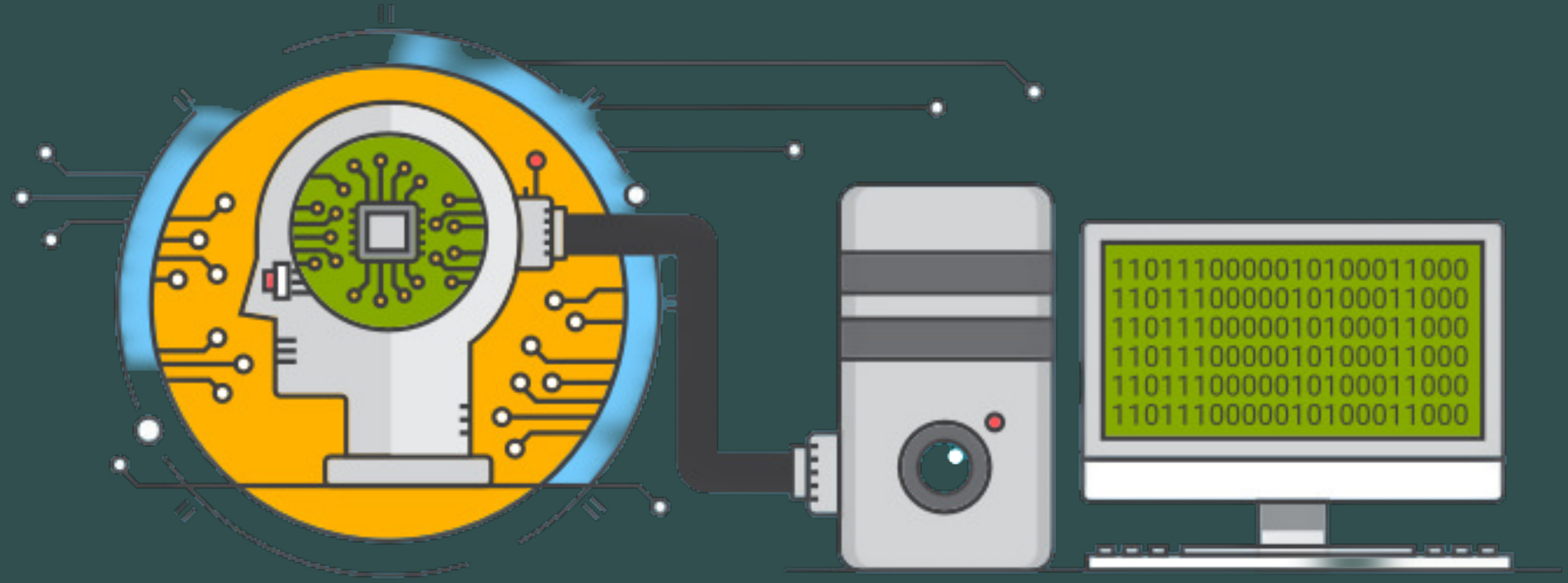
Real-Time Physics Analysis using AI Track Reconstruction Online

LDRD Proposal

G.Gavalian (Jefferson Lab)

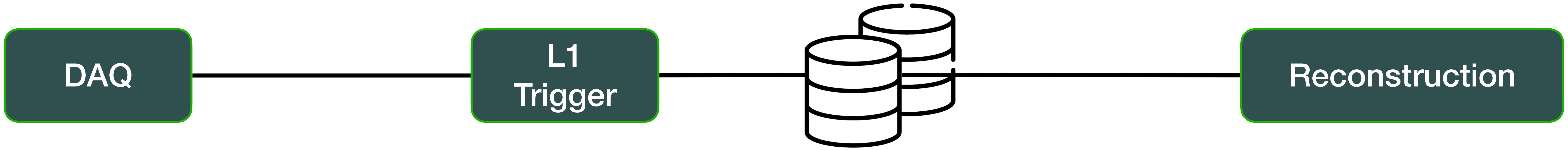


LDRD-2025

Outline

- What is the problem we are trying to solve?
- The goal of the project
- Feasibility studies
- Budget
- Milestones

DAQ

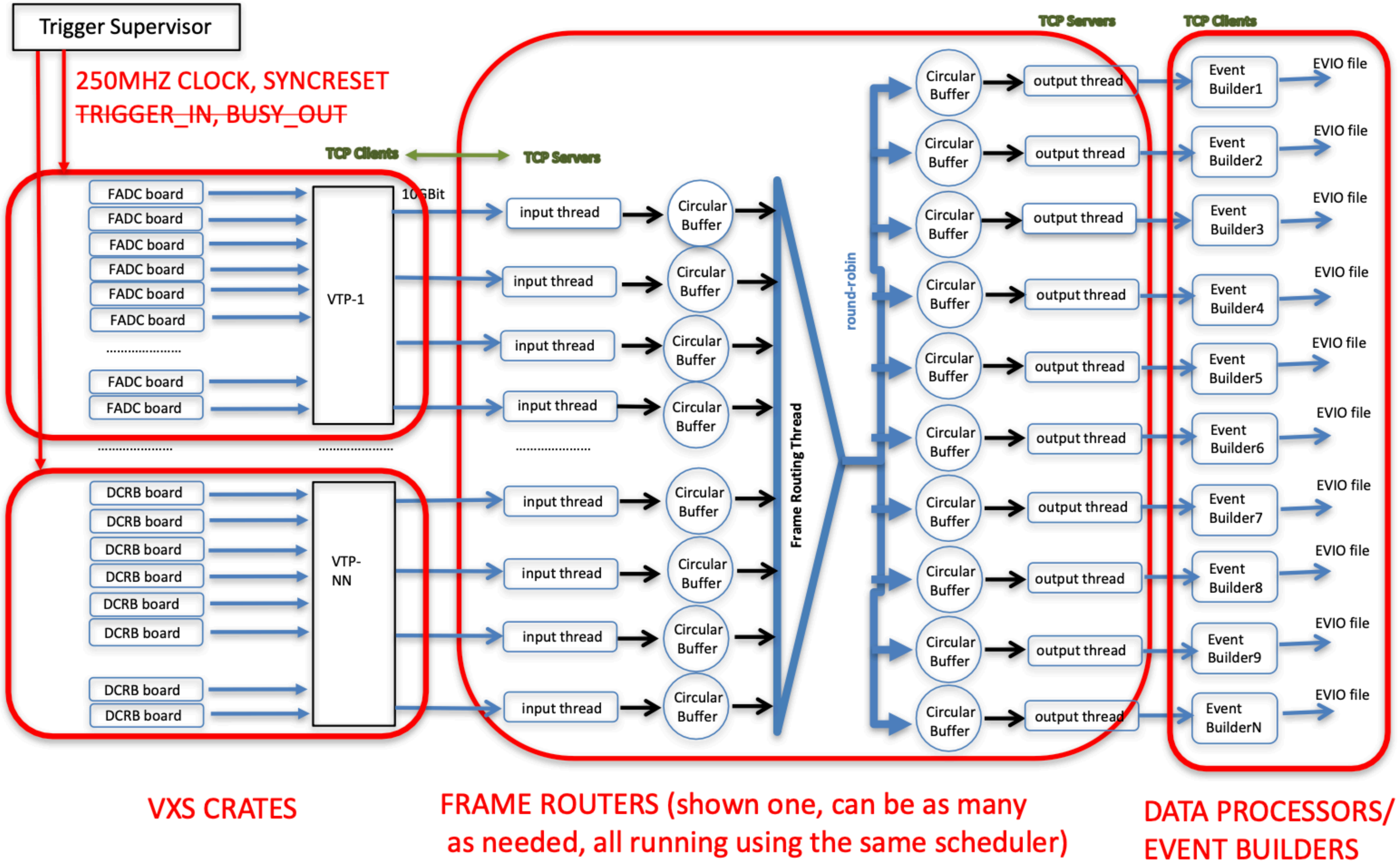


Current Running Conditions	16 kHz	400 MB/sec	40% contain electrons
Moving to High Luminosity	50 kHz	1.5 GB/sec	~15% contain electrons

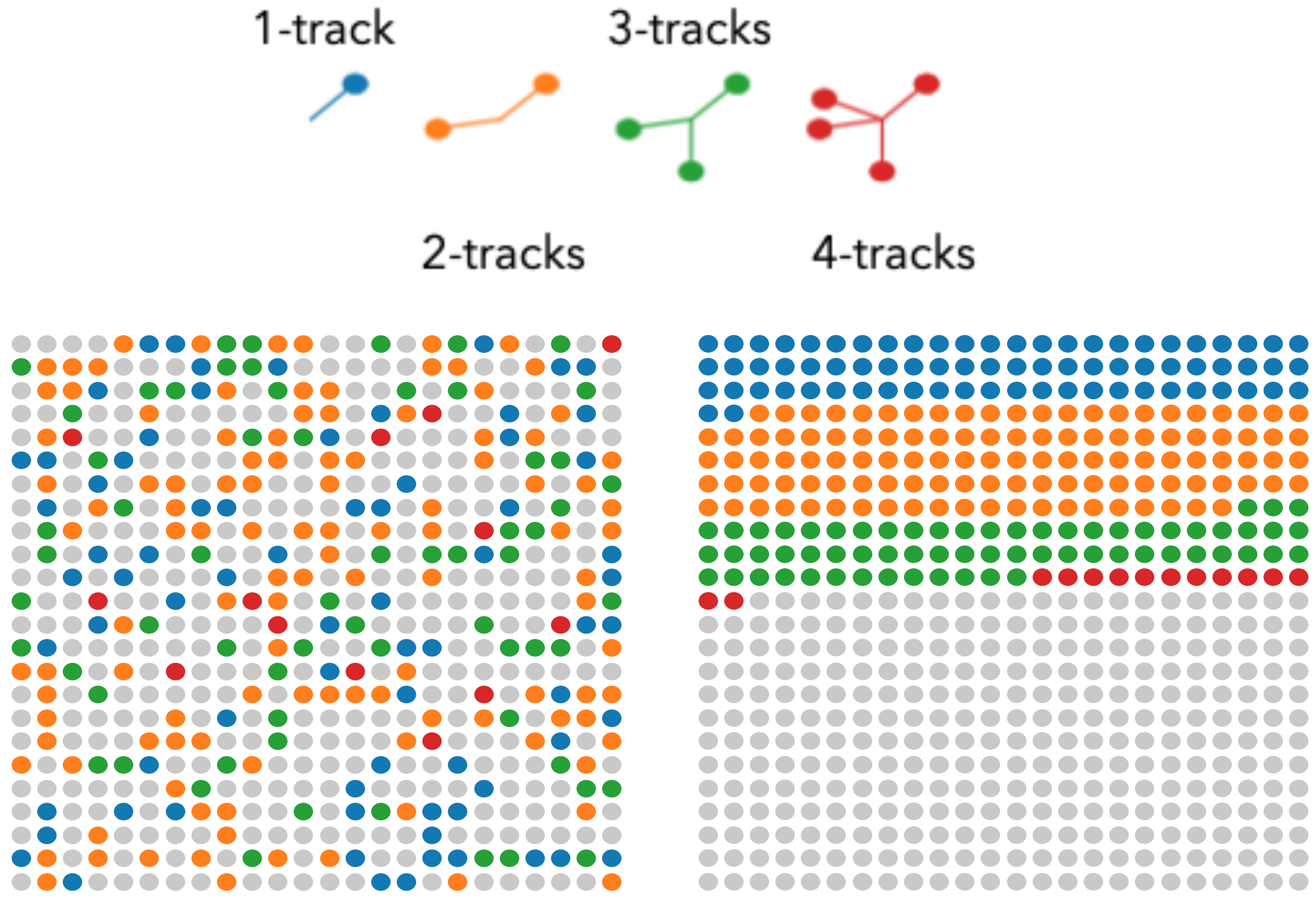
- Can we reduce the data in real time?
- Processing events with identified electrons is 7 times faster. (at 15% purity)
- Can we use AI to identify events that are worth keeping?
- What will happen with the transition to Streaming Readout? Are we ready?
- Can we do it at DAQ rates?

SRO

Hall B Streaming DAQ Diagram



- Streaming DAQ Hall-B
  - 2 sectors are implemented (only Forward)
  - Data rate 2.5 GB/sec (for 2 sectors Forward)
  - Expected Rate: 15 GB/sec
  - With Luminosity upgrade 50 GB/sec



Can we organize the output data by event topology in real-time?

The Goal

- **We propose to use AI to select events in the DAQ stream.**
  - Find segments in drift chambers from raw hits
  - Identify tracks based on combinations of segments
  - Find track parameters (momentum, direction, vertex)
  - Identify electrons by matching tracks to relevant detectors (ECAL and Cherenkov)
  - Select events that satisfy physics trigger requirements

What is missing?

1 GB/sec (SRO 50) GB/sec

300 GB/sec (SRO 3-5 GB/sec)

DAQ/SRO

L3 InstaRec

Reconstruction

Initial estimates suggest 92% purity

✗	Data Converter	A fast multithreaded converter to convert EVIO to HiPO Apply translation tables, construct the data for AI
✗	Segment Finder	Clustering algorithm to find segments in drift chambers Current segment finder can not keep up with data rates
✓	Track Finder	Track finder identifies correct segment combinations And outputs track candidates.
✓	Track Parameter Finder	For track candidates from AI track finder predicts tracks' momentum, direction, and vertex
✓ (yellow)	Electron identification	Develop AI infrastructure to associate tracks with Other detector components, train on data to identify electrons

12 kHz Per Core

Convert data to EVIO

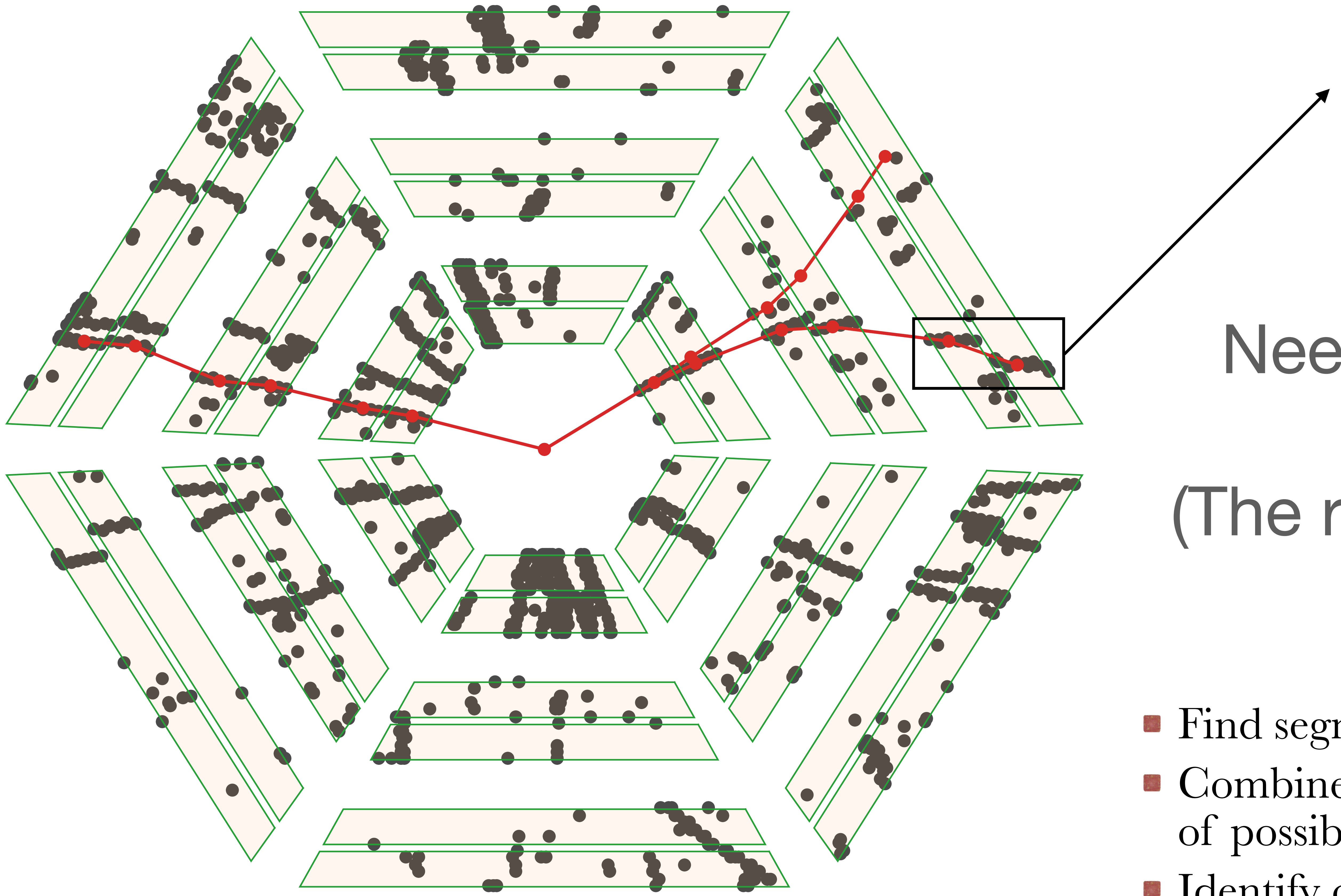
✓ (yellow) Feasibility studies were done

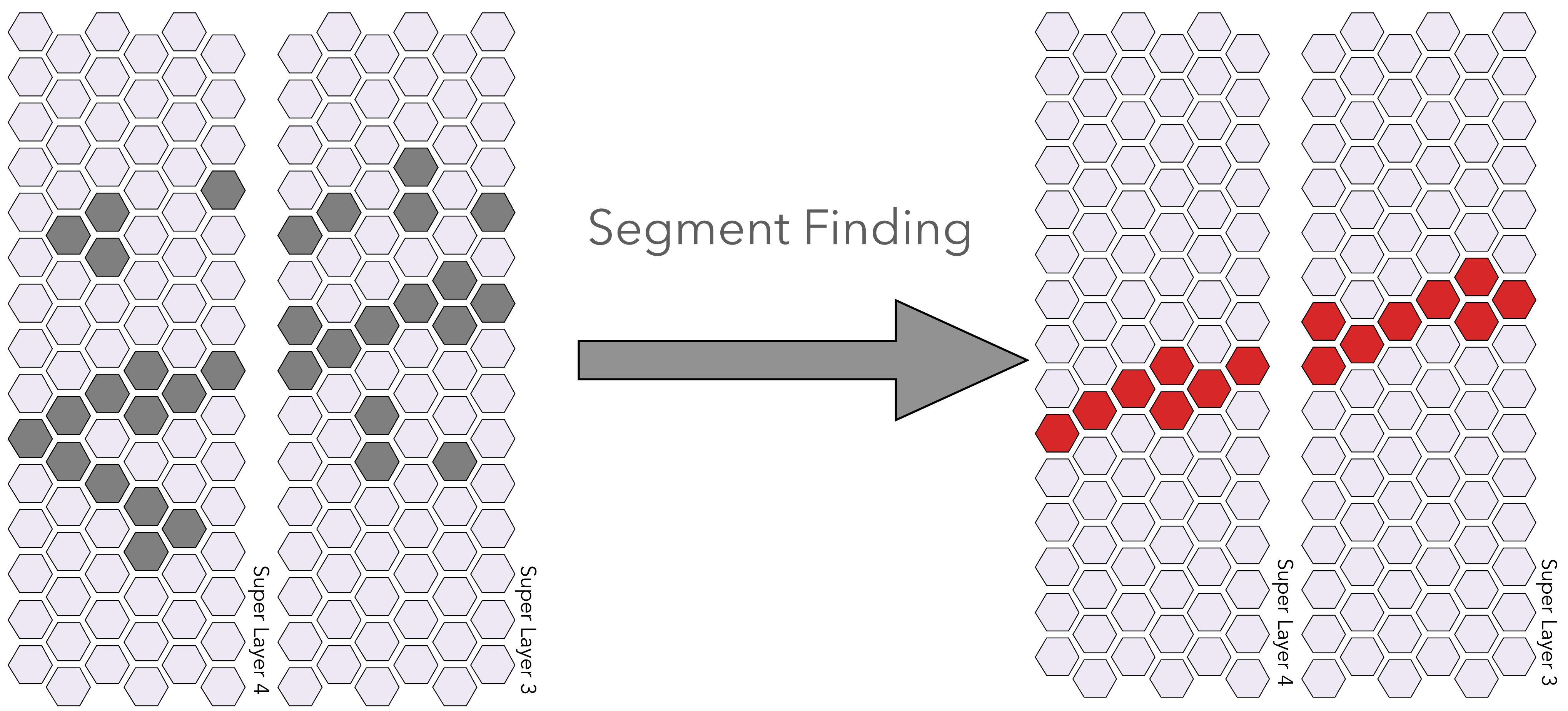
✓ (green) Components Exist

✗ To be developped

Segment Finding (to be developed)

- 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)

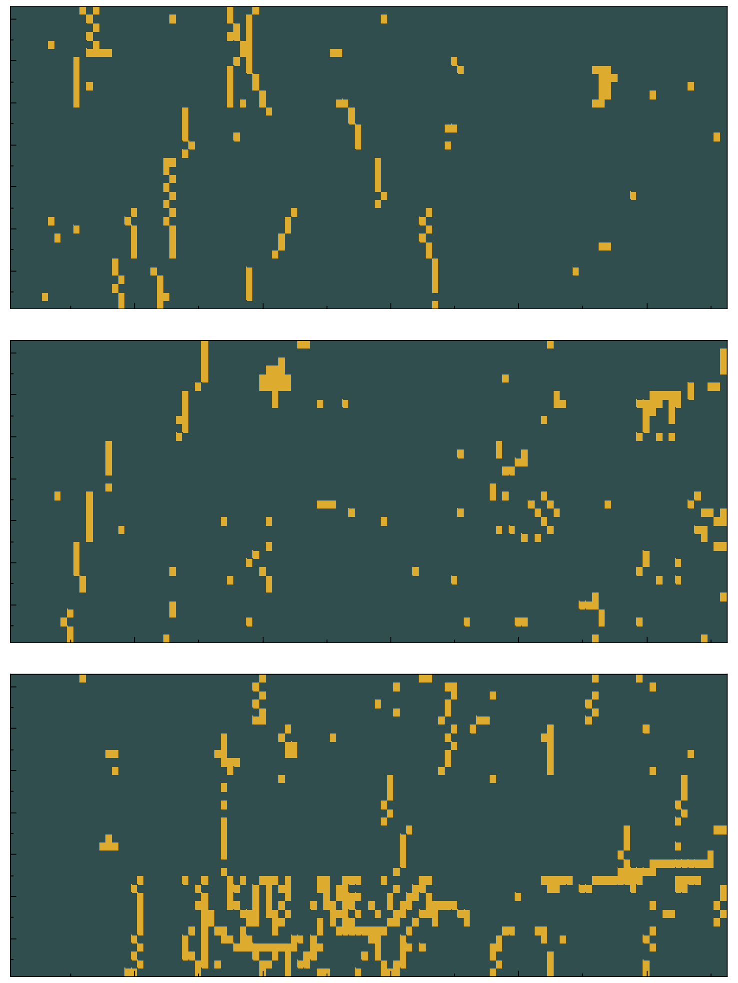


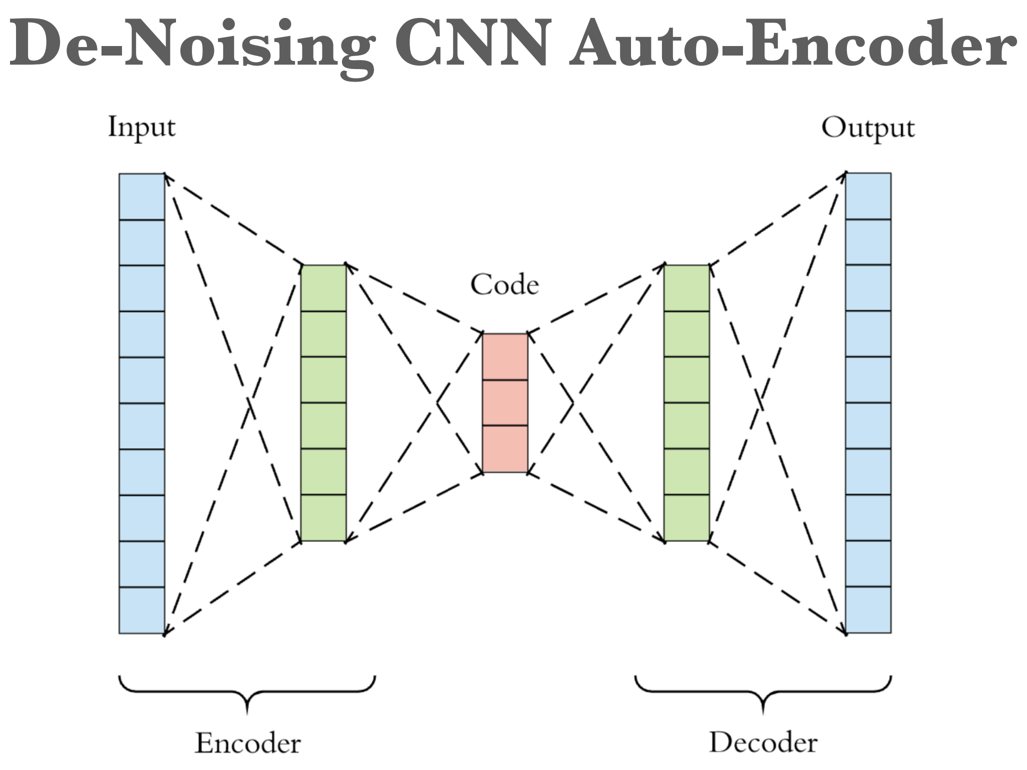


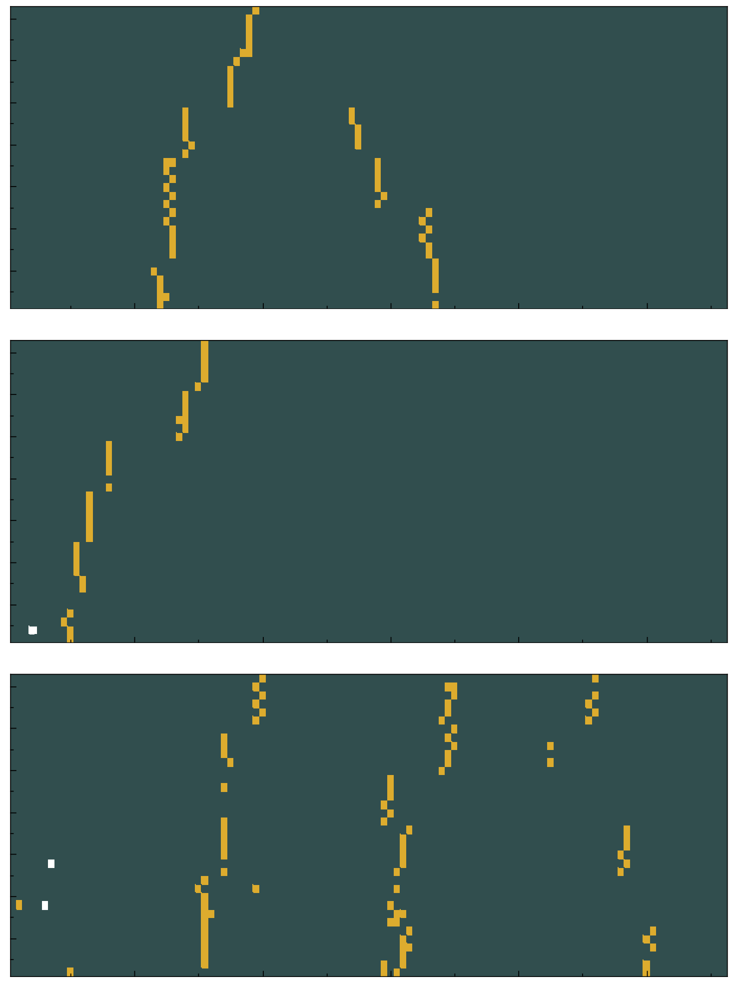
Need to develop fast AI based segment finding algorithm
(The reconstruction segment finder is 6 Hz)
Not suitable for online

- 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

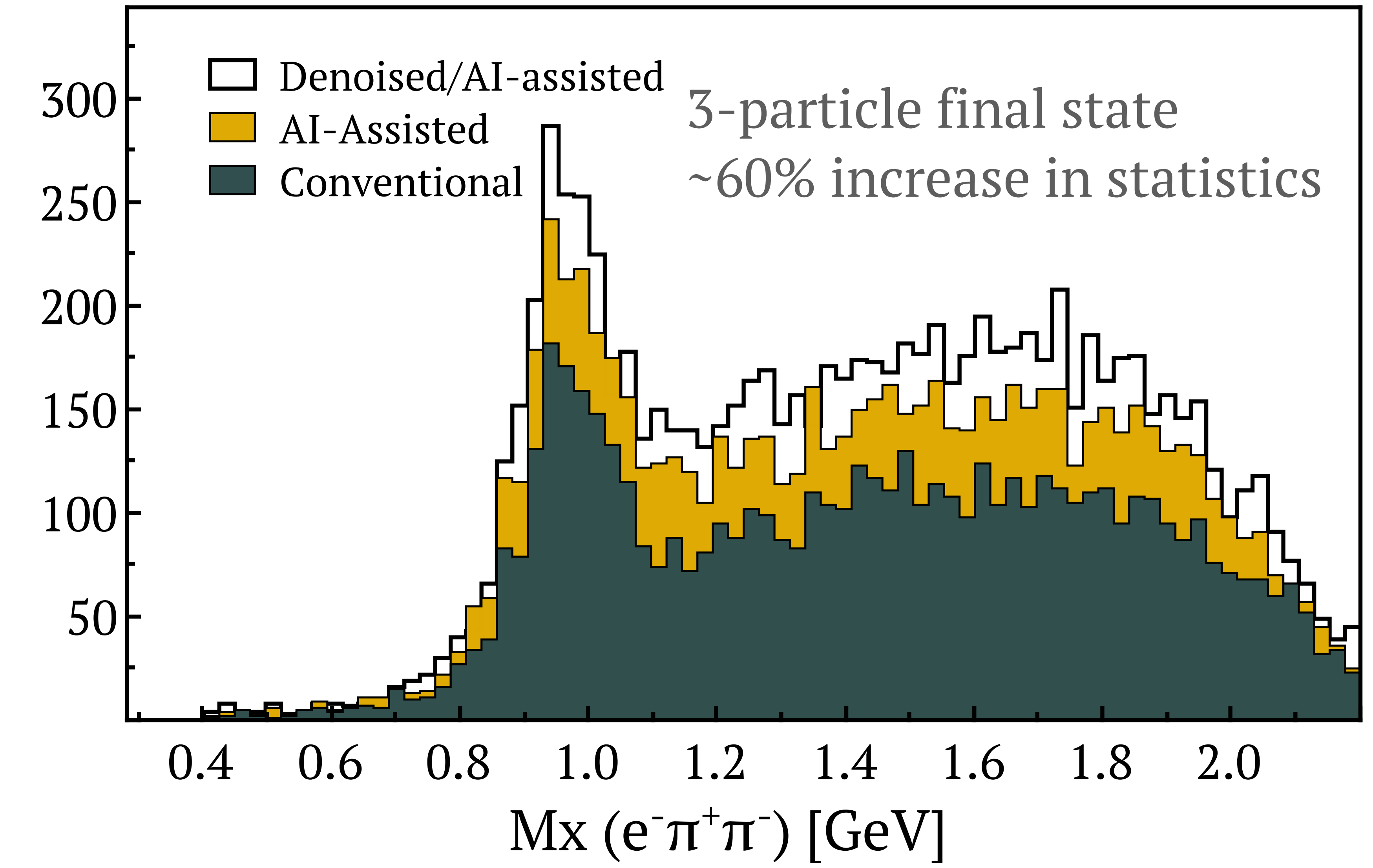




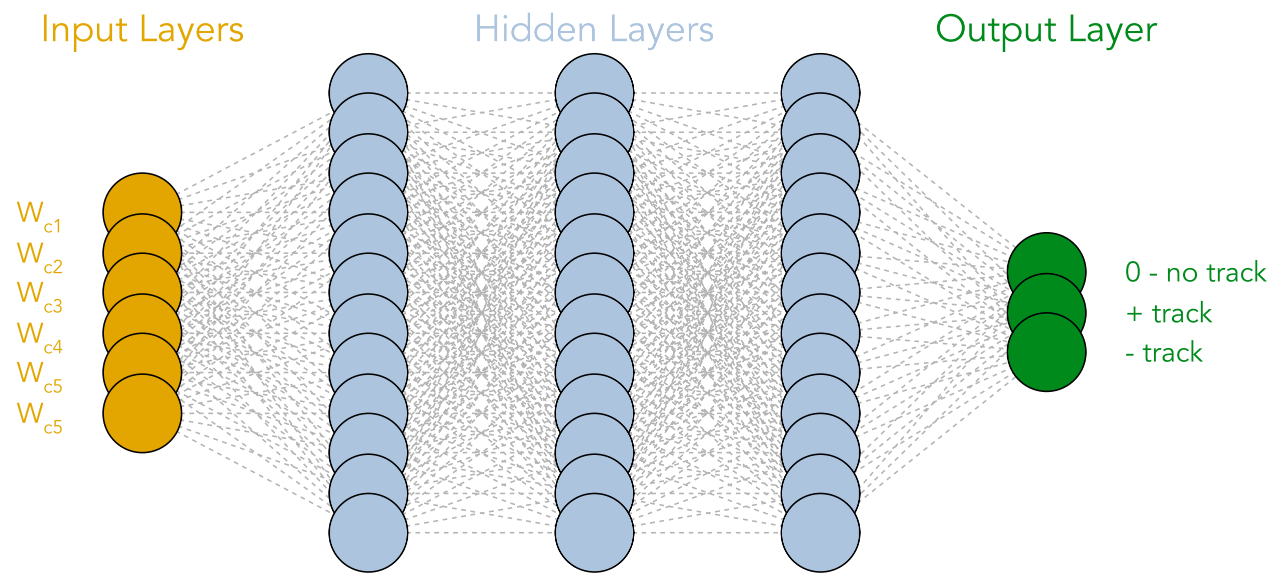


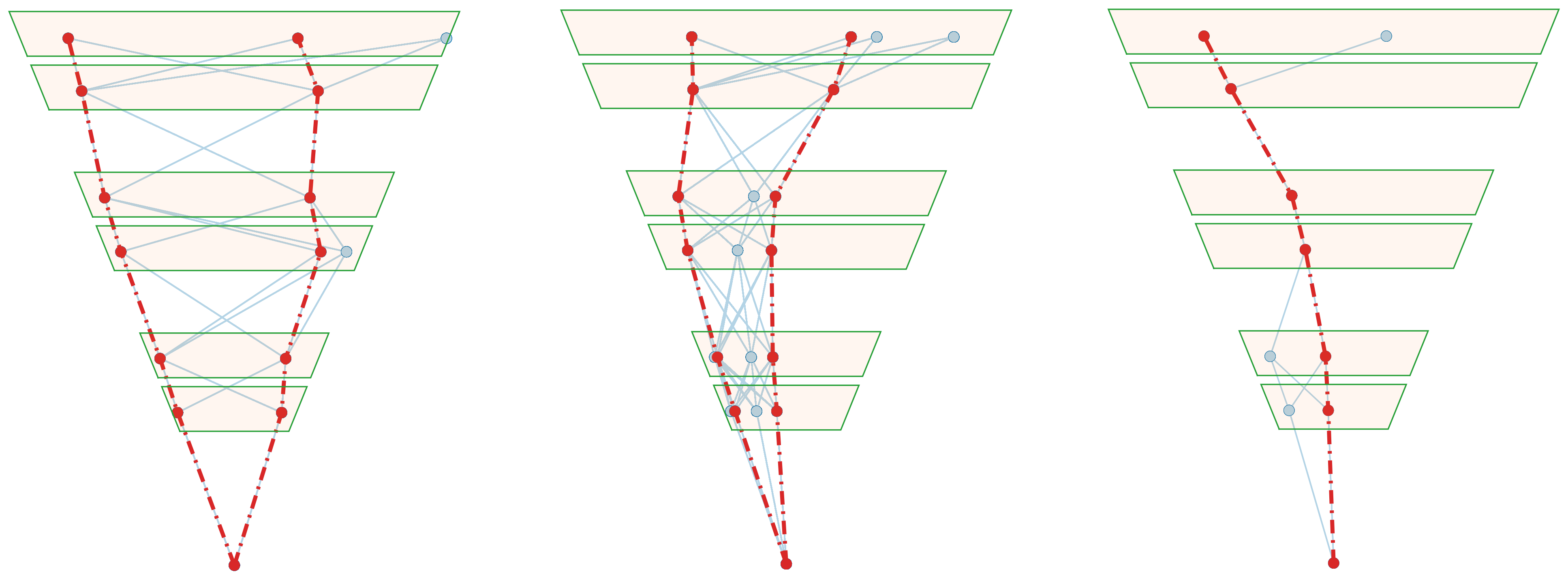
Retains >95% of hits belonging to tracks, and removes >90% of background hits

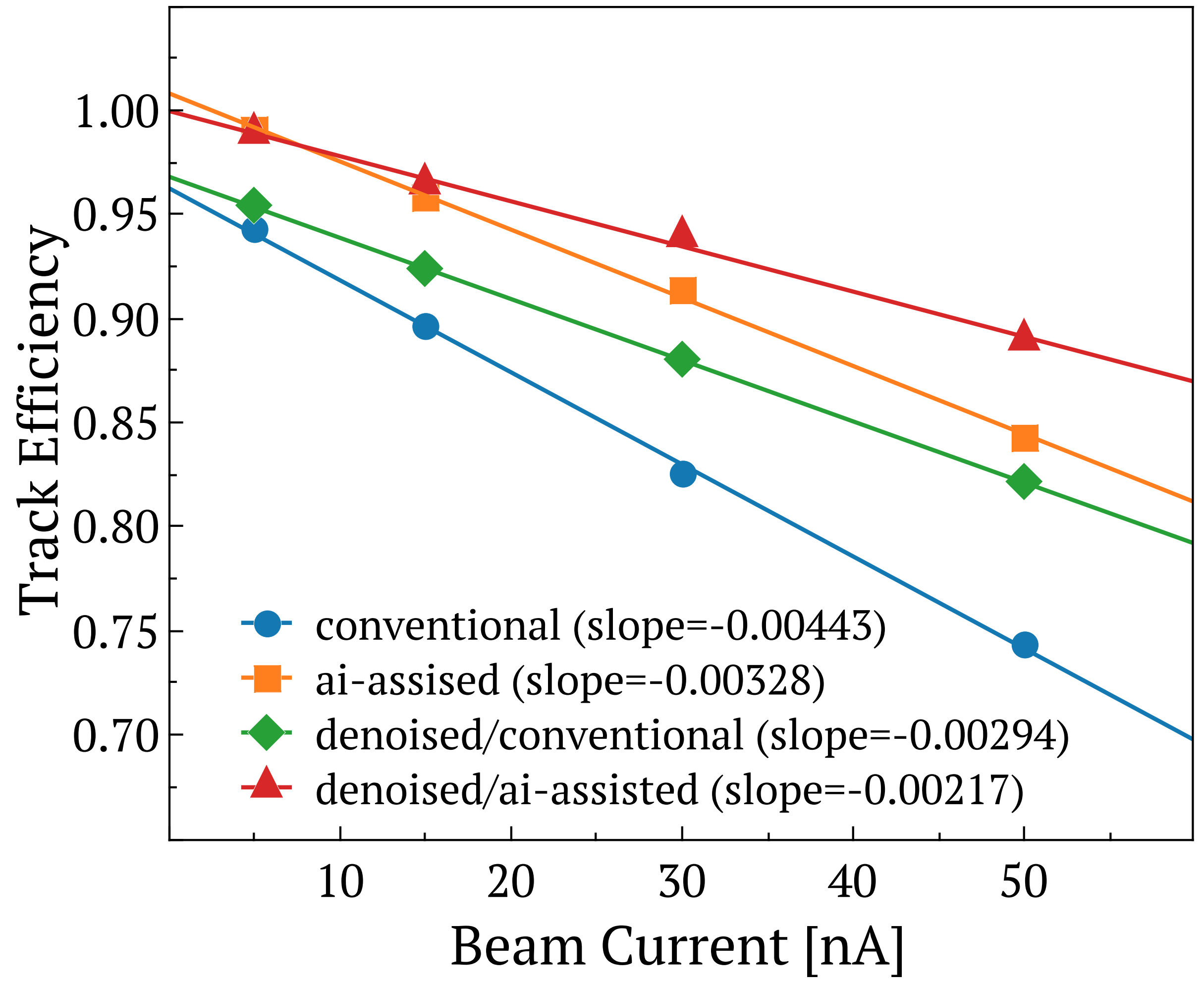
CLAS12 Track Reconstruction with Artificial Intelligence
Gagik Gavalian (Jefferson Lab), Pet all e-Print: 2205.02616 [physics.ins-det]



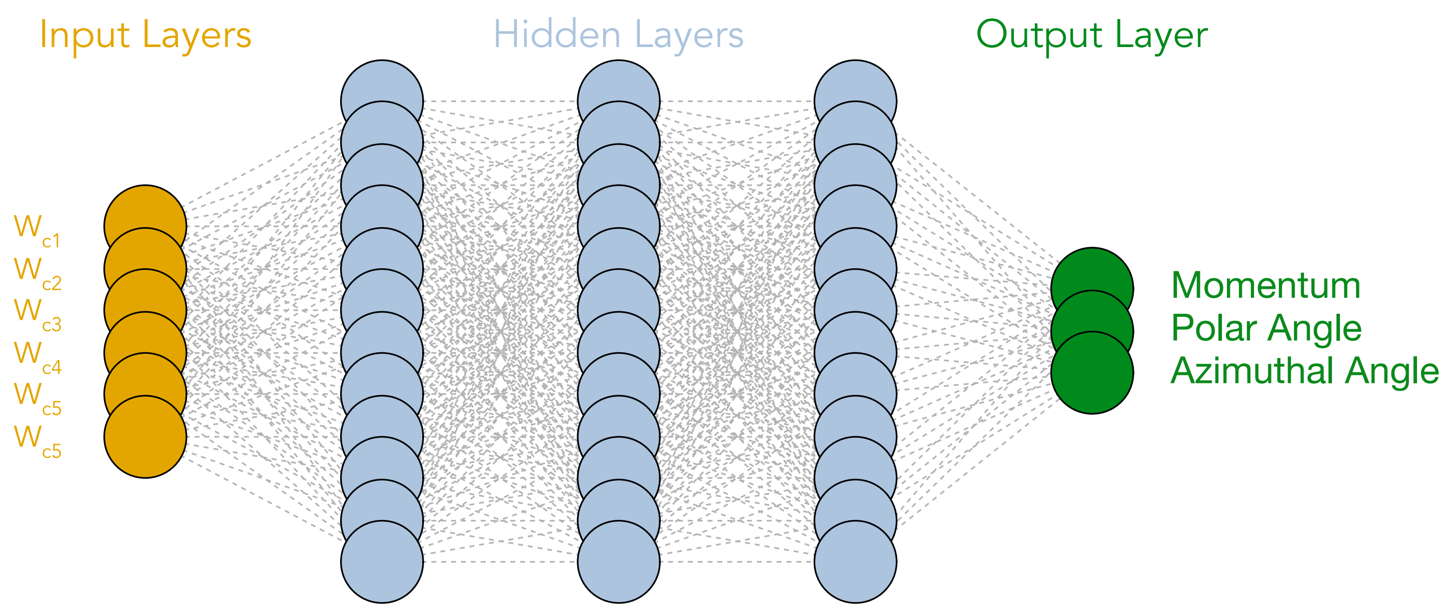
The Classifier network identifies tracks from segment combinations and identifies track charges. The AI-assisted track identification increased tracking efficiency by **15%-21%** (depending on luminosity)
Improvement of the efficiency slope as a function of luminosity.

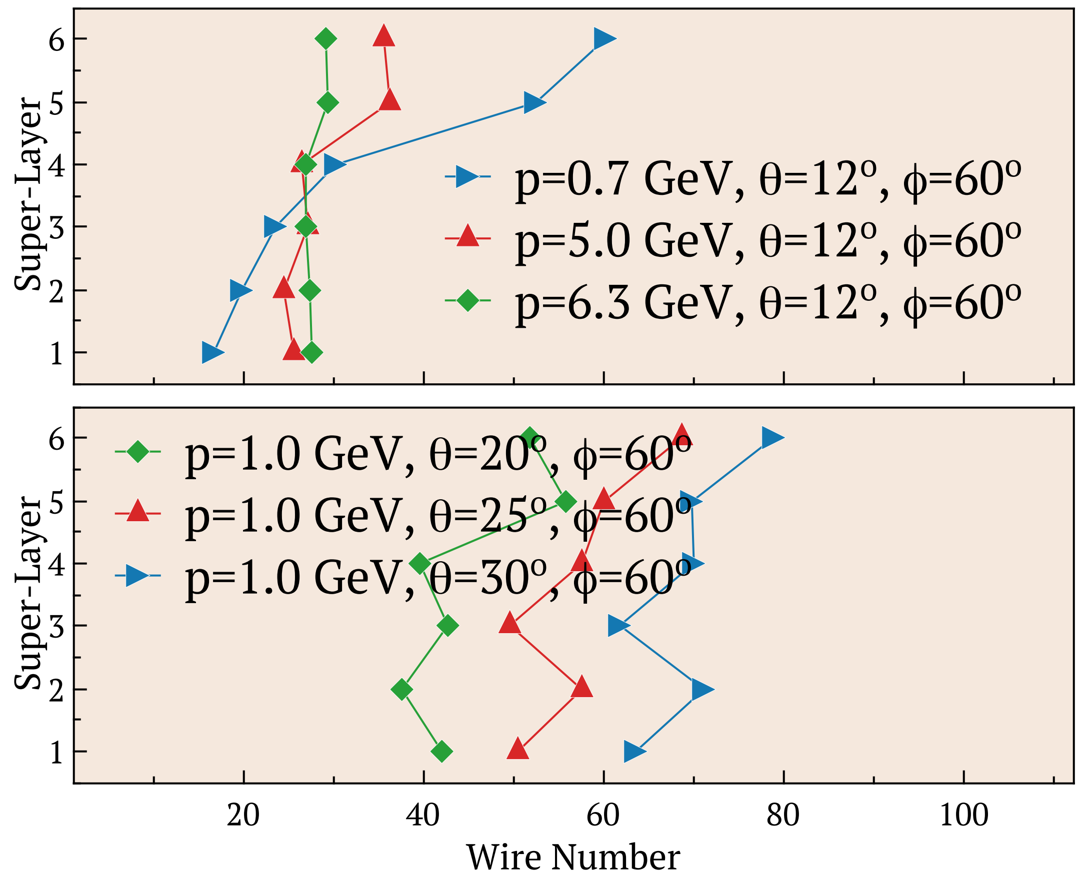






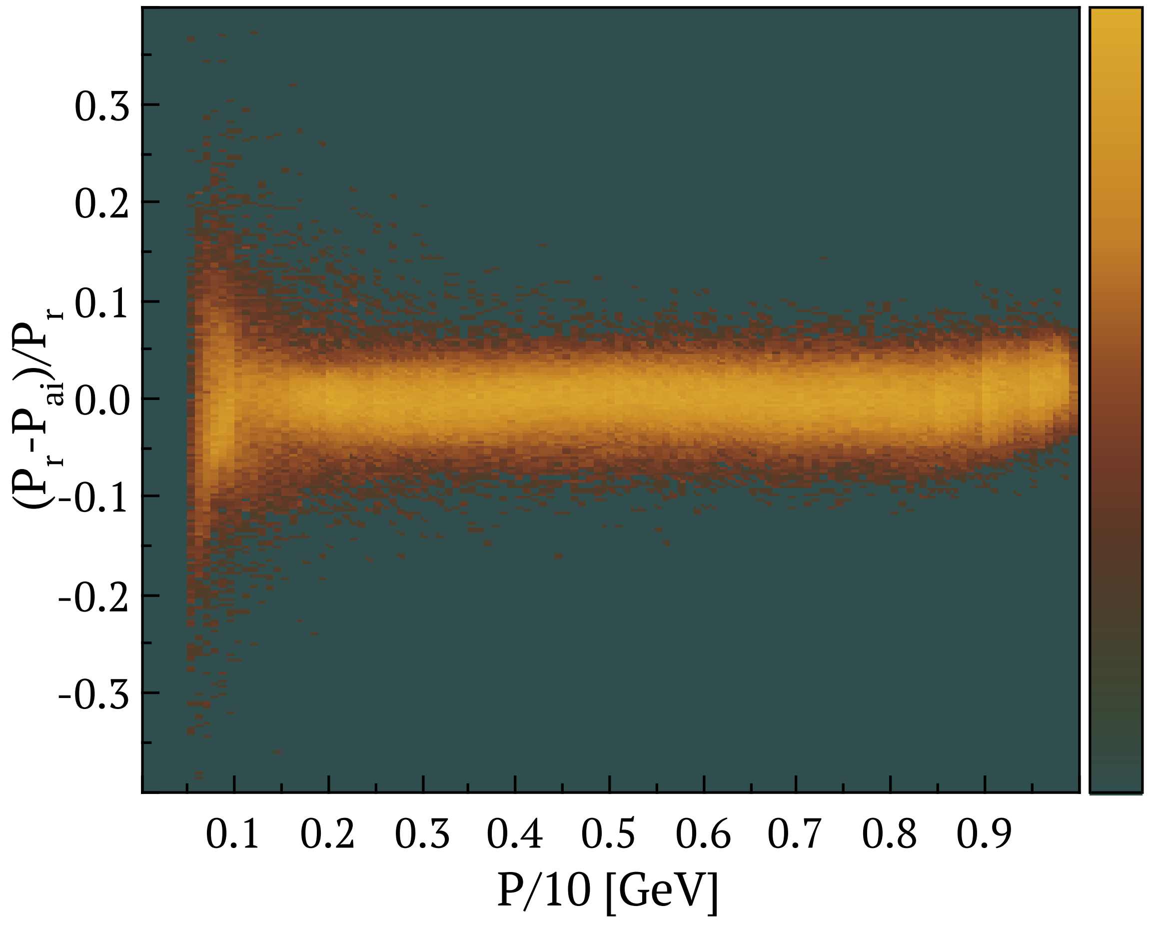
What do we have now

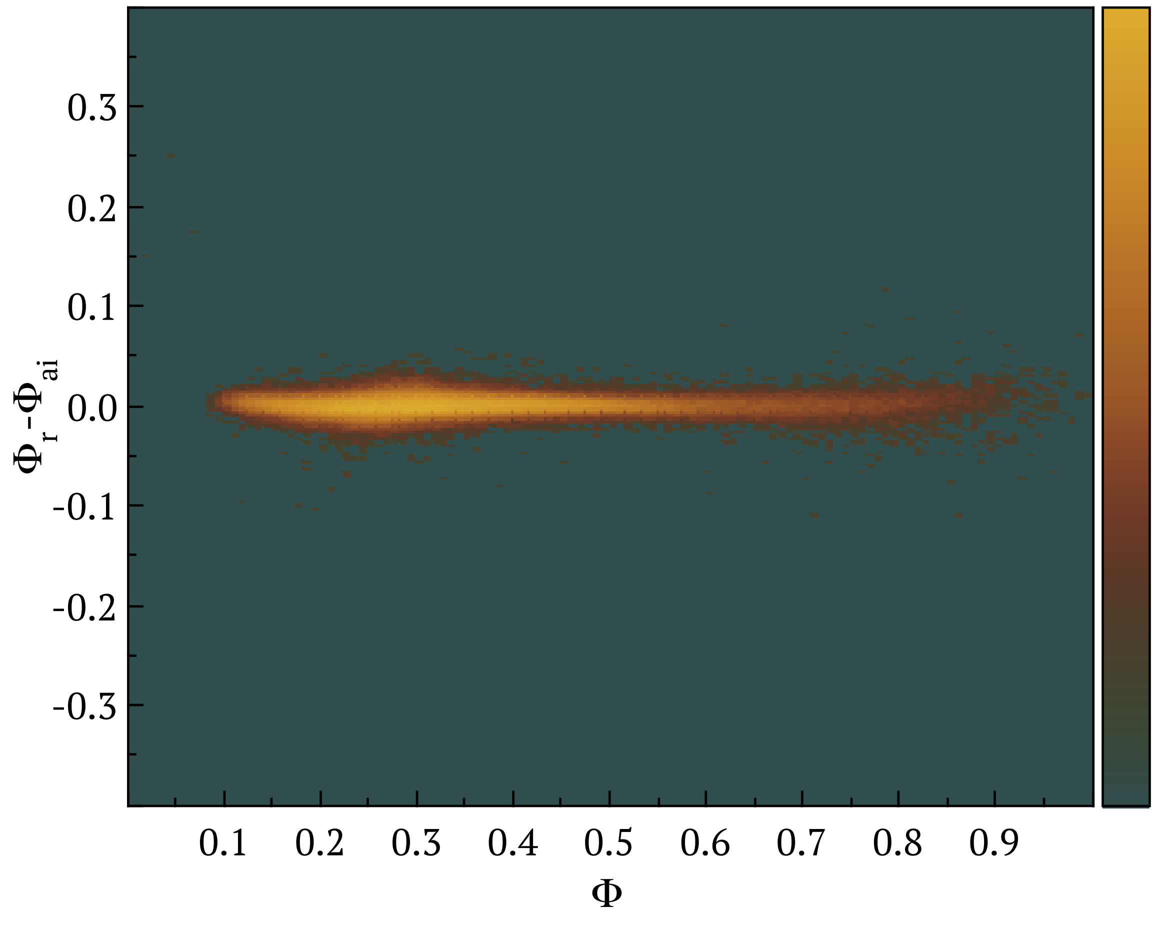


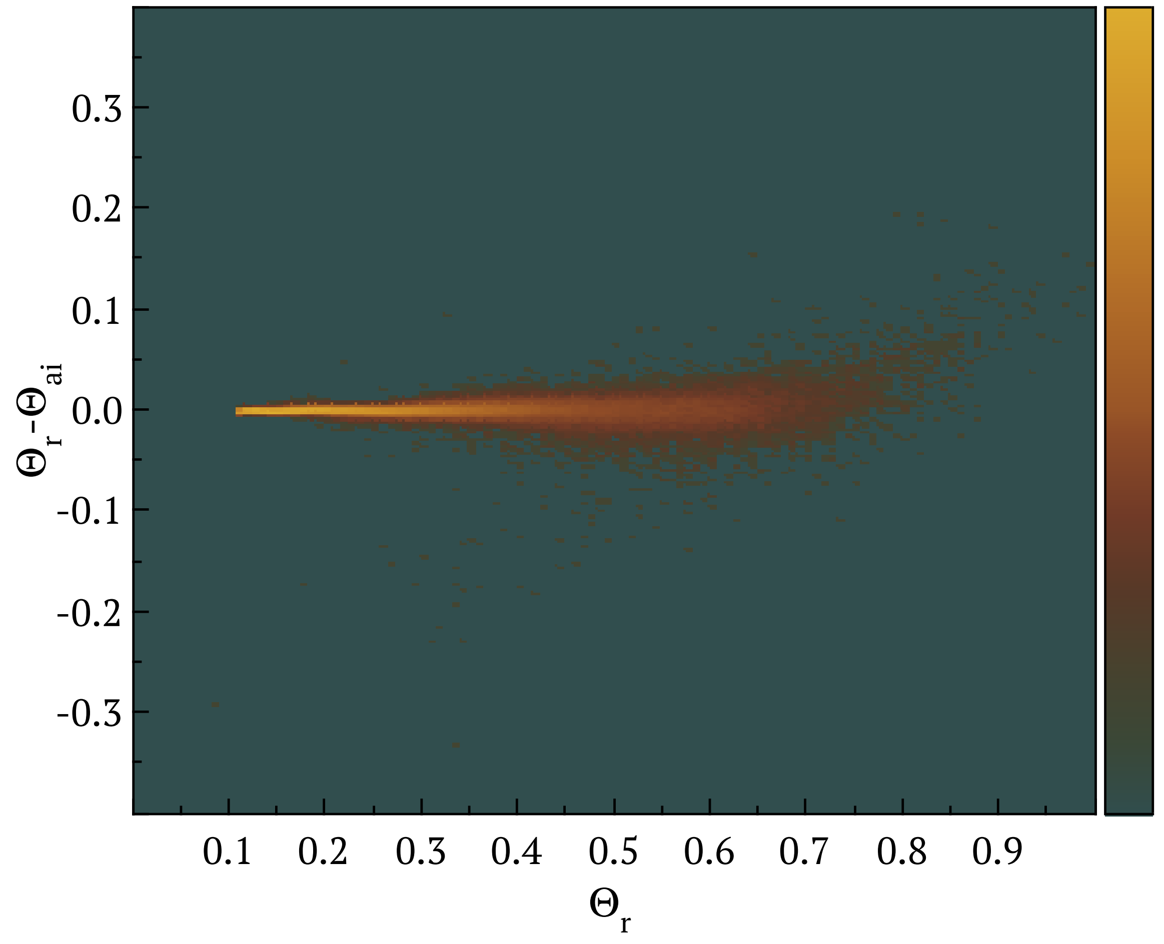


Regression Neural network to predict the track momentum and direction.
The track momentum is reconstructed with an accuracy of 1.4%-1.7%

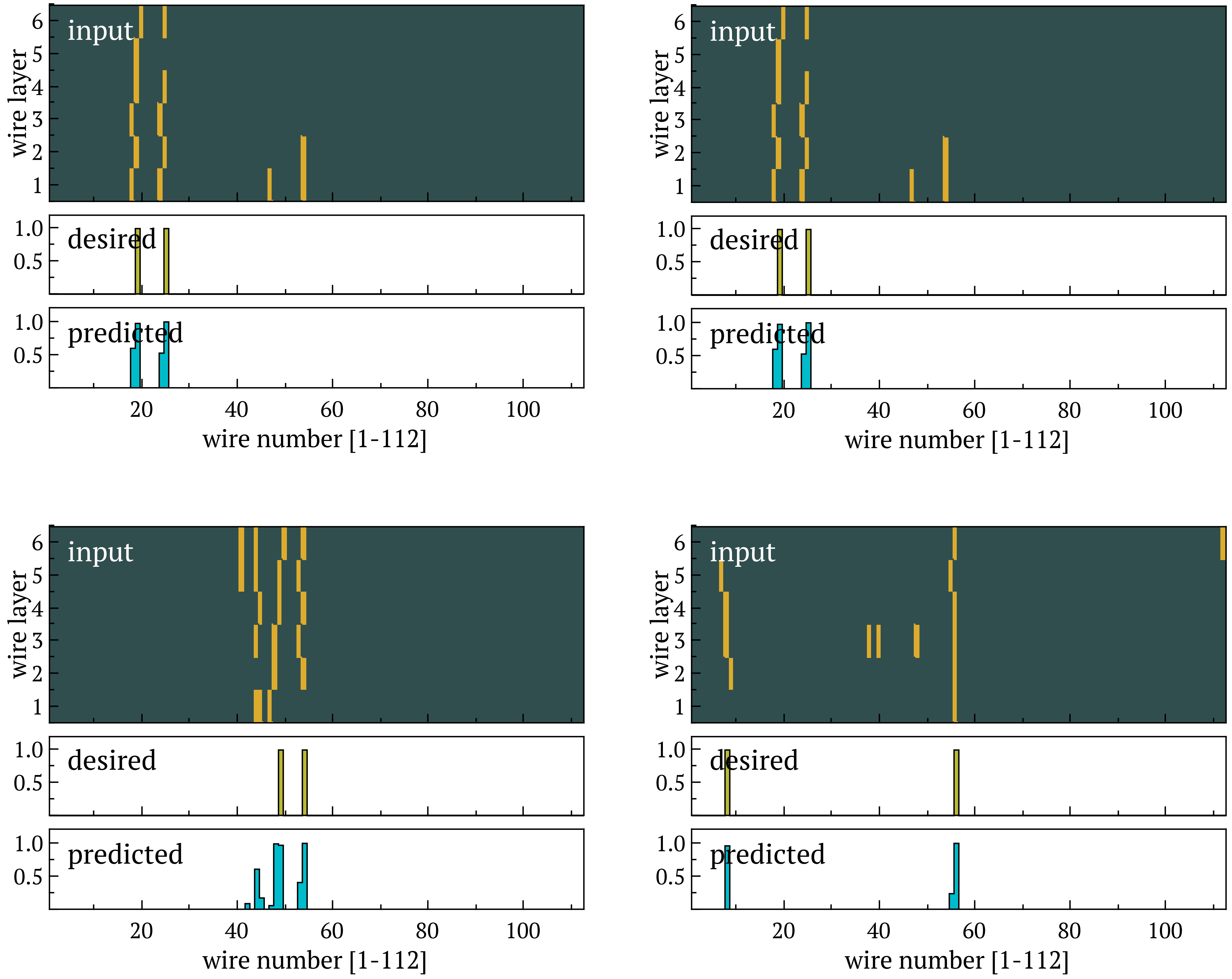
Event topologies can be cleanly identified using particle parameters inferred by the neural network.

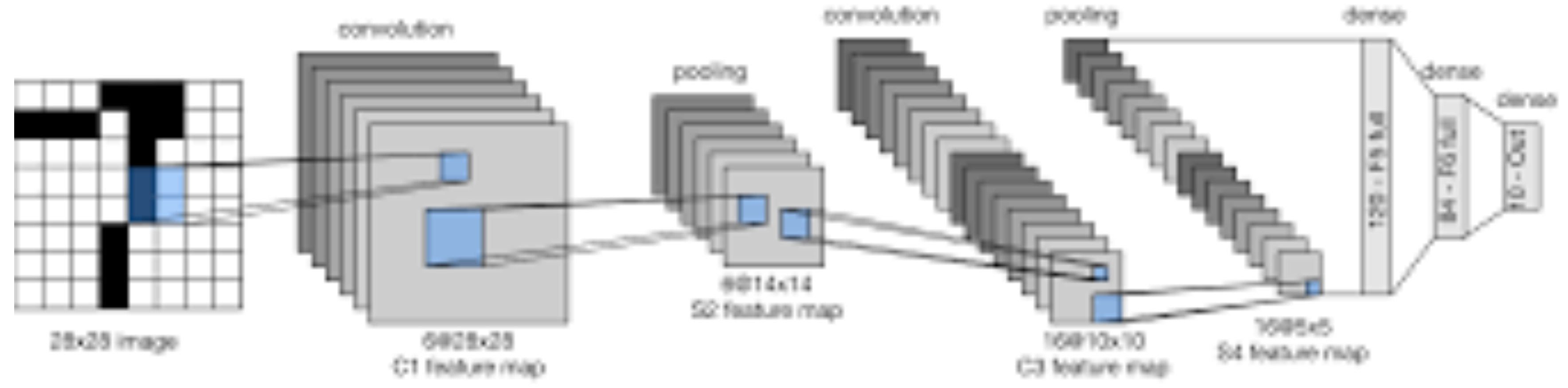






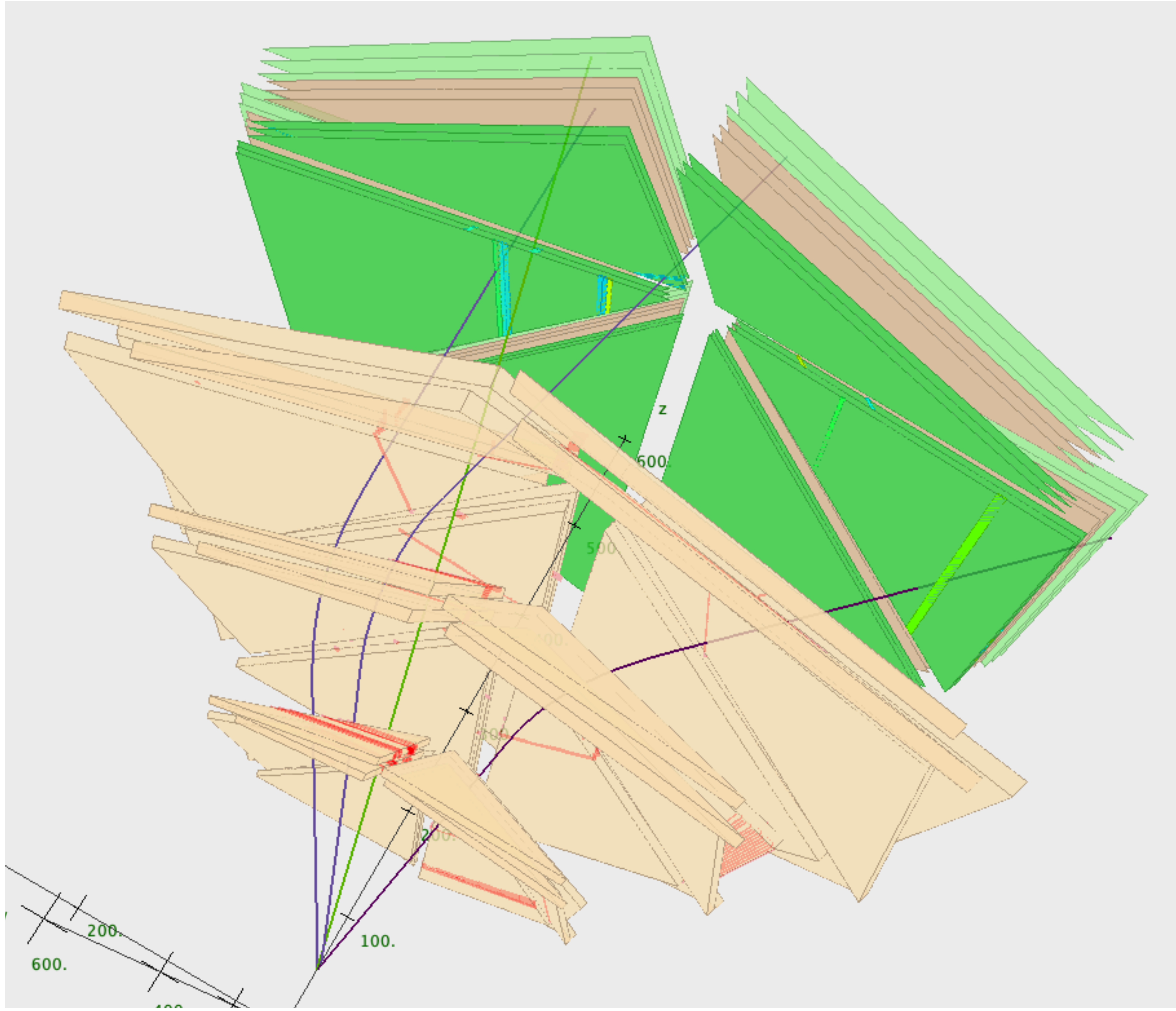
Feasibility tests



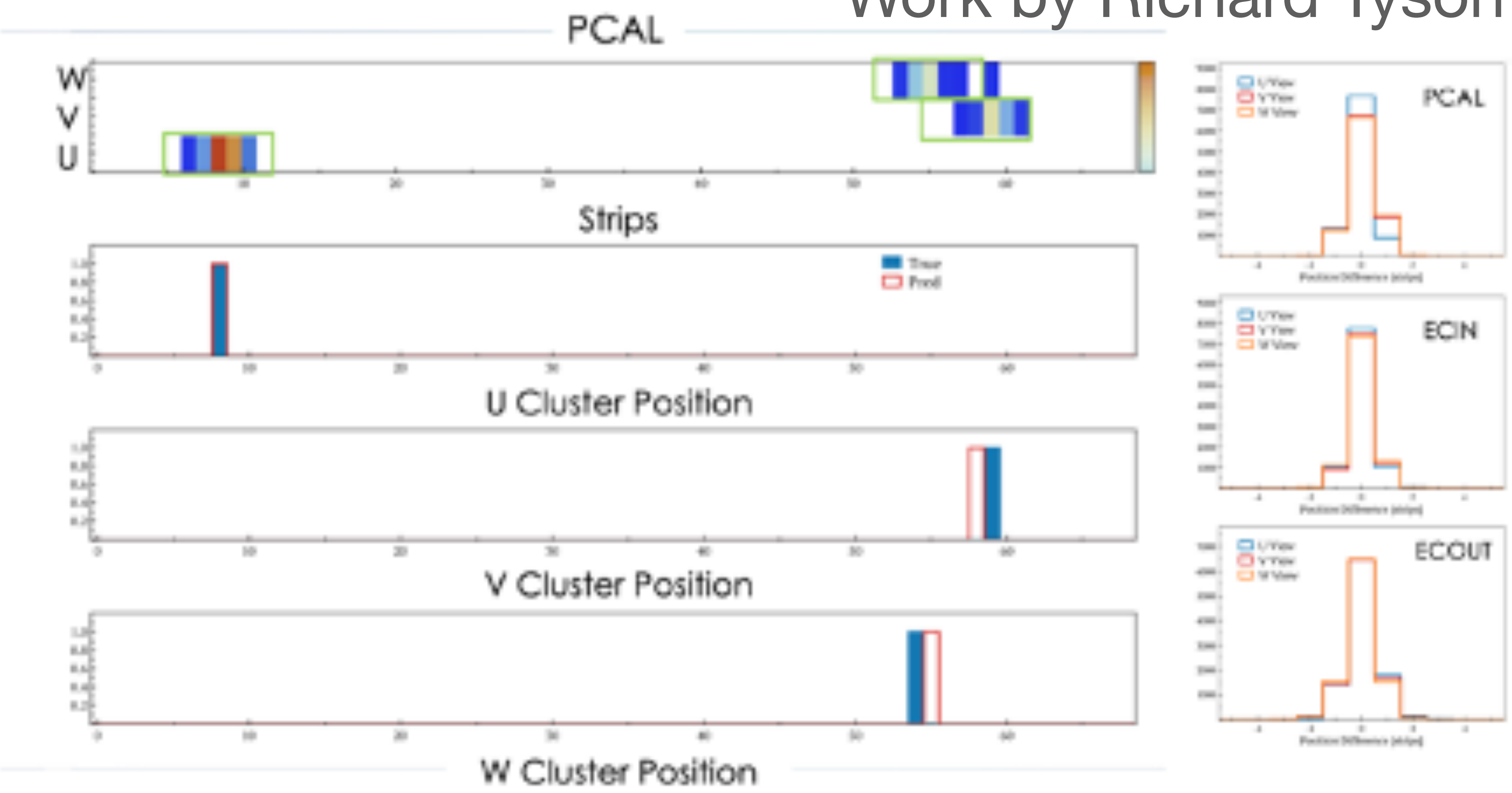


- Clustering Algorithm:
  - Convolutional Neural network with logistic regression to identify possible cluster positions.
  - Currently at 80% efficiency
- Proposed work:
  - Investigate algorithms to identify clusters fast and with higher efficiency

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

- Requested funds:
  - 25% of PI's time to coordinate the developments and implementation full chain of data processing and deployment. Development of data translation online.
  - 0.2 FTE for Hall-B post-doc (Richard Tyson CO-PI) to continue his work on detector matching for tracks and electron identification. Training the new hire with CLAS12 data and reconstruction procedures (1 year only)
  - **New hire postdoc with skills in AI:**
    - Develop clustering algorithm for drift chambers
    - Develop neural network architecture for particle identification from detector hits.
    - Help implement the workflow for analyzing streaming readout data and filtering the events.
    - Continue work on electron identification.
  - 0.3 FTE for student work to help write validation and monitoring software for checking online data selection and filtering efficiency.

Milestones

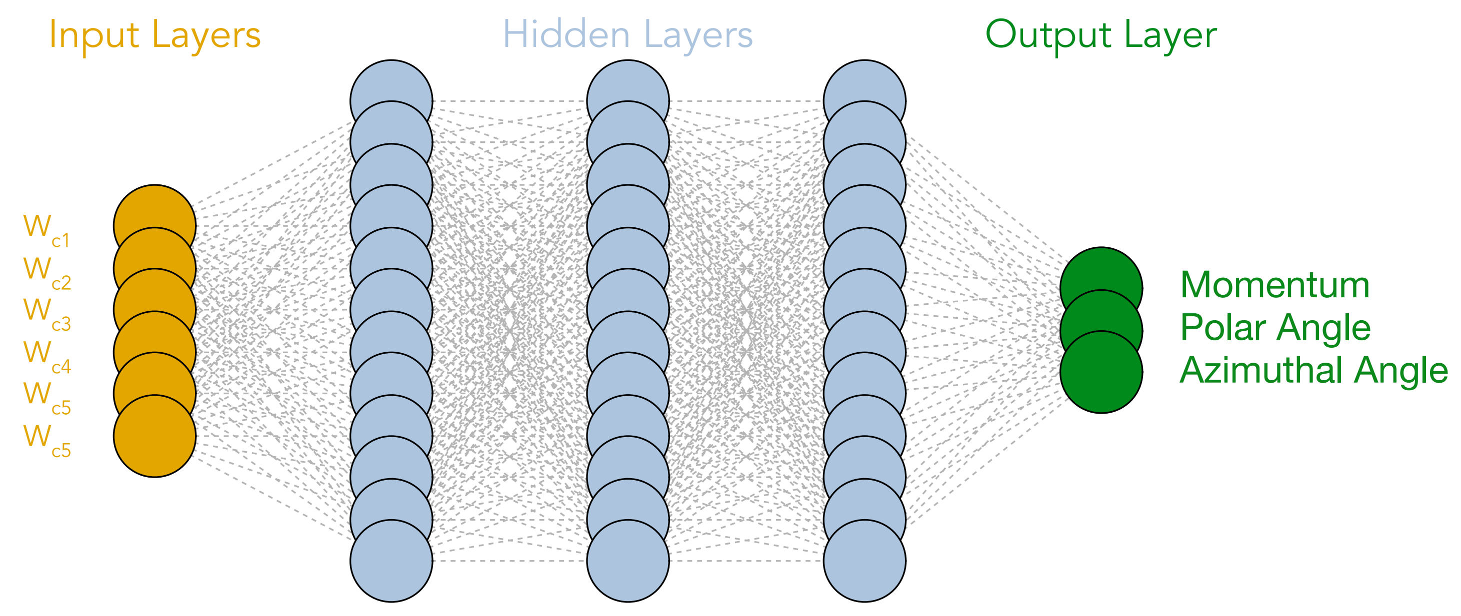
- **First Year (first 6 months):**
  - Develop online multithreaded data converter
  - Develop track segment finder (fast, multithreaded)
- **First Year (second 6 months):**
  - Work on neural networks for track detector matching
  - Implement electron identification neural networks
- **Second Year (third 6 months):**
  - Implement the track finder and electron identification online
  - Validate the results compared to conventional reconstruction
  - Implement SRO data reading and converting software
- **Second Year (fourth 6 months):**
  - Run the developed data reduction software on SRO output (hopefully taken next Spring)
  - Work on publication

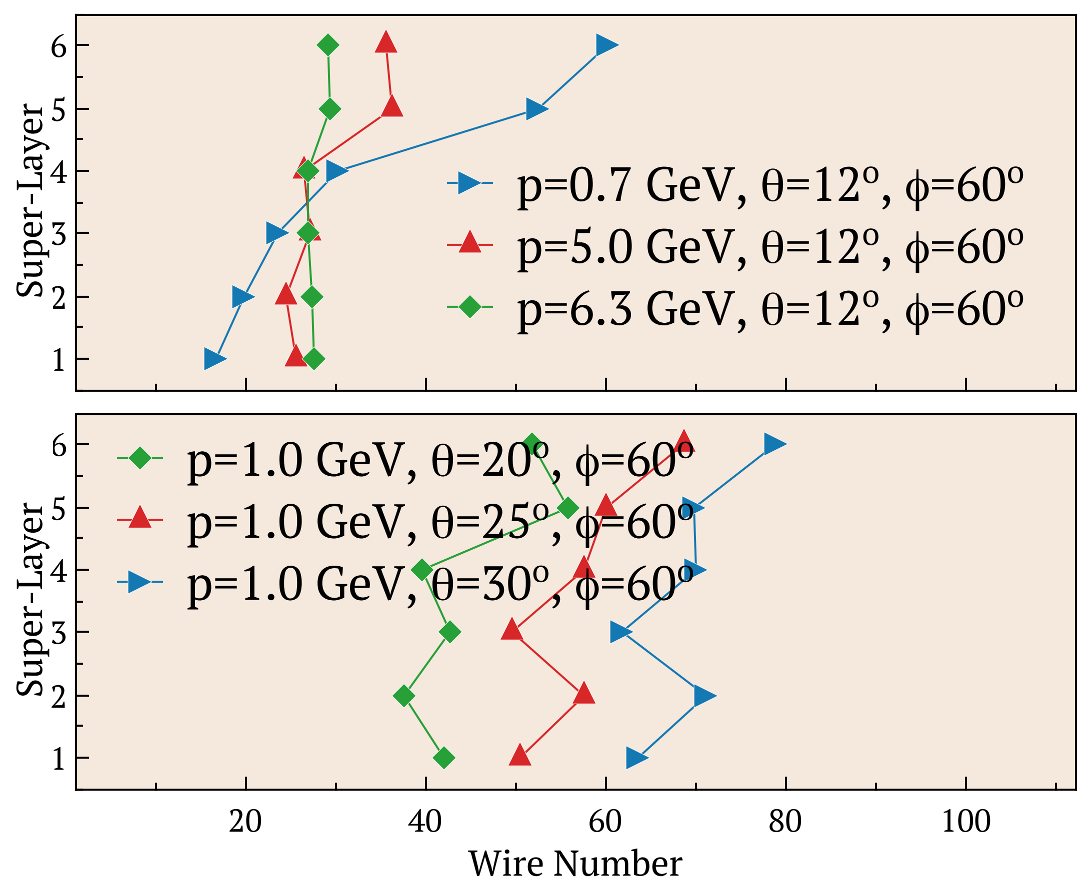
Summary

- **Impact and strategic value to the Laboratory's mission:**
  - Efficient data taking at current and higher luminosities
  - Ability to run at higher luminosities using Streaming Readout (significant reduction of data volume)
  - Much faster turnaround for data processing (from experimental data to physics results)
  - Sharing our experience with other Halls/Experiments
- **Level of Innovation:**
  - Hall-B is the only experiment using AI tracking in production (to the best of my knowledge) already leading to a >60% increase in statistics for physics observables.
  - Using AI for data reduction and event topology identification in real-time is a first at the lab.
- **Using State of the art Neural Networks:**
  - The choice of the neural network architectures in this work is based on the requirement of inference speed.
  - A more modern and powerful network will not be able to process data at a rate of 12 kHz/core (needed for streaming readout applications)

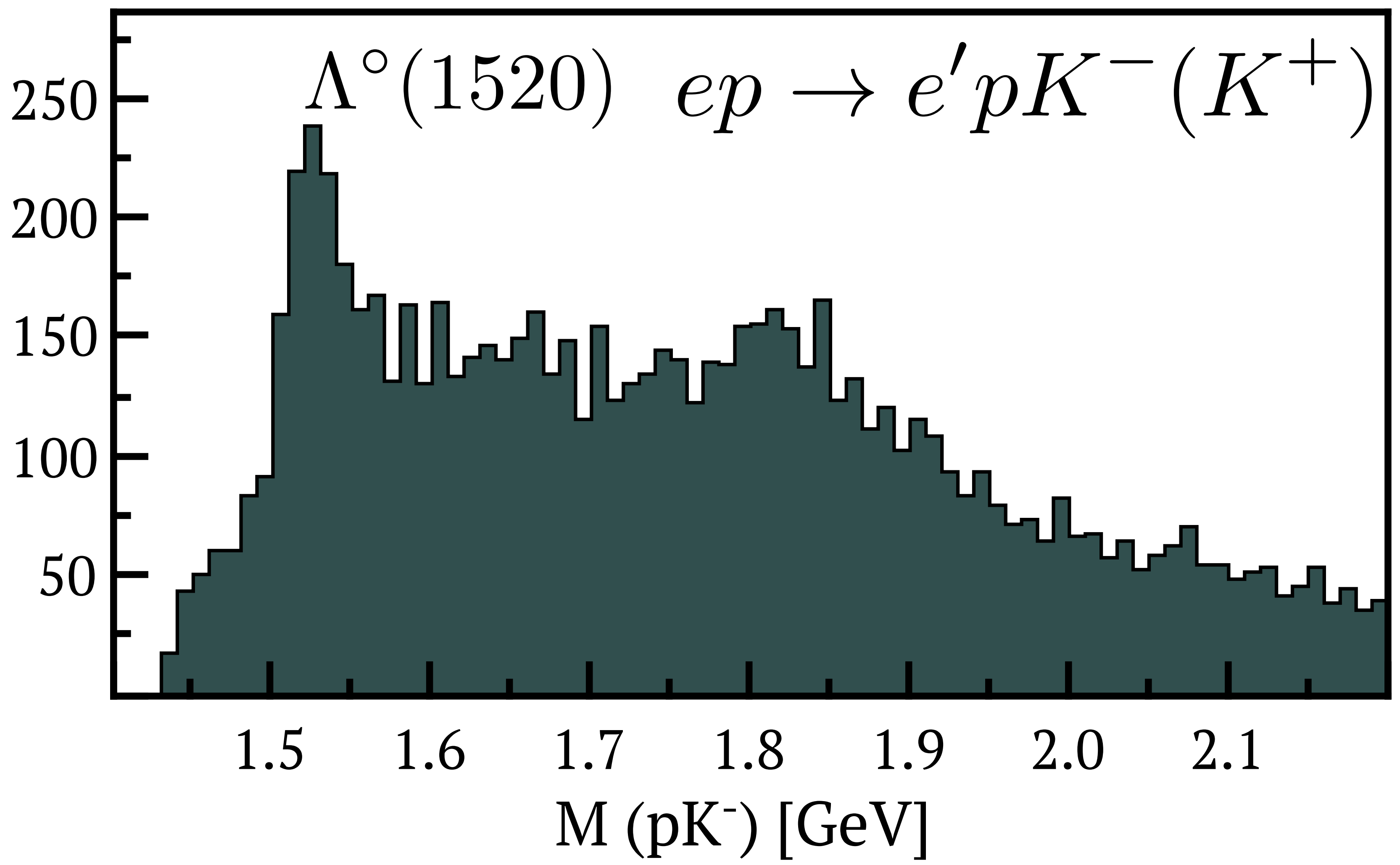
BACK UP SLIDES

What do we have now.





Distributions calculated from track reconstruction from RAW Drift Chamber hits
Inference speed 96 kHz on a laptop (MacBook M3)



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.