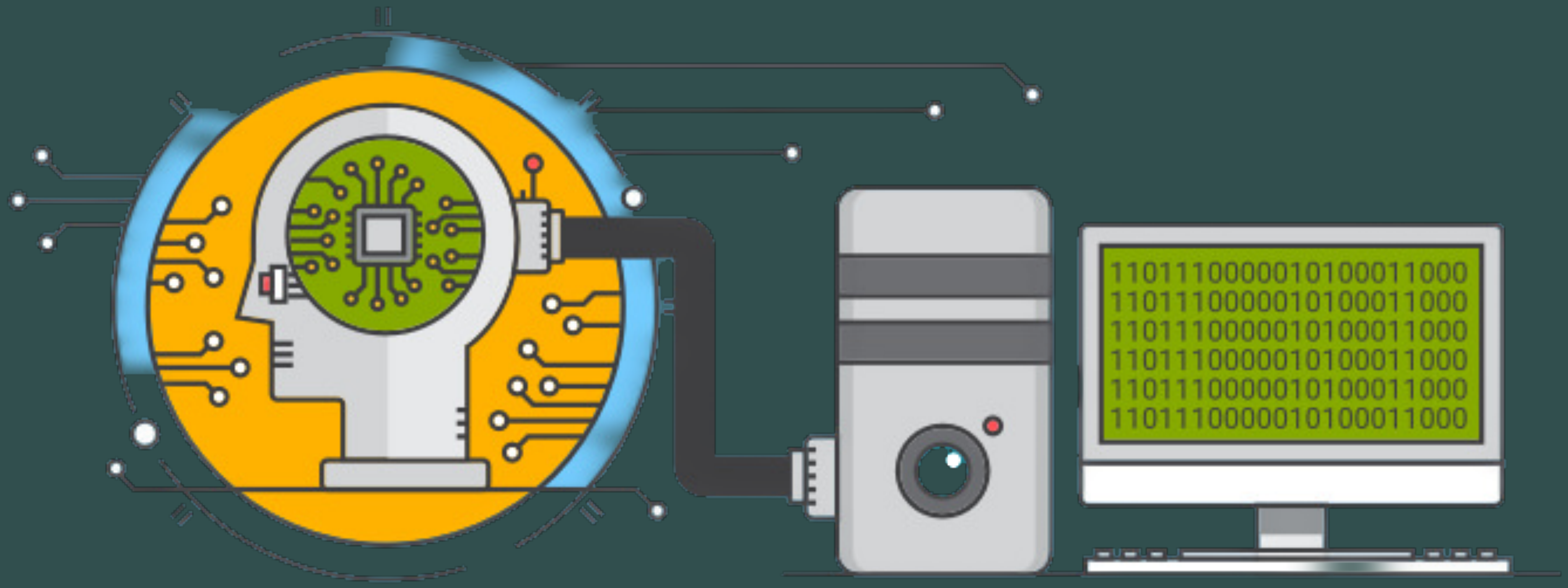
Real-Time Track Reconstruction with AI

CLAS12 Collaboration

G.Gavalian (Jefferson Lab)

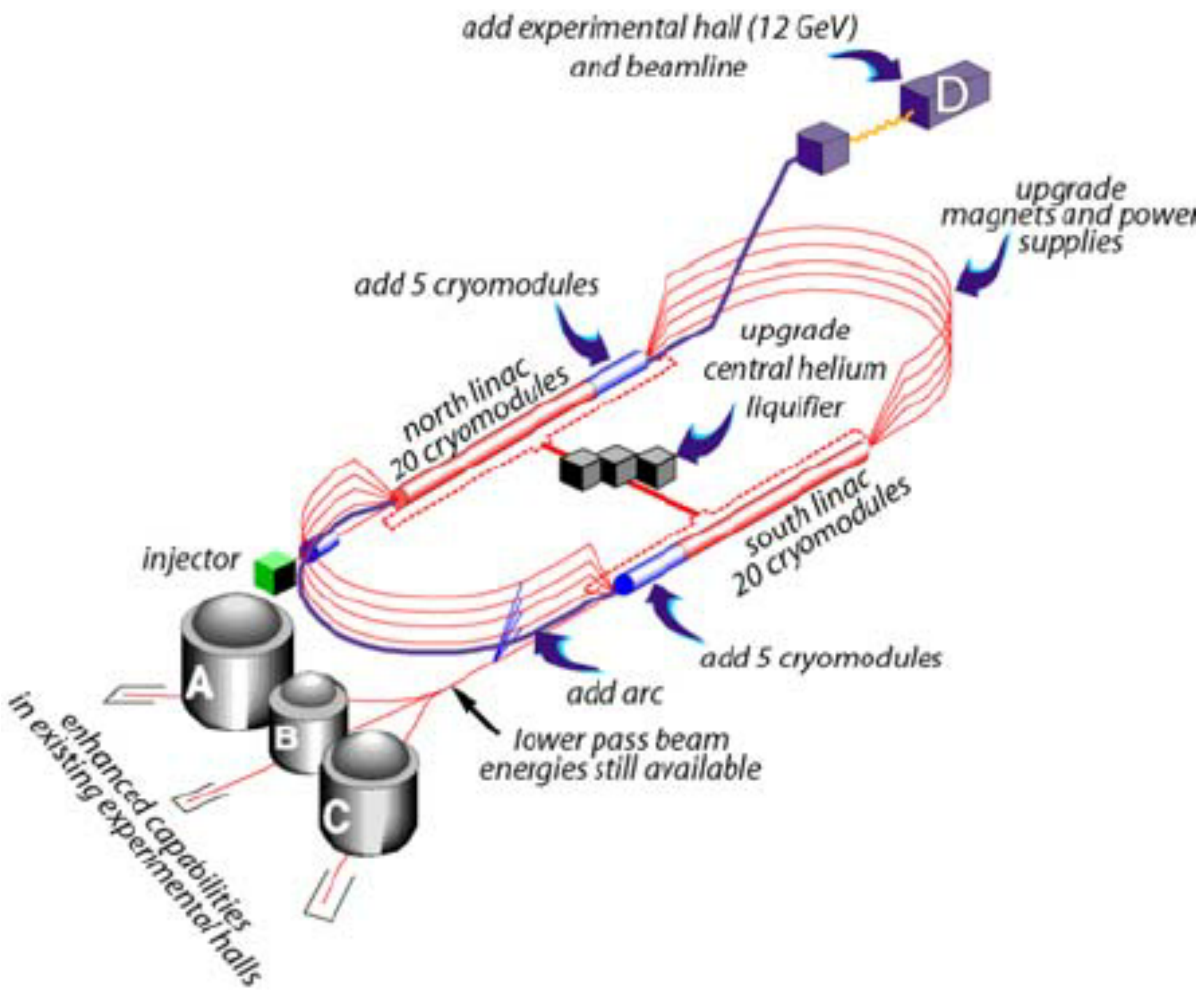


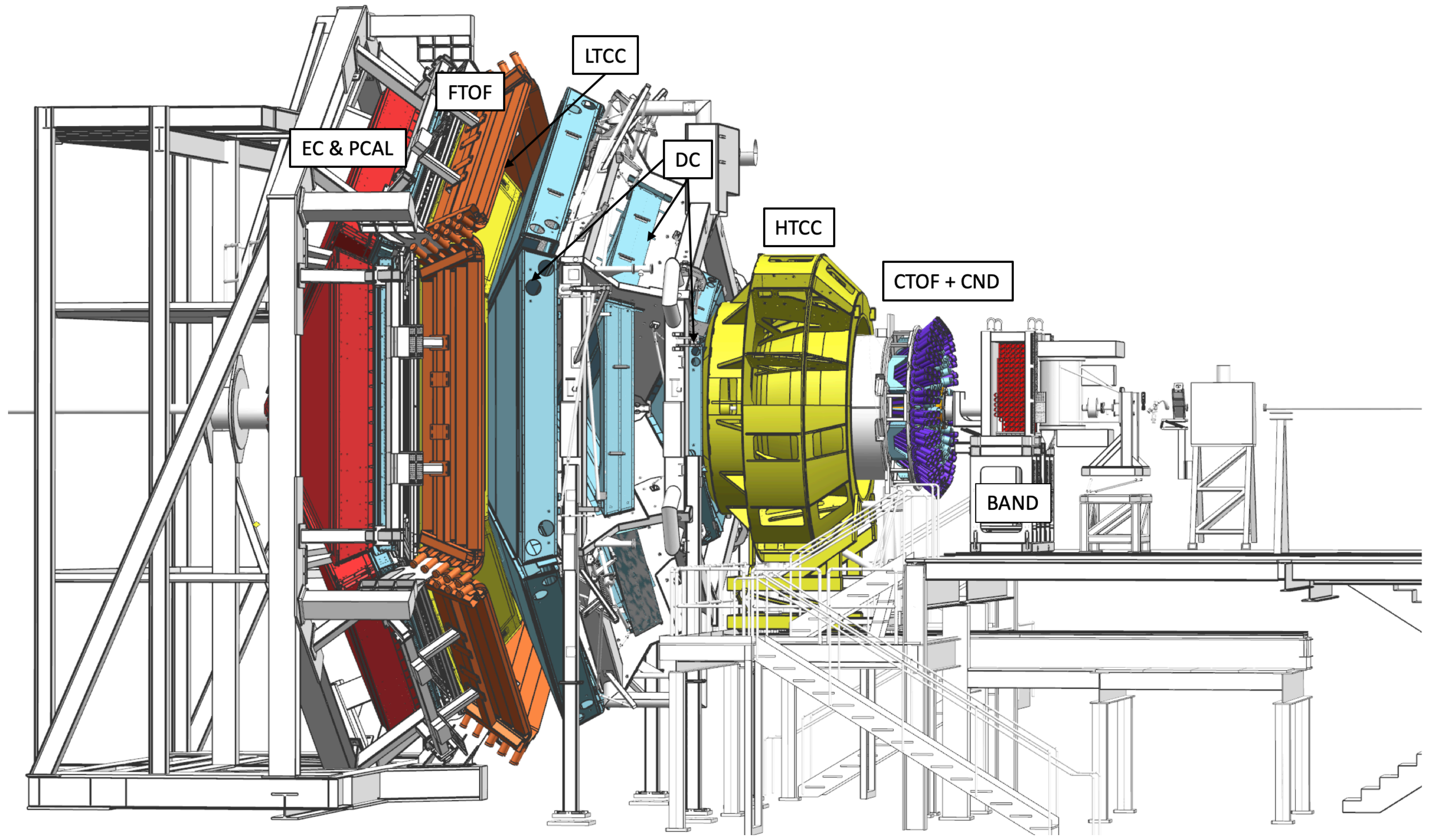
CHEP 2024

Outline

- CLAS12 setup
- Track identification using AI
- Track Parameter Estimation using AI
- Looking into the future

Jefferson Lab (Hall-B/CLAS12 Detector)





- **CEBAF**
  - 12 GeV electron beam distributed to 4 experimental hall
  - Each experimental hall contains a detector system for specific experiments
- **Hall-B:**
  - CEBAF Large Acceptance Spectrometer (CLAS12) Located in Hall-B
  - **Central Detector:**
    - Silicon Tracker
    - Time-Of-Flight
    - Neutron Detector
  - **Forward Detector:**
    - Drift Chambers
    - Time of Flight
    - High Threshold Cherenkov Counter
    - Ring Imaging Cherenkov Counter
    - Electromagnetic Calorimeter

Track Reconstruction (short history)

My First experience with Event Reconstruction

Rate: ~**0.0008** Hz (per person, assuming 20 min per event)

Earth Population: 4.767 billion

(2,135 kHz assuming 56% in the age bracket 21-65)

CLAS6 Detector (Jlab)

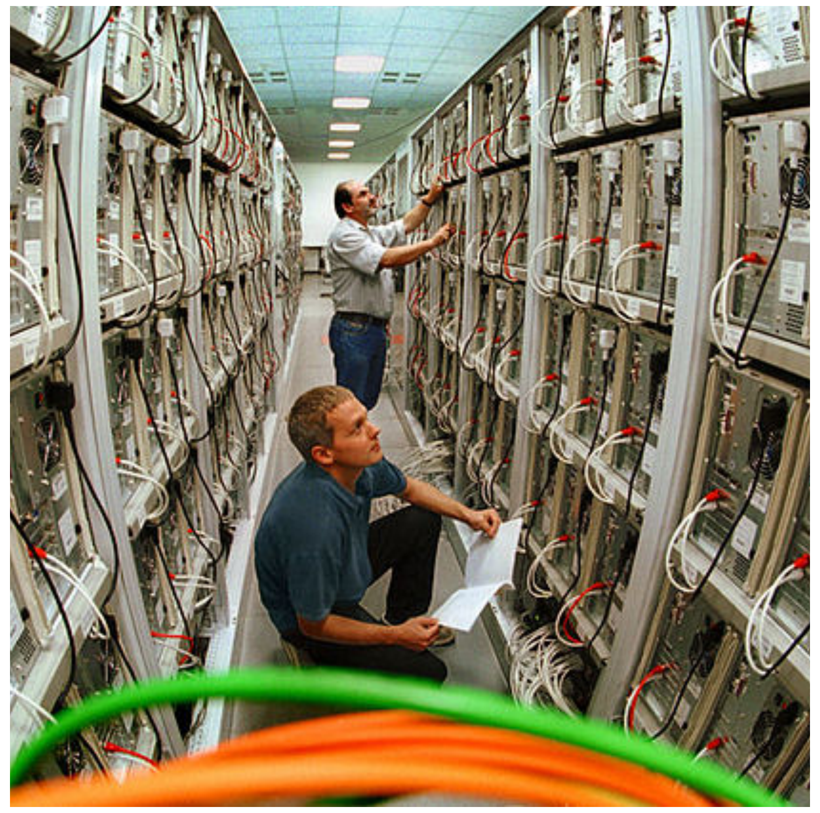
Rate: ~**2** Hz

Computers: 64 cores (2.6 GHz)

1983

2023

2003



CLAS6 Detector (Jlab)

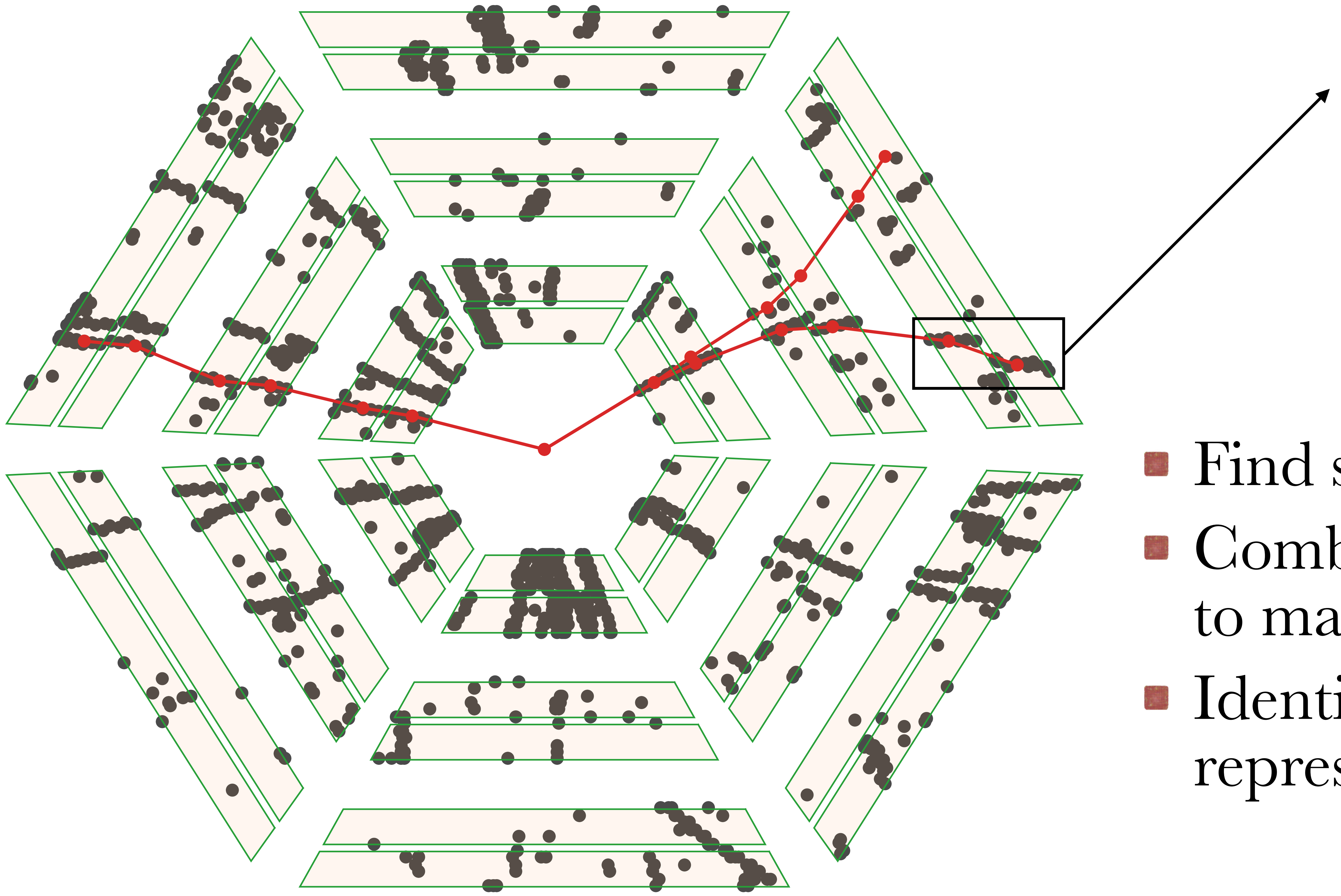
Rate: ~**4-8** Hz

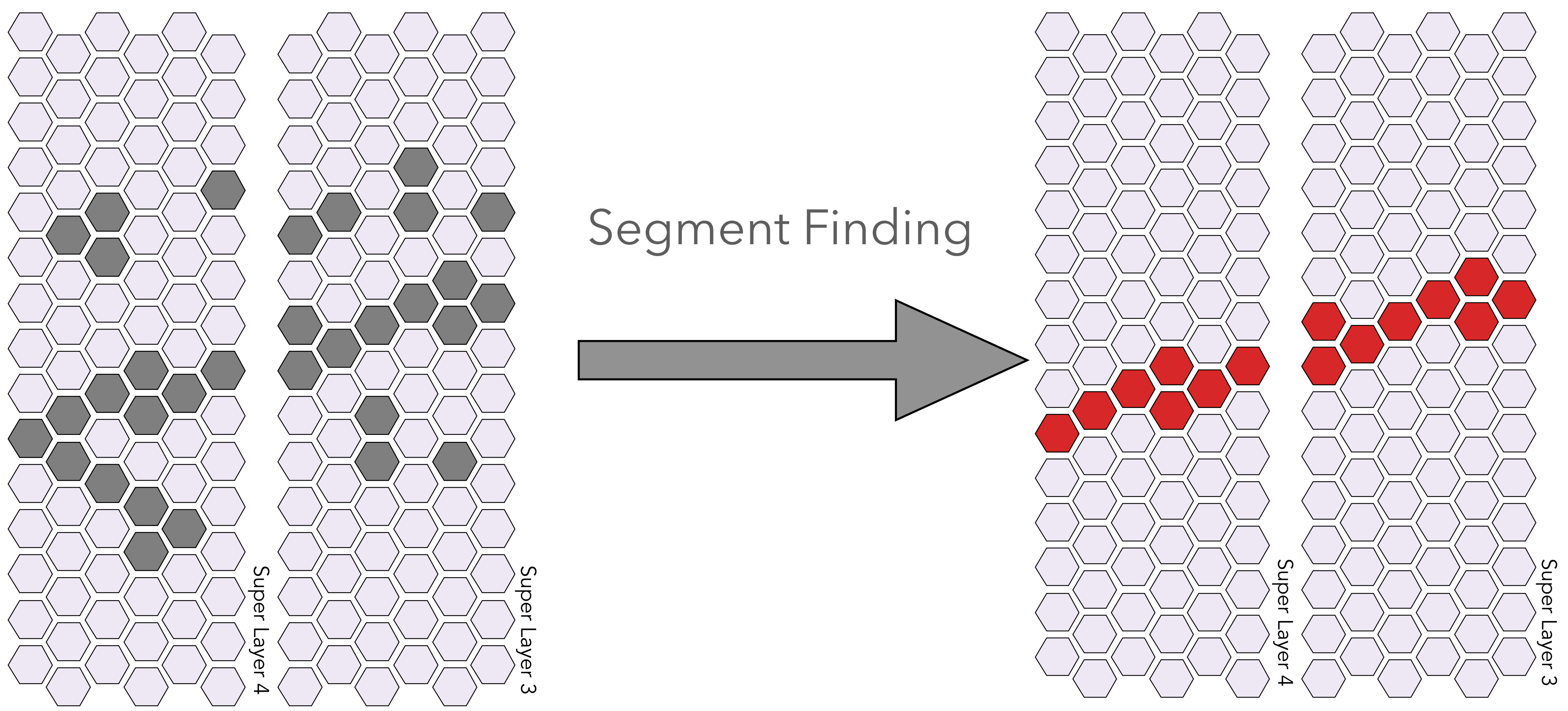
Computers: 8 core (2.4 GHz)

- Big jump in track reconstruction in first 20 years
- No change in the second 20 years
- What's in the future?

Track Finding Procedure

- 6 sectors with 6 chambers in each sector (called super-layers)
- 6 wire planes in each super layer with 6-degree tilt relative to each other, (112 wires in each plane)

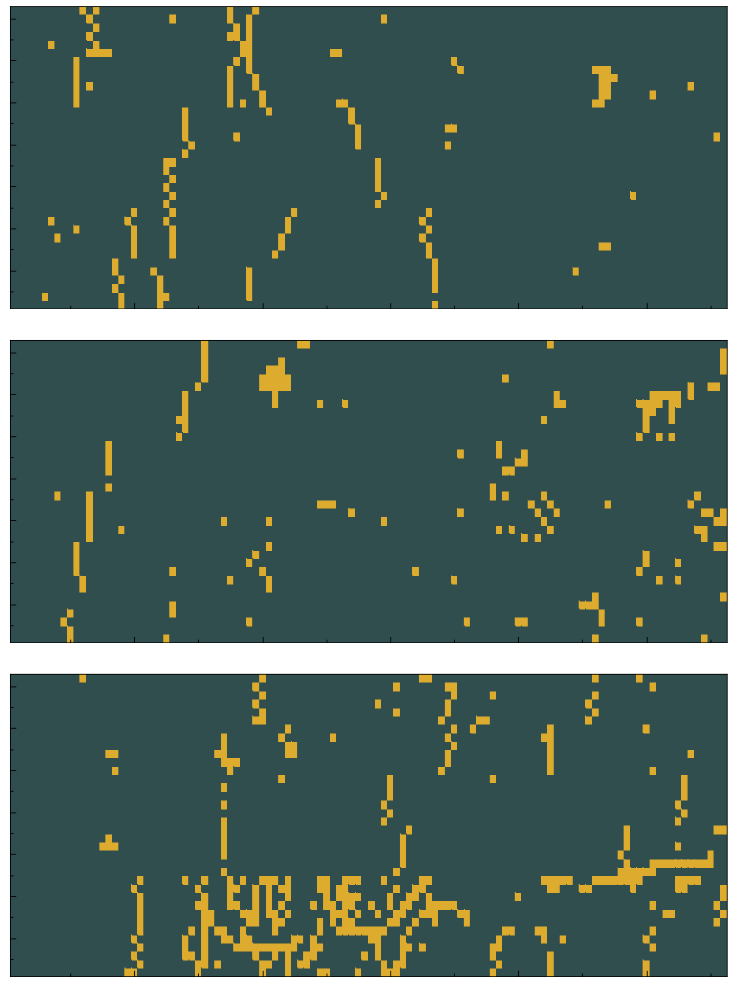




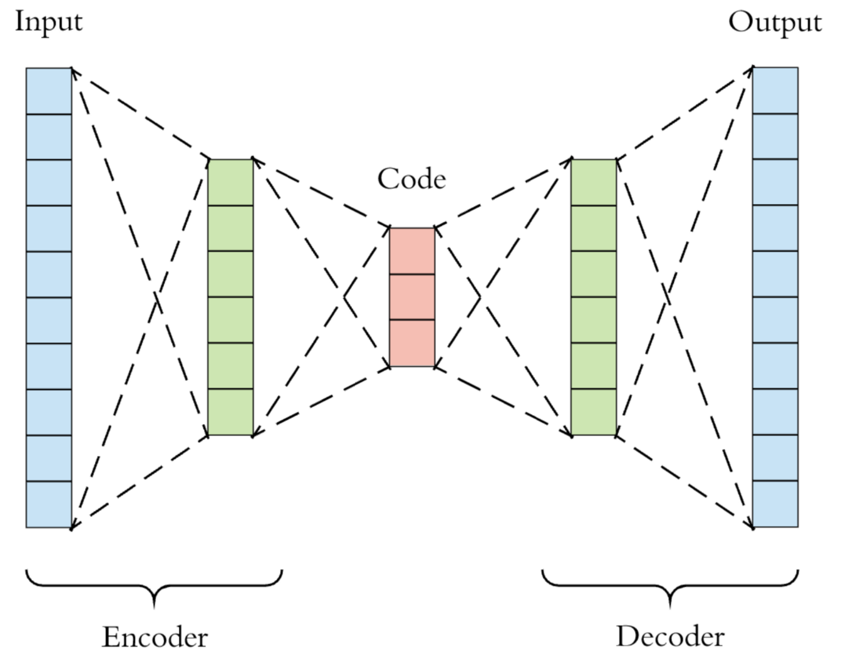
- Find segments in each super layer (remove noise)
- Combine 6 segments (one from each super layer) to make a list of possible tracks
- Identify correct combinations of segments that represent a track

What do we have now

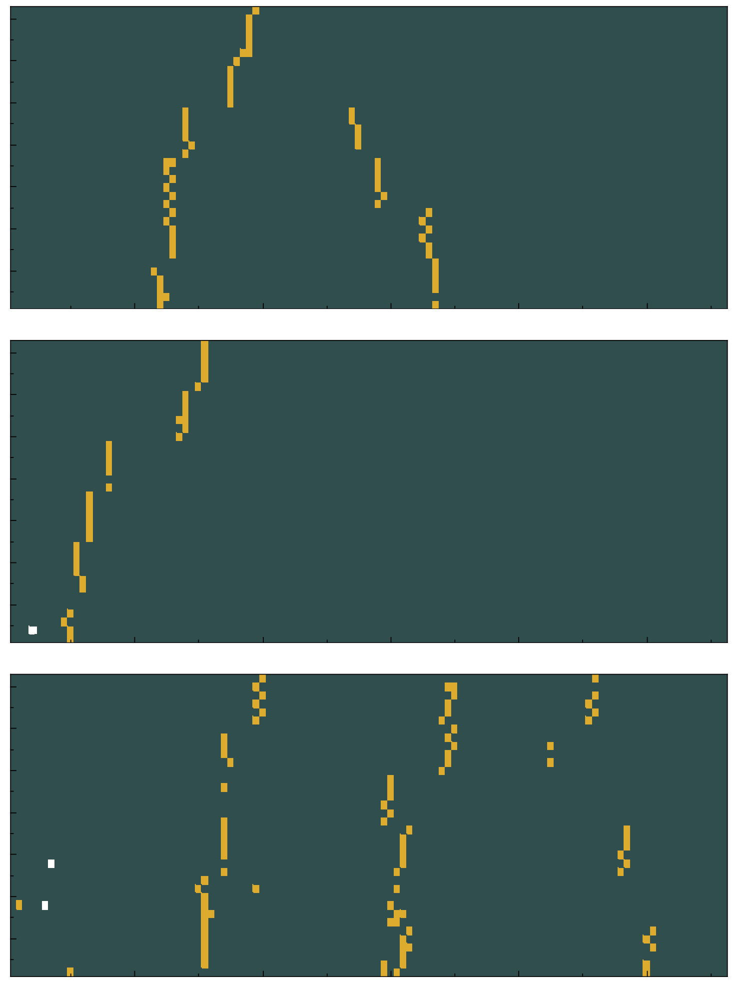
CLAS12 Track Reconstruction with Artificial Intelligence
Gagik Gavalian (Jefferson Lab), et all e-Print: 2205.02616 [physics.ins-det]

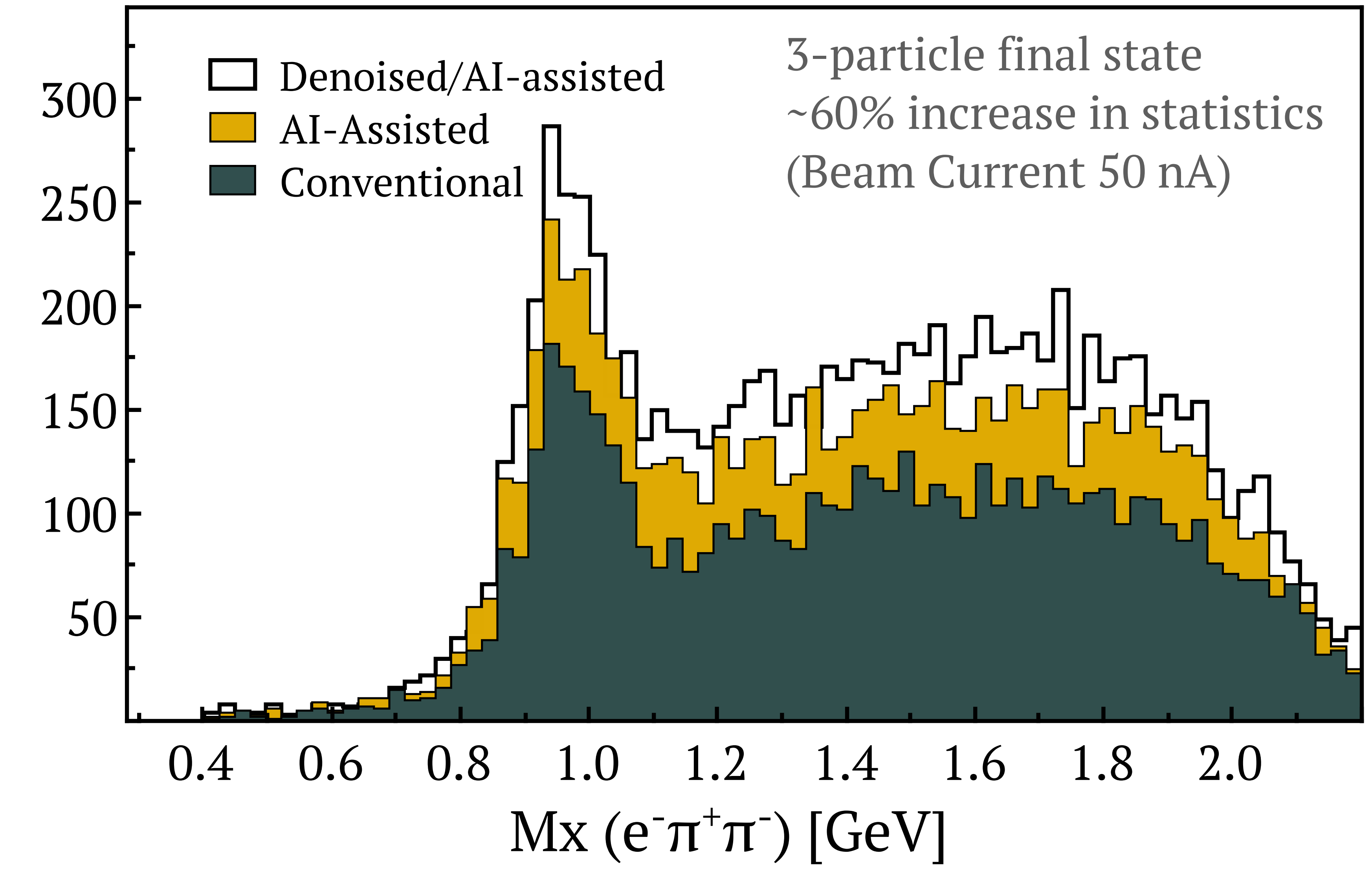


De-Noising CNN Auto-Encoder

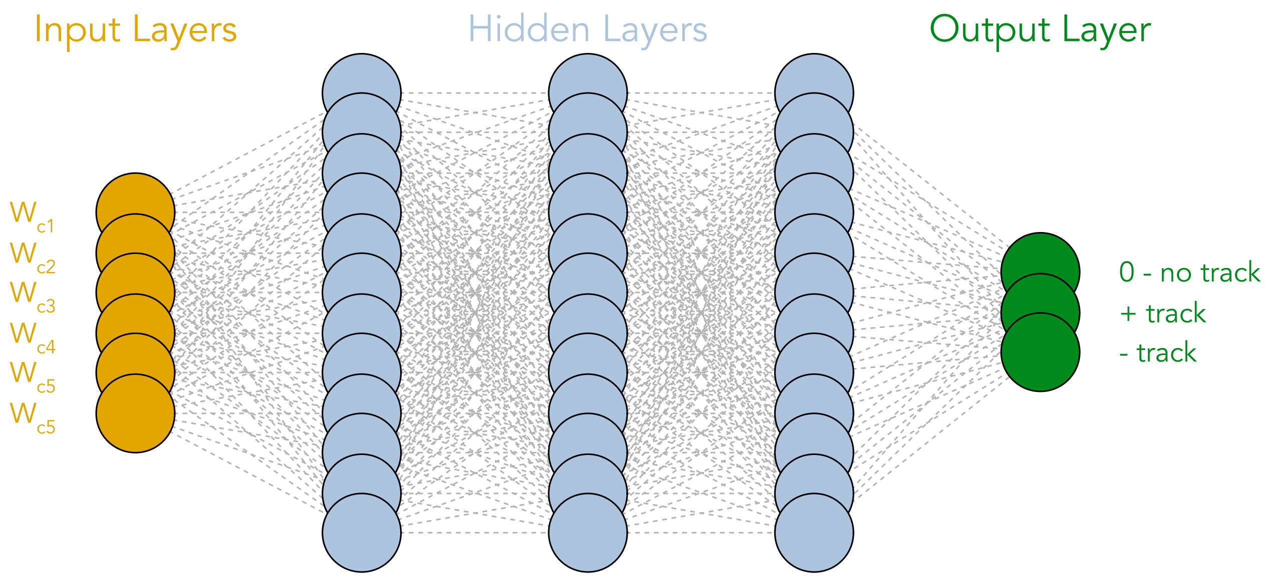


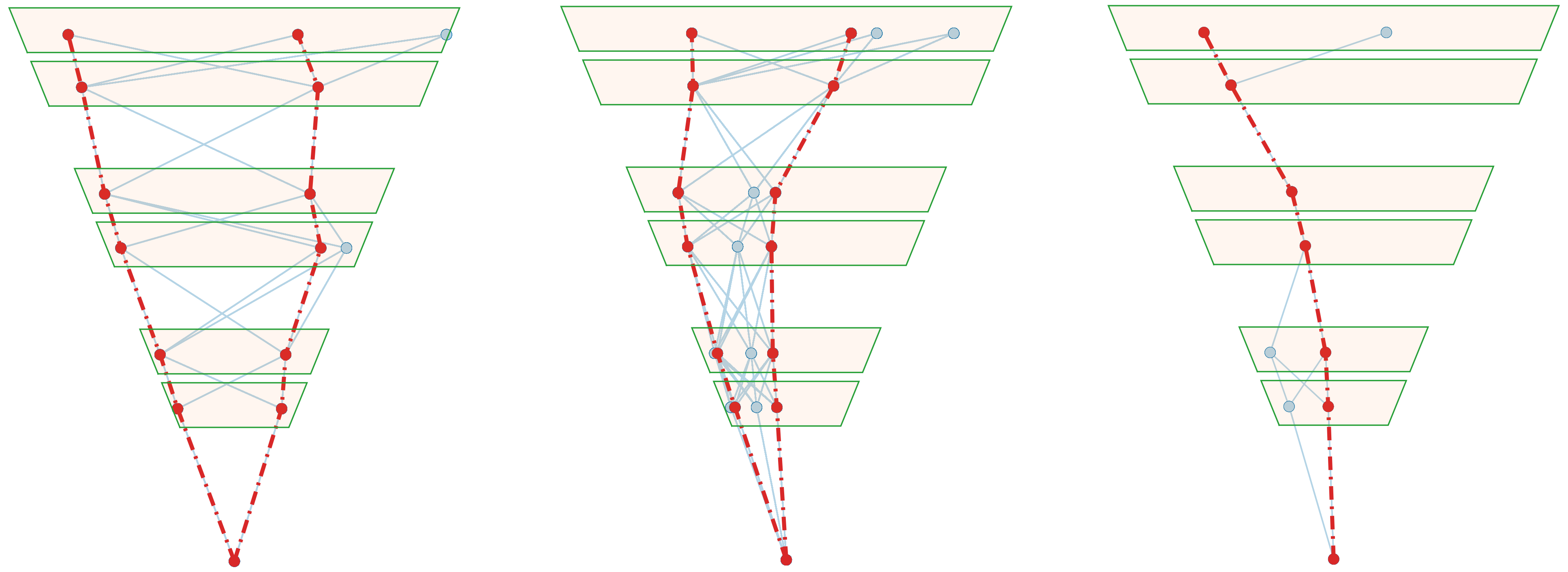
Retains >95% of hits belonging to tracks, and removes >90% of background hits

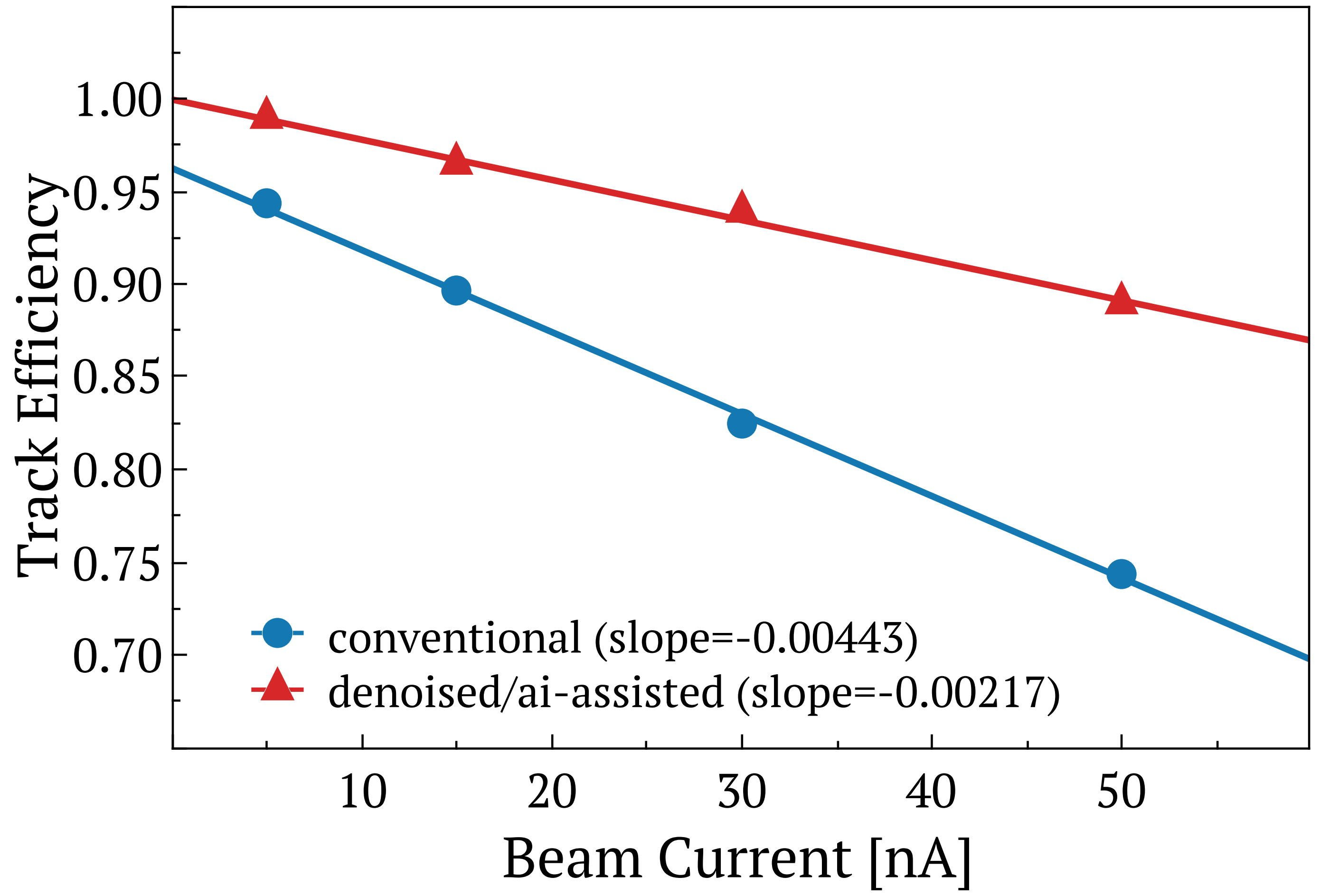


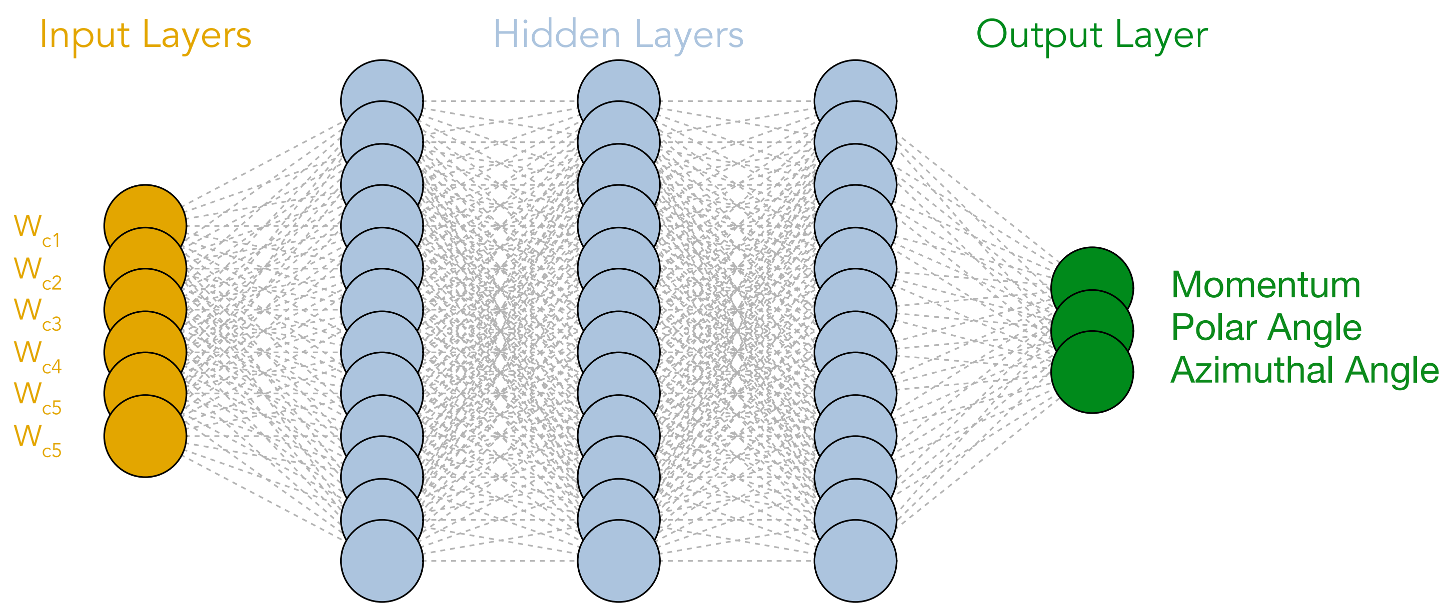


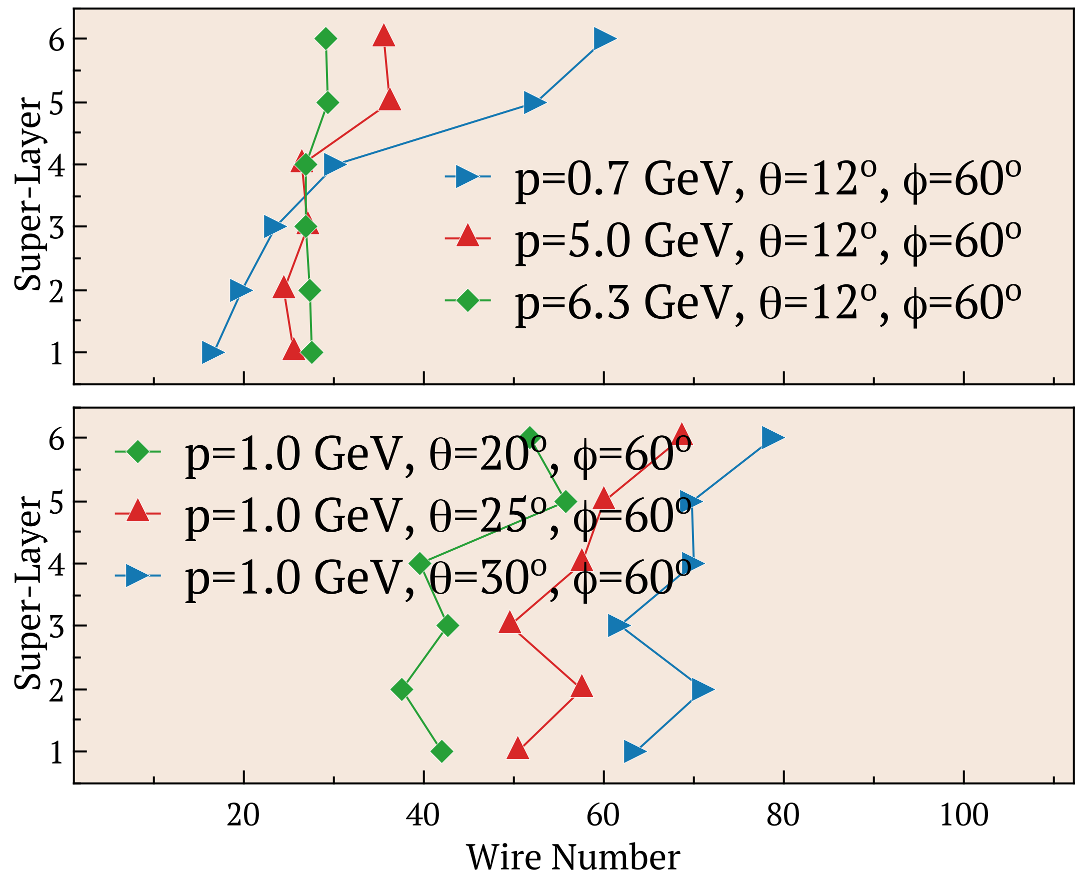
The Classifier network identifies tracks from segment combinations and identifies track charges. The AI-assisted track identification increased tracking efficiency by **15%-21%** (depending on luminosity)
Improvement of the efficiency slope as a function of luminosity.





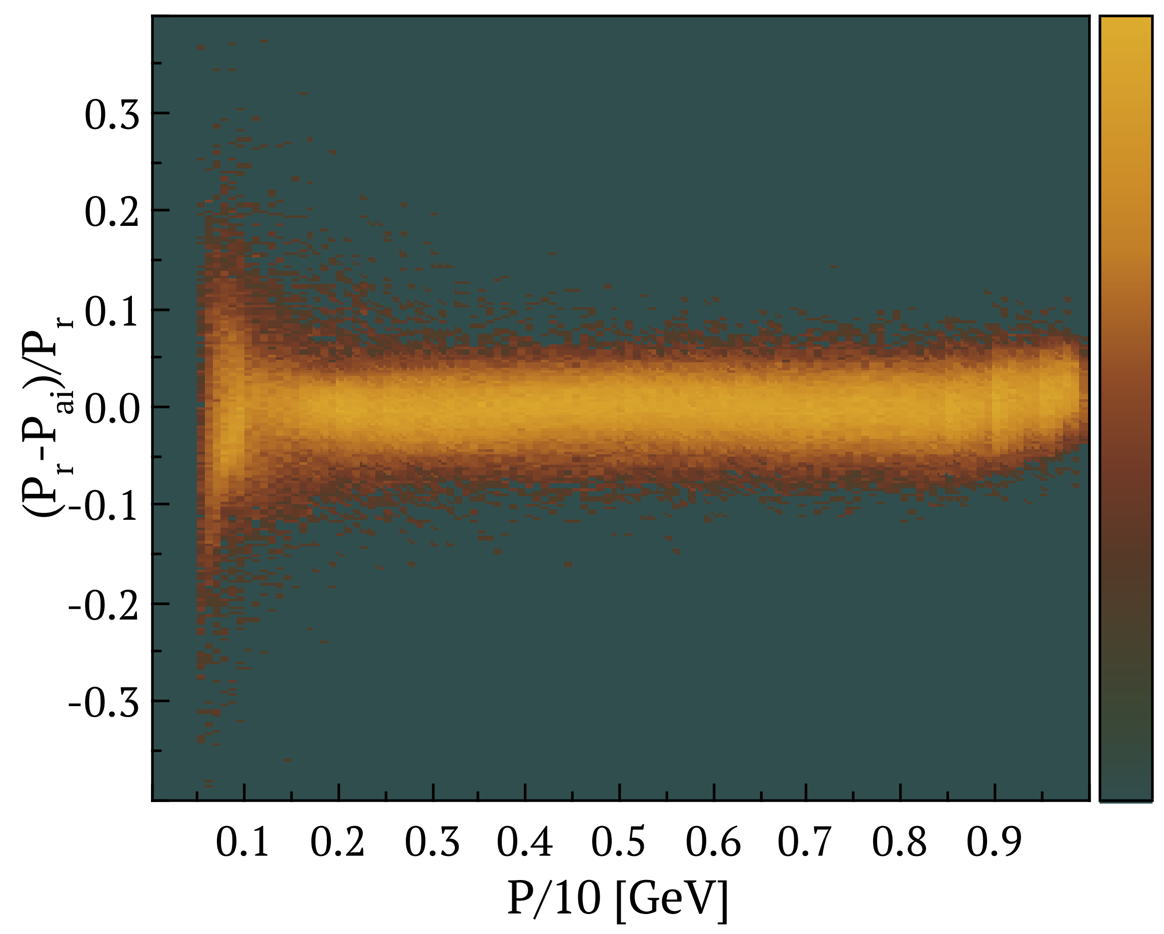


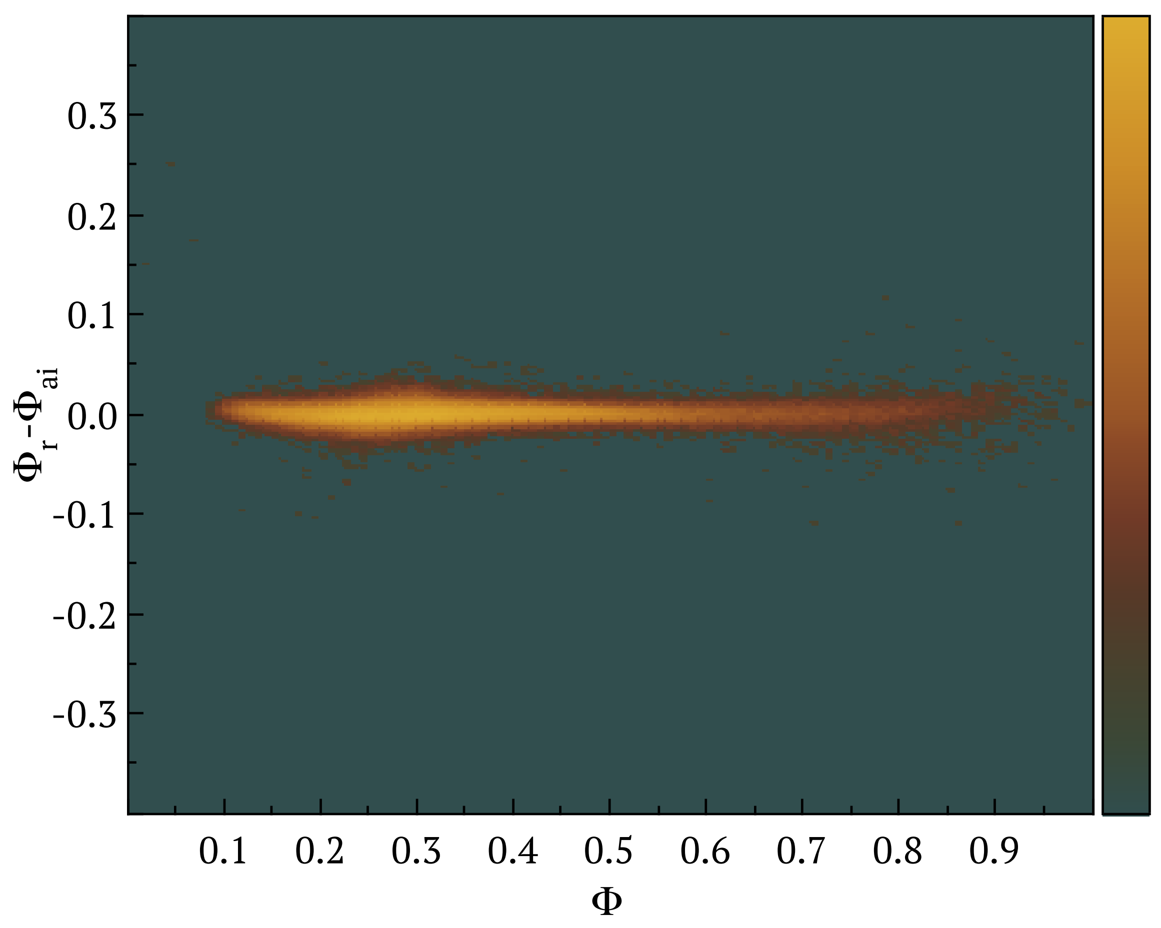


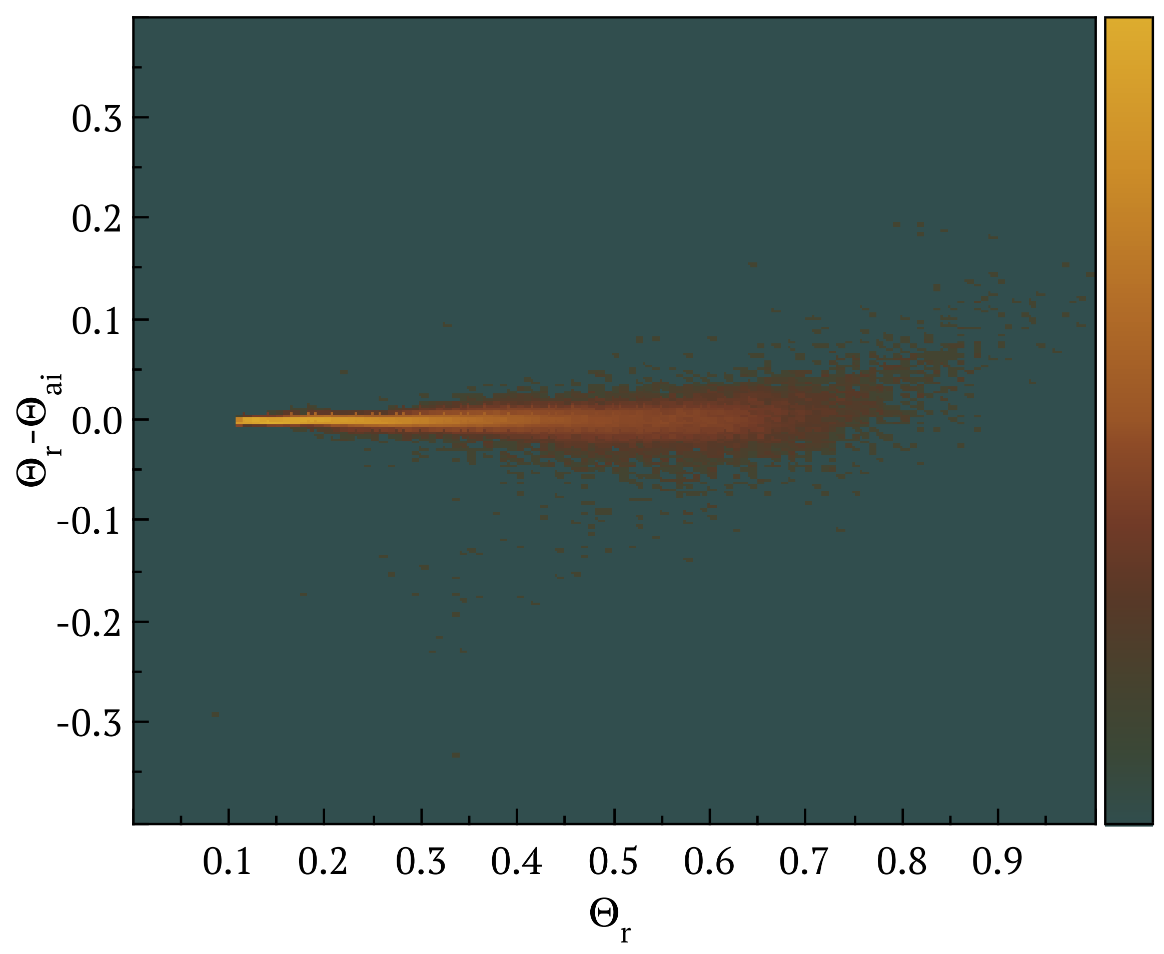


Regression Neural network to predict the track momentum and direction.
The track momentum is reconstructed with an accuracy of 1.4%-1.7%

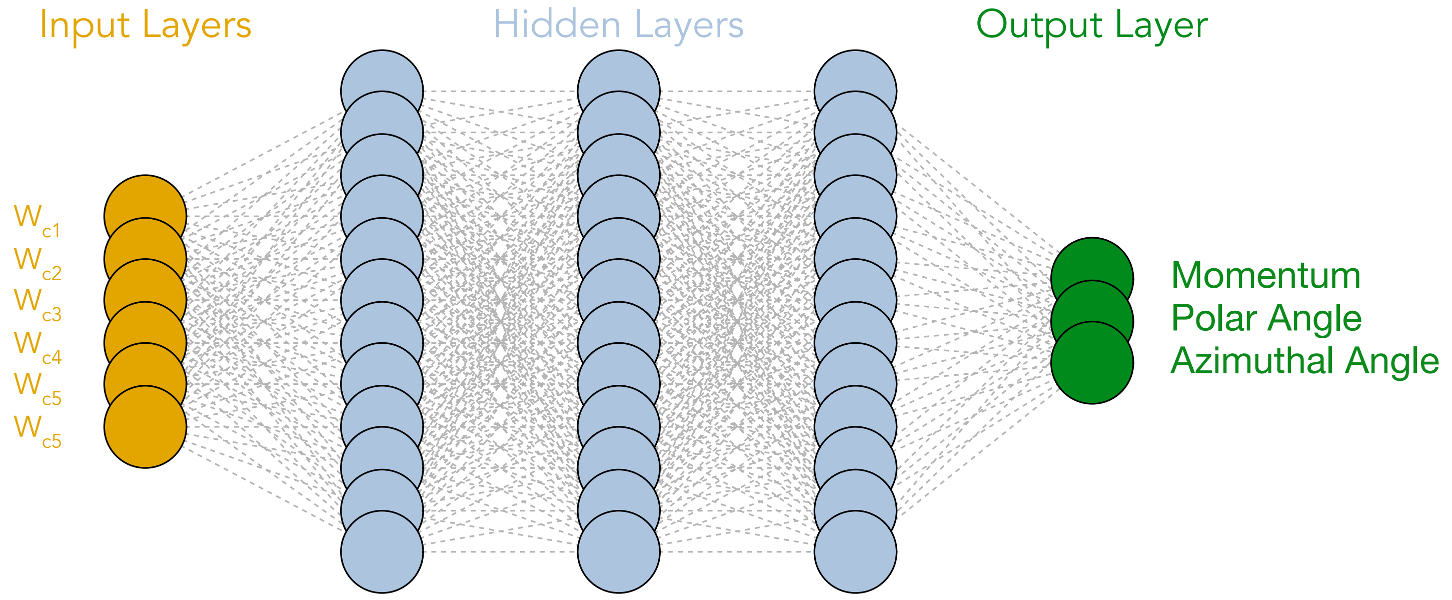
Event topologies can be cleanly identified using particle parameters inferred by the neural network.

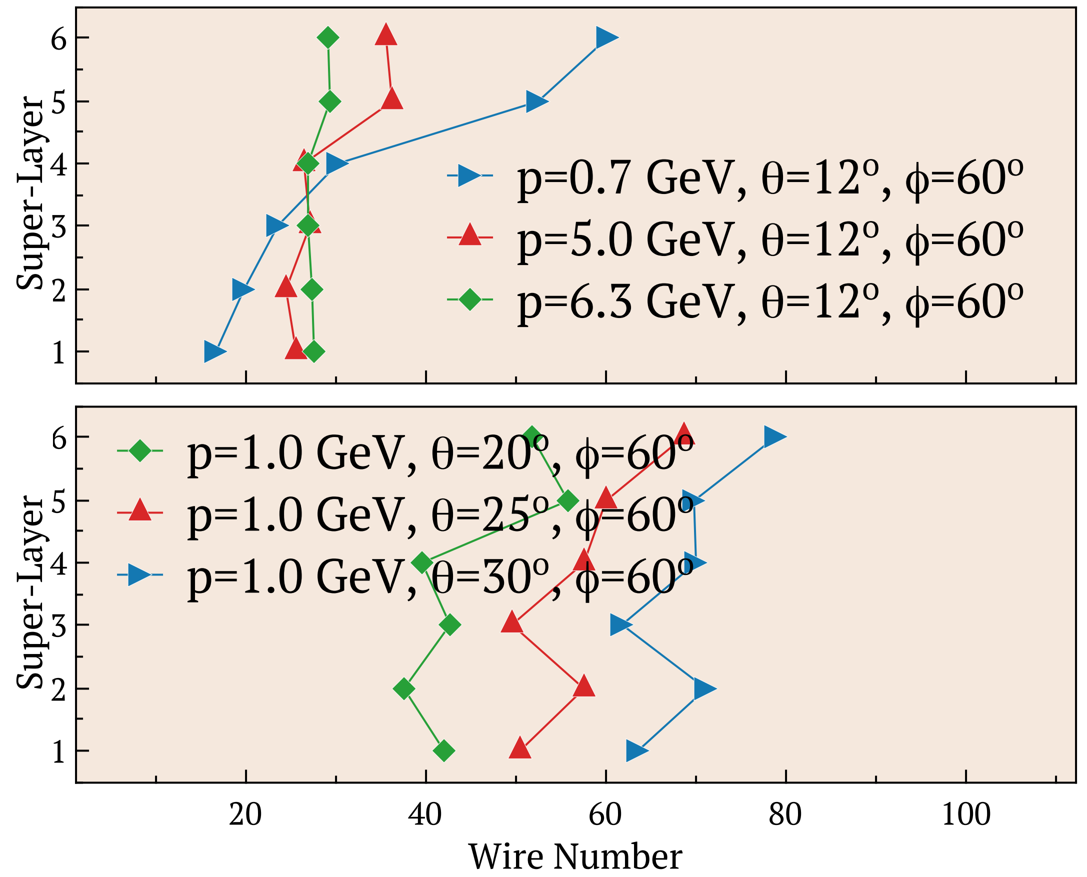






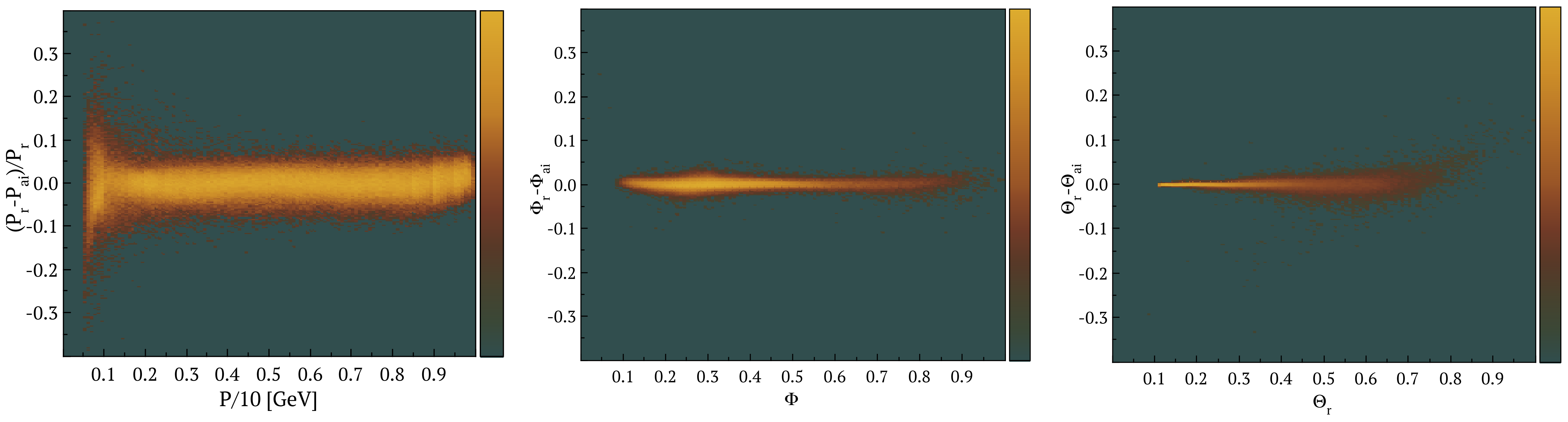
Track Reconstruction



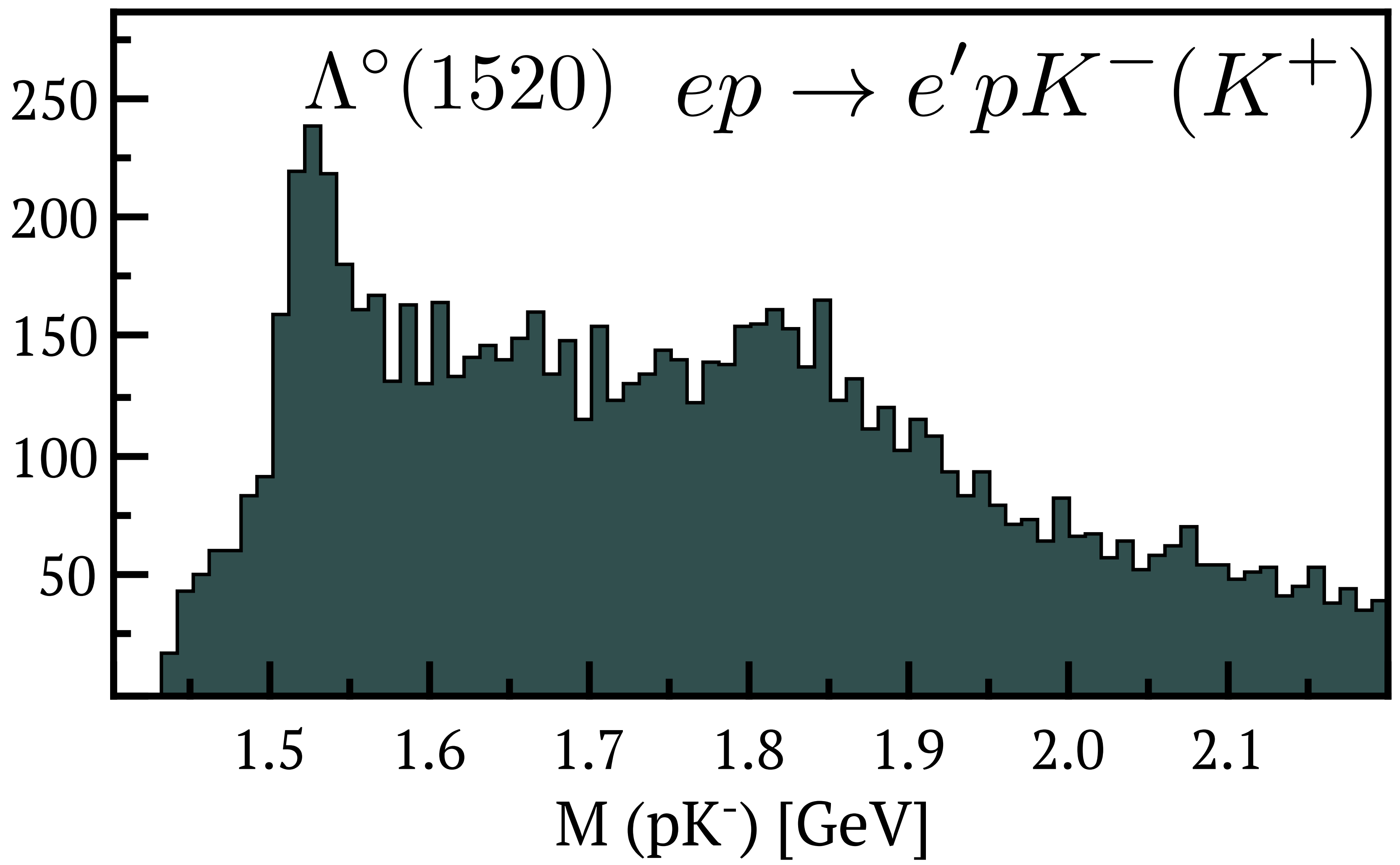


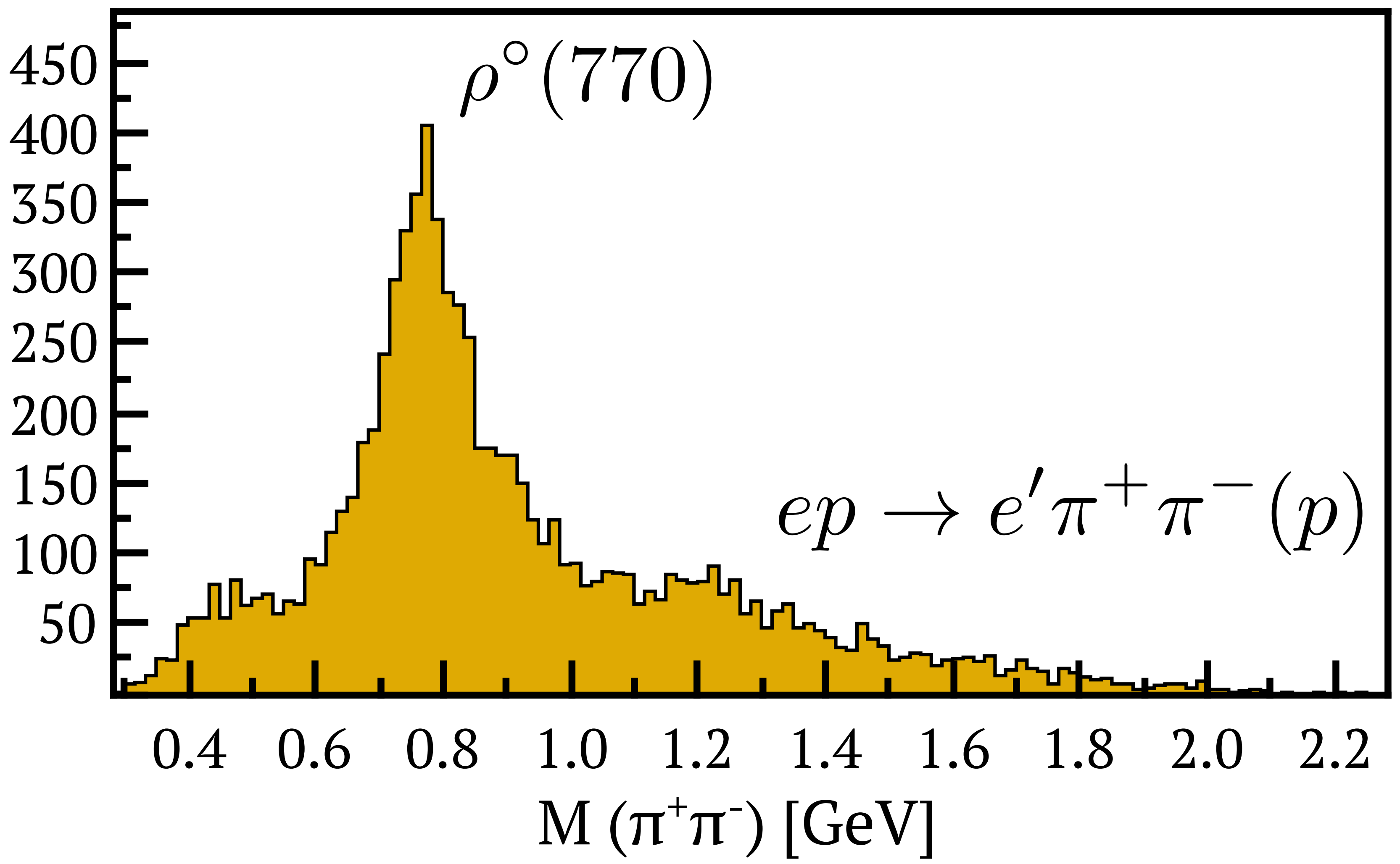
Regression Neural network to predict the track momentum and direction.
The track momentum is reconstructed with an accuracy of 1.4%-1.7%

Physics reactions can be cleanly identified using particle parameters inferred by the neural network.



Distributions calculated from track reconstruction from RAW Drift Chamber hits





Track Reconstruction (performance)

Data Used in this study:

100 M triggered events in ~4 hours
(CLAS12 DAQ ~12 kHz)

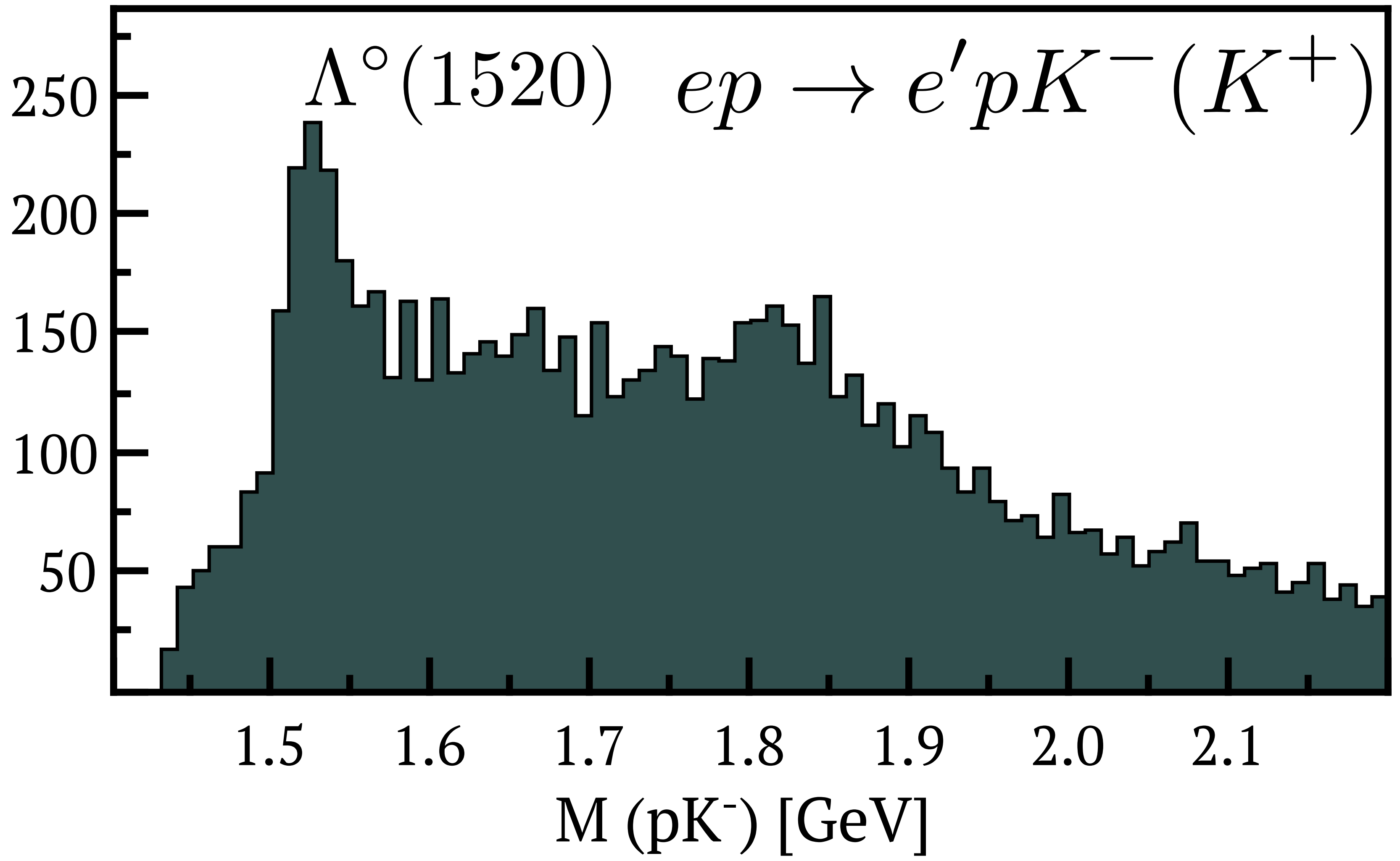
Conventional Data Processing:

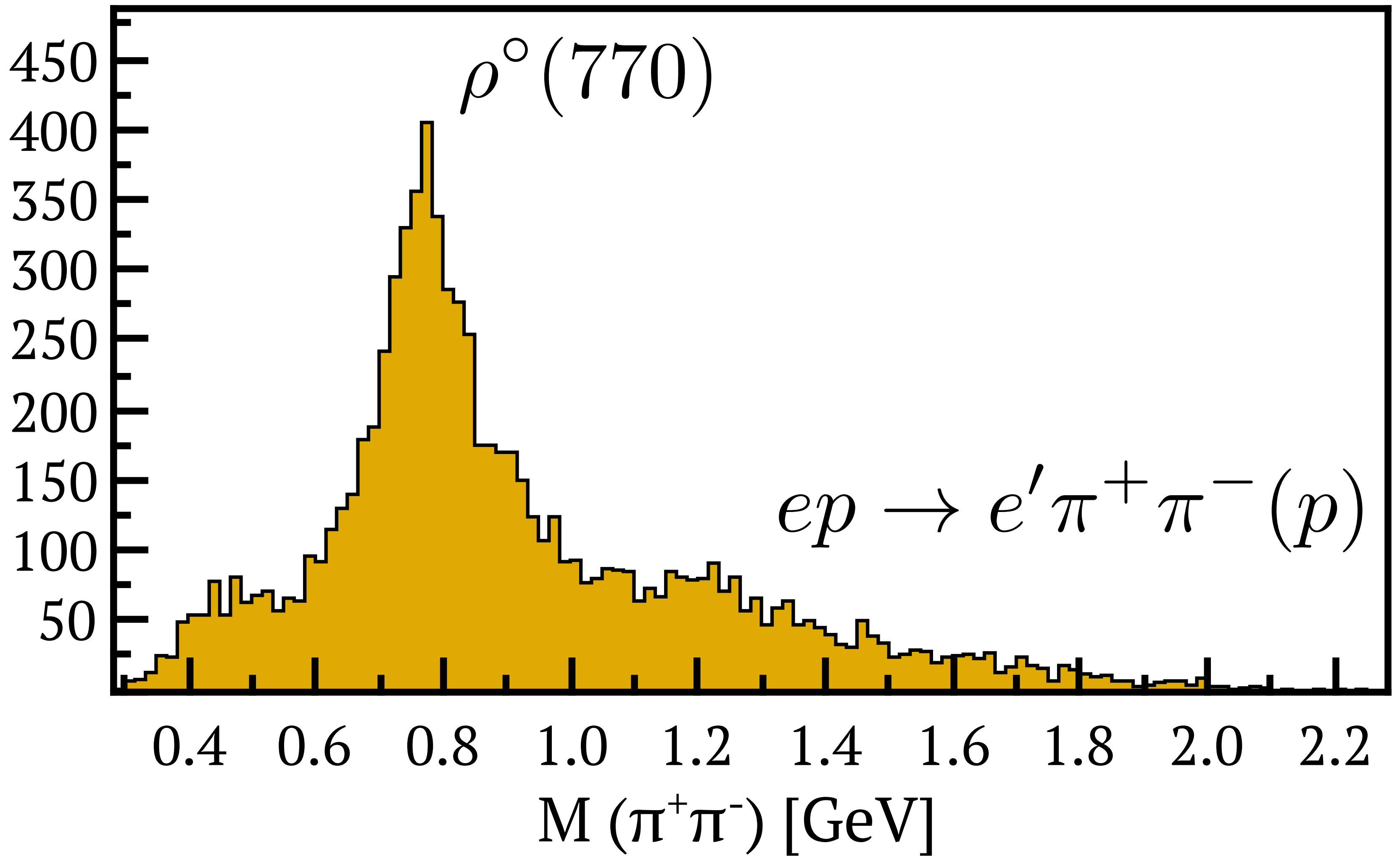
768 cores used
10 hours processing time

AI track Reconstruction:

12 cores (MacBook M3)
25 minutes processing time
(8 kHz per/core, 96 kHz on M3)

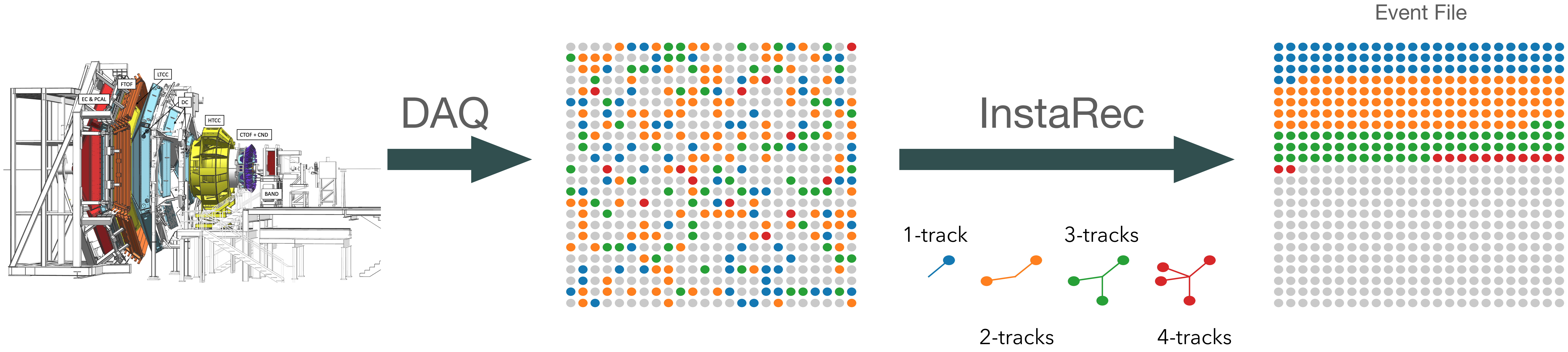
Distributions calculated from AI track reconstruction from RAW Drift Chamber hits (Obtained in 25 minutes on MacBook M3)



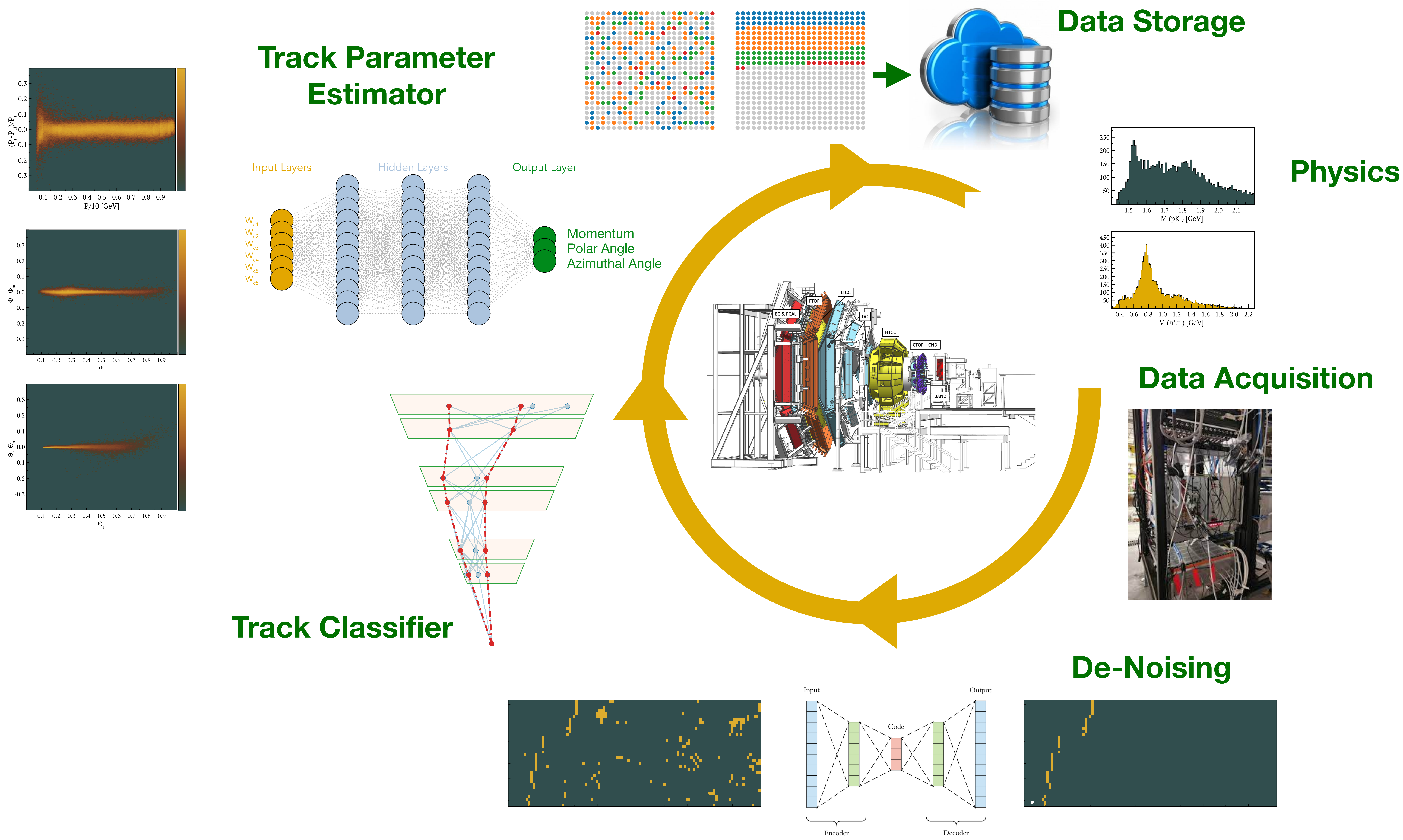


Data Sorting

- CLAS12 uses a custom data format (HIPO) with a bucket tagging feature that allows similar events to be grouped.
- The online track processor can sort the raw data by particle topology.
- The work on particle IDs is in progress; with PID we can also identify reactions and tag accordingly.
- This development will be crucial when transitioning to streaming readout.



Physics Online (aka the future)



Software

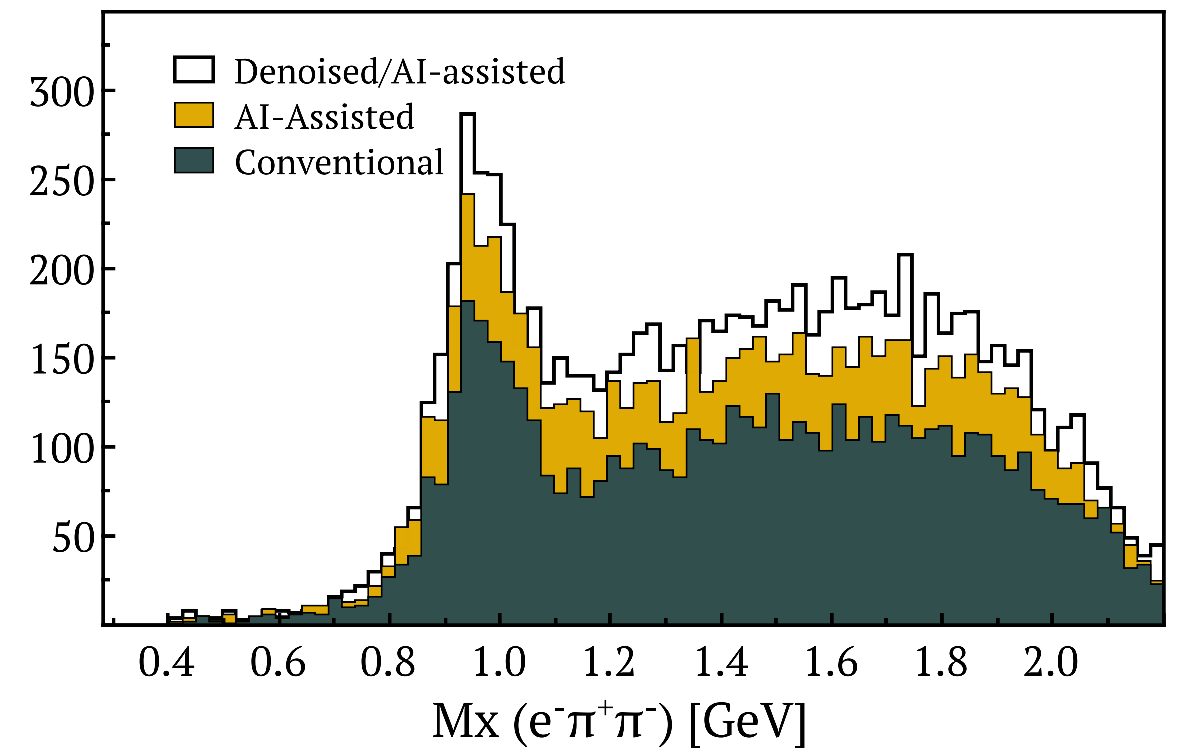
- Many (our collaboration included) think that all AI/ML work is ONLY in Python
- The track reconstruction project, track classification, and track parameter prediction are done in Java (both training and deployment)
- The CLAS12 Reconstruction software is in Java; the AI track classifier and track predictor are easily integrated
- The DeepNetts library was used for all of the developments (https://www.deepnetts.com)
- Java provides platform-independent software for training and validation software. Significantly simplifies the workflow.

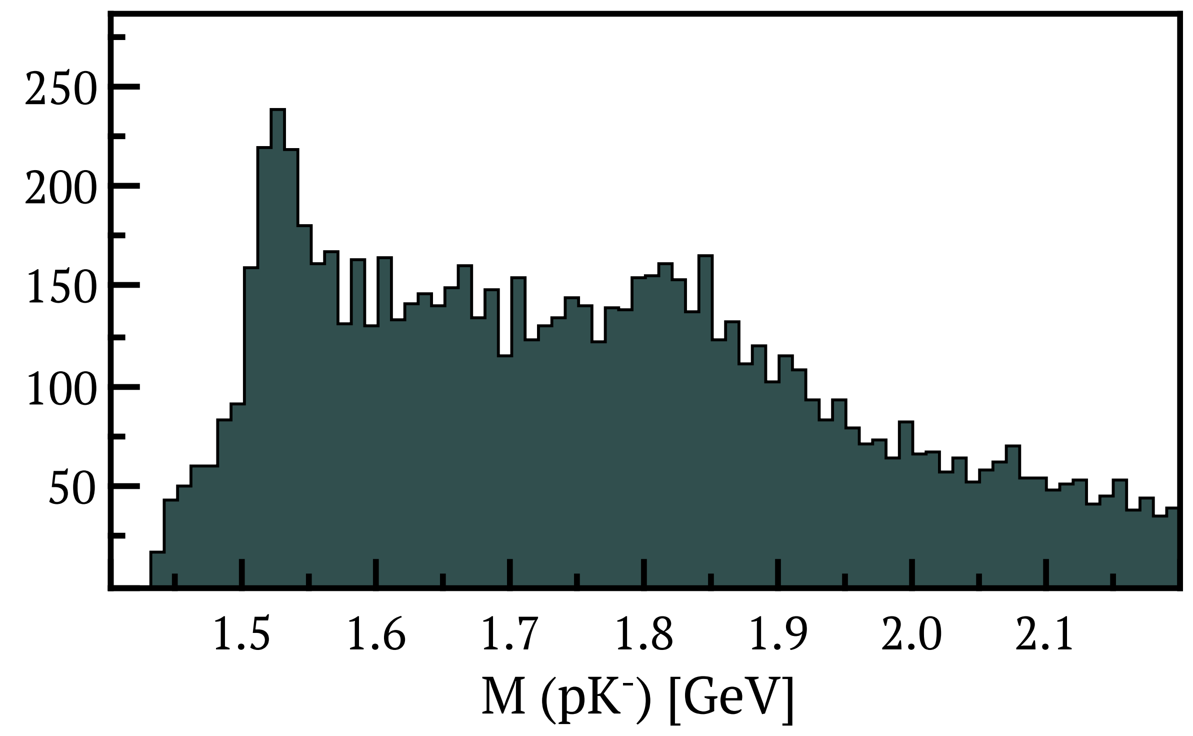
DEEP NETTS

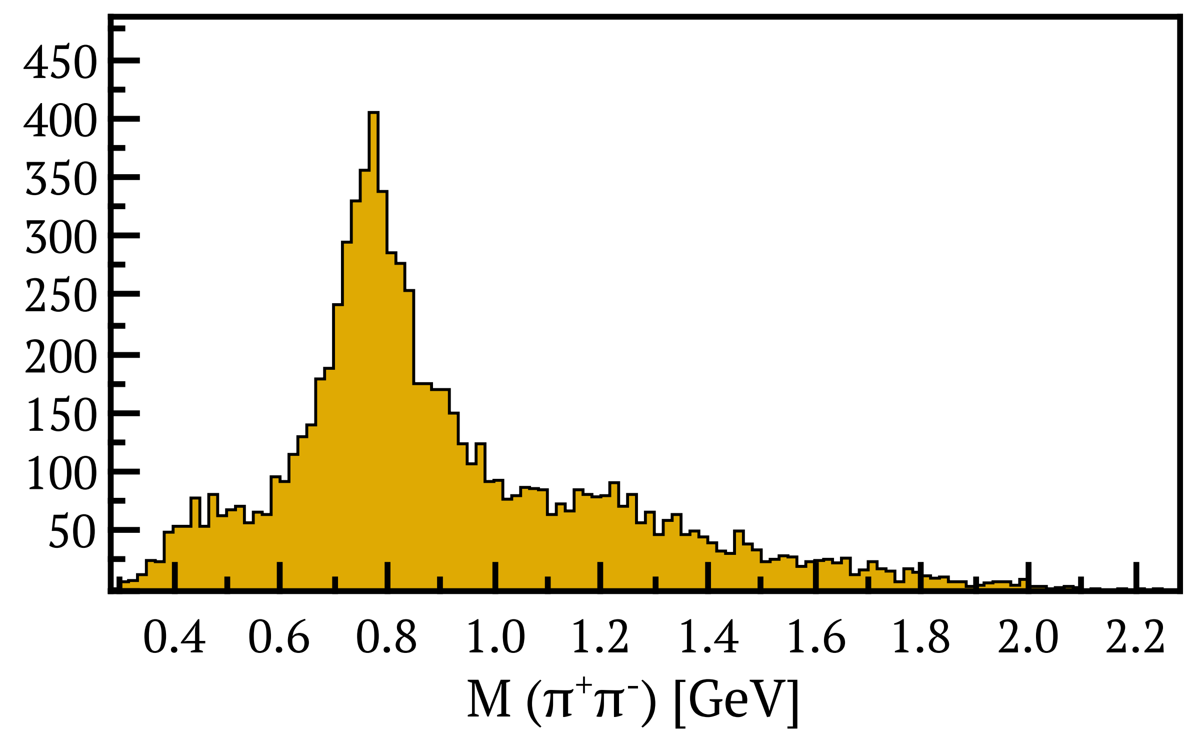
Summary

- AI track identification (including denoising) is now part of the reconstruction framework.
- Using AI in reconstruction increased the statistics of existing experiments by **~60%-75%**
- Due to improved single-particle efficiency, future experiments can run at higher luminosities, resulting in more data.
- The InstaRec reconstructs tracks at 8 kHz/core rate allowing physics online
- InstaRec allows us to select event topologies in real-time, improving data processing times.
- Further developments in particle identification will allow triggering in streaming data regimes and identify physics reactions at the DAQ level.

Physics with AI reconstruction Online

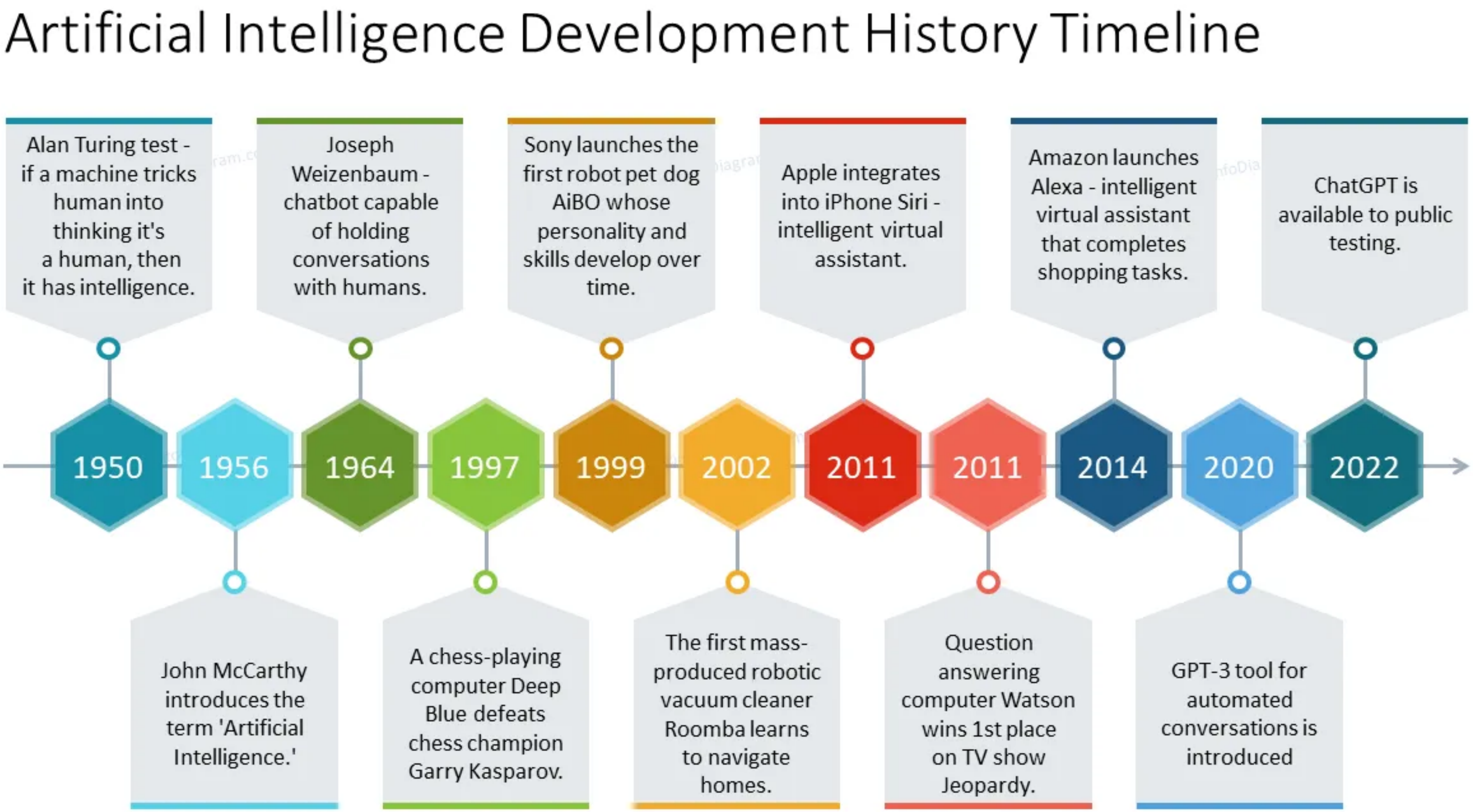




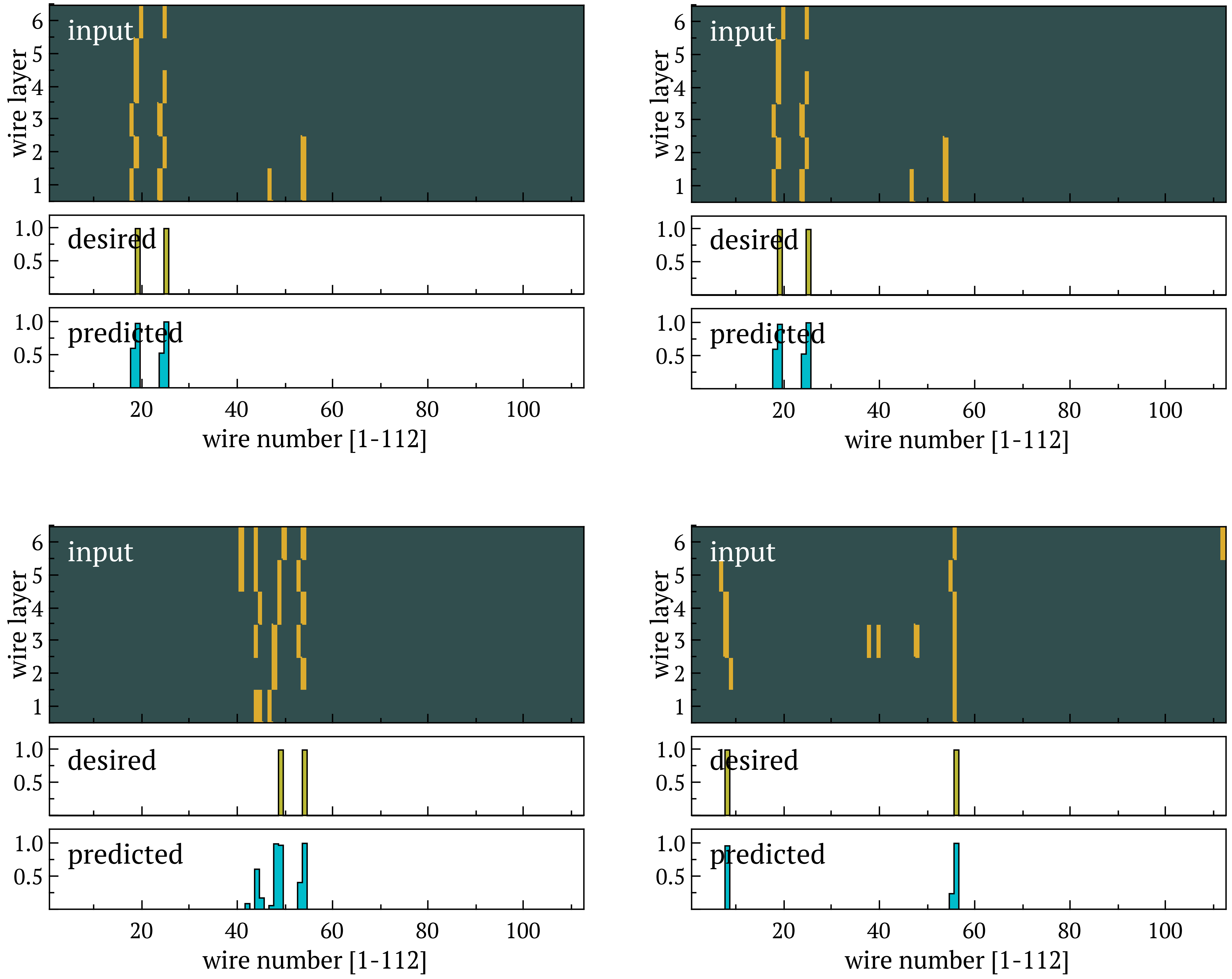


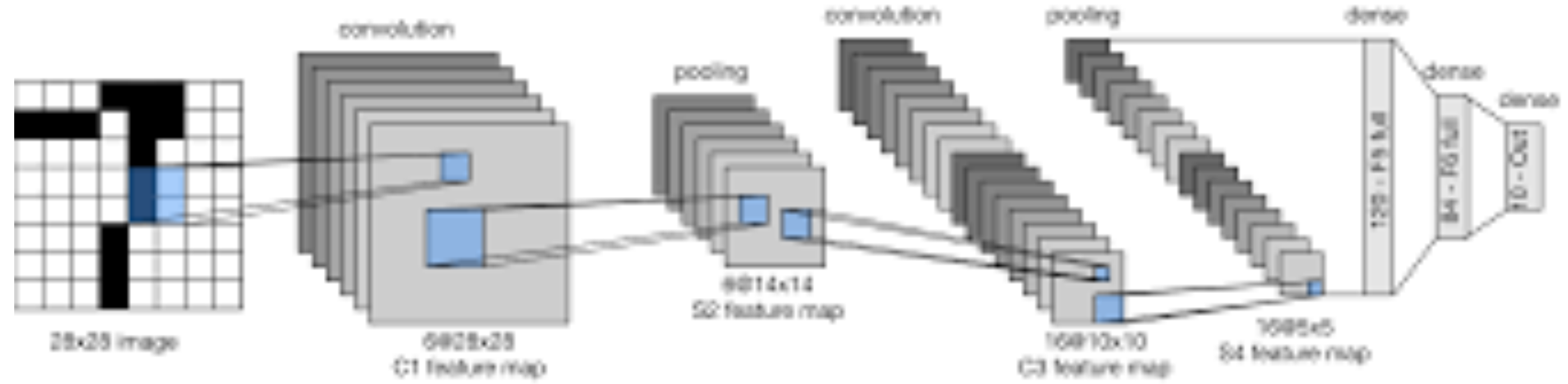
BACK UP SLIDES

Artificial Intelligence (short history)



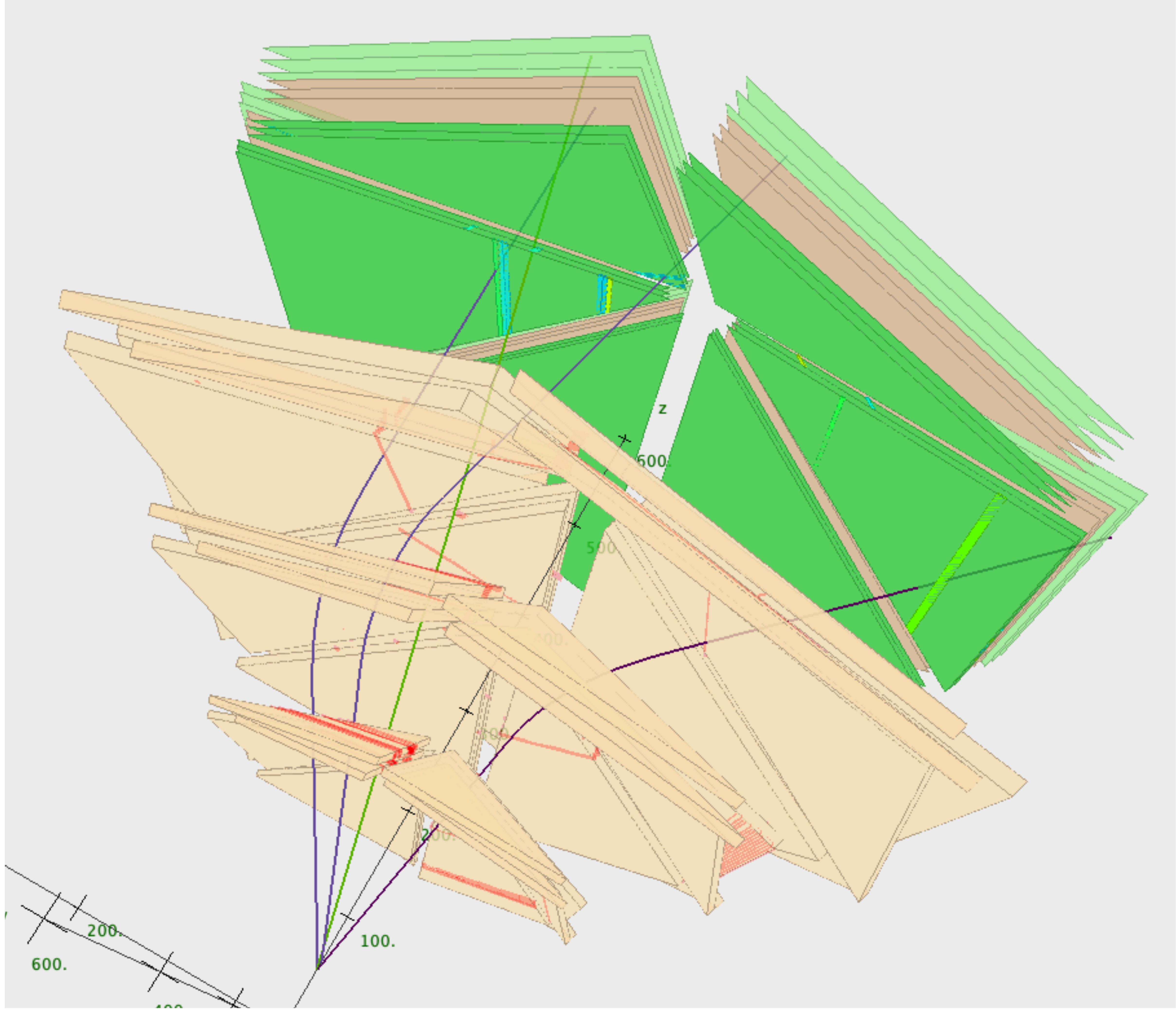
Feasibility tests



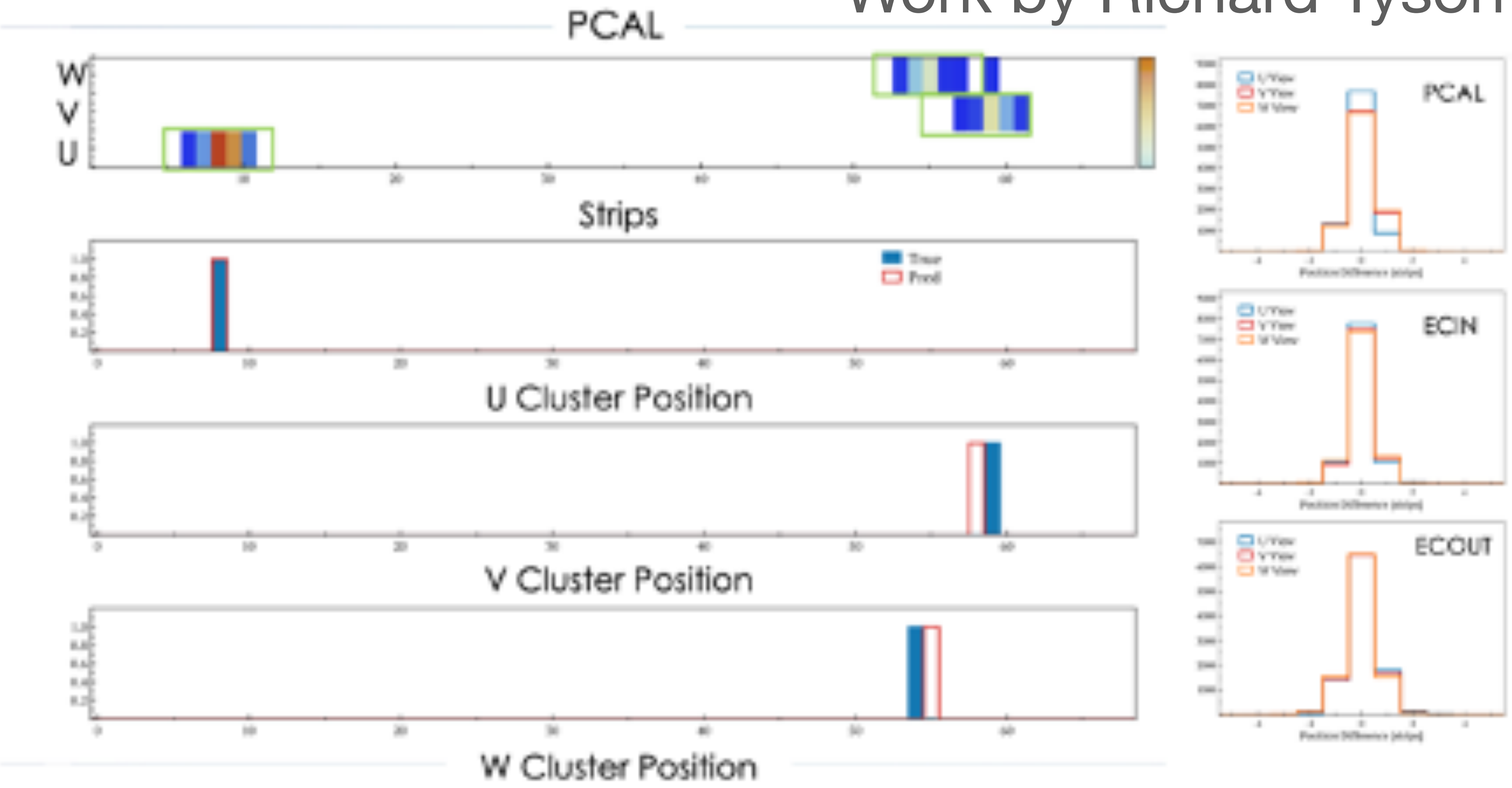


- Clustering Algorithm:
  - Convolutional Neural network with logistic regression to identify possible cluster positions.
  - Currently at 97% efficiency
- Proposed work:
  - Investigate algorithms to identify clusters fast and with higher efficiency

Feasibility tests

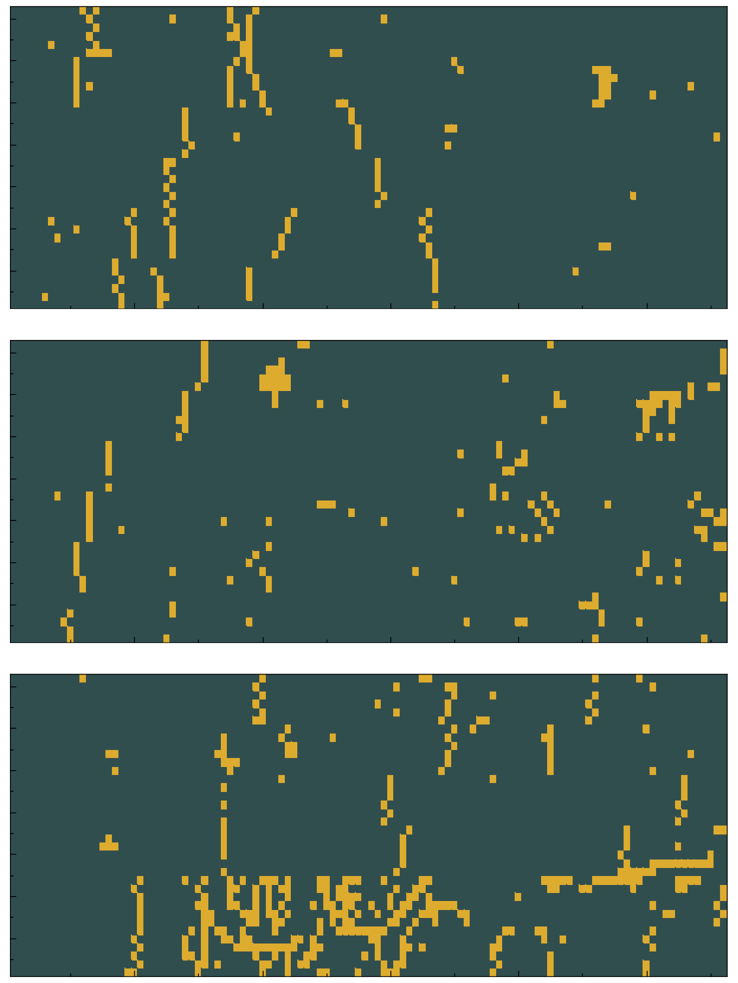


Work by Richard Tyson

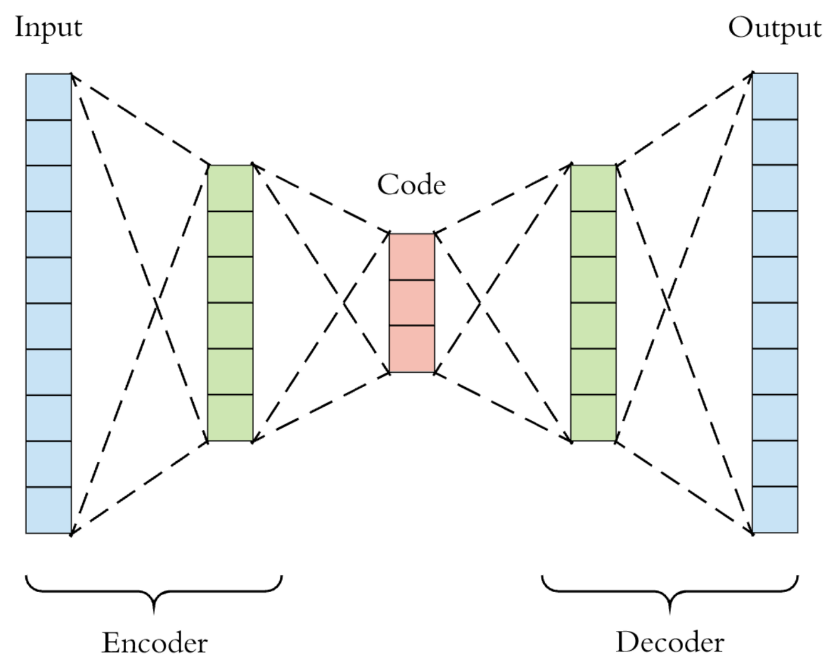


- Initial Tests:
  - A Neural Network to predict the track's impact point on the calorimeter surface.
  - Capable of predicting position within one strip
- Future Developments:
  - Extend the network to predict the track's impact position with all the detectors on the track's path
  - Construct dependency graph for detector responses
  - Develop electron identification algorithm (first)
  - Develop general particle identification

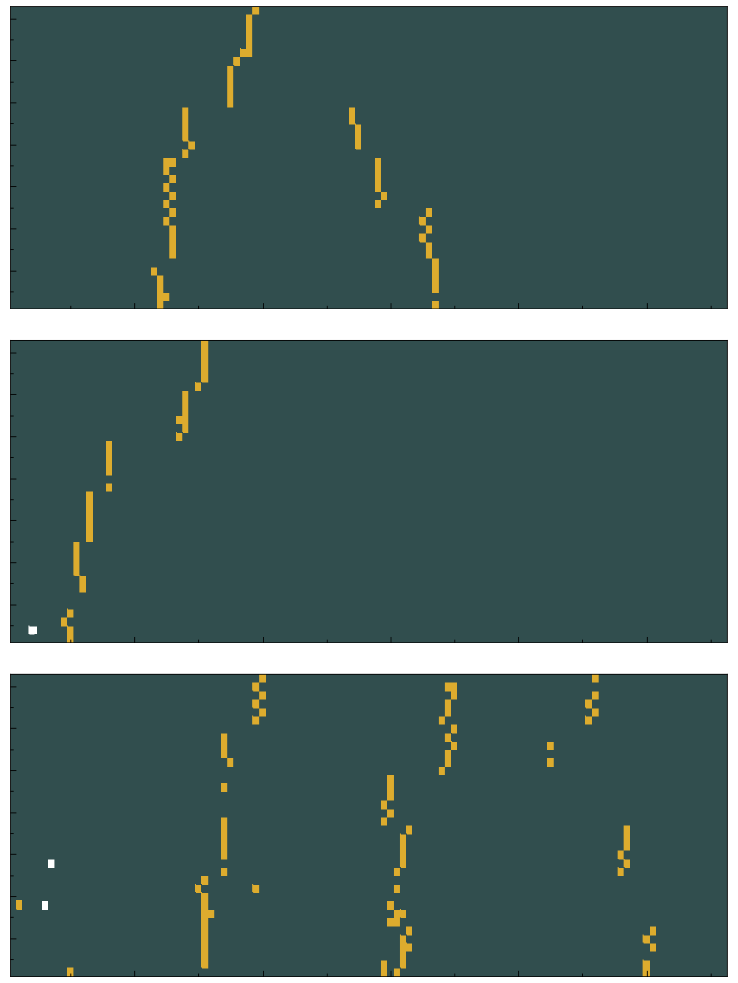
What do we have now



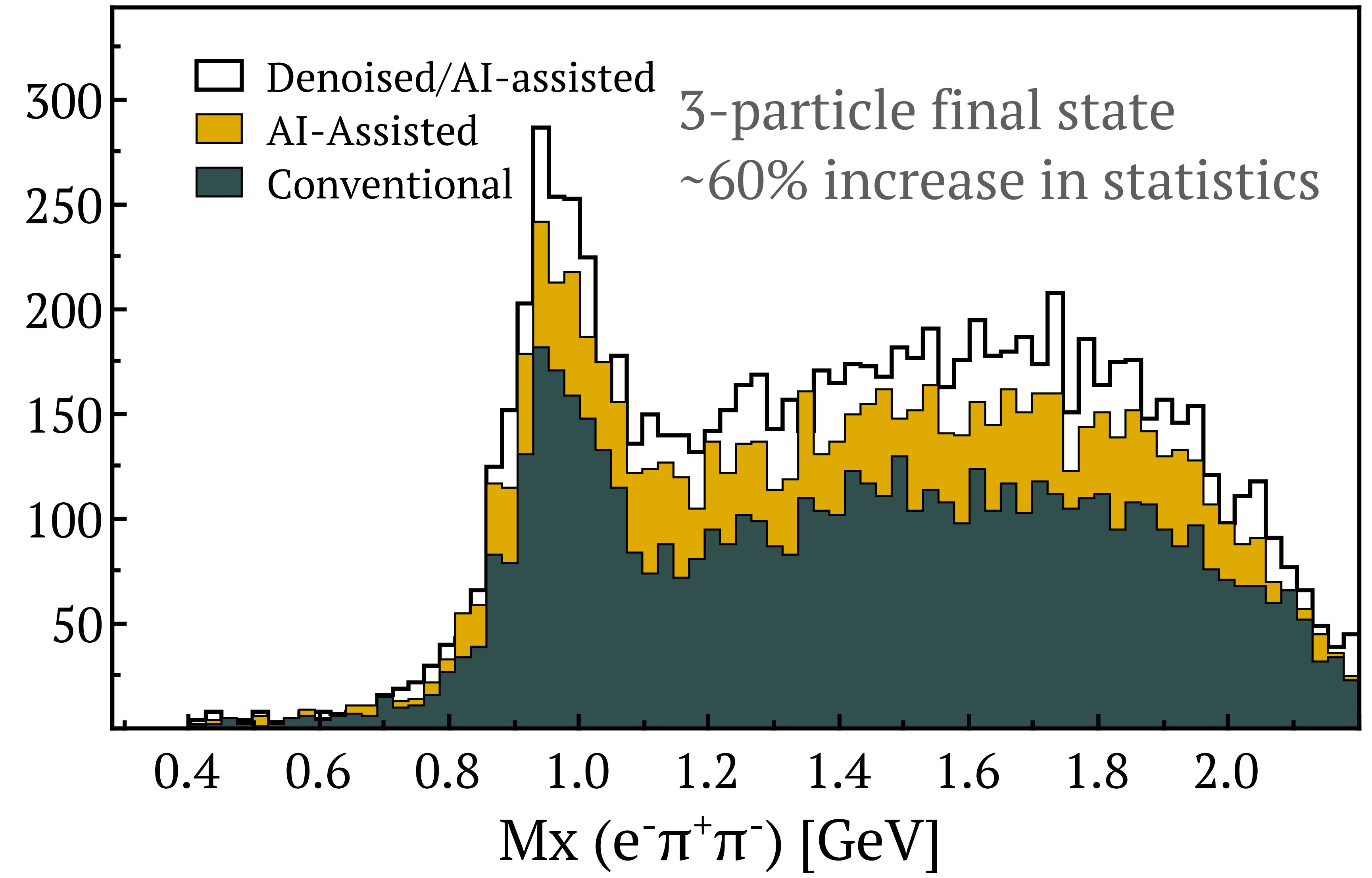
De-Noising CNN Auto-Encoder



Retains >95% of hits belonging to tracks, and removes >90% of background hits



CLAS12 Track Reconstruction with Artificial Intelligence
Gagik Gavalian (Jefferson Lab), Pet all e-Print: 2205.02616 [physics.ins-det]



The Classifier network identifies tracks from segment combinations and identifies track charges. The AI-assisted track identification increased tracking efficiency by **15%-21%** (depending on luminosity)
Improvement of the efficiency slope as a function of luminosity.