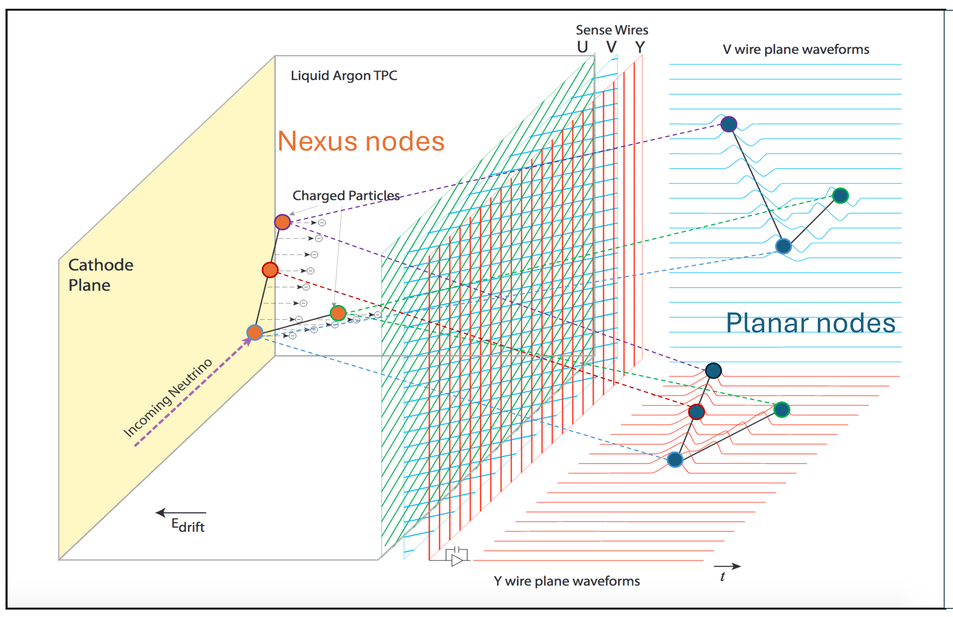
NuGraph2: A Graph Neural Network for Neutrino Event Reconstruction

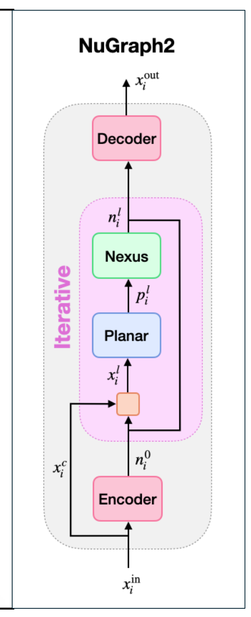
G. Cerati (FNAL) [cerati@fnal.gov],
M. Voetberg, J. Kowalkowski (FNAL), H. Meidani, V. Grizzi (UIUC), D. Caratelli, M. Fang (UCSB),
A. Aurisano, V Hewes (UCincinnati), D. Grzenda, K. Gumpula, X. Zhang (UChicago), C. S. Lee and W. Liao (NorthwesternU)

FERMILAB-POSTER-24-0108-CSAID

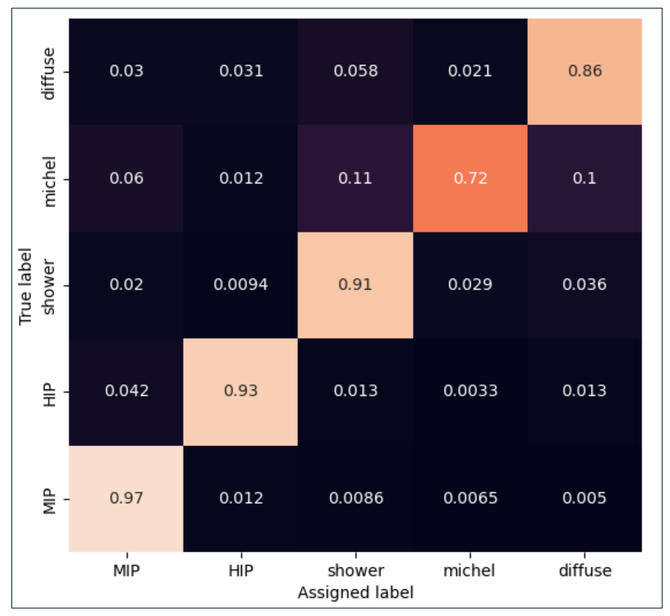
Network definition and performance on simulation

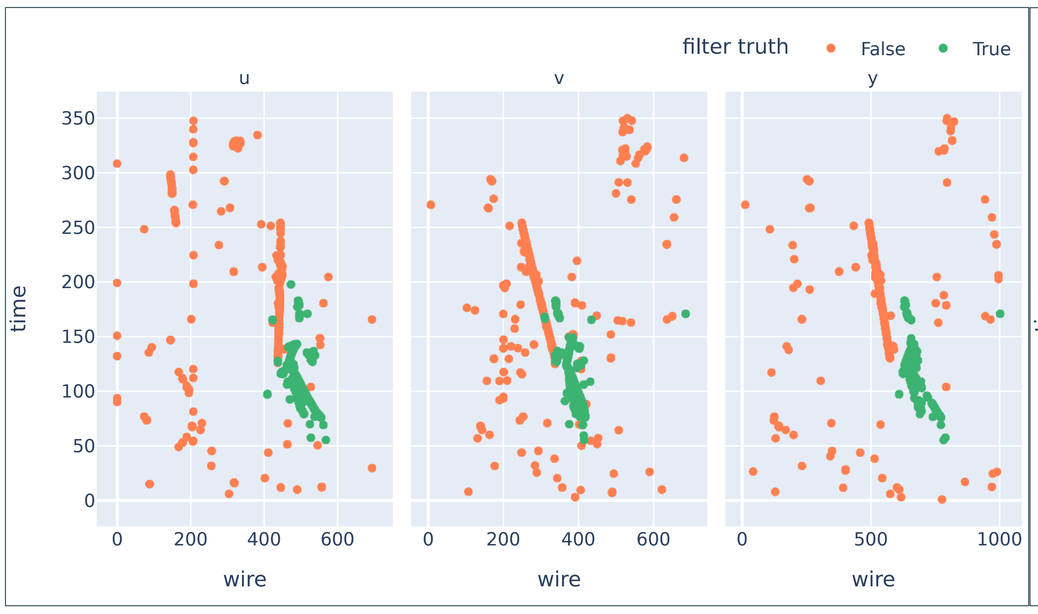
Paper accepted by Phys. Rev. D [arXiv:2403.11872]

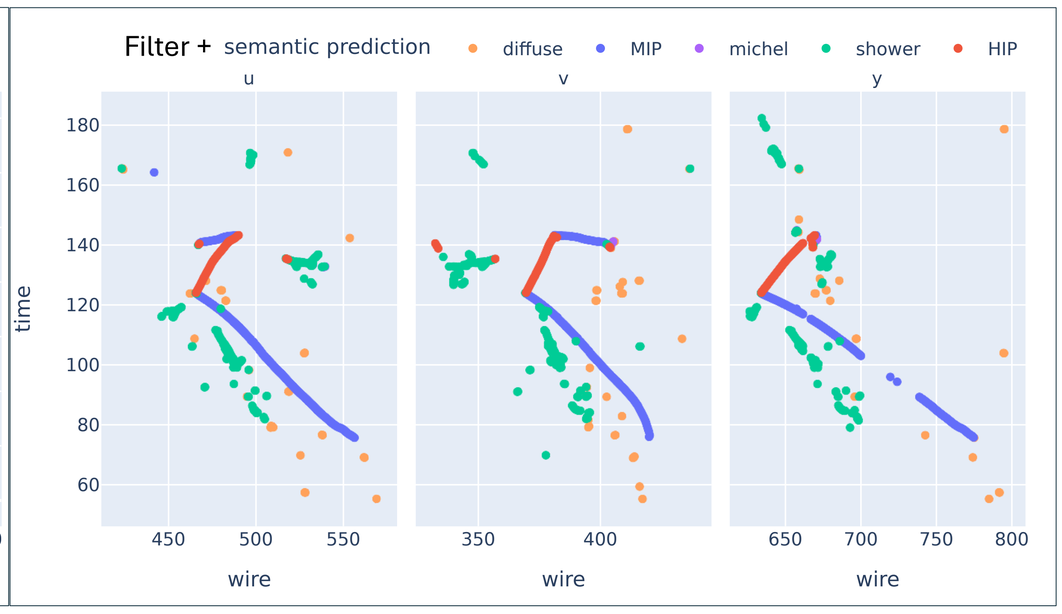




- NuGraph2 is a **Graph Neural Network** for event reconstruction in LArTPC experiments.
- A graph is built connecting **2D hits (nodes)** within planes using Delaunay triangulation. Hits across planes are connected to **3D "nexus" nodes** based on the Space Point Solver algorithm.
- **Input features** to the network encoder are: hit wire, peak time, RMS, integral.
- Information flows in 5 **message passing** iterations across the planar and nexus edges.
- After the last step two **decoders** extract the physics output: filter (neutrino vs background hit) and semantic (label hits by particle type).





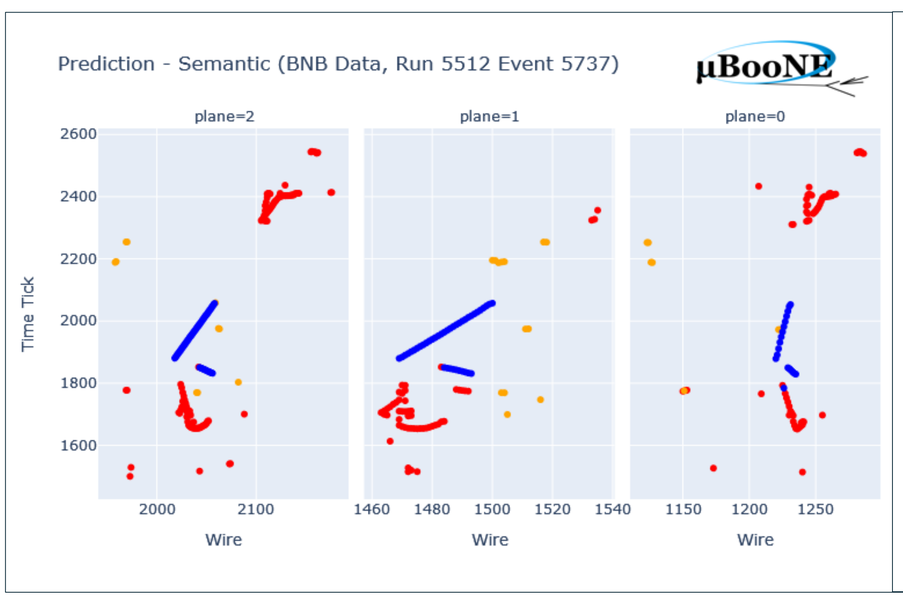


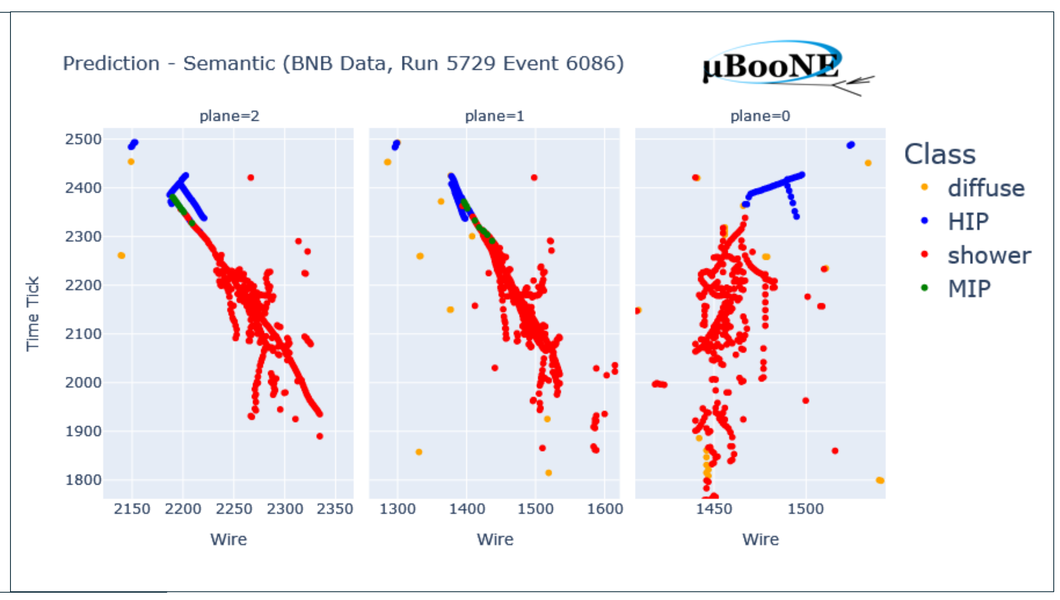
- NuGraph2 is trained on the **MicroBooNE open data** sets [arXiv:2309.15362].
- It achieves an **accuracy of 98% for filter and 95% for semantic** decoder.
- The nexus edges ensure a **consistent 3D interpretation** of the event at ≥95% level.
- Event displays show that NuGraph2 can identify cosmics close to the ν interaction and resolve non-trivial topologies.

See **poster #150** in this session NuGraph3 developments towards a full GNN-based event reconstruction in LArTPC detectors.

Integration in LArSoft and First Tests on MicroBooNE data

- Inference time on CPU is < 1 s/event, including graph construction. When running on GPU in batches, inference time is reduced down to 0.005 s/event.
- NuGraph2 is **integrated in LArSoft** (larrecodnn/NuGraph package), starting form v09_83_01 release. Requires libtorch v2_1_1, torch_scatter v2_1_2, delaunator v0_4_0.
- Integration enables NuGraph2 to run in reconstruction workflows of the LArTPC experiments and facilitates testing NuGraph2 on real data.

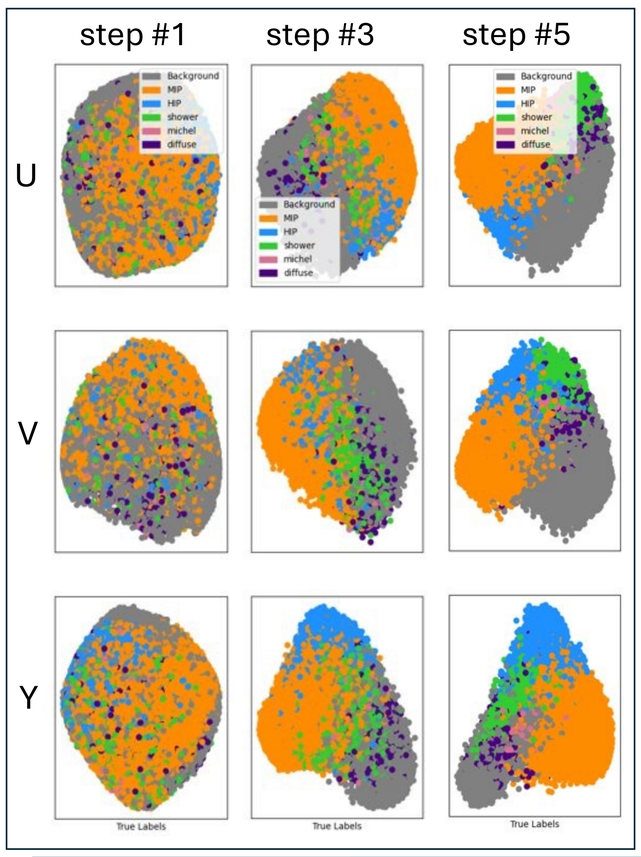




- First tests on MicroBooNE data events passing a loose ν_e CC preselection show **encouraging performance** for the **semantic decoder**: NuGraph2 correctly tags shower hits both from primary electrons (left plot) and photons (right) from π^0 background.
- The filter decoder seems to overly reject shower hits from the neutrino interaction, so **domain shifts** between the training data set and the application data are being investigated.

See **posters #130 and #389** in this session for applications of NuGraph to τ neutrino event reconstruction in DUNE.

Network Explainability and Injection of Physics Domain Knowledge

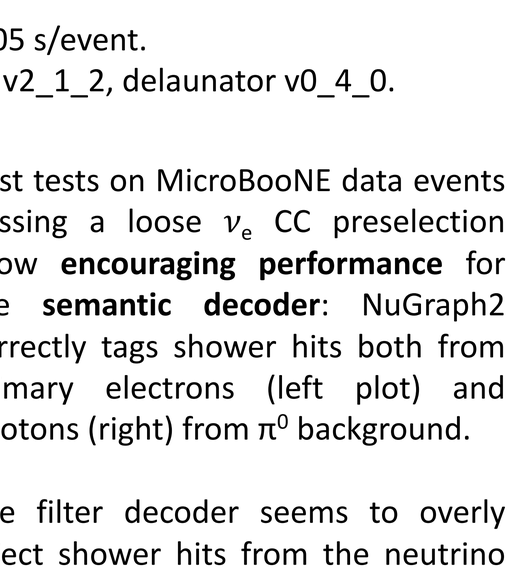


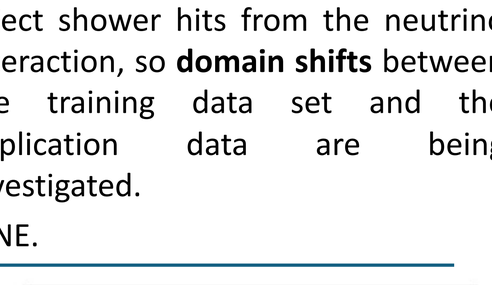
Explainability: *Goal is to "open the black box" to build confidence and drive developments.*

1. Visualization of latent space (left figure)
   - Cluster latent space features and project in 2D for visualization
   - Separation between different categories achieved by the last (5^{th}) network iteration
2. Understanding the role of "hub" nodes (right figures)
   - Detached nodes with large edge multiplicity connect nodes within and across objects
   - Pruning test: 12 is the lowest upper bound in multiplicity without affecting performance, when pruning edges uniformly in terms of length
   - Demonstrates that there is a large degree of redundancy and that both short and long edges matter. Can also lead to network speedups.

Domain Knowledge: *Goal is to drive the learning and mitigate training dataset limitations.*

1. Force learning correlations between semantic classes
   - Add a decoder to regress the fraction of hits in each semantic category
   - Evaluation of impact on the network performance is in progress.
2. Add non-local and topological node features
   - Add the following quantities to the input node features: Δtime and Δwire between the two closest nodes, distance to closest node, edge multiplicity
   - Improves the network accuracy by ~5% (relative) for the Michel category





This document was prepared by the Exa.TrkX/NuGraph Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, Office of High Energy Physics HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.