PROCEEDINGS OF SCIENCE

Photon identification with the ATLAS detector

Björn Wendland[a,*] **on behalf of the ATLAS Collaboration**

Technische Universität Dortmund,
Otto-Hahn Straße 4a, 44227 Dortmund, Germany

E-mail: bjoern.wendland@tu-dortmund.de

Good photon identification capabilities are important for many aspects of the ATLAS physics programme, including measurements of fundamental properties of the hard interaction in final states with one or more photons, possibly produced in association with jets or gauge bosons. The identification of prompt photons and the rejection of background coming mostly from photons from hadron decays relies on the high granularity of the ATLAS electromagnetic calorimeter. Several methods are used to measure with data the efficiency of the photon identification requirements, covering a broad energy spectrum. At low energy, photons from radiative Z decays are used. In the medium energy range, similarities between electrons and photon showers are exploited using $Z \to ee$ decays. At high energy, inclusive photon samples are used. The results of these measurements performed with pp collisions data at $\sqrt{s} = 13$ TeV in 2015-2018 corresponding to an integrated luminosity of 140 fb^{-1} are presented. The impact on the photon identification of the pile-up, especially large in the second part of 2017 data taking, is also discussed.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

*Speaker

PoS(EPS-HEP2021)332

1. Physics with photons at the ATLAS experiment

Measurements using topologies with one or multiple photons in the final state are important parts of the physics programme of the ATLAS experiment at the LHC [1]. Photon candidates are reconstructed from energy depositions in the calorimeter system that are combined with tracking information from the inner detector. Depending on whether reconstructed tracks can be matched to the clustered energy depositions, these candidates are classified as unconverted or converted photons, which are photons that undergo electron-positron pair production upstream of the calorimeters. Physics analyses target final states with prompt photons for testing the perturbative and non-perturbative regimes of QCD and the electroweak sector of the Standard Model. These final states are furthermore utilized to search for physics beyond the Standard Model. However, hadron collisions represent a particularly challenging environment for physics with photons as the vast majority of reconstructed photon candidates are of non-prompt origin. Therefore, the application of an identification algorithm offering excellent rejection of non-prompt photons while ensuring a sufficient selection efficiency of prompt photons is vital for physics analyses.

2. Photon identification with the ATLAS detector

The evolution of the electromagnetic shower in the calorimeter system is different for prompt and non-prompt photons. Therefore, the photon identification (PID) with the ATLAS detector employs so-called *shower shape variables* (SSVs) that describe the lateral and longitudinal development and exploit the fine granularity of the electromagnetic calorimeter [2]. Although the general features of the SSV distributions of photon candidates agree well between data and Monte Carlo (MC) simulation, systematic differences are observed especially for the mean values. This is corrected by shifting the MC distributions to match the data, where the shift values are obtained by χ^2-minimization in regions enriched by prompt photons [3]. However, slight mismodelling of the SSVs is still observed after the application of this correction. Rectangular cuts are applied to the SSVs of photon candidates, which are optimized with the goal to efficiently select prompt photons while providing an excellent rejection of non-prompt photons. These optimized cuts are derived separately in bins of the transverse energy of the photon candidate E_{T}^{γ} and in different regions of the absolute value of the pseudorapidity of the photon candidate $|\eta^{\gamma}|$ in order to account for the detector geometry. Three PID working points (WPs) are defined, *loose*, *medium* and *tight*, listed in decreasing order of prompt photon efficiency and increasing order of background rejection. While the former two WPs are mainly used for photon triggers, the latter is usually applied by physics analyses.

3. Measurements of the Photon identification efficiencies

The PID efficiency $\varepsilon_{\mathrm{ID}}$ is given by the fraction of prompt photons that pass the PID requirements N_{pass}^{S} of all prompt photons N_{all}^{S} in the considered phase space. In order to measure these quantities in data, the prompt photon purities after (P) and before (A) applying the PID requirements are extracted and multiplied by the corresponding observed numbers of photon candidates in data, N_{pass} and N_{all}, respectively. By comparing the measured efficiencies to the efficiencies predicted in MC,

data/MC scale factors (SF) are retrieved that are applied by physics analyses in order to correct the PID response in MC.

$$\varepsilon_{\mathrm{ID}} = \frac{N^S_{\mathrm{pass}}}{N^S_{\mathrm{all}}} = \frac{P\, N_{\mathrm{pass}}}{A\, N_{\mathrm{all}}}$$

In order to validate the tight PID WP for the use in physics analyses, the ATLAS collaboration makes use of three different data-driven methods to measure the efficiency of the tight PID WPs and to calibrate the efficiency of simulated prompt photons to data, which cover different regions of E_{T}^{γ}:

- **Radiative Z boson decays (RadZ)**: A sample containing radiative leptonic Z boson decays, i.e. $Z \to ee\gamma$ and $Z \to \mu\mu\gamma$, is selected, which is characterized by a high prompt photon purity ($\geq 90\,\%$). Photons from initial state radiation are rejected by requiring the invariant mass of the lepton pair to match $40\,\mathrm{GeV} < m_{\ell\ell} < 83\,\mathrm{GeV}$. The results obtained by extracting the prompt photon purities from data by using template fits to the $m_{\ell\ell\gamma}$ distributions are combined with the results from a fully data-driven method based on isolation distributions. This method is used in the range of $10\,\mathrm{GeV} < E_{\mathrm{T}}^{\gamma} < 100\,\mathrm{GeV}$.
- **Electron extrapolation method (EE)**: For this method, a pure electron sample is generated by selecting events containing $Z \to ee$ decays via a tag-and-probe algorithm. Smirnov transformations are applied to the SSVs of the probe electron in order to match the SSVs of prompt photon candidates. The parameters of this transformation are obtained by the comparison of the SSV distributions of electrons and photons in MC. The requirements of the PID are applied to the transformed SSVs of the probe electron. A template fit to the m_{ee} distributions is employed to extract the prompt photon purities. This method is used in the range of $25\,\mathrm{GeV} < E_{\mathrm{T}}^{\gamma} < 250\,\mathrm{GeV}$.
- **Matrix method (MM)**: Samples containing single photons mainly stemming from quark-antiquark-annihilation or Compton scattering are used. This method covers the largest region of photon transverse energy of $25\,\mathrm{GeV} < E_{\mathrm{T}}^{\gamma} < 1\,500\,\mathrm{GeV}$. It suffers from a large contamination of non-prompt photons mainly originating from dijet production in the low E_{T}^{γ} region, but offers high prompt photon purity and large statistics in the high E_{T}^{γ} region. The prompt photon purities are estimated by determining the track isolation efficiencies ($\hat{\varepsilon}$) in data and for prompt and non-prompt photons. While $\hat{\varepsilon}$ is extracted from MC for prompt photons, a data-driven estimation is performed for non-prompt photons employing regions enriched in non-prompt photons defined by inverted requirements on the SSVs.

4. Results

The results for the tight PID efficiencies for converted and unconverted photons in $|\eta^{\gamma}| < 0.6$ are shown in Figure 1 for all three methods. While results obtained with the full Run 2 dataset taken from 2015-2018 corresponding to an integrated luminosity ($\mathcal{L}_{\mathrm{int}}$) of $140\,\mathrm{fb}^{-1}$ are shown for the measurements with MM and RadZ, results from Ref. [2] are depicted for EE using a partial Run 2

dataset taken during 2015-2017 corresponding to $\mathcal{L}_{\text{int}} = 81\,\text{fb}^{-1}$. It is observed that the individual results are in good agreement with each other. The efficiencies increase as function of E_{T}^{γ} and saturate in the high E_{T}^{γ} region. The SFs obtained by the individual measurements are shown in the bottom panels of Figure 2 and agree well with each other within the uncertainties. These SFs are combined by using the BLUE method [5], where the uncertainties of the individual measurements are treated as uncorrelated. The combined SFs are close to 1 for the whole E_{T}^{γ} region implying a good modelling of prompt photons in MC and reach a precision of $\leq 2\,\%$ for $E_{\text{T}}^{\gamma} > 25\,\text{GeV}$.

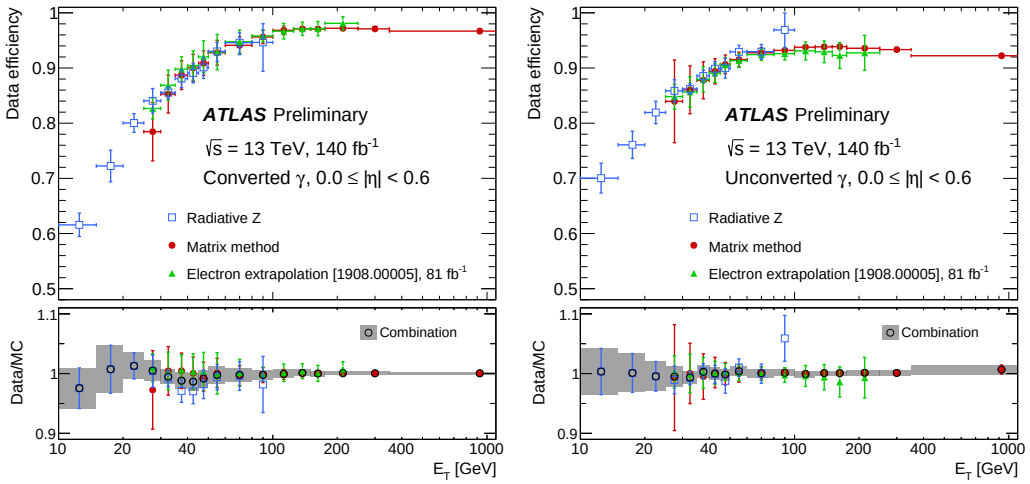

Figure 1: Results for the tight PID efficiency measurements as function of E_{T}^{γ} obtained by the three data-driven different methods for converted photons (left) and unconverted photons (right) fulfilling $|\eta| < 0.6$. The uncertainty bars represent the total uncertainty of the measurements. The bottom panel shows the corresponding measured SF from the individual measurements and the combination of the three individual values using the BLUE method. The grey band represents the total uncertainty [4].

The dependence of the tight PID efficiency on the mean number of interactions per bunch crossing $<\mu>$ is studied with photon candidates obtained from RadZ and is illustrated in Figure 2 for converted and unconverted photons fulfilling $20\,\text{GeV} < E_{\text{T}}^{\gamma} < 40\,\text{GeV}$. The data taken during 2017 corresponding to $\mathcal{L} = 44.3\,\text{fb}^{-1}$ is analyzed as it is highly suitable due to $<\mu>$ being especially large in the second part of this data taking period. While the bottom panel indicates a good agreement of this dependence between data and MC, it is observed that the tight PID efficiency drops by 10-15 % from low to high pile-up conditions.

5. Conclusions

The photon identification employs rectangular cuts on shower shape variables that describe the development of the electromagnetic shower in the calorimeter system of the ATLAS detector. It is optimized in order to efficiently select prompt photons while excellently rejecting non-prompt photons. The performance is evaluated with three different data-driven techniques, radiative Z boson decays, the electron extrapolation method and the matrix method. The results from the three measurements agree well within their uncertainties and indicate a consistent modelling in simulation that agrees well with data. The individual results are combined by using the BLUE method. The

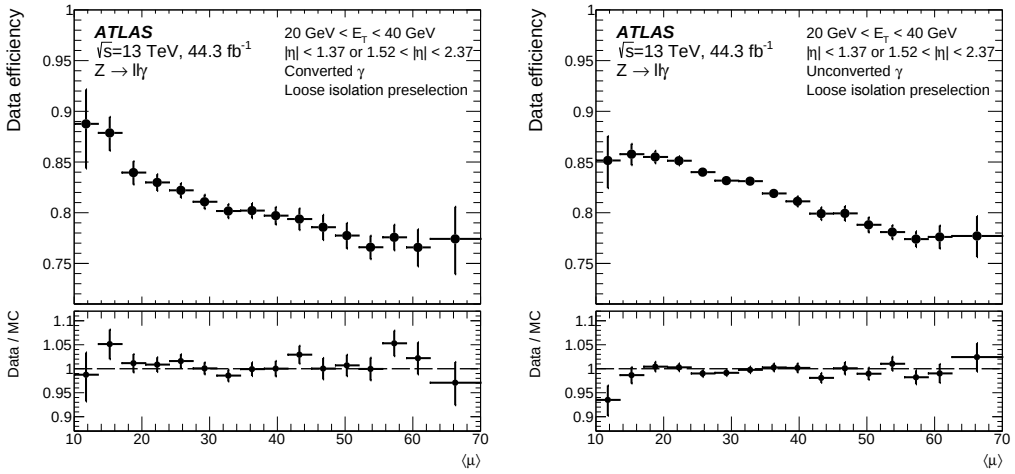

Figure 2: Measured tight PID efficiency as function of <μ> for converted (left) and unconverted photons (right) with 20 GeV< E_{T}^{γ} < 40 GeV. The bottom panel shows the corresponding SF. The results are obtained using RadZ and data taken during 2017 corresponding to $\mathcal{L}_{\mathrm{int}} = 44.3\,\mathrm{fb}^{-1}$ [2].

photon identification performance drops with increasing pile-up, while a good modelling of the pile-up dependence in simulation is observed.

Acknowledgements

The author acknowledges the financial support by the Federal Ministry of Education and Research of Germany in the framework of ATLAS (ErUM-FSP T02).

References

[1] ATLAS collaboration, *The ATLAS Experiment at the CERN Large hadron Collider*, *JINST* **3** (2008) S08003.

[2] ATLAS collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *JINST* **14** (2019) P12006.

[3] ATLAS collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, *Eur. Phys. J. C* **79** (2019) 205.

[4] ATLAS collaboration, *Photon identification efficiencies and scale factors in Run 2*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2021-01.

[5] L. Lyons, D. Gibaut, P. Clifford, *How to combine correlated estimates of a single physical quantity*, *Nucl. Instr. and Meth. A* **270** (1988) 110.