CMS electron and photon performance at Run 2 and prospects for Run 3

Livia Soffi[a] for the CMS Collaboration

[a]*Istituto Nazionale di Fisica Nucleare, Rome, Italy*

E-mail: livia.soffi@cern.ch

The performance of electron and photon reconstruction and identification in CMS during LHC Run 2 was measured using data collected in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2016–2018 corresponding to a total integrated luminosity of 136 fb^{-1}. The energy resolution and energy scale measurements, together with the relevant identification efficiencies, remain stable throughout the full Run 2 data-taking period (2016–2018) and are similar to that of Run 1, despite the increased pileup and radiation damage. A global timing resolution of 200 ps is also measured for electrons from Z decays with the full Run 2 collision data.

1. Introduction

Electrons and photons are reconstructed with high purity and efficiency in the CMS experiment, one of the two general-purpose detectors operating at the CERN LHC. They leave a distinctive signal in the electromagnetic calorimeter (ECAL) as an isolated energy deposit that is also associated with a trace in the silicon tracker in the case of electrons. In this paper, the performance of the reconstruction and identification of electrons and photons with the CMS detector in LHC Run 2 is presented. The results are based on proton-proton (pp) collision data collected during 2016–2018, and correspond to a total integrated luminosity of 136 fb^{-1} [1–3].

2. The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate (PbWO$_4$) crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside of the solenoid. The original (Phase 0) pixel detector had three layers in the barrel and two disks in each of the endcaps, whereas the Phase 1 pixel detector has one more layer and disk each in the barrel and endcaps, respectively, with a total of 124 million pixels. The ECAL consists of 75 848 PbWO$_4$ crystals, which cover the range $\eta < 1.48$ in the barrel region (EB) and $1.48 < \eta < 3.00$ in the two endcap regions (EE).
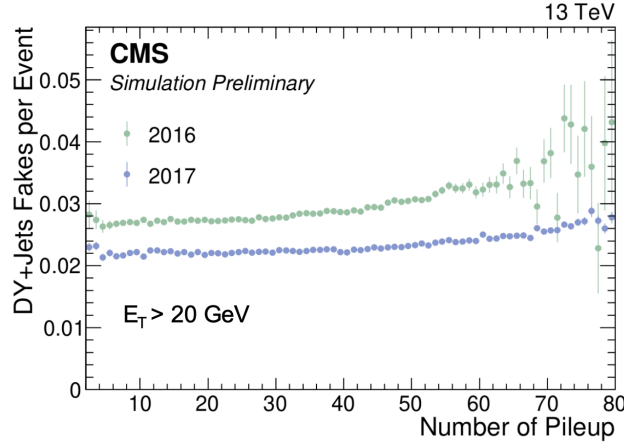
3. Electrons and photons reconstruction

Photons are identified as ECAL energy deposits (clusters) not linked to any extrapolated track. Electrons are identified as primary charged-particle tracks and potentially as ECAL energy clusters. These clusters correspond to the electron tracks extrapolated to the ECAL surface and to possible bremsstrahlung photons emitted by the electron when traversing through the tracker material. The signals in the ECAL crystals are reconstructed by fitting the signal pulse with multiple template functions to subtract the contribution from out-of-time pileup. This procedure [4] has been used for the whole LHC Run 2 data-taking period, for both the High-Level Trigger and offline event reconstruction.

As an electron or photon propagates through the material in front of the ECAL, it may interact with the material with the electron emitting bremsstrahlung photons and the photon converting into an electron-positron pair. A dedicated algorithm is used to combine the clusters from the individual particles into a single object to recover the energy of the primary electron or photon. A dedicated tracking algorithm, based on the Gaussian sum filter (GSF), is used for electrons to estimate the track parameters [5]. Energy deposits in several ECAL channels are clustered (superclustering, SC) with a procedure called "mustache" algorithm, which is particularly useful to properly measure low-energy deposits. The algorithm starts from a cluster above 1 GeV, called seed cluster. Additional clusters are added if falling into a zone, whose shape is similar to a mustache in the transverse plane. The mustache SCs can be refined using the information from detector subsystems beyond the ECAL. Additional conversion and bremsstrahlung clusters are recovered using information from the tracker [6].

Electron reconstruction efficiency is higher than 95% for transverse energy $E_T > 20$ GeV, and is compatible between data and simulation within 2%. The better the reconstruction algorithm, the lower the misidentification rate per event. The larger the number of multiple interactions in an event, the larger the misidentification rate.

Figure 1 shows the number of misidentified electron candidates (fakes) per event as a function of the number of pileup vertices, in simulated DY+Jets samples. The significant suppression of the misidentification rate in 2017 is due to the new pixel detector.


Figure 1: Number of misidentified electrons per event as a function of the number of generated vertices.

4. Energy corrections

The energy deposited by electrons and photons in the ECAL and collected by the superclustering algorithm is subject to losses through lateral and longitudinal shower leakages or in intermodule gaps or dead crystals or because of the energy lost in the tracker.

These losses result in systematic variations of the energy measured in the ECAL. Without any corrections, this would lead to a degradation of the energy resolution for reconstructed electrons and photons. Energy resolution in simulation is spreaded to match that observed in data. The energy scales are corrected by varying the scale in the data to match that observed in simulated events. The magnitude of the final correction is up to 1.5% with a total uncertainty estimated to be smaller than 0.1 (0.3)% in the barrel (endcap). The data-to-simulation agreement, after the application of residual scales to data and
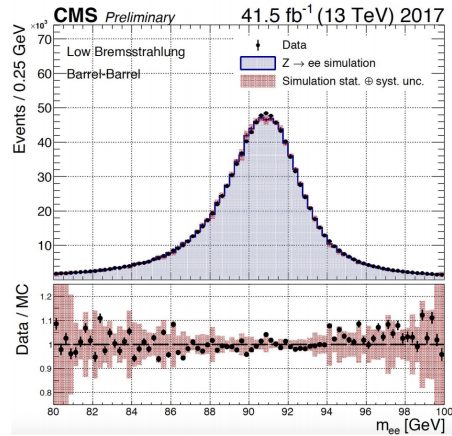

Figure 2: Invariant mass distribution of $Z \rightarrow ee$ events, after spreading is applied to simulation and scale corrections to data. The hatched regions show the combined statistical and systematic uncertainties in the simulation.

3

spreadings to simulated events, is shown in Fig. 2 for a representative category. The ultimate energy resolution after all the corrections (regression and scale corrections) ranges from 2 to 5%, depending on electron pseudorapidity and energy loss through bremsstrahlung in the detector material among the entire Run 2. The precise Higgs boson mass measurement in the diphoton channel [7] benefits from a refined calibration of the ECAL to constrain the uncertainty in the Higgs boson mass to $m_H = 125.78 \pm 0.26$ GeV [7]. The accuracy of the energy scale correction extrapolation in the energy range of interest of the $H \to \gamma\gamma$ search (between 45 and 65 GeV in E_T) is 0.05–0.1 (0.1–0.3)% for photons in the EB (EE) [7].

5. Identification strategies

The performance of the identification algorithms for electrons and photons is crucial for the physics reach of the CMS experiment. One of the most efficient ways to reject electron and photon backgrounds is the use of isolation energy sums of the reconstructed energy in a cone around electrons or photons in different subdetectors, corrected to mitigate the contribution from pileup. Another method to reject jets with high electromagnetic content exploits the shape of the electromagnetic shower in the ECAL. Two different techniques are used in CMS for the identification of electrons and photons. One is based on sequential requirements (cut-based), and the other is based on a multivariate discriminant (BDT). Although the latter is more suited for precision measurements and physics analyses with well-established final states, the former is largely used for model independent searches of nonconventional signatures. A comparison of the performance between cut-based identification and BDT identification for photons is shown in Fig. 3. The efficiency correction factors between data and simulation are stable within 3% over the full three years as shown in Figure 4 for a representative electron identification working point.
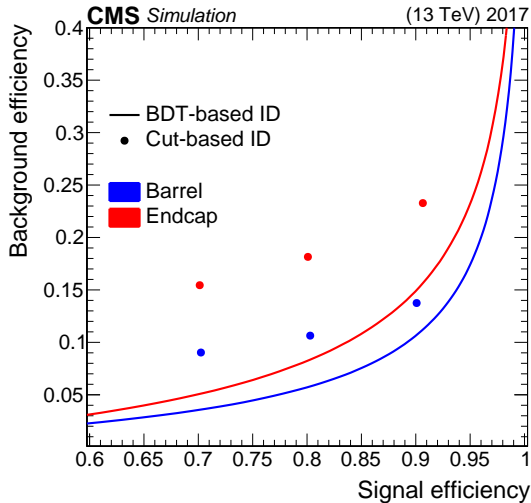
Figure 3: Performance of photon identification algorithms in 2017. Cut-based identification is shown for three working points.
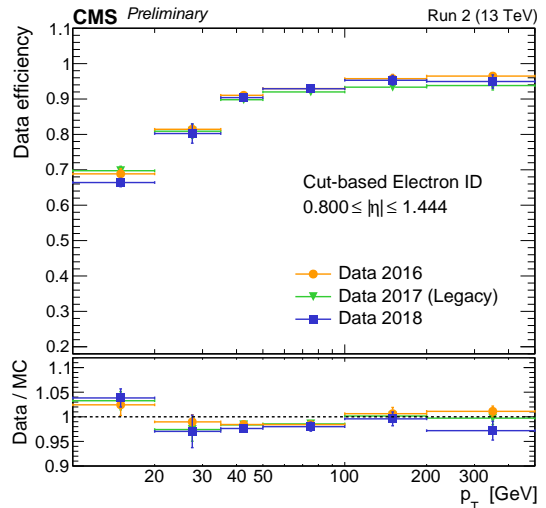
Figure 4: Cut-based loose electron identification efficiency in data and data-to-simulation efficiency ratios shown for Run 2.

6. Timing measurement

In addition to the energy measurement, the ECAL provides a time of arrival for electromagnetic energy deposits that can separate prompt electrons and photons from backgrounds with a broader time of arrival distribution [8]. Many physics analyses at CMS win simpathy for timing information as a discriminating variable, for example, in searches for nonconventional signatures with delayed objects. The resulting resolution for the full Run 2 inclusive data set is shown in Fig. 5 as a function
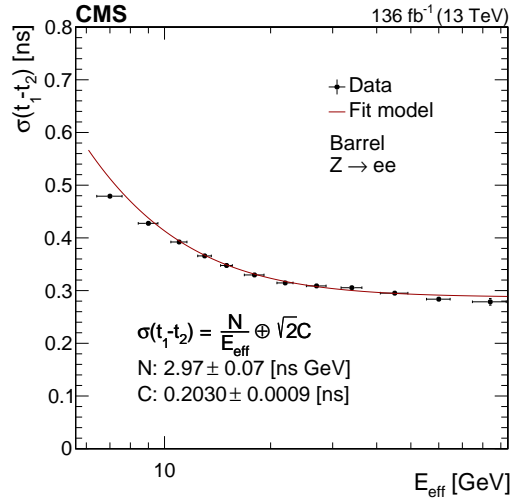
Figure 5: ECAL timing resolution in Run 2 data.

of the effective energy of the dielectron system, which depends on the individual energies of the two electrons measured in the two seed crystals as $E_{\mathrm{eff}} = E_1 E_2 / \sqrt{E_1^2 + E_2^2}$. The noise term N is very similar to that obtained prior to collisions [8] and the constant term C, which dominates at energies above 30-40 GeV, is about 200 ps.

7. Run 3 challenges

Run 3 will bring a harsher environment for electromagnetic object identification, compared to Run 2, mainly because of increasing pile-up. To mitigate the effect of increased noise in Run 3 in the ECAL subdetector, the noise threshold values have been optimized. Showers from real electrons or photons are narrower than those from jets. However, if crystals with noise are included in the definition of the shower shape variables, for example the $\sigma_{i\eta i\eta}$ variable [9], the shape of real electromagnetic objects becomes broader, and starts to look more like
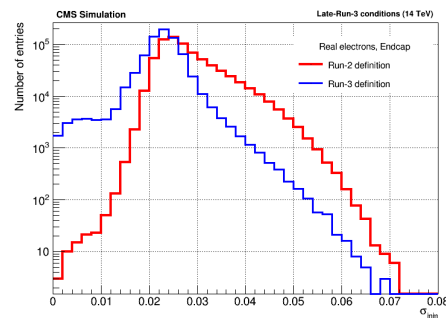
Figure 6: Comparison of distributions of $\sigma_{i\eta i\eta}$ in the endcaps, using Run 2 and Run 3 definitions.

that of jets. For the redefinition of $\sigma_{i\eta i\eta}$ in Run 3, only those crystals will be considered to construct the variable which have deposited energy greater than 1.25 times the noise threshold value. Comparison of distributions of $\sigma_{i\eta i\eta}$ in the endcaps, using Run 2 and Run 3 definitions, for reconstructed electrons in the $Z \rightarrow ee$ is shown in Figure 6. In the Run 3 definition, the distribution shifts towards smaller values as the shower shape becomes narrower after noise cleaning.

8. Conclusion

The performance of electron and photon reconstruction and identification in CMS during LHC Run 2 was measured using data collected in proton-proton collisions at $\sqrt{s}$ = 13 TeV in 2016–2018 corresponding to a total integrated luminosity of 136 fb^{-1}. The energy resolution and energy scale measurements, together with the relevant identification efficiencies, remain stable throughout the full Run 2 data-taking period (2016–2018) and are similar to that of Run 1, despite the increased pileup and radiation damage. A global timing resolution of 200 ps is also measured for electrons from Z decays with the full Run 2 collision data.

References

[1] CMS collaboration, *CMS luminosity measurements for the 2016 data-taking period*, Tech. Rep. CMS-PAS-LUM-17-001 (2017).

[2] CMS collaboration, *CMS luminosity measurement for the 2017 data-taking period at* $\sqrt{s}$ = 13 TeV, Tech. Rep. CMS-PAS-LUM-17-004 (2018).

[3] CMS collaboration, *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s}$ = 13 TeV, Tech. Rep. CMS-PAS-LUM-18-002 (2019).

[4] CMS collaboration, *Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions*, *JINST* **15** (2020) P10002 [2006.14359].

[5] W. Adam, R. Frühwirth, A. Strandlie and T. Todorov, *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, *J. Phys. G* **31** (2005) 9 [physics/0306087].

[6] CMS collaboration, *Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at* $\sqrt{s}$ = 8 TeV, *JINST* **10** (2015) P08010 [1502.02702].

[7] CMS collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, *Phys. Lett. B* **805** (2020) 135423.

[8] CMS collaboration, *Time reconstruction and performance of the CMS electromagnetic calorimeter*, *JINST* **5** (2010) T03011 [0911.4044].

[9] CMS collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at* $\sqrt{s}$ = 8 TeV, *JINST* **10** (2015) P06005 [1502.02701].