

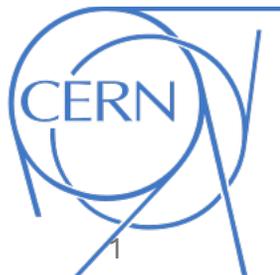
Lepton Photon 2025
Madison, US

Performance of the CMS ECAL and of electron and photon reconstruction in Run3

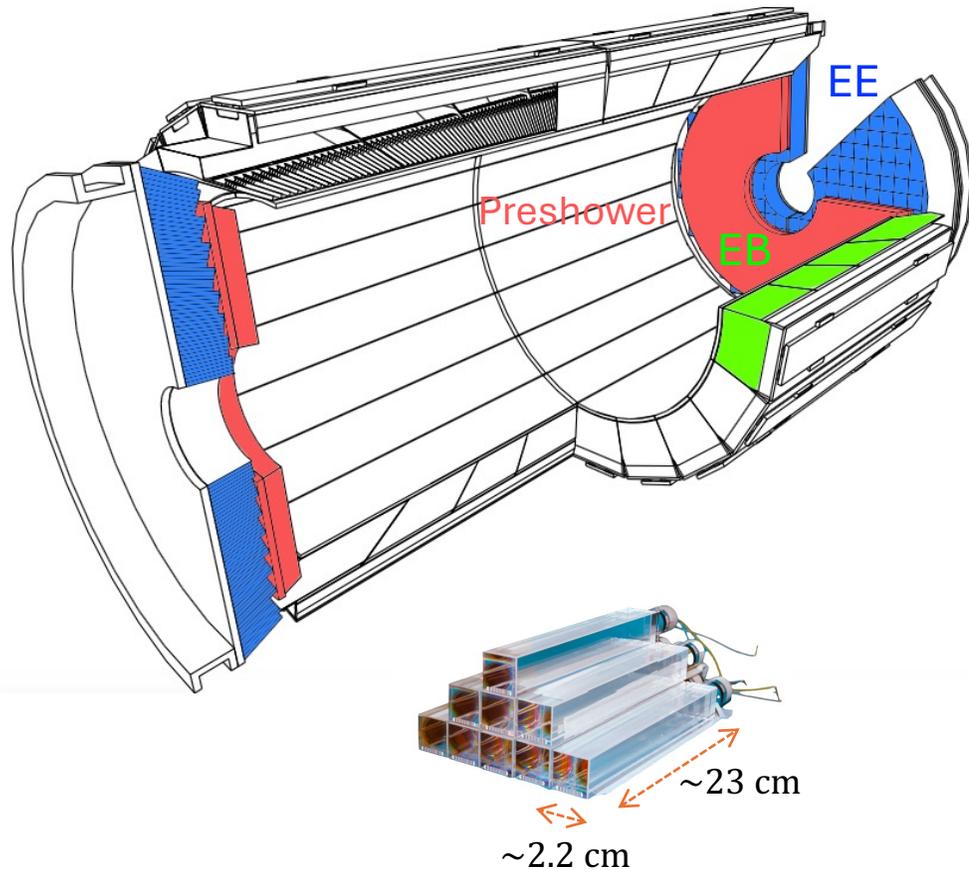
Riccardo Salvatico¹

on behalf of the CMS Collaboration

¹ CERN



The CMS electromagnetic calorimeter



Homogeneous calorimeter

- Barrel (EB) and endcaps (EE) made of $\sim 75k$ PbWO_4 crystals;
- Additional double-layer lead/silicon strip detector in front of EE (*preshower*) for better γ/π^0 discrimination;
- Placed within the volume of the CMS solenoid magnet.

Crystals and readout electronics

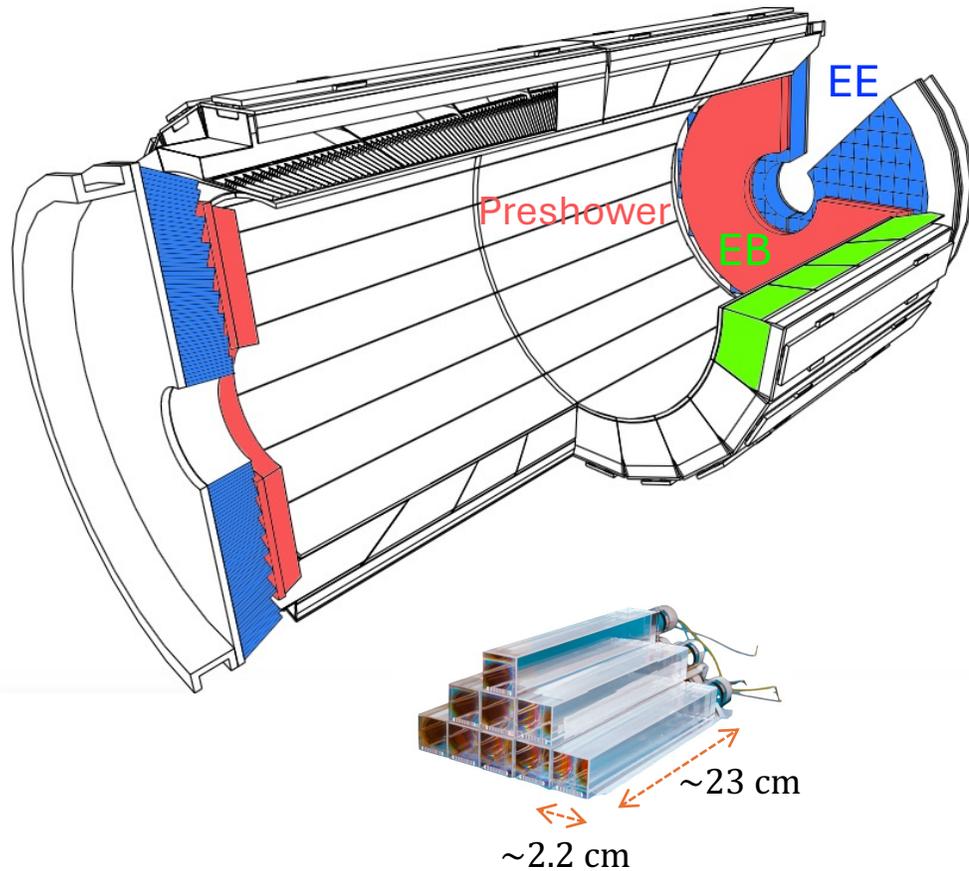
- Crystal properties: $X_0 = 0.89$ cm; $R_M = 2.2$ cm;
- Radiation-hard silicon photodetectors (APD in EB, VPT in EE) with high gain and low noise;
- fast electronics.

Operations

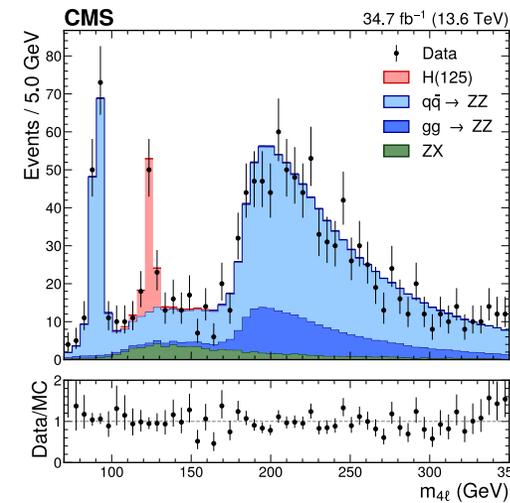
- Operated at 18°C to mitigate photodetector noise;
- Crystal transparency monitored using laser system;
- Continuously calibrated using “standard candle” physics processes ($Z \rightarrow ee, \pi^0 \rightarrow \gamma\gamma$).

The CMS electromagnetic calorimeter

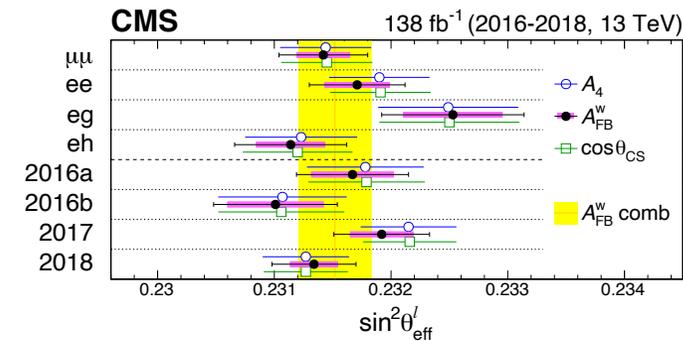
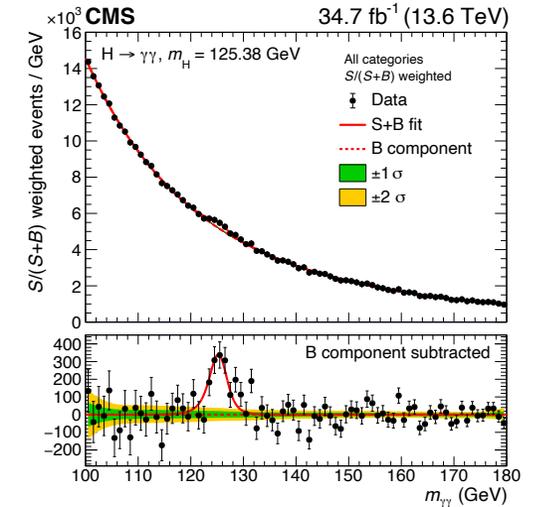
■ ■ ■ The **ECAL** and its reconstructed **electrons** and **photons** played and still play a crucial role in the **CMS physics program**.



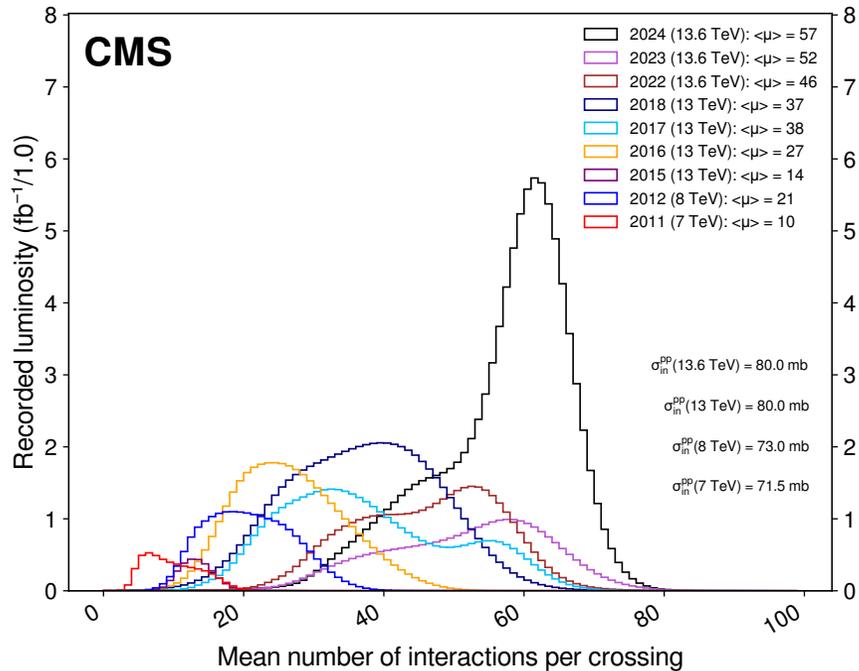
[CMS-HIG-23-014](#)



[CMS-HIG-24-013](#)



The Run 3 challenges



LHC initial design: luminosity up to 1×10^{34} , pileup up to 25

■ ■ ■ A record-breaking data taking

- Instantaneous luminosity 2×10^{34} (Run 2) $\rightarrow 2.6 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (Run 3)
- Center-of-mass energy 13 $\rightarrow 13.6 \text{ TeV}$
- Mean pileup increasing (46 in 2022, 52 in 2023, 57 in 2024)
- Lumi-leveling at PU 60-65, with hardware trigger rate up to $\sim 115 \text{ kHz}$
- Paramount to keep excellent detector performance (energy and time resolution) and operability (limit data acquisition dead time).

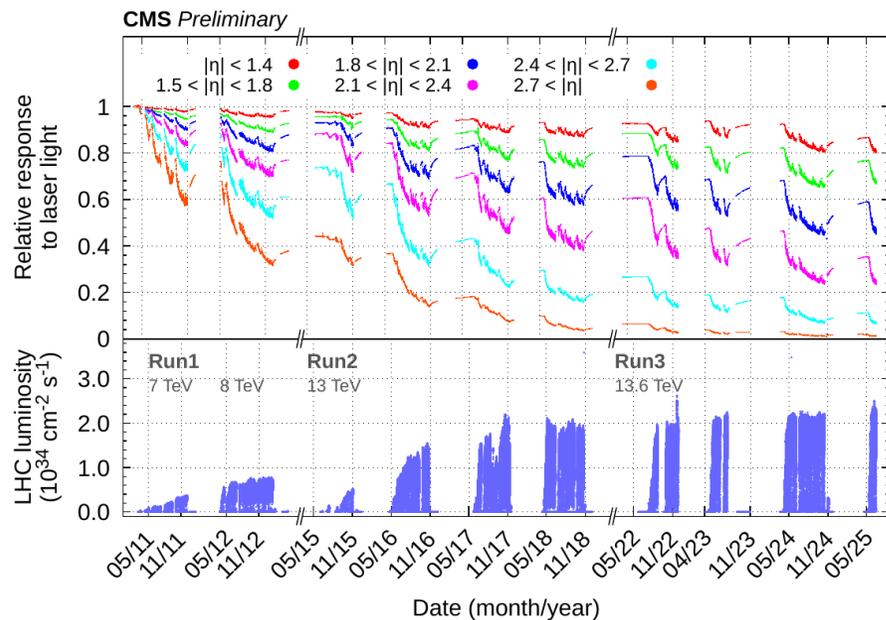
The Run 3 challenges

■ ■ ■ A record-breaking data taking

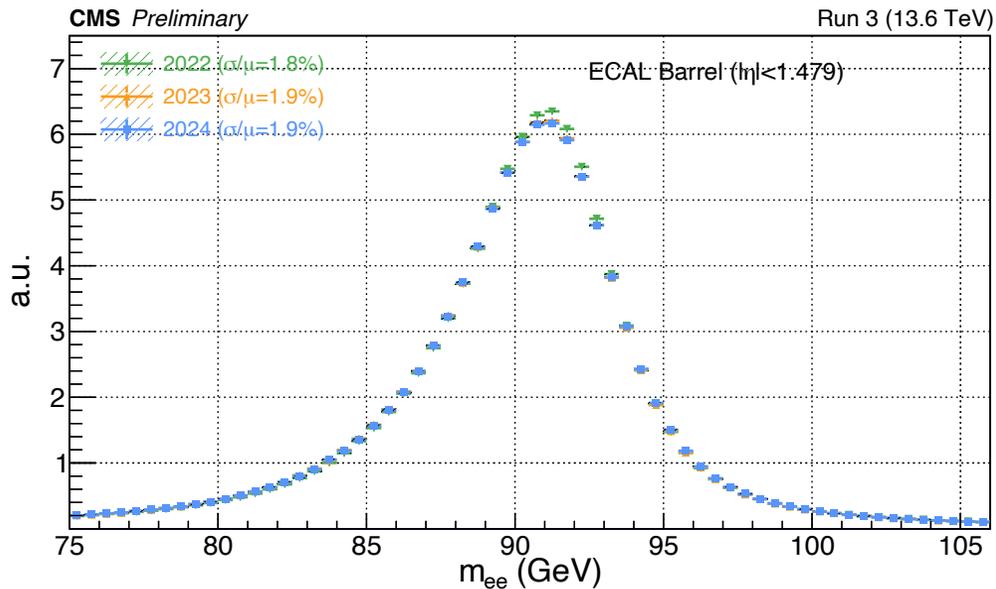
- Instantaneous luminosity 2×10^{34} (Run 2) \rightarrow $2.6 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ (Run 3)
- Center-of-mass energy **13** \rightarrow **13.6 TeV**
- Mean **pileup increasing** (46 in 2022, 52 in 2023, 57 in 2024)
- Lumi-leveling at **PU 60-65**, with hardware trigger rate up to **$\sim 115 \text{ kHz}$**
- Paramount to keep excellent detector performance (**energy and time resolution**) and operability (**limit data acquisition dead time**).

■ ■ ■ ECAL approach

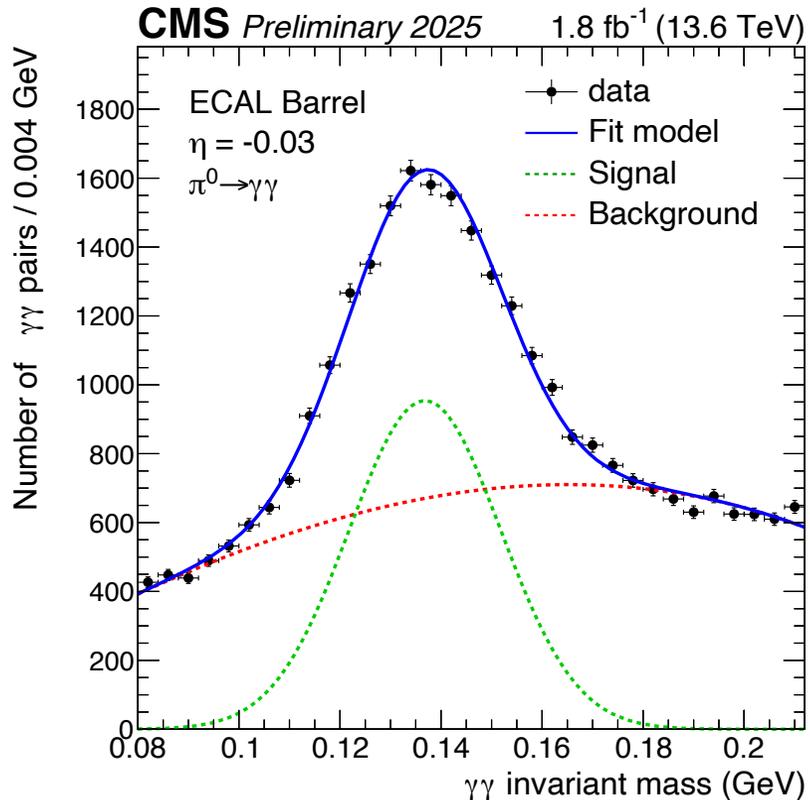
- Upgrade laser for transparency monitoring
- Automate several calibration procedures and increase their frequency
- Improve and speed up DAQ software procedures (e.g., SEU recovery)
- Develop new algorithms to improve timing and energy resolution



Intercalibration procedure



- Improve energy resolution by reducing the channel-to-channel response spread
 - Uniform response across the detector is obtained using different methods, including:
 - Invariant mass peak in $Z \rightarrow e^+e^-$ events
- stable over the 2022—2024 period, despite luminosity increase and detector ageing

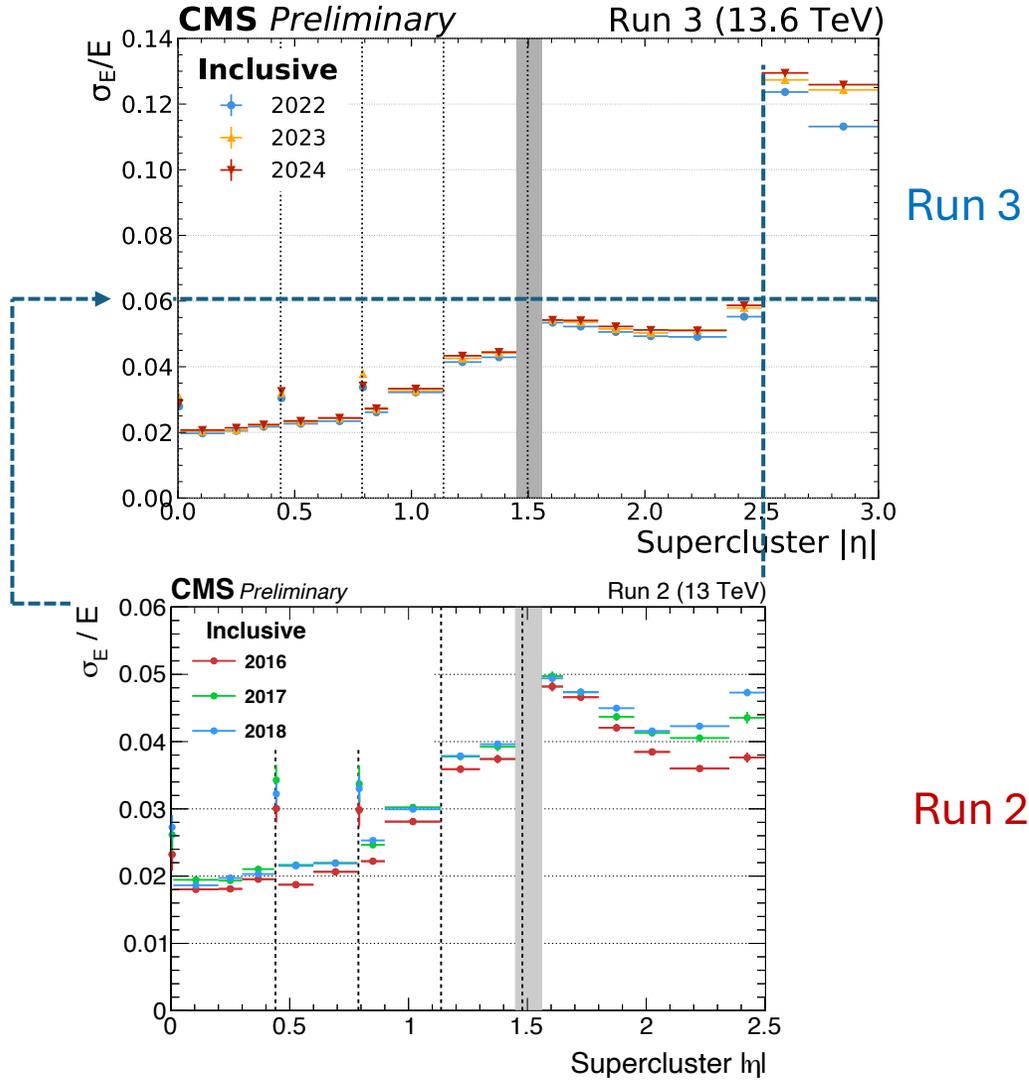


Intercalibration procedure

- Improve energy resolution by reducing the channel-to-channel response spread
- Uniform response across the detector is obtained using different methods, including:
 - Invariant mass peak in $Z \rightarrow e^+e^-$ events
→ stable over the 2022—2024 period, despite luminosity increase and detector ageing
 - Invariant mass peak in $\pi^0 \rightarrow \gamma\gamma$ events
→ large dataset thanks to dedicated high-rate, small-event content data stream, but suffers from low- p_T photon background

Thanks to the large datasets collected, intercalibration constants are updated every 25 fb⁻¹ in Run 3.

A peek at energy resolution



Relative electron energy resolution

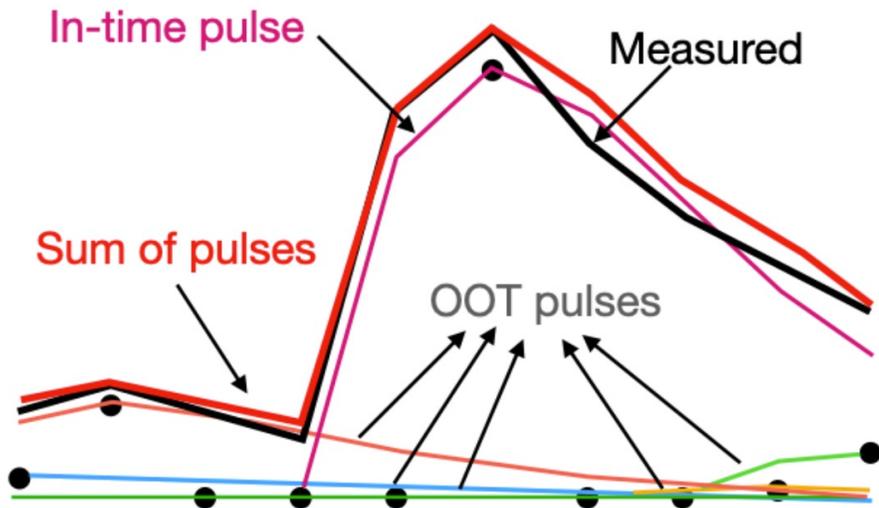
- Estimated using $Z \rightarrow e^+e^-$ events in several η intervals
- Essentially stable over the period 2022—2024
- Some loss wrt Run 2, especially in the endcaps, due to detector ageing, higher pileup, and increase in reconstruction thresholds
 - Partially recovered using energy regression
 - Between 2 and 6% within the tracker pseudorapidity coverage ($|\eta| < 2.5$)

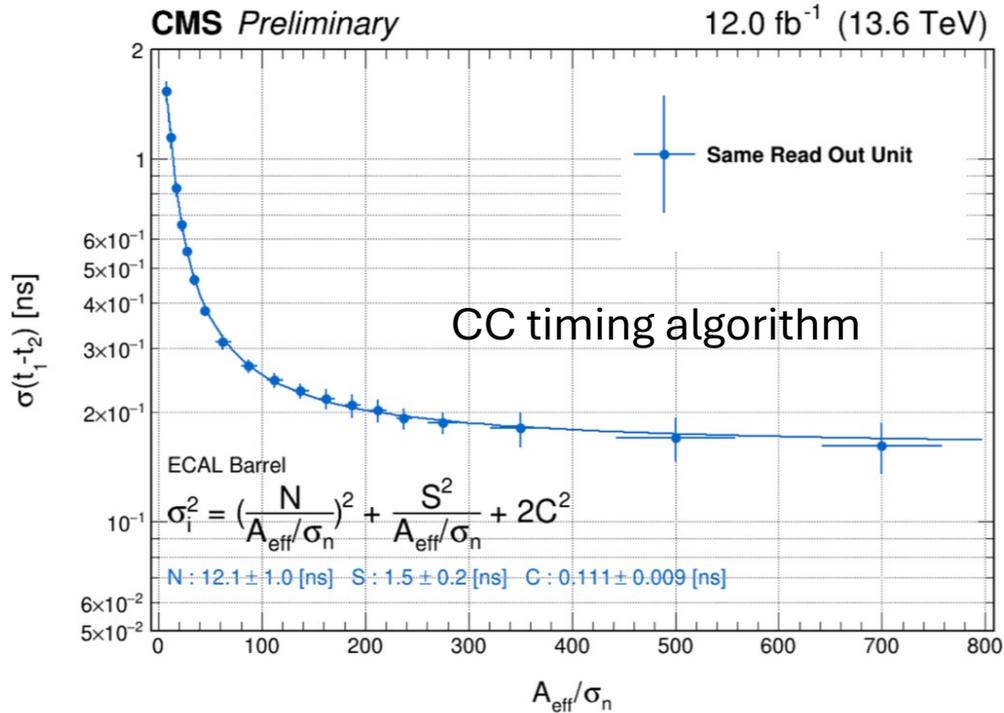
Ratio timing

- Time of a pulse measured in a crystal (wrt the LHC clock) is derived from the ratio of consecutive amplitudes $A(t)$ and $A(t+25 \text{ ns})$
- Default time reconstruction algorithm in Run 2 + 2022/23
- Still used in online reconstruction

Cross correlation (CC)

- New algorithm developed to cope with high pileup
- Uses templated pulse shapes to subtract out-of-time (OOT) pulses and extract the in-time one through cross-correlation fit
- Rejects non-scintillation signals and is OOT pileup-aware





Ratio timing

- Time of a pulse measured in a crystal (wrt the LHC clock) is derived from the ratio of consecutive amplitudes $A(t)$ and $A(t+25 \text{ ns})$
- Default time reconstruction algorithm in Run 2 + 2022/23
- Still used in online reconstruction

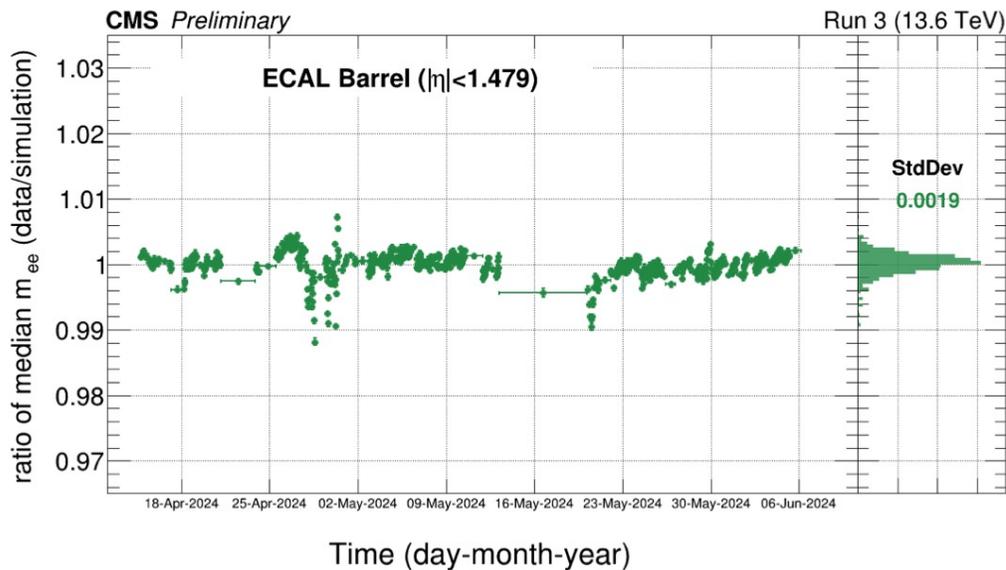
Cross correlation (CC)

- New algorithm developed to cope with high pileup
 - Uses templated pulse shapes to subtract out-of-time (OOT) pulses and extract the in-time one through cross-correlation fit
 - Rejects non-scintillation signals and is OOT pileup-aware
- **Significantly improves time resolution** wrt ratio method (expected $\sim 100\text{-}150 \text{ ps}$ in Run 3 vs 200 ps in Run 2 for e/γ objects)

Automation of calibrations



Monitoring of the calibration and reconstruction jobs



Automation framework operational since 2022

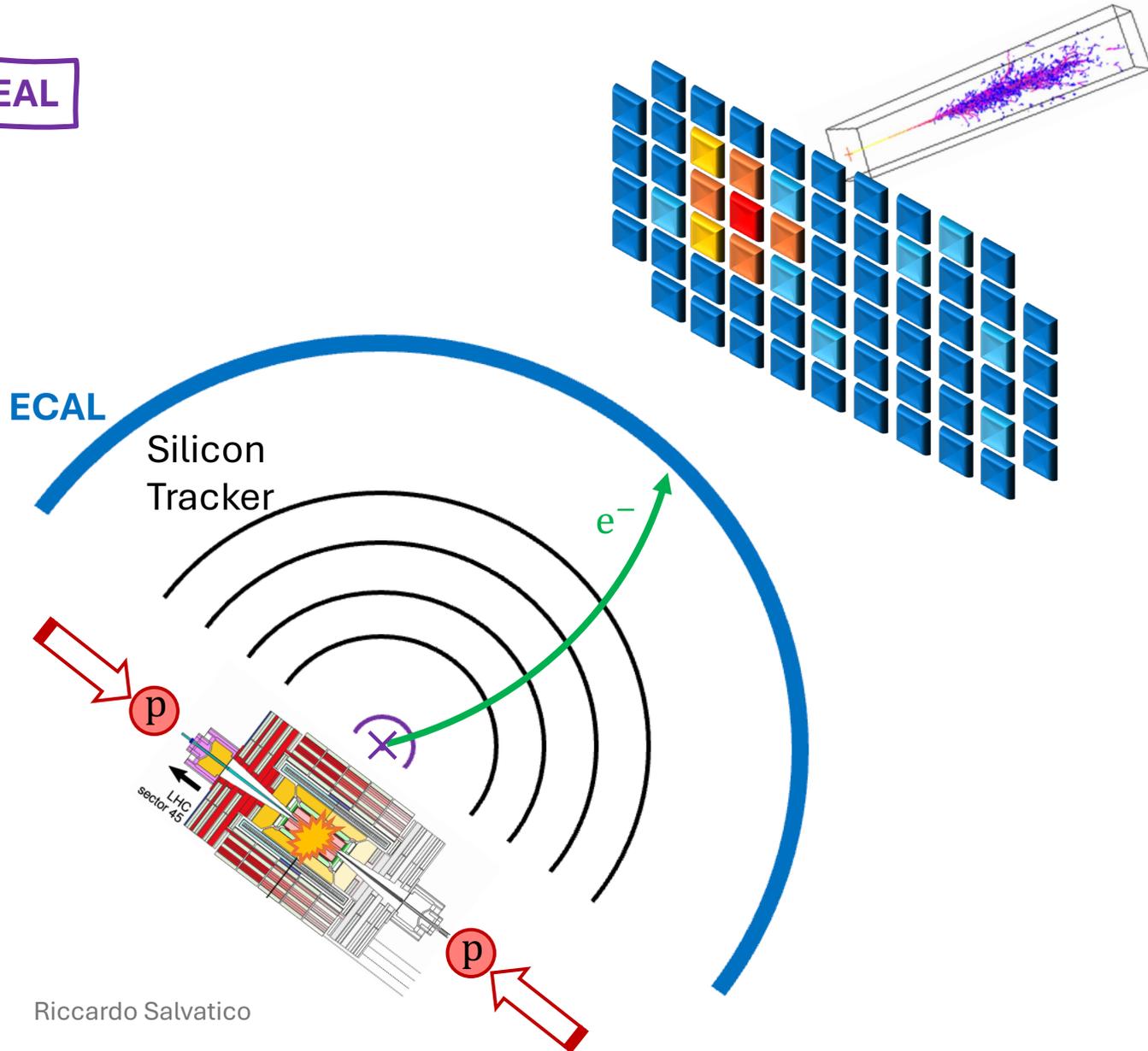
- Deliver **frequent and up-to-date calibrations online**, optimizing the necessary person-power
- Pre-determined input conditions and dataset types to derive:
 - alignment coefficients
 - pulse shape templates
 - timing calibrations
 - laser harness corrections
 - energy scales and intercalibrations
 - ϕ -symmetry reconstruction
 - m_{ee} and π^0 monitoring

Example: time stability of m_{ee}

- Comparing data and Run 3 simulation using $Z \rightarrow e^+e^-$
- Spread of the median ratio $\sim 0.2\%$ in 2022—2024

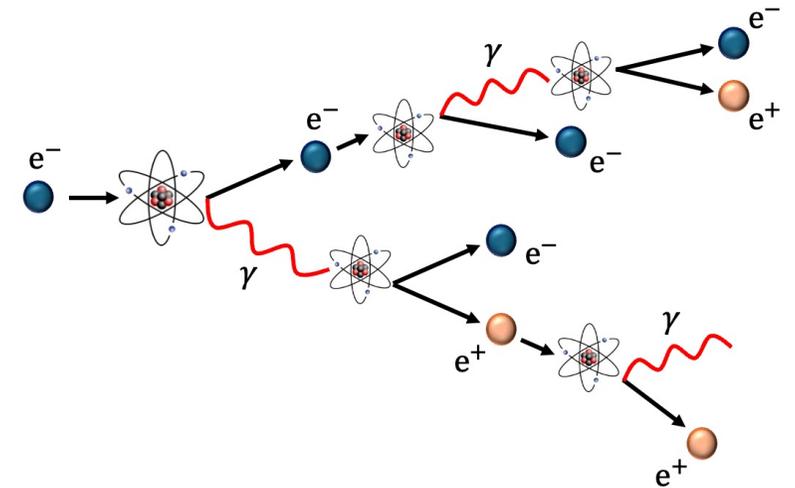
Electrons and photons at CMS

IDEAL



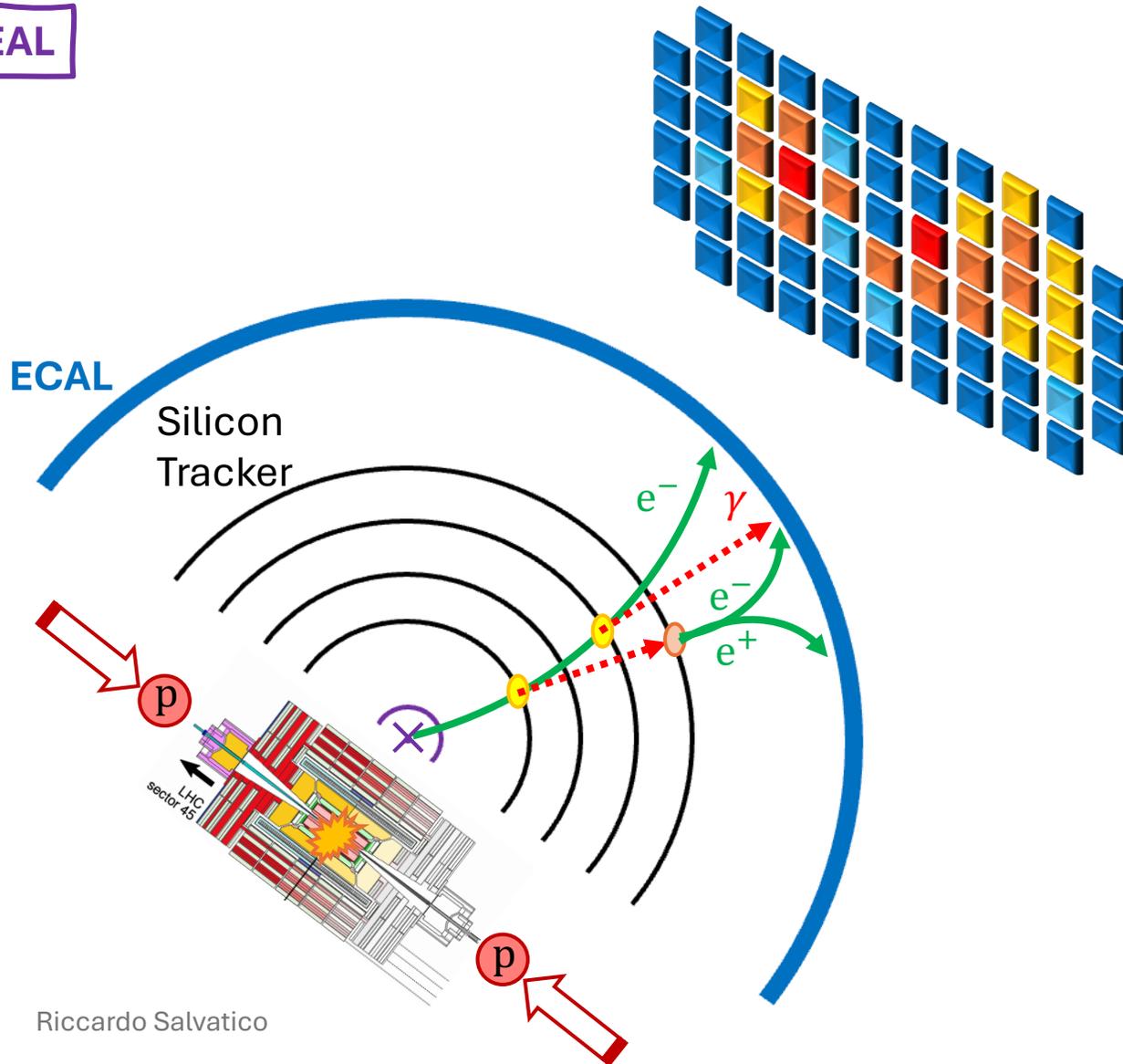
A hit in the ECAL

→ an electromagnetic shower in a small **cluster** of PbWO_4 crystals



Electrons and photons at CMS

REAL



Bremsstrahlung and photon conversions need to be associated to the same initial particle

→ **Supercluster**

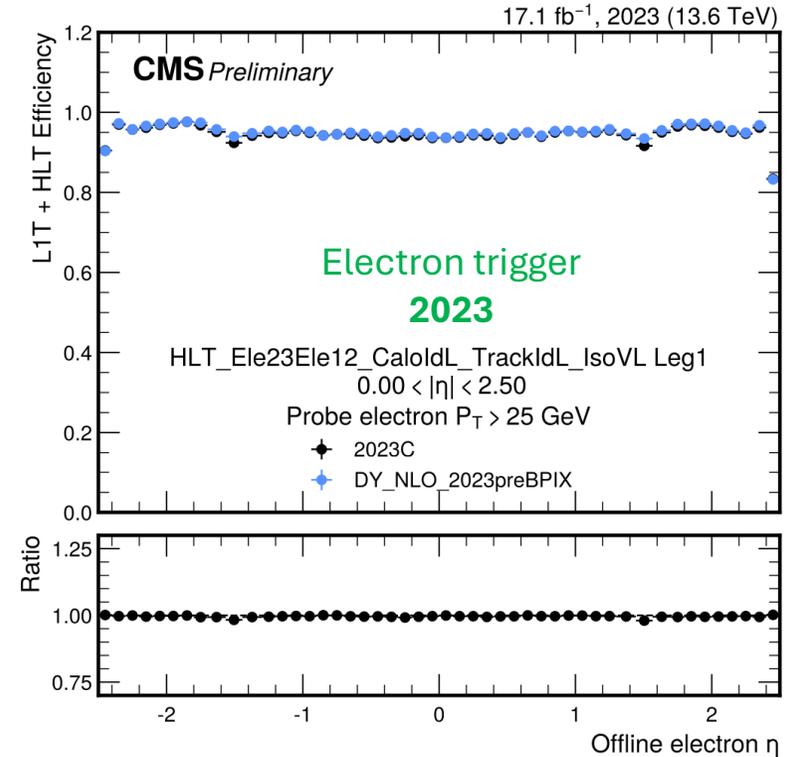
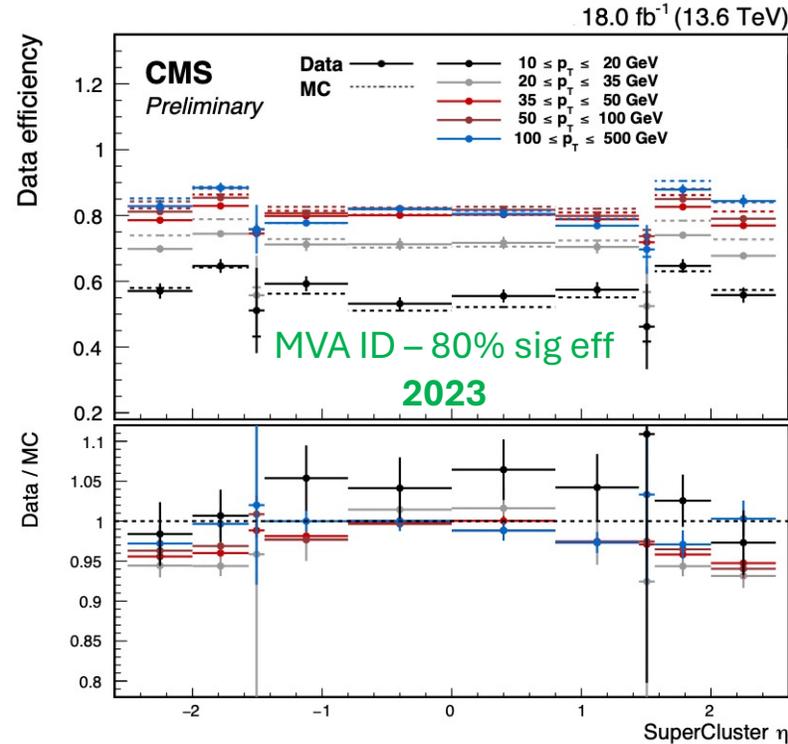
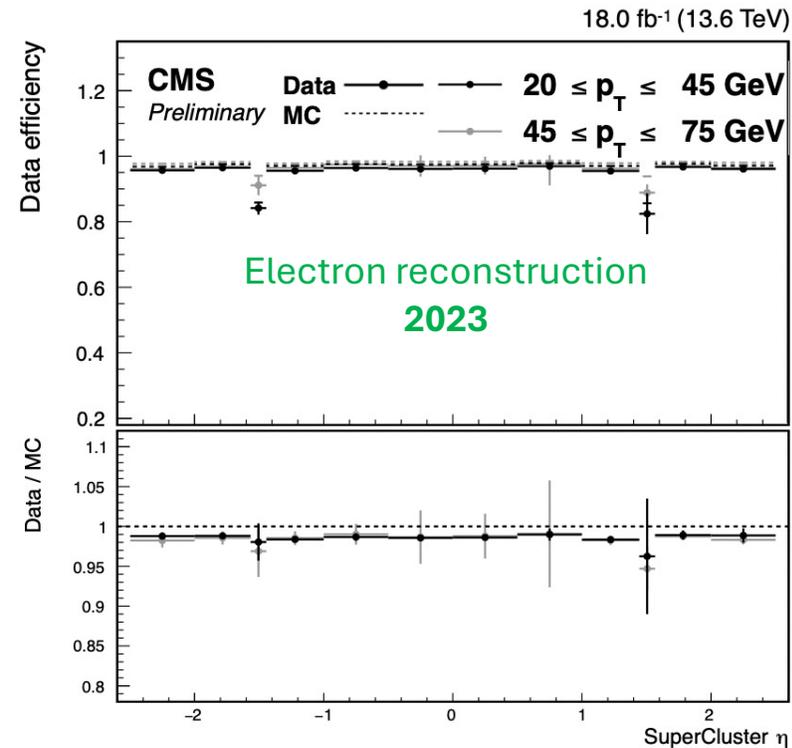
Energy deposit in the ECAL + track association

→ **electron/positron**

Using **special track reconstruction algorithm** (Gaussian-sum filter) to consider energy losses via photon radiation.

Performance and corrections

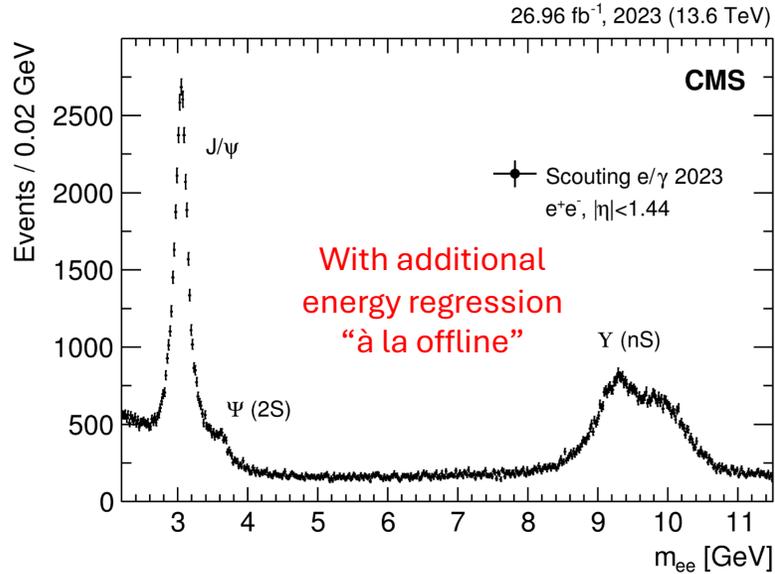
The reconstruction, identification, and trigger performance for electrons and photons is studied in collision data and simulation, and we correct them for discrepancies.



Efficiency to match a supercluster with a track in the tracker

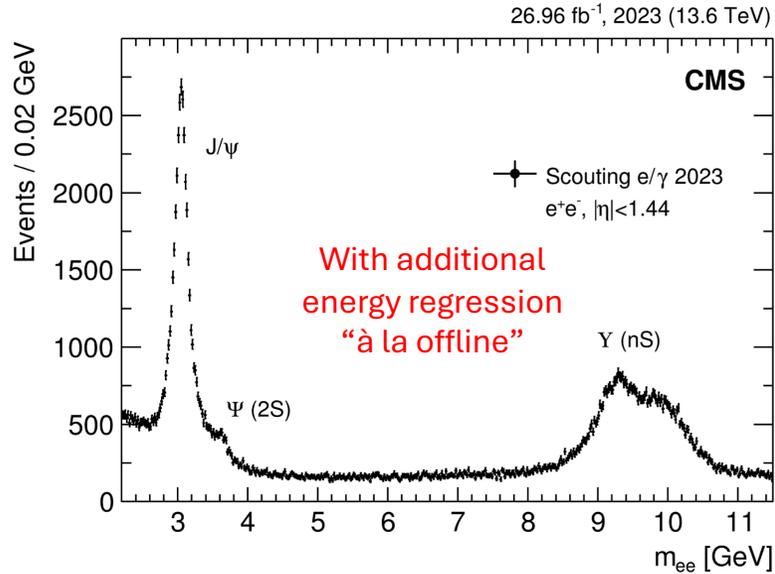
Efficiency to satisfy a set of identification criteria

Efficiency for offline-reconstructed objects to pass a certain trigger



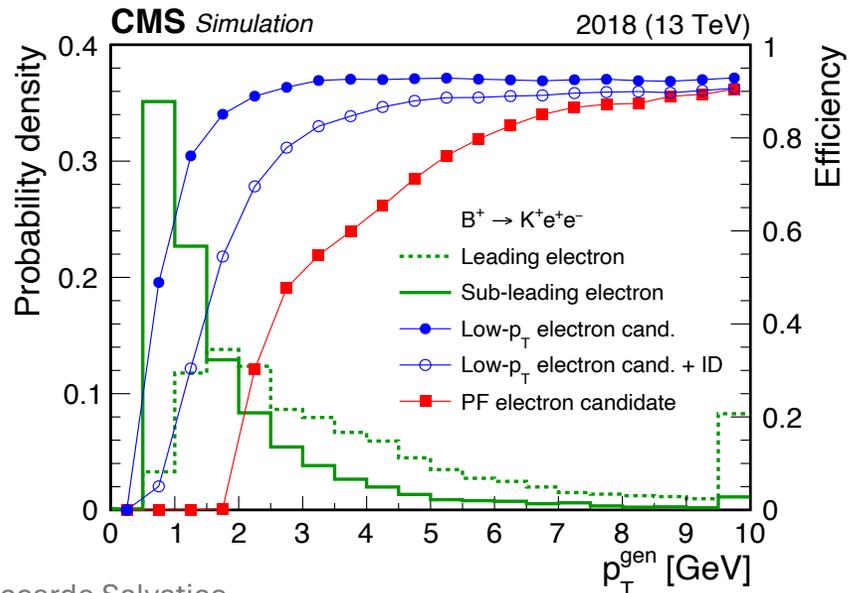
High-level trigger (HLT) scouting (*online*)

- HLT reconstruction with reduced event content; no offline reconstruction.
- Considerably reduce p_T thresholds wrt standard HLT, especially for single photon (200 GeV in 2022 and 2023).
- Minimal HLT selection (on supercluster) and thresholds:
 - 30 GeV for single e/γ → ~ 9.0 kHz HLT (2023)
 - 12 GeV for double e/γ → ~ 0.5 kHz HLT (2023)



High-level trigger (HLT) scouting (*online*)

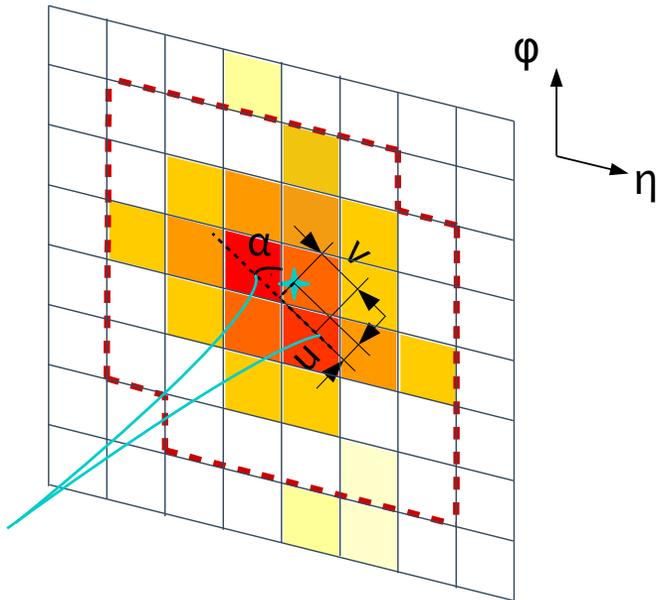
- HLT reconstruction with reduced event content; no offline reconstruction.
- Considerably reduce p_T thresholds wrt standard HLT, especially for single photon (200 GeV in 2022 and 2023).
- Minimal HLT selection (on supercluster) and thresholds:
 - 30 GeV for single e/γ → ~9.0 kHz HLT (2023)
 - 12 GeV for double e/γ → ~0.5 kHz HLT (2023)



Low- p_T electron collection (*offline*)

- A special, track-oriented reconstruction algorithm using BDTs to identify suitable track seeds before building full GSF tracks;
- Dedicated ID to suppress large background;
- Higher reconstruction efficiency wrt standard electrons within [1, 10] GeV;
- Users: B-physics, compressed SUSY, HZZ.

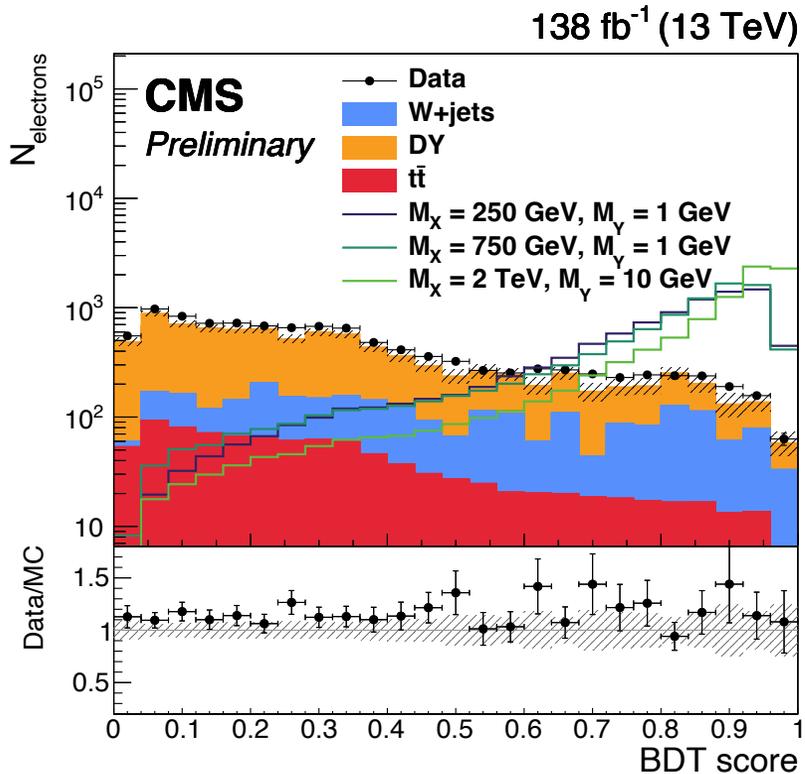
Merged di-electron ID



Targeting pairs of boosted electrons with arbitrary m_{ee}

- If $\Delta R(ee) \lesssim 0.1$ ($\sim 6 R_M$ in EB), their energy deposits may overlap and shift the SC's center-of-gravity, leading to SC-track matching inefficiencies
 - 1 reconstructed electron, 1 unmatched track
- If $\Delta R(ee) \approx O(10^{-3})$, the electron tracks can share hits and either of them may be removed by a trajectory cleaner algorithm
 - 1 reconstructed electron, 0 unmatched tracks

Merged di-electron ID

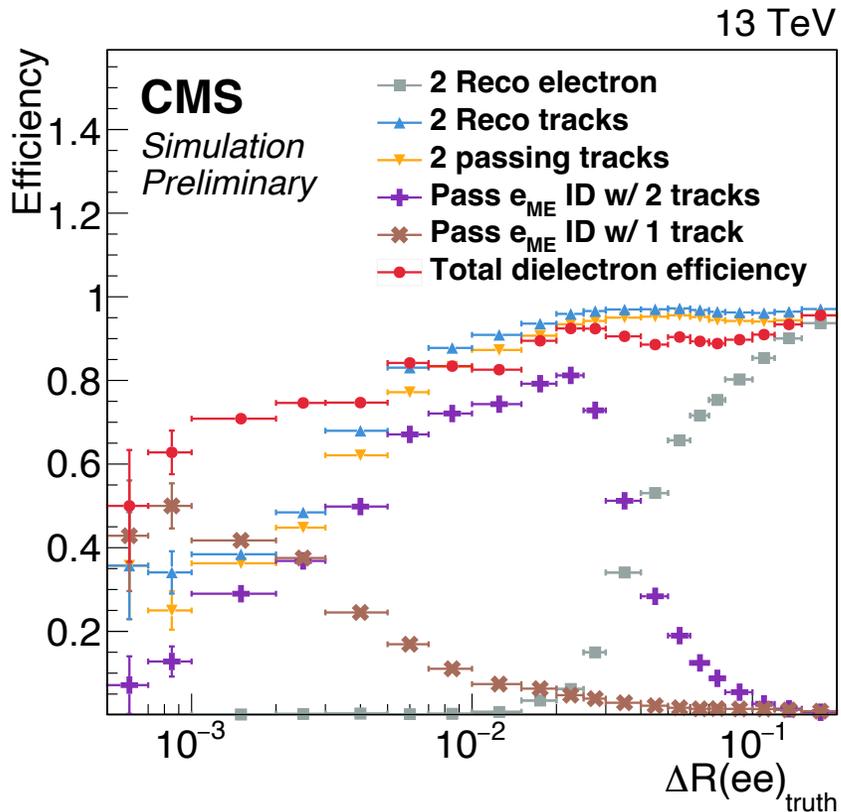


Targeting pairs of boosted electrons with arbitrary m_{ee}

Signatures with merged clusters, with or without a secondary track

1. Take standard electron candidates (SC + GSF track) as input
2. Train BDT using neighboring GSF/KF tracks and novel SC-track matching variables that are less sensitive to the SC's center-of-gravity shift

Merged di-electron ID



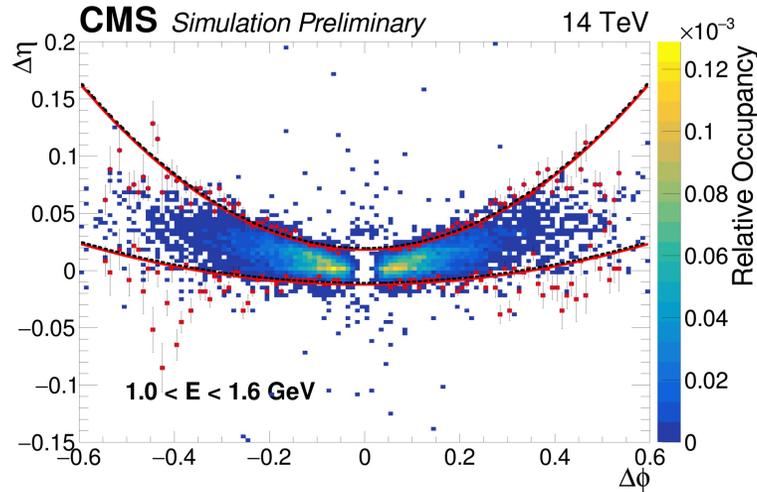
Targeting pairs of boosted electrons with arbitrary m_{ee}

Signatures with merged clusters, with or without a secondary track

Enhance electron reconstruction efficiency at small $\Delta R(ee)$ and facilitate searches such as high-mass resonances in leptonic final states ([CMS-PAS-EXO-24-006](#)).

Redesigning the supercluster

Base component in the reconstruction of photons and electrons, essential for the ECAL energy response calibration, and input to the [Particle Flow](#) global event reconstruction.

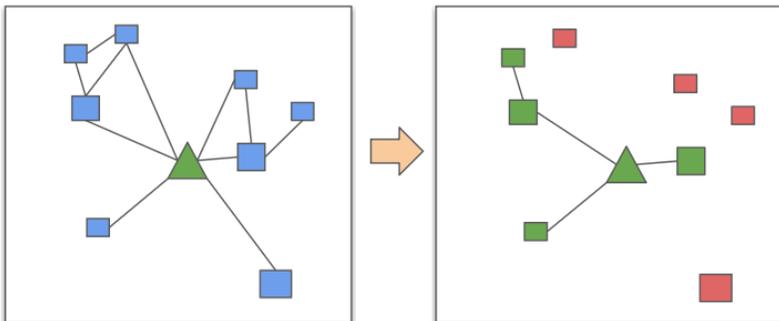


■ ■ ■ Moustache

- selects clusters defining a parabolic $\eta - \phi$ region parametrized by the seed position and the cluster transverse energy
- very efficient but subject to PU/noise contamination – resolution degrades with PU

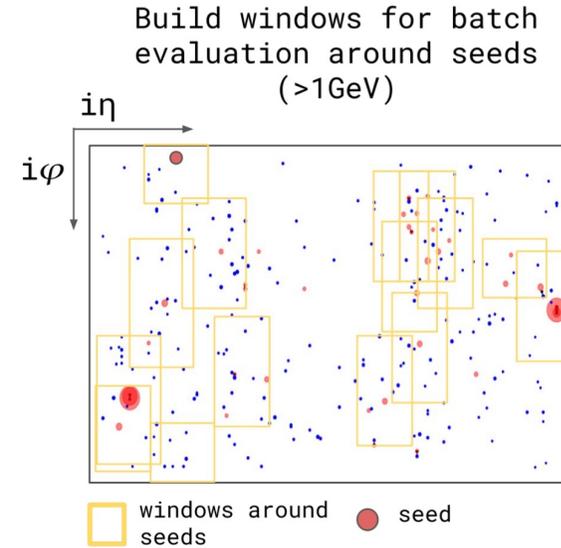
■ ■ ■ GNN supercluster (“DeepSC”)

- uses small clusters and single crystal hits in a window around the seed
- filters noise/PU *on a cluster-by-cluster basis*, improving the “raw” resolution



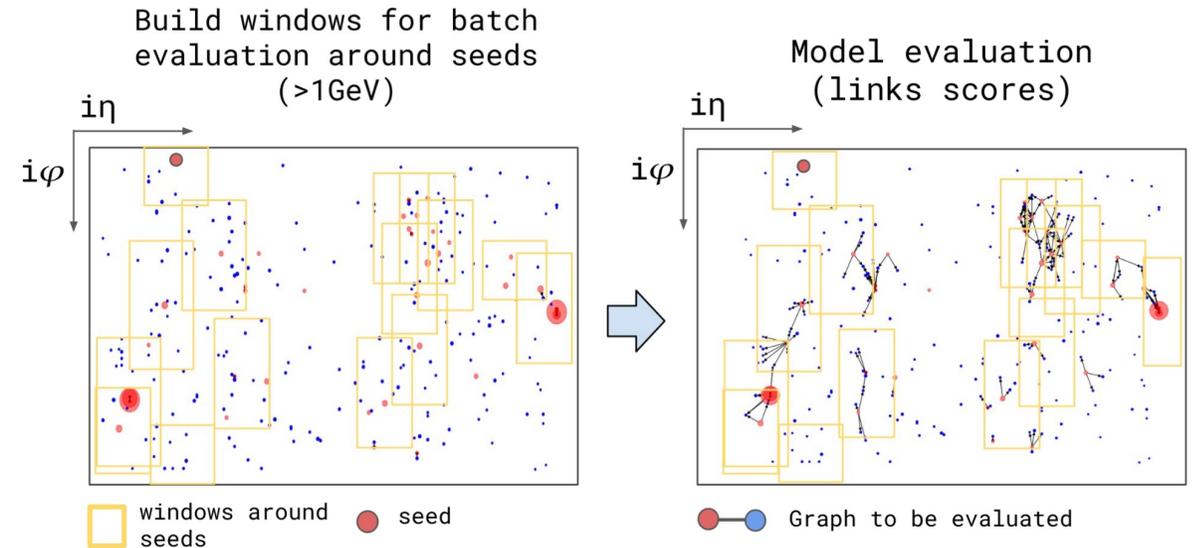
The “DeepSC”

1. Build detector windows around each cluster with $p_T > 1 \text{ GeV}$ (**seed**).



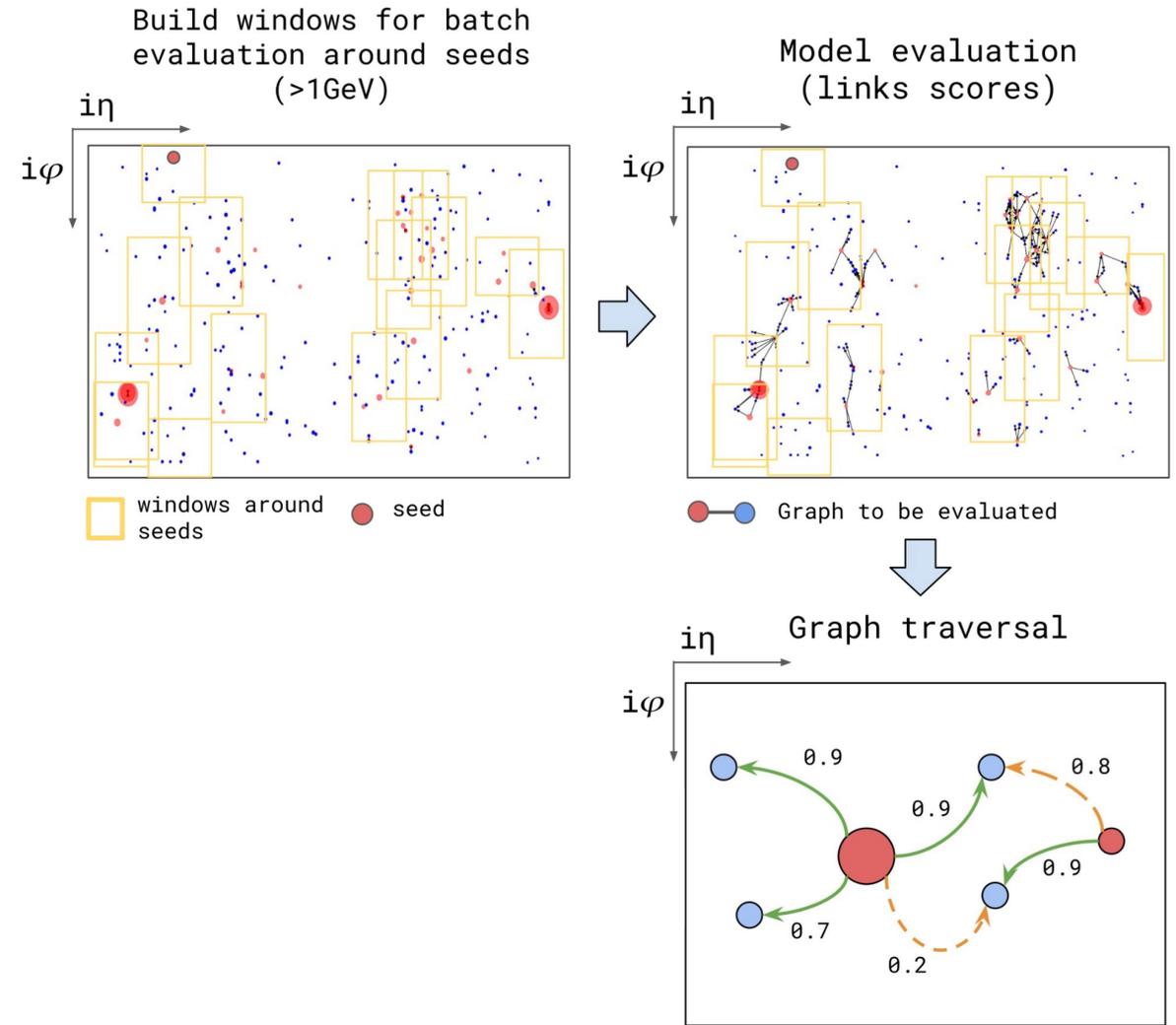
The “DeepSC”

1. Build detector windows around each cluster with $p_T > 1 \text{ GeV}$ (**seed**).
2. Connect the seed to all the clusters in each window to form a graph. Use basic cluster properties to evaluate a [TensorFlow](#) model on groups of detector windows: the output is the **probability** of each edge between the seed and the nearby clusters.



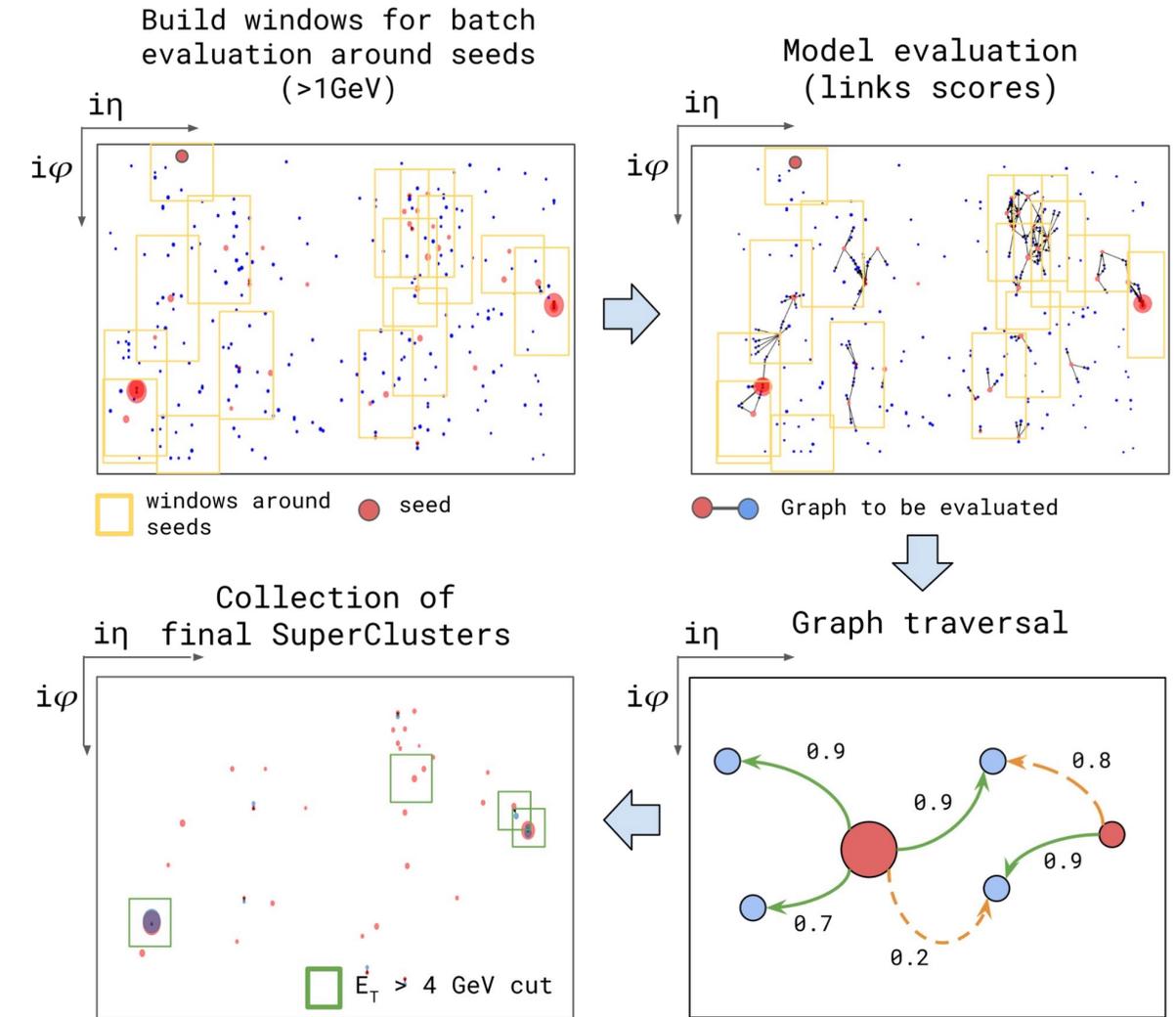
The “DeepSC”

1. Build detector windows around each cluster with $p_T > 1$ GeV (**seed**).
2. Connect the seed to all the clusters in each window to form a graph. Use basic cluster properties to evaluate a [TensorFlow](#) model on groups of detector windows: the output is the **probability** of each edge between the seed and the nearby clusters.
3. Traverse the graph starting from the highest energy seed to choose how to **assign clusters** that can potentially belong to several superclusters.



The “DeepSC”

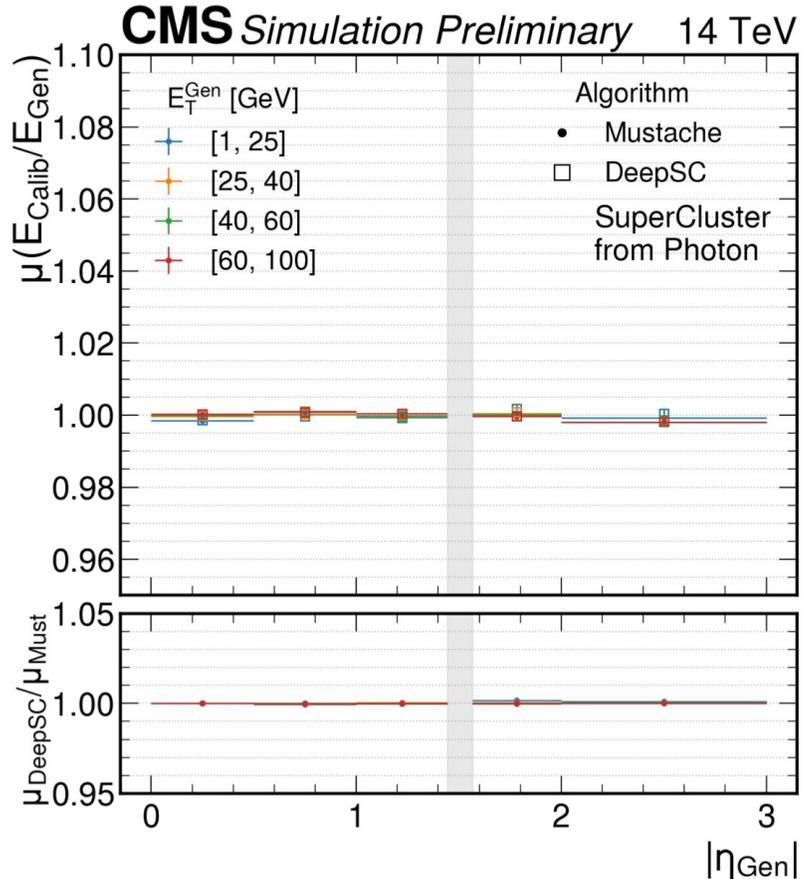
1. Build detector windows around each cluster with $p_T > 1$ GeV (**seed**).
2. Connect the seed to all the clusters in each window to form a graph. Use basic cluster properties to evaluate a [TensorFlow](#) model on groups of detector windows: the output is the **probability** of each edge between the seed and the nearby clusters.
3. Traverse the graph starting from the highest energy seed to choose how to **assign clusters** that can potentially belong to several superclusters.
4. **Apply a cut** of $p_T > 4$ GeV to remove supercluster formed mainly by detector noise and very low energy pileup.



A promising way forward

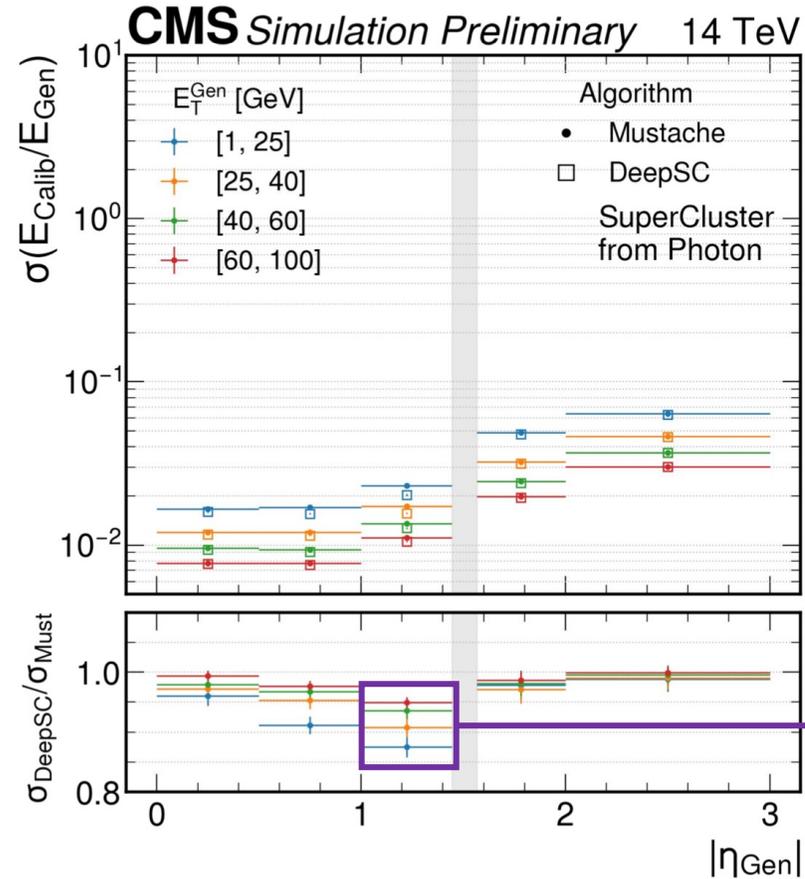
Energy scale

comparable to the moustache



Energy resolution

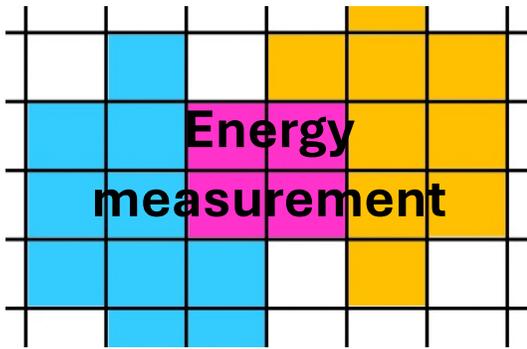
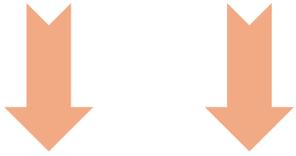
improvements observed across the detector, in particular in the most challenging regions



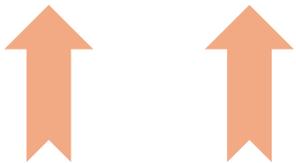
Region with the largest tracker material budget in front of the ECAL
 → largest secondary emissions

SC, electron, and photon energy corrections

Energy loss in the tracker
Leakage into the HCAL

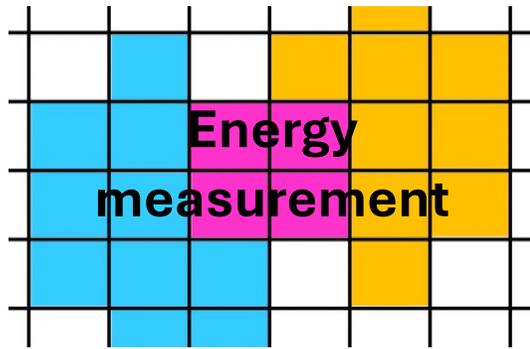
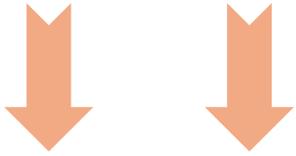


Clustering inefficiency
Finite noise thresholds



SC, electron, and photon energy corrections

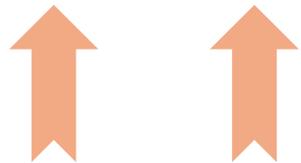
Energy loss in the tracker
Leakage into the HCAL



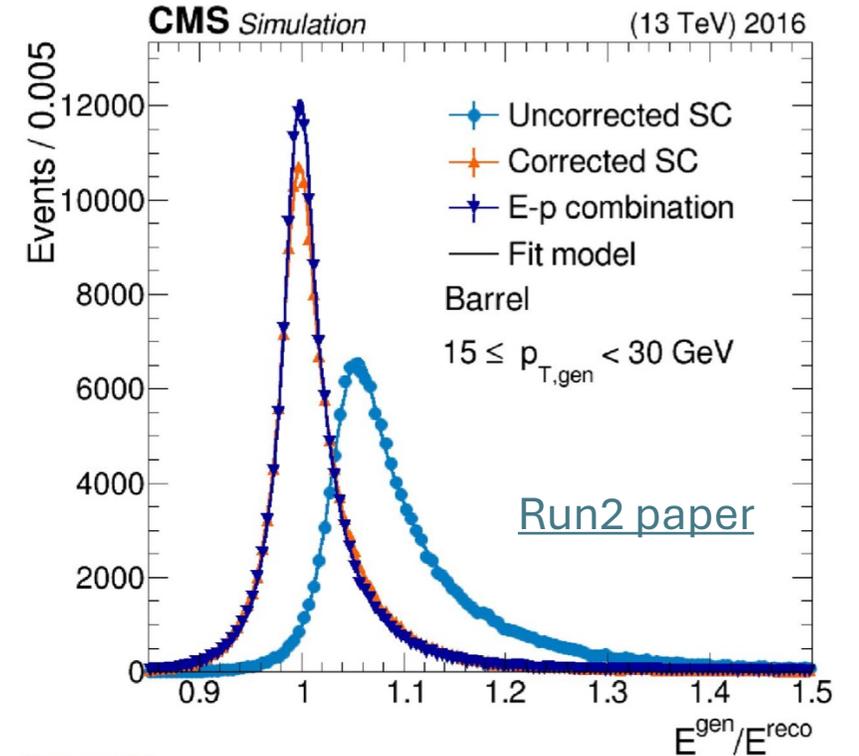
BDT-based **semi-parametric regression**
using ~ 30 high-level input variables



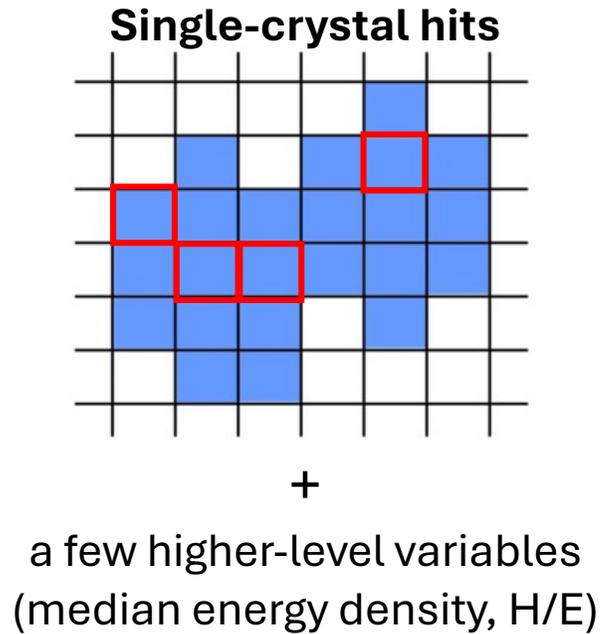
Provide per-object correction
(supercluster, electron, photon)



Clustering inefficiency
Finite noise thresholds



DRN energy regression

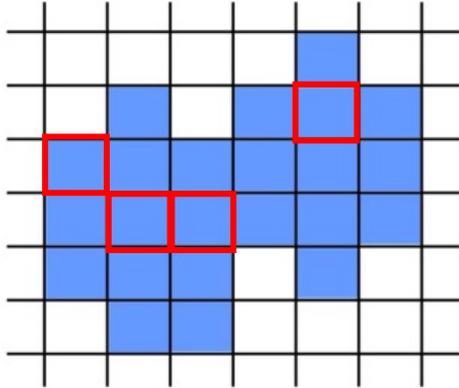


Dynamic Reduction Network (GNN)

- Allows for variable size input
- Can deal with unordered inputs
- Can handle complicated geometries
- Can handle representation in 4D (x, y, z, energy)

DRN energy regression

Single-crystal hits



+

a few higher-level variables
(median energy density, H/E)

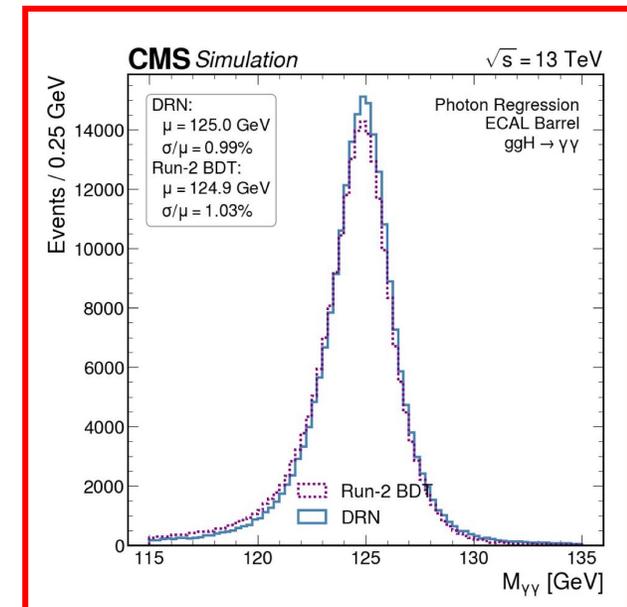
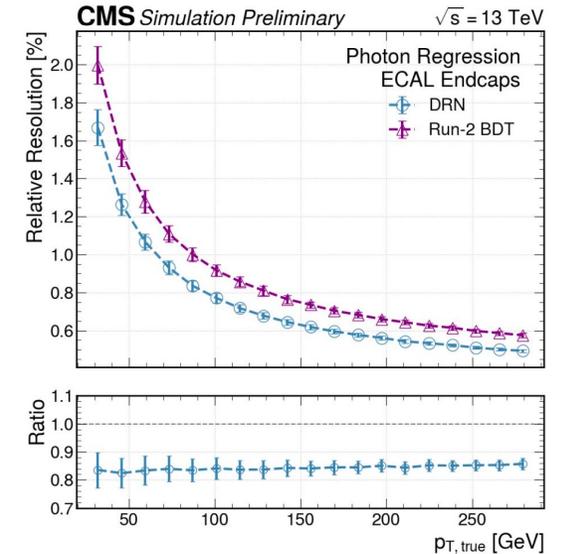
1. Hits are mapped to a multidimensional space



2. GNN learns high-level characteristics from the hits by grouping them based on their similarities

3. A further NN maps the learned features to the regression target

~5% improvement in the $M_{\gamma\gamma}$ resolution in the barrel region!



Summary

- Remarkable performance of the CMS ECAL in Run 3
 - ✓ Capable of coping with the highest instantaneous luminosity, pileup, and trigger rate in the CMS history
- New algorithms and frequent calibrations allowing us to match or improve the Run 2 performance
 - ✓ Substantial help from automation framework
- Excellent electron and photon reconstruction performance – offline and online
 - ✓ Calibrations and corrections are adopted to ensure the best precision
- Expanding the reconstruction and identification capabilities towards challenging phase spaces
 - ✓ Low pT: high trigger rate thanks to scouting, high reconstruction efficiency thanks to dedicated algorithms
 - ✓ Boosted regimes
- Ongoing studies entailing novel techniques are paving the way to more efficient SC, electron, and photon reconstruction and more precise energy measurements

BACKUP

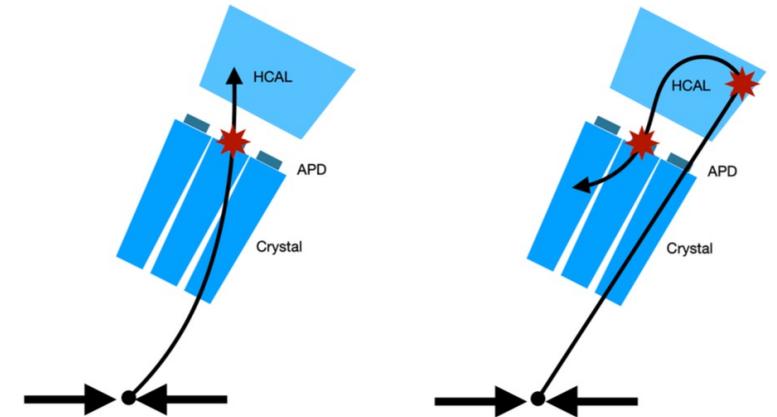
Spikes and the “spike killer”

Large apparent energy pulses (“**spikes**”) are induced by direct hadronic ionization of APDs

ECAL uses a combination of **online** and **offline** spike killing to reduce the impact of spikes

Online:

- **Strip Fine Grain Veto Bit (sFGVB)** exploits the spike deposit shape in a single TT
- If the readout tower energy is above threshold and sFGVB: no trigger



Offline:

- **Swiss Cross** variable: $1 - E_4/E_1$
- Spikes identified through anomalous energy sharing between crystals

