Graph Neural Networks for Calorimeter-Based Cluster Calibration at the ATLAS Experiment

Matthew Green
On behalf of the ATLAS Collaboration

ML4Jets 2026, 14th-18th of September, Vienna

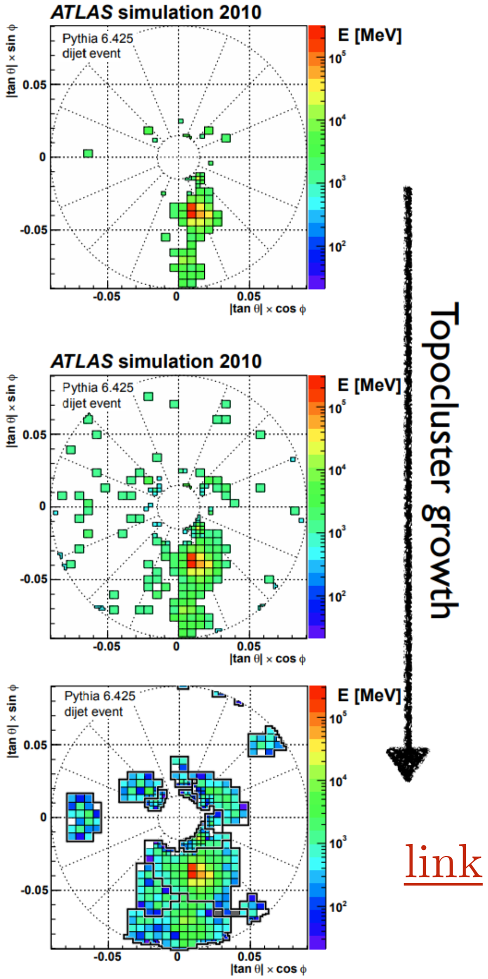
ATL-PHYS-PUB-2026-019

Jet constituent: Topo-clusters

- Topo-clusters are 3D connected-cell objects built from calorimeter cell energies via a topological clustering algorithm
  - Define significance for each cell:
    - Seed Cells: $\zeta_{clus}^{EM} > 4$
    - Neighbour Cells: $\zeta_{clus}^{EM} > 2$
    - Boundary Cells: $\zeta_{clus}^{EM} > 0$

$$\zeta_{\text{cell}} = \frac{E_{\text{cell}}}{\sigma_{\text{noise,cell}}}$$

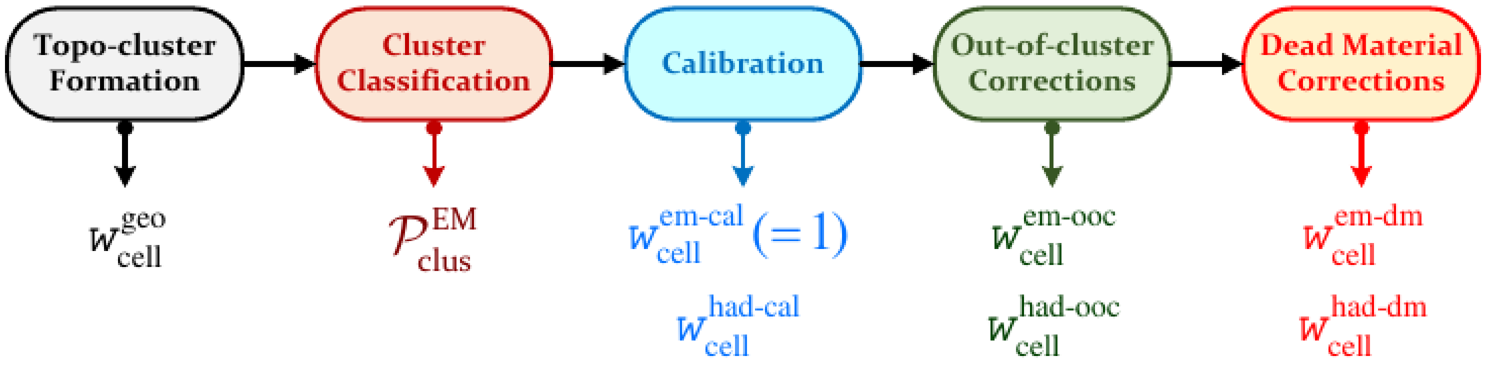
- Cell energies are calibrated at the electromagnetic (EM) scale, correctly measuring EM shower energy
  - Hadronic (HAD) showers deposit energy differently (nuclear breakup, invisible energy), so EM-scale calibration under-measures hadronic energy – motivating further calibration



link

Topo-cluster calibration

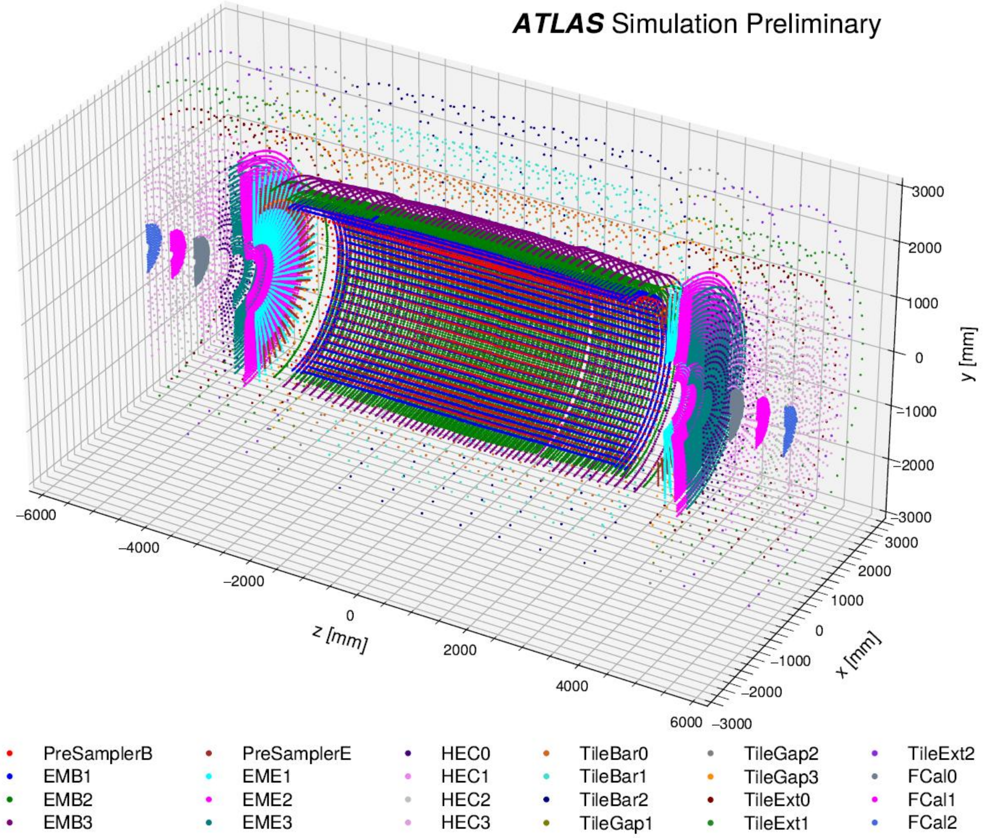
- Usual method: Local Cell Weighting (LCW)
  - Multi-step sequence which has look-up-table corrections
    - w.r.t $\mathcal{P}_{\text{clus}}^{\text{EM/HAD}}$ - likelihood of the topo-cluster being attributed to an EM/HAD-like energy deposit



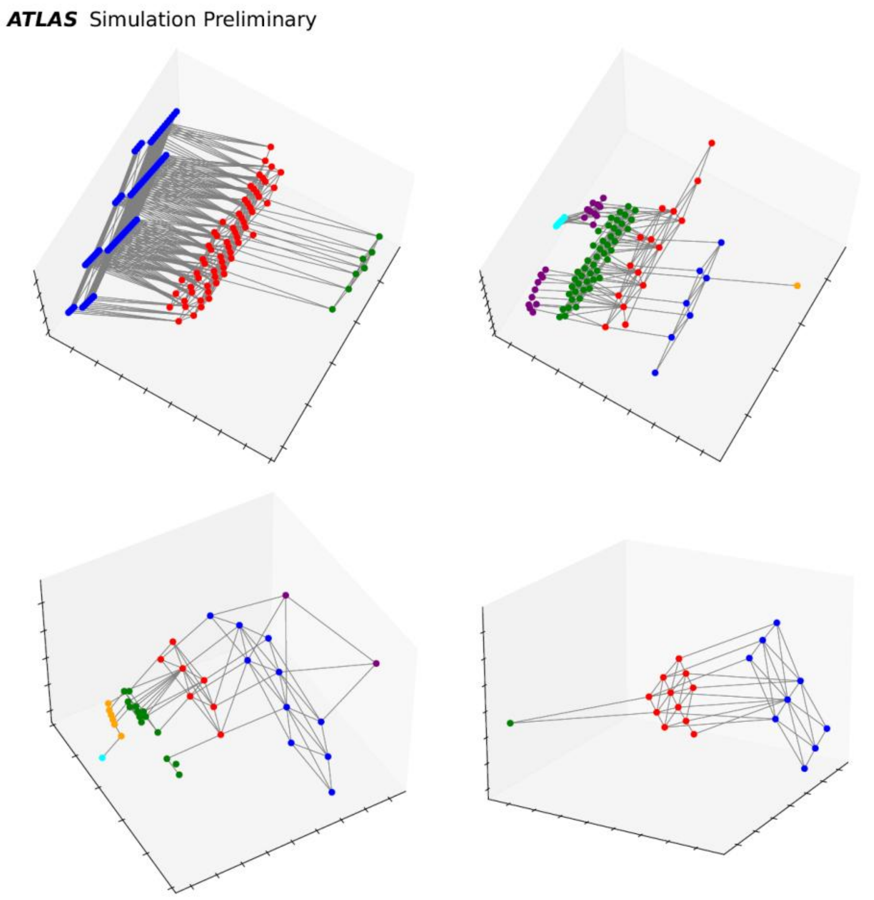
- Calibration factors are extracted from rigid multi-dimensional binned look-up tables and hence bin-edge effects limit performance.

Calorimeter geometry

- The ATLAS calorimeter system consists of 180,000+ individual cells, in a 3D space
- Can we exploit this natural geometry using graph neutral networks (GNNs) for cluster calibration?



Cluster Graphs



- Nodes ($\mathcal{V}$) – Cells of the cluster
  - Features: Cell Energy, Sampling layer, η, ϕ, $r_{\perp}$, $\Delta\eta$ & $\Delta\phi$ (from barycentre)
- Edges ($\mathcal{E}$) – Detector geometry

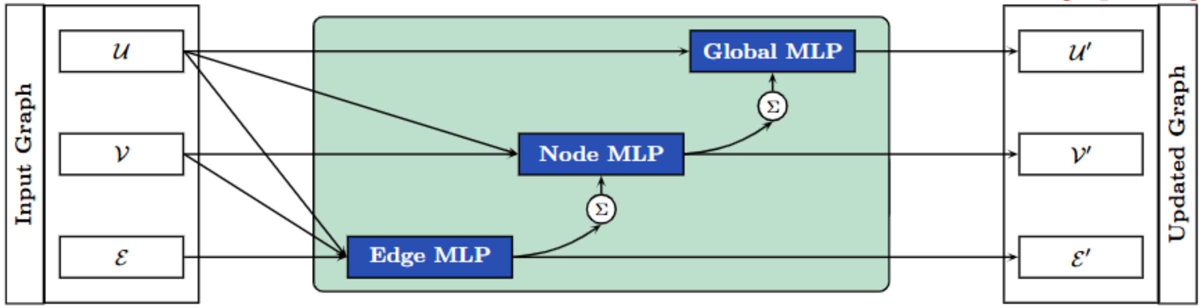
Edge type	Description
`prevInEta` / `nextInEta`	Adjacent in η, same layer
`prevInPhi` / `nextInPhi`	Adjacent in ϕ, same layer
`prevInSamp` / `nextInSamp`	Adjacent sampling layer
`prevSubDet` / `nextSubDet`	Adjacent across sub-detectors
`prevSuperCalo` / `nextSuperCalo`	Used for the densely packed FCAL
`corners2D`	Diagonal adjacency, combining `InEta` and `InPhi` relations

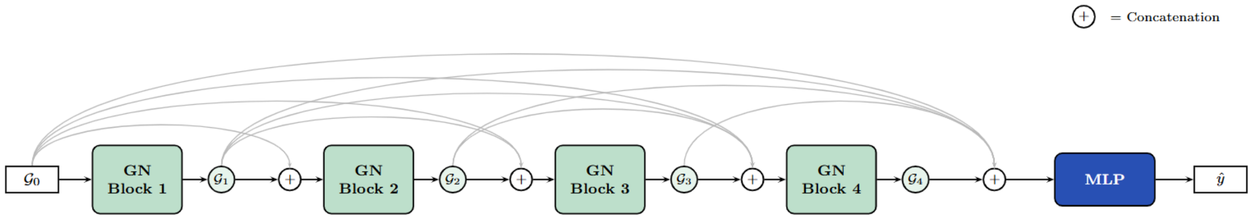
- Global ($\mathcal{U}$)– Reconstructed Cluster Energy (EM)

GNN Model

- 4 layers of a GN block
  - Utilising standard message-passing techniques
- Connect these GN blocks using Dense-Net-like connections
- Into a MLP for regression
- GNN will be predicting the <u>truth cluster energy</u>

Message-passing



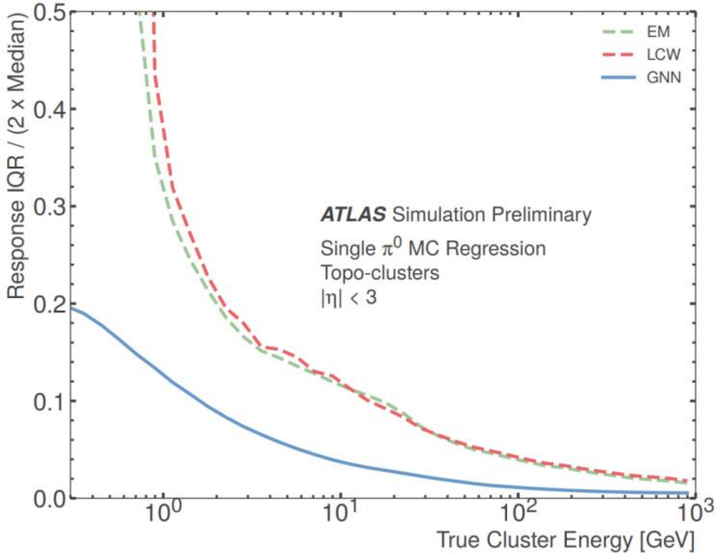


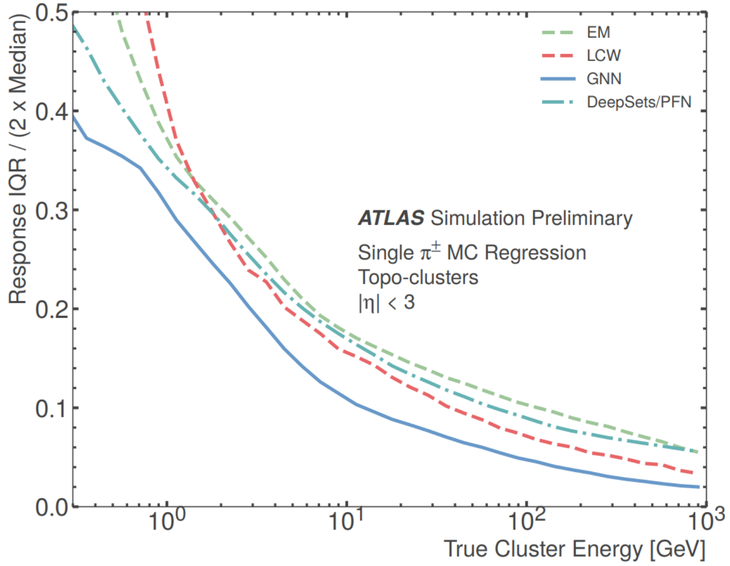
DenseNet

Hyperparameter	Value
Model architecture	
Number of blocks	4
Latent size	64
Number of MLP layers	4
Aggregation function	Mean
Training	
Batch size	1024
Epochs	100
Train/validation split	0.75
Inital learning rate	1×10^{-3}
Patience (epochs before LR reduction)	5
Learning rate reduction factor (Δ)	2
Early-stopping patience (epochs)	20
Loss	
Loss function	Mean absolute error

Previous results

- In 2022, ATLAS released a PUBNOTE focusing on clusters from π's





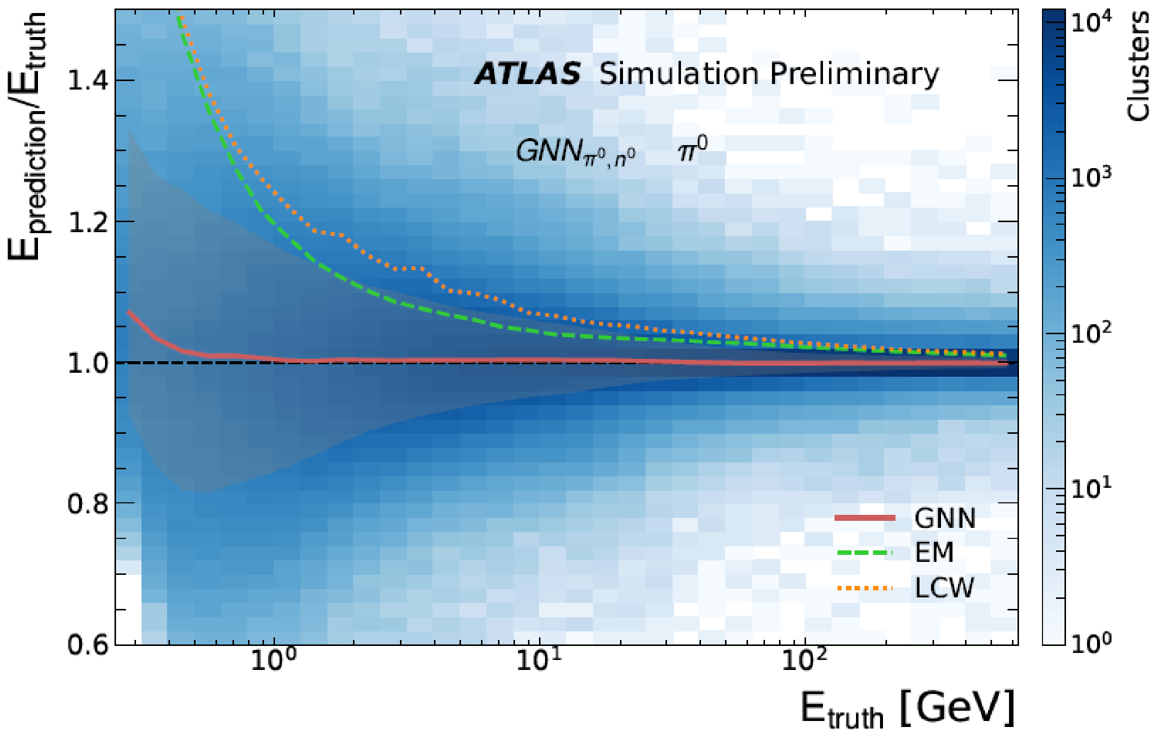
- The results I show today expand upon these results significantly by using the same GNN architecture, but applied to a wider range of particle types in more complex environments

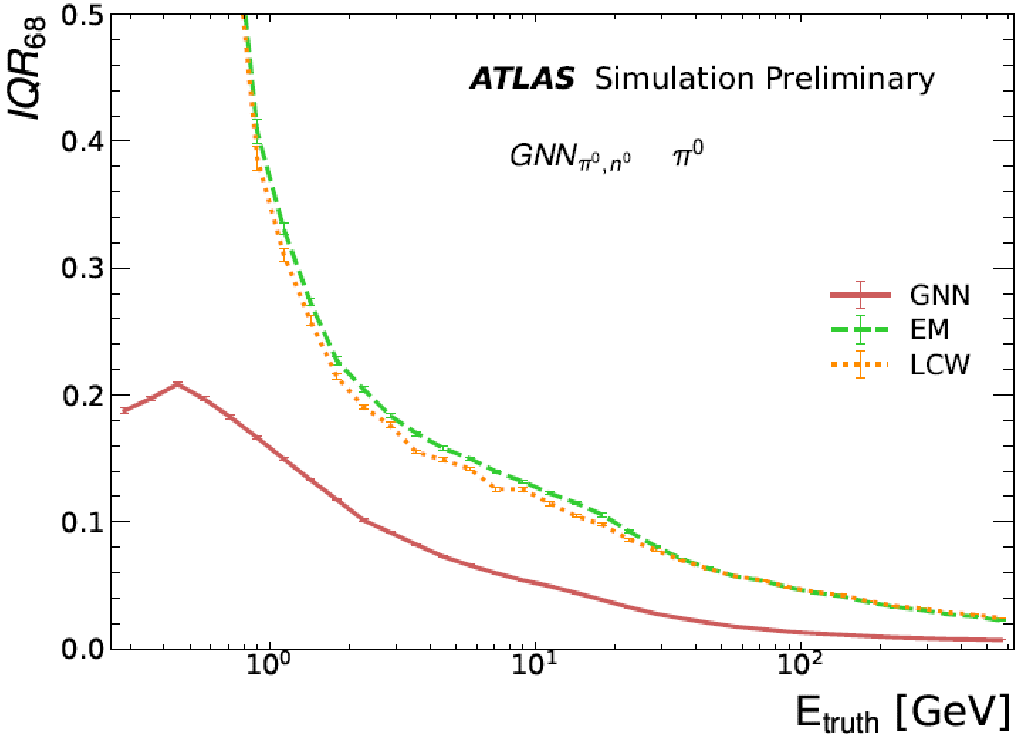
Results

- Several models are trained, each with inputs of more complex topologies for the clusters
- Each of these were simulated with full Run3 ATLAS Simulation
  - GNN_{π^0,n^0}: 1M clusters from each π^0 and n^0 ParticleGuns
  - $\mathrm{GNN}_{\mathrm{Dijet}}$: 3M clusters from QCD Dijets MC
  - $\mathrm{GNN}_{\mathrm{Dijet,PF}}$: 3M clusters from QCD Dijets MC, ParticleFlow applied
- 2 performance metrics are used:
  - (Median [$\mathcal{R}_{50}$]) Energy Response: $\mathcal{R} = \frac{E_{prediction}}{E_{truth}}$
  - Measure of spread, $\mathrm{IQR}_{68} = \frac{1}{2}\frac{\mathcal{R}_{84}-\mathcal{R}_{16}}{\mathcal{R}_{50}}$,
- Compared against the baselines of EM and LCW

NB: All samples are without pileup

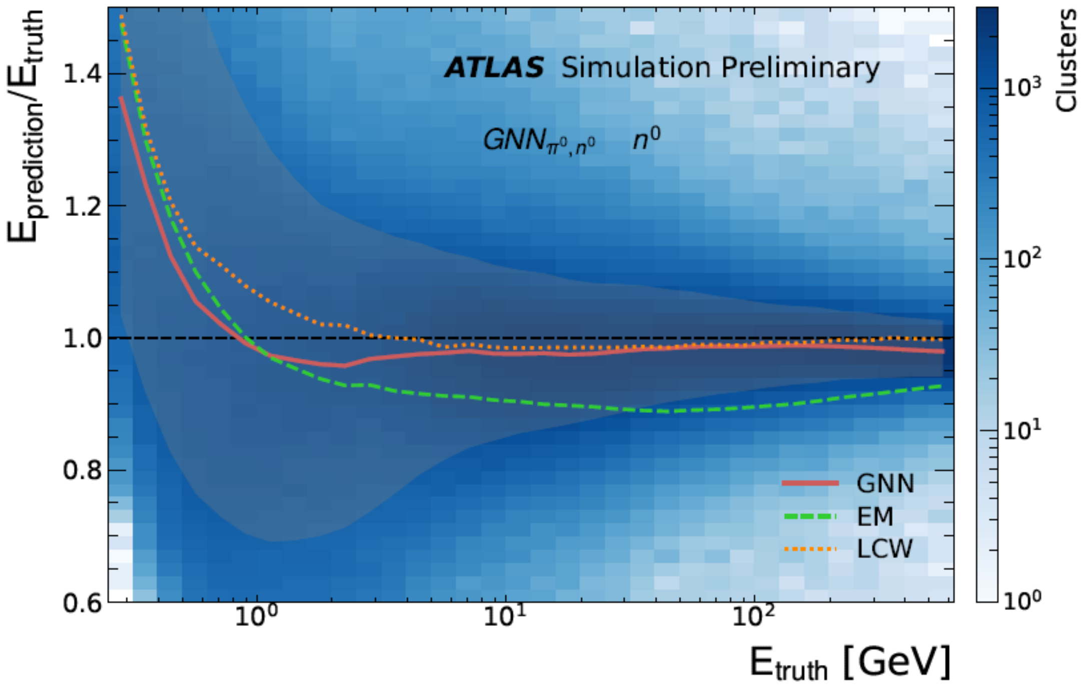
GNN_{π^0,n^0} for π^0

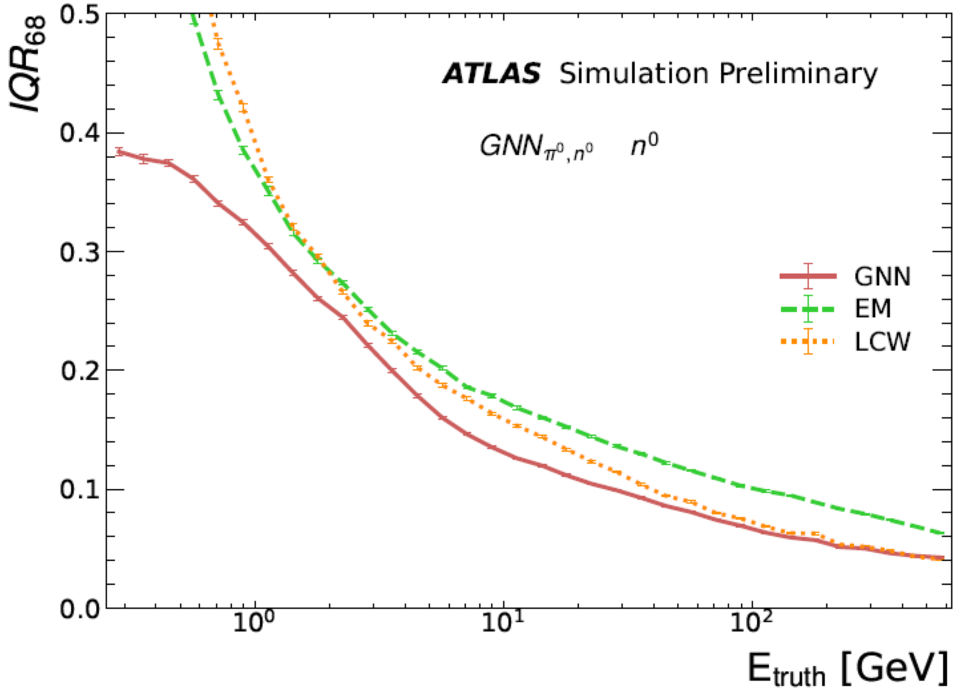




- Very good linearity compared to EM/LCW.
- Much better resolution (+30% in some cases)

GNN_{π^0,n^0} for n^0

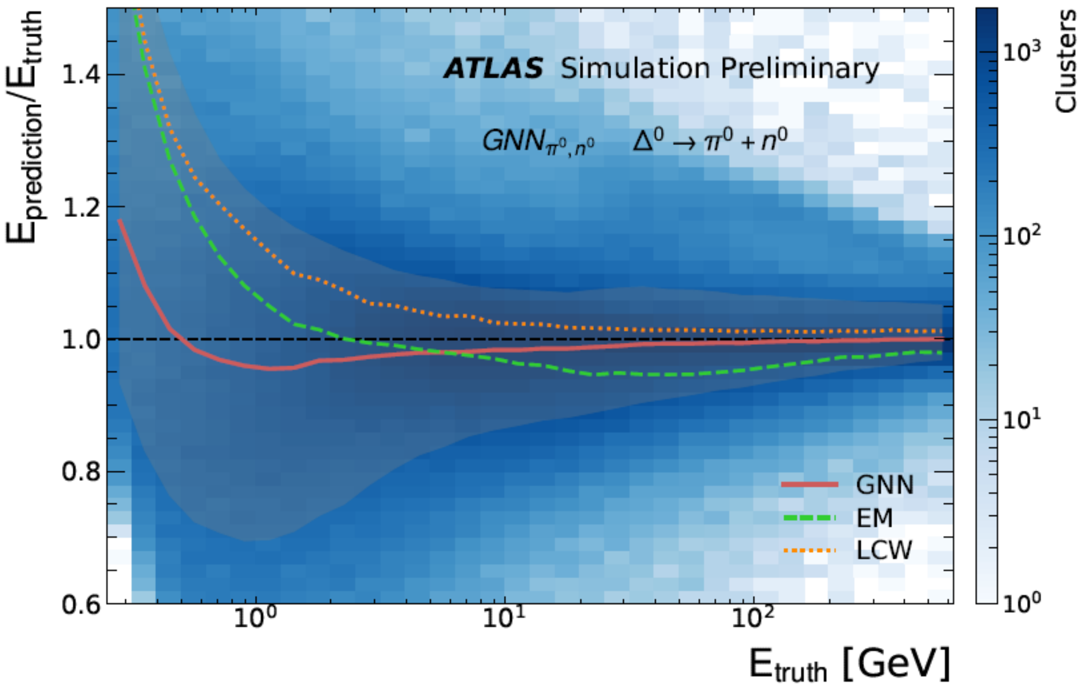


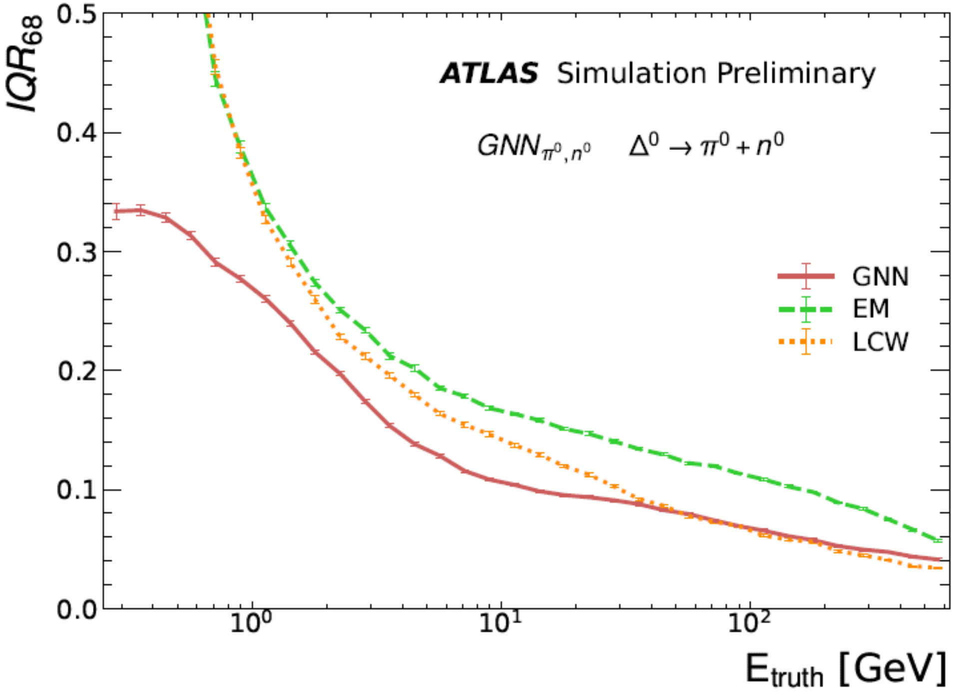


- Resolution still outperforms the baselines, especially at low E_{truth}
- Comparable to LCW at high E_{truth}

GNN_{π^0,n^0} for $\Delta^0 \rightarrow \pi^0 + n^0$

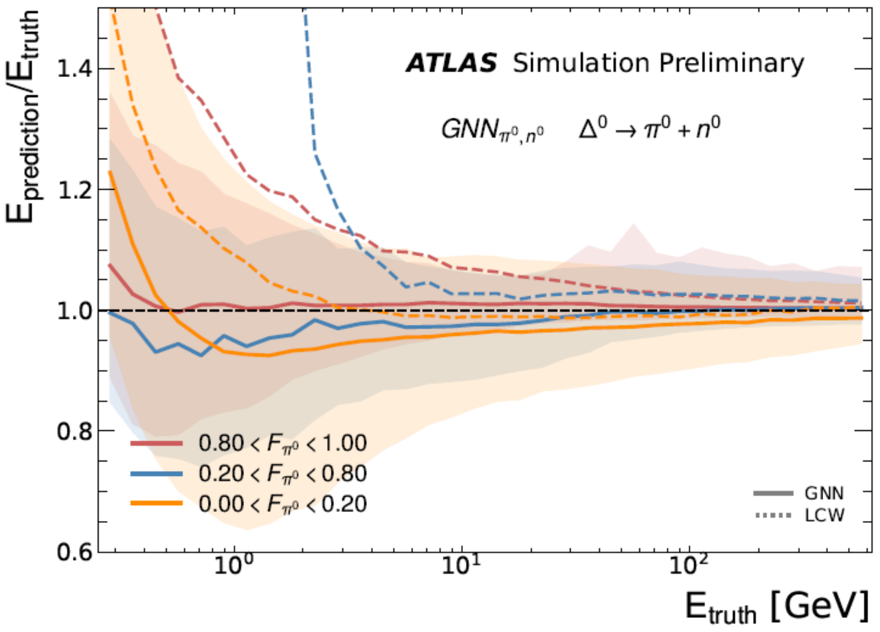
- ➢ To see if the GNN can generalise, inference on Δ-ParticleGun, enforcing the $\Delta^0 \rightarrow \pi^0 + n^0$ decay

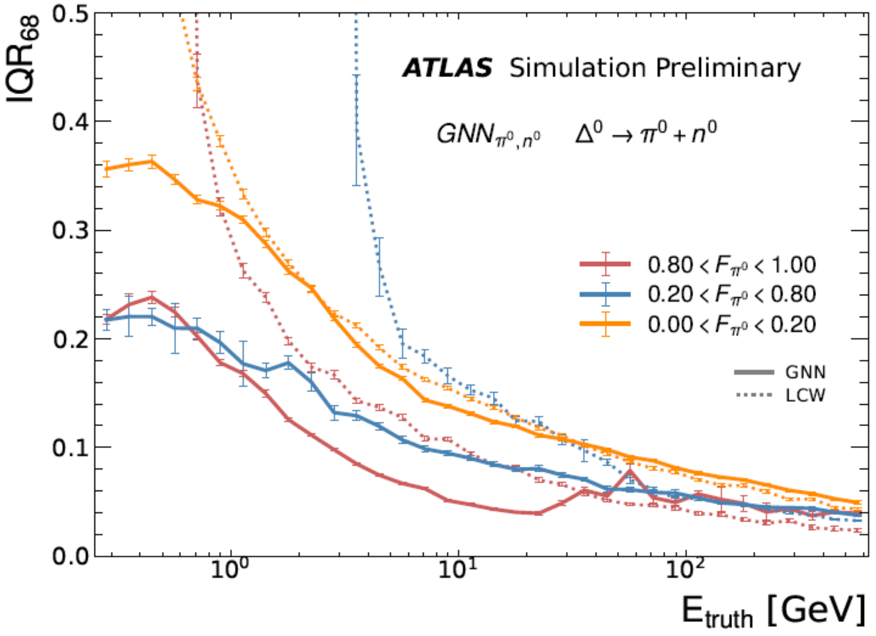




Composite clusters

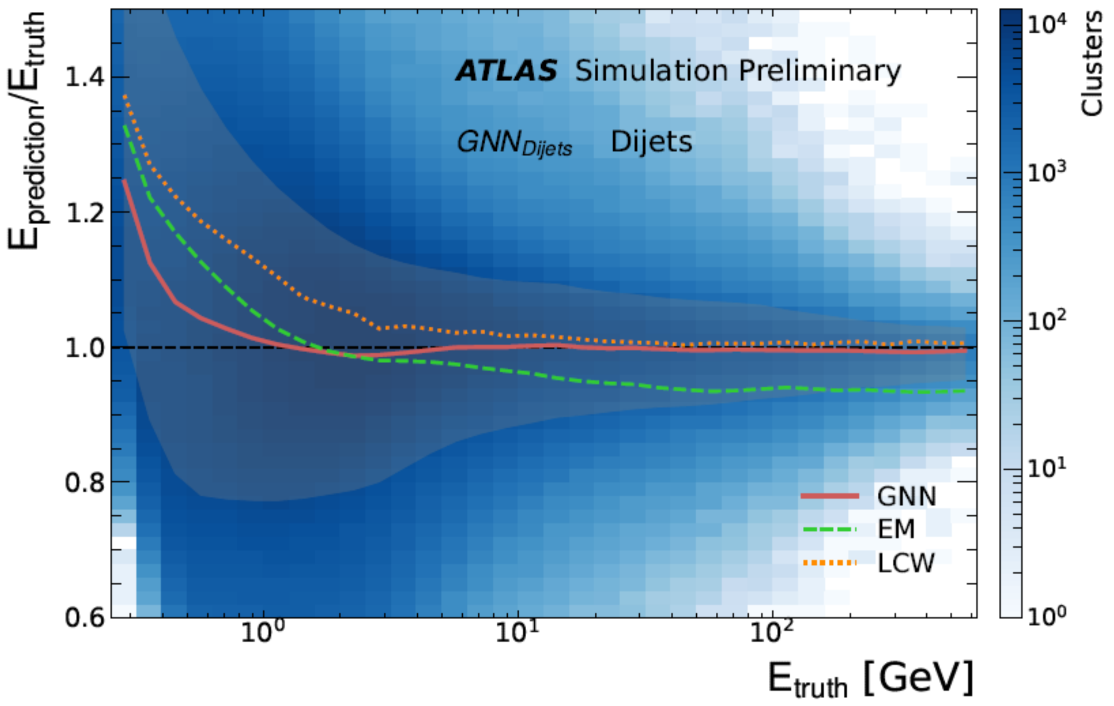
- F_{π^0} - Fraction of energy attributed to the π^0

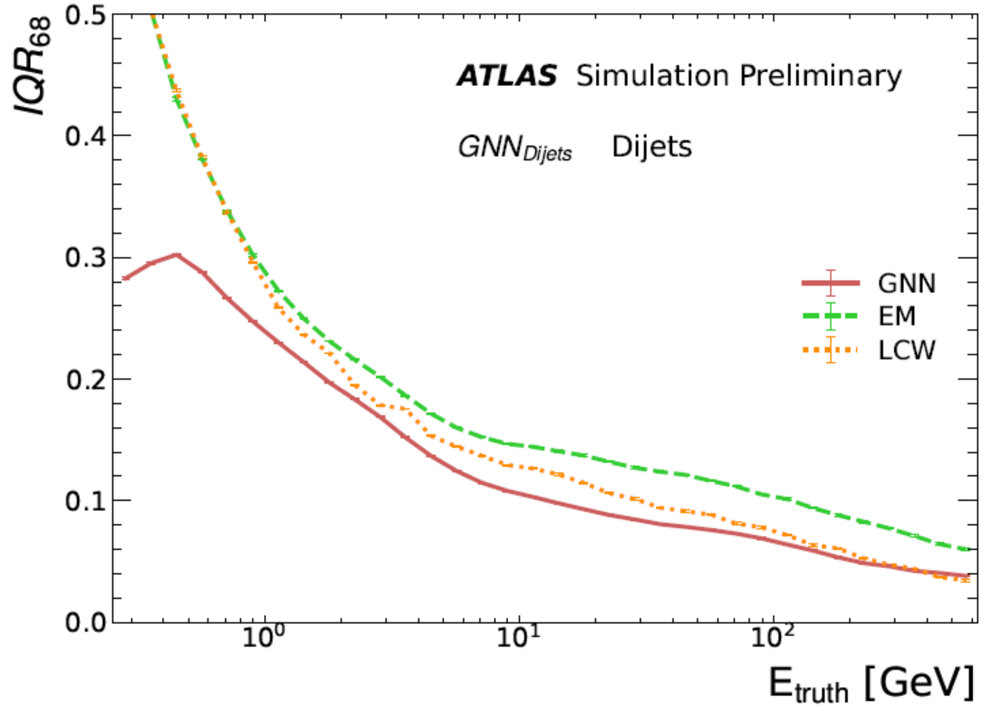




- Similar performance as the single-particle clusters
- Combined clusters severally outperform the LCW baseline
  - GNN can generalise to composite clusters

$\mathrm{GNN_{Dijet}}$

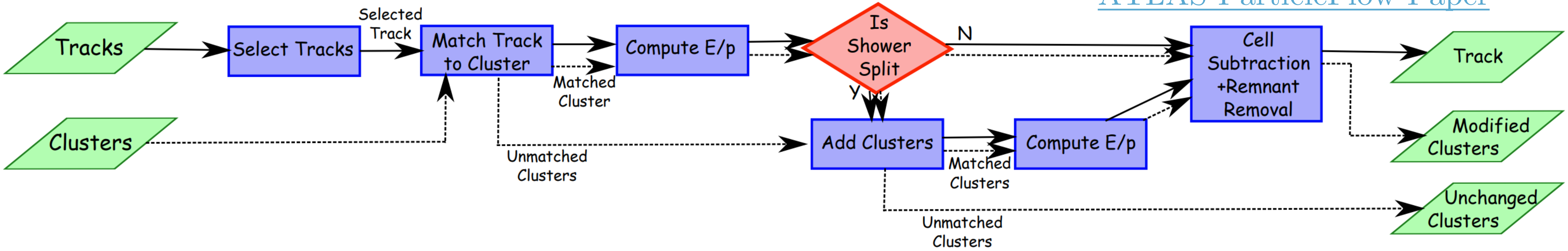




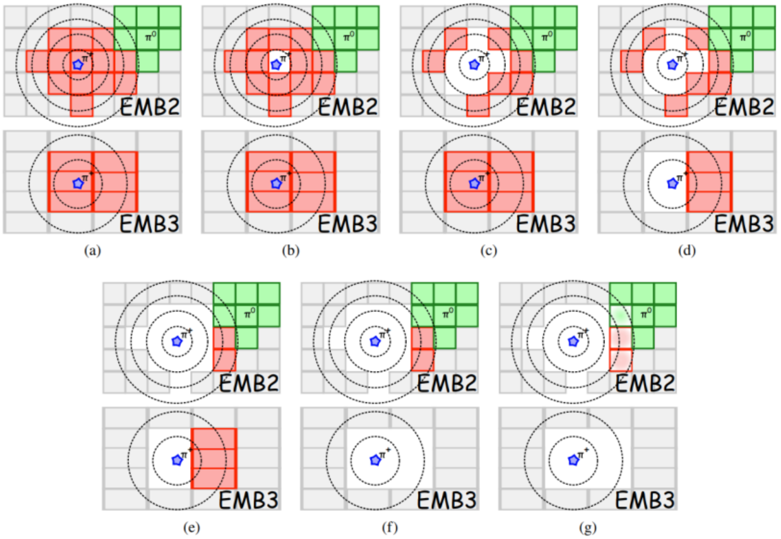
- To limit test the model, it was trained with clusters from QCD Dijets
  - Outperforms LCW on the lower E_{truth}, comparable at high E_{truth}

ATLAS ParticleFlow Overview

[ATLAS ParticleFlow Paper]

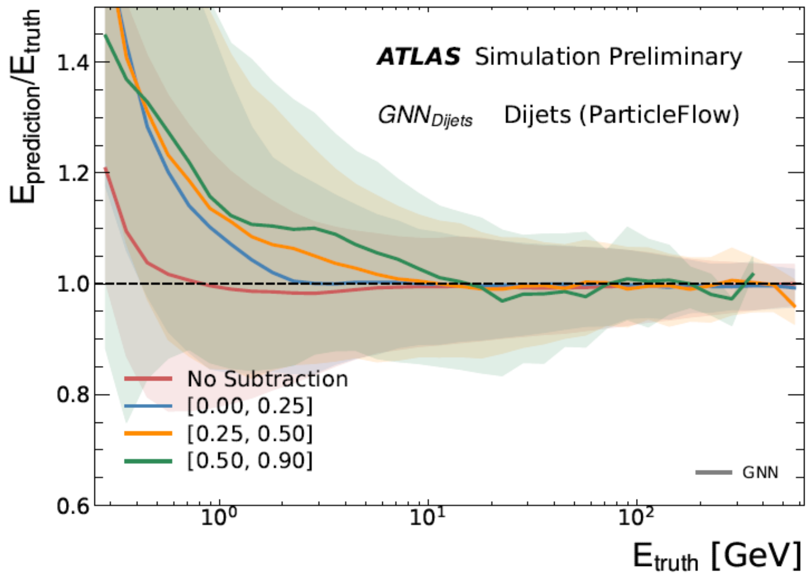


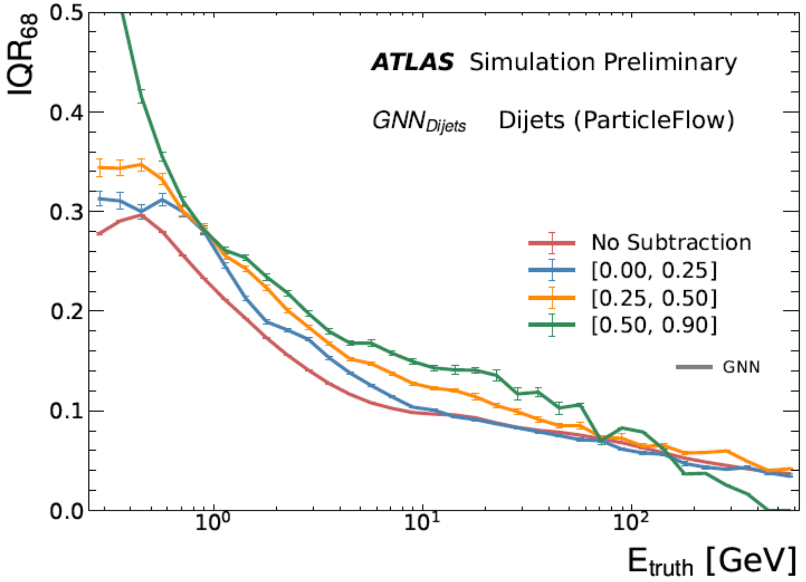
- ➢ Maximises the entire detector by leveraging the tracker information
- ➢ Removes the charged energy from the clusters via cell subtraction
  - Can the GNN be used on the cluster remnants?



$\text{GNN}_{\text{Dijet}}$, in bins of $F_{cells\,removed}$

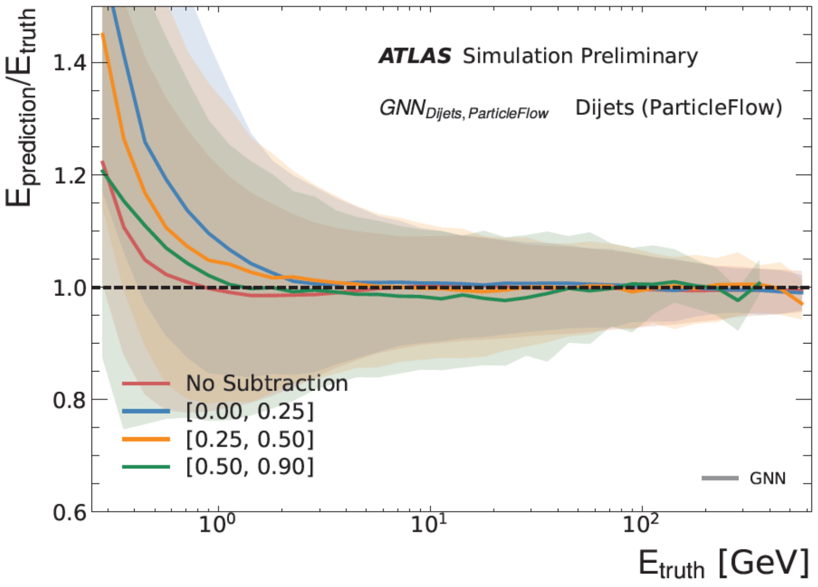
- Perform ATLAS ParticleFlow => inference on these clusters
- $F_{cells\,removed} = \dfrac{\text{number of cells after PFlow}}{\text{number of cells before PFlow}}$

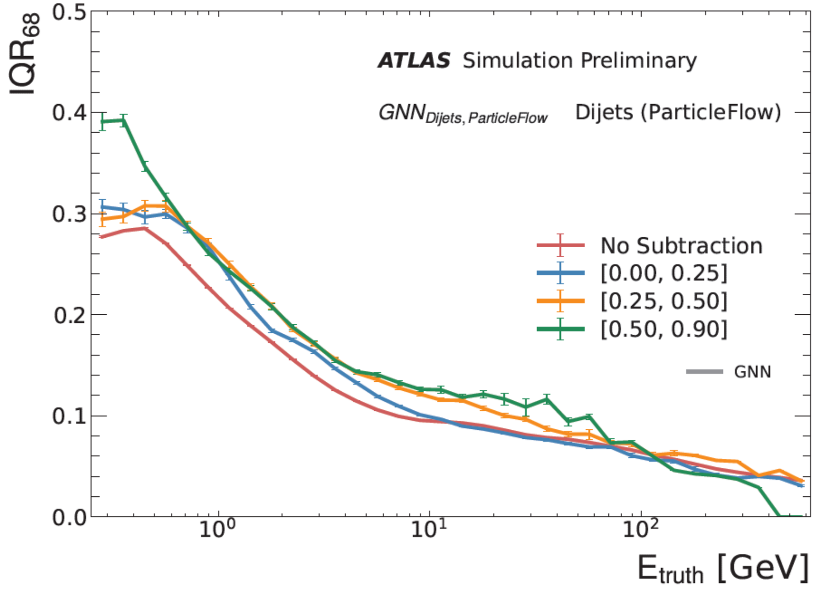




$\text{GNN}_{\text{Dijet,ParticleFlow}}$, in bins of $F_{cells\ removed}$

- Retrained the model using clusters that have gone through ParticleFlow





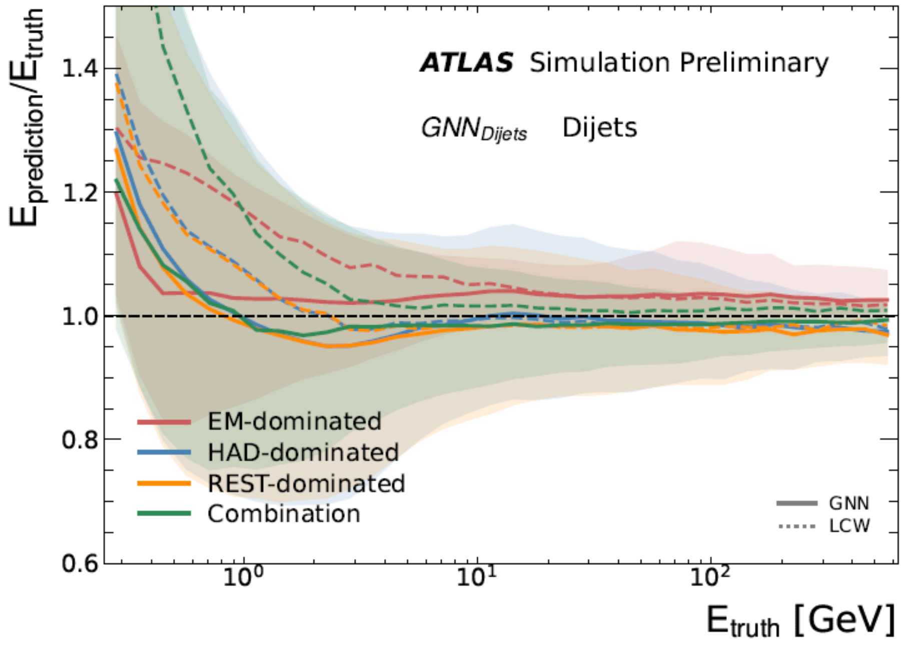
- Clear improvement and much better linearity in the response

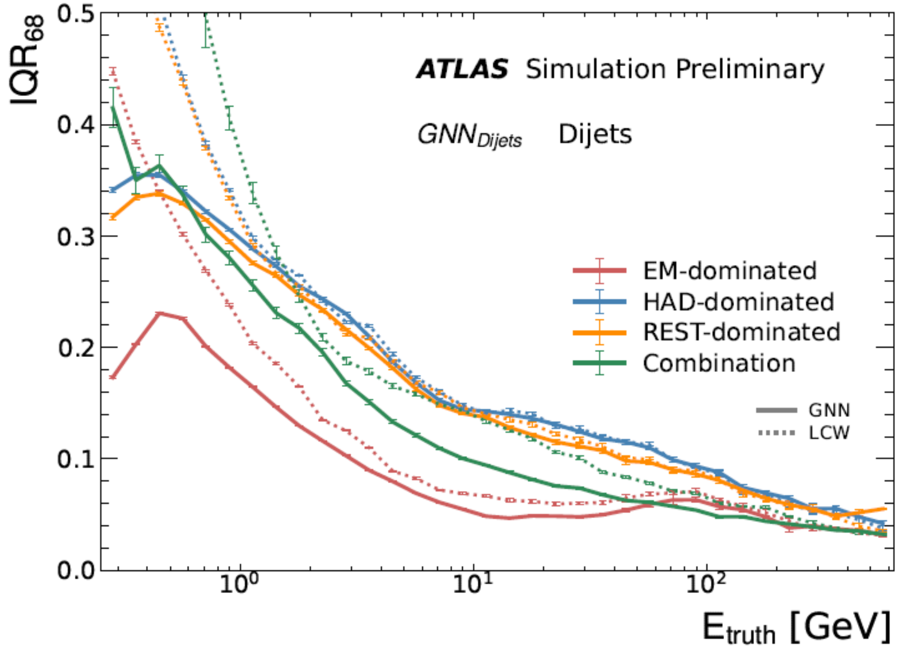
Conclusion

- Showed that using the GNN-based cluster calibration performs well for complex topologies:
  - Able to generalise more than the traditional LCW approach for composite clusters
  - Similar performance was shown for dijets, novel for this architecture as it was only previously used on ParticleGun samples
  - GNN can be used as a post-ParticleFlow scheme
    - Training on these clusters explicitly improves the calibration without spoiling the performance of non-subtracted clusters.

Backup

$\text{GNN}_{\text{Dijet}}$, in bins of EM/HAD/REST





“Dominated” - >90% of particle type
“Combination” - <70% of particle type

- EM: electrons, photons and π^0
- HAD: $\pi^{\pm}$
- REST: attributed to any other particle (n^0, K_L, etc.)

Truth Information: Calibration Hits

- Associated with every particle, and every calorimeter cell

Table 1: The four contributions to the truth energy of each stable truth particle recorded by calibration hits.

Contribution	Description
Electromagnetic	Energy deposited via electromagnetic interactions
Non-electromagnetic	Energy deposited via non-electromagnetic (hadronic) interactions
Invisible	Energy lost primarily to nuclear absorption, not detected
Escaped	Energy carried by particles that do not interact and escape the detector (neutrinos)