Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016

The ATLAS Collaboration

DOI
10.1140/epjc/s10052-019-6650-6

Publication date
2019

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016

The ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 12 October 2018 / Accepted: 5 February 2019 / Published online: 7 March 2019
© CERN for the benefit of the ATLAS collaboration 2019

Abstract The efficiency of the photon identification criteria in the ATLAS detector is measured using 36.1 fb\(^{-1}\) to 36.7 fb\(^{-1}\) of pp collision data at \(\sqrt{s} = 13\) TeV collected in 2015 and 2016. The efficiencies are measured separately for converted and unconverted isolated photons, in four different pseudorapidity regions, for transverse momenta between 10 GeV and 1.5 TeV. The results from the combination of three data-driven techniques are compared with the predictions from simulation after correcting the variables describing the shape of electromagnetic showers in simulation for the average differences observed relative to data. Data-to-simulation efficiency ratios are determined to account for the small residual efficiency differences. These factors are measured with uncertainties between 0.5% and 5% depending on the photon transverse momentum and pseudorapidity. The impact of the isolation criteria on the photon identification efficiency, and that of additional soft pp interactions, are also discussed. The probability of reconstructing an electron as a photon candidate is measured in data, and compared with the predictions from simulation. The efficiency of the reconstruction of photon conversions is measured using a sample of photon candidates from \(Z \rightarrow \mu\mu\gamma\) events, exploiting the properties of the ratio of the energies deposited in the first and second longitudinal layers of the ATLAS electromagnetic calorimeter.

1 Introduction

Processes with prompt photons in the final state, occurring in proton–proton collisions at the Large Hadron Collider (LHC), play a central role in the ATLAS physics programme. They encompass all phenomena where photons do not originate from hadron decays. These range from non-resonant QCD production, where prompt photons are produced in association with jets or in pairs with cross sections of the order of tens of nanobarns or picobarns respectively, to rarer processes where prompt photons arise from the decay of a heavy particle. The study of QCD prompt photon production at the LHC and the measurement of the corresponding production cross sections allow a test of perturbative and non-perturbative regimes of QCD, and can provide useful information about the proton’s parton distribution functions (PDFs) (see for instance Ref. [1] for a first measurement at \(\sqrt{s} = 13\) TeV). The excellent capability of the ATLAS detector to reconstruct, identify and calibrate prompt photons has proved fundamental to the discovery of the Higgs boson and to the precision measurement of its properties with the \(H \rightarrow \gamma\gamma\) decay [2,3]. Similarly, prompt photons are paramount to several searches for phenomena beyond the Standard Model (SM), where they would come from the decay of various new heavy states [4–6].

The identification of prompt photons in hadronic collisions is particularly challenging, since the overwhelming majority of reconstructed photon candidates arise from background non-prompt photons from hadron decays in jets, while a smaller fraction of fake candidates are associated with hadrons that deposit significant energy in the electromagnetic calorimeter, mimicking that of real photons. Prompt photons are identified in the ATLAS experiment by means of selections on quantities describing the shape and properties of the associated electromagnetic showers, and by requiring them to be isolated from other particles in the event. These selections are separately optimised for those photon candidates that convert into an electron–positron pair before reaching the electromagnetic calorimeter (converted photons), and those photon candidates that are not associated with a conversion (unconverted photons). As already observed using LHC data in Run 1 [7], the efficiency of the selection criteria is modelled by Monte Carlo (MC) simulation with only \(\mathcal{O}(2 - 5\%)\) accuracy, being mostly limited by an imperfect description of the electromagnetic shower development in the detector. The photon identification effi-
ciency can on the other hand be measured in data with a precision of $O(1\%)$. Corrections are thus applied to the MC-simulated samples in order to guarantee the highest possible accuracy for photon measurements.

In this paper, the reconstruction and identification of photons by the ATLAS detector are described, and the measurements of the identification efficiency using 36.1 fb$^{-1}$ to 36.7 fb$^{-1}$ of $pp$ collisions collected at $\sqrt{s} = 13$ TeV in 2015 and 2016 are reported. These measurements are based on the techniques developed for the photon identification efficiency measurement performed with $\sqrt{s} = 7$ TeV and 8 TeV data [7], while addressing the different beam conditions at $\sqrt{s} = 13$ TeV, and exploiting the larger integrated luminosity. The measurements reported in this paper focus on the identification criteria optimised for the data-taking period at $\sqrt{s} = 13$ TeV that have been revisited relative to those used for the $\sqrt{s} = 8$ TeV run, and reported in Ref. [7], in order to better cope with the larger average number of interactions per beam bunch crossing $\mu$.

To overcome the difficulties arising from the absence of a single, pure control sample of prompt-photons over a large range of transverse momentum, three different data-driven techniques are used. The first method selects photons from radiative decays of the $Z$ boson, i.e. $Z \rightarrow \ell\ell\gamma$. The second one extrapolates photon properties from electrons and positrons\(^1\) from $Z$ boson decays, by exploiting the similarity of the photon and electron interactions in the ATLAS electromagnetic calorimeter. The third approach exploits a track-based measure of photon isolation to determine the fraction of background present in a sample of isolated photon candidates. Each of these techniques can measure the photon identification efficiency in overlapping parts of complementary $E_T$ regions.\(^2\) The combination of different measurements in the overlapping regions further improves the photon efficiency precision, which is measured for candidates with transverse momentum ranging from about 10 GeV to about 1.5 TeV.

The measurement of the rate of misidentification of electrons as photon candidates, as well as the results of a novel technique for measuring the efficiency of reconstructing a photon conversion, only deployed for the $\sqrt{s} = 13$ TeV data taking, are reported.

This paper is organised as follows. An overview of the ATLAS detector is provided in Sect. 2. The photon reconstruction and identification algorithms used in ATLAS for the data taking at $\sqrt{s} = 13$ TeV are detailed in Sect. 3, highlighting the differences relative to the reconstruction procedure reported in Ref. [7] and the properties of the identification criteria optimised for the $\sqrt{s} = 13$ TeV data taking.

Section 4 summarizes the collision and simulated data samples used for the various measurements, and describes the corrections applied to the simulated photon shower shapes in order to improve agreement with the data. In Sect. 5 the three data-driven approaches to the measurement of the photon identification efficiency are described, listing their respective sources of uncertainty and the precision reached in the relevant $E_T$ ranges. Their combination in the overlapping $E_T$ regions is presented, as well as a comparison of the combined data-driven photon identification efficiency with the MC predictions. The impact of the isolation criteria on the photon identification efficiency, and that of additional soft $pp$ interactions, is also discussed. The measurement of the rate of misidentification of electrons as photon candidates is reported in Sect. 6. The efficiency of reconstructing a photon conversion is summarised in Sect. 7.

2 The ATLAS detector

The ATLAS experiment [8] uses a multipurpose particle detector with approximately forward–backward symmetric cylindrical geometry and nearly $4\pi$ coverage in solid angle.\(^3\) It consists of an inner tracking system surrounded by a thin superconducting solenoid producing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large toroid magnet assemblies.

The inner detector (ID) tracking system provides position measurements for charged particles in the range $|\eta| < 2.5$ by combining information from three subdetectors. The ID consists of a cylindrical central region (full coverage for $|\eta| < 1.5$) arranged around the beam pipe, and two end-caps. Disks in the endcap region are placed perpendicular to the beam axis, covering $1.5 < |\eta| < 2.5$. Starting from the interaction point, the high-granularity silicon pixel detector segmented in $r$-$\phi$ and $z$ covers the vertex region and typically provides four three-dimensional measurements per track. The ID includes a new innermost layer, the insertable B-layer [9], with a mean radius of 33 mm, while the remaining three layers of the pixel system are located at mean radii of 50.5, 88.5, and 122.5 mm respectively. The coverage in the endcap region is enhanced by three disks on either side.

\(^1\) In the rest of this paper, electrons will be used to refer to both electrons and positrons.

\(^2\) The photon transverse momentum $E_T$ is defined in footnote\(^3\).

\(^3\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upward. Cylindrical coordinates ($r$, $\phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln\tan(\theta/2)$. Angular distance is defined as $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. The transverse momentum of the photon candidates is defined as $E_T = E / \cosh(\eta)$, where $E$ is the candidate energy.
of the interaction point. A semiconductor tracker consisting of modules with two back-to-back silicon microstrip sensors with small-angle stereo readout surrounds the pixel detector, providing typically eight two-dimensional hits translating to four three-dimensional measurements, per track at intermediate radii (275 mm < r < 560 mm). The outermost region of the ID (563 mm < r < 1066 mm) is covered by a transition radiation tracker (TRT) consisting of straw drift tubes filled with a gas mixture consisting of about 70% Xe, 27% CO$_2$ and 3% O$_2$,$^4$ interleaved with polypropylene/polyethylene transition radiators. The inner detector allows an accurate reconstruction and transverse momentum measurement of tracks from the primary proton–proton collision region. It also identifies tracks from secondary vertices, permitting the efficient reconstruction of photon conversions up to a radial distance of about 80 cm from the beam-line.

The solenoid is surrounded by a high-granularity lead/liquid-argon (LAr) sampling electromagnetic (EM) calorimeter with an accordion geometry. The EM calorimeter (EMC) measures the energy and the position of electromagnetic showers with |η| < 3.2. It is divided into a barrel section, covering the pseudorapidity region |η| < 1.475, and two endcap sections, covering the pseudorapidity regions 1.375 < |η| < 3.2. The transition region between the barrel and the endcaps, 1.37 ≤ |η| < 1.52, has a large amount of material upstream of the first active calorimeter layer and is not considered for the precision photon measurements reported in this paper. The EM calorimeter is composed, for |η| < 2.5, of three sampling layers, longitudinal in shower depth. The first layer has a thickness of about 4.4 radiation lengths (X$_0$) at η = 0.$^5$ In the ranges |η| < 1.4 and 1.5 < |η| < 2.4, the first layer is segmented into high-granularity strips in the η direction, with a typical cell size of 0.003 × 0.0982 in Δη × Δϕ in the barrel. For 1.4 < |η| < 1.5 and 2.4 < |η| < 2.5 the η segmentation of the first layer is coarser, and the cell size is Δη × Δϕ = 0.025 × 0.0982.

The fine η granularity of the strips is sufficient to provide, for transverse momenta up to O(100 GeV), an event-by-event discrimination between single-photon showers and two overlapping showers originating from the decays of neutral hadrons, mostly π$^0$ and η mesons in jets, in the fiducial pseudorapidity region |η| < 1.37 or 1.52 ≤ |η| < 2.37. The second layer has a thickness of about 16 X$_0$ at η = 0, and a granularity of 0.025 × 0.0245 in Δη × Δϕ. It collects most of the energy deposited in the calorimeter by photon and electron showers. The third layer has a granularity of 0.05 × 0.0245 in Δη × Δϕ and a depth of about 2 X$_0$ at η = 0. It is used to correct for leakage of high-energy showers into material beyond the EM calorimeter. In front of the accordion calorimeter, a thin presampler (PS) layer, covering the pseudorapidity interval |η| < 1.8, is used to correct for energy loss upstream of the calorimeter. The PS consists of an active LAr layer with a thickness of 1.1 cm (0.5 cm) in the barrel (endcap) and has a granularity of Δη × Δϕ = 0.025 × 0.0982.

The material upstream of the PS has a thickness of about 2 X$_0$ for |η| < 0.6. In the region 0.6 < |η| < 0.8 this thickness increases linearly from 2 X$_0$ to 3 X$_0$. For 0.8 < |η| < 1.8 the material thickness is about or slightly larger than 3 X$_0$, with the exception of the transition region between the barrel and the endcaps and the region near |η| = 1.7, where it reaches 5–6 X$_0$. A sketch of a the EM calorimeter’s longitudinal and lateral segmentation around η = 0 is shown in Fig. 1.

The hadronic calorimeter (HCAL) surrounds the EM calorimeter. It consists of a steel/scintillator tile calorimeter in the central region (|η| < 1.7), and LAr sampling calorimeters with copper and tungsten absorbers in the endcap (1.5 < |η| < 3.2) and forward (3.1 < |η| < 4.9) regions. The muon spectrometer (MS) surrounds the calorimeters. It consists of three large superconducting air-core toroid magnets, each with eight coils, a system of precision tracking chambers (|η| < 2.7), and fast tracking chambers (|η| < 2.4) for triggering.

A two-level trigger system, custom hardware followed by a software-based level, is used for online event selection and to reduce the event rate to about 1 kHz for offline reconstruction and storage [10]. To reduce the data acquisition rate of low-threshold triggers, used for collecting various control samples, prescale factors N can be applied to each trigger, such that only one in N events passing the trigger causes an event to be accepted at that trigger level.

3 Photon reconstruction and identification

3.1 Photon reconstruction

The interactions of photons and electrons with the ATLAS EMC produce similar electromagnetic showers, depositing a significant amount of energy in a restricted number of neighbouring calorimeter cells. As photons and electrons have very similar signatures in the EMC, their reconstruction proceeds in parallel. The reconstruction of electron candidates, including a dedicated, cluster-seeded track-finding algorithm to increase the efficiency for the reconstruction of low-momentum electron tracks, is described in Ref. [11]. The reconstruction of unconverted and converted photons in Run 2 data collected in 2015 and 2016 is largely unchanged from the reconstruction used Run 1 and described in Ref. [7], and can be summarised as follows:

---

$^4$ During part of the 2016 data-taking some TRT layers were filled with argon instead of xenon.

$^5$ The depth of the calorimeter layers varies with η, generally increasing at higher pseudorapidity.
A sliding window with a size of $3 \times 5$ in units of $\Delta \eta \times \Delta \phi = 0.025 \times 0.0245$, corresponding to the granularity of the EM calorimeter middle layer, is used to search for electromagnetic cluster seeds as longitudinal towers with total cluster transverse energy above 2.5 GeV. The clusters are then formed around the seeds using a clustering algorithm [12] that allows for removal of duplicates. The cluster kinematics are reconstructed using an extended window depending on the cluster position in the calorimeter. The efficiency of the cluster search in simulation is higher than 99% for photons with $E_T > 20$ GeV.

Tracks reconstructed in the inner detector are loosely matched to seed clusters. Seed clusters that pass loose shower shape requirements in hadronic leakage and energy distribution in $\eta$ are used to create regions-of-interest (ROIs), within which standard track pattern reconstruction [13] is first performed. If the pattern recognition fails for a silicon track seed that is within an ROI, a modified pattern reconstruction algorithm is performed based on a Kalman filter formalism [14], allowing for up to 30% energy loss at each material intersection. Track candidates are then fitted with the global $\chi^2$ fitter [15], allowing for additional energy loss in cases where the standard track fit fails. Tracks with silicon hits loosely matched to EM clusters are re-fitted using a Gaussian-sum filter (GSF) fitter [16], a non-linear generalization of the Kalman filter, for improved track parameter estimation.

The loosely-matched tracks serve as input to the conversion vertex reconstruction. Tracks with silicon hits (referred to as $Si$ tracks) and tracks reconstructed only in the TRT (referred to as $TRT$ tracks) are used for the conversion reconstruction. Two-track conversion vertices are reconstructed from two tracks forming a vertex consistent with that of a massless particle, while single-track vertices are built from tracks without hits in the innermost sensitive layers. To increase the converted photon purity, the tracks used to build conversion vertices must
generally have a high probability to be electron tracks as determined by the TRT [17], especially for single-track vertices and conversion vertices constructed from TRT tracks. If there are multiple conversion vertices matched to a cluster, double-track conversions with two silicon tracks are preferred over other double-track conversions, followed by single-track conversions. Within each category, the vertex with the smallest conversion radius is preferred.

- An arbitration relying on the properties of the tracks and conversion vertices matched to a given electromagnetic cluster is performed, to determine whether an object is reconstructed as an electron, a converted or an unconverted photon, or both as an electron and a photon object in the ambiguous cases: clusters to which neither a conversion vertex candidate nor any track has been matched during the electron reconstruction are considered unconverted photon candidates; clusters matched to a conversion vertex candidate are considered converted photon candidates; converted photon candidates that are also reconstructed as electrons, the electron track is evaluated against the properties of the track(s) originating from the conversion vertex candidate matched to the same cluster; unconverted photon candidates are recovered from reconstructed electron candidates depending on the track hits, momentum and $E/p$ properties. This procedure is discussed in details in Ref. [7].

Since the analysis reported in Ref. [7], the reconstruction of converted photon candidates has undergone a few changes to improve both reconstruction efficiency and rejection of fake converted photons. Improvements are made especially in track reconstruction and conversion vertex building for TRT tracks:

- the reconstruction of tracks using the outside-in tracking algorithm [13] is restricted to ROIs defined by electromagnetic clusters;
- the efficiency for the reconstruction of double-track TRT conversions is improved by allowing the reconstruction of TRT tracks which share up to 70% of hits;
- the fraction of unconverted photons reconstructed as double- or single-track TRT conversions is reduced by tightening the requirements on the TRT tracks: the tracks are required to have at least 25% precision hits (a precision hit is defined as a hit with a track-to-wire distance within 2.5 times the drift circle uncertainty [17]);
- the determination of the probability of a track to be an electron using high-threshold hit information from the TRT is improved, taking into account the TRT occupancy as a measure of the pile-up level of an event [18].

With this improved reconstruction of converted photons, the efficiency to reconstruct a true converted photon is higher than 70% for simulated photons with true $E_T > 20$ GeV. This efficiency is higher at lower $\mu$ values, being greater than 75% at $\mu \sim 0$ and decreasing to about 65% at $\mu \sim 60$. The fraction of true unconverted photons in simulation that are erroneously reconstructed as converted photons is below 9% for $\mu = 60$, and decreases with $\mu$ to become smaller than 1% for $\mu < 24$.

The photon energy measurement is performed using information from the calorimeter. The photon energy calibration, which accounts for upstream energy loss and both lateral and longitudinal leakage, is based on the same procedure developed in Run 1 [19], but specifically tuned to the Run 2 detector configuration [20] in order to account for a different amount of material upstream of the EMC, due to the presence of the insertable B-layer. The energy of the electromagnetic clusters associated with the photon candidates is corrected in subsequent steps using a combination of simulation-based and data-driven correction factors, with the calibration regression being separately optimised for converted and unconverted photons. The uniformity corrections and the intercalibration of the longitudinal calorimeter layers are unchanged from to those determined in Run 1.

In the following the photon $E_T$ is computed from the photon cluster’s calibrated energy $E$ and the pseudorapidity $\eta$ of the barycentre of the cluster in the second layer of the EMC as $E_T = E / \cosh(\eta)$.

### 3.2 Photon identification

The identification of photon candidates in ATLAS relies on rectangular cuts using calorimetric variables which deliver good separation between prompt-photons and fake signatures from non-prompt-photons originating from the decay of neutral hadrons in jets, or QCD jets depositing a large energy fraction in the EMC. Such variables, listed in Table 1 and depicted in Fig. 2 with their respective definitions, characterize the lateral and longitudinal electromagnetic shower development in the EMC and the shower leakage fraction in the HCAL. Prompt-photons typically produce narrower energy deposits in the EMC and have smaller leakage to the HCAL compared to background photons from jets. Additionally, background candidates from $\pi^0 \rightarrow \gamma\gamma$ decays are often characterized by two separate local energy maxima in the finely segmented strips of the EMC first layer.

---

6 A true converted photon is defined as a photon undergoing a conversion into an electron–positron pair within a distance $r < 80$ cm from the interaction point.

7 The $R_{\text{had}}$ variable was initially used by ATLAS along the whole $\eta$ acceptance range [21]; however, the use of the normalised total hadronic energy $R_{\text{had}}$ is found to be more effective in discriminating hadronic showers in the region $0.8 < |\eta| < 1.37$ [22].
Table 1 Discriminating variables used for loose and tight photon identification

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Name</th>
<th>loose</th>
<th>tight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>$</td>
<td>\eta</td>
<td>&lt; 2.37$, with $1.37 \leq</td>
<td>\eta</td>
</tr>
<tr>
<td>Hadronic leakage</td>
<td>Ratio of $E_T$ in the first sampling layer of the hadronic calorimeter to $E_T$ of the EM cluster (used over the range $</td>
<td>\eta</td>
<td>&lt; 0.8$ or $</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>Ratio of $E_T$ in the hadronic calorimeter to $E_T$ of the EM cluster (used over the range $0.8 &lt;</td>
<td>\eta</td>
<td>&lt; 1.37$)</td>
<td>$R_{\text{had}}$</td>
</tr>
<tr>
<td>EM middle layer</td>
<td>Ratio of the energy in $3 \times 7 \eta \times \phi$ cells over the energy in $7 \times 7$ cells centered around the photon cluster position</td>
<td>$R_{\text{eta}}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Lateral shower width, $\sqrt{\left(\sum E_i \eta_i^2 / \sum E_i\right) - \left(\sum E_i \eta_i / \sum E_i\right)^2}$, where $E_i$ is the energy and $\eta_i$ is the pseudorapidity of cell $i$ and the sum is calculated within a window of $3 \times 5$ cells</td>
<td>$w_{\text{eta}}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Ratio of the energy in $3 \times 3 \eta \times \phi$ cells over the energy of $3 \times 7$ cells centered around the photon cluster position</td>
<td>$R_{\phi}$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>EM strip layer</td>
<td>Lateral shower width, $\sqrt{\left(\sum E_i(i-i_{\text{max}})^2 / \sum E_i\right)}$, where $i$ runs over all strips in a window of $3 \times 2 \eta \times \phi$ strips, and $i_{\text{max}}$ is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit</td>
<td>$w_{s3}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Total lateral shower width $\sqrt{\left(\sum E_i(i-i_{\text{max}})^2 / \sum E_i\right)}$, where $i$ runs over all strips in a window of $20 \times 2 \eta \times \phi$ strips, and $i_{\text{max}}$ is the index of the highest-energy strip measured in the strip layer</td>
<td>$w_{s\text{tot}}$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips</td>
<td>$f_{\text{side}}$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima</td>
<td>$\Delta E_i$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies</td>
<td>$E_{\text{ratio}}$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ratio of the energy in the first layer to the total energy of the EM cluster</td>
<td>$f_1$</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Two reference sets of cuts—loose and tight—are specifically defined for the $pp$ data collected at $\sqrt{s} = 13$ TeV in 2015 and 2016. While the same set of discriminating variables employed by the photon identification in Run 1 [7] are used, the selection cuts are tuned to reduce the dependency of the identification efficiency on pile-up, in order to cope with the harsher Run 2 conditions. This mostly results in looser selections for converted photons, where broader electromagnetic calorimeter in a cluster of properties $C$, identifying the number and/or properties of selected cells. $E_i$ is the energy in the $i$-th cell, $\eta_i$ the pseudorapidity centre of that cell.
shower shapes tend to be more affected by the larger number of interactions per beam bunch crossing.

The loose selection is based on shower shapes in the second layer of the electromagnetic calorimeter and on the energy deposited in the hadronic calorimeter. The tight selections add information from the finely segmented strip layer of the calorimeter, and are separately optimised for unconverted and converted photons, to account for the generally broader lateral shower profile of the latter. The thresholds of the selection criteria are different in seven intervals of the reconstructed photon $|\eta|$ (0.0–0.6, 0.6–0.8, 0.8–1.15, 1.15–1.37, 1.52–1.81, 1.81–2.01, 2.01–3.7) to account for the calorimeter geometry, and for different effects on the shower shapes from the material upstream of the calorimeter.

The distributions of the discriminating variables for both the prompt and background photons are affected by additional soft $pp$ interactions that may accompany the hard-scattering collision, referred to as in-time pile-up, as well as by out-of-time pile-up arising from bunches before or after the bunch where the event of interest was triggered. Pile-up collisions result in the presence of low-$E_T$ activity in the detector, including energy deposits in the EMC. A greater number of superimposed $pp$ events, $\mu$, would generally broaden the photon shower shapes because of these additional energy deposits in the calorimeter, thus resulting in a lower identification efficiency for larger $\mu$ values, as discussed in Sect. 5.5.

### 3.3 Photon isolation

The identification efficiencies presented in this paper are measured for photon candidates passing an isolation requirement, similar to those applied to reduce hadronic background in prompt-photon cross-section measurements [1], $H \rightarrow \gamma\gamma$ measurements [2, 24], or searches for exotic processes with photons [4–6]. The choice of a specific isolation criterion is determined by the actual physics analysis, since it depends on the different background sources, the signal-to-background ratio, and the background rejection needs. On the other hand, it is shown in Sect. 5.6 that the photon identification efficiency does not show a significant dependence on the chosen isolation criterion. Additionally, it is shown in Sect. 5.4 that the corrections meant to address the mismodelling by simulation of the photon identification efficiency measured in data (scale factors) do not depend, within uncertainties, on the physics process used to measure it and the isolation criterion of choice.

The definition of photon isolation in ATLAS is based on the transverse energy in a cone with angular size $\Delta R$ around the direction of the photon candidate. This transverse energy is characterized by two quantities, the calorimeter isolation and the track isolation. The calorimeter isolation $E^{\text{iso}}_T$ is obtained from the sum of transverse energies of topological clusters [12] in the calorimeters, after subtracting on an event-by-event basis the energy deposited by the photon candidate and the contribution from the underlying event and pile-up. This uses the method described in Refs. [25–27] and is discussed in more detail in Ref. [7]. The track isolation $p^{\text{iso}}_T$ is obtained by summing the transverse momenta of all the tracks with transverse momentum above 1 GeV and having a distance of closest approach to the primary vertex [28] along the beam axis $|z_0 \sin \theta| < 3$ mm, and excluding the tracks associated with photon conversions.

ATLAS analyses selecting final-state photons use a variety of isolation selection criteria. The most commonly adopted are a loose isolation requirement, based on both the calorimeter isolation and the track isolation, in both cases computed in a cone with $\Delta R = 0.2$:

\[
E^{\text{iso}}_T|_{\Delta R < 0.2} < 0.065 \cdot E_T \quad \text{and} \quad p^{\text{iso}}_T|_{\Delta R < 0.2} < 0.05 \cdot E_T;
\]

a tight isolation requirement, based on the calorimeter isolation computed in a cone with $\Delta R = 0.4$, and the track isolation computed in a cone with $\Delta R = 0.2$:

\[
E^{\text{iso}}_T|_{\Delta R < 0.4} < 0.022 \cdot E_T + 2.45 \text{ GeV} \quad \text{and} \quad p^{\text{iso}}_T|_{\Delta R < 0.2} < 0.05 \cdot E_T;
\]

an alternative version of the tight isolation requirement (calorimeter-only tight), based only on the calorimeter isolation:

\[
E^{\text{iso}}_T|_{\Delta R < 0.4} < 0.022 \cdot E_T + 2.45 \text{ GeV};
\]

and a legacy isolation requirement, requiring a fixed selection on the calorimeter isolation:

\[
E^{\text{iso}}_T|_{\Delta R < 0.4} < 4 \text{ GeV}.
\]

The data/MC corrections to the electromagnetic shower shape variables discussed in Sect. 4 are computed using photon candidates satisfying the calorimeter-only tight isolation criterion. The measurements of photon identification efficiency reported in Sect. 5 are performed for isolated photon candidates meeting the loose criterion, apart from the measurement using radiative $Z$ decays, which is nominally performed for the tight criterion, and repeated using the loose isolation and the calorimeter-only tight isolation criteria in order to evaluate the potential dependency of the identification efficiency on the photon isolation (Sect. 5.6). The measurement of the electron-to-photon fake rate discussed in Sect. 6 is performed for isolated photon candidates satisfying the loose criterion, and its dependency on the isolation selection verified on candidates meeting the tight and calorimeter-only tight criteria. The conversion reconstruction efficiency reported in Sect. 7 is per-
formed for isolated photon candidates satisfying the legacy criterion.

4 Collision and simulated data samples

The measurements presented in this paper use proton–proton (pp) collisions at √s = 13 TeV recorded by the ATLAS detector in 2015 and 2016 during the LHC Run 2. The data are required to pass good quality requirements on the detector performance and object reconstruction, leading to 36.1 fb⁻¹ of integrated luminosity. The inclusive photon measurement discussed in Sect. 5.3 relaxes the requirement on the performance of the ATLAS muon spectrometer, and uses 36.7 fb⁻¹ of integrated luminosity. In these datasets the mean number of interactions per bunch crossing is 13.5 in 2015 data and 24.9 in 2016 data.

Two of the methods used to measure the photon identification efficiency described in Sect. 5 rely on the use of Z boson decays into electron–positron pairs Z → e⁺e⁻ and on Z boson radiative decays Z → ℓℓγ (ℓ = e, μ); these events are selected in data collected with the lowest-threshold unprescaled lepton triggers. The single-electron trigger has a transverse momentum threshold of 24 GeV in 2015 and in most of 2016, and of 26 GeV in the last data-taking period of 2016; the single-muon trigger uses a transverse momentum threshold of 24 GeV in 2015 and in most of 2016, and of 10 GeV in the last data-taking period of 2016; the dimuon trigger has a transverse momentum threshold of 8 GeV in both 2015 and 2016. The dielectron trigger has a transverse momentum threshold of 9 GeV in 2015 and in most of 2016, and of 10 GeV in the last data-taking period of 2016; the dimuon trigger has a transverse momentum threshold of 8 GeV in both 2015 and 2016. The third method for measuring the photon identification efficiency uses a sample of single-photon candidates, selected in data from events collected with single-photon triggers with loose identification requirements and large prescale factors, thus exploiting only a fraction of the total luminosity. The lowest transverse momentum threshold of these single-photon triggers is 10 GeV in both 2015 and 2016.

Simulated MC samples of prompt-photon production were generated with PYTHIA8 [29,30]. Such samples include the leading-order γ + jet events from qg → qγ and qq̅ → gγ hard scattering, as well as prompt-photons from quark fragmentation in QCD dijet events. Samples of background photons in jets were produced by generating with PYTHIA8 all tree-level Z → 2 QCD processes, removing γ + jet events from quark fragmentation. Simulated samples of Z → ℓℓγ (ℓ = e, μ) events were generated with SHERPA [31] or with POWHEG-BOX [32,33] interfaced to PHOTOS [34] for the modelling of QED final-state radiation and to PYTHIA8 for showering, hadronization and modelling of the underlying event. Z(→ ℓℓ)+jet MC events were generated for both ℓ = e and ℓ = μ with SHERPA. All MC samples were processed through a full simulation of the ATLAS detector response [35] using GEANT4 [36]. Pile-up pp interactions in the same and nearby bunch crossings are included in the simulation. MC samples were reweighted to reproduce the distribution of μ observed in data.

4.1 Data-driven corrections to shower shapes in simulated data

The distributions of the photon transverse shower shapes in the ATLAS MC simulation do not perfectly describe those observed in data. While these distributions in simulation are rather similar in shape to those found in the data, small systematic differences in their average values are observed, pointing to a mismodelling in MC simulation of the lateral profile development of the electromagnetic showers, while, overall, the longitudinal electromagnetic shower profiles are well described. These differences between data and MC distributions are measured and parameterised as simple shifts to be applied to the MC-simulated values to align with the distributions observed in data. The shifts are calculated by minimizing the χ² between the data and the shifted MC distributions of photon candidates satisfying the tight identification criteria and the calorimeter isolation requirement described in the previous section. The shifts are computed in intervals of the reconstructed photon pseudorapidity and transverse momentum. The pseudorapidity intervals are the same as those used to define the photon selection criteria, while the E_T bin boundaries are 8, 15, 20, 25, 30, 40, 50, 60, 80, 100, 250, and 1000 GeV. Photon candidates from Z → ℓℓγ events are used for E_T < 50 GeV, while candidates from single-photon events are used for E_T > 50 GeV. The correction factors are measured from both types of events in the overlapping E_T region around 50 GeV and found to be compatible.

The typical size of the correction is 10% of the root-mean-square of the distribution of the corresponding variable in data. The corresponding correction to the photon efficiency predicted by simulation varies with pseudorapidity between −10 and −5% for photon transverse momenta close to 10 GeV, and approaches zero for transverse momenta above 50 GeV. Examples of the simulated discriminating variable distributions before and after corrections, for converted and unconverted photon candidates originating from Z boson radiative decays, are shown in Fig. 3. For comparison, the distributions observed in data for candidates passing the Z boson radiative decay selection illustrated in Sect. 5.1, are also shown. Improved agreement between the shower shape distributions in data and simulation after applying such corrections is clearly visible. Residual discrepancies are observed in the tail of the distributions. Their effect on the MC description of the photon identification efficiency is addressed with data/MC scale factors. Similarly, while there
Fig. 3 Distributions of the calorimetric discriminating variables $R_\eta$ and $w_{s,3}$ for converted and unconverted photon candidates with $E_T \in [10, 50]$ GeV and $|\eta| < 2.37$ (excluding $1.37 \leq |\eta| < 1.52$) selected from $\ell^+ \ell^- \gamma$ events (black dots). The distributions for true photons from simulated $Z \rightarrow \ell \ell \gamma$ events are also shown, for the uncorrected simulation (dashed red line) and the simulation corrected by the average shift between data and simulation distributions (solid blue line). The definition of the plotted variables is given in Table 1 and Fig. 2.

are correlations between variables and this might not be perfectly described by simulation, these simple corrections do not attempt to address such a potential mismodelling, whose impact would instead be collectively taken into account with the same data/MC scale factors mentioned above.

In all the analyses described in Sect. 5 the reference simulated samples are modified with the correction factors described above, while in Sect. 5.4 the measured photon identification efficiencies are compared with the values in both the uncorrected and corrected MC samples.

5 Measurements of the photon identification efficiency

The efficiency $\varepsilon_{ID}$ of the tight photon identification criterion described in Sect. 3.2 is measured in data using three methods:

- **Radiative $Z$ decays:** this method uses a clean sample of low-energy photons obtained from $Z \rightarrow \ell \ell \gamma$ decays ($\ell = e, \mu$). This allows measurements of $\varepsilon_{ID}$ from $E_T = 10$ GeV, below which photons are not reconstructed, to $E_T \sim 100$ GeV, beyond which event yields are insufficient. The method is described in detail in Sect. 5.1 below.

- **Electron extrapolation:** this method uses a sample of electromagnetic showers from electrons originating from $Z \rightarrow ee$ decays, identified using a tag-and-probe method. These showers are modified so that their shape information matches the properties of photon showers, and used to measure $\varepsilon_{ID}$ in the region $25 < E_T < 150$ GeV where sufficient numbers of $Z \rightarrow ee$ electron candidates are available. The method is described in detail in Sect. 5.2 below.
• **Inclusive photons:** this method uses an inclusive photon sample collected using single-photon triggers. The efficiency of a tight track-based isolation criterion is used to obtain the fraction of prompt-photons in the full sample and in the subsample satisfying the tight identification criterion, from which a measurement of $\varepsilon_{\text{ID}}$ can be derived. The measurement is performed over a wide kinematic range spanning $25 \text{ GeV} < E_T < 1.5 \text{ TeV}$. At low transverse energy this is limited by the prescaling of single-photon triggers below $E_T = 140 \text{ GeV}$, and at high transverse energy by limited event yields. The method is described in detail in Sect. 5.6 below.

The efficiencies are reported in each case for converted and unconverted photons separately, since their distributions for the discriminating variables listed in Table 1 typically differ due to differences in electromagnetic shower development. Efficiencies are measured in a two-dimensional interval grid in photon $E_T$ and $|\eta|$ with boundaries at 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 80, 100, 125, 150, 175, 250 and 1500 GeV and 0, 0.6, 1.37, 1.81 and 2.37 respectively; each method covers only a subspace of this region, as described above.

The various methods provide measurements covering overlapping kinematic ranges, where the measured values of the photon identification efficiency can be compared. Because of the different compositions of the prompt-photon samples used to measure the efficiencies (i.e. the varying fraction of photons originating from fragmentation processes), the three methods are not necessarily expected to provide the same efficiency values for the same photon. On the other hand, one expects any residual mismodelling of the photon identification efficiency by MC simulation to be independent of the physics process used to measure the efficiency, since it would mostly be due to an imperfect modelling of the detector response. For this reason, the corrections for this mismodelling are expected to be universal, and the correction values obtained from the various methods are therefore combined for increased precision (see Sect. 5.4).

### 5.1 Photons from $Z$ boson radiative decays

Radiative $Z$ decays are selected by requiring the presence of a photon candidate and an opposite-charge pair of electron or muon candidates. Photon candidates are required to have a transverse momentum $E_T > 10 \text{ GeV}$ and a pseudorapidity in the range $|\eta| < 1.37$ or $1.52 \leq |\eta| < 2.37$. The loose isolation selection described in Sect. 3.3 is applied for the nominal results, while the effect of applying alternative isolation criteria is studied in Sect. 5.6. No other selection is applied to the photon, in order to avoid biases due to the photon selection in the efficiency measurement.

The events are selected using unprescaled single-lepton and dilepton triggers with the lowest transverse momentum threshold.

Muons candidates are required to be reconstructed from hits in both the MS and the ID [37]. They must fulfill the conditions $E_T > 10 \text{ GeV}$ and $|\eta| < 2.5$, and the impact parameters of their track must be compatible with originating from the primary event vertex. They must fulfill the medium identification criterion [37], which is based on the overall quality of the track fit and the compatibility of track parameters measured in the ID and the MS, with selection cuts chosen to be 99% efficient. They must also satisfy the loose isolation criterion [37], defined similarly to those described for photons in Sect. 3.3.

Electron candidates are required to have $E_T > 10 \text{ GeV}$ and $|\eta| < 2.47$, excluding the barrel–endcap transition region, and their track must fulfill loose impact parameter selections. They are required to satisfy the medium identification criterion [11], which relies mainly on information about the shape of the associated cluster in the electromagnetic calorimeter and transition radiation emission in the TRT. Electrons are also required to meet the loose isolation criterion [11] similar to the one described in Sect. 3.3.

Radiative decays are selected by requiring $40 < m_{\ell\ell} < 83 \text{ GeV}$ and $80 < m_{\ell\ell\gamma} < 100 \text{ GeV}$, were $m_{\ell\ell}$ is the invariant mass of the dilepton system and $m_{\ell\ell\gamma}$ that of the two leptons and the photon. These selections are meant to isolate the radiative decays from events where the photon originates from initial state radiation. Distributions of the $m_{\ell\ell}$ and $m_{\ell\ell\gamma}$ quantities for the electron channel are shown in the left panel of Fig. 4. Separations $\Delta R > 0.2$ and $\Delta R > 0.4$ are required between the photon and the closest muon and electron candidate respectively, in order to avoid biases in the photon shower shape and isolation variables. About 170,000 unconverted photons and 60,000 converted photons are found to pass all selections in the $\mu\mu\gamma$ channel, and about 90,000 and 30,000 respectively in the $ee\gamma$ channel.

A small background contamination occurs due to $Z \rightarrow \ell\ell$ decays accompanied by a jet which is misidentified as a photon, particularly if the photon candidate has low transverse momentum. The size of these contributions is estimated using a fit of the $m_{\ell\ell\gamma}$ shape in data over the range $65 < m_{\ell\ell\gamma} < 105 \text{ GeV}$. The model uses signal and background shapes obtained in simulated samples of $Z \rightarrow \ell^+\ell^-\gamma$ and $Z \rightarrow \ell^+\ell^- + \text{jet}$ production respectively, described in Sect. 4. The signal and background yields are determined by the fit, an example of which is shown in the right panel of Fig. 4 for the electron channel. In the region $10 < E_T < 25 \text{ GeV}$, the purity of the selection, defined as the ratio of the $Z \rightarrow \ell\ell\gamma$ yield to the total sample size, is measured to be about 82% in the $\mu\mu\gamma$ channel and 86% in the $ee\gamma$ channel. After applying photon identification cuts, this purity becomes about 94% in the $\mu\mu\gamma$ channel and 97% in

---

8 Called Matrix Method in Ref. [7].
the \( ee \) channel. The photon identification efficiency is then computed as \( \varepsilon_{\text{ID}} = (P_{\text{pass}}/N_{\text{data}}) / (P_{\text{total}}/N_{\text{data}}) \), where \( N_{\text{pass}} \) and \( N_{\text{data}} \) are the numbers of events (purities) in the full sample and the subset in which the photon passes the identification cuts respectively. The computation is performed separately for the \( \mu \mu \gamma \) and \( ee \) channels in \( (E_{T}, |\eta|) \) intervals ranging up to \( E_{T} = 100 \text{ GeV} \). For \( E_{T} > 25 \text{ GeV} \), the purity in the full sample is above 96% and no correction is applied, but a systematic uncertainty is included to account for the residual background level. The \( \mu \mu \gamma \) and \( ee \) results are combined as discussed in Sect. 5.4.

The following sources of systematic uncertainties are considered:

- A closure test is made by performing the measurement on a sample consisting of known fractions of simulated signal and background events. This test is only performed for \( E_{T} < 25 \text{ GeV} \), because of the limited number of MC events at higher transverse momenta. Deviations from the true identification efficiency are included as systematic uncertainties. Their value is below 1% in all regions.
- An uncertainty in the level of background contamination is assessed by computing \( \varepsilon_{\text{ID}} \) with and without accounting for the background component, and using the differences between the two results in each region as a systematic uncertainty. Its values are less than 2.5%, except in the region \( E_{T} < 15 \text{ GeV} \) where they are as large as 8%.
- An uncertainty in the description of the detector in simulation is assessed by using an alternative geometry with additional inactive material in front of the calorimeter when obtaining the simulated signal distribution. The amount of additional material is chosen to be compatible with the measurements performed using Run 1 data [19]. The determination of \( \varepsilon_{\text{ID}} \) is repeated with this configuration, and the relative changes in the results for each region are counted as systematic uncertainties. Their values are typically below 2%, but up to 5% in the endcap and negligible for \( E_{T} > 25 \text{ GeV} \).

- Similarly to the above, the generator used in the signal simulation is changed from SHERPA to POWHEG-BOX. The impact of this change on the computed \( \varepsilon_{\text{ID}} \) is typically 3% or less except for \( E_{T} < 15 \text{ GeV} \) where it is as large as 10%, and is included as a systematic uncertainty.

The statistical uncertainty is obtained from the \( m_{\ell\ell\gamma} \) shape fit. It remains typically below 1% for \( E_{T} < 40 \text{ GeV} \) but rises to about 5% at 80 GeV. The total uncertainty reaches 5–15% for \( E_{T} < 15 \text{ GeV} \), about 5% for \( 15 < E_{T} < 25 \text{ GeV} \), 1% for \( 25 < E_{T} < 40 \text{ GeV} \) and then follows a rise driven by the statistical uncertainty. Results are shown in Figs. 7 and 8.

5.2 Electron extrapolation

The electron extrapolation method uses a clean sample of electron candidates from \( Z \rightarrow ee \) decays. The distributions of the shower shapes associated with electron candidates are then modified with a Smirnov transform, estimated from simulation and discussed below, to reproduce those associated with photon candidates.

Electrons are selected using a tag-and-probe method, in order to avoid selection biases in the electron shower shape distributions: most of the selections are applied to one of the electrons (the tag), while only a loose selection is applied to the other electron (the probe) from which shower shape
distributions are then obtained. The events are required to pass a single-electron trigger selection, and the trigger object must match the tag electron. Both electrons are required to have $E_T > 25$ GeV and $|\eta| < 2.37$, excluding the transition region between barrel and endcap calorimeters. The tag electron is required to pass the right identification requirement [11], while the probe electron is only required to have a track with at least seven track hits in the semiconductor tracker and at least one hit in the pixel detector. However, in order to match the photon selection described below, the probe electron is required to pass the loose isolation requirement as described in Sect. 3.3. There must also be no more than one jet (the one reconstructed from the energy deposited in the calorimeter by the electron) with $E_T > 20$ GeV and within $\Delta R = 0.4$ of the probe electron. The two electrons are required to have opposite charges and an invariant mass in the range $70 < m_{ee} < 110$ GeV.

The shower-shape variables listed in Table 1 are obtained from the electromagnetic clusters of the electron in the same way as for photons. Differences between the distributions of photon and electron shower shapes are corrected using simulation. A set of simulated probe electrons is selected by applying the selection above to $Z \rightarrow ee$ simulated signal events. A set of simulated photons is selected in single-photon simulated samples by applying the same $E_T$, $|\eta|$ and isolation selections as described above for the probe electrons, and requiring the photon candidate to be matched to a true photon object.

In each case, the distribution of each shower-shape variable $x_i$ is then obtained, with shifts applied to the photon distributions as described in Sect. 4.1, and a similar procedure applied to the electron distributions. Smirnov transforms $^9 S_i$ are defined by the relations:

$$x_{y,i} = S_i(x_{e,i}) \equiv F_{y,i}^{-1}(F_{e,i}(x_{e,i})),$$

where $F_{y,i}$ and $F_{e,i}$ are the cumulative distributions of $x_i$ for simulated photons and electrons respectively [38]. The transformations are such that for an input $x_{e,i}$ following the electron distribution, the output $x_{y,i}$ follows the photon distribution. They are therefore applied to the shower shape properties of data electrons in order to match the expected photon profiles. Transformations are separately computed for converted and unconverted photons for all discriminating variables. An example of the procedure is shown in Fig. 5 for the $R_{#phi}$ variable for converted photons.

The transformed variables are then used to apply the photon identification selection to the electrons in the same way as for photon candidates. The ratio of the number of transformed electron candidates passing the photon selections to the total number of electron candidates is used to estimate $\varepsilon_{ID}$ separately in each ($E_T$, $|\eta|$) bin, and separately for converted and unconverted photons.

The data sample includes a small contamination from processes where the probe is a fake electron, mainly from multijet and W+jets production. The size of this contamination is estimated in each $E_T$ and $|\eta|$ bin using a shape fit of the $m_{ee}$ variable over the range $70 < m_{ee} < 110$ GeV with a signal and a background component. The shape of the signal component is obtained from the simulated electron sample. The background shape is obtained from the data by requiring that the probe electron fail at least two of the loose electron identification selections of the cut-based selection defined in Ref. [11], as well as the calorimeter-based requirement of the loose isolation selection. The background contribution is estimated separately in each analysis bin and subtracted from both the numerator and denominator in the computation of $\varepsilon_{ID}$.

The following sources of systematic uncertainties are considered:

- A closure test of the Smirnov transform procedure is performed by comparing photon identification efficiencies for transformed MC electrons and MC photons. This check accounts in particular for a difference in the correlations of shower-shape variables between photons and electrons, which are not modified by the per-variable Smirnov transforms. The effect is found to be at most 1% for converted photons and 2% for unconverted photons.
- An uncertainty is assigned to the background subtraction technique by repeating the measurement while using the range $80 < m_{ee} < 100$ GeV for the template fit. The difference between this result and the nominal result is used as a systematic uncertainty.
- An uncertainty is assigned because of the difference in the fraction of converted photons between data and simulation, which impacts the simulated shower shapes used to derive the Smirnov transforms. The fraction of true converted photons in the simulated photon sample is varied by $\pm 10\%$, an amplitude which covers the differences between data and simulation reported in Sect. 7; the resulting change in $\varepsilon_{ID}$ is used to estimate the uncertainty. The effect is of the order of 0.2% or less, and up to about 1% in the first endcap $|\eta|$ bin.
- As described in Sect. 4.1, shifts are applied to simulated shower shape distributions to align them with those in data. These do not, however, capture the full difference between data and simulation if the shapes cannot be reconciled by simple shifts. The impact of the residual differences is accounted for by defining for each variable a range of shift values such that, for any value of the variable, the data distribution can be locally matched to the

---

$^9$ The Smirnov transform is also known as the inverse probability integral transform [38].
Fig. 5 Illustration of the Smirnov transform technique applied to photon and electron shower shapes. $R_\phi$ is chosen as an example of a shower shape which differs notably between electrons and photons. The $R_\phi$ distribution in each sample (top left) is used to calculate the corresponding cumulative distributions (top right). From the two cumulative distributions, a Smirnov transformation can be derived (bottom left). Applying the transformation leads to an $R_\phi$ distribution which closely matches the photon distribution. The definition of the $R_\phi$ is given in Table 1 and Fig. 2.

Simulated distribution by a shift belonging to the range of allowed shift values. The measurement is then repeated with the endpoints of the range replacing the nominal value of the shift for each variable. The sum in quadrature of the maximum changes relative to the nominal measurement for each variable is used as an uncertainty. The uncertainties are typically below 1% at low $E_T$. However, the relatively tight cut on the $f_{\text{side}}$ variable in the $1.81 \leq |\eta| < 2.37$ bin leads to uncertainties of about 5% for unconverted photons and 2% for converted photons.

- An uncertainty is assigned to the fraction of photons originating from fragmentation processes in the simulation. These photons are less isolated than direct photons and have broader showers, which affects the Smirnov transforms. The uncertainty is computed as the variation in $\varepsilon_{\text{ID}}$ when the number of fragmentation photons is varied by $\pm 50\%$ in simulation. The uncertainty is typically 0.3% or less, rising to 1% at high $E_T$.

- Finally, an uncertainty is assigned to account for statistical uncertainties in the simulation sample. The uncertainty is computed by iteratively resampling the simulated samples, recomputing $\varepsilon_{\text{ID}}$ for each iteration, and the uncertainty is extracted as the width of the resulting distribution. The uncertainties are typically 0.3%, and up to 0.6% at high $E_T$.

The statistical uncertainty is computed by iteratively resampling the data as described above for the simulated samples. It remains below 0.1% over the range $25 < E_T < 150$ GeV covered by this measurement. Overall, the total uncertainty reaches about 2% at low $E_T$, and is typically below 1% for $E_T > 40$ GeV. However, values of up to 5% are reached for unconverted photons in the bin $1.81 \leq |\eta| < 2.37$ due to the data–MC differences noted above in the $f_{\text{side}}$ variable. Results are shown in Figs. 7 and 8.

5.3 Inclusive photon method

This method is based on an inclusive photon sample collected by single-photon triggers. These triggers have thresholds ranging from 10 to 140 GeV and require loose photon identification selections. They are prescaled except at the 140 GeV threshold, but provide large photon datasets at high $E_T$, with a few hundred events in each bin for $E_T > 250$ GeV,
allowing efficiency measurements to be performed up to $E_T \sim 1.5$ TeV.

In addition to the trigger requirements, the photons are required to have $E_T > 25$ GeV and $|\eta| < 2.37$, excluding the region $1.37 \leq |\eta| < 1.52$, and to pass the loose isolation requirements described in Sect. 3.3. However, the purity of the sample, defined as the fraction of true photon candidates, is low, especially at low photon $E_T$, both with and without the identification cuts applied. The identification efficiency can be estimated as

$$\epsilon_{ID} = \frac{P_{pass} N_{pass}}{P_{total} N_{total}}$$

where $P_{total}$ and $P_{pass}$ are the purities in the full sample and the subset passing the tight photon identification selection respectively, and $N_{total}$ and $N_{pass}$ are the total number of photon candidates in each case. As described below, the purities are estimated using a tight isolation criterion which requires that no track with $p_T > 1$ GeV is within $0.1 < \Delta R < 0.4$ of the photon cluster, the lower bound in $\Delta R$ being introduced to avoid selecting conversion tracks.

Purities are obtained by comparing the efficiency for this selection in data before (after) tight photon identification cuts are applied, $\epsilon_{total}$ ($\epsilon_{pass}$), with reference efficiency values for the true photon component, $\epsilon_{\gamma}$ ($\epsilon_{\gamma}^{pass}$), and background component, $\epsilon_{bkg}$ ($\epsilon_{bkg}^{pass}$), as

$$P_{pass} = \frac{\epsilon_{pass} - \epsilon_{bkg}}{\epsilon_{\gamma} - \epsilon_{bkg}}.$$

(1)

The $\epsilon_{\gamma}$ are estimated from simulation, and the $\epsilon_{bkg}$ from data. The efficiency $\epsilon_{total}$ in the full sample is measured in the subset which fails the tight photon identification selection, in order to reduce the contamination from true photons. The efficiency $\epsilon_{bkg}$ after tight photon identification cuts, is similarly evaluated by inverting some photon identification cuts to reduce the contamination from true photons. The inverted cuts are chosen to be the criteria for the $w_{3,1}$, $f_{side}$, $E_\gamma$, and $E_{nT}$ quantities, which are measured in the finely segmented first layer of the calorimeter and thus expected to be largely uncorrelated with isolation. In both cases, the residual contamination from true photons is subtracted using identification and track isolation efficiencies obtained from simulation and a data-driven overall normalisation. The evolution of the tight isolation efficiencies and the sample purities as a function of photon $E_T$ is shown in Fig. 6 for unconverted photons in $|\eta| < 0.6$. The isolation efficiency for prompt-photons is nearly constant in $E_T$, while for fake photons the efficiency decreases with $E_T$ since higher-energy fake photons are typically associated with higher-energy jets, which are more likely to contain tracks. For unconverted photons, the purities $P_{total}$ before photon identification selections are found to range from about 30% at low $E_T$ to about 85% at high $E_T$. After applying the tight photon identification selection, this rises to about 50% at low $E_T$ and about 90% at high $E_T$. For converted photons the purities are lower, ranging from 20% to 60% before the identification requirement and 40% to 80% after it is applied.

Finally, since the efficiency $\epsilon$ is computed using a sample of photons which pass the loose identification selection, a correction obtained from prompt-photon simulation is applied to measure $\epsilon_{ID}$. 

![Fig. 6](image-url)
The following sources of systematic uncertainties are considered:

- A closure check on the computation of the background efficiencies $\epsilon_{\text{total}}$ and $\epsilon_{\text{pass}}$ is performed by determining their values as described above in a sample of simulated multijet production (see Sect. 4). The relative differences between these values and the true value is used as a systematic uncertainty. The uncertainty reaches 18% at low $E_T$ but is about 3% at 50 GeV and below 1% at high $E_T$.

- An uncertainty due to the tight isolation requirement is evaluated by changing the size of the isolation cone from 0.4 to 0.2. The uncertainty reaches 8% at low $E_T$, but is less than 1% above 50 GeV.

- An uncertainty due to the description of the detector in simulated samples is derived in the same way as for the method using radiative $Z$ decays, using samples with variations in the amount of inactive material in front of the calorimeter. The size of the uncertainty is typically 1% at low $E_T$ and at the per-mil level at high $E_T$, except for the unconverted photons in the first endcap bin where uncertainties reach 4%.

- The statistical uncertainties in the simulation samples are accounted for using the electron extrapolation method by iteratively resampling the corresponding datasets, and are typically 0.5% or less.

The statistical uncertainty is computed as the width of the distribution of results obtained when repeating the measurement on pseudo-datasets obtained by resampling the data and reach 1–2% for $E_T < 50$ GeV and typically 0.5% at higher $E_T$. The total uncertainty reaches to $7$–$18\%$ at $E_T = 25$ GeV, but $2$–$3\%$ at 40 GeV and 1% or less above 100 GeV except for unconverted photons for $E_T > 250$ GeV and $1.52 \leq |\eta| < 1.81$ where it reaches 4% as noted above. Results are shown in Figs. 7 and 8.
Fig. 8 Comparison of the data-driven measurements of the identification efficiency for converted photons as a function of $E_T$, for the four pseudorapidity intervals a $|\eta| < 0.6$, b $0.6 \leq |\eta| < 1.37$, c $1.52 \leq |\eta| < 1.81$, and d $1.81 \leq |\eta| < 2.37$. The error bars represent the sum in quadrature of the statistical and systematic uncertainties estimated in each method. The shaded areas correspond to the statistical uncertainties. The last bin extends to 1.5 TeV.

5.4 Results and combination

The efficiencies $\varepsilon_{ID}$ measured in data for each method are shown in Fig. 7 for unconverted photons, and in Fig. 8 for converted photons. The results obtained in Sect. 5.1 in the $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$ channels are found to be in good agreement and are therefore combined and shown as $Z \rightarrow \ell\ell\gamma$. The combination is performed following the method presented below.

The three methods show excellent agreement over the energy ranges where their measurements overlap, with photon $\varepsilon_{ID}$ values ranging from 50–60% at $E_T = 10$ GeV, to 95–99% (unconverted) and 88–96% (converted) for photons with $E_T$ above 250 GeV. The $\varepsilon_{ID}$ values obtained from the inclusive photon and electron extrapolation methods can be compared directly, since they both apply to photons from the $\gamma +$ jet prompt-photon production process. The radiative $Z$ method applies to the photon sample obtained from the $Z \rightarrow \ell\ell\gamma$ process, with a different $\varepsilon_{ID}$ value due to differences in the photon kinematics within each measurement bin, and because of the impact of photons from fragmentation processes in $\gamma +$ jet production, which are absent from $Z \rightarrow \ell\ell\gamma$. Fragmentation photons are typically identified with lower efficiency, due to higher hadronic activity around the photon candidate. However, they only constitute a small fraction of the total sample, about 10% at low $E_T$ and a few percent at higher $E_T$ [7], and their impact does not lead to significant differences between the $\varepsilon_{ID}$ values measured by the different methods.

Efficiency scale factors (SF) are computed as the ratios of the data-driven values of $\varepsilon_{ID}$ to the values obtained in simulation. To account for the differences in photon properties
Fig. 9 Comparison of the measurements of the data-driven identification efficiency for unconverted photons measurements obtained using the radiative Z method with the predictions from Z → ℓℓγ simulation as a function of photon $E_T$, for the four pseudorapidity intervals a $|\eta| < 0.6$, b $0.6 \leq |\eta| < 1.37$, c $1.37 \leq |\eta| < 1.81$, and d $1.37 \leq |\eta| < 2.37$. Predictions are shown for both the nominal simulation and with the corrections described in Sect. 4.1. The bottom panels show the ratio of the data-driven values to the MC predictions (referred to as efficiency corrections in the text).

discussed above, simulated $\gamma + \text{jet}$ production is used for the electron extrapolation and inclusive photon methods, while Z → $\ell\ell\gamma$ simulation is used for the radiative Z method. Results from the radiative Z method are shown in Figs. 9 and 10 for unconverted and converted photons respectively, while those from the other two methods are shown in Figs. 11 and 12. The SF values are all compatible with unity, and do not show significant trends as a function of $E_T$.

To increase the measurement precision in the ranges where measurements from different methods are available, a combination of the three measurements is performed. The SF values are combined rather than the $\varepsilon_{ID}$ directly, in order to account for the small differences discussed above in the composition of the photon samples which may affect the measured $\varepsilon_{ID}$ value. The combination is performed using the BLUE algorithm [39,40]: the combined scale factor in each measurement bin computed as a linear combination of the per-method scale factors, with coefficients that are obtained from the values of their statistical and systematic uncertainties and their correlations.

Statistical uncertainties are taken to be uncorrelated since the methods are applied to disjoint datasets. Systematic uncertainties in material modelling, which affect both the radiative Z and inclusive photon method, are expected to have correlated effects in these two measurements. However, since these methods provide measurements in largely
disjoint photon $E_T$ intervals, the correlation assumption has a negligible impact on the combined result. All sources of statistical and systematic uncertainties are therefore assumed to be uncorrelated. The combined results are shown in Figs. 13 and 14 for unconverted and converted photons respectively. The central values are again compatible within uncertainties, which are at the level of 1% or less in the region $30 < E_T < 150 \text{ GeV}$, as well as at higher $E_T$ excepting the region $1.52 \leq |\eta| < 1.81$ where they reach 4% at $E_T = 1.5 \text{ TeV}$. For $E_T < 30 \text{ GeV}$ the uncertainties rise due to the limited size of the photon samples, reaching 25% at $E_T = 10 \text{ GeV}$.

Fig. 10 Comparison of the measurements of the data-driven identification efficiency for converted photons measurements obtained using the radiative $Z$ method with the predictions from $Z \rightarrow \ell\ell\gamma$ simulation as a function of photon $E_T$, for the four pseudorapidity intervals a $|\eta| < 0.6$, b $0.6 \leq |\eta| < 1.37$, c $1.52 \leq |\eta| < 1.81$, and d $1.81 \leq |\eta| < 2.37$.

Predictions are shown for both the nominal simulation and with the corrections described in Sect. 4.1. The bottom panels show the ratio of the data-driven values to the MC predictions (referred to as scale factors in the text).

5.5 Dependency of the photon identification efficiency on pile-up

The presence of additional $pp$ interactions together with the hard-scattering process is expected to affect the photon isolation efficiency. To describe this effect, the simulated MC events are reweighted to reproduce the distribution of $\mu$ observed in data. Thus, if the dependence of the identification efficiency on $\mu$ is correctly described by simulation, there is no need to account for it in analyses using photons.

The dependency of the photon identification efficiency on $\mu$ is measured using photon candidates originating from the $Z \rightarrow \ell\ell\gamma$ radiative decay process, following the procedure
Fig. 11 Comparison of the measurements of the data-driven identification efficiency for unconverted photons obtained using the electron extrapolation and inclusive photon methods with the predictions from prompt-photon + jet simulation as a function of photon $E_T$, for the four pseudorapidity intervals $a$ $|\eta| < 0.6$, $b$ $0.6 \leq |\eta| < 1.37$, $c$ $1.52 \leq |\eta| < 1.81$, and $d$ $1.81 \leq |\eta| < 2.37$. Predictions are shown for both the nominal simulation and with the corrections described in Sect. 4.1. The bottom panels show the ratio of the data-driven values to the MC predictions (referred to as scale factors in the text). The last bin extends to 1.5 TeV.

The measurement is restricted to candidates with transverse momentum $E_T > 20$ GeV for which background contamination is lower, and no background subtraction procedure is performed. The requirement $E_T < 40$ GeV is also applied to avoid large statistical uncertainties. Figure 15 shows the evolution of photon identification efficiency for these events as a function of $\mu$ in different regions of pseudorapidity, as measured in data and in MC simulation. A clear dependency is observed, with the photon identification efficiency dropping by about 10–15% when going from $\mu \sim 5$ to $\mu \sim 40$, depending on the photon candidate’s pseudorapidity and conversion status. The ratio of the efficiency dependence on $\mu$ in data to that in simulation is modelled with a linear dependence on $\mu$ and found to be compatible with a constant value within the measurement uncertainties in all detector regions for both the unconverted and converted photons. Similar results are obtained with lower precision from the other measurement methods.

5.6 Dependency of the photon identification efficiency on isolation

All measurements of the photon identification efficiency discussed in Sect. 5 are performed for isolated photons satisfying $1.52 \leq |\eta| < 1.81$, and $d$ $1.81 \leq |\eta| < 2.37$. Predictions are shown for both the nominal simulation and with the corrections described in Sect. 4.1. The bottom panels show the ratio of the data-driven values to the MC predictions (referred to as scale factors in the text). The last bin extends to 1.5 TeV.

Any constant difference from unity in the data/MC ratio is accounted for by the efficiency scale factors.
Fig. 12 Comparison of the measurements of the data-driven identification efficiency for converted photons obtained using the electron extrapolation and inclusive photon methods with the predictions from prompt-photon+jet simulation as a function of photon $E_T$, for the four pseudo-rapidity intervals a $|\eta| < 0.6$, b $0.6 \leq |\eta| < 1.37$, c $1.52 \leq |\eta| < 1.81$, and d $1.81 \leq |\eta| < 2.37$. Predictions are shown for both the nominal simulation and with the corrections described in Sect. 4.1. The bottom panels show the ratio of the data-driven values to the MC predictions (referred to as scale factors in the text). The last bin extends to 1.5 TeV.

6 Measurement of the $e \rightarrow \gamma$ fake-rate

Although the photon and the electron reconstruction algorithms are designed to reduce the misidentification of electrons as photons (see Sect. 3.1 for details) a residual small fraction of electrons can still be reconstructed as photon candidates and pass the photon identification requirements. Since the reconstruction efficiency of an electromagnetic cluster is about 100%, the fraction of electrons wrongly reconstructed and misidentified as photon candidates is mainly due to tracking inefficiency or bad matching of the track with the electromagnetic cluster reconstructed in the calorimeter.
The electron-to-photon fake-rate is defined as the ratio of the probability of wrongly reconstructing and identifying an electron as a photon to the probability of correctly reconstructing an electron:

$$F_{e \rightarrow \gamma} = \frac{\epsilon_{e}(\text{truth} \rightarrow \gamma_{\text{reco}}) \epsilon_{\gamma}}{\epsilon_{e}(\text{truth} \rightarrow e_{\text{reco}}) \epsilon_{e}} = \frac{\rho \epsilon_{\gamma}}{1 - \rho \epsilon_{e}},$$

where $\epsilon_{\gamma}$ and $\epsilon_{e}$ are the identification efficiency for the photon and the electron respectively and $\rho$ is the fraction of true electrons reconstructed as photon candidates. The fake-rate $F_{e \rightarrow \gamma}$ can be used to estimate the background from single-electrons misidentified as photons in analyses using photons: a control region is defined using the same selection as for the signal region, except that the photon reconstruction and identification criteria are replaced by the electron selection criteria given below. The misidentified electron background yield is then obtained as the product of $F_{e \rightarrow \gamma}$ with the observed electron yield in this control region.

The fake-rate $F_{e \rightarrow \gamma}$ is measured in data from the numbers of electron–positron candidates from $Z \rightarrow ee$ decays that are reconstructed as $ee$ or $e\gamma$ pair. Such yields are estimated by comparing the observed $ee$ and $e\gamma$ invariant mass spectra around the $Z$ boson peak, after having subtracted the respective background contributions, estimated from the sidebands around the $Z$ boson peak. Events are required to contain at least two opposite-sign electron candidates, or an electron and a photon candidate, both having a transverse momentum greater than 25 GeV and pseudorapidity $|\eta| < 2.37$, excluding the transition region $1.37 < |\eta| < 1.52$. Electron candidates are required to satisfy the electron tight identification criterion and the electron tight isolation criterion [11]. Photon candidates are required to meet the tight identification criterion and the loose isolation criterion, both discussed in Sect. 3.3. When both an electron and a photon candidate are selected within a cone of radius $\Delta R = 0.4$, the photon is kept and the electron discarded. Events with $m_{ee}$ or $m_{e\gamma} \in [86, 96]$ GeV are selected. If multiple $ee$ and/or $e\gamma$ pairs in the same event pass the previous requirements, only the pair with invariant mass closest to the $Z$ boson world-average mass is kept [41]. The fake-rate is computed as a function of the transverse momentum and absolute value of the pseudorapidity of the electron candidate. The measured fake-rate is corrected for the different background contamination in the $ee$ and $e\gamma$ samples. The background contamination is estimated by a combined signal-plus-background fit of the $ee$ or $e\gamma$ invariant mass distribution, where the resonant signal is modelled by a double-sided Crystal Ball function [42] and the continuous background by an exponential function or a Bernstein polynomial, depending on the $\eta$ bin.

Fig. 13  Efficiency scale factors (SF) for each method and their combination for unconverted photons. The last bin extends to 1.5 TeV
In order to avoid bias associated with the spectra distortion introduced by the $p_T$ binning, the background is estimated in $p_T/m_{ee,\gamma \gamma}$ bins.

The following systematic uncertainties in the measured $e \to \gamma$ fake-rate are considered: the variation of the invariant mass range used to select the dielectron and electron–photon pair candidates, from $[88.5, 93.5]$ GeV to $[83.5, 98.5]$ GeV, leads to a $\sim 5\%$ variation of the measured fake-rate; the bias introduced by calibrating the energy of an electron as a photon is found to be $\sim 4\%$. The impact of the background contributions, estimated by ignoring these components or by increasing them by 100\%, results in a $\sim 3\%$ variation in the fake-rate.

Comparing the fake-rate measured in MC simulation to the true value yields good agreement in all regions of $\eta$ and $p_T$, with differences of at most 5\%.

The integrated $e \to \gamma$ fake-rate, averaged over $\eta$ and for $p_T > 25$ GeV, is $0.015 \pm 0.001$ for converted photons and $0.030 \pm 0.002$ for unconverted photons. Examples of the fake-rate values measured as a function of $p_T$ in two different $\eta$ bins are shown in Fig. 16 for converted and unconverted photons. The fake-rate amounts to a few percent in the central region and increases in the more forward ones, reaching a value of 10–20\% for converted photons in the most extreme cases. Figure 17 shows the measured fake-rates as a function of $\eta$, averaged over $p_T$ for $p_T > 25$ GeV, and compared to the MC predictions. The fake-rate of electrons reconstructed as converted (unconverted) photons is 2–4\% (1–2\%) in the barrel and 4–12\% (2–4\%) in the endcaps. The absolute uncertainty, dominated by the systematic component, is typically around 0.2\%. The MC simulation generally underestimates the fake-rate by values ranging from a few % to about 50\% in the most extreme case (converted photons, central region).

7 Measurement of the conversion reconstruction efficiency

The photon identification selections described in Sect. 3.2 are optimised separately for unconverted and converted photons, since differences in electromagnetic shower development between the two types lead to different distributions for the discriminating variables listed in Table 1. Differences between the true conversion status of a photon and the one determined in reconstruction, described in Sect. 3.1, can thus lead to the use of a suboptimal selection, if they are not properly described in the simulation samples used for the optimization.
Fig. 15 Photon identification efficiencies as a function of the number of $pp$ interactions per bunch crossing $\mu$ for reconstructed unconverted photons (left) and converted photons (right), in four pseudorapidity regions.
Fig. 16 Evolution of the $e \to \gamma$ fake-rate as a function of $p_T$ for electrons reconstructed as converted (blue squares) and unconverted (red circles) photons in the $|\eta|$ bins 0.0–0.6, 1.15–1.37 and 2.01–2.37. Symbols with error bars correspond to the results with their total uncertainties, and gray shaded areas correspond to systematic uncertainties.

Photon conversion reconstruction efficiencies are therefore measured in data using a clean photon sample obtained from $Z \to \mu\mu\gamma$ decays. The selection closely follows the one used for the determination of photon identification efficiency using radiative $Z$ events, described in Sect. 5.1, but requires in addition the legacy photon isolation requirement (see Sect. 3.3).

Photon conversions are probed through the ratio $E_1/E_2$ of energy deposited in the first and second layers of the EMC, which has generally lower values for the later-developing showers of unconverted photons than for the electrons resulting from conversions. Binned fits in $E_1/E_2$ are performed using templates describing true unconverted photons, true converted photons and background. The fits are performed simultaneously in samples of reconstructed unconverted and converted photon samples, and split into the same four $|\eta|$ bins as for the photon identification measurements. Signal shapes are obtained from simulated $Z \to \mu\mu\gamma$ events.
Fig. 17 Evolution of the $e \rightarrow \gamma$ fake-rate as a function of $\eta$ for electrons reconstructed as unconverted (left) and converted (right) photon candidates, in data (circles) and $Z \rightarrow ee$ MC events (black triangles). Symbols with error bars correspond to the results with their total uncertainties, and gray shaded areas correspond to systematic uncertainties. The bottom plot shows the ratio of the two measurements. The error bars correspond to the scaled data uncertainty and the gray dashed area corresponds to the relative MC uncertainty.

Fig. 18 Left: Distribution of $E_1/E_2$ for converted photons with $|\eta|<0.6$ in data (black points), for true converted photons (dashed green line), unconverted photons (dashed pink line), and background (dashed red line). The full model after the fit to data is also shown (solid blue line). Right: Distribution of $m_{\mu\mu\gamma}$ for converted photons with $|\eta|<0.6$ in data (black points), together with the model fit to data (blue line) and its signal (dashed green line) and background (dashed red line) components requiring the presence of a true converted or unconverted photon. Background levels are estimated by performing in each $E_1/E_2$ bin a template fit to the distribution of the invariant mass $m_{\mu\mu\gamma}$ of the $\mu\mu\gamma$ system, over the range $76 < m_{\mu\mu\gamma} < 106$ GeV. The signal shape is obtained from $Z \rightarrow \mu\mu\gamma$ simulation, and modelled by the sum of a Crystal Ball function [42] and a Gaussian distribution. The background is modelled using a second-order Chebychev polynomial, which is found to describe well the $m_{\mu\mu\gamma}$ shape in simulated $Z \rightarrow \mu\mu +$ jets production. The resulting background yield in each $E_1/E_2$ bin is used to describe the background shape in the $E_1/E_2$ template fit. Figure 18 shows the result of the $E_1/E_2$ fit for converted photons in the region $|\eta|<0.6$, and the $m_{\mu\mu\gamma}$ fit used to obtain the background yield in the bin $E_1/E_2 < 0.2$ for the same photon selection.

The true unconverted and converted photon yields in each bin are expressed in terms of the fraction $f_{\text{conv}}^{\text{total}}$ of all photons that are reconstructed as converted, the fraction $f_{\text{fake}}^{\text{conv}}$ of true
unconverted photons that are reconstructed as converted, and the fraction $f_{\text{conv}}$ of true converted photons that are reconstructed as converted, Results are shown in Table 2 in bins of $|\eta|.$

Systematic uncertainties include uncertainties in the signal shapes from variations in the amount of material upstream of the calorimeter (following the method described in Sect. 5.1) and from the use of an analytical function to describe the signal shapes, as well as statistical uncertainties in the simulation sample. Their combined impact on the measurement of $f_{\text{conv}}$ ranges from 1.3 to 4.9%, depending on the $\eta$ bin. Statistical uncertainties range from 0.6 to 1.5%.

### 8 Conclusion

The reconstruction of photon candidates measured in the ATLAS detector is improved in several ways for the 2015 and 2016 Run 2 data-taking periods at the LHC with $pp$ collisions at $\sqrt{s} = 13 \text{ TeV},$ building on what was implemented for Run 1. A more robust reconstruction of photon conversions is introduced to cope with the higher pile-up conditions. The photon identification criteria, based on rectangular cuts on calorimetric variables, are also tuned to reduce the dependence on pile-up.

The efficiency of the tight identification criteria is measured from from $36.1 \text{ fb}^{-1}$ to $36.7 \text{ fb}^{-1}$ of $pp$ collision data for isolated photon candidates, using three independent methods in different photon $E_T$ ranges. The three measurements agree within their uncertainties in the overlapping $E_T$ ranges, and are combined. The identification efficiency increases from 45–60% at $E_T = 10 \text{ GeV}$ to 95–98% for $E_T > 100 \text{ GeV},$ depending on the pseudorapidity region of the detector and on the photon candidate conversion status. The nominal MC simulation of prompt-photons in ATLAS predicts significantly higher identification efficiency values than those measured in some regions of the phase space, particularly at low $E_T$ and high $|\eta|.$ A simulation with shower shapes corrected for the average shifts observed relative to the data better reproduces the identification efficiencies measured in the entire $E_T$ and $\eta$ range accessible by the data-driven methods. The residual difference between the efficiencies in data and in the corrected simulation are taken into account by computing data-to-MC efficiency scale factors. These factors differ from one by up to 3–4% at $E_T = 10 \text{ GeV}$ and at most by 1–2% above $E_T = 40 \text{ GeV},$ with an uncertainty decreasing from about 10% at $E_T = 10 \text{ GeV}$ to $< 1%–2%$ at high $E_T.$ The dependence of the identification efficiency on pile-up is measured and found to be well described by simulation. The dependence of the identification efficiency on the photon isolation criterion is similarly measured, and found to be negligible compared with the measurement uncertainties.

The probability of reconstructing an electron as a photon candidate is also measured in data, and compared with the predictions from simulation. It is found to range from a few % in the central region and increase in the more forward ones, reaching a value of 10–20% for converted photons in the most extreme cases. The MC simulation generally underestimates the fake rate by values ranging from a few % to about 50% in the most extreme case (converted photons, central region).

The efficiency of the reconstruction of photon conversions is measured using a sample of photon candidates from $Z \rightarrow \mu \mu \gamma$ events, exploiting the properties of the ratio of energies deposited in the first and second longitudinal layers of the ATLAS electromagnetic calorimeter. The efficiency is found to be compatible with the prediction of the MC simulation.

### Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSEF, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, The Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNED/FVA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern.

---

**Table 2** Results from data and MC simulation of the $f_{\text{conv}}$ fraction of all photons that are reconstructed as converted, of the $f_{\text{conv}}$ fraction of true unconverted photons that are reconstructed as converted, and of the $f_{\text{conv}}$ probability for a true converted photon to be reconstructed as converted, in bins of $|\eta|.$

| $|\eta|$ region | $f_{\text{conv}}$ | $f_{\text{conv}}$ | $f_{\text{conv}}$ |
|----------------|------------------|------------------|------------------|
|                | total            | fake             | reco             | total            | fake             | reco             |
| $|\eta| < 0.6$    | 0.215 ± 0.014    | 0.185            | 0.053 ± 0.007    | 0.053            | 0.731 ± 0.040    | 0.882            |
| $0.6 \leq |\eta| < 1.37$  | 0.309 ± 0.021    | 0.328            | 0.036 ± 0.007    | 0.029            | 0.708 ± 0.043    | 0.778            |
| $1.52 \leq |\eta| < 1.81$  | 0.438 ± 0.031    | 0.439            | 0.001 ± 0.009    | 0.027            | 0.812 ± 0.052    | 0.907            |
| $1.81 \leq |\eta| < 2.37$  | 0.536 ± 0.014    | 0.521            | 0.003 ± 0.006    | 0.008            | 0.544 ± 0.014    | 0.614            |

---

**CERN**
and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, UK; DOE and NSF, USA. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Sklodowska-Curie Actions, European Union; Investissements d’ Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSS-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; The Royal Society and Leverhulme Trust, UK. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (The Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (The Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNLS (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. 43.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Author’s comment: “All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standards conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/).” This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit and indicate if changes were made. Funded by SCOAP3.

References


The ATLAS Collaboration

1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany, NY, USA
3 Department of Physics, University of Alberta, Edmonton, AB, Canada
4 (a) Department of Physics, Ankara University, Ankara, Turkey; (b) Istanbul Aydin University, Istanbul, Turkey; (c) Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA
7 Department of Physics, University of Arizona, Tucson, AZ, USA
8 Department of Physics, University of Texas at Arlington, Arlington, TX, USA
9 Physics Department, National and Kapodistrian University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Department of Physics, University of Texas at Austin, Austin, TX, USA
12 (a) Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey; (b) Faculty of Engineering and Natural Sciences, Istanbul Bilgi University, Istanbul, Turkey; (c) Department of Physics, Bogazici University, Istanbul, Turkey; (d) Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
13 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
14 Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
15 (a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b) Physics Department, Tsinghua University, Beijing, China; (c) Department of Physics, Nanjing University, Nanjing, China; (d) University of Chinese Academy of Science (UCAS), Beijing, China
16 Institute of Physics, University of Belgrade, Belgrade, Serbia
17 Department for Physics and Technology, University of Bergen, Bergen, Norway
18 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
19 Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
20 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
21 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
22 Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia
23 (a) Dipartimento di Fisica, INFN Bologna and Universita’ di Bologna, Bologna, Italy; (b) INFN Sezione di Bologna, Bologna, Italy
24 Physikalisches Institut, Universität Bonn, Bonn, Germany
25 Department of Physics, Boston University, Boston, MA, USA
26 Department of Physics, Brandeis University, Waltham, MA, USA
27 (a) Transilvania University of Brasov, Brasov, Romania; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d) Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (e) University Politehnica Bucharest, Bucharest, Romania; (f) West University in Timisoara, Timisoara, Romania
28 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
29 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
30 Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
31 Cavendish Laboratory, University of Cambridge, Cambridge, UK
32 (a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa; (c) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
33 Department of Physics, Carleton University, Ottawa, ON, Canada
34 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies-Université Hassan II, Casablanca, Morocco; (b) Centre National de l’Énergie des Sciences Techniques Nucléaires (CNESTEN), Rabat, Morocco; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech, Marrakech, Morocco; (d) Faculté des Sciences, Université Mohamed Premier and LPTPM, Oujda, Morocco; (e) Faculté des sciences, Université Mohammed V, Rabat, Morocco
35 CERN, Geneva, Switzerland
am Also at Manhattan College, New York, NY, USA
an Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
ao Also at National Research Nuclear University MEPhI, Moscow, Russia
ap Also at Physics Dept, University of South Africa, Pretoria, South Africa
aq Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
ar Also at School of Physics, Sun Yat-sen University, Guangzhou, China
as Also at The City College of New York, New York, NY, USA
at Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing, China
au Also at Tomsk State University, Tomsk, and Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
av Also at TRIUMF, Vancouver, BC, Canada
aw Also at Universidad de Granada, Granada (Spain), Spain
ax Also at Universita di Napoli Parthenope, Naples, Italy
* Deceased