SoLID Tracking

11/09/2021

Weizhi Xiong

Syracuse University

Outline

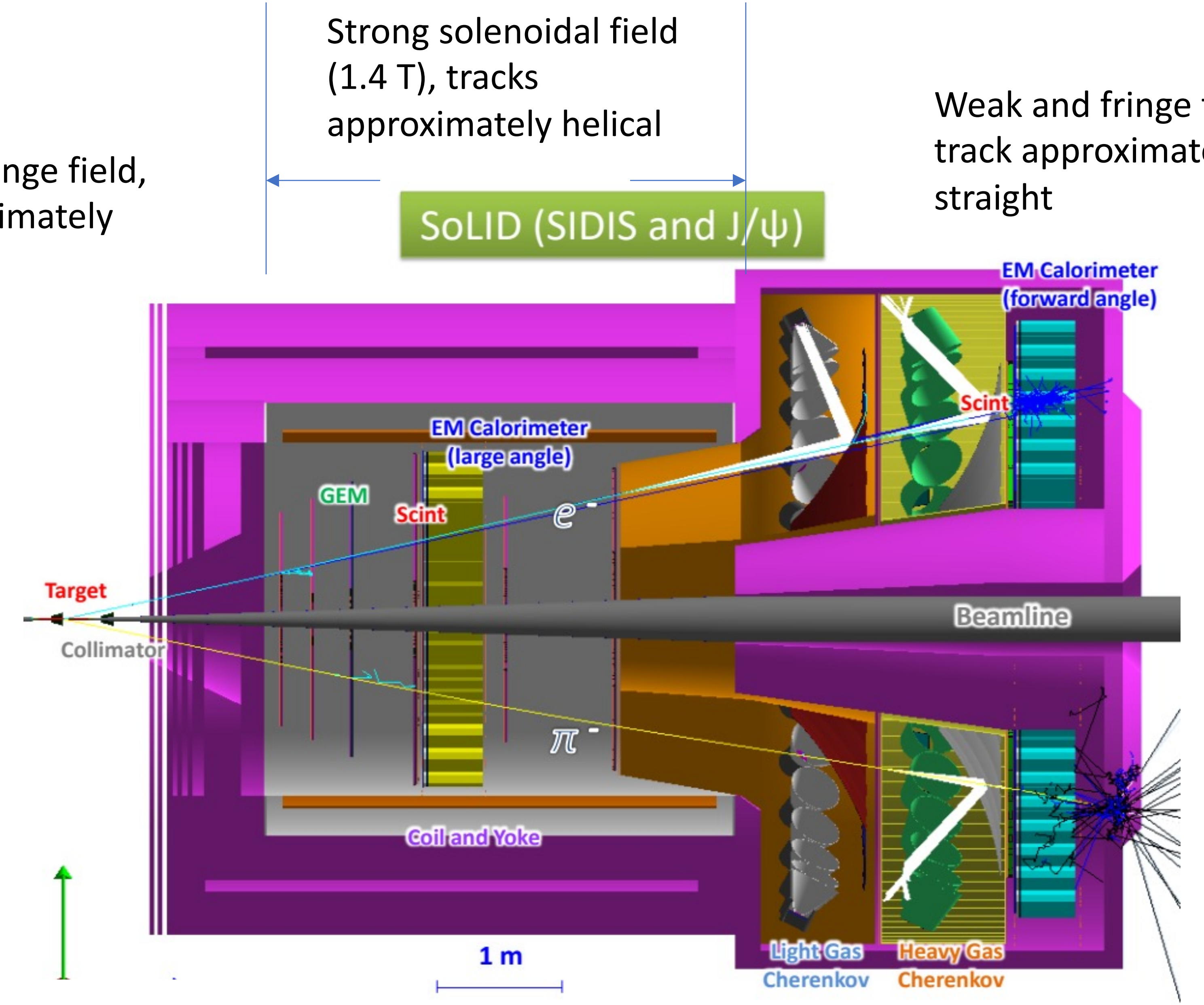
1. SoLID apparatus
2. GEM clustering
3. Tracking reconstruction
4. Current results
5. Summary

SoLID Apparatus

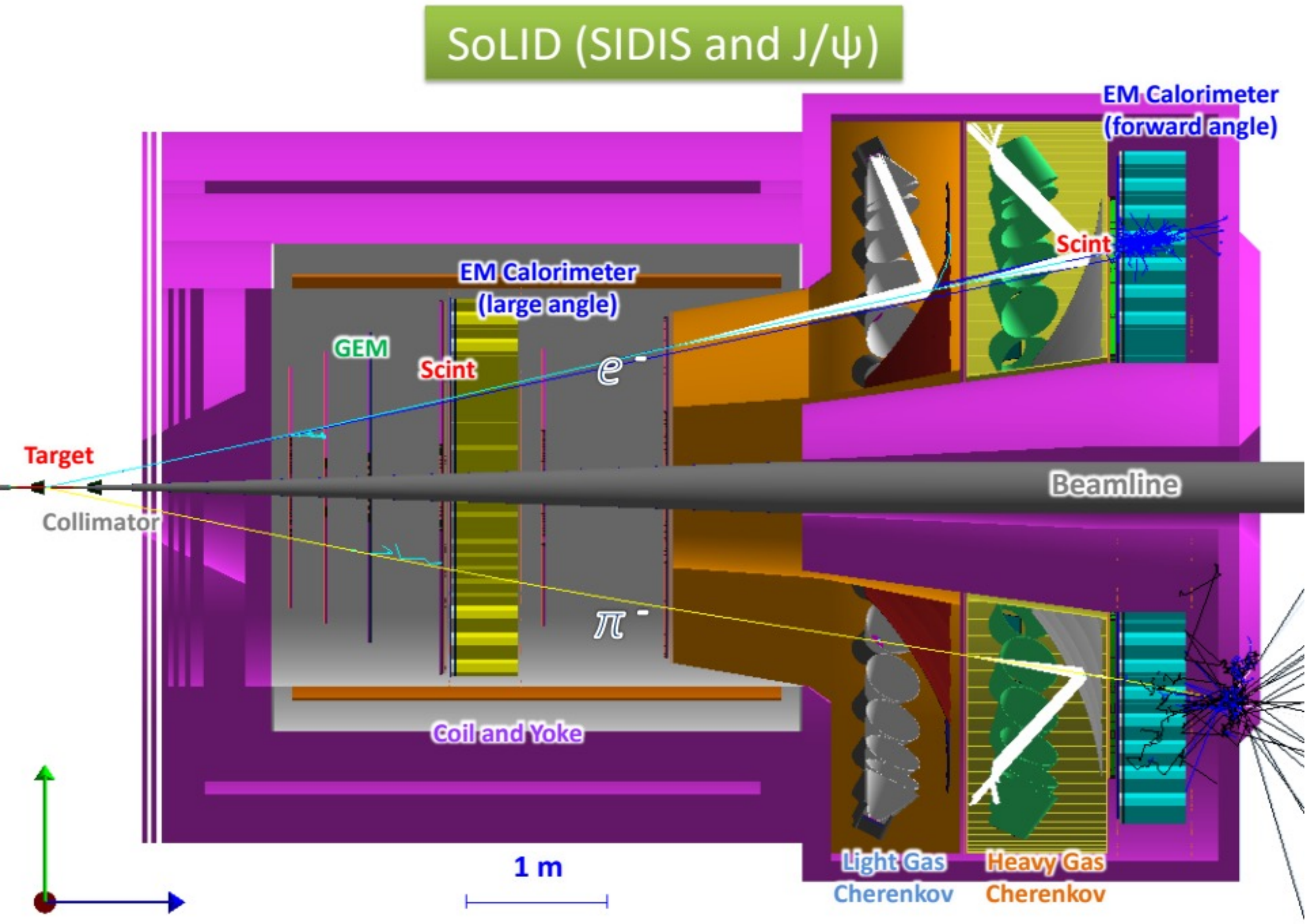
Strong solenoidal field (1.4 T), tracks approximately helical

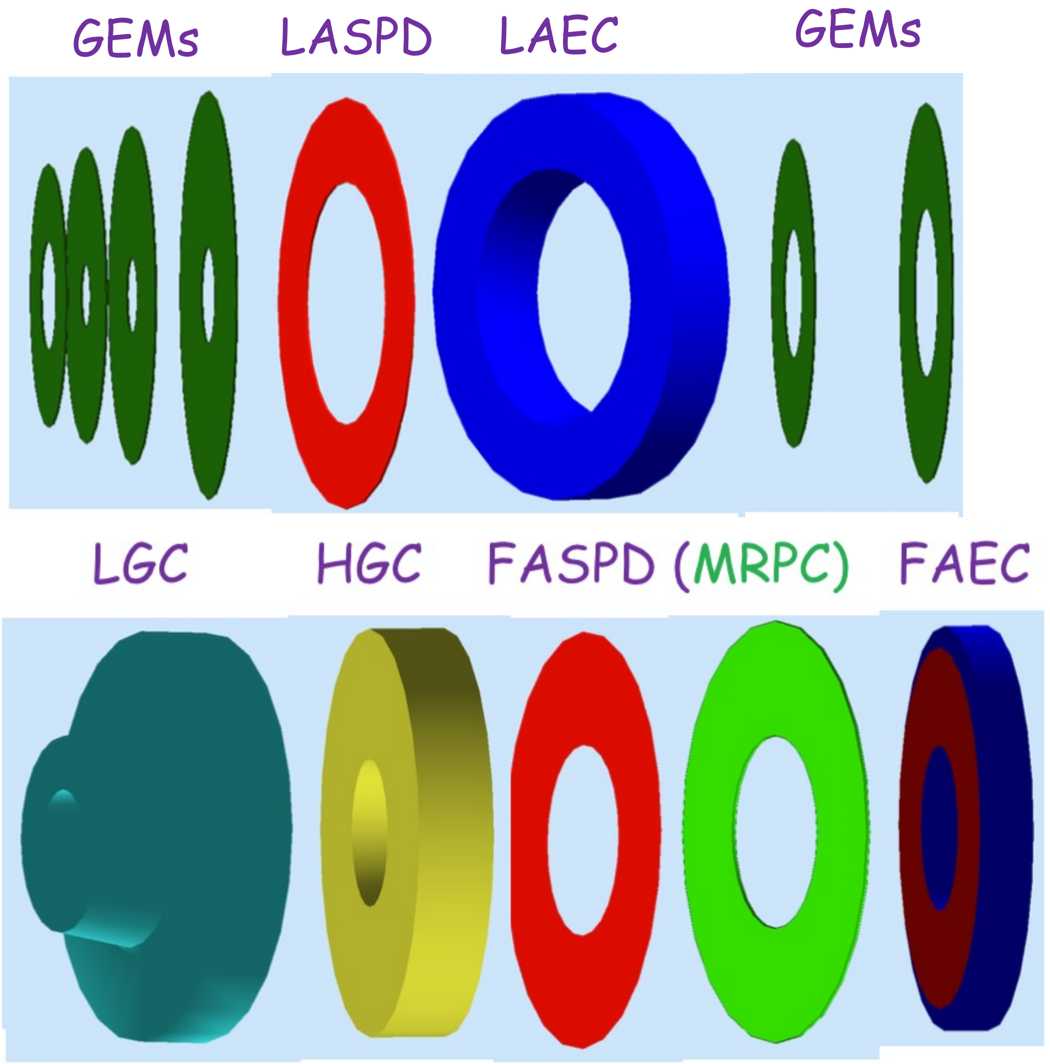
Weak and fringe field, track approximately straight

Weak and fringe field, track approximately straight



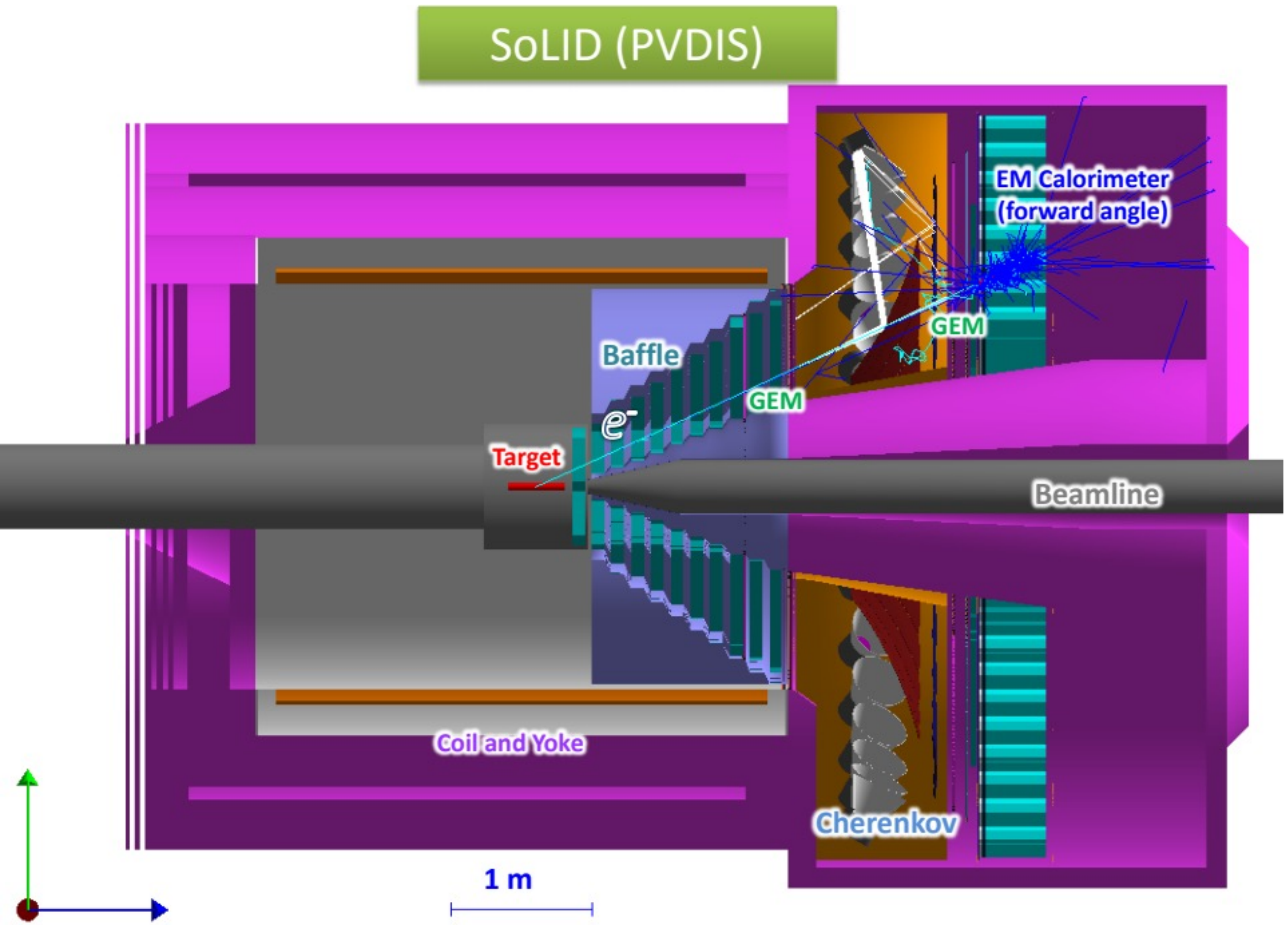
SoLID Apparatus – SIDIS and J/ψ Configuration

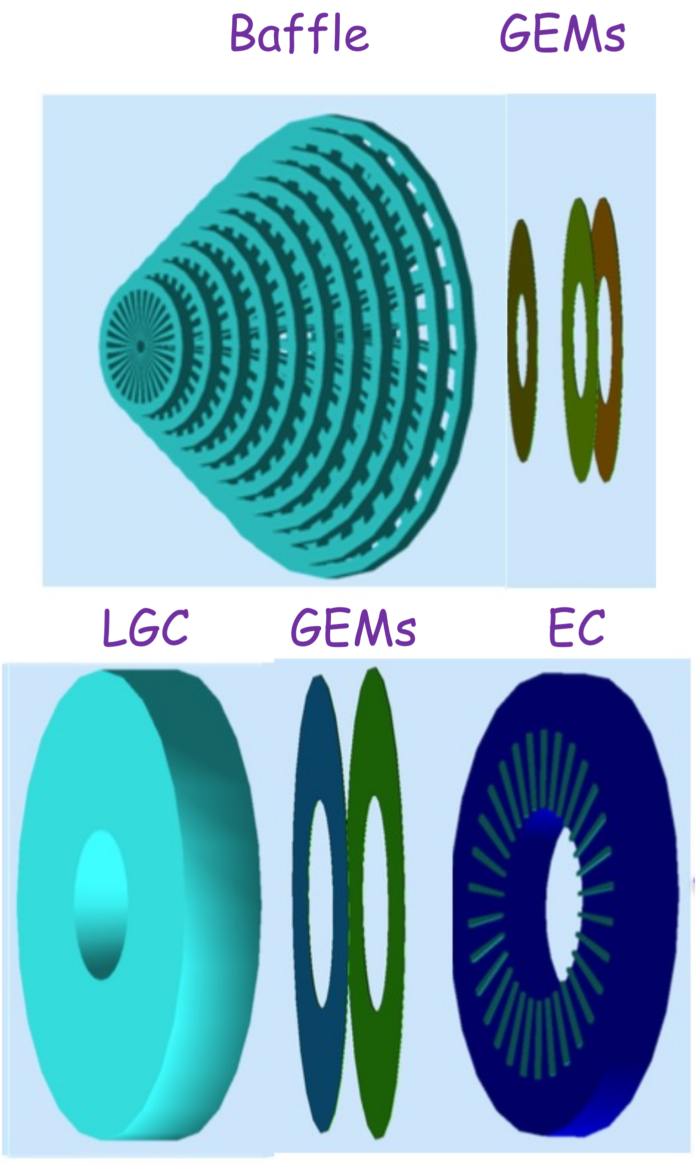




- SIDIS detects both electron and pion in the forward angle region, only electrons in the large angle region
- Jpsi detects electron, positron and proton in both forward and large angle region

SoLID Apparatus – PVDIS Configuration





- PVDIS only wants to detect the scattered electrons

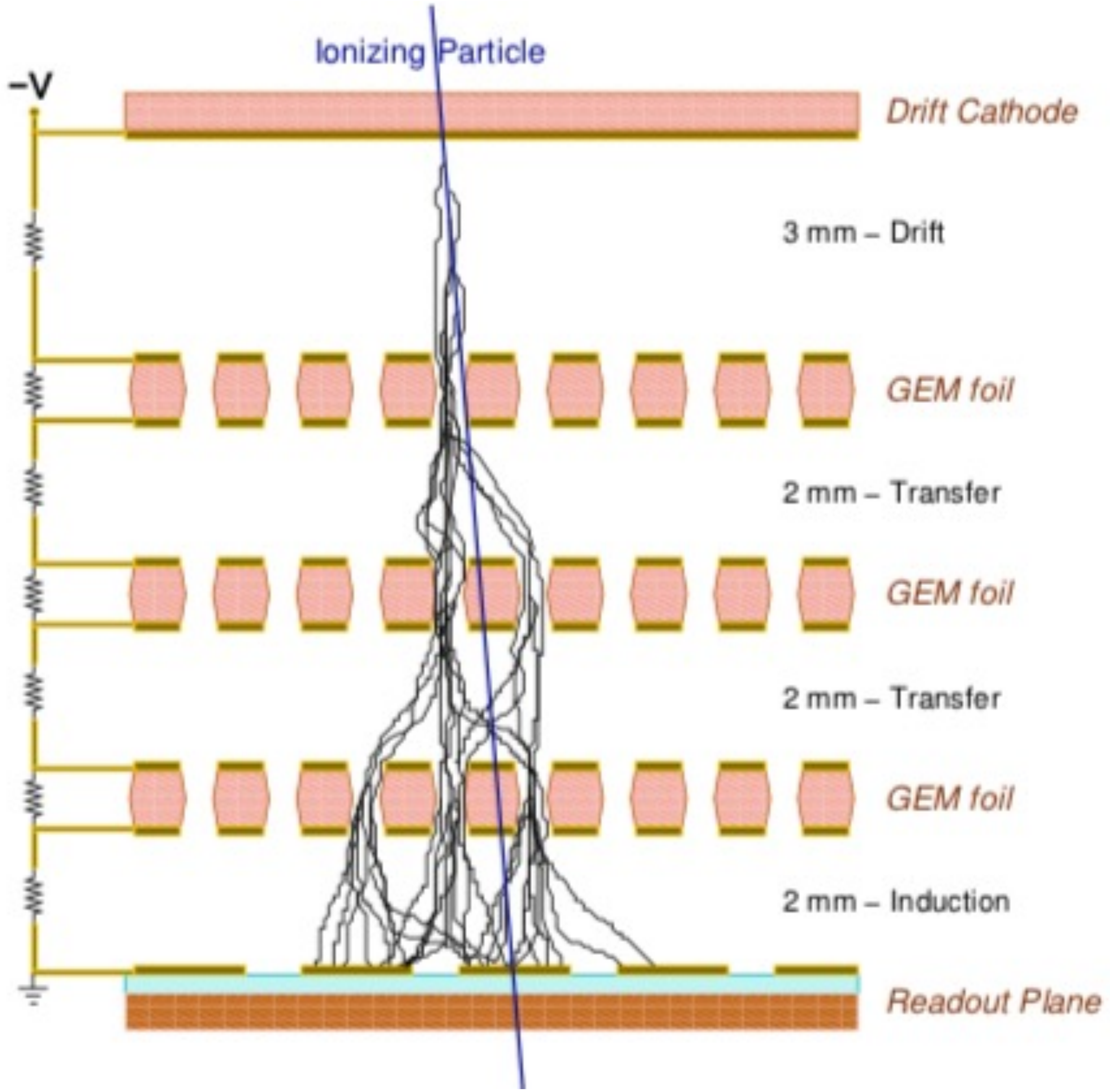
GEM detector

1. Primary ionization electrons produced in the drift layer

2. Signals are amplified by the triple-GEM foils (avalanche process)

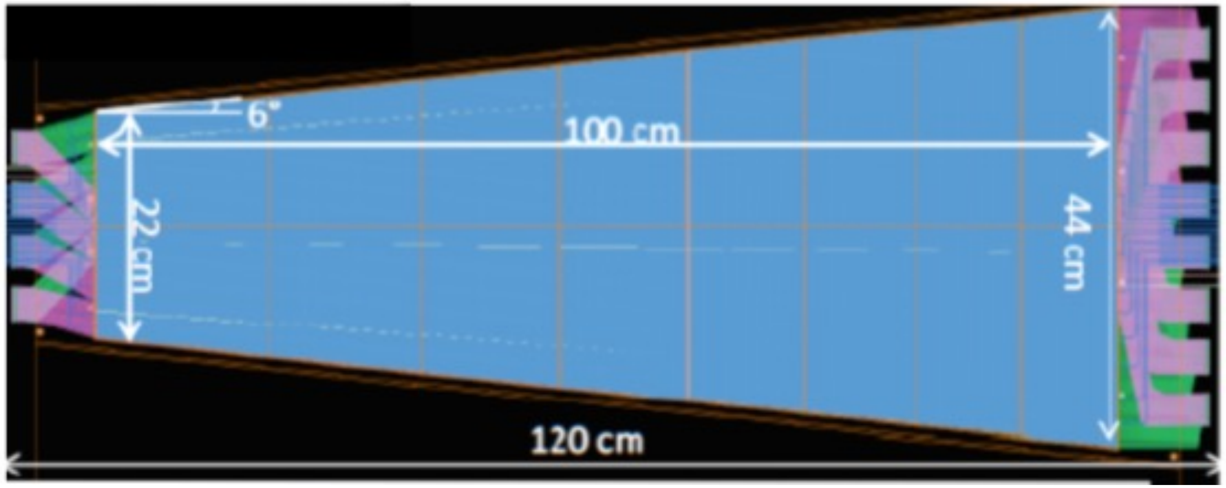
3. Amplified electrons collected by the two readout planes. Finally being read out by APV25 or VMM3 chips

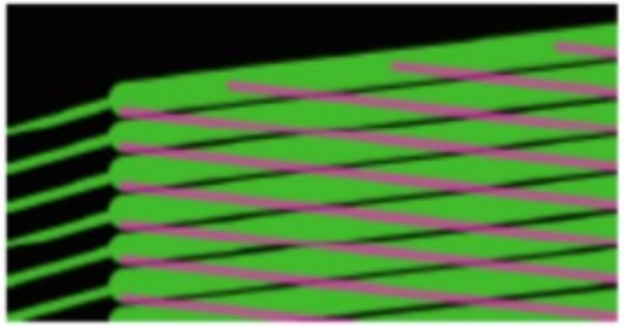
Signal shape can depend on the incident angles



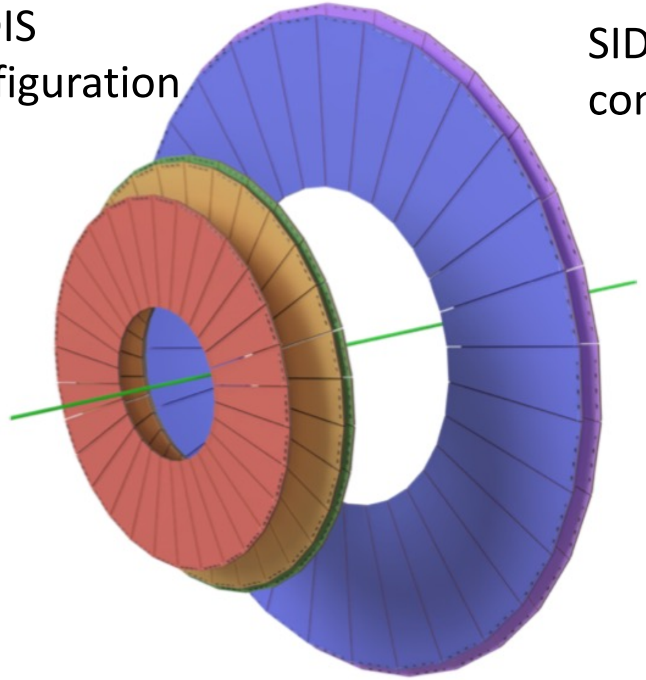
GEM detector

- 200k to 300k channels from all the GEM chambers for each configuration

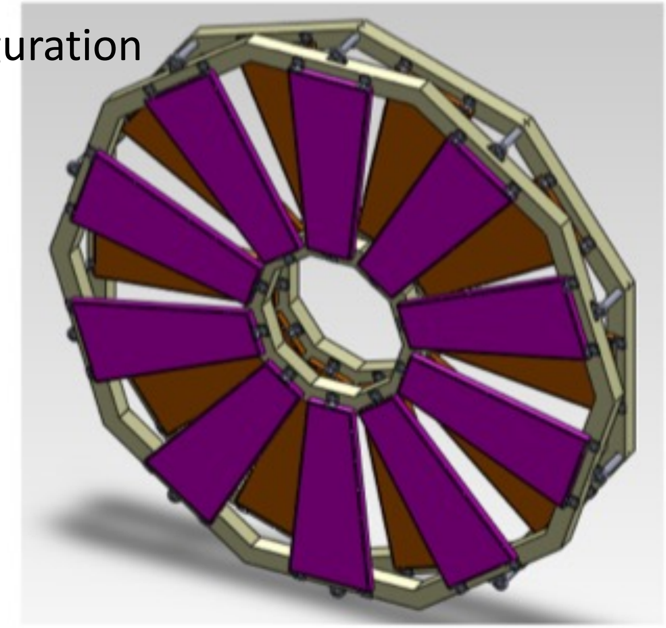


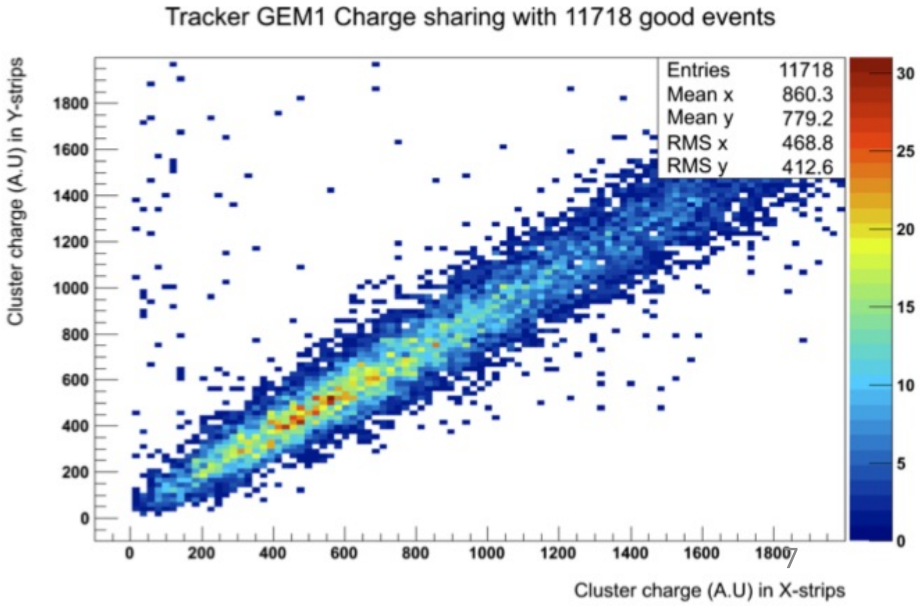


PVDIS configuration



SIDIS configuration



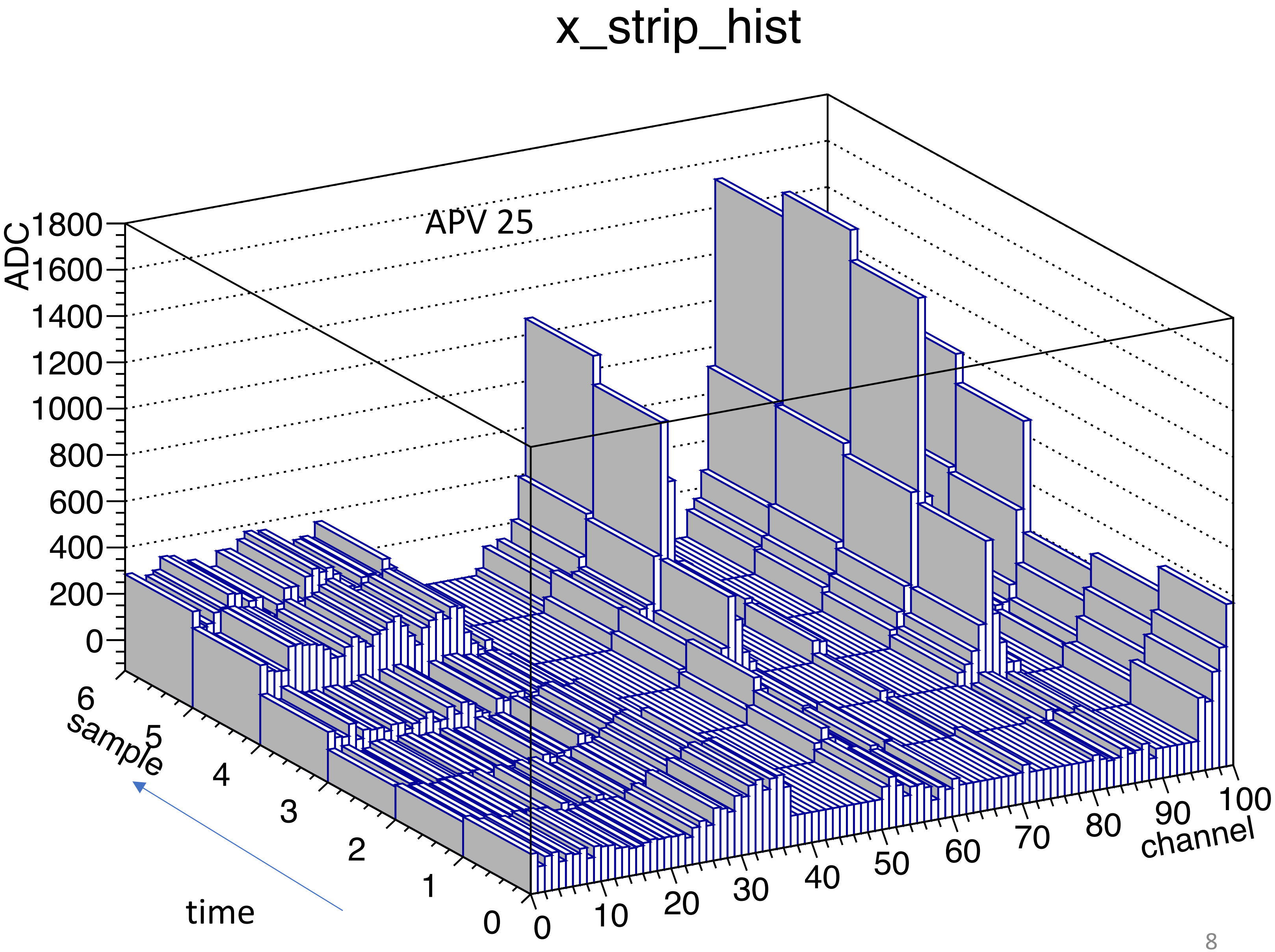


GEM clustering

- Currently, the GEM clustering algorithm uses a peak-valley splitting algorithm
  - If there are multiple "partially " overlapping clusters, split the clusters at the local minima. ADC at the local minima is shared "equally" between the two adjacent clusters.
  - Final position estimated using the charge weighted average

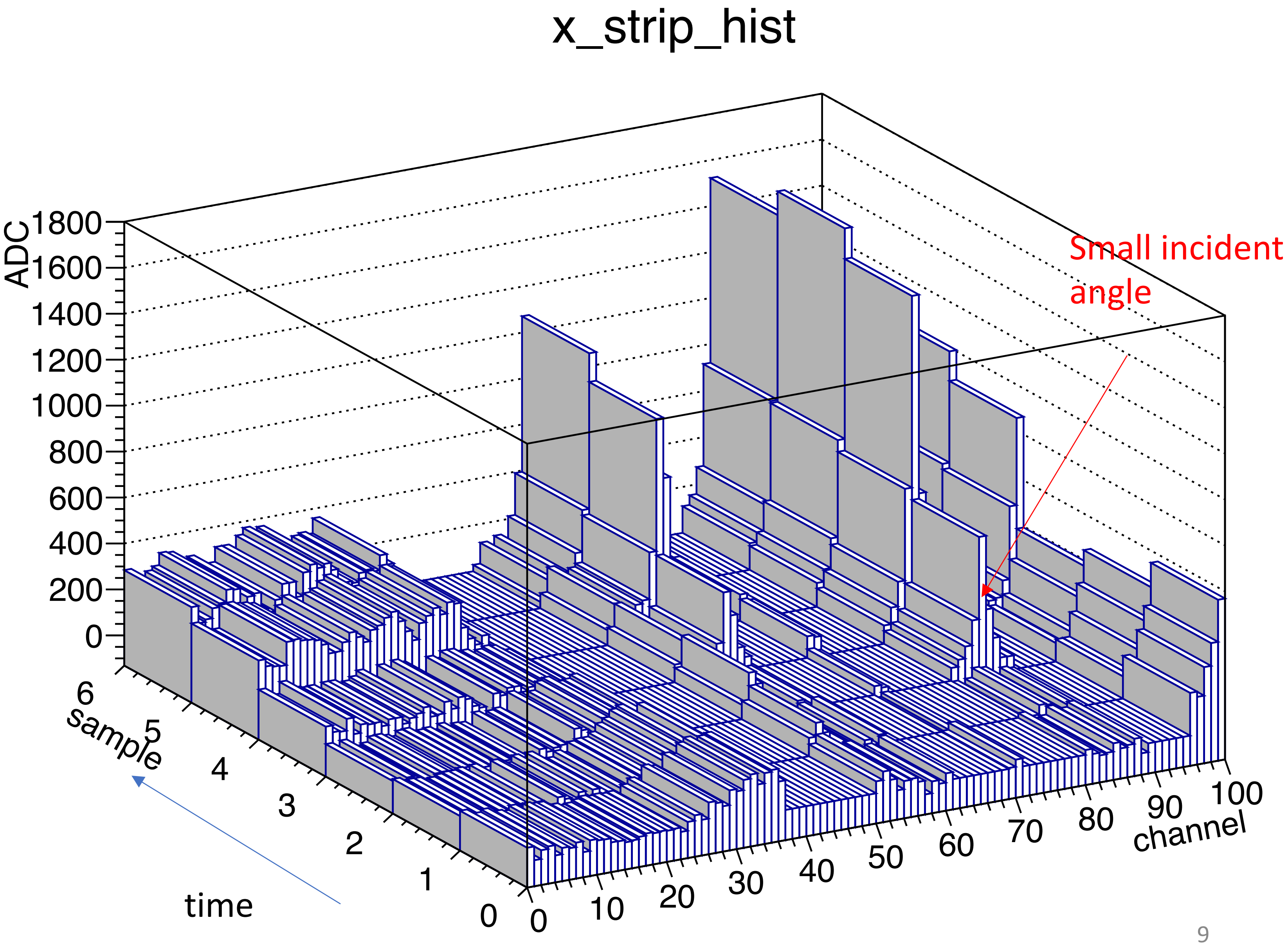
- Advantage: fast and easy to implement, hard to go very wrong

- Disadvantage: crude and no using fully the shape of the clusters



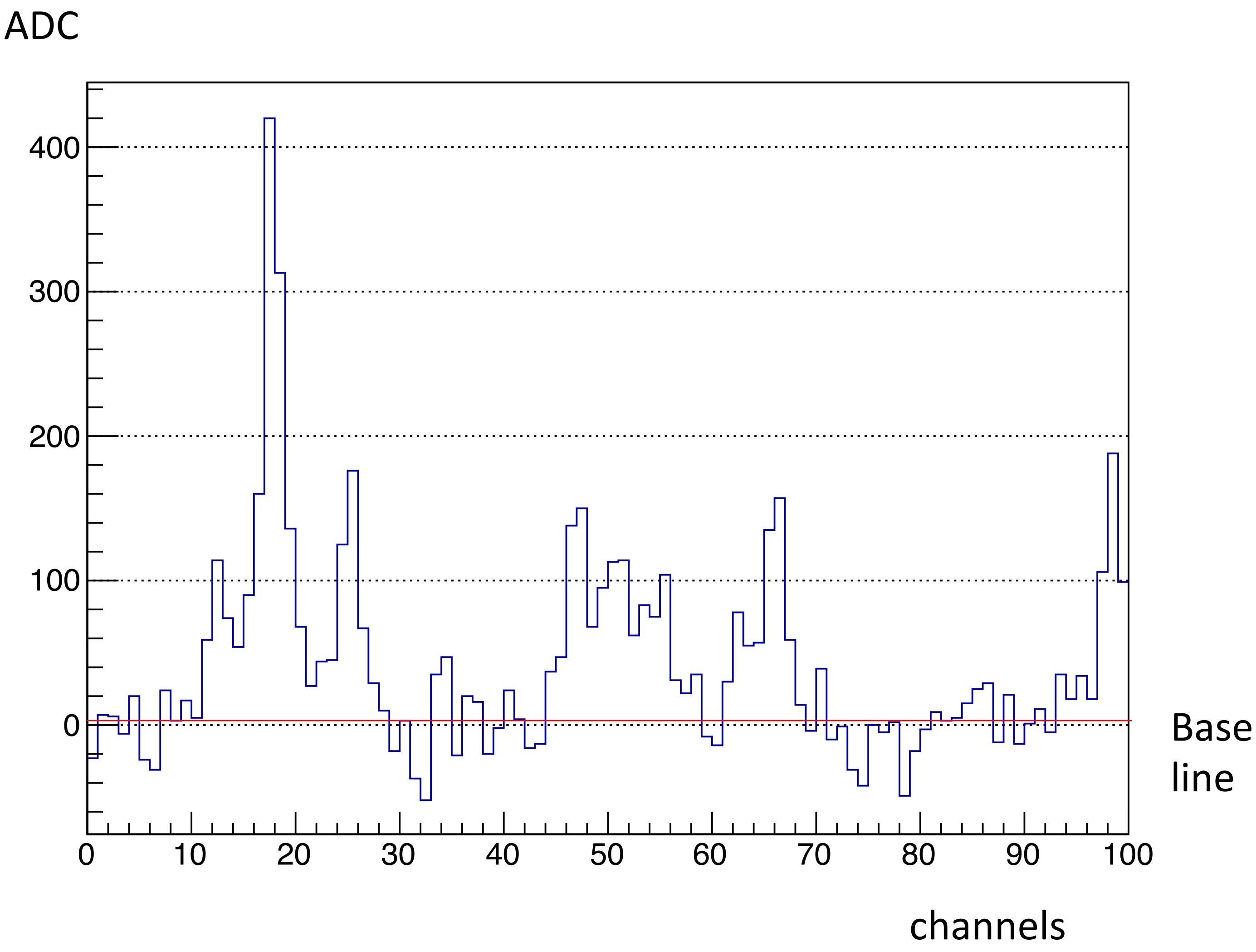
GEM clustering

- Possible improvement from ML:
  - Identify signals from particle with small incident angles from those with large incident angles (more likely backgrounds)
  - Help determine better the reconstructed hit position
  - Help determine the electronic noise and pedestal (depends on the readout chip)



GEM clustering

- For VMM3 we don't have enough information on the electronic noise yet. We will know this later
- For APV25, we already have lots of data and knowledge about the noises. And one of the more difficult noise is the "common mode" (the baseline where the signals sit on)
- To estimate the common mode, we need to take an average of all the "un-fired" channels, but which channels are really un-fired?



GEM Occupancy

- Occupancy numbers are averaged over the entire GEM plane
- Local occupancy could be significantly higher
- Numbers showing here are for the VMM3 chip, for APV, the numbers can be 2 to 3 times higher

PVDIS

	40ns
Plane 1	7.5%
Plane 2	4.3%
Plane 3	3.7%
Plane 4	1.6%
Plane 5	1.6%

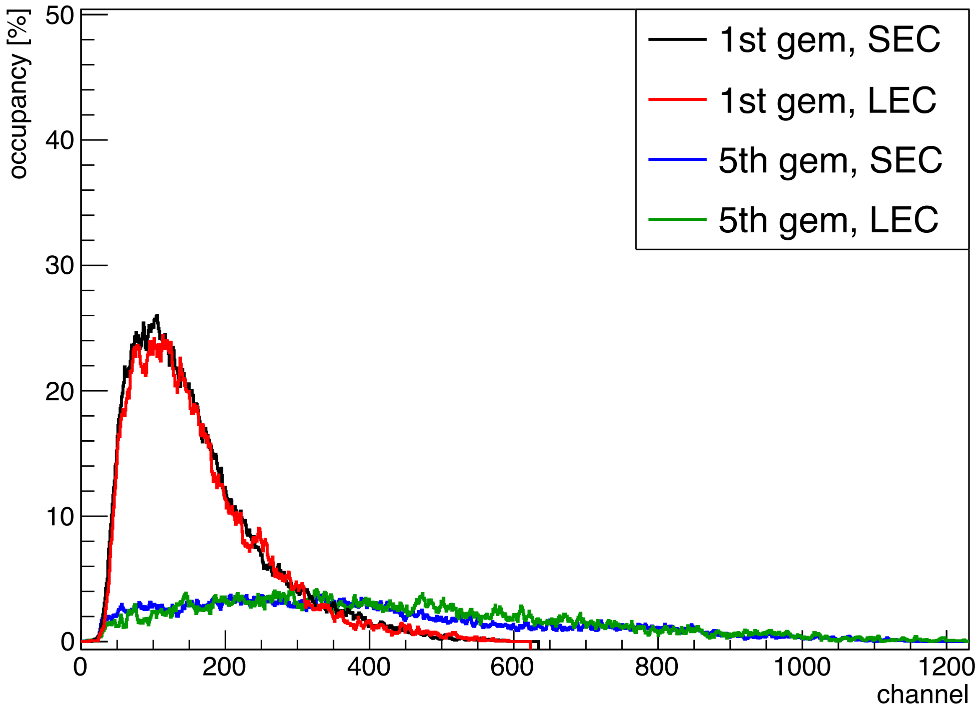
SIDIS

	40ns
Plane 1	1.3%
Plane 2	4.0%
Plane 3	1.9%
Plane 4	1.2%
Plane 5	1.2%
Plane 6	0.9%

JPsi

	40ns
Plane 1	3.1%
Plane 2	5.9%
Plane 3	3.7%
Plane 4	2.9%
Plane 5	2.8%
Plane 6	2.2%

Occupancy for each readout strip (PVDIS)



Kalman Filter Algorithm

Kalman Filter: a recursive fitting algorithm based on χ^2 minimization

1. Prediction: Predict state vector (track parameters) at next measurement site

2. Error Propagation: propagate covariance matrix to the next measurement site and calculate process noise matrix along the way

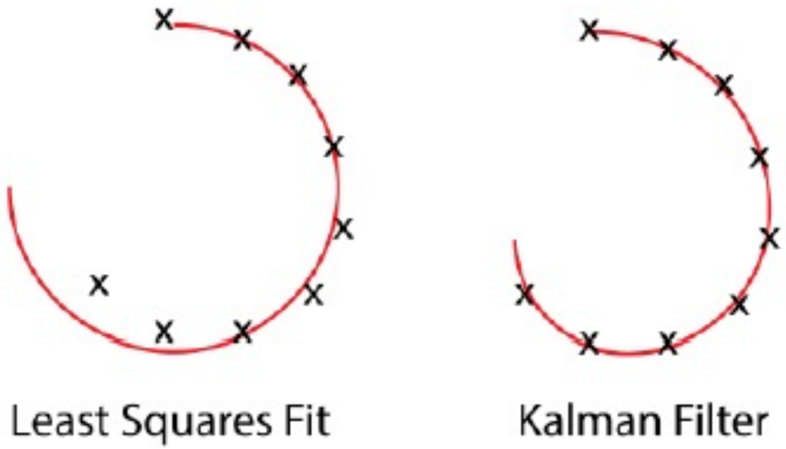
3. Filtering: the weighted mean of the predicted state vector and measurement vector on detector k is calculated

Using Kalman Filter as track finder

Arbitration: decide whether we want to accept the hit based on the prediction, the propagated covariance matrix and χ^2 increment if accept

Kalman Filter Algorithm

- Track representation or state vector (x, y, t_x, t_y, q/p)
  - Allow smooth transition between uniform and fringe field
  - Rely completely on accurate field map measurement

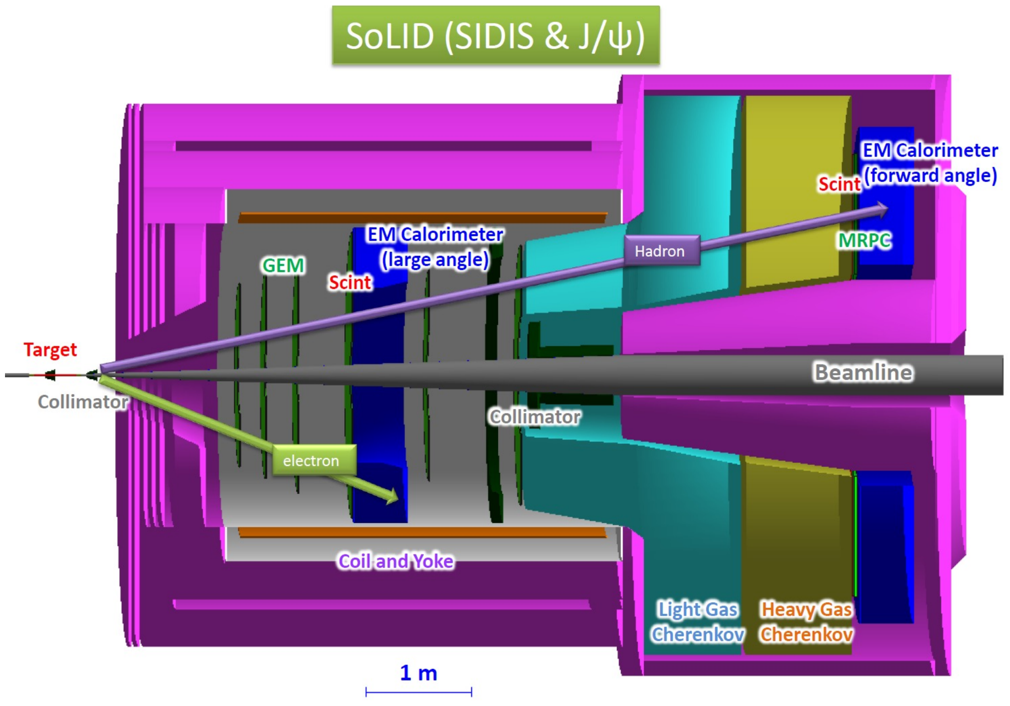


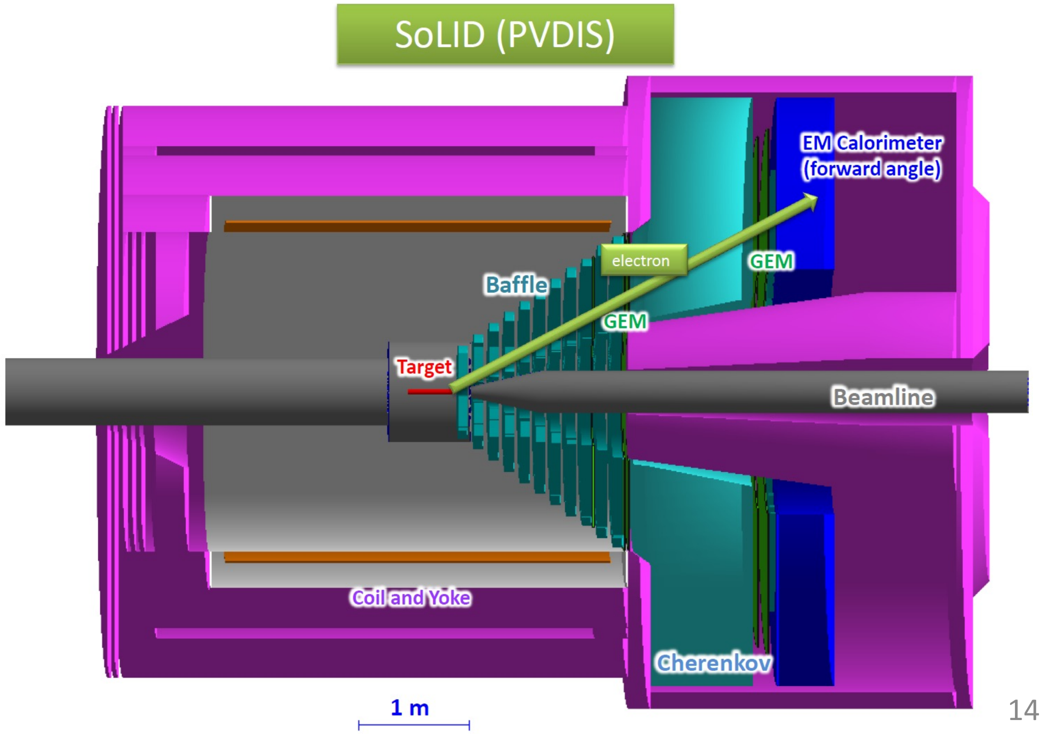
- Kalman Filter track finder advantages:
  - Evolution of track parameters, favors local information
  - Concurrent track finding and fitting
  - Discriminating power improved as more hits added

- Kalman Filter track finder disadvantages:
  - Relatively slow due to field propagation and large computation power requirement (5-D matrices propagation, multiplication and inversion)
  - Weak discriminating power at early stage
  - Rely on efficient seed finding

Kalman Filter Algorithm -- Seeding

- The first step for KF is seeding – making an initial guess of the track parameters
  - Loop through hits from the last three GEMs (lower occupancies or hit multiplicities) to form a track segment
  - Do a rough propagation to EC/SPD, as well as the target area to see if it is potentially a good seed
    - For PVDIS we only detect electrons, so we only care about the EC, for SIDIS and JPsi we want to detect pions and protons too, so we will look at the SPD





Kalman Filter Algorithm -- Seeding

- Usually we will end up with 10 to 1000 (most likely will be more) seeds for each particle we want to detect
- We need to propagate each seed to the upstream GEMs to see if there are enough hits to pick up. And whether the track can be propagated to the target region
- Each particle takes 1 to 10 ms to reconstruct, for multiple particles, we need to add the time together

- This is a rather time-consuming process. It would be great if we can reduce the number of seeds with help from ML
- Currently we know some successful examples from Hall B using ML algorithm like multi-layer perceptron

Track Selection rules

1. A track must have enough hit found (usually 4 out of 5 GEM planes)
2. χ^2 / ndf
3. Reconstructed vertex z, momentum and polar angle
4. Matching with hits from other detectors like Ecal and SPD
5. Charge symmetry for GEM hits
6. Coincidence vertex for multi particle tracks
7. …

- If more than one track found for a particle, we select the one with most hit and smallest χ^2
- This is essentially a multi-variant analysis problem, and ML may help to optimize the cuts

Current results

With VMM3 chip

PVDIS

Zero-track rate	Single-track rate	Multi-track rate	Single-track acc.
18.9%	80.7%	0.4%	91.9%

SIDIS

	Zero-track rate	Single-track rate	Multi-track rate	Single-track acc.
electron	6.6%	93.3%	0.0%	98.2%
pion	11.6%	88.2%	0.0%	97.0%

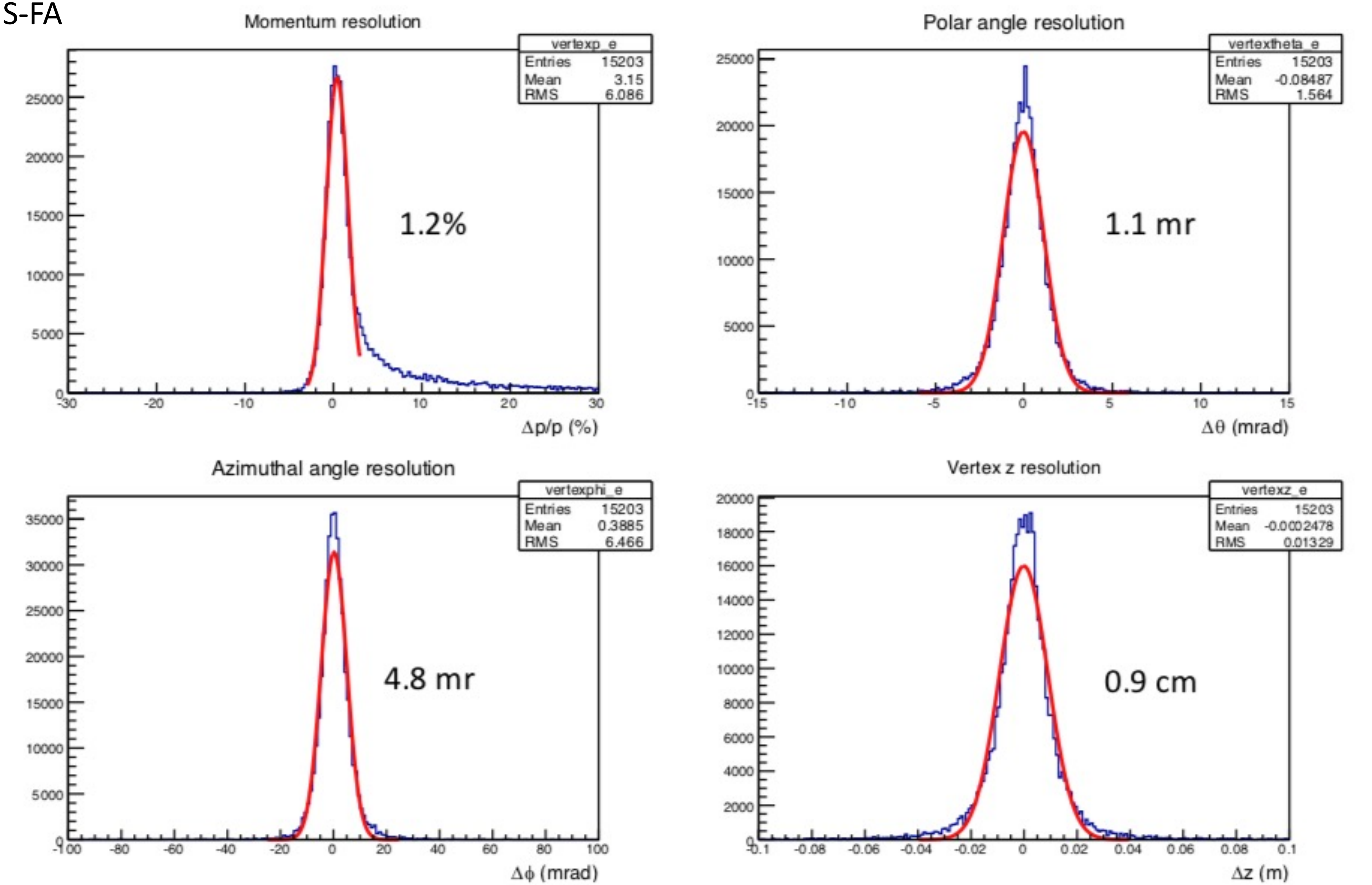
JPsi

Zero-track rate	Single-track rate	Multi-track rate	Single-track acc.
16.4%	83.5%	0.1%	94.2%

- A track is considered accurate if all its reconstructed hits are “good” hits
- “Good” hits are the closest reconstructed hit to the MC true hit, within +/- 1.2mm
- ML may help getting rid of the bad hit as well

Current results

SIDIS-FA



What we can provide for ML

- GEM digitized simulation: ADC for each GEM channel. We also can indicate which channel contains the signal we want to detect, as well as the true MC hit position

- For seeding: we can provide simulation and identify the true seed

- For the track selection, we can also provide the information for the true tracks, as well as bad tracks

Summary

- ML has the potential to enhance significantly the performance of tracking
    1. GEM clustering: identify good signals from noisy 1D histogram, help get rid of electronic noise
    2. Help finding good quality seeds: improve speed in the reconstruction, or maybe even KF
    3. Optimize the cuts used to select the final tracks: a multi-variant analysis problem

- We can provide digitized simulation with true MC info as the dataset for ML training