FERMILAB-SLIDES-24-0146

# SBND Analysis using ML Reconstruction Chain

***Castaly Fan***

*on behalf of the SBND collaboration*

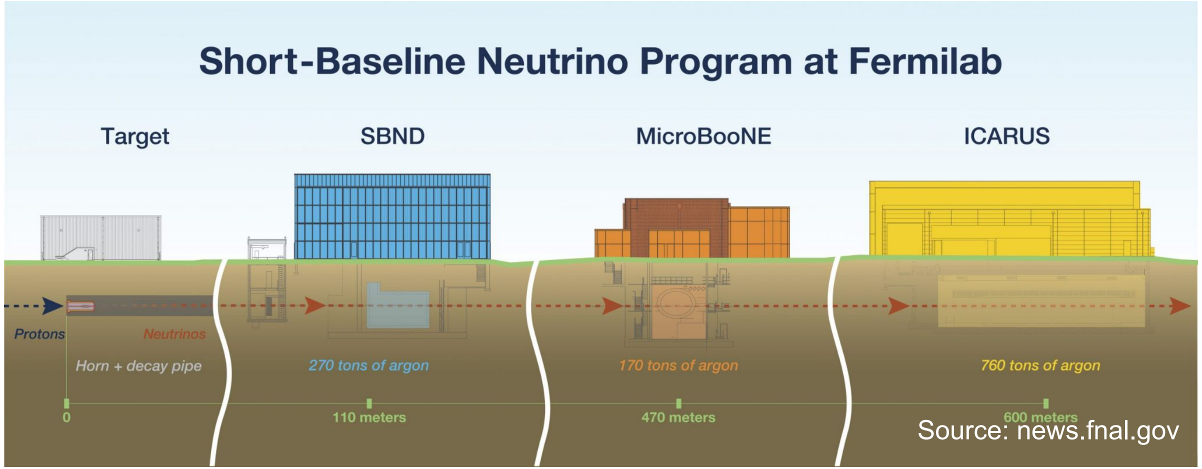
**New Perspectives 2024**

July 8, 2024

This document was prepared by the SBNS 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.

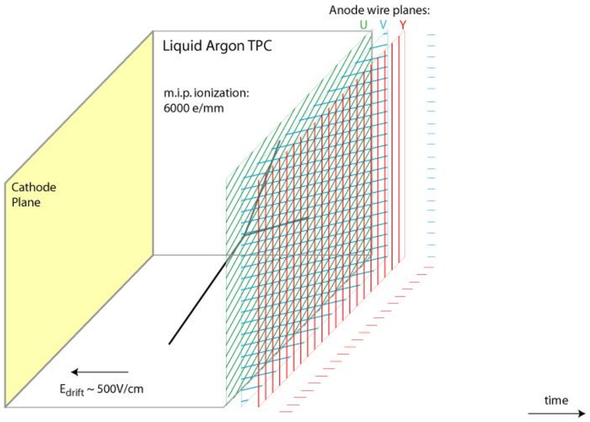
# From neutrinos searches to SBND

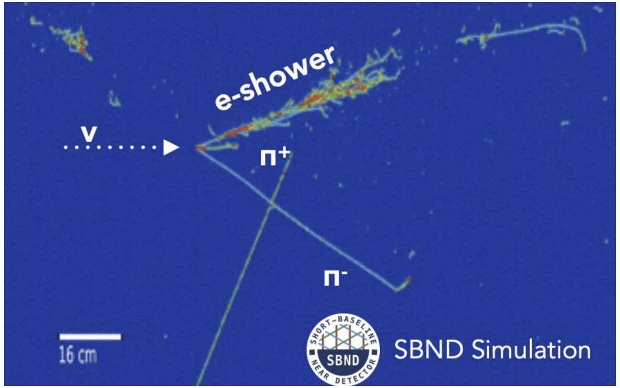
- **SBND**'s proximity (110 m) to the source (BNB) implies we'll collect *a large number* of neutrinos.
- We are here to measure the *un-oscillated* flux for the SBN (Short-Baseline Neutrino) program; cross section and BSM searches will also be studied.



# From SBND to particle identification

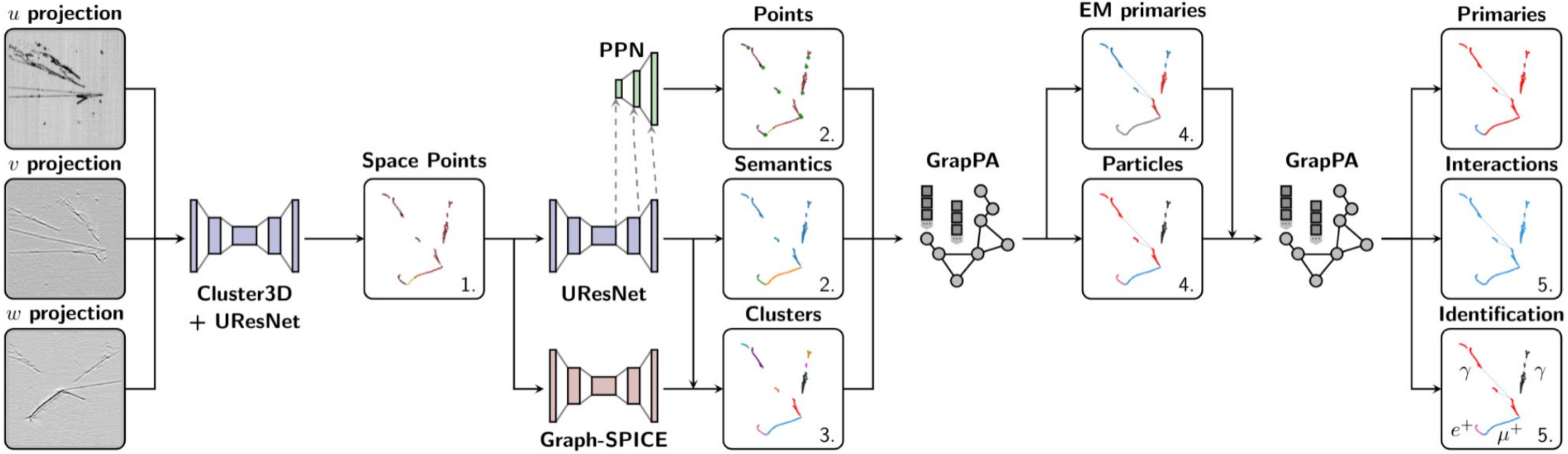
- SBND is a **LArTPC (Liquid Argon Time Projection Chamber)**, where charged particles passing through argon ionizes electrons.
- As the electrons drift toward the anode, the location, arrival time, and the deposited charge is recorded on the wire planes.
  - 2D images formed on the three wire planes will be used to construct 3D images later.



# ML Reconstruction Chain

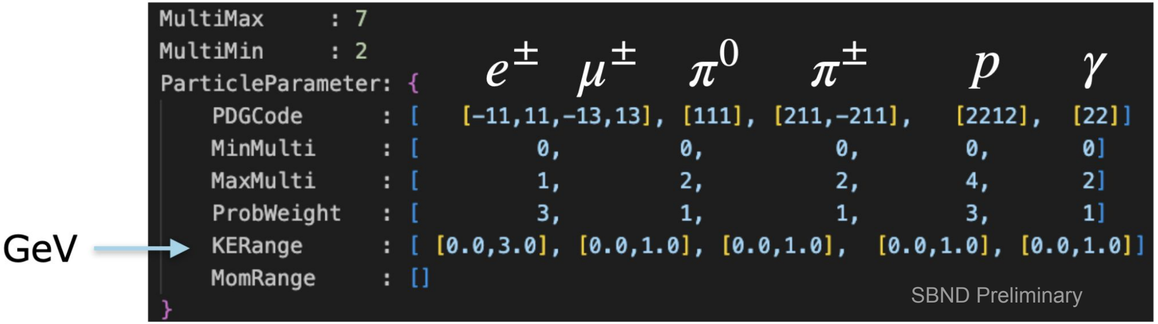
- The **S**calable **P**article **I**maging with **N**eural **E**mbeddings (SPINE) will be employed.




- F. Drielsma: arXiv:2102.01033

# Training and samples

- Multi-particle vertex multi-particle rain (**MPVMPR**) sample has 3 generators:
  - Out-of-time rain (MPR): trains for out-of-time cosmic activity
  - In-time rain (MPR): trains for in-time cosmic activity
  - Vertex (MPV): trains for neutrino activity
- 278k training, validation samples + ~50k testing samples



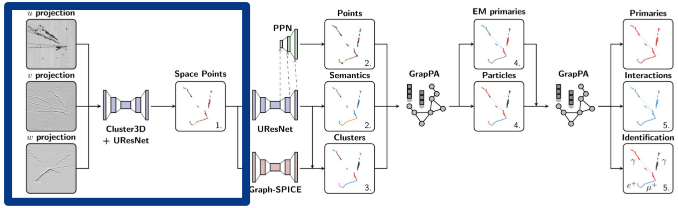
```
MultiMax      : 7
MultiMin      : 2
ParticleParameter: {
    PDGCode      : [   [-11,11,-13,13], [111], [211,-211],   [2212],  [22]]
    MinMulti     : [          0,         0,          0,        0,     0]
    MaxMulti     : [          1,         2,          2,        4,     2]
    ProbWeight   : [          3,         1,          1,        3,     1]
    KERange      : [ [0.0,3.0], [0.0,1.0], [0.0,1.0],  [0.0,1.0], [0.0,1.0]]
    MomRange     : []
}
```

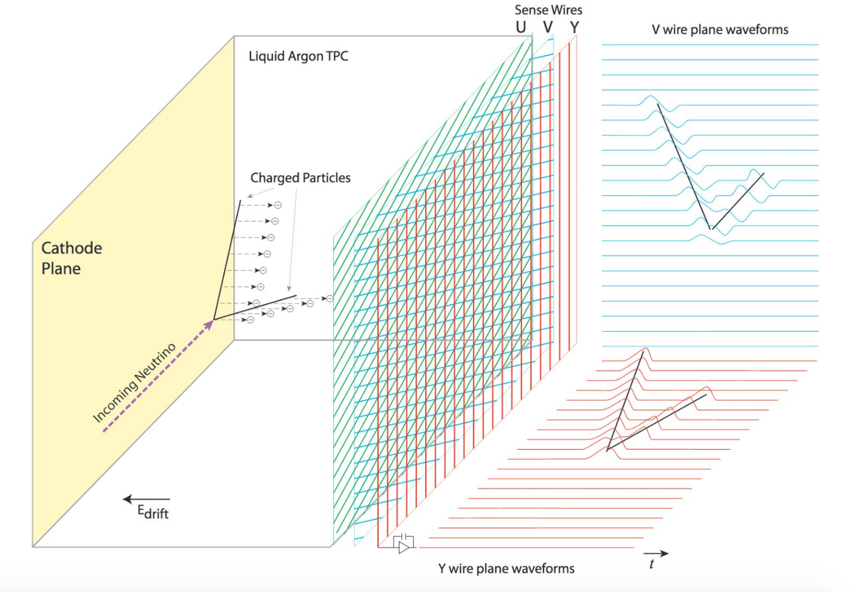


MPV v01 parameters (source: NPML talk by B. Carlson)

# From images to space points

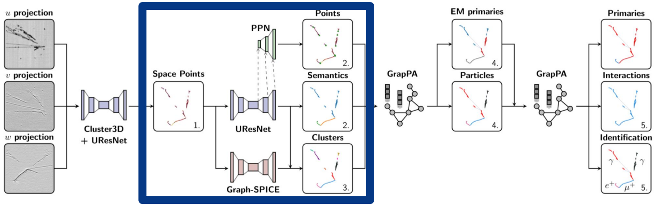
- **Tomographic reconstruction**:
  - Reconstruct the **space points** created by particles passing through the LArTPC
  - Create a 3D image of an object by combining three 2D images taken from wire planes using Cluster3D.
  - Spurious or false signals (e.g. electronic noises) will be identified and “deghosted” using UResNet.



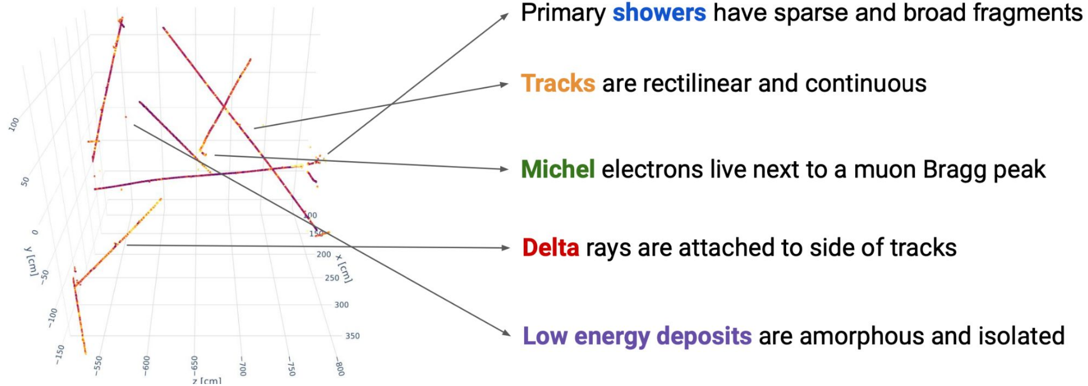




- F. Drielsma: arXiv:2102.01033
- L. Dominé and K. Terao: PhysRevD.102.012005

# From space points to fragments

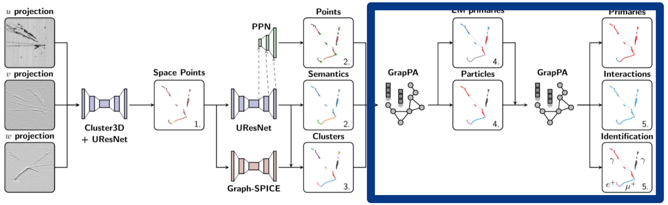



- UResNet → Locates and names the space points:
  - **Semantic Segmentation**: Classify the points into five classes (as listed below)
  - PPN identifies "points of interest": Identify start and end points.
- Graph-SPICE → **Clusters** points into particle fragments (and later into particles)
  - For the tracks/showers that *share the same vertex* → **Dense clustering** by Graph-SPICE
  - For the *breaking tracks* (e.g. cathode crosser) or *shower fragments* → **Aggregation** (later)

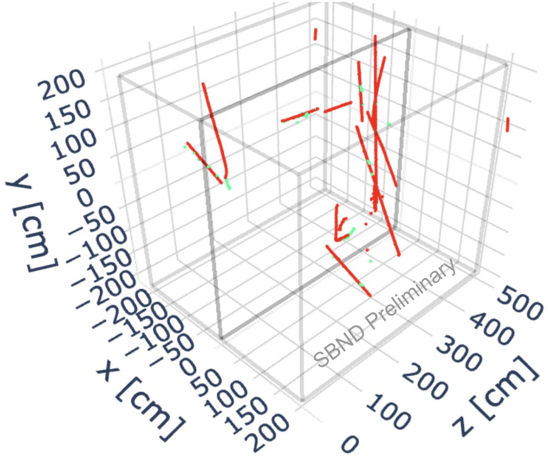



- Image from: SBN/ICARUS ML Workshop 2023
- F. Drielsma, *et al*.: Phys. Rev. D 104, 072004

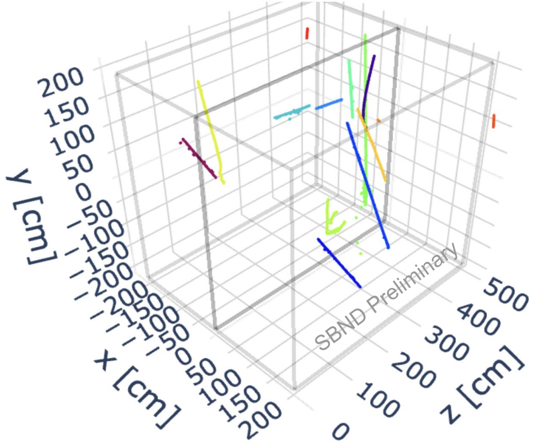
# From fragments to particles



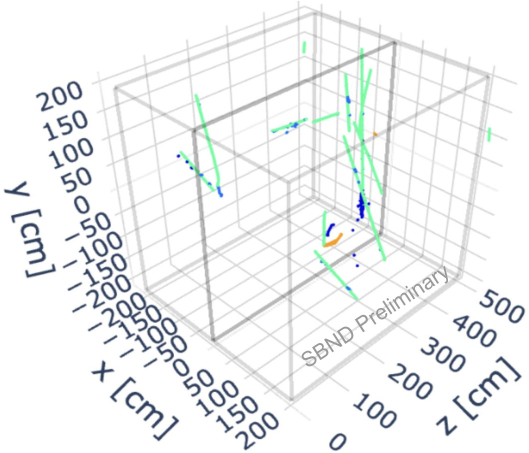

- We utilize <u>GrapPA</u> to do aggregation tasks:
  - Aggregate fragments (broken tracks & shower fragments) into **particles**.
  - Aggregate particles into the original **interactions**.
  - Classify particles into five species (**photons**, **electrons**, **muons**, **pions**, or **protons**)



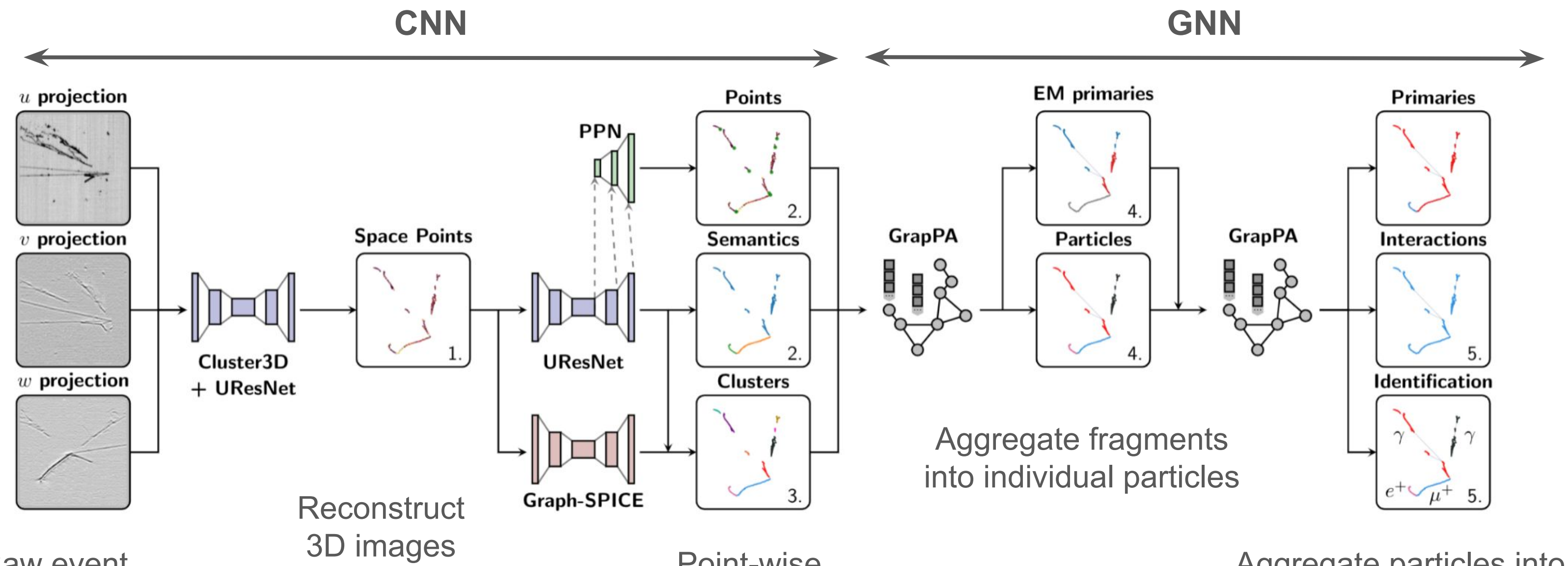

Primaries or not




Interactions (from the same parent or others)




Particle types (five species)



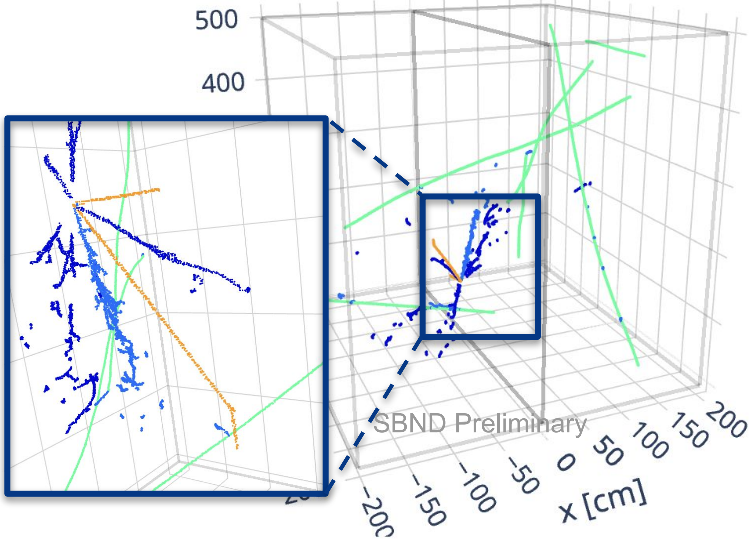

Raw event signals from the detector

Point-wise classification and fragment clustering

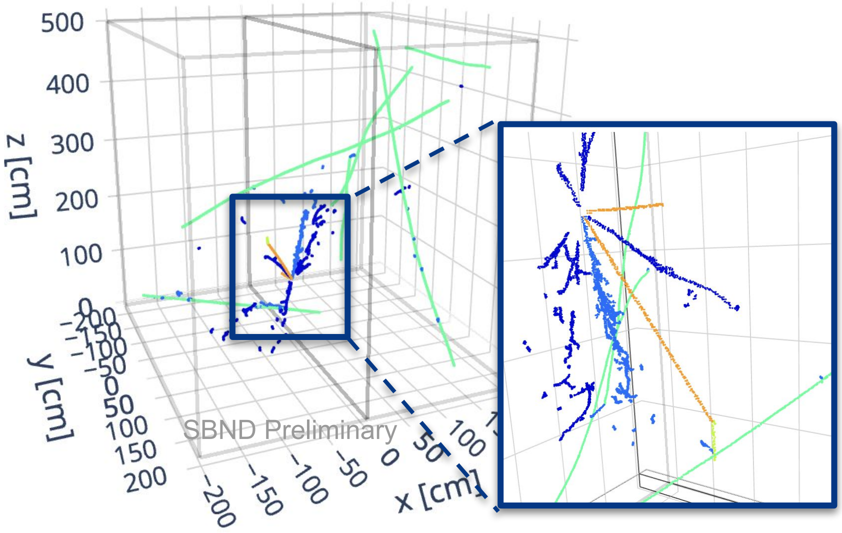
Aggregate particles into interactions & perform PID and primary identification

# Performance

- SPINE did an excellent job, even for the tricky tasks e.g. shower classification.
- Primary PID accuracy 85.5%
    - Electron confusion comes from a class imbalance (more primary photons than electrons) during training.



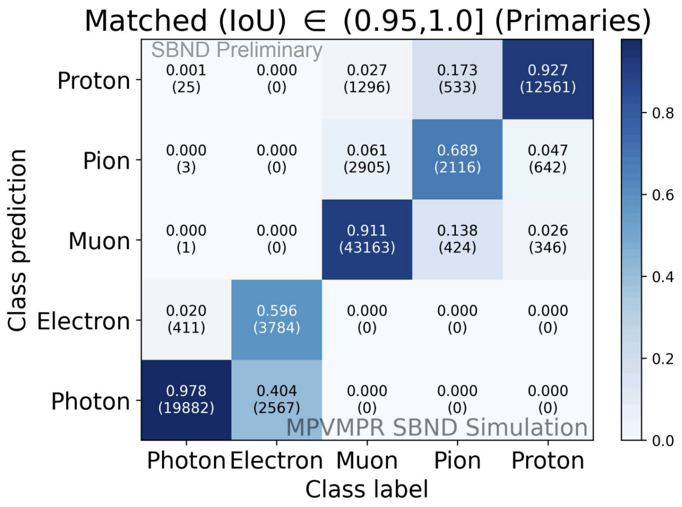

Reco



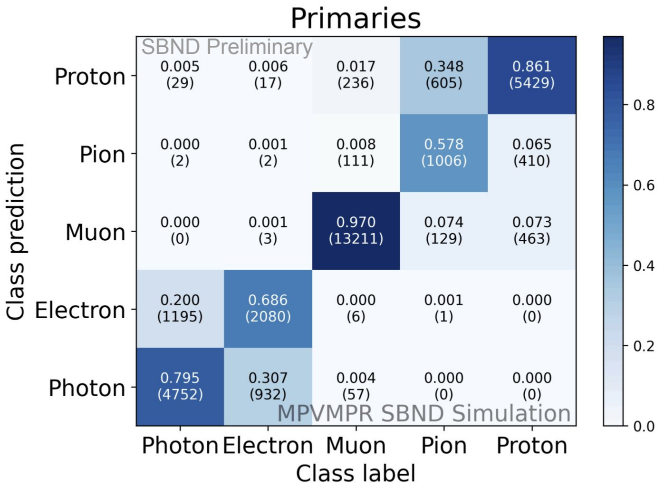

True

# Performance

- SPINE did an excellent job, even for the tricky tasks e.g. shower classification.
- Primary PID accuracy 85.5%
    - Electron confusion comes from a class imbalance (more primary photons than electrons) during training.



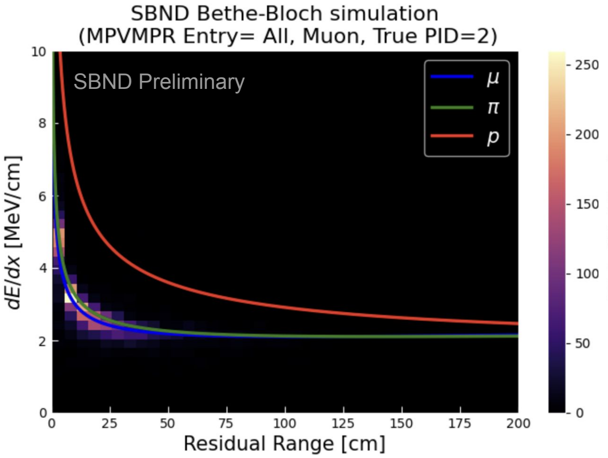

Purity from primaries
(from the *latest* MPVMPR sample)

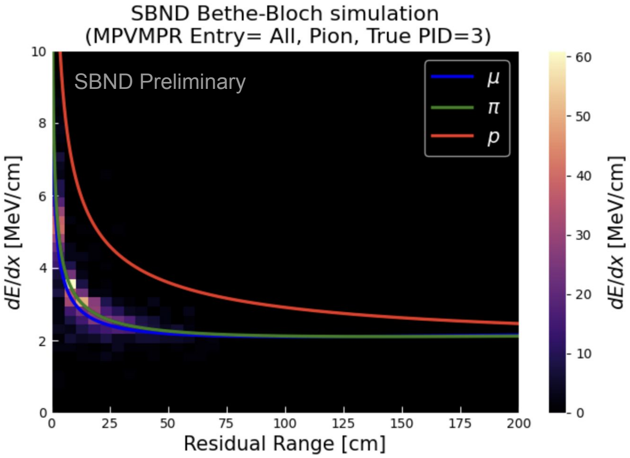


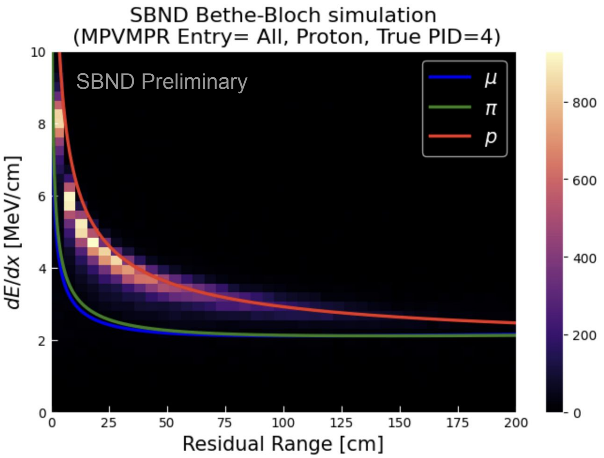

Purity from primaries
(from the *previous* MPVMPR sample)

# Example: dE/dx analysis

- The *dE/dx vs. residual range* plots among muons and pions:
  - SPINE achieved a great performance: They are basically matched the predicted curves.
  - Muons and pions have similar dE/dx profiles, but still distinguishable even with a tiny difference.
  - The discernible shift on proton’s plot comes from the ADC-MeV conversion factor for now.



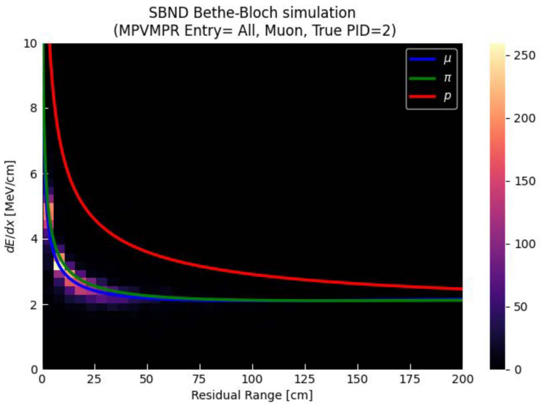
# Summary

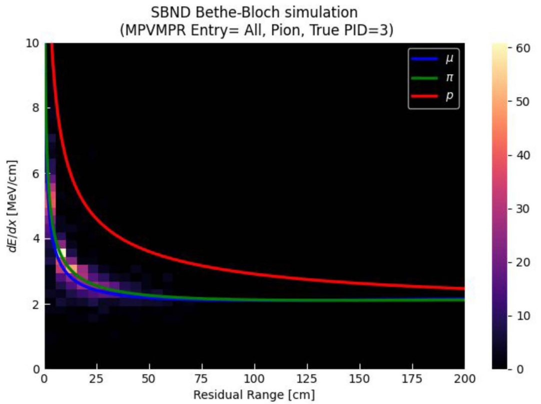
- SPINE is an ML reconstruction framework for event reconstruction, which is currently employed for ICARUS will be widely employed for SBND analyses.
- The reconstruction chain utilizes cutting-edge CNN and GNN to reconstruct the events.
- Particle and interaction types can be identified with high performance.
  - It’s “scalable” (performance won’t be affected by data amount/complexity)
  - And also, an “end-to-end” structure (a single integrated system)
- Resources
  - SPINE’s github page: DeepLearnPhysics/SPINE
  - Useful resources: ML workshop 2023
  - Original paper: arXiv:2102.01033
  - More about UResNet: PhysRevD.102.012005
  - More about GrapPA: Phys. Rev. D 104, 072004
  - SBND overview: NPML conference slides

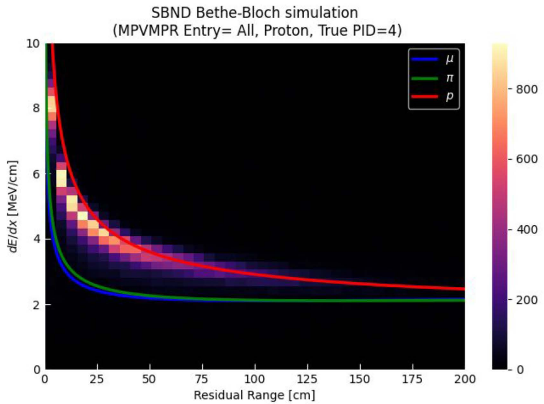
# Thank you!

# Backups

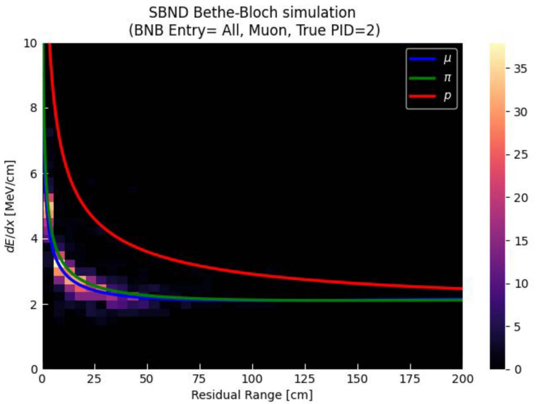
MPVMPR

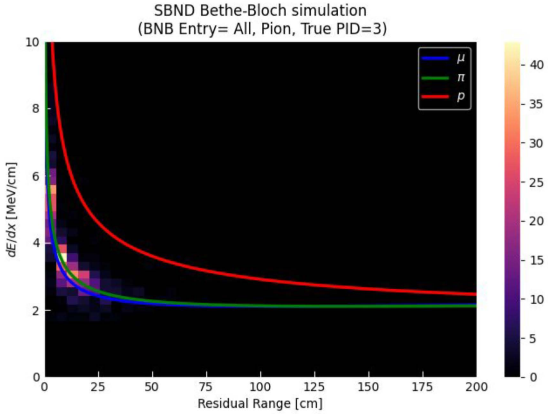

SBND Bethe-Bloch simulation
(MPVMPR Entry= All, Muon, True PID=2)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p

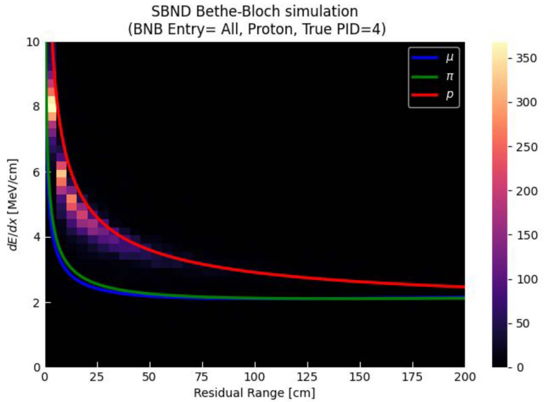

SBND Bethe-Bloch simulation
(MPVMPR Entry= All, Pion, True PID=3)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p


SBND Bethe-Bloch simulation
(MPVMPR Entry= All, Proton, True PID=4)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p

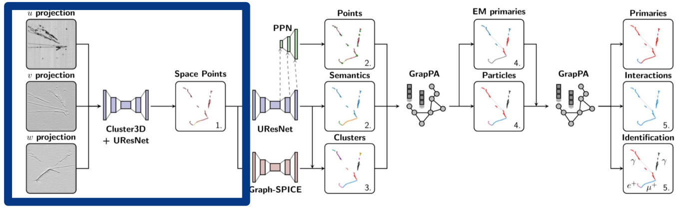
BNB


SBND Bethe-Bloch simulation
(BNB Entry= All, Muon, True PID=2)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p


SBND Bethe-Bloch simulation
(BNB Entry= All, Pion, True PID=3)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p

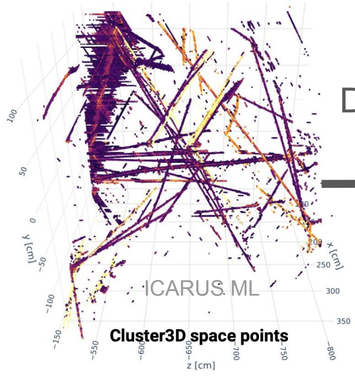

SBND Bethe-Bloch simulation
(BNB Entry= All, Proton, True PID=4)
dE/dx [MeV/cm]
Residual Range [cm]
μ
π
p

# Tomographic Reconstruction

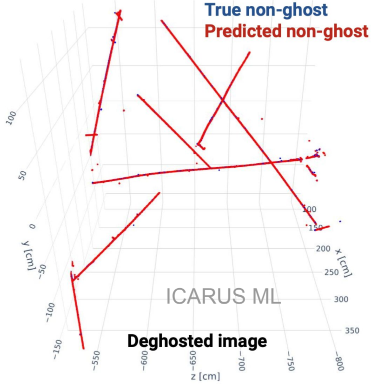



- “Ghosts”:
  - spurious or false signals (e.g. electronic noises) arise from matching plane images.
- **Deghost**:
  - Classify each voxel into two categories i.e. ghost and non-ghost → Identify and then bust them!
  - How? **UResNet** neural network will be used; it can help semantic segmentation too.

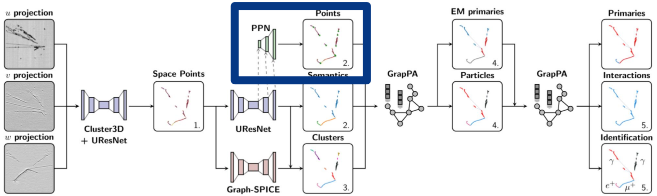
Tomographic reconstruction (Cluster3D)



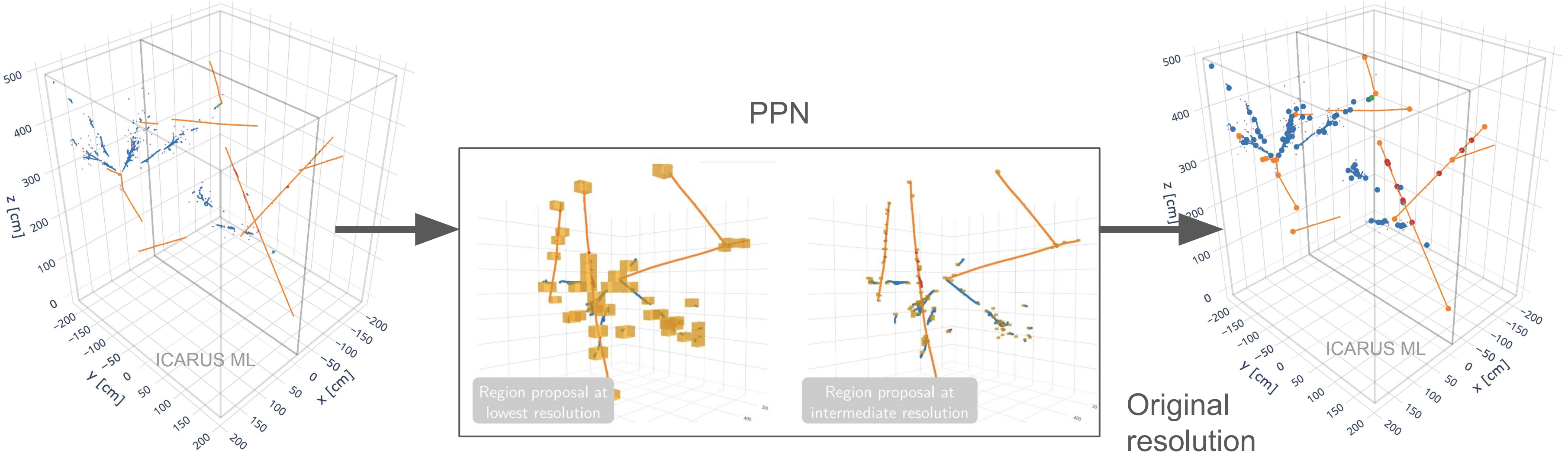

Deghost



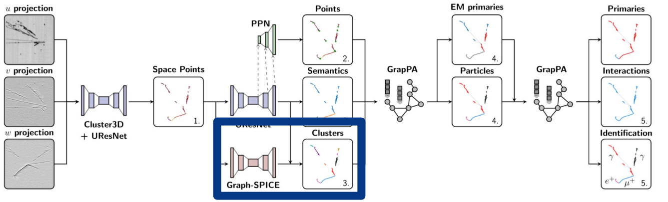
# Point Proposal Network (PPN)



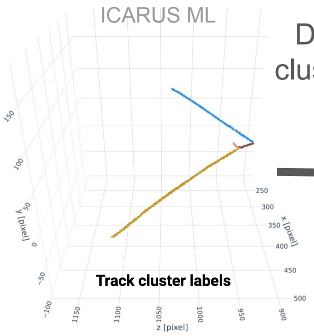

- **PPN** identifies "points of interest"
  - Using a UResNet.
  - Labels start and end points of tracks, deltas, and start points of showers.



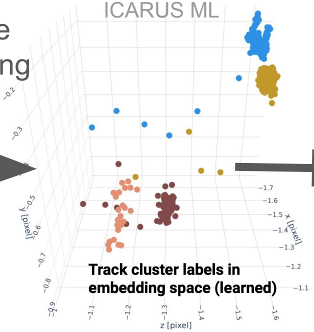
# (Dense) Clustering

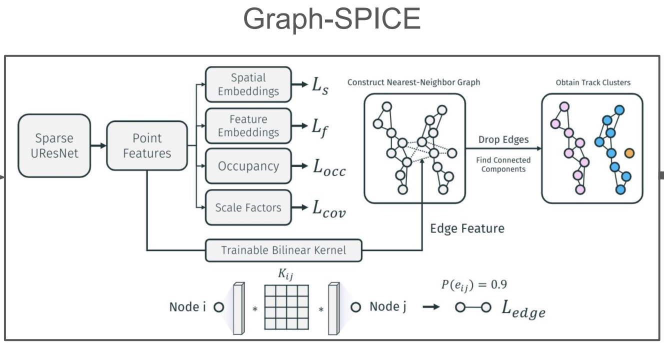


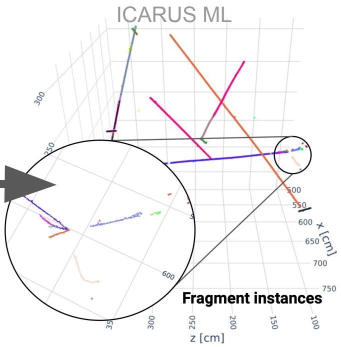

- But we have to know “which is which”
- → **clustering** will be required:
  - Finding connected voxels that touch each other will be necessary
    - However, some of the tracks/showers *share the same vertex* → **Dense clustering**
    - Besides, it would be challenging to identify some of the *breaking tracks* (e.g. cathode crosser) or gather the *shower fragments* → **Aggregation**



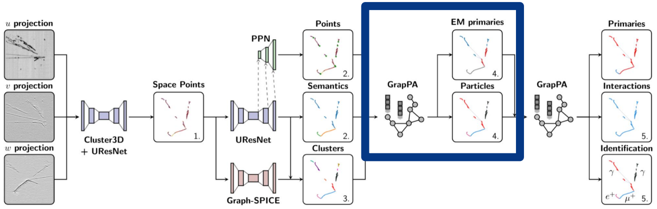

Dense clustering



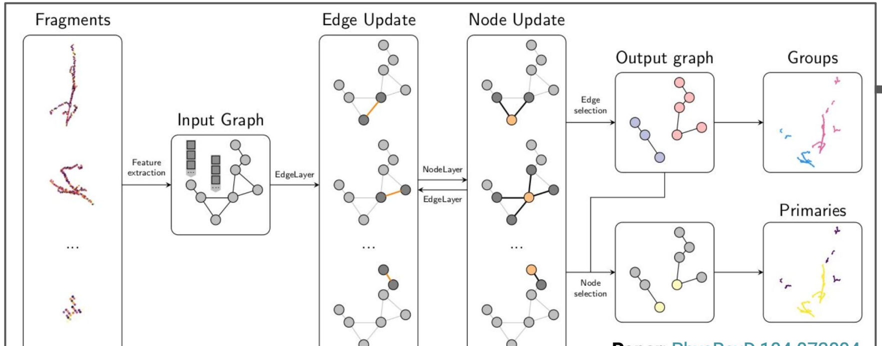



# Aggregation (for Particles)

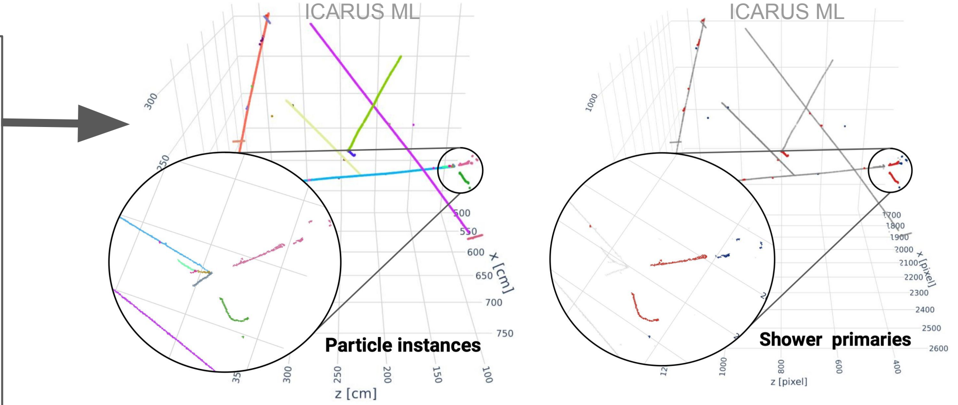



- Now, we need to cope with the particle fragments:
  - Broken tracks & shower fragments.
- **GrapPA** (Graph Particle Aggregator) can help.
  - Node features: Particle centroid, dQ/dx, PCA, covariance matrix, PPN end points, directions
  - Aggregate fragments into **particles**.
  - Classify shower into **primary/secondary**.

GrapPA



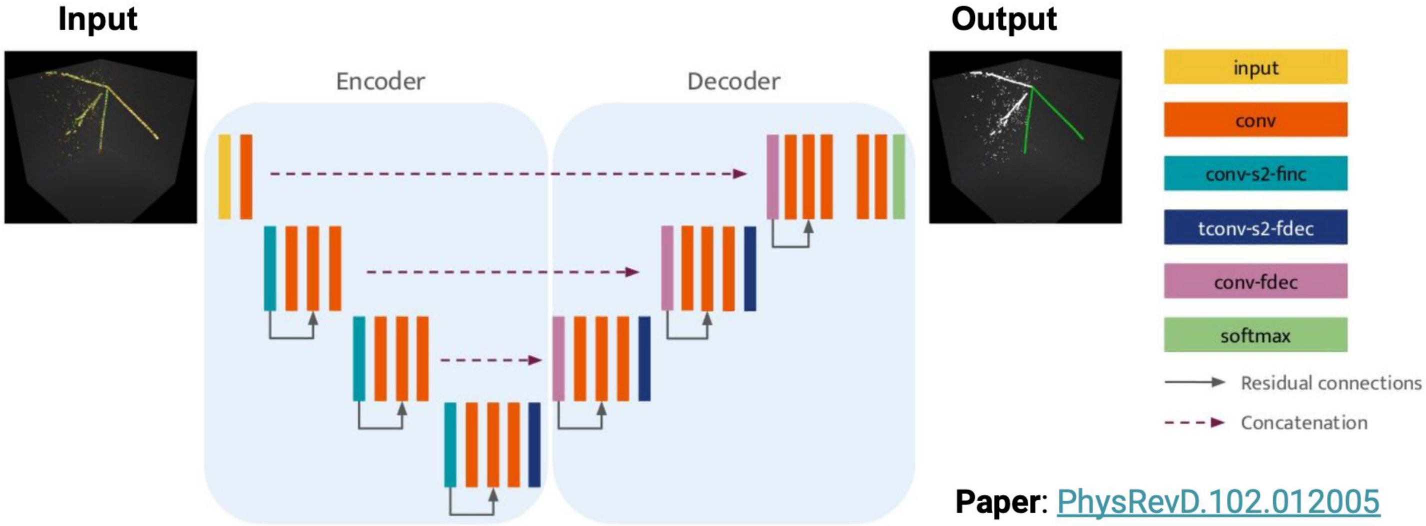

**Paper**: PhysRevD.104.072004



# Semantic Segmentation

## Backbone (L. Dominé)

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




**Paper**: PhysRevD.102.012005

ML-based Reconstruction for LArTPCs, F. Drielsma (SLAC) 14

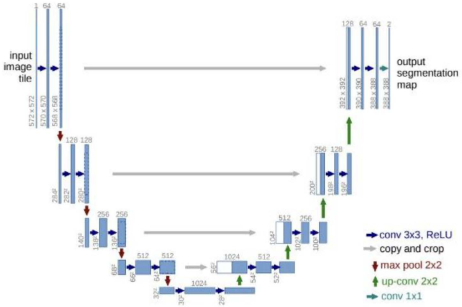
# Semantic Segmentation

## Backbone (L. Dominé)

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

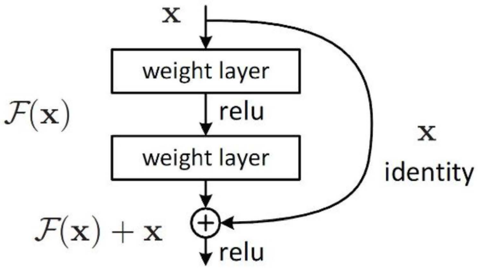
### UNet

- Downsizing -> expand receptive field
- Skip connections -> preserve resolution



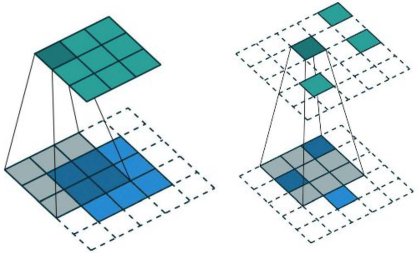

### ResNet

- Identity bypass + convolution -> learns residual transform
- Speeds up learning, enables deeper networks




### Sparse Convolutions

- Only applies convolutions on active pixels
- Saves memory, execution speed (dramatically)



ML-based Reconstruction for LArTPCs, F. Drielsma (SLAC)

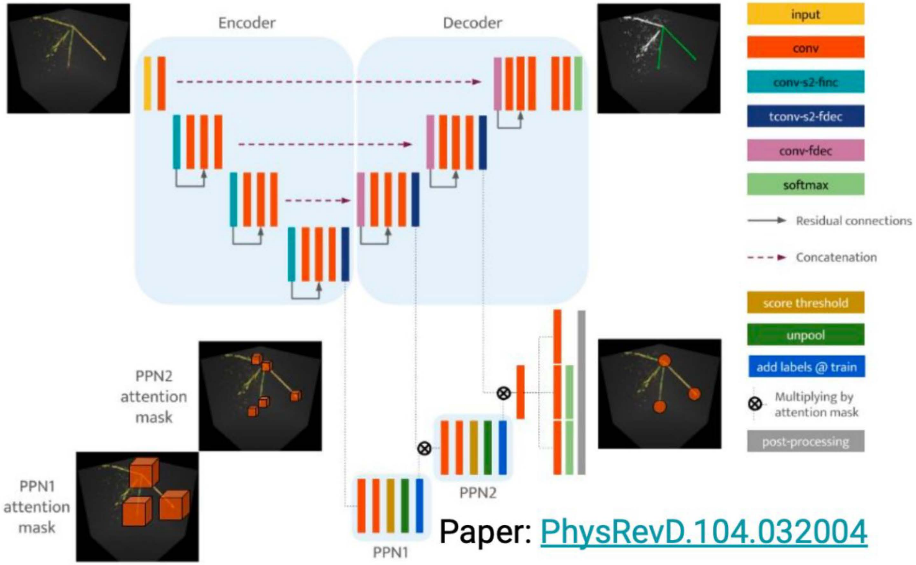
15

# Point Proposal Network (PPN)

## Architecture (L. Dominé)

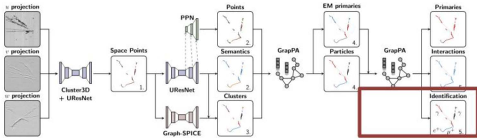
The Point Proposal Network (PPN) identifies **points of interest** using decoder features:

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

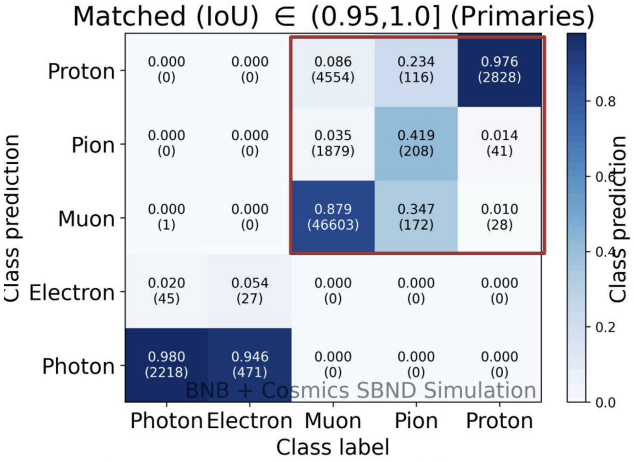



Paper: PhysRevD.104.032004

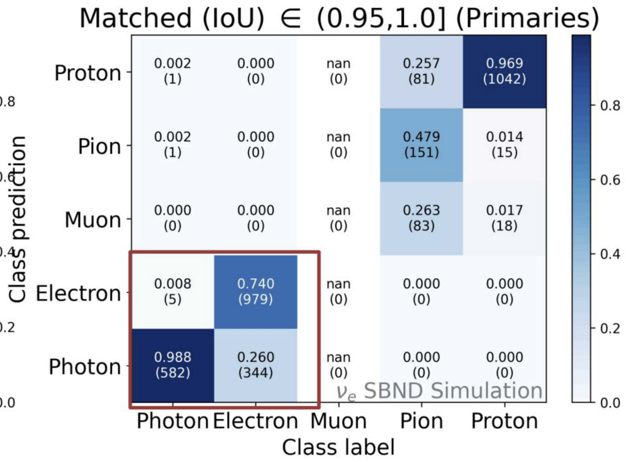
# Graph particle aggregator (GrapPA)




- Aggregates **particles** into **interactions** and identifies **primaries** and **PID**
- Primary PID accuracy **85.5%**
- Electron-photon confusion from poor class balancing during training




BNB + cosmics + rockbox




Intrinsic nu e