Electrons and Photons

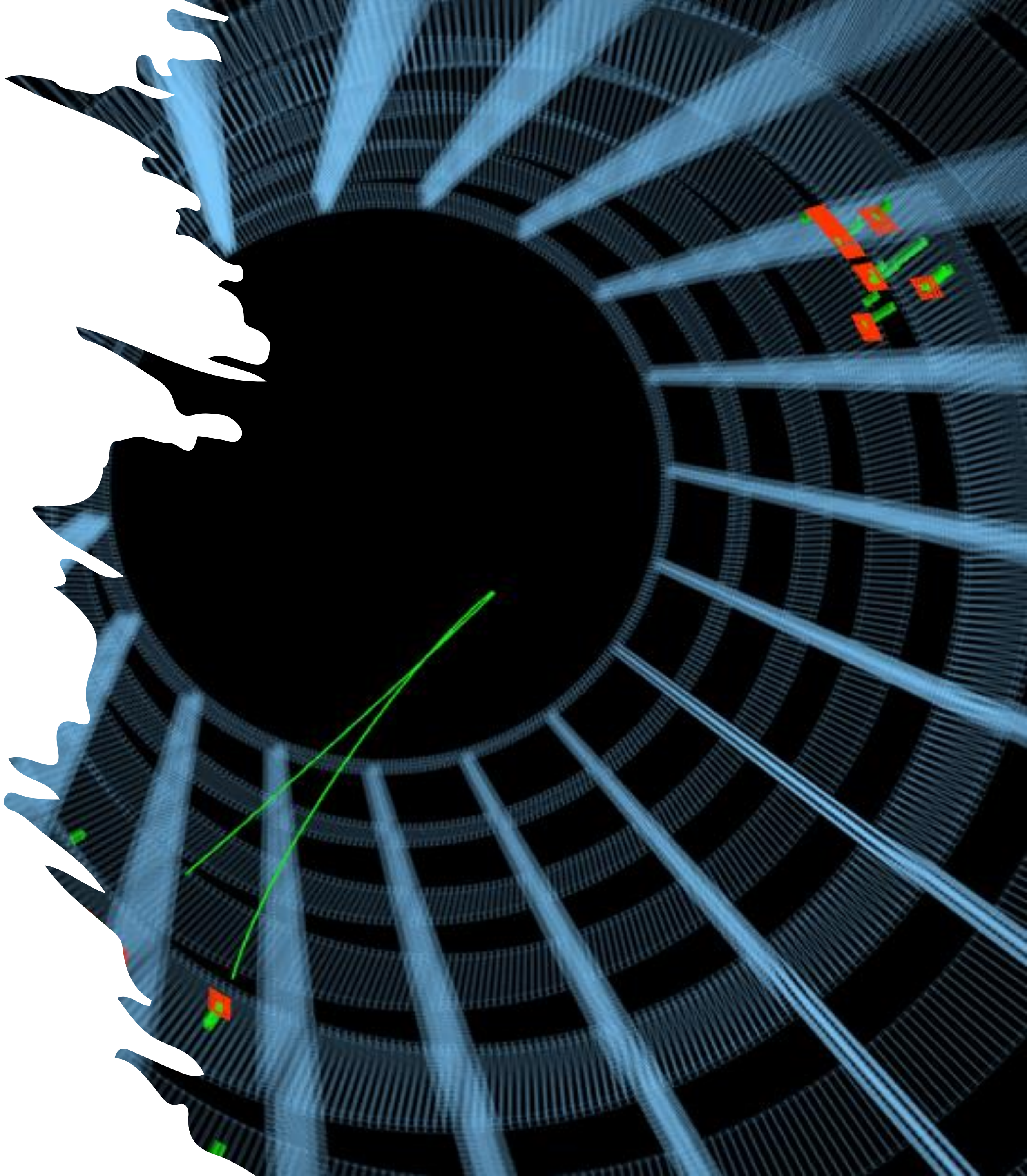
CMS China School 2024

Anshul Kapoor

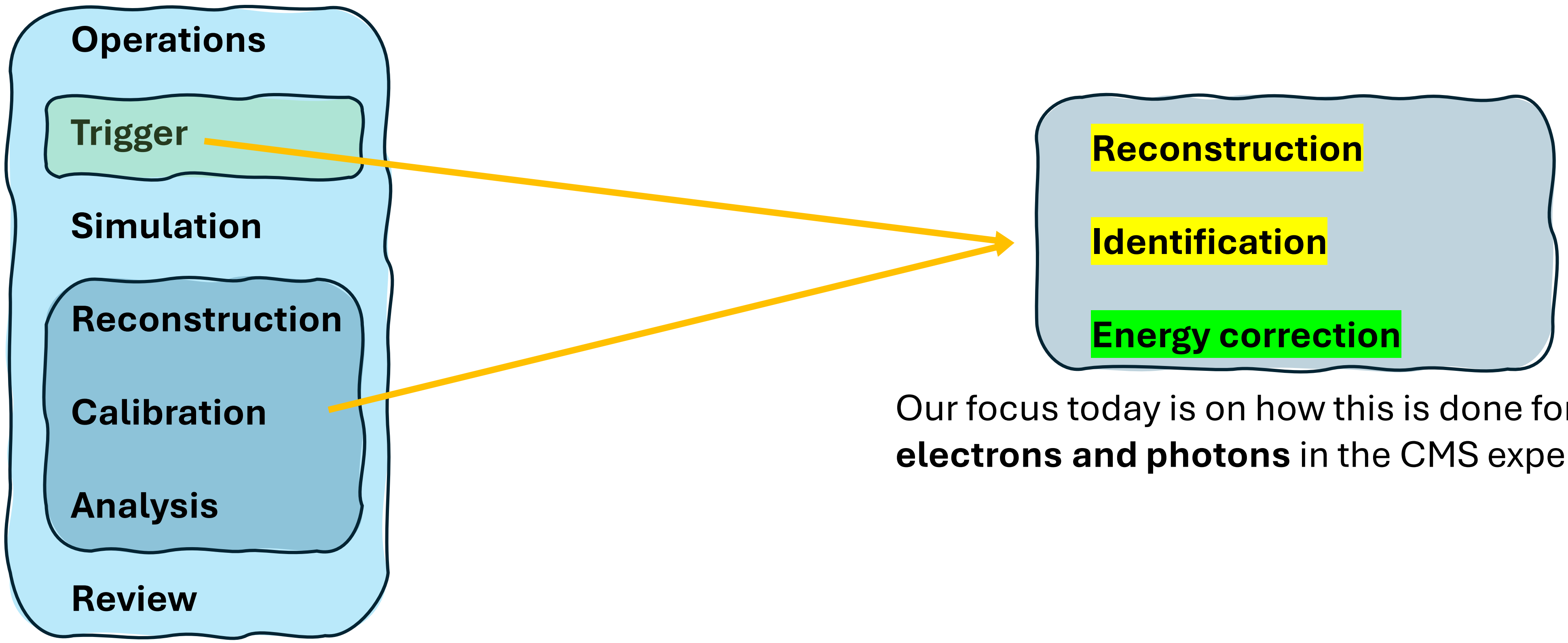
中国高等科学技术中心
China Center of Advanced Science and Technology

中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

21st Jan 2024



What is the focus today

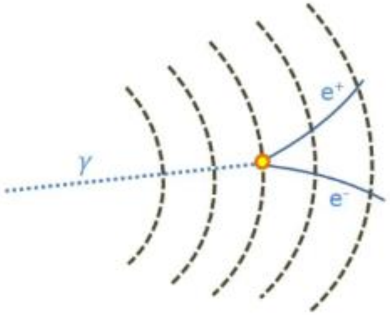


Our focus today is on how this is done for **electrons and photons** in the CMS experiment

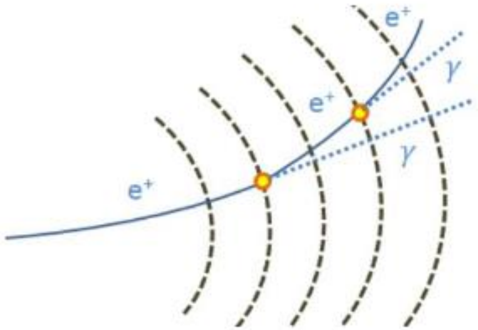
Interactions

How e/γ interact with matter

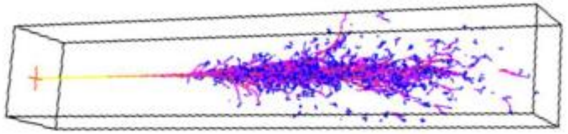
Pair production

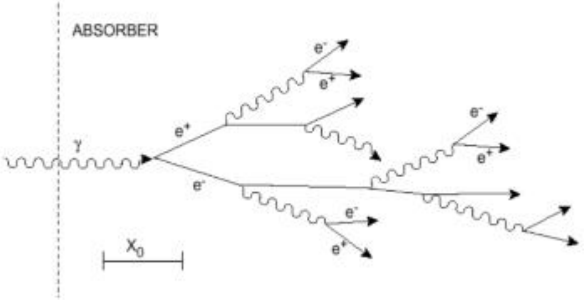


Bremsstrahlung

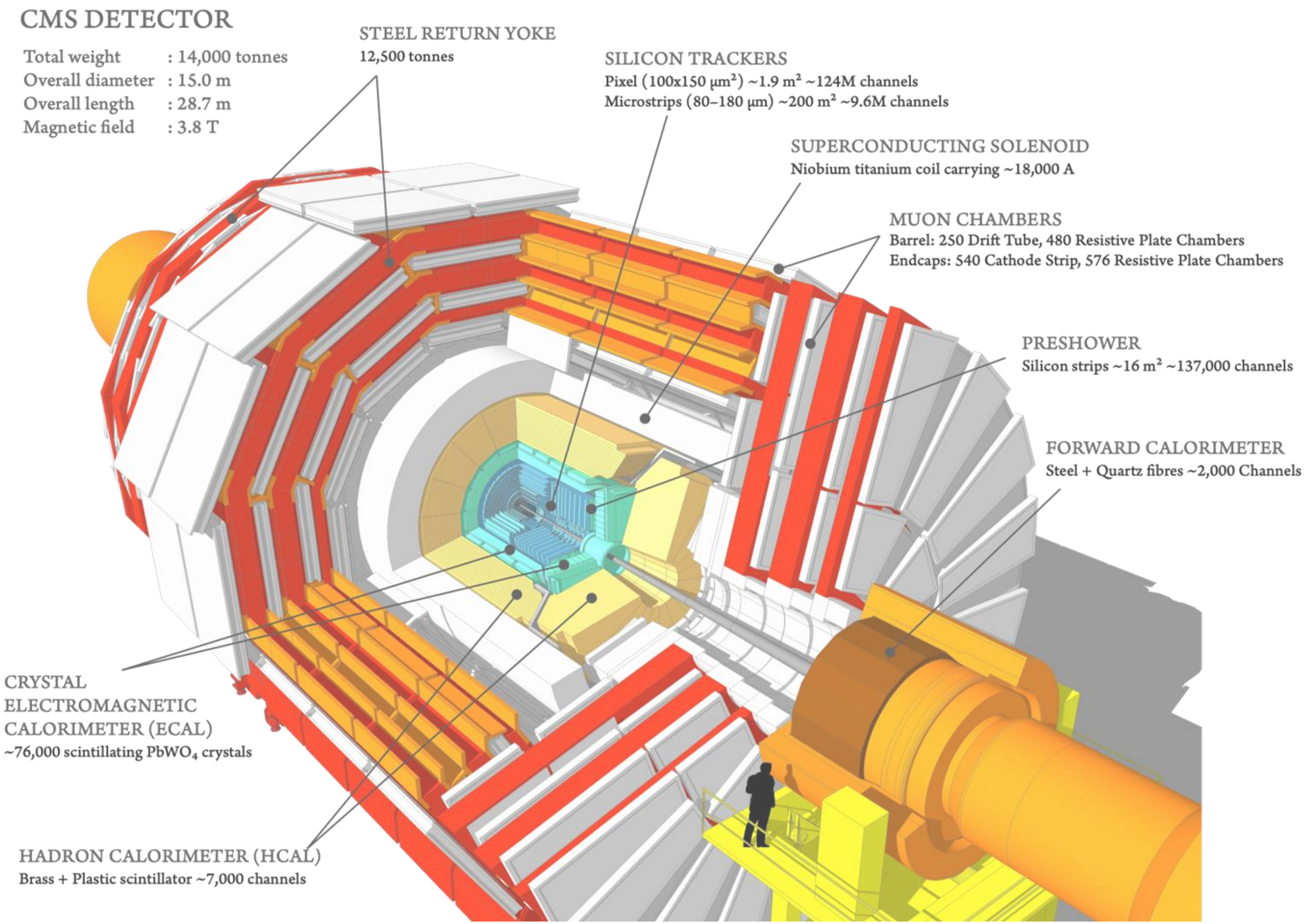


These two processes form the basis of the "electromagnetic shower"

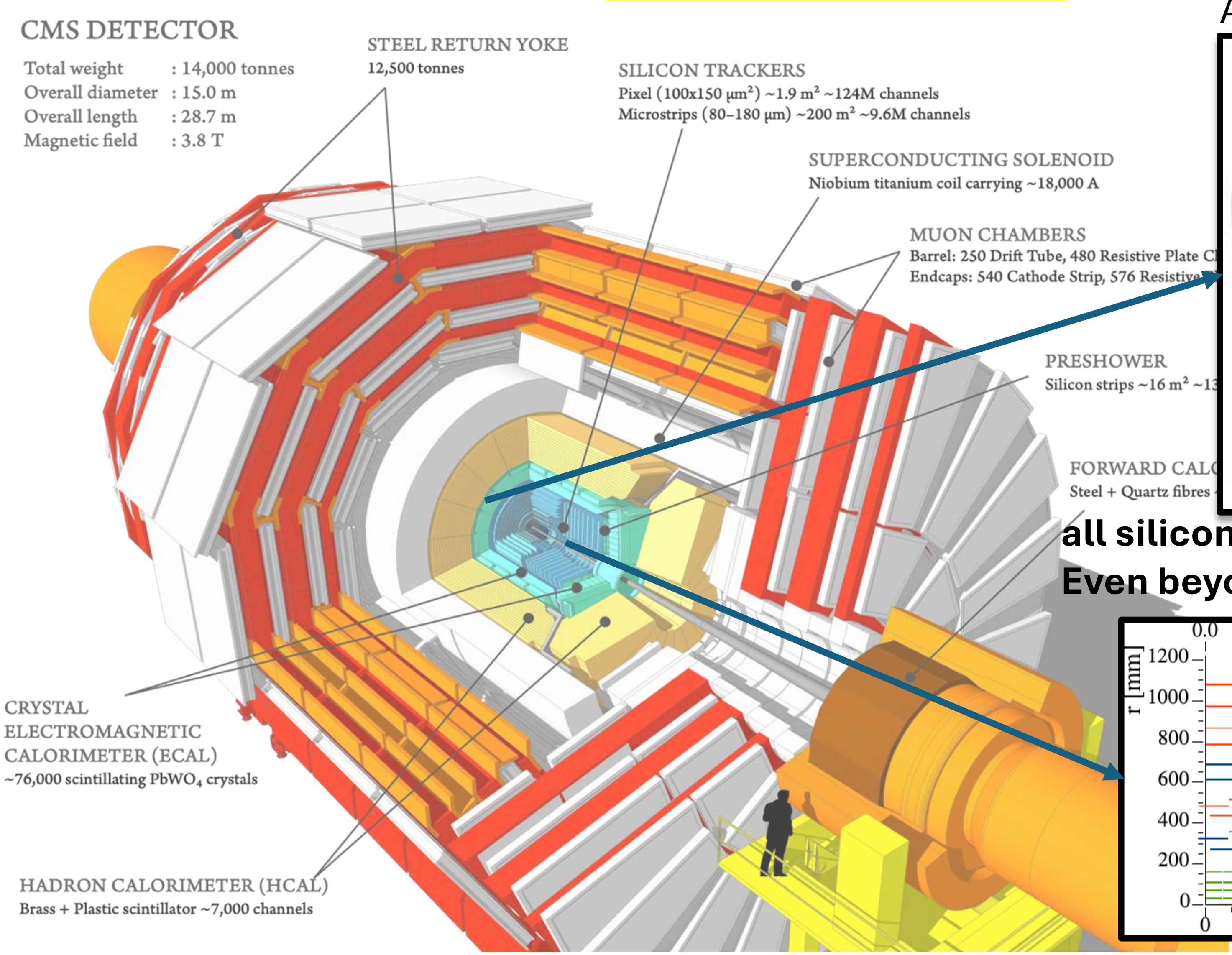




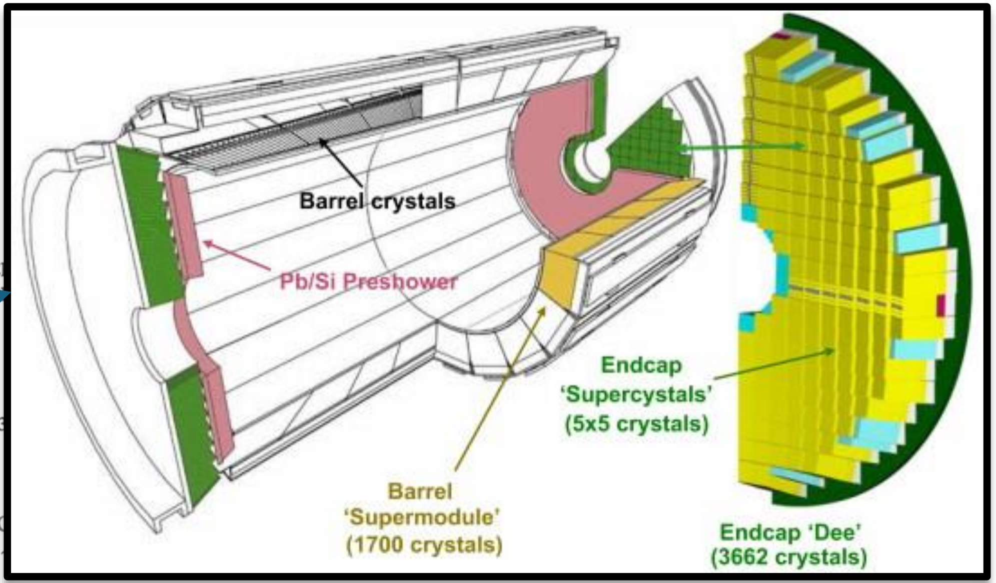
The detector



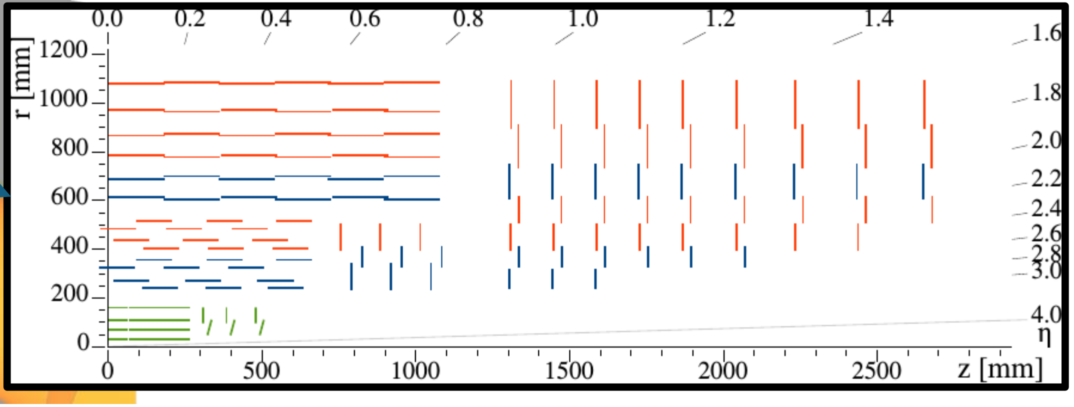
The detector



PbWO4 crystals **Hermetic Barrel & Endcap**
Additional preshower for endcap



all silicon, full coverage |η|<2.6
Even beyond, |η|>2.6, some coverage from the pixel detec



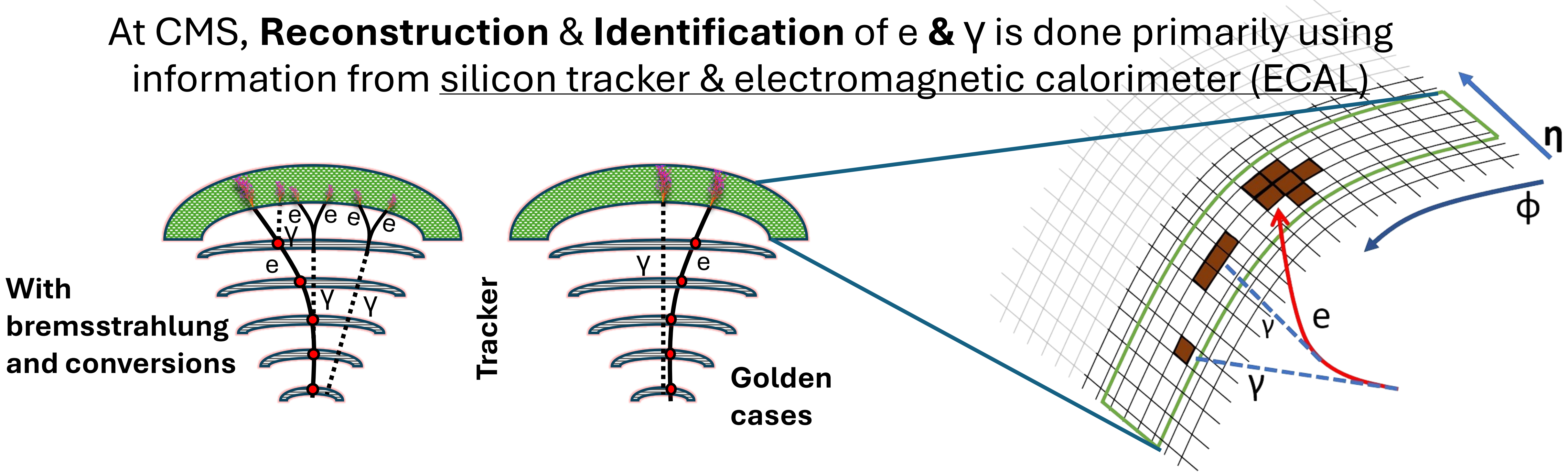
Electrons and Photons

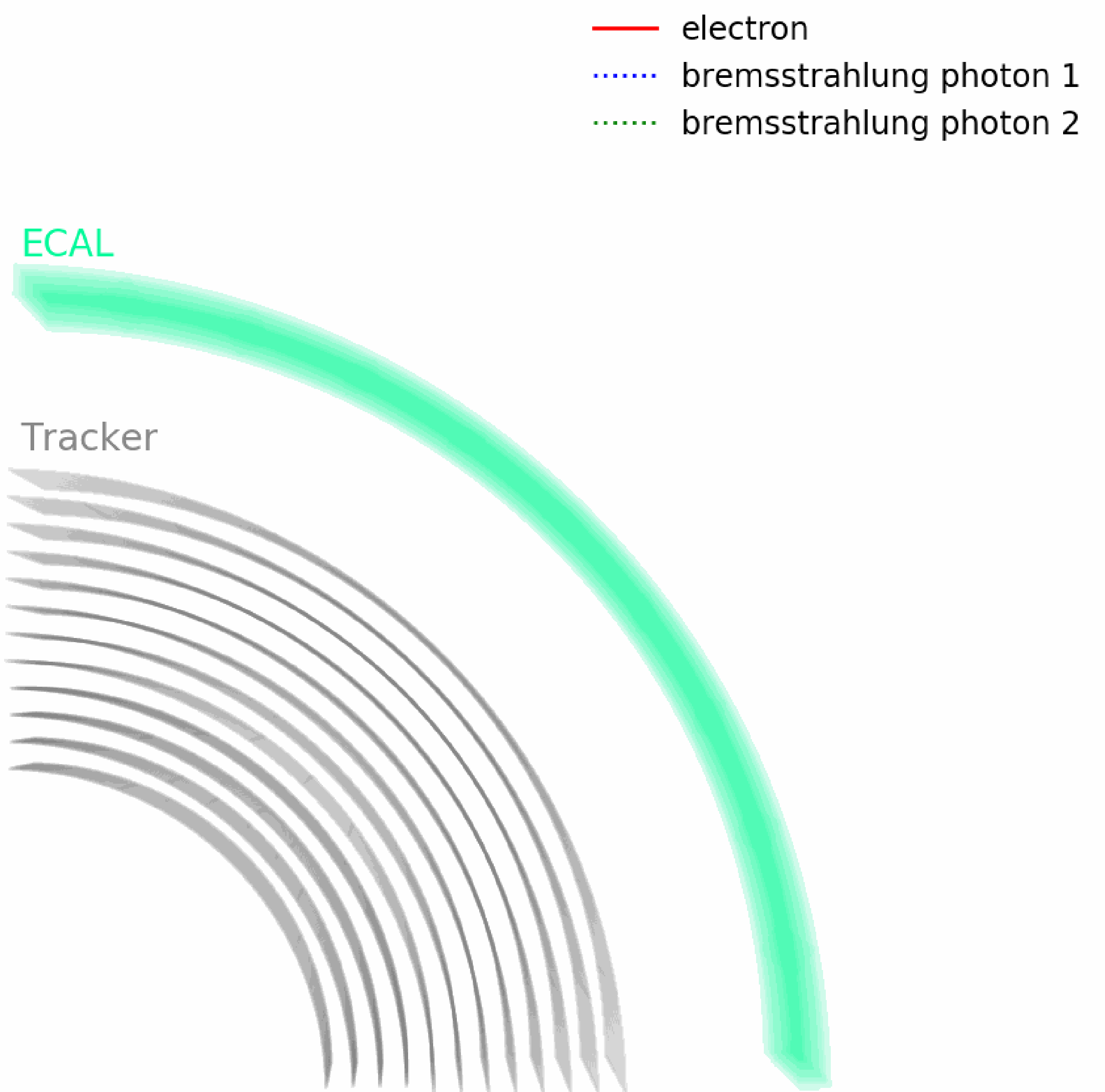
Electrons (e) **and photons** (γ) are critical to the experimental high energy physics program of the CMS experiment

e **&** γ appear in several new physics signatures

Also critical to standard model measurements

At CMS, **Reconstruction** & **Identification** of e **&** γ is done primarily using information from <u>silicon tracker & electromagnetic calorimeter (ECAL)</u>




electron
bremsstrahlung photon 1
bremsstrahlung photon 2
ECAL
Tracker

What does an electron look like?

Electrons can interact with the material before ECAL
→ emits bremsstrahlung Photon

✓ Leads to multiple energy deposits in ECAL

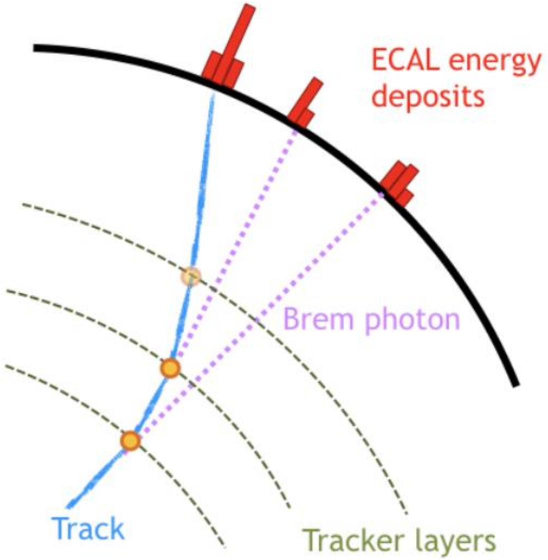
✓ Need to properly take into account energy of brems to reconstruct full energy of electron

Electron, after radiating a brem-photon in tracker, loses energy and starts to bend more in the magnetic field

✓ Electron track can have "kinks"

✓ Special tracking needed for electrons to take care of that

✓ Electron reconstruction algorithm takes into account all these complexities



Electron tracks vs other charged particles

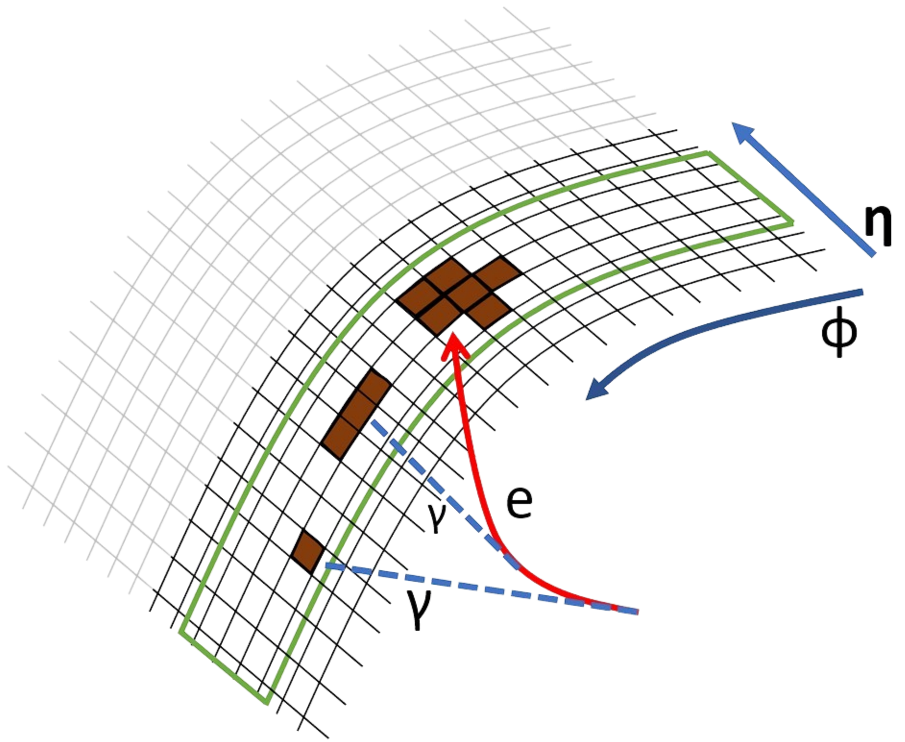
➤ For all other charged particle tracking, Kalman Filter is used-> Single Gaussians to model the probability for energy losses in the detector material when propagating from layer to layer

➤ Not a good choice for electrons, which have a high probability for large non-Gaussian energy losses due to bremsstrahlung

➤ Electrons use a dedicated tracking algorithm known as Gaussian Sum Filter (GSF). It is a non-linear extension of KF

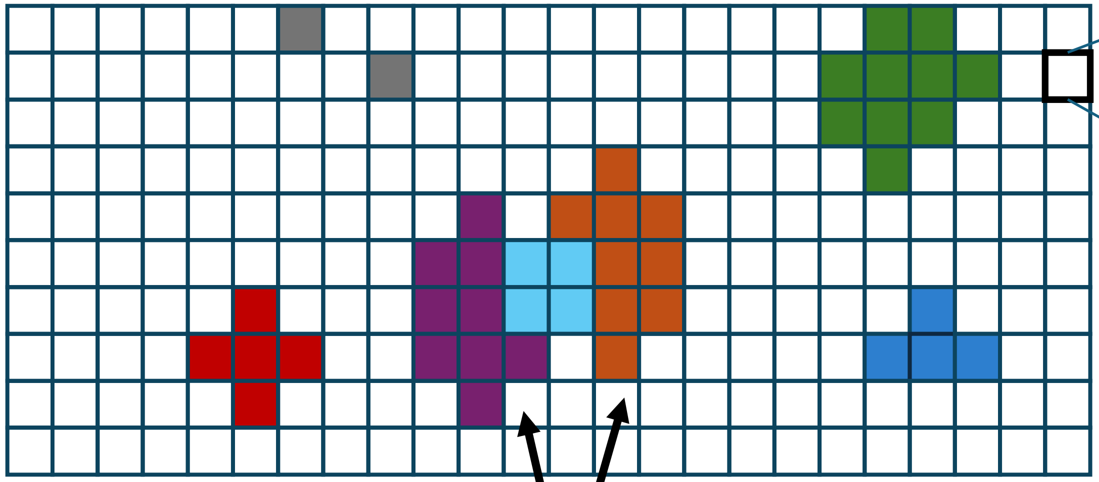
➤ GSF tracking takes into account radiative losses due to brem so we can measure pin and pout

➤ Use these measurements to be able to produce brem tangents, which are later used for refinement



Clustering

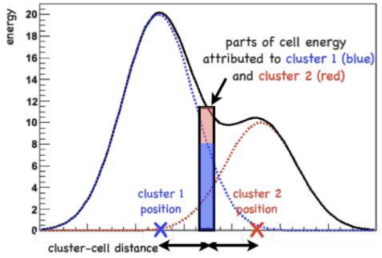
- Clustering of ECAL clusters

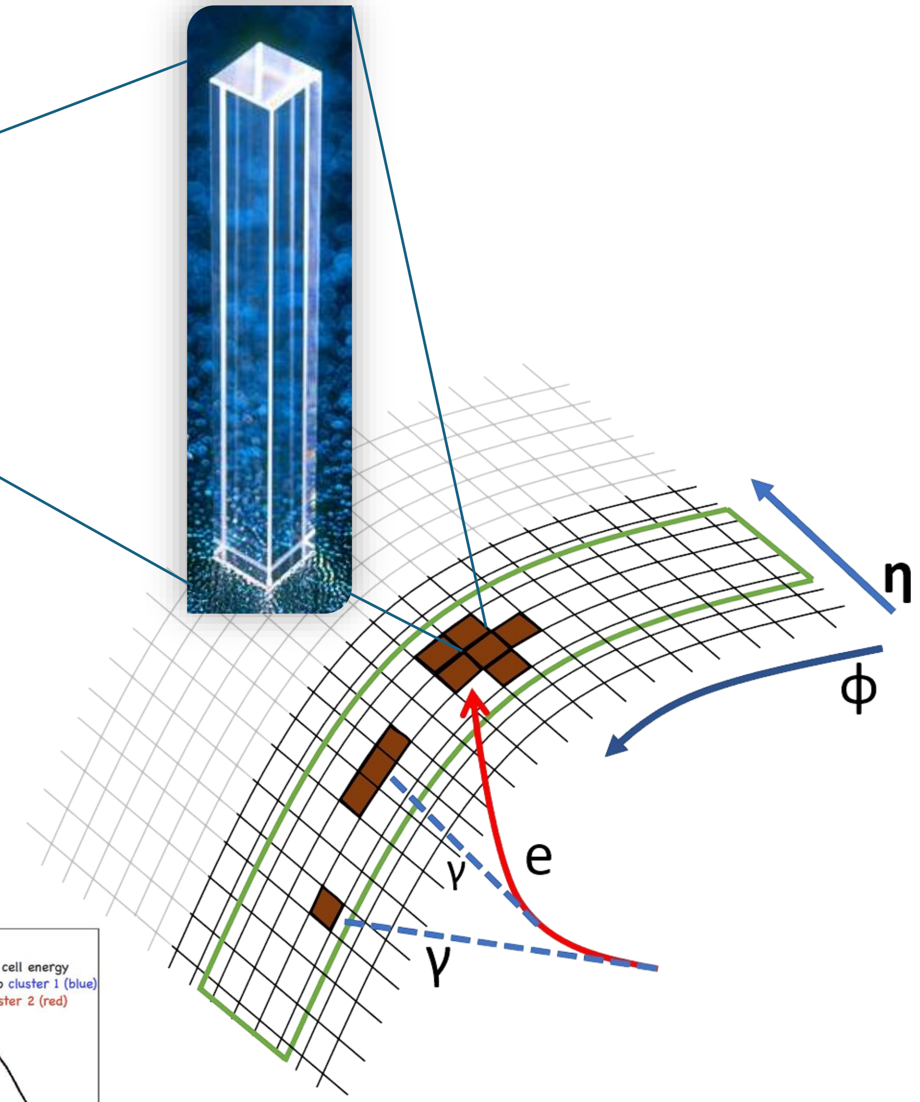


Clusters corresponding to electrons / photons

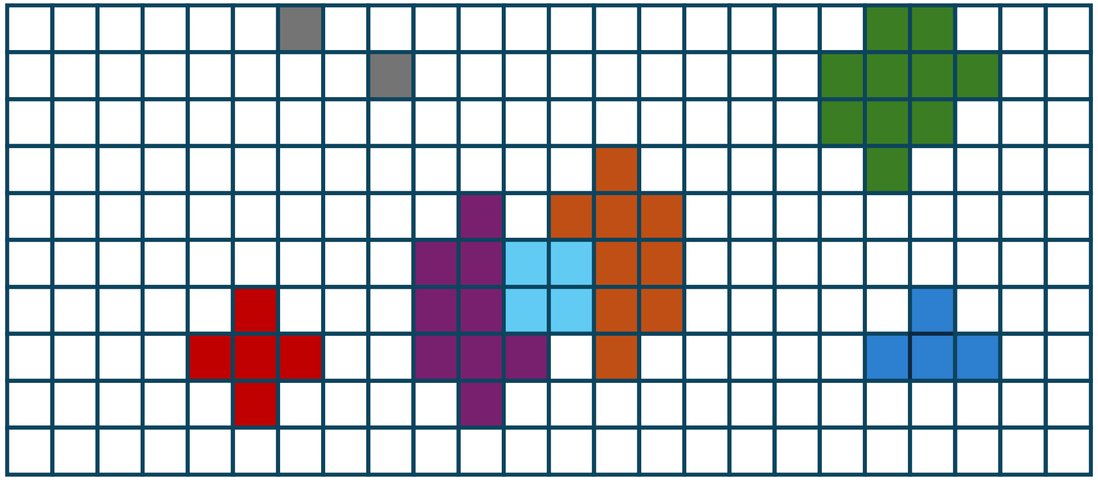
found 5 Clusters

these two clusters overlap, clustering algo shares energy of rec-hits between the two clusters according to a Gaussian energy profile, each gets a fraction of the rec-hit energy

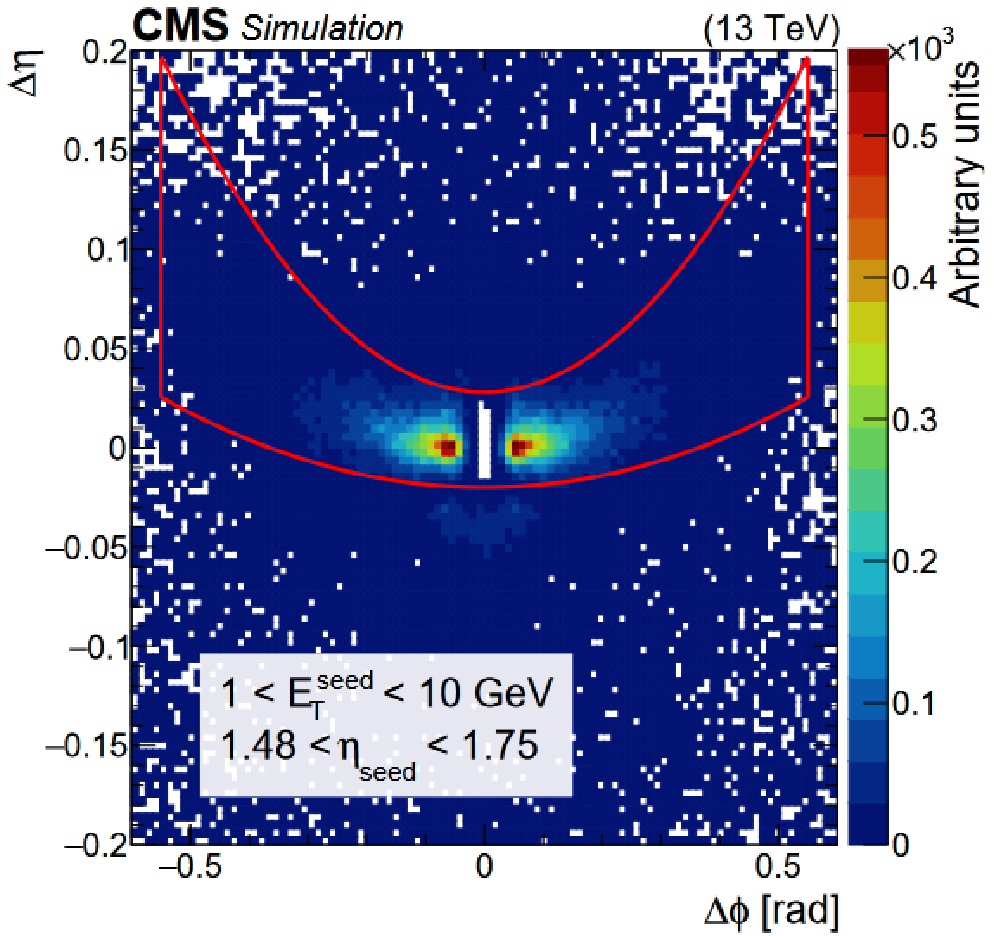




Clustering



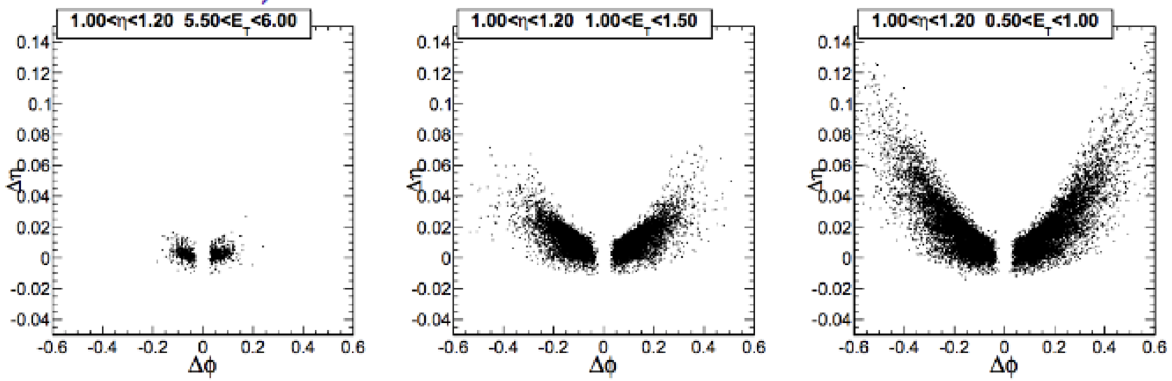
Moustache supercluster
A cluster of clusters



Clustering

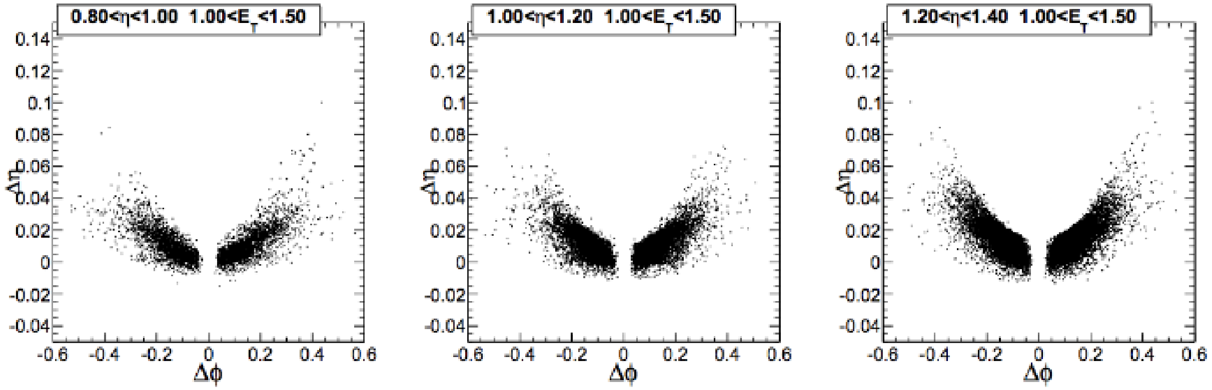
Clusters with smaller Et have larger allowed Δη/Δφ distances from seed.

Same eta, different Et

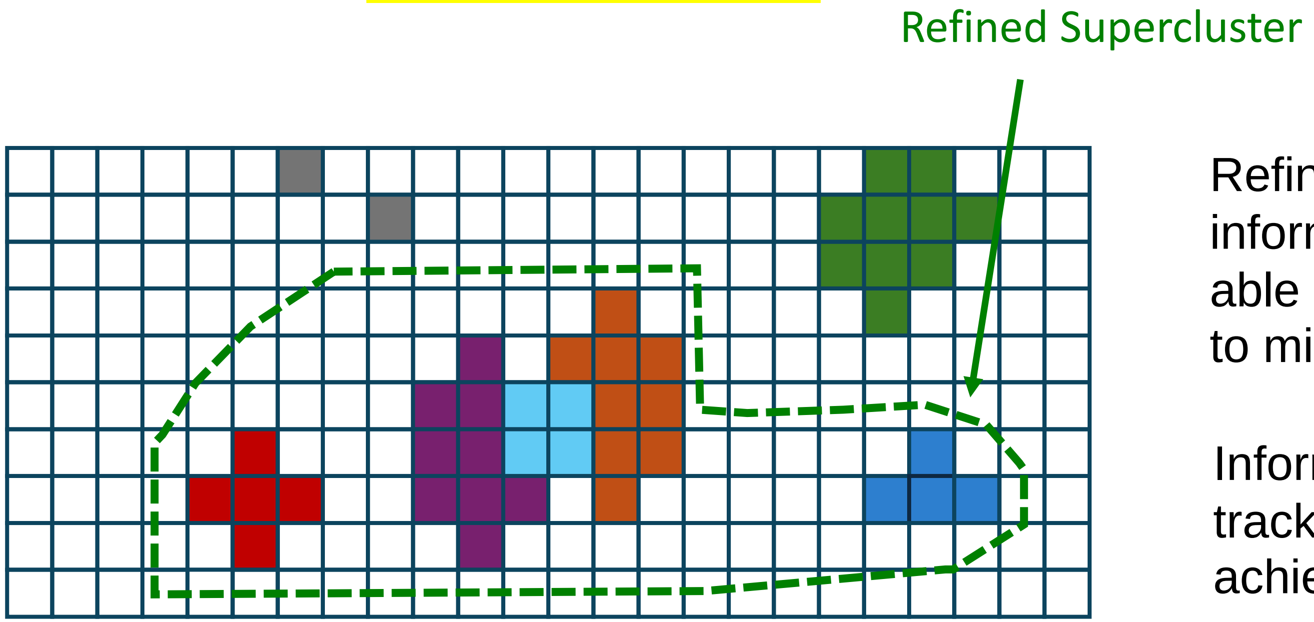


Allowed Δη/Δφ depends on cluster eta also.

Same Et, different eta



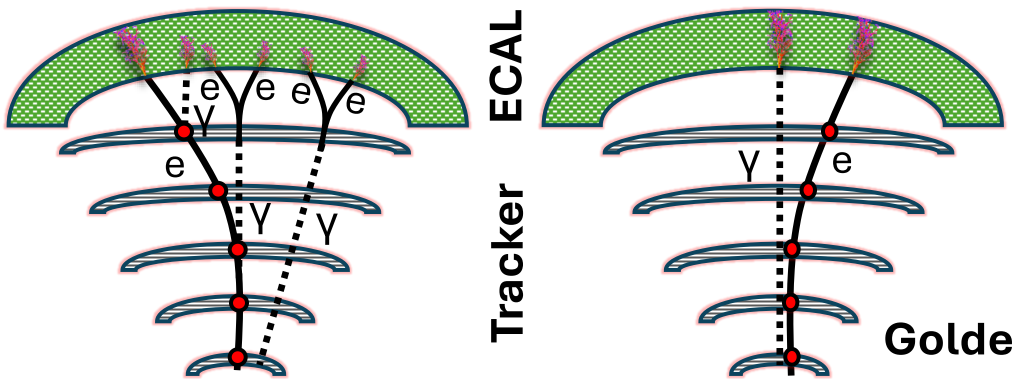
Clustering



Refined superclusters use the information from the tracker, to be able to link bremsstrahlung emissions to missed ECAL deposits

Information from clustering and tracking is used in tandem to achieve best resolution

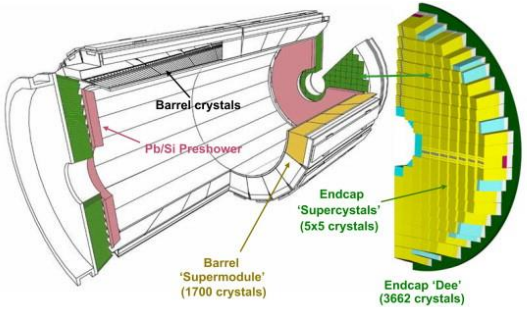
With bremsstrahlung and conversions

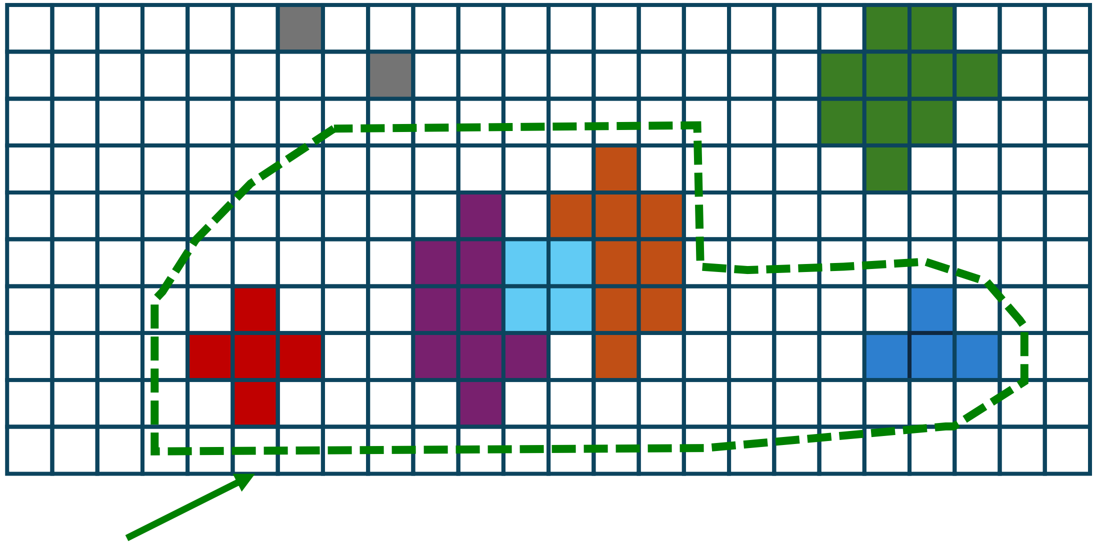


Golden cases

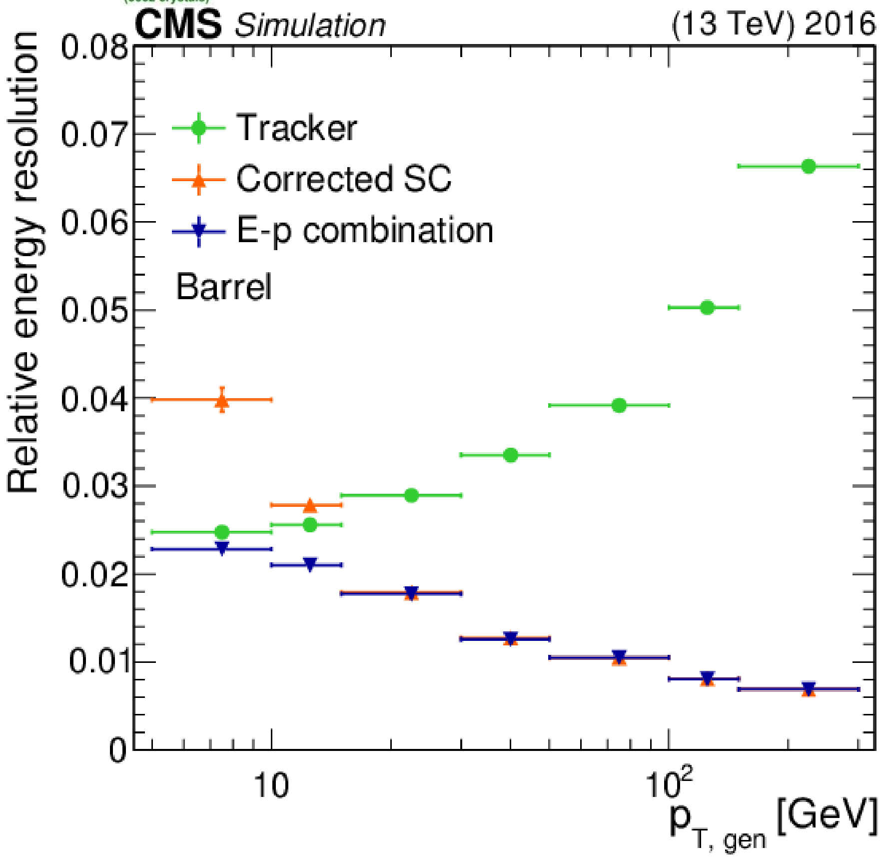
Clustering

- Clustering of ECAL clusters





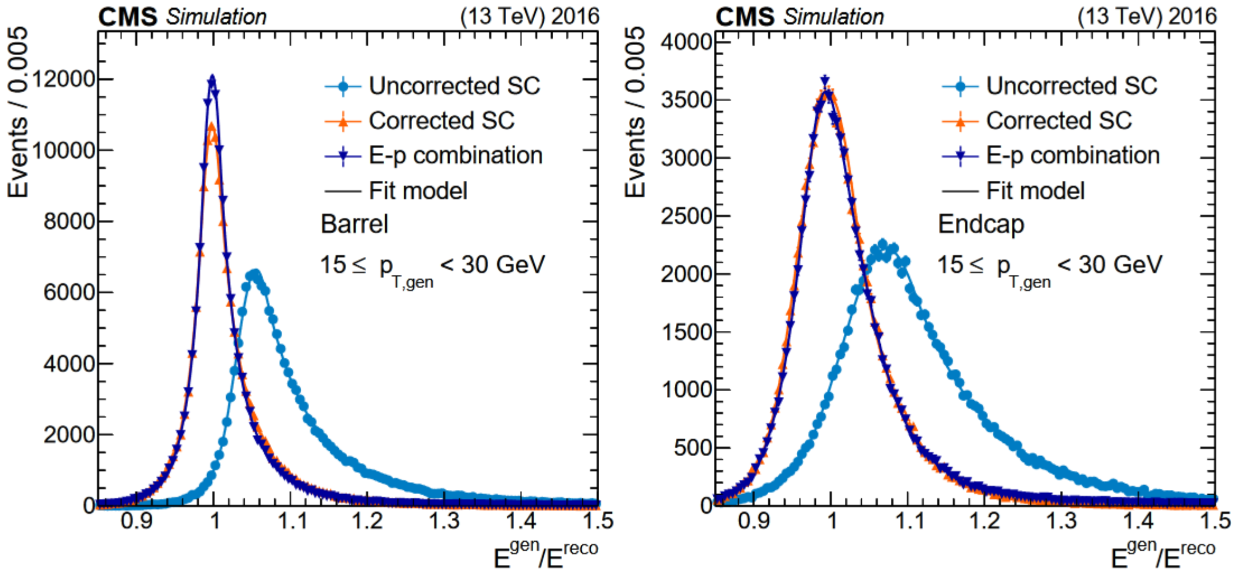
Refined superclusters use the information from the tracker, to be able to link bremsstrahlung emissions to missed ECAL deposits



There is also dedicated photon conversion recovery algorithm

Conversions can happen

- Several losses occur before electrons and photons deposit energy in the ECAL
- We calibrate the reconstructed energy back to the expected original energy using correction procedures
- Employ machine learning in tandem with algorithmic approaches
- Tracker information used for E-p combination



Conversions can happen

- Electrons can radiate hard photons in the tracker
- The brem photon can then convert to e+ e-
- It is not very rare (because of material budget in tracker)
- Sometimes, one of the converted e+ e- is not reconstructed.
- If converted brems are not correctly identified, those tracks can affect the global event description.
- That’s why we have a dedicated converted brem track finder

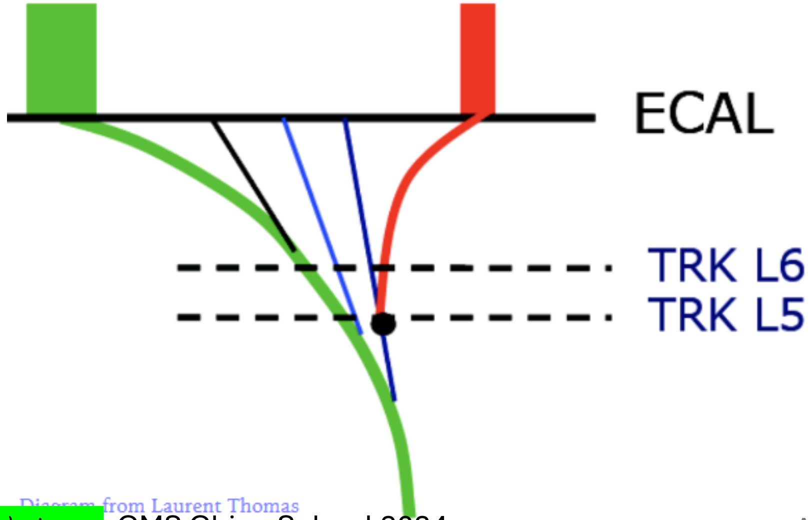
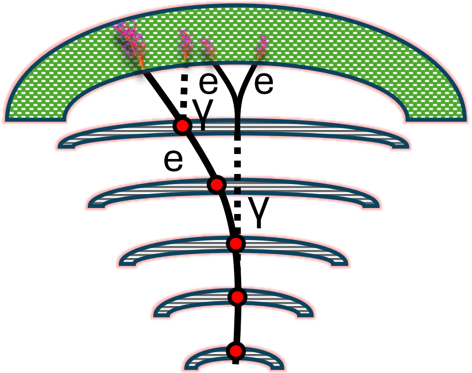


Diagram from Laurent Thomas

Conversions of photons within the “electron”

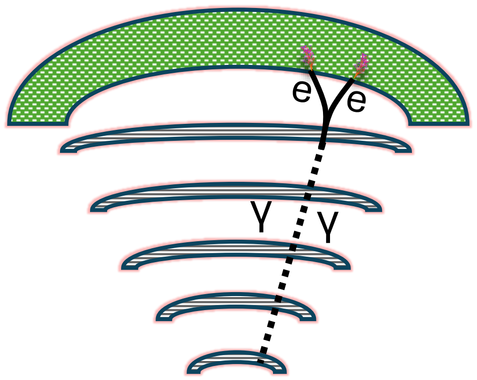


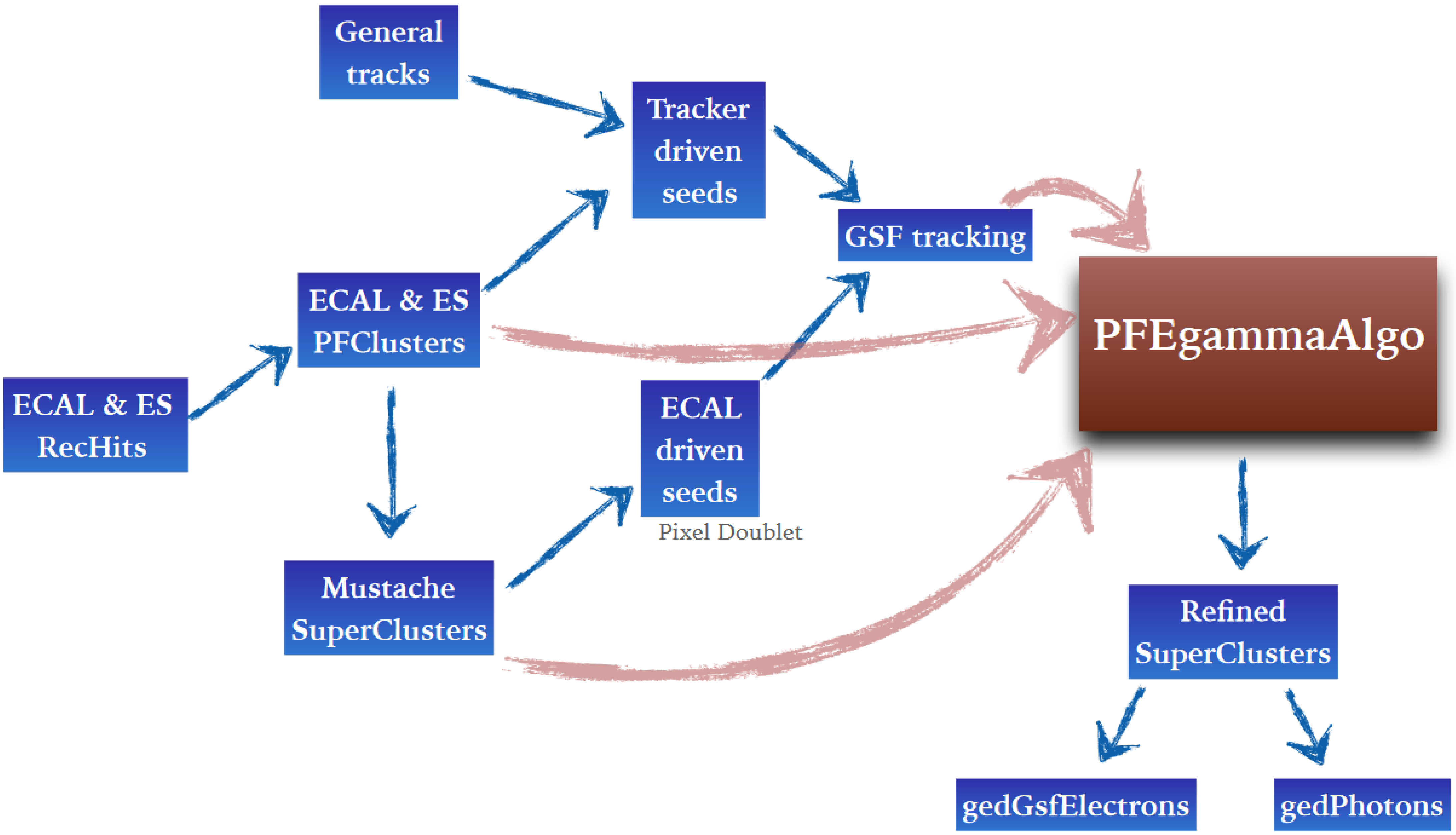
Conversions can happen

Conversions of photons

Conversion reconstruction is challenging:

- ➤ Displaced tracks to be reconstructed (issues with resolution and combinatorics)
- ➤ Trailing conversion leg may be very soft
- ➤ Conversion legs can radiate photon




General tracks
Tracker driven seeds
GSF tracking
ECAL & ES PFClusters
ECAL & ES RecHits
PFEgammaAlgo
ECAL driven seeds
Pixel Doublet
Mustache SuperClusters
Refined SuperClusters
gedGsfElectrons
gedPhotons

Identification

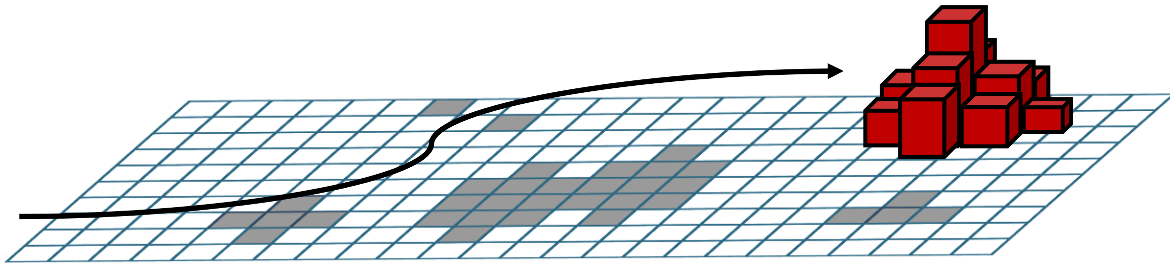
Two schemes are primarily used for identification:

- Via series of selections on various high-level properties
- Via machine learning based classifiers trained on these high level properties

What are high level properties?

- **Description of the electromagnetic shower**

(energy deposit pattern, lateral and longitudinal spread etc.)

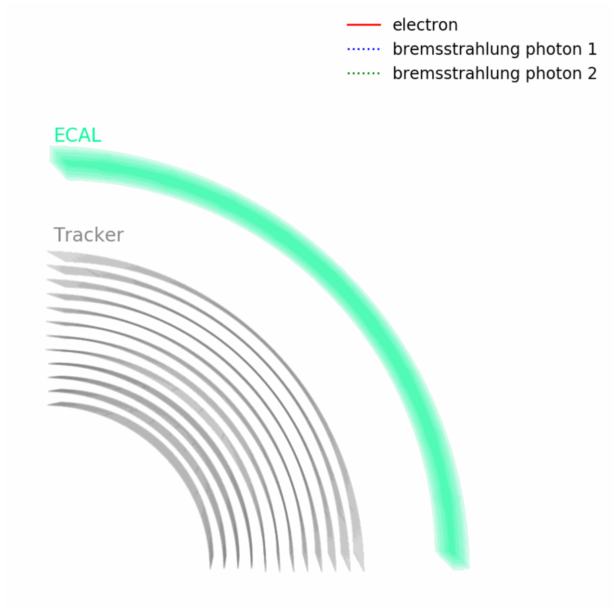


- **Tracking and clustering matching parameters**

(momentum trajectory extrapolated to ECAL considering the magnetic field etc.)

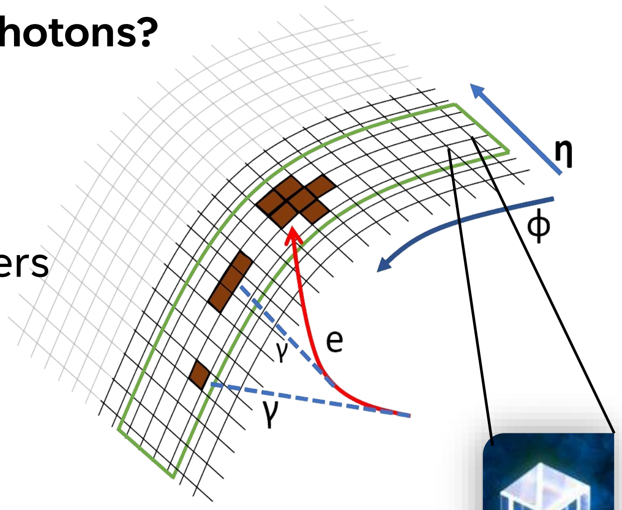
- **Quantification of isolation of these objects**

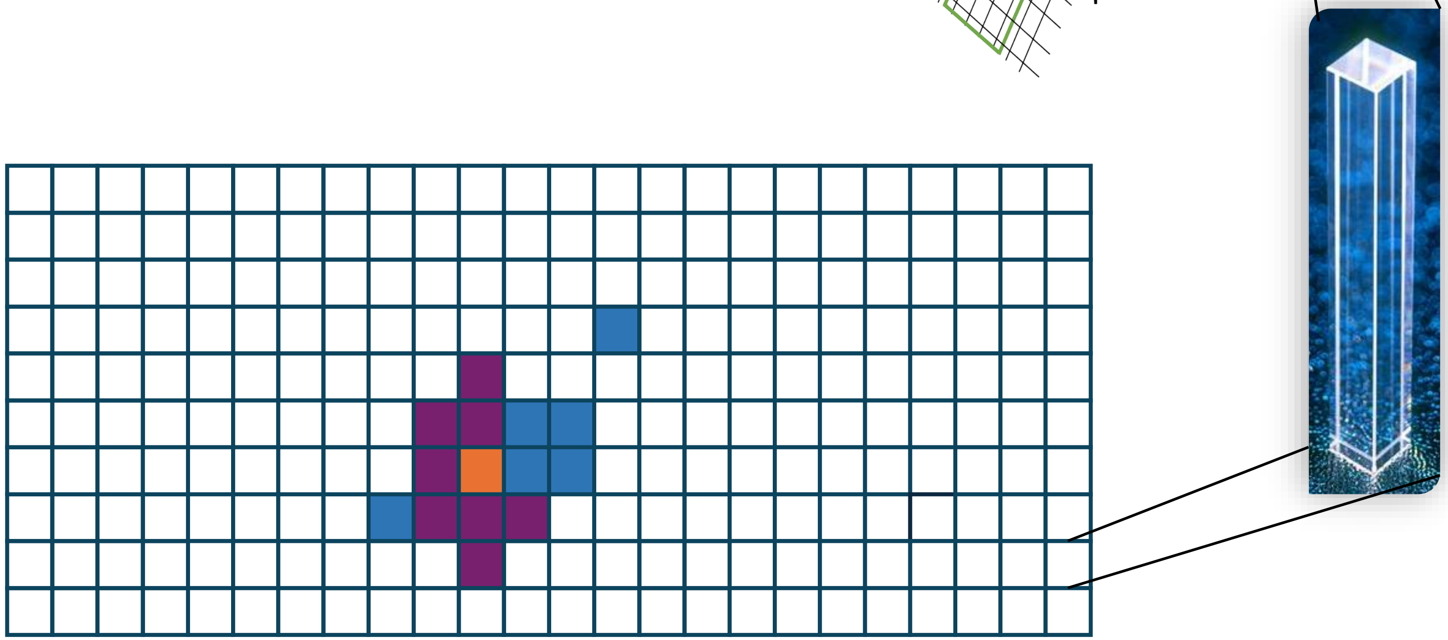
(Energy sums of crystals in ECAL in a defined area, leakage in HCAL etc.)



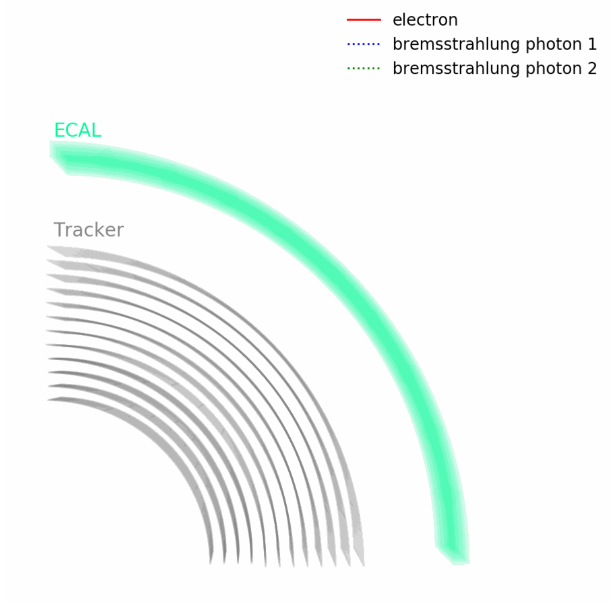
What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

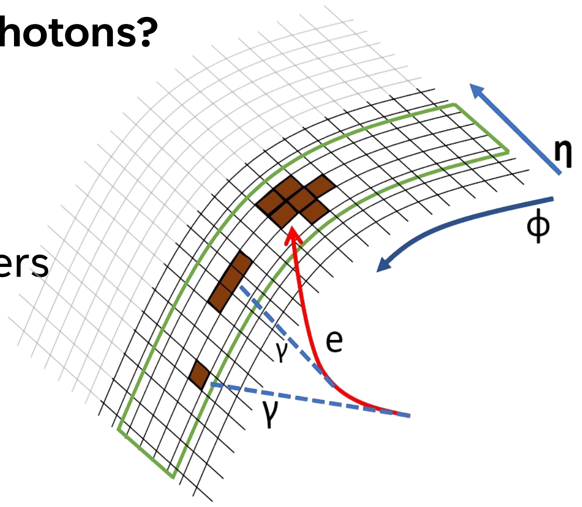


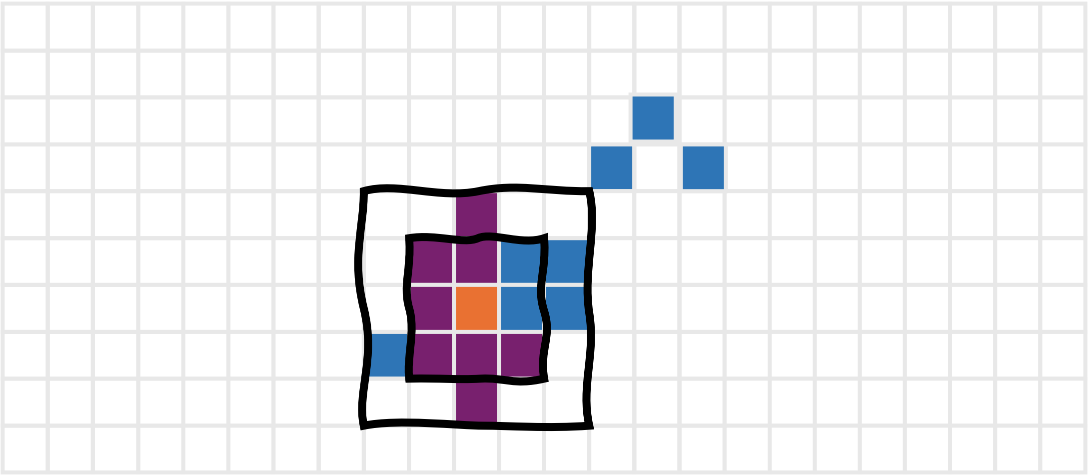


What can we use to identify electrons and photons?



- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

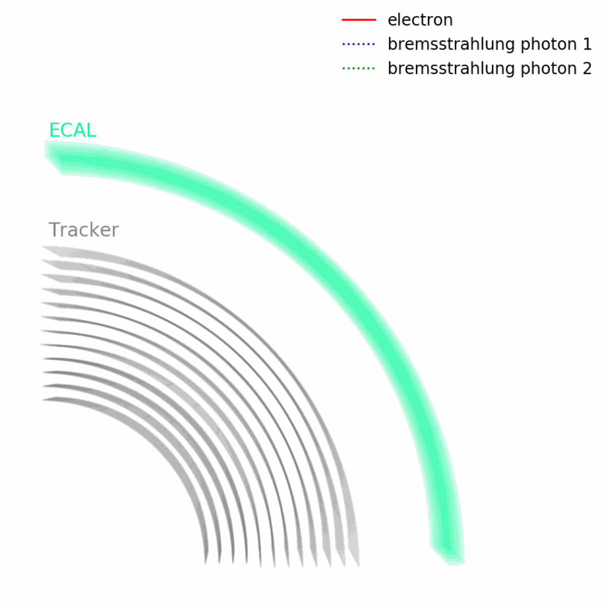


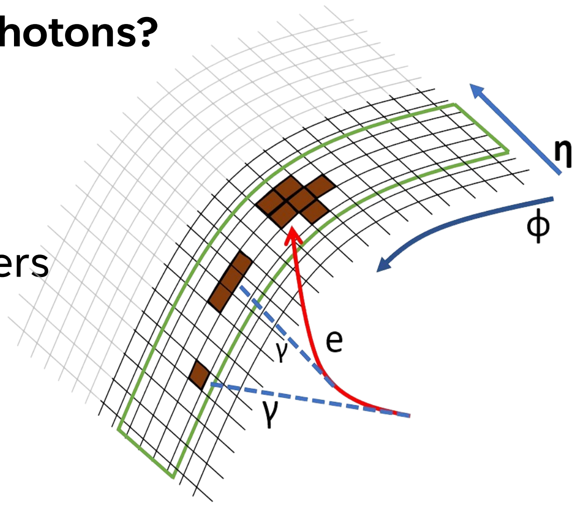


E_{sc} : Energy of supercluster
E_{2x2} : Energy contained in 2X2 crystals
E_{3x3} : Energy contained in 3X3 crystals
E_{5x5} : Energy contained in 5X5 crystals
E_{1x5} : Energy contained in 1X5 crystals
$d\eta_{sc}$: η width of supercluster
$d\Phi_{sc}$: Φ width of supercluster

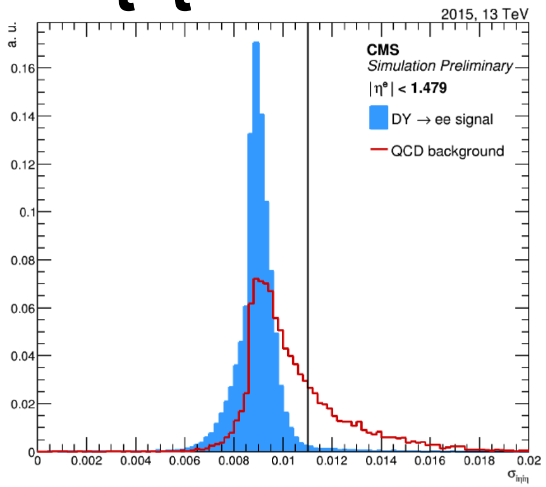
What can we use to identify electrons and photons?

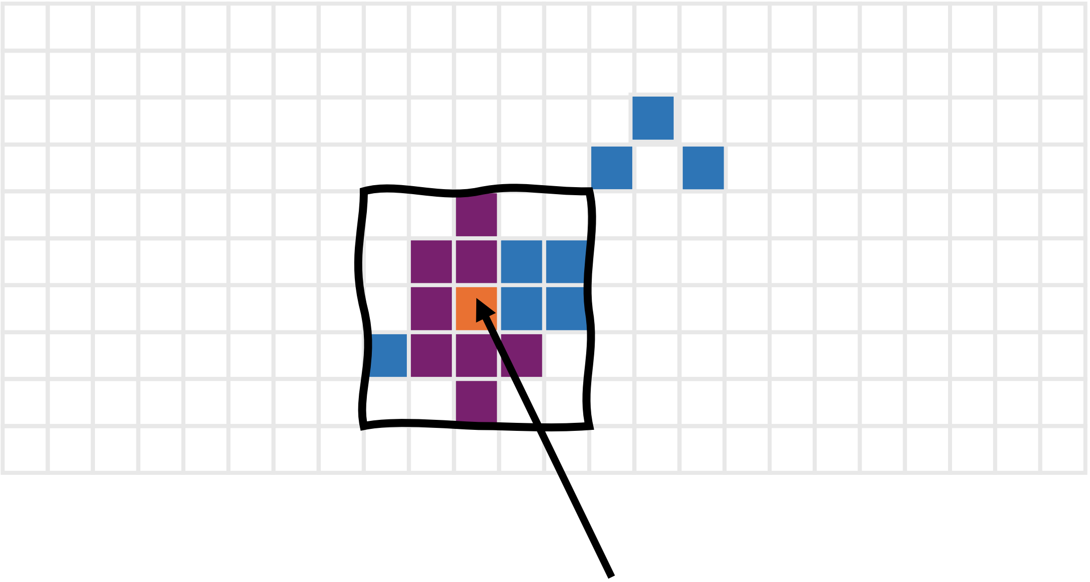
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects





$\sigma_{i\eta i\eta}$

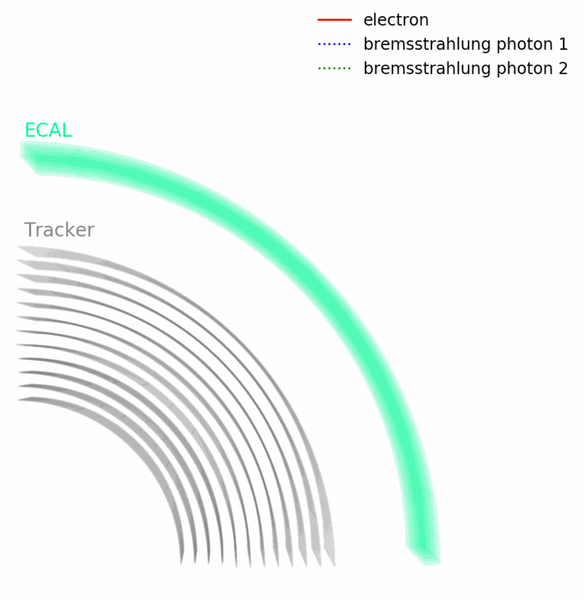


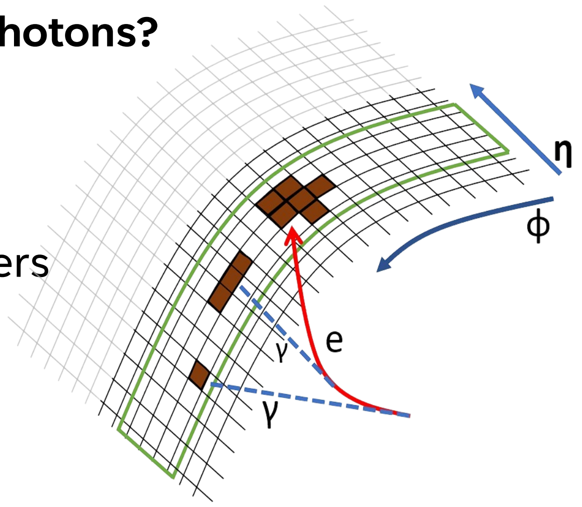


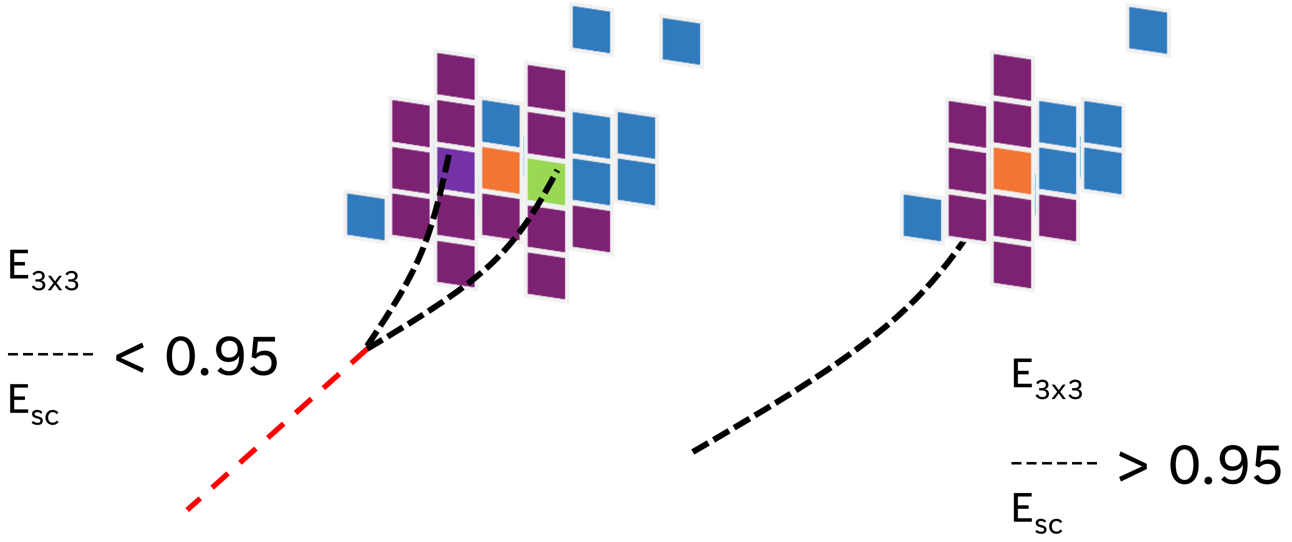
Most energetic crystal of 5x5 array

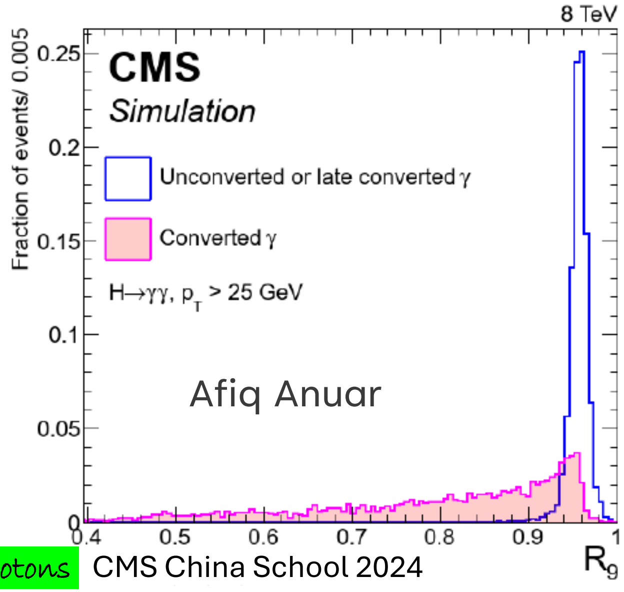
What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects



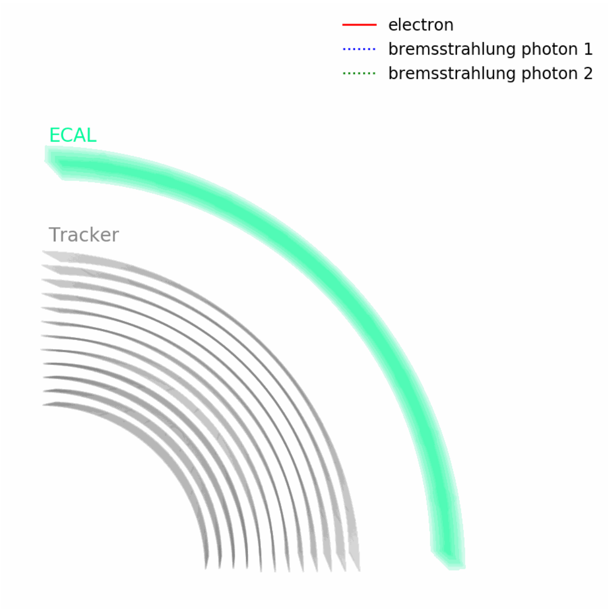




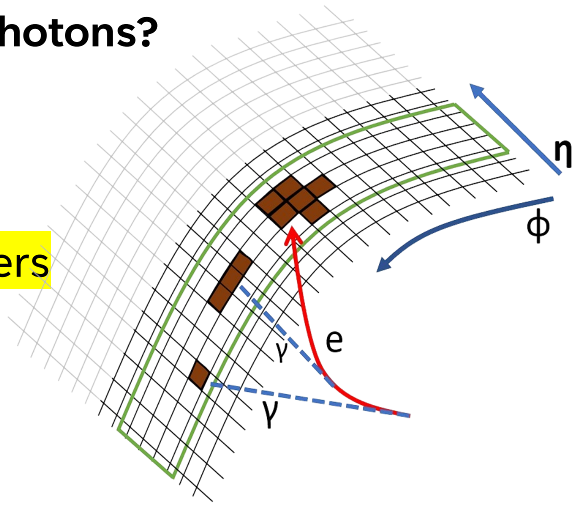


5x5 matrix contains ~96% (~97%) of unconverted photon energy in EB (EE)

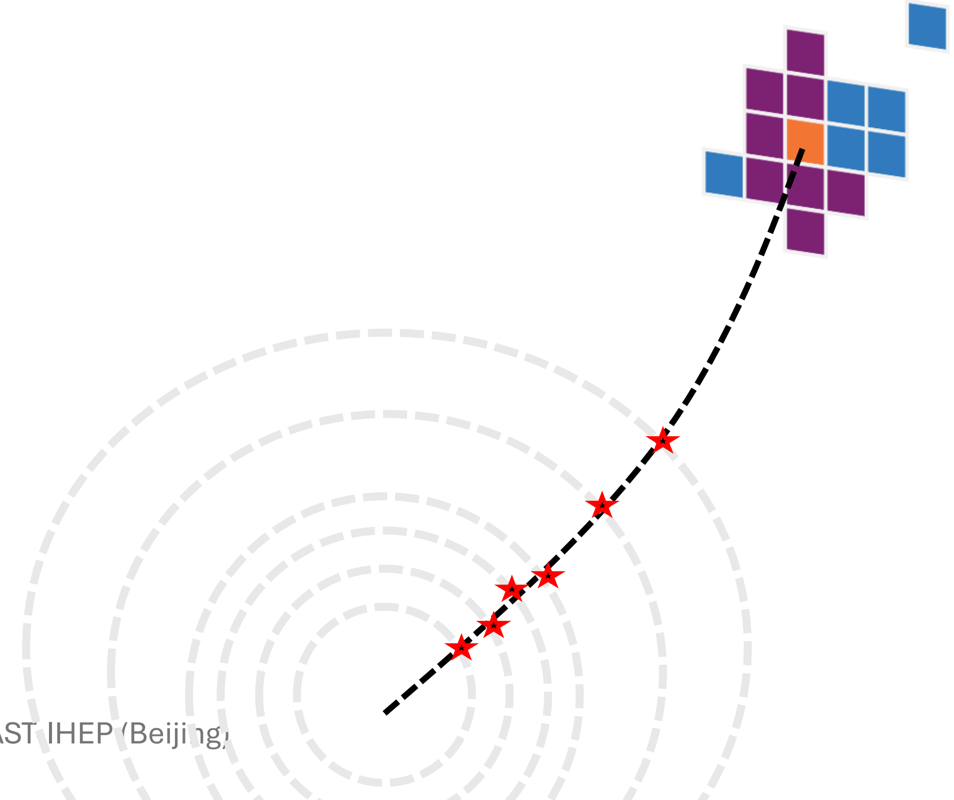
What can we use to identify electrons and photons?



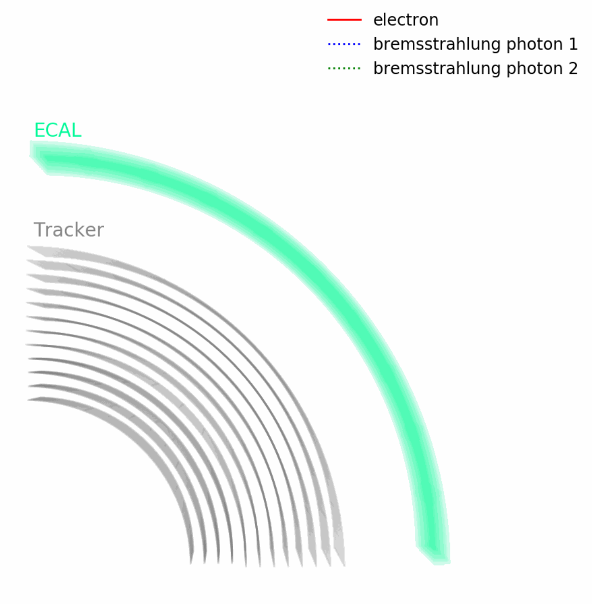
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects



N^{gsf}_{hits} : Hits in the “gsf” track
N^{kf}_{hits} : Hits in the “kf” track
E/p : Energy of supercluster/ momentum
χ^2: Track quality

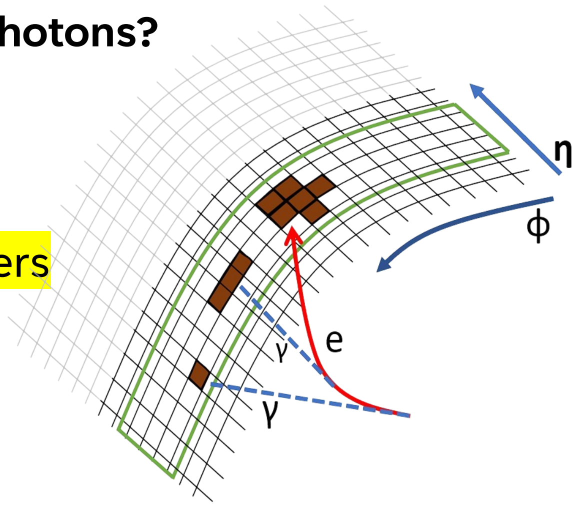


Not just for electrons,
Even for photons, good ones should not have a “track at all”
Or would have converted?



What can we use to identify electrons and photons?

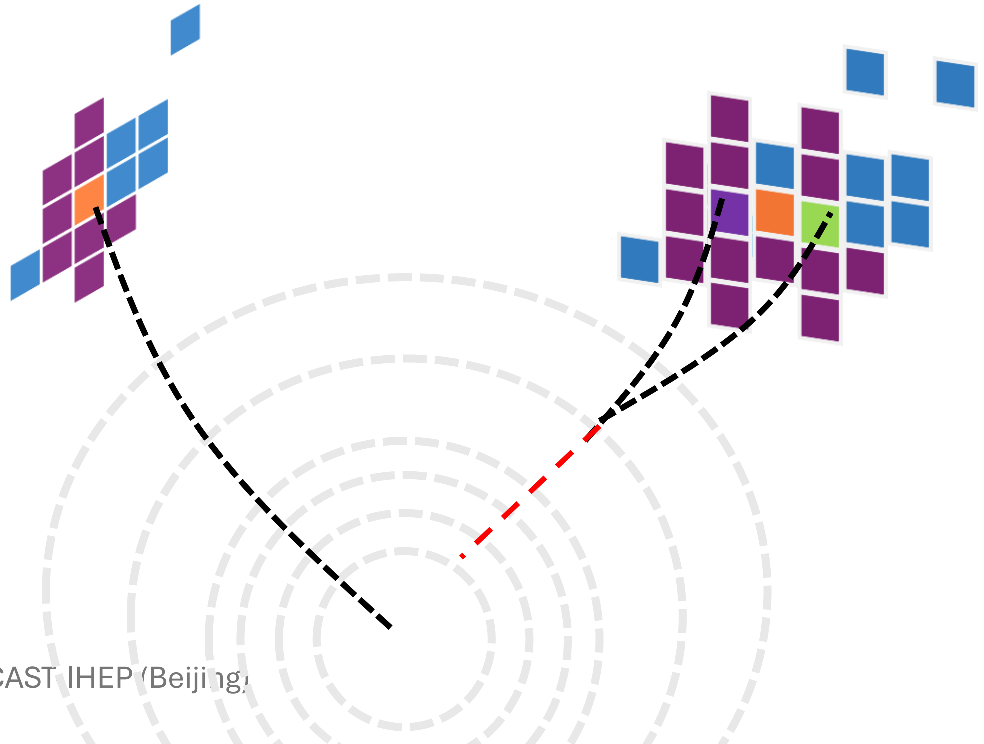
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

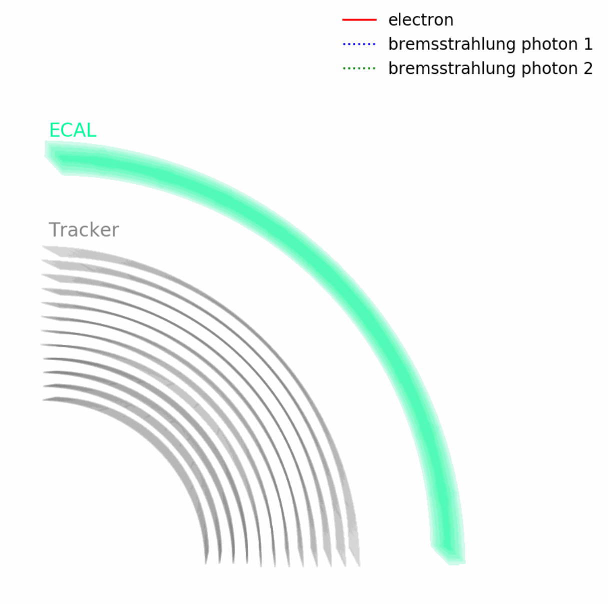


Conversion ID variables

How to differentiate between electrons in photon conversions from prompt electrons?

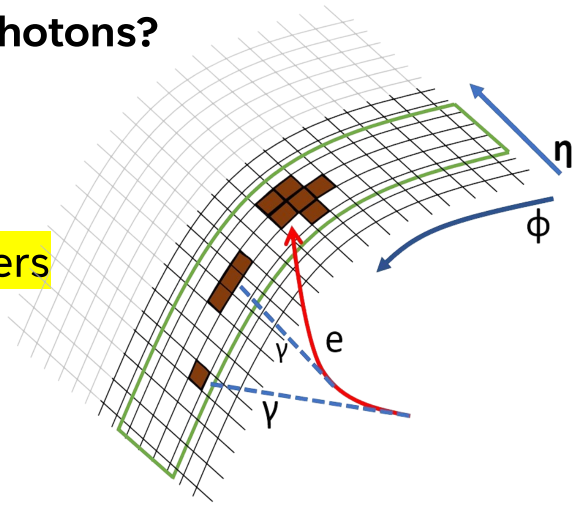
Conversion Safe Electron Veto
Pixel Veto





What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

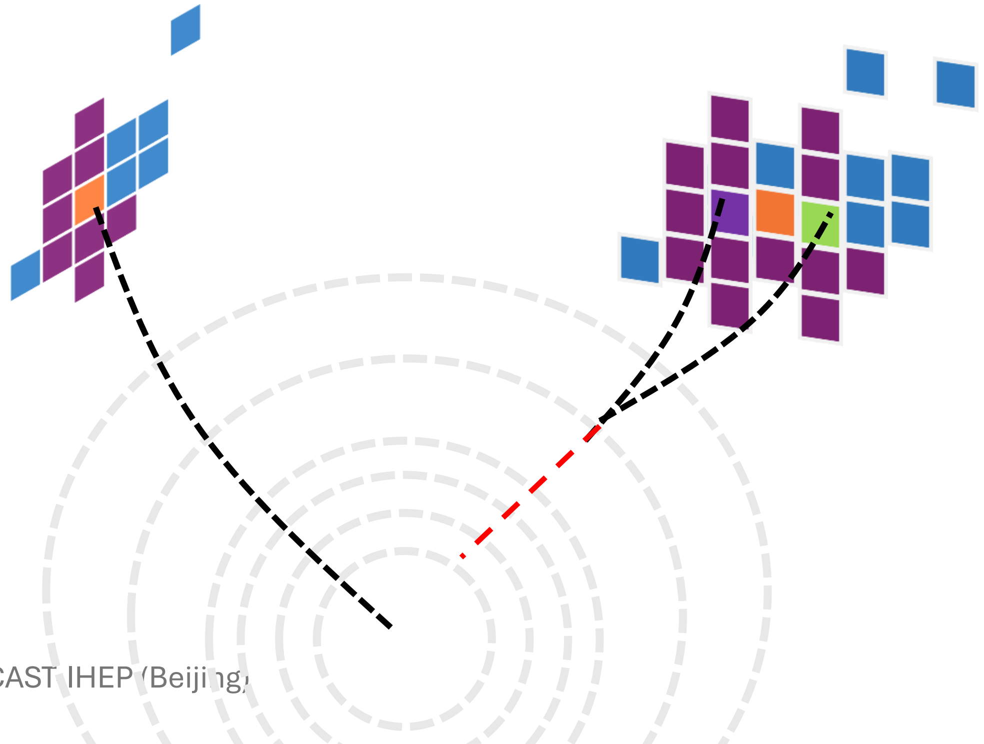


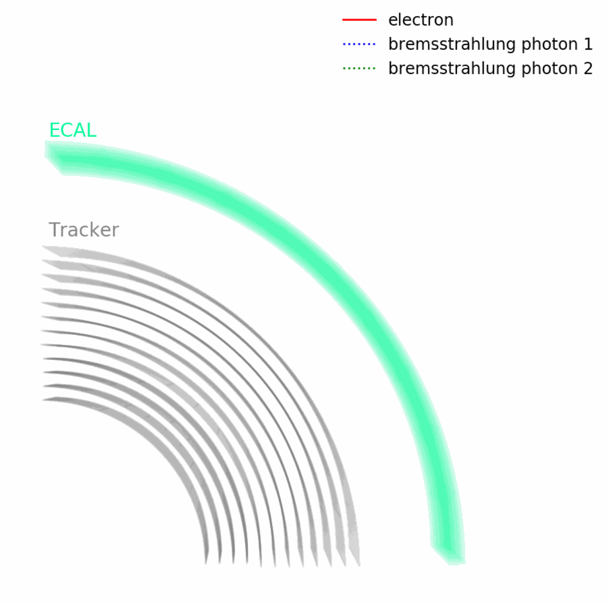
Conversion Safe Electron Veto

If a secondary vertex is found, this is not an electron!

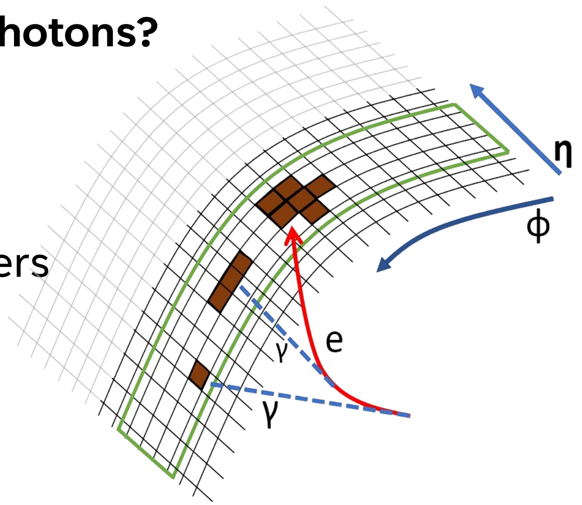
Pixel Veto

Track in pixel detector, this is not a photon



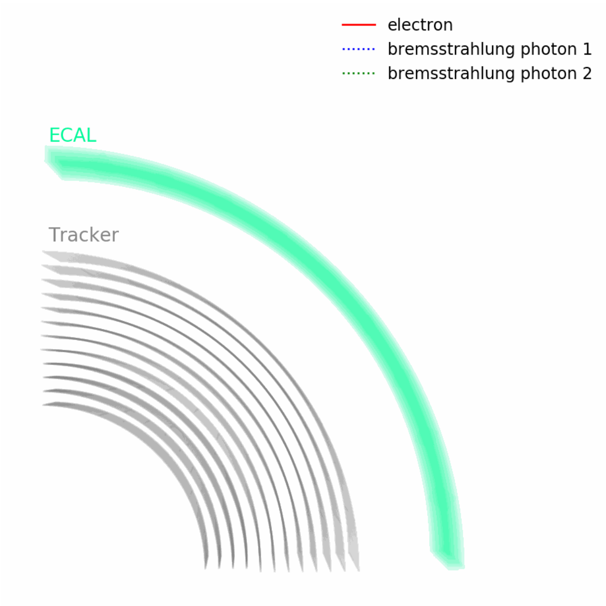


What can we use to identify electrons and photons?

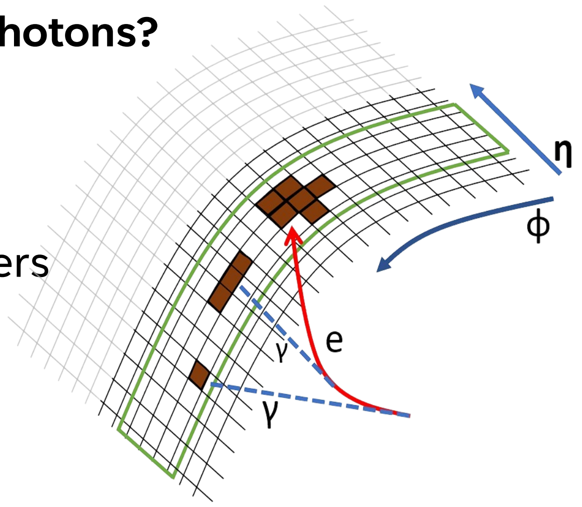


- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

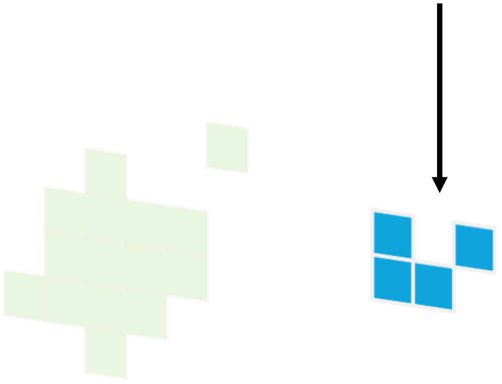
What can we use to identify electrons and photons?



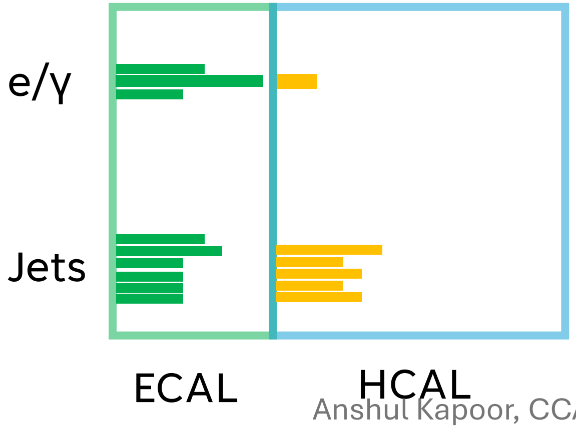
- Description of the EM shower shape
- Tracking and clustering matching parameters
- ==Quantification of isolation of these objects==

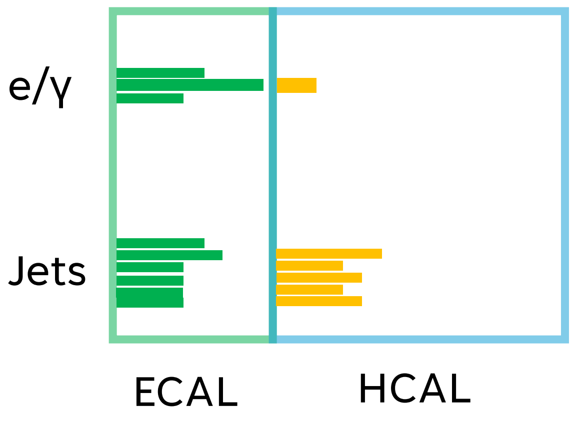


Leakage in HCAL



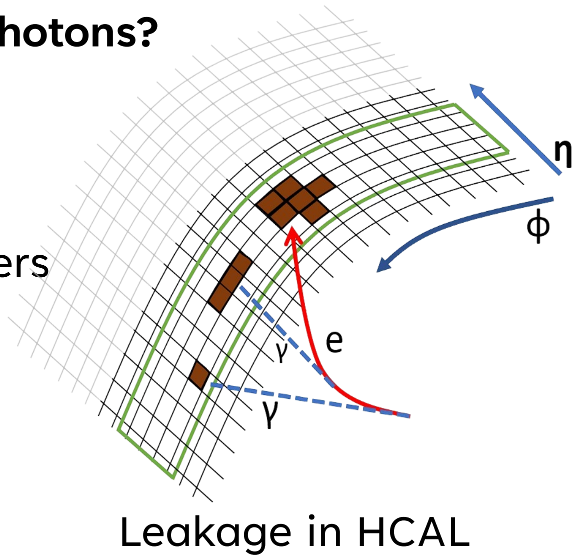
H/E : Energy leaked into HCAL / Energy in ECAL



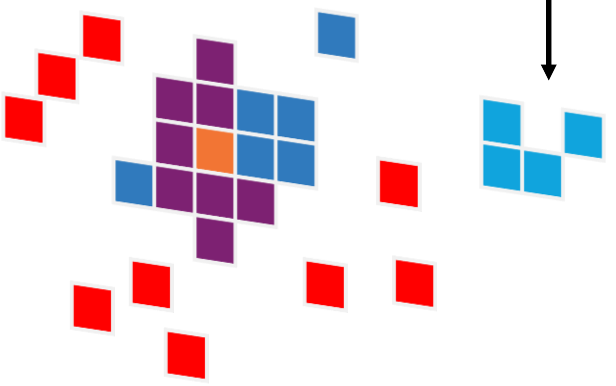


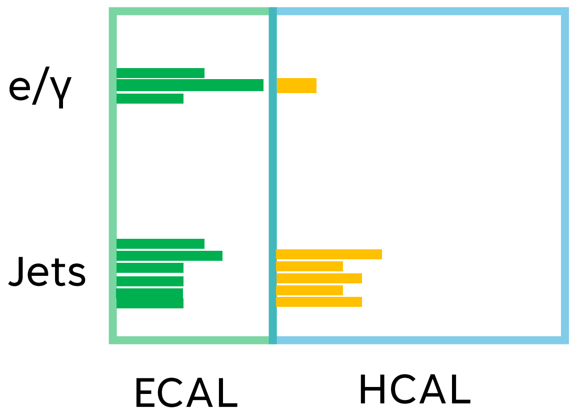
What can we use to identify electrons and photons?

- ➢Description of the EM shower shape
- ➢Tracking and clustering matching parameters
- ➢Quantification of isolation of these objects



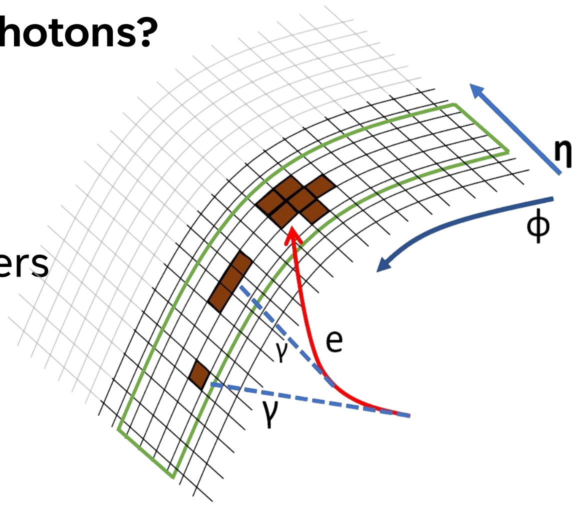
H/E : Energy leaked into HCAL / Energy in ECAL
Isolation: Other stuff around it?





What can we use to identify electrons and photons?

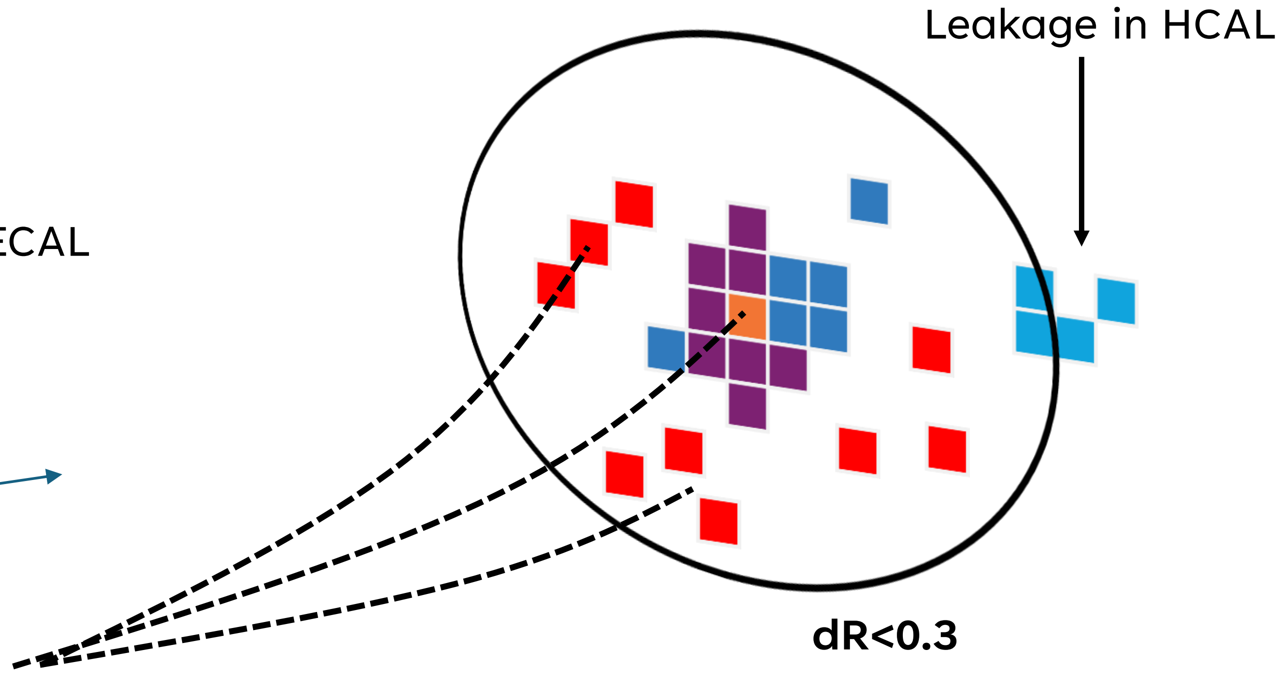
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

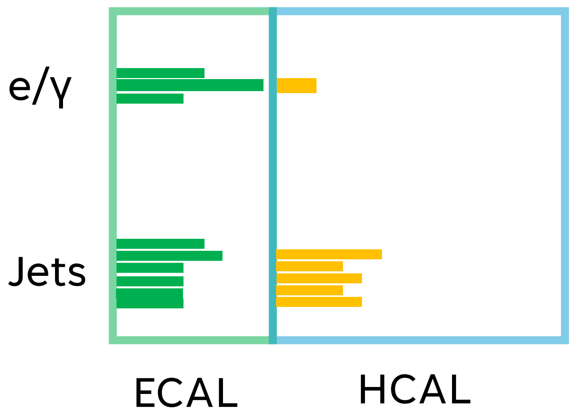


H/E : Energy leaked into HCAL / Energy in ECAL
Isolation: Other stuff around it?

Information provided by PF Is used here:

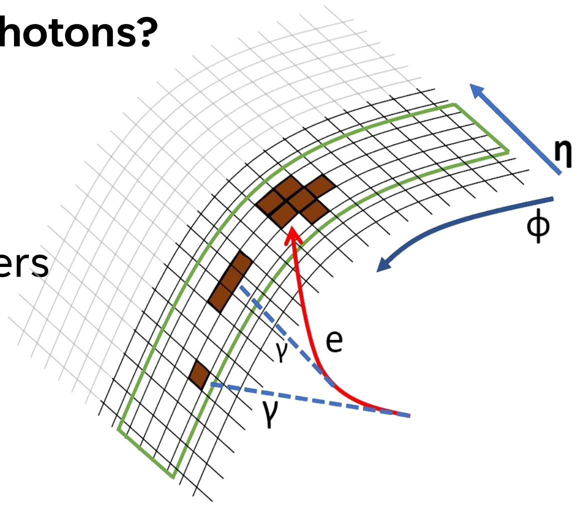
PF Cluster+Tracker isolations
HCAL clusters
ECAL clusters
Tracks





What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

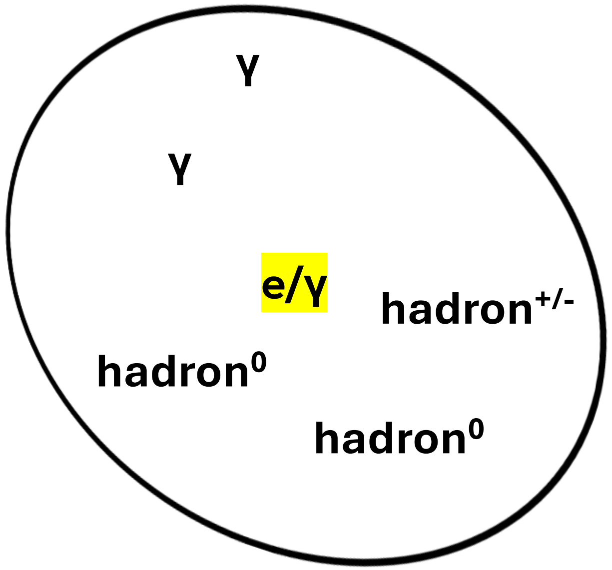


Correlated

H/E : Energy leaked into HCAL / Energy in ECAL
Isolation: Other stuff around it?

Information provided by PF Is used here:

PF Particle Isolations
Charged hadrons
Neutral hadrons
Photons



dR<0.3

Offline Analysis
Central E/Gamma Identification

Electron case as an example:

What should an electron look like? Here are some key indicators to consider:

•Isolated cluster: A prompt electron will result in an isolated cluster

•Electromagnetic cluster: characterized by low H/E

•EM shower typically appears as a concentrated cluster

•E_{SC} /pT: a track and cluster produced by the same particle, indicating a high likelihood of being an electron

•Matching cluster-track: comparing the angle of the cluster to the track can help distinguish against Bremsstrahlung radiation.

Offline Analysis

~70/80/90% signal efficiency

Tight

Medium

Loose

Cut based

Veto

Offline Analysis
Central E/Gamma Identification

Electron case as an example

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

Offline Analysis
Central E/Gamma Identification

Electron case as an example

Efficient against:
All hadronic background

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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Electron case as an example

Efficient against:
Jets

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
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Electron case as an example

Efficient against:
Pions, photons

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
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Tight

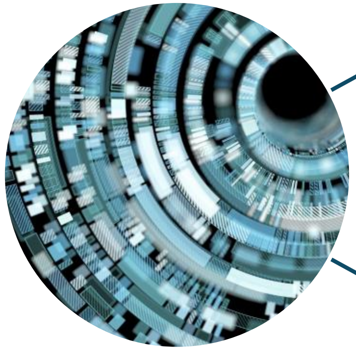
Medium

Loose

Cut based

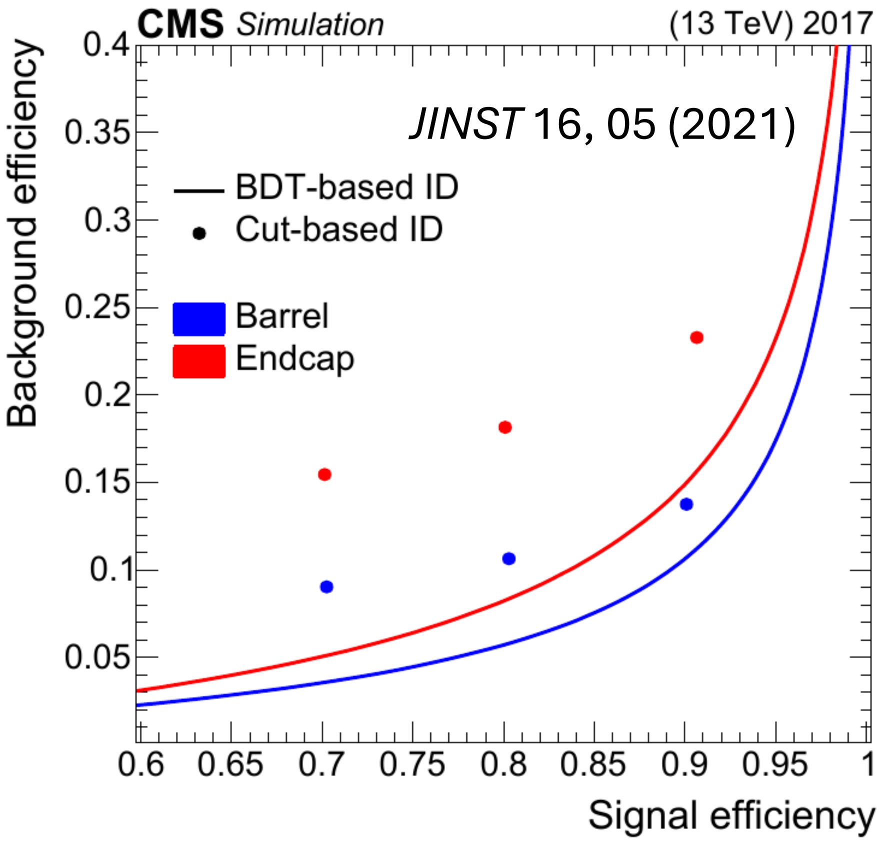
Veto

80% Signal efficiency



90% Signal efficiency

MVA based



Photon ID?

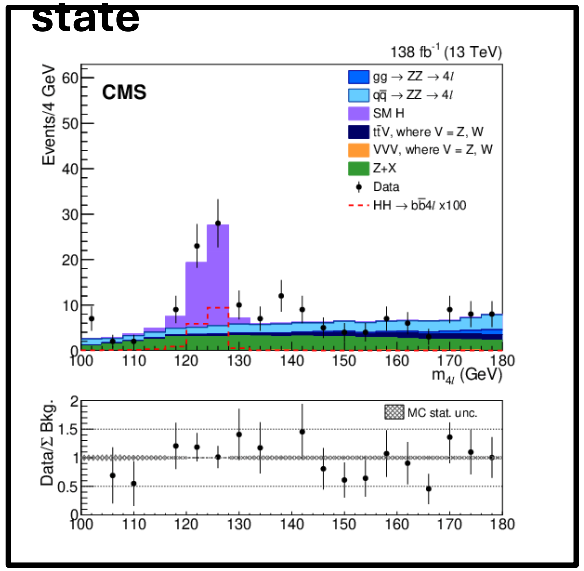
Junquan will now talk about how to identify photons using a hands-on exercise

BACKUP

Some plots on these slides have been picked up from common material of E/Gamma POG.

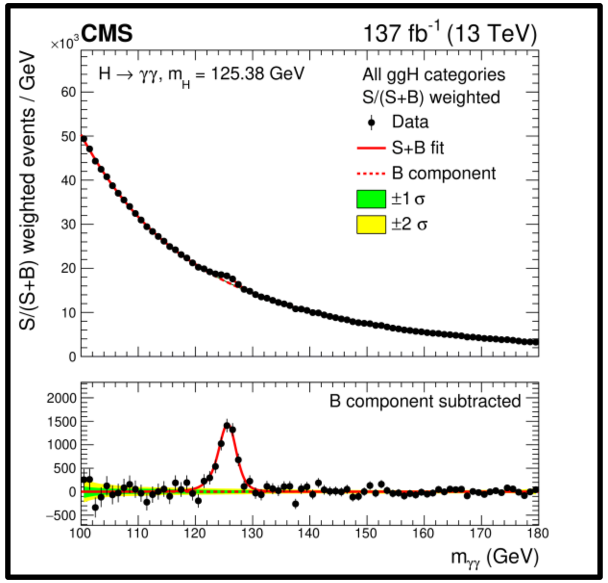
e/γ Identification is super critical for several analyses

Search for non-resonant Higgs boson pair production in the four leptons plus two b jets final state



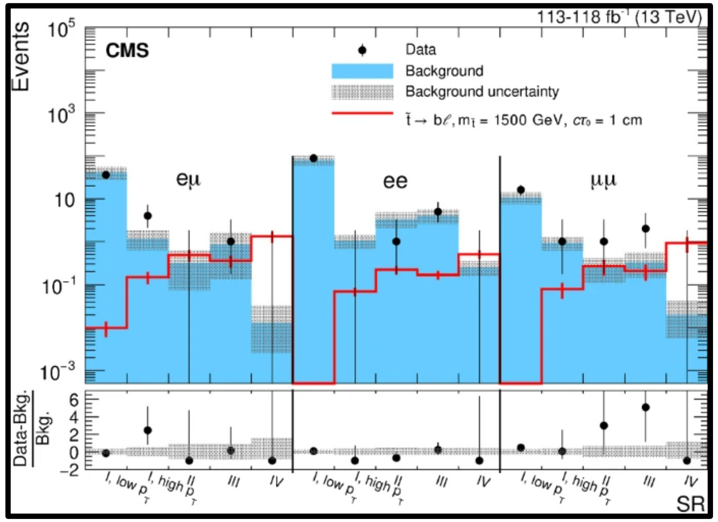
CERN-EP-2022-114

Submitted to the JHEP



JHEP 07 (2021) 027

Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel

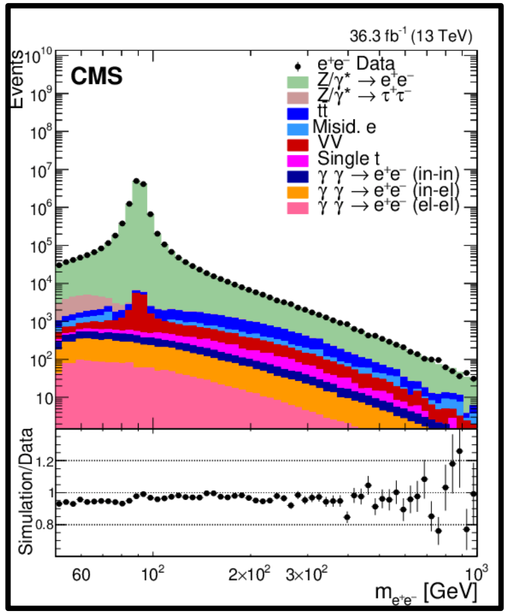


Eur.Phys.J.C 82(2022)1537

Search for long-lived particles decaying to displaced leptons

CERN-EP-2022-053

Submitted to the EPJ-C



Where to find the necessary information?

Parent Twiki:
https://twiki.cern.ch/twiki/bin/view/CMS/EgammaPOG

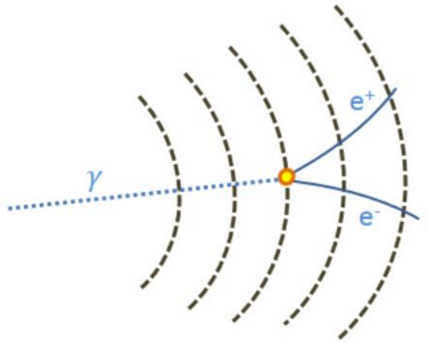
Run2:
https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2

Run3:
https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun3

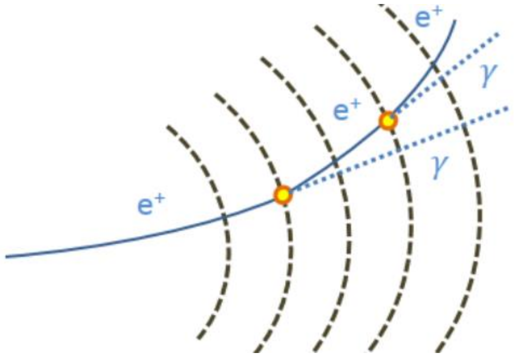
High pT Photons:
https://twiki.cern.ch/twiki/bin/view/CMS/EGMPhotonIDHighPtPhotons

How e/γ interact with matter

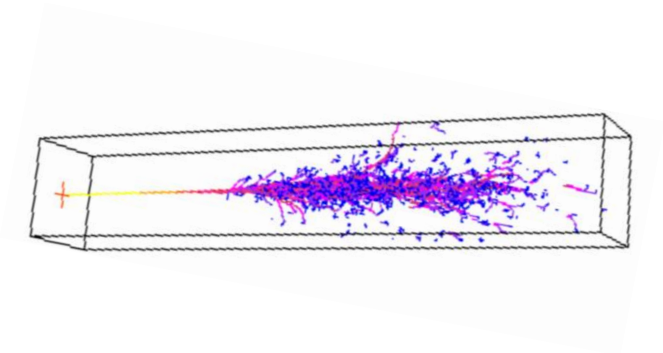
Pair production

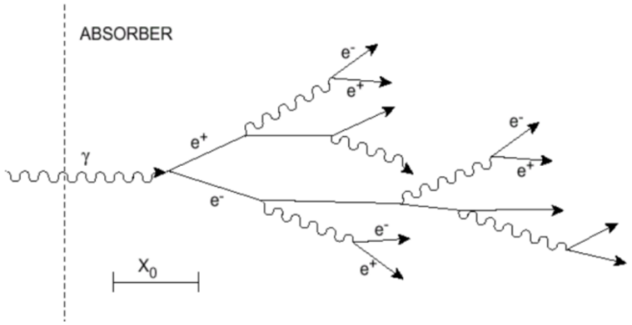


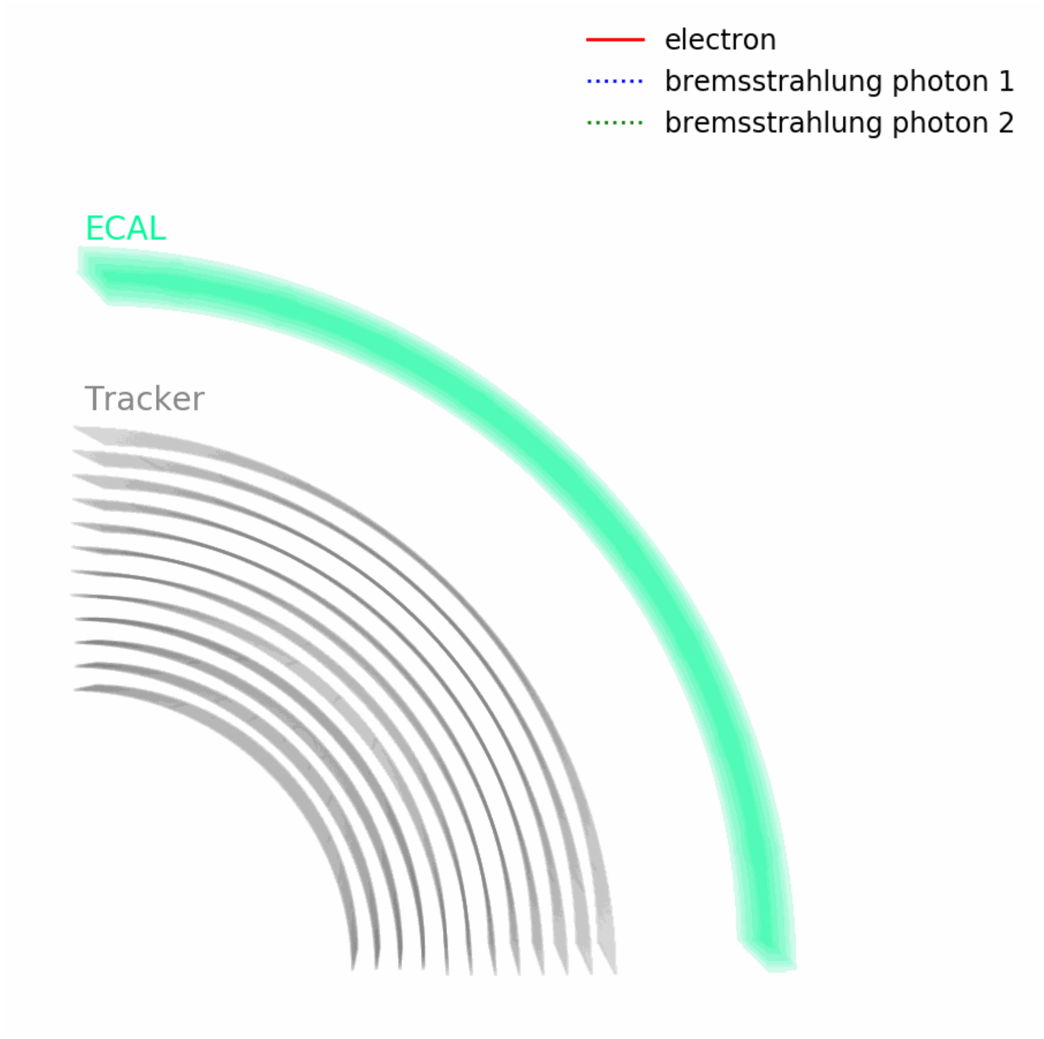
Bremsstrahlung

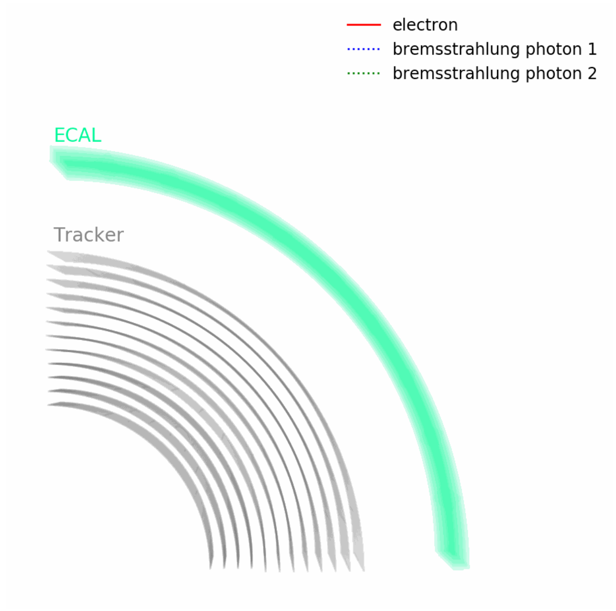


These two processes form the basis of the “electromagnetic shower”



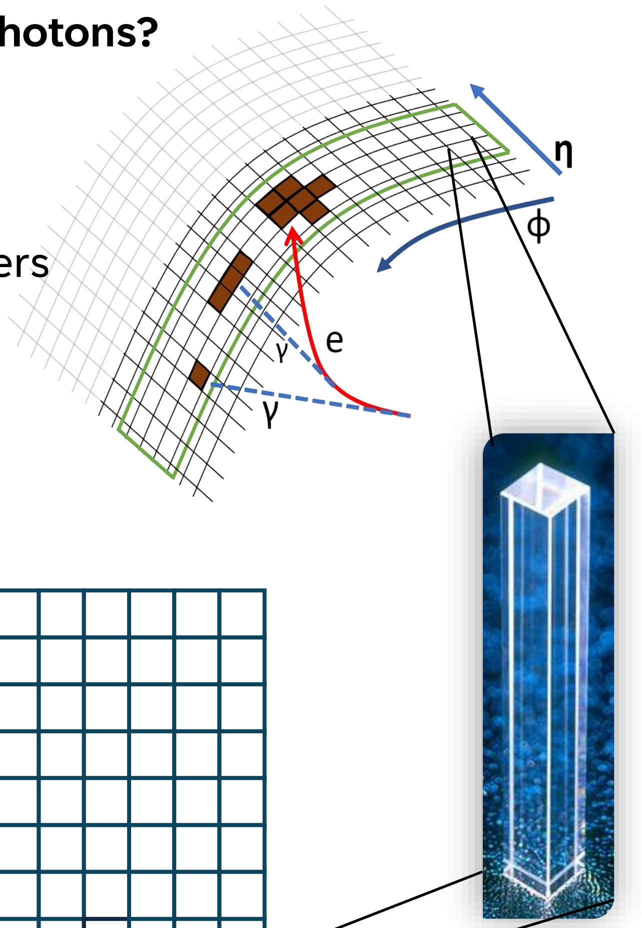


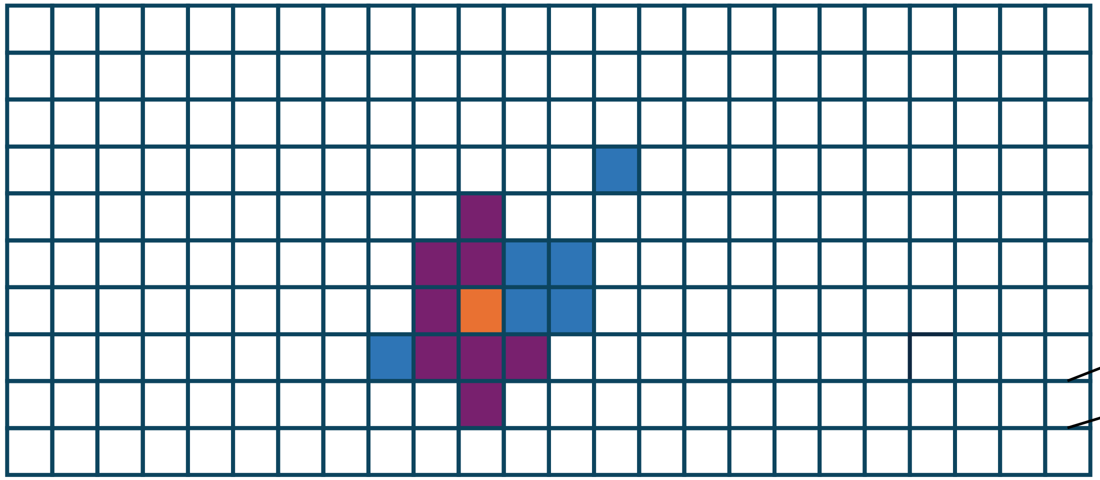

electron
bremsstrahlung photon 1
bremsstrahlung photon 2
ECAL
Tracker

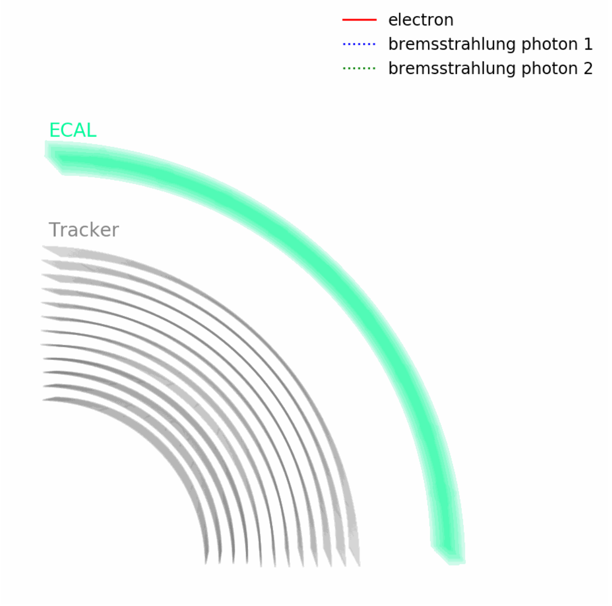


What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

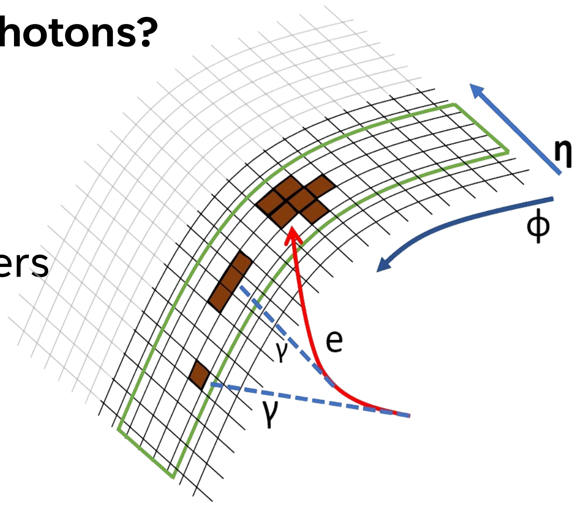


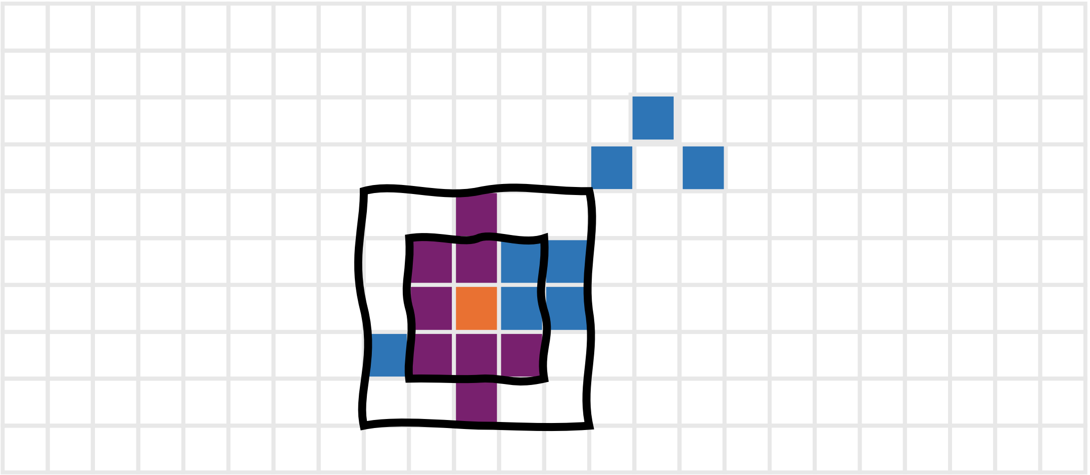




What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

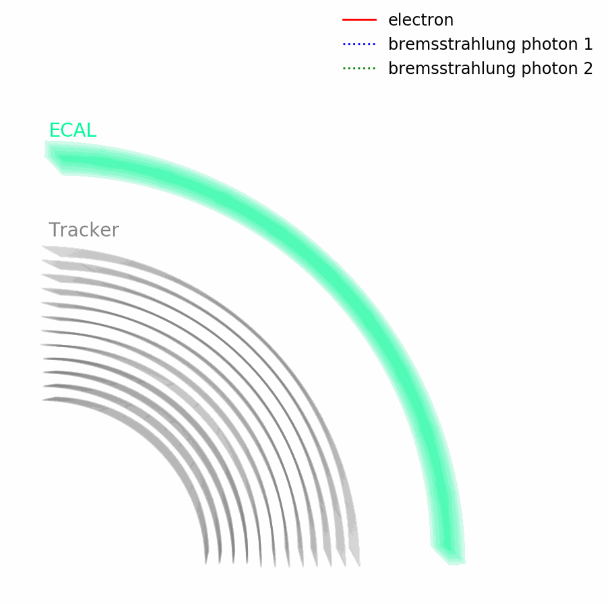


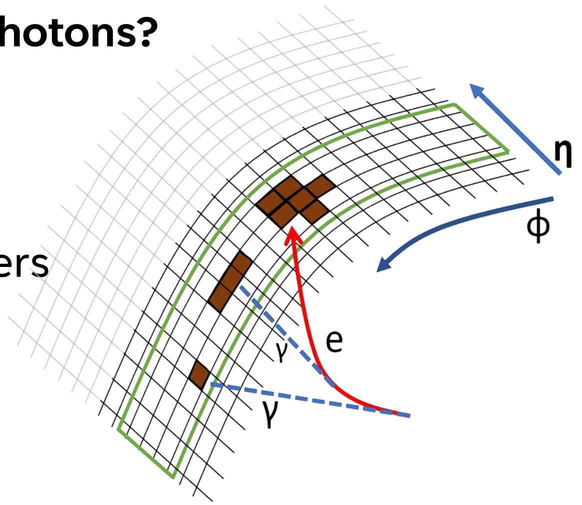


E_{sc} : Energy of supercluster
E_{2x2} : Energy contained in 2X2 crystals
E_{3x3} : Energy contained in 3X3 crystals
E_{5x5} : Energy contained in 5X5 crystals
E_{1x5} : Energy contained in 1X5 crystals
$d\eta_{sc}$: η width of supercluster
$d\Phi_{sc}$: Φ width of supercluster

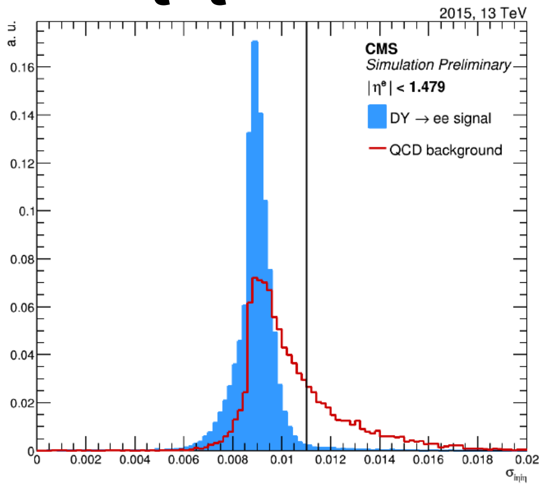
What can we use to identify electrons and photons?

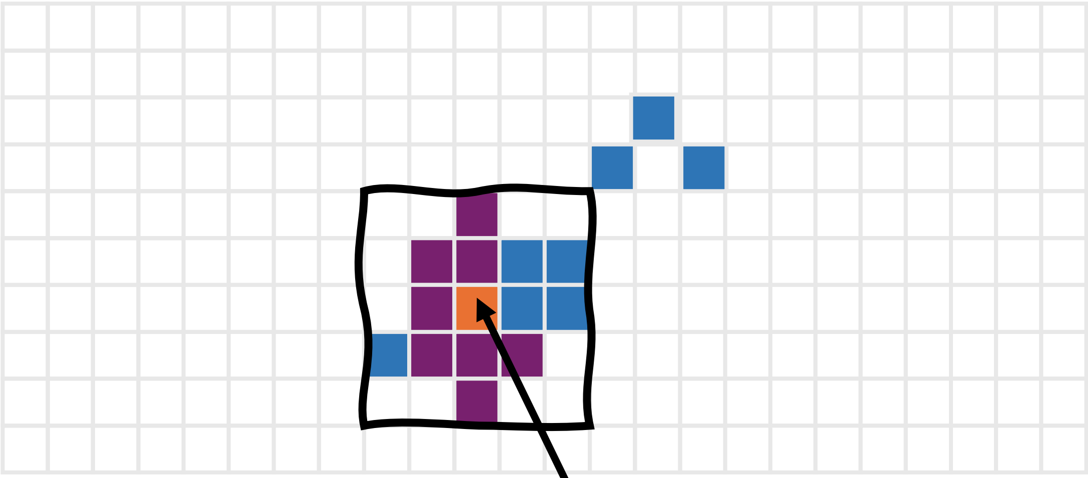
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects





$\sigma_{i\eta i\eta}$

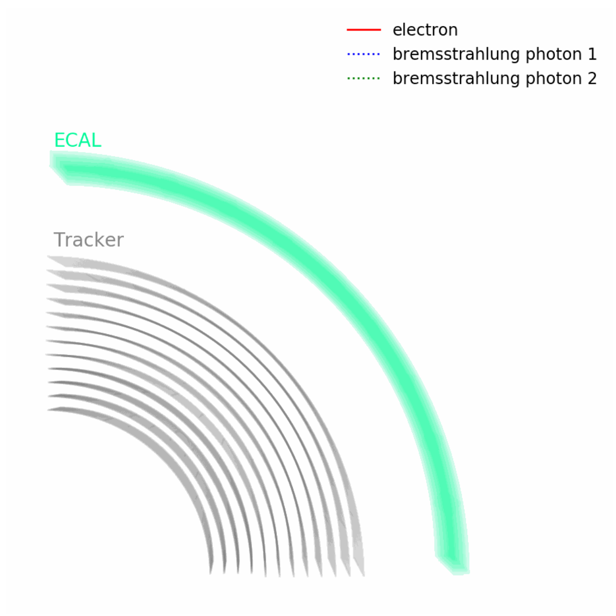


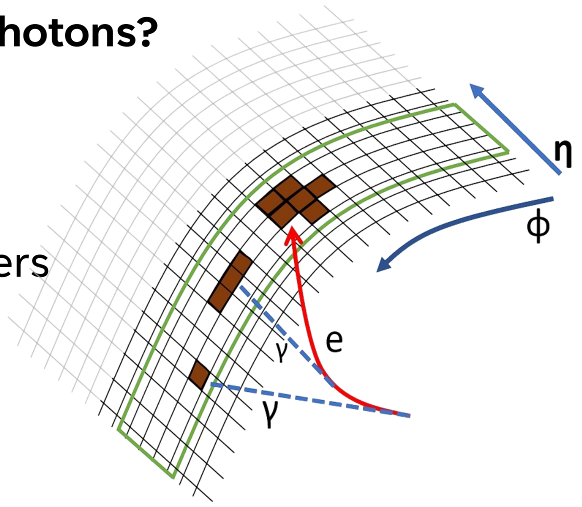


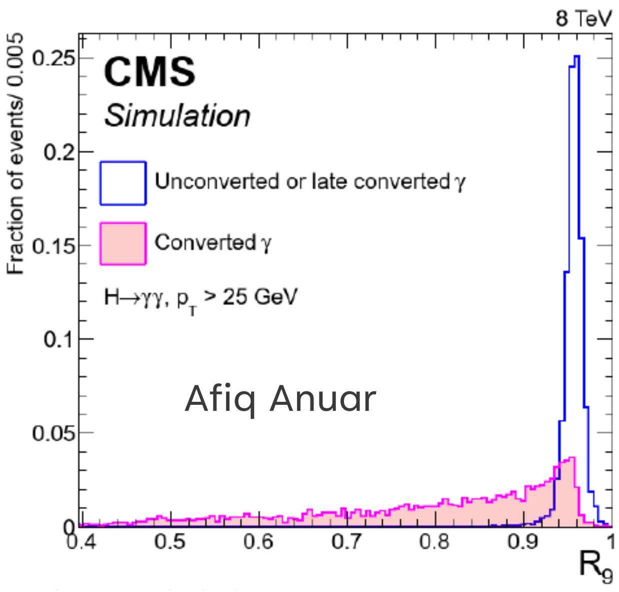
Most energetic crystal of 5x5 array

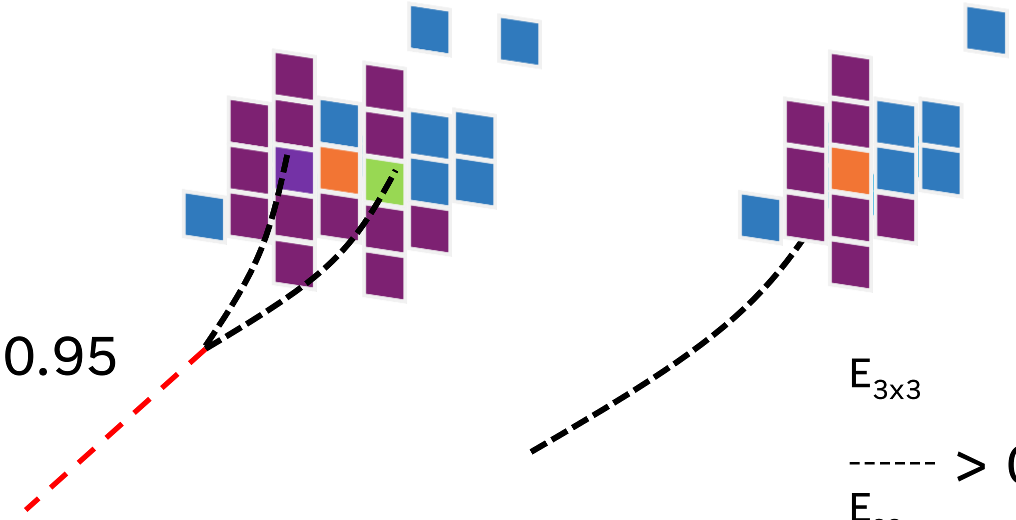
What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects





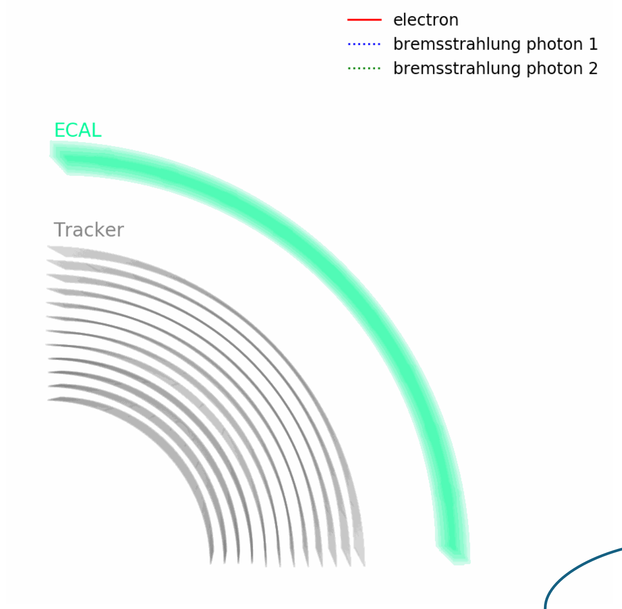




$$\frac{E_{3x3}}{E_{sc}} < 0.95$$

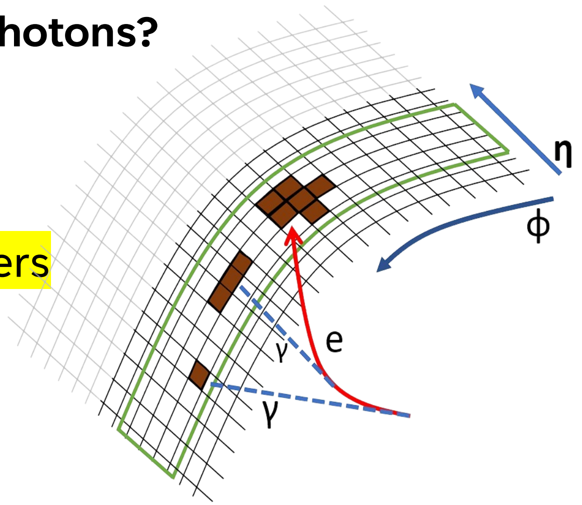
$$\frac{E_{3x3}}{E_{sc}} > 0.95$$

5x5 matrix contains ~96% (~97%) of unconverted photon energy in EB (EE)



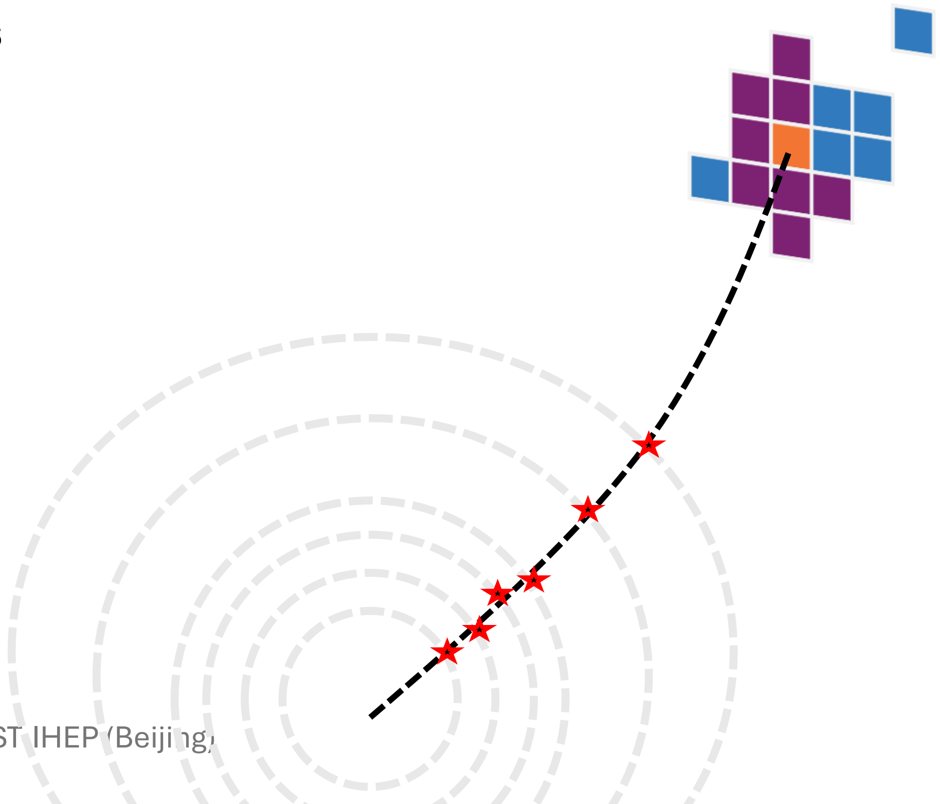
What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

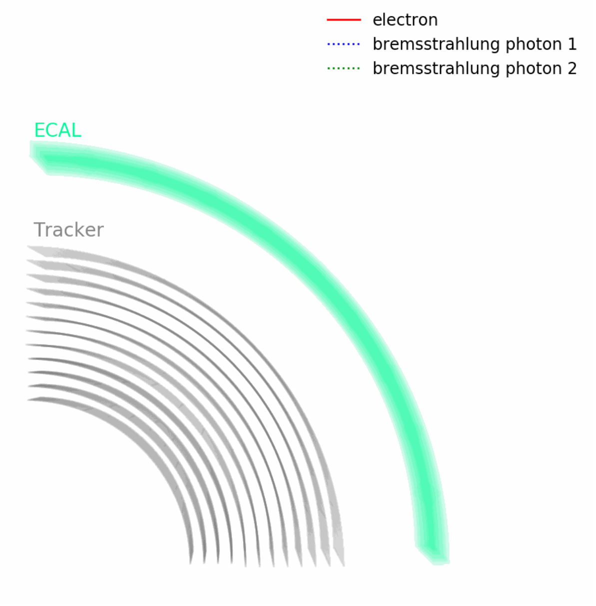


We heard about this in Wolfgang's talk in the morning

N^{gsf}_{hits} : Hits in the "gsf" track
N^{kf}_{hits} : Hits in the "kf" track
E/p : Energy of supercluster/ momentum
χ^2: Track quality

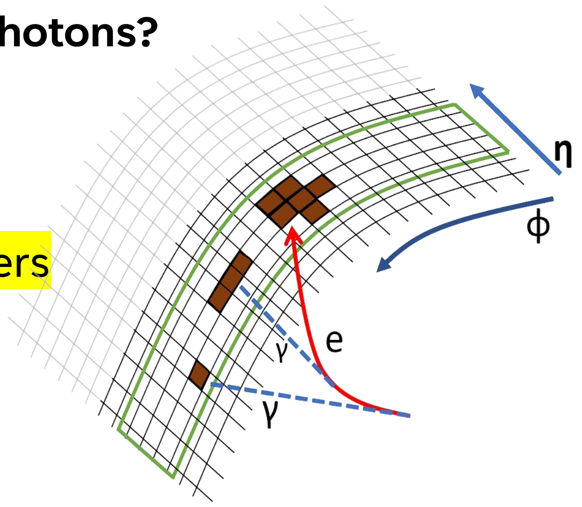


Not just for electrons,
Even for photons, good ones should not have a "track at all"
Or would have converted?



What can we use to identify electrons and photons?

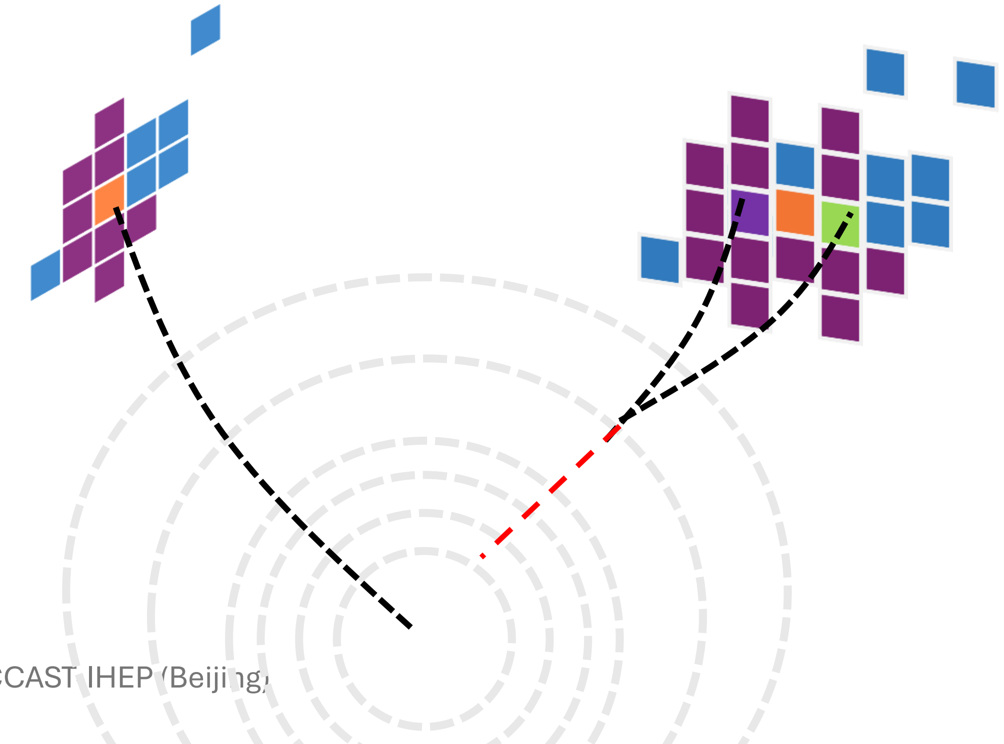
- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

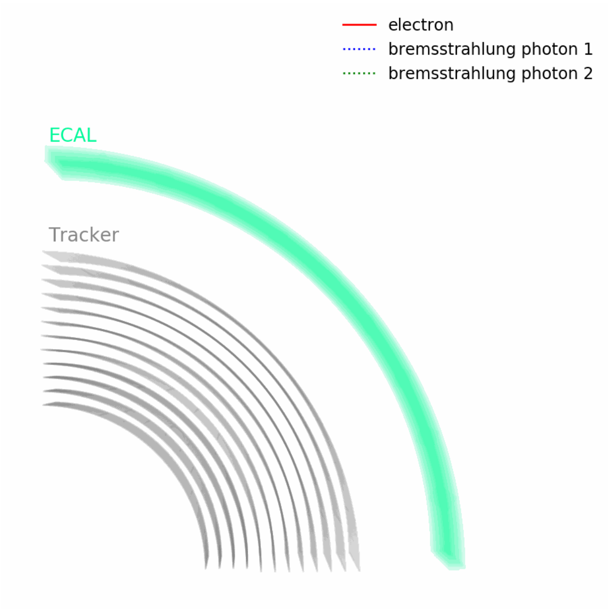


Conversion ID variables

How to differentiate between electrons in photon conversions from prompt electrons?

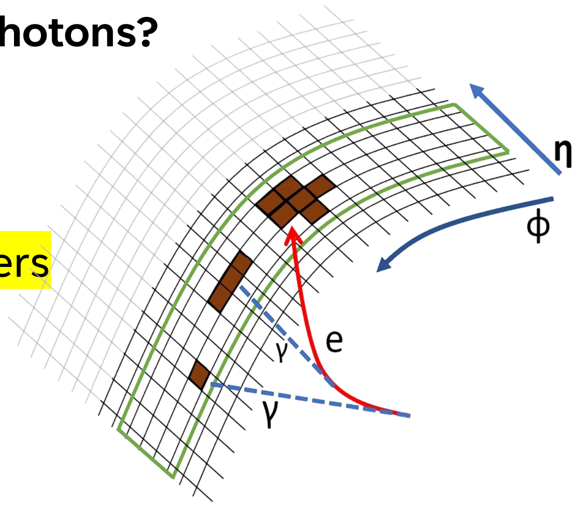
Conversion Safe Electron Veto
Pixel Veto





What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

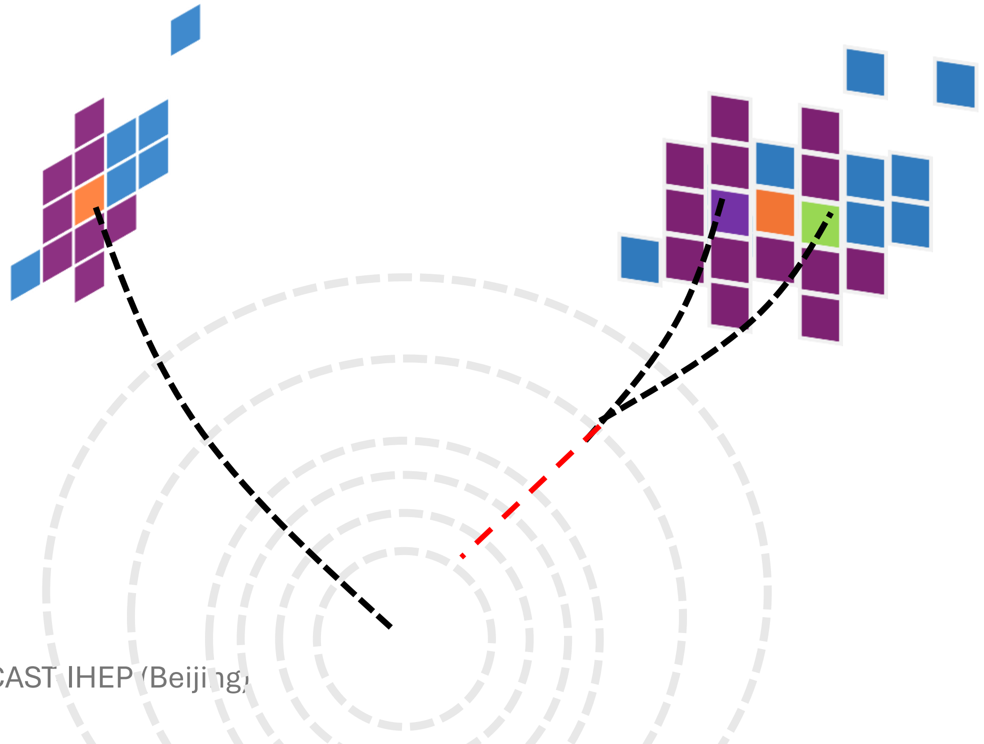


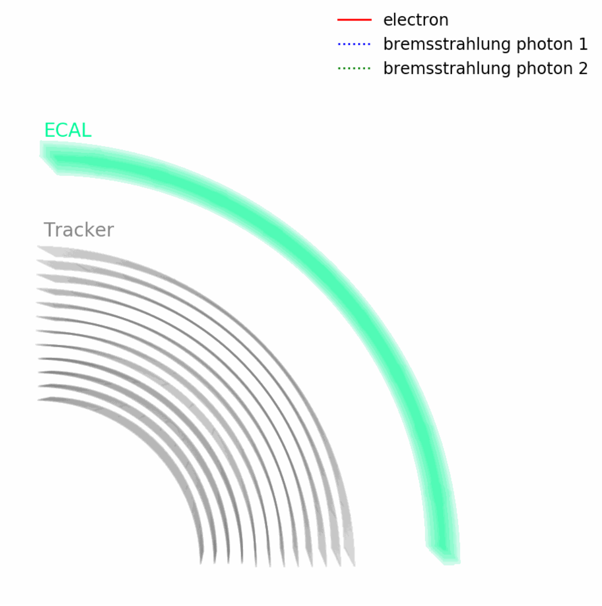
Conversion Safe Electron Veto

If a secondary vertex is found, this is not an electron!

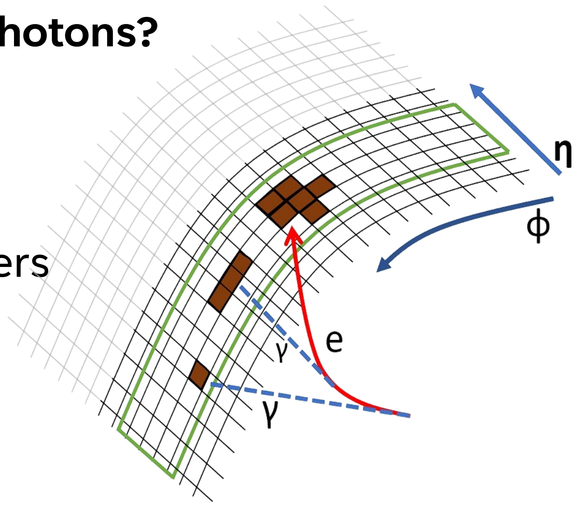
Pixel Veto

Track in pixel detector, this is not a photon



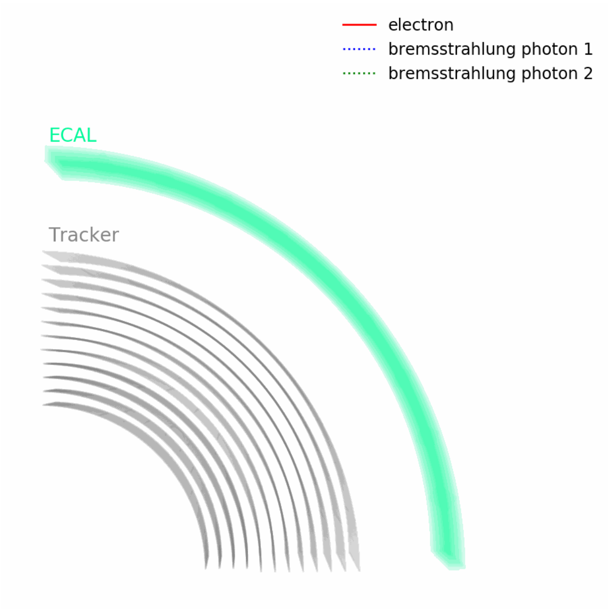


What can we use to identify electrons and photons?

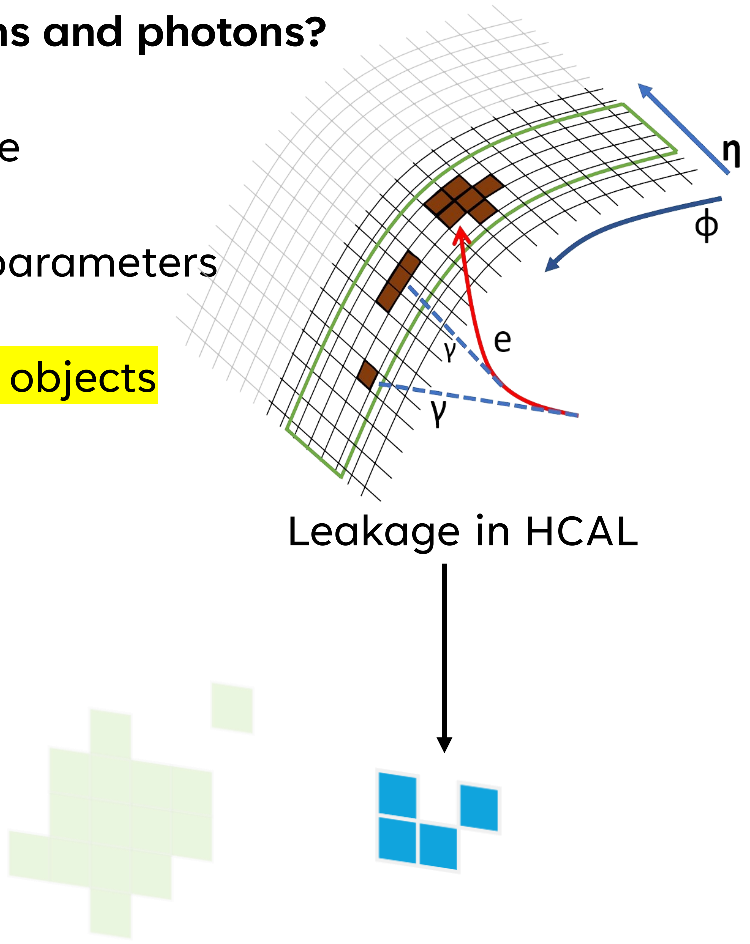


- ➢Description of the EM shower shape
- ➢Tracking and clustering matching parameters
- ➢Quantification of isolation of these objects

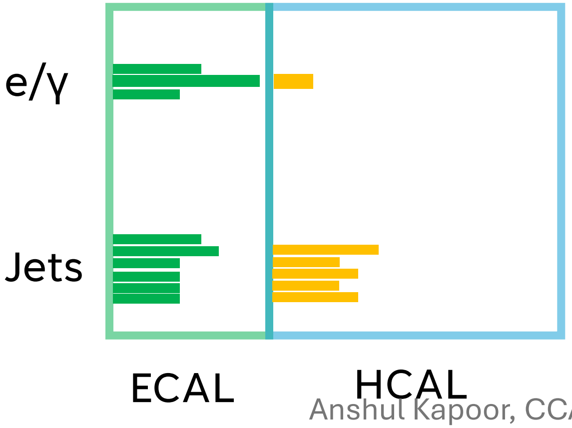
What can we use to identify electrons and photons?

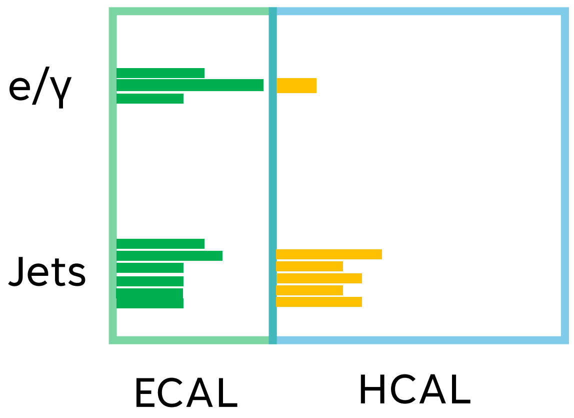


- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects



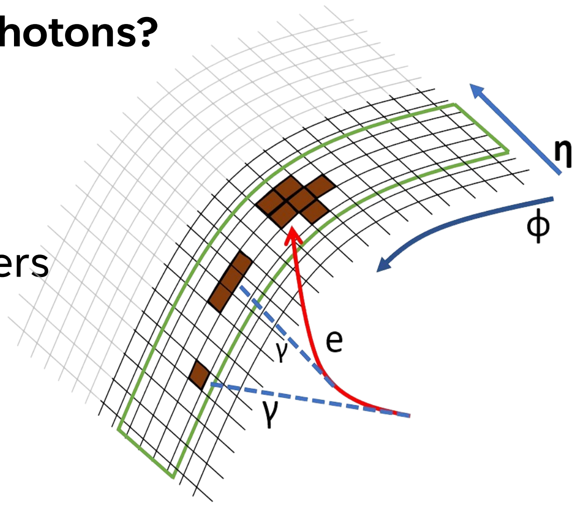
H/E : Energy leaked into HCAL / Energy in ECAL



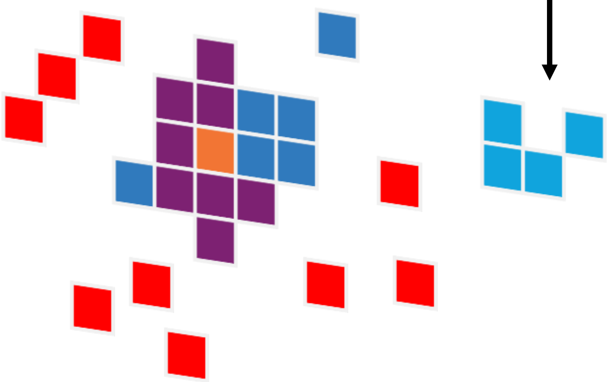


What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

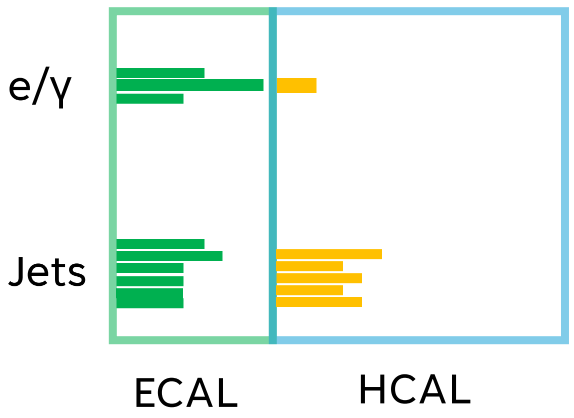


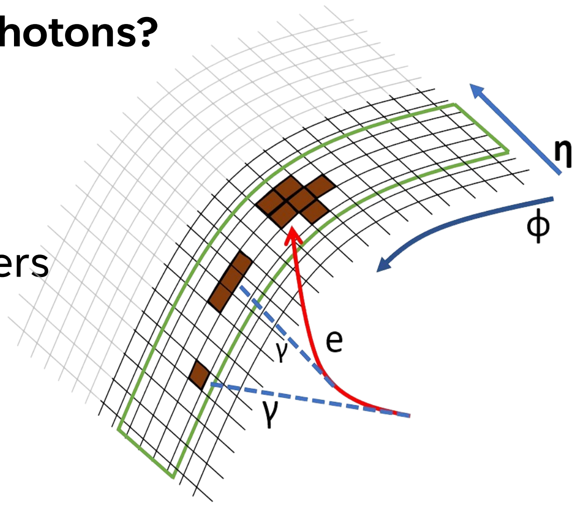
H/E : Energy leaked into HCAL / Energy in ECAL
Isolation: Other stuff around it?



What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects

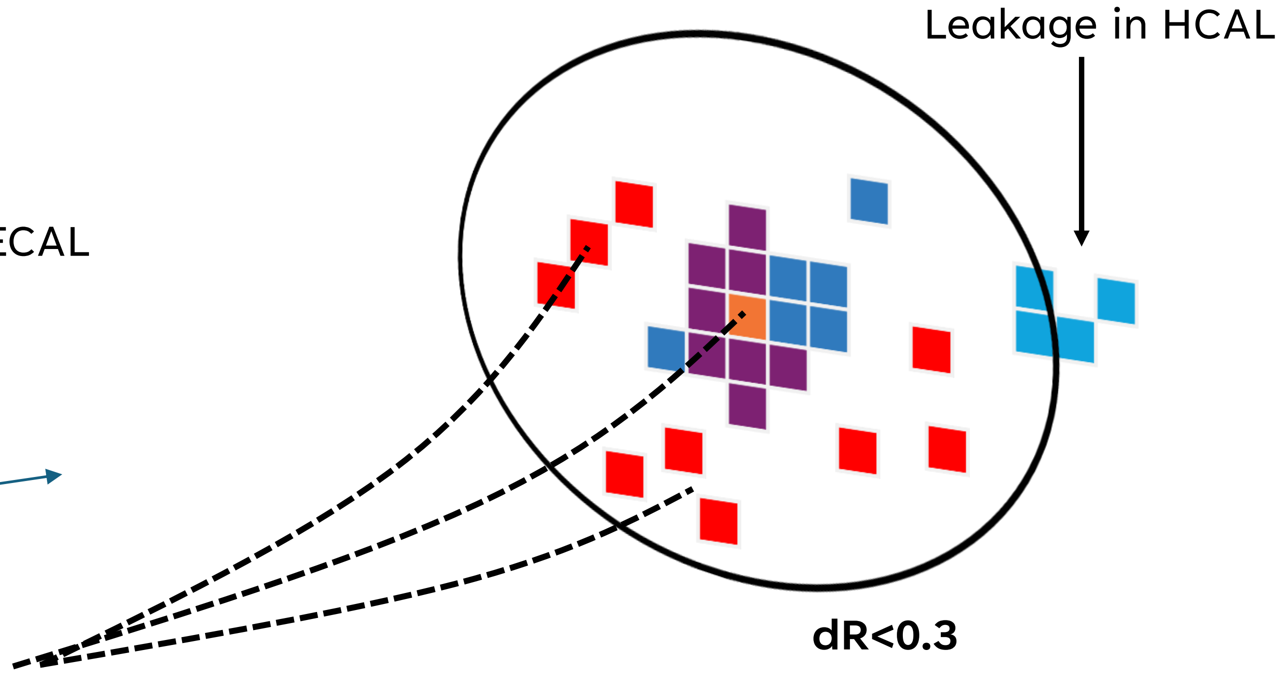


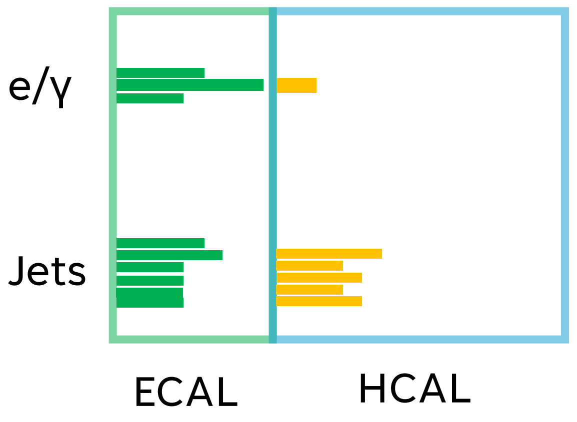


H/E : Energy leaked into HCAL / Energy in ECAL
Isolation: Other stuff around it?

Information provided by PF Is used here:

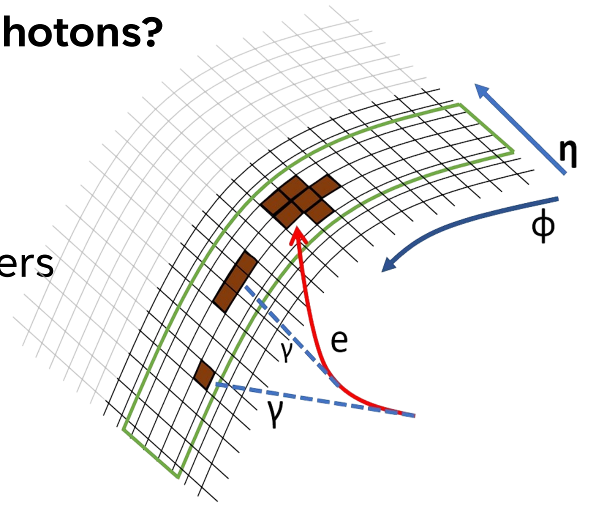
PF Cluster+Tracker isolations
HCAL clusters
ECAL clusters
Tracks





What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects



Correlated

H/E : Energy leaked into HCAL / Energy in ECAL

Isolation: Other stuff around it?

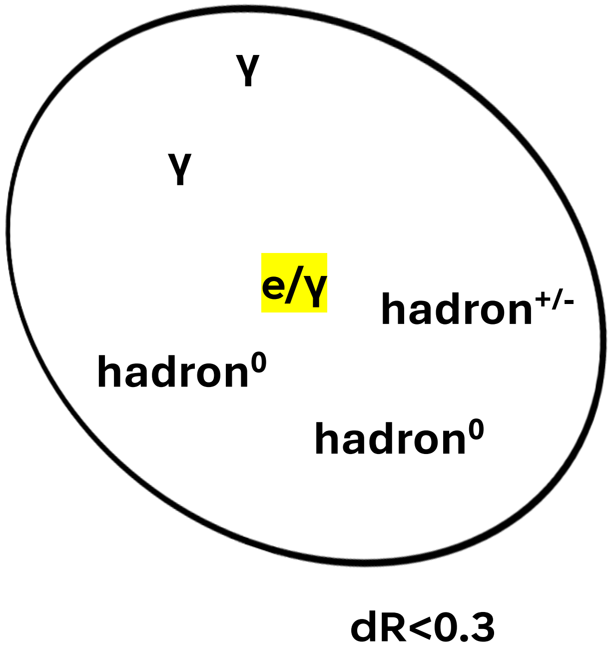
Information provided by PF Is used here:

PF Particle Isolations

Charged hadrons

Neutral hadrons

Photons



Central e/γ Identification

HLT
~99% signal efficiency

Offline Particle Flow
~98-99% signal efficiency

Offline Analysis
~70/80/90% signal efficiency

The basics that go into each of them are the same!

Yet, "signal of interest" is not EXACTLY the same

Gsf electrons

Central e/γ Identification

HLT
~99% signal efficiency

Offline Particle Flow
~98-99% signal efficiency

Offline Analysis
~70/80/90% signal efficiency

Part 2

13:30 Introduction to EGamma HLT and Tag-and-Probe setup

Riccardo Salvatico

The basics that go into each of them are the same!

Yet, "signal of interest" is not EXACTLY the same

Central e/γ Identification

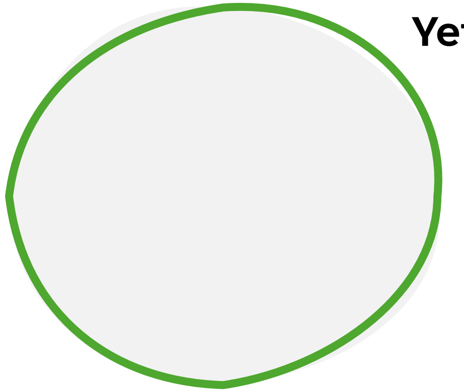
HLT
~99% signal efficiency

Offline Particle Flow
~98-99% signal efficiency

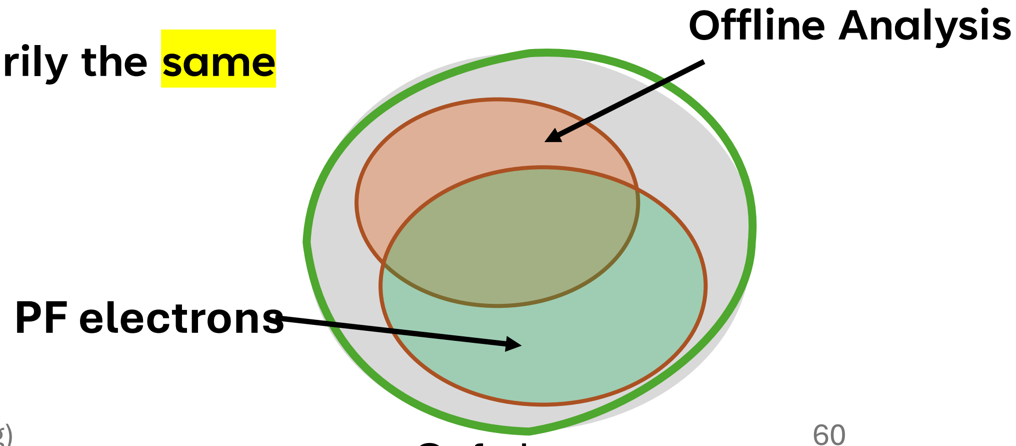
Offline Analysis
~70/80/90% signal efficiency

The basics that go into each of them are the same!

Yet, “signal of interest” is not necessarily the same



HLT Electrons

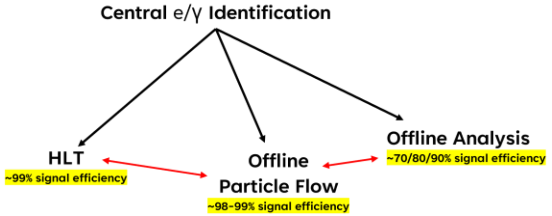


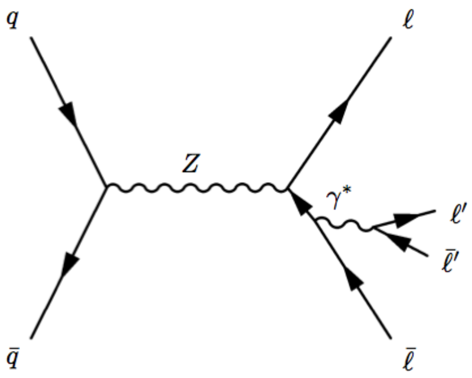
Gsf electrons

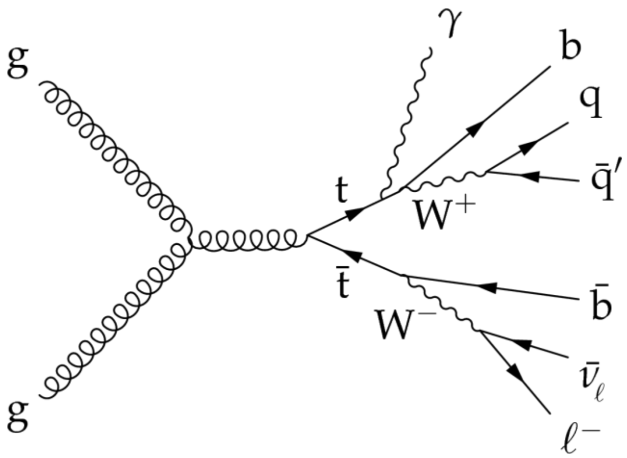
Good object/ Signal

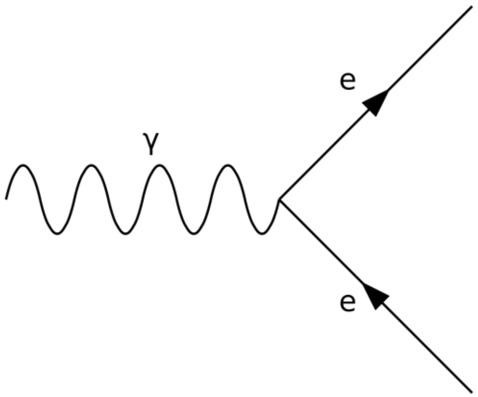
Bad object/ Background

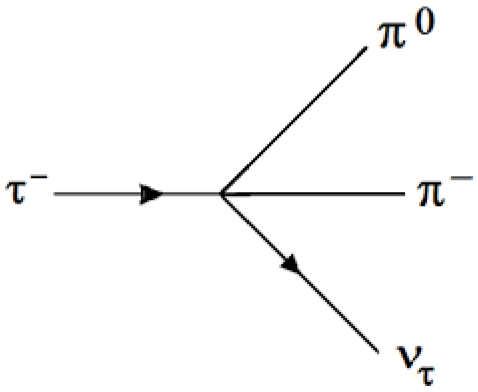
What is good and what is bad, can often be subjective!

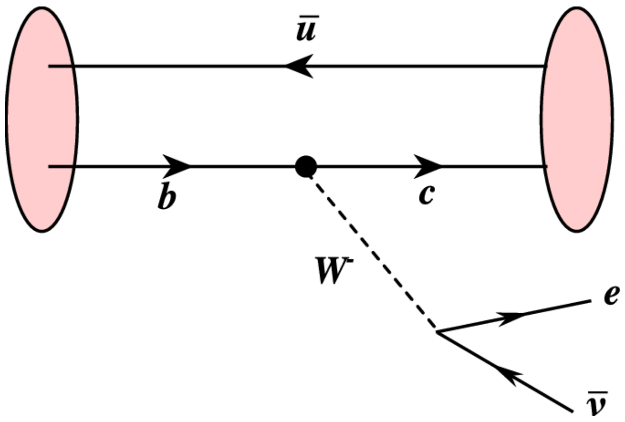








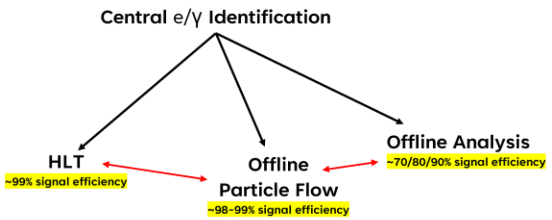


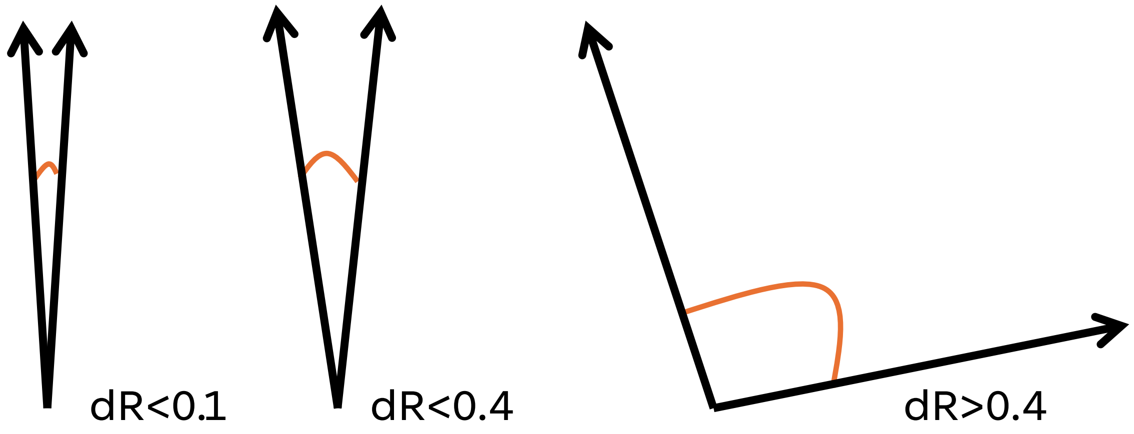


Good object/ Signal

Bad object/ Background

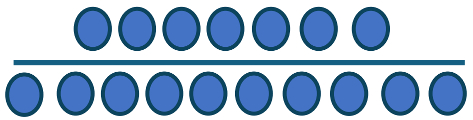
What is good and what is bad, can often be subjective!



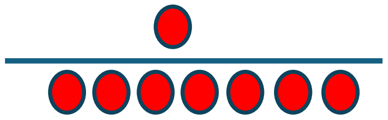


One identification criteria won’t work for everything!

Signal efficiency



Background efficiency



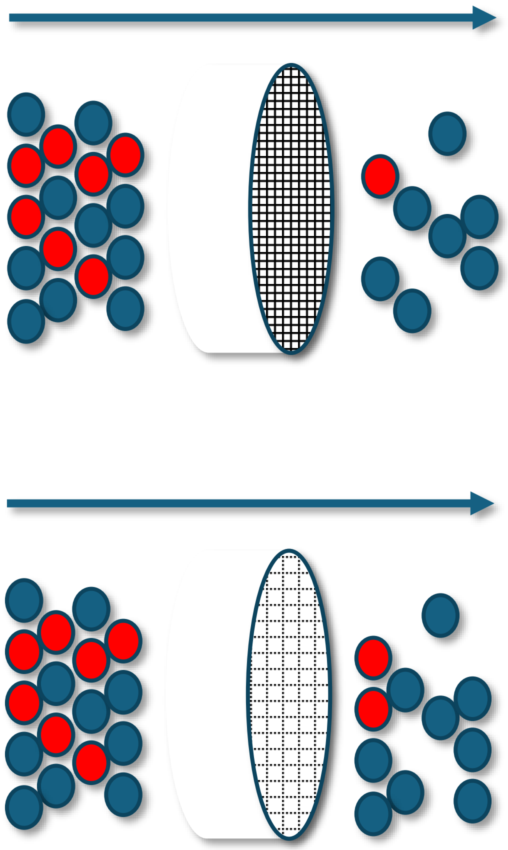
All identification criteria have two important metrics:
False positive / Background efficiency
True positive / Signal efficiency

High signal efficiency and low background efficiency are the most important requirements

Along with a few other necessary conditions!

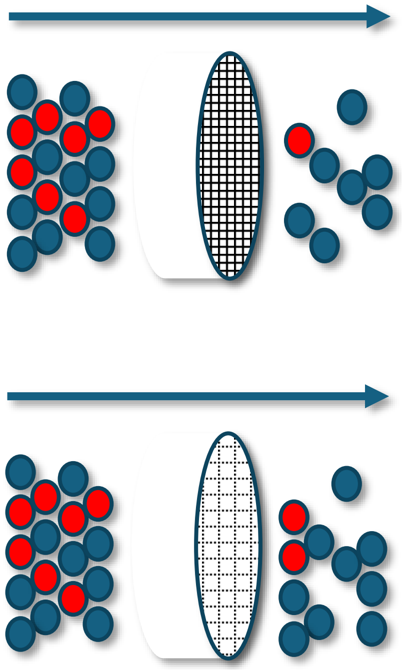
The basics: Working Points

The same kind of criteria can have a looser/tighter version



The goal is to design one or more criteria that most “good” objects pass and most “bad” objects fail

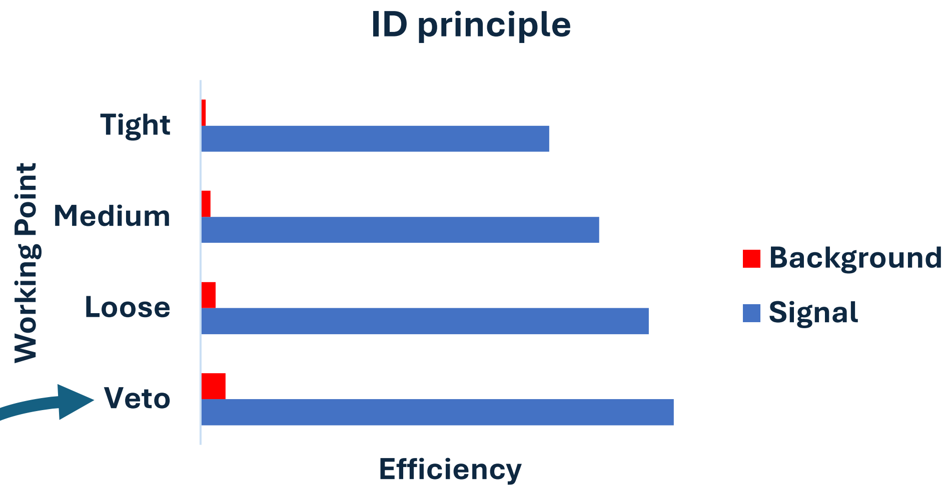
Purpose: Someone is happy with 70% signal but might want absolutely minimal background, whereas for another it is probably ok to have some background, but they want 90% signal



What is good and what is bad, can often be subjective!

Yet, for most purposes, good is always good, and bad is "mostly" bad.

With this principle in mind, E/Gamma POG designs the most "general" offline identification criteria



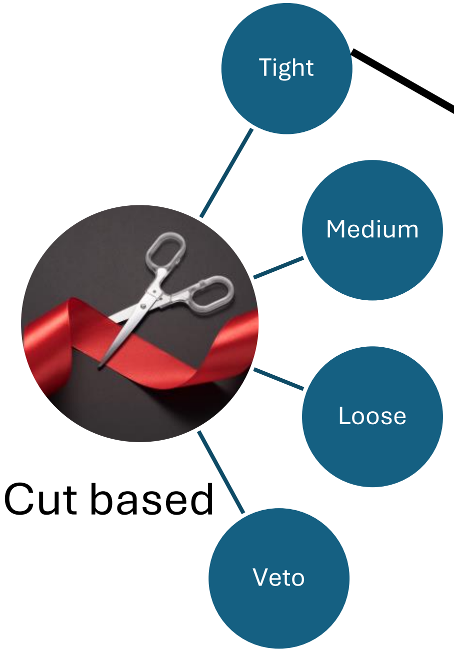
Selects almost every "good object", but significant background also passes

Offline Analysis
Central E/Gamma Identification

Electron case as an example:

What should an electron look like? Here are some key indicators to consider:

- Isolated cluster: A prompt electron will result in an isolated cluster
- Electromagnetic cluster: characterized by low H/E
- EM shower typically appears as a concentrated cluster
- E_{SC} /pT: a track and cluster produced by the same particle, indicating a high likelihood of being an electron
- Matching cluster-track: comparing the angle of the cluster to the track can help distinguish against Bremsstrahlung radiation.



Offline Analysis
Central E/Gamma Identification

Electron case as an example

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
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Electron case as an example

Efficient against:
All hadronic background

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
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Electron case as an example

Efficient against:
Jets

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
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
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
Pass conversion veto	Yes	Yes

Offline Analysis
Central E/Gamma Identification

Electron case as an example

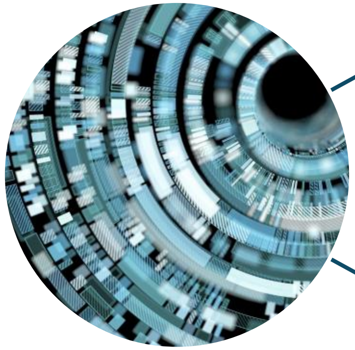
Variable	Barrel (tight WP)	Endcaps (tight WP)
$\sigma_{i\eta i\eta}$	<0.010	<0.035
$\lvert\Delta\eta_{\text{in}}^{\text{seed}}\rvert$	<0.0025	<0.005
$\lvert\Delta\phi_{\text{in}}\rvert$	$<0.022\,\text{rad}$	$<0.024\,\text{rad}$
H/E	$<0.026 + 1.15\,\text{GeV}/E_{\text{SC}}$ $+0.032\rho/E_{\text{SC}}$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

Efficient against:
Pions, photons

Offline Analysis
Central E/Gamma Identification

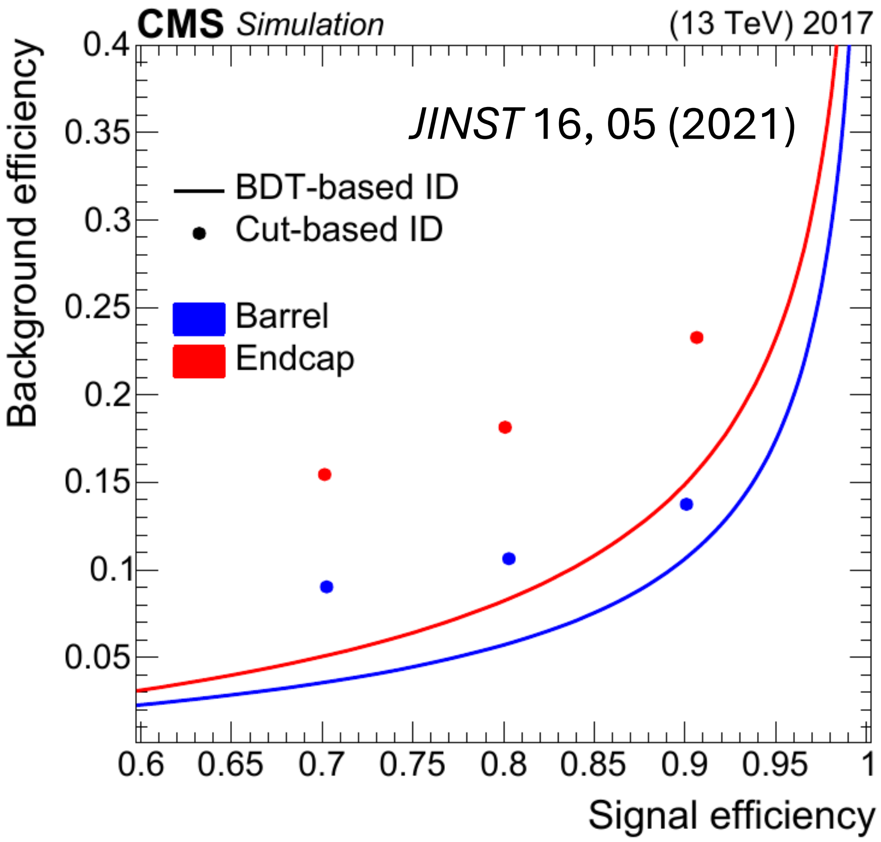
Cut based

- Tight
- Medium
- Loose
- Veto



MVA based

- 80% Signal efficiency
- 90% Signal efficiency



Offline Analysis

~70/80/90% signal efficiency

Offline Analysis
Central E/Gamma Identification

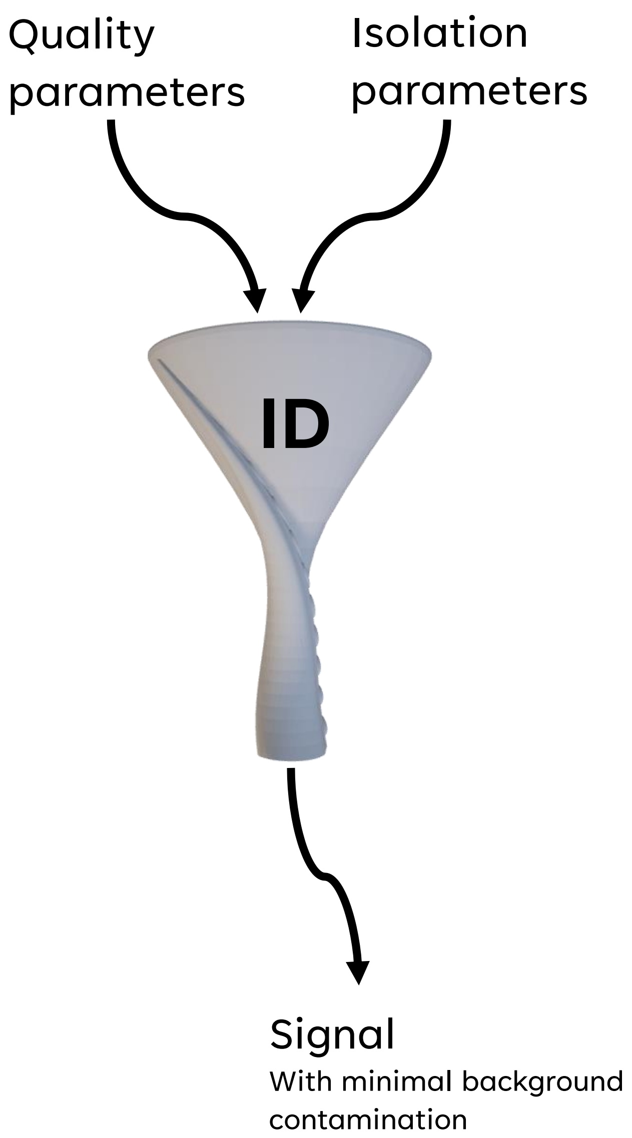
What is a good electron/photon?

For an experimentalist, the answer depends on whether one is

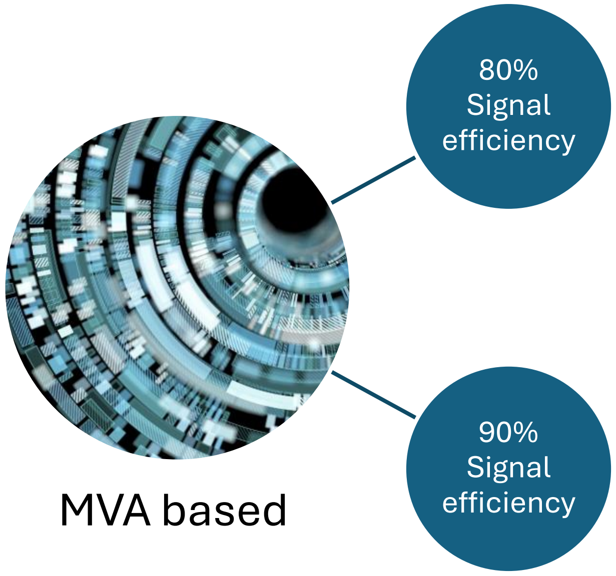
- Searching for a new physics process
- Isolating electrons from jets
- Correcting jet energy based on its leptonic content
- Identifying electrons in hadronically decaying tau leptons, and so on…

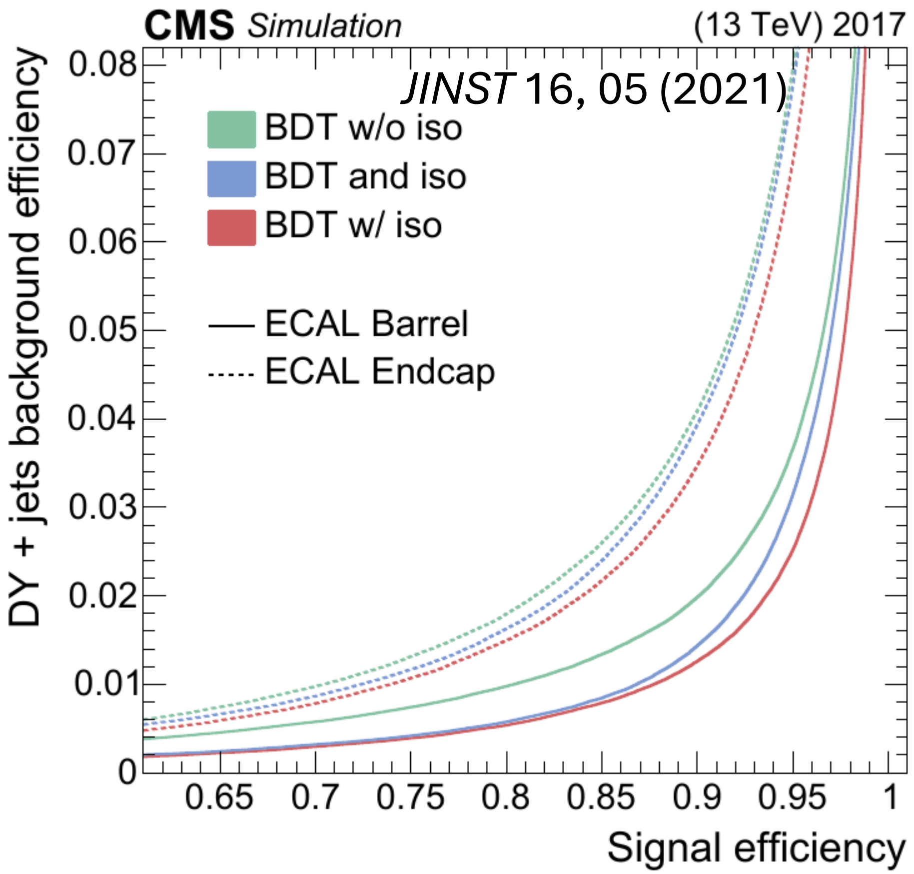
It is thus important to "isolate" the "isolation variables"

-> Stuff that depends on other objects around the main object of interest



Offline Analysis
Central E/Gamma Identification





Offline Analysis
Central E/Gamma Identification

Cut based	Veto	Loose	Medium	Tight
Electron	✓	✓	✓	✓
Photon		✓	✓	✓

MVA ID	90wp	80wp	
Electron (Isolated)	✓	✓	
Electron (Non-Isolated)	✓	✓	
Photon	✓	✓	
HEEP ID	✓	✓	

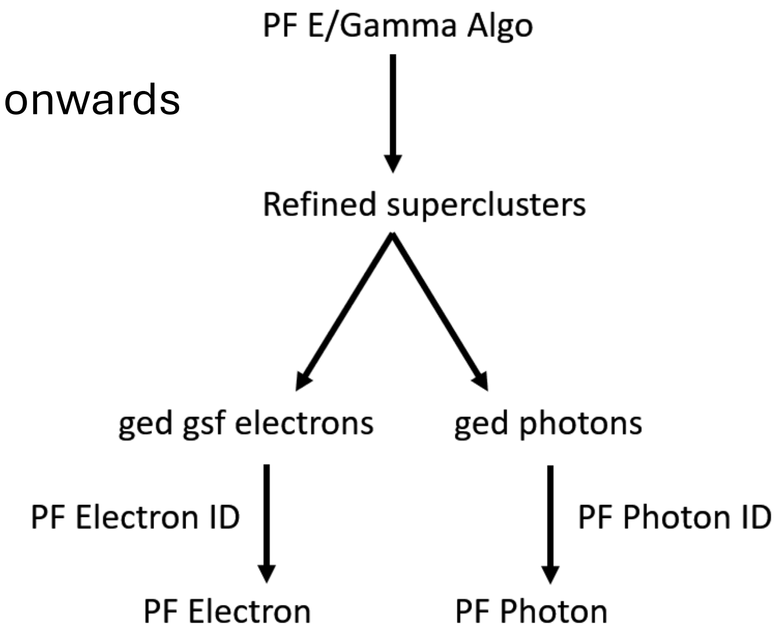
Offline Particle Flow

~98-99% signal efficiency

Particle Flow
Central E/Gamma Identification

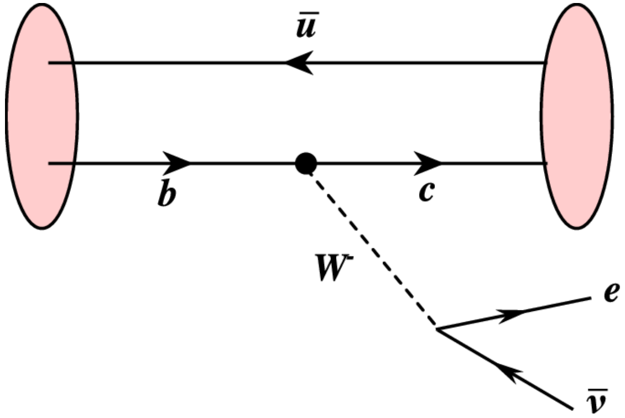
Correctly identifying **PF electrons** and **photons** is critical to PF reconstruction
- Indirect impact on PF quantities like isolation and on tau reconstruction

$\pi^0 \rightarrow$ **Photons** -> e+e- e+e-
π± can produce cascades of low pT electrons and photons
We already know 5% of 1prong+π0 tau leptons were lost Run1 onwards



~98-99% signal efficiency

Particle Flow
Central E/Gamma Identification



Good object/ Signal
Bad object/ Background

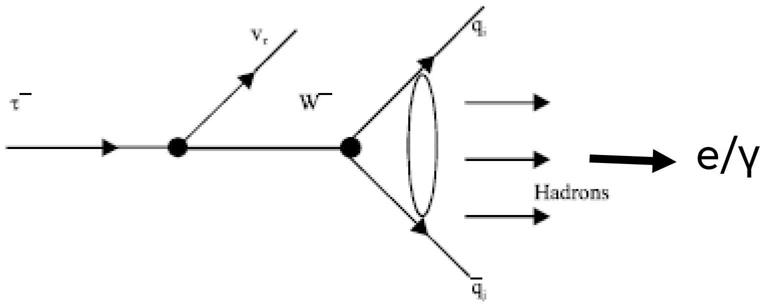
?

Most offline analyses will not consider this electron as a signal!

At the level of particle flow, this needs to identify as a proper electron.

Offline
Particle Flow
~98-99% signal efficiency

Particle Flow
Central E/Gamma Identification



Good object/ Signal
Bad object/ Background

?

If you reconstruct these as PF electrons or PF photons, the hadronic tau reconstruction may not have these “candidates” available for PF Tau reconstruction

How does one reject such leptons but still identify the one on the previous slide?

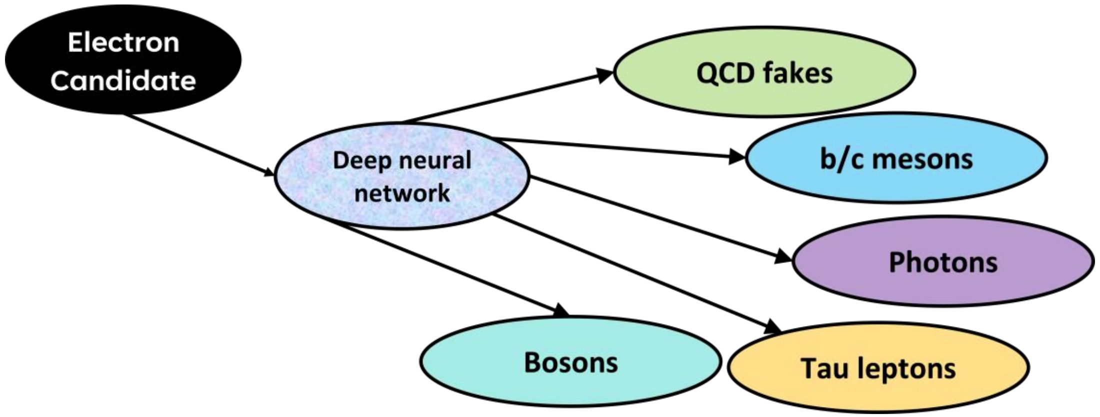
Offline

Particle Flow

~98-99% signal efficiency

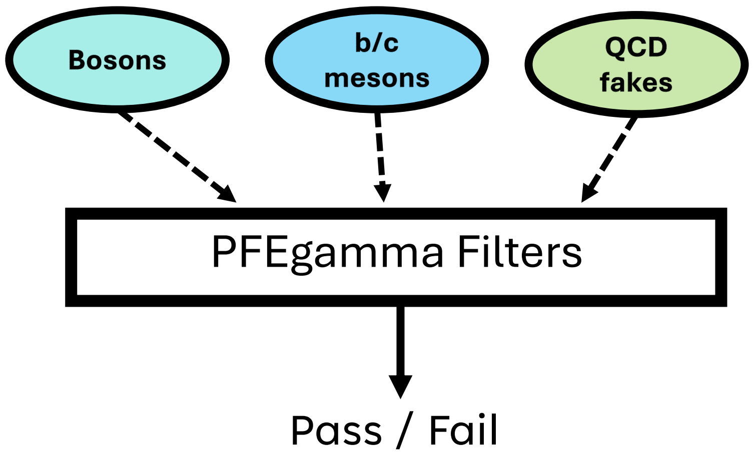
Particle Flow
Central E/Gamma Identification

New ID in Run3



A DNN-based multi-class identification is done for electrons.

Fixed "working point"



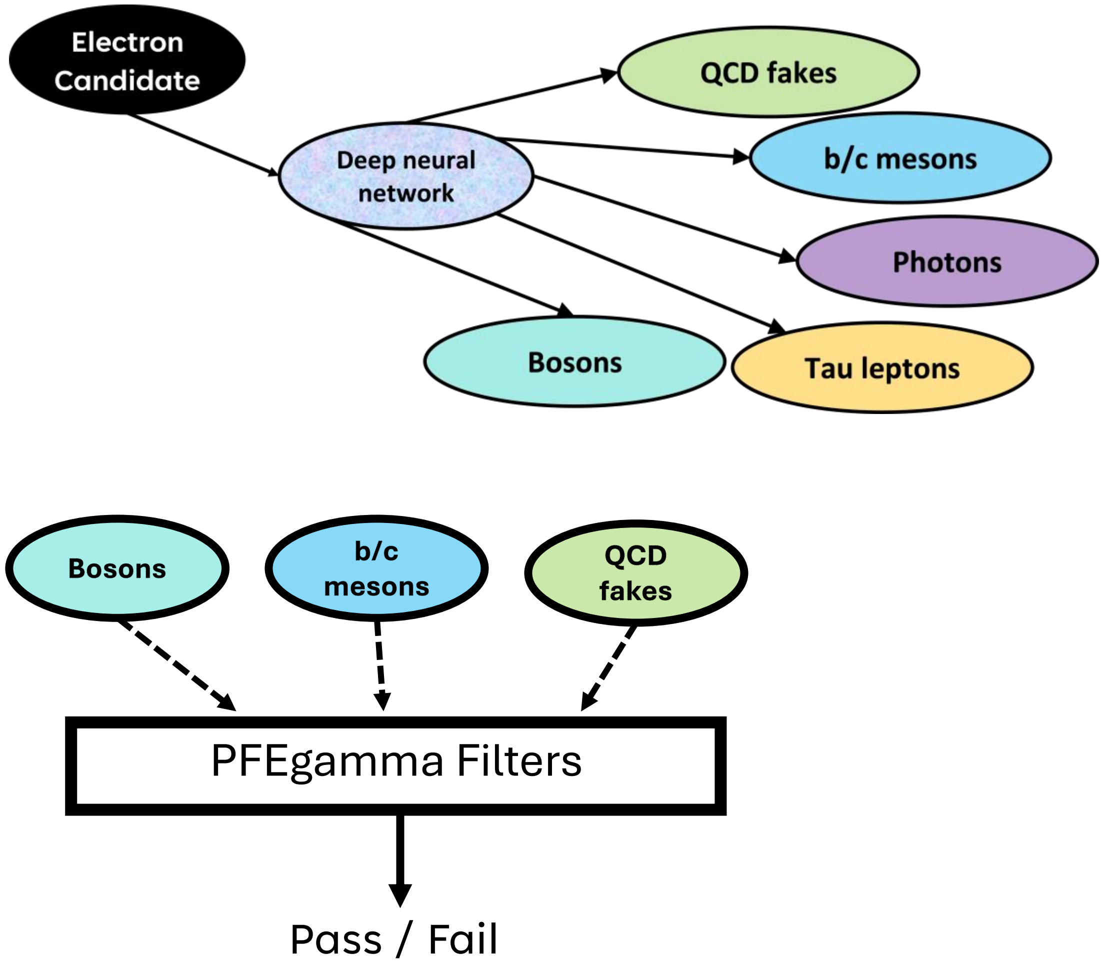
A neural network is used to assign a probability that an electron candidate came from a certain source

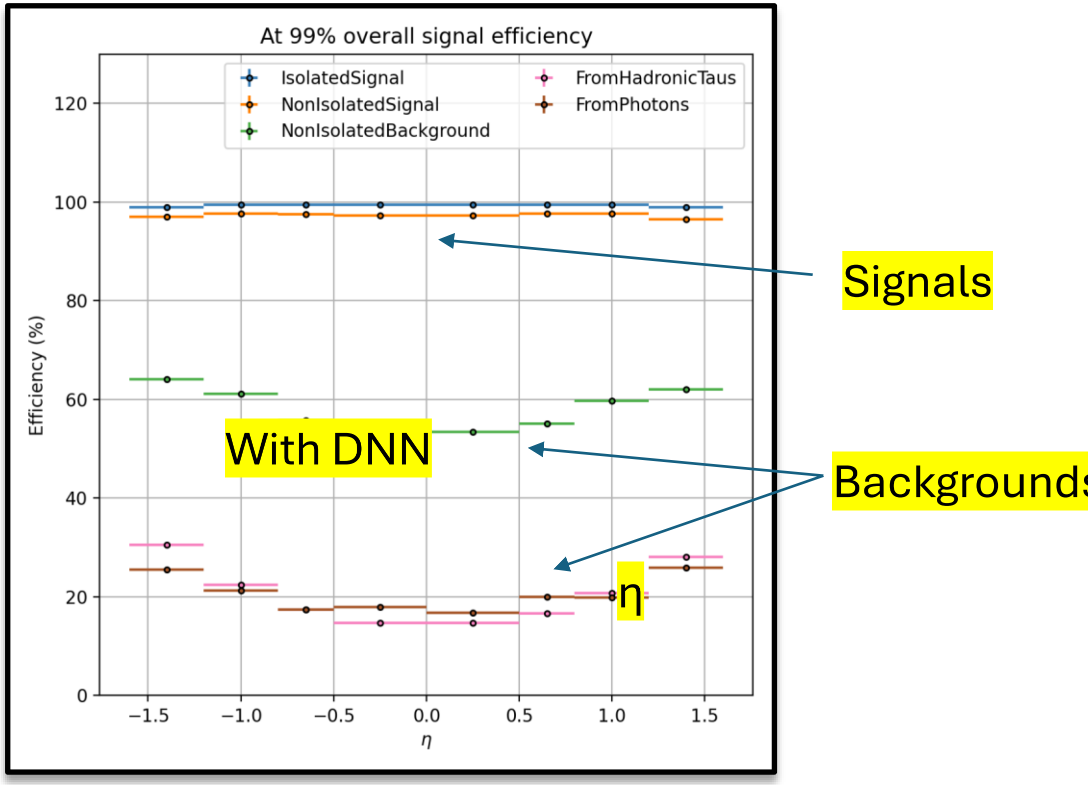
A discriminator is then constructed using 3 of the nodes

Offline Particle Flow

~98-99% signal efficiency

Particle Flow
Central E/Gamma Identification

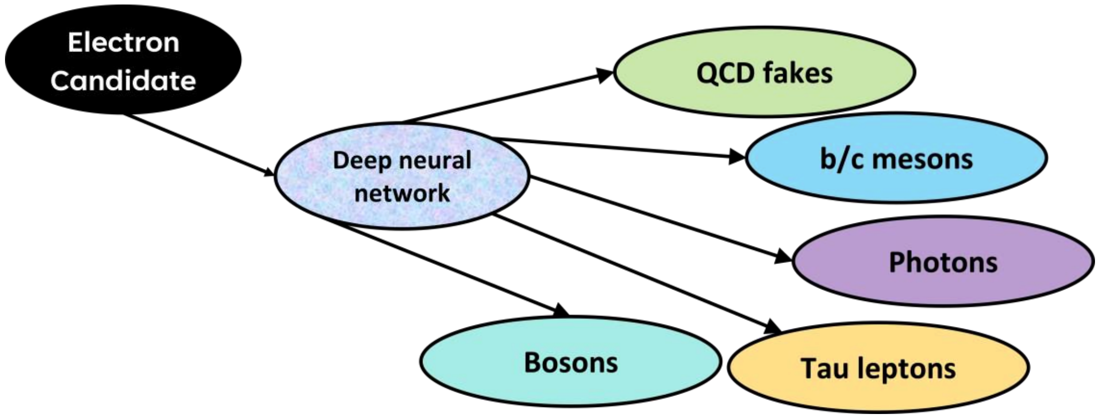


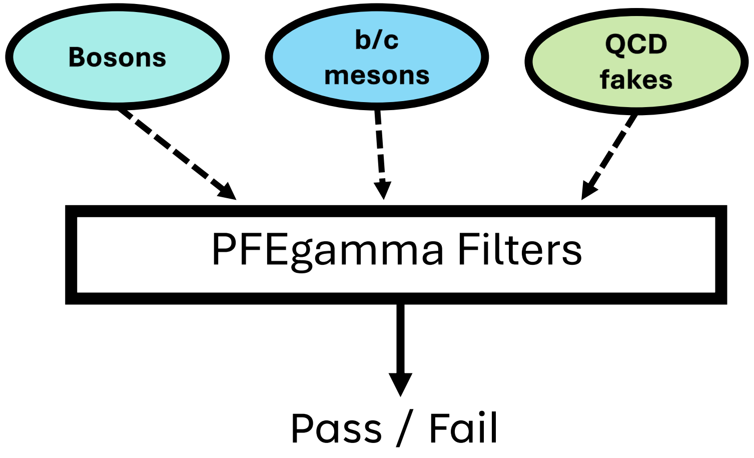


Particle Flow
Central E/Gamma Identification

Offline Particle Flow

~98-99% signal efficiency





Old ID

DNN off

Offline DM \ Gen DM	Other	π	$\pi\pi^0$s	$\pi\pi\pi$	$\pi\pi\pi\pi^0$s
$\pi\pi\pi\pi^0$s	0.12	0.01	0.02	0.07	0.32
$\pi\pi\pi$	0.08	0.01	0.02	0.41	0.20
$\pi\pi^0$s	0.51	0.16	0.56	0.07	0.07
π	0.10	0.76	0.25	0.02	0.00
Other	0.17	0.02	0.09	0.43	0.41
None	0.03	0.05	0.06	0.00	0.00

New DNN ID for Run3

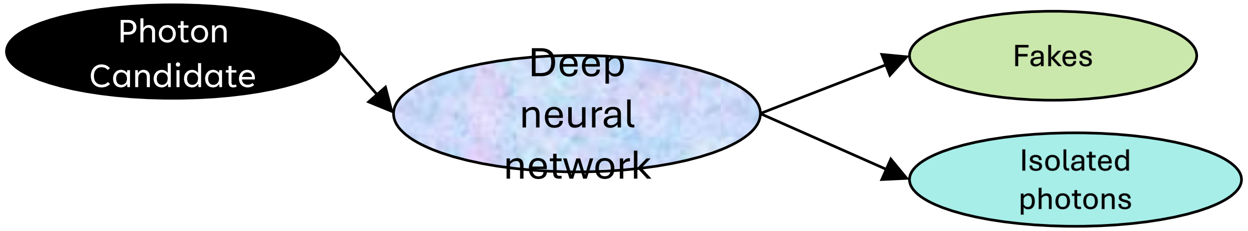
DNN on

Offline DM \ Gen DM	Other	π	$\pi\pi^0$s	$\pi\pi\pi$	$\pi\pi\pi\pi^0$s
$\pi\pi\pi\pi^0$s	0.12	0.01	0.02	0.07	0.32
$\pi\pi\pi$	0.08	0.00	0.02	0.41	0.19
$\pi\pi^0$s	0.54	0.17	0.67	0.07	0.07
π	0.07	0.76	0.15	0.02	0.00
Other	0.17	0.02	0.09	0.43	0.41
None	0.02	0.04	0.05	0.00	0.00

Migration of $\pi^\pm \pi^0$ is to $\pi^\pm$ now reduced by ~11%

Offline
Particle Flow
~98-99% signal efficiency

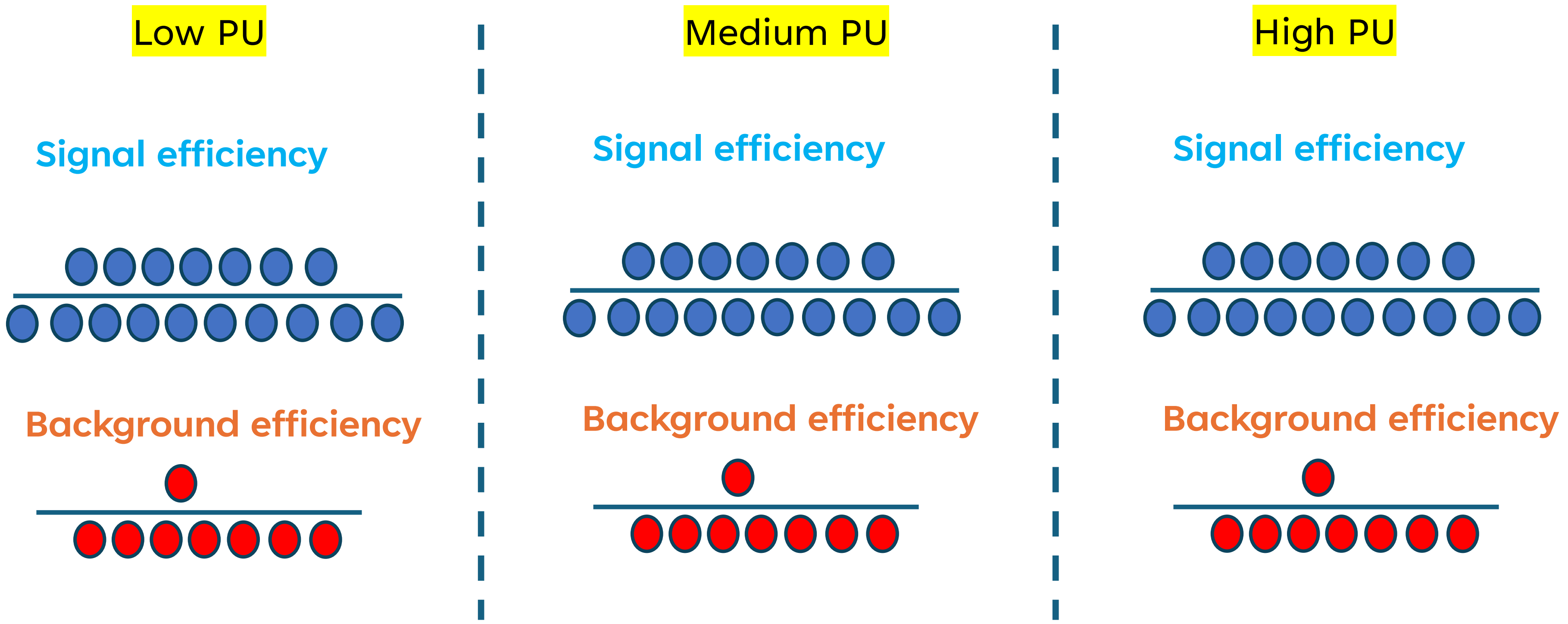
Particle Flow
Central E/Gamma Identification



Currently, a simple binary DNN is in place.
(**Scope for improvement**)

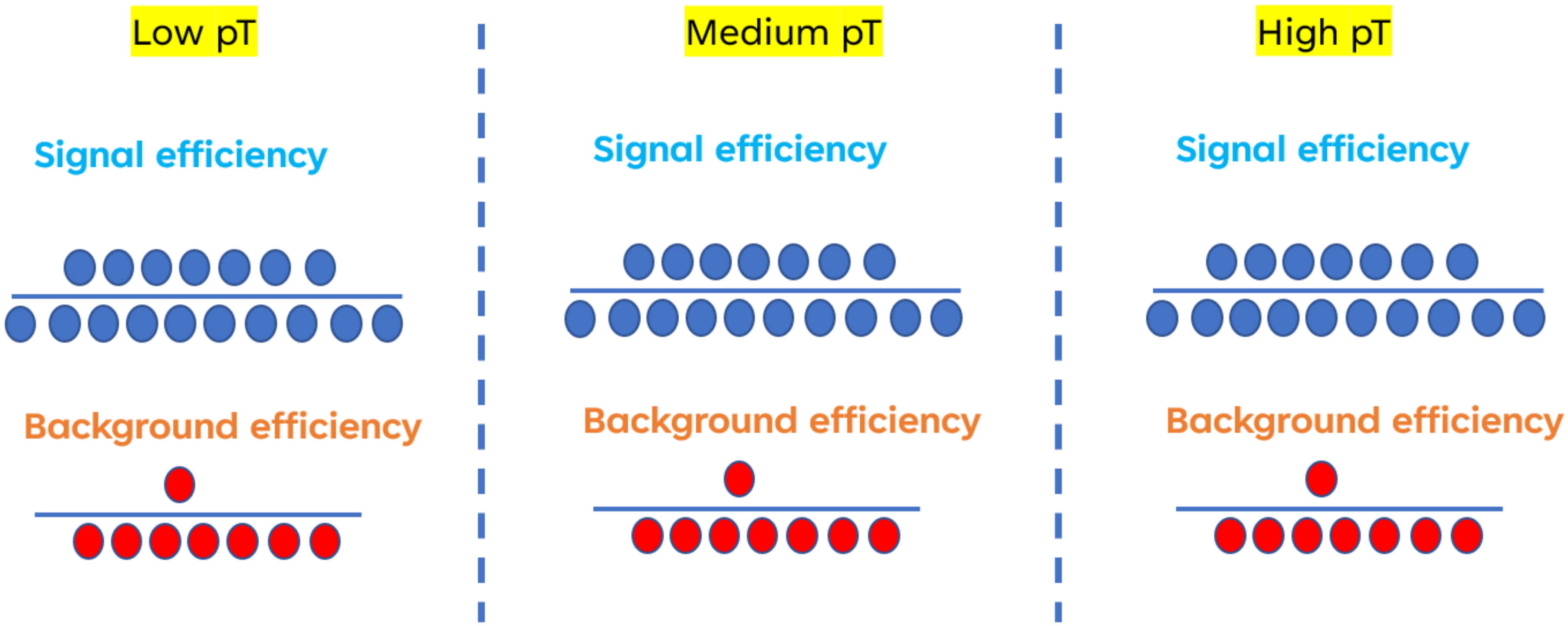
Robustness of identification criteria

An overall efficiency is as important as “stability” of efficiency



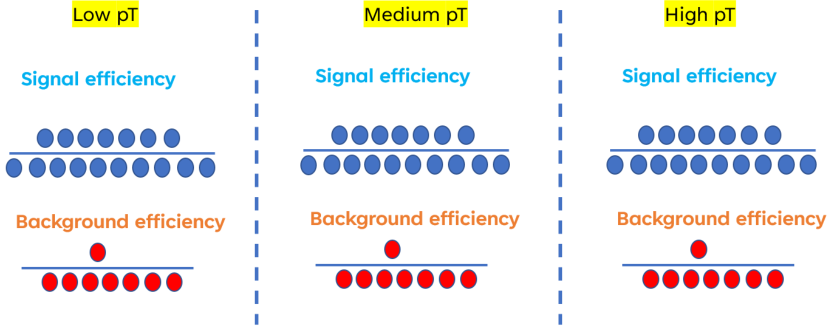
Robustness of identification criteria

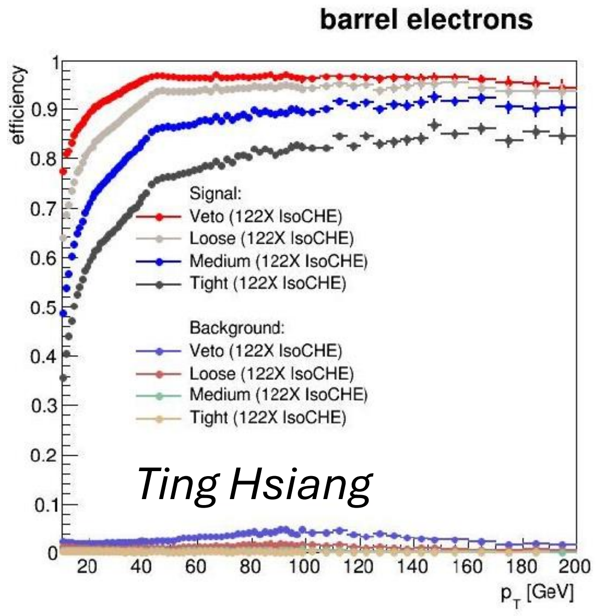
An overall efficiency is as important as "stability" of efficiency

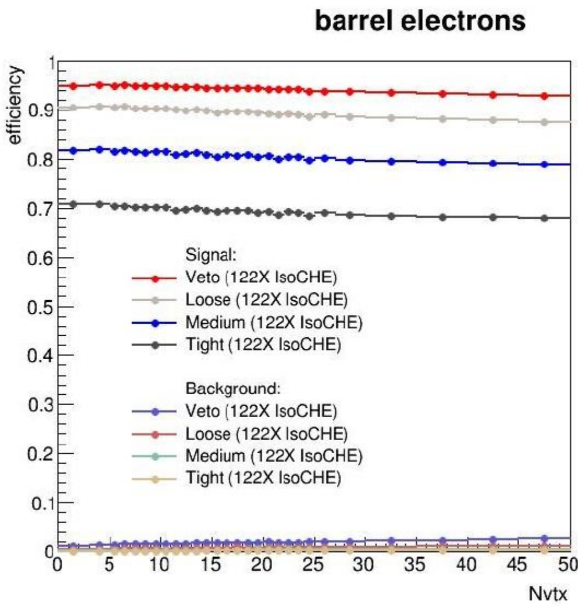


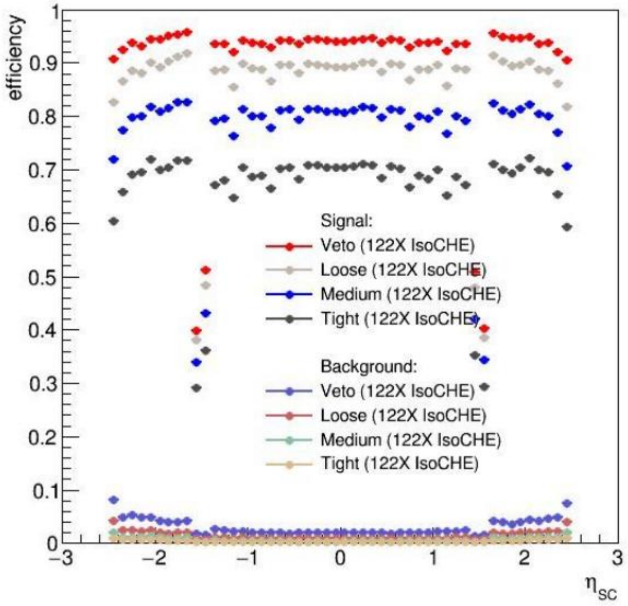
Robustness of identification criteria

An overall efficiency is as important as “stability” of efficiency



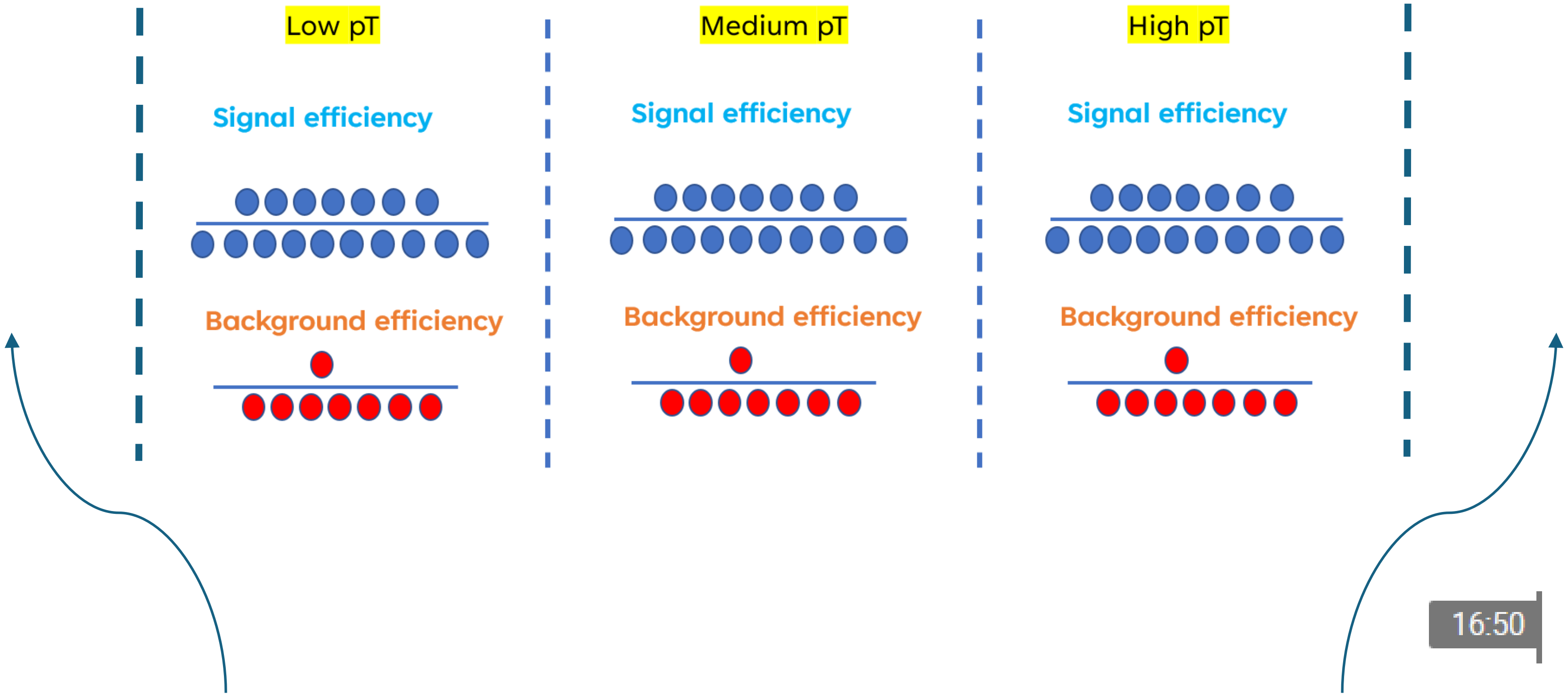






Robustness of identification criteria

An overall efficiency is as important as “stability” of efficiency



Part 3

16:50

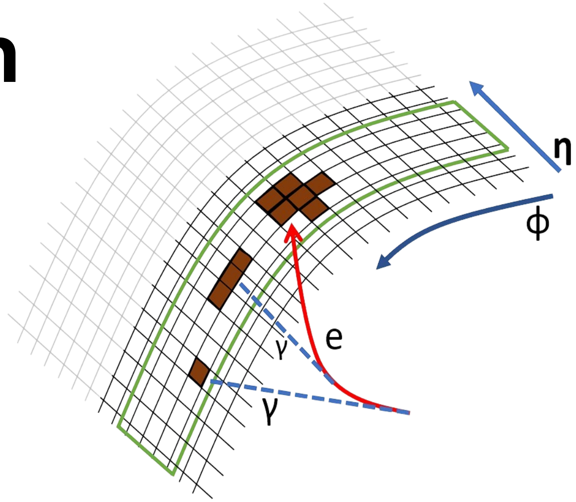
Introduction to HEEP ID

Sam Harper

In extreme scenarios, standard identification criteria become limited in performance and stability

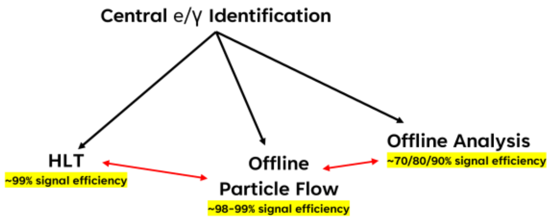
Offline Analysis
~70/80/90% signal efficiency

Conclusion



To identify electrons and photons we rely on:
- ➢Description of the EM shower shape
- ➢Tracking and clustering matching parameters
- ➢Quantification of isolation of these objects

In E/Gamma, identification occurs are three levels: **HLT**, **Particle Flow**, and **Analysis**
The particle flow IDs are “fixed working point IDs”
E/Gamma POG provides several working points and IDs for offline analyses



Additional note

When Central E/Gamma Identification is not sufficient

Note on custom IDs

"Please" consult PAG EGM contacts or directly E/Gamma POG, when "tinkering" with central IDs
Feel free to consult E/Gamma POG, when designing custom IDs (earlier is better!)

When "adding" or "removing" variables from cut-based IDs, please consult us before applying scale factors
In most cases, you will need approval from the contact or EGM POG anyway.

The central IDs are not to be applied blindly! Please check signal and background efficiencies when in doubt.

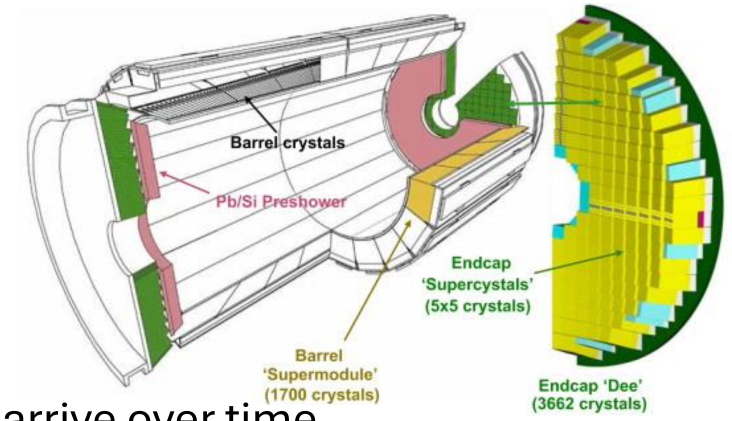
"Do not" blindly apply central ID scale factors if background efficiency is "high"

Contact us:
- **cms-phys-conveners-EGM(at)cern(dot)ch**(Send Email)
- **cms-egamma-l3-convenors(at)cern(dot)ch**(Send Email)
- **cmstalk+egm(at)dovecotmta(dot)cer(dot)ch**(Send Email)

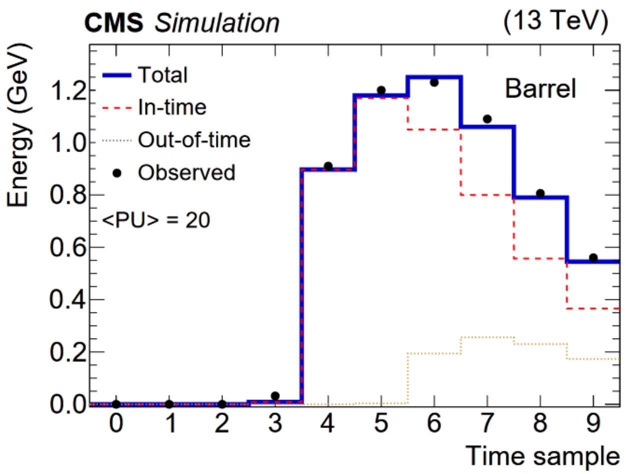
Hope that we can now have a **charged** and **enlightening** discussion!

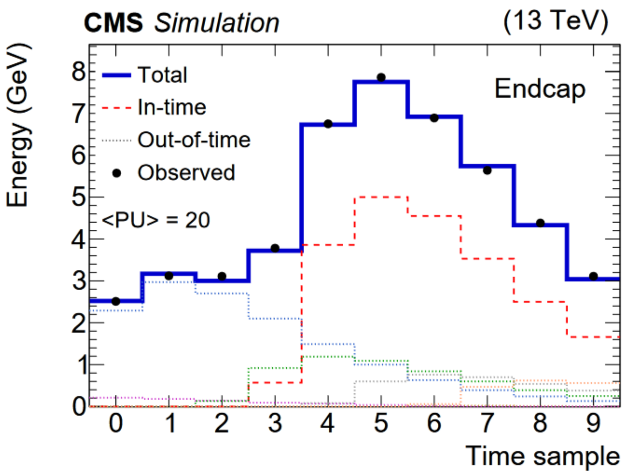
Backup

Ecal Energy reconstruction



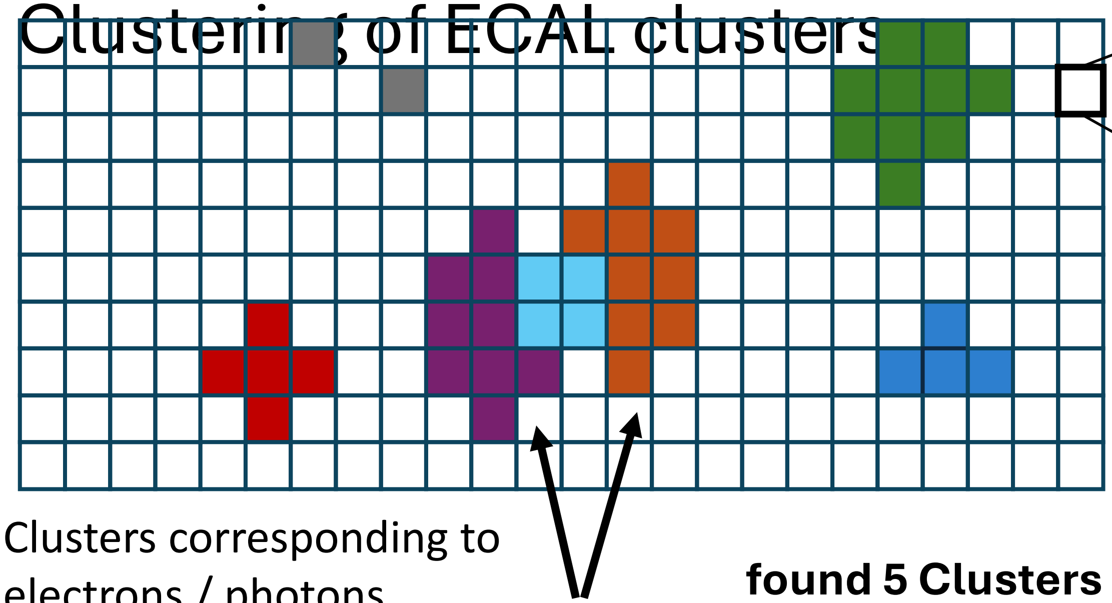
- Energy in the ECAL is observed as a pulse as the photons from the shower arrive over time
- This pulse shape is converted to energy (E ∝ A, signal channel amplitude) In case of out-of-time pile-up events, multiple pulses from different bunch crossings are generated
- They lead to an apparently increased amplitude measurement
- The amplitudes of different pulses are resolved by fitting the multiple pulse shapes simultaneously





Ecal Energy reconstruction

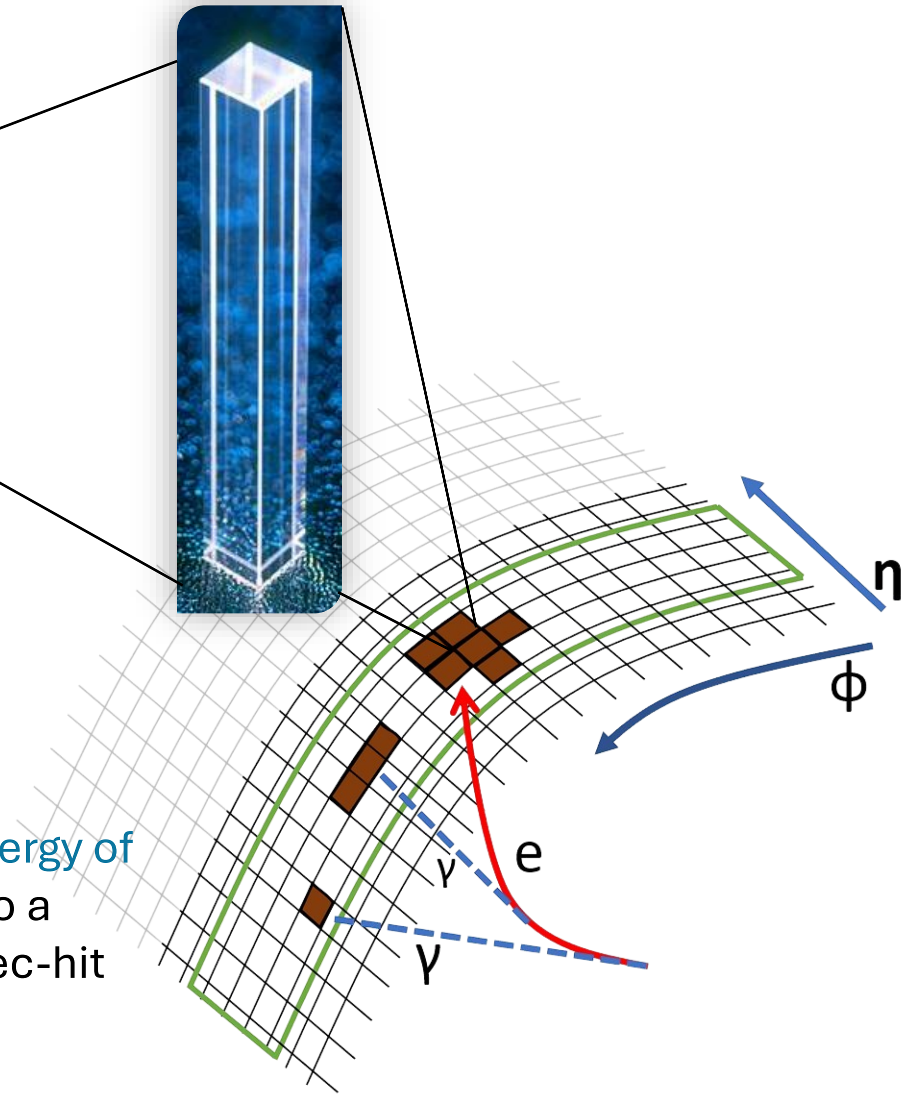
- Clustering of ECAL clusters



Clusters corresponding to electrons / photons

found 5 Clusters

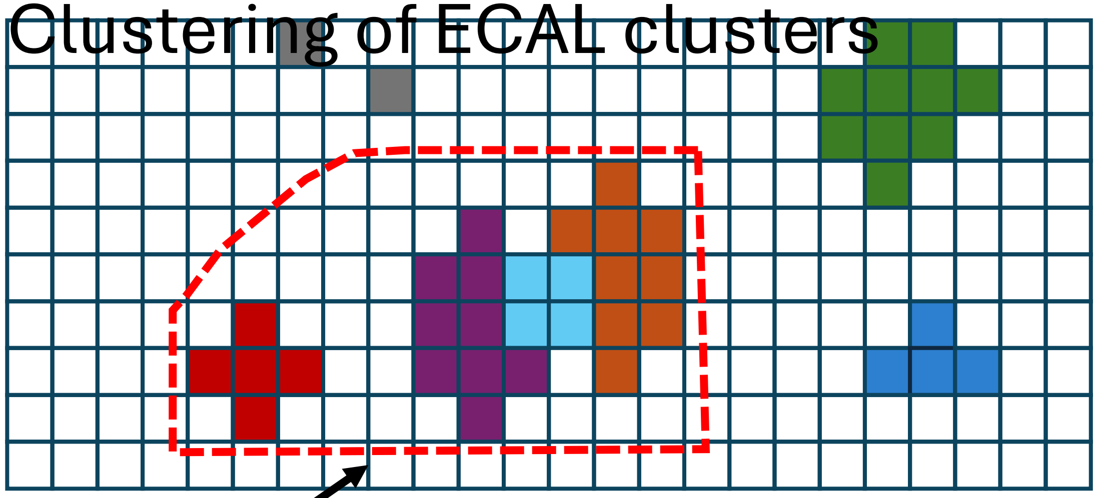
these two clusters overlap, clustering algo shares energy of yellow rec-hits between the two clusters according to a Gaussian energy profile, each gets a fraction of the rec-hit energy



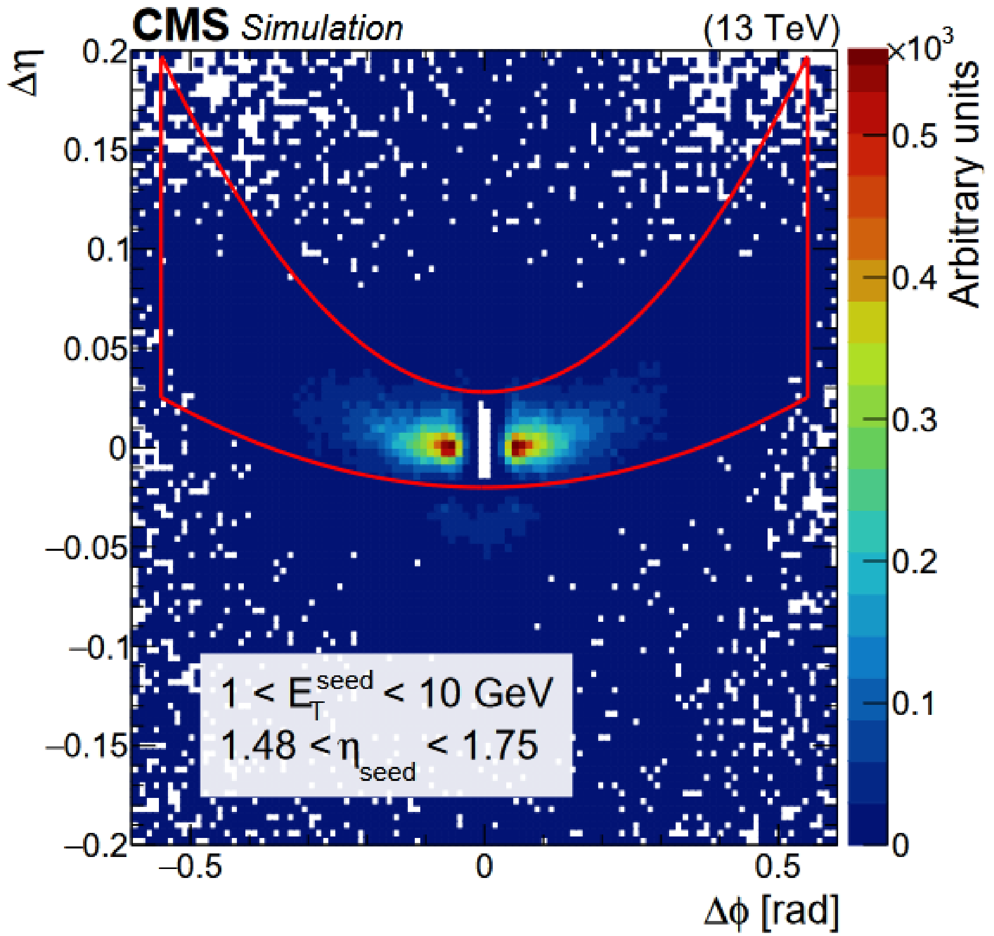
Ecal Energy reconstruction

JINST **16** P05014

- Clustering of ECAL clusters

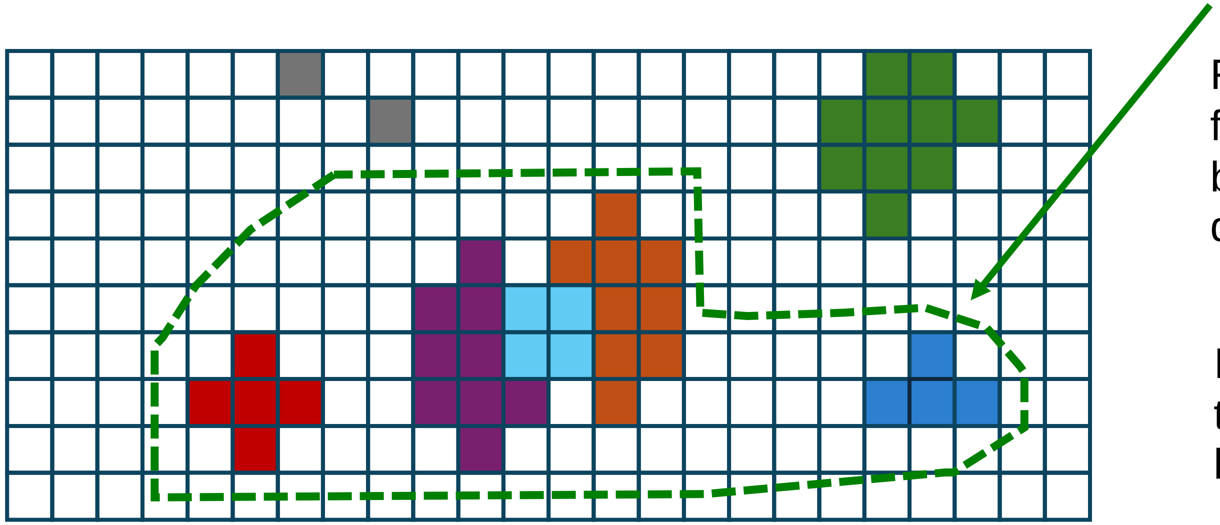


Moustache supercluster
A cluster of clusters



Ecal Energy reconstruction

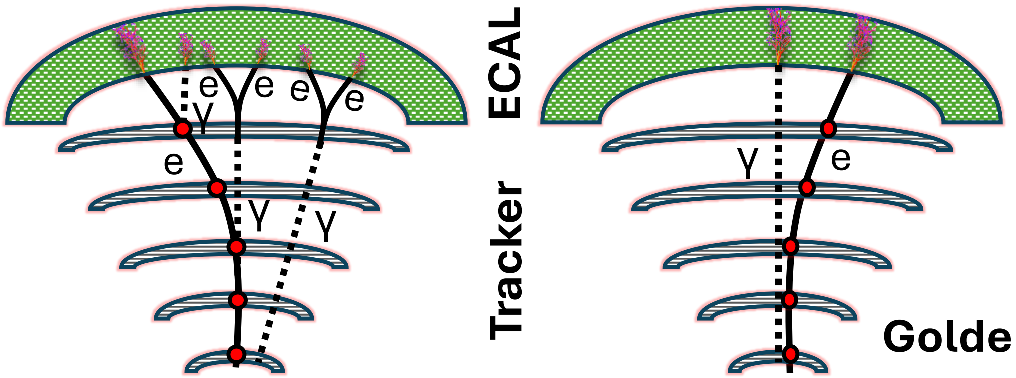
Refined Supercluster



Refined superclusters use the information from the tracker, to be able to link bremsstrahlung emissions to missed ECAL deposits

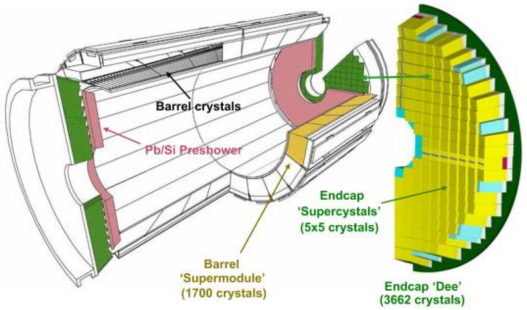
Information from clustering and tracking is used in tandem to achieve best resolution

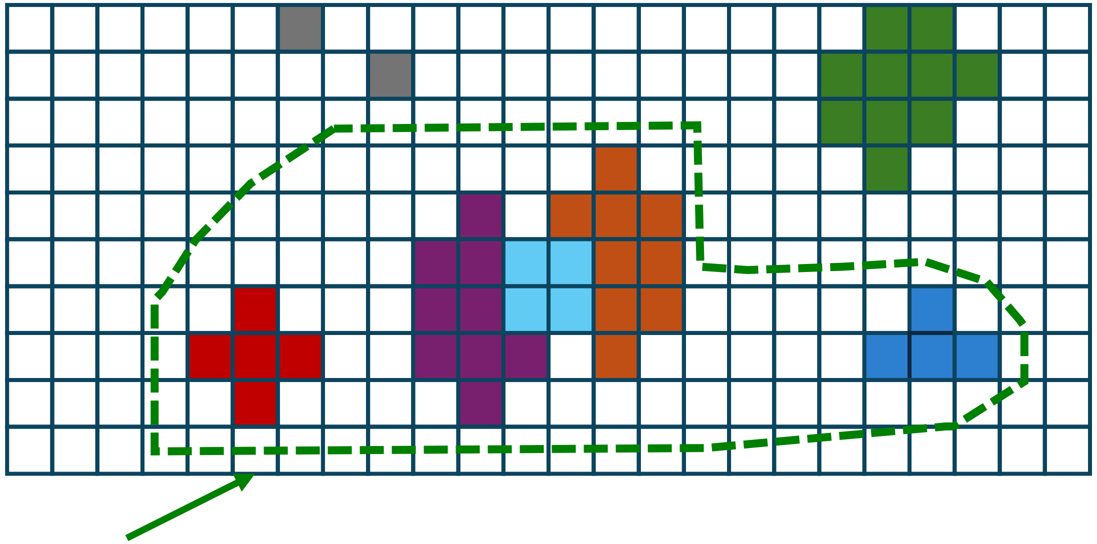
With bremsstrahlung and conversions



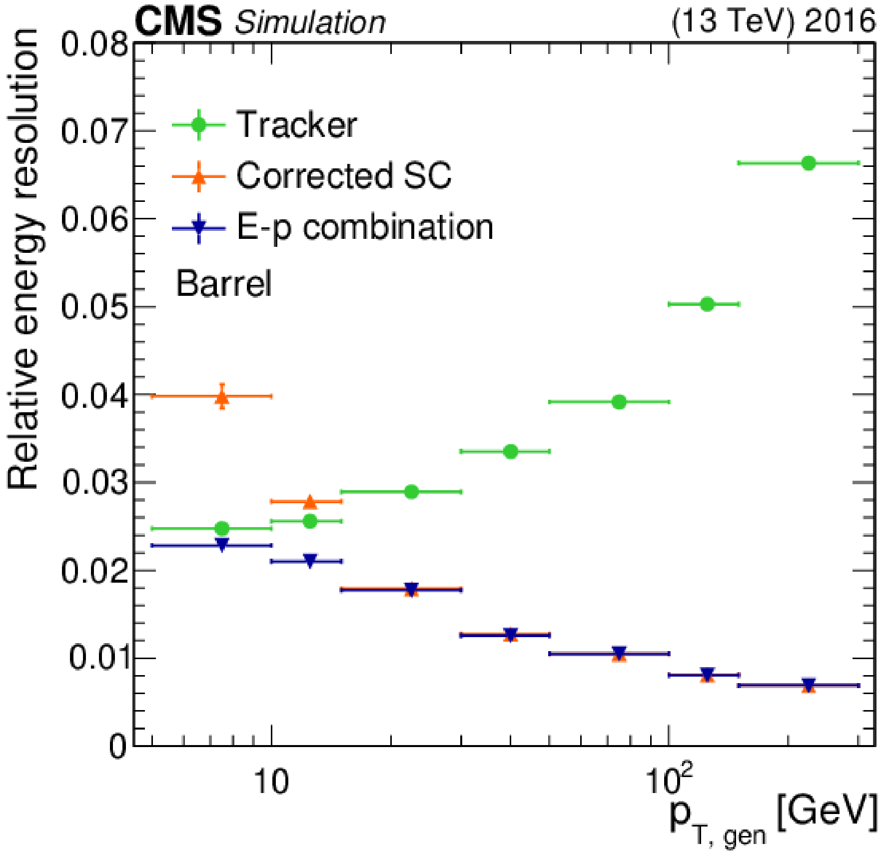
Golden cases

Ecal Energy reconstruction





Refined superclusters use the information from the tracker, to be able to link bremsstrahlung emissions to missed ECAL deposits



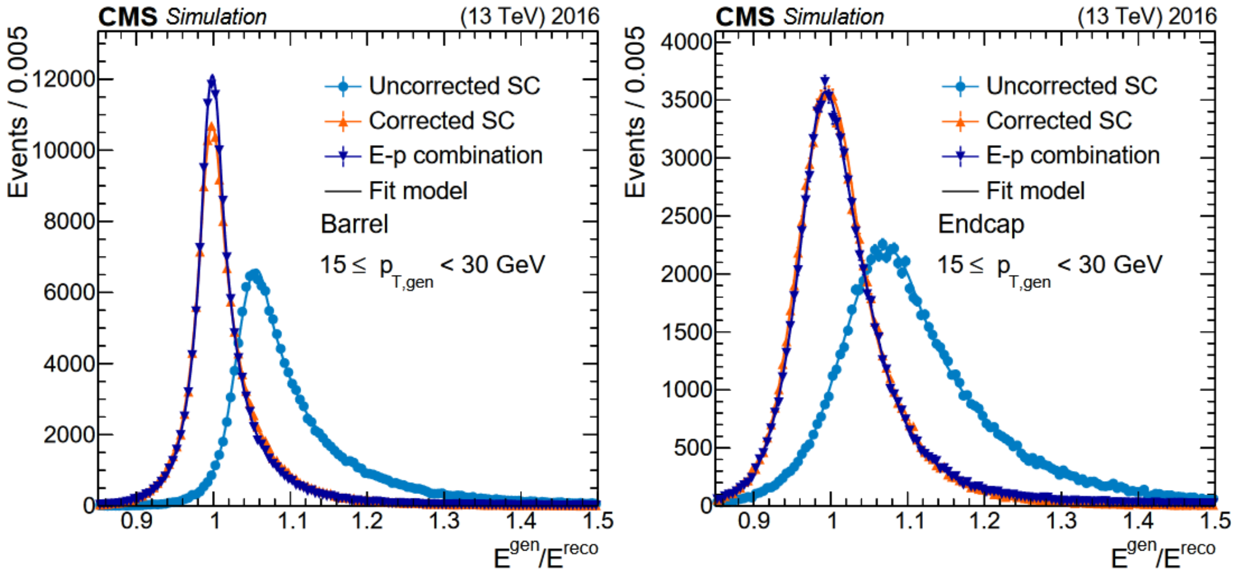
There is also dedicated photon conversion recovery algorithm

Ecal Energy Correction

Several losses occur before electrons and photons deposit energy in the ECAL

- We calibrate the reconstructed energy back to expected original energy using correction procedures
- Employ machine learning in tandem with algorithmic approaches
- Tracker information used for E-p combination

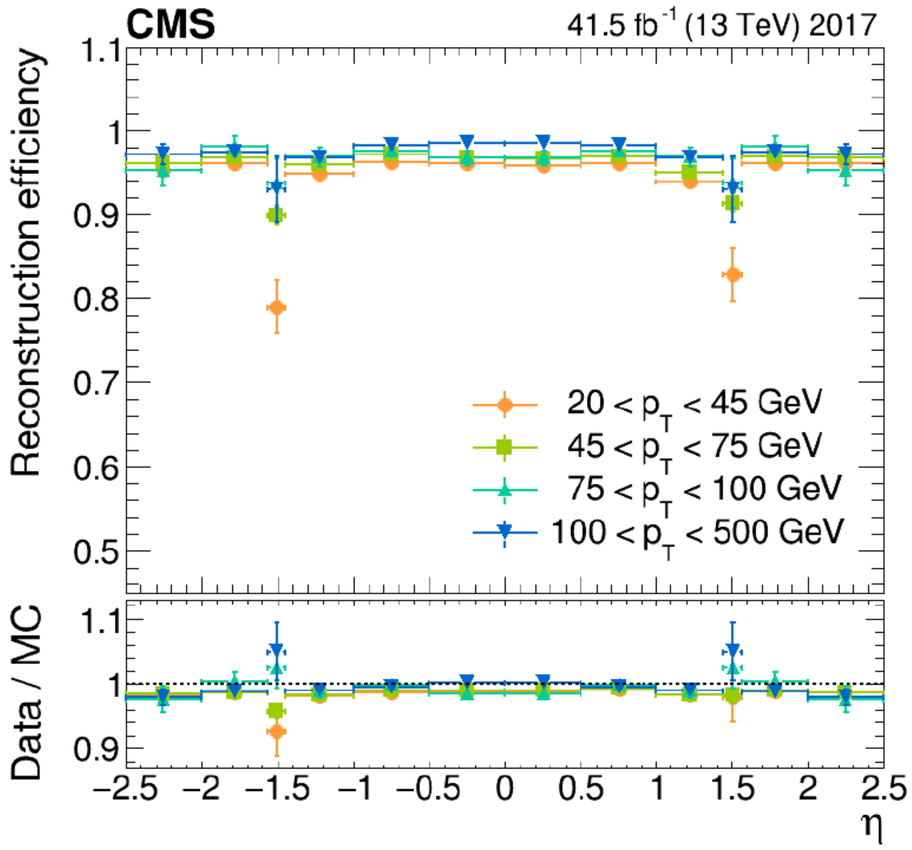
- Any residual energy corrections that maybe needed are derived using **Z→ee** events in data and MC



JINST **16** P05014

Reconstruction performance

Electron reconstruction efficiency is higher than 95% for E_T> 20 GeV and is compatible between data and simulation within 2%



Identification: The good and the bad

Several variables have been developed to separate electrons/photons from background (jets, photon conversions, particles from secondary vertices)

They exploit that electrons/photons are single objects which are almost fully contained in the ECAL

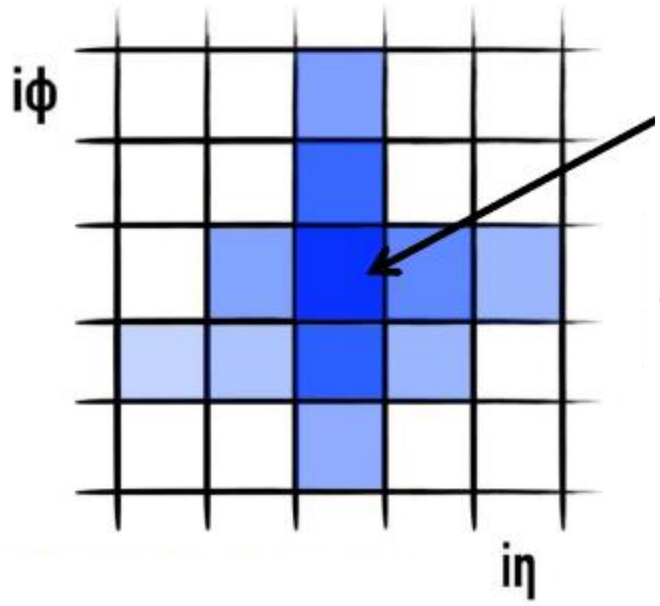
Many different types:

- **Shower-shape variables** → **Are the energy deposits in the calorimeters compatible with coming from a single electron/photon?**
- **Track matching variables** → **Does the ECAL deposit have a compatible track?**
- **Conversion ID variables** → **Are the tracks compatible with coming from the collision point? Or do they appear later on in the tracker?**
- **Isolation variables** → **Is there a large amount of other particles nearby the electron/photon?**

Identification: The good and the bad

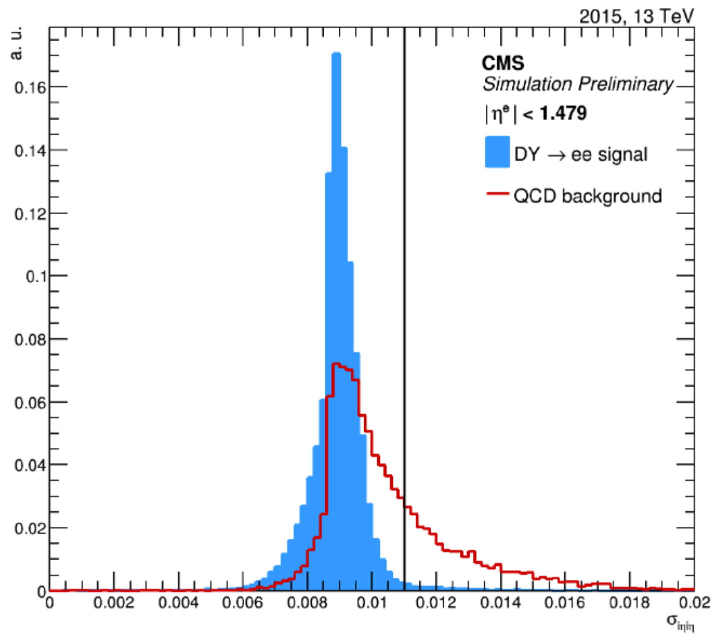
Many different types:

- **Shower-shape variables**
- **Track matching variables**
- **Conversion ID variables**
- **Isolation variables**



Most energetic crystal of 5x5 array

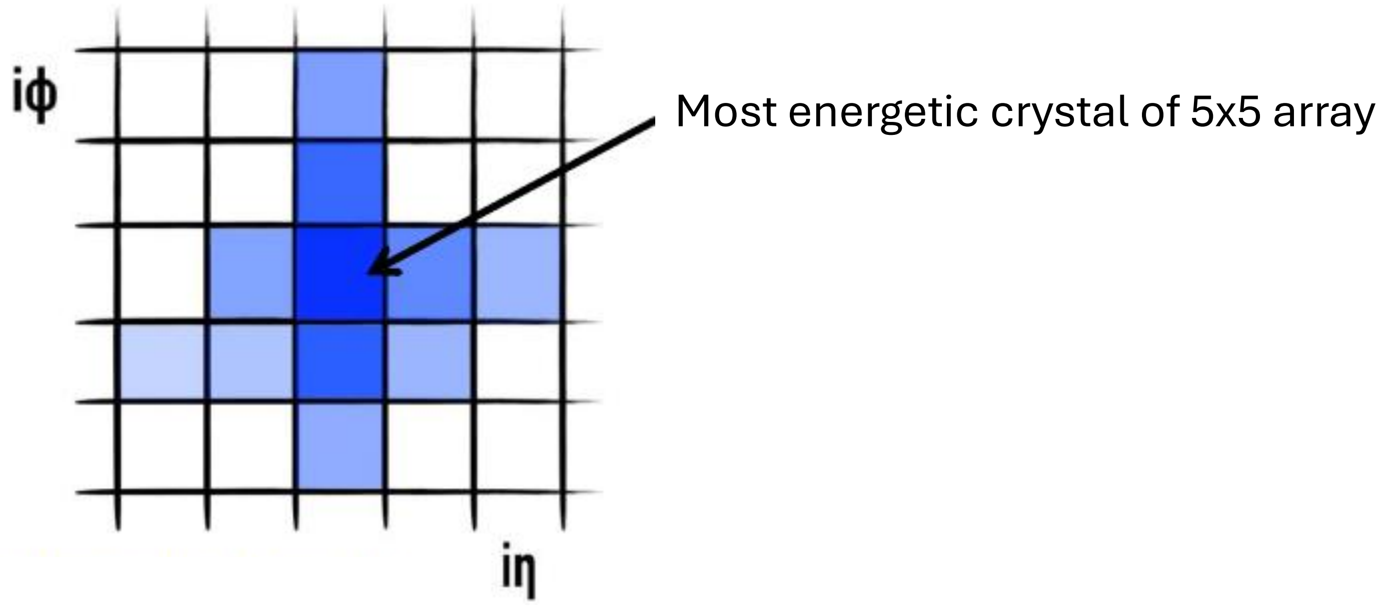
$$\sigma_{i\eta i\eta} = \sqrt{\left(\frac{\Sigma_i^{5\times5} w_i(\eta_i - \bar{\eta}_{5\times5})^2}{\Sigma_i^{5\times5} w_i}\right)}$$



Identification: The good and the bad

Many different types:

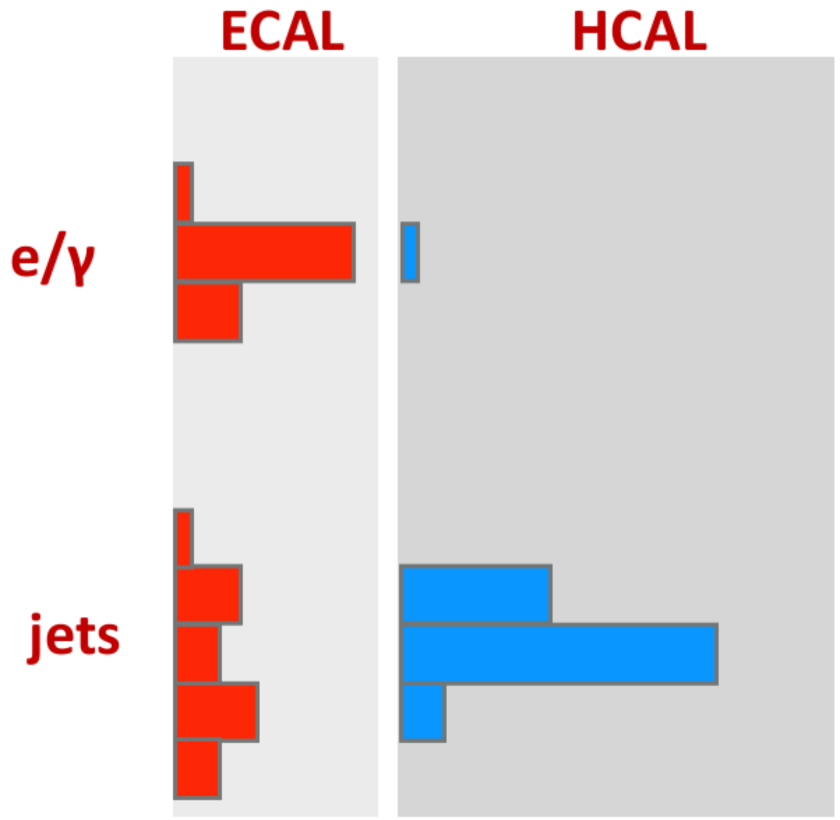
- **Shower-shape variables**
- **Track matching variables**
- **Conversion ID variables**
- **Isolation variables**



Identification: The good and the bad

Many different types:

- **Shower-shape variables**
- **Track matching variables**
- **Conversion ID variables**
- **Isolation variables**



H/E

Identification: The good and the bad

Many different types:

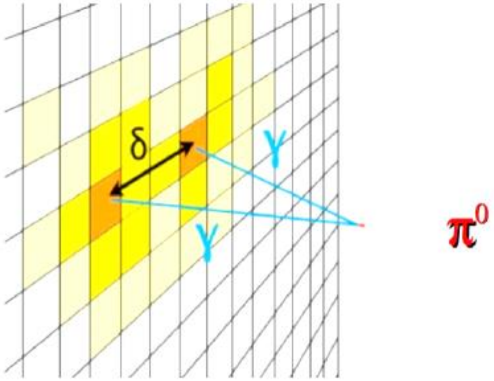
- **Shower-shape variables**
- **Track matching variables**
- **Conversion ID variables**
- **Isolation variables**

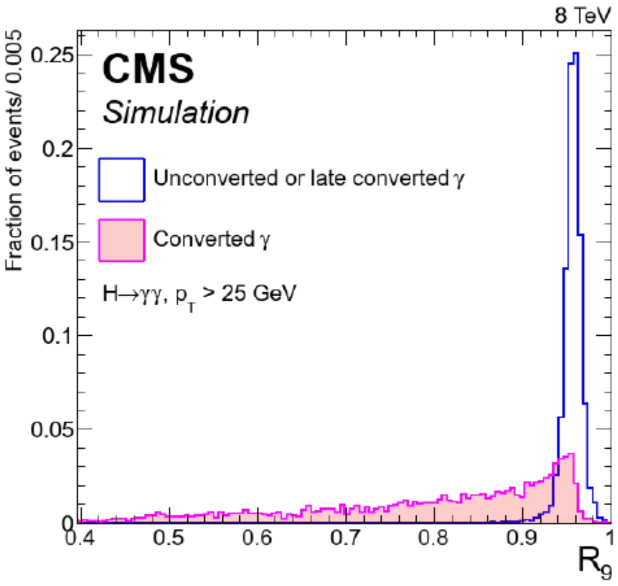
R_9

5x5 matrix contains ~96% (~97%) of unconverted photon energy in EB (EE)

R_9 is the energy sum of the 3×3 crystals centered on the most energetic crystal in the supercluster divided by the energy of the supercluster

R_9 helps in conversions identification and to distinguish real photons from pions





- $\sigma_{i\eta i\eta}$: the second moment of the log-weighted distribution of crystal energies in η, calculated in the 5×5 matrix around the most energetic crystal in the SC and rescaled to units of crystal size. The mathematical expression is given below:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_i^{5\times5} w_i(\eta_i - \overline{\eta}_{5\times5})^2}{\sum_i^{5\times5} w_i}}. \tag{7.1}$$

What can we use to identify electrons and photons?

- Description of the EM shower shape
- Tracking and clustering matching parameters
- Quantification of isolation of these objects