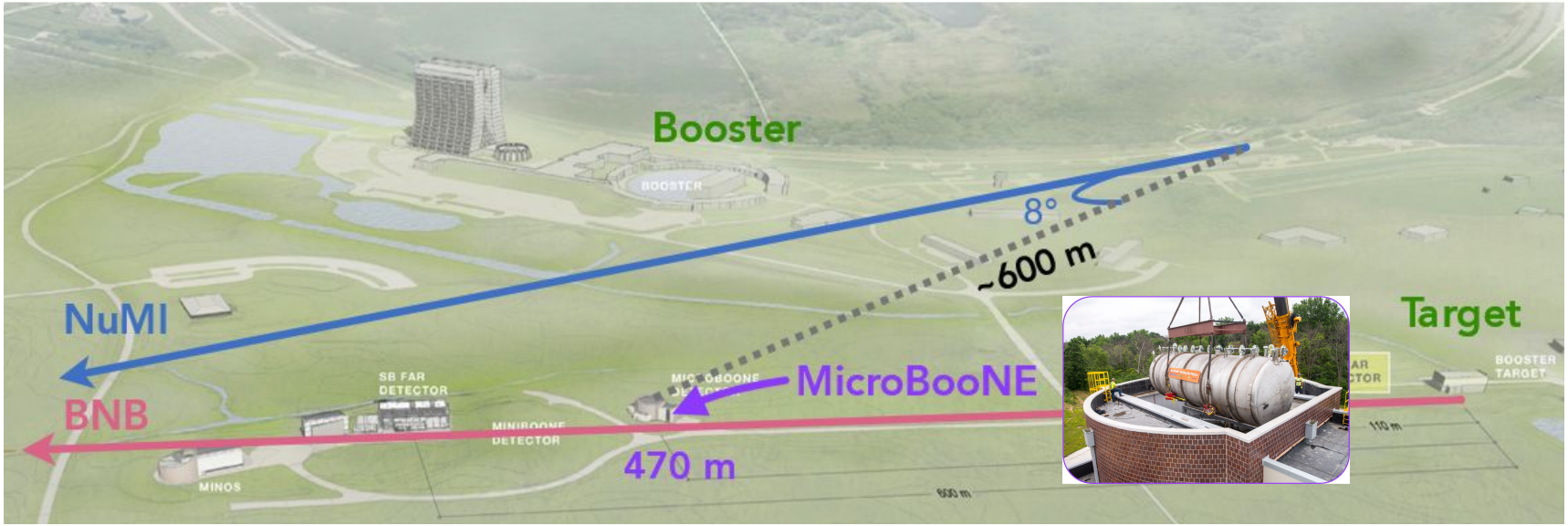
Integrating NuGraph2 in MicroBooNE Reconstruction

Chuyue “Michaelia” Fang
on Behalf of the MicroBooNE Collaboration
cfang@ucsb.edu
NPML 2026 @ UC Irvine
6/15/2026

Outline

- Introduction to MicroBooNE and NuGraph2
- Effort in deploying NuGraph2 on MicroBooNE data
  - challenges encountered during integration and corresponding mitigation strategies
- Impact on reconstruction and physics analyses
  - ν_e, π^0, p, $\pi^{\pm}$, e^+e^-...



- MicroBooNE was one of the first large LArTPC experiments to operate in a neutrino beam
  - on-axis to BNB beam, 8° off-axis to NuMI beam
- took data from 2015-2021 at Fermilab

MicroBooNE Physics Goals

- short-baseline oscillations
- ν-Ar cross-sections
- BSM searches
- …

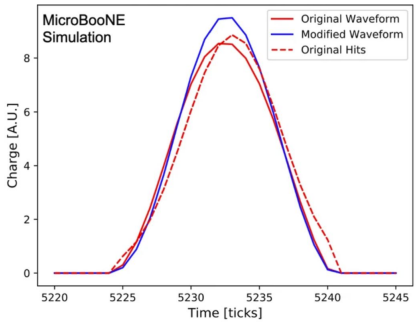
Requires Robust Reconstruction of Key Final States:

- ν_e key signature for oscillations
- π^0 common background to ν_e and BSM signatures
- p, $\pi^{\pm}$ hadronic final states in cross-section measurements
- e^+e^- BSM signature
- …

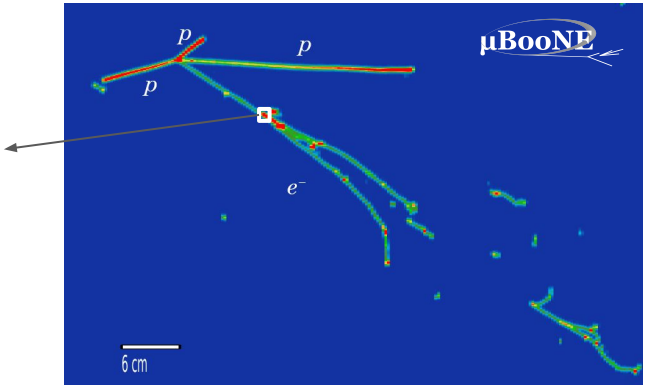
Introduction to NuGraph2

Phys. Rev. D **110**, 032008

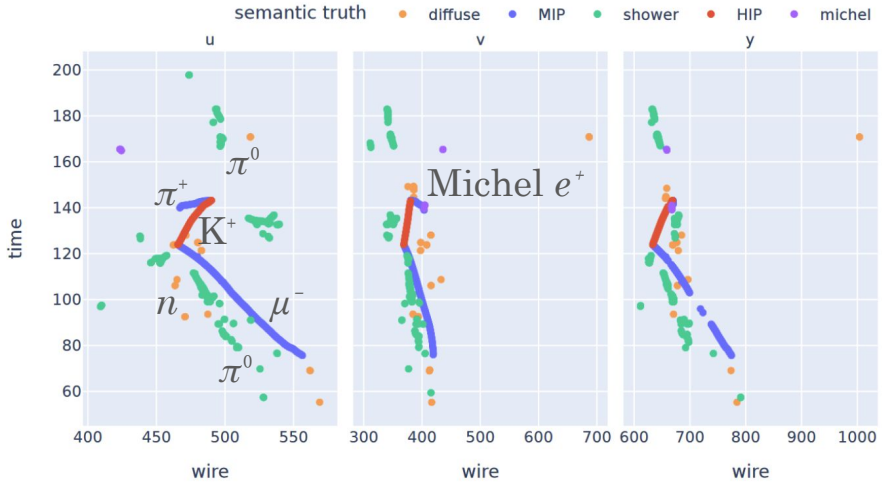
- a graph neural network for neutrino events reconstruction
  - each detector hit is considered a node
  - makes predictions on each hit for **cosmic backgrounds** and **particle types**
  - trained on MicroBooNE Public Datasets
  - takes 4 input features:



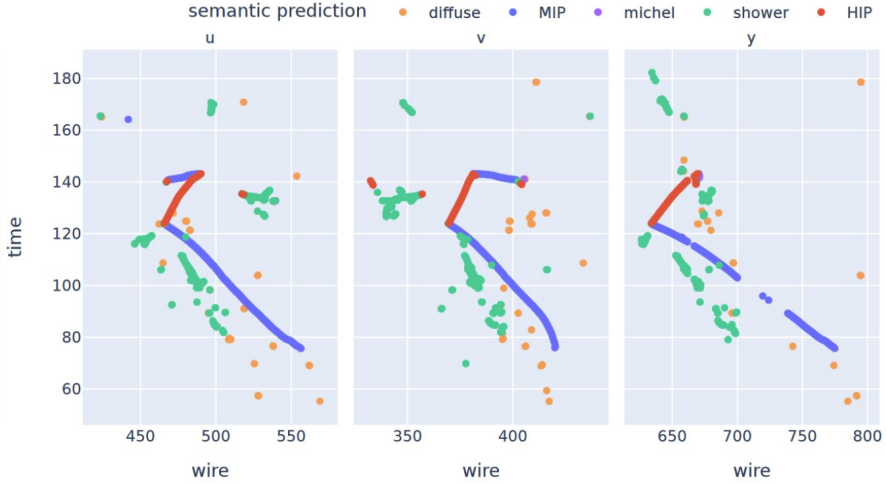
- wire index
- time coordinate
- integral
- RMS width



Introduction to NuGraph2



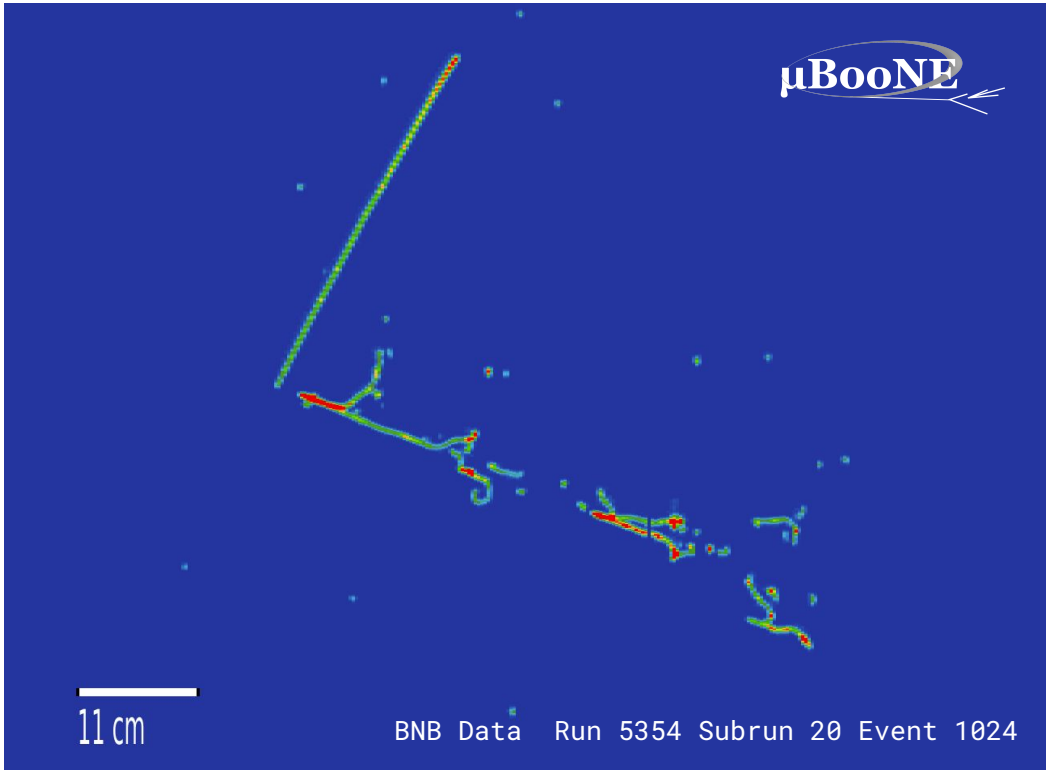
Truth particle type labels of hits



Predicted particle type labels of hits by NG2

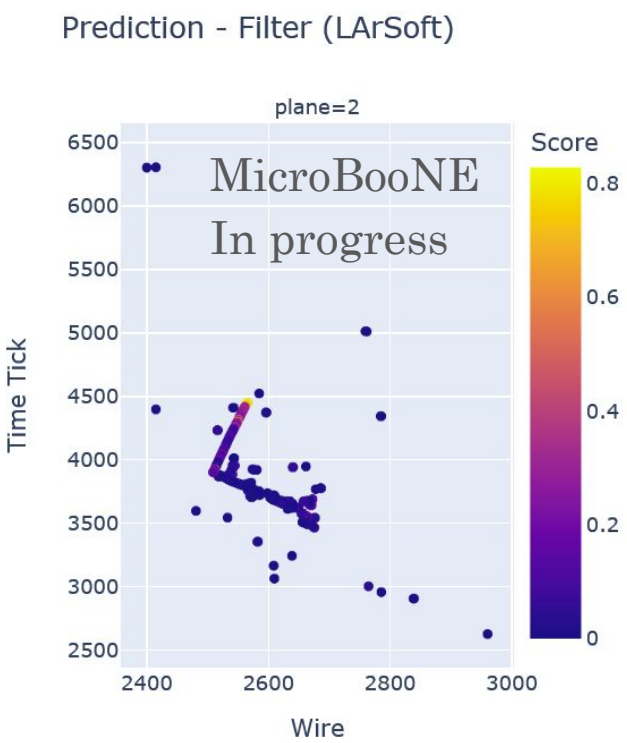
- great performance in simulation
  - cosmic filter: 98% accuracy
  - particle types: 95% accuracy

First Tests of NuGraph2 on MicroBooNE Data



an example event from MicroBooNE data with a proton track and a shower

First Tests of NuGraph2 on MicroBooNE Data

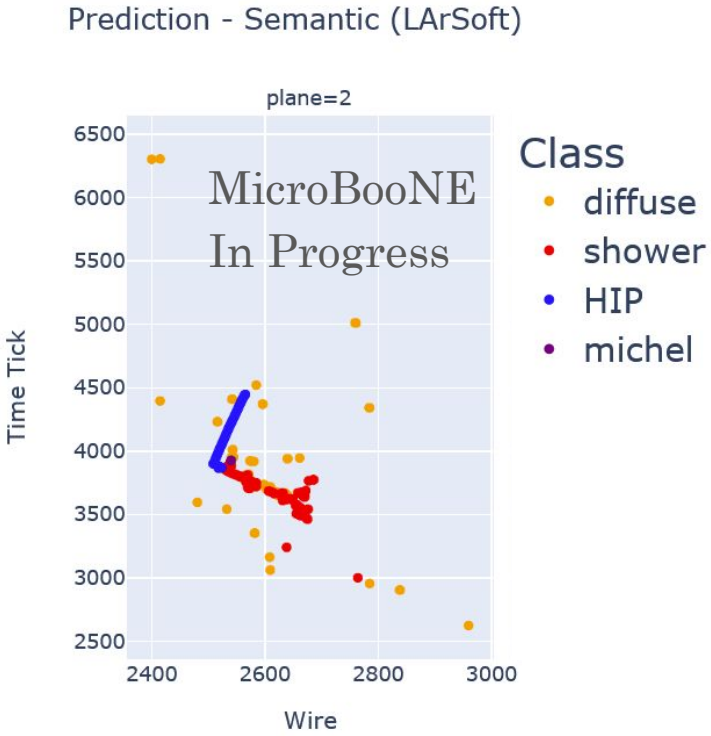


predicted to be from neutrino signal

predicted to be from cosmic background

a large fraction of the hits from neutrino interaction was predicted to be from cosmics!

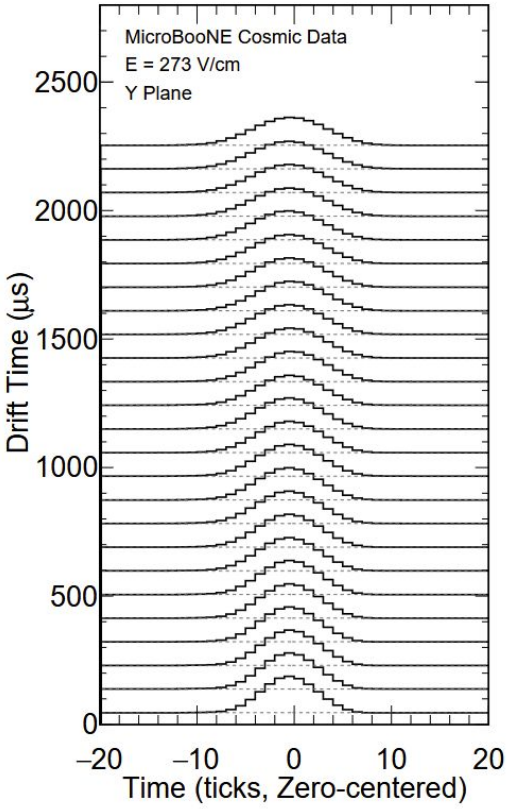
First Tests of NuGraph2 on MicroBooNE Data



- predictions for the particle types are correct for most hits

- Why are these hits from neutrino interaction predicted to be coming from cosmic backgrounds?

- Why are these hits from neutrino interaction predicted to be coming from cosmic backgrounds?
  - original model was trained with simulation which has known issues with diffusion modeling



JINST **16** P09025

diffusion modeling

- electron diffusion spread the ionization cloud as a function of drift time
- affects hit width (input feature “RMS” used in NuGraph2)

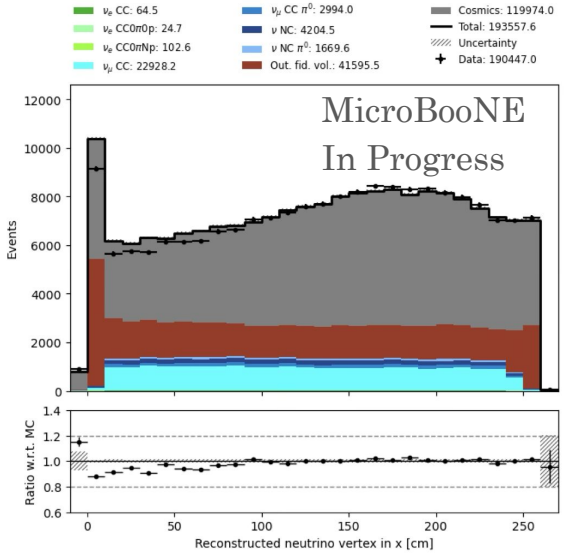
- Why are these hits from neutrino interaction predicted to be coming from cosmic backgrounds?
  - original model was trained with simulation which has known issues with diffusion modeling
  - model was likely trained to rely on data-MC discrepancy to make predictions for cosmic filters

Minimizing Simulation-induced Bias

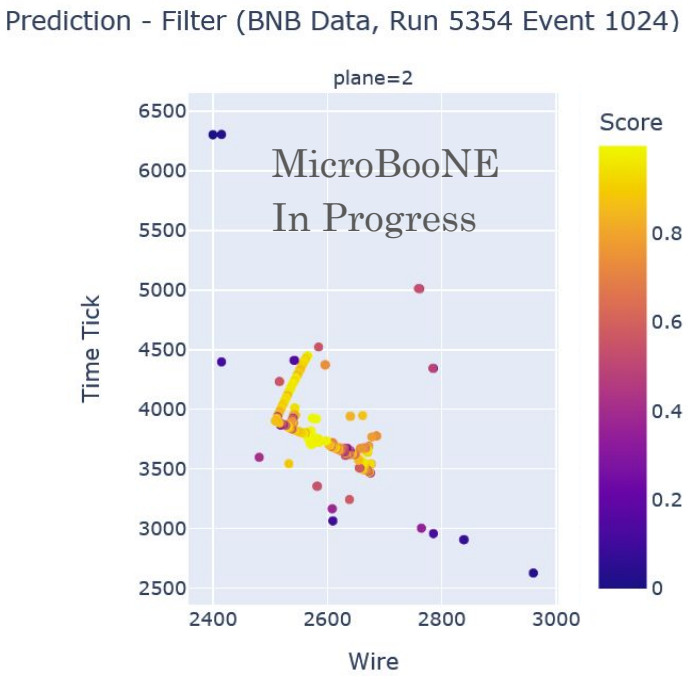
- first trained NuGraph2 with a different simulation using data-driven diffusion constant
  - great improvement in cosmic filters

Minimizing Simulation-induced Bias

- first trained NuGraph2 with a different simulation using data-driven diffusion constant
  - great improvement in cosmic filters
- during validation, still saw bias in cosmic rejection
  - completely removed RMS as an input feature to NuGraph2 with <2% loss in accuracy



Tests of NuGraph2 Trained with Different Simulation

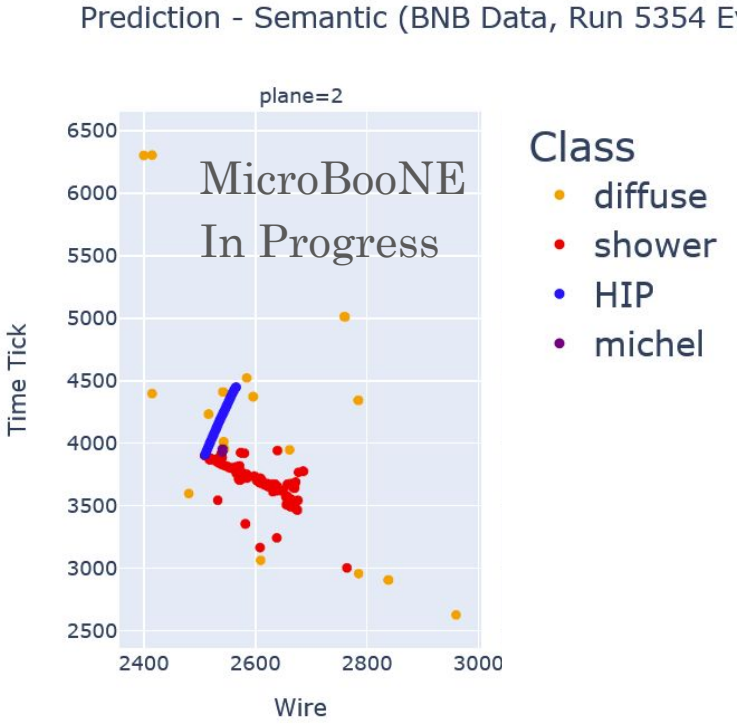


predicted to be from neutrino signal

predicted to be from cosmic background

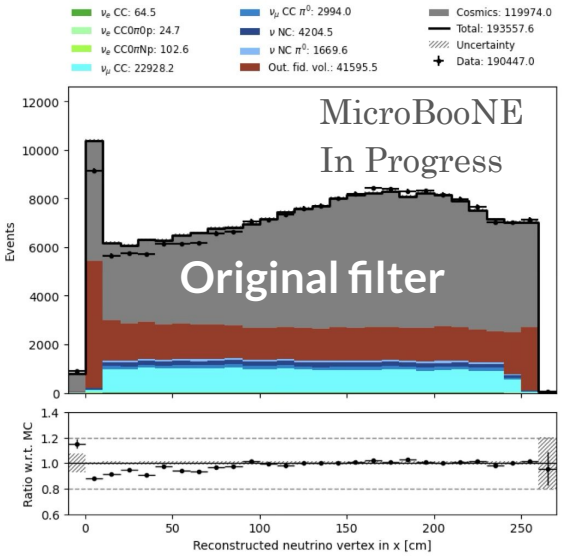
now, most hits are predicted correctly

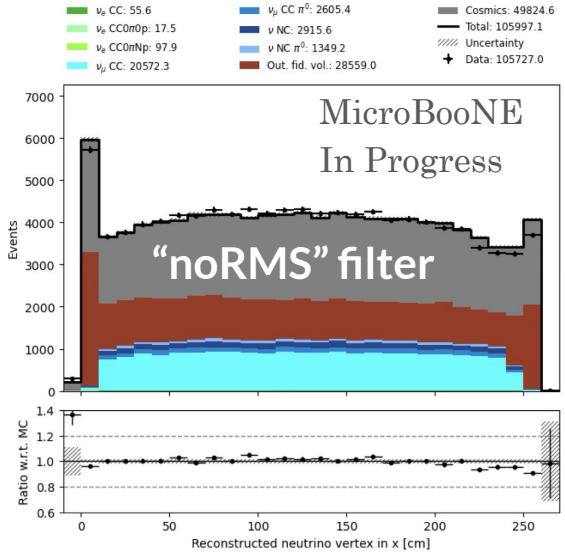
Tests of NuGraph2 Trained with Different Simulation



- predictions for the particle types are correct for most hits
- slight improvement for the start of the shower

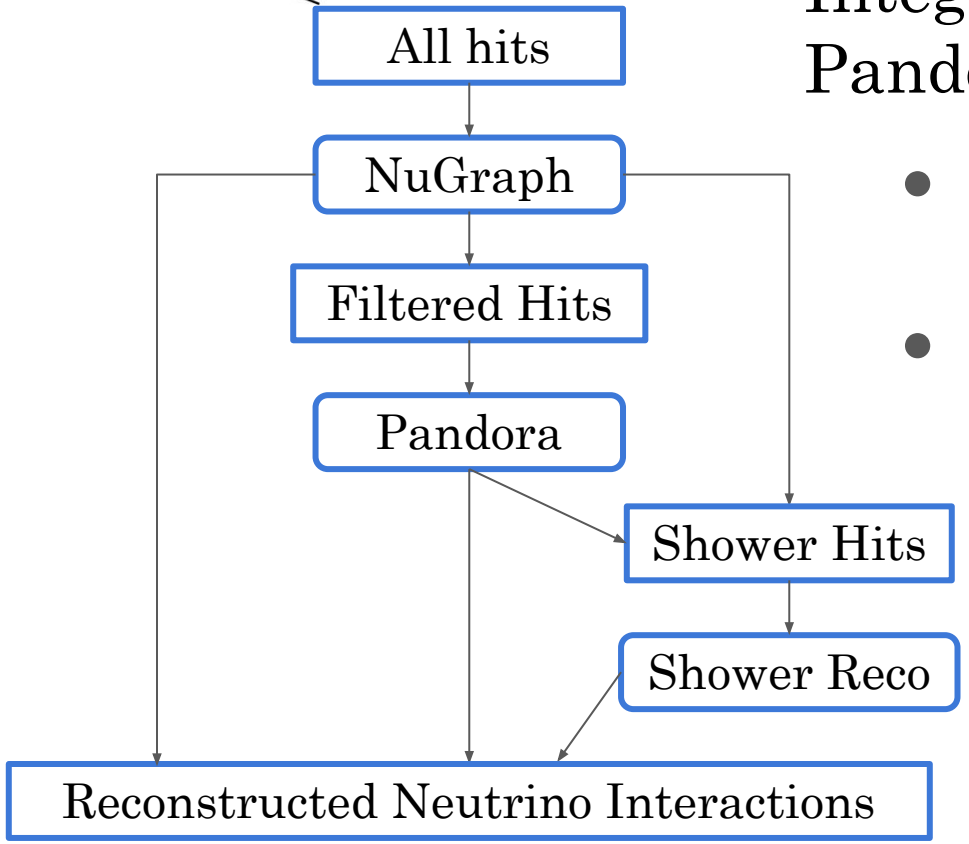
Reduced Bias in Cosmic Rejection





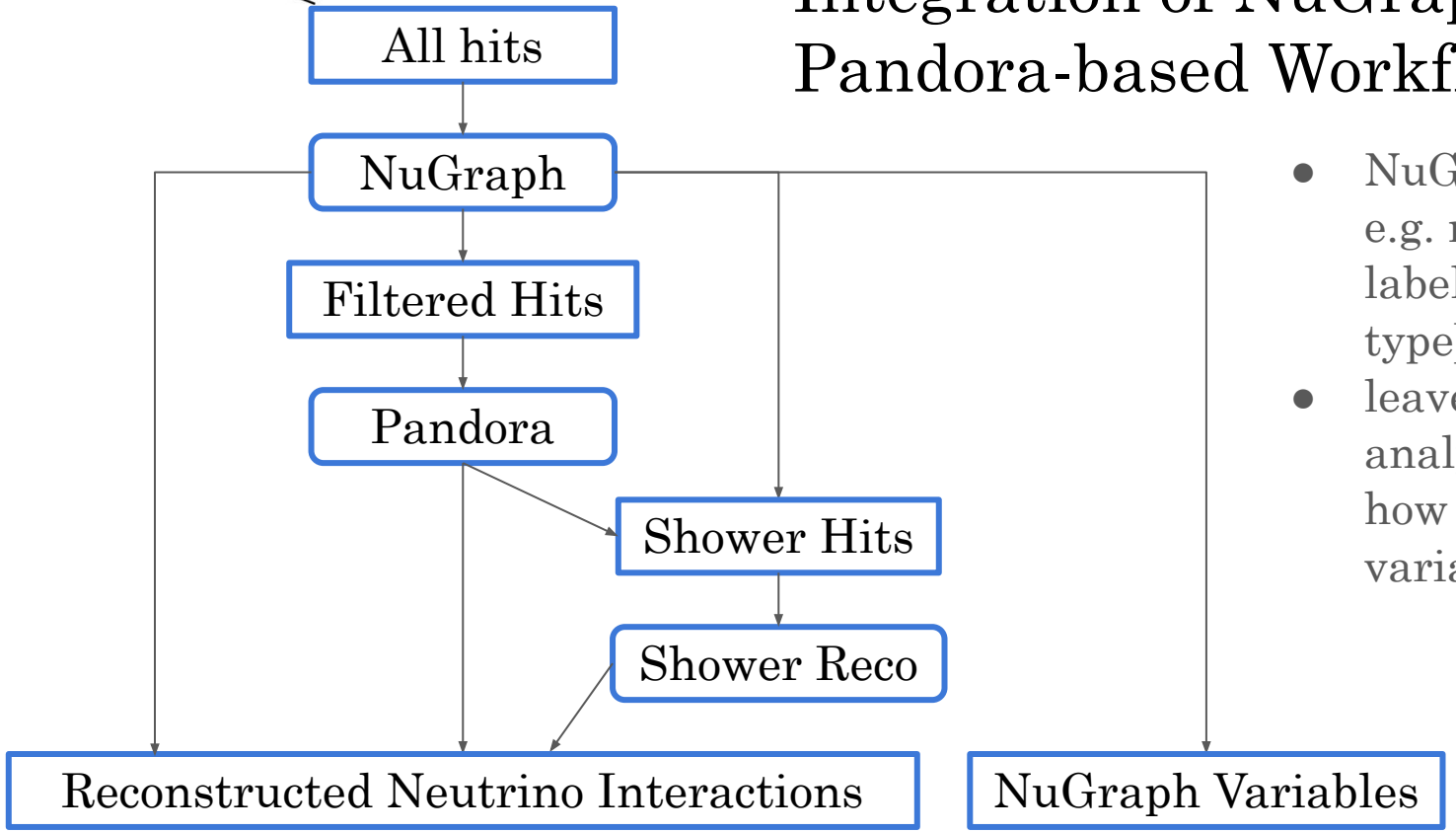
- noRMS filter largely reduced the bias in cosmic rejection

Integration of NuGraph2 in Pandora-based Workflow



- cosmic filter decoder used to reject cosmics
- semantic decoder used for particle identification

Integration of NuGraph2 in Pandora-based Workflow



- NuGraph variables e.g. number of hits labelled as [particle type] by NG2
- leave it to the analyzers to decide how they use these variables

Impact of NuGraph2 Integration

NuGraph2 Integration

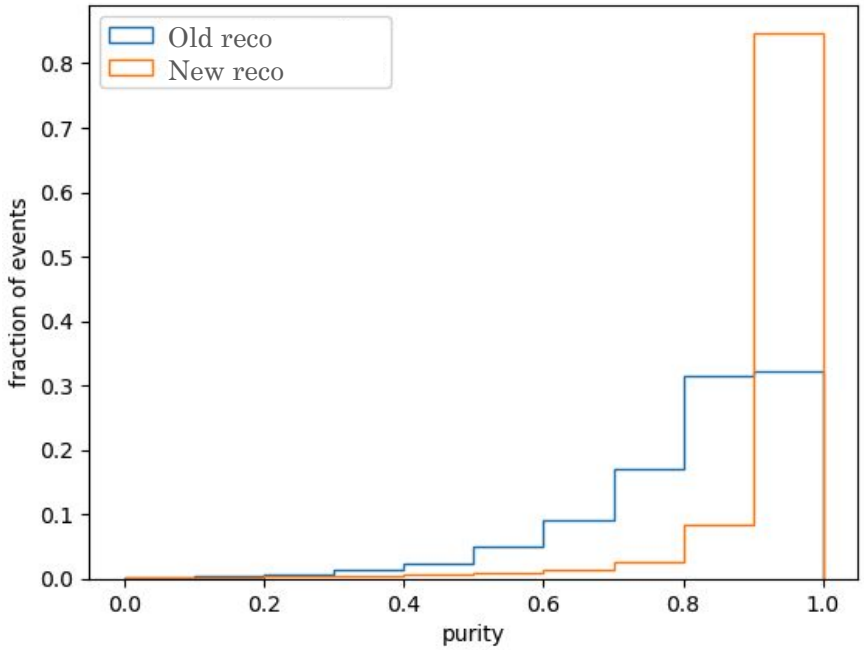
Reconstruction Improvements

- cosmic rejection
- object completeness and purity
- particle identification
- kinematic resolution
- …

Enhancements in Physics Analyses

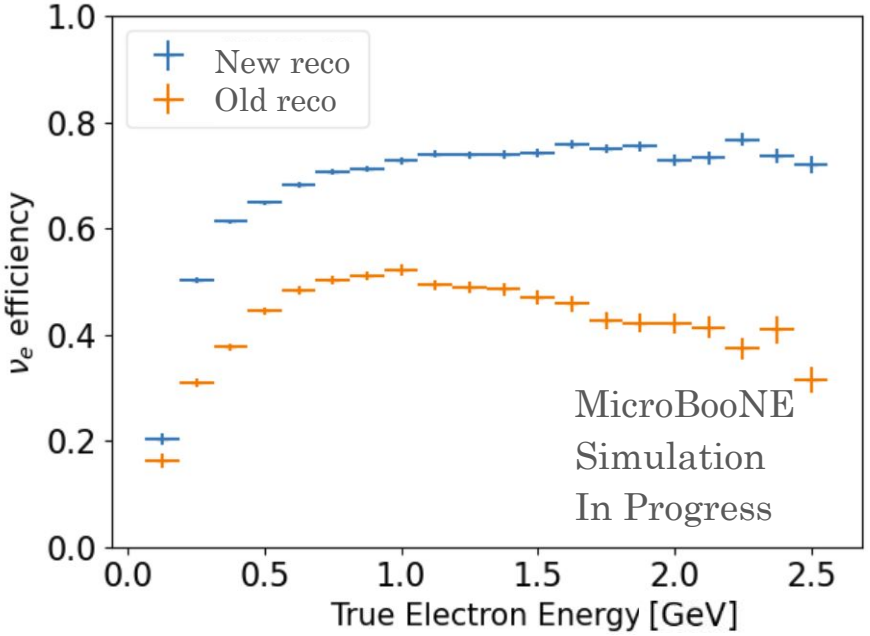
- 1e0p0$\pi^{\pm}$
- CC2$\pi^{\pm}$1p
- dark tridents
- HNL
- …

Cosmic Rejection

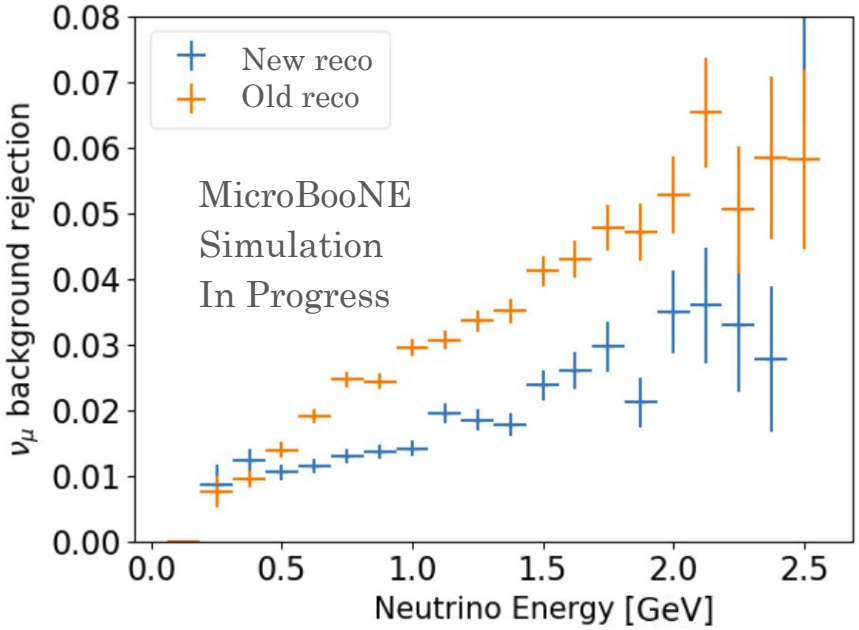


- applying NuGraph cosmic filter largely improved the purity of reconstructed neutrino interactions

ν_e Pre-selection

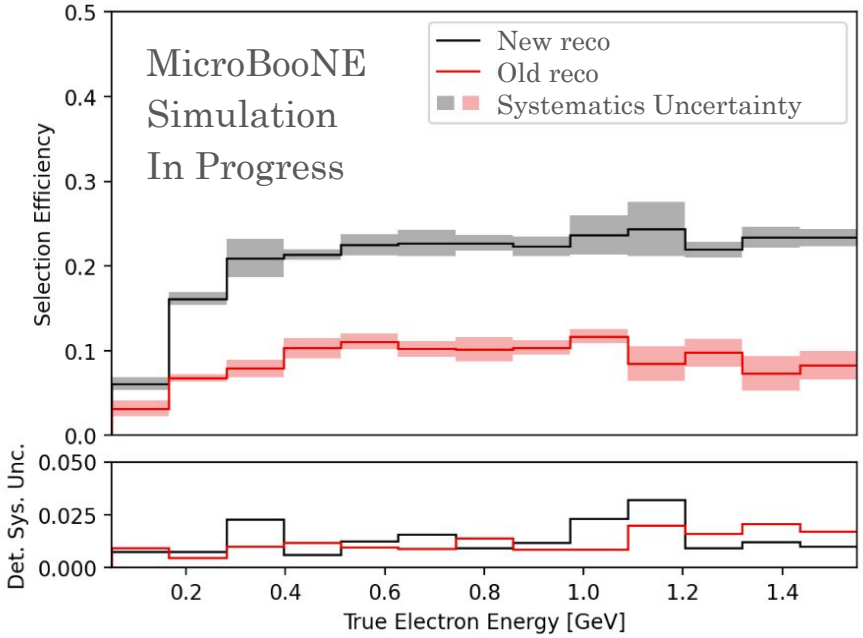


ν_e signal



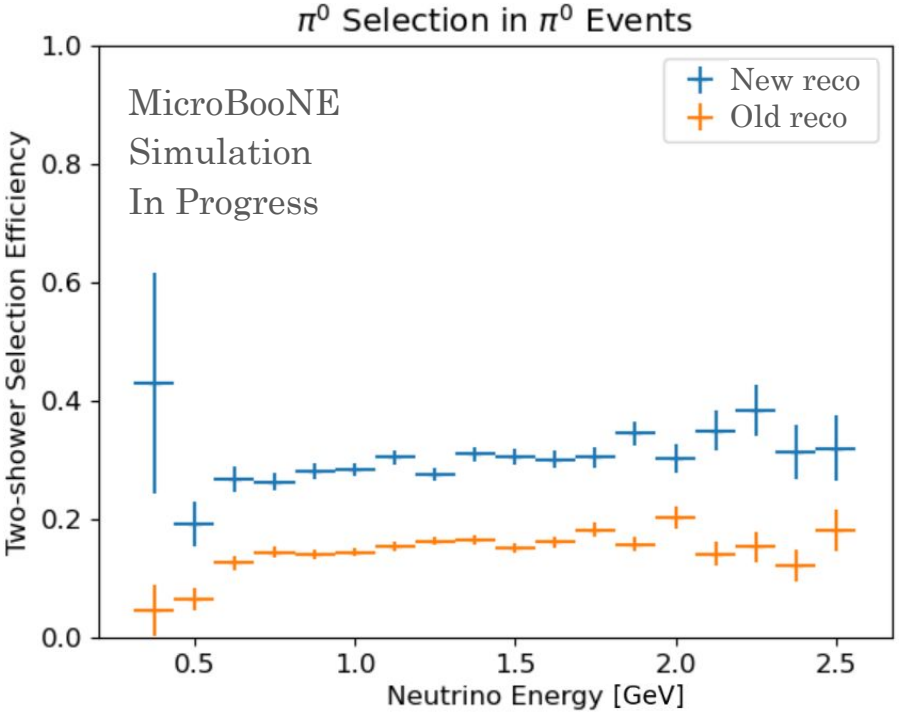
ν_μ background

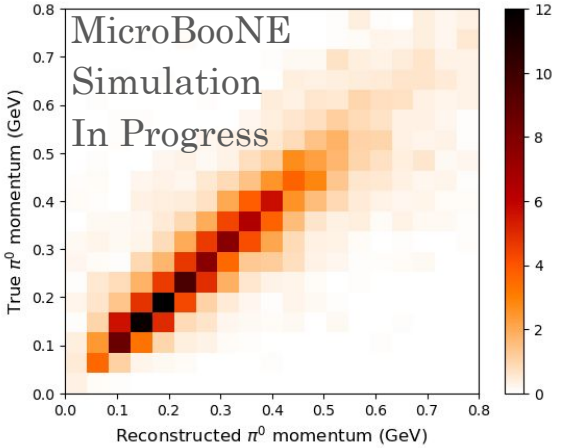
ν_e Selection



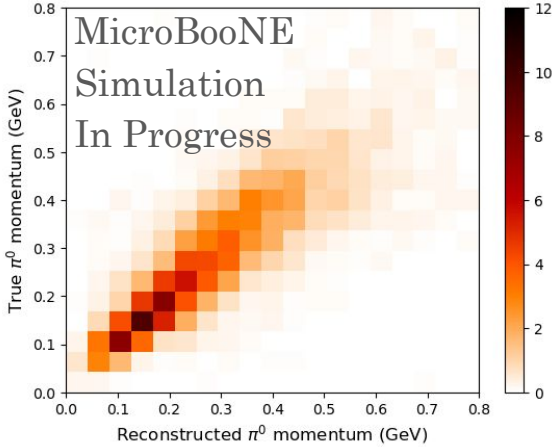
- direct comparison with old reconstruction ν_e selection shows improvement in efficiency while maintaining similar levels of systematics
  - not optimized selection
- demonstrates robustness against detector systematics

π^0 Selection and Momentum Resolution



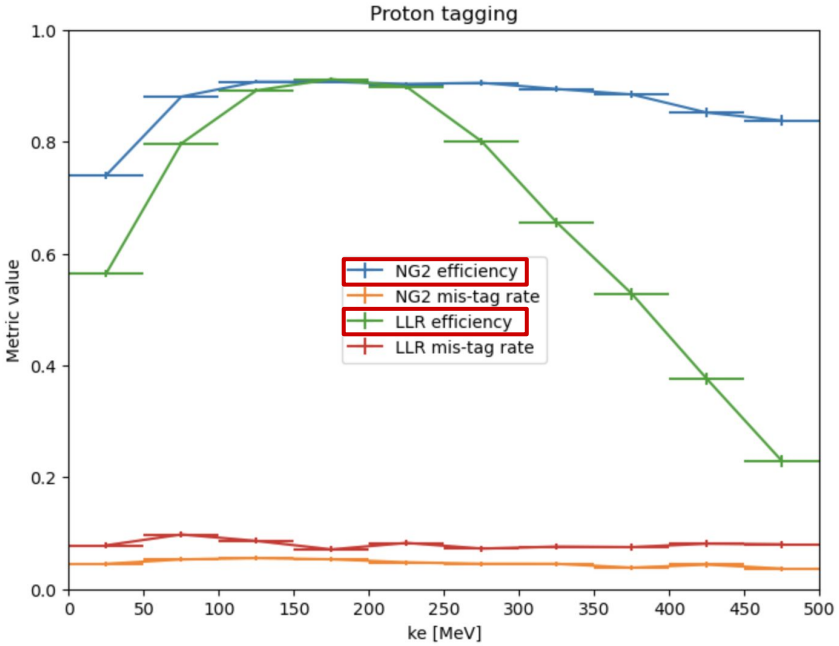


New reco
$\sigma = 15.9\%$



Old reco
$\sigma = 23.5\%$

Proton Tagging



- NuGraph brings higher efficiency in proton tagging
  - even when Bragg peak not clearly visible or well reconstructed

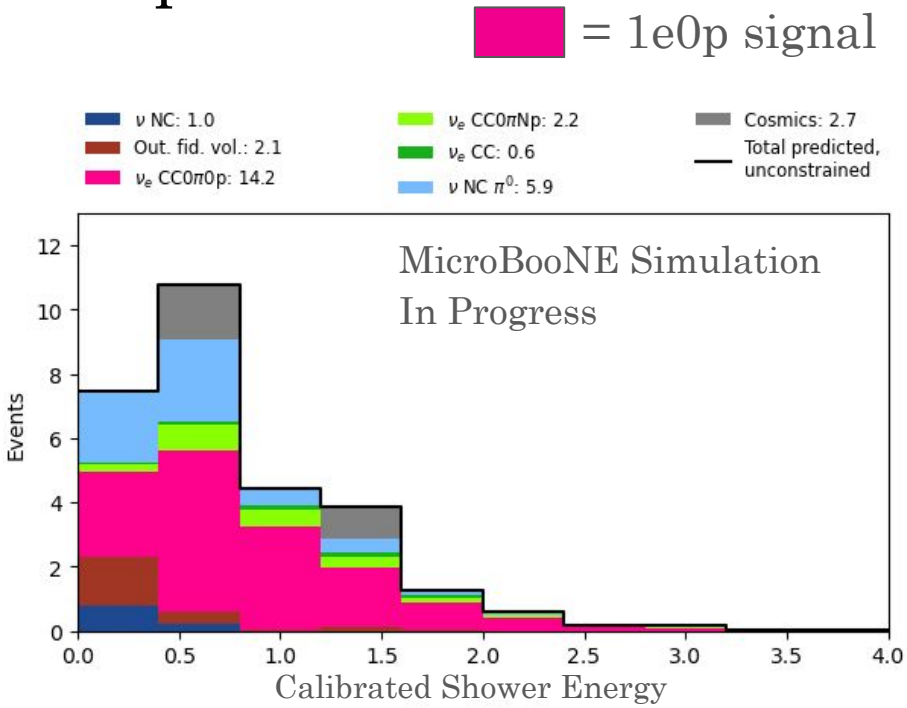
w/ NuGraph

w/o NuGraph

Enhancements in Physics Analyses

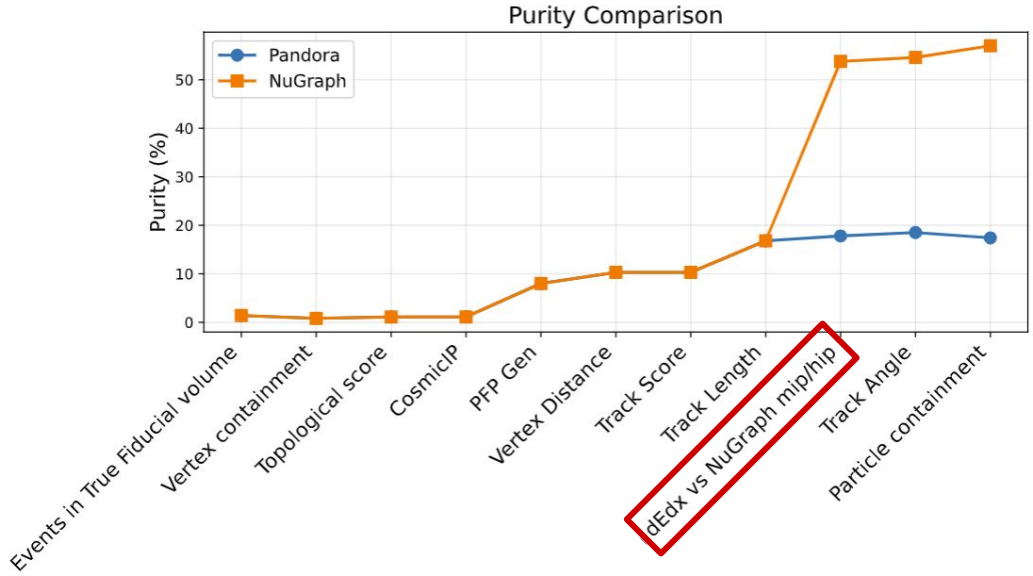
- Ongoing analyses see great improvements from the NuGraph integration
  - 1e0p0$\pi^{\pm}$
  - CC2$\pi^{\pm}$1p
  - dark tridents $\rightarrow e^{+}e^{-}$
  - HNL (heavy neutral leptons) $\rightarrow e^{+}e^{-}$
  - …

1e0p0$\pi^{\pm}$



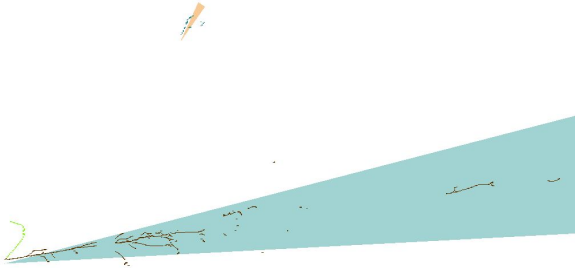
- loose selection efficiency: 37% → 63%
- in-progress BDT trained for 1e0p/1eNp separation ranked NuGraph HIP variable as the most important

$CC2\pi^{\pm}1p$

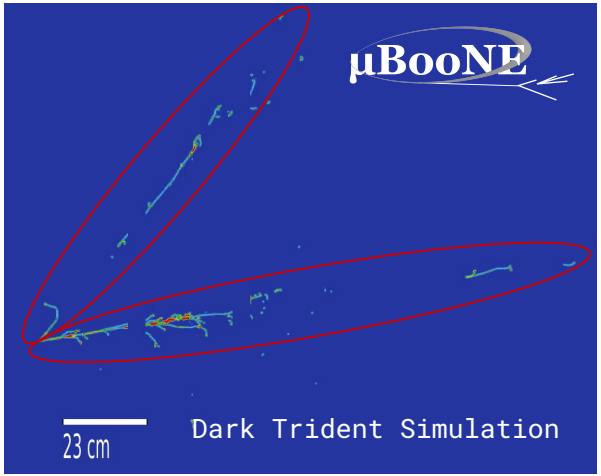


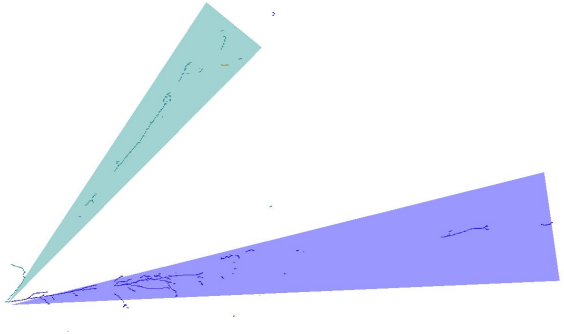
- changed the original dEdx cut to a cut based on NuGraph MIP/HIP variables
- purity: 17.8% → 53.8%

Dark Tridents → e^+e^-



Old reco





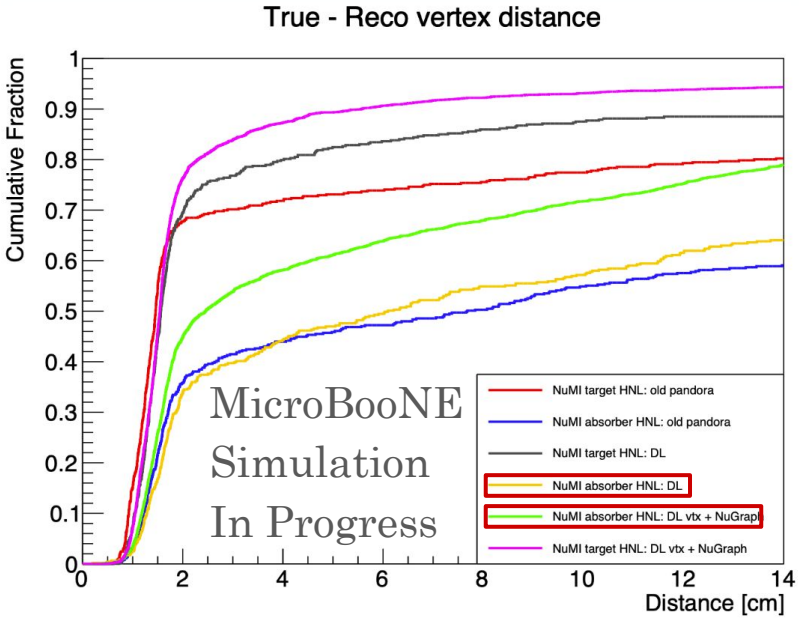
New reco

- early-stage dark trident analysis
- with a few customized changes for better e^+e^- reconstruction

$\mathrm{HNL} \rightarrow e^+e^-$

w/ NuGraph

w/o NuGraph



- improved vertex resolution from a combination of NuGraph and another ML vertexing tool
- e^+e^- topology was something NuGraph2 was not trained for

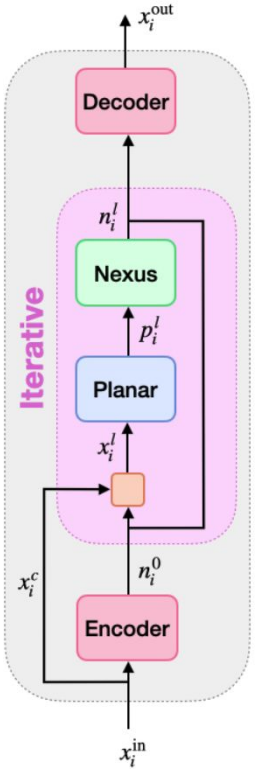
Conclusion

- in MicroBooNE, we have integrated NuGraph2 with Pandora to form a combined reconstruction framework
- encountered challenges with simulation-induced bias
  - took various strategies to minimize bias
- significantly improved performance across the board in reconstruction and ongoing physics analyses
  - ν_e, π^0, p, $\pi^{\pm}$, e^+e^-...

Thank you!

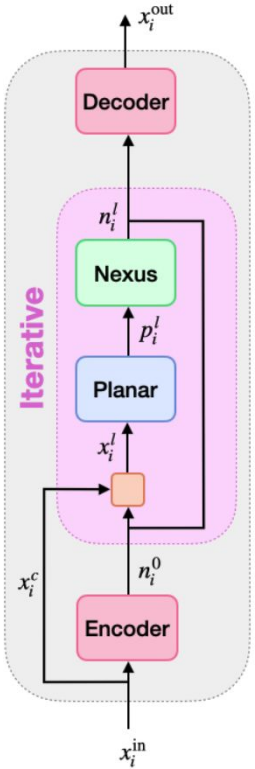
Backup

Backup Slide - NuGraph2 Network Structure



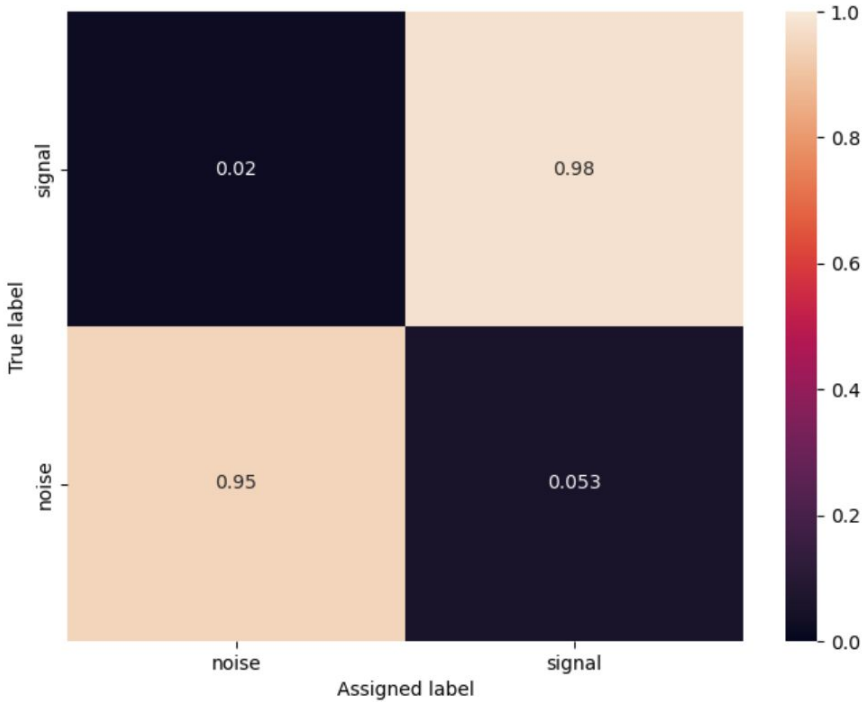
- only hits passing through **Pandora neutrino slice selection**
- a planar subgraph is generated for each plane
- a nexus subgraph is constructed utilizing SpacePointSolver, and it connects planar hits with the same spacepoint

Backup Slide - NuGraph2 Filter and Semantic Decoder



- **filter and semantic decoder outputs probability scores**
  - for filter outputs, cosmic filter predictions are determined with a probability threshold of 0.5
  - for semantic outputs, particle category predictions are made by selecting the class with the largest probability score

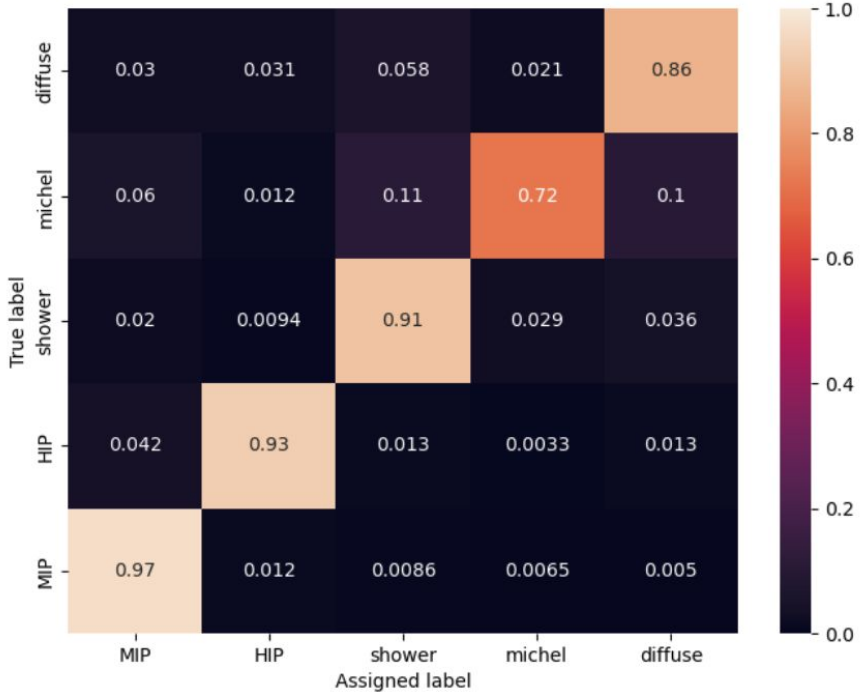
Performance in Simulation - Cosmic Filtering



Efficiency Confusion Matrix

- filters out hits which don't come from **primary interactions of neutrinos**
  - **efficiency: 0.98**
  - **purity: 0.98**

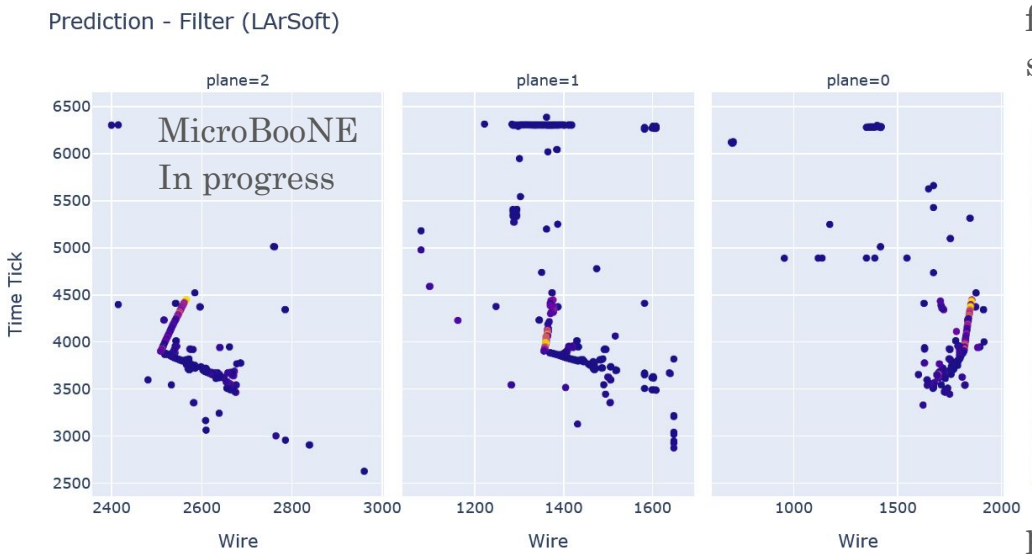
Performance in Simulation - Semantic Labeling



Efficiency Confusion Matrix

- 5 semantic categories: **MIP**, **HIP**, **EM showers, Michel electrons** and **diffuse activity**
  - **accuracy: 0.95**
- efficiency for each class was affected by number of hits in this class in dataset

First Tests of NuGraph2 on MicroBooNE Data

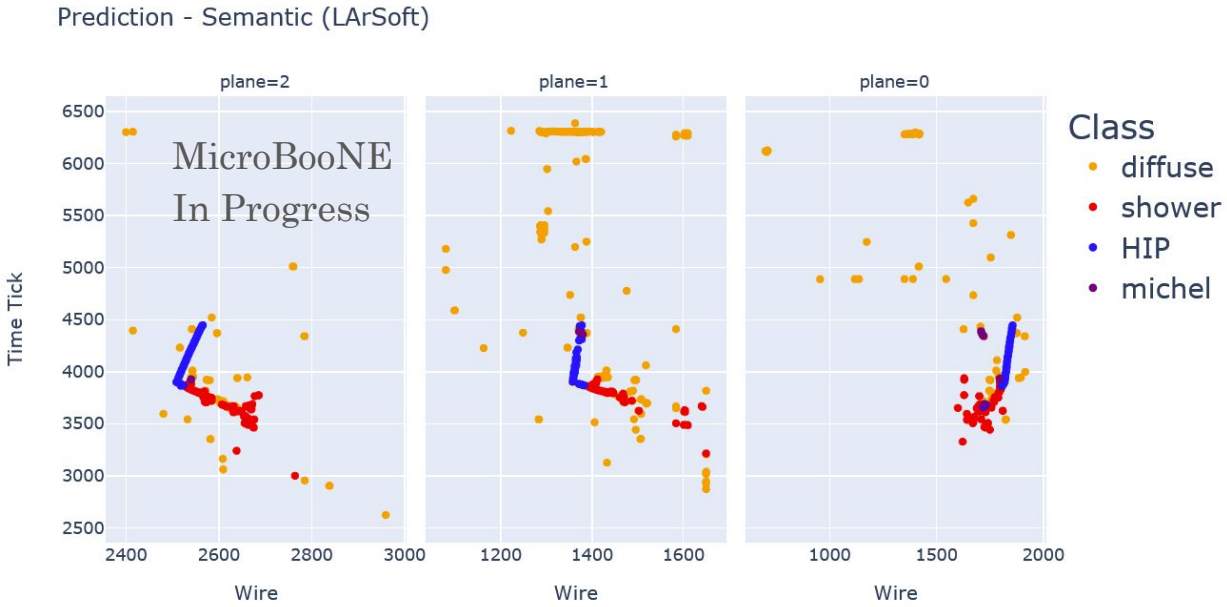


predicted to be from neutrino signal

predicted to be from cosmic background

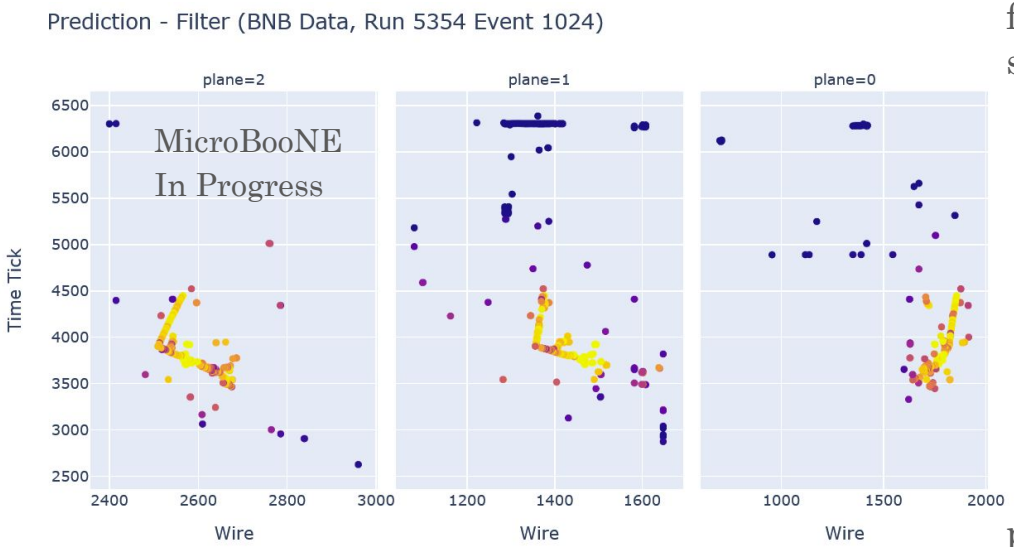
a large fraction of the hits from neutrino interaction was predicted to be from cosmics!

First Tests of NuGraph2 on MicroBooNE Data



- predictions for the particle types are correct for most hits

Tests of NuGraph2 Trained with Different Simulation

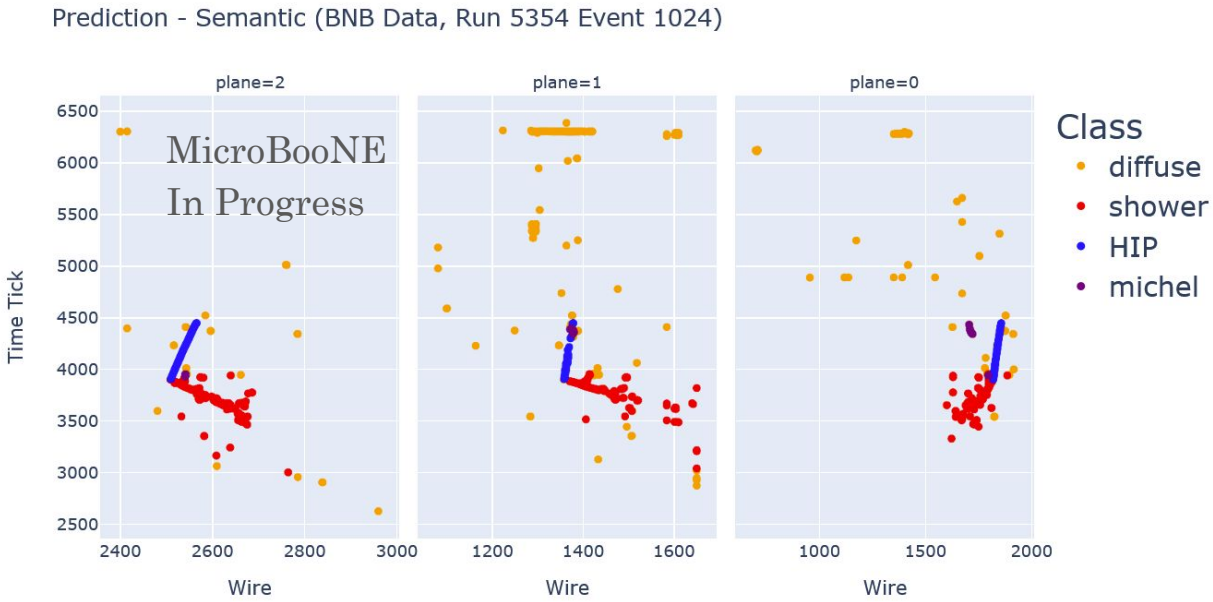


predicted to be from neutrino signal

predicted to be from cosmic background

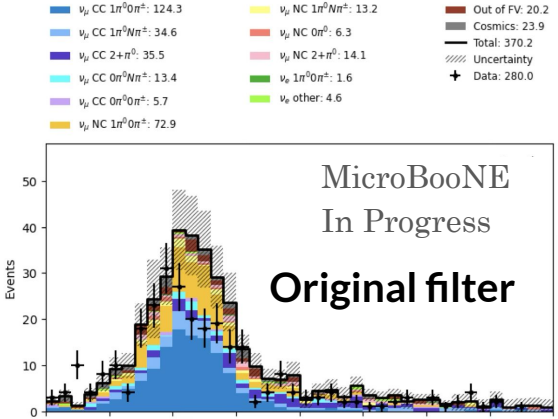
now, most hits are predicted correctly

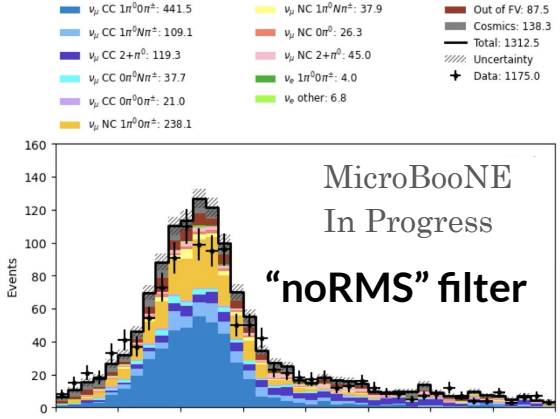
Tests of NuGraph2 Trained with Different Simulation



- predictions for the particle types are correct for most hits
- slight improvement for the start of the shower

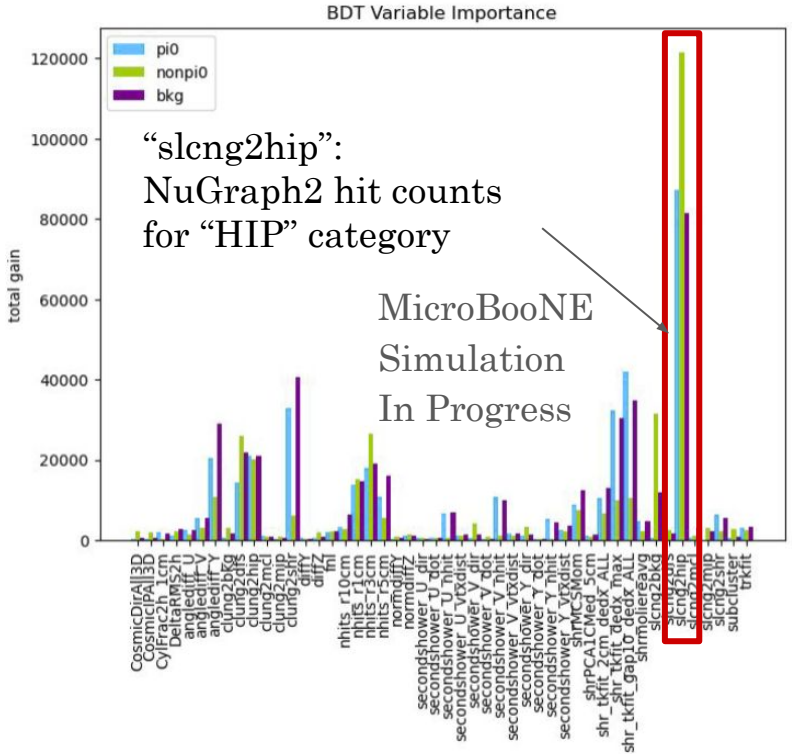
data/MC Discrepancy in π^0 Mass Plot





- original filter also showed data-mc discrepancy in π^0 mass plot
- noRMS filter also largely improved the data-mc agreement

1e0p0$\pi^{\pm}$



- Kate Pletcher's 1e0p analysis
- in-progress BDT trained for 1e0p/1eNp separation ranked NuGraph HIP variable as the most important

HNL - Updated Reco Vertex Plot