



## ATLAS Note

GROUP-2017-XX

26th May 2021



Draft version 0.1

1

2

3

4

# Study of CP properties of vector-boson-fusion produced Higgs boson in di-photon decay channel with ATLAS detector at the LHC

5

The ATLAS Collaboration

6

This note performs a study of VBF produced Higgs boson CP properties in diphoton decay channel. The analysis uses  $139.8 \text{ fb}^{-1}$  p-p collision data collected by ATLAS detector during 2015-2018 LHC RunII period at  $\sqrt{s} = 13\text{TeV}$ . A six dimension effective field theory model is considered to describe the contribution from CP-violating interaction between Higgs boson and electroweak vector boson, in which the CP violation magnitude can be expressed by one parameter  $\tilde{d}$ . By performing a maximum likelihood fit, the  $\tilde{d}$  is expected to be constrained in  $[-0.027, 0.027]$  interval at 68% confidence level(CL) and  $[-0.059, 0.060]$  interval at 95% CL based on Standard Model prediction.

14

© 2021 CERN for the benefit of the ATLAS Collaboration.

15 Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

---

## 16 **Contents**

17	<b>1 Introduction</b>	4
18	<b>2 Theoretical framework</b>	5
19	<b>3 ATLAS Detector</b>	6
20	<b>4 Data and Monte Carlo Samples</b>	8
21	<b>5 Event selection</b>	10
22	5.1 Object reconstruction	10
23	5.2 Event Selection	13
24	<b>6 Event Categorization</b>	15
25	6.1 New MVA based categorization	15
26	6.2 Optimal observable binning	17
27	<b>7 Signal modelling</b>	19
28	<b>8 Background Modelling</b>	22
29	8.1 Background modeling strategy	22
30	8.2 Background compositions	22
31	8.3 Background templates	24
32	8.4 Spurious signal test	32
33	8.5 Background modelling results	35
34	<b>9 Systematic uncertainties</b>	39
35	9.1 Theory uncertainties	39
36	9.2 Uncertainties from signal and background modeling	40
37	9.3 Experimental systematic uncertainties affecting the expected signal yields	40
38	<b>10 Fitting strategy</b>	42
39	<b>11 Results</b>	43
40	<b>12 Conclusion</b>	45
41	<b>Appendices</b>	48
42	<b>A Statistic only results including signal yields</b>	48
43	<b>B Performance of new BDT based VBF categories</b>	48
44	<b>C Binning optimization</b>	48
45	<b>D Background decomposition with 2x2D sideband method</b>	48

46	<b>E Full results of spurious signal test</b>	<b>48</b>
47	<b>F Statistics results based on 2D model</b>	<b>48</b>

## 48 1 Introduction

49 The discovery of Higgs boson by ATLAS and CMS experiment in 2012 opens a new era for particle  
50 physics, and provides a new opportunity to search for new physics beyond Standard Model, such as the  
51 new source of CP-violation required by the observed baryon asymmetry in the universe. In the Standard  
52 Model the expected Higgs boson is a scalar boson with  $J^{PC} = 0^{++}$ , till now all the experimental results  
53 in ATLAS and CMS are consistent with this prediction within the uncertainty. However, limited by the  
54 statistics, the direct Higgs CP measurements are performed only in very recent period [1][2][3][4]. All of  
55 these observations show no derivation from SM, but the restriction is not so satisfying.

56 The focus of this work is the CP property of Higgs boson in its interaction with gauge vector bosons.  
57 The test is performed through the largest HVV production mode in LHC, Vector Boson Fusion, in diphoton  
58 decay channel, with  $140\text{ fb}^{-1}$  p-p collision data collected by ATLAS detector at  $\sqrt{s} = 13\text{ TeV}$  during  
59 2015 to 2018. In order to have a more sensitive result and be independent with CP-even observables, the  
60 Optimal Observable is introduced as in previous di-tau channel analysis.

61 The Effective Field Theory and Optimal Observable are introduced in Sec. 2. Sec. 3 and Sec. 4 show a  
62 brief description about ATLAS detector and the simulated data we used in this analysis. In Sec. 5 Sec. 7 and  
63 Sec. 8, we presented the main analyses procedure, including the event selection, categorization and event  
64 modelling. Following in Sec. 9 is the experimental uncertainty considered, and statistic model constructed  
65 for parameters estimation is in Sec. 10. The results and conclusions are discussed in Sec. 11 and Sec. 12.

## 66 2 Theoretical framework

67 In order to be convenient for the future combination in different Higgs decay channel, the basic theory  
 68 considered in this analysis keeps consistency with previous Higgs CP test in di-tau channel. The effective  
 69 Lagrangian is the SM Lagrangian augmented with CP-odd operator of mass dimension 6, involving  
 70 the Higgs field and electroweak gauge field. The other interaction between Higgs/vector boson and  
 71 fermions/gluons is assumed to be the same with SM prediction. The Lagrangian could be written as:

$$72 \quad \mathcal{L}_{eff} = \mathcal{L}_{SM} + \tilde{g}_{HAA} H \tilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{HAZ} H \tilde{A}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HZZ} H \tilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} H \tilde{W}_{\mu\nu}^+ W^{-\mu\nu}$$

73 Where  $V^{\mu\nu}$  and  $\tilde{V}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma}$  (with  $V = W^\pm, Z, A$ ) denotes the vector field strength. Among the 4  
 74 coupling constants  $g_{HVV}$  only 2 of them are independent due to the  $U(1)^*SU(2)$  constrain, they can  
 75 be written into 2 dimensionless coupling  $\tilde{d}$  and  $\tilde{d}_B$ . As the different contribution from various vector  
 76 electroweak gauge boson fusion process could not be distinguished experimentally, an arbitrary choice  $\tilde{d} =$   
 77  $\tilde{d}_B$  is adopt. So the following relation could be derived:

$$78 \quad \tilde{g}_{HAA} = \tilde{g}_{HZZ} = \frac{1}{2}\tilde{g}_{HWW} = \frac{g}{2m_W}\tilde{d} \quad and \quad \tilde{g}_{HAZ} = 0$$

79 The parameter  $\tilde{d}$  becomes the only one describing the CP violation in VBF Higgs production. The  
 80 corresponding matrix element could be written as the sum of a CP-even SM contribution  $\mathcal{M}_{SM}$  and a  
 81 CP-odd contribution  $\mathcal{M}_{CP-odd}$  from EFT, and  $\tilde{d}$  could be factorized out:

$$82 \quad \mathcal{M} = \mathcal{M}_{SM} + \tilde{d} \cdot \mathcal{M}_{CP-odd}$$

83 The VBF cross section is proportional to the squared matrix element:

$$84 \quad |\mathcal{M}|^2 = |\mathcal{M}_{SM}|^2 + \tilde{d} \cdot 2Re(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd}) + \tilde{d}^2 \cdot |\mathcal{M}_{CP-odd}|^2$$

85 The first term and last term are both CP-even, only the second term from the interference of 2 matrix  
 86 element is CP-odd, representing a new source of CPV in Higgs field. This CP-odd term would vanish  
 87 after integrated after a CP-symmetric phase space, so it does not contribute to the total cross section and  
 88 observed event yields after applying CP-symmetric selection criteria. The third term could increase the  
 89 total cross section by an amount of  $\tilde{d}^2$ .

90 The VBF Higgs decay final state contains the reconstructed Higgs boson and two tagging jets from  
 91 VBF topology. This can be characterized in a seven-dimension phase space, when fixing the Higgs mass,  
 92 neglecting jet masses and exploiting the transverse momentum conservation. The Optimal Observable  
 93 combines the information from total phase space into one single variable, so it shows much better  
 94 performance than traditional operators. It is defined as the ratio of the interface term in the matrix element  
 95 to the SM contribution:

$$96 \quad OO = \frac{2Re(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd})}{|\mathcal{M}_{SM}|^2}$$

97     The Optimal Observable is a CP-odd variable, so it can be a probe of CPV in the HVV vertex. In the SM  
 98   the expectation value would vanish( $\langle OO \rangle = 0$ ), so any non-zero mean value or asymmetry in distribution of  
 99   Optimal Observable indicates the physics beyond SM, either stemming from CPV, or originating from  
 100   rescattering effects [Brehmer\_2018]. Plot 1 shows the distribution of Optimal Observable for several  $\tilde{d}$   
 101   VBF signal model, with obvious non-zero mean value and asymmetry shape.

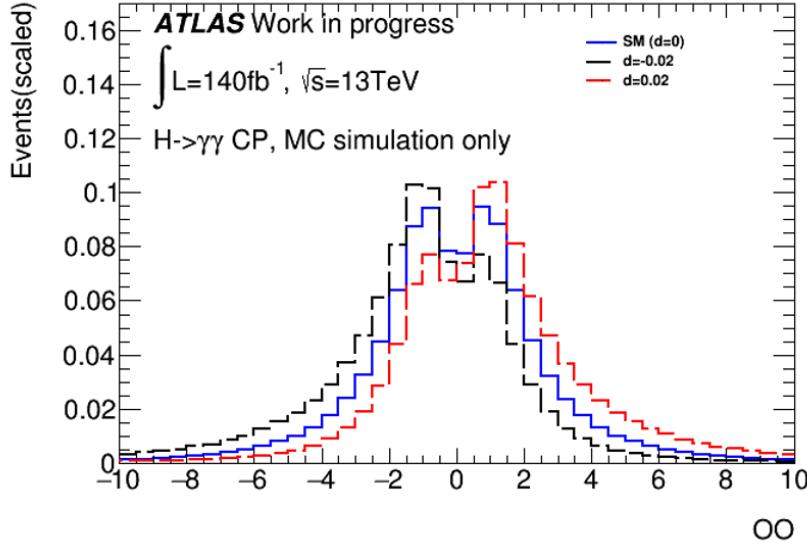


Figure 1: Optimal Observable distribution in VBF process, with  $\tilde{d} = -0.02$ ,  $\tilde{d} = 0$  (SM) and  $\tilde{d} = 0.02$

102   The package HAWK [HAWK\_2015] provides a calculation for VBF and VH process matrix element,  
 103   with the kinematic variables in final state (4-momenta of the Higgs boson and 2 tagging jets in VBF). For  
 104   the initial state, the momentum fraction Bjorken  $x_1(x_2)$  of parton can be derived by energy-momentum  
 105   conservation:

$$x_{1,2}^{reco} = \frac{m_{Hjj}}{\sqrt{s}} e^{\pm y_{Hjj}}$$

107   Where  $m_{Hjj}$  and  $y_{Hjj}$  are the invariant mass and pseudo rapidity of Higgs+2 tagging jets system. The  
 108   matrix element calculation performs a summation of all possible flavor configurations in initial and final  
 109   state  $jj \rightarrow klH$  weighted by CT10 leading order parton distribution functions(PDFs.).

$$\begin{aligned} 110 \quad 2Re(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd}) &= \sum_{i,j,k,l} f_i(x_1) f_j(x_2) 2Re((\mathcal{M}_{SM}^{ij \rightarrow klH})^* \mathcal{M}_{CP-odd}^{ij \rightarrow klH}) \\ 111 \quad |\mathcal{M}_{SM}|^2 &= \sum_{i,j,k,l} f_i(x_1) f_j(x_2) |\mathcal{M}_{SM}^{ij \rightarrow klH}|^2 \end{aligned}$$

### 112   3 ATLAS Detector

113   The ATLAS detector is one of two general detectors at LHC. It is normally forward-backward symmetric  
 114   and covers nearly all the solid angle around proton-proton interaction point. From inner it consists of an

115 inner tracker detector(ID), a superconducting solenoid, electromagnetic and hadronic calorimeter, and  
116 a muon spectrometer. 3 main component of Inner Detector, pixel detector, semiconductor tracker and  
117 transition radiation tracker provide precise measurement of transverse momentum, direction, charge for  
118 the charged particle, as well as a preliminary identification to b-jets in barrel region ( $|\eta| < 2.5$ ). The  
119 whole Inner Detector is immersed in a 2T axial magnetic field provided by the solenoid. The outer part,  
120 electromagnetic calorimeter is a high granularity lead/liquid-argon(LAr) sampling calorimeter, measures  
121 the electromagnetic shower in barrel ( $|\eta| < 1.475$ ) and endcap ( $1.375 < |\eta| < 3.2$ ) regions. Following  
122 hadronic calorimeter can reconstruct the hadronic shower with steel and scintillator tiles ( $|\eta| < 1.7$ ),  
123 copper/LAr ( $1.5 < |\eta| < 2.7$ ) or copper-tungsten/LAr ( $3.1 < |\eta| < 4.9$ ) respectively. The outermost  
124 layer muon spectrometer comprises 3 large superconducting eight-coil toroids, separate trigger chambers  
125 ( $|\eta| < 2.4$ ) and precision tracking chambers ( $|\eta| < 2.7$ ).

126 The ATLAS data-taking system uses a two-level system: a hardware-based level trigger(L1) to reduce  
127 the event rate to at most 100 kHz, and a software-based level trigger(L2) to reduce the event rate to  
128 approximately 1 kHz.

## 129 4 Data and Monte Carlo Samples

130 This analysis uses LHC proton-proton collision data collected by ATLAS detector from 2015 to 2018, at  
 131 central-of-mass energy  $\sqrt{s} = 13\text{TeV}$ . After requiring good data quality the data set amounts to integrated  
 132 luminosity  $139.8 \text{ fb}^{-1}$ . The mean number of interactions per bunch crossing,  $\mu$ , was 34 on average, varying  
 133 from 24 in 2015-2016 to 37 in 2018 data. Events were selected by a di-photon trigger with pT thresholds  
 134 of 35GeV and 25 GeV for leading and sub-leading photon. A tight photon identification and isolation  
 135 requirement are applied. On average the trigger efficiency can reach to 98% after the event selection.

Period	$\int \mathcal{L} dt [\text{fb}^{-1}]$
2015	3.2
2016	32.9
2017	43.8
2018	59.9
total	139.8

Table 1: Integrated luminosity in ATLAS RunII period. The combined luminosity uncertainty is 2.0%

136 The main Standard Model processes considered in this work are gluon-fusion (ggF) Higgs production,  
 137 Vector Boson Fusion (VBF) Higgs production, and continuum diphoton production. For the ggF and  
 138 VBF process, Monte-Carlo (MC) generator provides a precise simulation, the detailed generator, parton  
 139 distribution functions (PDFs) and perturbative order in QCD is summarized in Table 2.

Process	Generator	Showering	PDF sets	QCD order	$\sigma[\text{pb}] \sqrt{s} = 13\text{TeV}$
VBF	Powheg	PYTHIA8	PDF4LHC15	NNLO(QCD)+NLO(EW)	3.75
ggF	Powheg	PYTHIA8	PDF4LHC15	NNLO(QCD)+NLO(EW)	28.3
$\gamma\gamma$	Sherpa	Sherpa	CT10		

Table 2: Monte Carlo samples used in this analysis.

140

141 The complicated final state in di-photon + multi-jets process makes it hard to simulate in MC. The  
 142 Optimal Observable combines the information from 7-dimension phase space, any mismodelling would  
 143 lead to an biased estimation in OO. So sideband data becomes the best candidate for the background, for  
 144 both the modelling and event yield. While restricted by the statistics, di-photon continuum background  
 145 modelling could not be performed by sideband data, so a set of di-photon+jets process MC are generated  
 146 by Sherpa, with proper hadronization and reconstruction.

147 To simulate the CP violation phenomenon in HVV vertex, a matrix element based reweighting method  
 148 is performed in the SM VBF sample. The weight is defined by the square of matrix element value of  
 149 VBF process associated with a specific amount of CP mixing ( $\tilde{d}$ ) to the one obtained from SM. The  
 150 corresponding 2 matrix elements are calculated by HAWK with initial and final state information in each  
 151 event, assuming a  $2 \rightarrow 2 + H$  or  $2 \rightarrow 3 + H$  process. For convenience the weight is parameterized as a  
 152 function of  $\tilde{d}$ , so the CPV process in any mixing level could be simulated immediately.

153  $w = 1 + \tilde{d}w_1 + \tilde{d}^2w_2$

154 For all MC generated samples, a full simulation of ATLAS detector response with Geant4 program  
155 is performed, except the CPV VBF Higgs process is reweighted from reconstructed SM VBF process.  
156 Additional proton-proton interactions (pile-up) are produced using Pythia8 with A2 parameter set and  
157 MSTW2008LO PDF set. They are included in the simulation for all generated events such that the  
158 distribution of the mean number of interactions per bunch crossing reproduces that observed in the data.

---

## 159 5 Event selection

160 In this section, we provide a brief description of the different objects selection and reconstruction in photon  
 161 + jets final state. The main part is similar to previous  $H \rightarrow \gamma\gamma$  coupling work [ATLAS-CONF-2020-026],  
 162 and the relevant changes specifically for this analysis are highlighted.

### 163 5.1 Object reconstruction

#### 164 5.1.1 Photon Reconstruction and pre-selections

165 Photons are reconstructed using dynamic, variable-size energy clusters in the EM calorimeter [5], associated  
 166 with tracks reconstructed in the inner detector in case of conversion vertices are identified. The Loose  
 167 identification of photon-like object bases on shower shape variables defined with cells of middle and  
 168 back portions of the LAr according calorimeter. We pre-select loose photons requiring  $p_T > 25$  GeV and  
 169  $|\eta_{S2}| < 2.37$ , vetoing the transition region between  $1.37 < |\eta_{S2}| < 1.52$ . The selected loose photons are  
 170 calibrated using the latest Run-2 calibration corrections with 69 variations for the scale and 10 for the  
 171 resolution, which detailed in [Andari:2655306]. Additionally an isolation requirement is applied based on  
 172 both calorimeter-based and track-based requirements. The FixedCutLoose working point is used. During  
 173 the event selection, events are required to have at least two loose photons, and the two loose photons with  
 174 the highest pT define the Higgs candidate.

175

176 **NEED VERIFICATION** An NN based algorithm is used to choose the diphoton vertex after the selection  
 177 of two photon candidates. All the object kinematics are recomputed again: all the cuts reported in below  
 178 sections are related to the diphoton vertex.

#### 179 5.1.2 Jet reconstruction and selections

180 Jets are reconstructed using the anti- $k_t$  [Cacciari:2008gp] algorithm with a radius parameter  $R = 0.4$  as  
 181 implemented in the FastJet 3.2.2 [Fastjet, Cacciari:2005hq] software package. A particle flow (PFlow)  
 182 algorithm developed in Ref. [6] is considered to improve the energy and angular resolution, reconstruction  
 183 efficiency, and pile-up stability compared to calorimeter jets. PFlow jets were reconstructed at the derivation  
 184 step to ensure that only tracks coming from the primary vertex defined in the analysis are used. Quality  
 185 criteria are applied to the events in order to reject jets affected by noisy cells in the calorimeter or other bad  
 186 performance detector effects. Events with jets consistent with noise in the calorimeter or non-collision  
 187 background are vetoed.

188 Preselected jets must have  $p_T^{jet} > 25$  GeV and  $|y^{jet}| < 4.4$ . An overlap removal is also applied: jets are  
 189 rejected if they lie within a distance of  $\Delta R < 0.4$  of a selected photon or electron. To suppress pile-up, jets  
 190 a cut on the so-called “jet vertex tag” is applied. The JVT variable is required to be larger than 0.5 for jets  
 191 with  $|\eta| < 2.4$  in the region  $20 < p_T^{jet} < 60$  GeV. A fJVT cut is currently not used in diphoton analyses,  
 192 since fJVT calibration for the PFlow jet collection is still work-in-progress. From EMTopo fJVT studies,  
 193 it has been seen that the scale-factor is significantly deviating from one and also the use of fJVT in the  
 194 analysis brings a moderate improvement.

---

**195 5.1.3 Electron selections**

196 Electrons are reconstructed using dynamic, variable-size energy clusters in the EM calorimeter [5],  
 197 associated with tracks reconstructed in the inner detector. Electron candidates are required to have  
 198  $p_T > 10 \text{ GeV}$  and be in a region defined by  $|\eta| < 2.47$ . Electrons within the EM calorimeter transition  
 199 region  $1.37 < |\eta| < 1.52$  are not considered. In addition, cut for electron track to vertex association are  
 200 applied. Electrons are required to have a transverse impact parameter  $|d_0/\sigma(d_0)| < 5$  and a longitudinal  
 201 impact parameter  $|z_0 \sin(\theta)| < 0.5 \text{ mm}$ . Electron candidates with associated super-clusters affected by dead  
 202 front-end boards in the first or second sampling or by the presence high-voltage trips affecting the three  
 203 samplings or that includes a masked cell during in the core are considered as bad-quality electrons and not  
 204 considered in the analysis.

205 Electrons are selected using a likelihood (LH) based identification. The inputs to the LH include  
 206 measurements from the tracking system, the calorimeter system and quantities that combine both tracks  
 207 and calorimeter information. While the electron reconstruction has been moved to a super-cluster based  
 208 approach, the calorimetric variables describing the shape of the shower have not been changed and still using  
 209 a fixed-cluster approach [[Anastopoulos:2652163](#)]. A Medium LH identification is used in the analysis.

210 The calibration scheme comprises a simulation-based optimization of the energy resolution, corrections  
 211 accounting for differences between data and simulation, the adjustment of the absolute energy scale using  
 212  $Z$  boson decays into electron-positron pairs, and the validation of the energy scale universality using  
 213  $J/\psi$  decays decays into electron-positron pairs and radiative  $Z$  boson decays [[Andari:2655306](#)]. The  
 214 `es2018_R21_v0` energy-scale calibration model provided by the EGamma group is used.

215 The isolation of the electron candidates is achieved by applying requirements on variables using calorimeter  
 216 and tracks information. The calorimetric isolation variable is built summing the transverse energy of  
 217 positive energy topological clusters whose barycentre falls within a cone centered around the electron  
 218 cluster barycentre. These topo-clusters remain at the electromagnetic scale. The energy of the electron  
 219 is removed by subtracting the energy of a cluster build from the EM calorimeter cells contained in a  
 220  $\Delta\eta \times \Delta\phi = 5 \times 7$  around the electron barycenter. A correction to take into account the leakage from the  
 221 electron outside the core is applied. The effects of the pileup and underlying event contribution is subtracted  
 222 by using the jet-area method technique. In the analysis, the FCLoose isolation criteria is used. These  
 223 means that the calorimetric isolation is computed within a cone of radius  $R = 0.2$  and the maximum  
 224 energy allowed inside the cone is  $0.2 \cdot p_T (E_T^{\text{iso}}|_{\Delta R < 0.2} < 0.2 \cdot p_T)$  and that the track isolation requires  
 225  $p_T^{\text{iso}}|_{\Delta R < 0.2} < 0.15 \cdot p_T$ . With this working point, the  $p_T^{\text{iso}}$  quantity is computed summing the transverse  
 226 momentum of tracks (passing Loose track quality,  $p_T > 1$  and Loose vertex association) within a cone of  
 227 variable size centred around the electron track. The size of the cone depends on the  $p_T$  of the electron as:  
 228

$$\Delta R = \min \left( \frac{k_T}{p_T}, 0.2 \right) \quad (1)$$

229 Electrons and photons leave a very similar signature in the detector. Although photons reconstruction  
 230 algorithms are designed to reduce this mis-identification of electrons as photons, in order to keep the  
 231 photon and electron reconstruction efficiencies as high as possible, a strict overlap removal is not enforced.  
 232 In case of doubt, the same electromagnetic clusters, generated by the same particle, is reconstructed under  
 233 both the electron and the photon hypothesis. This implies that genuine electrons may end up reconstructed  
 234 as as photons. If the reconstructed object generated by a real electron passed the analysis offline selection  
 235 or photons, they will make up the so called  $e \rightarrow \gamma$  mis-identification background. The studies performed

236 show that the measured fake rate increase from < 2% in the central barrel to 7% in the endcaps and decrease  
 237 by 1 – 2% from 25 to 300 GeV.

238 **5.1.4 Muon selections**

239 Muon candidates are reconstructed from tracks in the inner detector and the muon spectrometer. Muons  
 240 are required to have  $p_T > 10 \text{ GeV}$  and  $|\eta| < 2.7$ . The transverse impact parameter is required to be  
 241  $|d_0/\sigma(d_0)| < 3$  while the requirement on the longitudinal parameter is  $|z_0 \sin(\theta)| < 0.5 \text{ mm}$ .

242 Muon identification is performed by applying a set of quality requirements based upon the specific features  
 243 of each of the reconstructed muon types. Such requirements aim to reject fake muons coming mainly from  
 244 pion and kaon decays and guarantee robust momentum measurement. Muon candidates originating from  
 245 in-flight decays of charged hadrons are often characterized by the presence of a distinctive “kink” topology  
 246 in the reconstructed track. The direct consequence is a poor compatibility within the uncertainties between  
 247 the measured momentum in the inner detector and the muon spectrometer (MS), and a poor  $\chi^2$  quality of  
 248 the ID-MS combined track fit. The “ $q/p$  significance”,  $\rho'$  and the normalized  $\chi^2$  of a combined track fit  
 249 variables are chosen to define the selection working points [Bugge:2665711]. The identification working  
 250 point used in this analysis is the default one in ATLAS, the Medium working point. This working point  
 251 aims to minimize the systematic uncertainties associated with muon reconstruction and calibration.

252 The muon momentum resolution and scale calibration model used in the analysis is 200202\_Precision\_r21.  
 253 Additionally a charge-dependent Sagitta bias correction is applied.

254 In order to combine the tracking and calorimeter information in the computation of the muon isolation,  
 255 the particle flow algorithm is used. The isolation criteria used in the analysis is PflowLoose\_FixedRad  
 256 which has different criteria depending on the  $p_T$  of the muon. The track variable is computed in a similar  
 257 way as was described for the electrons with some modifications in the track selection. Also the calorimetric  
 258 variable (`newisoflow`) is computed in a similar way but using the neutral energy flow objects instead of the  
 259 raw topological clusters as input:  $(\text{ptvarcone30\_TightTTVA}_\text{pt500} + 0.4 \cdot \text{neflowisol20})/p_T < 0.16$   
 260 for  $p_T < 50 \text{ GeV}$  and  $(\text{ptcone20\_TightTTVA}_\text{pt500} + 0.4 \cdot \text{neflowisol20})/p_T < 0.16$  for  $p_T > 50 \text{ GeV}$ .  
 261 The previous expressions follow the same notation as in [Klein:2665708].

262 **5.1.5 Overlap removal**

263 In order to avoid any possible double counting between objects an overlap removal is applied following  
 264 HGam strategy, that is, starting from the selected photons. This strategy follows the recommendations  
 265 on analysis harmonization detailed in [Adams:1700874]. In this approach, the two leading photons are  
 266 always kept. Electrons and muons in a cone of  $\Delta R = 0.4$  around any of the photons are discarded. Jets  
 267 which are closer than  $\Delta R = 0.2$  ( $\Delta R = 0.4$ ) of an electron (photon) are not considered. Also electrons at  
 268 a distance  $\Delta R < 0.4$  the remaining jets are removed . Additionally, muons with a distance smaller than  
 269  $\Delta R = 0.4$  to a jet are rejected.

---

270 **5.1.6 Cross-checks with previous MxAOD production**

271 The framework used for this analysis is based on a HGAM group common software package. This framework  
 272 uses HIGG1D1 derivation as input and outputs mini-xAOD (MxAOD) with all the calibrated objects and  
 273 user-level variables. The current tag of the framework is h026. Several updates were made with respect to  
 274 the previous tag h025 used in several HGAM analyses published during 2020. Some of the main changes  
 275 made in the code are: **Need detailed description of h026**

- 276   • New Tight photon ID efficiency calibration.  
 277   • Enabled PFlow fJVT.  
 278   • Added single photon triggers (g120\_loose, g140\_loose) in OR with the diphoton triggers.

279 The distribution of several basic kinematic variables and optimal observable were cross-checked against  
 280 h025. The distributions are shown in Figures ?? to ?? . A derivation in OO is observed and is believed to  
 281 be caused by jets.

282 **5.2 Event Selection**

283 **Trigger**

284 **2015+2016** The HLT\_g35\_loose\_g25\_loose diphoton trigger is used which requires at least two recon-  
 285 structed photons with  $E_T$  larger than 35 and 25 GeV passing *loose* identification requirements.

286 **2017+2018** The HLT\_g35\_medium\_g25\_medium\_L12EM20VH diphoton trigger is used which requires at  
 287 least two reconstructed photons with  $E_T$  larger than 35 and 25 GeV passing *medium* identification  
 288 requirements.

289 **Good Run List** Events must belong to the luminosity blocks specified in:

```
290     data15_13TeV.periodAllYear_DetStatus-v89-pro21-02
291     _Unknown_PHYS_StandardGRL_All_Good_25ns.xml
292
293     data16_13TeV.periodAllYear_DetStatus-v89-pro21-01
294     _DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml
295
296     data17_13TeV.periodAllYear_DetStatus-v99-pro22-01
297     _Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml
298
299     data18_13TeV.periodAllYear_DetStatus-v102-pro22-04
300     _Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml
```

301 for the data collected from 2015 to 2018, which require that all the subdetectors relevant for this  
 302 analysis are fully operative.

303 **Event Quality** Standard event cleaning is applied on data/MC based on the detector quality flags. This  
 304 means that events with data integrity errors in the calorimeters and incomplete events where some  
 305 detector information is missing are rejected.

306 **Primary vertex** At least one primary vertex is required to be reconstructed in the event. No additional  
307 requirements are applied to this vertex, other than the default reconstruction requirements.

308 **Trigger matching** The two offline photons are required to match the photons identified by the HLT.

309  **$\gamma\gamma$  event pre-selections** The  $\gamma\gamma$  event pre-selections require at least two preselected photon as described  
310 above. These photon candidates must satisfy the electron-photon ambiguity resolution and are used  
311 to reconstruct the diphoton primary vertex.

312  **$\gamma\gamma$  selection** After the events pre-selection, the two highest  $p_T$  loose photon candidates are further required  
313 to pass the Tight identification and to satisfy the FixedCutLoose isolation criteria, based on both  
314 the calorimetric and track isolation variables, built in cones of  $\Delta R = 0.2$ :

$$E_T^{\text{iso}}|_{\Delta R < 0.2} < 0.065 \cdot E_T \quad \text{and} \quad p_T^{\text{iso}}|_{\Delta R < 0.2} < 0.05 \cdot E_T;$$

315 Only tracks associated to the diphoton primary vertex are considered in the isolation calculation. If  
316 both the photons pass these requirements the event is finally selected, while events with less than two  
317 Tight and isolated photons are rejected.

318 **Photon relative  $p_T$  cut** The leading and subleading photons are required to have  $p_T/m_{\gamma\gamma}$  larger than 0.35  
319 and 0.25, respectively.

320  **$m_{\gamma\gamma}$  mass window requirement** The diphoton invariant mass must be in the range  $105 < m_{\gamma\gamma} < 160 \text{ GeV}$ .

### 321 5.2.1 VBF enriched category

322 The per-selection for VBF process is inherited from old STXS 1.0 framework in  $80 \text{ fb}^{-1}$  analysis, which  
323 is also based on the kinematic characters of the reconstructed particles [7]. A  $qq \rightarrow Hqq$  final state is  
324 required, with the rapidity of Higgs should be within the barrel region  $|\eta_H| < 2.5$ . The upper limit for  
325 jet energy is set to  $200 \text{ GeV}$ . Di-jet system should have a large invariant mass  $m_{jj} > 400 \text{ GeV}$  and large  
326 pseudo-rapidity separation  $|\Delta\eta_{jj}| > 2.0$ . A Zeppenfeld  $\eta$  cut is used to focus on VBF process  $|\eta^{ZepP}| < 5$ .  
327 An additional BDT based categorization specifically for this analysis would be described in Section 6.

### 328 5.2.2 Data and MC cut-flow of diphoton selection and expected signal yields

	VBF		ggF		Sideband data	
	yield	selection eff.	yield	selection eff.	yield	selection eff.
initial	1205.27		15475.44		15329817	
2 tight photons	435.7	36.15%	5470.36	35.35%	808643	5.27%
$N_j \leq 2$	229.389	19.03%	710.967	4.59%	137731	0.90%
$m_{jj} > 400 \text{ GeV}$	149.704	12.42%	117.629	0.76%	18647	0.12%
$ \Delta\eta  > 2$	148.812	12.35%	111.131	0.72%	15279	0.10%
$ \eta^{ZepP}  < 5$	148.747	12.34%	110.765	0.72%	15159	0.10%

Table 3: Cutflow of the expected VBF and ggF yields and sideband data.

## 329 6 Event Categorization

### 330 6.1 New MVA based categorization

331 The VBF-enriched categories in Higgs coupling analysis focus more on VBF significance and the topology  
 332 of final state to match the calculation in STXS framework, they are not suitable for this precise measurement  
 333 of Higgs CP. For this reason a new 2 dimension boosted decision tree(BDT) is trained to have a high  
 334 VBF purity signal region and reject both gluon-fusion and continuum background events. Both of these  
 335 two BDTs use 7 variables as input:  $pT_{Hjj}$ ,  $pT_t$ ,  $m_{jj}$ ,  $\Delta\eta_{jj}$ ,  $\delta\Phi_{\gamma\gamma,jj}$ ,  $\eta^{ZepP}$ ,  $\Delta R_{\gamma,j}^{min}$ . They are trained  
 336 independently with total templates. Figure 2 3 show the trained BDT distribution and ROC curve, more  
 337 details can be found in Appendix B

338 The  $BDT_{VBF/ggF}$  uses VBF purity  $p = \frac{N_{VBF}}{N_{VBF} + N_{ggF}}$  as criterion. Since the maximum value of VBF purity  
 339 appears at the edge of BDT range  $BDT_{VBF/ggF} = 1$ , the working point is manually chosen to have the  
 340 same VBF efficiency 34% as in Higgs coupling categorization. The VBF purity is increased from 87.7%  
 341 to 91.3%, and gluon-fusion efficiency is decreased from 6.5% to 4.5%.

342

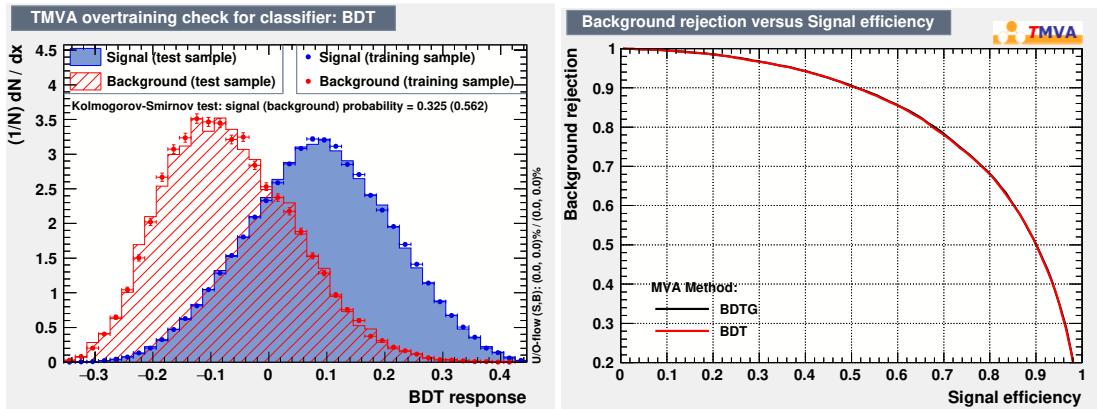


Figure 2: BDT distribution(left) and ROC curve(right) in VBF-ggF BDT training.

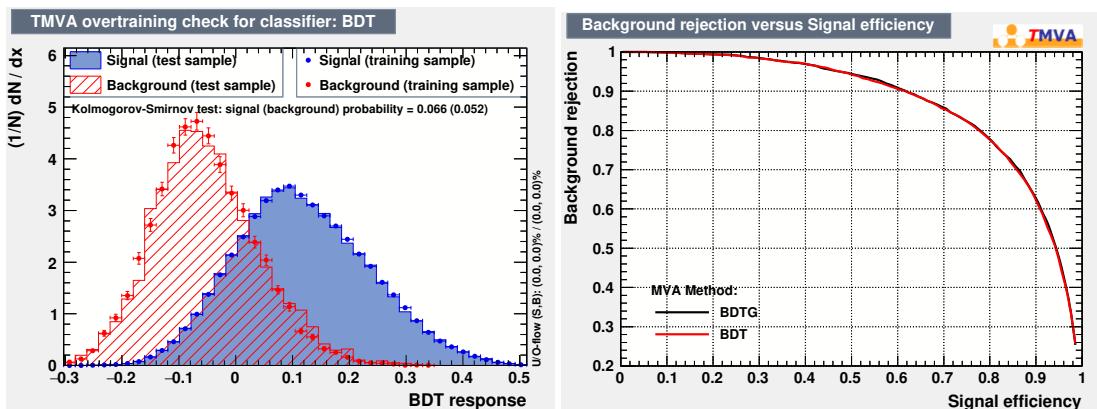


Figure 3: BDT distribution(left) and ROC curve(right) in VBF-MC $\gamma\gamma$  BDT training.

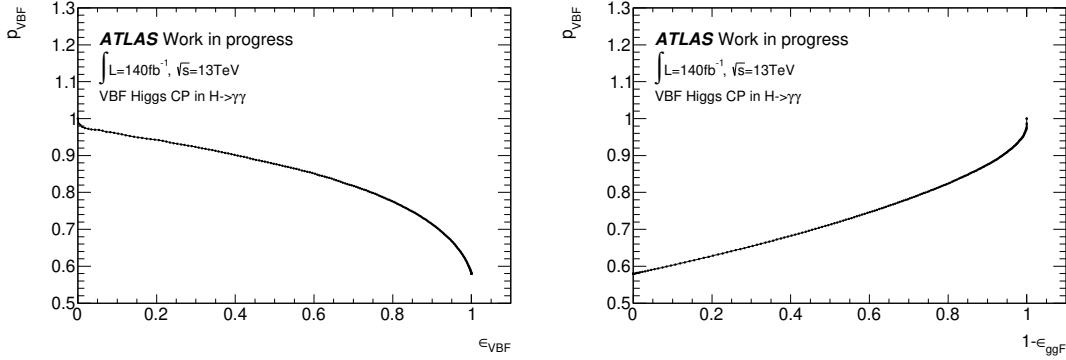


Figure 4: VBF purity with relationship of VBF efficiency(left) and ggF rejection(right). The working point is manually chosen with  $\epsilon_{VBF} = 34\%$ , corresponding VBF purity and ggF rejection could be gotten.

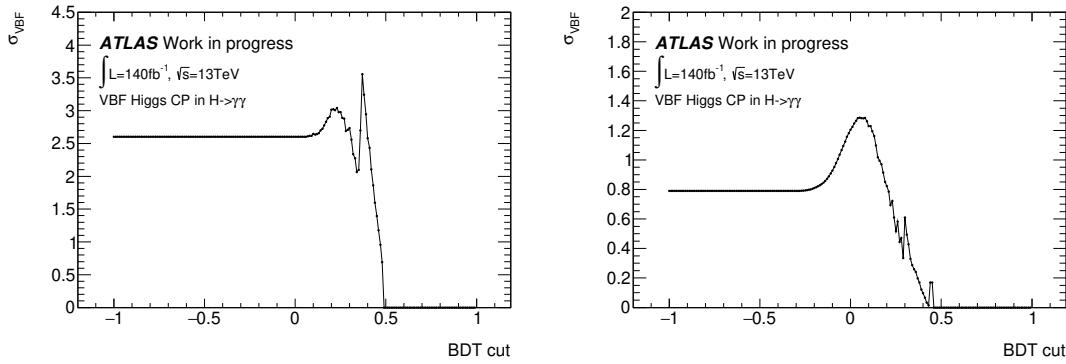
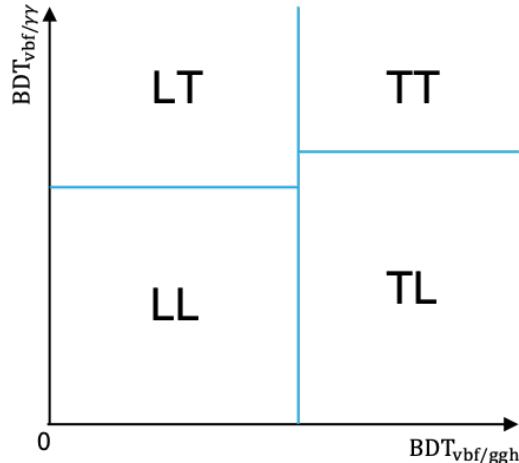


Figure 5: VBF significance in ggH\_tight(left) and ggH\_loose(right) regions with BDT cut value. In ggH\_tight the high fluctuation point has been ignored.

343 Two regions are defined by the former criterion, naming ggH\_tight and ggH\_loose temporarily. Scanning  $BDT_{VBF/\gamma\gamma}$  value in these two regions individually can provide a cut value at the global maximum  
 344 VBF significance point. Figure 6 summary the definition of totally 4 categories, naming TT, TL, LT and LL.  
 345  
 346

	TT	TL	LT	LL	sum
VBF	26.30	25.88	52.71	43.85	148.75
ggF	1.72	3.22	24.06	81.77	110.76
continuum	106	310	1886	16646	18949
[120, 130]GeV	21.25	62	377.25	3329.25	3789.7
significance	5.49	3.20	2.63	0.75	
combined			6.92		

Table 4: Expectend event yields and VBF significance in 4 categories. Continuum background yield is normalized to sideband data, [120, 130]GeV is estimated with background event number in full mass region times 0.2. VBF significance is calculated with background in mass window only.



(a)

Category	Description
TT	$BDT_{VBF/ggF} > 0.14, BDT_{VBF/\gamma\gamma} > 0.23$
TL	$BDT_{VBF/ggF} > 0.14, BDT_{VBF/\gamma\gamma} < 0.23$
LT	$BDT_{VBF/ggF} < 0.14, BDT_{VBF/\gamma\gamma} > 0.05$
LL	$BDT_{VBF/ggF} < 0.14, BDT_{VBF/\gamma\gamma} < 0.05$

(b)

Figure 6: The definition of 4 categories used in this analysis.

## 347 6.2 Optimal observable binning

348 The choice of optimal observable binning method can influence the sensitive of CP measurement. All events  
 349 are divided into 6 positive-negative symmetry bins to ensure enough statistics and match the symmetry  
 350 distribution of optimal observable. With this restriction the binning can be decided by scanning only 2  
 351 parameter  $p_1$  and  $p_2$  with step 0.5 to have maximum global VBF significance:

$$Z_{\tilde{d}} = \sqrt{(\sum_{i=0}^6 Z_i^2)} Z_i = Nvbf / \sqrt{Nvbf + Nggh + Nsr}$$

352 With  $Nsr$  is the extrapolated background event number from sideband data. For different CP-mixing model  
 353 the  $Z$  would be different, so in order to be model independent, an average  $Z$  value from  $d=-0.1$  to  $0.1$  is  
 354 chosen as the final criteria. Figure 7 shows the  $Z$  value for  $p_1$  and  $p_2$  scanning. The event number in every  
 355 BDT categories and bins are shown in Table ??

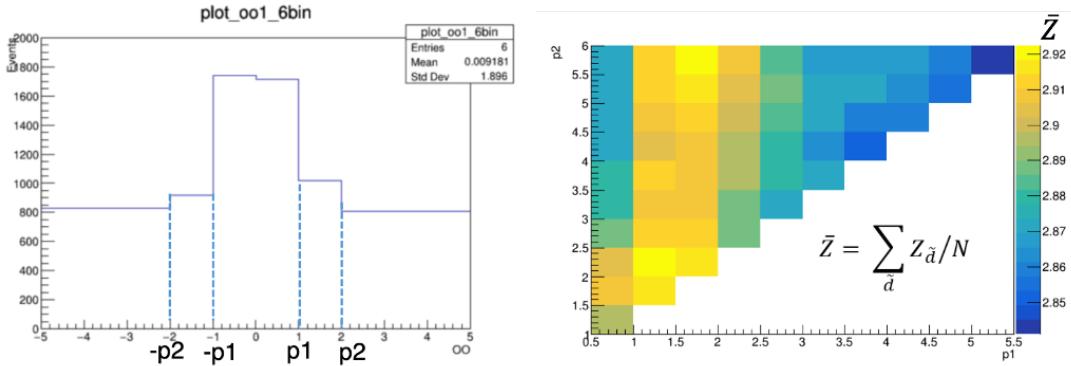


Figure 7: The optimal observable binning is determined with 2 parameters. Global VBF significance is calculated in scanning two parameters with step of 0.5. ( $p_1, p_2$ ) equal to (1, 2) and (1.5, 5, 5) show very similar significance, while in order to get enough statistics in edge bins, the former one is chosen.

VBF	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	sum
TT	6.29	3.63	3.22	3.20	3.63	6.32	26.30
TL	5.05	4.04	3.82	3.84	4.08	5.05	25.88
LT	8.96	7.72	9.78	9.65	7.70	8.90	52.71
LL	6.65	6.12	9.18	9.17	6.14	6.60	43.85

(a)

ggF	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	sum
TT	0.28	0.23	0.33	0.32	0.26	0.29	1.72
TL	0.52	0.43	0.65	0.69	0.42	0.50	3.22
LT	3.95	2.76	5.23	5.31	2.79	4.02	24.06
LL	12.10	8.89	19.83	20.06	8.80	12.09	81.77

(b)

Side-band data	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	sum
TT	18	13	13	12	13	16	85
TL	48	39	41	39	40	41	248
LT	216	203	329	325	214	222	1509
LL	2303	1539	2837	2702	1587	2349	13317

(c)

Table 5: Expected signal event yield and sideband data number in each bins and categories.

## 356 7 Signal modelling

The shape of the signal and background  $m_{\gamma\gamma}$  distributions is described with analytical functions. Like in the past analysis [Hasib:2238687], the shape of  $m_{\gamma\gamma}$  invariant mass distribution for signal events in each category is modelled with a *double-sided Crystal Ball* (DSCB) function. It is a composite function with 6 parameters formed by a Gaussian core, which models the peak, and two power-law tails, as detailed in Eq. 2:

$$f_{\text{DSCB}}(m_{\gamma\gamma}) = N \times \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{low} \leq t \leq \alpha_{high} \\ \frac{e^{-\frac{1}{2}\alpha_{low}^2}}{\left[\frac{1}{R_{low}}(R_{low}-\alpha_{low}-t)\right]^{n_{low}}} & \text{if } t < -\alpha_{low} \\ \frac{e^{-\frac{1}{2}\alpha_{high}^2}}{\left[\frac{1}{R_{high}}(R_{high}-\alpha_{high}+t)\right]^{n_{high}}} & \text{if } t > \alpha_{high} \end{cases} \quad (2)$$

357 where  $N$  is a normalization factor and the six parameters are

- 358 •  $\mu_{\text{CB}}$  and  $\sigma_{\text{CB}}$  describe the mean and the width of the Gaussian core, which are combined in  
359  $t = (m_{\gamma\gamma} - \mu_{\text{CB}}) / \sigma_{\text{CB}}$ ;
- 360 •  $\alpha_{low}$  and  $\alpha_{high}$  are the positions of the transitions with respect to  $\mu_{\text{CB}}$  from the Gaussian core to  
361 power-law tails, in unit of  $\sigma_{\text{CB}}$ , on the low and high mass sides respectively;
- 362 •  $n_{low}$  and  $n_{high}$  are the exponents of the low and high mass tails. With the  $\alpha$ 's, they define  $R_{low} = \frac{n_{low}}{\alpha_{low}^2}$   
363 and  $R_{high}$  similarly.

364 In the previous analysis [Hasib:2238687], the DSCB has showed a good  $\chi^2$  in the signal only fit and gives  
365 a slightly smaller bias on the fitted signal yield using injection tests on Asimov background and signal MC.  
366 The advantage of the DSCB is to well separate the contribution coming from the core and from the tails,  
367 making easier to apply systematic variations on the scale (on  $\mu_{\text{CB}}$ ) and on the resolution (to  $\sigma_{\text{CB}}$ ).  
368

369 Thanks to the very precise knowledge of the Higgs mass from Run1 from the combination of ATLAS and  
370 CMS measurements ( $m_H = 125.09 \pm 0.21$  (stat.)  $\pm 0.11$  (syst.) GeV) [8] and to the relatively small error of  
371 the mass scale systematics ( $< 1\%$ ), it is possible to parametrize the shape using just one set of MC samples  
372 simulated with the Higgs mass  $m_H$  fixed at 125 GeV. The MC sample combines all the production modes  
373 assuming Standard Model cross sections and the three MC types (mc16a+d+e) are weighted taking into  
374 account the relative luminosity of 2015/2016, 2017 and 2018 datasets, respectively. The generalization of  
375 the fitted model to any value of  $m_H$  (in a small interval close 125 GeV) is done with a simple shift of the  
376  $\mu_{\text{CB}}$  parameter ( $\mu_{\text{CB}} = m_H + \mu_{\text{CB}}^{125 \text{ GeV}} - 125 \text{ GeV}$ ).  
377

378 **Systematic from mH: Higgs mass:  $125.09 \pm 0.24$  GeV as photon energy scale (PES).**

379 The signal shapes are fitted in the fixed mass range 105 to 160 GeV while leaving floating all the six  
380 parameters of the DSCB function. The signal models for each category are shown in Figure 8. The same  
381 procedure is performed both for SM and BSM Monte Carlo samples, in every VBF categories, Optimal  
382 Observable bins and CP-mixing hypothesis the  $m_{\gamma\gamma}$  models are determined by their best fit results from  
383 MC.

384 In the final statistical workspace:

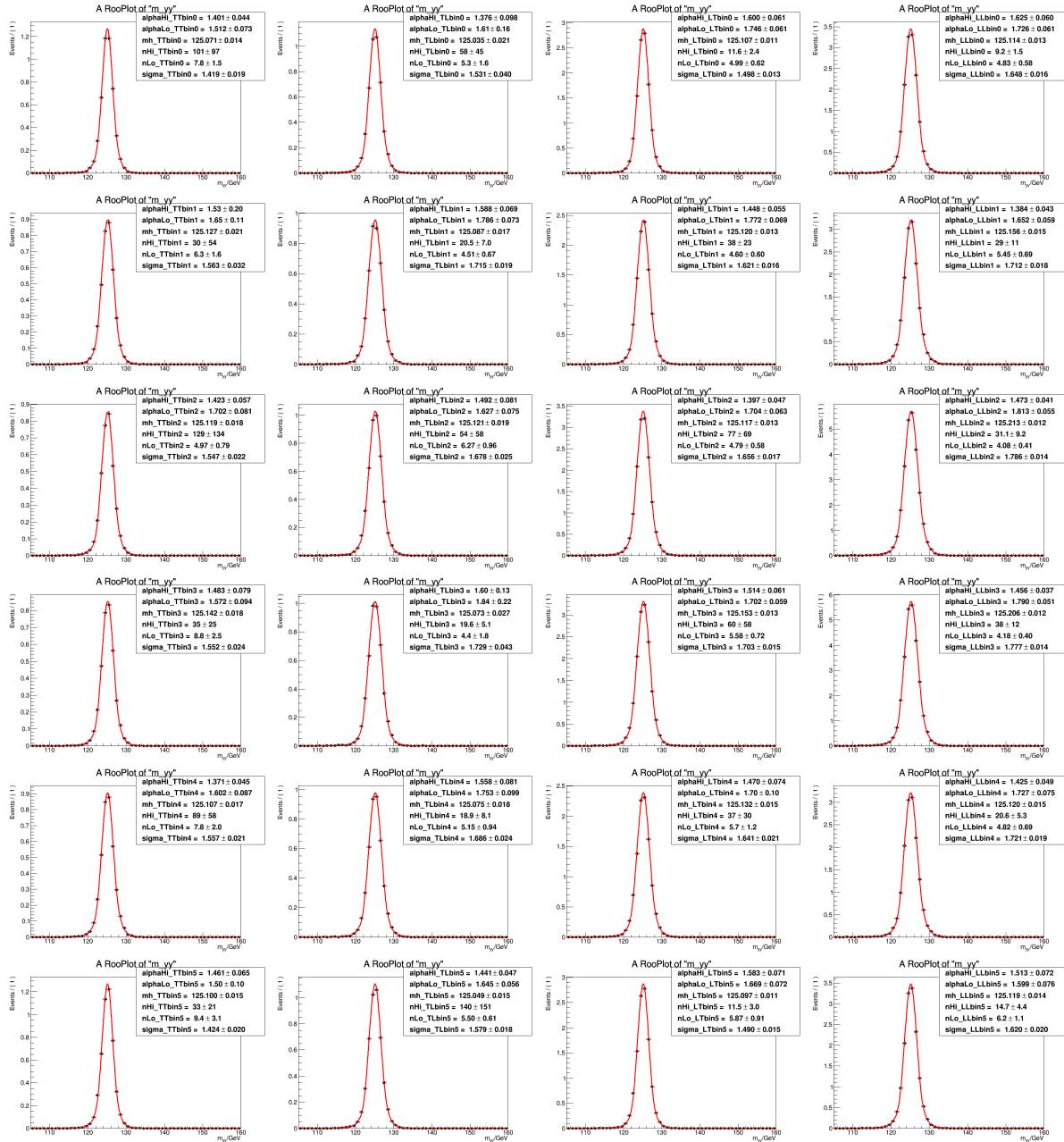


Figure 8: Fitted Doubles Sided Crystal ball shapes for all the reconstructed categories of the analysis.

- 385     • the value of  $m_H$  is constrained to the one of Run 1 measurement, taking into account its error;
- 386     • all the six parameters of the DSCB are kept fix during the fit
- 387     •  $\mu_{\text{CB}}$  and  $\sigma_{\text{CB}}$  are each multiplied by a response function to take into account photon energy scale  
388       and photon energy resolution systematics, respectively (see Section ??)

## 389 8 Background Modelling

### 390 8.1 Background modeling strategy

391 The strategy for the continuous background modelling follows what traditionally done in  $H \rightarrow \gamma\gamma$   
 392 measurements, with the addition of some dedicated procedures to cope with specific issues related  
 393 to insufficient MC statistics or very low-statistics categories. This section only contains a simplified  
 394 description, more details could be found in HGam coupling analysis ??.

395 The steps comprising the background modelling strategy can be summarised as follows:

- 396 • The composition of the continuous background in term of  $\gamma\gamma$ ,  $\gamma j$  and  $jj$  components is estimated  
 397 with data driven technique for each category entering the measurement, as detailed in Section 8.2.
- 398 • The background fractions are used to build  $m_{\gamma\gamma}$  background templates for each category entering  
 399 the measurement, where the  $\gamma\gamma$  component is obtained from MC, while the  $\gamma j$  and  $jj$  components  
 400 are measured from data driven control regions and used to reweight the  $\gamma\gamma$  MC, as detailed in  
 401 Section 8.3.
- 402 • In order to improve the background modeling and reduce the impact of limited MC statistics, a  
 403 smoothing procedure is being developed to be eventually applied to the background template, provided  
 404 that the improvements on the results outweigh the possible associated systematic uncertainties.
- 405 • The Spurious Signal approach is used in the smoothed  $\gamma\gamma$  templates to select a functional form to  
 406 describe each template, as well as the associated bias that will enter the measurement as a systematic  
 407 uncertainties, as detailed in Section 8.4. The Spurious Signal criteria are defined in order to cope  
 408 with the limited MC statistics that in some case leads to non-physical large bias measurements.
- 409 • In very low-statistic categories, defined as those categories with less than 10 events per bin in the  $m_{\gamma\gamma}$   
 410 sidebands, the smoothing procedure tends to introduce large non-physical structures in the templates,  
 411 and it is therefore not suited to be used for them. For this reason, for this categories we initially select  
 412 the statistically-justified function performing a Wald-test on the data sidebands using a family of  
 413 function based on a exponential of increasing complexity. We accept the simplest function justifies  
 414 by the Wald-test, and perform the Spurious Signal measurement with this function tolerating a bias  
 415 that could be larger than the criteria used for the high-statistic categories, but would have a small  
 416 impact on the overall sensitivity. This procedure is described in Section 8.4.1.

### 417 8.2 Background compositions

418 The dominant background entering the invariant mass spectrum comes from SM continuum diphoton  
 419 production. A large number of photons are produced inside jets due to decays of neutral mesons to  
 420 photon pairs. Thus, photon-jet and di-jet events in which the jets are mis-identified as photons represent a  
 421 non-negligible source of background. Photon-pairs can be also faked from Drell-Yan events in which both  
 422 electrons are misidentified as photons. However, this background only contributes with a small fraction  
 423 (< 1%).

424 The number of  $\gamma\gamma$ ,  $\gamma j$  and  $jj$  events entering each category after the final selection is estimated by means  
 425 of a double two-dimensional sideband ABCD method [ATL-COM-PHYS-2012-592]. This data-driven  
 426 method extrapolates the fraction of fake photons within the signal region from the composition of the

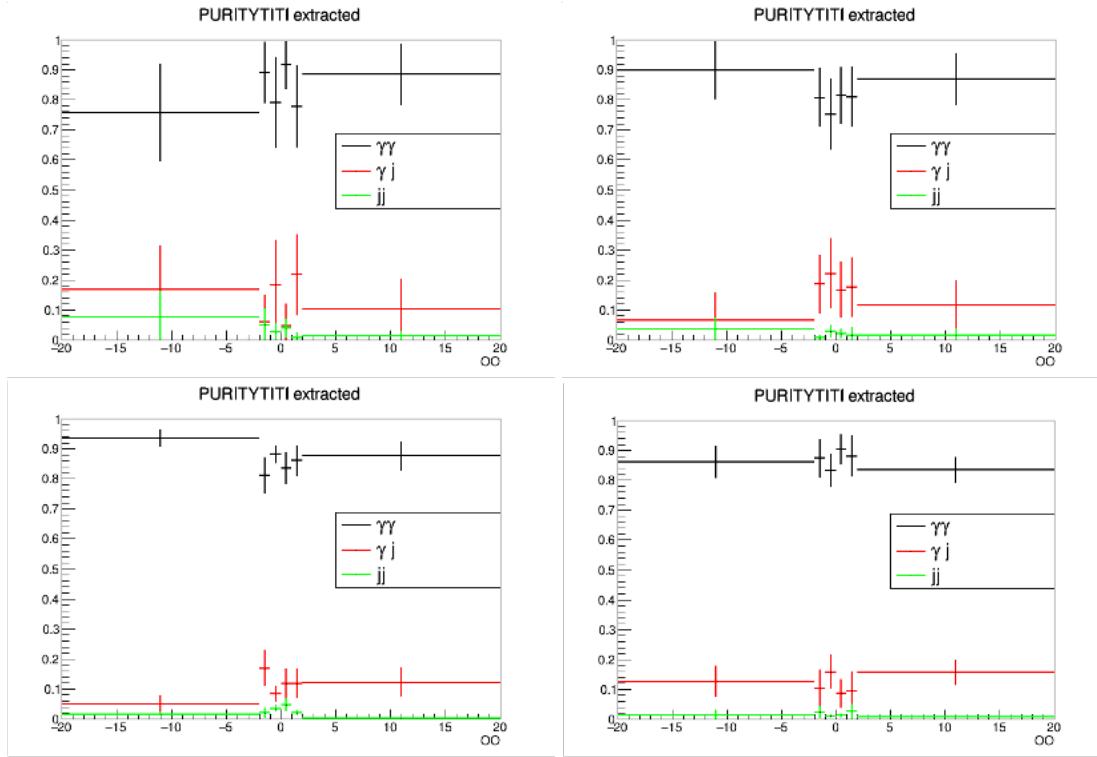


Figure 9:  $\gamma\gamma$ ,  $\gamma j$ ,  $jj$  fraction in TT (upper left), TL (upper right), LT (bottom left), LL (bottom right) categories, as a function of OO.

427 side-band control regions, built by inverting photon identification and isolation requirements. The fractions  
 428 are calculated individually in each categories and OO bins for the background template building. Table 6  
 429 lists the  $\gamma\gamma$  fraction in all categories. More details are shown in Appendix D.

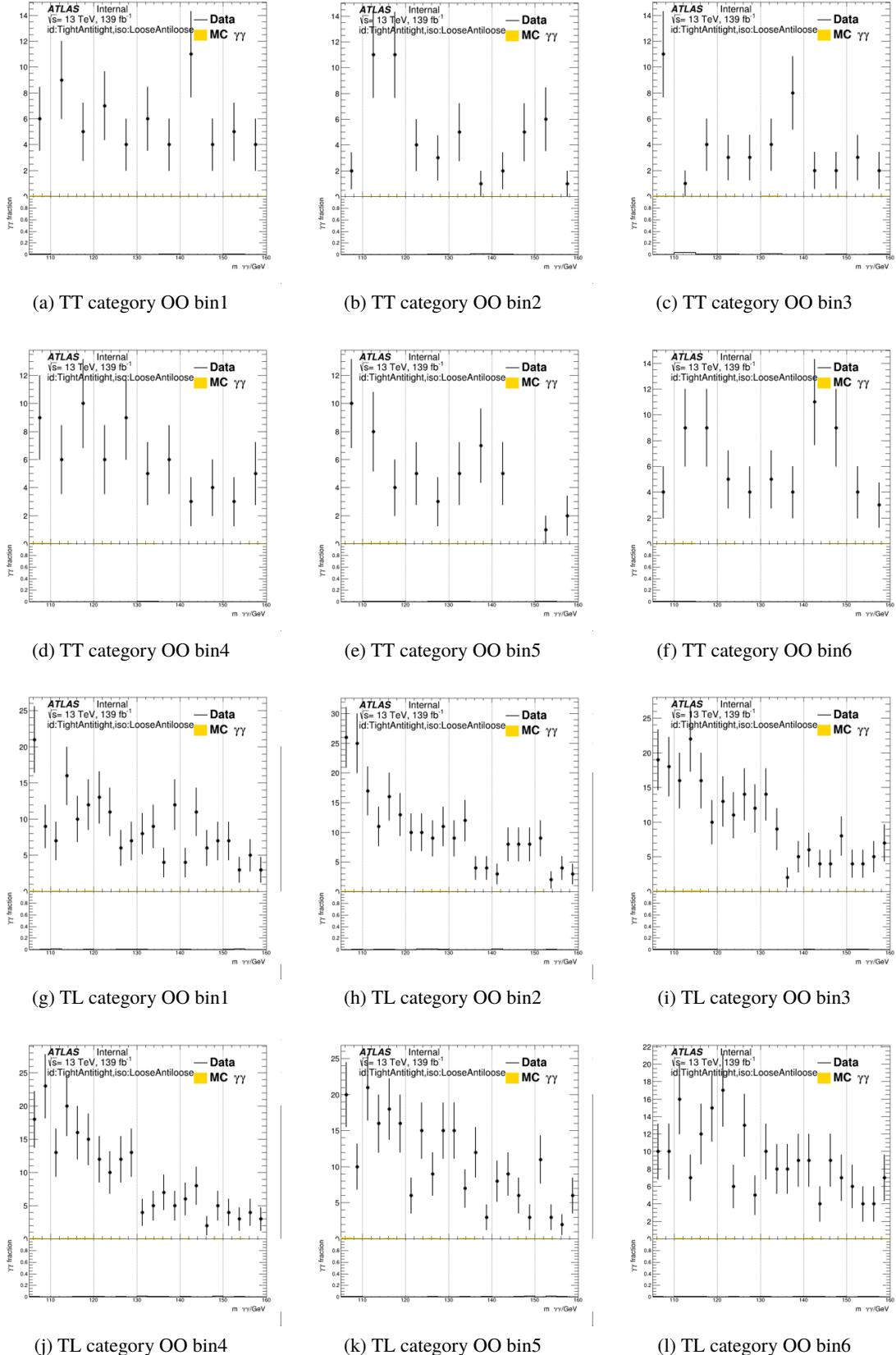
	TT	TL	LT	LL
OO bin1	$0.76 \pm 0.16$	$0.90 \pm 0.10$	$0.93 \pm 0.03$	$0.86 \pm 0.05$
OO bin2	$0.89 \pm 0.10$	$0.81 \pm 0.10$	$0.81 \pm 0.06$	$0.87 \pm 0.07$
OO bin3	$0.79 \pm 0.15$	$0.75 \pm 0.12$	$0.88 \pm 0.03$	$0.83 \pm 0.06$
OO bin4	$0.92 \pm 0.08$	$0.81 \pm 0.09$	$0.84 \pm 0.05$	$0.90 \pm 0.05$
OO bin5	$0.78 \pm 0.14$	$0.81 \pm 0.10$	$0.86 \pm 0.05$	$0.88 \pm 0.07$
OO bin6	$0.88 \pm 0.10$	$0.87 \pm 0.09$	$0.88 \pm 0.05$	$0.83 \pm 0.04$

Table 6:  $\gamma\gamma$  fraction calculated with double two-dimensional sideband ABCD method.  $\gamma j$  and  $jj$  components are merged together in following analysis, so the fraction is  $1 - f_{\gamma\gamma}$ .

**430 8.3 Background templates**

431 Considering the low  $jj$  fraction and similar performance of  $\gamma j$  and  $jj$ , those two components are merged  
432 together as  $\gamma j + jj$  for the background estimation. The shape is determined with a data control region, in TT,  
433 TL and LT categories it is defined as at least one of the two selected photons fail the tight identification and  
434 loose isolation requirement due to insufficient MC statistics, and in LL category it is defined as inverting  
435 the identification and isolation of exactly one of two photon candidates. Contamination from  $\gamma\gamma$  has  
436 been tested to be ignorable as shown in Figure 10, and a **linear** reweighting is derived to match the  $\gamma\gamma$   
437 and control region shapes. The bin width for ratio function smoothing varies from 5GeV, 2GeV, 1GeV  
438 and 0.5GeV respectively in TT, TL, LT, LL categories to match the MC statistics for an un-biased result.  
439 Uncertainty of this reweighted template contains two part: from the choice of smoothing function and  
440 from  $\gamma\gamma$  fraction, while the latter is dominant. These derived shapes are combined with the measured  
441 relative event fractions to obtain the total background shape. The  $\gamma\gamma$  fraction is measured as mentioned in 8.2  
442

443 The obtained templates show a good agreement when comparing them with events in data passing the nom-  
444 inal selection (with tight photon identification and isolation cuts, TI) in the sidebands, as shown in Figure 12.  
445



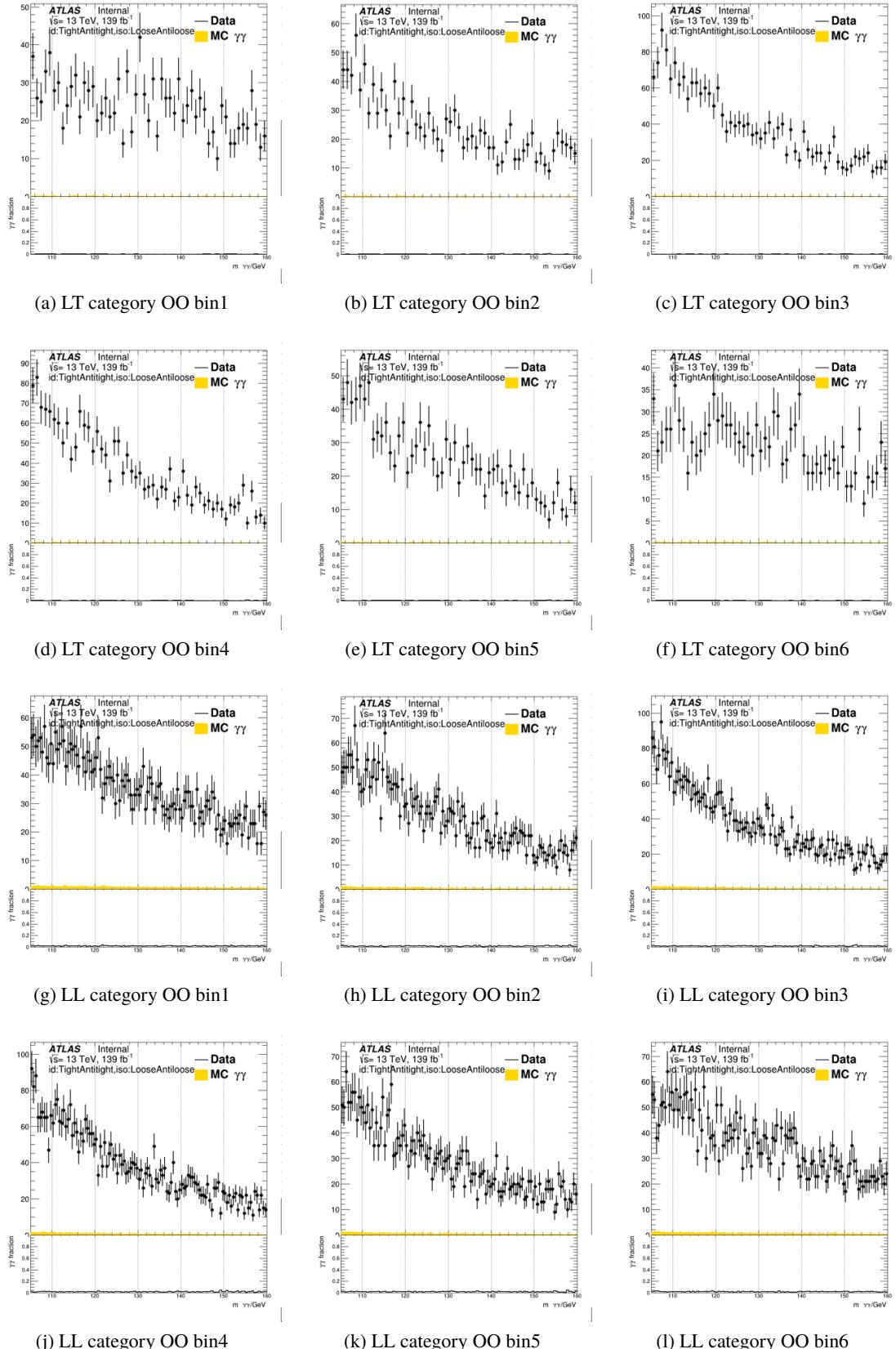
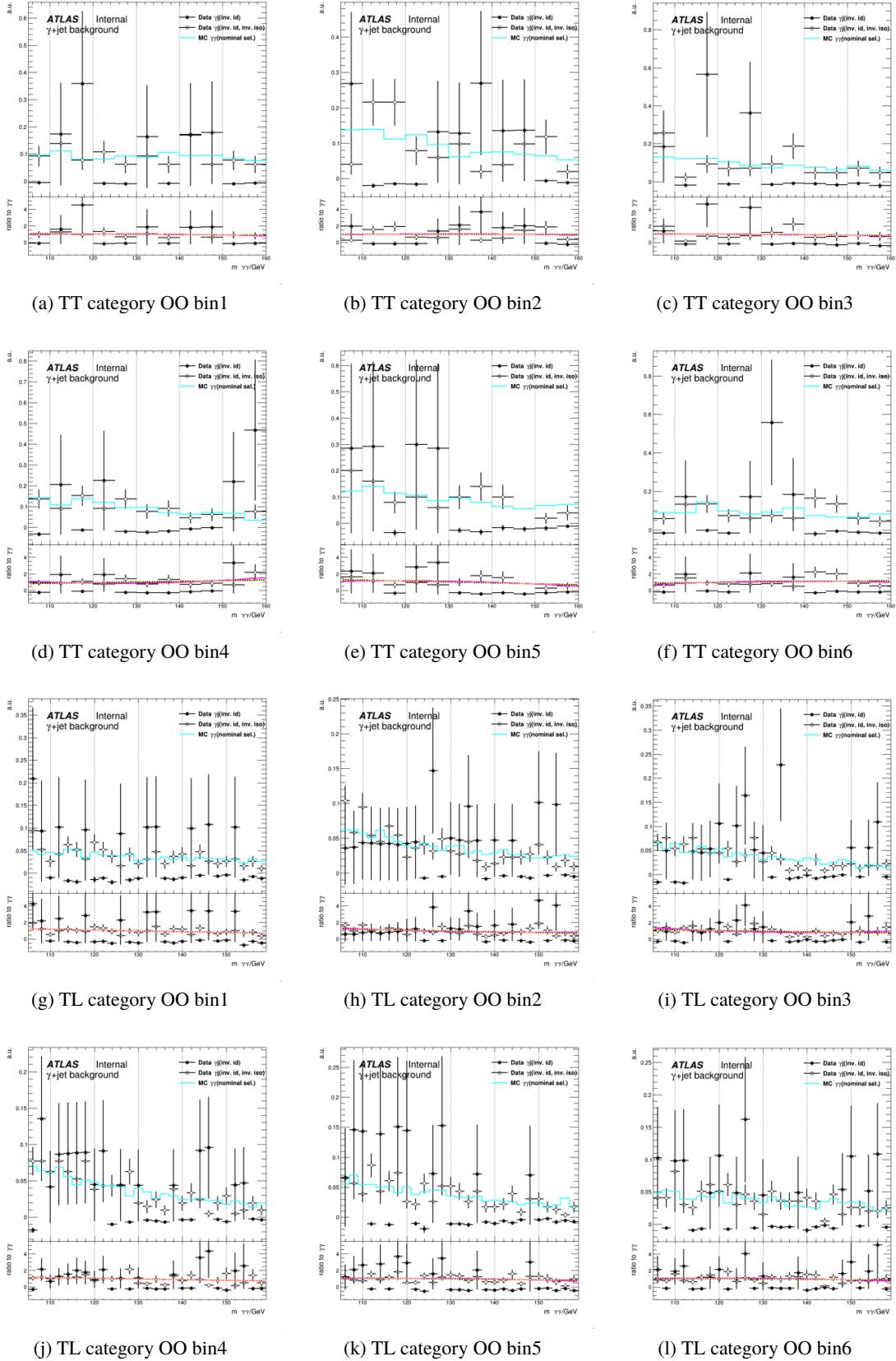
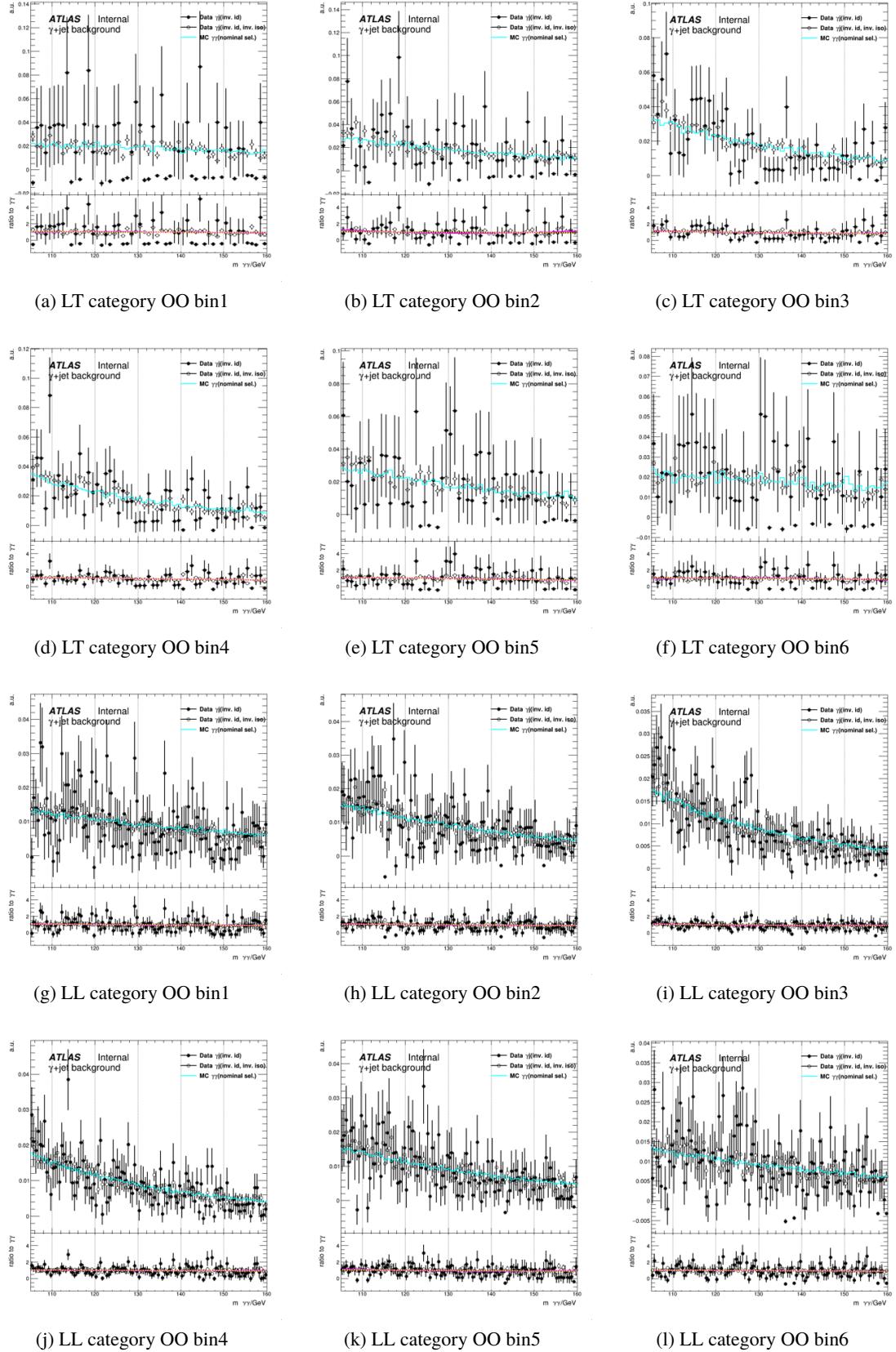
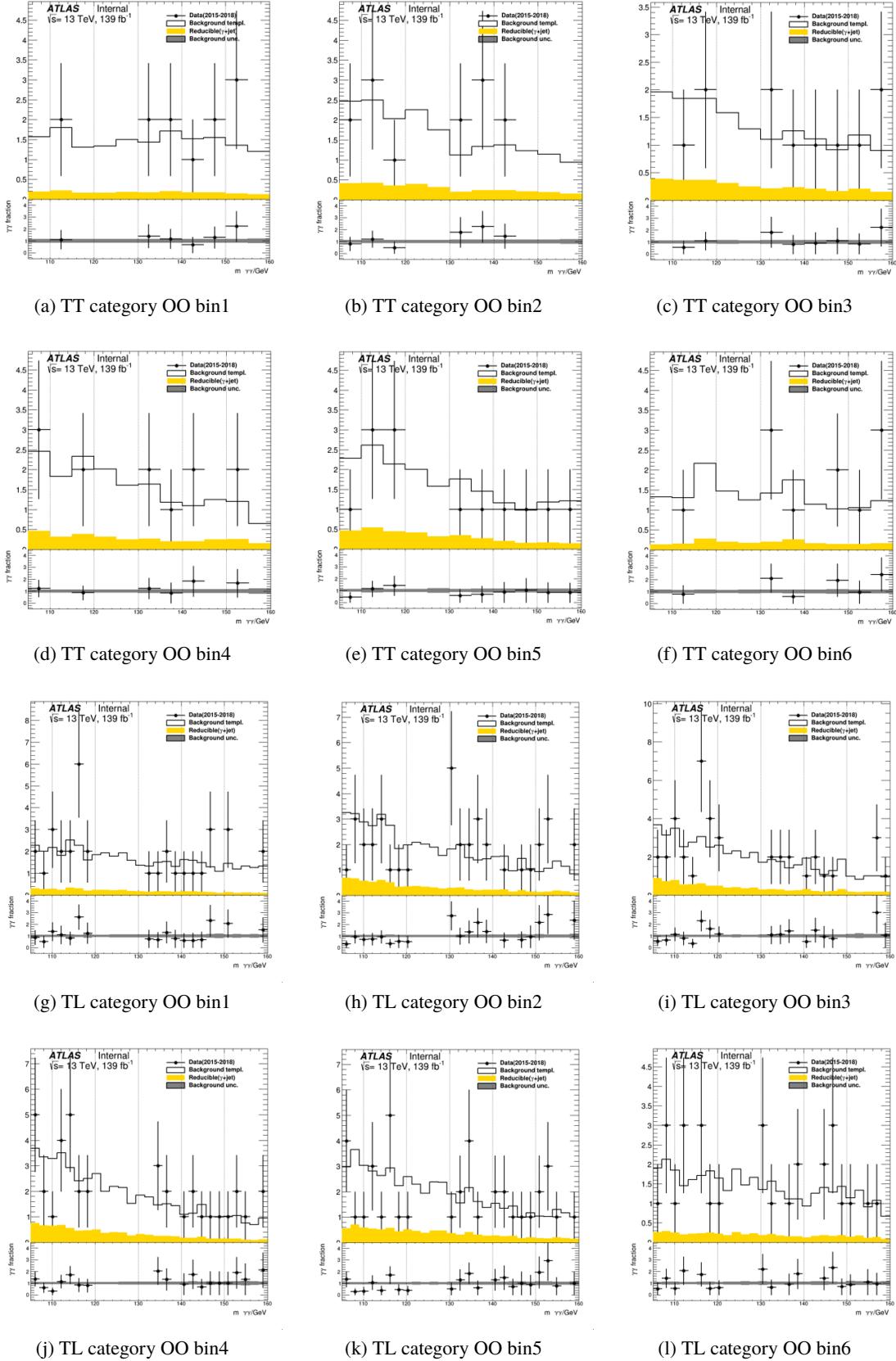


Figure 10: Control region data in all categories and OO bins.



Figure 11: Reweight  $\gamma\gamma$  MC with a linear function to match the  $\gamma j + jj$  shape in each bins and categories.



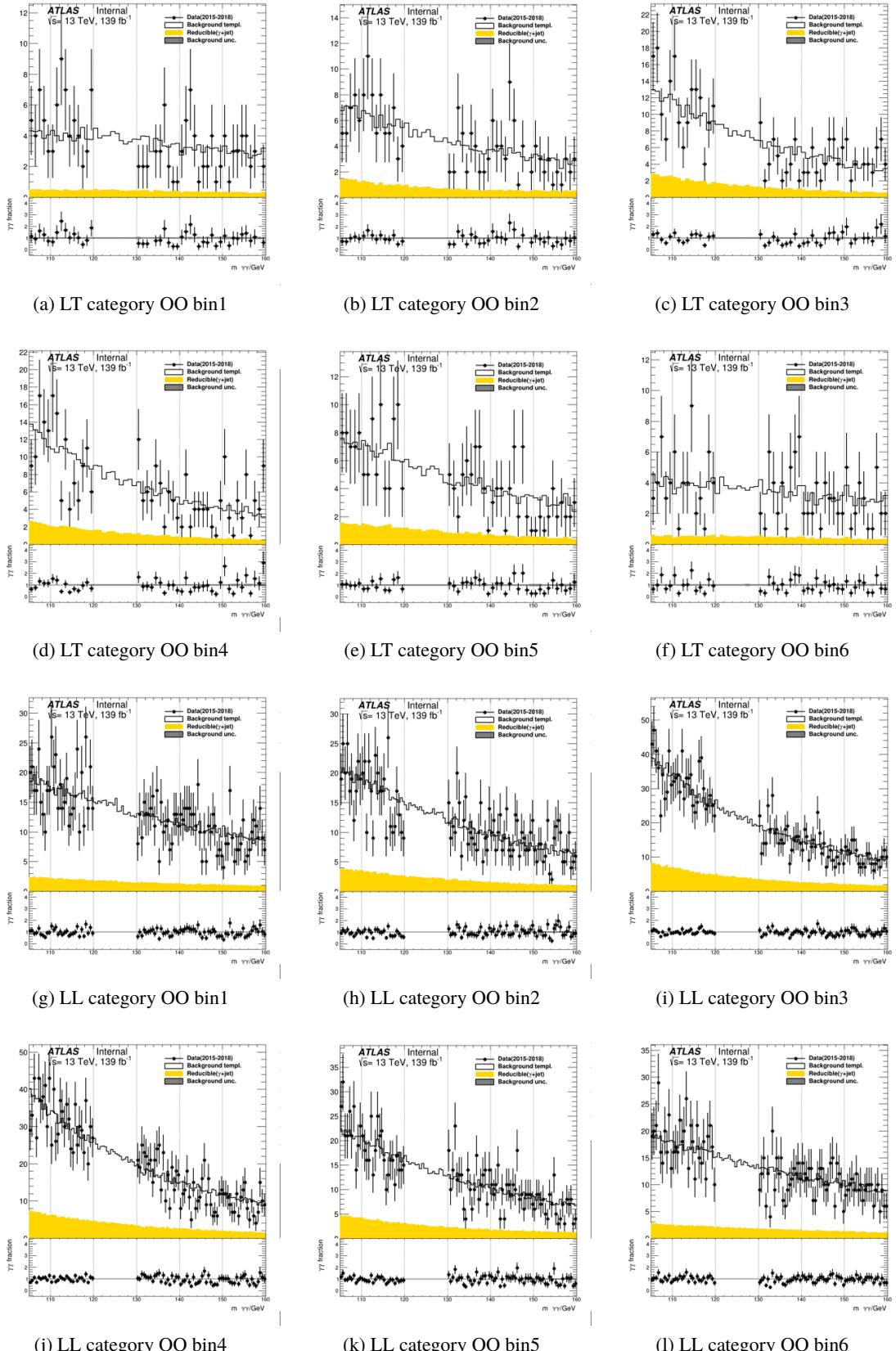


Figure 12: Final background templates.

446 **8.3.1 Background smoothing**

447 **Need Simplification**

448

449 Because of the limited statistics of the simulation samples used in the calculation, the value of the  
 450 spurious signal systematic uncertainty is subject to significant statistical fluctuations within many of the  
 451 analysis categories [Hyneman:2712576]. Statistical fluctuations in the background-only sample may  
 452 cause signal-like bumps, which are then fit as signal by the spurious signal procedure. These statistical  
 453 fluctuations do not capture the shape mismodeling from the analytic function, and they often drastically  
 454 inflate the value of the systematic.

455 Although simply producing additional simulation samples would alleviate the issue of statistical fluctuations  
 456 fitted as spurious signal, producing more simulated events is computationally expensive. Additionally,  
 457 producing events which fall into specific phase spaces is often highly inefficient. Therefore, an alternative  
 458 solution using the available simulation samples is preferred.

459 A Gaussian Process (GP) is defined as a set of random processes, where all finite subsets of these processes  
 460 have a multivariate normal distribution [ebden2015gaussian]. Given a finite dataset – such as the bin  
 461 contents of a smooth histogram – with corresponding mean and covariance matrices, a Gaussian Process  
 462 may be defined. The “correct” mean and the covariance, however, are not necessarily well defined, as  
 463 they encode specific assumptions about the underlying dataset. In practice, the two quantities are fit to a  
 464 finite dataset using a minimization algorithm. In the case of a one-dimensional histogram with a finite  
 465 number of bins, the mean can be interpreted as a “rough” description of the underlying shape. The diagonal  
 466 elements of the covariance matrix represent the error of each bin while the off-diagonal elements specify  
 467 how “similar” the bin content of two different bins should be.

468 The covariance matrix can be simplified through the introduction of a kernel, which analytically determines  
 469 the level of correlation between two distinct points (i.e., the length scales in X at which points are expected  
 470 to influence one another in Y). Two useful kernels are the Radial Basis Function (RBF) kernel and the  
 471 Gibbs kernel [3569, Gibbs].

472 The RBF kernel has one hyperparameter, the constant length scale  $l$ , and it is defined as

$$K_{\text{RBF}}(x, x') = \exp\left(\frac{-(x - x')^2}{2l^2}\right) \quad (3)$$

473 The RBF kernel is useful for mostly-flat functions. However, for smoothly-falling functions, it is likely  
 474 that nearby points will be more correlated in some regions than in others, so a constant length scale is a  
 475 suboptimal model. The Gibbs kernel allows the length scale  $l(x)$  to vary linearly as a function of  $x$ , and  
 476 thus has two hyperparameters: the initial length scale and the length scale slope. The Gibbs kernel function  
 477 is:

$$K_{\text{Gibbs}}(x, x') = \frac{\sqrt{2l(x)l(x')}}{l(x)^2 + l(x')^2} \cdot \exp\left(\frac{-(x - x')^2}{l(x)^2 + l(x')^2}\right) \quad (4)$$

478 The background templates used in the spurious signal test for the analysis categories are all smooth,  
 479 roughly exponentially falling distributions with statistical fluctuations. Fitting a background template  
 480 using Gaussian Process Regression (using the Gibbs kernel with the errors as determined by the initial

481 templates) offers a consistent method of estimating the underlying smooth shape of the template, without  
 482 the problematic fluctuations. Notably, the GP smoothing technique makes no assumption on the underlying  
 483 distribution other than that it is smooth and falling, hence the choice of functional form from the spurious  
 484 signal test will not be biased.

485 The hyperparameters (initial length scale and length scale slope) are allowed to vary over a range specified  
 486 by the user; the optimal hyperparameters within this range are determined by the Gaussian Process fitting  
 487 procedure.

488 A GP is fit to the original (noisy) background template in each category. The GP mean in the fits is  
 489 defined as an exponential function, the parameters of which are obtained by a fit to the original background  
 490 template. The exponential shape has been observed to be a sufficiently close guess for the categories used  
 491 by the analysis. However, in cases where the input template has very few statistics (less than about ten  
 492 events per bin on average), the resulting GP fit may be nearly identical to the mean exponential shape. This  
 493 issue occurs when the statistical uncertainties of the original template are so large that the template is  
 494 fully compatible with the preliminary exponential shape. Although the exponential shape is technically an  
 495 adequate descriptor of the template shape, the choice of the exponential mean does bias the functional  
 496 choice of the spurious signal test in this case. Therefore, a check has been added to re-perform the GP fit  
 497 using a flat mean in cases where the resulting GP shape and the mean exponential shape disagree with a  
 498  $\chi^2/DoF < 0.1$ .

499 The resulting smoothed shape obtained from the GP fit is then saved as a new histogram. This smoothed  
 500 histogram is passed as the background template to the spurious signal test, which then determines the  
 501 background functional form and spurious signal systematic uncertainty as described in Section 8.4.

502 Note that based on previous study, the GPR method can only remain effectively unbiased for smoothly-falling  
 503 templates containing more than an average of 20 Monte Carlo events/ bin [ATLAS-CONF-2020-026].  
 504 The TT and TL categories in this analysis even do not have sufficient statistics for GPR. Only LT and LL  
 505 categories used this method, and the treatment of TT and TL is discussed in Section 8.4.1.

506 Examples of the smoothed templates are presented in Figure 13 for a sample category with a high level of  
 507 statistics and for a category with a very low level of statistics. The original templates are shown as well, for  
 508 comparison. The data sidebands are also shown for validation, although the GP smoothing technique does  
 509 not take into account the data sidebands.

## 510 8.4 Spurious signal test

511 The background  $m_{\gamma\gamma}$  shape, for each analysis category, is described using an analytic function whose  
 512 parameters and normalization are fitted to data. The choices for the analytic function that has been  
 513 considered are:

- 514 • Exponential Function:  $f(m_{\gamma\gamma}) = e^{c \cdot m_{\gamma\gamma}}$
- 515 • Exponential Function of 2<sup>nd</sup> Order Polynomial:  $f(m_{\gamma\gamma}) = e^{c_1 \cdot m_{\gamma\gamma}^2 + c_2 \cdot m_{\gamma\gamma}}$
- 516 • Exponential Function of 3<sup>rd</sup> Order Polynomial:  $f(m_{\gamma\gamma}) = e^{c_1 \cdot m_{\gamma\gamma}^3 + c_2 \cdot m_{\gamma\gamma}^2 + c_3 \cdot m_{\gamma\gamma}}$
- 517 • Bernstein polynomial of order N:  $B_N(m_{\gamma\gamma}) = \sum_{i=0}^N c_i \cdot b_{i,N}$  with  $b_{i,N} = \binom{N}{i} m_{\gamma\gamma}^i (1 - m_{\gamma\gamma})^{N-i}$ , N=3,  
 518 4, 5

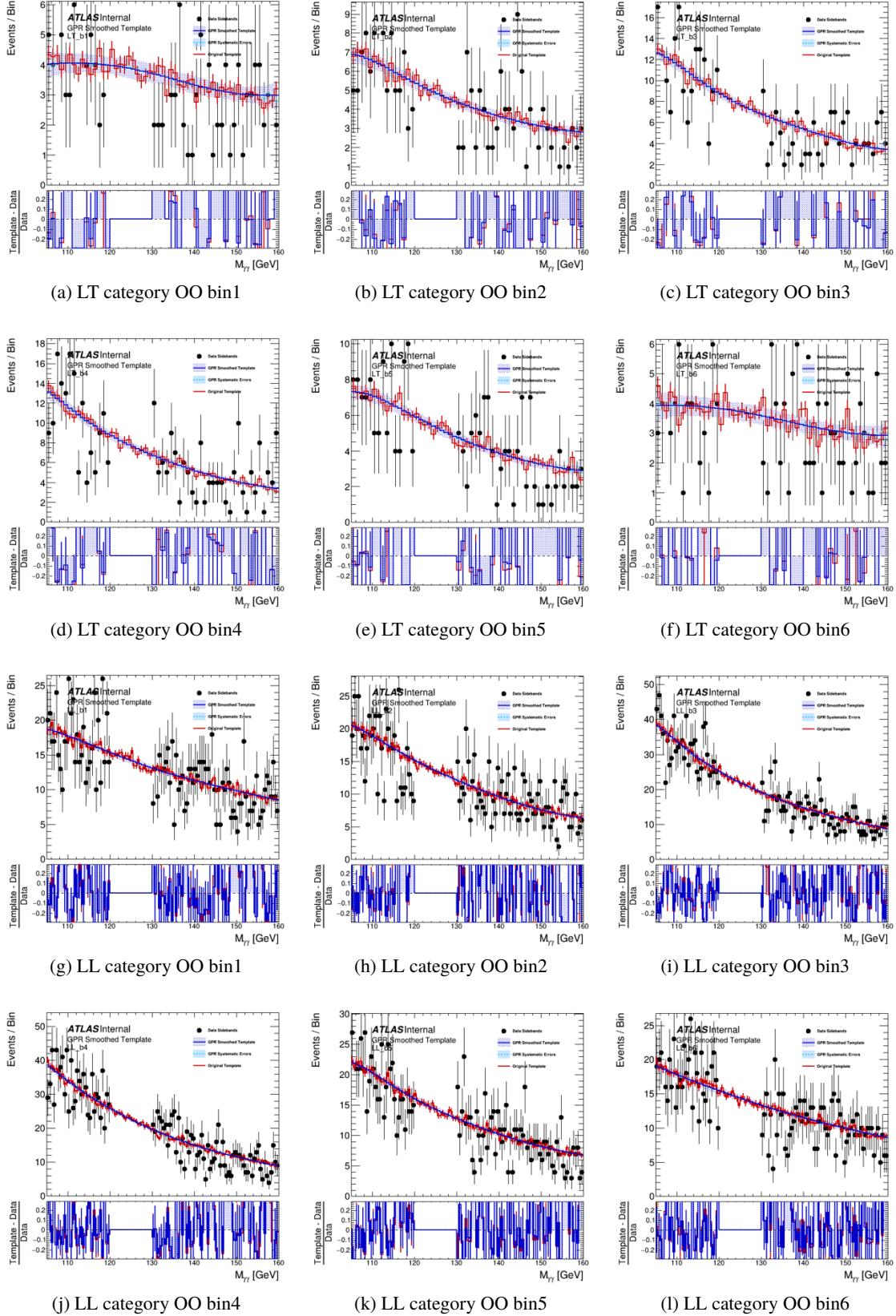


Figure 13: LT and LL category background templates after GPR smooth.

- 519 • First-Order Power Law Function:  $f(m_{\gamma\gamma}) = m_{\gamma\gamma}^c$

520 The method to select the functional form in each category is based on the spurious signal. The spurious  
521 signal test is described in more detail in [Hasib:2238687].

522 To perform the spurious signal test, the full analytic signal plus background model is fitted to a background-  
523 only template for each category separately. The fit is performed in the nominal diphoton mass range of  
524  $105 \leq m_{\gamma\gamma} \leq 160$  GeV. The number of fitted signal events as a function of the Higgs mass is computed  
525 in intervals of 1 GeV within the diphoton mass range of  $120 \leq m_{\gamma\gamma} \leq 130$  GeV. Negative fluctuations  
526 are taken into account. The number of spurious signal events  $N_{sp}$  is then defined as the maximum of  
527 the absolute value of the fitted number of signal events within the signal diphoton mass sub-range. The  
528 statistical error is always taken into account.

529 The choice of analytic function is done such that the function satisfies at least one of the following criteria:

- 530 •  $N_{sp} < 10\% N_{s,exp}$ , where  $N_{s,exp}$  is the expected number of signal events within the SM in the  
531 category ( $\mu_{sp} = N_{sp}/N_{s,exp}$ )
- 532 •  $N_{sp} < 20\% \sigma_{bkg}$ , where  $\sigma_{bkg}$  is the statistical uncertainty on the fitted number of signal events,  
533 when fitting the signal+background model to a background only Asimov dataset ( $Z_{sp} = N_{sp}/\sigma_{bkg}$ ).

534 If multiple functions pass the test, then the one with the least number of parameters is chosen.

A relaxed spurious signal criteria, that are made to accommodate  $2\sigma$  local statistical fluctuations in the MC template, in case the MC statistics is small compared to the expected signal. This is done by defining a new variable:

$$\zeta_{sp} = \begin{cases} N_{sp} + 2\Delta_{MC}, & N_{sp} + 2\Delta_{MC} < 0 \\ N_{sp} - 2\Delta_{MC}, & N_{sp} - 2\Delta_{MC} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

535 where  $\Delta_{MC}$  is a local statistical fluctuation of the MC background template. This is illustrated in Figure ??.  
536 This newly defined variable should then pass the criteria as  $N_{sp}$  before.

537 For categories with at least 100 data events in the sideband regions defined as  $105 < m_{\gamma\gamma} < 120$  GeV or  
538  $130 < m_{\gamma\gamma} < 160$  GeV, the background functions are selected to result in a  $|S_{\text{spur}}|$  that is smaller than either  
539 10% of the total expected Higgs boson signal event ( $S_{\text{exp}}$ ) or 20% of the statistical uncertainty of the fitted  
540 signal yield ( $\sigma_{\text{exp}}$ ). The fit of the analytic function to the background template is required to yield a  $\chi^2$   
541 probability of at least 1% <sup>1</sup>.

542 If multiple functions pass the requirements, the one with the smallest number of degrees of freedom is  
543 chosen. If no function passes, the spurious signal requirement is relaxed to  $\zeta_{sp} < 10\% N_{s,exp}$  or  $\zeta_{sp} <$   
544  $20\% \sigma_{bkg}$ . For 28 out of the 88 categories, these relaxed conditions were used. It's to be emphasized that  
545 even the relaxed condition is used, the background uncertainty is still  $N_{sp}$  from the nominal approach of  
546 spurious signal study.

---

<sup>1</sup> The  $\chi^2$  is computed with a background template uniformly binned over  $105 < m_{\gamma\gamma} < 160$  GeV. The number of bins is 22, and the degrees of freedom used in the computation is  $21 - N_{\text{pars}}$ , the number of free function parameters.

547 **8.4.1 Treatment of low-statistic categories**

548 The determination of analytical functions is part of the result of spurious signal study, however, for low  
 549 stastics categories with a background template suffering of very large statistical fluctuation we introduce  
 550 a preliminary step. The straight use of the spurious signal procedure on these templates, in fact, would  
 551 potentially lead to select a function with large number of degrees of freedom to accomodate these statistical  
 552 fluctuation, and these function can prove to be dangerous when used on data. For this reason, for categories  
 553 with less than 100 events in the  $m_{\gamma\gamma}$  sidebands, we restrict the selection of the modeling function to  
 554 exponential of first, second and third order, collectively described by:

$$f_i x = e^{\sum_{i=0}^2 a_i x^i} \quad (6)$$

555 where  $x = m_{\gamma\gamma}$ .

556 We accept as modeling function the  $f_i$  function with the lowest number of degrees of freedom for which  
 557 the addition of more degrees of freedom does not lead to a significant improvement in the quality of the fit  
 558 of the  $m_{\gamma\gamma}$  sidebands. To this purpose, we checked the likelihood ratio between the fits to the side-band  
 559 data with two different background models:

$$\lambda_{(1,2)} = -2 \ln(L_1/L_2) \quad (7)$$

560 Where  $L_1$  and  $L_2$  are the likelihood values of the two fits.

561 The distributions of these test statistics are computed with  $m_{\gamma\gamma}$  side-band data, fitting it with two models  
 562 and calculating the value of the test statistics. Once the distributions of the test-statistics are computed, we  
 563 can reject the simpler (less parameters) background model by checking the P-value  $P(\lambda \geq \lambda_{\text{data}})$  with the  
 564  $\lambda$  measured in data ( $\lambda_{\text{data}}$ ). This procedure to reject simpler background function is called Wald test.

565 Figure 14 shows plots of the Wald-test on the low-statistic categories. All of the categories selected  
 566 exponential function as the final decison. Where the lowest P-value among all categories are around  
 567 3%, corresponding to a significance of  $2.2\sigma$  standard deviations. The background uncertainty for these  
 568 categories will be the spurious signal from exponential function.

569 As a cross-check, a F-test is also performed in Appendix ??, and we draw the same conclusion that all  
 570 of the categories can select exponential function. The reason we use Wald test to decide the background  
 571 function is the fits on side-band data are un-bin fit (so that it's consistent with the final result), there might  
 572 be bias if we calculate the  $\chi^2$  values in an artifical bin-size for the F-test. In this case the Wald test based  
 573 on likelihood ratio is more reliable.

574 **8.5 Background modelling results**

575 The determined analytical functions modeling the baground in each category entering the measurement  
 576 are found in Table 7 , as well as the corresponsing Spurious Signal bias. The background only fit on the  
 577 background templates based on MC samples is shown in Figure 15

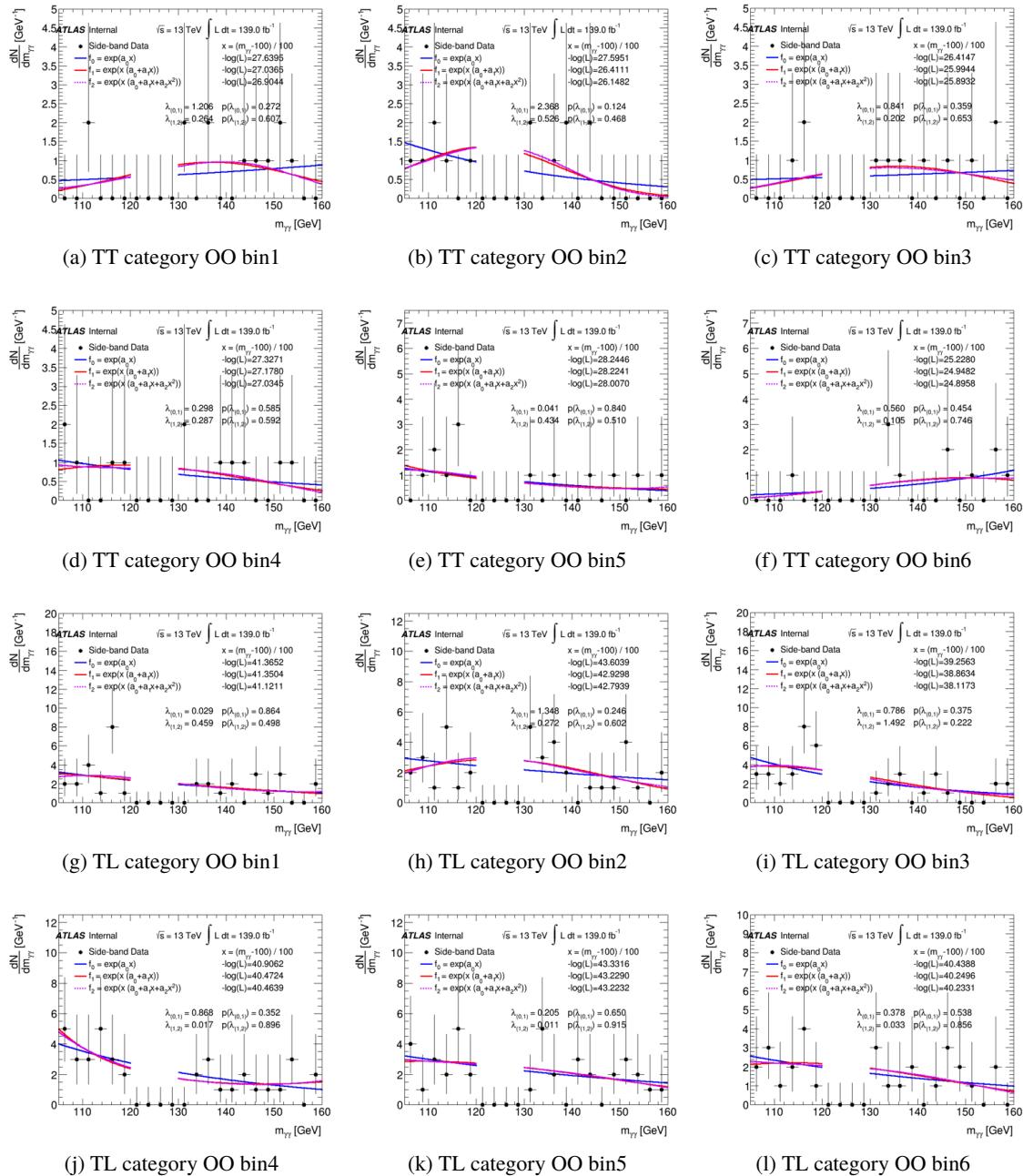
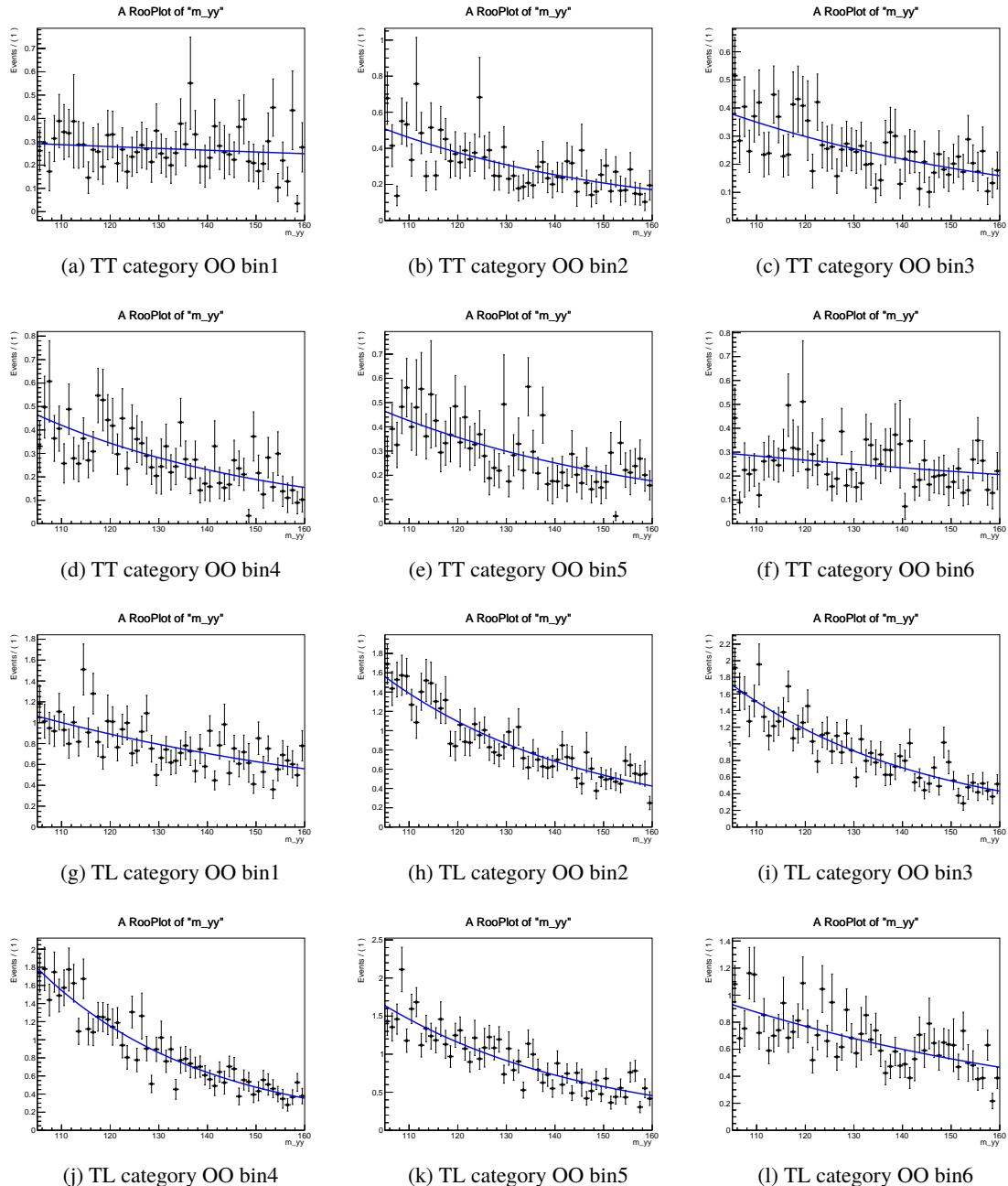


Figure 14: Wald test in TT and TL categories. In all OO bins the first order exponential function shows the lower p value and is selected.

	TT		TL		LT		LL	
	Selected func.	maxS						
bin0	Exp	-0.29	Exp	-0.468	Exp	0.87	Exp	0.885
bin1	Exp	0.582	Exp	-0.882	Exp	-0.534	Exp	1.54
bin2	Exp	0.235	Exp	-0.355	Pow	0.437	ExpPoly2	-1.64
bin3	Exp	0.409	Exp	-0.304	Pow	0.642	Exp	-0.718
bin4	Exp	-0.404	Exp	0.779	Exp	0.775	Exp	-1.19
bin5	Exp	0.293	Exp	0.408	Exp	0.627	Exp	0.52

Table 7: The final background modelling decision and the size of spurious signal uncertainties.



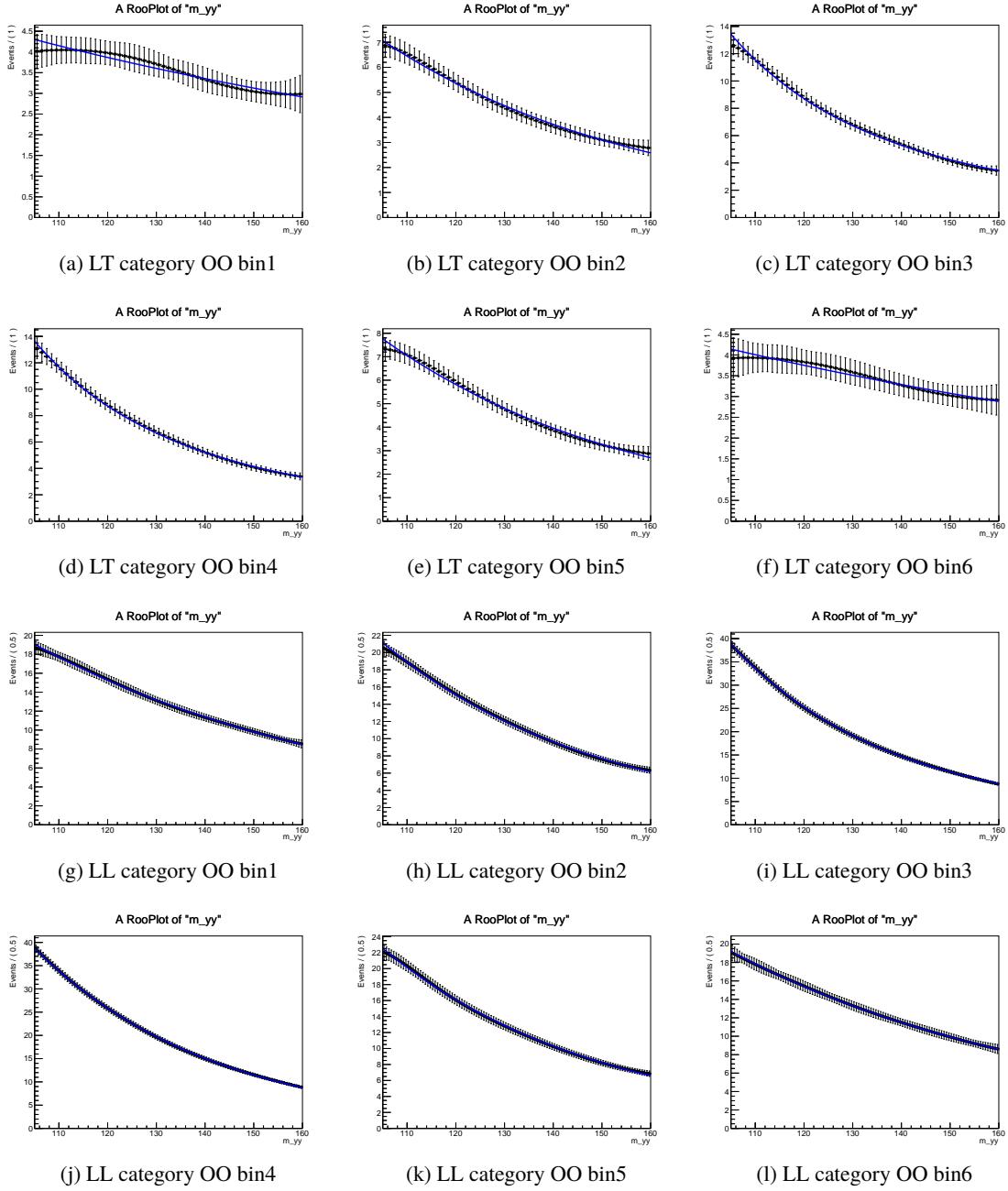


Figure 15: Background only fit on the background templates based on MC samples.

## 578 9 Systematic uncertainties

### 579 9.1 Theory uncertainties

580 The estimation of the CP BSM parameters from the OO shape requires the precise knowledge of the SM  
 581 contributions and related uncertainties. This includes both SM gluon-gluon fusion (background) and SM  
 582 VBF (signal). Therefore, the different theoretical uncertainties affecting the SM predictions are included in  
 583 the fitting likelihood. This includes uncertainties from PDF, QCD order and  $\alpha_s$ .

584 For the gluon-fusion production mode, nine uncertainty sources are used to model the QCD theory  
 585 uncertainties, following the recommendation of the LHC Higgs cross section working group [LHCXS\_4].  
 586 These sources are:

- 587 • two sources correspond to yield uncertainties related to the total cross section. Their magnitude is  
 588 taken from the STWZ-BLPTW predictions [ggF\_qcd\_unc\_1, ggF\_qcd\_unc\_2, LHCXS\_4] and their impact on the different bins is evaluated using NNLOPS.
- 589 • two sources correspond to migration uncertainties related to splitting the phase space by jet  
 590 multiplicity. Their magnitude and impact are derived similarly to the yield uncertainties
- 592 • two uncertainty sources are related to the  $p_T^H$  shape and are estimated from scale variations in  
 593 NNLOPS, including variations of the HNNLO input scales and the renormalization and factorization  
 594 scales in Powheg.
- 595 • two uncertainty sources related to the enhancement of uncertainties for events with typical VBF  
 596 topology (due to explicit or implicit third-jet vetos), and are estimated by scale variations in  
 597 MCFM [MCFM], and the corresponding uncertainties are estimated using the same procedure use  
 598 for yield and migration uncertainties.
- 599 • one uncertainty source is related to the treatment of  $m_t$  and is most important at large  $p_T^H$ .

600 Following the recommendations of PDF4LHC [pdf4lhc], the PDF uncertainties are evaluated using the 30  
 601 eigenvectors set and treating each of them as an uncorrelated source. One additional nuisance parameter  
 602 accounts for the uncertainties in  $\alpha_s$ .

603 For the  $VBF + VH$  production modes, QCD uncertainties are estimated as an envelope of the scale  
 604 variations available in Powheg [Nason:2009ai, VBFVH\_theoryUnc]. Uncertainties from the choice of  
 605 the PDF set and  $\alpha_s$  are evaluated similar to the gluon-fusion case. A summary of the magnitude of these  
 606 uncertainties in the different categories and OO bins is shown in the Figure 16 17.

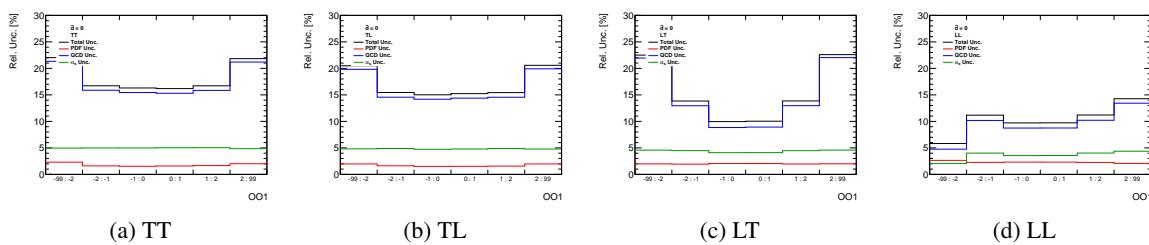


Figure 16: Theory uncertainty in different OO bins for ggF process

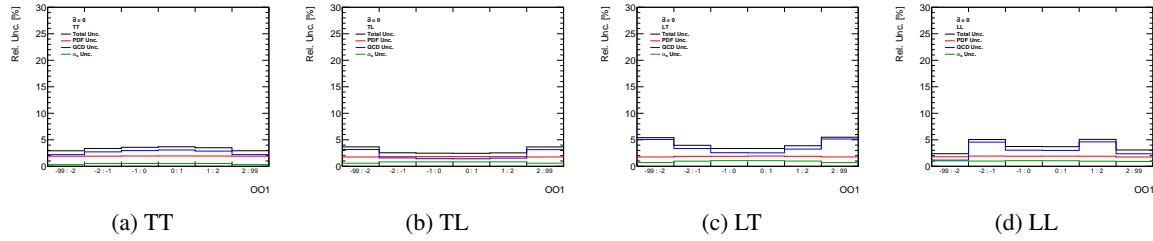


Figure 17: Theory uncertainty in different OO bins for VBF process

## 9.2 Uncertainties from signal and background modeling

The Higgs peak described by double-side crystal ball function can be influenced by several experimental effects. Photon energy scale uncertainties shift the position of signal peak by between  $\pm 0.2\%$  and  $\pm 0.4\%$ , while the photon resolution uncertainties change the width of signal peak by between  $\pm 6\%$  and  $\pm 15\%$ , following Ref [ref:phscaleres]. Uncertainty due to knowledge of Higgs boson mass of 0.24GeV [ref:mHerror] would also affect the signal peak position. These sources are considered independently into systematic uncertainties. # **NUMBER NEED CHECK**

The uncertainty due to the background choice is taken to be the spurious signal yield discussed in Section 8.4, and assumed to be uncorrelated in each bins.

### 9.3 Experimental systematic uncertainties affecting the expected signal yields

With data taken during 2015-2018, the uncertainty from integrated luminosity is 2.0%. Other sources of experimental uncertainty affecting the expected signal yields include: the efficiency of diphoton trigger, the photon identification and isolation efficiencies, the photon energy scale and resolution, the modelling of pile-up in the simulation, the jet energy scale and resolution, the efficiency of the jet vertex tagger. Among them the dominant one is jet flavor composition `ATLAS_JET_Flavor_Composition` which contributes 5.5%.

The considered systematic uncertainty terms are shown in Figure 18.

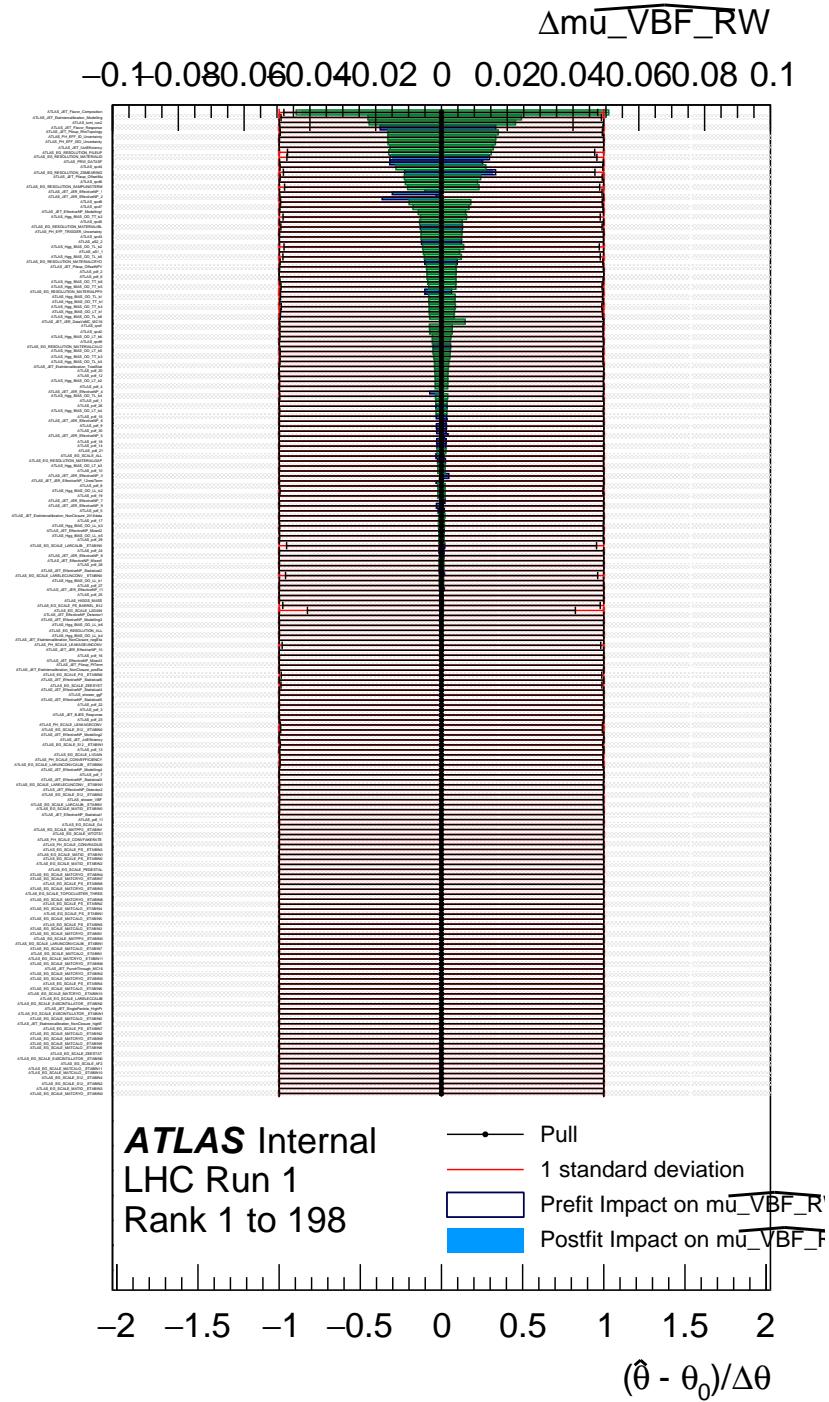


Figure 18: Ranking plot for all systematic uncertainty terms.

## 625 10 Fitting strategy

626 The estimation of CP-mixing parameter  $\tilde{d}$  uses a maximum-likelihood fit performed on  $m_{\gamma\gamma}$  distribution  
 627 simultaneously in all 6 OO bins. The likelihood function could be constructed as:

$$628 \quad \mathcal{L} = \mathcal{L}(\mathbf{x}|\boldsymbol{\theta}) = \prod_{bin} \prod_{j=1}^N f_{bin}(m_{\gamma\gamma})G(\theta)$$

629 Where  $\mathbf{x}$  represents dataset,  $\boldsymbol{\theta}$  is nuisance parameter,  $f_{bin}(m_{\gamma\gamma})$  is the diphoton invariant mass distribution  
 630 model for each bins,  $G(\theta)$  is constrain function for systematic uncertainties. The parameter of interest(POI),  
 631  $\tilde{d}$ , is embedded in  $m_{\gamma\gamma}$  model, so there is no analytic relation between likelihood value and POI. A set of  
 632 signal templates corresponding to different value of CP-mixing parameter  $\tilde{d}$  is created by reweighting the  
 633 SM VBF  $H \rightarrow \gamma\gamma$  to build the CP-mixing models, and then a template fit could be performed with data  
 634 and model with different  $\tilde{d}$  value (background model keeps consistent) to evaluate the likelihood function.  
 635 For the fit in each  $\tilde{d}$  model, VBF signal strength, continuum background yield and parameters describing  
 636 the background model are float, which means the effect of background mis-modelling is considered and  
 637 any constrain from possible model-dependent cross section information is not exploited. Other nuisance  
 638 parameters are fixed to their best-fit values  $\hat{\theta}$ .

639 A negative log-likelihood (NLL) curve can be constructed by calculating the NLL value for each  
 640  $\tilde{d}$  hypothesis. Best-estimated  $\tilde{d}$  as well as its central confidence interval at 68% (95%) confidence  
 641 level (CL) can be determined with the minimum value of NLL,  $NLL_{min}$  and the points at which  
 642  $\Delta NLL = NLL - NLL_{min} = 0.5(1.96)$ . An Asimov dataset is used to get expected sensitivity from this  
 643 method and is shown in Figure 19.

644 Since the Optimal Observable is the main sensitive variable for CP test in this analysis, another 2D  
 645 model for OO and  $m_{\gamma\gamma}$  was considered to construct the likelihood function. In some preliminary study  
 646 this 2D model did not show significant improvement with baseline method, and had some difficulty in  
 647 modelling, it was obsoleted in this analysis. Appendix F includes some results based on this.

	$\mu \pm \text{unc.}(stat.)$	$\tilde{d} \pm \text{unc.}(stat.)(68\% \text{interval})$	95% interval of $\tilde{d}$ (stat.)	95% interval of $\tilde{d}$ (stat.+syst.)
TT	$1.00 \pm 0.27(0.25)$	$0.00^{+0.039(0.039)}_{-0.039(0.038)}$	[-0.092, 0.096]	[-0.094, 0.099]
TL	$1.00 \pm 0.36(0.34)$	$0.00^{+0.062(0.061)}_{-0.062(0.061)}$	[-0.272, 0.254]	[-0.295, 0.273]
LT	$1.00 \pm 0.38(0.35)$	$0.00^{+0.062(0.061)}_{-0.061(0.060)}$	[-0.303, 0.315]	[-0.329, 0.337]
LL	$1.00 \pm 1.12(1.12)$	$0.00^{+1.407(1.197)}_{-1.293(1.148)}$	[-4.924, 5.071]	[-5.551, 6.023]
combined	$1.00 \pm 0.20(0.17)$	$0.00^{+0.027(0.027)}_{-0.027(0.026)}$	[-0.057, 0.059]	[-0.059, 0.060]

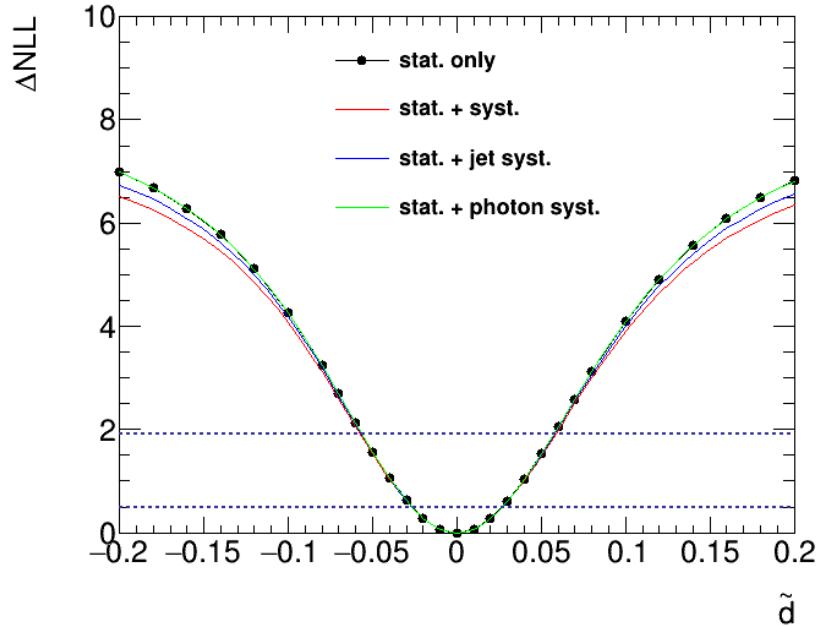
Table 8: Expected results for signal strength and  $\tilde{d}$ .

## 11 Results

A precise measurement of CP-mixing parameter  $\tilde{d}$  is performed with a maximum likelihood fit described in Section 10. The parameter of interest  $\tilde{d}$  is scanned with step of 0.01 and other nuisance parameters are adjusted within their allowed constraints. Table ?? summarized the best fit signal strength and  $\tilde{d}$  with statistical and systematic uncertainty. Expected results from MC is shown in Table 8.

653

The expected and observed  $\Delta NLL$  curve is shown in figure 19 as a function of  $\tilde{d}$ . In the expected result the signal strength is set to 1, and  $\tilde{d}$  is set to 0, corresponding to the best estimate of sensitivity of the analysis based on SM prediction. The contribution from 4 categories are shown with statistical only  $\Delta NLL$  curve in Figure 20.

Figure 19: Expected  $\Delta NLL$  curve for combined result.

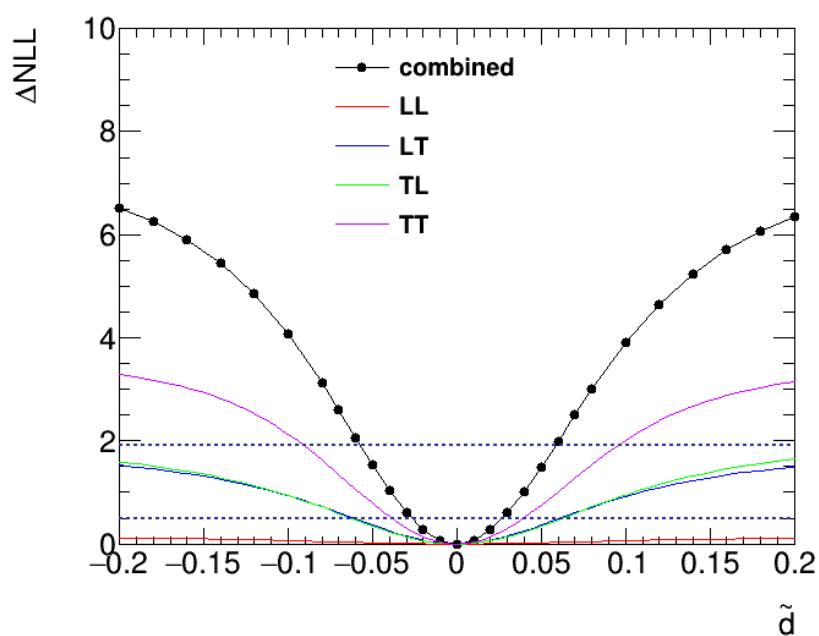


Figure 20: Statistical only  $\Delta NLL$  curve for 4 categories.

**658 12 Conclusion**

659 The CP invariance of Higgs boson coupling to vector bosons has been tested in VBF  $H \rightarrow \gamma\gamma$  process  
660 with  $140fb^{-1}$  13TeV proton-proton collision data collected by ATLAS detector in its full Run2 period.  
661 The measurement is performed in differential Optimal Observable region within SMEFT model, the BSM  
662 Higgs coupling contributing to CP are represented with  $c_{HW}$ , xxx, xxx.

663

664 #Results

---

## 665 References

- 666 [1] G. Aad et al., *Test of CP invariance in vector-boson fusion production of the Higgs boson in the*  
 667  *$H \rightarrow \tau\tau$  channel in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Physics Letters*  
 668 *B* **805** (2020) 135426, ISSN: 0370-2693, URL: <http://dx.doi.org/10.1016/j.physletb.2020.135426> (cit. on p. 4).
- 670 [2] G. Aad et al., *CP Properties of Higgs Boson Interactions with Top Quarks in the  $t\bar{t}H$  and  $tH$  Processes*  
 671 *Using  $H \rightarrow \gamma\gamma$  with the ATLAS Detector*, *Physical Review Letters* **125** (2020), ISSN: 1079-7114, URL:  
 672 <http://dx.doi.org/10.1103/PhysRevLett.125.061802> (cit. on p. 4).
- 673 [3] CMS Collaboration, *Constraints on anomalous HVV couplings from the production of Higgs bosons*  
 674 *decaying to  $\tau$  lepton pairs*, *Phys. Rev.* (2019), arXiv: [1903.06973 \[hep-ex\]](https://arxiv.org/abs/1903.06973) (cit. on p. 4).
- 675 [4] A. M. Sirunyan et al., *Measurements of  $t\bar{t}H$  Production and the CP Structure of the Yukawa Interaction*  
 676 *between the Higgs Boson and Top Quark in the Diphoton Decay Channel*, *Physical Review Letters*  
 677 **125** (2020), ISSN: 1079-7114, URL: <http://dx.doi.org/10.1103/PhysRevLett.125.061801>  
 678 (cit. on p. 4).
- 679 [5] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a*  
 680 *dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022, 2017, URL:  
 681 <https://cds.cern.ch/record/2298955> (cit. on pp. 10, 11).
- 682 [6] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS*  
 683 *Detector*, *Eur. Phys. J. C* **77** (2017) 466, arXiv: [1703.10485 \[hep-ex\]](https://arxiv.org/abs/1703.10485) (cit. on p. 10).
- 684 [7] ATLAS Collaboration, *Measurement of Higgs boson properties in the diphoton decay channel using*  
 685  *$80\text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS-CONF-2018-028,  
 686 2018, URL: <https://cds.cern.ch/record/2628771> (cit. on p. 14).
- 687 [8] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions*  
 688 *at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803,  
 689 arXiv: [1503.07589 \[hep-ex\]](https://arxiv.org/abs/1503.07589) (cit. on p. 19).

690 The supporting notes for the analysis should also contain a list of contributors. This information should  
691 usually be included in `mydocument-metadata.tex`. The list should be printed either here or before the  
692 Table of Contents.

693 **List of contributions**

694

695 **Appendices**

696 **A Statistic only results including signal yields**

697 **B Performance of new BDT based VBF categories**

698 **C Binning optimization**

699 **D Background decomposition with 2x2D sideband method**

700 **E Full results of spurious signal test**

701 **F Statistics results based on 2D model**

702 In an ATLAS note, use the appendices to include all the technical details of your work that are relevant for  
703 the ATLAS Collaboration only (e.g. dataset details, software release used). This information should be  
704 printed after the Bibliography.