Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Shower Separation for Highly Granular Calorimeters using Machine Learning

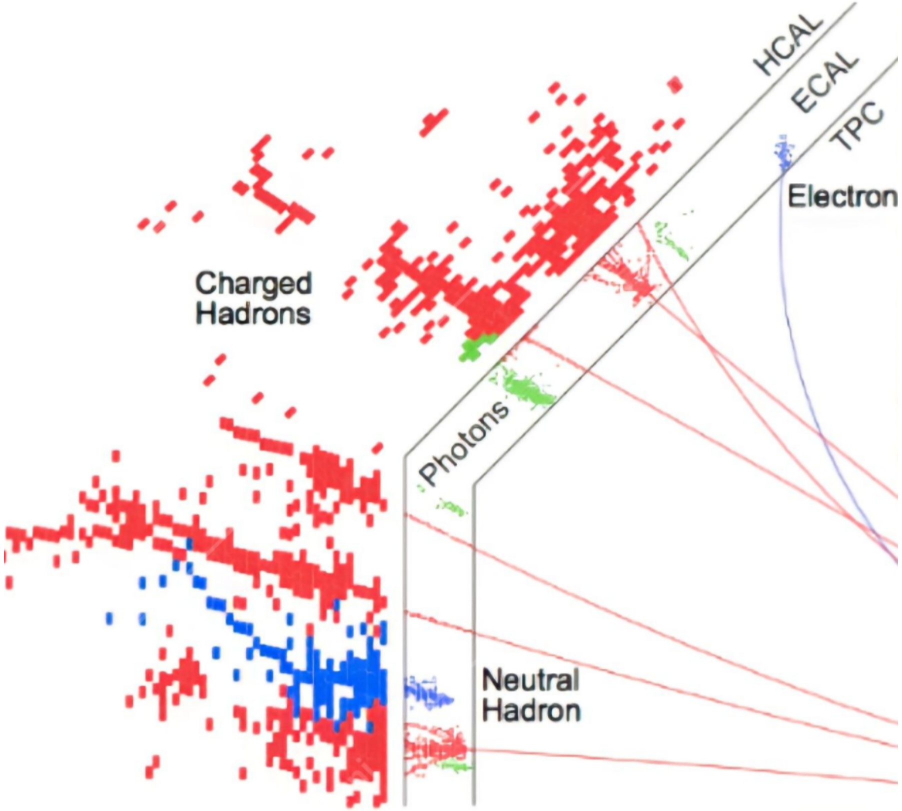
Jack Rolph, Erika Garutti
Institute for Experimental Physics, University of Hamburg

CALICE Collaboration Meeting, Göttingen

31/03/2023

Particle Flow Calorimetry

- Jet energy resolution at future precision e^+-e^- colliders must produce a di-jet invariant mass resolution of ~3% in the range of jet energies 50-200 GeV [1]
- **Problem:** typical jet energy resolution of 'traditional' calorimetry is much worse than required at ILC.
- **Solution:** Particle Flow Calorimetry (PFC) [2]:
  - measure momentum of charged particles (~60% of jet energy) using tracker;
  - use highly granular calorimeters to measure energy of photons and neural hadrons;
  - Use sophisticated clustering algorithms to associate tracks to energy deposits e.g. Pandora Particle Flow Algorithm, PPFA

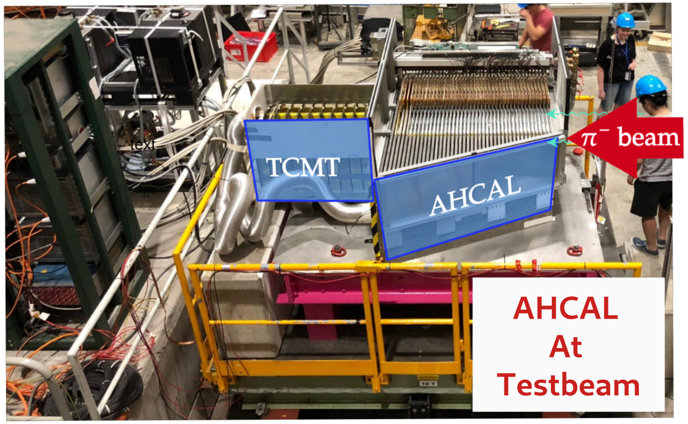


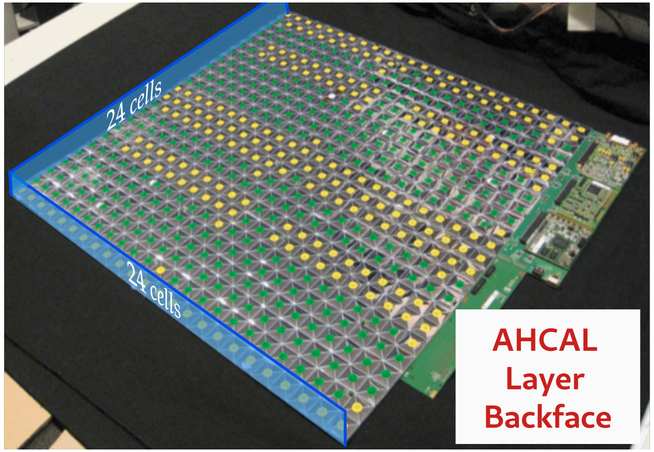
[1] M. A. Thomson. 'Particle Flow Calorimetry and the PandoraPFA Algorithm'. NIMA, pp. 25–40. doi: 10.1016/j.nima.2009.09.009.

CALICE Analogue Hadronic CALorimeter (AHCAL)

- AHCAL is a Fe-Sc highly granular calorimeter prototype designed for Particle Flow;
- Calorimeter has ~22,000 individual SiPM-on-tile readout channels → highly granular;
- AHCAL is a **five dimensional calorimeter:**
  - it measures energy density of hadron showers in **space and time;**
  - time resolution: up to 100 ps time resolution allowed by hardware;
- Spatial and temporal readout is expected to aid in clustering →

  improved sensitivity to hadron shower substructure and development;





TAKE-HOME MESSAGE:

spatial & temporal energy density information available from AHCAL mat improve PFC clustering performance.

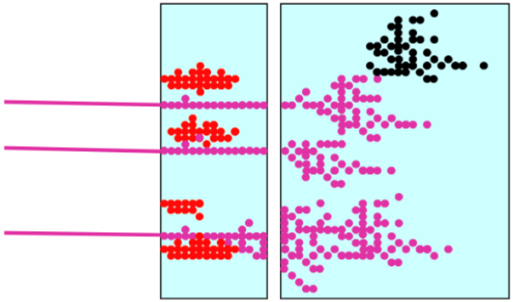
Machine Learning for PFC Clustering

- Confusion defined as:
  'the energy misallocated between clusters of energy deposits in PFC' [1];

- Can occur for two main reasons [1]:
  - insufficient sampling points in the calorimeter;
  - **lack of sophistication in the pattern recognition algorithms.**

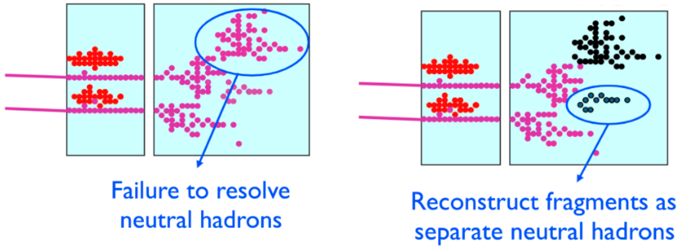
- Graph neural network techniques have demonstrated excellent performance for shower separation [2].

- However:
  - Influence of timing information on confusion unknown;
  - AHCAL very highly granular compared to [2] → unknown if models scale.



$$E_{JET} = E_{TRACK} + E_{\gamma} + E_{n}$$

Studied in this presentation

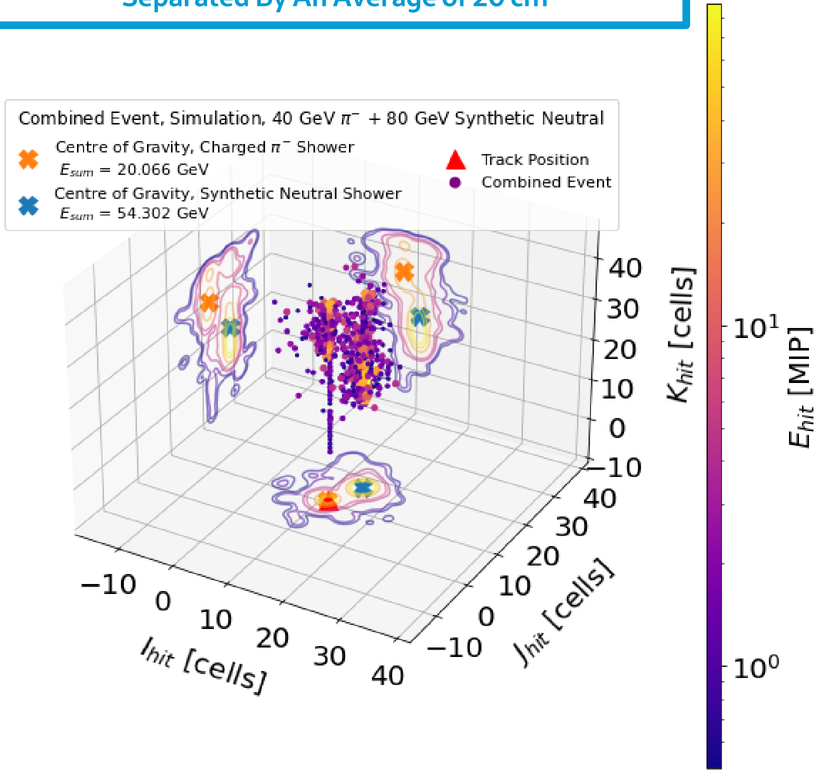


[2] Shah Rukh Qasim et al. 'Learning representations of irregular particle-detector geometry with distance-weighted graph networks'. In: The European Physical Journal C 79.7 (July 18, 2019), p. 608. doi:10.1140/epjc/s10052-019-7113-9.

Problem Statement

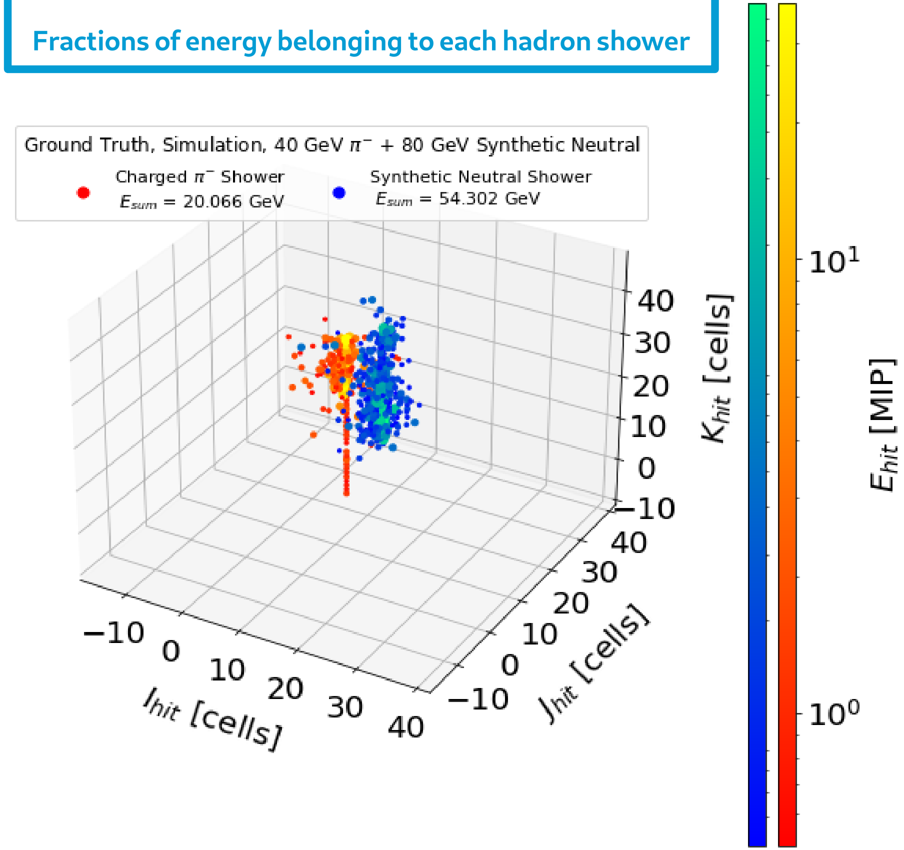
What We Have:

Synthetic 'Charged-Neutral' AHCAL Shower Events, Separated By An Average of 20 cm



What We Want:

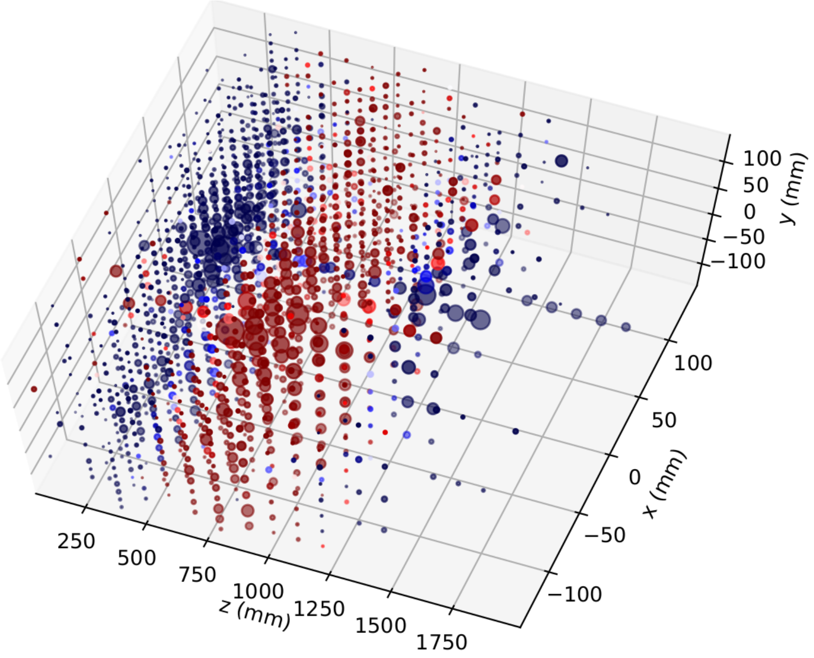
Fractions of energy belonging to each hadron shower



Shower Separation Models

- Three published neural network models are applied to shower separation for AHCAL:
  - **PointNet** [3]
  - **Dynamic Graph Convolutional Neural Network (DGCNN)** [4]
  - **GravNet** [2]
- Output → fraction of energy in each cell belonging to each shower;
- Models modified to be able to include full event information;
- Around $2\text{x}10^6$ weights overall, with 90-100 weights per sensor.

Event display of GravNet applied to hadron shower separation in simulated tungsten calorimeter [2]



[2] Shah Rukh Qasim et al. 'Learning representations of irregular particle-detector geometry with distance-weighted graph networks'. In: The European Physical Journal C 79.7 (July 18, 2019), p. 608. doi:10.1140/epjc/s10052-019-7113-9.
[3] Charles R. Qi et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Apr. 10, 2017. doi: 10 . 48550 / arXiv . 1612 . 00593.
[4] Yue Wang et al. Dynamic Graph CNN for Learning on Point Clouds. June 11, 2019. doi: 10.48550/arXiv. 1801.07829.

Training Details

- All networks were implemented in PyTorch Lightning;
- 6 models were trained, one for each model architectures, with and without timing information;
- Hyperparameters of each network optimised using Optuna;
- Events are synthesised from single hadron shower events:
  - cuts applied to remove punch-through pions, and for containment.
  - synthetic neutral hadrons produced using topological cut to remove minimum ionising track;
  - average shower distance chosen so that 80% of average shower energy integrated at that distance (20cm);
  - showers 'overlayed' from single hadron showers;
- Loss function modified from GravNet paper [2]:

$$\mathcal{L} = \sum_k \frac{\sum_i \sqrt{E_i t_{ik}} (p_{ik} - t_{ik})^2}{\sum_i \sqrt{E_i t_{ik}}}$$

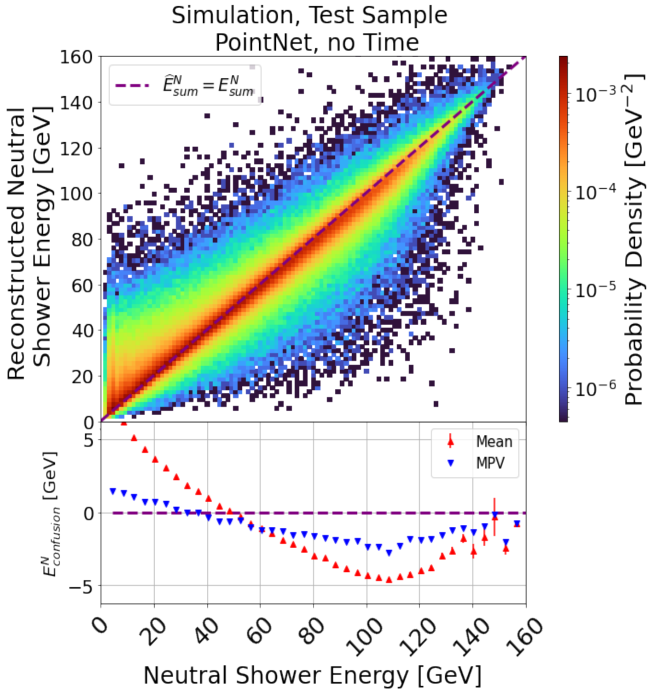
$t, p =$ true/predicted sensor energy fraction
$i, k =$ index of active cell/shower
$E_i =$ active cell energy of sensor i

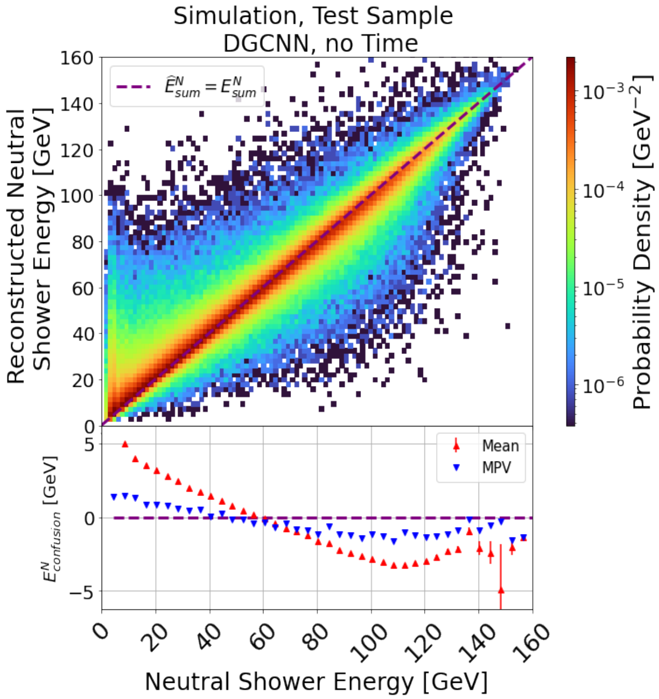
Simulation Properties	
Particle	π^- (negative pion)
Software	Geant4, dd4HEP, CALICESoft
Physics List	QGSP_BERT_HP
Based On	June 2018 SPS Testbeam
Particle Energies	5-120 GeV in steps of 5 GeV

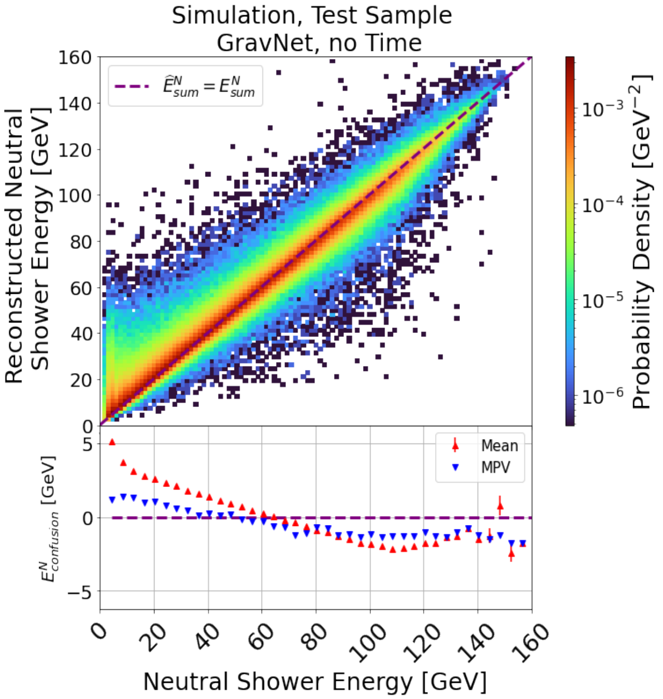
Samples of Charged-Neutral Pairs

Sample	# Events	#Events/ Particle Energy Combination
Training	7.2×10^6	~1250
Validation	8.0×10^5	~140
Testing	8.0×10^6	~1400

Results: Linearity







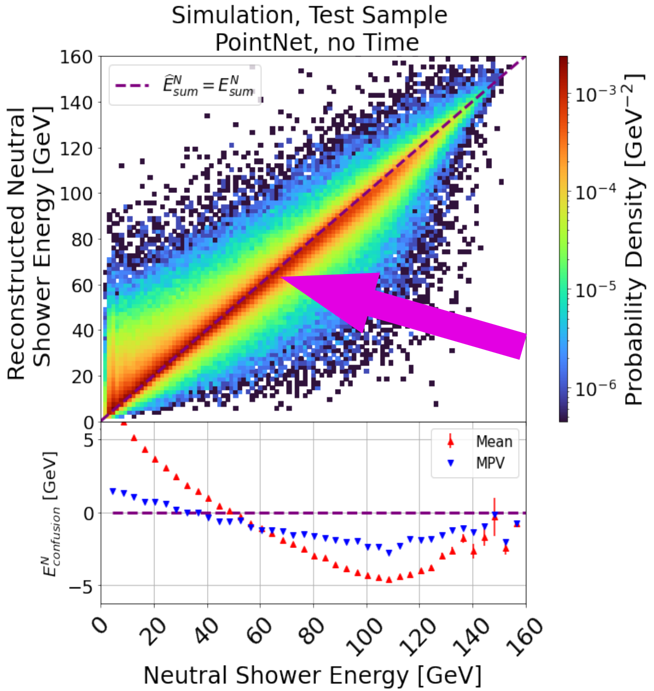
What is shown:

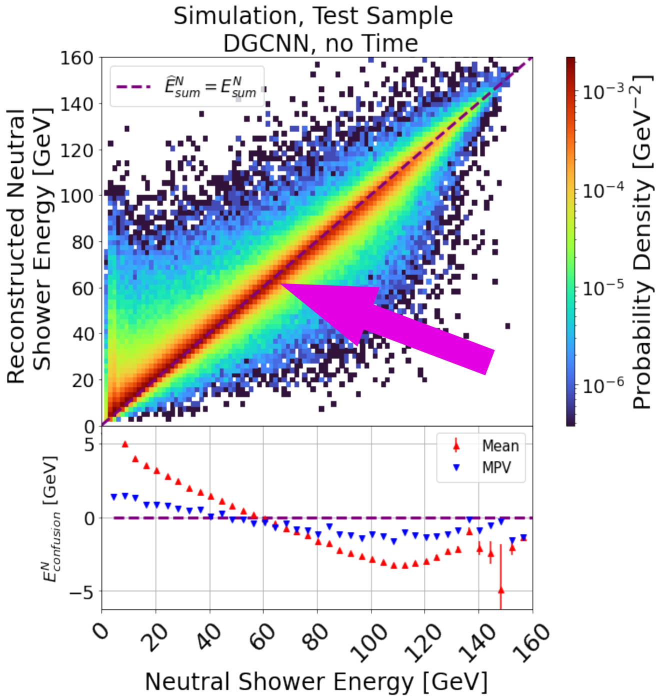
- **Plots:** 2D histograms of reconstructed vs. true neutral shower energy for each model.
- **Subplots:** 'most probable value' (MPV) and mean confusion energy.

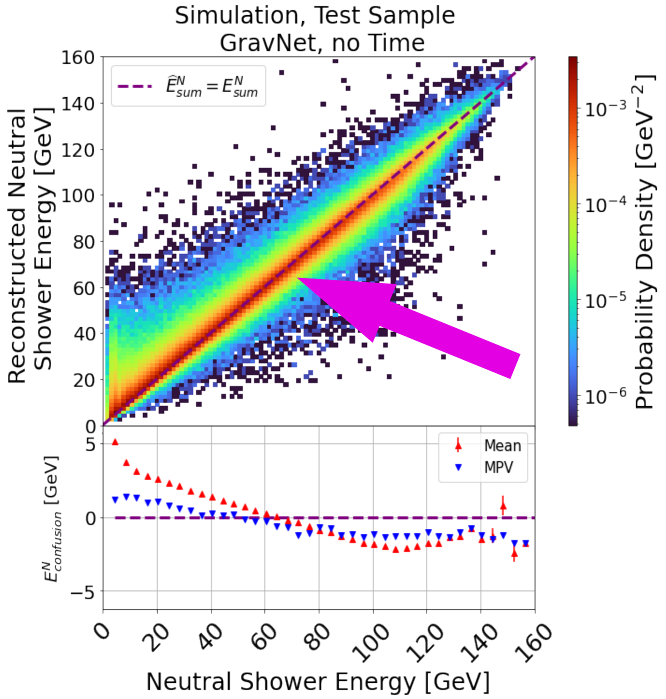
Reconstructed - True

$$E^N_{\mathrm{confusion}} = \hat{E}^N_{\mathrm{sum}} - E^N_{\mathrm{sum}}$$

Results: Linearity



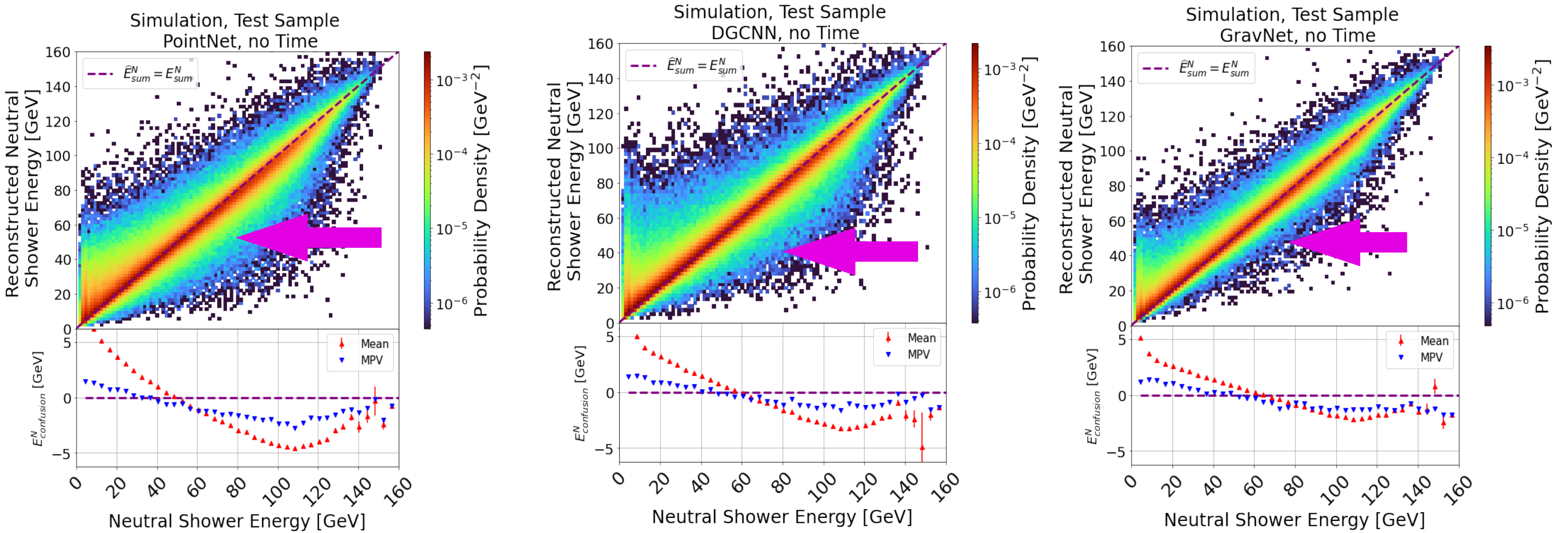




What we learn:

- Red Region close to purple dashed line and MPV of subplot:
  showers frequently well reconstructed (i.e. confusion energy most probably close to 0 GeV)

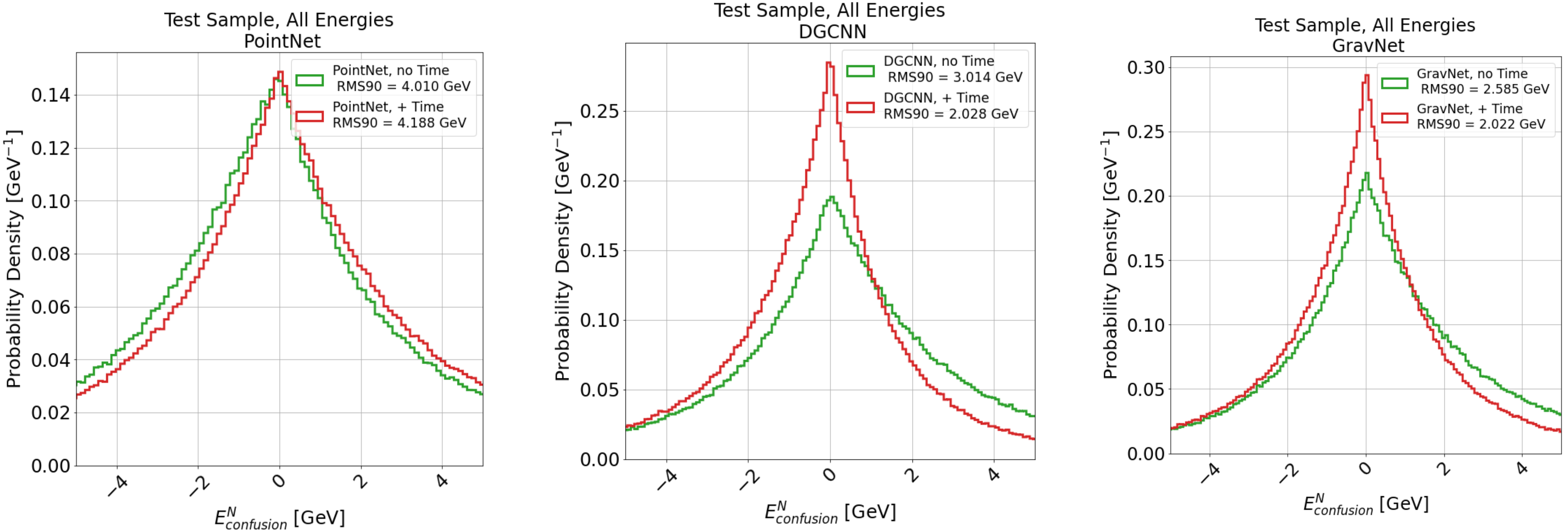
Results: Linearity



What we learn:

- Asymmetric green region and difference between MPV/Mean in subplot: distributions show skewness → there exists a bias in reconstruction.

Results: Confusion Energy Distributions



What is shown:

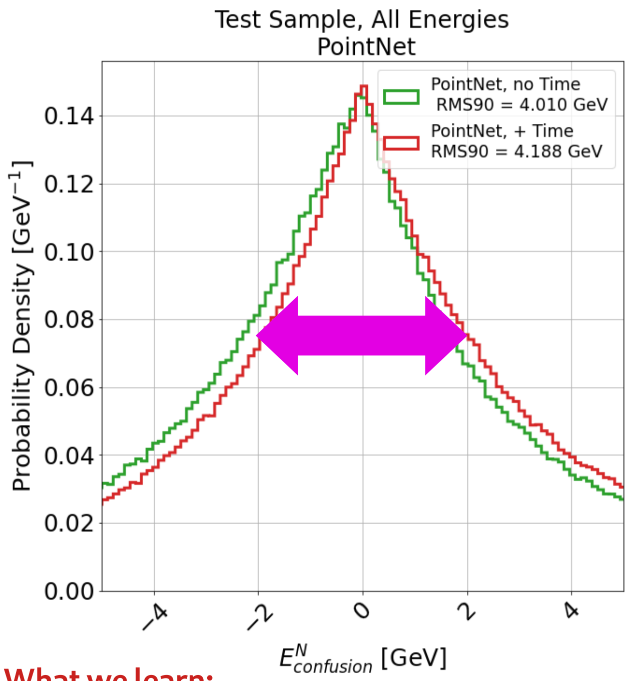
Plot: distributions of neutral confusion energy from each network. Spread measured with RMS_{90} and median absolute deviation (MAD)

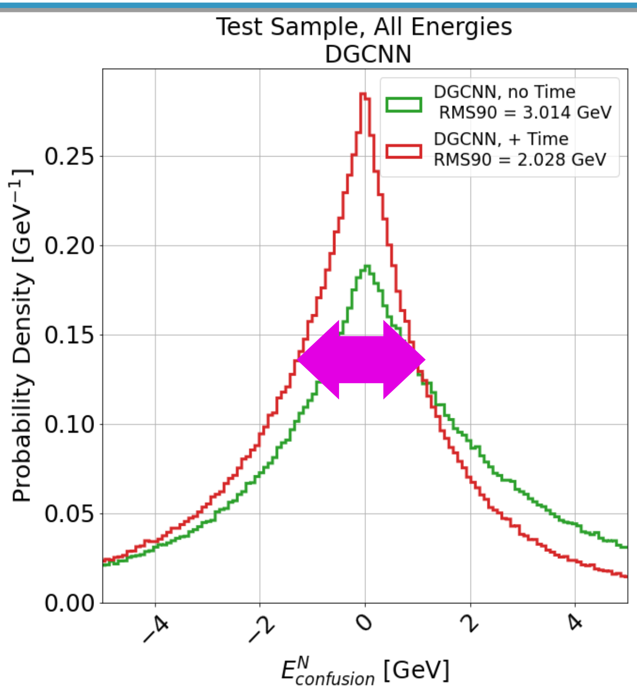
- **neutral network, no timing information applied** (green);
- **neural network, with timing information applied** (red).

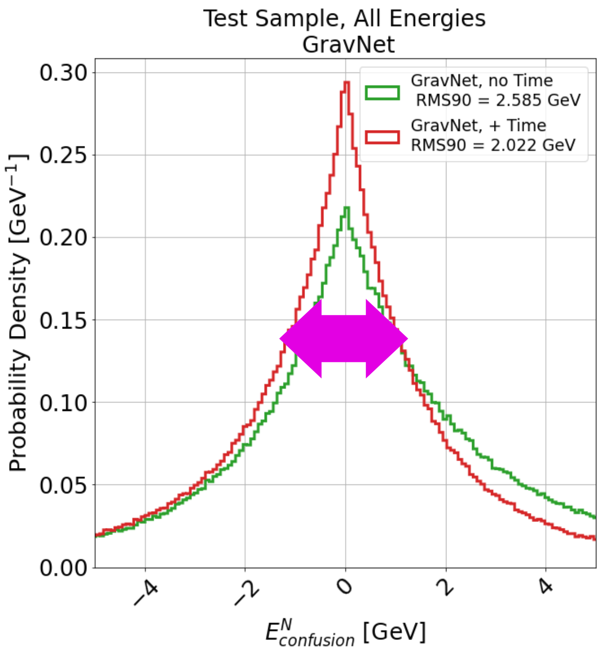
Reconstructed - True

$$E^N_{\text{confusion}} = \widehat{E}^N_{\text{sum}} - E^N_{\text{sum}}$$

Results: Confusion Energy Distributions







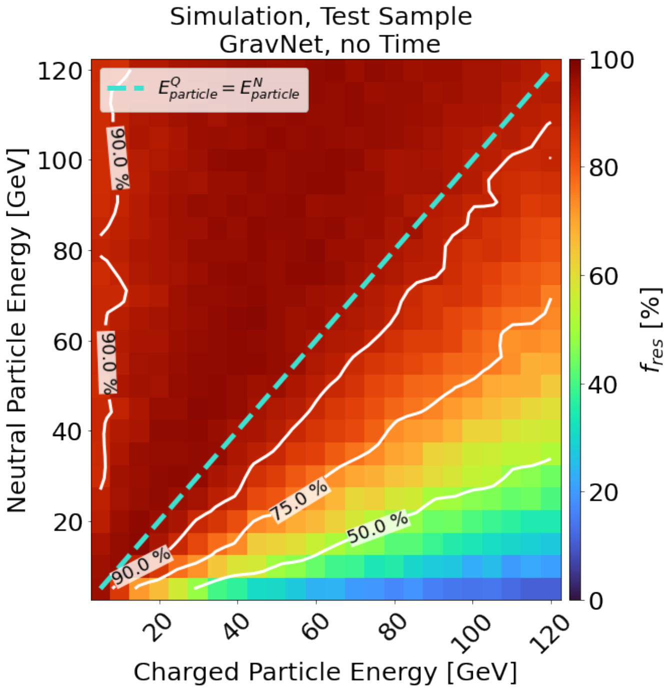
What we learn:

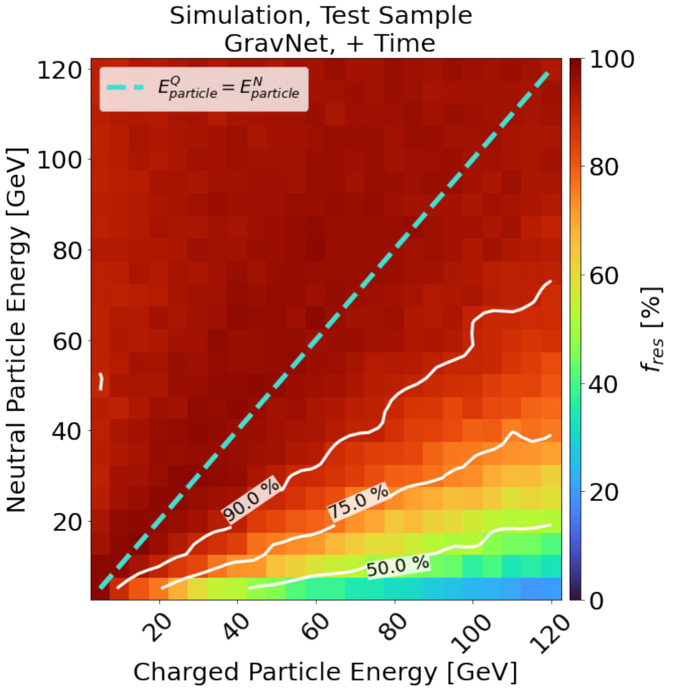
- No improvement using timing information in PointNet.
- DGCNN and GravNet show significant improvement using timing information;

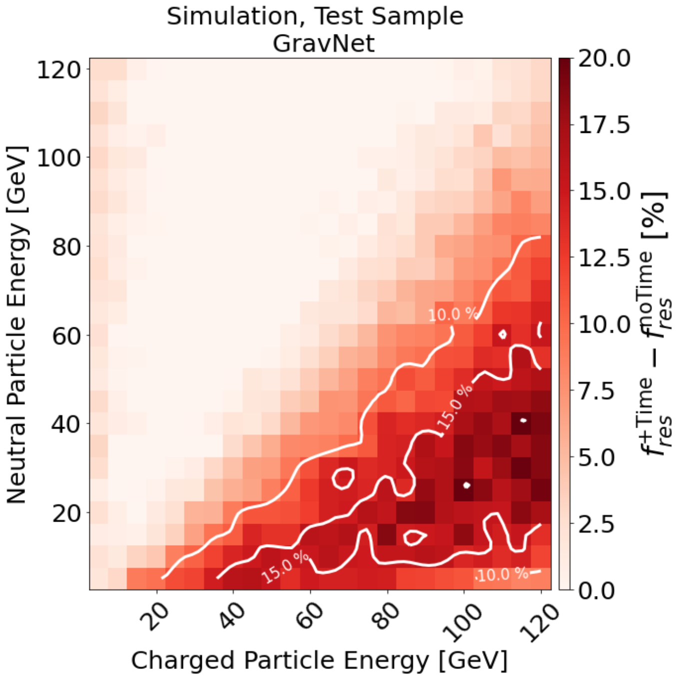
  → **21% reduction in MAD using time with GravNet**
- → **35% reduction in MAD using time with DGCNN**

- Quote from DGCNN Paper: *"Instead of working on individual points like PointNet, we exploit local geometric structures…'"*
- Tentative hypothesis:
  timing information provides a richer description of 'local energy density' (subshowers, decays, etc.)

Results: Clustering Performance







What is shown:

Left, Middle:
Matrices of the fraction of events of the sample reconstructed within the resolution of the AHCAL.

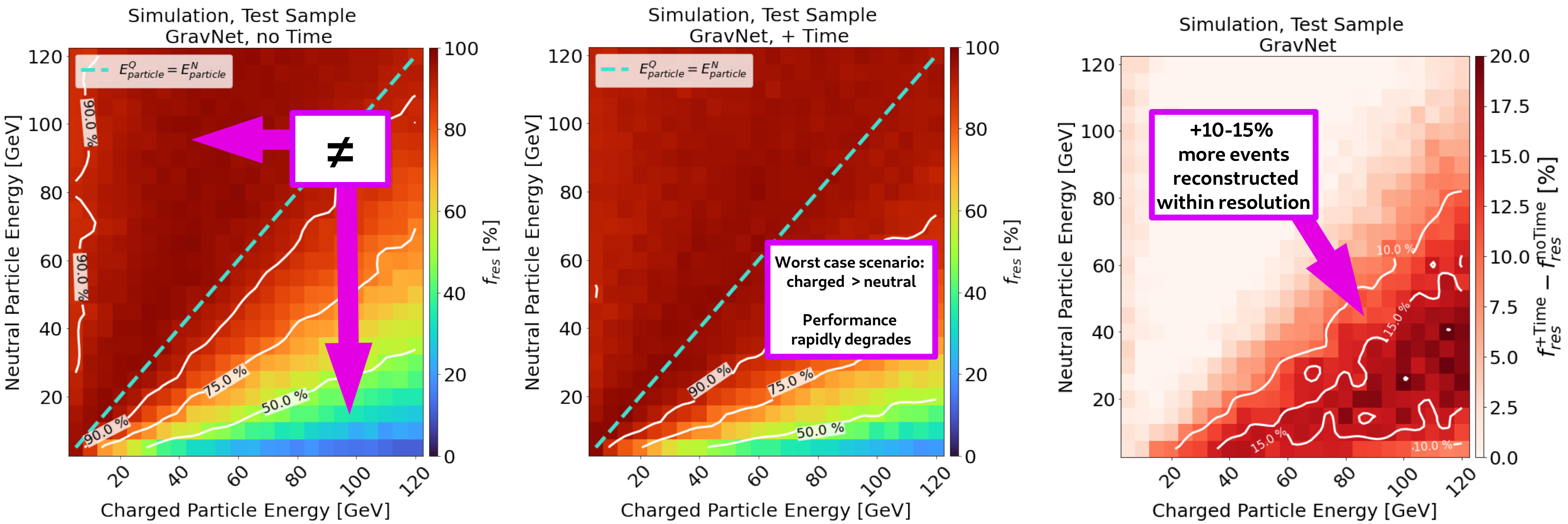
Right:
Additional events reconstructed with than without time.

Red means more events are reconstructed with time than without it.

Resolution of AHCAL In Simulation:

R = 49%/√E ⊕ 7%

Results: Clustering Performance



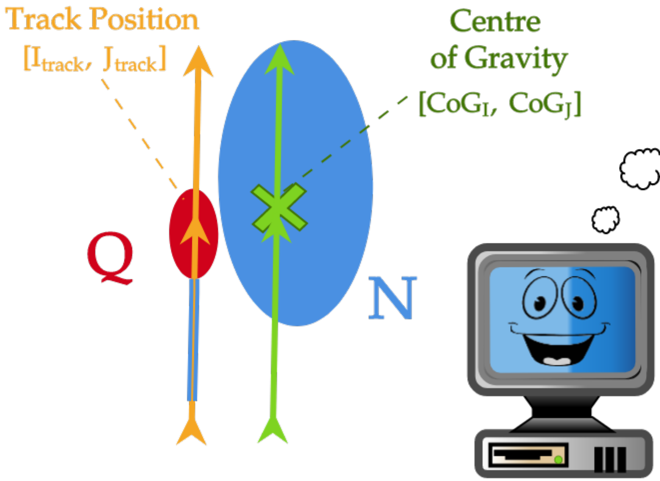
What we learn:

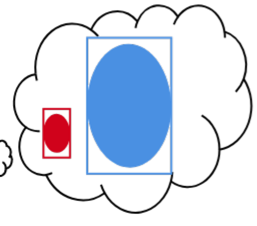
- Clustering performance depends on particle energy combination.
- Result indicates use of track to cluster charged particle → see next slide.
- Timing information helps significantly with the most challenging case of charged > neutral shower energy.

Results: Clustering Performance

$E^{N}_{\text{particle}} > E^{Q}_{\text{particle}}$

Track Position $\neq$ Centre of Gravity

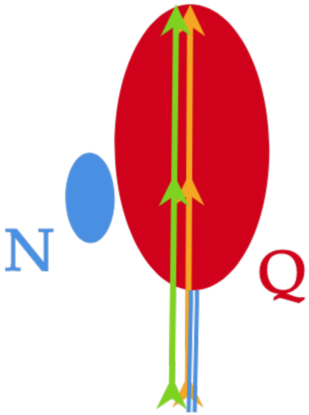


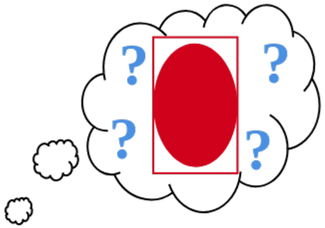


AI:
"Excellent! I can cluster around two axes, one of which is definitely my charged shower."

$E^{Q}_{\text{particle}} > E^{N}_{\text{particle}}$

Track Position $\simeq$ Centre of Gravity





AI:
"Darn! The track information is close to the centre of gravity! That only gives me one axis for clustering!

Now I have to figure out where the neural shower is... "

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Results: Reconstruction Bias

HIGH-D

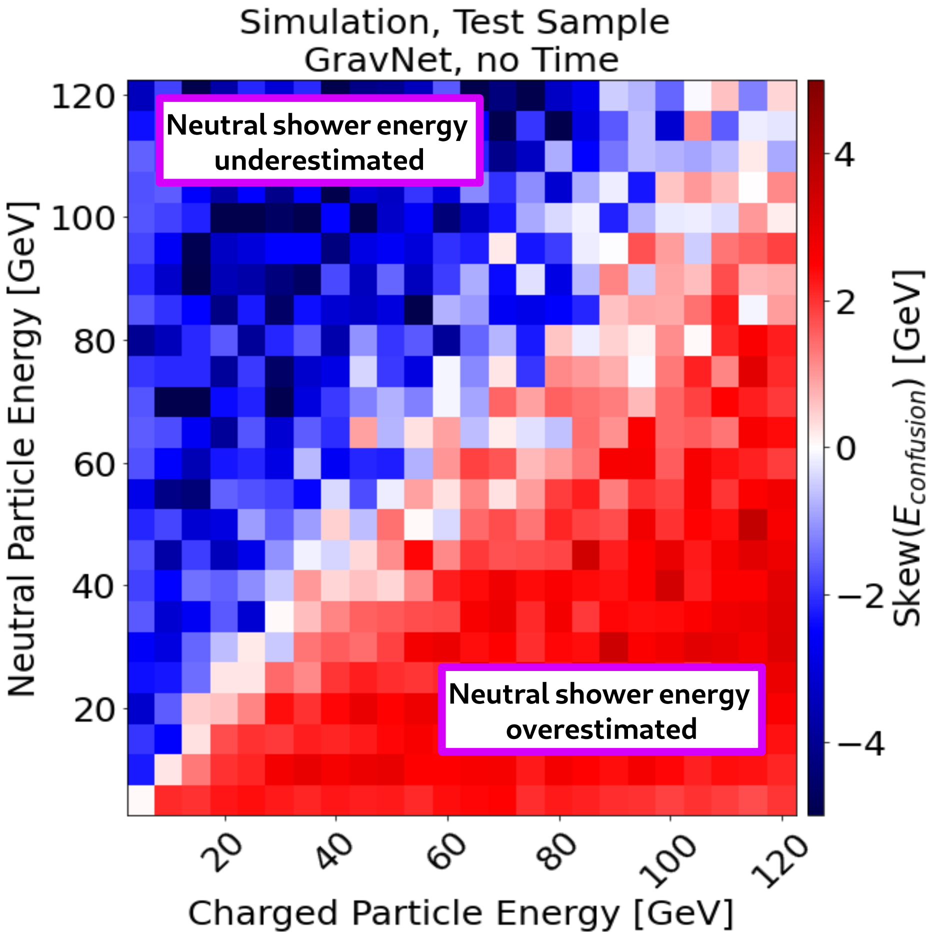
What is shown:

- Skewness of the confusion energy distributions as a function of particle energy;
- Skewness is the third statistical moment, and describes "*asymmetry of the distribution about it's mean*"
- Blue means left-tailed distribution → neutral shower energy **underestimated**
- Red means right-tailed distribution→ neutral shower energy **overestimated**

What we learn:
All studied networks more frequently donate energy from the shower with more energy to the one with less;

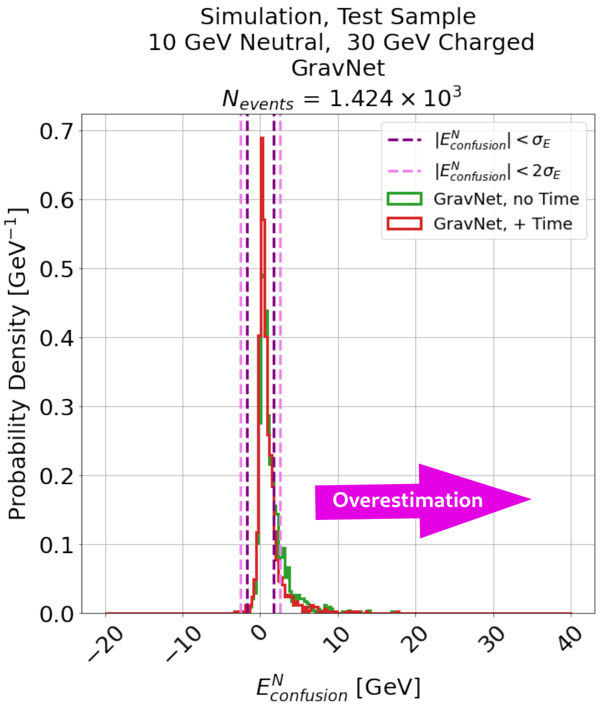
Reconstructed - True

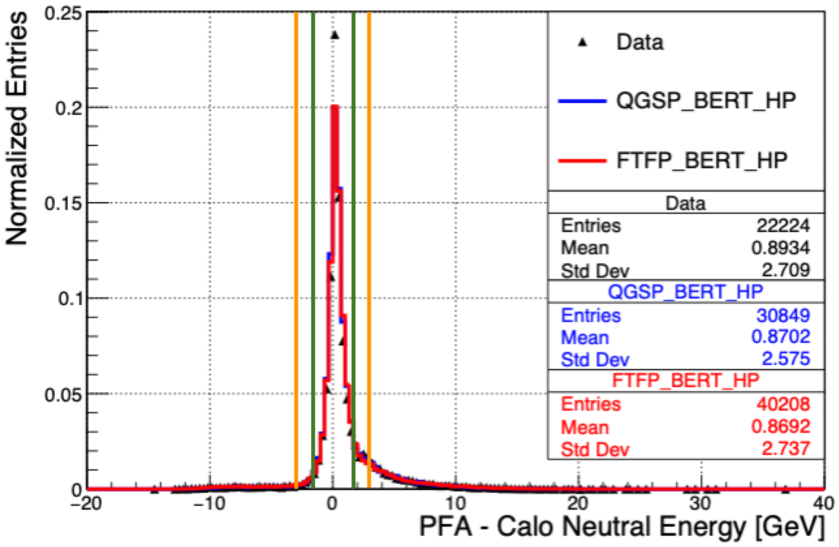
$$E^N_{\text{confusion}} = \widehat{E}^N_{\text{sum}} - E^N_{\text{sum}}$$



Results: Reconstruction Bias

- Note, from [1], on Pandora PFA: ***"...by design the initial clustering stage errs on the side of splitting up true clusters rather than merging energy deposits";***
- Very good agreement between network and in Pandora PFA using AHCAL [5];
- Suggests networks have learned similar strategy as state-of-the-art method;



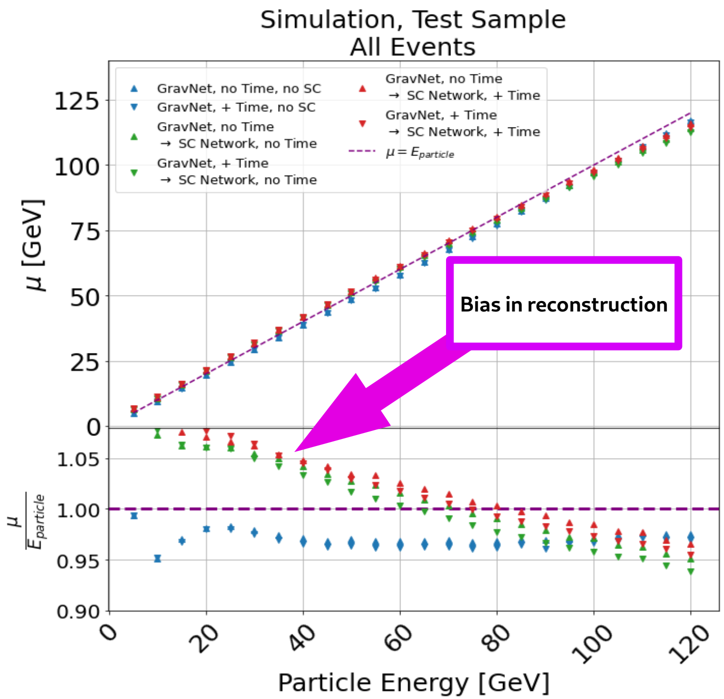


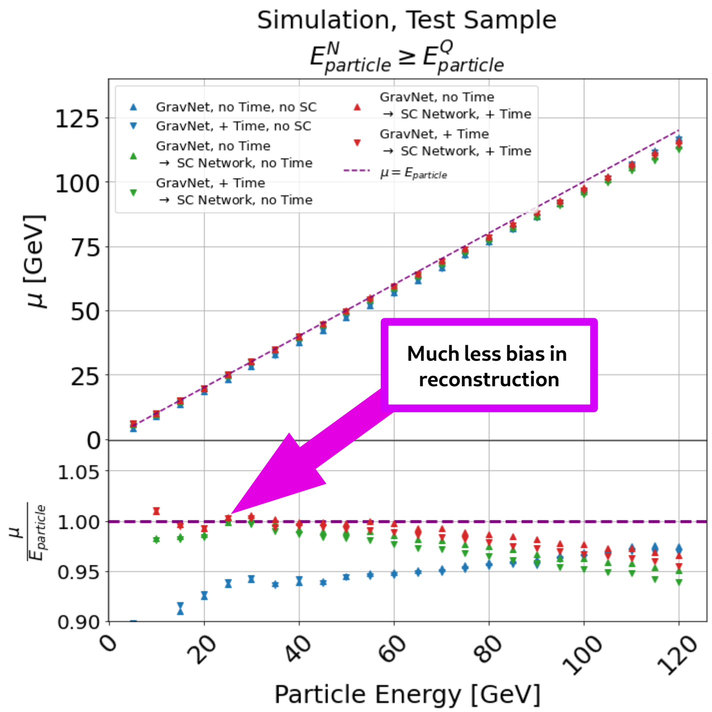
Reconstructed - True

$$E^N_{\text{confusion}} = \widehat{E}^N_{\text{sum}} - E^N_{\text{sum}}$$

[5] Daniel Heuchel. 'Particle Flow Studies with Highly Granular Calorimeter Data'. Place: Heidelberg. Dissertation. 2022. doi: 10.11588/heidok.00031794.

Results: Software Compensation





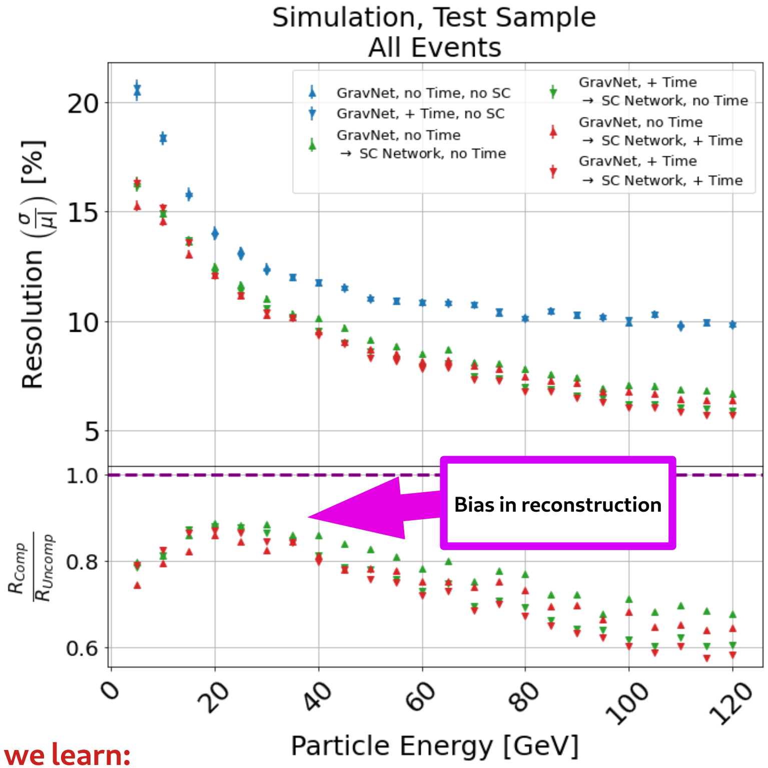
What we learn:

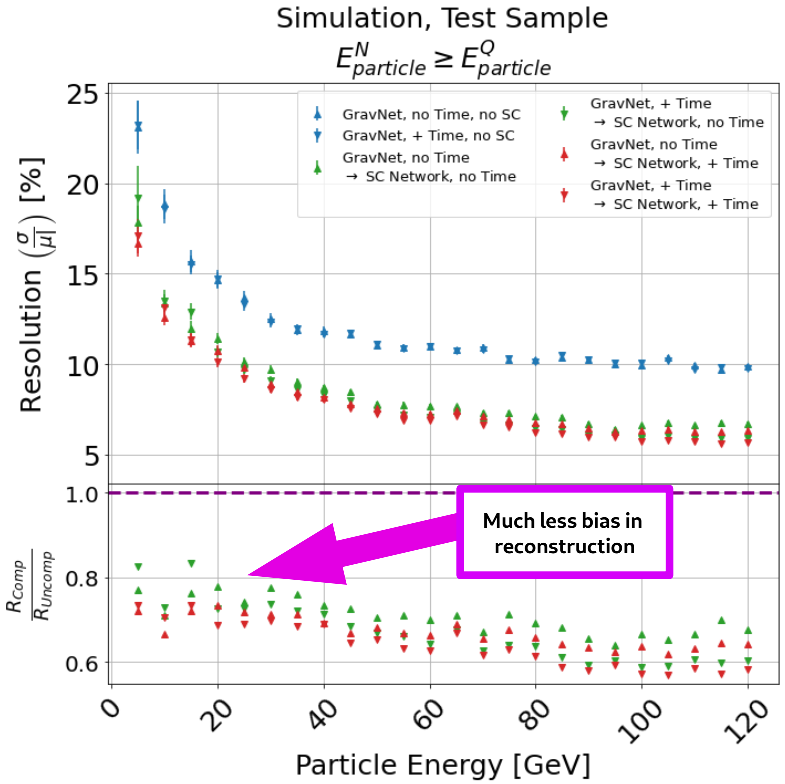
- ML software compensation can be applied after shower separation in the majority of cases.
- Confusion plays a role in the linearity of response.
- Confusion cannot be neglected from the training of SC algorithms → solution is to train the models together.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Results: Resolution

HIGH-D





What we learn:

- ML software compensation can be applied after shower separation in the majority of cases.
- Confusion plays a role in the linearity of response.
- Confusion cannot be neglected from the training of SC algorithms → **solution is to train the models together.**

Conclusion

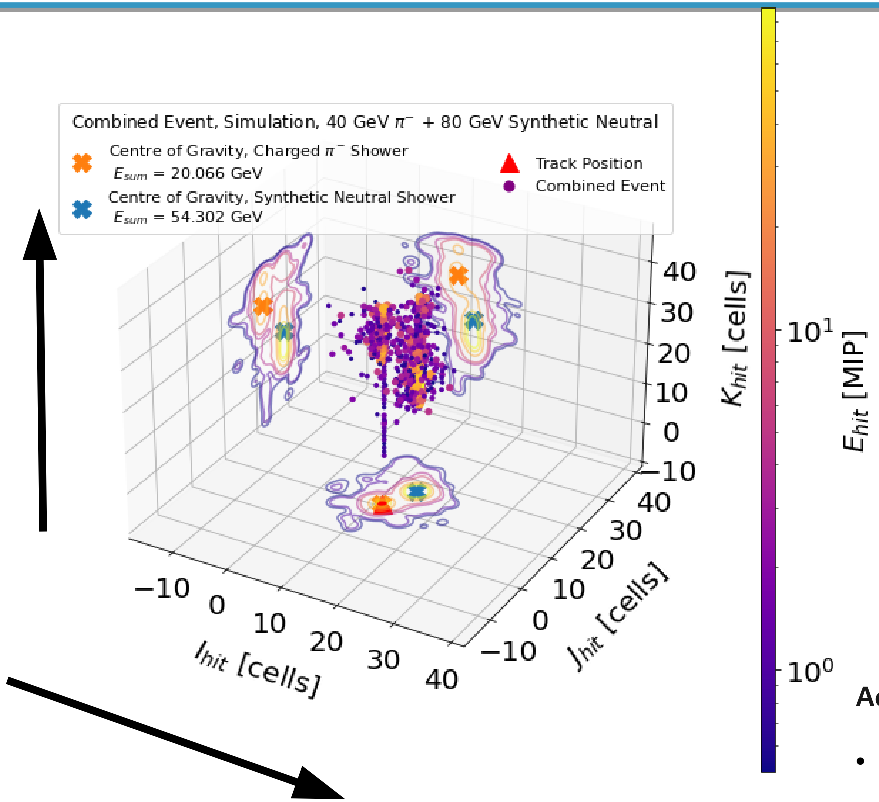
- Shower separation is critical to the performance of Particle Flow Calorimetery;
- Several neural network models were implemented for shower separation in AHCAL, exploiting track and timing information
- The following observations are made:
  - graph neural networks produce superior results to point-based model;
  - strong evidence for use of topological track clustering learned by models;
  - where neutral > charged particle energy:
    - more than 90% of events reconstructed within AHCAL resolution, with or without timing information;
  - where charged > neutral particle energy:
    - +10-15% more events are reconstructed within AHCAL resolution, if 100 ps timing resolution is available;
  - all neural networks prefer to separate clusters of energy than merge them, similarly to Pandora PFA.

- MAIN RESULTS:
  - **A strong case exists for the AHCAL temporal calorimeter→ improvement for charged > neutral particle energy;**
  - **Neural networks learn similar clustering strategies to Pandora PFA.**

What constitutes an ‘event’ for AHCAL?”

Z Axis:

- Layers of absorber, active material and sensors (cells);
- 38 active layers.



X-Y Axes:

- Matrices of sensors (cells);
- 24 x 24 cells per layer

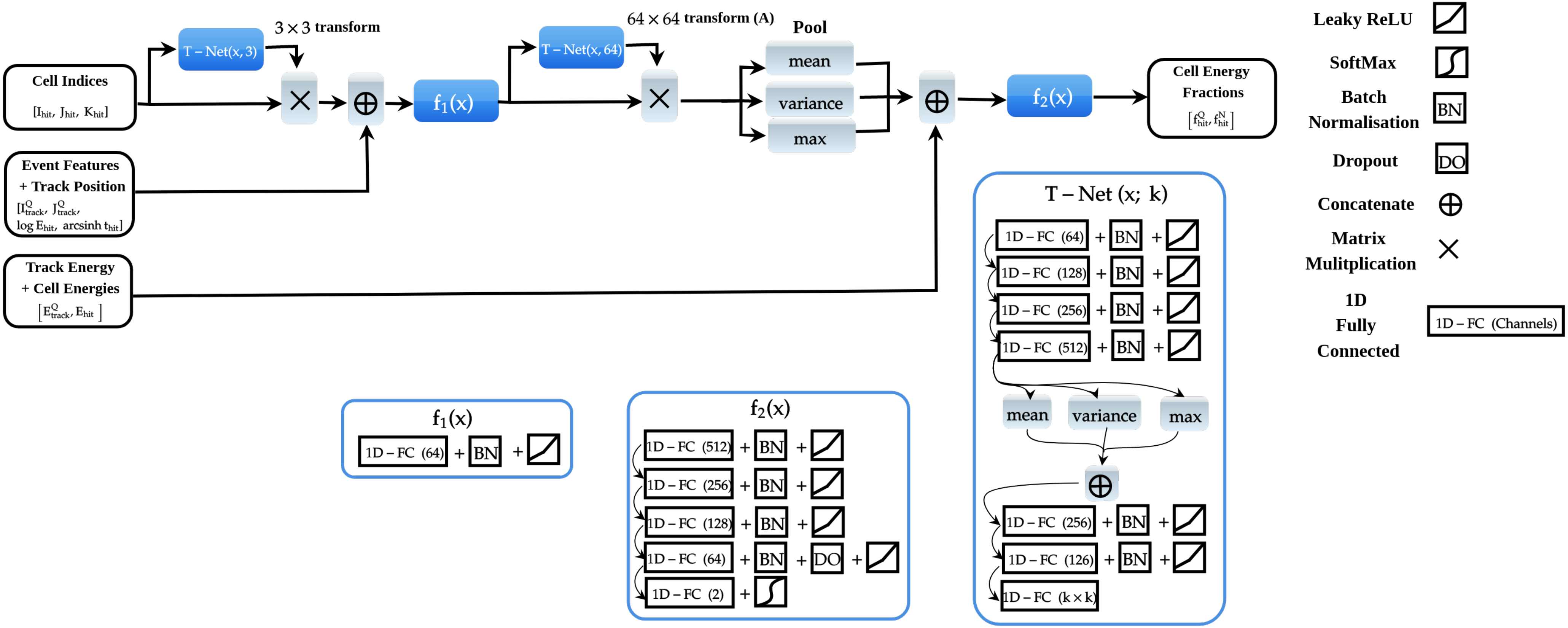
Color Axis:

- Energy of cell, in muon-calibrated ‘minimum ionising particle’ (MIP) units;
- 22,000 cells altogether;
- Sum of all the cell energies → reconstructed energy of hadron;

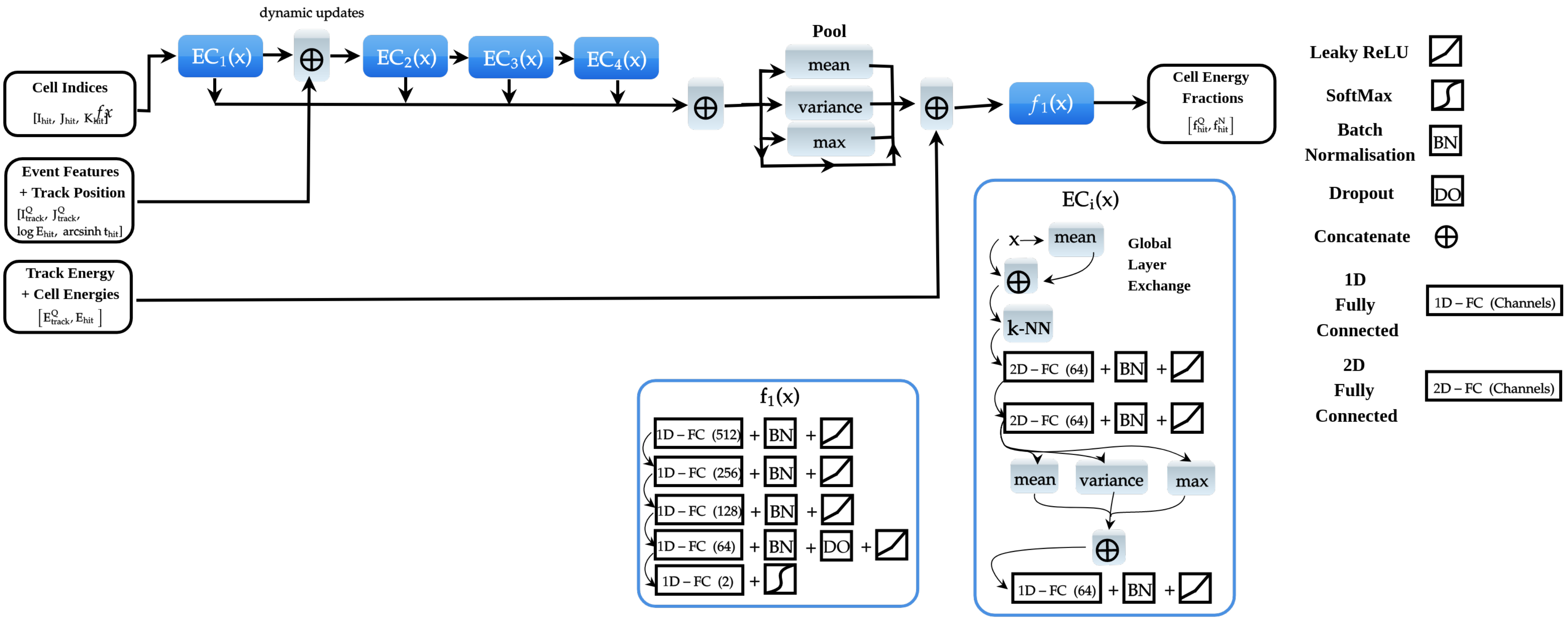
Additionally:

- Timing information for each cell in nanoseconds;
- Not shown in this event display.

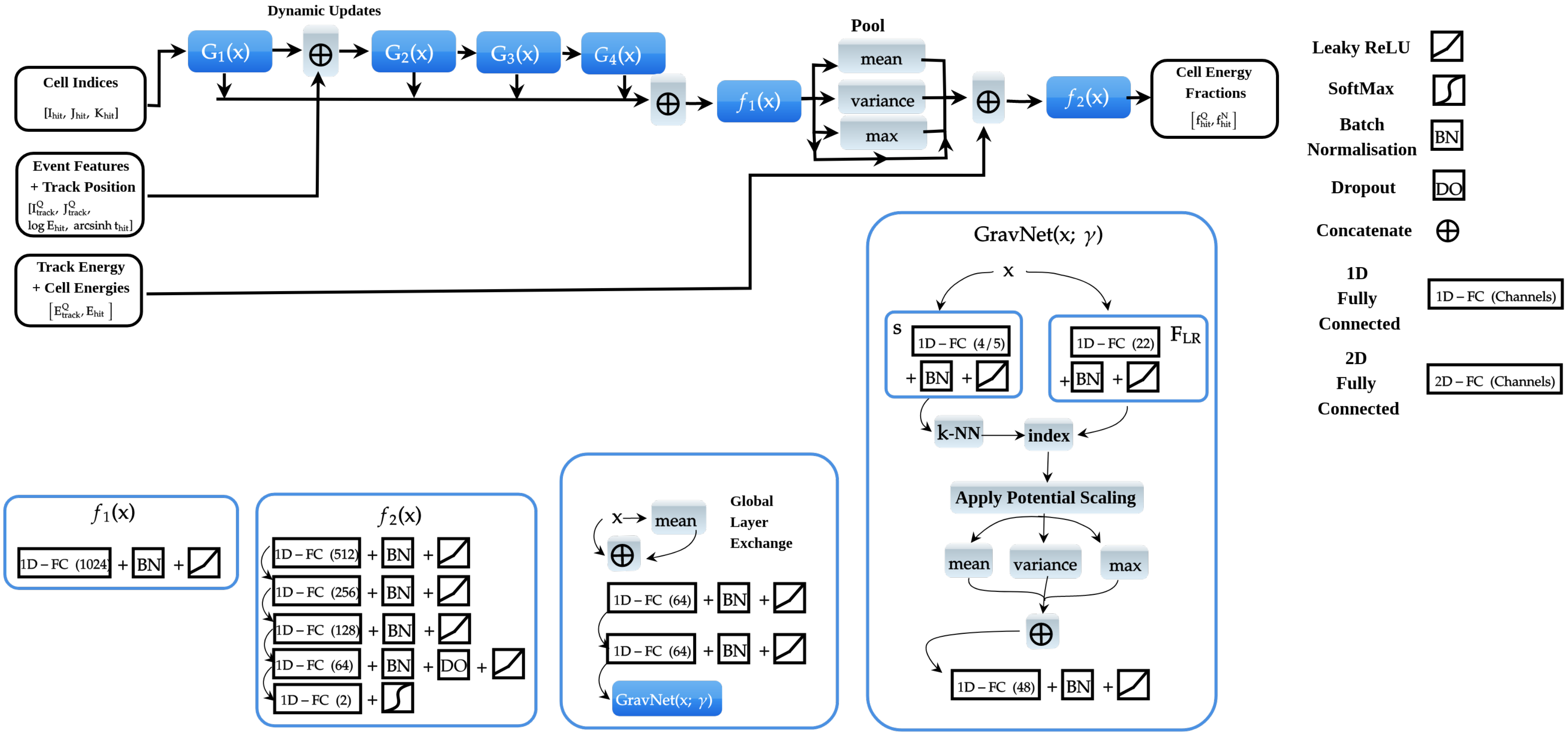
PointNet Model



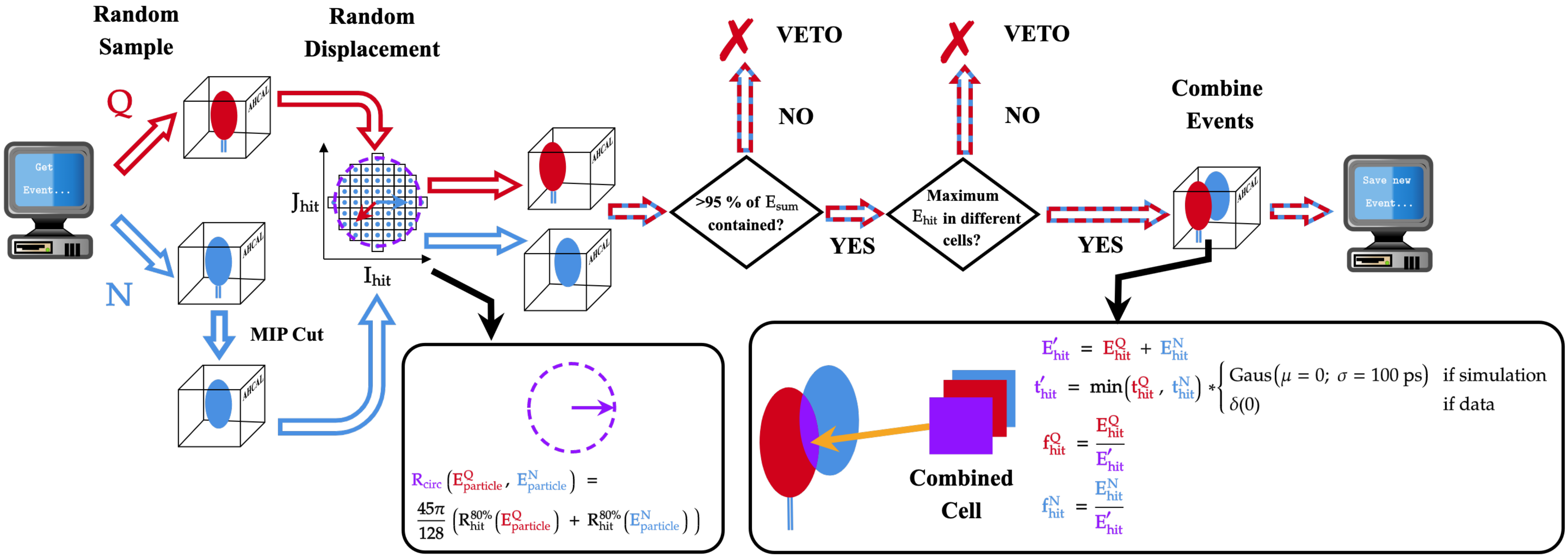
DGCNN Model


dynamic updates
Cell Indices
[I_hit, J_hit, K_hit]
Event Features
+ Track Position
[I^Q_track, J^Q_track,
log E_hit, arcsinh t_hit]
Track Energy
+ Cell Energies
[E^Q_track, E_hit]
EC1(x)
EC2(x)
EC3(x)
EC4(x)
Pool
mean
variance
max
f1(x)
Cell Energy
Fractions
[f^Q_hit, f^N_hit]
ECi(x)
x
mean
Global
Layer
Exchange
k-NN
2D – FC (64)
BN
2D – FC (64)
mean
variance
max
1D – FC (64)
f1(x)
1D – FC (512)
1D – FC (256)
1D – FC (128)
1D – FC (64)
DO
1D – FC (2)
Leaky ReLU
SoftMax
Batch
Normalisation
Dropout
Concatenate
1D
Fully
Connected
1D – FC (Channels)
2D
Fully
Connected
2D – FC (Channels)

GravNet Model



Synthetic Shower Overlay Algorithm

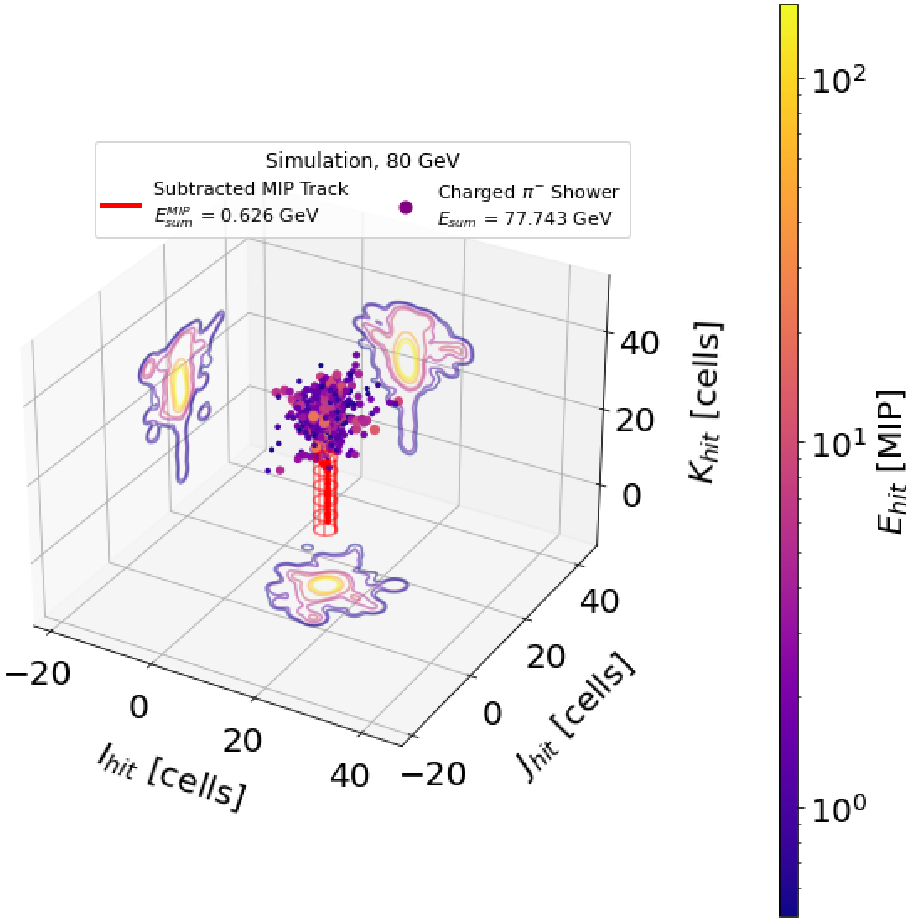


Synthetic Neutral Showers

Charged particles ionise dense matter. Neutral particles do not.

No difference is expected between hadron showers caused by neutral and charged hadrons after first hard interaction.

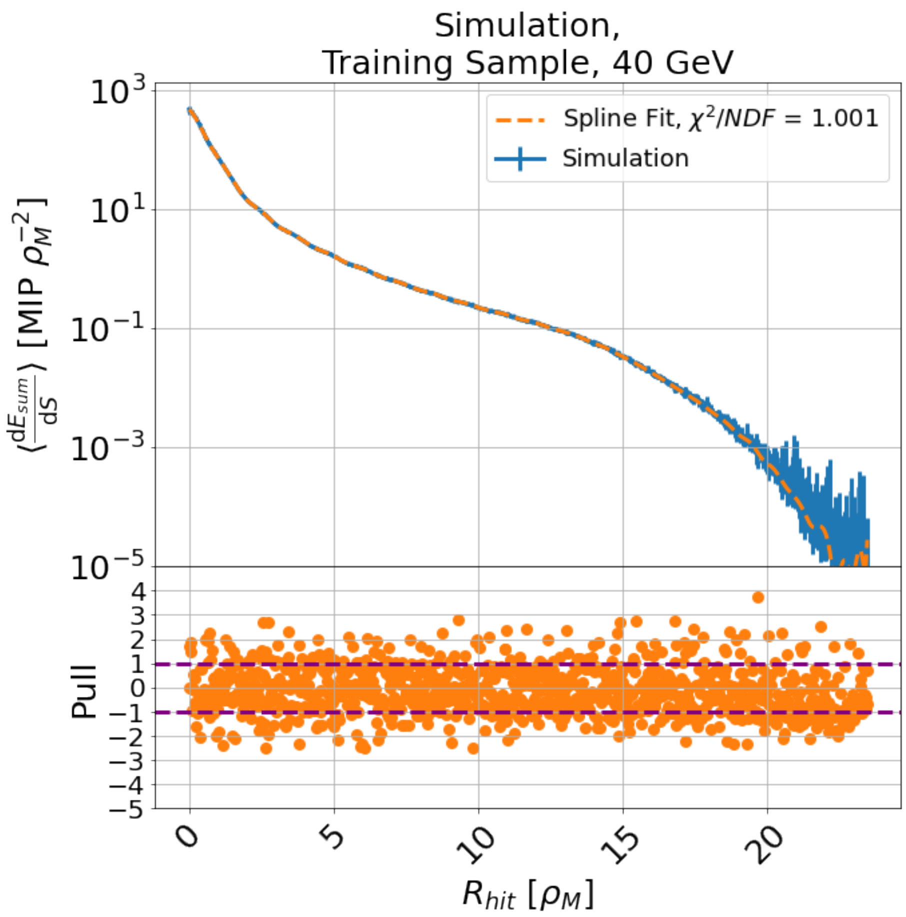
- Synthetic neutral showers can be produced by removing energy deposits from ionisation;
- Achieved using a topological cut on the event.
  - Active cells less than 60 mm from shower 'centre-of-gravity' (energy-weighted mean)
  - Active cells observed less than one layer before the shower starting layer;
  - Active cells with less than 3 calibrated MIP energy units;
- Preliminary studies using event classifier indicate this method is highly effective at producing synthetic neutral showers at the event level.



Training Shower Distance Procedure

Step 1:

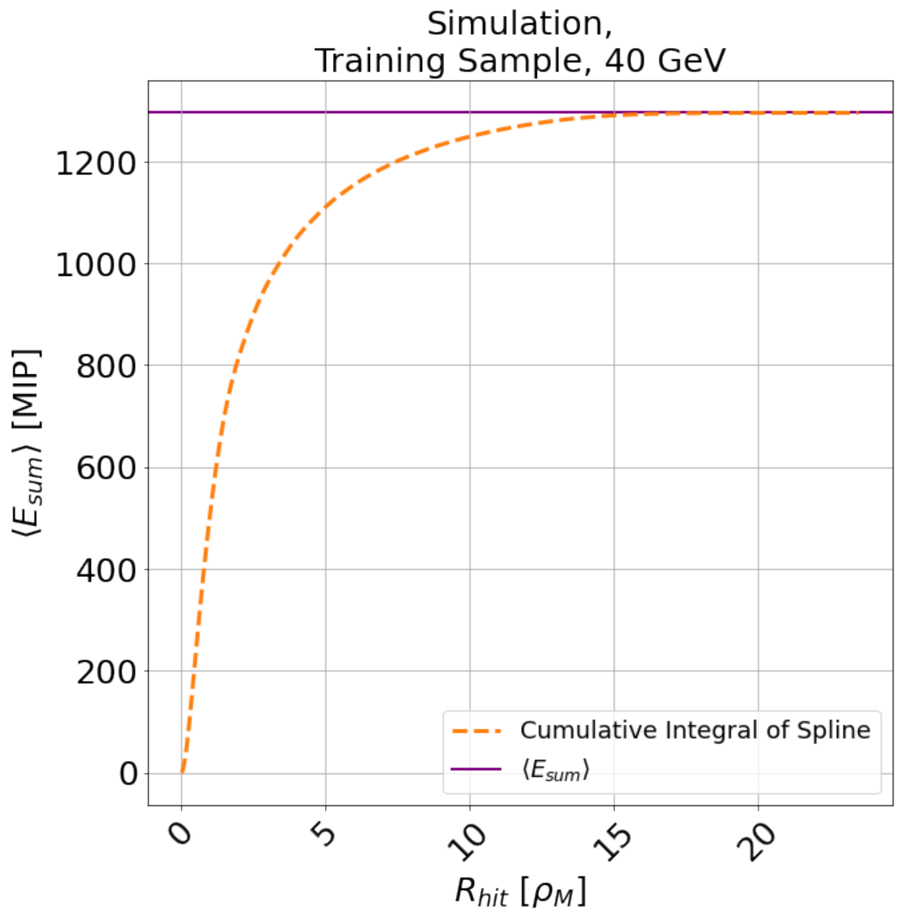
- Calculate differential energy deposited per unit area in a circle of radius R_{hit} around the centre-of-gravity;
- Fit the distribution with a cubic spline.



Training Shower Distance Procedure

Step 2:

- Calculate cumulative energy distribution;
- Currve saturates at the average energy measured by AHCAL for a particular shower energy;



Step 3:

- Distribution is inverted, and evaluated as a function of particle energy
- Curve follows logarithmic relationship:

$$R_{\text{hit}}^{80\,\%}(E_{\text{particle}}) = a_R + b_R \cdot \log E_{\text{particle}} + c_R \cdot E_{\text{particle}}$$

- Apply factor to determine correct radius of circle in which to displace events such that 80% of energy is integrated.
- For any particular combination of energies, the average shower distance is approximately 20 cm