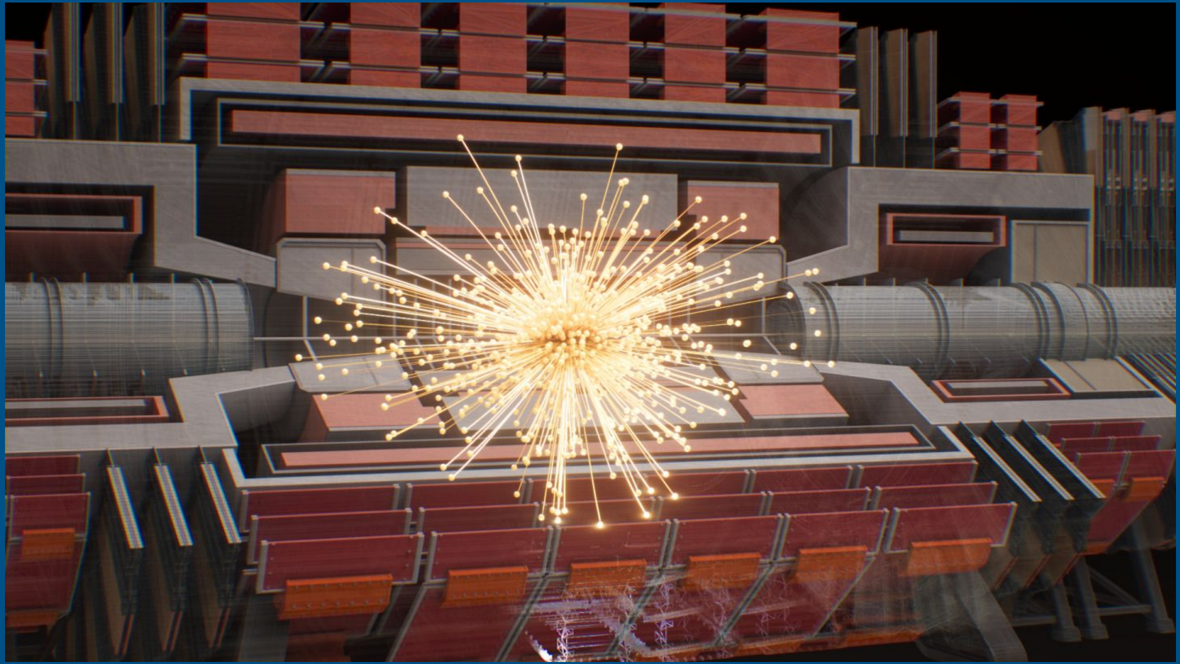
06.2025

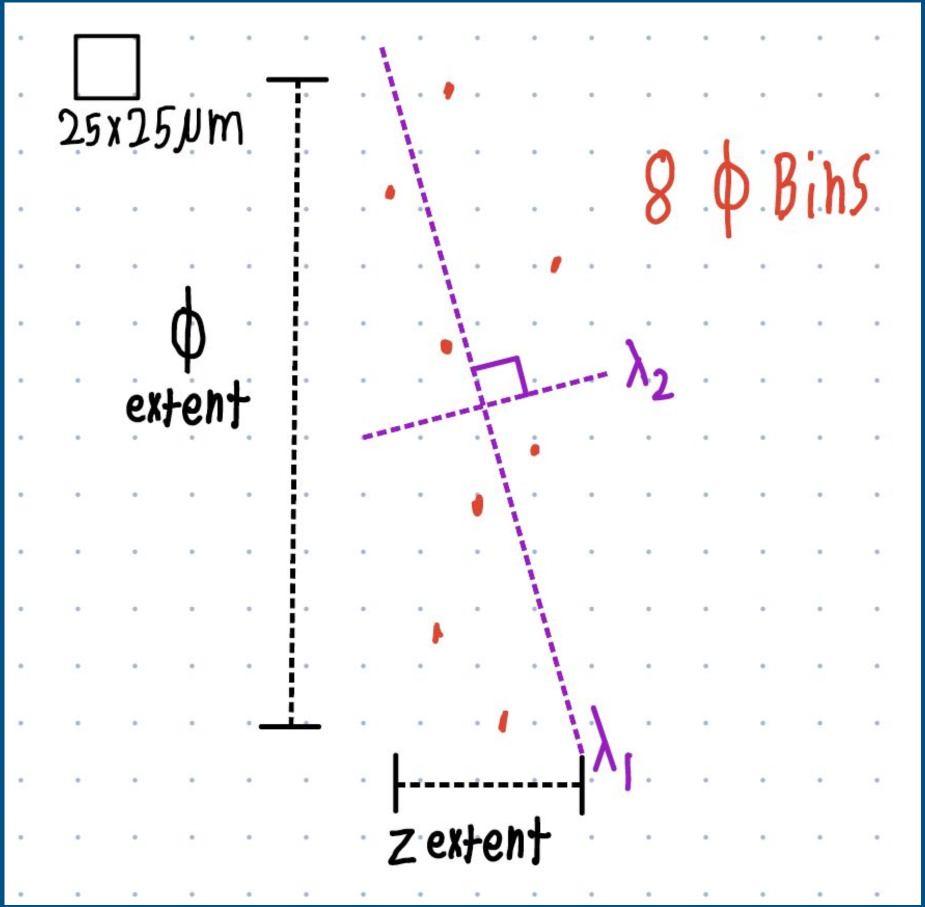
FCCee Geometric Clustering Characterization



Emmett Forrestel

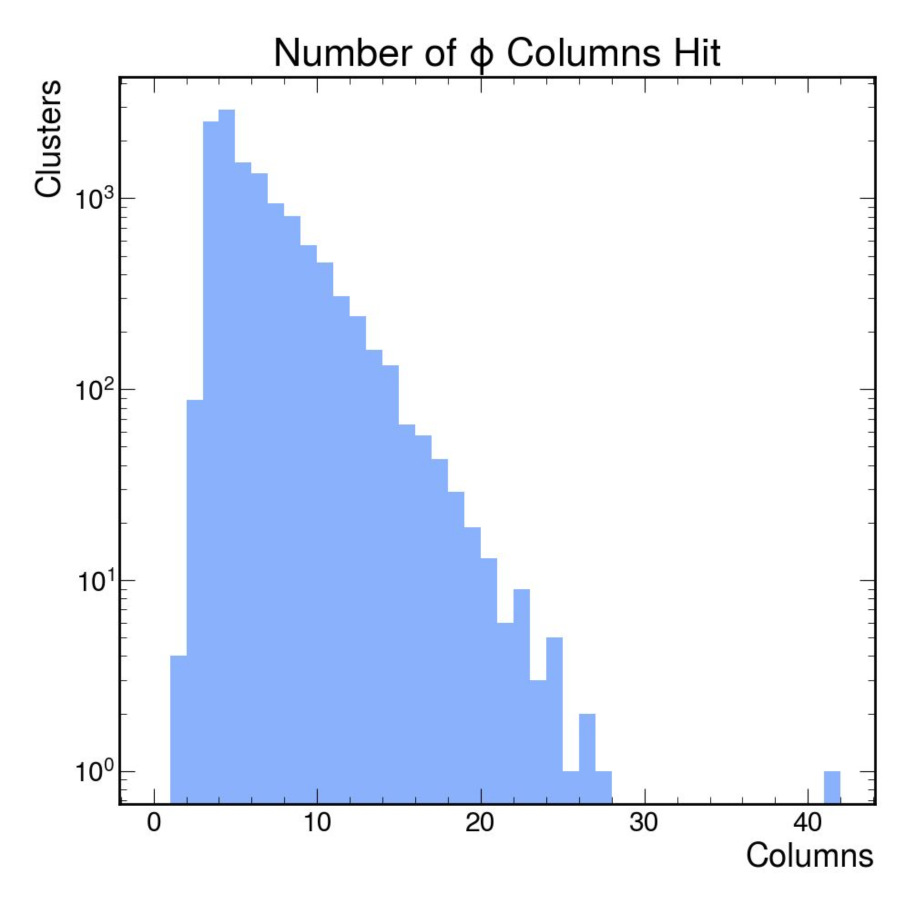
Introduction

- Analyzed hit clusters from a single MC particle in the first layer of the CLD vertex detector using simulation background data.
- Computed four geometric descriptors per cluster to study track shapes and detector response:
  a. Azimuthal spread ($\Delta\varphi$)
  b. Number of φ columns hit
  c. Z extent
  d. PCA elongation (λ_1/λ_2)



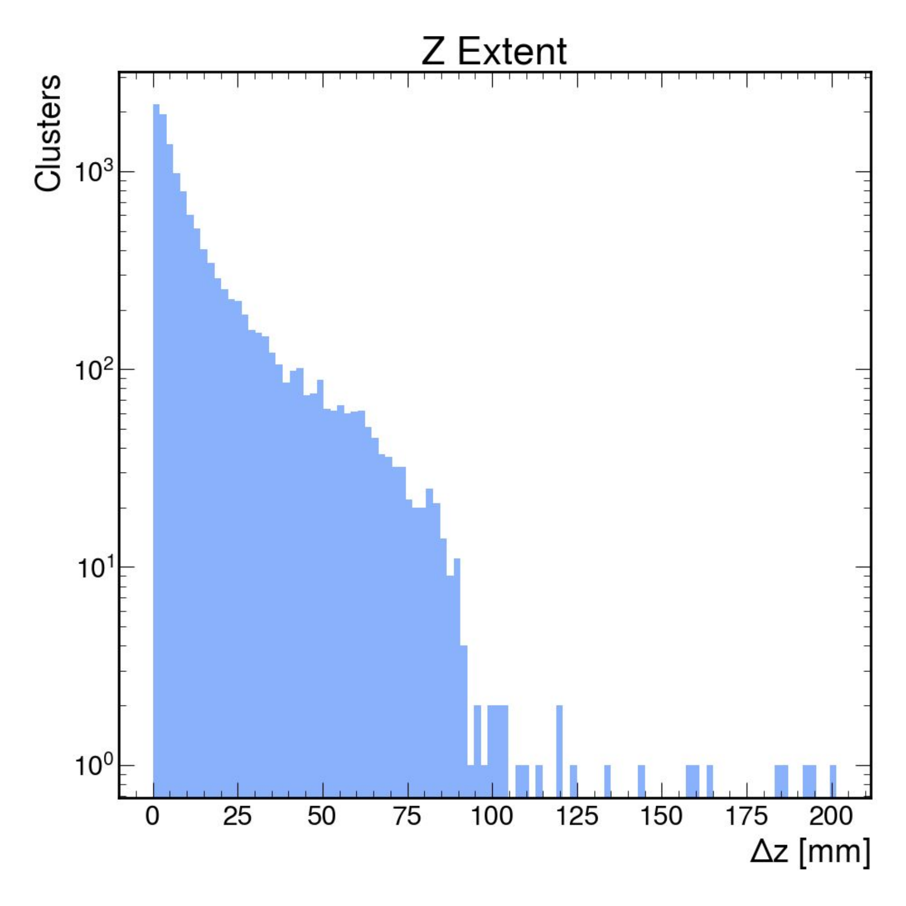
Φ Columns Hit

- Converts φ spread into discrete sensor bins, showing how many 25 µm-wide columns the cluster spans (from ARCADIA-MD3 chip).
- Gives a detector-relevant view of how many readout elements are activated by a single cluster.
- Helps connect simulation to real sensor granularity, affecting resolution, occupancy, and data rates.



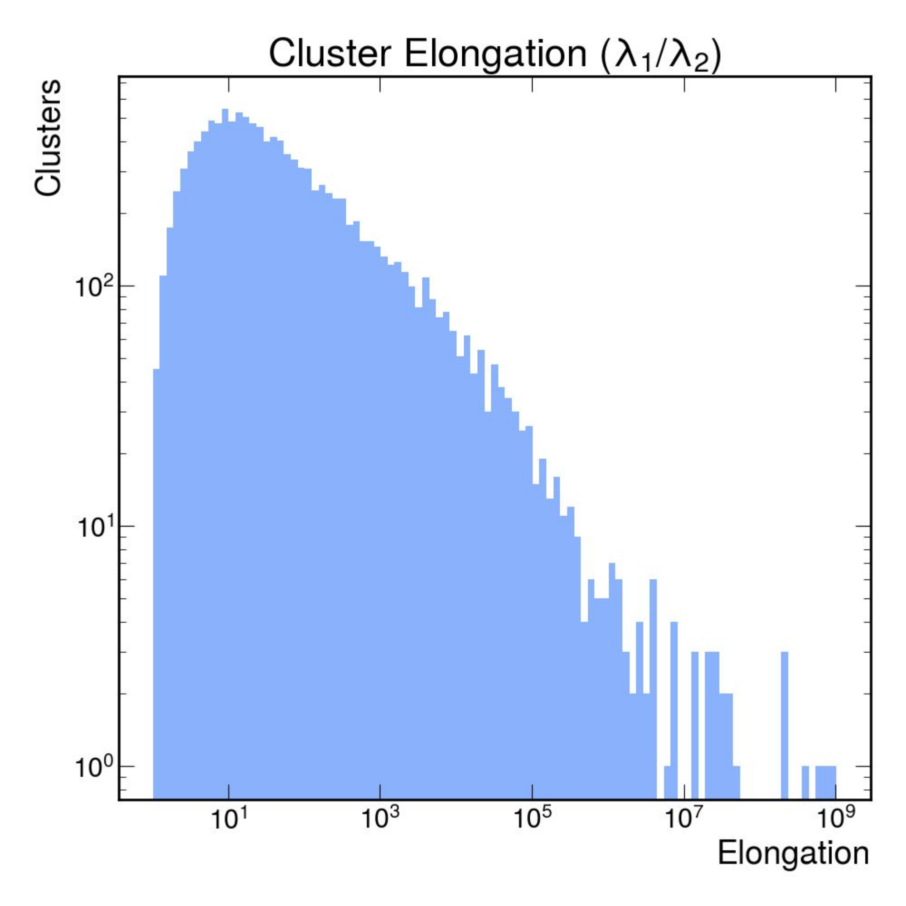
Z Extent

- Represents how long the cluster is along the beam axis, from minimum to maximum z hit.
- Sensitive to incident angle in the r–z plane, and also to physical effects like multiple scattering.
- Longer z extents may correspond to grazing angles, or distorted clusters.



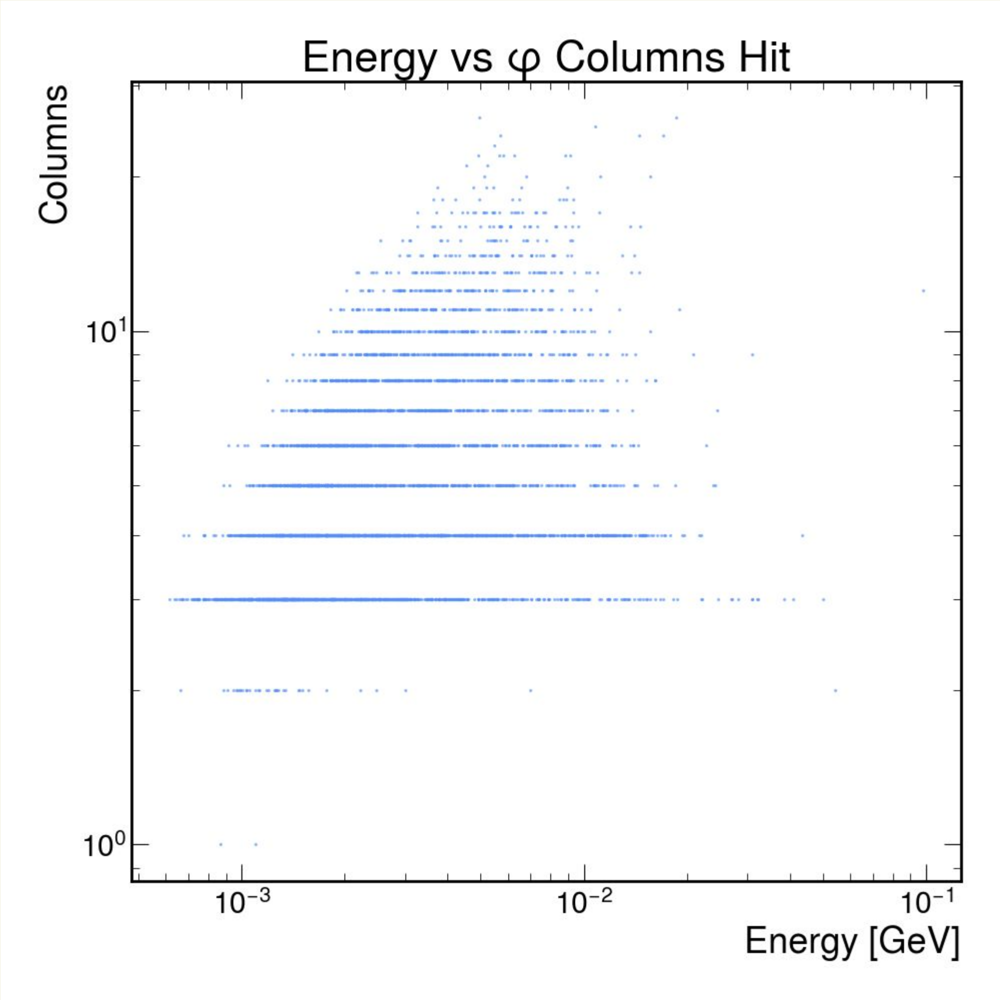
PCA Elongation

- Uses PCA to assess cluster shape in φ–z space: the ratio of primary to secondary variance, where λ_1 and λ_2 are eigenvalues of covariance matrix.
- High ratios indicate stretched clusters, while ratios near 1 indicate blob-like shapes.
- Offers a high-level summary of cluster geometry, useful for pattern recognition and potential ML-based classification.



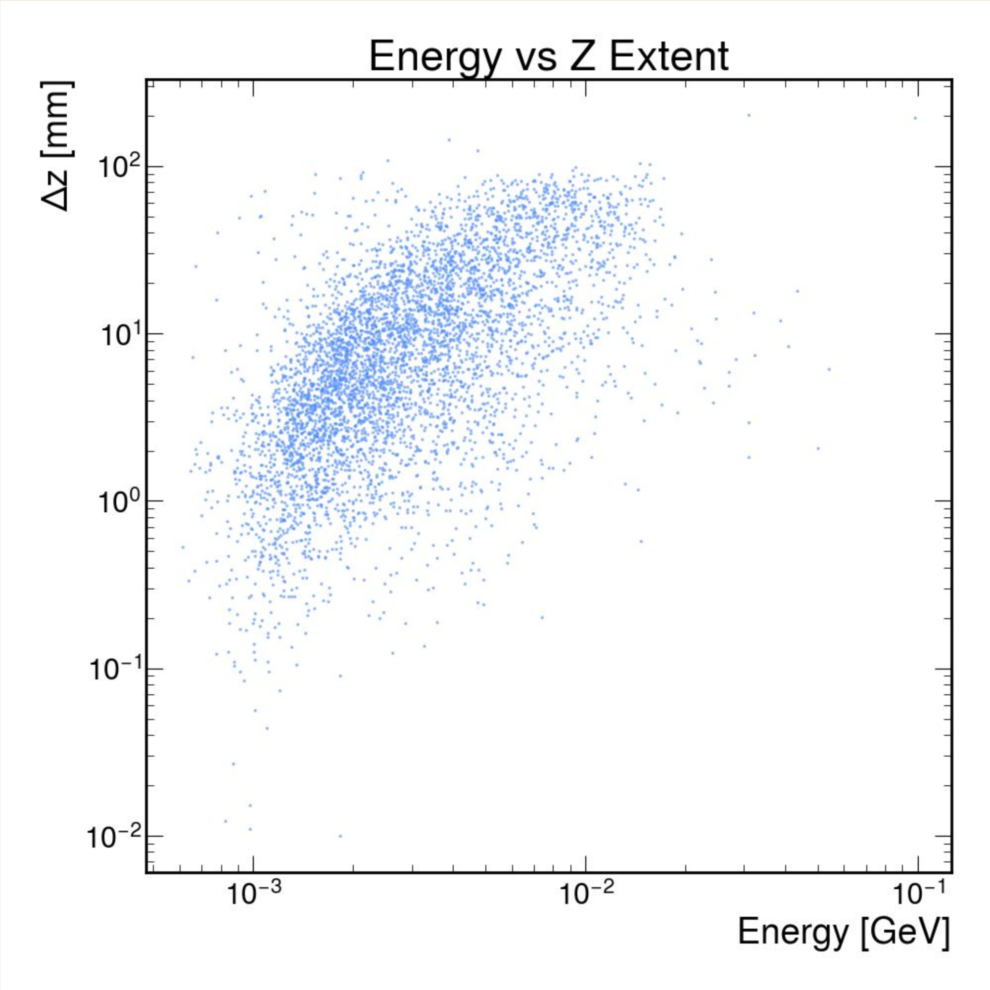
Energy vs Φ Columns

- Same discretized φ spread into 25 μm bins, according to the ARCADIA-MD3 chip architecture.
- All hits are low energy background, resulting in little correlation between the two.



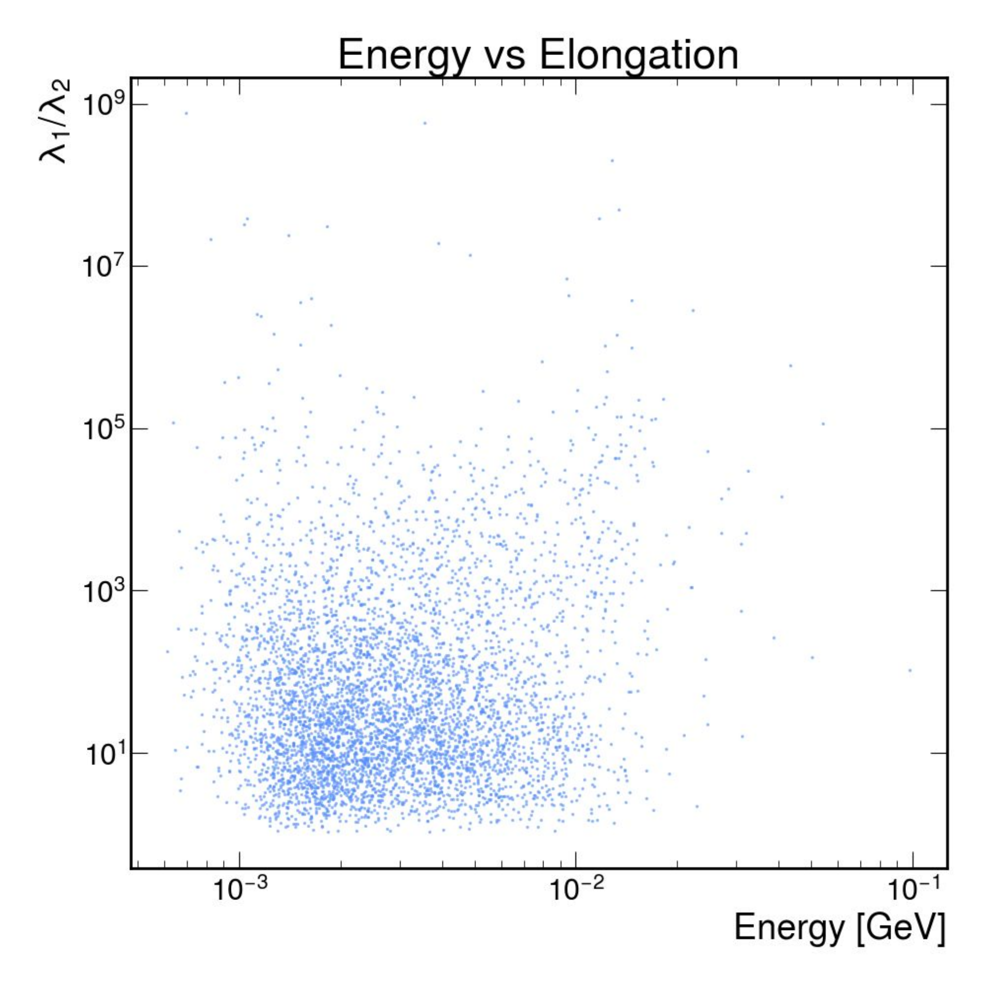
Energy vs. Z

- Plots Z_extent of background particles against energy.
- Again not much correlation is present, as all data is low energy background.
- Comparing with signal may provide useful correlations.



Energy vs. PCA Elongation

- Compares energy to the ratio of primary and secondary variance, λ_1 and λ_2 of covariance matrix.
- When compared with a signal sample, could provide a high-level summary of cluster geometry, useful for pattern recognition and potential ML-based classification.



Next Steps

- Apply the same analysis to signal data sets (ran on 500 background samples from FCCee_Z_4IP_04may23_FCCee_z).
- Further compare cluster shapes with particle properties to explore physical correlations, besides energy (e.g., charge, pseudorapidity).
- Use these metrics to inform or improve reconstruction algorithms, especially for low-energy or high-background regimes.