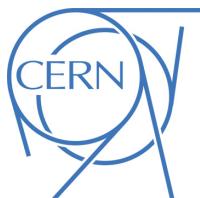




## ATLAS CONF Note

ATLAS-CONF-2019-029

15th July 2019



# Measurements and interpretations of Higgs-boson fiducial cross sections in the diphoton decay channel using $139 \text{ fb}^{-1}$ of $pp$ collision data at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector

The ATLAS Collaboration

Fiducial integrated and differential cross sections for the production of the Higgs boson decaying to two photons are measured using  $139 \text{ fb}^{-1}$  of proton–proton collision data recorded at  $\sqrt{s} = 13 \text{ TeV}$  by the ATLAS experiment at the Large Hadron Collider. The inclusive production cross section in a fiducial region closely matching the experimental selection of the photons is measured to be  $65.2 \pm 7.1 \text{ fb}$ , which is in good agreement with the Standard Model prediction of  $63.6 \pm 3.3 \text{ fb}$ . Differential measurements are performed for a set of variables that are related to the diphoton kinematics as well as the kinematics and multiplicity of the jets produced in association with the Higgs boson. The measurements are compared to various QCD calculations and are found to be in good agreement with the Standard Model predictions. The measurements are also used to probe the strength and tensor structure of the interactions of the Higgs boson using an effective Lagrangian which introduces additional CP-even and CP-odd interactions. In addition, an interpretation of the transverse momentum distribution of the Higgs boson is performed as an indirect probe of the Yukawa coupling of the Higgs boson to the charm quark. Resulting limits on the strength of anomalous interactions are presented for these two approaches.

# 1 Introduction

This note reports measurements of integrated and differential cross sections of the Higgs boson decaying into two photons using the full Run-2 proton-proton collision data set collected at a centre-of-mass energy  $\sqrt{s} = 13$  TeV by the ATLAS detector during the period 2015–2018. Fiducial integrated and differential cross sections were previously measured with smaller data sets at  $\sqrt{s} = 13$  TeV in the  $H \rightarrow \gamma\gamma$  decay channel by both the ATLAS [1] and the CMS [2] Collaborations, and were found to be consistent with Standard Model (SM) predictions. This note closely follows the procedures of the  $pp \rightarrow H \rightarrow \gamma\gamma$  study of the 2015–2017 ATLAS data [3], aiming to improve the precision of the measurements by benefiting from the larger data set and from improved analysis techniques, such as better photon identification and jet energy calibration, leading to reduced associated systematic uncertainties.

The diphoton cross section is measured in a fiducial region that follows closely the experimental reconstructed signature, defined by two isolated photons with transverse momentum greater than 35% and 25% of the diphoton invariant mass, each with absolute pseudorapidity<sup>1</sup>  $|\eta| < 1.37$  or  $1.52 < |\eta| < 2.37$ . For cross sections of jet-related observables, the jets are required to have transverse momentum  $p_T > 30$  GeV and rapidity  $|y| < 4.4$ .

In this note the following distributions are measured: the transverse momentum and absolute rapidity of the Higgs boson in its diphoton decay,  $p_T^{\gamma\gamma}$  and  $|y_{\gamma\gamma}|$ , the multiplicity of jets associated with the Higgs boson production,  $N_{\text{jets}}$ , the transverse momentum of the highest- $p_T$  (leading) jet,  $p_T^{j_1}$ , the invariant mass of the two leading jets,  $m_{jj}$ , and their azimuthal angular difference,  $\Delta\phi_{jj}$ <sup>2</sup>. No attempt is made to separate the different Higgs boson production modes; the measured cross sections are inclusive in the production.

Inclusive Higgs boson production is dominated by gluon–gluon fusion (ggF), for which the transverse momentum of the Higgs boson (at low transverse momentum) is largely balanced by the emission of soft gluons. Measuring  $p_T^{\gamma\gamma}$  probes the perturbative quantum chromodynamics (QCD) calculation of this production mechanism, which has some sensitivity to the bottom and charm quark Yukawa couplings of the Higgs boson [4]. The distribution at high transverse momentum is sensitive to new heavy particles coupling to the Higgs boson and to the treatment of the top-quark mass in the QCD calculations. The  $|y_{\gamma\gamma}|$  distribution is also sensitive to the modelling of the ggF production mechanism, as well as to the parton distribution functions (PDFs) of the colliding protons. The  $p_T^{j_1}$  distribution probes the perturbative QCD prediction and is sensitive to the relative contributions of the different Higgs production mechanisms. The angular variable  $\Delta\phi_{jj}$  is sensitive to the spin and CP quantum numbers of the Higgs boson. The  $m_{jj}$  distribution is sensitive to the vector boson fusion (VBF) production mechanism. All fiducial differential cross sections are reported with their full statistical and experimental correlations and are compared to multiple predictions, where possible.

The  $H \rightarrow \gamma\gamma$  signal yields are extracted from fits to the diphoton invariant mass spectrum. All the fits are performed under the assumption that the Higgs boson mass is  $125.09 \pm 0.24$  GeV [5]. The cross sections are determined by correcting these yields for experimental effects and by taking into account the integrated luminosity of the data.

<sup>1</sup> ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the  $z$ -axis along the beam pipe. The  $x$ -axis points from the IP to the centre of the LHC ring, and the  $y$ -axis points upward. Cylindrical coordinates  $(r, \phi)$  are used in the transverse plane,  $\phi$  being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle  $\theta$  as  $\eta = -\ln \tan(\theta/2)$ .

<sup>2</sup> The  $\Delta\phi_{jj}$  variable is calculated from the two leading jets in the event as the difference in azimuthal angle between the more forward jet minus that of the more central one.

In addition to the comparisons to the SM predictions, the model-independent measurement of the differential cross section can be used to probe physics beyond the SM via an effective field theory (EFT) approach. In this approach, the SM Lagrangian,  $\mathcal{L}_{\text{SM}}$ , is supplemented by additional dimension-six operators,  $O_i^{(6)}$ , such that in the gauge eigenbasis the Lagrangian is specified by

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} O_i^{(6)}, \quad (1)$$

where the  $c_i$  specify the strength of the new interactions and are known as the *Wilson coefficients*, and  $\Lambda$  is the scale of new physics. Non-zero values of these Wilson coefficients can modify the shapes of kinematic distributions or the overall production cross section of the Higgs boson. The cross-section measurements as function of  $p_{\text{T}}^{\gamma\gamma}$ ,  $N_{\text{jets}}$ ,  $m_{jj}$ ,  $p_{\text{T}}^{j_1}$  and  $\Delta\phi_{jj}$  and their correlations are used to set constraints on the Wilson coefficients of the operators of the Higgs Effective Lagrangian in the Strongly Interacting Light Higgs basis (SILH) [6] and the operators of the SMEFT Lagrangian [7] in the Warsaw basis (SMEFT) [8].

In addition to an interpretation in terms of an EFT, the differential cross section as a function of  $p_{\text{T}}^{\gamma\gamma}$  can be used to extract information about the Yukawa coupling of the Higgs boson to the charm quark. The coupling of the charm quark to the Higgs boson has a direct effect on ggF and quark-initiated Higgs boson production ( $q\bar{q} \rightarrow H$ ). Based on the expected effect on the shape of the inclusive  $p_{\text{T}}^{\gamma\gamma}$  differential distribution from ggF and  $c\bar{c} \rightarrow H$ , an analysis is performed to set limits on a potential modification of the Yukawa coupling to the charm quark compared to that predicted in the SM, given the measured differential cross section.

## 2 The ATLAS detector

The ATLAS detector [9, 10] is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry and almost  $4\pi$  coverage in solid angle. The inner tracking detector (ID) covers  $|\eta| < 2.5$  in pseudorapidity and consists of a silicon pixel detector, a silicon microstrip detector, and a transition radiation tracker. The ID is surrounded by a superconducting solenoid and a hermetic calorimeter system, which provides three-dimensional reconstruction of particle showers up to  $|\eta| = 4.9$ . The electromagnetic (EM) calorimeter is a lead/liquid-argon (LAr) sampling calorimeter, measuring electromagnetic showers in the barrel ( $|\eta| < 1.475$ ) and endcap ( $1.375 < |\eta| < 3.2$ ) regions. The hadronic calorimeter reconstructs hadronic showers using steel and scintillator tiles ( $|\eta| < 1.7$ ), copper/LAr ( $1.5 < |\eta| < 3.2$ ), or copper–tungsten/LAr ( $3.1 < |\eta| < 4.9$ ). ATLAS data-taking uses a two-level trigger system [11] with a hardware-based level to reduce the event rate to at most 100 kHz, and a software-based level to reduce the event rate to approximately 1 kHz for storage.

## 3 Data set and event simulation

The data set of  $\sqrt{s} = 13$  TeV proton-proton collisions was recorded by the ATLAS detector from 2015 to 2018 with a proton bunch spacing of 25 ns. The mean number of interactions per bunch crossing,  $\mu$ , was 34 on average, varying from 24 in 2015–2016 to 37 in 2017–2018 data. Events were selected by a diphoton trigger with  $p_{\text{T}}$  thresholds of 35 GeV and 25 GeV for the leading and subleading photon candidates, respectively. Loose photon identification requirements [12] were applied by this trigger in 2015–2016 and were tightened in 2017 to cope with a higher instantaneous luminosity. On average, the

trigger has an efficiency greater than 98% for  $H \rightarrow \gamma\gamma$  events that pass the diphoton event selection described in Section 4. After data quality requirements are applied to ensure good working condition of all detector components, the data set amounts to an integrated luminosity of  $(139.0 \pm 2.4) \text{ fb}^{-1}$  [13].

Event samples are produced with Monte Carlo (MC) simulation to study the modelling of the signal and background processes. The  $H \rightarrow \gamma\gamma$  signal simulation is used to estimate the mass shape and efficiency correction factors. The background simulation is used to choose a functional form to fit the data and for the estimation of systematic uncertainties in the signal extraction, see Section 5.

Table 1: Monte Carlo signal samples used in the analysis. The order listed for the QCD and EW calculations refers to the order to which the samples are normalised.

Process	Generator	Cross-section normalisation	$\sigma \times \text{BR}[\text{fb}]$
ggF	POWHEG NNLOPS	$\text{N}^3\text{LO}(\text{QCD})+\text{NLO}(\text{EW})$	110
VBF	POWHEG-BOX	approx. $\text{NNLO}(\text{QCD})+\text{NLO}(\text{EW})$	8.58
$W^+H$	POWHEG-BOX	$\text{NNLO}(\text{QCD})+\text{NLO}(\text{EW})$	1.90
$W^-H$	POWHEG-BOX	$\text{NNLO}(\text{QCD})+\text{NLO}(\text{EW})$	1.21
$q\bar{q} \rightarrow ZH$	POWHEG-BOX	$\text{NNLO}(\text{QCD})+\text{NLO}(\text{EW})$	1.73
$gg \rightarrow ZH$	POWHEG-BOX	$\text{NLO}(\text{QCD})+\text{NLO}(\text{EW})$	0.28
$t\bar{t}H$	POWHEG-BOX	$\text{NLO}(\text{QCD})+\text{NLO}(\text{EW})$	1.15
$b\bar{b}H$	POWHEG-BOX	5FS (NNLO), 4FS (NLO)	1.10

The signal samples used are listed in Table 1. The mass of the Higgs boson was set in the simulations to  $m_H = 125 \text{ GeV}$  with a width of  $\Gamma_H = 4.07 \text{ MeV}$  [14]. The samples are normalised to the latest available theoretical calculations of the corresponding SM production cross sections, summarised in Ref. [15], and the  $H \rightarrow \gamma\gamma$  branching ratio of 0.227% calculated with HDECAY [16–19] and PROPHECY4F [20–22] for a Higgs boson mass of 125.09 GeV.

The ggF sample is generated with NNLOPS [23–30] and achieves an accuracy in the description of the Higgs  $p_T$  and rapidity distribution that is compatible with that of next-to-next-to-leading order (NNLO) in QCD. It is normalised to a calculation at  $\text{N}^3\text{LO}$  (QCD) with additional NLO electroweak (EW) corrections [31–42], accounting also for the treatment of the quark mass effects, through a correction factor for the total cross section. The PDF4LHC15 NNLO PDF set [43] and the AZNLO tune [44] of PYTHIA8 [45, 46] are used.

The VBF sample is generated with Powheg [23–25, 47] and interfaced with PYTHIA8 for parton shower and non-perturbative effects. Powheg is accurate to NLO and tuned to match calculations with effects due to finite heavy-quark masses and soft-gluon resummations up to NNLL. The VBF sample is normalised to an approximate NNLO (QCD) calculation with NLO EW corrections [48–50]. The PDF4LHC15 PDF set and the AZNLO tune of PYTHIA8 are used.

Samples with a Higgs boson produced in association with a vector boson ( $WH$  or  $ZH$ , collectively referred to as  $VH$ ) are generated with Powheg-Box [51, 52]. Separate  $W^+H$  and  $W^-H$  samples are produced. The  $q\bar{q} \rightarrow ZH$  and  $gg \rightarrow ZH$  production modes are also produced separately. The samples are generated with NLO matrix elements matched to the parton shower, except for the  $gg \rightarrow ZH$ , which is generated at LO. All  $VH$  samples are produced using the PDF4LHC15 PDF set and the AZNLO tune of PYTHIA8, and are normalised to cross sections calculated at NNLO in QCD with NLO electroweak corrections for  $q\bar{q} \rightarrow VH$  and at NLO and next-to-leading-logarithm accuracy in QCD for  $gg \rightarrow ZH$  [53–60].

The sample of events for a Higgs boson produced in association with two top quarks ( $t\bar{t}H$ ) is generated using Powheg+PYTHIA8 [61]. The  $t\bar{t}H$  sample uses the PDF4LHC15 PDF set and the A14 generator

tune [62]. The cross section for its normalisation is taken from a calculation accurate to NLO in QCD with NLO electroweak corrections [63–66].

The  $b\bar{b}H$  sample is generated with PowHEG-Box [67] and contains additional NLO electroweak corrections, accounting for the treatment of the quark-mass effects. The PDF4LHC15 PDF set and the A14 NNPDF23LO generator tune of PowHEG+PYTHIA8 are used.

Background events from continuum  $\gamma\gamma$  production, neglecting any interference effects with the  $H \rightarrow \gamma\gamma$  signal, are generated using SHERPA 2.2.4 [68, 69] and merged with the SHERPA parton shower [70] according to the ME+PS@NLO prescription [71–74]. The PDF set used is the NNPDF3.0 NNLO [75].

The effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) is modelled by overlaying simulated events of inelastic  $pp$  interactions generated with PYTHIA8 using the NNPDF2.3LO set of PDFs [76] and the A3 tune [77] over the original hard-scattering event. Events in the MC simulation are weighted to reproduce the distribution of  $\mu$  observed in data.

The generated particles are passed through the ATLAS detector simulation [78] within the GEANT4 framework [79] and reconstructed with the same software as the data. The  $\gamma\gamma$  background sample is processed with a fast simulation in which the calorimeter response is simulated with a parametrisation [80]. Simulated events are corrected to reflect the photon and jet momentum scales and resolutions as well as the trigger, identification and isolation efficiencies measured in data.

Additional MC samples and theory computations are produced to be used as predictions in the comparison with the cross-section measurements and also for the purposes of the EFT and Yukawa-coupling interpretations. These samples are produced only at particle-level and details will be given in the corresponding sections.

## 4 Event reconstruction and selection

Collision events are reconstructed in the ATLAS detector by a series of algorithms. Objects reconstructed as photons, electrons and jets are utilised in this analysis. A summary of the  $e/\gamma$  reconstruction and performance and of the event selections follows, while detailed descriptions can be found in Ref. [12, 81] and Ref. [1], respectively.

### 4.1 Photon reconstruction and identification

Photons are reconstructed from clusters of cells in the EM calorimeter formed with a dynamical, topological cell clustering algorithm [81] and classified based upon requirements on the EM shower shape variables and matching tracks in the inner detector [12]. The photon candidate energy is calibrated using the procedure described in Ref. [82], optimised for 13 TeV data. Reconstructed photons must satisfy  $|\eta| < 2.37$  in order to fall inside the region of the EM calorimeter with a finely segmented first layer, and outside the range  $1.37 < |\eta| < 1.52$  corresponding to the transition region of the barrel and endcap EM calorimeters. Then shower-shape identification criteria [12] are used to suppress hadrons and photons from hadron decays (mostly  $\pi^0 \rightarrow \gamma\gamma$  in jets). Compared to previous ATLAS measurements which used  $p_T$ -independent photon identification criteria, the requirements are now optimised in sub-ranges of the photon's transverse momentum, as the EM shower depends significantly on the photon energy. The new selection provides an increase in efficiency at low  $p_T^\gamma$  and better background rejection at high  $p_T^\gamma$ , which is expected to improve

the cross-section measurements by a few percent. The total reconstruction and identification efficiency for a photon of  $p_T > 25$  GeV is greater than 82%.

To further suppress jets misidentified as photons, calorimeter- and track-based isolation requirements are applied. The calorimeter isolation variable is defined as the sum of the transverse energies of positive-energy topological clusters in the calorimeter within a cone of  $\Delta r = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.2$  around the photon candidate, subtracting the contribution of the photon candidate and also of the underlying event and pile-up [83]. The calorimeter-based isolation must be less than 6.5% of the photon transverse energy for each photon candidate. The track-based isolation variable is defined as the scalar sum of the transverse momenta of tracks within a cone of  $\Delta r = 0.2$  around the photon candidate. The tracks considered in the isolation variable are restricted to those with  $p_T > 1$  GeV that are associated to the primary vertex selected as explained in the next section. Both photons must have a track isolation less than 5% of the photon transverse energy.

## 4.2 Event and diphoton vertex selection

From all the reconstructed photon candidates in each event passing loose identification criteria and  $p_T > 25$  GeV, the two leading in  $p_T$  are selected as the potential products of a  $H \rightarrow \gamma\gamma$  decay. The diphoton production vertex is identified using a dedicated neural network that exploits the information of the two photon candidates from photon pointing or from conversion vertices and the reconstructed primary-vertex candidates of the event. The neural network is trained using simulated ggF  $H \rightarrow \gamma\gamma$  events to choose the vertex corresponding to the hard scatter event in which the Higgs boson is produced. According to simulation, the algorithm has a 76% chance of selecting the true vertex or a compatible one, in terms of position resolution [84], that is within a distance of 0.3 mm from the true vertex. Its performance was validated using  $Z \rightarrow ee$  events in data and simulation, ignoring the track information of the electrons and treating them as photon candidates. The selected vertex is used to correct the directional part of the four-momenta of the two photons so that they point to that vertex, improving the  $m_{\gamma\gamma}$  resolution. At this step, the leading and subleading photon candidates are required to satisfy  $p_T/m_{\gamma\gamma} > 0.35$  and 0.25, respectively. As a final step, the two preselected photon candidates are required to satisfy the tight identification and the isolation criteria described earlier.

The predicted efficiency of the selection criteria for a SM Higgs boson decaying to two photons in the full phase space is 34%. It is of specific interest to maintain a good rejection of reducible background from dijet ( $jj$ ) and photon+jet ( $\gamma j$ ) processes even in conditions with large pile-up. The event selection is found to be resilient to additional activity due to pile-up quantified by  $\mu$ . As demonstrated in Figure 1, the contribution of  $\gamma\gamma$  events to all selected events is reduced by less than 4% from  $\mu = 15$  to  $\mu = 50$ . It is estimated from data with the procedure detailed in Section 5.2.

## 4.3 Reconstruction and selection of hadronic jets

Jets are reconstructed from three-dimensional topological clusters of neighbouring cells in the calorimeters [85] using the anti- $k_t$  algorithm [86] with radius parameter 0.4. Jets are required to have  $|y| < 4.4$  and  $p_T > 30$  GeV. Jets with  $|\eta| < 2.5$  and  $p_T < 120$  GeV originating from pile-up collisions are suppressed using a jet vertex tagger multivariate discriminant [87, 88].

To avoid double-counting reconstructed electron or photon candidates as jets, an overlap removal procedure is performed based on the quantity  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ . First, electrons overlapping with any of the two

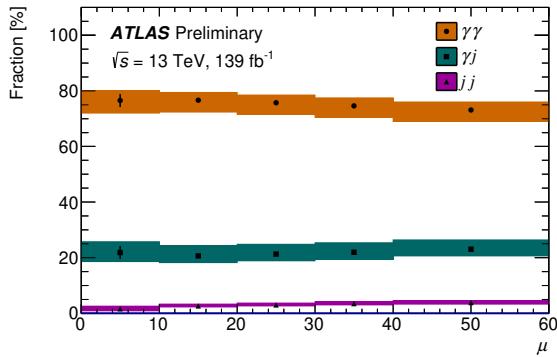


Figure 1: Fraction of  $\gamma\gamma$  events in data measured in bins of  $\mu$  (the mean number of interactions per bunch crossing) using the procedure described in Section 5.2. The fraction of events where one ( $\gamma j$ ) or both photon candidates originate from misidentified jets ( $jj$ ), are also shown. The error bars on the markers represent the statistical uncertainty and the filled rectangles the total uncertainty, which is dominated by systematic sources.

selected photons ( $\Delta R < 0.4$ ) are removed and then jets overlapping with any of the selected photons ( $\Delta R < 0.4$ ) and electrons ( $\Delta R < 0.2$ ) are removed. Electron reconstruction and identification is described in Ref. [89] and the electron candidates used in the overlap removal are required to pass  $p_T > 10$  GeV,  $|\eta| < 2.47$ , medium identification and loose isolation criteria.

## 5 Signal and background modelling of the diphoton mass spectrum

The Higgs boson signal yield is measured with an unbinned maximum-likelihood fit to the diphoton invariant mass spectrum in the inclusive selection and in each bin of a fiducial differential cross section, as further discussed in Section 6. The invariant mass range considered is  $105 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$ , wide enough to allow a reliable determination of the background shape from the data using the sidebands to both sides of the resonant signal and at the same time narrow enough to limit the uncertainties from the choice of the background parametrisation. Apart from the signal yield, the background model parameters are also freely floated in the fit.

### 5.1 Signal model

The signal distribution is empirically modeled as a double-sided Crystal Ball function [90], consisting of a Gaussian central part and power-law tails on both sides. The Gaussian core of the Crystal Ball function is parametrised by the peak position ( $m_H + \Delta\mu_{\text{CB}}$ ) and the width ( $\sigma_{\text{CB}}$ ). The parametric form used can be found in Ref. [1]. The parameters of the model that define the shape of the signal distribution are determined through fits to the simulated signal samples, taking into account all production modes according to their expected contributions. The parametrisation is derived separately for each bin considered in the cross-section measurement. As an example of the signal model, Figure 2 shows the parametrisations in the lowest and the highest  $p_T^{\gamma\gamma}$  bins considered in the measurement. The effective signal mass resolution of the two bins, defined as half the width containing 68% (90%) of the signal events, is 2.1 (3.8) GeV and 1.6 (2.7) GeV, respectively, and varies from low to high  $p_T^{\gamma\gamma}$  due to its dependence on the  $p_T$  of the photons.

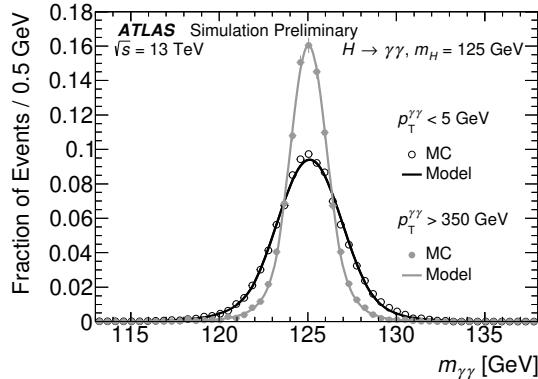


Figure 2: Shown is the signal pdf in the lowest and highest  $p_T^{\gamma\gamma}$  bins considered, demonstrating the variation of the signal  $m_{\gamma\gamma}$  resolution. The signal pdfs (solid curves) are compared against the  $m_{\gamma\gamma}$  distribution of the signal MC simulation (open and filled markers, labelled as MC) for events in the corresponding  $p_T^{\gamma\gamma}$  bins.

## 5.2 Background composition

The background consists of three components, separated according to the origin of the photon candidates. The  $\gamma\gamma$  component comprises the production of prompt and isolated diphotons; it is simulated with the SHERPA event generator as explained in Section 3. The  $\gamma j$  and  $jj$  components denote the cases where one or both photon candidates in the final state originate from hadronic jets misidentified as photons.

The measurement of the fraction of each background component in data is performed for each cross-section bin using a double two-dimensional sideband method [91]. The nominal identification and isolation requirements are loosened for both photon candidates, and the data are split into 16 orthogonal regions defined by diphoton pairs in which one or both photons satisfy or fail identification and/or isolation requirements. The region in which both photons satisfy the nominal identification and isolation requirements corresponds to the nominal selection of Section 4, while the other 15 regions provide control regions, whose  $\gamma\gamma$ ,  $\gamma j$  and  $jj$  yields are related to those in the signal region via the efficiencies for photons and for hadronic jets to satisfy the photon identification and isolation requirements. The  $\gamma\gamma$ ,  $\gamma j$  and  $jj$  yields in the signal region are thus obtained, together with the efficiencies for hadronic jets, by solving a system of equations using as inputs the observed yields in the 16 regions and the photon efficiencies estimated by the MC simulation. The systematic uncertainties in the measured background fractions are due to the definition of the background control regions and dominate the total uncertainty of the measurements. The fractions of these background sources in the inclusive diphoton selection are  $(75 \pm 4)\%$ ,  $(22 \pm 3)\%$  and  $(3 \pm 1)\%$ , respectively. The  $\gamma\gamma$  fraction changes smoothly across the bins of the differential measurements and ranges from 70% to 85%.

## 5.3 Background model

The functional form used to model the background  $m_{\gamma\gamma}$  distribution in the fit to the data is chosen, in each bin, to ensure the biases in the extracted signal yields are small with respect to the expected signal yield, following the procedure described in Ref. [92]. The  $m_{\gamma\gamma}$  distribution in all bins is found to be well modelled by at least one of the following functions: an exponential of a first- or second-order polynomial, a power law, or a Bernstein polynomial of third or fourth order. The potential bias (*spurious signal*) is estimated as

the maximum of the absolute value of the fitted signal yield in successive fits to background templates, using a signal model with resonant mass, scanning the range between 121 and 129 GeV. The background templates are built by adding the  $m_{\gamma\gamma}$  shapes of the  $\gamma\gamma$  and the reducible background components according to the relative fractions measured with data, as described in the previous section. The shape of the  $\gamma\gamma$  component is determined from simulated events passing the full event selection. The shape of the reducible background components is obtained from data in control regions formed by inverting the tight photon identification requirement on any of the two photons in the final state. The total background template is then normalised to match the data entries in the sidebands of the signal peak (*i.e.* excluding the range  $121 \text{ GeV} < m_{\gamma\gamma} < 129 \text{ GeV}$ ).

The size of the simulated background sample limits the precision of the spurious signal estimation. In particular, fluctuations in the template in the range  $121 \text{ GeV} < m_{\gamma\gamma} < 129 \text{ GeV}$  will lead to an overestimate of the associated systematics, as these cannot be distinguished from genuine mismodelling of the background shape. The spurious signal is one of the dominant sources of systematic uncertainty in many of the bins considered in the analysis. Further increasing the size of the samples is difficult due to the computational resources required to generate and simulate a sample that would need to consist of about  $10^9$  events. Therefore, to reduce the impact of the sample size on the estimation of the spurious signal and the corresponding systematic uncertainty, the template is smoothed with a Gaussian Process approach [93]. It was verified with pseudo-experiments that the smoothing procedure maintains the underlying shape of the  $m_{\gamma\gamma}$  distribution hence introducing no significant bias on spurious signal. The achieved reduction of the associated systematic uncertainties after smoothing is 20% on average, with the largest improvements seen in the low-yield bins, and little to no change seen in high-yield bins.

For a function to be considered as a good candidate for the fit to data, the spurious signal is required to be less than 10% of the expected SM signal yield or less than 20% of the expected statistical uncertainty of the SM signal yield. If none from the list of functions considered is found to pass any of these requirements, which is the case for only a few bins considered in the analysis, then the requirements are relaxed by incorporating the MC statistical uncertainty of the background template. In the case when two or more functions satisfy those requirements, the background model with the least number of parameters is chosen. For the inclusive case, from the list of functional forms considered, an exponential or a second-order polynomial is selected for the fit to the data.

Prior to the final fit to the data, an F-test is performed, testing the selected model against a model from the same family of functions but with one more degree of freedom (for instance, a third-order polynomial is tested against a fourth-order one) to check, using only events in the diphoton invariant mass sidebands, if the data favours a more complex model. A test statistic is built from the  $\chi^2$  values and number of degrees of freedom of two binned fits to the data with the two background models. The expected distribution of the test statistic is built from pseudo-data assuming that the function with fewer degrees of freedom is the true underlying model. The value of the test statistic obtained in the data is compared to such distribution, and the simpler model is rejected in favour of the more complex one if the  $p$ -value of such comparison is lower than 5%.

## 5.4 Statistical model

The statistical procedure used to interpret the data is described in Ref. [94] and also, in more detail, in Ref. [95]. For the inclusive fiducial measurement, the likelihood function is built from the diphoton invariant mass distribution of selected events in the range 105–160 GeV using the signal and background

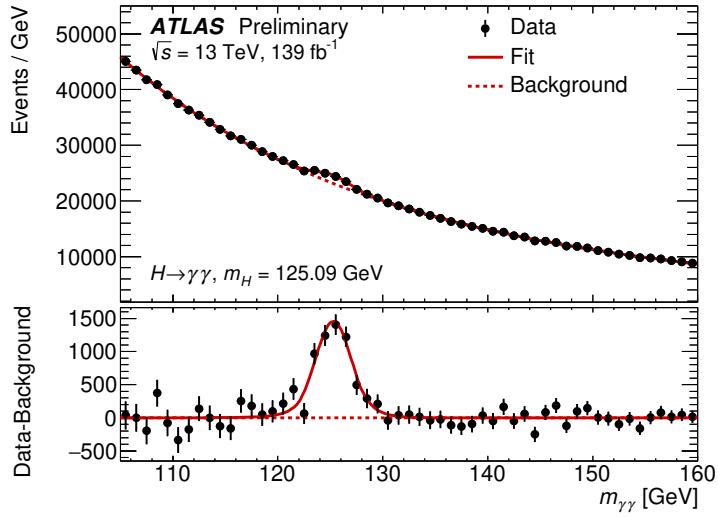


Figure 3: The diphoton invariant mass spectrum of events in the diphoton fiducial region. The solid red curve shows the fitted signal-plus-background model when the Higgs boson mass is constrained to  $125.09 \pm 0.24 \text{ GeV}$ . The background component of the fitted model is an exponential of a second-order polynomial and is shown with the dotted red curve. The bottom plot shows the residuals between the data and the background component of the fitted model.

parametrisations determined for the inclusive phase space. Similarly for each differential measurement, an extended likelihood function is constructed in each bin from the corresponding signal and background parametrisations and the product of the likelihood functions from each bin is taken as the full likelihood function for that measurement. Systematic uncertainties are incorporated into the likelihood function of each bin using a set of Gaussian or log-normal constraints on nuisance parameters. The full list and treatment of systematic uncertainties is discussed in Section 6.3.

Figure 3 illustrates the diphoton invariant mass spectrum of events in the diphoton fiducial region, fitted with the full statistical model. In total, the fit extracts  $6550 \pm 530 H \rightarrow \gamma\gamma$  candidates, with the quoted error including both the statistical and systematic uncertainty components of the fit. The Higgs boson mass is constrained to be  $125.09 \pm 0.24 \text{ GeV}$ .

## 6 Measurement of fiducial integrated and differential cross sections

Measurements of differential and integrated cross sections are presented in fiducial volumes defined at particle level that resemble the ATLAS detector acceptance and analysis selections. The measured cross sections are model independent and are compared with predictions of various QCD calculations.

### 6.1 Particle-level fiducial definition of the $H \rightarrow \gamma\gamma$ cross sections

Stable particles, defined as having a lifetime of  $\tau > 30$  ps, are considered for the particle-level selections. The two highest- $p_T$  photons, not originating from a hadron, that are in the detector acceptance,  $|\eta| < 2.37$  and outside the region  $1.37 < |\eta| < 1.52$ , are considered. They are required to have an invariant mass in the range  $105 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$  and pass a  $p_T^\gamma/m_{\gamma\gamma}$  kinematic threshold that is 0.35 for the leading- $p_T$  photon and 0.25 for the subleading one. A particle-level isolation requirement is also included in the fiducial definition in order to imitate the acceptance of the isolation requirement on reconstructed particles and reduce the dependence of the measurements on assumptions for the jet multiplicity of signal events. Each photon must satisfy  $\sum p_T^i/p_T^\gamma < 0.05$ , where  $\sum p_T^i$  is the scalar sum of transverse momenta of all charged stable particles with  $p_T > 1 \text{ GeV}$  inside a cone  $\Delta r = 0.2$  around the photon. This particle-level isolation requirement has an acceptance that is high for ggF events (about 98%) and becomes significantly smaller for events for high jet multiplicity, as in the case of  $t\bar{t}H$  production (about 90%).

Jets at particle level are constructed from all stable particles, excluding muons and neutrinos, using the anti- $k_t$  algorithm with a 0.4 radius parameter. Particle-level jets must have  $p_T > 30 \text{ GeV}$ ,  $|y| < 4.4$ , and must be well separated from the two photons ( $\Delta R_{j,\gamma} > 0.4$ ) and electrons ( $\Delta R_{j,e} > 0.2$ ). Electrons considered in this overlap check must not originate from a hadron, be well separated ( $\Delta R_{e,\gamma} > 0.4$ ) from the two photons and satisfy  $p_T > 10 \text{ GeV}$  and  $|\eta| < 2.47$ , excluding  $1.37 < |\eta| < 1.52$ .

A brief overview of the particle-level selections for the diphoton fiducial and differential cross section measurements described above is given in Table 2.

Table 2: Particle-level selections for the fiducial measurements. The photon isolation,  $\sum p_T^i/p_T^\gamma$ , is defined as the sum of the  $p_T$  of charged particles within  $\Delta r < 0.2$  of the photon.

Objects	Fiducial definition
Photons	$ \eta  < 2.37$ (excluding $1.37 <  \eta  < 1.52$ ), $\sum p_T^i/p_T^\gamma < 0.05$
Jets	anti- $k_t$ , $R = 0.4$ , $p_T > 30 \text{ GeV}$ , $ y  < 4.4$
Diphoton	$N_\gamma \geq 2$ , $105 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$ , $p_T^{\gamma_1}/m_{\gamma\gamma} > 0.35$ , $p_T^{\gamma_2}/m_{\gamma\gamma} > 0.25$

### 6.2 Integrated and differential cross section definitions

The cross section measurement is performed in two steps: first, the number of signal events in data is extracted in each bin using a fit to the diphoton invariant mass spectrum as described in Section 5. In the second step, the reconstruction-level data signal yields are unfolded to particle level using a bin-by-bin

correction factor. The cross section times branching ratio in the fiducial volume,  $\sigma_{\text{fid}}$ , or in a bin  $i$  of variable  $x$ ,  $(d\sigma/dx)_i$ , are given by

$$\sigma_{\text{fid}} = \frac{N^{\text{sig}}}{c_{\text{fid}} \mathcal{L}_{\text{int}}} \quad \text{and} \quad \left( \frac{d\sigma}{dx} \right)_i = \frac{N_i^{\text{sig}}}{c_{\text{fid},i} \Delta x_i \mathcal{L}_{\text{int}}},$$

where  $N^{\text{sig}}$  refers to the signal yield extracted from data,  $c_{\text{fid}}$  is the correction factor for detector efficiency and resolution effects,  $\mathcal{L}_{\text{int}}$  is the integrated luminosity of the data set and  $\Delta x_i$  is the bin width. The correction factors used to unfold the data are determined with the signal MC simulation described in Section 3. They are defined as the ratio of simulated  $H \rightarrow \gamma\gamma$  signal yields reconstructed in a bin of the cross section (or inclusively), including migrations from outside the fiducial region to the bin, over the true yield of signal events in the bin (or inclusively) passing the fiducial selections. As the inefficiency in the photon isolation due to jet activity is already considered in the particle-level selection as discussed in Section 6.2, the experimental correction factor is found to be rather insensitive to the Higgs boson production mode. The correction factor in the full fiducial volume is about 0.71 and varies by up to 1% for the different production modes. Its value across the bins of the differential distributions varies from 0.65 to 0.8.

The bias of the unfolding method was checked in various scenarios assuming non-SM kinematics and signal composition variations within the experimental constraints from previous measurements [96] and was found to be very small compared to the statistical and systematic uncertainties in each bin. This potential unfolding bias is covered by dedicated systematic uncertainties as explained in Section 6.3. An alternative unfolding method, based on the inversion of the full detector response matrix, has also been considered and leads to results that are similar to those of the baseline method but with different statistical uncertainties due to the conceptual differences of the two unfolding methods (see Appendix B).

### 6.3 Systematic uncertainties

Systematic uncertainties that affect the cross section measurements arise from the integrated luminosity of the analysed data set, the signal extraction from the data and various experimental and theoretical sources with impact on the unfolding correction factors.

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [13], obtained using the LUCID-2 detector [97] for the primary luminosity measurements.

Sources of uncertainty in the signal extraction relate to the signal and background model. Photon energy scale and resolution uncertainties are evaluated according to the methodology detailed in Ref. [82]. They affect the signal model by shifting the position of the signal peak by between  $\pm 0.2\%$  and  $\pm 0.4\%$  and changing the width of the signal peak from  $\pm 6\%$  to  $\pm 15\%$ . Another uncertainty affecting the signal peak position is from the knowledge of the Higgs boson mass, and it is considered to be 0.24 GeV [5]. The uncertainty due to the background function choice is taken to be the spurious signal yield estimated with the procedure described in Section 5.3 and assumed to be uncorrelated between bins. Photon energy scale, resolution and  $m_H$  uncertainties are implemented directly into the likelihood function (described in Section 5.4) which is used to fit the data, while the background modelling uncertainty is added quadratically to the resulting fit uncertainty. Scale and  $m_H$  uncertainties have a negligible impact on the measurements, while the resolution and background modelling are among the leading sources of uncertainty. Uncertainties from the physics modelling of the signal have a negligible impact on the signal model and are hence ignored.

The correction factors used to unfold the data are associated with both experimental and theoretical uncertainties. The experimental ones relate to the reconstruction and selection efficiencies and to the modelling of the migrations of events between bins, as well as into and out of the fiducial volume. Sources of experimental uncertainties considered include: the photon identification and isolation efficiencies [12], the photon energy scale and resolution, the efficiency of the diphoton trigger [11], the modelling of pile-up in the simulation, the jet energy scale and resolution [98] and the efficiency of the jet vertex tagger. Among these, the uncertainties with the largest impact are the jet energy scale and resolution uncertainties (6% to 25%) and pile-up modelling (1% to 5%) in the jet-related measurements, followed by photon identification and isolation efficiency uncertainties (about 1% each). The theoretical uncertainties on the correction factors are considerably smaller than the experimental ones and are calculated as in Ref. [1] by taking the envelope of the variations from different sources. The uncertainties considered are from variations in the relative contributions of the different Higgs boson production modes following the constraints of recent measurements [96], variations in the parton shower, underlying event and hadronisation modelling obtained by comparing samples produced with PYTHIA8 and HERWIG7 [99], and the variation due a potential mismodelling of the Higgs boson kinematics, evaluated by reweighting the  $p_T^{\gamma\gamma}$  and  $|y_{\gamma\gamma}|$  distributions to the data measurements. In addition, the contribution of Dalitz decays in the simulation is also varied to assign a small systematic uncertainty on the correction factor. These sources have a per-mille effect on the measured cross sections, except for the signal composition and parton showers which have a larger effect of up to 2% on the measurements of  $p_T^{j_1}$  and dijet observables. Also the uncertainty on potential mismodelling of the Higgs boson kinematics is up to a few % in some bins with a maximum of 3% at low  $p_T^{\gamma\gamma}$ .

Table 3 summarises the systematic uncertainties discussed above and shows their impact on the cross section measurement in the diphoton fiducial region compared to the statistical uncertainty. The uncertainty of the fiducial measurement is equally affected by statistical and systematic uncertainties; when splitting in bins for the differential measurements, the statistical uncertainties dominate, as seen in Figure 4 for the  $p_T^{\gamma\gamma}$  and  $N_{\text{jets}}$  differential measurements. It can also be seen that the systematics associated to the signal extraction (background modelling and photon energy resolution) are typically larger than those on the correction factors, except for measurements with  $N_{\text{jets}} > 1$  where the impact of jet energy scale and resolution uncertainties on the correction factor become equally significant. The breakdown of the systematic uncertainty on the correction factors is shown in Figure 5 for the  $p_T^{\gamma\gamma}$  and  $N_{\text{jets}}$  differential measurements. For the measurements inclusive in jets,  $p_T^{\gamma\gamma}$  and  $|y_{\gamma\gamma}|$ , the photon isolation efficiency, modelling of pile-up in simulation and photon identification efficiency uncertainties contribute the most to the correction factor uncertainty. For the jet-related measurements ( $N_{\text{jets}}$ ,  $p_T^{j_1}$ ,  $m_{jj}$  and  $\Delta\phi_{jj}$ ), the dominant uncertainty is that in the jet energy scale and resolution.

## 6.4 Results and comparisons with theoretical predictions

The measured inclusive fiducial cross section is compared to a SM prediction referenced in the text as the default MC prediction. This prediction includes a ggF component modelled using POWHEG NNLOPS and normalised to the N<sup>3</sup>LO(QCD)+NLO(EW) prediction of Refs. [15, 32, 34, 39, 40, 43] and the diphoton decay branching ratio. The particle-level VBF,  $VH$ ,  $t\bar{t}H$  and  $b\bar{b}H$  contributions (collectively called  $XH$ ) are taken from the simulated samples and normalised using the cross section predictions described in Section 3. A description of the uncertainties assigned on the default MC prediction can be found in Ref. [1].

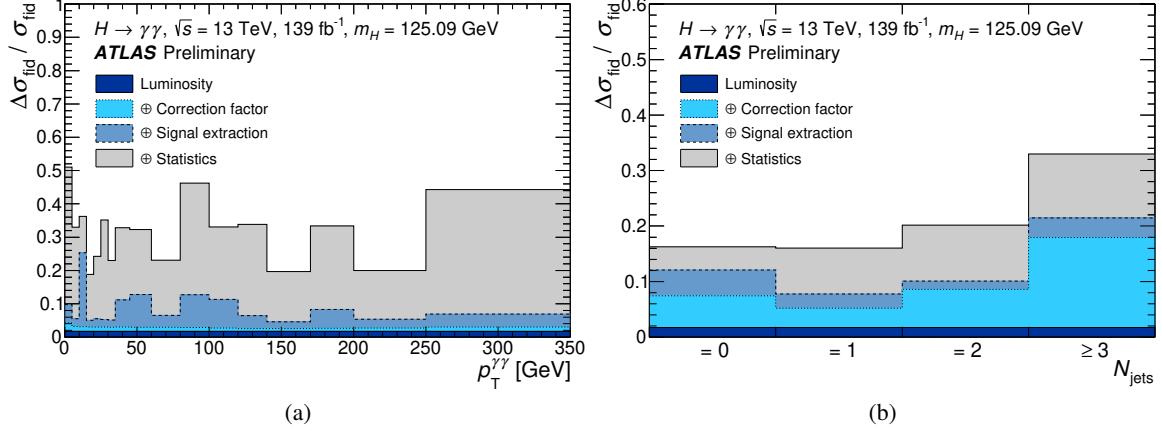


Figure 4: Statistical and systematic uncertainties relative to the differential cross sections measured in data, sequentially summed in quadrature, in each bin of the (a)  $p_T^{\gamma\gamma}$  and (b)  $N_{\text{jets}}$  differential distributions.

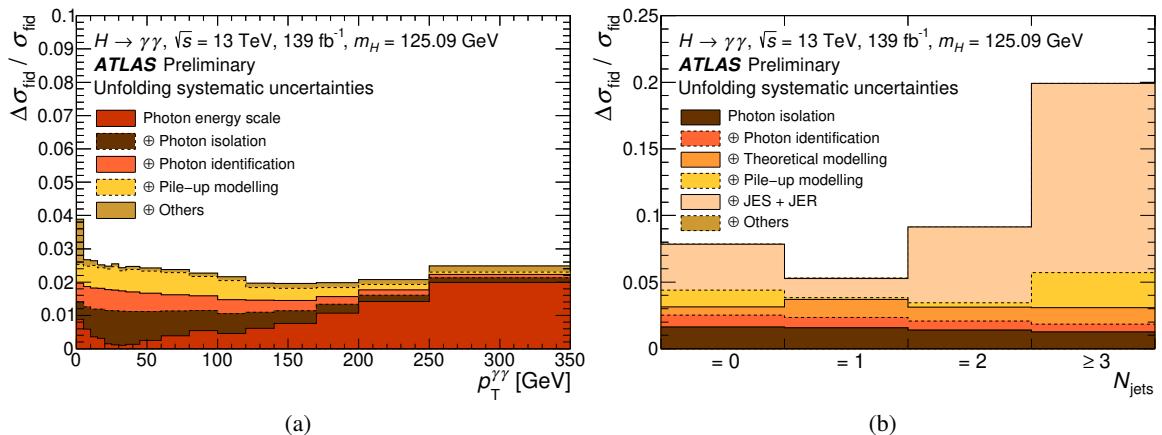


Figure 5: The systematic uncertainties of the unfolding correction factors, sequentially summed in quadrature, in each bin of the (a)  $p_T^{\gamma\gamma}$  and (b)  $N_{\text{jets}}$  differential distributions.

Table 3: The breakdown of uncertainties on the inclusive diphoton fiducial cross section measurement. The uncertainties from the statistics of the data and the systematic sources affecting the signal extraction are shown. The remaining uncertainties are associated with the unfolding correction factor and luminosity.

Source	Uncertainty (%)
Statistics	6.9
Signal extraction syst.	7.9
Photon energy scale & resolution	4.6
Background modelling (spurious signal)	6.4
Correction factor	2.6
Pile-up modelling	2.0
Photon identification efficiency	1.2
Photon isolation efficiency	1.1
Trigger efficiency	0.5
Theoretical modelling	0.5
Photon energy scale & resolution	0.1
Luminosity	1.7
Total	11.0

The inclusive fiducial cross section times the  $H \rightarrow \gamma\gamma$  branching ratio is measured to be

$$\sigma_{\text{fid}} = 65.2 \pm 4.5 \text{ (stat.)} \pm 5.6 \text{ (syst.)} \pm 0.3 \text{ (theo.) fb,}$$

which is within one standard deviation of the default SM prediction of  $63.6 \pm 3.3$  fb [15].

Figure 6 reports the unfolded differential cross section as a function of the diphoton kinematics,  $p_T^{\gamma\gamma}$  and  $|y_{\gamma\gamma}|$ . Figures 7 and 8 show the results for the corresponding one for the jet-related observables,  $N_{\text{jets}}$ ,  $p_T^{j_1}$ ,  $m_{jj}$  and  $\Delta\phi_{jj}$ . The first bin of the  $p_T^{j_1}$  distribution represents events that do not contain a jet passing the corresponding fiducial selections.

The unfolded differential distributions are compared to the default MC prediction for ggF and  $XH$  described earlier and also to additional theory predictions of ggF production, added to the same  $XH$  contributions, described below. All predictions are modified to include the effect of particle-level photon isolation efficiency by applying correction factors obtained from the POWHEG NNLOPS simulation.

The  $p_T^{\gamma\gamma}$  distribution is compared to NNLOJET+SCET [100], which provides predictions using a  $N^3LL$  resummation matched to an NNLO fixed-order calculation in the heavy top-quark mass limit. Corrections are applied for the fiducial selections of the analysis and are obtained from the POWHEG NNLOPS sample. The  $p_T^{\gamma\gamma}$  distribution reaches out to 350 GeV, a region where top-quark mass effects start to become sizeable. The statistical errors for the last bin prevent any conclusive statement about the presence of such effects in the data. The inclusive cross section for  $p_T^{\gamma\gamma} > 350$  GeV is measured to be  $0.23 \pm 0.14$  fb, with the uncertainty being predominantly statistical, and is in good agreement with the default prediction of about 0.21 fb. A finer binning has been chosen at lower  $p_T^{\gamma\gamma}$  to probe the region where resummation effects are important and to probe the charm quark Yukawa coupling.

The  $|y_{\gamma\gamma}|$  distribution is compared to SCETLIB+MCFM8, which provides predictions for  $|y_{\gamma\gamma}|$  at NNLO+NNLL'  $\varphi$  accuracy, derived by applying a resummation of the virtual corrections to the gluon form factor [101, 102]. The underlying NNLO predictions are obtained using MCFM8 with zero-jettiness

subtraction [103, 104]. The diphoton rapidity distribution is sensitive to the gluon distribution. Good agreement is observed over the full rapidity range.

The  $N_{\text{jets}}$  distribution is compared to

- The perturbative  $N^3\text{LO}$  QCD and NLO EW prediction of Refs. [15, 32, 34, 39, 40, 43], shown only for the total fiducial cross section (inclusive zero-jet).
- The perturbative JVE+ $N^3\text{LO}$  prediction of Ref. [105], which includes NNLL resummation in QCD of the  $p_{\text{T}}$  of the leading jet which is matched to the  $N^3\text{LO}$  total cross section, shown only for the inclusive one-jet cross section.
- The perturbative STWZ-BLPTW predictions of Refs. [106, 107], which include NNLL'+NNLO resummation for the  $p_{\text{T}}$  of the leading jet in QCD, combined with a NLL'+NLO resummation in QCD for the subleading jet.<sup>3</sup> The numerical predictions for  $\sqrt{s} = 13$  TeV are taken from Ref. [15]. This prediction is shown for the inclusive zero-, one- and two-jet cross sections as well as for the exclusive zero- and one-jet cross sections.
- The perturbative NNLOJET prediction of Refs. [108, 109] is a fixed-order NNLO prediction in QCD for inclusive  $H + 1$ -jet production. This prediction is shown for the inclusive one-, two-jet, and three-jet cross sections as well as for the exclusive one- and two-jet cross sections.
- The Powheg NNLOPS prediction already described in Section 3, not including its  $N^3\text{LO}$  normalisation factor.
- The perturbative GoSAM prediction of Refs. [52, 110], which provides the fixed-order loop contributions accurate at NLO in QCD in the inclusive  $H +$  zero-jet,  $H +$  one-jet,  $H +$  two-jet, and  $H +$  three-jet regions. The real-emission contributions at fixed order in QCD are provided by SHERPA [68].
- The SHERPA (MEPS@NLO) prediction of Refs. [69, 72, 111, 112] that is accurate to NLO in QCD in the inclusive  $H +$  zero-jet,  $H +$  one-jet,  $H +$  two-jet, and  $H +$  three-jet regions and includes top-quark mass effects.
- The MG5\_aMC@NLO prediction of Refs. [113, 114], which includes up to two jets at NLO accuracy using the FxFx merging scheme [115]. The central merging scale is taken to be 30 GeV. The generated events are passed to PYTHIA8 to provide parton showering and hadronisation to create the full final state (without underlying event). This prediction is shown for all measured inclusive and exclusive jet cross sections.

Good agreement is observed between the measured  $N_{\text{jets}}$  distributions and all predictions with precision better than NLO. The predictions of SHERPA and MG5\_aMC@NLO underestimate the inclusive and zero-jet cross section, but still give a reasonable description of the shape of the measured distributions. An overall scaling to the  $N^3\text{LO}$  prediction would yield good agreement with the data.

The  $p_{\text{T}}^{j_1}$  distribution is compared to:

- The NNLOJET prediction discussed above.
- SCETLIB(STWZ) [102, 106] which provides predictions for  $p_{\text{T}}^{j_1}$  at NNLL'+NNLO<sub>0</sub> accuracy by applying a resummation in  $p_{\text{T}}^{j_1}$ .

---

<sup>3</sup> The prime indicates that the leading contributions from  $N^3\text{LL}$  (resp. NNLL) are included along with the full NNLL (resp. NLL) corrections.

The  $p_T^{j_1}$  distribution covers the same kinematic range as the Higgs boson  $p_T^{\gamma\gamma}$  measurement, but coarser bins were chosen at low  $p_T$ . All predictions agree well with the data, with the NNLOJET prediction providing the greatest accuracy in the high  $p_T^{j_1}$  region.

The  $m_{jj}$  and  $\Delta\phi_{jj}$  distributions are compared to SHERPA (MEPS@NLO) and GoSAM described above, that are of NLO accuracy for this jet multiplicity. Good agreement is seen between the data and the predictions, including that of the default MC that is of LO accuracy for this jet multiplicity. In the higher  $m_{jj}$  bin that is more sensitive to VBF production, the data are in agreement within the prediction within the uncertainty of the measurement. The  $\Delta\phi_{jj}$  distribution that has sensitivity to the CP properties of the Higgs boson is in good agreement with the expected shape in the SM.

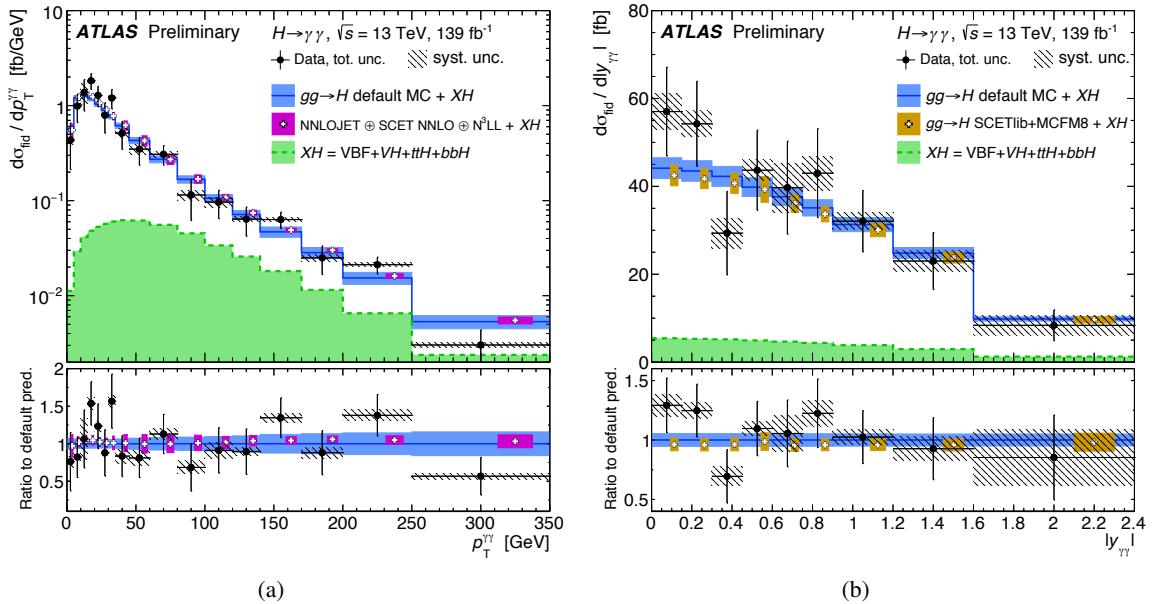


Figure 6: Cross sections measured as a function of the diphoton kinematics, (a)  $p_T^{\gamma\gamma}$ , (b)  $|y_{\gamma\gamma}|$ . The cross section as function of  $p_T^{\gamma\gamma}$  is shown in the range 0–350 GeV, while for  $p_T^{\gamma\gamma} > 350$  GeV it is measured to be  $0.23 \pm 0.14$  fb with the uncertainty being predominantly statistical. All measurements are compared to the default MC prediction in which ggF is modelled with PowHEG NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction, all described in Section 6.4. The measurement for  $p_T^{\gamma\gamma} > 350$  GeV agrees with the default prediction within less than one standard deviation.

The compatibility between the measured differential distributions and the default SM prediction is assessed using a  $\chi^2$  test. The  $\chi^2$  is computed using the covariance matrix constructed from the full set of uncertainties on the data measurements, taking into account correlations between bins, as well as the theory uncertainties on the SM prediction. Table 4 reports the  $p$ -values of the  $\chi^2$  between data and the default MC prediction introduced in Section 6.4 for all differential distributions. For all observables, the compatibility between the data and the SM prediction is good.

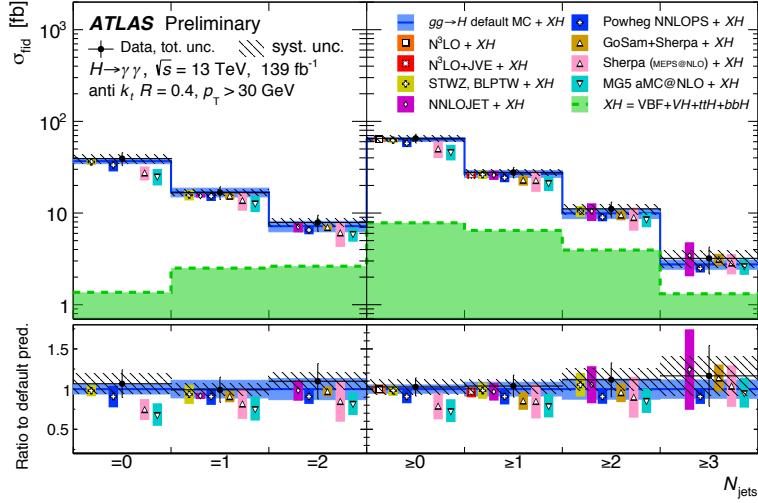


Figure 7: Cross sections measured as a function of the jet multiplicity,  $N_{\text{jets}}$ , in exclusive and inclusive bins. All measurements are compared to the default MC prediction in which ggF is modeled with Powheg NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction.

Table 4: Probabilities from a  $\chi^2$  compatibility test comparing data and the default SM prediction for each differential distribution. The  $\chi^2$  is computed using the covariance matrix constructed from the full set of uncertainties on the data measurements and the theory uncertainties on the SM prediction.

Distribution	$p(\chi^2)$ with Default MC Prediction
$p_T^{\gamma\gamma}$	44%
$ y_{\gamma\gamma} $	68%
$p_T^{J_1}$	77%
$N_{\text{jets}}$	96%
$\Delta\phi_{jj}$	82%
$m_{jj}$	75%

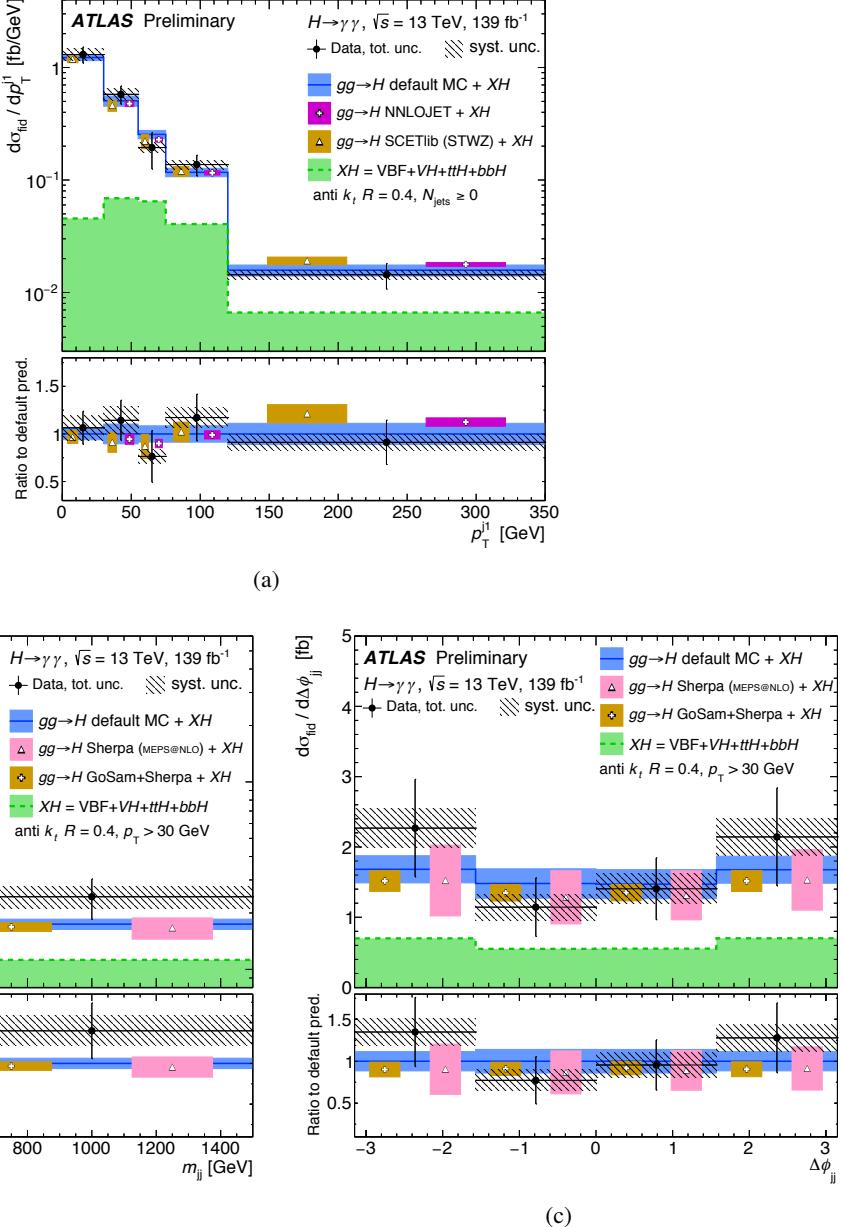


Figure 8: Cross sections measured as a function of jet kinematic observables, (a)  $p_T^{j1}$ , (b)  $m_{jj}$  and (c)  $\Delta\phi_{jj}$ . All measurements are compared to the default MC prediction in which ggF is modeled with Powheg NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction.

## 7 Search for anomalous Higgs-boson interactions using an effective field theory approach

The strength and tensor structure of the Higgs-boson interactions are investigated following an effective field theory (EFT) approach in which additional CP-even and CP-odd interactions can change the event rates, the kinematic properties of the Higgs boson, and associated jet spectra, from those predicted by the SM. Contributions from new physics in the differential cross sections are probed as non-zero values of the Wilson coefficients of the dimension-6 operators of an effective Langrangian [116]. Contributions from dimension-5 and dimension-7 operators are excluded assuming lepton and baryon number conservation. Operators with dimension 8 and higher are neglected as their effects are suppressed by at least  $1/\Lambda^2$  with respect to dimension-6 operators, where  $\Lambda$  is the scale of new physics. From the available bases for parametrising the dimension-6 operators, the SILH basis of the Higgs Effective Lagrangian [6] (referred to as SILH) is employed as well as the Warsaw basis of the SMEFT Lagrangian [7, 8] (referred to as SMEFT).

While new interactions between the Higgs boson and fermions would impact the inclusive production cross-section directly via the ggF mode, the differential  $H \rightarrow \gamma\gamma$  cross sections are also sensitive to operators that affect the Higgs-boson interactions with gauge bosons. In the SILH formulation, the relevant terms in the Lagrangian can be specified by

$$\mathcal{L}_{\text{eff}}^{\text{SILH}} \supset \bar{c}_g O_g + \bar{c}_\gamma O_\gamma + \bar{c}_{HW} O_{HW} + \bar{c}_{HB} O_{HB} + \tilde{c}_g \tilde{O}_g + \tilde{c}_\gamma \tilde{O}_\gamma + \tilde{c}_{HW} \tilde{O}_{HW} + \tilde{c}_{HB} \tilde{O}_{HB},$$

where  $\bar{c}_i$  and  $\tilde{c}_i$  are the dimensionless Wilson coefficients<sup>4</sup> specifying the strength of the new CP-even and CP-odd interactions, respectively, and the dimension-six operators  $O_i$  and  $\tilde{O}_i$  are those described in Refs. [116, 117]. The  $O_g$  ( $O_\gamma$ ) and  $\tilde{O}_g$  ( $\tilde{O}_\gamma$ ) operators introduce new interactions between the Higgs boson and two gluons (photons). The  $O_{HW}$ ,  $\tilde{O}_{HW}$  and  $O_{HB}$ ,  $\tilde{O}_{HB}$  operators introduce new  $HWW$ ,  $HZZ$  and  $HZ\gamma$  interactions and can be probed through VBF and  $VH$  production. Other operators in the full effective Lagrangian of Ref. [116] can also modify Higgs-boson interactions but are not considered here due to the lack of sensitivity of the  $H \rightarrow \gamma\gamma$  decay channel. Combinations of some of the CP-even operators have been constrained using global fits to experimental data from LEP and the LHC [116, 118, 119].

In the SMEFT formulation, a similar parametrisation is employed:

$$\mathcal{L}_{\text{eff}}^{\text{SMEFT}} \supset \bar{C}_{HG} O'_g + \bar{C}_{HW} O'_{HW} + \bar{C}_{HB} O'_{HB} + \bar{C}_{HWB} O'_{HWB} + \tilde{C}_{HG} \tilde{O}'_g + \tilde{C}_{HW} \tilde{O}'_{HW} + \tilde{C}_{HB} \tilde{O}'_{HB} + \tilde{C}_{HWB} \tilde{O}'_{HWB},$$

where all coefficients are dimensionless<sup>5</sup>. The coefficients  $\bar{C}_{HG}$  and  $\tilde{C}_{HG}$  determine the strength of operators that affect the ggF production and  $\bar{C}_{HW}$ ,  $\bar{C}_{HB}$ ,  $\bar{C}_{HWB}$  and their corresponding CP-odd counterparts,  $\tilde{C}_{HW}$ ,  $\tilde{C}_{HB}$ ,  $\tilde{C}_{HWB}$ , are for operators that impact VBF and  $VH$  production and the Higgs boson decay to photons. The operators in the SMEFT basis do not correspond to the same interactions as those in the SILH formulation, despite the similarity in the naming convention.

<sup>4</sup> Using the notation  $\bar{c}_i \equiv (c_i m_W^2)/(\Lambda^2 g)$  (and similarly for CP-odd ones) for the dimensionless coefficients, in the SILH formulation.

<sup>5</sup> Using the notation  $\bar{C}_i \equiv C_i v^2/\Lambda^2$  (and similarly for the CP-odd ones) for the dimensionless coefficients in the SMEFT formulation, where  $v$  is the vacuum expectation value of the Higgs field and  $\Lambda$  is the scale of new physics.

## 7.1 EFT predictions for SILH and SMEFT

The implementation of the effective Langrangian is done in FeynRules [117] for SILH<sup>6</sup>, and within the SMEFTsim package [120] for SMEFT.<sup>7</sup> Both implementations are interfaced to MADGRAPH5 [113] for event generation through the ggF, VBF and  $VH$  production modes with leading-order matrix elements. Other Higgs production modes, i.e.  $t\bar{t}H$  and  $b\bar{b}H$ , are assumed to occur as predicted by the SM, given that the cross sections measured in this analysis are inclusive in production mode and thus do not offer enough sensitivity to them.

The ggF Higgs-boson events are generated in MADGRAPH5 with up to two additional partons in the final state and are merged using the MLM matching scheme [121] to create the full final state. For each production mode, the Higgs boson mass is set to 125 GeV and events are generated using the NNPDF2.3LO PDF set [75] and the A14 parameter set [62]. The parton-level events are then passed to PYTHIA8 for parton showering, hadronisation and underlying event simulation. To apply the fiducial selections, calculate the observables and obtain the differential predictions, a RIVET [122] routine is used.

An assumption is made that higher-order QCD and electroweak corrections are the same for leading-order SM predictions and leading-order predictions that contain contributions from new physics. Thus, the leading-order differential ggF, VBF and  $VH$  predictions obtained from MADGRAPH5 for non-zero values of the Wilson coefficients are scaled by the ratio of the higher-order default SM differential predictions used in Section 6.4 over the differential predictions from MADGRAPH5 with Wilson coefficients set to zero.

To obtain cross sections at any given value of the Wilson coefficients, samples are produced for a limited set of values of the coefficients and then a multidimensional interpolation is performed based on the fact that the dependence of the cross-section on a single coefficient is known a priori. Specifically, the contributions to the cross-section can be separated into components for the SM, BSM and SM-BSM interference:

$$\sigma \propto |\mathcal{M}_{\text{EFT}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + |\mathcal{M}_{\text{d6}}|^2 + 2\text{Re}(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{d6}}) , \quad (2)$$

where the first term is the dimension-4 squared matrix element for the SM, the second term is the squared matrix element for dimension-6 EFT expansion that is of the order  $\frac{c_i^2}{\Lambda^4}$  ( $\frac{C_i^2}{\Lambda^4}$ ) and the last term represents the interference between the SM amplitude and the dimension-6 operators in the EFT expansion that of the order  $\frac{c_i}{\Lambda^2}$  ( $\frac{C_i}{\Lambda^2}$ ). For small values of the Wilson coefficients,  $c_i$ , the interference term is the dominant beyond-the-SM contribution to the cross-section. As it will be demonstrated later in this section, this is mostly the case in the SMEFT formulation for which the interference term can be studied separately. The predictions for SILH include both the interference and the squared dimension-6 term. Samples of 100,000 events are generated for each production mode for 11 values of each Wilson coefficient and used to derive the parametrisation. To study two coefficients simultaneously, the same procedure is repeated for selected combinations at 25 points in a  $5 \times 5$  grid. The parametrisation is obtained by interpolation between the produced samples with the PROFESSOR method [123].

The model implemented in FeynRules fixes the Higgs boson width to be that of the SM,  $\Gamma_H = 4.07$  MeV [14], while for non-zero values of the Wilson coefficients a change in the partial and total decay width are

---

<sup>6</sup> The implementation in Ref. [117] involves a redefinition of the gauge boson propagators that results in unphysical amplitudes unless certain physical constants are also redefined. The original implementation did not include the redefinition of these physical constants. However, the impact of redefining the physical constants is found to be negligible on the predicted cross sections across the range of Wilson coefficients studied.

<sup>7</sup> Using  $U(3)^5$  flavour symmetry with non-SM CP-violating phases and the  $\alpha$  scheme that uses the  $\alpha_{\text{ew}}$ ,  $m_Z$ ,  $G_F$  input parameters for the electroweak sector.

expected. For the SILH basis, the change in partial widths is determined for each Higgs-boson decay mode using the partial-width calculator in **MADGRAPH5** and normalised to reproduce the SM prediction from **HDECAY** [16]. The cross sections are scaled by  $\Gamma_H/(\Gamma_H + \Delta\Gamma)$ , where  $\Delta\Gamma$  is the change in partial widths due to a specific choice of Wilson coefficient. For the SMEFT basis, the modification of the total and partial width is obtained from Ref. [124].

The ratios of the expected differential cross sections to the SM predictions for some representative values of the Wilson coefficients of the SILH operators are shown in Figure 9. The impact of the  $\bar{c}_g$  and  $\tilde{c}_g$  coefficients is mainly on ggF, giving a large change in the overall cross-section normalisation. The  $\tilde{c}_g$  coefficient also changes the shape of the  $\Delta\phi_{jj}$  distribution, which is expected from consideration of the tensor structure of CP-even and CP-odd interactions [125, 126]. The impact of the  $\bar{c}_{HW}$ ,  $\bar{c}_{HB}$  and their CP-odd counterparts is mainly on VBF+VH production, giving large shape changes in all of the studied distributions. The  $\Delta\phi_{jj}$  distribution is of particular interest as it is known to discriminate between CP-even and CP-odd interactions in VBF production [127].

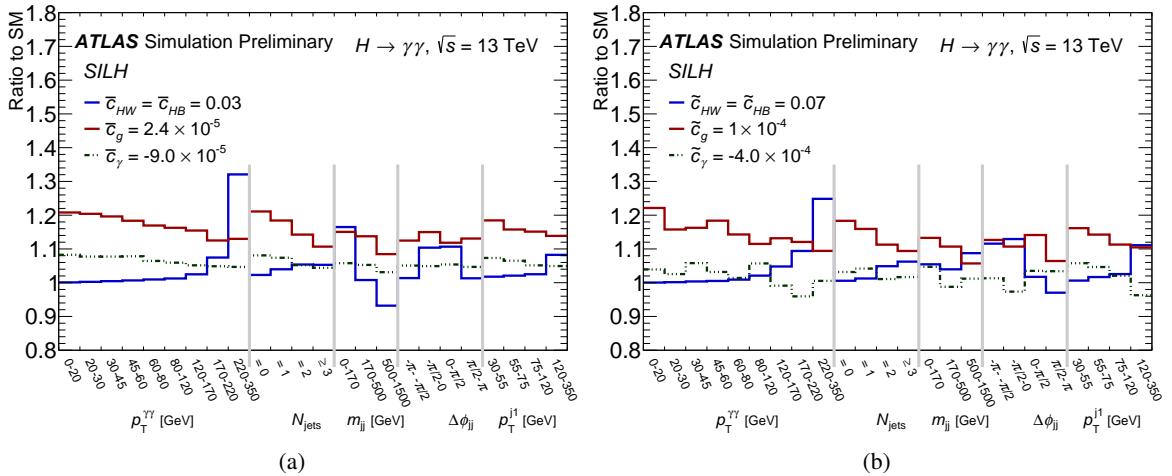


Figure 9: The effect on the five differential distributions used in the analysis of (a) the CP-even coefficients  $\bar{c}_g$ ,  $\tilde{c}_g$  and  $\bar{c}_{HW}$ , and (b) the CP-odd coefficients  $\tilde{c}_g$ ,  $\tilde{c}_y$  and  $\bar{c}_{HW}$  of the SILH effective Lagrangian for values of the coefficients close to the observed limits.

Figure 10 shows the corresponding modifications to the differential cross sections for SMEFT. The  $\bar{C}_{HG}$  and  $\tilde{C}_{HG}$  coefficients affect ggF production while  $\bar{C}_{HB}$ ,  $\bar{C}_{HW}$  and their CP-odd counterparts affect VBF+VH production. The main effect of  $\bar{C}_{HB}$ ,  $\bar{C}_{HW}$  and also of  $\bar{C}_{HWB}$ , however, is on the  $H \rightarrow \gamma\gamma$  decay rate, impacting the overall normalisation. The CP-odd coefficients, as seen in the figure, exhibit sensitivity only to the  $\Delta\phi_{jj}$  observable when only the interference term is considered [128].

## 7.2 Statistical interpretation

Limits on Wilson coefficients are set by constructing a likelihood function,

$$\mathcal{L} = \frac{1}{\sqrt{(2\pi)^k |C|}} \exp\left(-\frac{1}{2} (\vec{\sigma}_{\text{data}} - \vec{\sigma}_{\text{pred}})^T C^{-1} (\vec{\sigma}_{\text{data}} - \vec{\sigma}_{\text{pred}})\right), \quad (3)$$

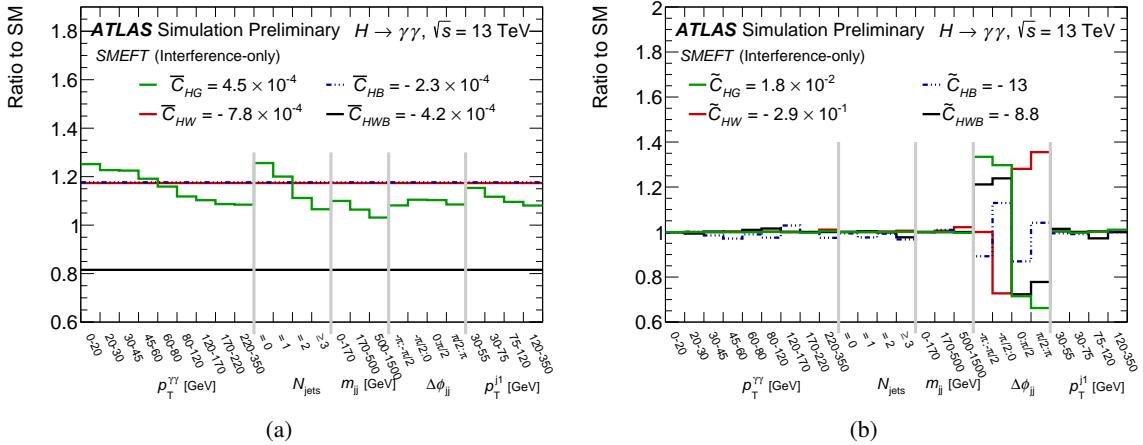


Figure 10: The effect on the five differential distributions used in the analysis of (a) the CP-even coefficients  $\bar{C}_{HG}$ ,  $\bar{C}_{HB}$ ,  $\bar{C}_{HW}$ ,  $\bar{C}_{HWB}$  and (b) the CP-odd coefficients  $\tilde{C}_{HG}$ ,  $\tilde{C}_{HB}$ ,  $\tilde{C}_{HW}$ ,  $\tilde{C}_{HWB}$  of the SMEFT effective Lagrangian for values of the coefficients close to the observed limits.

where  $\vec{\sigma}_{\text{data}}$  and  $\vec{\sigma}_{\text{pred}}$  are  $k$ -dimensional vectors from the measured and predicted differential cross sections of the five analysed observables, with  $k = 32$ ,  $C = C_{\text{stat}} + C_{\text{syst}} + C_{\text{theo}}$  is the  $k \times k$  total covariance matrix defined by the sum of the statistical, systematic and theoretical covariances, and  $|C|$  denotes its determinant.

The statistical covariance matrix is obtained with a bootstrapping technique similar to that described in Ref. [129] and the resulting correlation matrix shown in Figure 11. The matrix provides a measure of the statistical correlations between cross-section bins as the same events in data will populate the different observables used in the fit. Although the correlations refer to the associated uncertainties on the signal yields, they are practically dominated by the statistical fluctuations of the background under the signal peak due to the small signal-to-background ratio. For this reason, the bootstrapping is based on events in the data sidebands as they have the same correlations as events under the signal peak.

The covariance matrices for systematic and theoretical uncertainties are constructed from the uncertainties listed in Section 6.3. Additional theoretical uncertainties are obtained for the ggF, VBF and  $VH$  production modes using the default SM MC simulation to estimate the effect of QCD scale and PDF variations, and are considered to be independent of new physics. Identical sources are assumed to be fully correlated across bins and variables. In what follows, the likelihood function is numerically maximised to determine  $\mathcal{L}_{\text{max}}$  and confidence intervals for one or several Wilson coefficients are determined via

$$1 - \text{CL} = \int_{-2 \ln \mathcal{L}(c_i) + 2 \ln \mathcal{L}_{\text{max}}}^{\infty} dx f(x),$$

with  $\mathcal{L}(c_i)$  denoting the likelihood value evaluated for a given Wilson coefficient value  $c_i$ , and  $f(x)$  denoting the distribution of the test statistic,  $-2 \log(\mathcal{L}(c_i)/\mathcal{L}_{\text{max}})$ . The coverage of 68% and 95% CL limits using the likelihood ratio scan is validated by pseudo-experiments.

In Table 5, the expected and observed 95% CL limits are shown for the Wilson coefficients that are considered for the SILH formulation. The limit for  $\bar{C}_{HW}$  ( $\tilde{C}_{HW}$ ) is obtained after setting  $\bar{C}_{HB} = \bar{C}_{HW}$

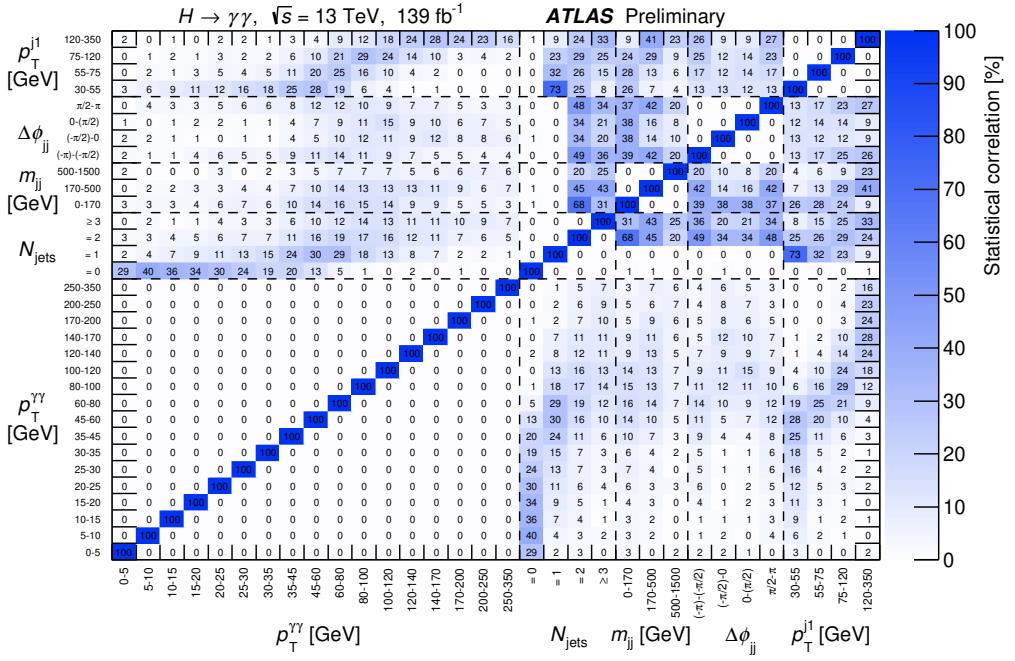


Figure 11: The observed statistical correlations between  $p_T^{YY}$ ,  $N_{\text{jets}}$ ,  $m_{jj}$ ,  $\Delta\phi_{jj}$  and  $p_T^{j1}$  are shown. These correlations are determined from 10,000 bootstrapped replicas of the data using the event counts in the  $m_{\gamma\gamma}$  sidebands.

( $\tilde{c}_{HB} = \tilde{c}_{HW}$ ) to ensure that the partial width for  $H \rightarrow Z\gamma$  is unchanged from the SM prediction<sup>8</sup> and setting all other Wilson coefficients to zero. For the rest, the limits are obtained assuming that all Wilson coefficients other than the one quoted are zero. For the simultaneous fit of  $\bar{c}_{HW}$  and  $\tilde{c}_{HW}$ , for which the 68% and 95% confidence regions are shown in Figure 12(a), the same approach is followed as described earlier. Figure 12(b) shows the corresponding limits from scanning  $\bar{c}_g$  and  $\tilde{c}_g$  simultaneously. Destructive interference causes the ggF production cross section to be zero around  $\bar{c}_g \sim -2.2 \cdot 10^{-4}$  for  $\tilde{c}_g \sim 0$ , which defines the structure seen in the observed limits in the two-dimensional parameter plane.

The corresponding results for the SMEFT coefficients follow. In accordance with the considerations detailed in Ref. [131], the results are provided for two scenarios where predictions are obtained using a) only the interference term and b) both the interference and the quadratic term. The limits on all the coefficients considered are shown in Table 6 for these two approaches. Since the interference terms dominate the predicted cross sections, the limits in the two approaches are very similar for coefficients of CP-even operators. Significant differences emerge for the CP-odd ones for which the interference term is vanishing. Nevertheless, neglecting the quadratic term can be considered as more conservative regarding the validity of the EFT and its dimension-6 truncation. Moreover, this approach simplifies significantly the interpolation procedure, as the cross-section has a linear dependence on the value of the Wilson coefficients. The limits on the Wilson coefficients obtained with the linear interpolation with the interference term only are shown in Figure 13 at both 68% and 95% CL.

<sup>8</sup> Values of  $|\bar{c}_{HW} - \bar{c}_{HB}| > 0.03$  lead to a very large decay rate for the  $H \rightarrow Z\gamma$  process that is contradicted by the experimental constraints reported by ATLAS [130].

Table 5: Observed allowed ranges at 95% CL for the  $\bar{c}_g$ ,  $\bar{c}_{HW}$ ,  $\bar{c}_\gamma$  Wilson coefficients of the SILH basis and their CP-conjugates. Limits on a coefficient are obtained by setting all others to zero. Limits on  $\bar{c}_{HW}$  and  $\tilde{c}_{HW}$  are derived by setting  $\bar{c}_{HB} = \bar{c}_{HW}$  and  $\tilde{c}_{HB} = \tilde{c}_{HW}$ , respectively, with remaining coefficients set to zero.

Coefficient	Observed 95% CL limit	Expected 95% CL limit
$\bar{c}_g$	$[-0.26, 0.26] \times 10^{-4}$	$[-0.25, 0.25] \cup [-4.7, -4.3] \times 10^{-4}$
$\tilde{c}_g$	$[-1.3, 1.1] \times 10^{-4}$	$[-1.1, 1.1] \times 10^{-4}$
$\bar{c}_{HW}$	$[-2.5, 2.2] \times 10^{-2}$	$[-3.0, 3.0] \times 10^{-2}$
$\tilde{c}_{HW}$	$[-6.5, 6.3] \times 10^{-2}$	$[-7.0, 7.0] \times 10^{-2}$
$\bar{c}_\gamma$	$[-1.1, 1.1] \times 10^{-4}$	$[-1.0, 1.2] \times 10^{-4}$
$\tilde{c}_\gamma$	$[-2.8, 4.3] \times 10^{-4}$	$[-2.9, 3.8] \times 10^{-4}$

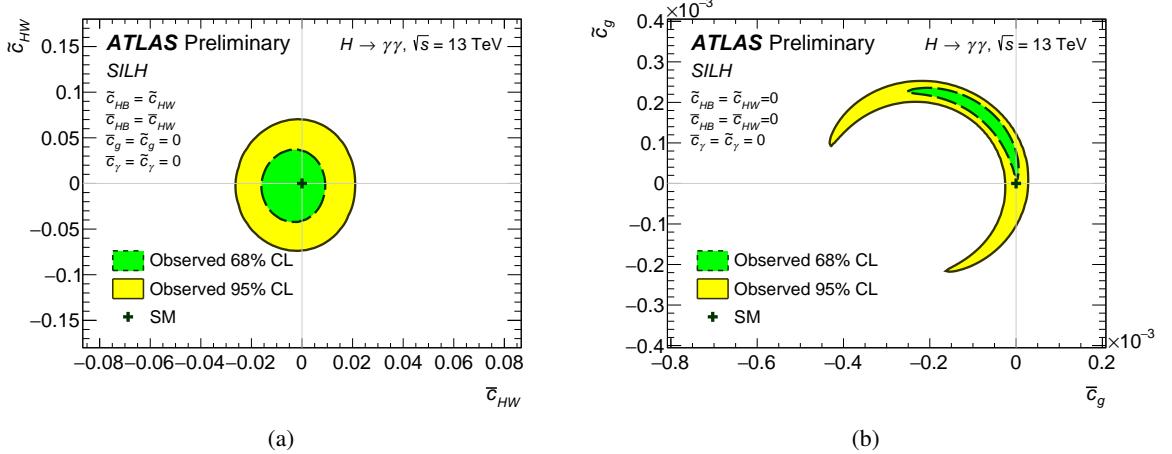


Figure 12: The observed 68% (green) and 95% (yellow) confidence level regions from the simultaneous fit to the  $\bar{c}_{HW}$  and  $\tilde{c}_{HW}$  Wilson coefficients of the SILH basis are shown in the left plot (a). The values of  $\bar{c}_{HB}$  and  $\tilde{c}_{HB}$  are set to be equal to  $\bar{c}_{HW}$  and  $\tilde{c}_{HW}$ , respectively, and all other Wilson coefficients are set to zero. The SM expectation at  $(0, 0)$  is also shown. The corresponding limits from the simultaneous fit to  $\bar{c}_g$  and  $\tilde{c}_g$ , setting all other Wilson coefficients to zero, are shown in the right plot (b).

Table 6: The 95% CL observed limits on the  $\bar{C}_{HG}$ ,  $\bar{C}_{HW}$ ,  $\bar{C}_{HB}$ ,  $\bar{C}_{HWB}$  Wilson coefficients of the SMEFT basis and their CP-odd counterparts using interference-only terms and using both interference and quadratic terms. Limits are derived fitting one Wilson coefficient at a time while setting the other coefficients to zero.

Coefficient	95% CL, interference-only terms	95% CL, interference and quadratic terms
$\bar{C}_{HG}$	$[-4.2, 4.8] \times 10^{-4}$	$[-6.1, 4.7] \times 10^{-4}$
$\tilde{C}_{HG}$	$[-2.1, 1.6] \times 10^{-2}$	$[-1.5, 1.4] \times 10^{-3}$
$\bar{C}_{HW}$	$[-8, 2, 7.4] \times 10^{-4}$	$[-8.3, 8.3] \times 10^{-4}$
$\tilde{C}_{HW}$	$[-0.26, 0.33]$	$[-3.7, 3.7] \times 10^{-3}$
$\bar{C}_{HB}$	$[-2.4, 2.3] \times 10^{-4}$	$[-2.4, 2.4] \times 10^{-4}$
$\tilde{C}_{HB}$	$[-13.0, 14.0]$	$[-1.2, 1.1] \times 10^{-3}$
$\bar{C}_{HWB}$	$[-4.0, 4.4] \times 10^{-4}$	$[-4.2, 4.2] \times 10^{-4}$
$\tilde{C}_{HWB}$	$[-11.1, 6.5]$	$[-2.0, 2.0] \times 10^{-3}$

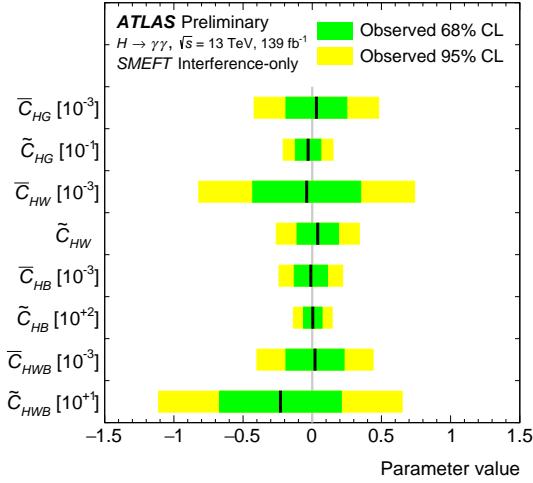


Figure 13: Observed 68% and 95% CL limits on SMEFT Wilson coefficients. Limits are derived fitting one Wilson coefficient at a time while setting the other coefficients to zero.

## 8 Limits on the $c$ -quark Yukawa coupling from the interpretation of the Higgs boson $p_T^{\gamma\gamma}$ spectrum

The Higgs boson  $p_T^{\gamma\gamma}$  spectrum is sensitive to the Yukawa couplings of the Higgs boson to the  $c$  and  $b$  quarks, both through quark-initiated ( $q\bar{q}$  and  $qg$ ) production of the Higgs boson and the contributions of  $c$  and  $b$  quarks to the loop-induced ggF production of the Higgs boson. Direct observations of the Higgs-boson coupling to  $b$  quarks [132, 133] provide stringent constraints on its possible modification with respect to the SM, while current searches for Higgs boson decays to charm final states [134, 135] still allow for a relatively large modification of the  $c$  quark coupling. The focus of this analysis is on the  $c$  quark coupling, given that the sensitivity that can be achieved for the  $b$  quark coupling is not competitive with the direct observations.

A modification in the coupling strength would impact the ggF and quark-initiated production and affect both the normalisation and the shape of the  $p_T^{\gamma\gamma}$  spectrum. The branching ratio for the  $H \rightarrow \gamma\gamma$  decay would also be affected. In this work, only the shape of the measured  $p_T^{\gamma\gamma}$  spectrum is used to set limits on the coupling modifier  $\kappa_c$  of the charm Yukawa coupling relative to that predicted by the SM, in order to be model-independent regarding a possible modification to the branching ratio.

The predictions for the modifications of ggF production are computed with RADISH at NNLL+NLO accuracy for the Higgs-boson  $p_T$  distribution [136, 137] using the PDF4LHC15 PDF set. The renormalisation, factorisation and resummation scales are chosen to be  $m_H/2$ . The assumption is made that higher order QCD corrections to the ggF cross section can be factorised from physics modifying the charm Yukawa coupling. Thus, the predictions for any  $\kappa_c$  value can be scaled by the ratio of the nominal N<sup>3</sup>LO ggF cross section used in this note over that provided by the RADISH prediction. The RADISH predictions are given for the total phase space. To obtain predictions in the fiducial phase space, bin-dependent correction factors derived using the PowHEG NNLOPS simulation have been applied. The fiducial corrections vary with the Higgs boson  $p_T$  from approximately 0.5 to 0.65. Variations of the charm Yukawa coupling affect the Higgs boson  $p_T$  distribution directly ( $\kappa_c^2$ ), and through interference effects with other heavy quarks in the loop,

with the latter being more significant.

Predictions for quark-initiated  $c\bar{c} \rightarrow H$  production are computed by generating events with **MADGRAPH5** at NLO, including the higher-order contribution  $cg \rightarrow Hc$ , in the five-flavour scheme using the PDF4LHC15 PDF. The renormalisation and factorisation scales are set to  $m_H/2$  and  $m_H/4$ . **PYTHIA8** with the A14 tune is used for the simulation of parton showers, hadronisation and underlying event, as well as the Higgs-boson decay. The **MADGRAPH5** prediction is scaled to the cross-section prediction from Ref. [138] times the square of  $\kappa_c$ . The shape of the Higgs-boson  $p_T$  distribution in the  $c\bar{c} \rightarrow H$  prediction is invariant when changing  $\kappa_c$ . However, since the distributions in ggF and  $c\bar{c} \rightarrow H$  productions differ, their sum will effectively have a different shape when increasing or decreasing the  $c\bar{c} \rightarrow H$  contribution. The impact of different values of  $\kappa_c$  is shown in Figure 14 separately for the cross section of the  $c\bar{c} \rightarrow H$  and ggF production modes. The  $b\bar{b} \rightarrow H$  predictions for this interpretation are obtained with **MADGRAPH5+PYTHIA8** in a similar setup as for the  $c\bar{c} \rightarrow H$  prediction, using a dedicated PDF set and the normalisation for the total cross from Ref. [139]. In the predictions, the top and  $b$ -quark Yukawa couplings are assumed to be those predicted in the SM.

For the ggF,  $c\bar{c} \rightarrow H$  and  $b\bar{b} \rightarrow H$  predictions, the theory uncertainties considered are related to missing higher-order QCD corrections, PDFs and the parton shower tune. The perturbative QCD uncertainties on the ggF,  $c\bar{c} \rightarrow H$  and  $b\bar{b} \rightarrow H$  predictions are determined by varying the renormalisation and factorisation scales up and down by a factor of two simultaneously. The effect of up and down variations is symmetrised by taking their quadratic sum and scaling by  $1/\sqrt{2}$ . The resummation scale for ggF is varied as well independently from the renormalisation and factorisation scales. The PDF uncertainty on the  $b\bar{b} \rightarrow H$  predictions is obtained from variations of the  $b$ -quark pole mass,  $m_b$  and the threshold above which the  $b$ -quark PDF is non-zero,  $\mu_q$ , by factors of 0.5 and 2 [139]. The PDF uncertainty on the  $c\bar{c} \rightarrow H$  predictions is obtained from the systematic variations provided by the used set. The uncertainty from the parton shower is evaluated from the systematic variations available for the A14 tune and is considered fully correlated between  $c\bar{c} \rightarrow H$  and  $b\bar{b} \rightarrow H$  predictions. The impact of the  $\kappa_c$  variation on the acceptance as a function of  $p_T^{\gamma\gamma}$  in the relevant range was studied, taking both ggF and  $c\bar{c} \rightarrow H$  components into account, and found to be negligible.

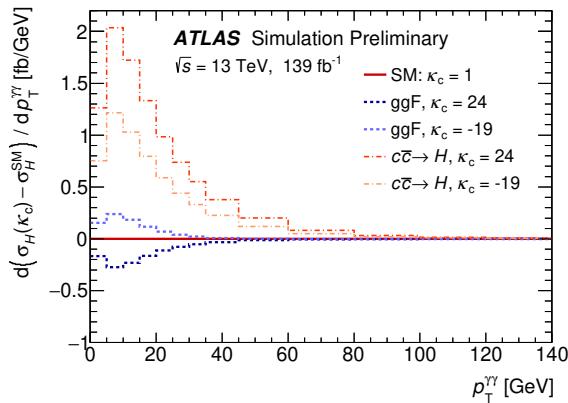


Figure 14: The modification of the  $p_T^{\gamma\gamma}$  differential cross section for different values of  $\kappa_c$ , shown separately for the cross sections of the ggF and  $c\bar{c} \rightarrow H$  production modes. As expected, for a given value of  $\kappa_c$ , the effect on the  $c\bar{c} \rightarrow H$  production cross section is larger than that on the ggF production.

The statistical interpretation of the  $p_T^{\gamma\gamma}$  distribution to set a limit on the value of  $\kappa_c$  is performed with the profile likelihood method. A likelihood similar to that of eq. 3 from Section 7 is built, where  $\vec{\sigma}_{\text{data}}$  and  $\vec{\sigma}_{\text{pred}}$  are the measured and predicted differential cross sections in  $p_T^{\gamma\gamma}$ , and the covariance matrix is constructed

from theoretical and experimental systematic uncertainties. Theory uncertainties are treated differently for those impacting VBF,  $VH$ ,  $t\bar{t}H$  and those affecting the  $c\bar{c} \rightarrow H$ ,  $b\bar{b} \rightarrow H$  and ggF production modes. The per-bin theoretical uncertainties on the ggF,  $c\bar{c} \rightarrow H$  and  $b\bar{b} \rightarrow H$  contributions are incorporated as nuisance parameters with Gaussian constraints into the likelihood function and are correlated between bins. Theory uncertainties on the VBF,  $VH$  and  $t\bar{t}H$  production modes from the QCD scale and PDF are used to construct the covariance matrix together with the experimental systematic uncertainties detailed in Section 6.3.

The differential cross section is used in the range of  $p_T^{\gamma\gamma}$  from zero up to 140 GeV which is the region most sensitive to variations of  $\kappa_c$ . The fit only uses shape information, while the normalisation is treated as a nuisance parameter and is not used in the interpretation. The profile likelihood ratio is shown in Figure 15 as a function of  $\kappa_c$ . The breakdown of uncertainties affecting the limit on  $\kappa_c$  is shown in Table 8 for the 68% CL interval. The observed and expected 95% confidence intervals are shown in Table 7. Figure 16 shows the data compared to predictions for two values of  $\kappa_c$  corresponding to the upper and lower limits at 95% CL.

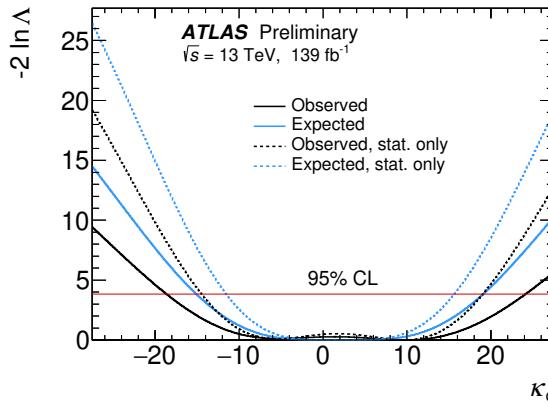


Figure 15: The profile likelihood ratio,  $\lambda$ , shown as a function of  $\kappa_c$  for the fit to the  $p_T^{\gamma\gamma}$  distribution. The intersection of the  $-2 \ln \lambda$  curve with the horizontal line provides the 95% confidence intervals.

Table 7: Observed and expected allowed ranges at 95% CL on modifications of the charm-quark Yukawa coupling to the Higgs boson,  $\kappa_c$ .

Coefficient	Observed 95% CL limit	Expected 95% CL limit
$\kappa_c$	$[-19, 24]$	$[-15, 19]$

The constraints obtained for  $\kappa_c$  with the presented approach are relatively loose, setting limits at 95% CL of  $[-19, 24]$ . The results are comparable to those reported in Ref. [2] following a similar approach of interpreting the Higgs-boson differential cross sections.

Table 8: Effect of statistical, experimental systematic and theoretical uncertainties on  $\kappa_c$  from the profile likelihood fit.

Source	$\delta\kappa_c$ ( $^{+\text{up}}_{-\text{down}}$ )
Stat.	$^{+10.1}_{-8.2}$
Exp. syst.	$^{+3.0}_{-2.7}$
QCD scale (ggF)	$^{+5.4}_{-5.4}$
QCD scale ( $c\bar{c} \rightarrow H$ & $b\bar{b} \rightarrow H$ )	$^{+0.8}_{-0.4}$
PDF (ggF)	$^{+0.5}_{-0.5}$
PDF ( $c\bar{c} \rightarrow H$ & $b\bar{b} \rightarrow H$ )	$^{+0.3}_{-0.1}$
Parton shower ( $c\bar{c} \rightarrow H$ & $b\bar{b} \rightarrow H$ )	$^{+1.4}_{-0.7}$
Total	$^{+12.1}_{-10.3}$

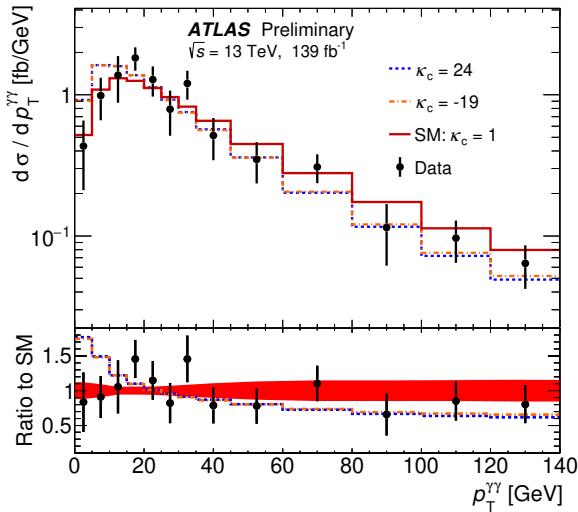


Figure 16: The cross section as function of  $p_T^{\gamma\gamma}$  shown for data against the predictions for different values of  $\kappa_c$  corresponding to the upper and lower limits at 95% CL. The SM prediction is also shown with the red continuous line. The bottom pad shows the ratio of the data and the different predictions to the SM prediction. The filled area corresponds to the shape component of the theoretical uncertainty of the SM prediction.

## 9 Summary and conclusions

Measurements of Higgs boson cross sections in the diphoton decay channel are performed using  $pp$  collision data recorded by the ATLAS experiment at the LHC. The data were taken at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV and correspond to the full Run 2 data set with an integrated luminosity of  $139\text{ fb}^{-1}$ .

Fiducial cross-section measurements are reported for a Higgs boson decaying into two isolated photons with transverse momentum greater than 35% and 25% of the diphoton invariant mass, and with  $|\eta| < 2.37$ , excluding the region of  $1.37 < |\eta| < 1.52$ . The integrated fiducial cross section times branching ratio is measured to be

$$\sigma_{\text{fid}} = 65.2 \pm 4.5 \text{ (stat.)} \pm 5.6 \text{ (syst.)} \pm 0.3 \text{ (theo.) fb},$$

and is in agreement with the SM expectation of  $63.6 \pm 3.3$  fb. With a total uncertainty of 11%, this result is the most precise  $H \rightarrow \gamma\gamma$  fiducial measurement to date.

Differential cross sections as a function of different observables sensitive to the Higgs boson production kinematics, the jet kinematics, the CP quantum numbers of the Higgs boson, and the VBF production mechanism are reported and compared to several theoretical predictions. The measurements do not exhibit any significant deviation from the SM expectation.

The strength and tensor structure of the Higgs boson interactions is investigated using five differential observables and an effective Lagrangian which introduces additional CP-even and CP-odd interactions. Two bases were used: the SILH basis of the Higgs Effective Lagrangian and the Warsaw basis of the SMEFT formulation. No significant new physics contributions are observed and the reported limits on such contributions have improved by a factor of two in comparison to the previous ATLAS measurement with the SILH basis, benefiting mainly from the larger size of the acquired data set.

The differential cross-section as a function the  $p_{\text{T}}$  of the diphoton system is used to probe the charm Yukawa coupling of the Higgs boson. Based on the observed shape, a limit is placed on the coupling that is found to be relatively loose and provides complementary constraints to the currently available limits from direct searches of this coupling. This analysis sets the 95% CL limit in the range  $[-19, 24]$  for the ratio of the charm coupling value relative to that predicted by the SM.

## References

- [1] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with  $36\text{fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector*, *Phys. Rev. D* **98** (2018) 052005, arXiv: [1802.04146 \[hep-ex\]](https://arxiv.org/abs/1802.04146).
- [2] CMS Collaboration, *Measurement and interpretation of differential cross sections for Higgs boson production at  $\sqrt{s} = 13\text{ TeV}$* , *Phys. Lett. B* **792** (2019) 369, arXiv: [1812.06504 \[hep-ex\]](https://arxiv.org/abs/1812.06504).
- [3] ATLAS Collaboration, *Measurement of Higgs boson properties in the diphoton decay channel using  $80\text{fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector*, ATLAS-CONF-2018-028, 2018, URL: <https://cds.cern.ch/record/2628771>.
- [4] F. Bishara, U. Haisch, P. F. Monni and E. Re, *Constraining Light-Quark Yukawa Couplings from Higgs Distributions*, *Phys. Rev. Lett.* **118** (2017) 121801, arXiv: [1606.09253 \[hep-ph\]](https://arxiv.org/abs/1606.09253).
- [5] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and  $8\text{ TeV}$  with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803, arXiv: [1503.07589 \[hep-ex\]](https://arxiv.org/abs/1503.07589).
- [6] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The strongly-interacting light Higgs*, *JHEP* **06** (2007) 045, arXiv: [hep-ph/0703164](https://arxiv.org/abs/hep-ph/0703164).
- [7] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1, arXiv: [1706.08945 \[hep-ph\]](https://arxiv.org/abs/1706.08945).
- [8] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, arXiv: [1008.4884 \[hep-ph\]](https://arxiv.org/abs/1008.4884).
- [9] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [10] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19, 2010, URL: <https://cds.cern.ch/record/1291633>, Addendum: ATLAS-TDR-19-ADD-1, 2012, URL: <https://cds.cern.ch/record/1451888>.
- [11] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317, arXiv: [1611.09661 \[hep-ex\]](https://arxiv.org/abs/1611.09661).
- [12] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, *Eur. Phys. J. C* **79** (2019) 205, arXiv: [1810.05087 \[hep-ex\]](https://arxiv.org/abs/1810.05087).
- [13] ATLAS Collaboration, *Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13\text{ TeV}$  using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: <https://cdsweb.cern.ch/record/2677054>.
- [14] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN-2013-004 (2013), arXiv: [1307.1347 \[hep-ph\]](https://arxiv.org/abs/1307.1347).
- [15] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: [1610.07922 \[hep-ph\]](https://arxiv.org/abs/1610.07922).
- [16] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, *Comput. Phys. Commun.* **108** (1998) 56, arXiv: [hep-ph/9704448](https://arxiv.org/abs/hep-ph/9704448).

- [17] M. Spira, *QCD Effects in Higgs physics*, *Fortsch. Phys.* **46** (1998) 203, arXiv: [hep-ph/9705337](#).
- [18] A. Djouadi, M. M. Mühlleitner and M. Spira, *Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-Interface)*, *Acta Phys. Polon. B* **38** (2007) 635, arXiv: [hep-ph/0609292](#).
- [19] A. Djouadi, J. Kalinowski, M. Mühlleitner and M. Spira, *HDECAY: Twenty<sub>++</sub> years after*, *Comput. Phys. Commun.* **238** (2019) 214, arXiv: [1801.09506 \[hep-ph\]](#).
- [20] A. Bredenstein, A. Denner, S. Dittmaier and M. M. Weber, *Radiative corrections to the semileptonic and hadronic Higgs-boson decays  $H \rightarrow WW/ZZ \rightarrow 4 \text{ fermions}$* , *JHEP* **02** (2007) 080, arXiv: [hep-ph/0611234](#).
- [21] A. Bredenstein, A. Denner, S. Dittmaier and M. M. Weber, *Precise predictions for the Higgs-boson decay  $H \rightarrow WW/ZZ \rightarrow 4 \text{ leptons}$* , *Phys. Rev. D* **74** (2006) 013004, arXiv: [hep-ph/0604011](#).
- [22] A. Bredenstein, A. Denner, S. Dittmaier and M. M. Weber, *Precision calculations for the Higgs decays  $H \rightarrow ZZ/WW \rightarrow 4 \text{ leptons}$* , *Nucl. Phys. Proc. Suppl.* **160** (2006) 131, arXiv: [hep-ph/0607060](#).
- [23] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146 \[hep-ph\]](#).
- [24] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, arXiv: [0709.2092 \[hep-ph\]](#).
- [25] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581 \[hep-ph\]](#).
- [26] J. M. Campbell et al., *NLO Higgs boson production plus one and two jets using the POWHEG BOX, MadGraph4 and MCFM*, *JHEP* **07** (2012) 092, arXiv: [1202.5475 \[hep-ph\]](#).
- [27] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: multi-scale improved NLO*, *JHEP* **10** (2012) 155, arXiv: [1206.3572 \[hep-ph\]](#).
- [28] K. Hamilton, P. Nason, C. Oleari and G. Zanderighi, *Merging  $H/W/Z + 0$  and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, *JHEP* **05** (2013) 082, arXiv: [1212.4504 \[hep-ph\]](#).
- [29] K. Hamilton, P. Nason, E. Re and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, *JHEP* **10** (2013) 222, arXiv: [1309.0017 \[hep-ph\]](#).
- [30] K. Hamilton, P. Nason and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, *JHEP* **05** (2015) 140, arXiv: [1501.04637 \[hep-ph\]](#).
- [31] U. Aglietti, R. Bonciani, G. Degrassi and A. Vicini, *Two-loop light fermion contribution to Higgs production and decays*, *Phys. Lett. B* **595** (2004) 432, arXiv: [hep-ph/0404071](#).
- [32] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, *Phys. Lett. B* **670** (2008) 12, arXiv: [0809.1301 \[hep-ph\]](#).
- [33] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NNLO computational techniques: The cases  $H \rightarrow \gamma\gamma$  and  $H \rightarrow gg$* , *Nucl. Phys. B* **811** (2009) 182, arXiv: [0809.3667 \[hep-ph\]](#).
- [34] C. Anastasiou, R. Boughezal and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, *JHEP* **04** (2009) 003, arXiv: [0811.3458 \[hep-ph\]](#).

- [35] A. Pak, M. Rogal and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, [JHEP \*\*02\*\* \(2010\) 025](#), arXiv: [0911.4662 \[hep-ph\]](#).
- [36] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*, [Phys. Lett. B \*\*679\*\* \(2009\) 467](#), arXiv: [0907.2997 \[hep-ph\]](#).
- [37] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, [JHEP \*\*11\*\* \(2009\) 088](#), arXiv: [0909.3420 \[hep-ph\]](#).
- [38] R. V. Harlander, H. Mantler, S. Marzani and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, [Eur. Phys. J. C \*\*66\*\* \(2010\) 359](#), arXiv: [0912.2104 \[hep-ph\]](#).
- [39] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, [Phys. Rev. Lett. \*\*114\*\* \(2015\) 212001](#), arXiv: [1503.06056 \[hep-ph\]](#).
- [40] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, [JHEP \*\*05\*\* \(2016\) 058](#), arXiv: [1602.00695 \[hep-ph\]](#).
- [41] F. Dulat, A. Lazopoulos and B. Mistlberger, *iHixs 2 – Inclusive Higgs cross sections*, [Comput. Phys. Commun. \*\*233\*\* \(2018\) 243](#), arXiv: [1802.00827 \[hep-ph\]](#).
- [42] M. Bonetti, K. Melnikov and L. Tancredi, *Higher order corrections to mixed QCD-EW contributions to Higgs boson production in gluon fusion*, [Phys. Rev. D \*\*97\*\* \(2018\) 056017](#), [Erratum: [Phys. Rev. D 97 \(2018\) 099906](#)], arXiv: [1801.10403 \[hep-ph\]](#).
- [43] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, [J. Phys. G \*\*43\*\* \(2016\) 023001](#), arXiv: [1510.03865 \[hep-ph\]](#).
- [44] ATLAS Collaboration, *Measurement of the  $Z/\gamma^*$  boson transverse momentum distribution in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, [JHEP \*\*09\*\* \(2014\) 145](#), arXiv: [1406.3660 \[hep-ex\]](#).
- [45] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 physics and manual*, [JHEP \*\*05\*\* \(2006\) 026](#), arXiv: [hep-ph/0603175](#).
- [46] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, [Comput. Phys. Commun. \*\*178\*\* \(2008\) 852](#), arXiv: [0710.3820 \[hep-ph\]](#).
- [47] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, [JHEP \*\*02\*\* \(2010\) 037](#), arXiv: [0911.5299 \[hep-ph\]](#).
- [48] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and Electroweak Corrections to the Production of Higgs + 2 Jets via Weak Interactions at the Large Hadron Collider*, [Phys. Rev. Lett. \*\*99\*\* \(2007\) 161803](#), arXiv: [0707.0381 \[hep-ph\]](#).
- [49] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC*, [Phys. Rev. D \*\*77\*\* \(2008\) 013002](#), arXiv: [0710.4749 \[hep-ph\]](#).
- [50] P. Bolzoni, F. Maltoni, S.-O. Moch and M. Zaro, *Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, [Phys. Rev. Lett. \*\*105\*\* \(2010\) 011801](#), arXiv: [1003.4451 \[hep-ph\]](#).
- [51] G. Luisoni, P. Nason, C. Oleari and F. Tramontano,  *$HW^\pm/HZ + 0$  and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, [JHEP \*\*10\*\* \(2013\) 083](#), arXiv: [1306.2542 \[hep-ph\]](#).

- [52] G. Cullen et al., *Automated one-loop calculations with GoSam*, *Eur. Phys. J. C* **72** (2012) 1889, arXiv: [1111.2034 \[hep-ph\]](https://arxiv.org/abs/1111.2034).
- [53] M. L. Ciccolini, S. Dittmaier and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, *Phys. Rev. D* **68** (2003) 073003, arXiv: [hep-ph/0306234 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0306234).
- [54] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Phys. Lett. B* **579** (2004) 149, arXiv: [hep-ph/0307206](https://arxiv.org/abs/hep-ph/0307206).
- [55] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, *Eur. Phys. J. C* **72** (2012) 1868, arXiv: [1111.0761 \[hep-ph\]](https://arxiv.org/abs/1111.0761).
- [56] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, *JHEP* **02** (2013) 078, arXiv: [1211.5015 \[hep-ph\]](https://arxiv.org/abs/1211.5015).
- [57] A. Denner, S. Dittmaier, S. Kallweit and A. Mück, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, *Comput. Phys. Commun.* **195** (2015) 161, arXiv: [1412.5390 \[hep-ph\]](https://arxiv.org/abs/1412.5390).
- [58] O. Brein, R. V. Harlander and T. J. E. Zirke, *vh@nnlo – Higgs Strahlung at hadron colliders*, *Comput. Phys. Commun.* **184** (2013) 998, arXiv: [1210.5347 \[hep-ph\]](https://arxiv.org/abs/1210.5347).
- [59] R. V. Harlander, A. Kulesza, V. Theeuwes and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, *JHEP* **11** (2014) 082, arXiv: [1410.0217 \[hep-ph\]](https://arxiv.org/abs/1410.0217).
- [60] R. V. Harlander, J. Klappert, S. Liebler and L. Simon, *vh@nnlo-v2: new physics in Higgs Strahlung*, *JHEP* **05** (2018) 089, arXiv: [1802.04817 \[hep-ph\]](https://arxiv.org/abs/1802.04817).
- [61] H. B. Hartanto, B. Jäger, L. Reina and D. Wackerth, *Higgs boson production in association with top quarks in the POWHEG BOX*, *Phys. Rev. D* **91** (2015) 094003, arXiv: [1501.04498 \[hep-ph\]](https://arxiv.org/abs/1501.04498).
- [62] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419>.
- [63] W. Beenakker et al., *NLO QCD corrections to  $t\bar{t}H$  production in hadron collisions*, *Nucl. Phys. B* **653** (2003) 151, arXiv: [hep-ph/0211352](https://arxiv.org/abs/hep-ph/0211352).
- [64] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackerth, *Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections*, *Phys. Rev. D* **68** (2003) 034022, arXiv: [hep-ph/0305087](https://arxiv.org/abs/hep-ph/0305087).
- [65] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen and L. Guo, *QCD NLO and EW NLO corrections to  $t\bar{t}H$  production with top quark decays at hadron collider*, *Phys. Lett. B* **738** (2014) 1, arXiv: [1407.1110 \[hep-ph\]](https://arxiv.org/abs/1407.1110).
- [66] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, *JHEP* **06** (2015) 184, arXiv: [1504.03446 \[hep-ph\]](https://arxiv.org/abs/1504.03446).
- [67] B. Jäger, L. Reina and D. Wackerth, *Higgs boson production in association with b jets in the POWHEG BOX*, *Phys. Rev. D* **93** (2016) 014030, arXiv: [1509.05843 \[hep-ph\]](https://arxiv.org/abs/1509.05843).
- [68] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622).
- [69] E. Bothmann et al., *Event Generation with Sherpa 2.2*, (2019), arXiv: [1905.09127 \[hep-ph\]](https://arxiv.org/abs/1905.09127).

[70] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038, arXiv: [0709.1027 \[hep-ph\]](#).

[71] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, arXiv: [1111.1220 \[hep-ph\]](#).

[72] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](#).

[73] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063, arXiv: [hep-ph/0109231](#).

[74] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, arXiv: [0903.1219 \[hep-ph\]](#).

[75] R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](#).

[76] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244, arXiv: [1207.1303 \[hep-ph\]](#).

[77] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, url: <https://cds.cern.ch/record/2206965>.

[78] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568 \[physics.ins-det\]](#).

[79] GEANT4 Collaboration, *GEANT4: a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[80] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010, url: <https://cds.cern.ch/record/1300517>.

[81] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022, 2017, url: <https://cds.cern.ch/record/2298955>.

[82] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, *JINST* **14** (2019) P03017, arXiv: [1812.03848 \[hep-ex\]](#).

[83] M. Cacciari, G. P. Salam and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065, arXiv: [0912.4926 \[hep-ph\]](#).

[84] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, ATL-PHYS-PUB-2015-026, 2015, url: <https://cds.cern.ch/record/2037717>.

[85] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490, arXiv: [1603.02934 \[hep-ex\]](#).

[86] M. Cacciari, G. P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#).

[87] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, url: <https://cds.cern.ch/record/1700870>.

[88] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, arXiv: [1510.03823 \[hep-ex\]](#).

[89] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. (2019), arXiv: [1902.04655 \[hep-ex\]](https://arxiv.org/abs/1902.04655).

[90] M. Oreglia, *A Study of the Reactions  $\psi' \rightarrow \gamma\gamma\psi$* , SLAC-R-0236 (1980), URL: [www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf](http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf).

[91] ATLAS Collaboration, *Measurement of the isolated diphoton cross section in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 012003, arXiv: [1107.0581 \[hep-ex\]](https://arxiv.org/abs/1107.0581).

[92] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214).

[93] M. Frate, K. Cranmer, S. Kalia, A. Vandenberg-Rodes and D. Whiteson, *Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes*, (2017), arXiv: [1709.05681 \[physics.data-an\]](https://arxiv.org/abs/1709.05681).

[94] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[95] ATLAS Collaboration, *Combined search for the Standard Model Higgs boson in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, Phys. Rev. D **86** (2012) 032003, arXiv: [1207.0319 \[hep-ex\]](https://arxiv.org/abs/1207.0319).

[96] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  channels using  $\sqrt{s} = 13$  TeV proton–proton collision data collected with the ATLAS experiment*, ATLAS-CONF-2017-047, 2017, URL: <https://cds.cern.ch/record/2273854>.

[97] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017.

[98] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: [1703.09665 \[hep-ex\]](https://arxiv.org/abs/1703.09665).

[99] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: [1512.01178 \[hep-ph\]](https://arxiv.org/abs/1512.01178).

[100] X. Chen et al., *Precise QCD Description of the Higgs Boson Transverse Momentum Spectrum*, Phys. Lett. B **788** (2019) 425, arXiv: [1805.00736 \[hep-ph\]](https://arxiv.org/abs/1805.00736).

[101] M. A. Ebert, J. K. L. Michel and F. J. Tackmann, *Resummation improved rapidity spectrum for gluon fusion Higgs production*, JHEP **05** (2017) 088, arXiv: [1702.00794 \[hep-ph\]](https://arxiv.org/abs/1702.00794).

[102] M. A. Ebert et al., *SCETlib: A C++ Package for Numerical Calculations in QCD and Soft-Collinear Effective Theory*, DESY-17-099, URL: <http://scetlib.desy.de>.

[103] R. Boughezal et al., *Color singlet production at NNLO in MCFM*, Eur. Phys. J. C **77** (2017) 7, arXiv: [1605.08011 \[hep-ph\]](https://arxiv.org/abs/1605.08011).

[104] J. Gaunt, M. Stahlhofen, F. J. Tackmann and J. R. Walsh, *N-jettiness Subtractions for NNLO QCD calculations*, JHEP **09** (2015) 058, arXiv: [1505.04794 \[hep-ph\]](https://arxiv.org/abs/1505.04794).

- [105] A. Banfi et al., *Jet-vetoed Higgs cross section in gluon fusion at  $N^3LO+NNLL$  with small- $R$  resummation*, *JHEP* **04** (2016) 049, arXiv: [1511.02886 \[hep-ph\]](#).
- [106] I. W. Stewart, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Jet  $p_T$  Resummation in Higgs Production at  $NNLL' + NNLO$* , *Phys. Rev. D* **89** (2014) 054001, arXiv: [1307.1808 \[hep-ph\]](#).
- [107] R. Boughezal, X. Liu, F. Petriello, F. J. Tackmann and J. R. Walsh, *Combining resummed Higgs predictions across jet bins*, *Phys. Rev. D* **89** (2014) 074044, arXiv: [1312.4535 \[hep-ph\]](#).
- [108] X. Chen, T. Gehrmann, E. W. N. Glover and M. Jaquier, *Precise QCD predictions for the production of Higgs + jet final states*, *Phys. Lett. B* **740** (2015) 147, arXiv: [1408.5325 \[hep-ph\]](#).
- [109] X. Chen, J. Cruz-Martinez, T. Gehrmann, E. W. N. Glover and M. Jaquier, *NNLO QCD corrections to Higgs boson production at large transverse momentum*, *JHEP* **10** (2016) 066, arXiv: [1607.08817 \[hep-ph\]](#).
- [110] G. Cullen et al., *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*, *Eur. Phys. J. C* **74** (2014) 3001, arXiv: [1404.7096 \[hep-ph\]](#).
- [111] S. Höche, F. Krauss and M. Schönherr, *Uncertainties in MEPS@NLO calculations of  $h+jets$* , *Phys. Rev. D* **90** (2014) 014012, arXiv: [1401.7971 \[hep-ph\]](#).
- [112] M. Buschmann et al., *Mass Effects in the Higgs-Gluon Coupling: Boosted vs Off-Shell Production*, *JHEP* **02** (2015) 038, arXiv: [1410.5806 \[hep-ph\]](#).
- [113] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, arXiv: [1405.0301 \[hep-ph\]](#).
- [114] R. Frederix, S. Frixione, E. Vryonidou and M. Wiesemann, *Heavy-quark mass effects in Higgs plus jets production*, *JHEP* **08** (2016) 006, arXiv: [1604.03017 \[hep-ph\]](#).
- [115] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061, arXiv: [1209.6215 \[hep-ph\]](#).
- [116] R. Contino, M. Ghezzi, C. Grojean, M. Mühlleitner and M. Spira, *Effective Lagrangian for a light Higgs-like scalar*, *JHEP* **07** (2013) 035, arXiv: [1303.3876 \[hep-ph\]](#).
- [117] A. Alloul, B. Fuks and V. Sanz, *Phenomenology of the Higgs Effective Lagrangian via FEYNRULES*, *JHEP* **04** (2014) 110, arXiv: [1310.5150 \[hep-ph\]](#).
- [118] A. Pomarol and F. Riva, *Towards the ultimate SM fit to close in on Higgs physics*, *JHEP* **01** (2014) 151, arXiv: [1308.2803 \[hep-ph\]](#).
- [119] J. Ellis, V. Sanz and T. You, *The effective Standard Model after LHC Run I*, *JHEP* **03** (2015) 157, arXiv: [1410.7703 \[hep-ph\]](#).
- [120] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070, arXiv: [1709.06492 \[hep-ph\]](#).
- [121] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013, arXiv: [hep-ph/0611129](#).
- [122] A. Buckley et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803, arXiv: [1003.0694 \[hep-ph\]](#).
- [123] A. Buckley, H. Hoeth, H. Lacker, H. Schulz and J. E. von Seggern, *Systematic event generator tuning for the LHC*, *Eur. Phys. J. C* **65** (2010) 331, arXiv: [0907.2973 \[hep-ph\]](#).

- [124] I. Brivio, T. Corbett and M. Trott, *The Higgs width in the SMEFT*, (2019), arXiv: [1906.06949 \[hep-ph\]](#).
- [125] G. Klamke and D. Zeppenfeld, *Higgs plus two jet production via gluon fusion as a signal at the CERN LHC*, *JHEP* **04** (2007) 052, arXiv: [hep-ph/0703202](#).
- [126] J. R. Andersen, K. Arnold and D. Zeppenfeld, *Azimuthal angle correlations for Higgs boson plus multi-jet events*, *JHEP* **06** (2010) 091, arXiv: [1001.3822 \[hep-ph\]](#).
- [127] T. Plehn, D. L. Rainwater and D. Zeppenfeld, *Determining the structure of Higgs couplings at the LHC*, *Phys. Rev. Lett.* **88** (2002) 051801, arXiv: [hep-ph/0105325](#).
- [128] A. Azatov, R. Contino, C. S. Machado and F. Riva, *Helicity selection rules and noninterference for BSM amplitudes*, *Phys. Rev. D* **95** (2017) 065014, arXiv: [1607.05236 \[hep-ph\]](#).
- [129] ATLAS Collaboration, *Constraints on non-Standard Model Higgs boson interactions in an effective Lagrangian using differential cross sections measured in the  $H \rightarrow \gamma\gamma$  decay channel at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, *Phys. Lett. B* **753** (2016) 69, arXiv: [1508.02507 \[hep-ex\]](#).
- [130] ATLAS Collaboration, *Searches for the  $Z\gamma$  decay mode of the Higgs boson and for new high-mass resonances in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *JHEP* **10** (2017) 112, arXiv: [1708.00212 \[hep-ex\]](#).
- [131] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the Validity of the Effective Field Theory Approach to SM Precision Tests*, *JHEP* **07** (2016) 144, arXiv: [1604.06444 \[hep-ph\]](#).
- [132] ATLAS Collaboration, *Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59, arXiv: [1808.08238 \[hep-ex\]](#).
- [133] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, *Phys. Rev. Lett.* **121** (2018) 121801, arXiv: [1808.08242 \[hep-ex\]](#).
- [134] ATLAS Collaboration, *Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment*, *Phys. Rev. Lett.* **120** (2018) 211802, arXiv: [1802.04329 \[hep-ex\]](#).
- [135] ATLAS Collaboration, *Searches for exclusive Higgs and  $Z$  boson decays into  $J/\psi\gamma$ ,  $\psi(2S)\gamma$ , and  $\Upsilon(nS)\gamma$  at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 134, arXiv: [1807.00802 \[hep-ex\]](#).
- [136] P. F. Monni, E. Re and P. Torrielli, *Higgs Transverse-Momentum Resummation in Direct Space*, *Phys. Rev. Lett.* **116** (2016) 242001, arXiv: [1604.02191 \[hep-ph\]](#).
- [137] F. Caola et al., *Bottom-quark effects in Higgs production at intermediate transverse momentum*, *JHEP* **09** (2018) 035, arXiv: [1804.07632 \[hep-ph\]](#).
- [138] R. V. Harlander, *Higgs production in heavy quark annihilation through next-to-next-to-leading order QCD*, *Eur. Phys. J. C* **76** (2016) 252, arXiv: [1512.04901 \[hep-ph\]](#).
- [139] M. Bonvini, A. S. Papanastasiou and F. J. Tackmann, *Matched predictions for the  $b\bar{b}H$  cross section at the 13 TeV LHC*, *JHEP* **10** (2016) 053, arXiv: [1605.01733 \[hep-ph\]](#).

# Appendix

## A Vertex selection

Figure 17 shows the efficiency of the neural-network based selection of the diphoton vertex for MC-simulated ggF  $H \rightarrow \gamma\gamma$  events and for data and MC-simulated  $Z \rightarrow ee$  events ignoring the tracks associated with the electron candidates and treating them as photon candidates. Good agreement between the data and the simulation is found, which is at the level of 1% or better in the range of  $\mu$  that is more relevant for the analysis of the data set,  $\mu > 20$ . The efficiency difference between  $Z \rightarrow ee$  and ggF events is due to the kinematics of the two processes.

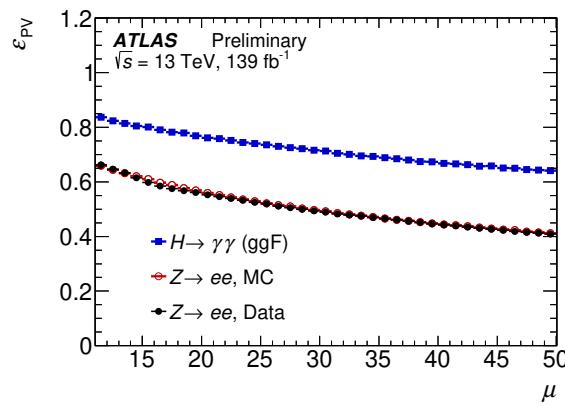


Figure 17: Efficiency to select a diphoton vertex within 0.3 mm of the production vertex as a function of the pile-up in the event. The plot shows the efficiency for simulated ggF events with two unconverted photons (blue markers) and for  $Z \rightarrow ee$  events with the electron tracks removed for the neural network-based identification of the vertex, both in data (black markers) and simulation (red markers).

## B Matrix inversion results

The cross sections as function of  $p_T^{\gamma\gamma}$ ,  $|y_{\gamma\gamma}|$ ,  $N_{\text{jets}}$ ,  $p_T^{j_1}$ ,  $\Delta\phi_{jj}$  and  $m_{jj}$  are obtained with the matrix-inversion unfolding method as a cross-check of the results with bin-by-bin unfolding shown in the main document. The results are shown in Figures 18, 19 and 20. In some bins, particularly for observables for which large migrations between bins are expected, the statistical uncertainties appear to be larger than those for the bin-by-bin correction-factor unfolding reported in the main part of the note as a result of large anti-correlations between bins (up to 50% for neighbouring bins of jet-related observables) that are introduced by the method.

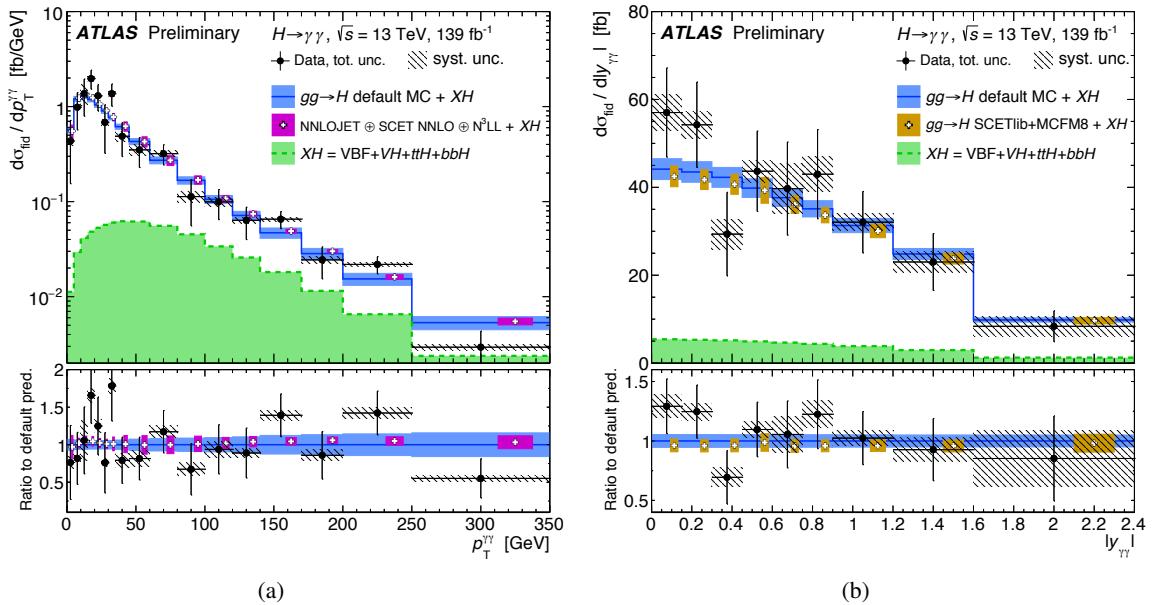


Figure 18: Cross sections unfolded with matrix inversion as a function of the diphoton kinematics, (a)  $p_T^{\gamma\gamma}$ , (b)  $|y_{\gamma\gamma}|$ . The cross section as function of  $p_T^{\gamma\gamma}$  is shown in the range 0–350 GeV. All measurements are compared to the default MC prediction in which ggF is modelled with PowHEG NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction, all described in Section 6.4.

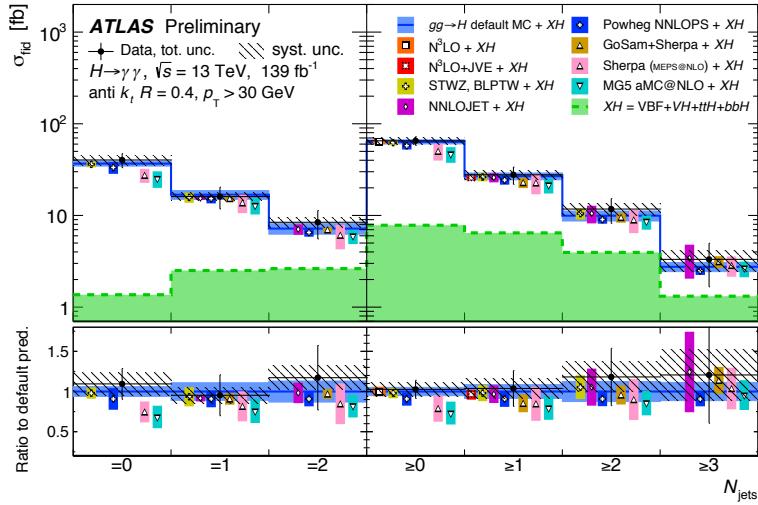


Figure 19: Cross sections unfolded with matrix inversion as a function of the jet multiplicity,  $N_{\text{jets}}$ , in exclusive and inclusive bins. All measurements are compared to the default MC prediction in which ggF is modeled with POWHEG NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction.

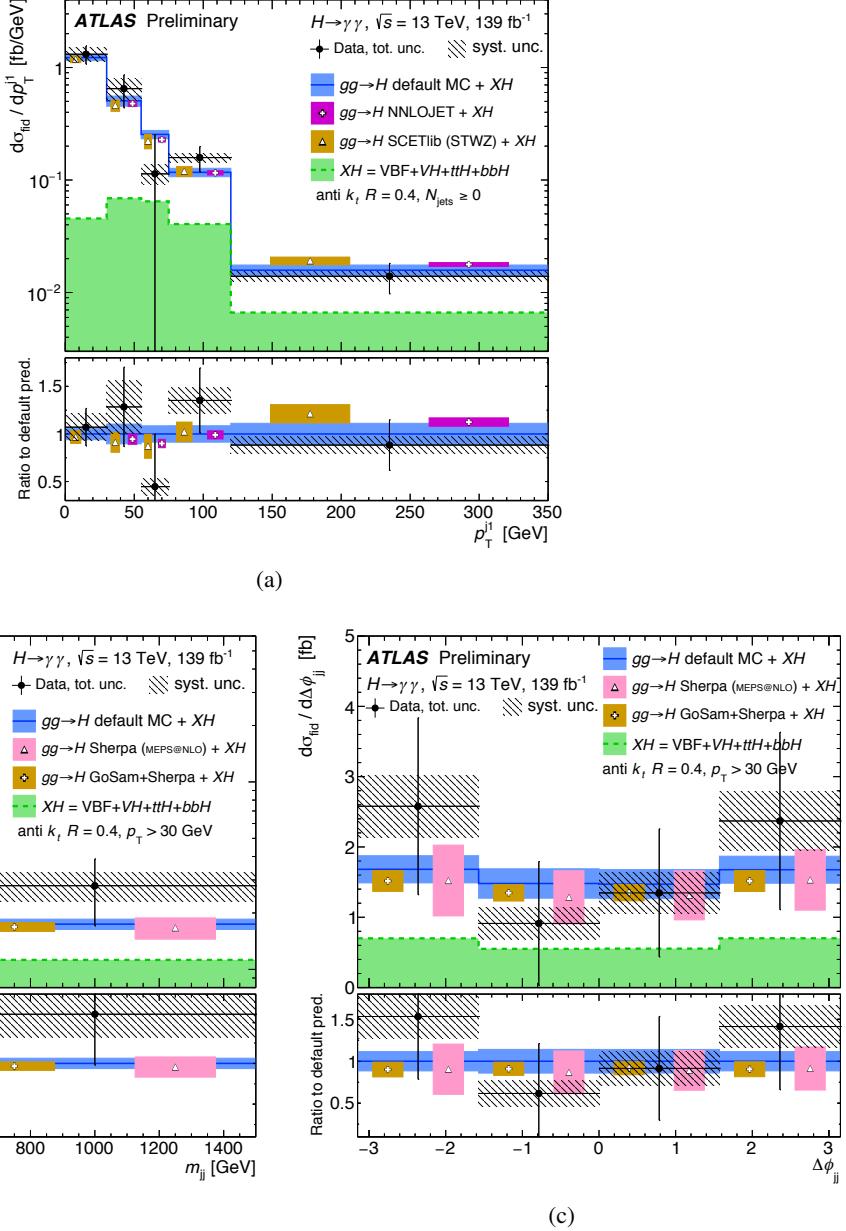


Figure 20: Cross sections unfolded with matrix inversion as a function of jet kinematic observables, (a)  $p_T^{j1}$ , (b)  $m_{jj}$  and (c)  $\Delta\phi_{jj}$ . All measurements are compared to the default MC prediction in which ggF is modeled with POWHEG NNLOPS and other Higgs production processes,  $XH$ , are modeled according to the descriptions of Section 3. Additional comparisons are also shown for different ggF components added to the same  $XH$  prediction.