e/γ reconstruction & identification in CMS

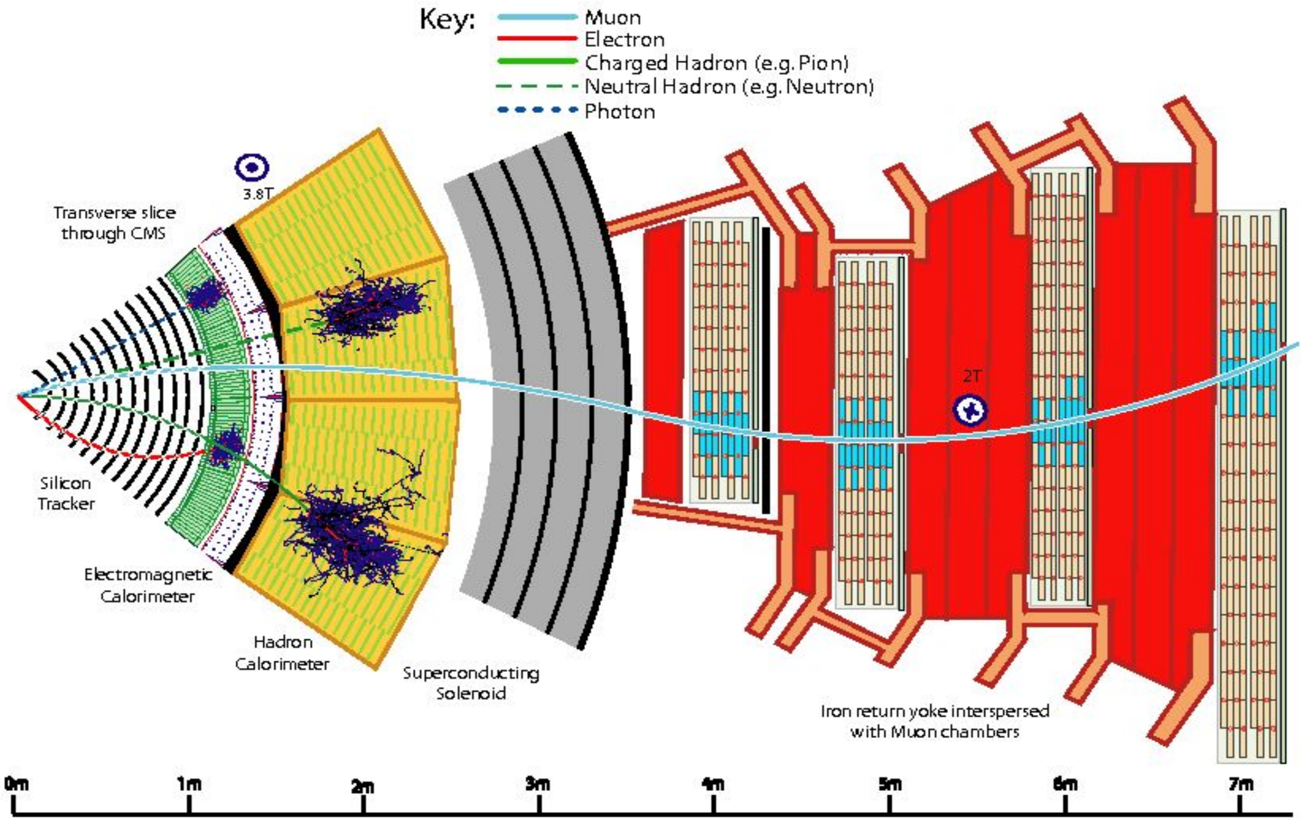
Anirban Bala
DHEP, TIFR

[arXiv: 2012.06888]

Introduction

- ➢ e/γ are reconstructed with high purity & efficiency in CMS.
- ➢ Distinctive signal in electromagnetic calorimeter (ECAL) as an isolated energy deposit associated with a trace in the 'Si' tracker in case of electrons.
- ➢ The main topics that will be covered in this talk:
  - Short description of the CMS detector
  - Offline e/γ reconstruction
  - Online e/γ reconstruction
  - Difference between offline & online reconstruction
  - Electron & photon selection
  - Photon identification
  - Electron identification

Cross-sectional view of the CMS detector



Short description of the CMS detector

- **The magnet:** central feature of the CMS detector is a large superconducting solenoid magnet. It delivers an axial & uniform magnetic field of 3.8T which helps to separate the calorimeter energy deposits of charged & neutral particles.
- **The silicon inner tracker:** provides a pure & efficient charged-particle trajectory reconstruction in jets with p_T upto around 1 TeV.
- **Highly-segmented ECAL:** homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals, sufficient to contain more than 98% of the energy of electrons & photons upto 1 TeV. Two thirds of the hadrons also start showering in the ECAL before entering the HCAL. A much finer-grained detector, called **preshower**, is installed in front of each endcap disk. The aim of the preshower ⇨ (1) to discriminate between prompt photons & photons coming from π^0 decay (2) to indicate the presence of electron or photon by requiring an associated signal in the preshower.
- **HCAL:** sampling calorimeter consisting of brass absorber & plastic scintillator. Charged & neutral hadrons may initiate a hadronic shower in ECAL, which is fully absorbed in the HCAL.
- **The muon detectors:** muons produce hits in additional tracking layers located outside the calorimeters.

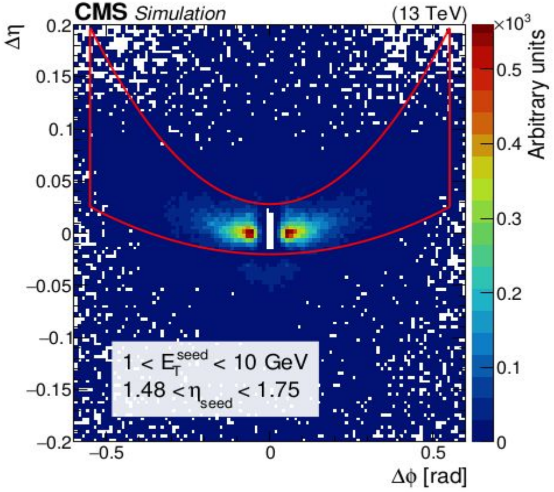
Offline e/γ reconstruction

Overview

- Electrons & photons deposit almost all of their energy in ECAL, whereas hadrons deposit in HCAL.
- In addition, electrons produce hits in the tracker.
- **Electrons** ➪ bremsstrahlung photons, **photons** ➪ electron positron pairs; thus when electron or photon reaches the ECAL, it may consist of a shower of particles.
- A dedicated algorithm to combine clusters of particles into a single object to recover the energy of primary electron or photon ➪ **Gaussian Sum Filter (GSF)**.
- The energy reconstruction algorithm starts with a formation of clusters (> 80 MeV in EB, > 300 MeV in EE). Seed consist of most energy deposits (> 1 GeV).
- Some neighbouring clusters around the SC to include pair production & brem. loss ➪ **superclustering**
- All objects with an associated GSF track ➪ electrons; without a GSF track ➪ photons

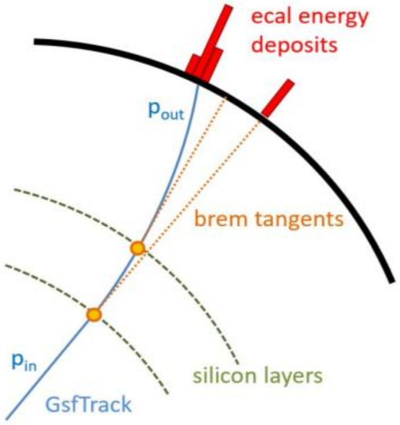
Superclustering in ECAL

- Because of showering the original object may consist of several e/γ produced from pair production/brem.
- The multiple ECAL clusters need to be combined into a single SC ⇨ **superclustering**
- There are mainly two algorithms ⇨
  - "Mustache" algorithm
  - "Refined" algorithm
- "Mustache" algorithm ⇨ useful for low-energy deposits. It uses information only from ECAL & preshower. The size of the mustache region depends on E_T
- "Refined" algorithm ⇨ It uses the tracking information. At each tracker layer, the trajectory of the GSF track is extrapolated to form a "brem. tangent", which can be linked to a compatible ECAL cluster. Clusters linked to the "brem. tangent" are then added to the refined SC.



Electron track reconstruction

- ➢ Electrons use the GSF tracking algorithm to include radiative losses from brem.
- ➢ The GSF track fitting algorithm is CPU intensive, so can't be run over all reconstructed hits in the tracker. The electron trajectory seed can be either "ECAL-driven" or "tracker-driven".
- ➢ ECAL driven seeding first selects mustache SCs with E_{SC} > 4 GeV & H/E_{SC} < 0.15. It performs better for high E_T electrons.
- ➢ Tracker driven seeding is performed only in offline reconstruction, not in HLT, as it is computationally expensive to reconstruct all tracks in an event. This approach is designed for low E_T electrons.
- ➢ The final collection of selected electron seeds is used to initiate the reconstruction of electron tracks.
- ➢ We have to use track-SC matching variables to associate a GSF track to an ECAL cluster. $|\Delta\eta| = |\eta_{SC} - \eta_{track}| < 0.02$ & $|\Delta\varphi| = |\varphi_{SC} - \varphi_{track}| < 0.15$



Online e/γ reconstruction

Overview

- e/γ candidates at L1 are based on ECAL trigger towers defined by arrays of 5x5 crystals.
- The central trigger tower with highest E_T (> 2 GeV) is designated as the seed tower.
- To recover energy loss from brem, clusters are built from surrounding towers with E_T > 1 GeV to form L1 candidates.
- No tracker information is available at L1, so electrons & photons are indistinguishable at this stage.
- The HLT electron & photon candidates are reconstructed from energy deposits in ECAL crystals grouped into clusters around L1 seed.
- Electron & photon selection at HLT relies on the identification & isolation criteria that will be discussed later.

Difference between online & offline reconstruction

- The differences between the HLT & offline reconstruction are mainly driven by:
  - The limited CPU time available at the HLT.
  - The lack of final calibrations which are not yet computed during the data-taking period.
  - More conservative selection criteria to avoid rejecting interesting events.
- Every electron candidate reconstructed at the HLT is ECAL driven.
- Offline tracker driven reconstruction is useful only for low energy electrons, which is not easy to trigger.

e/γ selection

Overview

- Two different techniques are used in CMS for the identification of electrons & photons.
    - Cut-based approach
    - MVA-based approach (will not be discussed in this talk)
- Different strategies are used to identify prompt electrons & photons.
- For prompt electrons, bkg can originate from photon conversion, hadrons misidentified as electrons, secondary electrons coming from semi-leptonic decays of b-quark.
- The main bkg to prompt photons are photons coming from π^0 decay.

Isolation criteria

- One of the methods to reject electron & photon bkg is the use of isolation energy sums.
- At first a cone is defined in η-φ plane; the distance with respect to the reconstructed electron or photon direction is defined by ΔR (= 0.3).
- The electron or photon itself is excluded from the isolation sum.

$$I_{\mathrm{PF}} = \frac{1}{p_{\mathrm{T}}}\left(\sum_{\mathrm{h}^{\pm}} p_{\mathrm{T}}^{\mathrm{h}^{\pm}} + \sum_{\gamma} p_{\mathrm{T}}^{\gamma} + \sum_{\mathrm{h}^{0}} p_{\mathrm{T}}^{\mathrm{h}^{0}}\right)$$

Shower shape criteria

- Three main shower shape variable that will be discussed:
  - H/E
  - $\sigma_{i\eta i\eta}$
  - R_9
- H/E ⇨ ratio between the energy deposited in HCAL in a cone of $\Delta R = 0.15$ around the SC & the energy of the e/γ candidate. HCAL noise, pileup, leakage of electrons or photons through the inter-modular gaps ⇨ 'H'. $H < X + Y\rho + JE$; X & Y represent the noise & pileup (PU) terms, J is scaling term for high energy electron or photon.
- $\sigma_{i\eta i\eta}$ ⇨

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_i^{5\times5} w_i(\eta_i - \overline{\eta}_{5\times5})^2}{\sum_i^{5\times5} w_i}}.$$

$w_i = \max(0, 4.7 + \ln(E_i/E_{5x5}))$

- R_9 ⇨ defined as E_{3x3}/E_{SC}. Showers of photons that convert before reaching the calorimeter have wider transverse profiles & lower values of R_9 than unconverted photons.

Photon identification

Cut-based photon identification

- The “loose” working point (WP) has an average signal efficiency of about 90%, used when bkg is low.
- The “medium” & “tight” WP have an average efficiency of about 80% & 70% respectively, used when bkg is larger.

Variable	Barrel (tight WP)	Endcap (tight WP)
H/E	<0.021	<0.032
$\sigma_{i\eta i\eta}$	<0.0099	<0.027
I_{ch}	$<0.65\,\mathrm{GeV}$	$<0.52\,\mathrm{GeV}$
I_{n}	$<0.32\,\mathrm{GeV} + 0.015E_{T} + 2.26\times10^{-5}E_{T}^{2}/\mathrm{GeV}$	$<2.72\,\mathrm{GeV} + 0.012E_{T} + 2.3\times10^{-5}E_{T}^{2}/\mathrm{GeV}$
I_{γ}	$<2.04\,\mathrm{GeV} + 0.0040E_{T}$	$<3.03\,\mathrm{GeV} + 0.0037E_{T}$

Electron identification

Cut-based electron identification

Variable	Barrel (tight WP)	Endcaps (tight WP)
$\sigma_{i\eta i\eta}$	<0.010	<0.035
$\lvert\Delta\eta_{\text{in}}^{\text{seed}}\rvert$	<0.0025	<0.005
$\lvert\Delta\phi_{\text{in}}\rvert$	<0.022 rad	<0.024 rad
H/E	$<0.026 + 1.15\,\text{GeV}/E_{\text{SC}}$ $+0.032\rho/E_{\text{SC}}$	$<0.019 + 2.06\,\text{GeV}/E_{\text{SC}}$ $+0.183\rho/E_{\text{SC}}$
$I_{\text{combined}}/E_{\text{T}}$	$<0.029 + 0.51\,\text{GeV}/E_{\text{T}}$	$<0.0445 + 0.963\,\text{GeV}/E_{\text{T}}$
$\lvert 1/E - 1/p\rvert$	$<0.16\,\text{GeV}^{-1}$	$<0.0197\,\text{GeV}^{-1}$
Number of missing hits	≤ 1	≤ 1

Summary

- First of all, a brief overview of CMS detector has been discussed.
- Then I discussed about the offline electron & photon reconstruction.
- Next online e/γ reconstruction has been described briefly.
- After that I discussed how e/γ can be selected using isolation & shower shape criteria.
- At last only the cut-based approach of e/γ identification has been discussed.