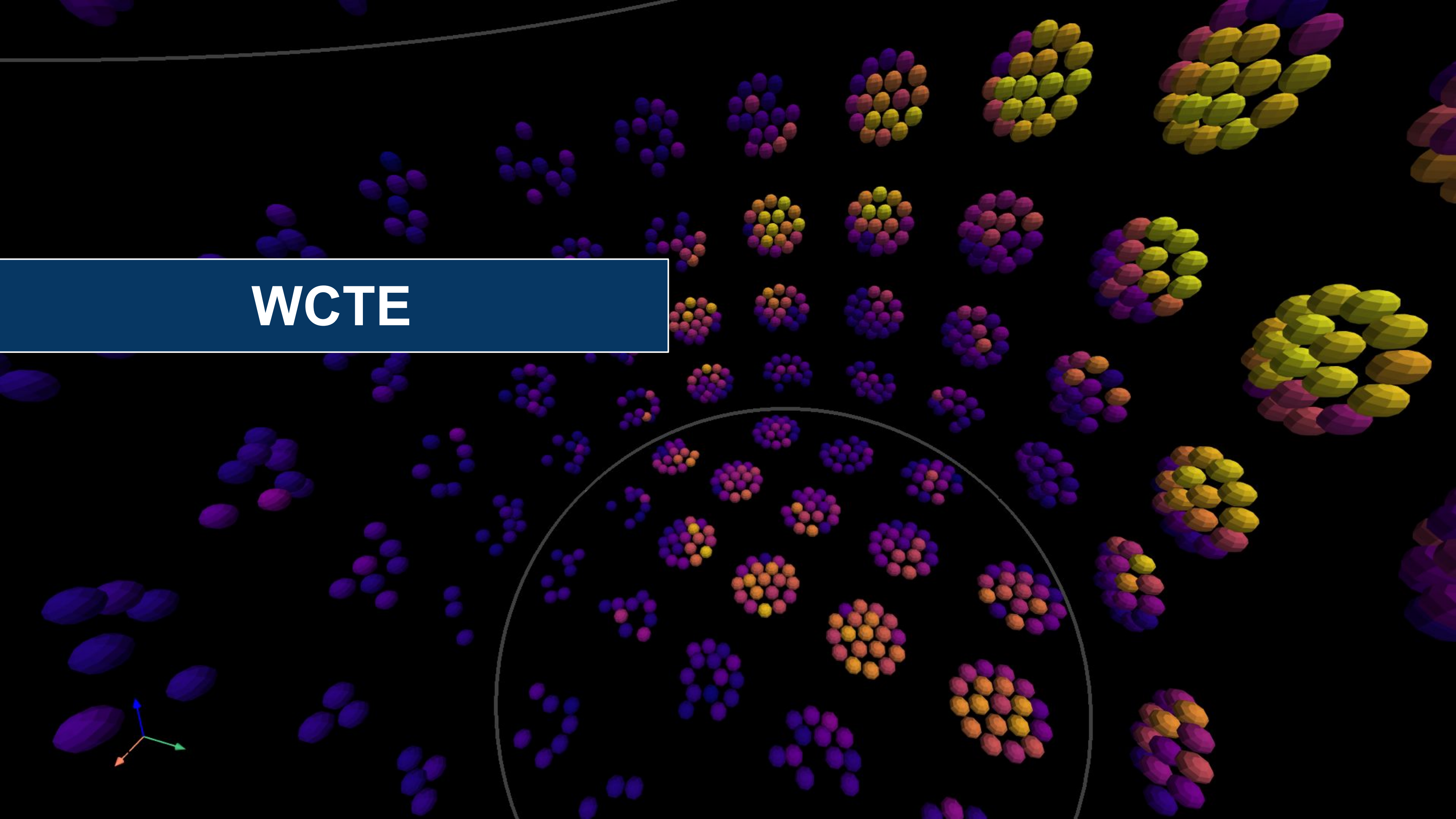
IRN Orsay
02/06/2026

The Water Cherenkov Test Experiment

A Playground for Machine Learning

Mathieu Ferey

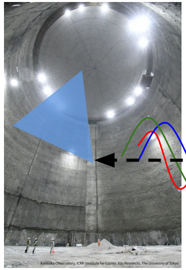
WCTE

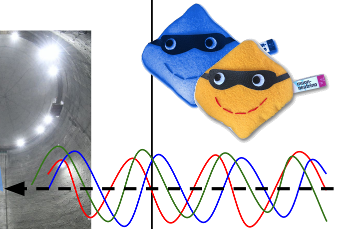


Hyper Kamiokande

Hyper Kamiokande

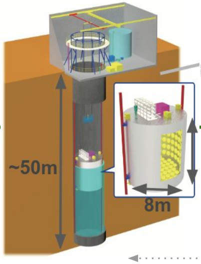
measure neutrino oscillation

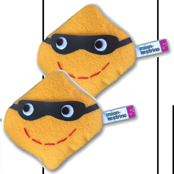




Intermediate Water Cherenkov Detector

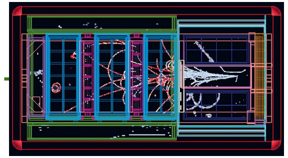
constrain water cross-sections on a wide energy spectrum

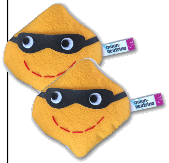




Near Detector

constrain cross-sections and flux

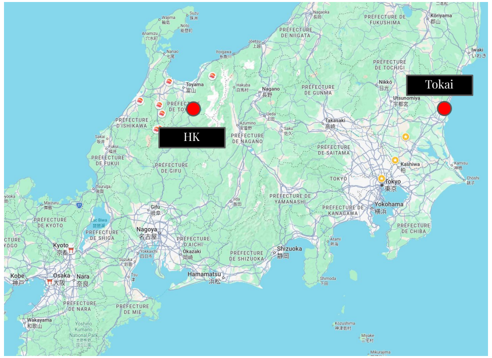




~300m

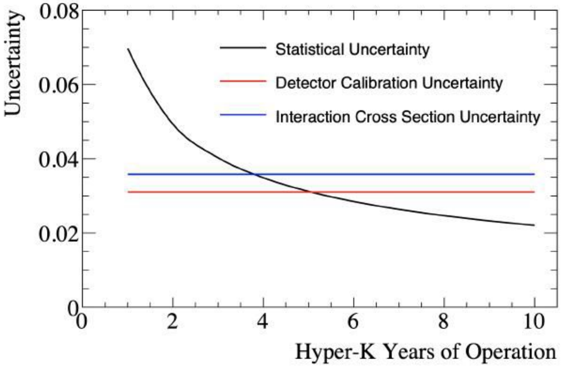
~1km

~300km



Systematic uncertainties are expected to dominate Hyper-Kamiokande CP violation analysis in ~5 years

Reduction of systematics to the percent level is crucial!



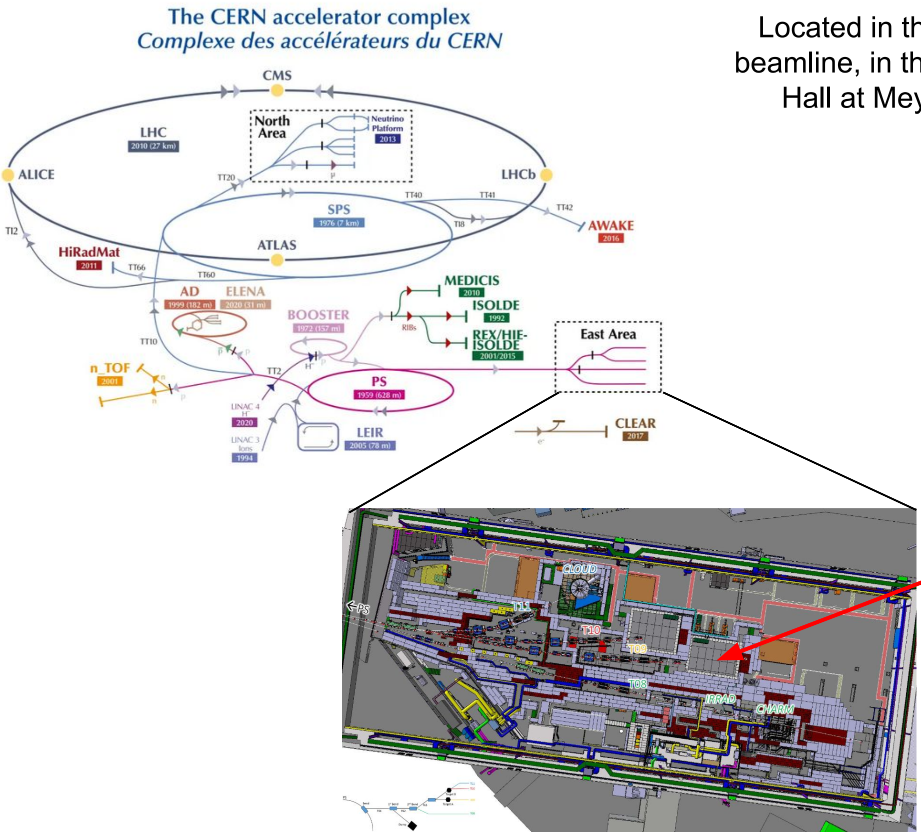
WCTE is a prototype for IWCD

- test new high granularity mPMTs
- new calibration techniques
- precise estimation of Water Cherenkov detector response
- explore new reconstruction methods

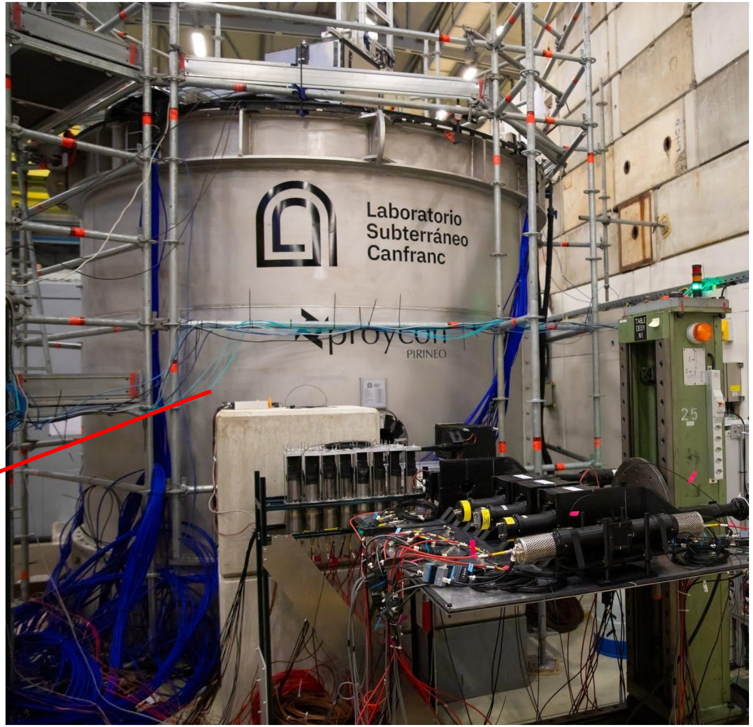
The detector

- 3.76 x 3.52m cylindrical tank
- 96 multi-PMTs - each with 19 3” PMTs
- mPMTs equipped with LEDs for timing calibration
- mechanical arm for precise use of calibration sources

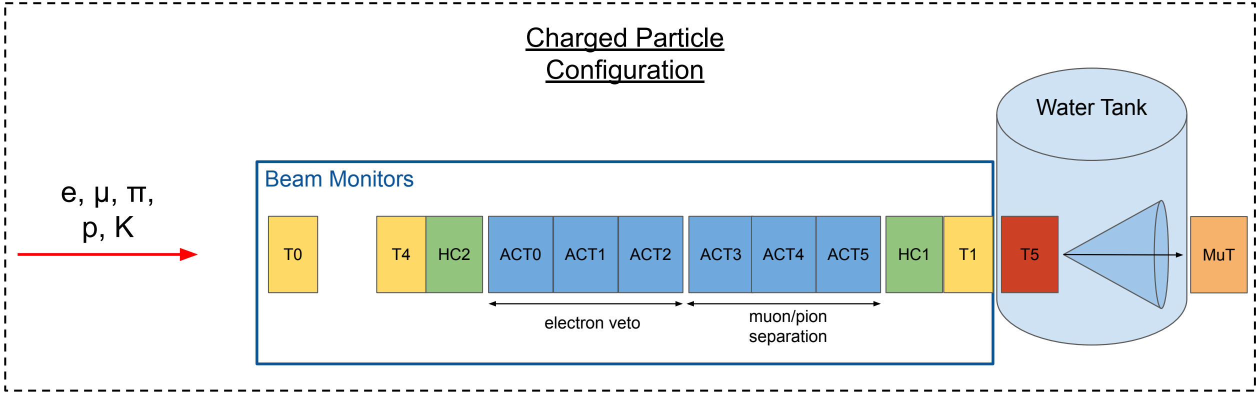
The Water Cherenkov Test Experiment

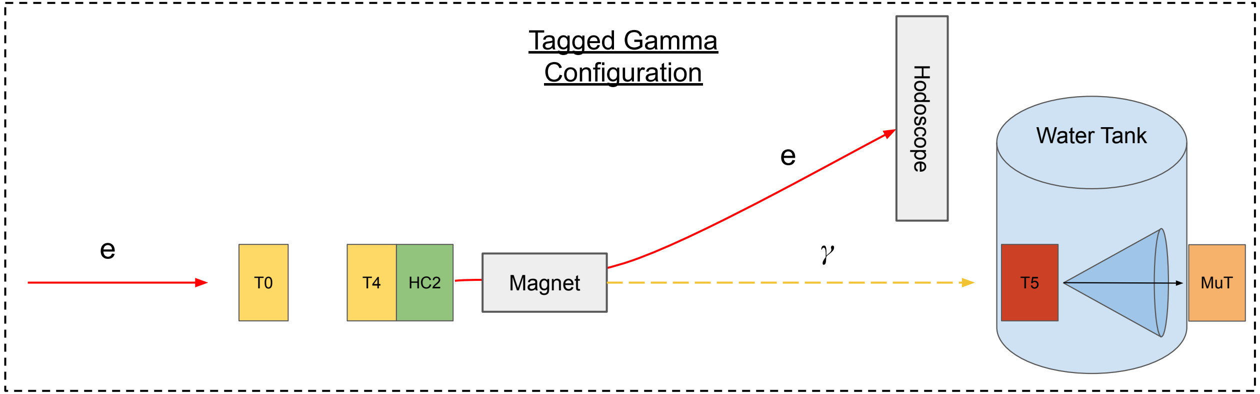


Located in the T9 beamline, in the East Hall at Meyrin



The WCTE beamline





- Protons accelerated in CERN's Proton Synchroton are shot on different targets, producing a beam of charged particles, ranging from ~0.2-1.2 GeV
- Beam monitors tasked with estimating particles PID and momentum before they enter the tank
- Study charge particles directly, or use a magnet to emit Bremsstrahlung gammas into the tank

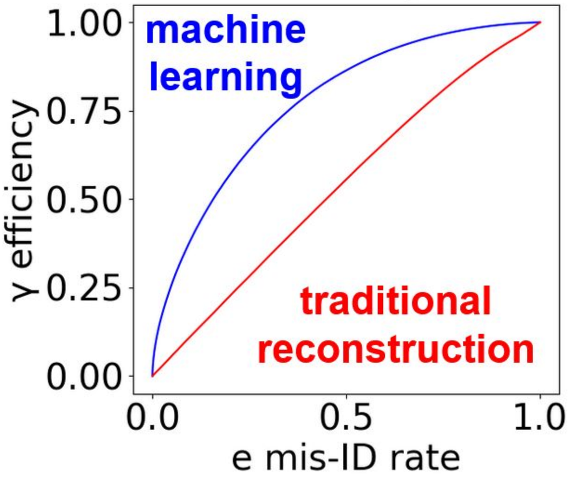
Machine Learning for Reconstruction

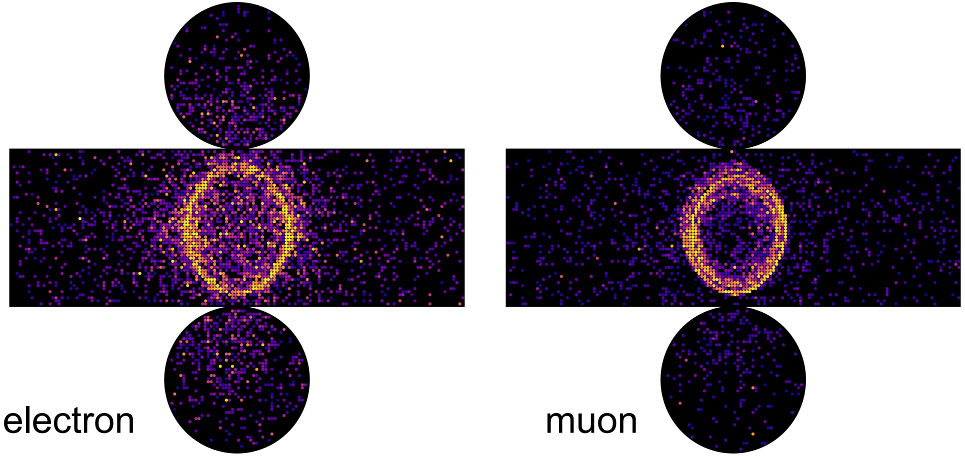
Traditional Reconstruction Algorithms still struggle:

- e/γ separation: central for NC background reduction
- μ/π classification
- low energy events
- events too close to the wall: loss of crucial fiducial volume
- multi-ring events
- evaluation is slow

Can Machine Learning improve challenging reconstruction tasks?

Nick Prouse

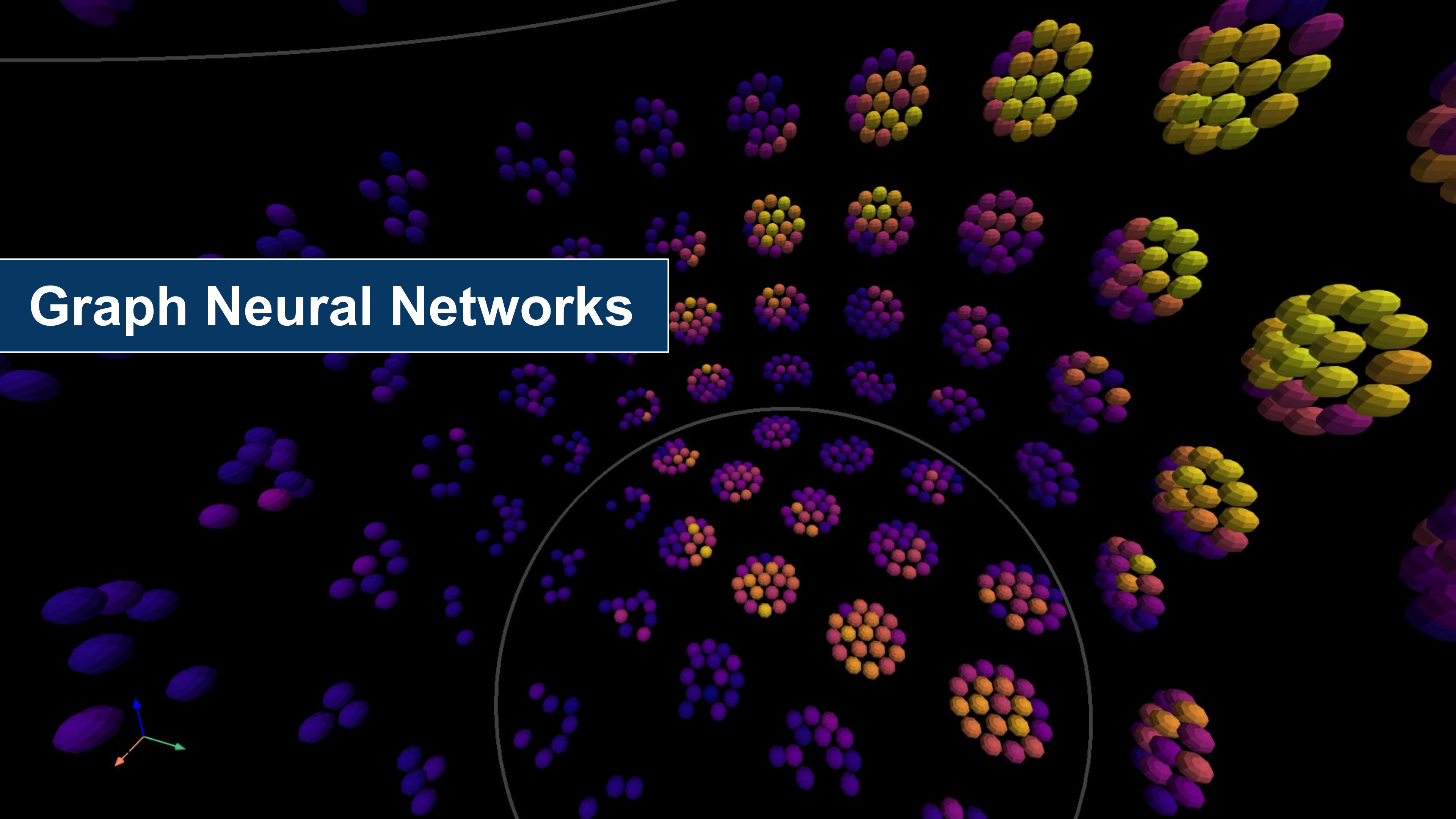


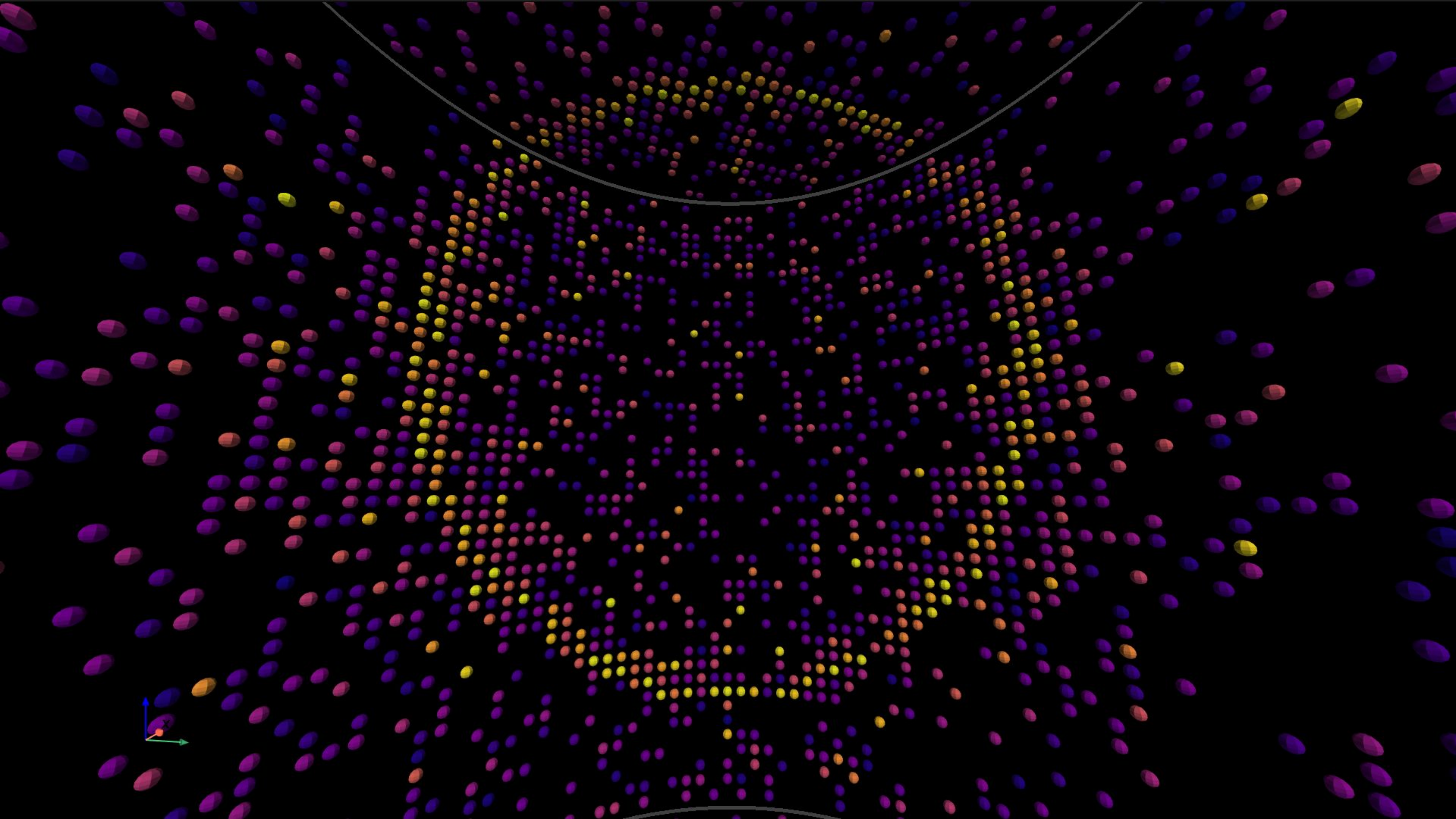


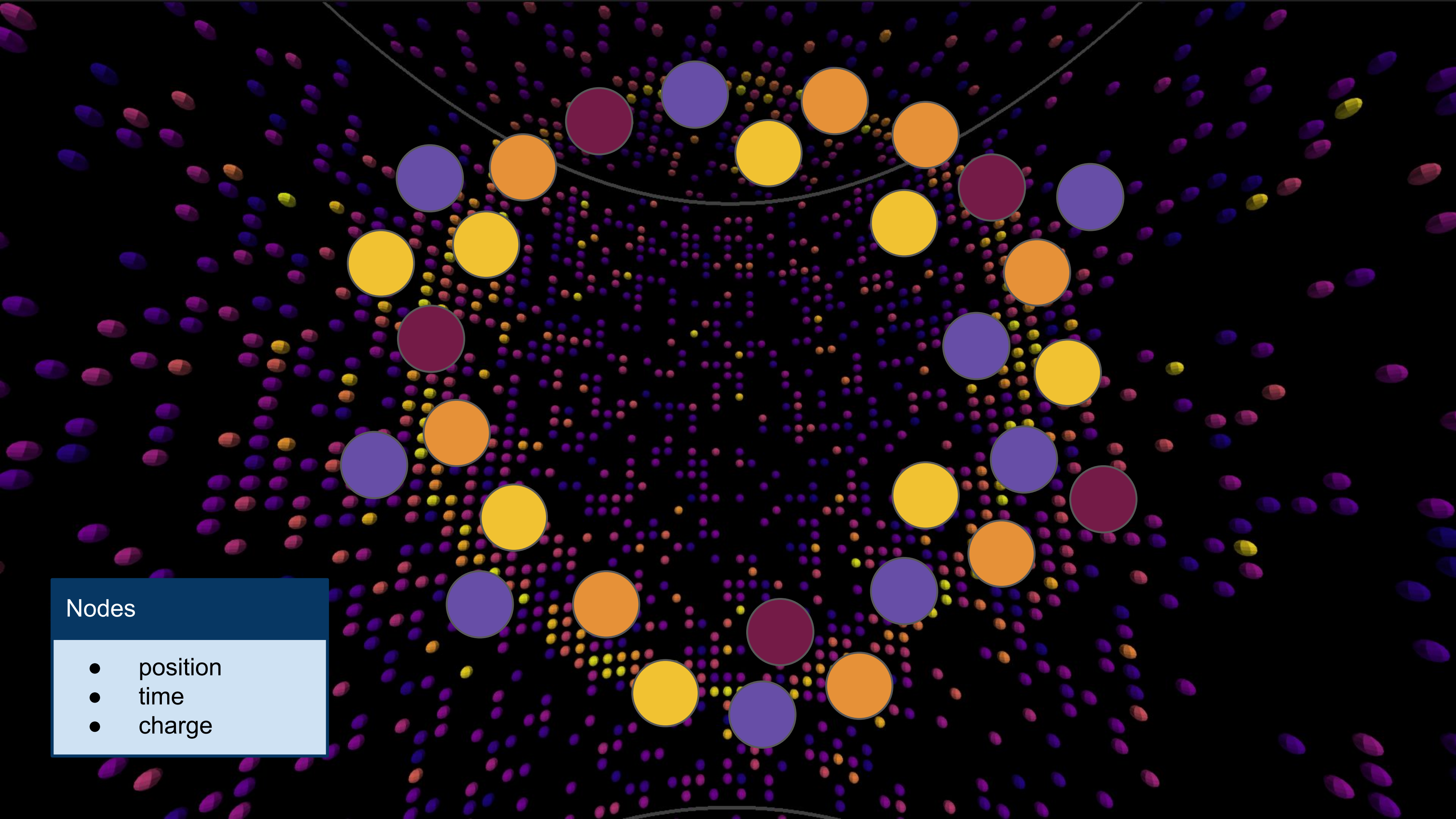
WCTE is a unique opportunity to develop and test ML reconstruction algorithms

- small detector: simulations are lightweight and training is fast
- availability of well understood beam data will allow to test the robustness of our algorithms

Graph Neural Networks

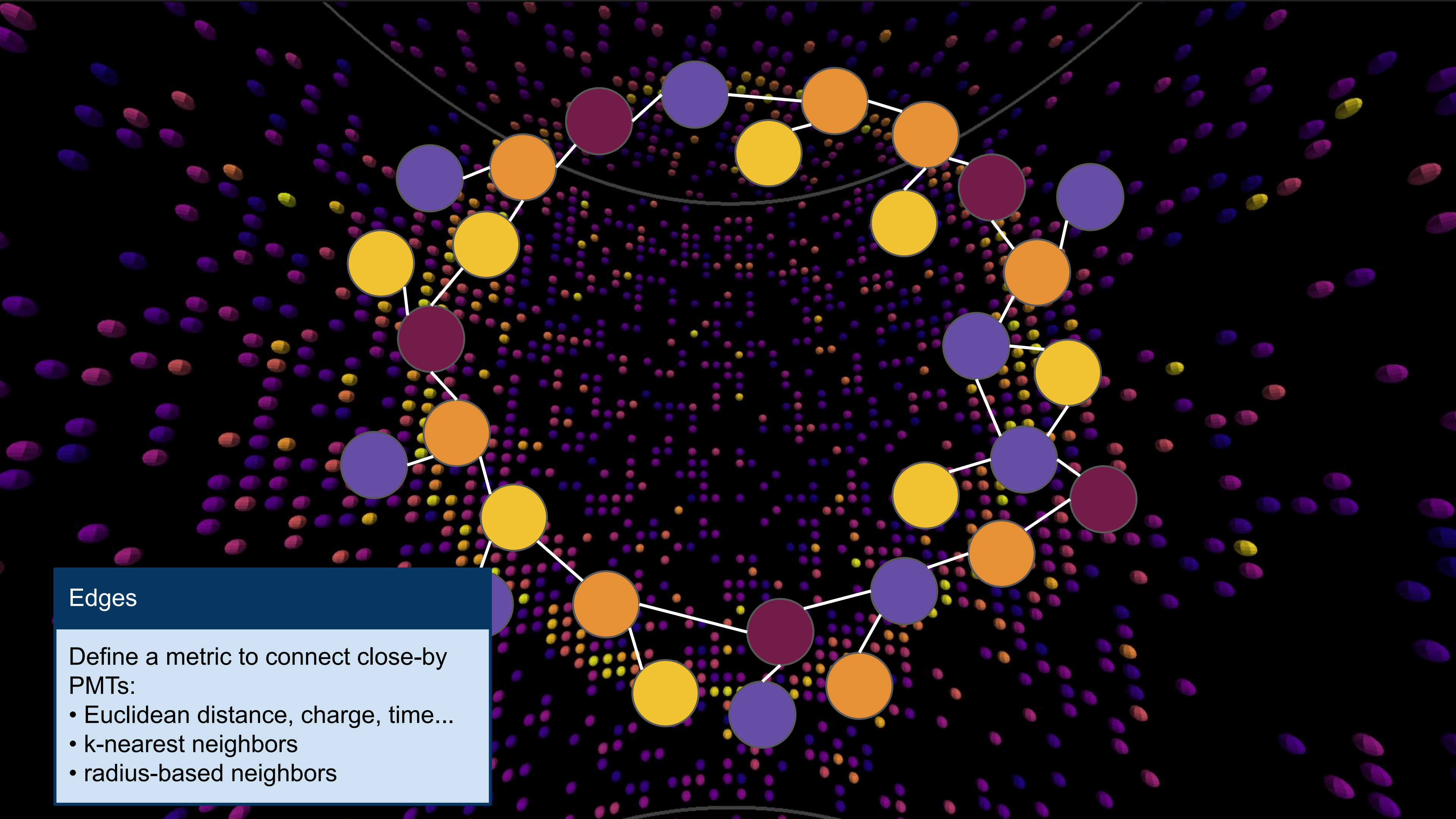






Nodes

- position
- time
- charge

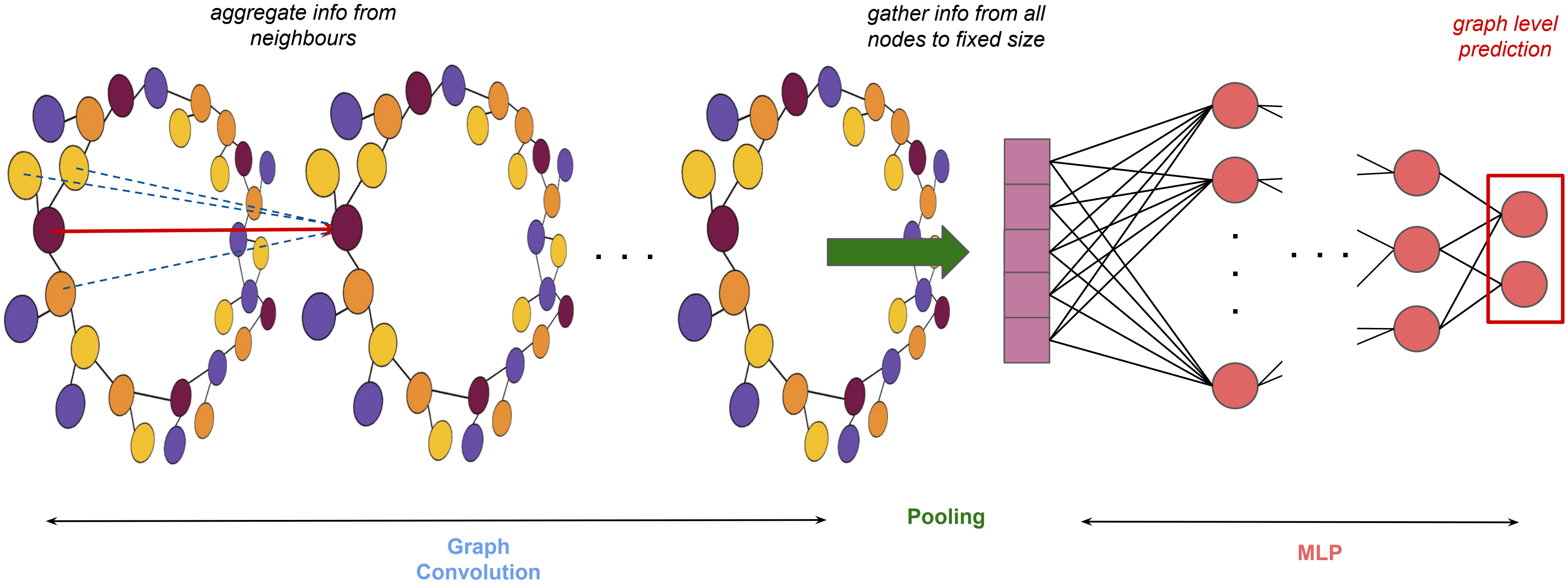


Edges

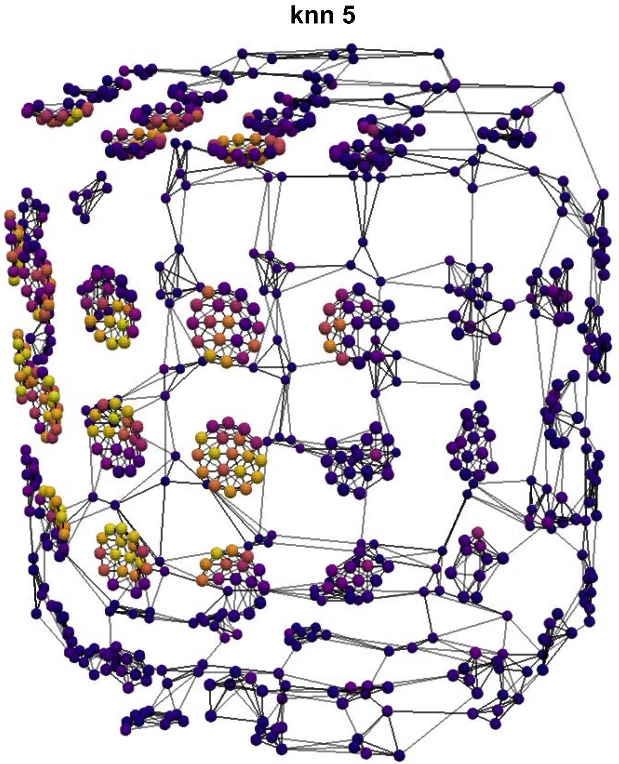
Define a metric to connect close-by PMTs:

- Euclidean distance, charge, time...
- k-nearest neighbors
- radius-based neighbors

GNN: how does it work



The mPMT problem



Because of their granularity, mPMTs in the main ring are isolated in the graph:

→no information is passed between PMTs!

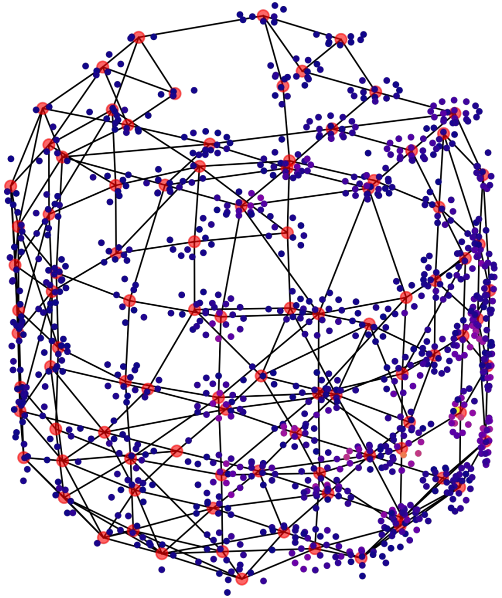
Increasing number of neighbours becomes rapidly intractable computing-wise

Hierarchical graphs

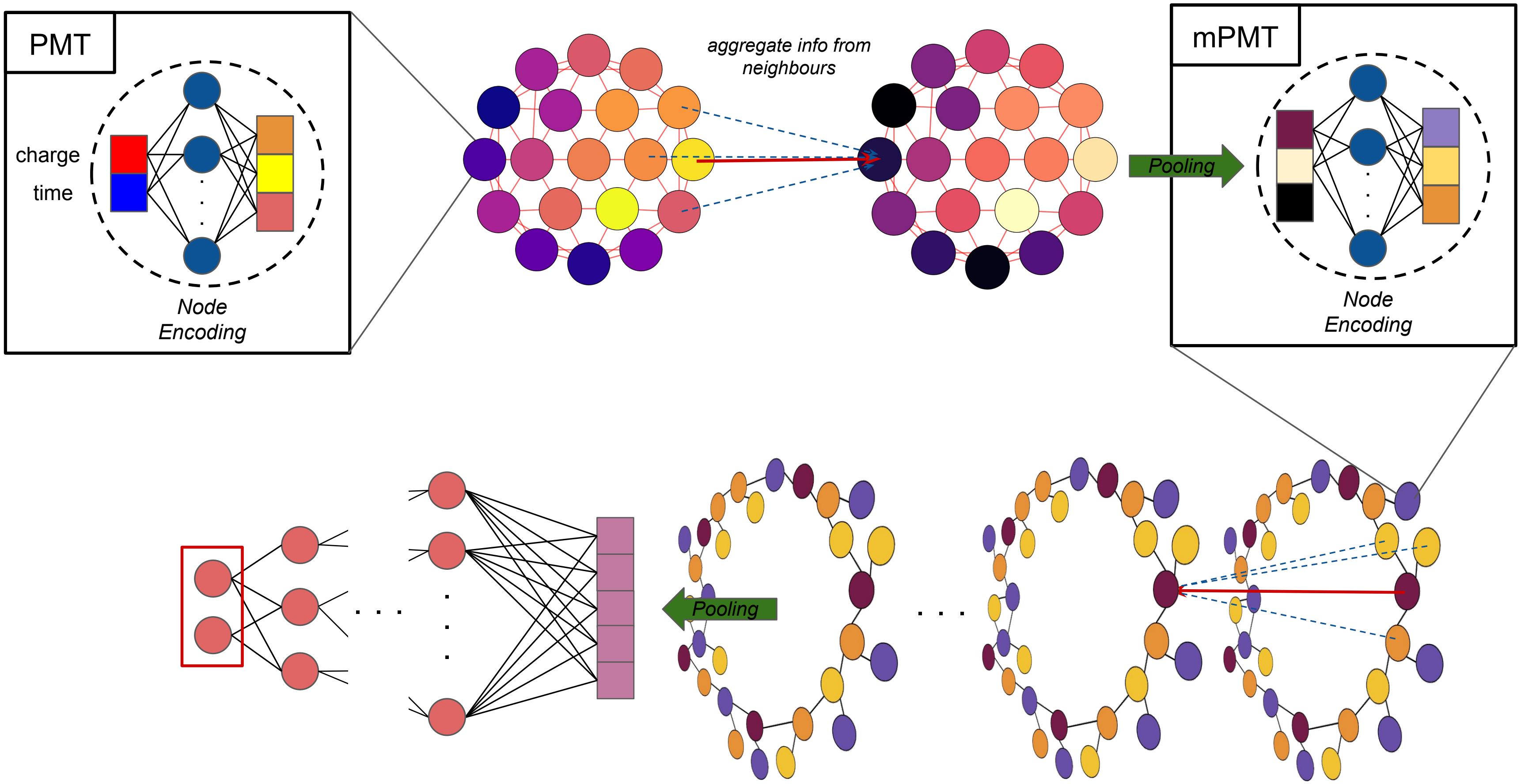
Instead, have two levels of graphs

Intra-mPMT
captures local information

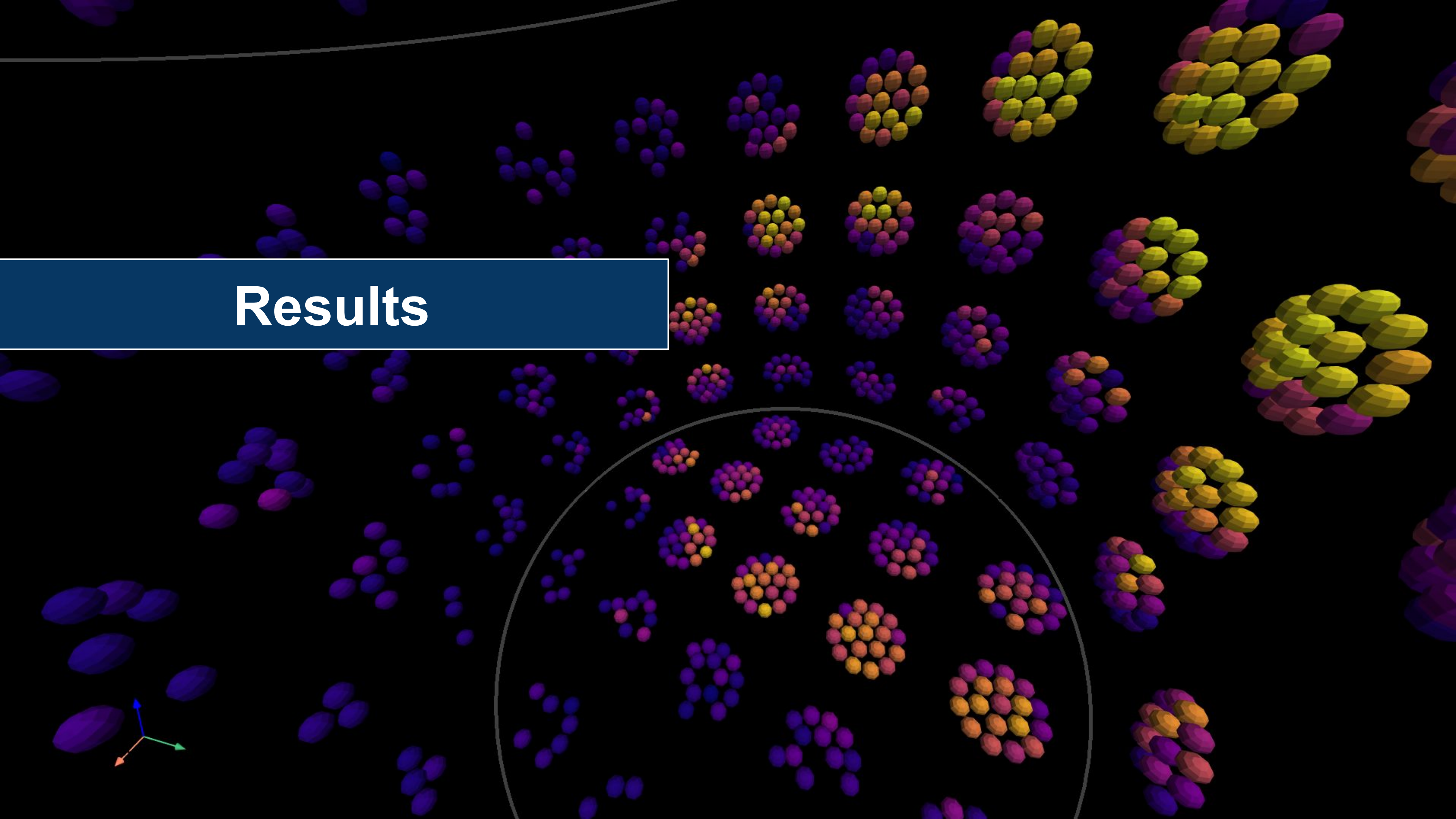
Inter-mPMT
captures global information



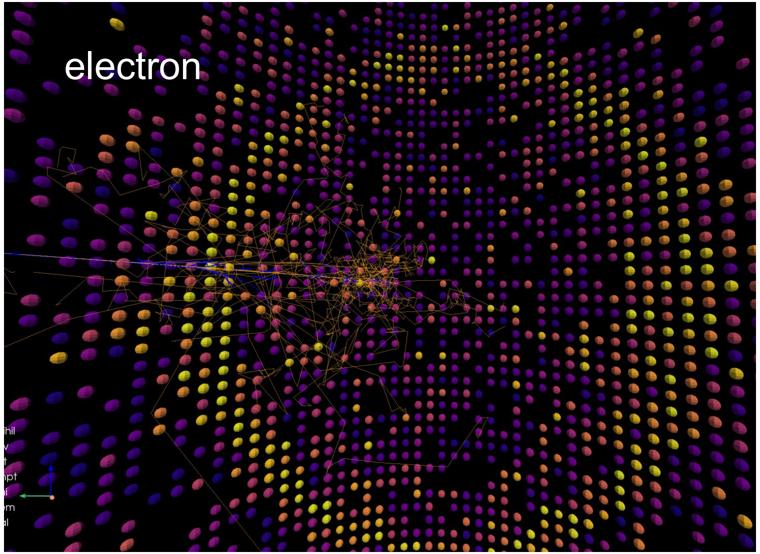
Hierarchical GNN



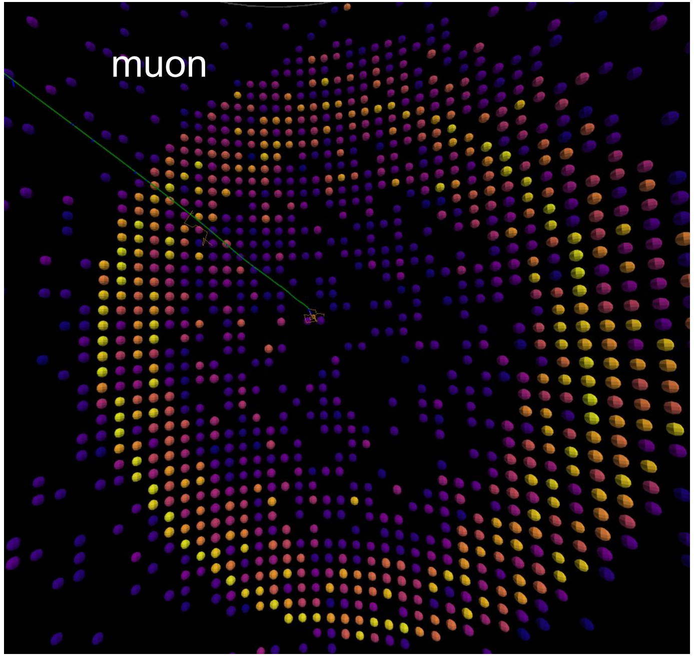
Results



Classification



Electron/Muon classification as a simple task to showcase GNN capabilities

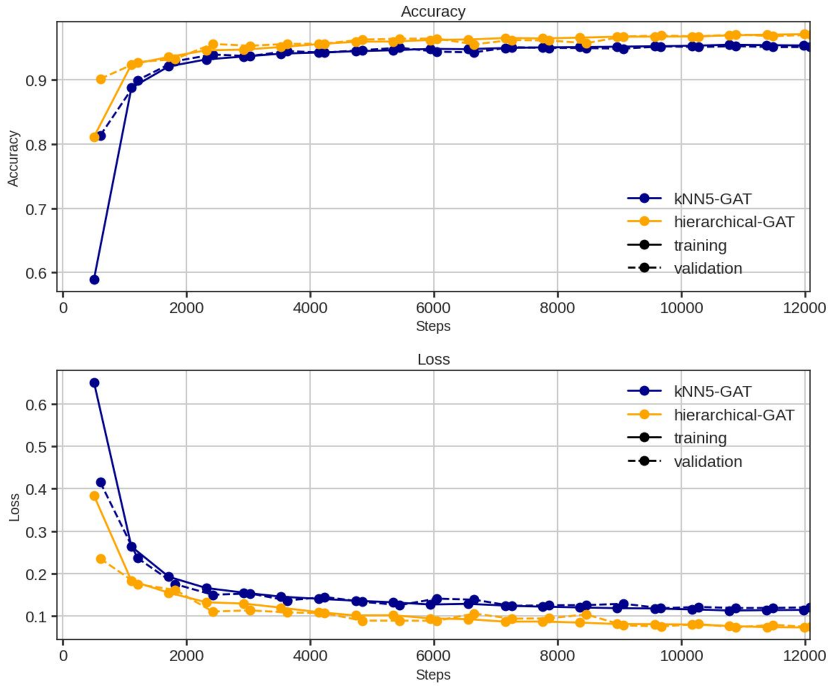


250,000 electrons, 250,000 muons

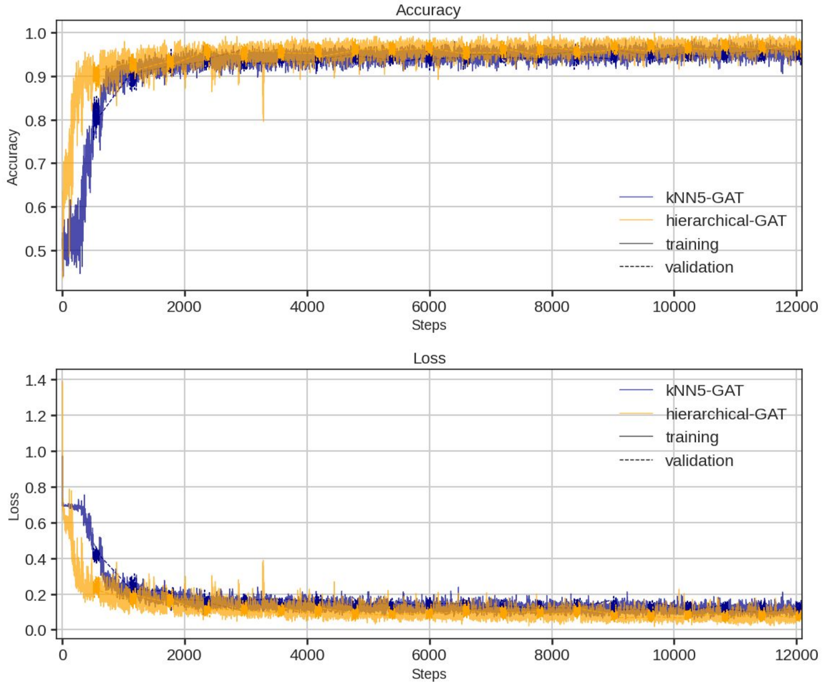
- training: 250,000
- validation: 50,000
- test: 200,000

Machine learning runs

Per epoch

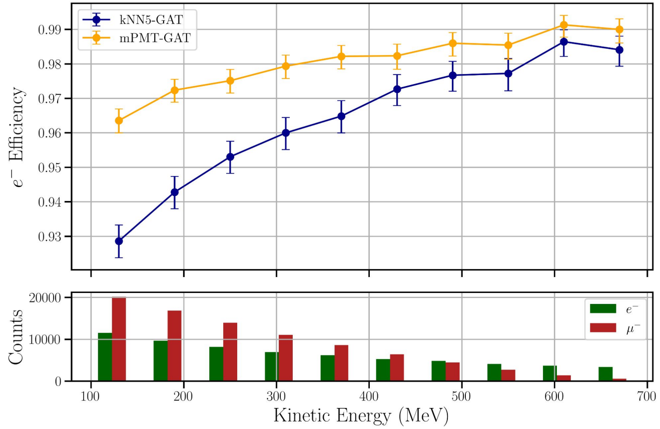


Per batch



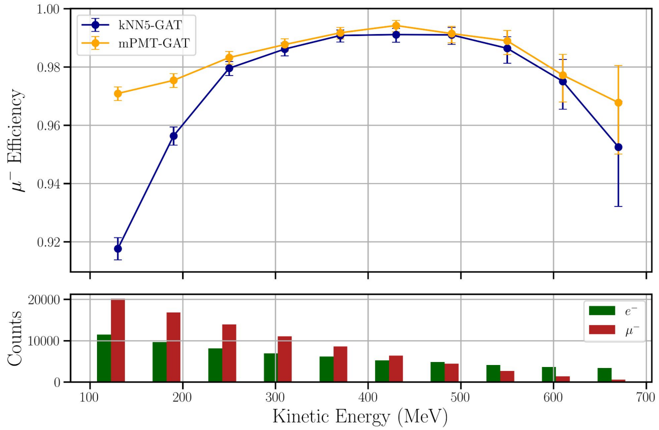
Training is stable and converges. Still prone to overfitting, so has to stop early

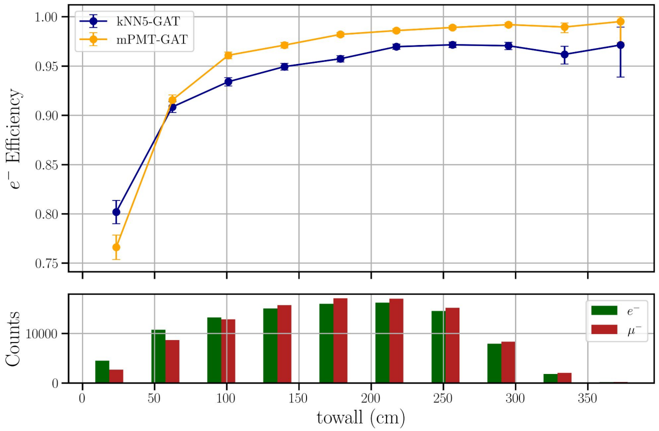
Efficiencies



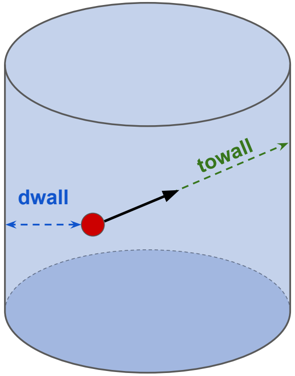
low energy or **low towall**

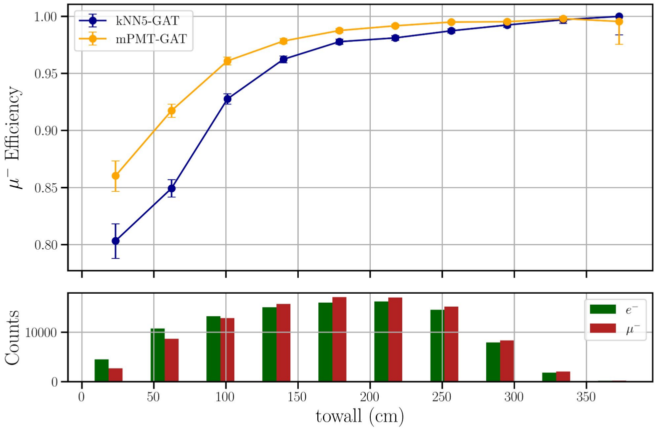
less hits, so less information,
worse performance



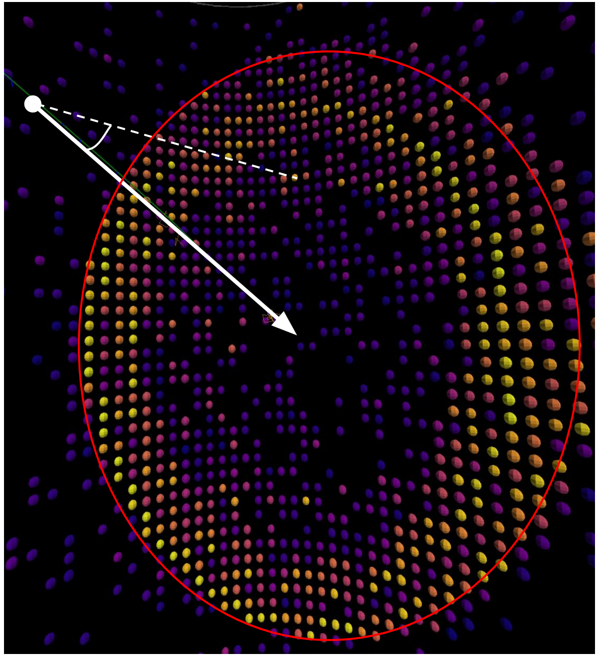


high energy muons:
what is happening?



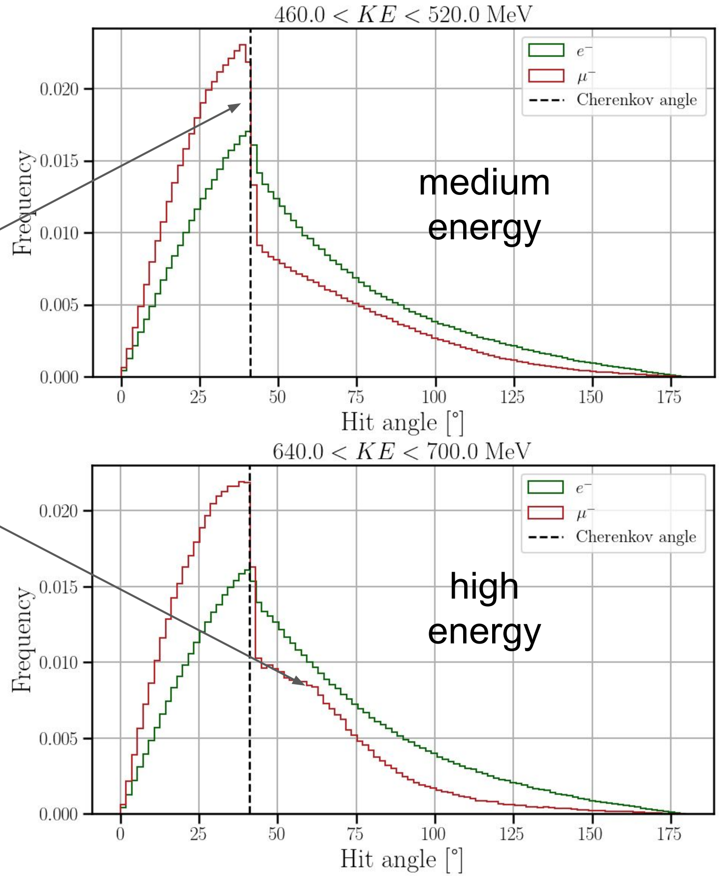


High Energy Muons



sharp muon ring at 41°

enhanced reflection shoulder at high energy



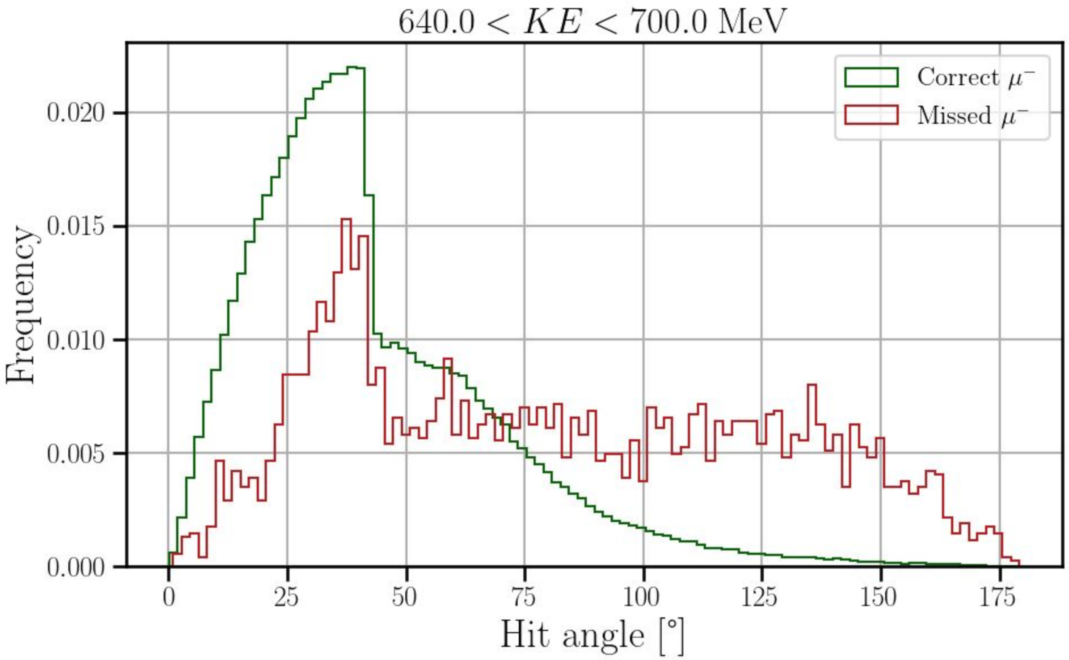
Muons loose little energy to ionisation

- long and straight path

Electrons experience a lot of bremsstrahlung and shower

- jagged a short path

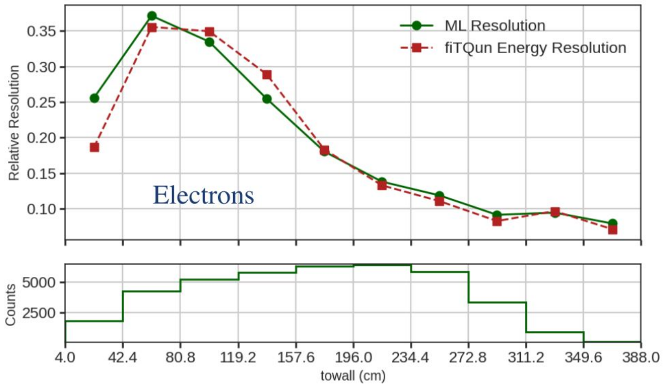
High Energy Muons

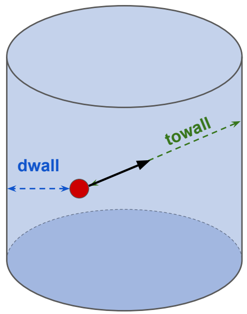


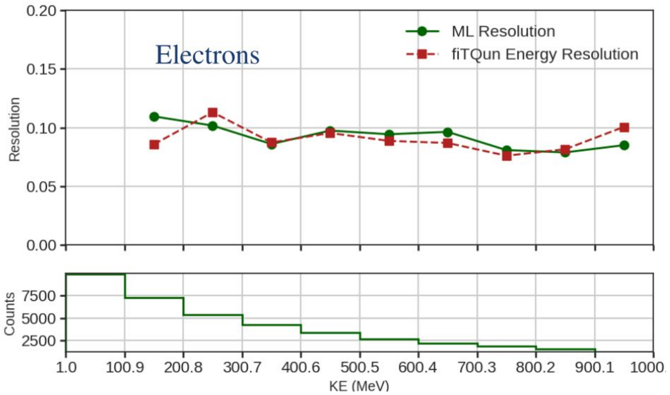
At high energy, miss-id muons show a very large reflection tail

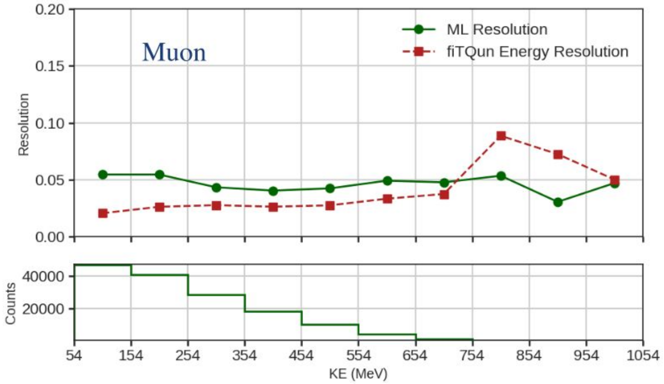
Energy resolution

Results by Viet Quoc NGUYEN

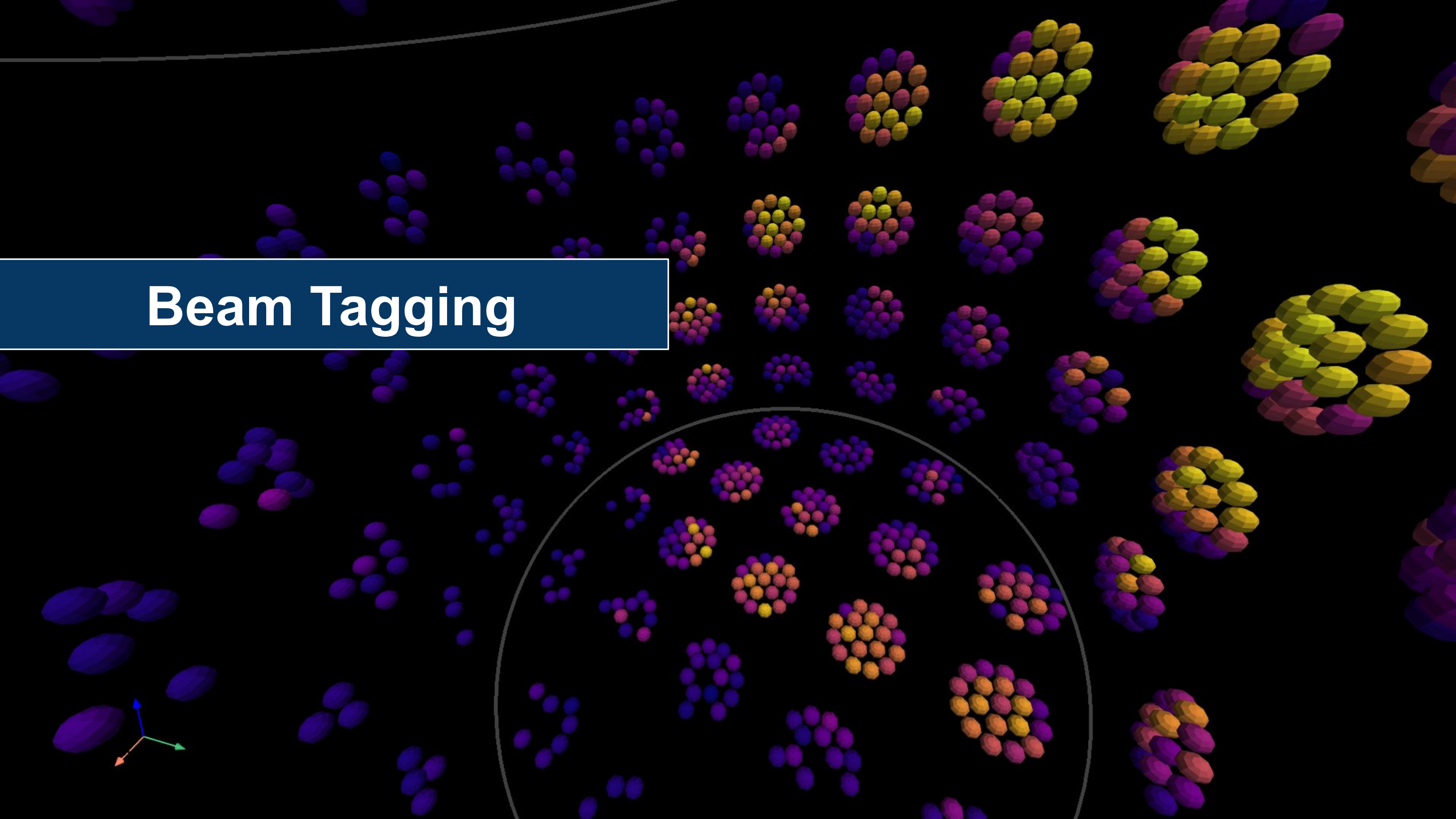








Beam Tagging



The beam monitors - charged particle configuration

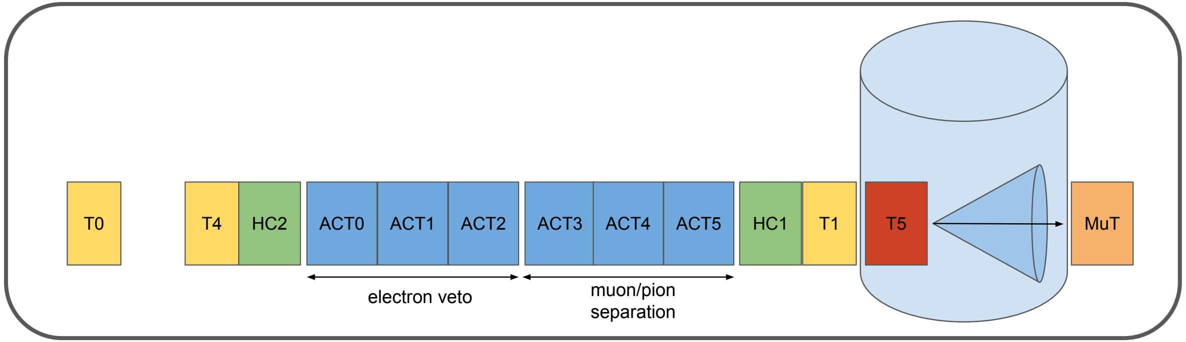
T0, T1, T4

Timing detectors:
- plastic scintillator coupled to 2/4 PMTs
- triggering and Time of Flight measurements

photos by B. Ferrazzi

ACT

Aerogel Cherenkov Threshold:
- gel with given refractive index, paired with two PMTs
- particle identification



T5

TOF detector:
- segmented plastic scintillator bars read by SiPMs
- triggering and Time of Flight measurements

HC

Halo Counters:
- plastic scintillator with a central hole, paired to a PMT
- veto particles deviating from the beam spot

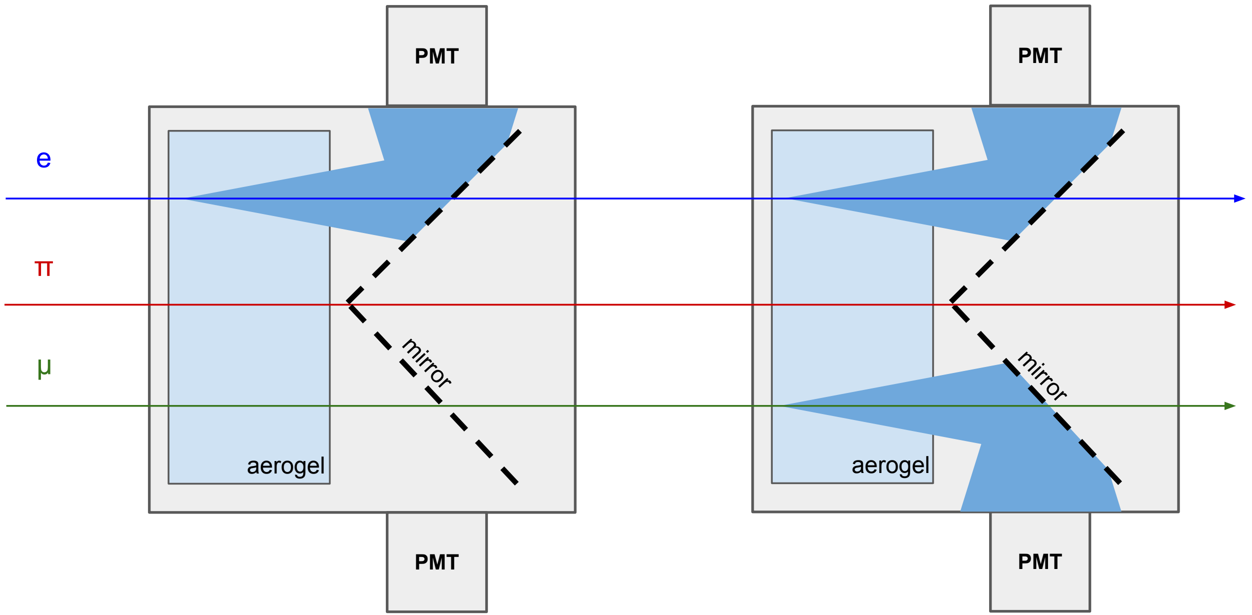
MuT

Muon Tagger:
- two scintillator plates equipped with PMTs
- tag high energy through going muons

Particle Identification

ACT0-1-2
electron veto

ACT3-4-5
muon/pion tagging

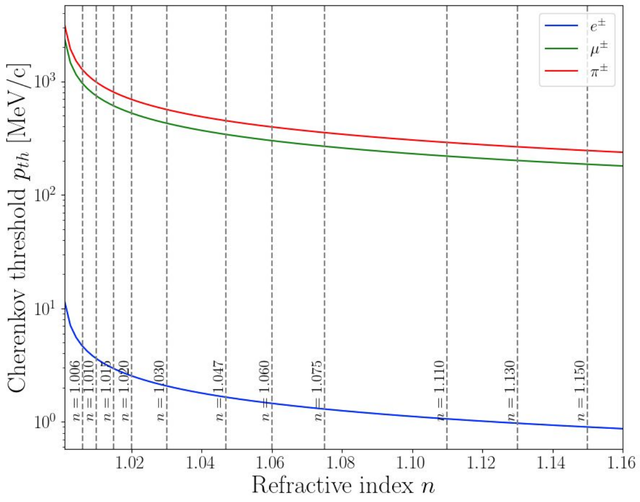


Choose index for electrons to be above threshold

Choose index for electrons and muons to be above threshold

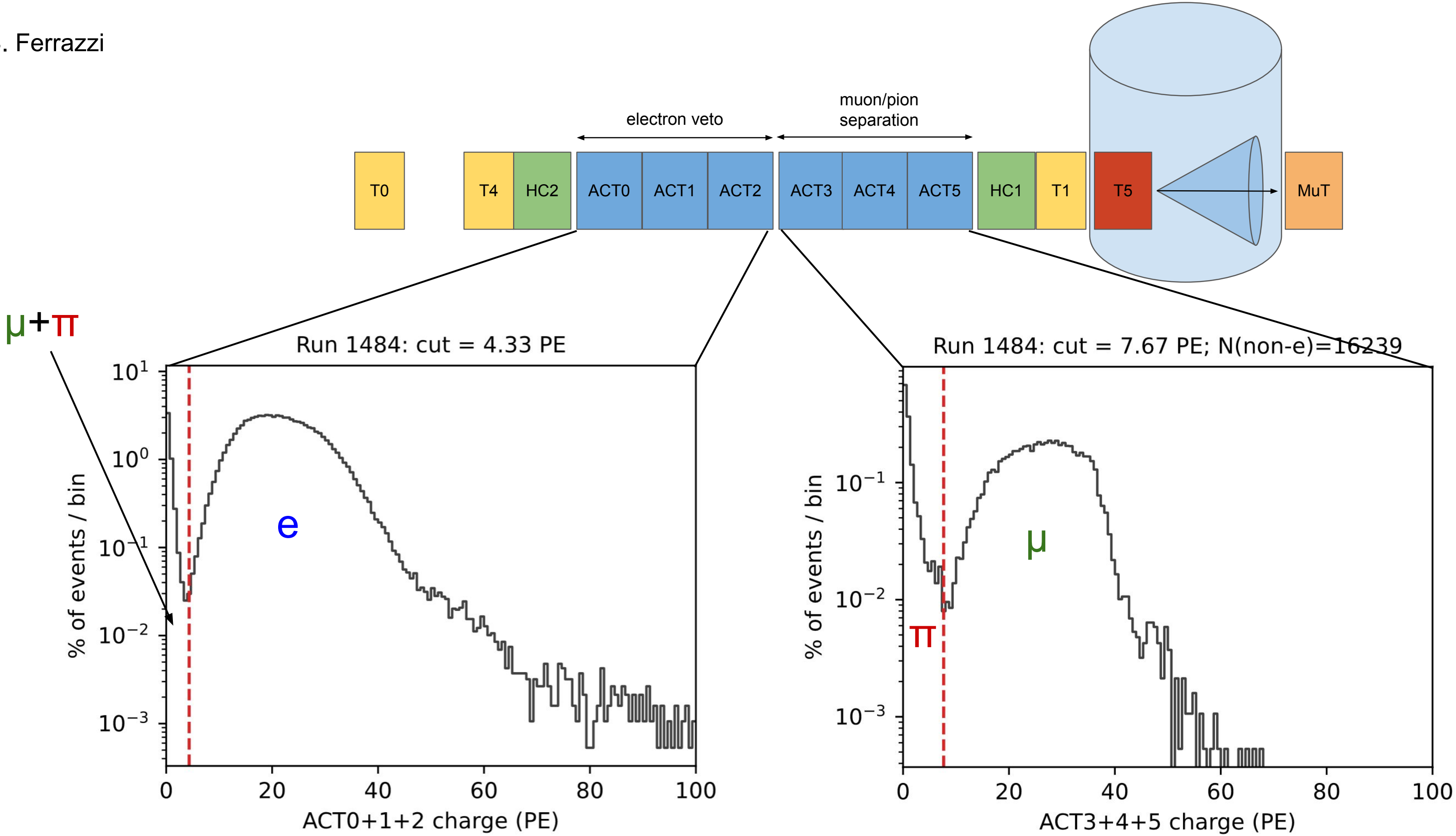
<u>Analysis Strategy</u>

- heavy particles (p, K, …) are cut off with a TOF cut
- use high charge in electron-veto ACTs to tag electrons
- once they are vetoed, use high charge in tagger ACTs to tag muons



Particle Identification

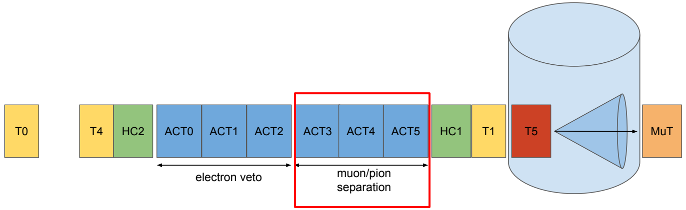
plots by B. Ferrazzi



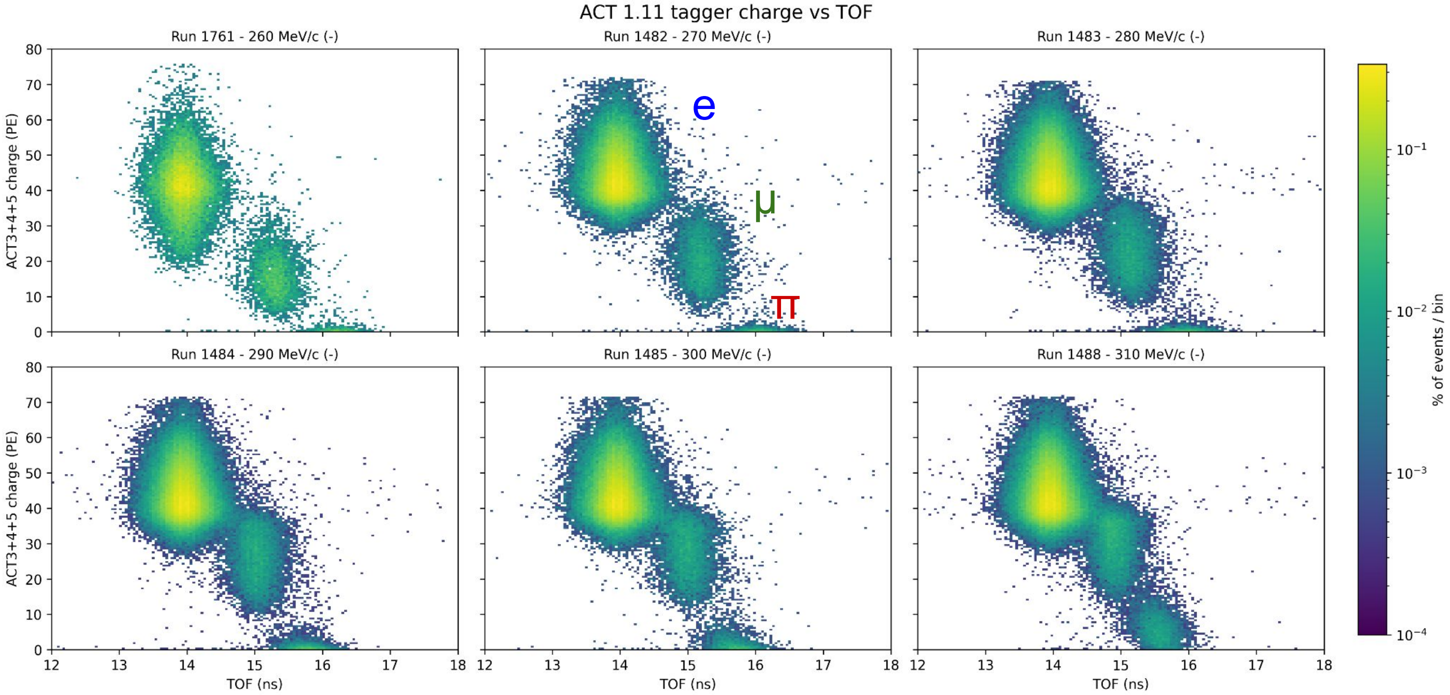
Charge vs Time-of-Flight

Time-of-Flight measured between T1 and T0

$$\mathrm{TOF} = \frac{L}{v} = \frac{L}{c}\sqrt{1 + \left(\frac{mc}{p}\right)^2}$$

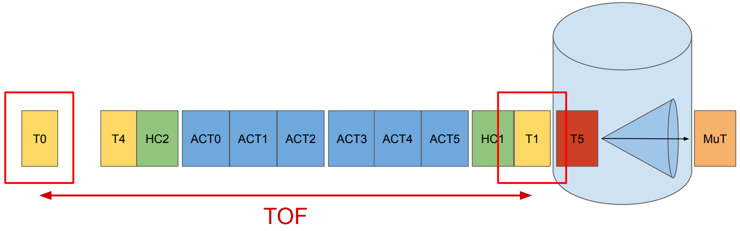


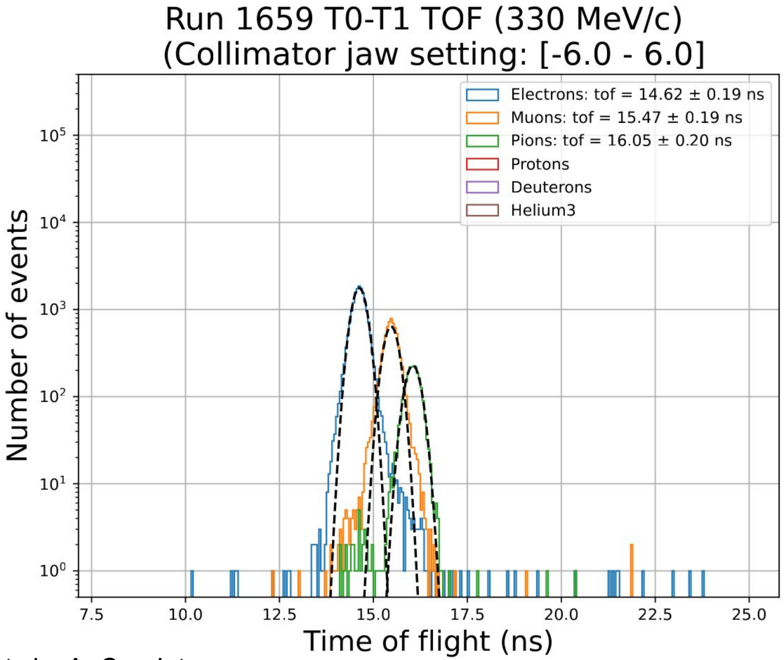
- ACT detectors allow for clear separation of e, μ and π
- estimation of sample purity is ongoing using complementary tank charge deposit



plots by B. Ferrazzi

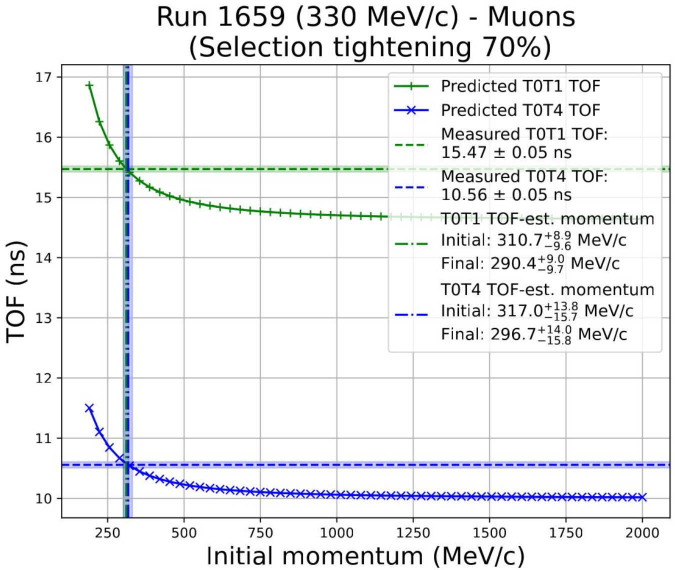
Momentum Estimation





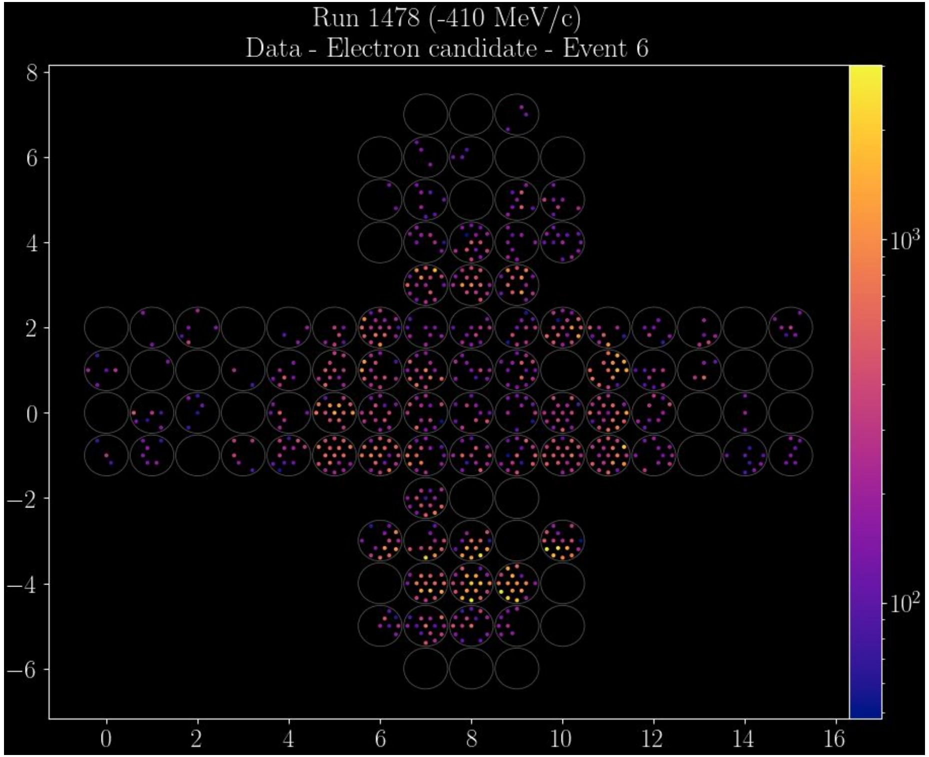
$$\mathrm{TOF} = \frac{L}{v} = \frac{L}{c}\sqrt{1 + \left(\frac{mc}{p}\right)^2}$$

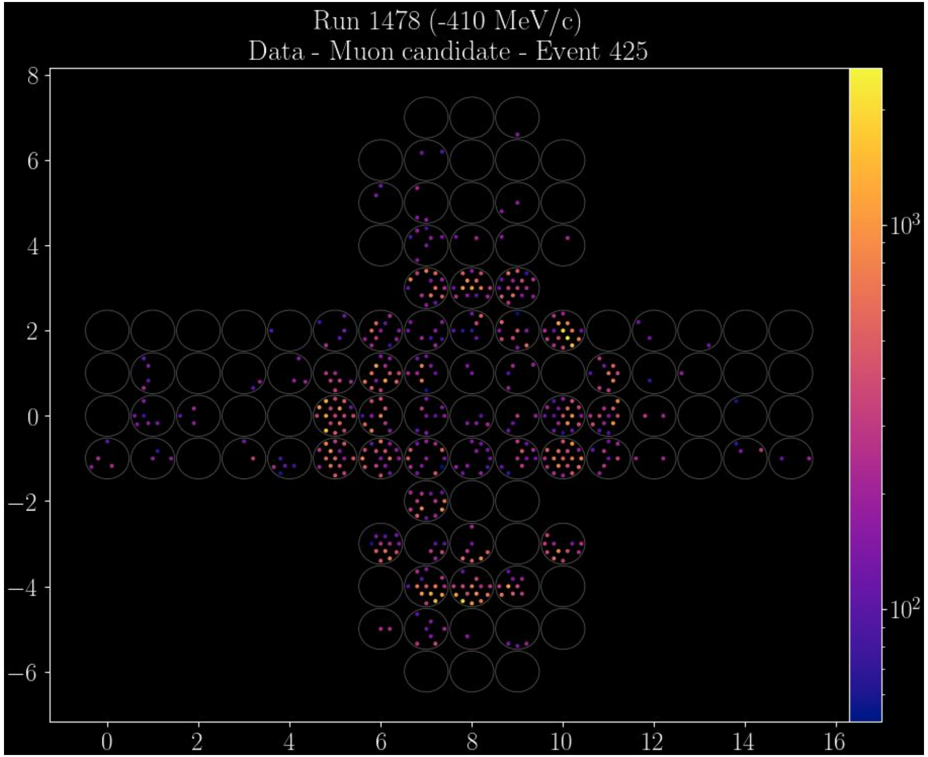
- model energy loss in beamline
- predict TOF as a function of momentum
- interpolate measured TOF to obtain momentum of muons and pions (electrons too light)



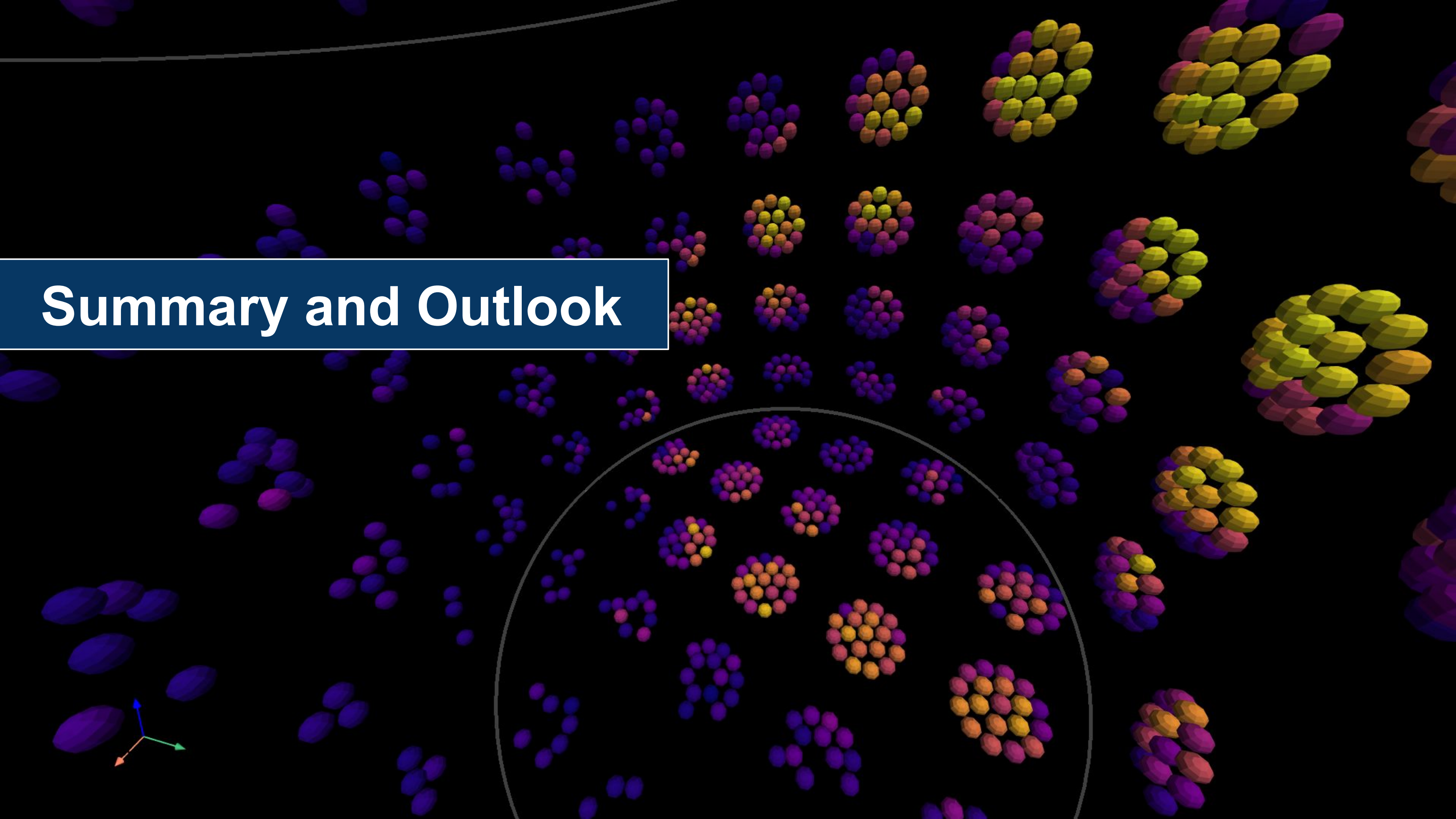
plots by A. Craplet

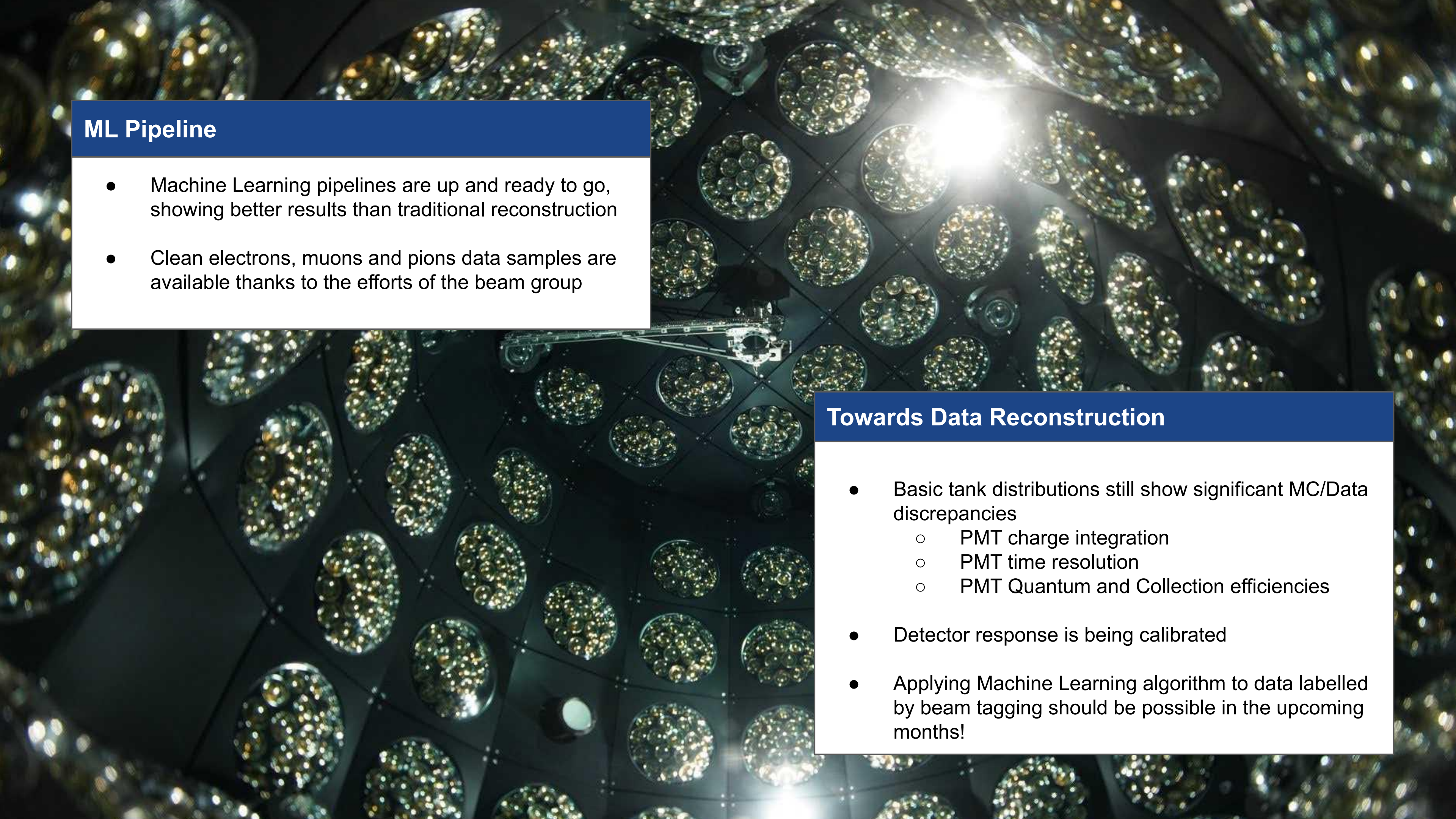
Events in the Tank!





Summary and Outlook





ML Pipeline

- Machine Learning pipelines are up and ready to go, showing better results than traditional reconstruction
- Clean electrons, muons and pions data samples are available thanks to the efforts of the beam group

Towards Data Reconstruction

- Basic tank distributions still show significant MC/Data discrepancies
  - PMT charge integration
  - PMT time resolution
  - PMT Quantum and Collection efficiencies
- Detector response is being calibrated
- Applying Machine Learning algorithm to data labelled by beam tagging should be possible in the upcoming months!