Machine Learning in Pandora

Andy Chappell

12/04/2024

The Second Wire-Cell Reconstruction Summit

WARWICK

ML in Pandora – Present and future

- Libtorch interface added to Pandora in October 2020
  - Allows us to run deep networks as a standard part of our algorithm chain
  - Libtorch v1.4.0 – fairly limited network architecture support - no sparse networks
- First fully integrated network, running in LArSoft July 2022
  - U-ResNet for vertex finding in DUNE Horizontal Drift far detector in Long Baseline workflow
  - Since demonstrated in various workflows and detector configurations at DUNE and MicroBooNE
  - Same architecture being re-used for signal/background separation in supernova neutrino samples at DUNE and neutrino/cosmic separation at MicroBooNE
  - Extending concept to secondary vertex finding
- LArSoft now supports Libtorch v2+
  - Provides support for GNNs and Transformers
- Expanding ML use in Pandora (all work in progress)
  - Identifying reconstruction errors (Transformer encoder)
  - Re-clustering (GNN)
  - Vertex refinement (GNN)
  - …

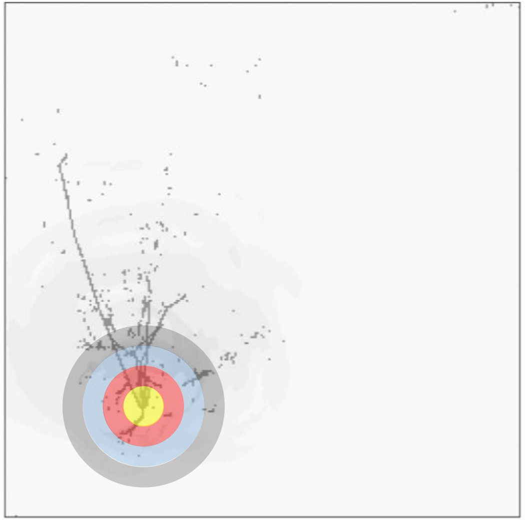
Vertex finding

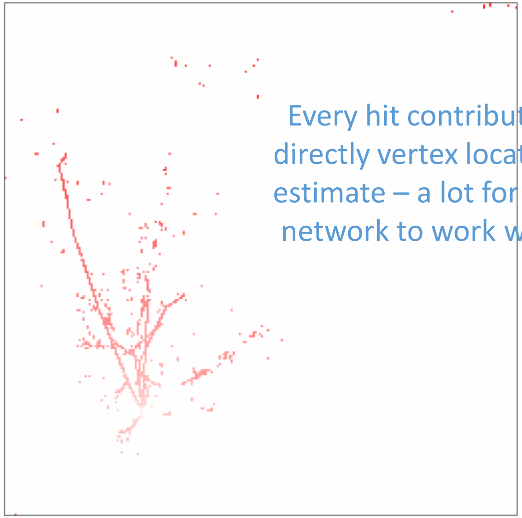
The concept

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

Network trained to learn those distances from input images

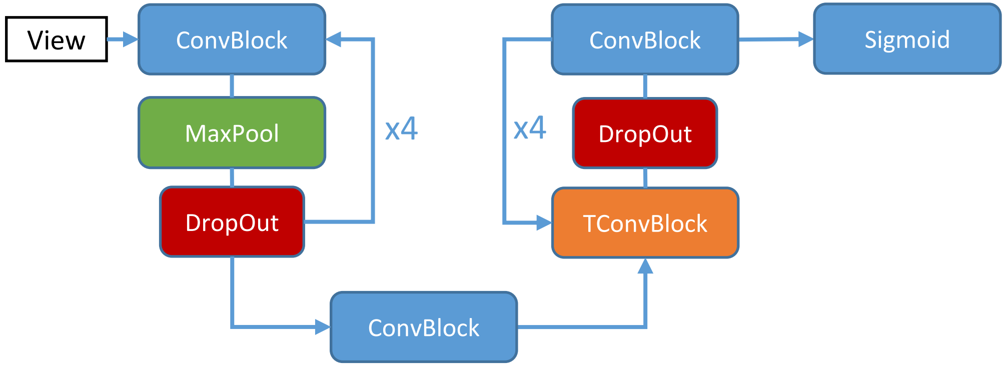
Network infers hit distances and resultant heat map isolates candidate vertex

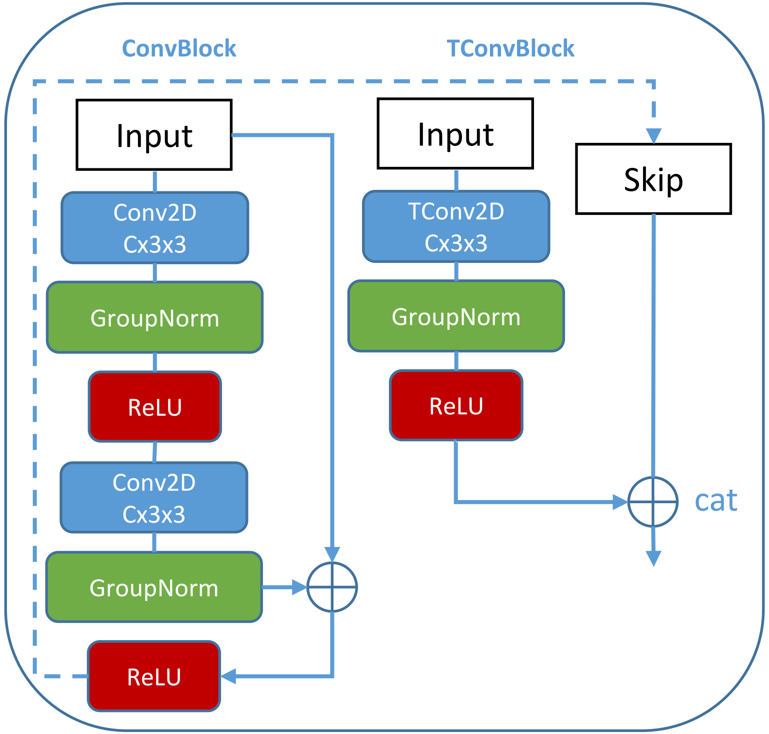




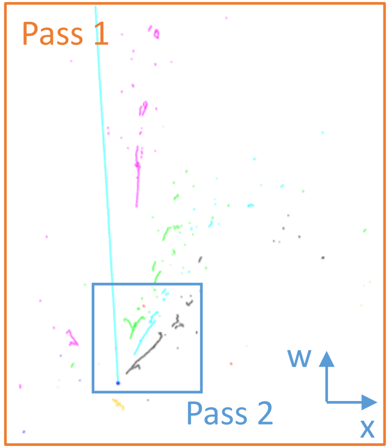
Network architecture

- U-ResNet structure for image segmentation
- Attempt to classify every pixel in an image



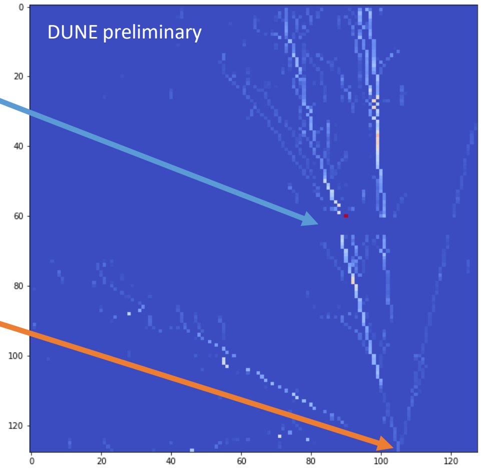


Two pass approach



- DUNE events can span a large physical region (many metres)
- 256x256 pixel pass 1 input to maintain computational tractability
- Pixels have low spatial resolution relative to DUNE's ~0.5 cm wire pitch
- Solution: Low resolution first pass, zoom in on RoI for second pass

Gap between anode plane assemblies



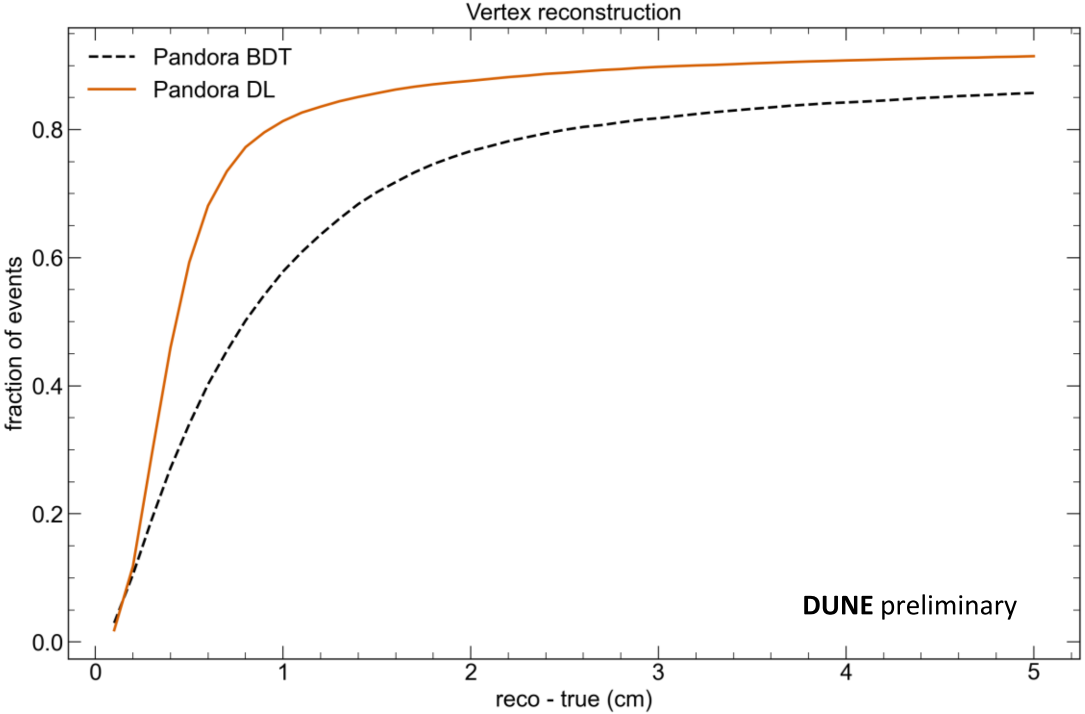
Pass 1 estimated vertex

- Use hit distribution around pass 1 estimated vertex to frame RoI to include as much context as possible
- 128x128 pixels for pass 2

Vertex reconstruction performance

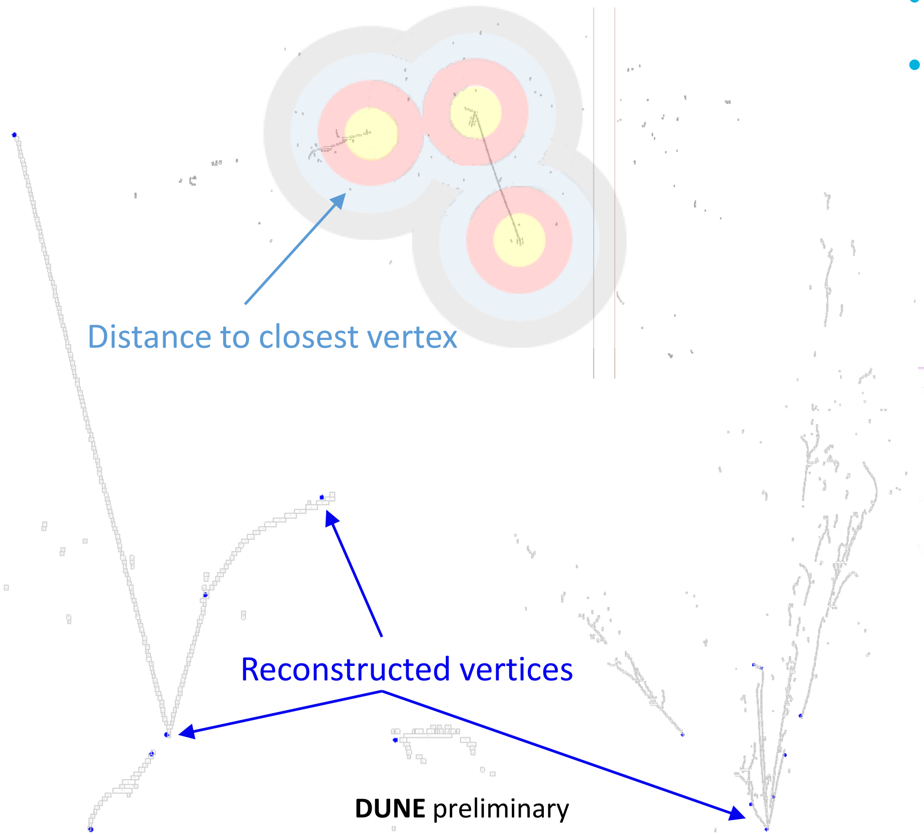
- Substantial improvement over previous BDT based vertex reconstruction
- Example here is from DUNE HD LBL workflow

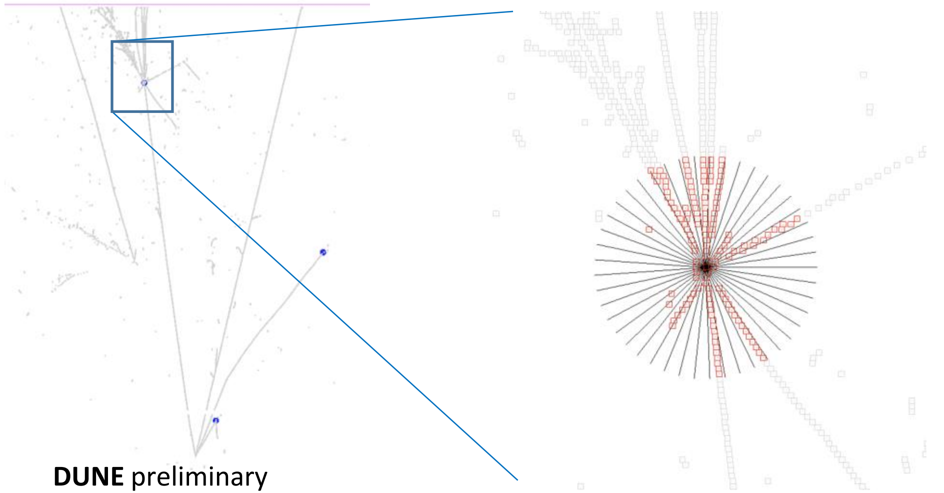
- About 10% of events (NC dominated) still have "catastrophic failures"
- Some evidence from MicroBooNE that GNNs can reduce such failures
- Potential to replace pass 1 with a GNN



Secondary vertex finding

- Isolate a region around the vertex
- Depending on vertex precision can either
  - Build stubs out from vertex
    - Grow from these seeds
  - Temporarily remove hits and reconstruct only well separated hits
    - Reconsider removed hits given context of surrounding reconstructed clusters

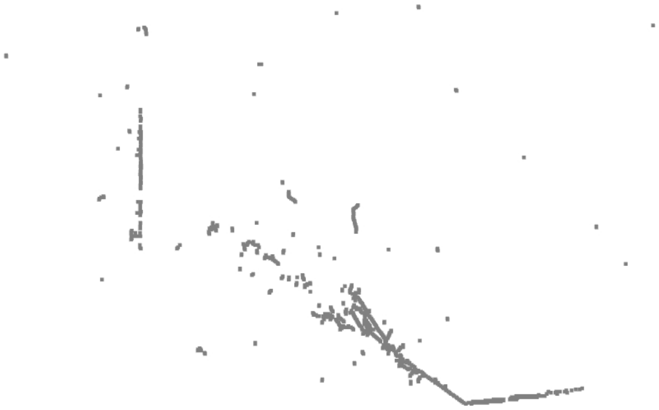




Vertexing in the presence of cosmics

The problem of cosmic remnants

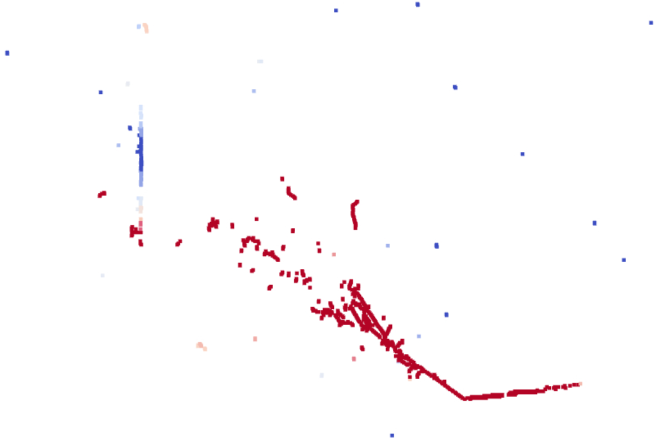
- In experiments like MicroBooNE neutrino interactions are crossed with many cosmics
- Pandora has algorithms to separate cosmics into slices separate from the neutrino, but sometimes contamination can remain
- This is a problem for vertexing because now there are hits with no meaningful relationship to the vertex
- Enough such hits can alter the estimated vertex position

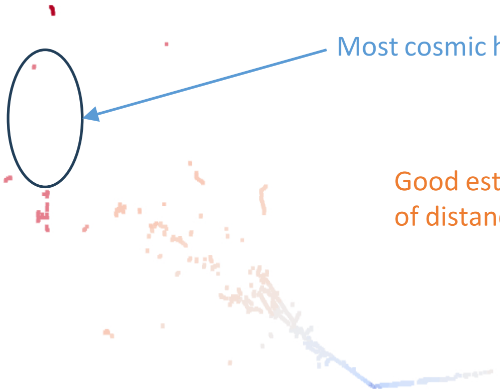


MicroBooNE preliminary

Leveraging the U-ResNet architecture

- A solution to this problem is to use the same U-ResNet architecture to tag cosmic and neutrino hits and remove the cosmics from consideration



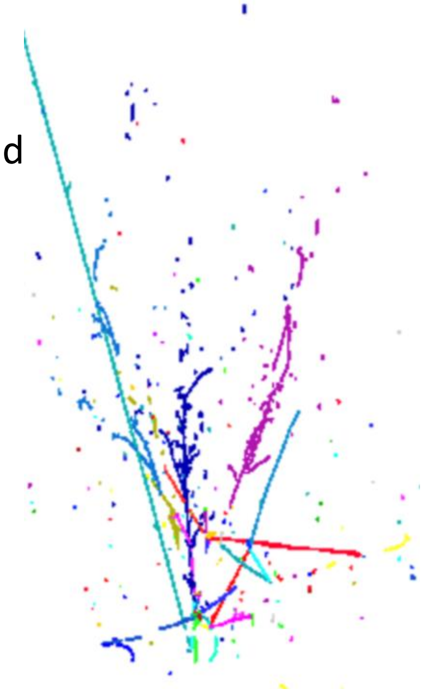


- We are looking at extending this procedure to improve the slicing process itself

Finding errors and re-clustering

Known problems

- Shower reconstruction is particularly challenging
- Fragmentation of shower clusters makes trade off between purity and completeness challenging
- Overlap, especially with 2D projections, adds to the challenge
- Pandora often over-merges showers as a result
- Two broad approaches
    1. Stop the problem happening in the first place
    2. Spot when the problem has occurred and fix it

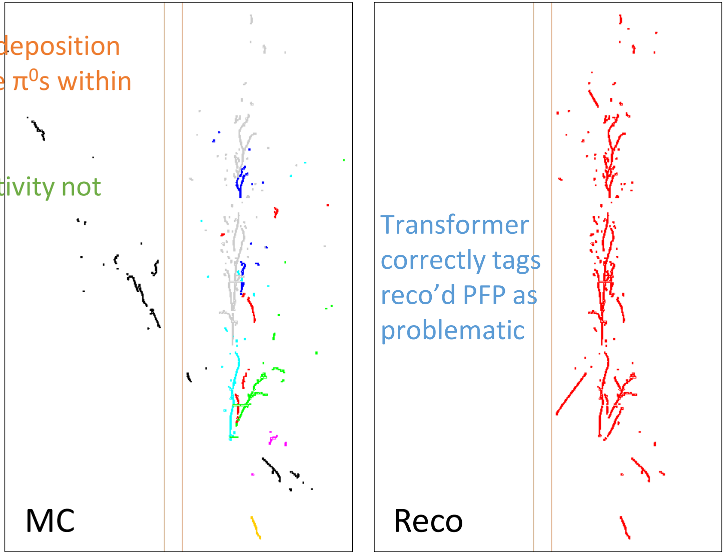


Transformer encoder

- Assess all PFPs reconstructed as shower-like by Pandora
- Train a transformer encoder to identify low purity showers
  - Any shower PFP where the main MC contribution is less than 80% of the total hit contribution
- Poorly reconstructed PFPs can be considered for re-clustering

Shower deposition from five π^0s within event

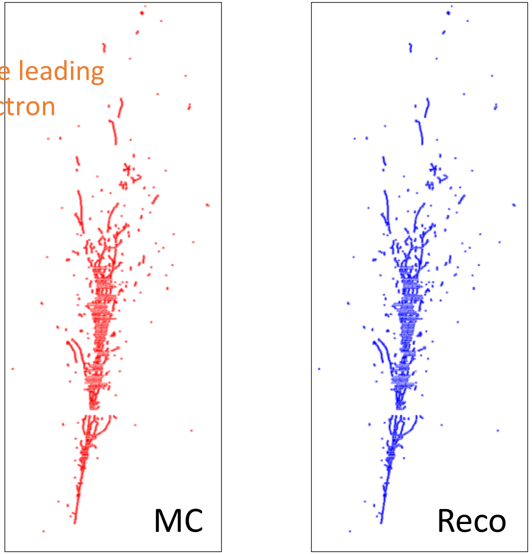
Track activity not shown



Transformer correctly tags reco'd PFP as problematic

DUNE preliminary

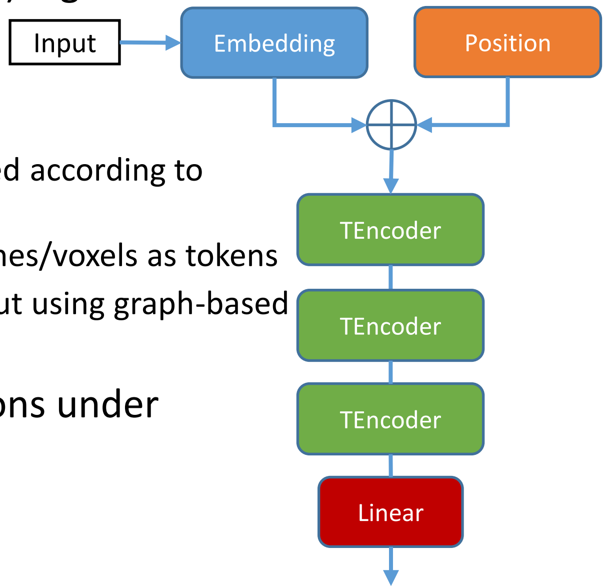
True leading electron



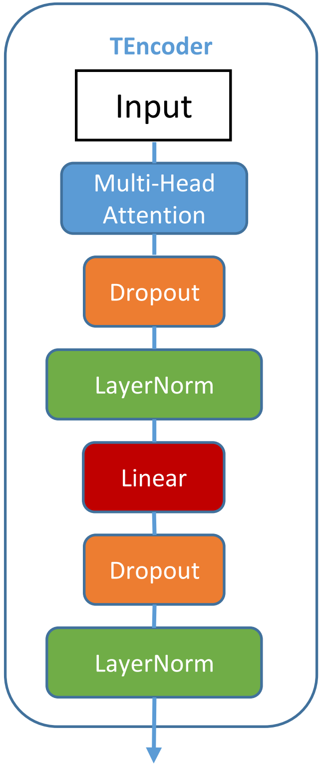
Transformer correctly tags reco'd PFP as correct

Transformer encoder network architecture

- Similar to encoding half of [Attention is All You Need] architecture
- Output is two class (good/bad) logits

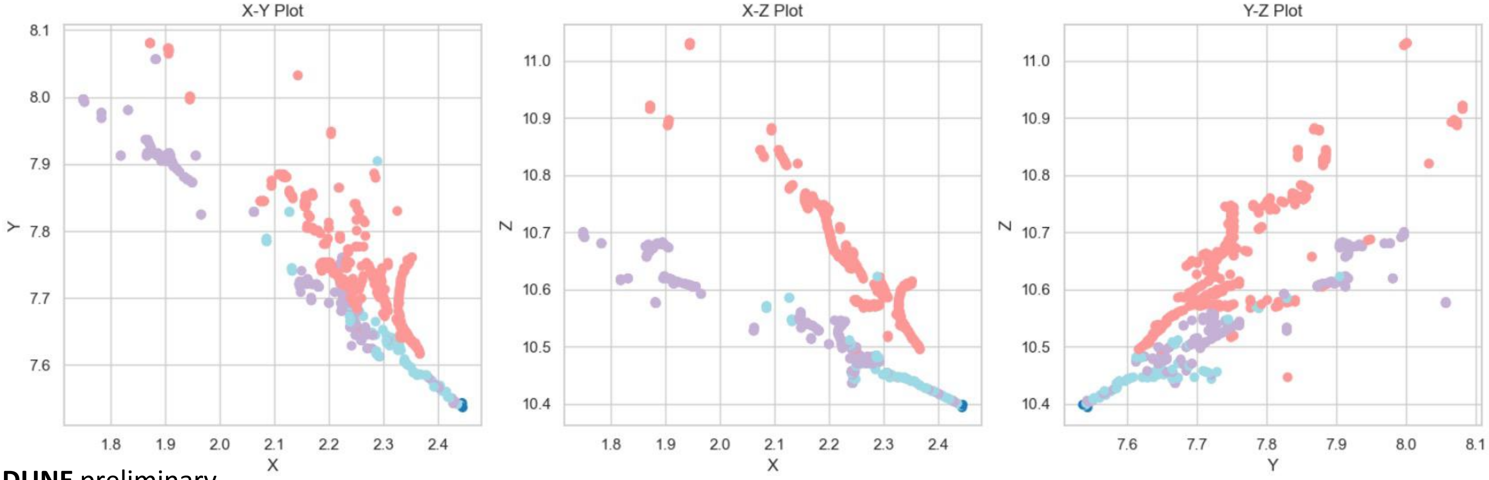


- Nature of input
  - PFP hits act as tokens sequenced according to distance from PFP reco vertex
  - Could consider PFP image patches/voxels as tokens
  - Could consider hits as tokens but using graph-based sequencing (later)
- Various hyperparameter options under consideration
  - Size of embedding
  - Number of encoders
  - Number of attention heads



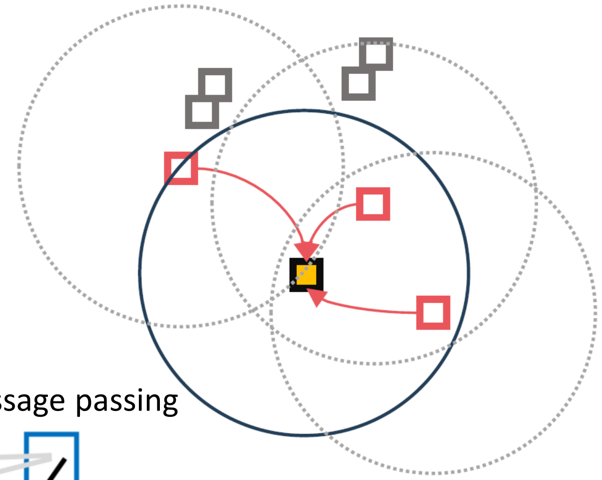
GraphSAGE for re-clustering

- Given a PFP tagged as being impure, consider re-clustering options
  - Intention is to allow support for multiple re-clustering options, but here we'll just look at a GNN
- Example training input where a single shower is actually composed of three separate, partially overlapping showers

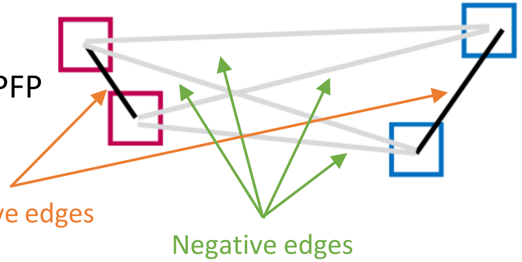


DUNE preliminary

GraphSAGE for re-clustering



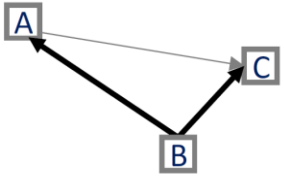
- Interpret 3D hits from PFP as a graph
    - 1 node per hit (Features: x, y, z, ADC)
    - 1 edge per node pair within 10 cm (Features: is from same particle?)
- GraphSAGE aggregates features among nearby nodes
    - Currently mean
    - Further aggregation is possibly for indirectly connected nodes via message passing
- Currently performs edge prediction
    - Decide if two hits belong to the same PFP
    - Node prediction also possible

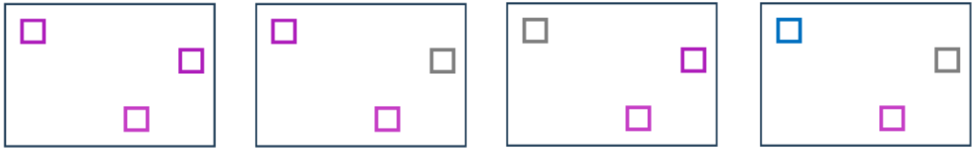


- Concatenate features of positively or negatively connected nodes and pass to fully connected layer to get out a final score

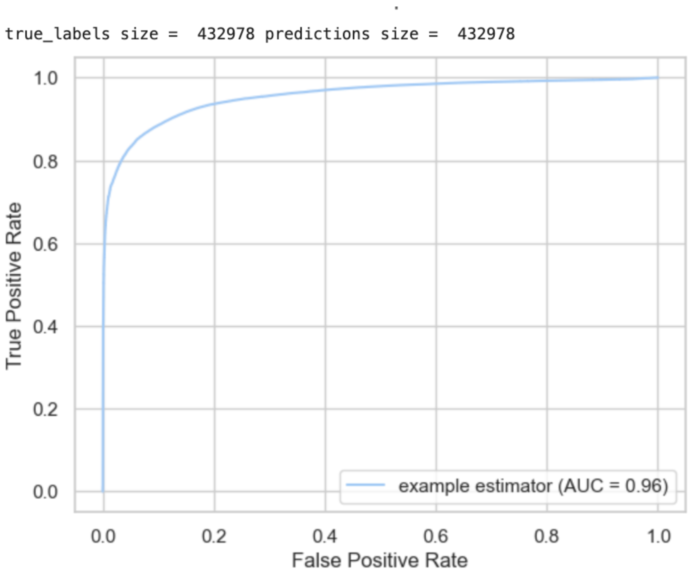
Performance and use

- Provisional, low stats training yields encouraging performance
- How to use it?
- Results can be ambiguous
  - Classify AB, BC connected
  - Classify AC not connected
  - Four choices:



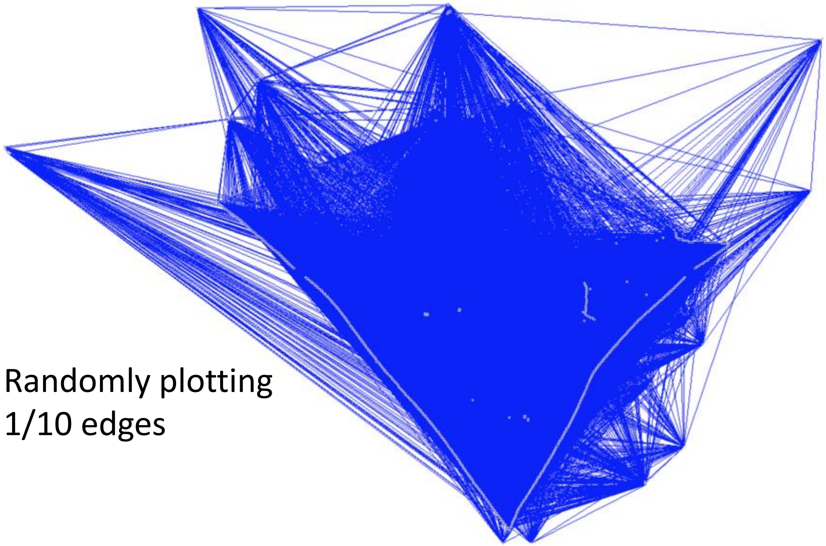


- Need a way to decide what the correct choice should be



An aside on graph and transformer inputs

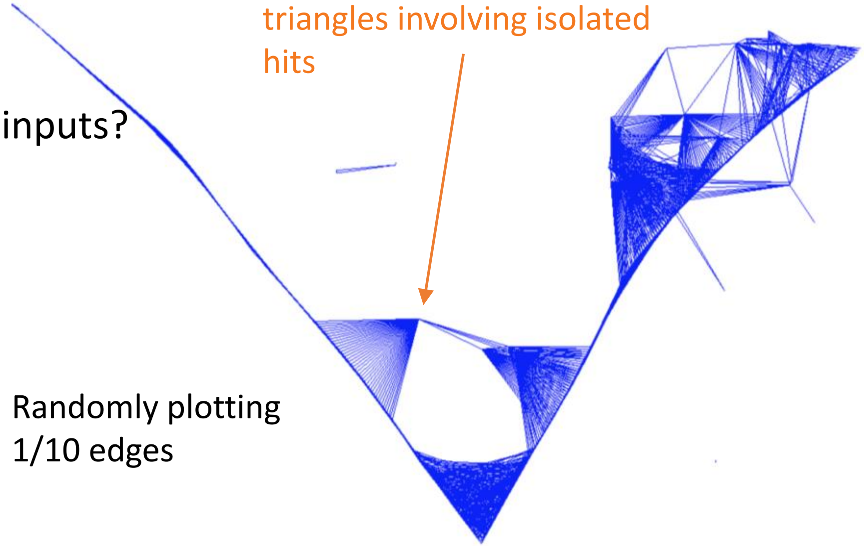
- One challenge of GNNs and transformers is the need to figure out input format
- Common approaches to graph inputs include distance-based connections and Delaunay triangulation
- Tessellation produced by Delaunay triangulation works well for nodes drawn from a spatially uniform distribution
- Our events don't look like that
- Hits are typically correlated, yields many often dubious edges
- There are, at least, no disconnected regions



Randomly plotting 1/10 edges

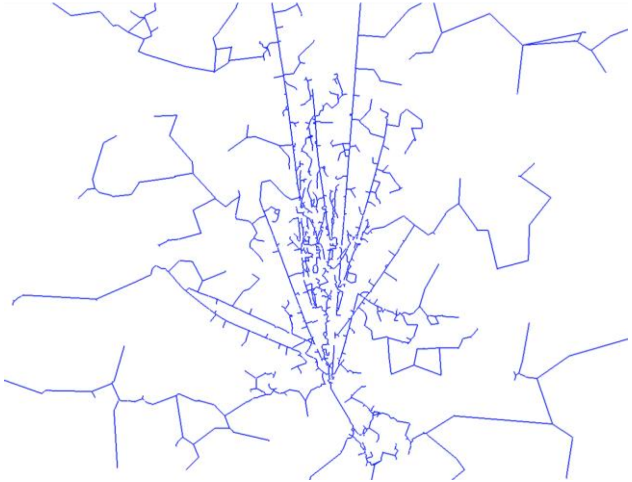
An aside on graph and transformer inputs

- We can prune the graph
- But still many dubious edges
- It's also a lot of wasted computation
- Can we do better and produce more useful inputs?

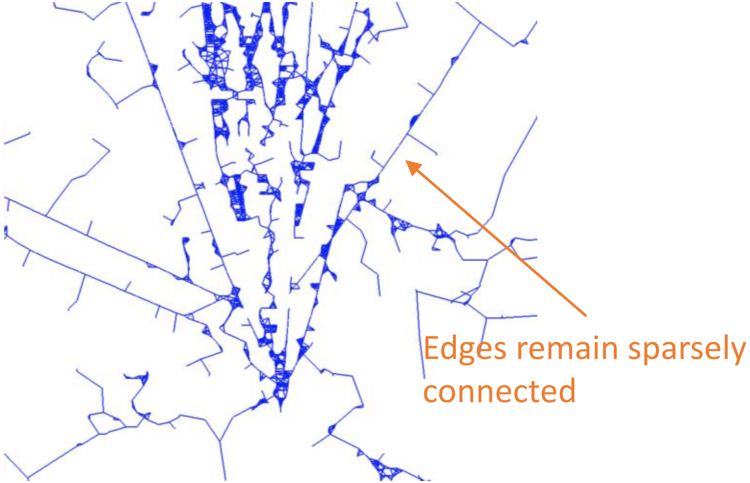


An aside on graph and transformer inputs

- With the help of vectorization we can build limited, locally connected graphs
- Extra step required to connect disconnected sub-graphs
- Edge multiplicity on left is probably too low for GNNs and could negatively impact message passing, but perhaps is useful to make sequencing choices for Transformers
- Allow for high local edge density but veto co-linearity



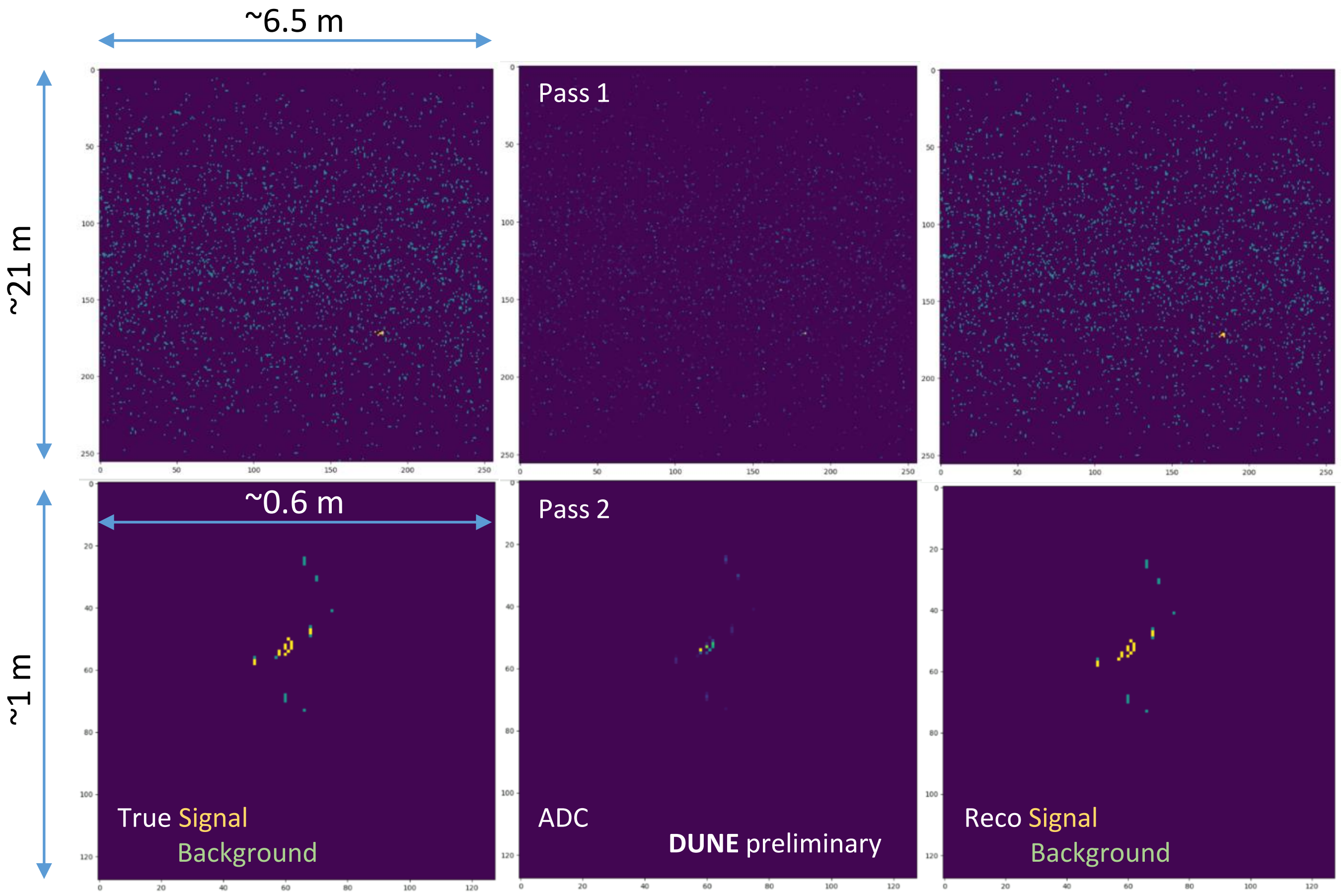
Zooming in on a dense shower



Supernova neutrino reconstruction

Example signal/background separation

- Need to extract supernova neutrino interaction from radiological background
- Expect U-ResNet to be well suited to this task
- Physical extent of input images very large
  - Need low-res region finding in pass 1
  - High-res classification in pass 2
- A lot of activity in pass 1 images, much less so in pass 2
- Provisional tests look promising



Conclusions

- Pandora has been capable of running neural networks for a few years
- ML-based vertex finding now well established in a number of workflows
  - Ongoing work to further refine performance
- Range of reconstruction tasks for which ML is being investigated now expanding
- ML solutions can fit neatly into Pandora's existing multi-algorithm paradigm