Higgs Couplings Measurement in the $H \to ZZ^* \to 4l$ Channel

Cheng Chi Lung

Department of Physics, The Chinese University of Hong Kong

Abstract

In this report, measurement of the spin-0 Higgs boson couplings with the Standard Model (SM) vector bosons except the photon in the $H \to ZZ^* \to l^+l^-l'^+l'^-$ decay channel, where $l, l' = e$ or μ, is presented. The analysis is performed using proton-proton collision data with an integrated luminosity of 36.1fb^{-1} at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ collected by the ATLAS detector at the LHC. Possible presence of beyond Standard Model (BSM) Higgs boson state $\mathcal{X}(J^P)$ with spin/parity assignments $J^P = 0^+, 0^-$ with possible $\mathcal{CP}$ mixing for the spin-0 scenario is investigated. The measurement employs the framework of Higgs Characterization based on an Effective Field Theory approach for the description of the Higgs interactions. To provide sensitive probes of the $\mathcal{CP}$ nature of the Higgs couplings for different production mechanisms, measurements are performed in separate categories of events with distinct experimental signatures. Statistical evaluation of the Higgs couplings with gluons and weak vector bosons is performed in the ggH and HVV production vertex with corresponding discriminant observables. The result shows no significant excess of BSM coupling contributions over the SM predictions.

I. Introduction

The discovery of a Higgs-like boson at a mass of 125 GeV at the Large Hadron Collider (LHC) has initiated a series of precision measurements of its properties for confirmation of its identity as the SM Higgs boson. It also opens up the search for possible hints to new physics beyond the SM.

Some of the key properties to be determined from the new boson are its spin and parity quantum numbers and its interactions with SM particles. The Higgs boson predicted by the SM is a $\mathcal{CP}$ even scalar particle with spin/parity $J^P = 0^+$. Alternative hypotheses may include $J^P = 0^-, 1^+, 1^-, 2^+, 2^-$. The observation of the two photon final states has strongly disfavored the spin-1 hypothesis as the decay is forbidden by the Landau-Yang theorem [19][20]. Moreover, measurements by the ATLAS and CMS collaborations [3][7] have strongly excluded these alternative spin/parity hypotheses in favor of the $\mathcal{CP}$ even scalar boson predicted by the SM. Nevertheless, many BSM theories such as the Two Higgs Doublet Model (2HDM) include an extended Higgs sector and several neutral Higgs boson states can exist. Instead of a pure $\mathcal{CP}$ even Higgs boson, there can be mixing of $\mathcal{CP}$ odd and $\mathcal{CP}$ even terms in the Higgs boson interactions, which could produce observable changes in the Higgs boson production and in the kinematics of the final-state particles.

Measurements on the Higgs boson couplings with possible $\mathcal{CP}$ mixing can therefore be performed by looking at the final state kinematics in specific Higgs production mechanisms to probe specific interactions with SM particles. In this report, a spin-0 scenario is studied and the theoretical framework of Higgs Characterization based on an Effective Field Theory (EFT) approach is used for describing the Higgs interactions which also allows a systematic measurement of its properties.

In the SM, the main production mechanisms of the Higgs boson is gluon gluon fusion (ggF), followed by vector boson fusion (VBF), vector boson associated production (VH) and top/bottom associated production ($t\bar{t}H/b\bar{t}H$). Their diagrams and production cross sections are shown in Fig. 1 and Fig. 2. The search for possible presence of BSM coupling terms in the ggH and HVV vertex is therefore performed.

This analysis is based on the $H \to ZZ^* \to 4l$ decay channel, diagram shown in Fig. 1, using pp collision data with an integrated luminosity of $36.1\,\text{fb}^{-1}$ at a center of mass energy $\sqrt{s} = 13\,\text{TeV}$ collected by the ATLAS experiment. The pure leptonic final states provide a

clean signature for full and accurate reconstruction of its kinematics. The decay of the Higgs resonance into neutral gauge bosons also allows the extraction of a rich set of information from a fully-reconstructed four-body final state [10] and has a reasonably high branching ratio as seen in Fig. 2.

The major background in the $H \to ZZ^* \to 4l$ channel is the irreducible $qq \to ZZ \to 4l$ continuum background which has similar kinematic distributions as the ggF process. Interference between the Higgs boson signal and the SM background processes is not considered in this analysis due to its negligible impact on the mass spectrum near the resonance in the $H \to ZZ^* \to 4l$ channel [17].

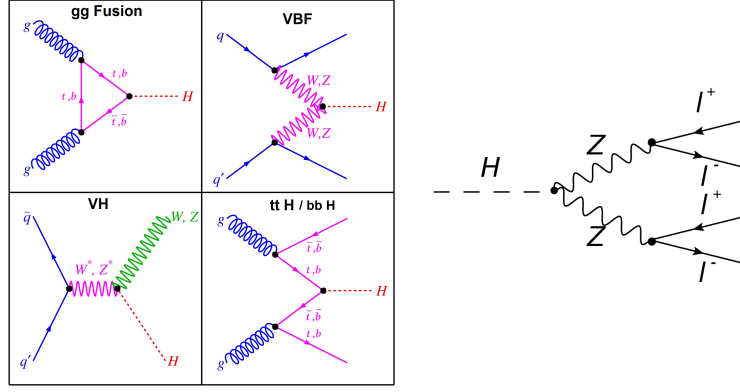Figure 1: Diagrams of the main Higgs boson production modes (left) and the $H \to ZZ^* \to l^+l^-l'^+l'^-$ decay process (right).
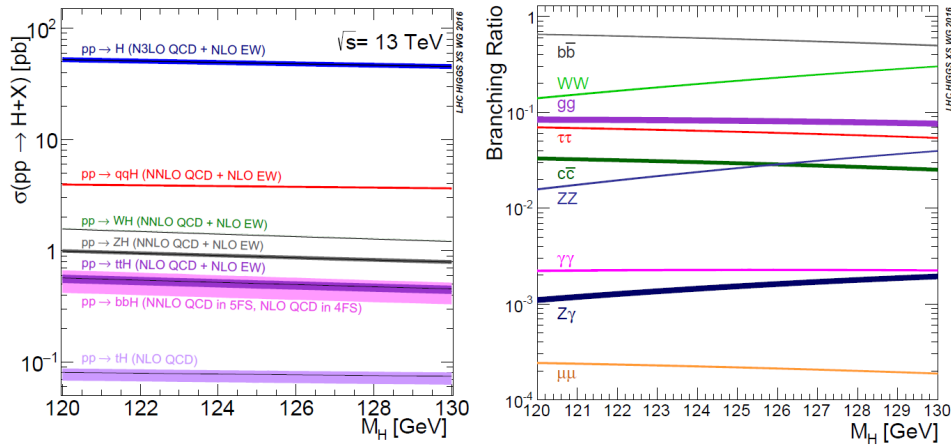

Figure 2: The SM Higgs boson production cross section (left) at $\sqrt{s} = 13\,\text{TeV}$ and the branching ratios (right) for the mass range around 125 GeV.

II. The ATLAS Detector

The ATLAS detector [2] is a multipurpose particle detector that consists of a tracking system (Inner Detector, or ID), sampling electromagnetic and hadronic calorimeters and a muon spectrometer. A sketch of the detector is shown in Fig. 3. A right-handed coordinate system is used with its origin at the nominal interaction point and the z-axis along the beam pipe. The detector is designed to have a large acceptance in pseudo-rapidity $\eta = -\ln\tan\theta/2$ with almost full azimuthal angle ϕ coverage.

The Inner Detector is contained in a central solenoid with a 2T axial magnetic field along the beam axis to bend the particles in the x-y plane. The primary purpose of the detector is to identify primary and secondary interaction vertices and measure the transverse momentum p_T of charged particles.

The ATLAS calorimeter system consists of inner electromagnetic calorimeters surrounded by hadronic calorimeters which are responsible for jet reconstruction. The electromagnetic calorimetry is provided by high granularity liquid-argon sampling calorimeters (LAr) which measure the energy and position of electromagnetic showers. The hadronic calorimetry is provided by a copper-liquid argon hadronic endcap calorimeter (HEC), a copper/tungsten liquid argon forward calorimeter (fCal) and a hadronic Tile calorimeter. Energy depositions of electromagnetic and hadronic showers in the calorimeters are recorded and several energy calibration techniques are used to relate the jet energy measured with the calorimeter to the true energy of the corresponding jet. Jet reconstruction typically induces larger systematic uncertainties such as the jet energy scale and pile up uncertainties [1].

The muon spectrometer encloses the ATLAS calorimeters and measures the muon tracks. It provides a precise tracking of high P_T muons with a good transverse momentum resolution, which allows an accurate reconstruction of final state decays into two or four muons (i.e. $H \to ZZ^* \to 2\mu 2l$ and $H \to ZZ^* \to 4\mu$).

The ATLAS trigger system uses a two-level trigger system for selecting interesting events for data analysis. A hardware-based Level 1 (L1) trigger uses a subset of detector information to reduce event rate to below $100\,\mathrm{kHz}$. This is followed by a software-based high-level trigger (HLT) which performs fast reconstruction of physical objects such as muons, electrons and jets to further reduce the event rate to about 1 kHz.
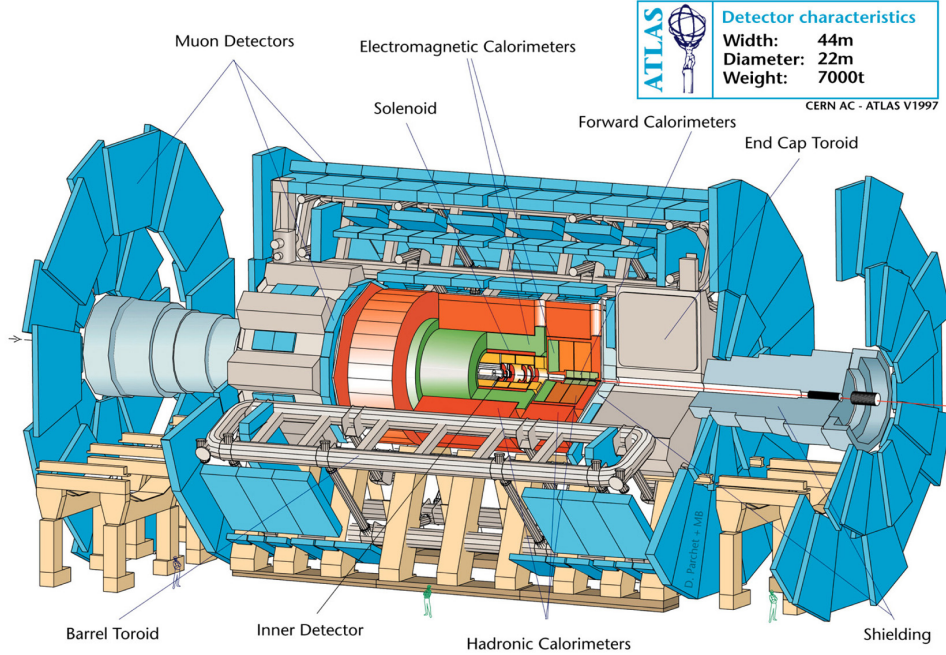
Figure 3: The major components of the ATLAS detector.

III. Effective Field Theory and Higgs Characterization Model

An effective field theory (EFT)[5] [9] is an approximation method for describing non SM interactions by introducing perturbative terms in the interaction Lagrangian as an expansion of operators in the mass dimension D with the leading order term being the SM Lagrangian. The theory does not require $\mathcal{CP}$ conservation and any new physics is assumed to be dominantly described by the subleading order operators. Such approximation is only valid up to an energy scale Λ above which the theory loses predictive power due to unitarity violation. The scale is taken to be 1 TeV in this analysis as no evidence of new physics at lower energy scales is observed by previous collider experiments.

In the Higgs Characterization model [4], one assumes that there is only one resonance with a mass of about 125 GeV and any other BSM particle exists at an energy scale larger than Λ. The width of the assumed Higgs boson is small with respect to experimental resolution and the effect of interference between the signal and SM backgrounds is neglected.

The Higgs Characterization model is implemented in the MadGraph5_aMC@NLO generator for creation of signal Monte Carlo samples. In this analysis, the observed resonance at 125 GeV is taken as a mixture of spin-0 CP-even and/or CP-odd states. In the case of CP mixing, the Higgs boson would be a mass eigenstate, but not a CP eigenstate [12].

IV. Coupling Measurements

Measurement of the Higgs couplings is performed in the ggH and HVV sector using the following effective Lagrangian:

$$\mathcal{L}_{HC,J} = \mathcal{L}_{SM-H} + \mathcal{L}_J \tag{IV.1}$$

where the first term on the R.H.S. describes the SM degrees of freedom except for the Higgs, and $\mathcal{L}_J$ contains the kinetic and interaction terms of the new bosonic state

The effective Lagrangian for the interaction of a spin 0 scalar and pseudoscalar states(H and A respectively) with vector bosons (gluon, Z and W) can be written as follows:

$$\begin{aligned}
\mathcal{L}_0^V = \Big\{ & c_\alpha \kappa_{\mathrm{SM}} [\frac{1}{2} g_{H_{ZZ}} Z_\mu Z^\mu + g_{H_{WW}} W_\mu^+ W^{-\mu}] \\
& - \frac{1}{4} [c_\alpha \kappa_{H_{gg}} g_{H_{gg}} G_{\mu\nu}^\alpha G^{\alpha,\mu\nu} + s_\alpha \kappa_{A_{gg}} g_{A_{gg}} G_{\mu\nu}^a \widetilde{G}^{a,\mu\nu}] \\
& - \frac{1}{4}\frac{1}{\Lambda} [c_\alpha \kappa_{H_{ZZ}} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{A_{ZZ}} Z_{\mu\nu} \widetilde{Z}^{\mu\nu}] \\
& - \frac{1}{2}\frac{1}{\Lambda} [c_\alpha \kappa_{H_{WW}} W_{\mu\nu}^+ W^{-\mu\nu} + s_\alpha \kappa_{A_{WW}} W_{\mu\nu}^+ \widetilde{W}^{-\mu\nu}] \Big\} \mathcal{X}_0,
\end{aligned} \tag{IV.2}$$

where $c_\alpha \equiv \cos\alpha$ and $s_\alpha \equiv \sin\alpha$ and α is the mixing angle. The terms V^μ ($V = Z, W^\pm$) represent the vector boson fields, $V_{\mu\nu}, G_{\mu\nu}^a$ are the (reduced) field strength tensors and $\widetilde{V}_{\mu\nu}$ is the dual tensor. Their definitions can be found in Appendix XIII A.

In the HVV vertex, mixing between the Standard model $\mathcal{CP}$ even (0^+), BSM $\mathcal{CP}$ odd (0^-) and BSM $\mathcal{CP}$ even (0^+) state is considered. The symbols κ_{SM}, $\kappa_{H_{VV}}$ and $\kappa_{A_{VV}}$ denote the coupling scale factor corresponding to the interaction of Standard Model 0^+, BSM $0^\pm$ particles, represented by the $\mathcal{X}_0$ field, with ZZ or WW pairs. For simplicity, it is assumed that $\kappa_{X_{WW}} = \kappa_{X_{ZZ}} = \kappa_{X_{VV}}$ ($X = A, H$). All couplings are assumed to be real to ensure the Lagrangian terms are Hermitian.

Similarly in the ggH vertex, mixing between the Standard model 0^+ and BSM 0^- state is considered. The coupling scale factors $\kappa_{H_{gg}}$ and $\kappa_{A_{gg}}$ correspond to the interaction of the $0^\pm$ states with the gluon. Here, H is taken to be the SM 0^+ state.

The $\mathcal{CP}$ mixing between $0^\pm$ states is completely described by the mixing angle α. $\mathcal{CP}$ violation will be implied for $\alpha \neq 0$ or $\alpha \neq \pi$ (with non-zero coupling constant terms).

For simplicity, when measurement of the couplings in a particular vertex is performed, the corresponding parts of the Lagrangian described in equation IV.2 is selected and all other contributions are set to zero.

While the SM predicts the ggF process to be the dominant production mode, presence of BSM terms in equation IV.2 can modify the relative contributions of VBF and VH processes. As noted in [12], the production cross section for the ggF process is generally proportional to κ_{BSM}^2 whereas that for VBF and VH associated Higgs production processes are proportional to κ_{BSM}^4; it is tempting to perform the coupling analysis with events separated into different production mechanisms. To achieve this, an event categorization scheme (discussed in section V) is introduced to classify each event from the collision data and the Monte Carlo generated samples according to their final state signatures. Discriminant observables (discussed in section VIII) from the event kinematics that are sensitive to each production mode with good background discriminating power and contain rich $\mathcal{CP}$ information about the Higgs boson will then be used to perform Likelihood fitting for finding the best fit value of the Higgs coupling.

V. Event Categorization

Following the Simplified Template Cross Section (STXS) framework suggested by the Higgs cross section working group [8] for performing Higgs boson coupling measurements, each production mechanism is split into exclusive regions of phase space called truth bins. A schematic view of the truth bins defined for each production mechanism is shown in Fig. 4. Due to the low sensitivity to the bbH process, truth bins for the bbH process are merged with ggF because of their similar acceptance. The main purpose of the framework is to maximize sensitivity of the measurements while minimize their theory dependence (hence the theoretical uncertainties).

Based on the truth bins defined in the STXS framework, each $H \rightarrow ZZ^* \rightarrow 4l$ candidate event that has passed the selection criteria with a reconstructed mass of the final state leptons in the range $118 < m_{4l} < 129$ GeV will be classified into one of the eight exclusive categories as shown in Fig. 4. Such classification allows the enrichment of events from a particular production mode. This can benefit the extraction of parity sensitive information about the Higgs boson from observables that have features specific to each production mechanism. Therefore, categorization of events may enhance the isolation of possible BSM effects in the statistical evaluation of the Higgs couplings.

The criteria for classifying the categories are mainly the jet multiplicity and the trans-
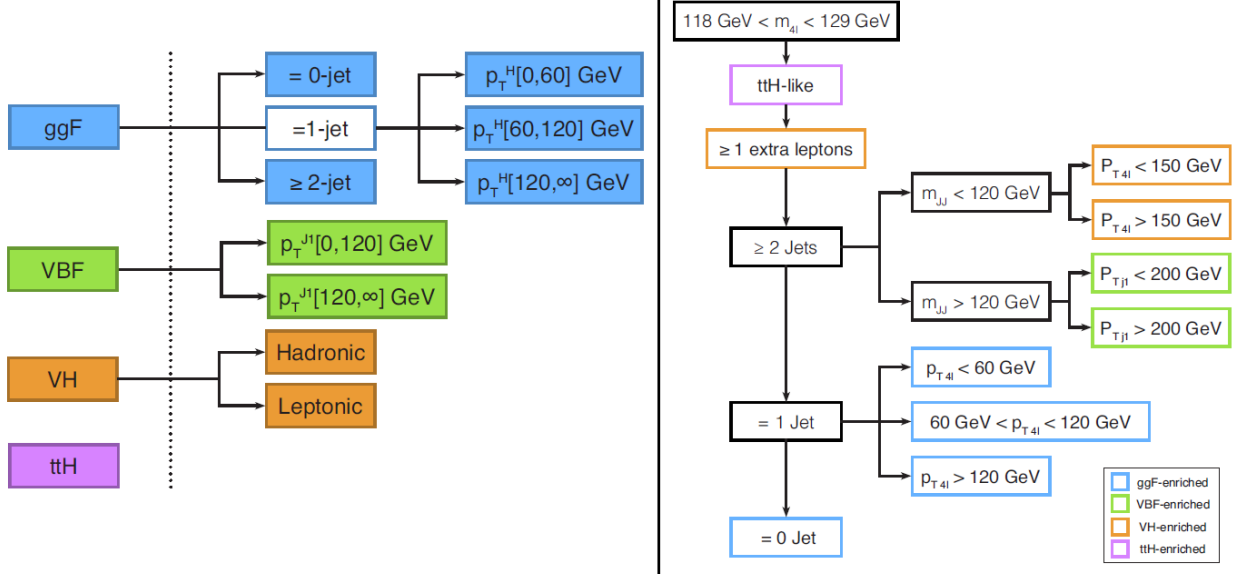
Figure 4: Truth bins defined for each production mode in the STXS framework (left). Event categorization scheme for coupling measurement (right).

verse momentum of the jets and leptons in a reconstructed event. The detailed criteria are summarized in Table I.

Category	enriched mode	Criteria
ttH	ttH	≥ 1 b-tagged jet; ≥ 4 jets or ≥ 2 jets + 1 extra lepton
VH leptonic	VH	≥ 1 extra leptons
2jet lowpt4l	VH	≥ 2 jets ; $m_{JJ} < 120\,\mathrm{GeV}$; $p_{T_{4l}} < 150\,\mathrm{GeV}$
2jet highpt4l	VH	≥ 2 jets ; $m_{JJ} < 120\,\mathrm{GeV}$; $p_{T_{4l}} > 150\,\mathrm{GeV}$
2jet lowptjet	VBF	≥ 2 jets ; $m_{JJ} > 120\,\mathrm{GeV}$; $p_{T_{jet}} < 200\,\mathrm{GeV}$
2jet highptjet	VBF	≥ 2 jets ; $m_{JJ} > 120\,\mathrm{GeV}$; $p_{T_{jet}} > 200\,\mathrm{GeV}$
1jet lowpt4l	ggF	1 jet; $p_{T_{4l}} < 60\,\mathrm{GeV}$
1jet mediumpt4l	ggF	1 jet; $60 < p_{T_{4l}} < 120\,\mathrm{GeV}$
1jet highpt4l	ggF	1 jet; $p_{T_{4l}} > 120\,\mathrm{GeV}$

Table I: Criteria for event categorization. The hierarchy of categorization runs from top to bottom of the table. For the notation: m_{JJ} is the mass of the leading and subleading jet, $p_{T_{4l}}$ is the transverse momentum of final state leptons, $p_{T_{jet}}$ is the transverse momentum of the leading jet.

It should be noted that in the final input Monte Carlo sample for statistical evaluation,

8

events created from Vector Boson Fusion (VBF) and VH production with hadronic decays are merged into a single sample because they are indistinguishable from each other at higher orders of calculation [8].

The relative composition of the production modes in the eight categories is shown in Fig. 5. The categorization scheme generally shows a good enrichment in the corresponding production mode for ggF, VH leptonic and ttH processes. A relatively high contamination from ggF processes is observed in the VBF and VH Hadronic enriched category. To further discriminate between the production modes, discriminant observables sensitive to each production modes can be used (See Section VIII).
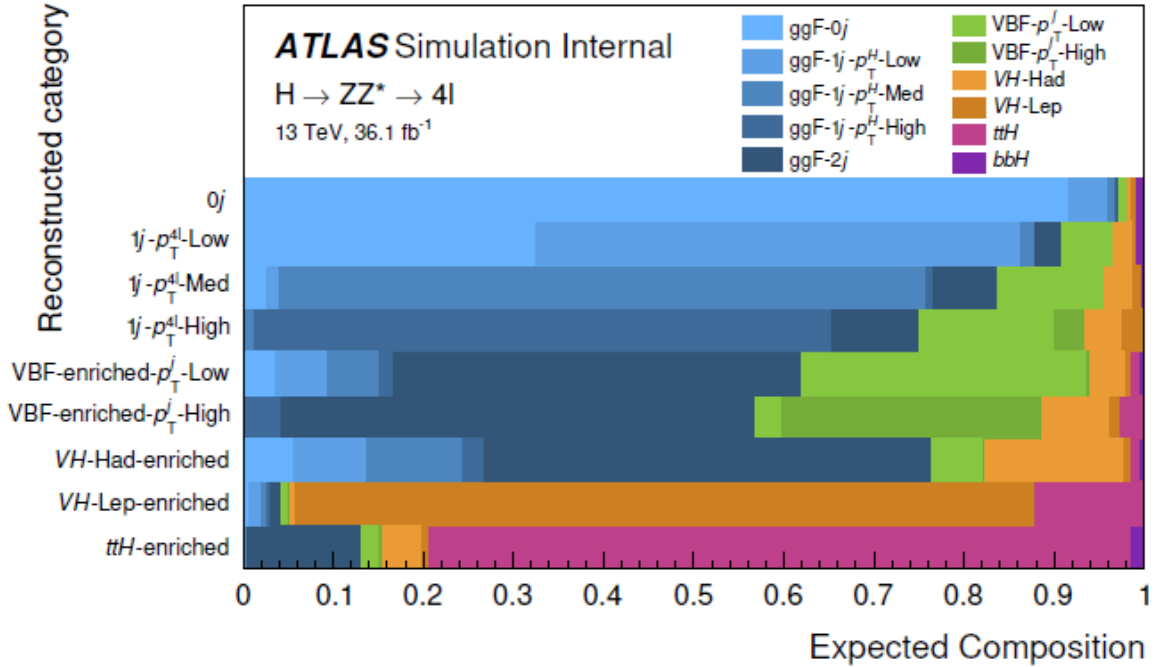

Figure 5: Relative composition in categories with truth bins in different production modes.

VI.   Event Generation

Monte Carlo simulated events are used to model the expected signal and background processes under different signal hypotheses (different input couplings) for comparison with the observed data collected by the ATLAS detector.

In this analysis, signal MC samples are generated under different coupling hypotheses using the MadGraph5_aMC@NLO event generator which incorporates the Higgs characterization model. A list of background and signal MC samples used in this analysis can be found in Appendix XIII B . All data samples are generated at $\sqrt{s} = 13\,\text{TeV}$. A detailed description of the event generation can be found in the analysis note [12].

The aspect of event reconstruction and selection is beyond the scope of this report and will not be discussed here. A detailed description on the event reconstruction and selection in the $H \to ZZ^* \to 4l$ channel can be found in the supporting note [13].

VII.   Kinematic Distributions

The major work of this study begins with an investigation of the event kinematics in the $H \to ZZ^* \to 4l$ decay channel using the fully reconstructed MC samples on top of the experimental data. The channel provides access to a wide range of kinematic variables, notably the angular distributions and transverse momentums of the jets and leptons in the final state as well as the invariant mass of the on shell and the off shell Z bosons from the Higgs decay. With a knowledge of how the distributions of individual observable are modified according to their production modes (which are roughly represented by their categories) and according to different coupling hypothesis, one can construct category based discriminant observables to probe possible BSM couplings in each production mode (i.e. in the ggH or HVV vertex). A list of kinematic observables used in this analysis and their definitions can be found in Appendix XIII C.

For a fair comparison between the distributions derived from simulated events based on the SM prediction (Null hypothesis) and the observed data collected by the detector, preliminary checks on their kinematic distributions are performed.

To reduce statistical errors, simulated signal and background events are usually generated with greater statistics than that collected by the experiment. Proper weighings of the

simulated events are therefore required to ensure a correct normalization with respect to the experimental luminosity of $36.1\,\text{fb}^{-1}$. Moreover, background modelings from MC simulation is usually carried out at a higher precision than signal modelings due to their much larger cross sections relative to the signals. Signal simulation at tree level or Next-to-Leading Order (NLO) precision usually underestimates the true signal yields which therefore ensure a conservative result when testing against the SM (Null) hypothesis. To account for higher order corrections in the signal cross sections, scale factors precise up to Next-3-Leading Order (N3LO) in QCD are applied to ggF and VBF signal processes due to their greater contributions in the signal region. In this analysis, the BSM signal events are assumed to have the same scale factors as the SM signal events. Details about the scaling factors are discussed in Appendix XIII E

Further treatment for reducible background recovery is carried out through data–driven background estimations which scale the reducible background yields (primarily for the Z+jets) in the simulated samples. Details about the data–driven background can be found in Appendix XIII D.

Several kinematic distributions are illustrated bellow to show some general features of the $H \rightarrow ZZ^* \rightarrow 4l$ channel and to test for agreement between simulated and observed data. MC simulated distributions are obtained from the luminosity reweighed data samples with and without corrections from signal scale factors and data driven background estimations. In general, contributions from data driven backgrounds are negligible with respect to the dominant ZZ^* background. No BSM contributions are included in these distributions.

The reconstructed four-lepton invariant mass distribution is shown in Fig. 6. A clear resonance of the assumed Higgs boson at a mass of about $125\,\text{GeV}$ can be observed. Improved agreement between simulation and observed data is reached in the signal region after the application of signal scale factors with dominant contribution from ggF production. The peak at $91\,\text{GeV}$ comes from the resonance of the Z boson. The predominant background is the $qq \rightarrow ZZ^*$ decay with minor contribution from the Z+jets backgrounds.

Fig. 7 shows the reconstructed mass distribution of the virtual Z^* boson. The major contribution from the Higgs boson comes from the small tail at the low mass region. A small deviation is observed at the Z boson resonance peak which is probably due to imperfect background modeling. As will be discussed in the next section, the precise shape of the signal contribution is dependent on the $\mathcal{CP}$ nature of the Higgs coupling.
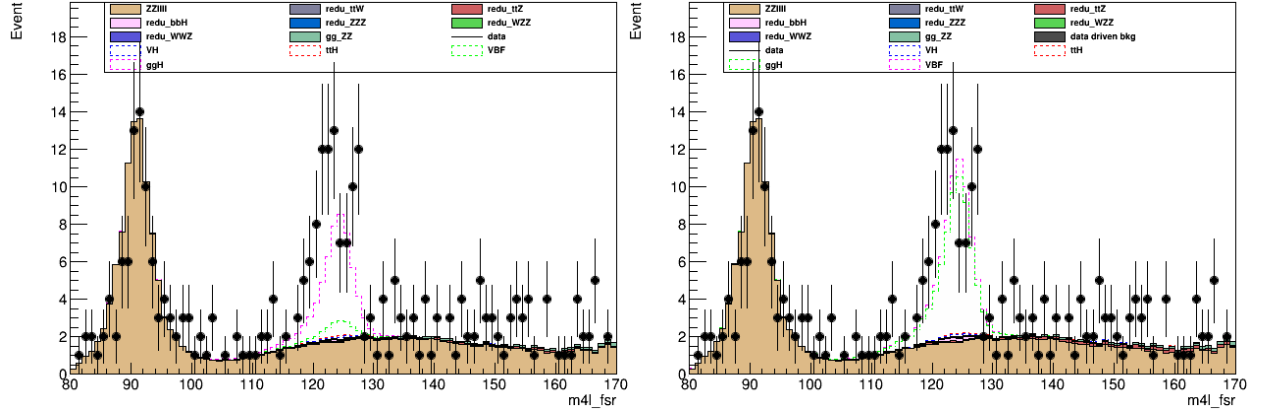
Figure 6: Distributions of the reconstructed four-lepton mass with final state radiation correction obtained from the raw MC samples (left) and from the scale factors reweighed and data driven background applied MC samples (right) at $\sqrt{s} = 13\,\mathrm{GeV}$. Observed data is presented as black dots with Poisson error.
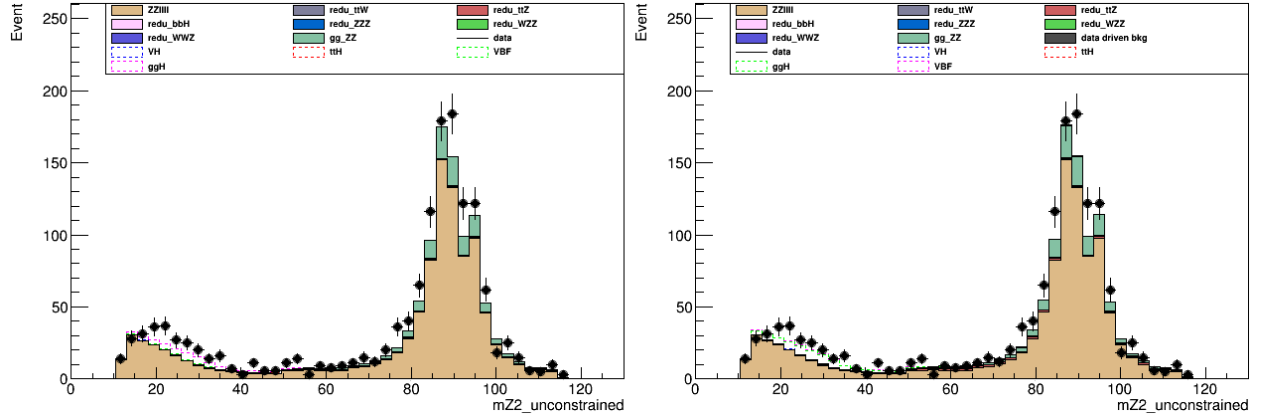

Figure 7: Distributions of the reconstructed off shell Z boson mass obtained from the raw MC samples (left) and from the scale factors reweighed and data driven background applied MC samples (right) at $\sqrt{s} = 13\,\mathrm{GeV}$.

For the jet kinematics, the leading jet angular distribution and the di–jet angular separations are illustrated in Fig. 8 and Fig. 9 respectively.

The distribution of the leading jet pseudorapidity η, presented in Fig. 8 describes the angular distance between the leading jet and the beam line axis. Some unexplained deviations from the observed data are seen at various values of η which may be attributed to backgrounds that are not properly modeled. A majority of the jets are created along the transverse direction so that they have η concentrated at 0. However, in a typical VBF event, two forward jets are produced from the showering of the two outgoing quarks which results

12

in a η distribution bulged at two ends. This feature will be demonstrated in the next section with application for the coupling measurement.

The distribution of the difference in azimuthal angular between two leading jets are illustrated in Fig. 9. As expected, the contribution from VBF process becomes larger in the signal region as most of the ggF events produce no jets or only 1 jet. The central peak shows the fact that most processes produce back-to-back jets with their angular separation being π. This will not be the case for VBF process. Therefore, two-jet observables remain a good discriminant for VBF process.
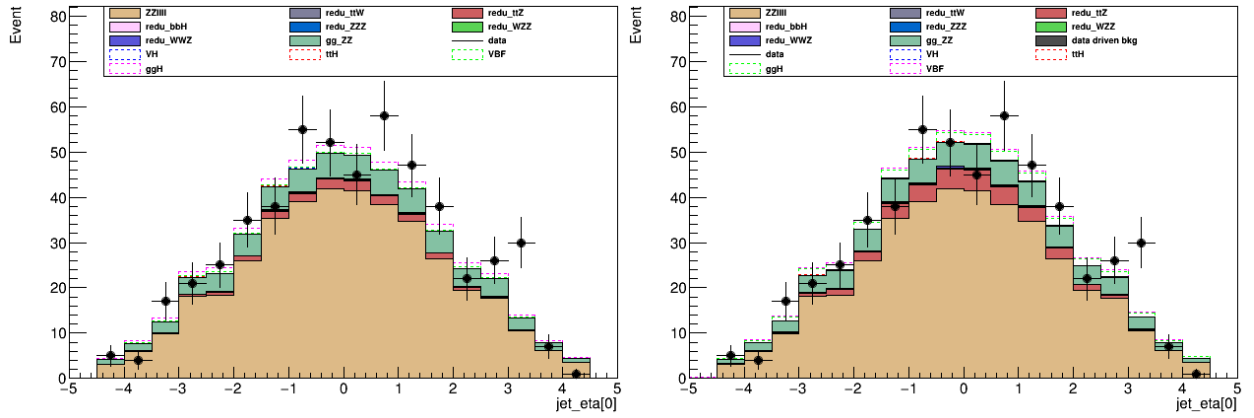
Figure 8: Distributions of the leading jet pseudorapidity obtained from the raw MC samples (left) and from the scale factors reweighed and data driven background applied MC samples (right) at $\sqrt{s} = 13\,\mathrm{GeV}$.

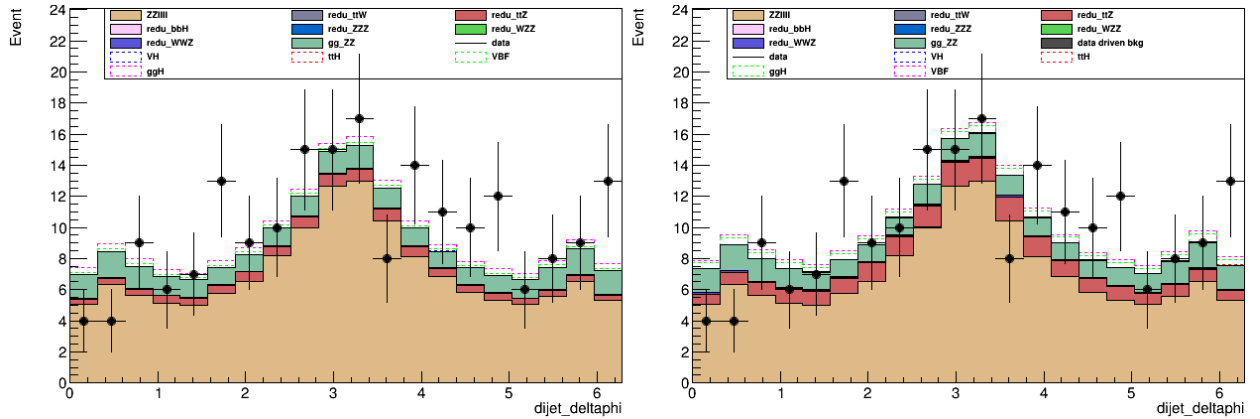Figure 9: Distributions of the difference in azimuthal angle of the two leading jets obtained from the raw MC samples (left) and from the scale factors reweighed and data driven background applied MC samples (right) at $\sqrt{s} = 13\,\mathrm{GeV}$.

13

VIII. Discriminant observables

In this section, discriminant observables sensitive to BSM $\mathcal{CP}$ odd and $\mathcal{CP}$ even couplings are discussed. As mentioned in various literatures, the angular distributions of the final state leptons [6] [11] and that of the jets from ggF and VBF processes [14][15][18] as well as the virtual Z boson invariance mass [11] are all potential discriminants that exhibit large modifications under different $\mathcal{CP}$ hypothesis.

As suggested by [12], the most effective way for constraining different BSM coupling hypothesis would be to construct a discriminant, specific to each category, that combines different sensitive observables to achieve optical discriminating power between signal–background processes, production modes as well as different coupling hypothesis. It utilities the Boosted Decision Tree (BDT) approach to train the best combination between the observables. One instance of the BDT discriminant in the 0–jet category is shown in Fig. 10. These discriminant are already established before this study so they will not be discussed further.
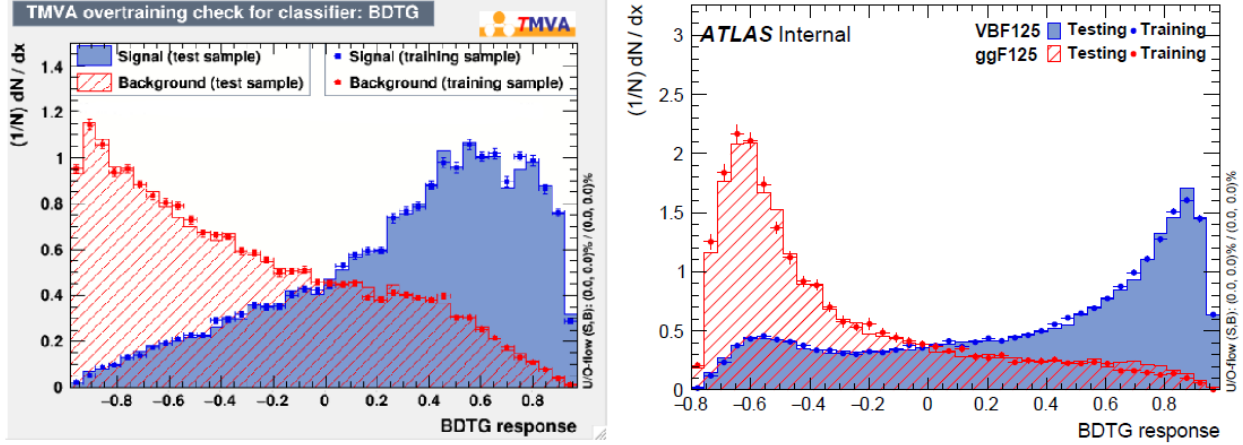
Figure 10: Superimposed training and testing samples for the BDT to separate ggF and ZZ^* events in the 0–jet category (left) and for the BDT to separate ggF and VBF events in the 2–jet category (right).

In this analysis, discriminating power of individual observables for both the $\mathcal{CP}$ nature of the couplings and their sign dependence are investigated. As a remark, the distribution from ggF also contribute to the $\kappa_{A/H_{VV}}$ couplings in the HVV vertex. In ggF production mode, couplings to vector bosons are possible at higher orders beyond tree level, as illustrated by the diagrams in Fig. 11. They contribute to jets arising from the ggF process which contaminate the categories associated with the VBF and VH–hadronic enriched truth bins.

14

Figure 11: Higher order diagrams of gluon gluon fusion process with Z and gluon associated production.

For the probing of BSM couplings, observables of interest include the virtual Z boson mass, difference in azimuthal and polar angle between the leading jets, polar angle of leading jets and the transverse momentum of the four-lepton final state. They are investigated in different production modes in the ggH and HVV vertex. In the following discussion, It should be n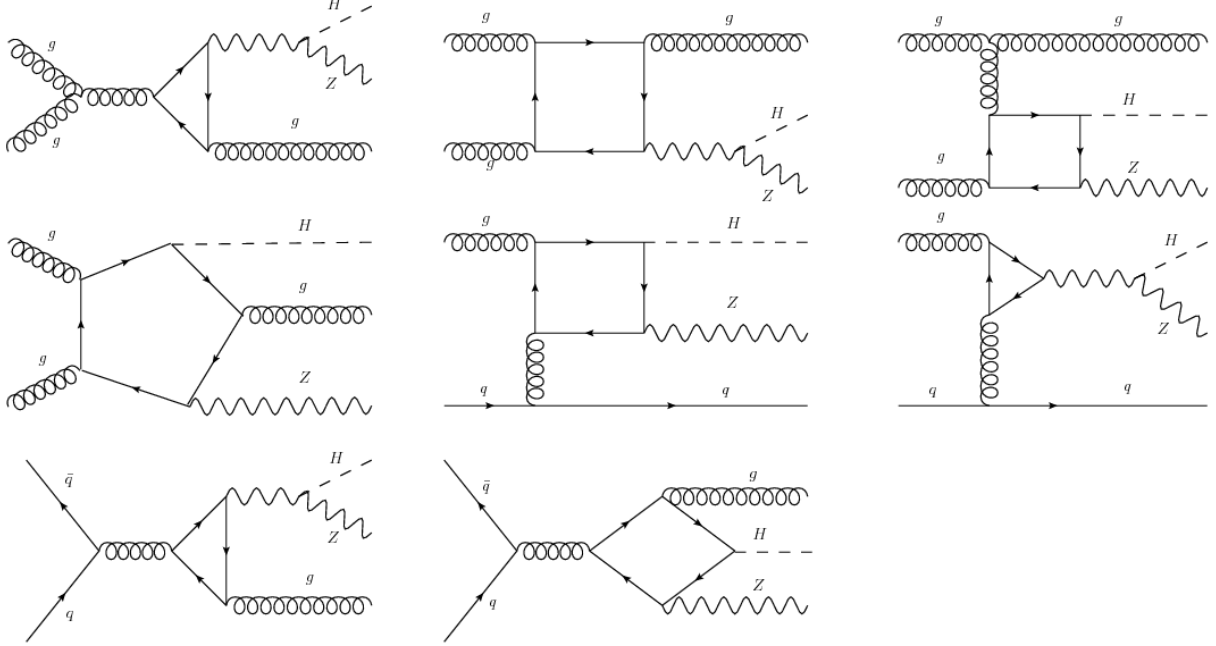oted that the hadronic VH production is merged with the VBF process since they are indistinguishable at higher order calculations. The leptonic VH production is then classified as a VH process without specifying its leptonic nature. For the ggH vertex, the $\mathcal{CP}$ even coupling is fixed to one to resemble a SM coupling in a heavy top quark approximation which reduces the vertex to an effective contact interaction.

Fig. 12 presents the distributions of the virtual Z boson mass in each of the production mechanisms relevant in the ggF and HVV vertex of the Higgs couplings. The presence of pure $\mathcal{CP}$ odd or $\mathcal{CP}$ even couplings both modify the distribution in the HVV vertex. A $\mathcal{CP}$ odd coupling $(H_{A_{VV}})$ generally shifts the virtual Z boson mass to a higher value compared to the SM whereas a $\mathcal{CP}$ even coupling $(H_{H_{VV}})$ has the opposite effect. In the ggF vertex, presence of either BSM couplings has little effect on the virtual Z boson mass distribution. It is noted that in general, most observables are not very sensitive to the BSM couplings in the ggF vertex. This discriminant is applicable to any categories since it does not require

the presence of jets.

For processes with two tagging jets, most observables, such as the transverse momentum and dijet invariant mass are strongly dependent on the form factor which is hard to predict [16]. The azimuthal angle between the two jets are on the other hand insensitive to form factor effects which make it a good candidate for discriminant observable in categories with 2 jets. Fig. 13 presents the distributions of the difference in azimuthal angle between the leading and subleading jets $\Delta\phi_{jj}$ for events with more than 2 jets. In the HVV vertex, the presence of pure $\mathcal{CP}$ odd couplings tend to produce less planar jets (jets at $\Delta\phi = 0$) in VBF processes but more planar jets in the presence of $\mathcal{CP}$ even couplings. For a more theoretical explanation, the behavior in the VBF process can attributed to the cancellation of epsilon tensor contracted with the four linearly dependent momentum vectors of the incoming and outgoing partons by the $\mathcal{CP}$ odd coupling which suppresses planar events **??**. Significant but non trivial modifications are also observed in ggF and VH processes at the HVV vertex as well as ggF process at the ggH vertex with the introduction of BSM couplings. This observable is useful for categories with two or more jets.

Fig. 14 presents the distributions of the difference in pseudorapidity between the leading and subleading jets, and the leading jet pseudorapidity for VBF process in the HVV vertex. Back-to-back jets become preferable with the presence of BSM coupling terms regardless of their $\mathcal{CP}$ status and the leading jet tends to be directed in the transverse direction with respect to the beam line axis. Significant modification from BSM coupling is not observed for processes other than VBF. Therefore, these observables may be used as discriminant variables for VBF specific processes.
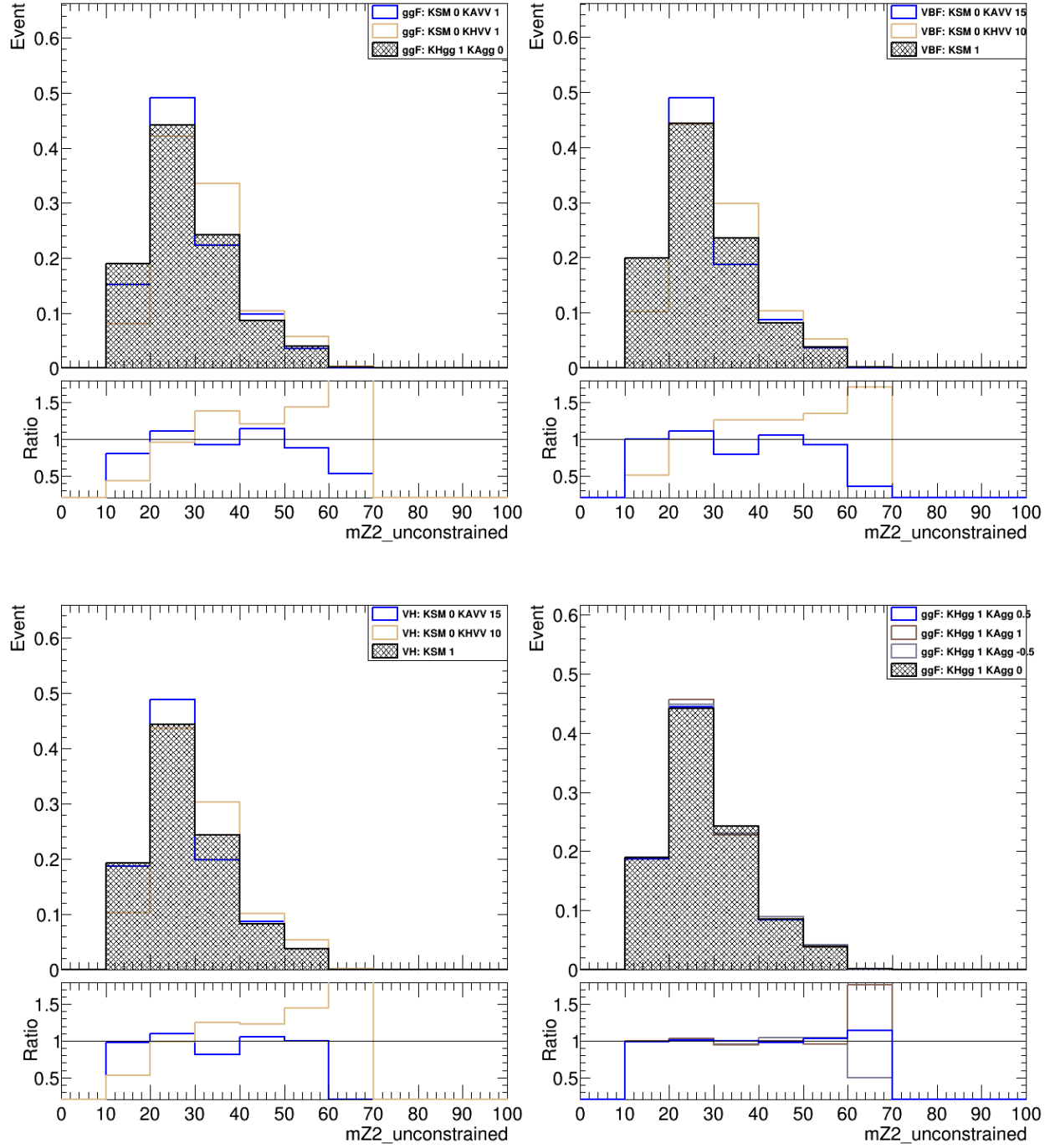
Figure 12: Normalized distributions of the virtual Z boson mass in the HVV vertex for ggF process (top left), VBF process (top right) and VH process (bottom left) and in the ggH vertex for ggF process (bottom right). The effect of pure $\mathcal{CP}$ odd or pure $\mathcal{CP}$ even couplings is presented.

Figure 13: Normalized distributions of the difference in azimuthal angle between the leading and subleading jets in the HVV vertex for ggF process (left), VBF process (center) and VH process (right). The effect of pure $\mathcal{CP}$ odd or pure $\mathcal{CP}$ even couplings is presented.
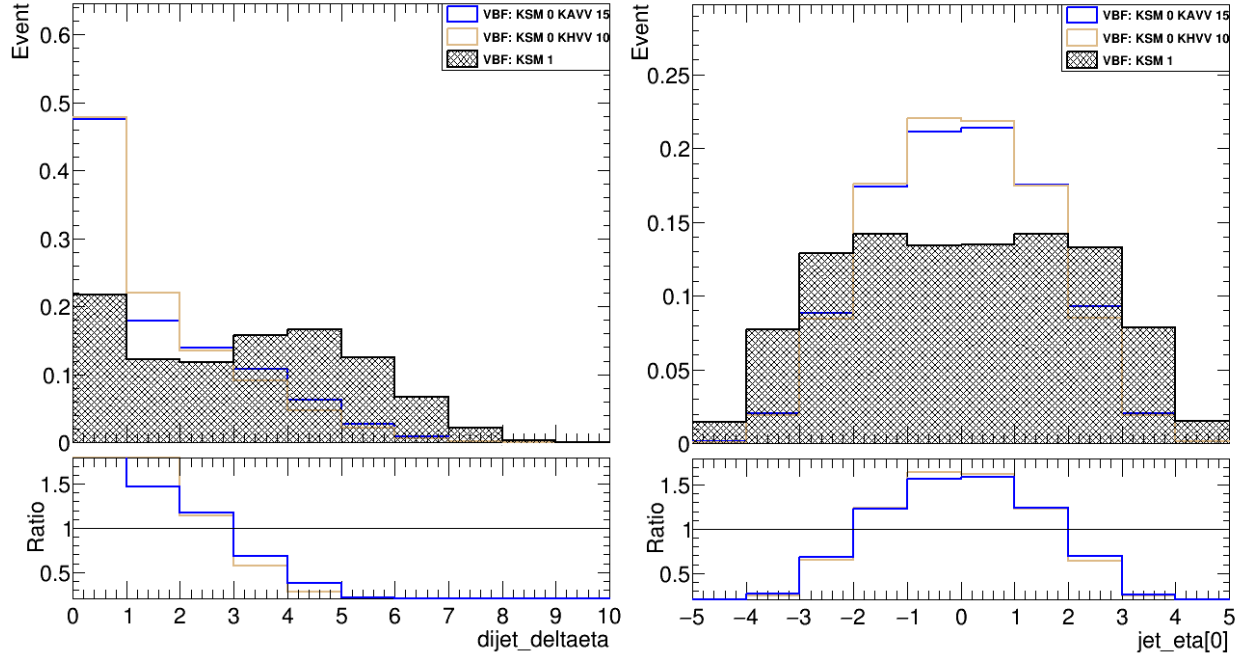
Figure 14: Normalized distributions of the difference in pseudorapidity between the leading and subleading jets in the HVV vertex for VBF process (left) and leading jet pseudorapidity in the HVV vertex for VBF process (right). The effect of pure $\mathcal{CP}$ odd or pure $\mathcal{CP}$ even couplings is presented.

Fig. 15 presents the distributions of the transverse momentum of the four-lepton final state for ggF, VBF and VH process in the HVV vertex and for ggF process in the ggH vertex. Presence of BSM couplings results in considerably larger transverse momentum of the four-lepton final state for weak vector boson processes but shows no effect for ggF process. The observable may be useful to probe BSM effect for weak vector boson processes.

For discrimination in the sign of BSM couplings, the observable $\Delta\phi_{jj}$ may again be used in the HVV vertex as shown in Fig. 16. Note that a SM coupling at its nominal value is also used which will be discussed in the next section. The perturbation to the shape due to the sign of BSM couplings is rather random for ggF and VH leptonic processes which may be easily washed out by statistical fluctuation in the data. On the other hand, the sign of BSM couplings pushes the distribution to opposite end of the spectrum for VBF process. A positive value of $\kappa_{H_{VV}}$ tends towards $\Delta\phi \sim 0$ whereas a negative value tends towards $\Delta\phi \sim |\pi|$. Both positive and negative values of $\kappa_{A_{VV}}$ tend towards $\Delta\phi \sim \pi/2$ but their peak has dependence on the jet transverse momentum (whether it is a leading or subleading jet).

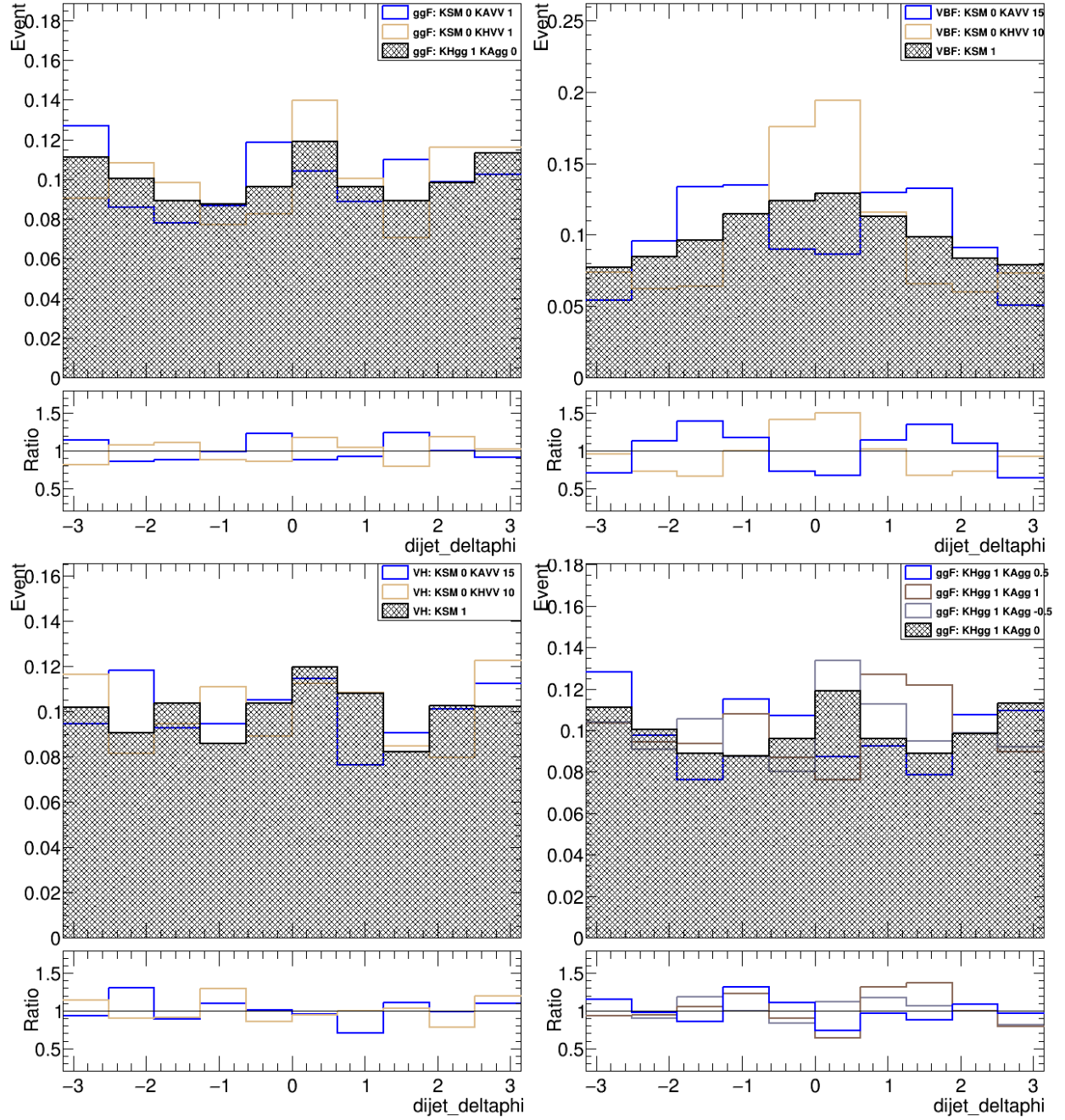Figure 15: Normalized distributions of the transverse momentum of the four-lepton final state in the HVV vertex for ggF process (top left), VBF process (top right) and VH process (bottom left) and in the ggH vertex for ggF process (bottom right). The effect of pure $\mathcal{CP}$ odd or pure $\mathcal{CP}$ even couplings is presented.

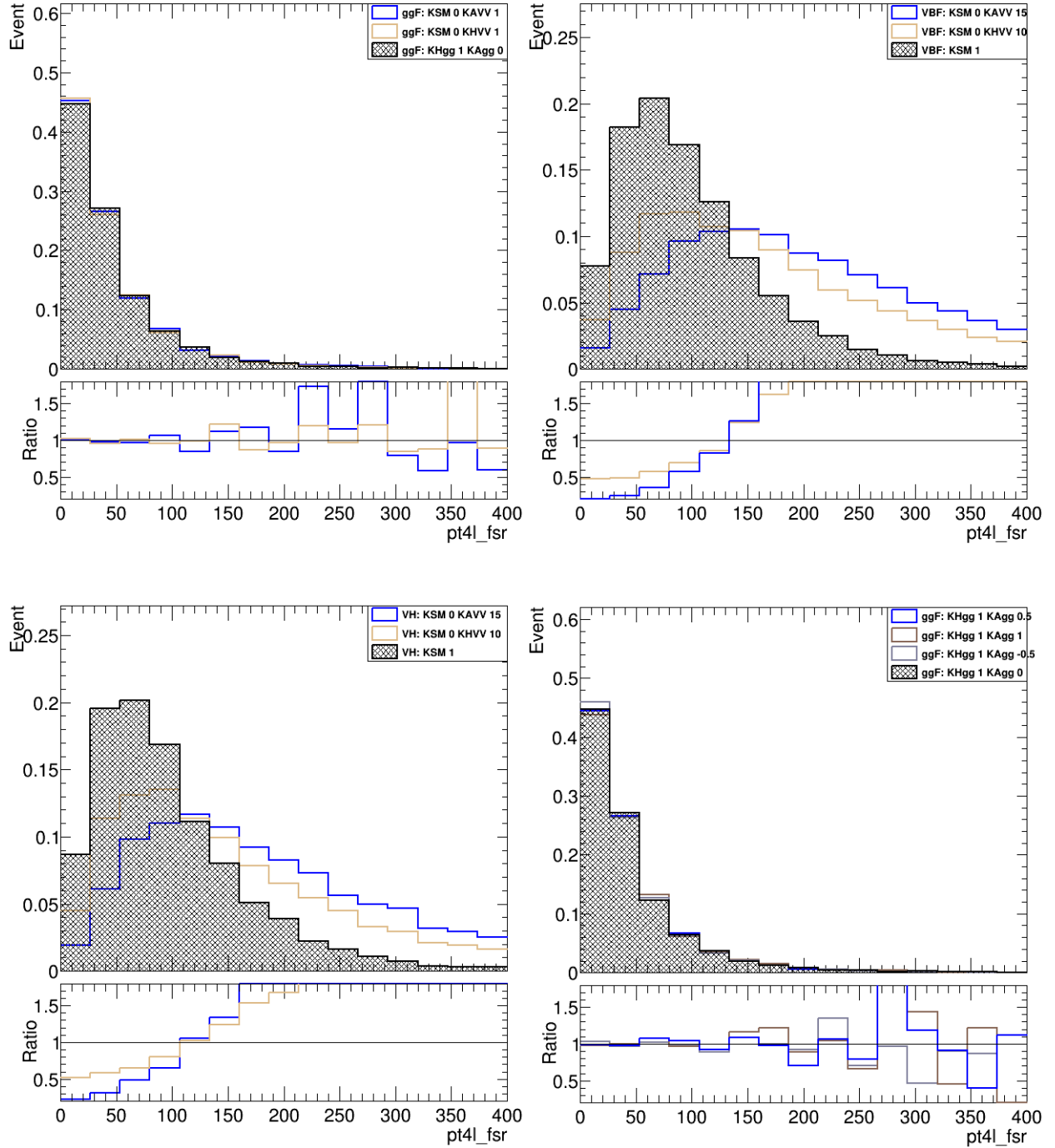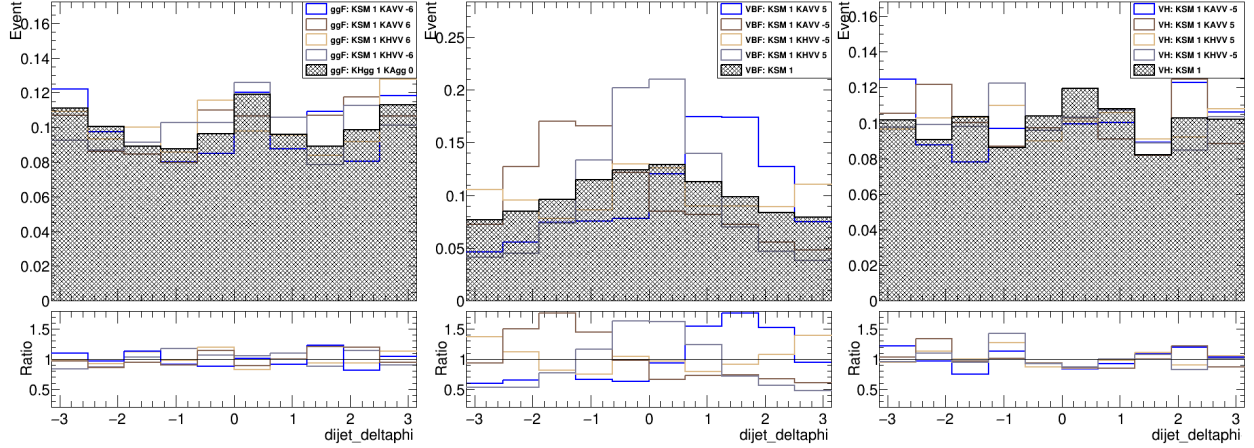Figure 16: Normalized distributions of the observable $\Delta\phi_{jj}$ in the HVV vertex for ggF process (top left), VBF process (top right) and VH process (bottom left) and in the ggH vertex for ggF process (bottom right). The effect of the sign of $\mathcal{CP}$ odd or $\mathcal{CP}$ even couplings mixed with SM couplings is presented.

IX.   Statistical Analysis

The statistical evaluation of the Higgs couplings relies on the distribution of discriminant observables in different categories enriched for specific production mechanisms. The binning of the distributions ranges from 1 to 4. For a one bin analysis, no shape information is derived from the observables, only their normalization. Signal distributions from different coupling hypothesis is chosen according to the production vertex. For analysis at the ggH vertex, MC samples of SM ggF as well as BSM ggF processes with different BSM couplings are used. For other production modes, only the SM samples will be used. The same holds for the analysis at the HVV vertex.

In this analysis, systematic uncertainties are not taken into account in the statistical evaluation. In general, the systematic uncertainties for the predicted signal and background yields can modeled by introduction of nuisance parameters $\vec{\theta} = \theta_1, \theta_2, ...\theta_N$. Suppose the expected signal yield for the a given hypothesis is $S_i^{(\vec{K_x})}(\vec{\theta})$ and that for the background is $B_i(\vec{\theta})$ in each bin i of the discriminant observable. The number of observed events N_i in the i-th bin follows the Poisson distribution

$$\mathrm{Pois}(N_i | S_i^{(\vec{K_x})}(\vec{\theta}) + B_i(\vec{\theta})) = \frac{(S_i^{(\vec{K_x})}(\vec{\theta}) + B_i(\vec{\theta}))^{N_i}}{N_i!} \exp\left[-(S_i^{(\vec{K_x})}(\vec{\theta}) + B_i(\vec{\theta}))\right] \qquad \text{(IX.1)}$$

A likelihood function that depends on the coupling hypothesis $\vec{K}_x$ of the signal, $\mathcal{L}(\vec{K}_x, \vec{\theta})$, is constructed as a product of conditional probabilities over the rate information in each category,

$$\mathcal{L}(\text{data}|\vec{K}_x, \vec{\theta}) = \prod_{i=1}^{N_{\text{categories}}} \text{Pois}(N_i|S_i^{(\vec{K}_x)}(\vec{\theta}) + B_i(\vec{\theta})) \times \mathcal{A}_i(\vec{\theta}), \qquad (\text{IX.2})$$

where $\vec{K}_x$ is the choice of the couplings for a specific hypothesis. $\mathcal{A}_i(\vec{\theta})$ represents the constraints of some nuisance parameters from auxiliary measurements.

Due to limited MC samples for different coupling hypotheses, intermediate values of the couplings that are not sampled are modeled using the morphing technique described in **??** which is able to reproduce the correct normalization and shape of discriminant observables for all coupling hypothesis from a few base MC samples.

A profile likelihood ratio is used as a test statistic,

$$q_\mu = -2\ln\frac{\mathcal{L}(\vec{K}_x|\hat{\hat{\vec{\theta}}}(\vec{K}_x))}{\mathcal{L}(\vec{K}_x|\hat{\vec{\theta}})}, \qquad (\text{IX.3})$$

Here $\hat{\hat{\vec{\theta}}}(\vec{K}_x)$ denotes the value of θ that maximizes the likelihood in the numerator under the coupling hypothesis K_x, and the denominator is the globally maximized likelihood. A higher value of q represents greater incompatibility between the observed data and the hypothesized coupling K_x.

Exclusion lines on the coupling hypotheses can then be drawn at 68% (1 σ) and 95% confidence levels (CL) which constrain the possible values of the couplings to within the range below the exclusion lines at the corresponding CL.

x.   Results

In this analysis, the mixing of SM coupling with either of the BSM $\mathcal{CP}$ odd or BSM $\mathcal{CP}$ even couplings is investigated. Couplings are independently measured in each of the ggH and HVV vertex, meaning that BSM couplings in the other vertex are set to zero. For simplicity, systematic uncertainties are not included in the statistical evaluation.

Different sets of category–specific discriminant observables are considered and their relative performance are compared. The distributions used for Likelihood fitting consist of SM backgrounds fixed in shape and normalization with the addition of SM signals and BSM signals which are to be fitted with two set of data, namely the Asimov data and the observed data. The Asimov data consist from purely SM backgrounds and signals at their nominal value generated from the same MC dataset which provides a check for the statistical procedure for the null hypothesis. The observed data is what actually constrains the BSM coupling hypothesis.

Several binnings are considered for the fitting, ranging from one to four bins in each distribution. Distributions with one bins are minimal and only provide information about the normalization (i.e. cross section) from a particular hypothesis whereas distributions with more than one bin provides both information about the shape and the normalization which should in principle give rise to better constraints on both the sign and the magnitude of a particular coupling hypothesis.

Four sets of discriminant observables are used in this analysis and their summaries are shown in Table II and III. For simplicity, same set of discriminant observables are used for measurement of couplings at different vertex.

Category	Set One
0jet	BDT_discriminant
$1\mathrm{jet}_{p_T^{4l}\mathrm{High}}$, $1\mathrm{jet}_{p_T^{4l}\mathrm{Medium}}$, $1\mathrm{jet}_{p_T^{4l}\mathrm{Low}}$	BDT_OneJet_jptetadR_discriminant
$2\mathrm{jet}_{\mathrm{VBF},p_T^{\mathrm{jet}}\mathrm{High}}$, $2\mathrm{jet}_{\mathrm{VBF},p_T^{\mathrm{jet}}\mathrm{Low}}$	BDT_TwoJet_discriminant
$2\mathrm{jet}_{\mathrm{VH},p_T^{4l}\mathrm{High}}$, $2\mathrm{jet}_{\mathrm{VH},p_T^{4l}\mathrm{Low}}$	BDT_VH_noptHjj_discriminant
ttH	M4l_constrained_HM
VH leptonic	M4l_constrained_HM

Table II: Summary of the observables used in discriminant "Set One".

Category	Set Two	Set Three	Set Four
0jet	$M_{\text{Z2, unconstrained}}$	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr
$1\text{jet}_{p_T^{4l}\text{High}}$, $1\text{jet}_{p_T^{4l}\text{Medium}}$, $1\text{jet}_{p_T^{4l}\text{Low}}$	leading jet η	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr
$2\text{jet}_{\text{VBF},p_T^{\text{jet}}\text{High}}$, $2\text{jet}_{\text{VBF},p_T^{\text{jet}}\text{Low}}$	$\Delta\eta_{jj}$	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr
$2\text{jet}_{\text{VH},p_T^{4l}\text{High}}$, $2\text{jet}_{\text{VH},p_T^{4l}\text{Low}}$	$\Delta\phi_{jj}$	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr
ttH	$p_{T,4l}$ fsr	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr
VH leptonic	p_{T_4l} fsr	$M_{\text{Z2, unconstrained}}$	$p_{T,4l}$ fsr

Table III: Summary of the observables used in discriminant "Set Two", "Set Three" and "Set Four".

In the following, results of the coupling analysis are presented. In general, a Likelihood scan of any BSM coupling hypothesis using the Asimov dataset should converge at a zero value corresponding to the SM expected value (Null hypothesis). A narrower Likelihood will exclude a large phase space of the BSM coupling hypothesis. Exclusions can be made when the coupling hypothesis lies outside the 95% CL line which corresponds to a value of 3.841 of the negative log Likelihood (which represents a chi square distribution). A symmetric Likelihood is expected for a $\mathcal{CP}$ odd coupling fitted using distributions with one bin only due to the fact that the cross sections of any production modes are generally proportional to κ^2 or κ^4 which results in equal yields regardless of the sign of the couplings. Asymmetries may be resulted when statistical fluctuation exists between the two signal distributions generated with $\mathcal{CP}$ odd couplings of opposite signs. The symmetry is expected to be lifted when distributions with more bins are used which contain information about the shape of the distributions. Asymmetry towards a positive value of $\mathcal{CP}$ even coupling is expected due to interference from SM signals. When the SM coupling is allowed to vary about its nominal value (which should generally be allowed if BSM effect actually exists), an extra degree of freedom is added to the fitting so more combinations of the SM coupling and the BSM coupling can reproduce the distribution that is closed to the observed one. Hence, fitting with a free SM coupling usually results in a worse result so that less constraint can be put on the BSM couplings. A summary of the Likelihood statistics can be found in Appendix XIII F.

In Fig. 17, Fig. 18 and Fig. 19, the results obtained using BDT discriminants (Set One)

with one, two and four bin distributions respectively are presented. Exclusions allowed for $\kappa_{A_{gg}}$ is generally much larger than for $\kappa_{A_{VV}}$ or $\kappa_{H_{VV}}$ due to much larger cross section for the ggF process (no couplings to vector bosons at leading order). A small asymmetry is observed for $\kappa_{A_{VV}}$ towards a negative value which is slightly lifted when more bins are introduced but the effect is probably too small to affect the exclusion at 95% CL beyond statistical uncertainties. As expected, increasing the number of binning improves the results by constraining more values of the BSM couplings for both $\kappa_{A_{VV}}$, $\kappa_{H_{VV}}$ and $\kappa_{A_{gg}}$ as the shape information is also included. The improvement is very small due to limited statistics (total yields of signal event is less than 100) so further increasing the binning may not increase the sensitivity by much. With increased binning, all coupling hypothesis converges towards the SM hypothesis indicating consistencies of the observed data with the SM prediction. It is noted that the result of $\kappa_{A_{gg}}$ is missing for four bin distributions due to convergence issues.

To compare the discriminating power of various discriminant observables, the results obtained using discriminant set two, set three and set four with one bin distributions are shown in Fig. 20, Fig. 22 and Fig. 23 respectively. The choice of the discriminants for set two is based on the observation in section VIII which show good sensitivity in the respective categories (or production modes). Some arbitrary choices are made for discriminants used in set three and set four with the primary purpose to investigate the discriminating power of a single observable. In general, the shapes of the Likelihood obtained from set two and set three are similar to that of the BDT discriminants which demonstrates that they all possess the discriminating power for separation of BSM effects. Still, the results obtained from BDT discriminants have narrower Likelihood showing that they have greater sensitivity compared to the set two and set three discriminants. It is noted that using the virtual Z boson mass alone as discriminants in set three actually produces reasonably good results particularly in the HVV vertex. This is expected as it the presence of BSM couplings shows largest modifications for ggF, VBF and VH processes in the HVV vertex but less in the ggH vertex. The results obtained using purely $p_T 4l$ as discriminants (set four) show no discriminating power in both HVV and ggH vertex. This is also expected as $p_T 4l$ only shows large modification for VBF and VH processes but their cross sections are far smaller than ggF process. Still, the interference with SM can be observed for $\kappa_{H_{VV}}$ which causes the best fit result to lie at a positive value.

Figure 17: Negative log likelihood plots without systematics for Asimov and observed data using one bin distributions from "Set One" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).

Figure 18: Negative log likelihood plots without systematics for Asimov and observed data using two bin distributions from "Set One" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).

Figure 19: Negative log likelihood plots without systematics for Asimov and observed data using four bin distributions from "Set One" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).

Finally, to study the discriminating power on the sign of the BSM couplings, results obtained using discriminant set two with four bin distributions are shown in Fig. 21. Disappointingly, the resulting Likelihoods have similar shapes compared to the one bin Likelihoods. It may be due to the fact that the bin resolution is far from enough to discriminate the difference in shapes obtained from BSM couplings of opposite signs. It is likely that the symmetry between positive and negative couplings will not be lifted further with the current available statistics. Again, convergence issues make the result for $\kappa_{A_{gg}}$ unavailable.
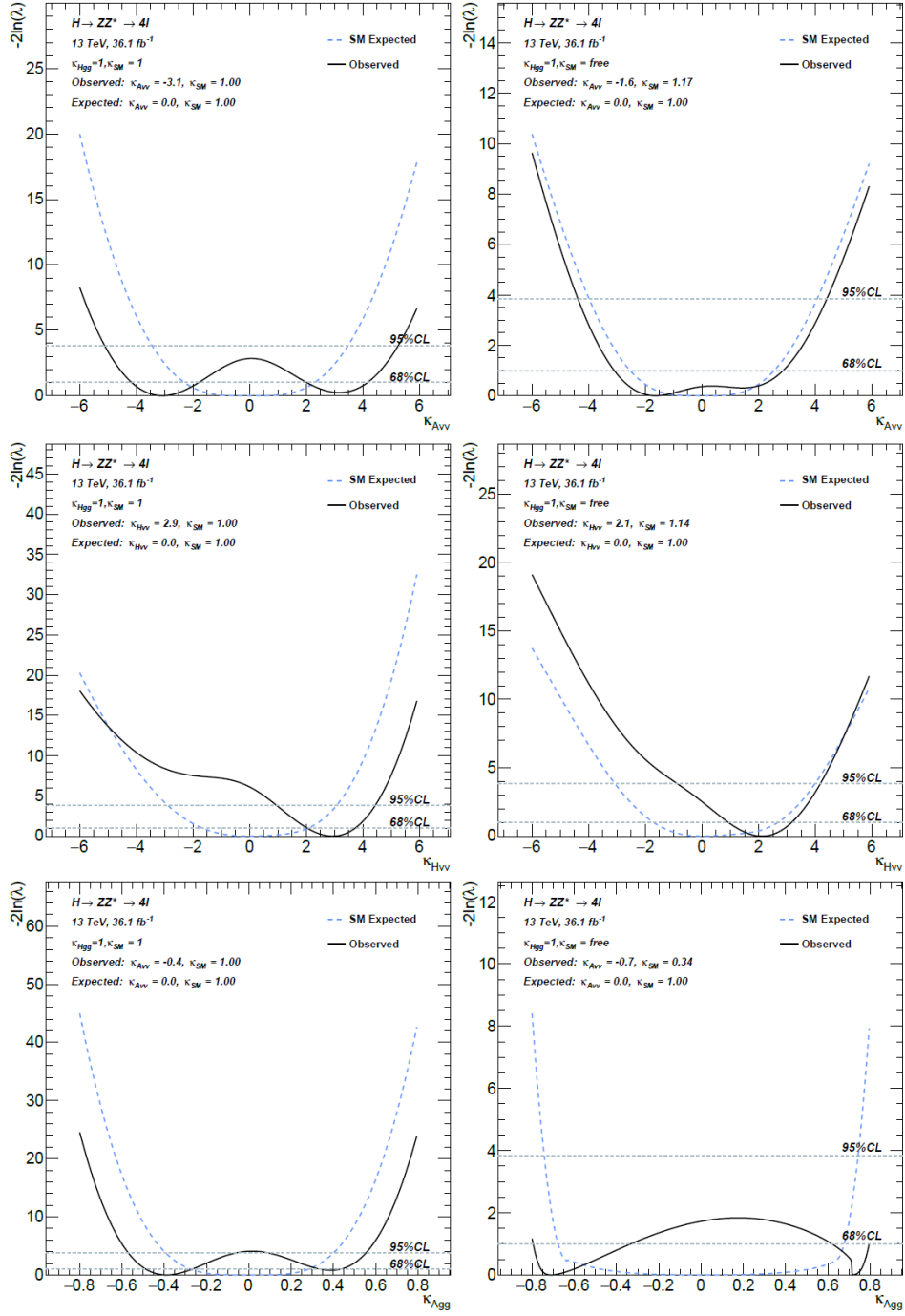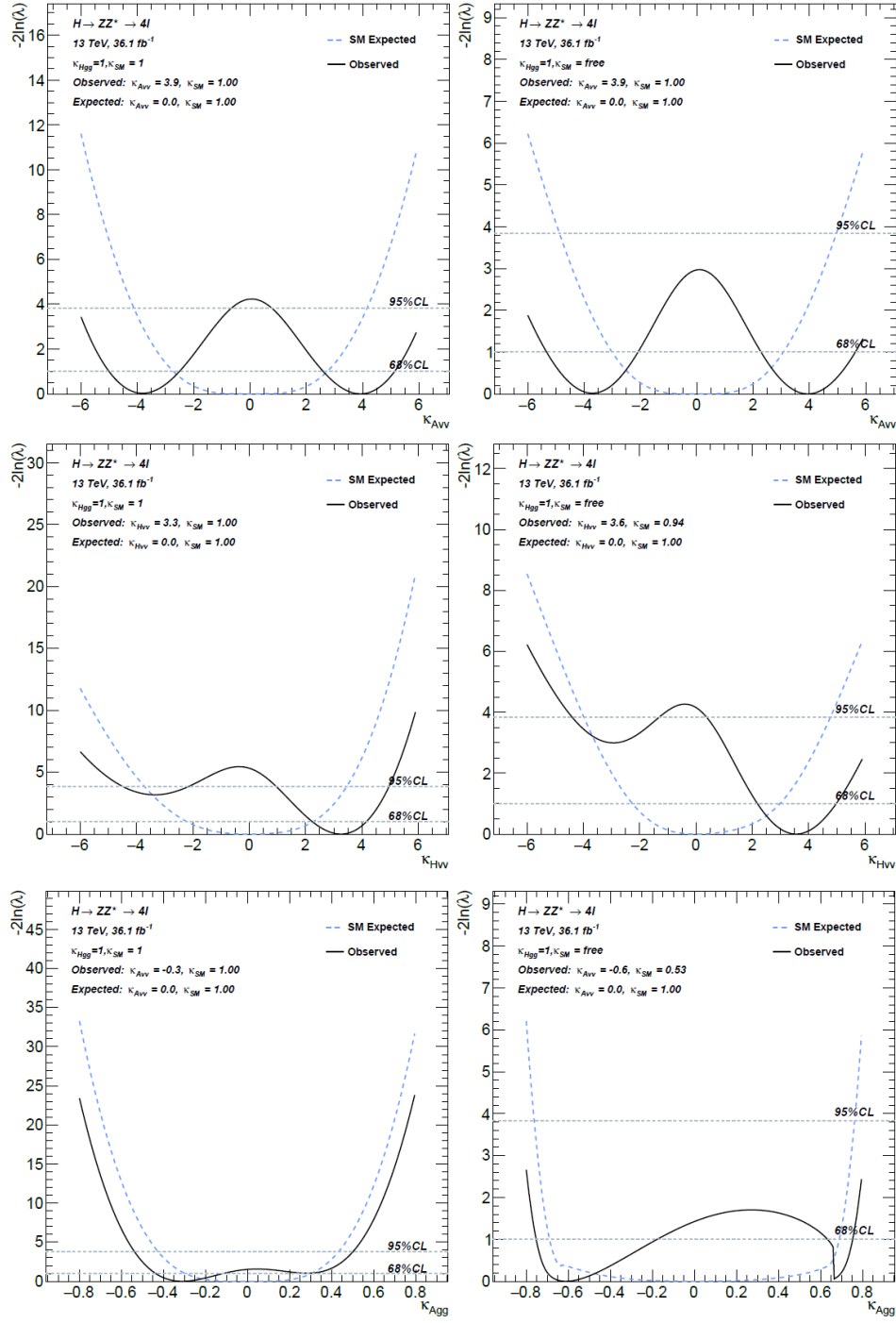
Figure 20: Negative log likelihood plots without systematics for Asimov and observed data using one bin distributions from "Set Two" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).
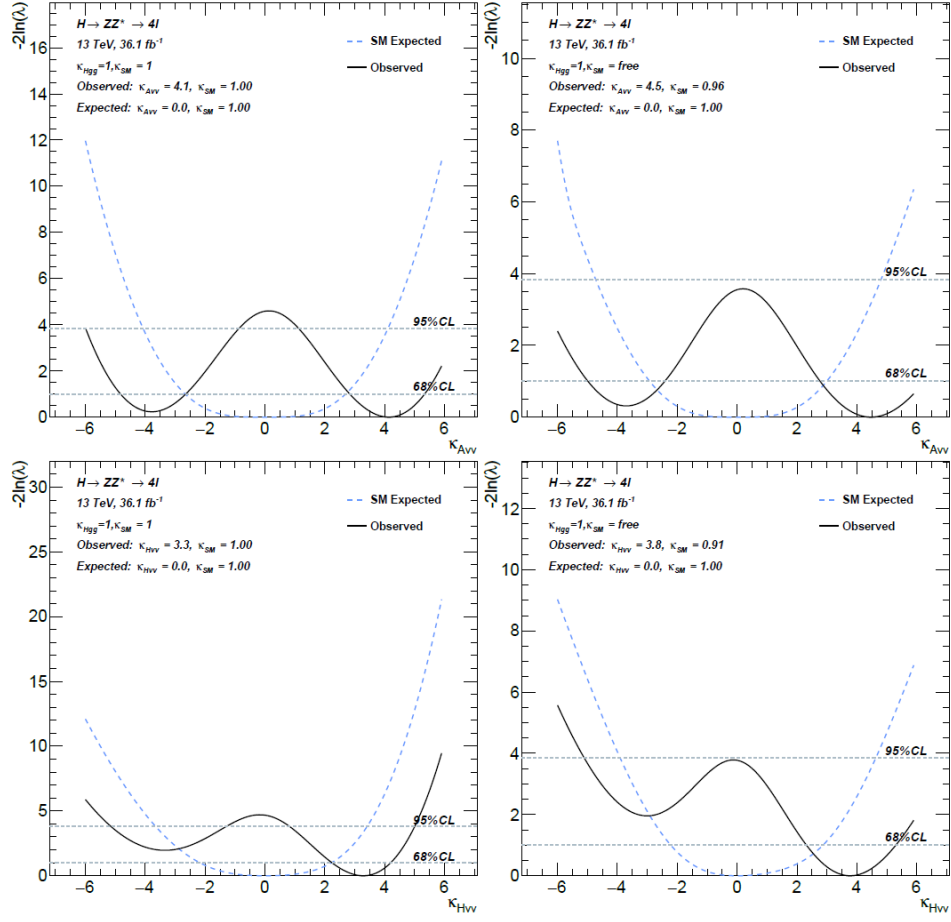
Figure 21: Negative log likelihood plots without systematics for Asimov and observed data using four bin distributions from "Set Two" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).
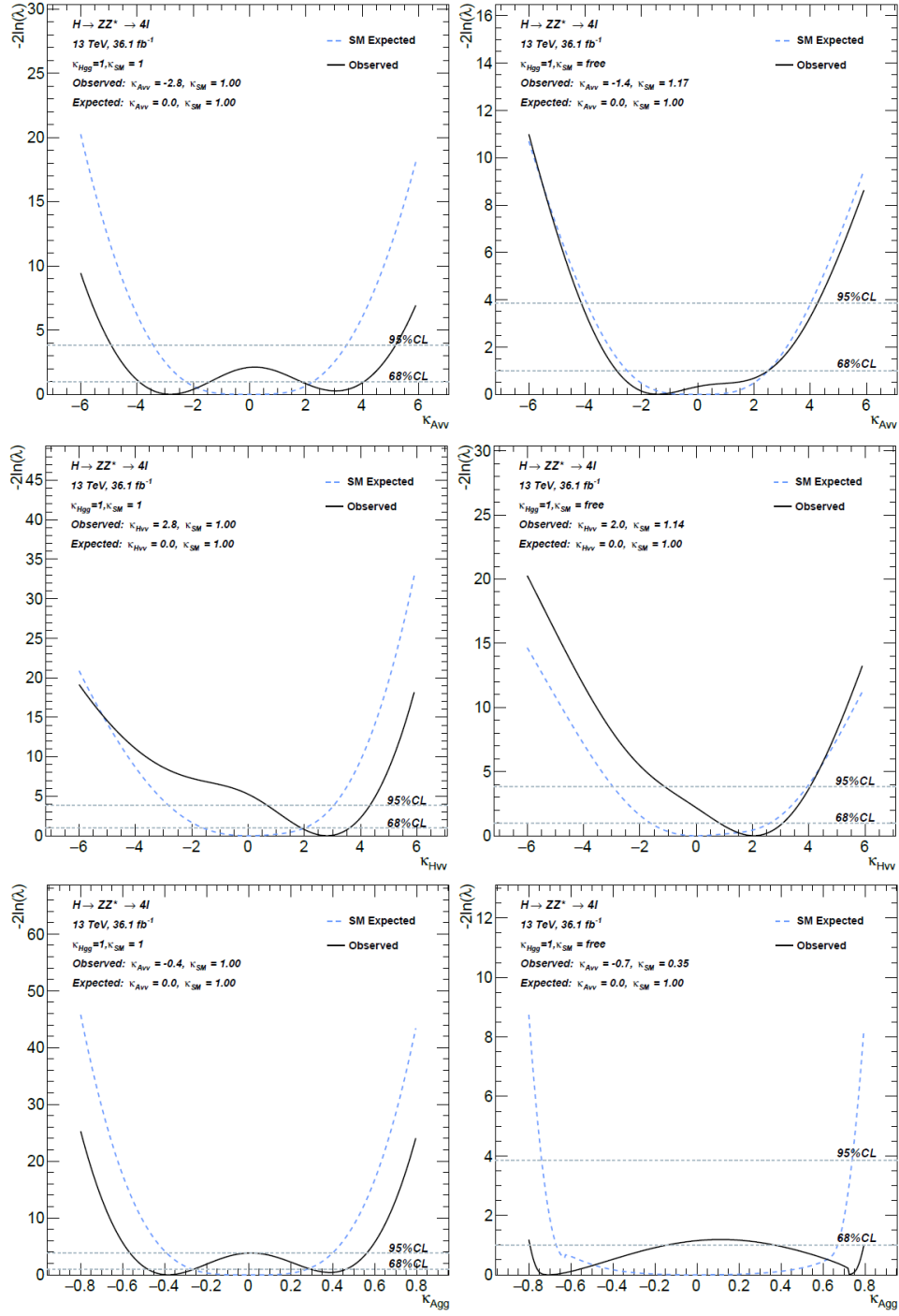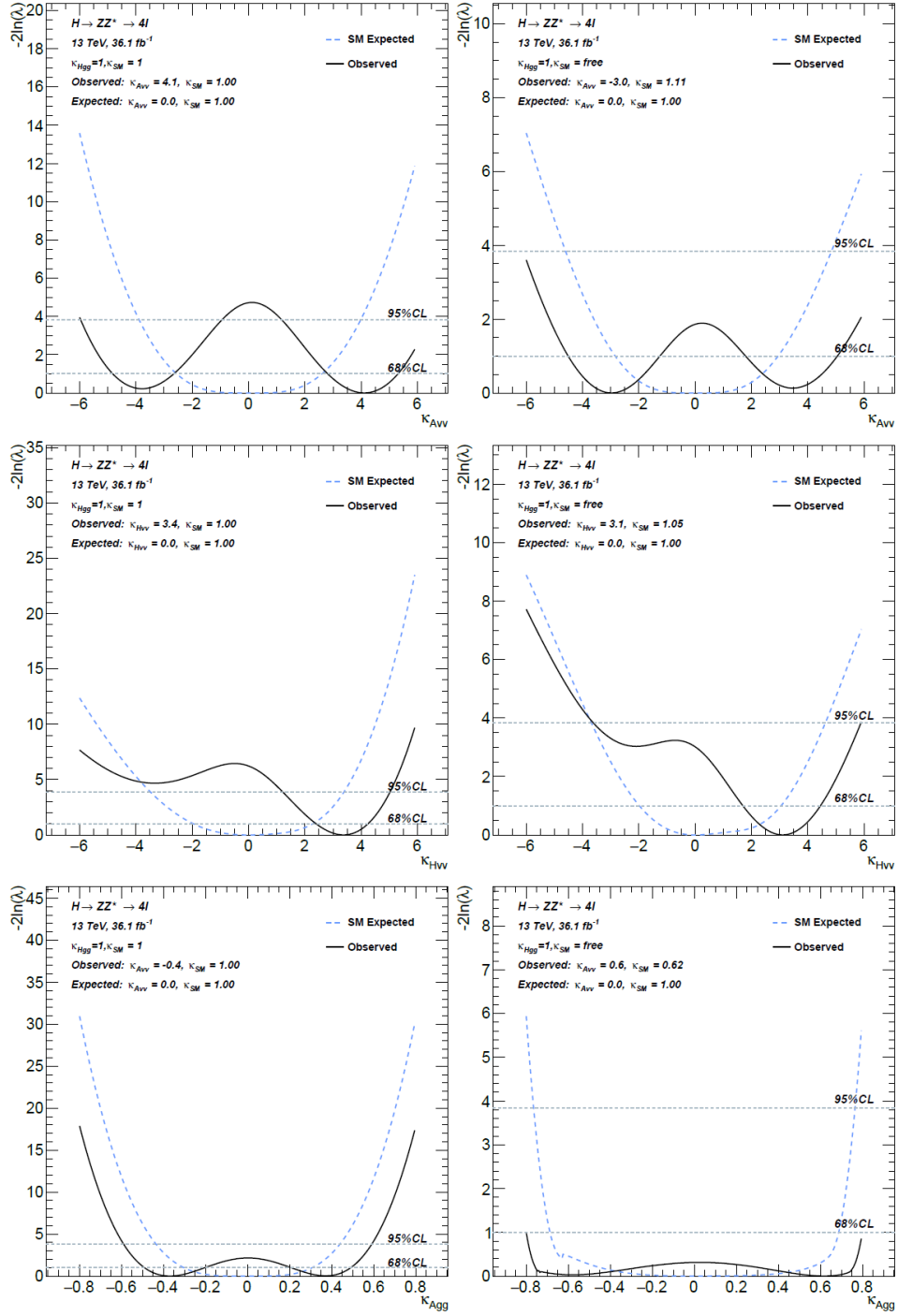
Figure 22: Negative log likelihood plots without systematics for Asimov and observed data using one bin distributions from "Set Three" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).
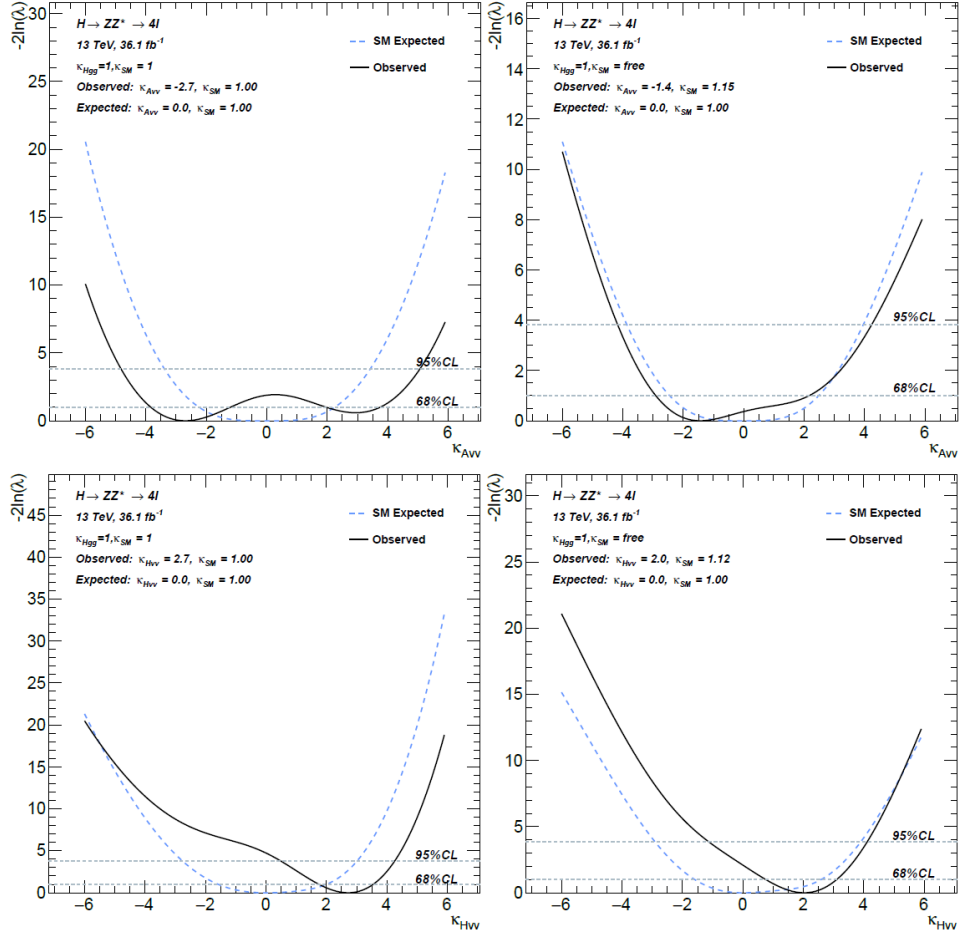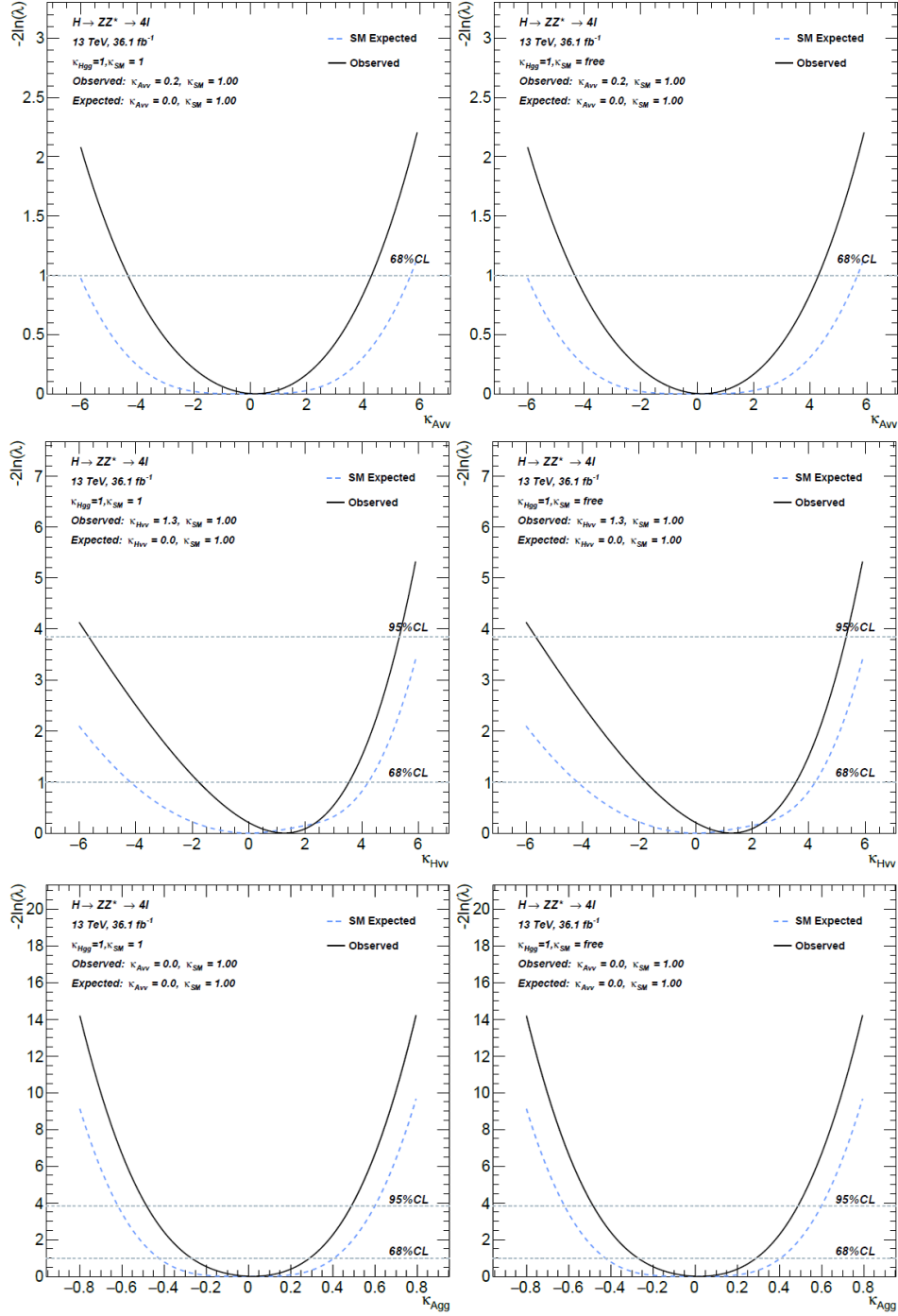
Figure 23: Negative log likelihood plots without systematics for Asimov and observed data using one bin distributions from "Set Four" discriminant observables. The scan for $\kappa_{A_{VV}}$ (top) and $\kappa_{H_{VV}}$ (center) in the HVV vertex and for $\kappa_{A_{gg}}$ (bottom) in the ggH vertex with the SM coupling (κ_{SM} in the HVV vertex and $\kappa_{H_{gg}}$ in the ggH vertex) being fixed to its nominal value (left) or held free (right).

XI. Conclusion

Based on the theoretical framework of the Higgs Characterization model, measurement of Higgs couplings with possible $\mathcal{CP}$ mixing is performed in the $H \rightarrow ZZ^* \rightarrow 4l$ decay channel using proton-proton collision data with an integrated luminosity of $36.1\,\mathrm{fb}^{-1}$ at a center–of–mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ collected by the ATLAS detector.

In particular, mixing of SM couplings with BSM $\mathcal{CP}$ odd and BSM $\mathcal{CP}$ even couplings is investigated in the ggH and HVV vertex for the ggF, VBF and VH production modes. In the analysis, all events are classified into one of the eight production mode enriched categories according to the STXS framework to enhance sensitivity to BSM effects and reduce theoretical uncertainties. Four sets of category–specific discriminant observables are introduced and their discriminating power is discussed.

In the statistical evaluation of the BSM couplings, it is generally found that the measurements for each production mode in the ggH and HVV vertex are consistent with SM expectations within 1 to 2.2 standard deviations. No systematic uncertainties are applied in this analysis.

For future work, systematic uncertainties will be included and further properties of the discriminants will be investigated. The results may be extrapolated to higher luminosities to study the impact of systematic uncertainties and the current limit of the sensitivity in the $H \rightarrow ZZ^* \rightarrow 4l$ decay channel. Investigations on the event selection scheme may also be carried out to better isolate Higgs candidates from background events.

XII. Acknowledgement

The author would like to acknowledge his supervisor Luis Roberto Flores Castillo and Postdoctoral Research Associate Valerio Bortolotto for their guidance and support throughout the project. This project will not be possible without their kindly advice and valuable discussions.

The author also acknowledges the Department of Physics at the Chinese University of Hong Kong for organizing the project.

XIII.   Appendix

A.   Interaction Lagrangian of Spin-0 State in the Effective Field Theory

Here we define $c_\alpha \equiv \cos\alpha$, $s_\alpha \equiv \sin\alpha$ and α is the $\mathcal{CP}$-mixing angle. The Standard Model, BSM CP-even and BSM CP-odd spin-0 Higgs fields are collectively denoted as $\mathcal{X}_0$. The constants $\kappa_{\mathcal{X}_j}$ represent the scale factors applied to the Higgs coupling for particle j which correspond to the deviations from the Standard Model Higgs coupling.

The effective Lagrangian for the interaction of scalar H and pseudoscalar A states with fermions is given by:

$$\mathcal{L}_0^f = - \sum_{f=t,b,\tau} \bar{\psi}_f (c_\alpha \kappa_{Hff} g_{Hff} + i s_\alpha \kappa_{Aff} g_{Aff} \gamma_5) \psi_f \mathcal{X}_0, \tag{XIII.1}$$

where g_{Hff} (g_{Aff}) denotes the strength of the scalar (pseudoscalar) coupling in the SM (in a 2 Higgs Doublet Model (2HDM) with $\tan\beta = 1$). ψ_f and $\mathcal{X}_0$ is the fermion field.

The effective Lagrangian for the interaction of scalar H and pseudoscalar A states with vector bosons is given by:

$$
\begin{aligned}
\mathcal{L}_0^V = \Big\{ & c_\alpha \kappa_{\mathrm{SM}} [\frac{1}{2} g_{HZZ} Z_\mu Z^\mu + g_{HWW} W_\mu^+ W^{-\mu}] \\
& - \frac{1}{4} [c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma} A_{\mu\nu} \widetilde{A}^{\mu\nu}] \\
& - \frac{1}{2} [c_\alpha \kappa_{HZ\gamma} g_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} g_{AZ\gamma} Z_{\mu\nu} \widetilde{A}^{\mu\nu}] \\
& - \frac{1}{4} [c_\alpha \kappa_{Hgg} g_{Hgg} G_{\mu\nu}^\alpha G^{\alpha,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg} G_{\mu\nu}^a \widetilde{G}^{a,\mu\nu}] \\
& - \frac{1}{4} \frac{1}{\Lambda} [c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} Z_{\mu\nu} \widetilde{Z}^{\mu\nu}] \\
& - \frac{1}{2} \frac{1}{\Lambda} [c_\alpha \kappa_{HWW} W_{\mu\nu}^+ W^{-\mu\nu} + s_\alpha \kappa_{AWW} W_{\mu\nu}^+ \widetilde{W}^{-\mu\nu}] \\
& - \frac{1}{\Lambda} c_\alpha [\kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial\mu Z^{\mu\nu} + (\kappa_{H\partial W} W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c.)] \Big\} \mathcal{X}_0,
\end{aligned} \tag{XIII.2}
$$

where the reduced field strength tensors are defined as:

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \quad (V = A, Z, W^\pm), \tag{XIII.3}$$

$$G_{\mu\nu}^a = \partial_\nu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c, \tag{XIII.4}$$

and the dual tensor is defined as:

$$\widetilde{V}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma} \tag{XIII.5}$$

The SM scenario can be obtained by setting $c_\alpha = 1$ and $\kappa_{H_{ff}} = 1$. A pseudoscalar state of a type-II CP-conserving 2HDM can be obtained by setting $s_\alpha = 1$ and $\kappa_{A_{ff}} = \cot\beta$ or $\kappa_{A_{ff}} = \tan\beta$ for up or down components of the SU(2) fermion doublet, respectively. $\mathcal{CP}$ violation will be implied when the mixing angle $\alpha \neq 0$ or $\alpha \neq \pi/2$ (and non-vanishing $\kappa_{H_{ff}}, \kappa_{A_{ff}}, \kappa_{H_{VV}}, \kappa_{A_{VV}}$).

B. Input Monte Carlo Samples

In this analysis, all input distributions for kinematic plots and statistical tests are generated from MadGraph5_aMC@NLO generator. Each Monte Carlo sample, denoted as Minitrees, represents a specific process containing events with associated kinematic observables.

Reducible Backgrounds

Sample Name	Process
mc15_13TeV.361621.Sherpa_CT10_WWZ_4l2v	WWZ
mc15_13TeV.361623.Sherpa_CT10_WZZ_5l1v	WZZ
mc15_13TeV.361625.Sherpa_CT10_ZZZ_6l0v	ZZZ
mc15_13TeV.361626.Sherpa_CT10_ZZZ_4l2v	ZZZ
mc15_13TeV.344973.aMcAtNloPythia8EvtGen_A14NNPDF23LO_bbH125_yb2_ZZ4lep_noTau	bbH
mc15_13TeV.344974.aMcAtNloPythia8EvtGen_A14NNPDF23LO_bbH125_ybyt_ZZ4lep_noTau	bbH
mc15_13TeV.410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee	ttZ
mc15_13TeV.410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu	ttZ
mc15_13TeV.410144.Sherpa_NNPDF30NNLO_ttW	ttW

Table IV: Monte Carlo samples used model the expected kinematic distributions for the reducible background contributions.

Irreducible Backgrounds

Sample Name	Process
mc15_13TeV.363490.Sherpa_221_NNPDF30NNLO_llll	ZZllll
mc15_13TeV.345108.Sherpa_221_NNPDF30NNLO_llll_m4l300	ZZllll
mc15_13TeV.345107.Sherpa_221_NNPDF30NNLO_llll_m4l100_300_filt100_150	ZZllll
mc15_13TeV.343212.Powheggg2vvPythia8EvtGen_gg_ZZ_bkg_2e2mu_13Tet	ggZZ
mc15_13TeV.343213.Powheggg2vvPythia8EvtGen_gg_ZZ_bkg_4l_noTau_13TeV	ggZZ
mc15_13TeV.342561.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_ttH125_4l	ttH

Table V: Monte Carlo samples used to model the expected kinematic distributions for the irreducible background contributions.

It is noted that bbH process may be included as either signal or background contribution. However, the number of bbH event is essentially zero so it has no effect on the final result.

Standard Model Signals

Sample Name	Process
mc15_13TeV.344158.aMcAtNloPythia8EvtGen_A14NNPDF23LO_ppx0_FxFx_Np012_SM	ggF
mc15_13TeV.343247.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0p	VBF+VH Had
mc15_13TeV.344135.MGPy8EG_A14NNPDF23LO_vhlep125_4l_0p	VH Lep

Table VI: Monte Carlo samples used to model the expected kinematic distributions for the Standard Model signal contributions.

BSM Signals

Sample Name (ggH vectex)	$\cos\alpha$	κ_{Hgg}	κ_{Agg}
mc15_13TeV.344158.aMcAtNloPythia8EvtGen_A14NNPDF23LO_ppx0_FxFx_Np012_SM	1	1	0
mc15_13TeV.344167.aMcAtNloPythia8EvtGen_A14NNPDF23LO_FxFx_kAgg0p5_cosa0p707	$\frac{1}{\sqrt{2}}$	1	0.5
mc15_13TeV.344168.aMcAtNloPythia8EvtGen_A14NNPDF23LO_FxFx_kAggn0p5_cosa0p707	$\frac{1}{\sqrt{2}}$	1	-0.5
Sample Name (HVV vertex, scalar Higgs)	$\cos\alpha$	κ_{SM}	$\kappa_{H_{VV}}$
mc15_13TeV.343247.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0p	1	1	0
mc15_13TeV.343249.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0phkHv10	1	0	10
mc15_13TeV.343260.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphkHv2p5	1	1	2.5
mc15_13TeV.343259.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphkHv5	1	1	5
mc15_13TeV.343261.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphkHvm5	1	1	-5
Sample Name (HVV vertex, pseudoscalar Higgs)	$\cos\alpha$	κ_{SM}	$\kappa_{A_{VV}}$
mc15_13TeV.343247.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0p	$\frac{1}{\sqrt{2}}$	1	0
mc15_13TeV.343252.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0mkAv15	$\frac{1}{\sqrt{2}}$	0	15
mc15_13TeV.343254.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphmkAv2p5	$\frac{1}{\sqrt{2}}$	1	2.5
mc15_13TeV.343253.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphmkAv5	$\frac{1}{\sqrt{2}}$	1	5
mc15_13TeV.343255.MadGraphPythia8EvtGen_A14NNPDF23LO_vbfhzz4l_0pphmkAvm5	$\frac{1}{\sqrt{2}}$	1	-5

Table VII: Monte Carlo samples used to model the expected kinematic distributions for the BSM signal contributions in the ggH and HVV vertex.

C. Observables

Signal and background distributions are constructed from different observables representing the kinematics of an event. These observables may include the reconstructed mass, momentum and angles of the detected particles of an event. The transverse momentum (momentum perpendicular to the beam line), denoted as p_T, is usually used so that there is no contribution from the left over momentum of the beam particles. In the $H \to ZZ^* \to 4l$ decay channel, the kinematics can be fully characterized by three invariant masses m_{ZZ^*}, m_{12}, and m_{34}, and six angles **??**. Five of these angles are illustrated in Fig. 24. The list of observables that are used in this analysis is summarized in Table VIII.

Observable	Description
m_{ZZ^*}	Mass of the on shell and off shell (Z and Z^*) Z boson
m_{12} / m_{Z1}	Mass of the lepton pair from the decay of the on shell Z boson
m_{34} / m_{Z2}	Mass of the lepton pair from the decay of the off shell Z^* boson
$p_T 4l_{\mathrm{fsr}}$	Transverse momentum of four leptons with final state photons correction
jet $p_T[\mathrm{i}]$	Transverse momentum of the $\mathrm{i^{th}}$ jets
ϕ, ϕ_1	Azimuthal angle ϕ (ϕ_1) between the beam axis and the plane formed by the lepton pair from the decay of the on (off) shell Z (Z^*) boson
θ, θ_1	Helicity angle θ (θ_1) of the on (off) shell Z (Z^*)boson
jet ϕ, ϕ_1	Azimuthal angle ϕ (ϕ_1)of the leading (sub-leading) jet for processes with two jets final state
jet $\eta[\mathrm{i}]$	pseudorapidity $\eta = -\ln[\tan(\theta/2)]$ of the $\mathrm{i^{th}}$ jet

Table VIII: Description of the kinematic observables used in this analysis

39
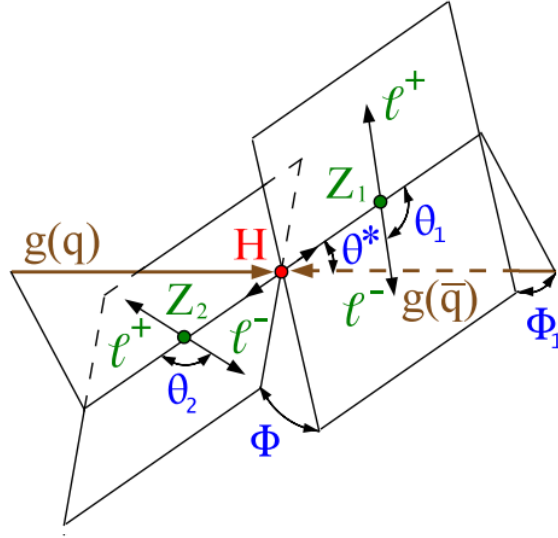
Figure 24: Illustration of the angular kinematics of an exotic X particle production and its decay in pp collision or $gg \to X \to ZZ^* \to 4l^{\pm}$.

D. Data Driven Background Estimation

The data driven background is used to model reducible background processes which contain fake and non-isolated leptons. It is implemented as category-wise distributions with their total yields fixed and their shape based on the total reducible background distributions. A summary of the data driven background events per category is shown in Table IX.

Category	Yields
0-jet	2.2373
1-jet, $p_T^{4\ell} < 60\,\mathrm{GeV}$	0.5720
1-jet, $60\,\mathrm{GeV} < p_T^{4\ell} < 60\,\mathrm{GeV}$	0.4200
1-jet, $p_T^{4\ell} > 120\,\mathrm{GeV}$	0.0874
2-jet VH, $p_T^{4\ell} < 150\,\mathrm{GeV}$	0.1634
2-jet VH, $p_T^{4\ell} > 150\,\mathrm{GeV}$	0.0164
2-jet VBF, $p_T^{j_1} < 200\,\mathrm{GeV}$	0.4617
2-jet VBF, $p_T^{j_1} > 200\,\mathrm{GeV}$	0.0846
VH-leptonc	0.0031
$t\bar{t}$H	0.0316

Table IX: Estimates of data driven background yields in each analysis category

E. Best Prediction Factors

To account for higher order QCD corrections in the signal samples generated by the Next-to-Leading Order (NLO) MadGraph5 generator, we introduce a "best prediction" scale factor f_{cat} for each Higgs boson production mode for each category. The scale factors for SM and BSM scenarios are assumed to be the same. A scale factor is defined as,

$$f_{\text{cat}} = \frac{\int S_{\text{Powheg,SM,cat}}}{\int S_{\text{MG5,SM,cat}}}. \tag{XIII.6}$$

where $S_{\text{Powheg,SM,cat}}$ and $S_{\text{MG5,SM,cat}}$ are the expected yields in a category obtained from the N^{3}LO generator Powheg (best SM prediction) and from the NLO generator MadGraph5.

A list of all scale factors per category can be found in Table X.

	Best prediction scale factors	
Category	ggF	VBF+VH
0-jet	1.5583	1.3529
1-jet, $p_T^{4\ell} < 60\,\text{GeV}$	1.4301	1.3272
1-jet, $60\,\text{GeV} < p_T^{4\ell} < 60\,\text{GeV}$	1.4592	1.4285
1-jet, $p_T^{4\ell} > 120\,\text{GeV}$	1.2010	1.5359
2-jet VH, $p_T^{4\ell} < 150\,\text{GeV}$	1.4391	1.3264
2-jet VH, $p_T^{4\ell} > 150\,\text{GeV}$	0.9024	1.3664
2-jet VBF, $p_T^{j_1} < 200\,\text{GeV}$	1.2464	1.0494
2-jet VBF, $p_T^{j_1} > 200\,\text{GeV}$	0.9992	1.0112
VH-leptonc	1.1548	1.3137
$t\bar{t}$H	0.9172	1.9469

Table X: List of best prediction scale factors for the effective Lagrangian analysis

F. Likelihood Scan Statistics

	Expected		Observed				
$\kappa_{A_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.30,2.29]	[-3.45,3.48]	[-4.17,-1.80,2.01,4.16]	[-3.45,3.48]	-3.10	-	1.33
κ_{SM} floating	[-2.54,2.55]	[-3.99,4.06]	[-3.10,2.87]	[-3.99,4.06]	-1.60	1.17	0.12
$\kappa_{H_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-1.67,2.07]	[-2.90,3.08]	[2.03,3.73]	[-2.90,3.08]	2.90	-	2.21
κ_{SM} floating	[-1.71,2.68]	[-3.08,3.99]	[0.93,3.17]	[-3.08,3.99]	2.10	1.14	1.21
$\kappa_{A_{gg}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{gg}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-0.28,0.28]	[-0.39,0.40]	[-0.49,-0.27,0.34,0.43]	[-0.39,0.40]	-0.39	-	1.71
κ_{SM} floating	[-0.68,0.68]	[-0.74,0.74]	[-0.80,-0.33,0.62]	[-0.74,0.74]	-0.71	0.34	0.88

Table XI: Likelihood scan result using one bin distributions of "Set One" discriminants. Systematic uncertainties are not included.

	Expected		Observed				
$\kappa_{A_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.28,2.26]	[-3.43,3.45]	[-3.94,-1.39,1.81,4.08]	[-3.43,3.45]	-2.80	-	1.05
κ_{SM} floating	[-2.51,2.51]	[-3.94,4.01]	[-2.88,2.50]	[-3.94,4.01]	-1.40	1.17	0.17
$\kappa_{H_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-1.61,2.02]	[-2.82,3.05]	[1.88,3.61]	[-2.82,3.05]	2.80	-	2.02
κ_{SM} floating	[-1.64,2.63]	[-2.96,3.95]	[0.81,3.08]	[-2.96,3.95]	2.00	1.14	1.08
$\kappa_{A_{gg}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{gg}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-0.27,0.28]	[-0.39,0.40]	[-0.48,-0.26,0.29,0.46]	[-0.39,0.40]	-0.39	-	1.64
κ_{SM} floating	[-0.67,0.67]	[-0.74,0.74]	[-0.79,-0.14,0.37,0.79]	[-0.74,0.74]	-0.71	0.35	0.57

Table XII: Likelihood scan result using two bin distributions of "Set One" discriminants. Systematic uncertainties are not included.

$\kappa_{A_{VV}}$	Expected 68% CL	95%CL	Observed 68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.24,2.25]	[-3.39,3.43]	[-3.84,-1.20,1.96,3.75]	[-3.39,3.43]	-2.70	-	0.96
κ_{SM} floating	[-2.45,2.48]	[-3.88,3.96]	[-2.91,2.16]	[-3.88,3.96]	-1.40	1.15	0.11

$\kappa_{H_{VV}}$	Expected 68% CL	95%CL	Observed 68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-1.58,2.00]	[-2.78,3.03]	[1.76,3.55]	[-2.78,3.03]	2.70	-	1.89
κ_{SM} floating	[-1.60,2.59]	[-2.91,3.90]	[0.75,3.10]	[-2.91,3.90]	2.00	1.12	1.04

Table XIII: Likelihood scan result using four bin distributions of "Set One" discriminants. Systematic uncertainties are not included.

$\kappa_{A_{VV}}$	Expected 68% CL	95%CL	Observed 68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.72,2.74]	[-4.13,4.16]	[-5.02,-2.49,2.58,5.15]	[-4.13,4.16]	3.90	-	1.75
κ_{SM} floating	[-3.01,3.07]	[-4.88,4.96]	[-5.35,-2.07,2.26,5.64]	[-4.88,4.96]	3.90	1.00	1.37

$\kappa_{H_{VV}}$	Expected 68% CL	95%CL	Observed 68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.21,2.27]	[-3.72,3.46]	[2.24,4.16]	[-3.72,3.46]	3.30	-	2.03
κ_{SM} floating	[-2.26,2.98]	[-3.99,4.77]	[2.18,4.98]	[-3.99,4.77]	3.60	0.94	1.73

$\kappa_{A_{gg}}$	Expected 68% CL	95%CL	Observed 68%CL	95%CL	$\kappa_{A_{gg}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-0.30,0.30]	[-0.43,0.43]	[-0.44,-0.12]	[-0.43,0.43]	-0.31	-	0.78
κ_{SM} floating	[-0.69,0.69]	[-0.76,0.76]	[-0.75,-0.18,0.63,0.75]	[-0.76,0.76]	-0.61	0.53	0.73

Table XIV: Likelihood scan result using one bin distributions of "Set Two" discriminants. Systematic uncertainties are not included.

	Expected		Observed				
$\kappa_{A_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.67,2.69]	[-4.07,4.11]	[-4.97,-2.48,2.68,5.46]	[-4.07,4.11]	4.10	-	1.79
κ_{SM} floating	[-2.92,2.99]	[-4.73,4.81]	[-5.26,-2.13,2.51]	[-4.73,4.81]	4.50	0.96	1.46
$\kappa_{H_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.17,2.22]	[-3.67,3.42]	[2.24,4.20]	[-3.67,3.42]	3.30	-	1.87
κ_{SM} floating	[-2.22,2.87]	[-3.90,4.63]	[2.32,5.30]	[-3.90,4.63]	3.80	0.91	1.63

Table XV: Likelihood scan result using four bin distributions of "Set Two" discriminants. Systematic uncertainties are not included.

	Expected		Observed				
$\kappa_{A_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-2.56,2.61]	[-3.89,3.99]	[-4.96,-2.49,2.66,5.43]	[-3.89,3.99]	4.10	-	1.83
κ_{SM} floating	[-2.84,2.94]	[-4.60,4.84]	[-4.51,-1.25,1.79,5.05]	[-4.60,4.84]	-3.00	1.11	0.94
$\kappa_{H_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-1.96,2.28]	[-3.47,3.38]	[2.40,4.24]	[-3.47,3.38]	3.40	-	2.24
κ_{SM} floating	[-1.99,3.05]	[-3.69,4.63]	[1.69,4.45]	[-3.69,4.63]	3.10	1.05	1.39
$\kappa_{A_{gg}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{gg}}$	κ_{SM}	$\sigma_{\text{SM, excl}}$
κ_{SM} fixed	[-0.30,0.31]	[-0.44,0.44]	[-0.49,-0.20,0.21,0.49]	[-0.44,0.44]	-0.37	-	1.07
κ_{SM} floating	[-0.69,0.69]	[-0.76,0.76]	-	[-0.76,0.76]	0.62	0.62	0.24

Table XVI: Likelihood scan result using one bin distributions of "Set Three" discriminants. Systematic uncertainties are not included.

	Expected		Observed				
$\kappa_{A_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{VV}}$	κ_{SM}	$\sigma_{\mathrm{SM,\ excl}}$
κ_{SM} fixed	[5.67]	-	[-4.34,4.29]	-	0.20	-	0.00
κ_{SM} floating	[5.67]	-	[-4.34,4.29]	-	0.20	1.00	0.00
$\kappa_{H_{VV}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{H_{VV}}$	κ_{SM}	$\sigma_{\mathrm{SM,\ excl}}$
κ_{SM} fixed	[-4.19,4.23]	-	[-2.12,3.74]	-	1.30	-	0.00
κ_{SM} floating	[-4.19,4.23]	-	[-2.12,3.74]	-	1.30	1.00	0.00
$\kappa_{A_{gg}}$	68% CL	95%CL	68%CL	95%CL	$\kappa_{A_{gg}}$	κ_{SM}	$\sigma_{\mathrm{SM,\ excl}}$
κ_{SM} fixed	[-0.43,0.41]	[-0.62,0.60]	[-0.27,0.29]	[-0.62,0.60]	0.02	-	0.00
κ_{SM} floating	[-0.43,0.41]	[-0.62,0.60]	[-0.27,0.29]	[-0.62,0.60]	0.02	1.00	0.00

Table XVII: Likelihood scan result using one bin distributions of "Set Four" discriminants. Systematic uncertainties are not included.

REFERENCES

[1] 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.*, D96(7):072002, 2017. doi:10.1103/PhysRevD.96.072002.

[2] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.

[3] Georges Aad et al. Evidence for the spin-0 nature of the Higgs boson using ATLAS data. *Phys. Lett.*, B726:120–144, 2013. doi:10.1016/j.physletb.2013.08.026.

[4] P. Artoisenet et al. A framework for Higgs characterisation. *JHEP*, 11:043, 2013. doi:10.1007/JHEP11(2013)043.

[5] S. Boselli, C. M. Carloni Calame, G. Montagna, O. Nicrosini, F. Piccinini, and A. Shivaji. Higgs decay into four charged leptons in the presence of dimension-six operators. 2017.

[6] S. Y. Choi, D. J. Miller, M. M. Muhlleitner, and P. M. Zerwas. Identifying the Higgs spin and parity in decays to Z pairs. *Phys. Lett.*, B553:61–71, 2003. doi:10.1016/S0370-2693(02)03191-X.

[7] Baradhwaj Coleppa, Kunal Kumar, and Heather E. Logan. Can the 126 GeV boson be a pseudoscalar? *Phys. Rev.*, D86:075022, 2012. doi:10.1103/PhysRevD.86.075022.

[8] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016. doi:10.23731/CYRM-2017-002.

[9] Adam Falkowski. Effective field theory approach to LHC Higgs data. *Pramana*, 87(3):39, 2016. doi:10.1007/s12043-016-1251-5.

[10] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Phys. Rev.*, D81:075022, 2010. doi:10.1103/PhysRevD.81.075022.

[11] Rohini M. Godbole, D. J. Miller, and M. Margarete Muhlleitner. Aspects of CP violation in the H ZZ coupling at the LHC. *JHEP*, 12:031, 2007. doi:10.1088/1126-6708/2007/12/031.

[12] ATLAS Group. Moriond 2017: Couplings and simplified cross section studies in the $h \to zz^* \to l^+l^-l^+l^-$ channel. `https://cds.cern.ch/record/2231596`, .

[13] ATLAS Group. Event selection and background estimation in the $h \to zz^* \to 4l$ channel at $\sqrt{s} = 13$ tev - moriond 2017 analysis. `https://cds.cern.ch/record/2231576`, .

[14] Kaoru Hagiwara, Qiang Li, and Kentarou Mawatari. Jet angular correlation in vector-boson fusion processes at hadron colliders. *JHEP*, 07:101, 2009. doi:10.1088/1126-6708/2009/07/101.

[15] V. Hankele, G. Klamke, and D. Zeppenfeld. Higgs + 2 jets as a probe for CP properties. In *Meeting on CP Violation and Non-standard Higgs Physics Geneva, Switzerland, December 2-3, 2004*, pages 58–62, 2006.

[16] V. Hankele, G. Klamke, D. Zeppenfeld, and T. Figy. Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC. *Phys. Rev.*, D74:095001, 2006. doi:10.1103/PhysRevD.74.095001.

[17] Nikolas Kauer. Inadequacy of zero-width approximation for a light Higgs boson signal. *Mod. Phys. Lett.*, A28:1330015, 2013. doi:10.1142/S0217732313300152.

[18] G. Klamke and D. Zeppenfeld. Higgs plus two jet production via gluon fusion as a signal at the CERN LHC. *JHEP*, 04:052, 2007. doi:10.1088/1126-6708/2007/04/052.

[19] L. D. Landau. On the angular momentum of a system of two photons. *Dokl. Akad. Nauk Ser. Fiz.*, 60(2):207–209, 1948. doi:10.1016/B978-0-08-010586-4.50070-5.

[20] Chen-Ning Yang. Selection Rules for the Dematerialization of a Particle Into Two Photons. *Phys. Rev.*, 77:242–245, 1950. doi:10.1103/PhysRev.77.242.