Software Compensation for Highly Granular Calorimeters using Machine Learning

Jack Rolph, Erika Garutti

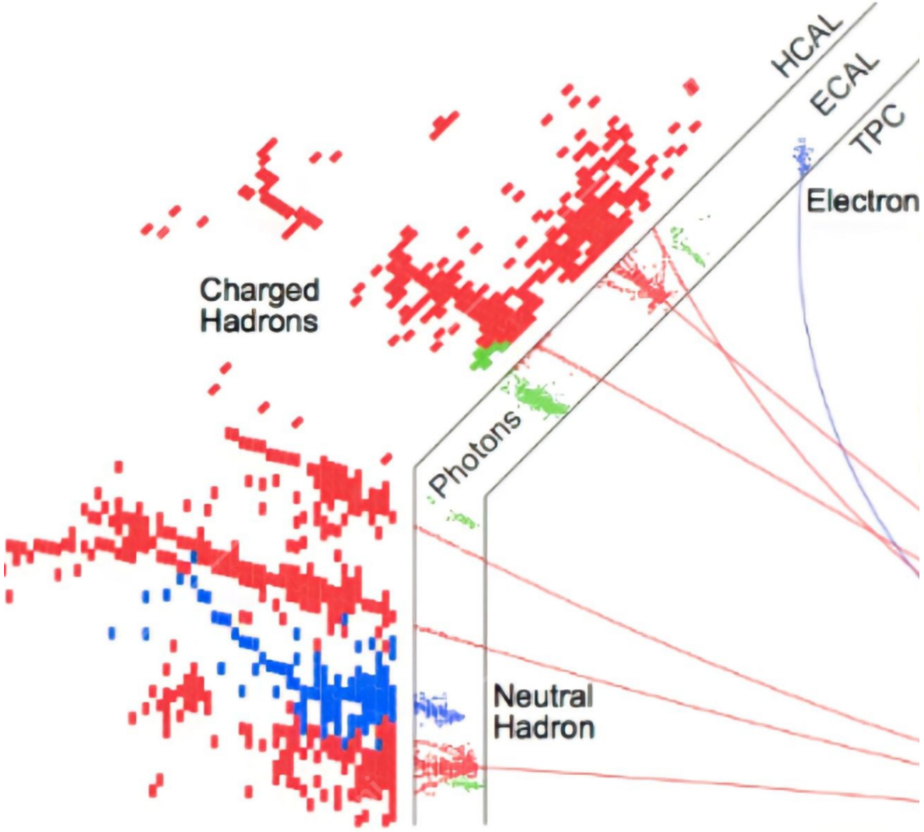
Institute for Experimental Physics, University of Hamburg

CALICE Collaboration Meeting, Göttingen

31/03/2023

Particle Flow Calorimetry

- Jet energy resolution at future precision e^+-e^- colliders such as ILC must produce a di-jet invariant mass resolution comparable to the weak vector bosons' decay width (~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 remainder of energy of photons and neural hadrons;
  - Use sophisticated clustering algorithms (e.g. Pandora Particle Flow Algorithm, PPFA) to associate tracks to energy depositions in the highly granular calorimeters;



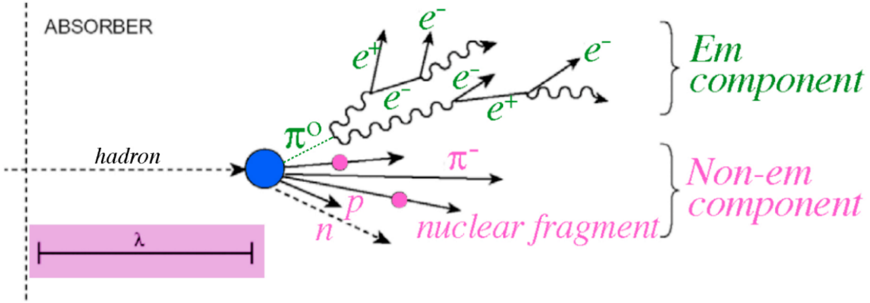
[1] M. A. Thomson. 'Particle Flow Calorimetry and the PandoraPFA Algorithm'. NIMA, pp. 25–40. doi: 10.1016/j.nima.2009.09.009.

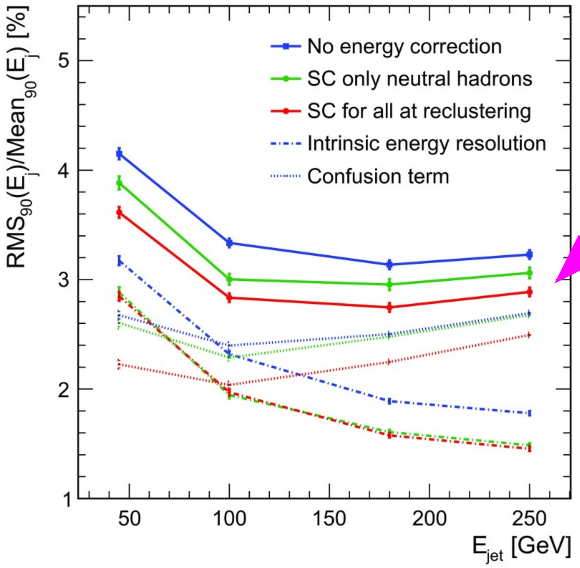
Hadron Calorimetry and Software Compensation

- When a hadron interacts with matter, it may undergo a hard interaction and induce a hadron shower;
- A hadron shower has two components: an electromagnetic (EM) and a hadronic (HAD) fraction;
- Up to 40% of the deposited energy in the HAD fraction is 'invisible', and fluctuates significantly from event to event;
- Calorimeter response to hadrons is split into two components: an EM response (e) and a HAD response (h). Typically, e>h;
- **Software compensation (SC)** is the equalisation of e and h using:
  - information from a recorded event and;
  - a specially designed algorithm.
- SC found to improve jet energy resolution achieved by Pandora PFA by enabling more accurate association of tracks to energy deposits [2];

TAKE HOME MESSAGE:

better SC →
more accurate energy clustering for PFC→
better jet energy resolution →
more precise measurements at future e^+-e^- colliders



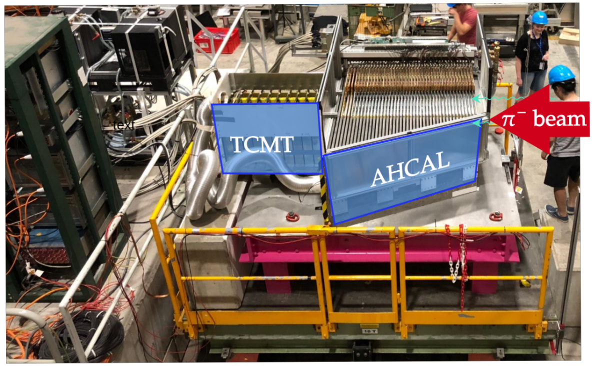


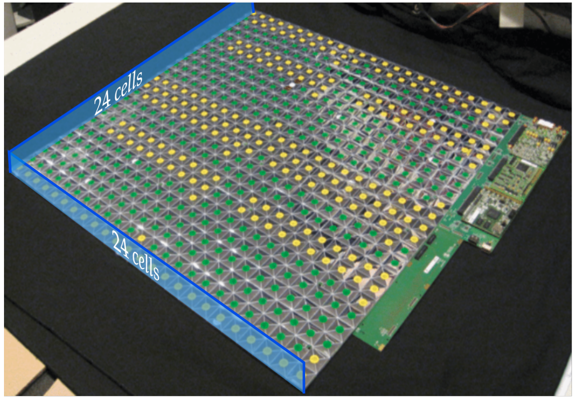
Red line < blue line → compensation improves jet energy resolution.

[2] CALICE Collaboration, "A new approach to software compensation for the CALICE AHCAL," tech. rep., 2010.

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, non-compensating calorimeter (e/π = 1.25-1.4):
- Five-dimensional calorimeter: measures energy density of hadron showers in space *and time*, with **up to 100 ps time resolution allowed by hardware** for each active cell (3 x 3 x 0.3 cm^3).
  - hadron showers develop with a dense EM core and sparse HAD halo →
    AHCAL can exploit spatial development of hadron showers for SC;
  - neutron fraction of hadron shower directly proportional to HAD fraction →
    indirect energy depositions from neutrons *delayed* compared to first nuclear interaction by hadron by 10-100 ns in steel →
    AHCAL can exploit temporal development of hadron showers for SC;





TAKE-HOME MESSAGE:

spatial & temporal energy density information

available from AHCAL may improve SC

Biasing of SC Models

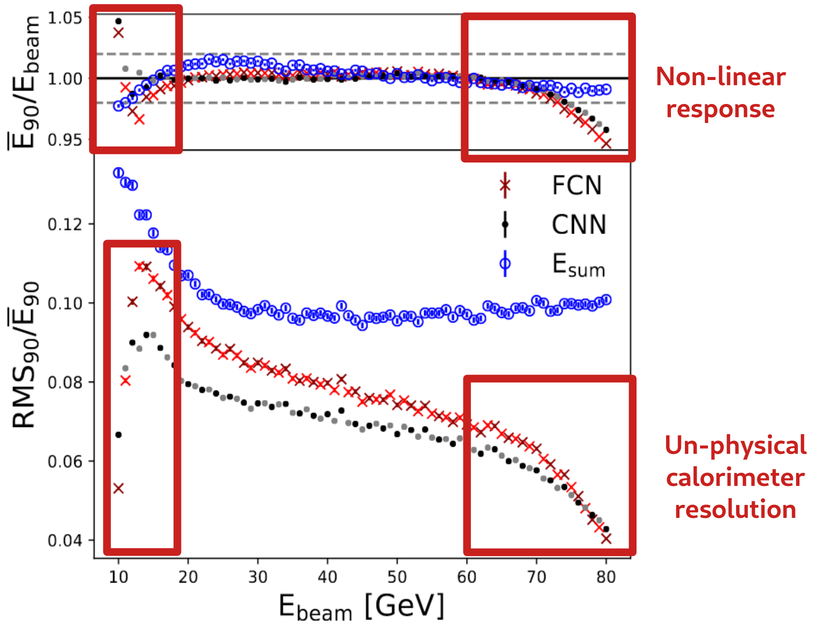
- Machine learning (ML) can be used to build a bespoke software compensation algorithm for a specific detector.

- Previous study using ML to build an SC algorithm for AHCAL [3] showed two undesirable effects:

  - **Interpolation Failure:** biasing of model to specific particle energies;
  - **Extrapolation Failure:** biasing to the lowest and highest of the training range of energies;

- Reason:

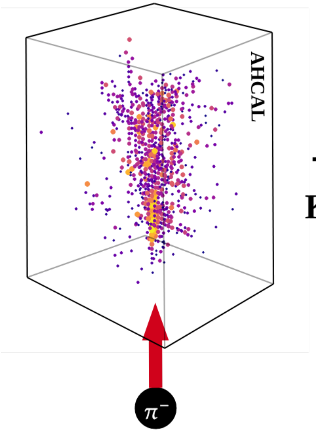
  - training samples of hadron shower data are always biased to the energies of the interacting particle;
  - EM/HAD fraction of event is unknown → inferred from the calorimeter response.
  - **Algorithm easily biases to the energy range upon which it is trained.**



[3] E. Buhmann and E. Garutti, "Deep learning based energy reconstruction for the CALICE AHCAL"

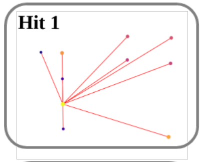
Explicitly Reducing Bias By Model Design

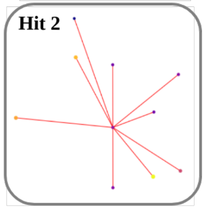
1. Hadron shower event in highly-granular calorimeter is split into individual active cells.

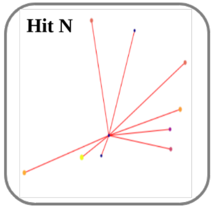


K-NN

2. K-nearest neighbour 'clusters' found for each individual active cell.







3. SC model operates independently on each 'cluster'.

$$\rightarrow f_{SC}(\mathcal{G}_1;\ \theta) \rightarrow \hat{E}_{\text{hit},\,1}$$

$$\rightarrow f_{SC}(\mathcal{G}_2;\ \theta) \rightarrow \hat{E}_{\text{hit},\,2}$$

$$\vdots$$

$$\rightarrow f_{SC}(\mathcal{G}_N;\ \theta) \rightarrow \hat{E}_{\text{hit},\,N}$$

4. Modified active cell 'energy' produced by model for each local cluster.

$$\hat{E}_{\text{sum}} = \sum_{i=1}^{N} \hat{E}_{\text{hit},\,i}$$

5. Compensated hadron shower energy produced from sum of model outputs.

TAKE-HOME MESSAGE:

- **The SC model is explicitly blinded to the mean response of the hadron shower, and cannot bias to it.**
- **The attention of the model is explicitly focused on the local distribution of energy in space and time → highly granular calorimeter information exploited by model**

Model and Training Details

- A basic neural network was implemented in PyTorch using the proposed solution of operating on individual clusters.
- The model was compared to an implementation of the CALICE standard software compensation method for AHCAL [3] (see upcoming CALICE Note for details).
- Control uses total calorimeter response for compensation → expected to be biased.
- 3 models were trained (colour coding shown):
  - **control method;**
  - **network method, without timing information;**
  - **network method, with timing information.**
- Models trained and tested using Geant4 simulation of π^- hadron showers in AHCAL showering in the first 4 layers, + containment cuts (see backup slides):
- Loss function: mean χ^2 goodness-of-fit of calorimeter response to known momentum.

Simulation Properties	
Particle	π^- (negative pion)
Software	Geant4, dd4HEP, CALICESoft
Physics List	QGSP_BERT_HP
Based On	June 2018 SPS Testbeam
Particle Energies	10-120 GeV in steps of 5 GeV

Samples of Charged Hadron Showers

Sample	# Events
Training (10 – 80 GeV, 10 GeV Steps Only)	1.8×10^5
Validation (10 – 80 GeV, 10 GeV Steps Only)	2.1×10^4
Testing (10 – 120 GeV, 5 GeV Steps)	4.0×10^5

[3] CALICE Collaboration, "A new approach to software compensation for the CALICE AHCAL," tech. rep., 2010.

Results: Distributions of Reconstructed Energy

What we expect:

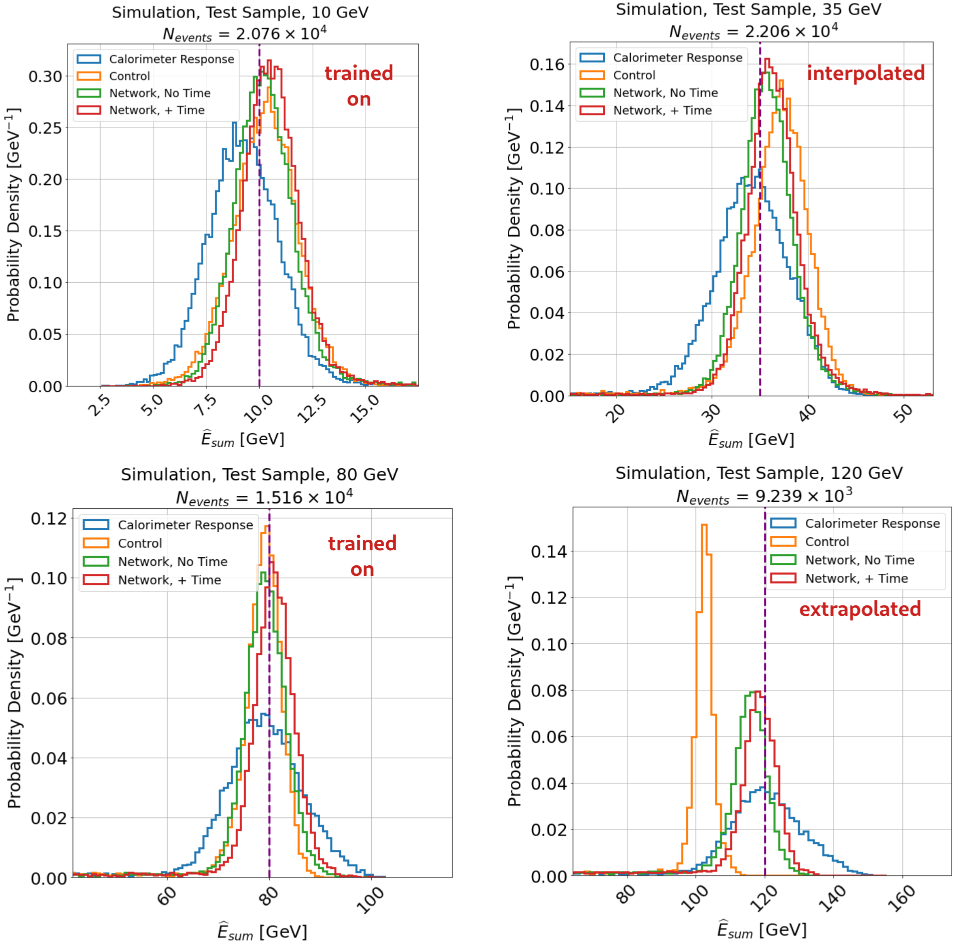
- compensation → the width of the response distributions of AHCAL to hadrons should reduce.
- The ML methods should be able to interpolate/extrapolate SC;

What is shown:

- The distributions of AHCAL response to the testing sample of π^- hadron showers for different particle energies, with:
  - **intrinsic calorimeter response** (blue);
  - **control method applied** (orange);
  - **neural network, no timing information applied** (green);
  - **neural network, with timing information applied** (red).

What we learn:

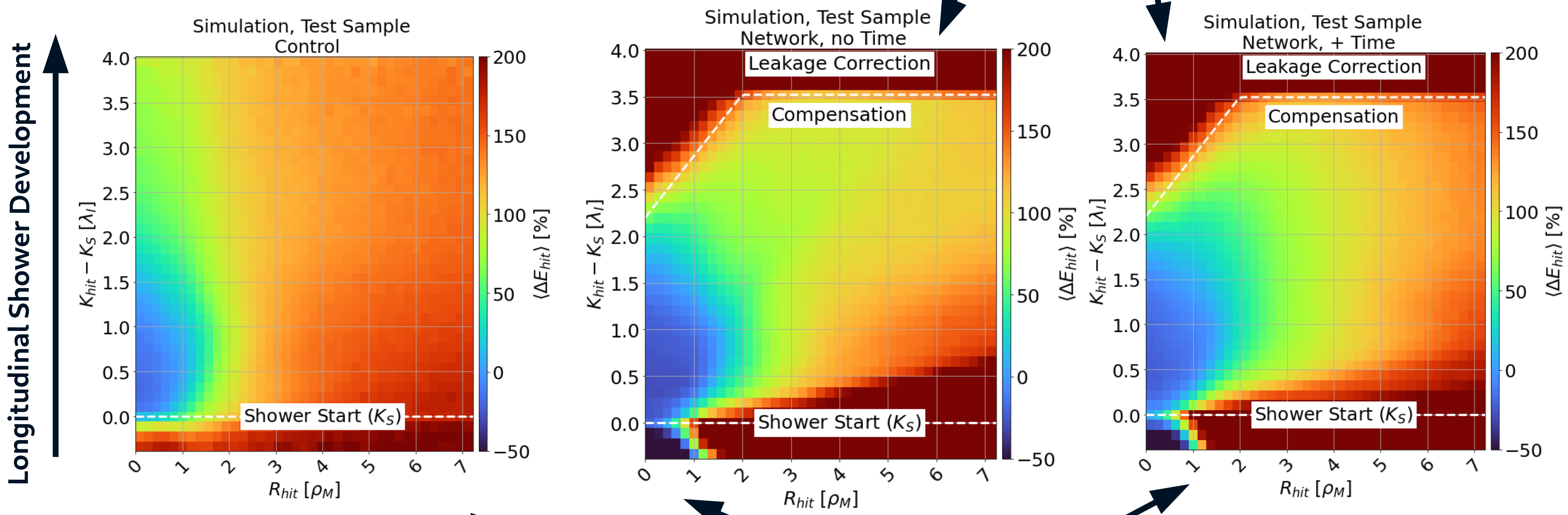
- Network method produces superior resolution for most of training range below < 60 GeV;
- Control method gives superior 'compensation' to neural network in range 60-80 GeV;
- Above 80 GeV, control method seen to bias to training range;
- Neural networks able to compensate effectively above the edge of the training range;



Results: Mean % Change in Cell Energy, Spatial Dependence

Blue: active cell energy attenuated
Red: active cell energy enhanced

leakage correction and stronger spatial dependence of weights



Lateral Shower Development

1 ρ_M = 24.9 mm, 1 λ_I =231.1 mm

MIP track subtraction

Temporal correlations in backup slides.

Results: Linearity of Response

What we expect:

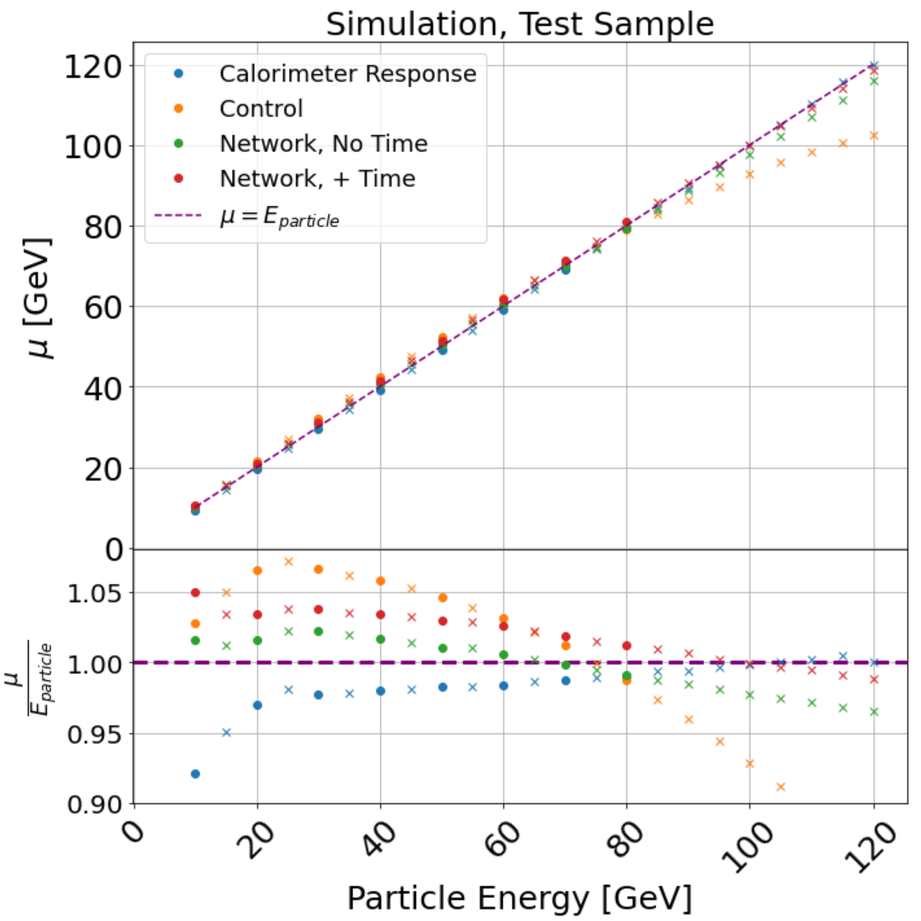
- mean calorimeter response after compensation = true particle energy

What is shown:

- Extracted mean (μ) of a Gaussian fit in a ±2 standard deviation range from the mean of the sample at each momentum;
- Colour coding as in previous slide.

What we learn:

- Network methods result in superior linearity of response than control method by around 2% on average;
- All SC methods over-correct the reconstructed energy;
- The control method is biased → above 80 GeV, the mean energy is not correctly reconstructed.
- The neural network methods reconstruct mean energy within 2% beyond the upper edge of the training range at 80 GeV.



Results: Calorimeter Resolution

What we expect:

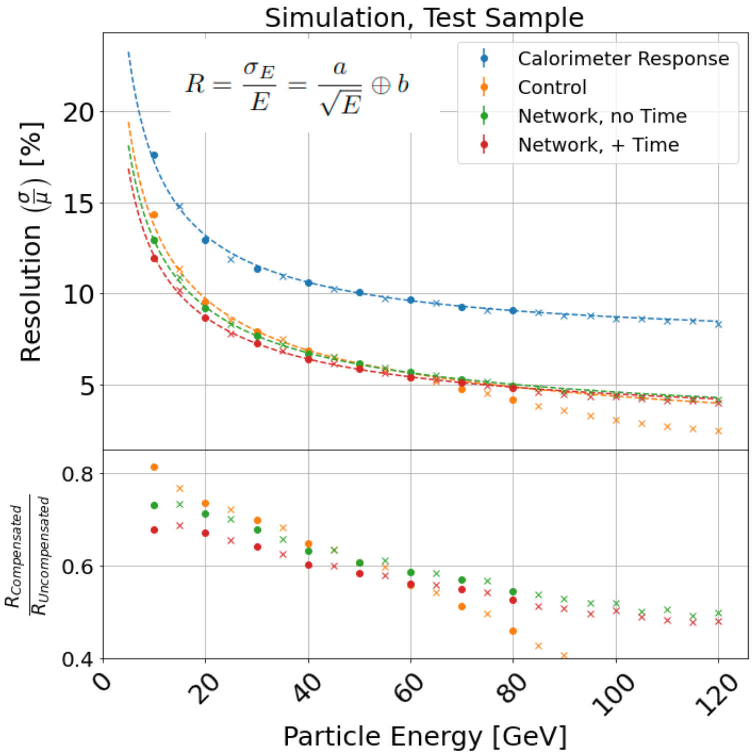
- The energy resolution of the calorimeter should improve by applying SC;
- Resolution function has two terms:
  - 'stochastic term', a → contribution of invisible energy fluctuations;
  - 'constant term', b → calibration quality of AHCAL detector;
  - both terms should reduce because of compensation.

What is shown:

- Extracted σ/μ of a Gaussian fit in a ±2 standard deviation range from the mean of the sample at each momentum studied;
- Fits of resolution equation shown in top-left of figure;
- Control fitted only up to 60 GeV → bias results in unphysical resolution.

What we learn:

- Neural networks show significant improvement in stochastic term, a:
  - - 3 % vs control using spatial information;
  - - 6 % vs control using spatial + 100 ps timing information;
- ~5% improvement in the constant term, b, suggests method also calibrates detector;
- Neural networks show excellent agreement with fit function by comparison to control.



	a [%]	b [%]	$\frac{\chi^2}{NDF}$
Calorimeter Response	49.516 ± 0.401	7.147 ± 0.067	4.575
Control	43.387 ± 0.119	0.010 ± 2.873	14.333
Network, No Time	40.236 ± 0.217	2.158 ± 0.087	0.857
Network, + Time	37.275 ± 0.208	2.448 ± 0.070	1.440

Bonus: Learning Physics From the Machine

What we expect:

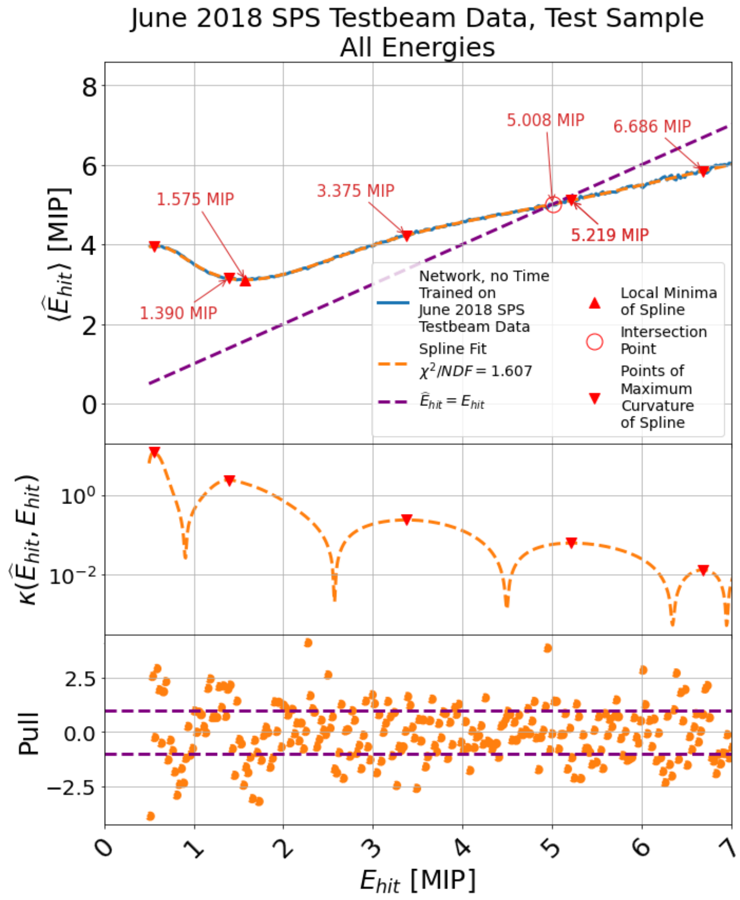
- The energy reconstruction should reflect properties of physics/detector;

What is shown:

- Average reconstructed active cell energy vs. original active cell energy, **trained on data;**
- Spline fit applied on orange dashed line.
- Markers:
  - Down arrow: 'elbow' points from curvature;
  - Up arrow: local minima of the curve;
  - Empty circle: intersection point with purple line.

What we learn:

- Attenuation occurs above 5 MIP, enhancement below:
  → 5 MIP is the AHCAL high gain/low gain switching mode;
- Curve highly nonlinear, with clearly different behaviour for different 'regions' of energy;
- Points of high curvature observed in strong agreement with the physics energy regions predicted in [4], from which the 'energy binning' SC idea is derived.



[4] V. Morgunov and A. Raspereza. Novel 3D Clustering Algorithm and Two Particle Separation with Tile HCAL. Dec. 17, 2004. doi: 10.48550/arXiv.physics/0412108.

Conclusion

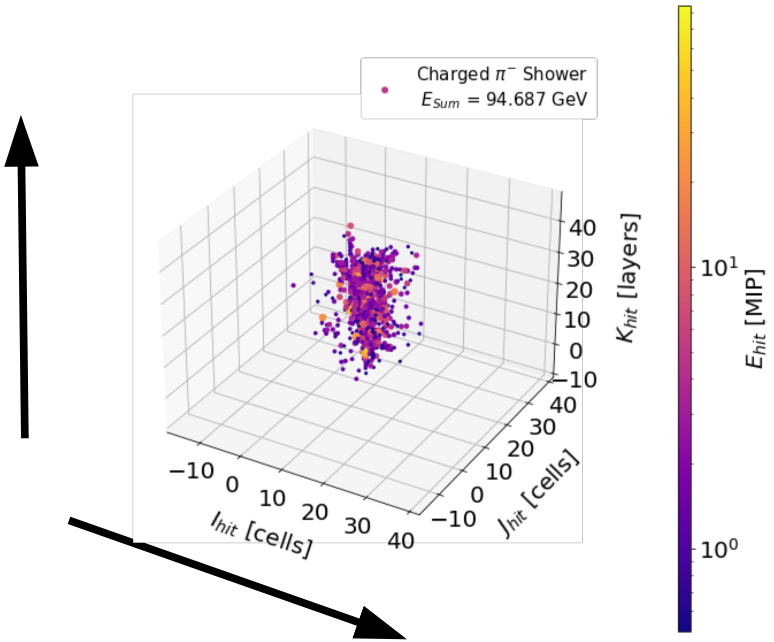
- Software compensation (SC) can improve jet energy resolution in Particle Flow Calorimetery;
- Data-driven SC models have been observed to bias to the particle momenta they are trained on;
- An neural network SC model was devised to both exploit the spatial and temporal energy density of the highly-granular AHCAL detector and to overcome the limitations of biasing;
- The network method outperformed the control:
  - superior linearity of response by around 2% with and without timing information;
  - superior stochastic resolution improved vs control method by:
    - 3% using spatial energy-density event information;
    - 6% using spatial + 100 ps resolution timing information
  - Network also improved detector calibration → possibility to apply method as a generic detector calibration tool?
- MAIN RESULT:
  - **The network method was found not to bias to the training range of energies →**

    method can be used to perform compensation with limited simulation/data in an experimental setting.

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

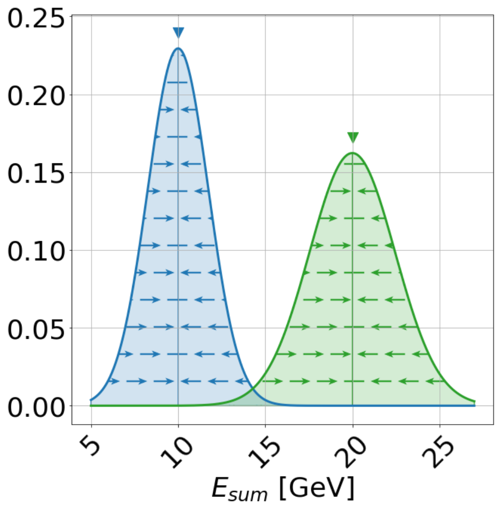
Color Axis:

- 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.

Backup: Biasing of SC Models



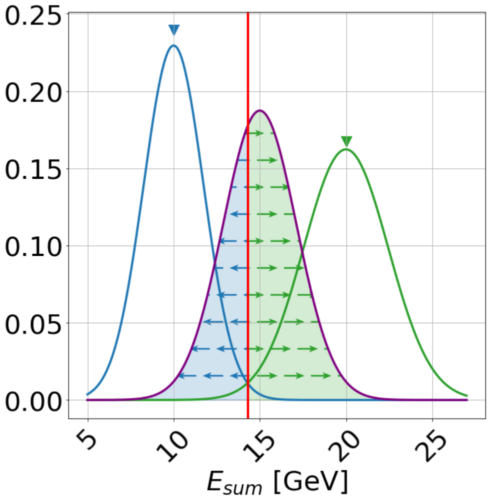
- Width of theoretical calorimeter response distribution proportional to Poissonian missing energy fluctuations (under certain assumptions);

- Ansatz made for training SC algorithms:

improvement in resolution → compensation

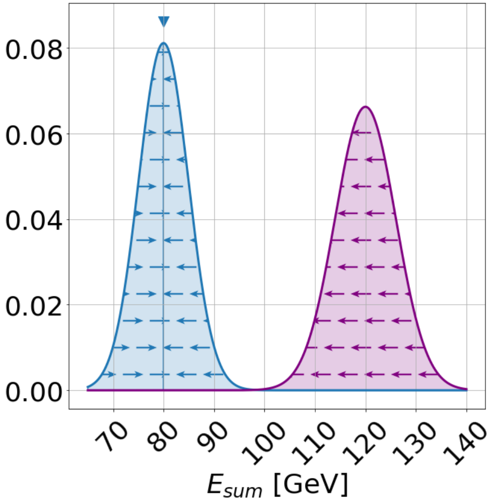
→ minimise mean χ^2 goodness-of-fit between calorimeter response and known particle momentum

Interpolation Failure



- Model learns to 'classify' hadron shower by mean response;

Extrapolation Failure



- Model learns the upper/lower edges of the training energy range;

TAKE-HOME MESSAGE:
SC algorithms tend to bias if the model is exposed to the mean responses of the training hadron shower events.

Control Method

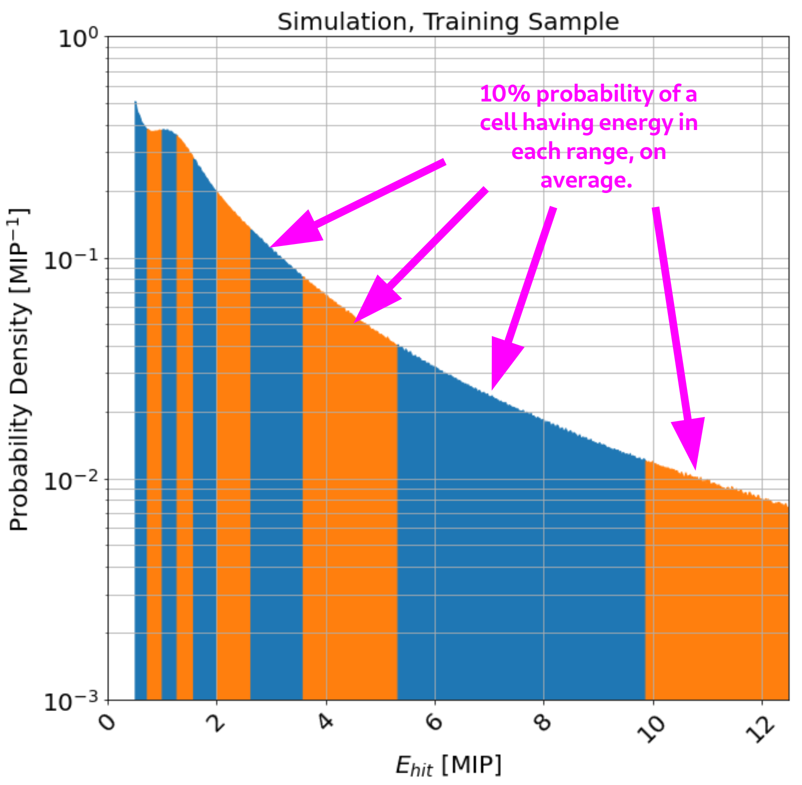
- Standard CALICE SC method used as a control [3];
- Control method is also designed to estimate the energy density of the hadron shower, event-by-event, but relies on calorimeter response to do so:
  - Individual cell energy distribution binned in deciles (10% probability an active cell energy will fall in any given bin, on average)
  - Compensation weight for each bin is calculated using a Chebyshev polynomial function approximator as a function of calorimeter response:

$$\omega_b(E_{\text{sum}}; S, \alpha_b, \beta_b, \gamma_b) = \alpha_b + \beta_b \cdot \frac{E_{\text{sum}}}{S} + \gamma_b \cdot \left(2\left(\frac{E_{\text{sum}}}{S}\right)^2 - 1\right)$$

$$S = 150\text{GeV} \qquad b = \text{bin} \qquad \alpha_b, \beta_b, \gamma_b = \text{weights}$$

  - Energy falling in each bin scaled by weight;

$$\widehat{E}_{\text{sum}} = \sum_b^{\text{bins}} \omega_b \cdot E_{\text{sum},b}$$



TAKE-HOME MESSAGE: algorithm weights fraction of energy falling into each bin as a function of the calorimeter response to hadrons.

[3] CALICE Collaboration, "A new approach to software compensation for the CALICE AHCAL," tech. rep., 2010.

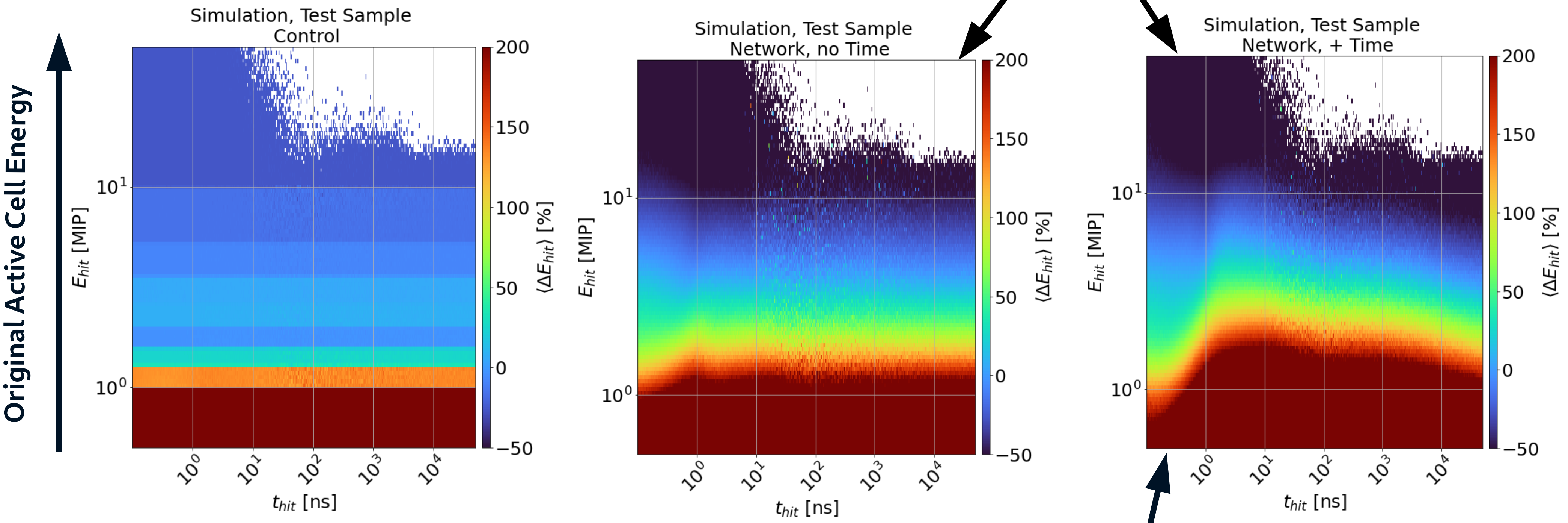
Backup: Mean % Change in Cell Energy, Energy/Temporal Dependence

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

HIGH-D

Blue: active cell energy attenuated
Red: active cell energy enhanced

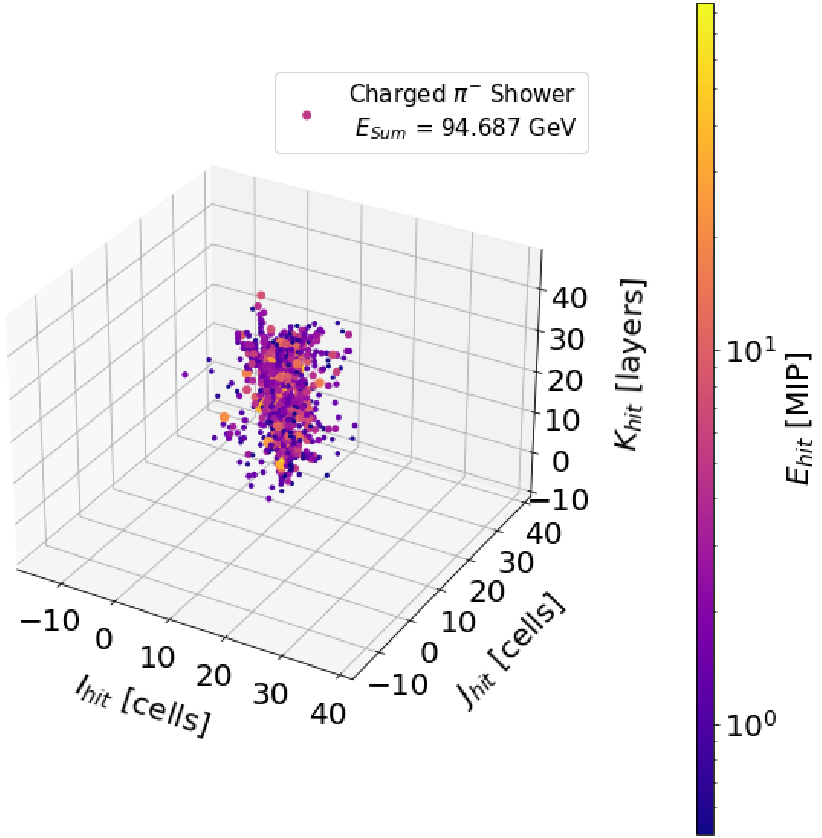
continuous weighting of active cell energy

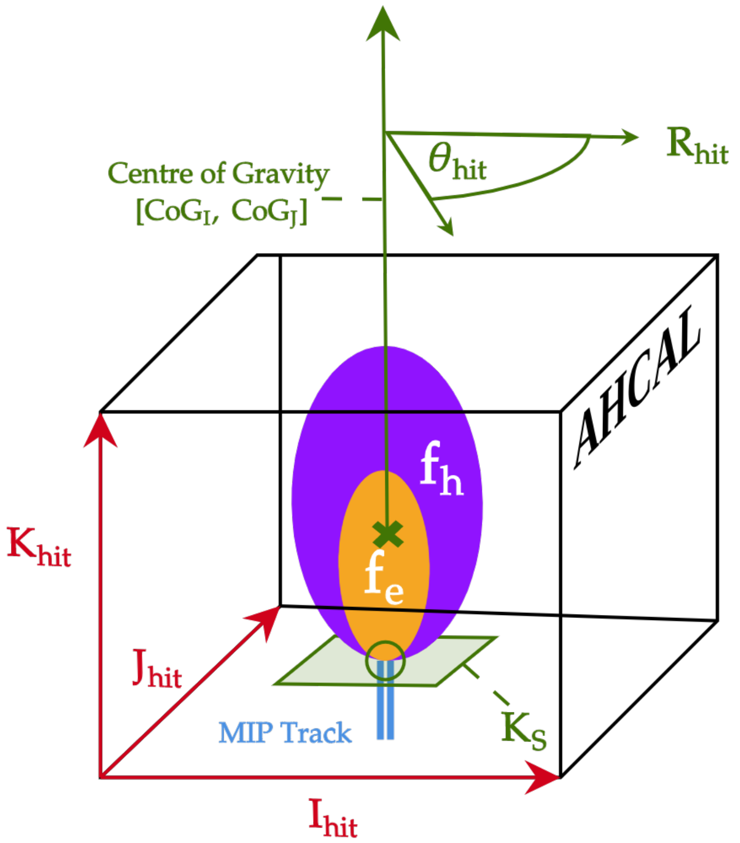


Temporal Development

threshold for weighting increased at >10 ns → sensitivity to neutron fraction

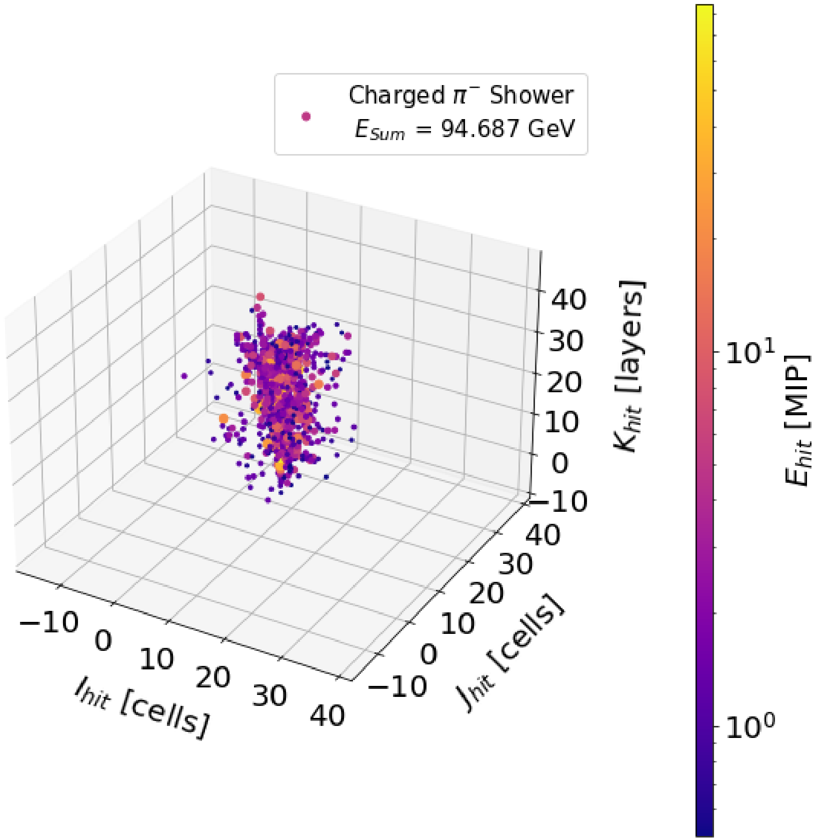
Backup: 'Natural' Shower Coordinate System



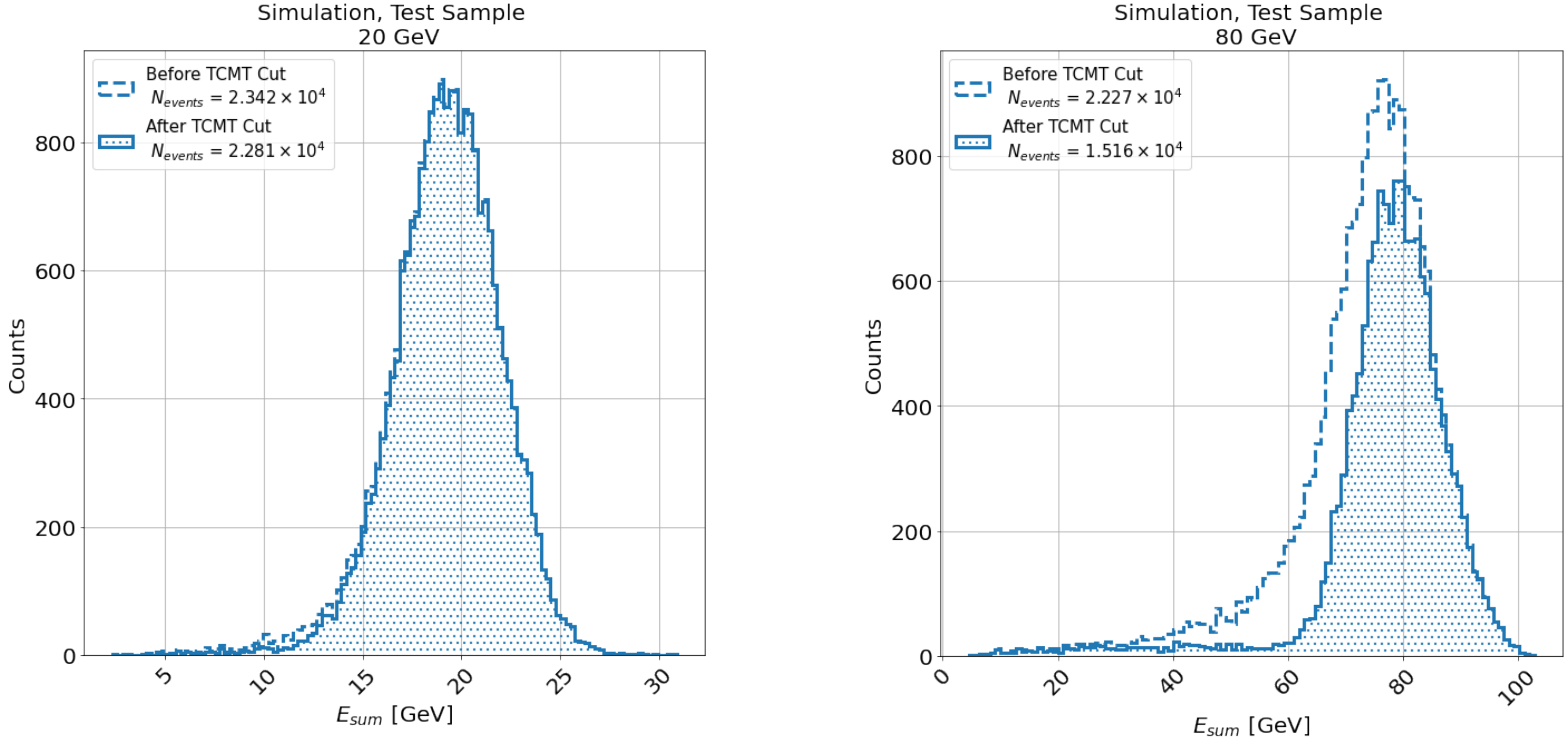


Backup: Simulation Details

- Physics list: QGSB_BERT_HP
- Particle: π^-
- Cuts:
  - + Shower Start: $1 < K_S \leq 5$
  - + Single track, with position 'inside' calorimeter: $1 < I_{\mathrm{Track}}/J_{\mathrm{Track}} \leq 24$
  - + PID MIP Cut: $P_{\mu\text{-like}} < 0.5\%$
  - (+ Tail Catcher Leakage Cut: $E^{TC}_{\mathrm{sum}} < 25\mathrm{MIP}$, (TCMT) for resolution/linearity measurement only)
- Events:
  - Training/Validation Sample:
    - 10-80 GeV, steps of 10 GeV;
    - Training: ~185,000 events (~20,000 events per step) (90% available)
    - Validation: ~ 20,000 events (~2,500 events per step) (10% available)
  - Testing Sample:
    - 10-120 GeV, steps of 5 GeV
    - Testing: ~ 497,000 events (~20,000 events per step) (all available)
    - Testing, + TC cut, ~384,000 events



Backup: TCMT Cut



What is shown:
Distributions of uncompensated calorimeter response, before and after tail-catcher cut, at 20 GeV and 80 GeV.

What we learn:
Effect of leakage reduced by application of tail-catcher cut at high particle momentum.