ATLAS Note

ANA-STDM-2018-54-INT1

6th March 2023

Draft version 0.1

$Z(\to \nu\bar{\nu})\gamma$+jets differential cross-section measurements and search for neutral triple gauge couplings in 13 TeV pp collisions with the ATLAS detector

Soldatov, Evgeny[a], Pyatiizbyantseva, Diana[a], Romaniouk, Anatoli[a], Semushin, Artur[a], Kazakova, Katerina[a], Kurova, Anastasia[a], Zubov, Dmitriy[a], Simbiryatin, Lev[a]

[a]*National Research Nuclear University MEPhI (RU)*

The production of Z bosons in association with a high-energy photon ($Z\gamma$ production) is studied in the neutrino decay channel of the Z boson using pp collisions at $\sqrt{s}$ = 13 TeV. The analysis uses a data sample with an integrated luminosity of 140 fb^{-1} collected by the ATLAS detector during Run2 of the LHC. Differential cross-section measurements of $Z\gamma$ production in association with hadronic jets are performed for the observables, including those sensitive to the hard scattering in the event and others which are sensitive to CP violation, expected for some of beyond Standard Model theories. Search of neutral triple gauge-boson couplings is performed in $Z\gamma$ production with photon E_T greater than X GeV. No excess is observed relative to the Standard Model expectation, and upper limits are set on the strength of $ZZ\gamma$ and $Z\gamma\gamma$ couplings both in vertex function and effective field theory approaches.

Contents

List of contributions

Soldatov, Evgeny	Analysis contact, supervisor of the group, MC and derivation requests, theoretical cross section
Pyatiizbyantseva, Diana	Analysis co-contact, data-driven background estimation ($j \rightarrow \gamma$, pile-up), statistical model, framework maintenance, Katerina's supervisor
Romaniouk, Anatoli	Team leader, global supervisor
Semushin, Artur	Anomalous TGC interpretations
Kazakova, Katerina	Alternative data-driven background estimation ($j \rightarrow \gamma$), beam-induced background suppression, cross-section extraction, unfolding
Kurova, Anastasia	Framework development and maintenance, analysis ntuples production manager, data-driven background estimation ($e \rightarrow \gamma$)
Zubov, Dmitriy	Optimisation of the signal phase space
Simbiryatin, Lev	Data-driven background estimation ($j \rightarrow E_{\mathrm{T}}^{\mathrm{miss}}$)

1 Introduction

In the absence of clear indications of new physics in collected data, precision measurements of Standard Model (SM) processes at LHC remain highly important. In particular, differential distributions play a key role to constrain parameters of both SM Lagrangian and its extensions.

This note presents integrated and differential cross-section measurements of the production of Z boson associated with an isolated photon of high energy in ATLAS experiment. Such study of the boson pairs production has been used in many experiments for high precision tests of the Standard Model's (SM) electroweak sector. This production is sensitive to the triple gauge-boson self-couplings (TGCs), which are the consequence of the non-Abelian nature of the electroweak symmetry group $SU(2)_L \times U(1)_Y$. The couplings of the Z boson to other bosons have been observed and they are in agreement with the SM predictions in LEP, Tevatron and LHC experiments. No experimental evidence has been reported for couplings of Z bosons to photons. These neutral couplings are absent in SM at tree level. Anomalous properties of the Z boson are often constrained in terms of limits on the triple gauge-boson couplings ($ZZ\gamma$ and $Z\gamma\gamma$). Such limits have been reported by many experiments at the LEP, the Tevatron and the LHC.

The measurement in this note uses full Run2 dataset of 140 fb^{-1} of proton–proton (pp) collisions data collected with the ATLAS detector at the CERN LHC which operates at a center of mass energy of 13 TeV.

The analysis uses $Z \to \nu\bar{\nu}$ decay channel. The $\nu\bar{\nu}\gamma$ final state in the SM mainly can be produced in the process with the photon emission by initial state quarks (left diagram in Figure 1). Example of the hypothetical triple gauge-boson coupling involving Z bosons and photons is shown by the right diagram in Figure 1. The absence of the final state photon radiation from neutrinos leads to the measurement of pure $Z + \gamma$ process (without presence of Z boson decay with photon in final state), which can be used as a probe of boson self-couplings. Also it leads to the simplification in the truth level definition, since one do not need to «dress» neutrinos with QED final state radiation photons as electrons or muons in the case of $Z(\ell^+\ell^-)\gamma$ channel. These advantages together with much higher Z boson branching ratio into neutrinos lead to the highest sensitivity of this channel to the neutral anomalous triple gauge couplings. Excellence over the hadronic channel of Z boson decay achieved due to much better background control.

The events are selected using high E_T single photon trigger. Integrated cross section measurement is made for the processes $pp \to \nu\bar{\nu}\gamma + X$ with the transverse energy of the photon higher than 150 GeV.

Cross section is measured differentially for the main kinematic observables and also for sensitive observables to the beyond SM theories manifestations. Differential cross sections are compared to the higher order SM predictions. The accuracy of the measurement at such high energies exceeds similar measurements done with charged lepton channels of Z boson decay. It allows to perform a crucial test for SM and its possible extensions.

The measured $Z\gamma$ production cross section at high values of the photon E_T is used to search for anomalous triple gauge-boson ($ZZ\gamma$ and $Z\gamma\gamma$) couplings (aTGC).

This note is structured as follows. The ATLAS detector is briefly described in Section 2. The signal and background simulation is presented in Section **??**. The object and event selections are described in Section 3. The methodologies for the estimation of various backgrounds in this study are discussed in Section 4. The discussion on different types of uncertainties for the measurement is given in Section **??**. The results extraction precedure of integrated and differential cross-section measurements and their comparison with the Standard Model predictions are presented in Sections 5 and **??**. The limits on the anomalous triple and quartic gauge-boson couplings are presented in Section **??**. Section 6 provides the conclusions.

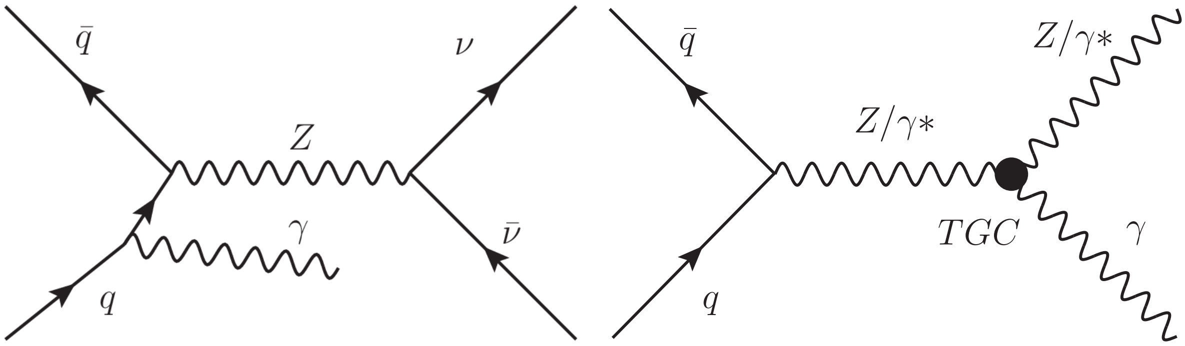


Figure 1: Feynman diagrams of $Z(\nu\bar{\nu})\gamma$ production: (left) initial-state photon radiation (ISR) and (right) hypothetical neutral triple gauge-boson coupling (TGC) vertex.

2 ATLAS detector

The ATLAS detector [1] is a multi-purpose particle detector with a forward–backward symmetric cylindrical geometry and coverage of nearly the entire 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 (ECAL) and hadronic (HCAL) calorimeters, and a muon spectrometer (MS).

The ID is used for precise measurements of charged-particle tracks. It is composed of two silicon detectors covering the pseudorapidity range $|\eta| < 2.5$: a pixel detector (including the insertable B-layer [2, 3]) and a silicon microstrip tracker, surrounded by a straw-tube transition radiation tracker (TRT) with an acceptance of $|\eta| < 2.0$, which also contributes to electron identification.

The ECAL is composed of high-granularity lead/liquid-argon (LAr) calorimeters in the region $|\eta| < 3.2$ and copper/LAr calorimeters in the region $3.2 < |\eta| < 4.9$. It plays a crucial role in photon identification, since photons are identified as narrow isolated showers in the ECAL. The HCAL consists of a steel/scintillator-tile calorimeter within $|\eta| < 1.7$ and two copper/LAr and tungsten/LAr forward calorimeters within $1.7 < |\eta| < 4.9$. The fine segmentation of the ATLAS calorimeter system allows efficient separation of jets from isolated prompt photons.

The MS comprises three large superconducting toroids, each having eight coils, as well as trigger and high-precision tracking chamber systems that cover the regions $|\eta| < 2.4$ and $|\eta| < 2.7$, respectively.

The ATLAS trigger system [4] has two levels, a hardware-based first-level trigger and a software-based high-level trigger (HLT). The trigger system selects events from the 40 MHz LHC proton bunch crossings at a rate of about 1 kHz.

An extensive software suite [5] 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.

3 Object and Event Selection

The following section describes the selection criteria for all objects and event candidates.

More specifically, the definition of photons, jets and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}/E_{\mathrm{T}}^{\mathrm{miss}}$ can be found in Sections 3.1, 3.2 and 3.4 respectively. Selection criteria for leptons used in background evaluations and for lepton veto are presented in Section 3.3. In addition to the basic object selection, it is possible that two or more reconstructed objects overlap in (η, ϕ) space. The definition and order of this overlap removal are explained in Section 3.5. For the MC predictions, the lepton and photon data/MC efficiency correction scale factors are included. The simulated lepton and photon four-momenta are also tuned via calorimeter energy scaling and momentum resolution smearing to reproduce the distributions observed in the data. The event selection criteria are described in Section 3.6.

[1] A right-handed coordinate system is used 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 $\eta = -\ln\tan(\theta/2)$. The angular distance between two physics objects is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3.1 Photons

Photons are reconstructed from the clusters measured in projective towers of $N_1 \times N_2$ cells in $\eta \times \phi$ of the second layer of the electromagnetic calorimeter. Clusters without matching tracks are classified as unconverted photon candidates. Clusters matched to tracks originating from conversion vertices in the inner detector or to tracks consistent with coming from a conversion are considered as converted photon candidates [6].

The photon energy reconstructed from data is known to be miscalibrated. The photon energy is calibrated differently depending on its classification as converted or unconverted. The energy is corrected using EgammaCalibrationAndSmearingTool [7]. The energy scales are measured on well-known resonances ($Z \to e^+e^-$ and $J/\Psi \to e^+e^-$) or E/p studies using isolated electrons from the $W \to e\nu$ process. The photon energy in the MC simulation is known to not reproduce the resolution of the ATLAS detector. The photon energy in the MC simulation is then smeared such that the resolution in the sample is found to match the ATLAS detector using also EgammaCalibrationAndSmearingTool.

The photon ambiguity resolver is used on reconstruction level to reject against possible electrons misidentified as converted photons.

The shower shapes produced in the electromagnetic calorimeter by photon candidates provide a handle to distinguish photons from hadronic decays which may mimic a prompt photon. These shower shapes are not perfectly modeled in the MC simulation. As such the fudge factors are used to correct the photons shower shapes, performing simple shifts of the shower shape distributions. Fudge factors are applied using ElectronPhotonShowerShapeFudgeTool [8] and following the recommendations indicated in Ref. [9].

After the showers are corrected in the MC or taken directly in the data, the egammaPID routine is used to apply rectangular cuts on the shower shape distributions to obtain identification quality of photons. The photon must have good object quality flags to reject against photon candidates which may arise due to detector noise or jets misidentified as photons. A photon is flagged as bad if it contains or edges dead/masked cells. There are two offline menus, *loose* and *tight*, with different quality of identification. *Loose* criterion, which is mostly in use for the trigger purposes, contains cuts on hadronic leakage variables and shower shapes in the Middle EMC layer. *Tight* selection criterion contains cuts on each shower shape variable. *Tight* selection provides high identification quality and can be used for the offline selection in most of physical analyses. *Tight* working point is used in this analysis. Identification efficiency for *tight* photons was found to be greater than 88% in Run 2 [10]. *Loose* photons are selected in order to help modelling jets misidentified as photons for the data-driven background estimation methods described below. Specific ID scale factors provided by PhotonEfficiencyCorrectionTool [11] are applied to photons in MC to match the efficiency measured in data.

The preselected photon candidate is required to have an $|\eta|$ position within 2.37. The additional requirement is that photon must not be found within the calorimeter transition region ($1.37 < |\eta| < 1.52$). The photon is required to have E_T^{γ} greater than 150 GeV. It is motivated by the high energy threshold of the single photon trigger (140 GeV) used in the analysis. A more detailed description of the trigger is given in Sec. 3.6.1.

Photon isolation is computed from both isolation energies of the tracker (which provides a more pile-up independent isolation) and of the calorimeter (which detects also neutral hadrons) [9]. A calorimeter-based variable E_T^{cone20} constructed as the sum of the transverse energies (at the electromagnetic scale) of positive energy topological clusters located within a distance $\Delta R = 0.2$ of the photon candidate. It is corrected for various contributions with a respect to the hardest vertex. A track-based variable p_T^{cone20} constructed as the sum of the transverse momenta of good tracks located within a distance $\Delta R = 0.2$ of the photon candidate. For converted photon candidates, tracks linked to associated photon vertex are excluded.

The isolation requirement on photon candidates in the current analysis corresponds to the *FixedCutLoose* working point defined as:

- $E_{\mathrm{T}}^{\mathrm{cone20}} < 0.065 p_{\mathrm{T}}$ [GeV];
- $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.05$.

The choice of this working point was made during the selection optimization, which is described in Section **??**. The details of the comparison of the working points are shown in Table 16 of Appendix A.1.

Finally, the absolute value of z coordinate pointed by the photon candidate with respect to the identified primary vertex is required to be less than 250 mm. This criterion was included in the photon candidates selection to suppress the anomalous rate of unconverted photons caused by beam-induced background in the *loose* and isolated region of the data-driven background estimation method for events with jets misidentification as photons [12].

3.2 Jets

Jets are reconstructed with the particle flow anti-k_{T} clustering algorithm [13] using a radius parameter $R = 0.4$. The particle flow algorithm utilizes the information from a list of tracks and a list of calorimetric topo-clusters to reconstruct the hadronic jets and soft activity (additional hadronic recoil used in the jet reconstruction). First, high quality tracks with $0.5 < p_{\mathrm{T}} < 40$ GeV, $|\eta| < 2.5$, with at least nine hits in the silicon detectors, and no missing Pixel hits are selected. Each of the selected tracks is attempted to match with a single topo-cluster. The particle flow algorithm additionally determines the probability of the particle energy to be deposited in more than one cluster and adds more topo-clusters if necessary to recover the full shower energy Preselected jets must have $p_{\mathrm{T}} > 20$ GeV. In order to suppress jets originating from pile-up, for jets with $20 < p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ the output of the jet vertex tagger (JVT) is required to be larger than 0.5.

3.3 Leptons

3.3.1 Leptons

Muon candidates are reconstructed by the algorithms that perform a statistical combination of a track, reconstructed in the muon spectrometer, with a corresponding track in the inner detector.

All muon identification selection criteria in this analysis come from recommendations proposed by the muon combined performance (CP) group [14]. Before the muon selection is done, momentum calibration is applied to the raw muons from the reconstruction. In MC samples, momentum smearing is applied to improve data-MC agreement. Basic selection is done using the MuonSelectionTool [15] and configured at the *Loose* working point.

The p_{T} of a preselected reconstructed muon is required to be larger than 4 GeV. The pseudorapidity of a reconstructed muon must have $|\eta| <2.7$. To be sure that the muons have to come from the primary vertex of the interaction the transverse and the longitudinal impact parameters were introduced. The transverse impact parameter, $|d_0|/\sigma(d_0)$, is determined as the significance of the distance of the closest approach to the primary vertex in the transverse plane, and it is required to be less than 3. The longitudinal impact parameter, $|z_0 \cdot \sin\theta|$, must be less than 0.5 mm, where θ is a track zenith angle.

Isolation requirements on the muons are also imposed. *FixedCutLoose* isolation working point is used for that purpose [16]. Scale factors correcting reconstruction and selections efficiencies in the MC to those in data are applied to muons in the form of object weights using the MuonEfficiencyScaleFactors tool [17] provided by the CP group. Scale factors coming from different sources are multiplied.

3.3.2 Electrons

An electron candidate is obtained from an energy cluster in the EM calorimeter associated with a reconstructed track in the Inner Detector (ID). The EgammaCalibrationAndSmearingTool from last CP recommendations [7] is applied for raw electrons to calibrate electron cluster energy in the data and electron energy resolution in the MC. The "preselected" electrons are required to pass a variant of the *loose* particle identification selection criterion called "LooseCutBL". This selection includes B-layer cut, which is very important in separation of prompt electrons and those coming from photon conversions.

The selection criterion requires that the preselected electrons transverse energy is greater than 4.5 GeV. To maintain good tracking requirements an electron candidate must be found with an $|\eta|$ position within 2.47, but out off the calorimeter transition region ($1.37 < |\eta| < 1.52$). To be sure that electrons have to come from the primary vertex of the interaction, $|d_0|/\sigma(d_0)$, is required to be smaller than 5, and $|z_0 \cdot \sin\theta|$ must be less than 0.5 mm.

Isolation requirements on electrons are also imposed. *FixedCutLoose* isolation working point is used for that purpose [16]. Scale factors correcting reconstruction and selections efficiencies in the MC to those in data are applied to electrons in the form of object weights using the AsgElectronEfficiencyCorrectionTool tool [17] provided by the CP group. Scale factors coming from different sources are multiplied.

3.3.3 Taus

"Taus" smeared

3.4 Missing Transverse Momentum and Energy

The missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is the vector of momentum imbalance in the transverse plane. The reconstruction of the direction and magnitude of the missing transverse momentum vector is described in Ref. [18]. $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as a magnitude of transverse momentum vector $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$.

Missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is a hallmark of neutrino production in hadron colliders. Due to the conservation of momentum, if all particles produced in the primary collision were detectable, then there should be no $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event unless it arises from detector-level effects e.g. resolution, material effects, or uninstrumented regions of the detector. However, events which were produced with the neutrinos carrying large amounts of p_{T} can be expected to have large $E_{\mathrm{T}}^{\mathrm{miss}}$.

METtool [19] is used for $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction . The magnitude of missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is based on the energy deposited in calorimeter cells up to $|\eta| < 4.9$ and muons with $p_{\mathrm{T}} > 4$ GeV.

Each cell in the calorimeter is associated with some object, which is then used to define a calibration for the signal observed in the cell. Preselected electrons with $p_{\mathrm{T}} > 4.5$ GeV, muons with $p_{\mathrm{T}} > 4$ GeV, photons with $p_{\mathrm{T}} > 10$ GeV and jets with $p_{\mathrm{T}} > 20$ GeV are given as input to the $E_{\mathrm{T}}^{\mathrm{miss}}$ rebuilding algorithm. Energy deposits/tracks not associated with any objects are also taken into account in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation [20].

Missing transverse momentum is calculated as the sum of the following terms:

$$E_{\mathrm{x(y)}}^{\mathrm{miss}} = E_{\mathrm{x(y)}}^{\mathrm{miss,e}} + E_{\mathrm{x(y)}}^{\mathrm{miss},\gamma} + E_{\mathrm{x(y)}}^{\mathrm{miss,jets}} + E_{\mathrm{x(y)}}^{\mathrm{miss,SoftTerm}} + E_{\mathrm{x(y)}}^{\mathrm{miss},\mu}, \tag{1}$$

where each term is calculated as the negative sum of the calibrated reconstructed objects, projected onto the x and y directions and from the "Soft Term". Soft Term is calculated in the present analysis using the tracks from the primary vertex that are not matched to selected hard objects (so called "Track Soft Term" or TST). It provides a more robust measurement against pileup [21].

3.5 Object Overlap Removal

Overlap removal (OR) is performed for two reasons. The first reason is to make sure that a single object is not double-counted as two different objects (e.g. that a photon is not counted as a jet). The second reason is that there are isolation requirements for leptons and photons and allowing for an additional object nearby would contaminate some of these isolation requirements.

The overlap requirements are defined below and applied one by one.

- All objects are reconstructed. Leptons are selected as described in Sec. 3.3. Jets are selected according to Sec. 3.2 apart from JVT cut, which is applied after OR according to recommendations in [22]. Photons are required to be loose not-isolated and satisfy other selection described in Sec. 3.1.
- Electrons within $\Delta R < 0.1$ of a muon are removed.
- Photons within $\Delta R < 0.4$ of either a muon or an electron are removed.
- Jets within $\Delta R < 0.3$ of a photon, muon, or electron are removed.

The overlap removal is applied after missing energy reconstruction since METtool has its own implementation of overlap removal.

3.6 Event Selection

The resulting selection criteria for such objects as photons, jets and leptons are described in Table 1.

The event selection criteria are chosen to provide precise cross-section measurements of $Z(\nu\nu)\gamma$ production and good sensitivities to anomalous gauge-boson couplings.

Photon	Jet	Muon	Electron
$E_T > 150$ GeV, cluster quality cut, ambiguity cut, $\|\eta\| < 2.37$, photon cleaning, Loose ID, crack region rejection, $\Delta R(\gamma, e/\mu) < 0.4$, FixedCutLoose isolation	AntiKt4EMPFlowJets, $p_T > 50$ GeV, $\|\eta\| < 4.5$, $\Delta R(jet, e/\mu) < 0.4$, JVT cut	Loose ID, $p_T > 4$ GeV, $\|\eta\| < 2.47$, $\|z_0 \cdot \sin\theta\| < 0.5$ mm, d_0 signif. < 3, isolation FixedCutLoose	Cluster quality cut, LooseCutBL ID, $p_T > 4.5$ GeV, $\|\eta\| < 2.47$, crack region excluded, $\|z_0 \cdot \sin\theta\| < 0.5$ mm, d_0 signif. < 5, isolation FixedCutLoose, $\Delta R(e, \mu) < 0.1$

Table 1: Definition of photons, jets and leptons.

3.6.1 Event Preselection

Since the final state contains one photon and the high-energetic region is very sensitive to aTGCs, the lowest E_T threshold unprescaled single photon trigger `HLT_g140_loose` is used for this analysis. This high-level single photon trigger requires a transverse energy E_T greater than 140 GeV and loose photon identification criteria.

The same trigger is simulated and applied in the MC signal and background production. Efficiency of this trigger for the photon candidates reconstructed offline and passing the tight identification selection is 100% for $E_T > 150$ GeV region [23]. The trigger SF is not applied in the analysis.

The preselection can be defined as follows. Candidate events are required to have one leading tight isolated photon with transverse energy $E_T^\gamma > 150$ GeV and at least two jets.

In order to reduce the contamination coming from events which do not contain real neutrinos (mainly γ + jet background with instrumental E_T^{miss}) it is required that the selected events have $E_T^{\text{miss}} > 120$ GeV.

To reduce the number of $W(l\nu)\gamma$ and $Z(ll)\gamma$ events, lepton veto is applied: events with any preselected electrons or muons are discarded.

The preselections before the optimization procedure:

- $E_T^{\text{miss}} > 120$ GeV;
- exactly one tight isolated photon with $E_T^\gamma > 150$ GeV;
- charged lepton (e/μ) veto in the event.

3.6.2 Selection Optimisation

The selection optimisation is applied starting from the preselected events and used for obtaining high statistical significance by finding optimal constraints on the variables. The statistical significance is defined as:

$$S = N_{\text{signal}}/\sqrt{N_{\text{signal}} + N_{\text{bkg}}}, \tag{2}$$

where N_{signal} and N_{bkg} are accumulated number of signal and background events respectively. The following selections are used in the optimisation procedure to find the optimal threshold values: $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $E_{\mathrm{T}}^{\mathrm{miss}}$, $|\Delta\phi(\gamma, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})|$, $|\Delta\phi(j_1, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})|$.

Statistical significance is considered as a function of variable constraints. To construct a multivariate dependence of statistical significance on thresholds on variables, multivariate histograms for signal and background processes were filled in, and statistical significance values were calculated for each variant of phase space limitation on their basis. Such multivariate histograms are implemented using the `THnSparse` class based on the ROOT package. The results of the optimisation procedure are presented in Table 2.

Selections	Cut Value
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 130 GeV
E_{T}^{γ}	> 150 GeV
Number of tight isolated photons	$N_{\gamma} = 1$
Lepton veto	$N_{\mathrm{e}} = 0$, $N_{\mu} = 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$ significance	> 11
$\lvert\Delta\phi(\gamma, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})\rvert$	> 0.7
$\lvert\Delta\phi(j_1, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})\rvert$	> 0.4

Table 2: Event selection criteria for $Z(\nu\bar{\nu})\gamma$ candidate events.

While selection described in Table 2 defines the SR, the $W\gamma$ CR is defined by keeping all of selection the same but requiring at least one lepton in the event.

4 Background estimation

4.1 Background induced by $e \to \gamma$ misidentification

It is known that electrons can be misidentified as converted photons, which produce closeby e^+e^- pairs and therefore leave tracks in ID. An average rate of such misidentification ranges from 2 to 18% for different momenta and pseudorapidity of the particle [24]. Given the fact that the main source of this background $W(\to e\nu)$ production has cross section of more than 2 orders larger than $Z(\to \nu\bar{\nu})\gamma$, the contribution from such misidentification is not negligible. Moreover, there are other processes, which have similar final state to $e\nu$, for example single-top and $t\bar{t}$ production with successive decay of one of the top-quarks to W+jets. MC prediction for $e \to \gamma$ misidentification is not reliable enough, therefore data-driven estimation is used instead. In order to derive contribution of the fake photons to the signal events yield two steps are made. The first step is to measure the probability of electron to photon misidentification (fake rate). This is done via tag-and-probe method, where tag-and-probe particle pairs are considered to have invariant mass close to Z boson mass. Good quality electron is considered as *tag* particle and another electron or photon – as *probe*. Then the fake rate will be determined as the number of events ratio of electron-photon pairs to electron-electron pairs.

The second step is to estimate the background contribution from processes, where electron faking photon. To obtain the final number of such events in the signal region the fake rate is used to scale probe-electron control region (e-probe CR). Detailed description of the procedure can be found in Sec. 4.1.2

4.1.1 Electron-to-photon fake rate estimation

To perform the electron-to-photon fake-rate calculation using forementioned tag-and-probe method the following selections are used:

- Electrons:
  - Both *tag* and *probe* electrons are selected as described in Sec. 3.3.2 except for the identification criteria and transverse momentum. *Tight* criteria is used for both *tag* and *probe* candidates. Transverse momentum larger than 25 GeV is required for the *tag* electron and larger than 150 GeV for the *probe* electron to fit the photon selection. Also *tag* and *probe* electrons should have charges of opposite sign.
  - For "real fake-rate" estimation using MC both *tag* and *probe* electrons were checked with MCTruthClassifier [25] to have association at truth level with an isolated electron (*truth type 2*), which comes from Z or W boson (*truth origin* 13 or 12)
- Photons
  - The *probe* photon is selected as described in Sec. 3.1
  - For "real fake rate" *probe* photon was also checked with MCTruthClassifier to have association at truth level with an isolated electron (*truth type 2*) or final state radiation (*truth type 15*) and also originate from Z/W boson or final state radiation (*truth origin* 13/12 or 40).

Event is selected if it contains one of tag-and-probe pair, which has invariant mass within 20 GeV window around Z boson mass, and passes $E_T^{miss} < 40$ GeV requirement to reduce $W\gamma$ background. Data-driven estimation supposes, that these events also should pass all basic event selection including trigger `HLT_g140_loose`.

Since the Z boson cannot decay to an electron and a photon, the most of $e\gamma$ events with mass near the Z pole must be made of electrons misidentified as photons. A first estimate of the electron-to-photon fake rate can be found by taking the ratio of reconstructed $e\gamma$ events to reconstructed e^+e^- events (as defined above). However as it is shown in Fig. 2 fake rate depends on η and p_T of the *probe* particle. To take these dependences into account fake-rate estimation was made in 3 regions. Photon pseudorapidity is divided on central ($|\eta| < 1.37$) and forward ($1.52 < |\eta| < 2.37$) regions[2]. Low-p_T ($150 < p_T < 250$ GeV) and high-p_T ($p_T > 250$ GeV) separation is made only in central region, since in forward one the fake rate distribution on p_T is flat within uncertainty.

Apart from $Z(\rightarrow e^+e^-)$ events, selected e^+e^- and $e\gamma$ pairs can be produced by Drell-Yann production of e^+e^- pairs. To improve the fake-rate estimation it is necessary to remove this background. This was done via side-band fit (see Appendix B.1)

There are 3 components of systematic uncertainty on the electron-to-photon fake rate (listed in ascending order by their magnitude in relation to the fake-rate value):

- Uncertainty on the bias caused by the Z peak mass window choice. It is estimated via variation of mass window width to 1σ of event yield inside it. Calculation is made on Z(ee) MC to avoid additional uncertainty caused by background subtraction (varies from 0.3% to 0.8% in different regions).

[2] Thinner division was found to be impossible due to lack of statistics (see App. C).

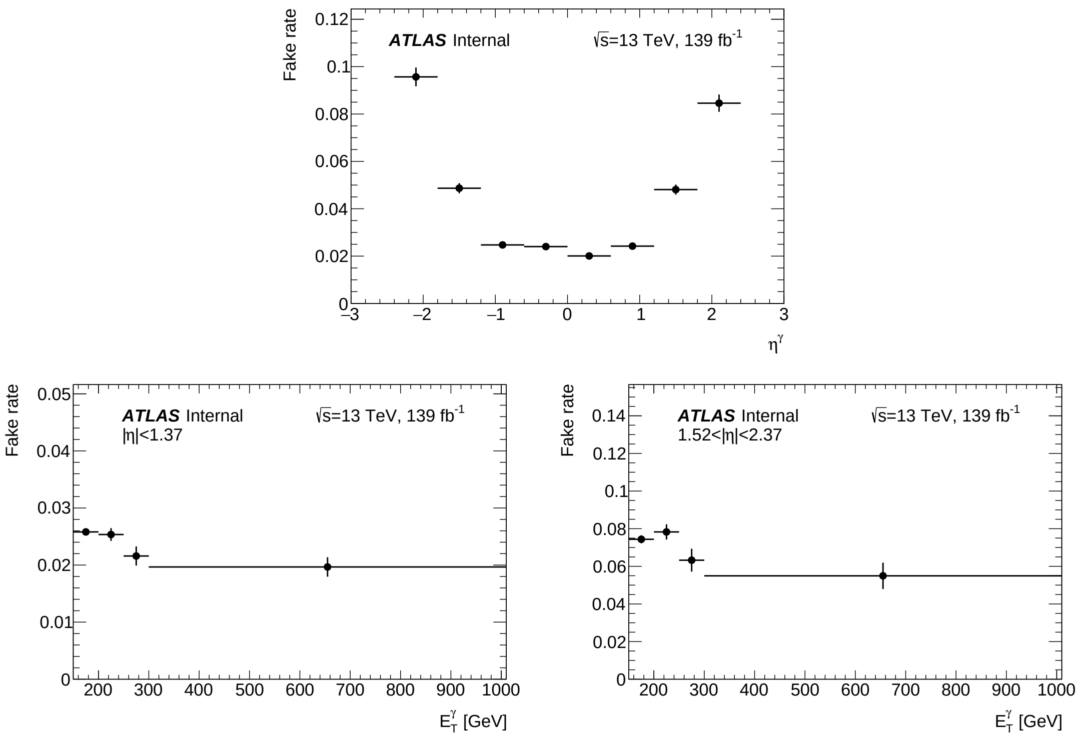


Figure 2: Fake-rate dependencies versus photon η (top) and p_T (bottom) without combinatorial background subtraction.

$e \to \gamma$ fake rates		
	$150 < E_T^\gamma < 250$ GeV	$E_T^\gamma > 250$ GeV
$0 < \lvert\eta\rvert < 1.37$	$0.0234 \pm 0.0006 \pm 0.0011$	$0.0195 \pm 0.0013 \pm 0.0041$
$1.52 < \lvert\eta\rvert < 2.37$	$0.0704 \pm 0.0018 \pm 0.0083$	

Table 3: Electron-to-photon fake rates from data in separate regions: splitted using photon E_T and η. The first error is statistical, the second one is systematical.

- The bias coming from estimation of combinatorial background under Z peak in data. It is estimated from variation of the fit extrapolation to the region under the peak from both sides of the tag-n-probe mass spectra away from the peak (varies from 3% to 11% in different regions).

- Uncertainty on the method itself is estimated via difference between "real fake rate" in Z(ee) MC and tag-and-probe method performed on $Z(ee)$ MC (varies from 2.8% to 18% in different regions).

The "real fake rate" is estimated in $Z(ee)$ basing on origin and type of the particles provided by MCTruthClassifier as described above.

Detailed description of fake-rate and "real fake rate" calculation for each uncertainty can be found in Appendix B. All systematic uncertainties described above are added in quadratures. Resulting values of electron-to-photon fake rate can be found in Table 3. First uncertainty is statistical, second is systematical. Systematical uncertainty on the fake rate does not exceed 21%. Values of fake rate from Table 3 are used to scale the e-probe regions with corresponding η and p_{T} of chosen electron.

4.1.2 Resulting background event yield

For estimation of $e \to \gamma$ background the following e-probe CR is constructed in the data: all the selection criteria are taken either from SR phase space, in which the *probe* electron is selected instead of photon. The e-probe CR is based on SR selection and will be used for cross-section measurement in cut-based analysis. Selection criteria for e-probe CR are summarized in Table 4. For $W\gamma$ CR the corresponding e-probe CR is formed in a similar way with some distinctions: $(N_\mu + N_e) \geq 2$ (where leptons are selected as described in Sec. 3.3) is required there is no restriction on the number of electrons passing e-probe selection in the event if there more of them than 0.

The shapes for this background in SR regionand Wgamma CR are taken from the corresponding e-probe CR in data (respectively e-probe CR and e-probe $W\gamma$ CR).

Selections	Cut Value
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 130 GeV
$E_{\mathrm{T}}^{e-probe}$	> 150 GeV
Number of loose non-isolated photons	$N_\gamma = 0$
Number of tight probe electrons	$N_{e-probe} = 1$
Number of jets	$N_{\mathrm{jets}} \geq 2$
Lepton veto	$N_\mu = 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$ significance	> 11
$\lvert\Delta\phi(e - probe, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})\rvert$	> 0.7
$\lvert\Delta\phi(j_1, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})\rvert$	> 0.4

Table 4: Event selection criteria for e-probe CR events.

Kinematic distributions in e-probe CR can be found in Figure 3. In these figures real $e + E_{\mathrm{T}}^{\mathrm{miss}}$ includes modeling of all non-negligible processes with real electrons and real $E_{\mathrm{T}}^{\mathrm{miss}}$, which includes $W \to e\nu$, $W \to \tau\nu$ (with leptonic decay), single-top, $t\bar{t}$, $t\bar{t}\gamma$ and $W\gamma$ processes. Fake $e + E_{\mathrm{T}}^{\mathrm{miss}}$ consists of $Z(\to \nu\bar{\nu})\gamma$, inclusive $Z(\to \nu\bar{\nu})$ and multijet modeling.

Event yield	real $e + E_{\mathrm{T}}^{\mathrm{miss}}$ (MC)	fake $e + E_{\mathrm{T}}^{\mathrm{miss}}$ (MC)	data
e-probe CR	85100 ± 4408	739 ± 41	86497

Table 5: Event yields for real $e + E_{\mathrm{T}}^{\mathrm{miss}}$ and fake $e + E_{\mathrm{T}}^{\mathrm{miss}}$ prediction and observed data in probe-electron control regions. Indicated uncertainties are statistical.

Shape disagreement between prediction and data was checked to be caused by insufficient statistics in $W(\rightarrow e\nu)$ MC even after combination of Sherpa 2.2.1 samples with existing Sherpa 2.2.7 (see App. G of [26]). Also $jet \rightarrow e$ misidentification prediction from MC was checked to be in agreement with data-driven estimation (see App. G of [26]).

Event yields for predicted real and fake $e + E_{\mathrm{T}}^{\mathrm{miss}}$ processes together with observed data are shown in Table 5. A purity of the control region is determined as a portion of data, which is not contaminated by fake $e + E_{\mathrm{T}}^{\mathrm{miss}}$ events. It counts $99.15 \pm 0.05\%$ in e-probe CR. The impurity of e-probe CR $0.85 \pm 0.05\%$ is considered as additional systematic uncertainty.

The statistical uncertainty on fake-rate value in every region is also taken into account as additional source of systematic uncertainty in corresponding region of η and p_{T}.

Then the measured event yield from this CR is scaled on the corresponding fake rate for different p_T and η ranges. The total number of events for this background equals to $3039 \pm 12 \pm 209$ in SR phase space. First uncertainty is statistical, second is systematical. Total systematics for this background does not exceed 6.8%. See systematics breakdown in Table 23 of App. B.3.

4.2 Background induced by $jet \rightarrow \gamma$ misidentification

Most of $jet \rightarrow \gamma$ misidentification comes from the $Z(\nu\bar{\nu}) + jets$ and multi-jet processes and $W(\tau\nu)$ process, where τ decays into hadrons. As well as the other kinds of misidentification it can not be properly modeled with the Monte Carlo. Therefore, two-dimensional sideband method [27] is used for the estimation of $jet \rightarrow \gamma$ background. Photon isolation and identification based on shower shape variables are taken as two discriminating variables. They are used as a basis of three control region and a signal region as shown in Figure 4.

- Tight and isolated region (region A – equivalent to $Z\gamma$ signal region described in Sec. 3.6): events have a leading photon candidate that is isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} < 0$ GeV) and passes the *tight* selection.
- Tight but not isolated region (control region B): events have a leading photon candidate that is not isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} >$ iso gap) and passes the *tight* selection.
- Non-tight and isolated region (control region C): events have a leading photon candidate that is isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} < 0$ GeV) and passes the *non-tight* selection.
- Non-tight and not isolated region (control region D): events have a leading photon candidate that is not isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} >$ iso gap) and passes the *non-tight* selection.

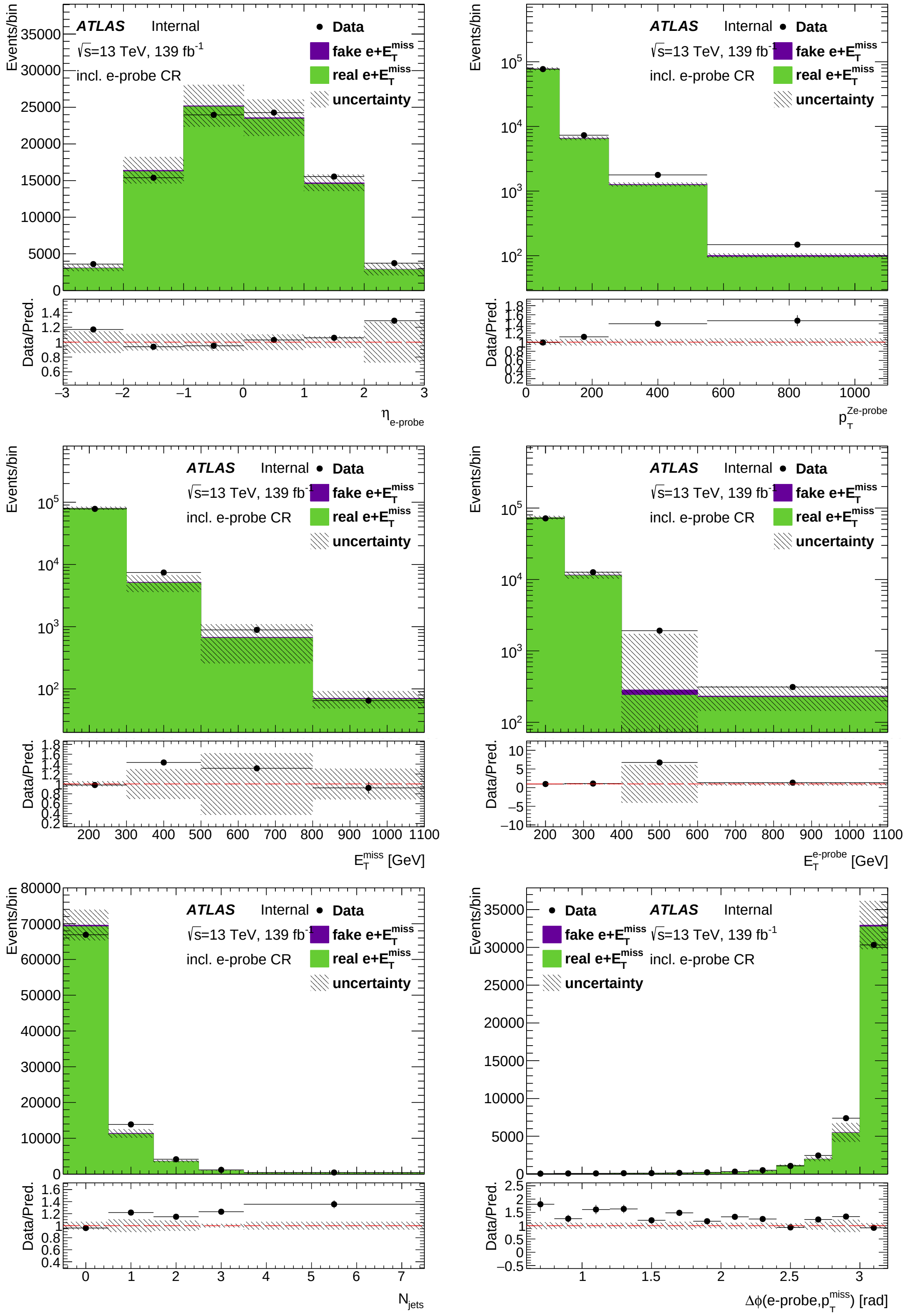


Figure 3: Kinematic distributions in e-probe control region. The dashed band represents the sum in quadrature of all statistical uncertainties for background and signal expectations.

Photon is required to pass $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} < 0.05$ track isolation in A and C isolated regions, and inverted track isolation ($p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} > 0.05$) in B and D non-isolated regions to have less correlated results. The procedure of the selection of isolation working point is described in Appendix F. An isolation gap of 2 GeV between isolated and non-isolated regions is used to reduce leakage of signal $Z(\nu\bar{\nu})\gamma$ process from signal A region to other control regions. The choice of this value is further described in Appendix G.

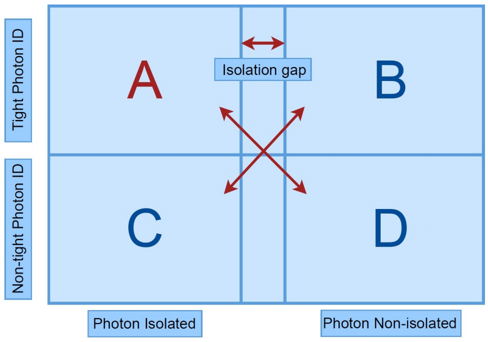


Figure 4: Schematic illustration of the two-dimensional plane based on photon isolation and identification variables with separation on A signal region and B, C and D control regions.

Non-tight photon passes not all shower shape requirements, which are defined for *tight*. There are few types of *non-tight* (*loose'*) working points available, in which at least one of the following EM shower shape criteria is not satisfied:

- *loose'2*: w_{s3}, F_{side}
- *loose'3*: w_{s3}, F_{side}, ΔE
- *loose'4*: w_{s3}, F_{side}, ΔE, E_{ratio}
- *loose'5*: w_{s3}, F_{side}, ΔE, E_{ratio}, w_{tot},

where w_{s3} is the shower width calculated from three strips around the strip with a maximum energy deposit, F_{side} is the energy outside the core of the three central strips but within seven strips divided by energy within the three central strips, ΔE is the difference between the energy associated with the second maximum energy deposit in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and the second maxima, E_{ratio} is the ratio of the energy difference associated with the largest and second-largest energy deposits to the sum of these energies and w_{tot} is the total lateral shower width. All these relaxed cuts are dedicated to strip layer variables.

In case of no correlation between the control regions a following equation should be valid: $\frac{N_{\mathrm{A}}^{\mathrm{jet}\to\gamma}}{N_{\mathrm{B}}} = \frac{N_{\mathrm{C}}}{N_{\mathrm{D}}}$, where $N_{\mathrm{A}}^{\mathrm{jet}\to\gamma}$ is a number of $jet \to \gamma$ events in A region, and N_{i} are the numbers of events in corresponding B, C and D regions. A basic assumption for this method to work is to have no correlation between *non-tight* working point and the isolation, therefore the least correlated *loose'* should be chosen. For this purpose, R correlation factor is estimated in the MC $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay as $R = \frac{N_{\mathrm{A}}N_{\mathrm{D}}}{N_{\mathrm{B}}N_{\mathrm{C}}}$, which should be equal to 1 in case of correlation absence. It should be noted that multi-jet MC is not used for R correlation factor calculation, since this dataset has extremely limited statistics, which causes a problem with normalization.

R factor in data can be estimated via extended ABCD method with the additional E and F regions as described in Appendix H. The resulting values of R correlation factor in data and MC for different *loose'* working points can be found in Table 6. In order to reduce systematic uncertainty the *loose'3* working point was chosen.

R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.11 ± 0.13	1.23 ± 0.12	1.34 ± 0.12	1.60 ± 0.13
Data-driven	0.97 ± 0.10	1.05 ± 0.10	1.05 ± 0.09	1.06 ± 0.08

Table 6: Estimated correlation factor R between isolation and different *loose'* photon identification working points in MC and data. Indicated uncertainties are statistical.

The two-dimensional sideband method is based on the assumption that signal A region is mainly consists of signal events, while three control regions (B, C and D) are mainly consist of background events, and the leakage of signal events into the control regions is well estimated by MC simulations. The number of events arising in each of the regions is estimated as shown below:

$$\begin{aligned} N_{\mathrm{A}} &= N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma} + N_{\mathrm{A}}^{\mathrm{bkg}} + N_{\mathrm{A}}^{\mathrm{jet}\to\gamma}; \\ N_{\mathrm{B}} &= c_{\mathrm{B}} N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma} + N_{\mathrm{B}}^{\mathrm{bkg}} + N_{\mathrm{B}}^{\mathrm{jet}\to\gamma}; \\ N_{\mathrm{C}} &= c_{\mathrm{C}} N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma} + N_{\mathrm{C}}^{\mathrm{bkg}} + N_{\mathrm{C}}^{\mathrm{jet}\to\gamma}; \\ N_{\mathrm{D}} &= c_{\mathrm{D}} N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma} + N_{\mathrm{D}}^{\mathrm{bkg}} + N_{\mathrm{D}}^{\mathrm{jet}\to\gamma}; \end{aligned}$$

where $N_{\mathrm{i}}^{\mathrm{bkg}}$ are the numbers of background events other than $jet \to \gamma$ that contributing in each region. They are estimated using corresponding MC samples or data-driven techniques. The signal leakage parameters c_{i} are derived from MC signal samples and defined as the ratio of the signal events in region B/C/D to the number of the signal events in region A:

$$\begin{aligned} c_{\mathrm{B}} &= \frac{N_{\mathrm{B}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}{N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}; \\ c_{\mathrm{C}} &= \frac{N_{\mathrm{C}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}{N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}; \\ c_{\mathrm{D}} &= \frac{N_{\mathrm{D}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}{N_{\mathrm{A}}^{\mathrm{Z}(\nu\bar{\nu})\gamma}}. \end{aligned}$$

After calculating the signal leakage parameters, removing the $N_{\mathrm{i}}^{\mathrm{bkg}}$ background contributions from each region and obtaining $\tilde{N}_{\mathrm{i}} = N_{\mathrm{i}} - N_{\mathrm{i}}^{\mathrm{bkg}}$ values and considering R factor in data, one can solve the equations above to find the following formula for signal yield:

$$N_{\mathrm{A}}^{Z(\nu\bar{\nu})\gamma} = \tilde{N}_{\mathrm{A}} - R(\tilde{N}_{\mathrm{B}} - c_{\mathrm{B}} N_{\mathrm{A}}^{Z(\nu\bar{\nu})\gamma}) \frac{\tilde{N}_{\mathrm{C}} - c_{\mathrm{C}} N_{\mathrm{A}}^{Z(\nu\bar{\nu})\gamma}}{\tilde{N}_{\mathrm{D}} - c_{\mathrm{D}} N_{\mathrm{A}}^{Z(\nu\bar{\nu})\gamma}}.$$

Which yields the following solution:

$$N_{\mathrm{A}}^{Z(\nu\bar{\nu})\gamma} = \frac{b - \sqrt{b^2 - 4ac}}{2a},$$

where a, b, and c have the following values:

$$a = c_{\mathrm{D}} - R c_{\mathrm{B}} c_{\mathrm{C}};$$
$$b = \tilde{N}_{\mathrm{D}} + c_{\mathrm{D}} \tilde{N}_{\mathrm{A}} - R(c_{\mathrm{B}} \tilde{N}_{\mathrm{C}} + c_{\mathrm{C}} \tilde{N}_{\mathrm{B}});$$
$$c = \tilde{N}_{\mathrm{D}} \tilde{N}_{\mathrm{A}} - R \tilde{N}_{\mathrm{C}} \tilde{N}_{\mathrm{B}}.$$

The resulting ABCD data yields and non-$jet \to \gamma$ background yields in all considered regions are reported in Table 7. The numbers of lepton $W(\tau\nu)$ decay, top, tt background events are derived from the data-driven technique for $e \to \gamma$ background described in link to section. The numbers of γ+jet background events are derived from MC, however, the numbers obtained using the data-driven technique described in link to section were checked and this change leads to negligible impact on $jet \to \gamma$ background estimation results.

	Data	$W\gamma$ QCD	$W\gamma$ EWK	$e \to \gamma$	$tt\gamma$	γ+jet	$Z(ll)\gamma$
A	26523 ± 163	3936 ± 23	136.3 ± 0.7	3039 ± 12	234 ± 3	5262 ± 53	285 ± 5
B	1475 ± 38	52 ± 4	1.86 ± 0.08	8.95 ± 0.03	1.3 ± 0.2	0.6 ± 0.4	1.0 ± 0.6
C	2568 ± 51	60 ± 2	2.16 ± 0.09	61.4 ± 0.2	4.2 ± 0.4	0.76 ± 6	4.8 ± 0.5
D	1443 ± 38	2.7 ± 0.6	0.17 ± 0.02	0.0715 ± 0.0002	0.35 ± 0.13	0 ± 0	0 ± 0

Table 7: Event yields for the data and non-$jet \to \gamma$ background processes considered in the ABCD method. Indicated uncertainties are statistical.

The central value of $jet \to \gamma$ background in signal A region is $N_{\mathrm{A}}^{\mathrm{jet}\to\gamma} = 2078$.

4.2.1 Uncertainties and differential distributions

To assess the statistical uncertainty, the event yields of four regions in data and non-$jet \to \gamma$ backgrounds are independently varied by $\pm 1\sigma$ (Table 7). The deviations from the nominal value are added up in quadrature. The statistical uncertainty of the signal leakage parameters is negligible and it is not used in calculation of the resulting uncertainty. The central value of $jet \to \gamma$ background and its statistical uncertainty is 2078^{+100}_{-97}.

A systematic uncertainty is estimated from variations of ABCD regions determination, i.e. non-tight definition variation and isolation gap variation until $\pm 1\sigma$ changes in data yield. Resulting deviations, as

Central value	2078^{+100}_{-97}
Loose'2	+327
Loose'4	−111
Loose'5	−173
Isolation gap +0.25 GeV	+48
Isolation gap −0.35 GeV	+29

Table 8: Central value of $jet \to \gamma$ background from data-driven estimation and deviations from variations of ABCD regions definition.

well as central value, can be found in Table 8. The largest deviation is 16% and it is taken as the systematic uncertainty.

Uncertainty coming from signal leakage parameters can be estimated using $Z\gamma$ QCD samples generated with two different generators: `Sherpa 2.2` and an alternative `MadGraph+Pythia8` (Table Table 9). The largest deviation is 0.8%

Signal leakage parameters	`MadGraph+Pythia8`, `Sherpa 2.2`	`MadGraph+Pythia8`, `MadGraph+Pythia8`	Relative deviation
c_{B}	0.00939 ± 0.00007	0.0155 ± 0.0004	39%
c_{C}	0.01536 ± 0.00010	0.0156 ± 0.0007	1.5%
c_{D}	0.00051 ± 0.00002	0.00077 ± 0.00009	34%
$jet \to \gamma$ estimation	2078^{+100}_{-97}	2061^{+100}_{-97}	0.8%

Table 9: Signal leakage with central value of $jet \to \gamma$ background in B, C and D control regions with their relative deviations for `Sherpa 2.2` and `MadGraph+Pythia8` $Z\gamma$ QCD samples.

Differences between MC exist also due to an imperfect photon iso/ID modeling. Therefore, systematics on signal leakage parameters can be derived from the iso/ID uncertainty on reconstruction photon efficiency $\delta^{\mathrm{eff}}_{\mathrm{iso/ID}}$ (relative). By definition, the photon isolation modeling only affects c_{B} and c_{D}, while the photon ID modeling only has effects on c_{C} and c_{D}, which gives the following σ for leakage parameters:

- $\sigma^{c_{\mathrm{B}}}_{\mathrm{iso}}(\text{relative}) = \delta^{\mathrm{eff}}_{\mathrm{iso}} * (c_{\mathrm{B}} + 1)/c_{\mathrm{B}}$
- $\sigma^{c_{\mathrm{C}}}_{\mathrm{ID}}(\text{relative}) = \delta^{\mathrm{eff}}_{\mathrm{ID}} * (c_{\mathrm{C}} + 1)/c_{\mathrm{C}}$
- $\sigma^{c_{\mathrm{D}}}_{\mathrm{iso}}(\text{relative}) = \delta^{\mathrm{eff}}_{\mathrm{iso}} * (c_{\mathrm{B}} + 1)/c_{\mathrm{B}}$
- $\sigma^{c_{\mathrm{D}}}_{\mathrm{ID}}(\text{relative}) = \delta^{\mathrm{eff}}_{\mathrm{ID}} * (c_{\mathrm{C}} + 1)/c_{\mathrm{C}}$

The largest deviation from this type of uncertainty is 1.3%. Thus the total systematics on $jet \to \gamma$ is found to be equal to 16%. The resulting number of $jet \to \gamma$ events in signal A region is $2078^{+100}_{-97} \pm 332$.

4.3 Slice method for $jet \to \gamma$ background estimation

The phase-space of $Z\gamma$ associated production is defined by the cuts, that are presented in section: link to section

Throughout this study, FixedCutLoose photon isolation working point is used as the isolation with maximum significance. Photon is required to pass $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} < 0.05$ track isolation in isolated regions.

To increase the statistics in non-isolated regions the inverted track isolation $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} > 0.05$ is applied. The background estimation method of $jet \to \gamma$ events relies on maximum likelihood fit to signal and background processes apart from $jet \to \gamma$. The fit can be performed for different variables, such as $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \gamma)|$, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_1)|$, in the phase-space region with relaxed cuts on these variables. The corresponding $jet \to \gamma$ background distributions are obtained from data applying the same kinematic selections. As a final step, the number of $jet \to \gamma$ events is obtained from the fits in the region with relaxed cuts and are extrapolated back to the signal region. The results are cross-checked by comparing the estimate to the two-dimentional sideband method (ABCD-method).

There are fit region, two control regions and a signal region defined by different kinematic cuts and isolation criteria as shown in Figure 5:

- Signal region (SR): events have a leading photon candidate that is isolated ($E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} < 0.065$) and passes signal kinematic selections.
- Fit region (FR): events have a leading photon candidate that is isolated ($E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} < 0.065$) and meets relaxed signal requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \gamma)|$, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_1)|$.
- Control region 1 (CR1): events have a leading photon candidate that is not isolated ($E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} > 0.065$) and meets relaxed signal requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \gamma)|$, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_1)|$.
- Control region 2 (CR2): events have a leading photon candidate that is not isolated ($E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} > 0.065$) and passes signal kinematic selections.

Photons in all four regions pass the *tight* selection. With these definitions, the SR is a subset of the FR, and CR2 is a subset of CR1.

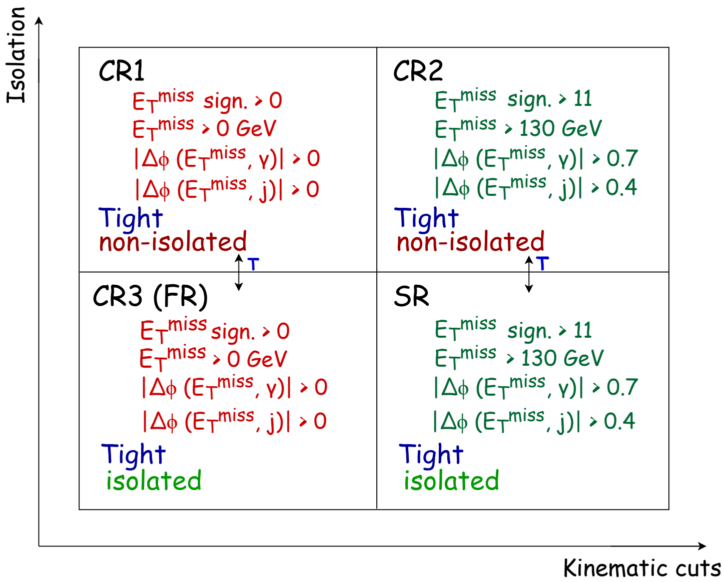


Figure 5: The definition of 4 regions used in $jet \to \gamma$ background estimation based on kinematic cuts and photon isolation requirements.

The main goal of the procedure is to estimate the number of $jet \to \gamma$ events in SR. The fit is performed in the FR, where the $jet \to \gamma$ process used for the fit is derived from CR1. The relaxed requirements in FR and CR1 are applied to dispose of enough statistics.

To study the dependence of the result on the isolation criterion, control regions CR1 and CR2 are split into successive intervals of the isolation variable, instead of a single, integrated anti-isolated region. Correspondingly, the isolation slices used for the $jet \to \gamma$ estimation can be defined as follows: [0.065, 0.080, 0.095, 0.115, 0.140].

In this way, the number of $jet \to \gamma$ background events for a given isolation slice i can be estimated as follows:

- The total number of $jet \to \gamma$ background events in each non-isolated slice (i) of CR1 ($N^{jet\to\gamma}_{\mathrm{CR1(i)}}$) is derived as follows:

$$N^{jet\to\gamma}_{\mathrm{CR1(i)}} = N^{\mathrm{data}}_{\mathrm{CR1(i)}} - N^{\mathrm{Z}(\nu\bar{\nu})\gamma}_{\mathrm{CR1(i)}} - N^{\mathrm{bkg}}_{\mathrm{CR1(i)}}$$

  where $N_{\mathrm{CR1(i)}}$ is obtained from any of the kinematic distributions used for the fit.

- The fit is performed in FR, characterized by relaxed kinematics cuts and signal isolation requirements. Thus, the total number of events in FR estimated from non-isolated slice of CR1 is given by:

$$N^{\mathrm{data}}_{\mathrm{FR(i)}} = \alpha \cdot (N^{\mathrm{Z}(\nu\bar{\nu})\gamma}_{\mathrm{FR(i)}} + N^{\mathrm{bkg}}_{\mathrm{FR(i)}}) + N^{jet\to\gamma}_{\mathrm{FR(i)}}.$$

  The fitting parameter $T_{(i)}$ gives the estimated number of $jet \to \gamma$ events in FR: $N^{jet\to\gamma}_{\mathrm{FR(i)}} \approx T_{(i)} \cdot N^{jet\to\gamma}_{\mathrm{CR1(i)}}$. An overall normalization factor α should be equal to 1 within the uncertainties. In this study, a parameter α is taken to be equal 1. The fit parameter $T_{(i)}$ is derived for each slice and kinematic variable.

- Finally, the fitted $jet \to \gamma$ yield is extrapolated to the signal region. The estimate for each slice and kinematic variable is determined by the equation: $N^{jet\to\gamma}_{\mathrm{SR(i)}} = T_{(i)} \cdot (N^{\mathrm{data}}_{\mathrm{CR2(i)}}$ - $N^{\mathrm{Z}(\nu\bar{\nu})\gamma}_{\mathrm{CR2(i)}}$ - $N^{\mathrm{bkg}}_{\mathrm{CR2(i)}})$.

4.4 Resulting background yields

A summary of estimated and observed event yields in all considered regions before the fit procedure described in Section 5 is presented in Table 10. Event yields after background only fit and fit to the observed data are presented in Tables 12 and 14 respectively.

Results for $e \to \gamma$, $j \to E_T^{miss}$ and $j \to \gamma$ backgrounds are obtained from data-driven estimated described above.

	$W\gamma$ CR	Signal region
$Z(\nu\bar{\nu})\gamma$ QCD	532.2 ± 1.9	$10\,583 \pm 8$
$Z(\nu\bar{\nu})\gamma$ EWK	11.59 ± 0.10	153.1 ± 0.8
$W\gamma$ QCD	4600 ± 30	3960 ± 30
$W\gamma$ EWK	242.6 ± 1.0	137.0 ± 1.0
$e \rightarrow \gamma$	334 ± 4	2874 ± 12
$j \rightarrow E_T^{miss}$	619 ± 18	5330 ± 60
$j \rightarrow \gamma$	100 ± 20	2070 ± 100
$Z(\ell\bar{\ell})\gamma$	280 ± 5	287 ± 5
$t\bar{t}\gamma$	691 ± 6	235 ± 3
Total	7410 ± 50	$25\,630 \pm 120$
Data	7892	

Table 10: Event yields for the signal and all of the background processes considered in this analysis the fit procedure described in Section 5. The yields are presented in the regions described in Section 3.6.2. The uncertainty of the expected yield consists of statistical and systematic uncertainties.

5 Template fit and results

To extract the $Z(\nu\bar{\nu})\gamma$ cross sections, a binned maximum-likelihood simultaneous fit [28] is performed using the photon transverse energy distribution in the SR and the $W\gamma$ CR. Two free parameters are introduced in the combined fit: a signal strength parameter, $\mu_{Z\gamma}$, used for the signal process and a normalisation factor $\mu_{W\gamma}$ used to scale the yields of the $W(\ell\nu)\gamma$ and $t\bar{t}\gamma$ processes due to their similar final states. Details of the likelihood fit procedure and its results are described in the following sections.

5.1 Likelihood function

The fit is performed with the `TRExFitter` [29] framework based on the `HistFactory` [30] package from `ROOT`.

The cross-section measurement is performed by estimating the *signal strength* also referred to as *parameter of interest* (POI), defined as

$$\mu_{Z\gamma} = \mu = \frac{\nu^s_{\text{meas.}}}{\nu^s_{\text{SM}}} = \frac{\sigma^s_{\text{fid, meas.}}}{\sigma^s_{\text{fid, SM}}}, \tag{3}$$

where ν^s_{SM} is the number of signal $Z\gamma$ events predicted by the SM and $\nu^s_{\text{meas.}}$ is the number of signal events measured from the observed data.

To account for the systematic uncertainties and background normalisation constraints a set of *nuisance parameters* (NPs) θ should also be included into the likelihood model.

The binned likelihood function used in this analysis is

$$\mathcal{L}(\mu,\theta) = \prod_r^{\text{regions}} \left[\prod_i^{\text{bins}\in r} \text{Pois}(N_i^{\text{data}} | \mu\nu_i^s \eta^s(\theta) + \nu_i^b \eta^b(\theta)) \right] \cdot \prod_i^{\text{nuis. par.}} \mathcal{L}(\theta_i), \tag{4}$$

where

- N_i^{data} is the number of the observed data events in the bin;
- ν_i is the expected number of the signal or background events in the bin ($\nu_i^s = \nu^s_{\text{SM}}$ from Eq. 3);
- $\eta(\theta)$ reflects the impact systematic uncertainties and normalisation constraints have on the number of the events in the bin through the set of NPs θ;
- $\mathcal{L}(\theta_i)$ is the likelihood function of the "subsidiary measurement" that reflects the nature of the systematic uncertainties (for normalisation constraints such measurement is performed in a dedicated CR region and included in the first product over the regions).

Measurement of the POI μ and NPs θ is performed by minimising "$-\ln\mathcal{L}(\mu,\theta)$". The corresponding values are denoted as $\hat{\mu}$ and $\hat{\theta}$, respectively.

The following statistic is used to compute the discovery significance and the uncertainties of $\hat{\mu}$ and $\hat{\theta}$ estimates:

$$q(\mu,\hat{\mu},\hat{\theta}) = -2\ln\lambda(\mu,\hat{\mu},\hat{\theta}) = -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\theta})}, \tag{5}$$

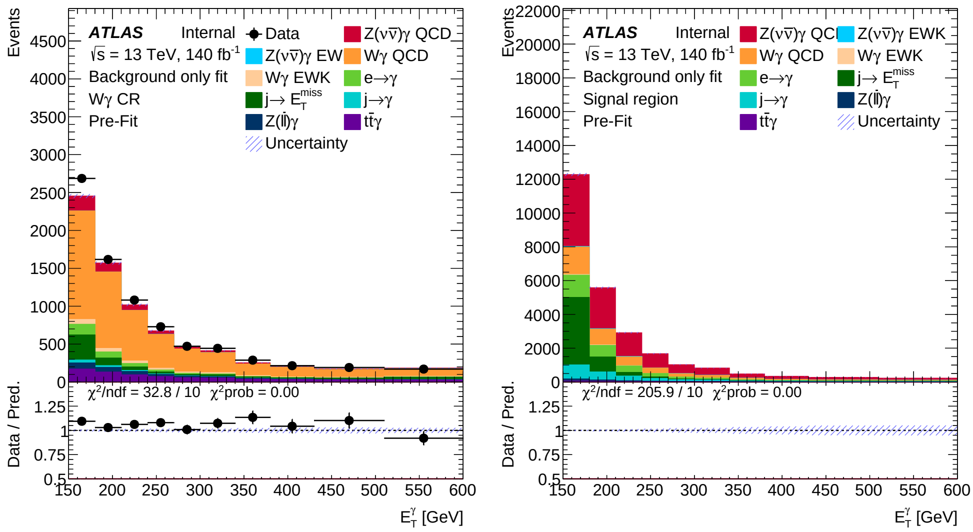


Figure 6: Pre background-only fit distributions. Expected distribution is compared to observed data in the $W\gamma$ CR. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

where $\lambda(\mu, \hat{\mu}, \hat{\theta})$ is known as *profile likelihood ratio* and $\hat{\hat{\theta}}(\mu)$ is the set of NP values, that minimises $-\ln \mathcal{L}(\mu, \theta)$ for any given μ.

According to Ref. [31], the discovery significance of the measurement and the expected median discovery significance can be calculated with

$$Z^{\text{meas.}}_{\text{disc}} = \sqrt{q(\mu = 0)}, \tag{6}$$

$$Z^{\text{exp.}}_{\text{disc}} = \sqrt{q(\mu = 1)_A}, \tag{7}$$

where $q(\mu = 1)_A$ is calculated using the Asimov dataset (defined in Section 5.2) instead of the observed data.

The 1, 2 and 3σ uncertainties for $\hat{\mu}$ estimate could be computed by finding such values of μ that $q(\mu, \hat{\mu}, \hat{\theta})$ differs from its minimum value by 1, 4 and 9 respectively.

5.2 Fit procedure

The signal and control regions are defined in Section 3.6.2. The photon transverse energy distribution template is used in the SR and the $W\gamma$ CR. Those templates are shown in Figure 6.

The shape of the signal $Z\gamma$ process is taken from the MC predictions. The shapes of $Z\gamma$, $W\gamma$ and $tt\gamma$ processes are also taken from the MC predictions and their normalisation is evaluated by assigning them normalisation coefficient $\mu_{W\gamma}$. The Table 11 summarises the use of the normalisation coefficients.

Systematic uncertainties for those processes are discussed in X. $e \to \gamma$, $j \to E_{\text{T}}^{\text{miss}}$, and $j \to \gamma$ processes distributions are obtained with data-driven techniques as described in Section 4. Shape and normalisation of $Z(\ell\ell) + \gamma$ process are taken from the MC predictions.

	SR	$W\gamma$ CR
$\mu_{Z\gamma}$	✓	
$\mu_{W\gamma}$	✓	✓

Table 11: Table of regions where the normalisation coefficients are used to calculate the likelihood function.

Since the extraction of the expected results requires the SR to be blinded this analysis uses different procedures to obtain expected and observed results.

To obtain the expected results a two-step fit procedure is used. At the first step *background-only* fit to the observed data is performed without considering the SR to get an estimate of the background normalisation coefficient and nuisance parameters.

At the second step both CR and SR are considered and $\mu_{Z\gamma}$ is used as a fit parameter. Since the signal region is blinded during the extraction of the expected results, so-called *Asimov dataset* is used instead of the observed data. An Asimov dataset is such dataset that when one uses it to evaluate the estimators for all parameters, one obtains the true parameter values. It is created by summing all of the expected event yields and taking into account the effect of the background normalisation coefficient and systematic NPs as estimated in the background-only fit [31]. Based on the definition of the Asimov dataset, this fit will yield the same results for the background normalisation coefficient and systematic uncertainties NPs mean values and errors as the background-only fit. However, this fit allows to also estimate the significance and the uncertainty of the POI.

The observed results are obtained in one step. The observed data is used in both CR and SR and a simultaneous fit of all of the parameters with $\mu_{Z\gamma}$ as a POI is done.

5.3 Fit results

5.3.1 Background only fit

Figure 7 shows the distributions after the first fit step — background-only fit. Figure 8 presents the summary of the processes event yields for all of the regions in the background-only fit. Tables 10 and 12 show the event yields before and after the first fit. The result for the background normalisation coefficient is $\mu_{W\gamma} = 1.08 \pm 0.02(\text{stat})$.

Values and errors of the corresponding NPs for background-only fit are presented in Figure 9. Figure 10 shows the correlation matrix for NPs with the highest value of correlation coefficients for the background-only fit.

5.3.2 Fit to Asimov data

Results of the fit to Asimov data are presented in Table 13. The values of the background normalisation coefficients are the same as for the background-only fit by the definition of the Asimov dataset. The $\mu_{Z\gamma\text{EWK}} = 1$ since it was not affected by the background-only fit.

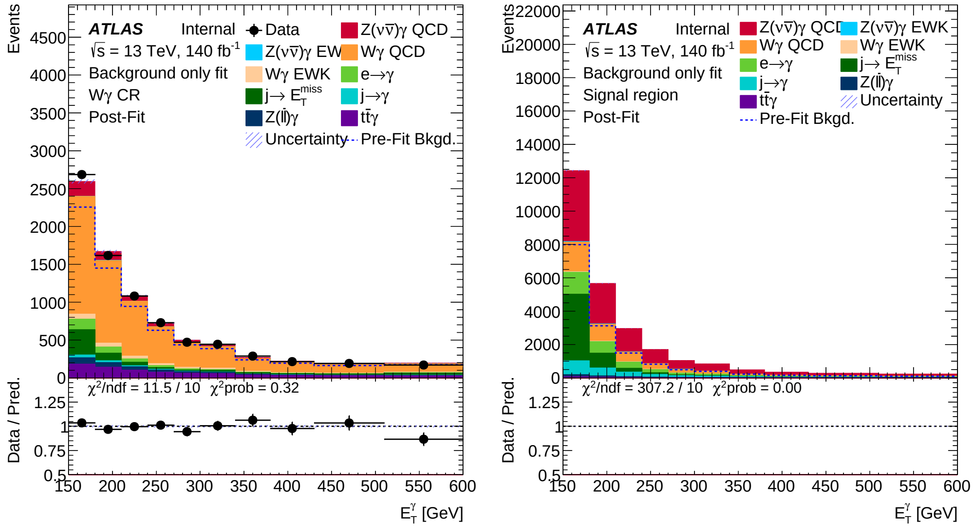


Figure 7: Post background-only fit distributions. Expected distribution are compared to observed data in the $W\gamma$ CR. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

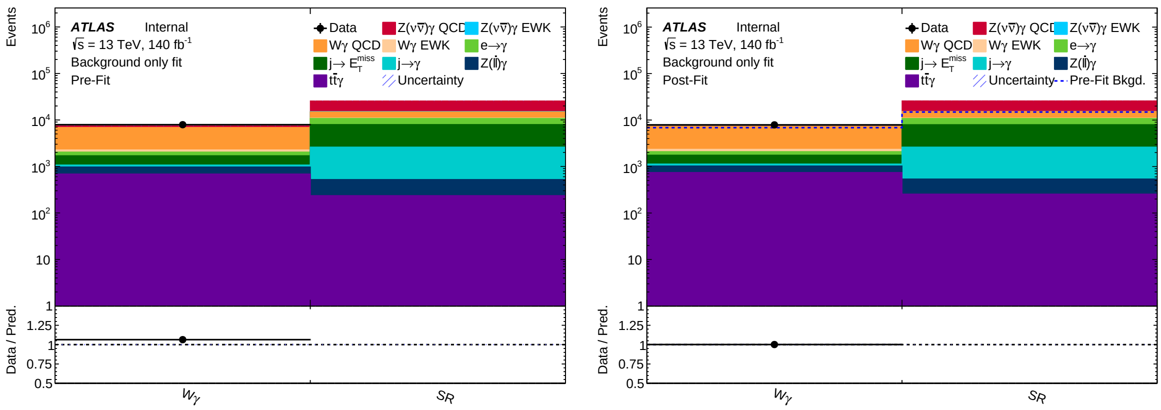


Figure 8: Summary of the pre and post background-only fit processes yield for all of the regions. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

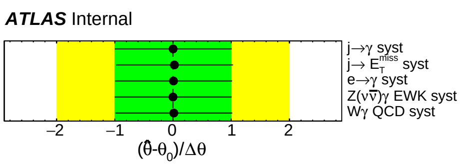


Figure 9: Nuisance parameters value and error (pull plot) after background only fit.

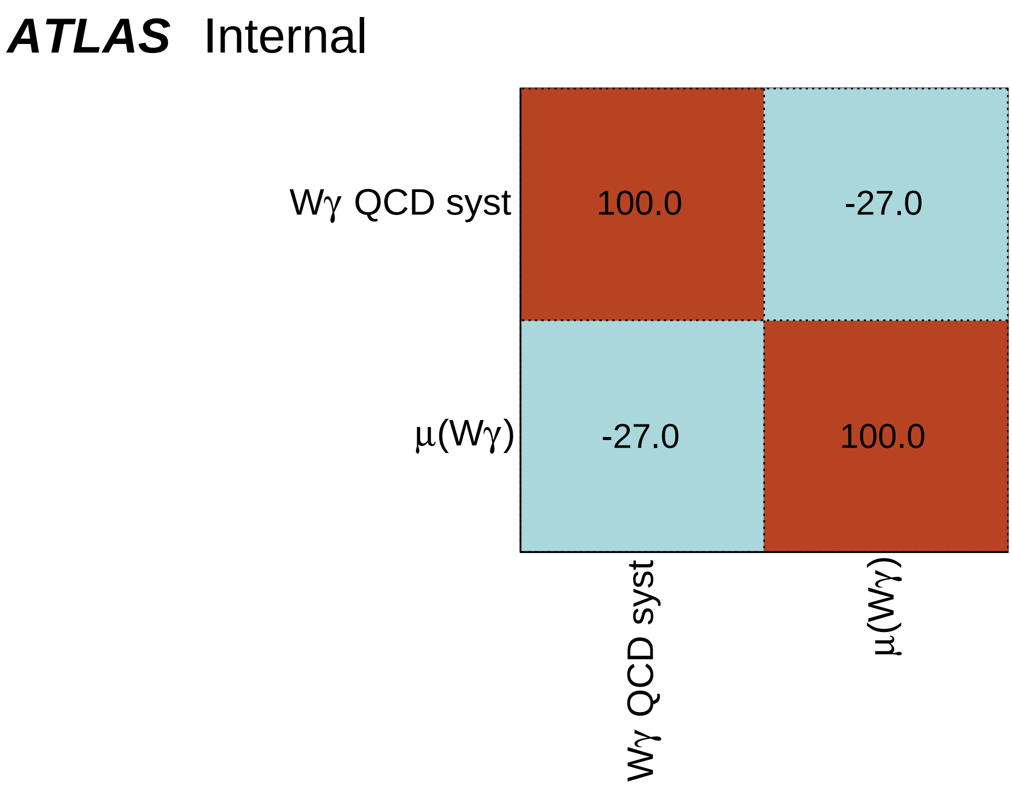


Figure 10: Correlation matrix of the background only fit parameters. Only parameters with the highest values of the correlation coefficient are shown.

	$W\gamma$ CR	Signal region
$Z(\nu\bar{\nu})\gamma$ QCD	532 ± 0	10 583 ± 0
$Z(\nu\bar{\nu})\gamma$ EWK	11.59 ± 0.06	153.1 ± 0.8
$W\gamma$ QCD	4990 ± 70	4300 ± 60
$W\gamma$ EWK	263 ± 4	149 ± 2
$e \to \gamma$	334.4 ± 1.7	2874 ± 0
$j \to E_T^{miss}$	619 ± 3	5330 ± 30
$j \to \gamma$	104.1 ± 0.5	2070 ± 10
$Z(\ell\bar{\ell})\gamma$	280 ± 0	287 ± 0
$t\bar{t}\gamma$	749 ± 11	255 ± 4
Total	7880 ± 90	26 000 ± 70
Data	7892	

Table 12: Event yields for the signal and all of the background processes considered in this analysis after the background only fit. The yields are presented in the regions described in Section 3.6.2. The uncertainty of the expected yield consists of statistical and systematic uncertainties.

$\mu_{Z\gamma}$	1.00 ± 0.02(stat)
$\mu_{W\gamma}$	1.08 ± 0.02(stat)
Expected significance, σ	70

Table 13: Values of normalisation coefficients and the expected median significance of the cross-section measurement after the fit to Asimov data

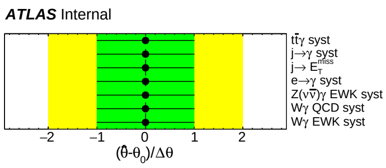


Figure 11: Nuisance parameters value and error (pull plot) after Asimov data fit.

Values and errors of the corresponding NPs for Asimov data fit are presented in Figure 11. Figure 12 shows the correlation matrix for NPs with the highest value of correlation coefficients for the Asimov data fit.

Figure 13 shows the estimations of the pre-fit and post-fit impact of the systematic uncertainties on the total systematic uncertainty of the POI in Asimov data fit.

5.3.3 Observed results

Figures 15 and 14 show the distributions for the fit to the observed data in both $W\gamma$ CR and SR before and after the fit, respectively. Figure 16 presents the summary of the processes event yields for all of the regions in the fit to the observed data in both $W\gamma$ CR and SR. Table 14 shows the event yields after the fit. Normalisation coefficients obtained in the fit to the observed data in both $W\gamma$ CR and SR are presented in Table 15.

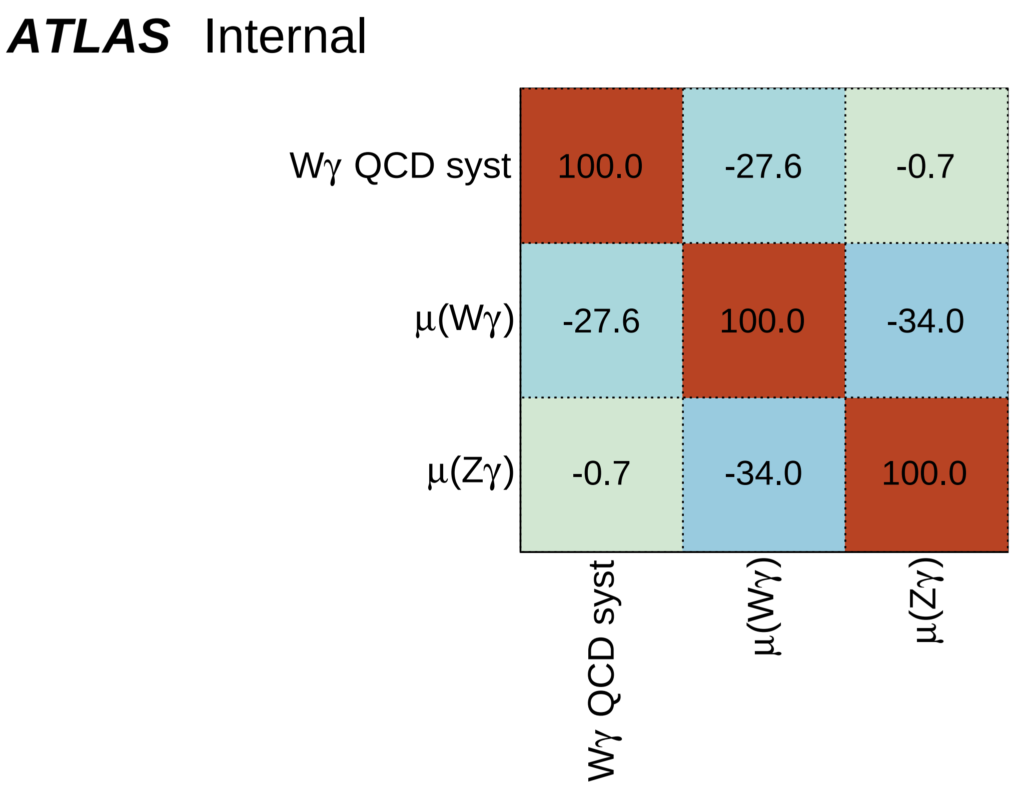


Figure 12: Correlation matrix of the Asimov data fit parameters. Only parameters with the highest values of the correlation coefficient are shown.

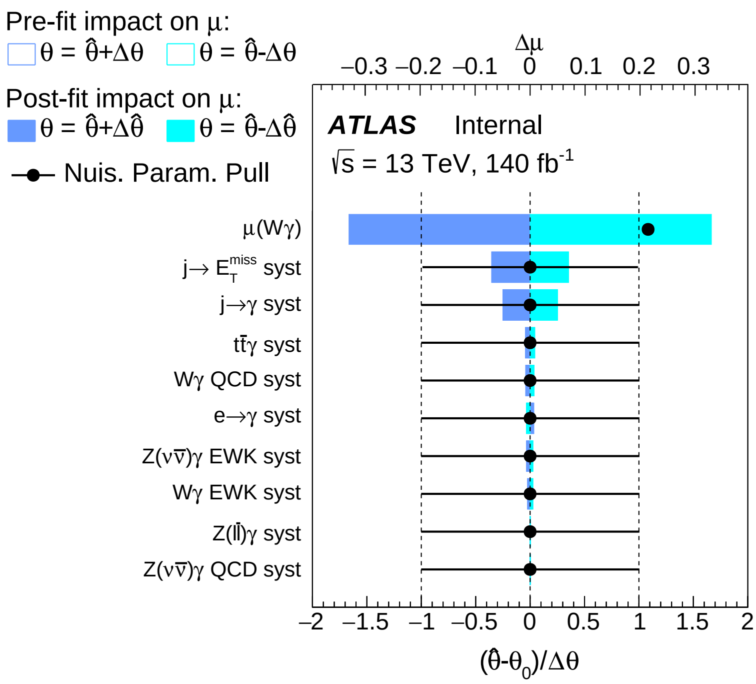


Figure 13: Estimations of the pre-fit (empty blue bars) and post-fit (filled blue bars) impact of the systematic uncertainties on the total systematic uncertainty of the POI (measured with the scale on the top part of the chart) and the values of the corresponding NPs and their errors (black points with error bars, measured with the scale on the bottom part of the chart). Results are presented for the fit to the Asimov data. The vertical dashed lines mark the -1 and 1 points on the bottom scale — expected span of the NP error. The systematic uncertainties are sorted by the estimation of the post-fit impact, only the first 30 systematic uncertainties are shown

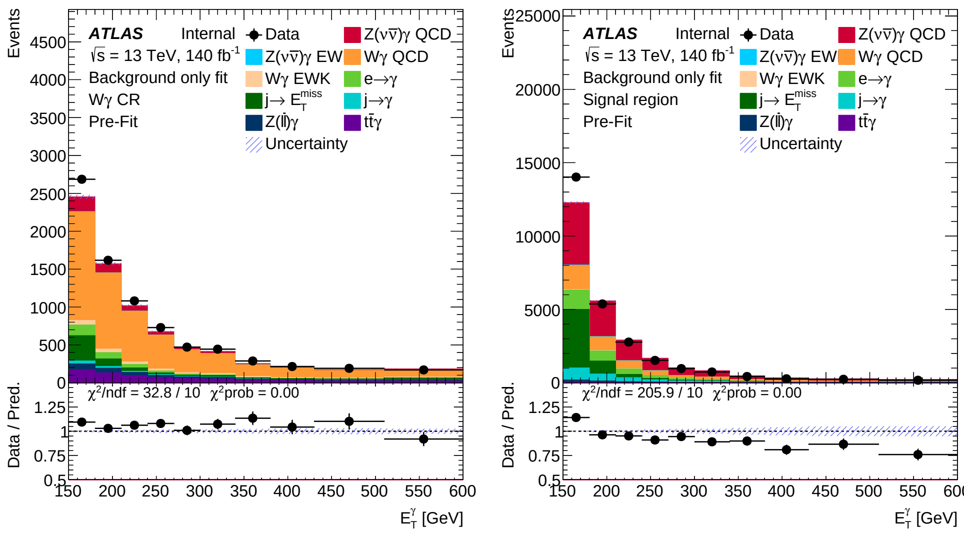


Figure 14: Distributions before the fit to the observed data in both $W\gamma$ CR and SR. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

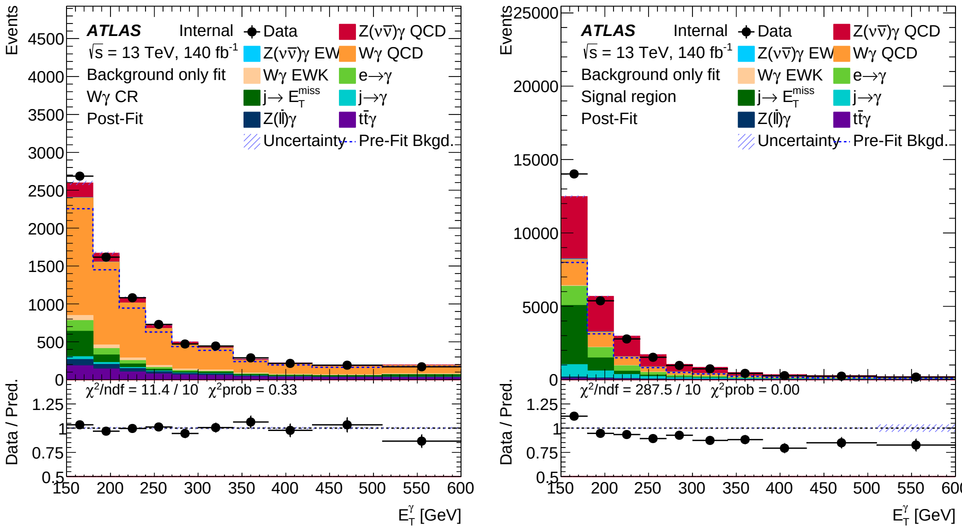


Figure 15: Distributions after the fit to the observed data in both $W\gamma$ CR and SR. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

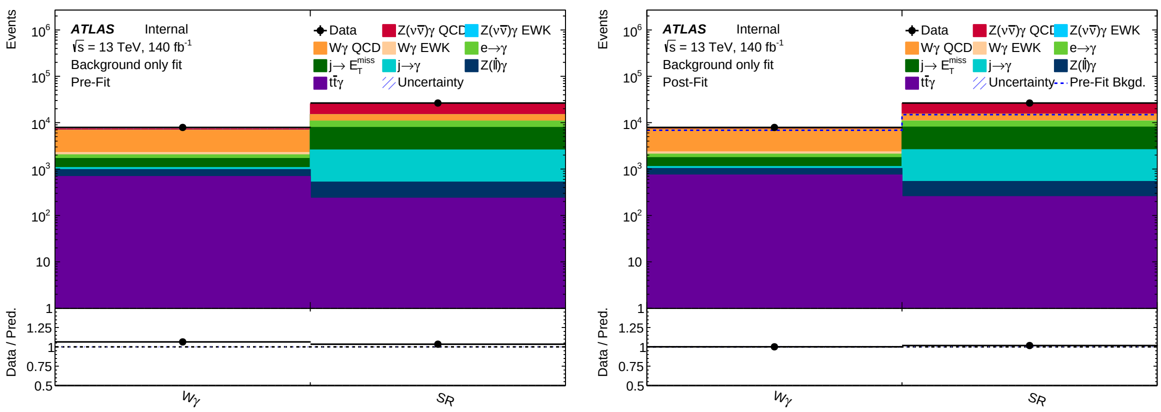


Figure 16: Summary of the processes yield for all of the regions before and after the fit to the observed data in both $W\gamma$ CR and SR. The vertical error bars on the data points correspond to the data statistical uncertainty. The dashed band corresponds to the combination of the MC statistical uncertainty and systematic uncertainties.

	$W\gamma$ CR	Signal region
$Z(\nu\bar{\nu})\gamma$ QCD	532 ± 5	10 600 ± 200
$Z(\nu\bar{\nu})\gamma$ EWK	11.59 ± 0.12	154 ± 3
$W\gamma$ QCD	4990 ± 90	4300 ± 80
$W\gamma$ EWK	263 ± 5	149 ± 3
$e \to \gamma$	334 ± 3	2870 ± 30
$j \to E_T^{miss}$	619 ± 6	5330 ± 50
$j \to \gamma$	104.1 ± 1.0	2070 ± 20
$Z(\ell\bar{\ell})\gamma$	280 ± 3	287 ± 3
$t\bar{t}\gamma$	749 ± 13	255 ± 5
Total	7890 ± 100	26 000 ± 200
Data	7892	26523

Table 14: Event yields for the signal and all of the background processes considered in this analysis after the unblinded background and signal fit. The yields are presented in the regions described in Section 3.6.2. The uncertainty of the expected yield consists of statistical and systematic uncertainties.

$\mu_{Z\gamma}$	1.01 ± 0.02(stat)
$\mu_{W\gamma}$	1.08 ± 0.02(stat)
Observed significance, σ	66

Table 15: Values of normalisation coefficients and the observed significance of the cross-section measurement after the fit to the observed data in both $W\gamma$ CR and SR.

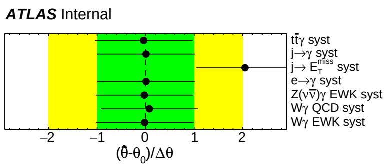


Figure 17: Nuisance parameters value and error (pull plot) after the fit to the observed data in both $W\gamma$ CR and SR.

Values and errors of the corresponding NPs for Asimov data fit are presented in Figure 17. Figure 18 shows the correlation matrix for NPs with the highest value of correlation coefficients for the Asimov data fit.

Figure 19 shows the estimations of the pre-fit and post-fit impact of the systematic uncertainties on the total systematic uncertainty of the POI in Asimov data fit.

6 Conclusion

Measurements of $Z\gamma$ production in $\sqrt{s}$ = 13 TeV pp collisions at the LHC are presented. The analysed data were collected with the ATLAS detector during full Run2 of the LHC and correspond to an integrated luminosity of 140 fb^{-1}. The events are selected using high E_T photon trigger. The dominant backgrounds are from γ+jet and $W\gamma$ processes and these are evaluated using two-dimensional sideband method and simultaneous fit to data respectively. Also significant backgrounds are from electrons-faking-photons

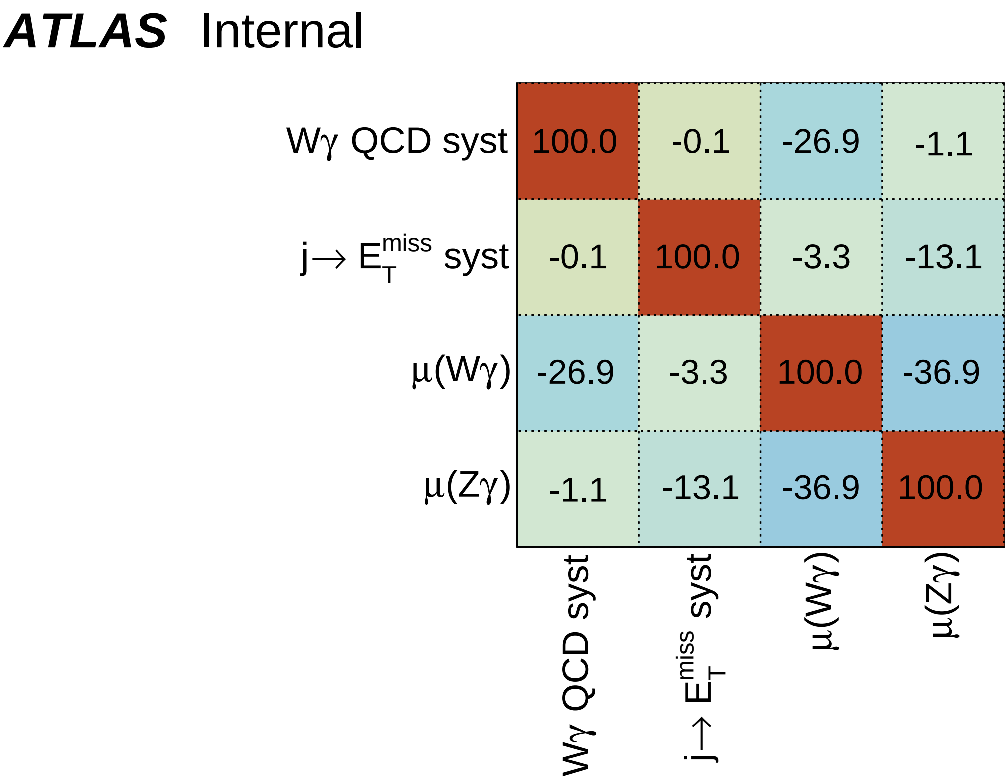


Figure 18: Correlation matrix of the parameters for the fit to the observed data in both $W\gamma$ CR and SR. Only parameters with the highest values of the correlation coefficient are shown.

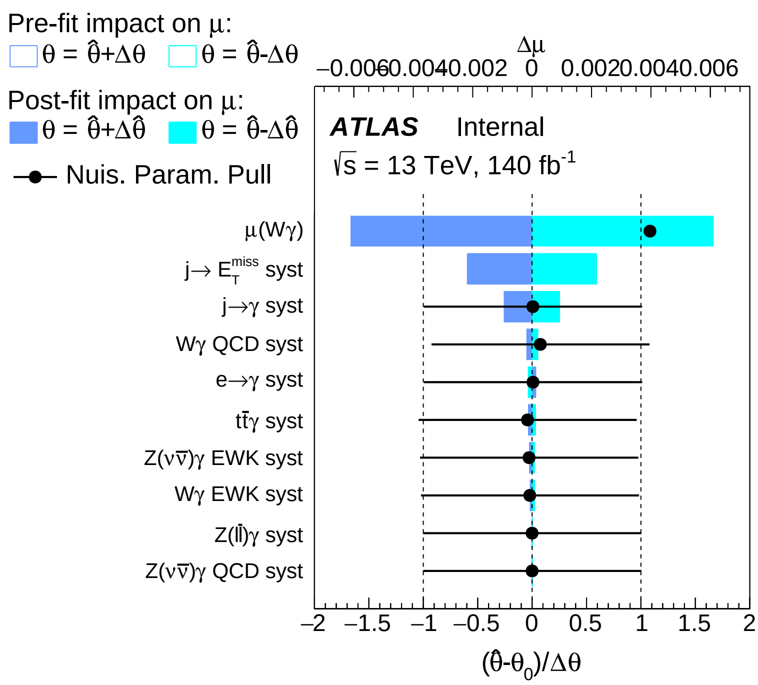


Figure 19: Estimations of the pre-fit (empty blue bars) and post-fit (filled blue bars) impact of the systematic uncertainties on the total systematic uncertainty of the POI (measured with the scale on the top part of the chart) and the values of the corresponding NPs and their errors (black points with error bars, measured with the scale on the bottom part of the chart). Results are presented for the fit to the observed data in both $W\gamma$ CR and SR. The vertical dashed lines mark the -1 and 1 points on the bottom scale — expected span of the NP error. The systematic uncertainties are sorted by the estimation of the post-fit impact, only the first 30 systematic uncertainties are shown

and jets-faking-photons and these are evaluated using other data-driven techniques. The measurement uses invisible decay mode of the gauge boson $Z \to \nu\bar{\nu}$ and is performed in a fiducial phase space closely matching the detector acceptance.

The cross section is measured integrally and differentially as a function of the transverse momentum p_{T} of the photon, pseudorapidity η of the photon, the missing transverse energy, the jet multiplicity and YYYY. The cross sections and kinematics are quoted for the sum of the three neutrino flavors. The measured cross sections and unfolded kinematic distributions are compared to SM predictions. The SM predictions agree well with the data. The NNLO parton-level generator MCFM, with a scale factor making a correction from the parton to the hadron level, is used for predictions of both the differential spectra and the absolute production cross sections. For further cross-check of the theoretical predictions, the computational framework POWHEG is used. Table **??** presents the cross-section comparison between data and the SM for $p + p \to \mathrm{Z} + \gamma + \mathrm{X}$.

Differential cross-section results can be found on the Figures **??**.

Having found no significant deviations from SM predictions, the data are used to put limits on triple anomalous couplings of photons to Z bosons from Z/γ^* s-channel production coupled to a final state Z boson and one photon (aTGC's). The limits obtained on the aTGC parameters h_3^V and h_4^V (V = Z or γ) are presented in Table **??**. The limits obtained with the EFT formalism aTGC parameters $C_{\tilde{B}W}/\Lambda^4$, C_{BW}/Λ^4, C_{BB}/Λ^4, C_{WW}/Λ^4, $G+$ and $G-$ are presented in Table **??**.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on p. 6).

[2] 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. 6).

[3] 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. 6), Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: `https://cds.cern.ch/record/1451888`.

[4] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: `2007.12539 [hep-ex]` (cit. on p. 6).

[5] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: `https://cds.cern.ch/record/2767187` (cit. on p. 6).

[6] 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 p. 7).

[7] *ElectronPhotonFourMomentumCorrection*, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronPhotonFourMomentumCorrection` (cit. on pp. 7, 9).

[8] *ElectronPhotonShowerShapeFudgeTool*, URL: `https://gitlab.cern.ch/atlas/athena/tree/master/PhysicsAnalysis/ElectronPhotonID/ElectronPhotonShowerShapeFudgeTool` (cit. on p. 7).

[9] C. Anastopoulos et al., *Photon identification pre-recommendations for run 2*, tech. rep. ATL-COM-PHYS-2015-496, CERN, 2015, URL: `https://cds.cern.ch/record/2022342` (cit. on p. 7).

[10] ATLAS Collaboration, *Photon identification in 2015 ATLAS data*, ATL-PHYS-PUB-2016-014, 2016, URL: `https://cds.cern.ch/record/2203125` (cit. on p. 7).

[11] *PhotonEfficiencyCorrection*, URL: `https://gitlab.cern.ch/atlas/athena/tree/master/PhysicsAnalysis/ElectronPhotonID/PhotonEfficiencyCorrection` (cit. on p. 7).

[12] L. Aubry, L. Carminati et al., *Search for new phenomena with the ATLAS detector in monophoton events from proton-proton collisions at sqrt(s) = 13 TeV*, tech. rep. ATL-COM-PHYS-2015-291, CERN, 2015, URL: `https://cds.cern.ch/record/2010478` (cit. on p. 8).

[13] A. Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, The European Physical Journal C **77** (2017), URL: `https://doi.org/10.1140%2Fepjc%2Fs10052-017-5031-2` (cit. on p. 8).

[14] *MCPAnalysisConsolidationMC16*, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisConsolidationMC16` (cit. on p. 8).

[15] *MuonSelectionTool*, URL: `https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionToolR21` (cit. on p. 8).

[16] *IsolationSelectionTool*, URL:
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool
(cit. on p. 9).

[17] *MuonEfficiencyCorrections*, URL: https://gitlab.cern.ch/atlas/athena/tree/master/PhysicsAnalysis/MuonID/MuonIDAnalysis/MuonEfficiencyCorrections (cit. on p. 9).

[18] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, tech. rep. ATL-PHYS-PUB-2015-027, CERN, 2015, URL: https://cds.cern.ch/record/2037904 (cit. on p. 9).

[19] *MissingETUtility*, URL:
https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MissingETUtility
(cit. on p. 9).

[20] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton–Proton Collisions recorded in 2012 at $\sqrt{s}$ = 8 TeV*, ATLAS-CONF-2013-082, 2013, URL: https://cds.cern.ch/record/1570993 (cit. on p. 10).

[21] B. H. Brunt et al., *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s}$ = 13 TeV*, tech. rep. ATL-COM-PHYS-2015-347, CERN, 2015, URL: https://cds.cern.ch/record/2013489 (cit. on p. 10).

[22] *JVTCalibration*,
URL: //twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations
(cit. on p. 10).

[23] R. White, F. Monticelli, A. Ruiz-Martinez and M. Backes, *Electron / photon trigger plots for LHCC*, tech. rep. ATL-COM-DAQ-2016-050, CERN, 2016,
URL: https://cds.cern.ch/record/2154228 (cit. on p. 11).

[24] G. M. Stefano Manzonia Leonardo Carminatia, *Electron-to-photon fake rate measurements - Supporting documentation for the paper Photon identification in 2015+2016 ATLAS data*, tech. rep. ATL-COM-PHYS-2017-1277, CERN, 2017,
URL: https://cds.cern.ch/record/2280801 (cit. on p. 12).

[25] *MCTruthClassifier*, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCTruthClassifier#Truth_particles_classification (cit. on p. 13).

[26] E. Soldatov et al., *Measurement of the electroweak $Z(\nu\bar{\nu})\gamma jj$ production cross section in pp Collisions at $\sqrt{s}$=13 TeV with the ATLAS Detector*, tech. rep., CERN, 2019,
URL: https://cds.cern.ch/record/2696408 (cit. on p. 16).

[27] A. Collaboration, Physical Review D **87** (2013),
URL: https://doi.org/10.1103%2Fphysrevd.87.112003 (cit. on p. 16).

[28] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, tech. rep., New York U., 2012, URL: https://cds.cern.ch/record/1456844 (cit. on p. 25).

[29] *TRExFitter*,
URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TtHFitter
(cit. on p. 25).

[30] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, (2012)
(cit. on p. 25).

[31] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: `1007.1727 [physics.data-an]` (cit. on pp. 26, 27), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[32] *NonCollisionBackgroundPublicResults*, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/NonCollisionBackgroundPublicResults` (cit. on p. 49).

Variable	1	2	3	4
$E_T^{miss} signif.$	> 11			—
$\Delta\phi(E_T^{miss}, \gamma)$	> 0.7			—
$\Delta\phi(E_T^{miss}, j_1)$	> 0.4			—
E_T^{miss}, GeV	>130			—
Signal				
$Z(\nu\nu)\gamma$QCD	9749 ± 8	9840 ± 8	10513 ± 8	13937 ± 9
$Z(\nu\nu)\gamma$EWK	138.3 ± 0.3	140.1 ± 0.3	152.1 ± 0.3	314 ± 0.4
Total signal	9887 ± 8	9980 ± 8	10675 ± 8	14251 ± 9
Background				
Wγ QCD	3600 ± 21	3645 ± 22	3936 ± 23	8838 ± 33
Wγ EWK	124.3 ± 0.7	126.1 ± 0.7	136.3 ± 0.7	430.7 ± 1.3
tt, top	186 ± 6	208 ± 6	241 ± 6	3217 ± 23
W(eν)	2942 ± 447	3259 ± 518	3430 ± 479	8569 ± 664
ttγ	210 ± 3	213 ± 3	234 ± 3	1075 ± 7
γ+j	4950 ± 51	5015 ± 52	5262 ± 53	45029 ± 145
Zj	213 ± 16	315 ± 20	403 ± 21	2575 ± 50
Z(ll)γ	266 ± 4	270 ± 4	285 ± 5	715 ± 7
W($\tau\nu$)	512 ± 75	697 ± 81	785 ± 79	5538 ± 144
Total bkg.	13003 ± 457	13747 ± 528	14712 ± 490	75988 ± 698
Stat. signif.	65.4 ± 0.7	64.8 ± 0.7	66.9 ± 0.6	47.44 ± 0.19
S/S_0	0.6938 ± 0.0007	0.7003 ± 0.0007	0.7491 ± 0.0007	1
B/B_0	0.171 ± 0.006	0.181 ± 0.007	0.194 ± 0.007	1
$1 - B/B_0$	0.829 ± 0.006	0.819 ± 0.007	0.806 ± 0.007	0
S/B	0.76 ± 0.03	0.73 ± 0.03	0.73 ± 0.02	0.1875 ± 0.0017

Table 16: The results of selection optimisation at three different working points *FixedCutTight*, *FixedCutTightCaloOnly*, *FixedCutLoose*.

Appendices

A Additional tests in the course of selection optimization

A.1 Selection optimisation at different working points

Table 16 shows the results of selection optimisation at three different working points *FixedCutTight*, *FixedCutTightCaloOnly*, *FixedCutLoose*. The optimal variable constraints are the same for the considered working points, but using the working point *FixedCutLoose* achieves greater statistical significance and preserves a larger number of signal events. The table also shows the event yields for each optimisation option and event yields before the constraints on the optimised variables are imposed.

	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$	$E_T^{miss}\ signif.$ E_T^{miss}, GeV $\Delta\phi(E_T^{miss},\gamma)$ $\Delta\phi(E_T^{miss},j_1)$ $\Delta\phi(E_T^{miss},j_2)$
Signal						
Z($\nu\nu$)γ QCD	9718 ± 8	11063 ± 8	9813 ± 8	9732 ± 8	9902 ± 8	9767 ± 8
Z($\nu\nu$)γ EWK	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Total signal	9718 ± 8	11063 ± 8	9813 ± 8	9732 ± 8	9902 ± 8	9767 ± 8
Background						
Wγ QCD	3143 ± 21	4413 ± 24	3236 ± 21	3206 ± 21	3510 ± 22	3299 ± 21
Wγ EWK	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
W(eν)	2936 ± 447	3733 ± 481	3141 ± 469	3190 ± 447	3248 ± 449	3197 ± 447
tt, top	177 ± 5	521 ± 9	180 ± 6	224 ± 6	235 ± 6	235 ± 6
ttγ	194 ± 3	537 ± 5	195 ± 3	197 ± 3	219 ± 3	214 ± 3
γ+j	7178 ± 77	17871 ± 120	7528 ± 78	7202 ± 77	8650 ± 81	7530 ± 78
Zj	211 ± 16	268 ± 17	213 ± 16	213 ± 16	223 ± 16	215 ± 16
Z(ll)γ	253 ± 4	335 ± 5	264 ± 4	254 ± 4	313 ± 5	267 ± 4
Total bkg.	14092 ± 455	27678 ± 497	14757 ± 477	14486 ± 455	16400 ± 457	14955 ± 455
Stat. signif.	63.0 ± 1.7	56.2 ± 1.8	62.6 ± 1.7	62.5 ± 1.7	61.1 ± 1.8	62.1 ± 1.7

Table 17: Comparison of statistical significance and event returns when each of the optimised variables is excluded. The excluded variable is highlighted in red.

A.2 Comparison of the efficiency of variable cuts

A.3 A test of the efficiency of the variable N_{b-jets}

A test of the efficiency of variable N_{b-jets} found that the optimal number of b-jets in an event does not exceed 1 (Table 18). However, it can be seen from Figure 20 as well as from Table 19 that the improvement in significance associated with the number of b-jets constraint does not exceed the statistical error. Thus, it was decided to abandon the phase space restriction for this variable.

A.4 A test of the efficiency of the variable $p_T^{SoftTerm}$

A test of the efficiency of the variable $p_T^{SoftTerm}$ showed that the optimal upper threshold value for this variable is 39. (Table 20). However, it can be seen from Figure 21 as well as from Table 20 that the improvement in significance associated with the $p_T^{SoftTerm}$ constraint does not exceed the statistical error. Thus, it was decided to abandon the phase space restriction for this variable.

Variable	Cut	Cut
$E_T^{miss} signif.$	> 11	—
$\Delta\phi(E_T^{miss}, \gamma)$	> 0.6	—
$\Delta\phi(E_T^{miss}, j_1)$	> 0.4	—
$\Delta\phi(E_T^{miss}, j_2)$	> 0.2	—
N_{b-jet}	< 2	—
E_T^{miss}, GeV	> 130	—
Stat. signif.	63.0 ± 1.7	41.9 ± 1.6
Total signal	9714 ± 8	12256 ± 9
Total bkg.	14039 ± 455	73234 ± 546

Table 18: Thresholds on variables obtained during optimisation as well as statistical significance and event yields before and after optimisation.

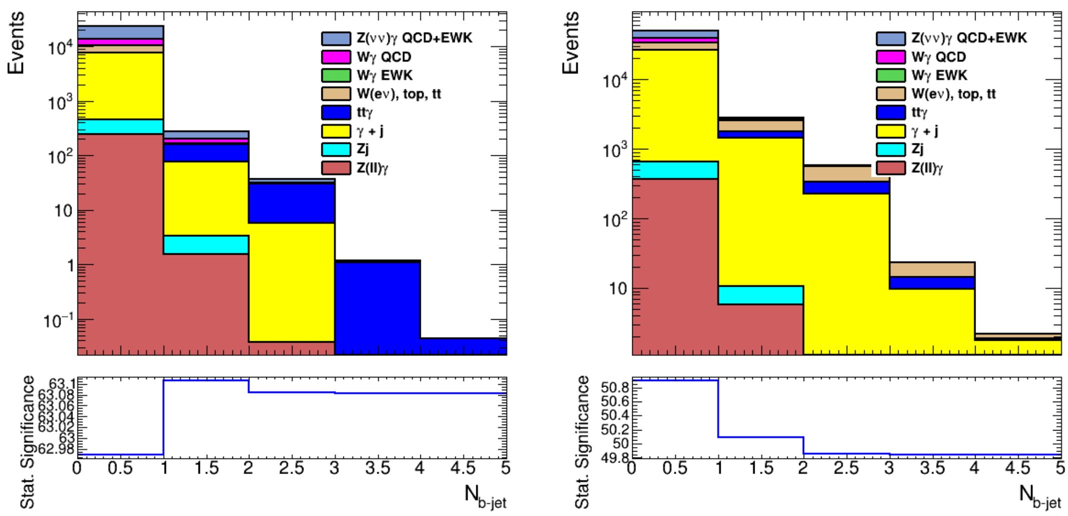


Figure 20: Distributions of the variable number of b-jets before (left) and after (right) optimisation. The left distribution is plotted in the phase space bounded by preselection conditions (Sec. 3.6.1). The right distribution is plotted in the phase space bounded by preselection conditions and optimal selections obtained during optimisation.

Signal				
	Before optimisation	$N_{b-jet} < \infty$	N_{b-jet}<2	N_{b-jet}<1
Z$(\nu\nu)\gamma$ QCD	12256 ± 9	9718 ± 8	9714 ± 8	9635 ± 8
Z$(\nu\nu)\gamma$ EWK	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Total signal	12256 ± 9	9718 ± 8	9714 ± 8	9635 ± 8
Background				
Wγ QCD	6528 ± 28	3143 ± 21	3142 ± 21	3105 ± 21
Wγ EWK	0 ± 0	0 ± 0	0 ± 0	0 ± 0
W(eν)	6952 ± 507	2936 ± 447	2936 ± 447	2928 ± 447
tt, top	1667 ± 17	177 ± 5	159 ± 5	88 ± 4
ttγ	874 ± 6	194 ± 3	167 ± 3	80.8 ± 1.9
γ+j	56284 ± 198	7178 ± 77	7171 ± 77	7091 ± 76
Zj	333 ± 18	211 ± 16	211 ± 16	210 ± 16
Z(ll)γ	596 ± 7	253 ± 4	253 ± 4	251 ± 4
Total bkg.	73234 ± 546	14092 ± 455	14039 ± 455	13753 ± 455
Stat. signif.	41.9 ± 1.6	63.0 ± 1.7	63.0 ± 1.7	63.0 ± 1.8

Table 19: Statistical significance and event yields before optimisation, with no restriction on the number of b-jets, with optimal restriction on the number of b-jets, and with full veto on b-jets

Variable	Cut	Cut
$E_T^{miss} signif.$	>11	>11
$\Delta\phi(E_T^{miss}, \gamma)$	> 0.6	> 0.6
$\Delta\phi(E_T^{miss}, j_1)$	> 0.4	> 0.4
$\Delta\phi(E_T^{miss}, j_2)$	> 0.2	> 0.2
E_T^{miss}, GeV	> 130	> 130
$p_T^{SoftTerm}$, GeV	—	< 39
Stat. signif.	63.0 ± 1.7	63.0 ± 1.8
Total signal	9718 ± 8	9627 ± 8
Total bkg.	14092 ± 455	13682 ± 455

Table 20: Results of selection optimisation with and without variable $p_T^{SoftTerm}$

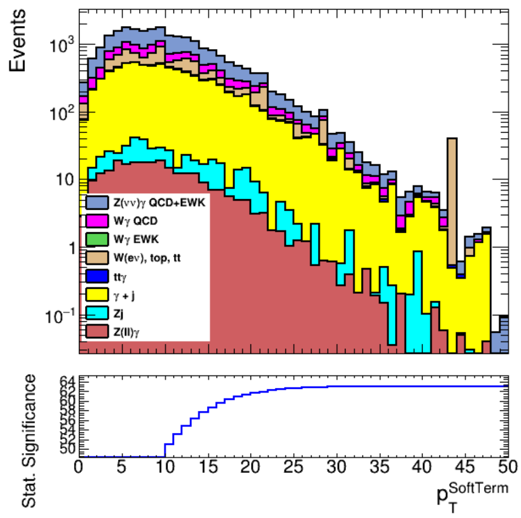


Figure 21: Distributions of the variable $p_T^{SoftTerm}$ after optimisation. The distribution is plotted in the phase space bounded by preselection conditions and optimal selections obtained during optimisation.

B Estimation of systematic uncertainties on $e \to \gamma$ fake rate

B.1 Background under Z peak evaluation and related systematics

As can be seen from Fig. 22 distributions on e-e and e-γ invariant mass have a different shapes of the background, therefore to estimate number of background events under Z peak the fit is performed in a different way for e-e and e-γ spectra.

The e-e spectra are fitted from the left and from the right sides of the Z peak separately. The ranges $(30, 60)$ GeV and $(120, 200)$ GeV are chosen to avoid the bump in the beginning of mass spectrum and not to get too close to the Z-peak. The exponential function with 2 parameters is used for the fit: $f(x) = exp(p_0 + p_1 \cdot x)$. The results can be seen in Fig. 22 on the left side plots. Then fit functions extrapolations from the left and from the right sides are used to calculate the integral under each of them in Z peak region $(M_Z - 10, M_Z + 10)$ GeV. It results in the maximum and the minimum estimation of the background: N^{bkg}_{max} and N^{bkg}_{min}. The average is taken for estimation of the nominal fake-rate value. N^{bkg}_{max} and N^{bkg}_{min} values are used as variations for obtaining systematics on background subtraction.

The e-γ spectra are fitted from the left and from the right sides of the Z peak by the same exponential function with 3 parametes: $f(x) = exp(p_0 + p_1 \cdot x + p_2 \cdot x^2)$. The ranges for the fit $(25, 70)$ GeV and $(110, 200)$ GeV are safely extended in comparison with e-e case to gain more statistics. The results of the fit can be seen in Fig. 22 on the right side plots. Fit function extrapolation to the region under the Z peak is used to obtain the integral in the region $(M_Z - 10, M_Z + 10)$ GeV, which is used as the background value for the estimation of the nominal fake-rate value. To estimate the systematic uncertainty on the integral the variations of fit parameters on their statistical uncertainties are performed.

To obtain the systematic uncertainty the largest deviations from the nominal fake-rate value coming from background variation in e-e and e-γ pairs are summed in quadrature.

B.2 Real fake-rate estimation

For estimation of "real fake rate" in $Z(ee)$ MC is calculated as ratio of all tag-n-probe $e - \gamma$ pairs to all tag-n-probe $e - e$ pairs, in which each particle is checked by MCTruthClassifier as described in Sec. 4.1.1 to be either electron coming from Z boson or misidentified photon coming from Z boson or final state radiation.

"Real fake rate" is estimated in the same 3 regions on η and p_T as fake-rate in data.

B.3 Resulting systematics

Fake rates obtained from $Z(ee)$ MC and used for systematic evaluation are listed in Table 21. For mass-window variation, the largest deviations from the nominal fake-rate value in $Z(ee)$ (line 1 of the forementioned table) in each region are listed in line 2. "Real fake rate" in $Z(ee)$ MC is presented in line 3.

Resulting values of each systematic component are listed in Table 22 in percents. Relative difference between nominal and the largest mass-window deviation fake-rates are shown in line 1. Relative systematic

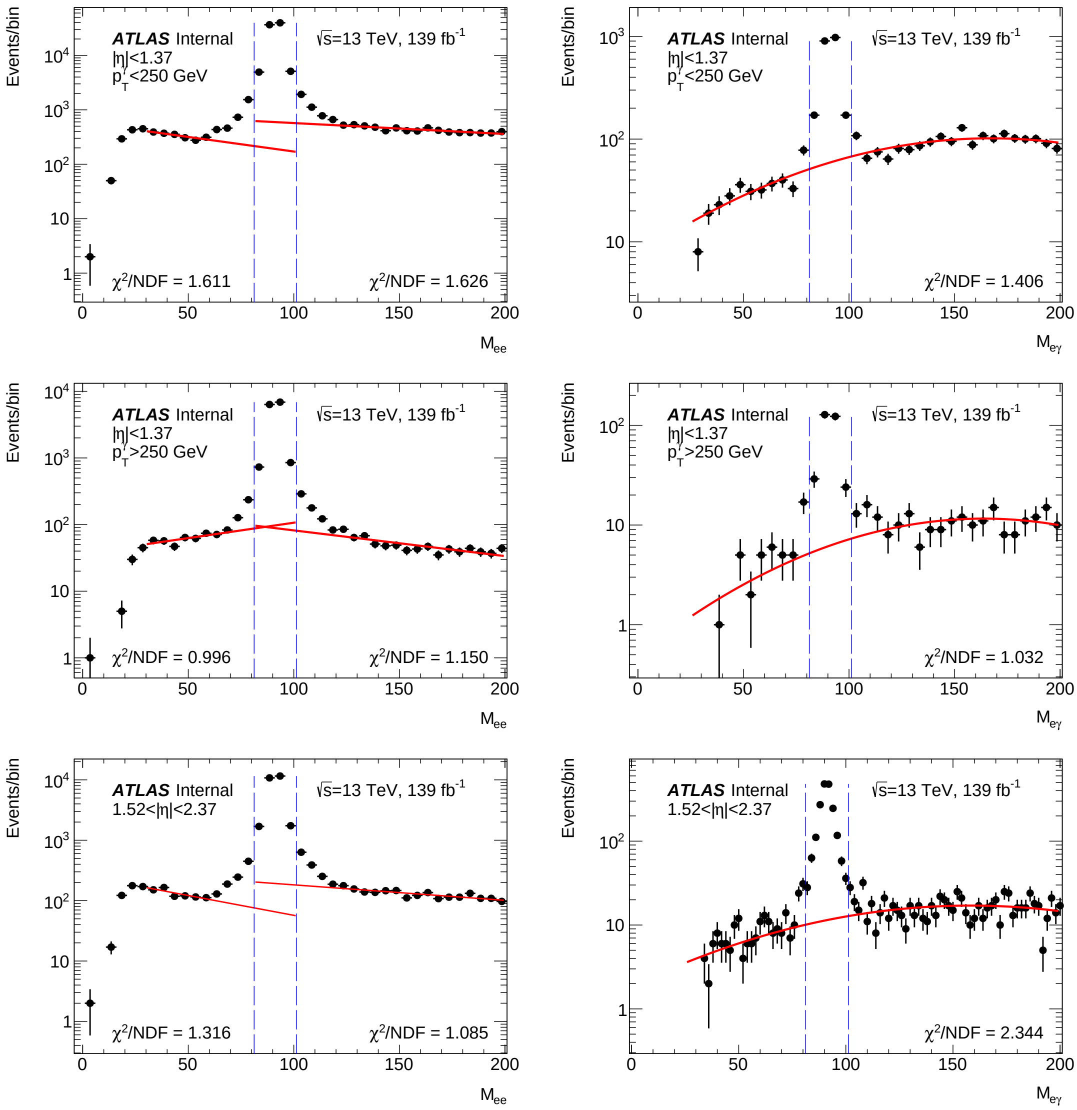


Figure 22: Fit of sideband regions and extrapolation of the combinatorial background to the region under Z peak. ee-pairs on the left, $e\gamma$-pairs on the right. Top – central low-p_T region, middle - central high-p_T region, bottom - forward region.

fake rate	$150 < E_T^\gamma < 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$E_T^\gamma > 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$1.52 < \lvert\eta\rvert < 2.37$
$Z(ee)$ MC tag-n-probe	0.0214 ± 0.0004	0.0199 ± 0.0006	0.0752 ± 0.0012
$Z(ee)$ MC mass window variation	0.0213 ± 0.0004	0.0200 ± 0.0006	0.0756 ± 0.0012
$Z(ee)$ MC "real"	0.022 ± 0.002	0.023 ± 0.002	0.084 ± 0.004

Table 21: Electron-to-photon fake rates estimated in MC.

fake rate	$150 < E_T^\gamma < 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$E_T^\gamma > 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$1.52 < \lvert\eta\rvert < 2.37$
syst. from mass window var.:	0.3%	0.8%	0.5%
syst. from tag-n-probe and real f.r.:	2.8%	18%	11%
Background fit variation	3.5%	11%	3%
Total syst.:	4.5%	21%	12%

Table 22: Electron-to-photon fake rate systematics components.

fake rate	$150 < E_T^\gamma < 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$E_T^\gamma > 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$1.52 < \lvert\eta\rvert < 2.37$	Total
syst. on fake-rate estimation.	4.5%	21%	12%	
syst. from stat. unc. on fake-rate	2.6%	6.7%	2.6%	
syst. from impurity of CR	0.85%	0.85%	0.85%	
Total rel. syst.	5.3%	22%	12.3%	
Event yield in (incl.) e-probe CR	51923	11971	22603	
Fake-rate	0.0234	0.0195	0.0704	
$e \to \gamma$ event yield in SR	1213.4	233.2	1592.2	3039
Total abs. syst.	64.3	51.3	195.8	209

Table 23: Systematics breakdown for $e \to \gamma$ background for SR.

uncertainty from difference in the "real fake rate" and the nominal fake rate in $Z(ee)$ MC is shown in line 2. Systematics on estimation of the background under Z peak described in Sec. B.1 is shown line 3.

Total systematic uncertainty on fake-rate is calculated as sum in quadrature of all four components and is shown in the bottom line of the table.

Total systematics breakdown on $e \to \gamma$ background in each of η and p_T regions are shown in Table 23. The numbers are shown for SR. Total systematics in SR is calculated as a sum in quadrature of the total absolute systematic uncertainties for every η and p_T regions. It counts 6.8% for SR.

C $e \to \gamma$ fake rate: forward region division in E_T^γ

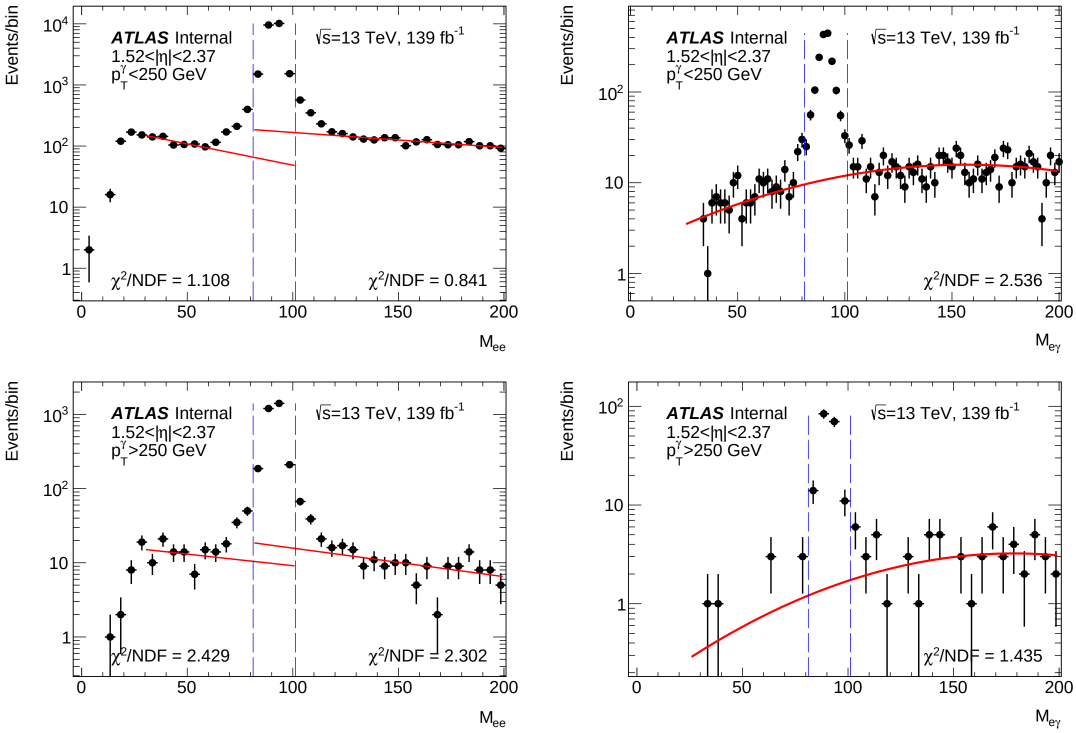


Figure 23: Fit of sideband regions and extrapolation of the combinatorial background to the region under Z peak. ee-pairs on the left, $e\gamma$-pairs on the right. Top – forward low-p_T region, bottom – forward high-p_T region.

D The suppression of beam-induced background events

This appendix consists in a summary of the studies done to understand the problem related to an anomalous rate of photons in the *loose'3* and isolated region of the data-driven background estimation method for events with jets misidentification as photons.

Figure 24 shows z coordinate pointed by the photon candidate with respect to the identified primary vertex $\Delta z = z_\gamma - z_{\rm vtx}$ distribution in data for different regions of ABCD method, described in Section 4.2.

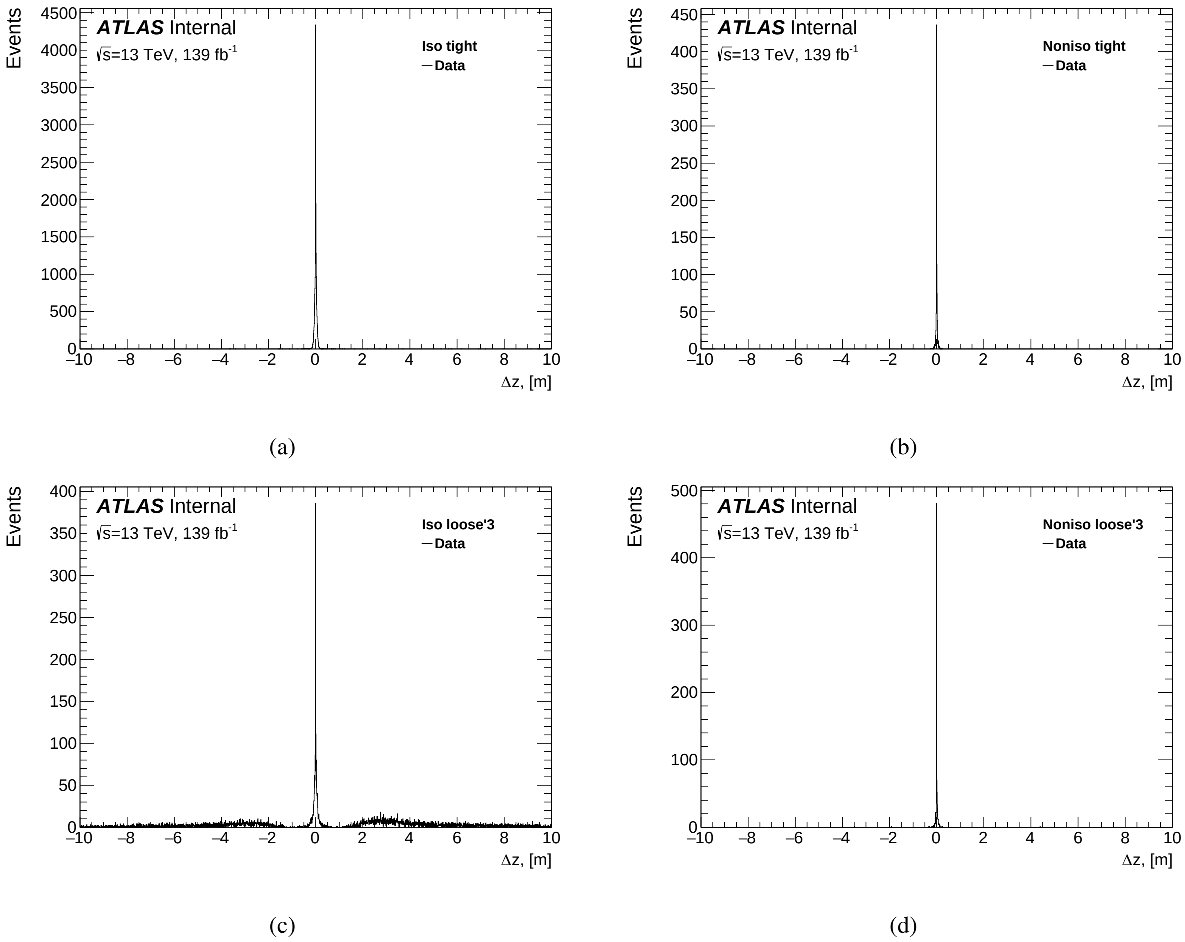
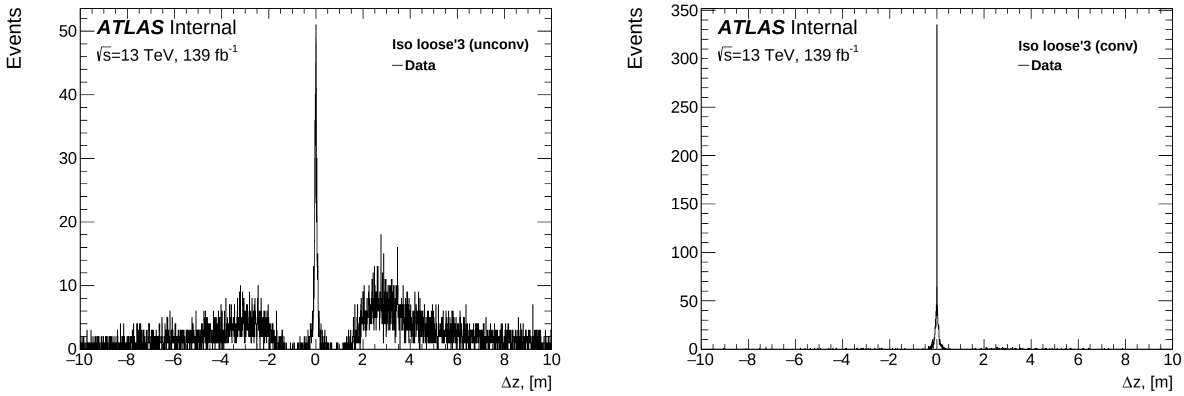


Figure 24: Δz distribution in data for (a) isolated *tight* photons, (b) non-isolated *tight* photons, (c) isolated *loose'3* photons and (d) non-isolated *loose'3* photons.

As for the tight photons the Δz is peaked at zero and for the region enriched in beam-induced background (BIB) [32] most of the events are concentrated at much higher values of Δz; Δz cut is thus used as a variable to discriminate between hard scatter photons and photons induced by beam background.

The *loose'3* and isolated region is the most affected region with the greatest contribution from unconverted photons as shown in Figure 25. Moreover, those photons are concentrated around $|\phi| = 0$, $|\phi| = 3$ and $|\eta| = 2$ as shown in Figure 26 pointing to beam induced background.



(a) (b)

Figure 25: Δz distribution in the *loose'3* and isolated region for (a) unconverted and (b) converted photons.

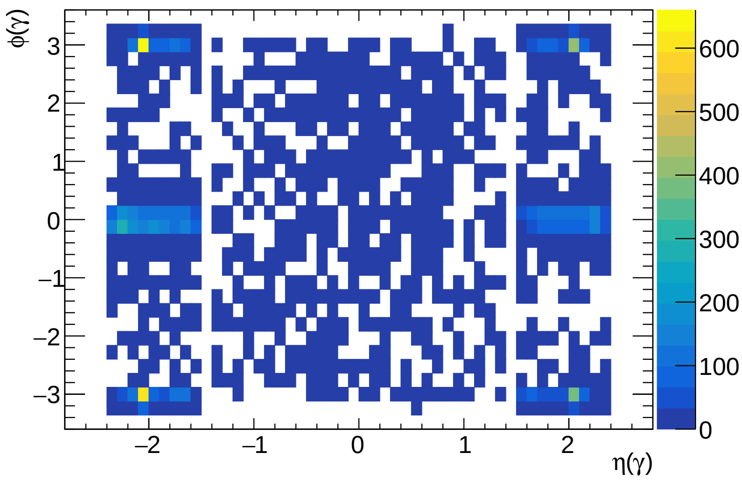


Figure 26: Distribution of photon ϕ versus photon η in the *loose'3* and isolated region.

By applying cuts on $|\phi| < 0.2$, $|\phi| \in [2.9, 3.2]$ and $|\eta| > 1.7$ for unconverted photons in the *loose'3* and isolated region one can be sure that beam induced background events concentrate in this area as reported in Figure 27.

A cut on Δz has been optimized directly for the analysis selections in order to reject the maximum number of events in the *loose'3* isolated region and at the same time accept the highest fraction of hard scatter photons in the *tight* isolated region. Figure 28(a) shows the acceptance efficiency of *tight* and isolated photons from the data. Figure 28(b) shows the rejection efficiency in the data for the *loose'3* isolated region as a function of the applied cut on Δz in mm.

When requiring $|\Delta z| < 250$ mm, the rejection efficiency is found to be $100 \pm 2\%$ in data, while the acceptance on signal photons is $99.7 \pm 0.9\%$, which means that this cut is optimal.

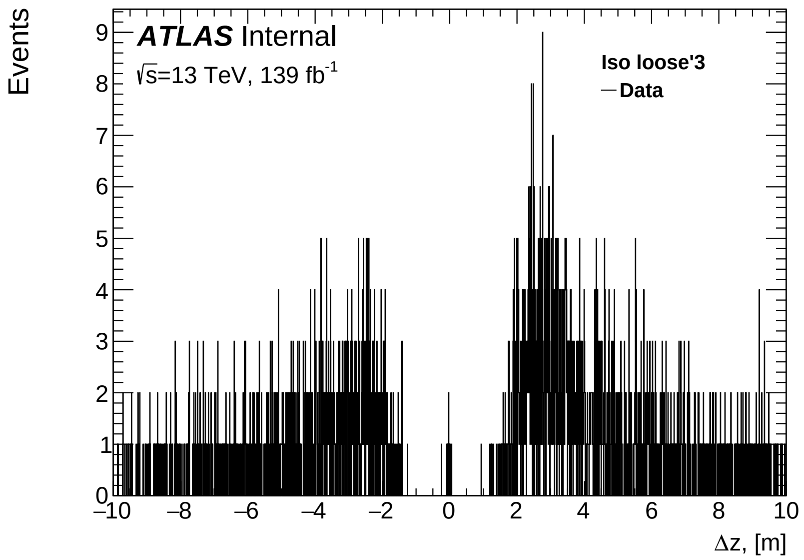
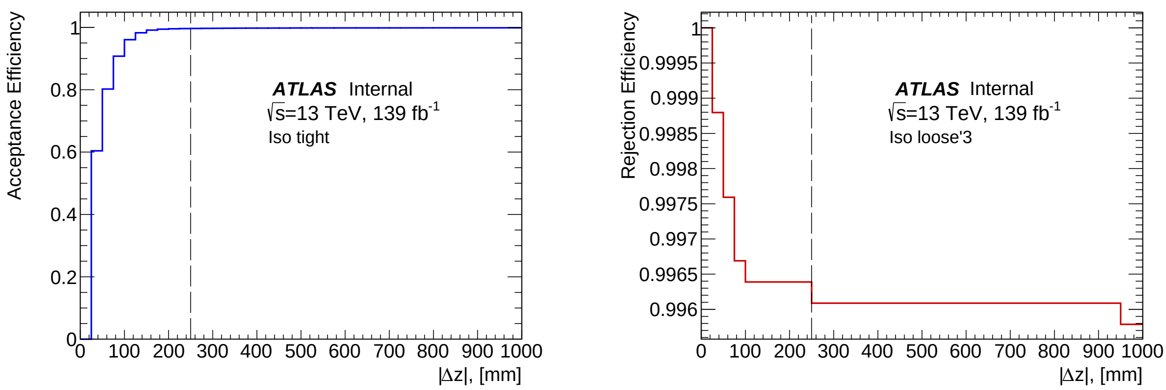


Figure 27: Δz distribution in the *loose'3* and isolated region for unconverted photons with cuts on $|\phi| < 0.2$, $|\phi| \in [2.9, 3.2]$ and $|\eta| > 1.7$.



(a)

(b)

Figure 28: (a) Acceptance efficiency in the *tight* isolated region and (b) rejection efficiency for *loose'3* isolated region in data.

E Pile-up background

F The choice of the isolation working point

Three isolation working points were considered: FixedCutTight, FixedCutTightCaloOnly and FixedCutLoose. Along with the nominal case, where a photon is required to pass $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} < 0.05$ isolation in all control regions, the track isolation can be additionally inverted $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\gamma} > 0.05$ in non-isolated regions to reduce correlations between photon identification and isolation variables. The distributions of isolation variables for each isolation working point and *loose'* are shown in Figure 29, Figure 30 and Figure 31.

The upper cut on the isolation energy can only be used only in FixedCutTight to prevent a deviation of R factor in data from 1. To define the regions where the R factor is stable, the cut $E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 p_{\mathrm{T}}^{\gamma} < 25.45$ GeV is applied.

The estimates for R factor in MC and data for all working points are reported in Table 24. The least correlation is observed for FixedCutLoose working point.

FixedCutTight working point				
R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.05 ± 0.14	1.21 ± 0.15	1.38 ± 0.16	1.65 ± 0.19
Data-driven	1.08 ± 0.07	1.14 ± 0.07	1.11 ± 0.06	1.15 ± 0.06
FixedCutTight working point with cut 25.45 GeV				
R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.07 ± 0.16	1.17 ± 0.17	1.18 ± 0.16	1.31 ± 0.17
Data-driven	1.11 ± 0.13	1.11 ± 0.12	1.14 ± 0.11	1.17 ± 0.11
FixedCutTightCaloOnly working point				
R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.18 ± 0.13	1.31 ± 0.13	1.37 ± 0.13	1.54 ± 0.14
Data-driven	1.14 ± 0.06	1.20 ± 0.06	1.19 ± 0.06	1.22 ± 0.06
FixedCutLoose with inverted track isolation				
R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.11 ± 0.13	1.23 ± 0.12	1.34 ± 0.12	1.60 ± 0.13
Data-driven	0.97 ± 0.10	1.05 ± 0.10	1.05 ± 0.09	1.06 ± 0.08

Table 24: Estimated correlation factor R between photon identification and isolation variables in MC and data for different isolation working points.

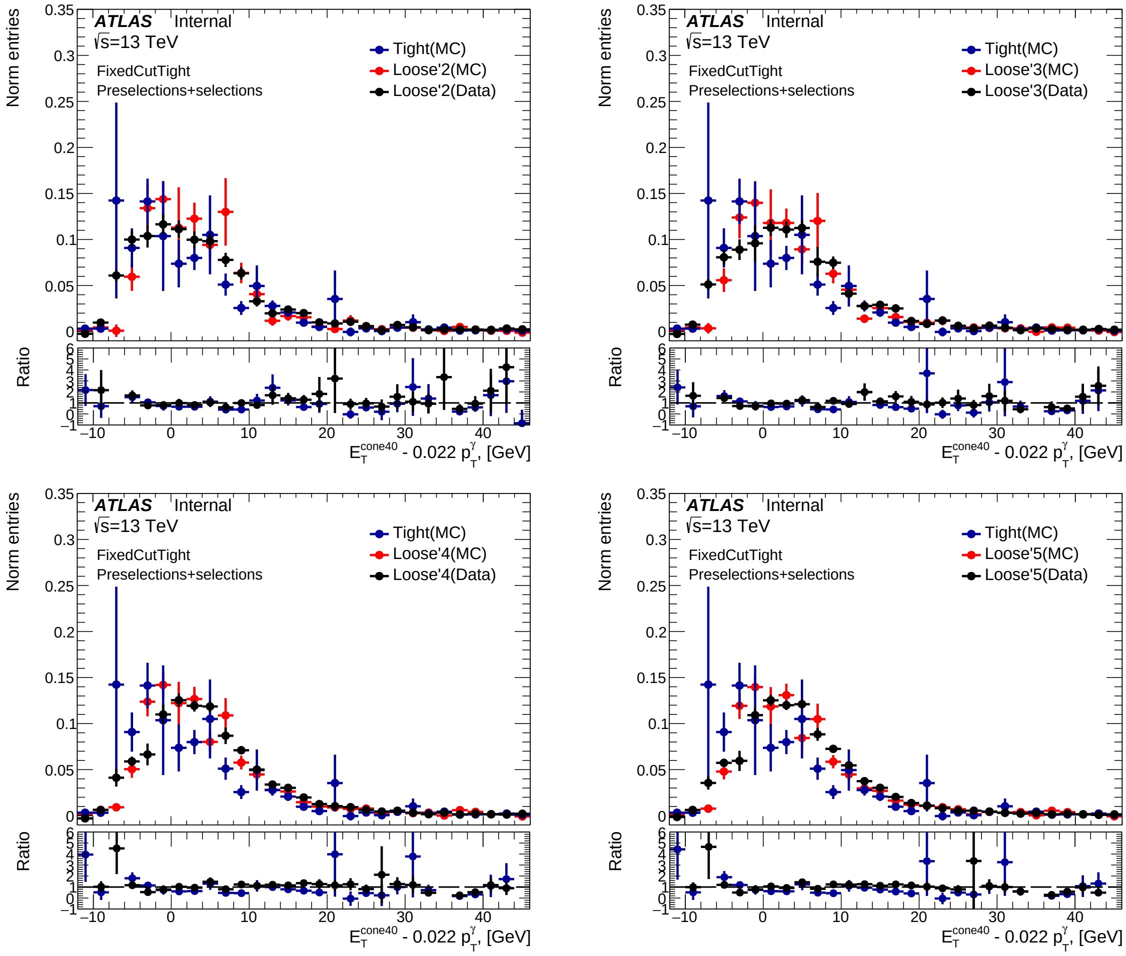


Figure 29: $E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 p_{\mathrm{T}}^{\gamma}$ distribution for FixedCutTight working point for $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay in data and MC. The bottom panel shows the ratio of tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay simulation and anti-tight photon candidates in data to the anti-tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay.

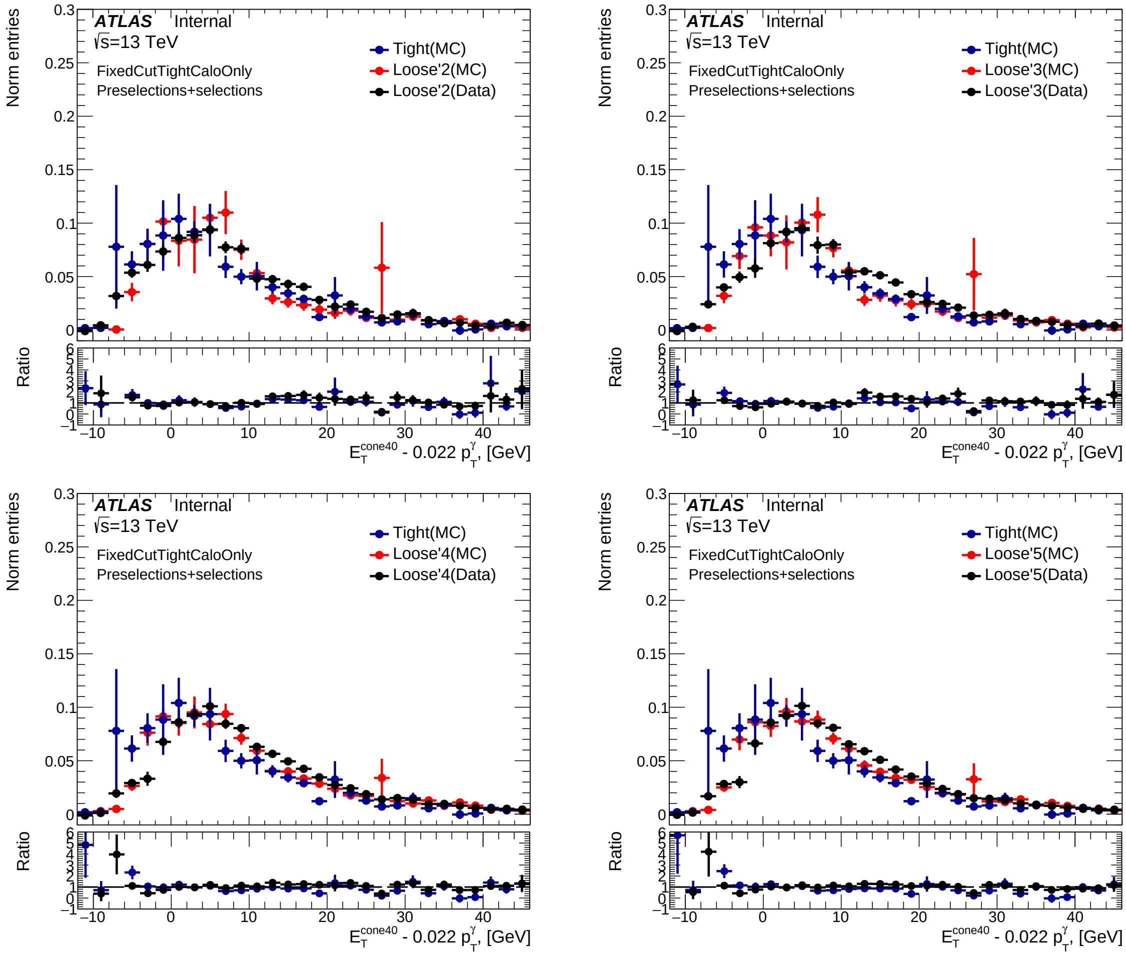


Figure 30: $E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 p_{\mathrm{T}}^{\gamma}$ distribution for FixedCutTightCaloOnly working point for $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay in data and MC. The bottom panel shows the ratio of tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay simulation and anti-tight photon candidates in data to the anti-tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay.

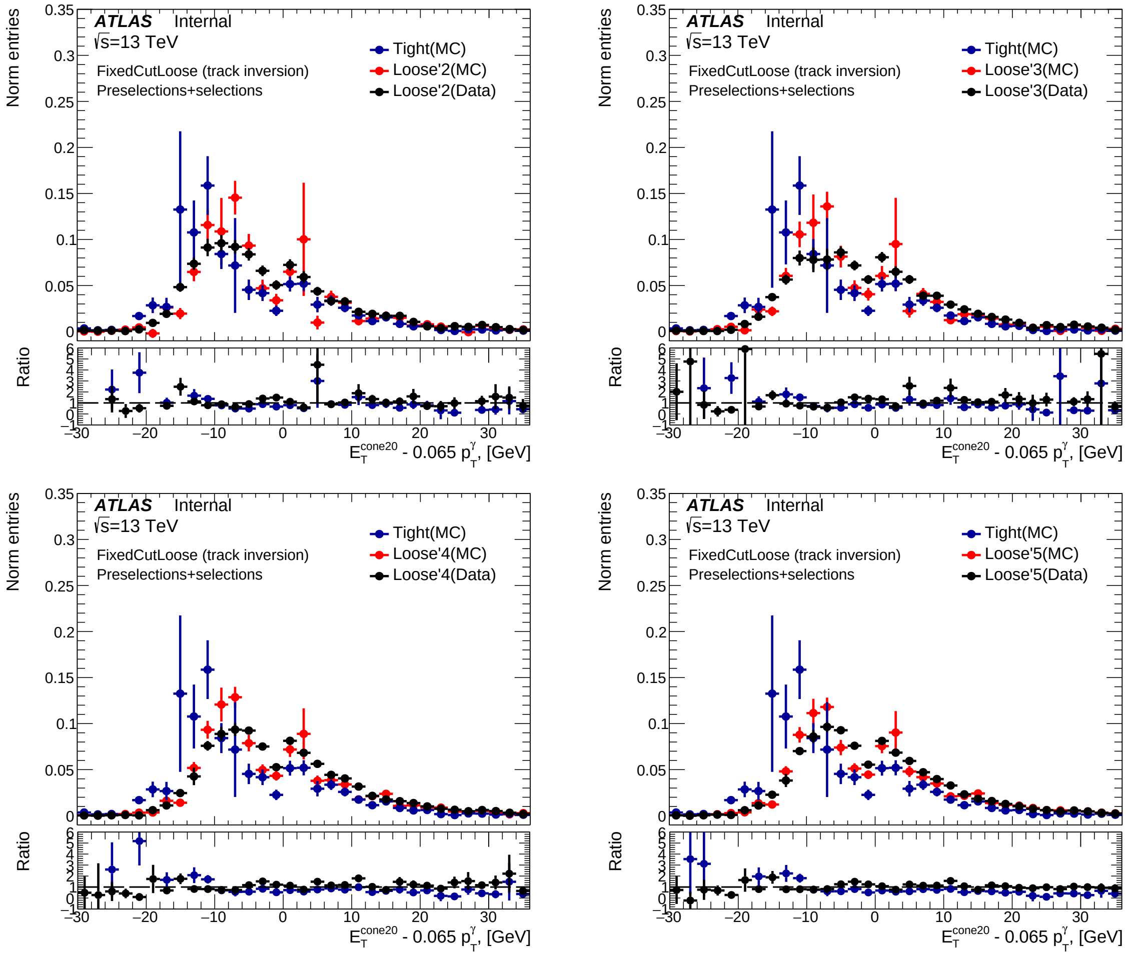


Figure 31: $E_{\mathrm{T}}^{\mathrm{cone40}} - 0.022 p_{\mathrm{T}}^{\gamma}$ distribution for FixedCutLoose working point for $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay in data and MC. The bottom panel shows the ratio of tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay simulation and anti-tight photon candidates in data to the anti-tight photon candidates from $Z(\nu\bar{\nu})$+jets and hadronic $W(\tau\nu)$ decay.

G The choice of the isolation gap value

The non-isolated control regions in the data-driven background estimation method for $jet \to \gamma$ events can be defined by the isolation gap between those control regions and the isolated regions which passes the *FixedCutLoose* isolation cut. This impacts mainly the signal leakage to control regions B and D. By increasing the isolation gap, the signal leakage becomes smaller, but the statistics in the control regions B and D also decrease. Table 25 shows the signal leakage parameters predicted by the signal MC for *loose'3* and different values of isolation gap.

Isolation gap	c_B	c_D
0 GeV	0.01003 ± 0.00008	0.00055 ± 0.00002
1 GeV	0.00970 ± 0.00007	0.00053 ± 0.00002
2 GeV	0.00939 ± 0.00007	0.00051 ± 0.00002
3 GeV	0.00908 ± 0.00007	0.00049 ± 0.00002
4 GeV	0.00879 ± 0.00007	0.00047 ± 0.00002

Table 25: Fraction of signal leakage to control regions B and D, c_B and c_D, for *loose'3* and different isolation gaps.

The isolation gap of 2 GeV is chosen as a baseline to ensure small enough signal leakage and also smaller uncertainties on the overall $Z(\nu\bar{\nu}) + jets$ and hadronic $W(\tau\nu)$ decay backgrounds prediction.

H Data-driven R factor estimation

To estimate the correlation between non-tight photon identification working point and isolation in data two additional regions are introduced as shown in Fig. 32:

- Tight and extra non-isolated region (control region E): events have a leading photon candidate that is extra non-isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} > 4.5$ GeV) and passes the *tight* selection.
- Non-tight and extra non-isolated region (control region F): events have a leading photon candidate that is extra non-isolated ($E_{\mathrm{T}}^{\mathrm{cone20}} - 0.065 p_{\mathrm{T}}^{\gamma} > 4.5$ GeV) and passes the *non-tight* selection.

In this case instead of formula

$$R = \frac{N_{\mathrm{A}}^{\mathrm{MC}} N_{\mathrm{D}}^{\mathrm{MC}}}{N_{\mathrm{B}}^{\mathrm{MC}} N_{\mathrm{C}}^{\mathrm{MC}}},$$

which was used for MC, one can use

$$R = \frac{N_{\mathrm{B-E}}^{\mathrm{data}} N_{\mathrm{F}}^{\mathrm{data}}}{N_{\mathrm{D-F}}^{\mathrm{data}} N_{\mathrm{E}}^{\mathrm{data}}},$$

where $N_{\mathrm{i}}^{\mathrm{data}}$ are the numbers of estimated $jet \to \gamma$ events in corresponding regions in data. To obtain these values, the numbers of events for other backgrounds and signal are subtracted from data yield in each region. Resulting R factors are shown in Table 6 of Section 4.2.

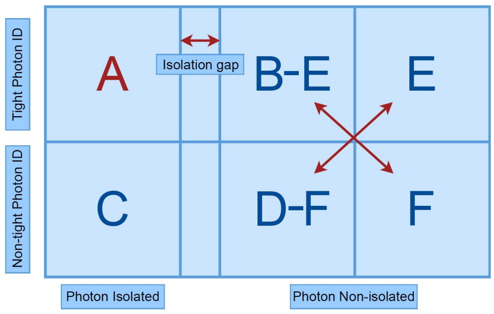


Figure 32: Schematic illustration of the two-dimensional plane based on photon isolation and identification variables with separation on A, B, C, D and extra E, F control regions.