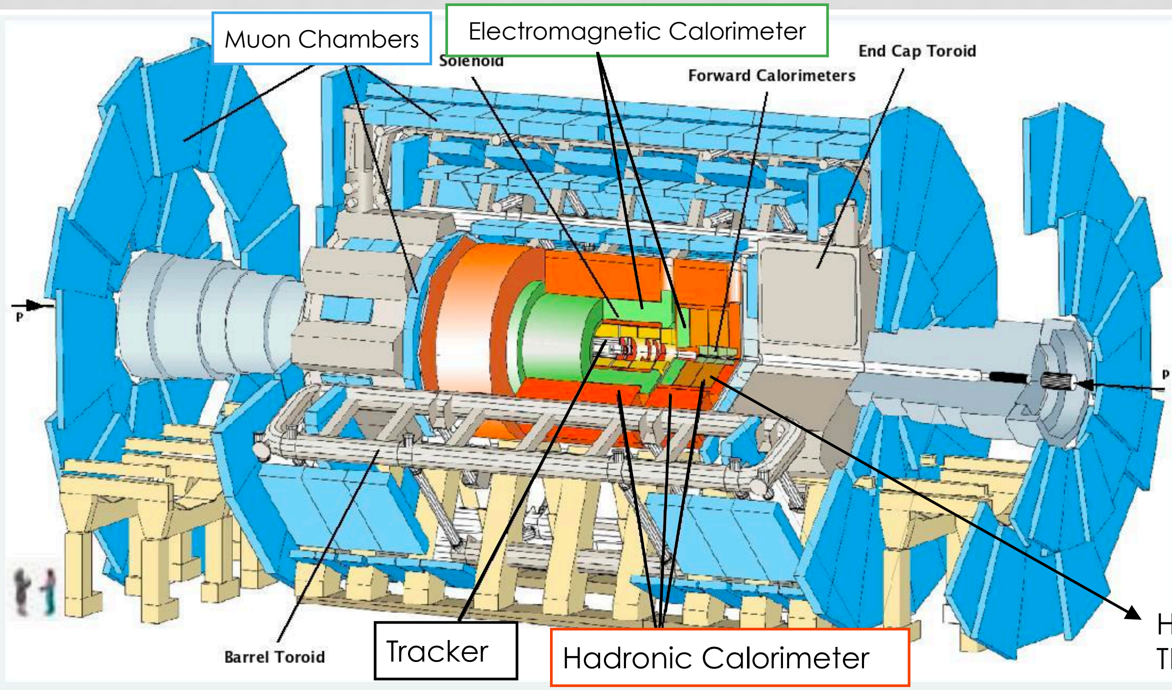
DEEP LEARNING METHODS FOR HADRONIC RECONSTRUCTION WITH THE ATLAS DETECTOR

DILIA MARÍA PORTILLO,

ALISON LISTER,MAX SWIATLOWSKI, WOJTEK FEDORKO, RUSSELL BATE

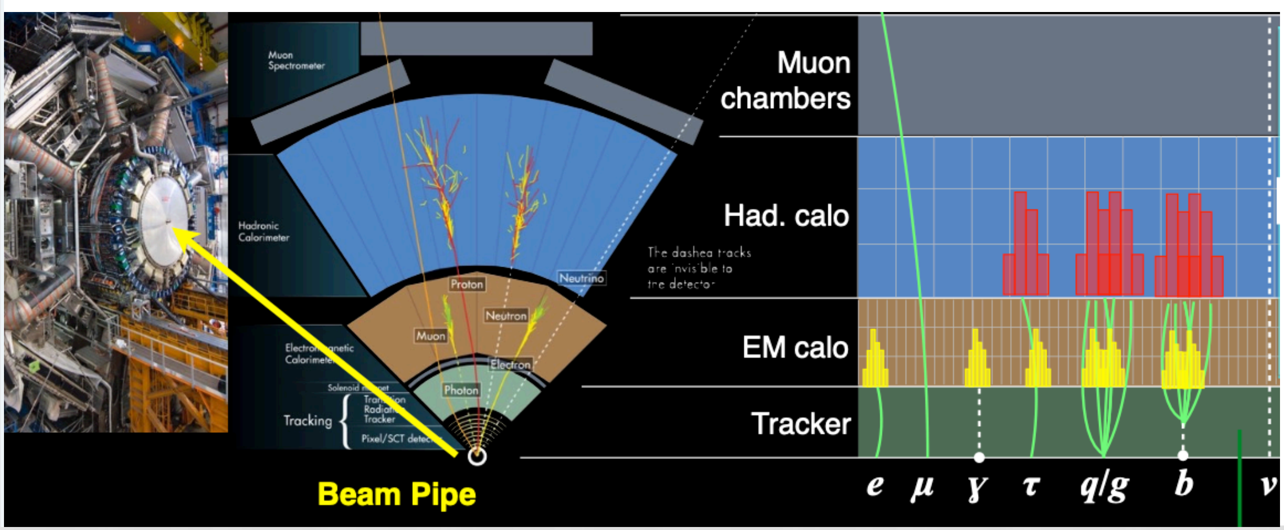
18-07-2022
TRIUMF SCIENCE WEEK 2022

The ATLAS detector



Hadronic endcap built by TRIUMF

- Multi-purpose detector
- Optimised for proton-proton interactions
- Onion-shell-like structure and covers almost the full 4π solid angle

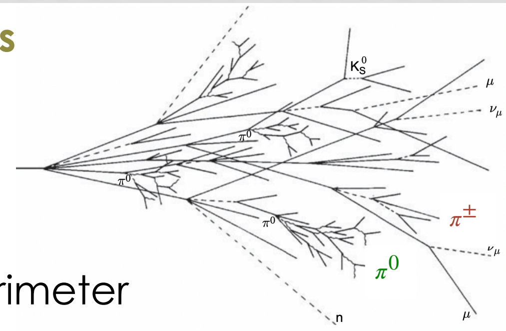


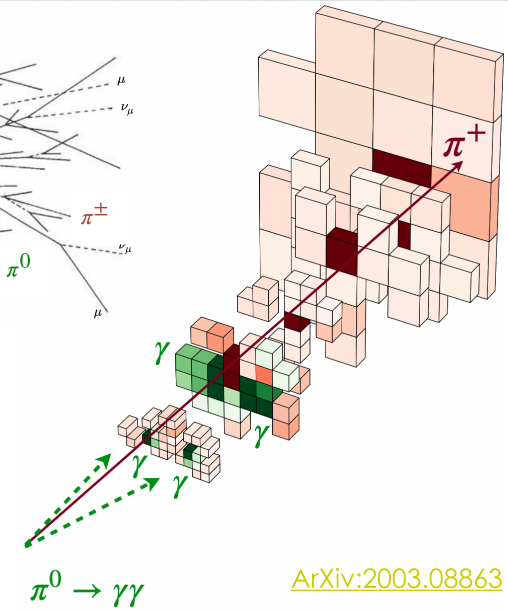
Calorimeters

- Each read-out unit of the calorimeter defines a cell
- Contain energy/ location information
- Each shower deposits energy in many cells

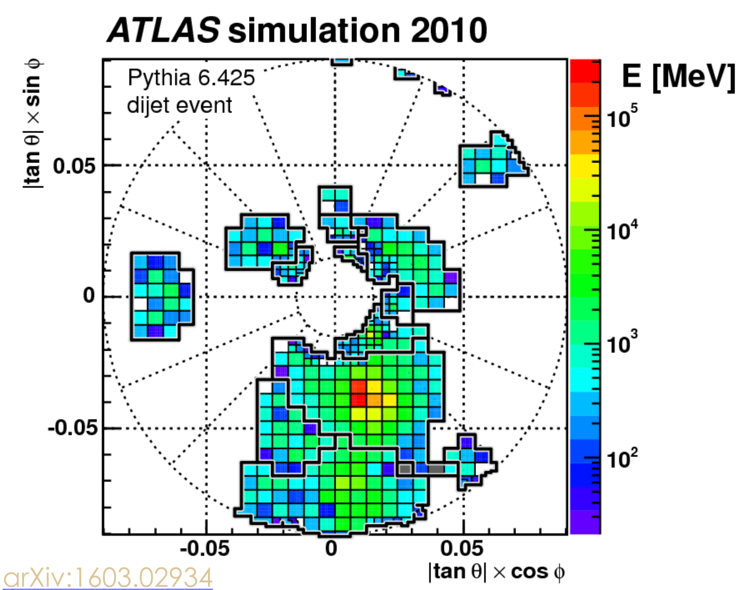
Hadronic reconstruction in ATLAS

- Hadronic showers are mostly composed of pions
  - **Neutral Pions** π^0:
    - Quickly decay to photons
    - Compact showers
    - Captured by the **electromagnetic** calorimeter
  - **Charged Pions** $\pi^{\pm}$:
    - Irregular showers
    - Require the dense material in the **hadronic** calorimeter to be stopped





ArXiv:2003.08863



arXiv:1603.02934

Topo-clusters

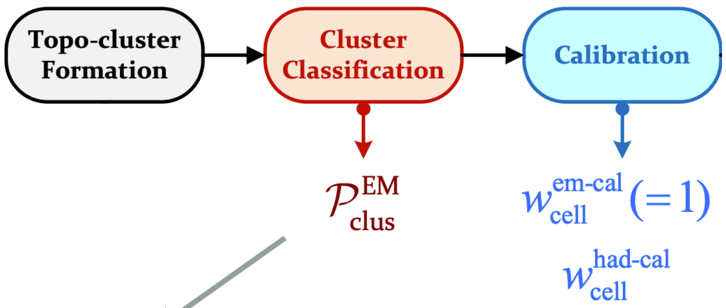
- Baseline hadronic reconstruction in ATLAS
- Uses clusters of calorimeter cells
  - 3D clusters of noise-suppressed calorimeter cells

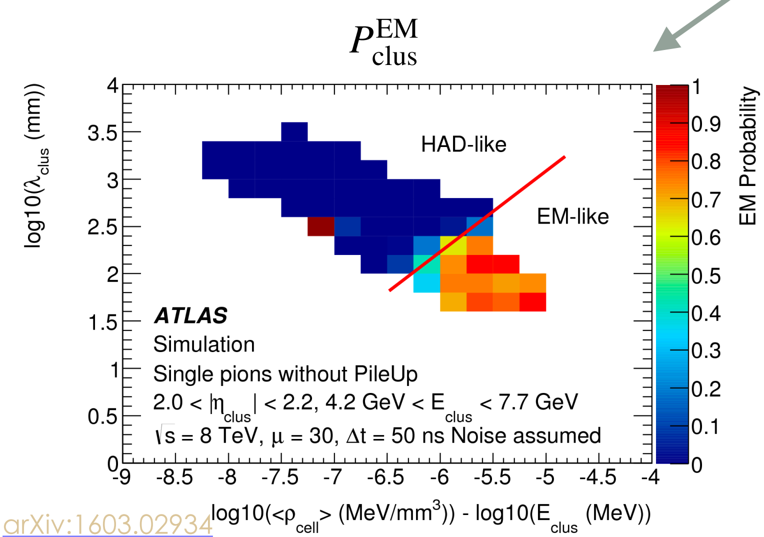
Hadronic calibration in ATLAS

- Topo-clusters needs to be calibrated:
  - Different detector response and measurement for π^0 vs. $\pi^{\pm}$ showers
- **Topo-cluster calibration:**

Local Cell Weighting (LCW)

1. **Classify** as **electromagnetic** or **hadronic** calculating the EM probability $P^{\mathrm{EM}}_{\mathrm{clus}}$





Hadronic calibration in ATLAS

- Topoclusters needs to be calibrated:
  - Different detector response and measurement for π^0 vs. $\pi^\pm$ showers
- **Topo-cluster calibration:**

Local Cell re-Weighting (LCW)

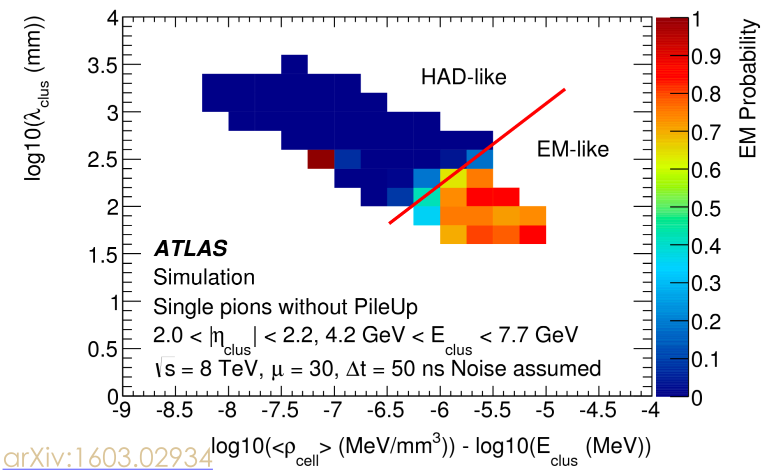
1. **Classify** as **electromagnetic** or **hadronic** calculating the EM probability $P^{\text{EM}}_{\text{clus}}$
2. **Calibrate** its energy to account for differences in response.

$\mathcal{P}^{\text{EM}}_{\text{clus}}$

$w^{\text{em-cal}}_{\text{cell}} (=1)$

$w^{\text{had-cal}}_{\text{cell}}$

$P^{\text{EM}}_{\text{clus}}$



$$E^{\text{LCW}}_{\text{clus}} = \sum_{i \in \text{cluster}} w^{\text{LCW}}_{\text{cell},i} \; E^{\text{EM}}_{\text{cell},i}$$

Each cell in the cluster is weighted

arXiv:1603.02934

CAN WE USE DEEP LEARNING TO IMPROVE THESE TECHNIQUES?

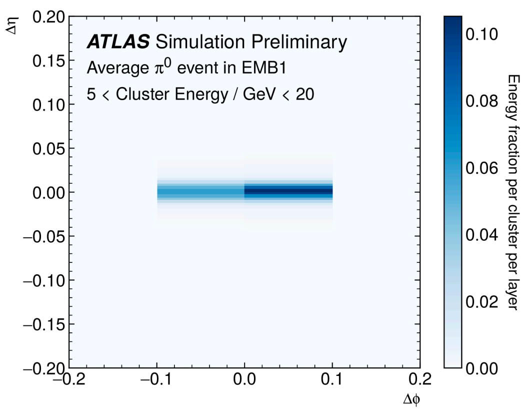
Topoclusters as images

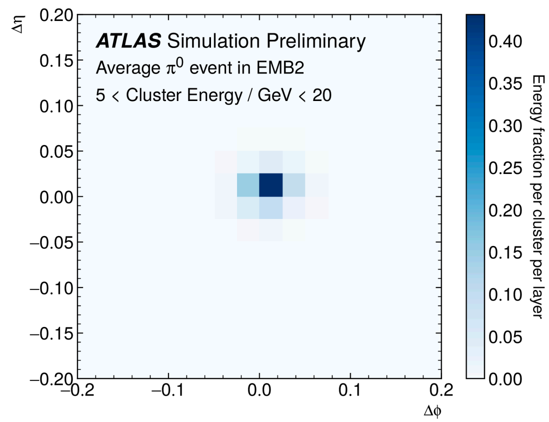
ATL-PHYS-PUB-2020-018

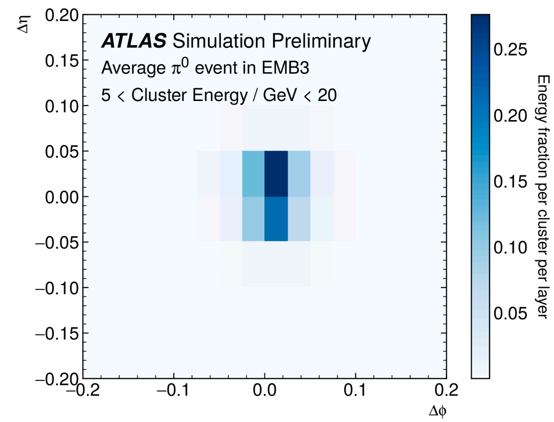
- Represent each **cluster as a pixelated image** per calorimeter layer using the appropriate cell granularity.
- Neural Networks trained using single-particle Monte Carlo simulations.

Calorimeter layer 1 | Calorimeter layer 2 | Calorimeter layer 3

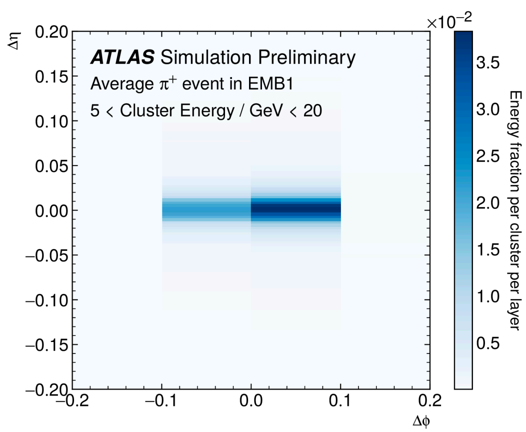
π^0

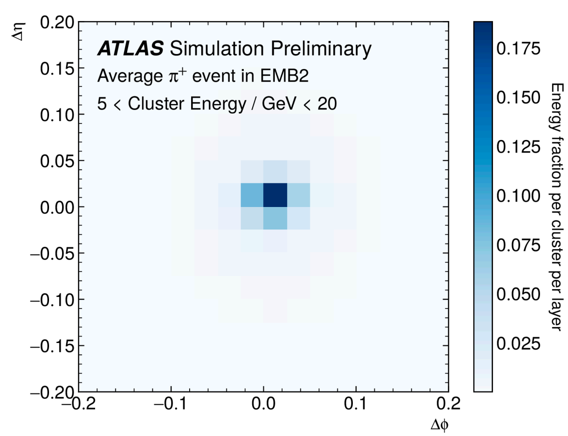


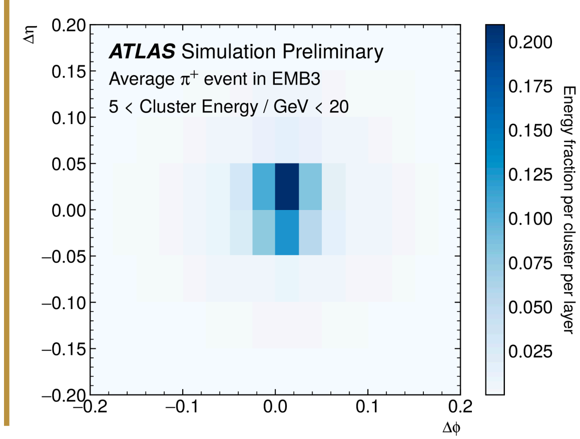




π^+





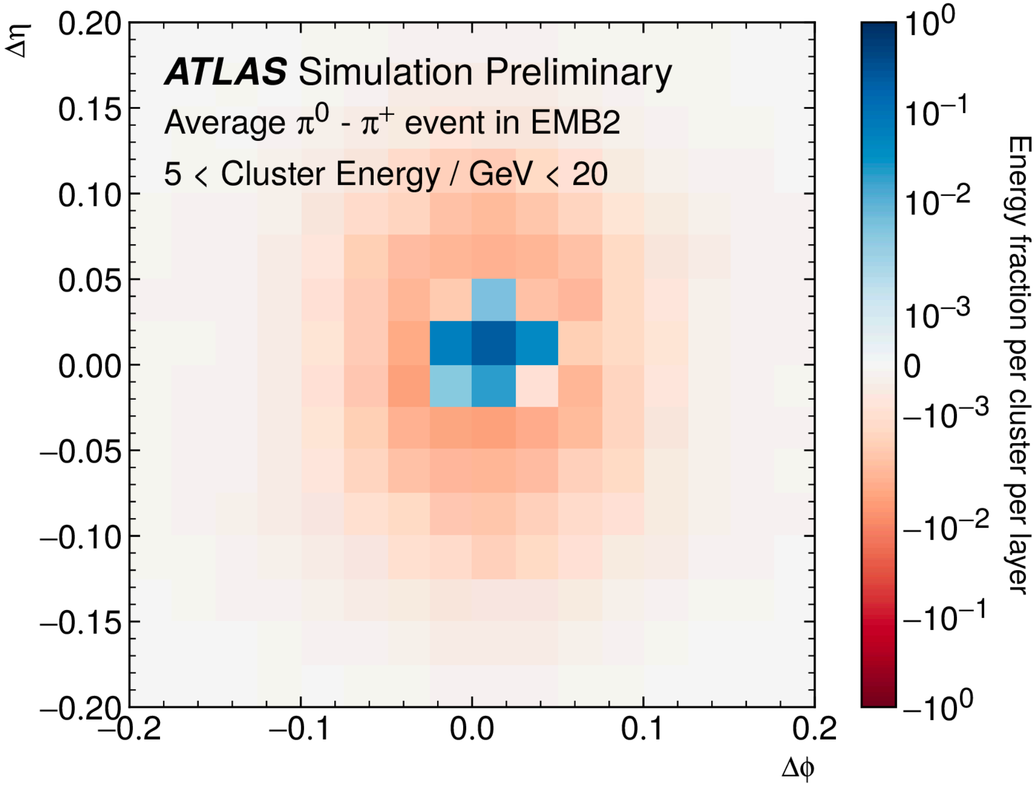


Topoclusters as images

ATL-PHYS-PUB-2020-018

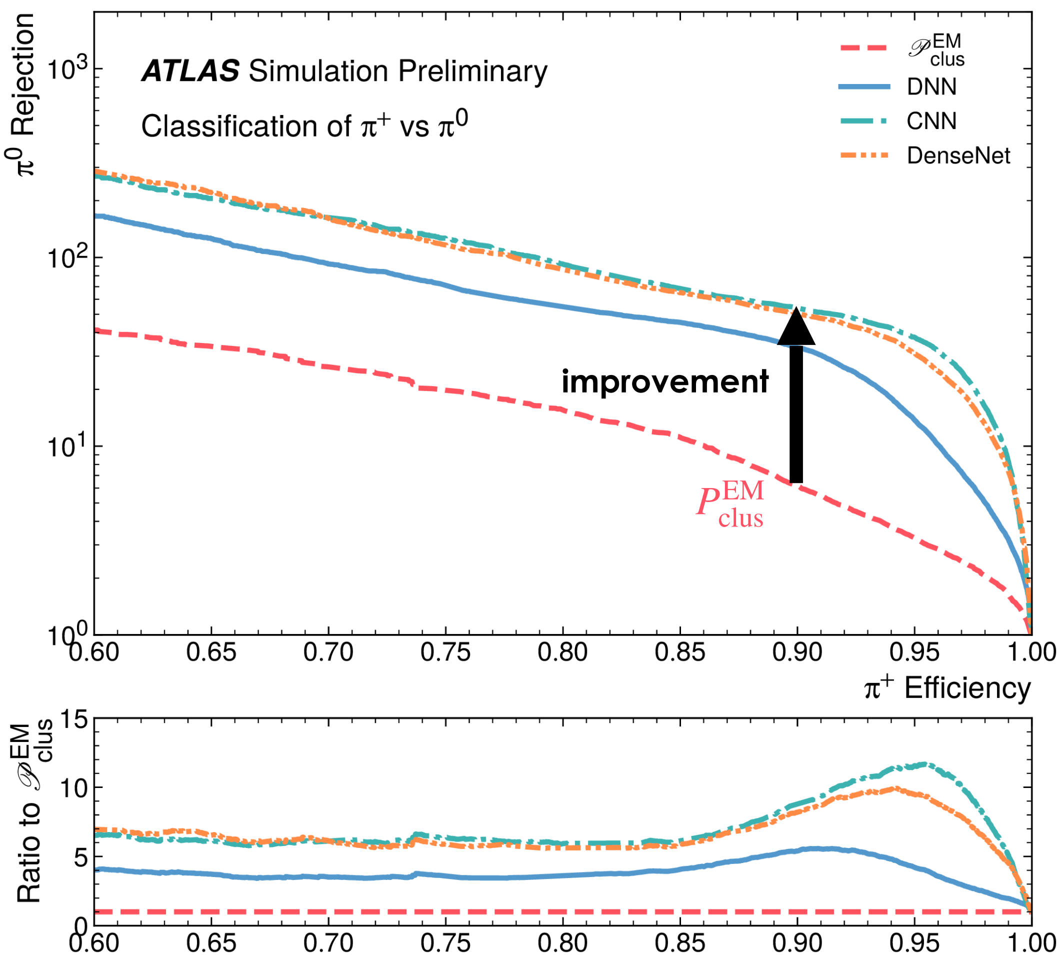
- Represent each **cluster as a pixelated image** per calorimeter layer using the appropriate cell granularity.
- Neural Networks trained using single-particle Monte Carlo simulations.

$\pi^0 - \pi^+$



Topocluster images: Pion Classification

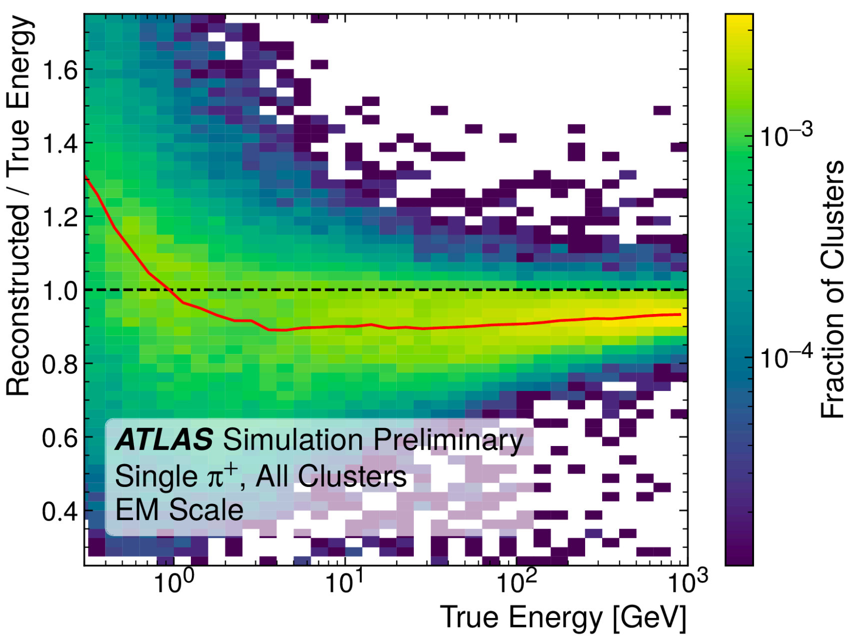
- Machine Learning techniques all do an excellent job of **distinguishing π^0 from $\pi^{\pm}$ showers**
  - Dramatic **improvements** compared to the current classification method using P^{EM}_{clus}



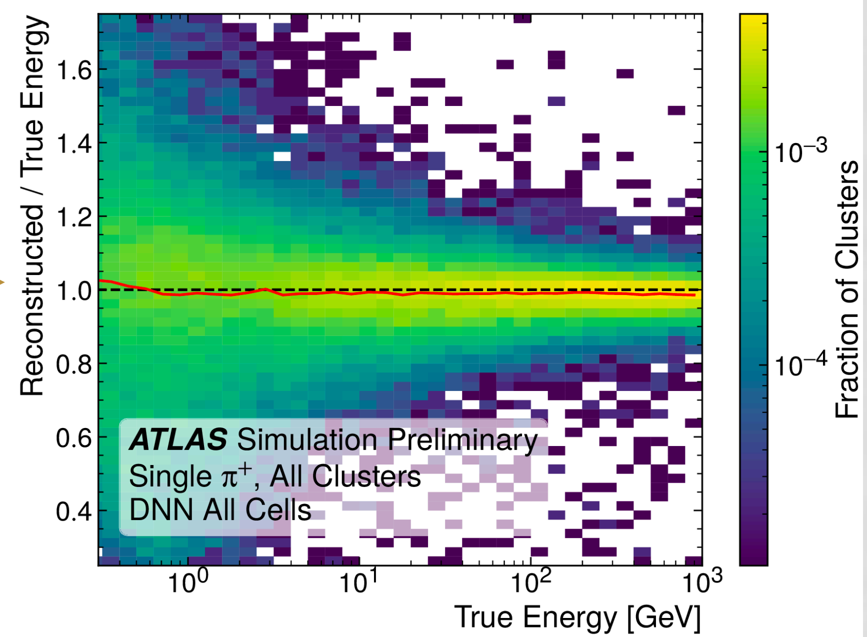
Topocluster images: Pion Energy Calibration

- After classifying a cluster, need to calibrate its energy
- **Energy calibration goal**: Correctly **predict the true energy** deposited in the cluster.
  - Quantified by measuring the cluster **energy response**: $R = \frac{E^{\mathrm{predicted}}}{E^{\mathrm{truth}}}$ that should be ~ 1

Regression performance for charged pions



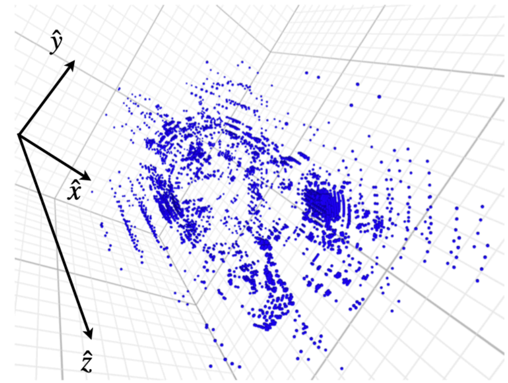
Before Calibration



After calibration

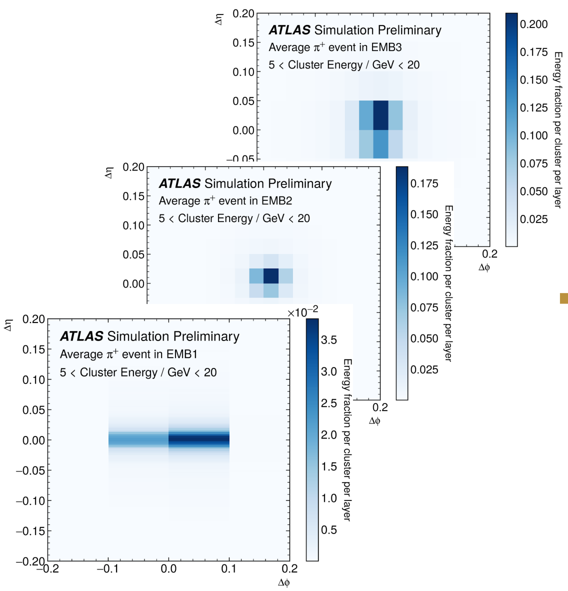
Topoclusters as point clouds

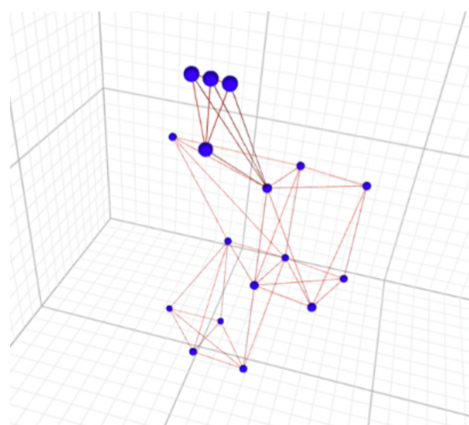
- **Point-clouds**: set of data points in space
- Use point clouds representation of clusters:
    - Each point in the set have features (E, η, ϕ, Calo layer) per cell



Advantages with respect image-based approach

- More **natural representation** of the 3D structure of calorimeter topo-clusters than a series of images
- More **flexible** as an input structure: Allows for the incorporation of track information. Doesn't require workarounds for the different layer geometries/granularities.



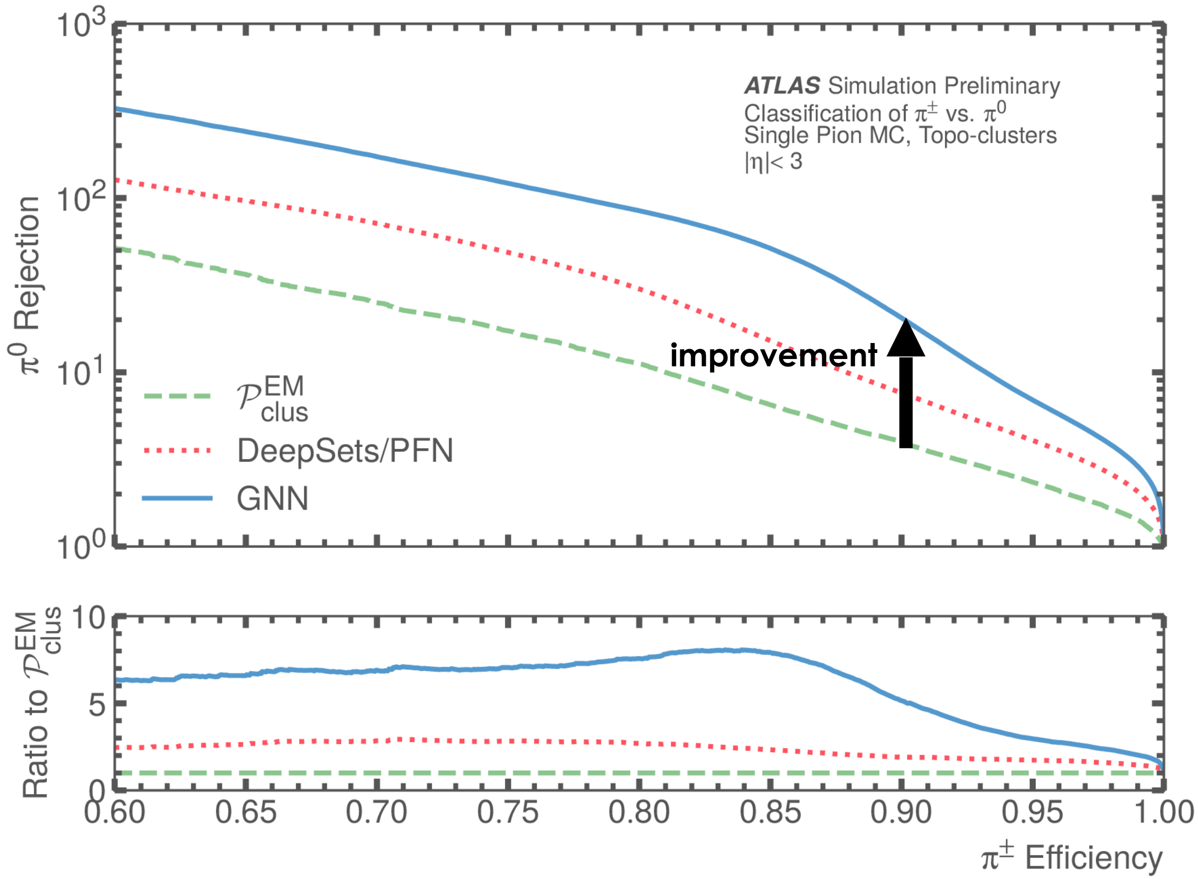


Graph Neural Network (GNN)

- Represent each pion topo-cluster as a graph
    - Nodes = individual cluster cell features
    - Edges = cell geometry information
    - Global feature = cluster energy

Topoclusters point clouds: Pion classification

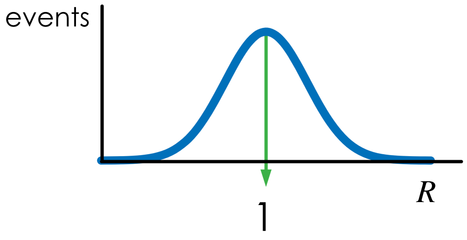
- New **point cloud** approaches (GNN & PFN) far **outperform** the baseline EM cluster probability ($P^{\mathrm{EM}}_{\mathrm{clus}}$)
- They also perform **on par with or better** than the **image-based** CNN approach for pion classification

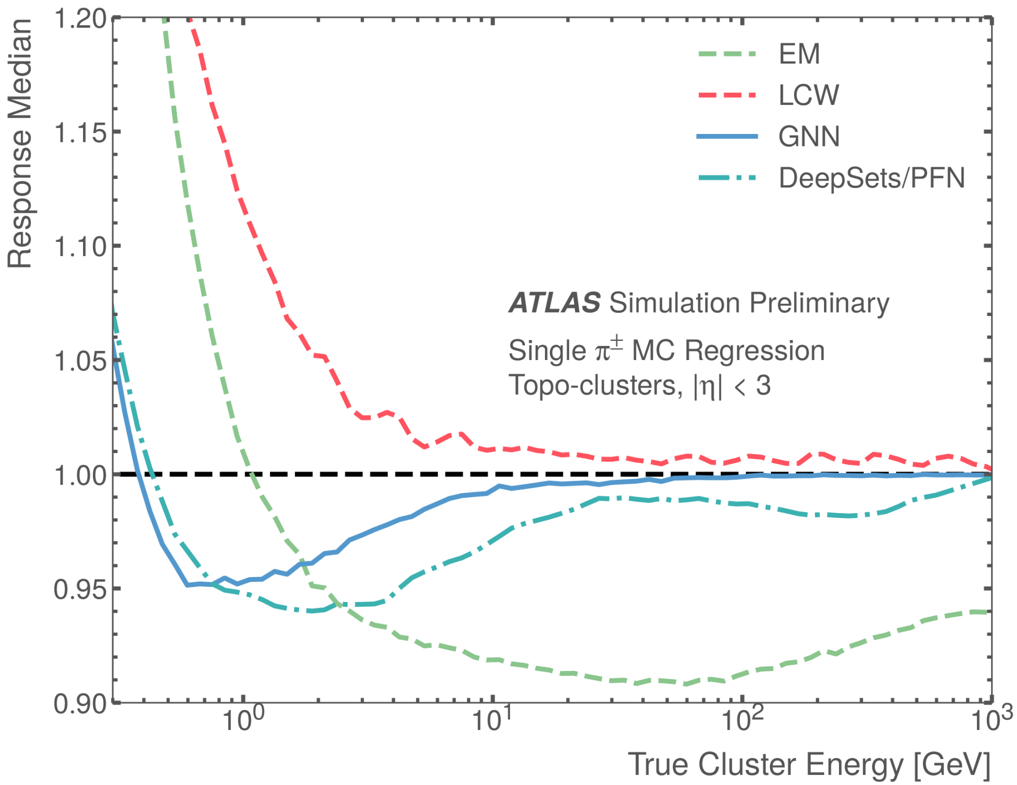


Topoclusters point clouds: Pion Energy Regression

Energy Response

After calibration: Median of the response $R = \dfrac{E^{\text{predicted}}}{E^{\text{truth}}}$ should be ~1



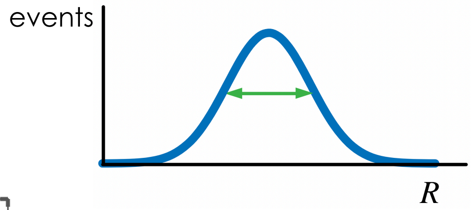


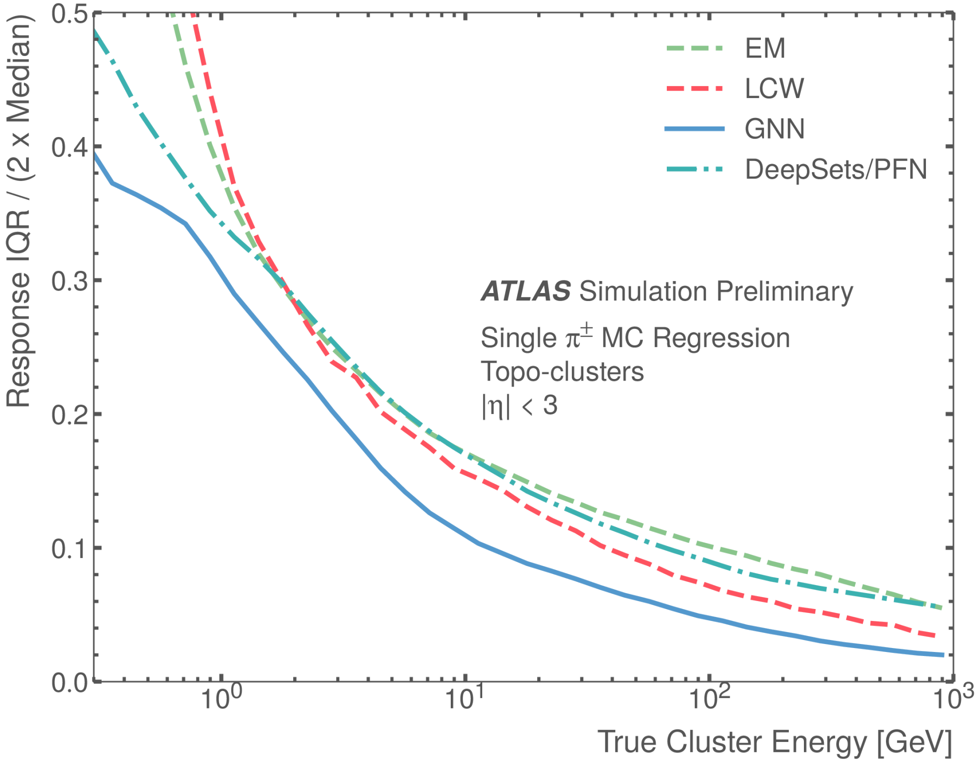
- **GNN** & **PFN** are closer to one than the **EM** scale (raw cluster energy) and outperform **LCW** calibration for low-energy clusters.

Topoclusters point clouds: Pion Energy Regression

Energy Resolution

After calibration: **Spread of the response** $R = \dfrac{E^{\text{predicted}}}{E^{\text{truth}}}$ around the media value to be as small as possible

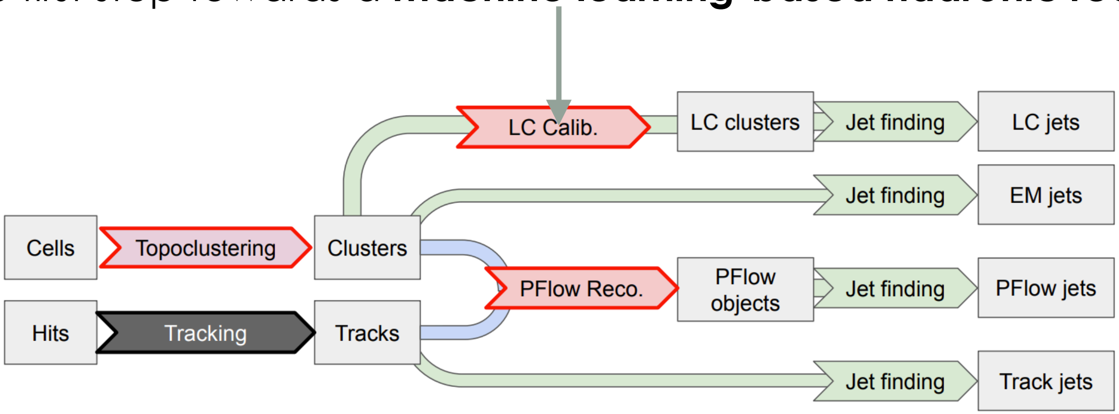




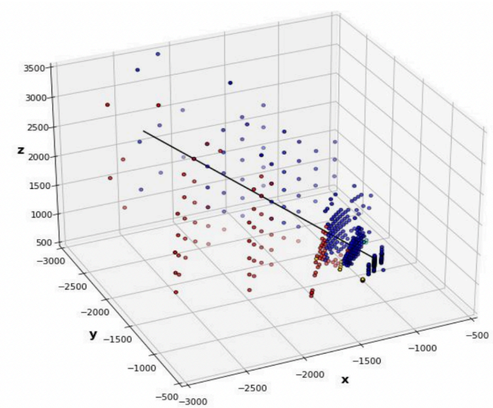
- The pion **energy resolution** of the GNN& PFN models indicate comparable or narrower response curves than the EM and LCW.

ML for Hadronic reconstruction: Summary and Outlook

- **Deep learning approaches outperform** the **classification** applied in the baseline hadronic calibration ($P^{\mathrm{EM}}_{\mathrm{clus}}$), and are able to **predict the pion energy** and **improve the energy resolution** for a wide range in particle momenta
- These results demonstrates the **potential of deep-learning-based low-level hadronic calibrations** to significantly improve the quality of **particle reconstruction** in the ATLAS calorimeter!
- This is the first step towards a **machine learning-based hadronic reconstruction**



- **Next steps:**
  - Add **tracking information** (complementary with calorimeters)
  - Study **environments closer to reality** (Multiparticles, pile-up, dense environments…jets!)



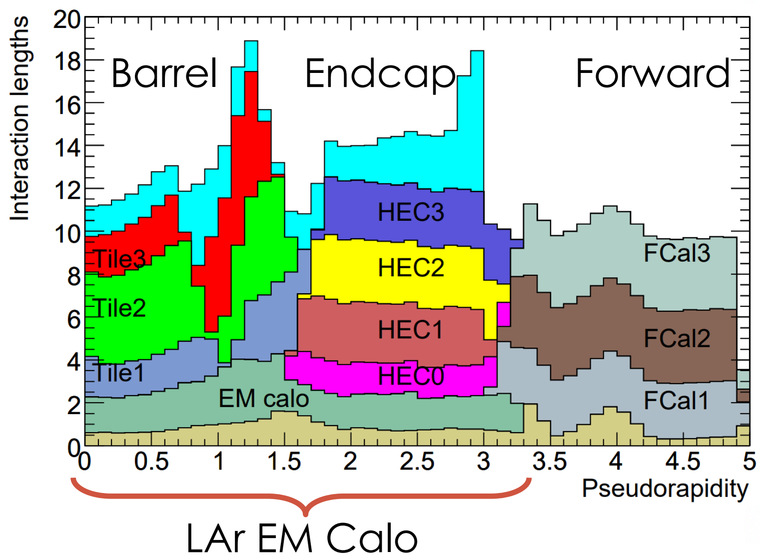
- Looking forward to implement a Particle Flow deep learning algorithm in ATLAS (uses tracks and calorimeter deposits that ideally will represent particles)

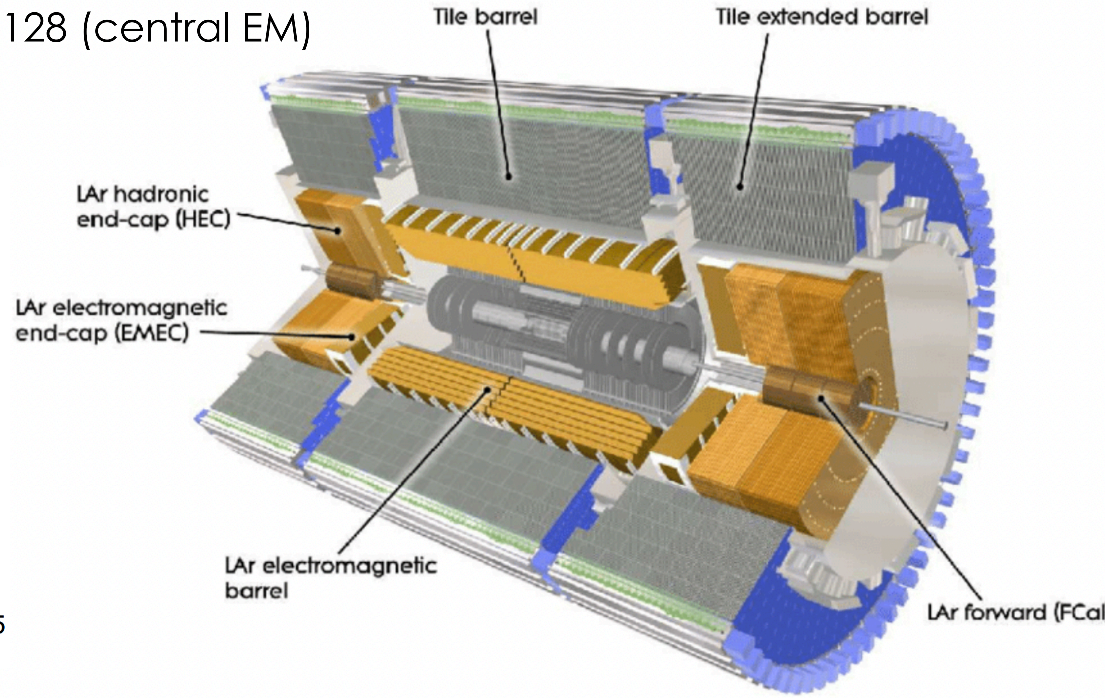
Met Significance
UFO
Mono-H
DJR
HH
LFV

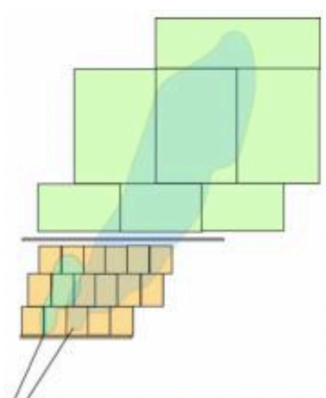
BACKUP

ATLAS CALORIMETERS

- Full coverage $|\eta| < 4.9$
- High granularity in $\Delta\eta \times \Delta\varphi = 0.025 \times \pi/128$ (central EM)
- Up to seven depth layers (samplings)



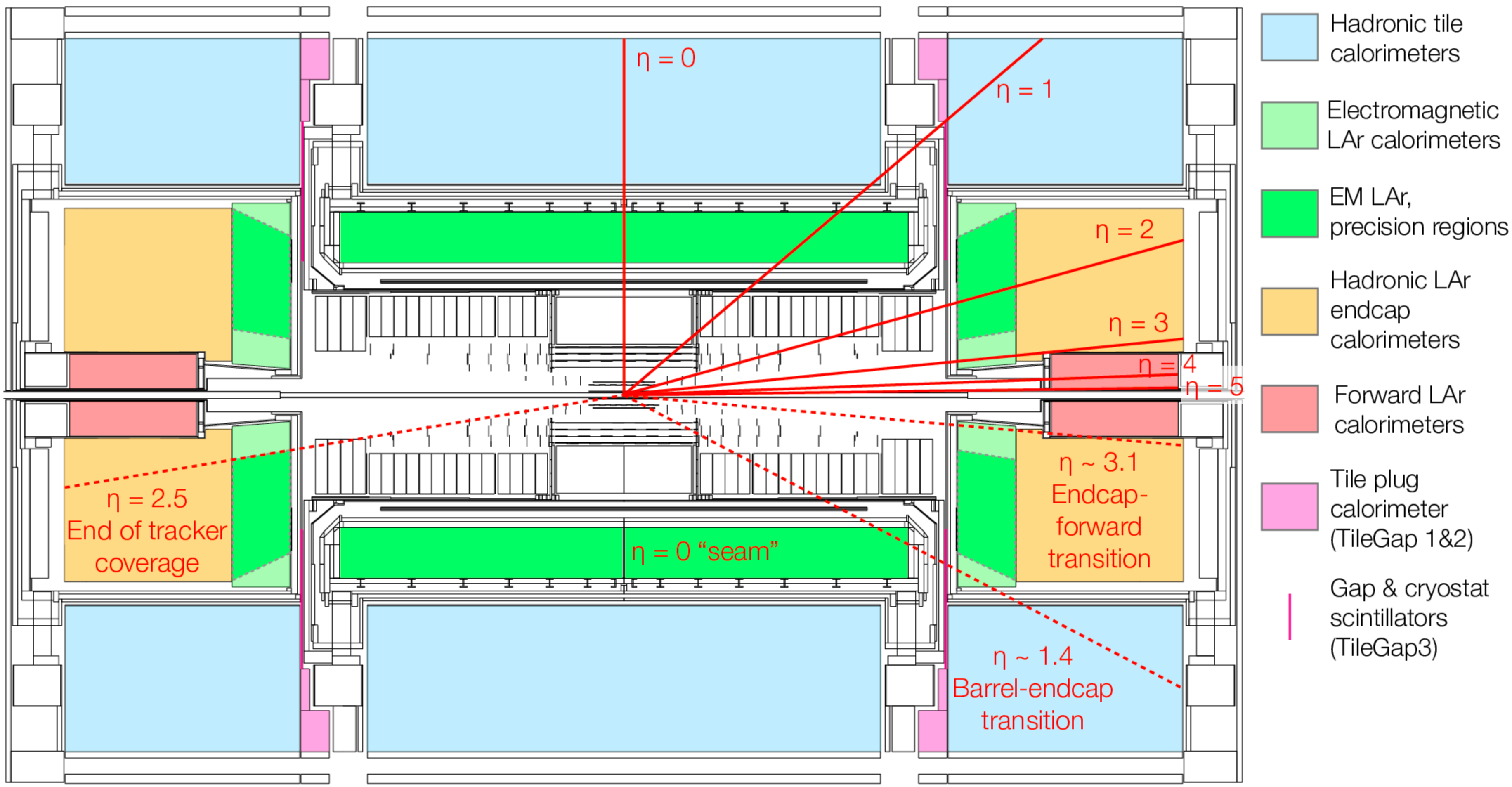




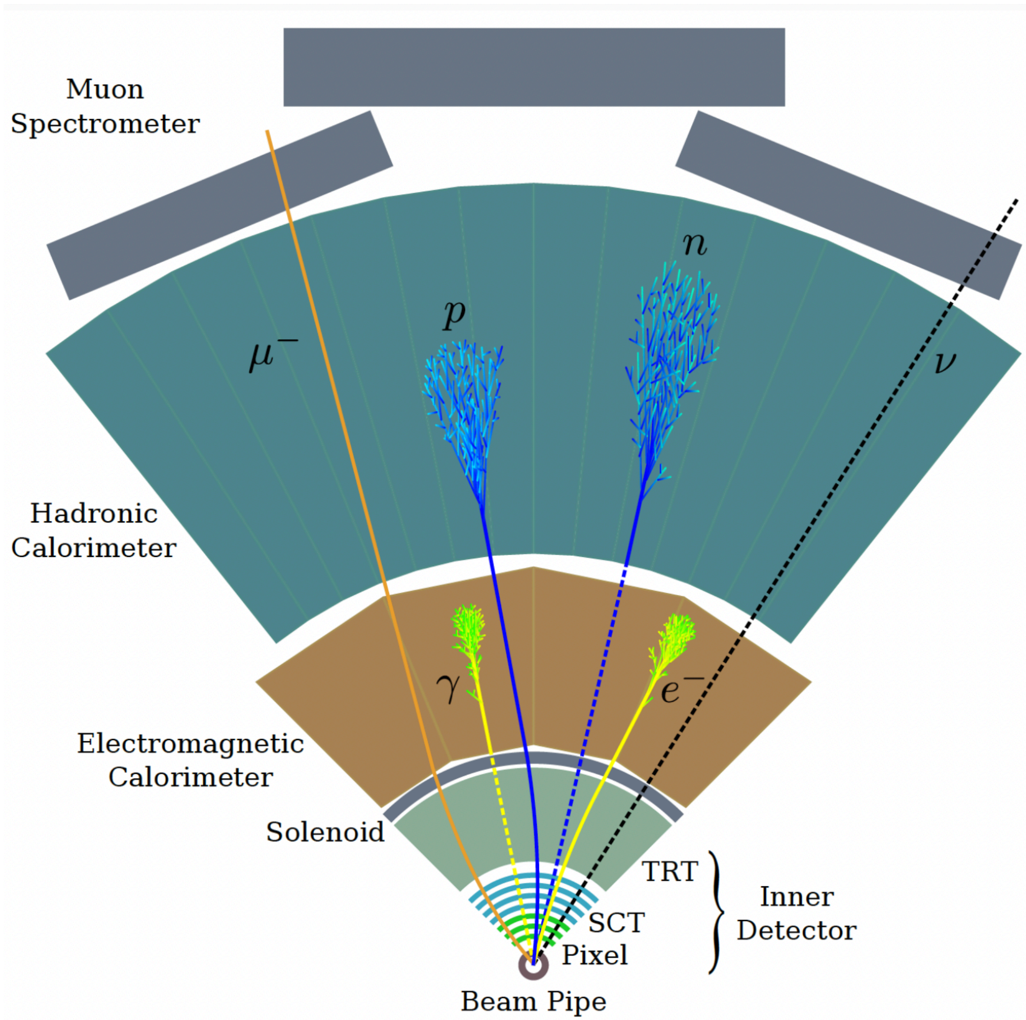
$|\eta| < 0.7$

Calorimeter Layer	$\Delta\eta$ Granularity	$\Delta\phi$ Granularity	Interaction Lengths
EMB1	$0.025/8 = 0.003125$	$\pi/32 \approx 0.1$	$\approx 4X_0$
EMB2	0.025	$\pi/128 \approx 0.025$	$\approx 16X_0$
EMB3	0.05	$\pi/128 \approx 0.025$	$\approx 2X_0$
Tile0	0.1	$\pi/32 \approx 0.1$	$\approx 1.5\lambda$
Tile1	0.1	$\pi/32 \approx 0.1$	$\approx 4\lambda$
Tile2	0.2	$\pi/32 \approx 0.1$	$\approx 2\lambda$

ATLAS calorimeters with pseudo rapidity



OBJECT RECONSTRUCTION



TOPO-CLUSTERS

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

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

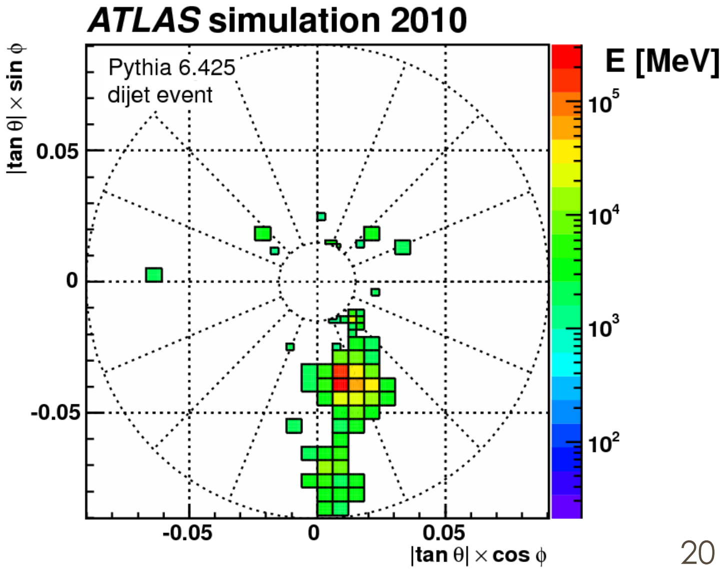
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

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

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$

Seed cells



TOPO-CLUSTERS

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

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm.**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

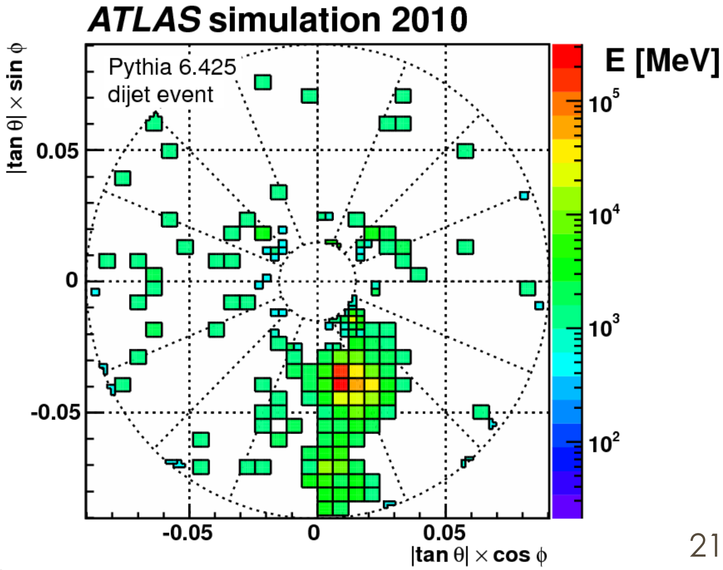
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

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

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$
- Expanded on neighbouring cells
  - All **Neighbors** with $|\zeta|>2$ are added

Growth cells



TOPO-CLUSTERS

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

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm.**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

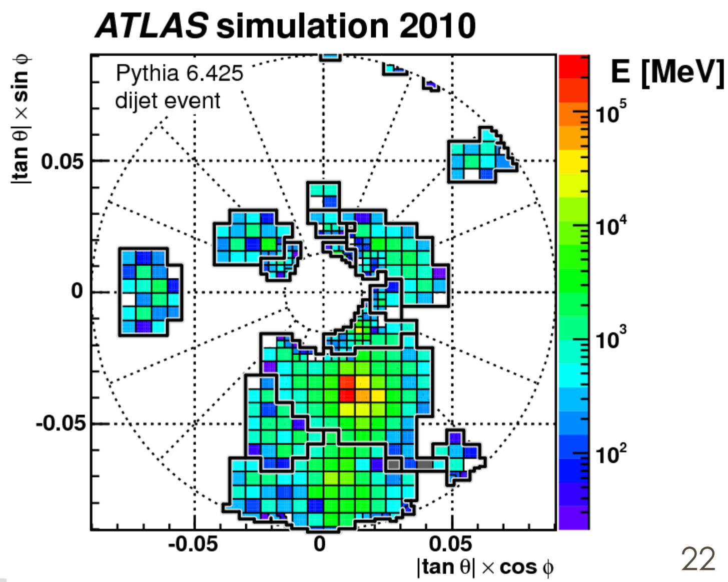
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

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

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$
- Expanded on neighbouring cells
  - All **Neighbors** with $|\zeta|>2$ are added
- **All neighbouring** cells are added regardless of the significance
  - $|\zeta|>0$
- Final cluster splitting step breaks up large topo-clusters with multiple local maxima

Boundary cells

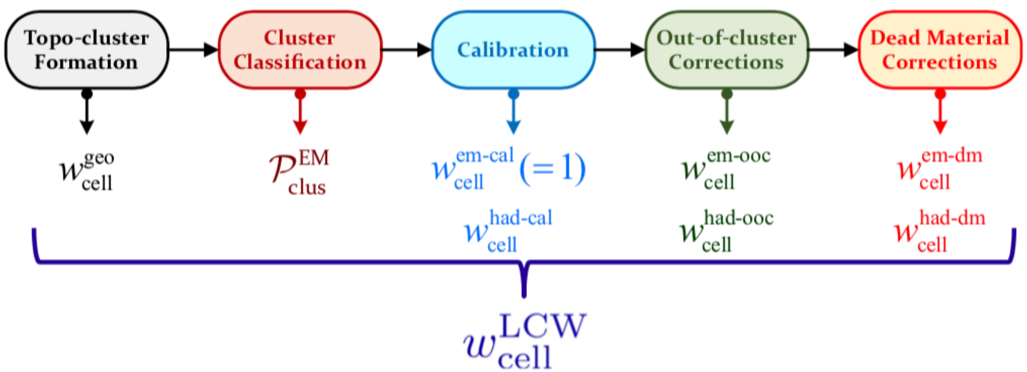


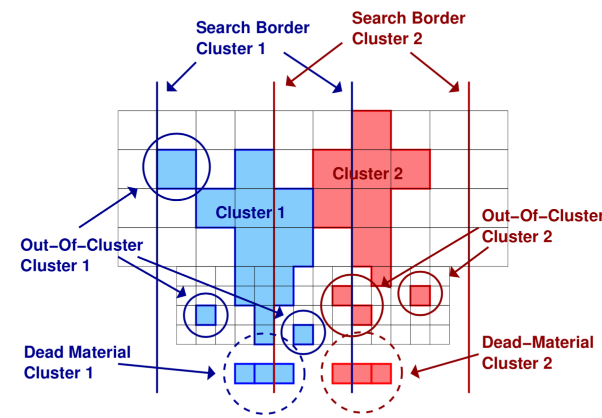
EM AND LCW SCALES

CERN-PH-EP-2011-191

Topo-clusters enter jet finding at one of two **scales**:

- **Electromagnetic** (EM) scale: same scale as the cells. Used for small-R jets.
- **Local cell weighted** (LCW) scale: Topo-clusters calibrated based on their properties. Used for large-R jets.
  - Topo-clusters are identified as either electromagnetic or hadronic. Weights are then assigned to account for
    - Differences in detector response (EM vs. HAD)
    - Energy falling in unclustered cells
    - Energy deposited in inactive (dead) regions of the detector

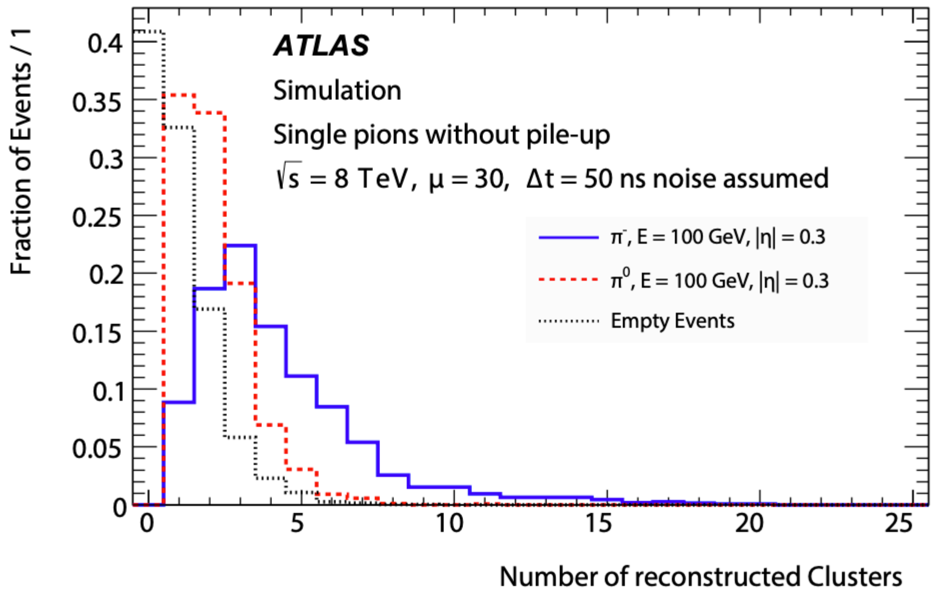




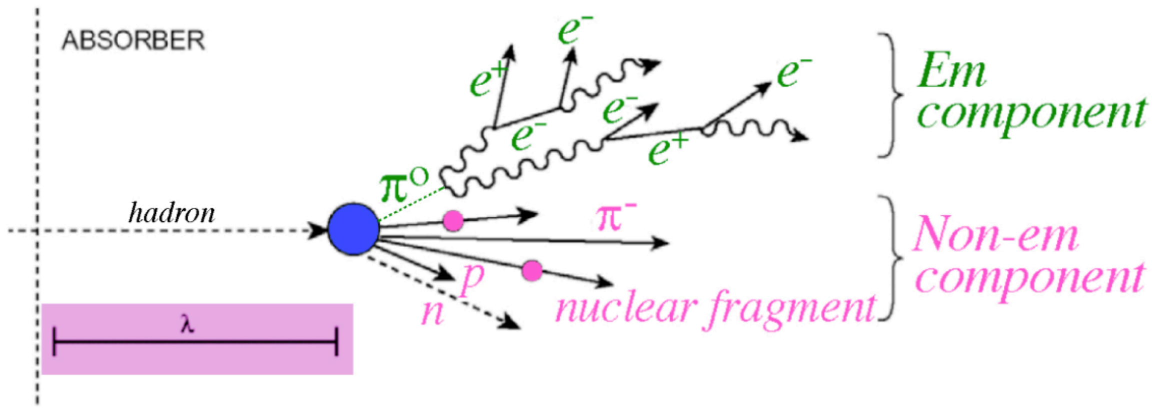
$$E_{\text{clus}}^{\text{LCW}} = \sum_{i \in \text{cluster}} w_{\text{cell},i}^{\text{LCW}} \, E_{\text{cell},i}^{\text{EM}}$$

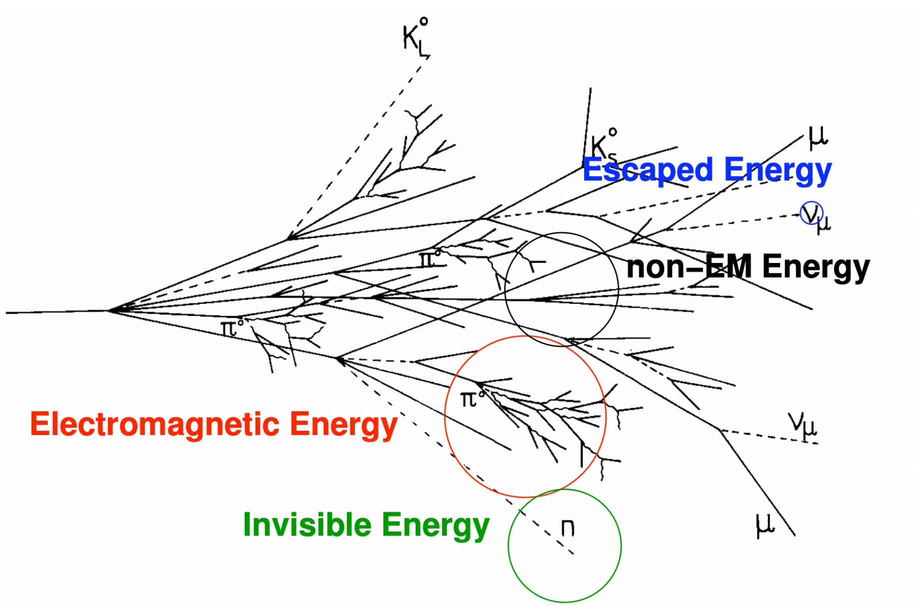
TOPO-CLUSTERS

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



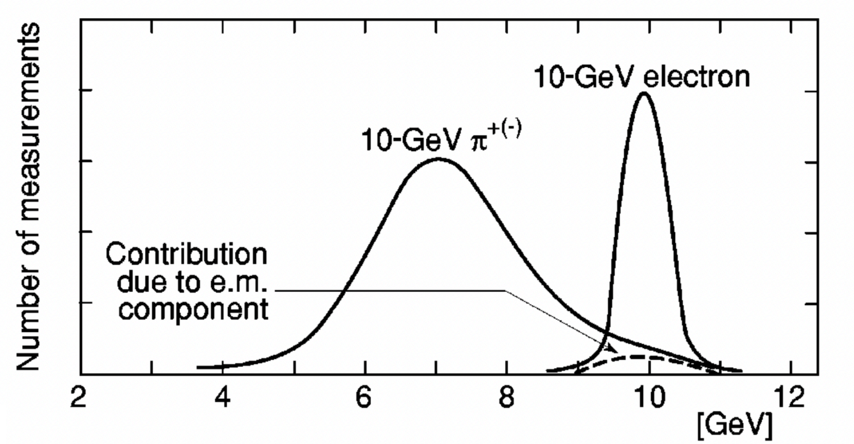
HADRONIC SHOWER





https://cds.cern.ch/record/1222464/files/0911.2639.pdf

https://cds.cern.ch/record/692252/files/RevModPhys.75.1243.pdf



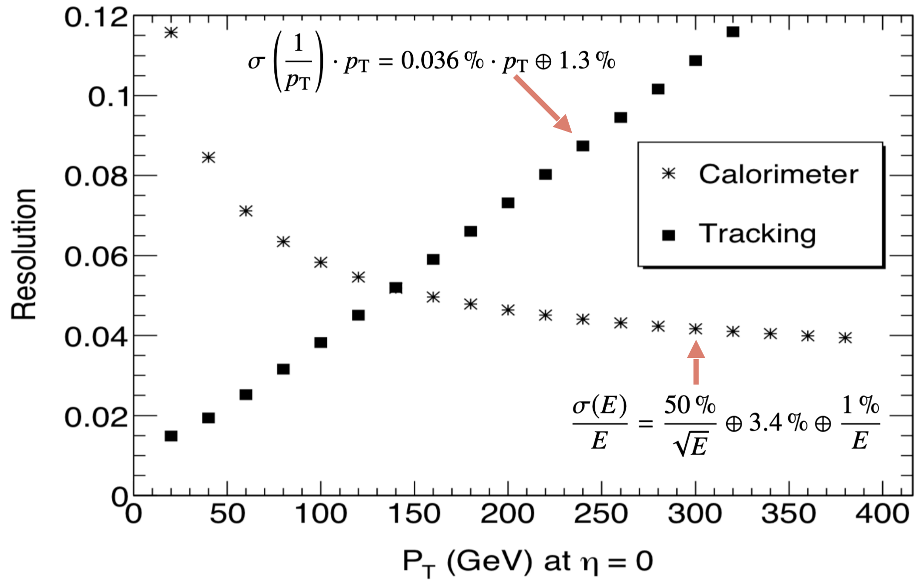
Signal (in energy units) obtained for a 10 GeV energy deposit

COMBINING CLUSTERS AND TRACKS

Calorimeter and tracker provide complementary information

Tracker:
- Sensitive to charged particles
- Better angular resolution
- Able to assign tracks to pile-up or hard-scatter vertex.
- Better reconstruction efficiency and momentum resolution at low pT

Calorimeters:
- Sensitive to both neutral and charged particles.
- Better energy resolution at high pT



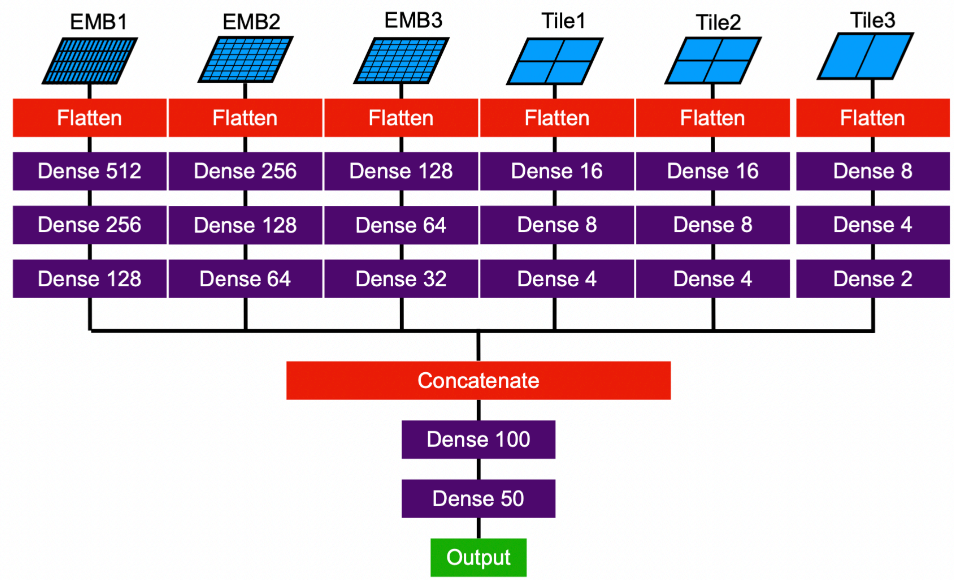
Topocluster images: Merged Deep Fully Connected Network (DNN)

- Images are flatten into one-dimensional vectors.

Classification

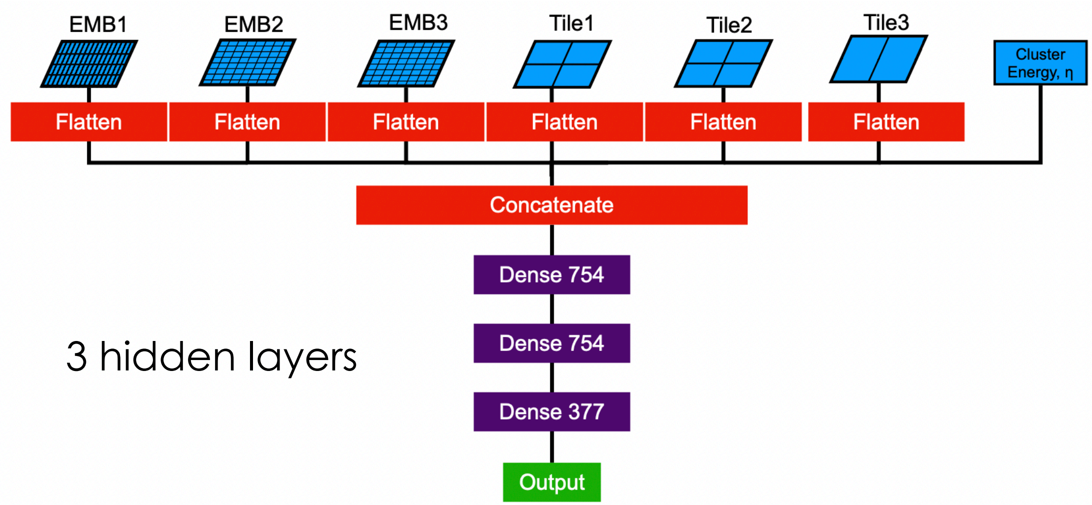
each calorimeter layer is considered as a separate input to the model

3 hidden layers



Regression

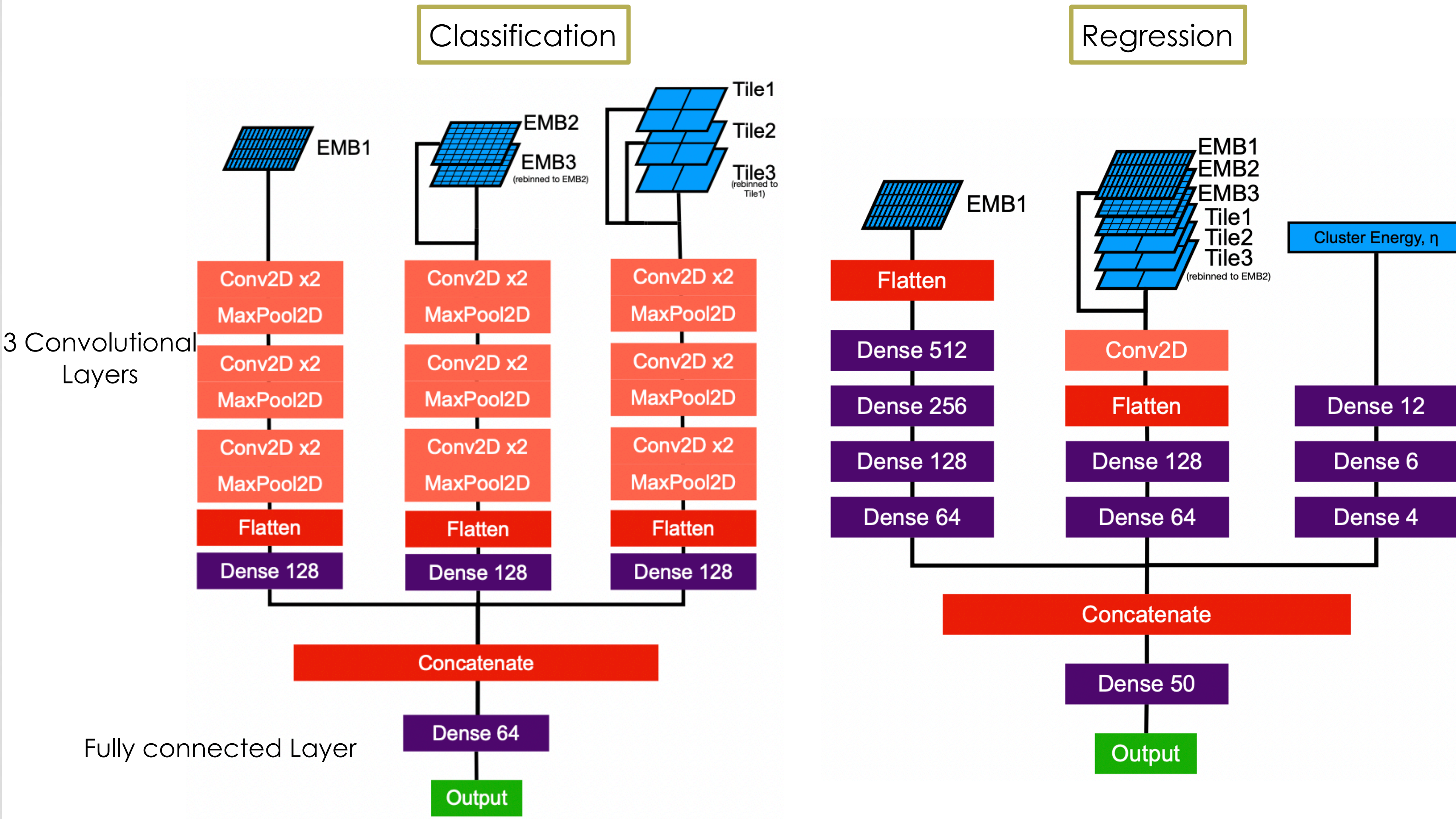
All cells are concatenated into a single vector of dimension 752



3 hidden layers

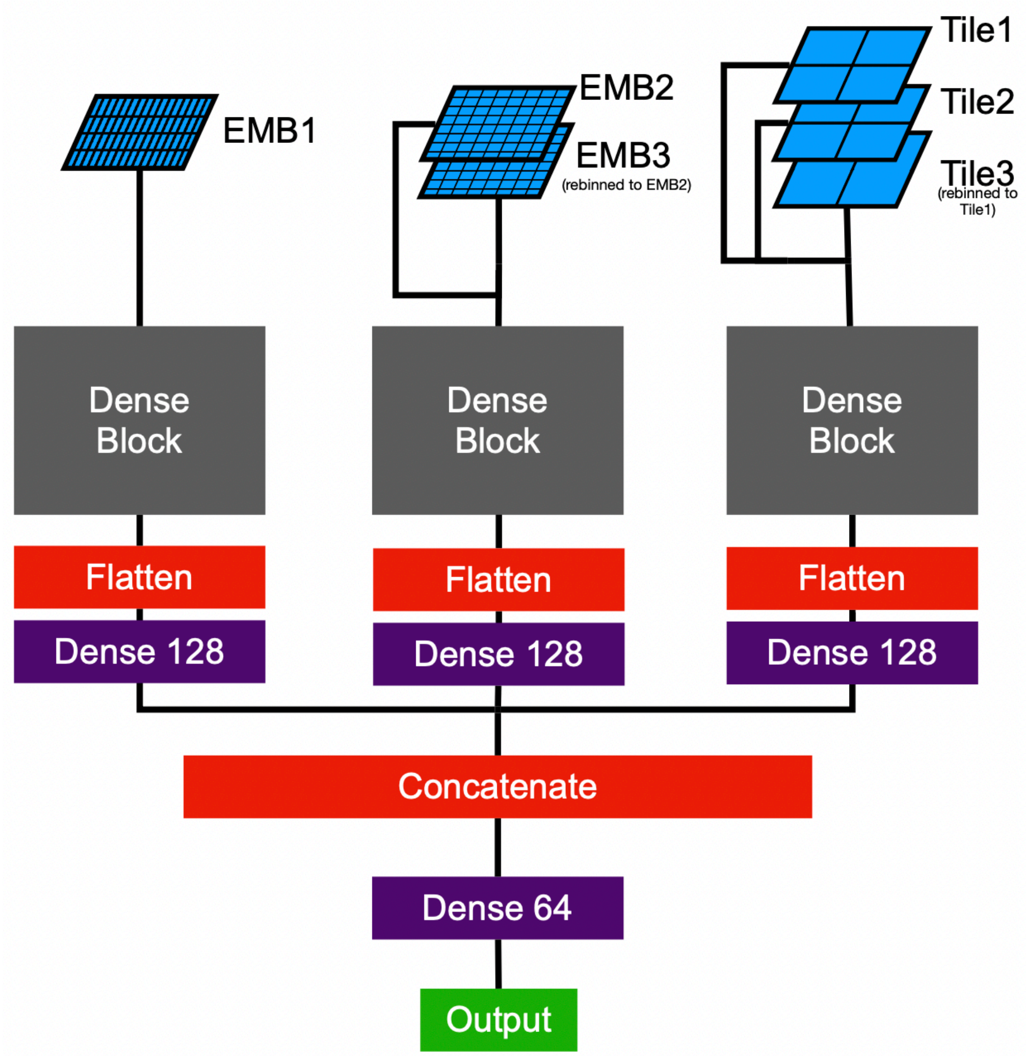
Topocluster images: Convolutional Neural Network (CNN)

- The entirety of the two-dimensional images are used as inputs to the model.
- The layers of the calorimeter can be thought of as color channels of traditional image classification problems



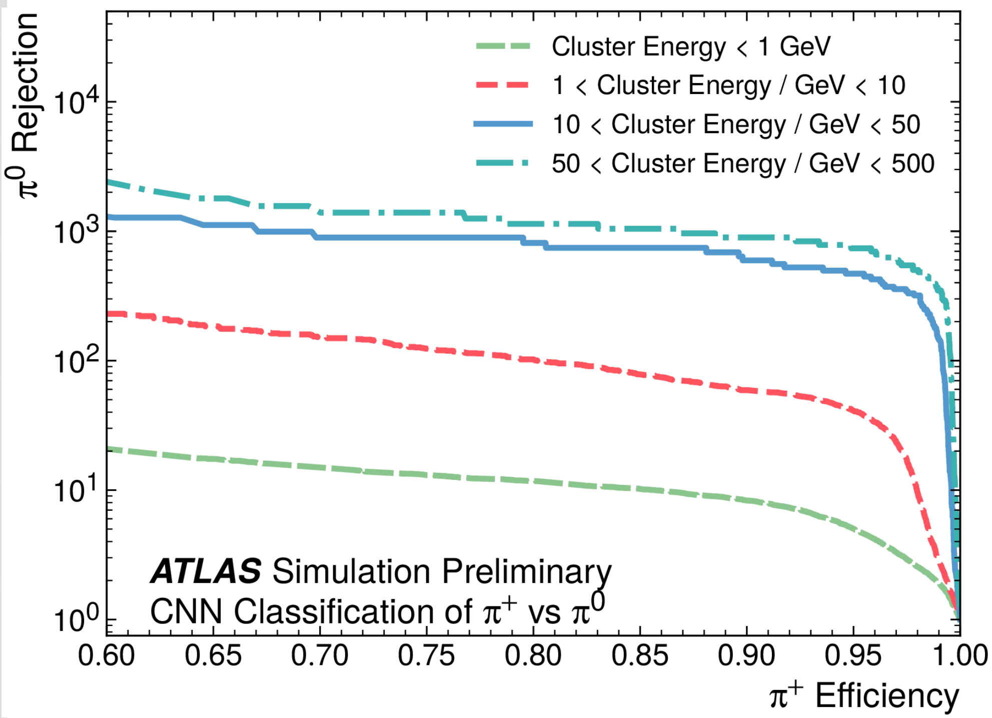
Topocluster images: Densely Connected Convolutional Networks (DenseNets)

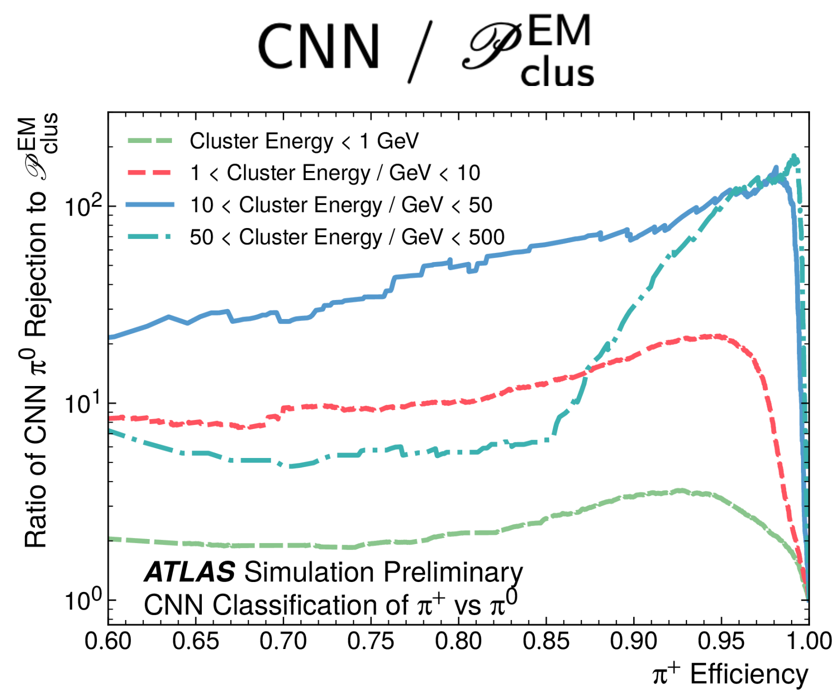
Every layer receives as inputs the concatenated feature maps from every previous layer



Topocluster images: CNN classification performance

- classifier performs increasingly better at higher energies
- CNN was not explicitly trained with energy as an input, but the shower shape's dependence on energy is sufficient to provide effective separation at all energies.

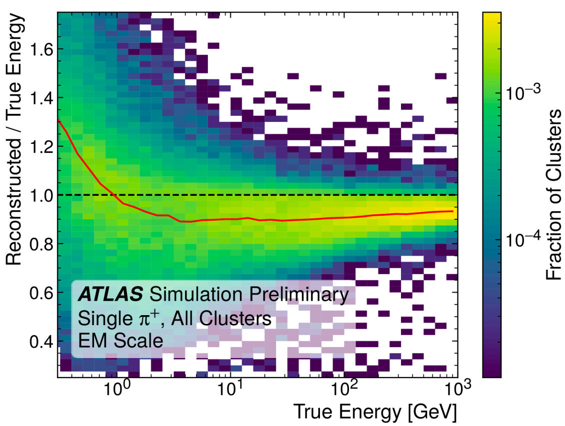




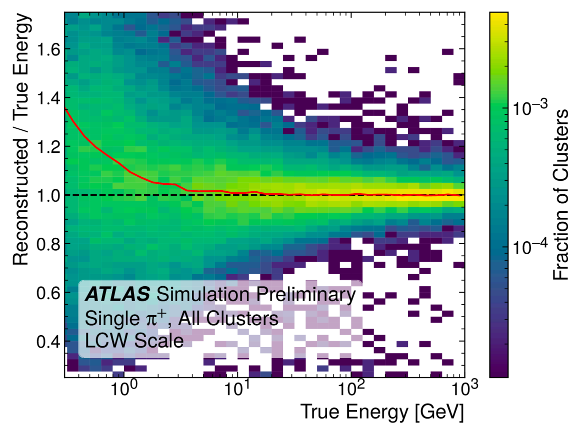
Topocluster images: Pion Energy Calibration

- After identifying a cluster as hadronic/EM, need to convert the signal into an energy measurement
- Energy regression goal: Correctly predict the true energy deposited in the cluster.
  - Quantified by measuring the cluster energy response: $R = \frac{E^{\text{reco}}}{E^{\text{truth}}}$ that should be ~ 1

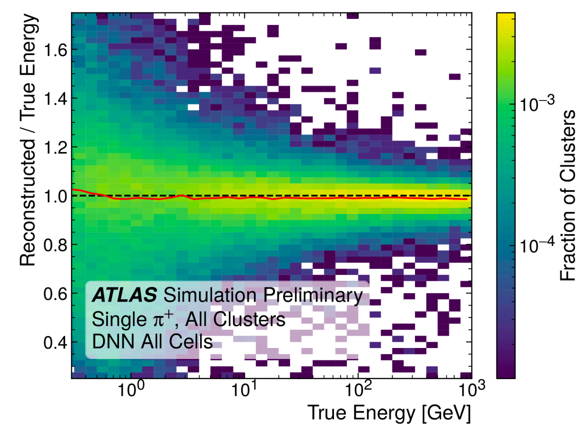
Regression performance for charged pions



Raw "EM" scale under-estimates R

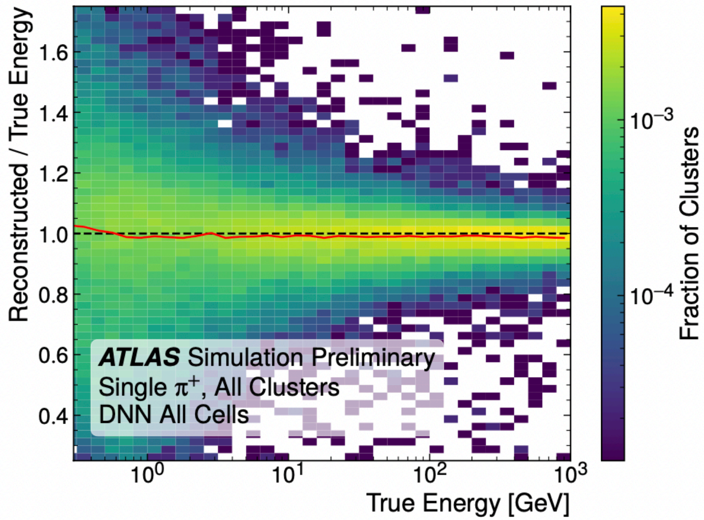


LCW over-estimates R at low-energy

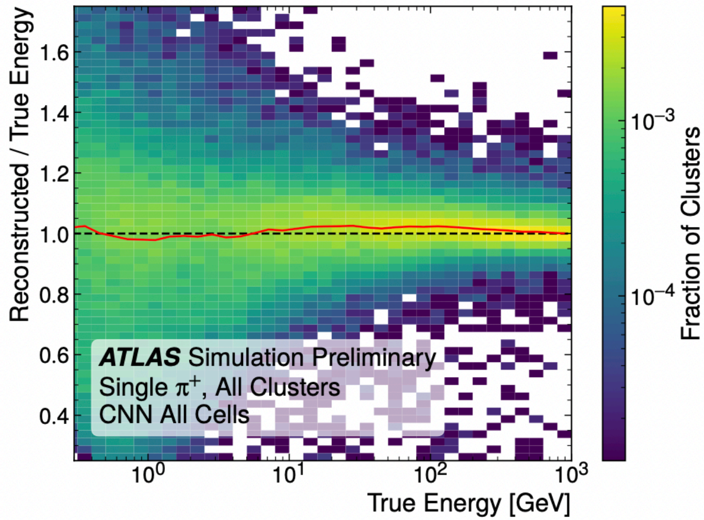


DNN regression does an excellent job nearly everywhere

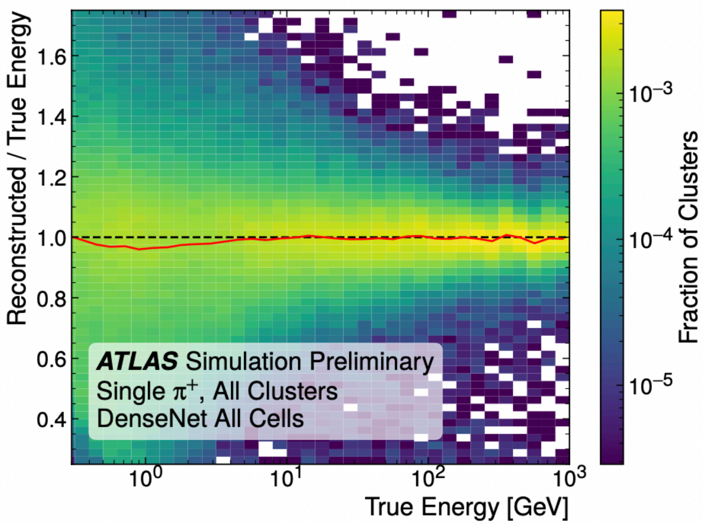
Topocluster images: Pion charged Energy Calibration



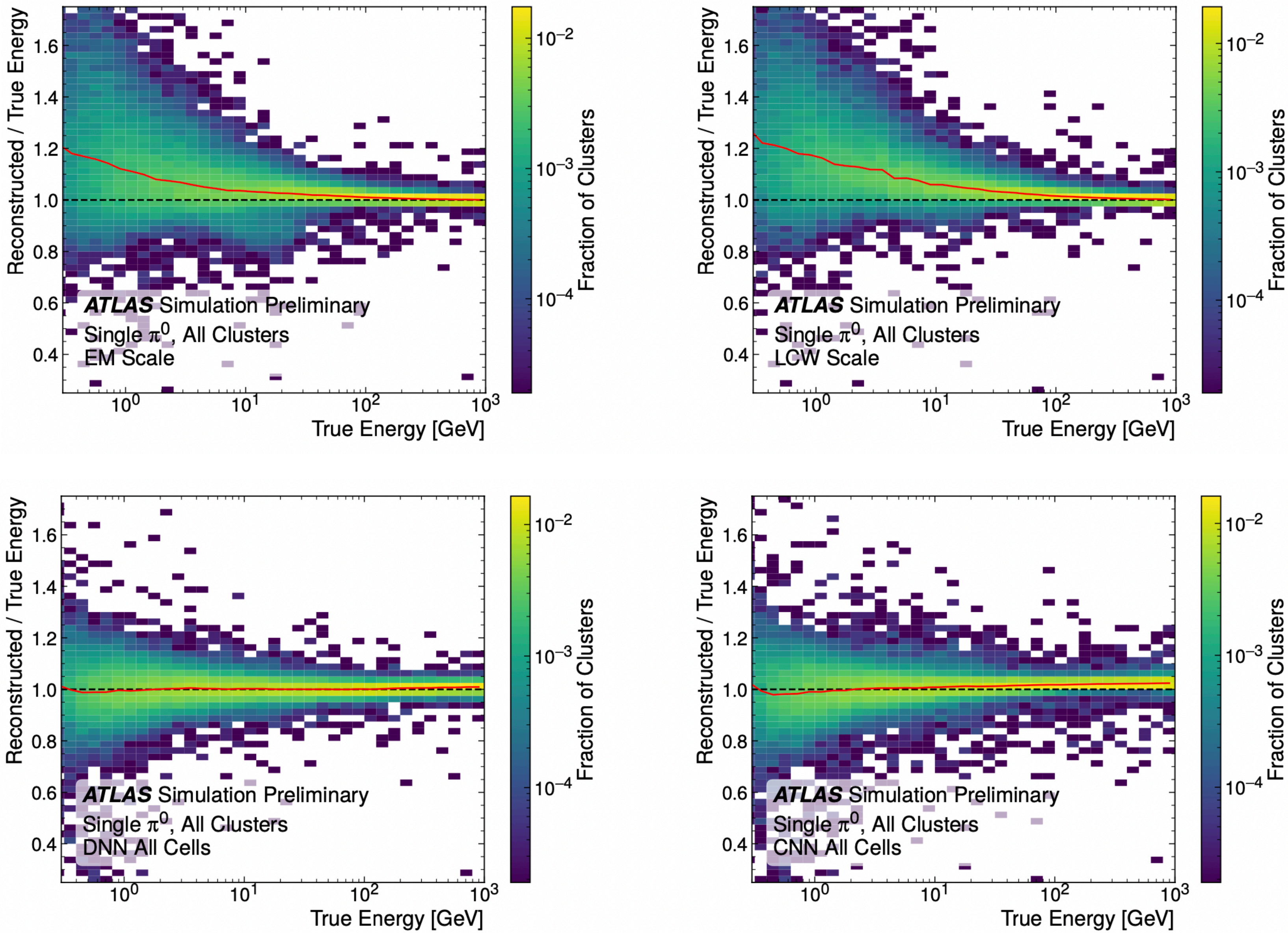
(c)



(d)



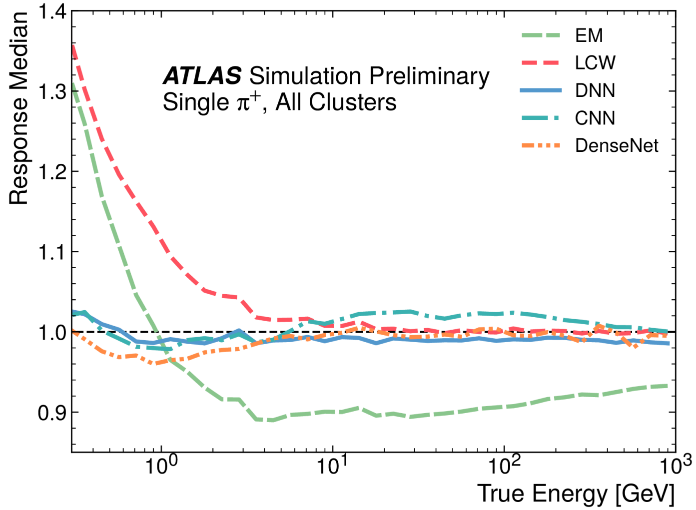
Topocluster images: Neutral Pion Energy Calibration

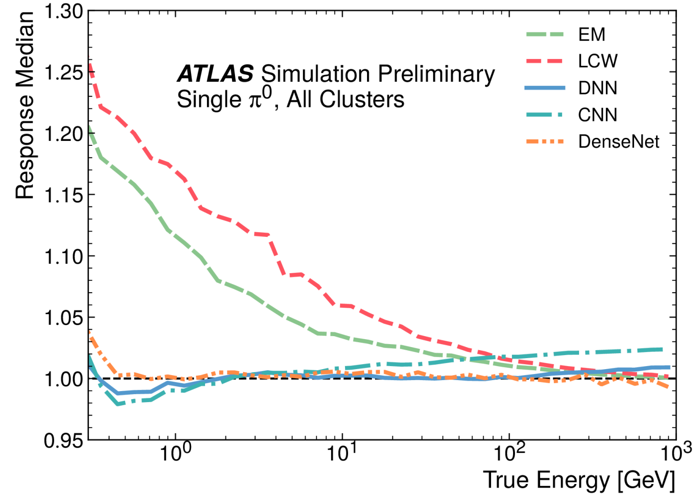


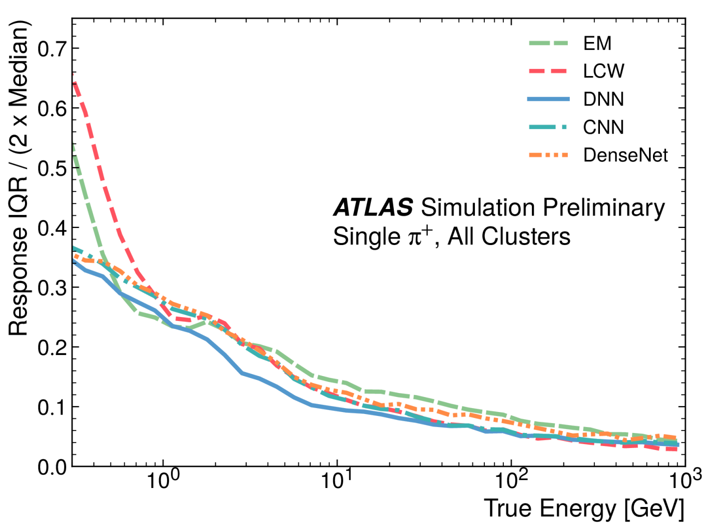
Pion Energy Calibration

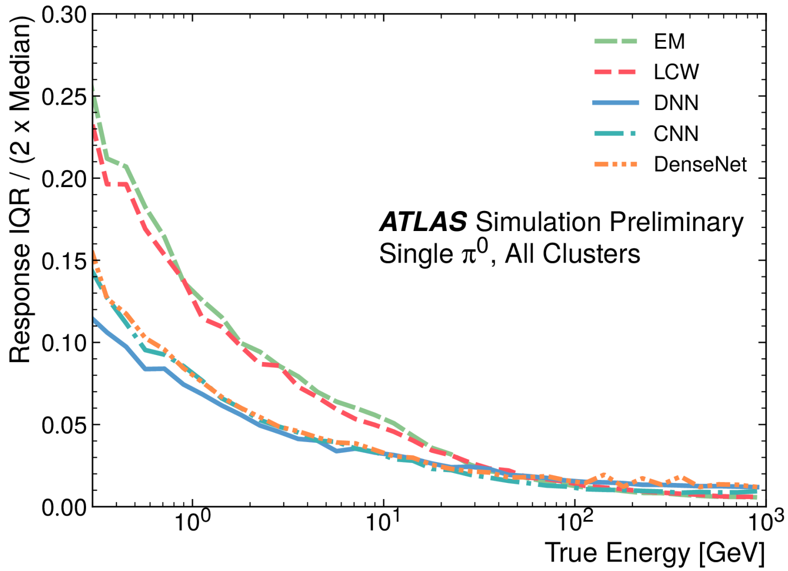
π^+

π^0



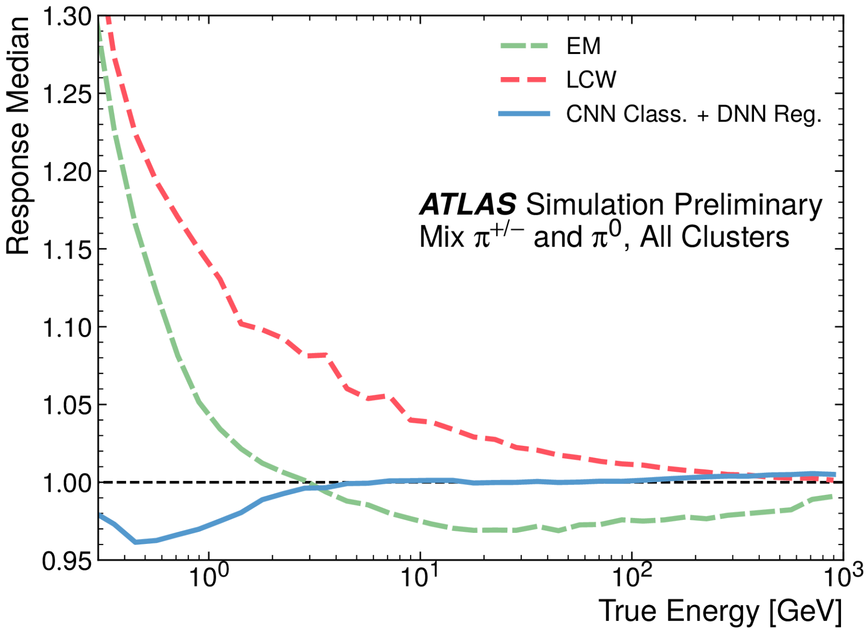


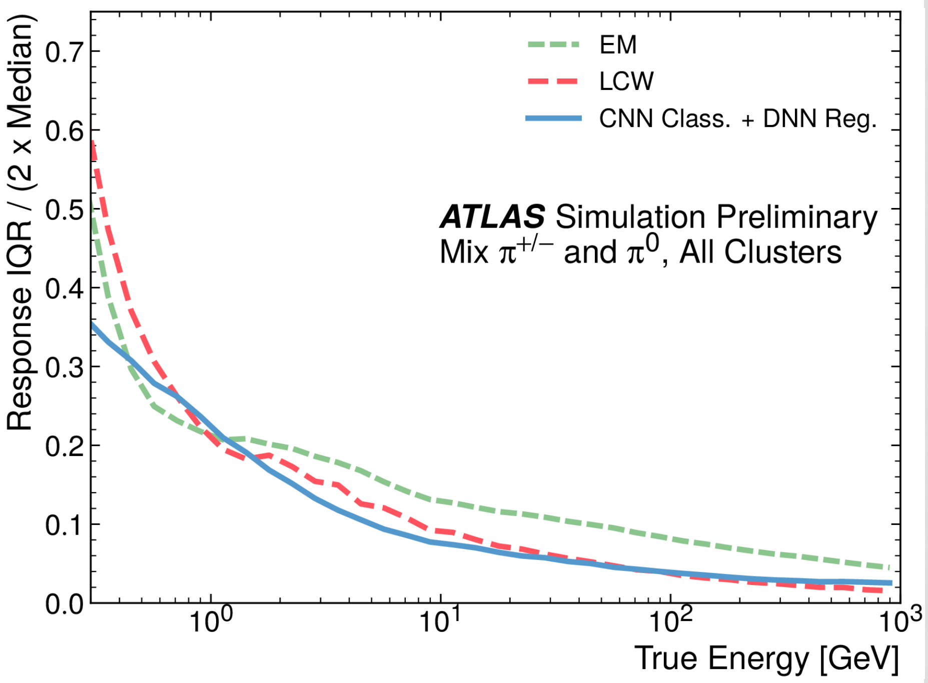




Mixed sample of π ± and π0

- First look at the **performance with jets**
  - π^+, π^- and π^0 mixed in a 1:1:1 ratio
  - Roughly correspond to the expected distribution in jets

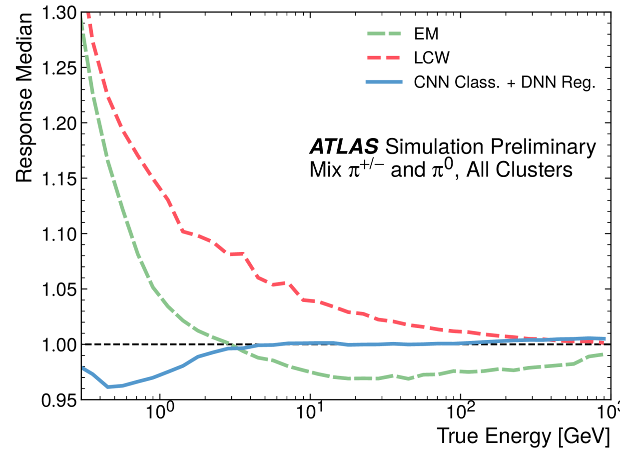




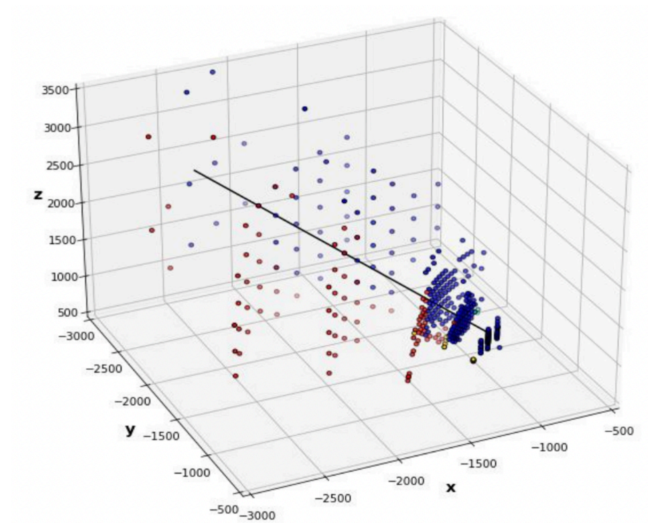
Outlook

Looking forward studying more complex scenarios:

- First look at the **performance with jets**
  - π^+, π^- and π^0 mixed in a 1:1:1 ratio
  - Roughly correspond to the expected distribution in jets



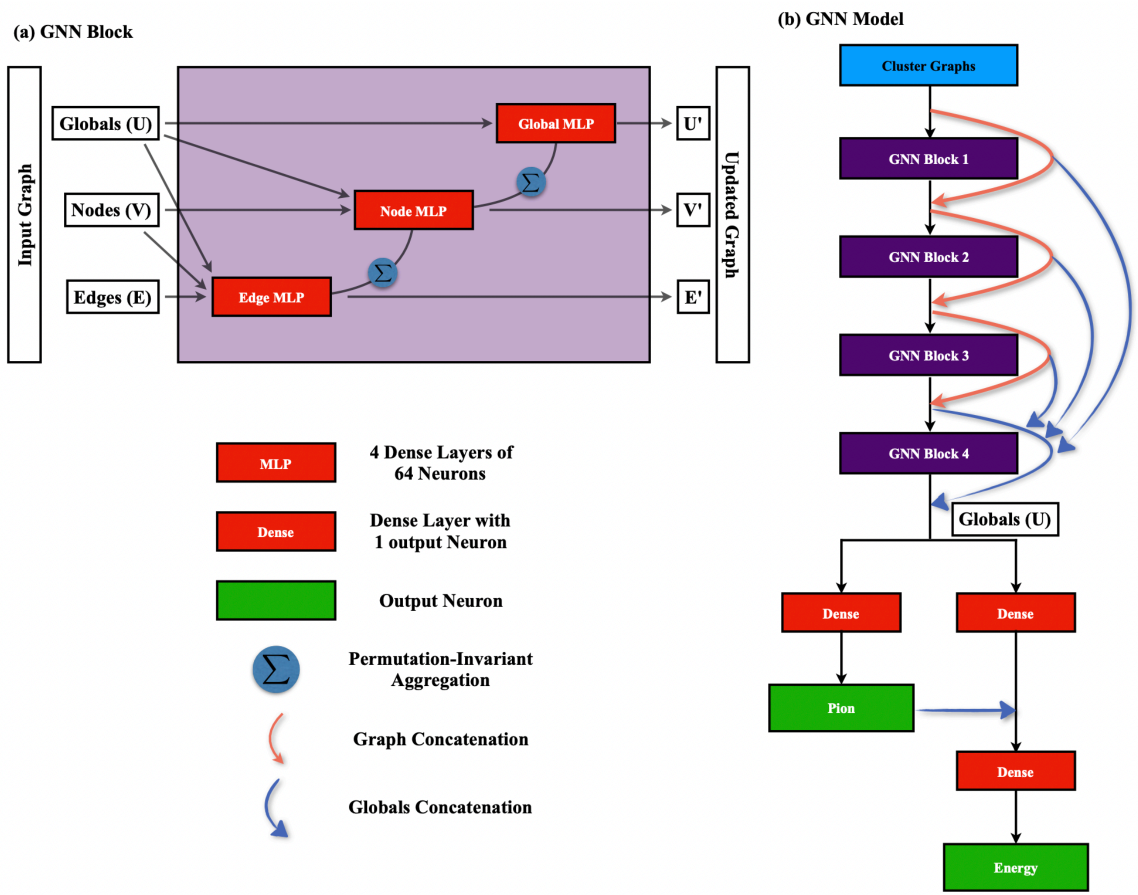
- Another handy way to represent energy deposits is as a **point-cloud**
  - Points contains cell info & cluster-level info.
  - Allows for combining signals from the inner detector (tracks) and from calorimeter (clusters)



Topoclusters point clouds: Graph Neural Network (GNN)

- The classification & regression losses are balanced together in the same model using a loss function that accommodates both tasks

$$\mathcal{L} = (1-\alpha)\mathcal{L}_{\text{classification}} + \alpha\mathcal{L}_{\text{Regression}}$$



Deep Sets

- Deep Sets are designed for permutation-invariant & variable-length data
- One can treat Topo-clusters as simple unordered "set"
- "Deep Sets" paradigm:

Observable Decomposition. *An observable $\mathcal{O}$ can be approximated arbitrarily well as:*

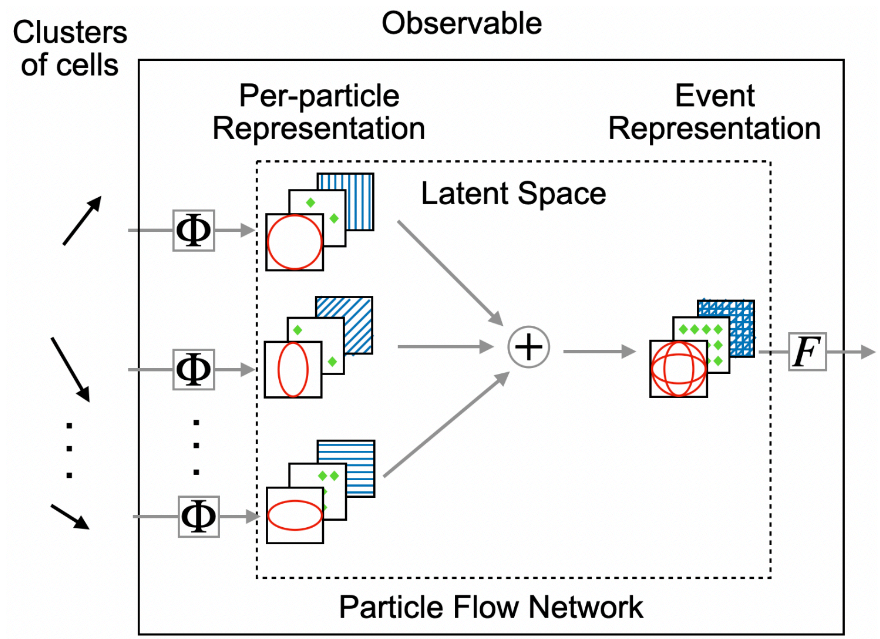
$$\mathcal{O}(\{p_1,\dots,p_M\}) = F\left(\sum_{i=1}^{M}\Phi(p_i)\right), \tag{1.1}$$

where $\Phi:\mathbb{R}^d\to\mathbb{R}^\ell$ is a per-particle mapping and $F:\mathbb{R}^\ell\to\mathbb{R}$ is a continuous function.

Particle Flow Network (PFN)

- Adapts the Deep Sets framework for particle physics data
- Each set point have features per cell

$$\text{PFN}: \text{O} = \sum_{\text{i}=1}^{\text{M}} \Phi(\text{E}, \eta_\text{i}, \phi_\text{i}, \text{SamplerID})$$



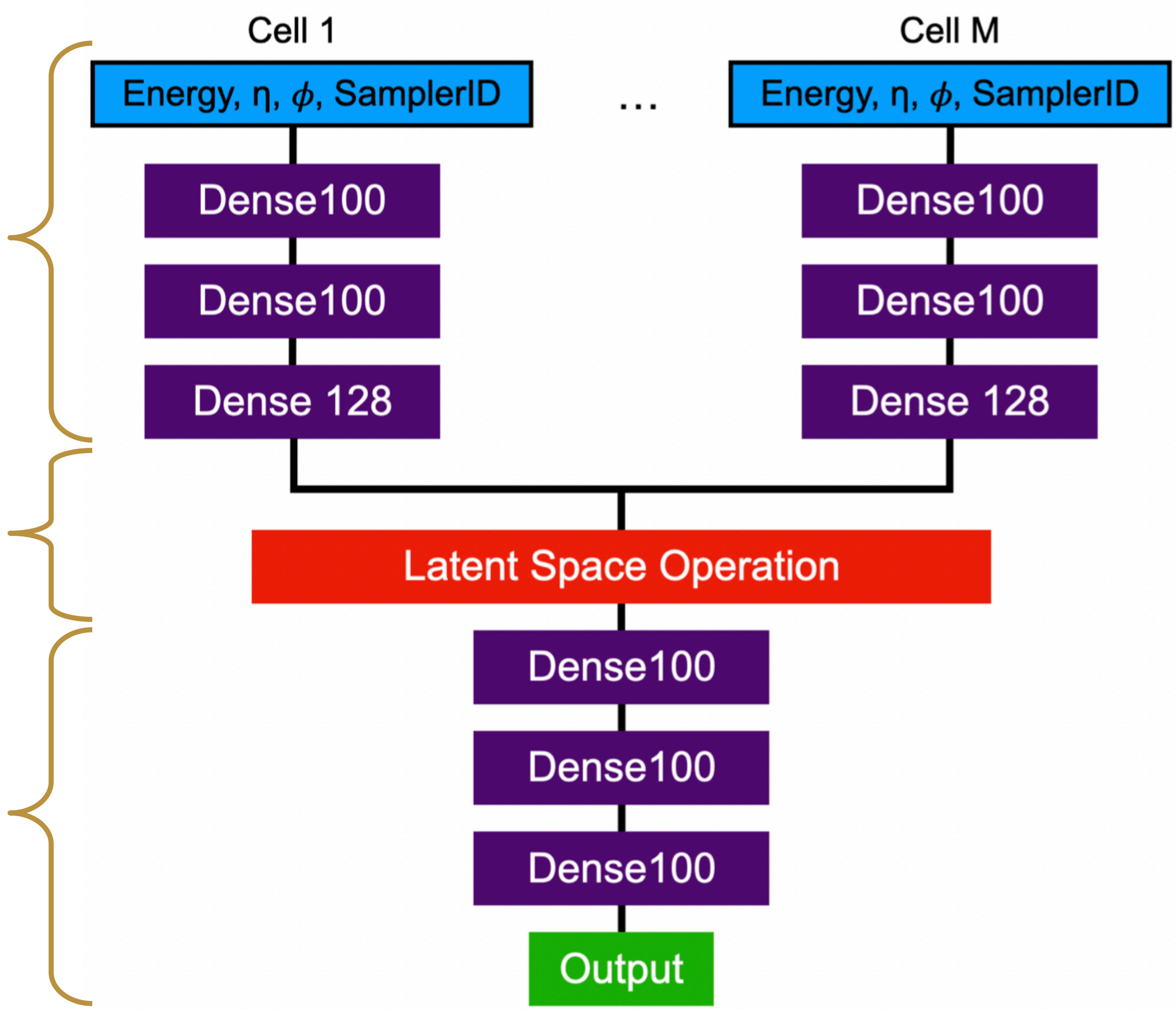
Particle Flow Network implementation

Approximate Φ and F with Neural Networks

Per-particle Φ dense networks

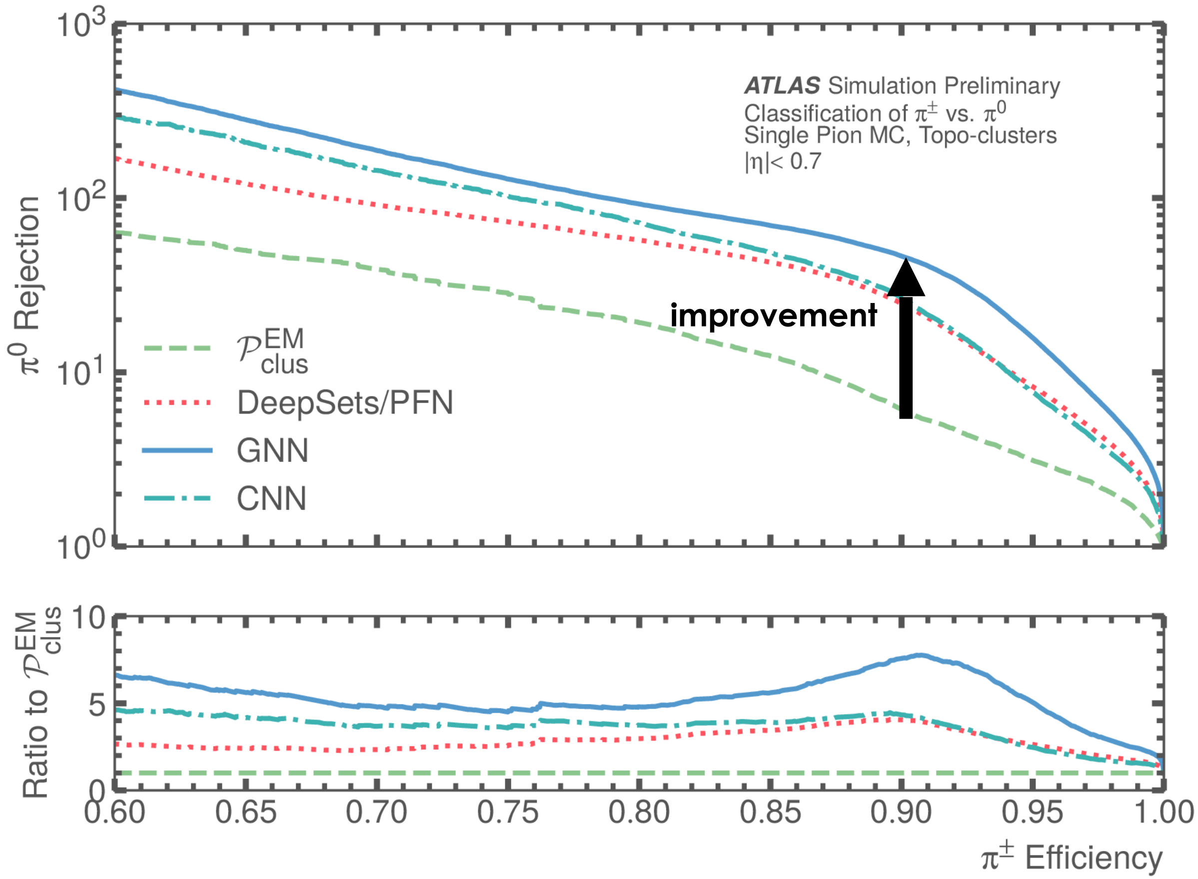
Direct sum (permutation-invariant)

Direct sum is passed through a final NN representing the global F



Topoclusters point clouds: Pion classification

- New **point cloud** approaches (GNN & PFN) far **outperform** the baseline EM cluster probability ($P_{\text{clus}}^{\text{EM}}$)
- They also perform **on par with or better** than the **image-based** CNN approach for pion classification



Pion classification with Particle Flow Network (PFN)