Transformer for Energy Calibration in the ATLAS Electromagnetic Calorimeter

Ryan Roberts for the ATLAS Collaboration

ML4Jets - November 5, 2024

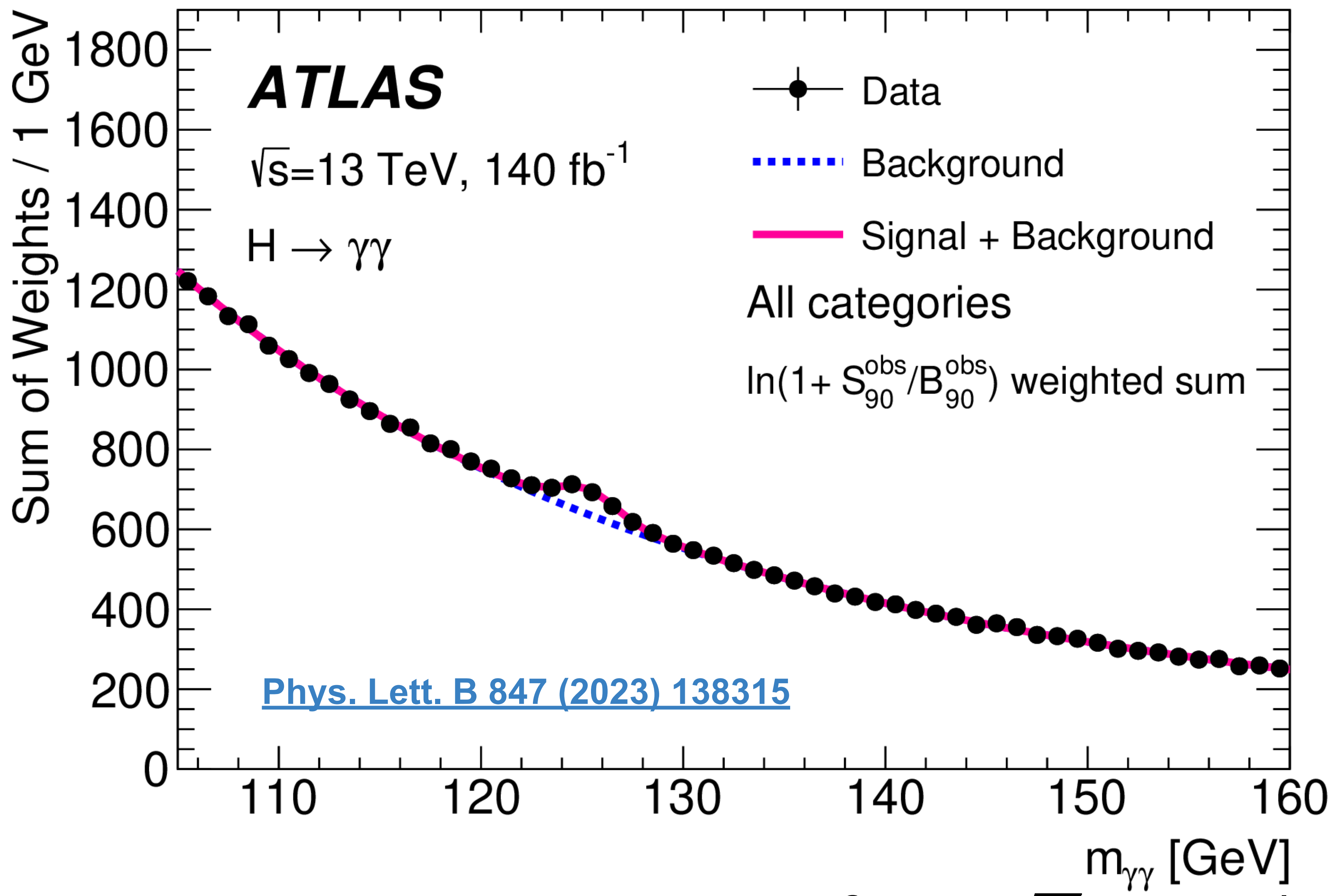
EGAM-2023-01

Electrons and Photons

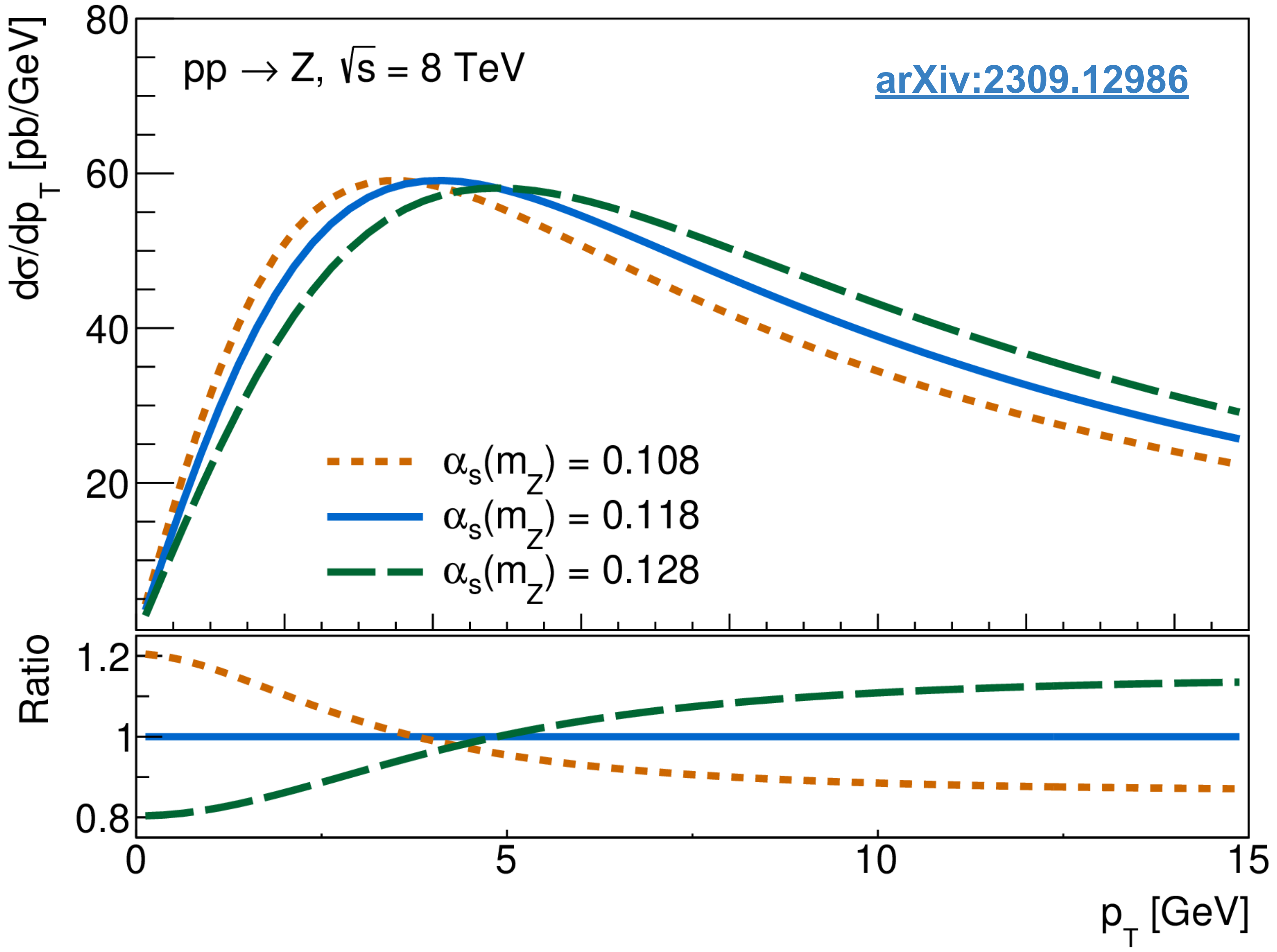
Physics Goals

- **Electrons** and **photons** are key to many important measurements at the LHC
  - Important to separate interesting processes from large QCD **backgrounds**
  - More **precisely measured** than hadronic states
  - Includes precision **measurements**, rare SM processes, **and searches**
- **Dedicated sub-detector** to contain and measure the **electromagnetic showers**
  - Longitudinal and radial segmentation to measure the position and shape of the shower
  - Ultimate goals are reconstruction and identification efficiency, energy resolution, and fake rejection

Higgs mass

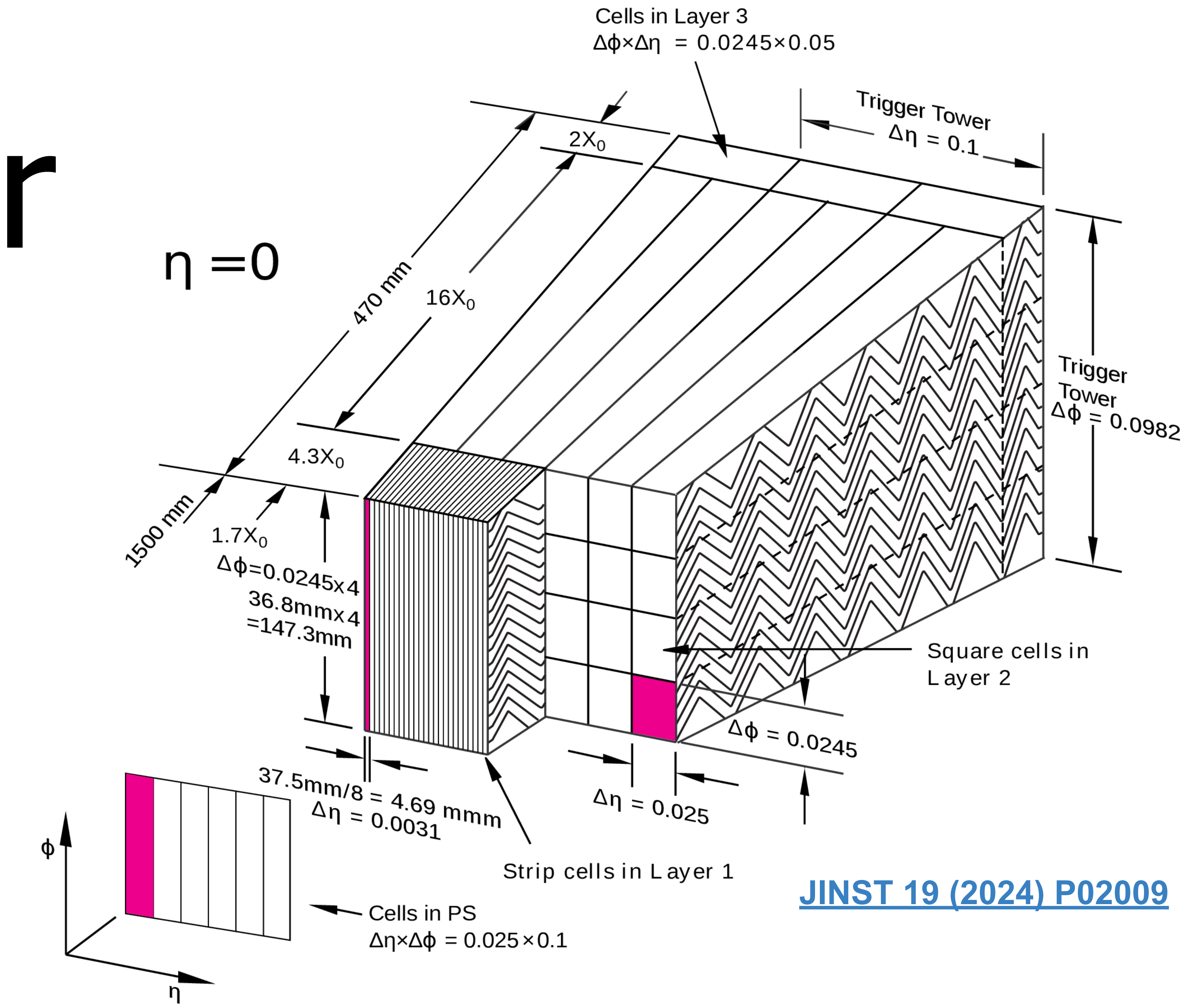


α_s from Z recoil

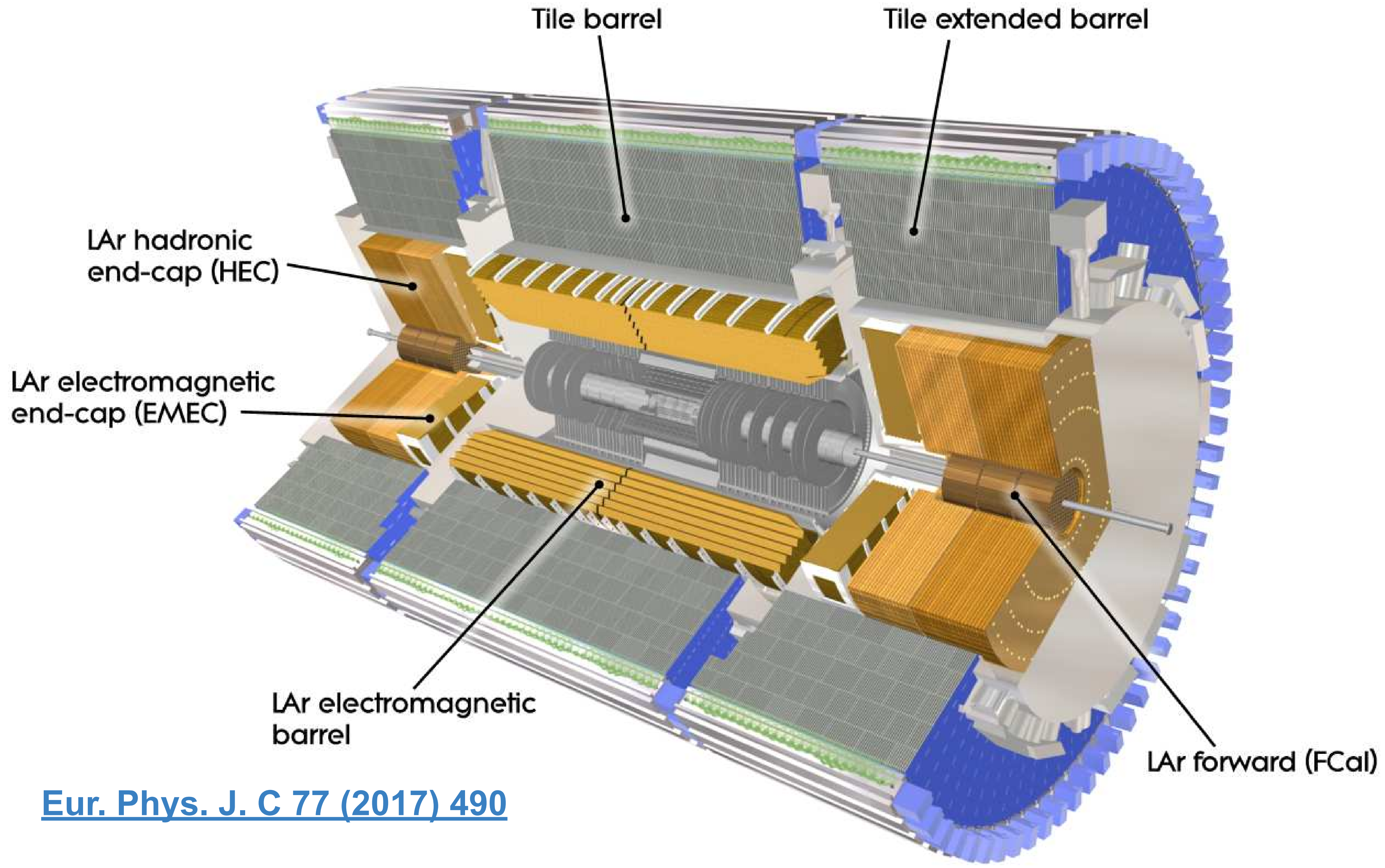


The ATLAS LAr EM Calorimeter

- Sampling calorimeter with lead absorber and liquid argon active medium. Accordion geometry with three layers of projective cells.
  - Additional **pre-sampler** (layer 0) to sample energy lost in the inner detector and passive material
  - Layer 1 - Finely segmented in η for γ/π_0 separation
  - Layer 2 - Large depth to contain most of the shower energy
  - Layer 3 - Thinner to measure the end of the shower and estimate punch-through
- Exact **cell sizes vary** in different regions of η
- Transition region of $1.37 < |\eta| < 1.52$ usually excluded from analysis



JINST 19 (2024) P02009

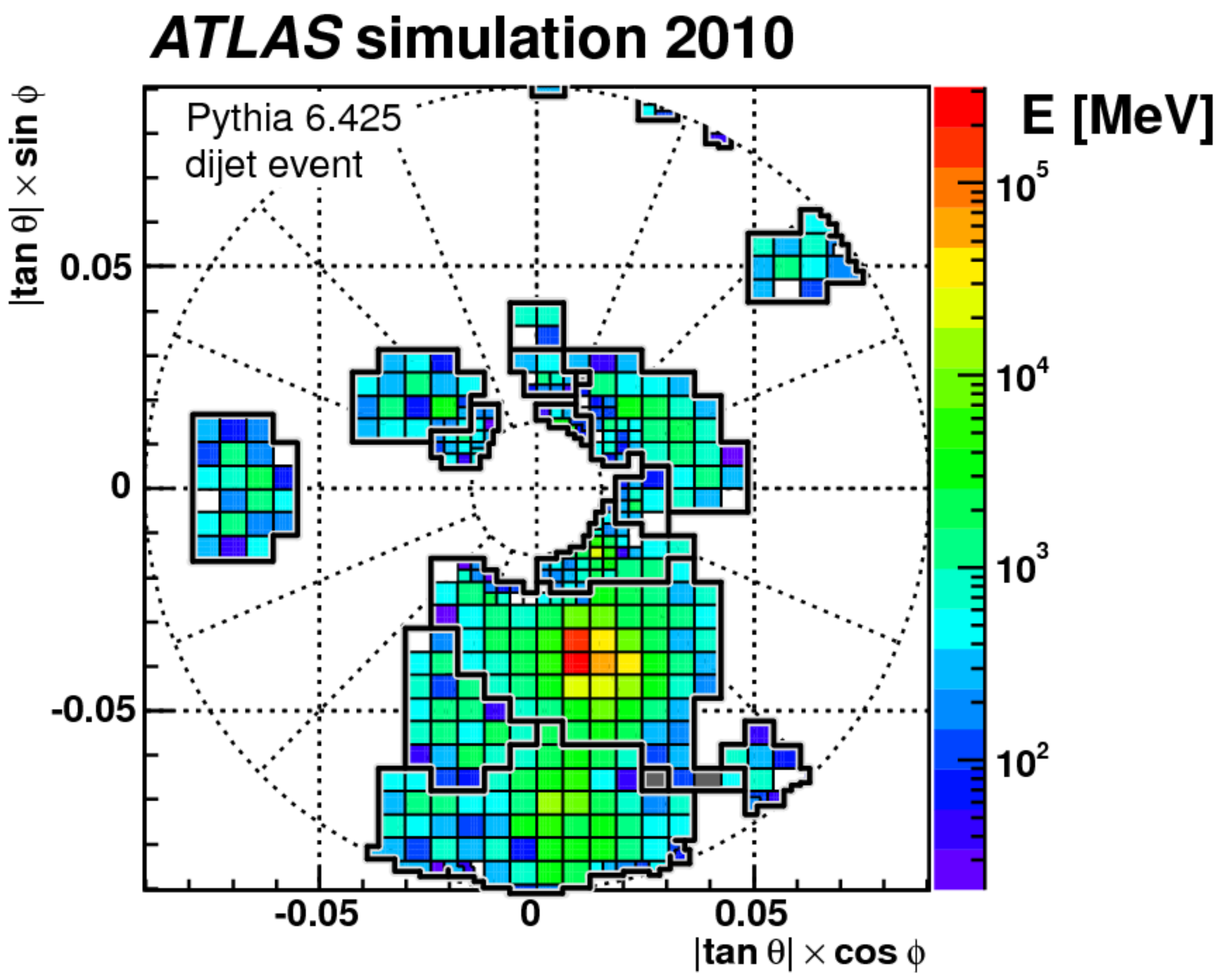


Eur. Phys. J. C 77 (2017) 490

ATLAS e/γ Calibration

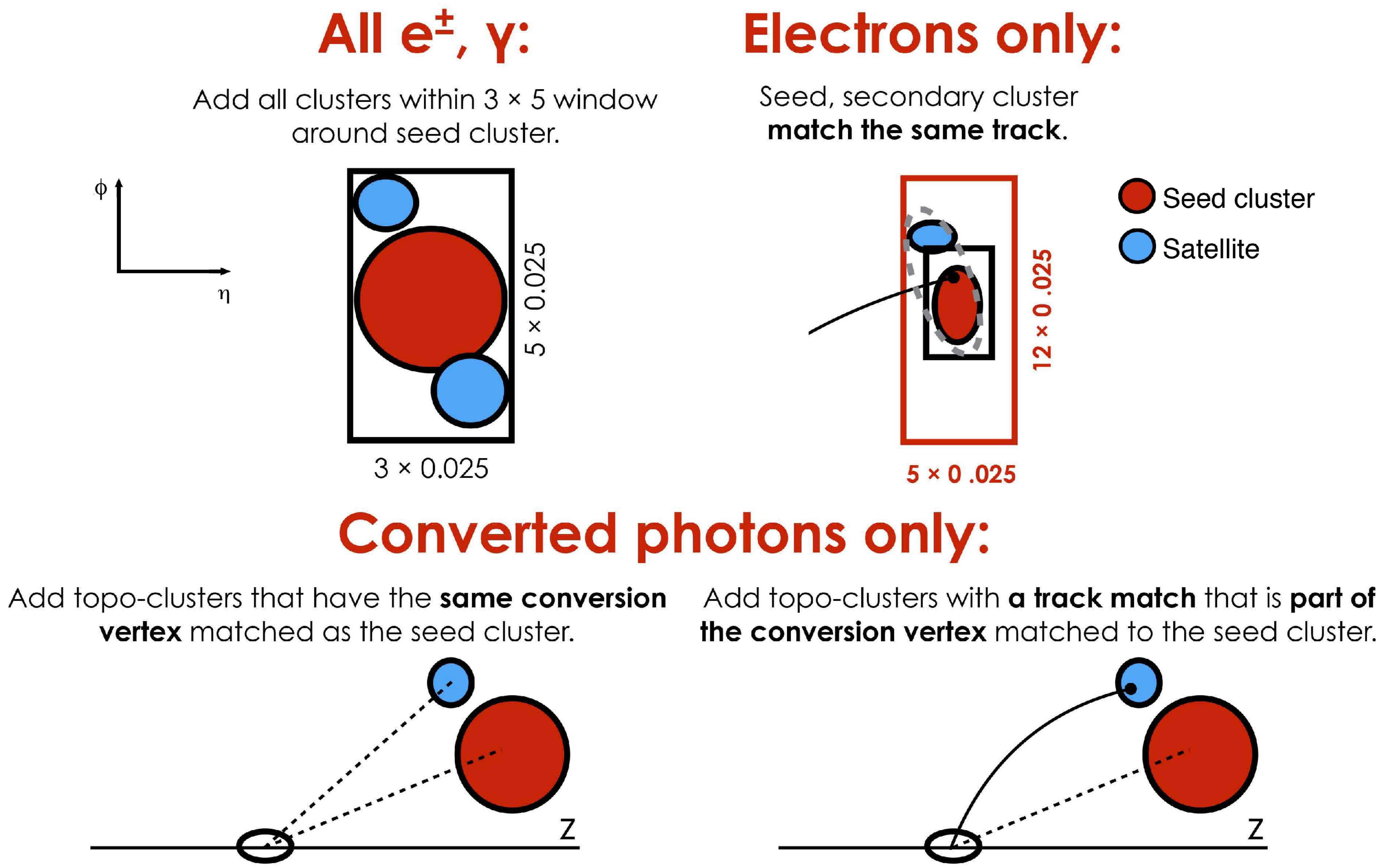
Why Calibrate?

- Topo-cluster algorithm forms connected clusters of cells with significant energy deposits
  - Nearby topo-clusters combined into super-clusters to form reconstructed $e^{\pm}$, γ.
- **Energy losses outside the cluster require a correction** to determine the true energy of the incident particle
  - Detector material upstream of the calorimeter
  - Passive material within the calorimeter
  - Punch-through to the hadronic calorimeter
- Other effects including energy deposits from in- and out-of-time pileup collisions and any effects from the clustering algorithm



Eur. Phys. J. C 77 (2017) 490

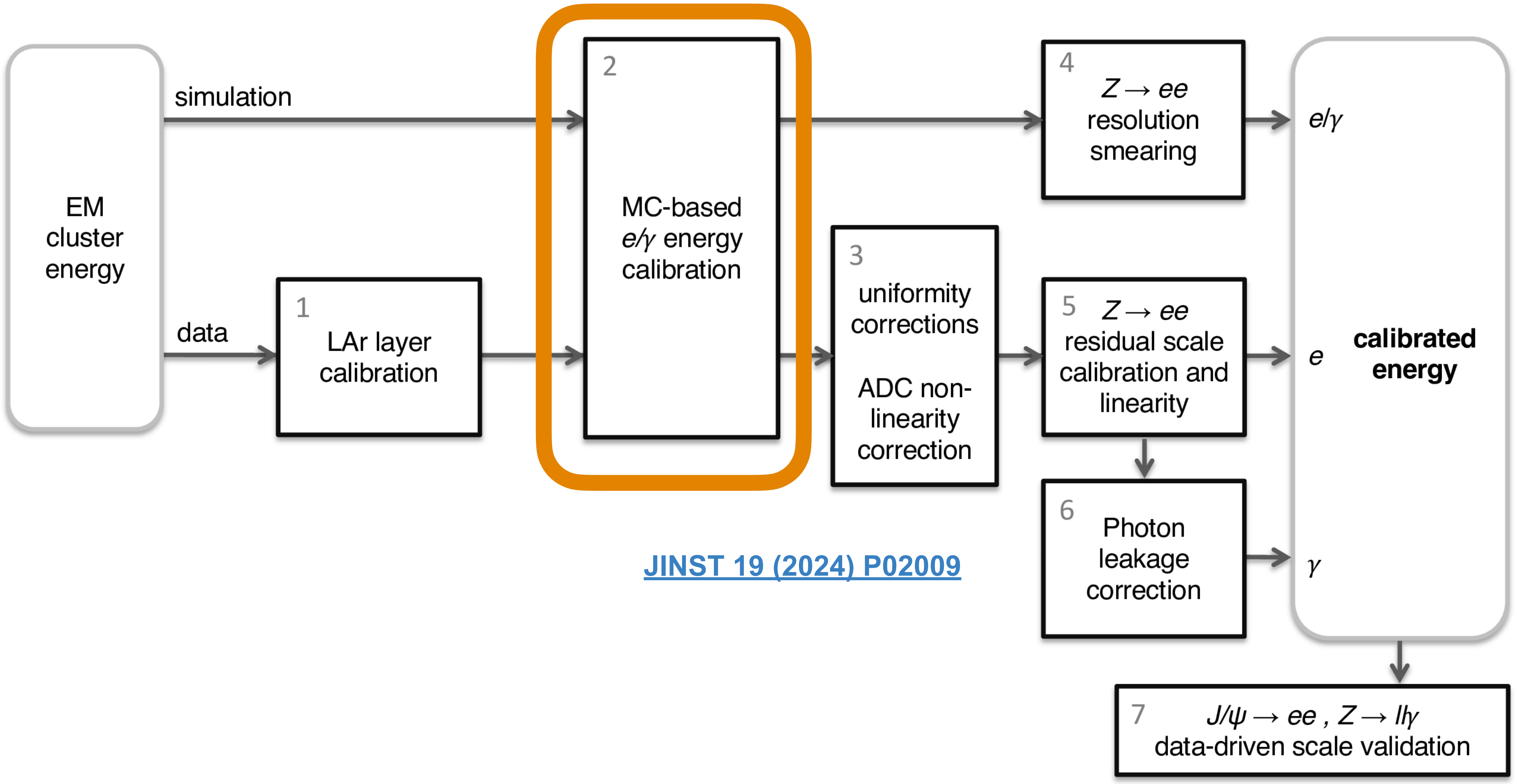
JINST 14 (2019) P12006



ATLAS e/γ Calibration

Current Method

- Calibration procedure must correct for complex detector effects and account for mis-modeling in the simulation of the EM shower and its digitization
- Today's focus: **MC calibration - simple regression from the total energy of all cells to the true energy of the simulated particle**
- Currently performed by separate BDTs for electrons, unconverted photons, and converted photons
  - Trained separately in around 100 (E, η) bins
  - Training variables derived from the total energy deposited in each layer and from the angular position of the cluster



JINST 19 (2024) P02009

GNN/Transformer Models for Calorimeters

- Many channels, semi-regular geometric pattern, typical physics signatures are compact.

- Goal is to measure total deposited energy, but the energy distribution carries useful information.

  - **Shower shape can indicate energy leakage and fluctuations from pileup collisions.**

- Many suggestions that GNNs/Transformers* are well-suited to ML tasks in calorimetry.

- Superclustering in CMS ECAL, HGCAL. High luminosity in Belle-II.

Point cloud representation of a calorimeter cluster.

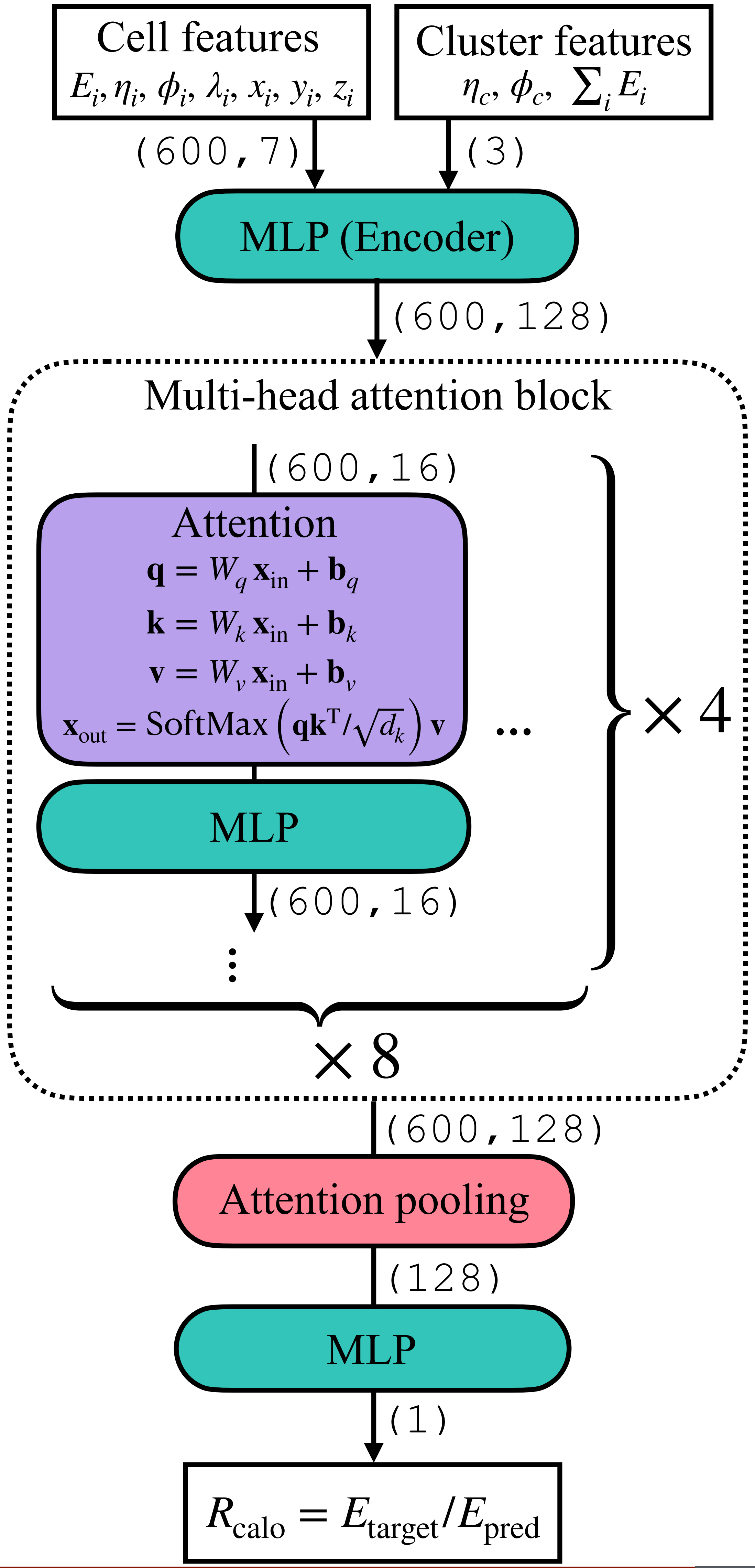
More GNNs and Transformers for Calorimeters at ML4Jets:
Clustering in CMS HGCAL
Simulation for CaloChallenge
Super Resolution and Denoising

*I will use these ~interchangeably. We are studying both, but the model I'll show today is best described as a transformer.

Transformer for e/γ Energy Calibration

Architecture, Setup, and Training

- Using the SALT framework developed for GNN based flavor tagging in ATLAS (GN1/GN2/GN2X)
- Cluster represented by **fully connected graph** formed from cells in all 3 accordion layers and the pre-sampler
- **Node features:** Cell energy, (x, y, z) coordinates, (η, ϕ) coordinates, and Cell Layer
- **Cluster features:** Cluster (E, η, ϕ)
- Extra tracking information for unconverted photons.
- Predict fractional correction to the cluster energy $E_{true}/E_{cluster}$ with MSE loss
- Separate trainings for electrons, unconverted photons, and converted photons based on single particle simulations with 2023-like pileup conditions

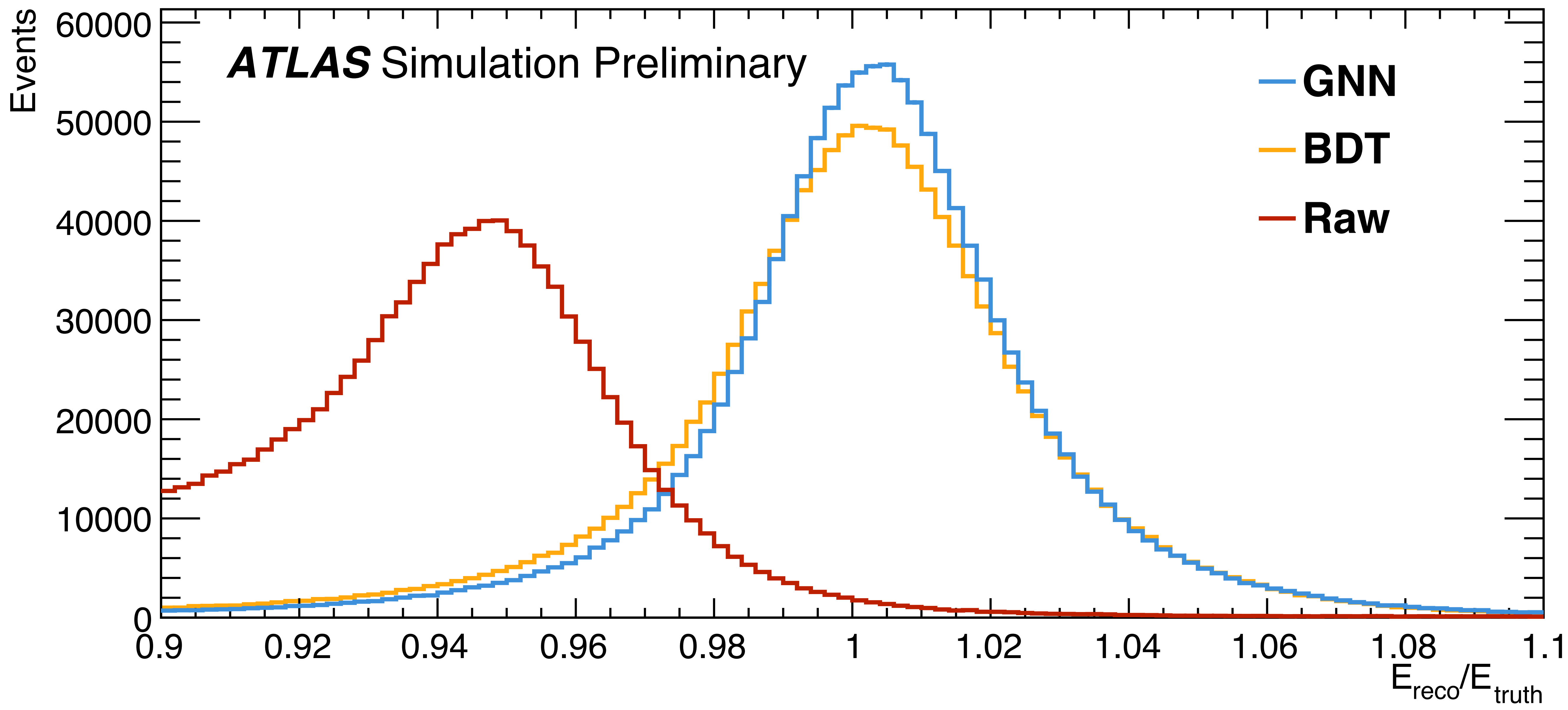


Training and Initial Performance Tests

- Each training performed on ~60M simulated clusters for 48 hours on 16 Nvidia A100 GPUs

- **BDT** and **GNN** correct the **raw** cluster energy distribution to match the true energy

- The core of the distribution is narrower for the **GNN** - improved resolution



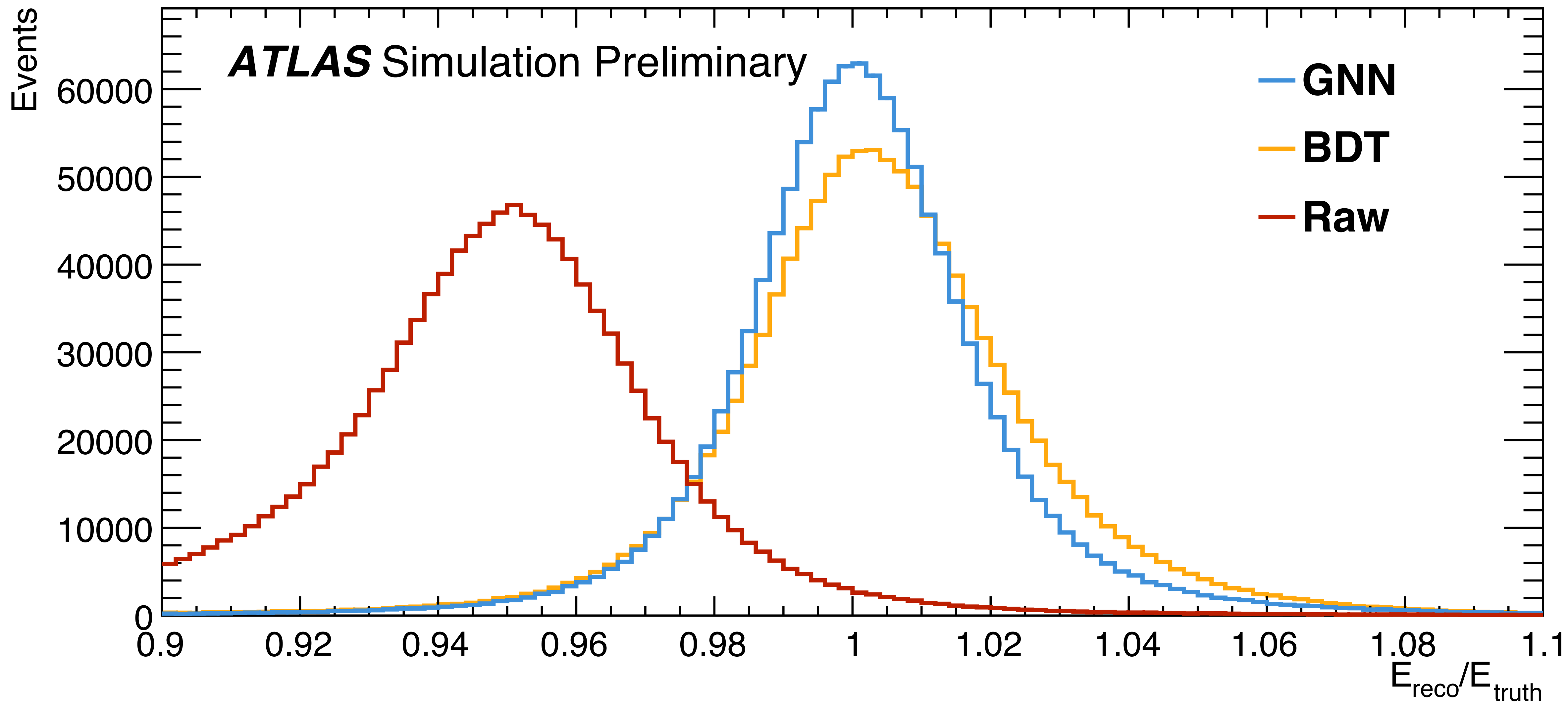
Ratio of calibrated to true energy for electrons with:

$30 < E < 100$ GeV, $|\eta| < 1.37$

Training and Initial Performance Tests

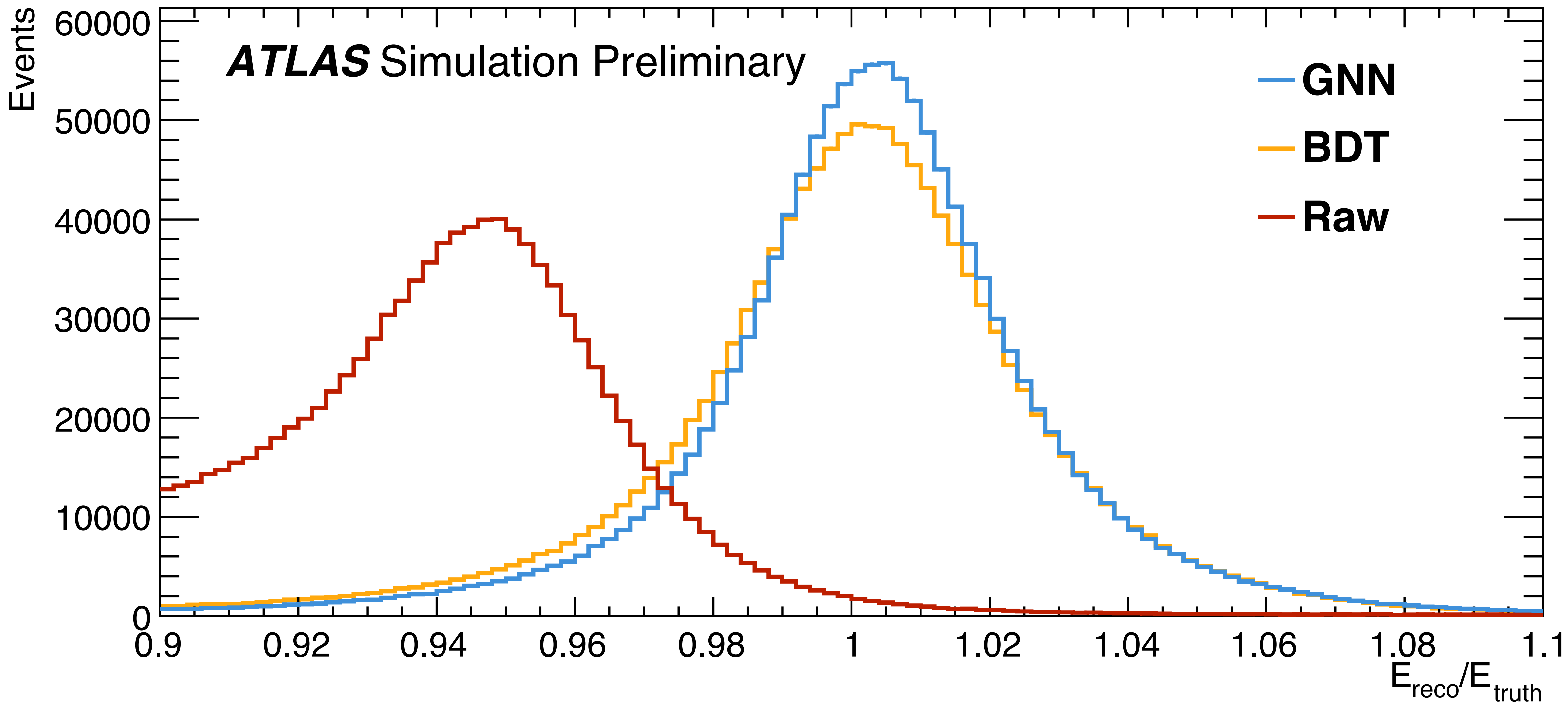
- Similar improvement for photons.
- Generally see that the performance gain for the **GNN** is less for converted photons.

Unconverted Photons

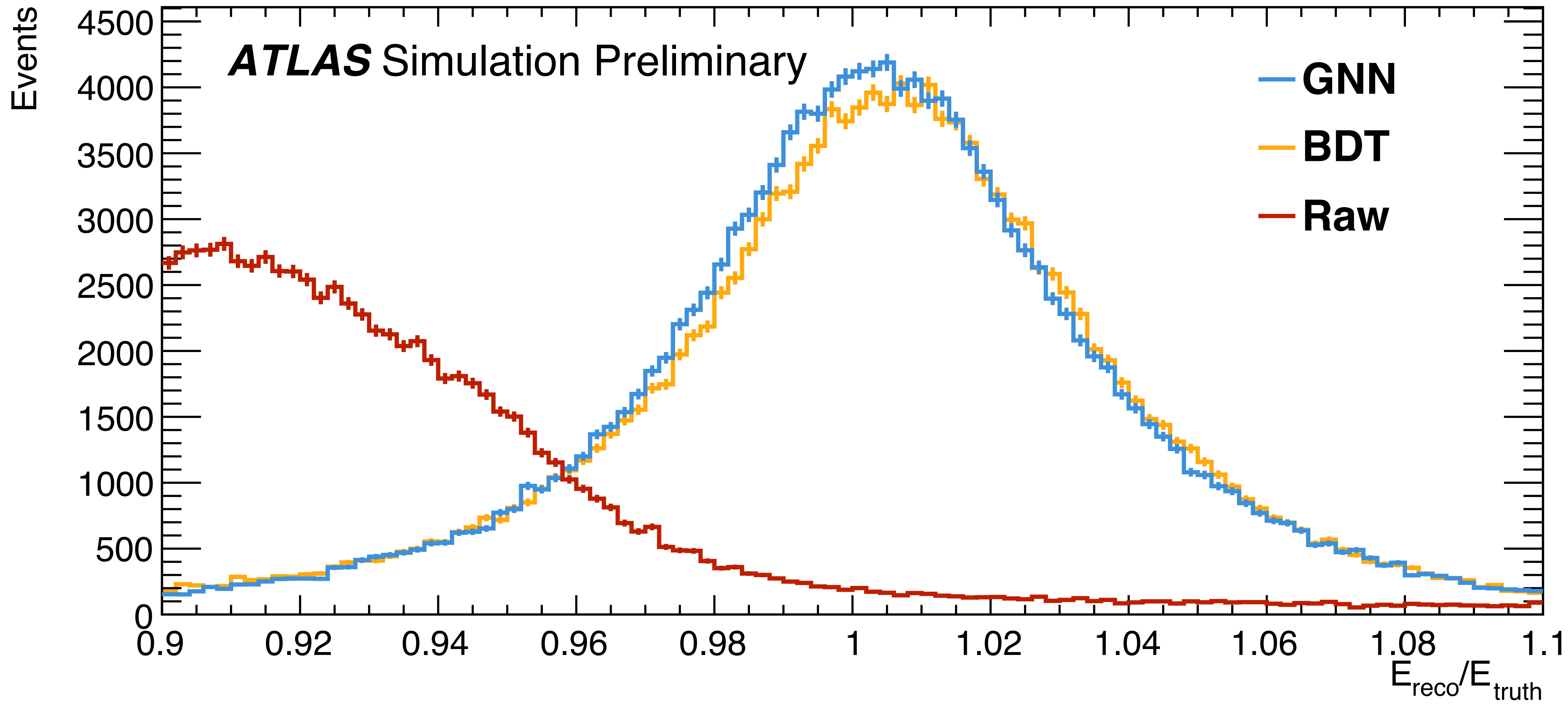


Ratio of calibrated to true energy for incident particles with $30 < E < 100\ \text{GeV},\ |\eta| < 1.37$

Electrons



Converted Photons

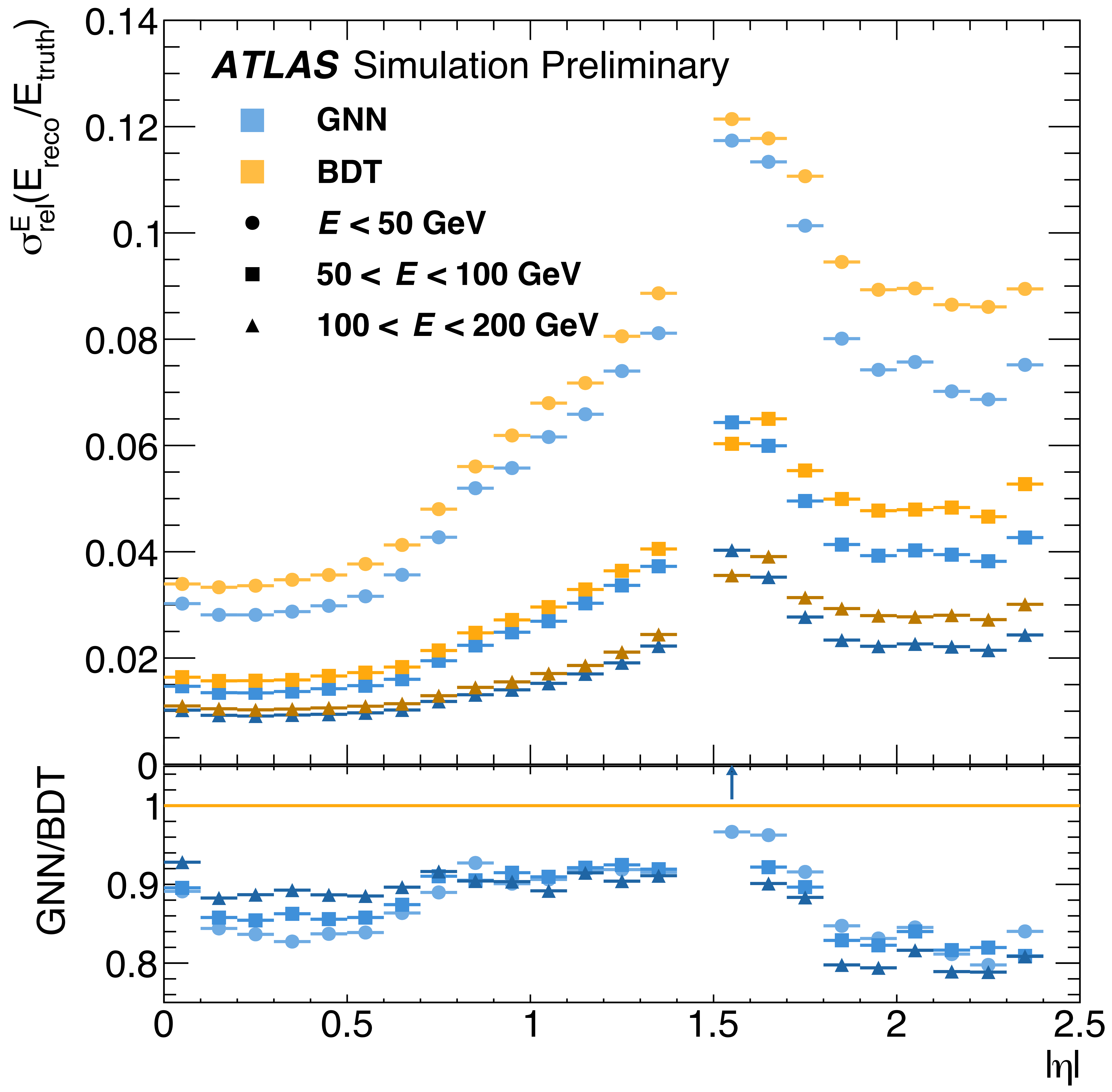


Resolution Improvements from Transformer Model

- Figure of merit is **resolution** — important to maintain performance in different detector regions (η) and energies.

- Measure the width of the $E_{\text{reco}}/E_{\text{truth}}$ peak with $\sigma_{rel}^{E} = \frac{\text{IQR}}{1.349\langle E_{\text{reco}}/E_{\text{truth}}\rangle}$ is equal to standard deviation for a normal distribution.

- **GNN** shows better resolution than **BDT** across most of the phase space.



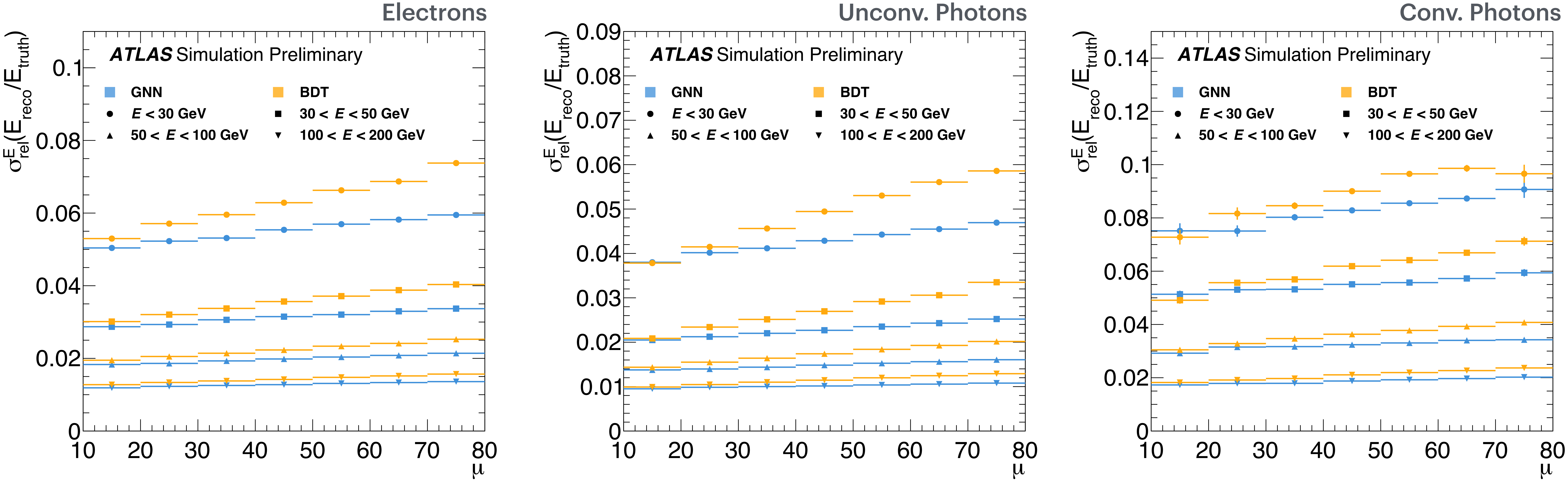
Resolution Improvements from Transformer Model

- Similar improvement for photons. Again see less improvement for converted photons, greater improvement in the endcap.

Pileup Dependence

GNN trained with Run 3-like pileup ($\langle\mu\rangle = 51$). **BDT** trained without pileup.

- Performance in high pileup conditions is crucial for both Run 3 and HL-LHC
- **GNN** calibration's energy resolution degrades less with increasing pileup



Outlook for GNN/Transformer e/ɣ Energy Calibration

- Transformer based energy calibration shows substantial resolution improvement over current methods — potential for significant impact on physics results
- **Next major challenge is to confront data/MC discrepancies** and work on the downstream calibration tasks
  - Some techniques, like layer energy corrections, not suited for GNN
  - Systematics model may need to be revisited
- Options for **tuning the architecture**, inputs, and training procedure — already early **evidence of possible performance gains**
  - Inclusion of edge features, loss function adjustments, usage of tile gap scintillators, feature importance studies
- Also working on technical details for **integration with the full ATLAS reconstruction software**

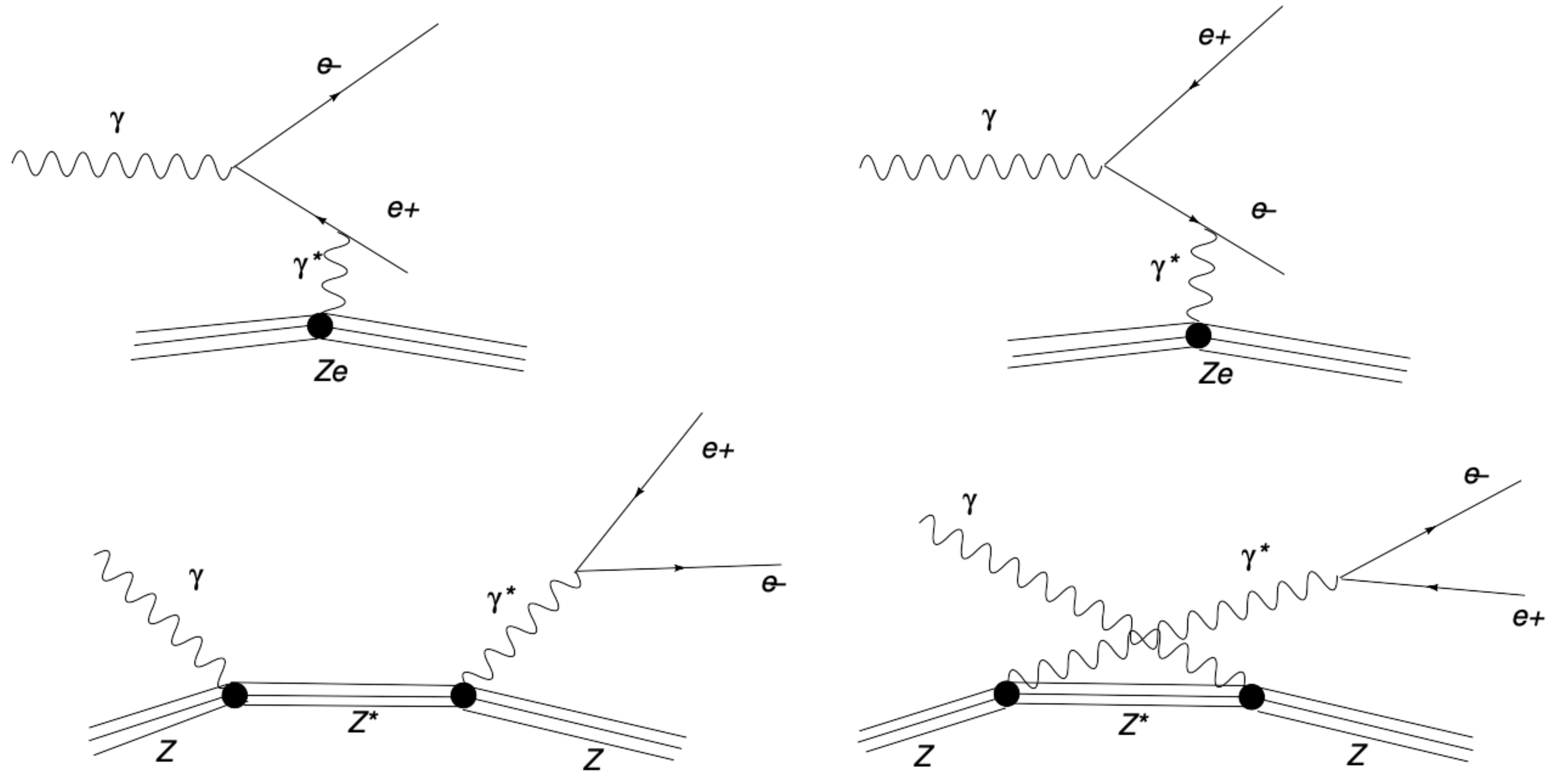
Thanks!

Backup

Transformer for e/ɣ Energy Calibration

Special Treatment of Converted Photons

- Photons which convert in the inner detector volume require special treatment to achieve reasonable performance
  - Important to preserve performance - 20-65% of photons depending on η
- BDT and GNN take use tracking inputs describing the reconstructed conversion vertex
  - Conversion radius
  - p_T ratio of conversion tracks
  - $E_T^{\text{cluster}}/p_T^{\text{vertex}}$



ATL-PHYS-PUB-2009-006