The CMS electromagnetic calorimeter

Rajdeep M Chatterjee , TIFR Mumbai
on behalf of the CMS Collaboration

ICNFP2024: XIII International Conference on New Frontiers in Physics

26 Aug - 4 Sep, 2024, OAC, Kolymbari, Crete, Greece

Outline

- The Role of the CMS ECAL in precision measurements and innovative searches.

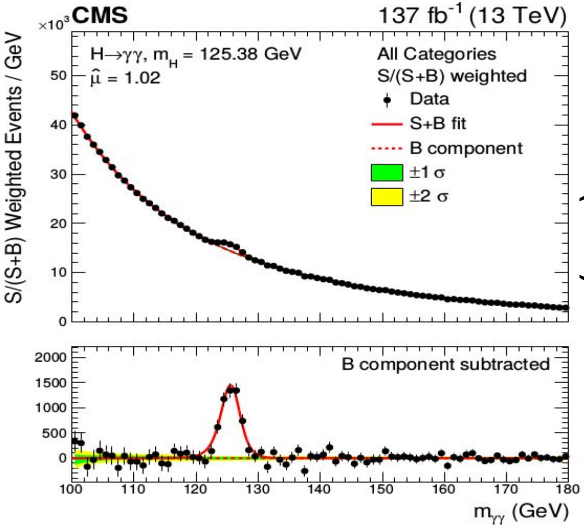
- Enhancing the precision of ECAL reconstruction
  - A GNN based strategy for ECAL superclustering
  - A GNN based strategy to optimize the energy resolution of electromagnetic objects
  - A methodology to correct the photon energy scale due to radiation damage of ECAL crystals

- An end-end regression strategy to reconstruct merged photons with the CMS ECAL

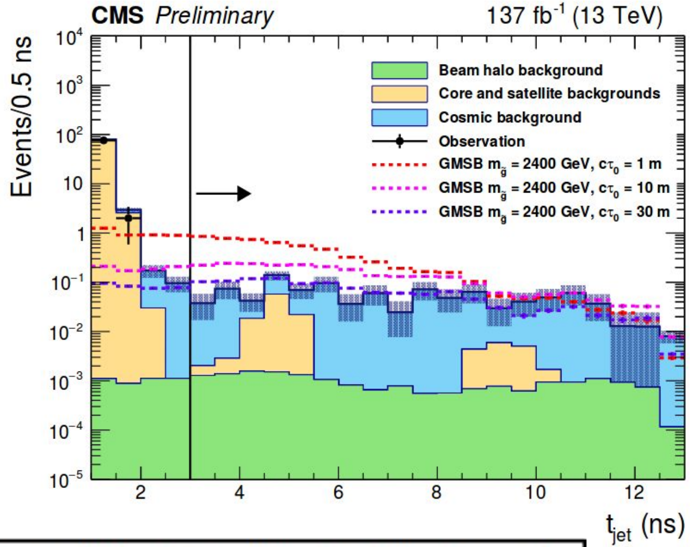
- An anomaly detection based ML strategy for online data quality monitoring with the CMS ECAL

Role of the CMS ECAL in physics measurements and searches

The discovery of the Higgs boson in the diphoton decay channel, followed by the measurement of its properties is one of the key drivers of the CMS ECAL design



JHEP 07 (2021) 027



PLB 797 (2019) 134876

- **Excellent energy resolution ([1.5-4]%)** and electron/photon ID played a crucial role in the discovery and the precise characterization of the Higgs Boson.

- **Good timing resolution (~200 ps)** key ingredient in searches for non conventional signatures like delayed jets.

EM Objects Energy Measurement & Calibration

FROM RAW DATA TO CALIBRATED ENERGIES

$$E_{e,\gamma} = \boxed{F_{e,\gamma}} \boxed{G(\eta, t) \sum_i c_i\,(t)\, s_i(t)\, A_i\,(t)}$$

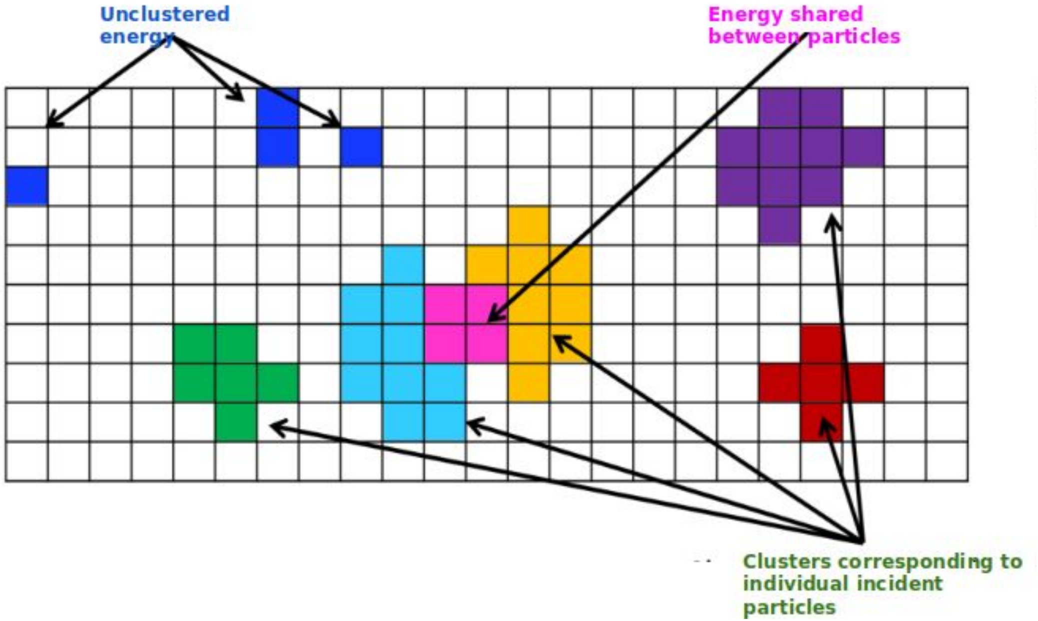
Different time granularity

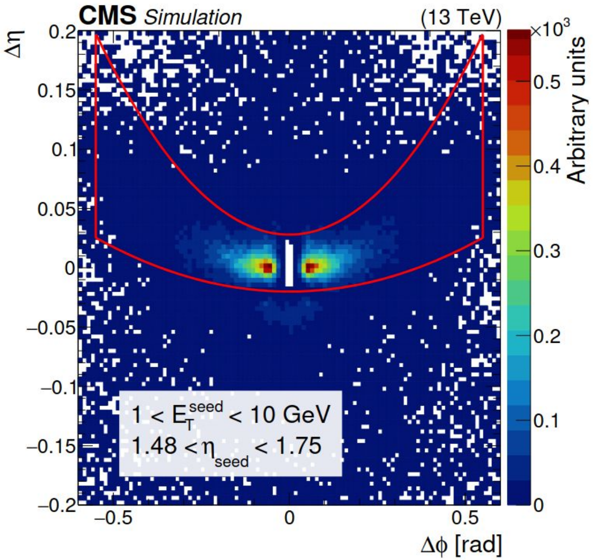
In this talk I will cover EM object level correction to ensure an optimal energy response across the detector.

Covered in the previous talk, these ingredients ensure that each crystal contributing to an EM cluster is calibrated.

ECAL clustering

The CMS ECAL clustering algorithm



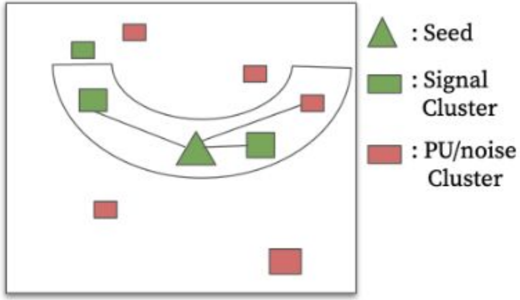


JINST 16(2021) P05014

- Start with pattern of detector hits
- Associate each incident particle with a collection of hits using a clustering algorithm
  - Hits are gathered together around the crystal with highest energy to form clusters
  - In the case of EM particles Bremsstrahlung and photon conversion before the ECAL, clusters have to be combined together to form a SuperCluster (SC)

Improving the ECAL SC algorithm

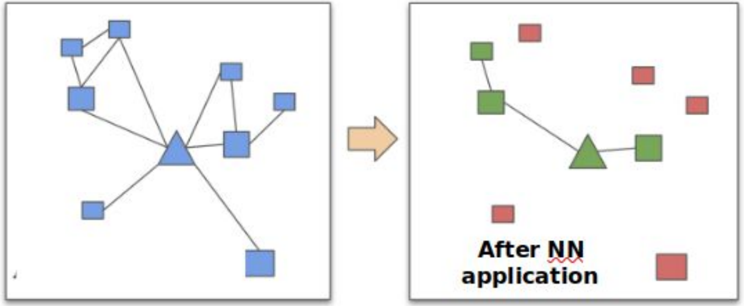
Mustache SC (currently used)



Purely geometrical algorithm

- Evaluated on one cluster at a time.
- Large efficiency for signal, but also large pileup and noise contamination.

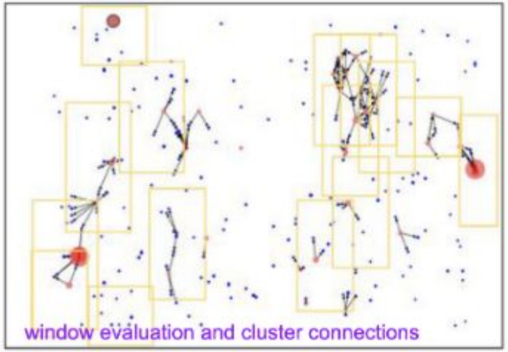
GNN based DeepSC (proposed)



Based on Graph Convolutional Networks (GCN)

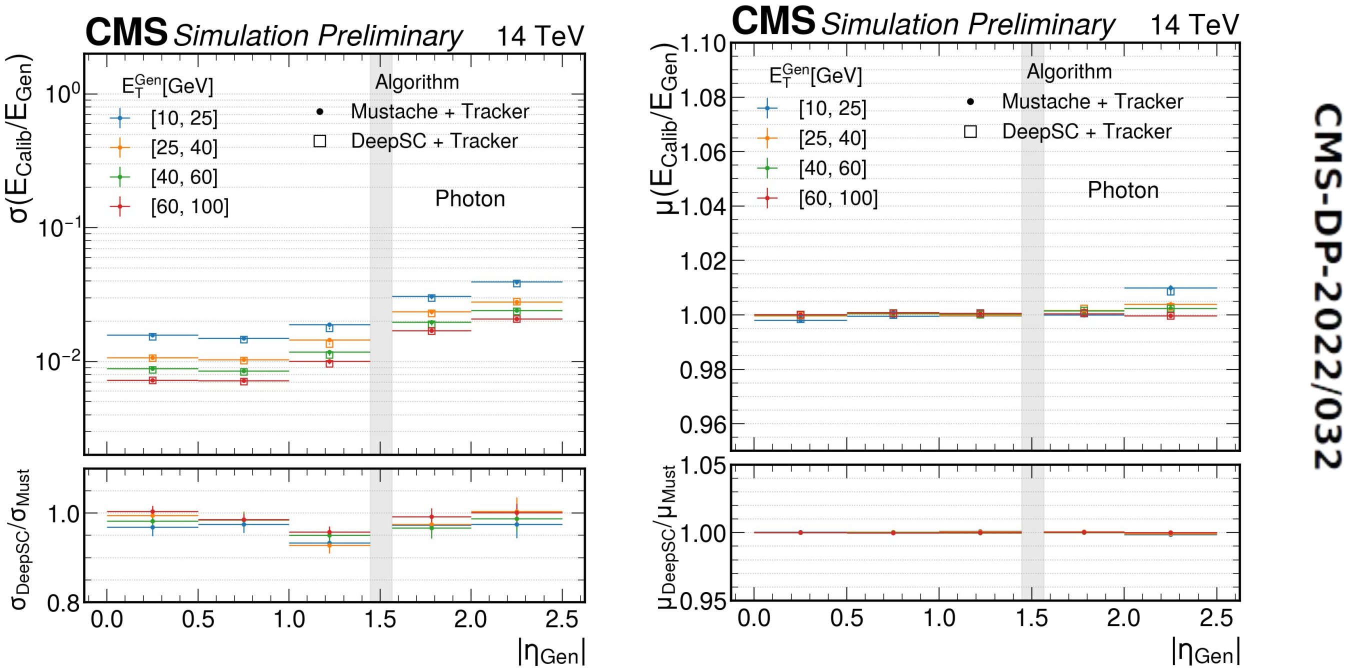
- Evaluated on all the clusters together in a region around the seed.
- Noise an pileup are filtered for each cluster :
  → For the same signal efficiency as the current algorithm provides clusters with higher purity.

GNN based DeepSC algorithm for the ECAL



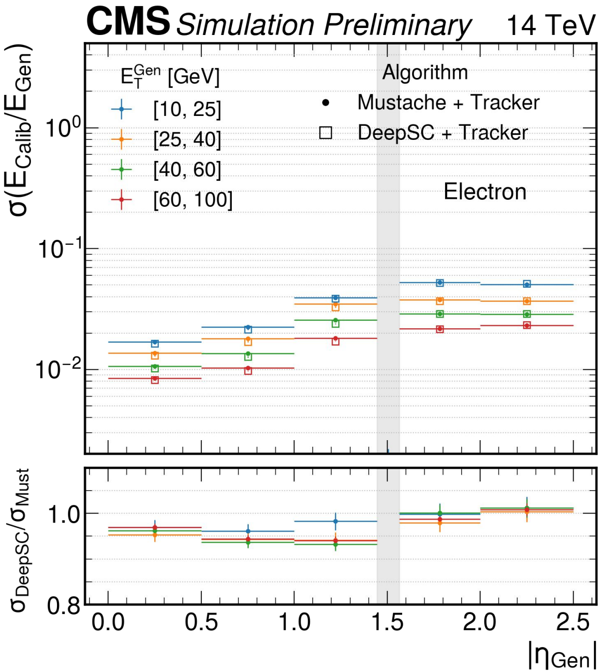
- Windows of variable dimensions along η are opened around all clusters with $E_T > 1$ GeV (seeds)
- **Input features**
  -- Cluster information : E, E_T, η, φ, z, #crystals; relative to the seed : ΔE, ΔE_T, Δη, Δφ
  -- List of Rechits for each cluster
  -- Summary of window features : min, max , mean of the crystal variables (E, E_T, ΔE, ΔE_T, Δη, Δφ)
- **Outputs :** Cluster classification (whether in/out of SC), window classification (electron/photon/jet)

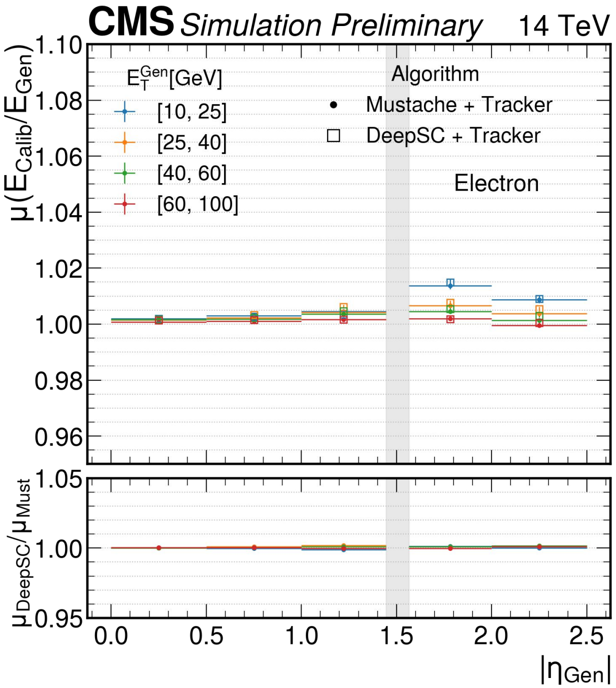
Energy scale and resolution performance: Photons



- The energy response of photons is estimated by fitting the calibrated photon energy divided by the true energy with a Cruijff function.
- The energy scale is found to be stable for all energies and the resolution either comparable or better compared to the current Mustache algorithm

Energy scale and resolution performance: Electrons





CMS-DP-2022/032

- The energy response of electrons is estimated by fitting the calibrated electron energy divided by the true energy with a Cruijff function.
- The energy scale is found to be stable for all energies and the resolution either comparable or better compared to the current Mustache algorithm

A regression to correct electron and photon energies

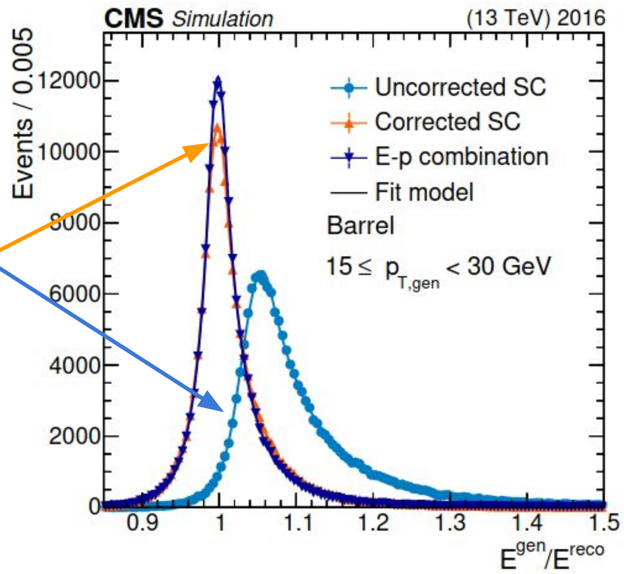
Why do we need this regression

- Electron/photon energies measured from the ECAL supercluster are subject to losses:
  - Energy lost in gaps and upstream material
  - Longitudinal energy leakage
  - Finite thresholds to suppress noise
  - Pileup and noise contamination

- Compensated per-particle using a ML regression Run-2 corrections implemented as a BDT
  - Uses ~30 high-level input features to describe EM shower

- BDT corrections first developed for Higgs discovery in 2012 and have supported all physics analyses using electrons/photons in CMS during LHC Run 2

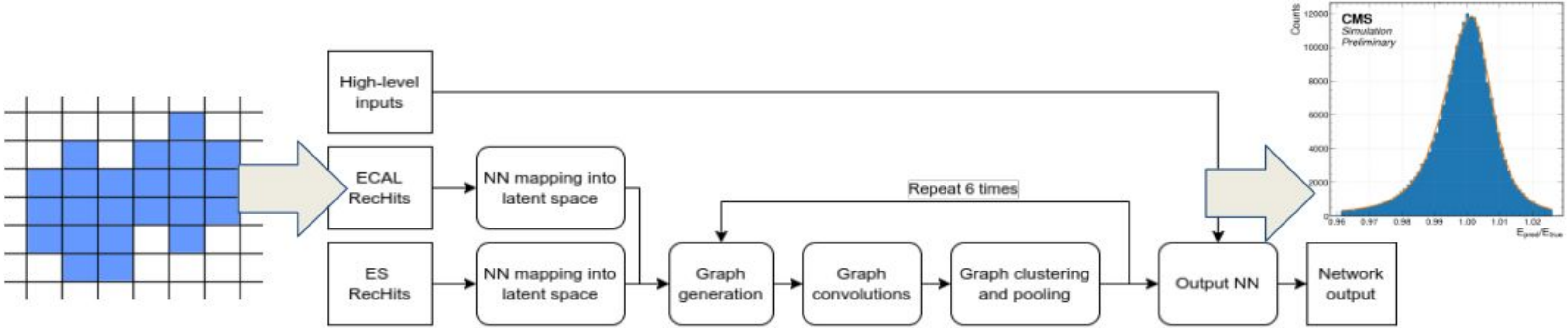
- **Can we do better using a GNN?**



JINST 16(2021) P05014

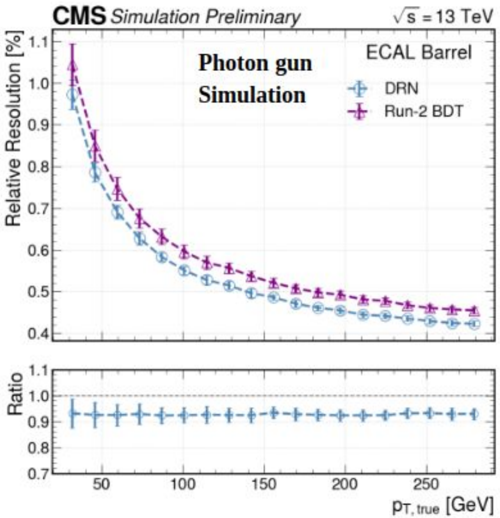
The Dynamic Reduction Network

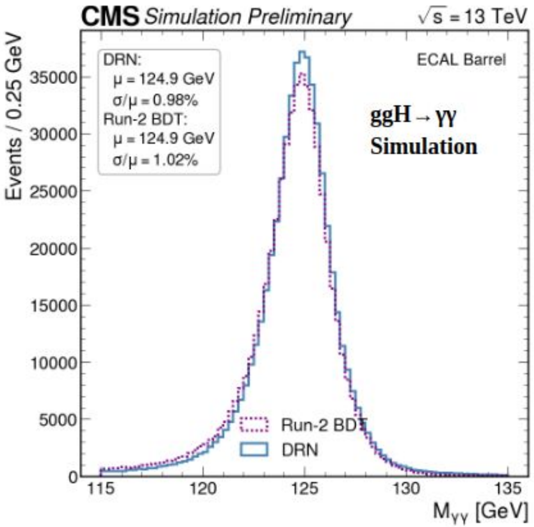
- New architecture: Dynamic Reduction Network (DRN)[1], based on dynamic GNNs
  - Graph operations take place in a high-dimensional latent space
  - Added clustering and pooling step to aggregate information across the graph
- Input includes hits from both ECAL and ECAL preshower, as well as additional features to describe information not encoded in the hit collection (pileup, leakage into HCAL)
- To account for the asymmetric response of the ECAL the regression targets are the parameters of a double-sided crystal ball probability density that models the energy response.

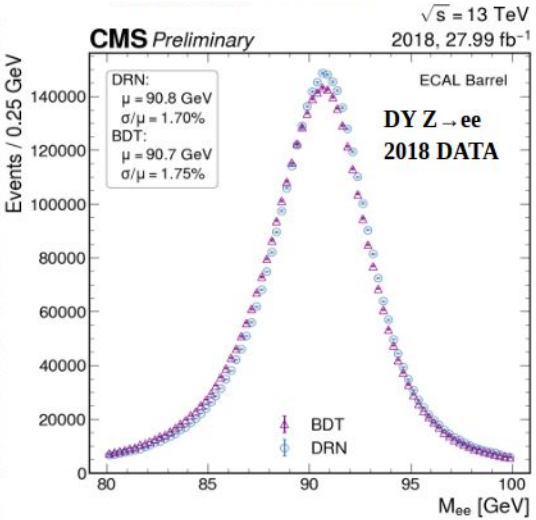


[1]: arXiv:2003.08013

Performance with the DRN







- Improved per-object energy resolution by ~10% for both electrons and photons
- Translates to improved invariant mass resolution by ~5%, confirmed in Run 2 data.
- First major change to energy regression algorithm since 2012
  - In process of being deployed for Run 3
- Similar algorithm extended for charged pion reconstruction with the CMS HGCAL.
  - Near 100% improvement in the energy resolution (arXiv:2406.11937)

Dominant systematic uncertainty in the m_H measurement in $H \rightarrow \gamma\gamma$

Source	Contribution (GeV)
Electron energy scale and resolution corrections	0.10
Residual p_T dependence of the photon energy scale	0.11
Modelling of the material budget	0.03
Nonuniformity of the light collection	0.11
Total systematic uncertainty	0.18
Statistical uncertainty	0.18
Total uncertainty	0.26

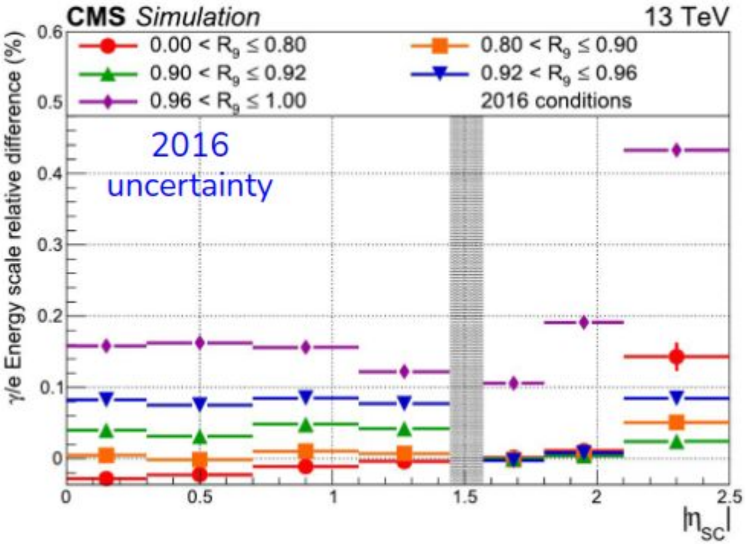
Phys. Lett. B 805 (2020) 135425

Dominant systematics:

→Residual energy scale

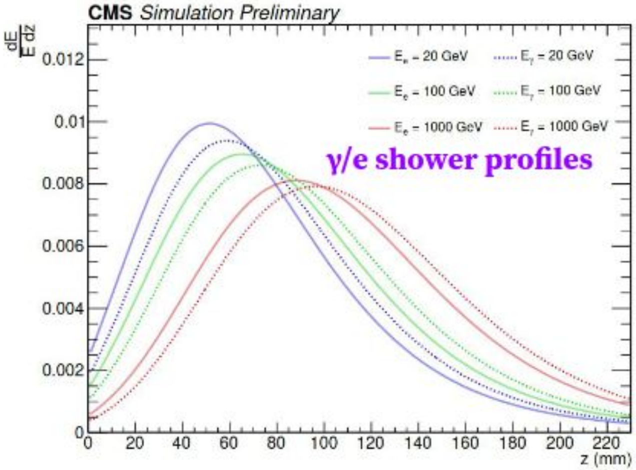
→**Non-uniformity of light collection**

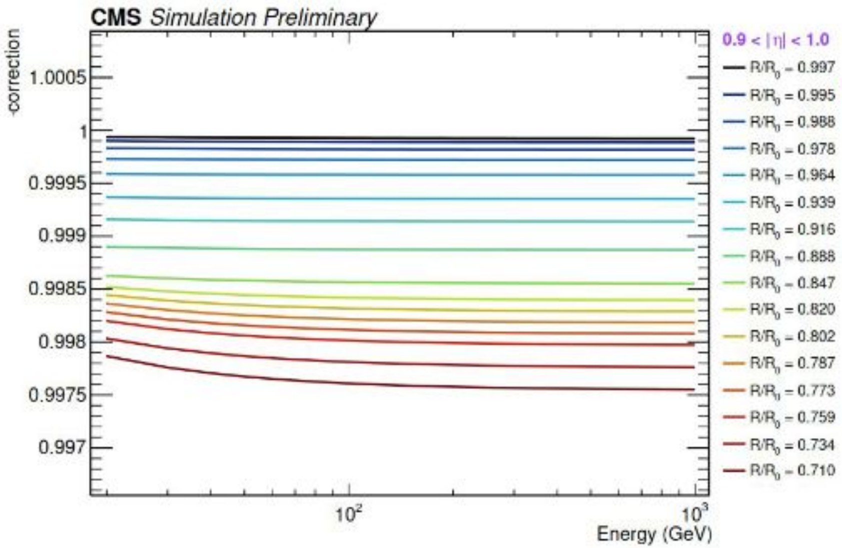
- **Non-uniformity of light collections** due to radiation damage:
  - **Energy scale corrections** derived in $Z \rightarrow ee$ events and are applied to photons
  - Photons penetrate ≈ $1X_0$ deeper than electrons in crystals
  - Radiation damage not simulated:
    - Impact of γ/e shower profile difference cannot be corrected by the energy regression
- **The difference in the energy scale** between electrons and photons is computed and **assigned as a systematic uncertainty.**



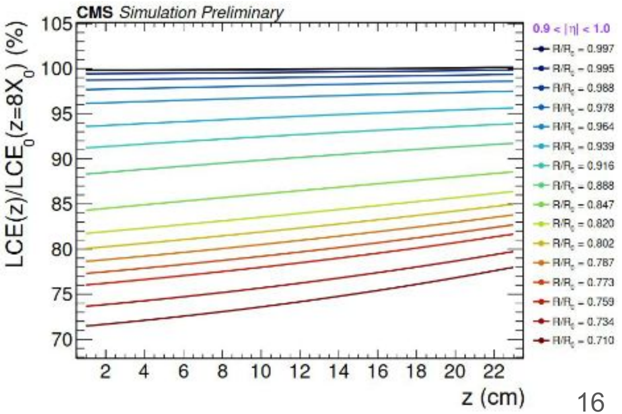
Impact of non-uniformity in light collection in the m_H measurement in $H \to \gamma\gamma$

- Model developed to predict light collection efficiency (LCE) using FLUKA+LITRANI **[CMS-DP-2024/045]**
- **New approach to reduce this uncertainty source** for the full Run 2 m_H measurement
  - **Compute and apply a correction to photons in data:**
    - **Assign an uncertainty to the correction** which is found to be significantly smaller that in the mass measurement with the 2016 dataset.



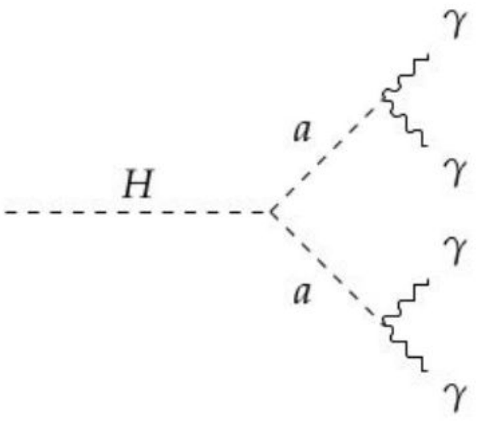


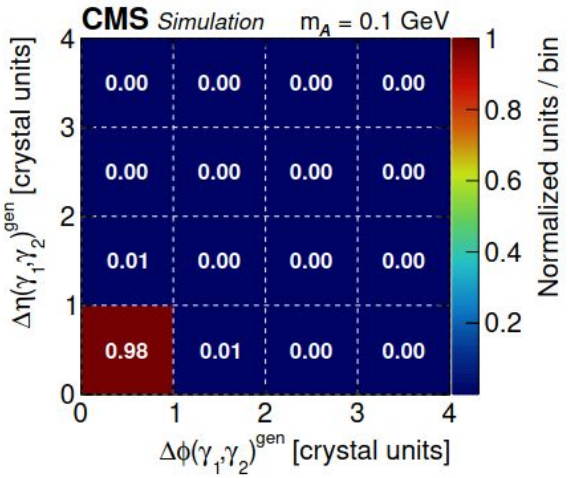
corrections and LCEs for different transparency scenarios (R/R_0)

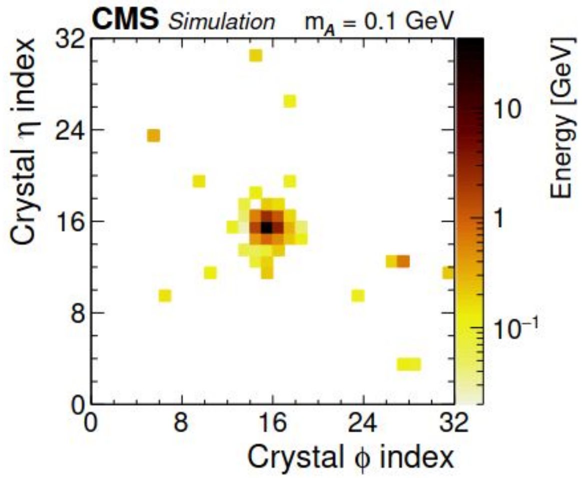


Merged photon reconstruction with the CMS ECAL

[Phys. Rev. D 108 (2023) 052002](https://doi.org/10.1103/PhysRevD.108.052002)



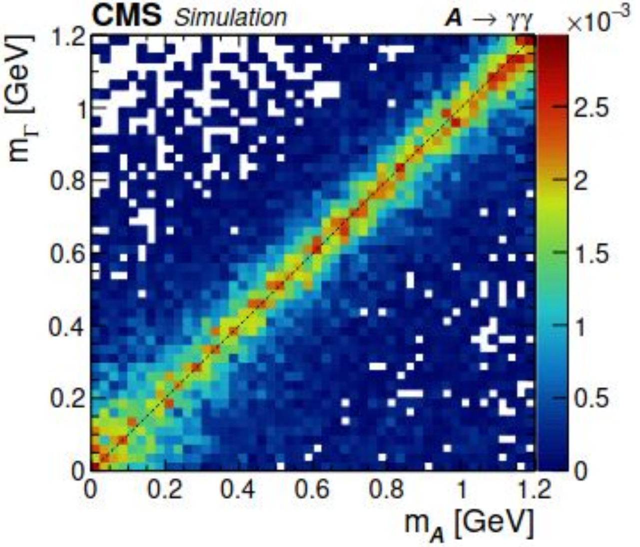


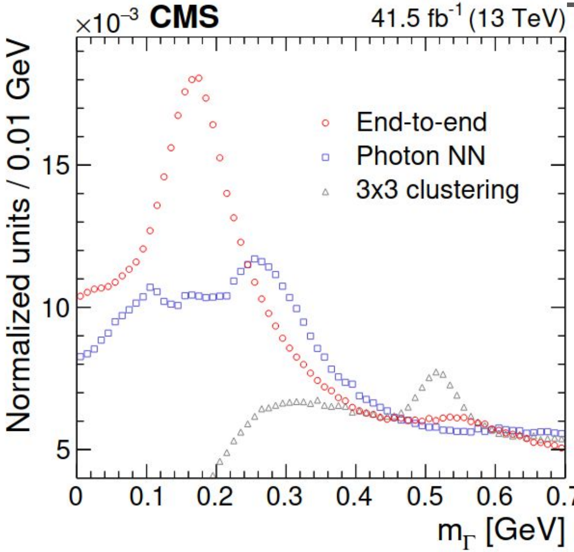


- In the context of new physics, eg ALPs from Higgs decay, OR SM particles, eg neutral pions, that are highly boosted, the diphoton pair they decay into appear as a single photon.
- In the case of ALPs like particles of mass ~100 MeV, from Higgs decay, nearly 98% of the two photons are contained within a single ECAL crystal at the SIM level.
- A CNN based end-end regression algorithm has been developed to reconstruct such topologies, training on the images of such decays in the ECAL, with the mass of the pseudo-scalar as the regression target.

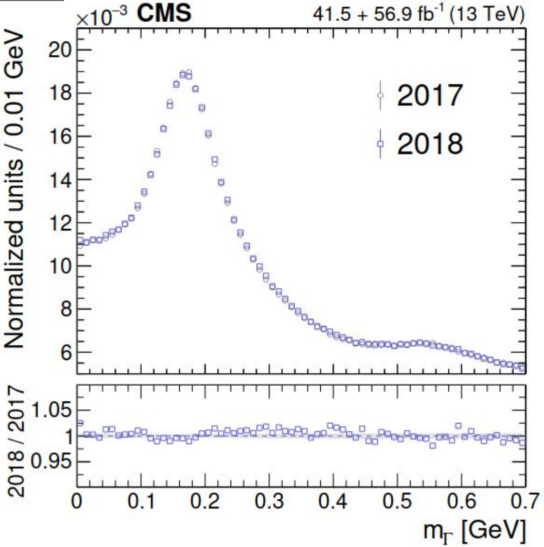
Merged photon reconstruction with the CMS ECAL

[Phys. Rev. D 108 (2023) 052002](https://doi.org/10.1103/PhysRevD.108.052002)





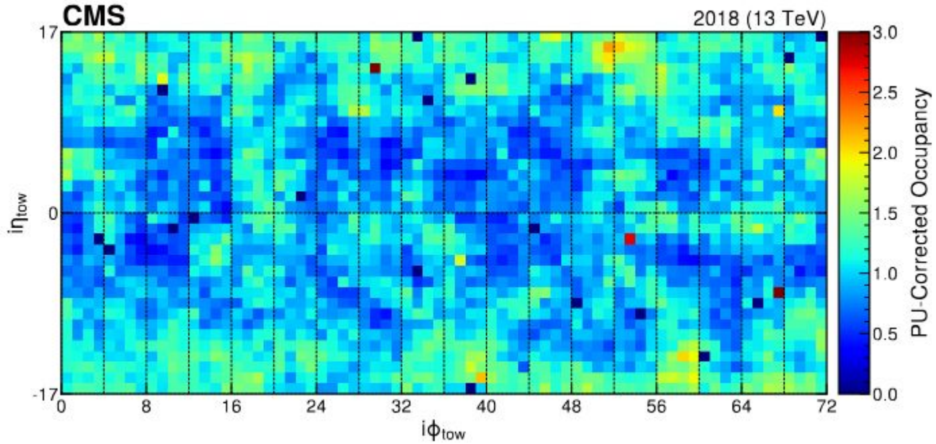
DATA

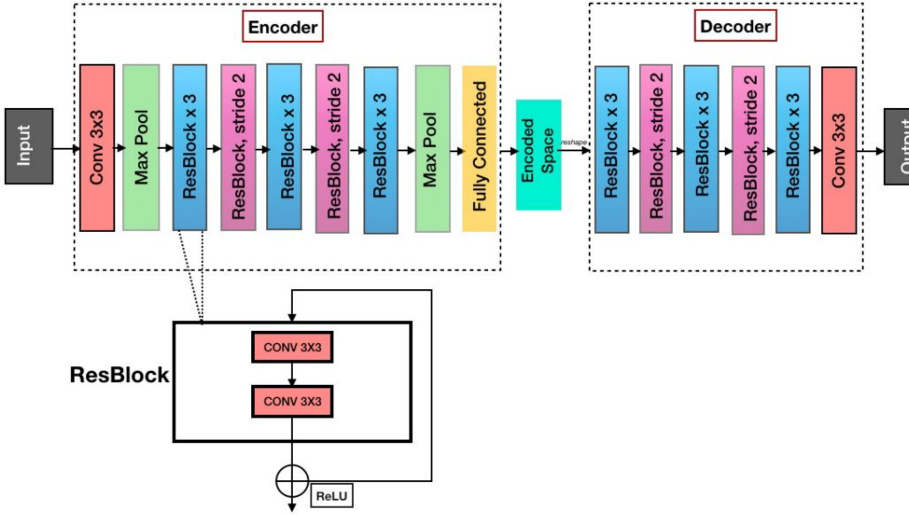


- While the algorithm is able to reconstruct ALPs from masses of 0.2 – 1.2 GeV, it is found to be able to reconstruct neutral pions from collision data and fairly resistant to pileup/noise.
- This algorithm is now being explored for the intercalibration of the ECAL using neutral pions in Run 3.

Anomaly detection based online data quality monitoring

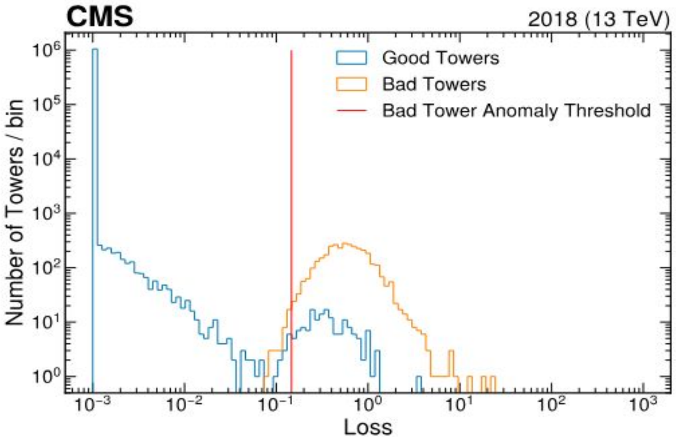
[Comput Softw Big Sci 8, 11 (2024)]





- Data Quality Monitoring (DQM) is a critical tool which allows detector experts to quickly identify and diagnose a broad range of detector issues that could affect quality of physics data being collected.
- An auto-encoder based anomaly detection tool has been developed to detect anomalies in real time
- The algorithm is trained on detector occupancy maps from runs that are certified as good

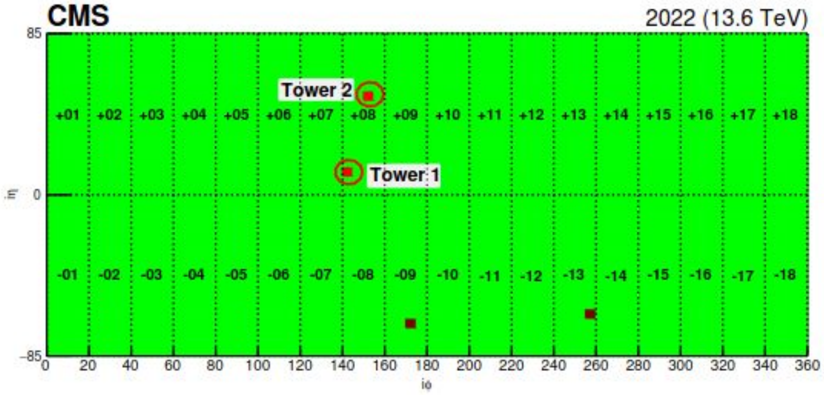
Performance of the anomaly detection algorithm

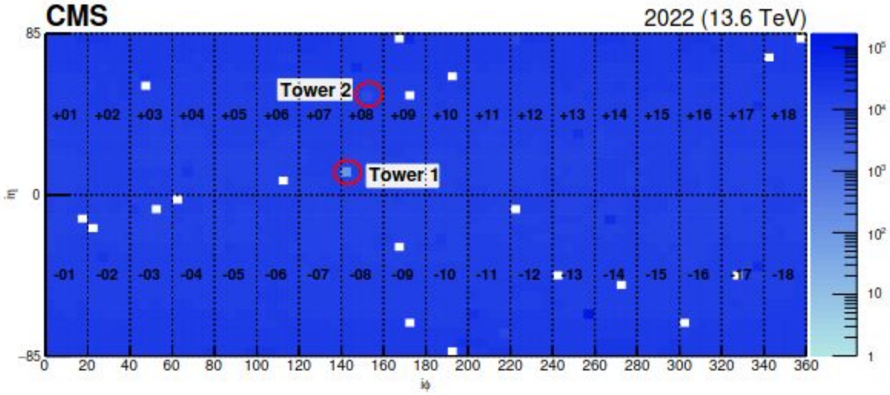


	FDR for 99% anomaly detection		
	Missing Supermodule (%)	Zero Occupancy Tower (%)	Hot Tower (%)
Baseline (no correction)	14	90	5.2
Baseline (after time correction)	5.9	80	< 0.01
AE (no correction)	3.6	51	2.8
AE (after spatial correction)	3.1	49	2.9
AE (after spatial and time corrections)	0.13	4.1	< 0.01

- A working point is chosen for 99% anomaly detection.
- With this working point the “False Discovery Rate” is evaluated for 3 distinct scenarios
- The AE based algorithm outperforms the baseline algorithm (cut-based algortihm) in all cases.

Anomaly detection based online data quality monitoring





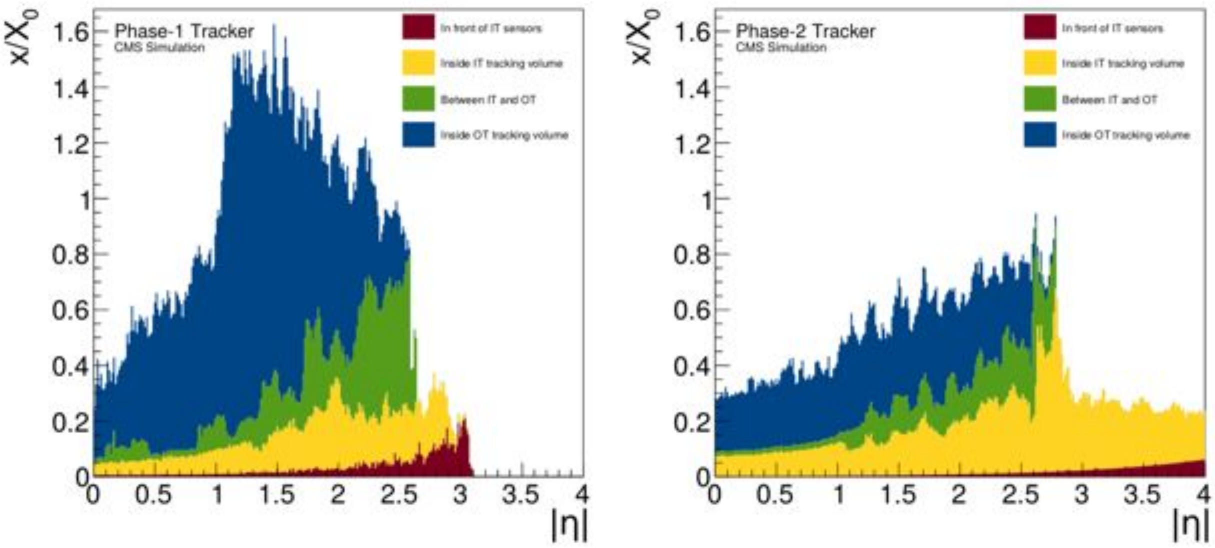
- This algorithm has been deployed in CMS for Run 3.
- This auto-encoder based DQM complements and further strengthens the existing DQM system.

Summary

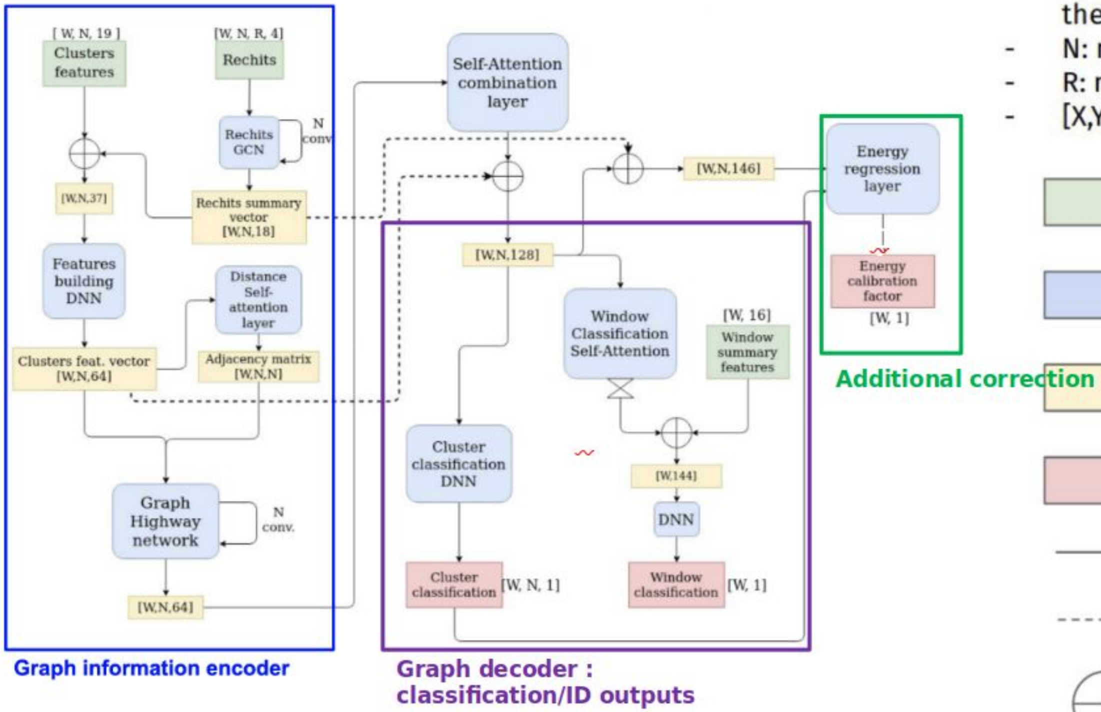
- The CMS ECAL has supported crucial high precision analyses and innovative searches in Run 2.
- The ECAL has also served as a nursery for the development of state-of-the art algorithms for the reconstruction of EM objects, exotic topologies, correcting for radiation damage, calibration and online monitoring.
- Several of these efforts have now sufficiently matured to significantly improve ECAL operations and physics performance in Run 3

BACKUP

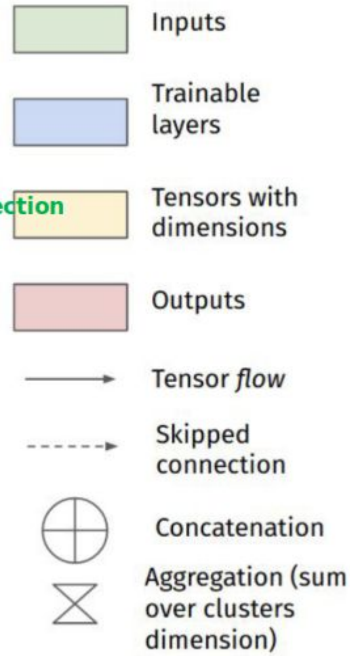
Material budget upstream of CMS ECAL



GNN architecture for the deep SC algorithm



- W : number of windows in the batch
- N: number of clusters
- R: number of rechits
- [X,Y,Z] tensor dimension



Correction to the photon energy scale to account for radiation damage

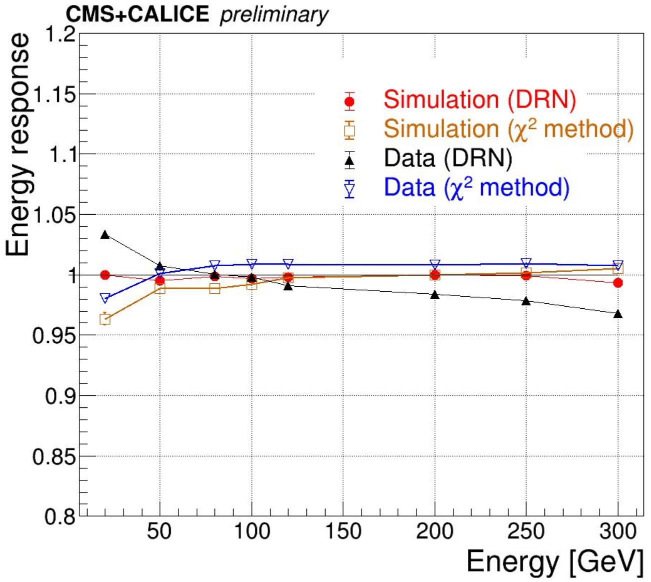
CMS-DP-2024/045

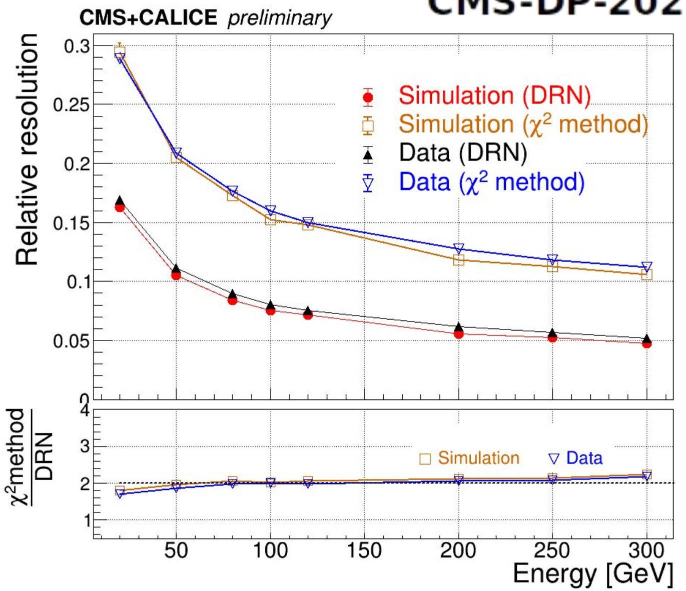
$$F = \frac{\frac{\int E^{e}_{dep}(z) LCE(z) dz}{\int E^{e}_{dep}(z) dz}}{\frac{\int E^{\gamma}_{dep}(z) LCE(z) dz}{\int E^{\gamma}_{dep}(z) dz}}$$

and applied to the calibrated energy of the photons in data. The correction is evaluated as function of the photon calibrated energy, η and the average R/R_0 of the crystals in the photon cluster, estimated using a transparency monitoring system based on laser light injection.

Energy response of charged pions with the HGCAL beam test prototype

CMS-DP-2022/022





- Use same Dynamic Reduction Network as in ECAL regressions
- Dramatic improvement in energy resolution w.r.t. χ^2 method
- Good agreement between the DRN predictions of the energy response and the relative resolution between beam test data and simulation