Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV $pp$ collision data collected by the ATLAS experiment

The ATLAS Collaboration

DOI
10.1007/JHEP11(2018)040

Publication date
2018

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Search for Higgs boson pair production in the $\gamma \gamma b \bar{b}$ final state with 13 TeV $pp$ collision data collected by the ATLAS experiment

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A search is performed for resonant and non-resonant Higgs boson pair production in the $\gamma \gamma b \bar{b}$ final state. The data set used corresponds to an integrated luminosity of 36.1 fb$^{-1}$ of proton-proton collisions at a centre-of-mass energy of 13 TeV recorded by the ATLAS detector at the CERN Large Hadron Collider. No significant excess relative to the Standard Model expectation is observed. The observed limit on the non-resonant Higgs boson pair cross-section is 0.73 pb at 95% confidence level. This observed limit is equivalent to 22 times the predicted Standard Model cross-section. The Higgs boson self-coupling ($\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$) is constrained at 95% confidence level to $-8.2 < \kappa_\lambda < 13.2$. For resonant Higgs boson pair production through $X \rightarrow HH \rightarrow \gamma \gamma b \bar{b}$, the limit is presented, using the narrow-width approximation, as a function of $m_X$ in the range 260 GeV $< m_X < 1000$ GeV. The observed limits range from 1.1 pb to 0.12 pb over this mass range.

KEYWORDS: Hadron-Hadron scattering (experiments), Higgs physics

ArXiv ePrint: 1807.04873
1 Introduction

The Higgs boson ($H$) was discovered by the ATLAS [1] and CMS [2] collaborations in 2012 using proton-proton ($pp$) collisions at the Large Hadron Collider (LHC). Measurements of the properties of the boson are in agreement with the predictions of the Standard Model (SM) [3, 4]. If SM expectations hold, the production of a Higgs boson pair in a single $pp$ interaction should not be observable with the currently available LHC data set. In
the SM, the dominant contributions to this process are shown in figures 1(a) and 1(b). However, some beyond-the-Standard-Model (BSM) scenarios may enhance the Higgs boson pair production rate.

Many BSM theories predict the existence of heavy particles that can decay into a pair of Higgs bosons. These could be identified as a resonance in the Higgs boson pair invariant mass spectrum. They could be produced, for example, through the gluon-gluon fusion mode shown in figure 1(c). Models with two Higgs doublets [5], such as the minimal supersymmetric extension of the SM [6], twin Higgs models [7] and composite Higgs models [8, 9], add a second complex scalar doublet to the Higgs sector. In general, the neutral Higgs fields from the two doublets will mix, which may result in the existence of a heavy Higgs boson that decays into two of its lighter Higgs boson partners. Alternatively, the Randall-Sundrum model of warped extra dimensions [10] predicts spin-0 radions and spin-2 gravitons that could couple to a Higgs boson pair.

In addition to the resonant production, there can also be non-resonant enhancements to the Higgs boson pair cross-section. These can either originate from loop corrections involving new particles, such as light, coloured scalars [11], or through non-SM couplings. Changes to the single Higgs boson production cross-section arising from such loop-corrections are neglected in this paper. Anomalous couplings can either be extensions to the SM, such as contact interactions between two top quarks and two Higgs bosons [12], or be deviations from the SM values of the couplings between the Higgs boson and other particles. In this work, the effective Higgs self-coupling, $\lambda_{HHH}$, is parameterised by a scale factor $\kappa_\lambda (\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}})$ where the SM superscript refers to the SM value of this parameter. The theoretical and phenomenological implications of such couplings for complete models are discussed in refs. [13, 14]. The Yukawa coupling between the top quark and the Higgs boson is set to its SM value in this paper, consistent with its recent direct observation [15, 16].

This paper describes a search for the production of pairs of Higgs bosons in $pp$ collisions at the LHC. The search is carried out in the $\gamma\gamma bb$ final state, and considers both resonant and non-resonant contributions. Previous searches were carried out by the ATLAS and
CMS collaborations in the $\gamma\gamma b\bar{b}$ channel at $\sqrt{s} = 8$ TeV [17, 18], as well as in other final states [19–22] at both $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV.

Events are required to have two isolated photons, accompanied by two jets with dijet invariant mass ($m_{jj}$) compatible with the mass of the Higgs boson, $m_H = 125.09$ GeV [3]. At least one of these jets must be tagged as containing a $b$-hadron; events are separated into signal categories depending on whether one or both jets are tagged in this way.

Loose and tight kinematic selections are defined, where the tight selection is a strict subset of the loose one. The searches for low-mass resonances and for non-SM values of the Higgs boson self-coupling both use the loose selection, as the average transverse momentum ($p_T$) of the Higgs bosons is lower in these cases [23]. The tight selection is used for signals where the Higgs bosons typically have larger average $p_T$, namely in the search for higher-mass resonances and in the measurement of SM non-resonant $HH$ production.

In the search for non-resonant production, the signal is extracted using a fit to the diphoton invariant mass ($m_{\gamma\gamma}$) distribution of the selected events. The signal consists of a narrow peak around $m_H$ superimposed on a smoothly falling background. Only the predominant gluon-gluon fusion production mode, which represents over 90% of the SM cross-section, is considered.

For resonant production, the signal is extracted from the four-object invariant mass ($m_{\gamma\gamma jj}$) spectrum for events with a diphoton mass compatible with the mass of the Higgs boson, by fitting a peak superimposed on a smoothly changing background. The narrow-width approximation is used, focusing on a resonance with mass ($m_X$) in the range $260$ GeV < $m_X$ < $1000$ GeV. Although this search is for a generic scalar decaying into a pair of Higgs bosons, the simulated samples used to optimise the search were produced in the gluon-gluon fusion mode.

The rest of this paper is organised as follows. Section 2 provides a brief description of the ATLAS detector, while section 3 describes the data and simulated event samples used. An overview of object and event selection is given in section 4, while section 5 explains the modelling of signal and background processes. The sources of systematic uncertainties are detailed in section 6. Final results including expected and observed limits are presented in section 7, and section 8 summarises the main findings.

2 ATLAS detector

The ATLAS detector [24] at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle. It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic (EM) and hadronic calorimeters, and a muon spectrometer (MS). The inner tracking detector, consisting of silicon pixel, sili-
con microstrip, and transition radiation tracking systems, covers the pseudorapidity range $|\eta| < 2.5$. The innermost pixel layer, the insertable B-layer (IBL) [25], was added between the first and second runs of the LHC, around a new, narrower and thinner beam pipe. The IBL improves the experiment’s ability to identify displaced vertices and thereby improves the performance of the $b$-tagging algorithms [26]. Lead/liquid-argon (LAr) sampling calorimeters with high granularity provide energy measurements of EM showers. A hadronic steel/scintillator-tile calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$), while a LAr hadronic endcap calorimeter provides coverage over $1.5 < |\eta| < 3.2$. The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The MS surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets, each with eight coils, and with bending power in the range of 2.0 to 7.5 T m. It includes a system of precision tracking chambers, covering the region $|\eta| < 2.7$, and fast detectors for triggering purposes, covering the range $|\eta| < 2.4$.

A two-level trigger system is used to select interesting events [27]. The first-level trigger is implemented in hardware and uses a subset of the total available information to make fast decisions to accept or reject an event, aiming to reduce the rate to around 100 kHz. This is followed by the software-based high-level trigger (HLT), which runs reconstruction and calibration software, reducing the event rate to about 1 kHz.

3 Data and simulated samples

3.1 Data selection

This analysis uses the $pp$ data sample collected at $\sqrt{s} = 13$ TeV with the ATLAS detector in 2015 and 2016, corresponding to an integrated luminosity of 36.1 fb$^{-1}$. All events for which the detector and trigger system satisfy a set of data-quality criteria are considered. Events are selected using a diphoton trigger, which requires two photon candidates with transverse energy ($E_T$) above 35 and 25 GeV, respectively. The overall trigger selection efficiency is greater than 99% for events having the characteristics to satisfy the event selection detailed in section 4.

3.2 Simulated event samples

Non-resonant production of Higgs boson pairs via the gluon-gluon fusion process was simulated at next-to-leading-order (NLO) accuracy in QCD using an effective field theory (EFT) approach, with form factors for the top-quark loop from HPAIR [28, 29] to approximate finite top-quark mass effects. The simulated events were reweighted to reproduce the $m_{HH}$ spectrum obtained in refs. [30, 31], which calculated the process at NLO in QCD while fully accounting for the top-quark mass. The total cross-section is normalised to 33.41 fb, in accordance with a calculation at next-to-next-to-leading order (NNLO) in QCD [32, 33].

Non-resonant BSM Higgs boson pair production with varied $\kappa_\lambda$ was simulated at LO accuracy in QCD [34] for eleven values of $\kappa_\lambda$ in the range $-10 < \kappa_\lambda < 10$. The total cross-sections for these samples were computed as a function of $\kappa_\lambda$ at LO accuracy in QCD.
A constant NNLO/LO $K$-factor (2.283) computed at $\kappa_\lambda = 1$, was then applied. As the amplitude for Higgs boson pair production can be expressed in terms of $\kappa_\lambda$ and the top quark’s Yukawa coupling, weighted combinations of the simulated samples can produce predictions for other values of $\kappa_\lambda$.

Resonant BSM Higgs boson pair production via a massive scalar, was simulated at NLO accuracy for ten different mass points (260, 275, 300, 325, 350, 400, 450, 500, 750 and 1000 GeV) using the narrow-width approximation. For all generated Higgs boson pair samples, both resonant and non-resonant, the branching fractions for $H \to b\bar{b}$ and $H \to \gamma\gamma$ and are taken to be 0.5809 and 0.00227 respectively [32].

This analysis is affected both by backgrounds from single-Higgs-boson production and by non-resonant backgrounds with continuum $m_{\gamma\gamma}$ spectra. Background estimation is carried out using data-driven methods whenever possible; in particular, data are used to estimate the continuum background contribution from SM processes with multiple photons and jets, which constitute the dominant background for this search. Monte Carlo event generators were used for the simulation of different signal hypotheses and the background from SM single Higgs boson production. The major single Higgs boson production channels contributing to the background are gluon-gluon fusion ($ggH$), associated production with a Z boson ($ZH$), associated production with a top quark pair ($t\bar{t}H$) and associated production with a single top quark ($tH$). In addition, contributions from vector-boson fusion (VBF H), associated production with a W boson ($WH$) and associated production with a bottom quark pair ($b\bar{b}H$) are also considered. Overall, the largest contributions come from $t\bar{t}H$ and $ZH$. More information about these simulated background samples can be found in ref. [35] and in table 1.

For all matrix element generators other than SHERPA, the resulting events were passed to another program for simulation of parton showering, hadronisation and the underlying event. This is either Herwig++ with the CTEQ6L1 parton distribution function (PDF) set [36] using the UEEE5 set of tuned parameters [37] or PYTHIA 8 with the NNPDF 2.3 LO PDF set [38] and the A14 set of tuned parameters [39]. For all simulated samples except those generated by SHERPA, the EvtGen v1.2.0 program [40] was used for modelling the properties of $b$- and $c$-hadron decays. Multiple overlaid $pp$ collisions (pile-up) were simulated with the soft QCD processes of PYTHIA 8.186 using the A2 set of tuned parameters [41] and the MSTW2008LO PDF set [42]. The distribution of the number of overlaid collisions simulated in each event approximately matches what was observed during 2015 and 2016 data-taking. Event-level weights were applied to the simulated samples in order to improve the level of agreement.

The final-state particles were passed either through a GEANT 4 [43] simulation of the ATLAS detector, or through the ATLAS fast simulation framework [44], which has been extensively cross-checked against the GEANT 4 model. The output from this detector simulation step is then reconstructed using the same software as used for the data. A list of the signal and dominant background samples used in the paper is shown in table 1.
Table 1. Summary of the event generators and PDF sets used to model the signal and the main background processes. The SM cross-sections $\sigma$ for the Higgs boson production processes with $m_H = 125.09$ GeV are also given separately for $\sqrt{s} = 13$ TeV, together with the orders of the calculation corresponding to the quoted cross-sections, which are used to normalise samples. The following generator versions were used: Pythia 8.212 [45] (event generation), Pythia 8.186 [46] (pile-up overlay); Herwig++ 2.7.1 [47, 48]; Powheg-Box r3154 (base) v2 [49–51]; MadGraph5_aMC@NLO 2.4.3 [52]; Sherpa 2.2.1 [53–56]. The PDF sets used are: CT10 NLO [57], CTEQ6L1 [36], NNPDF 2.3 LO [38], NNPDF 3.0 LO [58], PDF4LHC15 [59]. For the BSM signals, no cross-section is specified as it is the parameter of interest for measurement. For the Sherpa background, no cross-section is used, as the continuum background is fit in data.

4 Object and event selection

The photon selection and event selection for the present search follow those in another published ATLAS $H \rightarrow \gamma\gamma$ analysis [35]. The subsections below detail the selection and identification of all detector-level objects used in the analysis, followed by the event selection criteria and the classification into signal and background control categories.

4.1 Object selection

Photon candidates are reconstructed from energy clusters in the EM calorimeter [63]. The reconstruction algorithm searches for possible matches between energy clusters and tracks reconstructed in the inner detector and extrapolated to the calorimeter. Well-reconstructed tracks matched to clusters are classified as electron candidates, while clusters without matching tracks are classified as unconverted photon candidates. Clusters matched to a reconstructed conversion vertex or to pairs of tracks consistent with the hypothesis of a $e^+e^-$ conversion process are classified as converted photon candidates. Clusters matched to a reconstructed conversion vertex or to pairs of tracks consistent with the hypothesis of a $\gamma \rightarrow e^+e^-$ conversion process are classified as converted photon candidates. Photon energies are determined by summing the energies of all cells belonging to the associated cluster. Simulation-based corrections are then applied to account for energy losses and leakage outside the cluster [63]. The absolute energy scale and response resolution is calibrated using $Z \rightarrow e^+e^-$ events from data. For the photons considered in this analysis, the reconstruction efficiency for both the converted and unconverted photons is 97%. Photon identification is based on the lateral and longitudinal energy profiles of EM showers measured in the calorimeter [64]. The reconstructed photon candidates must satisfy tight photon identifi-
cation criteria. These exploit the fine granularity of the first layer of the EM calorimeter in order to reject background photons from hadron decays. The photon identification efficiency varies as a function of $E_T$ and $|\eta|$ and is typically 85–90% (85–95%) for unconverted (converted) photons in the range of $30 \text{ GeV} < E_T < 100 \text{ GeV}$.

All photon candidates must satisfy a set of calorimeter- and track-based isolation criteria designed to reject the background from jets misidentified as photons and to maximise the signal significance of simulated $H \rightarrow \gamma \gamma$ events against the continuum background. The calorimeter-based isolation variable $E_T^{\text{iso}}$ is defined as the sum of the energies of all topological clusters of calorimeter cells within $\Delta R = 0.2$ of the photon candidate, excluding clusters associated to the photon candidate. The track-based isolation variable $p_T^{\text{iso}}$ is defined as the sum of the transverse momenta ($p_T$) of all tracks with $p_T > 1 \text{ GeV}$ within $\Delta R = 0.2$ of the photon candidate, excluding tracks from photon conversions and tracks not associated with the interaction vertex. Candidates with $E_T^{\text{iso}}$ larger than 6.5% of their transverse energy or with $p_T^{\text{iso}}$ greater than 5% of their transverse energy are rejected. The efficiency of this isolation requirement is approximately 98%. Photons satisfying the isolation criteria are required to fall within the fiducial region of the EM calorimeter defined by $|\eta| < 2.37$, excluding a transition region between calorimeters ($1.37 < |\eta| < 1.52$). Among the photons satisfying the isolation and fiducial criteria, the two with the highest $p_T$ are required to have $E_T/m_{\gamma\gamma} > 0.35$ and 0.25, where $m_{\gamma\gamma}$ is the invariant mass of the diphoton system.

A neural network, trained on a simulated gluon-gluon fusion single-Higgs-boson sample, is used to select the primary vertex most likely to have produced the diphoton pair. The algorithm uses the directional information from the calorimeter and, in the case of converted photons, tracking information, to extrapolate the photon trajectories back to the beam axis. Additionally, vertex properties such as the sum of the squared transverse momenta or the scalar sum of the transverse momenta of the tracks associated with the vertex, are used as inputs to this algorithm. Due to the presence of two high-$p_T$ jets in addition to the two photons, the efficiency for selecting the correct primary vertex is more than 85%. All relevant tracking and calorimetry variables are recalculated with respect to the chosen primary vertex [35].

Jets are reconstructed via the FastJet package [65] from topological clusters of energy deposits in calorimeter cells [66], using the anti-$k_t$ algorithm [67] with a radius parameter of $R = 0.4$. Jets are corrected for contributions from pile-up by applying an event-by-event energy correction evaluated using calorimeter information [68]. They are then calibrated using a series of correction factors, derived from a mixture of simulated events and data, which correct for the different responses to EM and hadronic showers in each of the components of the calorimeters [69]. Jets that do not originate from the diphoton primary vertex, as detailed above, are rejected using the jet vertex tagger (JVT) [70], a multivariate likelihood constructed from two track-based variables. A JVT requirement is applied to jets with $20 \text{ GeV} < p_T < 60 \text{ GeV}$ and $|\eta| < 2.4$. This requirement is 92% efficient at selecting jets arising from the chosen primary vertex. Jets are required to satisfy $|\eta| < 2.5$ and $p_T > 25 \text{ GeV}$; any jets among these that are within $\Delta R = 0.4$ of an isolated photon candidate or within $\Delta R = 0.2$ of an isolated electron candidate are discarded.
The selected jets are classified as $b$-jets (those containing $b$-hadrons) or other jets using a multivariate classifier taking impact parameter information, reconstructed secondary vertex position and decay chain reconstruction as inputs [26, 71]. Working points are defined by requiring the discriminant output to exceed a particular value that is chosen to provide a specific $b$-jet efficiency in an inclusive $t\bar{t}$ sample. Correction factors derived from $t\bar{t}$ events with final states containing two leptons are applied to the simulated event samples to compensate for differences between data and simulation in the $b$-tagging efficiency [72]. The analysis uses two working points which have a $b$-tagging efficiency of 70% (60%), a $c$-jet rejection factor of 12 (35) and a light-jet rejection factor of 380 (1540) respectively. Muons [73] within $\Delta R = 0.4$ of a $b$-tagged jet are used to correct for energy losses from semileptonic $b$-hadron decays. This correction improves the energy measurement of $b$-jets and improves the signal acceptance by 5–6%.

4.2 Event selection and categorisation

Events are selected for analysis if there are at least two photons and at least two jets, one or two of which are tagged as $b$-jets, which satisfy the criteria outlined in section 4.1. The diphoton invariant mass is initially required to fall within a broad mass window of $105 \, \text{GeV} < m_{\gamma\gamma} < 160 \, \text{GeV}$. In order to remain orthogonal to the ATLAS search for $HH \rightarrow bbb\bar{b}$ [19], any event with more than two $b$-jets using the 70% efficient working point is rejected, before the remaining events are divided into three categories. The 2-tag signal category consists of events with exactly two $b$-jets satisfying the requirement for the 70% efficient working point.

Another signal category is defined using events failing this requirement but nevertheless containing exactly one $b$-jet identified using a more stringent (60% efficient) working point. Here the second jet, which is in this case not identified as a $b$-jet, is chosen using a boosted decision tree (BDT). Different BDTs are used when applying the loose and tight kinematic selections. These are optimised using simulated continuum background events as well as signal events from lower-mass or higher-mass resonances, respectively. For each event in the simulated samples used for training the BDTs, every pair of jets in the event is considered. A maximum of one of these jet pairs is correct; in the case of the background sample there are no correct pairs. The training is then performed on these correct and incorrect jet pairs. The BDTs use kinematic variables, namely jet $p_T$, dijet $p_T$, dijet mass, jet $\eta$, dijet $\eta$ and the $\Delta \eta$ between the selected jets, as well as information about whether each jet satisfied less stringent $b$-tagging criteria. The ranking of the jets from best to worst in terms of closest match between the dijet mass and $m_H$, highest jet $p_T$ and highest dijet $p_T$ are also used as inputs. The jet with the highest BDT score is selected and the event is included in the 1-tag signal category. The efficiency with which the correct jet is selected by this BDT is 60–80% across the range of resonant and non-resonant signal hypotheses considered in this paper. If the event contains no $b$-jet from either working point, the event is not directly used in the analysis, but is instead reserved for a 0-tag control category, which is used to provide data-driven estimates of the background shape in the signal categories.

Further requirements are then made on the $p_T$ of the jets and on the mass of the dijet system, which differ for the loose and tight selections. In the loose selection, the highest-$p_T$
jet is required to have $p_T > 40$ GeV, and the next-highest-$p_T$ jet must satisfy $p_T > 25$ GeV, with the invariant mass of the jet pair ($m_{jj}$) required to lie between 80 and 140 GeV. For the tight selection, the highest-$p_T$ and the next-highest-$p_T$ jets are required to have $p_T > 100$ GeV and $p_T > 30$ GeV, respectively, with $90 \text{ GeV} < m_{jj} < 140 \text{ GeV}$. Finally, in the resonant search, the diphoton invariant mass is required to be within 4.7 (4.3) GeV of the Higgs boson mass for the loose (tight) selection. This additional selection on $m_{\gamma\gamma}$ is optimised to contain at least 95% of the simulated Higgs boson pair events for each mass hypothesis.

For non-resonant Higgs boson pair production, among events in the 2-tag category, the efficiency with which the kinematic requirements are satisfied is 10% and 5.8% for the loose and tight selections, respectively. In the 1-tag category, the corresponding efficiencies are 7.2% and 3.9%, which are slightly lower than for the 2-tag category due to the lower probability of selecting the correct jet pair. For the resonant analysis, efficiencies range from 6% to 15.4% in the 2-tag category and from 5.1% to 12.3% in the 1-tag category for $260 \text{ GeV} < m_X < 1000 \text{ GeV}$.

Due to the differing jet kinematics, the signal acceptance is lower in all cases for the generated NLO signal than for a LO signal. The acceptance of the LO prediction is approximately 15% higher when using the tight selection and 10% higher when using the loose selection.

In the resonant analysis, before reconstructing the four-object mass, $m_{\gamma\gamma jj}$, the four-momentum of the dijet system is scaled by $m_H/m_{jj}$. As shown in figure 2, this improves the four-object mass resolution by 60% on average across the resonance mass range of interest. It also modifies the shape of the non-resonant background in the region below 270 GeV. After the correction, the $m_{\gamma\gamma jj}$ resolution is approximately 3% for all signal hypotheses considered in this paper. For the diphoton system, no such scaling is necessary due to the small (approximately 1.5%) diphoton mass resolution.

## 5 Signal and background modelling

Both the resonant and non-resonant searches for Higgs boson pairs proceed by performing unbinned maximum-likelihood fits to the data in the 1-tag and 2-tag signal categories simultaneously. The non-resonant search involves a fit to the $m_{\gamma\gamma}$ distribution, while the search for resonant production uses the $m_{\gamma\gamma jj}$ distribution. The signal-plus-background fit to the data uses parameterised forms for both the signal and background probability distributions. These parameterised forms are determined through fits to simulated samples.

As the loose selection is used for resonances with $m_X \leq 500$ GeV and the tight selection for resonances with $m_X \geq 500$ GeV, different ranges of $m_{\gamma\gamma jj}$ are used in each case. For the loose (tight) selection, only events with $m_{\gamma\gamma jj}$ in the range $245 \text{ GeV} < m_{\gamma\gamma jj} < 610 \text{ GeV}$ ($335 \text{ GeV} < m_{\gamma\gamma jj} < 1140 \text{ GeV}$) are considered. These ranges are the smallest that contain over 95% of all of the simulated signal sample events with $m_X$ below, or above, 500 GeV respectively.
Figure 2. Reconstructed $m_{\gamma\gamma jj}$ with (solid lines) and without (dashed lines) the dijet mass constraint, for a subset of the mass points used in the resonant analysis. The examples shown here are for (a) the 2-tag category with the loose selection and (b) the 1-tag category with the tight selection. The effect on the continuum background is also shown in (a).

5.1 Background composition

Contributions to the continuum diphoton background originate from $\gamma\gamma$, $\gamma j$, $j\gamma$ and $jj$ sources produced in association with jets, where $j$ denotes jets misidentified as photons and $\gamma j$ and $j\gamma$ differ by the jet faking the sub-leading or the leading photon candidate respectively. These are determined from data using a double two-dimensional sideband method (2x2D) based on varying the photon identification and isolation criteria [74, 75]. The number and relative fraction of events from each of these sources is calculated separately for the 1- and 2-tag categories. In each case the contribution from $\gamma\gamma$ events is in the range 80–90%.

The choice of functional form used to fit the background in the final likelihood models is derived using simulated events. Continuum $\gamma\gamma$ events were simulated using the SHERPA event generator as described in section 3. As this prediction from SHERPA does not provide a good description of the $m_{\gamma\gamma}$ spectrum in data, the mismodelling is corrected for using a data-driven reweighting function.

In the 0-tag control category, the number of events in data is high enough that the 2x2D method can be applied in bins of $m_{\gamma\gamma}$. The events generated by SHERPA can also be divided into $\gamma\gamma$, $\gamma j$, $j\gamma$ and $jj$ sources based on the same photon identification and isolation criteria as used in data. For each of these sources, the $m_{\gamma\gamma}$ distributions for both SHERPA and the data are fit using an exponential function and the ratio of the two fit results is taken as an $m_{\gamma\gamma}$-dependent correction function. The size of the correction is less than 5% for the majority of events.

These reweighting functions are then applied in the 1-tag and 2-tag signal categories to correct the shape of the SHERPA prediction. The fractional contribution from the different continuum background sources is fixed to the relative proportions derived in data with the 2x2D method. Finally, the overall normalisation is chosen such that, in the disjoint
sideband region $105 \text{ GeV} < m_{\gamma\gamma} < 120 \text{ GeV}$ and $130 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$, the total contribution from all backgrounds is equal to that from data.

The contribution from $\gamma\gamma$ produced in association with jets is further divided in accord with the flavours of the two jets (for example $bb$, $bc$, $c + \text{light jet}$). This decomposition is taken directly from the proportions predicted by the SHERPA event generator and no attempt is made to classify the data according to jet flavour. The continuum background in the 1-tag category comes primarily from $\gamma\gamma bj$ events ($\sim 60\%$) and in the 2-tag category from $\gamma\gamma bb$ events ($\sim 80\%$). A comparison between data in the 0-tag control category and this data-driven prediction of the total background can be seen in figure 3(a) for the $m_{\gamma\gamma}$ distribution from the tight selection and in figure 3(b) for the $m_{\gamma\gamma jj}$ distribution from the loose selection.

5.2 Signal modelling for the non-resonant analysis

The shape of the diphoton mass distribution in $HH \rightarrow \gamma\gamma jj$ events is described by the double-sided Crystal Ball function [35], consisting of a Gaussian core with power-law tails on either side. The parameters of this model are determined through fits to the simulated non-resonant SM $HH$ sample described in section 3.2.

5.3 Background modelling for the non-resonant analysis

For the non-resonant analysis, the continuum $m_{\gamma\gamma}$ background is modelled using a functional form obtained from a fit to the data. The potential bias arising from this procedure, termed ‘spurious signal’, is estimated by performing signal-plus-background fits to the
combined continuum background from simulation, including the $\gamma\gamma$, $\gamma j$, $j\gamma$ and $jj$ components [35]. The maximum absolute value of the extracted signal, for a signal in the range $121\text{ GeV} < m_{\gamma\gamma} < 129\text{ GeV}$, is taken as the bias. This method is used to discriminate between different potential fit functions — the function chosen is the one with the smallest spurious signal bias. If multiple functions have the same bias, the one with the smallest number of parameters is chosen. The first-order exponential function has the smallest bias among the seven functions considered and is therefore chosen. The background from single Higgs boson production is described using a double-sided Crystal Ball function, with its parameters determined through fits to the appropriate simulated samples.

5.4 Signal modelling for the resonant analysis
For each resonant hypothesis, a fit is performed to the $m_{\gamma\gamma jj}$ distribution of the simulated events in a window around the nominal $m_X$. The shape of this distribution is described using a function consisting of a Gaussian core with exponential tails on either side. A simultaneous fit to all signal samples is carried out in which each of the model parameters is further parameterised in terms of $m_X$. This allows the model to provide a prediction for any mass satisfying $260\text{ GeV} < m_X < 1000\text{ GeV}$, where these boundaries reflect the smallest and largest $m_X$ values among the generated samples described in section 3.2.

5.5 Background modelling for the resonant analysis
For the resonant analysis, a spurious-signal study is also carried out, using the $m_{\gamma\gamma jj}$ distribution for events within the $m_{\gamma\gamma}$ window described in section 4.2. The background used to evaluate the spurious-signal contribution is a combination of the continuum backgrounds together with the single Higgs boson backgrounds.

Due to the different $m_{\gamma\gamma jj}$ ranges used with the loose and tight selections, the shape of the $m_{\gamma\gamma jj}$ distribution differs between these two cases and hence different background functions are considered. For the loose (tight) mass selection, the Novosibirsk function\(^2\) (exponential function) has the smallest bias among the three (four) functions considered and is therefore chosen. As a result, for low-mass resonances both the signal and background fit functions have a characteristic peaked shape. This degeneracy could potentially introduce a bias in the extracted signal cross-section. In order to stabilise the background fit, nominal values of the shape parameters are estimated by fitting to the simulated events described in section 5.1. The shape is then allowed to vary in the likelihood to within the statistical covariance of this template fit. Experimental systematics on the background shape have a small effect and are neglected. The normalisation of the background is estimated by interpolating the $m_{\gamma\gamma}$ sideband data. Additionally, a simple bias test is performed by drawing pseudo-data sets from the overall probability distribution created by combining the Novosibirsk background function with the signal function. For each mass point and each value of the injected signal cross-section, fits are performed on the ensemble of pseudo-data sets and the median extracted signal cross-section is recorded. For resonances with masses below 400 GeV, a small correction is applied to remove the observed bias. The correction is less

\[ P(x) = e^{-0.5(q_0)^2} \Lambda^2 + x^2 \]

where $q_0 = 1 + \Lambda(x - x_0)/\sigma \times \frac{\sinh(\Lambda \sqrt{\ln 4})}{\Lambda \sqrt{\ln 4}}$ [76].

\[ 2 \]
than ±0.05 pb everywhere and a corresponding uncertainty of ±0.02 pb in this correction is applied to the extracted signal cross-section. The corresponding uncertainty in the number of events in each category is roughly half that of the spurious signal.

6 Systematic uncertainties

Although statistical uncertainties dominate the sensitivity of this analysis given the small number of events, care is taken to make the best possible estimates of all systematic uncertainties, as described in more detail below.

6.1 Theoretical uncertainties

Theoretical uncertainties in the production cross-section of single Higgs bosons are estimated by varying the renormalisation and factorisation scales. In addition, uncertainties due to the PDF and the running of the QCD coupling constant ($\alpha_S$) are considered. The scale uncertainties reach a maximum of $+20\% -24\%$ and the PDF+$\alpha_S$ uncertainty is not more than $±3.6\%$ [32]. An uncertainty in the rate of Higgs boson production with associated heavy-flavour jets is also considered. A 100% uncertainty is assigned to the $ggH$ and $WH$ production modes, motivated by studies of heavy-flavour production in association with top-quark pairs [77] and $W$ boson production in association with $b$-jets [78]. No heavy-flavour uncertainty is assigned to the $ZH$ and $t\bar{t}H$ production modes, where the dominant heavy-flavour contribution is already accounted for in the LO process. Finally, additional theoretical uncertainties in single Higgs boson production from uncertainties in the $H \to \gamma\gamma$ and $H \to b\bar{b}$ branching fractions are $+2.9\% -2.8\%$ and $±1.7\%$, respectively [32].

The same sources of uncertainty are considered on the SM $HH$ signal samples. The effect of scale and PDF+$\alpha_S$ uncertainties on the NNLO cross-section for SM Higgs boson pair production are 4–8% and 2–3% respectively. In addition, an uncertainty of 5% arising from the use of infinite top-quark mass in the EFT approximation is taken into account [30]. The cross-section, scale and PDF uncertainties are decorrelated between the single Higgs boson background and SM $HH$ signal.

In the search for resonant Higgs boson pair production, uncertainties arising from scale and PDF uncertainties, which primarily affect the signal yield, are neglected. For this search, the SM non-resonant $HH$ production is considered as a background, with an overall uncertainty on the cross-section of $±7\% -6\%$. Interference between SM $HH$ and the BSM signal is neglected.

For all samples, systematic differences between alternative models of parton showering and hadronisation were considered and found to have a negligible impact.

6.2 Experimental uncertainties

The systematic uncertainty in the integrated luminosity for the data in this analysis is 2.1%. It is derived following a methodology similar to that detailed in ref. [79], using beam-separation scans performed in 2015 and 2016.
The efficiency of the diphoton trigger is measured using bootstrap methods [27], and is found to be 99.4% with a systematic uncertainty of 0.4%. Uncertainties associated with the vertex selection algorithm have a negligible impact on the signal selection efficiency.

Differences between data and simulation give rise to uncertainties in the calibration of the photons and jets used in this analysis. As the continuum backgrounds are estimated from data, these uncertainties are applied only to the signal processes and to the single-Higgs-boson background process. Experimental uncertainties are propagated through the full analysis procedure, including the kinematic and BDT selections. The relevant observables are then constructed, before signal and background fits are performed as described in section 5. Changes in the peak location ($m_{\text{peak}}$), width ($\sigma_{\text{peak}}$) and expected yield in $m_{\gamma\gamma}$ ($m_{\gamma\gamma jj}$) for the non-resonant (resonant) model, relative to the nominal fits, are extracted. The tail parameters are kept at their nominal values in these modified fits. For the resonant analysis, systematic uncertainties are evaluated for each $m_X$ and the maximum across the range is taken as a conservative uncertainty.

The dominant yield uncertainties are listed in table 2. Uncertainties in the photon identification and isolation directly affect the diphoton selection efficiency; jet energy scale and resolution uncertainties affect the $m_{bb}$ window acceptance [69, 80, 81], while flavour-tagging uncertainties lead to migration of events between categories. Uncertainties in the peak location (width), which are mainly due to uncertainties in the photon energy scale (energy resolution), are about 0.2–0.6% (5–14%) for both the single-Higgs-boson and Higgs boson pair samples in the resonant and non-resonant analyses.

The spurious signal for the chosen background model, as defined in sections 5.3 and 5.5, is assessed as an additional uncertainty in the total number of signal events in each category. In the 2-tag (1-tag) category, the uncertainty corresponds to 0.63 (0.25) events for the non-resonant analysis, 0.58 (2.06) events for the resonant analysis with the loose selection, and 0.21 (0.89) events for the resonant analysis with the tight selection.

Finally, as described in section 5.5, an $m_X$-dependent correction to the signal cross-section, together with its associated uncertainty, is applied in the case of the resonant analysis at low masses to adjust for a small degeneracy bias.

### 7 Results

The observed data are in good agreement with the data-driven background expectation, as summarised in table 3. Across all categories, the number of observed events in data is compatible with the number of expected background events within the calculated uncertainties.

The signal and background models described in section 5 are used to construct an unbinned likelihood function which is maximised with respect to the observed data. In each case the parameter of interest is the signal cross-section, which is related in the likelihood model to the number of signal events after considering the integrated luminosity, branching ratio, phase-space acceptance and detection efficiency of the respective categories. The models for the statistically independent 1-tag and 2-tag categories are simultaneously fit to the data by maximizing the product of their likelihoods. The likelihood model also includes a number of nuisance parameters associated with the background shape and normalisation,
The loose and tight selections are not orthogonal.

The uncertainties on the continuum background are those arising from the fitting procedure. The uncertainties on the SM Higgs boson pair signal assume that the total production cross-section is 33 fb.

Sources marked ‘|’ and other sources not listed in the table are negligible by comparison. No systematic uncertainties related to the continuum background are considered, since this is derived through a fit to the observed data.

Table 3. Expected and observed numbers of events in the 1-tag and 2-tag categories for events passing the selection for the resonant analysis, including the $m_{\gamma\gamma}$ requirement. The event numbers quoted for the SM Higgs boson pair signal assume that the total production cross-section is 33.41 fb. The uncertainties on the continuum background are those arising from the fitting procedure. The uncertainties on the single-Higgs-boson and Higgs boson pair backgrounds are from Monte Carlo statistical error. The loose and tight selections are not orthogonal.
as well as the theoretical and experimental systematic uncertainties described in section 6. These nuisance parameters are included in the likelihood as terms which modulate their respective parameters, such as signal yield, along with a constraint term which encodes the scale of the uncertainty by reducing the likelihood when the parameter is pulled from its nominal value. In general the nuisance parameter for each systematic uncertainty has a correlated effect between 1-tag and 2-tag categories, with the exception of the spurious signal and background shape parameters, which are considered as individual degrees of freedom in each category.

Figure 4 shows the observed diphoton invariant mass spectra for the non-resonant analysis with the loose (top) and tight (bottom) selections. The best-fit Higgs boson pair cross-section is $0.04^{+0.43}_{-0.36} (-0.21^{+0.33}_{-0.25})$ pb for the loose (tight) selection. Figure 5 shows the observed four-body invariant mass spectra for the resonant analysis in the loose (top) and tight (bottom) selections. Maximum-likelihood background-only fits are also shown. While local fluctuations may appear in some of the categories shown in figure 5, only the combined two-category unbinned likelihood is considered for setting limits. The largest discrepancy between the background-only hypothesis and the data manifests as an excess at 480 GeV with a local significance of $1.2 \sigma$. The results are also interpreted as upper limits on the relevant Higgs boson pair production cross-sections.

Exclusion limits are set on Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state. The limits for both resonant and non-resonant production are calculated using the CL$_S$ method \cite{82}, with the likelihood-based test statistic $\tilde{q}_\mu$ which is suitable when considering signal strength $\mu \geq 0$ \cite{83,84}. Because both the expected and observed numbers of events are small in the case of the resonant analysis, test-statistic distributions are evaluated by pseudo-experiments generated by profiling the nuisance parameters of the likelihood model on the observed data, as described in ref. \cite{84}. Better limits on $\kappa_\lambda$ are expected with the loose selection, whereas for the SM value $\kappa_\lambda = 1$ the strongest limits on the Higgs boson pair cross-section are derived from the tight selection.

### 7.1 Exclusion limits on non-resonant $HH$ production

The 95% confidence level (CL) upper limit for the non-resonant Higgs boson pair cross-section is obtained using the tight selection. Figure 6(a) shows this upper limit, together with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands. The observed (expected) value is $0.73 (0.93)$ pb. As a multiple of the SM production cross-section, the observed (expected) limits are 22 (28). The limits and the $\pm 1\sigma$ band around each expected limit are presented in table 4.

### 7.2 Exclusion limits on $\lambda_{HHH}$

Varying the Higgs boson self-coupling, $\lambda_{HHH}$, affects both the total cross-section of the non-resonant Higgs boson pair production and the event kinematics, affecting the signal selection efficiency. In the non-resonant analysis, results are interpreted in the context of $\kappa_\Lambda$, using the loose selection, which is more sensitive for the range of $\kappa_\Lambda$ values accessible with this data set. As discussed in section 3.2, the samples used for this interpretation were generated at LO. The 95% CL limits on $\sigma_{gg\rightarrow HH}$ are shown together with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands around the expected limit in figure 6(b). The limits are calculated
Figure 4. For the non-resonant analysis, data (black points) are compared with the background-only fit (blue solid line) for $m_{\gamma\gamma}$ in the 1-tag (left) and 2-tag (right) categories with the loose (top) and tight (bottom) selections. Both the continuum $\gamma\gamma$ background and the background from single Higgs boson production are considered. The lower panel shows the residuals between the data and the best-fit background.
Figure 5. For the resonant analysis, data (black points) are compared with the background-only fit (blue solid line) for $m_{\gamma\gamma jj}$ in the 1-tag (left) and 2-tag (right) categories with the loose (top) and tight (bottom) selections. The lower panel shows the residuals between the data and the best-fit background.
<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Expected</th>
<th>$-1\sigma$</th>
<th>$+1\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{gg \rightarrow HH}$ [pb]</td>
<td>0.73</td>
<td>0.93</td>
<td>0.66</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Table 4. The 95% CL observed and expected limits on the Higgs boson pair cross-section in pb and as a multiple of the SM production cross-section. The ±1σ band around each 95% CL limit is also indicated.

![Observed and Expected Limits](image)

Figure 6. The expected and observed 95% CL limits on the non-resonant production cross-section $gg \rightarrow HH$ (a) for the SM-optimised limit using the tight selection and (b) as a function of $\kappa_\lambda$ using the loose selection. In (a) the red line indicates the 95% confidence level. The intersection of this line with the observed, expected, and ±1σ and ±2σ bands is the location of the limits. In (b) the red line indicates the predicted $HH$ cross-section if $\kappa_\lambda$ is varied but all other couplings remain at their SM values. The red band indicates the theoretical uncertainty of this prediction.

using the asymptotic approximation [83] for the profile-likelihood test statistic. Fixing all other SM parameters to their expected values, the Higgs boson self-coupling is constrained at 95% CL to $-8.2 < \kappa_\lambda < 13.2$ whereas the expected limits are $-8.3 < \kappa_\lambda < 13.2$.

### 7.3 Exclusion limits on resonant $HH$ production

The 95% CL limits on resonant Higgs boson pair production are shown in figure 7, utilising both the loose and tight selections. The SM $HH$ contribution is considered as part of the background in this search although its inclusion has a negligible impact on the results. For resonance masses in the range $260 \text{ GeV} < m_X < 1000 \text{ GeV}$, the observed (expected) limits range between 1.14 (0.90) pb and 0.12 (0.15) pb.
8 Conclusions

Searches for resonant and non-resonant Higgs boson pair production in the $\gamma \gamma b \bar{b}$ final state are performed using 36.1 fb$^{-1}$ of $pp$ collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC in 2015 and 2016. No significant deviations from the Standard Model predictions are observed. A 95% CL upper limit of 0.73 pb is set on the cross-section for non-resonant production, while the expected limit is 0.93 pb. This observed (expected) limit is 22 (28) times the predicted SM cross-section. The Higgs boson self-coupling is constrained at 95% CL to $-8.2 < \kappa_{\lambda} < 13.2$ whereas the expected limits are $-8.3 < \kappa_{\lambda} < 13.2$. For resonant production of $X \rightarrow HH \rightarrow \gamma \gamma b \bar{b}$, a limit is presented for the narrow-width approximation as a function of $m_X$. The observed (expected) limits range between 1.1 pb (0.9 pb) and 0.12 pb (0.15 pb) in the range $260 \text{GeV} < m_X < 1000 \text{GeV}$.

Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.
We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNISW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRTN, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [85].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[62] G. Luisoni, P. Nason, C. Oleari and F. Tramontano, HW±/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, JHEP 02 (2013) 083 [arXiv:1306.2542] [inspire].


The ATLAS collaboration


M. Besnou44, N. Bessou142, A. Bethan98, S. Bethke113, A. Betti24, A.J. Bevan90, J. Beyer113, R.M. Bianchi136, O. Biebel112, D. Biedermann19, R. Bielski35, K. Bierwagen97, N.V. Biesuz69a,69b, M. Biglietti29a, T.R.V. Billard107, M. Bindi51, A. Bingen124, C. Binit70a,70b,70c, S. Bioudi123b,23a,