Improving ATLAS Hadronic Object Performance with ML/AI Algorithms

Albert Kong, on behalf of the ATLAS Collaboration

Classifying and Calibrating

Most hadronic showers in proton-proton collisions delivered by the LHC are primarily generated by pions, so characterising calorimeter response to pions is a critical task!

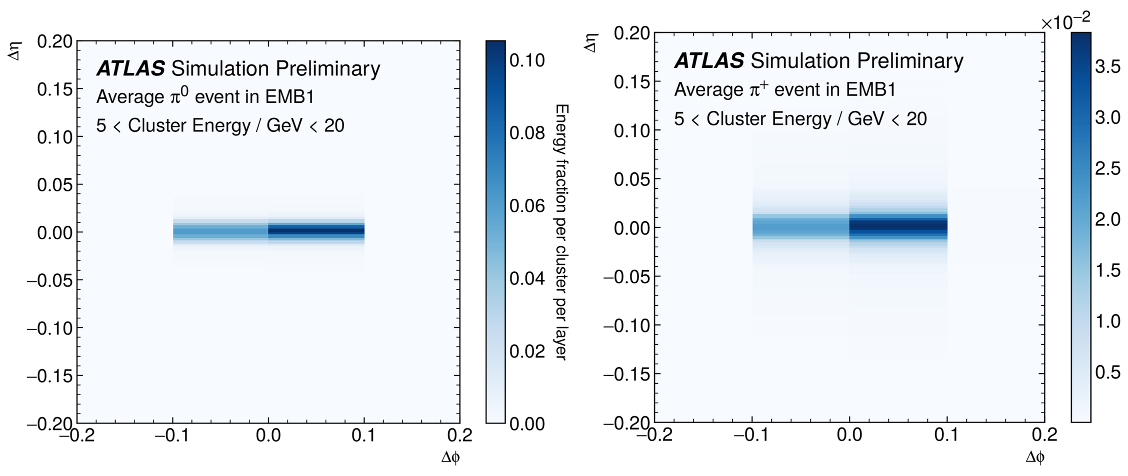

Figure 1: Average image for π^0 (left) and π^+ (right) samples in the first calorimeter layer (EMB1).

There are fundamental differences in shower shape and depth for charged and neutral pions which the current $\mathcal{P}^{\mathrm{EM}}_{\mathrm{clus}}$ classifier and Local Cell Weighting (LCW) calibration methods do not use, but machine learning methods are well suited to exploiting.

Machine Learning for Pions

We've recently explored deep learning techniques for pion classification and energy regression using both image-based datasets (figure 1) and point cloud datasets (figure 2) of simulated single pion events within the ATLAS detector.

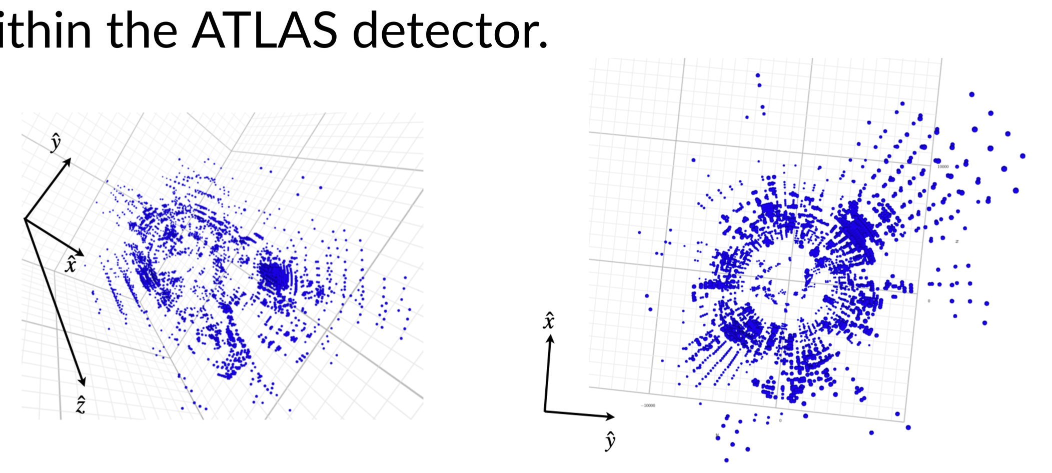

Figure 2: A dijet collision event rendered as a 3-dimensional point cloud of calorimeter cells, as seen from two orientations.

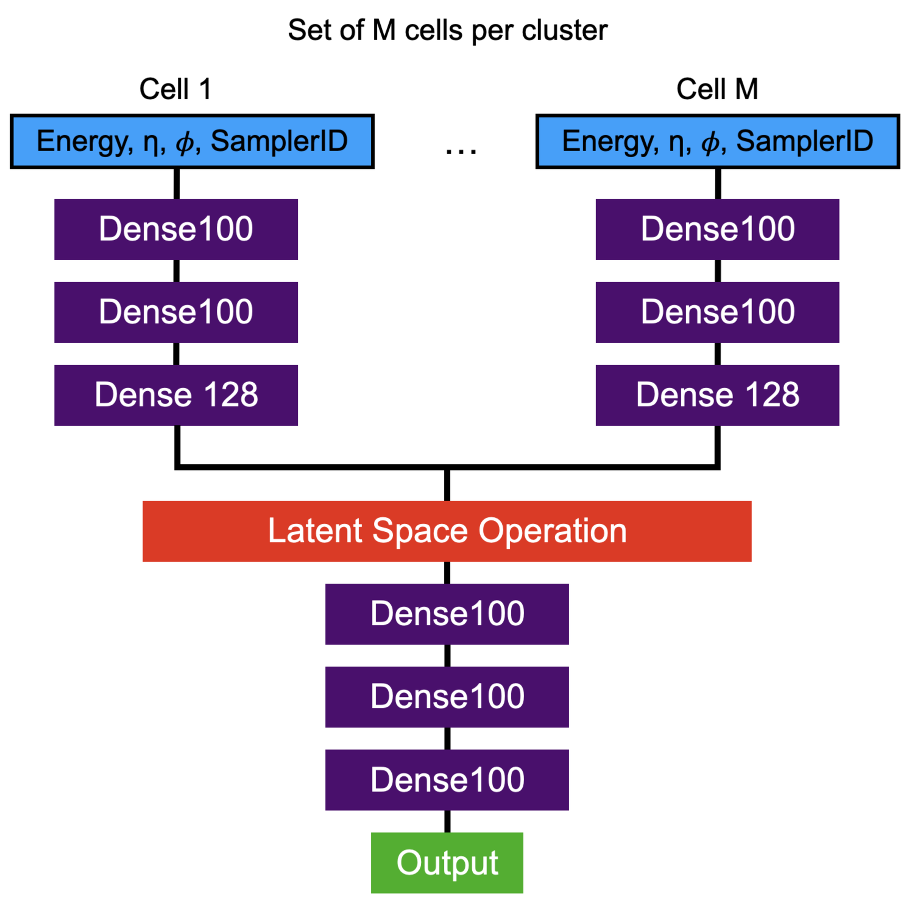

Figure 3: A diagram of the Deep Sets model architecture.

Deep Sets

Particle Flow Networks (PFNs) adapt the more general Deep Set framework to particle physics, and we used it to implement dense neural networks using the architecture seen in figure 3. The cell level features used were the (log) cell energy, sampling layer, η, and ϕ.

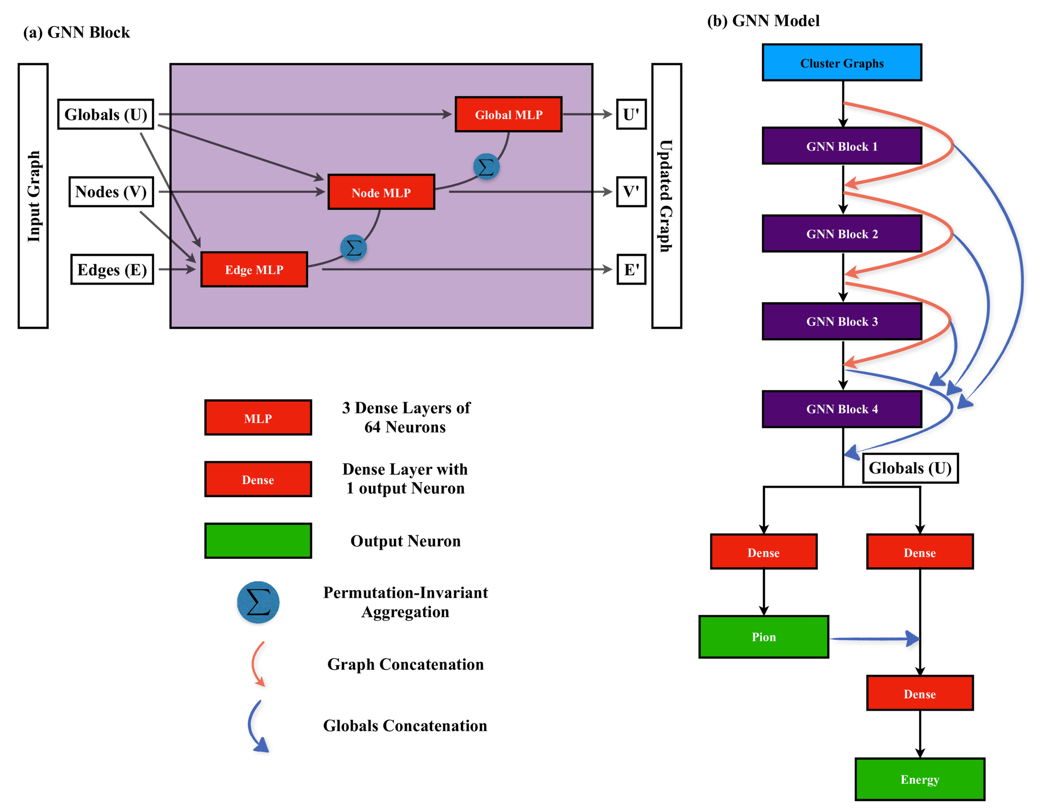

Figure 4: A schematic of the simultaneous classification and regression GNN model trained on topo-cluster information only.

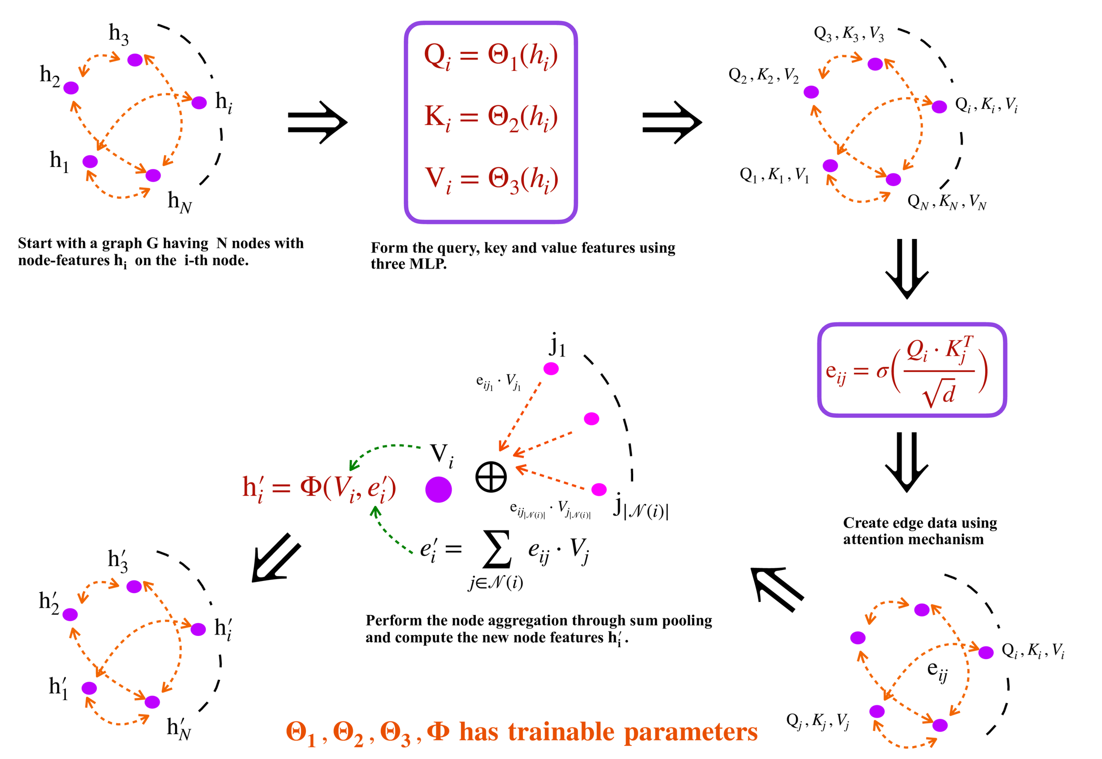

Figure 5: An overview of the transformer operation on the graph, implemented as a message passing network.

Graph Neural Networks

Our GNN was designed to perform the classification and energy regression tasks simultaneously, where each pion event is represented as a graph with calorimeter cells as nodes, edges defined by calorimeter geometry, and total topo-cluster energy as the only global feature. Node features used were cell energy, sampling layer, η, $\Delta\eta$, ϕ, $\Delta\phi$, and minimum radial distance of cell to shower axis $r_{\perp}$.

Transformer Networks

Our transformer model was implemented as a message passing layer that starts with calorimeter cells as nodes, using the same node features as with the GNN. The node features are transformed into three different latent representations, which are used to compute edges to form a graph using the self attenuation mechanism. Sum pooling is then used to compute the updated node features.

Classification Results

The following classification results were obtained using only calorimeter topo-cluster information (including cell geometry). For a fixed 90% charged pion efficiency, our best performing network (GNN) had almost 8 times better neutral pion rejection than the baseline $\mathcal{P}^{\mathrm{EM}}_{\mathrm{clus}}$ classification in the central η region, and approximately 5 times better rejection when considering the full η range.

Model	Rej. @ 90% Eff. for $\lvert\eta\rvert < 0.7$	Rej. @ 90% Eff. for $\lvert\eta\rvert < 3$
CNN	26.584	-
GNN	46.419	20.500
Deep Sets	24.814	7.608
$\mathcal{P}^{\mathrm{EM}}_{\mathrm{clus}}$	6.123	3.977

Table 1: Neutral pion rejection at a fixed charged pion efficiency of 90% for the various classification models.

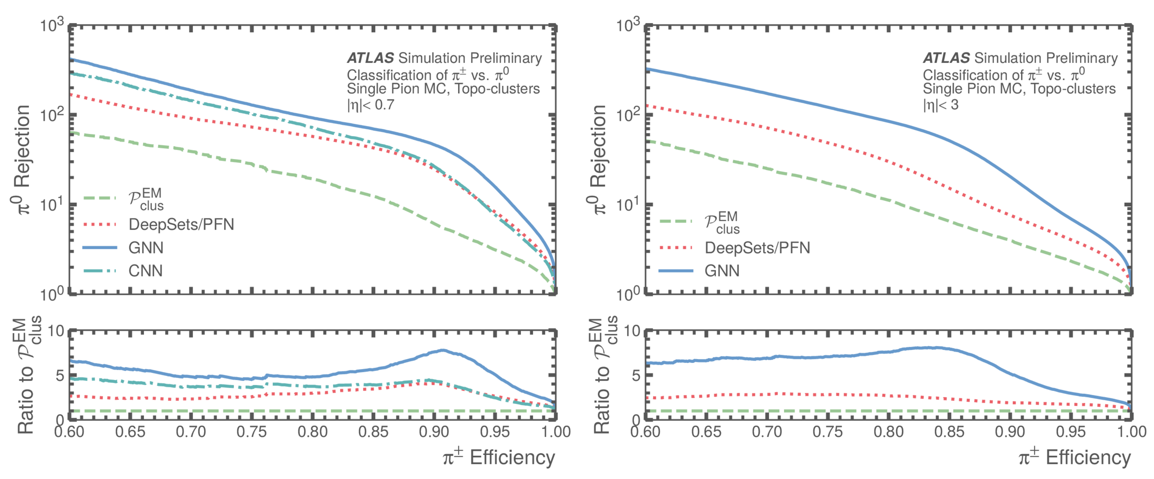

Figure 6: Comparison of classifier performance of all methods for $|\eta| < 0.7$ (left) and $|\eta| < 3$ (right). Higher values of rejection indicate improved classification performance for the same selection efficiency.

Regression Results

The following energy regression results were obtained using both calorimeter and tracking information. All models studied here significantly outperformed the baseline EM and LCW calibrations across the entire energy spectrum in terms of both median response (figure 7) and resolution (figure 8), in particular providing improved resolution over track-only results.

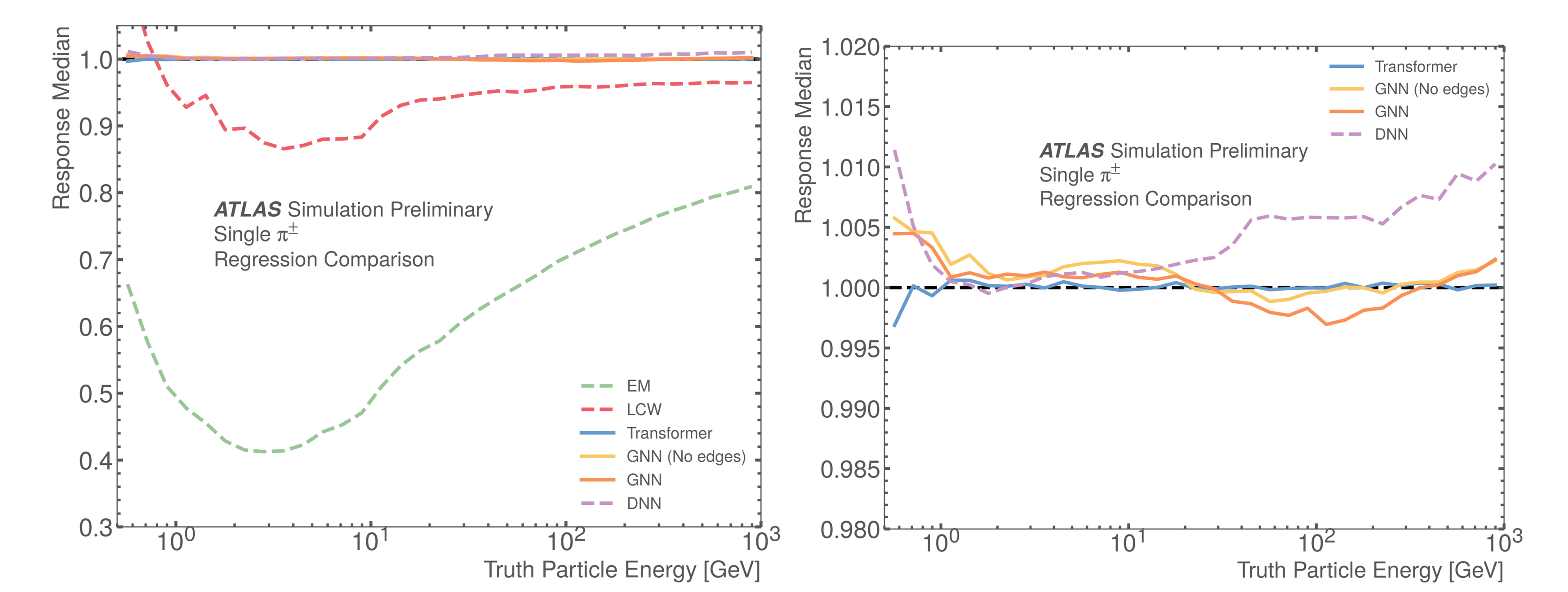

Figure 7: Comparison of mean ratios of predicted energy to truth pion energy for multiple charged pion energy regression methods, both including the baseline methods (left) and without (right). An ideal response is a median ratio of 1 across the entire energy range.

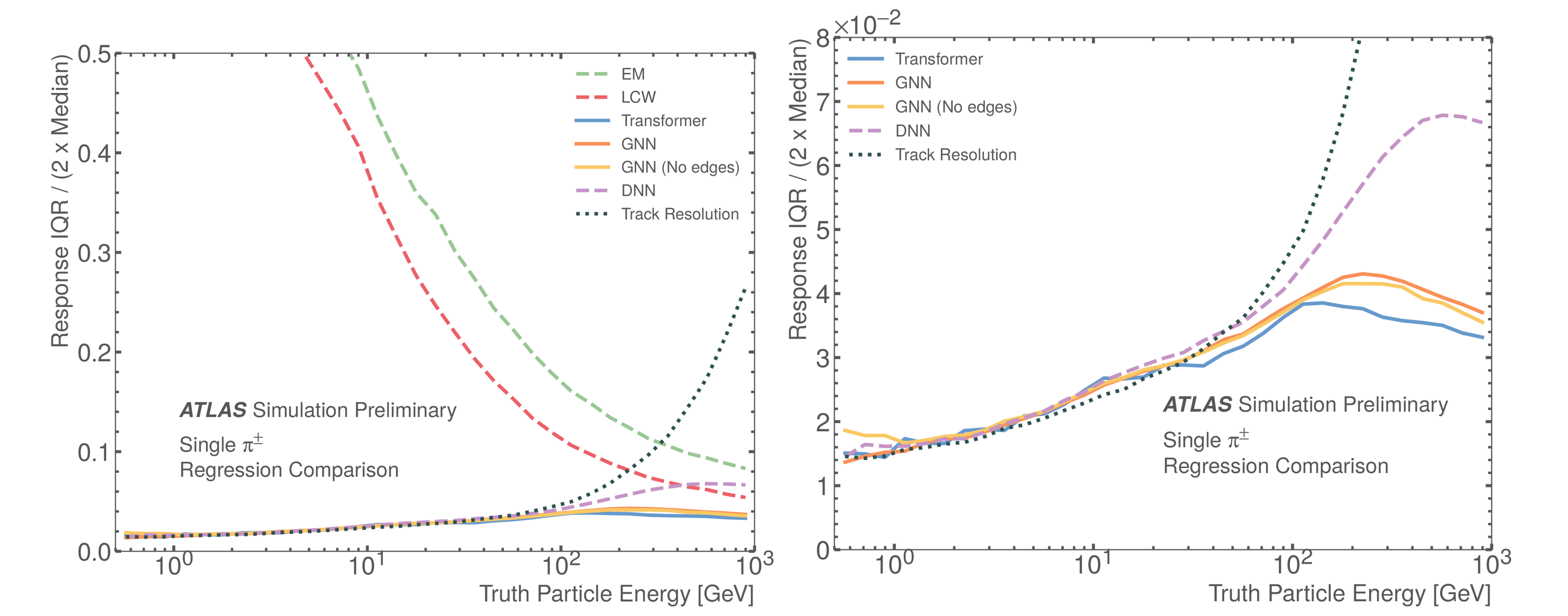

Figure 8: Comparison of half interquartile range (IQR) divided by median predicted energy for multiple charged pion energy regression methods, both including the baseline methods (left) and without (right). This can be seen as a measure of energy resolution, with lower values being better.

THE UNIVERSITY of ADELAIDE

ATLAS EXPERIMENT