ATLAS CONF Note

ATLAS-CONF-2023-010

23rd March 2023

Search for periodic signals in the dielectron and diphoton invariant mass spectra using 139 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector

The ATLAS Collaboration

A search for physics beyond the Standard Model inducing periodic signals in the dielectron and diphoton invariant mass spectra is presented using 139 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collision data collected by the ATLAS experiment at the LHC. Novel search techniques based on continuous wavelet transforms are used to infer the frequency of periodic signals from the invariant mass spectra and neural network classifiers are used to enhance the sensitivity to periodic resonances. In the absence of a signal, exclusion limits are placed at the 95% confidence level in the two-dimensional parameter space of the clockwork gravity model. Model-independent searches for deviations from the background-only hypothesis are also performed.

ATLAS-CONF-2023-010
31 March 2023

1 Introduction

Typical experimental searches for physics beyond the Standard Model (SM) at the Large Hadron Collider (LHC) search for new resonant peaks or non-resonant deviations in kinematic distributions, such as the invariant mass of two-particle systems. However, more complicated signatures are possible and have received little attention in experimental searches so far. One such possibility, which appears in several theoretical extensions to the SM, is a signal that gives rise to periodic structures. This type of signal is predicted by models with small mass splittings between a large number of resonances.

Extensions to the SM with these features include the continuum clockwork gravity model [1, 2], which has an identical five-dimensional spacetime metric to the linear dilaton scenario [3, 4]. This spacetime metric also approximates the dual of Little String Theory [5, 6]. As the same theory can be interpreted as either a linear dilaton model or as the continuum version of a clockwork model, this theory is referred to as the Clockwork/Linear Dilaton (CW/LD) model. These models are well motivated, as the continuum spacetime version of the clockwork gravity model can address the Higgs naturalness problem [1]. The CW/LD model is also related to previous proposed solutions to the hierarchy problem, including theories with large extra dimensions [7, 8] and Randall-Sundrum (RS) models where the extra dimension is warped [9].

The CW/LD model predicts a narrowly-spaced spectrum of resonances in mass. The mass spectrum and couplings of this tower of Kaluza-Klein (KK) gravitons are described by two parameters, k and M_5. The parameter k is a mass parameter that determines the onset of the KK graviton spectrum and M_5 is the 5D reduced Planck mass, the fundamental scale of the theory. As the cross-section for the CW/LD model approximately scales inversely with M_5, $\sigma \propto 1/M_5^3$, a limit on the signal strength in this search can be directly translated to a limit on M_5 for a given value of k.

When looking at the invariant mass of the decay products of these resonances in experiments, the narrowly-spaced spectrum of resonances predicted by the model may appear as a long-range semi-periodic structure. The narrow spacing of the mass spectrum suggests that an excellent detector resolution is needed to resolve these resonances, a requirement which can be satisfied in final states with two electrons or two photons. This paper presents a search for periodic contributions to the diphoton and dielectron invariant mass spectra using 139 fb^{-1} of $\sqrt{s}$ = 13 TeV proton-proton (pp) collision data recorded by the ATLAS detector from 2015 to 2018 at the LHC. Previous searches for physics beyond the SM in high-mass diphoton final states and dielectron final states during Run 2 of the LHC using pp collisions at $\sqrt{s}$ = 13 TeV have been conducted by the ATLAS and CMS experiments [10–14]. The CMS experiment performed a search for the same CW/LD signal considered in this analysis, using only diphoton events, in a search for a non-resonant excess at high invariant masses. The CMS result excludes values of M_5 in the range 10 TeV to 1 TeV for values of k in the range 0.1 GeV to 5 TeV [13].

2 Analysis strategy

This analysis is based on two previous ATLAS resonance searches performed in the dilepton [10] and diphoton [12] channels in the mass range of 225–6000 GeV and 150–5000 GeV, respectively. Particularly, this analysis uses the same dataset and background estimation methods, which in both cases were performed using data-driven techniques. The two previous analyses also estimated the background from simulation and found an excellent compatibility with the estimations from data. This analysis, therefore, takes advantage of the precise data-driven background estimation while also benefitting from the availability of a

simulation-based background estimation. The background estimation from simulation includes a complete set of systematic uncertainties in both analysis channels, which is important for this analysis because the "spurious-signal" uncertainty derived in Refs. [10, 12] is applicable only to searches which consider a single narrow resonance. The simulation-based uncertainties from the previous searches are used instead and their effect on the background estimation is incorporated using the methodology developed in a previous ATLAS search for non-resonant signatures in the dilepton channel [11]. Finally, this analysis also relies on transformational and statistical methods developed in Ref. [15] to deal with data featuring periodic structures.

The background estimation methods are discussed briefly in Section 7. The sources of systematic uncertainties are discussed in Section 8, while the uncertainty estimation process is discussed in Section 9.

This analysis performs a search for generic periodic features in the dielectron and diphoton invariant mass spectra, in the same mass ranges considered in Refs. [10, 12]. The dielectron and diphoton channels are analysed separately due to potential overlaps in their event selections. This search uses a continuous wavelet transform (CWT) to analyse these mass spectra in the frequency domain.[1]. The output of the CWT is a two-dimensional image of the wavelet amplitude in the frequency versus mass space, referred to as a scalogram. The potential periodicity of a signal can be revealed in the image, for example as a local "blob" around a small range of frequencies and/or masses. While conventional resonant and non-resonant searches may be suboptimal, particularly in the case of small CW/LD-like signals, the CWT may still provide an enhancement of a new signal's semi-periodic nature. This enhancement provides a clear separation of a signal from the background and suggests that the search should be conducted in the CWT space rather than in the mass space alone. The CWT method is described briefly in Section 10.

In the limit of infinite statistical precision, the local features of a potential signal in the CWT images can be clearly visible and separable from the background-only case. When accounting for realistic statistical fluctuations due to the finite-size distributions in the mass space, a level of blurring is imposed to the CWT images and the separation power becomes significantly smaller. Therefore, machine-learning techniques are applied to distinguish periodic contributions to the diphoton and dielectron CWT images made using the invariant mass spectra. Model-independent results are provided using an autoencoder-based anomaly detection procedure to search for generic periodic deviations in the scalograms. The CW/LD model is used as a benchmark model for a model-specific search using the CWT images. A neural network binary classifier is trained on the CWT images made from background-only and signal-plus background mass distributions to provide a test statistic for discovery of potential periodic resonances with specific k and M_5 values. In both the model-independent and model-specific searches, the test-statistic used during the statistical inference portion of the analysis is based on the machine-learning outputs themselves. As the CW/LD signal may introduce a non-resonant enhancement above the SM background in the mass space, the statistical inference in the two searches can be dominated by the non-resonant signal features instead of the semi-periodic contribution. Therefore, the two searches are performed with and without specific thresholds imposed to reduce potential non-resonant contributions and focus the sensitivity on the periodic features. The statistical analysis of the two methods and the related thresholds are discussed in Section 11. The model-independent results and the model-specific exclusion limits set in the k–M_5 plane of the CW/LD model are given in Section 12.

[1] In this analysis, the transformed variable is the invariant mass rather than time.

3 ATLAS detector

The ATLAS experiment [16–18] is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and nearly 4π coverage in solid angle.[2] It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner detector (ID) covers the pseudorapidity range $|\eta| < 2.5$ and 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 steel/scintillating-tile hadron sampling calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters, covering the region $|\eta| < 2.7$ and is based on three large superconducting air-core toroidal magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector acceptance. A two-level trigger system is used to select events. The first-level trigger is implemented in custom hardware and uses a subset of the detector information to accept events at a rate below 100 kHz. A software-based trigger then reduces the accepted event rate to an average of 1 kHz for offline storage [19]. An extensive software suite [20] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

4 Data and simulated event samples

The dataset used for the search consists of the pp collision data recorded by the ATLAS experiment at $\sqrt{s} = 13$ TeV during the 2015 to 2018 LHC data-taking period. After requiring stable beam conditions and data quality selections with all ATLAS subsystems operational [21], the dataset corresponds to an integrated luminosity of 139.0 ± 2.4 fb^{-1} [22]. The LUCID-2 detector [23] was used for the primary luminosity measurements.

The events used in this analysis were recorded using a set of diphoton and dielectron triggers. Events in the diphoton channel were recorded using a diphoton trigger that required at least two energy clusters in the EM calorimeter with transverse energies (E_{T}) greater than 35 and 25 GeV for the E_{T}-ordered leading and subleading photon candidates, respectively. Both clusters were required to fulfil photon identification criteria based on the shower shapes in the EM calorimeters. The triggers used in 2015 and 2016 required both photons to satisfy the *loose* identification requirement [24]. In 2017 and 2018, due to the greater instantaneous luminosity, the diphoton trigger requirements were tightened and both photons were required to fulfil the *medium* identification requirement. The efficiency of the diphoton trigger with respect to the event selection given in Section 5 is over 99% for the 2015–2016 data and above 98% for the 2017-2018 data [25].

Events in the dielectron channel were collected using several dielectron triggers. The trigger used in 2015 required both electrons to pass the *loose* identification criteria and E_{T} thresholds of 12 GeV for both electrons. In 2016, the E_{T} thresholds were increased to 17 GeV for both electrons and the electrons were

[2] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, 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 $\eta = -\ln\tan(\theta/2)$.

required to pass the *very loose* identification criteria. In 2017 and 2018, the identification criteria were left unchanged and the E_{T} thresholds were increased to 24 GeV for both electrons [25].

Although the background in this analysis is estimated using data-driven methods, simulated Monte Carlo (MC) events are used to optimise the analysis selections, determine appropriate fit functions for the data, estimate background compositions, and evaluate signal acceptances and efficiencies. Interference effects between the resonant signals and all background processes in both the diphoton and dielectron channels are neglected in this analysis.

For the diphoton channel, the largest background comes from the production of two prompt photons which represents the irreducible background in this search channel. Smaller background contributions come from events containing a photon and a jet and events with two jets, where the jets are misidentified as photons. These smaller backgrounds are estimated using a data-driven technique, the two-dimensional sideband method, described in Ref. [26].

Events with two prompt photons were simulated using the Sherpa 2.2.4 [27, 28] event generator. Matrix elements were calculated with up to one additional parton at next-to-leading-order (NLO) and including two or three additional partons at leading order (LO) in perturbative QCD (pQCD). These matrix elements were merged with the Sherpa parton-shower simulation using the ME+PS@NLO prescription [29–32]. The NNPDF3.0nnlo parton distribution function (PDF) set [33] was used and paired with a dedicated parton-shower tune in the Sherpa generator.

For the dielectron channel, the main prompt backgrounds arise from Drell-Yan (DY), top-quark pair ($t\bar{t}$), single-top-quark, and diboson production. The background contribution from non-prompt electrons from multi-jet and W+jets processes is estimated using a data-driven technique, the matrix method, as described in Ref. [34].

The DY [35] sample was generated using the Powheg Box v1 [36–39] event generator with the CT10 PDF set [40] and interfaced with the Pythia 8.186 [41] parton shower program. Next-to-next-to-leading-order (NNLO) corrections in pQCD and NLO corrections in electroweak (EW) theory were calculated and applied to the simulated DY events. The pQCD corrections were computed with VRAP v0.9 [42] and the CT14 NNLO PDF set [43]. The EW corrections were computed with MCSANC [44] which accounts for quantum electrodynamic effects due to initial state radiation, interference between initial and final state radiation, and Sudakov logarithm single-loop corrections.

The diboson [45] processes with fully leptonic and semi-leptonic final states were simulated using Sherpa 2.2.1 with the CT10 PDF set. Matrix elements were calculated at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. The matrix element calculations were matched and merged with the Sherpa parton shower based on Catani–Seymour dipole factorisation [46, 47] using the ME+PS@NLO prescription. The diboson and DY backgrounds were generated in slices of dilepton mass in order to enhance the MC statistical precision in the high-mass region.

The $t\bar{t}$ and single-top-quark samples were generated with Powheg Box v2 [36–38, 48–51] at NLO using the NNPDF3.0nlo PDF [33] in the matrix element and interfaced to Pythia 8.230 [52] for modelling the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune [53] and using the NNPDF2.3lo set of PDFs [54]. These samples were normalised to the theoretical cross-sections calculated at NNLO in pQCD and include resummation of the next-to-next-to-leading logarithmic soft gluon terms as provided by Top++ 2.0 [55].

Resonant single-graviton MC samples were generated using a Randall-Sundrum model [9]. These MC samples were generated at LO in pQCD using Pythia 8 with the NNPDF2.3lo PDF set and the A14 tune. In these samples, only the lightest KK graviton excitation was generated. For the diphoton decay channel, the samples were generated with a KK graviton mass m_{G^*} in the range 150–5000 GeV. A fixed coupling value of $k/\overline{M}_{\mathrm{Pl}} = 0.01$, where $\overline{M}_{\mathrm{Pl}} = M_{\mathrm{Pl}}/\sqrt{8\pi}$ is the reduced Planck scale, is used to ensure a sufficiently narrow-width signal. For the dielectron decay channel, samples were generated with masses in the range 200–6000 GeV and with a fixed coupling value of $k/\overline{M}_{\mathrm{Pl}} = 0.1$. These samples are used to determine the acceptance and efficiency of selecting the CW/LD signal.

Additional MC samples incorporating a series of RS graviton resonances were generated at LO in pQCD using Pythia 8.244 with the NNPDF2.3lo PDF set and the A14 tune. In these samples, only the direct graviton decays to dielectron and diphoton final states were simulated. Samples were generated for k values of 300 GeV, 500 GeV, 1 TeV, and 2 TeV with M_5 fixed to a value of 6 TeV. These samples are used to validate the analytic signal templates of the CW/LD signals, which are discussed in detail in Section 6.

The effects of multiple pp interactions in the same bunch crossing as the hard scatter and in neighbouring bunch crossings (pile-up) are included in all simulated samples. Pile-up collisions were generated with Pythia 8.186 using the NNPDF2.3lo PDF set and the ATLAS A3 set of tuned parameters [56]. Simulated event samples were weighted to reproduce the distribution of the average number of interactions per bunch crossing observed in data [57].

The spin-2 diphoton simulated signal events, and DY, $t\bar{t}$, single-top-quark, and diboson background events were processed using a detailed simulation of the ATLAS detector [58] based on Geant4 [59]. The irreducible prompt $\gamma\gamma$ background and spin-2 dielectron signals were processed using a fast simulation of the ATLAS detector [60], where the full simulation of the calorimeter is replaced with a fast parameterisation of the calorimeter response. All simulated events were reconstructed with the same reconstruction algorithms as those used for data.

Generator-level-only MC samples of the NLO DY background samples are used for the modelling studies described in Section 7. These samples could not be produced with the ATLAS detector simulation due to the large number of events required [10].

5 Object and event selection

The complete description of the object definitions and event selections are given in Ref. [10] for the dielectron channel and Ref. [12] for the diphoton channel. The criteria are identical to the ones applied in this analysis and a brief description is given here.

The event selection in the diphoton channel requires at least two reconstructed photon candidates with $E_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.37$, excluding candidates in the transition region $1.37 < |\eta| < 1.52$ between the barrel and endcap EM calorimeters. The two highest-E_{T} photons are used to form the diphoton candidate and are used with additional information from the tracking detectors to identify the diphoton primary vertex [61]. After the diphoton primary vertex is identified, the leading and subleading photons are required to have $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ and 0.25 respectively. The diphoton invariant mass, $m_{\gamma\gamma}$, is required to be greater than 150 GeV. To reduce the background from jets, photon candidates are required to pass the *tight* identification criteria based on shower shapes in the EM calorimeter [24].

Events in the dielectron channel are selected by requiring at least one pair of reconstructed electron candidates. Each event is required to contain at least one reconstructed pp interaction vertex, where the primary vertex is defined as the vertex with the highest sum of track transverse momenta squared. Electron candidates are reconstructed from ID tracks associated to energy clusters deposited in the EM calorimeter [24, 62]. All selected electrons are required to have $E_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.47$. As in the diphoton channel, electron candidates in the transition region ($1.37 < |\eta| < 1.52$) are not considered. The final selection requires that the electrons pass the *medium* identification working point.

Electron candidate tracks are required to satisfy $|d_0/\sigma(d_0)| < 5$ and $|z_0 \sin\theta| < 0.5$ mm, where d_0 and z_0 are the transverse and longitudinal impact parameters defined relative to the primary vertex position and $\sigma(d_0)$ is the uncertainty on d_0. This selection ensures that the electron candidate track is consistent with the primary vertex of the event.

If there are more than two electrons in the same event, the two electrons with the highest E_{T} are selected to form the dielectron pair. An opposite-sign electron charge requirement is not applied due to a high probability of charge misidentification for high-E_{T} electrons. The reconstructed invariant mass of the dielectron pair, m_{ee}, is required to be above 225 GeV to avoid the region dominated by decays of the Z boson which cannot be described using the same background parameterisation as the high-mass region.

To suppress the backgrounds from misidentified jets, the photon and electron candidates are required to pass calorimeter- and track-based isolation criteria. The photons are required to pass the "Tight" isolation which is 90–95% efficient for the analysis selection [24]. Electrons are required to pass the "Gradient" isolation which is 99% efficient for the analysis selection [24].

6 Signal modeling

The CW/LD signals are modelled using analytic invariant mass templates constructed from a range of inputs including PDF information, branching ratio, and detector resolution modelling derived from the previous ATLAS diphoton and dielectron analyses.

For each reconstructed mass bin, the expected number of signal events is determined by computing the contribution from the entire KK graviton tower. The estimate of the expected number of signal events for a given reconstructed mass bin i is given by

$$N_i^{\mathrm{reco}} = L_{\mathrm{int}} \cdot \sum_G \sigma_G \cdot \mathcal{B}_G(G \to XX) \cdot P(M_i|M_G) \cdot (A \times \epsilon)_G \tag{1}$$

where the index G runs over the KK graviton modes whose masses are within the search range. In Eq. (1), L_{int} is the integrated luminosity, σ_G is the production cross-section, $\mathcal{B}_G$ denotes the branching ratio, and $(A \times \epsilon)_G$ is the acceptance times efficiency for a given mass, M_G, of the KK graviton. The $P(M_i|M_G)$ term represents the probability to reconstruct an event in invariant mass bin i for events with a truth mass of M_G that pass the selection requirements and is referred to as the transfer function (TF). The TF provides a smooth transformation between the truth and reconstructed invariant mass spectrum by modelling the detector effects. The convolution becomes a product as the true masses M_G are discrete and the natural widths can be neglected as the resonances are described by the narrow-width-approximation (NWA) over the range of truth masses considered in this analysis.

The cross-sections and branching ratios for the gravitons in the CW/LD signal are taken from Ref. [2], where the effects of heavy graviton cascade decays to lighter graviton pairs are included in the total width

when calculating the diphoton and dielectron branching ratios. The additional contributions to the signal from these cascade decays into diphoton and dielectron final states are not considered.

The TF and acceptance times efficiency terms are derived following the methodologies of the previous ATLAS dielectron and diphoton resonance searches [10, 12] and are briefly outlined below. These terms are derived using the DY MC samples and the resonant single-graviton MC samples for the dielectron and diphoton channels, respectively.

For the dielectron channel, the detector response is defined in the TF with respect to the relative dilepton mass resolution $(m_{\ell\ell} - m_{\ell\ell}^{\text{true}})/m_{\ell\ell}^{\text{true}}$, where $m_{\ell\ell}^{\text{true}}$ is the generated dilepton mass at Born level before final-state radiation. The mass resolution is parametrised as the sum of a Gaussian distribution and a Crystal Ball function. The Gaussian component in this function describes the central peak of the detector response and the Crystal Ball component is used to model the effects of bremsstrahlung in the dielectron system. The parameterisation of the relative dilepton mass resolution is determined using a simultaneous fit of the resolution function to the DY MC events. The DY MC sample is separated into 200 bins of equal size in $m_{\ell\ell}^{\text{true}}$ based on a logarithmic scale in the dielectron mass range from 130 GeV to 6000 GeV. The mass resolution fit is repeated to evaluate the uncertainty on the fit parameters by shifting individually the lepton energy and momentum scale and resolutions by their uncertainties. Although the resolution function is derived on spin-1 events, the results are found to be compatible with the truth response coming from the spin-2 gravitons.

For the diphoton channel, the TF describes the detector resolution of the invariant mass using a double-sided Crystal Ball (DSCB) function [63, 64] with the parameters expressed as a function of the resonance mass, m_X. The DSCB parameters are extended to lower masses with respect to Ref. [12] as the CW/LD signals considered in this analysis can contain graviton masses below 500 GeV, which were not considered in that analysis.

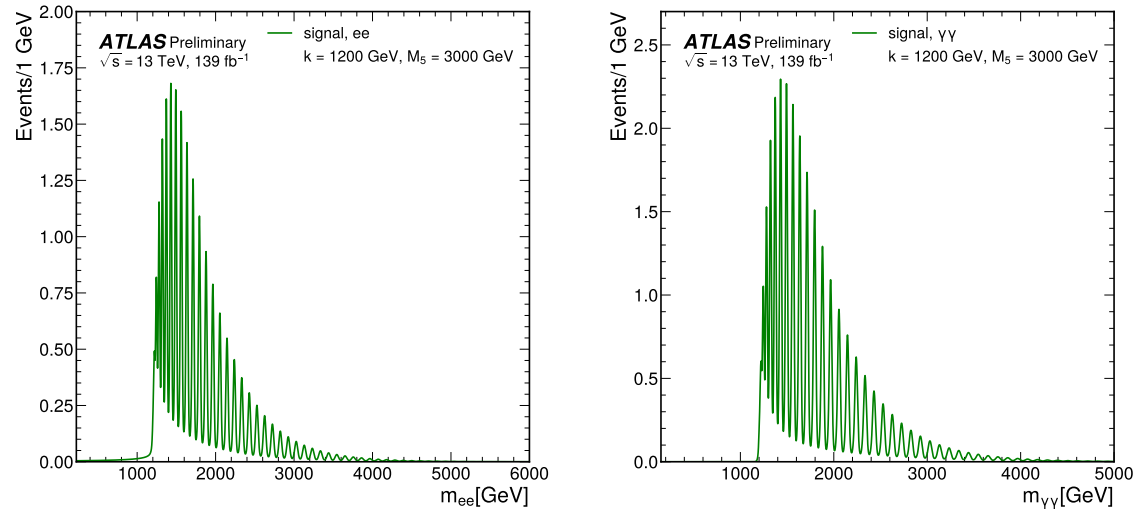
Analytic templates allow for signal shapes to be created for any choice of k and M_5. These templates are confirmed to agree with the MC samples generated at the specific points in the k–M_5 plane discussed in Section 4. Examples of the expected signal shapes in each analysis channel are shown in Figure 1 for $k = 1200$ GeV, $M_5 = 3000$ GeV and for $k = 2000$ GeV, $M_5 = 8500$ GeV.

7 Background modeling

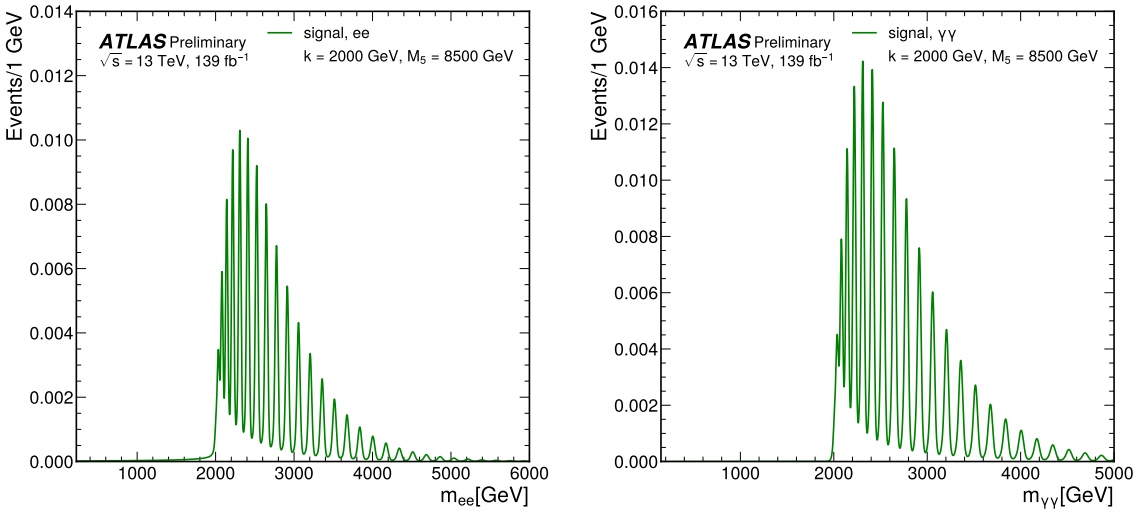
The diphoton and dielectron search channels in this analysis use the same data-driven background estimation technique based on a functional form fit to the data. Each analysis channel uses a single functional form to model the background across the full invariant mass spectrum. The analytic form for the background in each channel is determined from a fit to the background expectation derived from MC simulation and control-region data.

For the dielectron channel, the functional form is taken from the analysis in Ref. [10], where the fit was performed on a background template with contributions from the MC estimation of the DY, $t\bar{t}$, diboson, and single-top processes and the data-driven estimation of the reducible backgrounds arising from the multi-jet and W+jet processes. The function is parametrised as a function of $m_{\ell\ell}$ and is given by:

$$f_{ee}(m_{\ell\ell}) = f_{\text{BW},Z}(m_{\ell\ell}) \cdot (1-x)^b \cdot x^{\sum_{i=0}^{3} p_i \log(x)^i} \tag{2}$$



(a) (b)



(c) (d)

Figure 1: The expected signal invariant mass shape in the (a) ee and (b) $\gamma\gamma$ channels, where the CW/LD parameter values $k = 1200$ GeV and $M_5 = 3000$ GeV are used to generate the analytic signal template shapes. The expected signal invariant mass shape for CW/LD parameter values $k = 2000$ GeV and $M_5 = 8500$ GeV are shown in the (c) ee and (d) $\gamma\gamma$ channel. From these two different signal points, it can be seen that while the parameter k determines the onset of the KK graviton spectrum, the signal cross-section scales inversely with M_5^3.

where $x = m_{\ell\ell}/\sqrt{s}$ and $f_{\mathrm{BW},Z}(m_{\ell\ell})$ is a non-relativistic Breit–Wigner function with $M_Z = 91.1876\,\mathrm{GeV}$ and $\Gamma_Z = 2.4952\,\mathrm{GeV}$ [65]. The b and p_i parameters are fixed to the values from the fit to data in Ref. [10] and the $(1-x)^b$ term ensures that the background fit tends to zero as $x \to 1$, consistent with the expectation from the collision energy of the LHC. The background function is normalised such that its

integral corresponds to the total number of events in the signal region.

For the diphoton channel, the largest background component arises from the non-resonant production of photon pairs ($\gamma\gamma$ events). Smaller contributions to the total background come from events containing a photon and a jet misidentified as a photon (γj events) and from events with two jets where both jets are misidentified as photons (jj events). The relative contribution from each of these processes is determined using a two-dimensional sideband method described in Ref. [26]; the overall $\gamma\gamma$ purity increases as a function of $m_{\gamma\gamma}$ from around 89% at 150–200 GeV to around 97% above 400 GeV. The relative uncertainty on the purity ranges from 0.5–3% and the remainder of the background is dominated by γj events.

The analytic form for the diphoton channel is chosen from the family of functions previously used to describe the diphoton invariant mass spectrum [66], following the procedure described in Ref. [12]. The background function is described by the following form

$$f(x; b, a_0, a_1) = (1 - x^{1/3})^b x^{a_0 + a_1 \log(x)} \tag{3}$$

where b, a_0, a_1 are free parameters and $x = m_{\gamma\gamma}/\sqrt{s}$.

8 Uncertainties

The previous dilepton [10] and diphoton [12] data-driven analyses extracted full MC-based uncertainties, which are used in this search to construct systematically-varied distributions as discussed in Section 9. The variations considered in both the dielectron and diphoton channels come from theoretical and experimental uncertainties in the simulated backgrounds, theoretical and experimental uncertainties on the detector response and signal yield, as well as uncertainties induced by the data-driven background estimates.

The following theoretical uncertainties are considered for variations in the DY and $\gamma\gamma$ components: the variations of the nominal PDF set, variations of the PDF scales, and the effect of choosing different PDF sets. Additional theoretical uncertainties on the DY component from the variation of the strong coupling, electroweak corrections, and photon-induced corrections [67] are also considered.

The experimental uncertainties considered in the diphoton (dielectron) analysis channel include uncertainties on the luminosity determination, pile-up profile, trigger efficiency, photon (electron) energy scale and resolution, and the efficiency of photon (electron) identification and isolation requirements. More details on the determination of the experimental systematic uncertainties can be found in Refs. [10, 12]

The systematic uncertainties on the detector response are determined by varying the energy scale and energy resolution of the electrons and photons. The effect of the experimental uncertainties on the signal yield, based on the variations in the acceptance times efficiency, are also considered. An additional uncertainty is also derived for the signal based on the internal variations of the PDF set.

The largest data-driven uncertainty is a shape uncertainty derived in the diphoton channel that corresponds to the uncertainty on the relative contribution of the reducible γj background component. This uncertainty is estimated by varying the measured fraction of the $\gamma\gamma$ background according to its uncertainty from the two-dimensional sideband method mentioned in Section 7.

In this analysis, each systematic uncertainty is either naturally two-sided or is symmetrised to provide a two-sided systematic uncertainty. All uncertainty variations in both channels are smoothed using a sliding window method [10] with a window of 300 GeV. The value of the variation at each point is taken as the

average of 150 previous and 150 next bins (with bin size 1 GeV) and for the points near the edges, the input is extended by replicating the value at the edge.

The impact of the systematic uncertainties on the final result is quantified by changes in the test statistic values. The largest impact on the test statistic values comes from the theoretical uncertainties on the DY and diphoton continuum backgrounds due to the variations of the underlying PDF. The experimental uncertainties are sub-dominant to the theoretical uncertainties across the full mass range considered in each analysis channel.

9 Generation of pseudo-experiments

The analysis makes use of pseudo-datasets as an input for the CWT, where one scalogram is generated for each pseudo-dataset. These scalograms are the basic ingredient for the statistical procedures used in this analysis and are also used for estimating the systematic uncertainties. The pseudo-datasets (or "toys") are generated with background-only models based on the functional forms from Section 7 and with signal-plus-background models using the signal models discussed in Section 6 and the background-only models from Section 7. These toys are produced with finely-binned (1 GeV) mass spectra separately for the diphoton and dielectron final states. For the toys containing a signal, a specific choice of CW/LD model parameters is used. The generated pseudo-datasets are hereafter called statistical-only toys (stat-toys) or systematic toys (syst-toys) depending on whether systematic variations are considered when constructing the toy.

To make the stat-toys, the analytic shapes are constructed containing either a background-only shape or a signal-plus-background shape. The background component is given a normalisation corresponding to the data normalisation in the signal region of each analysis channel. Next, a Poisson-fluctuated toy distribution is generated from the analytic shape chosen. Each of these toys is then treated as a proxy for the real data in the subsequent statistical analysis discussed in Section 11.

To make the syst-toys, a similar procedure is followed as for the stat-toys, but starting from an alternative description of the background-only and signal-plus-background shapes. This analysis follows the procedure introduced in Ref. [11] and is described briefly below.

To make one syst-toy in the background-only case, the alternative background shape (called hereafter an uncertainty template) is built from the systematic variations that arise from the sources of both experimental and theoretical uncertainties as discussed in Section 8. In the background-only case, these systematic variations are given in the invariant mass space as shape variations around the background description from simulation, as obtained in Refs. [10, 12]. A relative variation is defined as the relative difference between the up or down variation and the background description from simulation. To obtain the absolute systematic difference with respect to the nominal background description from data defined in Sections 6 and 7, the relative variation is multiplied by the nominal background description from data. Each uncertainty template is constructed from the nominal background description from data, summed with the weighted absolute systematic differences. The weights are randomly sampled from a Gaussian distribution, with a mean of zero and standard deviation of one. The number of such weights in each uncertainty template is equal to the number of systematic variations considered.

To replicate the procedures given in Refs. [10, 12], the resulting uncertainty template is Poisson-fluctuated to generate one uncertainty pseudo-dataset. This uncertainty pseudo-dataset is then fit with the background model from either Eq. (2) or Eq. (3), depending on the analysis channel. This procedure results in one

smooth alternative background description, replacing the fits done in Ref. [10, 12], which are used to generate one respective background-only uncertainty pseudo-dataset.

To make one syst-toy in the signal-plus-background case, a similar procedure is followed with minor modifications. The procedure starts from the signal systematic variations discussed in Section 8. The nominal signal shape is described by Eq. (1). The absolute systematic difference for each variation is calculated directly with respect to the nominal signal shape. The signal uncertainty template is constructed from the nominal signal shape, summed with the Gaussian-weighted absolute systematic differences. Each signal-plus-background uncertainty template is constructed as the sum of one smooth alternative background description, as discussed above, and one signal uncertainty template. The result is Poisson-fluctuated to generate one respective signal-plus-background uncertainty pseudo-dataset.

These procedures are repeated to generate background-only and signal-plus-background syst-toys ensembles, where each toy in these ensembles reflects a different (random) combination of all uncertainties. The syst-toys ensembles are used in the subsequent statistical analysis discussed in Section 11.

10 Continuous wavelet transforms

Continuous wavelet transforms are used in this search to convert the diphoton and dielectron invariant mass spectra into mass versus frequency information in the form of scalograms prior to applying the machine learning techniques discussed in Section 11. An overview of the CWT method used in this search is detailed below.

The CWT is a measure of similarity between a chosen wavelet and a signal. The wavelet is used to scan the signal for different frequencies of the wavelet throughout the invariant mass spectrum.

The wavelet $\psi(x)$ is a basis function localised in both mass and frequency space and the CWT of a signal $f(m)$ at a scale[3] a and translational parameter b is given by a projection on $\psi(x)$ for different a and b values [15]:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(m)\, \psi^* \left(\frac{m-b}{a} \right) dm \tag{4}$$

A scalogram can be produced by taking the norm of the coefficient $W(a,b)$ for all values of a and b. Therefore, the CWT in this analysis defines how much of a certain frequency is present in the signal at a given invariant mass bin.

The *Morlet wavelet* [68] is used in this analysis as the choice of $\psi(x)$ in Eq. (4). The Morlet wavelet is chosen because it is Gaussian-shaped in the frequency domain, which minimises possible edge effects that could be interpreted as signal, as opposed to wavelets with sharp boundaries that can introduce edge effects [69]. It consists of a localised wave packet and is given by:

$$\psi(x) \equiv \frac{1}{\sqrt{B\pi}}\, e^{-x^2/B} \left(e^{i2\pi C x} - e^{-\pi^2 B C^2} \right), \tag{5}$$

where B and C are constants which are chosen to be 2 and 1 respectively. This choice of C ensures that the CWT of a signal reaches a maximum when the signal wavelength approximately equals the scale a.

[3] The scale a is inversely proportional to frequency, where higher scales represent the low frequency component and vice versa.

Example scalograms for the dielectron and diphoton channels can be seen in Figure 2. These scalograms are created after applying the CWT on a background-only distribution (Figures 2(a) and 2(b)) and on a signal-plus-background distribution (Figures 2(c) and 2(d)) with statistical fluctuations. The red saturated area in the high-scale region occurs because the background distribution is not flat. Therefore, the background may still appear as periodic for very large scales, i.e. those scales with small frequencies. The sharp transition seen between blue and red from low $|W(a,b)|$ to high $|W(a,b)|$ occurs because the $|W(a,b)|$ values change quickly relative to the scale chosen for the z-axis.

The signal contribution in Figures 2(c) and 2(d) is clearly seen as a local small "island", discernible from the continuum of the background shown in Figures 2(a) and 2(b), even when realistic statistical fluctuations are included. This signal-island represents the locality of the signal in both the mass and the scale (or frequency) spaces shown in Figure 1. Fixing M_5 and increasing the value of k shifts the signal-island in two ways. The signal-island is shifted horizontally to higher masses because the signal turn-on point in mass is roughly equal to the value of k. The signal-island is also shifted vertically towards higher scales because the spacing in the KK tower increases with k. Fixing k and changing M_5 only determines the prominence of the island, i.e. it becomes more distinct with decreasing M_5 because of the inverse relationship between M_5 and the signal cross section. Changing the mass resolution due to detector effects (within the uncertainties) only affects the $|W(a,b)|$ values of the signal-island and this change is effectively equivalent to changing M_5.

11 Data analysis

The diphoton and dielectron invariant mass distributions from the toys detailed in Section 9 are transformed into images using the CWT method described in Section 10 before being used to train a neural network (NN). The diphoton and dielectron images are treated as independent channels and are trained separately. Two types of convolutional NNs are used in this search, a classifier NN and an autoencoder (AE), which are discussed in Section 11.1 and Section 11.2, respectively. In both cases, the training is done using stat-toys, while the prediction is performed using syst-toys to evaluate the impact of the systematic uncertainties on the NN responses. As discussed in Section 9, the syst-toys already include a proper statistical representation of the data.

The classifier NN is used to probe for the periodic signals of the CW/LD model specifically (denoted here as model-dependent). On the other hand, the AE NN is used to search for more general anomalies in the data (denoted here as model-independent). Both of the neural networks used in this analysis are based on the setups given in Ref. [15]. The neural networks are implemented in Keras [70] using the TensorFlow back end [71]. The ADAM [72] optimiser is used to minimise the loss functions in each NN setup. The search is performed using a general form of a test statistic derived from the NN output itself.

The loss functions associated with each of the NNs play the role of the likelihood function in a typical analysis. The loss function used in the classifier NN is the Binary Cross-Entropy (BCE) [73], while the AE NN uses the Mean Squared Error (MSE) loss function. The MSE loss function is further discussed in Section 11.2. The training and validation loss are recorded as a function of the number of epochs and are inspected to verify that each NN gives good agreement in the output of the loss function when comparing the training and validation datasets.

In this analysis, the NN output in the classifier setup and the value of the loss function in the AE setup are used as test statistics. To set the exclusion limits on the model-dependent parameter space, the modified

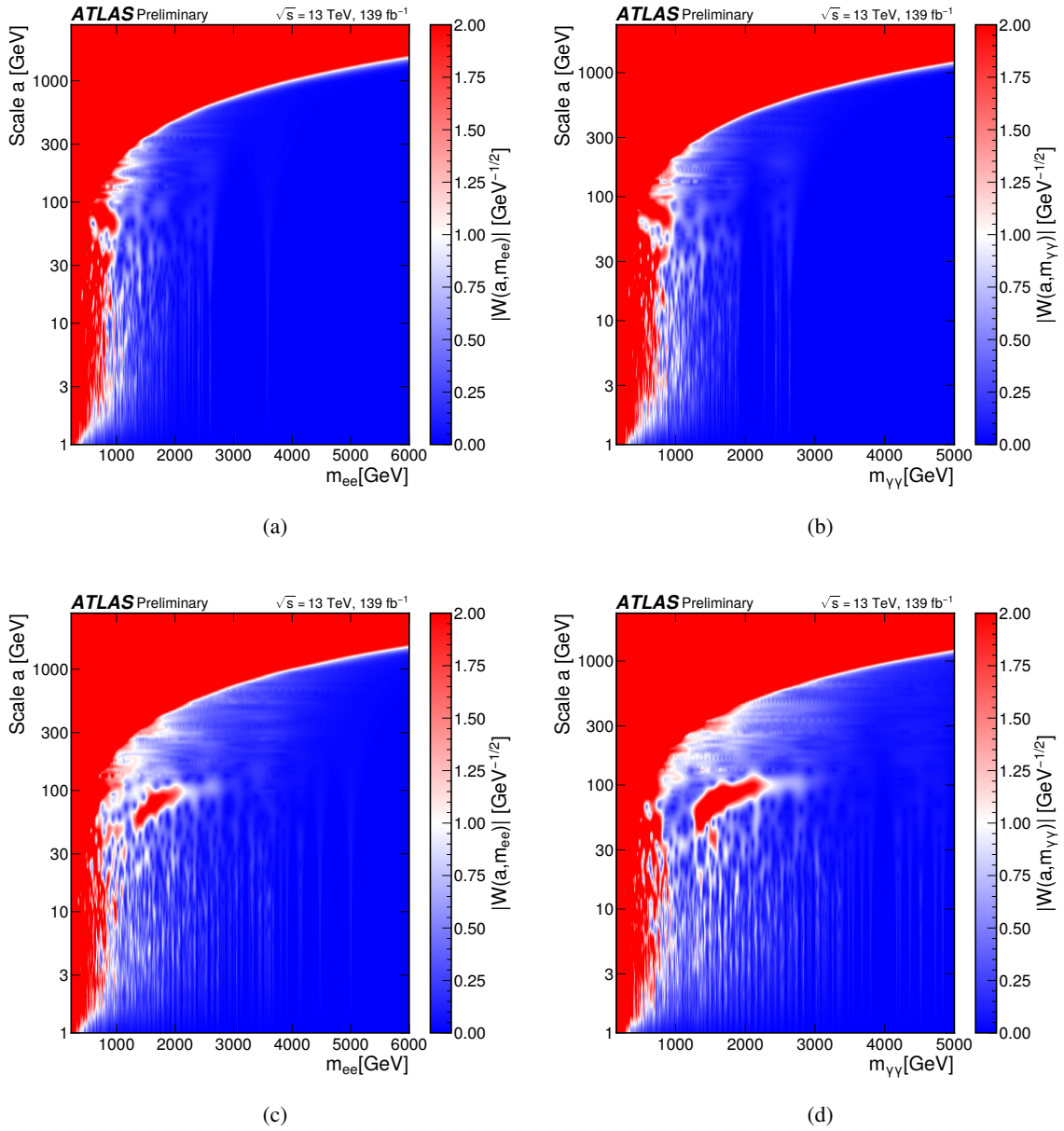


Figure 2: Scalogram output of the CWT of (a) dielectron and (b) diphoton background-only toy experiments and (c) dielectron and (d) diphoton signal-plus-background toy experiments with $k = 1200$ GeV and $M_5 = 3000$ GeV. The signal contribution in (c) and (d) manifests as a localised "island" in mass and scale, discernible from the continuum of the background shown in (a) and (b). The scalograms are provided here with a mass binning of 1 GeV. Here a is the CWT scale parameter and $W(a, b)$ are the wavelet coefficients defined in Eq. (4).

frequentist method, commonly known as the CL_s method [74–76], is used. This choice results in more conservative limits than those obtained by the standard p-value procedure [74, 75]. The dielectron and diphoton channels are analysed separately due to potential overlaps in their event selections.

11.1 Classifier NN

One possible way to discover a specific signal in a scalogram is with a classifier NN. In this analysis, the convolutional NN classifier setup from Ref. [15] is implemented for the model-dependent search. The CWT scalogram of the mass spectrum for each stat-toy, syst-toy, and the signal region data is produced and rebinned to ensure the inputs to the NN remain manageable. The stat-toy scalograms of the background-only case and the signal-plus-background case are used for training the binary classifier convolutional NN, where the output of the network is between 0 (background-like) and 1 (signal-plus-background like).

In addition to the signal periodicity of the CW/LD model, these models can introduce a non-resonant tail above the SM background at high dielectron or diphoton invariant masses. For these high-mass ranges, the periodicity may no longer be as apparent due to the widening of the resonances with mass. This happens because of the merging of the mass peaks due to the experimental resolution. This analysis focuses on the periodic features of the signal by adjusting the procedure such that the subsequent inference does not consider the effects of the non-resonant tail. To achieve this adjustment, a simple cutoff is applied to the dielectron or diphoton mass distribution to remove the high-mass events above a certain point in mass. This mass threshold is determined by calculating a signal's local significance, S/σ_B, where S is the signal yield and σ_B is the local per-bin uncertainty. This uncertainty includes the statistical uncertainty and the systematic uncertainties described in Section 8, added in quadrature, of the background. The bin in the invariant mass distribution on a local signal peak where the per-bin-significance falls below 50% of its maximal value is chosen as the threshold. While the cutoff is independent of M_5 by construction, it does rise monotonically with k. Therefore, the cutoff values are calculated for a few points in k and are then fit with a second-order polynomial to allow for interpolation between the k points. As the long-range signal shape falls rapidly after reaching its maximum, alternative choices close to the nominal choice of 50% of the maximum local significance have little impact on the result. A lower significance choice, for example $S/\sigma_B \approx 10\%$ of the maximal value, effectively introduces no cutoff and the subsequent inference remains affected by the presence of the non-resonant tail feature. A higher significance choice, e.g. $S/\sigma_B \approx 90\%$ of the maximum value, removes the non-resonant tail effectively but also removes a large mass range where the periodicity is still apparent and reduces the search sensitivity to the intended feature. For the dielectron channel, the 50% threshold mass value ranges from 1400 to 6000 GeV for k in the range of 200 GeV to 5850 GeV. For the diphoton channel, the mass threshold ranges from 2000 to 5000 GeV for k in the range of 150 GeV to 4900 GeV.

For completeness, the results of the classifier search without mass thresholds are also provided. In this case, the non-resonant feature of the signal may be present, while the periodicity may be unresolvable in the high mass region. The CWT and NN procedure can still provide some discrimination power between the signal-plus-background and background-only distributions due to the non-resonant feature and additional oscillations available to describe the signal. Therefore, the sensitivity for low-k ($k \lesssim 1000$ GeV) signals is expected to be higher than for the case with mass thresholds. For high-k values, $k \gtrsim 1000$ GeV, the sensitivities of the two methods are expected to be almost identical.

11.2 Autoencoder NN

A model-independent approach searching for anomalies in scalograms is performed using AEs. This technique has previously been applied to jet images in Refs. [77, 78] and was implemented and suggested for this type of search in Ref. [15]. The AE compresses a scalogram to a smaller set of parameters, which are then used to reconstruct the input scalogram. Instead of using the original background-only scalograms,

each bin in the scalogram is remapped monotonically to uniformly distributed values in the range 0 to 1. The negative logarithm of the remapped bin value is used to build a new two-dimensional input. This procedure is identical to the procedure in Ref. [15] and ensures that the loss function is not dominated by the region with high statistics. The AE is trained on background-only stat-toy remapped scalograms to reproduce the original remapped scalograms ($\mathbf{y}$) by minimising the MSE loss function [78], L_{MSE}, as defined below:

$$L_{\text{MSE}}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^{n} |\mathrm{y}_i - \hat{\mathrm{y}}_i|^2 \,, \tag{6}$$

where n is the number of bins used in the scalogram and $\hat{\mathbf{y}}$ is the output remapped scalogram from the AE.

After the AE model is trained, the AE should be able to approximately reproduce the original remapped scalogram if applied to a typical background-only syst-toy scalogram. Conversely, the AE may fail to reproduce the scalogram if applied to a data scalogram that contains a signal. This metric, L_{MSE}, is used as a test statistic. In this analysis, the AE is particularly sensitive to signals with periodic structures.

The L_{MSE} is also calculated per scalogram bin and at each point an ensemble of the L_{MSE} values for the background predictions is generated. From these ensembles, a local p-value is calculated for every scalogram bin. The statistically significant bins are then visible with p-values close to 0.

As discussed in Section 11.1, to focus the search on the periodicity features and avoid possible inference that is based only on non-resonant signals, two sets of selections are introduced. Unlike in Section 11.1, these selections cannot be based on a specific signal and should be characterised only based on information pertaining to the SM backgrounds. In the first method, the non-resonant features are removed by using the statistical error at the tail of the background template shape, $R = 1/\sqrt{B(m_{\text{min}})}$, where $B(m_{\text{min}})$ is the integral of the background template shape above some minimum invariant mass m_{min}. Different m_{min} thresholds are derived for different R values, where for each, the tail of the distribution above the threshold is removed before calculating the CWT. The thresholds and the respective invariant mass ranges for each channel are found in Table 1.

Table 1: Different precision-driven thresholds considered for each analysis channel and the corresponding mass ranges probed in the model-independent analysis. For the no threshold scenario, the upper limit in mass is chosen such that the range goes 300 GeV beyond the last observed data point in the dielectron and diphoton mass spectra.

Threshold	Dielectron Mass Range [GeV]	Diphoton Mass Range [GeV]
$R < 10\%$	225–1520	150–1380
$R < 50\%$	225–2700	150–2400
No threshold	225–4400	150–2700

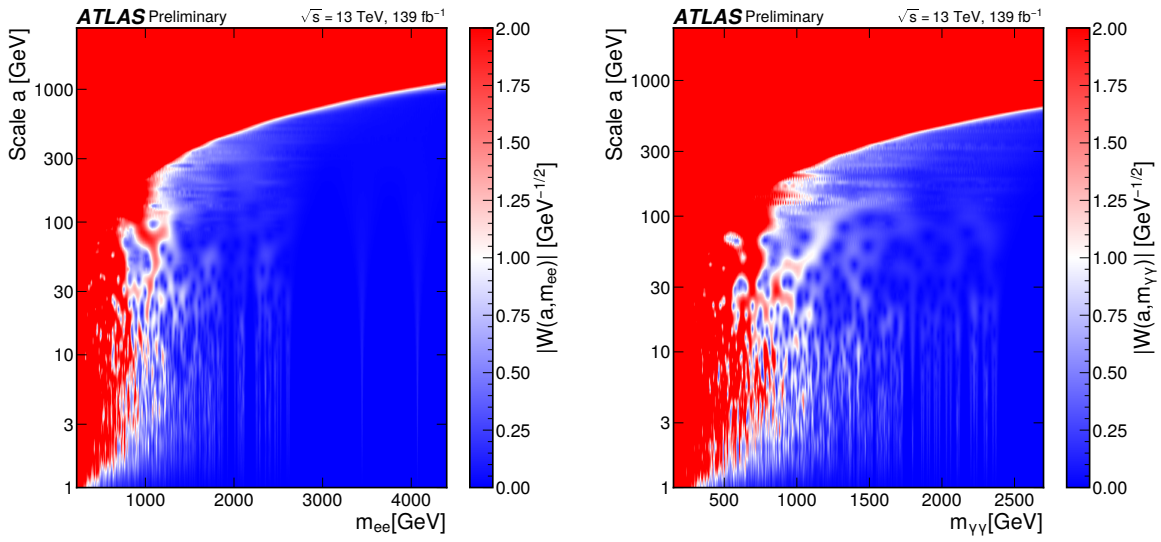
This procedure also prevents large statistical fluctuations in the high-mass tail of the background distribution from affecting the L_{MSE} of the AE predictions.

In the second approach, the scale range of the scalograms is instead modified and the mass range remains unrestricted. This approach, referred to as scale thresholding, avoids the large-scale region of the scalogram that may include continuum contributions from a signal. In this method, the bin content in the scalogram is replaced with zero if the scale a in that bin satisfies $a > m/4$, where m denotes the mass value of that bin.

In a similar way, to exclude the scales that are short and cannot be resolved experimentally, the bin content in the scalogram is replaced with zero if the bin satisfies $a < m/100$. The results from both methods are given in Section 12.

12 Results

The scalograms for the data events passing all selections for the diphoton and dielectron channels are shown in Figure 3.



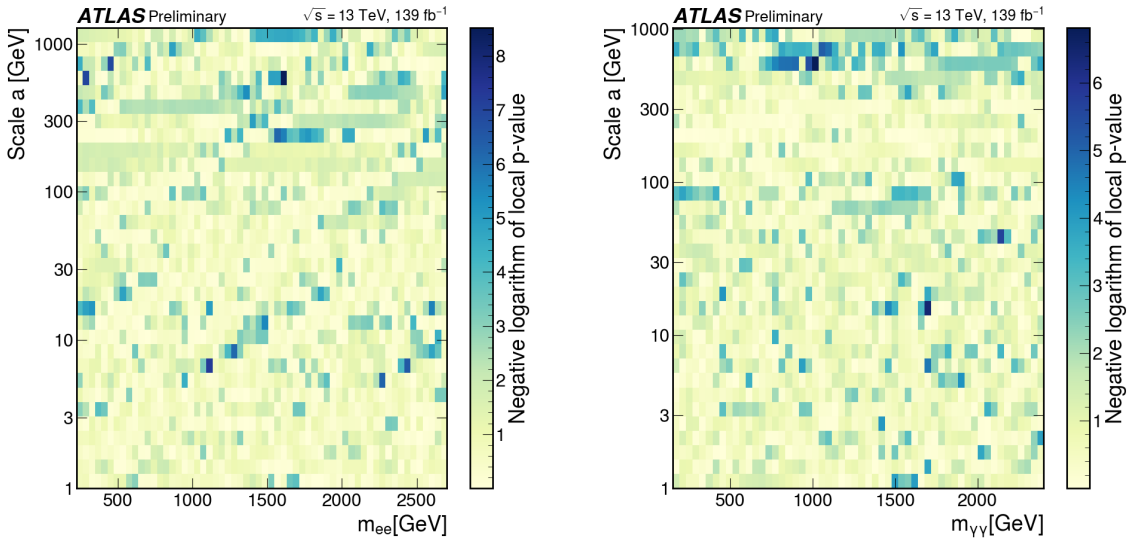
(a) (b)

Figure 3: The scalogram output of the CWT for the (a) dielectron and (b) diphoton search channels for the observed signal region data. The scalograms are shown here with a mass binning of 1 GeV and are rebinned to a coarser binning prior to running the neural networks. Here a is the CWT scale parameter and $W(a,b)$ are the wavelet coefficients defined in Eq. (4). The range of m_{ee} and $m_{\gamma\gamma}$ is chosen such that the plots cover approximately 300 GeV after the last observed data point.

The model-independent results from the AE-based search with $R < 50\%$ are presented as the negative logarithm of the local p-values corresponding to these data scalograms and are shown in Figure 4.

To obtain the significance from the model-independent results, the L_{MSE} is calculated for each original and predicted scalogram pair, with the sum in Eq. (6) running over all the scalogram bins. An ensemble of L_{MSE} values for the background prediction is generated from many original and predicted scalogram pairs. From this ensemble, the significance is calculated for the data scalogram.

The significances for the model-independent results in each analysis channel are given in Table 2. No significant deviations from the background-only hypothesis are observed. The largest deviation from the background-only (SM) hypothesis is found in the dielectron channel with the 50% threshold and the scale threshold, with an excess corresponding to a significance of 1.5σ.



(a) (b)

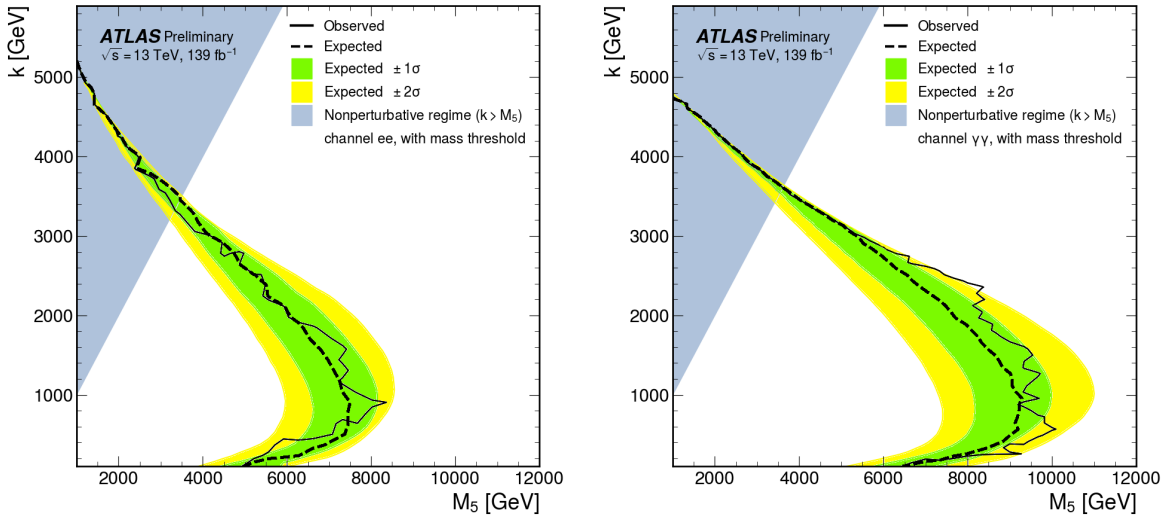
Figure 4: The negative logarithm of the local p-values calculated for the observed data mass distributions in the (a) dielectron and (b) diphoton search channels for the $R < 50\%$ scenario. Here a is the CWT scale parameter.

Table 2: The observed significance in each analysis channel for the different thresholds considered in the model-independent analysis. Each significance is signed with positive (negative) values indicating the loss from data is higher (lower) than the median loss from the expected background.

Threshold	Dielectron Significance	Diphoton Significance
$R < 10\%$	0.4	-1.8
$R < 50\%$	1.5	-0.2
No threshold	0.7	-0.7
Scale threshold	1.5	-0.6

The model-dependent results are also presented for each analysis channel. No significant deviation from the background-only hypothesis is seen in either of the analysis channels. In the absence of a clear signal, limits at 95% CL_s are set on the CW/LD model in the k–M_5 plane. The limits for the case with mass thresholding, as discussed in Section 11.1, are shown in Figure 5. Additional exclusions are also shown for the case without mass thresholds in Figure 6. As expected, the sensitivity and corresponding limits are stronger for the case without mass thresholds in the region of $k \lesssim 1000$ GeV. In each of the results, a shaded region with $k > M_5$ illustrates the region of phase space where the CW/LD theory becomes non-perturbative.

For the case with mass thresholds, the maximum excluded M_5 value in the diphoton channel is approximately 10 TeV for values of $k \approx 600$ GeV, while the maximum excluded M_5 value in the dielectron channel is approximately 8.4 TeV for values of $k \approx 900$ GeV. For the case without mass thresholds, the maximum excluded M_5 value in the diphoton channel is approximately 11 TeV for values of $k \approx 200 - 1000$ GeV,



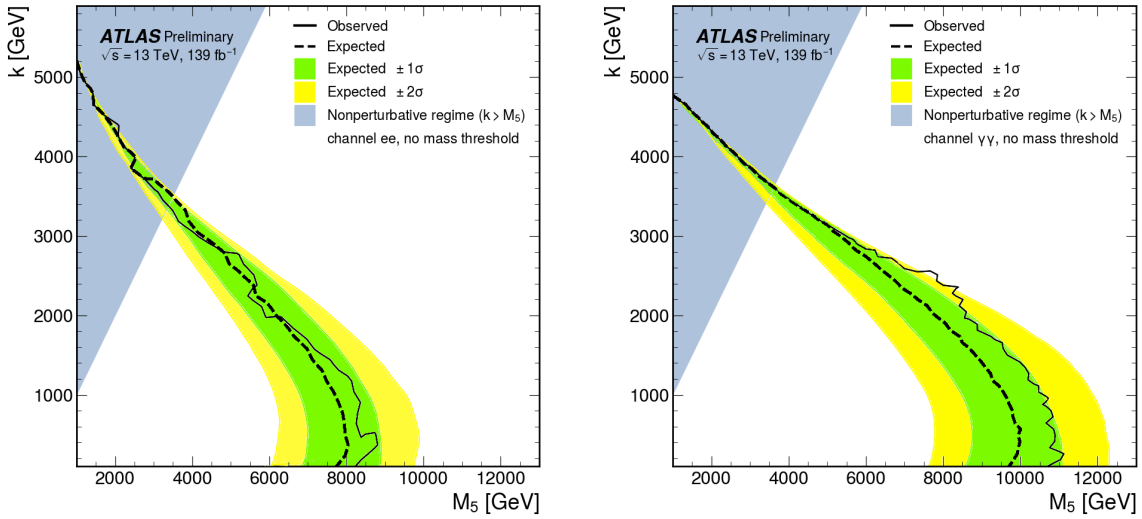
(a) (b)

Figure 5: The expected and observed exclusion limits at 95% CL for the clockwork gravity model projected in the k–M_5 parameter space for the (a) ee and (b) $\gamma\gamma$ channels, both for the case with mass thresholds. The green (yellow) shaded band represents the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainty on the expected limit. The shaded area with $k > M_5$ illustrates the region of parameter space where the CW/LD theory becomes non-perturbative.

while the maximum excluded M_5 value in the dielectron channel is approximately 8.8 TeV for values of $k \approx 400$ GeV.

The systematic uncertainties mainly impact the classifier result in the range $500 < k < 1500$ GeV for both analysis channels and statistical uncertainties dominate in the higher k ranges. The sensitivity of the classifier in the presence of systematic uncertainties is at most ~ 1 TeV weaker in M_5 exclusion. The most dominant uncertainty contribution in both channels is due to the theory uncertainties on the background modelling.

As the limit contour does not always follow the location of the lattice of points where the NNs were trained, it is worthwhile to mention that the NNs which were trained on the lattice of points in the k–M_5 plane (where a dedicated training for each point was used) are verified to be similarly effective for the models in-between those lattice points.



(a) (b)

Figure 6: The expected and observed exclusion limits at 95% CL for the clockwork gravity model projected in the k–M_5 parameter space for the (a) ee and (b) $\gamma\gamma$ channels, both for the case without mass thresholds. The green (yellow) shaded band represents the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainty on the expected limit. The shaded area with $k > M_5$ illustrates the region of phase space where the CW/LD theory becomes non-perturbative.

13 Conclusion

Searches for periodic new phenomena in the dielectron and diphoton invariant mass spectra with the ATLAS experiment at the LHC are presented. The searches are conducted with 139 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV. The model-dependent result is optimised for spin-2 periodic resonances (KK gravitons) in the clockwork/linear-dilaton model. Model-independent results sensitive to more general periodic contributions are also presented.

The data observed in the ee and $\gamma\gamma$ invariant mass spectra are found to be consistent with the SM expectations. In the absence of a significant excess, 95% CL$_s$ exclusion limits are set on the parameter space of the CW/LD model. This result excludes values of M_5 in the range 11 TeV to 1 TeV for values of k in the range 100 GeV to 5 TeV.

Appendix

Table 3: Structure of the classifier NN (based on Ref. [15]). For the input layer, the MergeX merges the mass binning by some factor and the MergeY merges the scale binning. The number of scale bins is calculated using the formula shown, which is based on the method used in the CWT package [79, 80]. The values used in this analysis for MergeX and MergeY are 40 and 4 respectively. In cases where a fractional number of bins is predicted, the floor function $\lfloor$ bins $\rfloor$ value is used and the bin count is padded to the nearest number divisible by 4.

Layer	Parameters
Input layer	$\frac{\text{Mass range}}{\text{MergeX}}$ mass bins $\times \left(12\left(\frac{\ln(\text{Mass range})}{\ln(2)} - 1\right)/\text{MergeY}\right)$ scale bins
Convolutional layer 1	# filters = 4
	kernel size = (3, 3)
	Activation: Elu
MaxPooling 1	Pooling size = (2, 2)
Convolutional layer 2	# filters = 8
	kernel size = (3, 3)
	Activation: Sigmoid
MaxPooling 2	Pooling size = (2, 2)
Convolutional layer 3	# filters = 16
	kernel size = (3, 3)
	Activation: Sigmoid
Dense 1	# of nodes = 200
	Activation: Sigmoid
Dense 2	# of nodes = 100
	Activation: Sigmoid
Output layer	# of nodes = 1

Table 4: Training parameters used for the classifier neural network based search. All parameters not specified here are left at their default KERAS values. The values for the parameters specified here are based on the values in Ref. [15].

Setting	Choice
Optimizer	ADAM
Loss function	BCE
# training experiments	4000
Validation split	0.2
Batch size	1000
# epochs	650
Callback	Smallest validation loss function

Table 5: Structure of the neural network used in the autoencoder approach (based on Ref. [15]). For the input and outputs layers, the MergeX merges the mass binning by the specified factor and the MergeY merges the scale binning. The number of scale bins is calculated using the formula shown, which is based on the method used in the CWT package [79, 80]. The values used in this analysis for MergeX and MergeY are 40 and 4 respectively. In cases where a fractional number of bins is predicted, the floor function $\lfloor$ bins $\rfloor$ value is used and the bin count is padded to the nearest number divisible by 4.

Layer	Parameters
Input layer	$\frac{\text{Mass range}}{\text{MergeX}}$ mass bins $\times \left(12 \left(\frac{\ln(\text{Mass range})}{\ln(2)} - 1\right) / \text{MergeY}\right)$ scale bins
Convolutional layer 1	# filters = 128
	kernel size = (3, 3)
	Activation: Elu
MaxPooling 1	Pooling size = (2, 2)
Convolutional layer 2	# filters = 128
	kernel size = (3, 3)
	Activation: Elu
MaxPooling 2	Pooling size = (2, 2)
Convolutional layer 3	# filters = 128
	kernel size = (3, 3)
	Activation: Elu
Dense 1	# of nodes = 40
	Activation: Elu
Dense 2 (*Encoded*)	# of nodes = 20
	Activation: Elu
Dense 3	# of nodes = 40
	Activation: Elu
Convolutional layer 4	# filters = 128
	kernel size = (3, 3)
	Activation: Elu
UpSampling 1	Upsampling factors = (2, 2)
Convolutional layer 5	# filters = 128
	kernel size = (3, 3)
	Activation: Elu
UpSampling 2	Upsampling factors = (2, 2)
Convolutional layer 6	# filters = 1
	kernel size = (3, 3)
	Activation: Elu
Output layer	Same as input layer

Table 6: Training parameters used for the autoencoder based search. All parameters not specified here are left at their default KERAS values. The values for the parameters specified here are based on the values in Ref. [15].

Setting	Choice
Optimizer	ADAM
Loss function	MSE
# training experiments	5000
Validation split	0.2
Batch size	1000
Padding	Same
# epochs	200
Callback	Smallest validation loss function

Table 7: The significance in each analysis channel for a specific signal injection for the no threshold scenario in the model-independent analysis. Each significance is signed with positive values indicating an excess and negative values indicating a deficit.

Injected signal	Dielectron Significance	Diphoton Significance
k = 500 GeV, M_5 = 4500 GeV	> 5	> 5
k = 500 GeV, M_5 = 6000 GeV	3.2	> 5
k = 500 GeV, M_5 = 8500 GeV	1.2	2.3
k = 1000 GeV, M_5 = 4500 GeV	> 5	> 5
k = 1000 GeV, M_5 = 6000 GeV	3.0	> 5
k = 1000 GeV, M_5 = 8500 GeV	1.2	2.0
k = 2000 GeV, M_5 = 4500 GeV	3.3	3.5
k = 2000 GeV, M_5 = 6000 GeV	1.5	1.8
k = 2000 GeV, M_5 = 8500 GeV	0.6	0.6

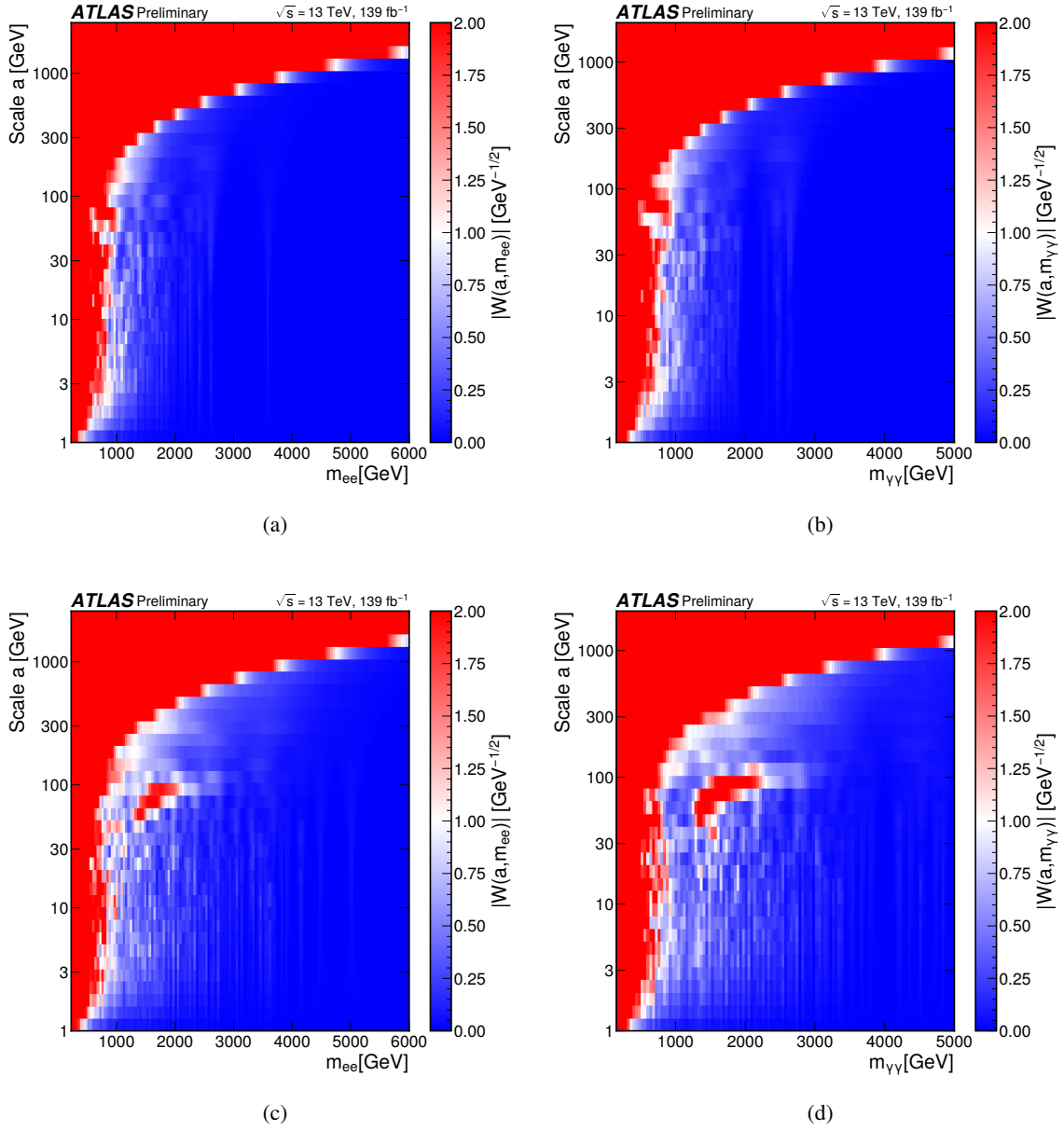
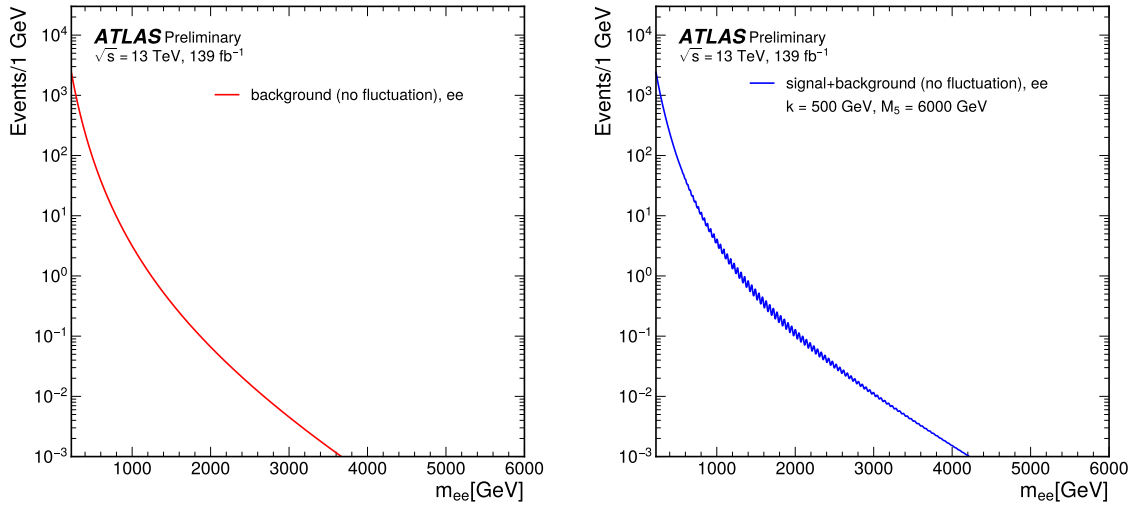
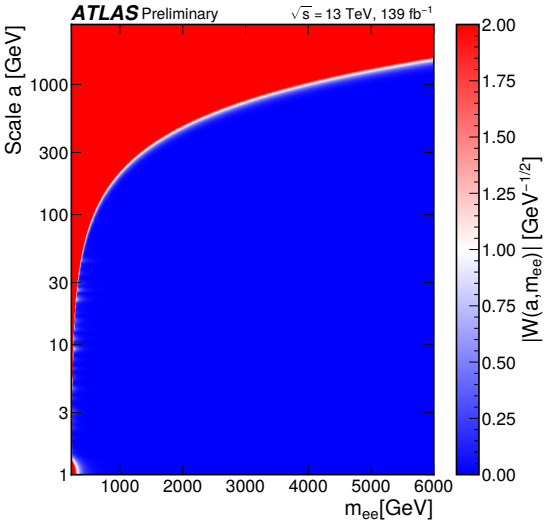


Figure 7: Scalogram output of the CWT of (a) dielectron and (b) diphoton background-only toy experiments and (c) dielectron and (d) diphoton signal-plus-background toy experiments with $k = 1200$ GeV and $M_5 = 3000$ GeV. The scalograms are provided here with the binning used for the NN inputs.

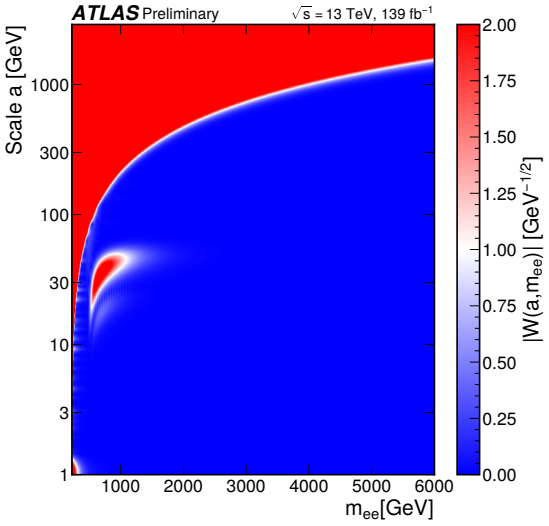


(a)

(b)



(c)



(d)

Figure 8: The dielectron invariant mass distribution for (a) background-only and (b) signal-plus-background (k = 500 GeV, M_5 = 6000 GeV) samples. Here the distributions are not Poisson-fluctuated. The corresponding scalogram plots for (c) background-only and (d) signal-plus-background samples are also shown. One can look at the scalogram plots to distinguish background-only distribution from signal-plus-background distribution.

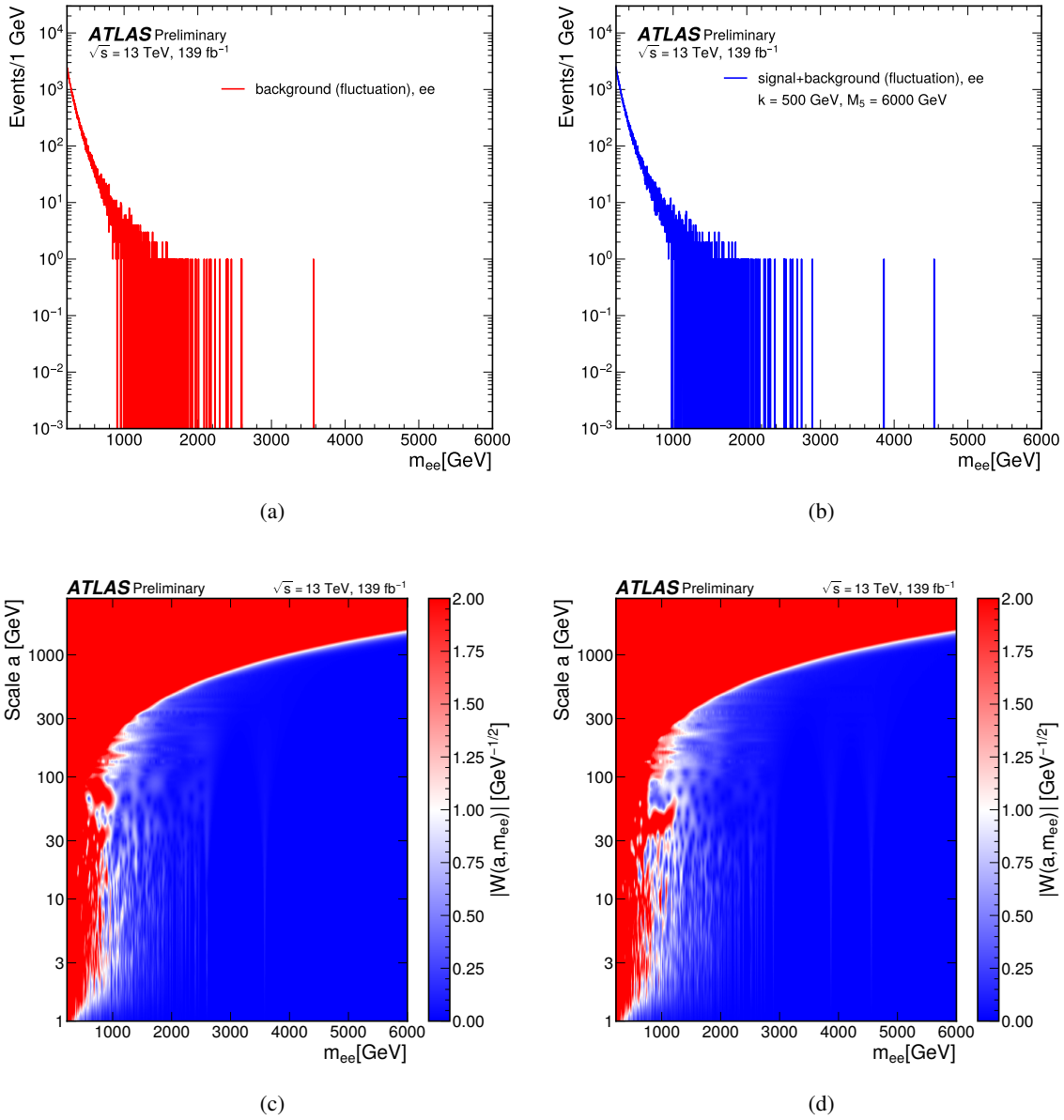
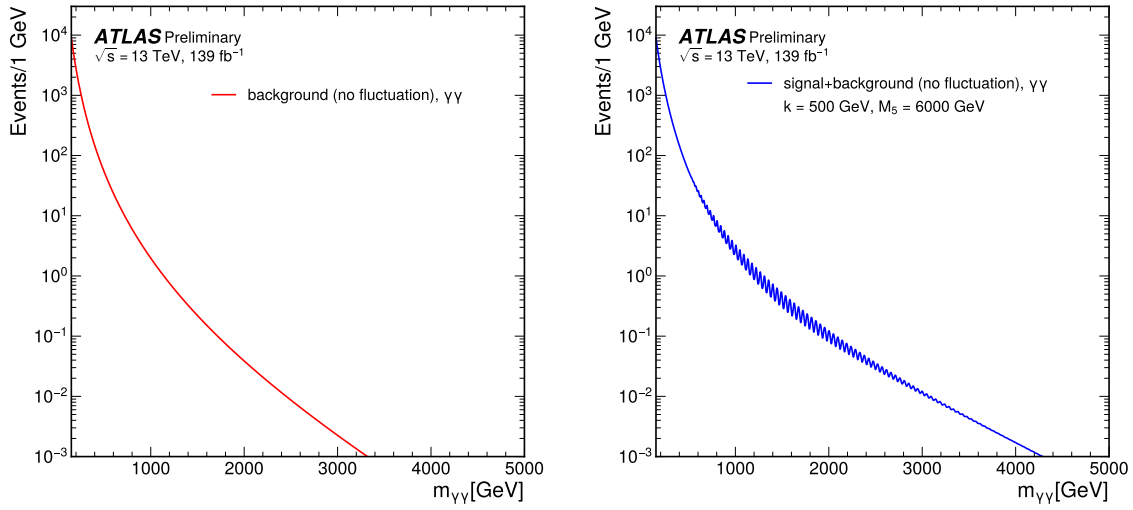
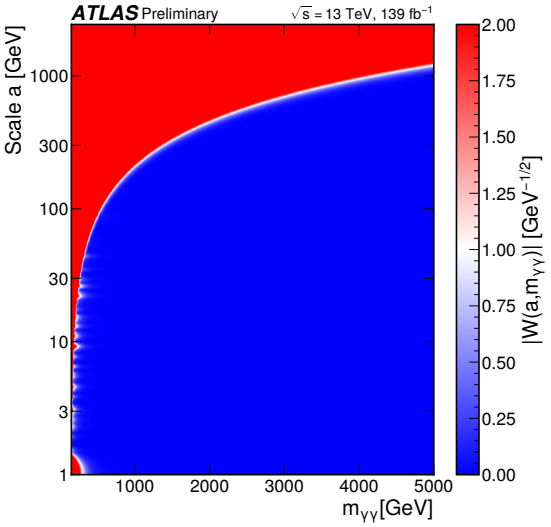


Figure 9: The dielectron invariant mass distribution for (a) background-only and (b) signal-plus-background (k = 500 GeV, M_5 = 6000 GeV) samples. Here the distributions are Poisson-fluctuated. The corresponding scalogram plots for (c) background-only and (d) signal-plus-background samples are also shown. It is not easy to distinguish background-only scalogram plot from that of the signal-plus-background.

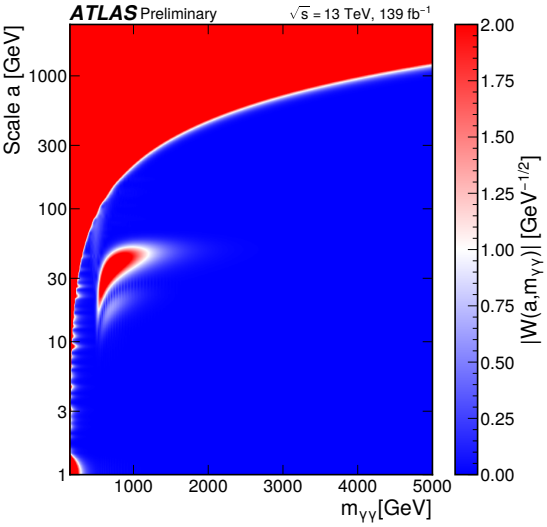


(a)

(b)

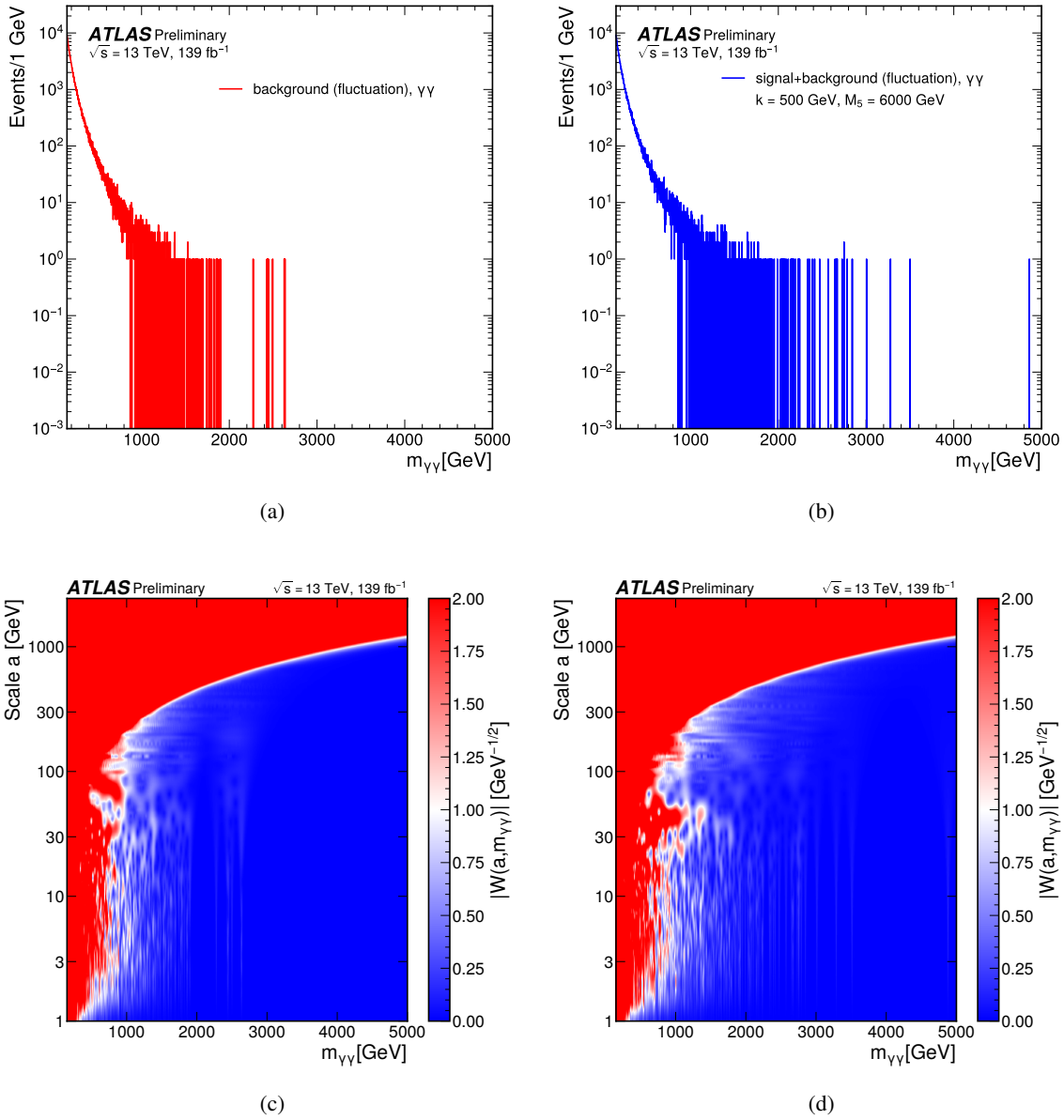


(c)



(d)

Figure 10: The diphoton invariant mass distribution for (a) background-only and (b) signal-plus-background (k = 500 GeV, M_5 = 6000 GeV) samples. Here the distributions are not Poisson-fluctuated. The corresponding scalogram plots for (c) background-only and (d) signal-plus-background samples are also shown. One can look at the scalogram plots to distinguish background-only distribution from signal-plus-background distribution.



(a) (b)

(c) (d)

Figure 11: The diphoton invariant mass distribution for (a) background-only and (b) signal-plus-background (k = 500 GeV, M_5 = 6000 GeV) samples. Here the distributions are Poisson-fluctuated. The corresponding scalogram plots for (c) background-only and (d) signal-plus-background samples are also shown. It is not easy to distinguish background-only scalogram plot from that of the signal-plus-background.

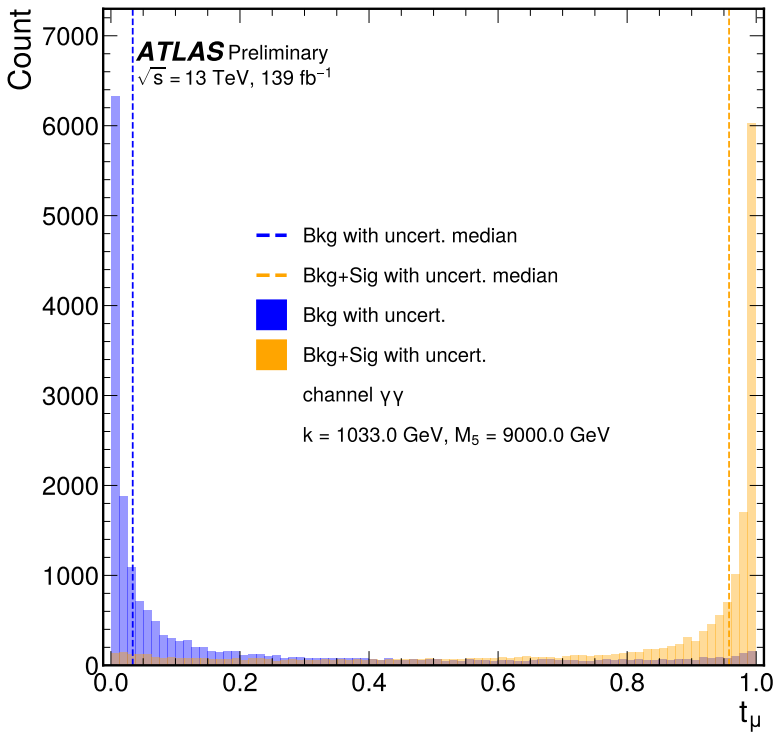


Figure 12: The distribution of the test statistic t_μ under the assumption of a signal model μ (orange) and under the assumption of the Standard Model (blue) for a signal with $k = 1033$ GeV and $M_5 = 9000$ GeV in the diphoton channel. The median of each distribution is displayed as a dashed line. The pseudo-experiments used to generate the test statistic include the effects from the statistical uncertainty and all systematic uncertainties.

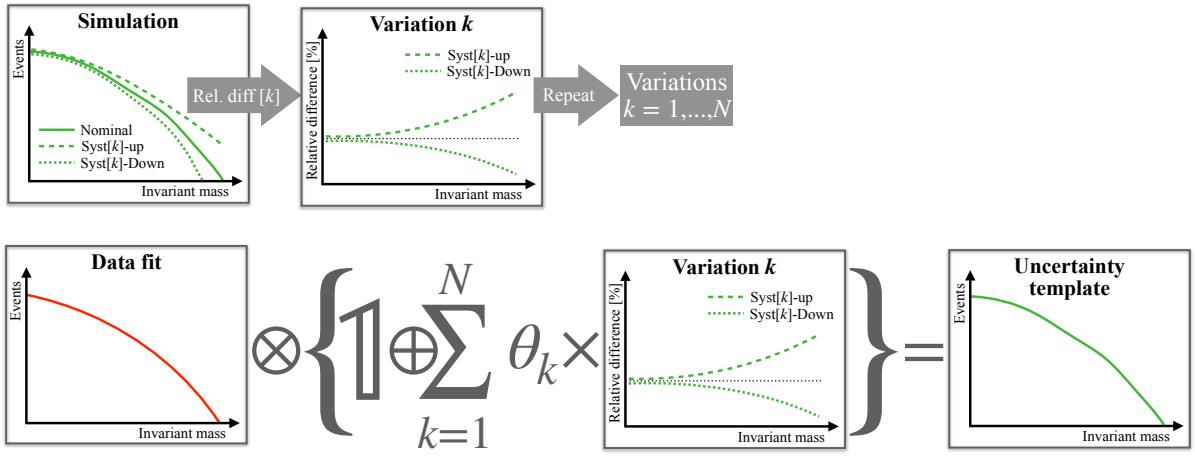


Figure 13: A simple illustration outlining the procedure to prepare the toy uncertainty mass distribution templates. Here, the relative variation in uncertainties with respect to the nominal sample is obtained from MC samples. The absolute variation is calculated by multiplying this distribution with the data-fit. The absolute variation for each uncertainty k is added to the data-fit with a weighting factor θ_k drawn randomly from a Gaussian distribution with mean 0 and standard deviation 1 to produce the uncertainty template. This uncertainty template is used to produce a syst-toy as shown in Figure 14.

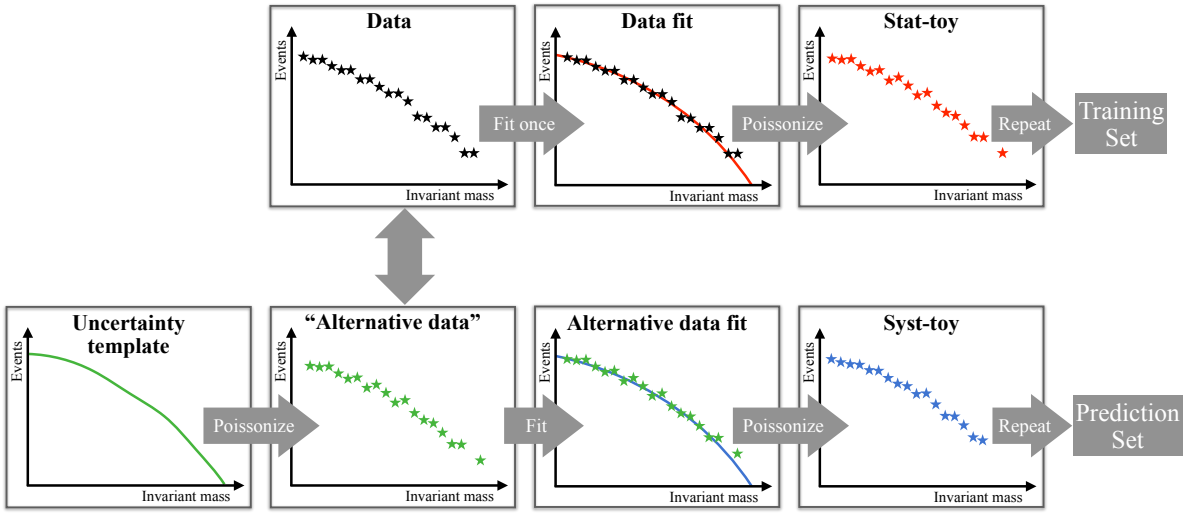
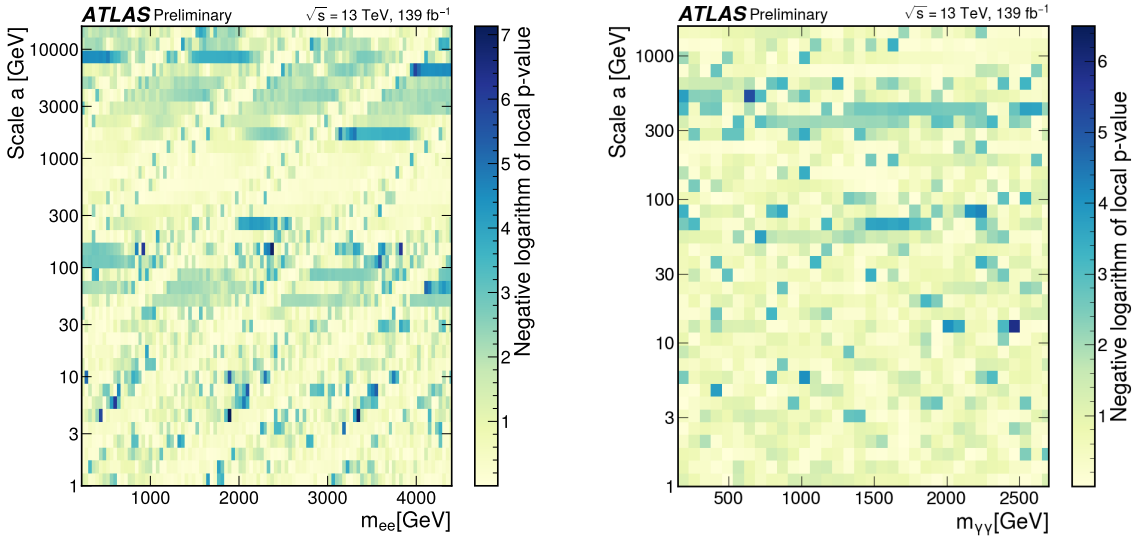


Figure 14: A simple illustration outlining the procedure to prepare stat-toys and syst-toys. In this analysis, stat-toys are used to train the NN and syst-toys are used to make the prediction set.



(a) (b)

Figure 15: The negative logarithm of the local p-values calculated for the observed data mass distributions in the (a) dielectron and (b) diphoton search channels for the case where no precision-driven threshold is applied. The range of m_{ee} and $m_{\gamma\gamma}$ is chosen such that the plots cover approximately 300 GeV after the last observed data point.

References

[1] G. F. Giudice and M. McCullough, *A Clockwork Theory*, JHEP **02** (2017) 036, arXiv: 1610.07962 [hep-ph] (cit. on p. 2).

[2] G. F. Giudice, Y. Kats, M. McCullough, R. Torre and A. Urbano, *Clockwork/linear dilaton: structure and phenomenology*, JHEP **06** (2018) 009, arXiv: 1711.08437 [hep-ph] (cit. on pp. 2, 7).

[3] I. Antoniadis, S. Dimopoulos and A. Giveon, *Little string theory at a TeV*, JHEP **05** (2001) 055, arXiv: hep-th/0103033 (cit. on p. 2).

[4] I. Antoniadis, A. Arvanitaki, S. Dimopoulos and A. Giveon, *Phenomenology of TeV Little String Theory from Holography*, Phys. Rev. Lett. **108** (2012) 081602, arXiv: 1102.4043 [hep-ph] (cit. on p. 2).

[5] M. Berkooz, M. Rozali and N. Seiberg, *Matrix description of M-theory on T^4 and T^5*, Phys. Lett. **B408** (1997) 105, arXiv: hep-th/9704089 (cit. on p. 2).

[6] N. Seiberg, *New theories in six-dimensions and matrix description of M-theory on T^5 and T^5/Z_2*, Phys. Lett. **B408** (1997) 98, arXiv: hep-th/9705221 (cit. on p. 2).

[7] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. B **429** (1998) 263, arXiv: hep-ph/9803315 (cit. on p. 2).

[8] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, *New dimensions at a millimeter to a Fermi and superstrings at a TeV*, Phys. Lett. B **436** (1998) 257, arXiv: hep-ph/9804398 (cit. on p. 2).

[9] L. Randall and R. Sundrum, *Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) 3370, arXiv: hep-ph/9905221 (cit. on pp. 2, 6).

[10] ATLAS Collaboration, *Search for high-mass dilepton resonances using $139\,fb^{-1}$ of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **796** (2019) 68, arXiv: 1903.06248 [hep-ex] (cit. on pp. 2, 3, 6, 8–12).

[11] ATLAS Collaboration, *Search for new non-resonant phenomena in high-mass dilepton final states with the ATLAS detector*, JHEP **11** (2020) 005, arXiv: 2006.12946 [hep-ex] (cit. on pp. 2, 3, 11), Erratum: JHEP **04** (2021) 142.

[12] ATLAS Collaboration, *Search for resonances decaying into photon pairs in $139\,fb^{-1}$ of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **822** (2021) 136651, arXiv: 2102.13405 [hep-ex] (cit. on pp. 2, 3, 6, 8, 10–12).

[13] 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] (cit. on p. 2).

[14] CMS Collaboration, *Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s} = 13$ TeV*, JHEP **07** (2021) 208, arXiv: 2103.02708 [hep-ex] (cit. on p. 2).

[15] H. Beauchesne and Y. Kats, *Searching for periodic signals in kinematic distributions using continuous wavelet transforms*, Eur. Phys. J. C **80** (2020) 192, arXiv: 1907.03676 [hep-ph] (cit. on pp. 3, 12, 13, 15, 16, 21–23).

[16] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on p. 4).

[17] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: https://cds.cern.ch/record/1291633 (cit. on p. 4), Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

[18] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det] (cit. on p. 4).

[19] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 4).

[20] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187 (cit. on p. 4).

[21] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on p. 4).

[22] 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 (cit. on p. 4).

[23] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 4).

[24] 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] (cit. on pp. 4, 6, 7).

[25] 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] (cit. on pp. 4, 5).

[26] 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 [hep-ex] (cit. on pp. 5, 10).

[27] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 5).

[28] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 5).

[29] 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] (cit. on p. 5).

[30] 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] (cit. on p. 5).

[31] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 (cit. on p. 5).

[32] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on p. 5).

[33] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 5).

[34] ATLAS Collaboration, *Search for new high-mass phenomena in the dilepton final state using* $36\,fb^{-1}$ *of proton–proton collision data at* $\sqrt{s} = 13\,TeV$ *with the ATLAS detector*, JHEP **10** (2017) 182, arXiv: 1707.02424 [hep-ex] (cit. on p. 5).

[35] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or* Z/γ^* *Boson in Association with Jets at ATLAS in Run 2*, ATL-PHYS-PUB-2016-003, 2016, URL: https://cds.cern.ch/record/2120133 (cit. on p. 5).

[36] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 5).

[37] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 5).

[38] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 5).

[39] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO vector-boson production matched with shower in POWHEG*, JHEP **07** (2008) 060, arXiv: 0805.4802 [hep-ph] (cit. on p. 5).

[40] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph] (cit. on p. 5).

[41] 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] (cit. on p. 5).

[42] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, Phys. Rev. D **69** (2004) 094008, arXiv: hep-ph/0312266 (cit. on p. 5).

[43] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006, arXiv: 1506.07443 [hep-ph] (cit. on p. 5).

[44] A. Arbuzov et al., *Update of the MCSANC Monte Carlo integrator, v. 1.20*, JETP Lett. **103** (2016) 131, arXiv: 1509.03052 [hep-ph] (cit. on p. 5).

[45] ATLAS Collaboration, *Multi-boson simulation for* 13 *TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986 (cit. on p. 5).

[46] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 5).

[47] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 5).

[48] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 5).

[49] S. Alioli, P. Nason, C. Oleari and E. Re,
NLO single-top production matched with shower in POWHEG: s- and t-channel contributions,
JHEP **09** (2009) 111, arXiv: 0907.4076 [hep-ph] (cit. on p. 5), Erratum: JHEP **02** (2010) 011.

[50] R. Frederix, E. Re and P. Torrielli,
Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO,
JHEP **09** (2012) 130, arXiv: 1207.5391 [hep-ph] (cit. on p. 5).

[51] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*,
Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph] (cit. on p. 5).

[52] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159,
arXiv: 1410.3012 [hep-ph] (cit. on p. 5).

[53] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014,
URL: https://cds.cern.ch/record/1966419 (cit. on p. 5).

[54] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph] (cit. on p. 5).

[55] M. Czakon and A. Mitov,
Top++: A program for the calculation of the top-pair cross-section at hadron colliders,
Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph] (cit. on p. 5).

[56] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965 (cit. on p. 6).

[57] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at* $\sqrt{s} = 13$ *TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002,
arXiv: 1606.02625 [hep-ex] (cit. on p. 6).

[58] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det] (cit. on p. 6).

[59] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 6).

[60] 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 (cit. on p. 6).

[61] 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] (cit. on p. 6).

[62] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022, 2017,
URL: https://cds.cern.ch/record/2298955 (cit. on p. 7).

[63] M. Oreglia, *A Study of the Reactions* $\psi' \to \gamma\gamma\psi$, SLAC-R-0236 (1980),
URL: www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf (cit. on p. 8).

[64] M. Oreglia et al., *Study of the reaction* $\psi' \to \gamma\gamma\frac{J}{\psi}$, Phys. Rev. D **25** (9 1982) 2259,
URL: https://link.aps.org/doi/10.1103/PhysRevD.25.2259 (cit. on p. 8).

[65] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001 (cit. on p. 9).

[66] ATLAS Collaboration,
Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector,
JHEP **09** (2016) 001, arXiv: 1606.03833 [hep-ex] (cit. on p. 10).

[67] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne,
Parton distributions incorporating QED contributions, Eur. Phys. J. C **39** (2005) 155,
arXiv: hep-ph/0411040 (cit. on p. 10).

[68] J. Morlet, G. Arens, E. Fourgeau and D. Giard, *Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media*, Geophysics **47** (1982) 203 (cit. on p. 12).

[69] M. X. Cohen, *A better way to define and describe Morlet wavelets for time-frequency analysis*,
NeuroImage **199** (2019) 81, ISSN: 1053-8119,
URL: https://www.sciencedirect.com/science/article/pii/S1053811919304409
(cit. on p. 12).

[70] F. Chollet et al., *Keras*, 2015, URL: https://keras.io (cit. on p. 13).

[71] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*,
(2016), arXiv: 1603.04467 [cs.DC] (cit. on p. 13).

[72] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, (2014),
arXiv: 1412.6980 [cs.LG] (cit. on p. 13).

[73] T. Cohen, M. Freytsis and B. Ostdiek, *(Machine) Learning to Do More with Less*,
JHEP **02** (2018) 034, arXiv: 1706.09451 [hep-ph] (cit. on p. 13).

[74] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693
(cit. on p. 14).

[75] T. Junk, *Confidence level computation for combining searches with small statistics*,
Nucl. Instrum. Meth. A **434** (1999) 435, arXiv: hep-ex/9902006 (cit. on p. 14).

[76] P. A. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) 083C01 (cit. on p. 14).

[77] T. Heimel, G. Kasieczka, T. Plehn and J. M. Thompson, *QCD or What?*, SciPost Phys. **6** (2019) 030,
arXiv: 1808.08979 [hep-ph] (cit. on p. 15).

[78] M. Farina, Y. Nakai and D. Shih, *Searching for New Physics with Deep Autoencoders*,
Phys. Rev. D **101** (2020) 075021, arXiv: 1808.08992 [hep-ph] (cit. on pp. 15, 16).

[79] G. R. Lee et al., *PyWavelets - Wavelet Transforms in Python*,
https://github.com/PyWavelets/pywt, 2019 (cit. on pp. 21, 22).

[80] G. R. Lee, R. Gommers, F. Wasilewski, K. Wohlfahrt and A. O'Leary,
PyWavelets: A Python package for wavelet analysis,
Journal of Open Source Software **4(36)** (2019) 1237 (cit. on pp. 21, 22).