



Search for boosted diphoton resonances in the 10 to 70 GeV mass range using 138 fb^{-1} of 13 TeV pp collisions with the ATLAS detector

The ATLAS Collaboration

A search for diphoton resonances in the mass range between 10 and 70 GeV with the ATLAS experiment at the Large Hadron Collider (LHC) is presented. The analysis is based on pp collision data corresponding to an integrated luminosity of 138 fb^{-1} at a centre-of-mass energy of 13 TeV recorded from 2015 to 2018. Previous searches for diphoton resonances at the LHC have explored masses down to 65 GeV, finding no evidence of new particles. This search exploits the particular kinematics of events with pairs of closely spaced photons reconstructed in the detector, allowing examination of invariant masses down to 10 GeV. The presented strategy covers a region previously unexplored at hadron colliders because of the experimental challenges of recording low-energy photons and estimating the backgrounds. No significant excess is observed and the reported limits provide the strongest bound on promptly decaying axion-like particles coupling to gluons and photons for masses between 10 and 70 GeV.

1 Introduction

One of the primary goals of the Large Hadron Collider (LHC) experiments is to investigate any ‘beyond the Standard Model’ (BSM) scenarios where deviations from Standard Model (SM) behaviour appear in the high-energy regime of observables accessible in hadron collisions. This work focuses on diphoton final states, looking for new phenomena in the form of narrow resonances X in the mass range between 10 and 70 GeV, below that targeted in previous diphoton searches at the LHC [1–4], and extending the experimental coverage of the LHC to BSM scenarios where i) all the heavy states are beyond the reach of the LHC, ii) a scalar which is a singlet of the SM gauge group is naturally lighter than the EW scale, iii) the singlet scalar is abundantly produced in proton–proton (pp) collisions and iv) the singlet scalar decays promptly into a pair of SM particles, generically with a narrow width.

The natural targets of this search are pseudo Nambu–Goldstone bosons (pNGBs) associated with spontaneously broken approximate global symmetries at the TeV scale or above, often referred to as axion-like particles (ALPs). A light ALP coupled to gluons would be abundantly produced in pp collisions, and even though it would decay mostly into dijets, its suppressed coupling to photons presents experimental advantages that make the diphoton final state more desirable to look for at the LHC. Light ALPs with these phenomenological features are predicted by a wide variety of BSM scenarios, such as ‘heavy’ QCD axion models [5–17], the R-axion in SUSY-breaking models [18–20] and Composite Higgs models [21–23]. Light ALPs can also be the mediator of dark-matter freeze out [24] and trigger baryogenesis [25]. In all of these models, the mass of the ALP is a free parameter without any particularly strong theory expectation, and for this reason the experimental coverage should span all possible masses.

This search uses the diphoton invariant mass distribution as the discriminating variable. It therefore closely follows the strategy used in previous diphoton searches [2, 26], but it also exhibits some clear differences. Selecting photons with transverse energies close to the trigger thresholds sculpts the diphoton invariant mass distribution at its lower edge, limiting the lowest accessible mass and requiring an accurate description of the background shape to avoid drastically limiting the final sensitivity. This challenge is overcome here by selecting events with pairs of closely spaced photons and large diphoton transverse momentum, typically arising from recoil against a jet. Concretely, requiring photon pairs with transverse momentum $p_T^{\gamma\gamma}$ larger than 50 GeV flattens the background shape, making it easier to describe with simple analytic functions.

A model composed of analytic functions is used to describe the signal and background components and fit the diphoton invariant mass spectrum to search for narrow resonances. The impact of the uncertainty arising from the analytic function chosen to describe the background shape, estimated by fitting the analytic model to a representative sample of simulated background events, is mitigated by applying a smoothing procedure using Gaussian Processes to the simulated background sample. In the scenario in which no significant signal excess is observed, limits are placed on the production cross-section times branching ratio, $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$, as a function of the resonance mass m_X . Additionally, the observed limits are recast in the ALP parameter space in terms of the mass of an ALP, m_a , and its decay constant, f_a .

2 ATLAS detector

The ATLAS experiment [27–29] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle.¹ It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer (MS). The ID covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron sampling calorimeter composed of steel and scintillator tiles covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for EM and hadronic energy measurements up to $|\eta| = 4.9$. For $|\eta| < 2.5$, the EM calorimeter is divided into three longitudinal layers, which are finely segmented in η and ϕ , particularly in the η direction in the first layer. This segmentation allows the measurement of the lateral and longitudinal shower profile, and the calculation of shower shapes [30] used for electron and photon identification. The longitudinal segmentation of the EM calorimeter is also exploited to calibrate the energy response of photon candidates [30]. The MS surrounds the calorimeters and is equipped with three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. A two-level trigger system is used to select events to be recorded [31]. The first-level trigger is implemented in hardware and uses a subset of the detector information. It selects events of interest within the event rate limitation of 100 kHz. This is followed by a software-based trigger which runs algorithms similar to those in the offline reconstruction software, reducing the event rate to approximately 1 kHz. An extensive software suite [32] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

3 Data and simulated event samples

The search is performed using the $\sqrt{s} = 13$ TeV pp collision dataset with a bunch spacing of 25 ns collected from 2015 to 2018 by the ATLAS detector, referred to as the full Run 2 dataset in the following. Only events with stable beam conditions and all ATLAS subsystems operational are considered [33], corresponding to an integrated luminosity of 138 fb^{-1} [34]. The data were recorded using a diphoton trigger [35] that required two electromagnetic energy clusters satisfying identification criteria based on the expected shape of the two electromagnetic showers and with transverse energies E_T above a certain threshold that varied across the data-taking period to cope with the increase in instantaneous luminosity over the years. In 2015 and the first portion of 2016, the E_T threshold was 20 GeV, while in the remainder of 2016 an E_T threshold of 22 GeV was imposed. During 2017 and 2018, the E_T threshold was lowered to 20 GeV and combined with an additional requirement on the calorimetric isolation transverse energy. Data collected with alternative and prescaled diphoton triggers with looser identification criteria for the EM shower shapes, corresponding to an integrated luminosity of 6.3 fb^{-1} , are used for background estimation purposes.

Monte Carlo (MC) simulated events are used to optimize the analysis selections and to characterize the signal and background shapes.

¹ 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 upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Signal event samples were generated for a hypothetical resonance produced in gluon–gluon fusion in association with up to two additional jets for resonance masses between 10 and 80 GeV. The decay width Γ_a was set to 4 MeV, negligible compared to the experimental resolution, to describe a hypothetical resonance in the narrow-width approximation (NWA). The samples were generated using the effective-field-theory approach [36] implemented in MADGRAPH5_AMC@NLO [37] with the NNPDF2.3LO PDF set [38], and using the A14 set [39] of tuned parameters and PYTHIA 8.240 [40] to simulate parton showering, hadronization and the decay of the resonance into a pair of photons.

Background events with two prompt photons and associated jets were simulated using the SHERPA 2.2.4 [41, 42] event generator. Matrix elements were calculated in perturbative QCD (pQCD) at next-to-leading order (NLO) for up to one additional parton, and at LO for two or three partons, and merged with the SHERPA parton-shower simulation using the MEPS@NLO prescription [43–46]. The NNPDF3.0NNLO PDF set was used in conjunction with a dedicated parton-shower tune in the SHERPA generator. Interference effects between the resonant signal and all background processes are expected to be small for narrow-width signals and are neglected in this analysis.

The effects of multiple pp interactions in the same bunch crossing as the hard scatter and in neighbouring ones (defined as pile-up) are included using simulated events generated with PYTHIA 8. Simulated events were weighted to reproduce the distribution of the average number of interactions per bunch crossing observed in data.

All simulated signal events were processed using a full simulation of the ATLAS detector [47] based on GEANT4 [48]. The background $\gamma\gamma$ events were processed using a fast simulation of the ATLAS detector [49], where the full simulation of the calorimeter is replaced with a parameterization of the calorimeter response. All simulated events were reconstructed with the same reconstruction algorithms as those used for data.

4 Object and event selection

Photon candidates are reconstructed from topological clusters of energy deposited in the EM calorimeter, as well as from charged-particle tracks and conversion vertices reconstructed in the inner detector, and they are calibrated as described in Ref. [30]. The event selection requires at least two photon candidates with transverse energies larger than 22 GeV and $|\eta| < 2.37$, excluding the barrel-to-endcap transition regions of the calorimeter, $1.37 < |\eta| < 1.52$. The transverse energy requirement is chosen to mitigate the effect of the trigger efficiency turn-on from the trigger thresholds discussed in Section 3. The properties of the EM clusters associated with the two highest- E_T photons and additional information from the tracking systems are used to identify the diphoton production vertex [50], which is used to correct the photon direction, resulting in improved $m_{\gamma\gamma}$ resolution.

To reduce the background from jets, photon candidates are required to satisfy *tight* identification criteria based on the shape of EM showers in the LAr calorimeter and energy leakage into the hadronic calorimeter [30]. Events with one or both photon candidates passing a looser identification are kept for background estimations. The *tight* identification is optimized in ranges of photon E_T and $|\eta|$, and has an identification efficiency that increases with E_T from 70% at 22 GeV to 90% above 50 GeV.

To further improve the rejection of jets misidentified as photons, the candidates are required to be isolated using information from both the calorimeter and tracking subsystems. The calorimeter isolation transverse energy $E_T^{\text{iso,calo}}$ is required to be smaller than $0.065E_T^\gamma$, where $E_T^{\text{iso,calo}}$ is defined as the sum of the transverse energies of positive-energy topological clusters [51] within a cone of size $\Delta R = 0.2$ around the photon

candidate, excluding the photon transverse energy E_T^γ and correcting for pile-up and underlying-event contributions [52–54]. The track isolation transverse energy $E_T^{\text{iso, trk}}$ is required to be less than $0.05E_T^\gamma$, where $E_T^{\text{iso, trk}}$ is defined as the scalar sum of the transverse momenta of tracks with $p_T > 1$ GeV in a $\Delta R = 0.2$ cone around the photon candidate, and which satisfy some loose track-quality criteria, are not associated with a photon conversion, and originate from the diphoton production vertex. The combined isolation efficiency for pairs of photons fulfilling the identification requirement in simulated signal samples increases with $m_{\gamma\gamma}$ from 80% at 10 GeV to 90% at 90 GeV.

The diphoton invariant mass is computed using the transverse energies of the leading and subleading photon candidates and their angular separation in both azimuth ϕ and pseudorapidity η , determined from their positions in the calorimeter and the diphoton production vertex.

An additional kinematic selection is placed on the transverse momentum of the diphoton system, $p_T^{\gamma\gamma}$, requiring events to have a diphoton pair with $p_T^{\gamma\gamma} > 50$ GeV. This requirement is motivated by the fact that the analysis targets diphoton pairs with low masses, down to about half the trigger energy thresholds, and such pairs are typically highly boosted with respect to the ATLAS detector rest frame. The $p_T^{\gamma\gamma}$ requirement is chosen in order to reach the best compromise between the statistical uncertainty in the lowest part of the spectrum and sculpting effects on the background shape from the trigger efficiency turn-on, the modelling of which would result in large systematic uncertainties.

In total, 1 166 636 data events with $m_{\gamma\gamma} < 80$ GeV are selected.

Following the detector-level selection, the measurement of the signal production cross-section is performed in a fiducial volume defined from the simulated samples by requiring two photons at particle level with $E_T > 22$ GeV, $|\eta| < 2.37$ and $p_T^{\gamma\gamma} > 50$ GeV. The particle isolation, defined as the scalar sum of the p_T of all the stable particles (except muons and neutrinos) found within a $\Delta R = 0.2$ cone around the photon direction, is required to be less than $0.05E_T^\gamma$. This isolation requirement is chosen to reproduce the detector-level selection.

5 Signal modelling

The shape of a possible signal in the diphoton invariant mass distribution is modelled by a double-sided Crystal Ball (DSCB) function, composed of a Gaussian core with power-law tails [1, 55], whose parameter values evolve linearly with respect to the mass m_X . The parameters of the DSCB function are extracted from fits to the MADGRAPH simulated signal samples. The width of the Gaussian core is entirely determined by the resolution of the detector and ranges from 0.2 to 1.2 GeV, as shown in Figure 1(a). Good agreement between the signal parameterization and the simulated signal samples is found, with differences below 1% of the fitted signal yield. An example of the simulated resonance overlaid with the signal model parameterization is shown in Figure 1(b).

6 Background estimates

The dominant background components consist of continuum $\gamma\gamma$ production, and of photon–jet (γj) and jet pair (jj) events where one or more jets are misidentified as photons. Other backgrounds arising from electrons faking photons in Z boson decays are found to be negligible in the mass range of this

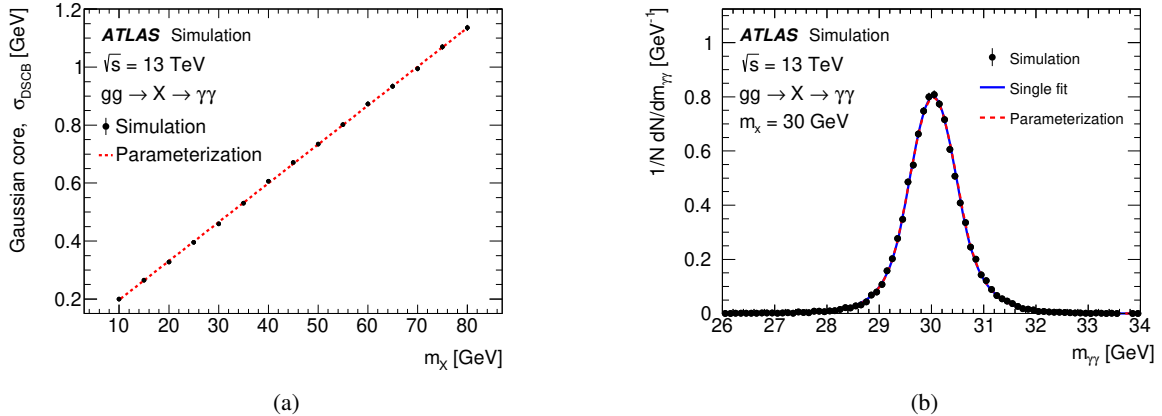


Figure 1: (a) Width of the Gaussian core of the DSCB function as a function of m_X (solid markers) overlaid with the linear parameterization (dashed line). (b) Simulated diphoton invariant mass distribution of a narrow-width signal resonance with $m_X = 30$ GeV (solid markers) overlaid with the DSCB function obtained from a single fit (solid line) and from the signal model parameterization (dashed line).

search and are not considered. The analysis makes use of a data-driven background estimate in which the continuum background shape is parameterized by an analytic function. The chosen analytic functional form is described in Section 6.1. The uncertainty arising from the choice of background model is based on signal-plus-background fits to background-only template histograms with a binning of 100 MeV, following the methodology described in Ref. [56] and further described in Section 6.1. The background modelling uncertainty is found to be dominated by the limited size of available simulated event samples. In order to reduce the impact of the background modelling uncertainty on the analysis, the background templates are smoothed using a Gaussian Processes fit. This technique is described further in Section 6.2, and its impact on the analysis is detailed in Section 6.3.

6.1 Background template modelling

The background-only template has two components. The $\gamma\gamma$ component is built from the simulated $\gamma\gamma$ samples described in Section 3, and the γj and jj components are built from control samples obtained from data, in which one or both photons must fail the *tight* identification requirements while passing a looser set of identification cuts. The two components are combined according to their relative fractions. The relative contribution of each of these processes is shown in Figure 2 and is estimated using the two-dimensional sideband method described in Ref. [57]. The purity of the diphoton sample, defined as the fraction of $\gamma\gamma$ events, increases with $m_{\gamma\gamma}$ from 50% at 10 GeV to 70% at 80 GeV with an overall uncertainty of 3% dominated by its statistical component arising from the limited size of the data sample collected with the prescaled triggers described in Section 4. No significant difference in the diphoton purity is observed between the various LHC data-taking periods of Run 2.

The goal of this analysis is to reach the lowest invariant mass possible, including the ‘turn-on’ region for masses below 20 GeV. The resulting background shape needs to be described by a more complex analytic form than in previous diphoton resonance searches [2, 26] and is constructed as the combination of two pieces: one capable of describing the turn-on shape and a second used to describe the smoothly decreasing

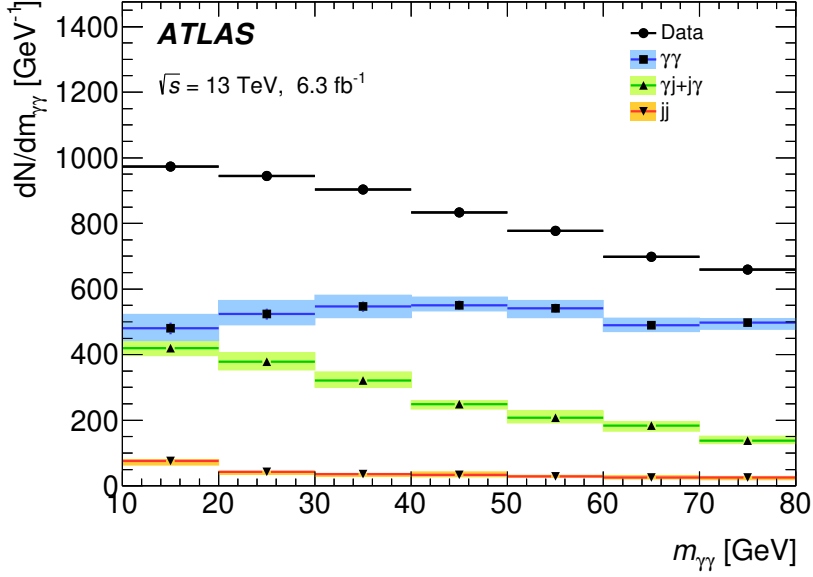


Figure 2: Diphoton invariant mass distributions of the data after event selection and their decomposition into contributions from diphoton ($\gamma\gamma$), photon+jet (γj and $j\gamma$) and dijet (jj) events as determined using the two-dimensional sideband method. The total uncertainties, including statistical and systematic components added in quadrature, are shown as error bands.

part. The two-parts analytic function described below was found to adequately model the background shape across the full invariant mass range of the search.

The turn-on region (TO) is described by the following function:

$$h_{\text{TO}}(m_{\gamma\gamma}; f_0, \tau_{\text{TO}}) = 1 - (1 - f_0) e^{-\frac{m_{\gamma\gamma}}{\tau_{\text{TO}}}},$$

where f_0 corresponds to the value of the function at $m_{\gamma\gamma} = 0$ and τ_{TO} is the length scale of the turn-on.

The smoothly falling region beyond 30 GeV is described by a power-law function multiplied by an ‘activation’ function to increase its flexibility in the high mass region (above 50 GeV). For this activation term, an exponential function times a ‘Fermi–Dirac-like’ function is chosen, and the total function is:

$$h_{\text{High}}(m_{\gamma\gamma}; c_1, a_0, c_0, \delta_{\text{tail}}, \tau_{\text{tail}}, \delta_{\text{thresh}}, \tau_{\text{thresh}}) = \underbrace{\left(1 - \left(\frac{m_{\gamma\gamma}}{c_1}\right)^{a_0}\right)^{c_0}}_{\text{Power-law}} \underbrace{\left(1 + \frac{e^{\frac{m_{\gamma\gamma} - \delta_{\text{tail}}}{\tau_{\text{tail}}}}}{1 + e^{-\frac{m_{\gamma\gamma} - \delta_{\text{thresh}}}{\tau_{\text{thresh}}}}}\right)}_{\text{Activation function}},$$

where c_1 , a_0 and c_0 are the parameters of the power-law term, and δ_{tail} , τ_{tail} , δ_{thresh} and τ_{thresh} describe the activation function. The power-law part is qualitatively described by its endpoints, being 1 at $m_{\gamma\gamma} = 0$, and 0 at $m_{\gamma\gamma} = c_1$, and a fixed value of $c_1 = 115$ GeV can be set with no impact on the flexibility of the complete model. The activation function only plays a role above δ_{thresh} , below which its value is practically 1.

The complete functional form is obtained by adding the two components and has ten parameters in total:

$$f(m_{\gamma\gamma}; \boldsymbol{\theta}_{\text{TO}}, \boldsymbol{\theta}_{\text{High}}, f_{\text{TO}}) = h_{\text{TO}}(\boldsymbol{\theta}_{\text{TO}}) \cdot f_{\text{TO}} + (1 - f_{\text{TO}}) \cdot h_{\text{High}}(\boldsymbol{\theta}_{\text{High}}), \quad (1)$$

where f_{TO} is the parameter that describes the relative contribution of the turn-on component and $\boldsymbol{\theta}_i$ are the sets of parameters belonging to h_{High} and h_{TO} . To reduce the potentially large correlations between the ten parameters, a subset of them are fixed to ensure the convergence of the fit. The choice of free parameters is based on the results of stability tests using generated pseudo-datasets from the best fit to the background template. The chosen configuration has the largest number of floating parameters, with only three fixed parameters ($c_1, \delta_{\text{thresh}}, \tau_{\text{thresh}}$) while the remaining seven parameters are free to vary.

Variations of the nominal background template are built to validate the flexibility of the chosen functional form and subsequently to estimate background modelling systematic uncertainties in Section 6.3. They are constructed by i) modifying the fractions of the background components, referred to as variations of the $\gamma\gamma$ fraction ($f_{\gamma\gamma}$); ii) varying the identification criteria used to define the γj and $j j$ control regions, referred to as variations of the control region; and iii) altering the templates by varying the $p_{\text{T}}^{\gamma\gamma}$ cut by 10%, referred to as $p_{\text{T}}^{\gamma\gamma}$ variations. These variations change the steepness of the turn-on by up to 20%, and the slope in the high mass region by $\pm 5\%$, with respect to the nominal background template.

6.2 Gaussian Processes to mitigate statistical fluctuations in background templates

The bias arising from the choice of background model is evaluated from signal-plus-background fits to background-only templates: any fitted signal yield N_{SS} is referred to as a ‘spurious signal’ and it is considered as a systematic uncertainty in the modelling of the background shape. The functional form of Eq. (1) provides an acceptable maximal spurious signal that is below 30% of the statistical uncertainty in the mass range between 10 and 75 GeV, and therefore it is chosen as the background model.

The estimation of this systematic uncertainty requires the shape of the background template to be as close as possible to the shape of the data distribution, but with an event count large enough for the statistical fluctuations of the template to be negligible. When evaluating the uncertainty, if the background model perfectly describes the representative background sample, then the number of signal events fitted by the signal-plus-background model will be zero. However, the representative background-only sample for this analysis is constructed using a limited number of simulated diphoton events, and the presence of statistical fluctuations in the sample introduces large statistical fluctuations in the number of fitted signal events, regardless of the quality of the background model. This issue was addressed previously [26] by using simulated datasets with much larger event counts than in data, which leads to smaller statistical fluctuations in the background shape but is computationally expensive. In order to meet the aforementioned requirements, an alternative approach to easing statistical effects on the background modelling uncertainty by using Gaussian Processes is followed instead.

A Gaussian Process (GP) is a flexible Bayesian machine-learning technique which may be used to obtain a non-parametric fit to an input dataset [58]. The analysis uses the scikit-learn GP implementation [59] to fit a GP to the representative background sample histogram; the posterior mean of the GP fit is used as a smoothed background template. The combined signal and background model fit is then performed on the smoothed template instead of the original representative background sample. The degree of smoothing applied is controlled through the choice of kernel and its hyperparameters. A radial basis function (RBF) kernel [58] with an additional constant noise component is utilized here. The RBF kernel includes a length-scale hyperparameter that encodes the correlation between the event counts for different bins in invariant mass. The contents of bins which are less than the length scale apart in invariant mass are highly

correlated, while the contents of those which are much further apart than the length scale are essentially uncorrelated. Physically, the length scale encodes a minimum feature size expected in the background shape, which in this analysis corresponds to the 1–2 GeV width of the trigger efficiency turn-on region and thus is also greater than the 100 MeV bin width of the original background histogram. The kernel hyperparameter values are determined in the fit to the representative background sample. Notably, because GPs are non-parametric in nature, the GP smoothing technique is not expected to significantly bias the shape of the resulting smoothed template towards a specific choice of analytic background model.

GPs may introduce mis-modelling at the edges of the diphoton invariant mass distribution, since edge points are only constrained by their correlation with other data points on one side. To mitigate these edge effects, the GP fit is performed using an extended invariant mass range of 7–80 GeV, while the combined signal and background model fit is performed using the nominal analysis invariant mass range of 9–77 GeV.

Figure 3 shows an example of a pseudo-dataset generated from the nominal background modelling function, as well as the GP-smoothed pseudo-dataset. The smoothed pseudo-dataset is observed to reproduce the nominal background modelling function shape with relatively high accuracy, and without the bin-scale statistical fluctuations of the original pseudo-dataset. The smoothed pseudo-dataset shows some oscillatory behaviour beyond the turn-on region; these features are an artefact of the steep slope in the turn-on region pulling the fitted length scale to a smaller value than would otherwise be needed to model the remaining mass range. Similarly, some mis-modelling of the smoothed pseudo-dataset is observed in the turn-on region because of the length scale being pulled to a larger value by the higher mass region. The magnitude of the fluctuations in the smoothed pseudo-dataset is significantly smaller than that of statistical fluctuations in the unsmoothed pseudo-dataset.

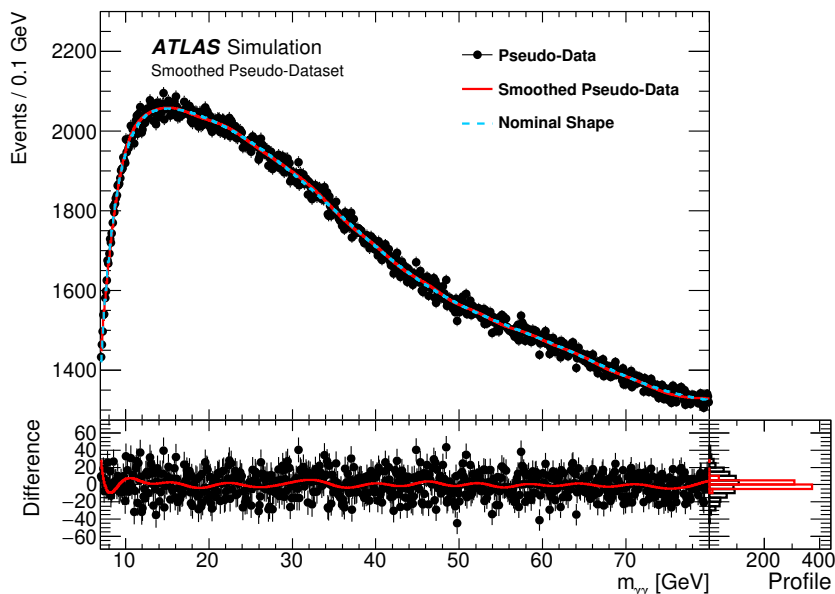


Figure 3: The upper panel shows a single pseudo-dataset (solid markers) generated from the nominal background modelling function described in Eq. (1) (blue dashed line). The GP-smoothed pseudo-dataset is shown with the red solid line. The bottom panel shows the difference between the unsmoothed and smoothed pseudo-datasets with respect to the nominal background shape. The horizontal axis of the plot utilizes a wider diphoton invariant mass range than the one used in the analysis in order to mitigate the impact of edge effects from the GP smoothing technique. The lower right panel shows the profile of the difference between the unsmoothed (black) and smoothed (red) pseudo-datasets with respect to the nominal background shape.

6.3 Impact of smoothing on the background modelling uncertainty

In order to verify that the GP smoothing technique does not introduce any significant bias into the background histogram shape, its effect is checked on an ensemble of pseudo-datasets generated from a known background shape. Ensembles are generated using both the nominal background modelling function (provided in Eq. (1)) and a set of analytic forms capable of describing the turn-on feature in the background shape. This set is composed of functional forms built similarly to the nominal functional form in which either the turn-on component or the smoothly falling component is replaced by other analytic forms, such as different Fermi–Dirac-like functions for the turn-on or sums of exponential functions for the smoothly falling component. The parameters of the analytic forms are determined by fitting them to the original simulated background sample, and each histogram in the ensemble is generated with the same effective event count as in the original simulated background sample.

The background modelling uncertainty is then evaluated for each pseudo-dataset, with and without smoothing. The aforementioned functional forms are used to probe for potential smoothing bias in both the cases where the analytic background model did or did not properly describe the pseudo-dataset. The bias that arises from the GP smoothing technique is defined as the difference between the observed spurious signals in the smoothed template and the unsmoothed template. The uncertainty associated with the GP smoothing technique is observed to be roughly 20% of the background modelling uncertainty for masses below 20 GeV, stabilizing at 5% for larger masses. This bias is added in quadrature to the background modelling uncertainty.

The final background modelling uncertainty is computed as the envelope of the maximal fitted signal yields over all the background template variations defined previously in this section after smoothing. Figure 4(a) shows the number of spurious-signal events N_{SS} , taken as the background modelling uncertainty, relative to its statistical uncertainty δS for the unsmoothed and smoothed templates. Applying the GP smoothing procedure to the background template leads to a reduction of at least 50% in this background modelling uncertainty relative to the unsmoothed case. The uncertainty arising from the GP smoothing technique is found to be small compared to the decrease in background modelling uncertainty due to the reduction of statistical noise. The magnitude of the smoothing uncertainty, as well as the remaining background modelling uncertainty, is presented as a function of the diphoton invariant mass in Figure 4(b).

7 Statistical analysis

The data are interpreted by following the statistical procedure described in Ref. [60]. A binned likelihood function is built from the observed diphoton invariant mass distribution and the analytic functions discussed in Sections 5 and 6, describing the signal and background components in the 9 to 77 GeV mass range. The search is performed in the 10 to 70 GeV mass range to avoid edge effects, based on the different diphoton invariant mass resolutions at these values, as illustrated in Figure 1(a).

The parameter of interest to be extracted from the likelihood fit is the fiducial production cross-section times branching ratio $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$. Since the measurement is performed in a fiducial volume (defined in Section 4) to allow easier reinterpretation of the results, the fiducial cross-section includes a correction factor C_X to account for the signal detection efficiency:

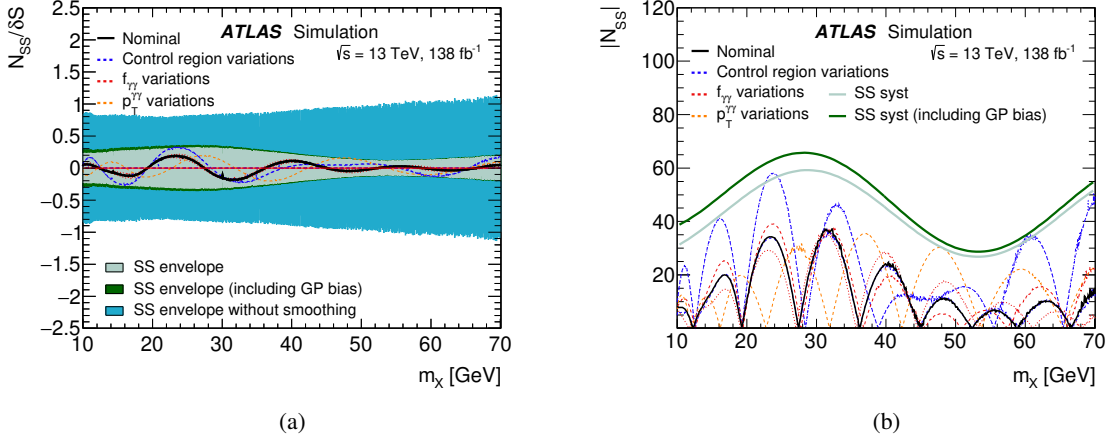


Figure 4: (a) Number of spurious-signal events N_{SS} relative to its uncertainty δS as a function of the resonance mass m_X for smoothed variations of the nominal template obtained from different control region definitions, diphoton purity variations and different kinematic selections. The light green line encloses the local maxima of the distribution. The dark green envelope includes the bias observed from the GP fit. The light blue envelope shows the relative background modelling uncertainty obtained from a parameterization of the unsmoothed template. (b) Absolute number of spurious-signal events $|N_{SS}|$ as a function of m_X evaluated for smoothed variations of the nominal template obtained from different control region definitions, diphoton purity variations and different kinematic selections. The light green line encloses the local maxima of the distribution. The dark green line includes the bias observed from the GP fit.

$$\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma) = \frac{N_S}{C_X \mathcal{L}}, \quad \text{with} \quad C_X = \frac{N_{\text{MC}}^{\text{det}}}{N_{\text{MC}}^{\text{fid}}},$$

where N_S is the number of signal events fitted in data, \mathcal{L} is the integrated luminosity, $N_{\text{MC}}^{\text{det}}$ is the number of reconstructed and selected signal events in the simulation and $N_{\text{MC}}^{\text{fid}}$ is the number of simulated signal events present within the fiducial volume. The C_X values are computed from the simulated signal samples described in Section 3 and range from 0.2 to 0.5 as a function of m_X .

The theoretical uncertainties affecting the measurement of $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$ arise from variations of the renormalization and factorization scales affecting the signal efficiencies evaluated in simulated samples. The experimental uncertainties directly impacting the signal yield include those involved in the luminosity determination, the modelling of pile-up interactions in simulation, the trigger efficiency, and photon identification and isolation. Another systematic uncertainty in the trigger arises from the boosted topology of the diphoton system and the capability of the trigger to identify two closely spaced electromagnetic showers. Uncertainties in the signal shape parameterization from the modelling and the determination of the photon energy resolution and scale are also accounted for, with mild impact on the signal yield.

The systematic uncertainties are implemented in the likelihood function as nuisance parameters constrained by Gaussian penalty terms, except for the background modelling systematic uncertainty, which is implemented as an additional signal component. All sources of systematic uncertainties are summarized in Table 1.

The compatibility of the observed data and the background-only hypothesis for a given signal hypothesis m_X is tested by estimating a local p -value based on a profile-likelihood-ratio test statistic, detailed in

Ref. [60]. The global significance of a given event excess is computed using background-only generated pseudo-datasets to account for the look-elsewhere effect [61].

In the absence of a signal, the expected and observed 95% confidence level (CL) exclusion limits on the cross-section times branching ratio are evaluated using the modified frequentist approach CL_s [62, 63] with the asymptotic approximation to the test-statistic distribution.

Table 1: Summary of the main sources of systematic uncertainty. Their impact on the fiducial production cross-section of a hypothetical resonant signal is shown, except for the background modelling uncertainty, which is expressed both as a number of events and relative to the expected statistical uncertainty δS of a fitted signal. Unless written otherwise, numbers are mass independent.

Source	Uncertainty
On $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$ [%]	
Pile-up modelling	± 3.5 (at 10 GeV) to ± 2 (beyond 15 GeV), mass dependent
Photon energy resolution	± 2.5 to ± 2.7 , mass dependent
Scale and PDFs uncertainties	± 2.5 to ± 0.5 , mass dependent
Trigger on closely spaced photons	± 2 (at 10 GeV) to < 0.1 (beyond 35 GeV), mass dependent
Photon identification	± 2.0
Isolation efficiency	± 2.0
Luminosity (2015–2018)	± 1.7
Trigger	± 1.0
Signal shape modelling	< 1
Photon energy scale	negligible
<i>Background modelling</i>	
Spurious signal (relative to δS)	30–65 events (10%–30%), mass dependent

8 Results

The diphoton invariant mass distribution of events passing the analysis selection is shown in Figure 5, along with the background-only fit performed in the 9 to 77 GeV mass range.

The result of the p -value scan as a function of the hypothesized resonance mass m_X is shown in Figure 6. The most significant deviation from the background-only hypothesis is observed for a mass of 19.4 GeV, corresponding to a local significance of 3.1σ . The global significance of such an excess is 1.5σ , computed using the methodology described in Section 7. Therefore, no significant deviations from the SM are observed.

The observed and expected upper limits on $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$ as a function of the resonance mass are shown in Figure 7. The expected upper limit is nearly constant across almost the whole invariant mass range of the search because of the competing effects of the increase of signal efficiencies and widening of the signal widths with increasing mass.

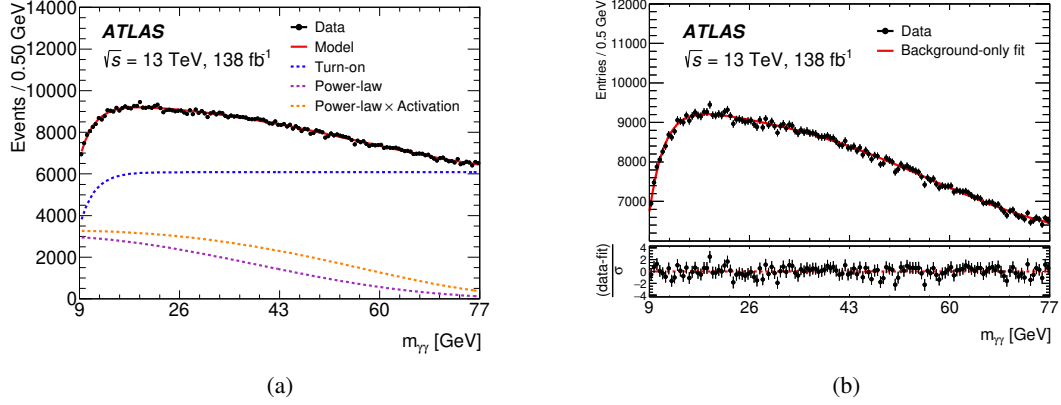


Figure 5: (a) Distribution of the diphoton invariant mass for all events passing the analysis selections in the full Run 2 dataset with the background-only fit superimposed. The functional form is decomposed into the different pieces described in detail in Section 6.1. (b) Vertically enlarged version of Figure 5(a). The normalized residuals between the data and the fit are shown in the bottom panel.

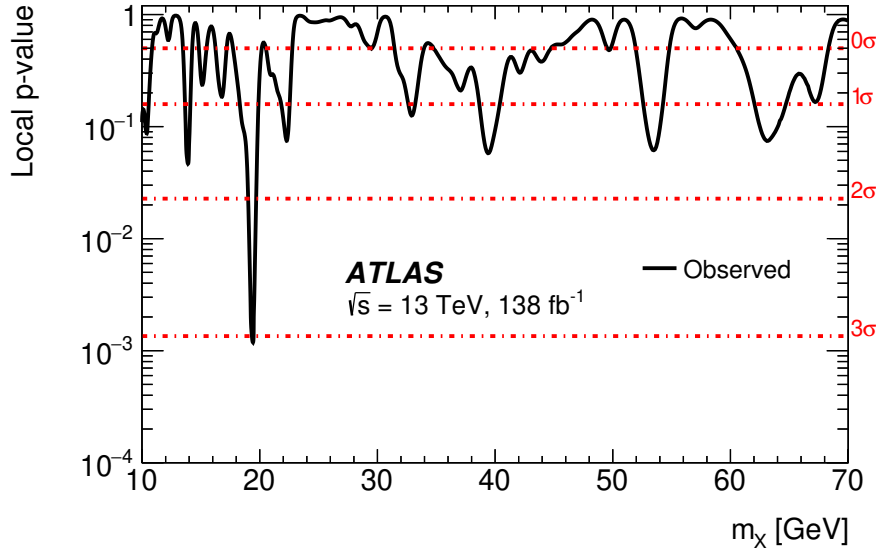


Figure 6: Scan of the observed p -value as a function of the diphoton invariant mass for the background-only hypothesis for the full Run 2 dataset.

9 Phenomenological interpretation

In this section the observed limits on the fiducial cross-section for a hypothetical resonance X are recast in the parameter space of an ALP a . For simplicity a ‘KSVZ-ALP’ is considered, inspired by the simplest QCD axion model [64–66], described by the following effective Lagrangian:

$$\frac{a}{4\pi f_a} \left[\alpha_3 c_3 G^a \tilde{G}^a + \alpha_2 c_2 W^i \tilde{W}^i + \alpha_1 c_1 B \tilde{B} \right] + \frac{1}{2} m_a^2 a^2, \quad (2)$$

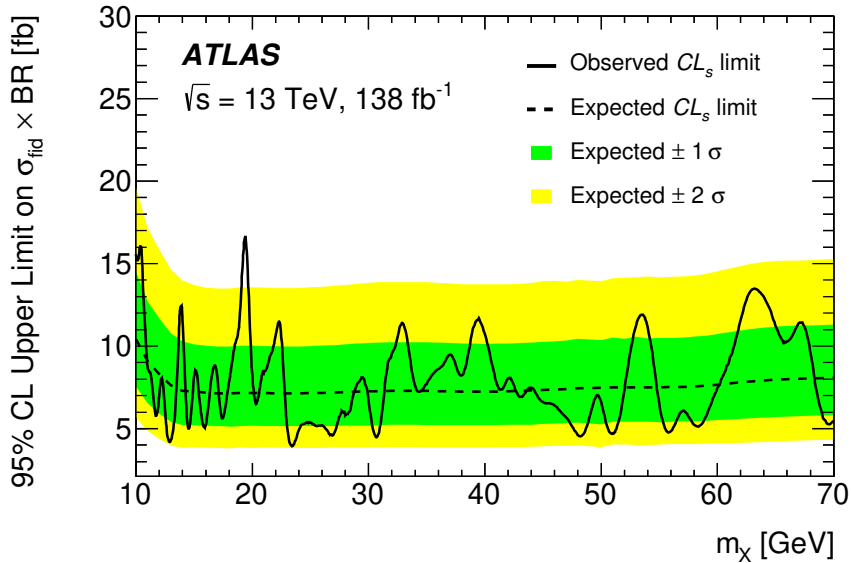


Figure 7: Expected and observed upper limits on the fiducial production cross-section times branching ratio to photon pairs for a scalar resonance in the NWA as a function of the resonance mass m_χ , for the full Run 2 dataset.

where a and m_a are the ALP field and its mass, and f_a corresponds to the decay constant that governs its coupling with the SM fields. The QCD and EW field strengths are denoted by G , W and B with $\tilde{F}^{\mu\nu} = (1/2) \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$ for all field strengths. The coupling constants α_3 , α_2 and $\alpha_1 = (5/3) \alpha_Y$ (where α_Y stands for the weak SM hypercharge Y coupling constant) set the strength of the strong and EW interactions in the SM. The coefficients c_i encode the anomalies of the global symmetry non-linearly realized by the ALP with the SM gauge group. These anomalies are generated by integrating out heavy fermions which are charged under the SM gauge group at the scale f_a . This Lagrangian is equivalent to the one used to generate the simulated signal samples described in Section 3.

The ALP a under consideration, being the pNGB of an approximate global symmetry, remains naturally light well below the scale of new physics. Considering m_a , the mass of the ALP, to be much smaller than m_Z , the relevant two-body decays of a are to photons and to jets, with widths which can be found, e.g., in Ref. [67]. In the 10 to 70 GeV mass range and for the choice of anomalies $c_1 = c_2 = c_3 = 10$, the branching ratio $\mathcal{B}(a \rightarrow \gamma\gamma)$ varies from $0.6 \cdot 10^{-3}$ to $1.6 \cdot 10^{-3}$. Choosing to set the magnitude of the c_i parameters to be the same is motivated by gauge coupling unification in a Grand Unified Theory scenario. While the specific value of 10 is arbitrary, the rescaling of the results to a different anomaly parameter choice would be trivial. The ALP Lagrangian in Eq. (2) is implemented in Feynrules [68], and the production cross-section at the LHC for the process $gg \rightarrow ag$ is computed at leading order with MADGRAPH [69], where the gluon is explicitly required to boost the ALP. A constant K -factor $K = 2$ is applied to this cross-section to account for NLO corrections, which were computed for a similar signal topology in Ref. [70]. ALPs that couple to gluons decay promptly over the entire mass range of interest for this study (recent studies of displaced ALP decays can be found in Refs. [70, 71]). Because of the large hierarchy between f_a and m_a , and the loop suppression of the coupling, the ALP total width is dominated by its coupling to gluons and is always small compared to its mass. As a consequence, the narrow-width approximation always applies and finite-width effects can be safely neglected.

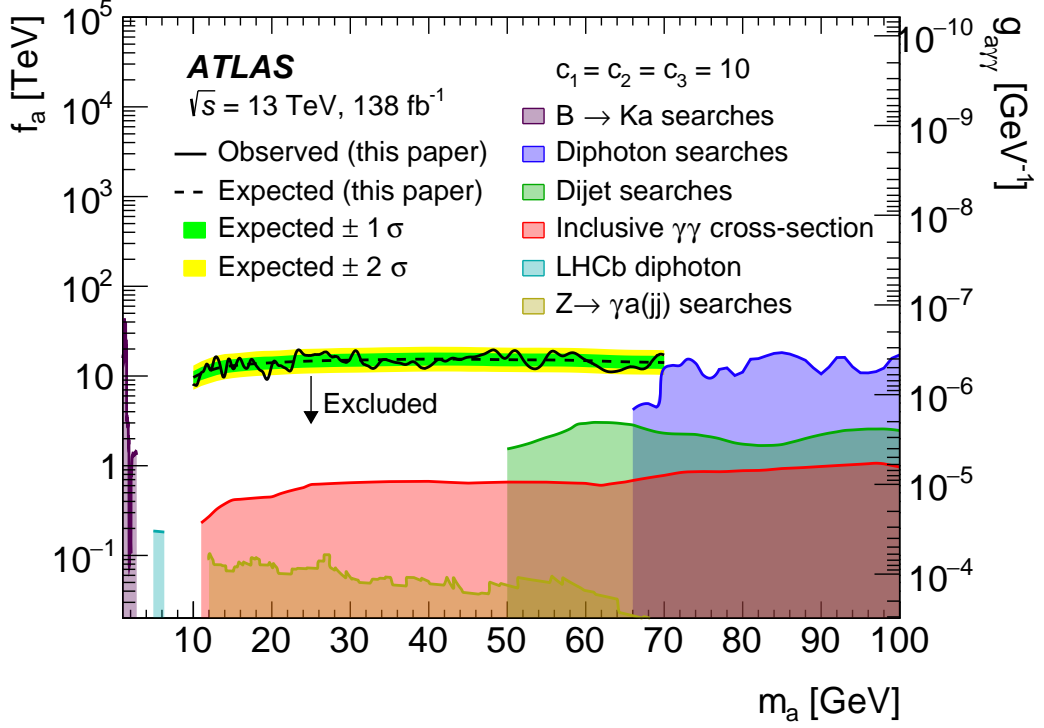


Figure 8: Parameter space of the ALP for $c_1 = c_2 = c_3 = 10$ (Eq. (2)). The observed and expected lower bounds on the ALP decay constant derived from this analysis are shown in black solid and dashed lines respectively. *BABAR* bounds on $B \rightarrow Ka$ derived in Ref. [72] are shown in purple; in green the LHC bounds on boosted dijet resonances [73] and in blue the LHC searches for diphoton resonances taken from Ref. [67]. The red bounds are derived from Tevatron [74] and LHC [57, 75, 76] diphoton cross-section measurements, following the method described in Ref. [67]. Weaker constraints covering lower invariant masses are obtained from LHCb diphoton measurements [77] and from LEP searches for $Z \rightarrow \gamma a(jj)$ [78], in cyan and yellow respectively. On the right, the y -axis shows the ALP–photon coupling $g_{a\gamma\gamma} \equiv \alpha_{\text{em}} E / \pi f_a$ ($E = c_2 + \frac{5}{3}c_1$), a standard QCD axion notation.

The recasting is done by comparing the theoretical signal yield obtained from the ALP model of Eq. (2), after applying the particle-level selection described in Section 4, with the bounds on the fiducial cross-section in Figure 7. The signal cross-section times branching ratio can be written as $1/f_a^2$ times a weakly varying function of the ALP mass. The upper limit on the cross-section then results in a lower limit on f_a , which is shown in Figure 8 for a specific choice of the c_i coefficients.

Figure 8 shows how the expected sensitivity of the search presented here covers a large portion of the unexplored ALP parameter space where the heavy colour states generating the ALP coupling to gauge bosons are in the multi-TeV range and therefore inaccessible at the LHC. Any production mechanism other than gluon–gluon fusion suffers from a smaller production cross-section, and the decoupling of the heavy states inducing the ALP coupling to SM states would require further study.

Constraints from $\Upsilon \rightarrow \gamma a(jj)$ [79], constraints from Z boson width measurements [80], and ALP production in light-by-light scattering in heavy-ion collisions [81, 82] are too weak to appear in the plot.

10 Conclusion

A search for new narrow-width boosted resonances is performed in the diphoton invariant mass spectrum ranging from 10 to 70 GeV, using 138 fb^{-1} of pp collision data collected at a centre-of-mass energy of 13 TeV with the ATLAS detector at the Large Hadron Collider. The data are consistent with the SM background expectation. Limits are set on the fiducial cross-section times branching ratio in a fiducial region defined to mimic the detector-level selection. The observed limits on $\sigma_{\text{fid}} \cdot \mathcal{B}(X \rightarrow \gamma\gamma)$ range from 4 to 17 fb, with variations mainly due to statistical fluctuations of the data. The dominant uncertainties arise from the limited number of pp collisions collected and the background modelling uncertainty. The impact of the latter is reduced by smoothing the simulated background-only sample with Gaussian Processes in order to reduce the statistical fluctuations in the sample. Furthermore, the observed limits are recast in the parameter space of an axion-like particle, covering a longstanding gap in diphoton resonance searches.

References

- [1] ATLAS Collaboration, *Search for Scalar Diphoton Resonances in the Mass Range 65–600 GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV*, *Phys. Rev. Lett.* **113** (2014) 171801, arXiv: [1407.6583 \[hep-ex\]](#).
- [2] ATLAS Collaboration, *Search for resonances decaying into photon pairs in 139fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **822** (2021) 136651, arXiv: [2102.13405 \[hep-ex\]](#).
- [3] CMS Collaboration, *Search for diphoton resonances in the mass range from 150 to 850 GeV in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **750** (2015) 494, arXiv: [1506.02301 \[hep-ex\]](#).
- [4] CMS Collaboration, *Search for physics beyond the standard model in high-mass diphoton events from proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **98** (2018) 092001, arXiv: [1809.00327 \[hep-ex\]](#).
- [5] B. Holdom and M. E. Peskin, *Raising the Axion Mass*, *Nucl. Phys. B* **208** (1982) 397.
- [6] K. Choi, C. W. Kim and W. K. Sze, *Mass Renormalization by Instantons and the Strong CP Problem*, *Phys. Rev. Lett.* **61** (1988) 794.
- [7] B. Holdom, *Strong QCD at high energies and a heavy axion*, *Phys. Lett. B* **154** (1985) 316, [Erratum: *Phys.Lett.B* 156, 452 (1985)].
- [8] M. Dine and N. Seiberg, *String theory and the strong CP problem*, *Nucl. Phys. B* **273** (1986) 109.
- [9] J. M. Flynn and L. Randall, *A computation of the small instanton contribution to the axion potential*, *Nucl. Phys. B* **293** (1987) 731.
- [10] V. A. Rubakov, *Grand unification and heavy axion*, *JETP Lett.* **65** (1997) 621, arXiv: [hep-ph/9703409](#).
- [11] K. Choi and H. D. Kim, *Small instanton contribution to the axion potential in supersymmetric models*, *Phys. Rev. D* **59** (1999) 072001, arXiv: [hep-ph/9809286](#).
- [12] Z. Berezhiani, L. Gianfagna and M. Giannotti, *Strong CP problem and mirror world: The Weinberg-Wilczek axion revisited*, *Phys. Lett. B* **500** (2001) 286, arXiv: [hep-ph/0009290](#).
- [13] A. Hook, *Anomalous Solutions to the Strong CP Problem*, *Phys. Rev. Lett.* **114** (2015) 141801, arXiv: [1411.3325 \[hep-ph\]](#).
- [14] H. Fukuda, K. Harigaya, M. Ibe and T. T. Yanagida, *Model of visible QCD axion*, *Phys. Rev. D* **92** (2015) 015021, arXiv: [1504.06084 \[hep-ph\]](#).
- [15] S. Dimopoulos, A. Hook, J. Huang and G. Marques-Tavares, *A collider observable QCD axion*, *JHEP* **11** (2016) 052, arXiv: [1606.03097 \[hep-ph\]](#).
- [16] P. Agrawal and K. Howe, *Factoring the strong CP problem*, *JHEP* **12** (2018) 029, arXiv: [1710.04213 \[hep-ph\]](#).
- [17] M. K. Gaillard, M. B. Gavela, R. Houtz, P. Quilez and R. Del Rey, *Color unified dynamical axion*, *Eur. Phys. J. C* **78** (2018) 972, arXiv: [1805.06465 \[hep-ph\]](#).
- [18] A. E. Nelson and N. Seiberg, *R symmetry breaking versus supersymmetry breaking*, *Nucl. Phys. B* **416** (1994) 46, arXiv: [hep-ph/9309299](#).

- [19] H.-S. Goh and M. Ibe, *R-axion detection at LHC*, *JHEP* **03** (2009) 049, arXiv: [0810.5773 \[hep-ph\]](#).
- [20] B. Bellazzini, A. Mariotti, D. Redigolo, F. Sala and J. Serra, *R-Axion at Colliders*, *Phys. Rev. Lett.* **119** (2017) 141804, arXiv: [1702.02152 \[hep-ph\]](#).
- [21] G. Ferretti and D. Karateev, *Fermionic UV completions of composite Higgs models*, *JHEP* **03** (2014) 077, arXiv: [1312.5330 \[hep-ph\]](#).
- [22] A. Belyaev et al., *Singlets in composite Higgs models in light of the LHC 750 GeV diphoton excess*, *Phys. Rev. D* **94** (2016) 015004, arXiv: [1512.07242 \[hep-ph\]](#).
- [23] G. Ferretti, *Gauge theories of Partial Compositeness: Scenarios for Run-II of the LHC*, *JHEP* **06** (2016) 107, arXiv: [1604.06467 \[hep-ph\]](#).
- [24] J. Mardon, Y. Nomura and J. Thaler, *Cosmic signals from the hidden sector*, *Phys. Rev. D* **80** (2009) 035013, arXiv: [0905.3749 \[hep-ph\]](#).
- [25] K. S. Jeong, T. H. Jung and C. S. Shin, *Axionic electroweak baryogenesis*, *Phys. Lett. B* **790** (2019) 326, arXiv: [1806.02591 \[hep-ph\]](#).
- [26] ATLAS Collaboration, *Search for resonances in the 65 to 110 GeV diphoton invariant mass range using 80 fb⁻¹ of pp collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2018-025, 2018, URL: <https://cds.cern.ch/record/2628760>.
- [27] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [28] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: <https://cds.cern.ch/record/1291633>, Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: <https://cds.cern.ch/record/1451888>.
- [29] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, *JINST* **13** (2018) T05008, arXiv: [1803.00844 \[physics.ins-det\]](#).
- [30] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, arXiv: [1908.00005 \[hep-ex\]](#).
- [31] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317, arXiv: [1611.09661 \[hep-ex\]](#).
- [32] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATLAS-SOFT-PUB-2021-001, 2021, URL: <https://cds.cern.ch/record/2767187>.
- [33] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003, arXiv: [1911.04632 \[physics.ins-det\]](#).
- [34] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: <https://cds.cern.ch/record/2677054>.
- [35] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47, arXiv: [1909.00761 \[hep-ex\]](#).
- [36] P. Artoisenet et al., *A framework for Higgs characterisation*, *JHEP* **11** (2013) 043, arXiv: [1306.6464 \[hep-ph\]](#).

- [37] 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 \[hep-ph\]](#).
- [38] R. D. Ball et al., *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](#).
- [39] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419>.
- [40] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, arXiv: [0710.3820 \[hep-ph\]](#).
- [41] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](#).
- [42] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034, arXiv: [1905.09127 \[hep-ph\]](#).
- [43] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, arXiv: [1111.1220 \[hep-ph\]](#).
- [44] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](#).
- [45] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, arXiv: [hep-ph/0109231](#).
- [46] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, arXiv: [0903.1219 \[hep-ph\]](#).
- [47] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568 \[physics.ins-det\]](#).
- [48] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [49] 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>.
- [50] ATLAS Collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, *Phys. Rev. D* **90** (2014) 112015, arXiv: [1408.7084 \[hep-ex\]](#).
- [51] 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 \[hep-ex\]](#).
- [52] ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Phys. Rev. D* **83** (2011) 052005, arXiv: [1012.4389 \[hep-ex\]](#).
- [53] M. Cacciari, G. P. Salam and G. Soyez, *The catchment area of jets*, *JHEP* **04** (2008) 005, arXiv: [0802.1188 \[hep-ph\]](#).

- [54] M. Cacciari, G. P. Salam and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065, arXiv: [0912.4926](https://arxiv.org/abs/0912.4926) [[hep-ph](#)].
- [55] M. Oreglia, *A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$* , PhD thesis, 1980, URL: <https://www.slac.stanford.edu/cgi-bin/getdoc/slac-r-236.pdf>.
- [56] ATLAS Collaboration, *Recommendations for the Modeling of Smooth Backgrounds*, ATL-PHYS-PUB-2020-028, 2020, URL: <https://cds.cern.ch/record/2743717>.
- [57] ATLAS Collaboration, *Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **01** (2013) 086, arXiv: [1211.1913](https://arxiv.org/abs/1211.1913) [[hep-ex](#)].
- [58] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2005, ISBN: 026218253X.
- [59] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, *Journal of Machine Learning Research* **12** (2012) 2825, URL: <https://arxiv.org/abs/1201.0490>.
- [60] 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](https://arxiv.org/abs/1007.1727) [[physics.data-an](#)], Erratum: *Eur. Phys. J. C* **73** (2013) 2501.
- [61] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *Eur. Phys. J. C* **70** (2010) 525, arXiv: [1005.1891](https://arxiv.org/abs/1005.1891) [[physics.data-an](#)].
- [62] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instrum. Meth. A* **434** (1999) 435, arXiv: [hep-ex/9902006](https://arxiv.org/abs/hep-ex/9902006).
- [63] A. L. Read, *Presentation of search results: the CL_s technique*, *J. Phys. G* **28** (2002) 2693.
- [64] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440.
- [65] J. E. Kim, *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103.
- [66] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Can confinement ensure natural CP invariance of strong interactions?*, *Nucl. Phys. B* **166** (1980) 493.
- [67] A. Mariotti, D. Redigolo, F. Sala and K. Tobioka, *New LHC bound on low-mass diphoton resonances*, *Phys. Lett. B* **783** (2018) 13, arXiv: [1710.01743](https://arxiv.org/abs/1710.01743) [[hep-ph](#)].
- [68] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250, arXiv: [1310.1921](https://arxiv.org/abs/1310.1921) [[hep-ph](#)].
- [69] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128, arXiv: [1106.0522](https://arxiv.org/abs/1106.0522) [[hep-ph](#)].
- [70] Y. Gershtein, S. Knapen and D. Redigolo, *Probing naturally light singlets with a displaced vertex trigger*, *Phys. Lett. B* **823** (2021) 136758, arXiv: [2012.07864](https://arxiv.org/abs/2012.07864) [[hep-ph](#)].
- [71] A. Hook, S. Kumar, Z. Liu and R. Sundrum, *High Quality QCD Axion and the LHC*, *Phys. Rev. Lett.* **124** (2020) 221801, arXiv: [1911.12364](https://arxiv.org/abs/1911.12364) [[hep-ph](#)].

- [72] S. Chakraborty, M. Kraus, V. Loladze, T. Okui and K. Tobioka, *Heavy QCD axion in $b \rightarrow s$ transition: Enhanced limits and projections*, *Phys. Rev. D* **104** (2021) 055036, arXiv: [2102.04474 \[hep-ph\]](#).
- [73] CMS Collaboration, *Search for low mass vector resonances decaying into quark–antiquark pairs in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **01** (2018) 097, arXiv: [1710.00159 \[hep-ex\]](#).
- [74] CDF Collaboration, *Measurement of the Cross Section for Prompt Isolated Diphoton Production Using the Full CDF Run II Data Sample*, *Phys. Rev. Lett.* **110** (2013) 101801, arXiv: [1212.4204 \[hep-ex\]](#).
- [75] CMS Collaboration, *Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **74** (2014) 3129, arXiv: [1405.7225 \[hep-ex\]](#).
- [76] ATLAS Collaboration, *Measurement of the production cross section of pairs of isolated photons in pp collisions at 13 TeV with the ATLAS detector*, *JHEP* **11** (2021) 169, arXiv: [2107.09330 \[hep-ex\]](#).
- [77] X. Cid Vidal, A. Mariotti, D. Redigolo, F. Sala and K. Tobioka, *New axion searches at flavor factories*, *JHEP* **01** (2019) 113, [Erratum: *JHEP* 06, 141 (2020)], arXiv: [1810.09452 \[hep-ph\]](#).
- [78] L3 Collaboration, *Isolated hard photon emission in hadronic Z^0 decays*, *Phys. Lett. B* **292** (1992) 472.
- [79] BaBar Collaboration, *Search for Hadronic Decays of a Light Higgs Boson in the Radiative Decay $\Upsilon \rightarrow \gamma A^0$* , *Phys. Rev. Lett.* **107** (2011) 221803, arXiv: [1108.3549 \[hep-ex\]](#).
- [80] Particle Data Group, *Review of Particle Physics*, *Chin. Phys. C* **40** (2016) 100001.
- [81] ATLAS Collaboration, *Measurement of light-by-light scattering and search for axion-like particles with 2.2 nb^{-1} of $Pb+Pb$ data with the ATLAS detector*, *JHEP* **03** (2021) 243, arXiv: [2008.05355 \[hep-ex\]](#), Erratum: *JHEP* **11** (2021) 050.
- [82] CMS Collaboration, *Evidence for light-by-light scattering and searches for axion-like particles in ultraperipheral $PbPb$ collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Lett. B* **797** (2019) 134826, arXiv: [1810.04602 \[hep-ex\]](#).