Available on CMS information server

CMS DP -2020/034

The Compact Muon Solenoid Experiment

CMS Performance Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

17 July 2020

Redefinition of Electron and Photon identification variables for Run-3 in the CMS Experiment

CMS Collaboration

Abstract

Run-3 will bring about a harsher environment for electromagnetic object identification, compared to Run-2, leading to a reduction in discrimination power for some of the identification variables. This note summarizes the re-optimization of the variables. As a result, a similar performance and discrimination power as in Run-2 can be retained.

Redefinition of Electron and Photon identification variables for Run-3 in the CMS Experiment

Detector Performance Summary Note

The CMS Collaboration
July 2020
cms-phys-conveners-egm@cern.ch

Introduction

- Efficient identification of electrons and photons is of key importance for many physics analyses.
- Run-3 will bring about a harsher environment for electromagnetic object identification, compared to Run-2, mainly because:
  - The number of pile-up interactions will increase.
  - The noise in the electromagnetic calorimeter (ECAL) will increase, especially in the endcaps.
- The discrimination power of some of the identification variables is found to be affected by these changes.
- This note summarizes the reoptimization of some of these variables to retain the discrimination power and thus maintain the Run-2 performance as much as possible.

Shower-shape variable $\sigma_{i\eta i\eta}$

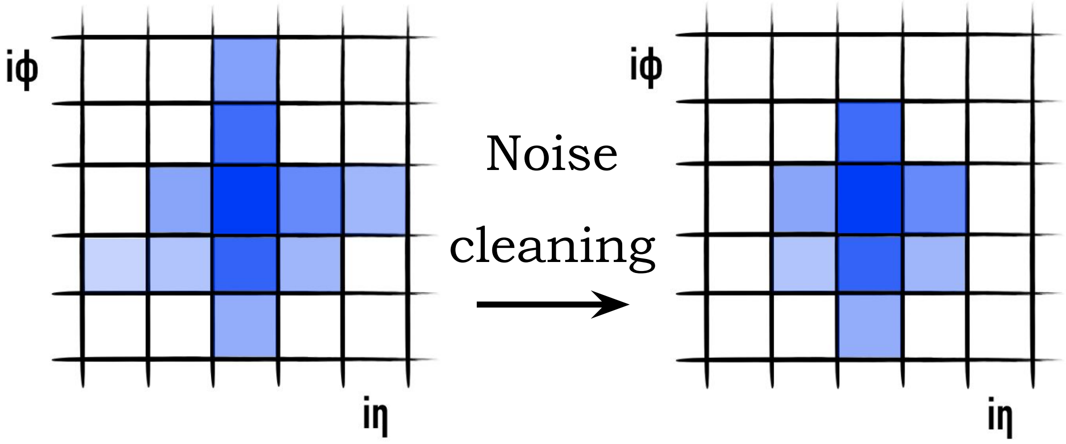
- $\sigma_{i\eta i\eta}$ is a lateral shower-shape variable, providing a measure of the spread of an electromagnetic shower in η direction.
    - It is defined as the logarithm of the energy-weighted RMS of the electromagnetic shower in units of crystals.
- This is one of the most important variables for identification of electrons and photons, because of its excellent discrimination power between electron/photon(signal) and jets(background).
- This variable is sensitive to the noise in the ECAL detector. This is particularly true for low to medium energy electrons and photons.

Redefinition of $\boldsymbol{\sigma}_{i\eta i\eta}$

- To mitigate the effect of increased noise in Run-3 in the ECAL sub-detector, the noise threshold values have been optimized.
  - A threshold level of 2σ of the noise for any crystal up to $|\eta|<2.5$, and 3σ beyond it, attenuates the effect of noise without affecting the physics performance. Here 1σ of the noise is defined as the RMS of the pedestal.
- For the redefinition of $\boldsymbol{\sigma}_{i\eta i\eta}$ in Run-3, only those crystals will be considered to construct the $\boldsymbol{\sigma}_{i\eta i\eta}$ variable which have deposited energy greater than 1.25 times the noise threshold value.
- The effect and the performance of the redefined variable are documented in the following slides.
- The performance of $\boldsymbol{\sigma}_{i\eta i\eta}$ for several scenarios of ECAL noise is compared:
  - **Late-Run-2 conditions**: ECAL noise level observed in 2018, this corresponds to integrated luminosity of 180 fb^{-1}
  - **Early-Run-3 conditions**: ECAL noise level expected in 2022
  - **Late-Run-3 conditions**: ECAL noise level expected in 2024
- We expect the noise to increase by a factor 1.8, 3 and 4 at $|\eta|=1.5$, 2.5 and 3 respectively, for an increase of integrated luminosity by 270 fb^{-1}, expected by late Run-3.

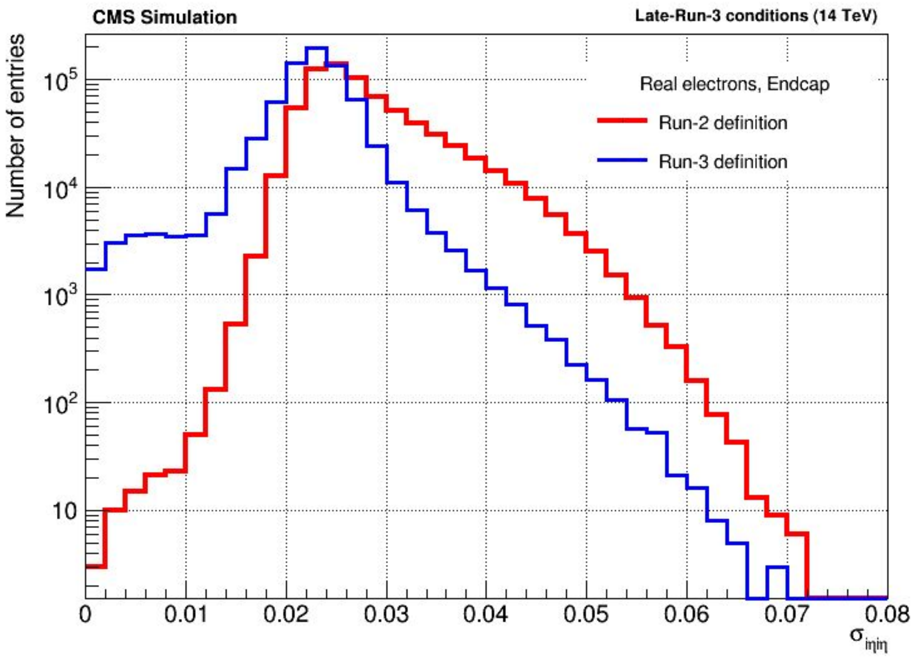
Effect of $\boldsymbol{\sigma}_{i\eta i\eta}$ redefinition

Showers from real electrons or photons are narrower than those from jets. However, if crystals with noise are included in the $\boldsymbol{\sigma}_{i\eta i\eta}$ definition, the shower shape of real electromagnetic objects becomes broader, and starts to look more like that of jets. Therefore the discrimination power of $\boldsymbol{\sigma}_{i\eta i\eta}$ decreases with increasing ECAL noise. By excluding the crystals that do not pass the noise threshold, the discrimination power can be restored.



(Left) All crystals having some energy deposit in the 5x5 area around the most energetic crystal are considered for the $\boldsymbol{\sigma}_{i\eta i\eta}$ definition. This leads to a broader shower. (Right) Crystals not passing noise thresholds are dropped. This leads to a narrower shower.

$\sigma_{i\eta i\eta}$: Run-2 vs Run-3 definitions

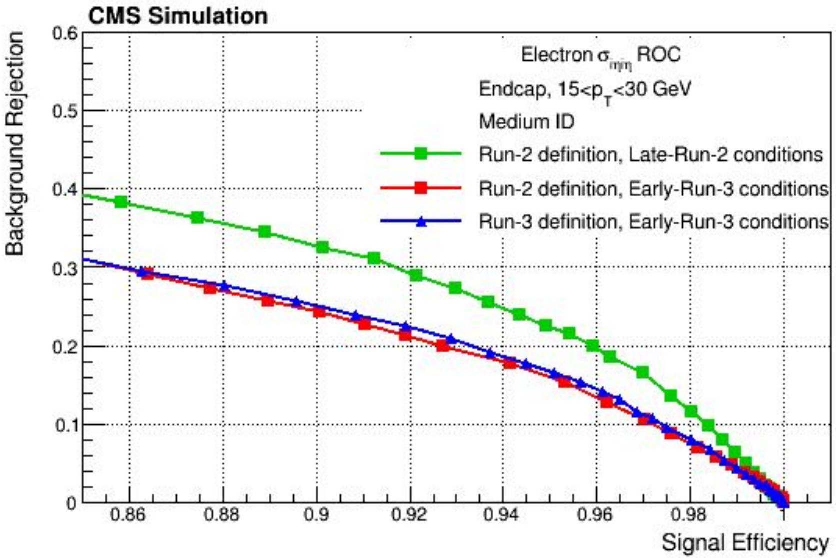


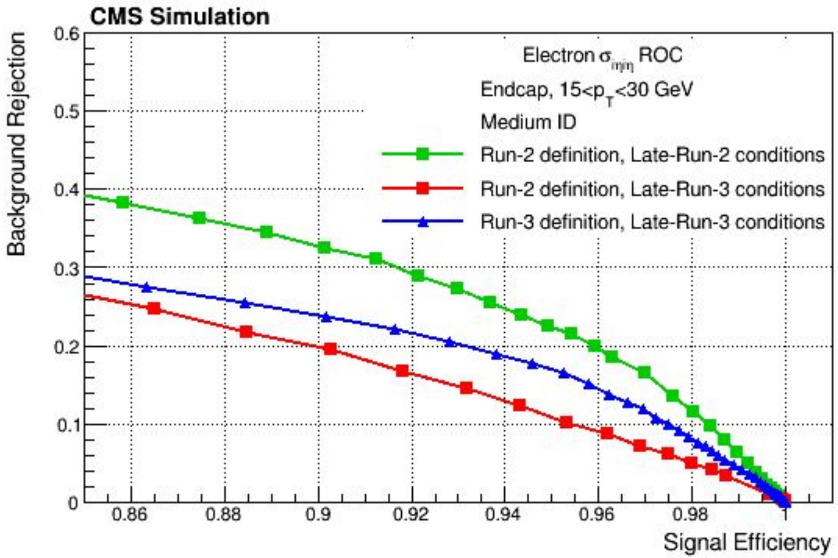
Comparison of distributions of $\sigma_{i\eta i\eta}$ in the endcaps, using Run-2 and Run-3 definitions, for reconstructed electrons in the Z->ee MC sample, geometrically matched to true electrons, in **late-Run-3 conditions**. The shower shape in the Run-2 definition is broader. In the Run-3 definition, the distribution shifts towards smaller values of $\sigma_{i\eta i\eta}$ as the shower shape becomes narrower after noise cleaning.

Technical details of $\sigma_{i\eta i\eta}$ studies

- In the next two slides, ROC curves are presented showing the background rejection as a function of signal efficiency for varied cut values on the $\sigma_{i\eta i\eta}$ distribution after applying electron/photon ID cuts (medium working point) on the other variables.
    - For electron ID, the other variables are longitudinal shower-shape variable H/E, isolation variables, and electron track-related variables.
    - For photon ID, the track-related variables are not used. Photons, being electrically neutral, pass through the tracking detector undetected. Photon cut-based ID thus solely relies on shower-shape and isolation variables.
- Signal (background) are reconstructed electrons or photons which are geometrically matched (unmatched) to true electrons or photons in simulated MC samples.
- For electrons, the signal electrons are taken from Z(ee)+jets and top quark pair production processes and background is from the same processes and additionally from QCD.
- For photons, signal and background are taken from photon+jets MC samples.
- Run-2 (Run-3) MC samples are simulated for pp collision at center-of-mass energies of 13 (14) TeV.

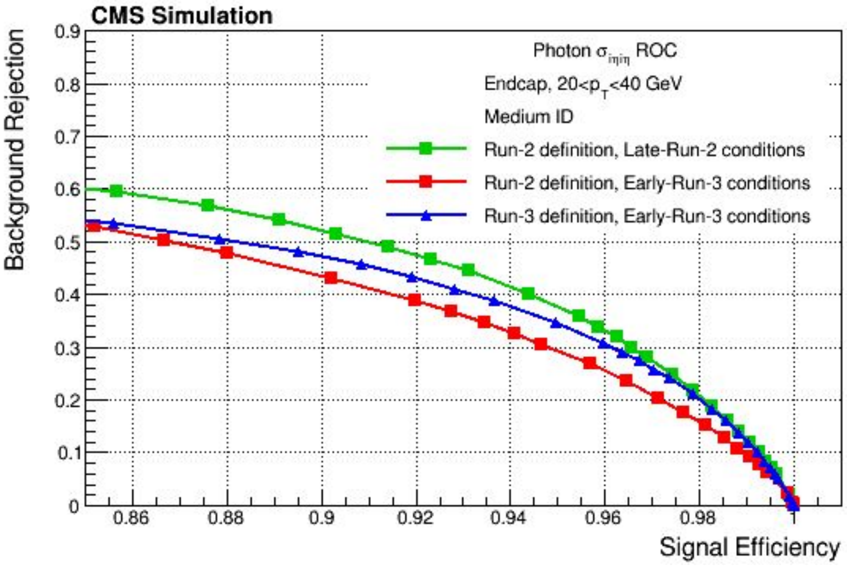
Effect of $\sigma_{i\eta i\eta}$ redefinition: electrons

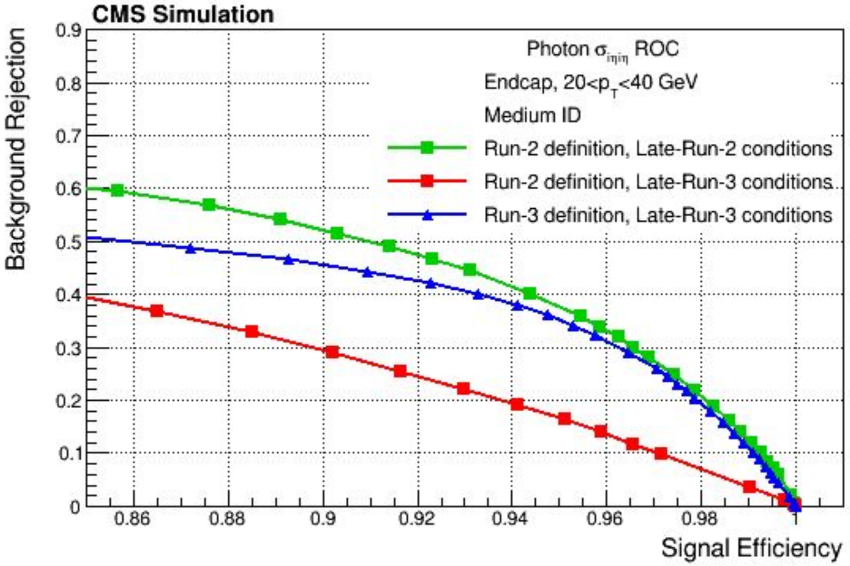




Redefinition of $\sigma_{i\eta i\eta}$ shows substantial performance improvements in the endcaps in Late-Run-3 conditions (right plot), where the noise is high. The noise cleaning brings less improvement for Early-Run-3 conditions (left plot), where the noise level is comparatively low. The decrease of discrimination power of $\sigma_{i\eta i\eta}$ due to the harsher environment in Run-3 can be partially regained with the redefinition.

Effect of $\sigma_{i\eta i\eta}$ redefinition: photons



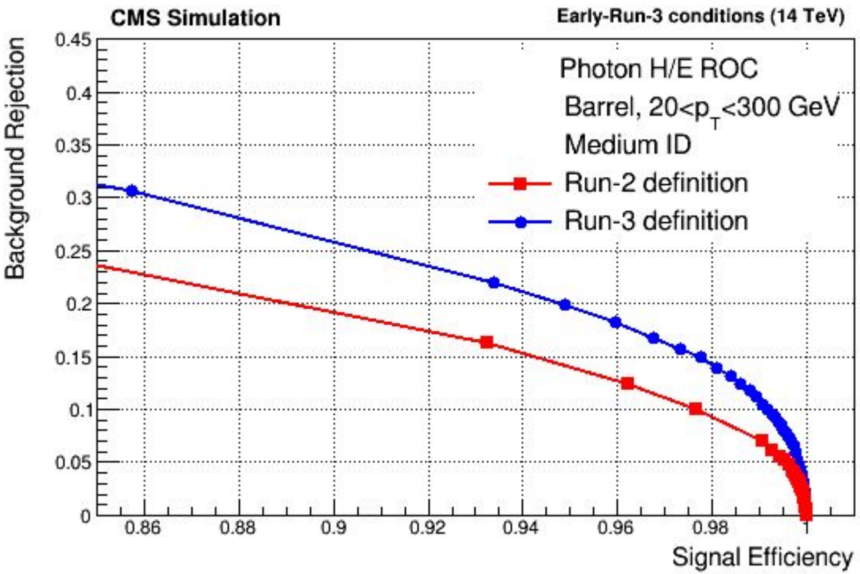


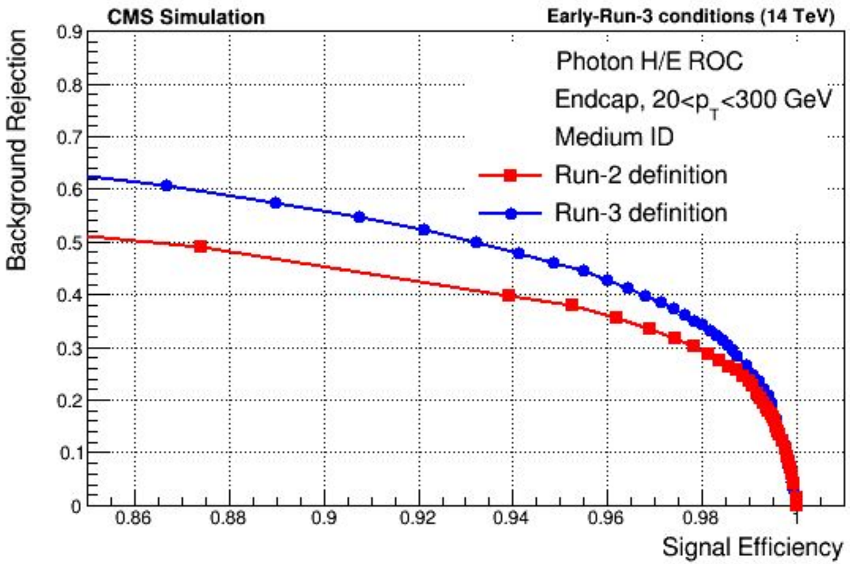
Since photon ID relies more on shower-shape variables compared to electrons, due to the absence of tracker related variables for photons, the redefinition of $\sigma_{i\eta i\eta}$ brings more improvement for photons compared to electrons.

Shower-shape variable H/E

- H/E is a longitudinal shower-shape variable, where H and E are the energy deposited by an object in HCAL and ECAL respectively.
- This is an important variable used for identification of electrons and photons, because of its excellent discrimination power between electromagnetic objects, which deposit very little energy in HCAL, and jets, which deposit substantial energy in HCAL.
- For electrons, H is defined as HCAL energy within a cone of $\Delta R<0.15$ around the position of the electron in the ECAL.
- For photons, H is defined as energy deposited in one HCAL tower directly behind the position of the photon in the ECAL.
- The definition used for electrons was found to be better than the definition used for photons, and in Run-3 the same cone-based H/E is proposed to be used for photons as well.

Photon H/E redefinition





- ROC comparing current single-tower based H/E definition for photons ($20<p_T<300$ GeV) with cone-based H/E definition for use in Run-3, for barrel (left) and endcaps (right) using simulated MC samples for **Early-Run-3 conditions**.
- The cone-based definition shows significant improvement in both cases.

Summary

- The performance of two discriminant variables, $\boldsymbol{\sigma}_{i\eta i\eta}$ and H/E, is presented in this note.
- In preparation of Run-3, the previous definition of $\boldsymbol{\sigma}_{i\eta i\eta}$ is improved in its robustness against increased noise levels in the ECAL. The redefined $\boldsymbol{\sigma}_{i\eta i\eta}$ is expected to be used for electron and photon identification in Run-3 in the endcaps, where ECAL noise will significantly increase compared to Run-2.
- For Run-2, different H/E definitions were used for electrons and photons. The cone-based definition used for electrons was found to be better in performance. In Run-3, photons will also use the cone-based definition for H/E, instead of the single-tower based definition used in the past.