Electron reconstruction in the CMS Phase-2 High Granularity Calorimeter Endcap

Théo Cuisset
Laboratoire Leprince-Ringuet, École Polytechnique

Journées de Rencontre Jeunes Chercheurs
Saint-Jacut-de-la-Mer
November 25th, 2024

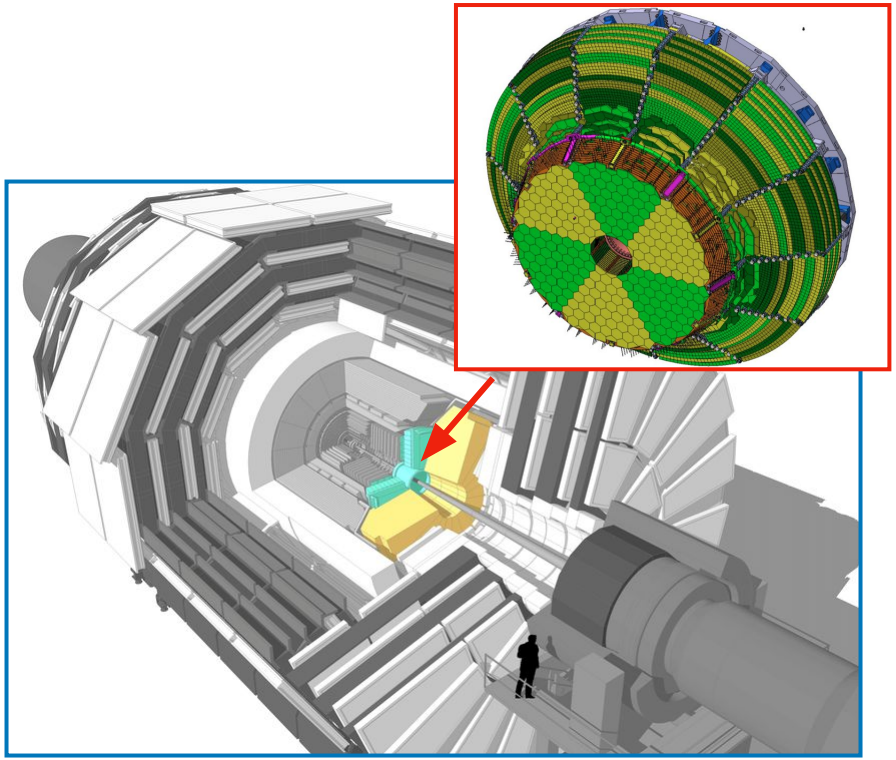
The High Granularity Calorimeter (HGCAL)

- To enhance the potential for discoveries at the LHC, CERN plans to upgrade the accelerator to the **high luminosity LHC** (HL-LHC) by 2029.

- At the end of LHC Phase 1 operations, the end-cap calorimeters will have suffered irrecoverable radiation damage.

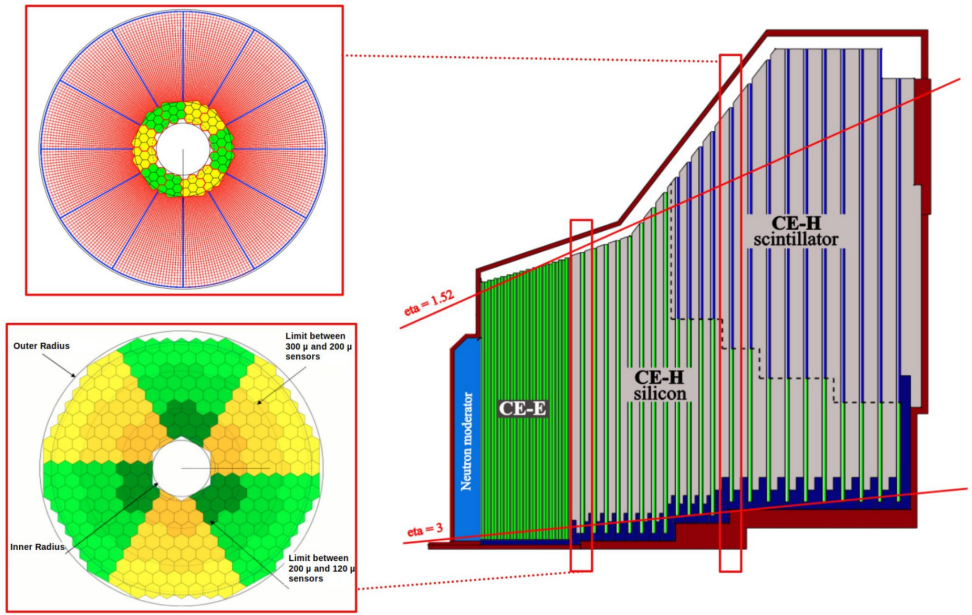
The CMS collaboration is replacing its endcap calorimeters with a new sampling

High Granularity Calorimeter.



The CMS detector and position of HGCAL

The High Granularity Calorimeter (HGCAL)

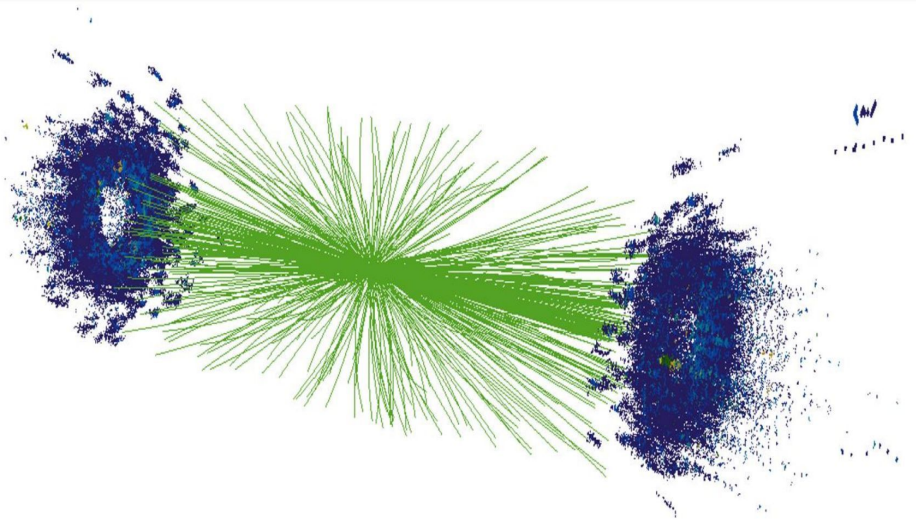


Longitudinal cross section of the upper half of one of the HGCAL (right), mixed silicon and scintillator layers (top left) and a full silicon layer (bottom left).

- First **imaging calorimeter** to be implemented in high-energy physics.
- **Position, energy and timing** information.
- **26** layers of electromagnetic section and **21** layers of hadronic section.
- Designed for the HL-LHC target integrated luminosity of **3000 fb^{-1}.**

Machine Learning in HGCAL

HGCAL : High granularity calorimeter designed for Particle Flow (→ combining tracker and calorimeter information)



Simulated event display of HGCAL at 200 pileup. The interesting hard process will be contaminated by hadrons coming from soft p-p interactions.

Challenges:

- Environment from the **high pile-up** : number of interactions per bunch crossing will be increased to **140-200.**
- **6 million** read-out channels and 5D information from them.

Novel **machine-learning based algorithms** to tackle these challenges!

Hybrid approach : combining Machine Learning and traditional algorithms to keep execution time at the **minimum**

Reconstruction in HGCAL : TICL framework

TICL : The Iterative CLustering is a new framework is designed to fully exploit the high spatial resolution and precise timing information provided by HGCAL

Rechits
Sensor readout
***O*(1M)** per event

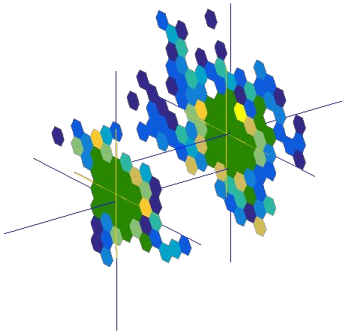
CLUE algorithm
on **GPU**

Layer clusters
Transverse section of a shower
***O*(100k)** per event

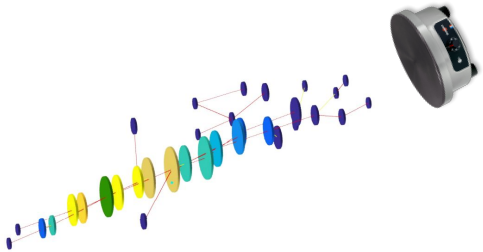
CLUE3D algorithm
soon on **GPU**

Tracksters
Full reconstructed shower of a particle entering the calorimeter
***O*(1k)** per event

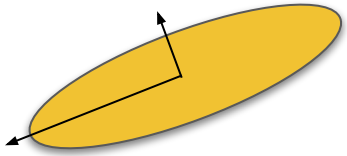
Position x,y,z
Energy, time



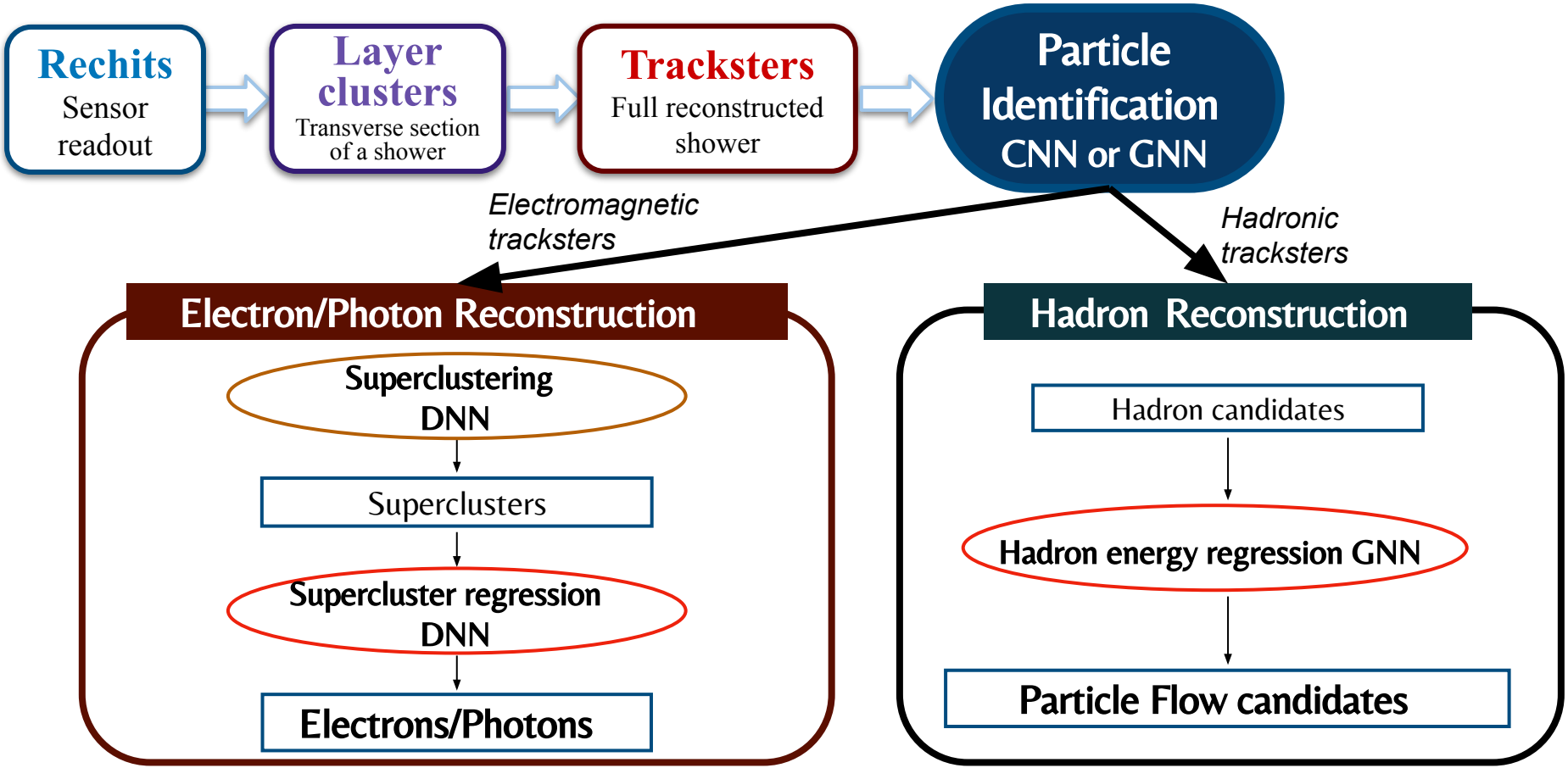
Position x,y, layer
Energy
Average time



Position (η-φ)
Pointing angle
Energy
Average time
Transverse & longitudinal spread



Reconstruction in HGCAL : TICL framework



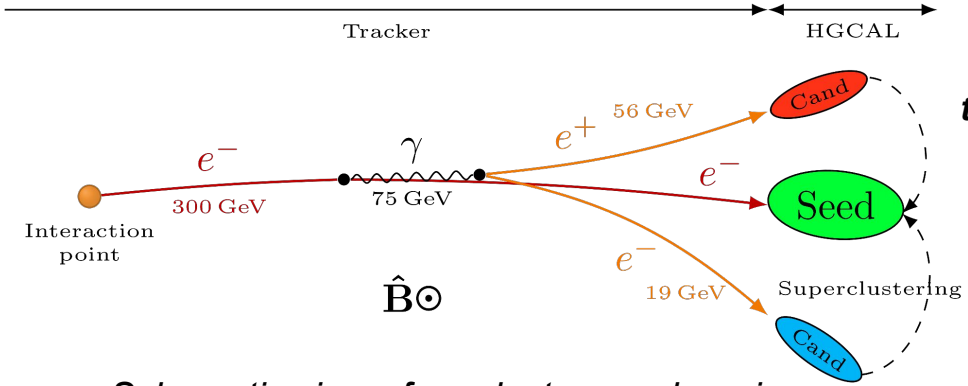
Electrons in CMS

An **electron** produced at the interaction point has to travel through the **tracker** before reaching the calorimeter:

- significant **material budget** in the tracker
- **magnetic field** of 3.8T

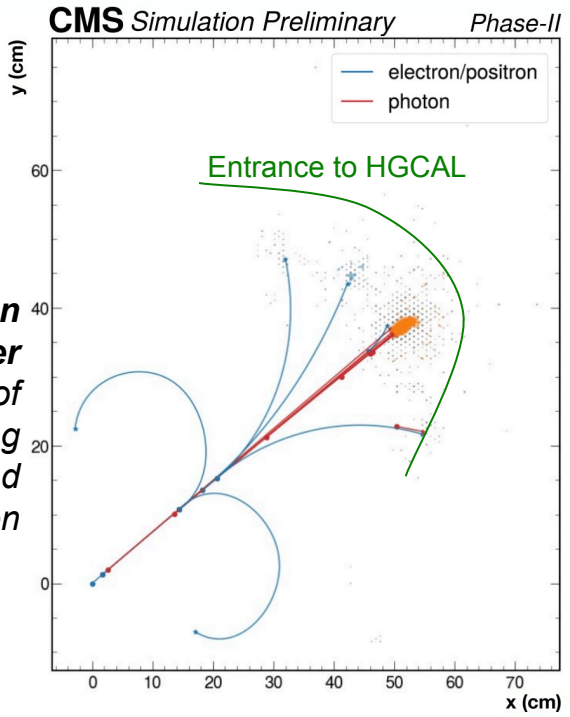
Thus an electron emits several **bremsstrahlung photons** whilst travelling through the tracker

➞ Several particles enter the calorimeter, spread in φ due to the magnetic field.



Schematic view of an electron undergoing bremsstrahlung

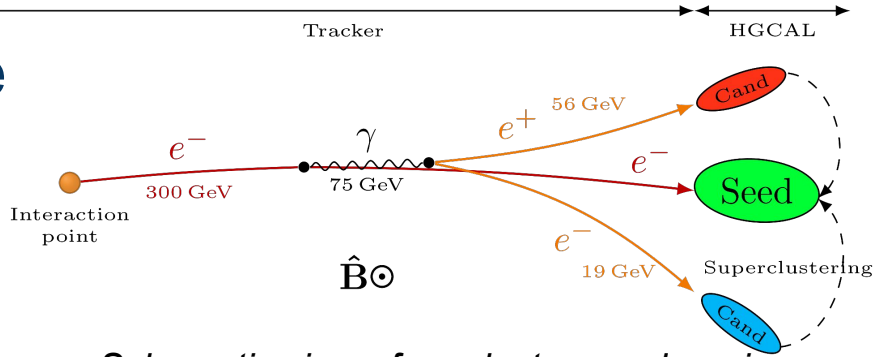
*Display of an **electron travelling in the tracker** until the entrance of HGCAL, undergoing bremmstrahlung and pair-production*



Electron superclustering: principle

For each electron, **several particles enter the calorimeter**, spread in φ due to the magnetic field. They need to be **reconnected to the parent electron**, to:

- seed tracker electron reconstruction
- collect all the electron energy, minimizing pileup contribution



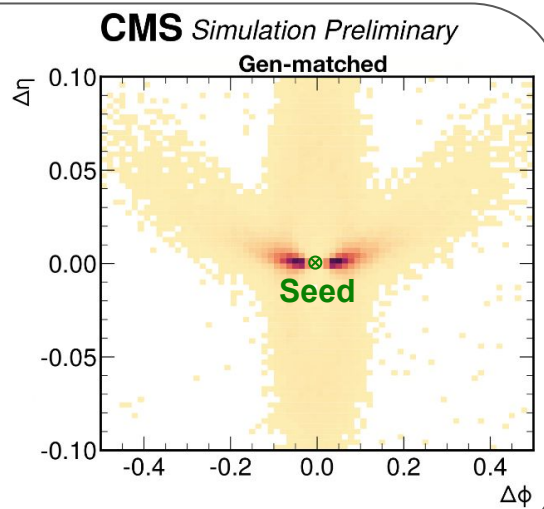
Schematic view of an electron undergoing bremsstrahlung

Need to strike a balance between:

- Seeding too many electrons: will lead to slow reconstruction, as **tracker electron reconstruction is resource-intensive**
- Missing some clusters: might lose some electrons

Spread in η-φ of the bremsstrahlung clusters in the calorimeter
(overlapping multiple events, **centered on the seed**)

Most bremsstrahlung spread in φ (red areas)
The elliptical halo is due to imperfect clustering

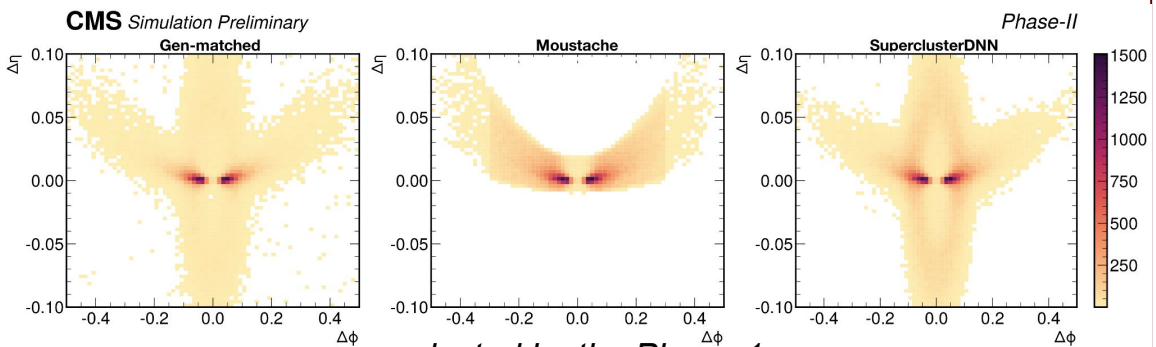


Electron superclustering algorithms

Phase-1 superclustering algorithm: **geometrical algorithm** known as Mustache. Uses solely $\Delta\eta$-$\Delta\varphi$: no longer suited to high pile-up environment of HGCAL

New **Deep Neural Network**-based algorithm for Phase-2, using the **imaging capabilities of HGCAL**

Spread in η-φ of the bremsstrahlung clusters in the calorimeter
(overlapping multiple events, centered on the seed)



at generator level

selected by the Phase-1 geometrical algorithm (Mustache)

selected using the new DNN approach

Superclustering algorithm

Input: **tracksters** of the event
Output : **superclusters** = collection of tracksters representing a calorimeter electron with all its brem deposits

DNN-based superclustering

Inference is run on **pairs of tracksters**, using position and angular variables.

Input features to the DNN

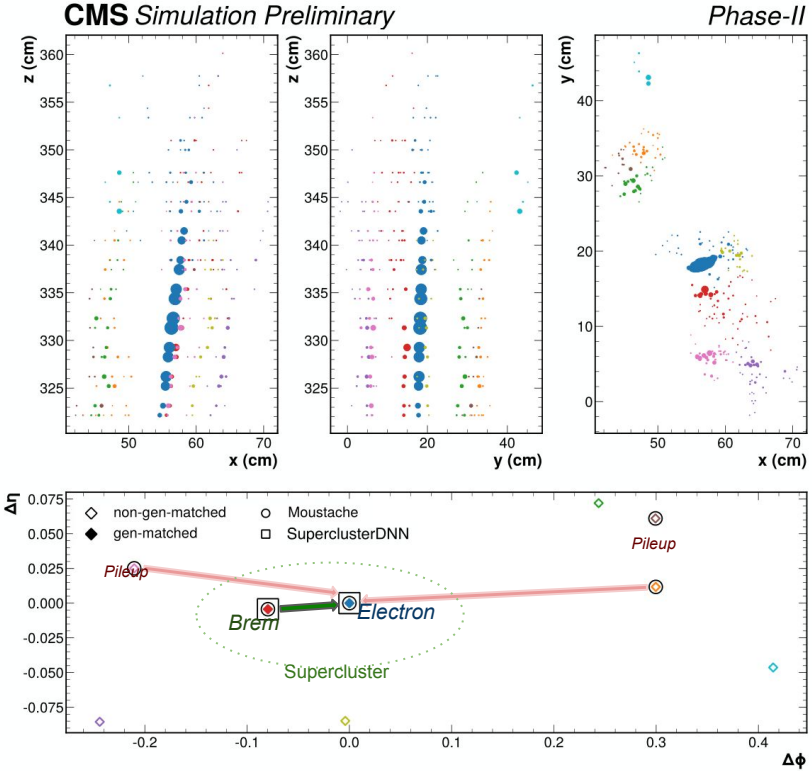
$\lvert\eta_{\text{cand}}\rvert - \lvert\eta_{\text{seed}}\rvert$	E_{cand}	ϕ_{seed}	$\theta_{yz}^{\text{seed frame}}$
$\phi_{\text{cand}} - \phi_{\text{seed}}$	E_{seed}	$\theta^{\text{CMS frame}}$	$\max_{i\in\{1\ldots3\}} \lambda_i$
η_{cand}	$p_{T,\text{cand}}$	$\theta_{xy}^{\text{seed frame}}$	$\frac{\max_{i\in\{1\ldots3\}} \lambda_i}{\sum_{j\in\{1\ldots3\}} \lambda_j}$
η_{seed}	$p_{T,\text{seed}}$	$\theta_{xz}^{\text{seed frame}}$	

Geometrical features

Energies & momentum

Angular variables

Eigenvalues from Principal Component Analysis



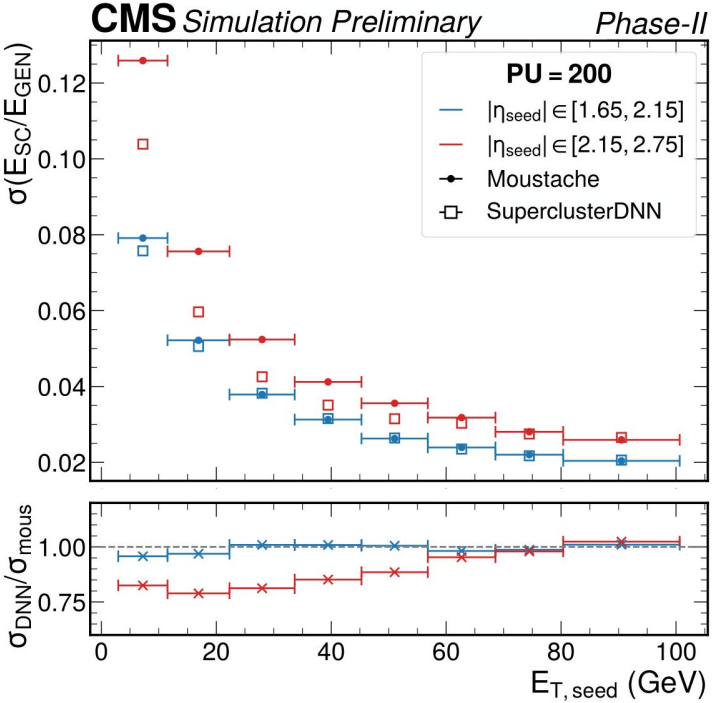
Superclusters are built iteratively, placing a threshold on the score.

Electron superclustering performance

Dataset : electrons in 200 pileup

Supercluster energy is compared to the **true Monte Carlo energy** in the calorimeter. **Energy resolution plots** are obtained, comparing the superclustering DNN to the Moustache geometric algorithm.

The DNN shows **improved resolution at low energy and high η**, thanks to pileup rejection.



The Deep Neural Network allows for **better energy resolution** and **pileup rejection**, allowing better reconstruction efficiency and lower energy thresholds

Conclusion

The CMS collaboration is replacing its **endcap calorimeters** with the HGCAL for HL-LHC, in order to cope with the **high pile-up environment.**

The electron reconstruction is being upgraded, with new algorithms making use of machine learning techniques.

A new **Deep Neural Network-based electron superclustering** algorithms allows for improved pile-up rejection and lower energy thresholds in electron reconstruction.

Future plans

- Develop use of **timing information** provided by HGCAL for pileup rejection.
- Use more complex ML architectures for offline reconstruction.