Implementation of NuGraph to $p \rightarrow K^{+}\bar{\nu}$ for Event Classification & Explorations of Nuclear Effects within DUNE

Tyler D. Stokes

On Behalf of the DUNE Collaboration

October 29th, 2024

DEEP UNDERGROUND NEUTRINO EXPERIMENT

Louisiana State University

Goals of DUNE

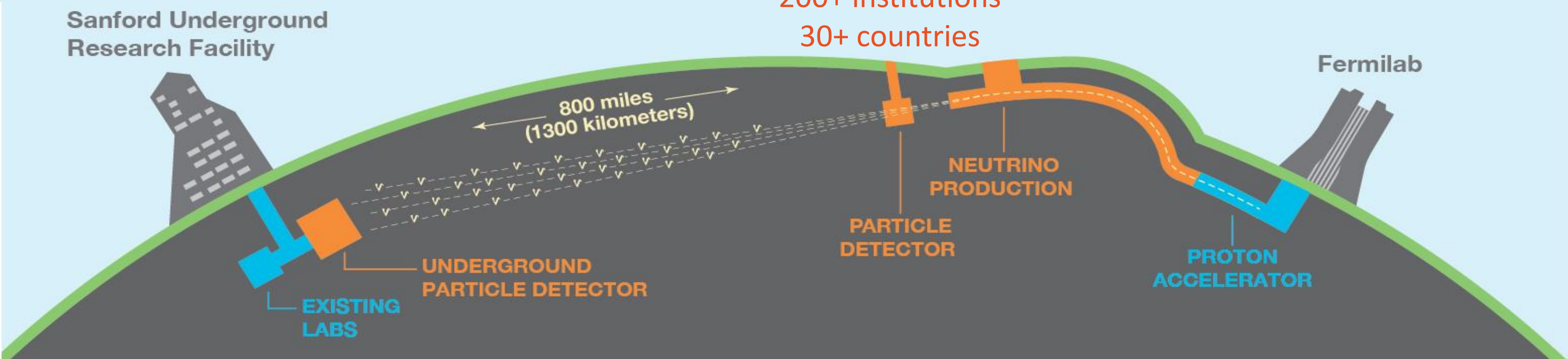
- Extract ν oscillation parameters
- Search for BSM Physics

DEEP UNDERGROUND NEUTRINO EXPERIMENT

Design of DUNE

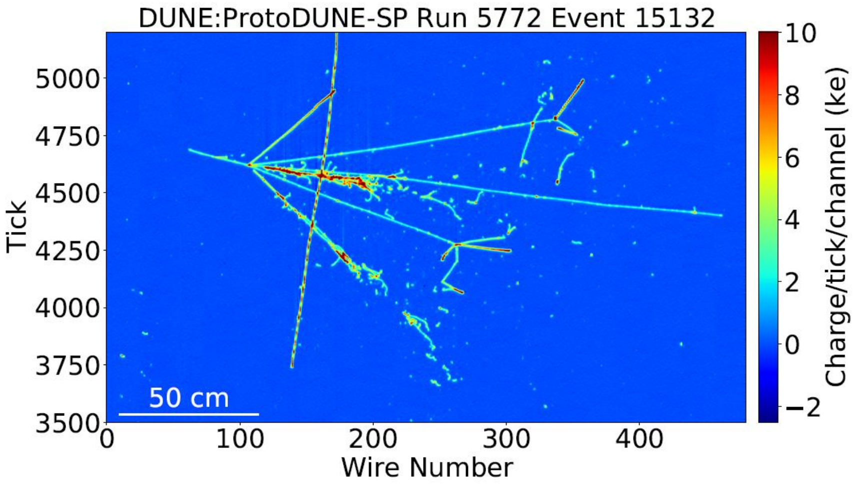
- Broad-band neutrino beam at Fermilab (1.2 MW upgradeable to >2.0 MW), 1300 km baseline
- **Near detectors**: suite of detector technologies
- **Far detectors:** Liquid Argon Time Projection Chambers (LArTPCs), total mass: **4x17kt**

1400+ collaborators
200+ institutions
30+ countries

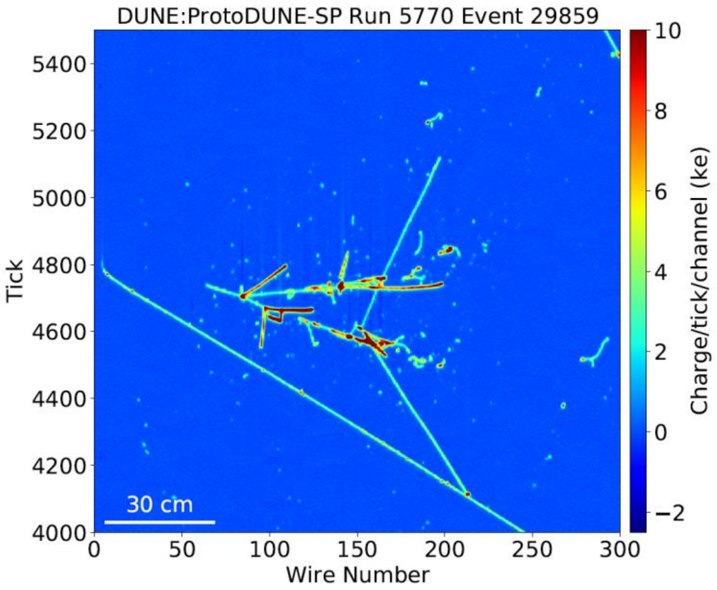


Liquid Argon Time Projection Chambers

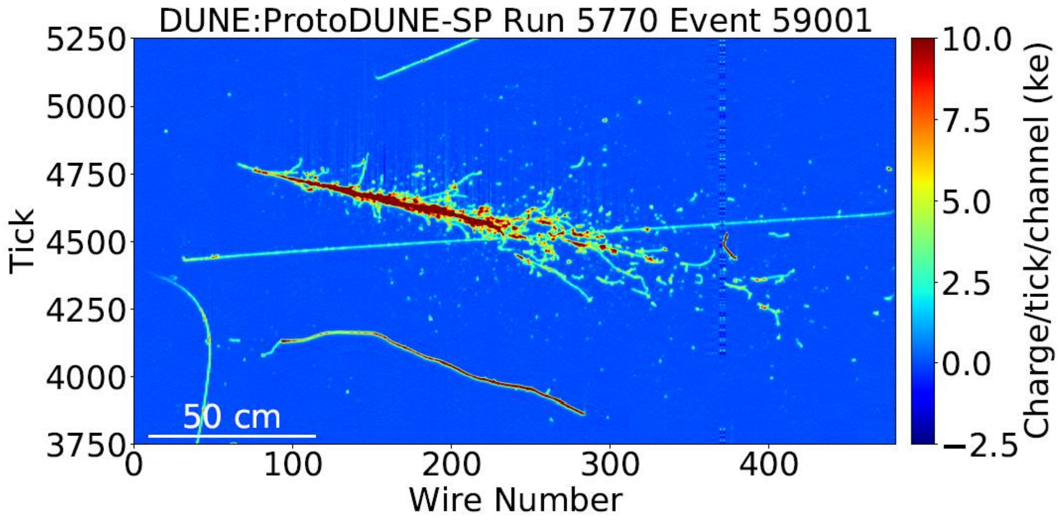
- 3D Bubble Chamber-like images
- Can classify complex topologies
- Example event displays from ProtoDUNE-SP
  - DUNE FD prototype at surface
- Can reconstruct $K^{\pm}$ from nucleon decay events



A 6 GeV/c Pion candidate



6 GeV/c Kaon candidate

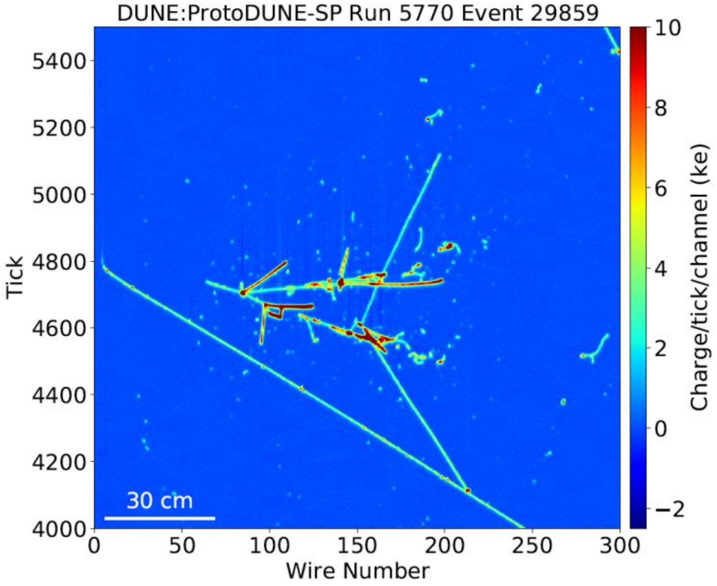


6 GeV/c electron candidate

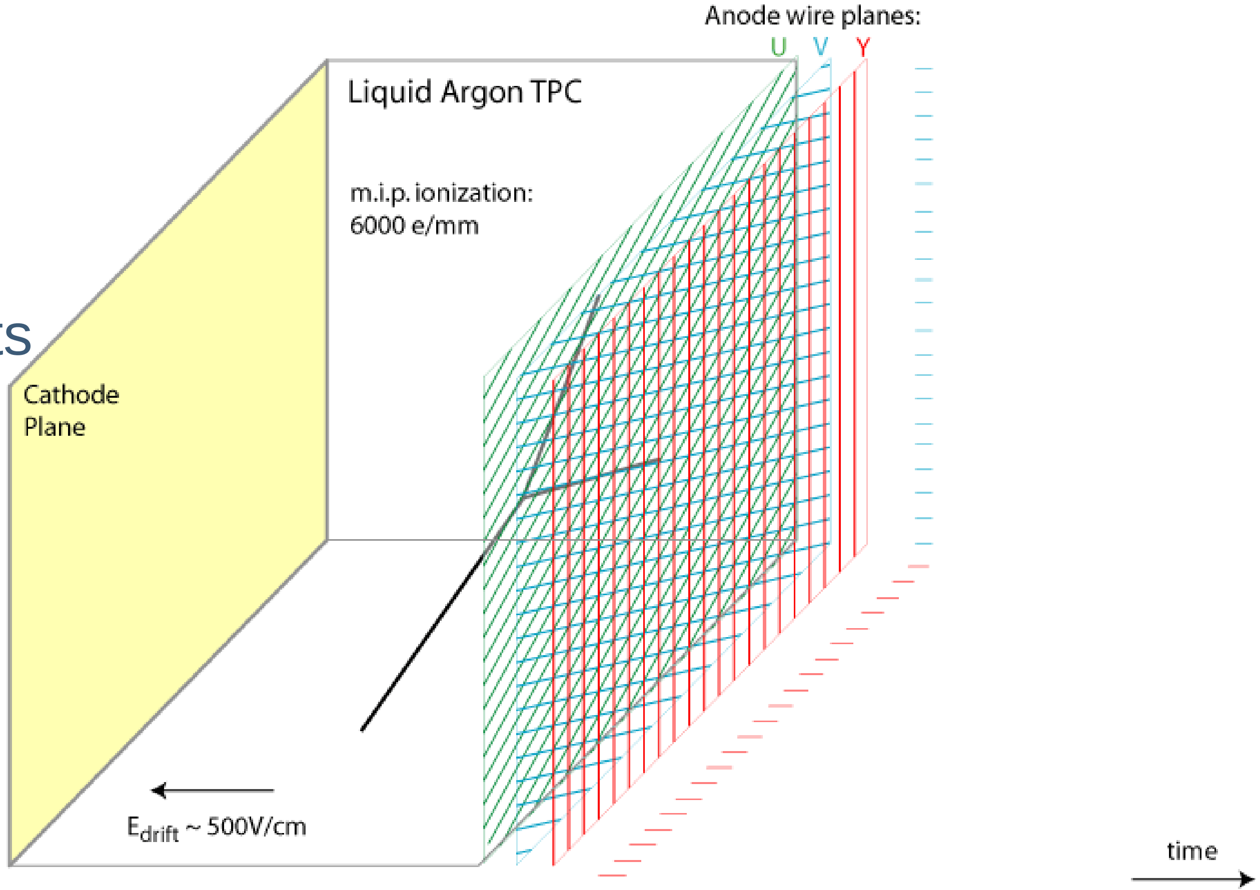
https://iopscience.iop.org/article/10.1088/1748-0221/15/12/P12004

Liquid Argon Time Projection Chambers

- 3D Bubble Chamber-like images
- Can classify complex topologies
- Example event displays from ProtoDUNE-SP
  - DUNE FD prototype at surface
- Can reconstruct $K^{\pm}$ from nucleon decay events



6 GeV/c Kaon candidate

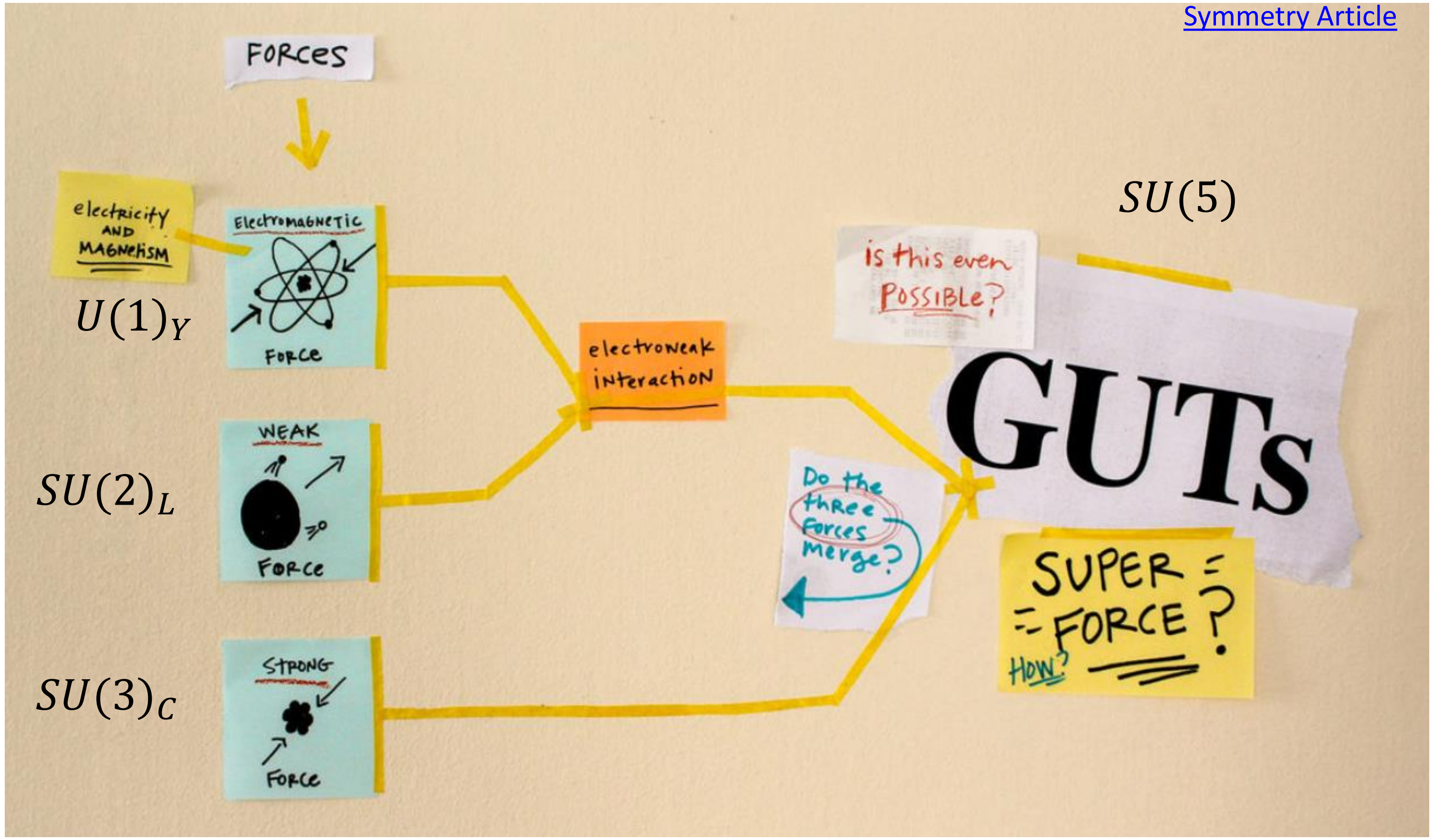


https://iopscience.iop.org/article/10.1088/1748-0221/15/12/P12004

$p \to K^+\bar{\nu}$: Baryon Number Violation?

- In the Standard Model Baryon Number conservation is not fundamental; it occurs accidentally
  - Arises as a byproduct of the SM Gauge Structure
- Grand Unified Theories (GUTs) naturally predict Baryon Number Violation

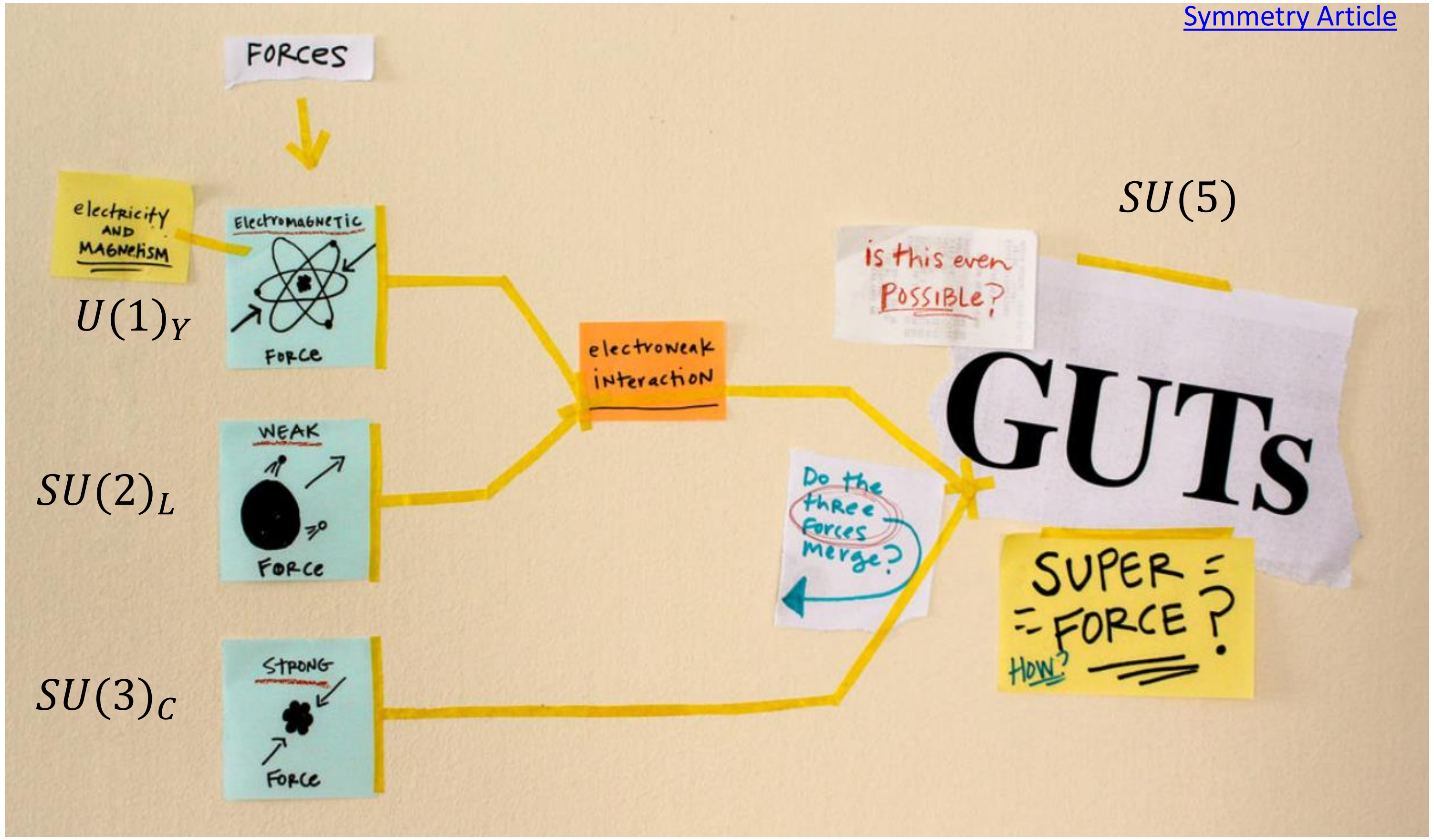
Symmetry Article



Baryon Number Violation?

- In the Standard Model Baryon Number conservation is not fundamental; it occurs accidentally
  - Arises as a byproduct of the SM Gauge Structure
- Grand Unified Theories (GUTs) naturally predict Baryon Number Violation

Symmetry Article



Unification of Forces

↓

Possibility of quark to lepton transition

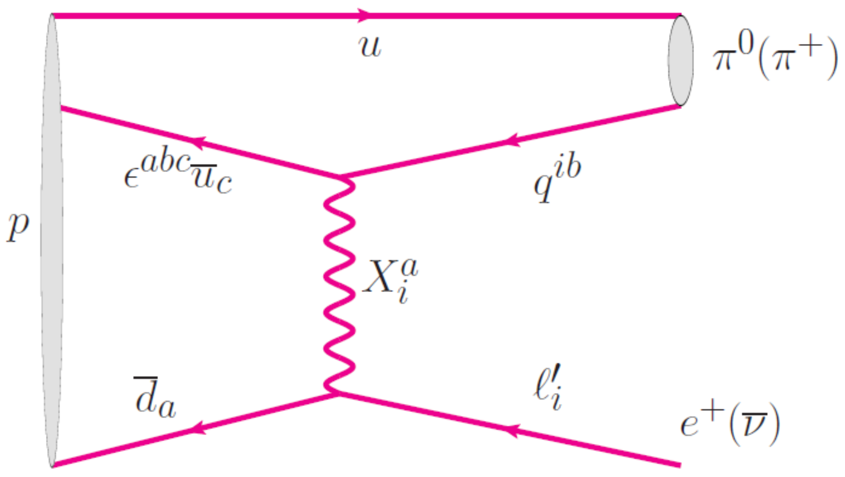
↓

Baryon Number Violation → Proton Decay

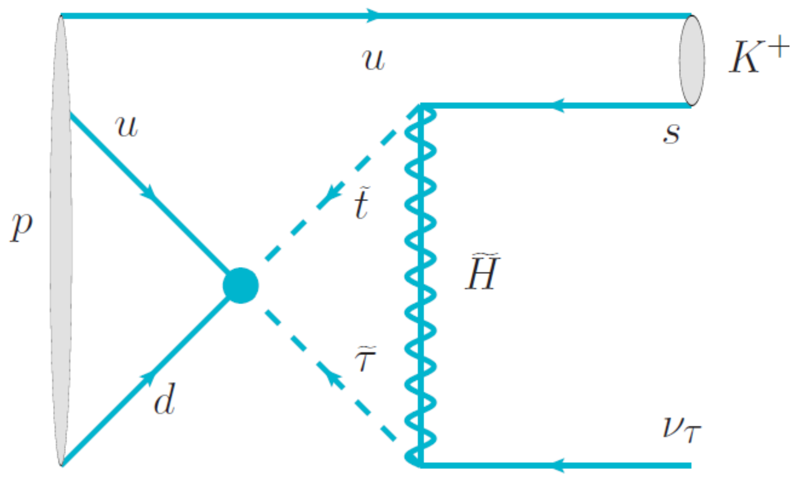
Baryon Number Violation?

- In the Standard Model Baryon Number conservation is not fundamental; it occurs accidentally
  - Arises as a byproduct of the SM Gauge Structure
- Grand Unified Theories (GUTs) naturally predict Baryon Number Violation

10.1088/1361-6471/ad1658



Non-SUSY GUT: $p \rightarrow e^+\pi^0$



SUSY GUT: $p \rightarrow K^+\bar{\nu}$

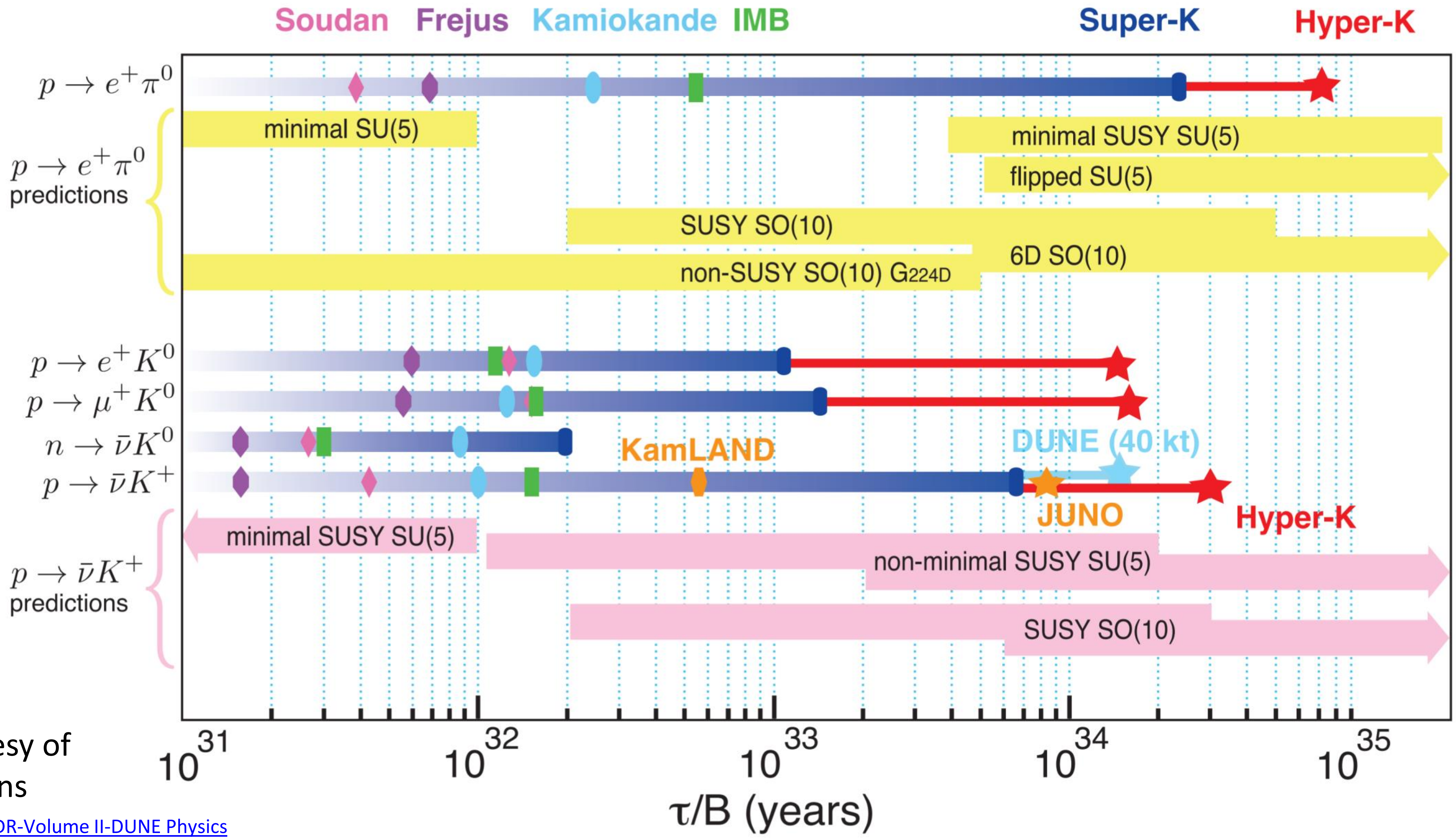
Unification of Forces

Possibility of quark to lepton transition

Baryon Number Violation → Proton Decay

- Detecting either signature would provide strong evidence for GUTs

Expected PDK Lifetimes

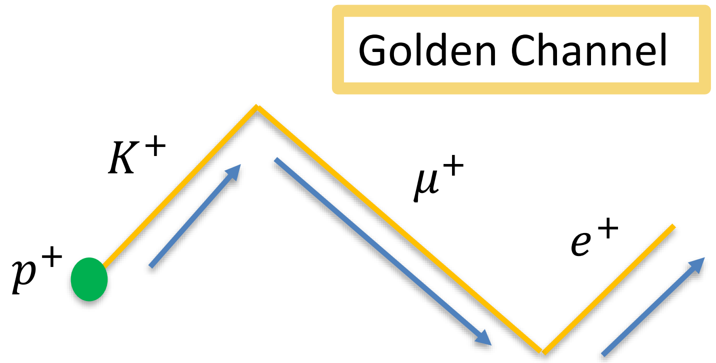


Plot Courtesy of
Ed Kearns

DUNE TDR-Volume II-DUNE Physics

Baryon Number Violation: Nucleon Decay

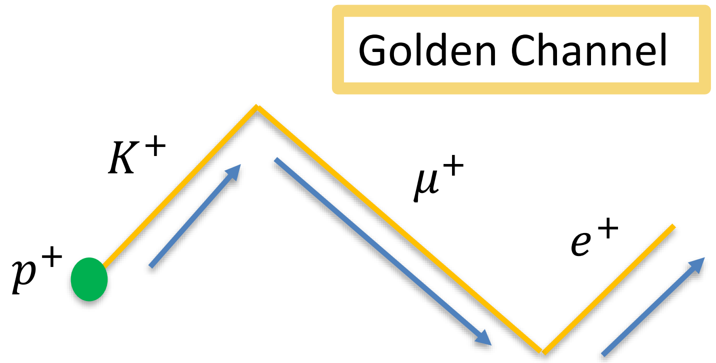
The Golden Channel: $p \rightarrow K^{+}\bar{\nu}$



- LArTPC has an advantage with charged Kaons over Water Cherenkov detectors
- Charged kaon can be fully reconstructed in LArTPC
  - Falls below Cherenkov threshold in water

Baryon Number Violation: Nucleon Decay

The Golden Channel: $p \rightarrow K^{+}\bar{\nu}$



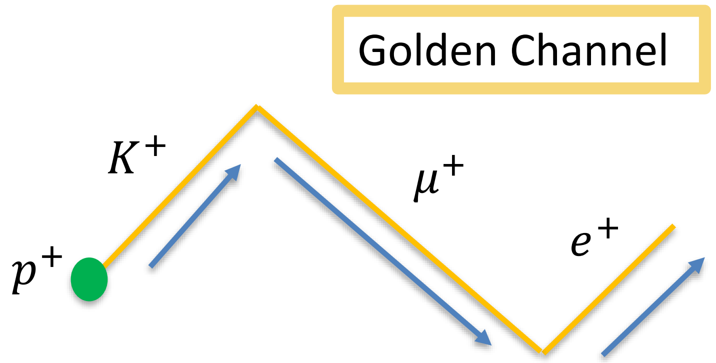
- LArTPC has an advantage with charged Kaons over Water Cherenkov detectors
- Charged kaon can be fully reconstructed in LArTPC
  - Falls below Cherenkov threshold in water

Key Features:

- Kaon decay daughters create a distinct signal
- Kaon Bragg peak near muon vertex

Baryon Number Violation: Nucleon Decay

The Golden Channel: $p \rightarrow K^{+}\bar{\nu}$



- LArTPC has an advantage with charged Kaons over Water Cherenkov detectors
- Charged kaon can be fully reconstructed in LArTPC
  - Falls below Cherenkov threshold in water

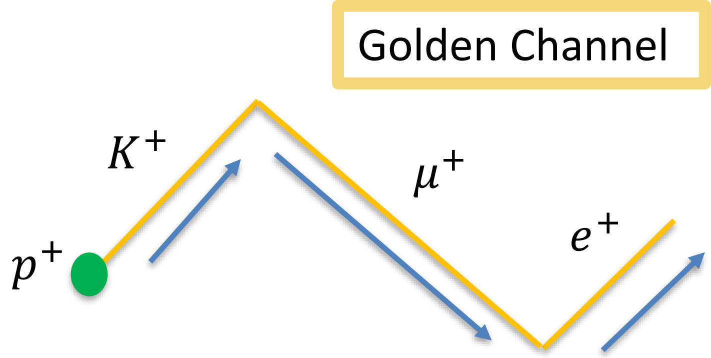
Key Features:

- Kaon decay daughters create a distinct signal
- Kaon Bragg peak near muon vertex

- **Key Difficulties:**
  - Final State Interactions (FSI) in Argon can affect decay products
    - Kaon energy loss complicates reconstruction

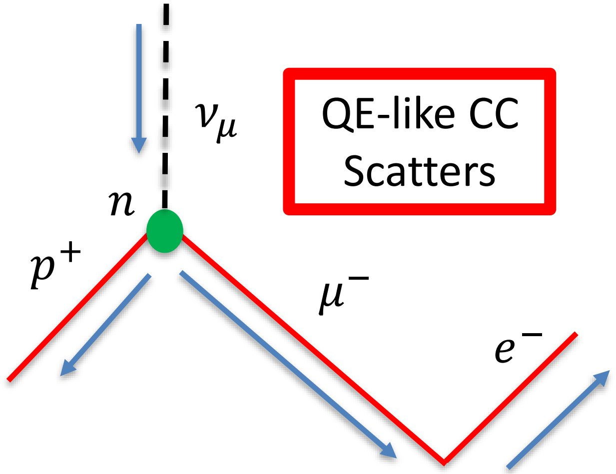
Baryon Number Violation: Nucleon Decay

The Golden Channel: $p \rightarrow K^{+}\bar{\nu}$



- LArTPC has an advantage with charged Kaons over Water Cherenkov detectors
- Charged kaon can be fully reconstructed in LArTPC
  - Falls below Cherenkov threshold in water

Background: atmospheric neutrino CC and NC interactions



Key Features:

- Kaon decay daughters create a distinct signal
- Kaon Bragg peak near muon vertex

- **Key Difficulties:**
  - Final State Interactions (FSI) in Argon can affect decay products
    - Kaon energy loss complicates reconstruction

Event Simulation

- Samples created with GENIEv3.0.6
- Proton Decay Module implemented within GENIE
  - Utilizes the proton decay mode of: $p \rightarrow K^{+}\overline{\nu}$
- Atmospheric Neutrino Sample is fully oscillated
  - Utilizes the Honda Flux at Homestake Mine during solar maximum

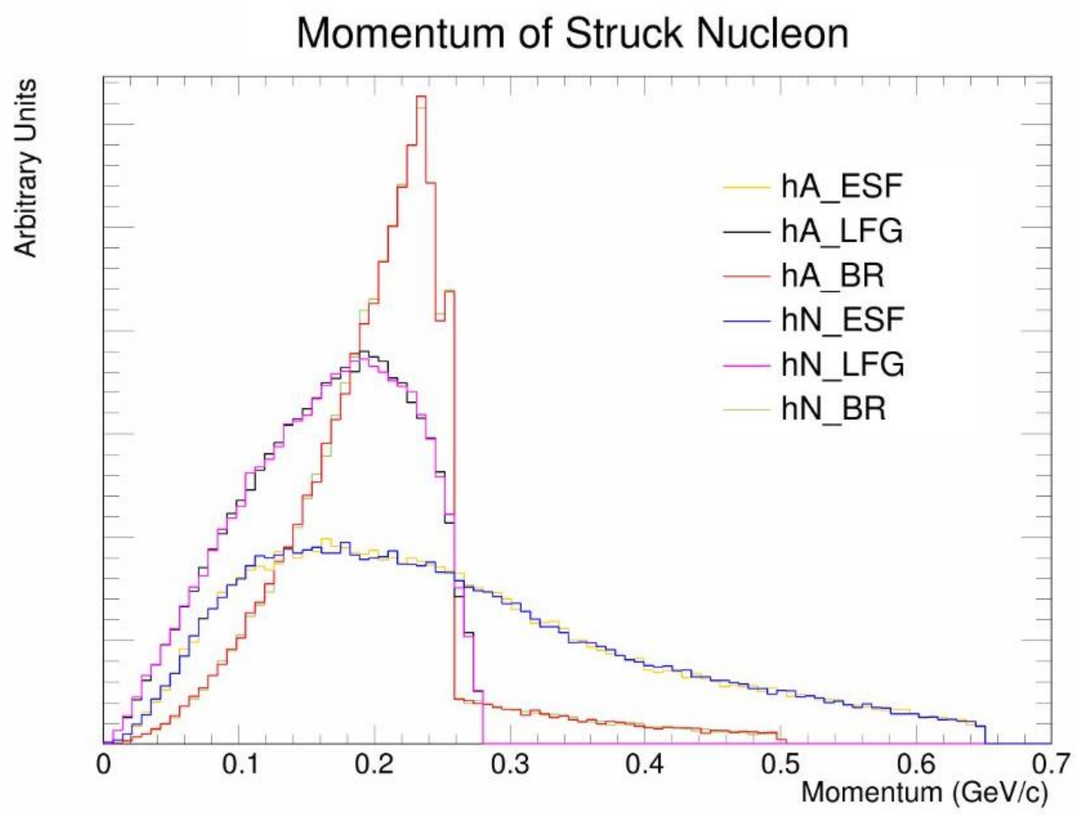
- There are three main nuclear models of Fermi motion:
  1. Bodek-Ritchie (relativistic) *nonlocal* Fermi gas
  2. *Local* (nonrelativistic) Fermi Gas
  3. Effective Spectral Function (*nonlocal*)
- There are two main intranuclear cascades available:
  1. hA2018 (single effective interaction)
  2. hN2018 (full intranuclear cascade model)

1. hA_BR
2. hA_LFG
3. hA_ESF

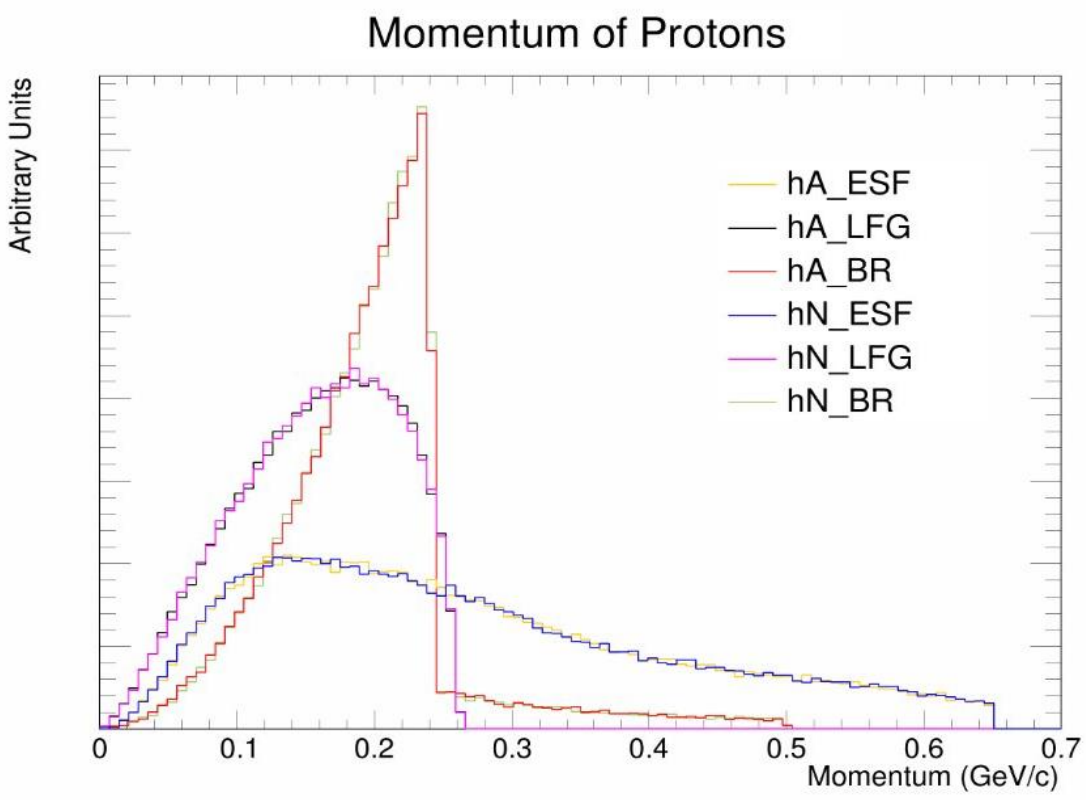
4. hN_BR
5. hN_LFG
6. hN_ESF

Initial Nucleon Momentum Distributions

- Initial State Preparation for Atmospherics and $p \rightarrow K^{+}\bar{\nu}$
- Fermi Momentum of nucleons within GENIE v3.0.6



Momentum of nucleons struck by atmospheric neutrinos



Momentum of protons decaying

Final State Interactions

The two main intranuclear cascades:

1. hA2018 (single effective interaction)
   - No step-by-step interaction modeling
   - Uses hadron+nucleus data for final state prediction

2. hN2018 (full intranuclear cascade model)
   - Models stochastic interactions within the nucleus
   - Simulates the full sequence of hadronic interactions

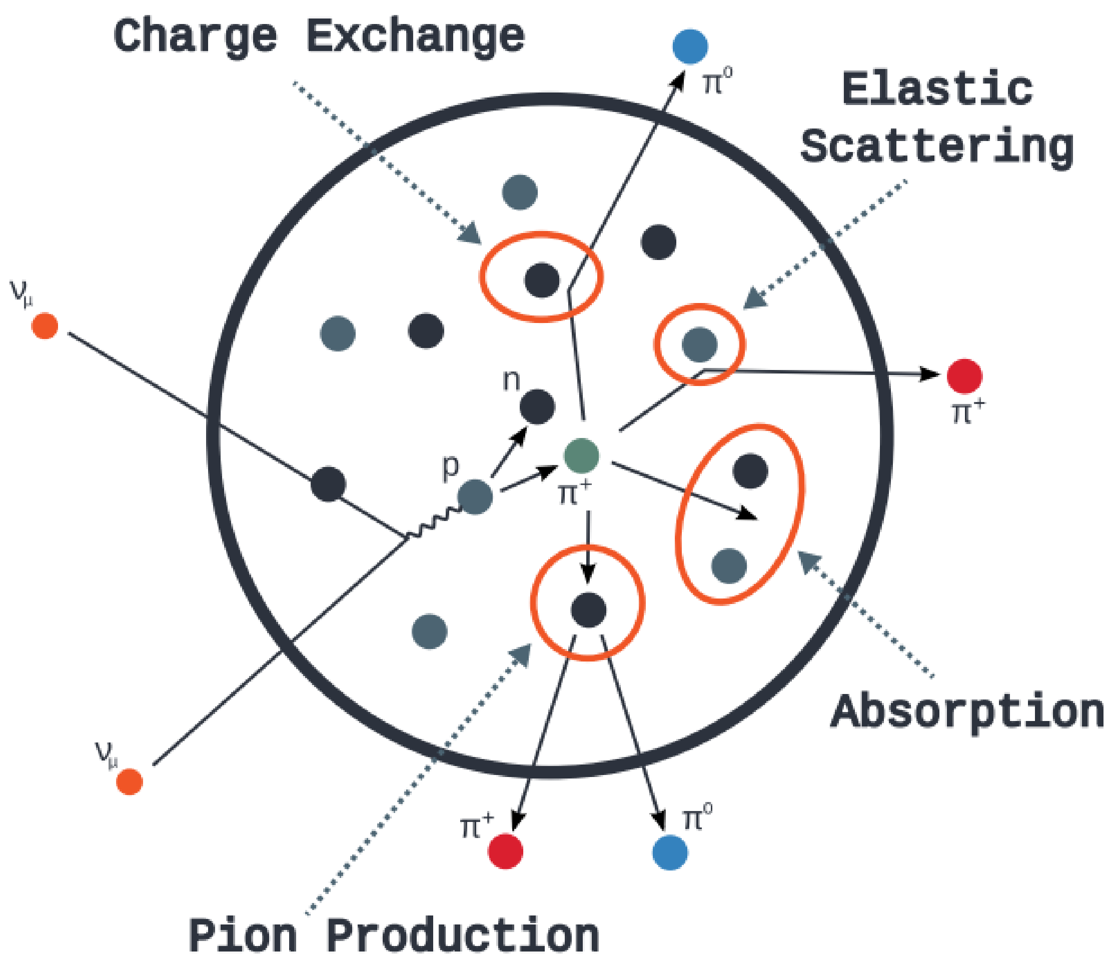
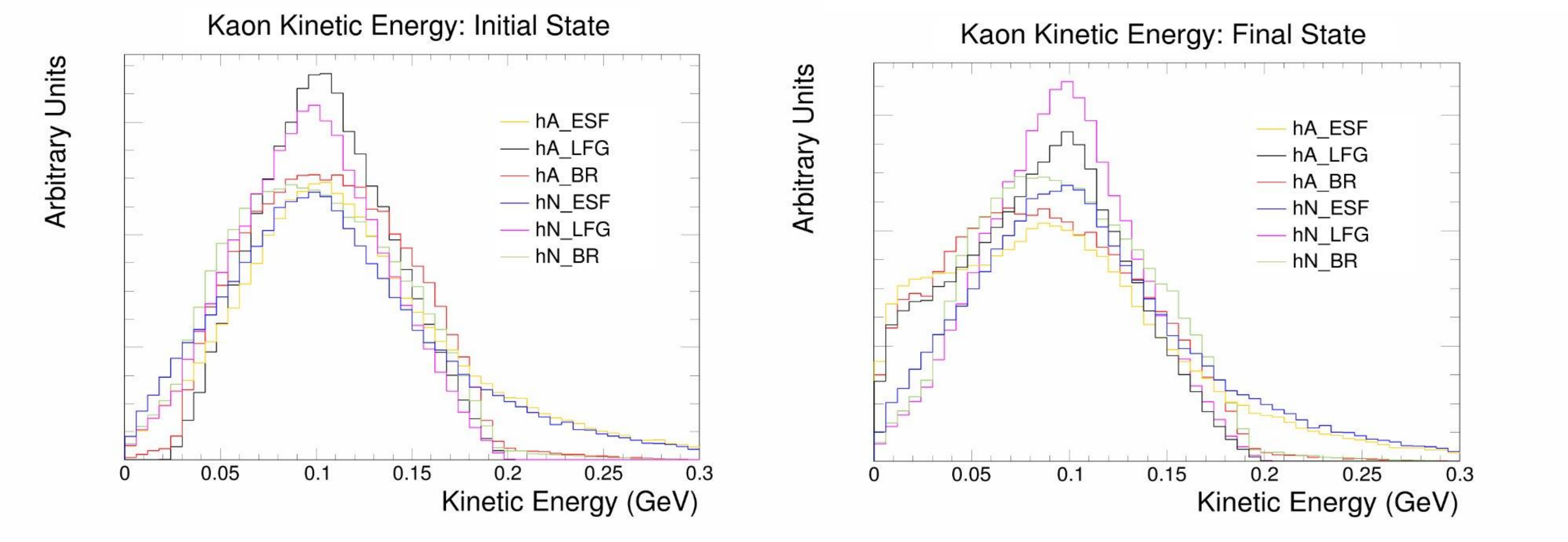


Figure from: Tomasz Golan
https://indico.fnal.gov/event/15286/contributions/30851/attachments/19320/24158/nustec_mc_02.pdf

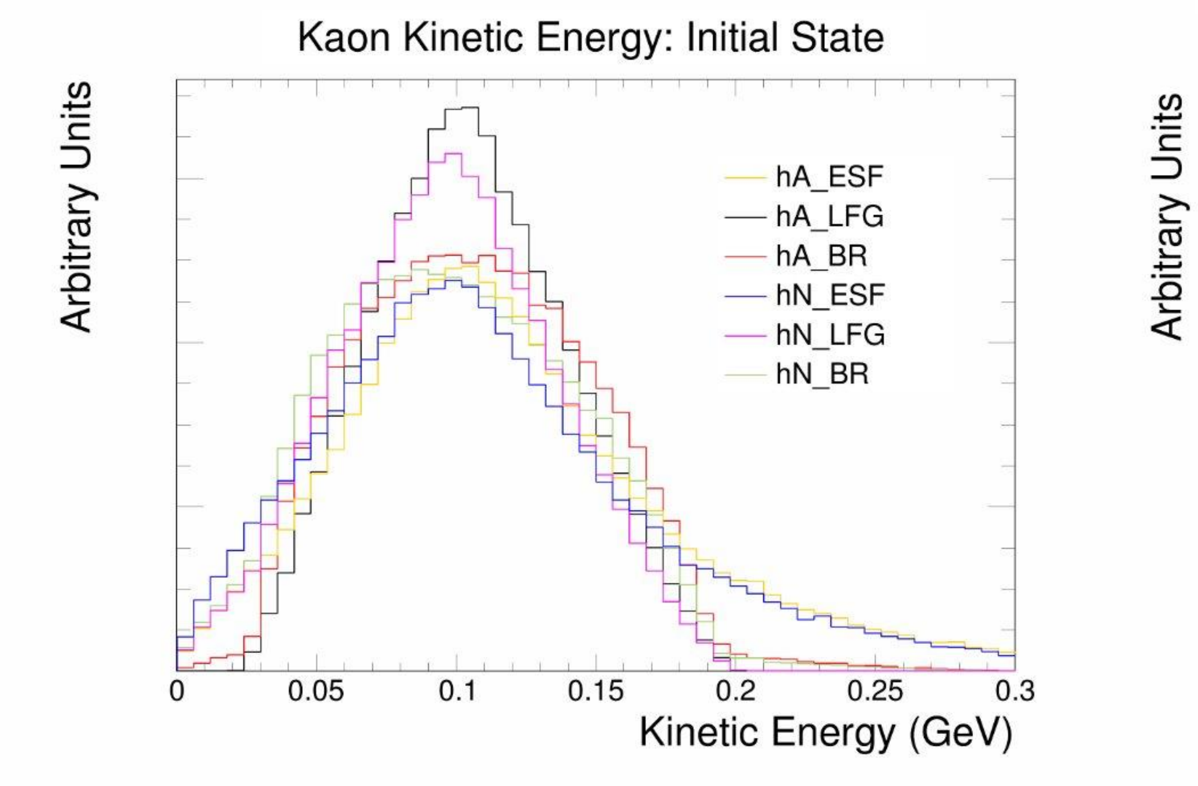
Nuclear Modeling Effects on Kaon Energy

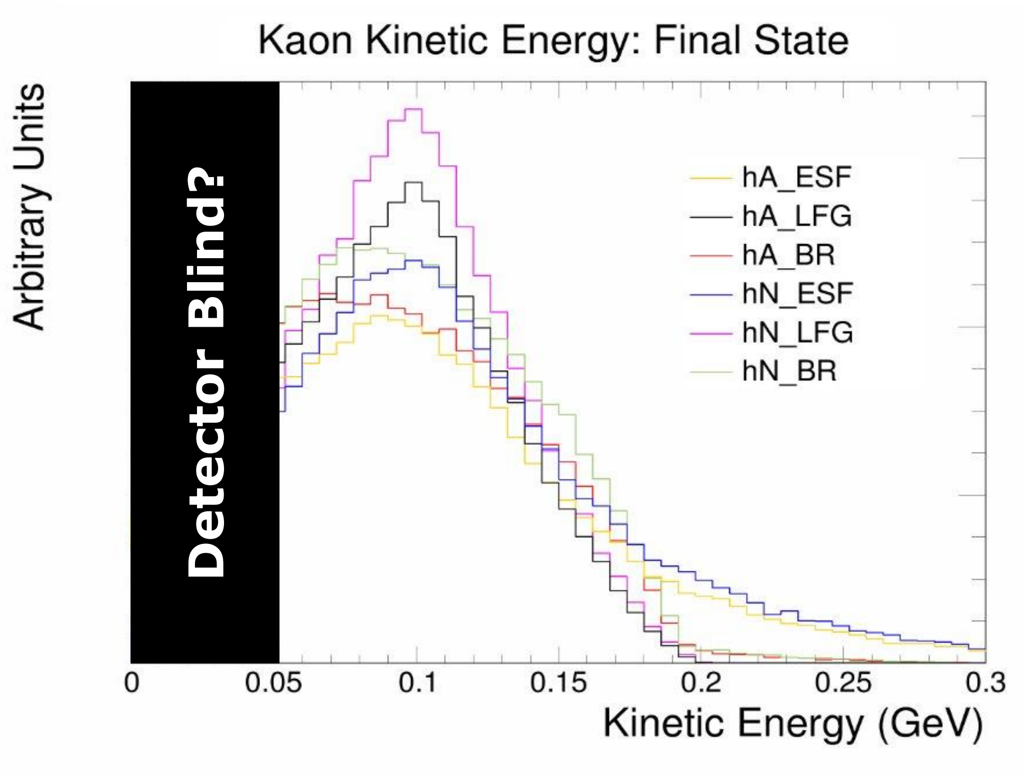
- K^+ generated directly from decaying p in nucleus
  - Initial momentum from fermi motion and rest mass
  - FSI effects of hA and hN Intranuke 2018 show differences
    - hA has a distinct shift towards lower energies



Nuclear Modeling Effects on Kaon Energy

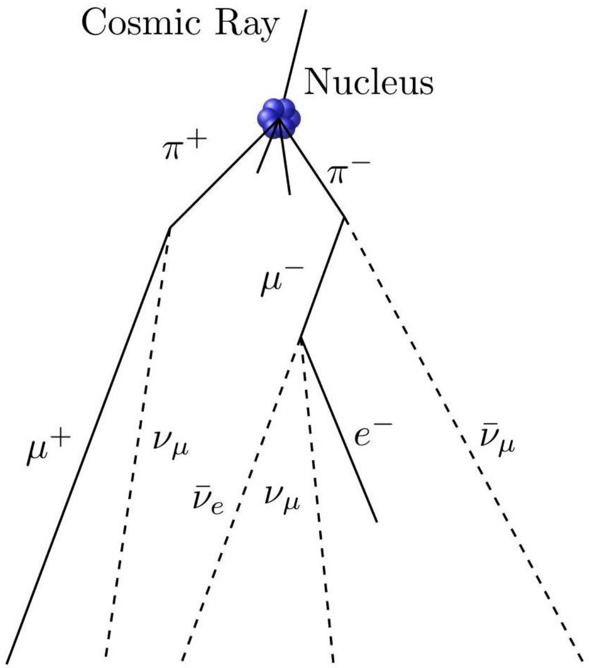
- K^+ generated directly from decaying p in nucleus
  - Initial momentum from fermi motion and rest mass
  - FSI effects of hA and hN Intranuke 2018 show differences
    - hA has a distinct shift towards lower energies

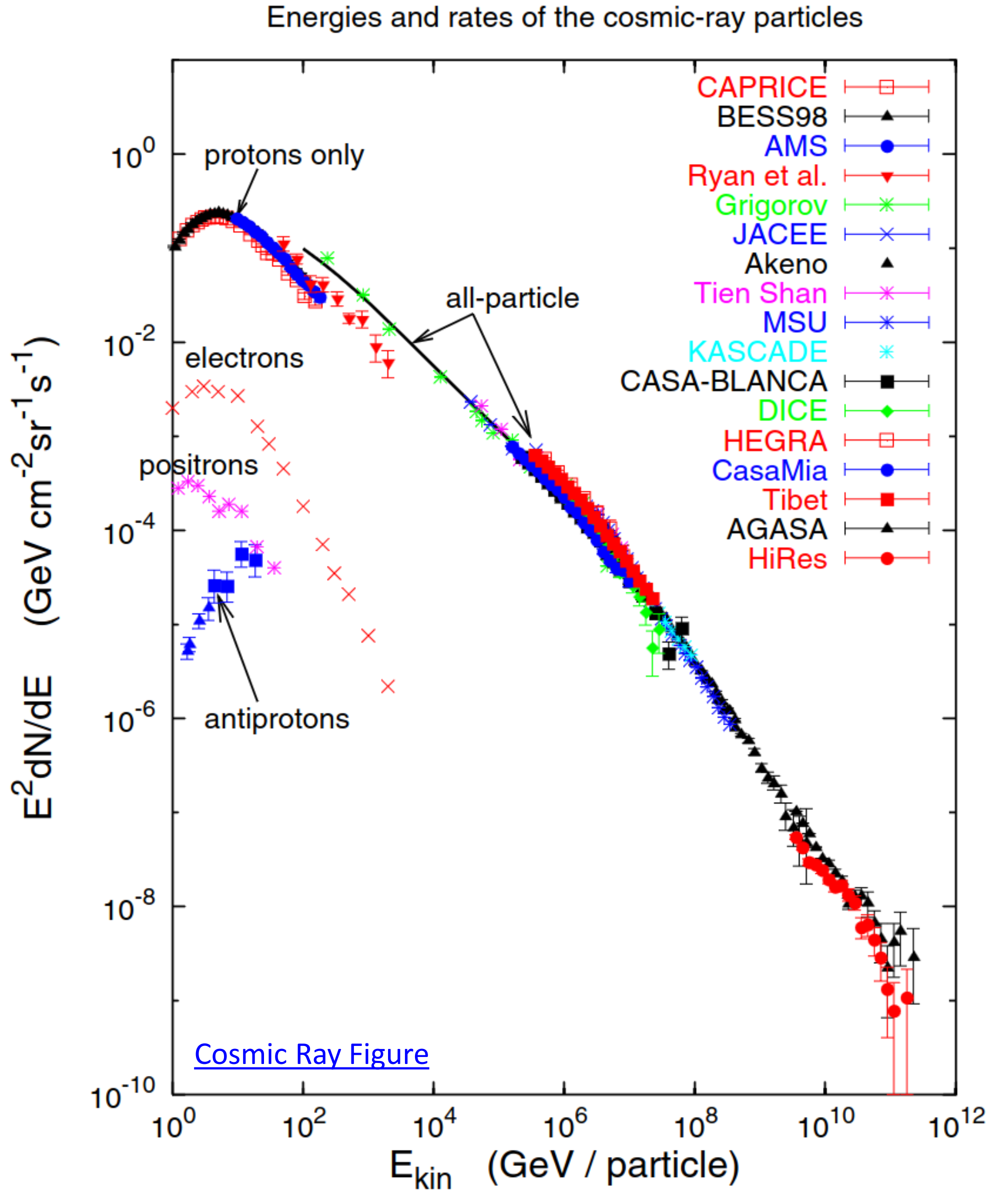




Atmospheric Neutrinos at Homestake Mine

- Atmospheric neutrinos are the main background for proton decay searches in DUNE
  - Generated when cosmic rays hit nuclei in upper atmosphere
  - Can produce pions which will decay into neutrinos
  - Hard to distinguish from proton decay signals
  - Being in a mine reduces cosmic ray noise, but not atmospheric neutrinos!

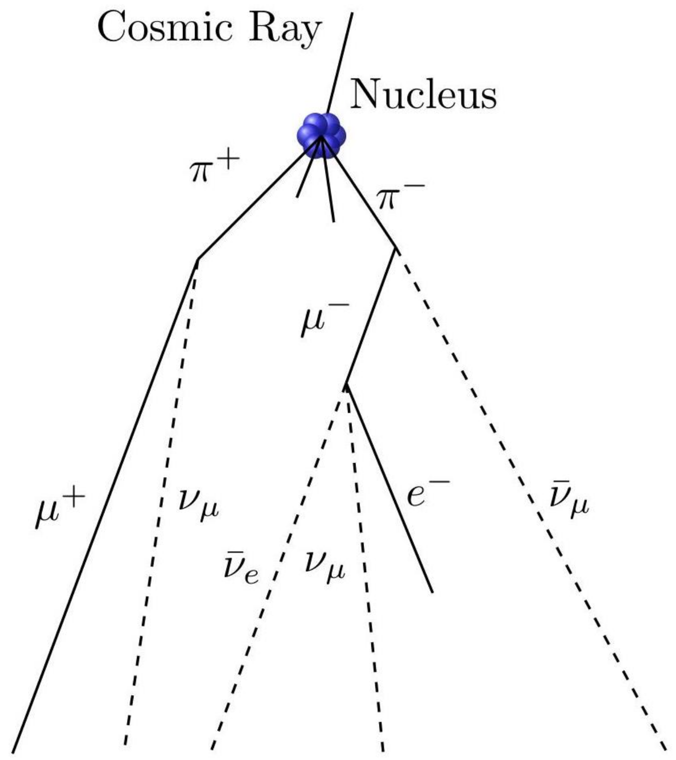




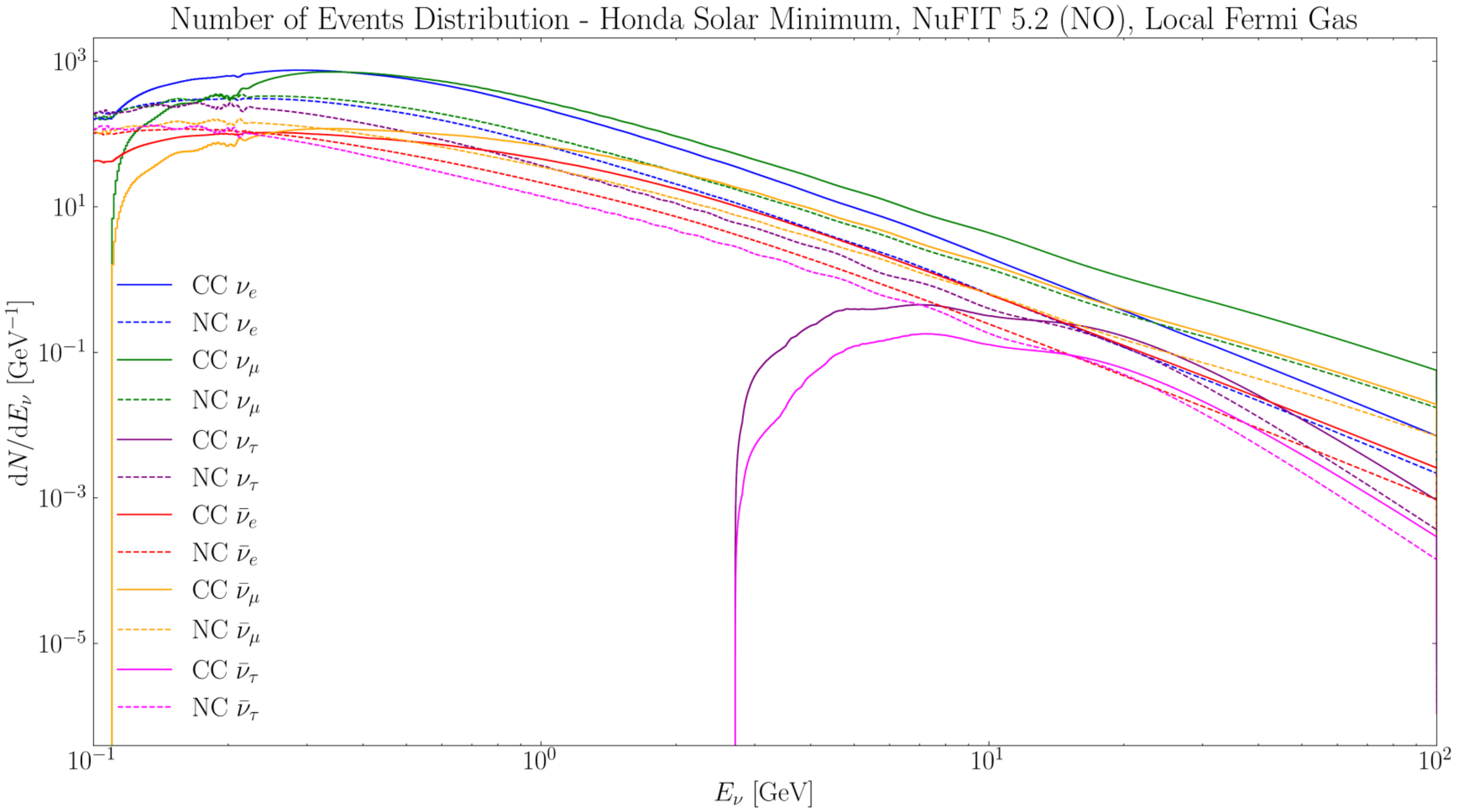
Credit: M. Oliveira-Ismerio

Atmospheric Neutrinos at Homestake Mine

- Expected $\boldsymbol{\nu}_{\mathbf{atm}}$ count rates are a key input for the proton decay background



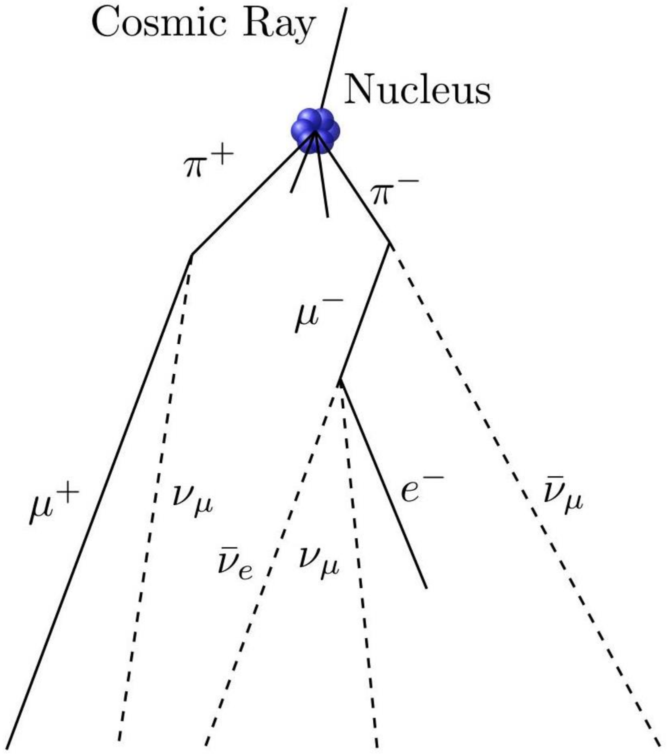
Credit: M. Oliveira-Ismerio

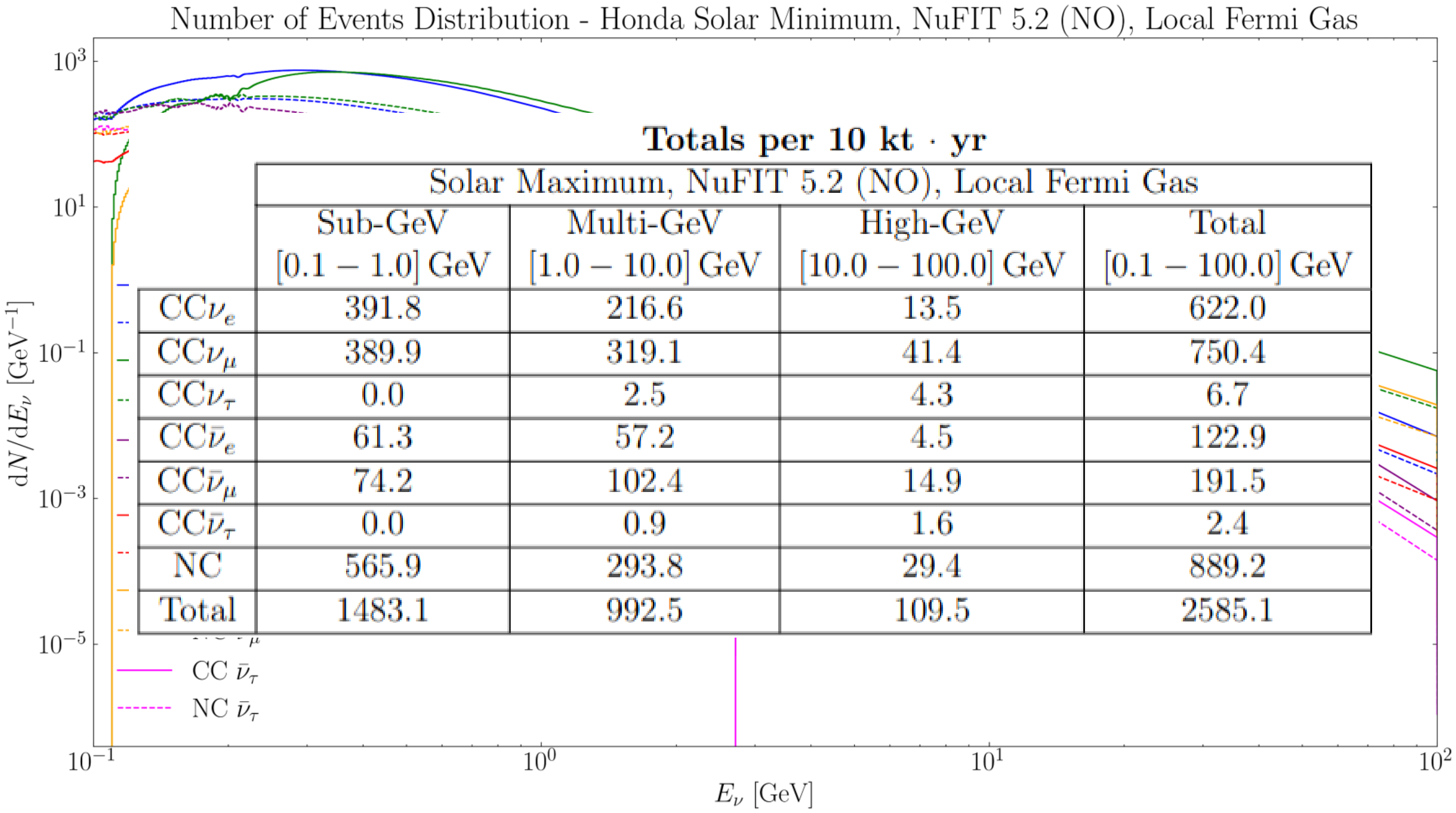


Atmospheric Neutrinos at Homestake Mine

- Expected $\boldsymbol{\nu_{\mathrm{atm}}}$ count rates are a key input for the proton decay background
  - Can be obtained via integration!

$$\phi(E) = 10^{\log_{10}\phi(E_1) + \frac{\log_{10}\phi(E_2) - \log_{10}\phi(E_1)}{\log_{10}E_2 - \log_{10}E_1}(\log_{10}E - \log_{10}E_1)}$$





Totals per 10 kt · yr

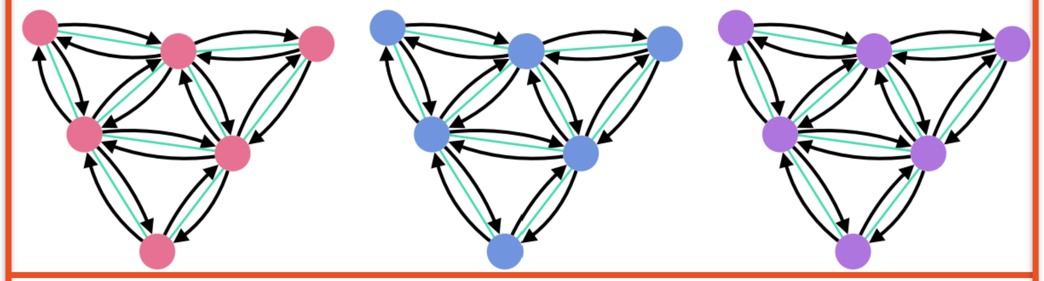
	Solar Maximum, NuFIT 5.2 (NO), Local Fermi Gas			
	Sub-GeV [0.1 – 1.0] GeV	Multi-GeV [1.0 – 10.0] GeV	High-GeV [10.0 – 100.0] GeV	Total [0.1 – 100.0] GeV
CCν_e	391.8	216.6	13.5	622.0
CCν_μ	389.9	319.1	41.4	750.4
CCν_τ	0.0	2.5	4.3	6.7
CC$\bar{\nu}_e$	61.3	57.2	4.5	122.9
CC$\bar{\nu}_\mu$	74.2	102.4	14.9	191.5
CC$\bar{\nu}_\tau$	0.0	0.9	1.6	2.4
NC	565.9	293.8	29.4	889.2
Total	1483.1	992.5	109.5	2585.1

Credit: M. Oliveira-Ismerio

NuGraph: A Graphical Neural Network

- A state-of-the-art graph neural network for semantically labelling detector hits in neutrino physics experiments
- Works natively on detector hits without modifying structure

3D graph nodes/convolutions perform message-passing independently in each detector view



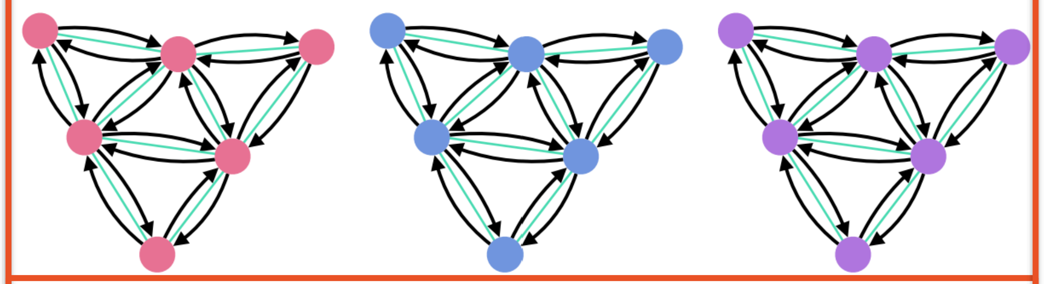
3D graph nodes/convolutions add additional 3D steps to the standard message passing loop

V Hewes CHEP 2023

NuGraph: A Graphical Neural Network

- A state-of-the-art graph neural network for semantically labelling detector hits in neutrino physics experiments
- Works natively on detector hits without modifying structure
- Utilizes a nexus convolution block
  - Combines information from all planes and injects it back into each plan at specific intervals

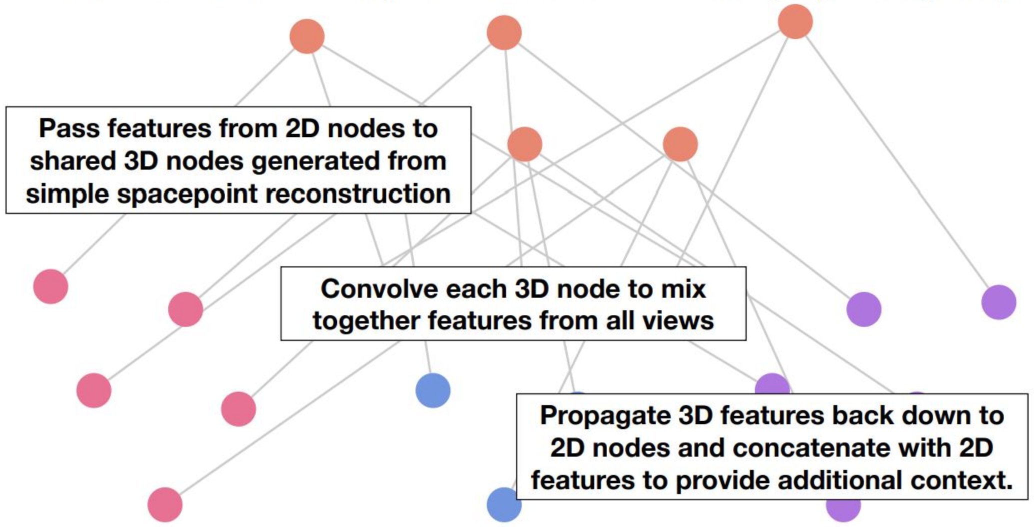
3D graph nodes/convolutions perform message-passing independently in each detector view



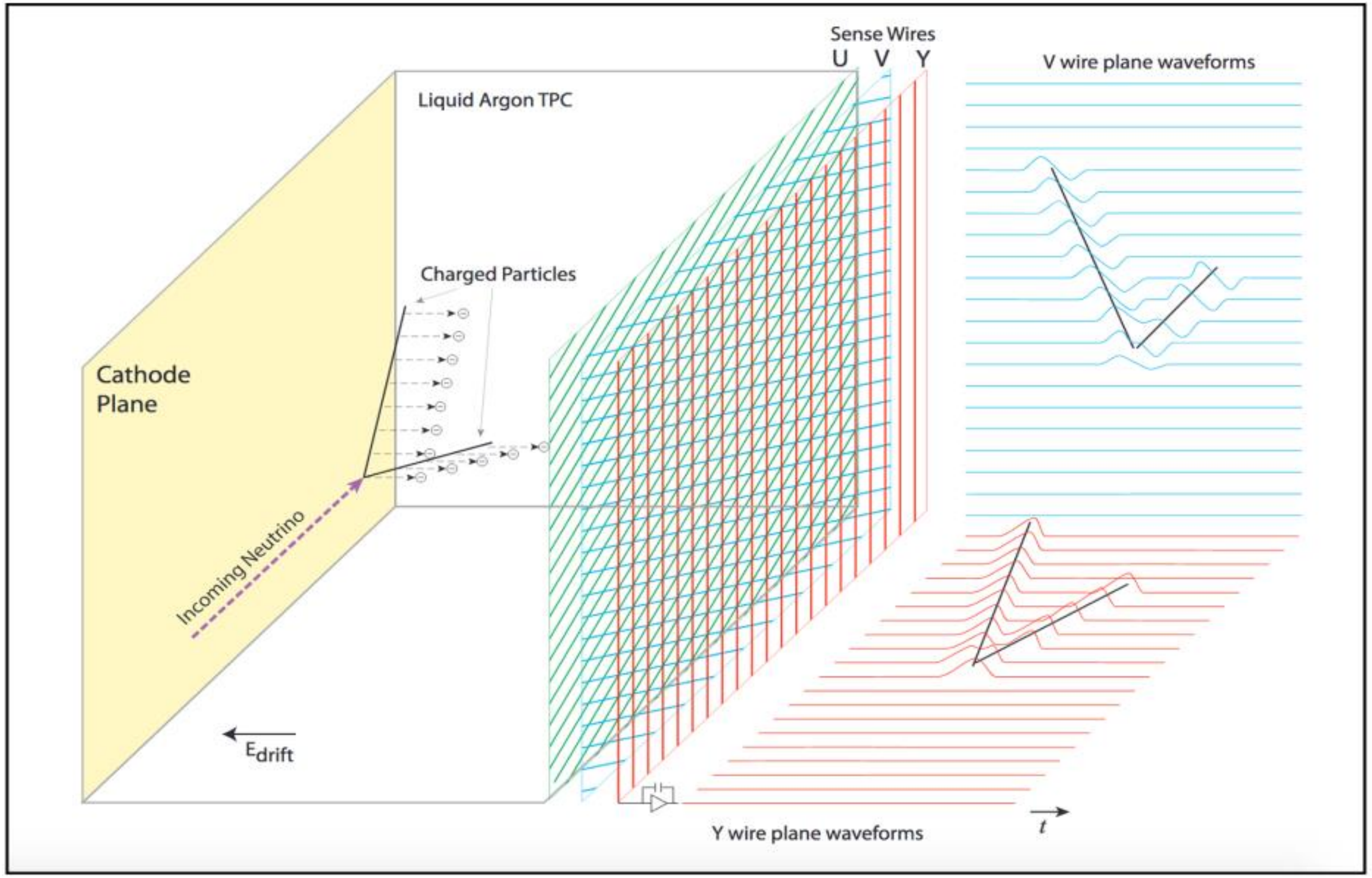
3D graph nodes/convolutions add additional 3D steps to the standard message passing loop

[V Hewes CHEP 2023]

3D Nexus Convolutions
- Add additional 3D step to the standard message-passing loop

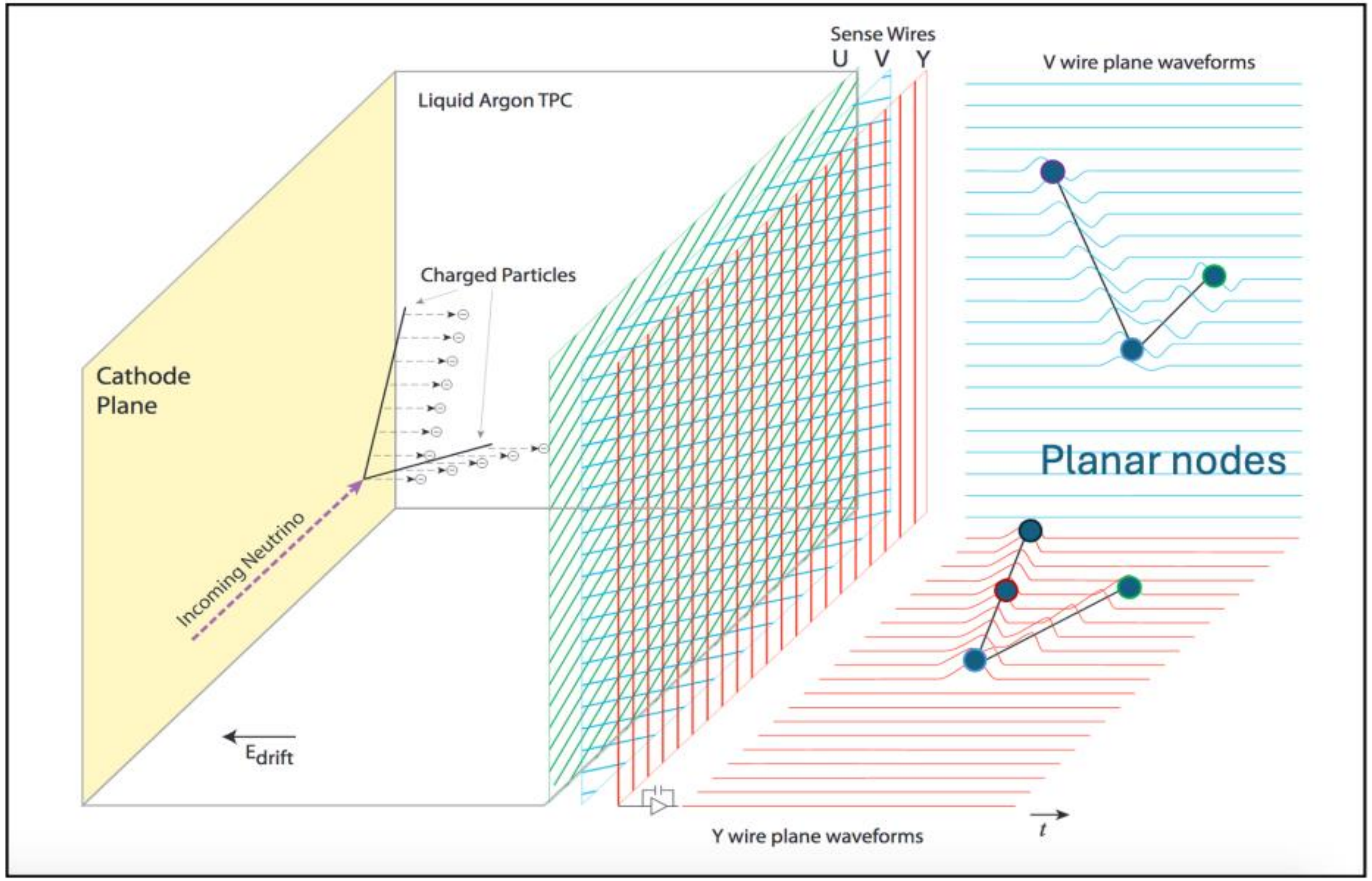


NuGraph: Operation



[ExaTrkx Phase II Kickoff: Cerati] [PhysRevD.110.032008] [Neutrino ML 2024: Cerati]

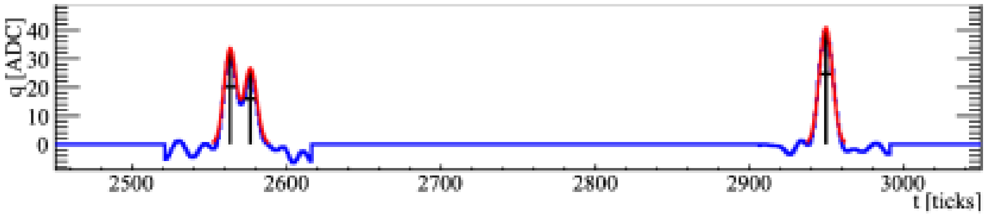
NuGraph: Operation



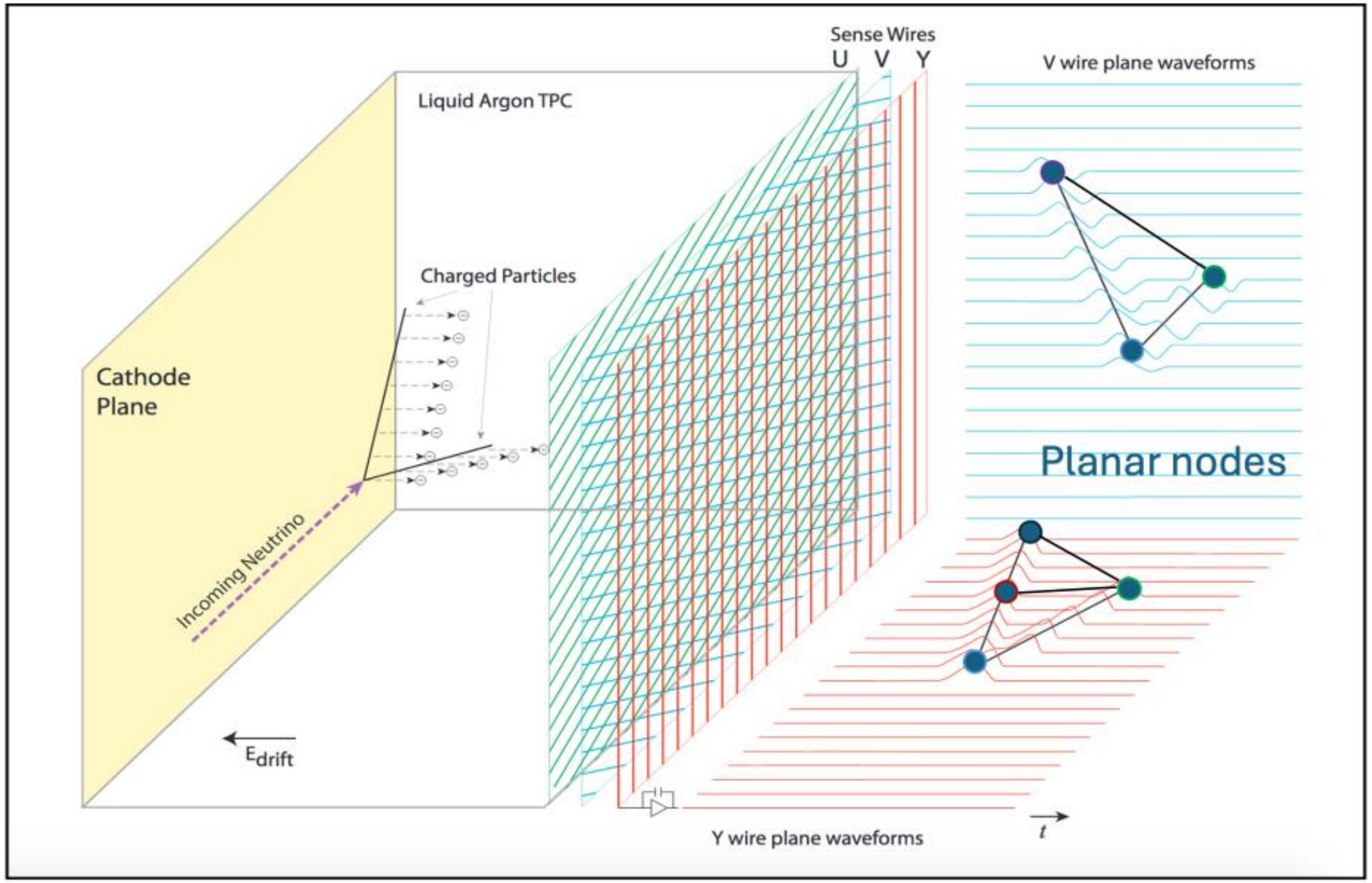
ExaTrkx Phase II Kickoff: Cerati PhysRevD.110.032008 Neutrino ML 2024: Cerati

1. Main input to the GNN are Hits
   - hits are Gaussian fits to waveforms
   - features: wire, peak time, integral, RMS

JINST 17 (2022) 01, P01026



NuGraph: Operation



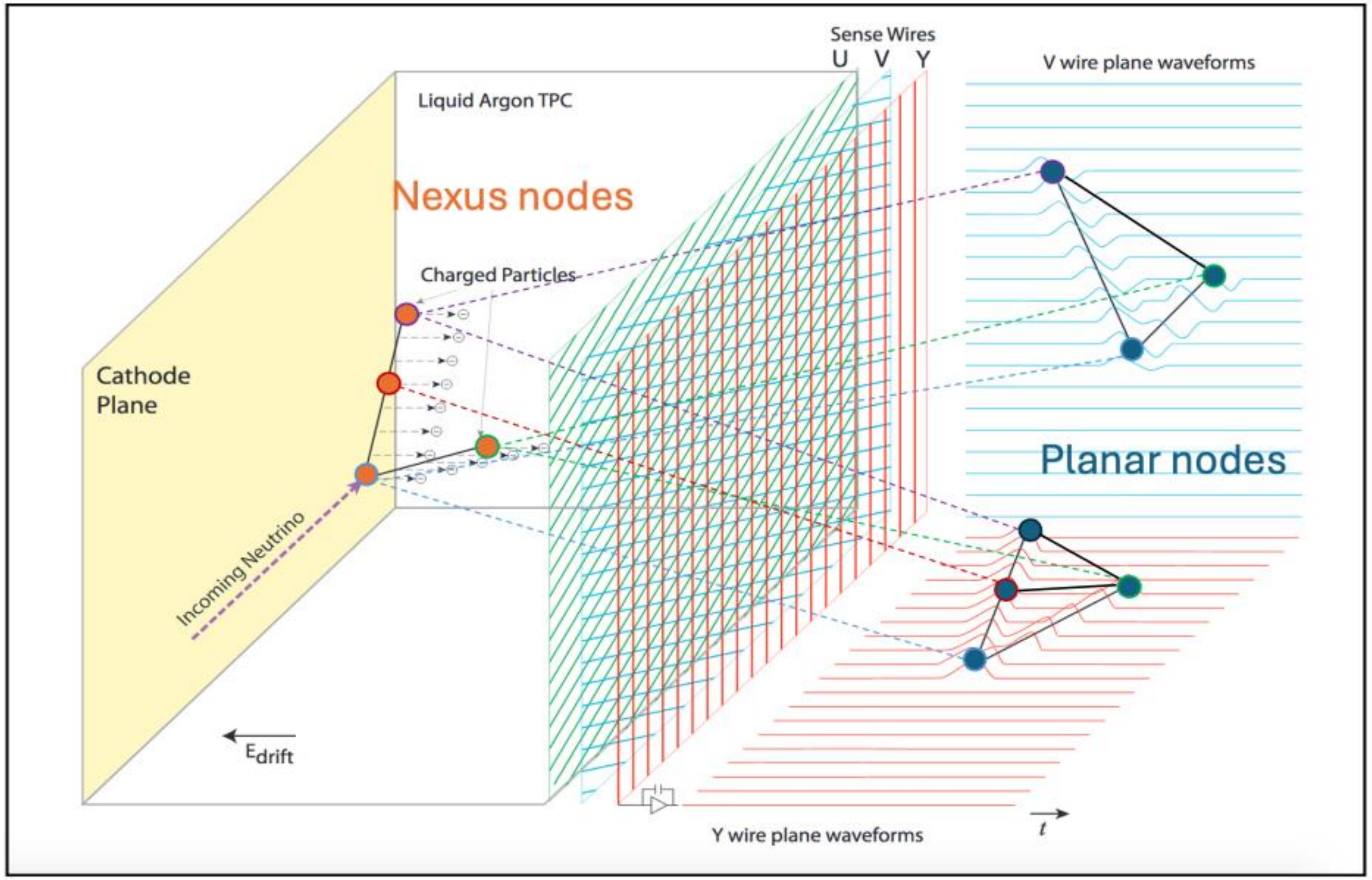
[ExaTrkx Phase II Kickoff: Cerati](#) [PhysRevD.110.032008](#) [Neutrino ML 2024: Cerati](#)

- Main input to the GNN are Hits
  - hits are Gaussian fits to waveforms
  - features: wire, peak time, integral, RMS
- Within each plane hits are connected in a graph using Delaunay Triangulation
  - Fully connected graph
  - Long and short connections
  - Connect across unresponsive wire regions

2023 J. Phys.: Conf. Ser. 2438 012091

Edge Type	Data Type	Labeling Scheme	Accuracy
Delaunay	**2D**	**Simple**	**86.24%**
Window	2D	Simple	76.9%
kNN	2D	Simple	81.14%
Radius	2D	Simple	78.58%

NuGraph: Operation



ExaTrkx Phase II Kickoff: Cerati PhysRevD.110.032008 Neutrino ML 2024: Cerati

- Main input to the GNN are Hits
  - hits are Gaussian fits to waveforms
  - features: wire, peak time, integral, RMS
- Within each plane hits are connected in a graph using Delaunay Triangulation
  - Fully connected graph
  - Long and short connections
  - Connect across unresponsive wire regions
- Hit associations with 3D SpacePoints create “nexus” connections across graphs in each plane
  - SPs are not connected among themselves
  - No input features for SPs

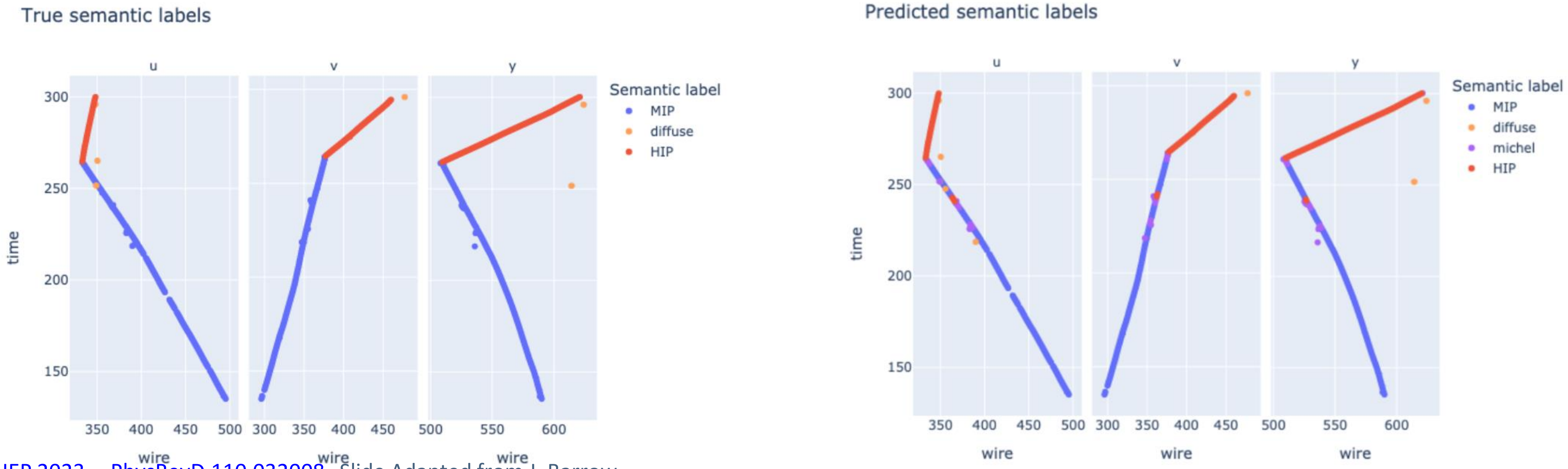
Why NuGraph?

Why NuGraph? MicroBooNE Open Data Example

Semantic Labeling Performance in: Example ν_μ Interaction

- 94.9% overall hit efficiency
- With 3D connections, consistency between views improved from 70% to 98%

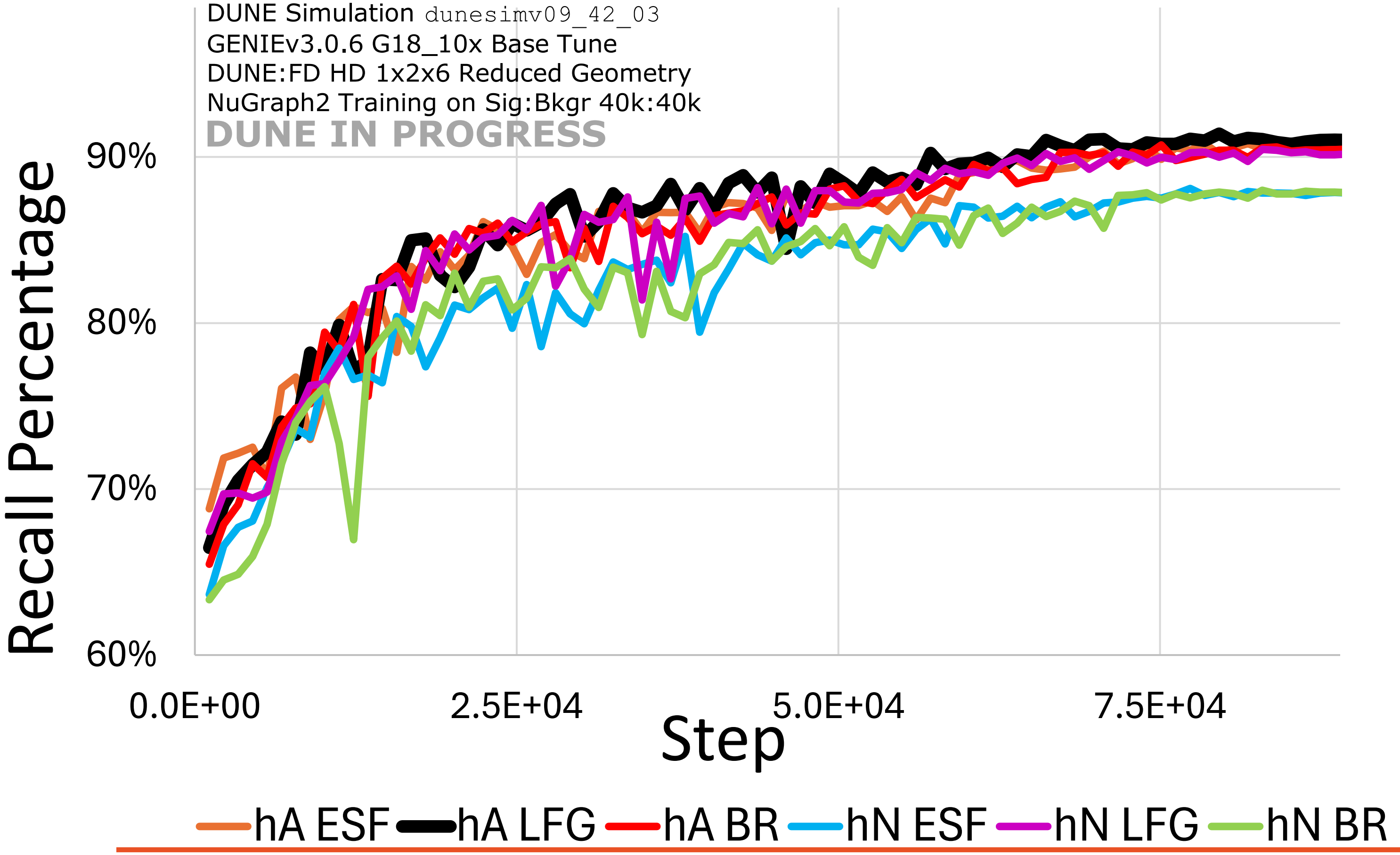
- NuGraph can semantically label detector hits well
  - Particle identification & topology is critical
    - Central in separating $p \rightarrow K^+\bar{\nu}$ signal from atmospheric ν background
- **NuGraph can also operate as a binary classifier for event classification**
  - **Direct access to signal-like or background-like discriminator**

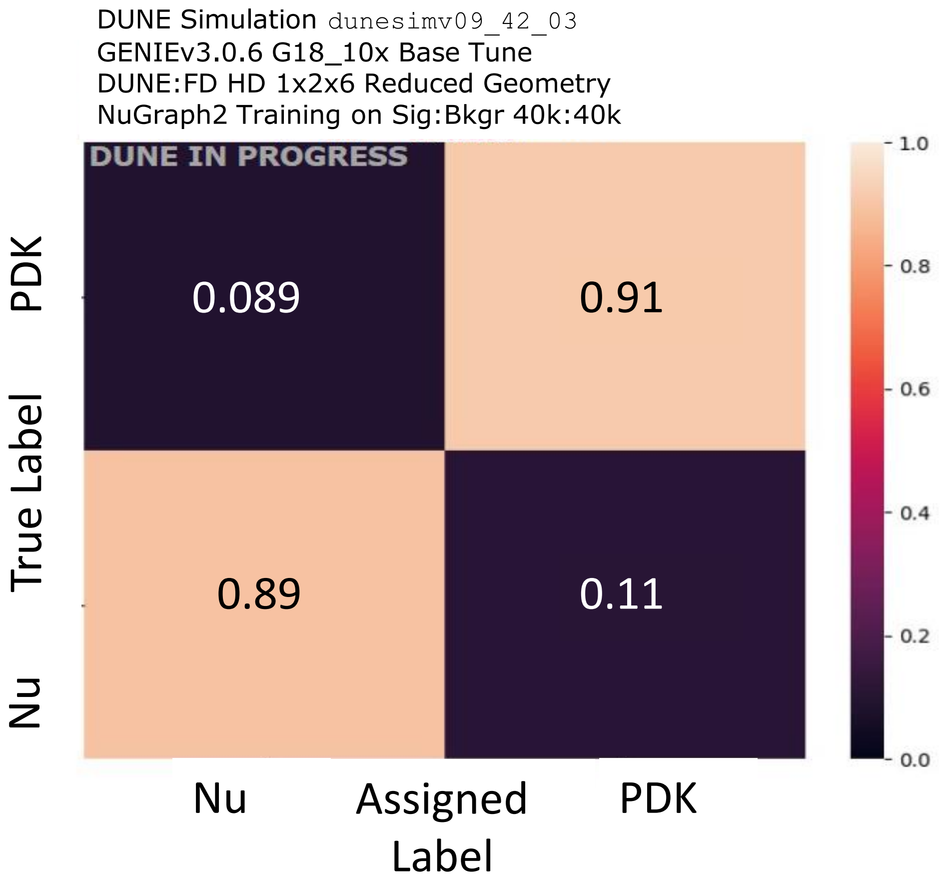


V Hewes CHEP 2023 PhysRevD.110.032008 Slide Adapted from J. Barrow

NuGraph Event Classification

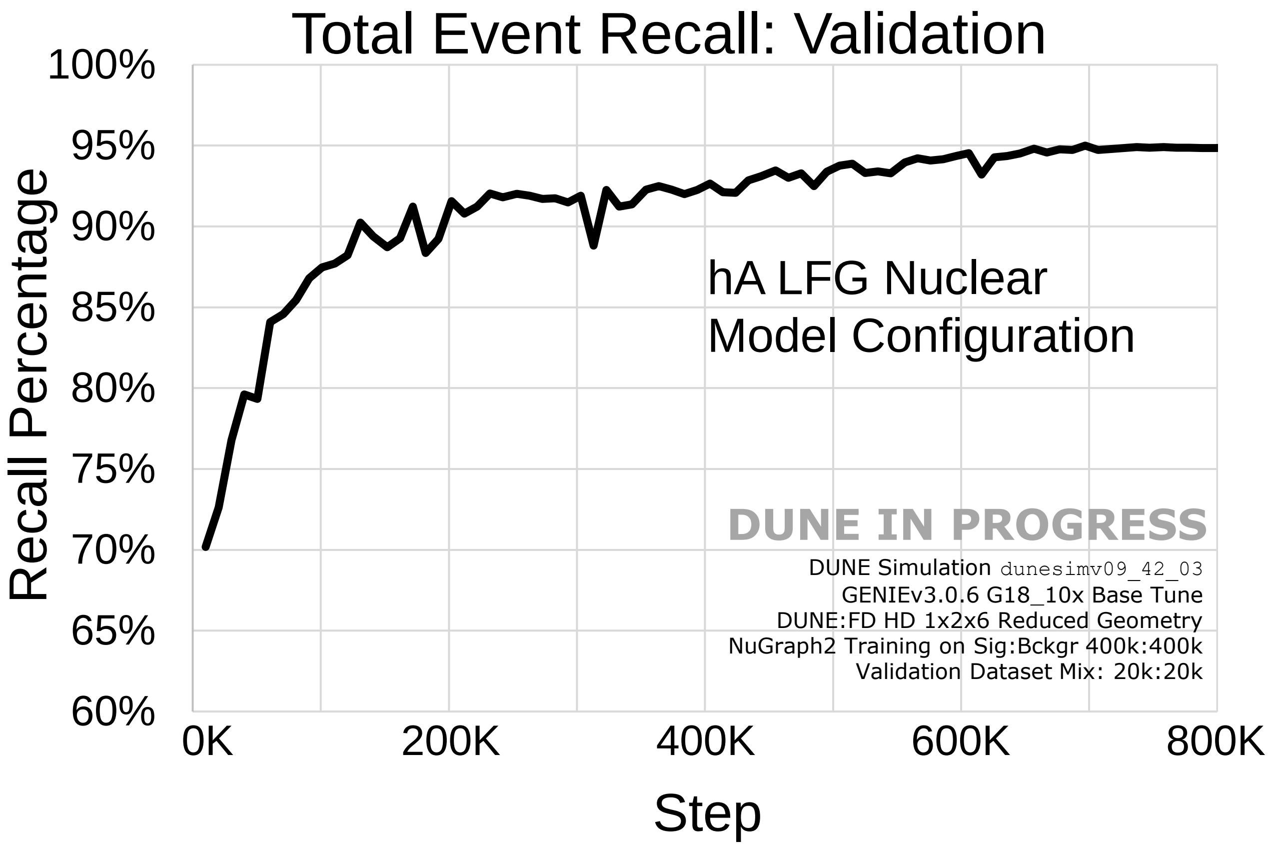
- Training on 40K:40K Signal : Background Events
- 1 Pre-Selection: 20 total hits in the event
- Score Cut Value: 0.5 [0.49 score → Bkg. 0.51 score → Signal]

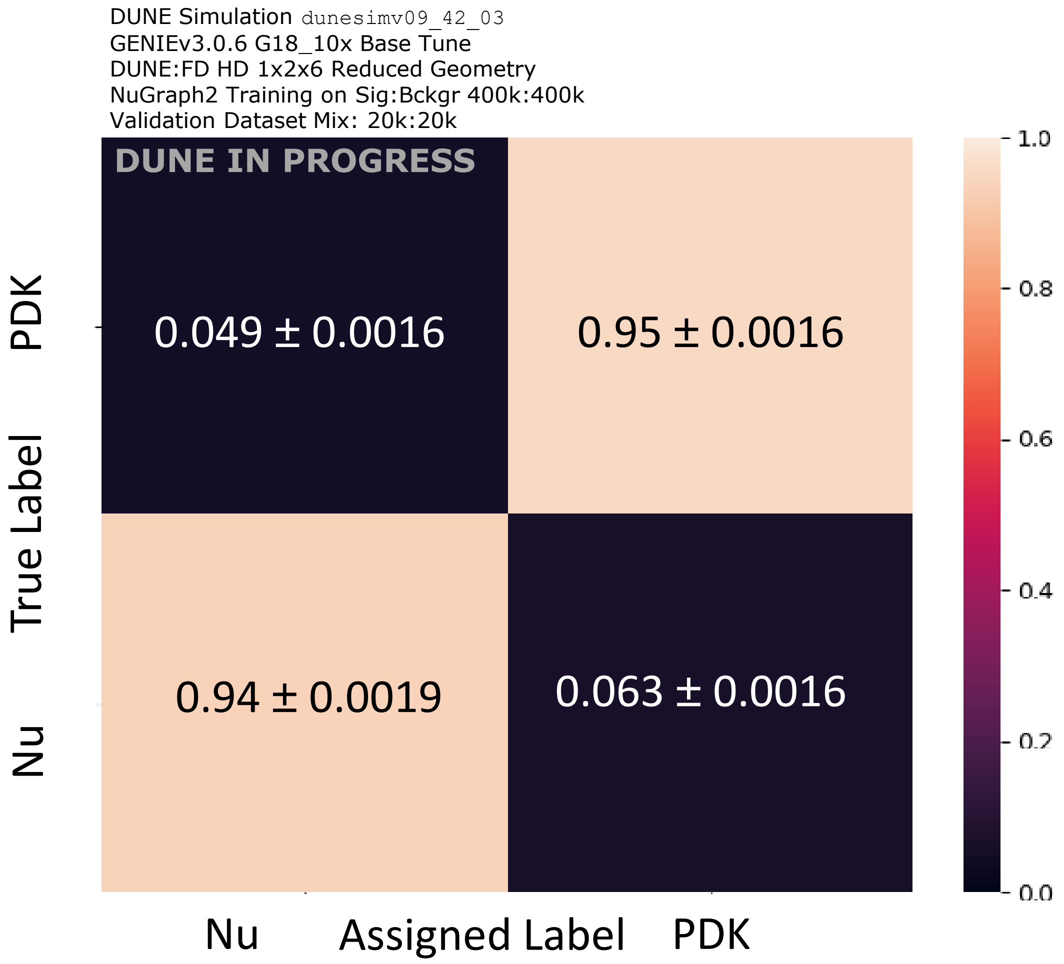




NuGraph Performance: Larger Dataset

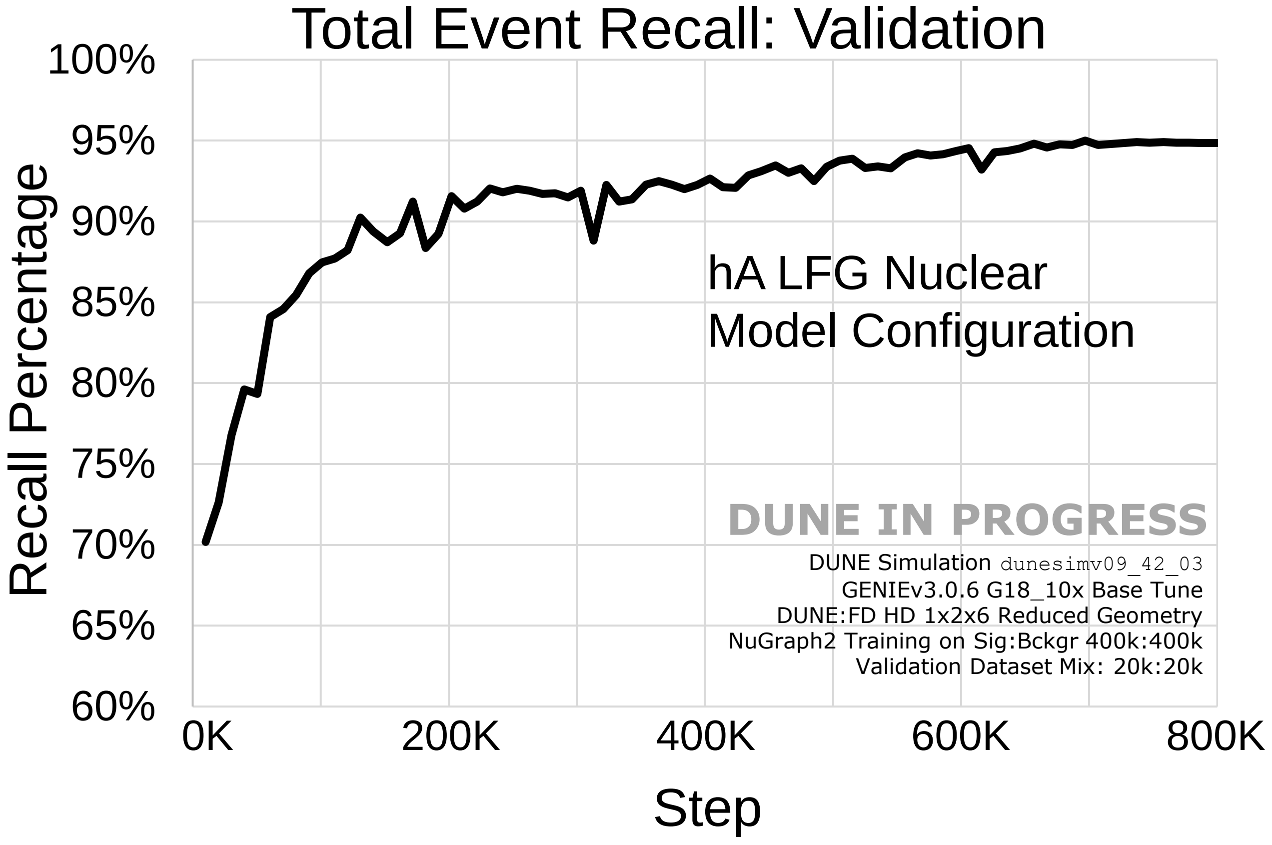
- Dataset: 400K hA_LFG Proton Decay and 400K hA_LFG Atmospheric Neutrinos
- 1 Pre-Selection: 20 total hits in the event
- 94.88% Event Recall

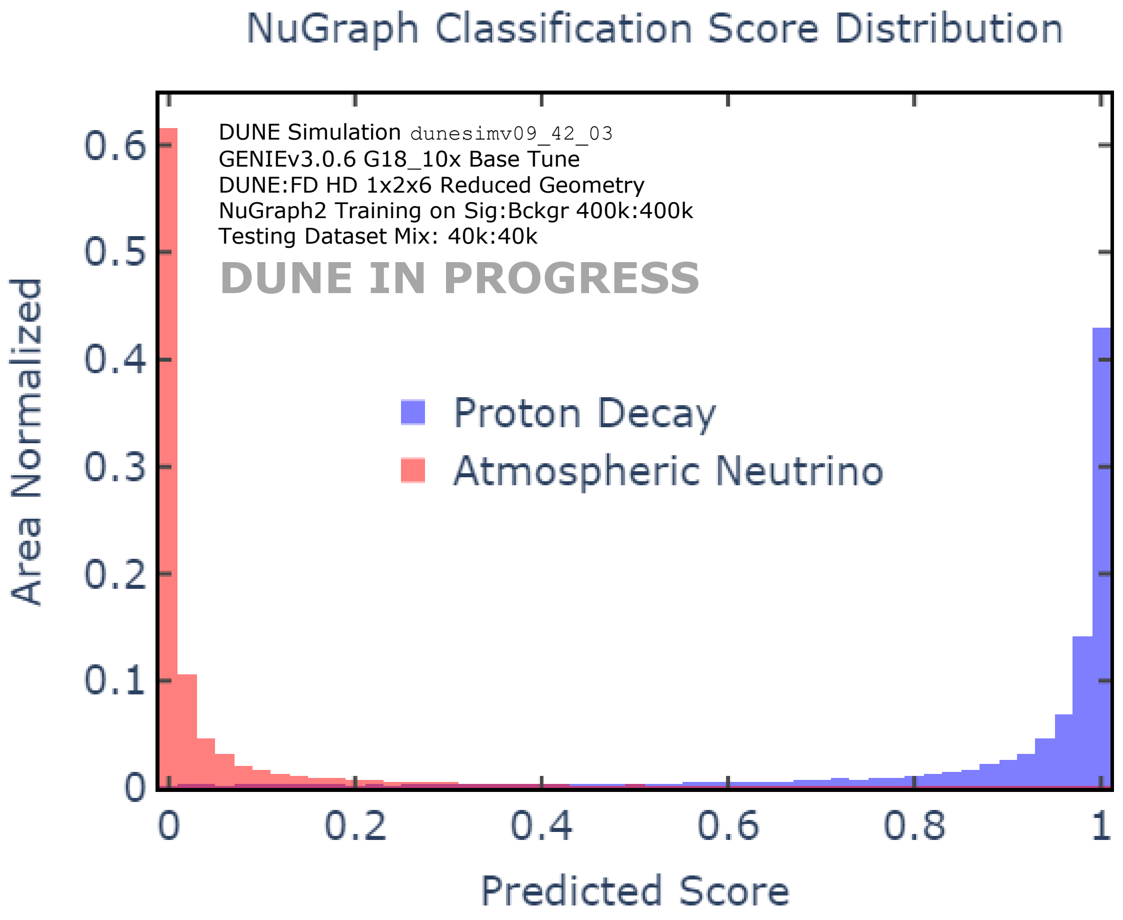




NuGraph Performance: Larger Dataset

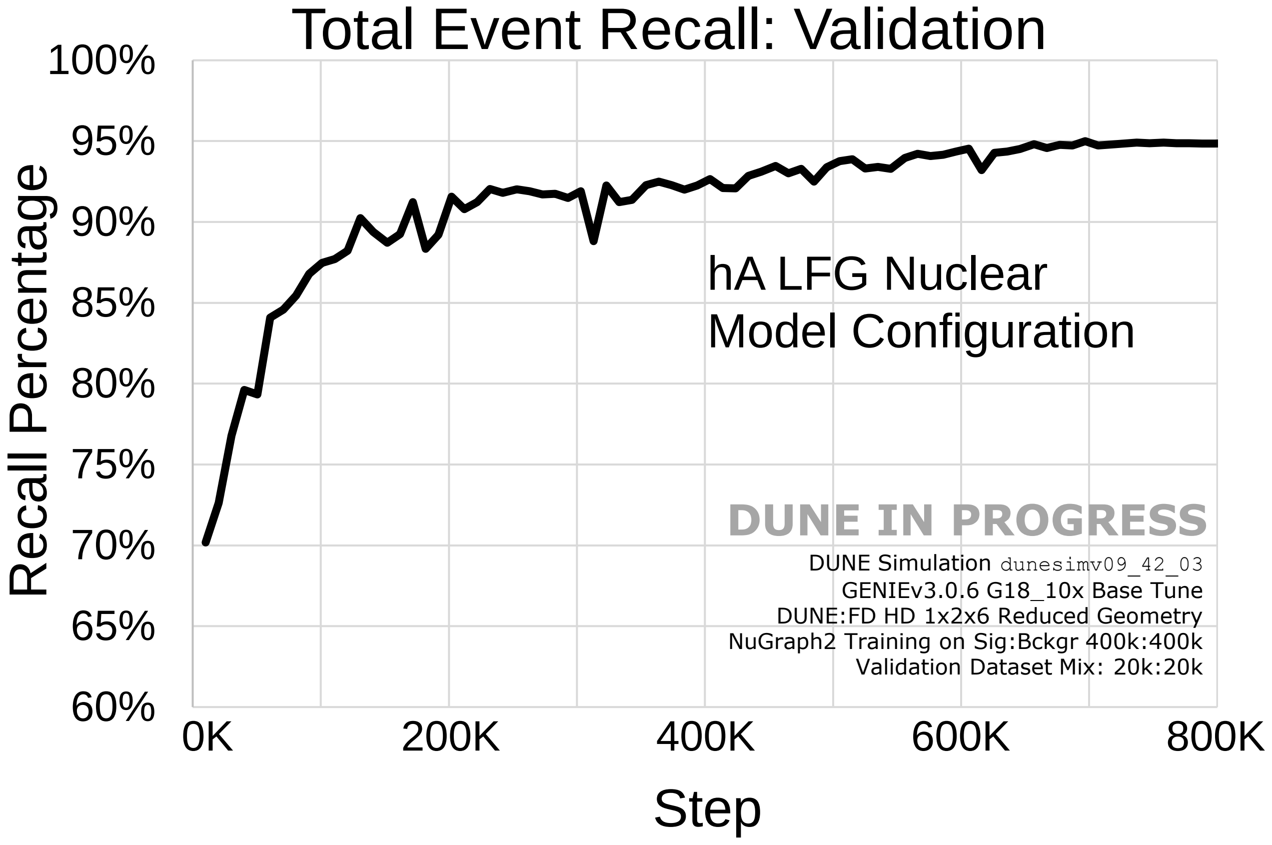
- Dataset: 400K hA_LFG Proton Decay and 400K hA_LFG Atmospheric Neutrinos
- 1 Pre-Selection: 20 total hits in the event
- 94.88% Event Recall

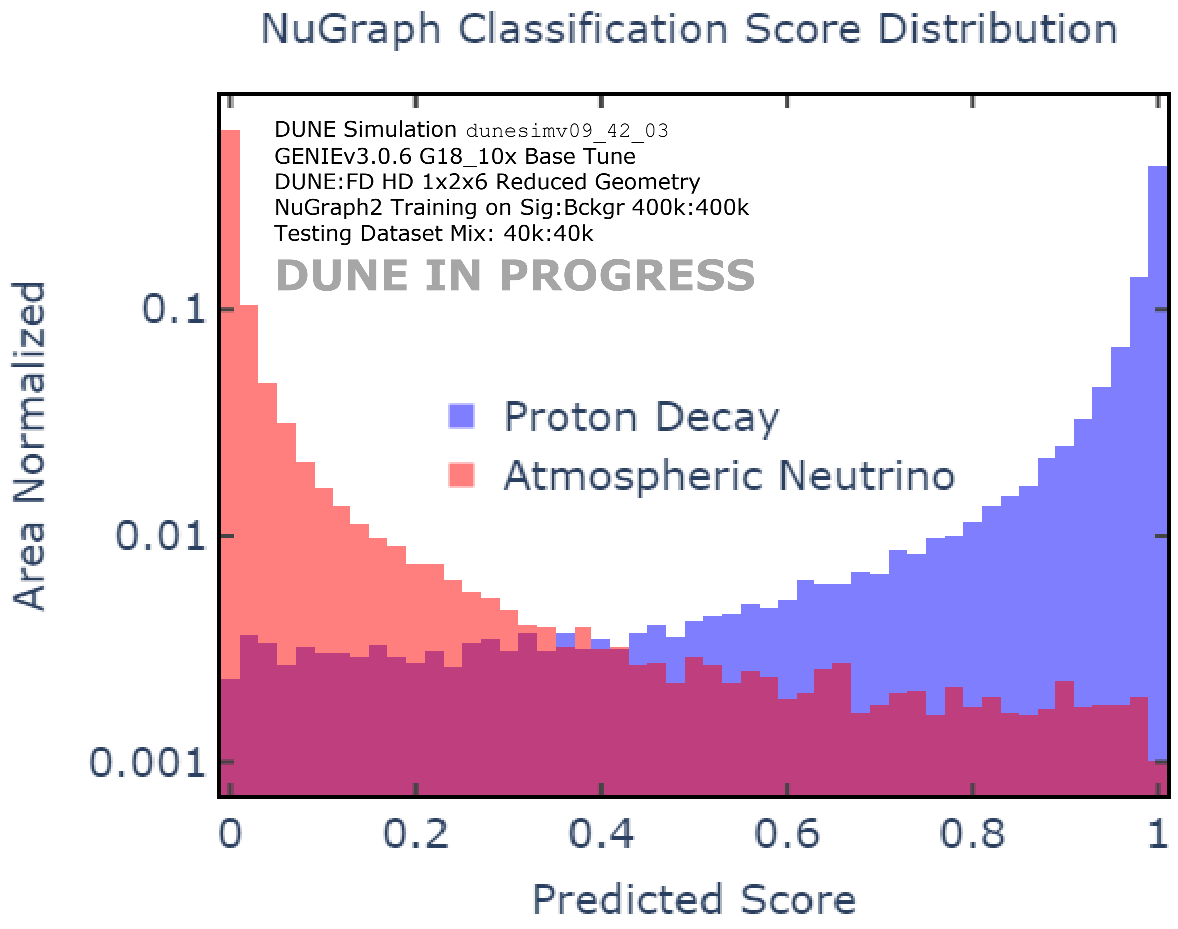




NuGraph Performance: Larger Dataset

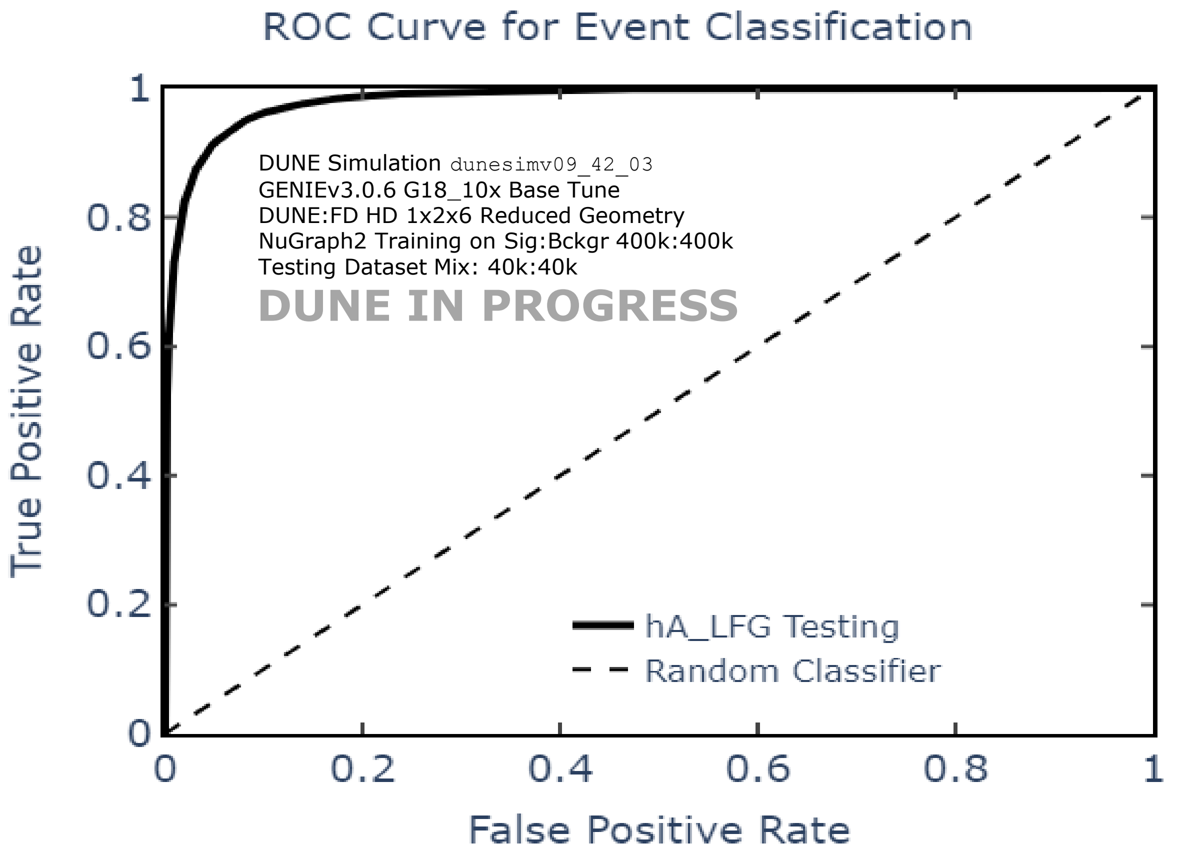
- Dataset: 400K hA_LFG Proton Decay and 400K hA_LFG Atmospheric Neutrinos
- 1 Pre-Selection: 20 total hits in the event
- 94.88% Event Recall

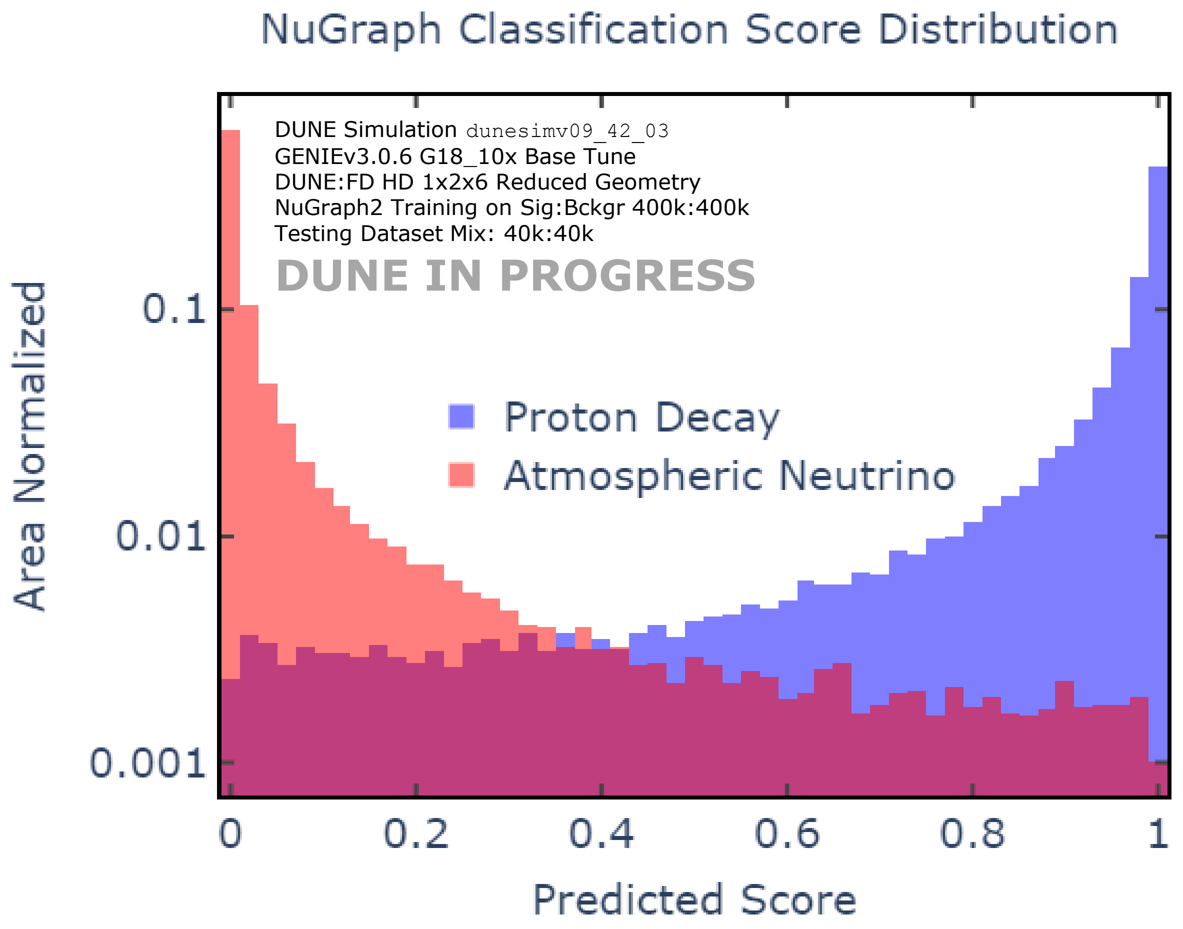




NuGraph Performance: Larger Dataset

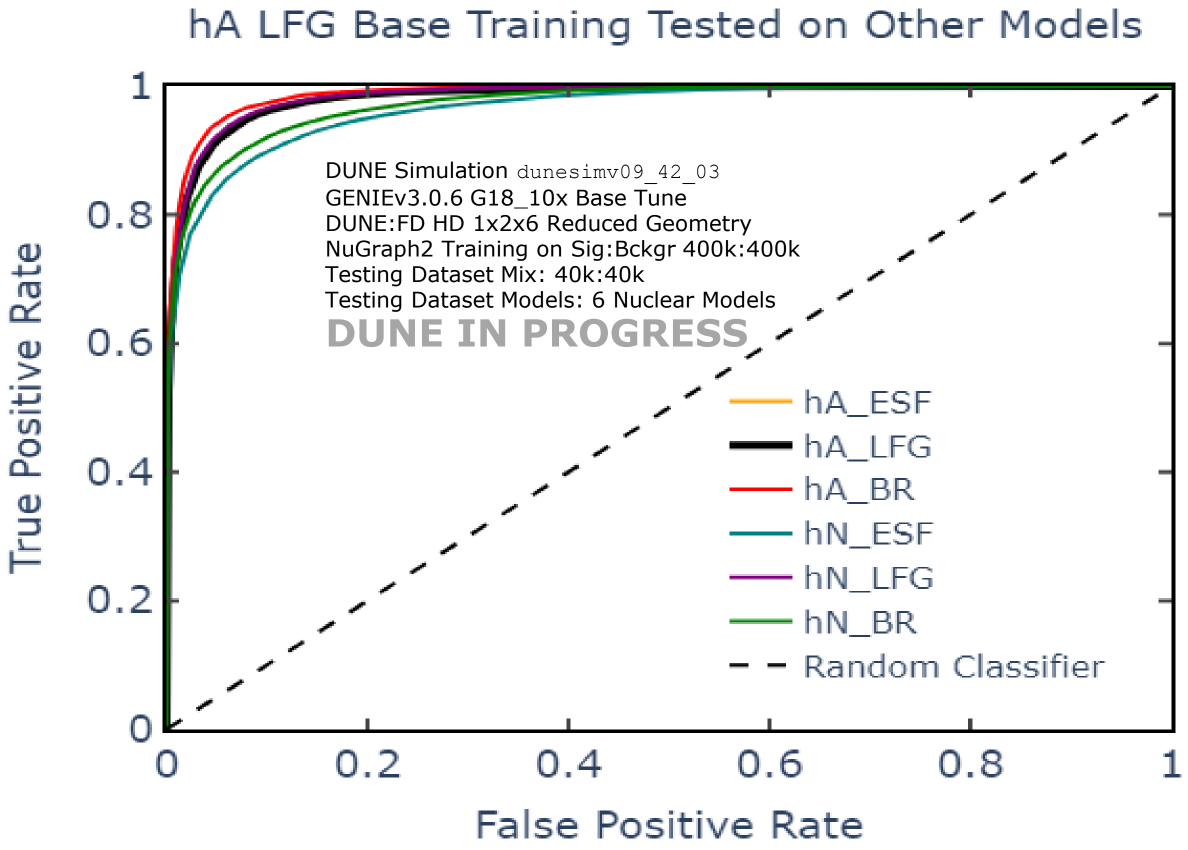
- Dataset: 400K hA_LFG Proton Decay and 400K hA_LFG Atmospheric Neutrinos
- 1 Pre-Selection: 20 total hits in the event
- 94.88% Event Recall | 0.98 AUC

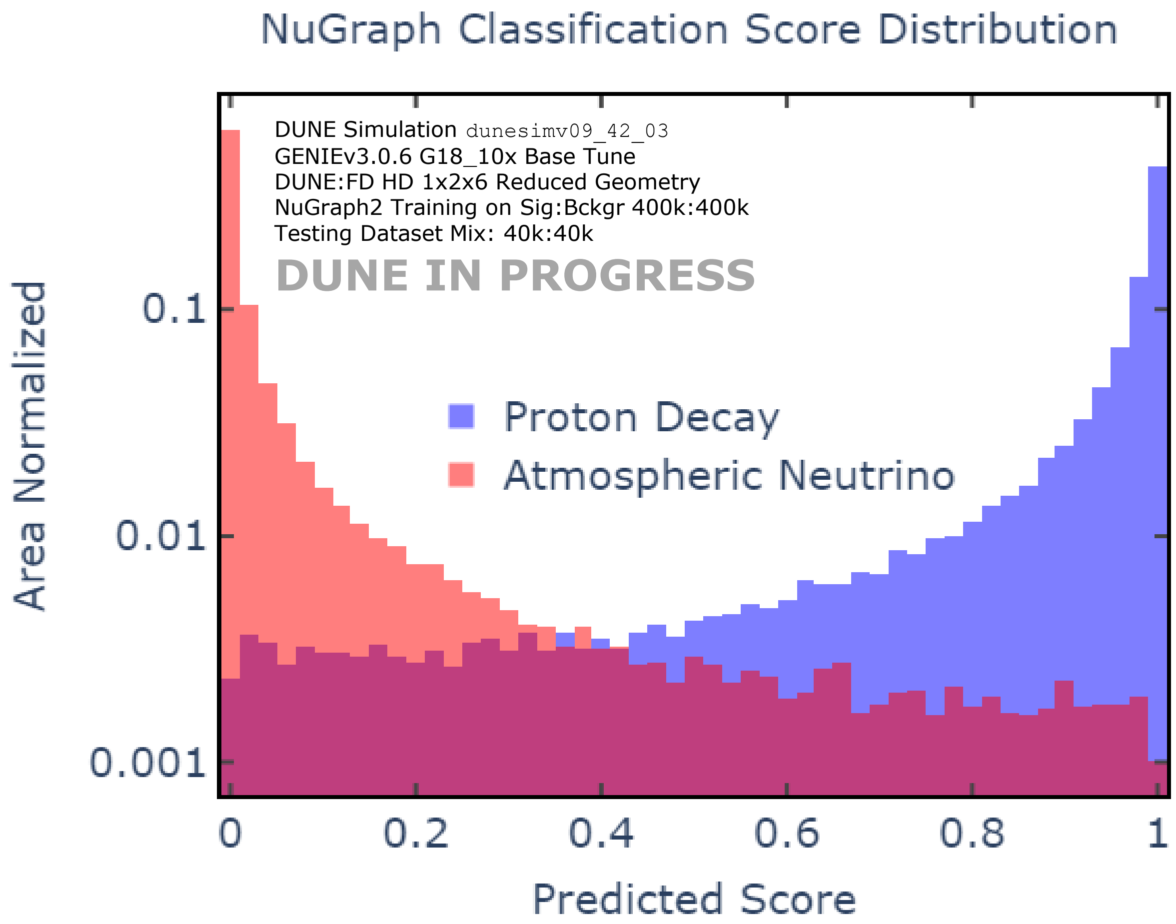




NuGraph Performance: Larger Dataset

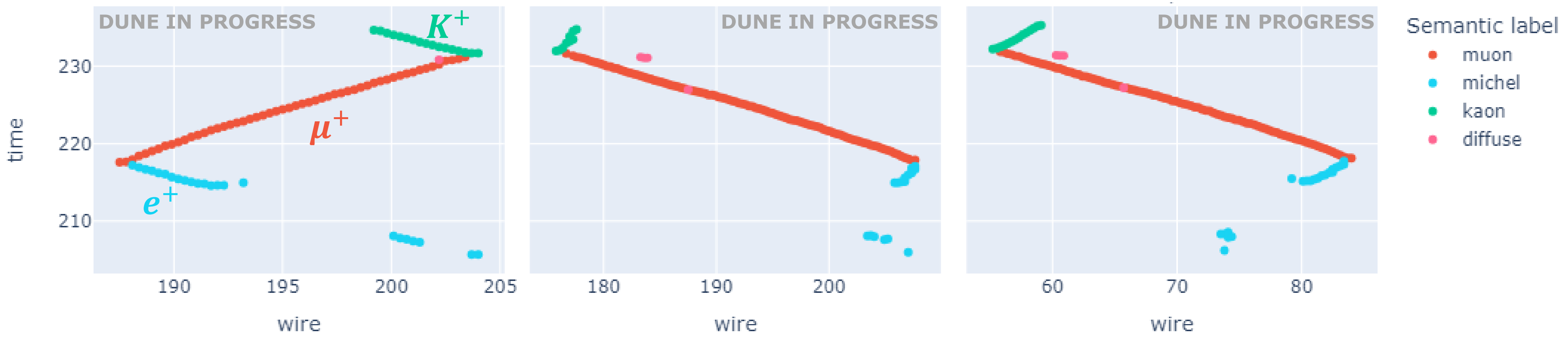
- Dataset: 400K hA_LFG Proton Decay and 400K hA_LFG Atmospheric Neutrinos
- 1 Pre-Selection: 20 total hits in the event
- 94.88% Event Recall | 0.96 to 0.98 AUC



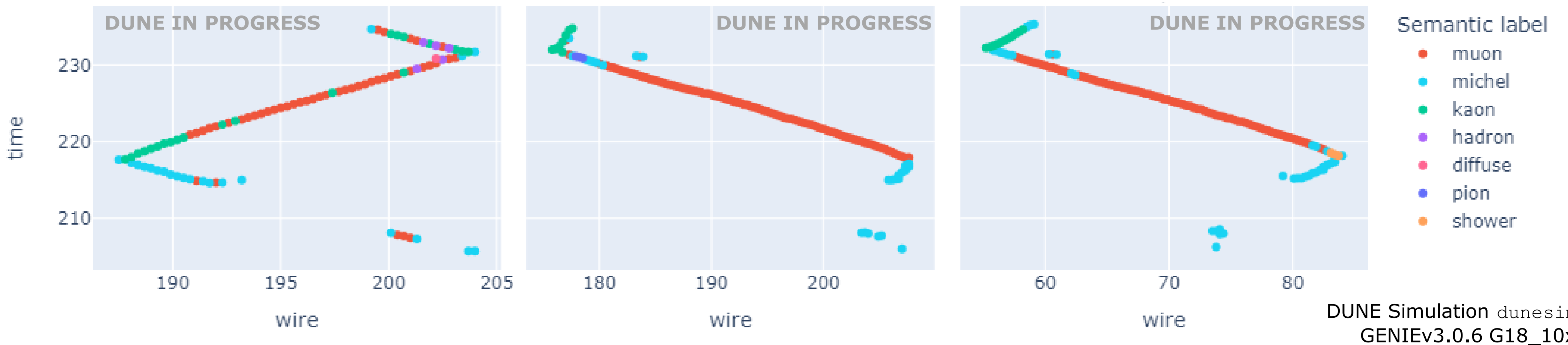


True Positive

True semantic labels for signal-like true $p \to K^+\bar{\nu}$ signal



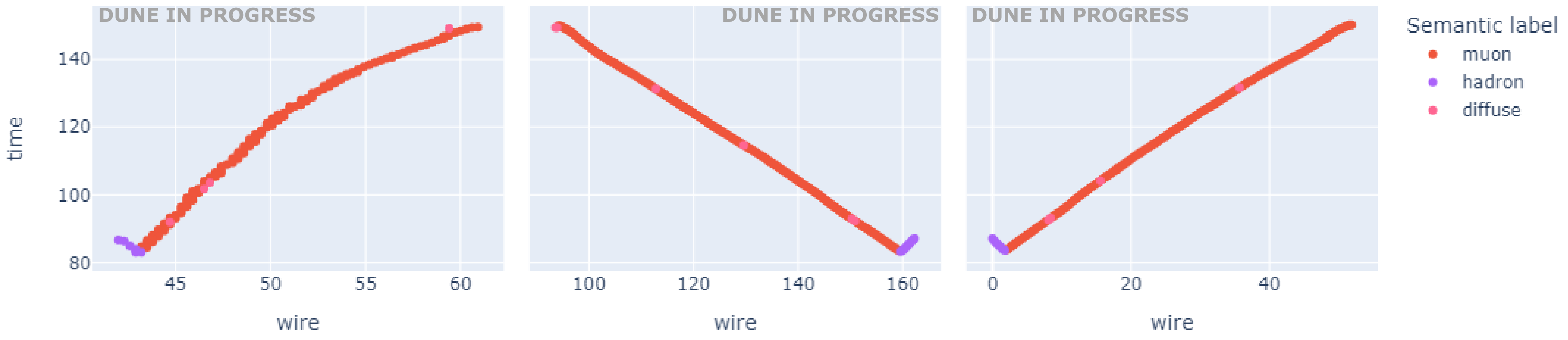
Predicted semantic labels for signal-like true $p \to K^+\bar{\nu}$ signal



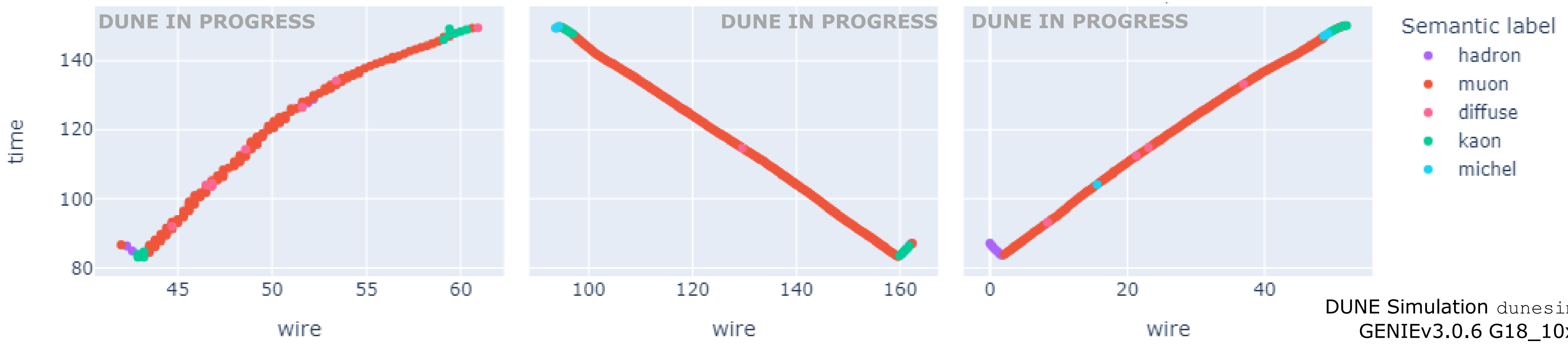
DUNE Simulation `dunesimv09_42_03`
GENIEv3.0.6 G18_10x Base Tune
DUNE:FD HD 1x2x6 Reduced Geometry
NuGraph2 Training on hA LFG Sig:Bckgr 40k:40k

True semantic labels for signal-like true atmospheric neutrino background

False Positive



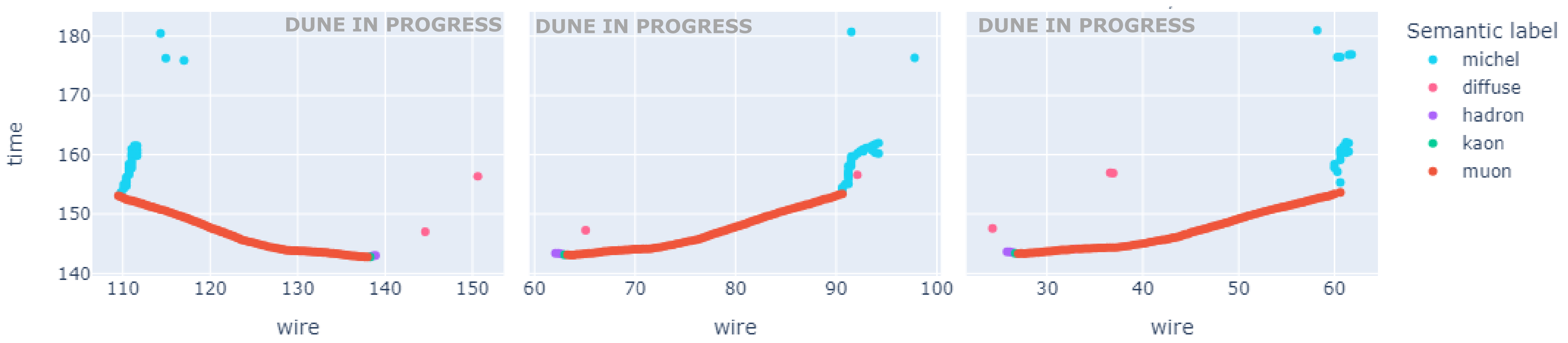
Predicted semantic labels for signal-like true atmospheric neutrino background



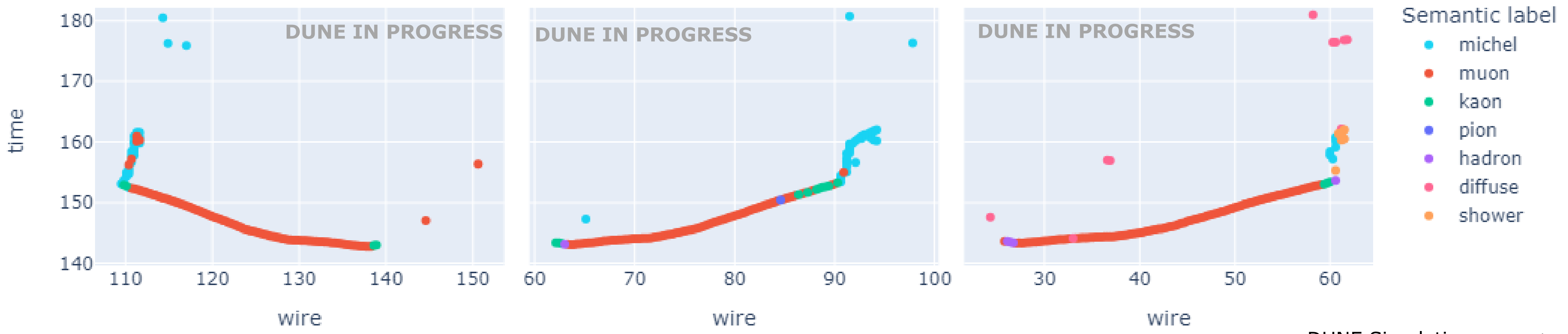
DUNE Simulation `dunesimv09_42_03`
GENIEv3.0.6 G18_10x Base Tune
DUNE:FD HD 1x2x6 Reduced Geometry
NuGraph2 Training on hA LFG Sig:Bckgr 40k:40k

False Negative

True semantic labels for background-like true $p \rightarrow K^{+}\bar{\nu}$ signal



Predicted semantic labels for background-like true $p \rightarrow K^{+}\bar{\nu}$ signal



DUNE Simulation `dunesimv09_42_03`
GENIEv3.0.6 G18_10x Base Tune
DUNE:FD HD 1x2x6 Reduced Geometry
NuGraph2 Training on hA LFG Sig:Bckgr 40k:40k

Conclusion

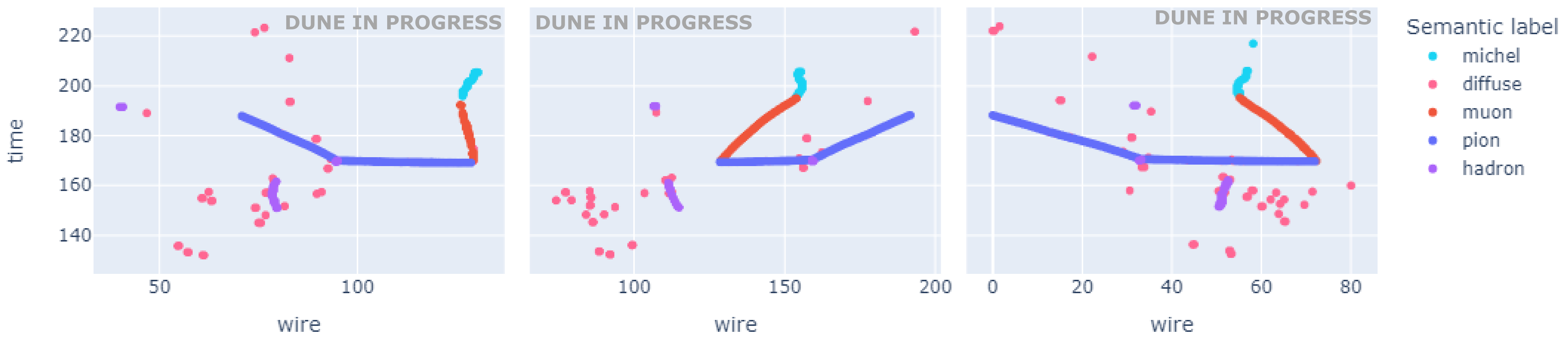
- NuGraph is a GNN originally designed for semantically classifying detector hits
- The network has been expanded to perform event classification for proton decay
- Iteration over nuclear model configurations can provide a conservative estimate of nuclear model uncertainty
- NuGraph performance is very encouraging! (0.98 AUC)
  - Yields similar classification performance as current BDT methods in DUNE
- Past studies have utilized a CNN + BDT combined approach to yield best results
  - Building on this, a combined GNN + BDT approach is being explored
- A lot of other ongoing work in DUNE on proton decay is also taking place!
  - See Natsumi's poster this afternoon

Thank you for your time!

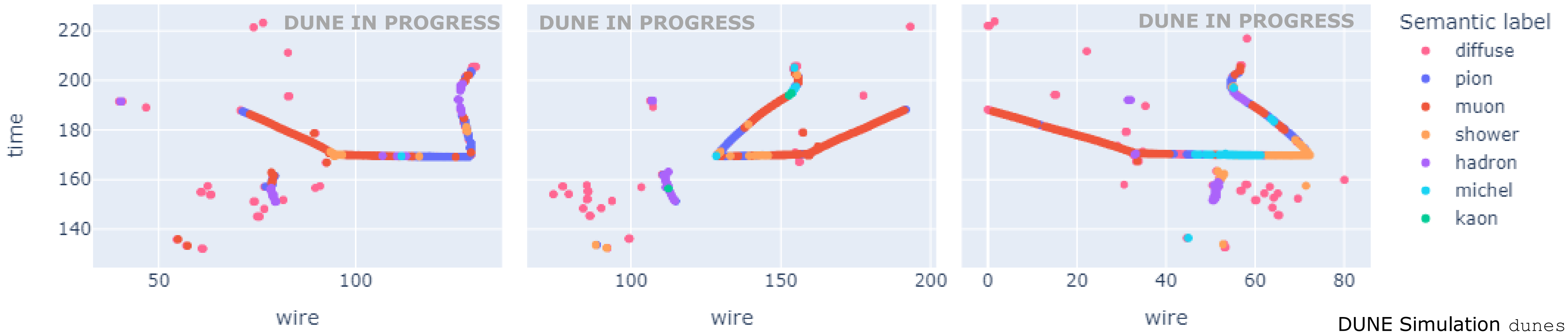
Contact Info

True semantic labels for background-like true atmospheric neutrino background

True Negative



Predicted semantic labels for background-like true atmospheric neutrino background

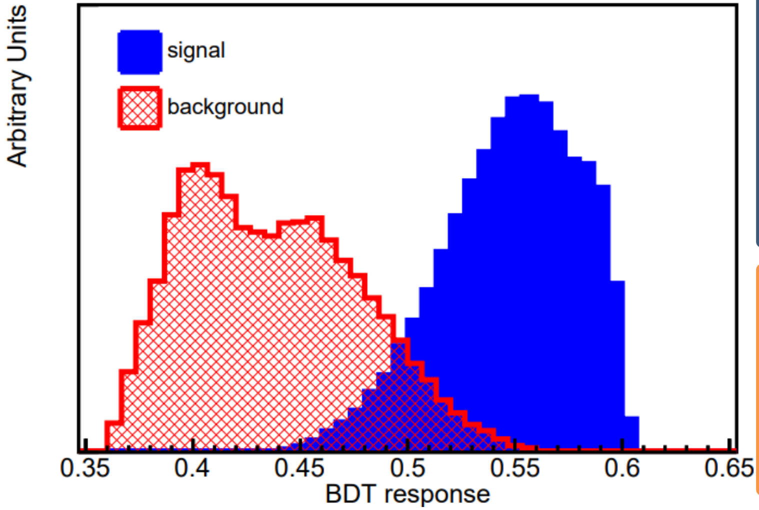


DUNE Simulation `dunesimv09_42_03`
GENIEv3.0.6 G18_10x Base Tune
DUNE:FD HD 1x2x6 Reduced Geometry
NuGraph2 Training on hA LFG Sig:Bckgr 40k:40k

$p \rightarrow K^+\bar{\nu}$ Sensitivity

https://link.springer.com/article/10.1140/epjc/s10052-021-09007-w

- Kaon tracking efficiency: 58%
  - With improved reconstruction this can be greatly improved
- 30% signal efficiency
  - Main limiting factor in signal efficiency is K/p separation
- 3×10^{-6} background suppression
  - 1 background per Mton-year or 25 years of data
- Systematics:
  - 2% on signal from FSI uncertainties
  - 20% on background from neutrino flux and cross-section uncertainties



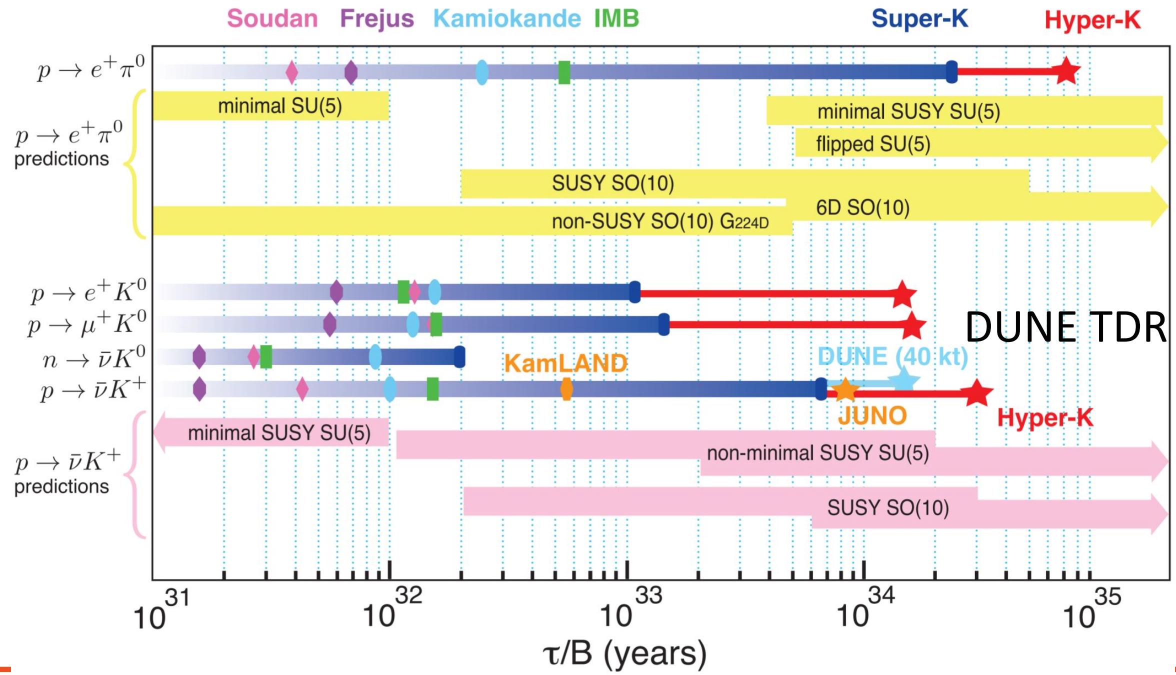
Expected Sensitivity

400 kt-year exposure with no observed events → a limit of 1.3×10^{34} years

Current Limit by SK

5.9×10^{33} years

https://journals.aps.org/prd/abstract/10.1103/PhysRevD.90.072005

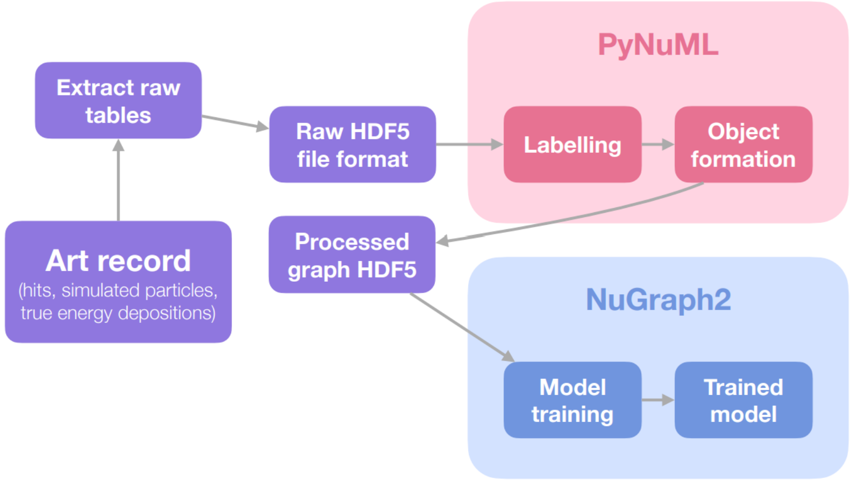


How Does NuGraph2 Work?

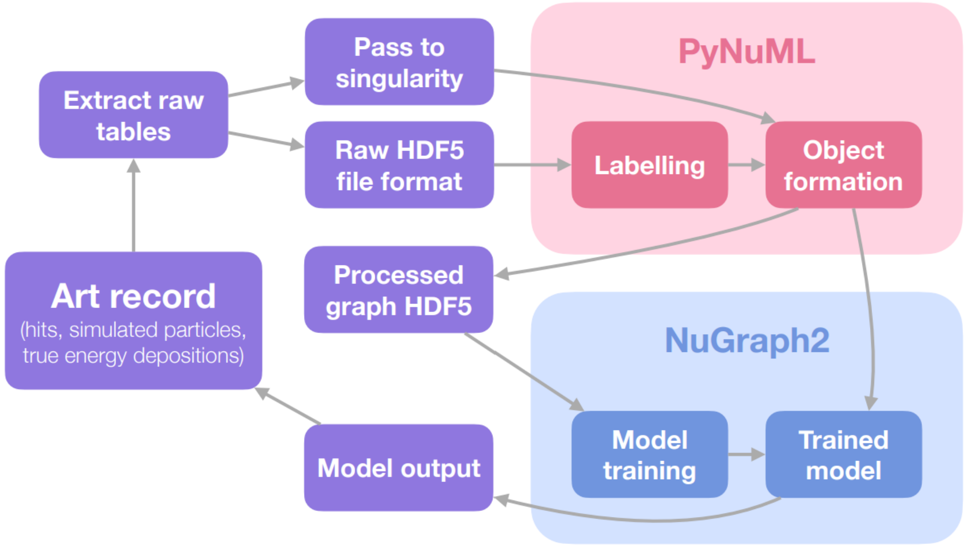
- NuML: ROOT files converted to HDF5
  - Holds low level information
    - Simulated particle labels, hits, true energy deposition, etc.
  - GitHub: https://github.com/nugraph/numl

- PyNuML: Provides efficient, flexible solution for tasks leveraging ML in particle physics
  - Defines ground truth labels
  - Arranges detector hits in ML objects (graphs)
  - Makes pixel maps, etc.
  - GitHub: https://github.com/nugraph/pynuml

Producing Graphs for Model Training



Inference in Production

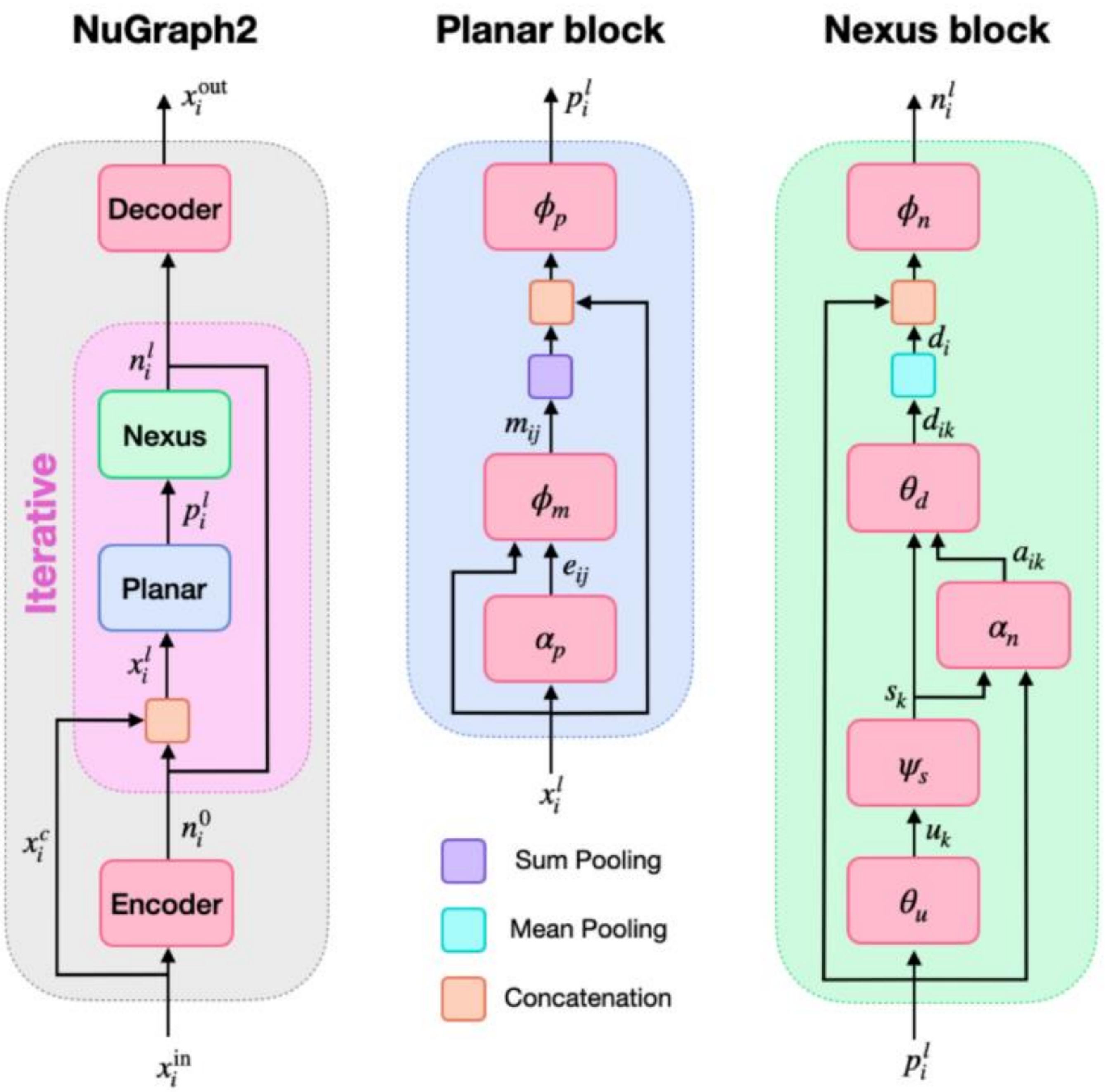


Network Architecture

- NuGraph2 architecture consists of iterative message passing

- Each message passing iteration consists of two blocks:
  - Planar Block: internal passing within each plane
  - Nexus Block: pass up to 3D nexus nodes to share context information amongst the Planar Blocks

- Messages are embedded on a categorical embedding
  - Each semantic category is provided with a separate set of embedded features, which are convolved independently
  - Context information is exchanged between different particle types via a categorical cross-attention mechanism



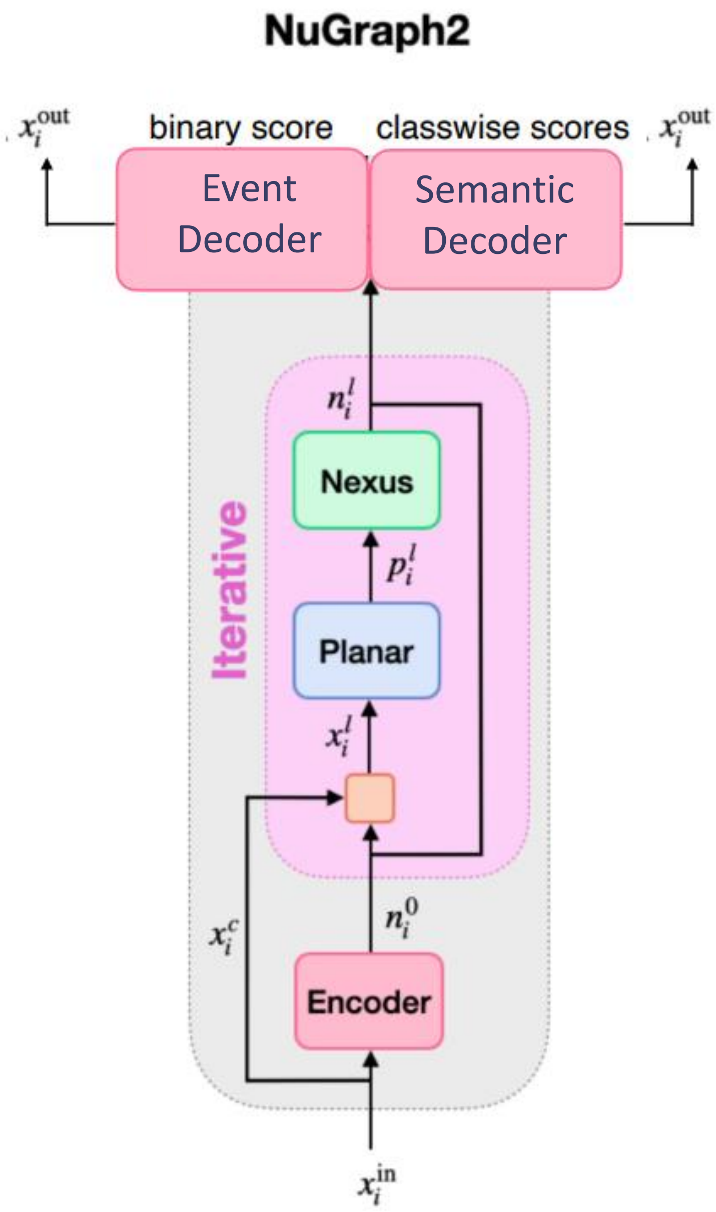
ExaTrkx Phase II Kickoff: Cerati PhysRevD.110.032008 Neutrino ML 2024: Cerati

Decoders

- At the end of the message passing are decoders

- Semantic: classify each hit by particle type
- Event: classify each event by type (pdk or atmospheric neutrino)
- Output both class-wise scores from the semantic decoder and a binary score from the event decoder
- Same learned features are used as input to all decoders

- Other decoders are being explored as well!



[ExaTrkx Phase II Kickoff: Cerati](#) [PhysRevD.110.032008](#) [Neutrino ML 2024: Cerati](#)