Investigating Far Forward Neutron Reconstruction in the ePIC Detector using Machine Learning

Vani Mahajan, IISER Mohali

Mentor: Dr. Stephen Kay, University of York
Supervisor: Prof. Sadhana Dash, IIT Bombay

Exclusive, Diffraction, & Tagging Meeting, 9 March 2026

Introduction

- ePIC is the main detector designed for the future Electron-Ion Collider to discover and understand the emergent phenomena of Quantum Chromo-Dynamics (QCD).
- Combines multiple sub-detectors for tracking, calorimetry, and particle ID.
- The B0 Calorimeter plays a crucial role in detecting forward going particles.
- Accurate neutron detection is vital for some reaction studies and momentum transfer (-t) reconstruction.

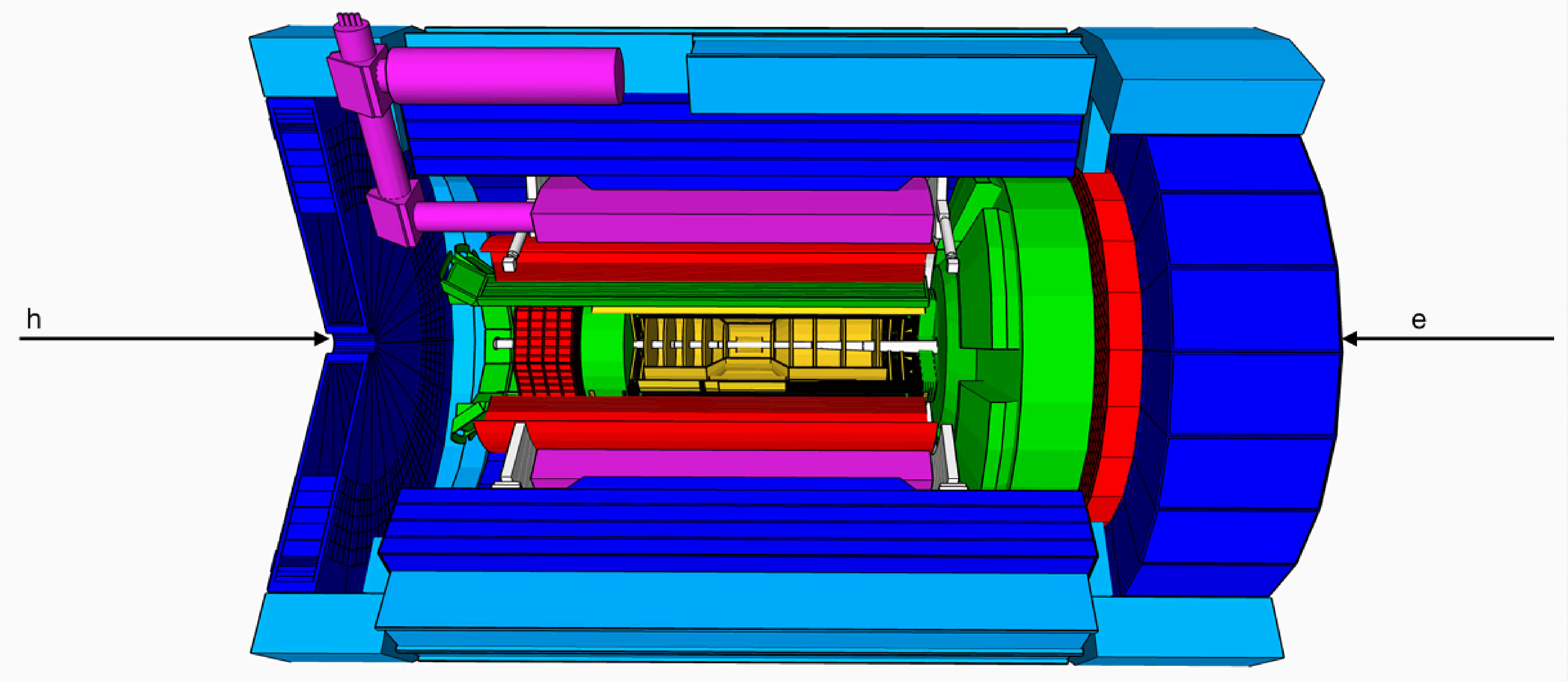


Fig 1: The electron-proton/ion collider (ePIC) detector

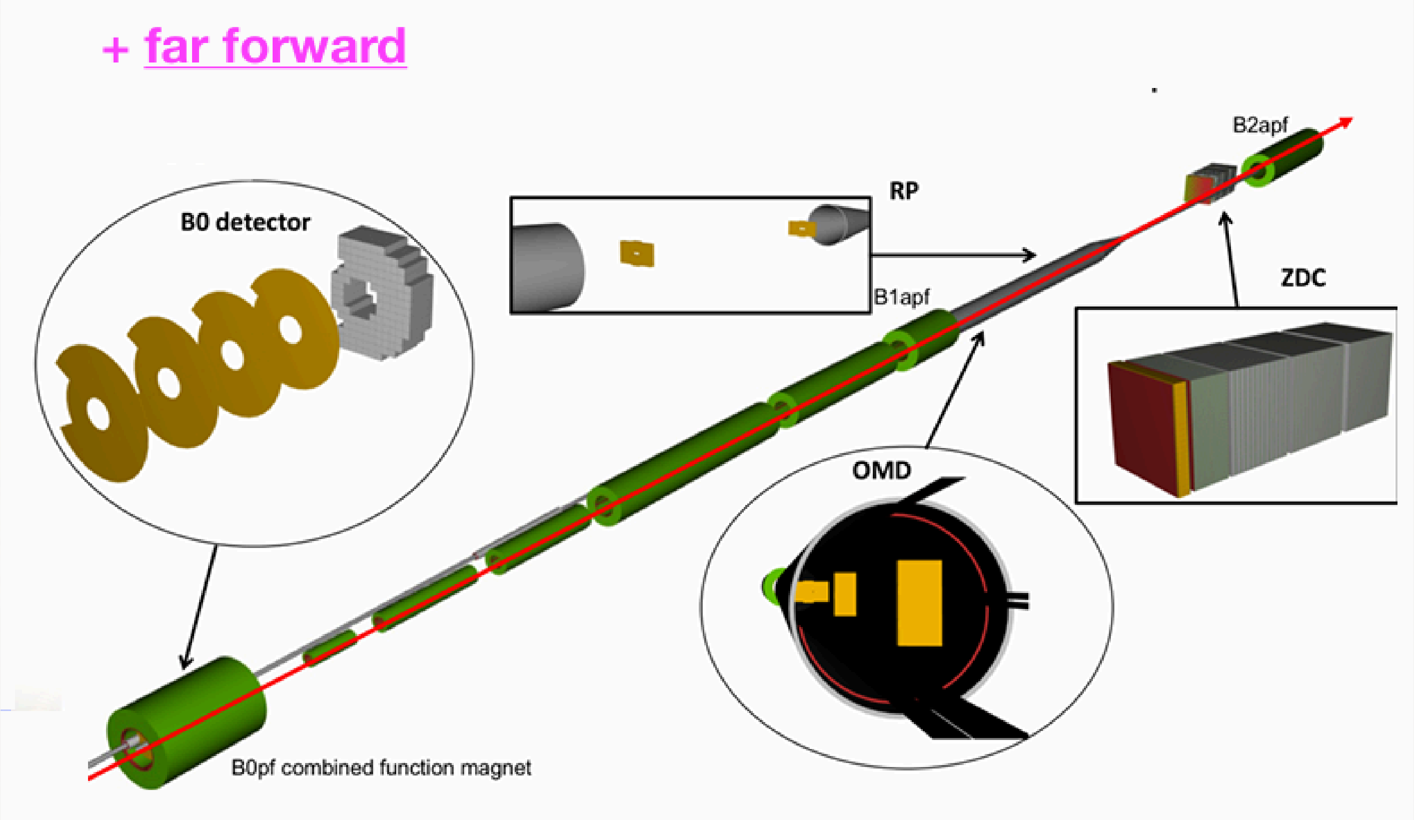


Fig 2: The zoomed in far-forward part of the ePIC detector.

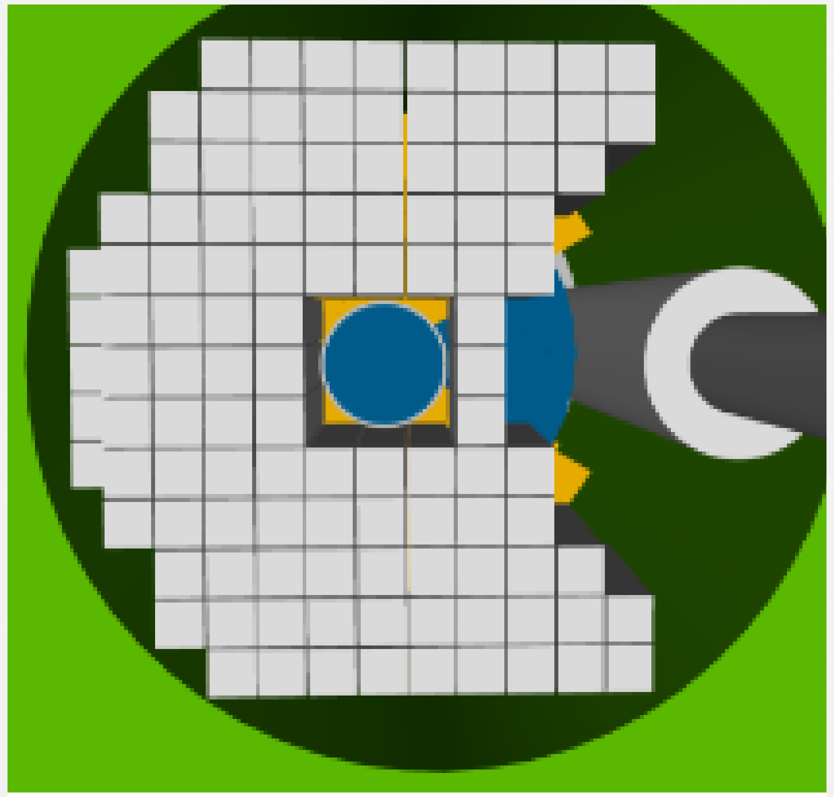


Fig 3: B0ECal calorimeter detector

Reference: [2]

B0ECal Detector Geometry

- Built from PbW04 crystals (2x2x20 cm3) read out by SiPMs.
- Measures forward particles with high precision.
- Primarily intended for photon detection.

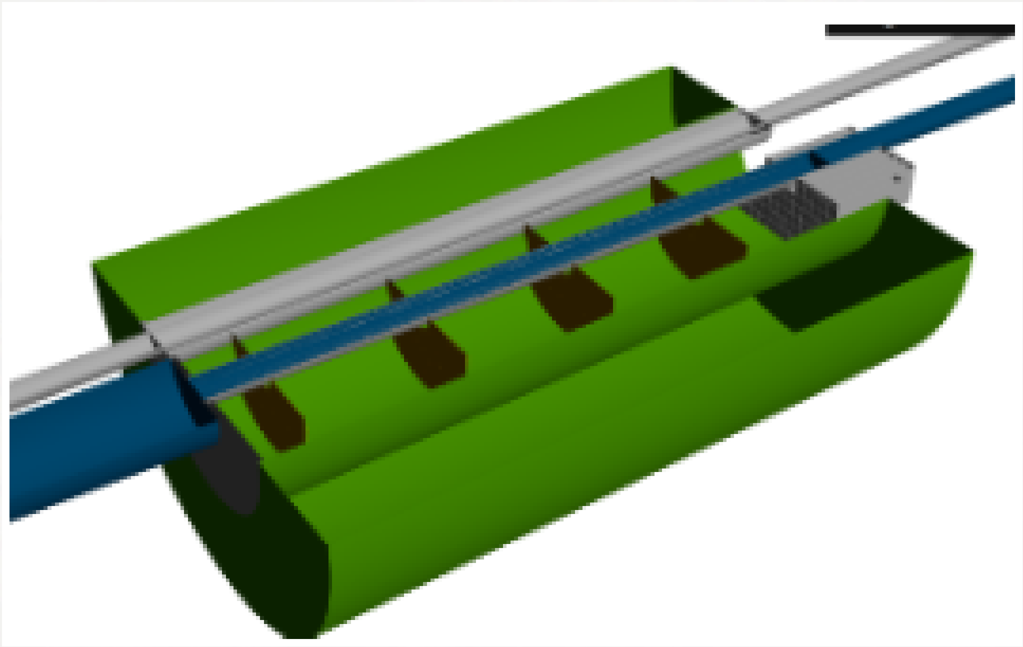


Fig 4: B0 Detector Geometry

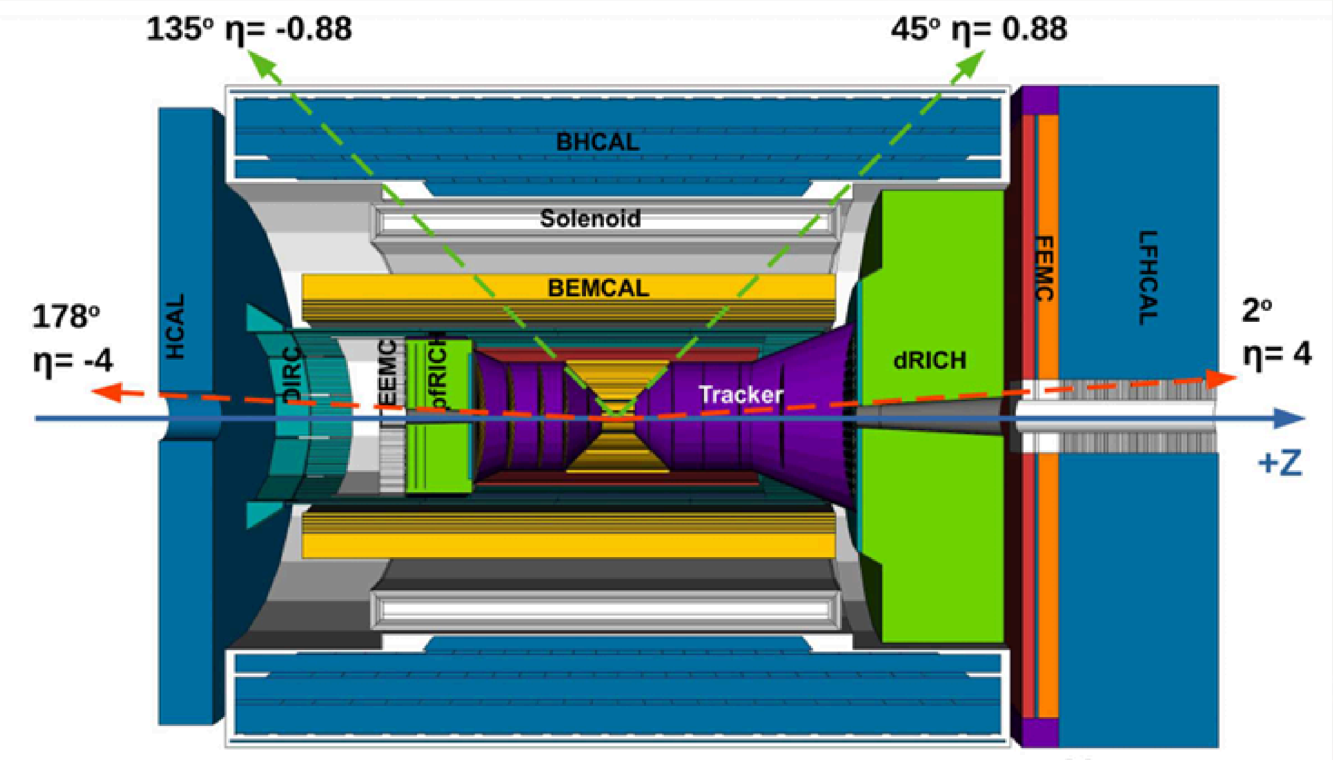


Fig 5: ePIC Detector Geometry

- **Barrel detectors:** $|\eta| < 1.5$
- **Electron endcap:** $-3.5 < \eta < -1$
- **Hadron endcap:** $1 < \eta < 3.5$
- **B0 region / forward detectors:** $\eta \approx 4–6$

Problem Statement

Problem: Charged particles were expected, but neutrons dominated in the reconstructed particle hits in **DEMP** interaction of the **B0ECal** detector simulation.

- Can neutrons be distinguished from photon hits?
- Develop a data-driven approach to separate genuine neutron clusters from background using ML and DNN.

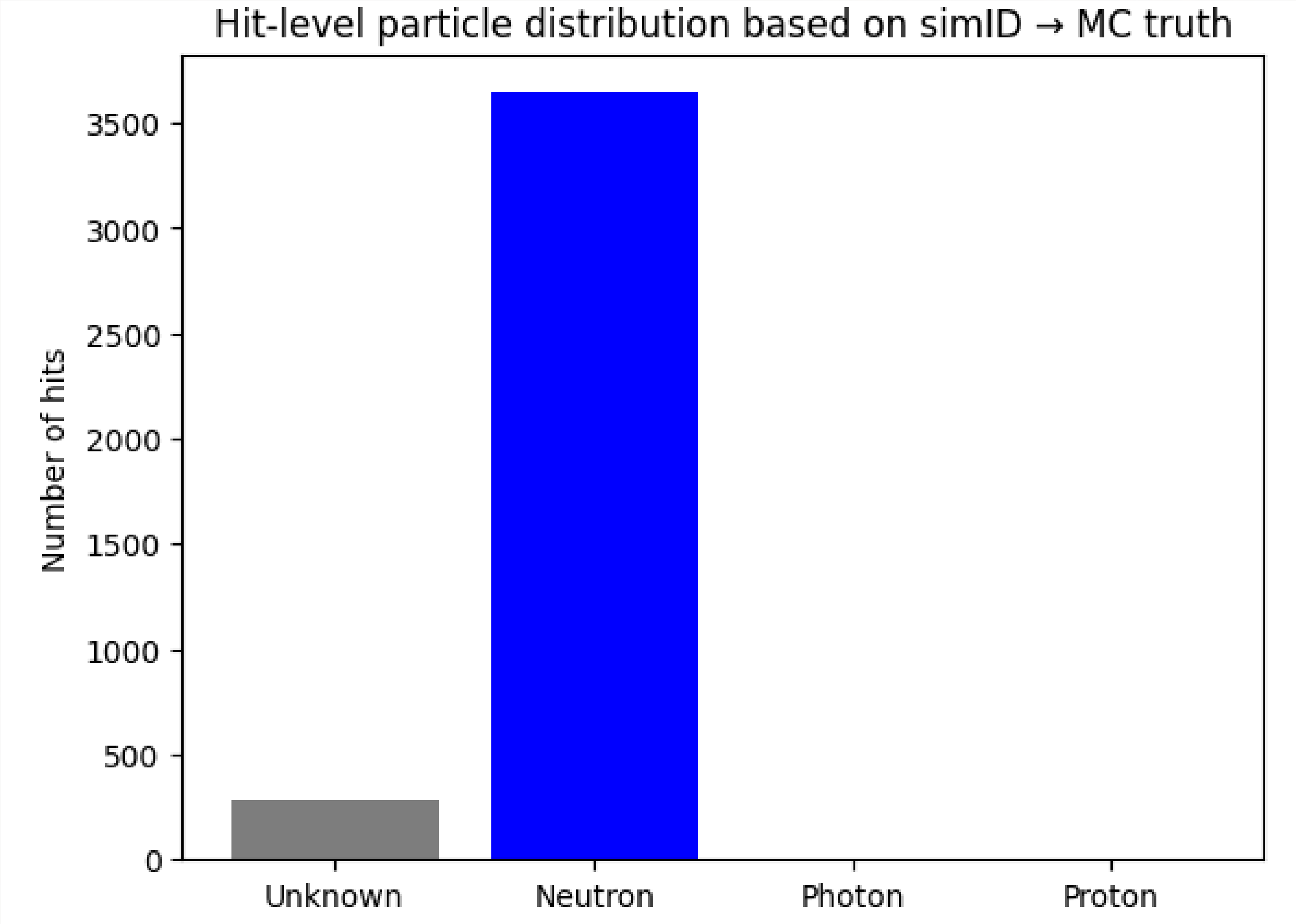


Fig 7: Hit-level particle distribution

Project Motivation and Goals

Motivation: To enhance neutron identification, leading to more accurate reconstruction of DEMP events in the simulated B0ECal detector.

Outline:

1. Basic analysis of cluster properties
2. ML-based classification
3. DNN-based optimization
4. Verification from Simulation

Methodology

- Used MC truth matching of reconstructed data to identify genuine neutron clusters.

- Observed:

1. Large number of DEMP neutrons reconstructing in B0 (for 5x41).
2. Some neutron clusters misreconstructed due to nearby hits.

- Defined a set of “good” neutron clusters for training and testing.

Data & Basic Analysis

Data:

- Simulation data used: ‘/volatile/eic/sjdkay/Mar2025_DEMP_RecoOut’ (Remote access) from GEANT4 based simulation
- Total files used: 6288
- Total events actually extracted: 17000

Variables studied:

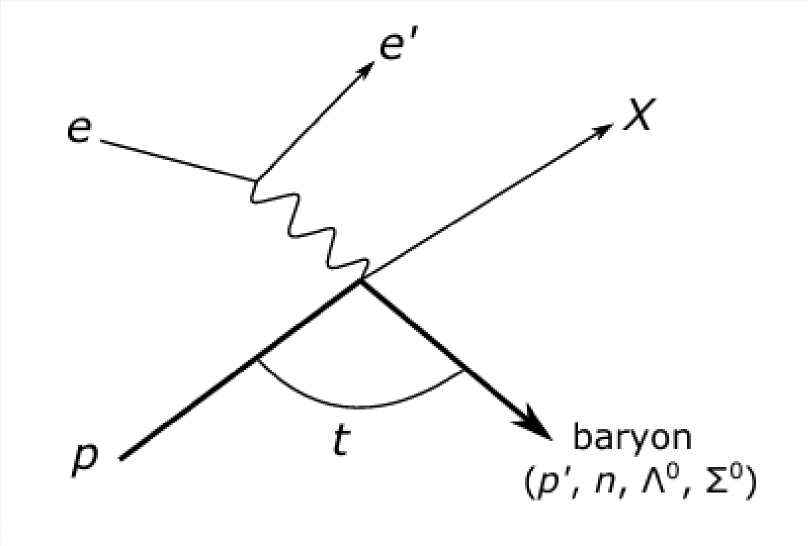
- Cluster energy (E)
- Number of hits
- Shower width / shape
- Cluster position (η, φ)

-t Distribution Graphs

The –t distribution represents the squared four-momentum transferred from the initial proton to the final-state neutron in the exclusive reaction.

BABE Method Equation used:

$$-t = (p^{\mu}_{\mathrm{BA}} - p^{\mu}_{\mathrm{BE}})^2 \text{ or } t = |\vec{p}^{\,\mathrm{T}}_{\mathrm{BA}}|^2$$



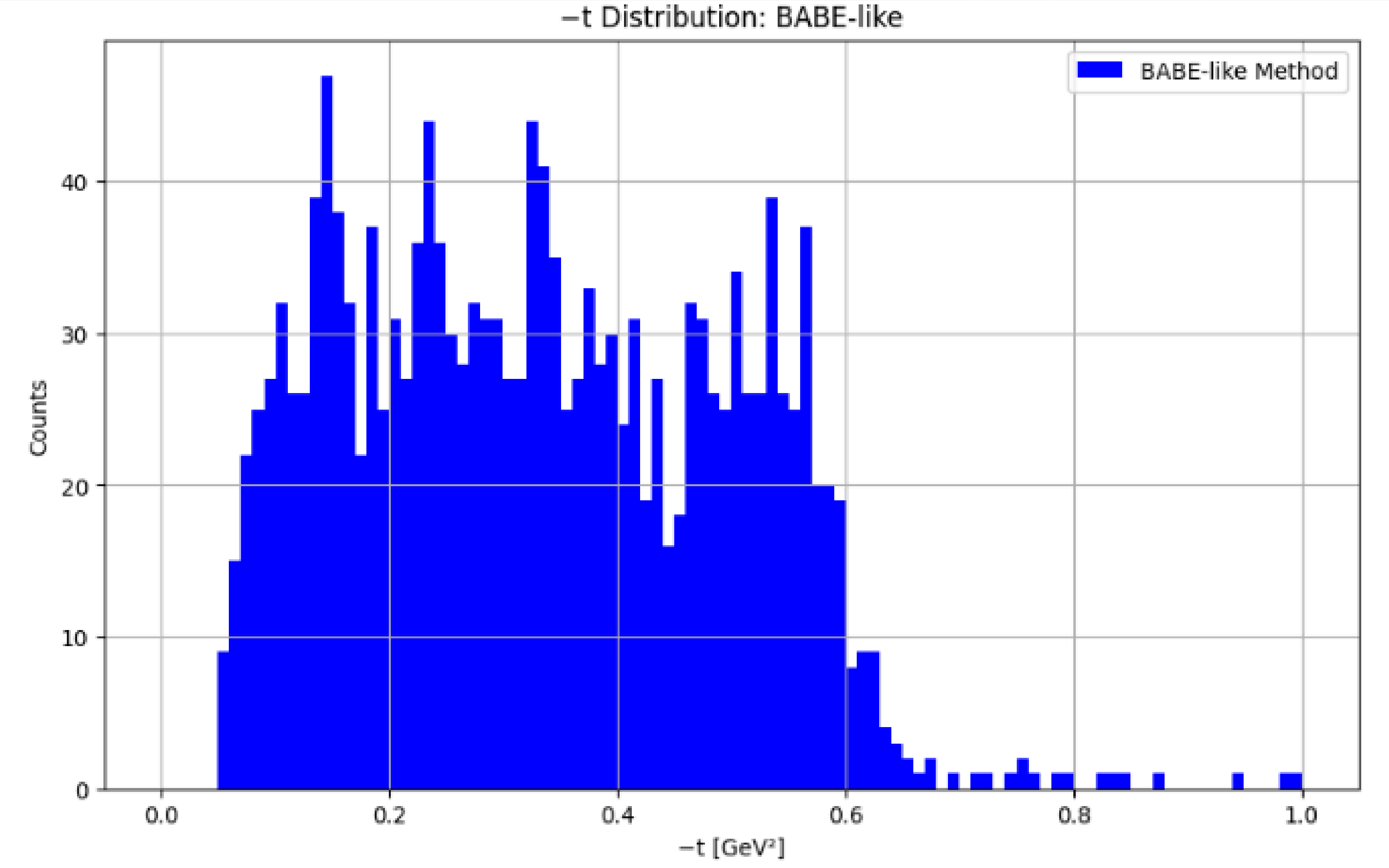


Fig 8: Momentum transfer -t graph for BABE method

Energy and Distance graphs

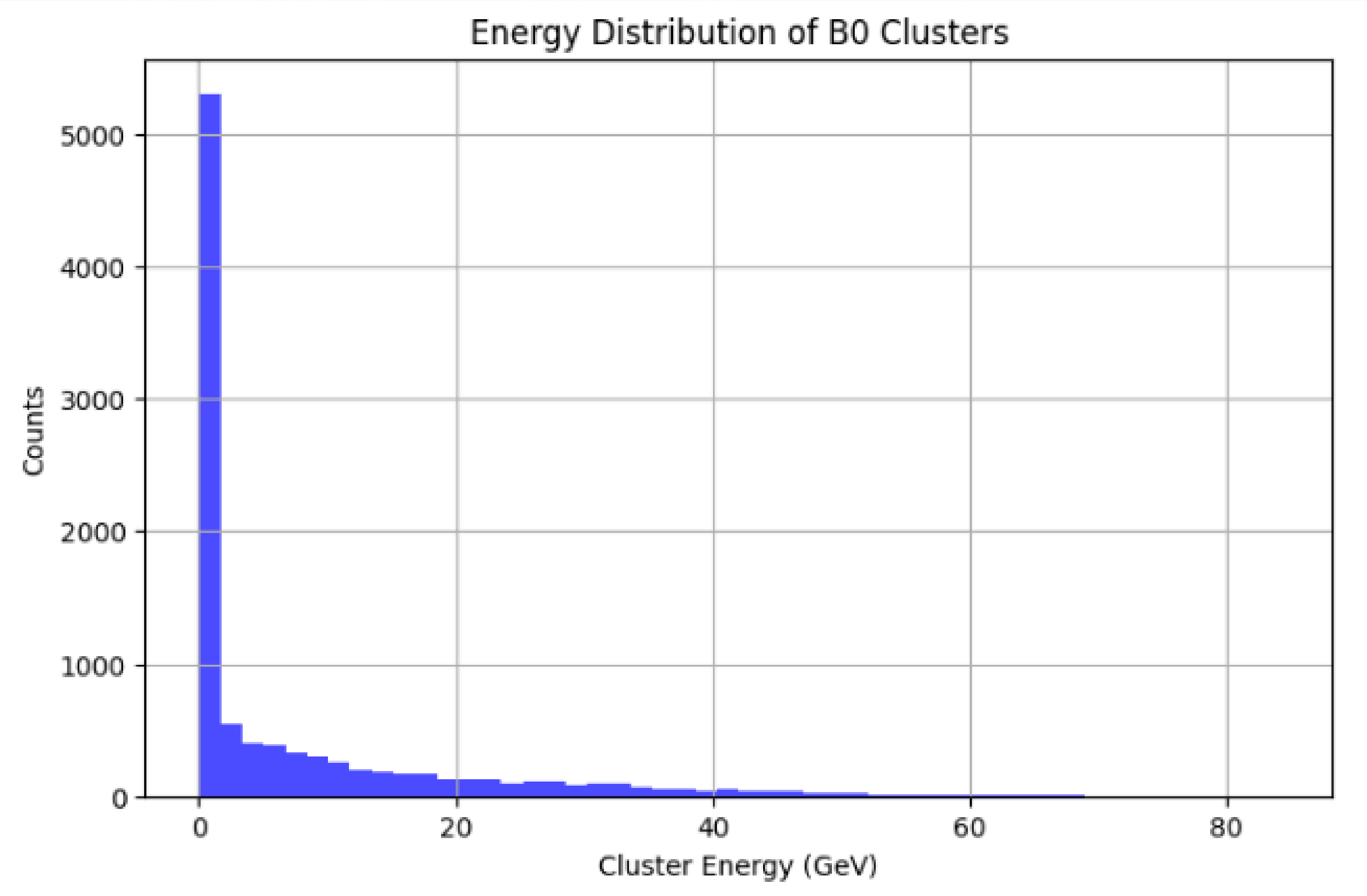


Fig 10: Most of the clusters have low energy

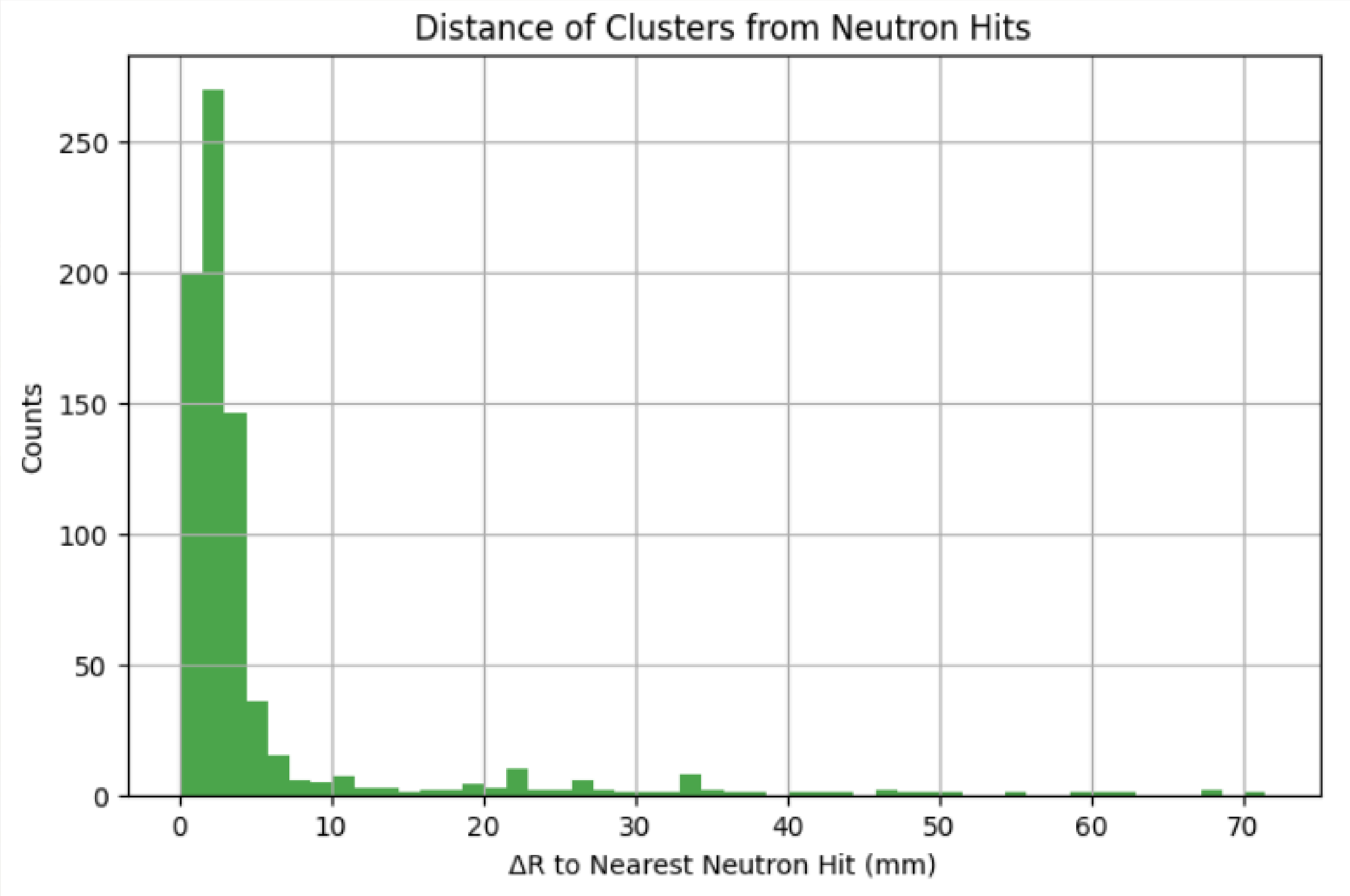


Fig 11: The distance between clusters is below a certain threshold.

Machine Learning Approach

Algorithm: Random Forest and Gradient Boosted Decision Trees

Input features

Energy, position, cluster RMS, shower length, etc.

Training

80% training / 20% validation split.

Output

Probability of being a neutron cluster.

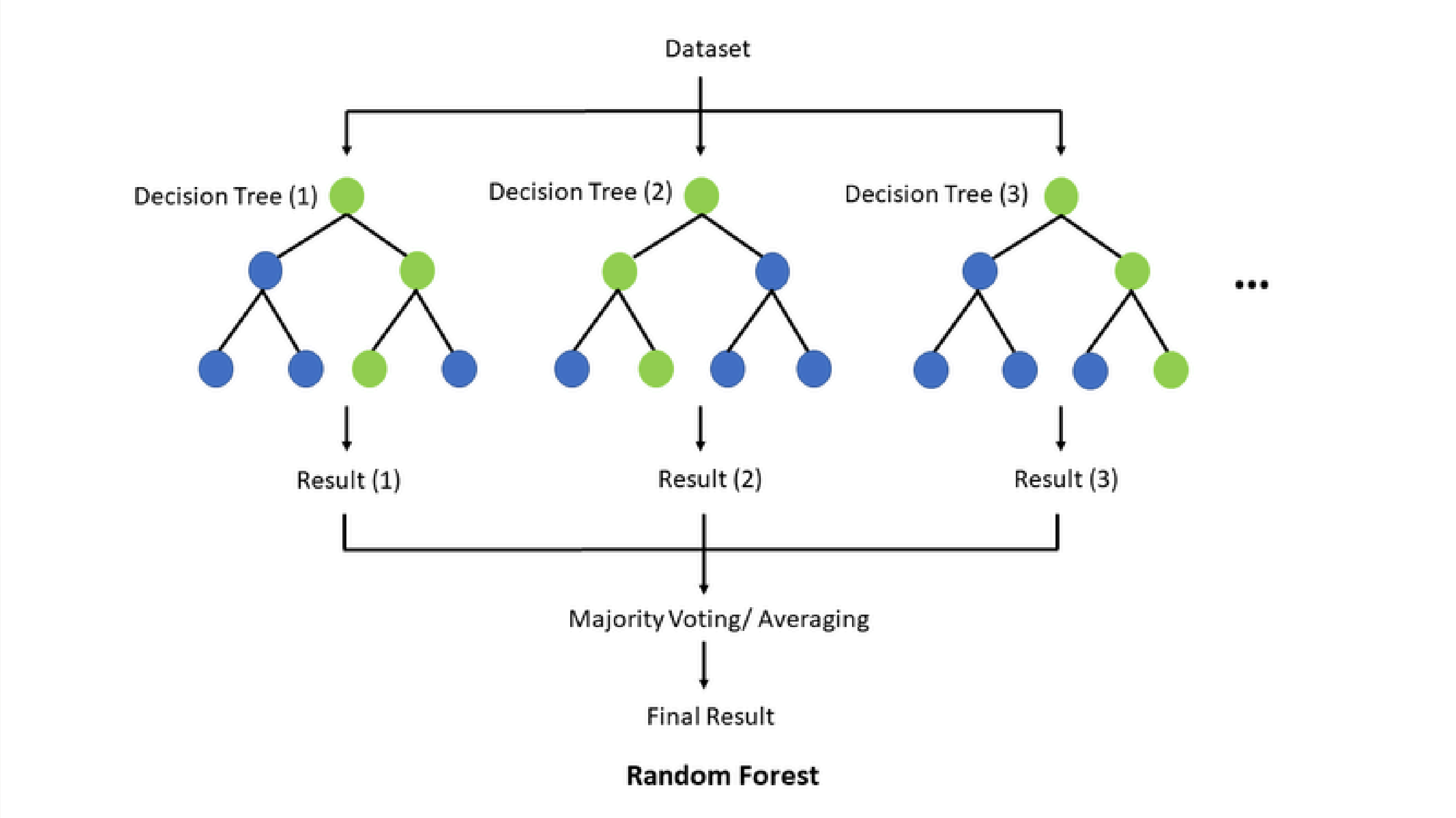


Fig 12: Random Forest

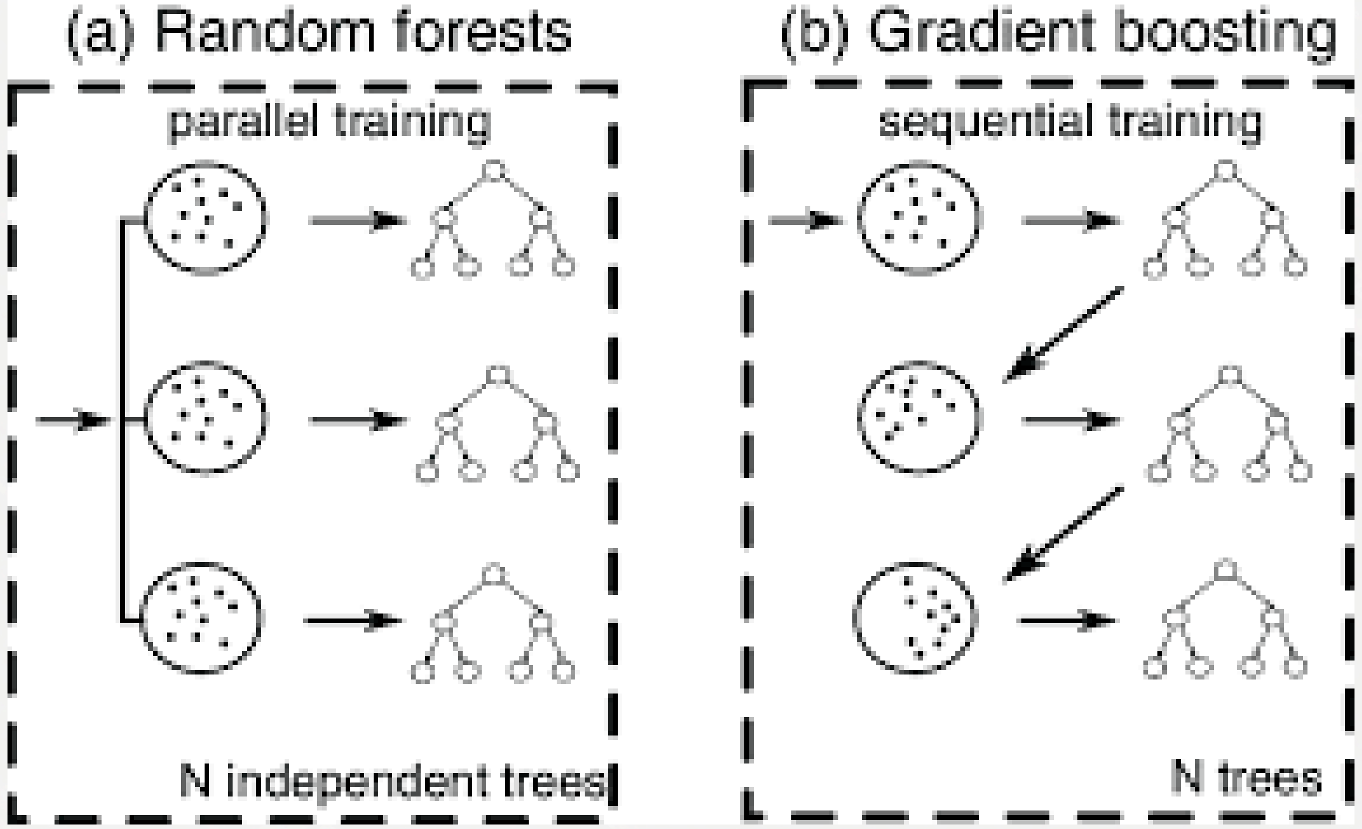


Fig 13: Gradient Boost

Labelled clusters on the basis of Distance

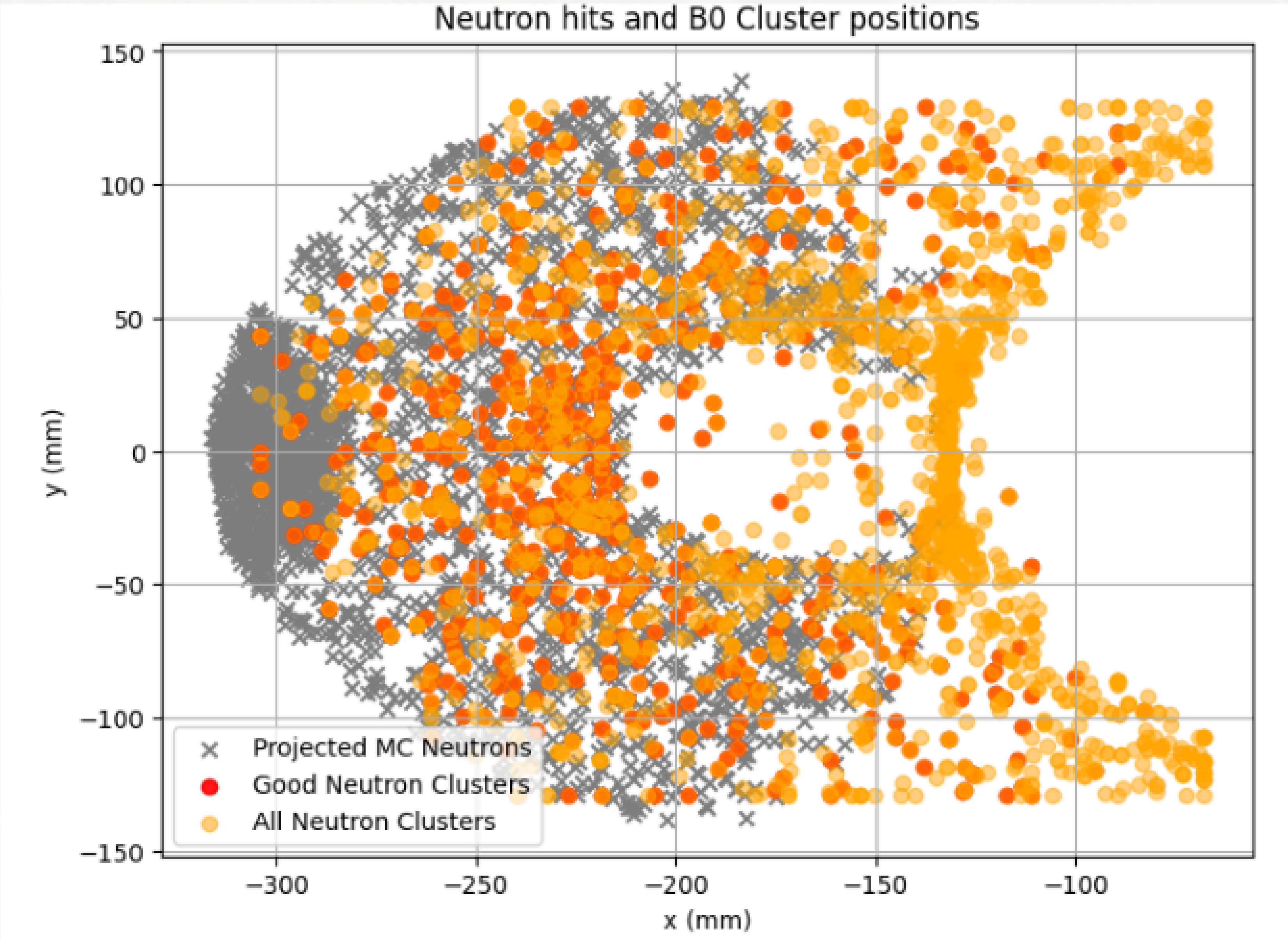


Fig 14: Neutron clusters selected

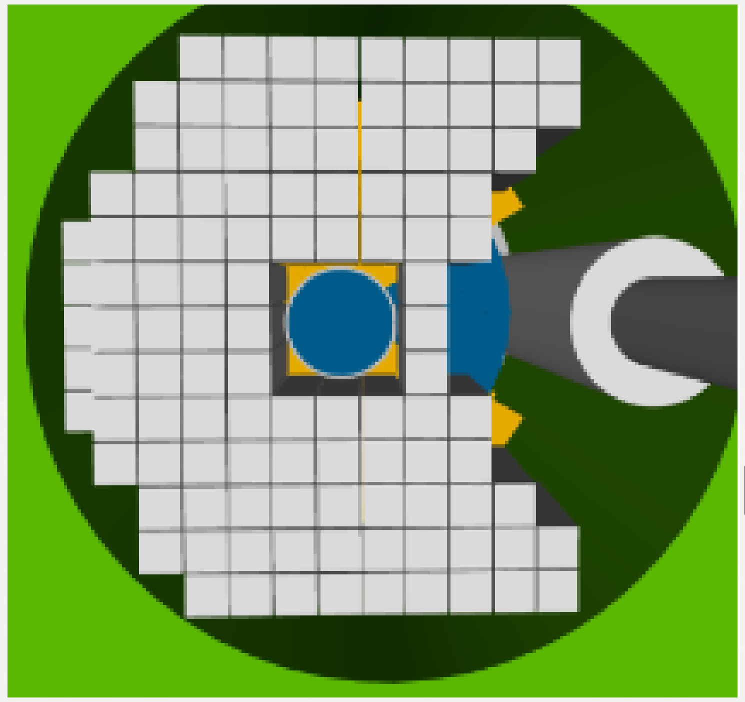


Fig 15: B0ECal Detector shape

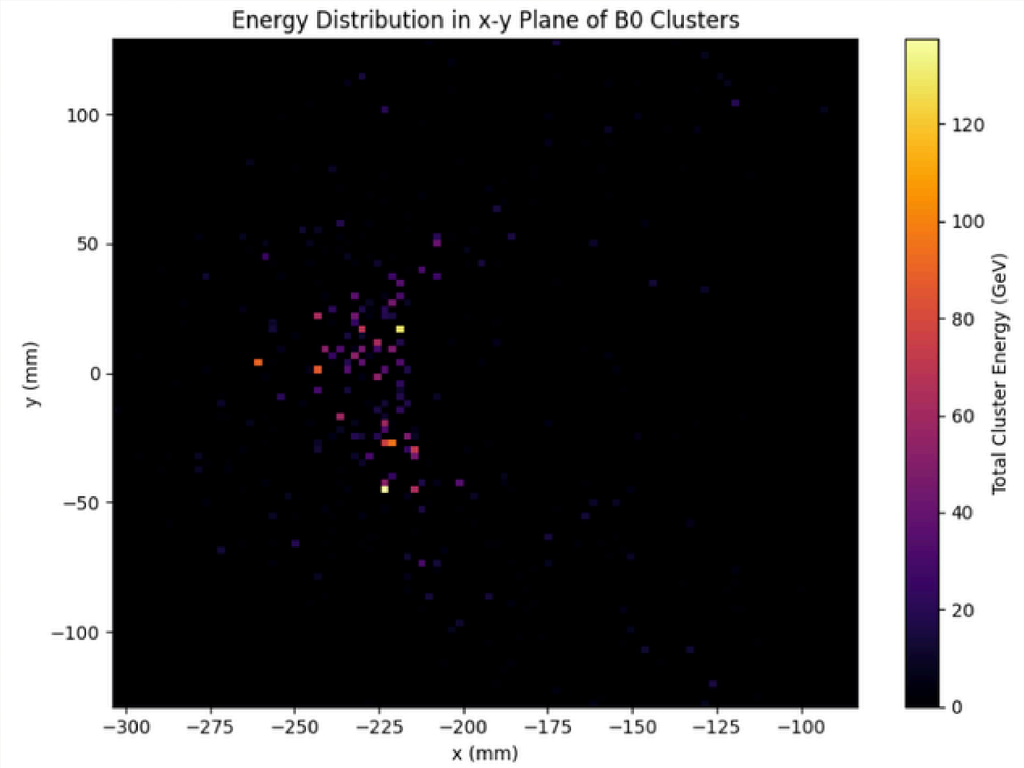


Fig 16: Energy distribution heatmap

Comparison of selected neutron clusters in the B0ECal detector shape.

It perfectly takes the shape of the detector.

Deep Neural Network (DNN) Approach

- Input features: Reconstructed hit energy (E), and position (x, y, z).
- Goal: Classify each hit as neutron, photon, or proton.
- Model: Simple feedforward DNN with two hidden layers (64 neurons each, ReLU activation).
- Output layer: 3-class softmax for particle type prediction.
- Training:
- Optimizer: Adam
- Loss: Sparse categorical cross-entropy
- 20 epochs, 80/20 train-validation split

Result: Achieved good accuracy in separating hit types based on calorimeter energy–position patterns.

Results & Comparisons

Random Forest

- Accuracy = 0.757
- Used direct neutron cluster data

Gradient Boost Classifier

- Accuracy = 0.767
- Used direct neutron cluster data

DNN

- Accuracy = 1.0
- Used neutron hits data
- Overfitting

Photon Gun Simulation

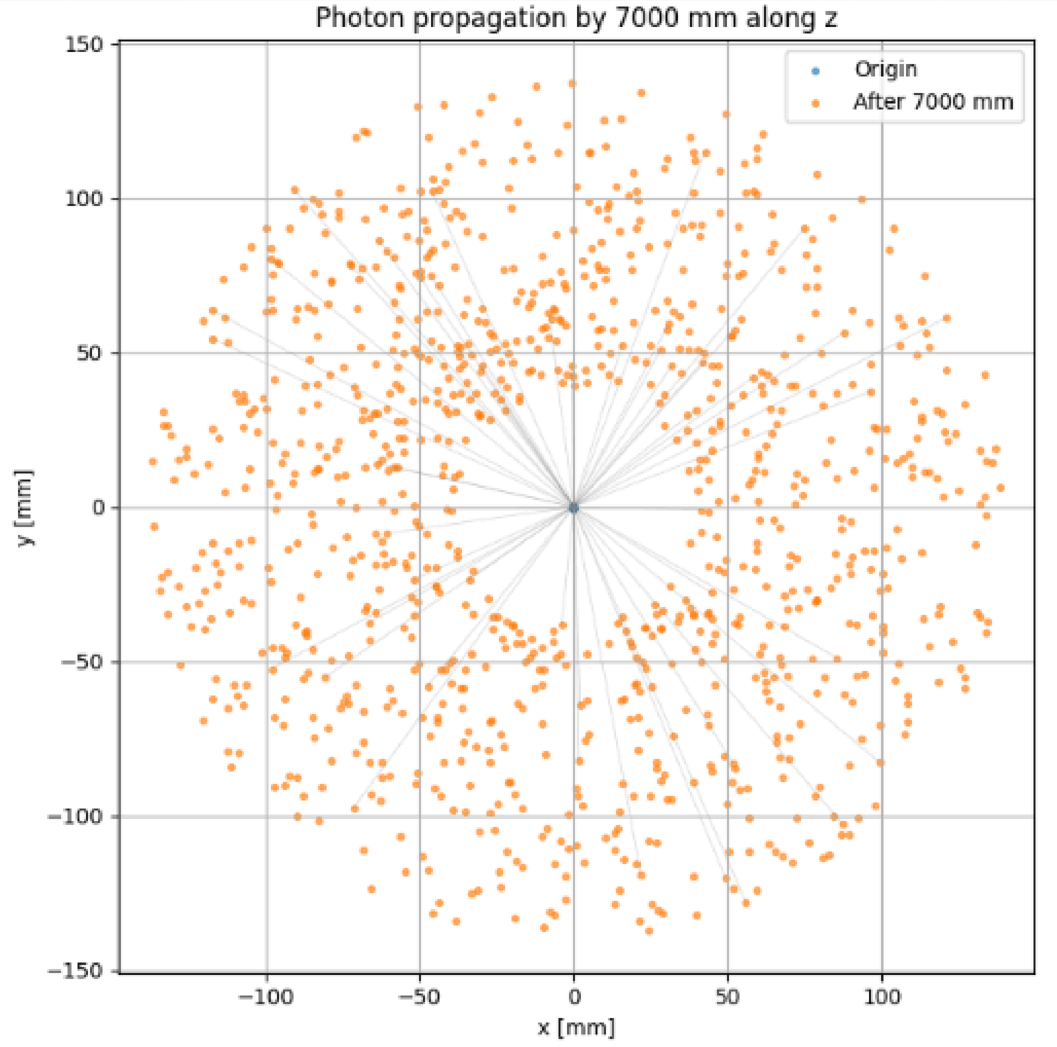


Fig 17: Photon Gun Beam

EIC Software

GEANT4 Simulation

No. of events : 1×10^6

B0ECal Angle: 0.0055 to 0.020 rad

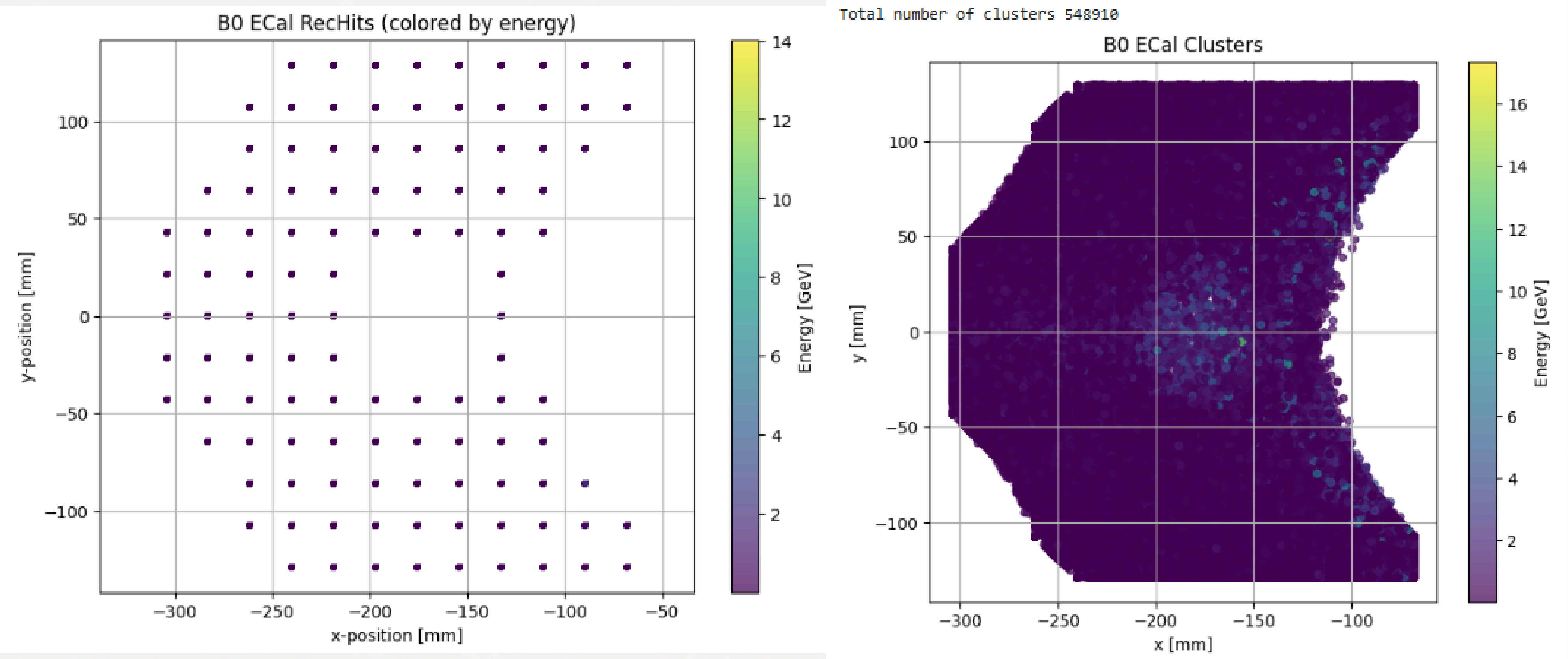


Fig 18: Reconstructed Photon Hits and Clusters

Cluster Size Comparison

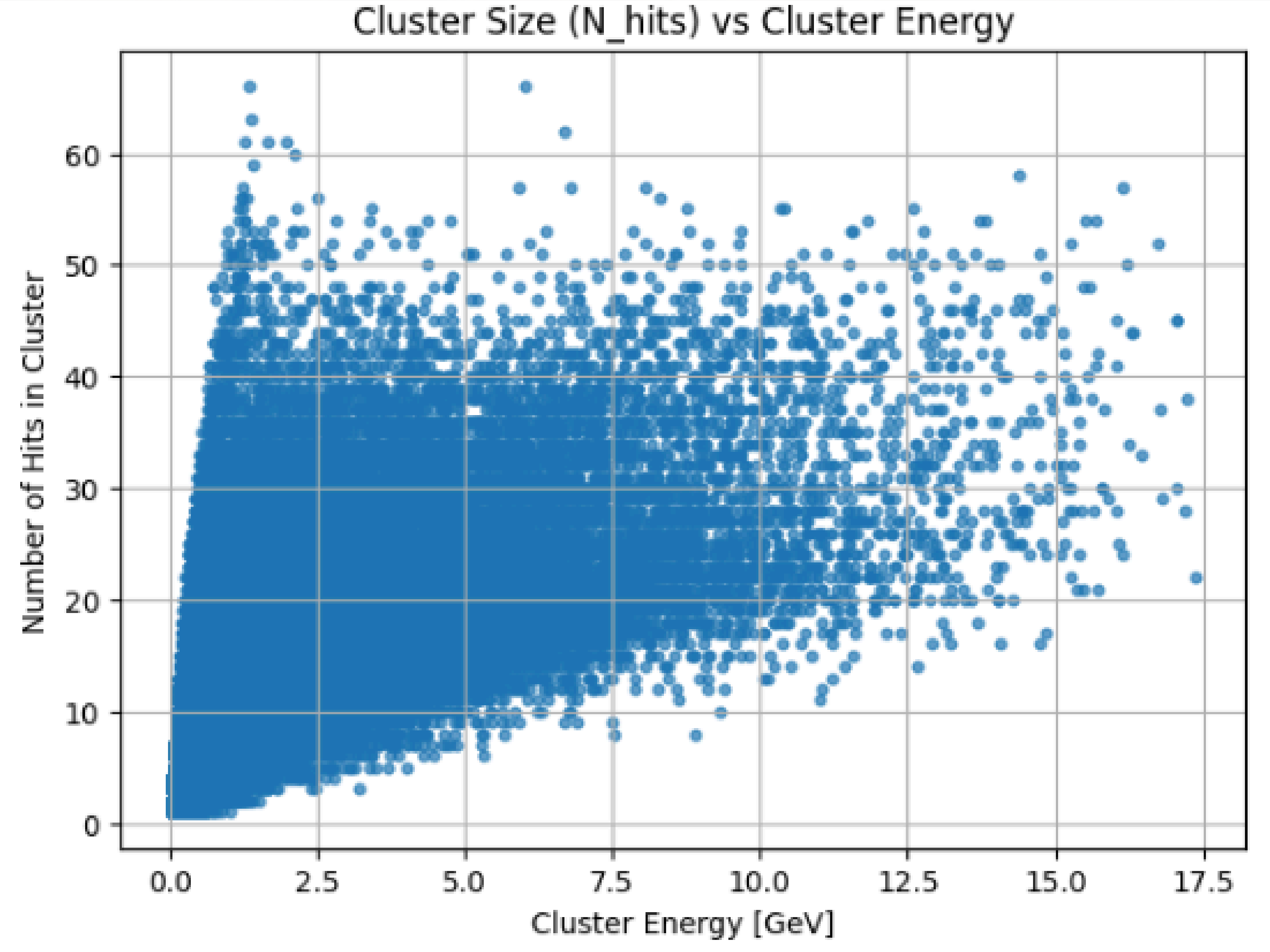


Fig 19: Cluster size vs energy for photons

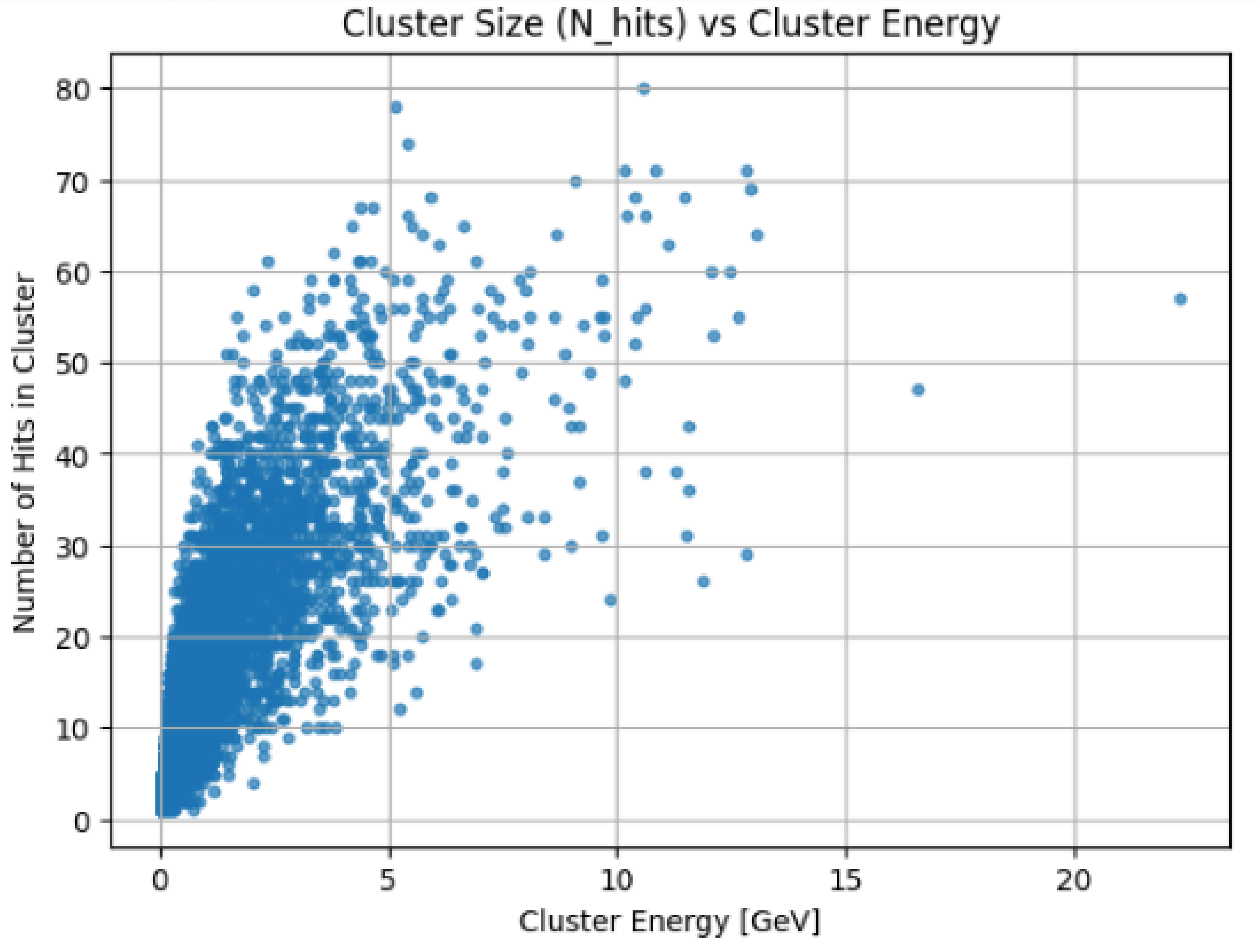


Fig 20: Cluster size vs energy for neutrons

Comparison between Photon and Neutron

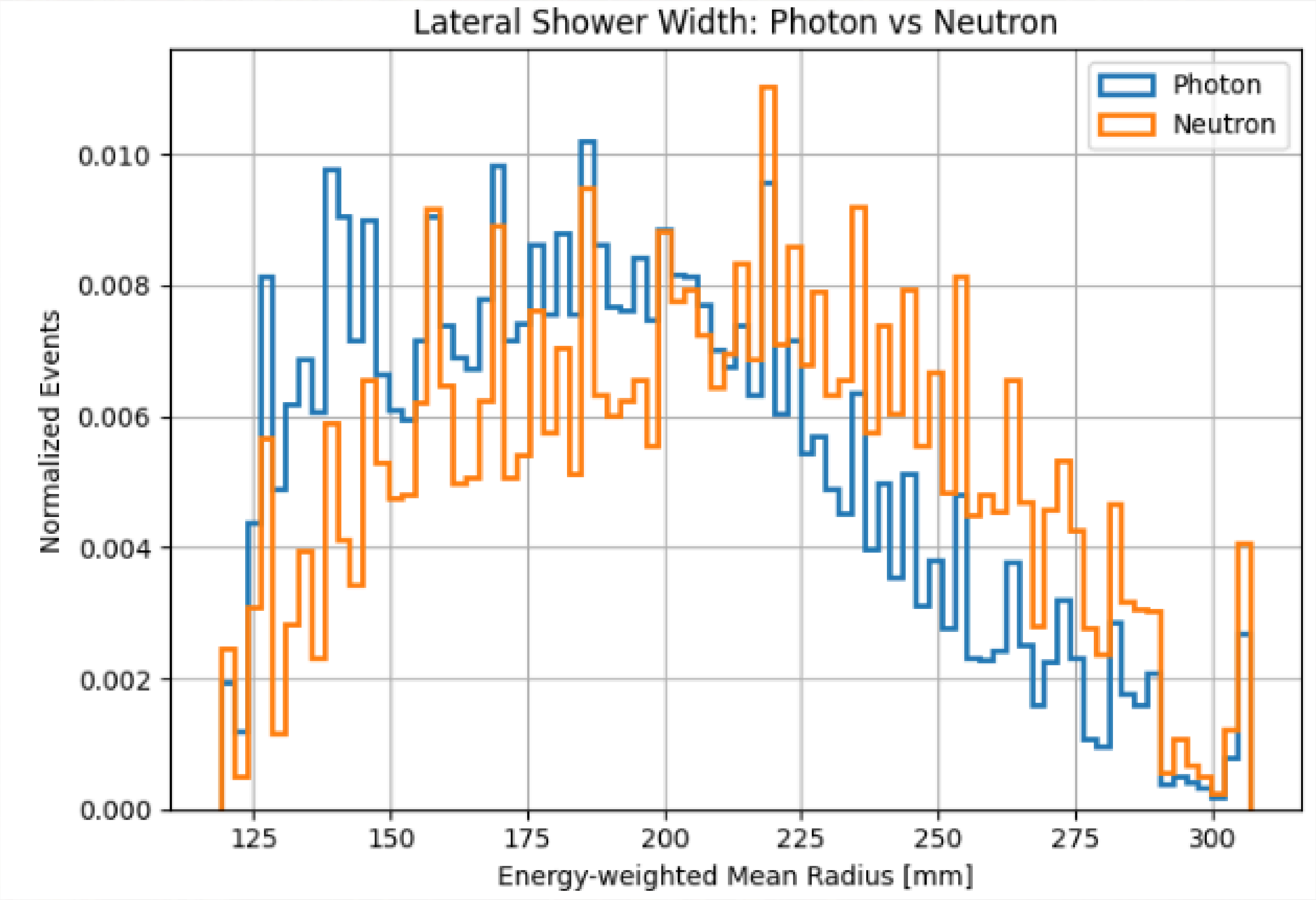


Fig 21: Lateral Shower Width

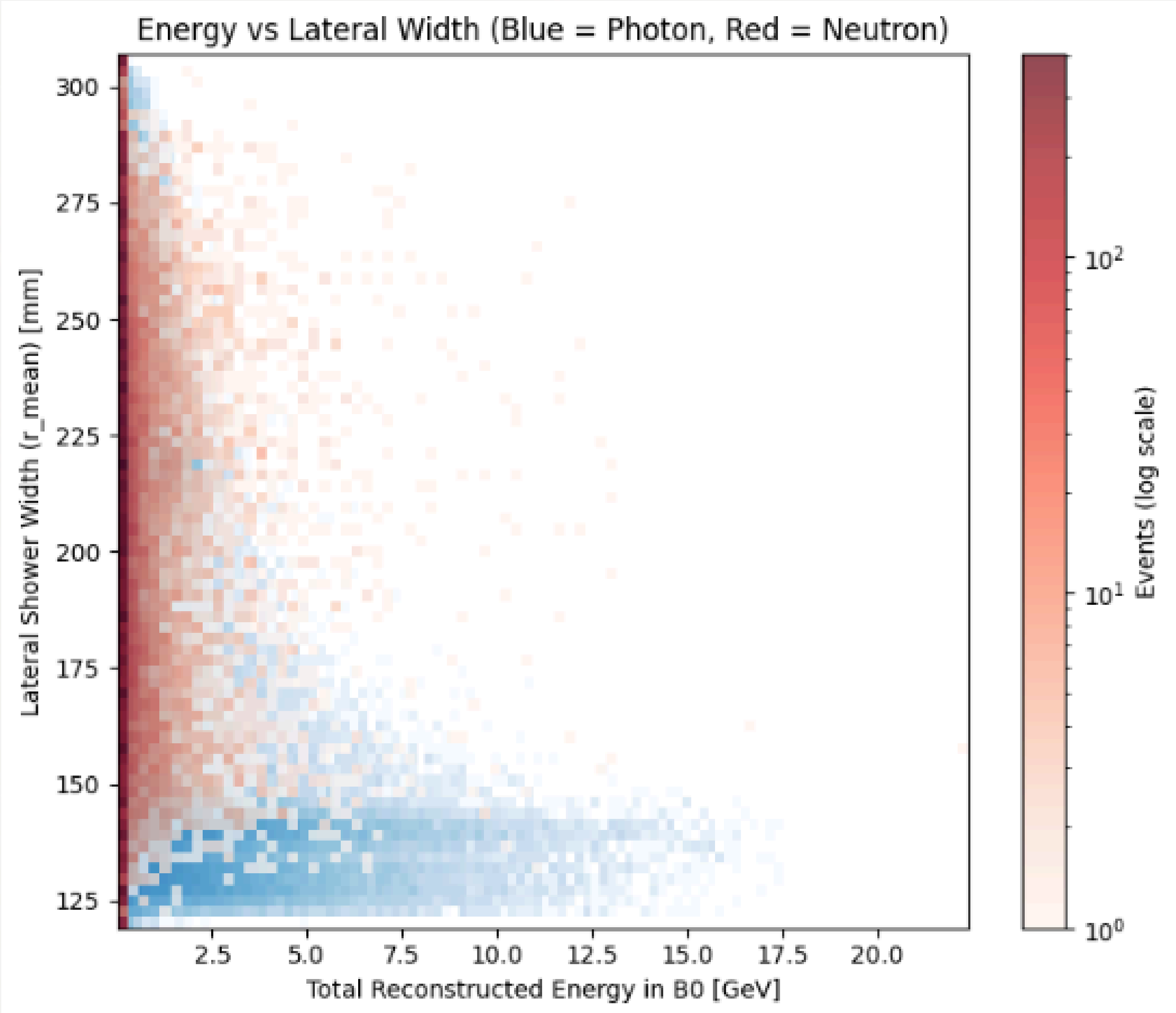


Fig 22: Energy vs Lateral Width

Neutron Detection Efficiency vs Hit Radius

ϵn = True neutrons / Reconstructed neutrons

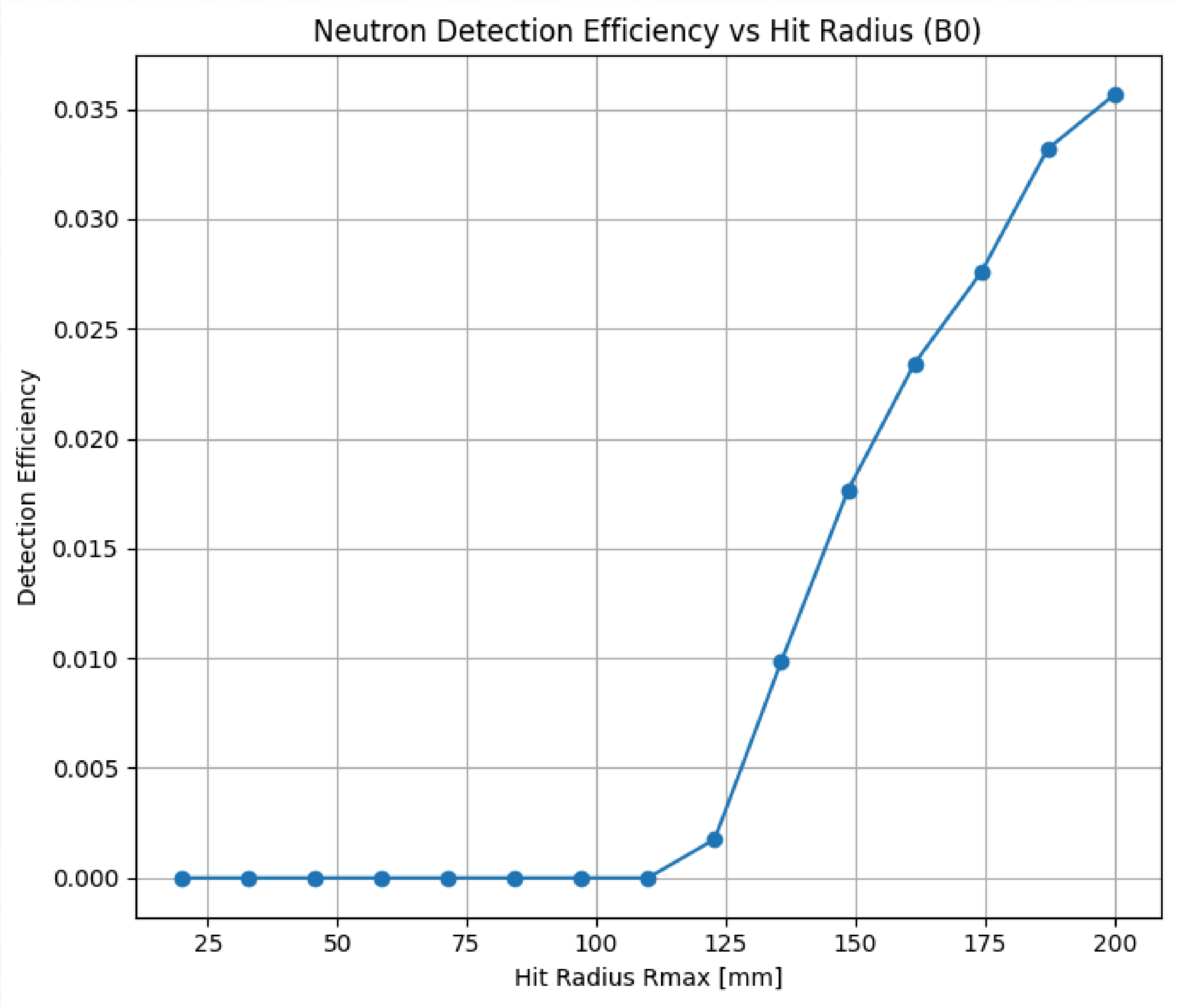


Fig 23: Efficiency for Neutron

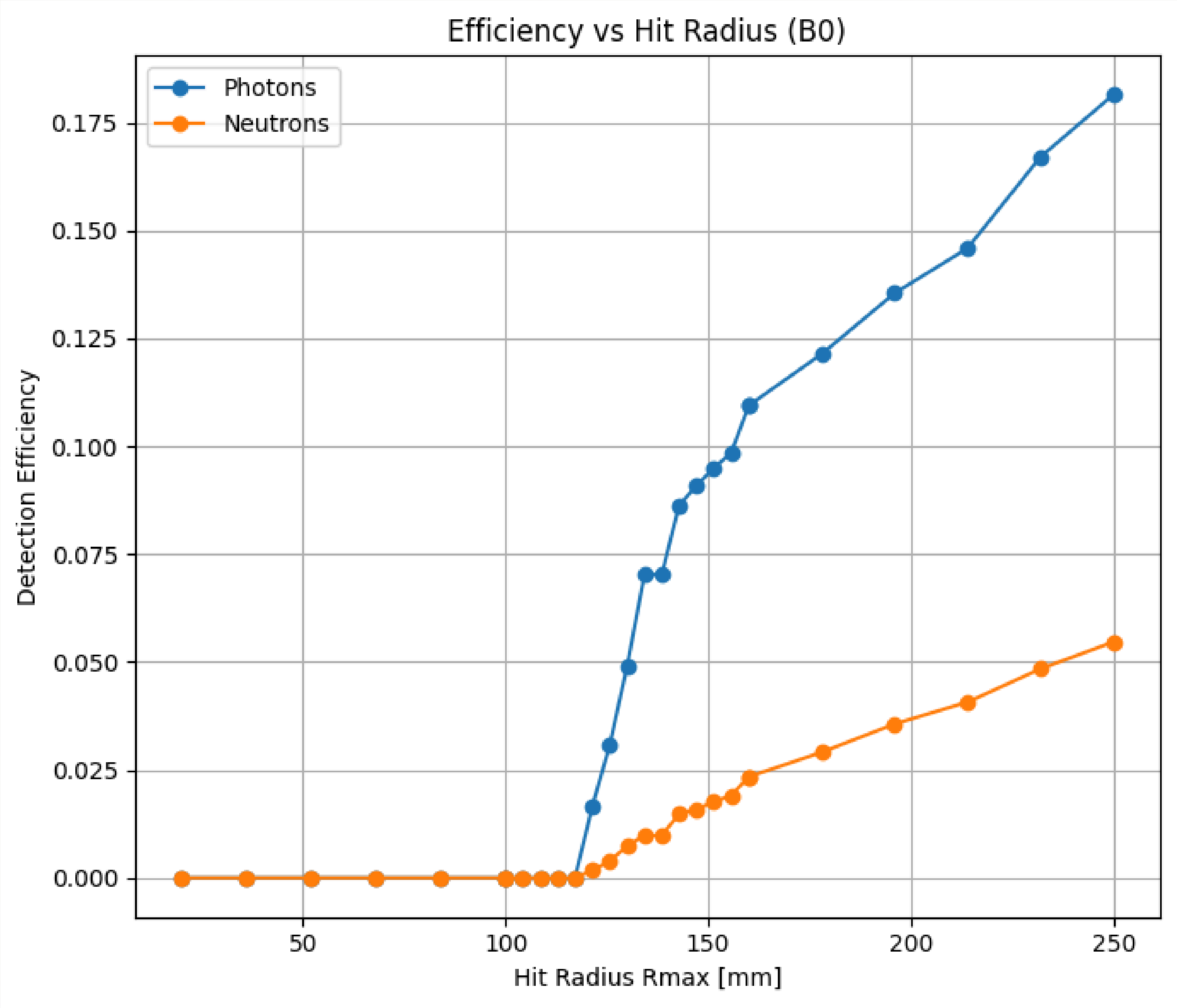


Fig 24: Efficiency comparison for Neutron and Photon

Energy vs Shower Width Separation

- 2D plot: Cluster Energy vs Shower Width
- Photons → narrow width at given energy
- Neutrons → broader distribution
- Partial separation visible in specific energy window
- Photons > 7.5 GeV
- Neutrons > 5 GeV and width > 180 mm

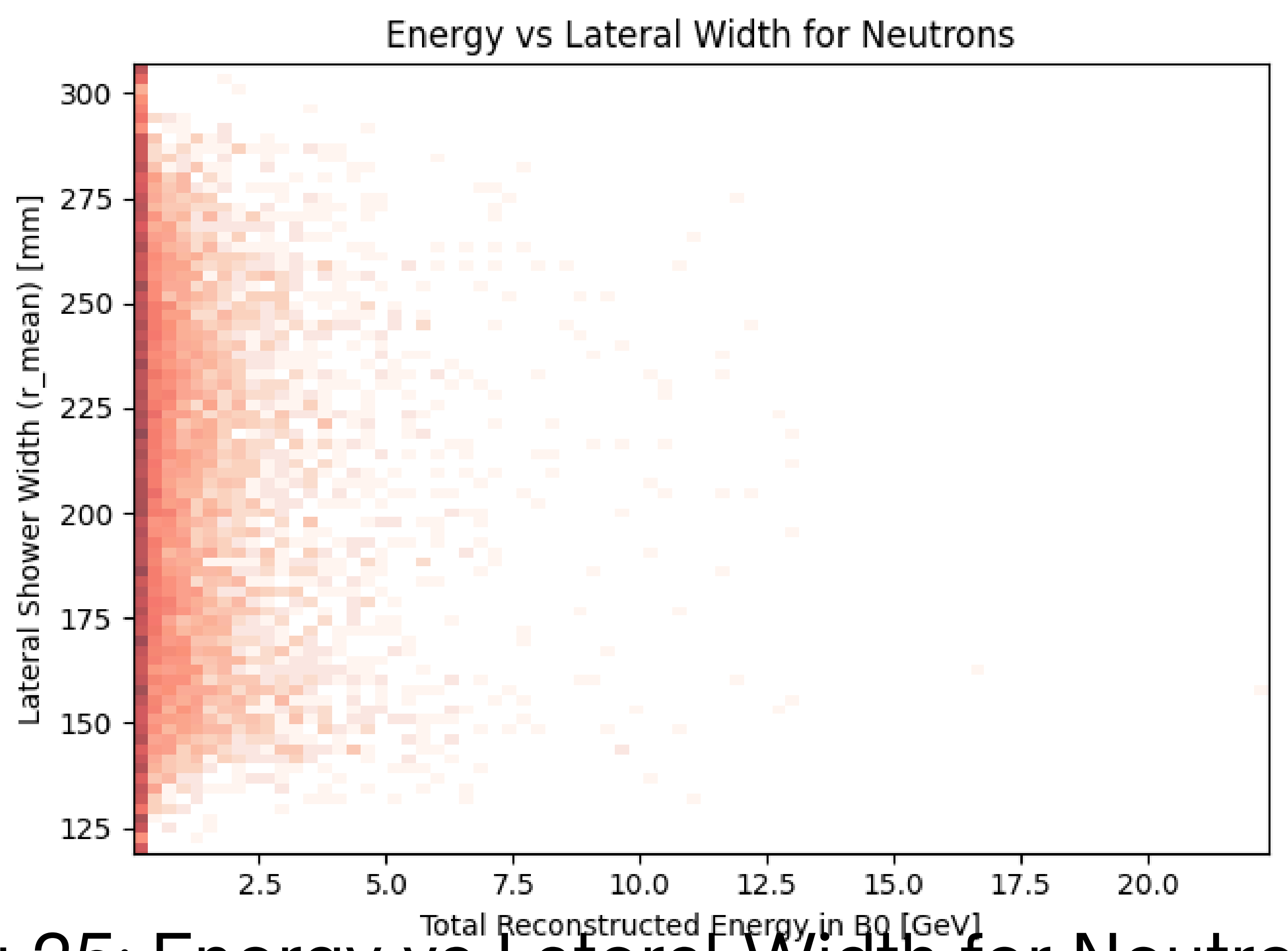


Fig 25: Energy vs Lateral Width for Neutrons

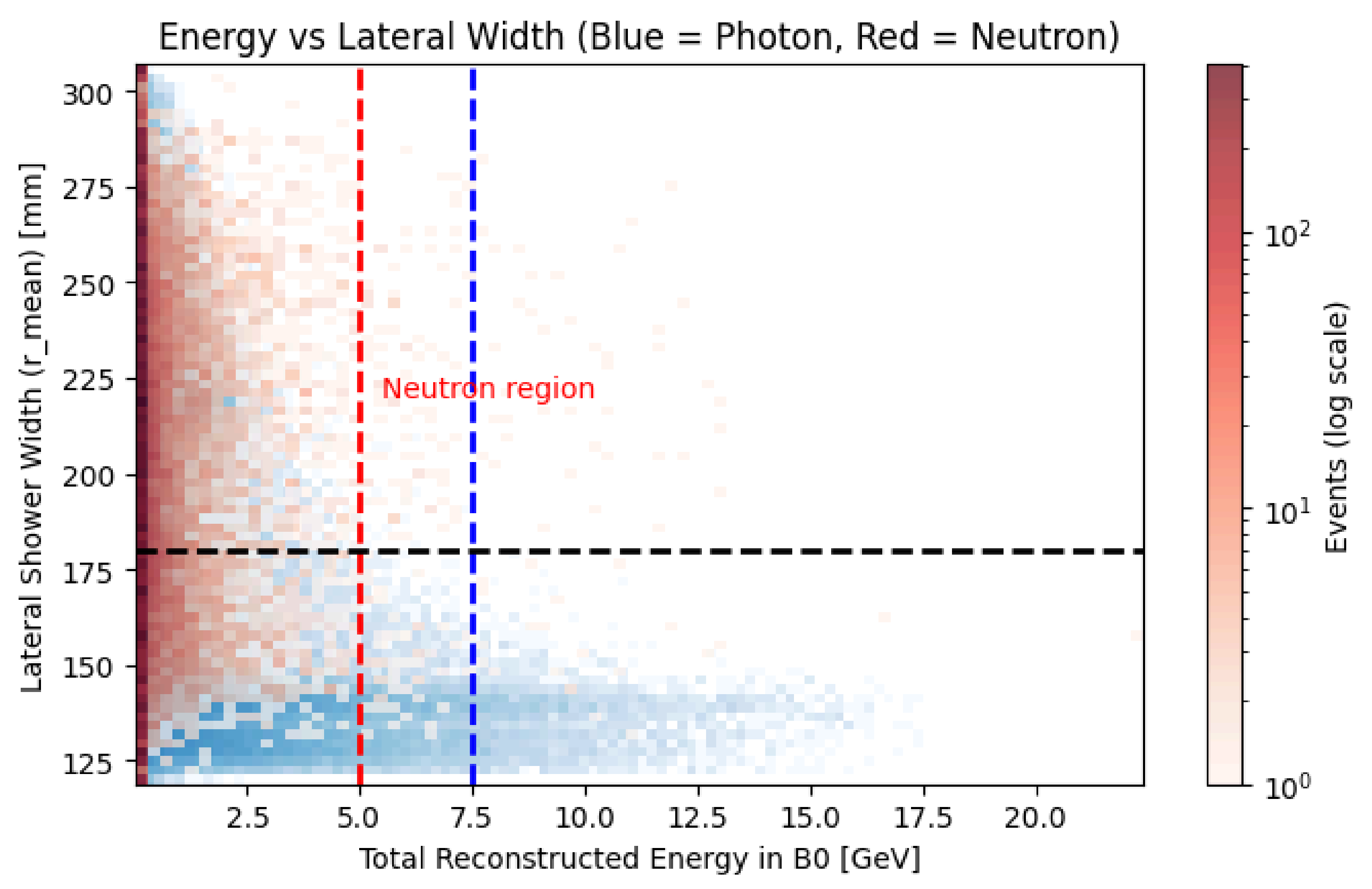


Fig 26: Energy vs Lateral Width

Conclusion

- Separation between neutron and photon in B0 is challenging
- Neutrons leave signals, but typically weaker and more diffuse.
- ML improves neutron-photon separation.
- Shower width is a key observable for discrimination.

Photons > 7.5 GeV show narrow EM showers → separable from neutrons.
Neutrons > 5 GeV and width > 180 mm distinguishable from photons.

- Effective separation achievable in this energy and shower-width range.

Future Work

Phase 1
Integrate timing and cluster topology features.

Phase 2
Integrate lateral shower width within the DNN model.

Phase 3
Include photon/proton gun studies for better background modeling.

Phase 4
Apply DNN outputs to exclusive reaction analysis:
$(e + p \rightarrow e' + \pi^{+} + n)$.

References

[1] Z. Ahmed, R.S. Evans, et. al., DEMPgen: Physics event generator for Deep Exclusive Meson Production at Jefferson Lab and the EIC, Computer Physics Communications, Volume 308, 2025, 109444, ISSN 0010-4655, https://doi.org/10.1016/j.cpc.2024.109444.

[2] E. C. Aschenauer, V. Batozskaya, et. al., Study of deeply virtual Compton scattering at the future electron-ion collider, 112, 10.1103/fy8y-bjc9, Physical Review D

[3] Gregory Matousek, Anselm Vossen, AI-Assisted Object Condensation Clustering for Calorimeter Shower Reconstruction at CLAS12, https://doi.org/10.48550/arXiv.2503.11277

Thankyou

Vani Mahajan