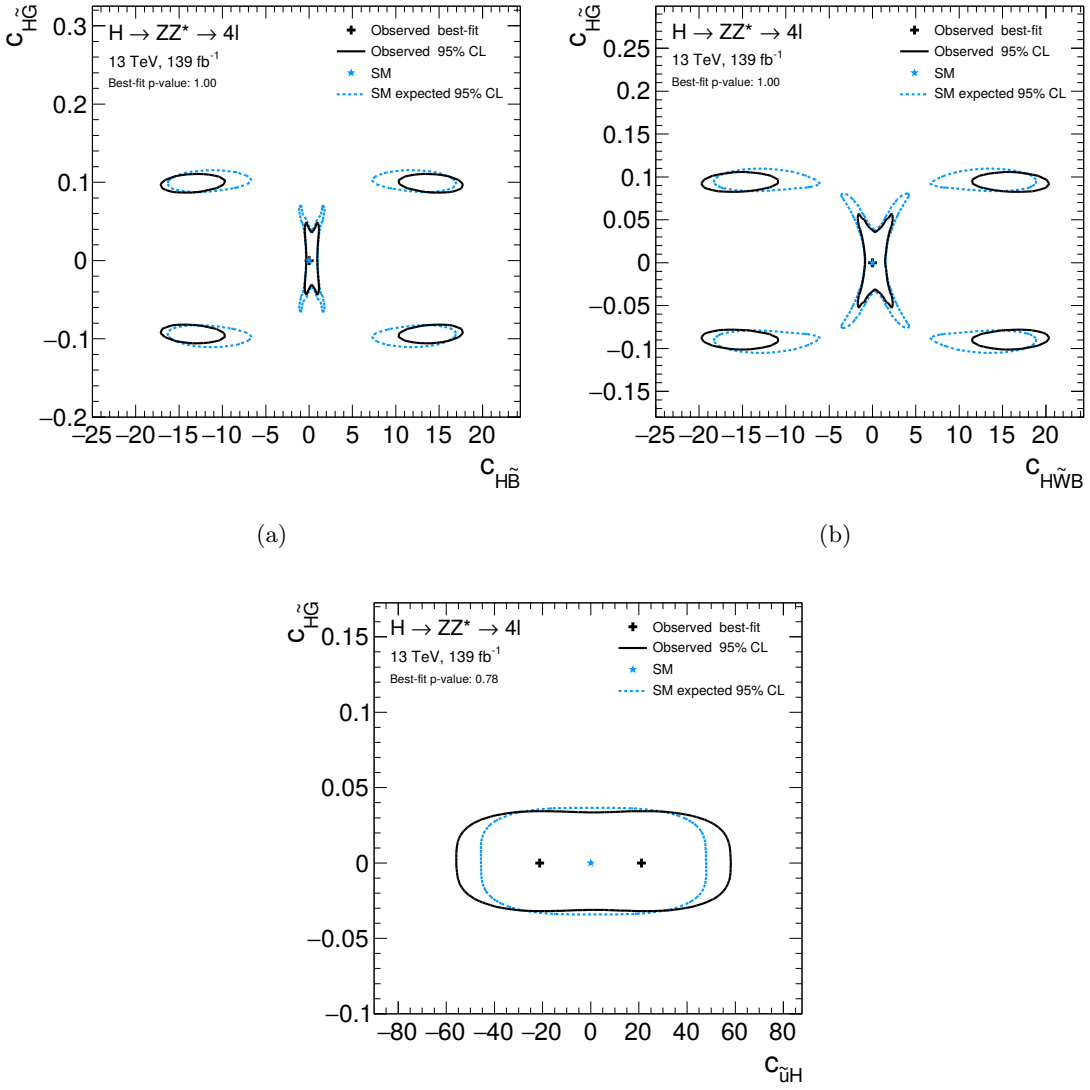


(c)



(d)

Figure 7.49: Expected and observed likelihood contours for the two-dimensional fit of the CP-odd BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV. The values of the Wilson coefficients, which are not among the two parameters of interest, are set to zero.



(a) (b)

(c)

Figure 7.50: Expected and observed likelihood contours for the two-dimensional fit of the CP-odd BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV. The values of the Wilson coefficients, which are not among the two parameters of interest, are set to zero.

7.4 Comparison with Previous Measurements

The two studies to probe the tensor structure of Higgs boson couplings presented in this chapter, the HC analysis (Section 7.2) and the interpretation within the SMEFT (Section 7.3), use different effective field theories to introduce the BSM contributions to the Higgs boson interaction. The former one is formulated in mass eigenstates after the electroweak symmetry breaking (HC framework), while the latter one is formulated before the electroweak symmetry breaking (SMEFT framework). A direct comparison, i.e. a translation between the SMEFT and the HC model, is only possible under certain assumptions. In the SMEFT model BSM contributions to the XVV vertex are described by three coupling parameters (c_{HW},c_{HB},c_{HWB} or $c_{H\widetilde{W}}$,$c_{H\widetilde{B}}$,$c_{H\widetilde{W}B}$), while these correspond to one in the HC model (κ_{HVV} or κ_{AVV}). Therefore, a comparison is only possible if all BSM coupling parameters in the SMEFT model are measured at the same time which was not possible in the presented study due to less amount of data. In addition, with the assumption in the HC model that the BSM coupling parameters to the Z bosons have the same correlation to corresponding couplings to W bosons as in the SM ($\kappa_{XWW} = \kappa_{XVV}$) a translation from HC to SMEFT is not possible. Thus, the results of the HC analysis and the interpretation within the SMEFT are not compared to each other.

However, the measurement of the tensor structure of Higgs boson couplings to weak vector bosons and gluons has also been performed by several other analyses in different Higgs boson decay channels with the ATLAS and CMS experiment (see Table 7.25). As mentioned in the beginning of this chapter, the BSM coupling contributions in the interaction vertex can be probed with different approaches: using total production and decay rates in different Higgs boson production modes, either by measuring directly the event yields in reconstructed event categories (see Section 7.2) or by interpreting the measured cross sections in each bin of the STXS framework (see Section 7.3); by means of inclusive differential distributions of kinematic properties without classification into event categories or particle level production bins; and from the shape of the distributions related to the kinematic properties of final state particles (see Chapter 8). The analyses summarised in Table 7.25 use one or more of these approaches to set constraints on BSM coupling parameters. All Higgs boson tensor structure measurements are in agreement with the SM expectation.

In the following, the constraints on BSM coupling parameters obtained in this thesis are compared to those from other analyses which use a compatible EFT framework, i.e. formulated either before or after the electroweak symmetry breaking. In some of the decay channels the same analysis is performed with different data sets and the latest measurement is used for the comparison.

Table 7.25: Summary of the measurements of the tensor structure of Higgs boson couplings in different analysis channels

Decay mode	Experiment	Data set	Centre-of-mass energy	Integrated luminosity	Observable	Reference
$H \to ZZ^* \to 4\ell$	ATLAS	Run 1	7,8 TeV	4.5, 20.3 fb^{-1}	Kinematic of decay	[28]
		Run 2	13 TeV	36.1 fb^{-1}	Production rate information	[182]
					Cross section in STXS production bins	[182]
				139 fb^{-1}	Cross section in STXS production bins	[201]
$H \to WW^* \to e\nu\mu\nu$	ATLAS	Run 1	8 TeV	20.3 fb^{-1}	Kinematic of decay	[28]
$H \to \gamma\gamma$	ATLAS	Run 2	13 TeV	36.1 fb^{-1}	Differential cross sections	[205]
				139 fb^{-1}		[206]
$H \to \tau\tau$	ATLAS	Run 1	8 TeV	20 fb^{-1}	Kinematic of production	[207]
		Run 2	13 TeV	36.1 fb^{-1}		[208]
$H \to VV \to 4\ell$	CMS	Run 1	7,8 TeV	5.1, 19.7 fb^{-1}	Kinematic of production and decay	[27]
		Run 1, Run 2	7,8,13 TeV	5.1, 19.7, 38.6 fb^{-1}		[209]
				5.1, 19.7, 80.2 fb^{-1}		[210]
$H \to \tau\tau$	CMS	Run 2	13 TeV	35.9 fb^{-1}	Kinematic of production and decay	[211]

7.4.1 Tensor Structure Measurements in the Higgs Characterisation Framework

In the $H \to ZZ^* \to 4\ell$ decay channel, the tensor structure of the Higgs boson coupling to weak vector bosons has already been probed with the Run 1 proton-proton collision data recorded with the ATLAS detector at a centre-of-mass energy of 7 and 8 TeV corresponding to 4.5 fb^{-1} and 20.3 fb^{-1}, respectively.

In contrast to the production rate information used in the HC analysis from Section 7.2, only the kinematic properties of final state particles are used in the Run 1 study. Constraints on BSM coupling parameters are set by means of the first- and second-order Optimal Observables similar to those as in Section 8.2.2. These matrix element based observables combine the full information of the decay products into two variables and the shapes of their distributions are used in the fit as final discriminants. To suppress the large ZZ^* background, a BDT discriminant is used as an additional observable in all fits.

Admixtures of BSM contributions to the XVV vertex are described using the HC framework as in Section 7.2, but the constraints are set relative to the SM-like coupling parameter κ_{SM}. The CP-even and CP-odd BSM coupling parameters are defined as

$$\frac{\widetilde{\kappa}_{HVV}}{\kappa_{SM}} = \frac{v}{4 \cdot \Lambda} \cdot \frac{\kappa_{HVV}}{\kappa_{SM}} \quad \text{and} \quad \frac{\widetilde{\kappa}_{AVV}}{\kappa_{SM}} \cdot \tan(\alpha) = \frac{v}{4 \cdot \Lambda} \cdot \frac{\kappa_{AVV}}{\kappa_{SM}} \cdot \tan(\alpha), \tag{7.36}$$

respectively. When including the production rate information the κ_{XVV} and κ_{SM} have to be probed independently of each other, since the total cross section depends on both BSM

and SM coupling parameters. Therefore, the comparison with the results of this thesis is only possible for a specific configuration with the SM coupling parameters fixed to their SM values, i.e. $c_\alpha\kappa_{SM} = 1$ and $c_\alpha\kappa_{Hgg} = 1$. To increase the sensitivity, the results of the $H \to ZZ^* \to 4\ell$ Run 1 analysis are combined with those from the $H \to WW^* \to e\nu\mu\nu$ decay channel. This analysis is also based on kinematic properties of the final state particles, while using BDT output instead of Optimal Observables as final discriminant.

Table 7.26 summarises the observed and expected 95% CL intervals as well as the best-fit values of BSM coupling parameters from both measurements. The Run 1 results are expressed in terms of the BSM coupling parameters used in the HC analysis from Section 7.2. The shapes of kinematic distributions of final state particles in the Run 1 analysis are sensitive to the sign of the CP-odd BSM coupling parameter $s_\alpha\kappa_{AVV}$, as opposed to the Run 2 analysis. Both analyses are sensitive to the sign of the CP-even BSM coupling parameter $c_\alpha\kappa_{HVV}$ due to the interference of the SM and BSM CP-even coupling contributions. Comparing the 95% CL intervals, the Run 1 analysis is over an order of magnitude less sensitive. Both analyses are dominated by statistical uncertainties and the sensitivity, therefore, increases with the fourth square root of the integrated luminosity. In addition, in Run 2 the number of signal and background events increases by a factor of two due to an increased centre-of-mass energy. After taking into account both the increase of the integrated luminosity and the centre-of-mass energy, the Run 2 analysis is still two times more sensitive than the combined Run 1 $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ analysis, since the total production and decay rates are much more sensitive to the presence of BSM coupling parameters. However, these observables cannot be used for the direct test of CP-violation since they cannot distinguish between the CP-even and CP-odd BSM contributions.

The XVV tensor structure is also studied with the ATLAS detector in VBF Higgs boson production using the $H \to \tau\tau$ decay channel. The last published analysis is performed with proton-proton collision data corresponding to an integrated luminosity of 36.1 fb^{-1} collected at $\sqrt{s} = 13$ TeV in the years 2015 and 2016.

The analysis employs a similar strategy as the Run 1 analysis in the $H \to ZZ^* \to 4\ell$ decay channel based on the shape of distributions of an Optimal Observable calculated from matrix elements (see Section 8.2.2). However, the study employs only the first-order Optimal Observables defined by the four-momenta of the Higgs boson candidates and of the two jets produced in the VBF, i.e. using only the production vertex. The BSM dependence of the production rates is not considered. The background to the VBF signal is suppressed by a threshold requirement on BDT discriminants trained separately for each of the considered final states, with leptonic and hadronic τ decays.

Table 7.26: Comparison of the expected and observed confidence intervals at 95% CL for CP-even and CP-odd BSM coupling parameters $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ in the XVV interaction vertex, as obtained from the Run 1 analysis in the $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ decay channel and from the Run 2 analysis in the $H \to ZZ^* \to 4\ell$ decay channel

BSM coupling	Fit	95% CL interval		Observed
κ_{BSM}	configuration	Expected	Observed	best-fit value
$H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$				
Run 1, 4.5 fb^{-1} and 20.3 fb^{-1} at $\sqrt{s}$ = 7 and 8 TeV				
$c_\alpha\kappa_{HVV}$	–	$[-8.9, 77.9]$	$[-11.9, 10.2]$	$-$ 7.9
$s_\alpha\kappa_{AVV}$	–	$[-37.9, 37.4]$	$[-35.4, 13.5]$	-11.0
$H \to ZZ^* \to 4\ell$				
Run 2, 36.1 fb^{-1} at $\sqrt{s}$ = 13 TeV				
$c_\alpha\kappa_{HVV}$	$c_\alpha\kappa_{SM} = 1$	$[-2.9, 3.2]$	$[-0.8, 4.5]$	2.9
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM} = 1$	$[-3.5, 3.5]$	$[-5.2, 5.2]$	$\pm$ 2.9

The effective Lagrangian is defined as the SM Lagrangian with CP-odd dimension-six operators which involve the Higgs and the electroweak gauge fields. No BSM CP-even dimension-six operators are taken into account since the final CP-sensitive observable is shown not to be sensitive to such contributions. After the electroweak symmetry breaking the Lagrangian can be written as [212]

$$\begin{aligned}\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} &+ \tilde{g}_{HAA} H \widetilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{HAZ} H \widetilde{A}_{\mu\nu} Z^{\mu\nu} \\ &+ \tilde{g}_{HZZ} H \widetilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} H \widetilde{W}^{+}_{\mu\nu} W^{-\mu\nu},\end{aligned} \tag{7.37}$$

in the mass basis of the Higgs boson H, the photon A and the weak gauge bosons $W^\pm$ and Z. $V^{\mu\nu}$ and $\widetilde{V}_{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma}$ $(V = W^\pm, Z, A)$ denote the field strength and dual field strength tensors, respectively. The couplings $\tilde{g}_{HVV}$ can be expressed in terms of the two dimensionless couplings $\tilde{d}$ and $\tilde{d}_B$ [213, 214]

$$\begin{aligned}\tilde{g}_{HAA} &= \frac{g}{2m_W}\left(\tilde{d}\sin^2\theta_W + \tilde{d}_B\cos^2\theta_W\right), \qquad \tilde{g}_{HAZ}\frac{g}{2m_W}\sin 2\theta_W\left(\tilde{d} - \tilde{d}_B\right), \\ \tilde{g}_{HZZ} &= \frac{g}{2m_W}\left(\tilde{d}\cos^2\theta_W + \tilde{d}_B\sin^2\theta_W\right), \qquad \tilde{g}_{HWW}\frac{g}{m_W}\tilde{d},\end{aligned} \tag{7.38}$$

where g is the electroweak coupling and θ_W the weak mixing angle. The arbitrary choice

$\widetilde{d} = \widetilde{d}_B$ leads to the relations

$$\widetilde{g}_{HAA} = \widetilde{g}_{HZZ} = \frac{1}{2}\widetilde{g}_{HWW} = \frac{g}{2m_W}\widetilde{d} \quad \text{and} \quad \widetilde{g}_{HAZ} = 0. \tag{7.39}$$

With this assumption, the strength of CP-violation in the VBF production is described by a single parameter $\widetilde{d}$. The CP-odd BSM coupling parameter $\widetilde{d}$ can be translated into the CP-odd parameter of the HC model via the relation

$$\widetilde{d} = \frac{v \cdot \kappa_{AVV}}{4 \cdot \Lambda \cdot \kappa_{SM}} \tan(\alpha). \tag{7.40}$$

With the current available amount of data, no constraints can be set on $\widetilde{d}$ at 95% CL with the observed data in the $H \to \tau\tau$ channel. Therefore, the 68% CL intervals are compared to those obtained for the CP-odd BSM coupling parameter $c_\alpha \kappa_{AVV}$ in the HC analysis with value of the SM-like coupling parameter fixed to one. The respective observed and expected confidence intervals at 68% CL intervals are summarised in Table 7.27 for both analysis. The results of the $H \to \tau\tau$ analysis are expressed in terms of the parameters in the HC model. Although the $H \to \tau\tau$ analysis uses only shape information from kinematic distributions in the VBF Higgs boson production mode, while the HC analysis employs the production rates in all production modes, the sensitivity of both measurements is of the same order of magnitude. The largest sensitivity of the HC analysis is reached in the VBF production mode. The number of selected VBF signal events is however about ten times larger in the $H \to \tau\tau$ decay channel, while the signal-to-background ratio remains similar, such that a better constraint at 68% CL is achieved in the $H \to \tau\tau$ decay channel. With the $H \to \tau\tau$ analysis, it is also possible to distinguish the sign of the CP-odd BSM coupling, while no confidence intervals at 95% CL were obtained.

Several measurements were performed by the CMS collaboration to probe for anomalous XVV couplings in $H \to ZZ/Z\gamma^*/\gamma^*\gamma^* \to 4\ell$ decays. The studies were continuously updated using the increasing amount of available data. The latest published result is obtained with the proton-proton collision data collected during Run 1 and Run 2 at centre-of-mass energies of 7,8 and 13 TeV corresponding to 5.1 fb^{-1}, 19.7 fb^{-1}and 80.2 fb^{-1}, respectively.

The analysis employs the full kinematic information for the Higgs boson production and decay vertex without using the production rate information. The larger amount of data in Run 2 allows for a discrimination between the VBF and VH production modes. Therefore, the events collected in 2016 and 2017 are categorised into three categories: *VBF-tagged*, *VH-tagged* and *Untagged*. The data collected during Run 1 and in 2015 are assigned to the *Untagged* category. For each of the coupling parameters in each of the reconstructed categories a three-dimensional matrix element based discriminant

Table 7.27: Comparison of the expected and observed confidence intervals at 68% CL for CP-odd couplings $s_\alpha\kappa_{AVV}$ in the XVV interaction vertex from the Run 2 analysis in the $H\to\tau\tau$ decay channel and the $H\to ZZ^*\to 4\ell$ decay channel

BSM coupling	Fit	68% CL interval		Observed
κ_{BSM}	configuration	Expected	Observed	best-fit value
		$H\to\tau\tau$		
		Run 2, 36.1 fb^{-1} at $\sqrt{s}=$ 13 TeV		
$s_\alpha\kappa_{AVV}$	–	$[-0.6, 0.5]$	$[-1.5, 0.6]$	-0.2
		$H\to ZZ^*\to 4\ell$		
		Run 2, 36.1 fb^{-1} at $\sqrt{s}=$ 13 TeV		
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM}=1$	$[-2.3, 2.3]$	$[-4.0,-1.5]\cup[1.5, 4.0]$	± 2.9

is constructed (two dimensions for the signal process and one for the ZZ^* background suppression). The ratio of the observed and SM expected number of events is taken into account by means of signal strength parameters which are left free in the fit: μ_V is used for VBF and VH production and μ_F for all other production modes.

Possible BSM contributions to the XVV vertex are described with a generic parametrisation of the scattering amplitude called the anomalous couplings approach [215]. The general scattering amplitude including all possible tensor structures that are consistent with gauge and Lorentz invariance is

$$A(X_{J=0}\to VV)=\frac{1}{v}\left(a_1 m_V^2\epsilon_1^*\epsilon_2^* + a_2 f_{\mu\nu}^{*(1)} f^{*(2),\mu\nu} + a_3 f_{\mu\nu}^{*(1)}\widetilde{f}^{*(2),\mu\nu}\right), \tag{7.41}$$

describing the interactions of the Higgs boson (X) with a pair of gauge bosons ($VV = ZZ, WW, Z\gamma, \gamma\gamma$ and gg). The field strength tensor is defined as $f^{(i),\mu\nu}=\epsilon_i^\mu q_i^\nu - \epsilon_i^\nu q_i^\mu$ with the momentum vector q_i and the polarisation vector ϵ_i of the gauge boson. The corresponding conjugate is $\widetilde{f}^{(i),\mu\nu}={}^1\!/_2\epsilon^{\mu\nu\alpha\beta}f_{\alpha\beta}$. The momentum dependent form factors a_i ($i=1,2,3$) denote the coupling strengths. The SM coupling of the Higgs boson to weak bosons is described by the first term with vanishing couplings a_2 and a_3. SM loop processes and anomalous CP-even BSM couplings are introduced with the second term and CP-odd BSM couplings with the third. In general, the coupling strength a_i can be complex due to contributions of new light particles. Since no new physics has been observed so far, a_i are assumed to be real and constant. With this assumption, the couplings can be expressed in terms of the parameters of the HC model via the relations

$$\frac{a_2}{a_1}=\frac{v}{4\cdot\Lambda}\cdot\frac{\kappa_{HVV}}{\kappa_{SM}} \quad\text{and}\quad \frac{a_3}{a_1}=\frac{v}{4\cdot\Lambda}\cdot\frac{\kappa_{AVV}}{\kappa_{SM}}\cdot\tan(\alpha), \tag{7.42}$$

for the CP-even and CP-odd BSM coupling parameter, respectively.

The results of the CMS analysis expressed in terms of the parameters of the HC model are compared to the results of the HC analysis in Table 7.28. Since the CMS collaboration measures the ratios of coupling parameters, a translation to the HC model is only possible assuming the SM-like couplings in the HC analysis to be equal one. Despite the larger amount of data analysed in the CMS measurement, tighter constraints on BSM coupling parameters are set with the ATLAS HC analysis. The reason is that the HC analysis employs mainly the production and decay rate information, while for the measurement performed from the CMS collaboration the shapes of kinematic distributions defined by the properties of final state particles are employed.

Table 7.28: Comparison of the expected and observed confidence intervals at 95% CL for CP-even and CP-odd BSM couplings $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ in the XVV interaction vertex from the CMS analysis and the ATLAS Run 2 analysis in the $H \to ZZ^* \to 4\ell$ decay channel

BSM coupling	Fit	95% CL interval	
κ_{BSM}	configuration	Expected	Observed
		CMS $H \to 4\ell$	
		Run 1 and Run 2, 5.1 fb^{-1}*, 19.7 fb*$^{-1}$ *and 80.2 fb*$^{-1}$ *at* $\sqrt{s}$ *=7,8 and 13 TeV*	
$c_\alpha\kappa_{HVV}$	–	$[-3.9, 5.0]$	$[-1.9, 4.2]$
$s_\alpha\kappa_{AVV}$	–	$[-12.3, 12.3]$	$[-18.3, 13.0]$
		ATLAS $H \to ZZ^* \to 4\ell$	
		Run 2, 36.1 fb^{-1} *at* $\sqrt{s}$ *= 13 TeV*	
$c_\alpha\kappa_{HVV}$	$c_\alpha\kappa_{SM} = 1$	$[-2.9, 3.2]$	$[-0.8, 4.5]$
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM} = 1$	$[-3.5, 3.5]$	$[-5.2, 5.2]$

Anomalous XVV interactions are also studied by the CMS collaboration in the $H \to \tau\tau$ decay channel using proton-proton collision data collected in 2016 at a centre-of-mass energy of 13 TeV corresponding to 35.9 fb^{-1}.

The analysis employs the full kinematic information from the Higgs boson production and decay without using the production rate information. The events selected in the different final states ($e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$) are categorised into three event categories: the *0-jet* and *VBF* category, targeting the ggF and VBF production and remaining events collected in the *Boosted* category. As in the case of the $H \to 4\ell$ analysis, matrix element based discriminants are used. In each of the categories two discriminants are

used separating SM and BSM processes. As in the case of the CMS $H \to 4\ell$ analysis, the observed yield is absorbed by the free signal strength parameters (μ_V and μ_F) in the fit.

In the CMS $H \to \tau\tau$ analysis, possible BSM contributions are introduced via the anomalous couplings approach introduced above (Equation 7.41), which can be translated into the HC model using Equation 7.42.

With the available amount of data no constraints can be obtained at 95% CL in the $H \to \tau\tau$ channel for both the CP-even and CP-odd BSM coupling parameters. Therefore, the results are combined with the CMS $H \to 4\ell$ analysis described above. Table 7.29 shows the combined results translated into the ATLAS HC model as well as the results of the HC analysis. Since the CMS results are reported as ratios of coupling parameters, the SM-like coupling parameter in the HC analysis is set to one. Due to a larger amount of data, the combined measurement in the $H \to 4\ell$ and $H \to \tau\tau$ decay channels provides a similar sensitivity as the ATLAS HC analysis, even through it employs the information from the shapes of kinematic distributions.

Table 7.29: Comparison of the expected and observed confidence intervals at 95% CL for CP-even and CP-odd BSM couplings $c_\alpha\kappa_{HVV}$ and $s_\alpha\kappa_{AVV}$ in the XVV interaction vertex from the combined CMS analysis in the $H \to 4\ell$ and the $H \to \tau\tau$ decay channel and the ATLAS Run 2 analysis in the $H \to ZZ^* \to 4\ell$ decay channel

BSM coupling	Fit	95% CL interval	
κ_{BSM}	configuration	Expected	Observed
CMS $H \to 4\ell$ and $H \to \tau\tau$			
Run 1 and Run 2, 5.1 fb^{-1}, 19.7 fb^{-1} and 80.2 fb^{-1} at $\sqrt{s}$ =7,8 and 13 TeV			
and Run 2, 35.9 fb^{-1} at $\sqrt{s}$ =13 TeV			
$c_\alpha\kappa_{HVV}$	–	$[-1.8, 1.8]$	$[-0.9, 1.6]$
$s_\alpha\kappa_{AVV}$	–	$[-1.5, 1.5]$	$[-13.2, 5.0]$
ATLAS $H \to ZZ^* \to 4\ell$			
Run 2, 36.1 fb^{-1} at $\sqrt{s}$ = 13 TeV			
$c_\alpha\kappa_{HVV}$	$c_\alpha\kappa_{SM} = 1$	$[-2.9, 3.2]$	$[-0.8, 4.5]$
$s_\alpha\kappa_{AVV}$	$c_\alpha\kappa_{SM} = 1$	$[-3.5, 3.5]$	$[-5.2, 5.2]$

7.4.2 Tensor Structure Measurements within the Standard Model Effective Field Theory

The tensor structure of the Higgs boson interaction is also probed in $H \to \gamma\gamma$ decays using the framework of the standard model effective field theory (SMEFT). The latest study is performed with the full Run 2 data set of proton-proton collision data (139 fb^{-1}) collected with the ATLAS detector at a centre-of-mass energy of 13 TeV.

The study uses a similar approach as the interpretation of the $H \to ZZ^* \to 4\ell$ within the SMEFT (Section 7.3), but instead of employing the measured the cross sections in each bin of the STXS framework, several measured differential cross section are employed. The $H \to \gamma\gamma$ cross section is measured in bins of several kinematic variables: the transverse momentum $p_{\gamma\gamma}$ of the di-photon system, number of jets N_j in the final state, the invariant mass m_{jj} of the two leading jets, the transverse momentum p_T^{j1} of the leading jet and the angular distance $\Delta\phi_{jj}$ between the two leading jets. The cross section in each of these bins is parametrised as a function of BSM coupling parameters normalised to the corresponding SM prediction.

As in the study presented in this thesis Section 7.3), the $H \to \gamma\gamma$ analysis employs the SMEFT framework including the quadratic matrix element terms. Constraints on each individual BSM coupling parameter describing the Higgs boson interaction with the weak vector bosons or gluons are obtained under the assumption that all other BSM coupling parameters equal to zero. This allows for a direct comparison of the results. The effective top Yukawa interaction is not probed in $H \to \gamma\gamma$ decays.

The observed 95% CL intervals for both analysis are summarised in Table 7.30. Expected CL intervals are not shown, since they are not reported for the measurement in the $H \to \gamma\gamma$ decay channel. The measurement in $H \to \gamma\gamma$ decays is about one order of magnitude more sensitive for the Wilson coefficients describing the effective Higgs to gluon coupling (c_{HG} and $c_{H\widetilde{G}}$) compared to the interpretation performed in this thesis. This is due to a larger amount of reconstructed Higgs boson candidates in $H \to \gamma\gamma$ decays. For the other probed Wilson coefficients, the measurement in $H \to ZZ^* \to 4\ell$ decays is about four orders of magnitude less sensitive. This is on the one hand caused by the strong BSM dependence of the detector acceptance in the $H \to ZZ^* \to 4\ell$ decay channel which decreases the sensitivity and on the other hand a smaller number of reconstructed $H \to 4\ell$ Higgs boson candidate events.

Table 7.30: The expected and observed confidence intervals at 95% CL for the BSM coupling parameters of the SMEFT model for an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV from the analyses of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ decay channels

BSM coupling parameter	95% CL interval Observed	BSM coupling parameter	95% CL interval Observed
	$H \to \gamma\gamma$		
	Run 2, 139 fb^{-1} at $\sqrt{s} = 13$ TeV		
c_{HG}	$[-6.1, 4.7] \times 10^{-4}$	$c_{H\widetilde{G}}$	$[-1.5, 1.4] \times 10^{-3}$
c_{HW}	$[-8.3, 8.3] \times 10^{-4}$	$c_{H\widetilde{W}}$	$[-3.7, 3.7] \times 10^{-3}$
c_{HB}	$[-2.4, 2.4] \times 10^{-4}$	$c_{H\widetilde{B}}$	$[-1.2, 1.1] \times 10^{-3}$
c_{HWB}	$[-4.2, 4.2] \times 10^{-4}$	$c_{H\widetilde{W}B}$	$[-2.0, 2.0] \times 10^{-3}$
	$H \to ZZ^* \to 4\ell$		
	Run 2, 139 fb^{-1} at $\sqrt{s} = 13$ TeV		
c_{HG}	$[-83, 68] \times 10^{-4}$	$c_{H\widetilde{G}}$	$[-29, 29] \times 10^{-3}$
c_{HW}	$[-3.4, 2.1]$	$c_{H\widetilde{W}}$	$[-2.4, 2.4]$
c_{HB}	$[-0.62, 0.59]$	$c_{H\widetilde{B}}$	$[-0.56, 0.56]$
c_{HWB}	$[-1.06, 0.99]$	$c_{H\widetilde{W}B}$	$[-1.03, 1.03]$

The results of the STXS Stage-0 measurement in the ggF production bin obtained using the $H \to ZZ^* \to 4\ell$ decay channel with the 36.1 fb^{-1} of data are also interpreted within the effective field theory approach using the HC model.

For this purpose, the cross section in the ggF production bin is parameterised as a function of the CP-odd BSM coupling parameter $s_\alpha \kappa_{Agg}$. Since for this interaction vertex no additional assumptions are made and since $s_\alpha \kappa_{Agg}$ is the only parameter describing the effective Higgs boson coupling to gluons, it can be translated into the SMEFT coupling parameter $c_{H\widetilde{G}}$ via the relation

$$-\frac{\alpha_s^2}{8 \cdot \pi} \cdot s_\alpha \kappa_{Agg} = c_{H\widetilde{G}}, \tag{7.43}$$

where α_s corresponds to the strong coupling constant.

The expected and observed 95% CL intervals obtained with the 36.1 fb^{-1} and 139 fb^{-1} data sets are shown in Table 7.31. The results obtained with the HC model are expressed in terms of the coupling parameter $c_{H\widetilde{G}}$ in the SMEFT model. The interpretation of the measured STXS cross sections within the SMEFT (Section 7.3) is about one order of magnitude more sensitive than the HC analysis. Since both measurements are dominated by statistical uncertainties the exclusion limits improve with the fourth squared of the integrated luminosity. Taking this into account, the 95% CL intervals obtained with the 36.1 fb^{-1} data set improve by a factor of two, which is still about one order of magnitude less sensitive compared to the constraints obtained with the 139 fb^{-1}. This is because there is only one production bin (ggF) employed with 36.1 fb^{-1} of data, while seven ggF production bins are employed in the interpretation of the 139 fb^{-1} data set within the SMEFT. The larger amount of production bins defined by the range of the Higgs boson transverse momentum p_T^H increases the sensitivity to BSM physics in the Xgg interaction vertex.

Table 7.31: Comparison of the expected and observed confidence intervals at 95% CL for the CP-odd BSM couplings $c_{H\widetilde{G}}$ in the Xgg interaction vertex from the measurement in the $H \to ZZ^* \to 4\ell$ decay channel with the 36.1 fb^{-1} and 139 fb^{-1} data sets

BSM coupling	95% CL interval	
	Expected	Observed
Run 2, 36.1 fb^{-1} at $\sqrt{s} = 13$ TeV		
$c_{H\widetilde{G}}$	$[-0.41, 0.41]$	$[-0.56, 0.56]$
Run 2, 139 fb^{-1} at $\sqrt{s} = 13$ TeV		
$c_{H\widetilde{G}}$	$[-0.031, 0.031]$	$[-0.029, 0.029]$

8

Test of CP-Violation in Vector Boson Fusion Production

In the previous chapter, the tensor structure of the Higgs boson couplings is mainly probed by the event rates measured for different Higgs boson production modes. In an alternative approach, the tensor structure can also be tested in a rate-independent way, by relying only on the shapes of the distributions of kinematic variables of final state particles. Such observables are in general less sensitive to BSM couplings than the production rates. However, the shape observables can be chosen such that they are insensitive to CP-even BSM couplings providing a direct probe of the CP-odd couplings and, therefore, of CP-violation in the Higgs sector.

In this chapter, a sensitivity study for such a test for CP-violation at the XVV vertex in $H \to ZZ^* \to 4\ell$ decays for the full Run 2 data set in VBF Higgs boson production is performed.

8.1 Analysis Strategy

The BSM signal is modelled within the Higgs Characterisation (HC) framework and the analysis is performed with the same CP-odd BSM signal and the same background samples (see Section 4.2.2) that were also used for the EFT-based HC analysis with the 36.1 fb^{-1}data set (Section 7.2). The signal events were generated with the MadGraph5_aMC@NLO generator. The sensitivity of the CP-violation measurement is evaluated for the integrated luminosity of 139 fb^{-1} of the full Run 2 data set.

The Higgs boson candidates are selected as for inclusive analysis using the 36.1 fb^{-1} data set described in Section 4.3. The 4ℓ mass window is adjusted to 115 GeV $< m_{4\ell} <$ 130 GeV to agree with the one used in the most recent coupling studies with 139 fb^{-1} (Section 7.3). The VBF Higgs boson production mode is targeted by selecting Higgs boson candidates which fall into the VBF-enriched category requiring at least two jets in the final state with an invariant mass of the two leading jets larger than 120 GeV.

The signal in this event category is modelled with simulated BSM VBF+VH-Had samples taking into account the contaminations from the background and the ggF Higgs boson production. In order to apply the best-prediction-scaling on the VBF+VH-Had samples simulated at LO in QCD (see Section 7.2.2.2), the samples are combined with the corresponding VH-Lep samples and compared to the sum of the nominal VBF and WH and ZH POWHEG samples. The signal event yield in the VBF-enriched category is then scaled to the best-prediction according to Equation 7.12, with a resulting scale factor of

$$\mathrm{SF}(\mathrm{VBF}+VH) = 1.044410. \tag{8.1}$$

The observables sensitive to the CP-odd XVV contributions are defined using only the production vertex information (see Section 8.2). Thus, it can be assumed that there is no impact of BSM contributions to the contaminating ggF production mode.

The distributions of CP-sensitive observables are modelled with the morphing method as described in Section 7.2.2.1 with the difference that each bin of the observable distribution is modelled instead of just a single event yield value in the reconstructed category. The same VBF+VH-Had and VH-Lep input samples are used as in the case of the one-dimensional scan of the Higgs Characterisation parameter $s_\alpha\kappa_{AVV}$ (see Table 7.4).

The constraints on possible CP-violating contributions are obtained from the binned likelihood fit with a likelihood function $\mathcal{L}\,(\mathrm{data} \mid \boldsymbol{\kappa})$ constructed out of Poisson distributions P for the observed distribution of the observable T in the VBF-enriched reconstructed category given the signal and background expectations $s\,(\boldsymbol{\kappa})$ and b, respectively,

$$\mathcal{L}\,(\mathrm{data} \mid \boldsymbol{\kappa}) = P\,(T \mid s\,(\boldsymbol{\kappa}) + b)\,. \tag{8.2}$$

The likelihood function depends on the parameter of interest $\boldsymbol{\kappa} = \{c_\alpha\kappa_{SM}, s_\alpha\kappa_{AVV}\}$. The parameter $c_\alpha\kappa_{SM}$ is related to the SM-like CP-even interaction terms, while $s_\alpha\kappa_{AVV}$ represents the CP-odd ones. The sensitivity study is performed without taking systematic uncertainties into account. Therefore, the nuisance parameters $\boldsymbol{\theta}$ are not included into the likelihood function. The exclusion limits on $\boldsymbol{\kappa}$ are obtained from the $\boldsymbol{\kappa}-$scan of the test statistic q based on the ratio of likelihoods [193],

$$q = -2\ln\frac{\mathcal{L}\,(\boldsymbol{\kappa})}{\mathcal{L}\,(\hat{\boldsymbol{\kappa}})} = -2\ln(\lambda)\,. \tag{8.3}$$

8.2 CP-Violating Observables

As mentioned previously, the production rates are sensitive to both CP-even and CP-odd BSM contributions. Such variables, therefore, cannot be employed as an unambiguous probe of CP-odd admixtures to the SM Higgs boson coupling. On the other hand, the interference between the SM and BSM CP-odd interactions introduce modifications

of certain kinematic distributions in the final state, while keeping the production rate constant. The distributions of the corresponding CP-odd observables $\mathcal{O}_{\mathrm{CP}}$, therefore, provide a direct test for CP-violation.

In order to remove any sensitivity on CP-even BSM interactions, only the shape and no normalisation of a given $\mathcal{O}_{\mathrm{CP}}$ distribution is considered. In case of the CP-invariance, the mean value of the CP-odd observable vanishes, $\langle \mathcal{O}_{\mathrm{CP}} \rangle = 0$, while a non-vanishing mean value would be a clear sign of CP-violation.

Two CP-odd observables are discussed in the following: the azimuthal angle between the two tagging jets in the final state ($\Delta\phi_{jj}^{\mathrm{sign}}$) and the first-order Optimal Observable for VBF production ($\mathcal{OO}_{1,jjH}$). The latter is based on a matrix element calculation which combines the information of jet and Higgs boson four-momenta, thus providing a higher sensitivity compared to the former one.

8.2.1 Azimuthal Angle Between Tagging Jets

A rather straightforward and model independent variable sensitive to CP-violation is the signed difference of the azimuthal angle of the two leading jets $\Delta\phi_{jj}^{\mathrm{sign}}$ [216, 217], defined on particle level as,

$$\epsilon_{\mu\nu\rho\sigma} b_+^\mu p_+^\nu b_-^\rho p_-^\sigma = 2p_{T,+}p_{T,-}\sin(\phi_+ - \phi_-) = 2p_{T,+}p_{T,-}\sin\left(\Delta\phi_{qq}^{\mathrm{sign}}\right), \tag{8.4}$$

where b_+^μ and b_-^μ correspond to the normalised four momenta of the two proton beams and p_+^μ and p_-^μ to the four momenta of the two leading quarks, with $p_+(p_-)$ pointing into the same detector hemisphere as $b_+^\mu(b_-^\mu)$. Since this definition is invariant under the interchange $(b_+, p_+) \leftrightarrow (b_-, p_-)$, $\Delta\phi_{qq}^{\mathrm{sign}}$ is a CP-odd observable.

The corresponding $\Delta\phi_{jj}^{\mathrm{sign}}$ observable can be calculated from the reconstructed final state particles by rewriting Equation 8.4:

$$\Delta\phi_{jj}^{\mathrm{sign}} = \begin{cases} \phi_{j1} - \phi_{j2} & \text{for } y_{j1} > y_{j2} \\ \phi_{j2} - \phi_{j1} & \text{for } y_{j2} > y_{j1} \end{cases}, \tag{8.5}$$

where $j1$ and $j2$ are the leading (highest-p_T) and the subleading (second-highest p_T) jet, respectively; y_j is the jet rapidity and ϕ_j the jet azimuthal angle. The distributions of the CP-odd observable $\Delta\phi_{jj}^{\mathrm{sign}}$ in the VBF-enriched reconstructed event category are shown in Figure 8.1 for the SM hypothesis and for CP-odd BSM admixtures to the SM. The expected distributions for the SM hypothesis is symmetric around zero, i.e. the mean value is vanishing $\langle \mathcal{O}_{\mathrm{CP}} \rangle = 0$. In contrast to that, the $\Delta\phi_{jj}^{\mathrm{sign}}$ distribution is asymmetric and sensitive to the sign of the CP-odd BSM coupling parameters. Positive (negative) values of the CP-odd BSM coupling parameters induce shifts with $\Delta\phi_{jj}^{\mathrm{sign}} > 0$ ($\Delta\phi_{jj}^{\mathrm{sign}} < 0$).

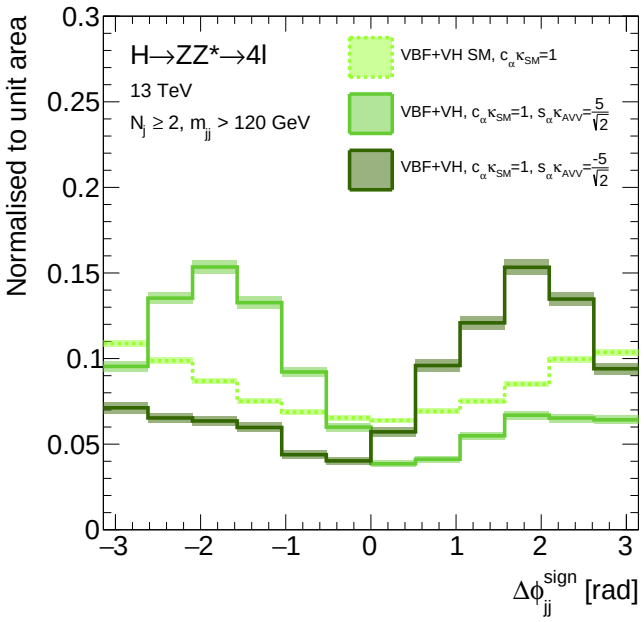


Figure 8.1: The expected distributions of the CP-odd observable $\Delta\phi_{jj}^{\text{sign}}$ in the VBF-enriched reconstructed category ($N_j \geq 2, m_{jj} > 120$ GeV) for the SM signal alone and with CP-odd BSM admixture with positive and negative CP-odd couplings.

In order to set constraints on the CP-violating contributions, either the observed mean value of the $\Delta\phi_{jj}^{\text{sign}}$ distribution is compared to data or a maximum likelihood fit to the entire $\Delta\phi_{jj}^{\text{sign}}$ distribution is performed. While the first approach is most model-independent, the latter approach is used in this thesis since the full shape information provides a better signal sensitivity.

8.2.2 First-Order Optimal Observable for VBF Production

Another CP-odd observable studied in this thesis is defined based on the concept of Optimal Observables [218–222]. Assuming momentum conservation in the transverse plane of the proton-proton collision and neglecting the jet mass, the final state with the Higgs boson and two tagging jets can be fully described by seven phase space variables. The concept of the Optimal Observable method combines in a well defined optimum way the full information of this multi-dimensional phase space into one single variable, the Optimal Observable ($\mathcal{OO}$), which is calculated from the leading-order matrix elements for the VBF production.

The leading-order matrix element $\mathcal{M}_{\text{Mix}}$ for the VBF production in the presence of CP-odd BSM contributions can be written as the sum of the CP-even SM matrix element $\mathcal{M}_{\text{SM}}$ and the CP-odd contribution $\mathcal{M}_{\text{CP-odd}}$,

$$\mathcal{M}_{\text{Mix}} = \mathcal{M}_{\text{SM}} + \frac{v}{4 \cdot \Lambda} \cdot \tan(\alpha) \cdot \frac{\kappa_{AVV}}{\kappa_{SM}} \cdot \mathcal{M}_{\text{CP-odd}} = \mathcal{M}_{\text{SM}} + \tilde{d} \cdot \mathcal{M}_{\text{CP-odd}}, \tag{8.6}$$

which leads to three components of the square of the matrix element:

$$|\mathcal{M}_{\text{Mix}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \widetilde{d}\cdot 2\mathfrak{Re}\,(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}}) + \widetilde{d}^2\cdot|\mathcal{M}_{\text{CP-odd}}|^2\,. \tag{8.7}$$

The two terms $|\mathcal{M}_{\text{SM}}|^2$ and $\widetilde{d}^2\cdot|\mathcal{M}_{\text{CP-odd}}|^2$ transform CP-even, while the interference term $\widetilde{d}\cdot 2\mathfrak{Re}\,(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}})$ is CP-odd and, therefore, directly sensitive to CP-violation. The integral of this term over a CP-symmetric part of the phase space vanishes yielding to no contribution to the total cross section. The SM cross section is only modified by the third term in Equation 8.7, which increases it proportional to $\widetilde{d}^2$.

The second and third terms in Equation 8.7 are used to compute the Optimal Observables. In general, any value of $\widetilde{d}$ can be used for the construction. For simplification, the value $\widetilde{d}=1$ is used, corresponding to a parameter configuration in the HC model of $\Lambda = 1$ TeV, $v = 246.22$ GeV, $\kappa_{SM} = \sqrt{2}$ and $s_\alpha\kappa_{AVV} = {}^{22.975}/_{\sqrt{2}}$. The first-order Optimal Observable is defined as

$$\mathcal{OO}_{1,jjH} = \frac{|\mathcal{M}_{\text{Mix}}|^2 - |\mathcal{M}_{\text{SM}}|^2 - |\mathcal{M}_{\text{CP-odd}}|^2}{|\mathcal{M}_{\text{SM}}|^2}, \tag{8.8}$$

while the second-order Optimal Observable is

$$\mathcal{OO}_{2,jjH} = \frac{|\mathcal{M}_{\text{CP-odd}}|^2}{|\mathcal{M}_{\text{SM}}|^2}. \tag{8.9}$$

In the case of CP-invariance, the distribution of the first-order Optimal Observable $\mathcal{OO}_{1,jjH}$ has mean zero, while the mean is shifted to positive (negative) values in the case of CP-violation with the positive (negative) sign of the CP-odd BSM coupling parameter $s_\alpha\kappa_{AVV}$. The second-order Optimal Observable $\mathcal{OO}_{2,jjH}$ affects the total cross section and is, therefore, not used in the present analysis in order to avoid sensitivity of the analysis to CP-even BSM contributions.

The matrix elements for the $\mathcal{OO}_{1,jjH}$ are calculated with MadGraph5_aMC@NLO using the NNPDF23lo PDF set as for the generation of the HC BSM signal samples (Section 4.2.2.2). For each reconstructed event, the matrix element is calculated from the four-momenta of the Higgs boson (from the four-momenta of the four leptons) and of the two leading jets. Since in total 60 Feynman diagrams contribute to the VBF process and because the flavour of the incoming and outgoing partons cannot be determined experimentally, the observable is summed over all possible flavour configurations $ij \to klH$ weighted by the parton distribution functions $f_i(x)$. The SM matrix element is then given by

$$|\mathcal{M}_{\text{SM}}|^2 = \sum_{i,j,k,l} f_i(x_1) f_j(x_2)\,|\mathcal{M}_{\text{SM}}|^2\,(ij \to klH)\,, \tag{8.10}$$

and the interference term as

$$2\mathfrak{Re}\,(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}}) = \sum_{i,j,k,l} f_i(x_1) f_j(x_2) 2\mathfrak{Re}\,(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}})\,(ij \to klH)\,, \tag{8.11}$$

with the Bjorken momentum fractions x_1 and x_2 of the incoming partons. Due to energy-momentum conservation, the Bjorken values can be reconstructed from the reconstructed final state observables,

$$x_1^{reco} = \frac{m_{jjH}}{\sqrt{s}} e^{y_{jjH}} \quad \text{and} \quad x_2^{reco} = \frac{m_{jjH}}{\sqrt{s}} e^{-y_{jjH}}, \tag{8.12}$$

where m_{jjH} is the mass and y_{jjH} the rapidity of the vectorial sum of the four-momenta of the Higgs boson and the two tagging jets. Figure 8.2 shows the distribution of $\mathcal{OO}_{1,jjH}$ for the SM hypothesis and for the CP-odd BSM admixtures to the SM in the VBF-enriched event category from simulation. As for $\Delta\phi_{jj}^{\text{sign}}$, the distribution of the first-order Optimal Observable for the SM CP-even hypothesis is symmetric around zero, while the CP-odd observable is asymmetric and distinguishes between positive and negative CP-odd BSM contributions. Positive coupling parameters induce a shift of the distribution to positive values, while negative coupling parameters prefer negative values.

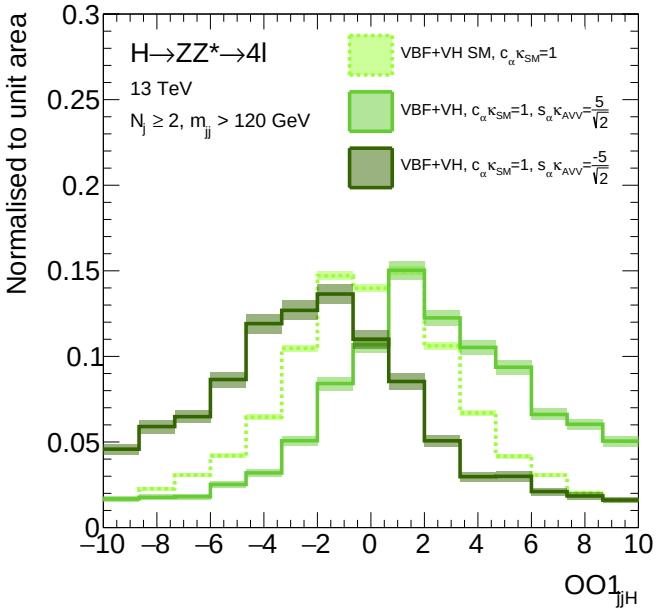


Figure 8.2: The expected distributions of the CP-odd observable $\mathcal{OO}_{1,jjH}$ in the VBF-enriched reconstructed category ($N_j \geq 2, m_{jj} > 120$ GeV) for the SM signal alone and with CP-odd BSM admixture with positive and negative CP-odd couplings.

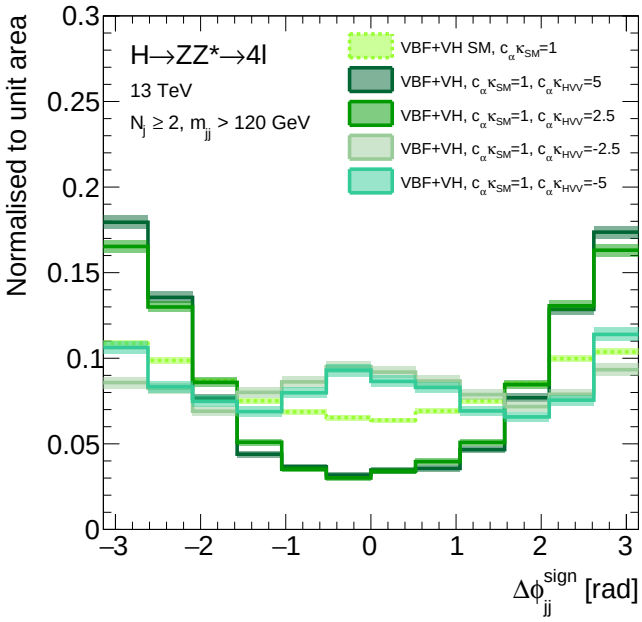
8.2.3 Validation of the CP-odd Observables

In order to verify that the CP-odd observables introduced above are not affected by CP-even BSM contributions to the XVV interaction vertex, the limits on the CP-odd $s_\alpha \kappa_{AVV}$ coupling parameter are evaluated for a CP-even BSM contribution to the interaction in the Monte Carlo Simulation. The signal samples used are summarised in Table 8.1.

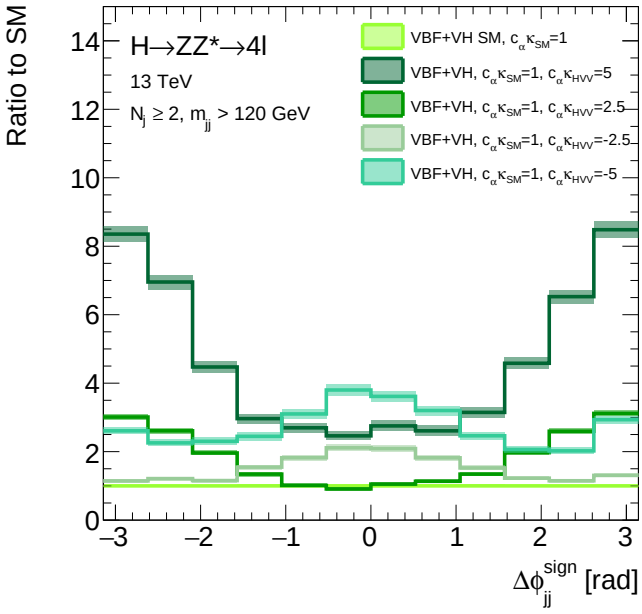
The distributions of the CP-odd observables $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$ for the SM and different BSM CP-even hypotheses are shown in Figure 8.3(a) and Figure 8.4(a). For both observables, the distributions from the different CP-even BSM hypotheses can be distinguished from each other and also from the SM hypothesis. This dependence of the shape of the CP-odd observables on the CP-even BSM coupling parameter is clearly visible in the ratio to the SM (Figure 8.3(b) and Figure 8.4(b)). For the $\Delta\phi_{jj}^{\text{sign}}$ observable, the ratio on average increases for increasing absolute values of BSM coupling parameters. In case of positive (negative) values of CP-even BSM coupling parameters, the ratio increases (decreases) with increasing absolute value of $\Delta\phi_{jj}^{\text{sign}}$. In case of the $\mathcal{OO}_{1,jjH}$ observable, the ratio to the SM also increases with increasing absolute values of CP-even BSM couplings. Different than for $\Delta\phi_{jj}^{\text{sign}}$, the ratio increases with increasing $\mathcal{OO}_{1,jjH}$ values for both the positive and negative BSM coupling parameters. Despite of this dependence, the measurement of the CP-invariance is not expected to be significantly biased by CP-even BSM contributions, because all distributions are symmetric around zero with a vanishing mean value of $\langle \mathcal{O}_{\text{CP}} \rangle = 0$.

Table 8.1: The VBF+VH Monte Carlo samples with CP-even BSM admixtures to the SM, used for the validation of the CP-odd observables $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$. The corresponding cross section times branching ratio values are also shown. In the sample name, P (N) denotes that positive (negative) coupling parameters were used for the generation.

Sample	c_α	κ_{SM}	κ_{HVV}	$\sigma \cdot \mathcal{B}$ [fb]
Mixture_1P	1	1	5	2.76
Mixture_1N	1	1	−2.5	1.31
Mixture_2P	1	1	2.5	0.92
Mixture_2N	1	1	−5	2.57

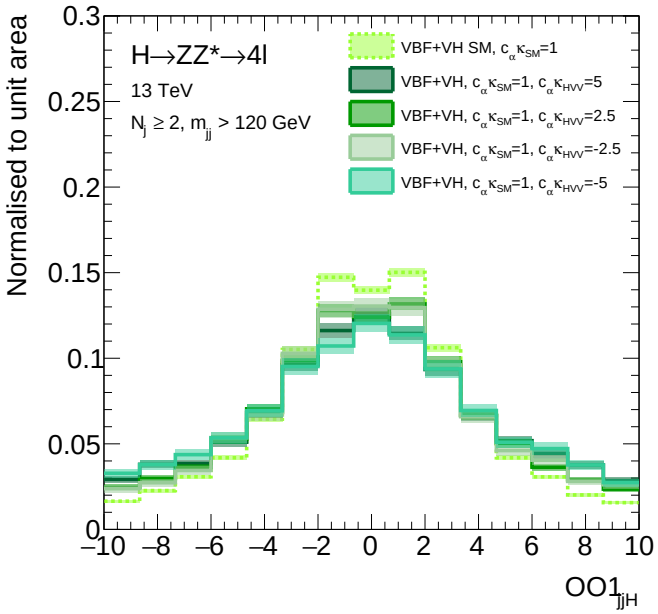


(a)

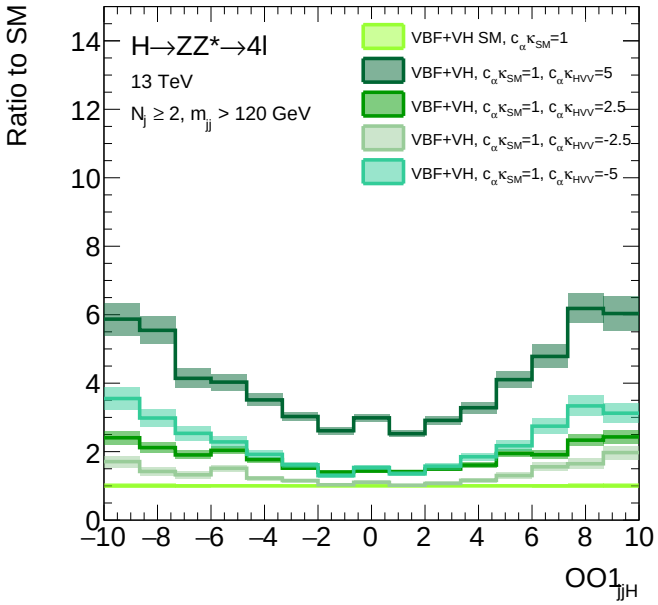


(b)

Figure 8.3: The (a) expected distributions of the CP-odd observable $\Delta\phi_{jj}^{\text{sign}}$ as well as (b) their ratio to the SM in the VBF-enriched reconstructed category ($N_j \geq 2, m_{jj} > 120$ GeV) in case of the SM hypothesis and assuming CP-even BSM admixtures to the SM.



(a)



(b)

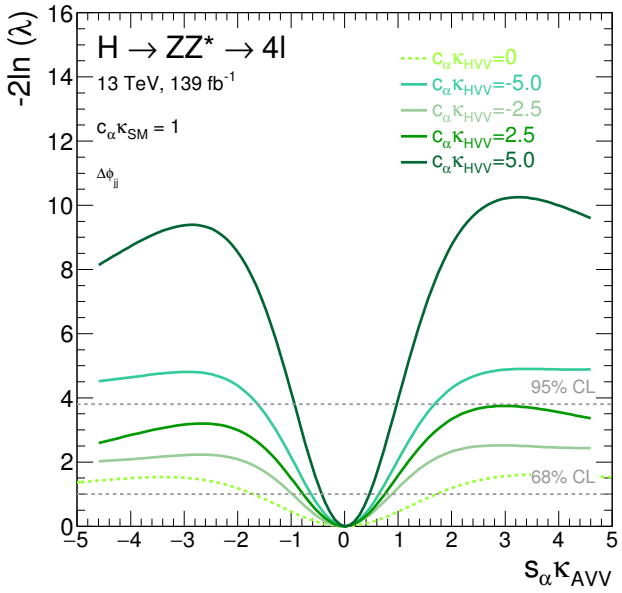
Figure 8.4: The (a) expected distributions of the CP-odd observable $\mathcal{OO}_{1,jjH}$ as well as (b) their ratio to the SM in the VBF-enriched reconstructed category ($N_j \geq 2, m_{jj} > 120$ GeV) in case of the SM hypothesis and assuming CP-even BSM admixtures to the SM.

In addition to the comparison of the shape of the distributions, a maximum likelihood fit of the full distribution is performed to directly evaluate the impact on the CP-odd BSM coupling $s_\alpha \kappa_{AVV}$. For this purpose, pseudo-data are introduced corresponding to the hypothesis with CP-even BSM admixtures and a scan of the test statistics q is performed over a range of the CP-odd BSM coupling parameter values $s_\alpha \kappa_{AVV}$ using this pseudo-data.

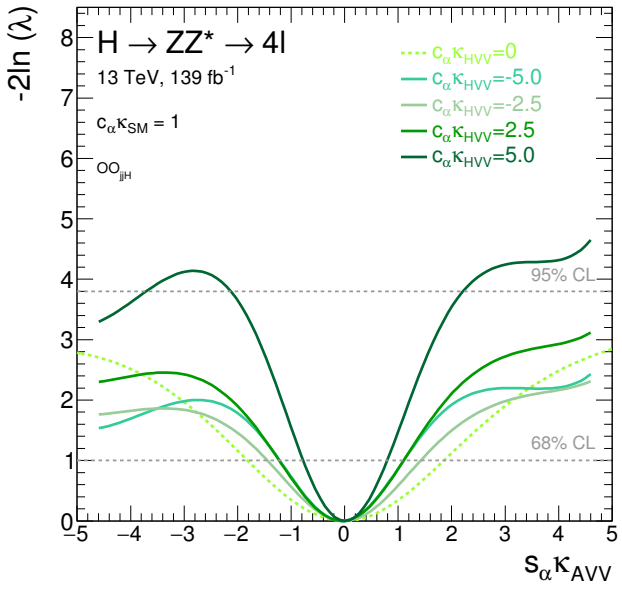
The expected values of the test statistics q of the scans are shown in Figure 8.5 for the SM hypothesis as well as for the four different CP-even BSM admixtures. The expected best-fit results and the 68% and 95% CL are summarised in Table 8.2. Within the statistical precision, the best-fit values obtained with all tested CP-even BSM hypotheses is zero for both CP-odd observables. Therefore, CP-even BSM contribution do not bias the measurement of the CP-odd coupling parameter value, verifying the eligibility of the $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$ observables for the study of the CP-invariance in the XVV vertex.

Table 8.2: The expected 68% and 95% CL intervals and best-fit values from one-dimensional likelihood scans of the test statistics q over the CP-odd BSM coupling parameter for the SM and various CP-even BSM hypotheses. The scans are performed with pseudo-data assuming an integrated luminosity of 139 fb^{-1} at a centre-of-mass energy of 13 TeV. The value of the SM-like coupling parameter $c_\alpha \kappa_{SM}$ is set to one in all scans.

BSM coupling	Confidence interval (CL)		Best-fit
κ_{BSM}	68%	95%	$c_\alpha \widehat{\kappa}_{\text{AVV}}$
Observable: $\Delta\phi_{jj}^{sign}$			
SM $c_\alpha \kappa_{HVV} = 0$	$[-1.69, 1.69]$	$[-, -]$	0.00
BSM $c_\alpha \kappa_{HVV} = -5$	$[-0.62, 0.64]$	$[-1.65, 1.70]$	0.00
BSM $c_\alpha \kappa_{HVV} = -2.5$	$[-0.97, 0.88]$	$[-, -]$	-0.04
BSM $c_\alpha \kappa_{HVV} = 2.5$	$[-0.77, 0.74]$	$[-, -]$	-0.01
BSM $c_\alpha \kappa_{HVV} = 5$	$[-0.42, 0.46]$	$[-0.94, 0.98]$	0.02
Observable: $\mathcal{OO}_{1,jjH}$			
SM $c_\alpha \kappa_{HVV} = 0$	$[-1.80, 1.84]$	$[-, -]$	0.00
BSM $c_\alpha \kappa_{HVV} = -5$	$[-1.20, 1.09]$	$[-, -]$	-0.05
BSM $c_\alpha \kappa_{HVV} = -2.5$	$[-1.43, 1.41]$	$[-, -]$	-0.01
BSM $c_\alpha \kappa_{HVV} = 2.5$	$[-1.22, 1.09]$	$[-, -]$	-0.05
BSM $c_\alpha \kappa_{HVV} = 5$	$[-0.77, 0.79]$	$[-2.17, 2.25]$	0.00



(a)



(b)

Figure 8.5: Expected values of test statistic $q = -2\ln(\lambda)$ in dependence on the CP-odd coupling parameter $s_\alpha \kappa_{AVV}$, for (a) the CP-odd observables $\Delta\phi_{jj}^{\text{sign}}$ and (b) $\mathcal{OO}_{1,jjH}$ assuming a signal with CP-even BSM contributions in the pseudo-data at an integrated luminosity of 139 fb^{-1} at a centre-of-mass energy of 13 TeV. The horizontal lines indicate the 68% and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.

Since the observables $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$ are symmetric around zero in case of CP-even BSM contributions, it is expected that also the exclusion limits would be symmetric. This is not fully the case in the presented study, due to fluctuations caused by the small number of expected events in the VBF-enriched category.

The exclusion limits on $s_\alpha\kappa_{AVV}$ obtained for different CP-even BSM hypotheses are more stringent compared to those for the SM hypothesis. The limits are increasingly tighter for the increasing ratio of the distributions from SM and CP-even BSM hypotheses (see Figure 8.3(b) and Figure 8.4(b)). For both observables, the largest ratio is expected for $s_\alpha\kappa_{AVV} = 5$ leading to the tightest confidence intervals of $[-0.77, 0.79]$ and $[-2.17, 2.25]$ at 95% CL compared to the SM limits of $[-1.69, 1.69]$ and $[-1.80, 1.84]$ for the observables $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$, respectively. Since in the case of the $\Delta\phi_{jj}^{\text{sign}}$ observable, the ratio to the SM is larger as for $\mathcal{OO}_{1,jjH}$ the corresponding difference in exclusion limits is also larger. Thus, the presence of CP-even BSM admixtures to the SM XVV vertex lead to tighter constraints on the CP-odd BSM parameters.

A more model independent measurement can be provided by a fit of the expected mean value of the observable to data or by using a one-bin asymmetry value as the final discriminant in the fit instead of the full distribution.

8.3 Expected Sensitivity for Full Run 2

The expected number of events after the inclusive analysis selection of Higgs boson candidates in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV and after the classification into the reconstructed VBF-enriched category are shown in Table 8.3 for an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV . The numbers differ slightly from those quoted in Section 7.3, since they are obtained from simulated samples used in the analysis of the 36.1 fb^{-1} data set. In addition, the inclusive event selection for the 139 fb^{-1} data set has been slightly changed compared to 36.1 fb^{-1} as described in Section 4.3.

The observables introduced for the test of CP-invariance are probing the VBF production vertex. However, since the BSM HC Monte Carlo samples include in addition to the VBF also VH-Had production mode, the VBF+VH-Had samples are combined with the VH-Lep samples to build the VBF+VH signal model. Processes with no XVV vertex in the production, such as the ggF, ttH and bbH production modes are taken into account as background processes, since this contributions are not modified by CP-odd BSM couplings. In addition, the SM ZZ^* production is taken into account as the largest background contributions, while minor contributions from other backgrounds are neglected. The expected distributions of the CP-odd observables $\Delta\phi_{jj}^{\text{sign}}$ and $\mathcal{OO}_{1,jjH}$ and back-

ground processes are shown in Figure 8.6 and Figure 8.6, respectively, assuming the SM signal hypothesis. All distributions are symmetric around zero. The $\Delta\phi_{jj}^{\text{sign}}$ distributions of the background processes have a similar parabolic shape as the distribution from the SM Higgs boson signal and in case of the $\mathcal{OO}_{1,jjH}$ observable all of the background distributions peak at zero while for the VBF+VH signal a double peak structure is expected.

Table 8.3: Expected number of events at an integrated luminosity of 139 fb^{-1} and $\sqrt{s} = 13$ TeV after the $H \to ZZ^* \to 4\ell$ event selection in the mass range of 115 GeV $< m_{4\ell} <$ 130 GeV, as well as after the categorisation in the reconstructed VBF-enriched category, assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV. Only statistical uncertainties are quoted.

Event selection stage	VBF+VH	ggF	ttH+bbH	ZZ^*	Total
115 GeV $< m_{4\ell} <$ 130 GeV	24.53 $\pm$ 0.05	186.9 $\pm$ 0.5	3.76 $\pm$ 0.05	100.4 $\pm$ 0.5	315.5 $\pm$ 0.7
VBF-enriched	10.10 $\pm$ 0.03	16.83 $\pm$ 0.12	0.507 $\pm$ 0.015	5.98 $\pm$ 0.11	33.41 $\pm$ 0.16

In order to estimate the sensitivity to CP-violation using the full Run 2 data set the distribution of an CP-sensitive observable is fitted to the expected data using a binned maximum likelihood fit. Only the shape and no rate information is employed. Two hypotheses are tested for the expected data: the SM hypothesis and a BSM hypothesis with the CP-odd BSM coupling parameter set to $s_\alpha \kappa_{AVV} = 0.43$. To evaluate the influence of the various background processes on the exclusion limits, the ggF, the ttH+bbH and the ZZ^* background contributions are cumulatively taken into account.

The expected 68% CL intervals on the BSM parameter $s_\alpha \kappa_{AVV}$ as well as the best-fit values are shown in Table 8.4 for both tested hypotheses. The corresponding distributions of the test statistics are shown in Figure 8.8. As expected stronger exclusion limits are obtained with the $\mathcal{OO}_{1,jjH}$ observable, since this observable incorporates the full information on the kinematic properties of final state particles, while the $\Delta\phi_{jj}^{\text{sign}}$ observable includes only the jet information.

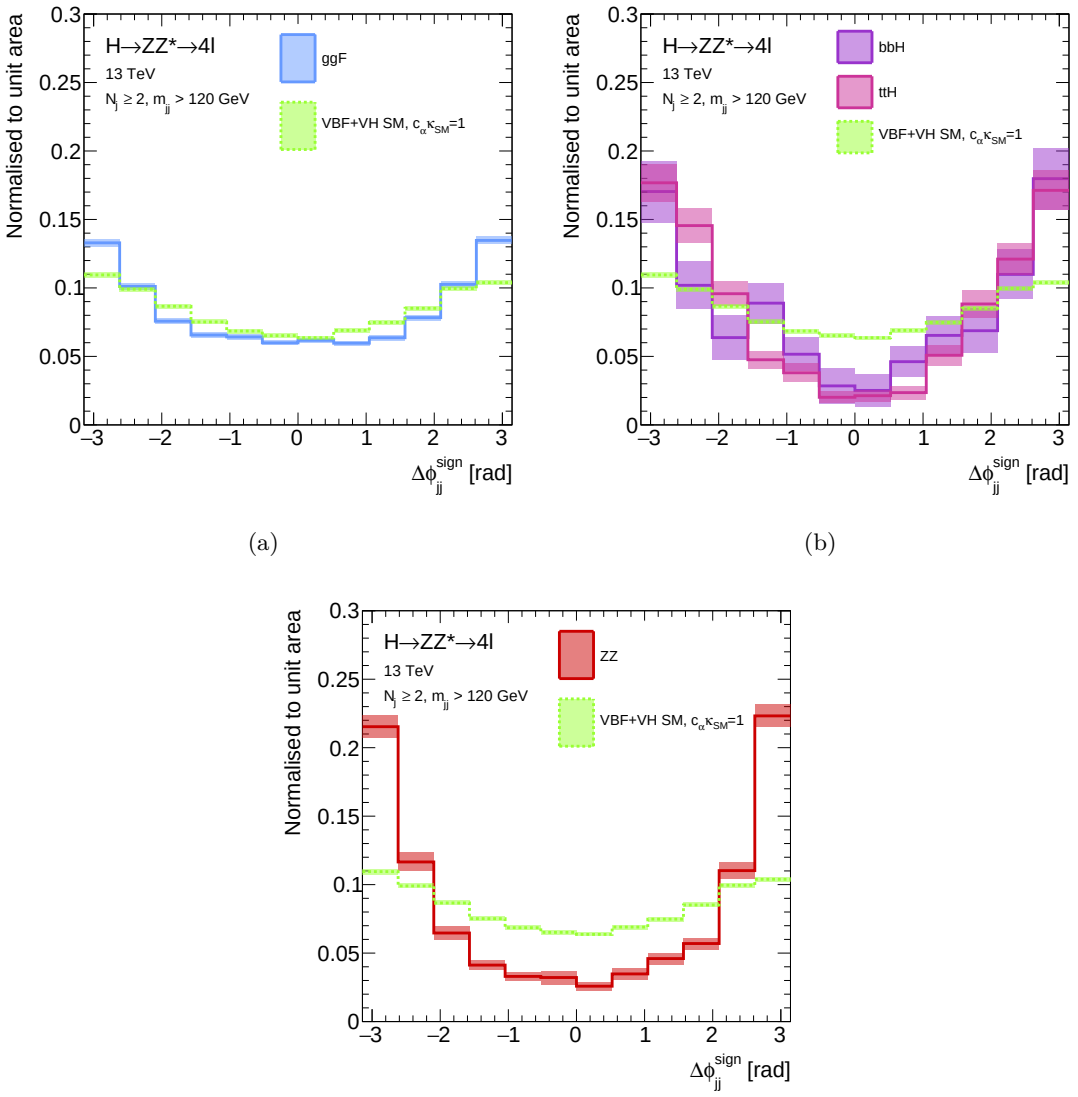
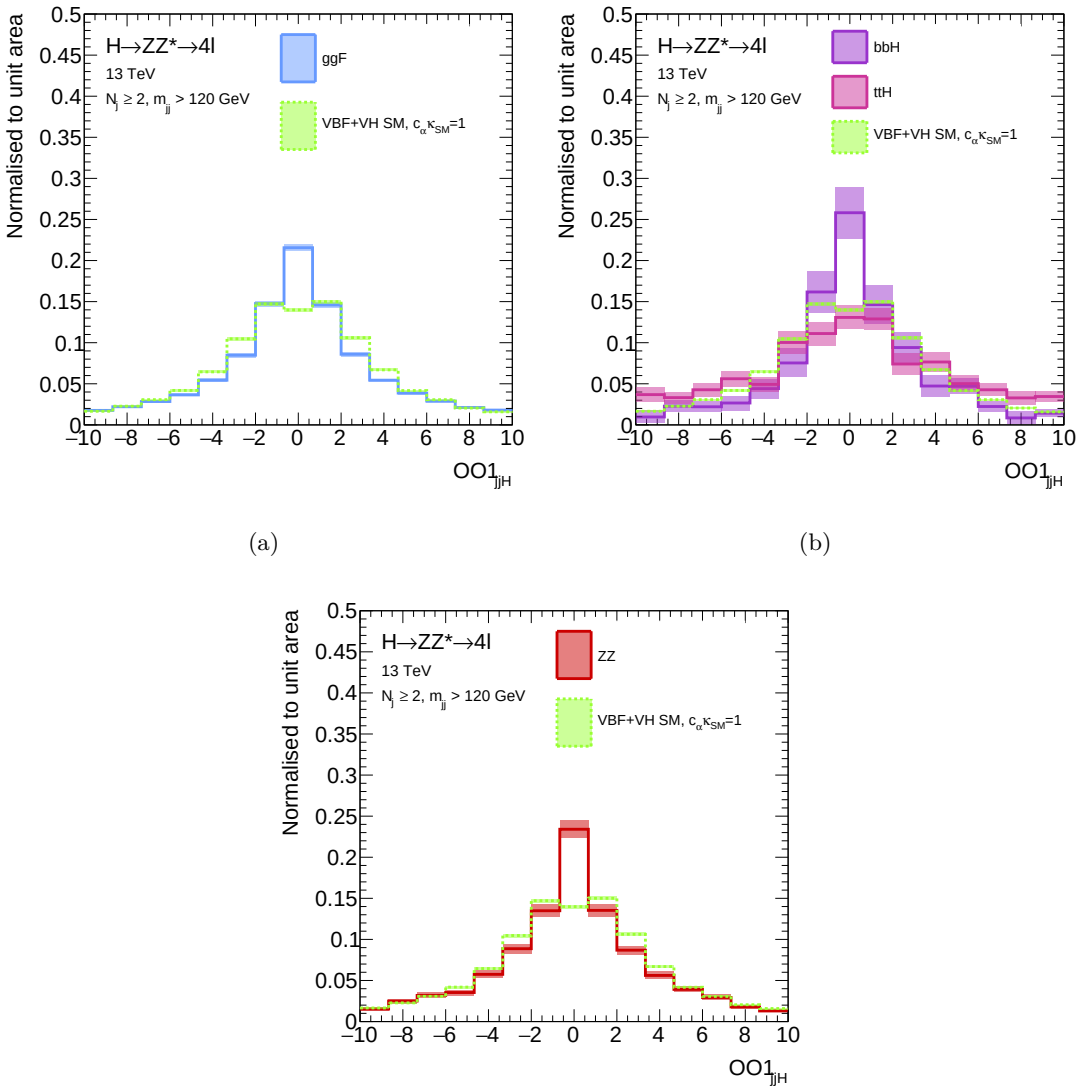


Figure 8.6: The expected distributions of the CP-odd observable $\Delta\phi_{jj}^{\text{sign}}$ for (a) ggF, (b) ttH+bbH and (c) ZZ^* background. The SM expectation for the VBF+VH signal is shown for comparison in all cases.



(a)

(b)

(c)

Figure 8.7: The expected distributions of the CP-odd observable $\mathcal{OO}_{1,jjH}$ for (a) ggF, (b) ttH+bbH and (c) ZZ^* background. The SM expectation for the VBF+VH signal is shown for comparison in all cases.

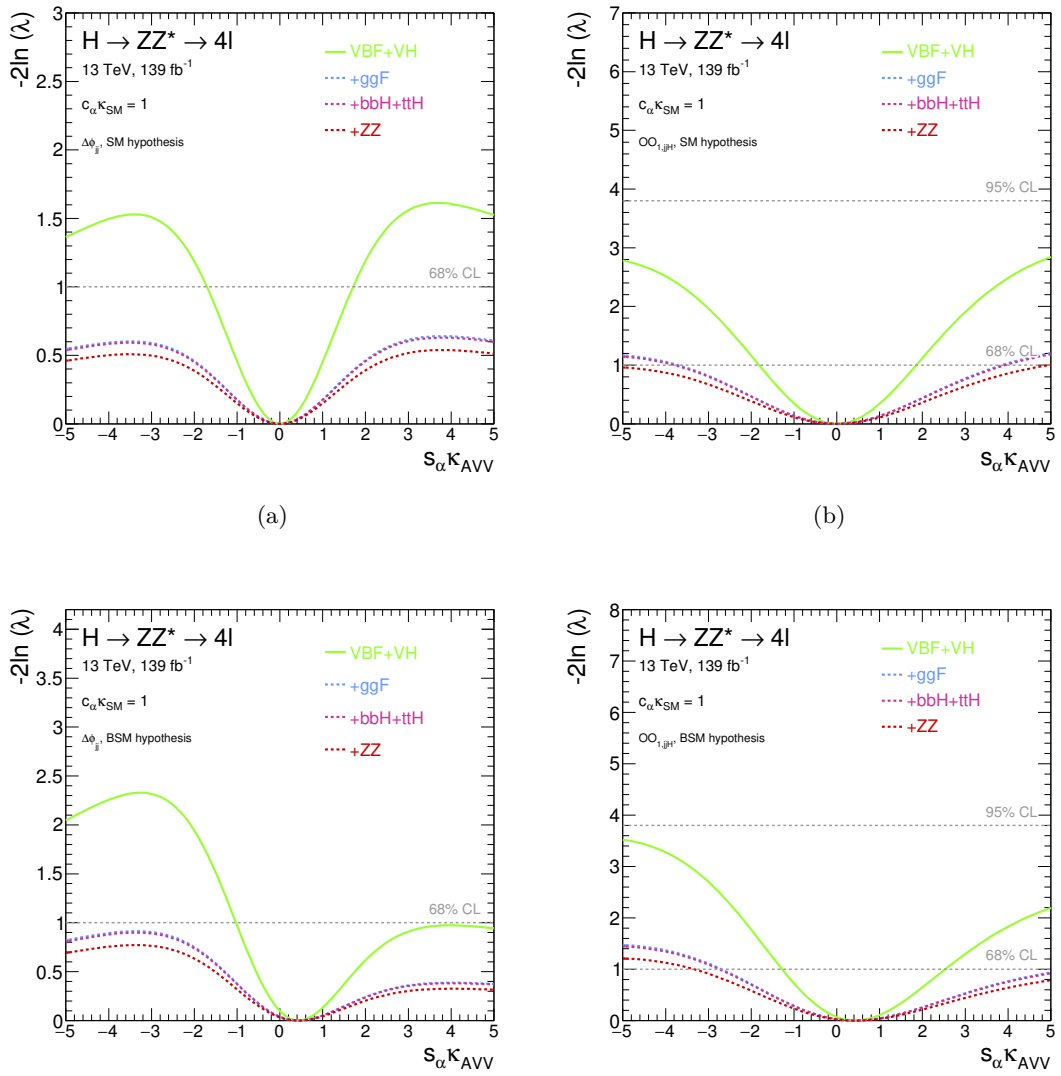


Figure 8.8: Expected values of the test statistic $q = -2\ln(\lambda)$ scanned over $s_\alpha \kappa_{AVV}$ using the (a) and (b) SM and (c) and (d) pseudo-data for distributions $\Delta\phi_{jj}^{\text{sign}}$ (left-hand-side) and $\mathcal{OO}_{1,jjH}$ (right-hand-side) for 139 fb^{-1} at a centre-of-mass energy of 13 TeV. The horizontal lines indicate the value of the profile likelihood ratio corresponding to the 68% and 95% CL intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.

Table 8.4: The expected confidence intervals at 68% CL and the best-fit values of the CP-odd BSM coupling parameter $s_\alpha \kappa_{AVV}$ for one-dimensional likelihood scans with the SM and CP-odd BSM pseudo-data at an integrated luminosity of 139 fb^{-1}at a centre-of-mass energy of 13 TeV. The SM coupling parameter $c_\alpha \kappa_{SM}$ is set to one in all scans.

Contributing processes	SM pseudo-data		BSM pseudo-data	
	68% CL interval	Best-fit	68% CL interval	Best-fit
	Observable: $\Delta\phi_{jj}^{sign}$			
VBF+VH	$[-1.69, 1.69]$	0.00	$[-1.02, -]$	0.43
+ ggF	$[-, -]$	0.00	$[-, -]$	0.43
+ ttH+bbH	$[-, -]$	0.00	$[-, -]$	0.43
+ ZZ^*	$[-, -]$	0.00	$[-, -]$	0.43
	Observable: $\mathcal{OO}_{1,jjH}$			
VBF+VH	$[-1.80, 1.84]$	0.00	$[-1.27, 2.52]$	0.43
+ ggF	$[-3.65, 3.80]$	0.00	$[-2.66, 5.41]$	0.43
+ ttH+bbH	$[-3.74, 3.88]$	0.00	$[-2.72, 5.56]$	0.43
+ ZZ^*	$[-4.90, 4.85]$	0.00	$[-3.30, -]$	0.43

In the case of the $\Delta\phi_{jj}^{\text{sign}}$ observable the exclusion limits at 68% CL can only be obtained for a pure VBF+VH signal and no background contributions. The corresponding 68% CL interval is $[-1.69, 1.69]$. By adding the ggF contribution or other background processes no exclusion limits can be set, neither for the SM nor for the BSM hypothesis.

The expected 68% CL interval in case of the $\mathcal{OO}_{1,jjH}$ observable and by considering only the VBF+VH signal is $[-1.80, 1.84]$ about 10% less significant than the one achieved with the $\Delta\phi_{jj}^{\text{sign}}$ observable. However, the $\mathcal{OO}_{1,jjH}$ observable provides in contrast to the $\Delta\phi_{jj}^{\text{sign}}$ observable a 68% CL interval even when all background processes are taken into account. By including the ggF production mode the exclusion limit degrades by a factor of two, due to a large ggF contribution of about 50% in the VBF-enriched category. Since the ttH+bbH production have only a small contribution of about 1.5% to the total signal yield only a slight degradation of the exclusion limit is expected from these processes. The irreducible background contributes with 18% to the event yield, leading to a further degradation of the exclusion limits by 30%. Finally, after taking all background processes into account, the expected exclusion limit on $s_\alpha \kappa_{AVV}$ is $[-4.90, 4.85]$ at 68% CL.

The maximum likelihood fit under the CP-odd BSM hypotheses always converges to the correct value of the hypothesised signal parameter of $s_\alpha \kappa_{AVV} = 0.43$. The values of the test statistics are smaller for positive than for negative BSM coupling parameters, leading to a one sided constrain on $s_\alpha \kappa_{AVV}$ of $[-3.30, -]$ at 68% CL from the $\mathcal{OO}_{1,jjH}$ observable after accounting for all background contributions.

In order to further improve the measurement sensitivity, the relatively large contribution of the background of about 70% has to be suppressed and discriminated from the VBF+VH signal. One possibility is to employ boosted decision trees as for the SM cross section measurement described in Section 5.3. This and other possible improvements are discussed in the following section.

8.4 Optimisation of the Measurement

The major background contribution in the VBF-enriched category originates from the ggF production mode with 50% of all events, followed by the irreducible ZZ^* background and the ttH+bbH production with 18% and 1.5%, respectively. As discussed in the previous section, these background contributions decrease the analysis sensitivity.

In order to provide a discrimination between the VBF+VH signal and the background processes, boosted decision trees introduced in Section 5.3 can be employed. A requirement on the BDT discriminant is optimised such to obtain the best signal enhancement. The BDT discriminant (BDT_{VBF}) used in the VBF-enriched category in the SM cross section measurement (Section 5.3.4) is employed.

The expected distribution of the BDT_{VBF} output discriminant in the VBF-enriched category is shown in Figure 8.9 for the SM ggF and SM VBF+VH production, as well as for VBF+VH signal with non-vanishing CP-odd BSM parameters. The distribution of the VBF+VH signal is shifted to positive values, while negative values are preferred by the ggF background. Thus, with a cut on BDT_{VBF} the ggF background contamination can be reduced.

However, the shape of the BDT_{VBF} discriminant for the VBF+VH processes is affected by the size of the CP-odd BSM coupling parameter. The larger absolute value of the CP-odd BSM coupling parameter, the stronger the shift of the distribution forwards to negative values. This dependence reduces the separation between signal and ggF and is due to the fact that some of the input variables used in the BDT training are dependent on BSM couplings.

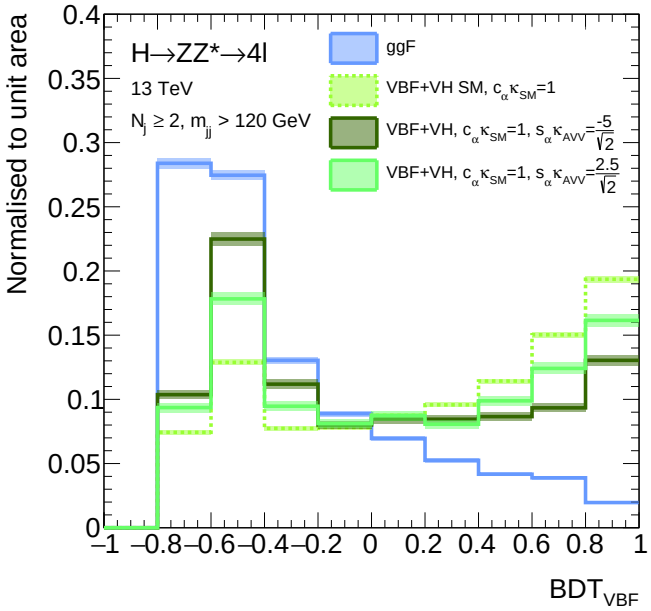
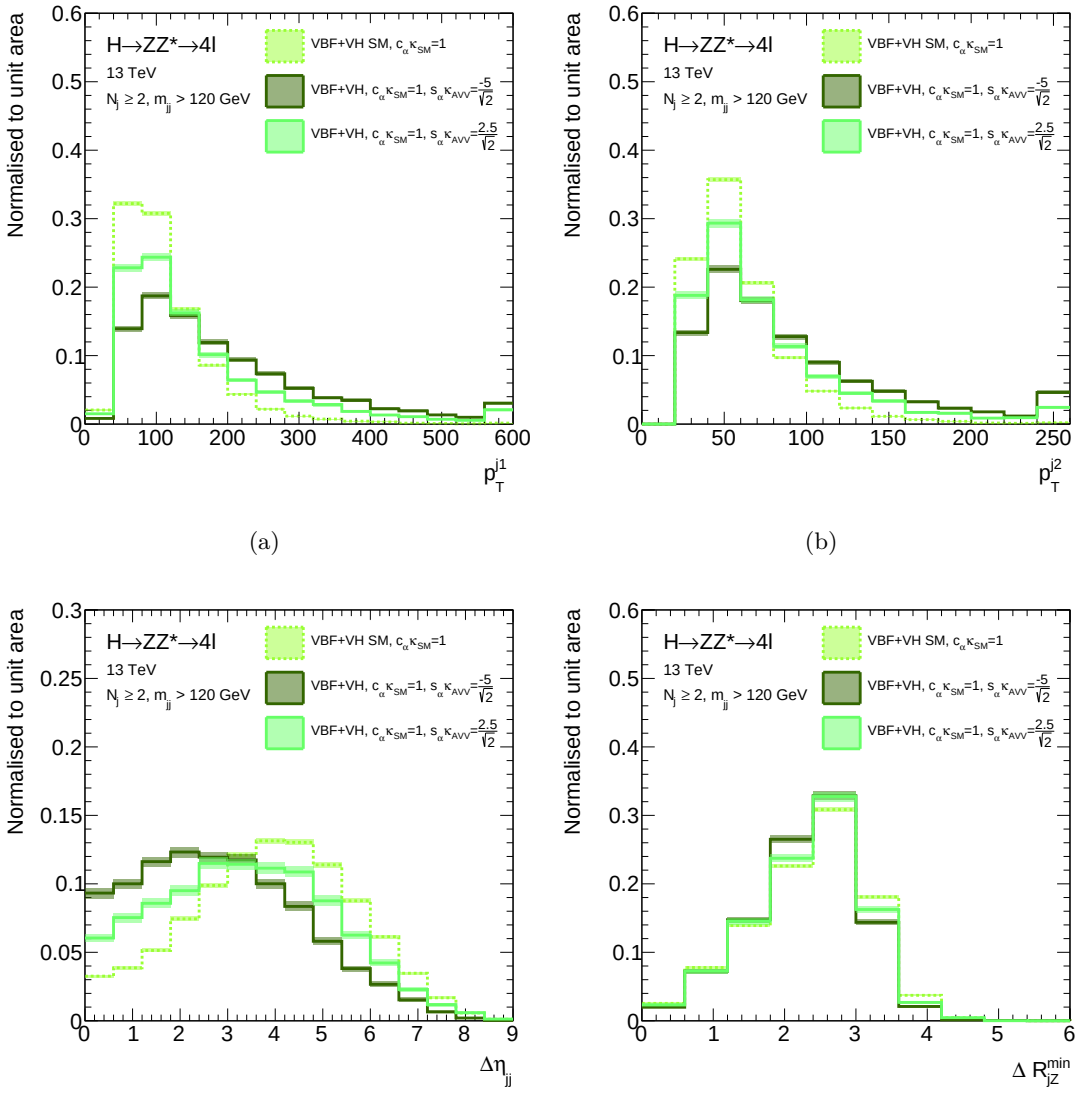


Figure 8.9: The expected distribution of the BDT response in the VBF-enriched reconstructed event category for the ggF and VBF+VH production predicted by the SM, as well as for the signal with different values of BSM CP-odd coupling parameters.

The distributions of input variables with largest BSM dependence (p_T^{j1}, p_T^{j2}, $\Delta\eta_{jj}$ and ΔR_{jZ}^{min}) are shown in Figure 8.10 for the VBF+VH processes with the SM and CP-odd BSM coupling parameters. The spectra of the transverse momentum of the two leading jets tend to be harder in the presence of BSM couplings (further away from ggF), while the jets are less separated in pseudorapidity compared to the SM expectation (i.e. closer to ggF). The minimal angle between the leading jet and the reconstructed Z boson ΔR_{jZ}^{min} is also smaller in case of the BSM couplings, i.e. it becomes similar to ggF.

Therefore, the threshold on the $\mathrm{BDT_{VBF}}$ response not only reduces the ggF background contribution, but also the BSM signal contribution. In order to remove potential biases of BDT discriminants on the final result, the final threshold value is chosen such to reject a similar fraction of SM and BSM VBF+VH signal events.

The fraction of signal events in the VBF-enriched category surviving a given requirement on the $\mathrm{BDT_{VBF}}$ threshold is shown in Table 8.5. Since the minimal value of the $\mathrm{BDT_{VBF}}$ is -0.8, the first cut is applied at $\mathrm{BDT_{VBF}} = -0.6$, rising in steps of 0.2 up to $\mathrm{BDT_{VBF}} > 0.8$. The fraction of SM signal events is 19.36% for a requirement of $\mathrm{BDT_{VBF}} > 0.8$ while it is even smaller, down to 16.15% and 12.62% for $s_\alpha\kappa_{AVV} = 2.5$ and $s_\alpha\kappa_{AVV} = 5$, respectively. In order to optimise the threshold value, the sensitivity of the analysis is studied in dependence on the threshold. The threshold values up to 0.4 are considered, corresponding to a relative difference in the fraction with respect to the SM of about 30% for the CP-odd BSM coupling parameter $s_\alpha\kappa_{AVV} = 5$.



(a) (b)

(c) (d)

Figure 8.10: The expected distributions of discriminating variables (a) p_T^{j1}, (b), p_T^{j2}, (c) $\Delta\eta_{jj}$ and (d) ΔR_{jZ}^{min} used as an input for the boosted decision tree discriminant $\mathrm{BDT_{VBF}}$ in the VBF-enriched reconstructed event category for the VBF+VH SM expectation and BSM CP-odd coupling parameters.

The expected number of signal and background events in the VBF-enriched event category are shown in Table 8.6 for several different thresholds on the $\mathrm{BDT_{VBF}}$ discriminant. The corresponding signal composition in each of these categories is shown in Table 8.7 and visualised in Figure 8.11. The number of SM VBF+VH signal events reduces by up to a factor of two for a BDT threshold of 0.4. The ggF background decreases by a larger factor of ten, and the ttH+bbH production and the ZZ^* background by a factor of 30 and 20, respectively. Since the background is rejected more strongly compared to the signal, the signal purity can be enhanced up to 72%.

For each of the tested BDT thresholds, a maximum likelihood fit of the $\Delta\phi_{jj}^{\mathrm{sign}}$ and $\mathcal{OO}_{1,jjH}$ distribution is performed for the SM hypothesis and for the BSM hypothesis with the CP-odd BSM coupling parameter $s_\alpha\kappa_{AVV} = 0.43$. The result is shown in Table 8.8. The respective values of the test statistic are shown in Figure 8.12.

Table 8.5: Expected fractions of VBF+VH signal events in the reconstructed VBF-enriched category after applying a requirement on the $\mathrm{BDT_{VBF}}$ response after the full event selection at an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV for the SM and CP-odd BSM admixtures with two different CP-odd BSM coupling parameters. In addition, the difference of the fraction $\Delta_{rel}^{\mathrm{BSM}} = \frac{\mathrm{fraction_{SM}} - \mathrm{fraction_{BSM}}}{\mathrm{fraction_{SM}}}$ to the SM prediction is shown.

VBF-enriched category	VBF+VH SM [%]	$s_\alpha\kappa_{AVV} = 2.5$ [%]	$\Delta_{rel}^{s_\alpha\kappa_{AVV}=2.5}$ [%]	$s_\alpha\kappa_{AVV} = 5$ [%]	$\Delta_{rel}^{s_\alpha\kappa_{AVV}=5}$ [%]
No $\mathrm{BDT_{VBF}}$ cut	100.00	100.00	0	100.00	0
$\mathrm{BDT_{VBF}} > -0.6$	92.57	90.64	2.08	89.06	3.79
$\mathrm{BDT_{VBF}} > -0.4$	79.69	72.81	8.63	67.27	15.59
$\mathrm{BDT_{VBF}} > -0.2$	71.95	63.35	11.95	56.07	22.07
$\mathrm{BDT_{VBF}} > 0.0$	64.14	55.24	13.88	47.87	25.37
$\mathrm{BDT_{VBF}} > 0.2$	55.38	46.51	16.02	39.77	28.19
$\mathrm{BDT_{VBF}} > 0.4$	45.81	38.45	16.08	31.69	30.83
$\mathrm{BDT_{VBF}} > 0.6$	34.39	28.56	16.93	22.53	34.48
$\mathrm{BDT_{VBF}} > 0.8$	19.36	16.15	16.54	12.62	34.79

Table 8.6: Expected number of VBF+VH signal and background events after the full event selection in the reconstructed VBF-enriched event category at an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV for different thresholds on the BDT$_{\text{VBF}}$ response assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV. Only statistical uncertainties are quoted.

VBF-enriched category	VBF+VH	ggF	ttH+bbH	ZZ^*	Total
No BDT$_{\text{VBF}}$ cut	10.10 ± 0.03	16.83 ± 0.12	0.507 ± 0.015	5.98 ± 0.11	33.41 ± 0.16
BDT$_{\text{VBF}} > -0.6$	9.31 ± 0.03	12.05 ± 0.10	0.375 ± 0.013	4.16 ± 0.09	25.89 ± 0.14
BDT$_{\text{VBF}} > -0.4$	8.03 ± 0.03	7.43 ± 0.08	0.136 ± 0.008	2.07 ± 0.06	17.66 ± 0.11
BDT$_{\text{VBF}} > -0.2$	7.31 ± 0.02	5.24 ± 0.07	0.081 ± 0.007	1.26 ± 0.05	13.88 ± 0.09
BDT$_{\text{VBF}} > 0.0$	6.62 ± 0.02	3.74 ± 0.06	0.049 ± 0.005	0.802 ± 0.039	11.22 ± 0.07
BDT$_{\text{VBF}} > 0.2$	5.87 ± 0.02	2.57 ± 0.05	0.027 ± 0.004	0.503 ± 0.031	8.97 ± 0.06
BDT$_{\text{VBF}} > 0.4$	5.05 ± 0.02	1.69 ± 0.04	0.018 ± 0.003	0.319 ± 0.025	7.07 ± 0.05

Table 8.7: Expected VBF+VH signal and background fractions after the full event selection in the reconstructed VBF-enriched event category at an integrated luminosity of 139 fb^{-1} at $\sqrt{s} = 13$ TeV for different thresholds on the BDT$_{\text{VBF}}$ response assuming the SM Higgs boson signal with a mass $m_H = 125$ GeV

VBF-enriched category	VBF+VH [%]	ggF [%]	ttH+bbH [%]	ZZ^* [%]
No BDT$_{\text{VBF}}$ cut	30.22	50.36	1.52	17.90
BDT$_{\text{VBF}} > -0.6$	35.95	46.55	1.45	16.06
BDT$_{\text{VBF}} > -0.4$	45.47	42.07	0.77	11.70
BDT$_{\text{VBF}} > -0.2$	52.65	37.72	0.58	9.05
BDT$_{\text{VBF}} > 0.0$	59.06	33.35	0.44	7.15
BDT$_{\text{VBF}} > 0.2$	65.45	28.65	0.30	5.60
BDT$_{\text{VBF}} > 0.4$	71.39	23.85	0.25	4.51

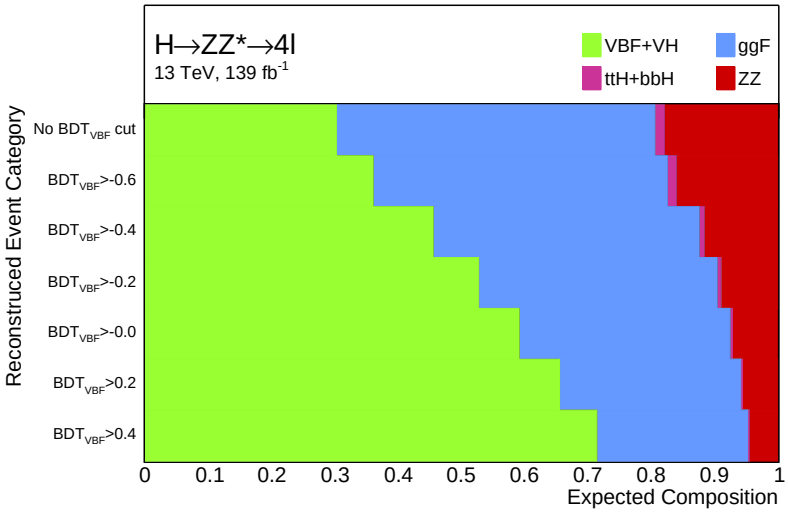


Figure 8.11: Expected signal composition in the VBF-enriched reconstructed event category with different thresholds on the $\mathrm{BDT_{VBF}}$ discriminant.

For both CP-odd observables the sensitivity is significantly improved by the BDT requirement. For the CP-violating observable $\Delta\phi_{jj}^{\mathrm{sign}}$, the SM hypothesis cannot be excluded if no $\mathrm{BDT_{VBF}}$ cut is applied. The 68% CL interval for the CP-odd coupling parameter improves to $[-2.39, 2.42]$ with a cut $\mathrm{BDT_{VBF}} > 0.4$.

For the CP-odd observable $\mathcal{O}\mathcal{O}_{1,jjH}$ the exclusion limits improve by a factor of two from $[-4.90, 4.85]$ to $[-2.47, 2.58]$ after applying the cut $\mathrm{BDT_{VBF}} > 0.4$. For $\mathrm{BDT_{VBF}}$ cuts above -0.2 the sensitivity improves only slightly, since the statistical uncertainty increases.

The fit performed for the hypothesis of a CP-odd BSM signal gives similar results as the one obtained for the SM expectation. The limits improve with larger thresholds on $\mathrm{BDT_{VBF}}$ and the relative improvement for thresholds above -0.2 is only minor. The best-fit values of $s_\alpha \kappa_{AVV}$ obtained from the fits agree with the values assumed for the BSM hypothesis, $s_\alpha \kappa_{AVV} - 0.43$.

Table 8.8: The expected confidence intervals at 68% CL and best-fit values of the CP-odd BSM parameter $s_\alpha \kappa_{AVV}$, obtained from the one-dimensional likelihood scan with the SM and the CP-odd BSM hypotheses at an integrated luminosity of 139 fb^{-1}at $\sqrt{s} = 13$ TeV. The results are shown for different BDT$_{\text{VBF}}$ cuts. In all scans the value of the SM coupling parameter $c_\alpha \kappa_{SM}$ is set to the SM prediction of one.

Employed	SM pseudo-data		BSM pseudo-data	
threshold	68% CL interval	Best-fit	68% CL interval	Best-fit
	Observable: $\Delta\phi_{jj}^{sign}$			
No BDT$_{\text{VBF}}$ cut	$[-,-]$	0.00	$[-,-]$	0.43
BDT$_{\text{VBF}} > -0.6$	$[-,-]$	0.00	$[-,-]$	0.43
BDT$_{\text{VBF}} > -0.4$	$[-3.56, 4.65]$	0.00	$[-1.96,-]$	0.43
BDT$_{\text{VBF}} > -0.2$	$[-2.90, 2.55]$	0.00	$[-1.73,-]$	0.43
BDT$_{\text{VBF}} > 0.0$	$[-2.51, 2.47]$	0.00	$[-1.64,-]$	0.43
BDT$_{\text{VBF}} > 0.2$	$[-2.40, 2.37$	0.00	$[-1.62,-]$	0.43
BDT$_{\text{VBF}} > 0.4$	$[-2.39, 2.42]$	0.00	$[-1.61, 4.36]$	0.43
	Observable: $\mathcal{OO}_{1,jjH}$			
No BDT$_{\text{VBF}}$ cut	$[-4.90, 4.85]$	0.00	$[-3.30,-]$	0.43
BDT$_{\text{VBF}} > -0.6$	$[-3.52, 3.75]$	0.00	$[-2.65, 5.02]$	0.43
BDT$_{\text{VBF}} > -0.4$	$[-2.81, 3.03]$	0.00	$[-2.18, 3.92]$	0.43
BDT$_{\text{VBF}} > -0.2$	$[\mathbf{-2.56, 2.64}]$	**0.00**	$[\mathbf{-1.96, 3.39}]$	**0.43**
BDT$_{\text{VBF}} > 0.0$	$[-2.46, 2.59]$	0.00	$[-1.91, 3.30]$	0.43
BDT$_{\text{VBF}} > 0.2$	$[-2.43, 2.52]$	0.00	$[-1.90, 3.19]$	0.43
BDT$_{\text{VBF}} > 0.4$	$[-2.47, 2.58]$	0.00	$[-1.94, 3.25]$	0.43

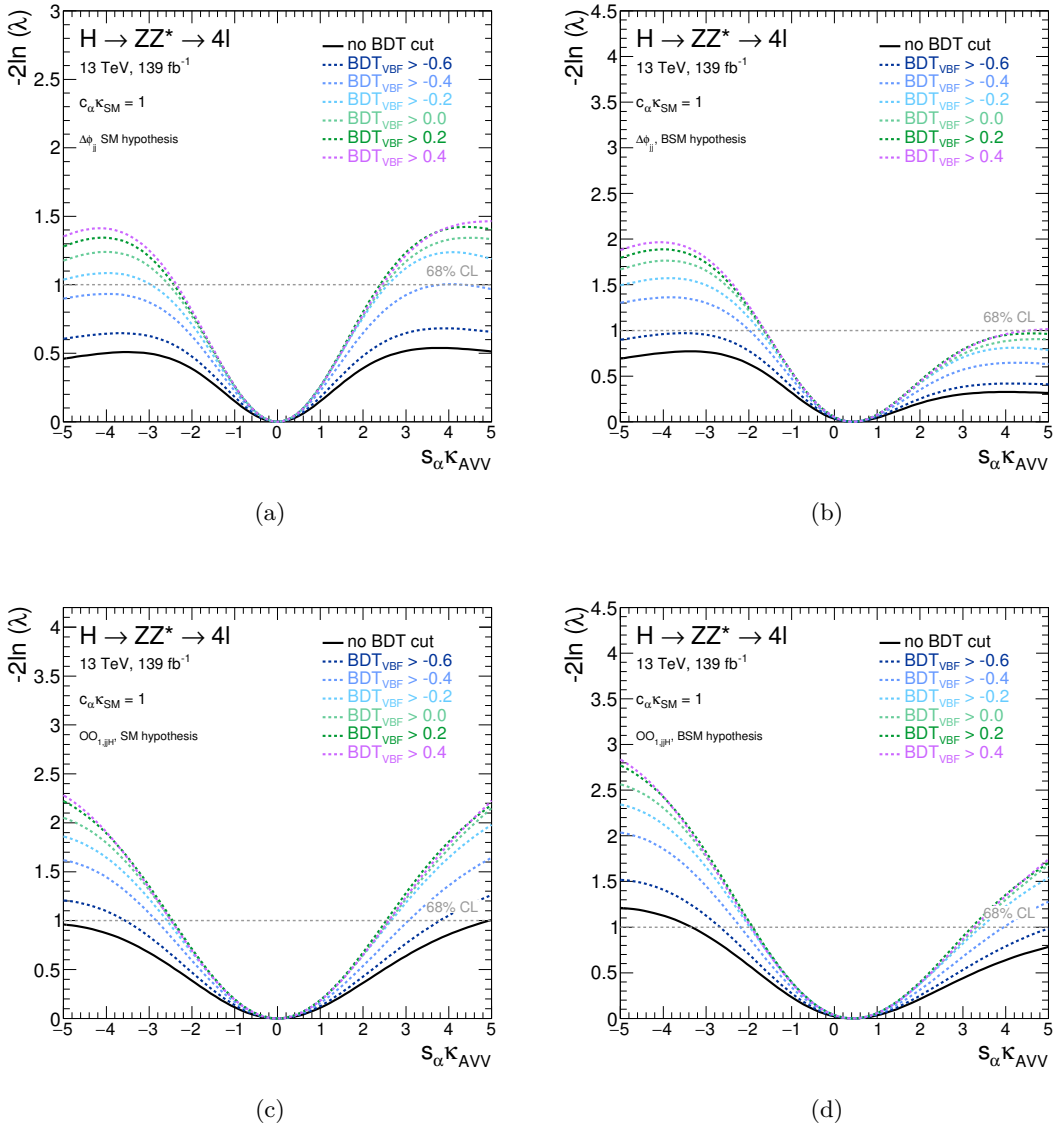


Figure 8.12: Expected values of the test statistic $q = -2\ln(\lambda)$ for the SM (left-hand-side) and BSM (right-hand-side) hypothesis, obtained from the scans over the CP-odd BSM coupling parameter $s_\alpha \kappa_{AVV}$ at an integrated luminosity of 139 fb^{-1} at a centre-of-mass energy of 13 TeV. The likelihood curves are shown for different $\mathrm{BDT_{VBF}}$ cuts, in (a) and (b) using the $\Delta\phi_{jj}^{\mathrm{sign}}$ observable and in (c) and (d) the $\mathcal{OO}_{1,jjH}$ observable. The horizontal lines indicate the value of the profile likelihood ratio corresponding to the 68% CL interval for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic.

As a consequence of these optimisation studies a cut of $\mathrm{BDT_{VBF}} > -0.2$ is employed. The corresponding relative difference in the BSM signal fraction relative to the SM signal is up to 20% (compare Table 8.5) for the BSM coupling parameter $s_\alpha \kappa_{AVV} = 5$ and the signal purity is 50%. The expected confidence intervals at 68% CL are $[-2.90, 2.55]$ and $[-2.56, 2.64]$ for the $\Delta\phi_{jj}^{\mathrm{sign}}$ and the $\mathcal{OO}_{1,jjH}$ observable, respectively.

To summarise, the use of BDT discriminants suppress the background contribution and allows for an improvement of exclusion limits on the CP-odd BSM couplings. However, the currently used $\mathrm{BDT_{VBF}}$ discriminant has a strong dependence on the size of the BSM couplings, reducing the discrimination power for large BSM coupling values. Possible solution is to train a new BDT discriminant dedicated for the measurement of CP-violation. The input variables should be chosen such to provide a good separation between the VBF+VH and ggF production mode, while at the same time minimising the BSM dependence.

Another possibility of the background suppression is the employment of recurrent neural network (rNN) [223] instead of BDTs. This deep learning method provides in general a better separation of the signal from the background for the same input variables. However, the training of the rNN requires larger input data sets.

Further improvements can be obtained by fitting the entire distribution of the multivariate discriminating variable instead of selecting only events above a certain threshold. In this way, the statistical precision is not reduced. The sensitivity can also be increased by using a variable binning for the distributions of the CP-odd observables. In this way, interesting structures in the shapes of the distribution can be resolved, while less sensitive regions like the side-band regions can be combined into one single bin, thus increase the statistical precision.

As discussed in Section 8.2.3, BSM CP-even contributions can influence the results of the CP-invariance measurement. To be agnostic against such contributions the amount of asymmetry of the distribution could be used as a single parameter instead of the full distribution. Another possibility is to take the shape information into account by using a binned asymmetry.

8.5 Comparison with Previous Measurements

The test of CP-invariance in the XVV vertex has been performed by several other analysis summarised in Table 8.9. No evidence for CP-violation has been found in any of the mentioned measurements. In particular, the CP-invariance was probed in the decay vertex of $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ decays using the full set of Run 1 data recorded with the ATLAS detector [28]. The XVV production vertex has also been probed with Run 1 and first Run 2 data with the ATLAS experiment in the VBF production mode using the similar Optimal Observable method in the $H \to \tau\tau$ decay channel [207, 208].

Since the analyses in the $H \to \tau\tau$ decay channel use the same approach to investigate the CP-invariance as presented in this thesis, this study is described in more detail. The results of each analysis are compared with the results obtained in this thesis.

Table 8.9: Summary of ATLAS analyses testing the CP-invariance in the XVV vertex

Decay mode	Data set	Centre-of-mass energy	Integrated luminosity	Reference
$H \to ZZ^* \to 4\ell$	Run 1	7,8 TeV	24.8 fb^{-1}	[28]
$H \to WW^* \to e\nu\mu\nu$	Run 1	8 TeV	20.3 fb^{-1}	[28]
$H \to \tau\tau$	Run 1	8 TeV	20 fb^{-1}	[207]
$H \to \tau\tau$	Run 2	13 TeV	36.1 fb^{-1}	[208]

The test of the CP-invariance in the XVV decay vertex of $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ decays is performed with proton-proton collision data of the full Run 1 data set corresponding to an integrated luminosity of 24.3 fb^{-1} collected with the ATLAS detector at $\sqrt{s} = 7$ and 8 TeV. Possible CP-violating contributions are modelled within the HC framework. The CP-odd BSM coupling under study is defined relative to the SM-like coupling parameter κ_{SM},

$$\frac{\widetilde{\kappa}_{AVV}}{\kappa_{SM}} \cdot \tan(\alpha) = \frac{v}{4 \cdot \Lambda} \cdot \frac{\kappa_{AVV}}{\kappa_{SM}} \cdot \tan(\alpha). \tag{8.13}$$

The measurement is based on the shape of the CP-sensitive kinematic distributions of the Higgs boson decay products. Constraints on the CP-odd BSM coupling are set separately for $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ decays and their combination.

The expected confidence intervals at 68% and 95% CL for the combined measurement are $(\widetilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan(\alpha) \in [-1.2, 1.2]$ and $(\widetilde{\kappa}_{AVV}/\kappa_{SM}) \cdot \tan(\alpha) \in [-2.33, 2.30]$. Assuming

the SM prediction for the SM-like coupling parameter, $c_\alpha \kappa_{SM} = 1$, the intervals translate into $s_\alpha \kappa_{AVV} \in [-19.5, 19.5]$ at 68% CL and $s_\alpha \kappa_{AVV} \in [-37.9, 37.4]$ at 95% CL. The limits improve by a factor $\sqrt[4]{L}$ and $\sqrt{2}$ due to the larger amount of data and increased centre-of-mass energy, respectively, leading to an expected limit of [-8.9,8.9] at 68% CL and [-17.4,17.2] at 95% CL for 139 fb^{-1} and $\sqrt{s} = 13$ TeV.

The exclusion limit obtained with the study presented in this thesis is [-2.56,2.64] at 68% CL, which is about a factor of three better. The CP-invariance measurement in $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ decays is also based on the concept of Optimal Observables but employing only the measured four-momenta of the charged leptons from the Higgs boson decay for the calculation of the matrix elements without taking into account the information of the leading final state jets from the VBF production. The decay vertex information and a better signal-to-background ratio allows to set the exclusion limits at 95% CL, which is not possible in the study presented in this thesis, which relies on the information from the VBF production vertex. Thus, the presented study can be further improved by taking into account the information from the decay vertex.

The CP-sensitive $H \to \tau\tau$ analysis is performed with proton-proton collision data corresponding to an integrated luminosity of 36.1 fb^{-1} collected with the ATLAS detector at $\sqrt{s} = 13$ TeV in the years 2015 and 2016. The effective Lagrangian considered in the $H \to \tau\tau$ decay channel is the SM Lagrangian augmented with CP-odd dimension-six operators, as introduced in Section 7.4.1. The strength of CP-violation in the VBF production is described by a single parameter $\widetilde{d}$. The CP-odd coupling parameter $\widetilde{d}$ can be translated into the HC model via the relation

$$\widetilde{d} = \frac{v \cdot \kappa_{AVV}}{4 \cdot \Lambda \cdot \kappa_{SM}} \tan(\alpha). \tag{8.14}$$

The CP-invariance measurement in the $H \to \tau\tau$ decay channel employs the same definition of the first-order Optimal Observable for the VBF production as introduced in Section 8.2.2, allowing for a direct comparison of the two measurements.

The analysis in the $H \to \tau\tau$ channel has a larger sensitivity to the VBF production. Due to a larger number of events much stronger VBF selection criteria can be applied: $N_j \geq 2$, $p_T^{j2} > 30$ GeV, $m_{jj} > 300$ GeV, $|\eta_{jj}| > 3$ and $p_T^{j1} > 40$ GeV. In addition, the signal-to-background ratio in the signal regions is enhanced by means of dedicated BDT discriminants. Furthermore, an optimised variable binning is used for the $\mathcal{OO}_{1,jjH}$ observable.

The measurement in the $H \to \tau\tau$ is performed with all possible combinations of final states, with τ leptons decaying either leptonically ($\tau \to \ell\nu\bar{\nu}$ with $\ell = e, \mu$) or hadronically ($\tau \to \nu$ hadrons). Depending on the decay mode of the τ lepton the events are separated

into four analysis channels, the dileptonic same-flavour ($\tau_{lep}\tau_{lep}$ SF), the dileptonic different flavour ($\tau_{lep}\tau_{lep}$ DF), the semileptonic ($\tau_{lep}\tau_{had}$), and the fully hadronic ($\tau_{had}\tau_{had}$) channel. In each final state a BDT discriminant is trained using different discriminating input variables. The signal regions are built by a threshold requirement on the corresponding BDT response. With the introduction of these criteria a similar signal-to-background ratio of about 2 as in the presented study can be achieved.

The analysis results are obtained from a binned maximum-likelihood fit of the $\mathcal{OO}_{1,jjH}$ distribution in the signal regions fitted simultaneously with dedicated discriminants in the control regions, which are used to constrain the background normalisations and the systematic uncertainties.

For an integrated luminosity of 36.1 fb^{-1} the expected 68% CL interval for the parameter $\widetilde{d}$ after combining all analysis channels is $\widetilde{d} \in [-0.035, 0.033]$, which translates into $s_\alpha \kappa_{AVV} \in [-0.40, 0.38]$ in the HC model. The limits scale with to an expected interval of $[-0.28, 0.27]$ at 139 fb^{-1}. The exclusion limit obtained with the study presented in this thesis is $[-2.56, 2.64]$, which is worse by a factor of about ten. This is mainly due to the larger number of signal events in the $H \to \tau\tau$ decay channel. With more the pure VBF category and the BDT observables a similar signal-to-background ratio is achieved in the $H \to \tau\tau$ as in the $H \to ZZ^* \to 4\ell$. However, the number of signal events is larger by a factor of ten. Furthermore, the variable binning of the $\mathcal{OO}_{1,jjH}$ distribution also increases the sensitivity.

9

Summary

The Higgs boson as last missing particle predicted by the Standard Model has been discovered in 2012 by the ATLAS and CMS experiments in proton-proton collisions at the Large Hadron Collider (LHC). The measurements of the spin and CP quantum numbers of the discovered particle confirmed that it is in deed the Standard Model Higgs boson. So far, the measurements of cross sections for different Higgs boson production modes and decay rates are well compatible with the Standard Model prediction. The measurement precision still does not exclude the most favoured models of physics beyond the Standard Model (BSM). For instance, it is still possible that the tensor structure of the Higgs boson couplings is modified by small CP-even or CP-odd BSM admixtures to the pure CP-even Higgs state of the Standard Model. Precise measurements of the Higgs boson production cross sections and of the tensor structure of the Higgs boson couplings, including the direct search for CP-violation in the Higgs sector, are very important as a window to new physics.

In this thesis, such measurements have been performed in the $H \to ZZ^* \to 4\ell$ ($\ell = e, \mu$) decay channel to probe the Higgs boson coupling structure. The cross sections for the main Higgs boson production modes have been measured inclusively in different phase space regions. These measurements have also been extrapolated to 3000 fb^{-1} of luminosities to estimate the ultimate precision at the high-luminosity LHC (HL-LHC). The tensor structure of the Higgs boson couplings to weak vector bosons, gluons and the top quark have been probed using two different analysis approaches. The sensitivity for CP-violation in the vector boson fusion production of the Higgs boson has been evaluated. The measurements have been performed with Run 2 data sets of the ATLAS detector at the LHC at a centre-of-mass energy of 13 TeV. The data sets used correspond to integrated luminosities of 36.1 fb^{-1}, 79.8 fb^{-1} and 139 fb^{-1}. The numbers of selected Higgs boson decays to four leptons in these data sets are 95, 195 and 316, respectively.

The Higgs boson production cross sections for the main production modes have been measured with 79.8 fb^{-1} of data. For this purpose, the selected Higgs boson decays have been classified according to their production process. Multivariate discriminants were employed to improve the sensitivity to the production modes in the different categories. The measured cross sections times branching ratio for $H \to ZZ^* \to 4\ell$ decays are in good

agreement with the Standard Model prediction. In the rapidity range $|y_H| < 2.5$ the measured value is 1.57 ± 0.18 pb which agrees within 1.5 standard deviations, with the Standard Model prediction of 1.33 ± 0.08 pb.

The above measurements have been extrapolated to the maximum integrated luminosity of 3000 fb^{-1} expected at the HL-LHC taking into account the increase in the centre-of-mass energy to 14 TeV and a reduction of theoretical and experimental uncertainties. The overall measurement precision is expected to improved by a factor of about four for ggF production and by a factor of about six for VBF and VH production at the end of the HL-LHC data taking. While with 79.8 fb^{-1} of data only an upper limit could be set on the ttH production cross section, it can be measured with a precision of 20% at the end of the HL-LHC.

The tensor structure of the Higgs boson couplings have been probed using two different effective field theory approaches in which deviations from the Standard Model predictions are parametrised in terms of BSM coupling parameters. The first analysis, performed with 36.1 fb^{-1} of data, probed for the contributions of anomalous CP-even and CP-odd contributions in the Higgs production and decay using the Higgs Characterisation framework. Possible BSM contributions to the Higgs boson couplings to vector bosons (HVV interaction vertex) and to gluons (Hgg interaction vertex) have been studied independently. Limits have been placed on CP-even (CP-odd) BSM coupling contributions to the HVV interaction vertex described by the coupling parameter $\cos\alpha \cdot \kappa_{HVV}$ ($\sin\alpha \cdot \kappa_{AVV}$) and on CP-odd BSM contributions to the Hgg interaction vertex described by the parameter $\sin\alpha \cdot \kappa_{Agg}$. Input to the analysis are the reconstructed Higgs boson candidates selected by the inclusive $H \to ZZ^* \to 4\ell$ analysis classified according to the different production modes. The limits $-0.6 < \cos\alpha \cdot \kappa_{HVV} < 4.2$, $-4.4 < \sin\alpha \cdot \kappa_{AVV} < 4.4$ and $-0.68 < \sin\alpha \cdot \kappa_{Agg} < 0.68$ have been set on the BSM coupling parameters at 95% CL.

The second analysis has been performed with the full Run 2 data set corresponding to an integrated luminosity of 139 fb^{-1}. In addition to the Higgs boson couplings to vector bosons and gluons, also the top Yukawa interaction vertex ttH has been probed in ttH production for CP-even and CP-odd BSM contributions modelled using the Standard Model Effective Field Theory (SMEFT). In the SMEFT, the CP-even (CP-odd) BSM couplings at the Hgg interaction vertex are described by coupling parameters c_{HG} ($c_{H\widetilde{G}}$). The corresponding CP-even (CP-odd) SMEFT parameters for the HVV interaction vertex are c_{HW}, c_{HB}, c_{HWB} ($c_{H\widetilde{W}}, c_{H\widetilde{B}}, c_{H\widetilde{W}B}$) and c_{uH} ($c_{\tilde{u}H}$) for the ttH vertex. Cross sections have been measured in exclusive phase space regions defined at particle level and fitted with the SMEFT prediction, rather than interpreting directly the observed event yield as in the previous approach. The most stringent limits have been set on the CP-even BSM coupling parameter to gluons, $-0.0074 < c_{HG} < 0.0080$, and on the CP-odd BSM coupling parameters to the weak vector bosons, $-2.4 < c_{H\widetilde{W}} < 2.4$, $-0.56 < c_{H\widetilde{B}} < 0.56$

and $-1.03 < c_{H\widetilde{W}B} < 1.03$. The constraints on the CP-even BSM coupling parameters are weaker, $-3.4 < c_{HW} < 2.1$, $-0.62 < c_{HB} < 0.59$ and $-1.06 < c_{HWB} < 0.99$.
These are the first constraints on SMEFT coupling parameters in the $H \to ZZ^* \to 4\ell$ channel. They will be combined with the measurements in other Higgs boson decay channels.

The sensitivity to CP-violation in vector boson production of the Higgs boson has been evaluated using fully simulated $H \to ZZ^* \to 4\ell$ candidates of the full Run 2 data set of 139 fb^{-1}. The kinematic properties of the final state particles are combined into a single discriminant in a well defined optimum way (Optimal Observable). Expected limits for the full Run 2 data set from Monte Carlo simulation have been derived for the CP-odd BSM coupling parameter $\sin\alpha \cdot \kappa_{AVV}$ of the Higgs Characterisation framework. The expected 68% CL interval is $-2.56 < \sin\alpha \cdot \kappa_{AVV} < 2.64$ while no limits can yet be set at 95% CL.

A

Four-Lepton Triggers

A.1 Electron Trigger

Table A.1: Electron trigger requirements for the $H \to ZZ^* \to 4\ell$ analysis

Data Period	Single-electron	Di-electron	Tri-electron
2015	HLT_e24_lhmedium_L1EM18VH (*D*)	HLT_2e12_lhloose_L12EM10VH	HLT_e17_lhloose_2e9_lhloose
D,E,F,G,H,J	HLT_e24_lhmedium_L1EM20VH (*E,F,G,H,J*)		
	HLT_e60_lhmedium e120_lhloose		
2016	HLT_e24_lhtight_nod0_ivarloose	HLT_2e15_lhvloose_nod0_L12EM13VH	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
A	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2016	HLT_e24_lhtight_nod0_ivarloose	HLT_2e15_lhvloose_nod0_L12EM13VH	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
B-D3	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2016	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
D4-E	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2016	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
F-G2	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2016	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
G3-I3	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2016	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0	HLT_e17_lhloose_nod0_2e9_lhloose_nod0
I4-L	HLT_e60_lhmedium_nod0		
	HLT_e60_medium		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2017	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0_L12EM15VHI	HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
	HLT_e60_lhmedium_nod0	HLT_2e24_lhvloose_nod0	
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		
2018	HLT_e26_lhtight_nod0_ivarloose	HLT_2e17_lhvloose_nod0_L12EM15VHI	HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
	HLT_e26_lhtight_nod0	HLT_2e24_lhvloose_nod0	
	HLT_e60_lhmedium_nod0		
	HLT_e140_lhloose_nod0		
	HLT_e300_etcut		

A.2 Muon Trigger

Table A.2: Muon trigger requirements for the $H \to ZZ^* \to 4\ell$ analysis

Data Period	Single-muon	Di-muon	Tri-muon
2015	HLT_mu20_iloose_L1MU15	HLT_2mu10	HLT_3mu6
D,E,F,G,H,J	HLT_mu40	HLT_mu18_mu8noL1	HLT_3mu6_msonly
	HLT_mu60_0eta105_msonly		HLT_mu18_2mu4noL1
2016	HLT_mu24_ivarloose_L1MU15	HLT_2mu10	HLT_mu20_2mu4noL1
A	HLT_mu24_iloose_L1MU15	HLT_2mu10_nomucomb	HLT_3mu4
	HLT_mu40	HLT_mu20_mu8noL1	HLT_mu6_2mu4
	HLT_mu50	HLT_mu20_nomucomb_mu6noL1_nscan03	HLT_mu20_nomucomb_mu6noL1_nscan03
			HLT_mu11_nomucomb_2mu4noL1_nscan03_L1MU11_2MU6
			HLT_mu20_msonly_mu10noL1_msonly_nscan05_noComb
2016	HLT_mu24_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
B-D3	HLT_mu24_imedium	HLT_2mu14_nomucomb	HLT_3mu6
	HLT_mu50	HLT_mu20_mu8noL1	HLT_mu6_2mu4
		HLT_mu20_nomucomb_mu6noL1_nscan03	HLT_mu20_nomucomb_mu6noL1_nscan03
			HLT_mu11_nomucomb_2mu4noL1_nscan03_L1MU11_2MU6
			HLT_mu20_msonly_mu10noL1_msonly_nscan05_noComb
2016	HLT_mu24_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
D4-E	HLT_mu24_imedium	HLT_mu20_mu8noL1	HLT_3mu6_msonly
	HLT_mu26_ivarmedium	HLT_mu22_mu8noL1	
	HLT_mu26_imedium		
	HLT_mu50		
2016	HLT_mu26_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
F-G2	HLT_mu26_imedium	HLT_mu22_mu8noL1	HLT_3mu6_msonly
	HLT_mu50		
2016	HLT_mu26_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
G3-I3	HLT_mu50	HLT_mu22_mu8noL1	HLT_3mu6_msonly
2016	HLT_mu26_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
I4-L	HLT_mu50	HLT_mu22_mu8noL1	HLT_3mu6_msonly
			HLT_3mu4
2017	HLT_mu26_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
	HLT_mu50	HLT_mu22_mu8noL1	HLT_3mu6_msonly
	HLT_mu60_0eta105_msonly	HLT_mu22_mu8noL1_calotag_0eta010	HLT_3mu4
		HLT_3mu6	
		HLT_4mu4	
2018	HLT_mu26_ivarmedium	HLT_2mu14	HLT_mu20_2mu4noL1
	HLT_mu50	HLT_mu22_mu8noL1	HLT_3mu6
	HLT_mu60_0eta105_msonly		

A.3 Electron-Muon Trigger

Table A.3: Electron-muon trigger requirements for the $H \to ZZ^* \to 4\ell$ analysis

Data Period	Electron-muon	Data Period	Electron-muon
2015	HLT_e17_lhloose_mu14	2016	HLT_e17_lhloose_nod0_mu14
D,E,F,G,H,J	HLT_2e12_lhloose_mu10	*G3-I3*	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1
	HLT_e12_lhloose_2mu10		HLT_e7_lhmedium_nod0_mu24
	HLT_e24_medium_L1EM20VHI_mu8noL1		HLT_e12_lhloose_nod0_2mu10
	HLT_e7_medium_mu24		HLT_2e12_lhloose_nod0_mu10
2016	HLT_e17_lhloose_nod0_mu14	2016	HLT_e17_lhloose_nod0_mu14
A	HLT_e24_lhmedium_nod0_L1EM20VHI_mu8noL1	*I4-L*	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1
	HLT_e7_lhmedium_nod0_mu24		HLT_e7_lhmedium_nod0_mu24
	HLT_e12_lhloose_nod0_2mu10		HLT_e12_lhloose_nod0_2mu10
	HLT_2e12_lhloose_nod0_mu10		HLT_2e12_lhloose_nod0_mu10
2016	HLT_e17_lhloose_nod0_mu14	2017	HLT_e17_lhloose_nod0_mu14
B-D3	HLT_e24_lhmedium_nod0_L1EM20VHI_mu8noL1		HLT_e26_lhmedium_nod0_mu8noL1
	HLT_e7_lhmedium_nod0_mu24		HLT_e7_lhmedium_nod0_mu24
	HLT_e12_lhloose_nod0_2mu10		HLT_e12_lhloose_nod0_2mu10
	HLT_2e12_lhloose_nod0_mu10		HLT_2e12_lhloose_nod0_mu10
2016	HLT_e17_lhloose_nod0_mu14	2018	HLT_e17_lhloose_nod0_mu14
D4-E	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1		HLT_e26_lhmedium_nod0_mu8noL1
	HLT_e7_lhmedium_nod0_mu24		HLT_e7_lhmedium_nod0_mu24
	HLT_e12_lhloose_nod0_2mu10		HLT_e12_lhloose_nod0_2mu10
	HLT_2e12_lhloose_nod0_mu10		HLT_2e12_lhloose_nod0_mu10
2016	HLT_e17_lhloose_nod0_mu14		
F-G2	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1		
	HLT_e7_lhmedium_nod0_mu24		
	HLT_e12_lhloose_nod0_2mu10		
	HLT_2e12_lhloose_nod0_mu10		

B

EFT Parametrisation with CP-even BSM Coupling Parameters

B.1 Generated BSM Signal Samples

Table B.1: MadGraph5_aMC@NLO syntax of the different Higgs boson production processes used for the event generation. Where not otherwise stated the four flavour scheme (4FS) is used for simulation of the Monte Carlo sample. 5FS denotes the five flavour scheme.

Production process	MadGraph5 syntax	Number of Monte Carlos samples

ggF+bbH	`define jb = j b b~`
	`generate p p > h QED=1`
	`add process p p > h jb QED=1`
	`add process p p > h jb jb QED=1`
		3

VBF+VH-Had	`generate p p > h j j QCD=0`	10
ZH-Lep	`generate p p > h l+ l-` `add process p p > h vl vl~`	10
WH-Lep	`generate p p > h l+ vl` `add process p p > h l- vl~`	3
ttH	`generate p p > h t t ~`	6
$tHjb$	`generate p p > h t b~` `add process p p > h t~ b j`	6

tHW	`define p = p b b ~`
	`generate p p > h t w -`
	`add process p p > h t~ w+`
		6

Table B.2: MadGraph5_aMC@NLO syntax of the considered decay modes for the parametrisation of the total decay width, as well as for the partial decay width in the $H \to ZZ^* \to 4\ell$ decay. In addition, the number of required simulated samples is quoted.

Decay process	MadGraph5 syntax	Number of Monte Carlos samples
$H \to \gamma\gamma$	`generate h > a a`	
$H \to Z\gamma$	`add process h > z a`	
$H \to bb$	`add process h > b b~`	
$H \to gg$	`add process h > g g`	
$H \to W\ell\nu$	`add process h > w+ l- vl~`	12
	`add process h > w- l+ vl`	
$H \to Wjj$	`add process h > w+ j j`	
	`add process h > w- j j`	
$H \to Z\ell\ell$	`add process h > z l+ l-`	
$H \to Zjj$	`add process h > z j j`	
$H \to Z\nu\nu$	`add process h > z vl vl~`	
$H \to 4\ell$	`add process h > l+ l- l+ l-`	9

Table B.3: Configuration of the simulated BSM signal samples with MadGraph5_aMC@NLO for ggF+bbH, VBF+VH-Had and VH-Lep production

Monte Carlo sample type		c_{HG}	c_{uH}	c_{HW}	c_{HB}	c_{HWB}	Syntax
ggF+bbH	SM	0	0	0	0	0	NP^2==0
	INT term c_{HG}	1	0	0	0	0	NP^2==1
	BSM term c_{HG}	1	0	0	0	0	NP^2==2
VBF+VH-Had and VH-Lep	SM	0	0	0	0	0	NP^2==0
	INT term c_{HW}	0	0	1	0	0	NP^2==1
	INT term c_{HB}	0	0	0	1	0	NP^2==1
	INT term c_{HWB}	0	0	0	0	1	NP^2==1
	BSM term c_{HW}	0	0	1	0	0	NP^2==2
	BSM term c_{HB}	0	0	0	1	0	NP^2==2
	BSM term c_{HWB}	0	0	0	0	1	NP^2==2
	BSM term c_{HW}, c_{HB}	0	0	1	1	0	NP^2==2
	BSM term c_{HW}, c_{HWB}	0	0	1	0	1	NP^2==2
	BSM term c_{HB}, c_{HWB}	0	0	0	1	1	NP^2==2

Table B.4: Configuration of the simulated BSM signal samples with MadGraph5_aMC@NLO for WH-Lep, ttH, $tHjb$ and tHW production

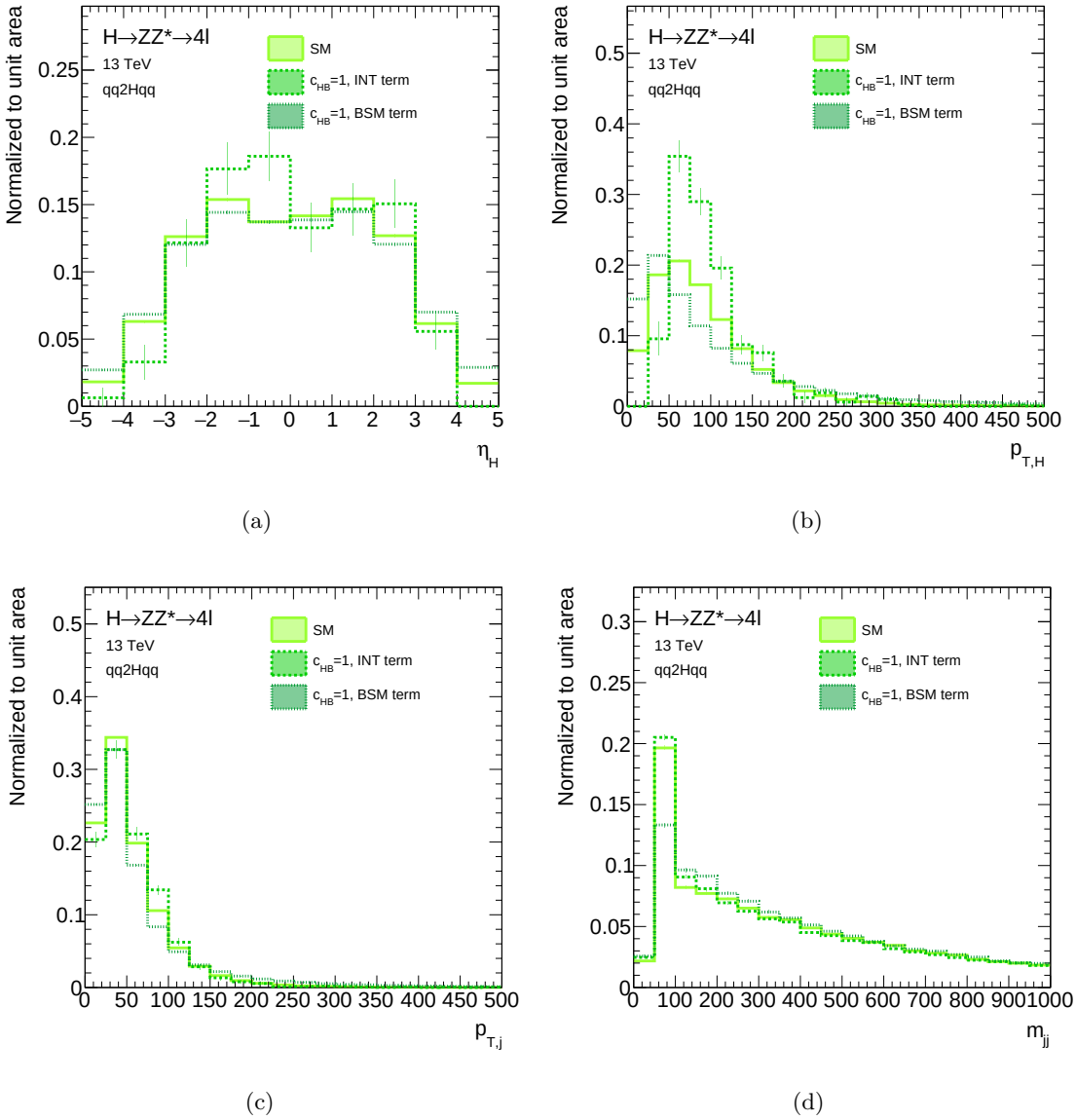
Monte Carlo sample type		c_{HG}	c_{uH}	c_{HW}	c_{HB}	c_{HWB}	Syntax
WH-Lep	SM	0	0	0	0	0	NP^2==0
	INT term c_{HW}	0	0	1	0	0	NP^2==1
	BSM term c_{HW}	0	0	1	0	0	NP^2==2
ttH, $tHjb$ and tHW	SM	0	0	0	0	0	NP^2==0
	INT term c_{HG}	1	0	0	0	0	NP^2==1
	INT term c_{uH}	0	1	0	0	0	NP^2==1
	BSM term c_{HG}	1	0	0	0	0	NP^2==2
	BSM term c_{uH}	0	1	0	0	0	NP^2==2
	BSM term c_{HG}, c_{uH}	1	1	0	0	0	NP^2==2

B.2 Production Cross Sections, Partial Decay and Total Width of the Higgs Boson

Table B.5: Inclusive cross section for all production modes as well as the total and partial Higgs boson decay width for CP-even BSM coupling parameters. NP1 denote the MadGraph5_aMC@NLO syntax `NP^2==1` and corresponds to the SM-BSM interference term, while NP2 denotes `NP^2==2` and corresponds to the pure BSM term.

ggF+bbH	SM	$c_{HG}=1$, NP1	$c_{HG}=1$ NP2									
Cross section [pb]	27.64	826.65	7077.00									
VBF+VH-Had	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2		
Cross section [pb]	4.14980	0.54030	0.53309	0.01442	0.11343	0.18065	0.08263	0.70511	0.56707	0.10768		
ZH-Lep	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2		
Cross section [pb]	0.1506	0.10958	0.03852	0.01327	0.00472	0.04880	0.00962	0.04351	0.07219	0.02207		
WH-Lep	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2									
Cross section [pb]	0.23775	0.21016	0.08283									
ttH	SM	$c_{HG}=1$, NP1	$c_{HG}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP1	$c_{HG}=1$, $c_{uH}=1$, NP2						
Cross section [pb]	0.41310	0.22906	0.28077	−0.04893	0.00145	0.27764						
$tHjb$	SM	$c_{HG}=1$, NP1	$c_{HG}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP1	$c_{HG}=1$, $c_{uH}=1$, NP2						
Cross section [pb]	0.06476	0.00045	0.00268	−0.00199	0.00273	0.00273						
tHW	SM	$c_{HG}=1$, NP1	$c_{HG}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP2	$c_{HG}=1$, $c_{uH}=1$, NP2						
Cross section [pb]	0.01651	0.00814	0.00596	−0.00170	0.00016	0.00585						
$H\to 4\ell$	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2		
Width [MeV]	$5.02\cdot10^{-4}$	$-1.00\cdot10^{-4}$	$3.78\cdot10^{-4}$	$-5.62\cdot10^{-5}$	$1.34\cdot10^{-3}$	$9.07\cdot10^{-5}$	$3.81\cdot10^{-4}$	$1.69\cdot10^{-3}$	$1.13\cdot10^{-4}$	$1.02\cdot10^{-3}$		
$H\to all$	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HG}=1$, NP1	$c_{HG}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2
Width [MeV]	4.99	−0.269	0.809	−0.377	6.03	0.309	1.78	7.58	74.4	9.42	0.392	2.17

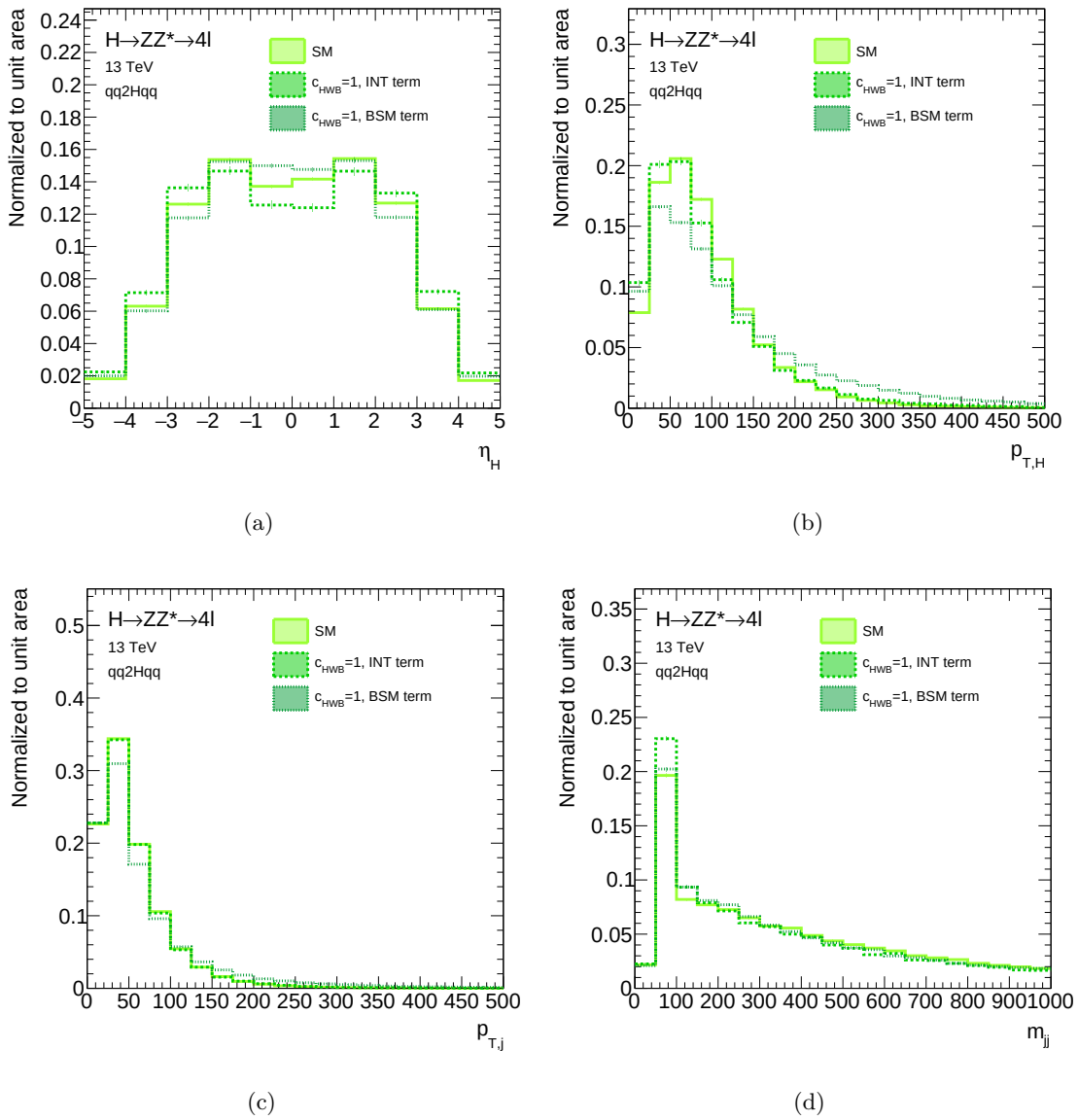
B.3 Particle Level Distributions



(a) (b)

(c) (d)

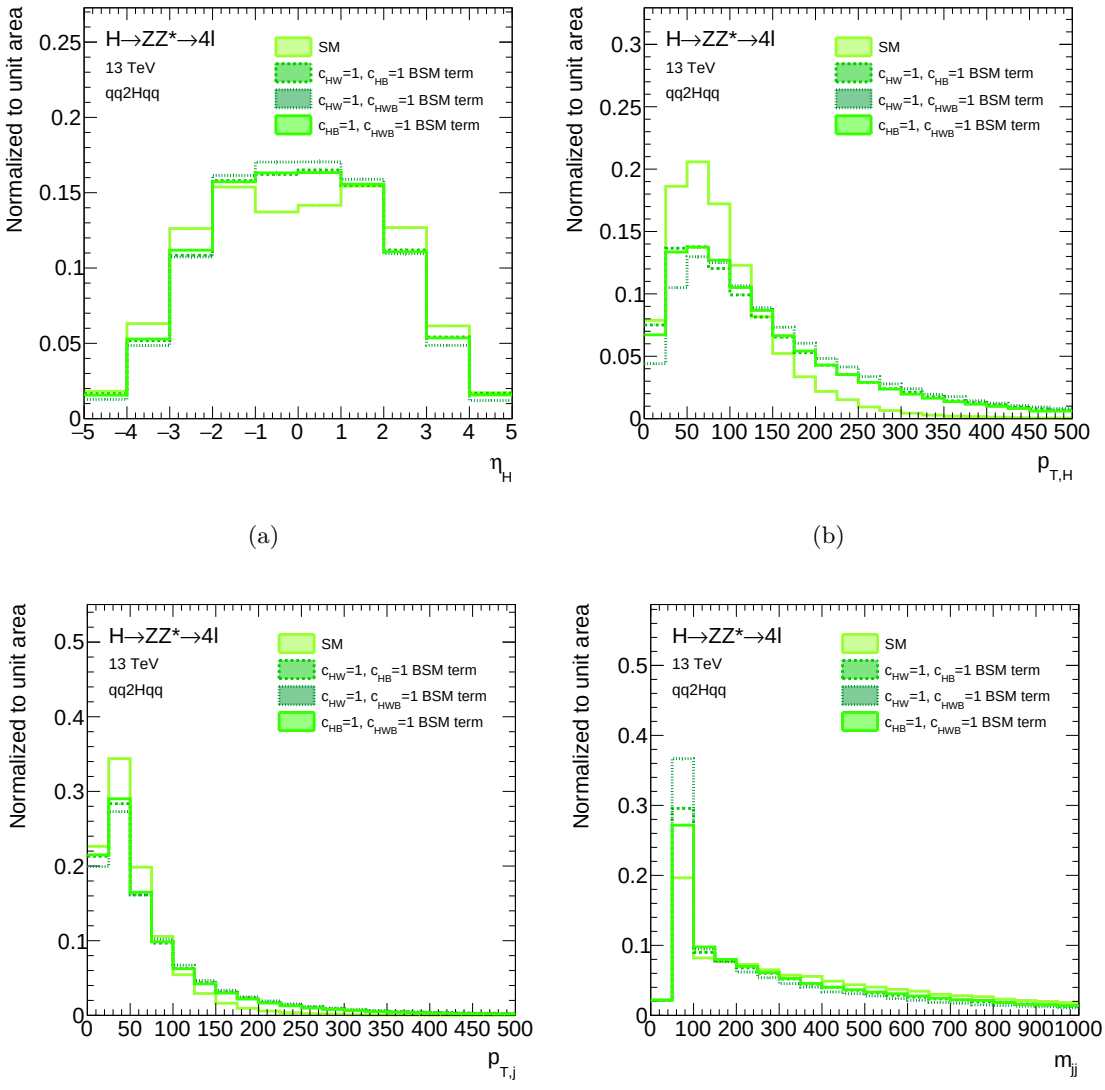
Figure B.1: The expected distributions of (a) η_H, (b) p_T^H, (c) p_T^j and (d) m_{jj} in the qq2Hqq production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HB} = 1$.



(a) (b)

(c) (d)

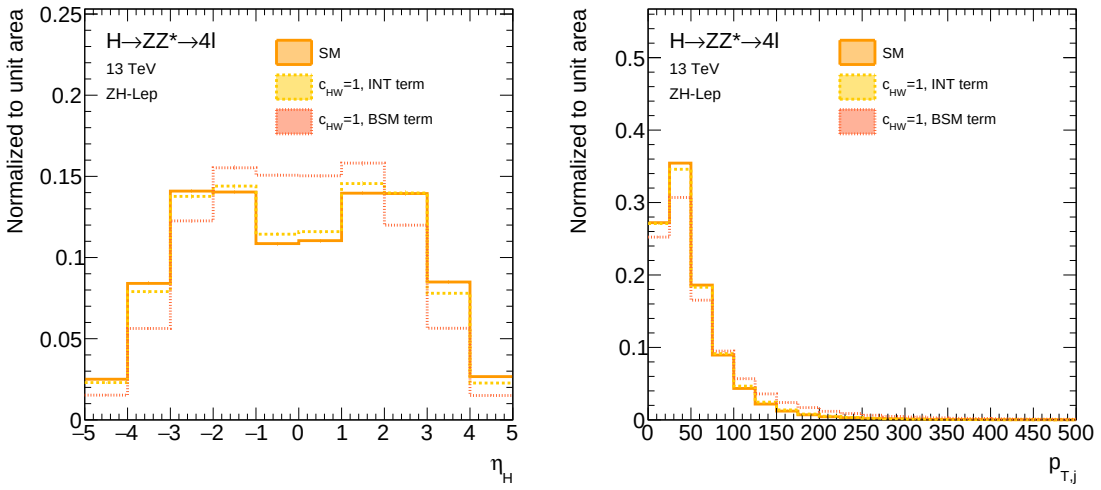
Figure B.2: The expected distributions of (a) η_H, (b) p_T^H, (c) p_T^j and (d) m_{jj} in the qq2Hqq production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HWB} = 1$.



(a) (b)

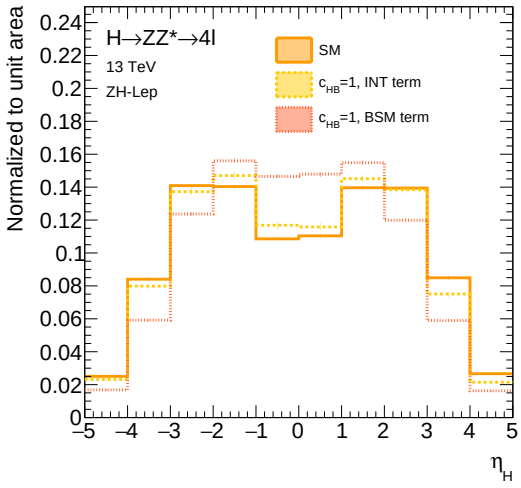
(c) (d)

Figure B.3: The expected distributions of (a) η_H, (b) p_T^H, (c) p_T^j and (d) m_{jj} in the qq2Hqq production mode shown separately for the SM and the mixed BSM terms.

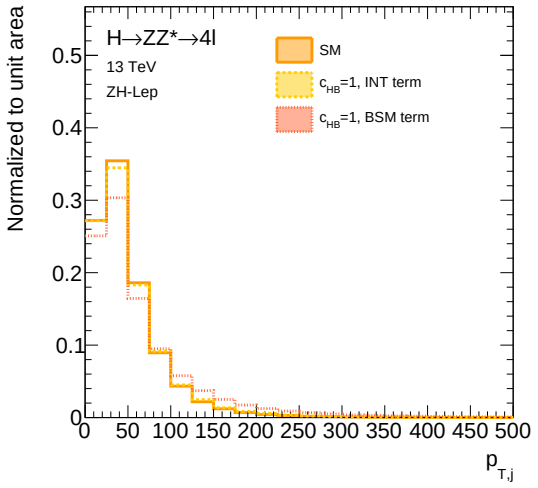


(a)

(b)

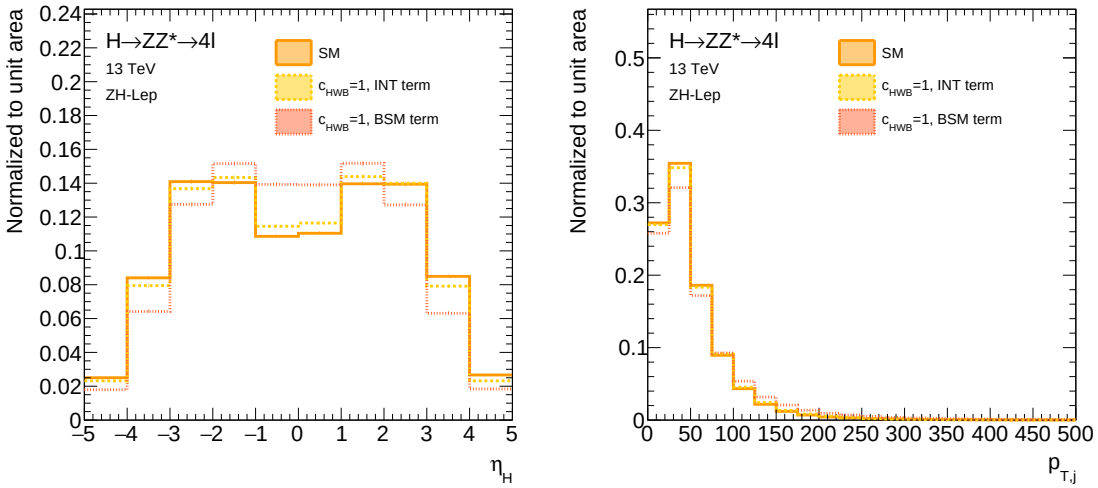


(c)



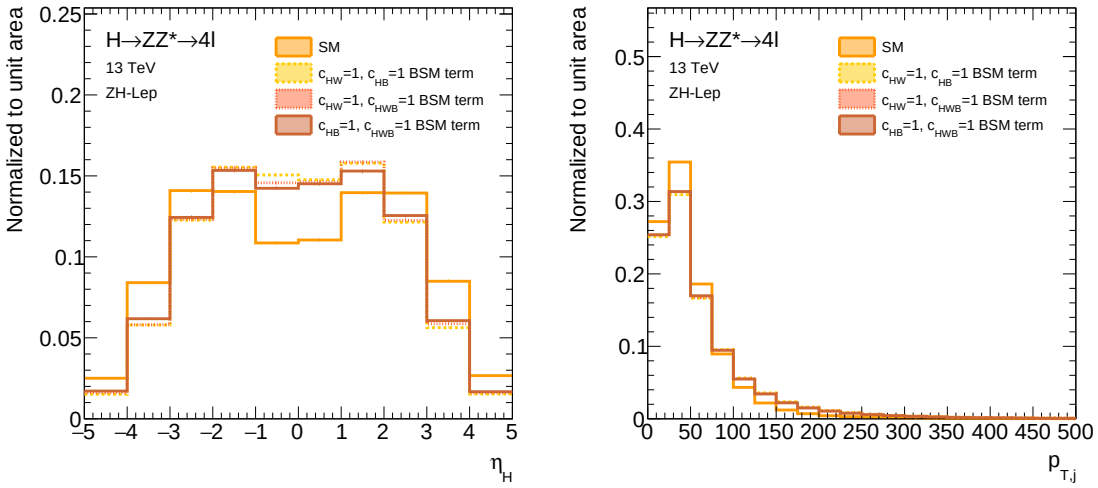
(d)

Figure B.4: The expected distributions of (a) and (c) η_H and (b) and (d) in the ZH-Lep production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HW} = 1$ and $c_{HB} = 1$, respectively.



(a)

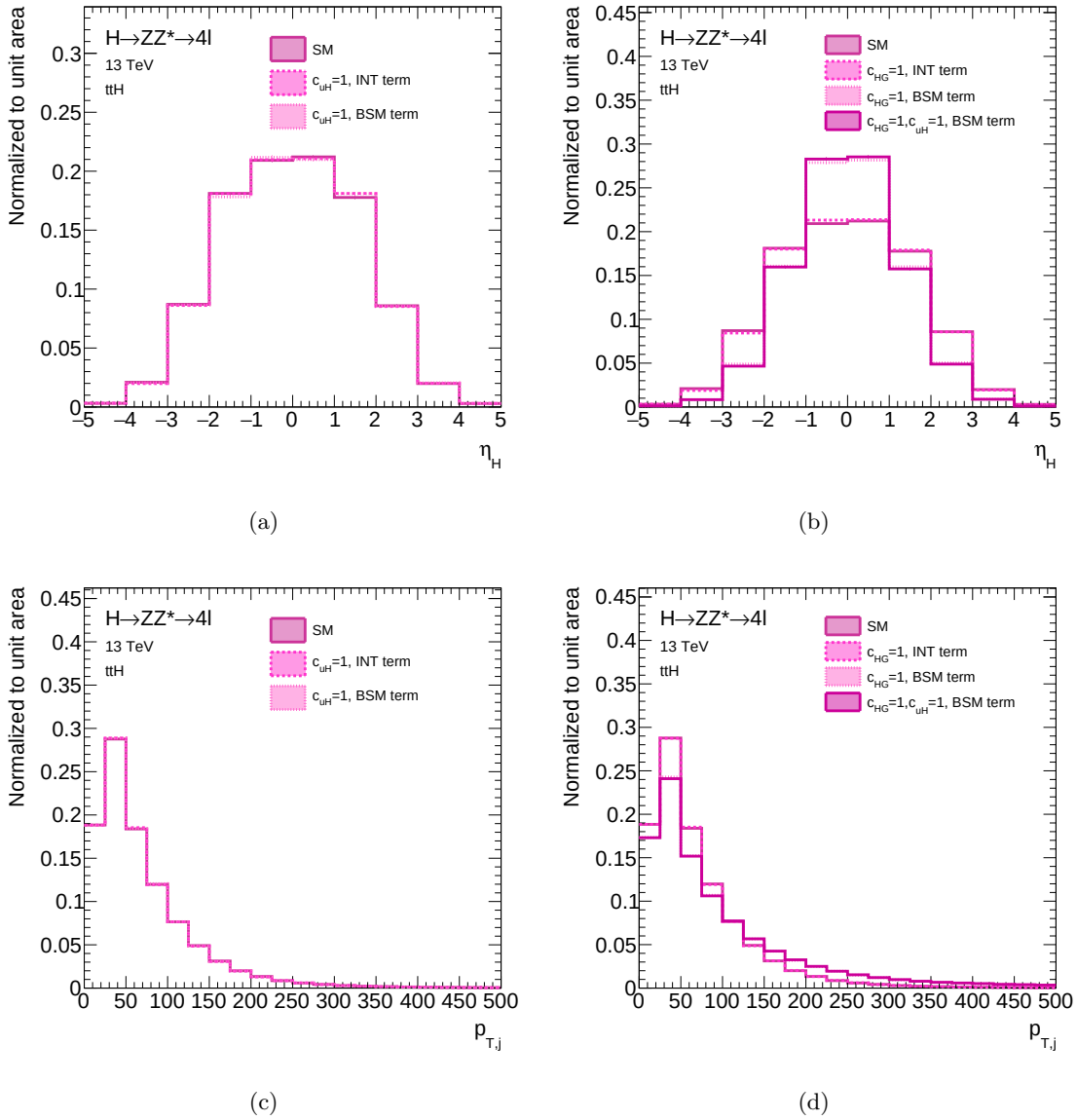
(b)



(c)

(d)

Figure B.5: The expected distributions of (a) and (c) η_H and (b) and (d) in the ZH-Lep production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HWB} = 1$ and for the SM and the mixed BSM terms, respectively.



(a) (b)

(c) (d)

Figure B.6: The expected distributions of (a) and (b) η_H and (c) and (d) p_T^j in the ttH production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{uH} = 1$ and $c_{HG} = 1$, respectively.

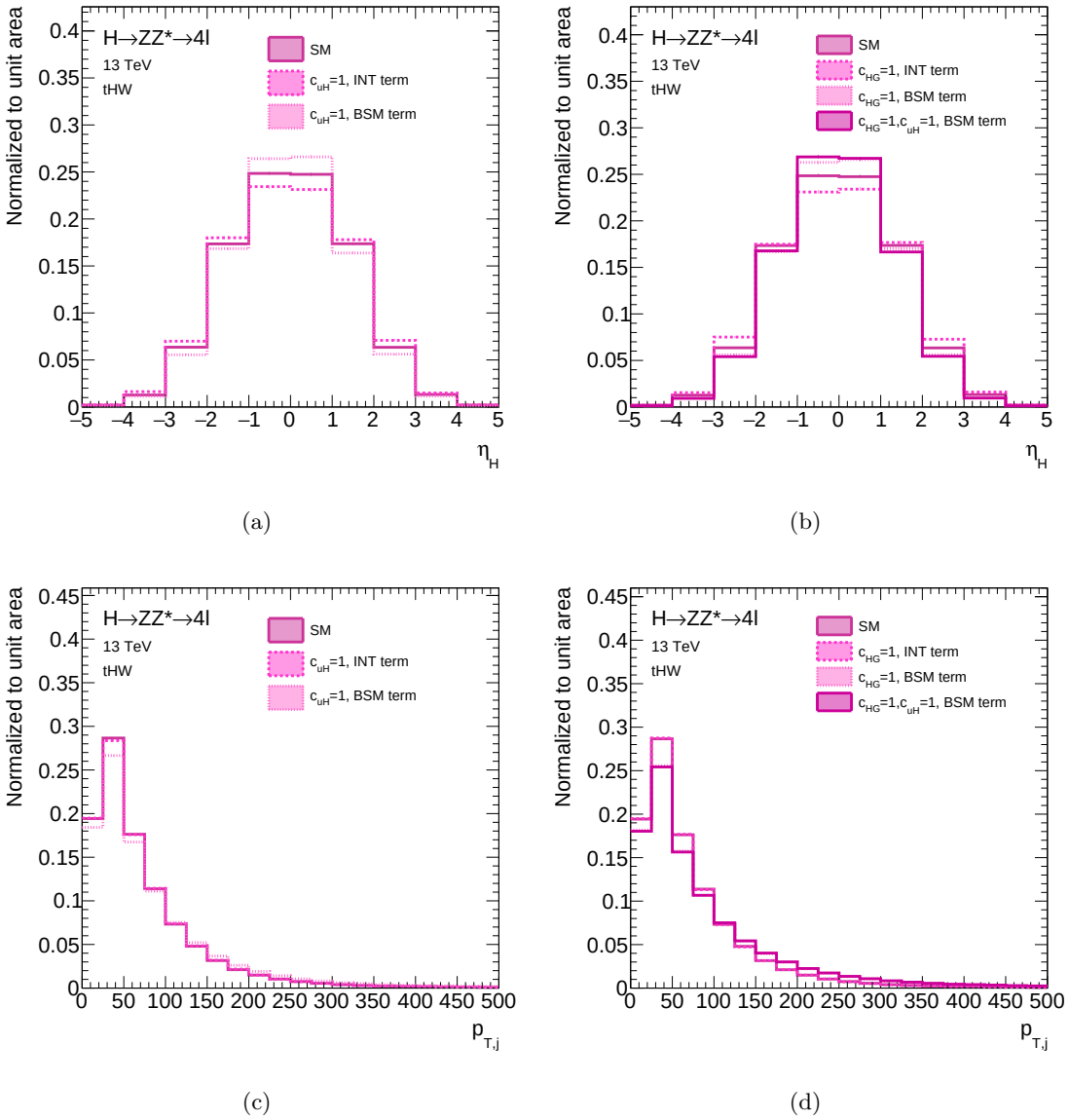
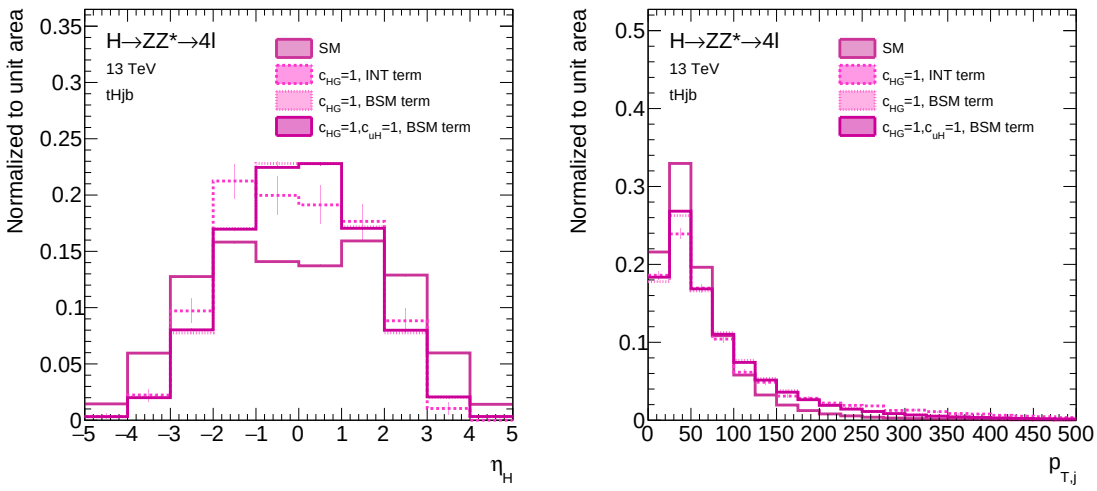


Figure B.7: The expected distributions of (a) and (b) η_H and (c) and (d) p_T^j in the tHW production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{uH} = 1$ and $c_{HG} = 1$, respectively.



(a) (b)

Figure B.8: The expected distributions of (a) η_H and (b) p_T^j in the tHB production mode shown separately for the SM, the interference and BSM term with the BSM coupling parameter $c_{HG} = 1$.

B.4 Calculated Fractions for the Cross Section Parametrisation

Table B.6: Fractions of the cross section in each STXS bins for CP-even BSM coupling parameters. NP1 denote the MadGraph5_aMC@NLO syntax `NP^2==1` and corresponds to the SM-BSM interference term, while NP2 denotes `NP^2==2` and corresponds to the pure BSM term.

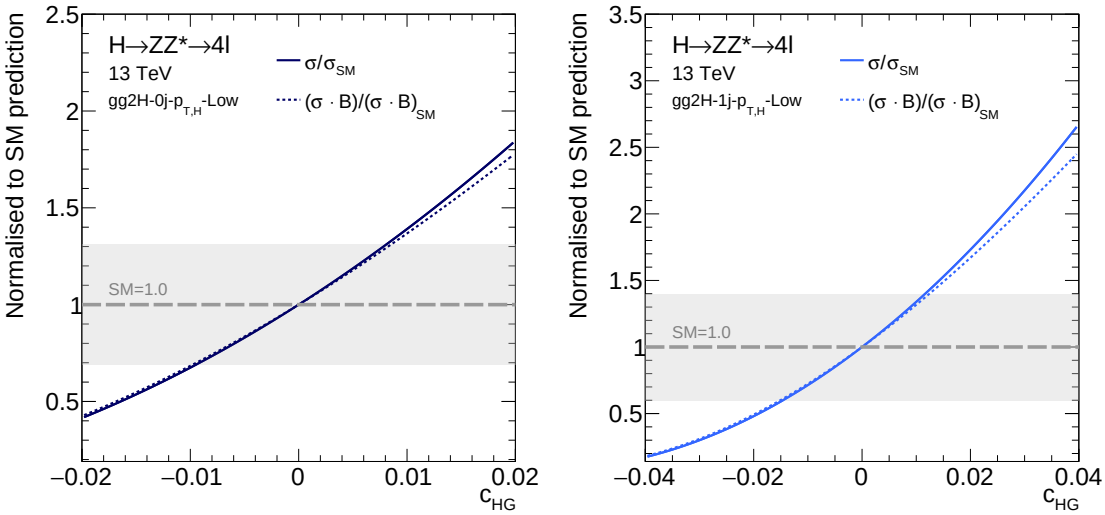
STXS bin ggF	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2							
gg2H-0j-p_T^H-Low	0.1182	0.1415	0.1506							
gg2H-0j-p_T^H-High	0.3271	0.3864	0.4076							
gg2H-1j-p_T^H-Low	0.1097	0.1147	0.1133							
gg2H-1j-p_T^H-Med	0.0924	0.0913	0.0852							
gg2H-1j-p_T^H-High	0.0235	0.0220	0.0207							
gg2H-2j	0.1927	0.1135	0.0895							
gg2H-p_T^H-High	0.0414	0.0246	0.0263							
STXS bin VBF+VH-Had	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2
qq2Hqq-VBF	0.7769	0.3523	0.557	0.3493	0.77	0.6618	0.6987	0.6145	0.5373	0.6289
qq2Hqq-VH-Like	0.0996	0.4537	0.2004	0.5566	0.0522	0.2008	0.1141	0.1631	0.2301	0.1628
qq2Hqq-BSM	0.0393	0.0572	0.1547	−0.0245	0.0826	0.0376	0.0952	0.1331	0.1423	0.117
STXS bin ZH-Lep	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2	$c_{HB}=1$, NP1	$c_{HB}=1$, NP2	$c_{HWB}=1$, NP1	$c_{HWB}=1$, NP2	$c_{HW}=1$, $c_{HB}=1$, NP2	$c_{HW}=1$, $c_{HWB}=1$, NP2	$c_{HB}=1$, $c_{HWB}=1$, NP2
qq/gg2HLep	0.8768	0.8812	0.8997	0.8864	0.9084	0.881	0.8961	0.902	0.8972	0.8968
STXS bin WH-Lep	SM	$c_{HW}=1$, NP1	$c_{HW}=1$, NP2							
qq/gg2HLep	0.866	0.872	0.8929							
STXS bin ttH	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP1	$c_{H\widetilde{G}}=1$, $c_{uH}=1$, NP2				
ttH	0.9828	0.9784	0.9809	0.9842	0.9837	0.9816				
STXS bin $tHjb$	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP1	$c_{H\widetilde{G}}=1$, $c_{uH}=1$, NP2				
ttH	0.9059	1.0375	0.9574	1.0276	0.9789	0.9585				
STXS bin tHW	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{HG}=1$ NP2	$c_{uH}=1$, NP1	$c_{uH}=1$, NP2	$c_{HG}=1$, $c_{uH}=1$, NP2				
ttH	0.9933	0.9872	0.9787	0.9931	0.9908	0.9795				

B.5 Simulated Monte Carlo Samples for the Parametrisation of the Acceptance

Table B.7: Simulated truth level samples with different BSM values for parametrisation of the acceptance with the corresponding value of the acceptance ratio $\mathcal{A}_{\text{BSM}}/\mathcal{A}_{\text{SM}}$.

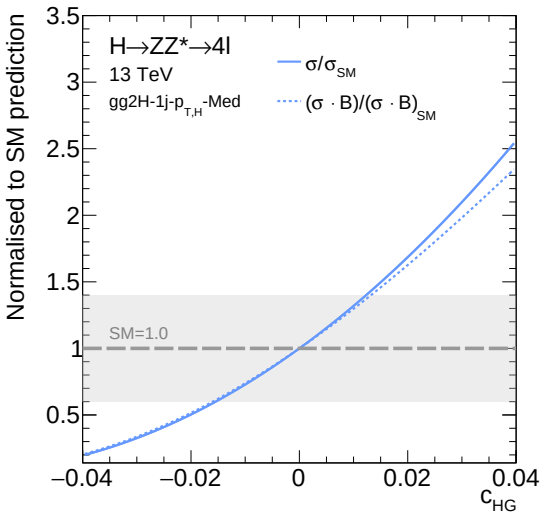
c_{HW}	c_{HB}	c_{HWB}	$\mathcal{A}_{\text{BSM}}/\mathcal{A}_{\text{SM}}$	c_{HW}	c_{HB}	c_{HWB}	$\mathcal{A}_{\text{BSM}}/\mathcal{A}_{\text{SM}}$	c_{HW}	c_{HB}	c_{HWB}	$\mathcal{A}_{\text{BSM}}/\mathcal{A}_{\text{SM}}$
0.0	0.0	0.0	1.0	0.3	−0.3	0	0.838	10	−10	0	0.176
−3	0	0	0.270	0.8	−0.8	0	0.462	3	3	0	0.153
−2	0	0	0.367	0.3	0.3	0	0.843	10	10	0	0.136
−0.8	0	0	0.687	0.8	0.8	0	0.415	−10	0	−10	0.334
−0.5	0	0	0.827	−0.8	0	−0.8	0.925	−3	0	−3	0.543
−0.3	0	0	0.912	−0.3	0	−0.3	0.990	−10	0	+10	0.148
0.3	0	0	0.999	−0.8	0	+0.8	0.425	−3	0	+3	0.173
0.5	0	0	0.937	−0.3	0	+0.3	0.780	3	0	−3	0.179
0.8	0	0	0.804	0.3	0	−0.3	0.898	10	0	−10	0.147
2	0	0	0.400	0.8	0	−0.8	0.488	3	0	3	0.530
3	0	0	0.288	0.3	0	0.3	0.991	10	0	10	0.328
5	0	0	0.214	0.8	0	0.8	0.920	0	−10	−10	0.158
10	0	0	0.183	0	−0.8	−0.8	0.564	0	−3	−3	0.209
0	−3	0	0.192	0	−0.3	−0.3	0.914	0	−10	+10	0.146
0	−2	0	0.236	0	−0.8	+0.8	0.363	0	−3	+3	0.175
0	−0.8	0	0.490	0	−0.3	+0.3	0.732	0	3	−3	0.159
0	−0.5	0	0.677	0	0.3	−0.3	0.768	0	10	−10	0.147
0	−0.3	0	0.844	0	0.8	−0.8	0.361	0	3	3	0.179
0	0.3	0	0.842	0	0.3	0.3	0.840	0	10	10	0.143
0	0.5	0	0.672	0	0.8	0.8	0.492	−10	−10	−10	0.142
0	0.8	0	0.458	−0.8	−0.8	−0.8	0.615	10	−10	−10	0.168
0	2	0	0.215	0.8	−0.8	−0.8	0.425	−10	10	−10	0.173
0	3	0	0.173	−0.8	0.8	−0.8	0.391	−10	−10	10	0.142
0	5	0	0.156	−0.8	−0.8	0.8	0.299	10	10	−10	0.134
0	10	0	0.150	0.8	0.8	−0.8	0.301	10	−10	10	0.168
0	0	−3	0.261	0.8	−0.8	0.8	0.404	−10	10	10	0.155
0	0	−2	0.381	−0.8	0.8	0.8	0.360	10	10	10	0.139
0	0	−0.8	0.790	0.8	0.8	0.8	0.589	−3	−3	−3	0.220
0	0	−0.5	0.928	−0.3	−0.3	−0.3	0.910	3	−3	−3	0.187
0	0	−0.3	0.988	0.3	−0.3	−0.3	0.831	−3	3	−3	0.185
0	0	0.3	0.908	−0.3	0.3	−0.3	0.776	−3	−3	3	0.152
0	0	0.5	0.815	−0.3	−0.3	0.3	0.654	3	3	−3	0.145
0	0	0.8	0.671	0.3	0.3	−0.3	0.726	3	−3	3	0.195
0	0	2	0.337	0.3	−0.3	0.3	0.782	−3	3	3	0.178
0	0	3	0.238	−0.3	0.3	0.3	0.921	3	3	3	0.181
0	0	5	0.178	0.3	0.3	0.3	0.901				
0	0	10	0.149	−10	−10	0	0.142				
−0.8	−0.8	0	0.428	−3	−3	0	0.171				
−0.3	−0.3	0	0.784	−10	+10	0	0.172				
−0.8	+0.8	0	0.415	−3	+3	0	0.193				
−0.3	+0.3	0	0.782	3	−3	0	0.203				

B.6 EFT Parametrisation



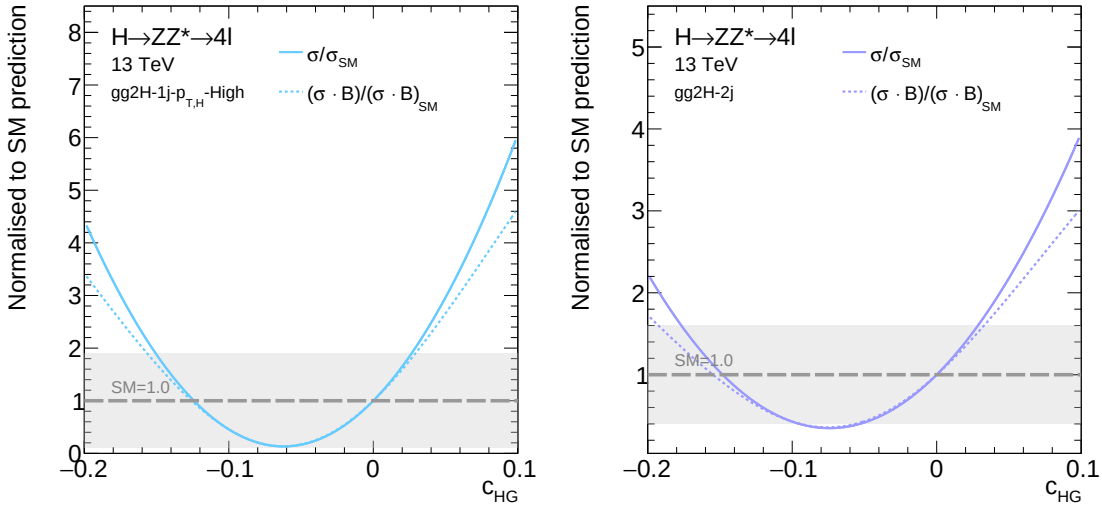
(a)

(b)



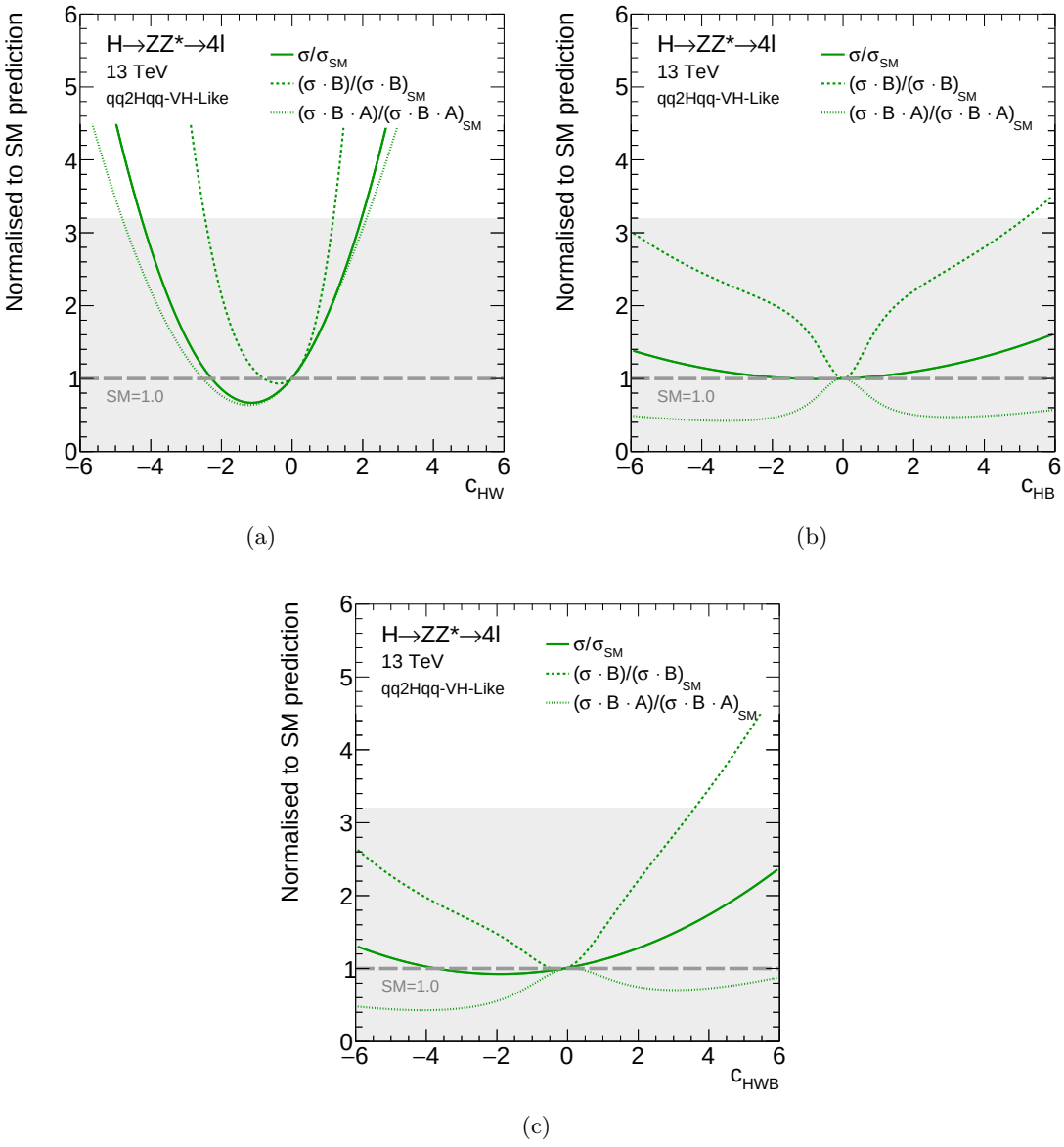
(c)

Figure B.9: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter c_{HG} in the (a) gg2H-0j-p_T^H-Low, (b) gg2H-1j-p_T^H-Lowand (c) gg2H-1j-p_T^H-Med production bin. The grey band indicates the expected sensitivity at 1σ level from the SM production cross section measurement.



(a) (b)

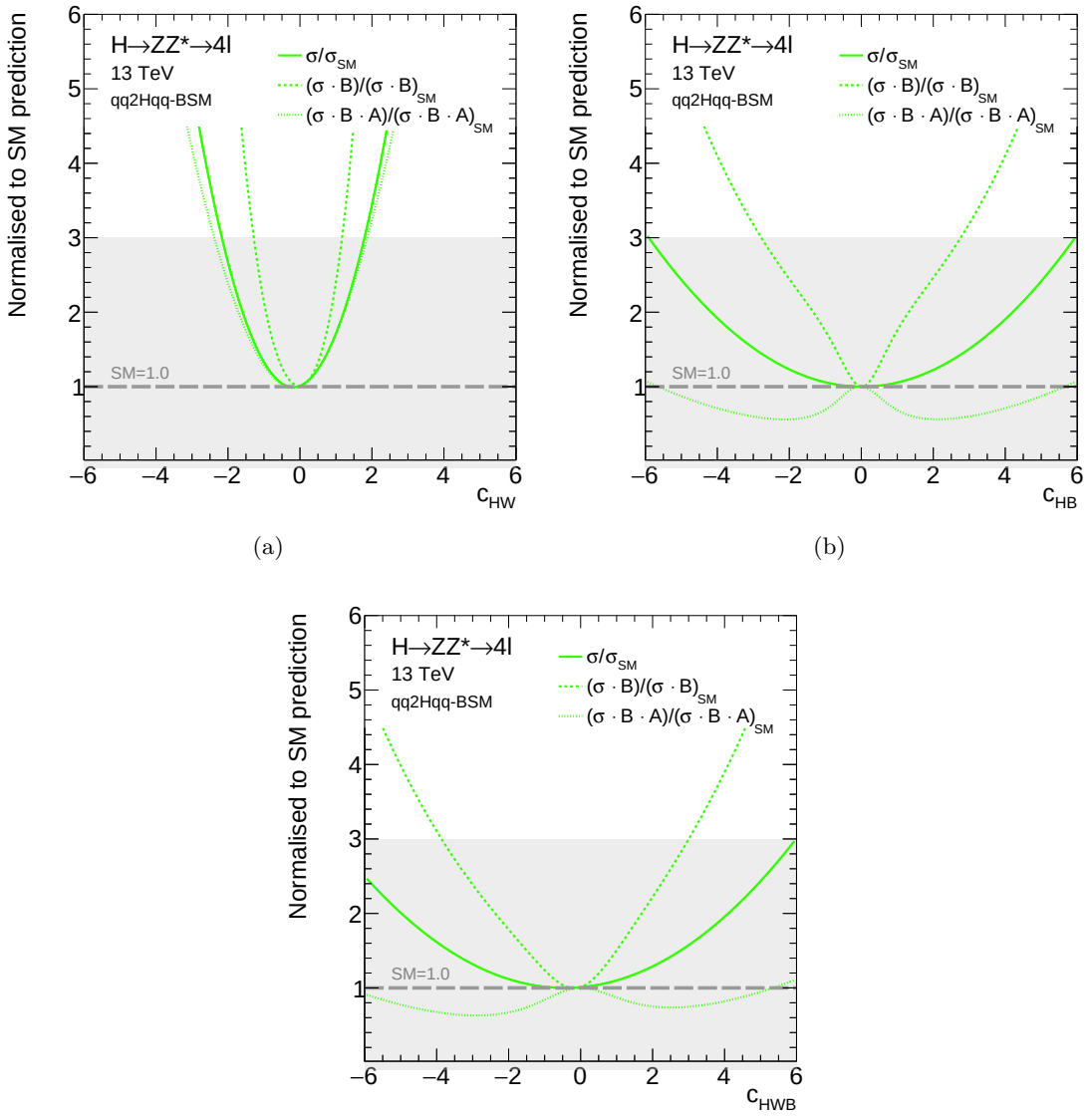
Figure B.10: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter c_{HG} in the (a) gg2H-1j-p_T^H-High and (b) gg2H-2j production bin. The grey band indicates the expected sensitivity at 1σ level from the SM production cross section measurement.



(a) (b)

(c)

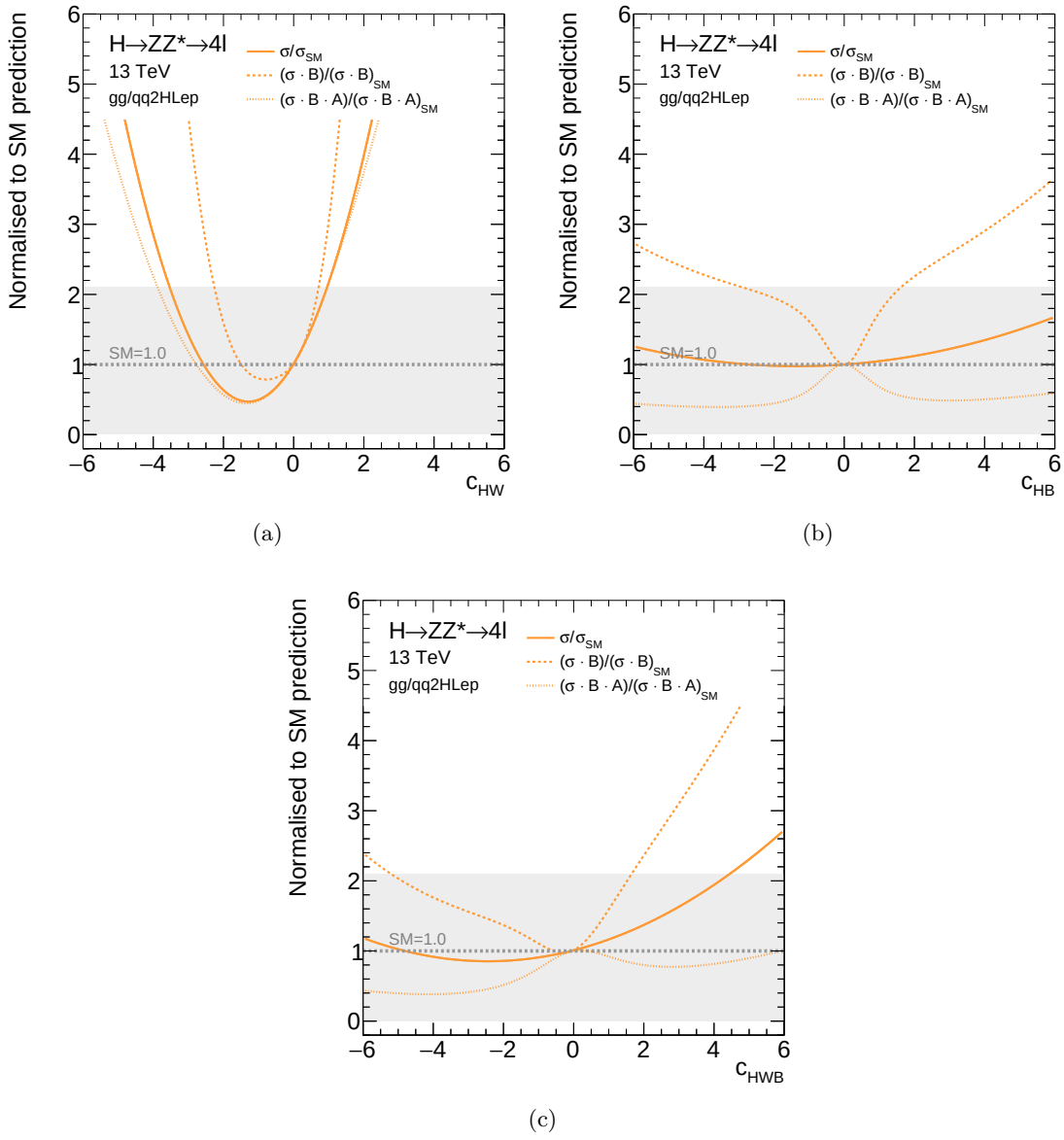
Figure B.11: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter (a) c_{HW}, (b) c_{HB} and (c) c_{HWB} in the qq2Hqq-VH-Like production bin. The grey band indicates the expected sensitivity at 1σ level from the SM production cross section measurement.



(a) (b)

(c)

Figure B.12: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter (a) c_{HW}, (b) c_{HB} and (c) c_{HWB} in the qq2Hqq-BSM production bin. The grey band indicates the expected sensitivity at 1σ level from the SM production cross section measurement.

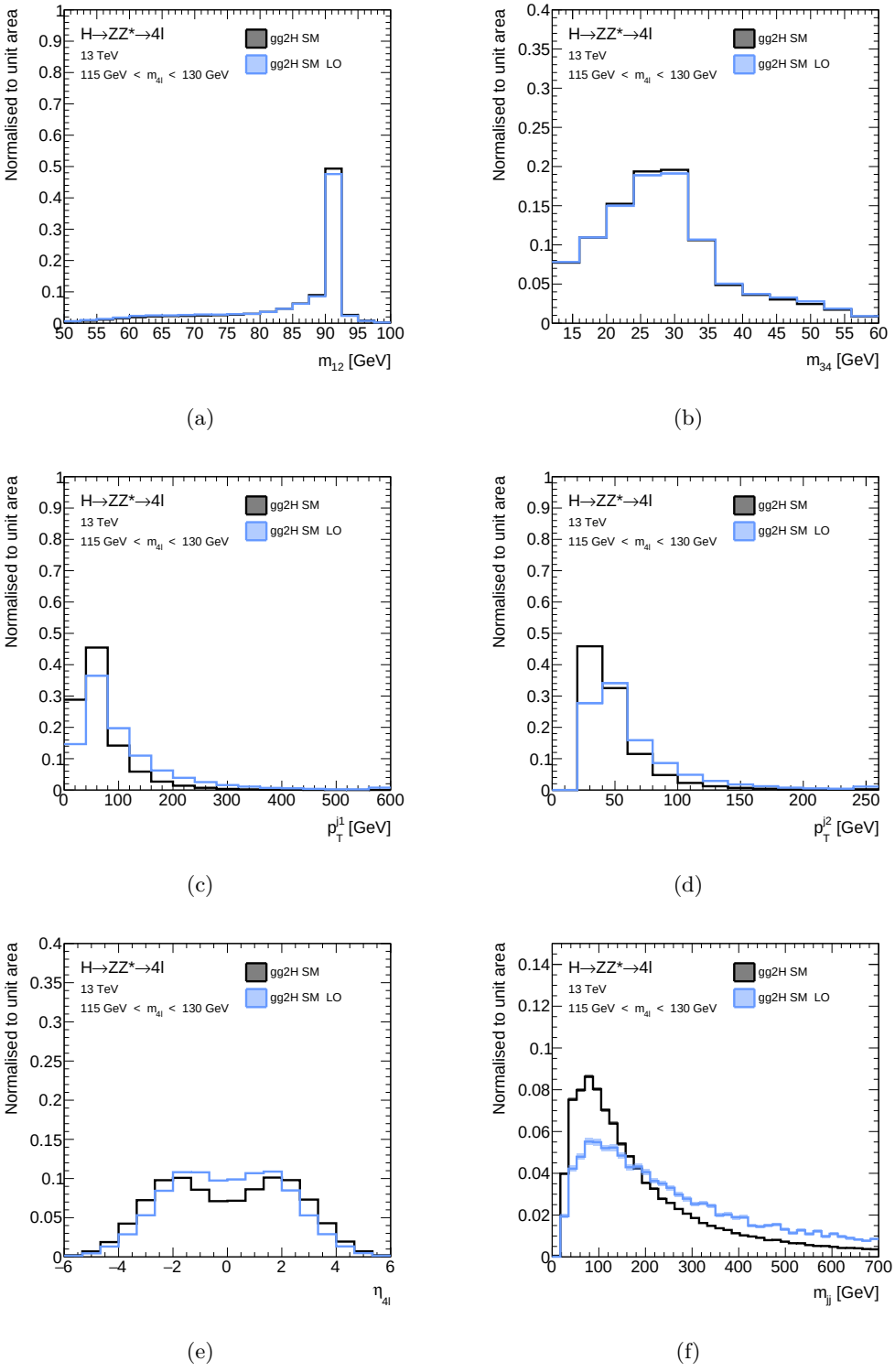


(a) (b)

(c)

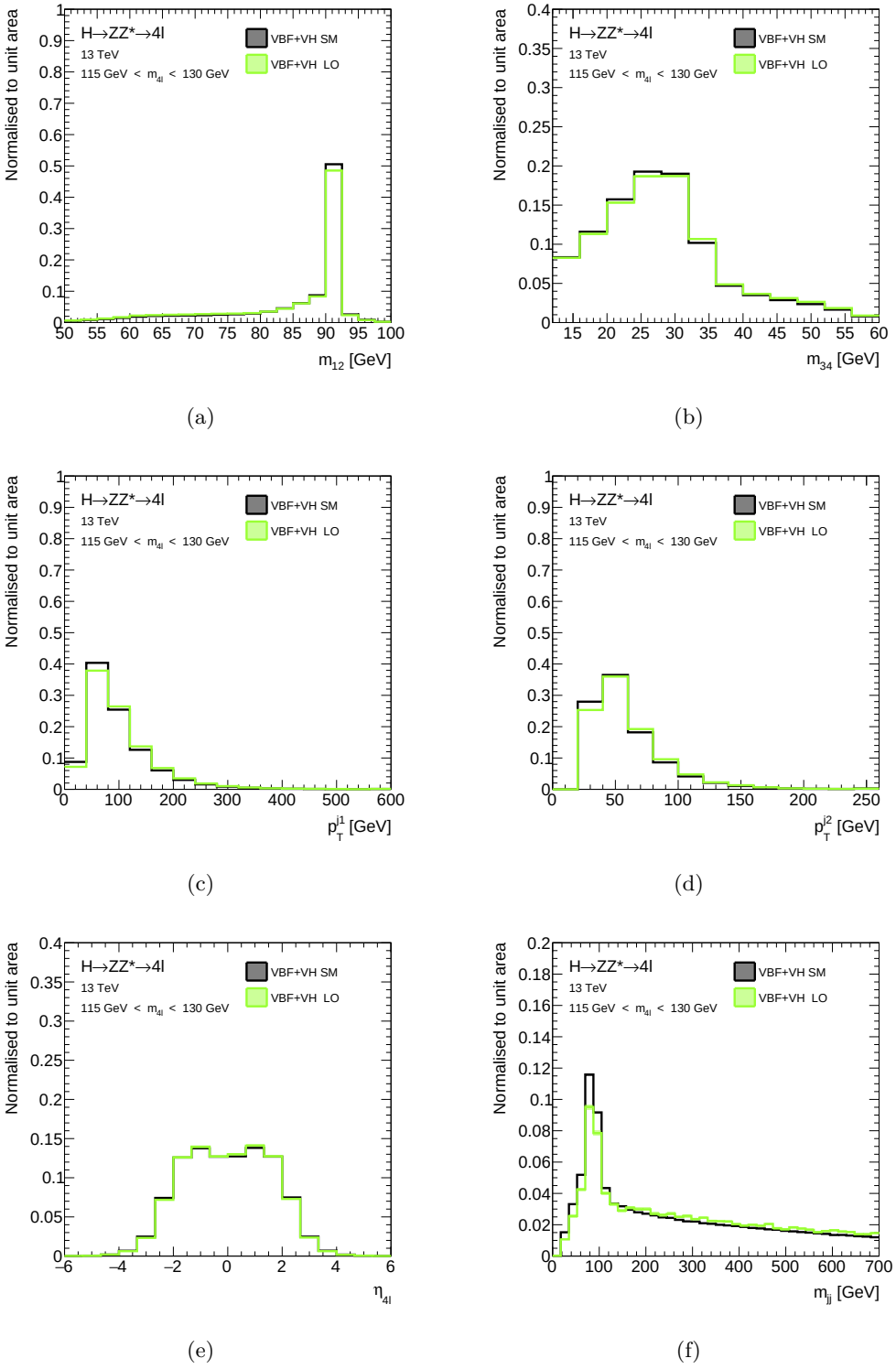
Figure B.13: The expected dependence of the event yield relative to the SM prediction in dependence on the BSM coupling parameter (a) c_{HW}, (c) c_{HB} and (c) c_{HWB} in the qq/gg2HLep production bin. The grey band indicates the expected sensitivity at 1σ level from the SM production cross section measurement.

B.7 SMEFT Signal Samples



(a) (b) (c) (d) (e) (f)

Figure B.14: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for the SM NNLO and SM SMEFT LO sample for gg2H production.



(a) (b)

(c) (d)

(e) (f)

Figure B.15: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for the SM NLO and SM SMEFT LO sample for VBF+VH production.

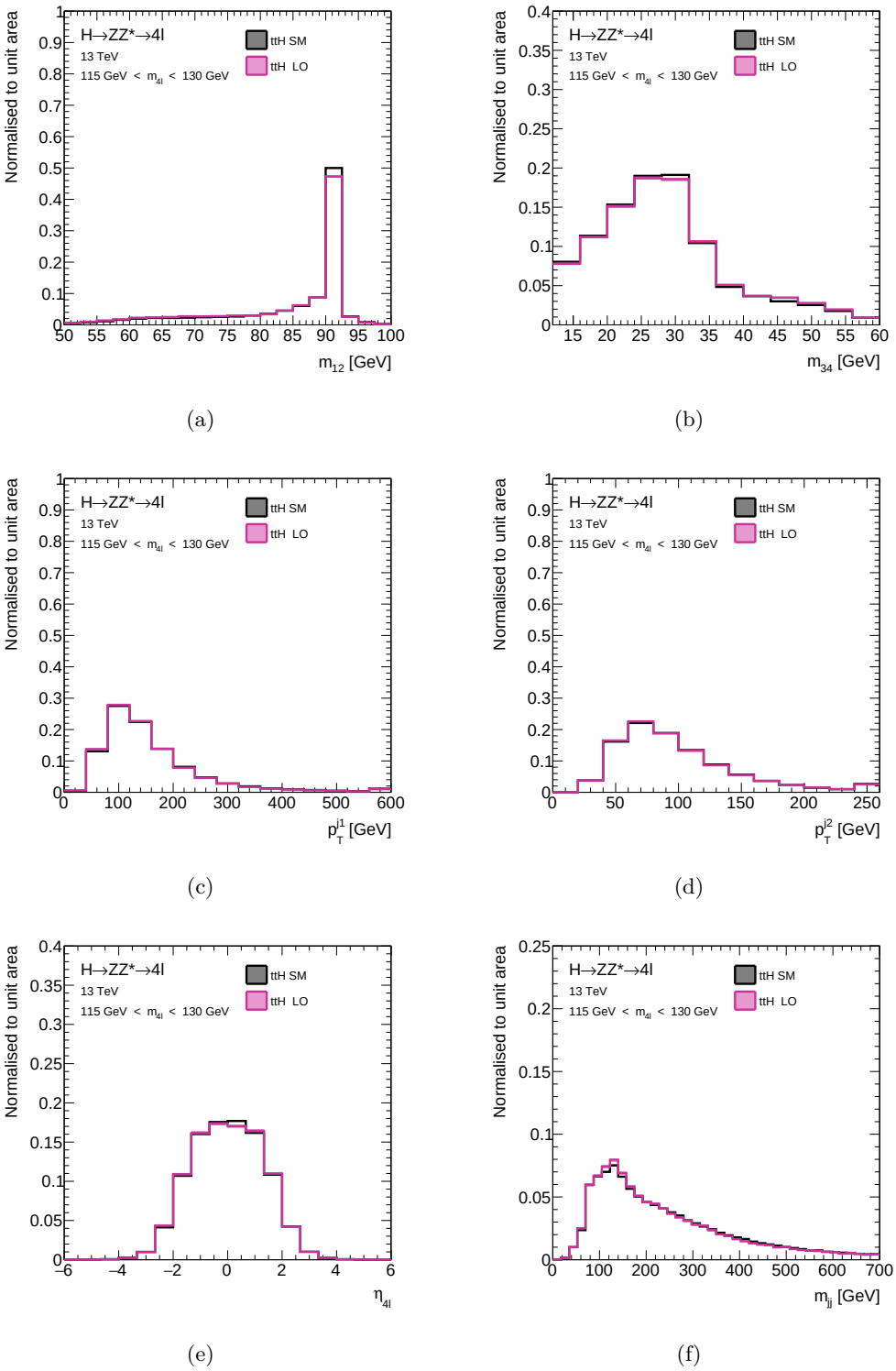


Figure B.16: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for the SM NLO and SM SMEFT LO sample for ttH production.

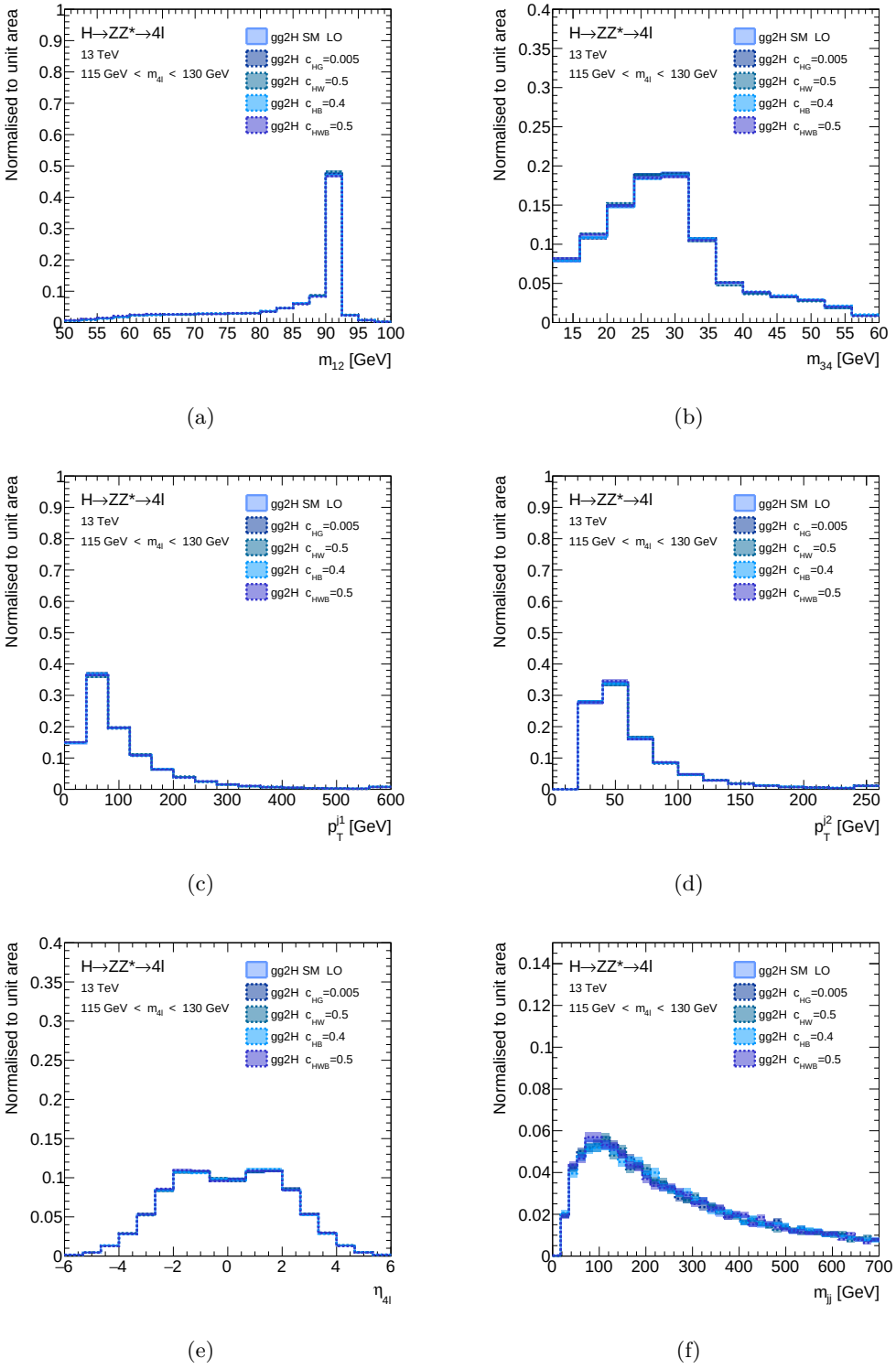
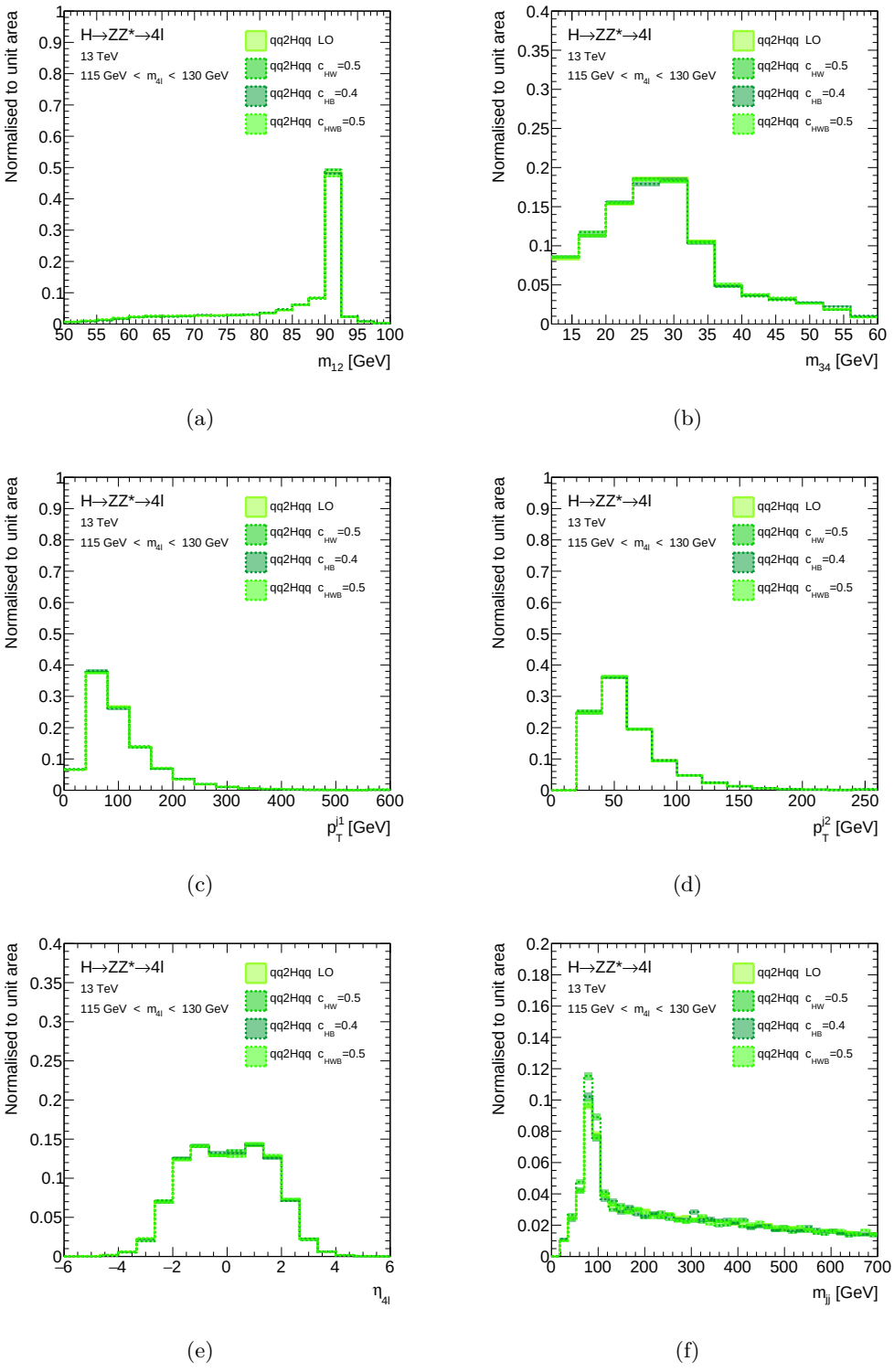
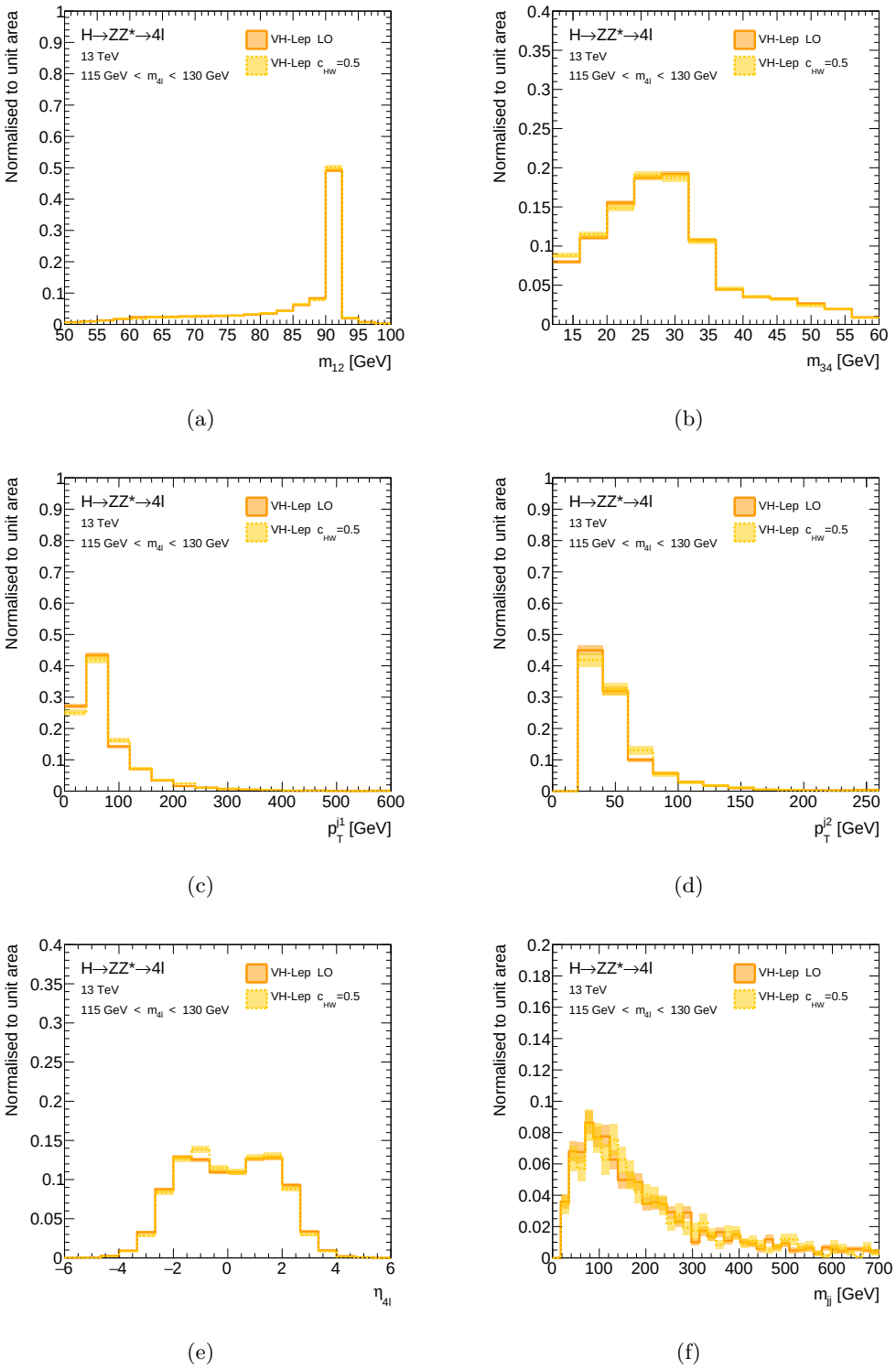


Figure B.17: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for SM and selected BSM Higgs boson signals for gg2H production.



(a) (b) (c) (d) (e) (f)

Figure B.18: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for SM and selected BSM Higgs boson signals for qq2Hqq production.

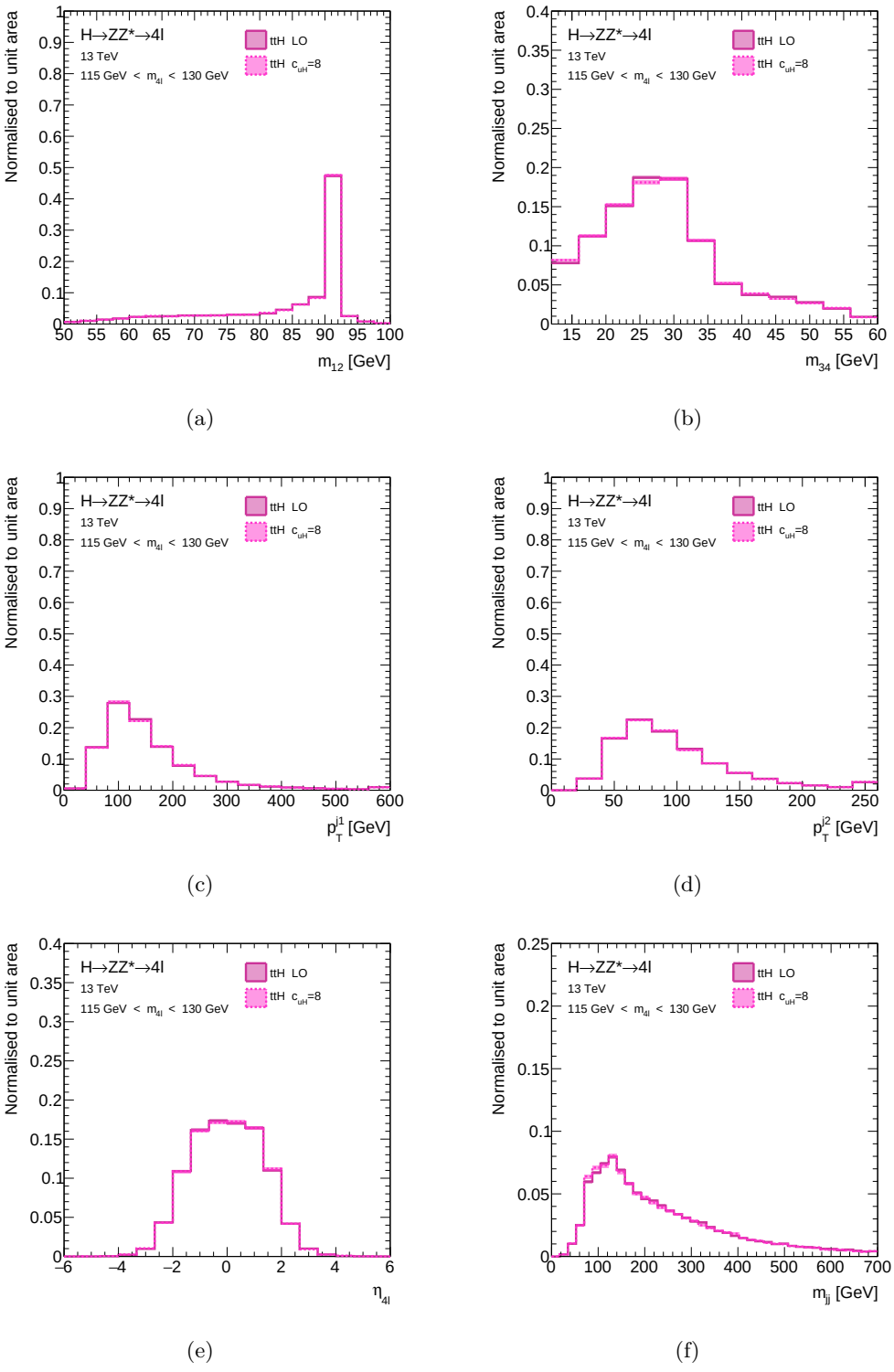


(a) (b)

(c) (d)

(e) (f)

Figure B.19: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for SM and selected BSM Higgs boson signals for VH-Lep production.



(a) (b) (c) (d) (e) (f)

Figure B.20: Expected distribution of (a) m_{12}, (b) m_{34}, (c) p_T^{j1}, (d) p_T^{j2}, (e) $\eta_{4\ell}$ and (f) m_{jj} for the SM and several BSM signal samples for ttH production.

C

EFT Parametrisation with CP-odd BSM Coupling Parameters

C.1 Generated BSM Signal Samples

Table C.1: Configuration of the simulated BSM samples with MadGraph5_aMC@NLO for all production processes taken into account

Monte Carlo sample type		$c_{H\widetilde{G}}$	c_{uH}	$c_{H\widetilde{W}}$	$c_{H\widetilde{B}}$	$c_{H\widetilde{W}B}$	Syntax
ggF+bbH	SM	0	0	0	0	0	`NP^2==0`
	BSM term $c_{H\widetilde{G}}$	1	0	0	0	0	`NP^2==2`
VBF+VH-Had and VH-Lep	SM	0	0	0	0	0	`NP^2==0`
	BSM term $c_{H\widetilde{W}}$	0	0	1	0	0	`NP^2==2`
	BSM term $c_{H\widetilde{B}}$	0	0	0	1	0	`NP^2==2`
	BSM term $c_{H\widetilde{W}B}$	0	0	0	0	1	`NP^2==2`
	BSM term $c_{H\widetilde{W}}$, $c_{H\widetilde{B}}$	0	0	1	1	0	`NP^2==2`
	BSM term $c_{H\widetilde{W}}$, $c_{H\widetilde{W}B}$	0	0	1	0	1	`NP^2==2`
	BSM term $c_{H\widetilde{B}}$, $c_{H\widetilde{W}B}$	0	0	0	1	1	`NP^2==2`
WH-Lep	SM	0	0	0	0	0	`NP^2==0`
	BSM term $c_{H\widetilde{W}}$	0	0	1	0	0	`NP^2==2`
ttH, $tHjb$ and tHW	SM	0	0	0	0	0	`NP^2==0`
	BSM term $c_{H\widetilde{G}}$	1	0	0	0	0	`NP^2==2`
	BSM term $c_{\tilde{u}H}$	0	1	0	0	0	`NP^2==2`
	BSM term $c_{H\widetilde{G}}$, $c_{\tilde{u}H}$	1	1	0	0	0	`NP^2==2`

C.2 Production Cross Sections, Partial Decay and Total Width of the Higgs Boson

Table C.2: Inclusive cross section for all production modes as well as the total and partial Higgs boson decay width for BSM CP-odd coupling parameters. NP1 denote the MadGraph5_aMC@NLO syntax `NP^2==1` and corresponds to the SM-BSM interference term, while NP2 denotes `NP^2==2` and corresponds to the pure BSM term.

ggF+bbH	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2								
Cross section [pb]	27.6	$<10^{-3}$	7048.6								
VBF+VH-Had	SM	$c_{H\widetilde{W}}=1$, NP1	$c_{H\widetilde{W}}=1$, NP2	$c_{H\widetilde{B}}=1$, NP1	$c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}B}=1$, NP1	$c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{B}}=1$, $c_{H\widetilde{W}B}=1$, NP2	
Cross section [pb]	4.1503	$<10^{-4}$	0.3900	$<10^{-4}$	0.1068	$<10^{-4}$	0.0673	0.5530	0.3821	0.0756	
ZH-Lep	SM	$c_{H\widetilde{W}}=1$, NP1	$c_{H\widetilde{W}}=1$, NP2	$c_{H\widetilde{B}}=1$, NP1	$c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}B}=1$, NP1	$c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{B}}=1$, $c_{H\widetilde{W}B}=1$, NP2	
Cross section [pb]	0.1505	$<10^{-4}$	0.0247	$<10^{-4}$	0.0050	$<10^{-4}$	0.0057	0.0294	0.0384	0.0139	
WH-Lep	SM	$c_{H\widetilde{W}}=1$, NP1	$c_{H\widetilde{W}}=1$, NP2								
Cross section [pb]	0.2382	$<2\cdot10^{-4}$	0.0528								
ttH	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2	$c_{\tilde{u}H}=1$, NP1	$c_{\tilde{u}H}=1$, NP1	$c_{H\widetilde{G}}=1$, $c_{\tilde{u}H}=1$, NP2					
Cross section [pb]	0.4128	$<10^{-4}$	0.2850	$<10^{-4}$	0.0006	0.3064					
$tHjb$	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2	$c_{\tilde{u}H}=1$, NP1	$c_{\tilde{u}H}=1$, NP1	$c_{H\widetilde{G}}=1$, $c_{\tilde{u}H}=1$, NP2					
Cross section [pb]	0.06476	–	0.00268	–	0.00024	0.00292					
tHW	SM	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$ NP2	$c_{\tilde{u}H}=1$, NP1	$c_{\tilde{u}H}=1$, NP2	$c_{H\widetilde{G}}=1$, $c_{\tilde{u}H}=1$, NP2					
Cross section [pb]	0.01651	–	0.00597	–	0.00012	0.00646					
$H\to4\ell$	SM	$c_{H\widetilde{W}}=1$, NP1	$c_{H\widetilde{W}}=1$, NP2	$c_{H\widetilde{B}}=1$, NP1	$c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}B}=1$, NP1	$c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{B}}=1$, $c_{H\widetilde{W}B}=1$, NP2	
Width [MeV]	$5.02\cdot10^{-7}$	$<10^{-9}$	$3.52\cdot10^{-7}$	$<10^{-9}$	$1.30\cdot10^{-6}$	$<10^{-9}$	$4.07\cdot10^{-7}$	$1.68\cdot10^{-6}$	$9.43\cdot10^{-8}$	$9.85\cdot10^{-7}$	
$H\to all$	SM	$c_{H\widetilde{W}}=1$, NP1	$c_{H\widetilde{W}}=1$, NP2	$c_{H\widetilde{B}}=1$, NP1	$c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}B}=1$, NP1	$c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{G}}=1$, NP1	$c_{H\widetilde{G}}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{B}}=1$, NP2	$c_{H\widetilde{W}}=1$, $c_{H\widetilde{W}B}=1$, NP2	$c_{H\widetilde{B}}=1$, $c_{H\widetilde{W}B}=1$, NP2
Width [MeV]	$5.00\cdot10^{-3}$	$<10^{-9}$	$8.07\cdot10^{-4}$	$<10^{-9}$	$6.03\cdot10^{-3}$	$<10^{-10}$	$1.78\cdot10^{-3}$	$<10^{-9}$	$7.44\cdot10^{-2}$	$9.42\cdot10^{-3}$	$3.90\cdot10^{-4}$	$2.17\cdot10^{-3}$

C.3 Calculated Fractions for the Cross Section Parametrisation

Table C.3: Fractions of the cross section in each STXS bins for BSM CP-odd coupling parameters. NP1 denote the MadGraph5_aMC@NLO syntax `NP^2==1` and corresponds to the SM-BSM interference term, while NP2 denotes `NP^2==2` and corresponds to the pure BSM term.

STXS bin ggF	SM	$c_{H\widetilde{G}} = 1$ NP2					
gg2H-0j-p_T^H-Low	0.1194	0.1488					
gg2H-0j-p_T^H-High	0.3255	0.4084					
gg2H-1j-p_T^H-Low	0.1075	0.1148					
gg2H-1j-p_T^H-Med	0.0935	0.0874					
gg2H-1j-p_T^H-High	0.0231	0.0212					
gg2H-2j	0.1906	0.0852					
gg2H-p_T^H-High	0.0434	0.0276					
STXS bin VBF+VH-Had	SM	$c_{H\widetilde{W}} = 1$, NP2	$c_{H\widetilde{B}} = 1$, NP2	$c_{H\widetilde{W}B} = 1$, NP2	$c_{H\widetilde{W}} = 1$, $c_{H\widetilde{B}} = 1$, NP2	$c_{H\widetilde{W}} = 1$, $c_{H\widetilde{W}B} = 1$, NP2	$c_{H\widetilde{B}} = 1$, $c_{H\widetilde{W}B} = 1$, NP2
qq2Hqq-VBF	0.7806	0.5127	0.7683	0.7138	0.5964	0.4874	0.6231
qq2Hqq-VH-Like	0.0975	0.1986	0.0475	0.0746	0.1507	0.2264	0.1325
qq2Hqq-BSM	0.0385	0.2107	0.0875	0.1241	0.1684	0.2089	0.1633
STXS bin ZH-Lep	SM	$c_{H\widetilde{W}} = 1$, NP2	$c_{H\widetilde{B}} = 1$, NP2	$c_{H\widetilde{W}B} = 1$, NP2	$c_{H\widetilde{W}} = 1$, $c_{H\widetilde{B}} = 1$, NP2	$c_{H\widetilde{W}} = 1$, $c_{H\widetilde{W}B} = 1$, NP2	$c_{H\widetilde{B}} = 1$, $c_{H\widetilde{W}B} = 1$, NP2
qq/gg2HLep	0.8763	0.9158	0.9512	0.9432	0.9219	0.9092	0.9313
STXS bin WH-Lep	SM	$c_{H\widetilde{W}} = 1$, NP1	$c_{H\widetilde{W}} = 1$, NP2				
qq/gg2HLep	0.8689	0.9043					
STXS bin ttH	SM	$c_{H\widetilde{G}} = 1$ NP2	$c_{\tilde{u}H} = 1$, NP2	$c_{H\widetilde{G}} = 1$, $c_{\tilde{u}H} = 1$, NP2			
ttH	0.9835	0.9822	0.9909	0.9821			
STXS bin $tHjb$	SM	$c_{H\widetilde{G}} = 1$ NP2	$c_{\tilde{u}H} = 1$, NP2	$c_{H\widetilde{G}} = 1$, $c_{\tilde{u}H} = 1$, NP2			
ttH	0.9059	0.9582	0.9828	0.9608			
STXS bin tHW	SM	$c_{H\widetilde{G}} = 1$ NP2	$c_{\tilde{u}H} = 1$, NP2	$c_{H\widetilde{G}} = 1$, $c_{\tilde{u}H} = 1$, NP2			
ttH	0.9933	0.9791	0.9909	0.9797			

C.4 EFT Parametrisation

Table C.4: EFT parametrisation of the cross section ratio $\sigma/\sigma_{\mathrm{SM}}$ for each of particle-level production bin of the Reduced-Stage-1.1 scheme and of the ratio of decay widths $\Gamma/\Gamma_{\mathrm{SM}}$ in dependence on the CP-odd BSM coupling parameters

Production bin	Cross section parametrisation, $\sigma/\sigma_{\mathrm{SM}}$
gg2H-0j-p_T^H-Low	$1+317.69c^2_{H\widetilde{G}}$
gg2H-0j-p_T^H-High	$1+319.85c^2_{H\widetilde{G}}$
gg2H-1j-p_T^H-Low	$1+272.23c^2_{H\widetilde{G}}$
gg2H-1j-p_T^H-Med	$1+238.29c^2_{H\widetilde{G}}$
gg2H-1j-p_T^H-High	$1+233.95c^2_{H\widetilde{G}}$
gg2H-2j	$1+113.95c^2_{H\widetilde{G}}$
gg2H-p_T^H-High	$1+162.11c^2_{H\widetilde{G}}$
qq2Hqq-VH-Like	$1 + 0.1912c^2_{HW} + 0.0125c^2_{H\widetilde{B}} + 0.0124c^2_{H\widetilde{W}B} + 0.0022c_{HW}c_{H\widetilde{B}} + 0.0101c_{HW}c_{H\widetilde{W}B} - 0.0002c_{H\widetilde{B}}c_{H\widetilde{W}B}$
qq2Hqq-BSM	$1 + 0.5138c^2_{HW} + 0.0585c^2_{H\widetilde{B}} + 0.0523c^2_{H\widetilde{W}B} + 0.0105c_{HW}c_{H\widetilde{B}} - 0.0665c_{HW}c_{H\widetilde{W}B} - 0.0334c_{H\widetilde{B}}c_{H\widetilde{W}B}$
qq2Hqq-VBF	$1 + 0.0617c^2_{HW} + 0.0253c^2_{H\widetilde{B}} + 0.0148c^2_{H\widetilde{W}B} + 0.0148c_{HW}c_{H\widetilde{B}} - 0.0190c_{HW}c_{H\widetilde{W}B} - 0.0256c_{H\widetilde{B}}c_{H\widetilde{W}B}$
qq/gg2HLep	$1 + 0.2075c^2_{HW} + 0.0141c^2_{H\widetilde{B}} + 0.0158c^2_{H\widetilde{W}B} - 0.0009c_{HW}c_{H\widetilde{B}} + 0.0203c_{HW}c_{H\widetilde{W}B} + 0.0083c_{H\widetilde{B}}c_{H\widetilde{W}B}$
ttH/tH	$1+0.599c^2_{H\widetilde{G}}+0.0020c^2_{uH}+0.043c_{H\widetilde{G}}c_{uH}$
Decay process	**Decay width parametrisation, $\Gamma/\Gamma_{\mathrm{SM}}$**
$\Gamma^{4\ell}/\Gamma^{4\ell}_{\mathrm{SM}}$	$1 + 0.702c^2_{HW} + 2.586c^2_{H\widetilde{B}} + 0.811c^2_{H\widetilde{W}B} + 0.050c_{HW}c_{H\widetilde{B}} - 1.325c_{HW}c_{H\widetilde{W}B} - 1.434c_{H\widetilde{B}}c_{H\widetilde{W}B}$
$\Gamma^{\mathrm{tot}}/\Gamma^{\mathrm{tot}}_{\mathrm{SM}}$	$1 + 0.161c^2_{HW} + 1.206c^2_{H\widetilde{B}} + 0.357c^2_{H\widetilde{W}B} 14.888c^2_{H\widetilde{G}} + 0.518c_{HW}c_{H\widetilde{B}} - 0.440c_{HW}c_{H\widetilde{W}B} - 1.129c_{H\widetilde{B}}c_{H\widetilde{W}B}$

C.5 EFT Parametrisation of the Signal Acceptance

Table C.5: Values of acceptance parameters obtained from the fit of the three-dimensional Lorentzian function (Equation 7.31) to the simulated BSM CP-odd predictions

Acceptance parameter	Fit result	Acceptance parameter	Fit result
α_0	0.118 ± 0.001	$\delta_{H\widetilde{W}}$	0.572 ± 0.003
α_1	0.853 ± 0.001	$\delta_{H\widetilde{B}}$	2.022 ± 0.004
α_2	0.826 ± 0.002	$\delta_{H\widetilde{W}B}$	0.644 ± 0.001
$\beta_{H\widetilde{W}}$	-0.001 ± 0.001	$\delta_{H\widetilde{W},H\widetilde{W}B}$	-1.070 ± 0.004
$\beta_{H\widetilde{B}}$	-0.001 ± 0.001	$\delta_{H\widetilde{B},H\widetilde{W}B}$	-1.085 ± 0.006
$\beta_{H\widetilde{W}B}$	0.001 ± 0.001	$\delta_{H\widetilde{W},H\widetilde{B}}$	-0.010 ± 0.008
		$\delta_{H\widetilde{W},H\widetilde{B},H\widetilde{W}B}$	-0.060 ± 0.088

List of Figures

List of Tables

Bibliography

[1] M. E. Peskin and D. V. Schroeder,
An Introduction To Quantum Field Theory (Frontiers in Physics),
Westview Press, 1995.

[2] M. D. Schwartz, *Quantum Field Theory and the Standard Model*,
Cambridge University Press, 2014.

[3] L. H. Ryder, *Quantum Field Theory 2nd ed.* Cambridge University Press, 1996.

[4] D. H. Perkins, *Introduction to High Energy Physics 4th ed.*
Cambridge Universisty Press, 2000.

[5] M. Böhm, A. Denner and H. Joos,
Gauge theories of strong and electroweak interactions; 3rd ed.
Stuttgart: B. G. Teubner, 2001.

[6] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*,
Physics textbook, New York, NY: Wiley, 2008.

[7] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*,
Annalen der Physik **354**, 1916 769.

[8] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13**, 9 1964 321.

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
Phys. Rev. Lett. **13**, 16 1964 508.

[10] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*,
Phys. Lett. B **716**, 2012 1, arXiv: `1207.7214 [hep-ex]`.

[11] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716**, 2012 30,
arXiv: `1207.7235 [hep-ex]`.

[12] S. L. Glashow, *Partial Symmetries of Weak Interactions*,
Nucl. Phys. **22**, 1961 579.

[13] A. Salam, *Weak and Electromagnetic Interactions*,
Conf. Proc. **C680519**, 1968 367.

[14] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 21 1967 1264.

[15] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12**, 1964 132.

[16] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 4 1966 1156.

[17] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13**, 20 1964 585.

[18] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155**, 5 1967 1554.

[19] J. Bernstein, *Spontaneous symmetry breaking, gauge theories, the Higgs mechanism and all that*, Rev. Mod. Phys. **46**, 1 1974 7.

[20] G. '. Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44**, 1972 189.

[21] G. 'tHooft, *Renormalization of massless Yang-Mills fields*, Nucl. Phys. B **33**, 1971 173.

[22] G. Hooft, *Renormalizable Lagrangians for massive Yang-Mills fields*, Nucl. Phys. B **35**, 1971 167.

[23] Y. Nambu and G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I*, Phys. Rev. **122**, 1 1961 345.

[24] *Particle data group Live*, `http://pdglive.lbl.gov/` (visited on 02/15/2019).

[25] H. E. Logan, *TASI 2013 lectures on Higgs physics within and beyond the Standard Model*, 2014, arXiv: `1406.1786 [hep-ph]`.

[26] ATLAS Collaboration, G. Aad et al., *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B **726**, 2013 120, arXiv: `1307.1432 [hep-ex]`.

[27] CMS Collaboration, V. Khachatryan et al., *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**, 2015 012004, arXiv: `1411.3441 [hep-ex]`.

[28] ATLAS Collaboration, G. Aad et al., *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. C **75**, 2015 476, arXiv: `1506.05669 [hep-ex]`.

[29] LHC Higgs Cross Section Working Group, J. R. A. et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, 2013, arXiv: `1307.1347 [hep-ph]`.

[30] A. Djouadi, *The anatomy of electroweak symmetry breaking: Tome I: The Higgs boson in the Standard Model*, Phys. Rep. **457**, 2008 1.

[31] *LHC Higgs Cross Section Working Group*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.

[32] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, JHEP **12**, 2017 070, arXiv: 1709.06492 [hep-ph].

[33] P. Artoisenet et al., *A framework for Higgs characterisation*, JHEP **1311**, 2013 043, arXiv: 1306.6464 [hep-ph].

[34] B. Grzadkowski et al., *Dimension-Six Terms in the Standard Model Lagrangian*, JHEP **10**, 2010 085, arXiv: 1008.4884 [hep-ph].

[35] W. Buchmueller and D. Wyler, *Effective lagrangian analysis of new interactions and flavour conservation*, Nuclear Physics B **268**, 1986 621, http://www.sciencedirect.com/science/article/pii/0550321386902622.

[36] S. Weinberg, *Baryon- and Lepton-Nonconserving Processes*, Phys. Rev. Lett. **43**, 21 1979 1566, https://link.aps.org/doi/10.1103/PhysRevLett.43.1566.

[37] F. Wilczek and Zee, *Operator Analysis of Nucleon Decay*, Phys. Rev. Lett. **43**, 1979 1571.

[38] L. F. Abbott and B. M. Wise, *Effective Hamiltonian for nucleon decay*, Phys. Rev. D **22**, 9 1980 2208, https://link.aps.org/doi/10.1103/PhysRevD.22.2208.

[39] L. Lehman, *Extending the Standard Model Effective Field Theory with the Complete Set of Dimension-7 Operators*, Phys. Rev. **D90**, 2014 125023, arXiv: 1410.4193 [hep-ph].

[40] L. Lehman and A. Martin, *Low-derivative operators of the Standard Model effective field theory via Hilbert series methods*, Journal of High Energy Physics **2016**, 2016 81, https://doi.org/10.1007/JHEP02(2016)081.

[41] B. Henning et al., *2, 84, 30, 993, 560, 15456, 11962, 261485, . . .: higher dimension operators in the SM EFT*, Journal of High Energy Physics **2017**, 2017 16, https://doi.org/10.1007/JHEP08(2017)016.

[42] F. Deppisch, J. Harz and W. Huang, *Falsifying High-Scale Baryogengesis with Neutrinoless Double Beta Decay and Lepton Flavour Violation*, Phys. Rev. **D92**, 2015 036005, arXiv: 1503.04825 [hep-ph].

[43] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. D **8**, 4 1973 1226, https://link.aps.org/doi/10.1103/PhysRevD.8.1226.

[44] ATLAS and CMS Collaborations, G. Aad et al.,
Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV, JHEP **08**, 2016 045, arXiv: 1606.02266 [hep-ex].

[45] Particle Data Group, T. et al., *Review of Particle Physics*,
Phys. Rev. D **98**, 3 2018 030001,
https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[46] ATLAS and CMS Collaborations, G. Aad et al.,
Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments,
Phys. Rev. Lett. **114**, 2015 191803, arXiv: 1503.07589 [hep-ex].

[47] CMS Collaboration, A. M. Sirunyan et al.,
Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV, JHEP **11**, 2017 047,
arXiv: 1706.09936 [hep-ex].

[48] CMS Collaboration, A. M. Sirunyan et al.,
Measurements of Higgs boson production via gluon fusion and vector boson fusion in the diphoton decay channel at $\sqrt{s} = 13$ TeV, 2019,
https://cds.cern.ch/record/2667225.

[49] ATLAS Collaboration, G. Aad et al.,
Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector,
Physics Letters B **784**, 2018 345?366, ISSN: 0370-2693,
http://dx.doi.org/10.1016/j.physletb.2018.07.050.

[50] ATLAS Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **90**, 2014 052004, arXiv: 1406.3827 [hep-ex].

[51] CMS Collaboration, V. Khachatryan et al., *Limits on the Higgs boson lifetime and width from its decay to four charged leptons*, Phys. Rev. D **92**, 2015 072010, arXiv: 1507.06656 [hep-ex].

[52] L. Dixon and M. S. Siu,
Resonance-Continuum Interference in the Diphoton Higgs Signal at the LHC,
Phys. Rev. Lett. **90**, 25 2003 252001,
https://link.aps.org/doi/10.1103/PhysRevLett.90.252001.

[53] S. P. Martin,
Shift in the LHC Higgs diphoton mass peak from interference with background,
Phys. Rev. D **86**, 7 2012 073016,
https://link.aps.org/doi/10.1103/PhysRevD.86.073016.

[54] L. J. Dixon and Y. Li, *Bounding the Higgs Boson Width through Interferometry*,
Phys. Rev. Lett. **111**, 11 2013 111802,
https://link.aps.org/doi/10.1103/PhysRevLett.111.111802.

[55] ATLAS Collaboration, G. Aad et al.,
Estimate of the m_H shift due to interference between signal and background processes in the $H \to \gamma\gamma$ channel, for the $\sqrt{s} = 8$ TeV dataset recorded by ATLAS,
2016, https://cds.cern.ch/record/2146386.

[56] J. Campbell et al.,
Interference in the $gg \to h \to \gamma\gamma$ On-Shell Rate and the Higgs Boson Total Width,
Phys. Rev. Lett. **119**, 2017 181801, [Addendum: Phys. Rev. Lett.119,no.19,199901(2017)], arXiv: 1704.08259 [hep-ph].

[57] ATLAS Collaboration, G. Aad et al.,
Projections for measurements of Higgs boson signal strengths and coupling parameters with the ATLAS detector at a HL-LHC, 2014,
http://cds.cern.ch/record/1956710.

[58] N. Kauer and G. Passarino,
Inadequacy of zero-width approximation for a light Higgs boson signal,
JHEP **08**, 2012 116, arXiv: 1206.4803 [hep-ph].

[59] F. Caola and K. Melnikov,
Constraining the Higgs boson width with ZZ production at the LHC,
Phys. Rev. **D88**, 2013 054024, arXiv: 1307.4935 [hep-ph].

[60] J. M. Campbell, R. K. Ellis and C. Williams, *Bounding the Higgs width at the LHC using full analytic results for gg →e + e − μ + μ−*,
Journal of High Energy Physics **2014**, 2014 60,
https://doi.org/10.1007/JHEP04(2014)060.

[61] J. M. Campbell, R. K. Ellis and C. Williams,
Bounding the Higgs width at the LHC: Complementary results from $H \to WW$,
Phys. Rev. D **89**, 5 2014 053011,
https://link.aps.org/doi/10.1103/PhysRevD.89.053011.

[62] CMS Collaboration, A. M. Sirunyan et al.,
Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state, 2019,
arXiv: 1901.00174 [hep-ex].

[63] ATLAS and CMS Collaborations, M. Aaboud et al.,
Constraints on off-shell Higgs boson production and the Higgs boson total width in $ZZ \to 4\ell$ and $ZZ \to 2\ell 2\nu$ final states with the ATLAS detector,
Physics Letters B **786**, 2018 223?244, ISSN: 0370-2693,
http://dx.doi.org/10.1016/j.physletb.2018.09.048.

[64] ATLAS and CMS Collaborations, M. Cepeda et al.,
Higgs Physics at the HL-LHC and HE-LHC, 2019, arXiv: 1902.00134 [hep-ph].

[65] ATLAS collaboration, G. Aad et al., *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} =$ 13 TeV collected with the ATLAS experiment*, 2019,
https://cds.cern.ch/record/2668375.

[66] CMS Collaboration, A. M. S. et al., *Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **79**, 2018 421. 67 p, Submitted to Eur.Phys.J.,
https://cds.cern.ch/record/2640611.

[67] ATLAS Collaboration, G. Aad et al., *Study of the double Higgs production channel $H(\to b\bar{b})H(\to \gamma\gamma)$ with the ATLAS experiment at the HL-LHC*, 2017,
https://cds.cern.ch/record/2243387.

[68] CMS Collaboration, A. M. S. et al.,
Higgs pair production at the High Luminosity LHC, 2015,
https://cds.cern.ch/record/2063038.

[69] ATLAS Collaboration, G. Aad et al.,
Prospects for measuring Higgs pair production in the channel $H(\to \gamma\gamma)H(\to b\bar{b})$ using the ATLAS detector at the HL-LHC, 2014,
http://cds.cern.ch/record/1956733.

[70] L. D. Landau, *On the angular momentum of a system of two photons*,
Dokl. Akad. Nauk Ser. Fiz. **60**, 1948 207.

[71] C. N. Yang,
Selection Rules for the Dematerialization of a Particle into Two Photons,
Phys. Rev. **77**, 2 1950 242,
https://link.aps.org/doi/10.1103/PhysRev.77.242.

[72] L. Evans and P. Bryant, *LHC Machine*,
Journal of Instrumentation **3**, 2008 S08001,
https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[73] O. Bruening et al., *LHC Design Report*,
CERN Yellow Reports: Monographs, 2004,
https://cds.cern.ch/record/782076.

[74] E. Mobs, *The CERN accelerator complex*, 2016, General Photo, https://cds.cern.ch/record/2197559.

[75] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**, 2008 S08003.

[76] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716**, 2012 30, arXiv: 1207.7235 [hep-ex].

[77] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3**, 2008 S08002.

[78] LHCb Collaboration, A. A. Alves et al., *The LHCb Detector at the LHC*, Journal of Instrumentation **3**, 2008 S08005, https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08005.

[79] ATLAS Collaboration, M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76**, 2016 653, arXiv: 1608.03953 [hep-ex].

[80] *Sherpa and Open Science Grid: Predicting the emergence of jets*, https://sciencenode.org/feature/sherpa-and-open-science-grid-predicting-emergence-jets.php (visited on 07/08/2016).

[81] C. J. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Ann. Rev. Nucl. Part. Sci. **37**, 1987 383.

[82] V. Gribov and L. Lipatov, *Deep inelastic electron scattering in perturbation theory*, Phys. Lett. B **37**, 1971 78.

[83] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126**, 1977 298.

[84] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. **D93**, 2016 033006, arXiv: 1506.07443 [hep-ph].

[85] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75**, 2015 204, arXiv: 1412.3989 [hep-ph].

[86] R. D. Ball et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, Nucl. Phys. **B849**, 2011 296, arXiv: 1101.1300 [hep-ph].

[87] The NNPDF collaboration et al., *Parton distributions for the LHC run II*, Journal of High Energy Physics **2015**, 2015 40, https://doi.org/10.1007/JHEP04(2015)040.

[88] J. B. et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. **G43**, 2016 023001, arXiv: 1510.03865 [hep-ph].

[89] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70**, 2010 823, arXiv: 1005.4568 [physics.ins-det].

[90] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506**, 2003 250.

[91] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72**, 2012 1849, arXiv: 1110.1530 [hep-ex].

[92] ATLAS Collaboration, G. Aad et al., *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016, https://cds.cern.ch/record/2136007.

[93] ATLAS Collaboration, G. Aad et al., *ATLAS inner detector: Technical Design Report, 1*, Technical Design Report ATLAS, 1997, https://cds.cern.ch/record/331063.

[94] ATLAS Collaboration, S. Haywood et al., *ATLAS inner detector: Technical Design Report, 2*, Technical Design Report ATLAS, 1997, https://cds.cern.ch/record/331064.

[95] ATLAS Collaboration, M. . Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, 2010, https://cds.cern.ch/record/1291633.

[96] ATLAS Collaboration, G. Aad et al., *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical Design Report ATLAS, 1996, https://cds.cern.ch/record/331061.

[97] ATLAS Collaboration, G. Aad et al., *ATLAS tile calorimeter: Technical Design Report*, Technical Design Report ATLAS, 1996, https://cds.cern.ch/record/331062.

[98] ATLAS Collaboration, G. Aad et al., *ATLAS muon spectrometer: Technical Design Report*, Technical Design Report ATLAS, 1997, https://cds.cern.ch/record/331068.

[99] ATLAS Collaboration, H. Kroha, R. Fakhrutdinov and A. Kozhin, *New high-precision drift-tube detectors for the ATLAS muon spectrometer*, Journal of Instrumentation **12**, 2017 C06007, https://doi.org/10.1088%2F1748-0221%2F12%2F06%2Fc06007.

[100] J. Pequenao and P. Schaffner, *An computer generated image representing how ATLAS detects particles*, 2013, https://cds.cern.ch/record/1505342.

[101] C. Grupen and B. Schwartz, *Particle detectors*, Cambridge University Press, 2008.

[102] ATLAS Collaboration, M. Aaboud et al., *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024, 2016, http://cds.cern.ch/record/2157687.

[103] ATLAS Collaboration, M. Aaboud et al., *Electron identification measurements in ATLAS using $\sqrt{s}$ = 13 TeV data with 50 ns bunch spacing*, 2015, https://cds.cern.ch/record/2048202.

[104] ATLAS Collaboration, M. Aaboud et al., *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV*, Submitted to: Eur. Phys. J., 2019, arXiv: 1902.04655 [physics.ins-det].

[105] ATLAS Collaboration, W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, https://cds.cern.ch/record/1099735.

[106] ATLAS Collaboration, G. Aad et al., *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, 2017, https://cds.cern.ch/record/2298955.

[107] ATLAS Collaboration, G. Aad et al., *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, https://cds.cern.ch/record/1449796.

[108] *Electron efficiency measurements in 2017 data and electron identification discriminating variables from 2016 data*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2018-002/index.html (visited on 06/24/2019).

[109] ATLAS collaboration, M. Aaboud et al., *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, Journal of Instrumentation **14**, 2019 P03017, https://doi.org/10.1088%2F1748-0221%2F14%2F03%2Fp03017.

[110] ATLAS Collaboration, G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C **74**, 2014 3130, arXiv: 1407.3935 [hep-ex].

[111] ATLAS Collaboration, G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton−proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. **C76**, 2016 292, arXiv: 1603.05598 [hep-ex].

[112] *ATLAS Muon Combined Performance with 2017 and 2016 dataset*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-002/index.html (visited on 06/28/2019).

[113] M. Cacciari, G. P. Salam and G. Soyez, *The anti-ktjet clustering algorithm*, Journal of High Energy Physics **2008**, 2008 063, https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063.

[114] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, The European Physical Journal C **72**, 2012 1896, https://doi.org/10.1140/epjc/s10052-012-1896-2.

[115] ATLAS Collaboration, M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96**, 7 2017 072002, https://link.aps.org/doi/10.1103/PhysRevD.96.072002.

[116] ATLAS Collaboration, G. Aad et al., *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, https://cds.cern.ch/record/1700870.

[117] ATLAS Collaboration, G. Aad et al., *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76**, 2016 581, arXiv: 1510.03823 [hep-ex].

[118] ATLAS Collaboration, M. Aaboud et al., *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08**, 2018 089, arXiv: 1805.01845 [hep-ex].

[119] ATLAS Collaboration, G. Aad et al., *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, https://cds.cern.ch/record/2037697.

[120] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010**, 2010 43, https://doi.org/10.1007/JHEP06(2010)043.

[121] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, Journal of High Energy Physics **2007**, 2007 070, https://doi.org/10.1088%2F1126-6708%2F2007%2F11%2F070.

[122] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, Journal of High Energy Physics **2004**, 2004 040, https://doi.org/10.1088%2F1126-6708%2F2004%2F11%2F040.

[123] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*,
Journal of High Energy Physics **2010**, 2010 37,
https://doi.org/10.1007/JHEP02(2010)037.

[124] G. Luisoni et al., *HW±/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*,
Journal of High Energy Physics **2013**, 2013 83,
https://doi.org/10.1007/JHEP10(2013)083.

[125] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462**, 2001 152, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines,
http://www.sciencedirect.com/science/article/pii/S0168900201000894.

[126] S. Catani and M. Grazzini,
Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-Boson Production at the Large Hadron Collider,
Phys. Rev. Lett. **98**, 22 2007 222002,
https://link.aps.org/doi/10.1103/PhysRevLett.98.222002.

[127] M. Grazzini, *NNLO predictions for the Higgs boson signal in the $H\to WW\to l\nu l\nu$ and $H\to ZZ\to 4l$decay channels*, Journal of High Energy Physics **2008**, 2008 043,
https://doi.org/10.1088%2F1126-6708%2F2008%2F02%2F043.

[128] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07**, 2014 079, arXiv: 1405.0301 [hep-ph].

[129] H.-L. Lai et al., *New parton distributions for collider physics*,
Phys. Rev. D **82**, 7 2010 074024,
https://link.aps.org/doi/10.1103/PhysRevD.82.074024.

[130] The NNPDF collaboration et al., *Parton distributions for the LHC run II*,
Journal of High Energy Physics **2015**, 2015 40,
https://doi.org/10.1007/JHEP04(2015)040.

[131] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*,
Computer Physics Communications **178**, 2008 852,
http://www.sciencedirect.com/science/article/pii/S0010465508000441.

[132] ATLAS Collaboration, G. Aad et al., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09**, 2014 145, arXiv: 1406.3660 [hep-ex].

[133] ATLAS Collaboration, G. Aad et al., *ATLAS Run 1 Pythia8 tunes*,
ATL-PHYS-PUB-2014-021, 2014, https://cds.cern.ch/record/1966419.

[134] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58**, 2008 639, arXiv: 0803.0883 [hep-ph].

[135] A. Bredenstein et al., *Precise predictions for the Higgs-boson decay H —> WW/ZZ —> 4 leptons*, Phys. Rev. **D74**, 2006 013004, arXiv: hep-ph/0604011 [hep-ph].

[136] A. Bredenstein et al., *Radiative corrections to the semileptonic and hadronic Higgs-boson decays H —> W W / Z Z —> 4 fermions*, JHEP **02**, 2007 080, arXiv: hep-ph/0611234 [hep-ph].

[137] A. Djouadi, M. Spira and P. M. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. B **264**, 1991 440.

[138] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. B **359**, 1991 283.

[139] M. Spira et al., *Higgs boson production at the LHC*, Nucl. Phys. B **453**, 1995 17, arXiv: hep-ph/9504378.

[140] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88**, 2002 201801, arXiv: hep-ph/0201206.

[141] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. B **646**, 2002 220, arXiv: hep-ph/0207004.

[142] V. Ravindran, J. Smith and W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. B **665**, 2003 325, arXiv: hep-ph/0302135.

[143] C. Anastasiou et al., *Higgs boson gluon-fusion production at threshold in N^3LO QCD*, Phys. Lett. B **737**, 2014 325, arXiv: 1403.4616 [hep-ph].

[144] C. Anastasiou et al., *Higgs boson gluon-fusion production beyond threshold in N^3LO QCD*, JHEP **03**, 2015 091, arXiv: 1411.3584 [hep-ph].

[145] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05**, 2016 058, arXiv: 1602.00695 [hep-ph].

[146] U. Aglietti et al., *Two loop light fermion contribution to Higgs production and decays*, Phys. Lett. B **595**, 2004 432, arXiv: hep-ph/0404071.

[147] S. Actis et al.,
NLO electroweak corrections to Higgs boson production at hadron colliders,
Phys. Lett. B **670**, 2008 12, arXiv: 0809.1301 [hep-ph].

[148] D. de Florian and M. Grazzini,
Higgs production at the LHC: updated cross sections at $\sqrt{s}$=8 TeV,
Phys. Lett. B **718**, 2012 117, arXiv: 1206.4133 [hep-ph].

[149] C. Anastasiou et al., *Inclusive Higgs boson cross-section for the LHC at 8 TeV*,
JHEP **04**, 2012 004, arXiv: 1202.3638 [hep-ph].

[150] J. Baglio and A. Djouadi, *Higgs production at the LHC*, JHEP **03**, 2011 055,
arXiv: 1012.0530 [hep-ph].

[151] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC*,
Phys. Rev. Lett. **99**, 2007 161803, arXiv: 0707.0381 [hep-ph].

[152] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*,
Phys. Rev. D **77**, 2008 013002, arXiv: 0710.4749 [hep-ph].

[153] K. Arnold et al.,
VBFNLO: A parton level Monte Carlo for processes with electroweak bosons,
Comput. Phys. Commun. **180**, 2009 1661, arXiv: 0811.4559 [hep-ph].

[154] P. Bolzoni et al., *Higgs production via vector-boson fusion at NNLO in QCD*,
Phys. Rev. Lett. **105**, 2010 011801, arXiv: 1003.4451 [hep-ph].

[155] T. Han and S. Willenbrock,
QCD correction to the $pp \to WH$ and ZH total cross-sections,
Phys. Lett. B **273**, 1991 167.

[156] O. Brein, A. Djouadi and R. Harlander,
NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders,
Phys. Lett. B **579**, 2004 149, arXiv: hep-ph/0307206.

[157] M.L. Ciccolini, S. Dittmaier and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*,
Phys. Rev. D **68**, 2003 073003, arXiv: hep-ph/0306234.

[158] W. Beenakker et al., *Higgs radiation off top quarks at the Tevatron and the LHC*,
Phys. Rev. Lett. **87**, 2001 201805, arXiv: hep-ph/0107081.

[159] W. Beenakker et al.,
NLO QCD corrections to $t\bar{t}H$ production in hadron collisions,
Nucl. Phys. B **653**, 2003 151, arXiv: hep-ph/0211352.

[160] S. Dawson et al., *Next-to-leading order QCD corrections to $pp \to t\bar{t}h$ at the CERN Large Hadron Collider*, Phys. Rev. D **67**, 2003 071503,
arXiv: hep-ph/0211438.

[161] S. Dawson et al., *Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections*, Phys. Rev. D **68**, 2003 034022, arXiv: hep-ph/0305087.

[162] R. Harlander, M. Krämer and M. Schumacher, *Bottom-quark associated Higgs-boson production: Reconciling the four- and five-flavour scheme approach*, 2011, arXiv: 1112.3478 [hep-ph].

[163] P. de Aquino and K. Mawatari, *Characterising a Higgs-like resonance at the LHC*, 2013, arXiv: 1307.5607 [hep-ph], http://www.slac.stanford.edu/econf/C130213.1/pdfs/mawatari.pdf.

[164] *The Higgs Characterisation model*, http://feynrules.irmp.ucl.ac.be/wiki/HiggsCharacterisation (visited on 11/27/2019).

[165] A. Alloul et al., *FeynRules 2.0 - A complete toolbox for three-level phenomenology*, Computer Physics Communications **185**, 2014 2250, http://www.sciencedirect.com/science/article/pii/S0010465514001350.

[166] C. Degrande et al., *UFO - The Universal FeynRules Output*, Computer Physics Communications **183**, 2012 1201, http://www.sciencedirect.com/science/article/pii/S0010465512000379.

[167] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **2014**, 2014 79, https://doi.org/10.1007/JHEP07(2014)079.

[168] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, JHEP **05**, 2002 046, arXiv: hep-ph/0112284 [hep-ph].

[169] *Standard Model Effective Field Theory – The SMEFTsim package*, https://feynrules.irmp.ucl.ac.be/wiki/SMEFT (visited on 11/27/2019).

[170] T. Gleisberg et al., *Event generation with SHERPA 1.1*, Journal of High Energy Physics **2009**, 2009 007, https://doi.org/10.1088%2F1126-6708%2F2009%2F02%2F007.

[171] T. Gleisberg and S. Haeche, *Comix, a new matrix element generator*, Journal of High Energy Physics **2008**, 2008 039, https://doi.org/10.1088%2F1126-6708%2F2008%2F12%2F039.

[172] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108**, 11 2012 111601, https://link.aps.org/doi/10.1103/PhysRevLett.108.111601.

[173] S. Schumann and F. Krauss,
A parton shower algorithm based on Catani-Seymour dipole factorisation,
Journal of High Energy Physics **2008**, 2008 038,
https://doi.org/10.1088%2F1126-6708%2F2008%2F03%2F038.

[174] S. Höche et al., *QCD matrix elements + parton showers. The NLO case*,
Journal of High Energy Physics **2013**, 2013 27,
https://doi.org/10.1007/JHEP04(2013)027.

[175] B. Biedermann et al., *Electroweak Corrections to $pp \to \mu^+\mu^-e^+e^- + X$ at the LHC: A Higgs Boson Background Study*, Phys. Rev. Lett. **116**, 16 2016 161803,
https://link.aps.org/doi/10.1103/PhysRevLett.116.161803.

[176] B. Biedermann et al., *Next-to-leading-order electroweak corrections to the production of four charged leptons at the LHC*,
Journal of High Energy Physics **2017**, 2017 33,
https://doi.org/10.1007/JHEP01(2017)033.

[177] N. Kauer, C. O'Brien and E. Vryonidou, *Interference effects for $H \to W\,W \to \ell\nu q\bar{q}'$ and $H \to ZZ \to \ell\bar{\ell}q\bar{q}$ searches in gluon fusion at the LHC*,
JHEP **10**, 2015 074, arXiv: 1506.01694 [hep-ph].

[178] F. Caola et al., *QCD corrections to ZZ production in gluon fusion at the LHC*,
Phys. Rev. D **92**, 9 2015 094028,
https://link.aps.org/doi/10.1103/PhysRevD.92.094028.

[179] F. Caola et al., *QCD corrections to W^+W^- production through gluon fusion*,
Physics Letters B **754**, 2016 275,
http://www.sciencedirect.com/science/article/pii/S0370269316000630.

[180] K. Melnikov and M. Dowling,
Production of two Z-bosons in gluon fusion in the heavy top quark approximation,
Physics Letters B **744**, 2015 43,
http://www.sciencedirect.com/science/article/pii/S0370269315001938.

[181] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*,
Journal of High Energy Physics **2014**, 2014 79.

[182] ATLAS Collaboration, M. Aaboud et al.,
Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector, JHEP **03**, 2018 095,
arXiv: 1712.02304 [hep-ex].

[183] ATLAS Collaboration, M. Aaboud et al.,
Measurements of the Higgs boson production, fiducial and differential cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector, 2018,
https://cds.cern.ch/record/2621479.

[184] ATLAS Collaboration, M. Aaboud et al.,
Measurements of the Higgs boson inclusive, differential and production cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector, 2019,
http://cds.cern.ch/record/2682107.

[185] K. P. Murphy, *Machine learning : a probabilistic perspective*,
Cambridge, Mass. [u.a.]: MIT Press, 2013.

[186] A. H. et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007,
arXiv: physics/0703039 [physics.data-an].

[187] *ROOT - Data Analysis Framework*,
https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html (visited on 12/02/2019).

[188] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*,
Eur. Phys. J. **C75**, 2015 132, arXiv: 1412.7420 [hep-ph].

[189] LHC Higgs Cross Section Working Group, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 2016,
arXiv: 1610.07922 [hep-ph].

[190] J. M. Campbell, R. K. Ellis and C. Williams,
Vector boson pair production at the LHC, JHEP **07**, 2011 018,
arXiv: 1105.0020 [hep-ph].

[191] S. Mrenna and P. Skands, *Automated parton-shower variations in pythia 8*,
Phys. Rev. D **94**, 7 2016 074005,
https://link.aps.org/doi/10.1103/PhysRevD.94.074005.

[192] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*,
The European Physical Journal C **76**, 2016 196,
https://doi.org/10.1140/epjc/s10052-016-4018-8.

[193] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*,
The European Physical Journal C **71**, 2011 1554,
https://doi.org/10.1140/epjc/s10052-011-1554-0.

[194] ATLAS Collaboration, M. Aaboud et al.,
Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC, 2018,
http://cds.cern.ch/record/2652762.

[195] B. H. et al., *Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020*, 2019, 254 p, http://cds.cern.ch/record/2691414.

[196] ATLAS Collaboration, M. Aaboud et al.,
Expected performance of the ATLAS detector at the High-Luminosity LHC, 2019,
https://cds.cern.ch/record/2655304.

[197] R. A. Khalek et al.,
Towards ultimate parton distributions at the high-luminosity LHC,
The European Physical Journal C **78**, 2018 962,
https://doi.org/10.1140/epjc/s10052-018-6448-y.

[198] ATLAS Collaboration, K. M. Ecker,
Measurement of the Higgs boson tensor coupling in $H \to ZZ^* \to 4\ell$ decays with the ATLAS detector - How odd is the Higgs boson?, Presented 22 May 2018,
2018, https://cds.cern.ch/record/2624277.

[199] ATLAS Collaboration, G. Aad et al., *A morphing technique for signal modelling in a multidimensional space of coupling parameters*, ATL-PHYS-PUB-2015-047,
2015, https://cds.cern.ch/record/2066980.

[200] J. Ellis, C. W. Murphy, V. S., T. You,
Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data,
10.1007/JHEP06(2018)146, 2018.

[201] ATLAS Collaboration, M. Aaboud et al.,
Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector, 2020,
arXiv: 2004.03447 [hep-ex].

[202] D. Guest, K. Cranmer and D. Whiteson,
Deep Learning and Its Application to LHC Physics,
Annual Review of Nuclear and Particle Science **68**, 2018 161,
eprint: https://doi.org/10.1146/annurev-nucl-101917-021019,
https://doi.org/10.1146/annurev-nucl-101917-021019.

[203] M. Abadi et al.,
TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems,
2015, Software available from tensorflow.org, https://www.tensorflow.org/.

[204] C. Degrande et al., *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*,
The European Physical Journal C **77**, 2017 262,
https://doi.org/10.1140/epjc/s10052-017-4793-x.

[205] ATLAS Collaboration, M. Aaboud et al.,
Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector,
Phys. Rev. **D98**, 2018 052005, arXiv: 1802.04146 [hep-ex].

[206] ATLAS Collaboration, M. Aaboud et al., *Measurements and interpretations of Higgs-boson fiducial cross sections in the diphoton decay channel using 139 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, 2019,
http://cds.cern.ch/record/2682800.

[207] ATLAS Collaboration, G. Aad et al., *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. C **76**, 2016 658, arXiv: 1602.04516 [hep-ex].

[208] ATLAS Collaboration, M. Aaboud et al., *Test of CP invariance in vector-boson fusion production of the Higgs boson in the $H \to \tau\tau$ channel in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, 2019, https://cds.cern.ch/record/2693960.

[209] CMS Collaboration, A. M. Sirunyan et al., *Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state*, Phys. Lett. B **775**, 2017 1, arXiv: 1707.00541 [hep-ex].

[210] CMS Collaboration, A. M. Sirunyan et al., *Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state*, Phys. Rev. D **99**, 2019 112003. 33 p, https://cds.cern.ch/record/2652808.

[211] CMS Collaboration, A. M. Sirunyan et al., *Constraints on anomalous HVV couplings from the production of Higgs bosons decaying to τ lepton pairs*, Phys. Rev. D **100**, 2019 112002. 33 p, Submitted to Phys.Rev., https://cds.cern.ch/record/2667207.

[212] P. Achard et al., *Search for anomalous couplings in the Higgs sector at LEP*, Physics Letters B **589**, 2004 89, http://www.sciencedirect.com/science/article/pii/S0370269304005040.

[213] W. Buchmueller and D. Wyler, *Effective lagrangian analysis of new interactions and flavour conservation*, Nuclear Physics B **268**, 1986 621, http://www.sciencedirect.com/science/article/pii/0550321386902622.

[214] V. Hankele et al., *Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC*, Phys. Rev. D **74**, 9 2006 095001, https://link.aps.org/doi/10.1103/PhysRevD.74.095001.

[215] I. Anderson et al., *Constraining anomalous HVV interactions at proton and lepton colliders*, Phys. Rev. D **89**, 3 2014 035007, https://link.aps.org/doi/10.1103/PhysRevD.89.035007.

[216] T. Plehn, D. L. Rainwater and D. Zeppenfeld, *Determining the Structure of Higgs Couplings at the LHC*, Phys. Rev. Lett. **88**, 2001 051801. 7 p, http://cds.cern.ch/record/502043.

[217] V. Hankele et al.,
Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC,
Phys. Rev. **D74**, 2006 095001, arXiv: hep-ph/0609075 [hep-ph].

[218] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form factors of the top quark via* $e^+e^- \to t\bar{t}$,
Phys. Rev. D **45**, 7 1992 2405,
https://link.aps.org/doi/10.1103/PhysRevD.45.2405.

[219] M. Diehl, O. Nachtmann and F. Nagel, *Triple gauge couplings in polarised* $e^-e^+ \to W^-W^+$ *and their measurement using optimal observables*,
The European Physical Journal C - Particles and Fields **27**, 2003 375,
https://doi.org/10.1140/epjc/s2002-01096-y.

[220] M. Diehl and O. Nachtmann, *Optimal observables for the measurement of three gauge boson couplings in* $e^-e^+ \to W^-W^+$,
The European Physical Journal C - Particles and Fields **62**, 1994 397,
https://doi.org/10.1007/BF01555899.

[221] M. Diehl and O. Nachtmann, *Anomalous three gauge boson couplings in* $e^-e^+ \to W^-W^+$ *and optimal strategies for their measurement*,
The European Physical Journal C **1**, 1998 177,
https://doi.org/10.1007/BF01245807.

[222] M. Diehl, O. Nachtmann and F. Nagel, *Triple gauge couplings in polarised* $e^-e^+ \to W^-W^+$ *and their measurement using optimal observables*,
The European Physical Journal C - Particles and Fields **27**, 2003 375,
https://doi.org/10.1140/epjc/s2002-01096-y.

[223] D. Guest, K. Cranmer and D. Whiteson,
Deep Learning and its Application to LHC Physics,
Ann. Rev. Nucl. Part. Sci. **68**, 2018 161, arXiv: 1806.11484 [hep-ex].