AI Tracking for ALERT hyperbolic drift chamber

Mathieu Ouillon (Mississippi State University)

6/18/25

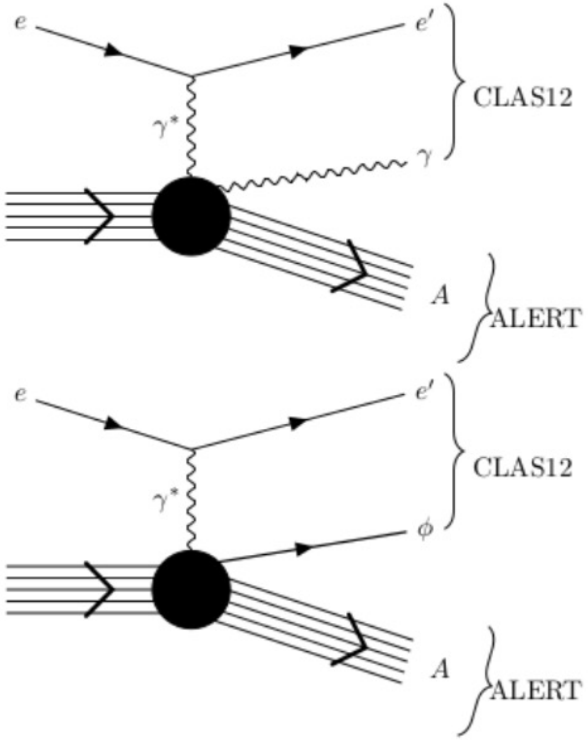
Outline

1. ALERT Physics Program
2. ALERT Experimental setup
3. Track Reconstruction
4. Hits Clustering
5. AI-assisted Model Description
6. AI-assisted Model Evaluation On Simulation
7. AI-assisted Model Evaluation On Data
8. Summary and outlook

ALERT Physics Program

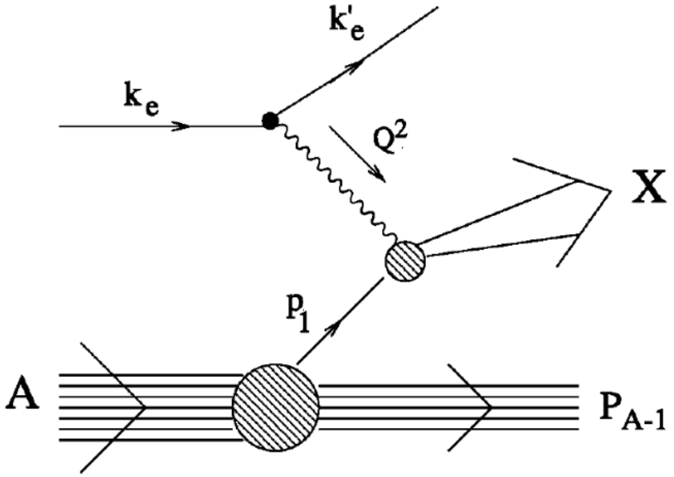
Comprehensive studies QCD in nuclei and associated medium modifications

Coherent DVCS on ^{4}He



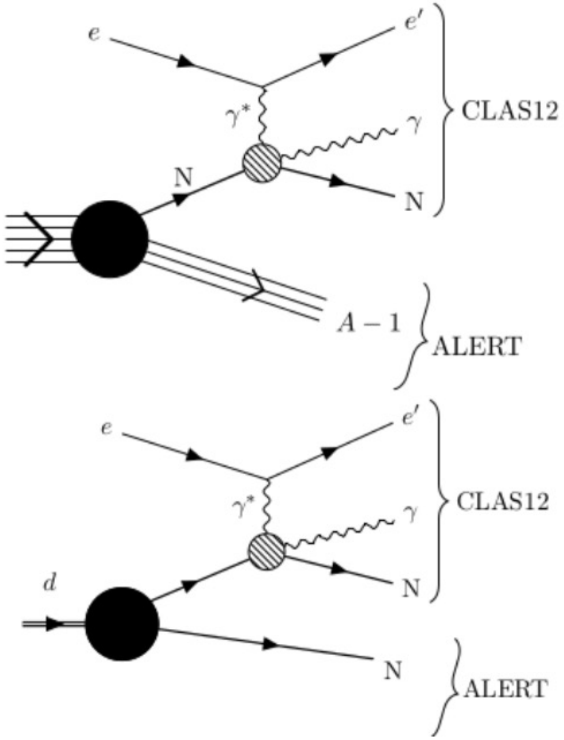
Explore the partonic structure of ^{4}He with Generalized Parton Distributions

DIS on ^{4}He and ^{2}H



Test the Final State Interactions and rescaling model

Incoherent DVCS on ^{4}He (^{2}H)

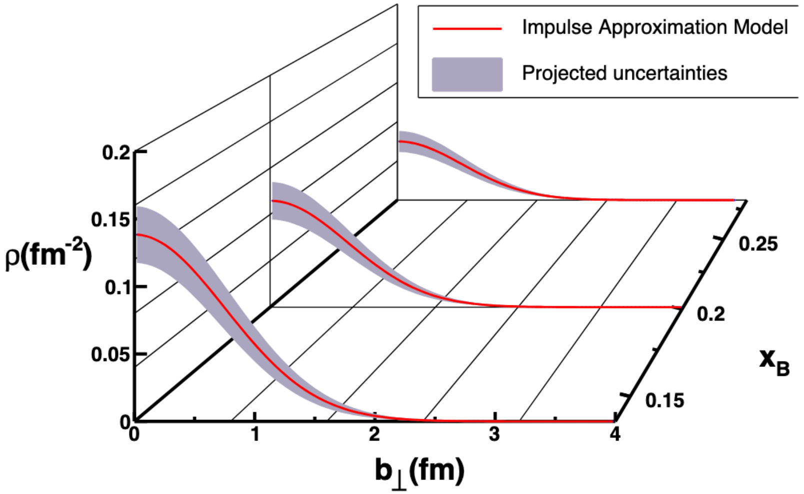


Explore the 3-D structure of modified nucleons in light ions ^{2}H,^{3}H, and ^{3}He (^{1}H or neutron)

ALERT Physics Program

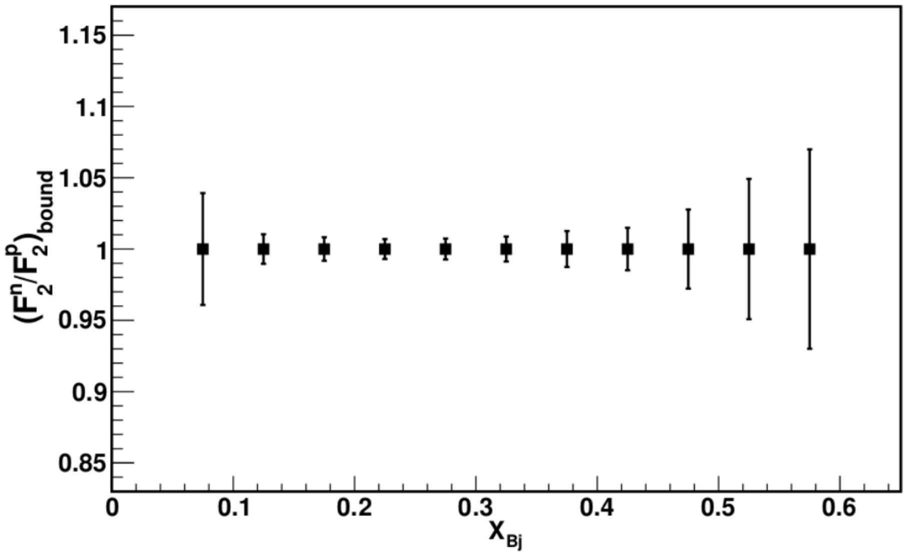
Comprehensive studies QCD in nuclei and associated medium modifications

Coherent Process on ^4He



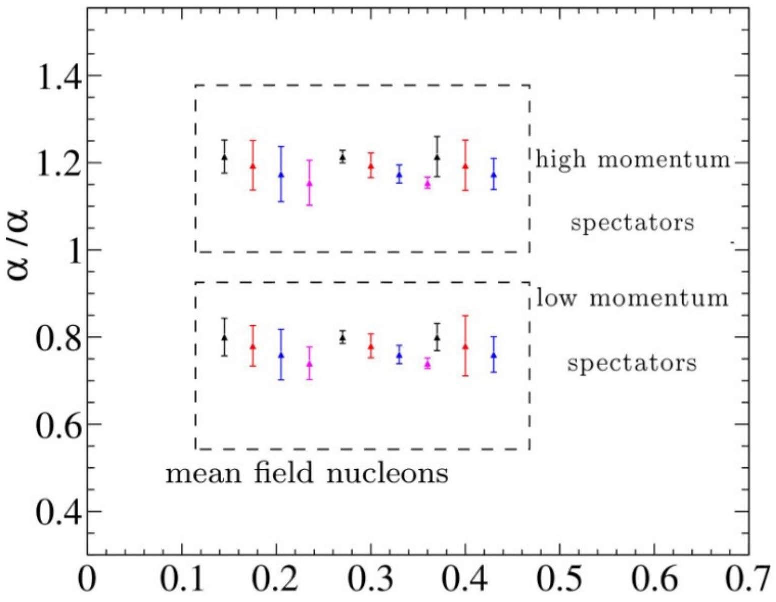
Extract quark and gluon GPDs in a dense nucleus; in this case, GPD H is obtained for both quarks and gluons in ^4He

DIS on ^4He and ^2H: Tagged EMC effect



Measure the F_2 structure function of a weakly bound nucleon in ^4He and compare it to the ^2H case. Control FSIs with tagged fragments

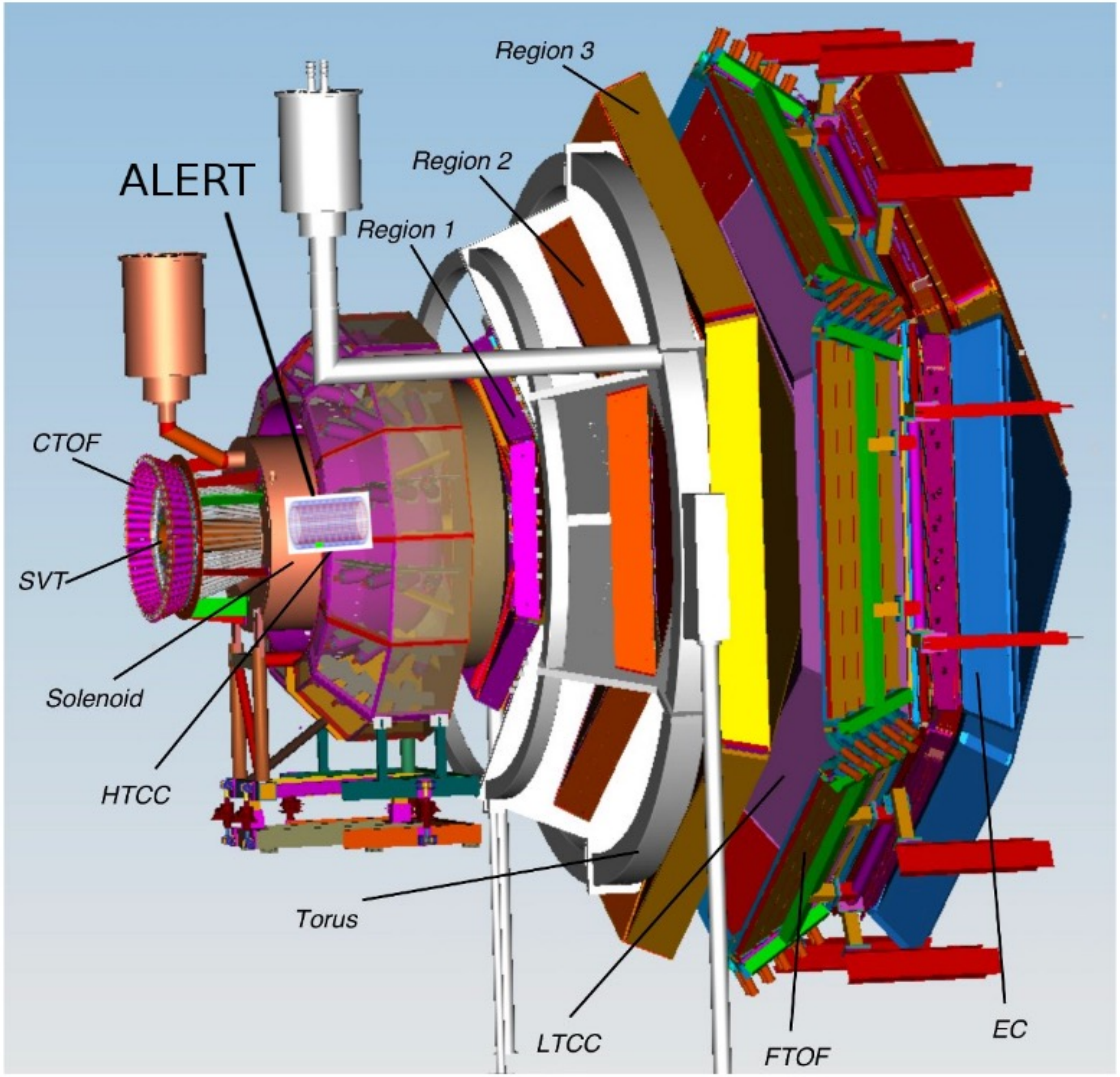
Incoherent Process on ^4He and ^2H



Extract quarks GPDs for bound nucleons and thus understand the effect of FSIs

ALERT Experiment Setup

- The ALERT experiment will take place in Hall B at Jefferson Lab:
  - CLAS12: detect scattered electrons and forward-scattered hadrons
  - ALERT: detect recoil spectators or coherently scattered nuclei
- ALERT Goals:
  - Aim to identify light ions: p, ^{2}H, ^{3}H, ^{3}He, and ^{4}He
  - Detect the lowest momentum possible, down to 70 MeV/c for proton
  - Handle high CLAS12 rates and luminosities (10^{35} cm^{2}s^{-1})



ALERT Experiment Setup

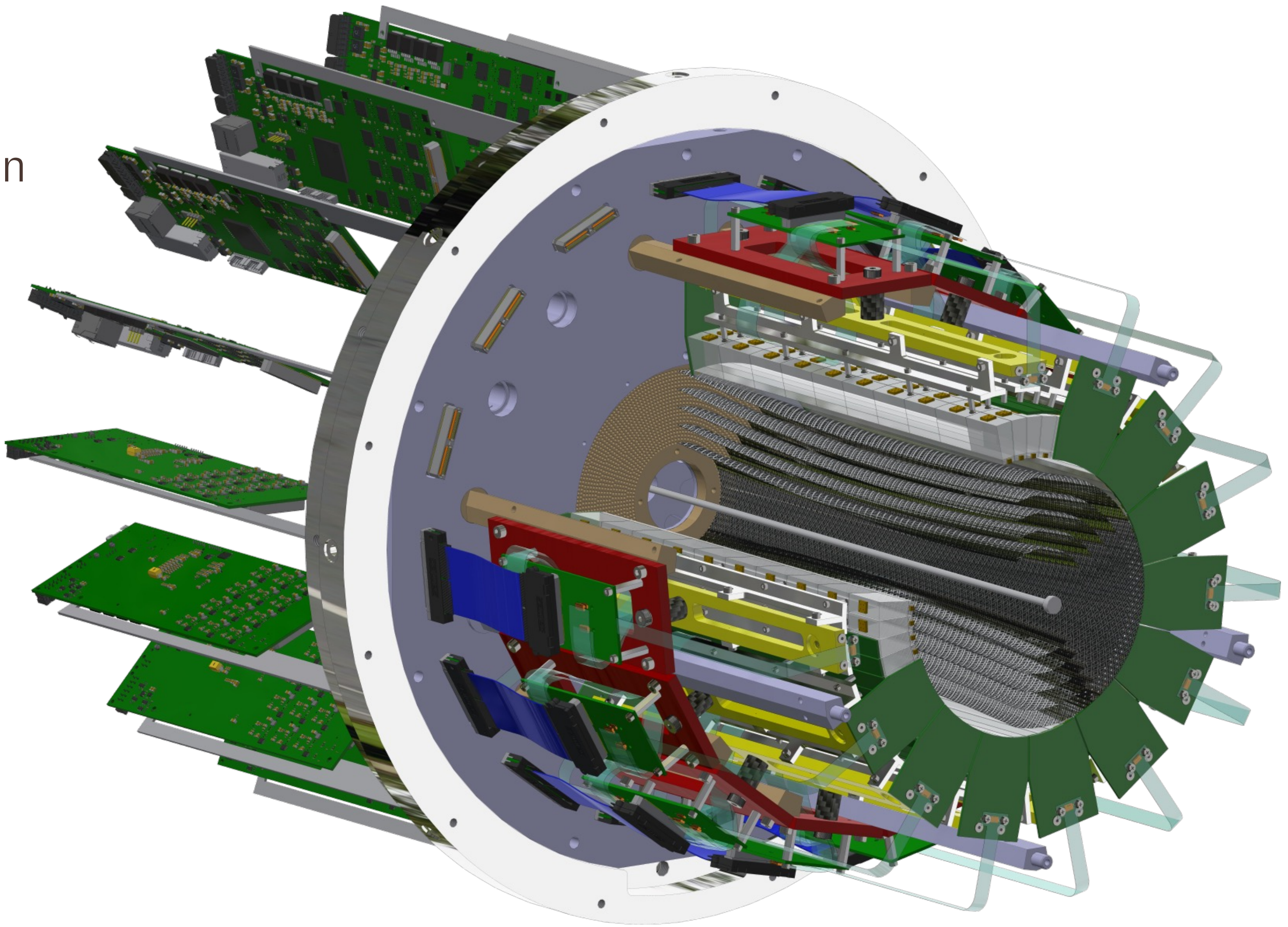
- ALERT have two sub-detectors:
  - A Hyperbolic Drift Chamber (AHDC) and A Time of Flight (ATOF)

ATOF

- Time of flight: used for Particle IDentification
- Small barrel of segmented scintillators
- The TOF measurement is degenerate for ^{2}H and ^{4}He, but dE/dx can distinguish the two nuclei bands

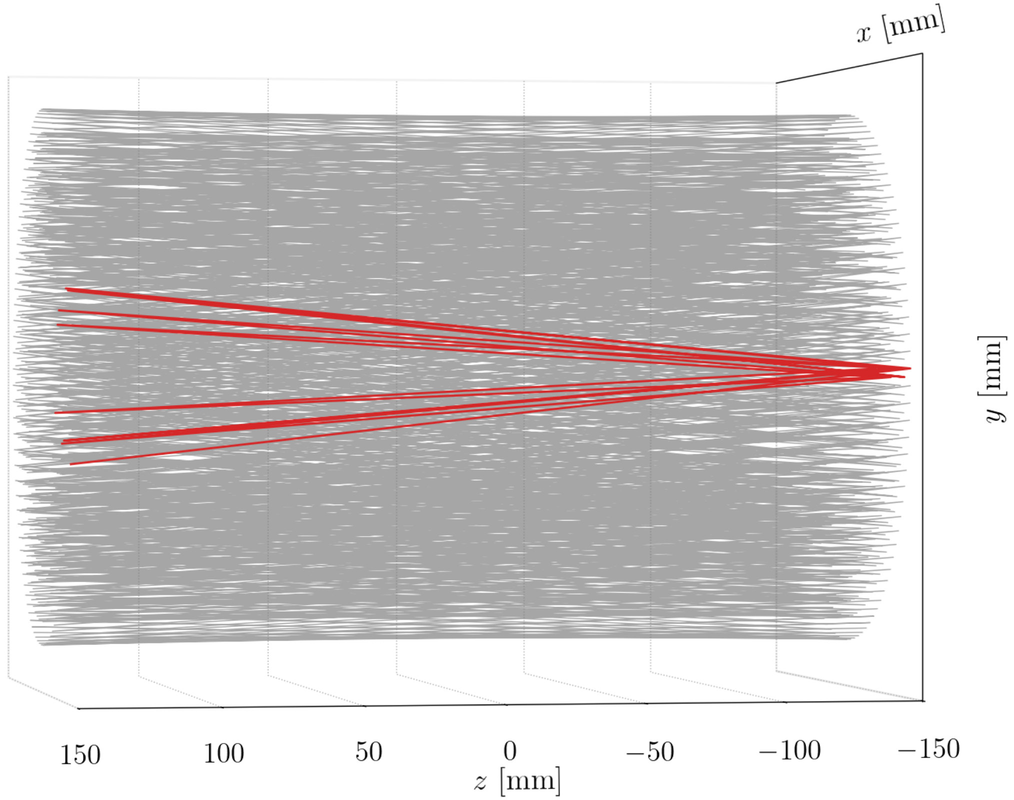
AHDC

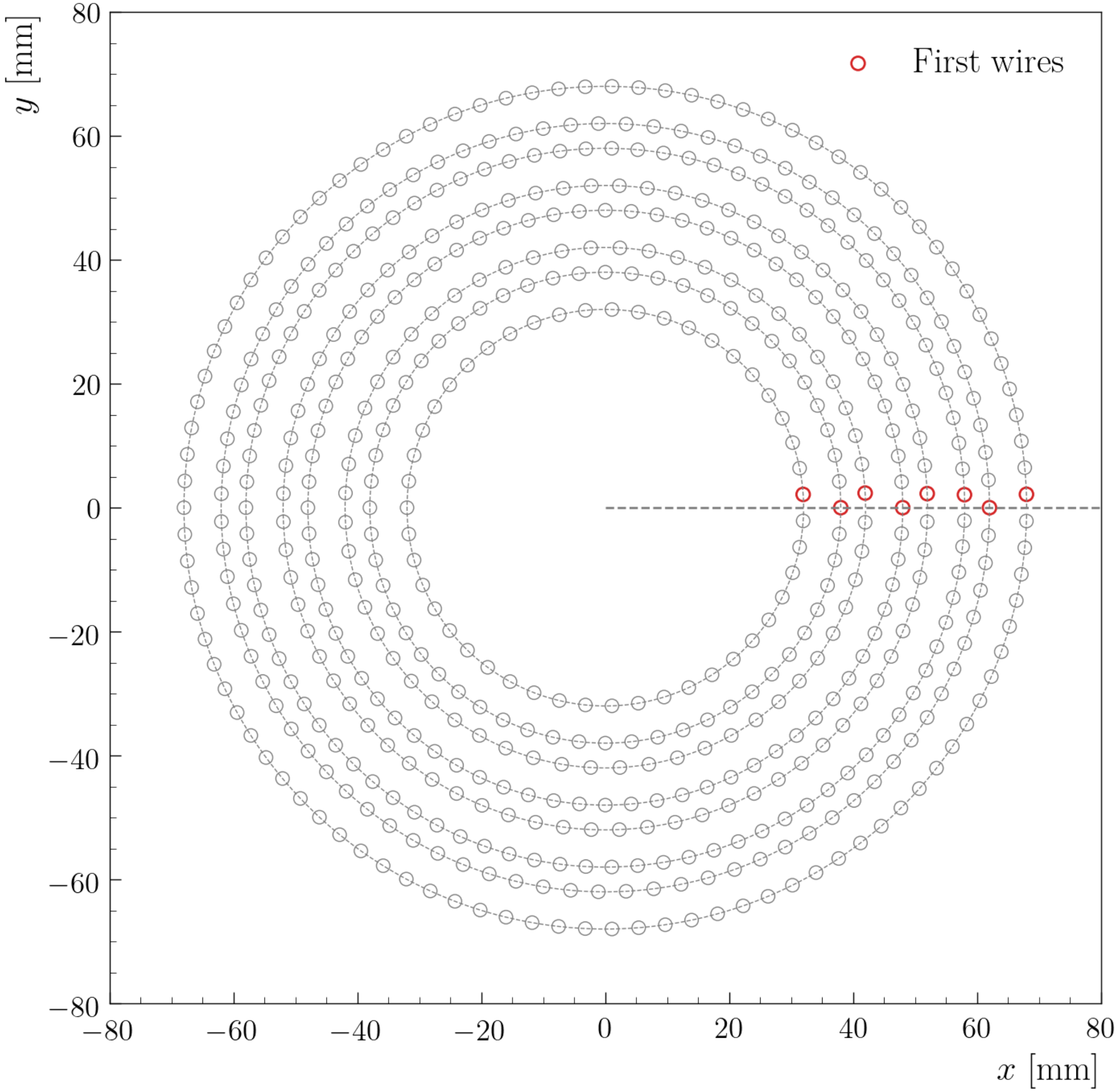
- Aluminum wire: 2 mm apart
- 20-degree stereo angle (hyperbolic shape)
- 5 superlayers, each composed of 2 layers
- 576 signal wires (6 ground wires of each signal)



AHDC Geometry

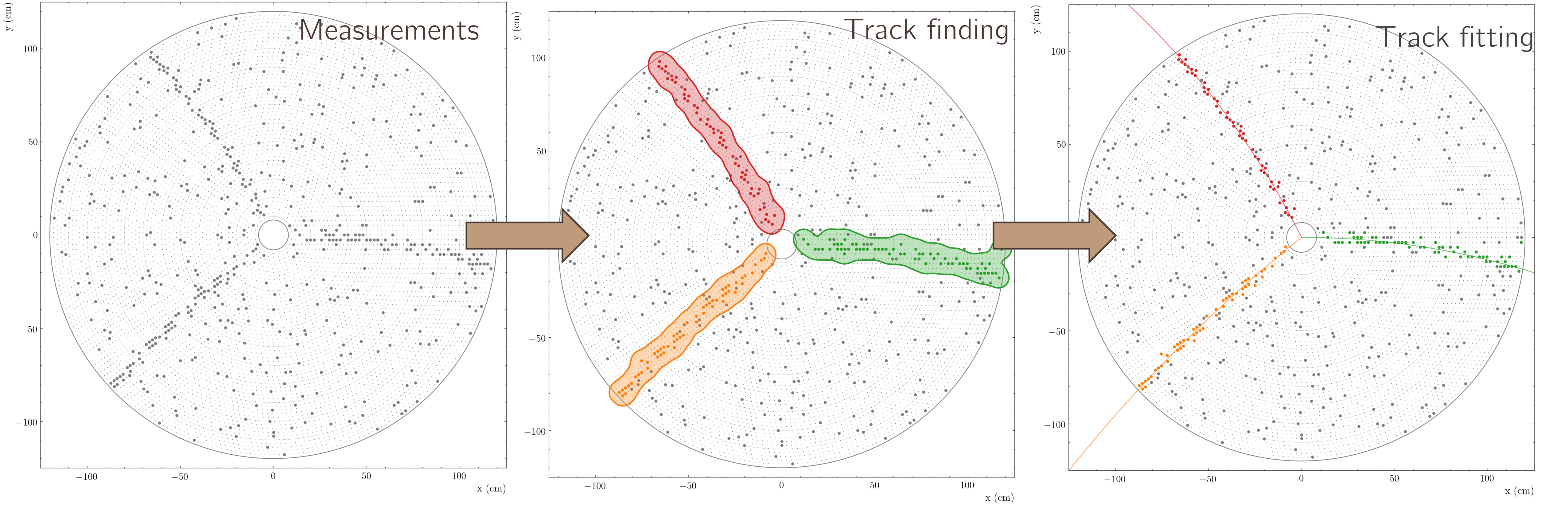
- The AHDC is composed of 576 wires
- The wires in the first layer in superlayer 2, 3, and 4 are shifted by half a cell
- The distance between wires is constant





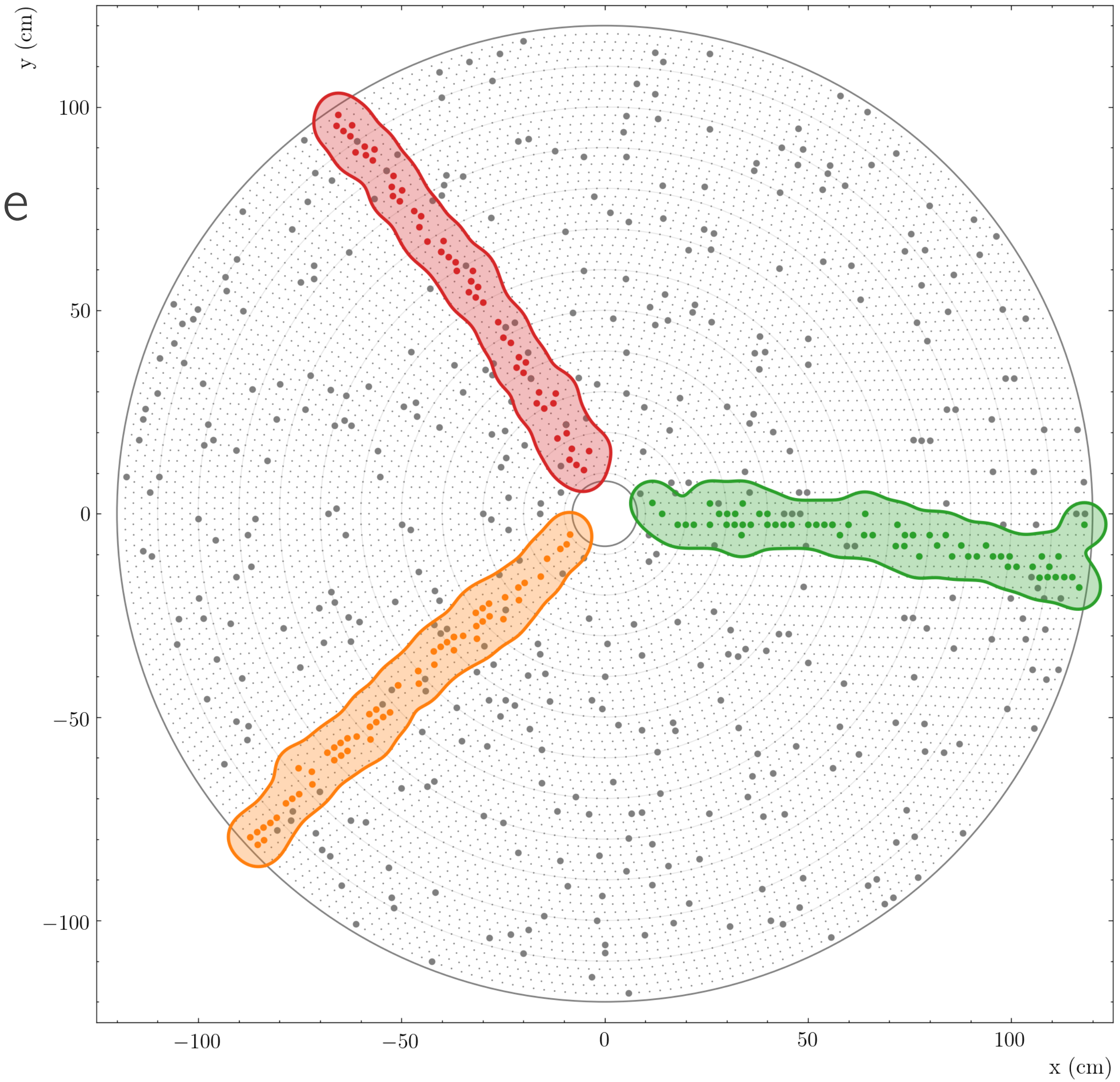
Track Reconstruction

- Aim to reconstruct the momentum and trajectory of charged particles
- Two stages of track reconstruction:
  - Track finding: Identifying which hits came from the same charged particle
  - Track fitting: Fitting hits to a single track to extract track parameters (momentum and position)



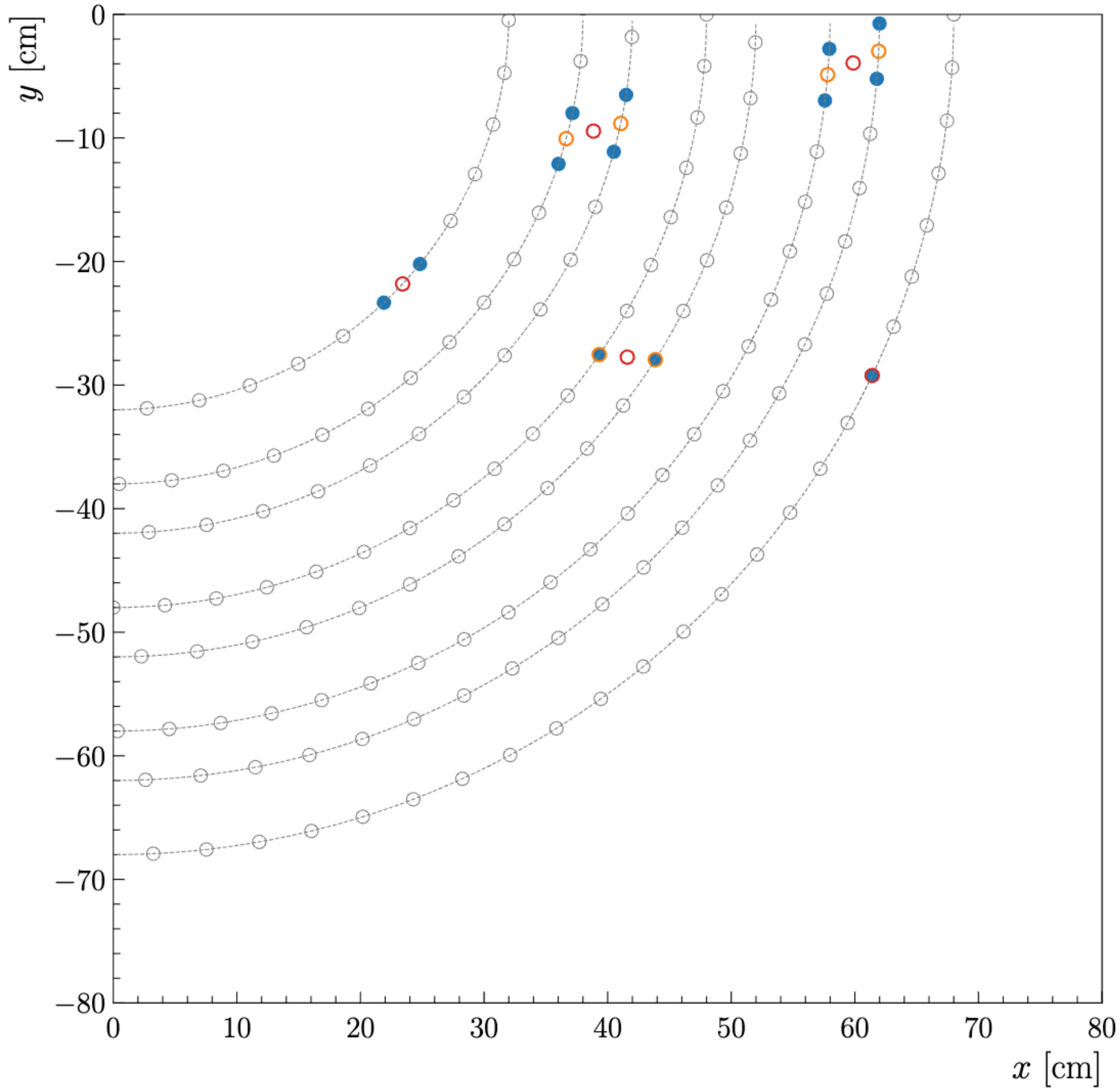
Track Finding

- Track finding is a clustering problem:
  - Set of points (hits) ⇒ cluster in sets (tracks) originating from the same particle
  - Hits: particles deposit energy when interacting with the detector material
  - Tracks: reconstructed sequences of hits representing charged particle trajectories
- Different algorithms:
  - Distance between hits + fit
  - Hough transform
  - Combinatorial Kalman Filter
  - Artificial Intelligence models (MLP, GNN...)



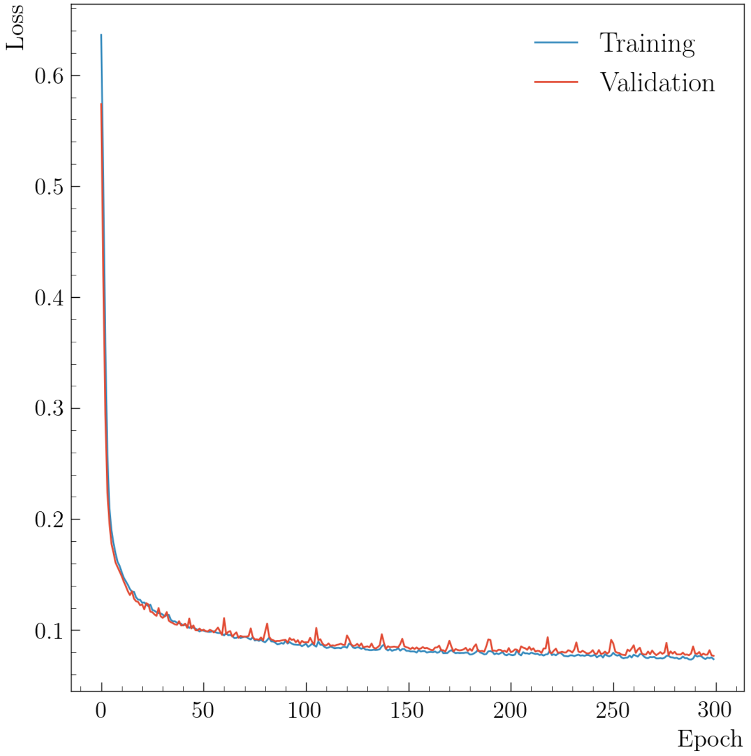
Clustering and Track Candidates

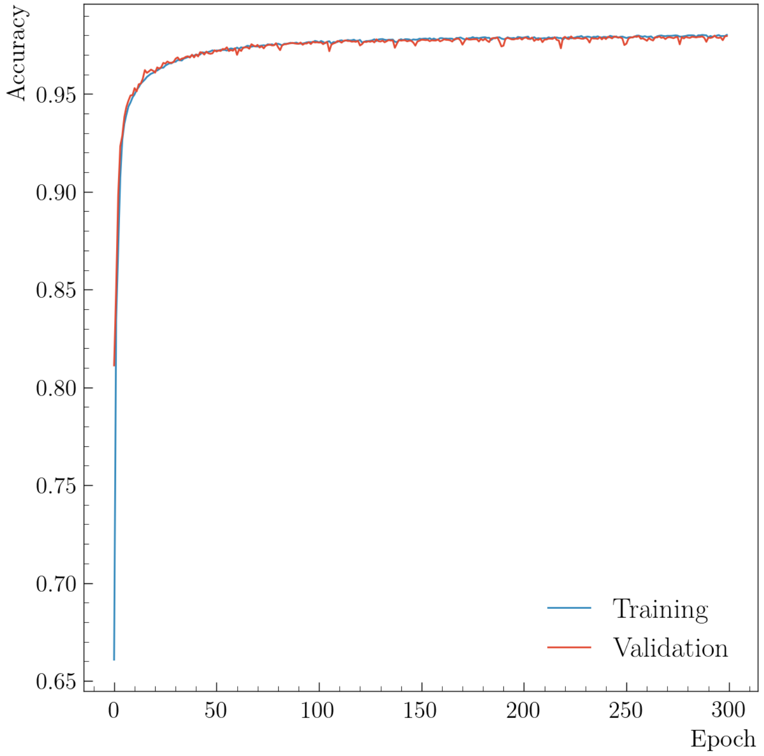
- First step, find all track candidates:
  - Clustering ⇒ merging hits close in the *x*-*y* plane to reduce the combinatorial background
  - Merge hits on the same layer that are one wire apart into intra-cluster
  - Merge intra-clusters in the same superlayer that are less than 8mm apart into inter-cluster
  - Generate all track candidates with 5 inter-clusters (one on each superlayer)
- Generate around 40/100k track candidates per event with inter-clusters/raw hits

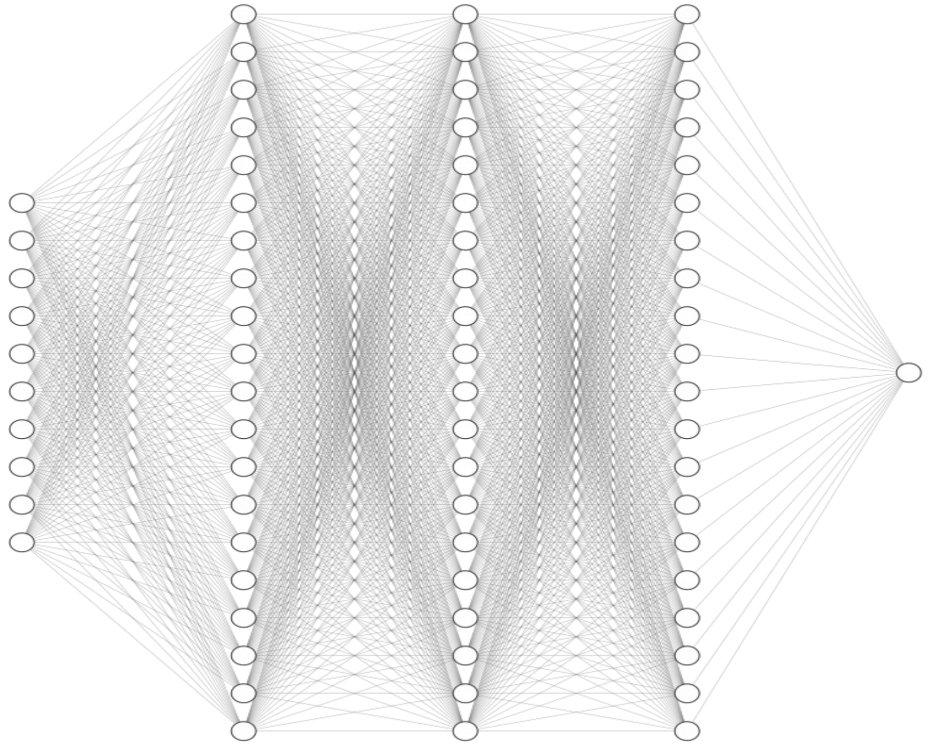


AI-assisted model: Description and training

- Model: MultiLayer Perceptron, 10 inputs, 1/3/5 hidden layer (15/20/100 neurons), 1 output
- Inputs: x and y values of the five inter-clusters
- For the training ⇒ Need good and bad tracks:
  - Good tracks: GEANT4 simulation (particle with p ∈ [0.07, 1.5] GeV/c, φ ∈ [0, 360]°, θ ∈ [30, 150]° and Vz ∈ [-15, 15] cm)
  - False tracks: Interchanging randomly up to two inter-clusters with another event
  - Generate 5M events composed of all light nuclei (flat distribution for all variables)
  - Output: Number between 0 and 1, with 0/1 means bad/good track



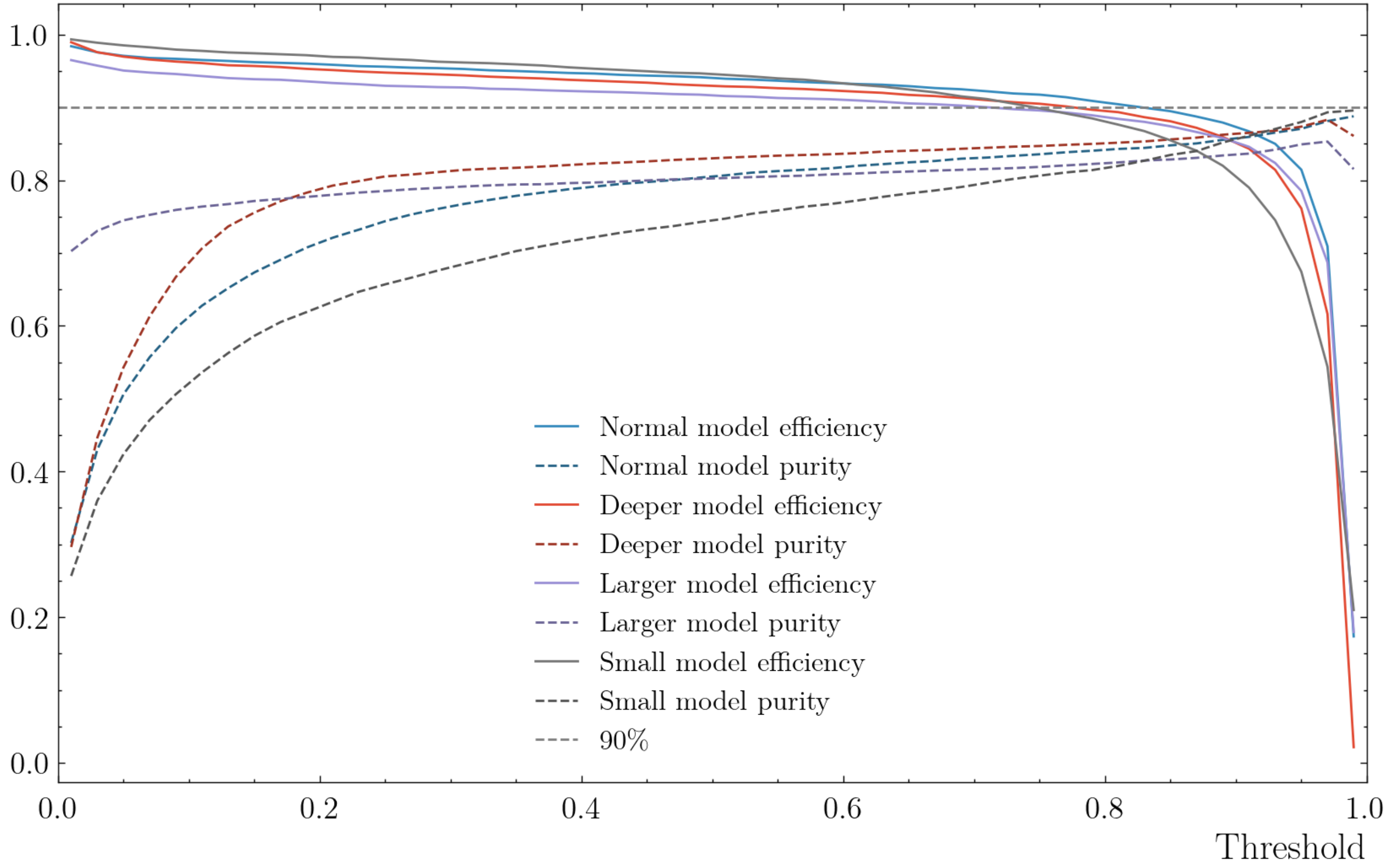




Evaluation on simulation: Efficiency and Purity vs. Threshold

- Threshold: if output above/lower than the threshold ⇒ good/bad tracks
- To evaluate the model:
  - Efficiency: Number of good tracks classified as good normalized by the number of events.
  - Purity: Number of good tracks classified as good normalized by the number of tracks (good or bad) classified as good.
- Events need to have at least one track candidate.
- Set the threshold to 0.2 to have a higher efficiency

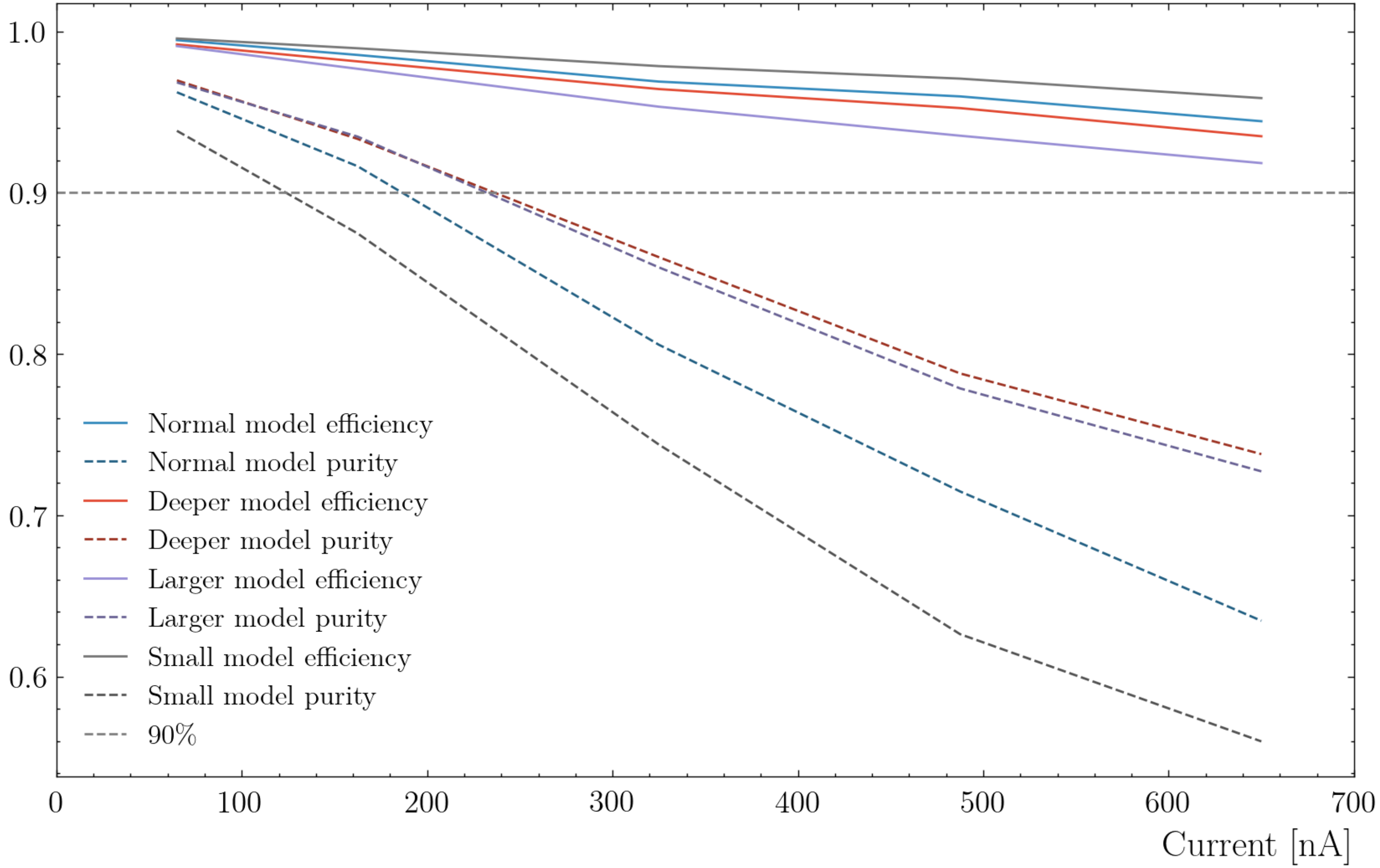
- Blue: model with 20 neurons in 3 hidden layers
- Violet: model with 100 neurons in 3 hidden layers
- Red: model with 20 neurons in 5 hidden layers
- Gray: model with 15 neurons in 1 hidden layer



Evaluation on simulation: Efficiency and Purity vs. Current

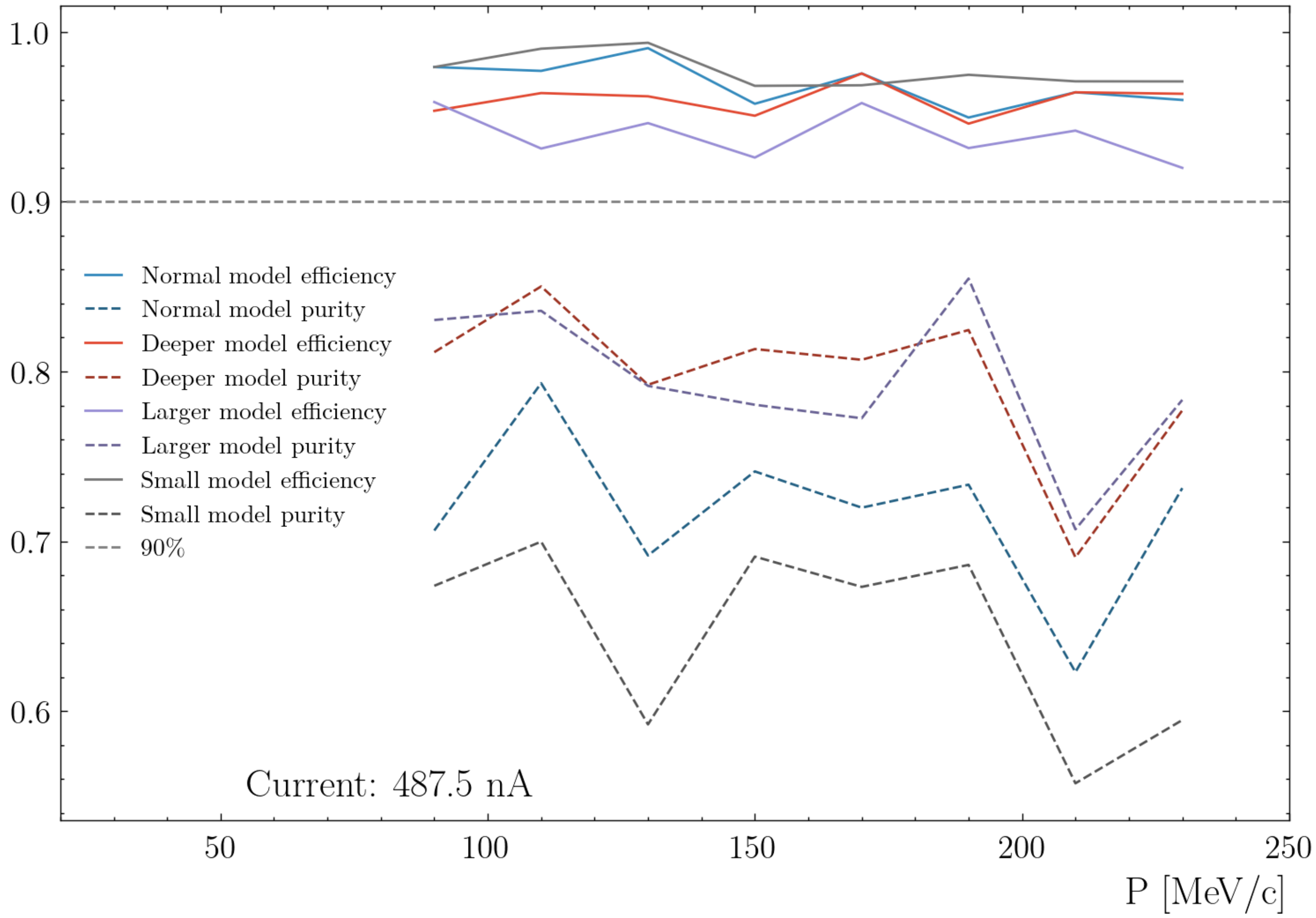
- Efficiency is always higher than 90% and the purity is between 55% and 95%
- More current means more background

- Blue: model with 20 neurons in 3 hidden layers
- Violet: model with 100 neurons in 3 hidden layers
- Red: model with 20 neurons in 5 hidden layers
- Gray: model with 15 neurons in 1 hidden layer



Evaluation on simulation: Efficiency and Purity vs. Momentum

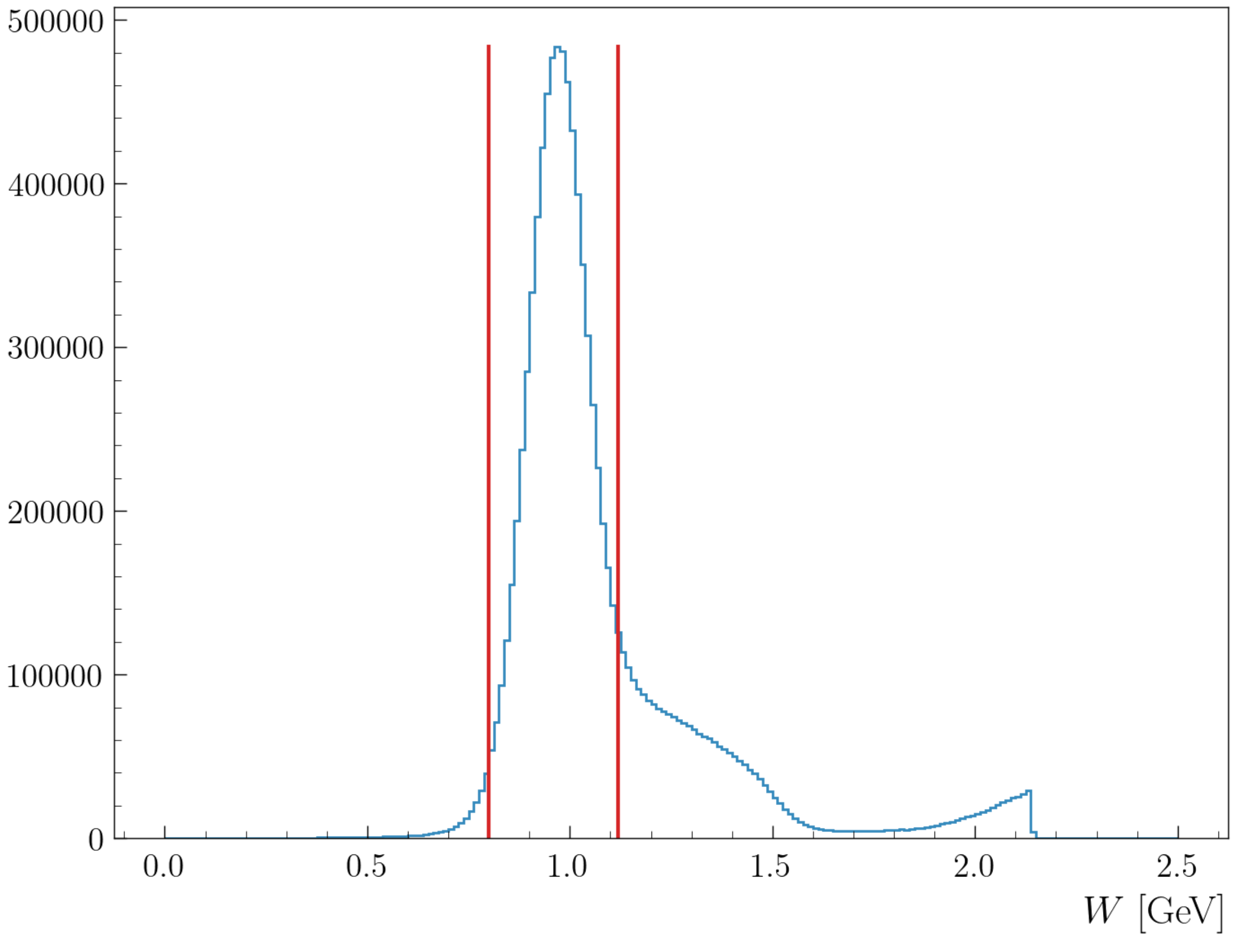
- Background is generated with current I = 487.5 nA
- Constant efficiency and purity across the momentum range

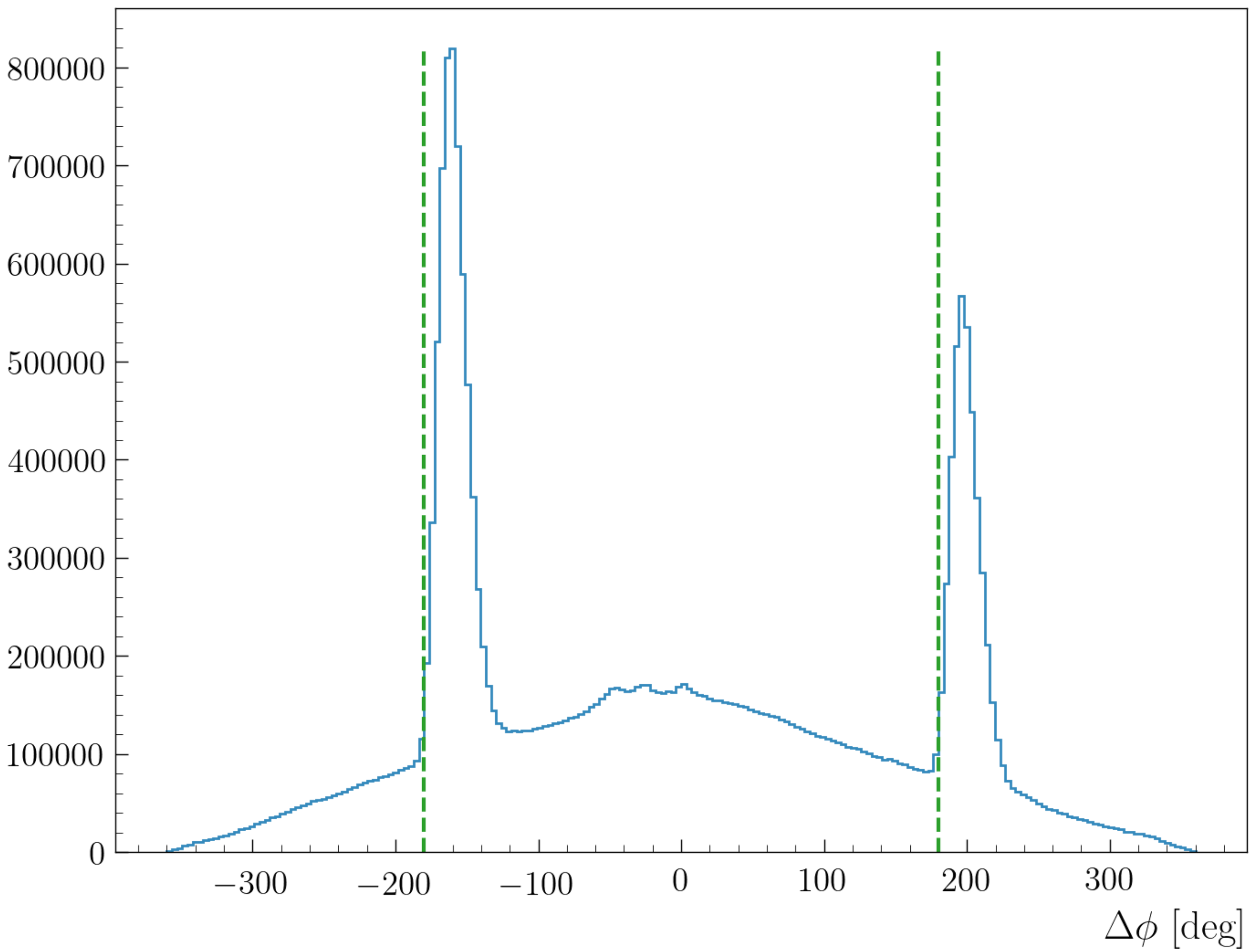


- Blue: model with 20 neurons in 3 hidden layers
- Violet: model with 100 neurons in 3 hidden layers
- Red: model with 20 neurons in 5 hidden layers
- Gray: model with 15 neurons in 1 hidden layer

Evaluation On Data: elastic scattering

- To evaluate the performance of the AI, use elastic scattering
- For the AHDC, want the low momentum proton ⇒ use electron at low θ
  Compute Δϕ using electron and AHDC hits ⇒ shift in Δϕ approx. 20°





Test on real data: AI efficiency and purity

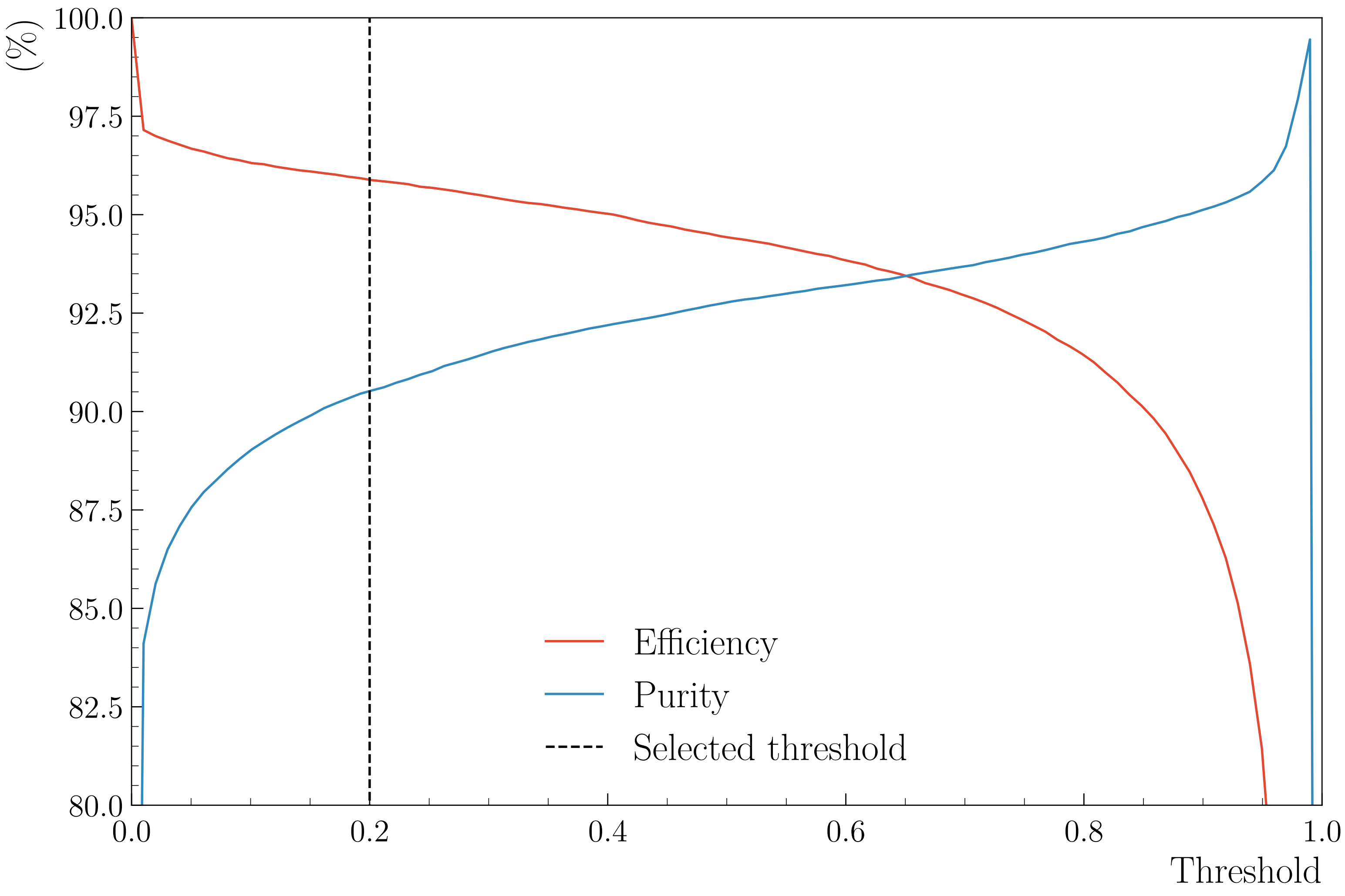
- Efficiency and purity as a function of the threshold:

- Threshold: if output above/lower than the threshold ⇒ good/bad tracks

- Efficiency: Number of good tracks classified as good normalized by the number of events.

- Purity: Number of good tracks classified as good normalized by the number of tracks (good or bad) classified as good.

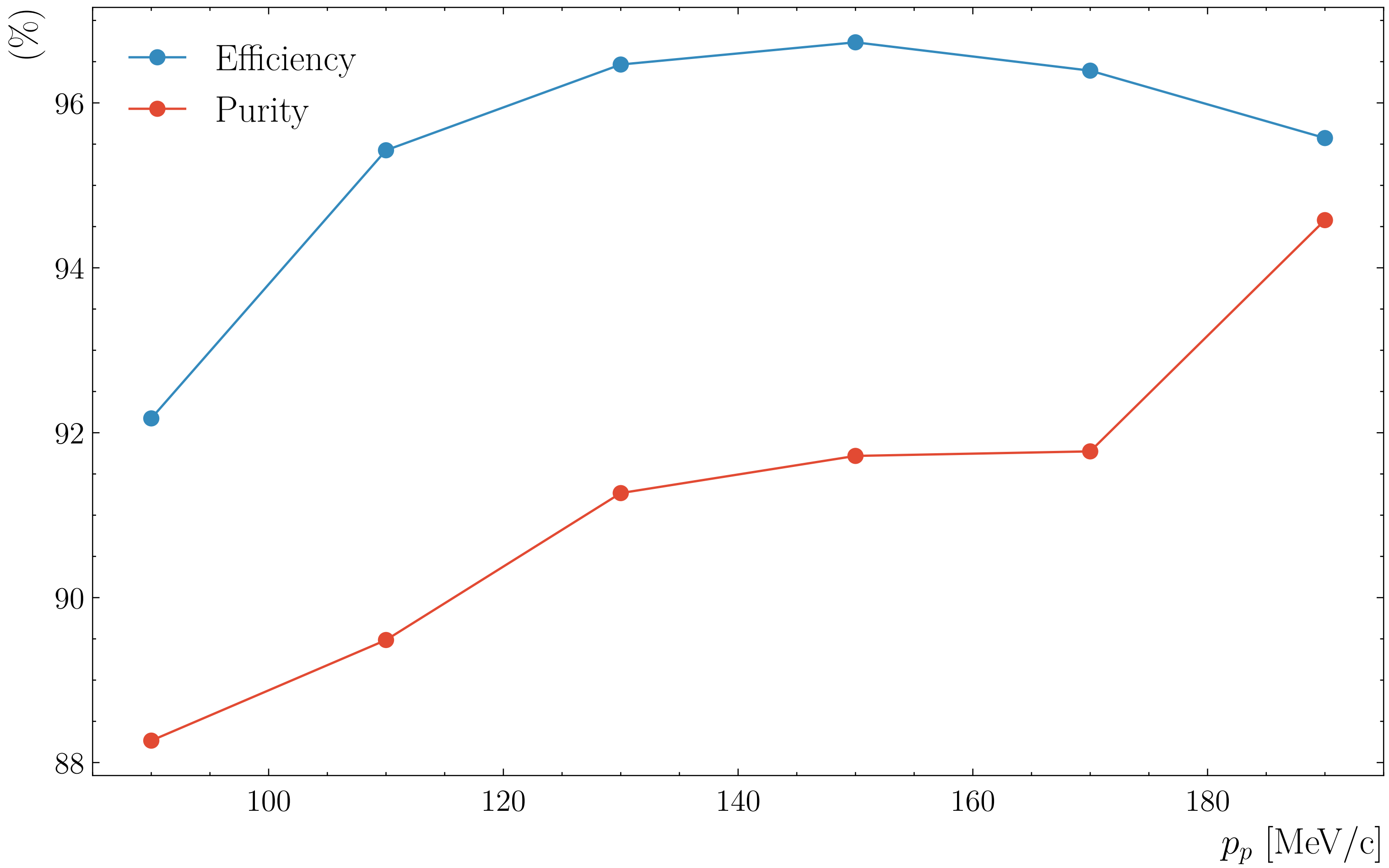
- We have an efficiency of 96%, and a purity of 90% at 0.2



Test on real data: AI efficiency and purity

- Efficiency and purity as a function of the proton momentum:
- Momentum is computed with θ_e

- Both the efficiency and purity decrease when the momentum decrease



Summary and Outlook

- ALERT physics program:
  - Tagged processes will provide insight into the origin of the EMC effect
  - Tagged DIS measurements will help differentiate between models
  - Tagged DVCS will bridge the gap between partonic and nucleonic interpretations of the EMC ratio

- We have developed an AI-assisted MLP for track finding:
  - Evaluated efficiency and purity on simulation as a function of momentum, threshold, and current
  - Evaluated efficiency and purity on elastic data as a function of momentum, threshold
  - Efficiency is always higher than 90%
- Future work:
  - Check the efficiency and purity of ^{4}He using elastic scattering
  - Improve the model with other information as input (energy deposited and angle between hits)
  - Match hit in the ATOF with track in the AHDC using AI

Thanks

This work is supported in part by the U.S. DOE award #: DE-FG02-07ER41528