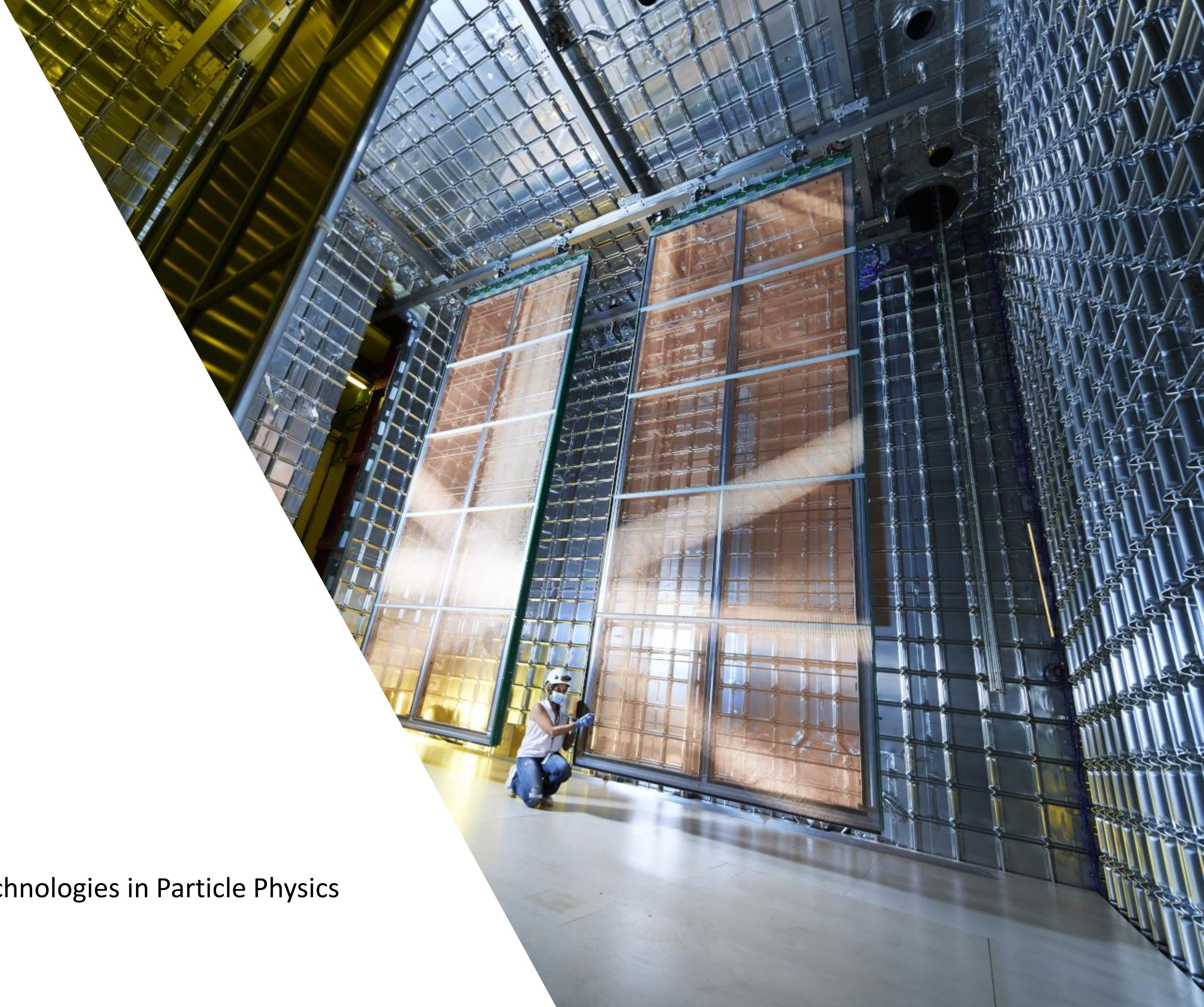
WARWICK
THE UNIVERSITY OF WARWICK

ML USAGE IN LARTPCS



17 September 2025

Artificial Intelligence (AI) and Emerging Technologies in Particle Physics

Outline

- Updates on ML usage in LArTPCs since the last meeting
  - Publication
  - Exploratory work with transformers
  - ML for pileup in neutrino beams
- Other ML activity in LArTPCs
- Resource availability

Pandora reconstruction in LArTPCs

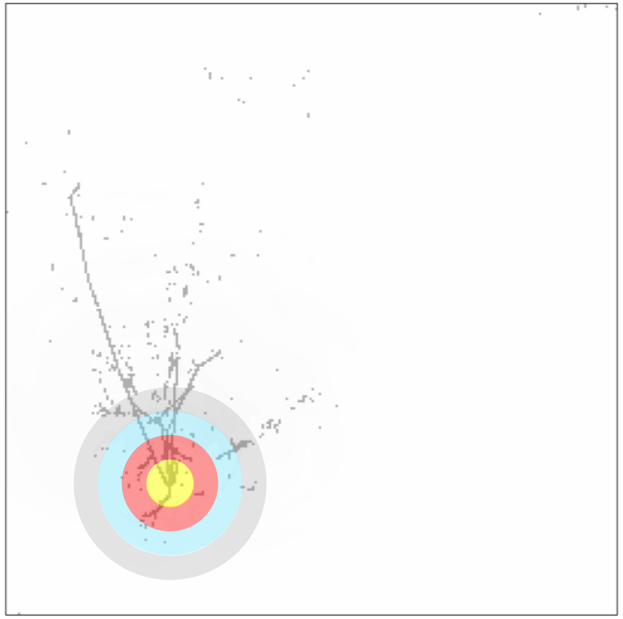
- This talk will give an overview of ML use in LArTPC neutrino detectors
  - I won’t be discussing work in water Cherenkov detectors
  - Key feature of **most** LArTPCs is that they produce 2D projections of neutrino interactions that are well suited for CNNs, and can also be represented as sparse hits suitable for GNNs, transformers, etc
- This talk will also focus on efforts related to ML use in the Pandora reconstruction package
  - I will briefly list some other ML projects in the LArTPC field

- Pandora is a multi-algorithm approach to reconstruction
  - Originally developed for colliders
  - Most effort in recent years focused on LArTPC development
  - Recent interest in further development for the FCC (FCC-ee Workshop on Particle Flow Reconstruction at CERN)
  - Interface to LibTorch allows algorithms using neural networks to run in our workflows
  - Anticipate a move to Inference-as-a-Service

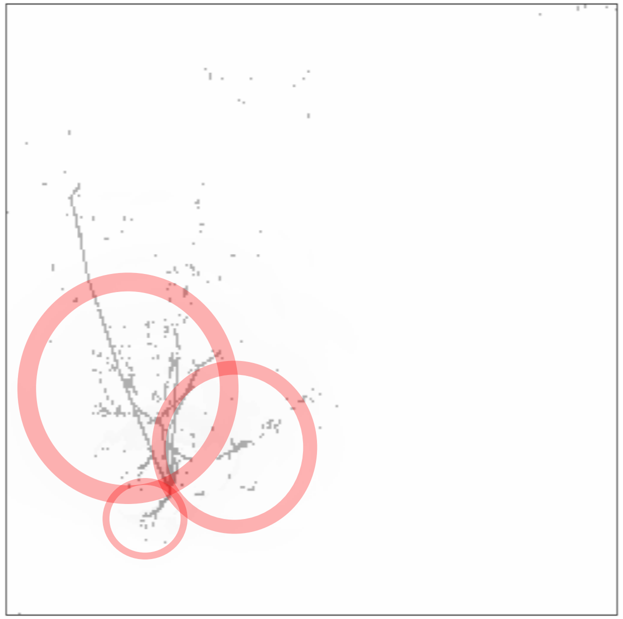
Publication

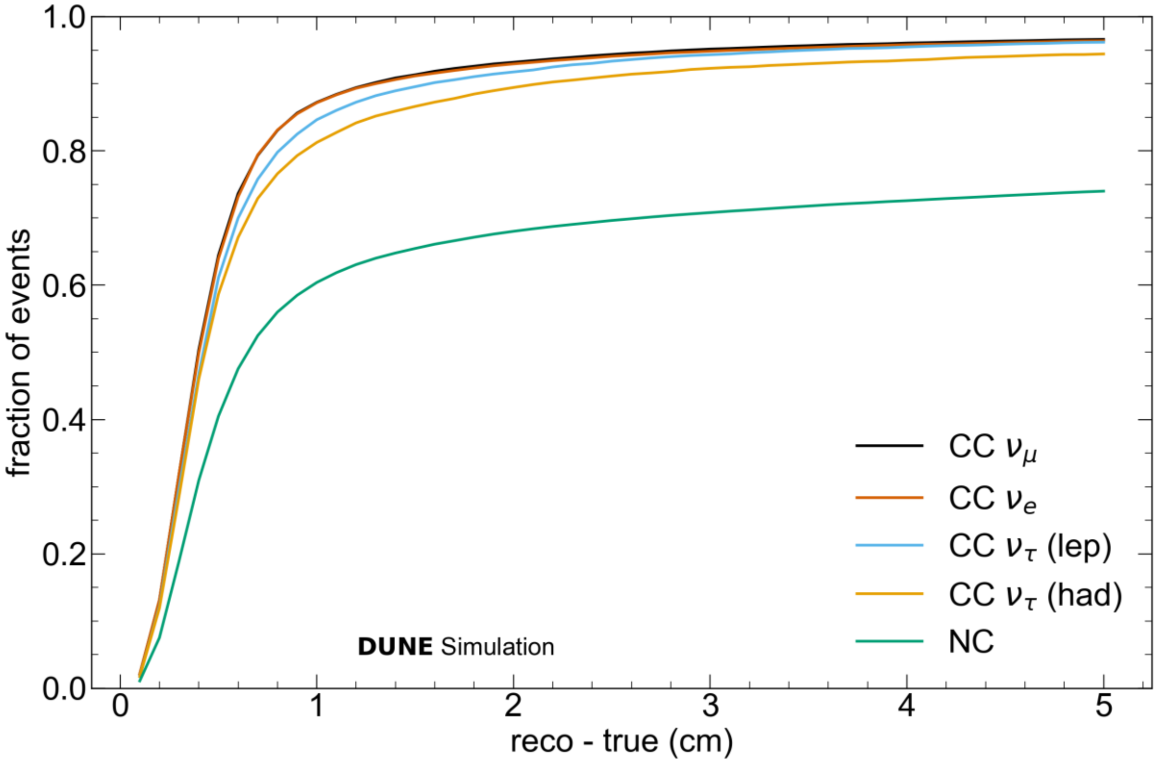
- At the last meeting I talked about vertex finding in LArTPCs via a U-ResNet
  - That work is now published:

In training hits are assigned a class according to distance from true vertex



Network infers hit distances and resultant heat map isolates candidate vertex





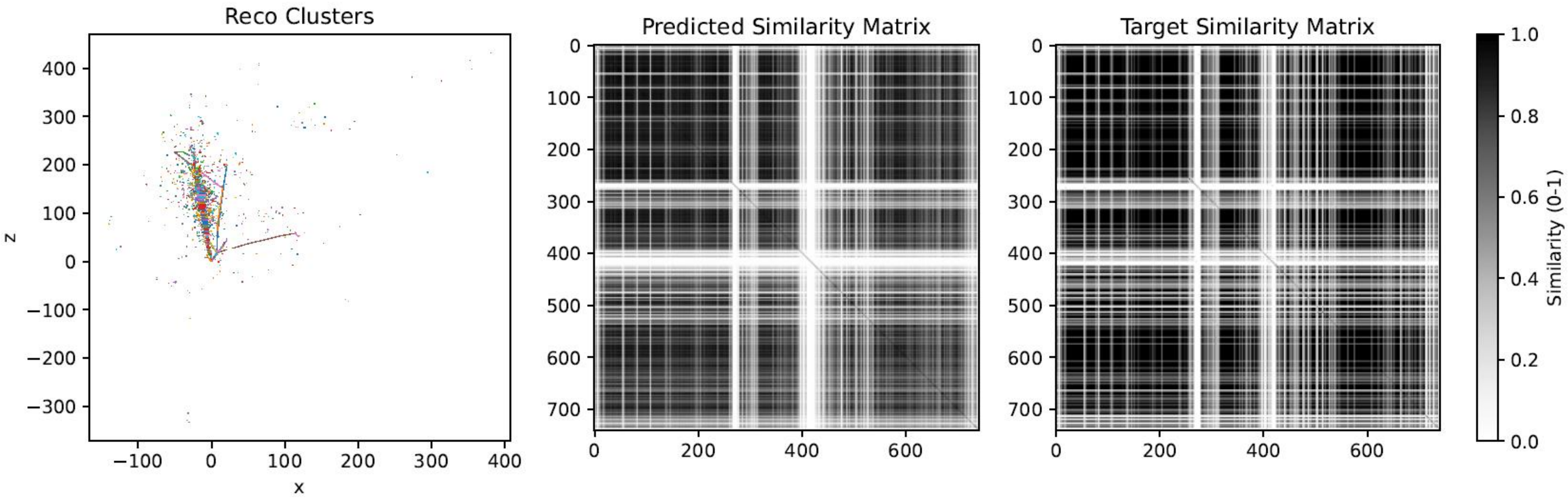
Eur. Phys. J. C **85**, 697 (2025)

Transformers for particle characterisation

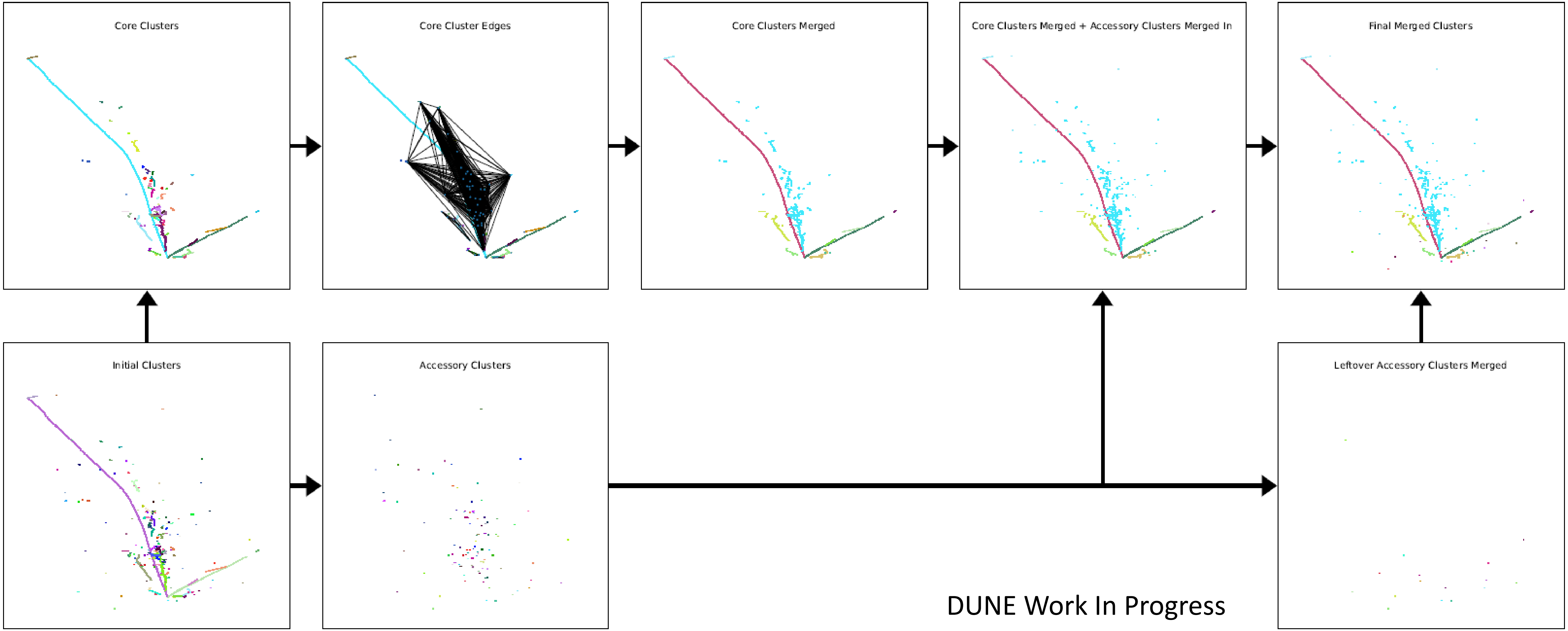
- Exploring the use of a transformer encoder to classify hits (MIP, HIP, Shower, …)
  - Each hit is represented by a token with a set of features
    - (plane, r, cosθ, sinθ, width, charge)
  - A hit's characterisation ought to be informed most by its local environment
  - Use an attention window with tokens order by distance to reconstructed vertex
  - Attention window then represents shells about the reconstructed neutrino vertex
- Compute salience to explore network behaviour
  - Classify target hit (red)
  - Network attends to shower-like hits nearby (blue)

Transformers for shower clustering

- Reconstructing showers in LArTPCs is particularly challenging
- Fragmented clusters and diffuse activity, often with many overlapping particles projected onto a 2D plane
- Hand-engineering algorithms for such topologies is very difficult
- Use the attention mechanism of transformers to learn and reconstruct similarity matrices for these clusters



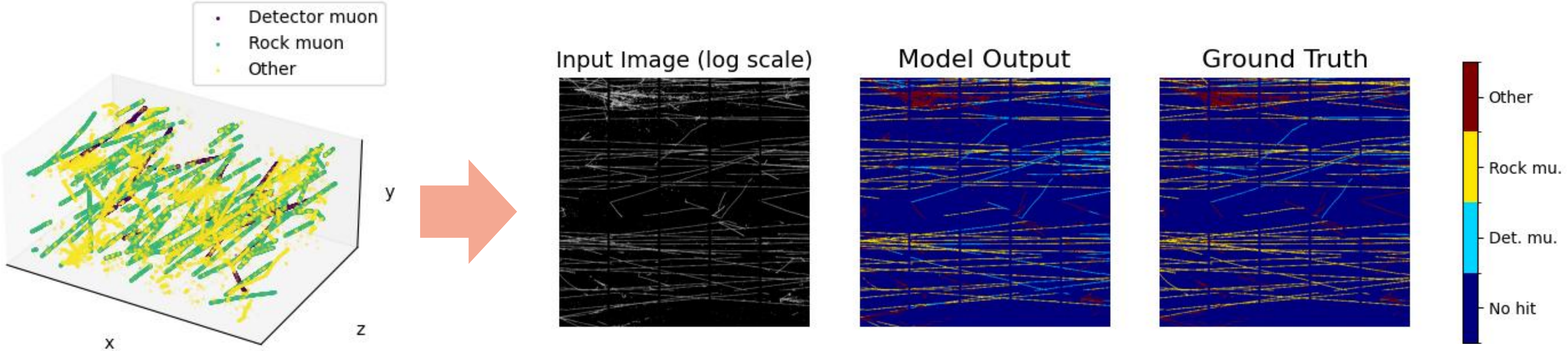
Transformers for shower clustering (2)



DUNE Work In Progress

Tagging rock muons

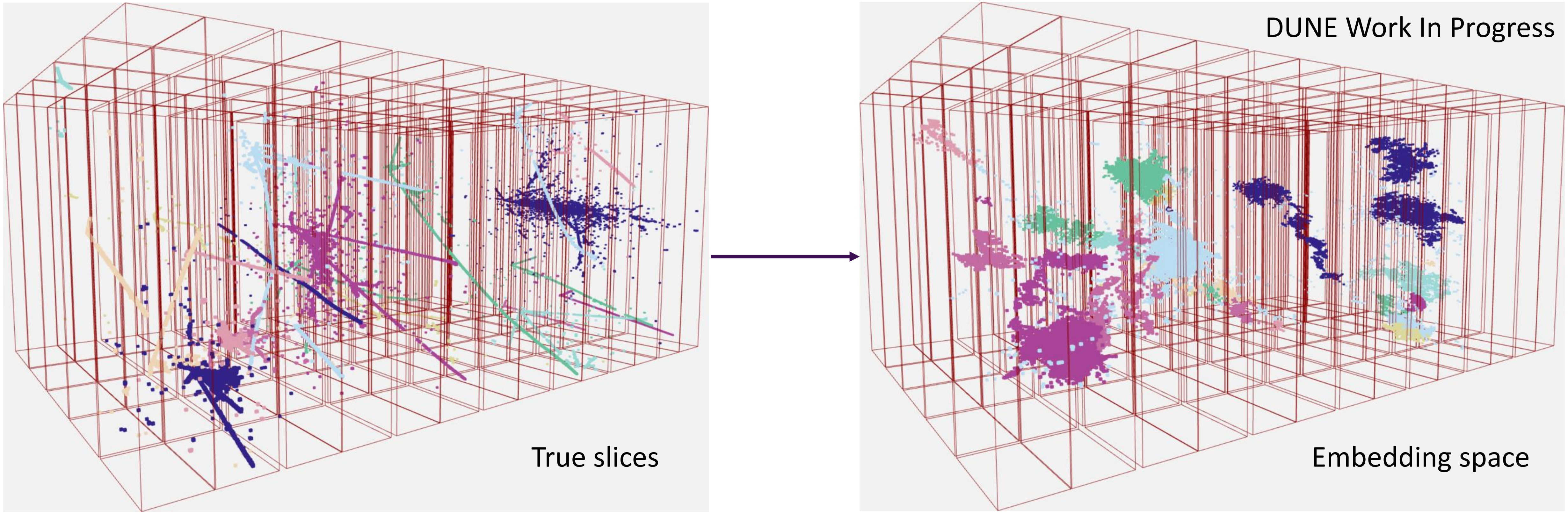
- DUNE’s LAr-based near detector will be exposed to a significant background of muons from interactions in the rock surrounding the detector (bottom left)
- Aim to tag and remove this background to retain a pure sample of neutrino interactions in the near detector LAr
- One avenue of work re-uses previously developed U-ResNets to tag these ‘rock muons’



DUNE Work In Progress

GNN to handle neutrino pile up

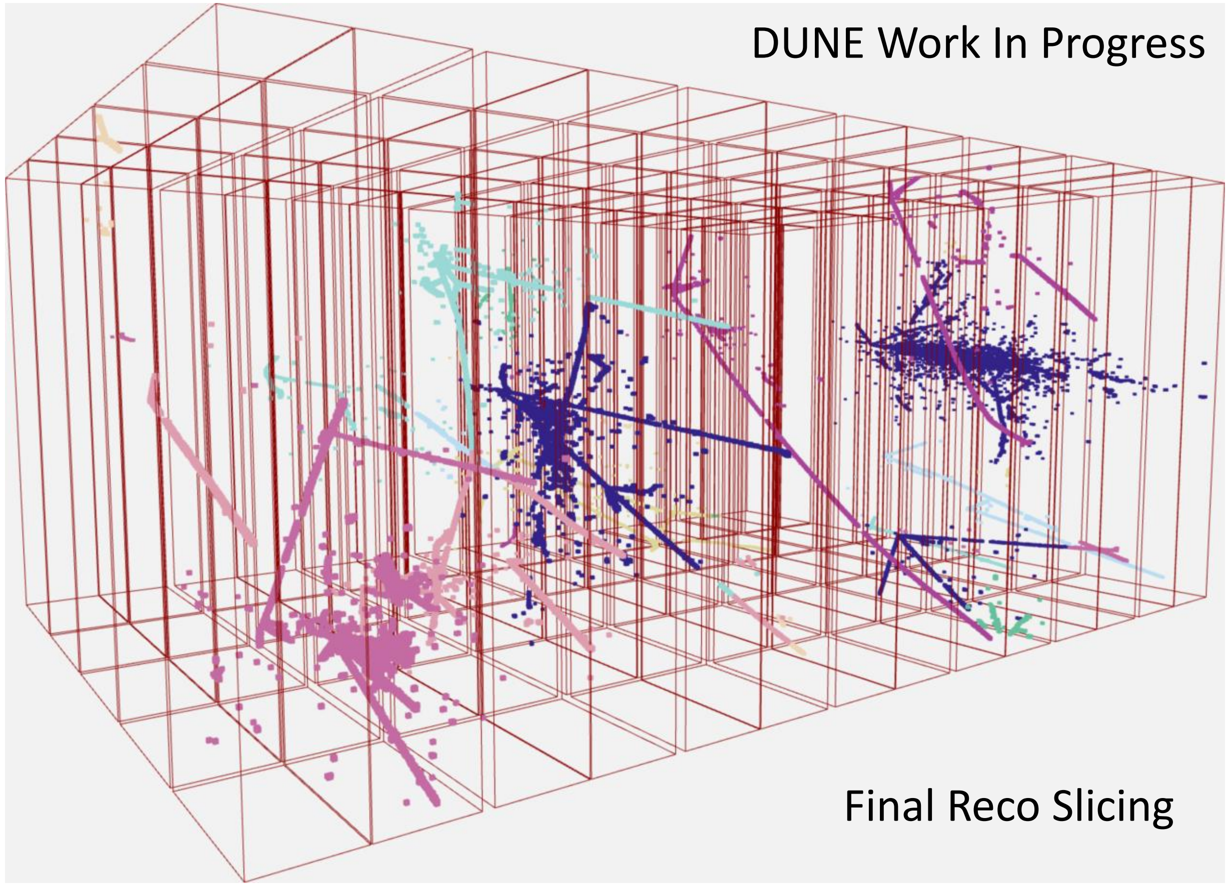
- Near detector output is natively 3D, so makes sense to exploit the additional information by using a GNN
- Typically expect to see O(30) neutrino interactions in the active volume per beam spill in DUNE's LAr-based near detector
- Each interaction can represent a rather complex hierarchy of final state, secondary and tertiary (etc) particles
- Aim to separate each individual neutrino-induced interaction and separately reconstruct it
- Construct a GNN using object condensation
    - Predict an interaction center for every hit, cluster the resulting blob-like hits



GNN to handle neutrino pile up (2)

- Unfold the clustering from the preceding steps to tag the reconstructed slices
- Promising results, but a little too much merging currently (currently final clustering undertaken via DBSCAN)
- Each of these slices would then be isolated and passed to downstream algorithms to reconstruct the individual particles in each slice

- Currently this assumes rock muons have been removed, but could potentially extend this method to separate the rock muons as well
- Such work can also be deployed to handle cosmic ray background in surface detectors such as SBND, MicroBooNE and the ProtoDUNEs



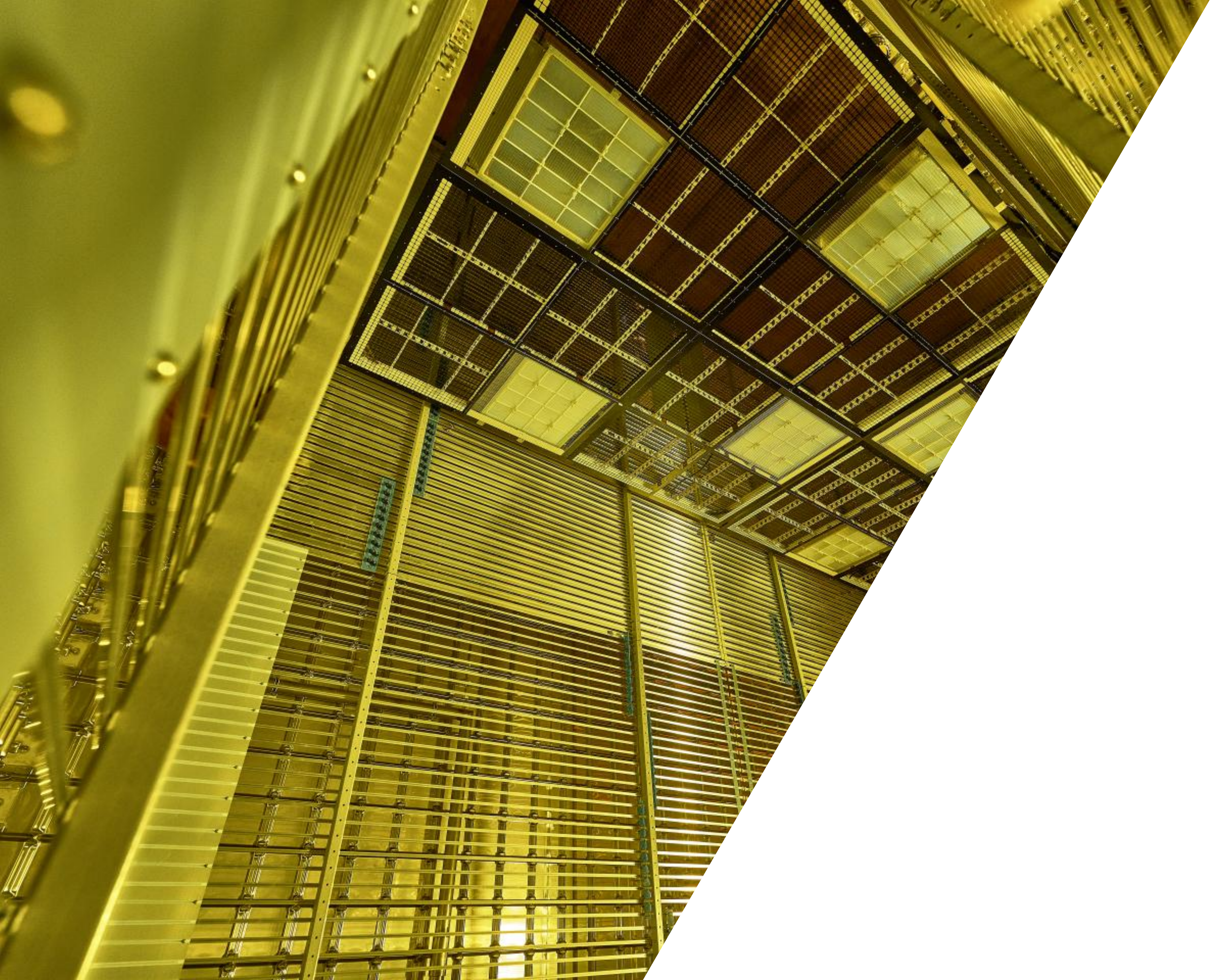
Other ML activity in LArTPCs

- A number of US-based groups are developing end-to-end ML reconstruction chains
  - SPINE multiple architecture approach (https://arxiv.org/abs/2102.01033)
  - NuGraph GNN approach (https://journals.aps.org/prd/abstract/10.1103/PhysRevD.110.032008)
- Flavour tagging at DUNE is provided by a CNN (https://journals.aps.org/prd/abstract/10.1103/PhysRevD.102.092003)
- Wire-Cell uses a deep neural network for region of interest finding in signal processing (https://iopscience.iop.org/article/10.1088/1748-0221/16/01/P01036)
- Numerous other efforts in beam design, simulation, DAQ and triggering

- DUNE LLM
  - Development of an AI agent unifying databases and documents
  - Provide referenced answers/source code for DUNE collaborators

Resource availability

- Warwick Blythe and Avon HPC clusters
  - 48 Nvidia Quadro RTX 6000 GPUs with 24 GB RAM
  - 48 Nvidia L40 GPUs with 48 GB RAM
  - 4 dedicated interactive nodes with Nvidia A40 GPUs with 48 GB RAM
  - (in practice, at any given moment, 10 GPU jobs are runnable by a user)
- Lancaster
  - HEC Cluster: 24 NVidia V100 GPUs, 8 NVidia L40 GPUs
  - 2 dedicated Nvidia A6000 ADAs for DUNE UK RS&DC project
- RAL PPD Cluster
  - 7 Nvidia A100s, 6 Nvidia Titan RTXs, 2 AMDGPU MI210s
  - + some older ones
- Non-UK resource
  - Cincinnati: 8 Nvidia V100 GPUs
  - Fermilab Elastic Analysis Facility: 8 Nvidia A100 GPUs with 80 GB RAM
- Broadly, we have good resource availability, but can always expand to more for things like hyperparameter exploration, etc



Backup