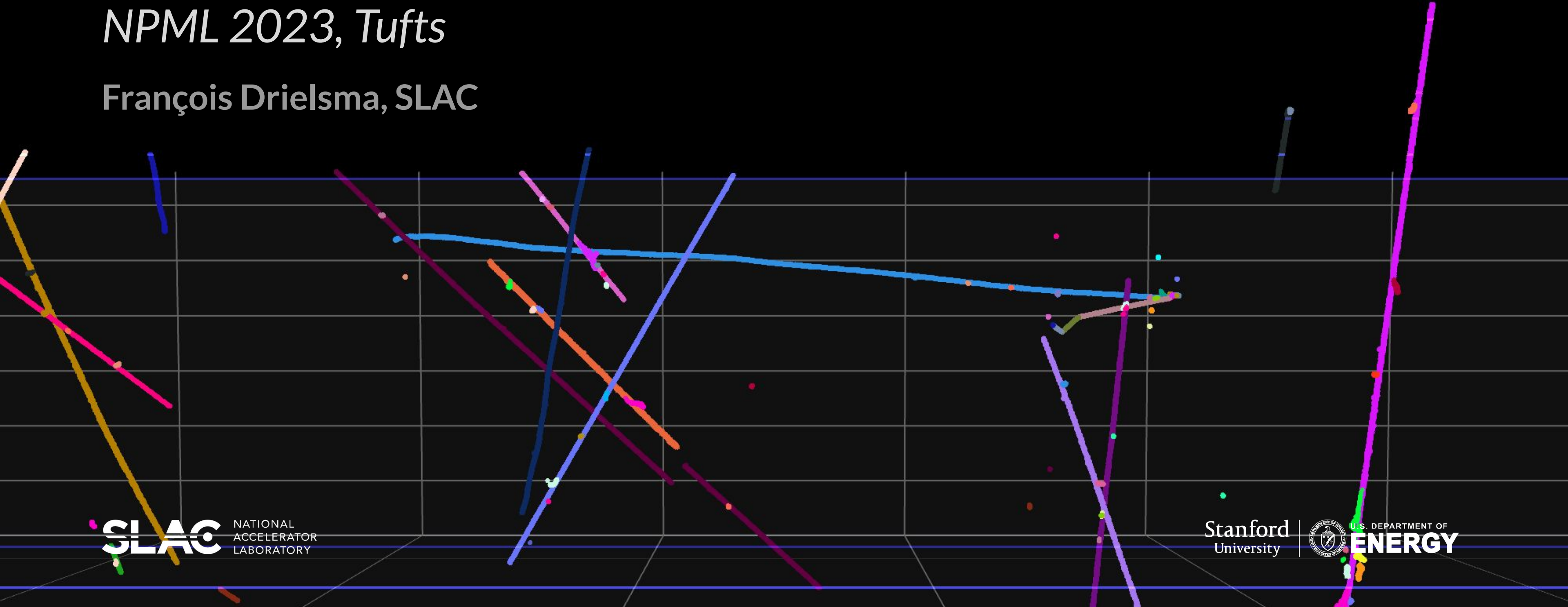
End-to-end, ML-based Reconstruction Chain for the Short Baseline Neutrino Program

NPML 2023, Tufts

François Drielsma, SLAC

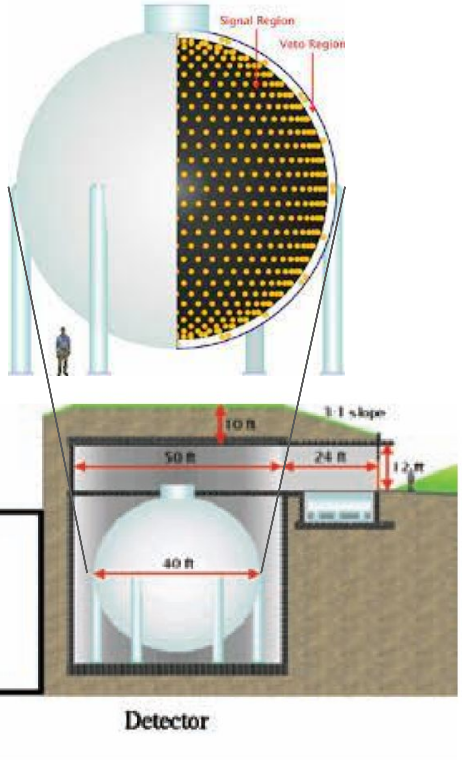


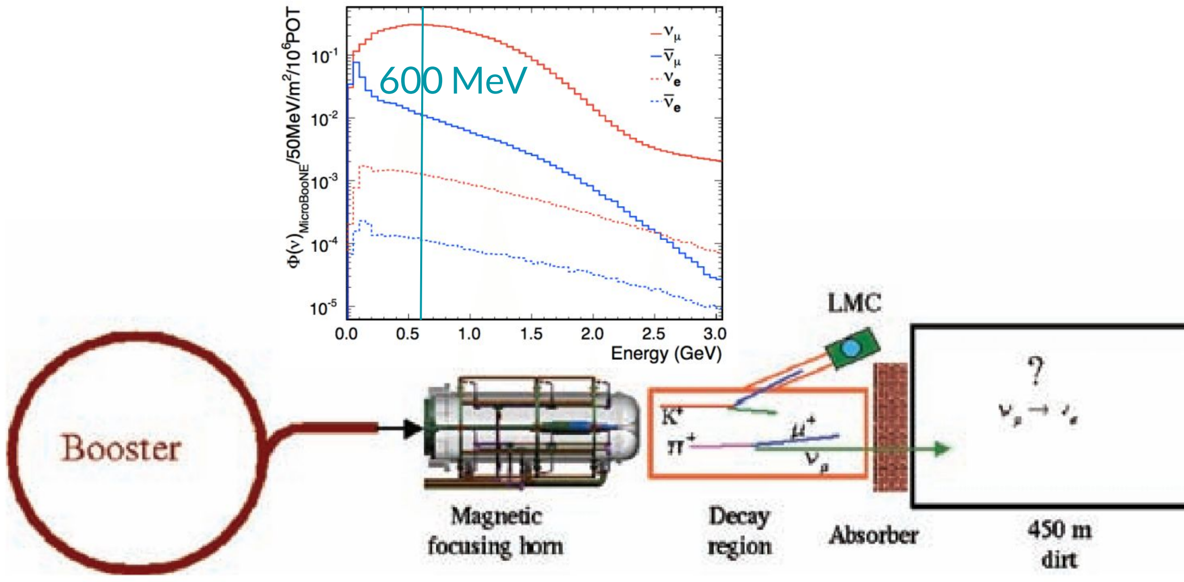
SLAC NATIONAL ACCELERATOR LABORATORY
Stanford University
U.S. DEPARTMENT OF ENERGY

The MiniBooNE Low Energy Excess

MiniBooNE was a **short baseline** neutrino experiment

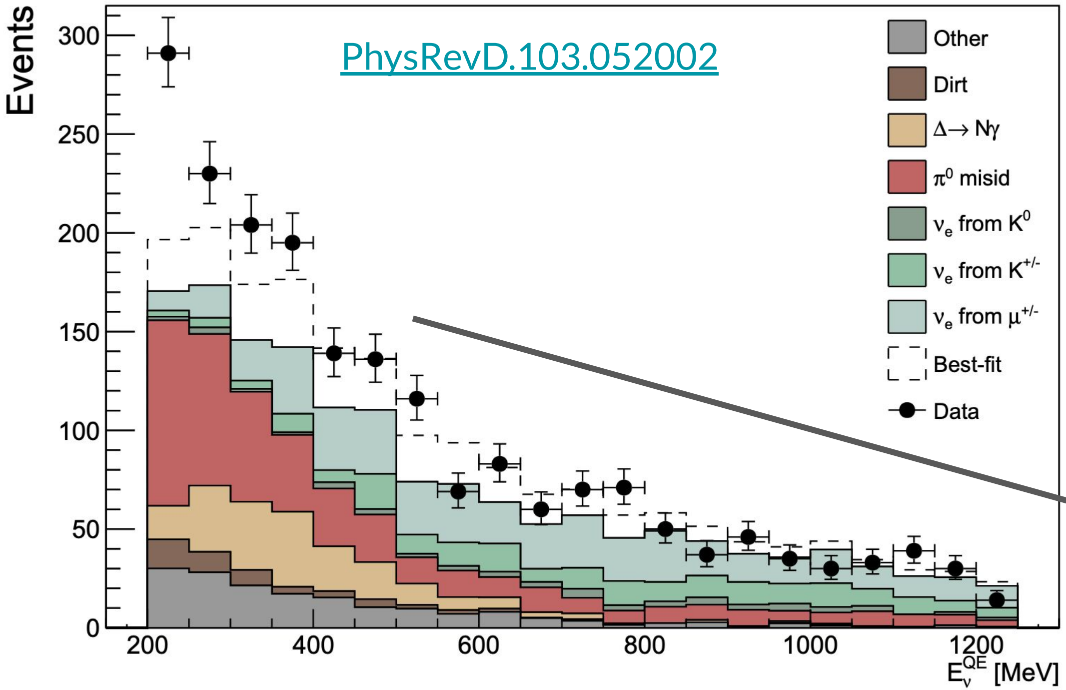
- **Booster Neutrino Beam** (BNB) at Fermilab
- Scintillator-based **Cherenkov detector**





The MiniBooNE Low Energy Excess

MiniBooNE observed **excess** of "**electron-like**" **neutrino events** (LSND-like)



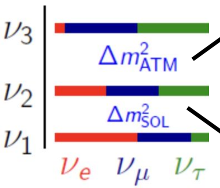
PhysRevD.103.052002

4.8 sigma excess…

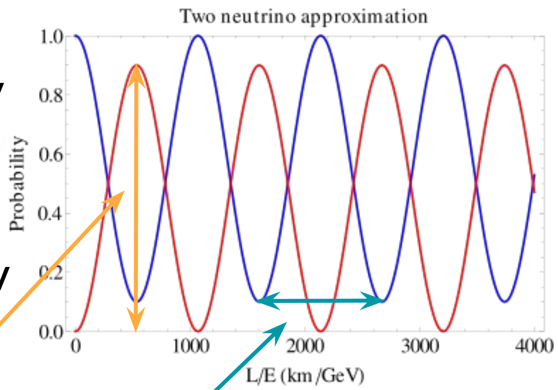
The MiniBooNE Low Energy Excess

Potential interpretation: excess of ν_e in the BNB ?

Peak: ~500 km/GeV



Peak: 1.6e4 km/GeV

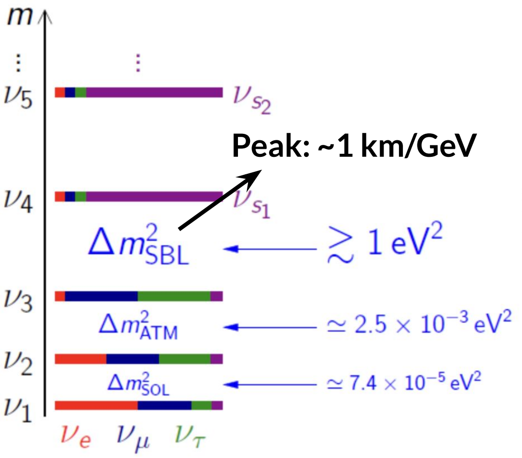


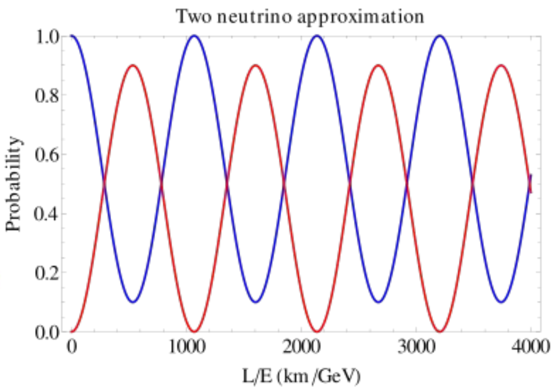
$$P_{\alpha\to\beta,\alpha\neq\beta} = \boxed{\sin^2(2\theta)}\,\sin^2\left(\boxed{1.27\,\frac{\Delta m^2}{E}}\!\!\frac{L}{}\ \frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

The MiniBooNE Low Energy Excess

Potential interpretation: excess of ν_e in the BNB ?

- In isolation, might be explained by >1 new **sterile ν eigenstate(s)**



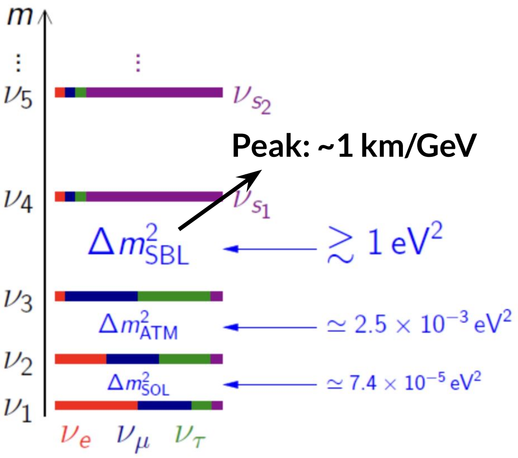


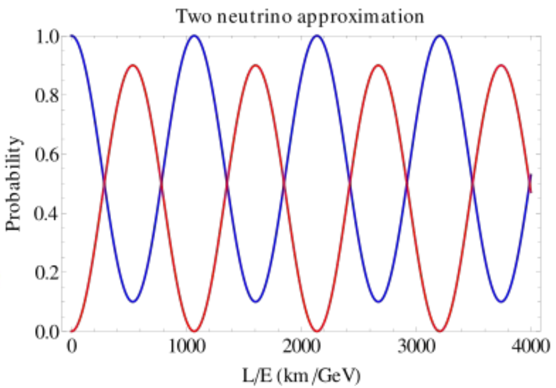
$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m^2 L}{E}\,\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

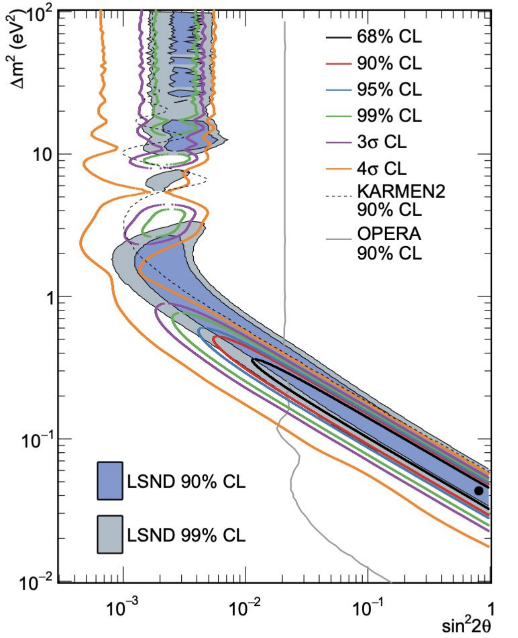
The MiniBooNE Low Energy Excess

Potential interpretation: excess of ν_e in the BNB ?

- In isolation, might be explained by >1 new **sterile ν eigenstate(s)**



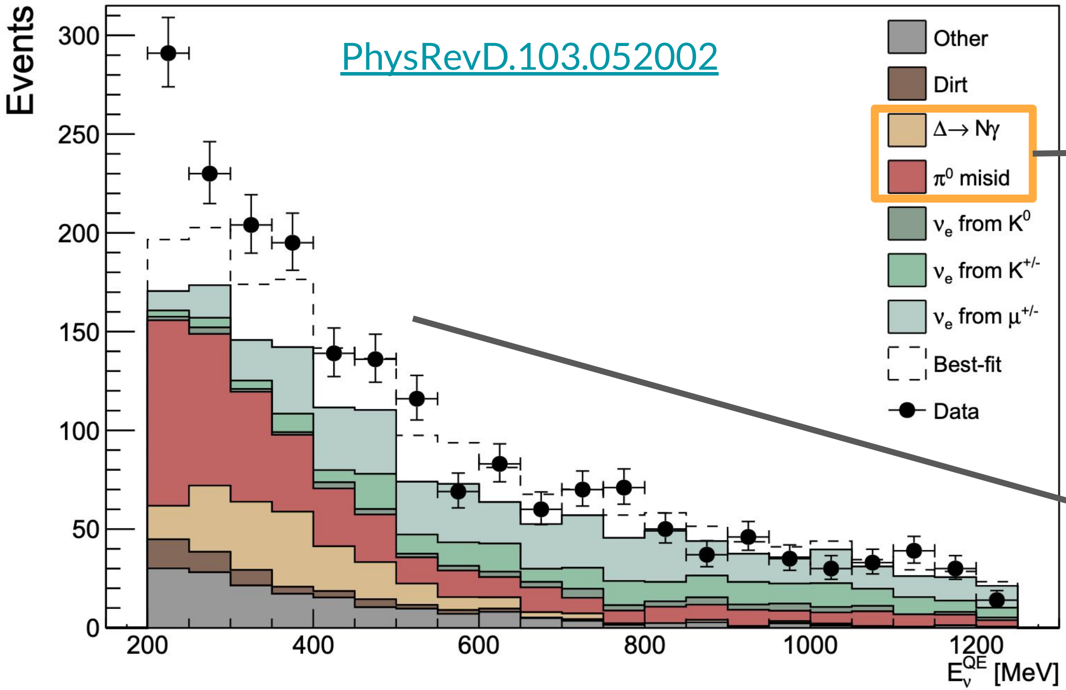




$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m^2 L}{E}\,\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

The MiniBooNE Low Energy Excess

Other interpretation: we just don't understand neutrino cross-sections…

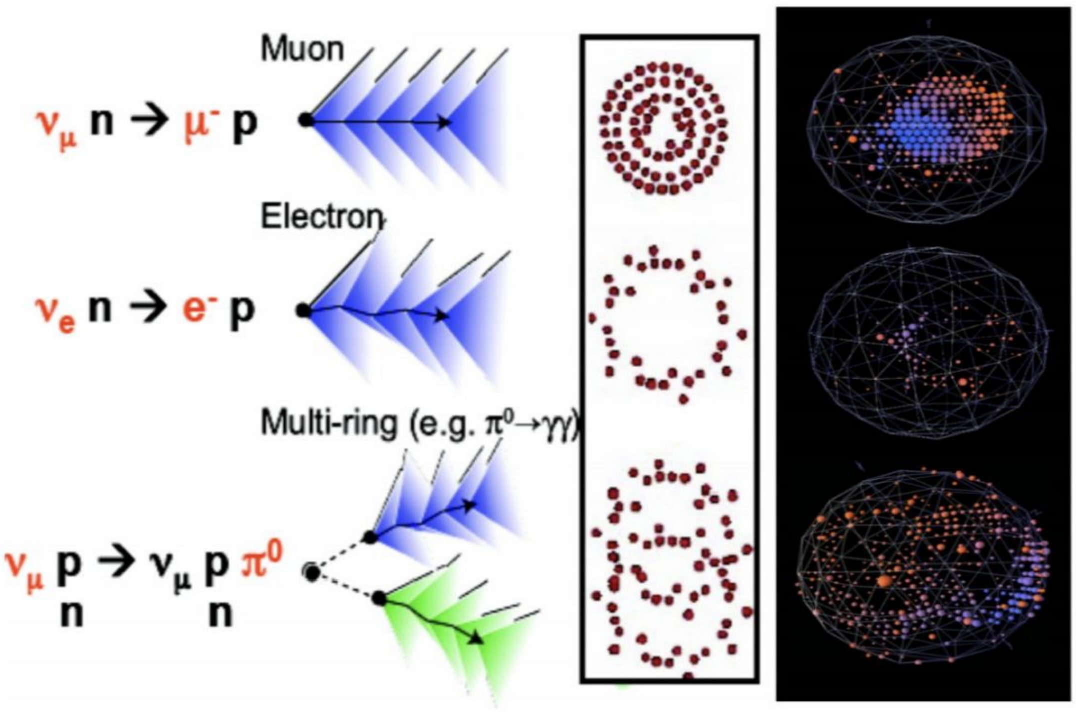


Low-energy "electron-like" events **dominated by ν_μ NC-γ and CC-π^0 background**

4.8 sigma excess…

The MiniBooNE Low Energy Excess

MiniBooNE's limitations: Cannot tell electrons from photons



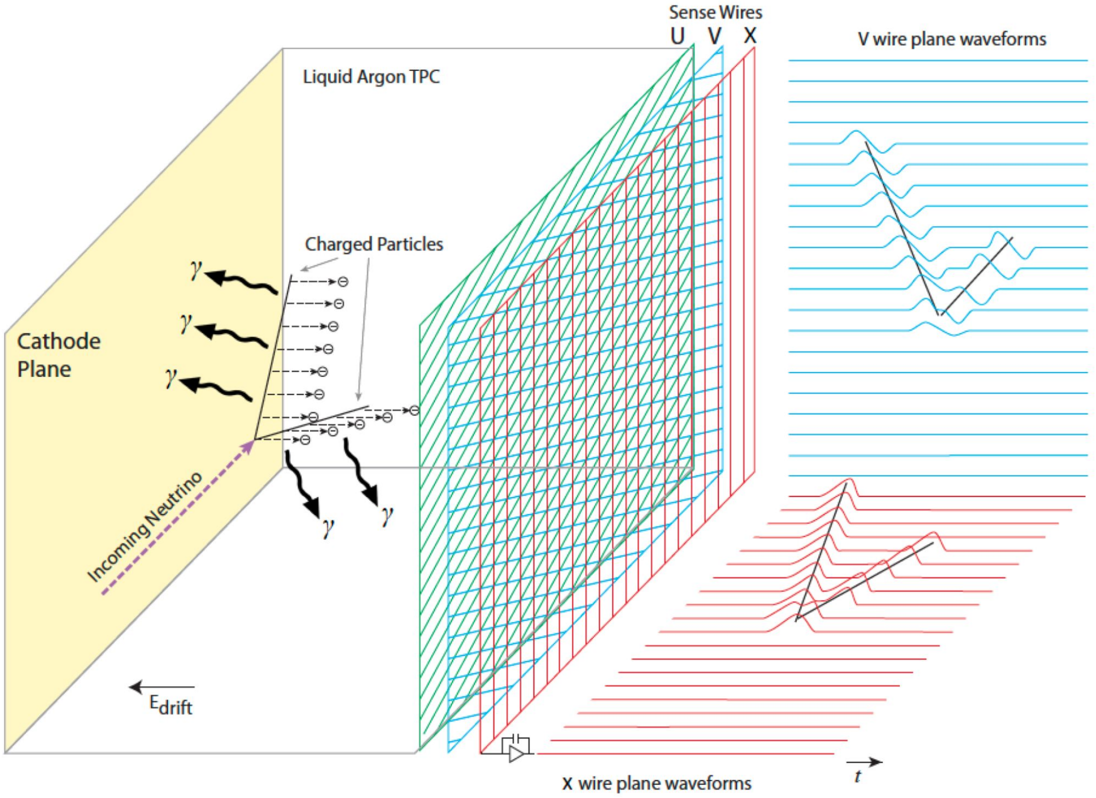
μ/e separation **reliable**

Single e and single-γ events **indistinguishable**

$\pi^0 \rightarrow \gamma\gamma$ events **indistinguishable** from e if one gamma missing

LAr Time Projection Chamber (LArTPC)

SLAC NATIONAL ACCELERATOR LABORATORY

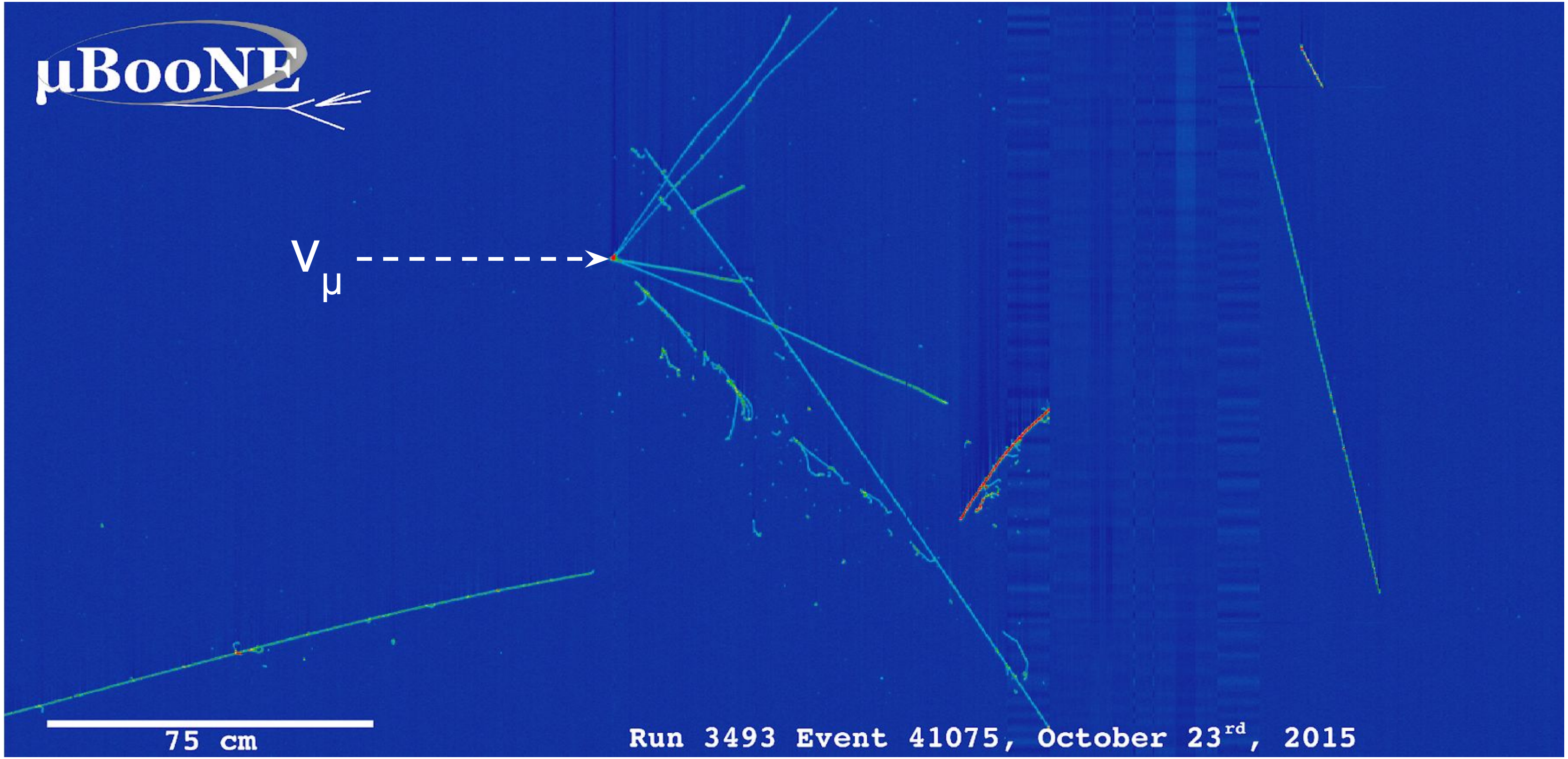


Advantages:

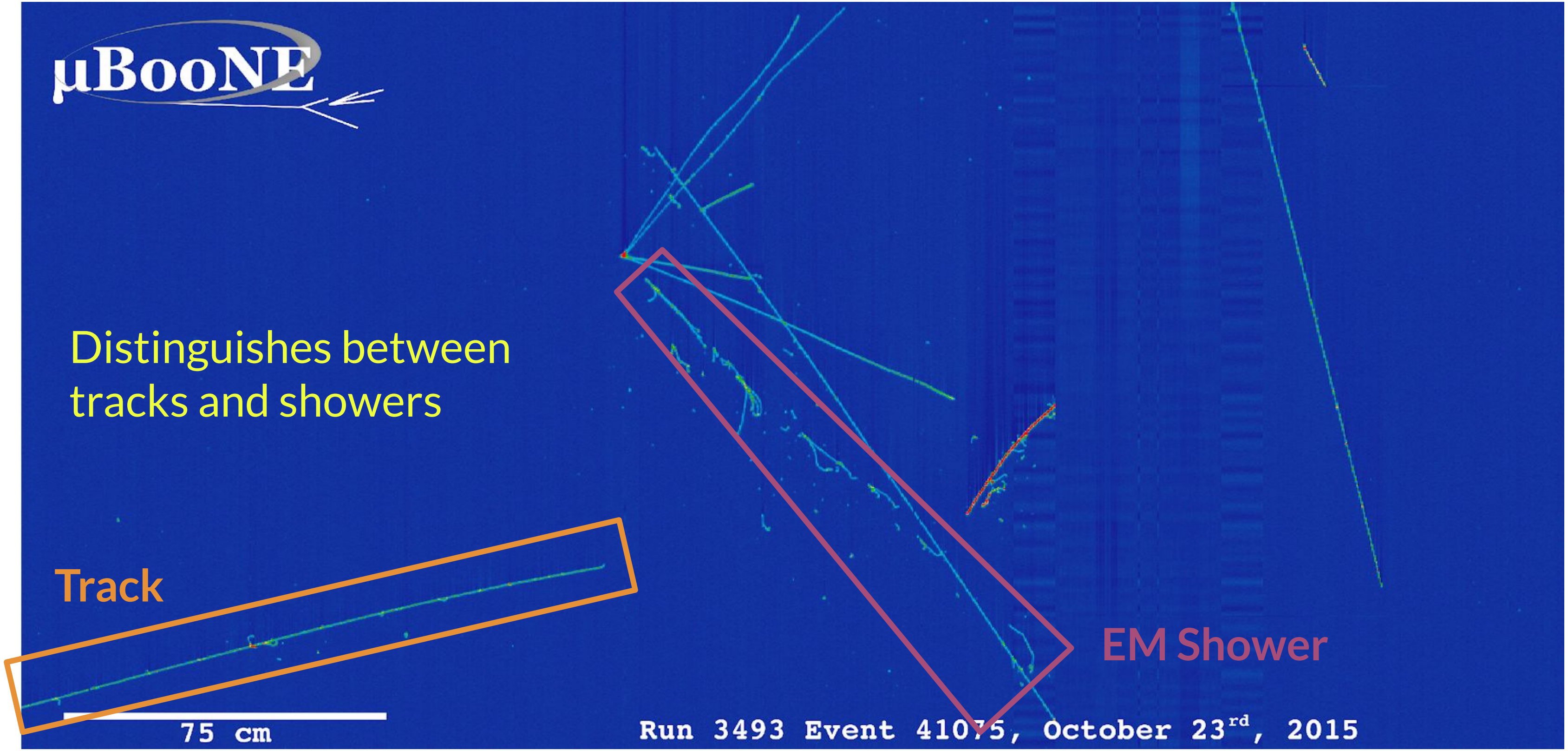
- **Detailed**: mm-scale resolution (wire pitch, readout freq.)
- **Calorimetry**
- **Dense**: high rate of ν interactions
- **Scalable**: detector up to O(10) kt

LArTPCs chosen as **the beam ν detector** in the US for the next 20 years

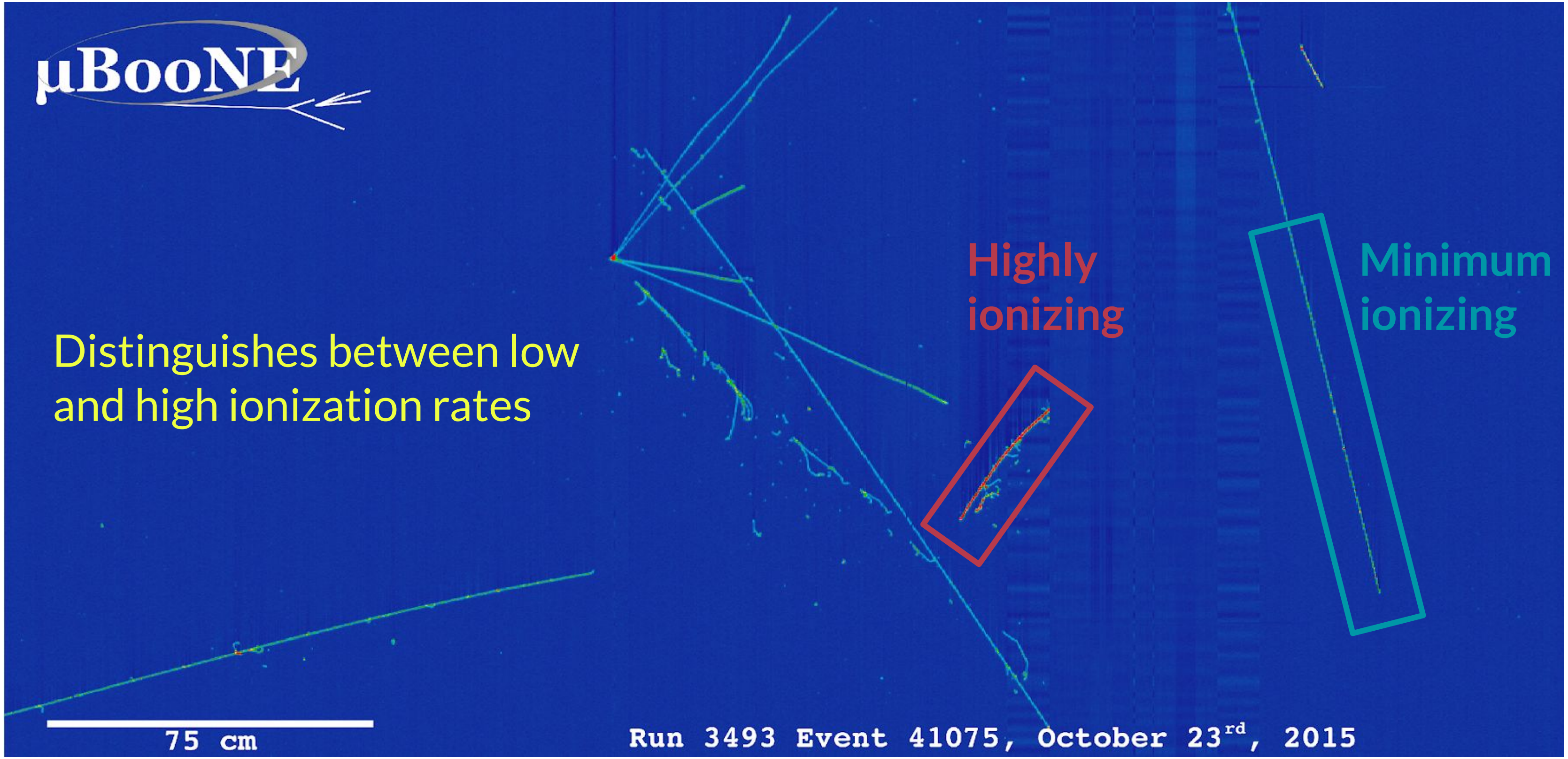
The Resolution Power of LArTPC Images



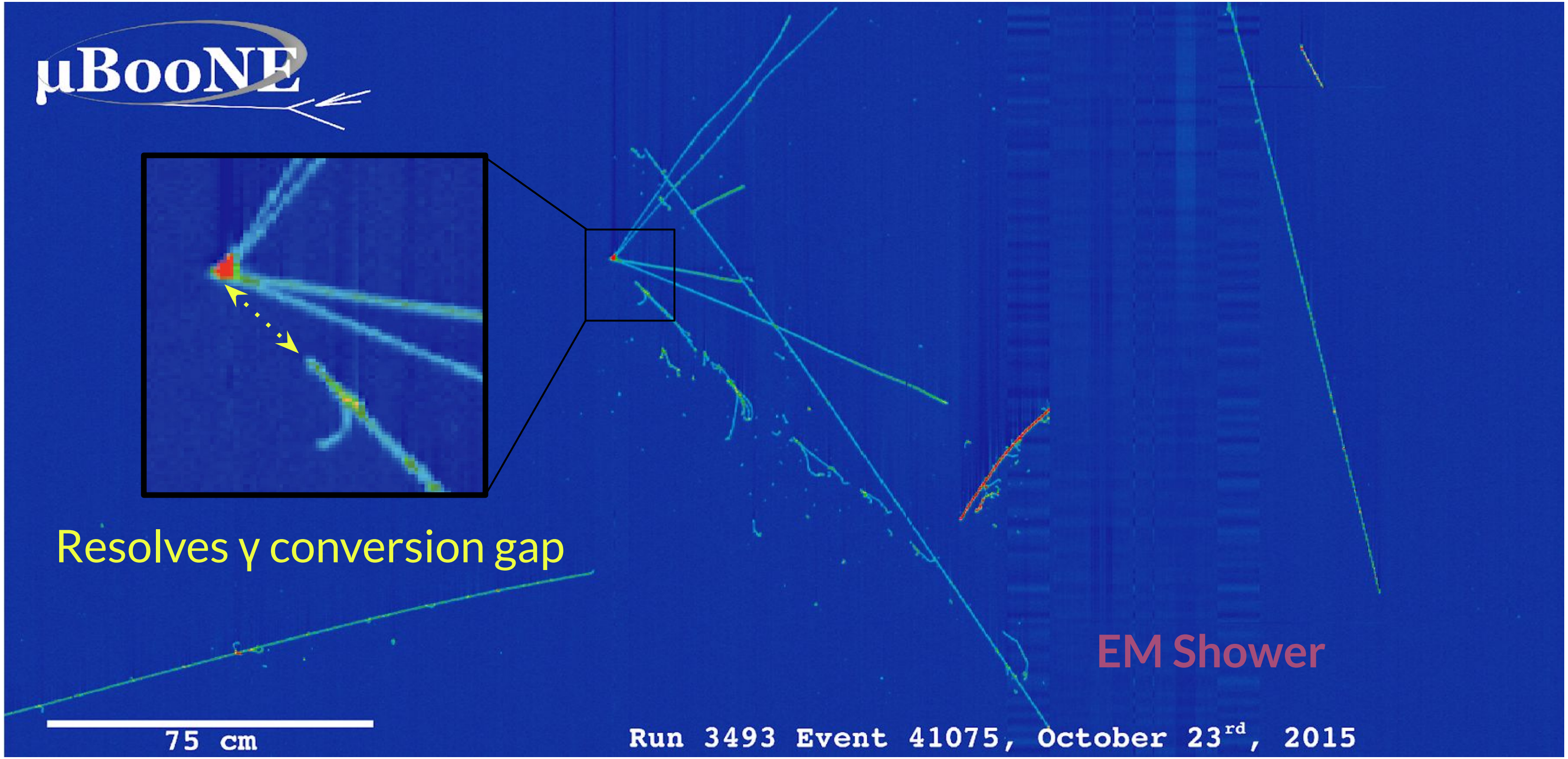
The Resolution Power of LArTPC Images



The Resolution Power of LArTPC Images



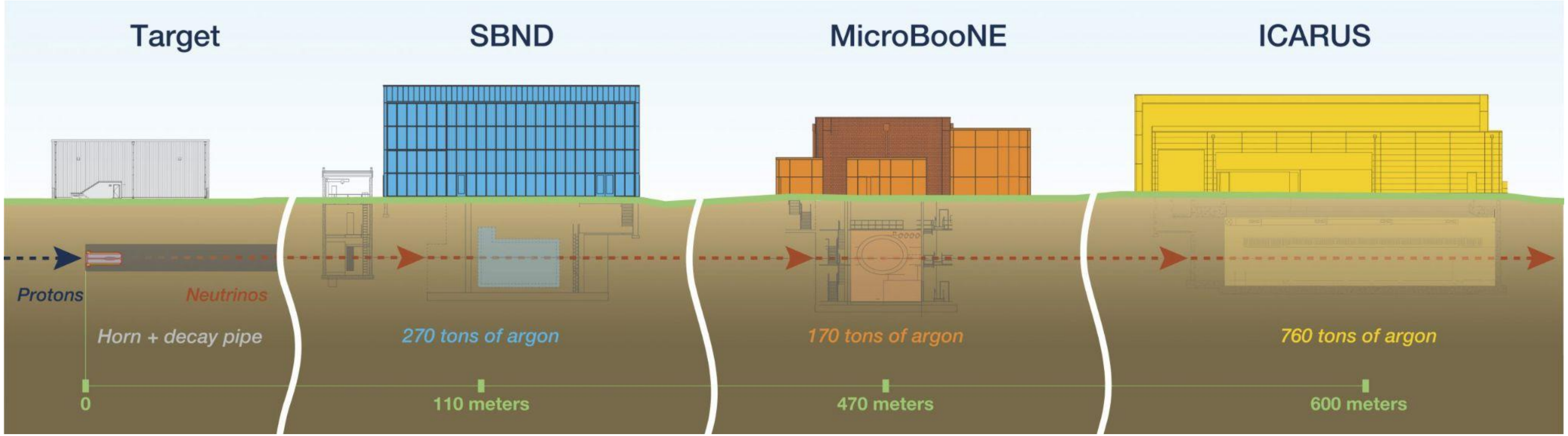
The Resolution Power of LArTPC Images



The Short Baseline Neutrino (SBN) Program

Suite of **three LArTPCs** at three **short baselines**

- Leverages **e-γ separation** power of LArTPCs to resolve MiniBooNE anomaly
- Perfect baseline for hypothetical **short baseline oscillations**
- Large number of ν-Ar events for a **breadth of XS measurements**
- **Technological test bed** for **DUNE**

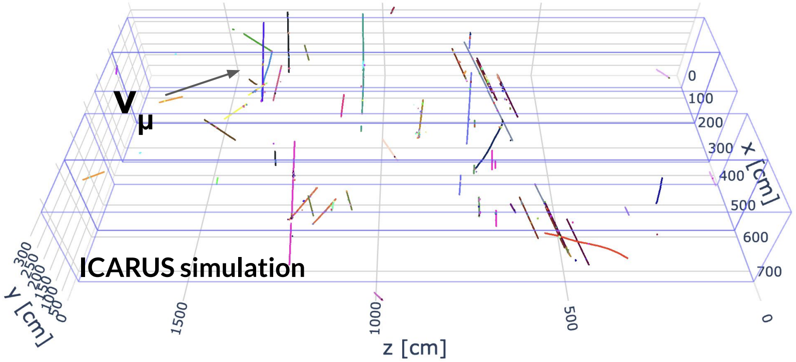


Challenges in LArTPCs

Dense medium → Slow

Electron drift velocity O(1) mm/μs

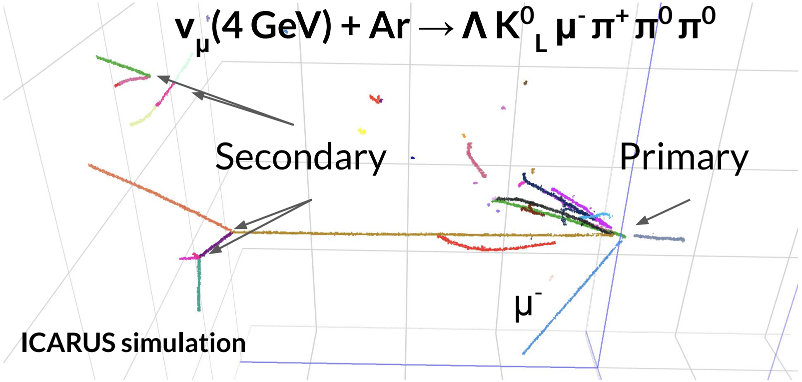
- **Long** (O(1) ms) readout **window**
- Need **light association** for timing



High Z material → Messy

Argon has a large nucleus (Z=18)

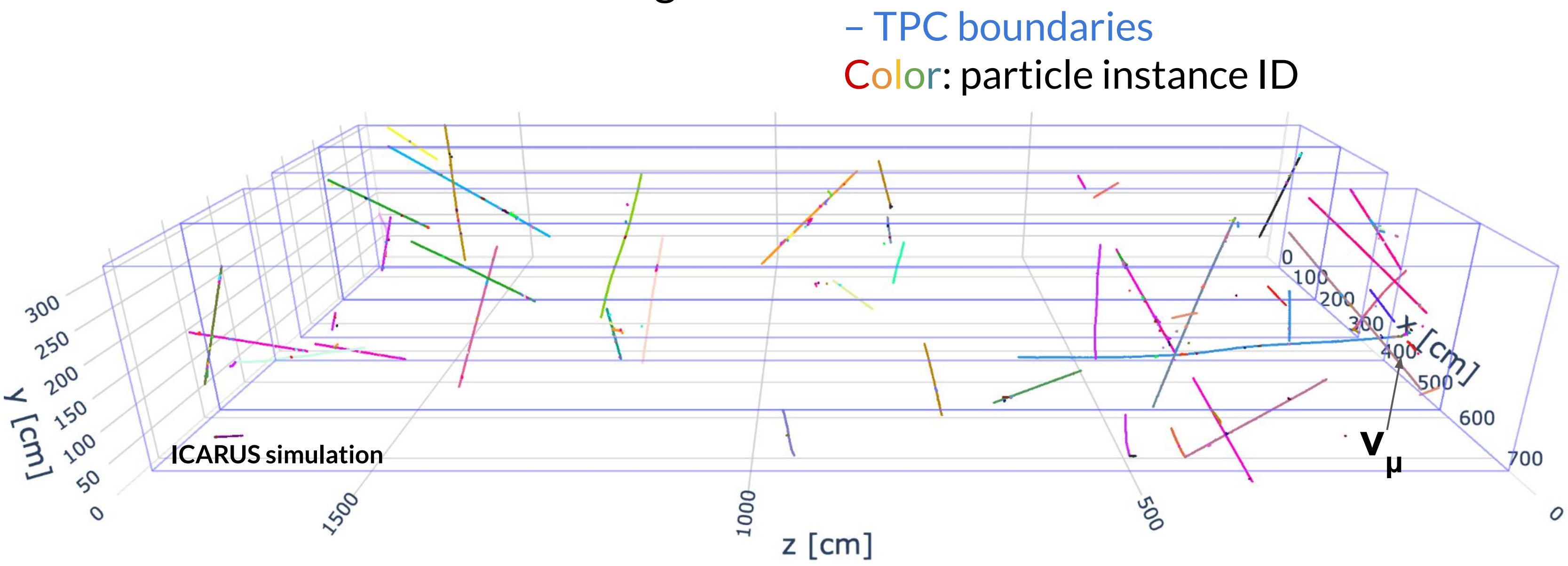
- **Complicated** nuclear physics
- **Secondary** interactions



LArTPC Simulation Test Case

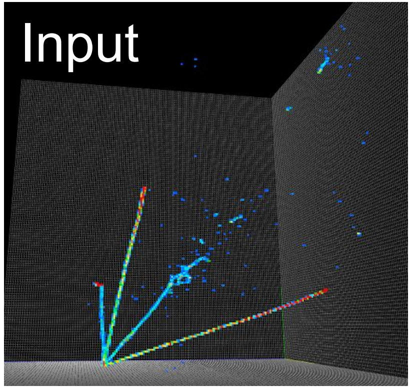
Realistic **BNB ν_μ + Cosmic** ICARUS simulation as a **benchmark**

- **One ν_μ** + Ar interaction/image
- **~25 cosmic** interactions/image



Hierarchical Feature Extraction

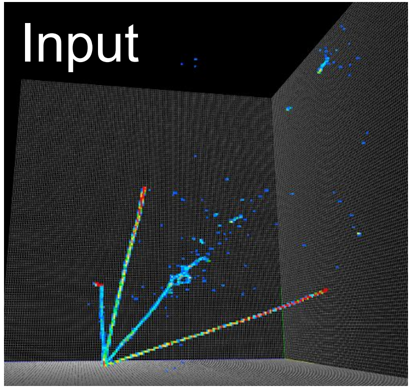
What is relevant to pattern recognition in a detailed interaction image?

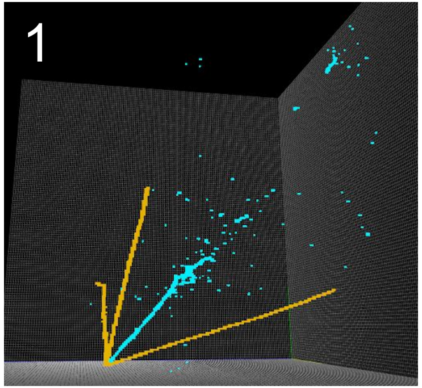


Hierarchical Feature Extraction

What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**

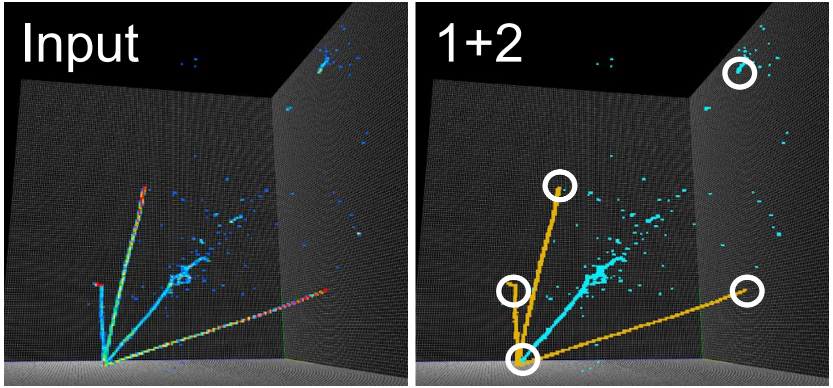




Hierarchical Feature Extraction

What is relevant to pattern recognition in a detailed interaction image?

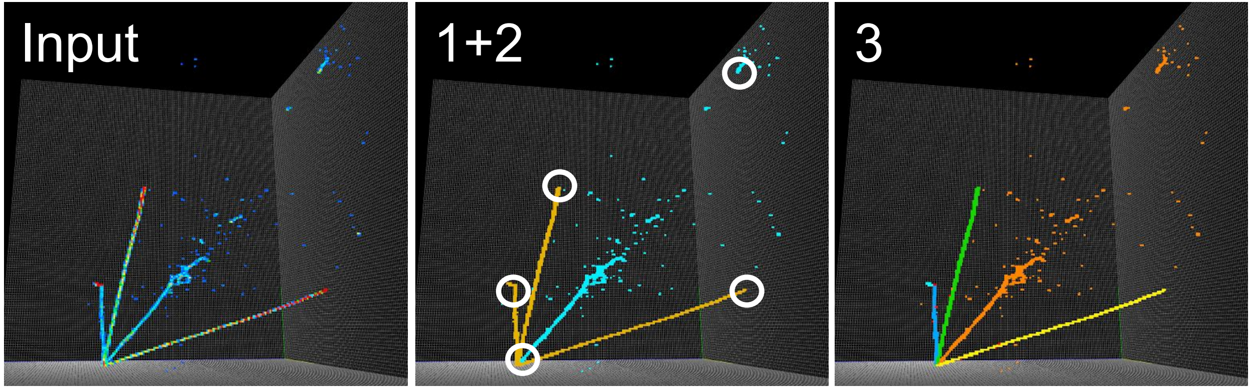
1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)



Hierarchical Feature Extraction

What is relevant to pattern recognition in a detailed interaction image?

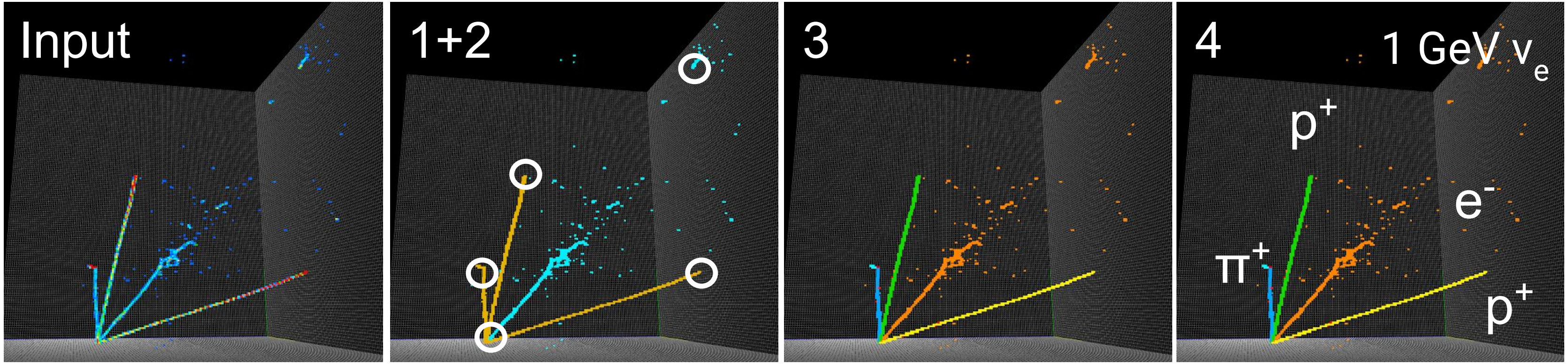
1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)
3. Cluster individual **particles** (tracks and full showers)



Hierarchical Feature Extraction

What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)
3. Cluster individual **particles** (tracks and full showers)
4. Cluster **interactions**, identify **particle properties** in context



Hierarchical Feature Extraction

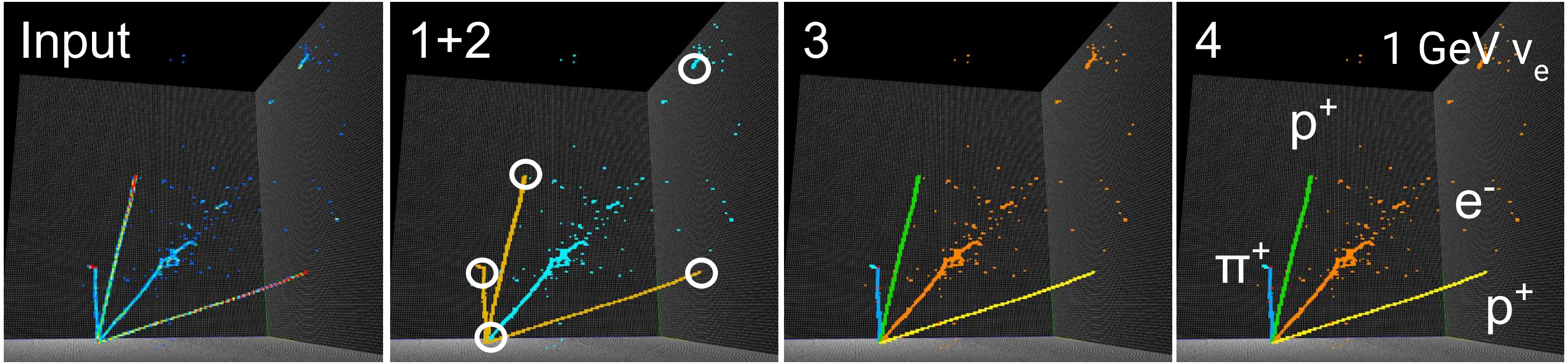
What is relevant to pattern recognition in a detailed interaction image?

1. Separate topologically distinguishable **types of activity**
2. Identify **important points** (vertex, start points, end points)

→ **Pixel-level**

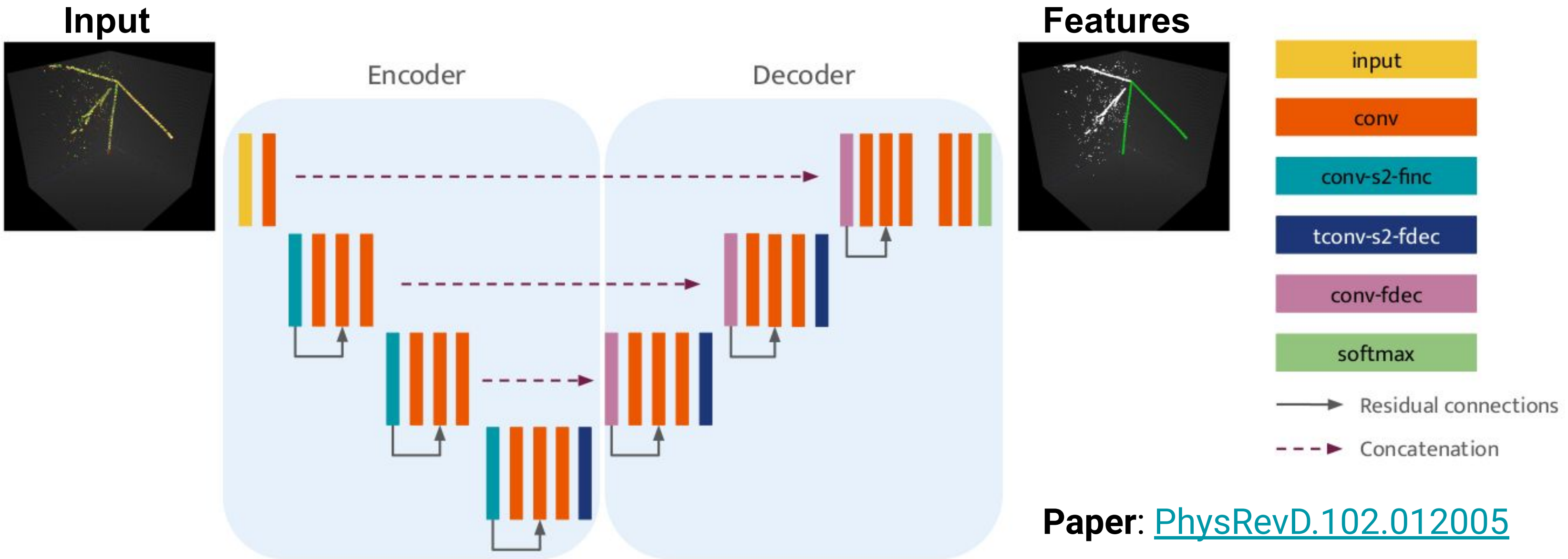
3. Cluster individual **particles** (tracks and full showers)
4. Cluster **interactions**, identify **particle properties** in context

→ **Cluster-level**



Pixel-Level Feature Extraction

UResNet (UNet + ResNet + Sparse Conv.) as the **backbone feature extractor**

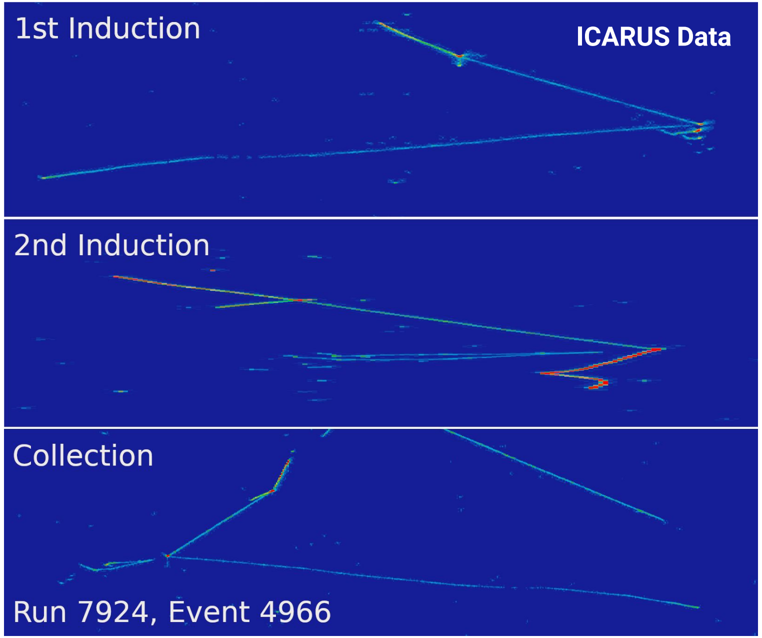


Paper: PhysRevD.102.012005

Tomographic Reconstruction

In a **wire TPC**, we do not get 3D images, but rather 3 x 2D projections

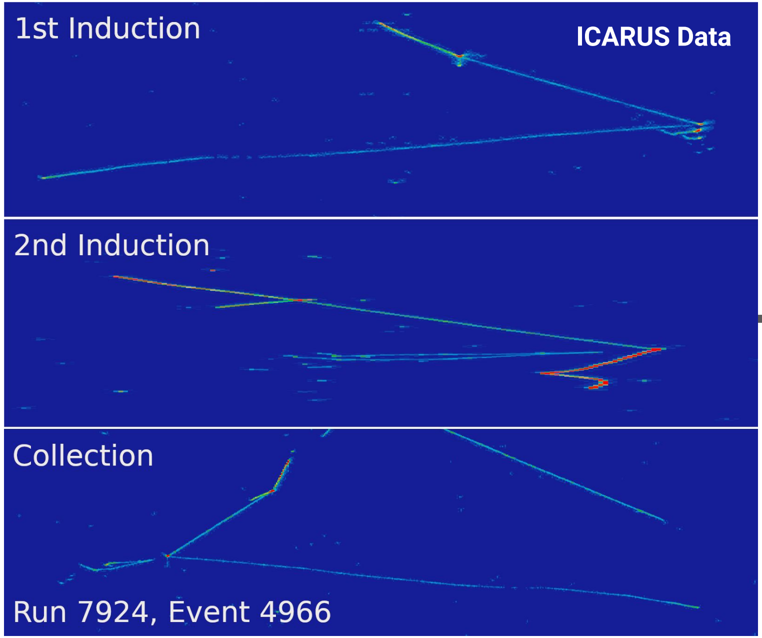
- First task: **combine projections into one 3D image**



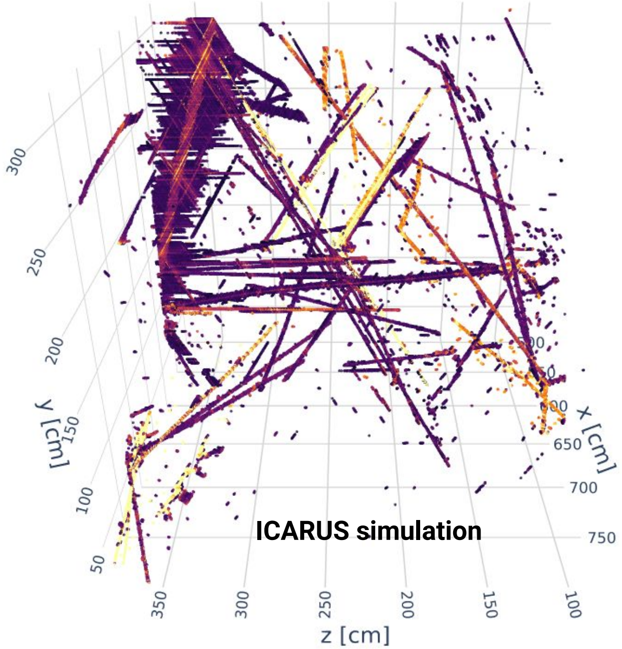
Tomographic Reconstruction

In a **wire TPC**, we do not get 3D images, but rather 3 x 2D projections

- First task: **combine projections into one 3D image**



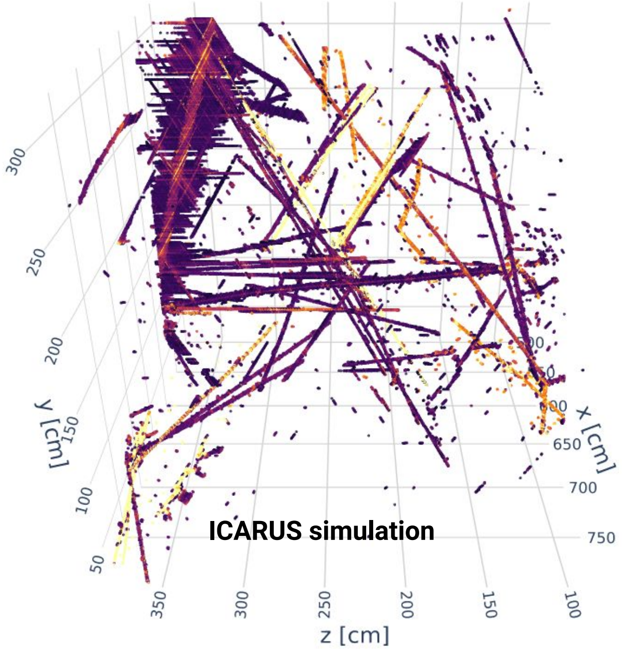
Make all valid combinations of 2 plane hits



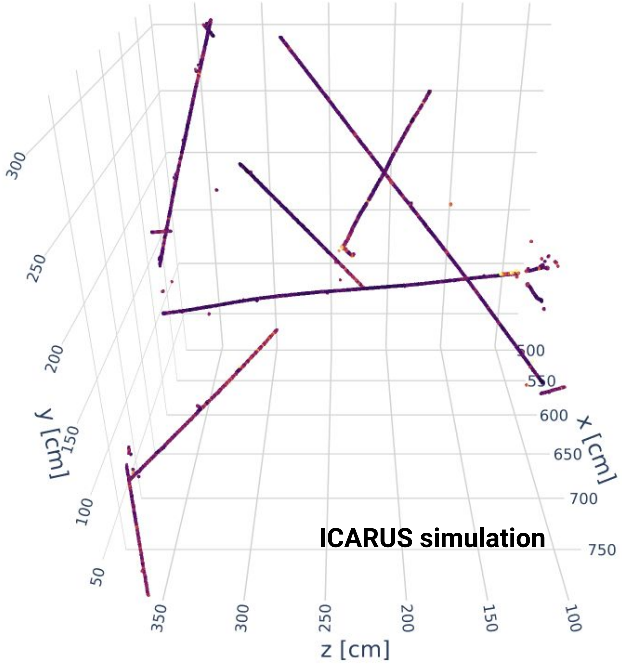
Tomographic Reconstruction

In a **wire TPC**, we do not get 3D images, but rather 3 x 2D projections

- First task: **combine projections into one 3D image**



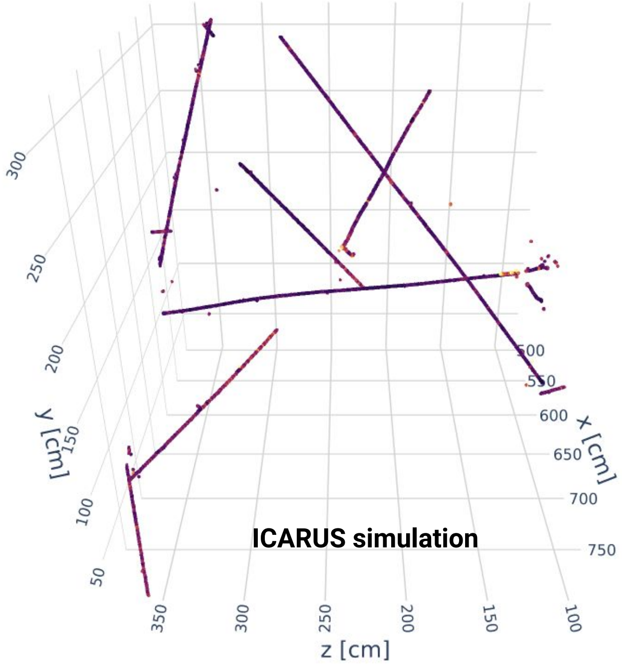
Remove artifacts with UResNet →

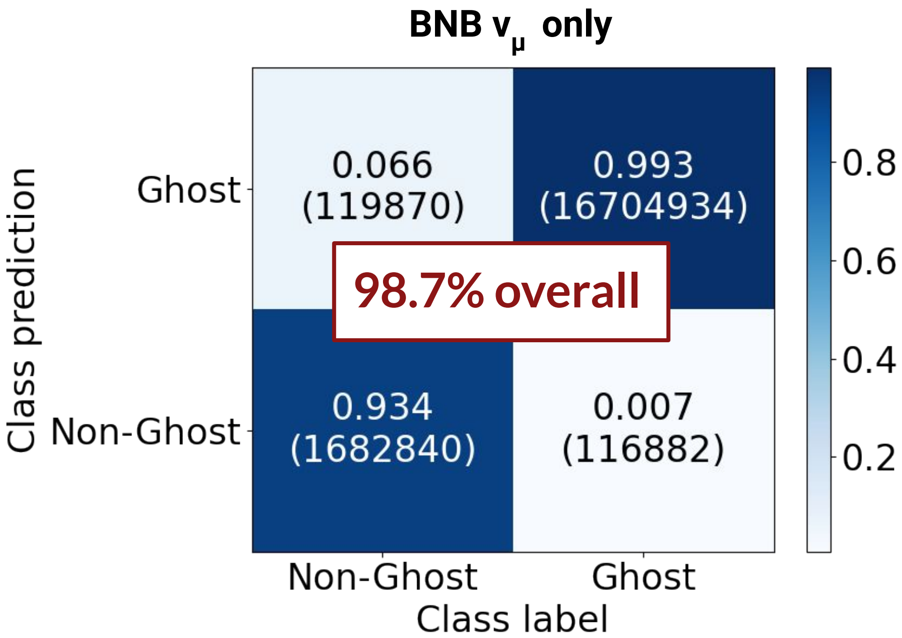


Tomographic Reconstruction

In a **wire TPC**, we do not get 3D images, but rather 3 x 2D projections

- First task: **combine projections into one 3D image**



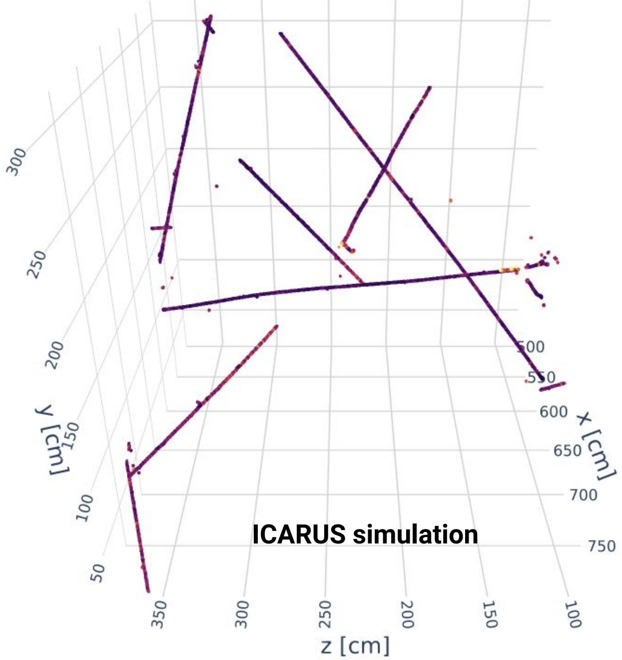


Semantic Segmentation

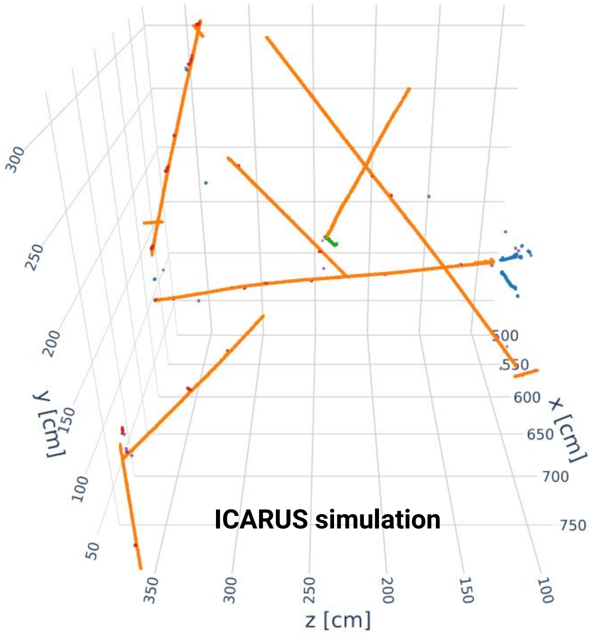
SLAC NATIONAL ACCELERATOR LABORATORY

Separate **topologically different** types of activity

- Tracks, Showers, delta rays, Michel electrons, low energy blips



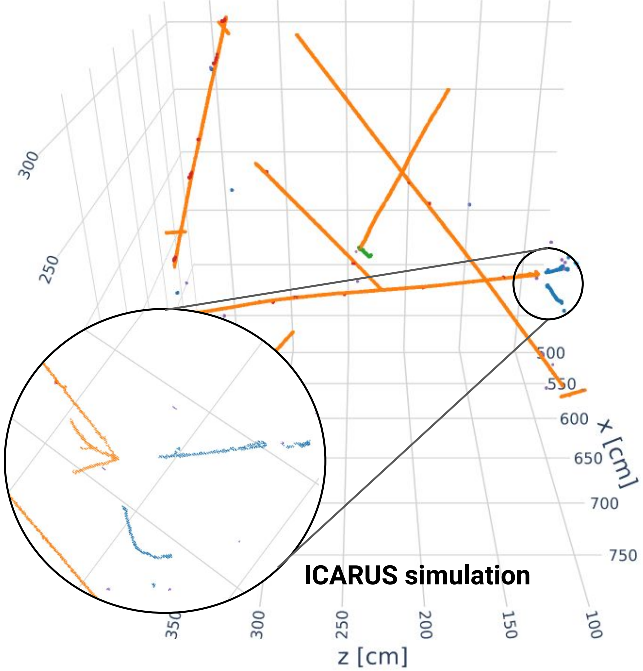
Classify voxels into categories with UResNet



Semantic Segmentation

Separate **topologically different** types of activity

- Tracks, Showers, delta rays, Michel electrons, low energy blips



BNB ν_μ only

Class prediction \ Class label	Shower	Track	Michel	Delta	LE
LE	0.011	0.000	0.001	0.001	0.850
Delta	0.001	0.001	0.007	0.851	0.001
Michel	0.003	0.000	0.877	0.005	0.001
Track	0.005	0.997	0.038	0.121	0.055
Shower	0.981	0.002	0.077	0.023	0.094

Colour scale: 0.2, 0.4, 0.6, 0.8

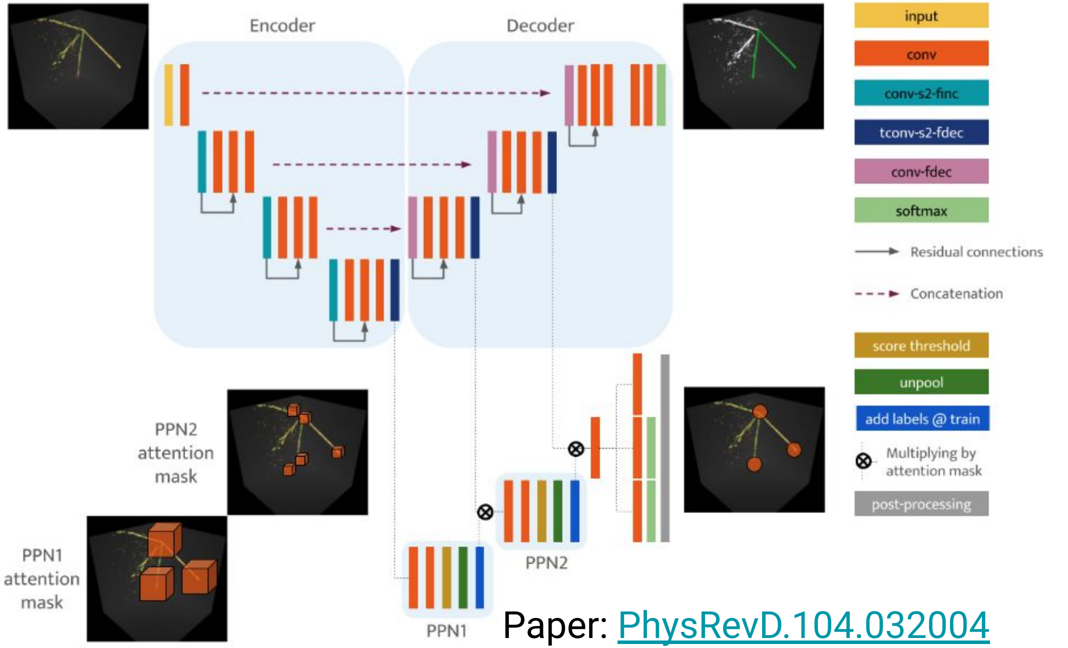
99.1% overall

Paper: PhysRevD.102.012005

Points of Interest

The Point Proposal Network (PPN) uses decoder features

- Three CCN layers to narrow ROI
- Last layer reconstructs:
  - Relative position to voxel center of active voxel
  - Point type
- Post-processing aggregates nearby points

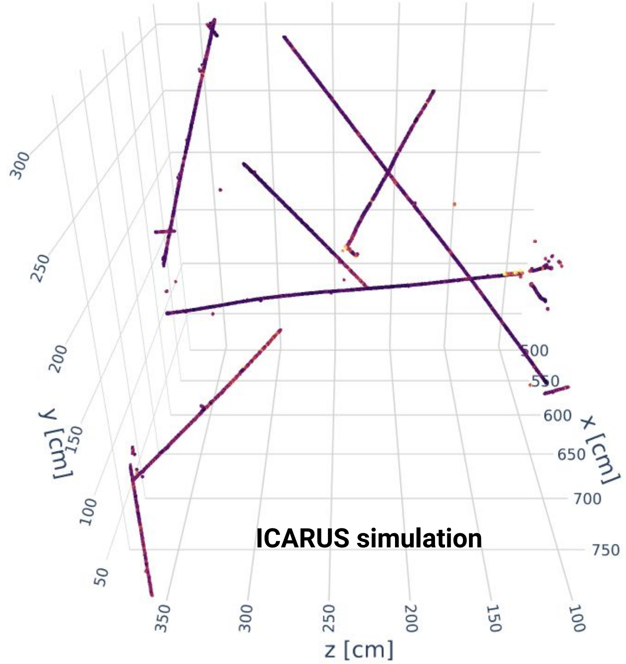


Paper: PhysRevD.104.032004

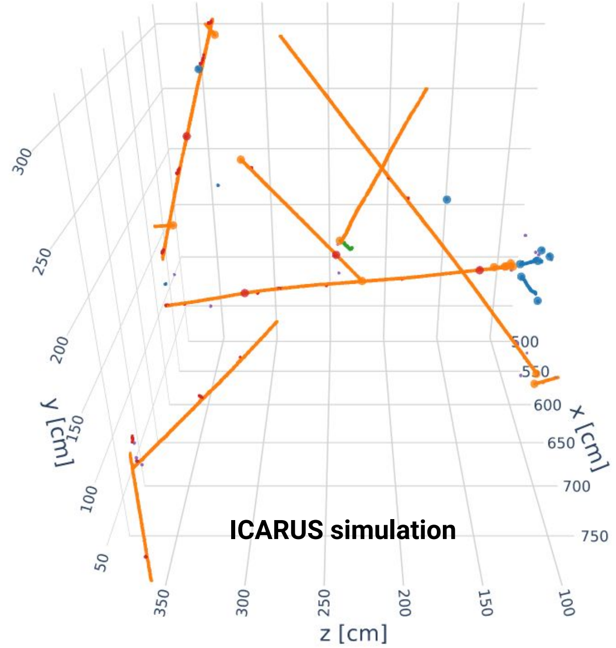
Points of Interest

Narrow down a region proposal all the way to a **point**

- Predict **masks** at different scales with UResNet, predict **position** in voxel



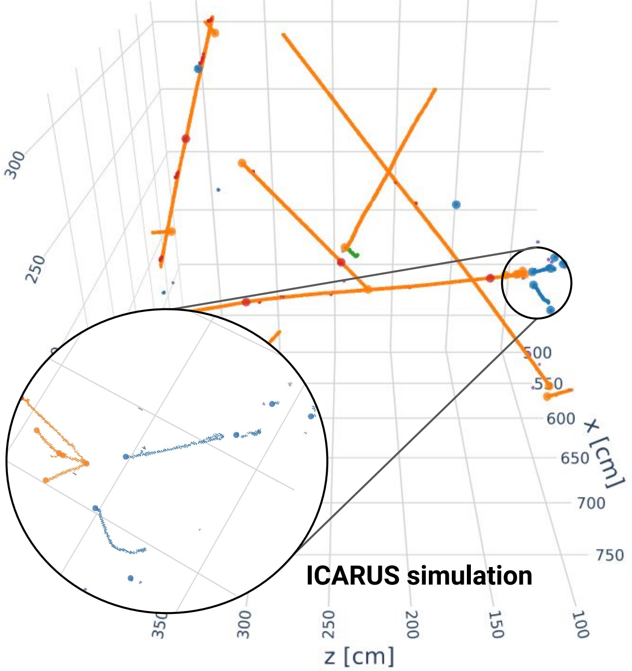
Identify end points of tracks, shower starts



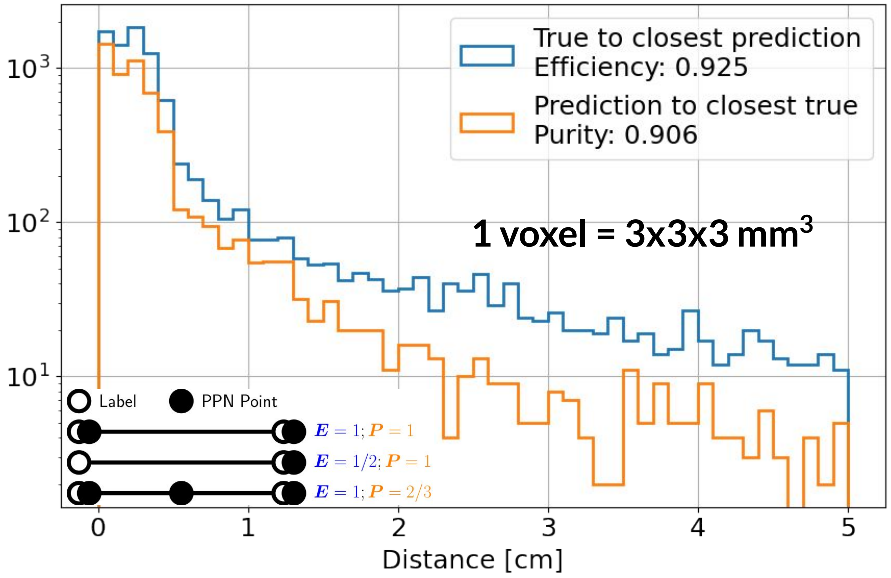
Points of Interest

Narrow down a region proposal all the way to a **point**

- Predict **masks** at different scales with UResNet, predict **position** in voxel



BNB ν_μ only

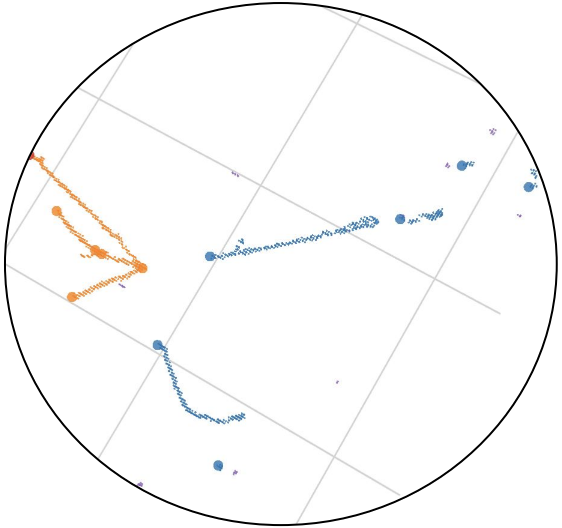


Paper: PhysRevD.104.032004

Clustering Strategy

At this point, we must **do away with pixel-level predictions**

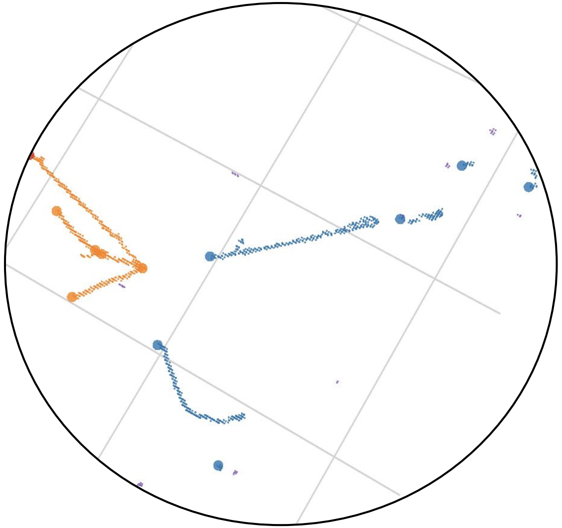
- Number of target clusters: **unknown**
- Cluster label: **non-unique** (permutation-invariant)



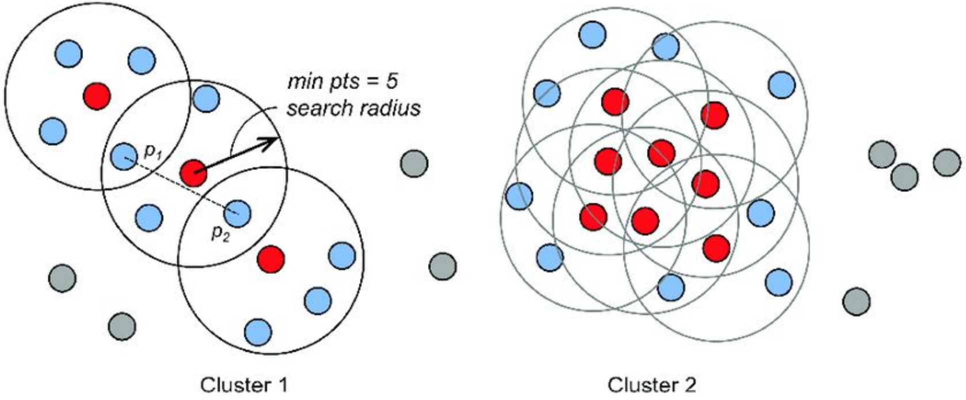
Clustering Strategy

At this point, must **do away with pixel-level predictions**

- Number of target clusters: **unknown**
- Cluster label: **non-unique** (permutation-invariant)



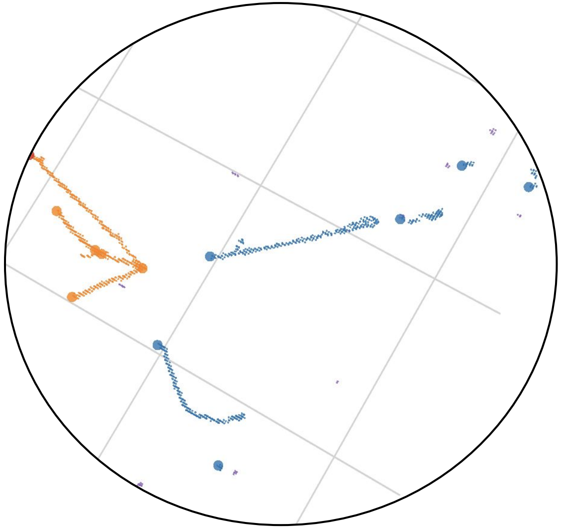
Density based clustering (DBSCAN)?



Clustering Strategy

At this point, must **do away with pixel-level predictions**

- Number of target clusters: **unknown**
- Cluster label: **non-unique** (permutation-invariant)



Density based clustering (DBSCAN)? Yes, but…

1. How to break tracks?

   → **Dense problem**

Spatial Embedding

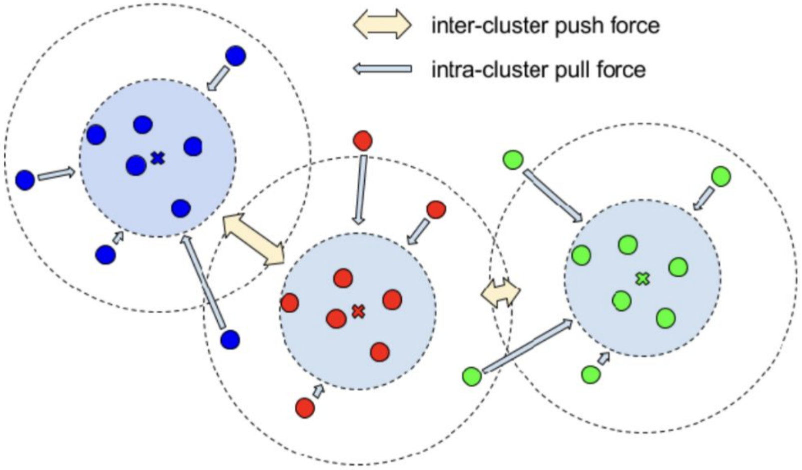
First: learn a **transformation** to a **separable space**

$$L = \alpha L_{var} + \beta L_{dist} + \gamma L_{reg},$$

$$L_{var} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{i=1}^{N_c}[\max(0, \|\mu_c - x_i\| - \delta_v)]^2$$

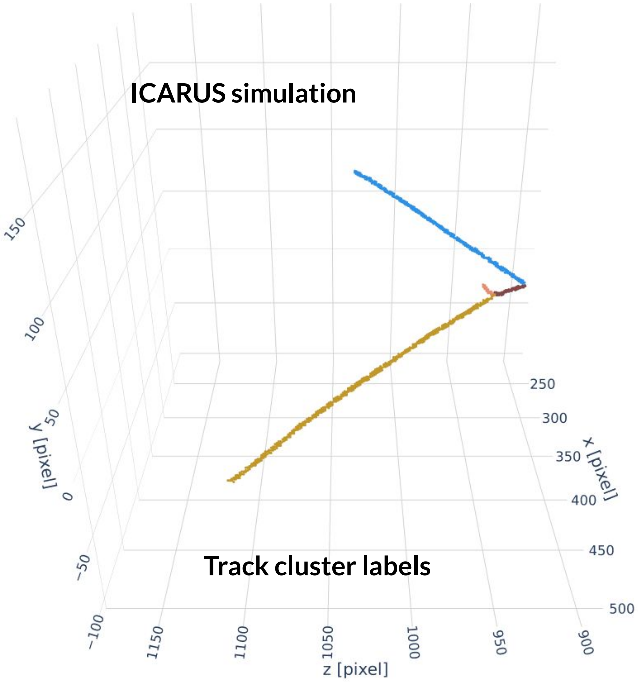
$$L_{dist} = \frac{1}{C(C-1)}\sum_{\substack{c_A, c_B = 1\\ c_A \neq c_B}}^{C}[\max(0, 2\delta_d - \|\mu_{c_A} - \mu_{c_B}\|)]^2$$

$$L_{reg} = \frac{1}{C}\sum_{c=1}^{C}\|\mu_c\|$$



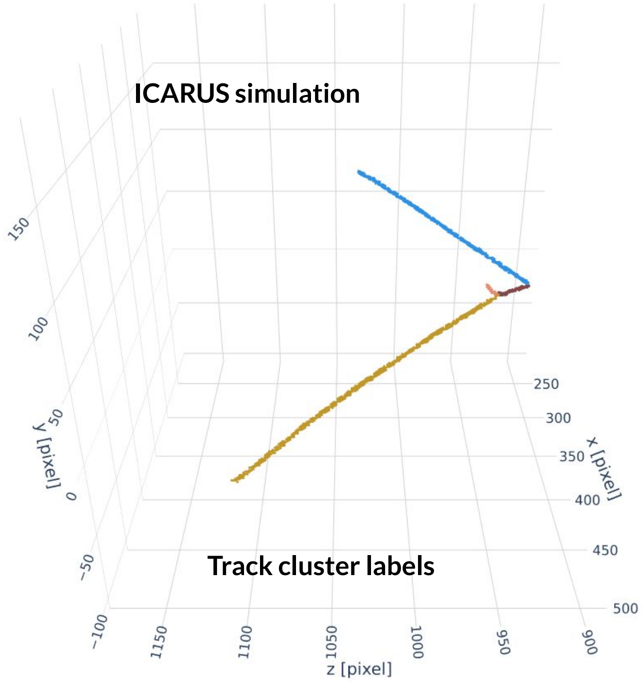
Spatial Embedding

First: learn a transformation to a **separable space**

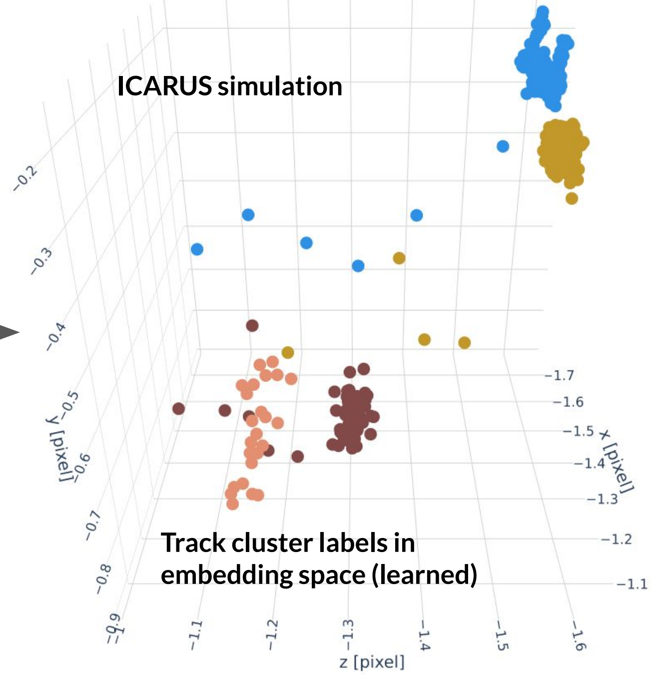


Spatial Embedding

First: **learn** a transformation to a **separable space**

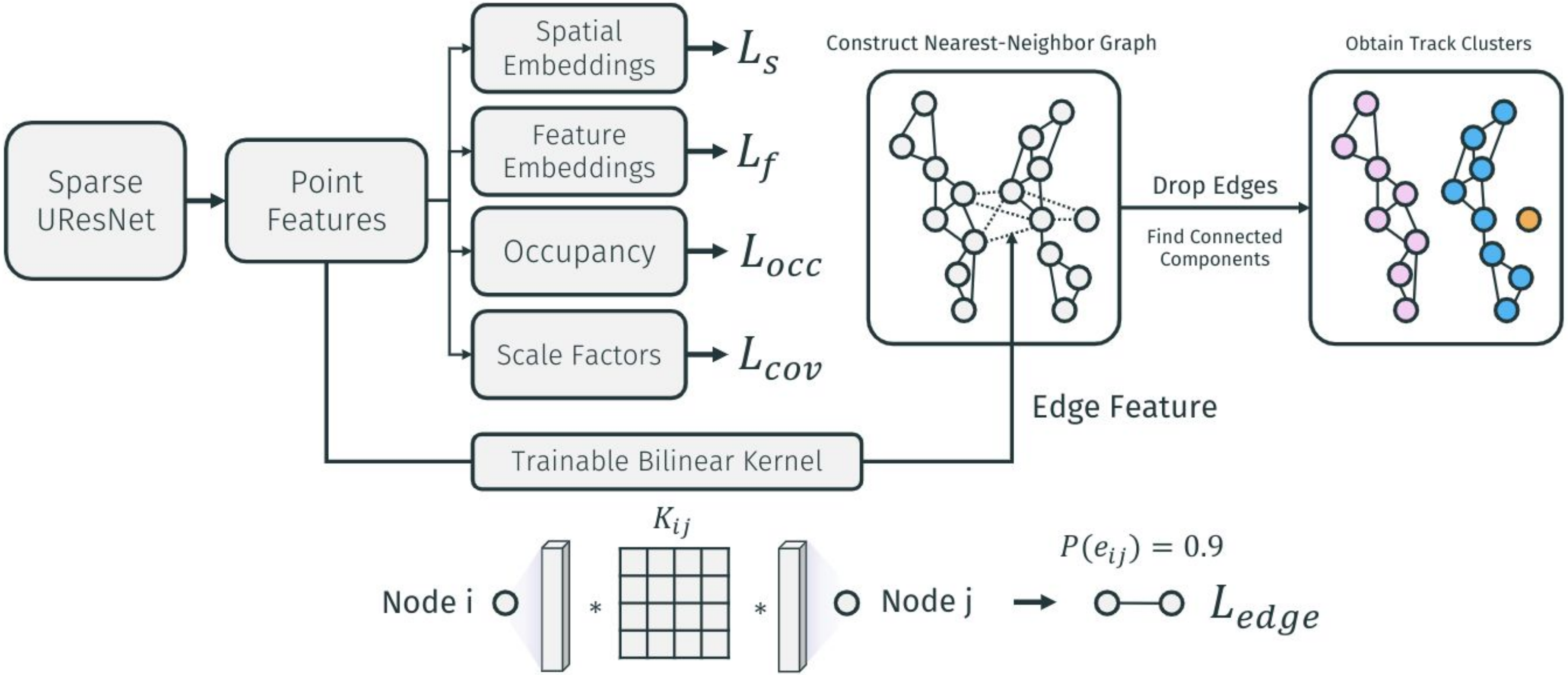


Map to embedded space with UResNet



Graph-SPICE

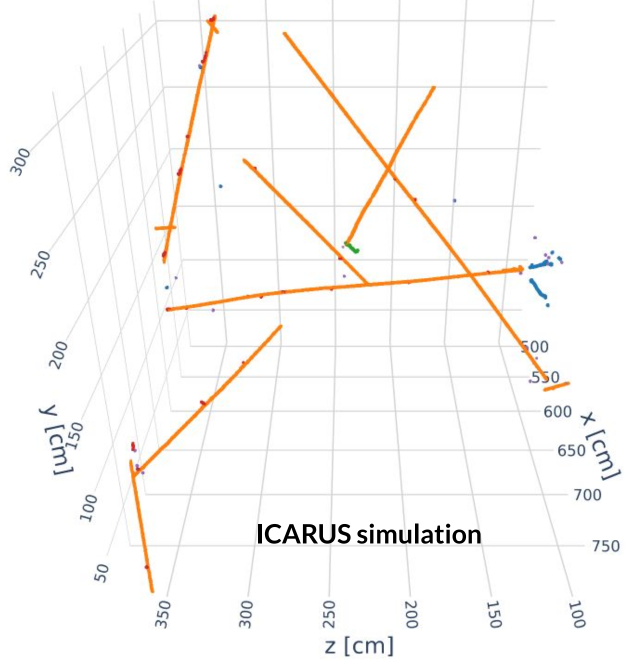
Second: learn a **smart** version of **DBSCAN** (connected components)



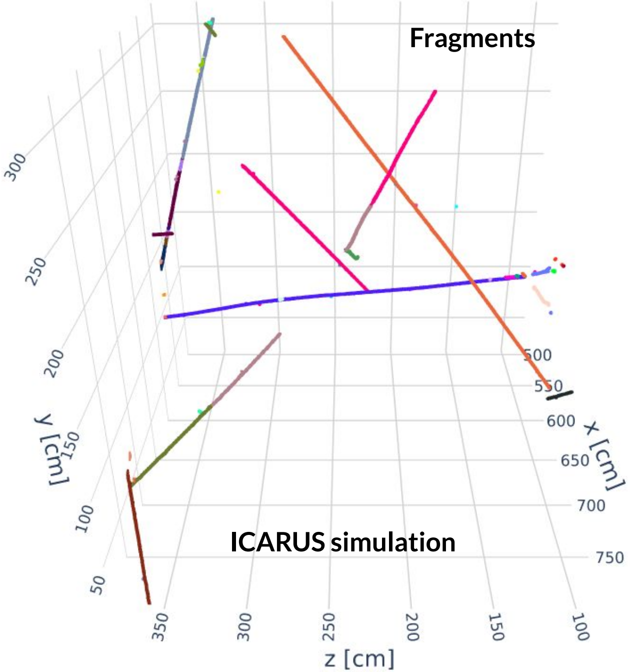
Dense Fragment Formation

Break track/shower fragment instances **where they touch**

- Cluster track/shower **fragments** at this stage



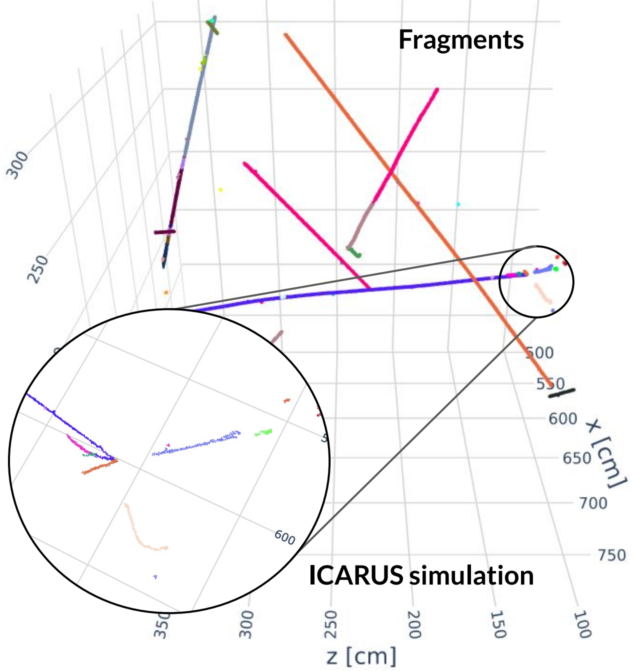
Classify voxels into categories with UResNet

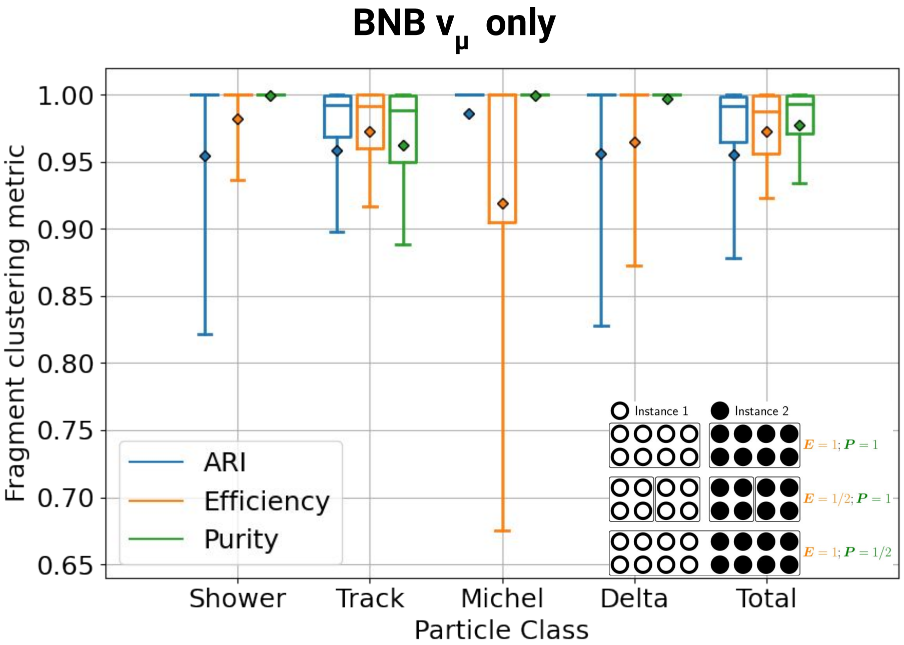


Dense Fragment Formation

Transform coordinates to an space in which tracks are spatially separated

- Cluster track/shower **fragments** at this stage



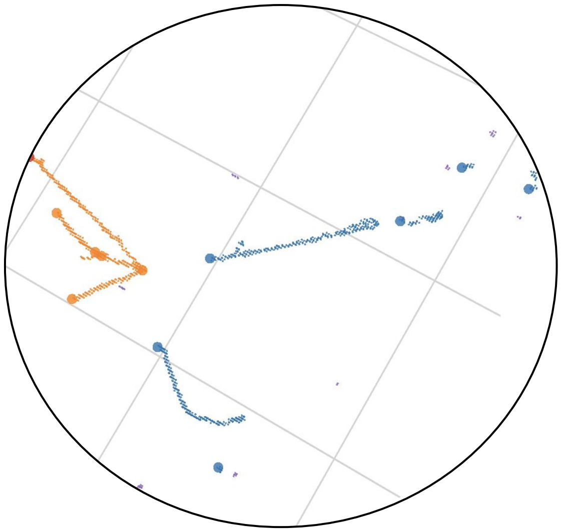


Paper: arXiv:2007.03083

Clustering Strategy

At this point, must **do away with pixel-level predictions**

- Number of target clusters: **unknown**
- Cluster label: **non-unique** (permutation-invariant)



Density based clustering (DBSCAN)? Yes, but…

1. How to break tracks?

   ~~→ **Dense problem**~~ ✓

2. How to aggregate shower fragments and broken up track fragments?

   → **Aggregation problem**

Fragment Graph Representation

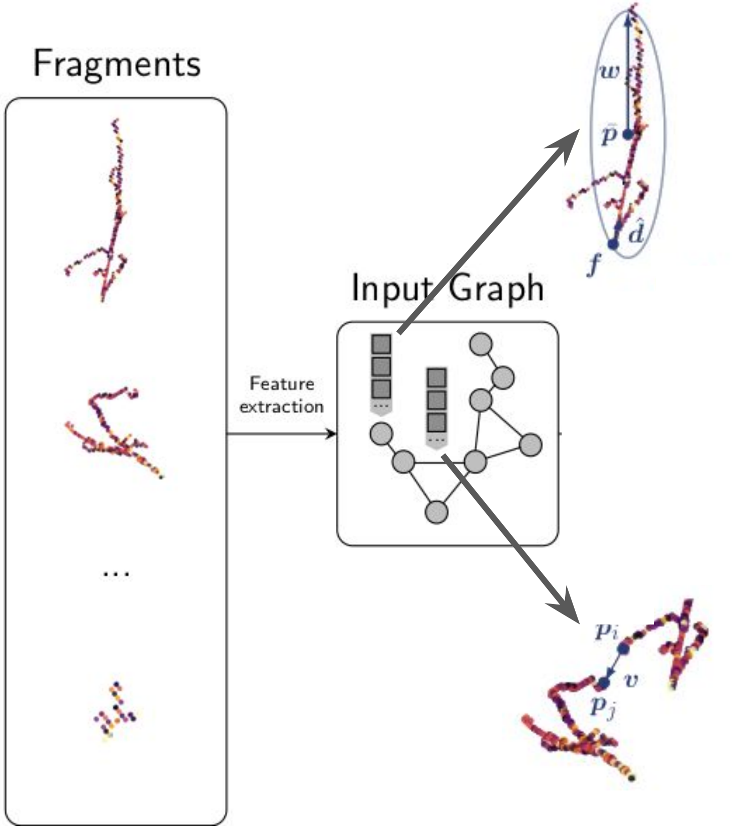
We now represent the set of fragments as a **set of nodes in a graph** where **edges represent correlations**

Node features:

- Centroid
- Covariance matrix
- Start point/direction
- . . .

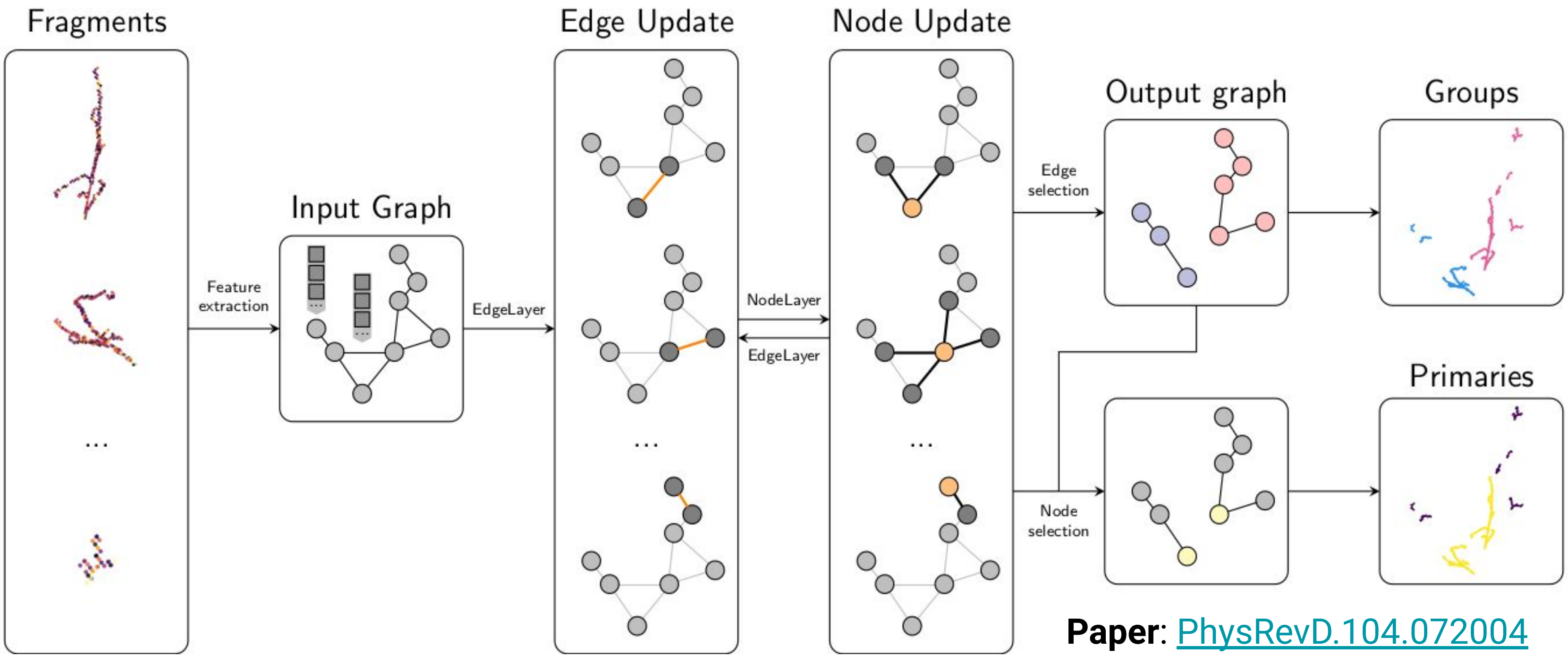
Edge features:

- Displacement vector
- . . .



Aggregation

Graph Neural Network: develop features useful to node+edge classification

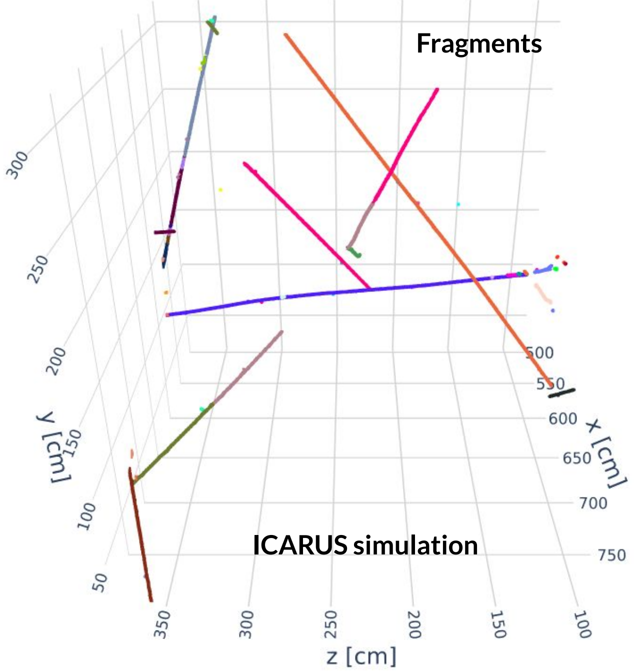


Paper: PhysRevD.104.072004

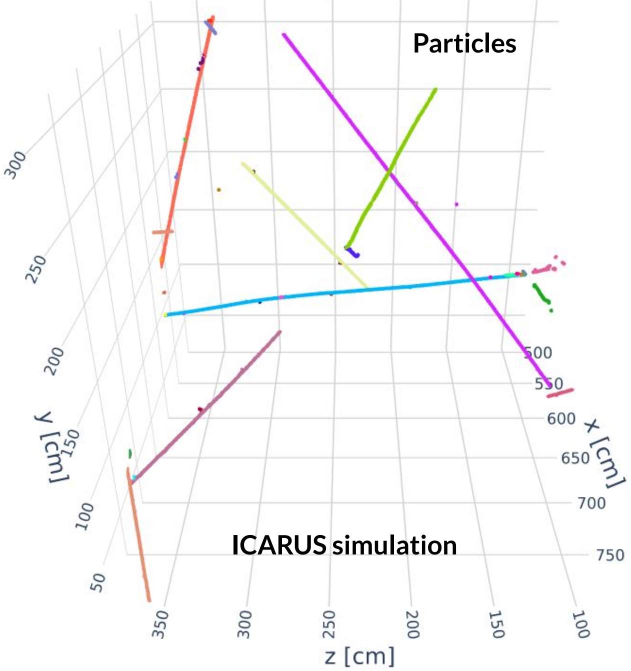
Particle Aggregation

Aggregate track/shower fragment instances into **particles**

- Find edges that connect **fragments that belong together**



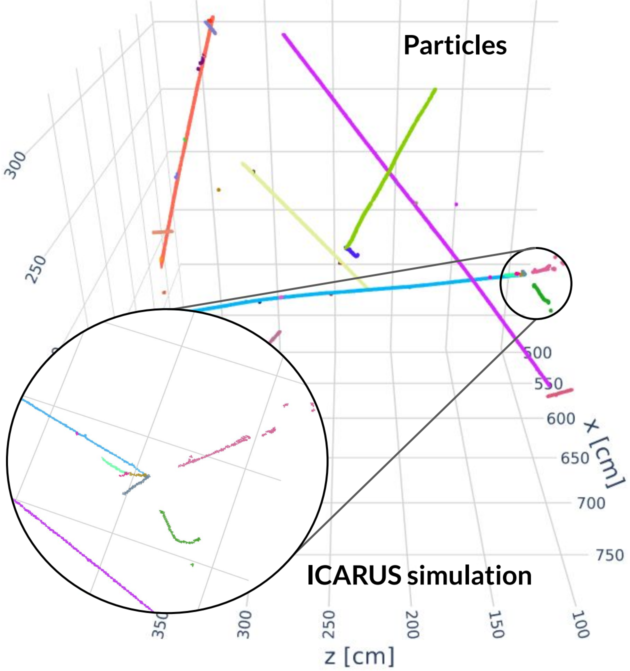
Classify voxels into categories with UResNet

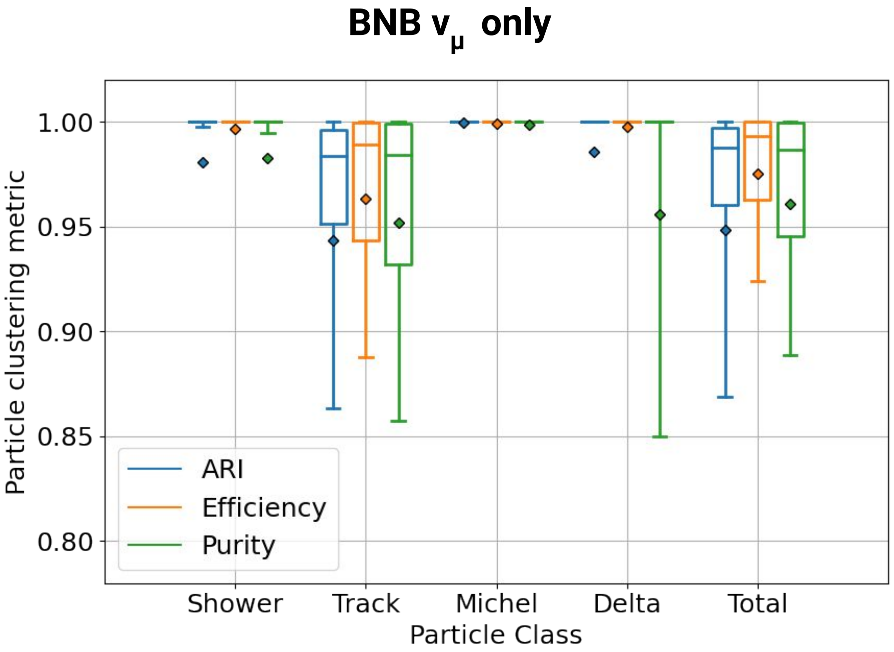


Particle Aggregation

Aggregate track/shower fragment instances into **particles**

- Find edges that connect **fragments that belong together**



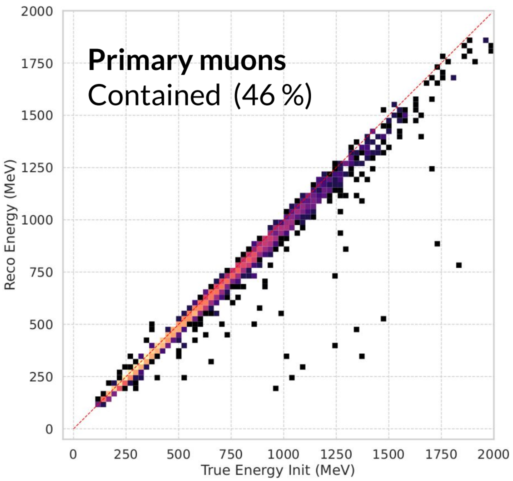


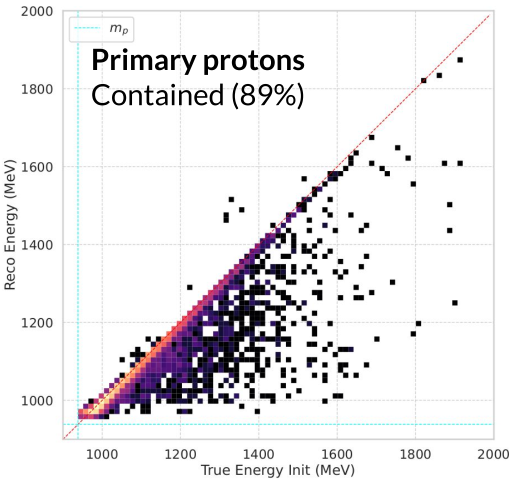
Paper: PhysRevD.104.072004

Particle energy reconstruction

Currently using **traditional techniques** for particle **energy reconstruction**:

- Range-based energy reconstruction of muons and protons look good



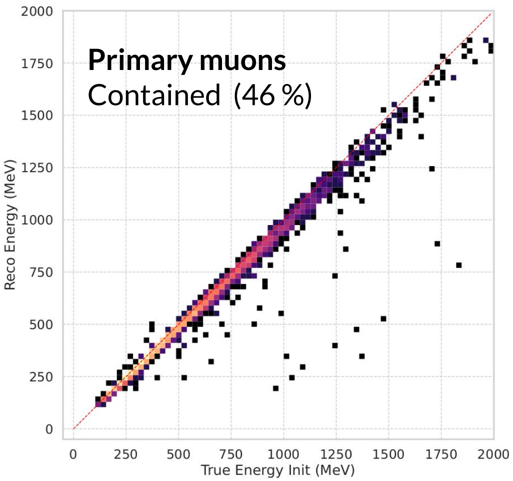


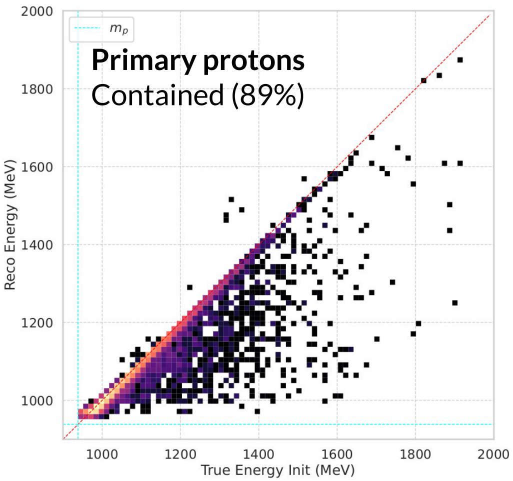
Particle energy reconstruction

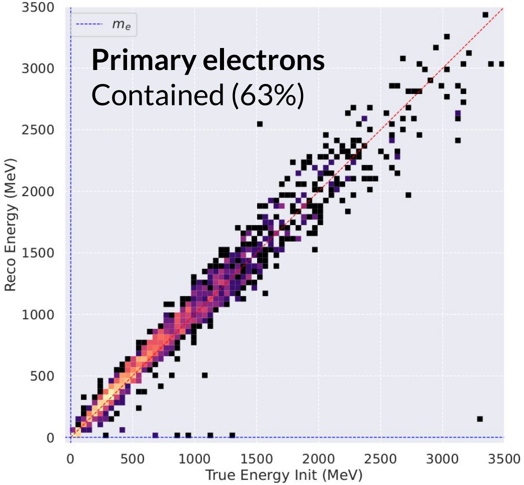
Currently using **traditional techniques** for particle **energy reconstruction**:

- Range-based energy reconstruction of muons and protons look good
- Calorimetric energy reconstruction of electrons also solid

More detail in Dae Heun’s talk this afternoon



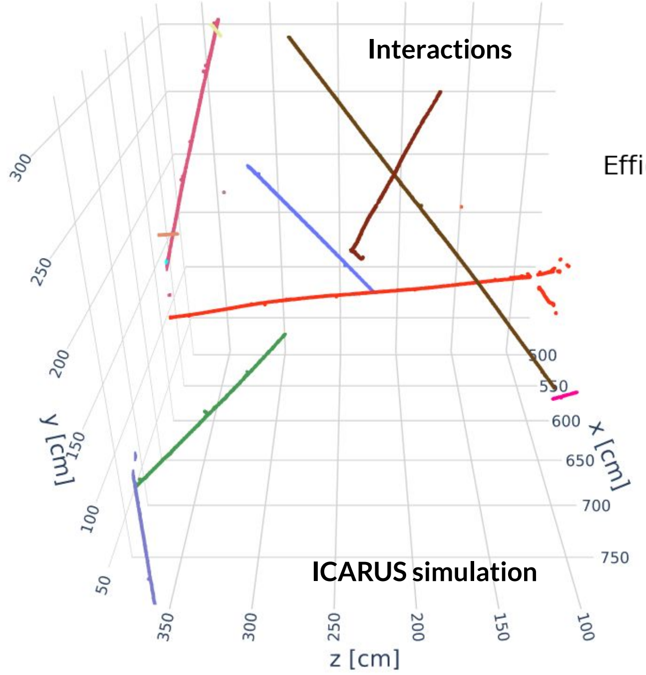


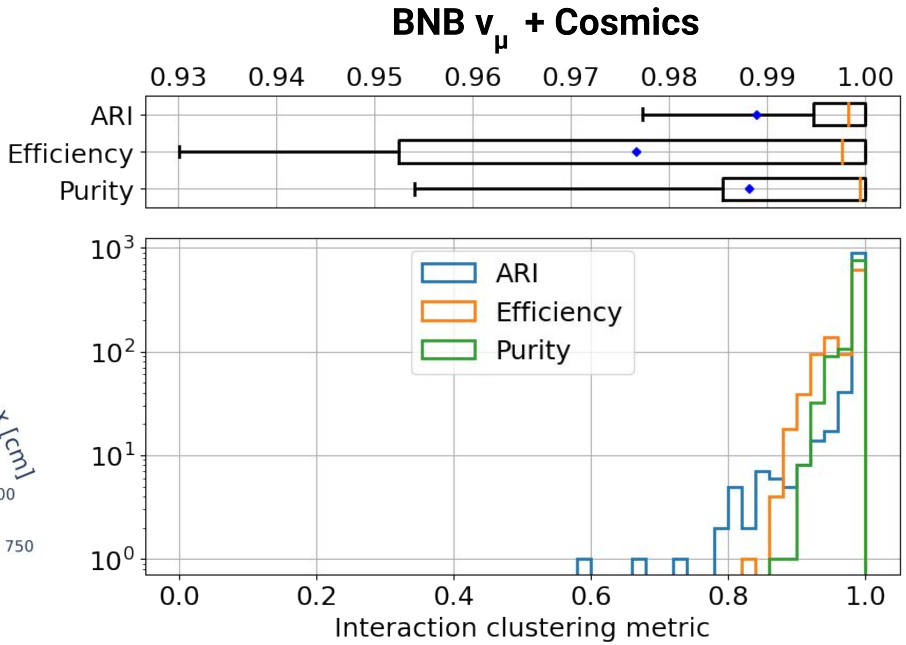


Interaction Aggregation

Aggregate track/shower particle instances into **interactions**

- Find edges that connect ~~fragments~~ **particles that belong together**



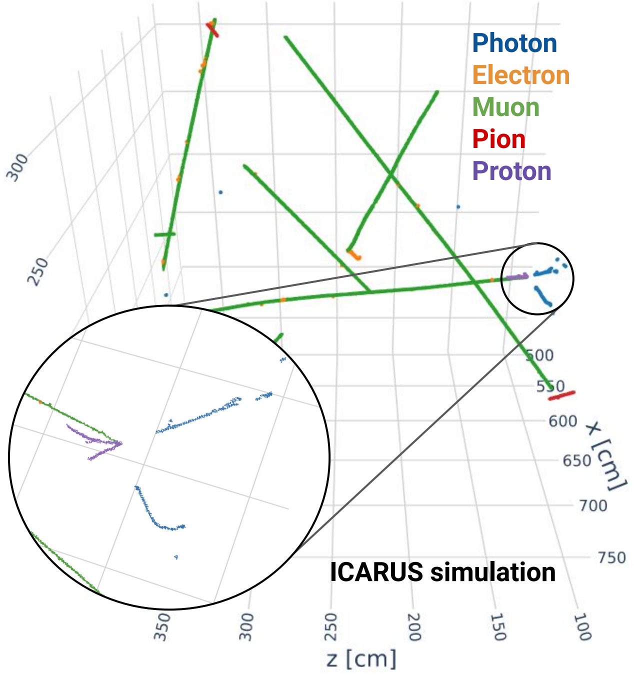


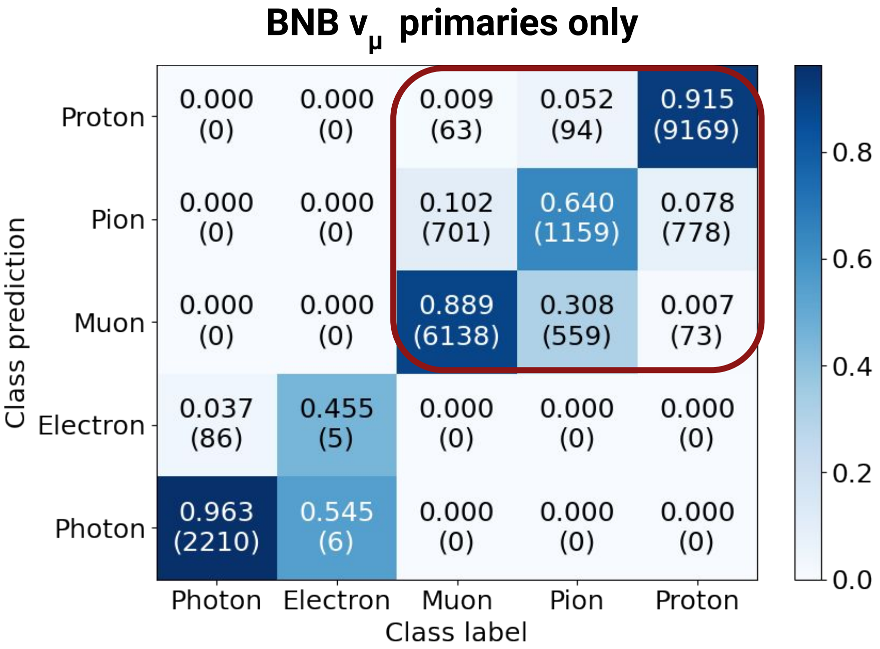
Paper: PhysRevD.104.072004

Particle Identification

Particle species much easier to **infer in context**

- Shower conversion gaps, secondary hadrons, Michel decays, etc.

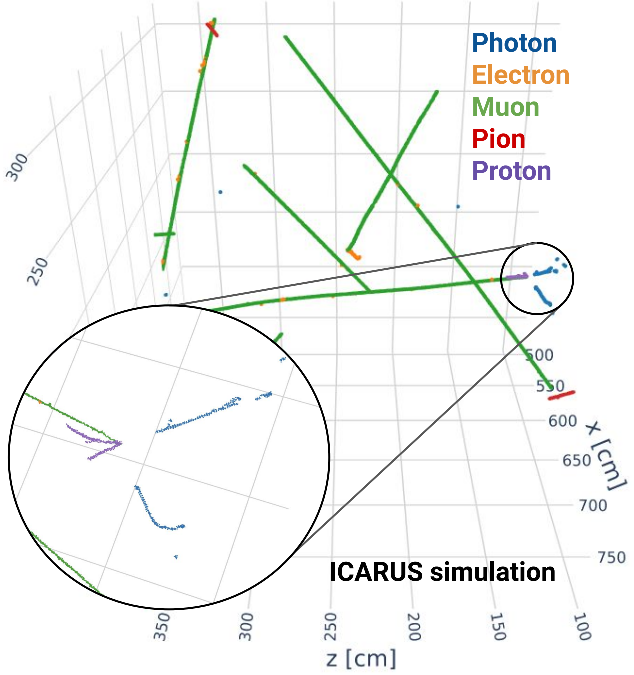


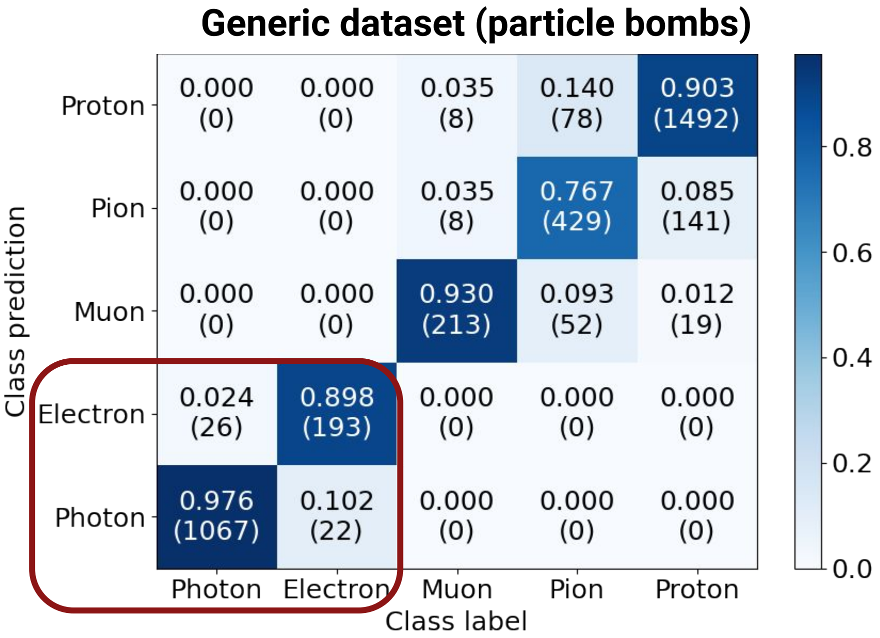


Particle Identification

Particle species much easier to **infer in context**

- Michel decays, secondary hadrons, shower conversion gaps, etc.

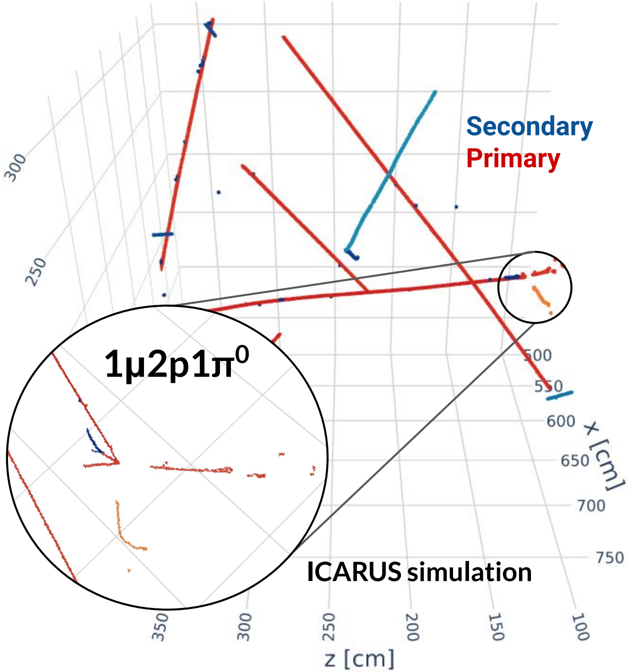


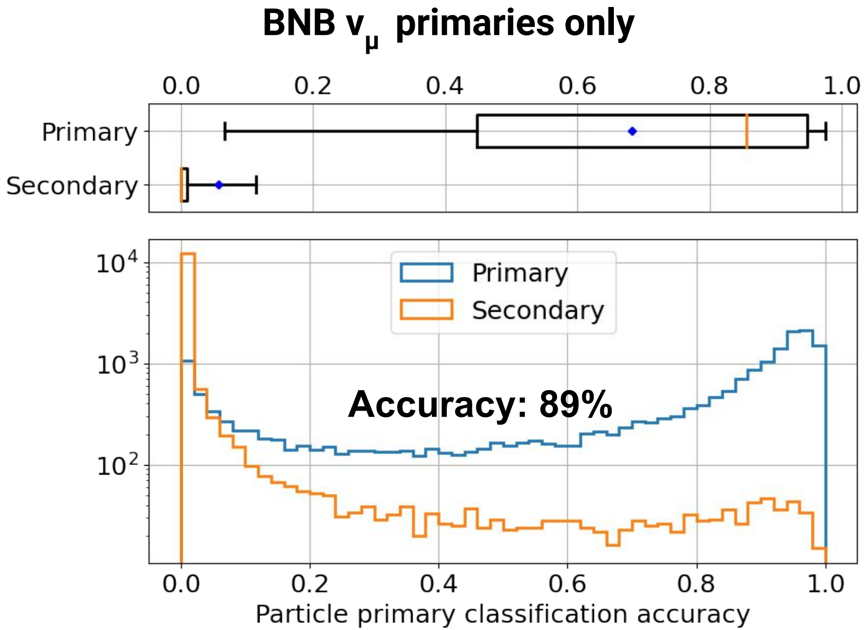


Primary Identification

Important to know **which particle originate from the vertex**

- Central to any **exclusive analysis** (study specific interaction channels)



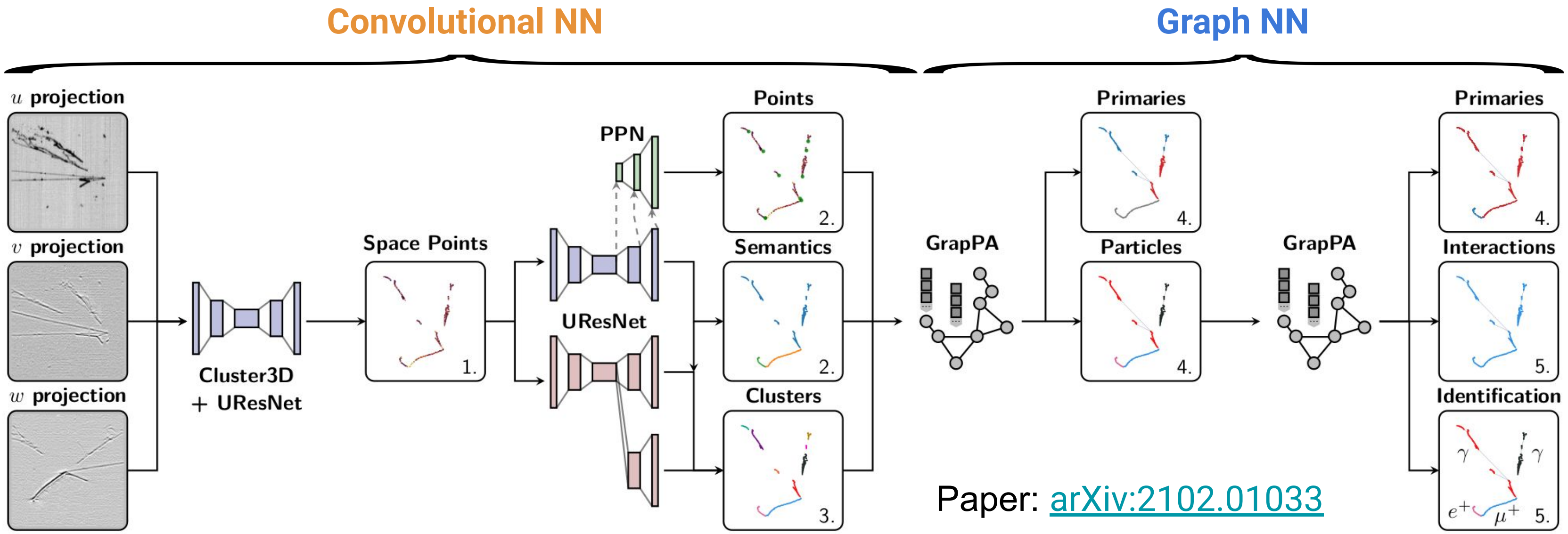


Reconstruction in LArTPCs

SLAC NATIONAL ACCELERATOR LABORATORY

End-to-end ML-based reconstruction chain

- **UResNet** for pixel feature extraction, **GrapPA** for superstructure formation

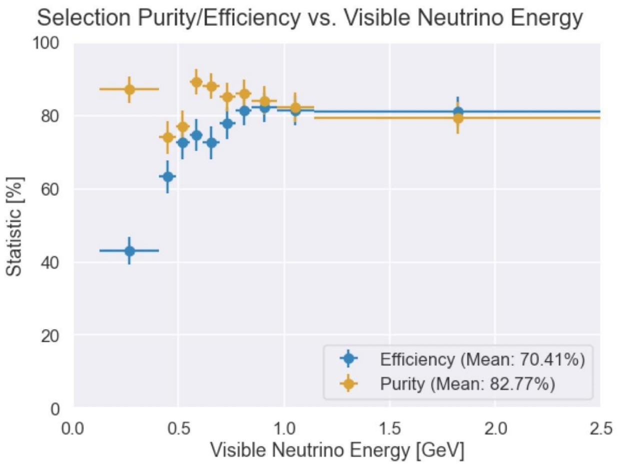


Paper: [arXiv:2102.01033](https://arxiv.org/abs/2102.01033)

ICARUS Analyses Underway

Several physics analyses underway in ICARUS using this ML chain:

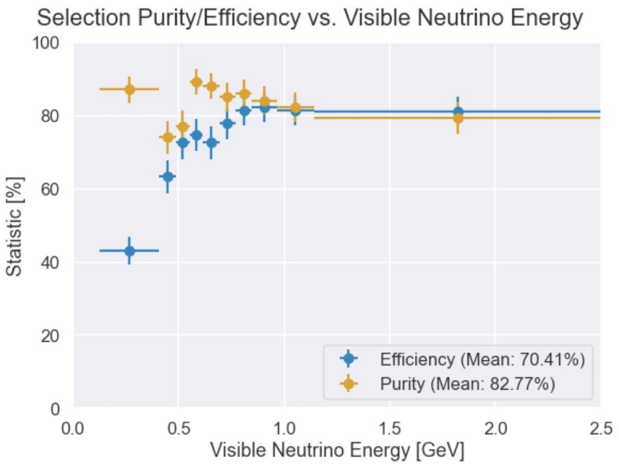
- BNB CCQE 1μ1p selection (J. Mueller): **ν_μ disappearance**/cross section

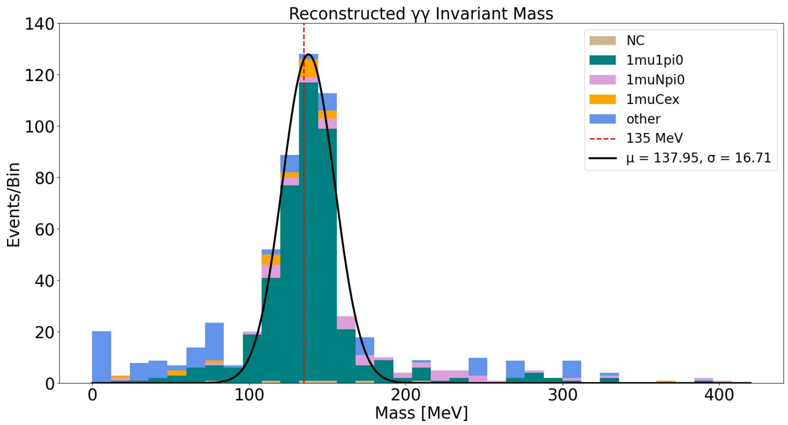


ICARUS Analyses Underway

Several physics analyses underway in ICARUS using this ML chain:

- BNB CCQE 1μ1p selection (J. Mueller): **ν_μ disappearance**/cross section
- BNB CC-π^0 selection (L. Kashur): **shower energy scale**/cross section

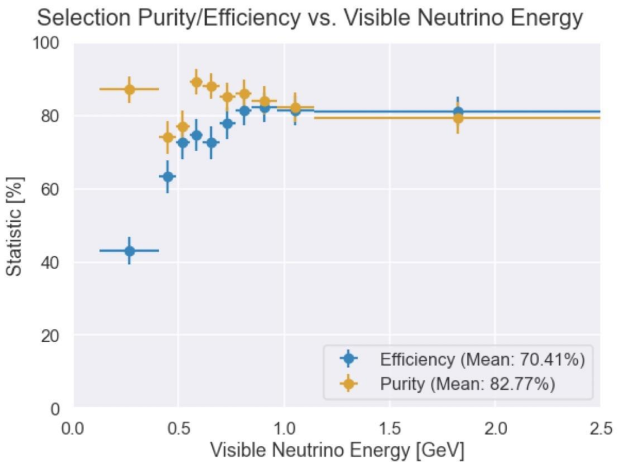


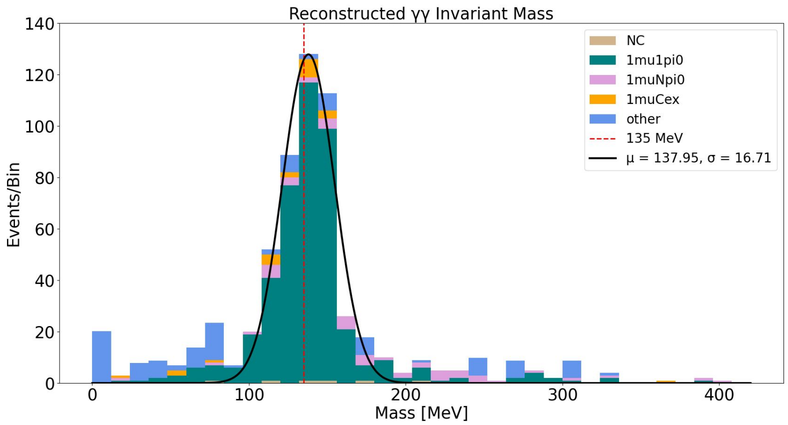


ICARUS Analyses Underway

Several physics analyses underway in ICARUS using this ML chain:

- BNB CCQE 1μ1p selection (J. Mueller): **ν_μ disappearance**/cross section
- BNB CC-π^0 selection (L. Kashur): **shower energy scale**/cross section
- BNB ν_e selection (D.H. Koh): **low energy excess, see his talk this afternoon!**

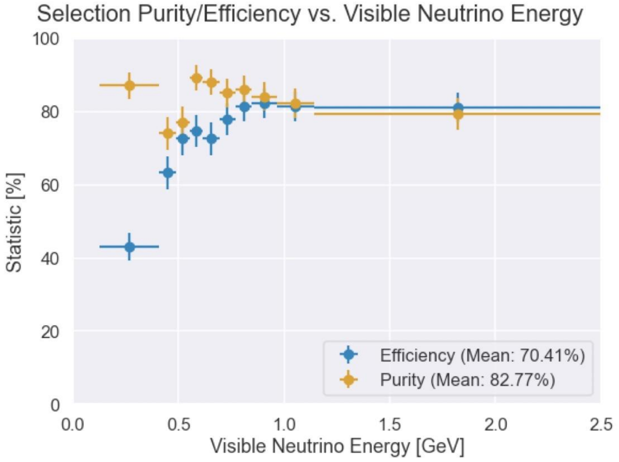


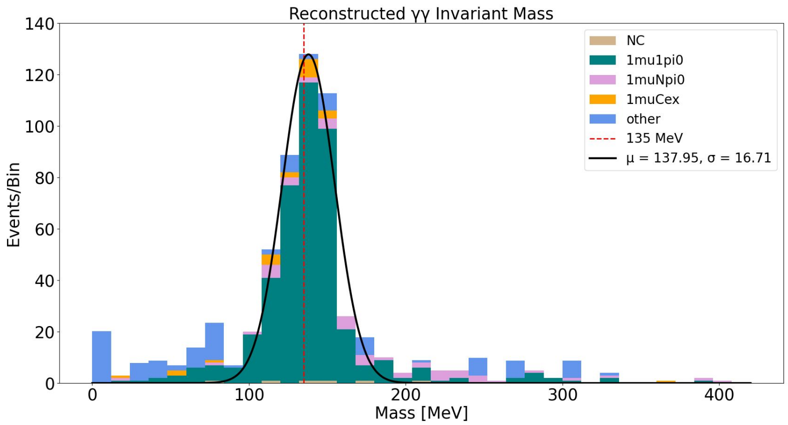


ICARUS Analyses Underway

Several physics analyses underway in ICARUS using this ML chain:

- BNB CCQE 1μ1p selection (J. Mueller): **ν_μ disappearance**/cross section
- BNB CC-π^0 selection (L. Kashur): **shower energy scale**/cross section
- BNB ν_e selection (D.H. Koh): **low energy excess, see his talk this afternoon!**
- NuMI ν_e selection: **cross-section**
- Higgs-portal scalar decays S→ee (J. Dyer): **BSM physics**

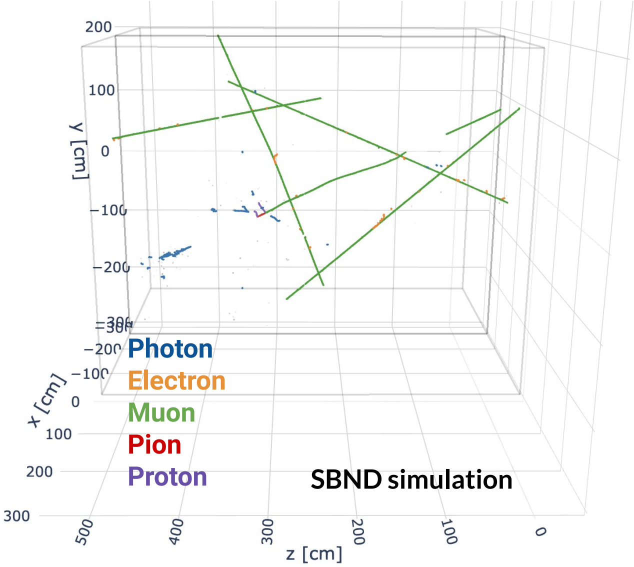


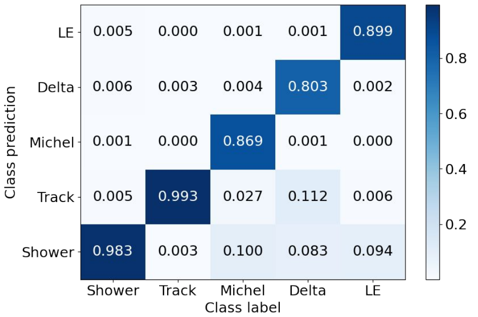


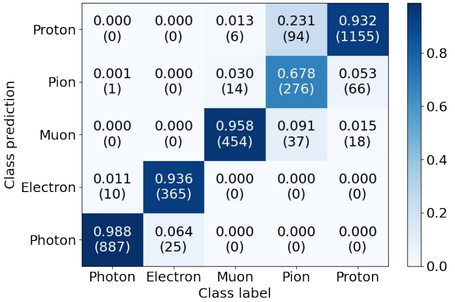
Short Baseline Near Detector (SBND)

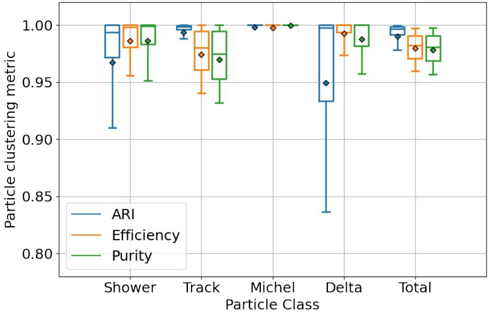
Heavy lifting by **B. Carlson** to port work to SBND

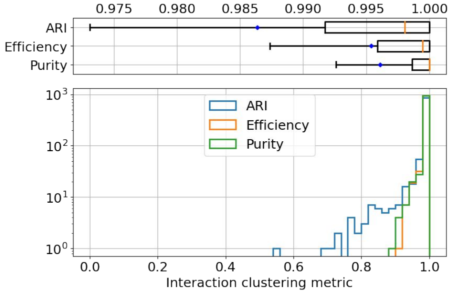
- First training sample produced, **preliminary training completed yesterday**
- **Performance** looks **as expected**, ready for **simulation-based analyses!**







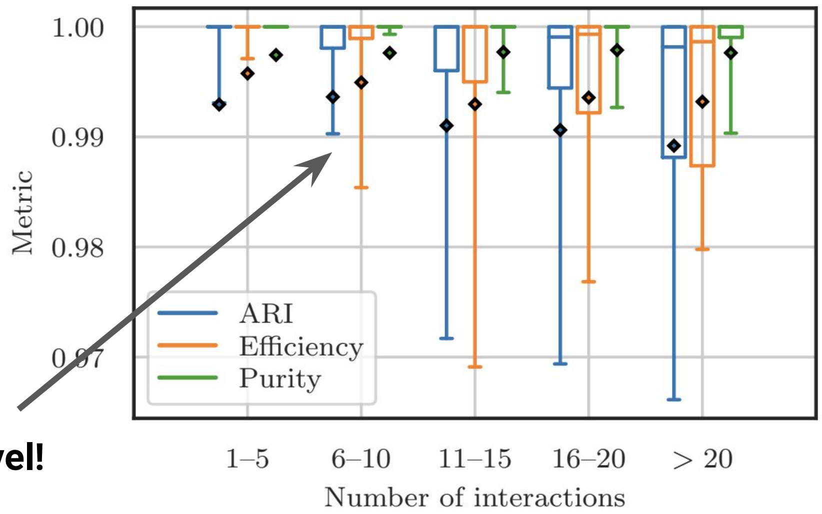




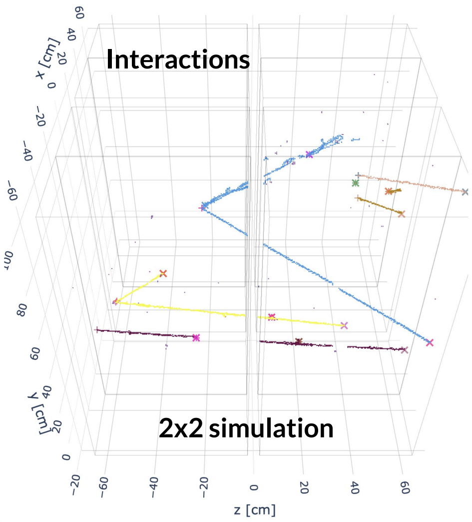
Beyond SBN

DUNE-ND will witness the **highest rate of ν interactions in a LArTPC**

- ~20 ν/spill + 30 rock muons (100 m^3 of LAr)
- Simulations show that **we can deal with that rate**
- 2x2 demonstrator online by EoY, more to come!



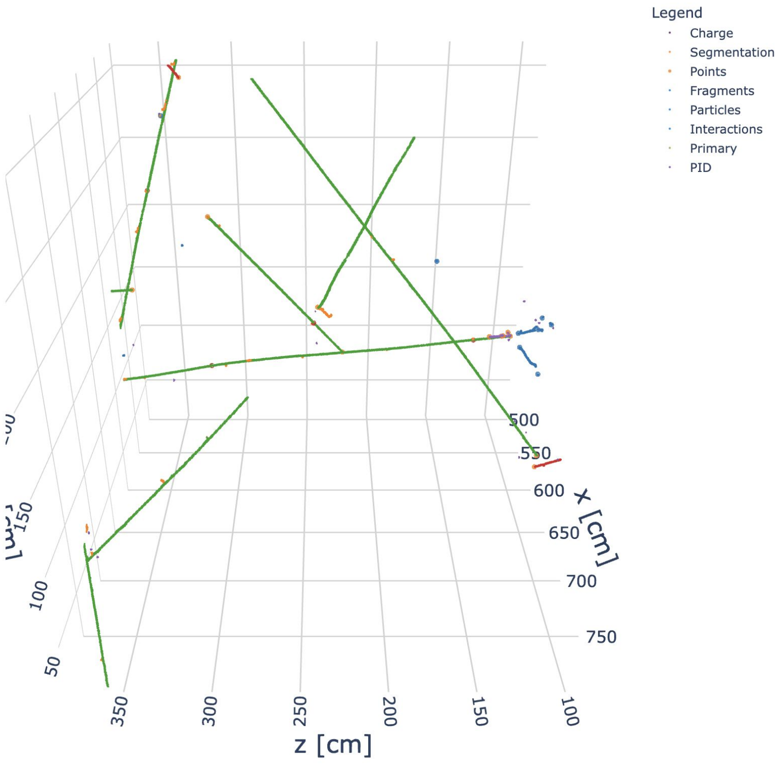
DUNE-ND pile-up level!



Conclusions

End-to-end ML-based reconstruction chain mature and functional:

- **UResNet** for pixel feature extraction, **GrapPA** for superstructure formation
- Used on **ICARUS** sim./data, **SBND** and **DUNE-ND** (high neutrino pileup) sim. **today**! Stay tuned…
- Check out this ICARUS [interactive reconstructed event]()



Backup Slides

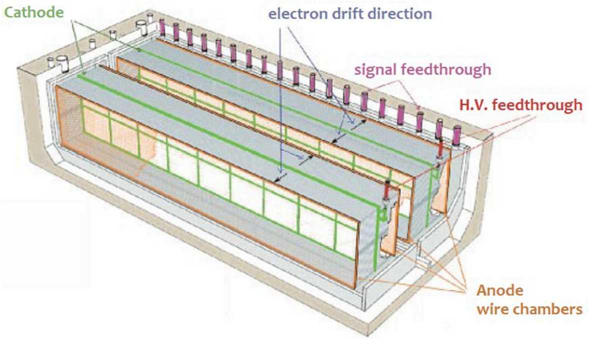
Liquid Argon Time-Projection Chambers

SLAC NATIONAL ACCELERATOR LABORATORY

Case study: Detector

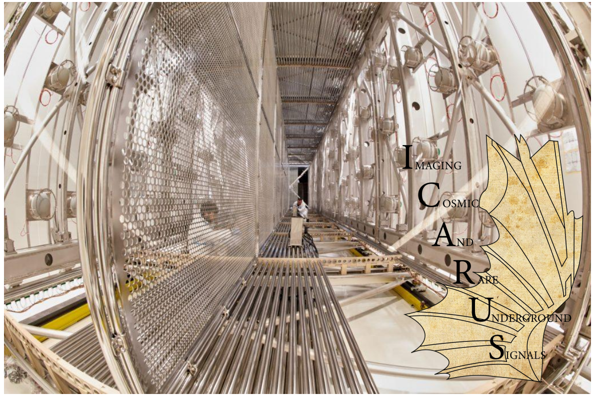
The **largest LArTPC** in operation is **ICARUS**

- **Surface**-level detector
- **500 t** fiducial mass (2 cryos, 4 TPCs)
- **Physics**: sterile neutrinos (MiniBooNE / Neutrino-4), cross sections, BSM



Event rates

- BNB beam: ~ 0.03 Hz neutrinos
- NuMI off-axis: ~ 0.015 Hz neutrinos
- In-time cosmic activity: ~ 0.25 Hz



Low-rate neutrino experiment with a **significant cosmic background**

Message passing

Two **feature update steps**

1. **Edge update**

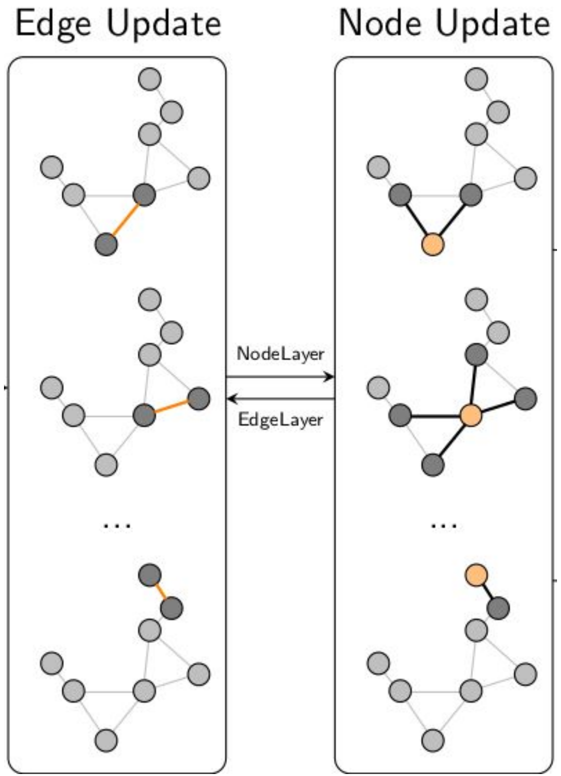
$$\mathbf{e}'_{ij} = \phi_\Theta(\mathbf{x}_i,\, \mathbf{x}_j,\, \mathbf{e}_{ij})$$

2. **Node update**

$$\mathbf{m}_{ji} = \chi_\Theta(\mathbf{x}_j,\, \mathbf{e}_{ji})$$

$$\mathbf{x}'_i = \psi_\Theta(\mathbf{x}_i,\, \square_{j \in \mathcal{N}(i)} \mathbf{m}_{ji})$$

Repeat **n** times (depth)



Edge Selection

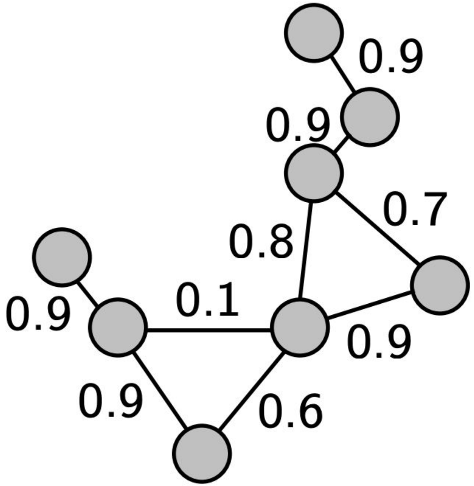
The GNN gives you a list of **edge scores**, not a partition

For the **best partition**, ĝ, we must select edges which minimizes the **partition CE loss**

Classification at the partition level!

$$\mathcal{L}_{\hat{g}} = -\frac{1}{N_e} \sum_{(i,j)\in E} \Big[\delta_{\hat{g}_i,\hat{g}_j} \ln(s_{ij}) + (1-\delta_{\hat{g}_i,\hat{g}_j}) \ln(1-s_{ij})\Big]$$

Edge scores

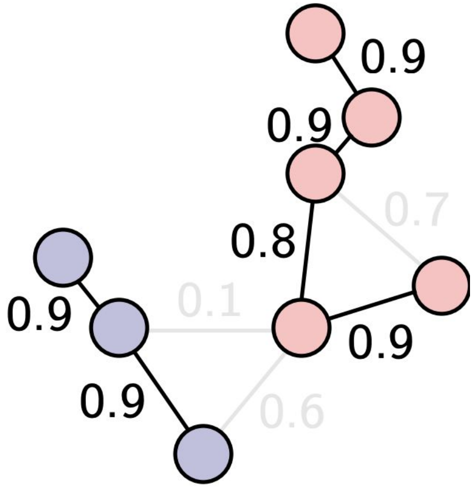


Edge Selection

Instead, **iterate:**

1. Compute partition **loss** for the empty graph
2. Add the **most likely edge**, compute loss again
3. If $L_{n+1} < L_n$, **update partition**
4. Repeat until the next best edge has $s_{ij} < 0.5$

Optimized partition



$L \simeq 2.13$