

SPINE for DUNE and its Prototypes

Nived Puthumana Meleppattu

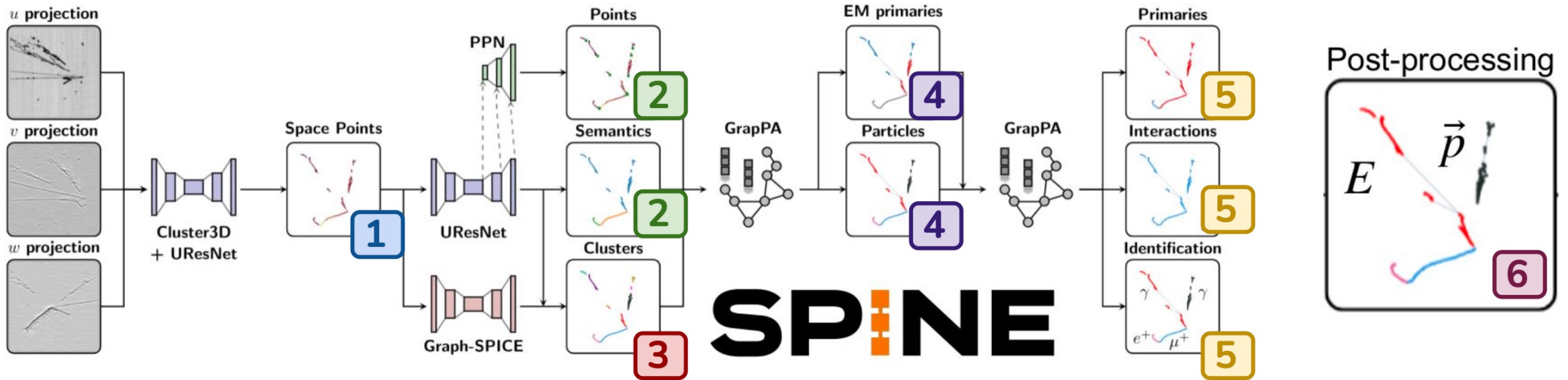
DUNE France meeting
01/04/2026

SPINE

DUNE


iJC Lab
Irène Joliot-Curie
Laboratoire de Physique
des 2 Infinis

Scalable Particle Imaging with Neural Embeddings

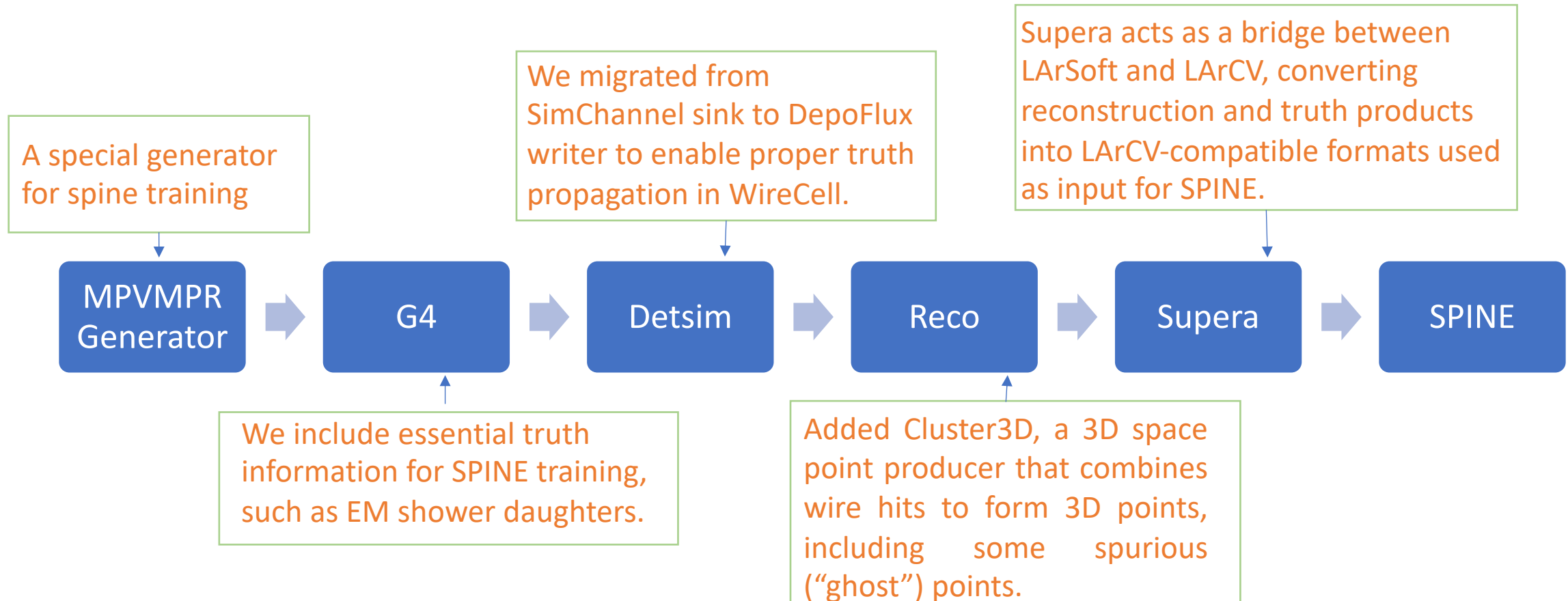


A fully trainable pipeline that takes raw 3D detector hits, groups them into higher level structures like particles or interaction vertices, and assigns each structure a semantic or particle type label.

- All the work presented in this presentation was shared by me at the SPINE Workshop 2026 held at CERN in February.

<https://github.com/DeepLearnPhysics/spine-workshop-2026/tree/main>

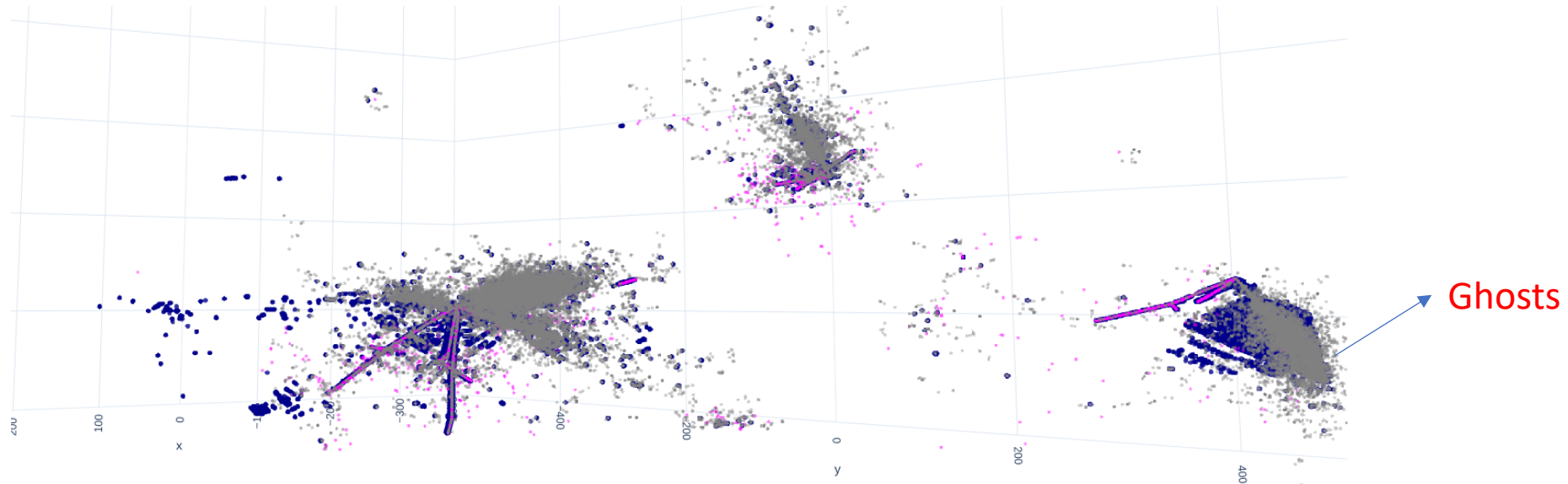
Updated Simulation and reco Pipeline



The updated configs are available in the latest dunesw release <https://github.com/DUNE/dunesw/tree/develop/fcl/protodunevd>

MPVMPR generator

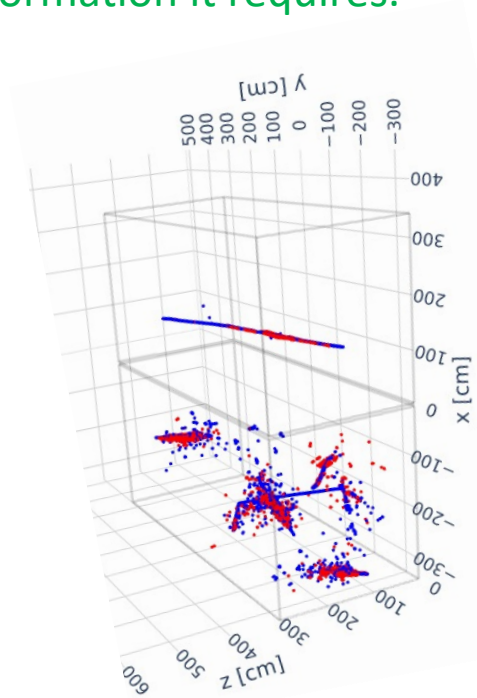
- MPVMPR is a custom event generator designed specifically for SPINE training.
- It produces particle guns and bombs in random detector directions, avoiding bias toward any specific generator or energy range.
- An example FDHD training sample image from this generator is shown below.
- The configuration for FDHD is available at /exp/dune/data/users/nivedpm/ML/spine_workshop/dunefd/



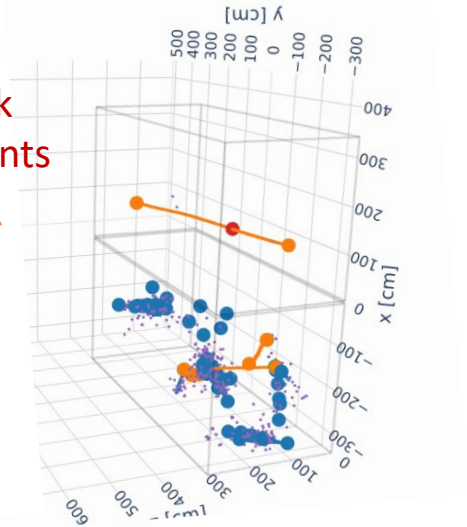
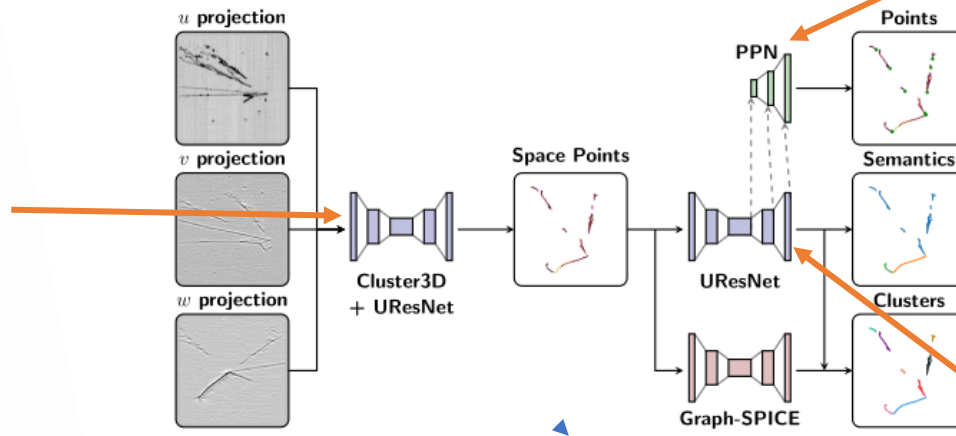
Protodune VD – Inputs for spine

Each SPINE branch is trained using the specific truth information it requires.

Vertices and track start and end points



3D Spacepoints with ghosts

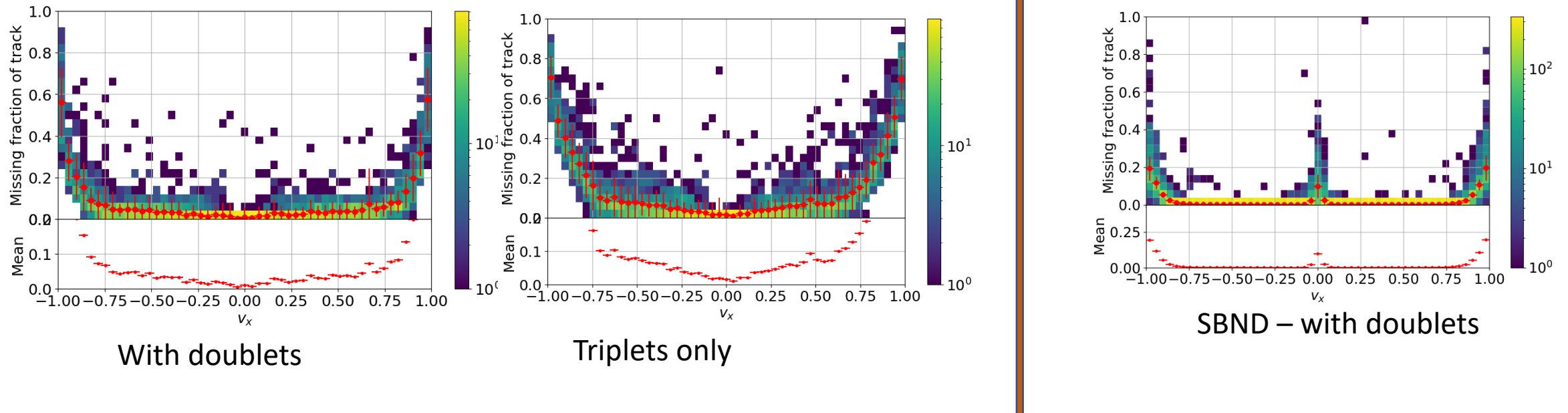


- Semantic classes
- Shower
 - Track
 - Michel
 - Delta
 - Low energy

Track incompleteness - PDVD

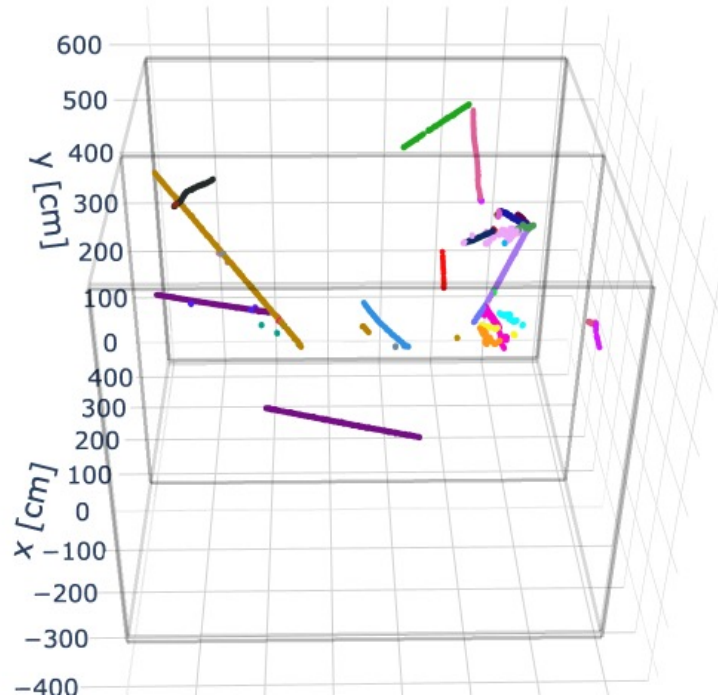
We analyse track incompleteness to assess the “gappiness” of tracks, using it as a metric to evaluate the performance of our spacepoint producer and ghost labelling algorithm.

Missing fraction of track = gap length / total length of track

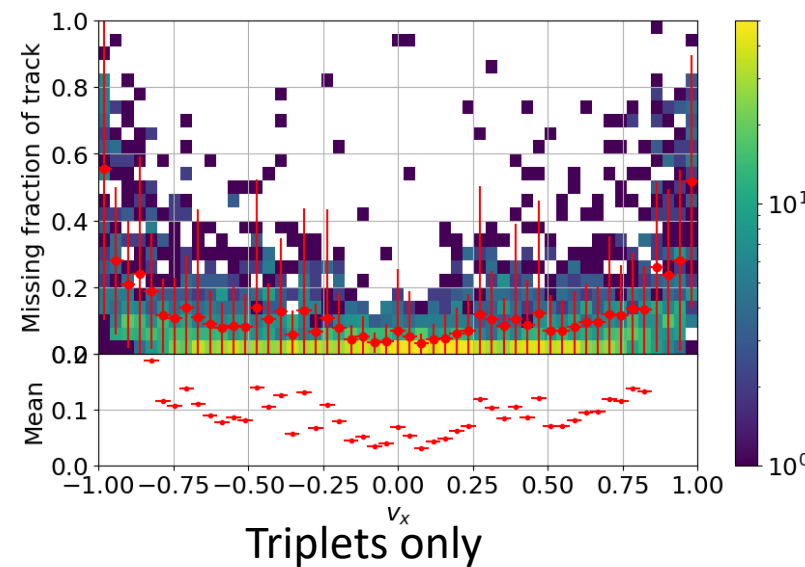
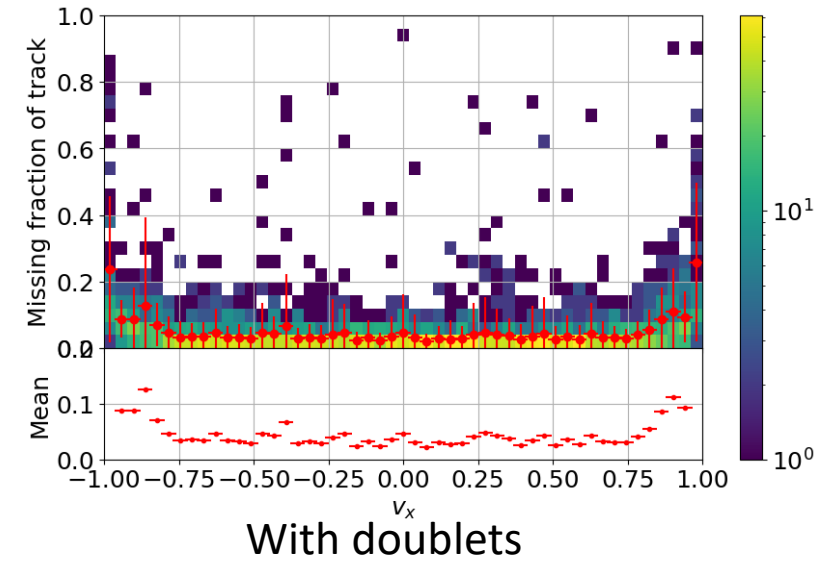


- We include doublets to recover 3D points when one wire plane is inefficient or missing hits
- Compared to SBND results, our initial training sample performs slightly worse but remains acceptable

Protodune SP

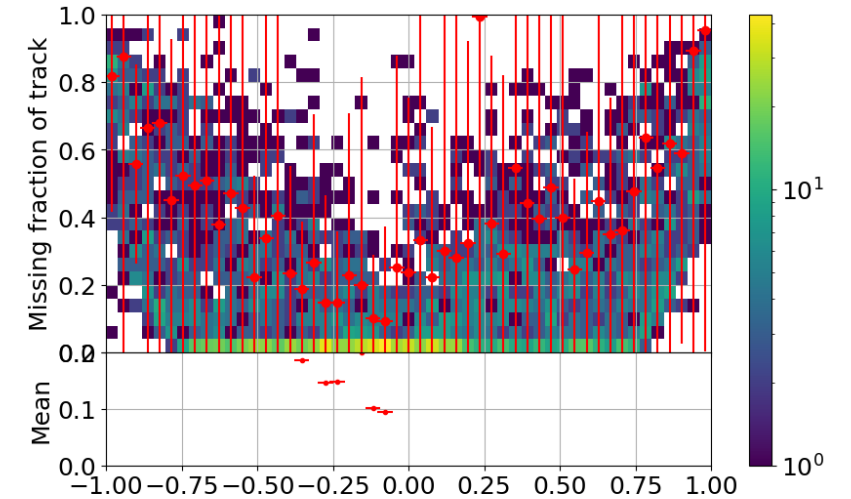
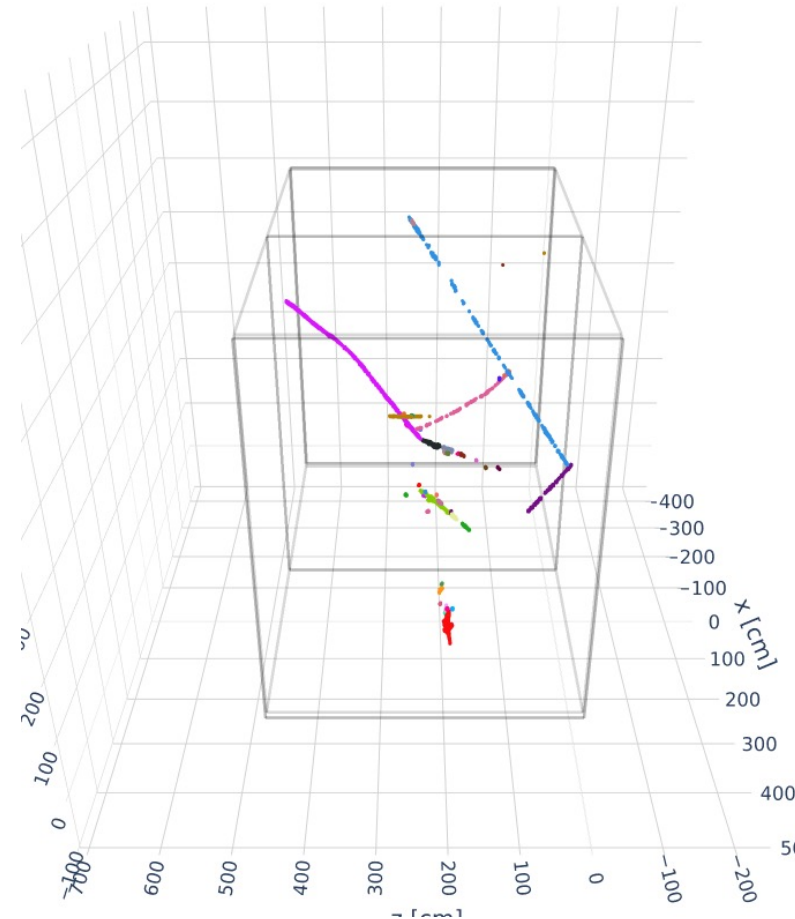


- PDSP produces slightly better tracks with less gappiness, likely due to its higher resolution.
- We plan to reprocess PDSP data, perform a cosmic+beam simulation, and compare the reconstruction performance with other pipelines.

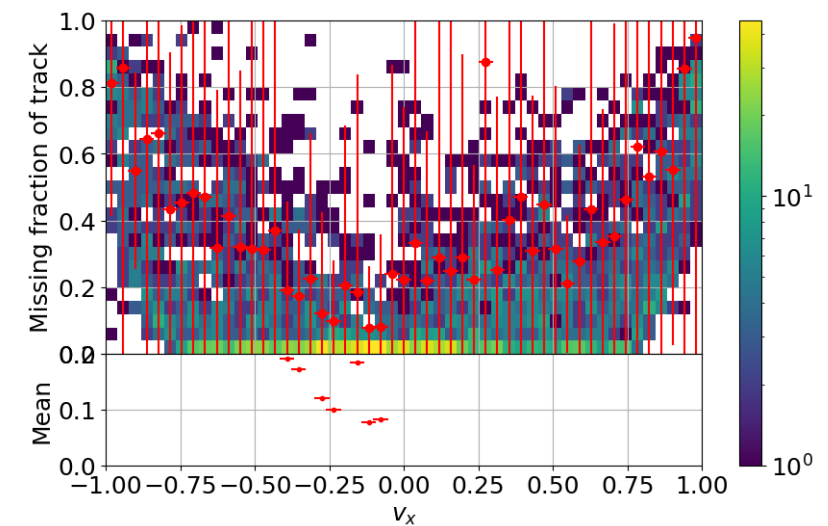


Protodune HD

- The faulty APA 1 in PDHD leads to a significant fraction of tracks being highly gappy.
- Such degraded track quality is suboptimal for SPINE training.
- This issue is deferred for now, pending further reconstruction improvements.



With doublets

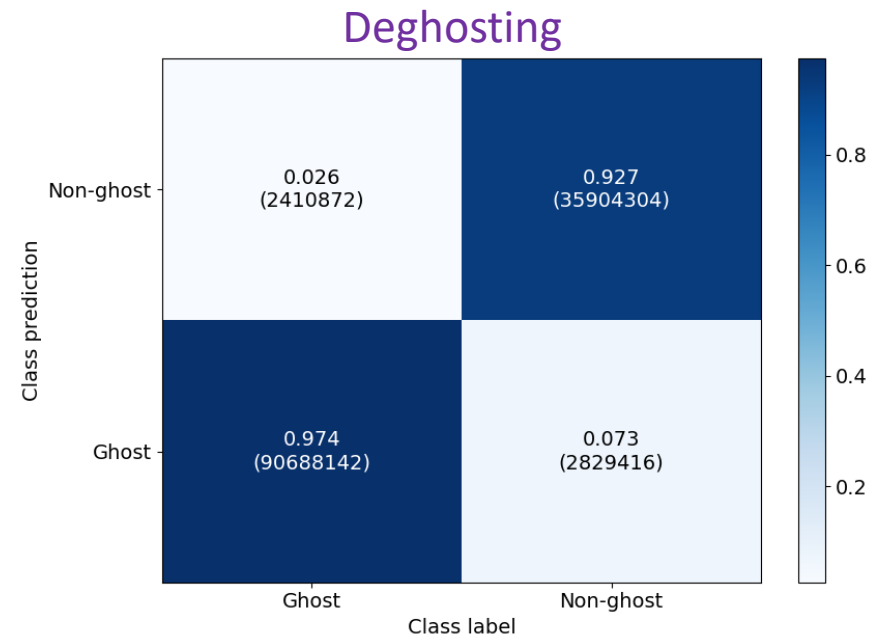


Triplets only

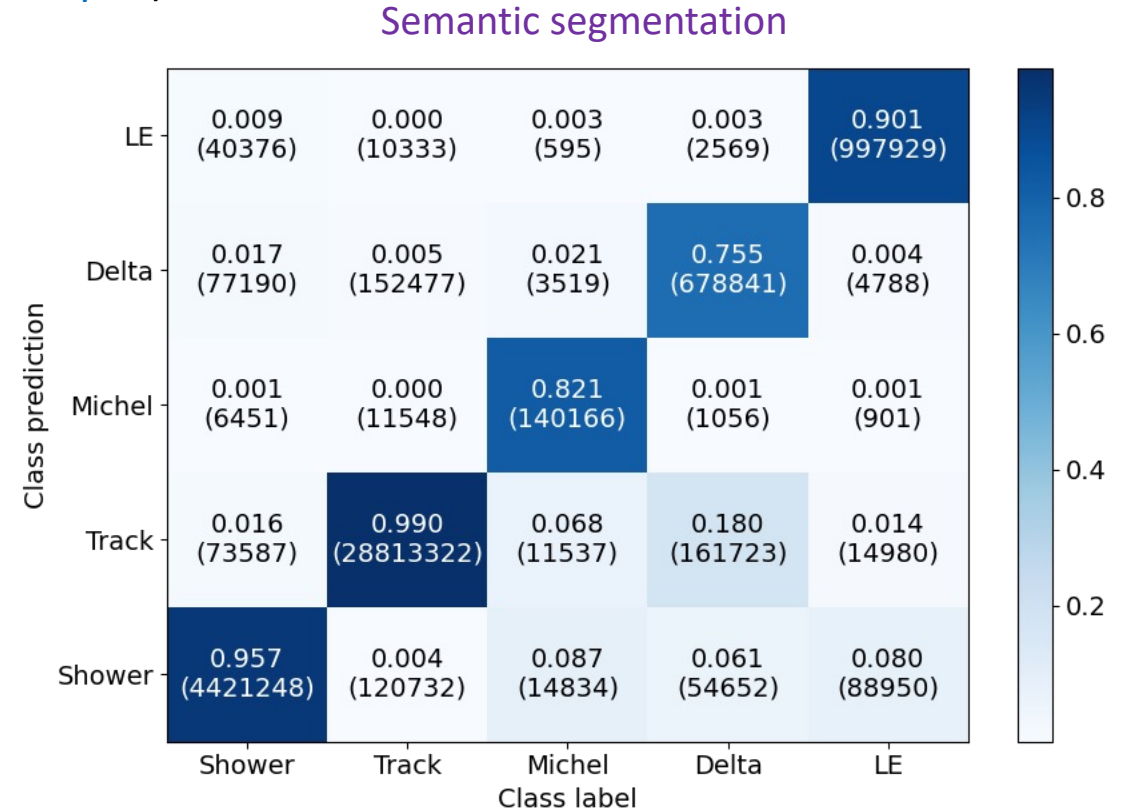
TRAINING SPINE FOR PDVD

- The network was trained with 350K mpvmpr events
- (dataset available at </pnfs/dune/persistent/physicsgroups/dunepd/>)

Results



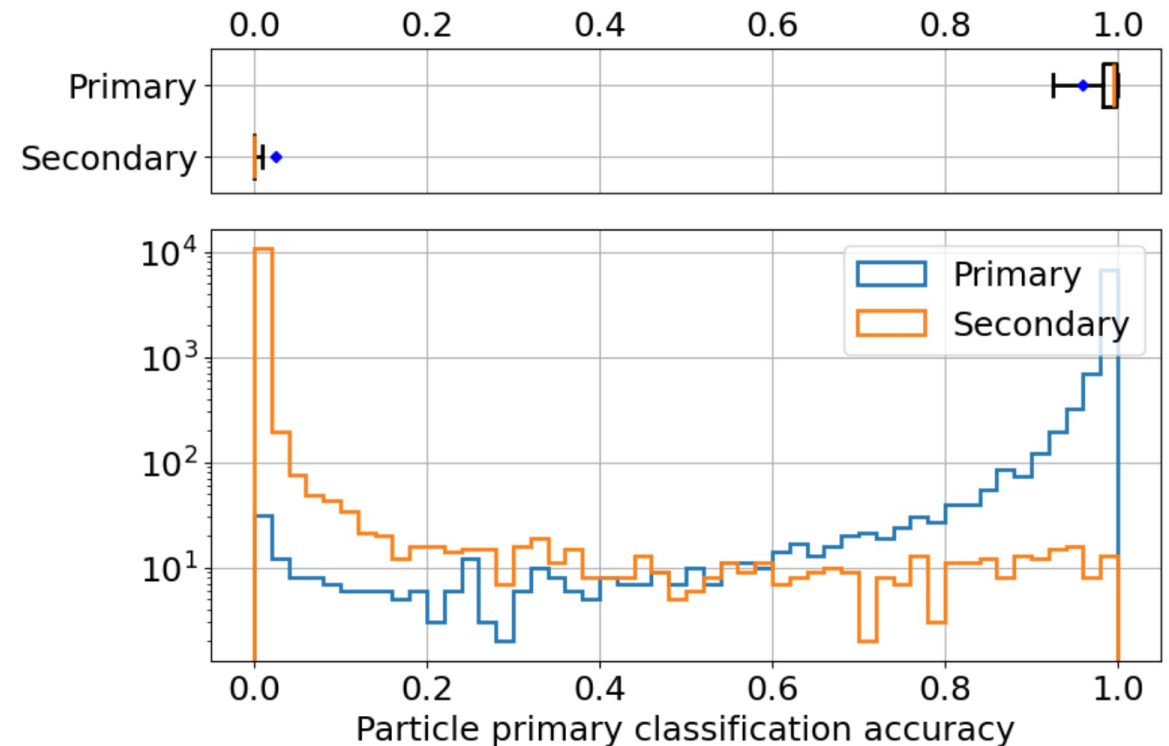
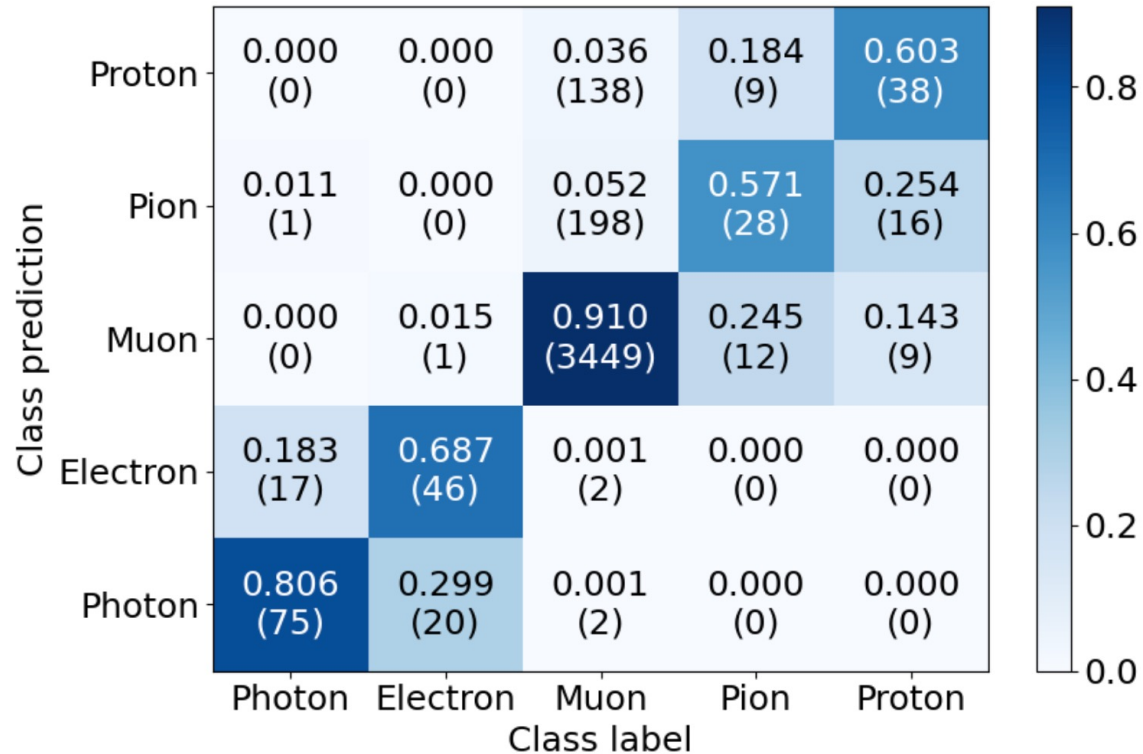
Ghost point separation



Separation into topologically different types of activity

Particle clustering and PID

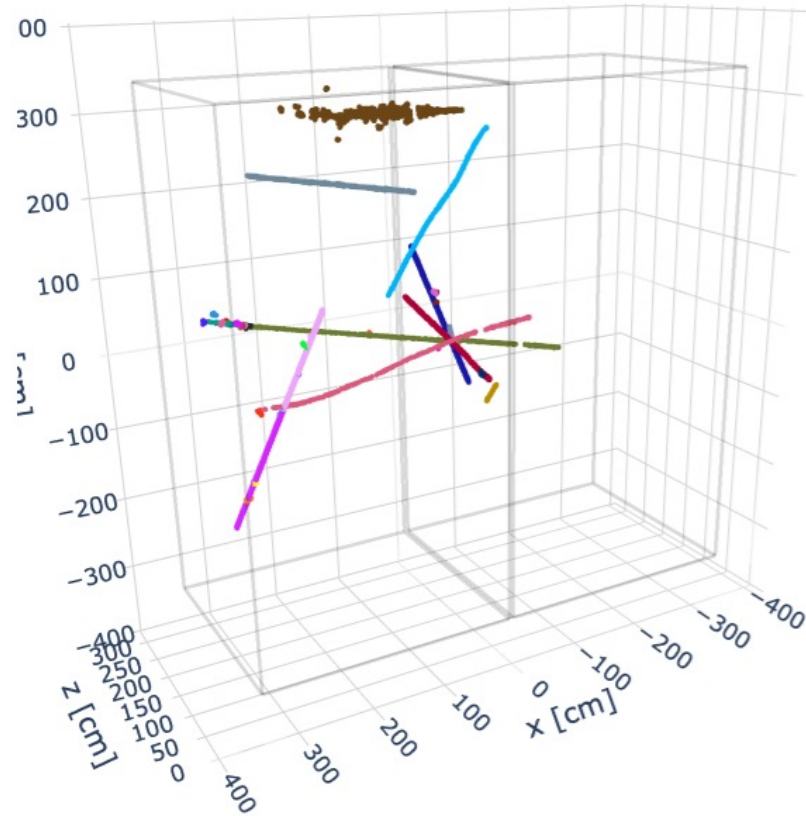
The final SPINE branch (GRAPPA-Inter) provides PID and primary particle performance.



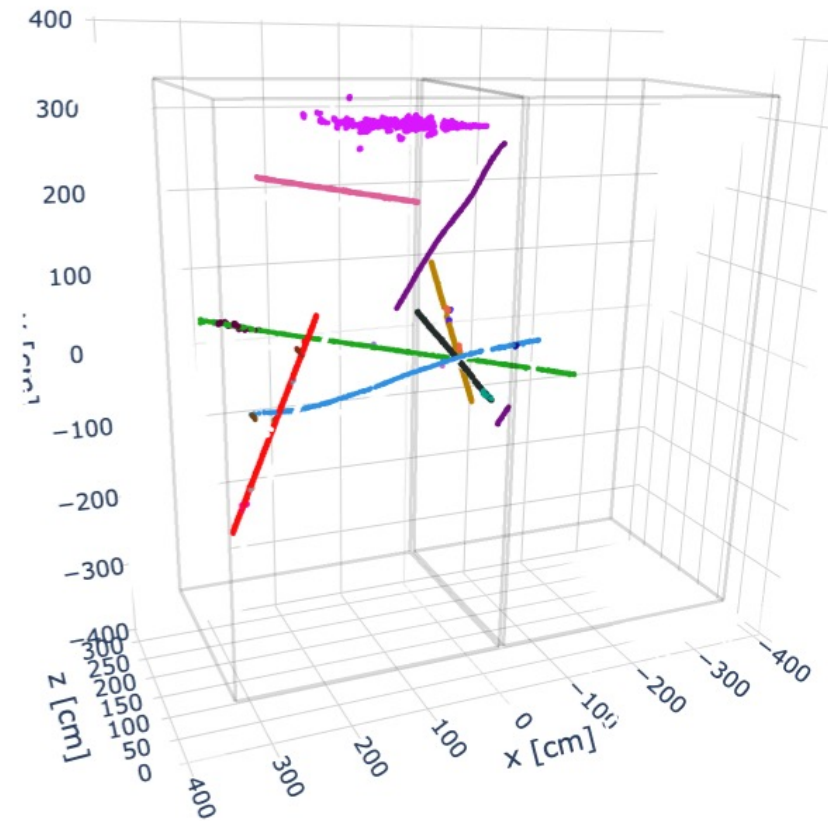
Primary performance $\approx 97\%$

(percentage of correctly identified primaries)

Reconstructed particles



Truth particles



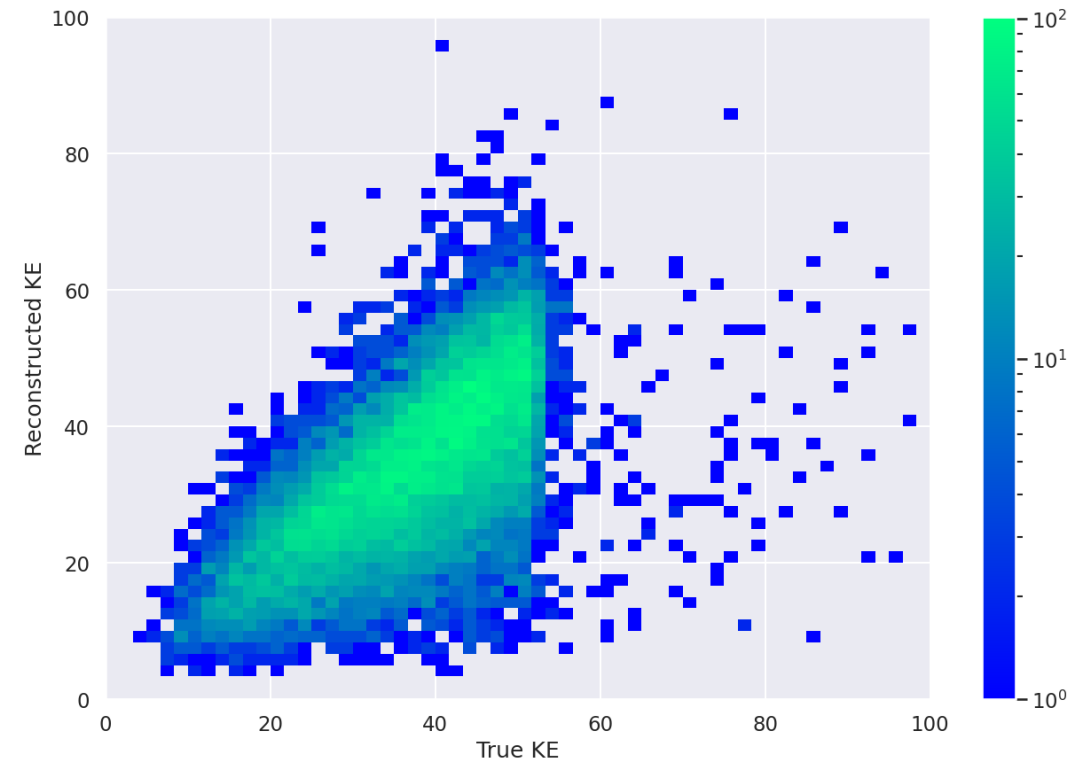
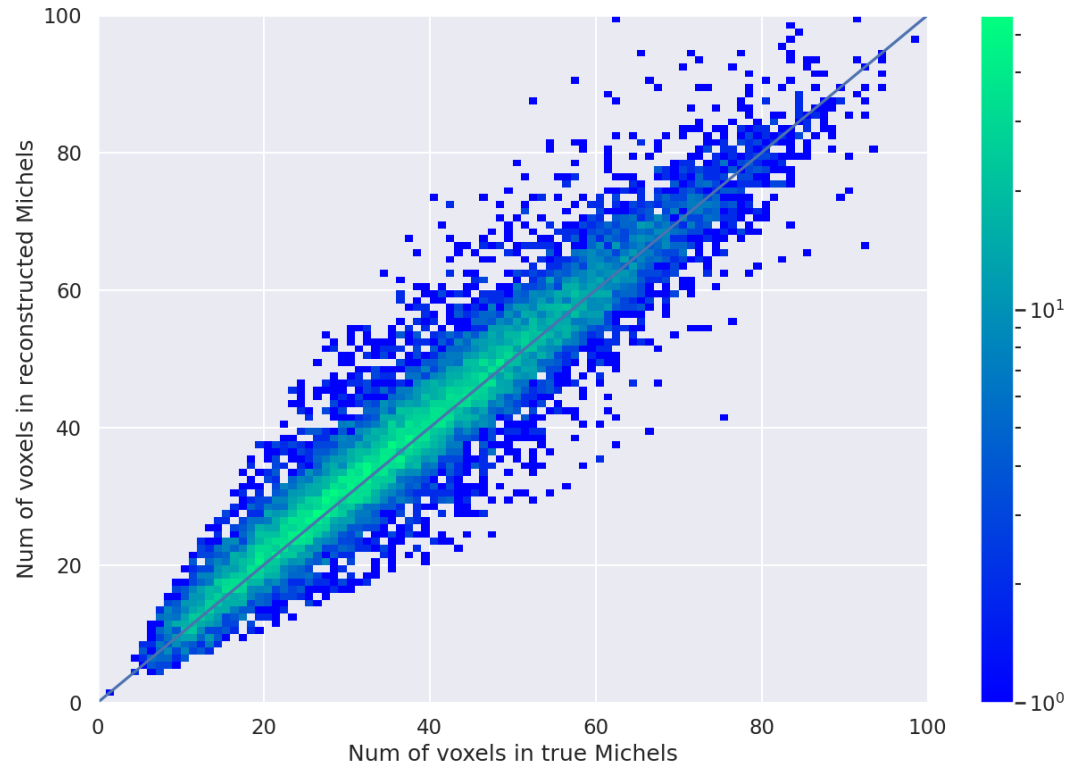
Legend

- Reco particles
- TPC
- Truth particles
- TPC

True vs. reconstructed particles in PDVD MPVMMPR events using SPINE.

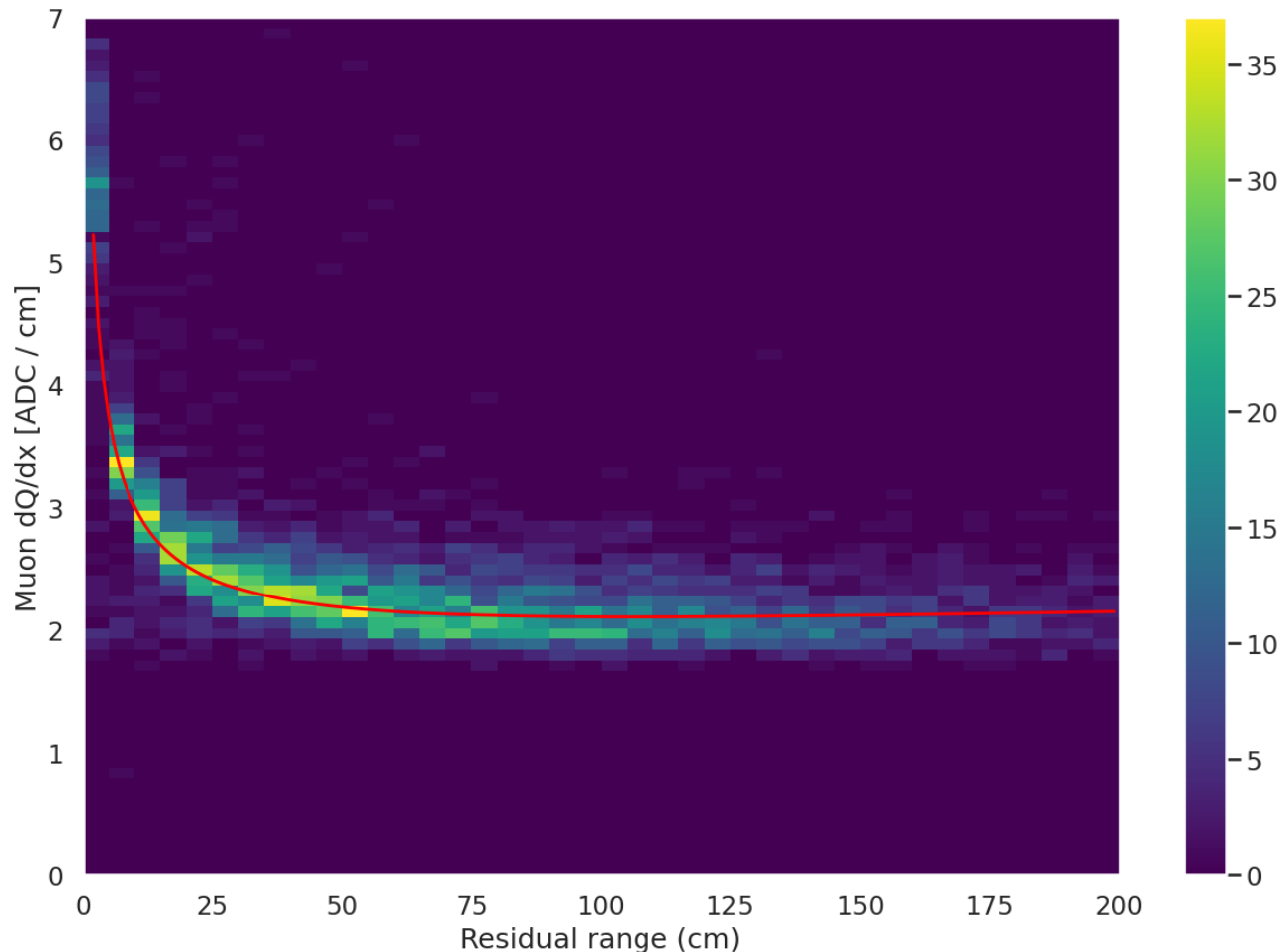
Michel tagging – PDVD Cosmic sim

We test higher-level calibration parameters, such as Michel electron reconstruction efficiency, at both spacepoint and track levels, showing the expected Michel energy results.



<https://github.com/DeepLearnPhysics/spine-workshop-2026/tree/main/reconstruction/michel>

Stopping track dQ/dx



- Clear **minimum ionizing plateau** at large residual range
- The distribution demonstrates **physically consistent behavior across the full track length**, supporting the validity of the reconstruction and calibration chain

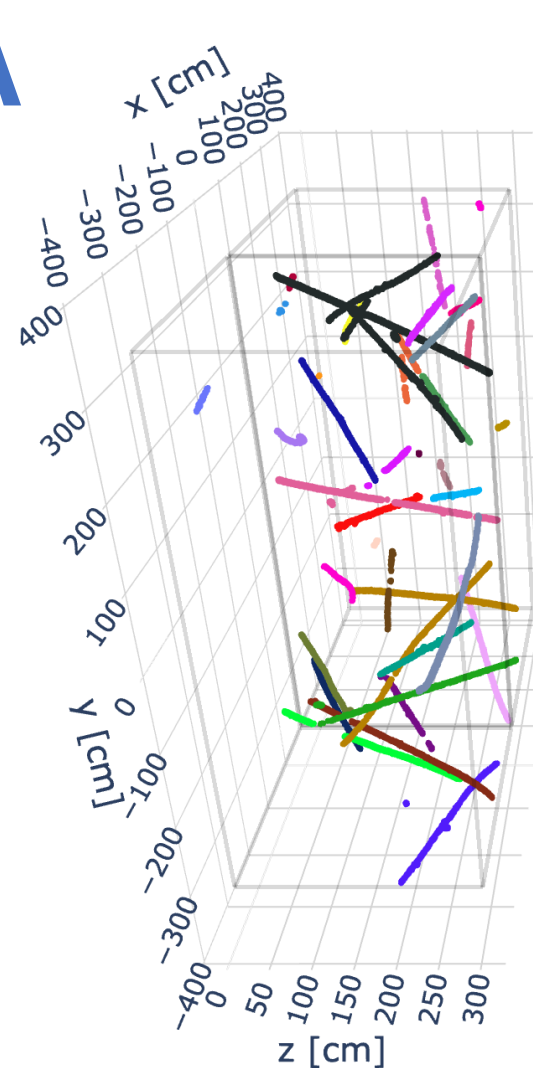
https://github.com/DeepLearnPhysics/spine-workshop-2026/blob/main/reconstruction/stopping_track_dqdx/dEdx_analysis.ipynb

PDVD COSMIC DATA

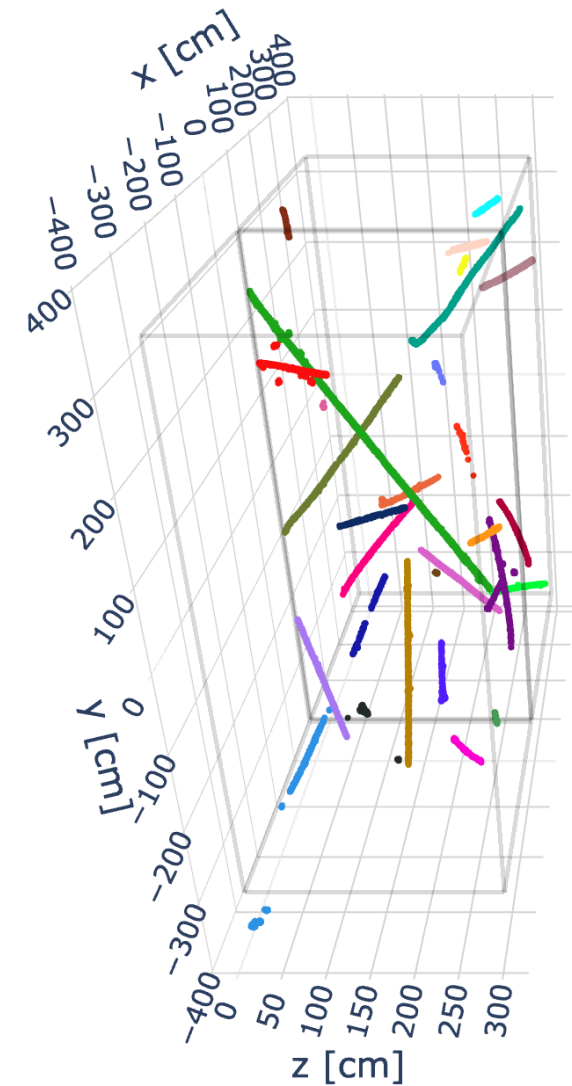
We reconstruct some PDVD cosmic data. It gives good results, and even tracks with gaps are reconstructed correctly.

Some analysis done on this data is given at

https://github.com/DeepLearnPhysics/spine-workshop-2026/blob/main/reconstruction/michel/michel_data.ipynb



Event 1

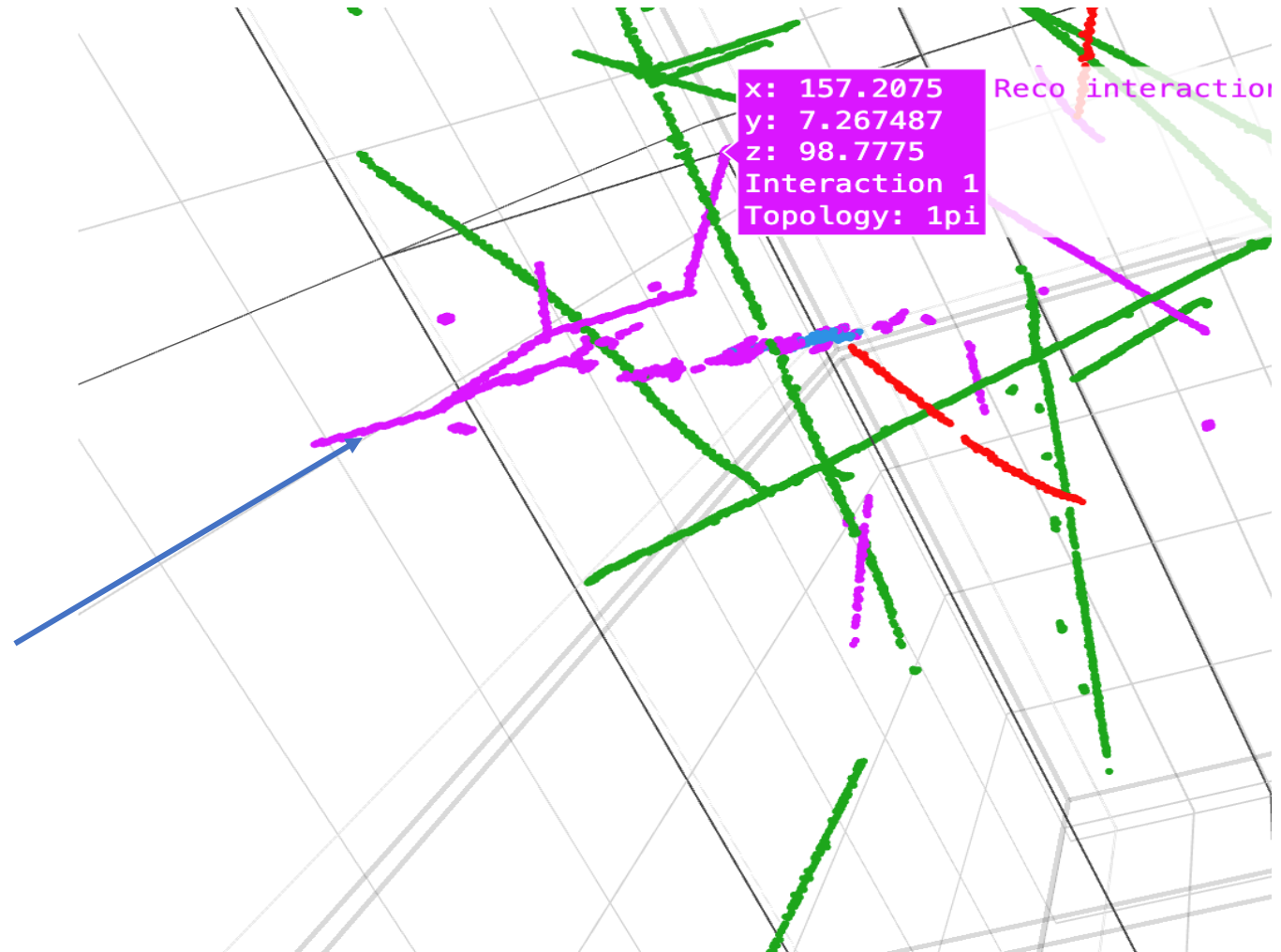


Event 2

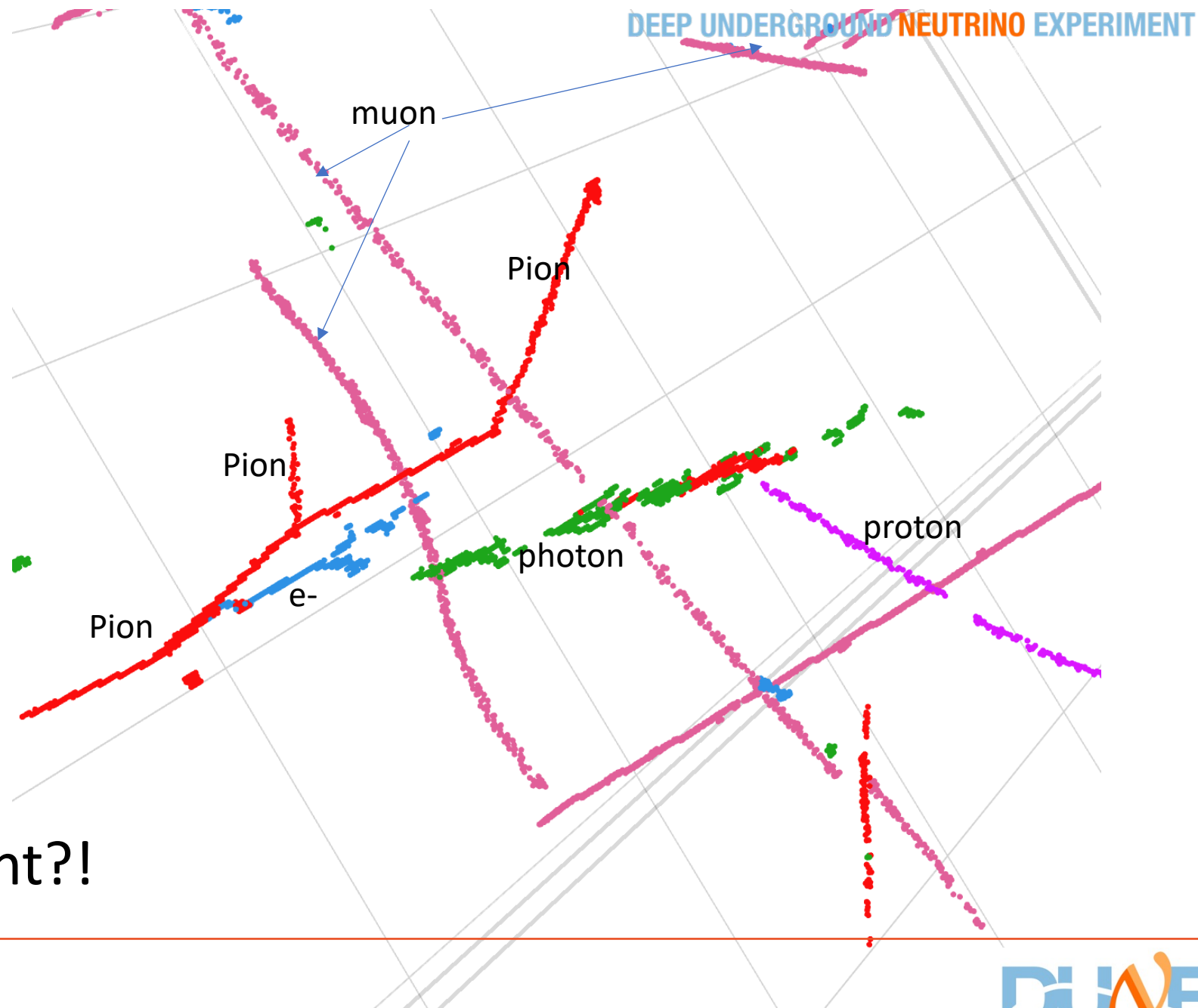
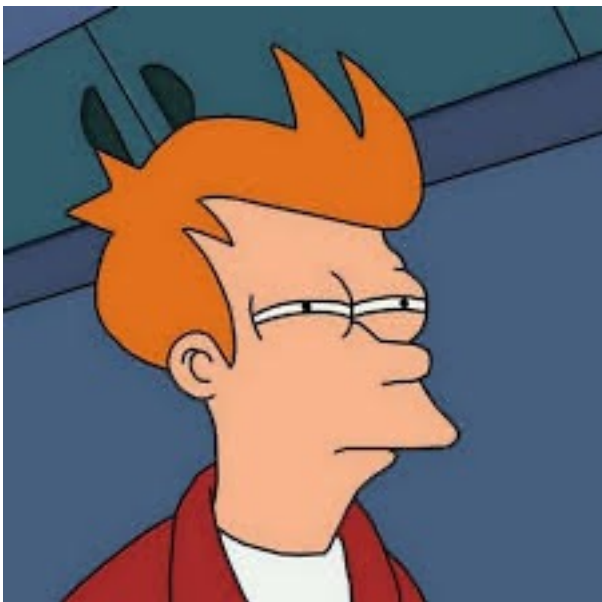
A Beam event!

We look at a PDVD beam event and the reconstructed products

Correctly identified interaction

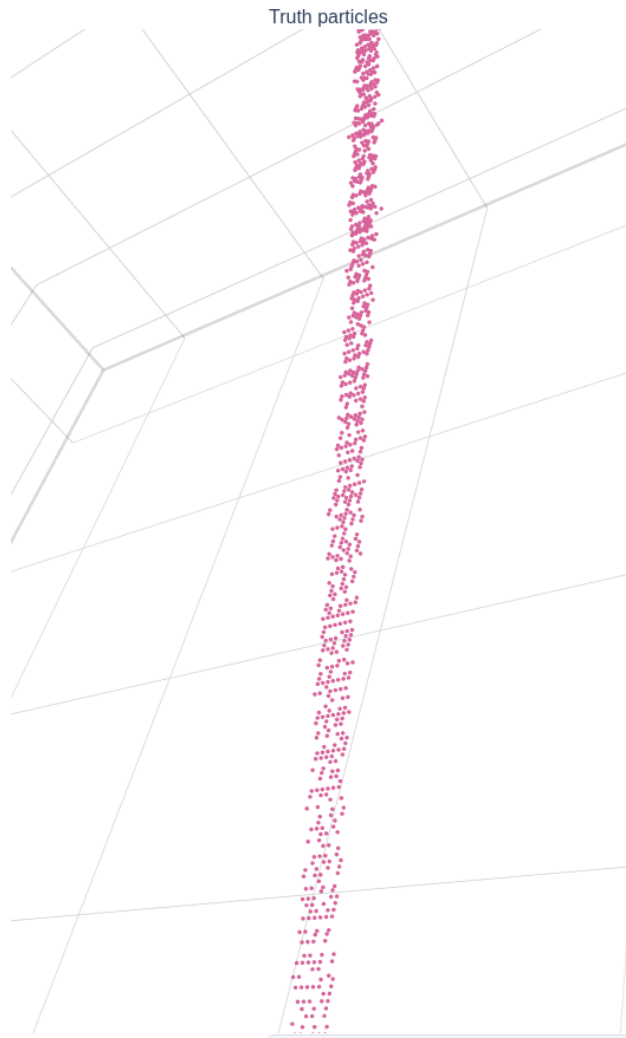


A Beam event!



PID needs improvement?!

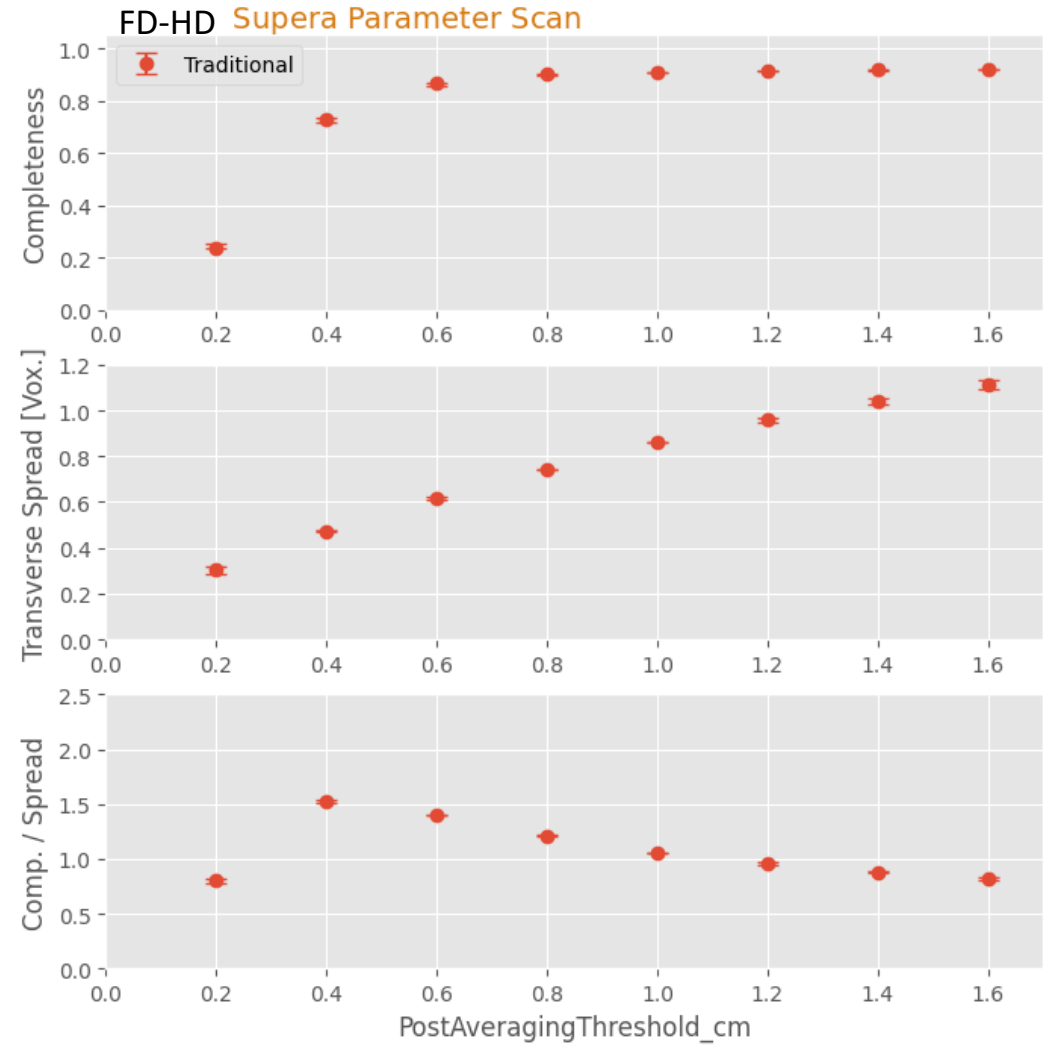
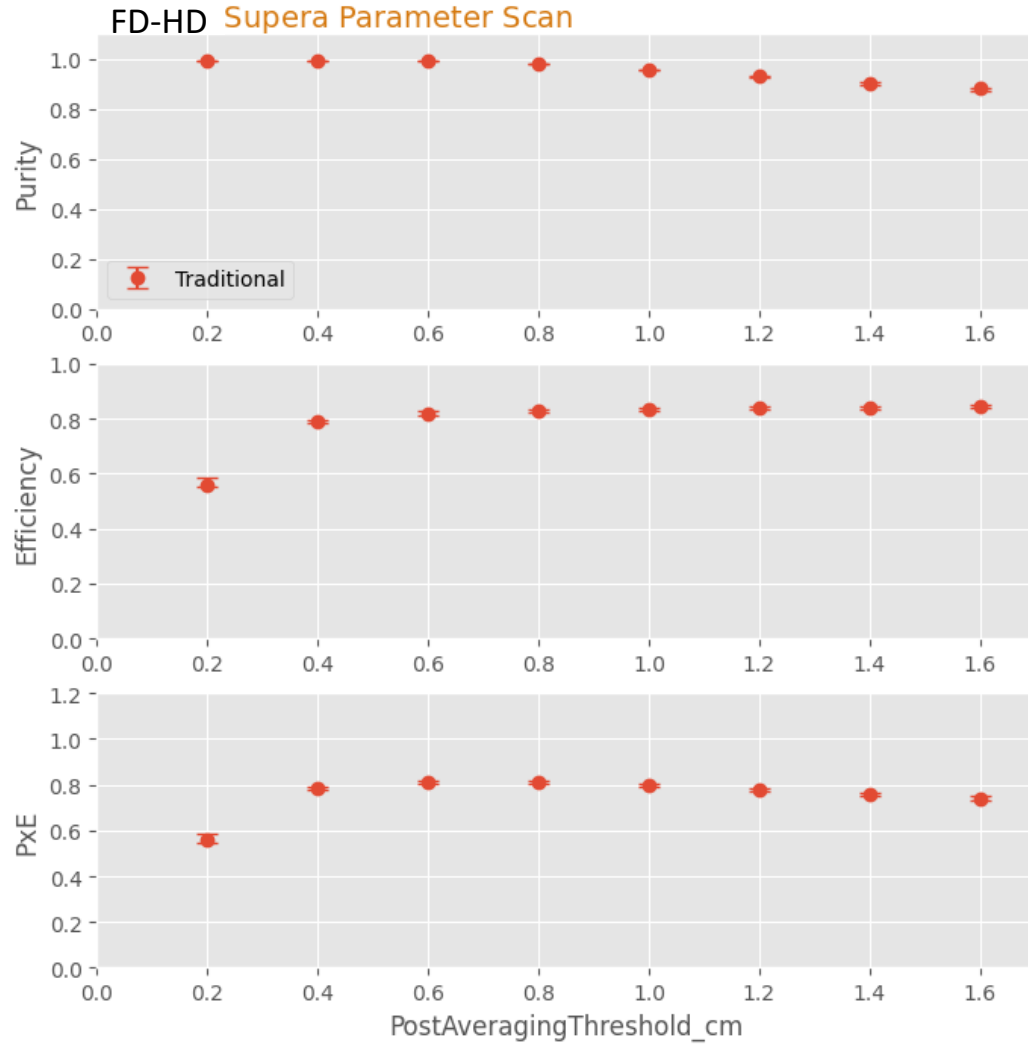
Supera parameter tuning - DUNE-FD example



- Initial training revealed clear room for improvement in model performance and stability.
- High track width and incompleteness required tuning of ghost labelling parameters before proceeding to the second training phase.

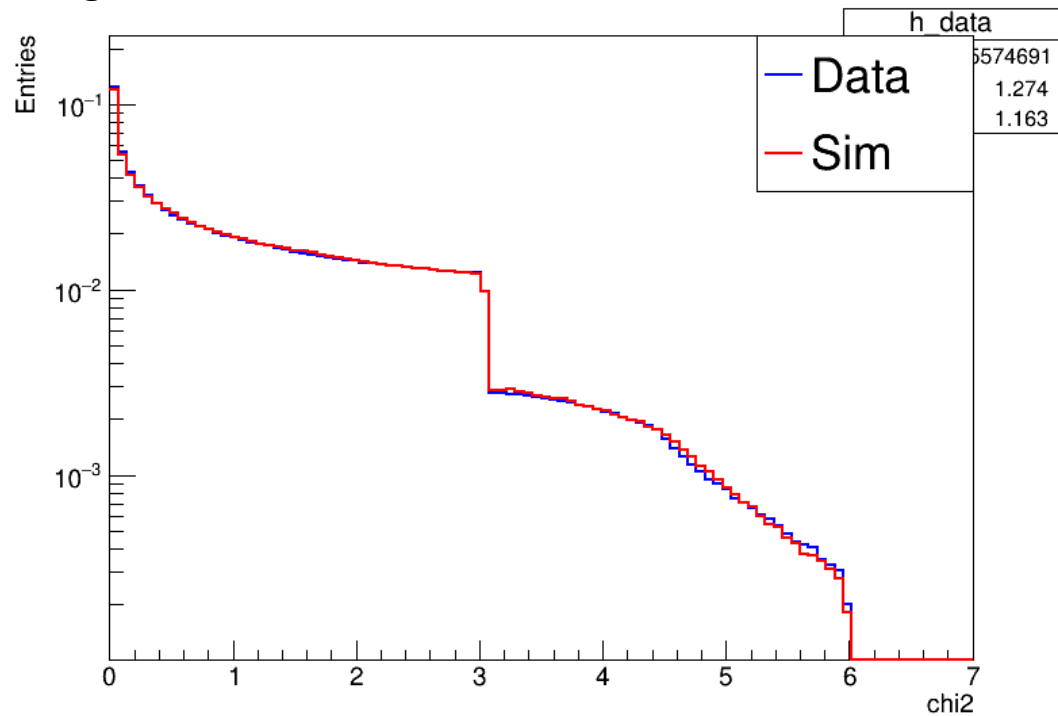
- **HitThresholdNe**: only reconstructed hits above this charge (electrons) are considered
- **HitWindowTicks**: reconstructed hits are tagged as non-ghost if matched to a true hit within this time window
- **PostAveragingThreshold_cm (optional)**: true energy is shared with nearby reconstructed voxels; all voxels within this distance are tagged as non-ghost

→ We tune these parameters across all detectors (FDHD, PDVD, PDSP) and find that the most influential parameter is the post-averaging threshold. An example of the FDHD results is shown below.

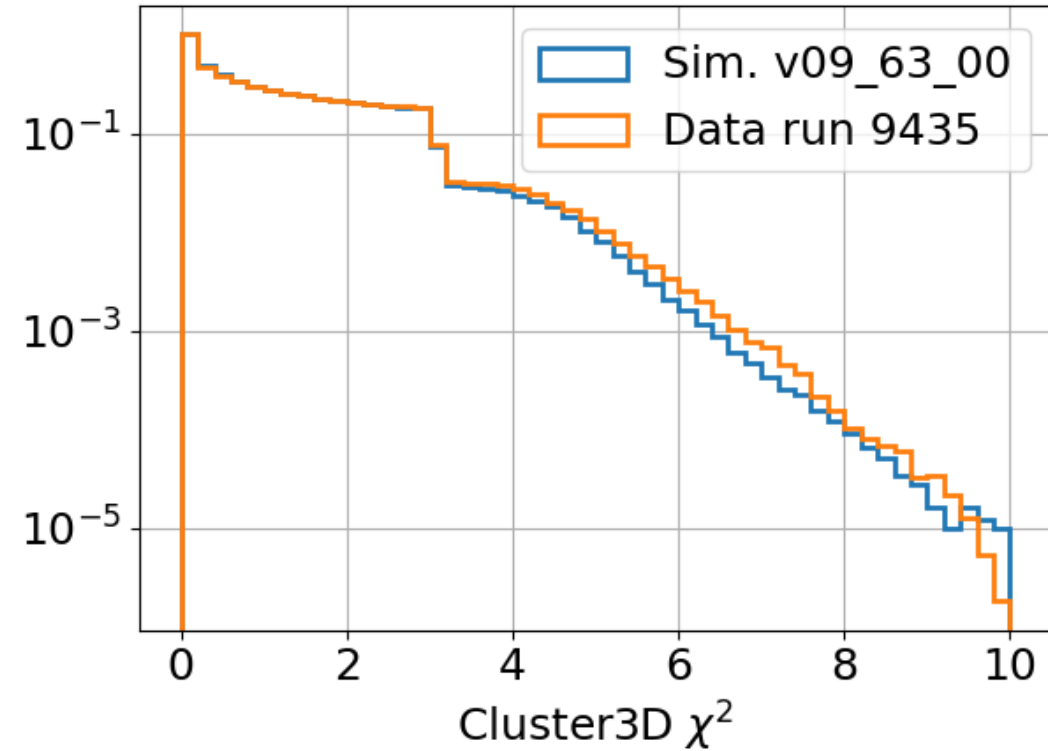


PDSP sim vs data

Here, we examine charge and χ^2 to assess the quality of our spacepoints in both data and simulation. A good overlap will give better data reconstruction.



PD-SP



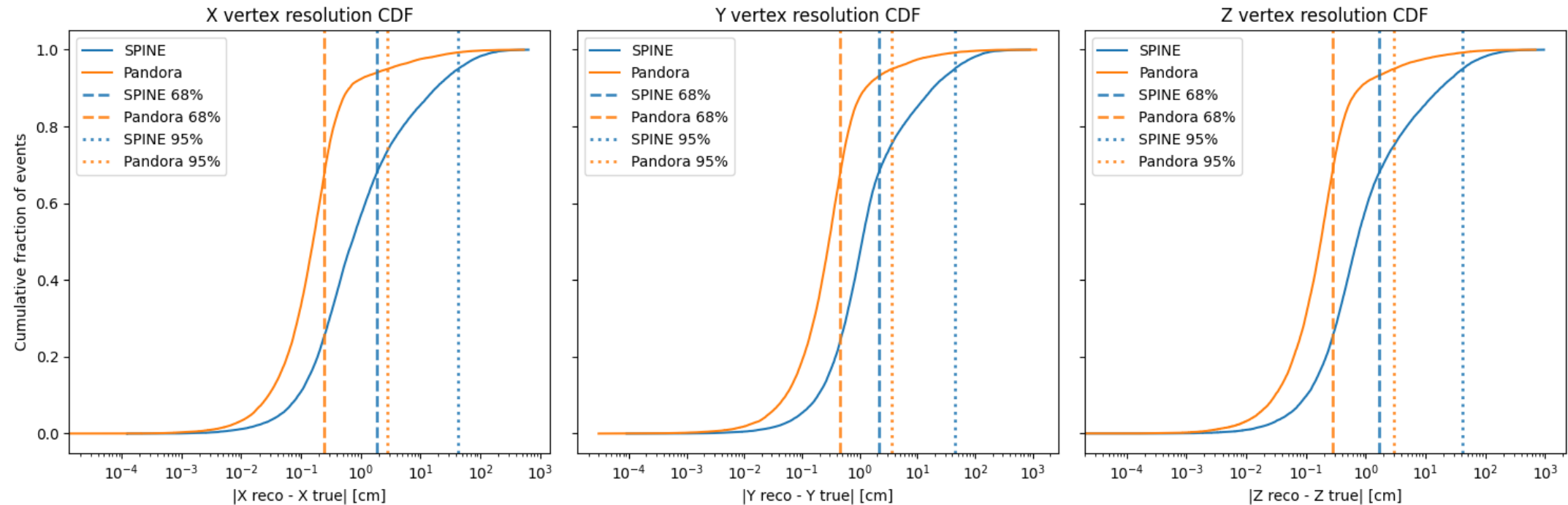
ICARUS

9

We conjecture that pdsp training should provide results as good as icarus

A peek into FDHD status

We study the vertex resolution by comparing SPINE and PANDORA using a small sample of atmospheric neutrino simulations.



Credit : Pierre granger

The vertex matching in FDHD using SPINE is currently performing worse than in Pandora. This is expected, as it is likely due to suboptimal ghost labelling, which we plan to address in the next training iteration. Additionally, SPINE uses an IOU-based matching method rather than vertex proximity, so this difference has minimal impact on our final results.

Next Steps

- Getting ready for retraining → FDHD samples are in production
- PDVD/PDSP files are going into production soon
- A large beam+cosmic simulation will be in production for PDSP and PDVD for accurate assessment of SPINE.
- Integrate spine inference step into standard reco --> will need gpus
- Mitigating the PDHD - APA1 issue for training

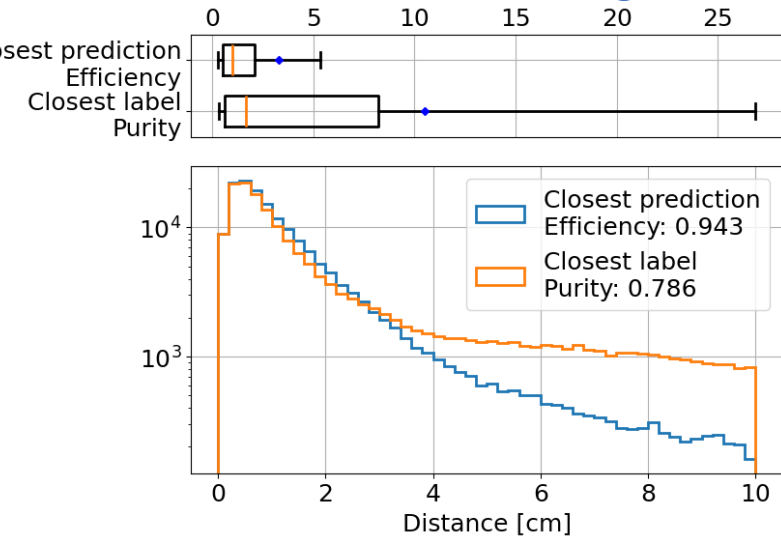
THANK YOU FOR LISTENING!

SPINE

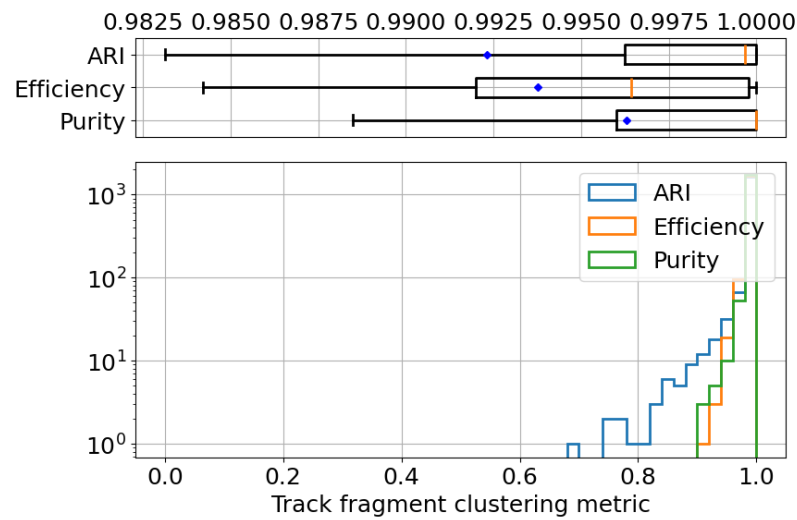


BACKUP SLIDES

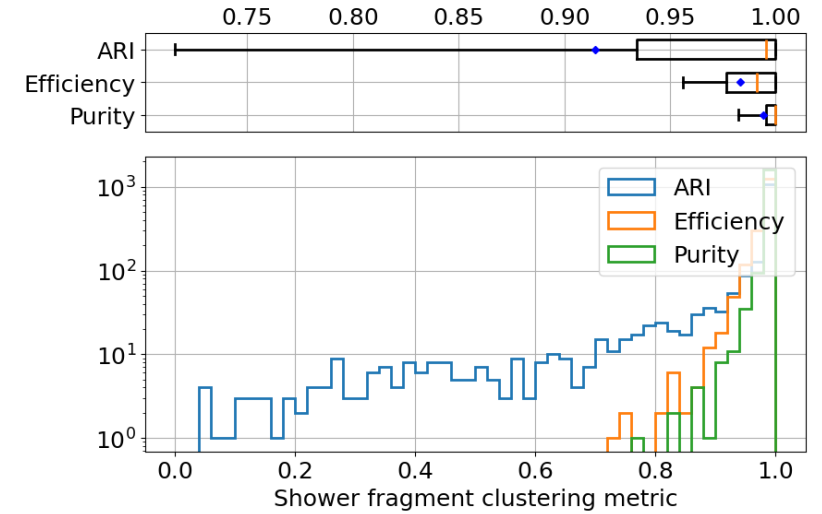
PPN - vertex finding



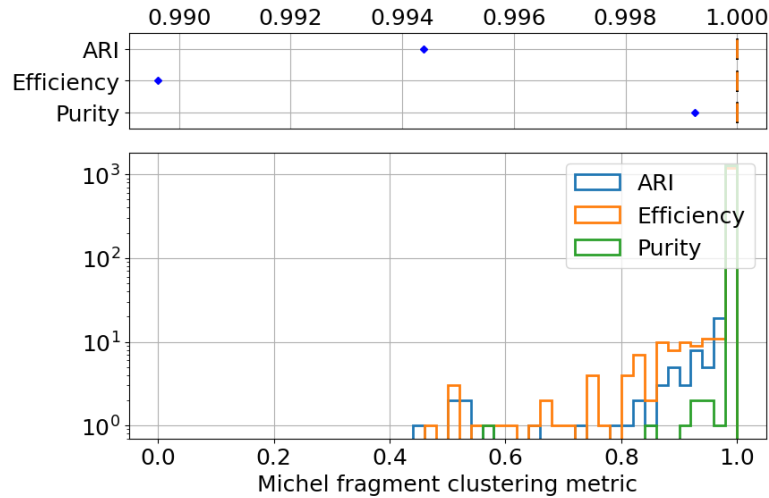
Fragment clustering - tracks



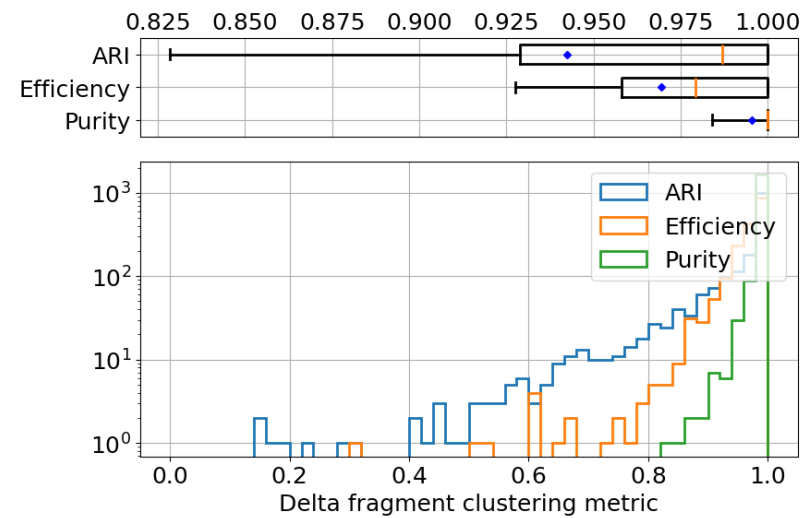
Fragment clustering - shower



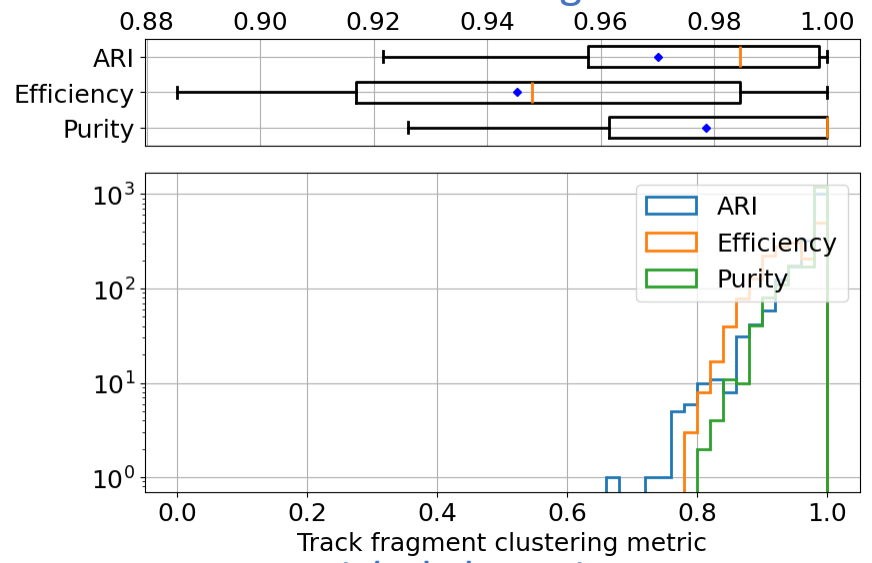
Fragment clustering - michels



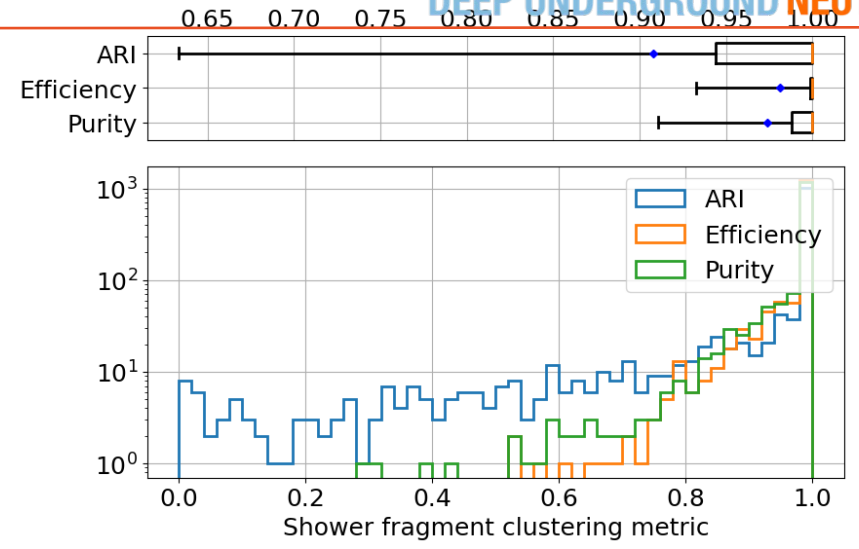
Fragment clustering - delta



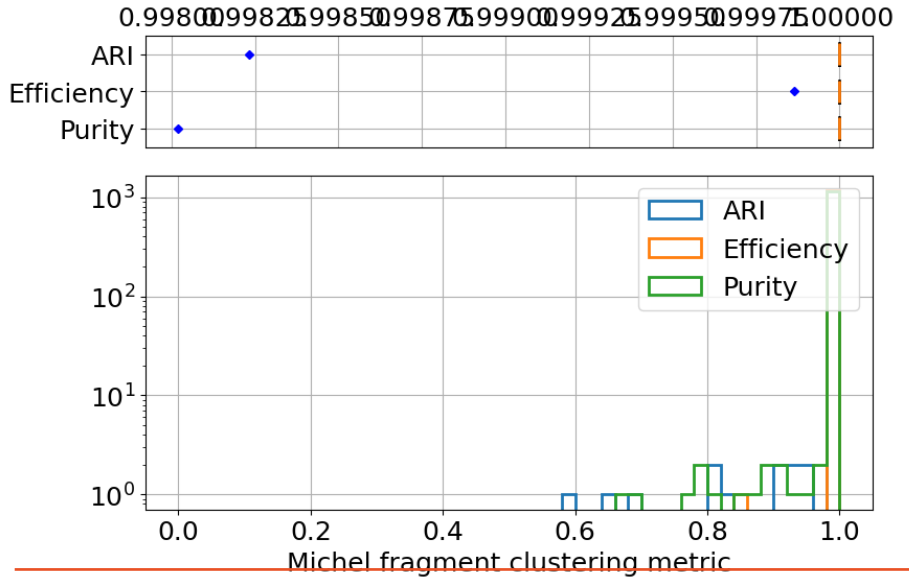
Track clustering



shower clustering



micel clustering



delta clustering

