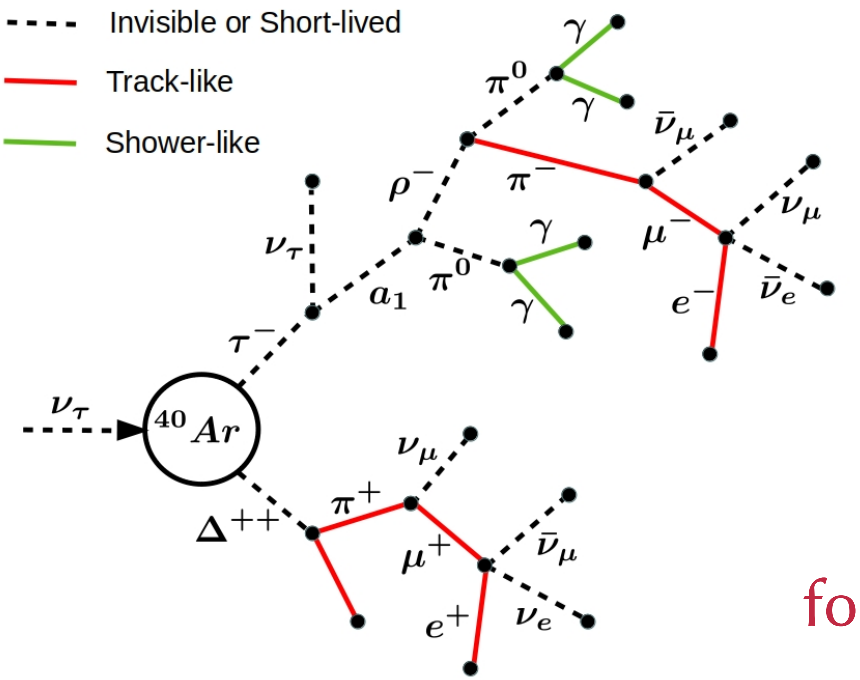
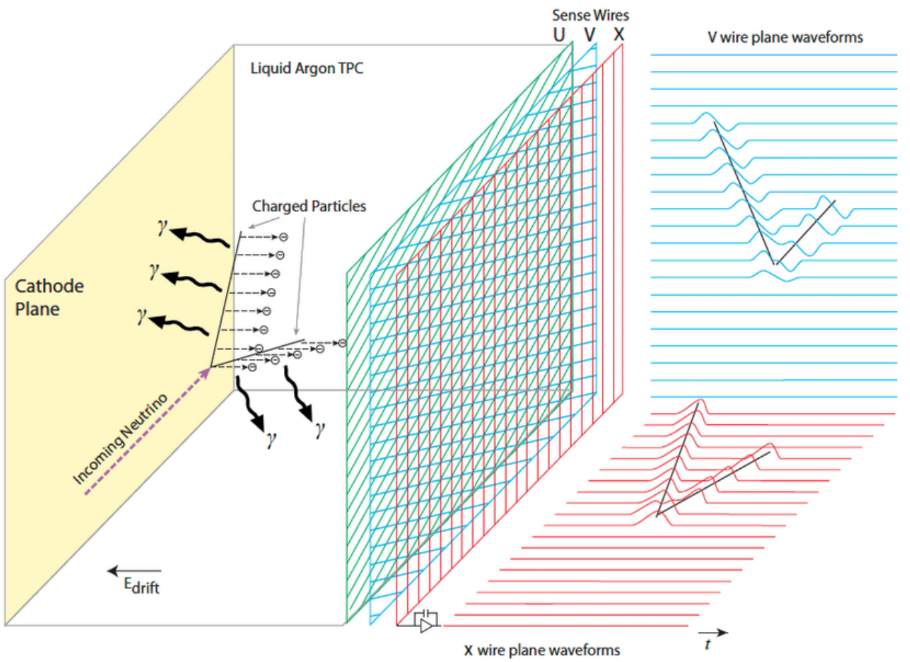
NuGraph3: Toward Full LArTPC Reconstruction using GNNs



Adam Aurisano
University of Cincinnati
for the Exa.TrkX/NuGraph Collaboration

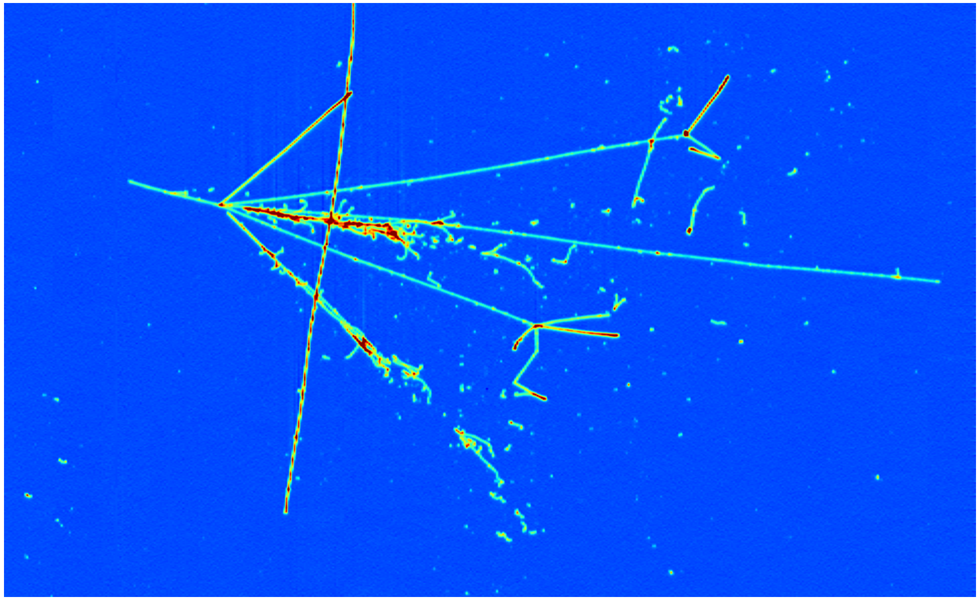
Neutrino Physics and Machine Learning 2024
Zurich, Switzerland
27 June 2024

Liquid Argon Time Projection Chambers



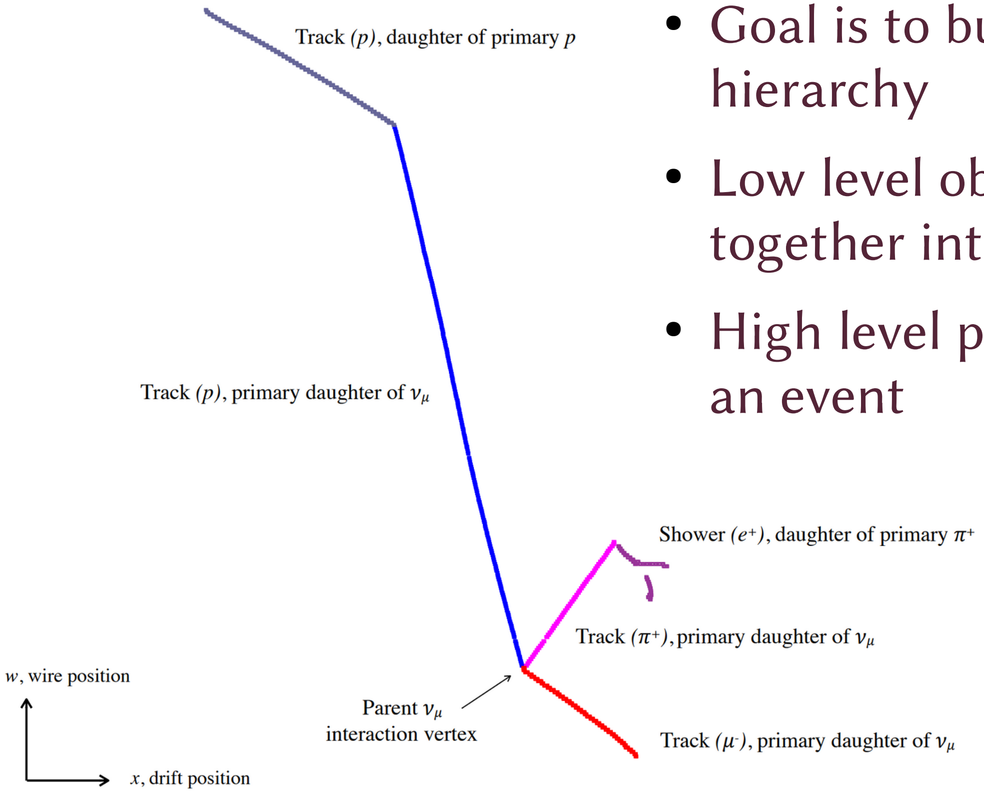
- LArTPCs are currently heavily used in neutrino physics
  - Now: MicroBooNE, Icarus, SBND
  - Future: DUNE (70 kT far detector deep underground)
- Charged particles ionize liquid argon as the travel
- Ionization electrons drift due potential between cathode and anode planes
- Closely spaced wires (~3 mm) at anode provide high-resolution image of neutrino interaction
- Multiple wire planes provide 3D information

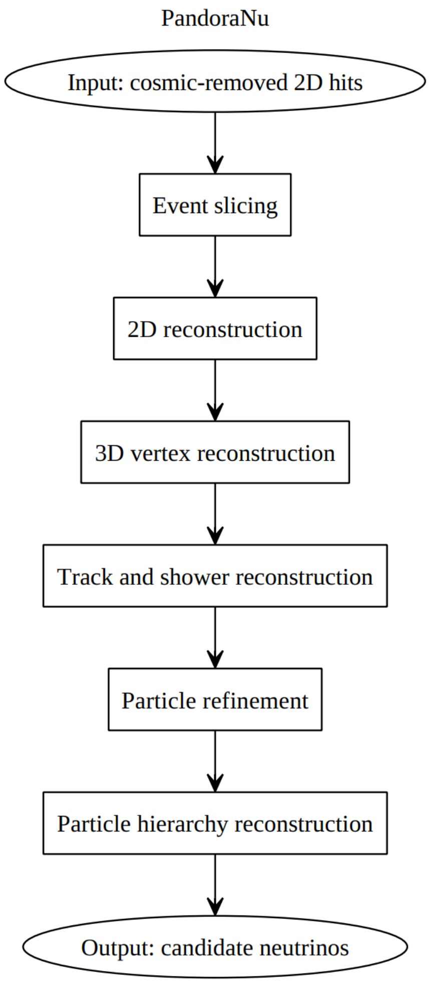
- High resolution images are blessing and curse
- Would like to
  - Cluster hits into objects
  - Classify objects according to the particle that created it
  - Assemble the objects into an event
  - Determine type and kinematic properties of the event



Pandora – Particle Flow

- Standard reconstruction package for LArTPC experiments is Pandora
- Large number of pattern recognition algorithms applied adaptively in a series of stages
- Goal is to build a full particle hierarchy
- Low level objects are stitched together into high level particles
- High level particles are stitched into an event

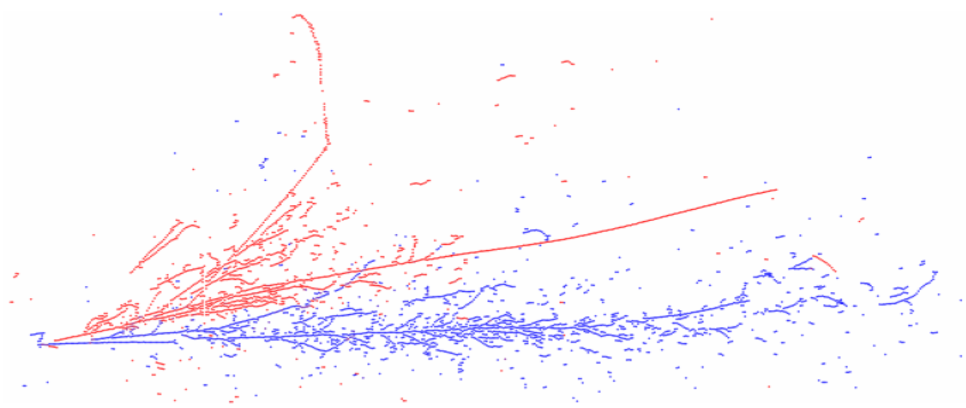




R. Acciarri et al, arXiv:1708.03135

Pathologies in Traditional Reconstruction

- Particle flow is a powerful technique, but it is subject to some pathologies
- Starts with 2D reconstruction
  - Some ambiguities cannot be resolved in 2D
- Serial reconstruction steps can lead to compounding errors
  - Some errors cannot be recognized until later in the reconstruction chain
- Pandora attempts to recover from these pathologies by iteratively rerunning algorithms



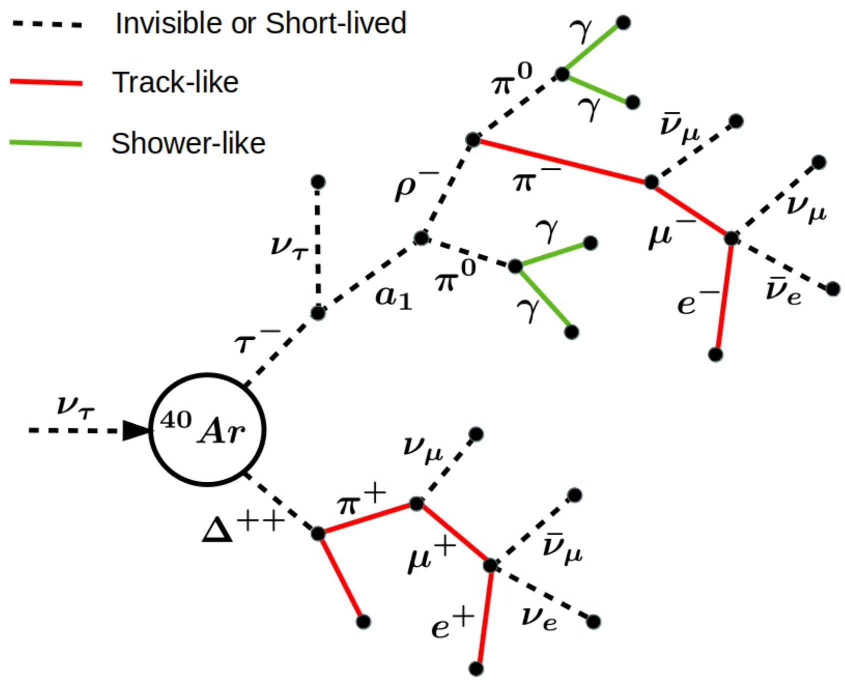
- Tau neutrino events are an important analysis target in the DUNE era
  - Frequently high multiplicity
  - Separating from other interactions requires excellent reconstruction of internal kinematics
    - Particle content is not sufficient
- Success depends on minimizing reconstruction pathologies

Graphs

- A graph is a mathematical structure that represents objects and binary relationships between them
  - Nodes: represent objects
    - Can hold associated information like spatial or temporal coordinates, or other features
  - Edges: connections between nodes
    - Relationship can be directed or undirected
    - Can have associated features
- Ideal structure for understanding physics data
  - Naturally sparse
  - Hits have a causal structure that can easily be modeled by edges
  - Accommodates relationships beyond nearest neighbor

$$G = (V, E)$$

Nodes → V, Edges → E



- Particle tree of a tau-neutrino interaction on argon
- Can be represented as a graph in several different ways
  - Particle tree
    - Nodes = particles
    - Edges = parentage relationship
  - Tracking
    - Nodes = hits
    - Edges = adjacent hits caused by same particle

Graph Neural Networks

- GNNs are an extension of the idea of CNNs
  - Instead of extracting features from patches in a regular grid, extract features from neighbors of node
- Iteratively learn a smart embedding of graph structure
- Encode geometric information by passing and aggregating messages from neighbors
- Learned edge weights can dynamically scale the importance of messages
- Used to great success by Exa.TrkX project for fast tracking at the LHC

Initial embeddings = node features

$$\mathbf{h}_v^0 = \mathbf{x}_v$$

Average of neighbors' previous embeddings

Previous embedding of v

$$\mathbf{h}_v^k = \sigma\left(\mathbf{W}_k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{|N(v)|} + \mathbf{B}_k \mathbf{h}_v^{k-1}\right), \forall k > 0$$

New embedding

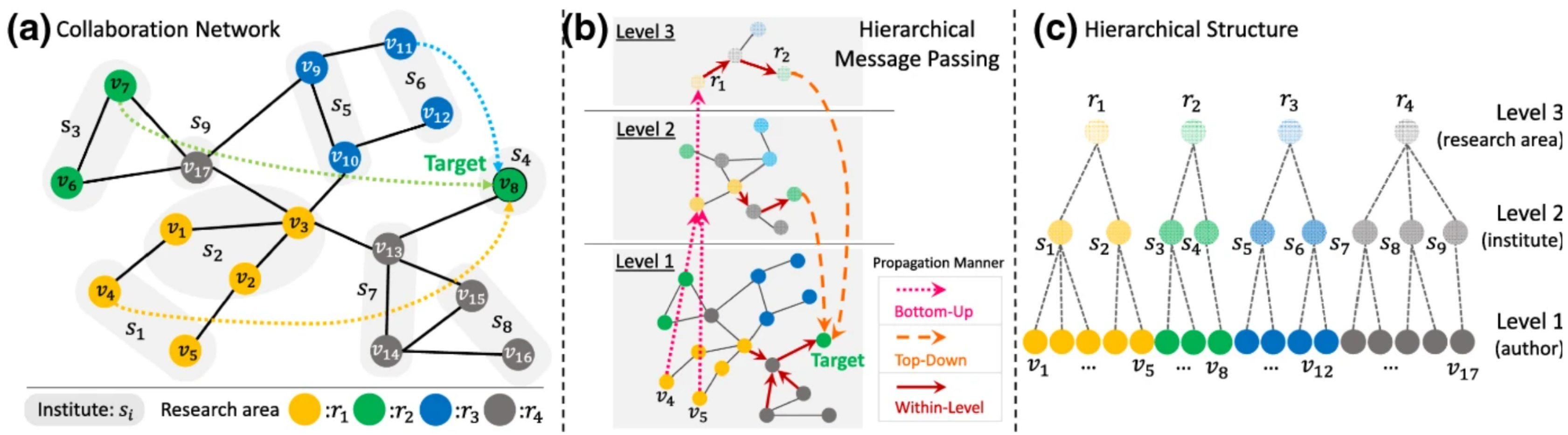
Shallow neural networks

Weaknesses of Flat GNNs

- Flat message-passing GNNs are powerful but have some weaknesses
  - Each message-passing iteration expands distance between connected nodes
  - Too many iterations degrades messages
    - Oversquashing
- Weaknesses were seen with early versions of NuGraph2
  - Attempted to find trajectories by iteratively improving edge weights
  - Initial graphs were kNN or ε-ball
  - Flip-flopping behavior in identifying track types
    - Tracks would be broken into segments alternatively classified as MIP or HIP
- NuGraph2 solves through Delaunay triangulation, but this makes the edges not physically interpretable

Hierarchical GNNs

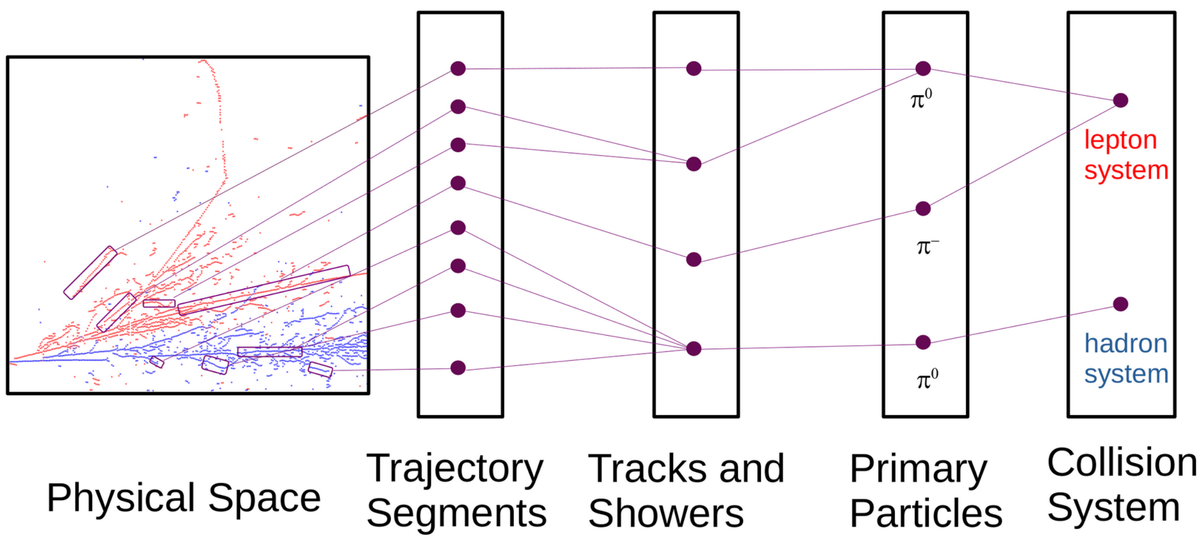
- Hierarchical GNNs solve "oversquashing" problem by allowing long-distance information flow through different hierarchical layers
- Layers capture rich, multi-scale information in a natural way
  - Can be used to better reflect inductive bias of the problem
- Message passing can occur both between and within levels



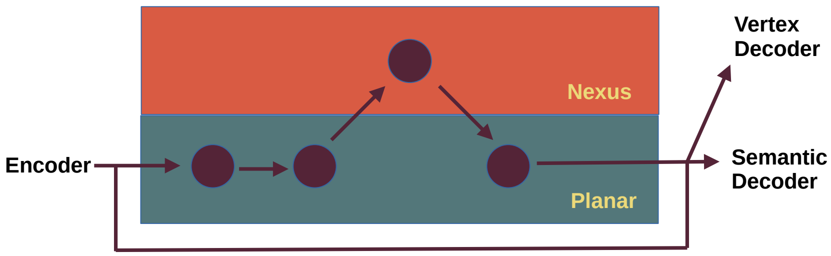
Z. Zhong, C. Li, J. Pang, arXiv:2009.03717

NuGraph3 Concept

- GNN-based particle flow reconstruction using NuGraph2 as starting point
- Similar to Pandora, consider series of reconstruction stages
- Each stage connects elements from stage before to produce higher level objects
  - Reconstruction chain expressible as a hierarchical graph with each level representing a reconstruction stage
- Avoid lossy serial steps by keeping many plausible reconstruction hypotheses and resolving them simultaneously
  - Expressible through fuzzy membership
    - Nodes on level L-1 can be connected to more than one node on level L
- Hierarchical message passing iteratively improves the particle tree reconstruction by choosing a reconstruction hypotheses using information from all stages simultaneously

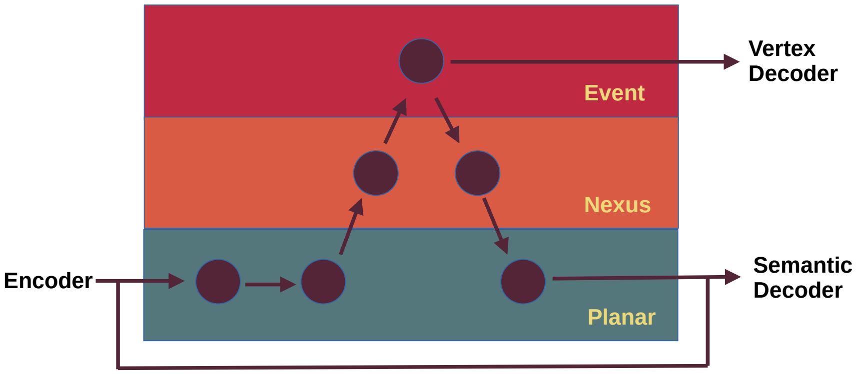


Hierarchical Message Passing

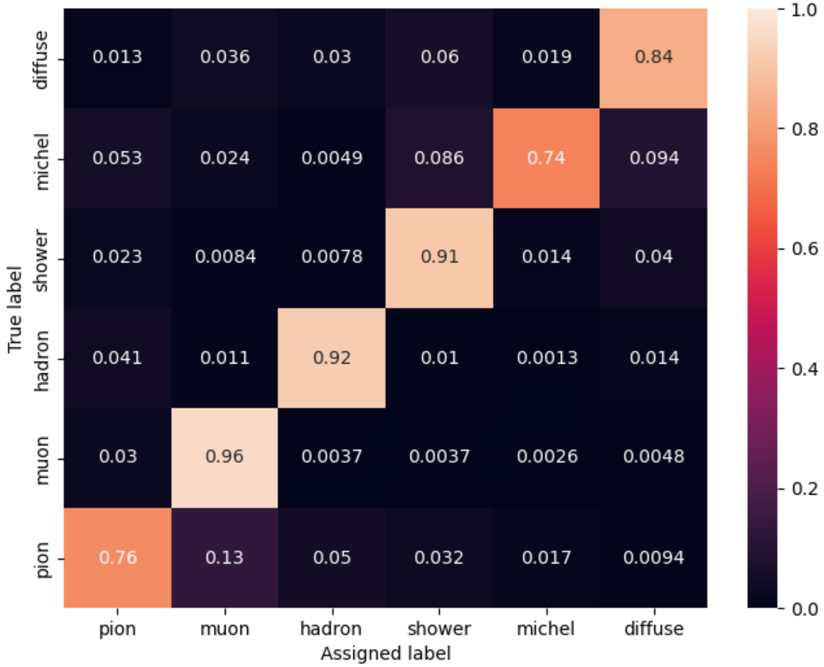


- NuGraph2 consisted of planar and nexus nodes connected in a pseudo-hierarchical fashion
- Nexus nodes primarily provided a way for enforcing consistency between semantic segmentation in each view
- Predicting event-level information was only possible through an aggregation layer (LSTM, transformer, etc)

- To test hierarchical message passing, added an event layer with a single node
- Message passing with learned edge weights between nexus nodes and the event node allows for lightweight and smart aggregation

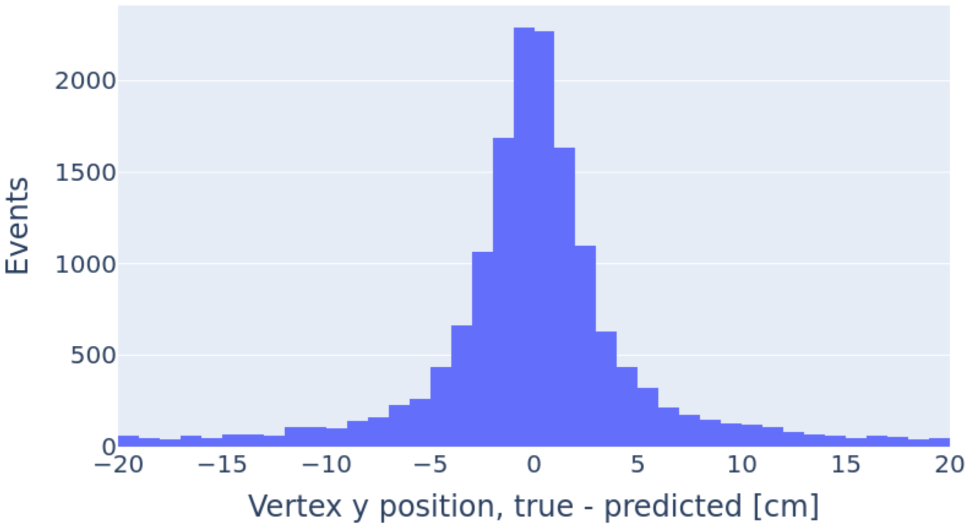


Hierarchical Message Passing



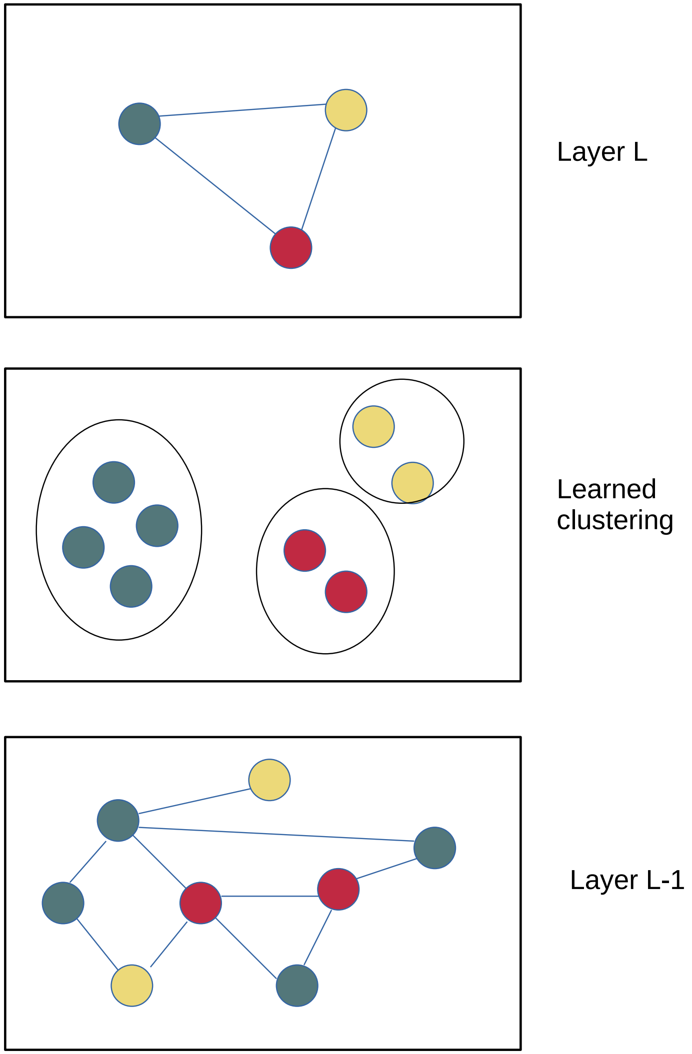
- Semantic performance of NuGraph3 is comparable to NuGraph2 despite breaking MIP category into muons and pions
  - Hierarchical message passing does not diminish performance of NuGraph2

- Features generated at the event node are ideal for extracting reconstructed quantities associated with the full graph
- Regressing interaction vertex position yields excellent resolution and light tails



Dynamic Graph Generation

- Building the hierarchical structure will require dynamic graph generation
- Message-passing iterations in L-1 layer produce predictions for coordinates inside a clustering space based on an object ground truth defined for that layer
- Nodes are clustered together in clustering space
  - Each cluster corresponds to a node in layer L
  - Nodes in L-1 can belong to multiple nodes in L
  - Edge weights between L-1 and L reflect relative certainty of cluster membership
- Generate edges within layer L
  - Number of nodes decreases sharply as L increases, so fully connected graphs may be feasible
- Continue constructing levels to match desired structure to extract



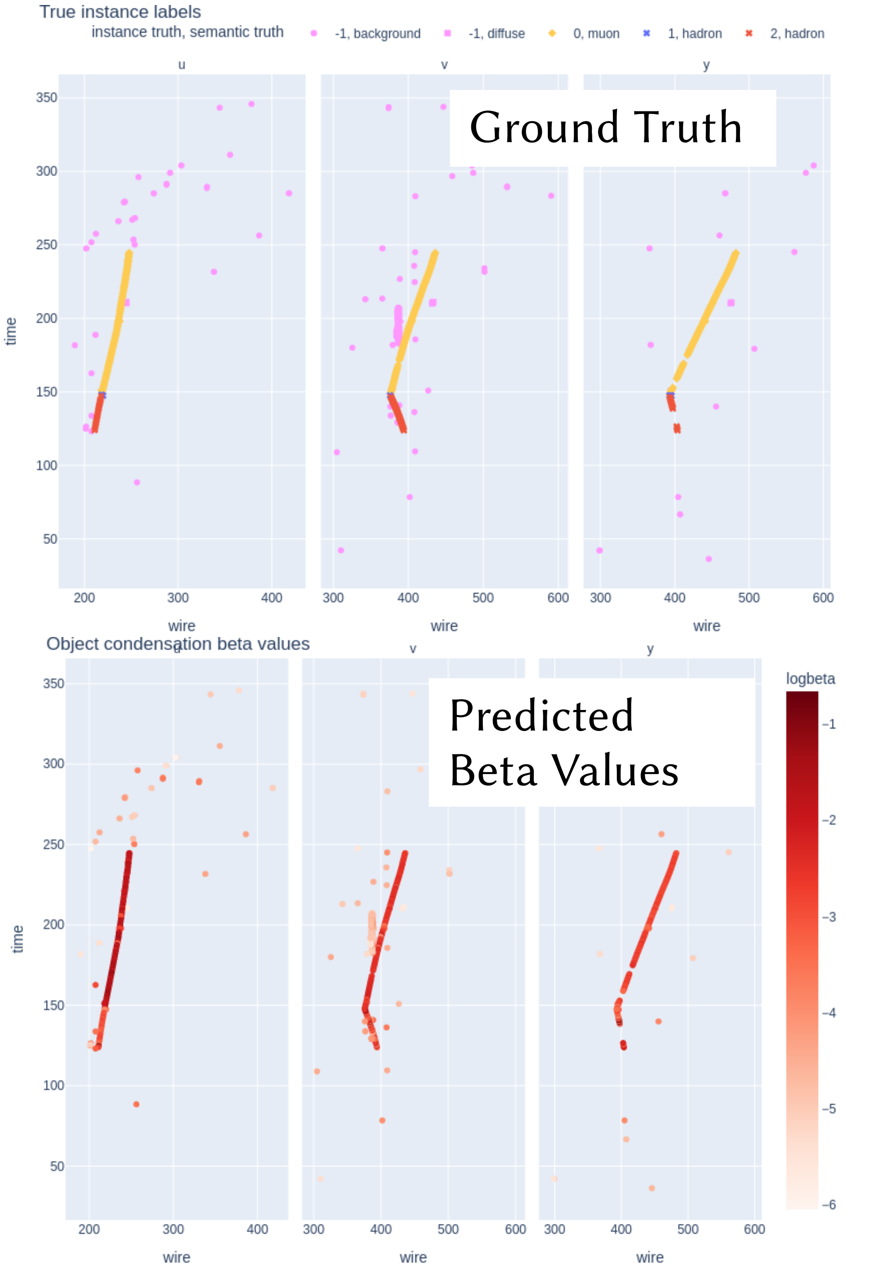
Object Condensation

- Object condensation is a grid-free approach based on an electro-static analog
- Predict a quantity β_i between 0 and 1 for each node
- This quantity will be used to assign a charge
- Points with maximum charge will be used as condensation points
  - Representative points around which clusters will be formed
- A loss is added which encourages a single condensation point per object

$$q_i = \text{artanh}^2 \beta_i + q_{\text{min}}$$

$$L_\beta = \frac{1}{K}\sum_k (1 - \beta_{\alpha k}) + \frac{s_B}{N_B}\sum_i^N n_i \beta_i$$

J. Kieseler, arXiv:2002.03717



Object Condensation

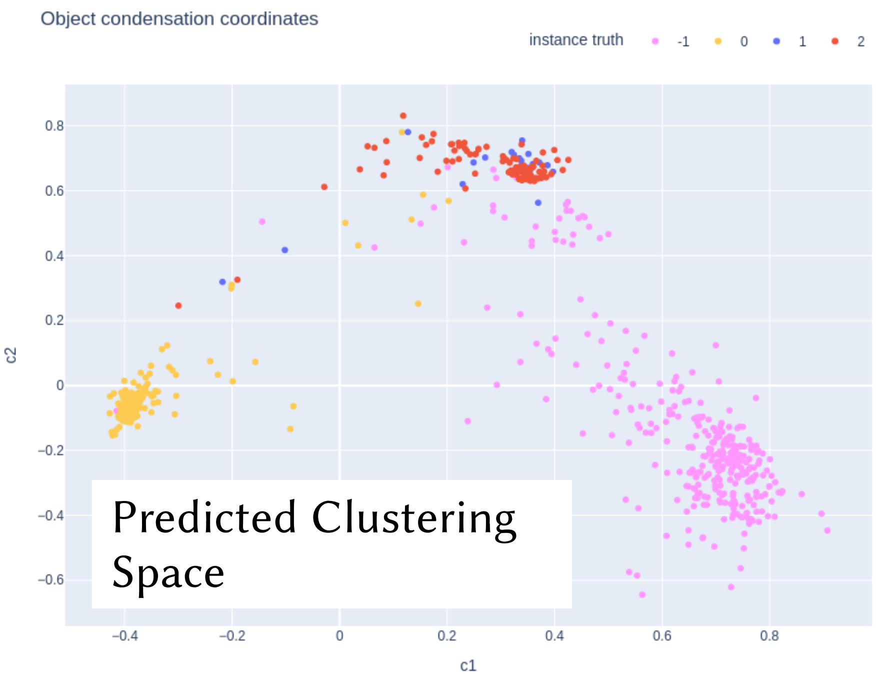
- Predict coordinates of each node in an abstract clustering space
- Attractive and repulsive potentials are defined such that nodes belonging to the same object are attracted and those from different objects are repelled
- Points with distance < 1 from a condensation point are clustered together

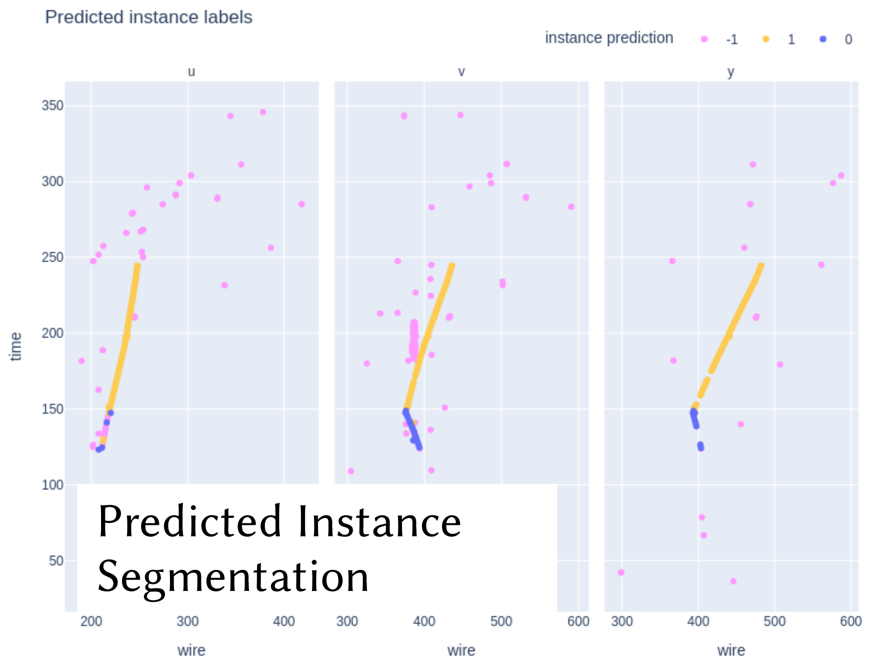
$$A_k = \| x - x_\alpha \|^2 \, q_{\alpha k}$$

$$R_k = max(0, 1 - \| x - x_\alpha \|) q_{\alpha k}$$

$$L_V = \frac{1}{N} \sum_{j=1}^{N} q_j \sum_{k=1}^{K} \left(M_{jk} A_k(x_j) + (1 - M_{jk}) R_k(x_j) \right)$$

M_{jk} = 1 if node j in object k
M_{jk} = 0 otherwise





Summary

- NuGraph2 is a multi-purpose GNN architecture for reconstructing neutrino interactions in MicroBooNE
  - Efficiently reject background detector hits
  - Classify detector hits according to particle type
- Next generation NuGraph3 to focus on full "particle flow" reconstruction
- Adding event layer efficiently aggregates information across full graph
- Use NuGraph2 as a starting point while adding hierarchical structure
  - Use object condensation to dynamically generate the initial hierarchical graph with structure that matches our understanding of the structure of neutrino interactions
  - Hierarchical message passing refines the dynamically generated structure to infer true particle tree
  - Condensation points can be use for inferring particle properties at different hierarchical levels
- First attempt at using object condensation to generate particle instances is encouraging
  - Clustering to create a particle instance layer is being implemented now