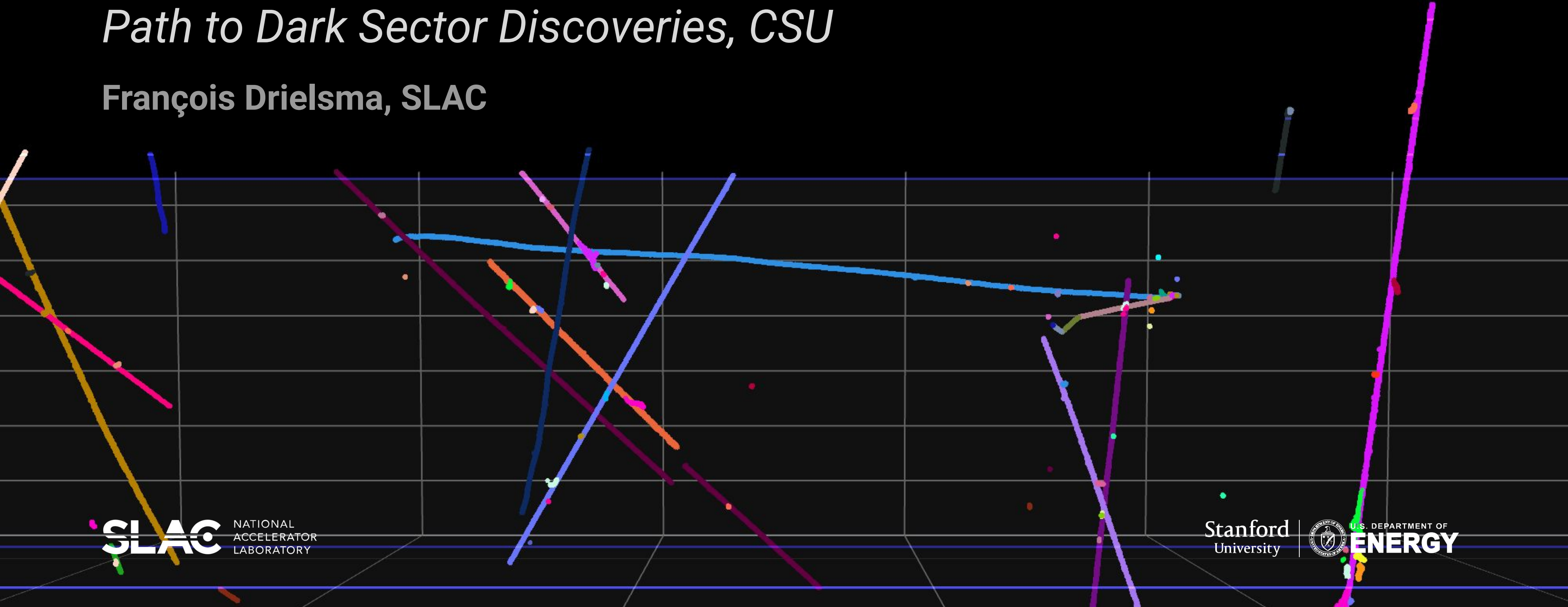
Scalable, End-to-end, ML-Based Reconstruction Chain in LArTPCs

Path to Dark Sector Discoveries, CSU

François Drielsma, SLAC

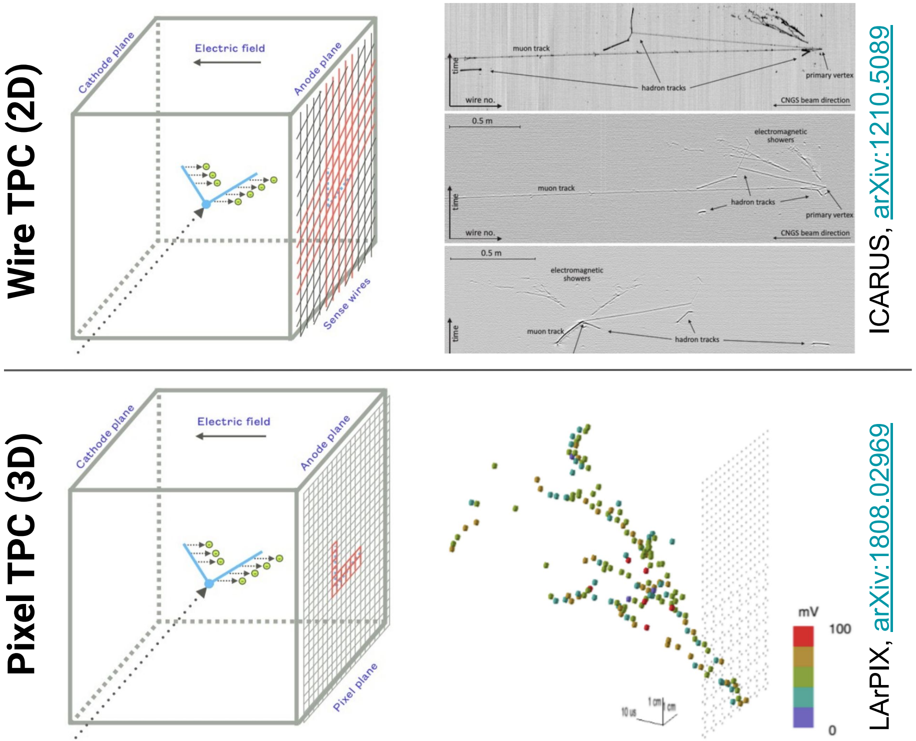


SLAC NATIONAL ACCELERATOR LABORATORY

Stanford University | U.S. DEPARTMENT OF ENERGY

Liquid Argon Time-Projection Chambers

The modern Particle Imaging Detector



LArTPC are at the center stage of **beam ν physics** in the US

Short Baseline Neutrino program

- µBooNE, **ICARUS**, SBND

DUNE long-baseline experiment

- **Wire**: DUNE FD
- **Pixel**: DUNE ND-LAr

Advantages:

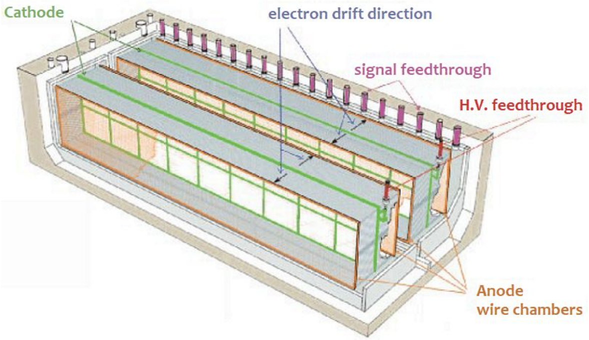
- **Detailed:** O(1) mm resolution, precise calorimetry
- **Scalable**: Up to tens of kt

Liquid Argon Time-Projection Chambers

Case study: Detector

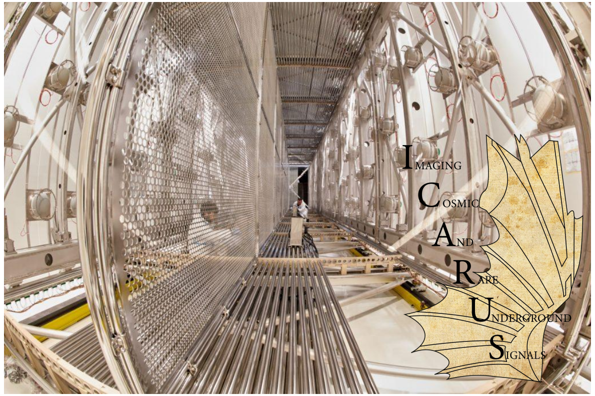
The **largest LArTPC** in operation is **ICARUS**

- **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**

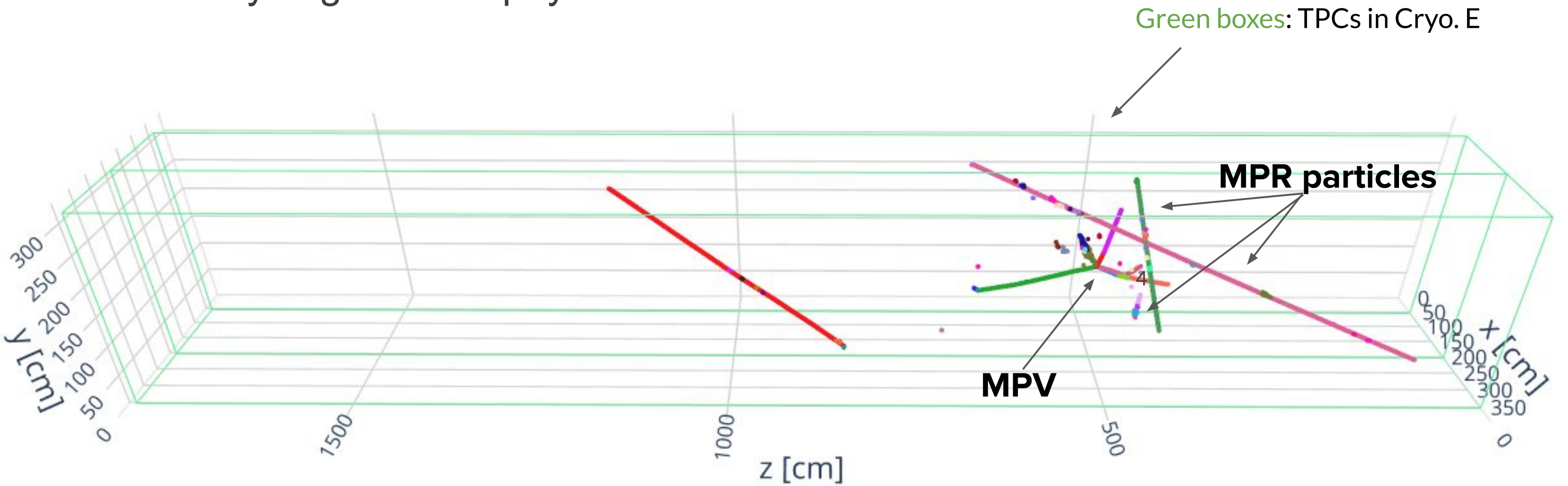
Liquid Argon Time-Projection Chambers

SLAC NATIONAL ACCELERATOR LABORATORY

Case study: Datasets

Generic **simulated** dataset used for **optimization** and **testing**:

- Isotropic mix of 1 set of particles sharing a vertex + 5-9 localized single particles
  - **Covers phase-space** of neutrino interactions + cosmics, but…
  - … stays agnostic to physics → unbiased



Liquid Argon Time-Projection Chambers

SLAC NATIONAL ACCELERATOR LABORATORY

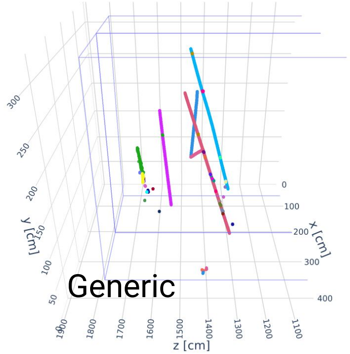
Case study: Datasets

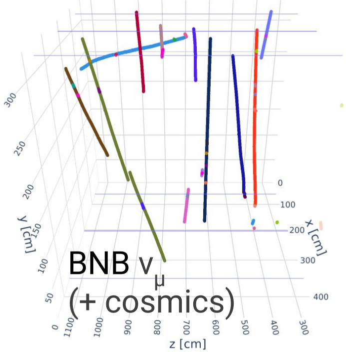
Generic **simulated** dataset used for **optimization** and **testing**:

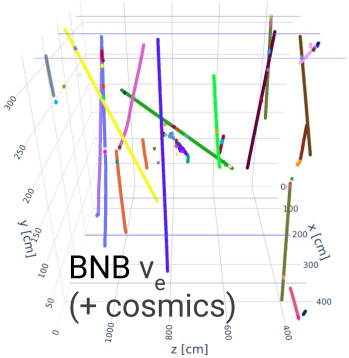
- Isotropic mix of 1 set of particles sharing a vertex + 5-9 localized single particles
    - **Covers phase-space** of neutrino interactions + cosmics, but…
    - … stays agnostic to physics → unbiased

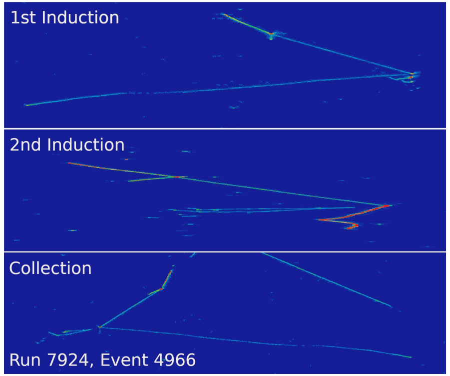
Specific datasets used for **validation**:

- Simulated **BNB ν_μ** and **BNB ν_e** + **hand-scanned data events** (C. Farnese et al.)





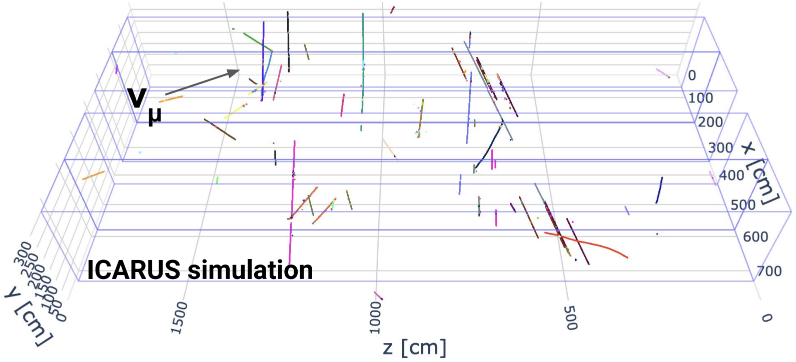


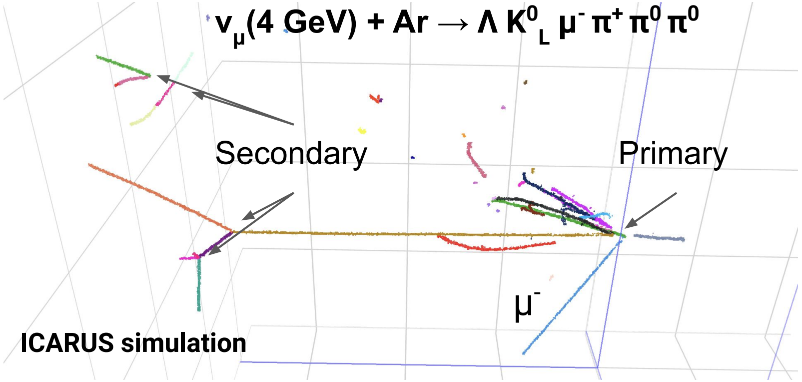


Reconstruction in LArTPCs

Challenges with LAr

Dense medium → Slow	High Z material → Messy
Electron drift velocity O(1) mm/μs • **Long** (O(1) ms) readout **window** • Need **light association** for timing	Argon has a large nucleus (Z=18) • **Complicated** nuclear physics • **Secondary** interactions





Reconstruction in LArTPCs

Machine Learning Group at SLAC

LArTPC ML effort **started** at SLAC by **K. Terao**

- Funded by **DoE ECA and ML grants**, many synergies with **ML initiative**
- **Primary Goal**: Implement **full** ML-based reco. **chain** for LArTPCs
- **Experiments**: µBooNE, **ICARUS**, pDUNE-SP, **pDUNE-ND**, DUNE-ND

ML Convener

T. Usher
ICARUS

F. Drielsma
ICARUS/ND

Y.F. Chen
DUNE-ND

L. Dominé
ICARUS

P. Tsang
DUNE

Y-J. Jwa
ICARUS

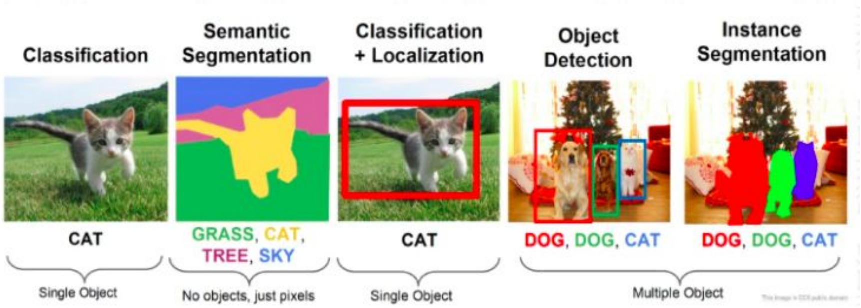
D.H. Koh
ICARUS

Z. Hulcher
DUNE-ND

Machine Learning in Computer Vision (CV)

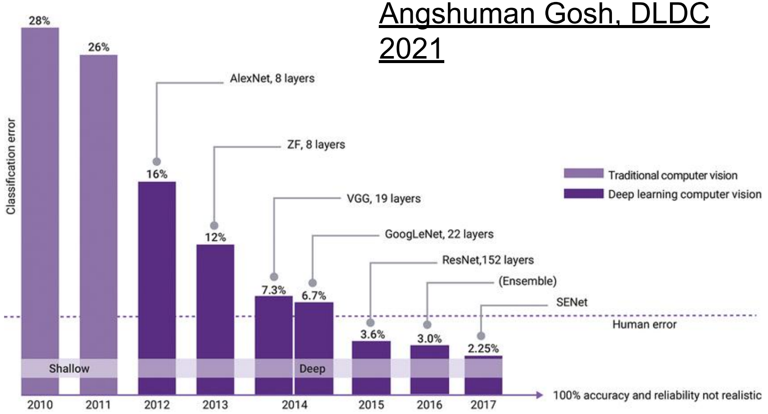
ML is the state-of-the-art in CV, i.e. extracting high level information from images

- ML revolutionized accuracy on image processing tasks
- Should **leverage those techniques in HEP**



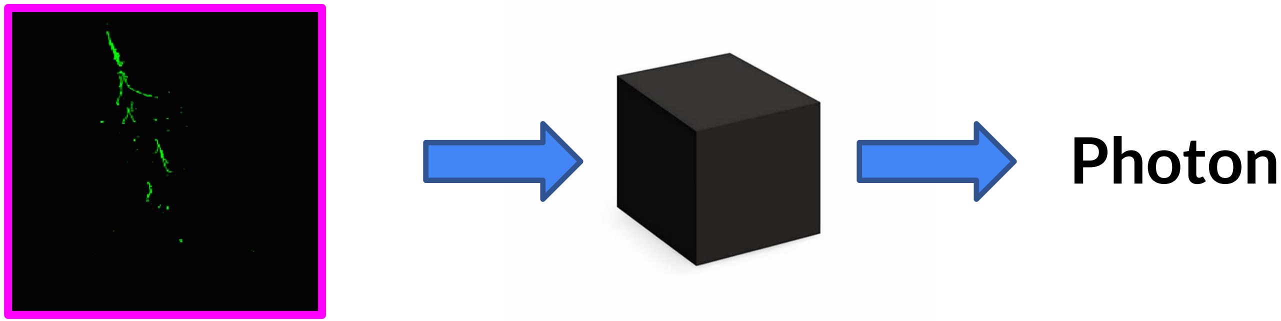
Stanford, CS231

Angshuman Gosh, DLDC 2021



Reconstruction in LArTPCs

Neural Network Primer



Reconstruction in LArTPCs

Neural Network Primer

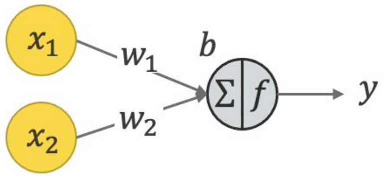
$$\mathbf{x} = \begin{bmatrix} 0 & 20.78 & \cdots & 0 \\ 0 & 56.34 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 86.3 \end{bmatrix} \Rightarrow F(\underbrace{\mathbf{x}}_{\text{Input}}, \underbrace{\mathbf{w}}_{\text{Parameters}}) \Rightarrow \mathbf{y} = \begin{array}{cc} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{array}{c} \gamma \\ e \\ \mu \\ \pi \\ p \end{array} \end{array}$$

Reconstruction in LArTPCs

Neural Network Primer

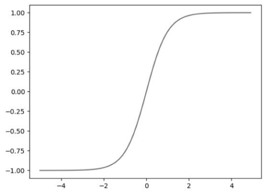
$$\mathbf{x} = \begin{bmatrix} 0 & 20.78 & \cdots & 0 \\ 0 & 56.34 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 86.3 \end{bmatrix} \Rightarrow F(\mathbf{x}, \mathbf{w}) \Rightarrow \mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{matrix} \gamma \\ e \\ \mu \\ \pi \\ p \end{matrix}$$

What is **F ?** In ML, typically a **neural network (NN)**, a **universal function approximator**



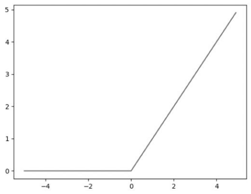
$$y = f(w_1 x_1 + w_2 x_2 + b)$$
$$f = \text{Tanh, ReLU, ...}$$

Tanh

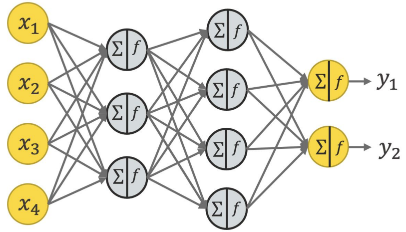


$$f(a) = \frac{e^{2ha} - 1}{e^{2ha} + 1}$$

Rectified Linear Unit (ReLU)



$$f(a) = max\{0, ha\}$$



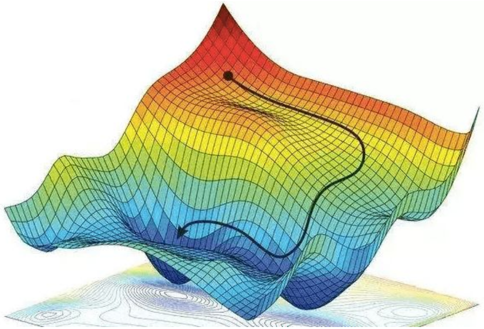
Machine Learning

Brief Primer

$$\mathbf{x} = \begin{bmatrix} 0 & 20.78 & \cdots & 0 \\ 0 & 56.34 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 86.3 \end{bmatrix} \Rightarrow F(\mathbf{x}, \mathbf{w}) \Rightarrow \mathbf{y} = \begin{array}{cc} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{array}{c} \gamma \\ e \\ \mu \\ \pi \\ p \end{array} \end{array}$$

How does it learn ? NN are fully differentiable. Define loss, optimize by **gradient descent**

$$L = ||F(\mathbf{x}, \mathbf{w}) - \mathbf{y}|| = \begin{bmatrix} 0.8 \\ 0.18 \\ 0.01 \\ 0.01 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

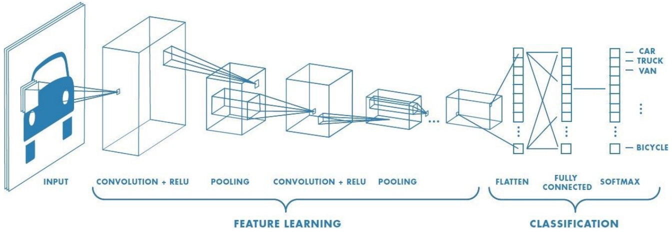


$$\mathbf{w}_i = \mathbf{w}_{i-1} - \alpha \nabla_{\mathbf{w}} L$$

Reconstruction in LArTPCs

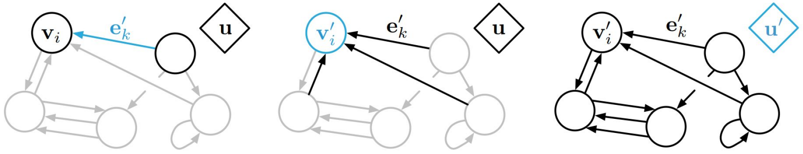
Convolutional Neural Network (CNN)

- Receptive field (kernel): **pixel neighborhood**
- Kernel shared in image: **translation**



Graph Neural Network (GNN)

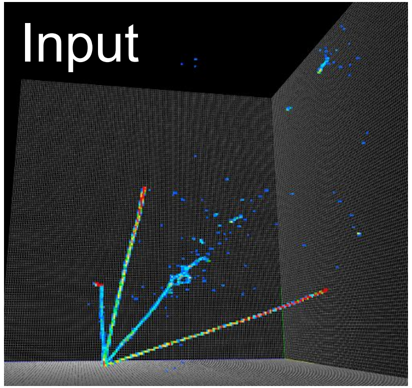
- Receptive field: **graph neighborhood**
- Agnostic to ordering: **permutation invariant**

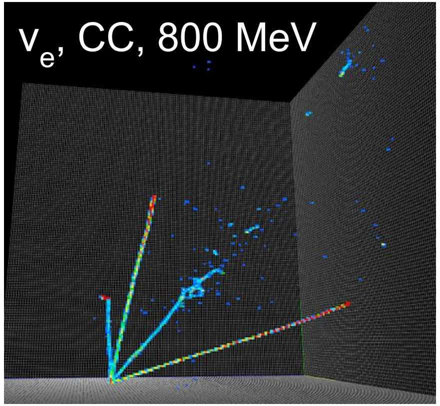


(a) Edge update (b) Node update (c) Global update

Reconstruction in LArTPCs

Hierarchical feature extraction





Reconstruction in LArTPCs

Hierarchical feature extraction

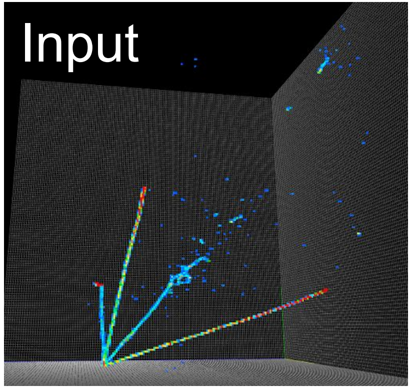
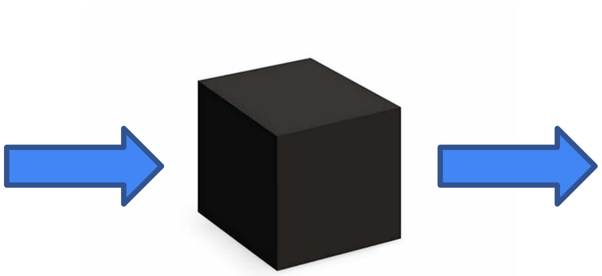
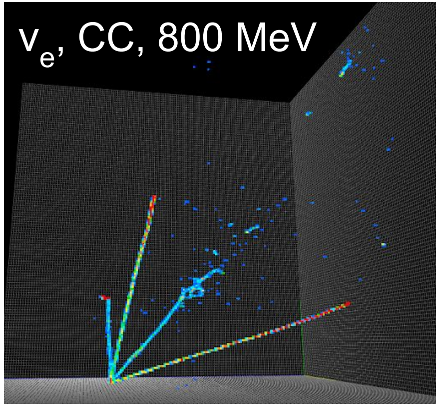


Image Classifier (CNN)





Reconstruction in LArTPCs

Hierarchical feature extraction

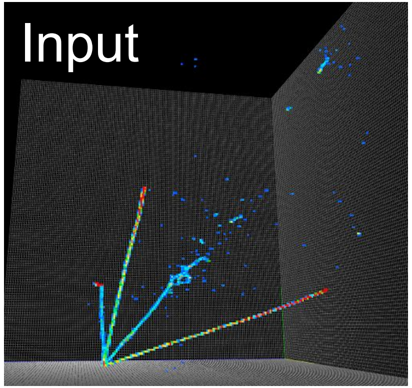
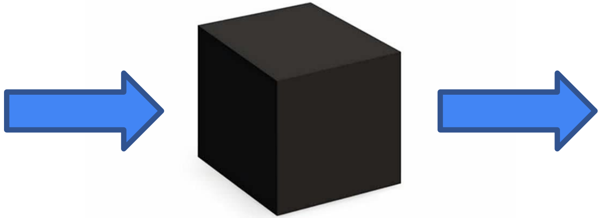
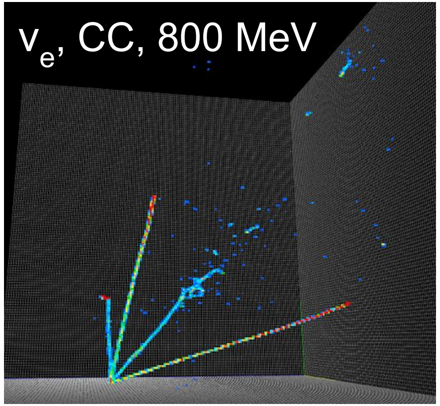


Image Classifier (CNN)

- What to do with > 1 interaction ?
- What if it fails ? Why ?
- What behavior if unknown interaction?

16

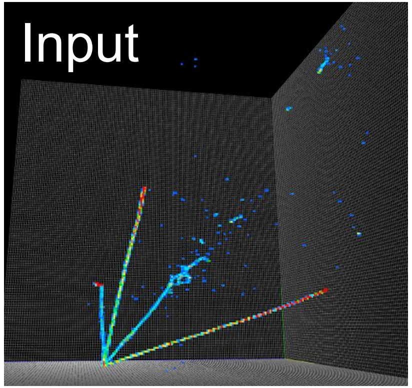




Reconstruction in LArTPCs

Hierarchical feature extraction

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

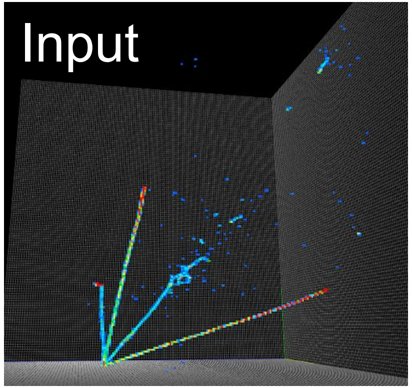


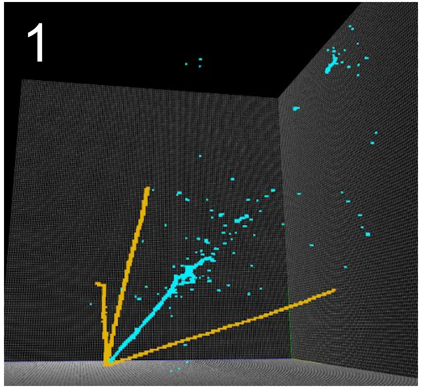
Reconstruction in LArTPCs

Hierarchical feature extraction

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

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



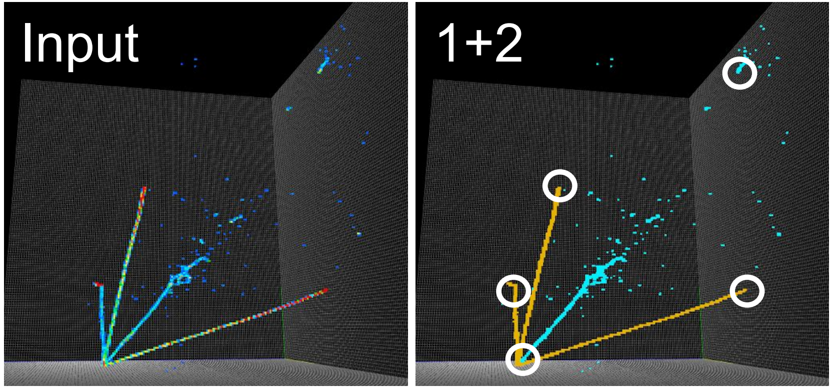


Reconstruction in LArTPCs

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)

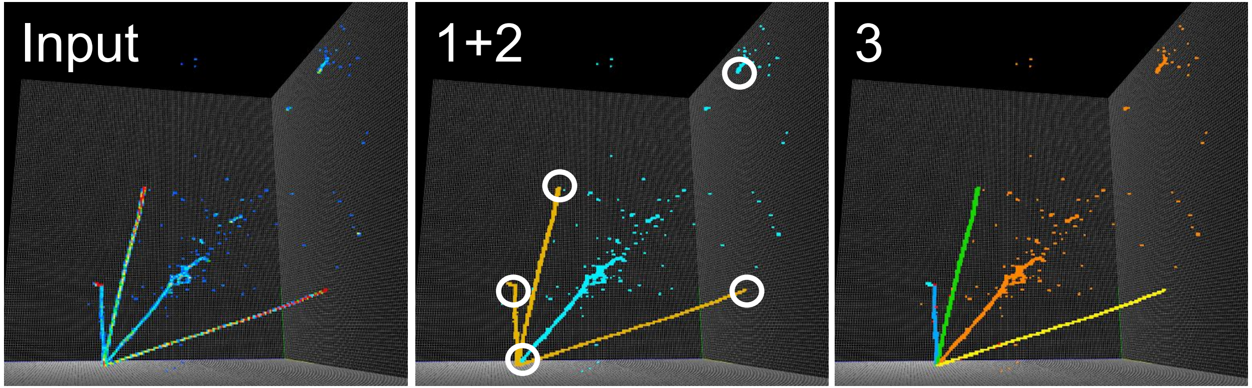


Reconstruction in LArTPCs

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)



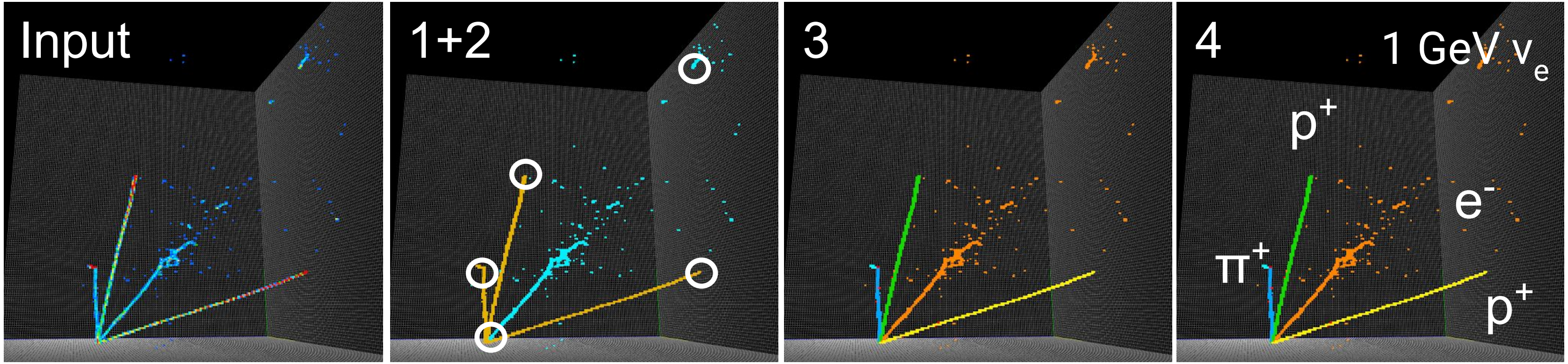
Reconstruction in LArTPCs

SLAC NATIONAL ACCELERATOR LABORATORY

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



Reconstruction in LArTPCs

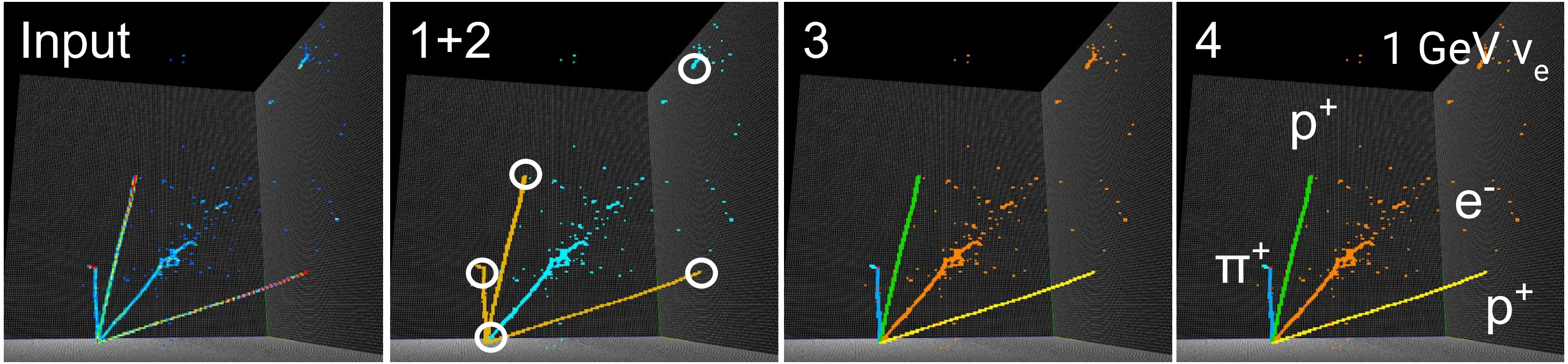
Hierarchical feature extraction

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

Pixel-level (1–2)

Cluster-level (3–4)

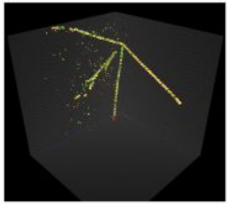


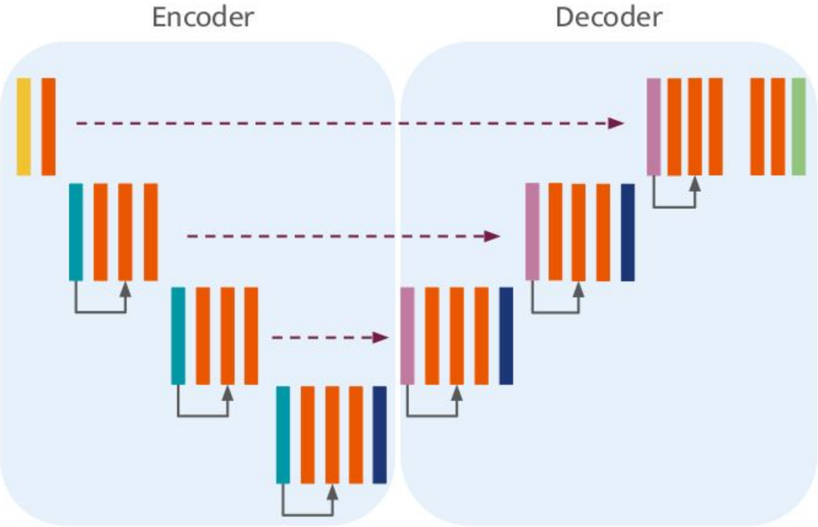
Pixel-Level Feature Extraction

Backbone

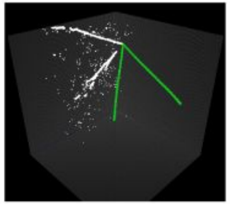
UResNet (UNet + ResNet + Sparse Conv.) as the **backbone feature extractor**

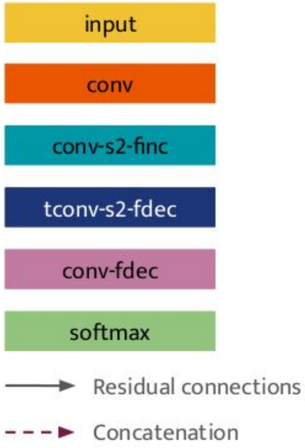
Input





Features





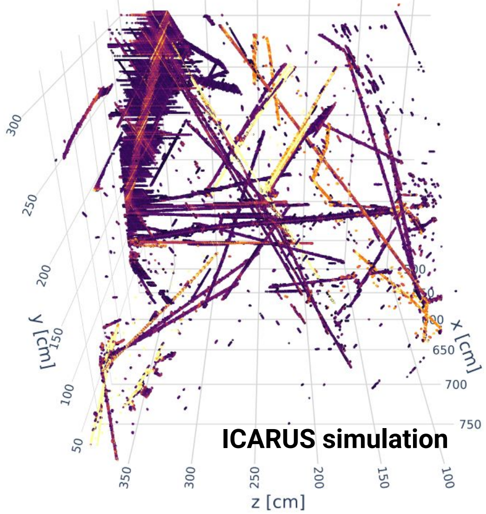
Paper: PhysRevD.102.012005

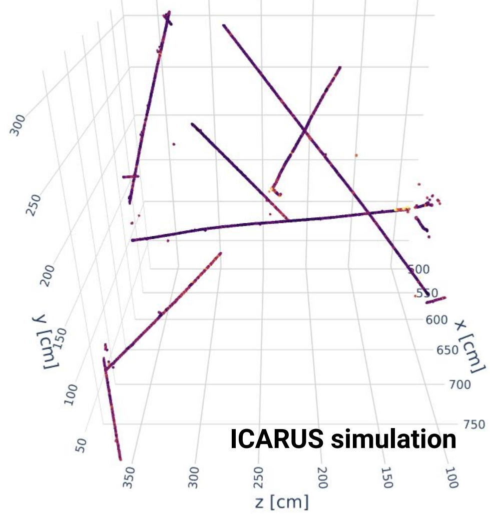
Tomographic Reconstruction

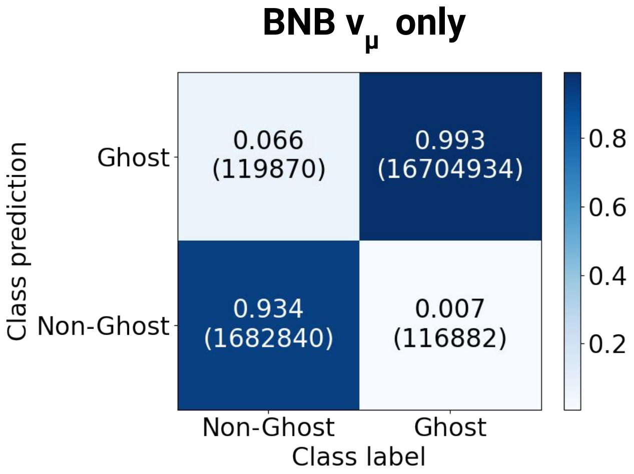
Ghost buster

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

- Find **valid combinations of 2D hits**: legitimate + artifacts (**ghosts**)
- Classify artificial space points as such: **ghost removal** (busting)





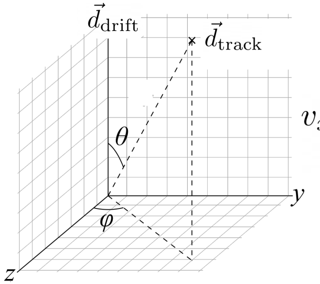


Tomographic Reconstruction

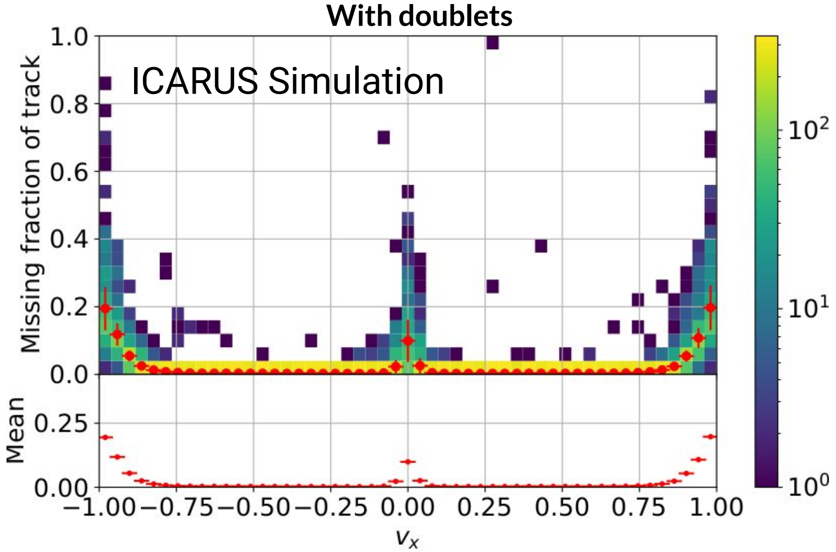
Track completeness

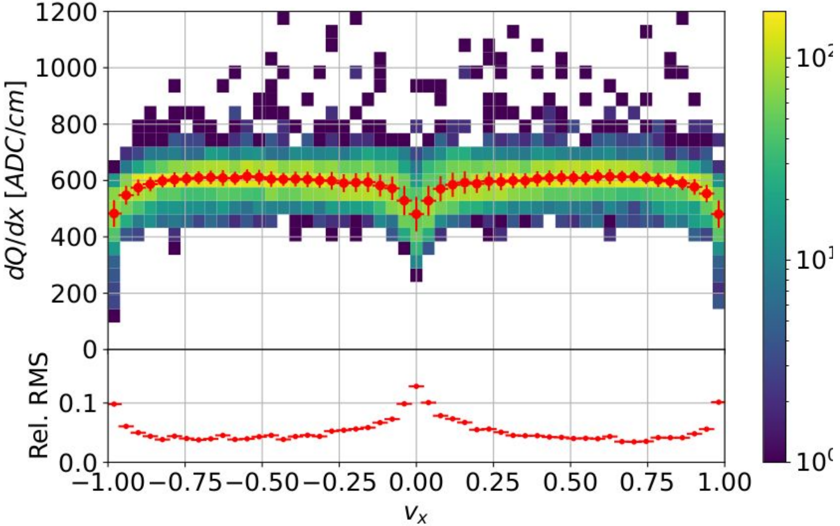
Definition: (total length of gaps)/(length of track)

- **Excellent track completeness with doublets**
- Overall **dQ/dx mostly flat w.r.t. angle**



$$v_x = \vec{d}_{\rm track} \cdot \vec{d}_{\rm drift} = \cos\theta$$





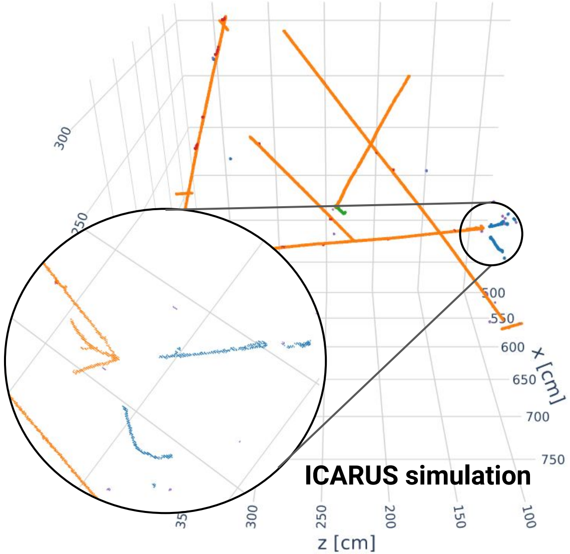
Semantic Segmentation

SLAC NATIONAL ACCELERATOR LABORATORY

Particle voxel class classification

Separate **topologically different** types of activity

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



BNB ν_μ only

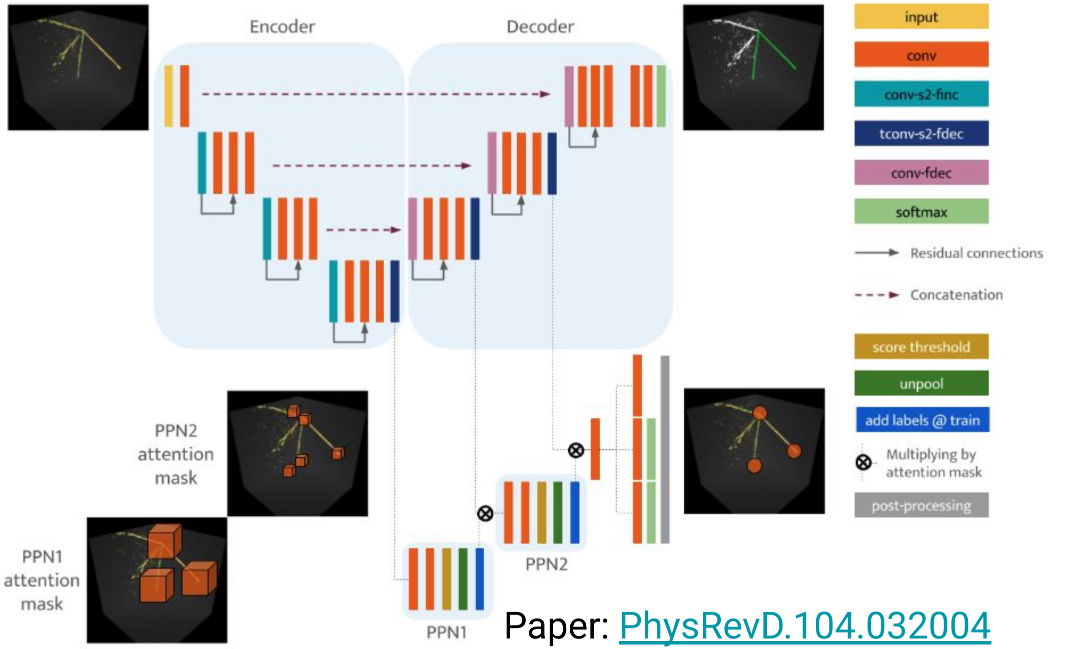
Paper: PhysRevD.102.012005

Point Proposal Network (PPN)

Architecture

The Point Proposal Network (PPN) uses 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



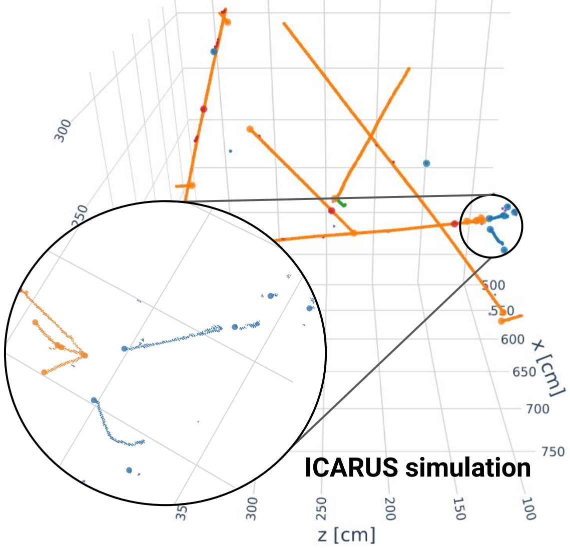
Paper: PhysRevD.104.032004

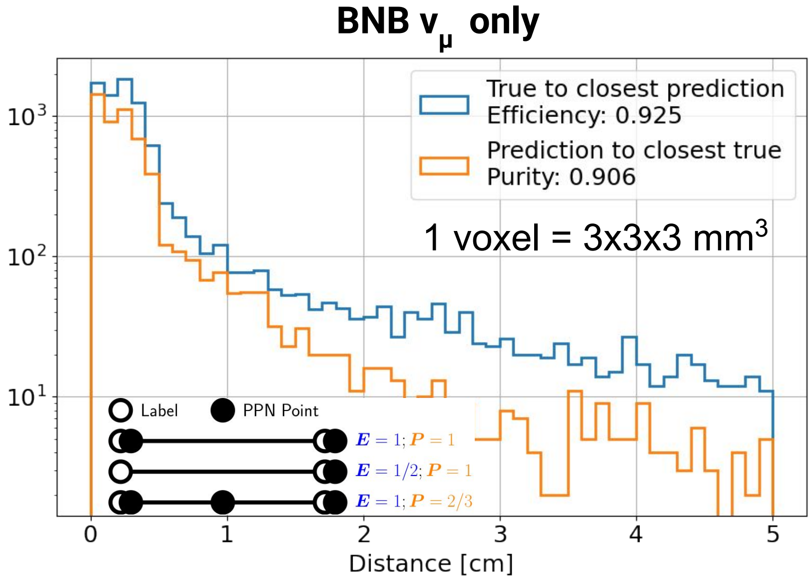
Point Proposal Network (PPN)

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

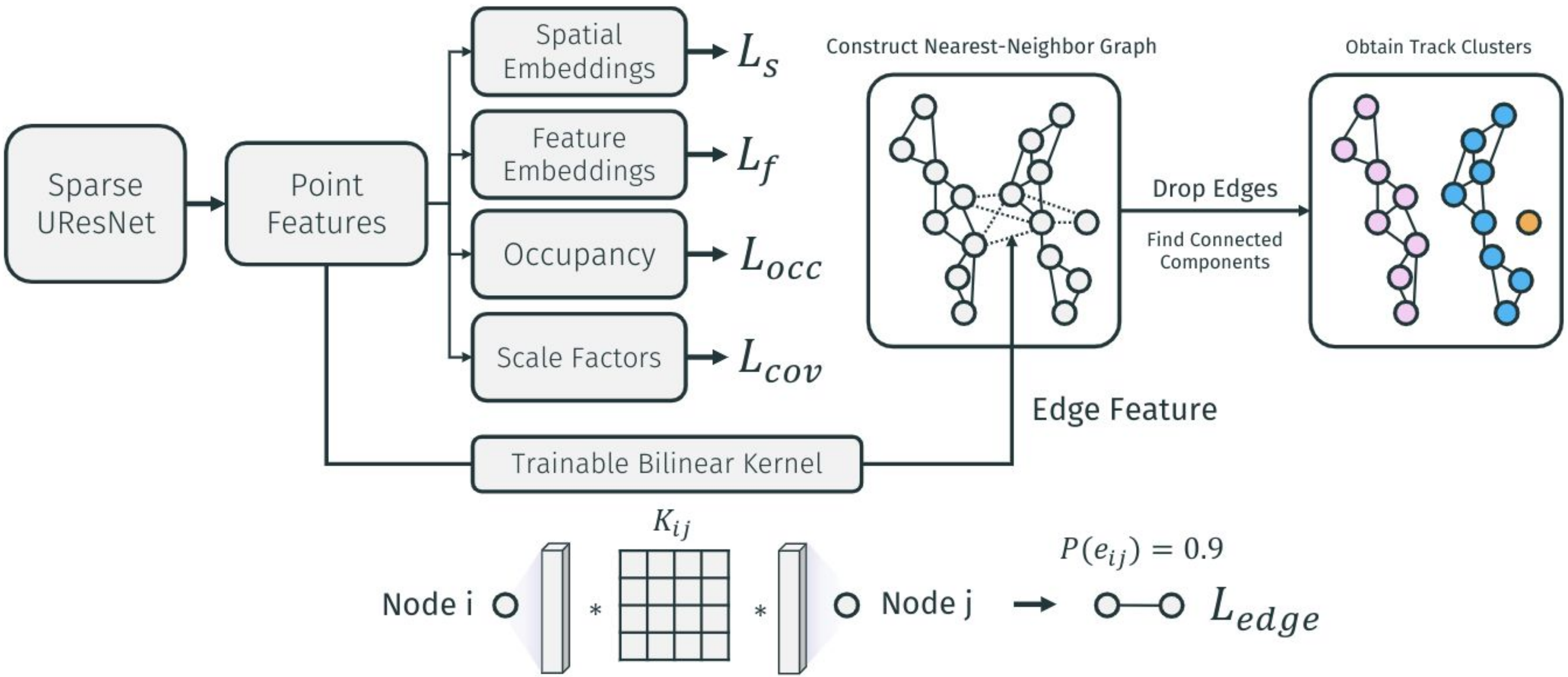




Paper: PhysRevD.104.032004

Graph-SPICE

Architecture

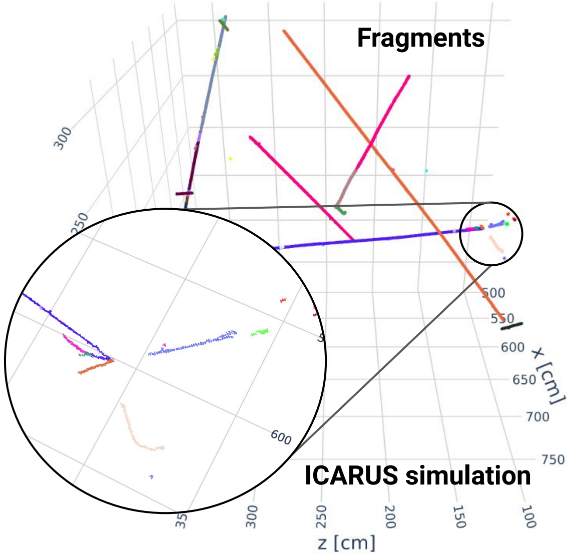


Dense Fragment Formation

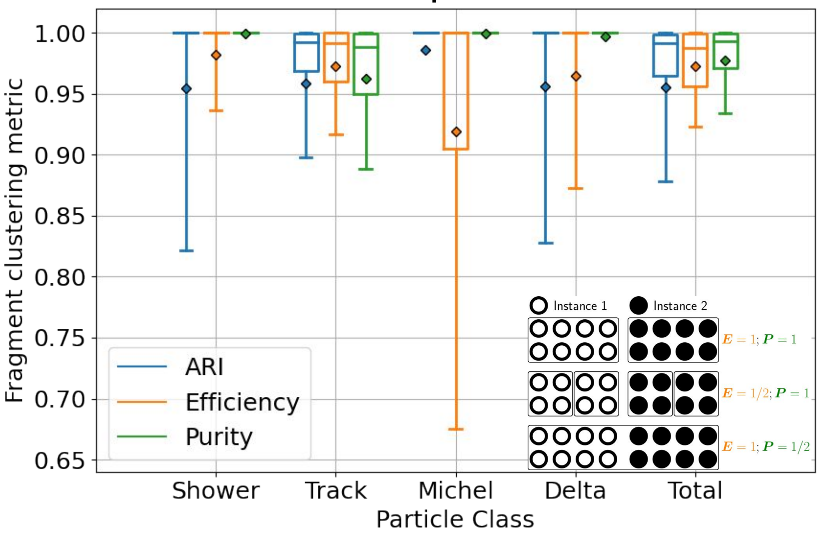
Spatial embedding transformation

Transform coordinates to an space in which tracks are spatially separated

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



BNB ν_μ only

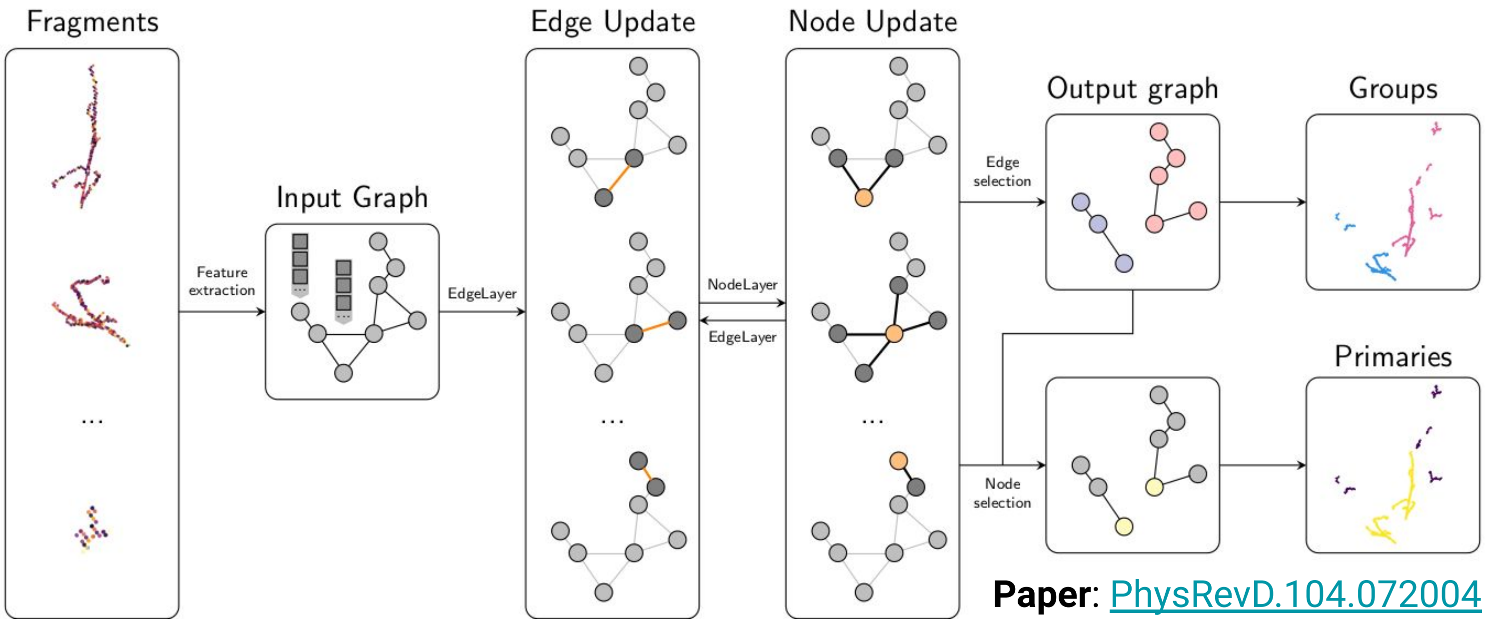


Paper: arXiv:2007.03083

Particle-Level Aggregation

Graph Particle Aggregator (GrapPA)

Graph Neural Network: fragments/particles (**nodes**), correlations (**edges**)



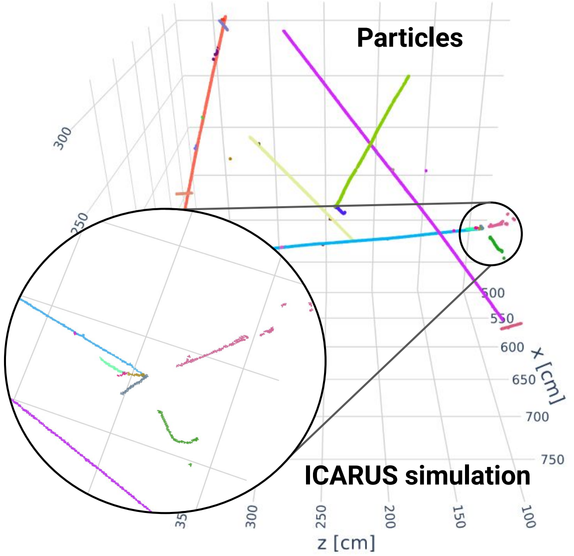
Paper: PhysRevD.104.072004

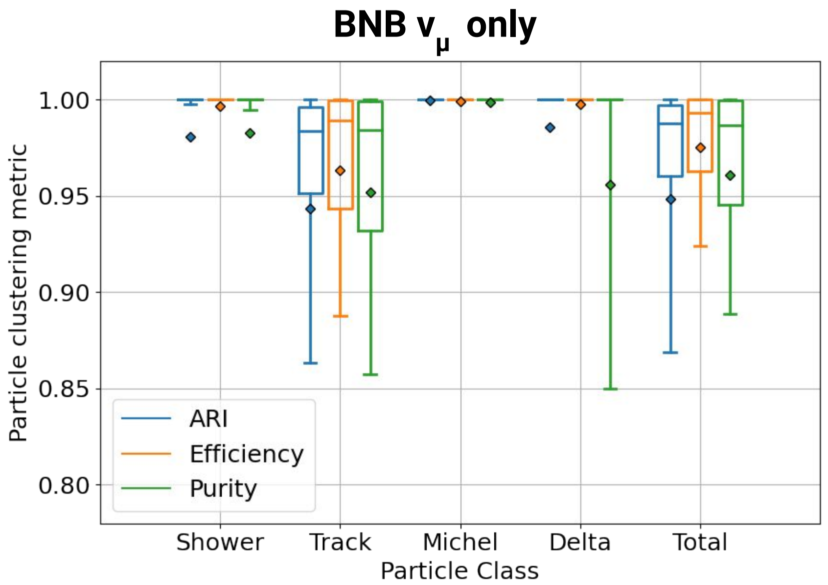
Aggregation

Graph edge classification

Two aggregation steps: **fragments** → **particles** → **interactions**

- **Select edges** in the graph that minimize loss, find **connected components**





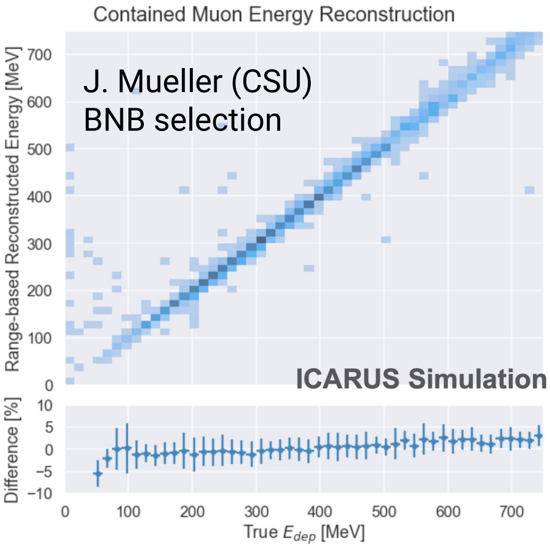
Paper: [PhysRevD.104.072004]

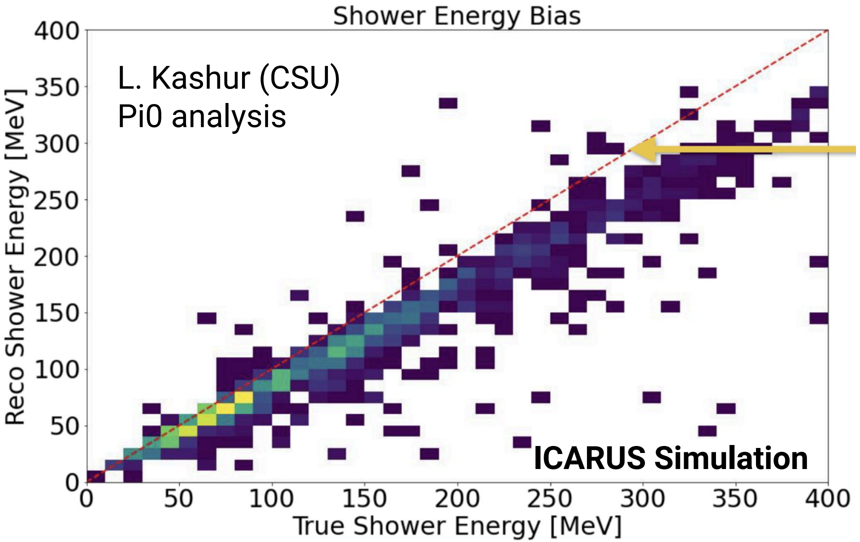
Energy reconstruction

Traditional techniques

Played around with regression NNs, but…

- **Range-based** momentum estimation of tracks is hard to beat
- **Calorimetric** energy reconstruction of showers is also hard to beat





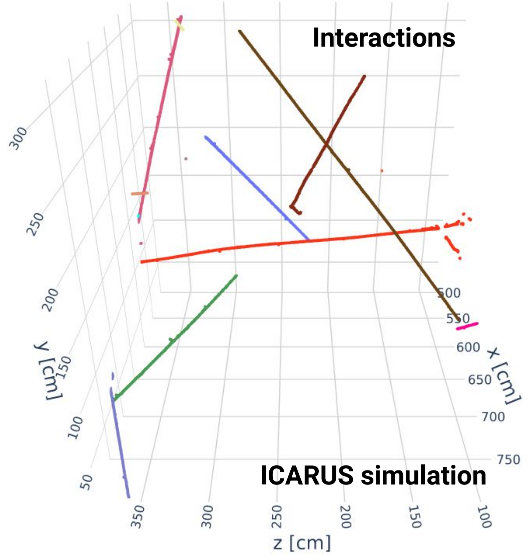
Offset is an overall scaling factor (tiny fragments missing)

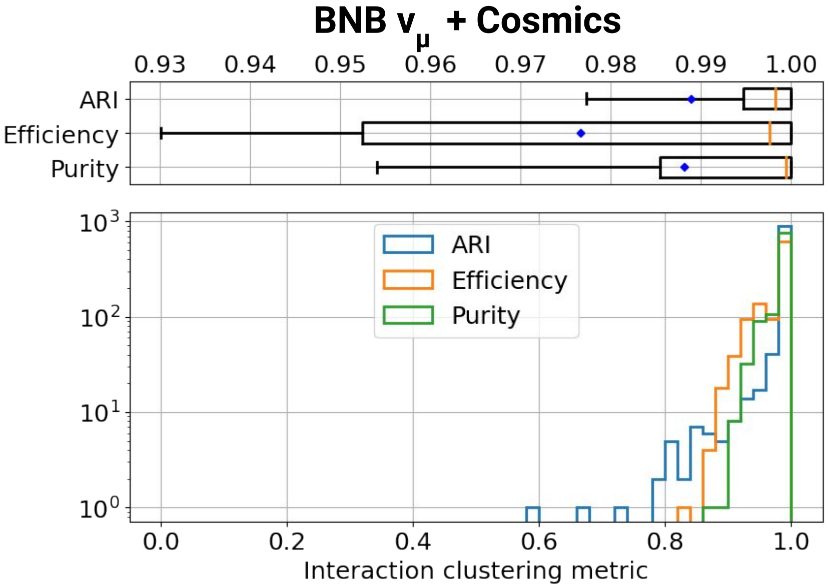
Aggregation

Graph edge classification

Two aggregation steps: **fragments** → **particles** → **interactions**

- **Select edges** in the graph that minimize loss, find **connected components**





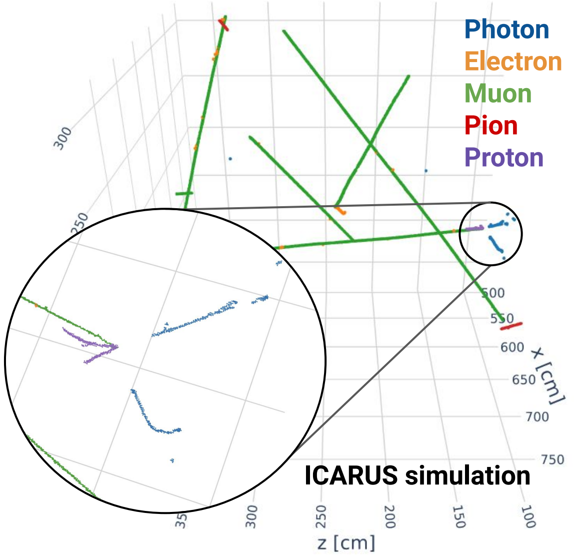
Paper: PhysRevD.104.072004

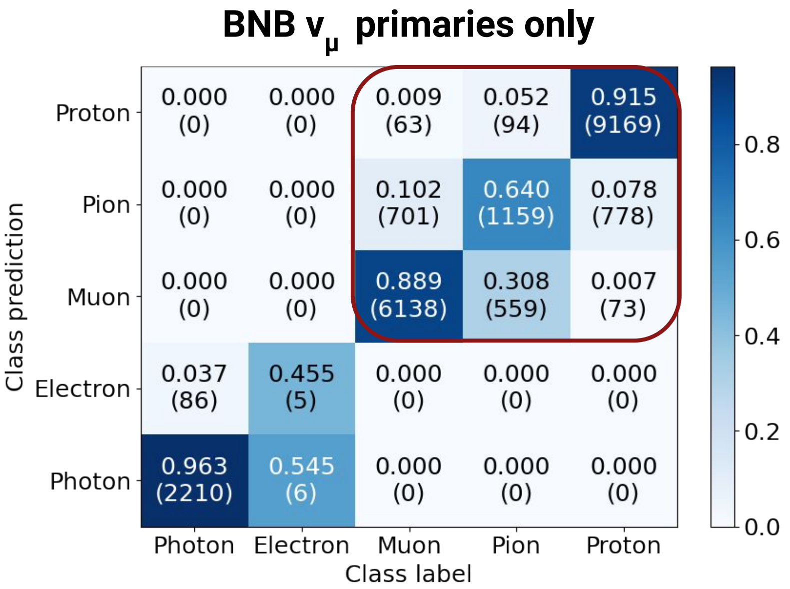
Particle Identification

Graph node classification

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

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



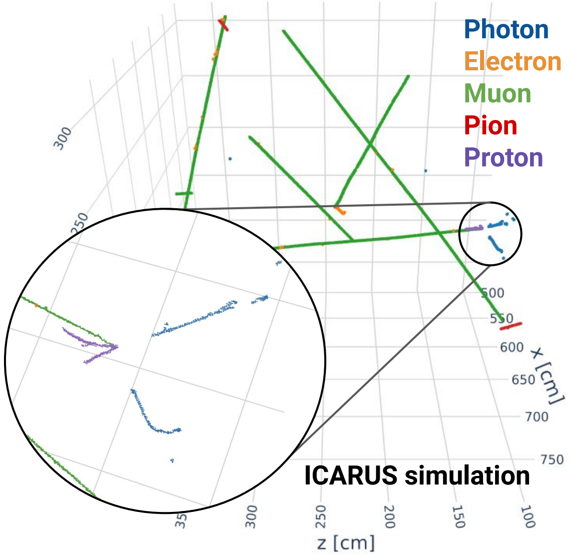


Particle Identification

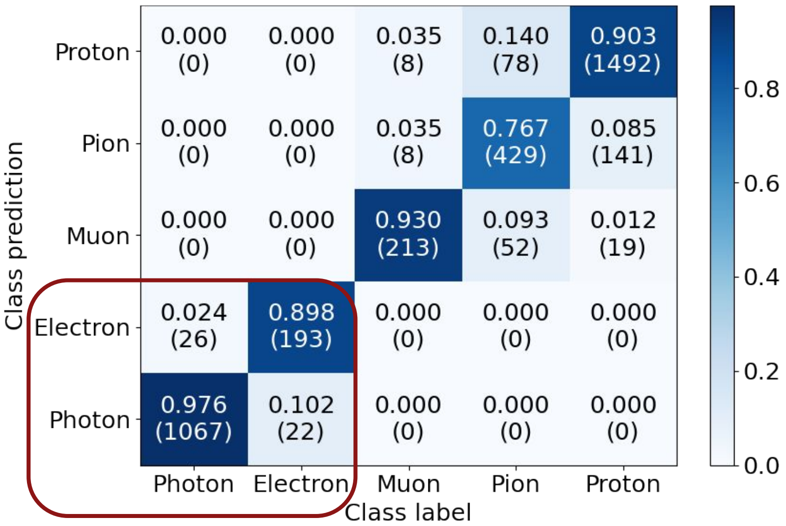
Graph node classification

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

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



Generic dataset (particle bombs)

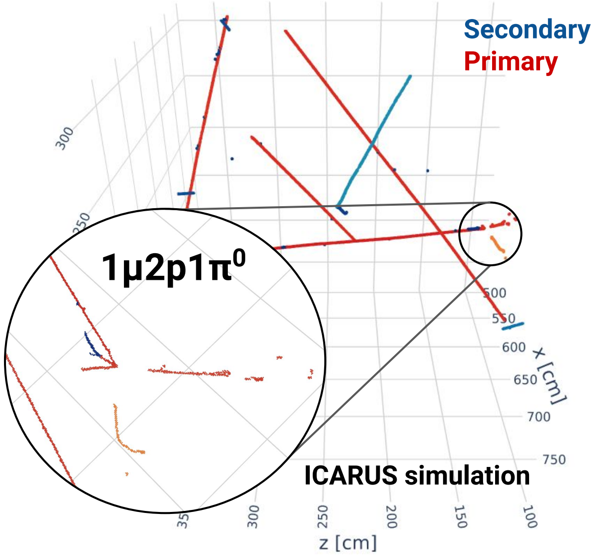


Primary Identification

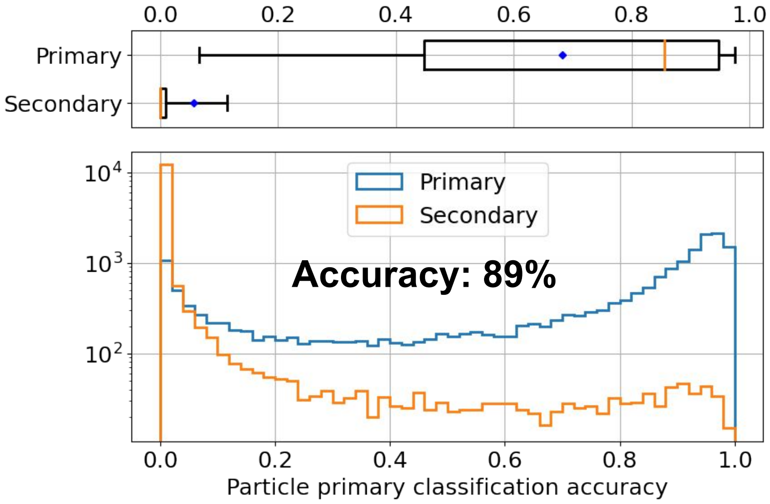
Graph node classification

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

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



BNB ν_μ primaries only

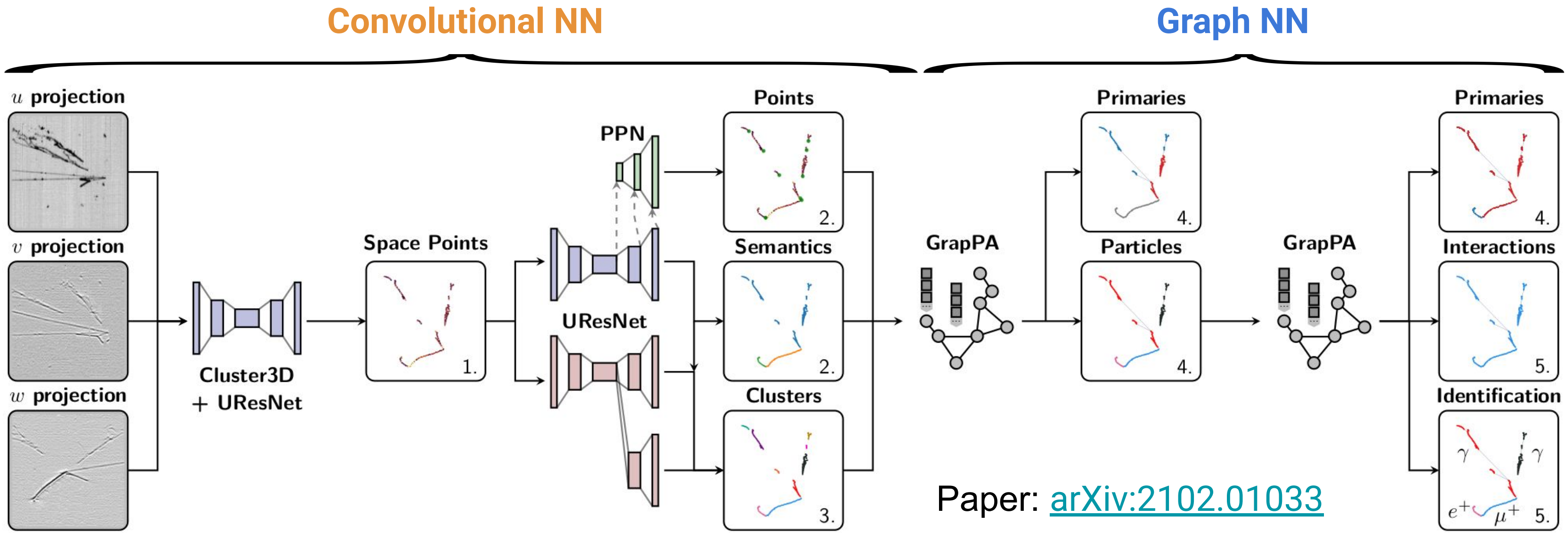


Reconstruction in LArTPCs

Full Reconstruction Chain Architecture

End-to-end ML-based reconstruction chain

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



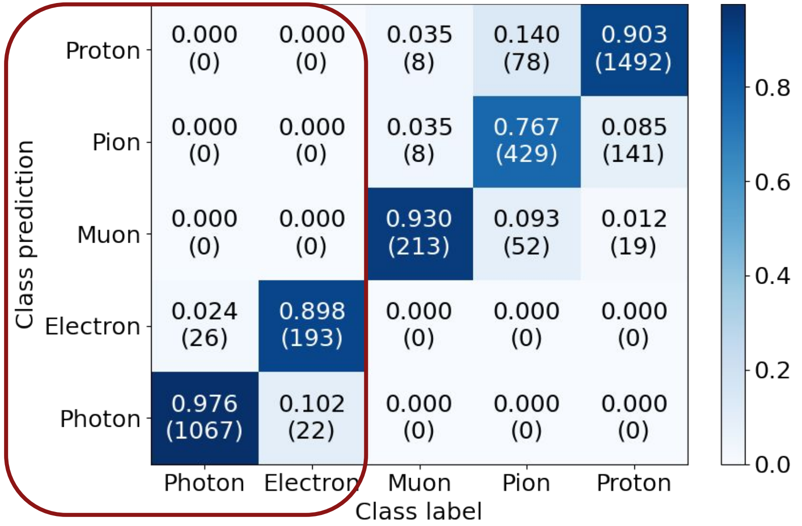
Paper: [arXiv:2102.01033](https://arxiv.org/abs/2102.01033)

S->ee searches

Challenges

What do we need for this search?

- Shower ID, e/gamma separation
  - Previous analyses see p->e confusion, **none in this ML scheme**

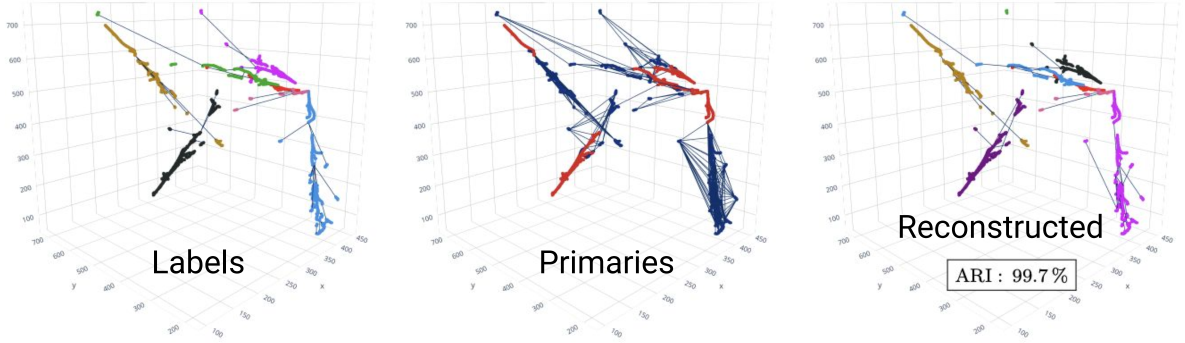


S->ee searches

Challenges

What do we need for this search?

- Shower ID, e/gamma separation
  - Previous analyses see p->e confusion, none in this ML scheme
- Reliable shower clustering (collinear showers are hard)
  - Need a dedicated study to see how far we can take it (J. Dyer)
  - Previous studies on pile-up indicate we do well at this



Generic simulation
6 showers in 10 m^3

Paper:
PhysRevD.104.072004

Low Energy Searches

Early days

Not much on this:

- Need to study blip reco. efficiency (secondary for oscillations)
- Some handle on low energy using Michel reconstruction as a benchmark
  - Currently much better than other techniques on the market

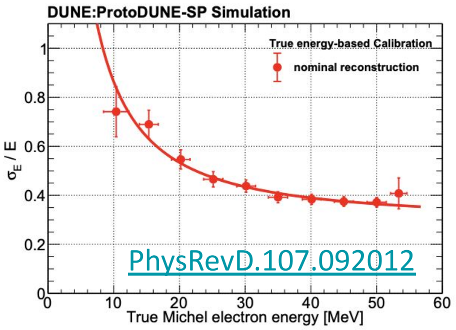
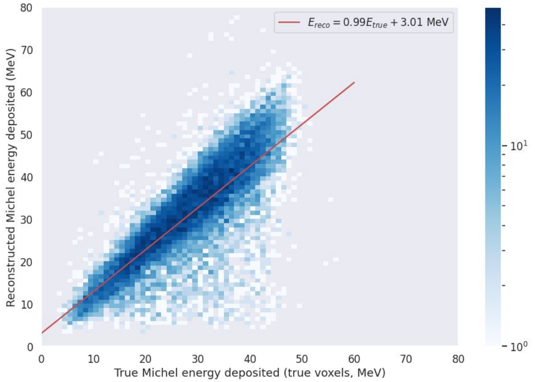
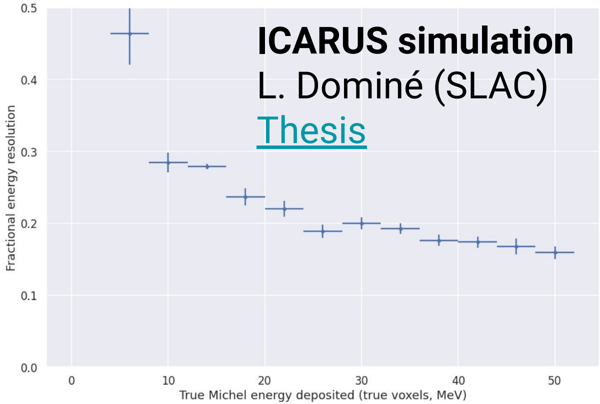


FIG. 10: Michel electron energy resolution as a function of Michel electron true energy using muon-based calibration (top); and the Michel electron energy resolution as a function of Michel electron true energy using true energy-based calibration (bottom).





ICARUS simulation
L. Dominé (SLAC)
Thesis

Reconstruction in LArTPCs

Scalability

A fair bit of work invested in speeding up the execution speed. On **ICARUS**:

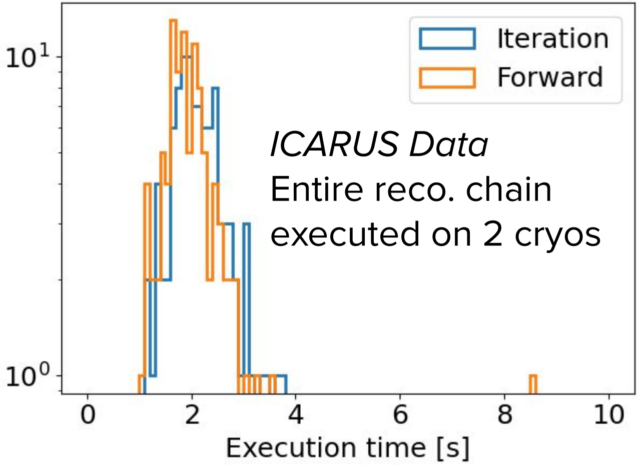
- ~2 M input space points/event
- O(1) M edges in the aggregator graphs
- **TPC reco: 2 s/event on an A100**

Very cheap to run on large datasets:

- **1 year of ICARUS beam-on data** can be reconstructed **in 1 day** with <200 A100s
- Perlmutter (NERSC): > 6000 A100s

Only **scales** with **space point count**

- Very cheap to run on DUNE-FD

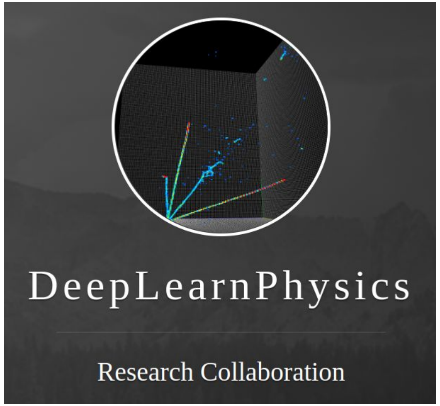


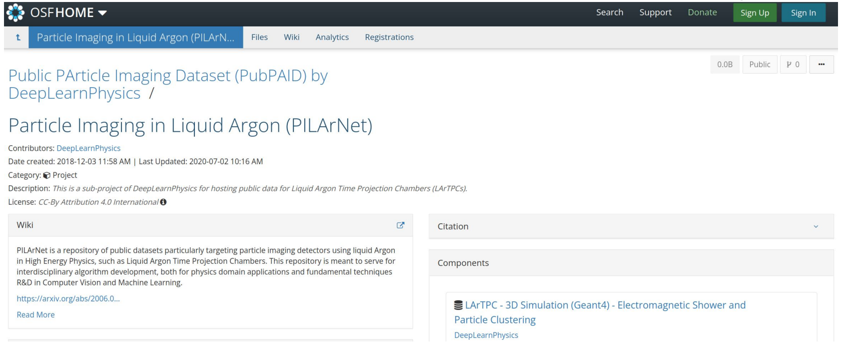
Reconstruction in LArTPCs

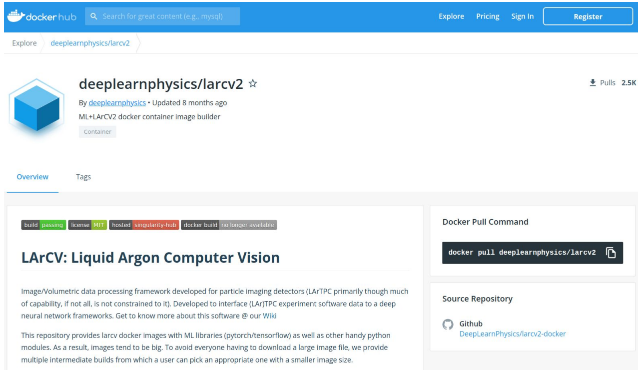
Open-source ecosystem

DeepLearnPhysics collaboration (ML techniques R&D)

- Public LAr simulation
  - Potential for open real data from prototypes
- Shared software dependencies with Docker/Singularity
- Open reconstruction software on GitHub
- Fully **reproducible results**
  - Readers have reproduced PhysRevD.102.012005







Other ML reconstruction efforts in LArTPCs

Very brief overview of the landscape

Many specific-purpose ML algorithms sprinkled in many places:

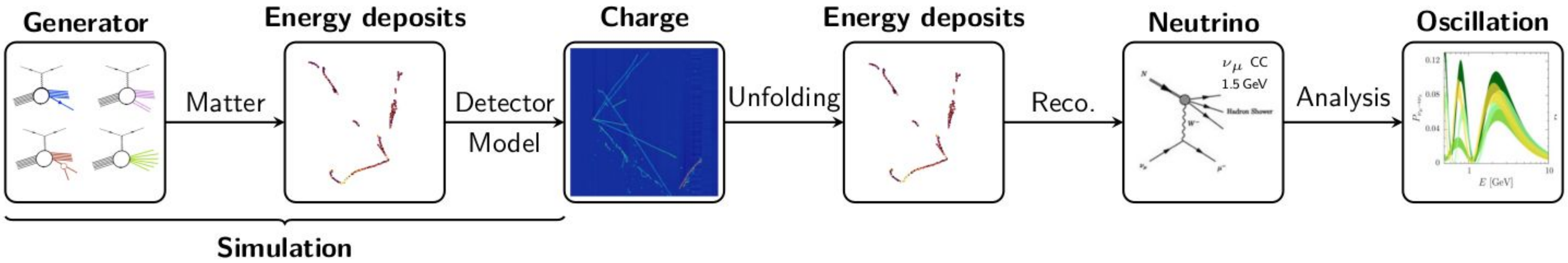
- BDTs are commonplace for S/B separation, particle type, etc.
- Some targeted ML-based effort to reconstruct the vertex location using CNNs
- Semantic segmentation in 2D in uBooNE (used in one of the LEE analyses)
- MPID algorithm based on CNNs in uBooNE (particle composition)
- CVN for image classification at the DUNE-FD

V. Hewes at U. Cincinnati leading another end-to-end effort based on GNNs from hits

- Early days, performance on semantic segmentation promising but worse
- Small networks (GNNs are shallower)
- Fast inference speed
- Nothing yet on clustering

Non-Reconstruction ML Efforts in LArTPCs

Future prospects



So far, we have tackled the **reconstruction challenge**, what's next?

- Can we go beyond “most likely” prediction and **quantify an uncertainty** ?
- Can we **mitigate differences** between simulation and data ?
- Can we **optimize detector modeling from data** and remove the issue altogether ?
- Can we **unfold detector effects** directly ? Yes, **learn inverse function automatically!**
- Can we **learn physics** (generators) **from data** ? Yes and no

Conclusions

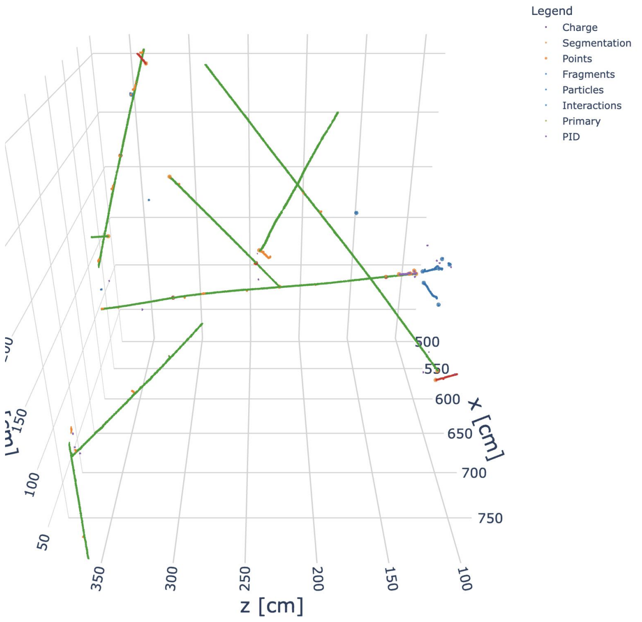
Takeaways

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

- Used on **ICARUS** sim./data and **DUNE-ND** (high neutrino pileup) sim. **today**
- Check out this ICARUS [interactive reconstructed event] !

In need of people to study performance on specific dark searches targets

- Promising early results on adjacent issues
- Many thanks to Jamie for her pioneering work in the ICARUS ML group



Backup Slides

GrapPA Aggregation Method

Edge selection procedure

What the network gives you:

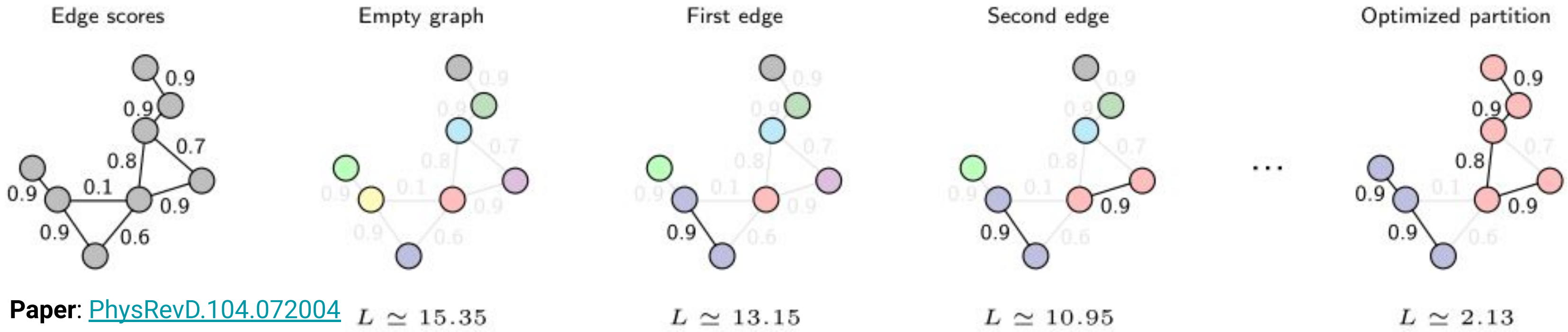
- Likelihood that an edge connects two objects in the same group

Target:

- Find the optimal partition

Method:

- Iteratively add the most likely edge to optimize CE loss



Paper: PhysRevD.104.072004

GrapPA Aggregation Method

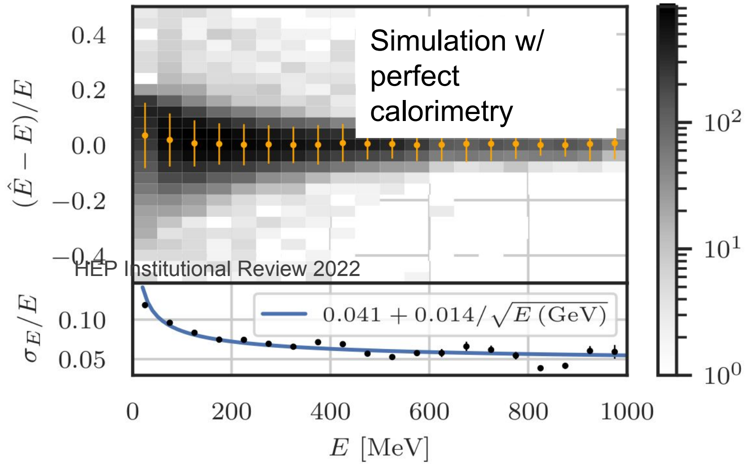
Shower energy reconstruction

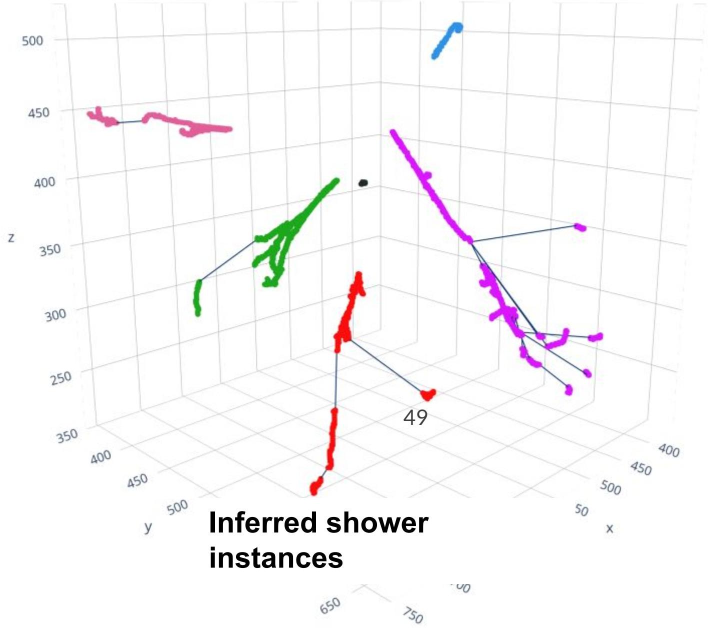
Identify **correlations** between shower fragments, **aggregate** them, identify **primaries**

Paper:

- [PhysRevD.104.072004]

Energy resolution at 1GeV: 5.5 %





Inferred shower instances

GrapPA Aggregation Method

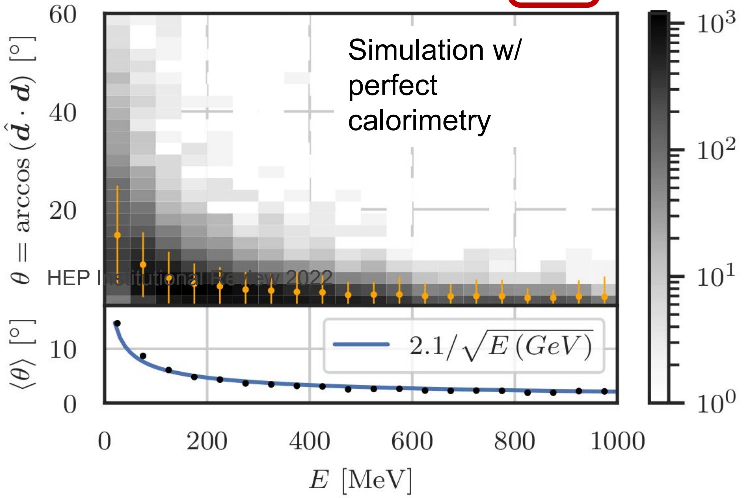
Shower angle reconstruction

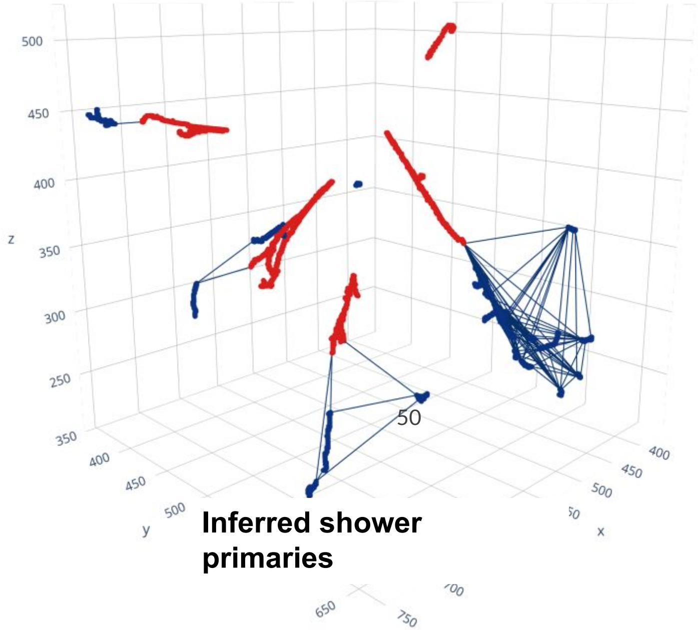
Identify **correlations** between shower fragments, **aggregate** them, identify **primaries**

Paper:

- [PhysRevD.104.072004]

Angular resolution at 1GeV: $\boxed{2.1°}$



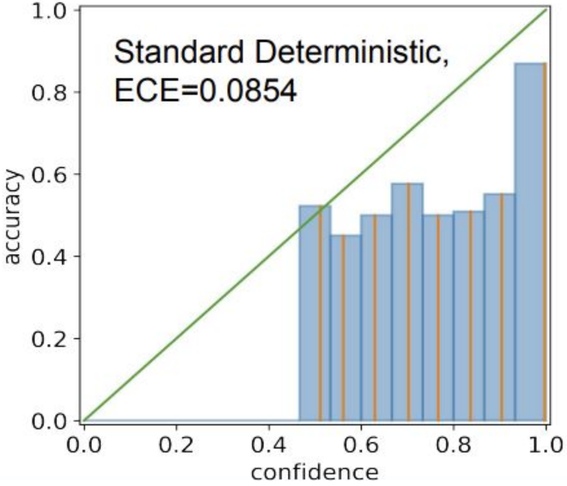


Uncertainty Quantification

Overview

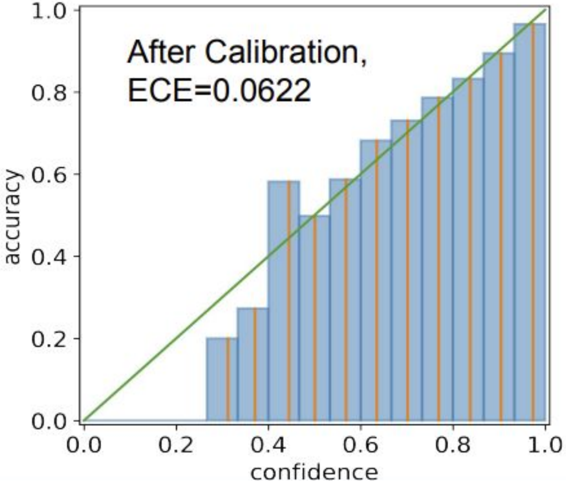
Goals of Uncertainty Quantification in Probabilistic Models:

- **Calibration**: Score p in [0,1] <=> probability p to be correct
- **Error detection**: Low confidence <=> large uncertainty



$$\hat{p}(x,T) = \max_i \frac{e^{l_i(x)/T}}{\sum_j e^{l_j(x)/T}}$$

Temperature Scaling



Uncertainty Quantification

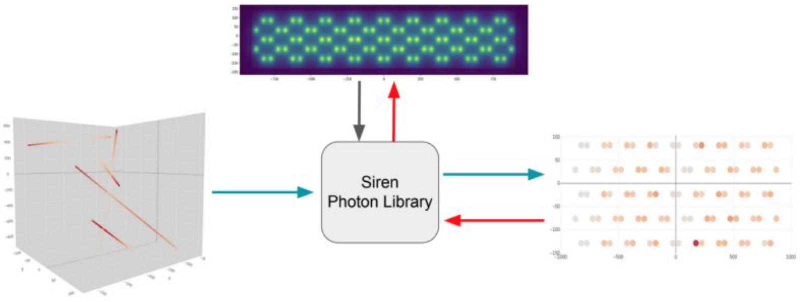
Photon visibility map

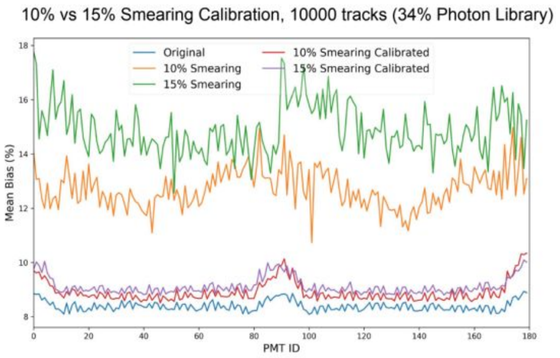
Can we make the **light simulation** differentiable ?

- Photon library maps x = (x, y, z) to visibility in each PMT (number of photons)
- Learn photon library using **scene representation** (SIREN): F (x, θ) differentiable

Calibration process: bias in library (offset in the actual visibility): θ ' = θ + δ

- Compare observed visibility to predicted visibility, use gradient descent to find θ ' !



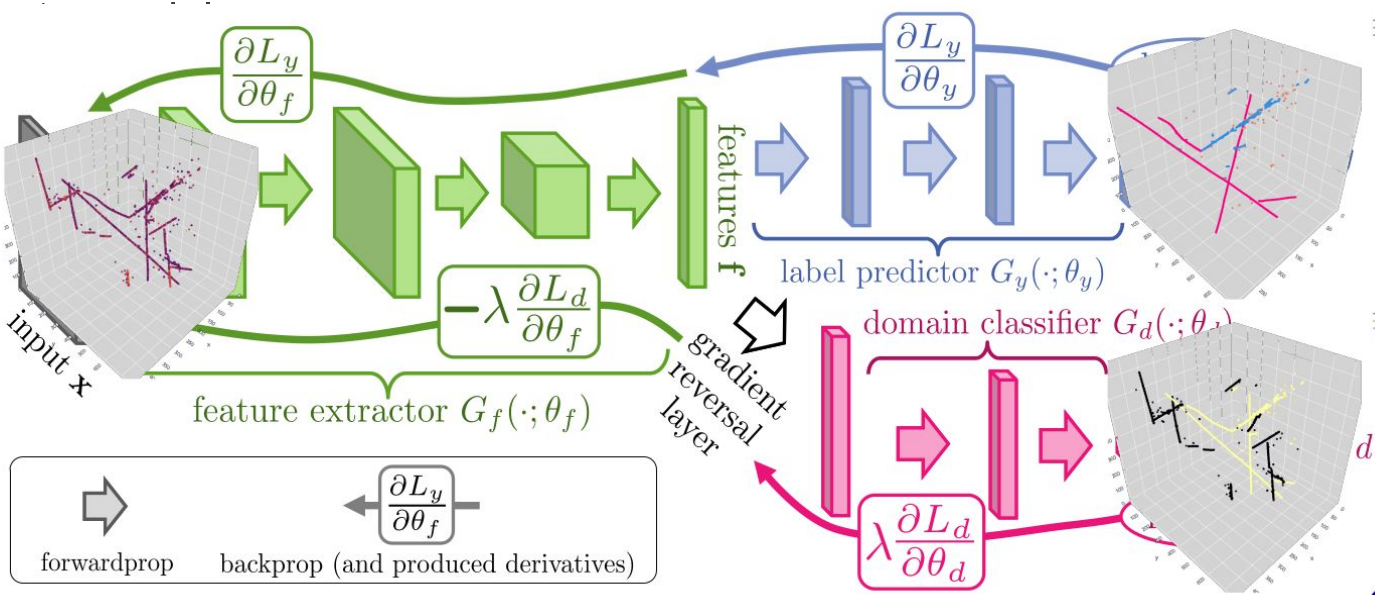


Domain Adversarial Training

Overview

Basics of **Domain Adversarial Networks**:

- Penalized for producing features that are different between

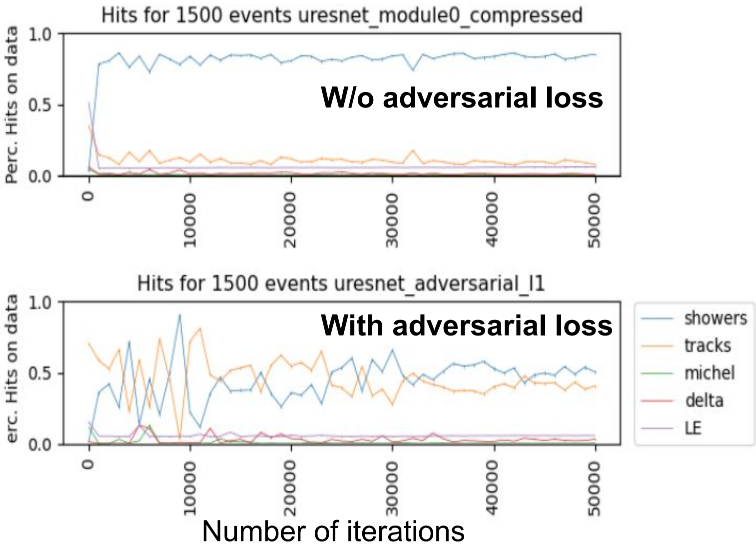


Domain Adversarial Training

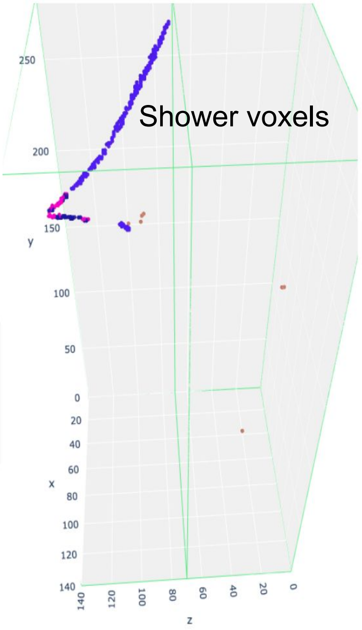
Overview

Basics of **Domain Adversarial Networks**:

- Penalized for producing features that are different between sim. and data



W/o adversarial loss



With adversarial loss