Wire-Cell: A High Quality Automated LArTPC Reconstruction for Neutrino Experiments

Haiwang Yu (BNL)

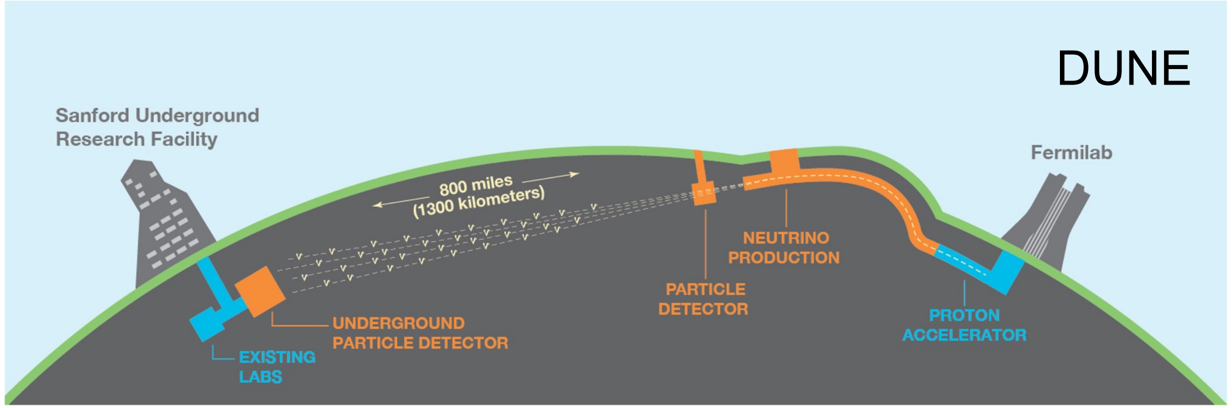
ACAT24

Stony Brook University

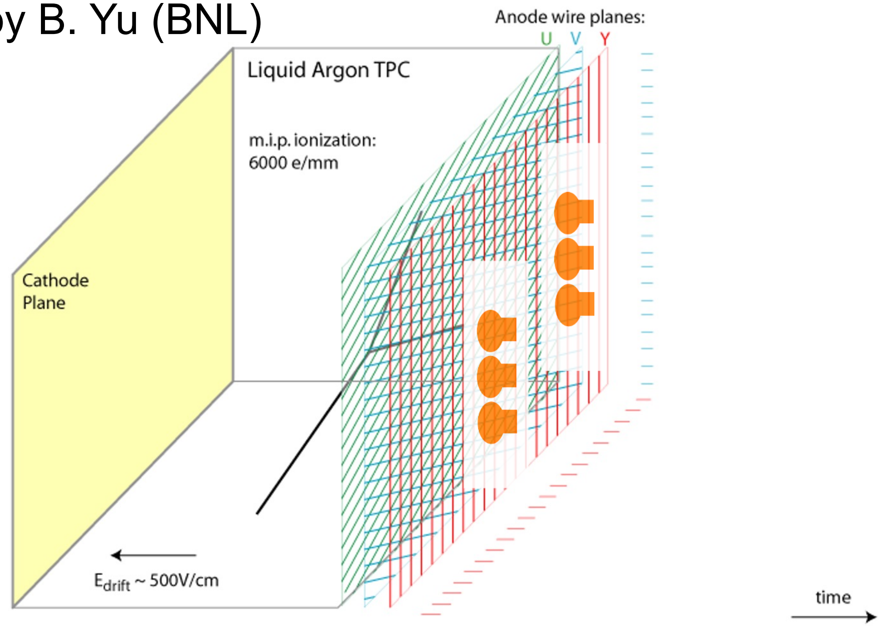
2024-03-11

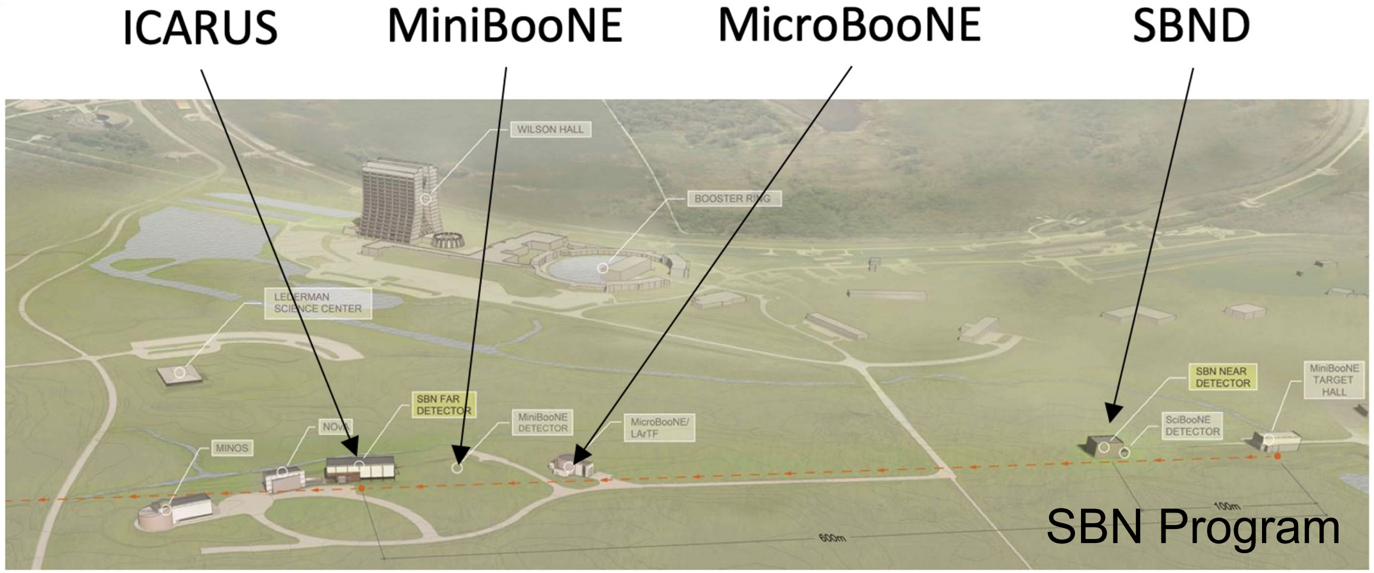
Liquid Argon TPC

- ~mm scale position resolution with multiple 1D wire readouts
- Particle identification (PID) with energy depositions and topologies



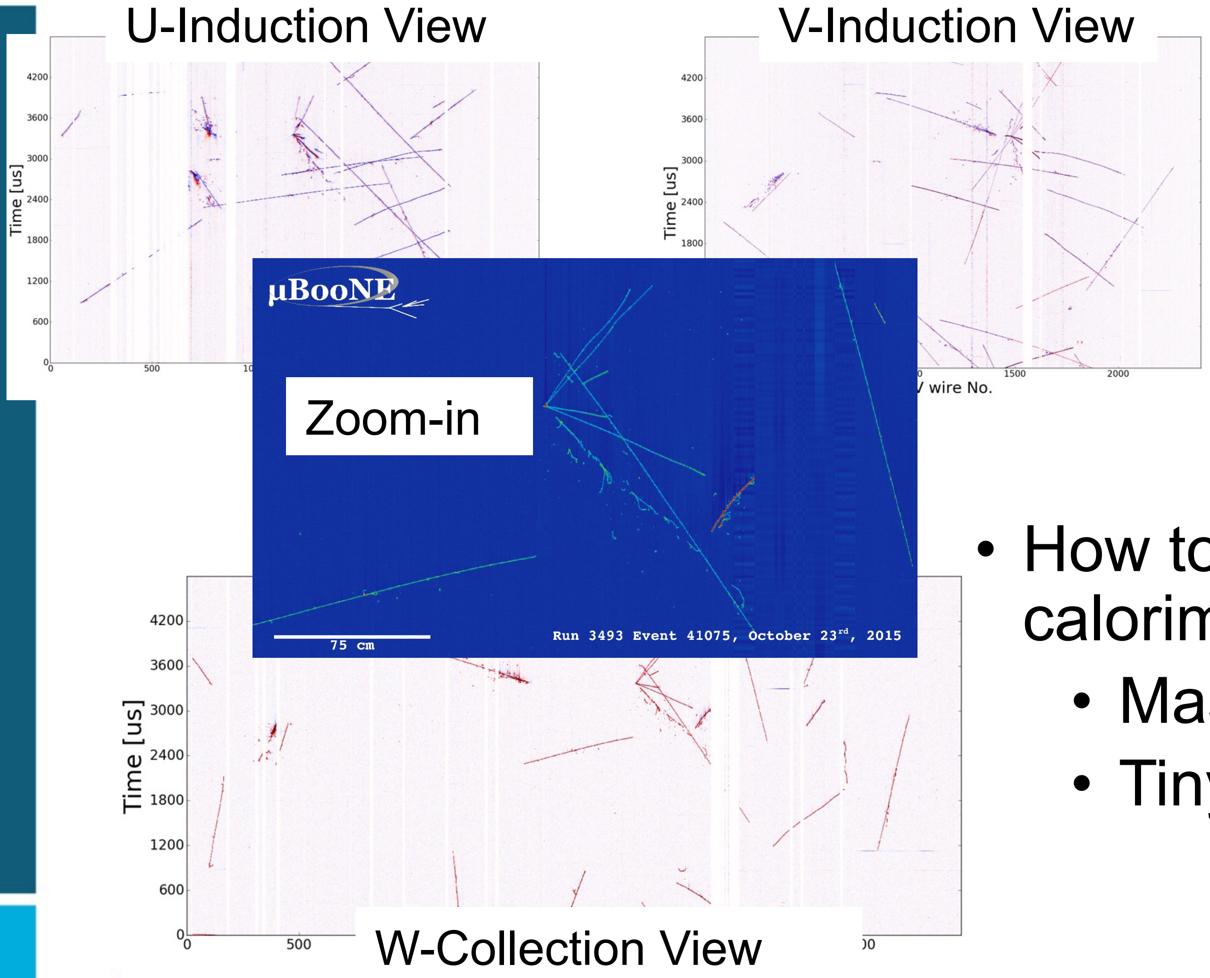
LArTPC Signal Formation Illustration
by B. Yu (BNL)

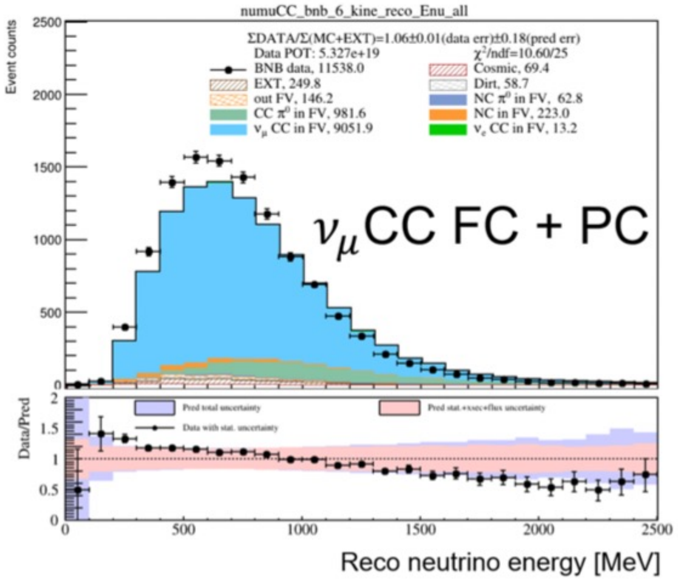




National Laboratory

Challenge in Automated Event Reconstruction





- How to convert the excellent resolution and calorimetry to rigorous physics results?
  - Massive amount of information across multi-scales
  - Tiny signal to background ratio

Wire-Cell Event Reconstruction

TPC simulation

noise filtering

signal processing

3D imaging

clustering

charge-light matching

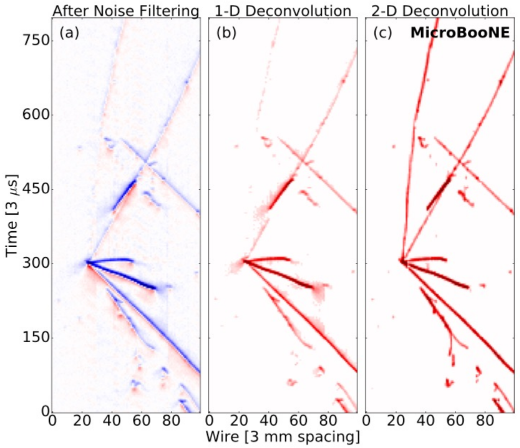
3D trajectory & dQ/dx fitting

cosmic muon tagger

multi-track fitting

DL-3D vertexing

particle identification

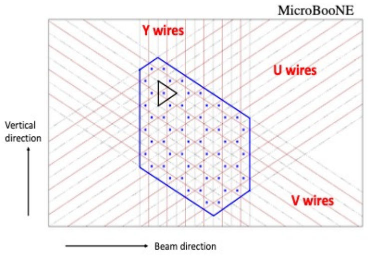


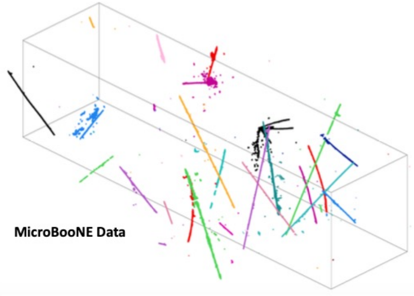
JINST 12 P08003 (2017)

JINST 13 P07006 (2018)

JINST 13 P07007 (2018)

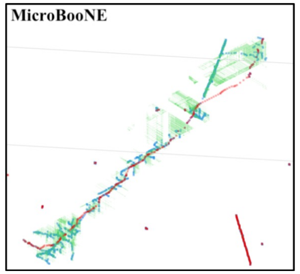
JINST 16 P01036 (2020)

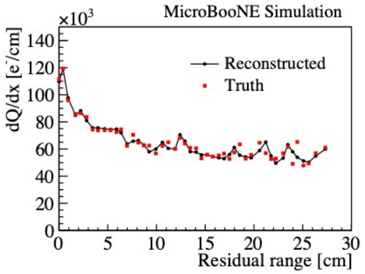




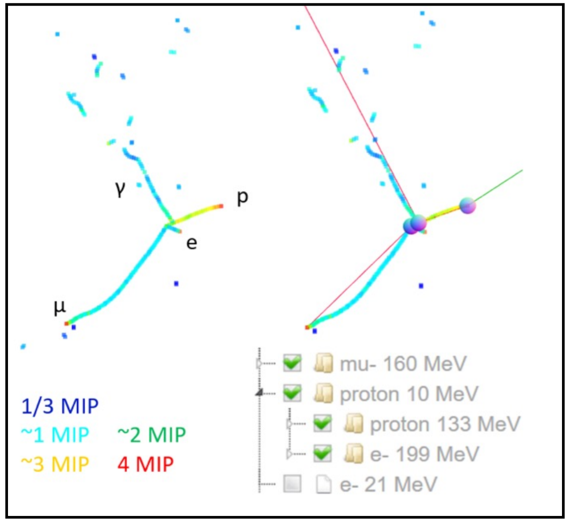
JINST 13 P05032 (2018)

JINST 16 P06043 (2021)





Phys. Rev. Applied 15, 064071 (2021)

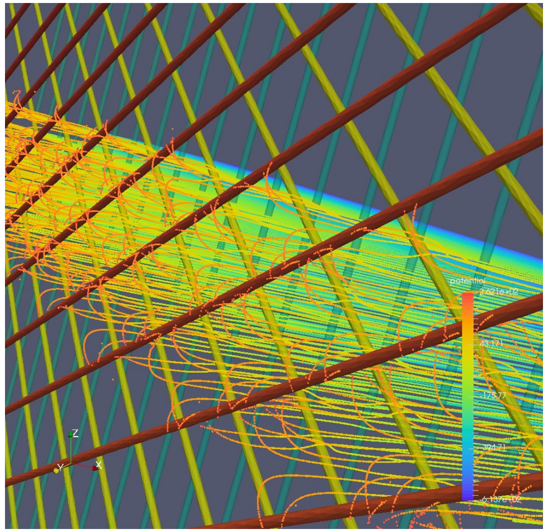


JINST 17 P01037 (2022)

Brookhaven National Laboratory

2D-Convolution based LArTPC Simulation

$$\text{Ramo's theorem: } i = -q\,\vec{E}_w \cdot \vec{v}_q$$

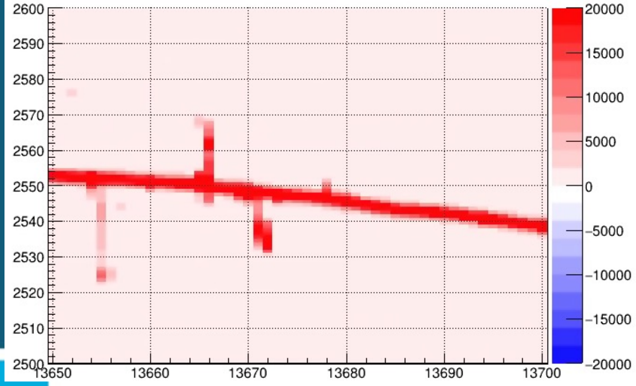


2D: approximate translational symmetry along the wire direction

LArTPC wire-readout measures induced charge ⊗ response

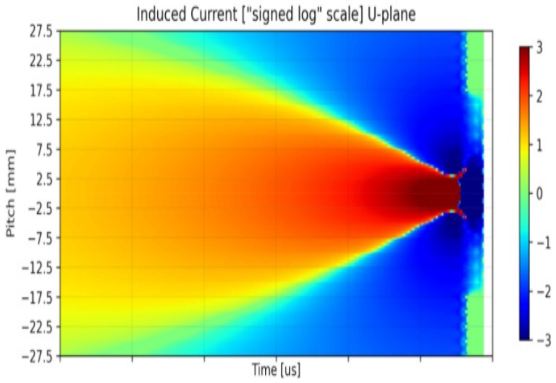
$$M(t',x') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(t,t',x,x') \cdot S(t,x)\,dt\,dx + N(t',x')$$

Energy depo + diffusion + rasterization



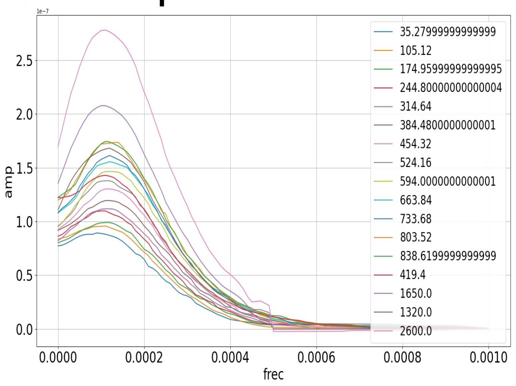
⊗

Long-range and position-dependent field



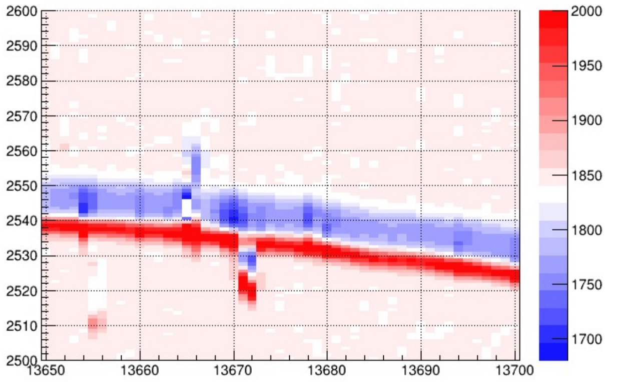
⊕

Noise Spectrum



⟶

Final Signal

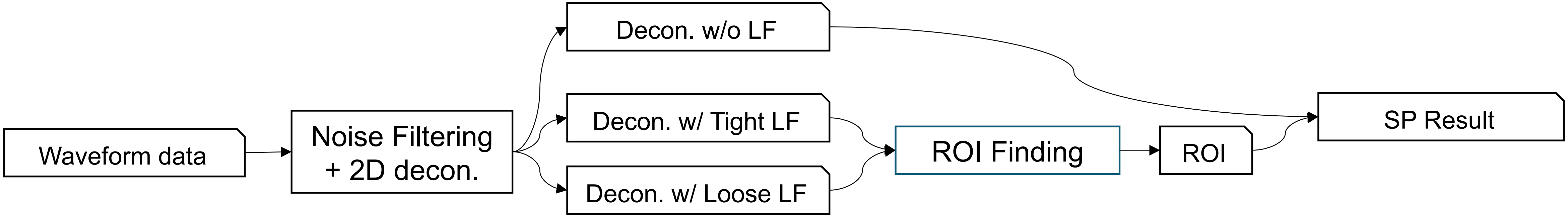


2D-Convolution based Signal Processing

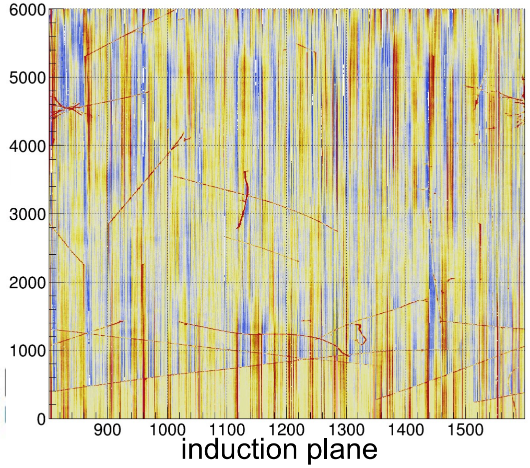
Signal Processing (SP) of LArTPC resolves charge from the original measurement:

$$S(\omega_t, \omega_x) \sim \frac{F(\omega_t, \omega_x) \cdot M(\omega_t, \omega_x)}{R(\omega_t, \omega_x)} \xrightarrow{IFT} S(t, x)$$

- *"2D deconvolution": assuming translational symmetry in the third dimension*
- *Utilize the signal/noise separation in both frequency and time domain*

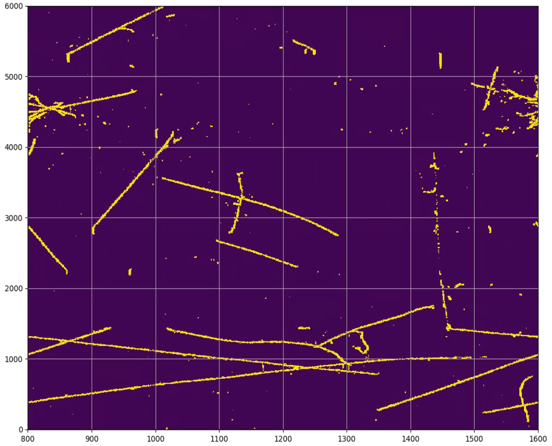


Decon. w/o LF (low frequency) filter
Waveform ⟶ charge, dense



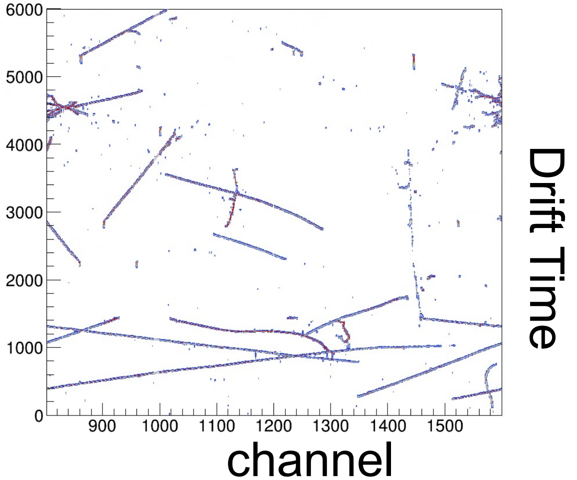
\+

ROI (region of interest):
Hit finding, sparsify



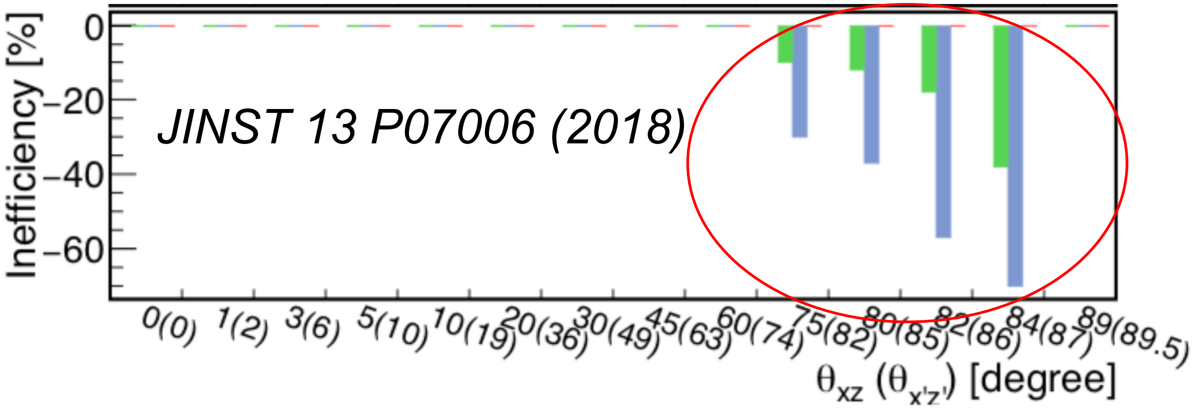
➡

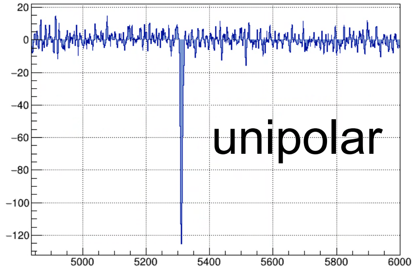
SP (signal processing) result:
Sparse, ionization charge

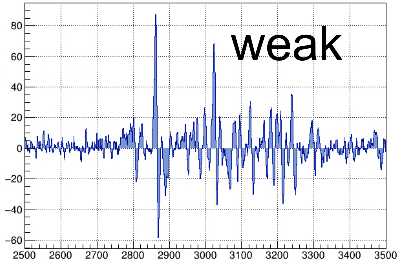


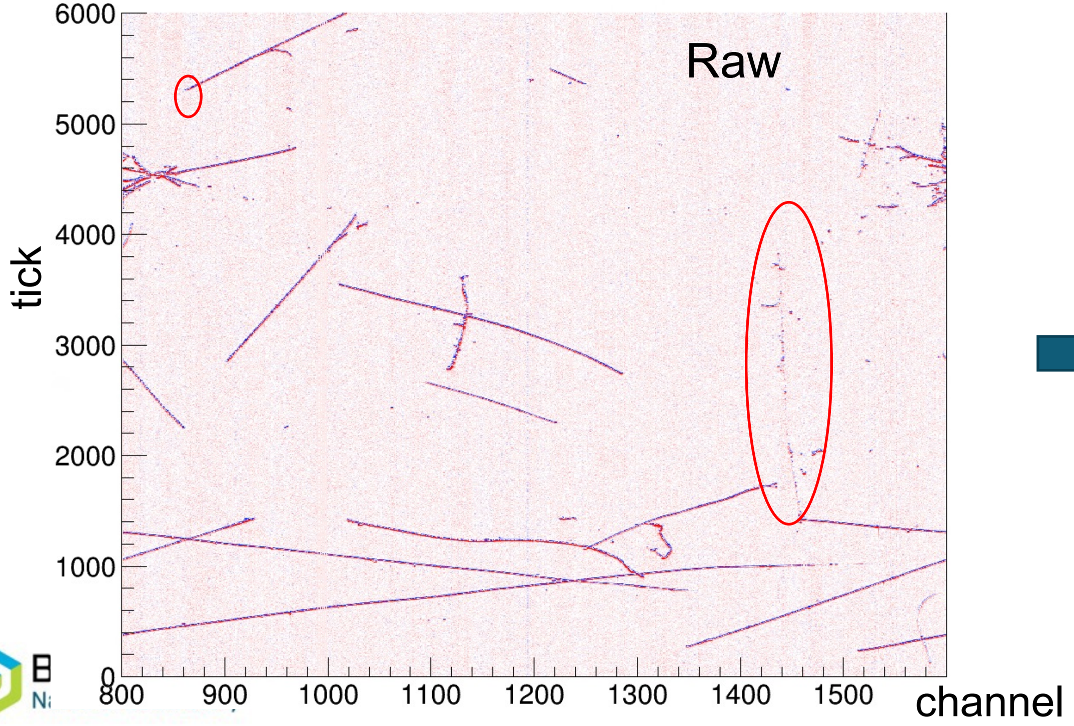
DNN ROI finding to improve LArTPC Signal Processing

- “Prolonged Track” – weak signal
- “Tear Drop” - distorted waveform
- Noisy dots - noise

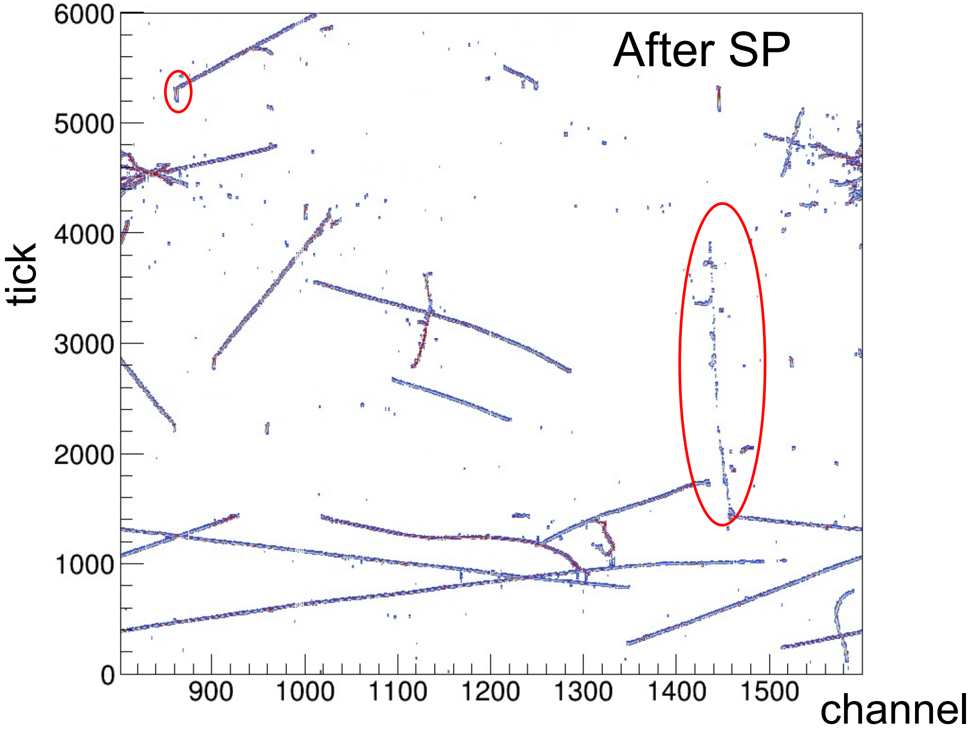








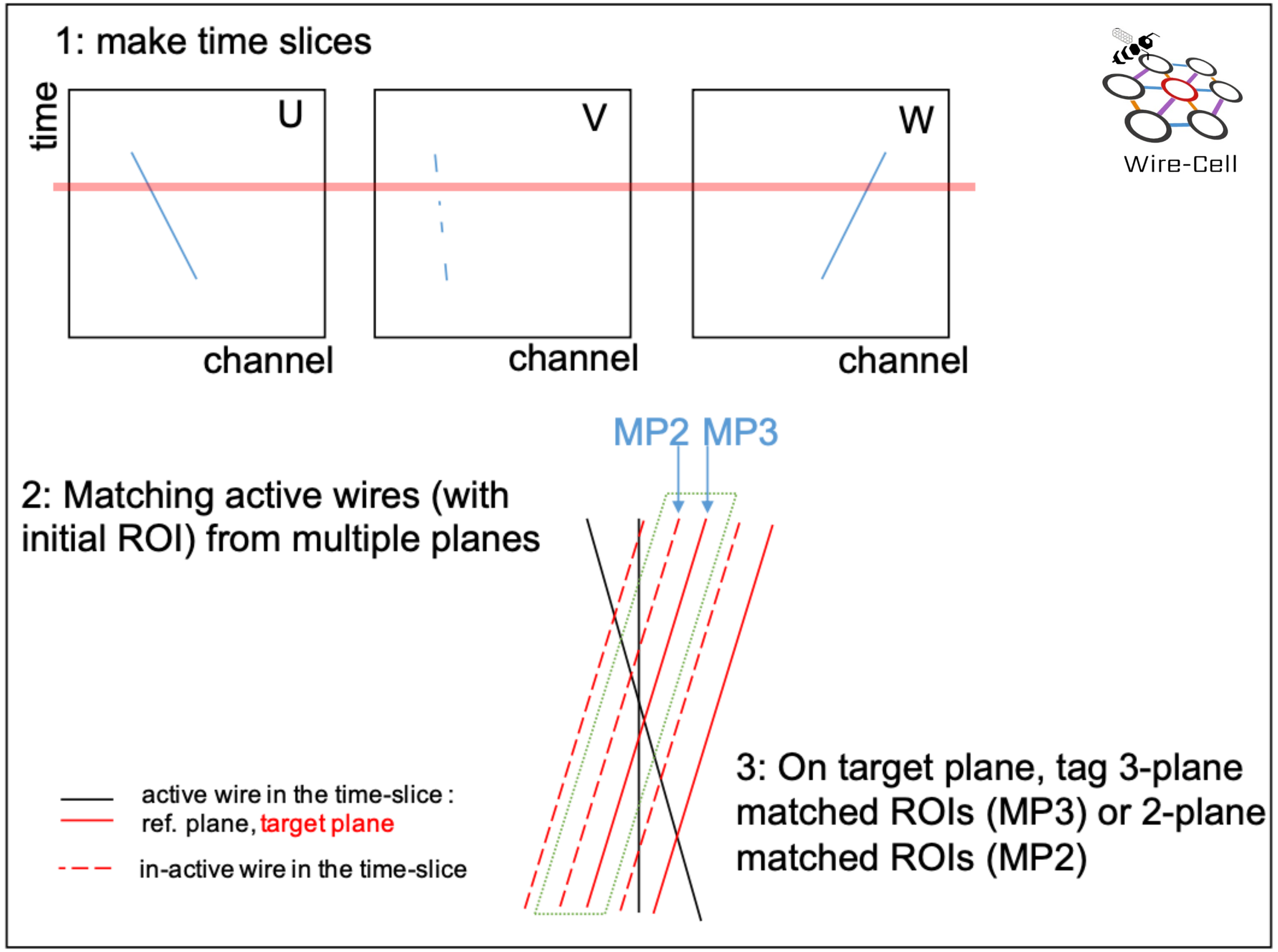
SP



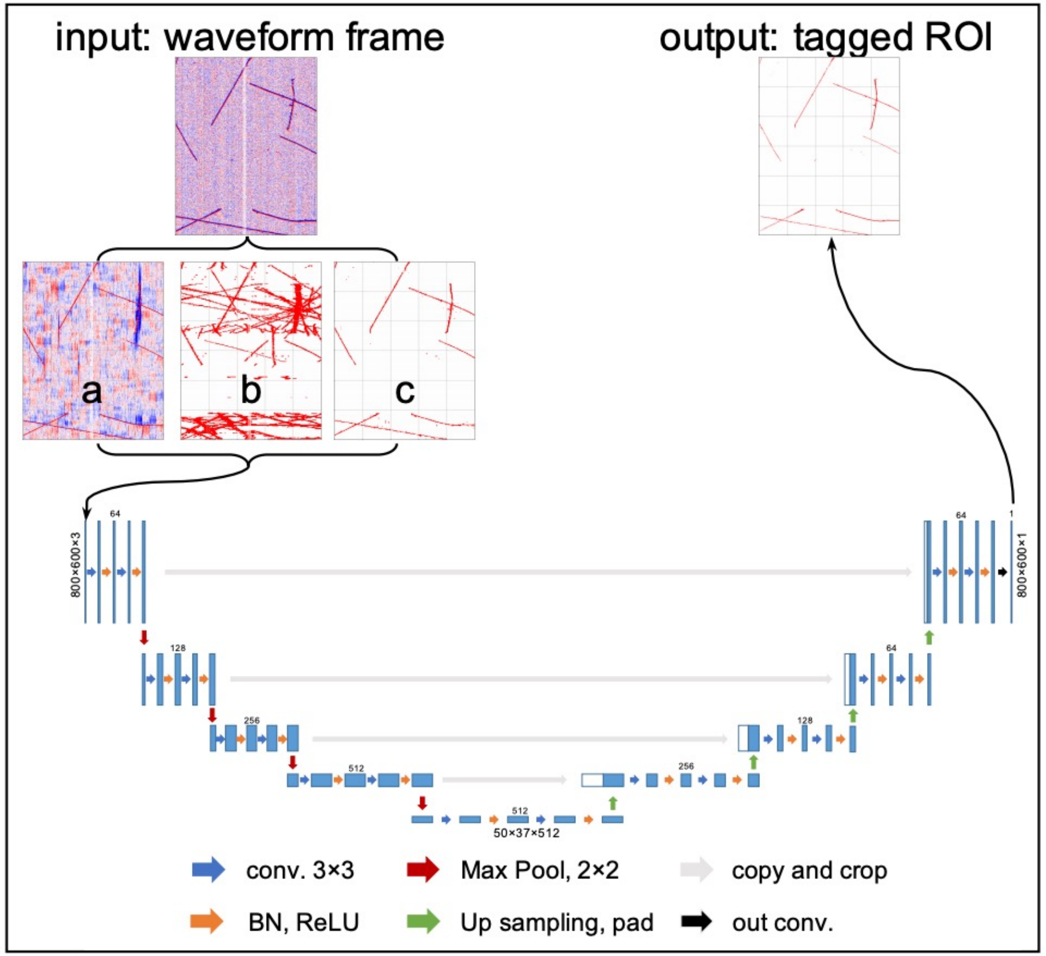
DNN ROI finding with multi-plane information

JINST 16 P01036 (2021)

Multi-plane information in Signal Processing

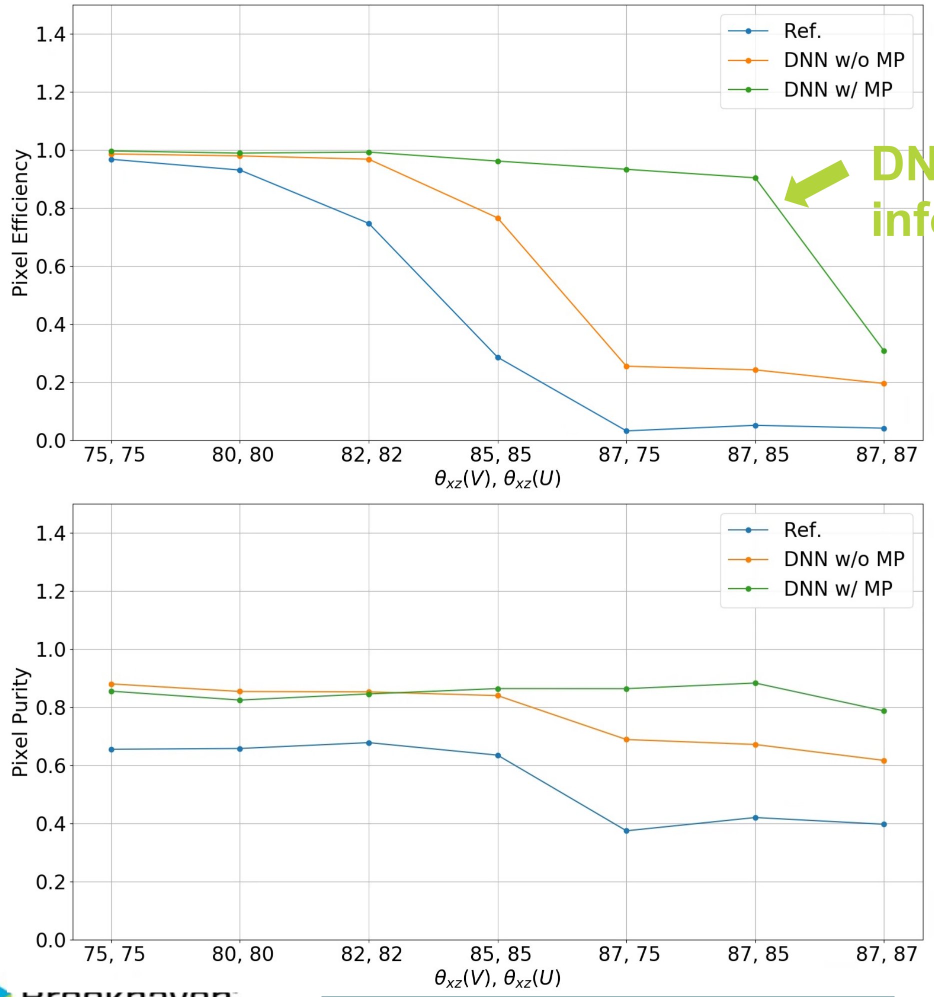


DNN ROI finding with multiple input channel



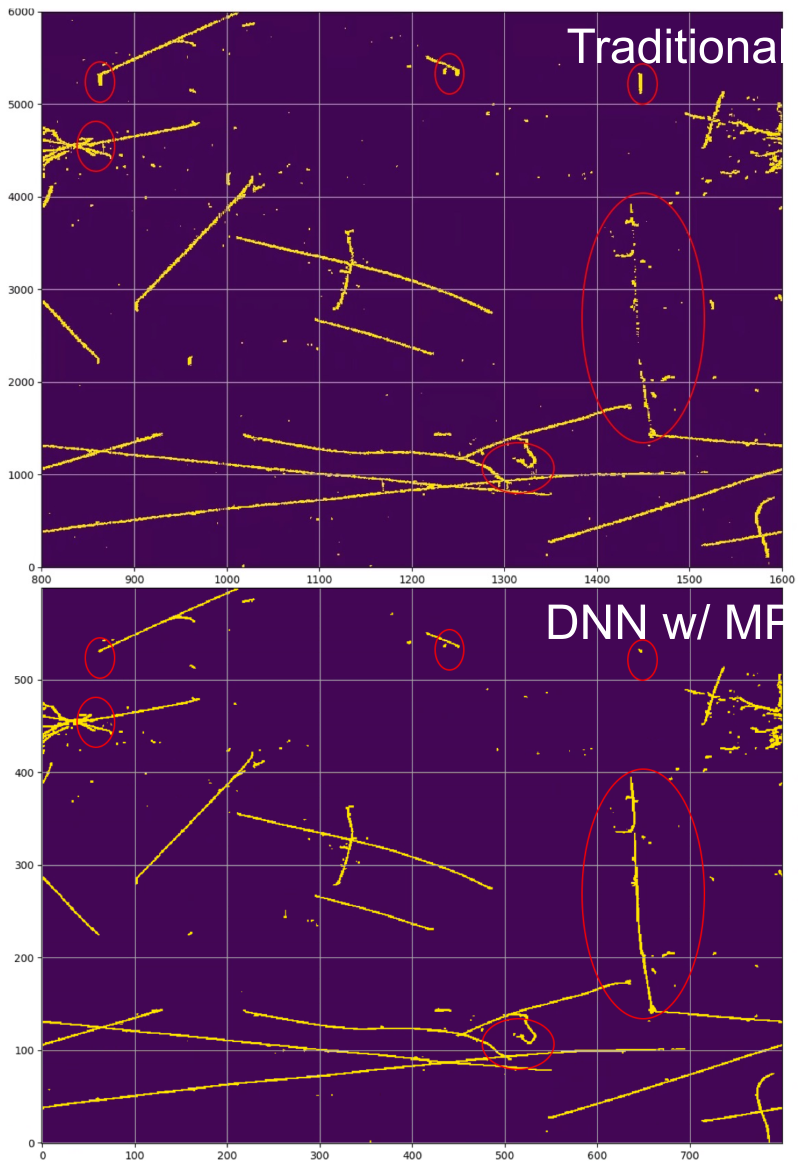
Trained using 450 images

ProtoDUNE simulation
ROI finding on V plane (2nd induction)



DNN With 3-plane information

tested on ProtoDUNE data



Brookhaven National Laboratory

[JINST 16 P01036 (2021)](#)

Wire-Cell Tomographic Event Reconstruction

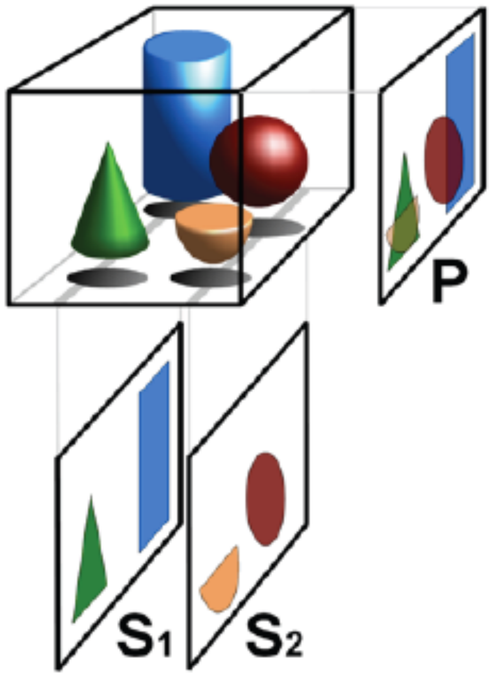
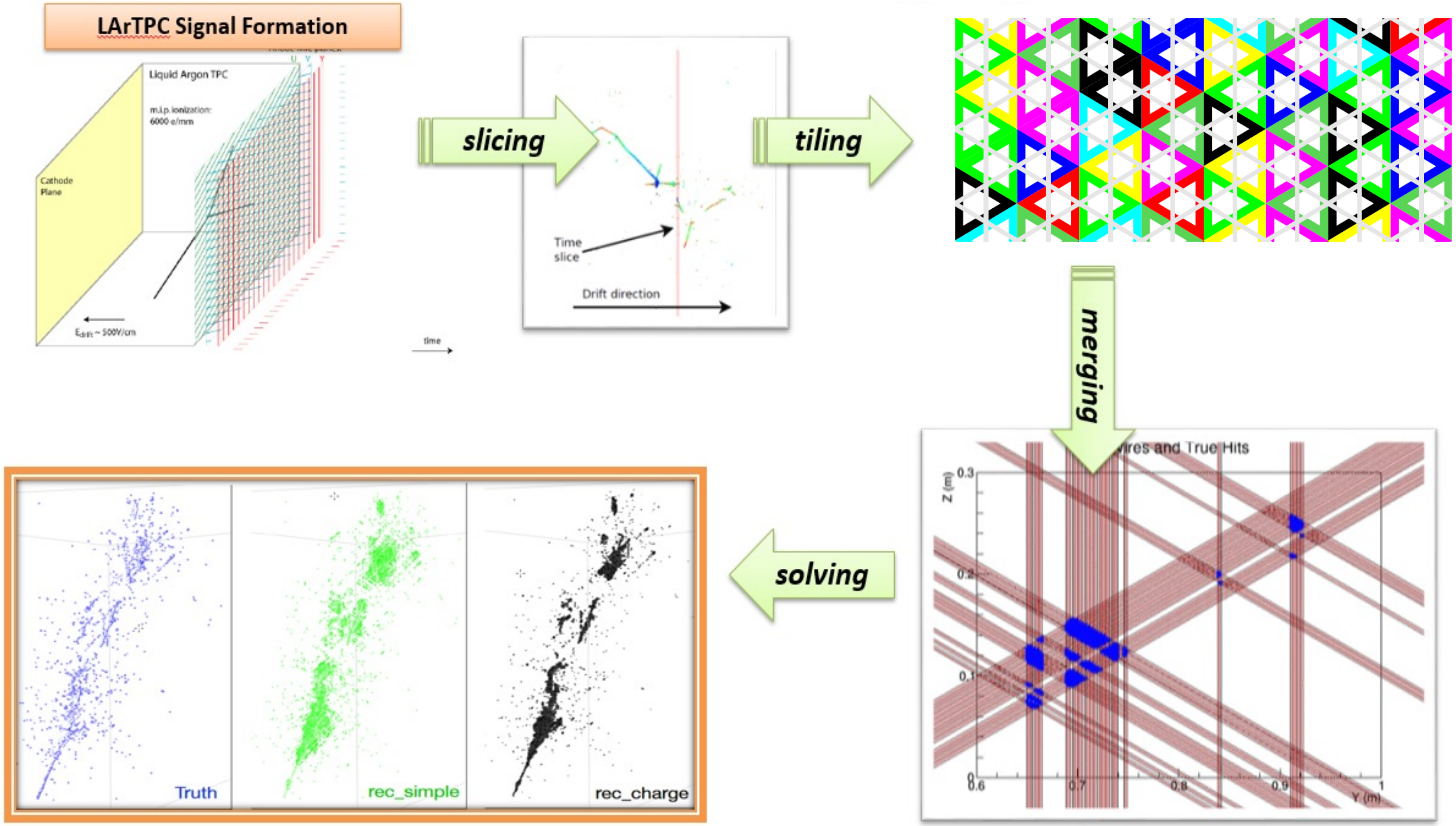


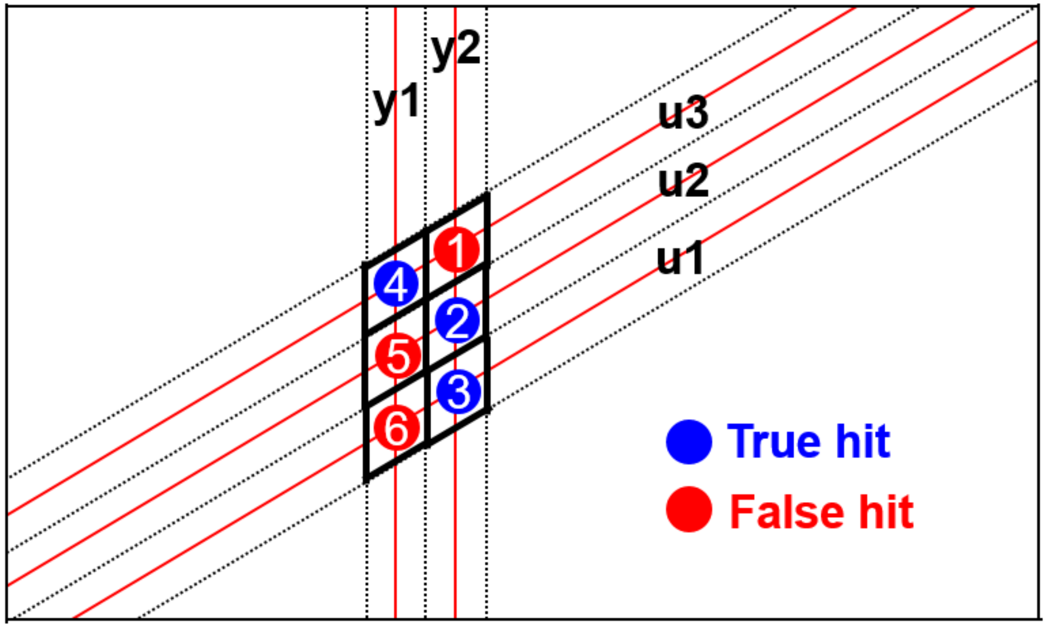
Fig.1:Basic principle of tomography: superposition free tomographic cross sections S1 and S2 compared with the projected image P

https://en.wikipedia.org/wiki/Tomography



"Three-dimensional Imaging for Large LArTPCs", JINST 13, P05032 (2018)

Solving: usage of Charge, Sparsity, Positivity, Proximity



measured charges on Wires

$$y = A \cdot X$$

true charge to be resolved

$$\begin{pmatrix} y1 \\ y2 \\ u1 \\ u2 \\ u3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & a & a & a \\ a & a & a & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & a \\ 0 & a & 0 & 0 & a & 0 \\ a & 0 & 0 & a & 0 & 0 \end{pmatrix} \begin{pmatrix} H1 \\ H2 \\ H3 \\ H4 \\ H5 \\ H6 \end{pmatrix}$$

matrix determined by geometry, a=1

- The goal is to differentiate the true hits from fake ones by using the charge information
  - ~ large charge → true hits
  - ~ zero charge → fake hits
- Sparsity, positivity, and proximity information are added through compressed sensing (L1 regularization)

L1 reg. $O(N!) \rightarrow O(m \times N)$

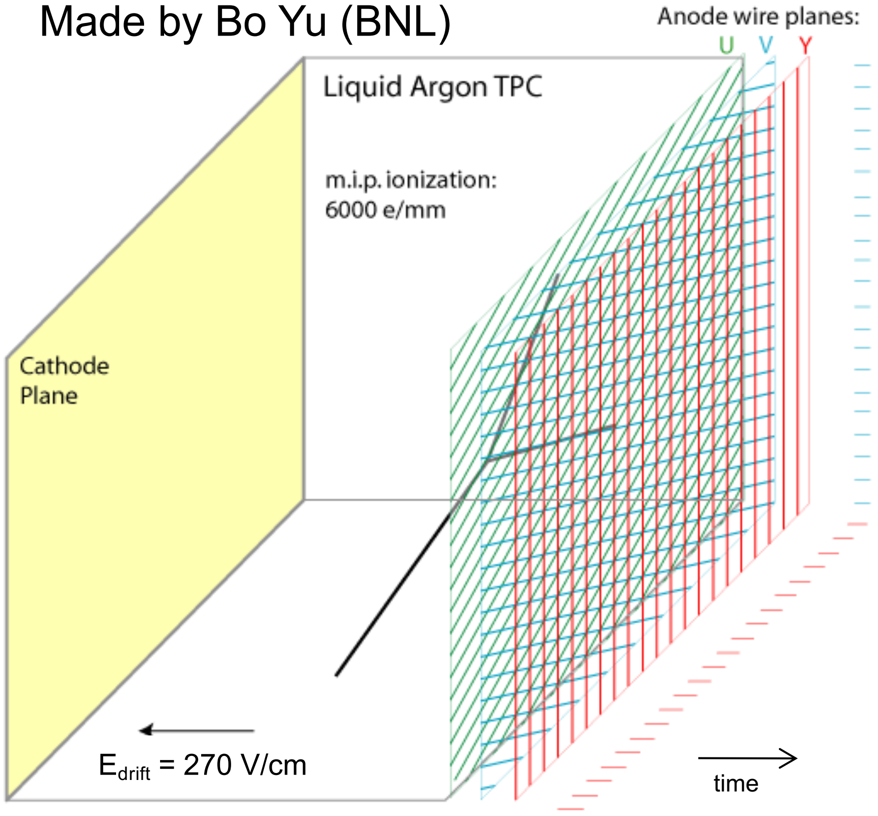
$$\chi^2 = (y - A \cdot x)^2 + \lambda \cdot \sum_i |x_i|$$

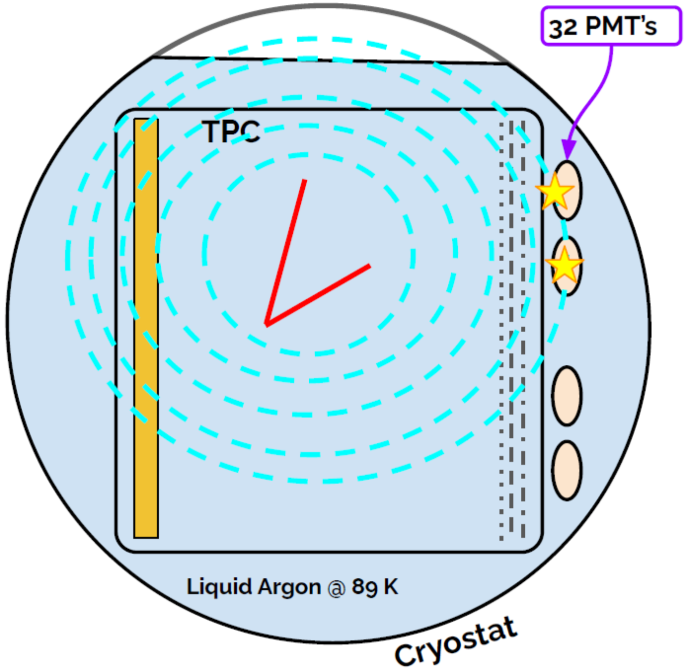
E. Candes, J. Romberg, T. Tao
arXiv-math/0503066

Brookhaven National Laboratory

Cluster-flash (light) Matching

Made by Bo Yu (BNL)





PMTs detect the scintillation light, time ~ns

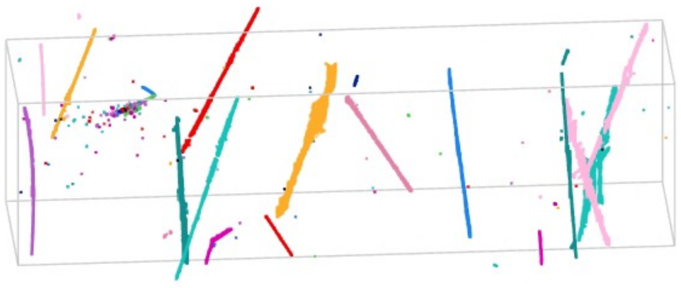
Drift velocity 1.1 mm/μs → several ms drift time

- In LArTPC, the light (PMT) readout and charge (TPC) readout systems are decoupled
- The identification of neutrino interaction candidate requires matching the charge signal with the light signal in order to obtain the event time

Brookhaven National Laboratory

Matching Principle

JINST 16 P06043 (2021)



Core Charge-Light Matching Algorithm

$\chi^2 = \sum_i \sum_j \chi^2_{ij} + \chi^2_{p1} + \chi^2_{p2} + \chi^2_{p3}$	Overall test statistics to be minimized
$\chi^2_{ij} = \frac{(M_{ij} - \sum_k a_{ik} \cdot P_{ikj} - b_i \cdot M_{ij})^2}{\delta M^2_{ij}}$	Comparison of the measured and predicted light pattern
Rule 1st $\chi^2_{p1} = \sum_i \frac{(\sum_k a_{ik} - 1)^2}{c_1^2}$	Each charge cluster can only be used once
Rule 2nd $\chi^2_{p2} = \sum_i \frac{b_i^2}{c_2^2}$	Observed light flash may not correspond to any charge cluster
$\chi^2_{p3} = \lambda \cdot \sum_i \sum_k a_{ik}$	Compressed sensing to select the best pairs

M: Measured Light Pattern
P: Predicted Light Pattern
δ : Uncertainty
i : ith Light Flash
j : jth PMT
k : kth Charge Cluster

Aggressively pursue charge-light matching
Additional cuts to examine the "light mismatch" events

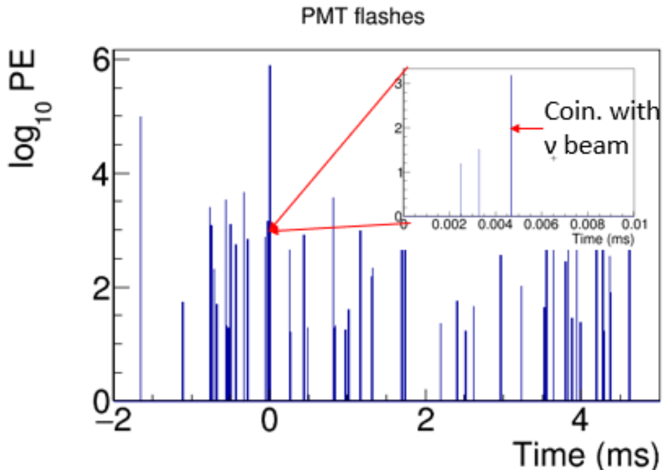
Hypotheses Selection

Mat… Hyp…

Reconstr…

- **Light signal proportional to (reconstructed 3D) charge**
- Known light acceptance given position
- Predicted vs. Measured light pattern with Compressed Sensing

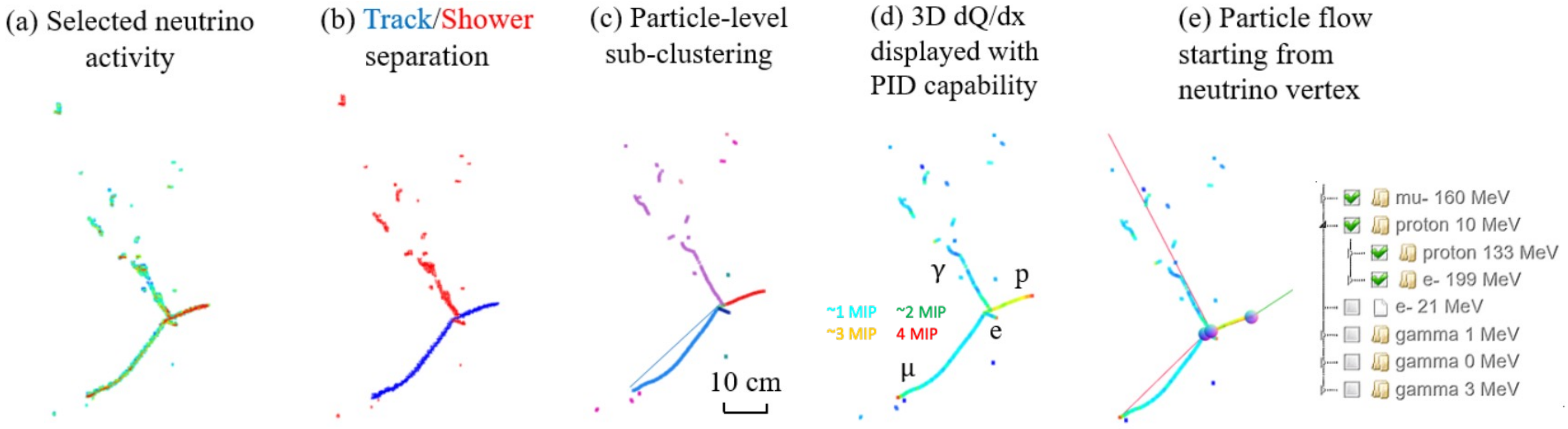
All possible hypotheses
- One cluster → at most one flash (inefficiency in the light system)
- One flash → many or zero TPC clusters within corresponding active volume (activities in inactive volume)

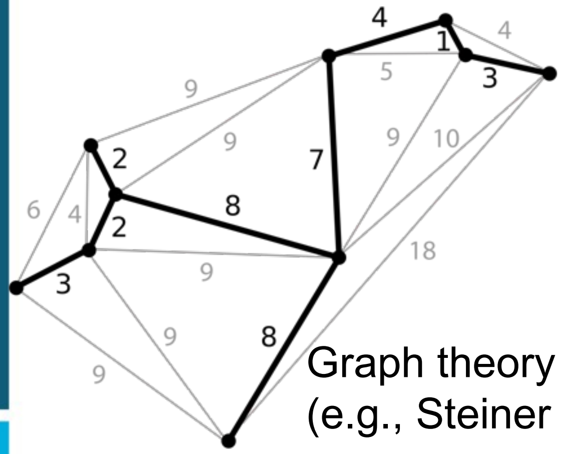


40-50 PMT activities

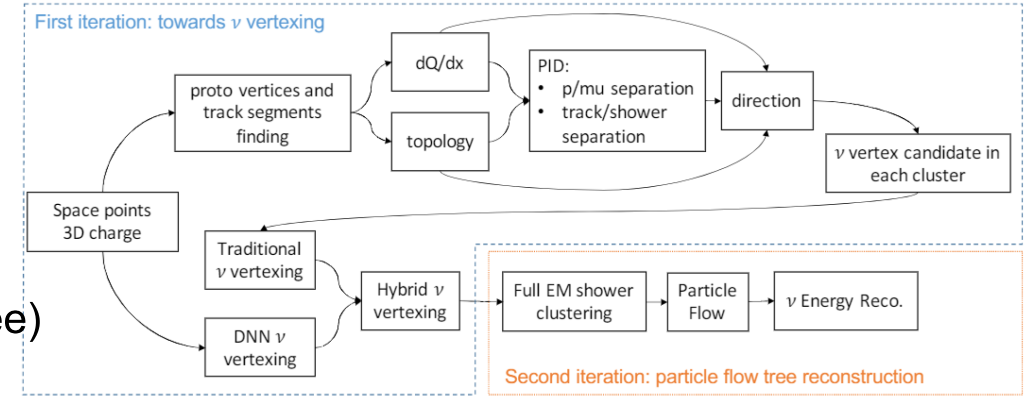
Brookhaven National Laboratory

3D Pattern Recognition

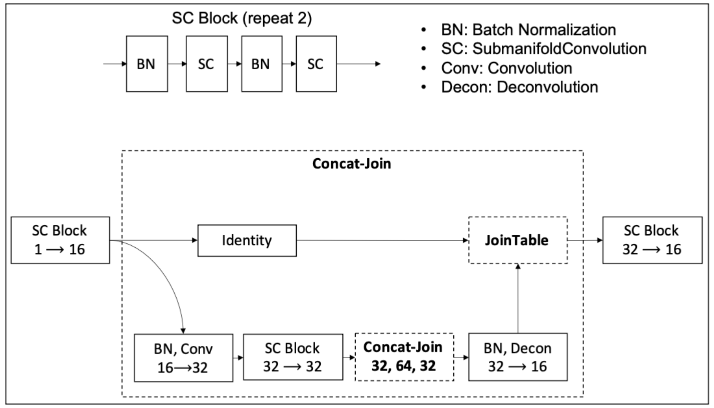




Graph theory
(e.g., Steiner tree)



Sparse Regression U-Net



Brookhaven National Laboratory

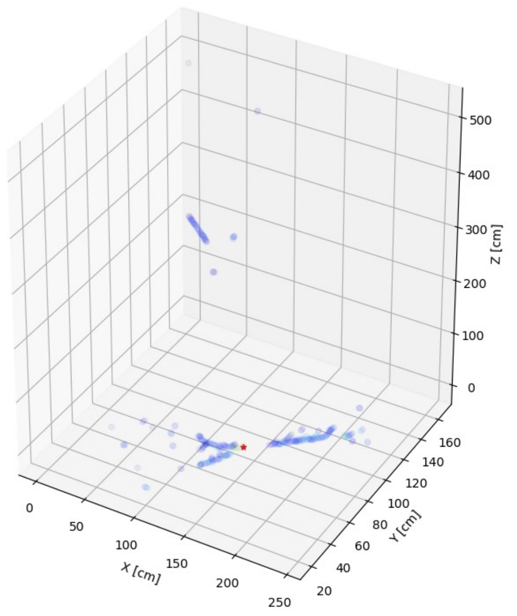
JINST 17 P01037 (2022)

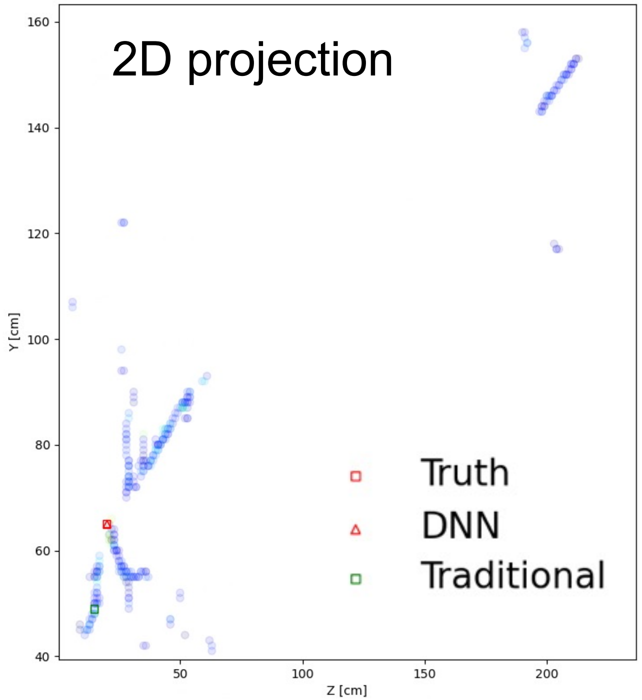
Deep Learning based Neutrino Interaction Vertex Finding

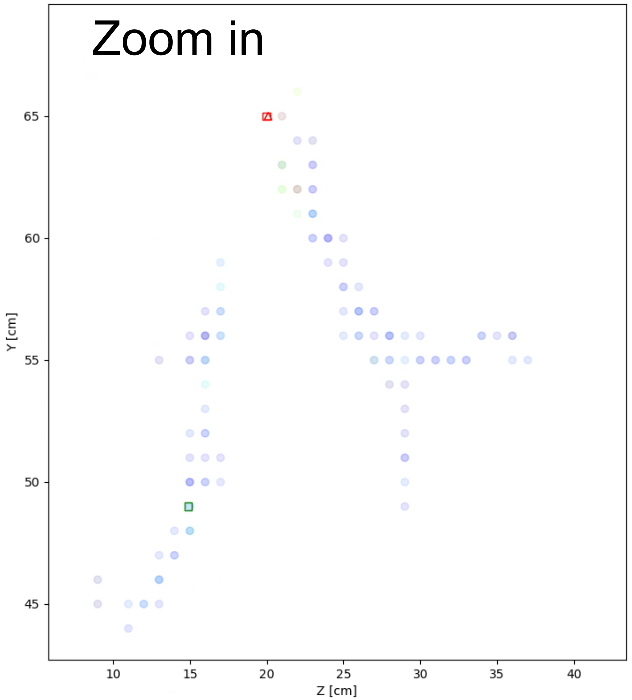
Regressional segmentation with a sparse U-Net

- U-Net: efficiently use geometry info which is critical
  - compared to graph networks
- Regressional loss on distance based "confidence map" to use a region of points instead of only one
  - otherwise, data is highly imbalanced (Z. Cao etc, arXiv:1812.08008)
- Sparse: boosted computing efficiency with our sparse 3D data
  - Submanifold Sparse Convolutional Networks (B. Graham etc, arXiv:1706.01307)

3D points from Wire-Cell







Regressional segmentation

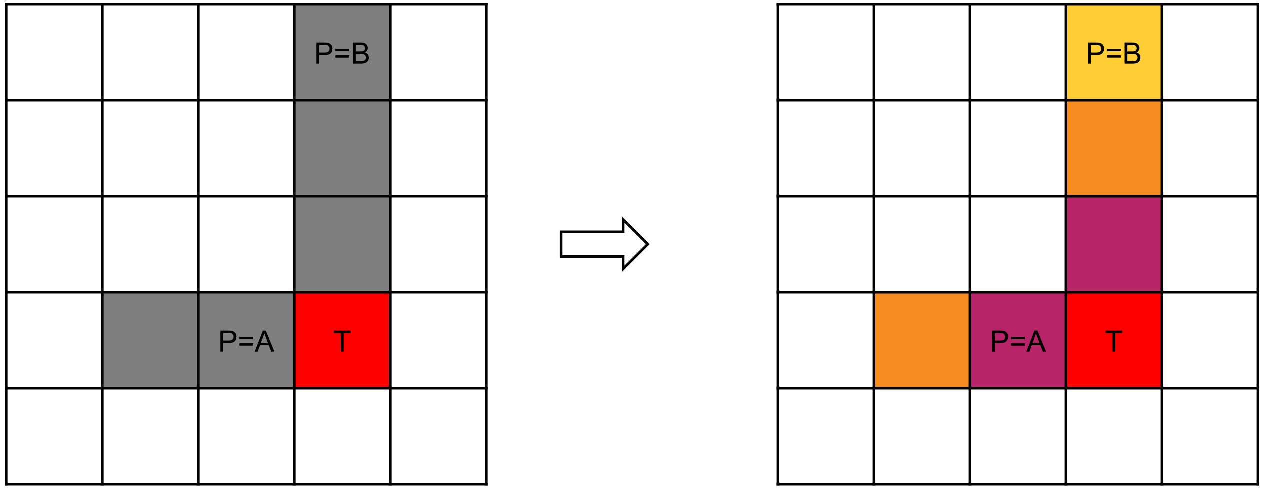
Initially we used Cross Entropy loss

- effectively only use the vertex information for one space point
- doesn’t care about the distance between the prediction and the target.
  - while our main metric is this distance.

⟶ encode the distance information for a region of points

- predicting the full “confidence map” instead of only one point
- current mapping: $\mathrm{Conf}_{\mathrm{truth}} = \exp\left(-\frac{\|\vec{x}-\vec{v}_{\mathrm{truth}}\|^2}{2\sigma^2}\right)$

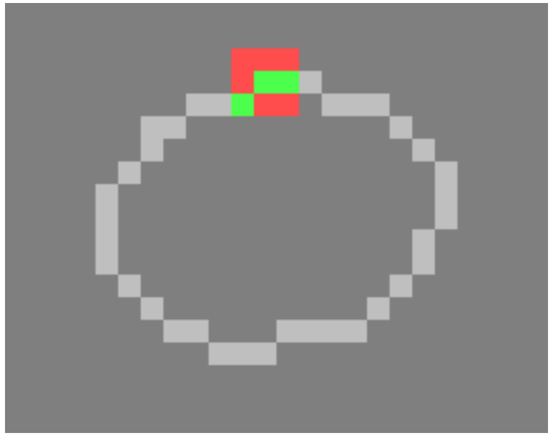
OpenPose:
https://arxiv.org/pdf/1812.08008.pdf



Brookhaven National Laboratory

Network structure and improvements

SparseConvNet

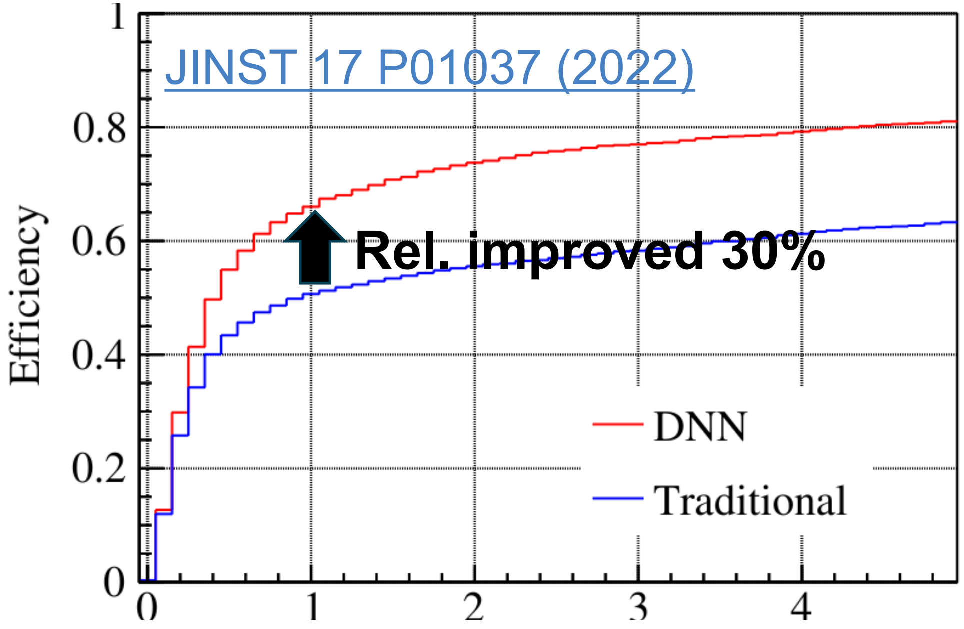


Used *SparseConvNet* to realized 3D sparse conv. DNN
https://github.com/facebookresearch/SparseConvNet

This work: github, paper

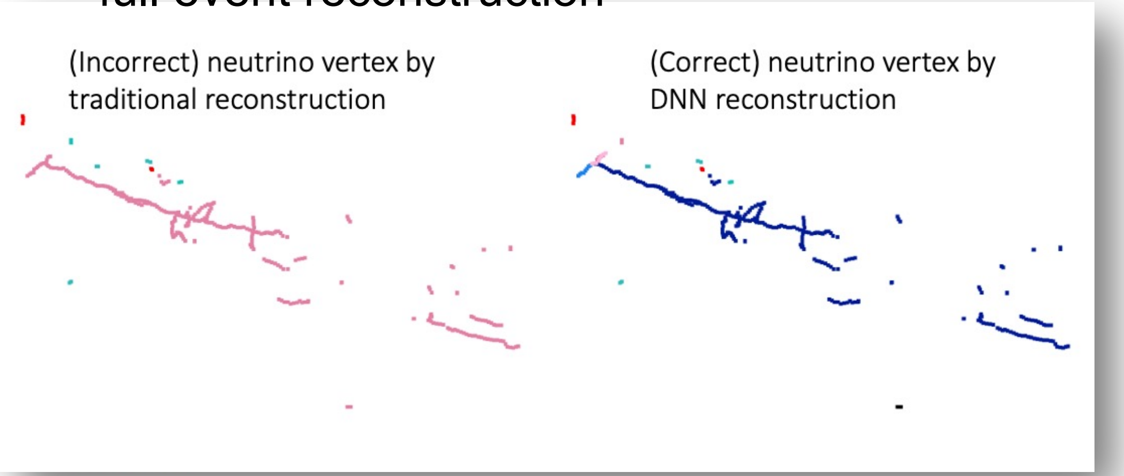
3D charge reco'ed from Wire-Cell → Sparse U-Net → Fully connected + sigmoid → confidence value for each voxel

ν_eCC vertex identification efficiency



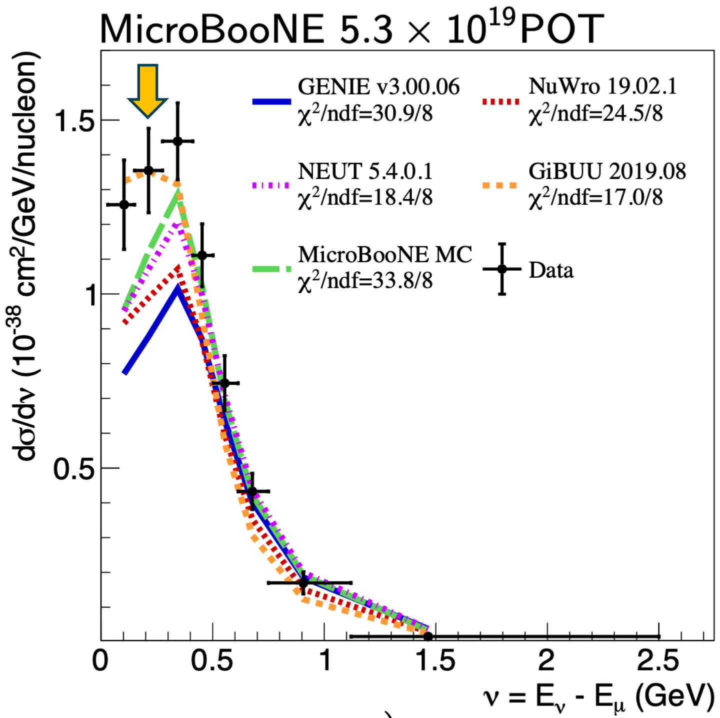
Distance between reco. and true vertex (cm)

Illustration of impact of vertex ID on the full event reconstruction



Application of Wire-Cell in Physics Analyses

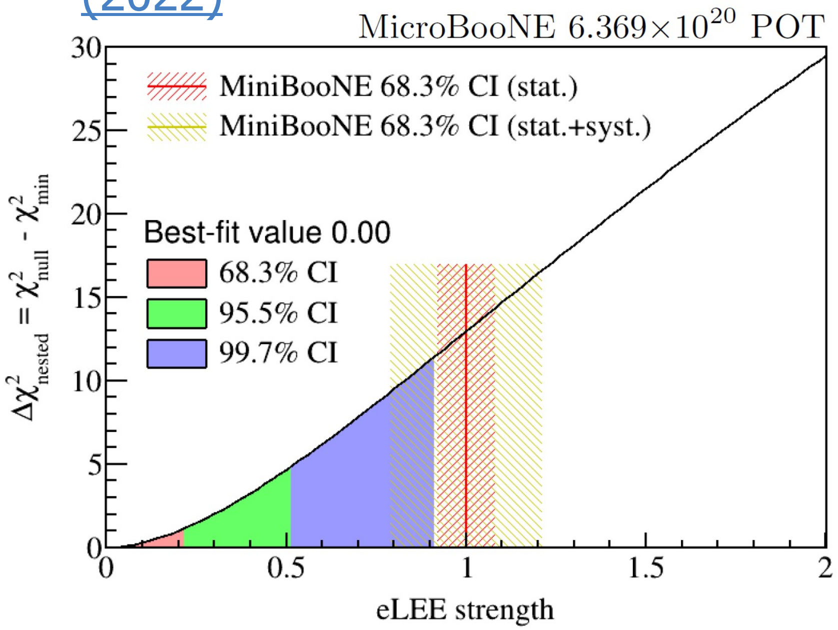
Energy-dependent Cross Section
Phys. Rev. Lett. **128**, 151801 (2022)



c)

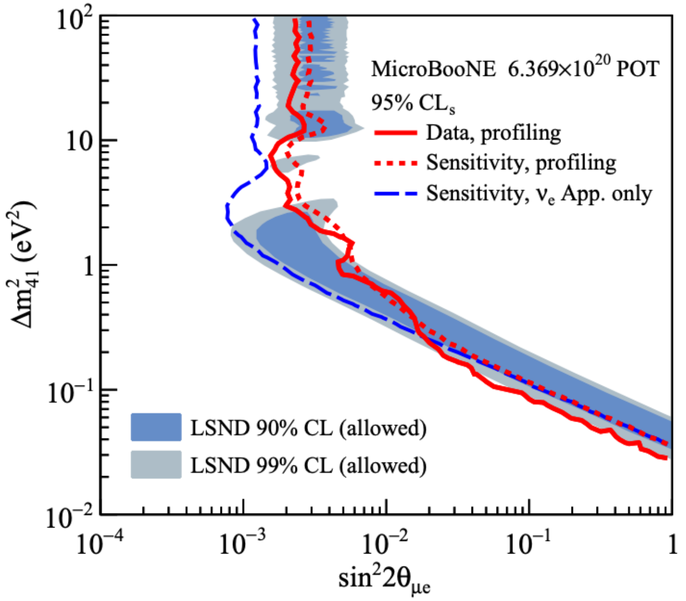
- Good separation power of model predictions from different generators
- GiBUU's central prediction gives best agreement at low energy transfer for Ar $\Longrightarrow$ more contribution of 2p2h

Search for ν_e Low Energy Excess
Phys. Rev. D **105**, 112005 (2022)
Phys. Rev. Lett. **128**, 241801 (2022)



- 68% stat-only (full) uncer. MiniBooNE CI is disfavored at over 3σ (2.6σ)
- ve cannot be the sole explanation of MiniBooNE LEE!

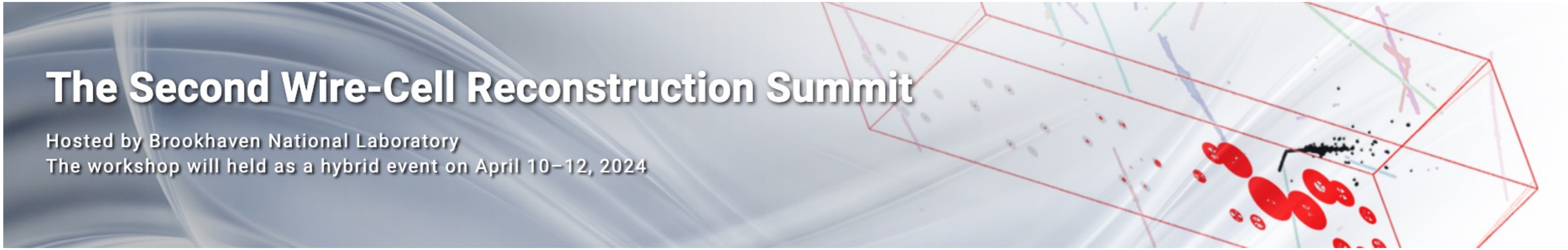
Constraints on light Sterile neutrino oscillations, Phys. Rev. Lett. 130, 011801 (2023), submitted to PRL



- No hint of light sterile neutrino oscillations has been found yet, and so we set exclusion contours (plotted at the 95% confidence level)

Outlook

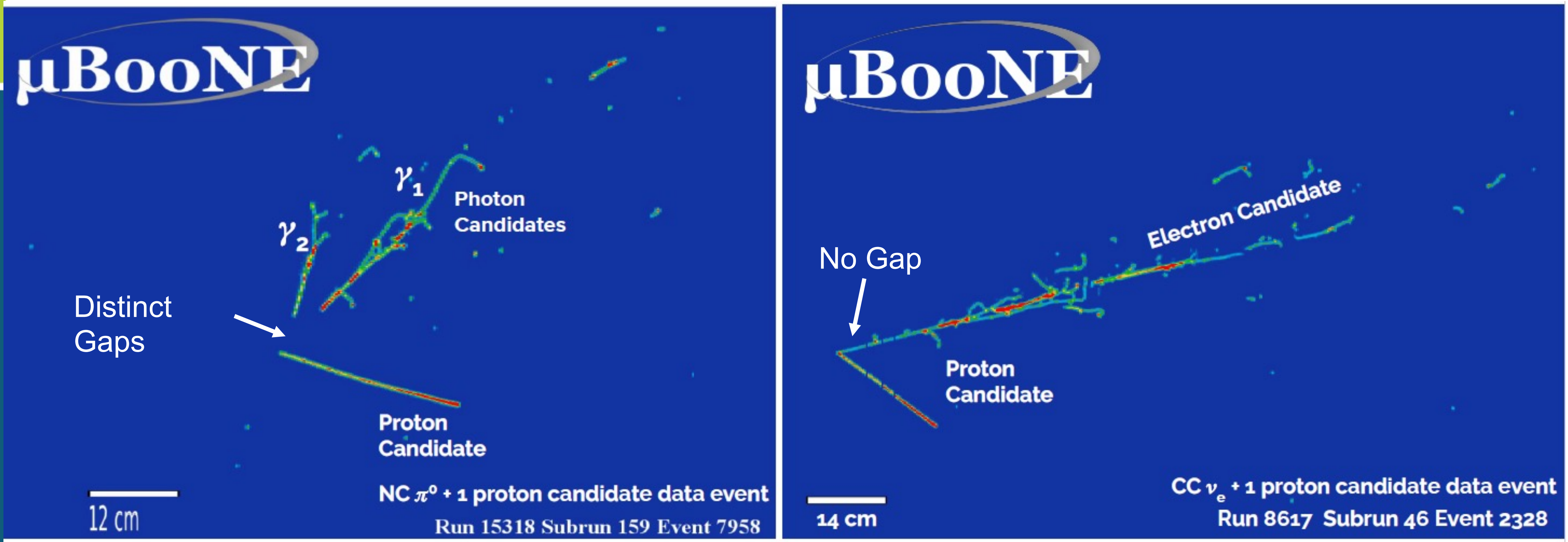
- The Wire-Cell team has developed a fully automated reconstruction chain for LArTPC reconstruction for neutrino experiments
- Its good performance was demonstrated in MicroBooNE analyses
- https://www.bnl.gov/wirecellsummit/



Brookhaven National Laboratory

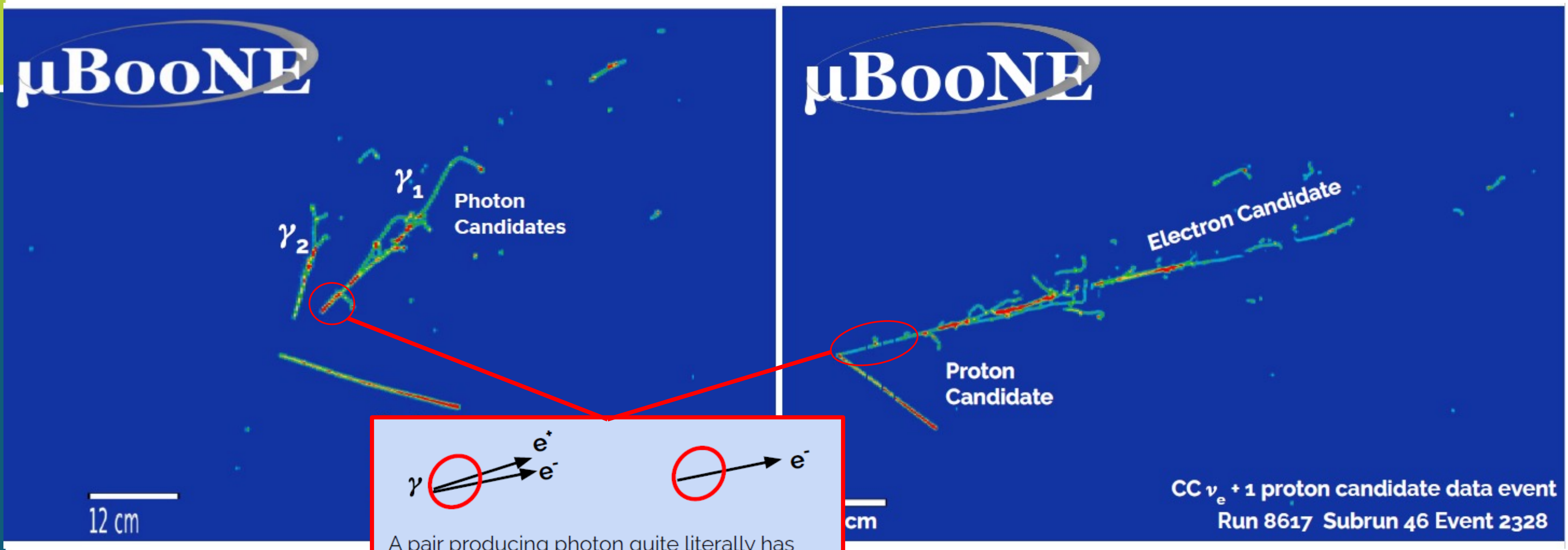
Thanks!

Separation of e and γ in LArTPC



- Event topology to separate EM showers (e/γ) from tracks (proton, muon)
- Separation of e and γ : Gap Identification

Separation of e and γ in LArTPC



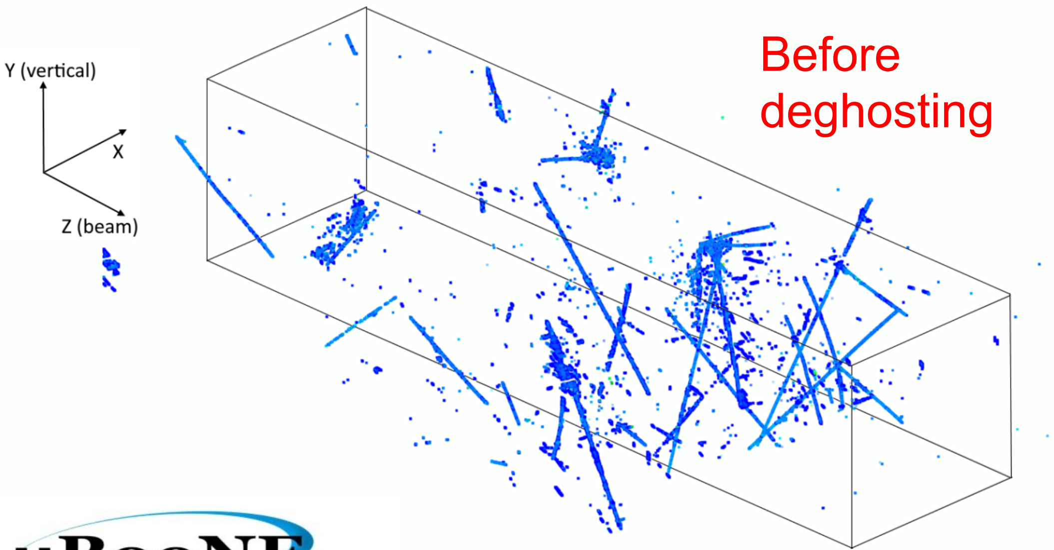
A pair producing photon quite literally has twice the number of minimally ionizing electrons, so has **twice the charge deposited at the start of the shower (dE/dx)**

- Event topology t… e/γ) from tracks (proton, muon)
- Separation of e and γ : Gap Identification + dE/dx
- Unique capability to identify ν_e charge-current (CC) interactions in LArTPC

Overcome Challenges of 10% non-functional channels

- Impact of 10% non-functional channels is reduced from **~30% → ~3% dead volume** by requiring only 2 out 3 wire planes in reconstruction when necessary
  - Utilizing coverage of 3 planes, but generating a lot of fake 3D activities (ghosts)
  - Dedicated algorithm in deghosting, clustering, charge solving etc. have to be developed

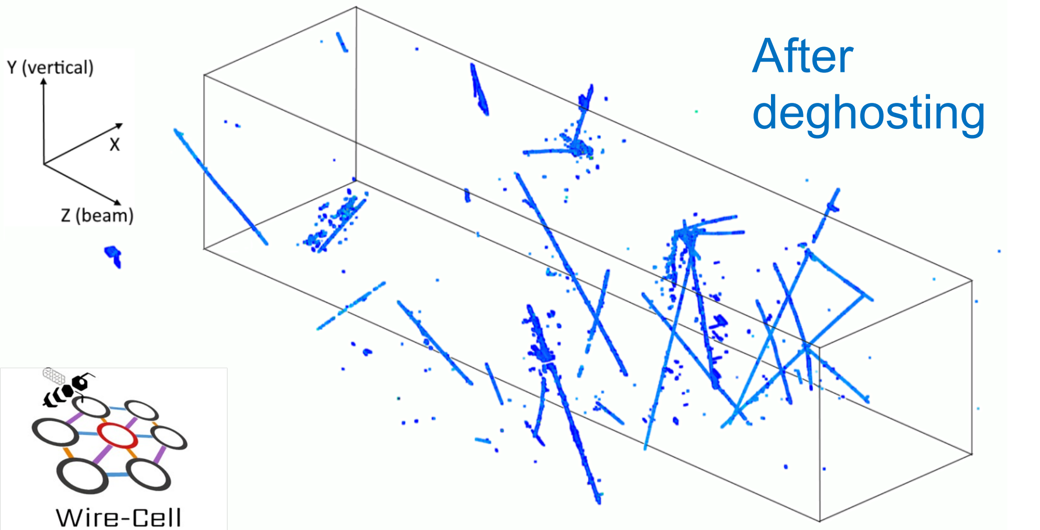
Before deghosting



μBooNE

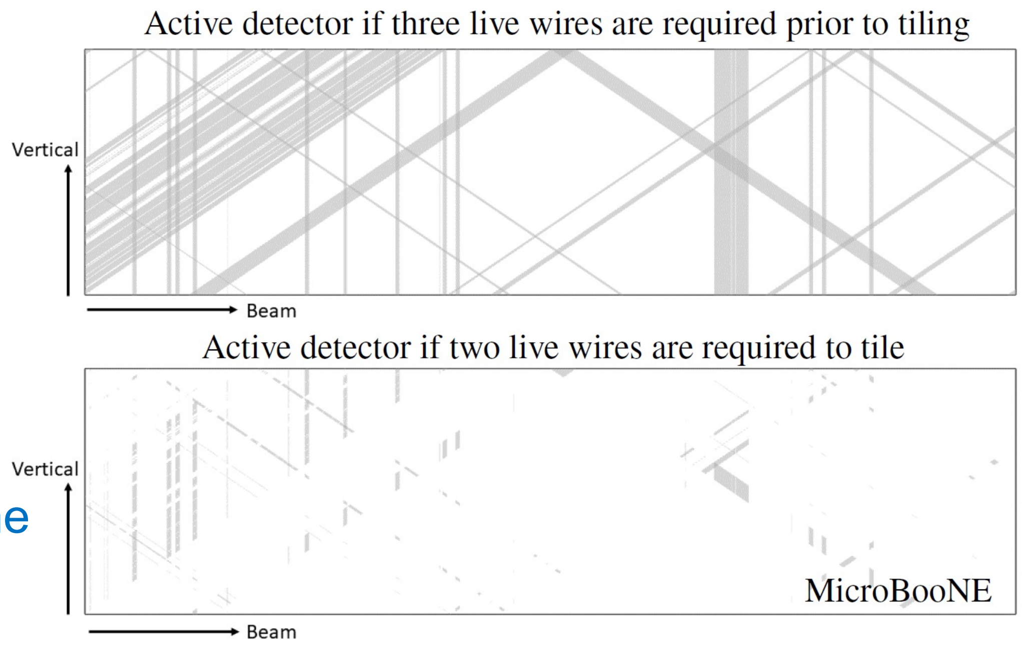
[JINST 16, P06043](#)

After deghosting



30% dead volume

3% dead volume

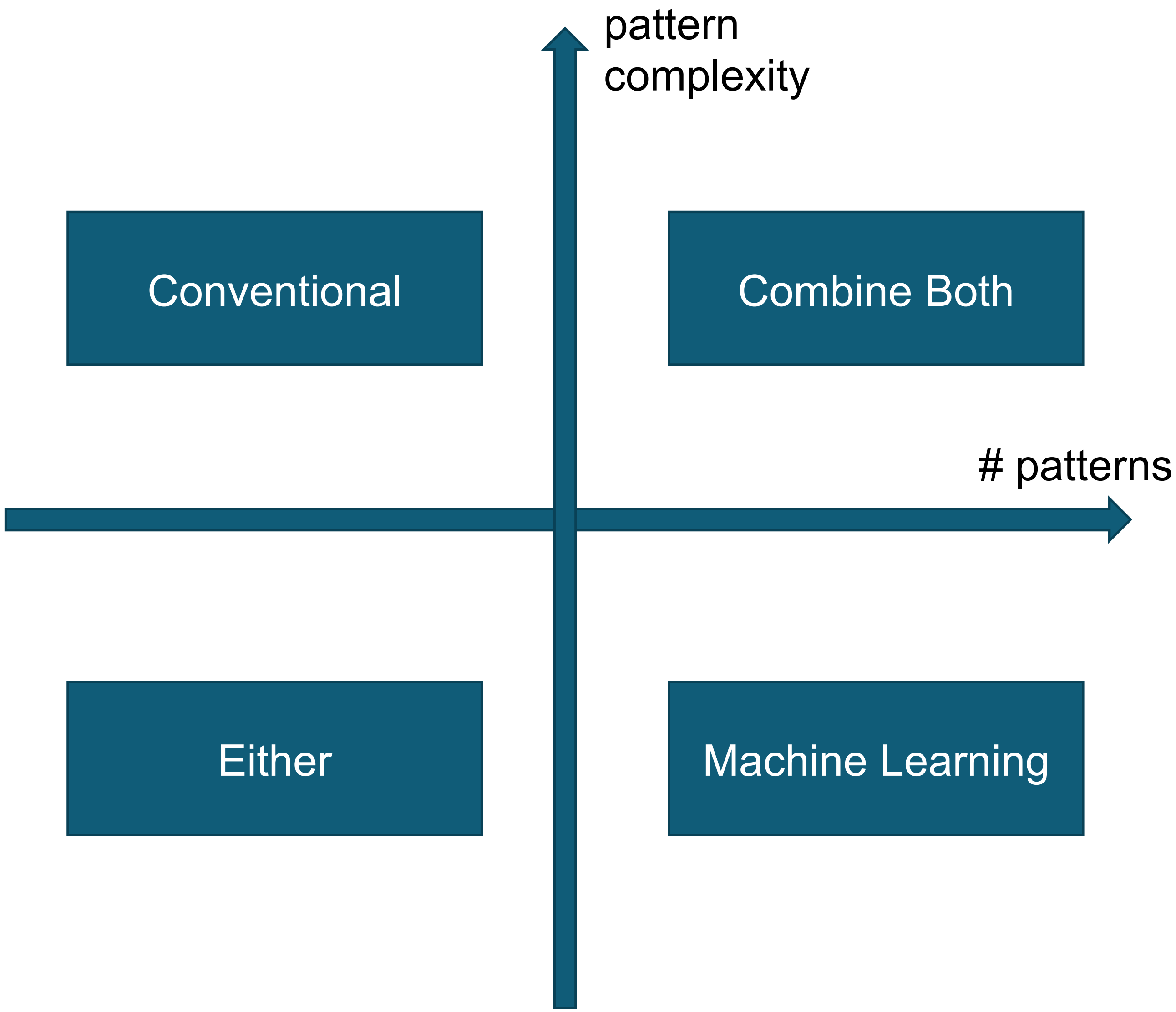


Wire-Cell

Human to machine

- We learned that some tasks in the chain fit better for conventional alg. while some others fit better for ML alg.

- I believe combining both would give us the best performance with limited data and computing resources



Brookhaven National Laboratory