Investigation of Fast Hadron Shower Simulation Methods with the CALICE AHCAL Prototype

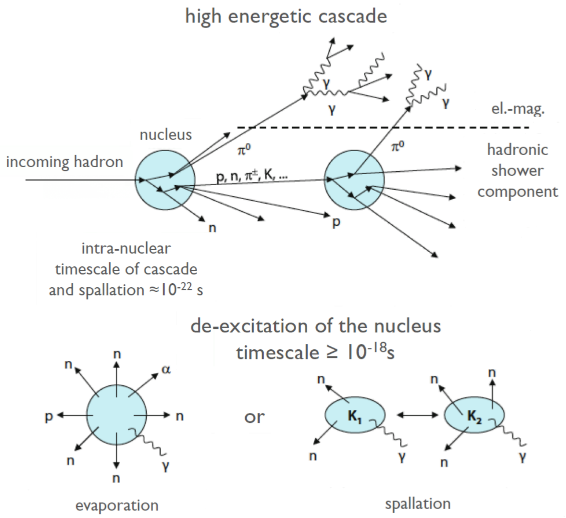
André Wilhahn
In consultation with Stan Lai and Julian Utehs

II. Physikalisches Institut, Georg-August-Universität Göttingen
CALICE Collaboration Meeting

March 29 - 31, 2023

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

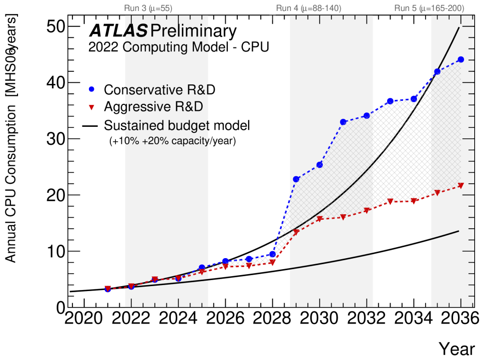
Motivation

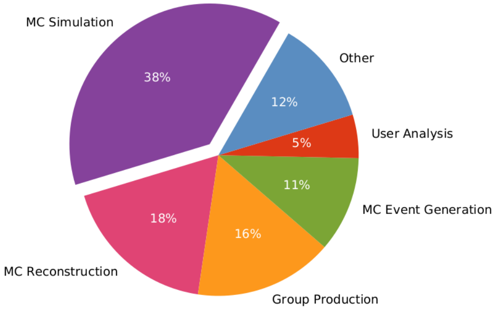


- Hadronic showers are very chaotic
- Energy resolution can vary strongly between events
  $\Rightarrow$ Limited by electromagnetic fraction
- Study of single showers helps to understand behaviour of hadronic showers in highly granular calorimeters
- Can also develop fast simulation by studying single showers

Motivation

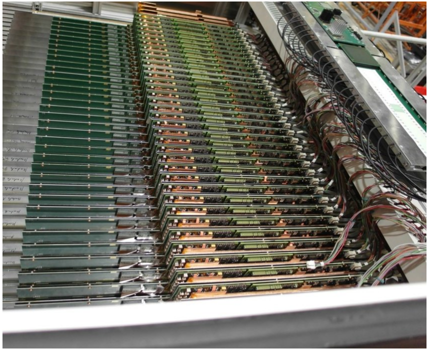
- CPU consumption of MC simulations increases with occupancy/granularity
- Up to 90 % of calculation time is needed for the calorimeter
- Saving of computational resources will become necessary sooner or later

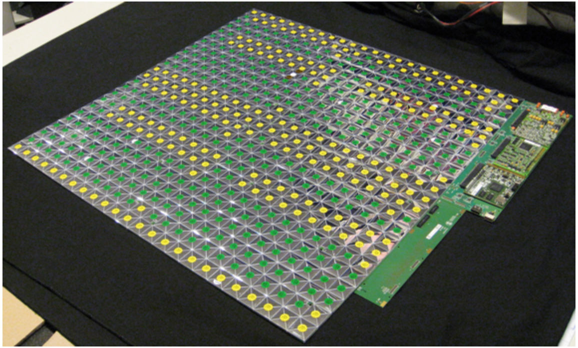




AHCAL Technological Prototype

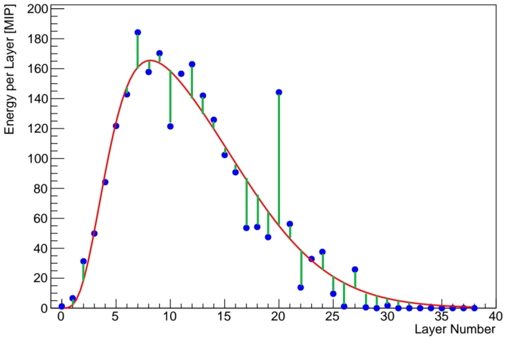
- 38 active layers of 24×24 scintillator tiles ($3 \times 3\,\mathrm{cm}^2$) embedded in stainless-steel absorber structure, each individually read out via silicon photomultipliers
- Build **data-based fast simulation** based on pion shower dataset recorded at CERN $\Rightarrow$ 10 GeV, 20 GeV, 30 GeV, 40 GeV, 60 GeV, 80 GeV, 120 GeV, 160 GeV, 200 GeV

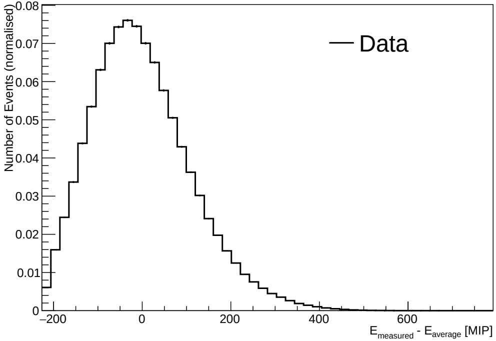




Differences between Single and Average Pion Showers

- Consider differences in energy deposition between single and average shower instead of absolute energies
- Investigate energy deposition layerwise relative to shower start layer





Kernel Density Estimators

- Want to find PDF of dataset x_1, x_2, ..., x_n
- Define Kernel Density Estimator (KDE) with bandwidth h as:

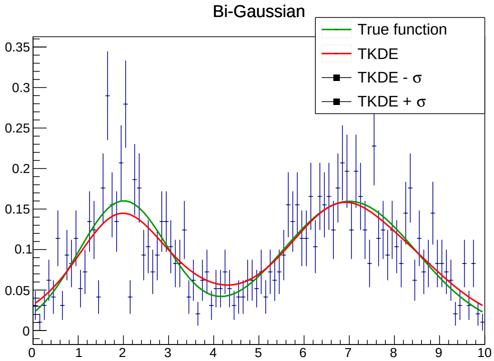
$$f(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

with

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

- PDF = sum of all Gaussian kernels
- Choice of bandwidth determines smoothness of PDF
- Apply KDE of energy differences simultaneously on layer groups

Kernel Density Estimators



- Generalise to d dimensions:

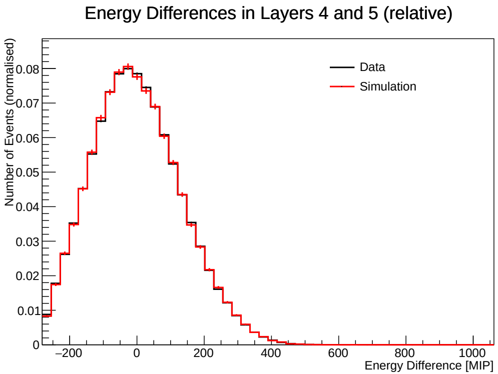
$$f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} |\mathbf{H}|^{-1/2} K\left(\mathbf{H}^{-1/2}(\mathbf{x}-\mathbf{x}_i)\right)$$

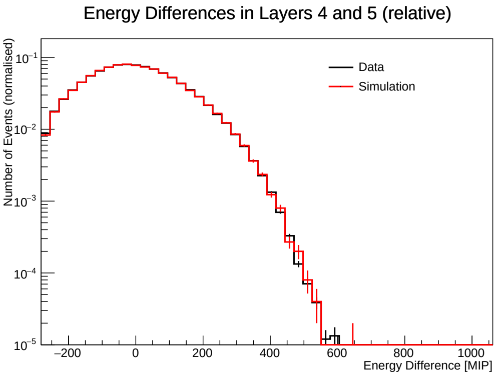
- $\mathbf{x}$: d-dimensional data vector; $\mathbf{H}$: $d \times d$ bandwidth matrix
- Have chosen $\mathbf{H} = h^2\mathbf{C}$ where $\mathbf{C}$ is the covariance matrix of the dataset

Simulated Energy Differences for 60 GeV Pions

- Generate 100 000 events with KDEs

  ⇒ Each event containing eight energy differences corresponding to eight layer groups

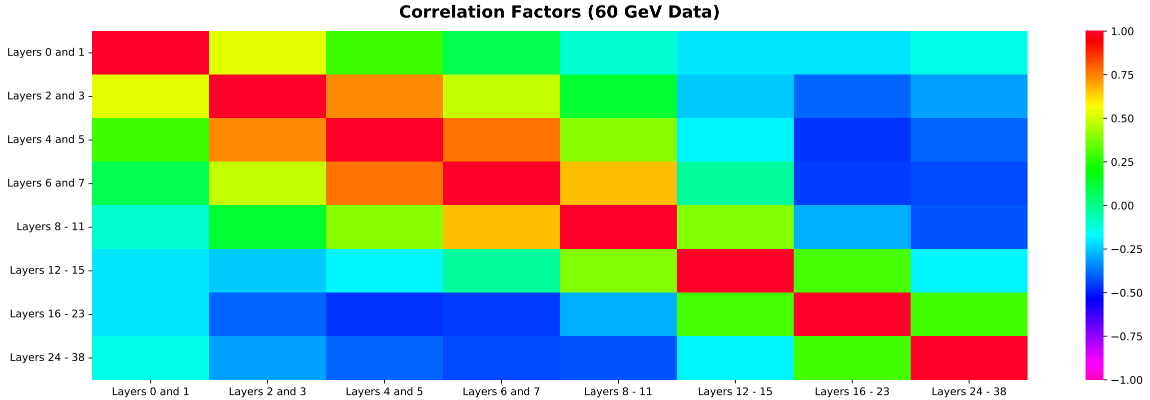




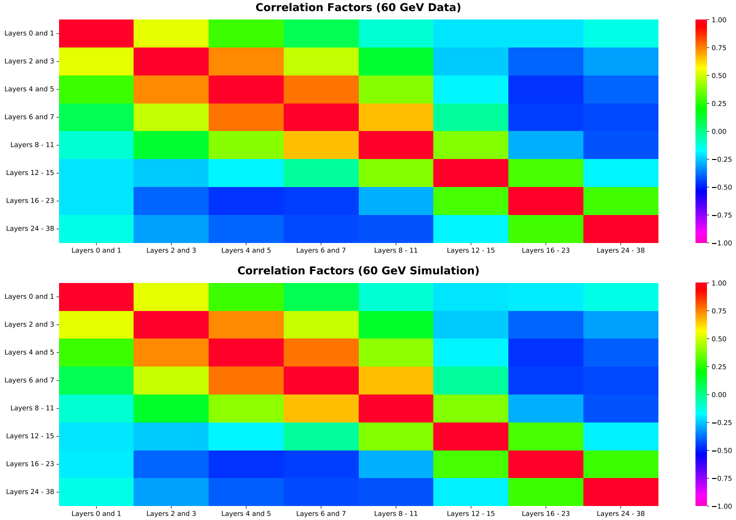
Excellent agreement between data and simulation!

Correlation Factors between Calorimeter Layers

- Simulation must preserve (anti-)correlations between detector layers
- Strong correlations for neighbouring layers; strong anticorrelations for layers far apart



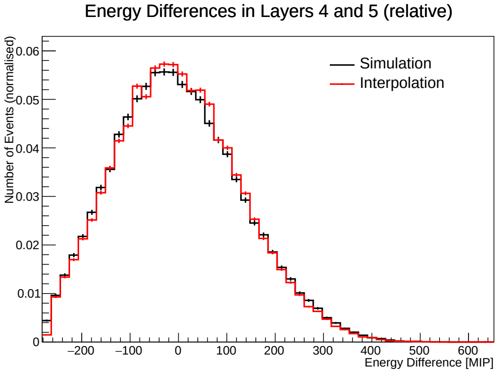
Comparison of Correlation Factors for 60 GeV Pions

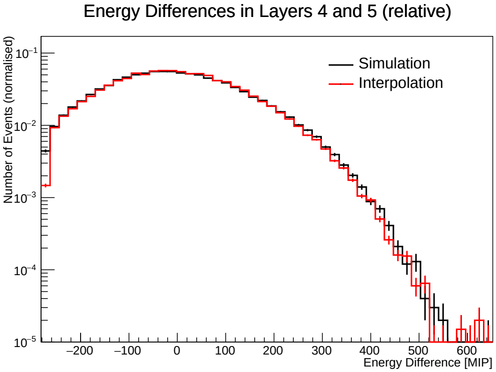


- Create correlation factor heatmaps for data and simulation
- Comparison shows negligible differences between simulation and expectation
- Proof that KDEs are well suited as pion shower simulation method

Interpolated Energy Differences for 60 GeV Pions

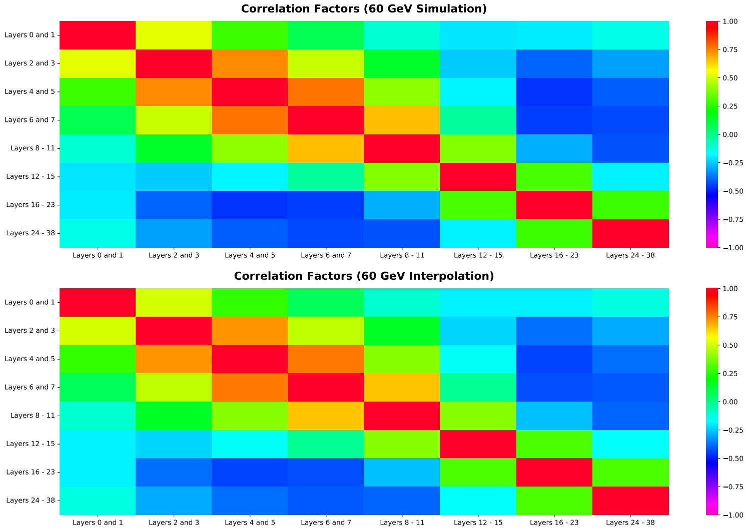
- Algorithm developed to interpolate energy difference PDFs between initial beam energies (40 GeV → 60 GeV ← 80 GeV)
- Compare with directly simulated events





Excellent agreement between simulation and interpolation as well!

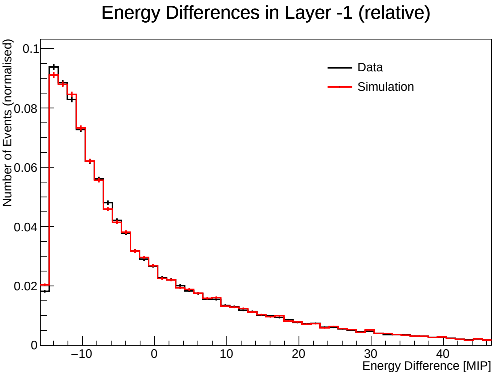
Interpolated Correlation Factors for 60 GeV Pions

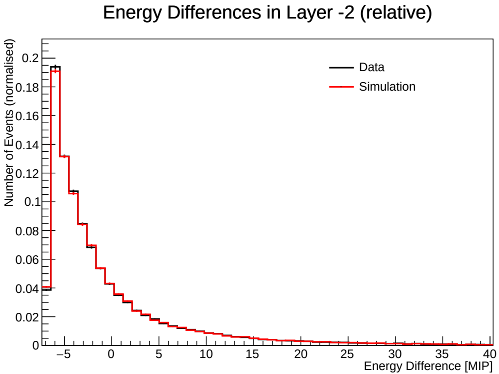


- Very good agreement between correlation factors of directly simulated and interpolated energy differences
- Interpolation works for both equi- and nonequidistant initial energies

Pre-Shower-Start Energy Differences for 60 GeV Pions

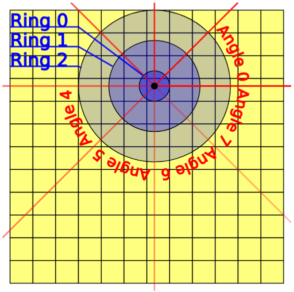
- Shower start finding algorithm not 100 % precise
  $\Rightarrow$ Uncertainty in shower start layer; might not include all hits in fast simulation
- Investigate and simulate distributions of energy differences **before** shower start layer
- Investigate relative layers -1 and -2 individually; combine all layers ≤ -3 into one

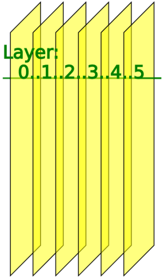


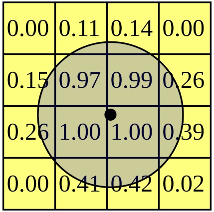


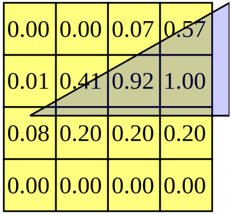
Intersection Area Calculation

- Intersection area calculation is needed for radial and angular simulation
- Examples of the percentages of the intersection areas for one radial and angular case



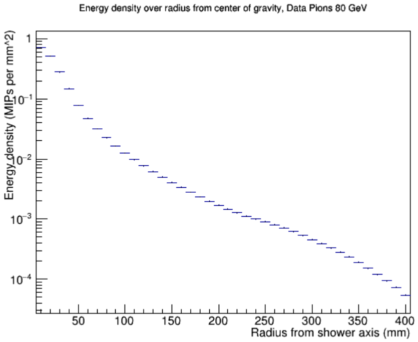


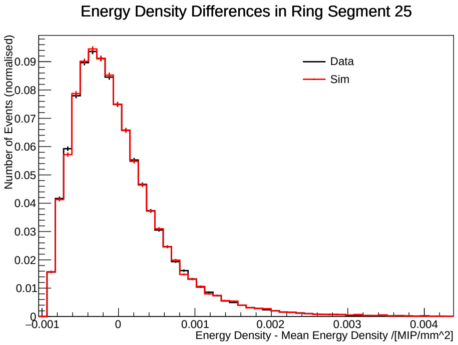




Radial Simulation - Distributions (80 GeV Pion Data)

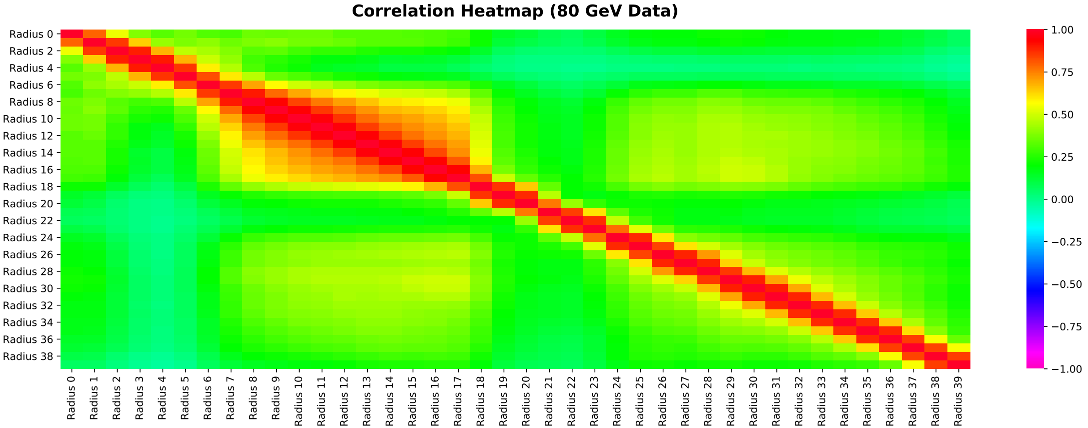
- Radial energy deposition decreases steeply from shower centre
- Investigate energy density differences from mean value per ring from shower centre



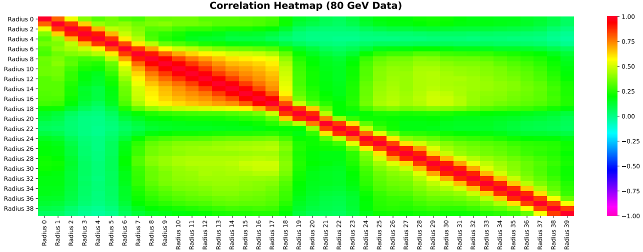


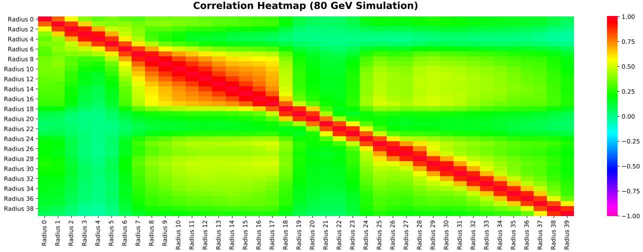
Radial Simulation - Correlation Factors (80 GeV Pion Data)

- Correlation factor heatmaps show a strong correlation between neighbouring rings
- Heatmaps show "blobs" of correlations → reason why yet unknown

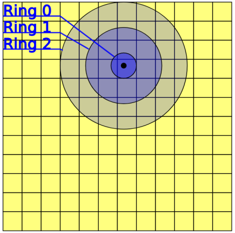


Radial Simulation - Correlation Factors (80 GeV Pion Data)



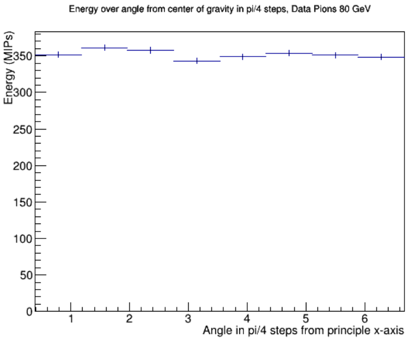


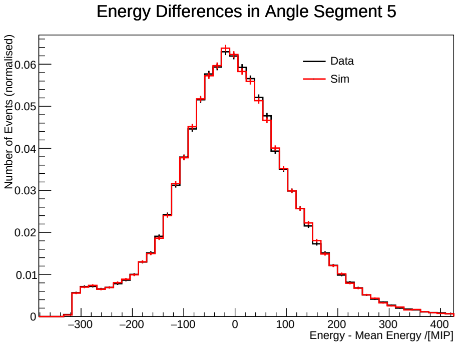
- Correlation factor heatmaps show good agreement between data and simulation
- KDEs preserve correlations and anticorrelations between rings



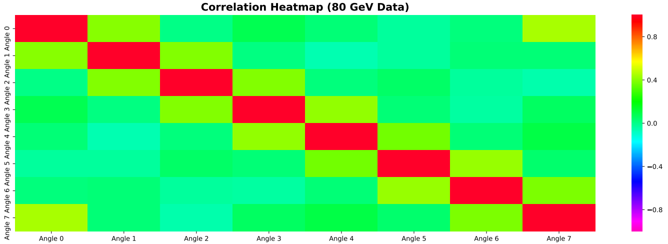
Angular Simulation - Distributions (80 GeV Pion Data)

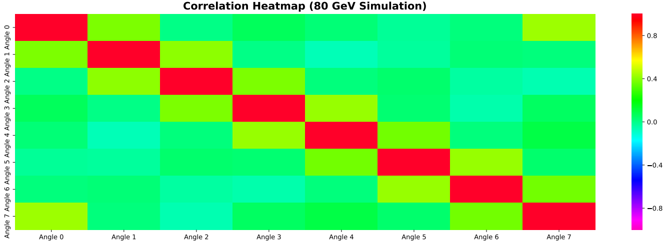
- Angular energy deposition is almost constant in every direction
- Distribution of energy differences within angle segment is nearly Gaussian



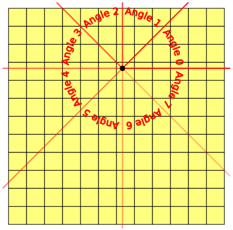


Angular Simulation - Correlation Plots (80 GeV Pion Data)

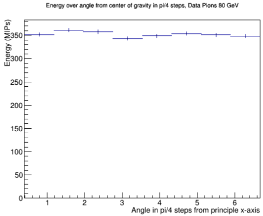


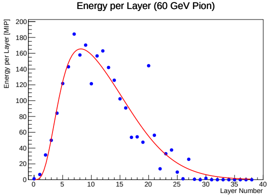


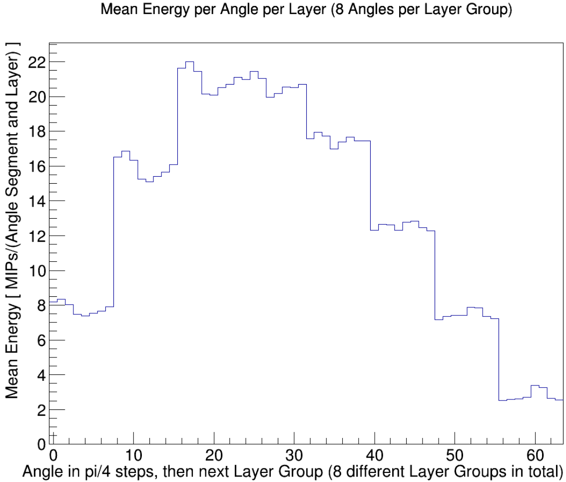
- Strong correlations between neighbouring angle segments; anticorrelations between opposing segments
- Simulation and data are in good agreement



Combining Angular and Longitudinal Simulation

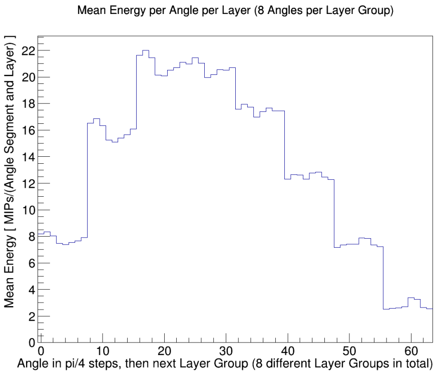


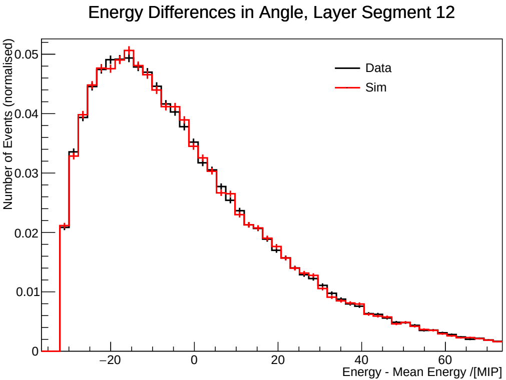




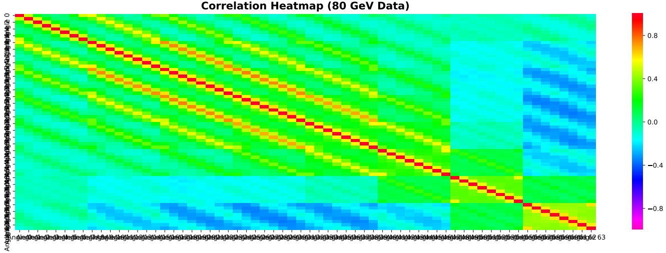
Energy Differences per Angle per Layer (80 GeV Pion Data)

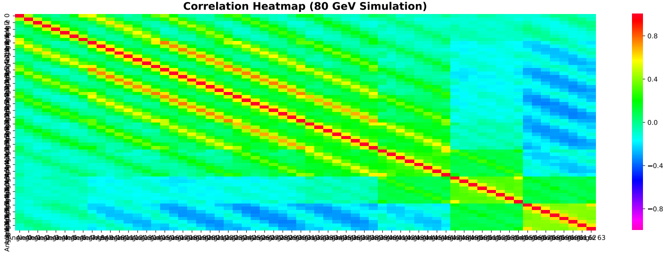
- Longitudinal layer configuration: 0-1, 2-3, 4-5, 6-7, 8-11, 12-15, 16-23, 24-38
- Simulation can reproduce the expected curves





Correlations of Angular-Longitudinal Segments (80 GeV Pion Data)





- Correlations of combination of angular and longitudinal energy differences are preserved
- Demonstrates that KDEs are well suited for the calculation of PDFs, even if several input parameters are used

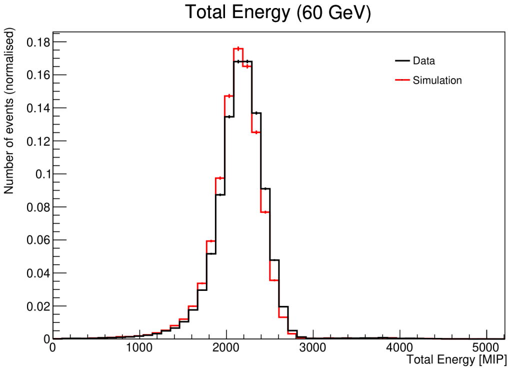
Conclusion

Summary:

- Simulated energy difference distributions and correlations match expectations very well
- Interpolation algorithm gives expected results

Outlook:

- Implement algorithm for energy extrapolation
- Merge longitudinal and radial simulations
- Provide events in same data format as test beam data



Thank you for your attention!

Event Selection

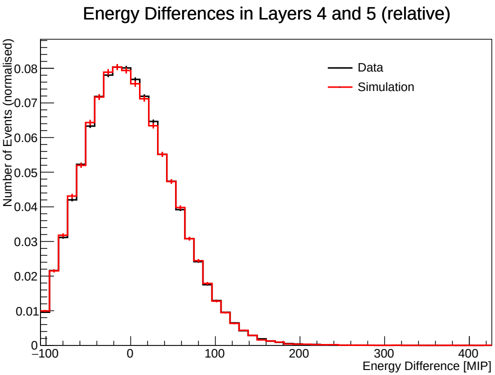
Selection:

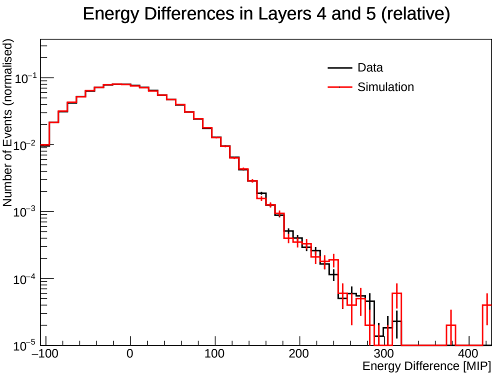
- Applied particle identification using BDT-techniques to remove beam contamination
- Exclude first physical AHCAL layer in order to minimise uncertainty in shower start finding algorithm

Event Selection:

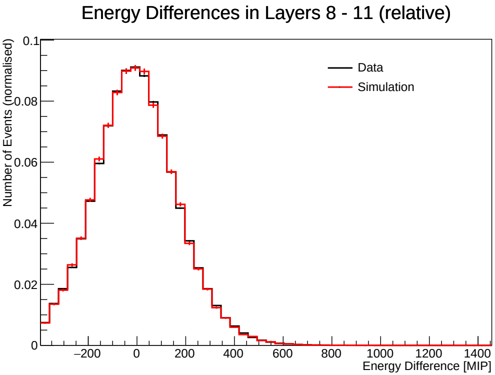
- Apply low-energy cut to remove muons from dataset (initial muons as well as those coming from $\pi^- \to \mu^- \bar{\nu}_\mu$)
- Exclude events starting in layer 11 or later to minimise leakage

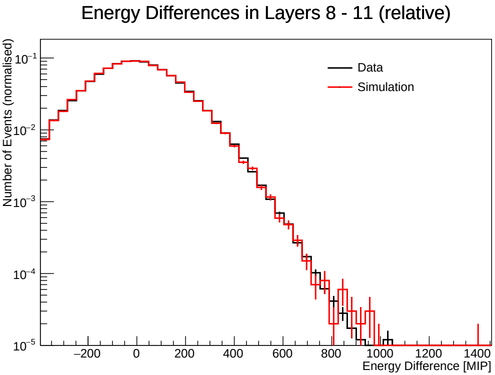
Simulated Energy Differences for 20 GeV Pions (Longitudinal)



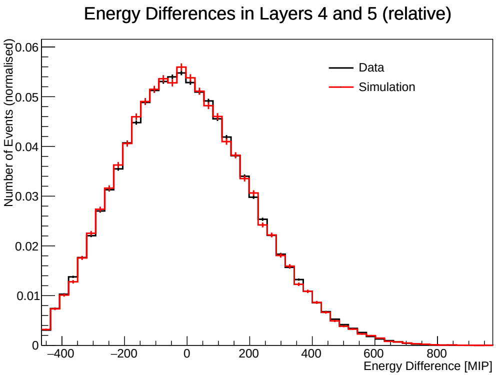


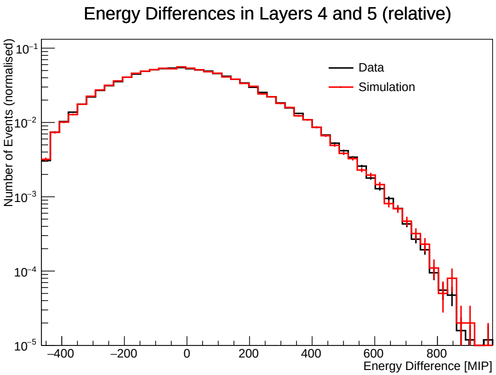
Simulated Energy Differences for 60 GeV Pions (Longitudinal)





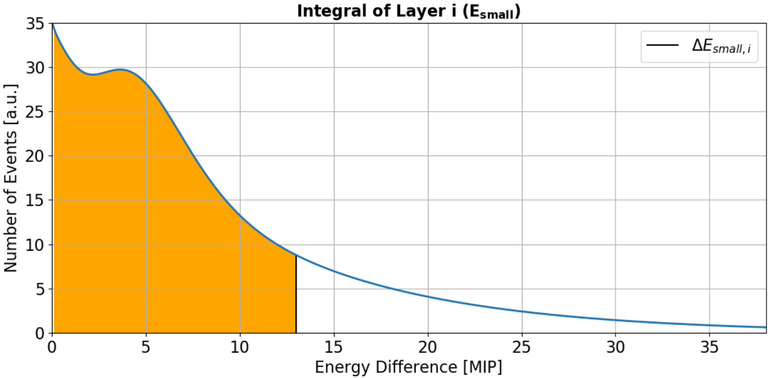
Simulated Energy Differences for 120 GeV Pions (Longitudinal)





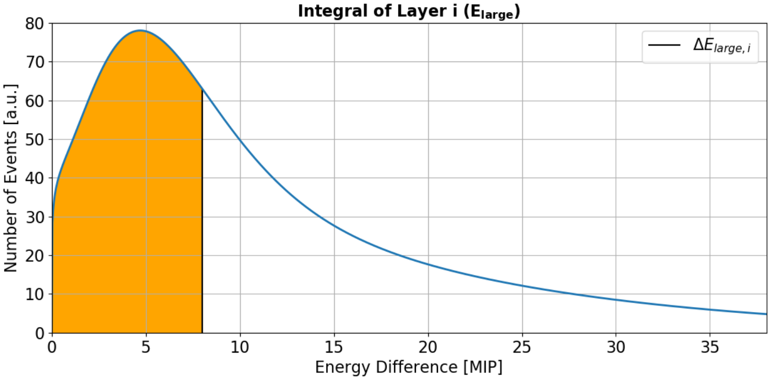
Mathematical Approach to Interpolations

- Interpolate from two energies (E_{small} and E_{large}) to target energy $E_{\text{interpolate}}$
- Generate one event randomly (one value for each layer group of E_{small})
- Integrate PDF of layer group i until $\Delta E_{\text{small},\, i}$ is reached (save area A_i)



Mathematical Approach to Interpolations

- Integrate PDF of same layer but for E_{large}, until area equals A_i (areas equal at $\Delta E_{\text{large},\, i}$)
- Interpolated energy difference: $\Delta E_{\text{interpolate},\, i} = \frac{\Delta E_{\text{small},\, i} + \Delta E_{\text{large},\, i}}{2}$
- Repeat vice versa ("small" and "large" interchanged)



Mathematical Approach to Interpolations

- Generalise interpolation to nonequidistant initial energies:

$$\Delta E_{\text{interpolate},\, i} = w_{\text{small}} \Delta E_{\text{small},\, i} + w_{\text{large}} \Delta E_{\text{large},\, i}$$

- Weights must fulfil

$$w_{\text{small}} + w_{\text{large}} = 1$$

- Weights depend on “distance” between initial energies:

$$w_i = 1 - \frac{|E_{\text{interpolate}} - E_i|}{E_{\text{large}} - E_{\text{small}}}$$

- Index i either “small” or “large”
- Larger weights for closer energies, and vice versa

Simulated Energy Differences (Radial)