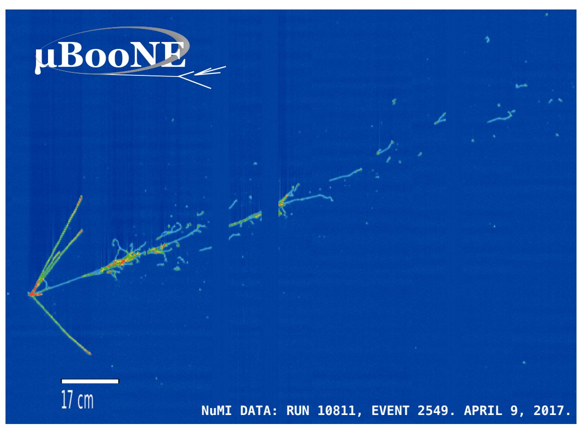
Image Based Reconstruction and Deep Learning Tools in MicroBooNE

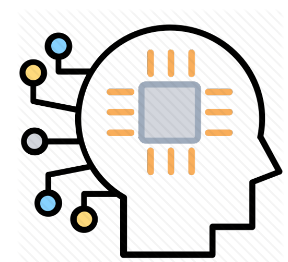
Rui An @ *Illinois Institute of Technology*
09-18-2019, DESY
On behalf of MicroBooNE Collaboration
Reconstruction and Machine Learning in Neutrino Experiments

Outline

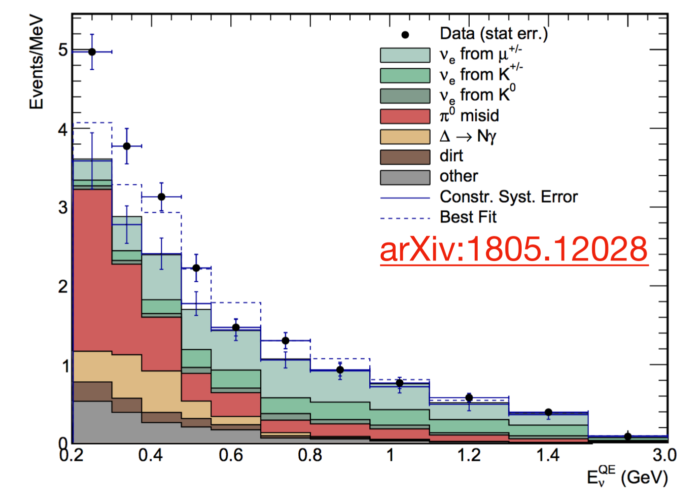
- MicroBooNE developed an *imaged based event reconstruction chain*,
  - ★ taking advantage of high resolution image data and computer vision software OpenCV,
  - ★ targeting the exclusive 1$lepton$-1Proton topologies in the low energy excess region,



+ OpenCV +

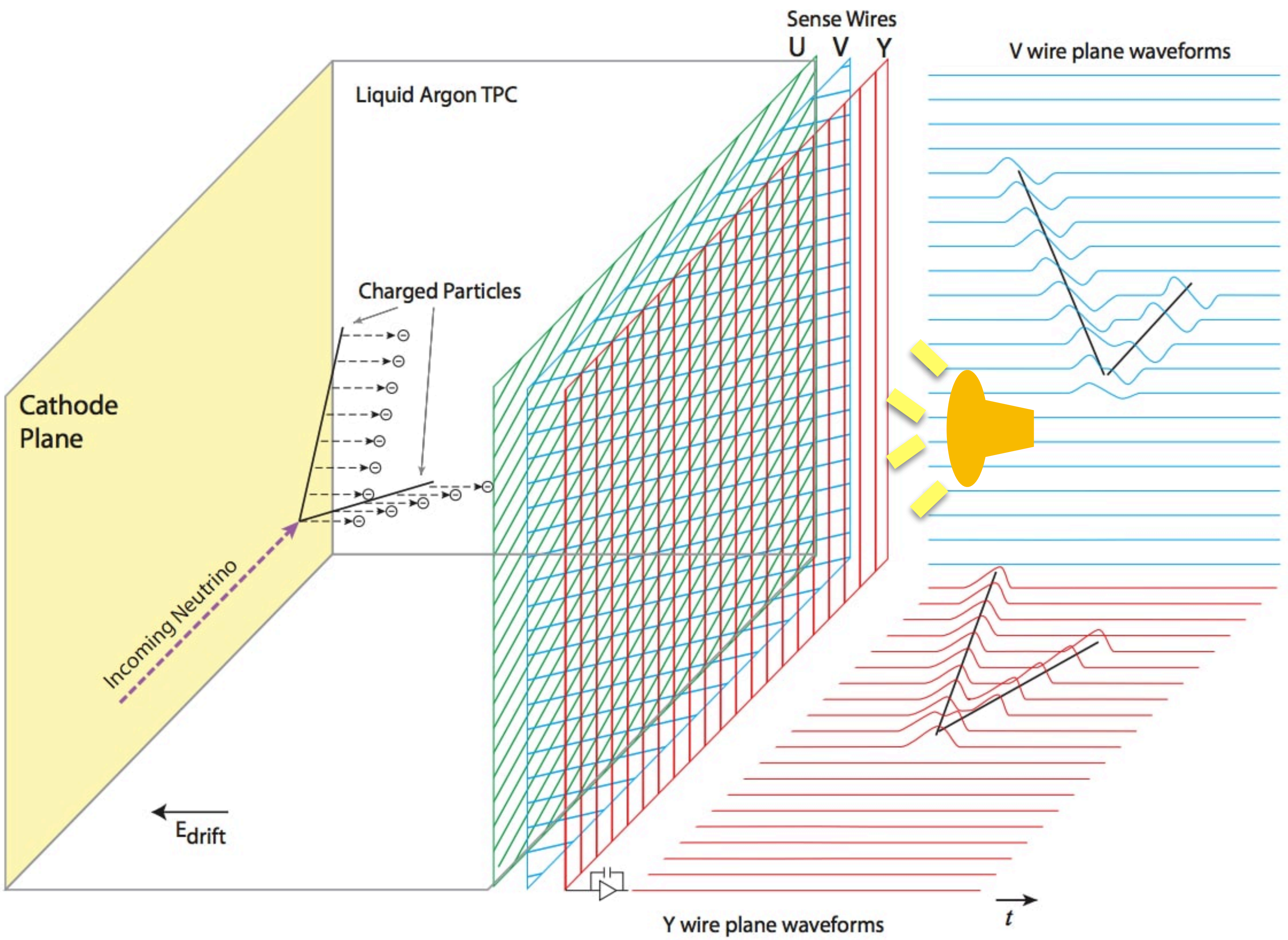


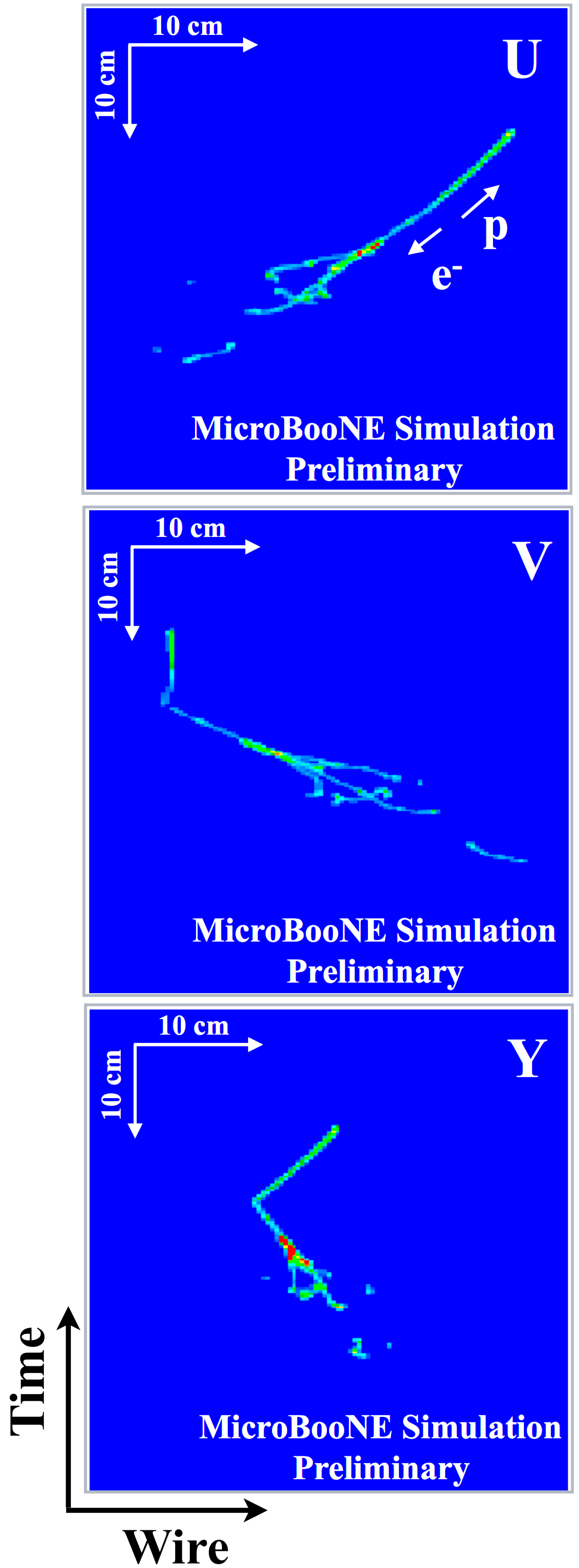
→



- Various deep learning networks applied by MicroBooNE,
  - ★ Semantic segmentation network→pixel clustering→**SSNet**,
  - ★ Classification network→particle identification→**MPID**,
  - ★ Instance segmentation network→cosmic removal→**MaskRCNN.**

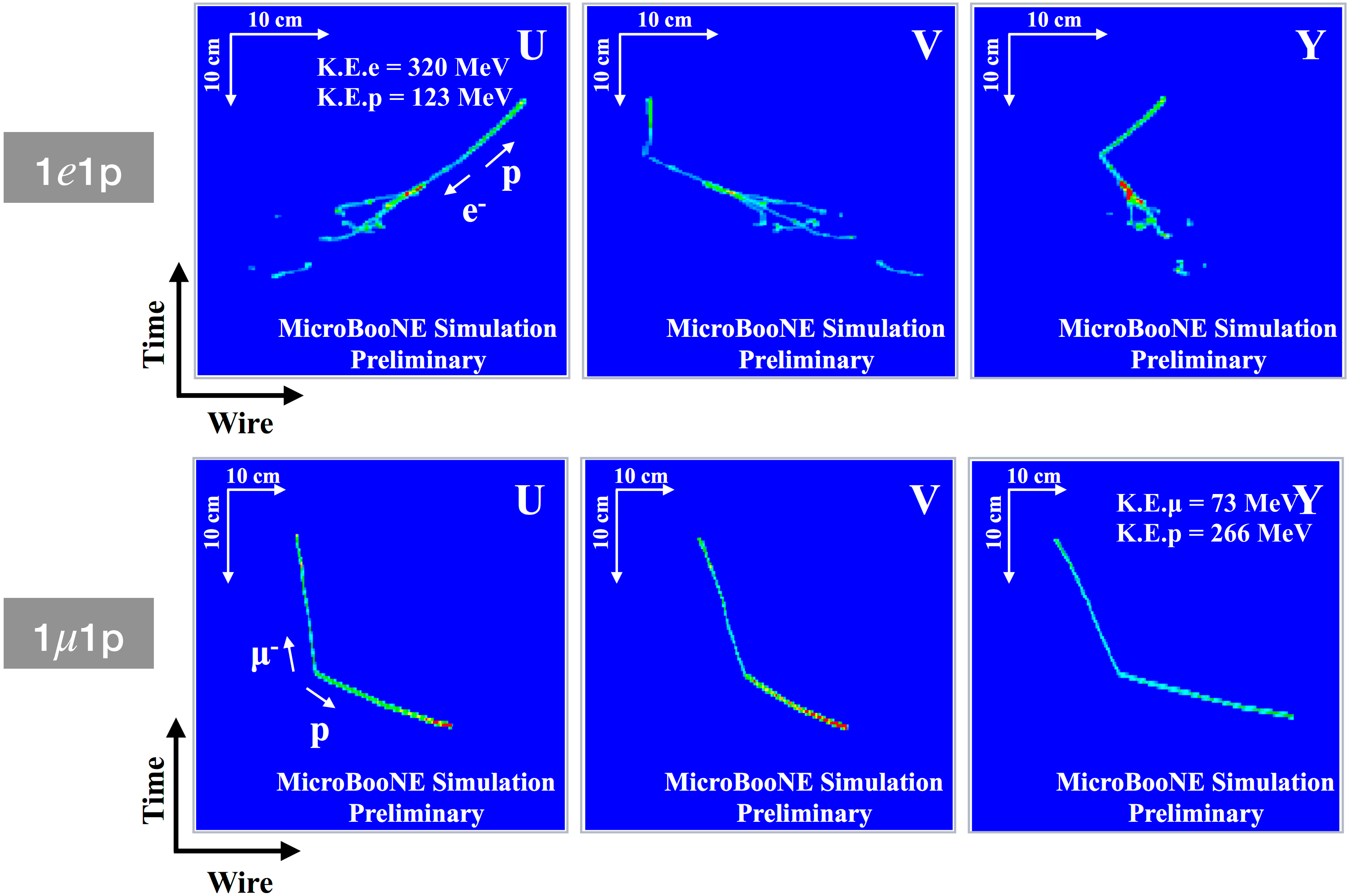
MicroBooNE LArTPC and Image





2,400 wire on U,V and 3,456 wires on Y plane.
4.8 ms readout.
~[wires]x[ticks /6] pixels after applying compression factor on wires and time.

1l1Proton Event Topology



- Benefits: easier to reject cosmic and non-CCQE in LEE region using topology parameters.

- Visible requirement: proton > 60 MeV and lepton > 35 MeV.

Reconstruction Chain

Cosmic Tagging

↓

Shower / Track Pixel Labeling

↓

2D & 3D Reconstruction

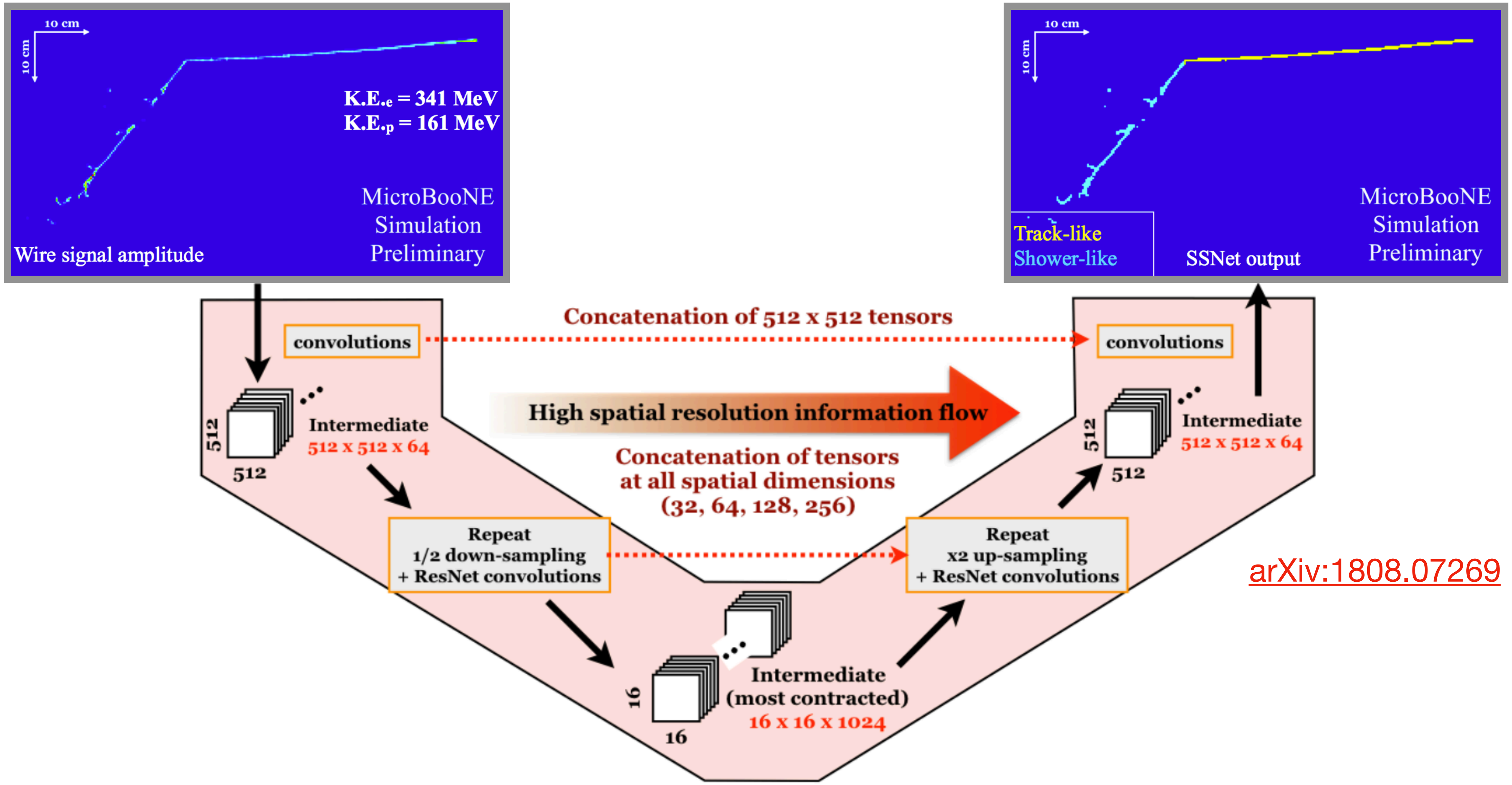
↙ ↘

PID using CNN → $1e1p$

$1\mu1p$

Semantic Segmentation Network

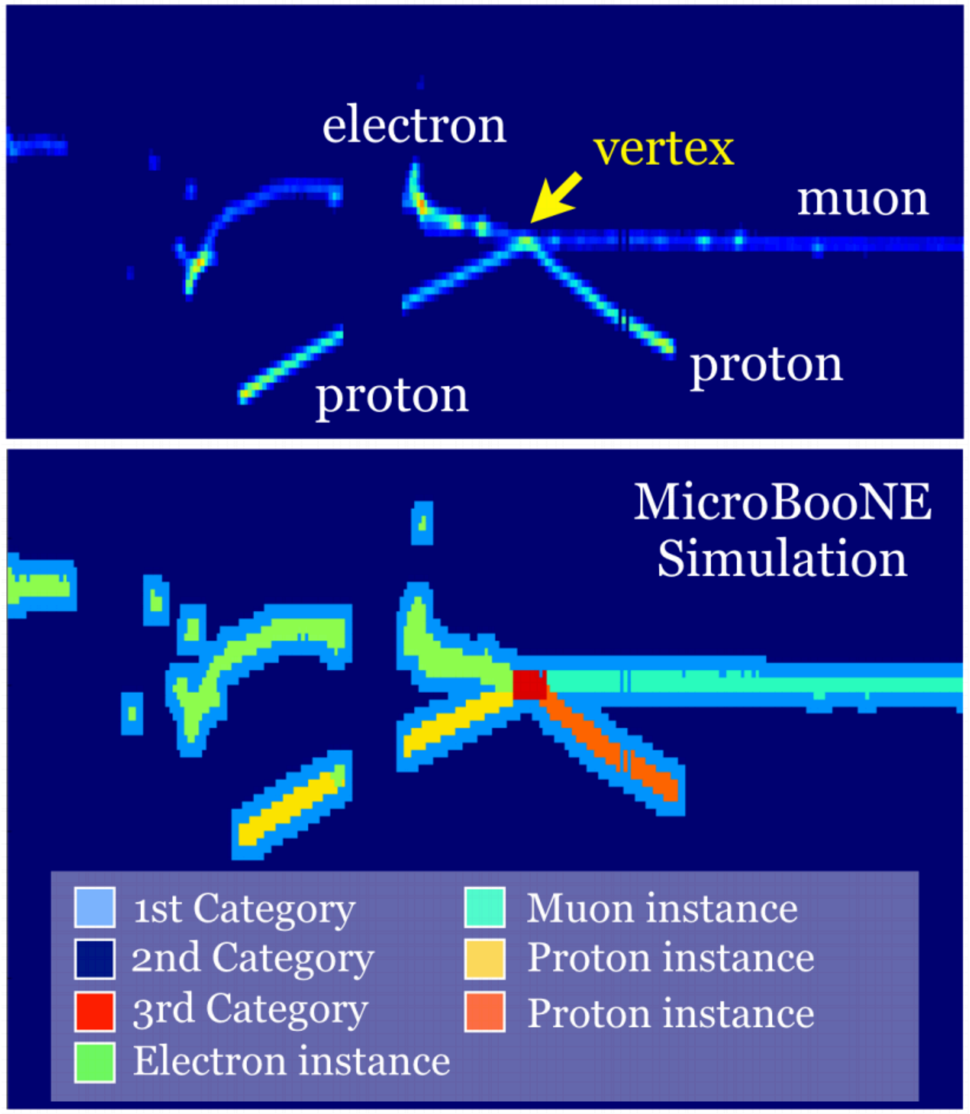
μBooNE



arXiv:1808.07269

- The goal is to make pixel-level label decision between track and shower for LArTPC,
- SSNet applies a combination of U-Net and ResNet. Feature maps concatenated between down and up sampling steps,
- Applied the SparseNet, training speed boosted by 5,
- Training sample size: 100,000 events. Test sample size: 20,000 events.

Semantic Segmentation Network

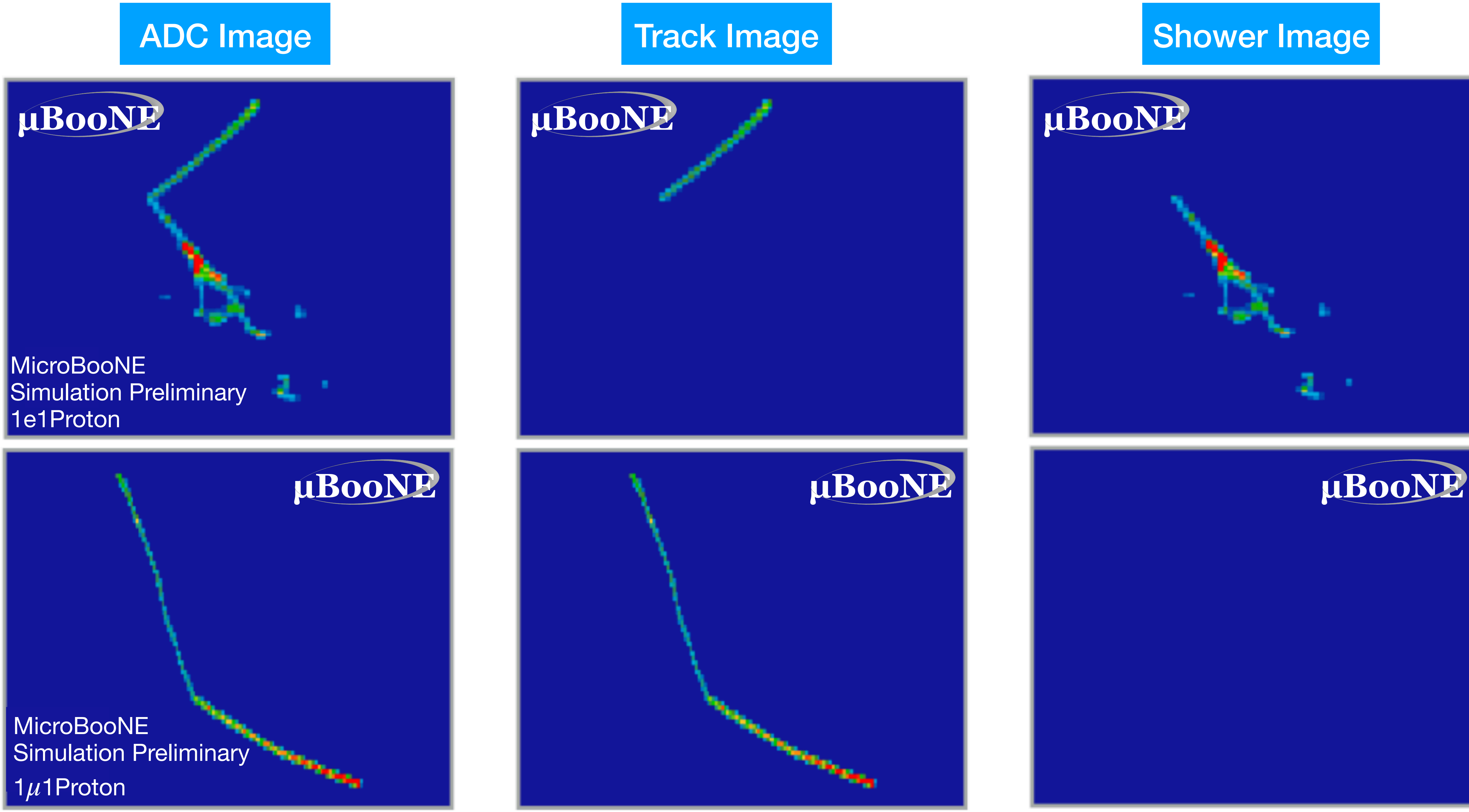


Sample	ICPF mean	ICPF 90%	Shower	Track
Test	1.9	4.6	4.1	2.6
ν_e	6.0	13.8	7.6	13.8
ν_μ	3.9	4.5	14.2	4.3
1e1p	2.2	5.7	2.8	4.0
1μ1p-LE	2.3	2.2	6.2	2.4
1e1p-LE	3.9	11.5	3.8	8.0

* ICPF: Incorrectly Classified Pixel Fraction

- SSNet applied a weighted loss for the LArTPC image sparsity.
- Achieved very low error rates across different samples.

SSNet Output



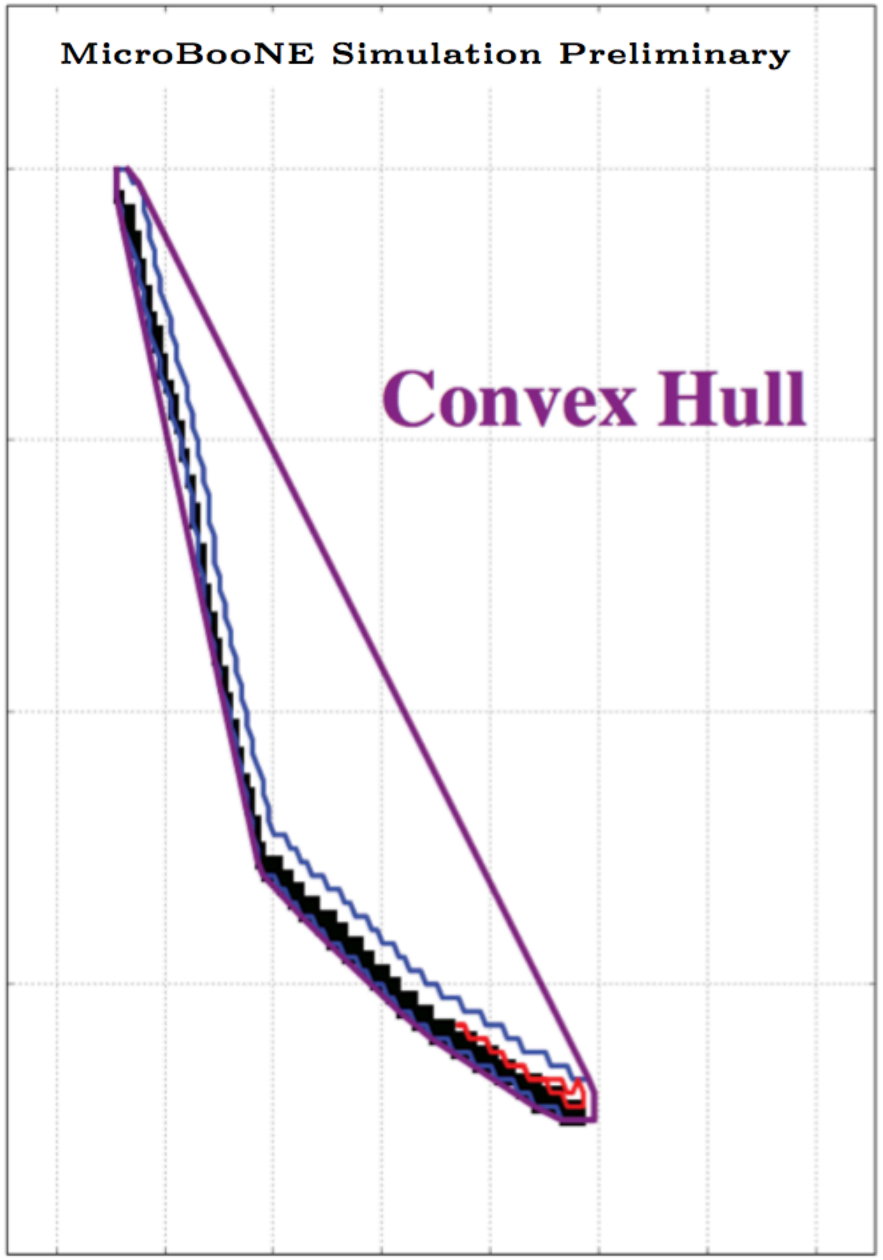
- SSNet is used to provide vertex seed.

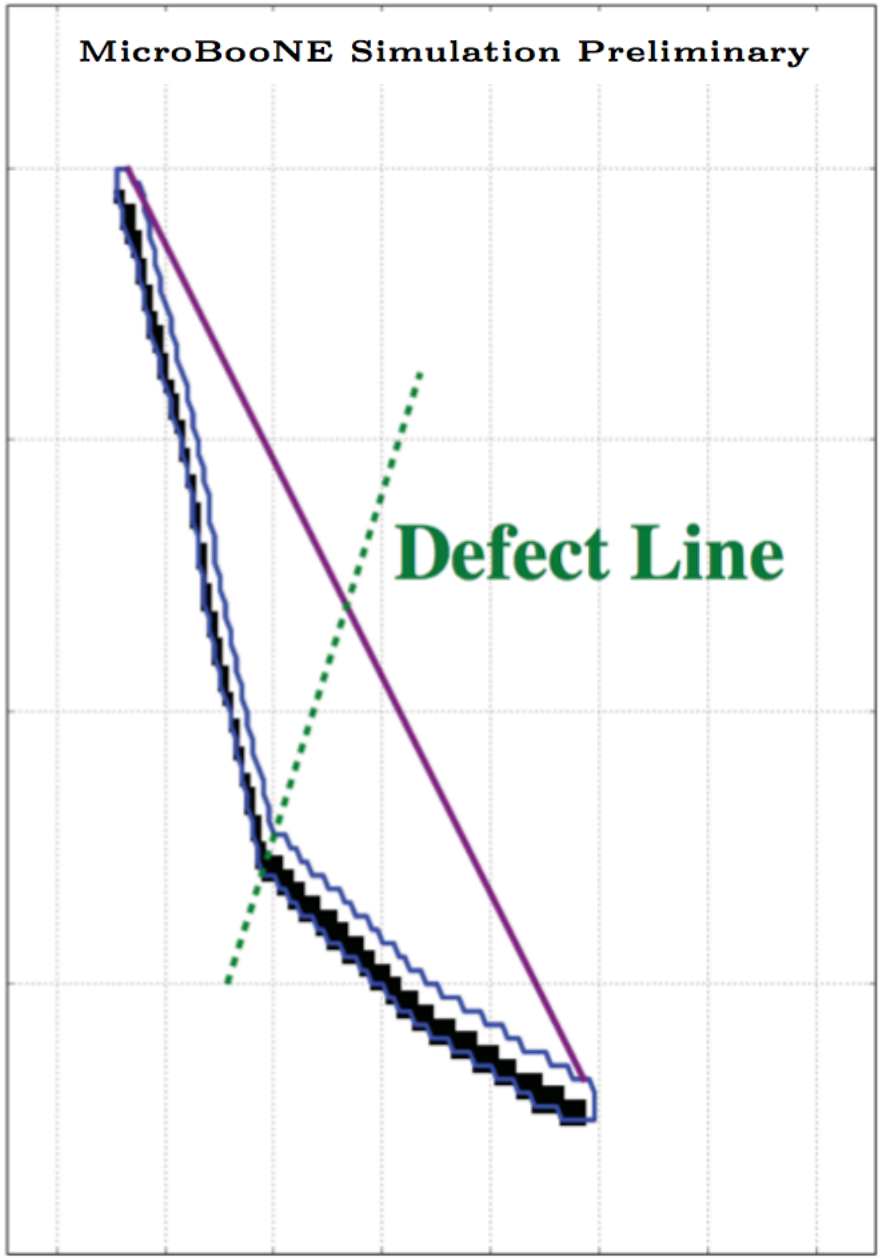
$1l1P$ Reconstruction

- We developed two reconstruction chains respectively for,
  - ★ 1μ1Proton (this talk) → using track pixel only images.
  - ★ $1e$1Proton (backup slides) → using track and shower pixel images.

- Two libraries developed:
  - ★ Liquid Argon Open Computer Vision(LArOpenCV),
    - ✳ library for image manipulation, reconstruction and analysis,
    - ✳ interface between LArTPC data and OpenCV,
  - ★ Liquid Argon Computer Vision (LArCV),
    - ✳ Bridge LArTPC data and TensorFlow, PyTorch and caffe etc,
    - ✳ Image data format and processing tool.

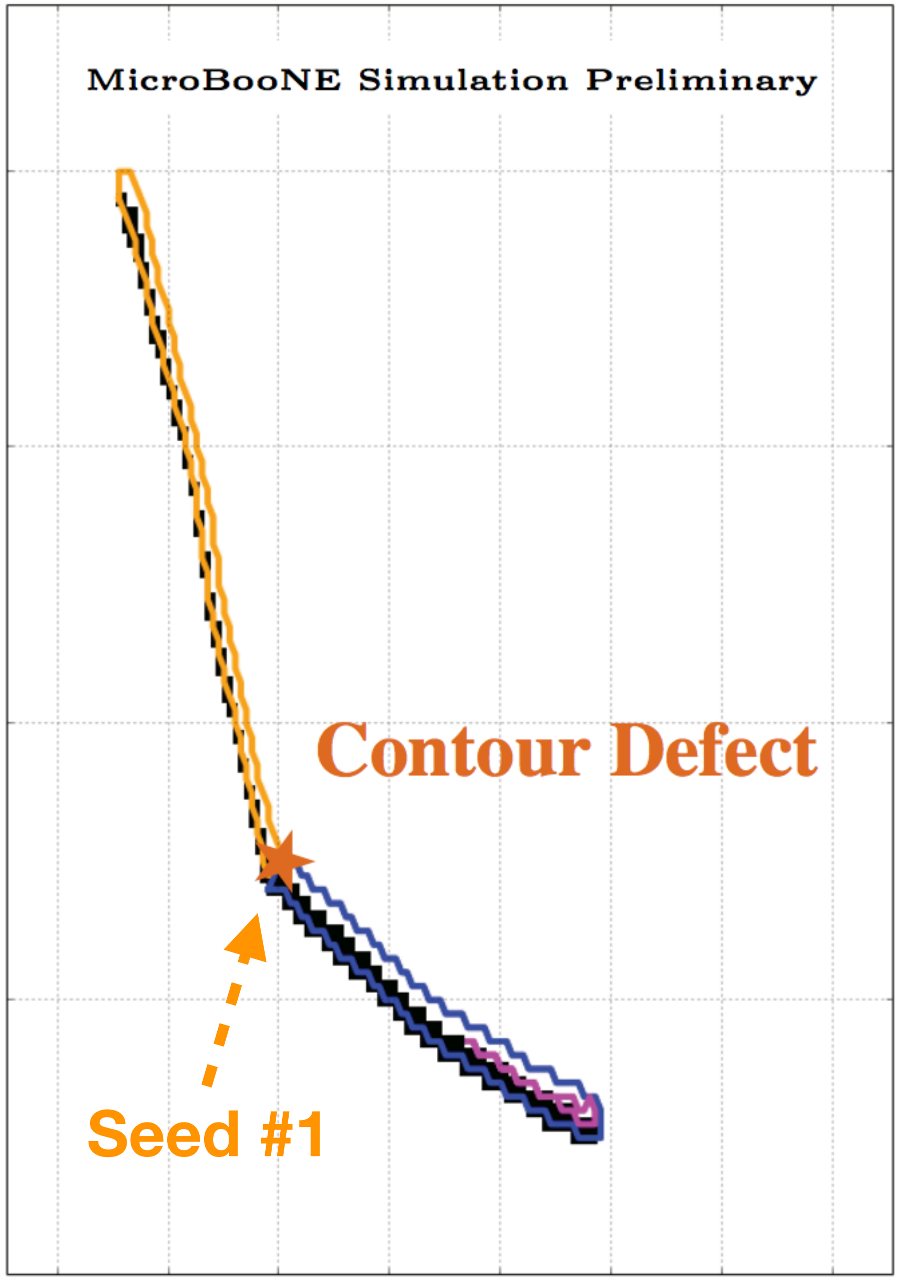
$1\mu1$P Vertex Candidates Examples

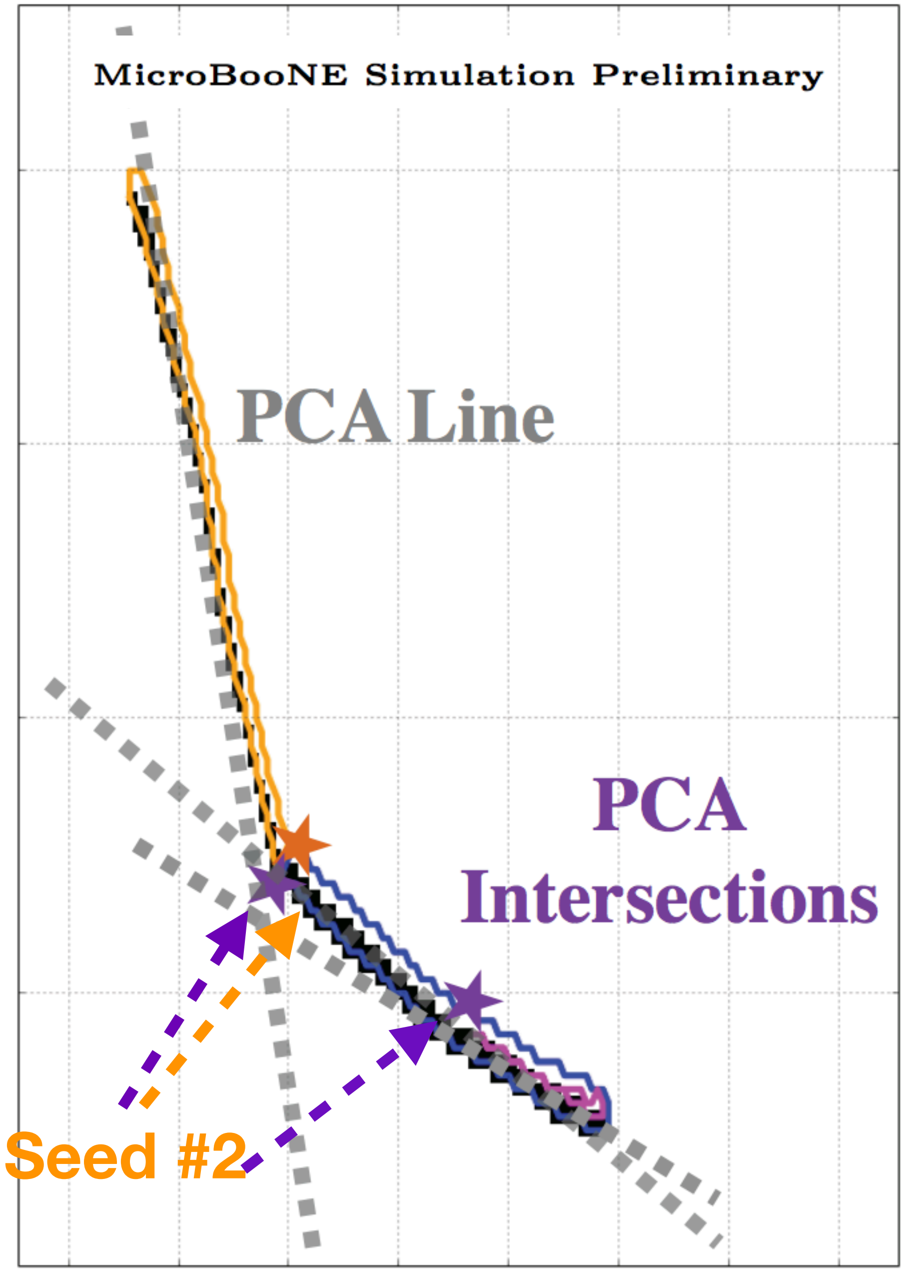




- $1\mu1$Proton vertex finding uses track-only images,
- Calculated with OpenCV.

$1\mu1P$ Vertex Candidates

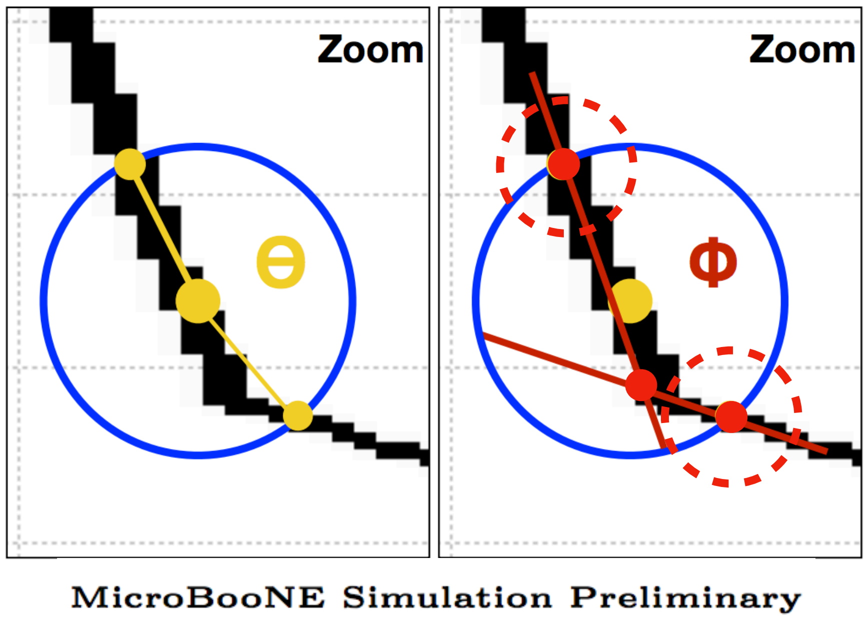




- Vertex Candidate #1: contour defect points,
- Vertex Candidate #2: PCA (principal component analysis) crossing points.

$1\mu 1P$ Angular Metric

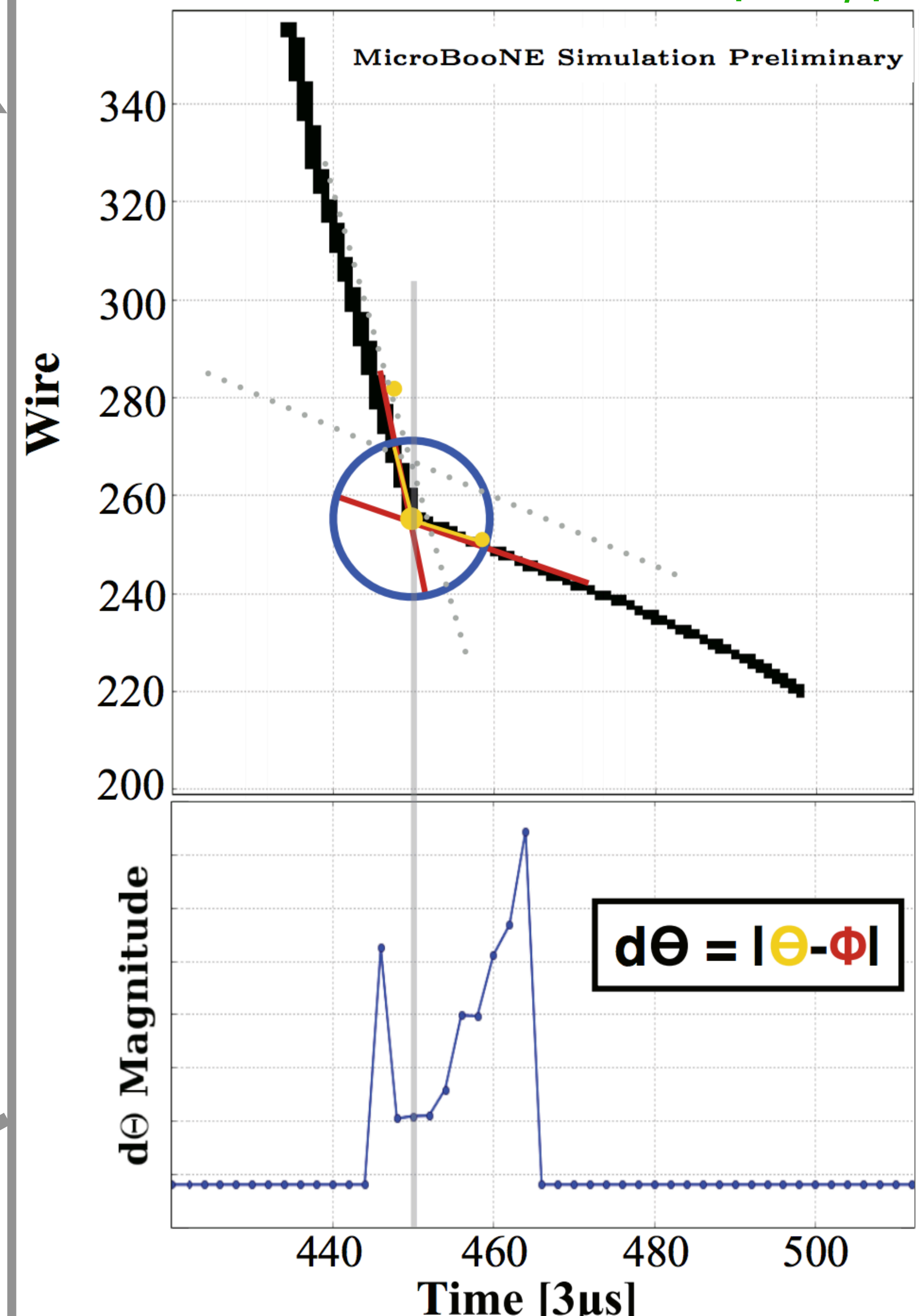
Angle Definition



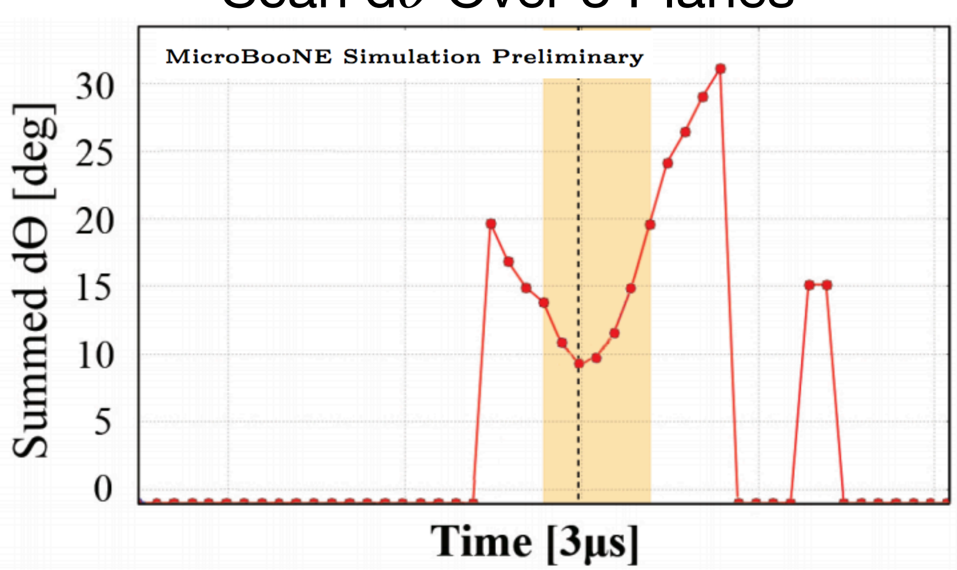
θ: angle of seed and crossing points.

ϕ: angle of PCA crossing points.

Find local minimum of $|\theta - \phi|$



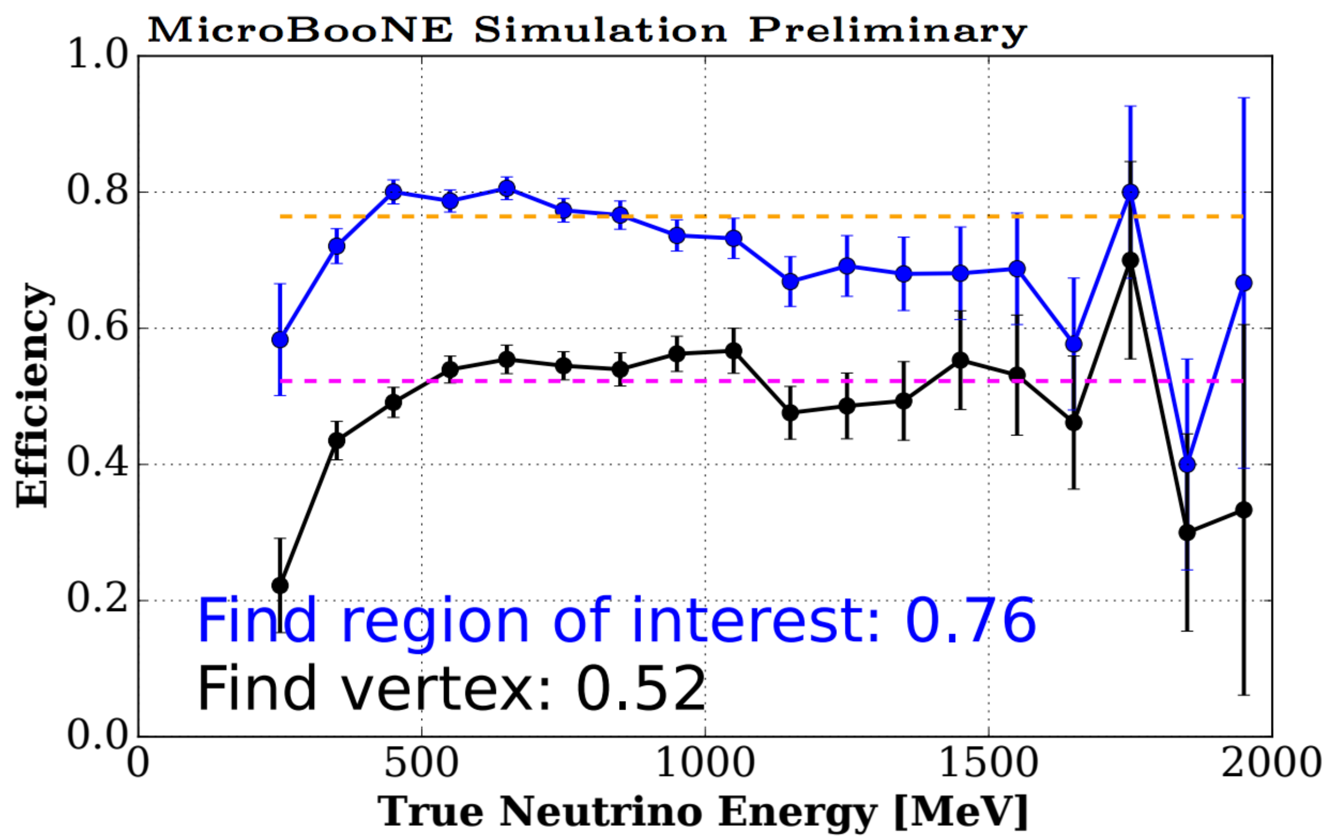
Scan dθ Over 3 Planes

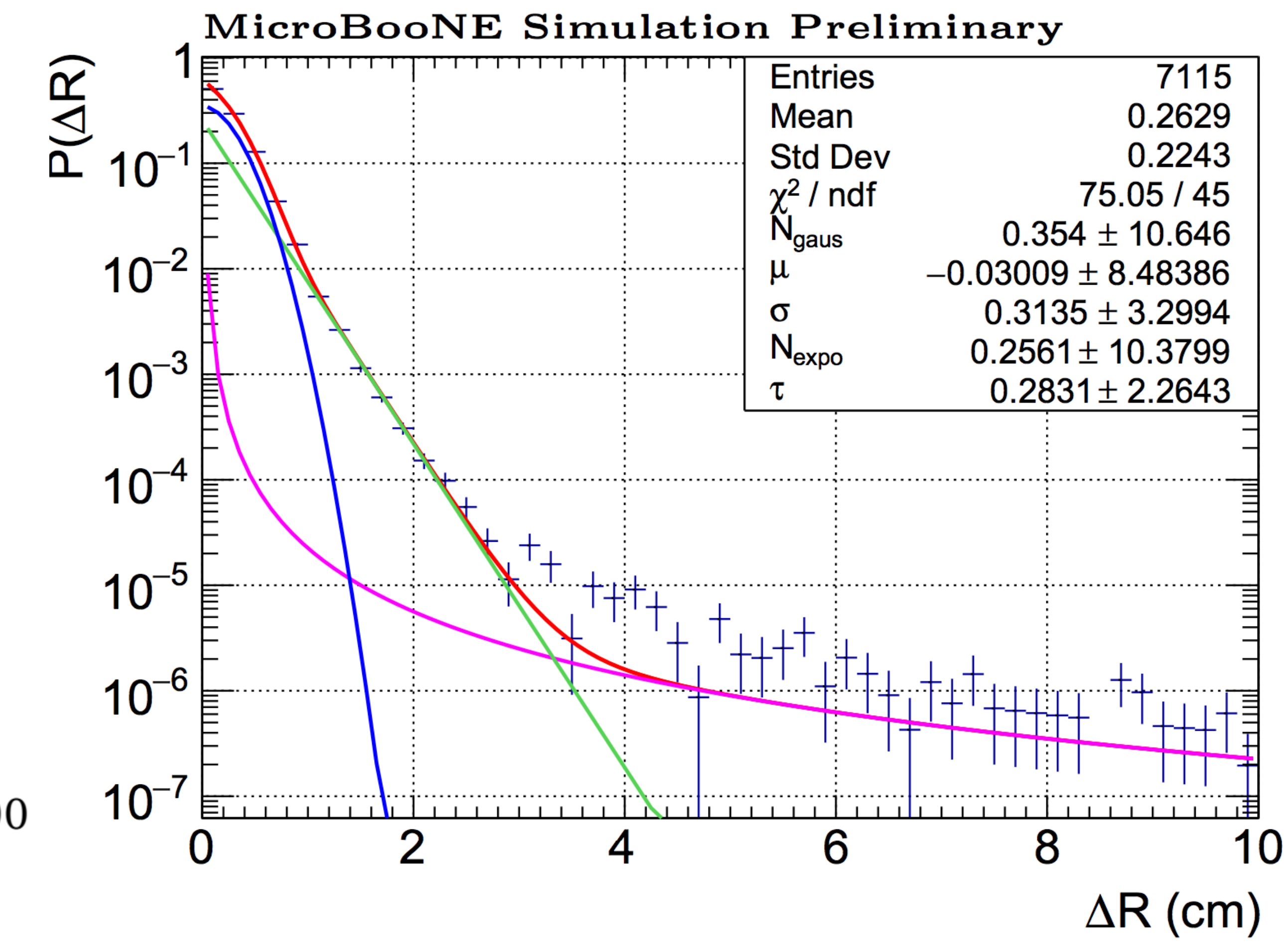


$1\mu1P$ Vertex Finding Performance

µBooNE

MICROBOONE-NOTE-1042-PUB

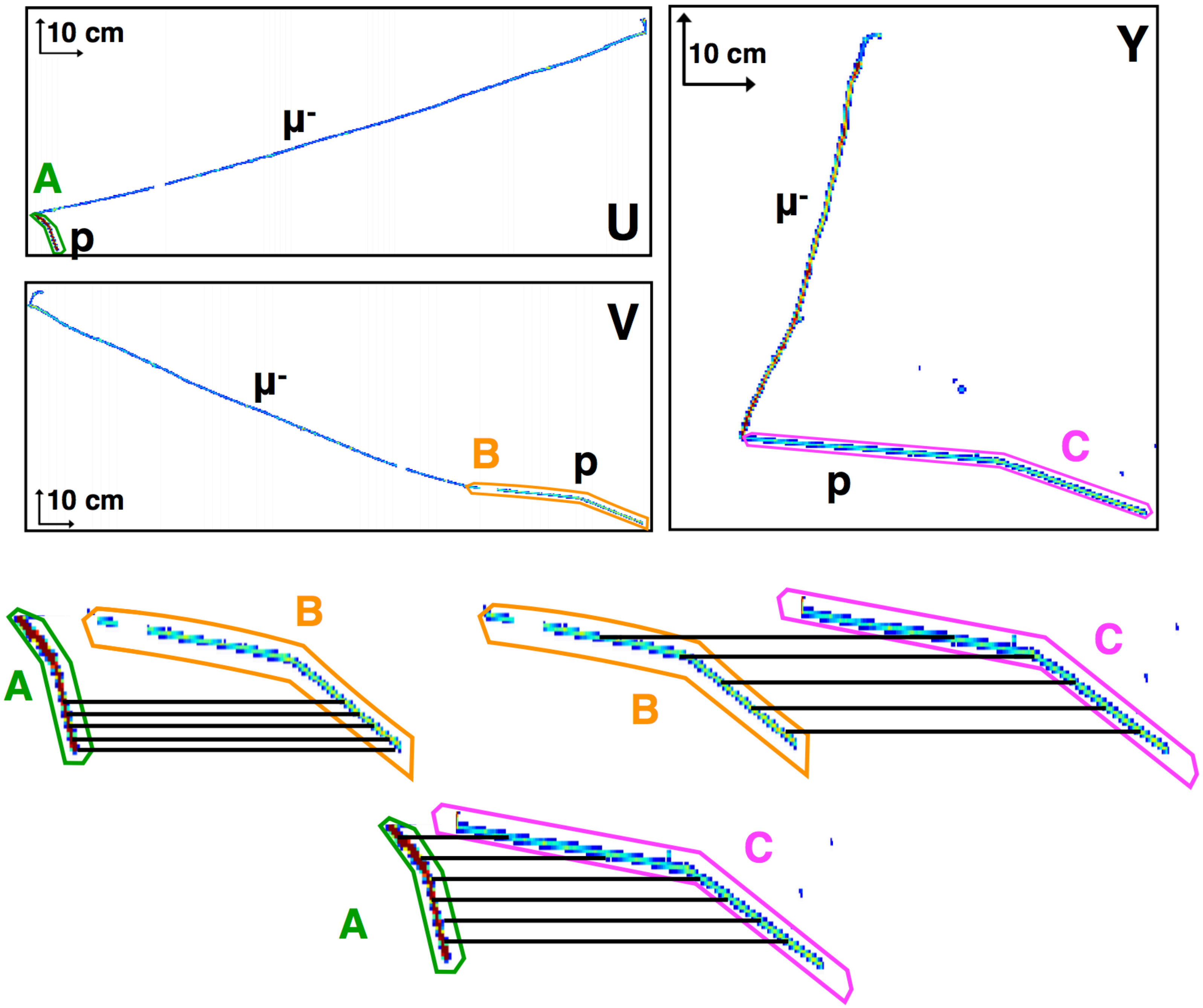




- Average efficiency of finding vertex at ~52%.
- Vertex finding resolution for $1\mu1p$ → less than 0.73cm for 68% events.
- Shower vertex finding procedure in backup slides.

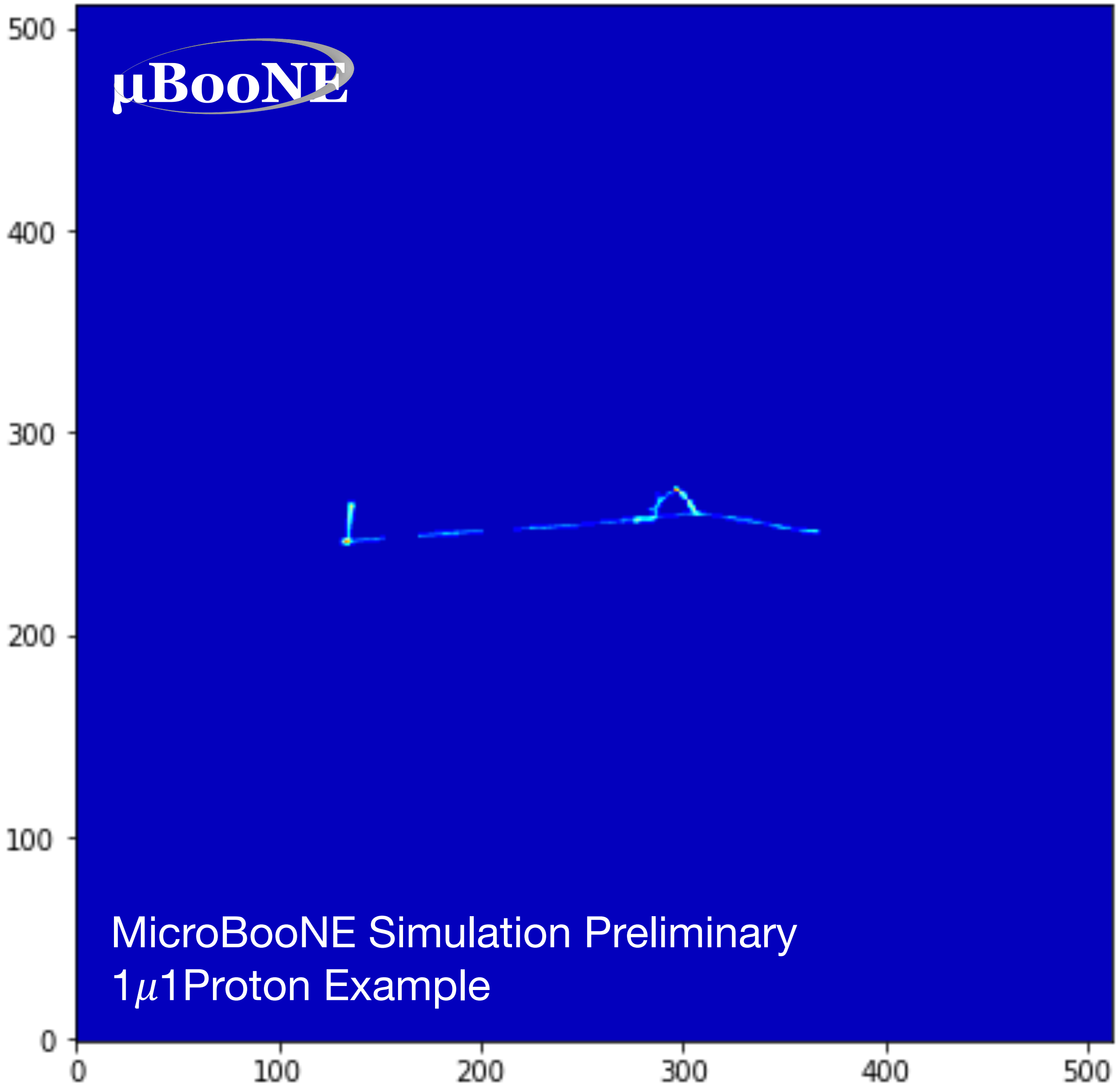
Track Particle Clustering

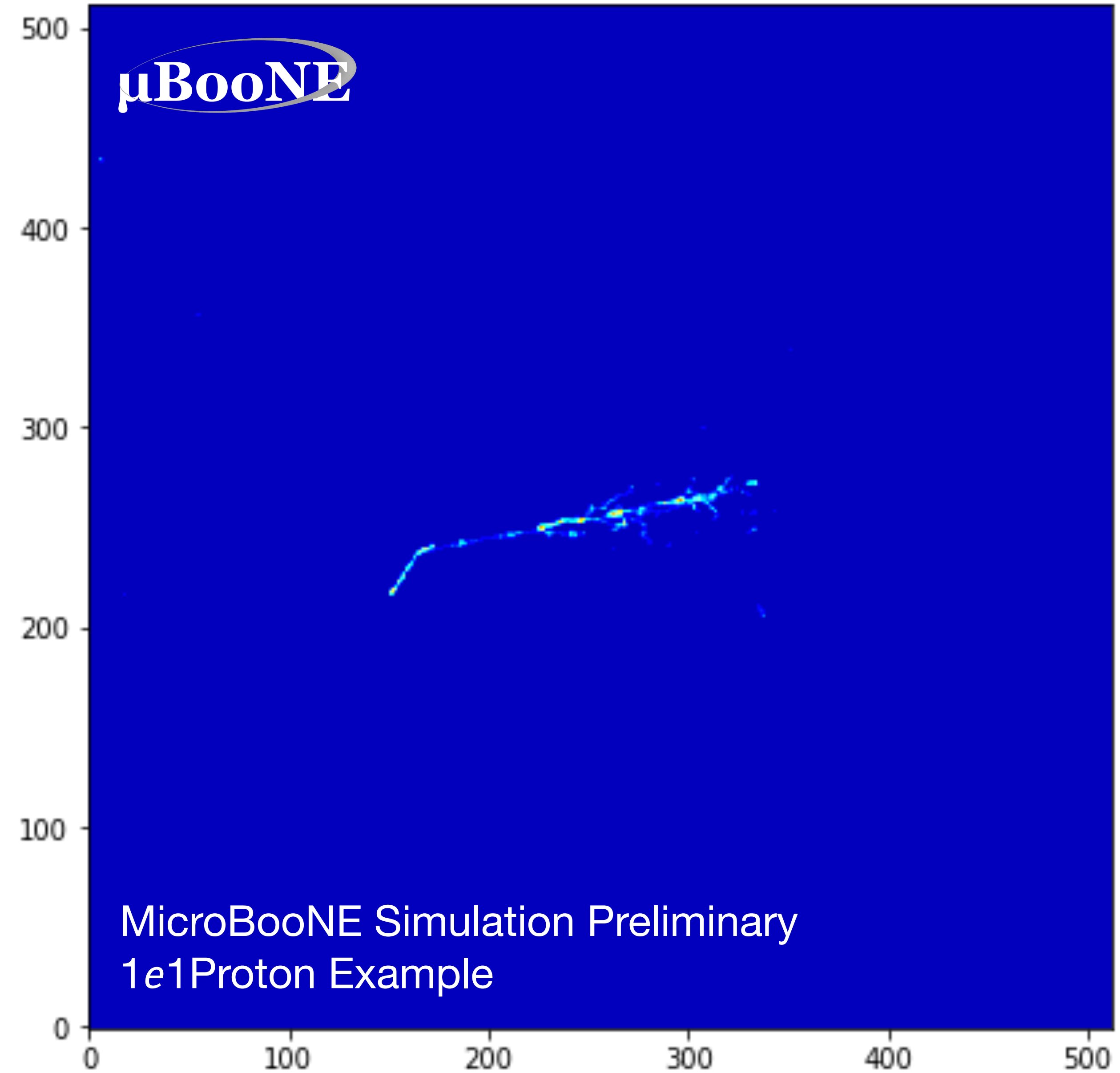
MicroBooNE Simulation Preliminary



- Pixels are clustered based on their distance and angle under the polar coordinate.
- Clusters across three planes are further evaluated by matching pixels over time tick and wire crossing.

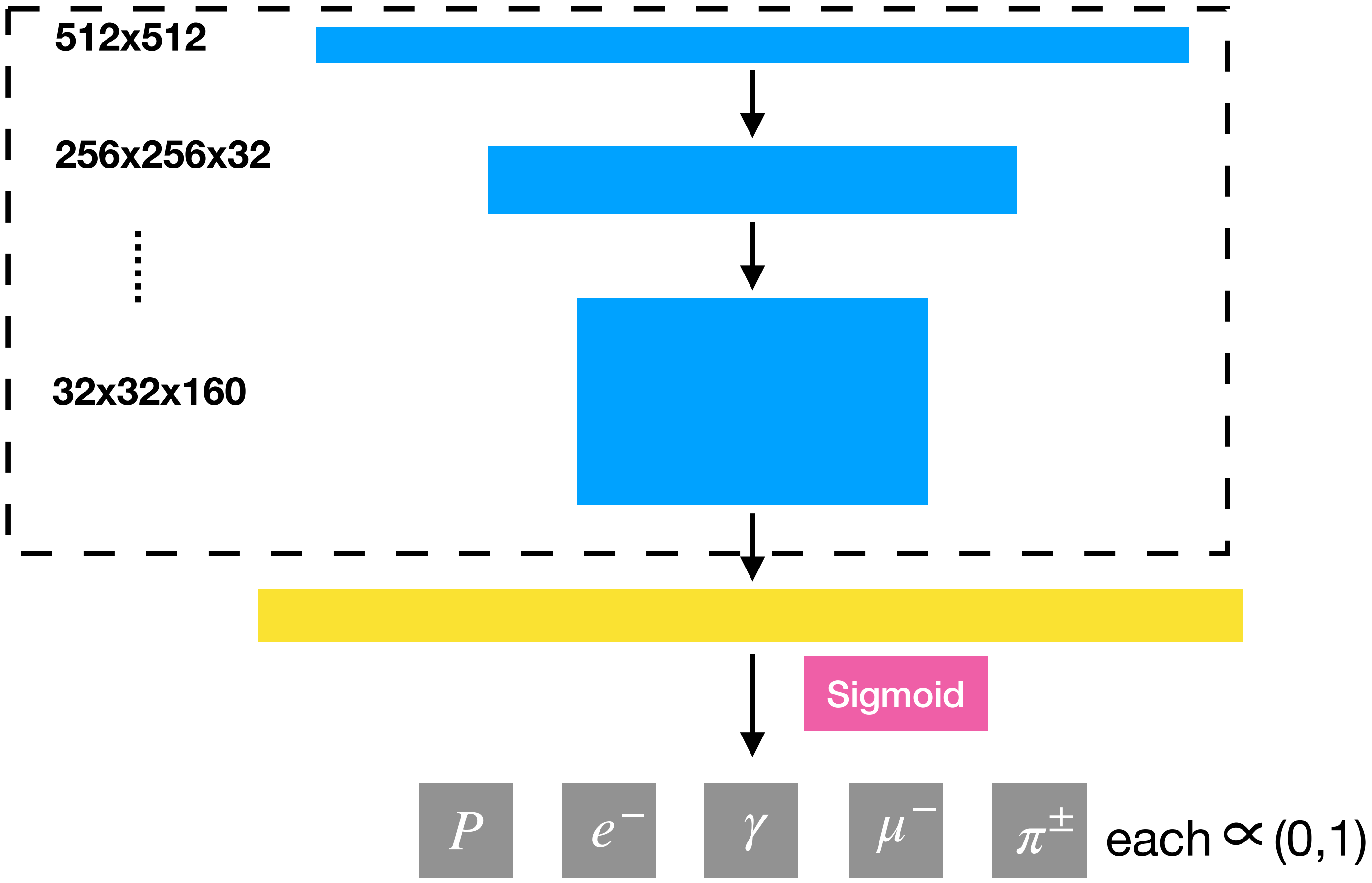
Reconstruction Example

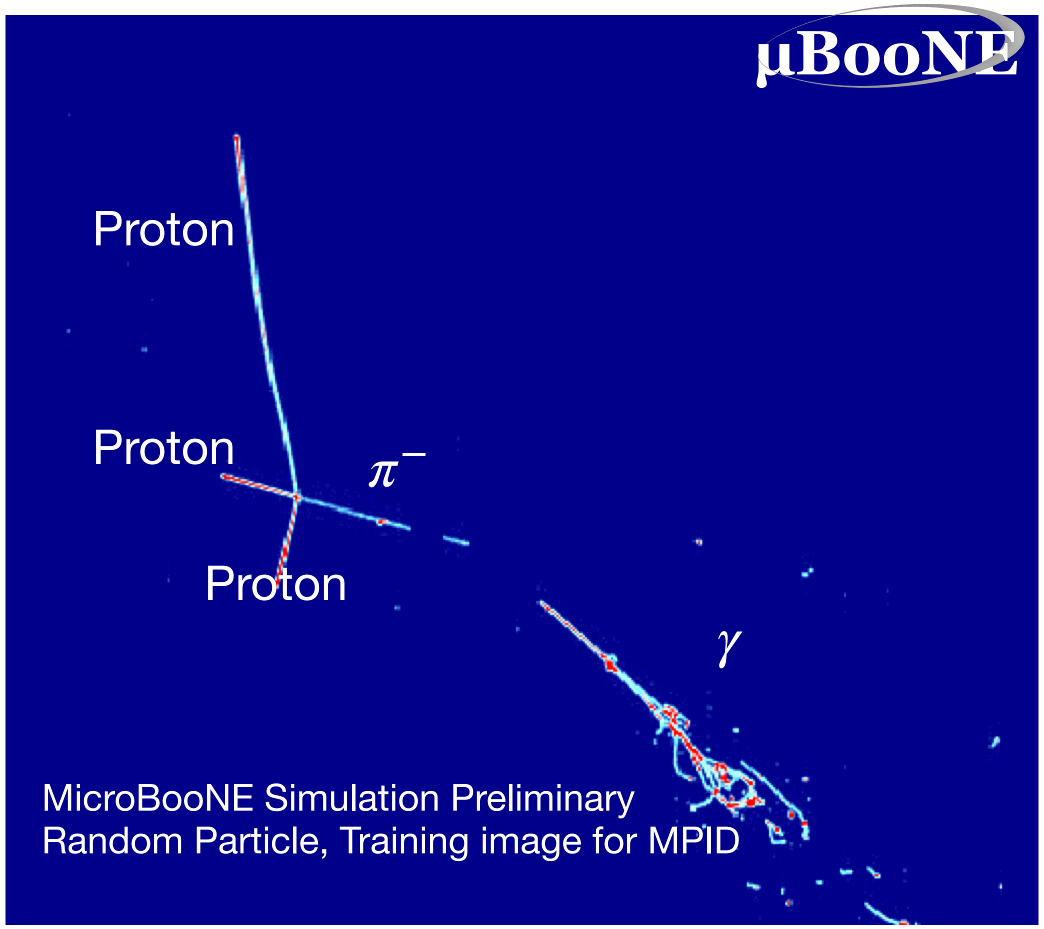




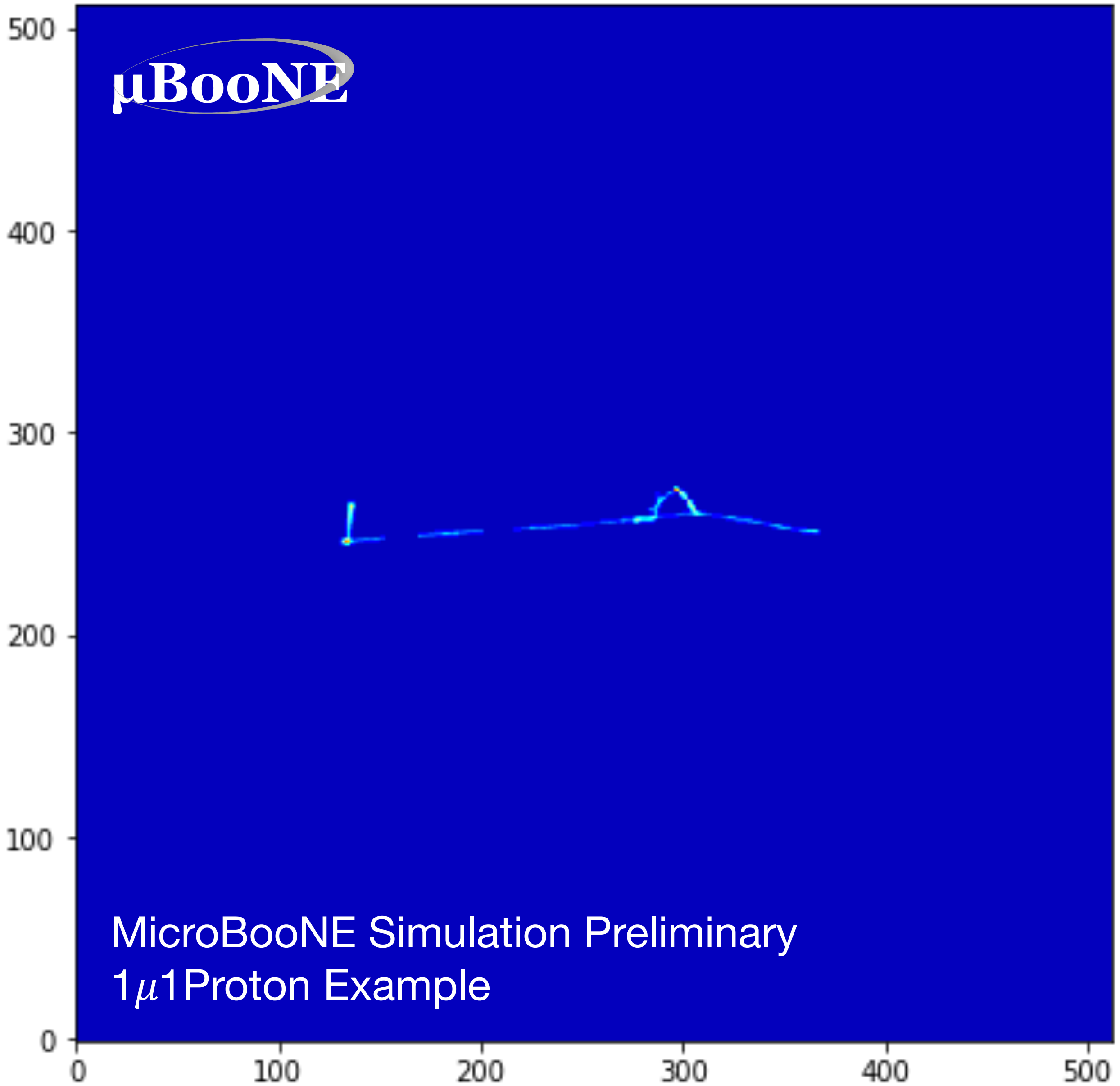
Multi-particle PID Network

- Multi-particle PID network is a CNN network application, extended work of the single-particle PID network in arXiv: 1611.05531.
- Final layer in a sigmoid function and predicts probabilities of particles in the input image.
- Training image, simulated multiple particles coming from one vertex,
  - ⋆ avoid bias from neutrino mode,
  - ⋆ learns more from richer topology information.
- Training sample size: 50,000 events. Test sample size: 40,000 events.
- Better as,
  - ⋆ does not require a vertex resolution,
  - ⋆ does better on Pi0 present events (hard to reconstruct detached shower).

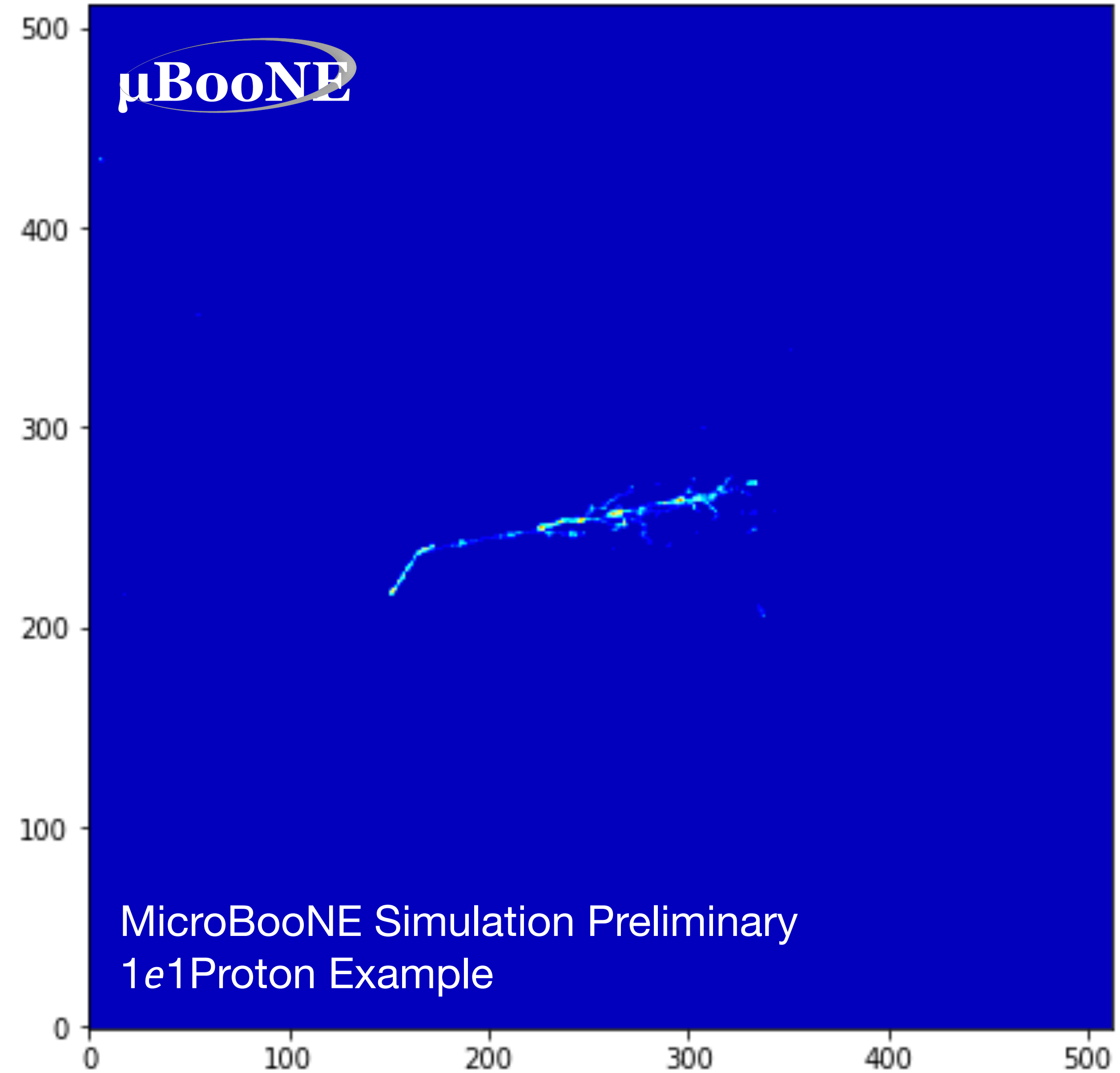




Multi-particle PID Network

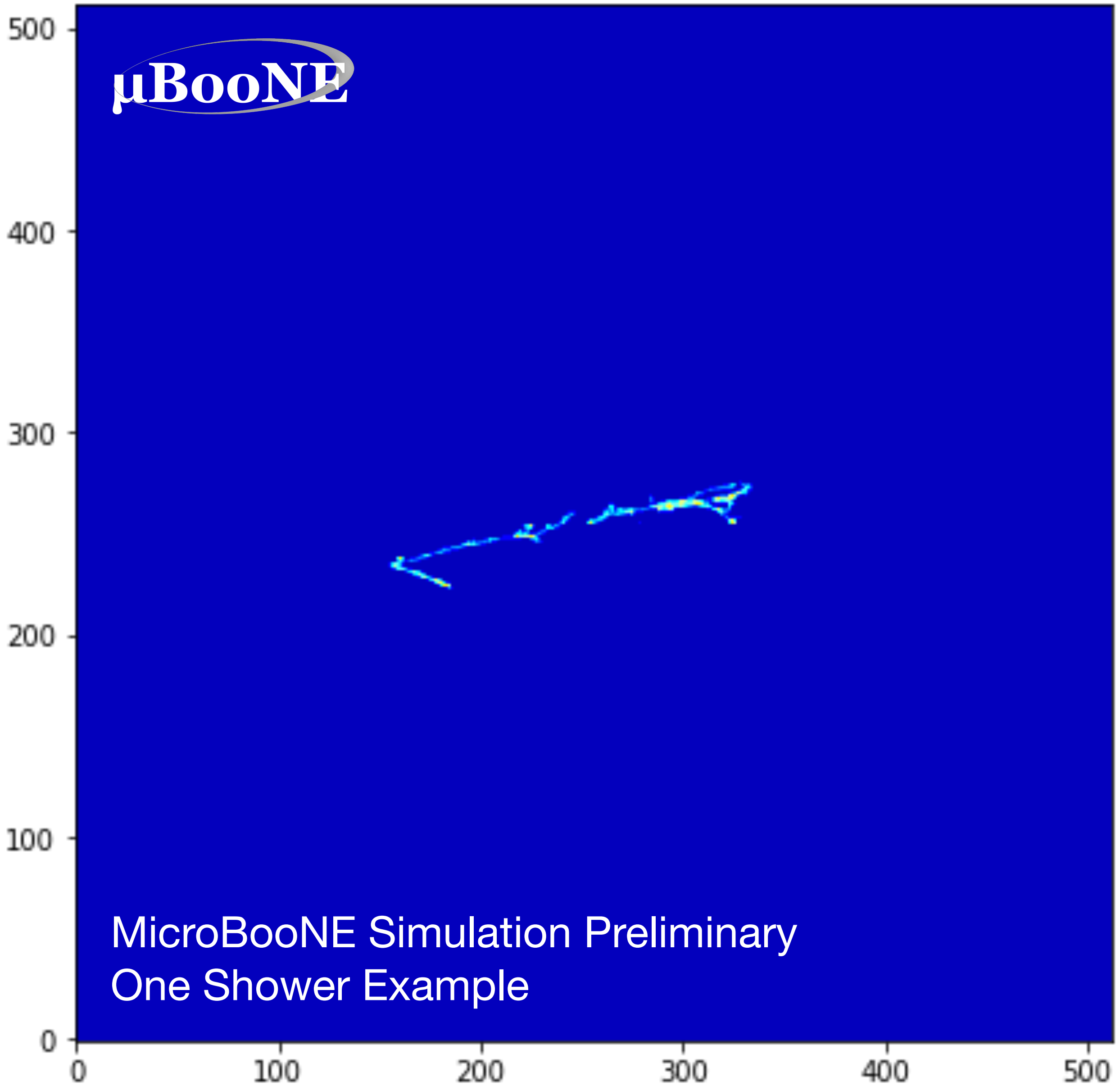


e^-	γ	μ^-	$\pi^\pm$	P^+
0.1	0.16	0.86	0.38	0.77

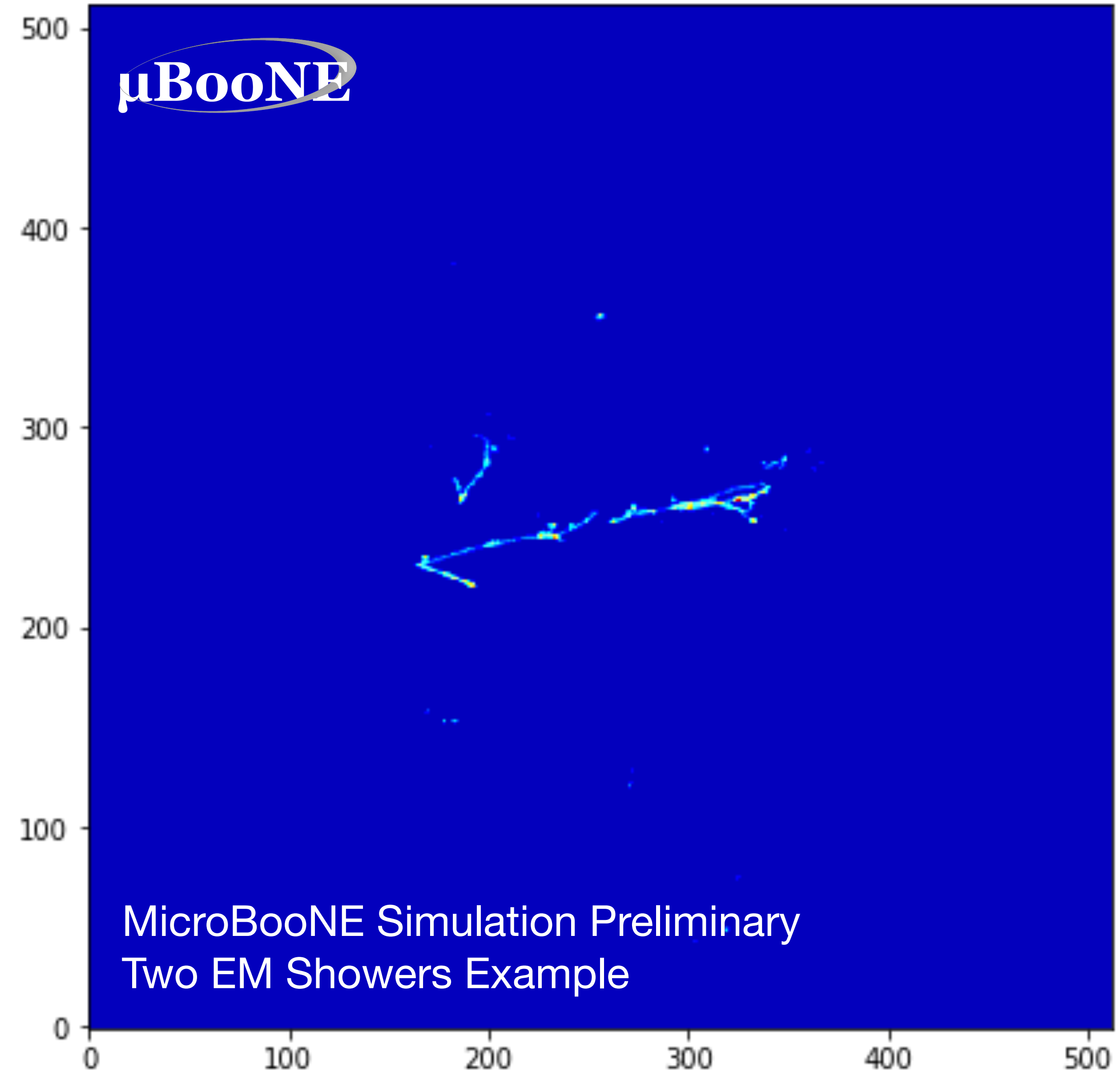


e^-	γ	μ^-	$\pi^\pm$	P^+
0.92	0.23	0.35	0.07	0.74

Multi-particle PID Network

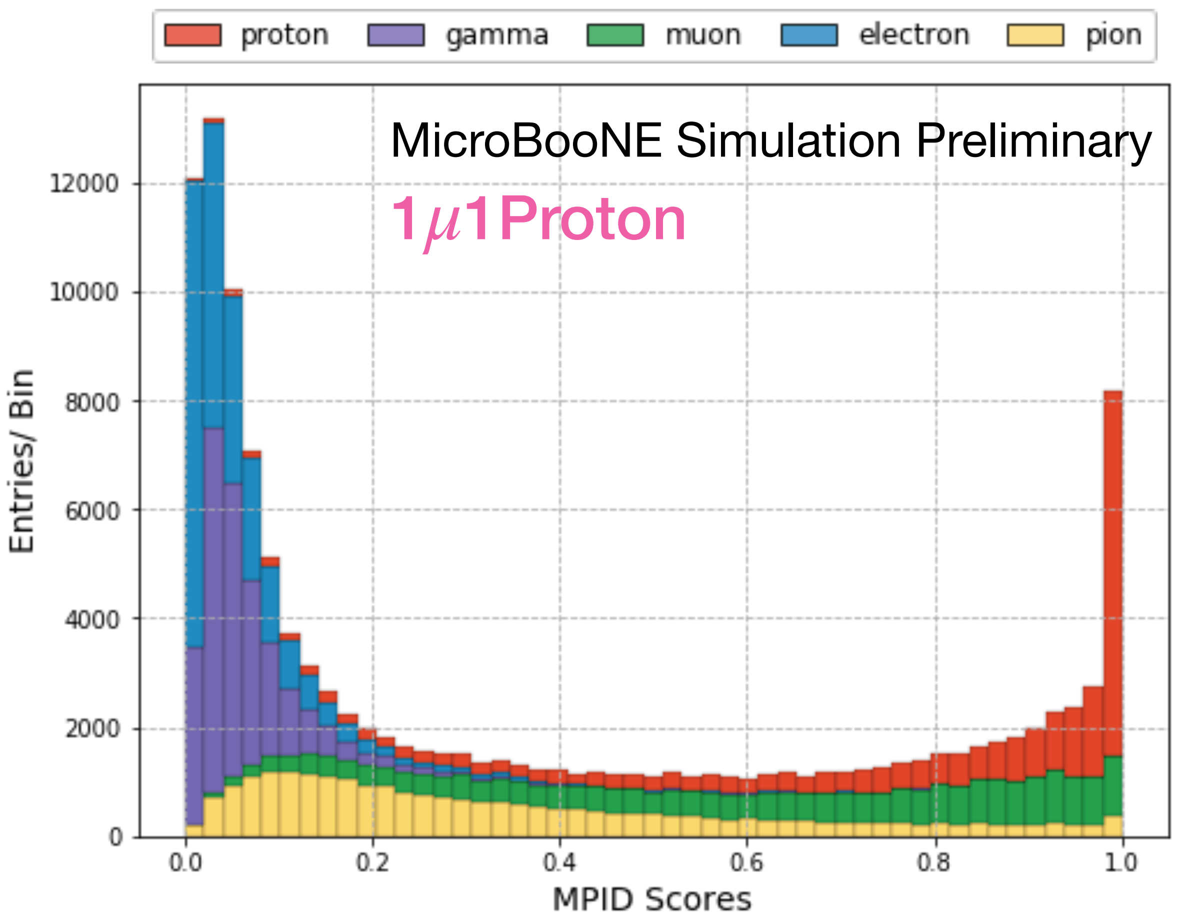


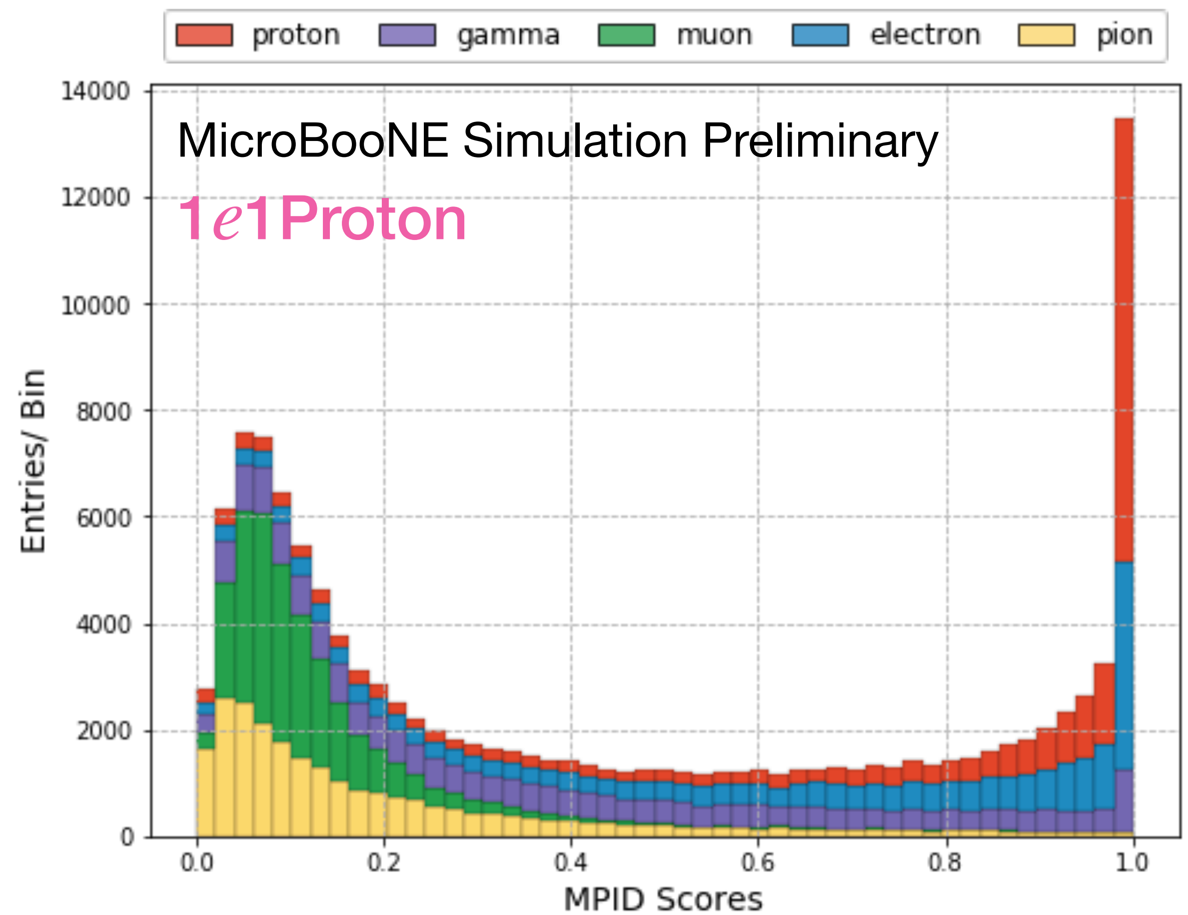
e^-	γ	μ^-	$\pi^\pm$	P^+
0.82	0.4	0.15	0.02	0.72



e^-	γ	μ^-	$\pi^\pm$	P^+
0.64	0.71	0.12	0.23	0.5

MPID on Simulation

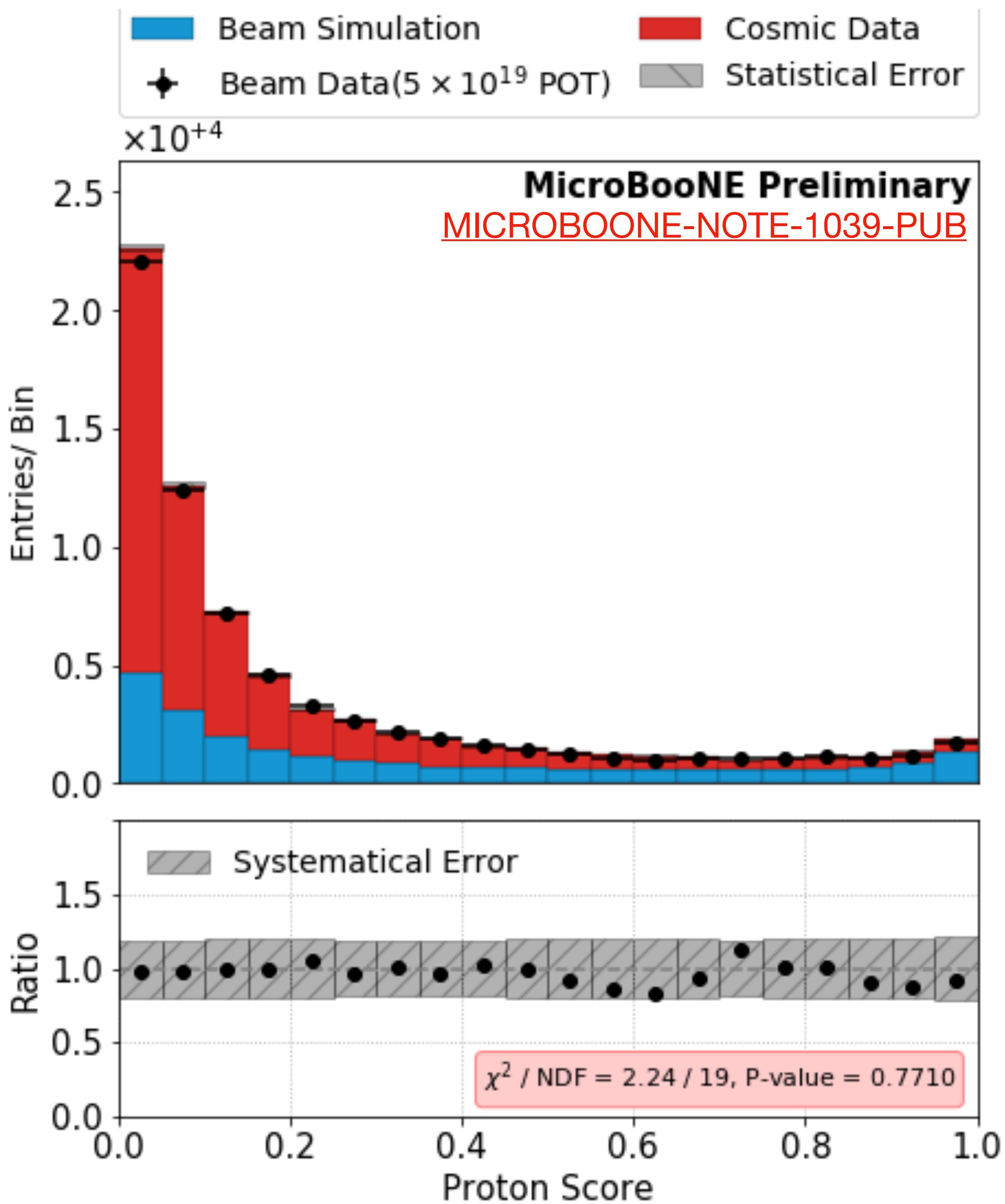




- Good separation between shower-like and track-like particles,
- Good separation between track-like particles:
  - ★ proton, muon and charged pion.
- Good separation between track-like particles:
  - ★ electron and gamma,

Data Simulation Comparison for MPID

μBooNE



- Using $1\mu1P$ as a sideband study for MPID data MC comparison.
- Good comparison between MicroBooNE's open data and beam simulation + cosmic.
- Syst. error includes applying MPID to same events with various E-field, SCE, channel noise etc.
- In proton score distribution, there is a bump at 1 introduced by neutrino events.

New Reconstruction Chain

µBooNE

Cosmic Tagging

↓

Shower / Track Pixel Labeling

↓

2D & 3D Reconstruction

↙ ↘

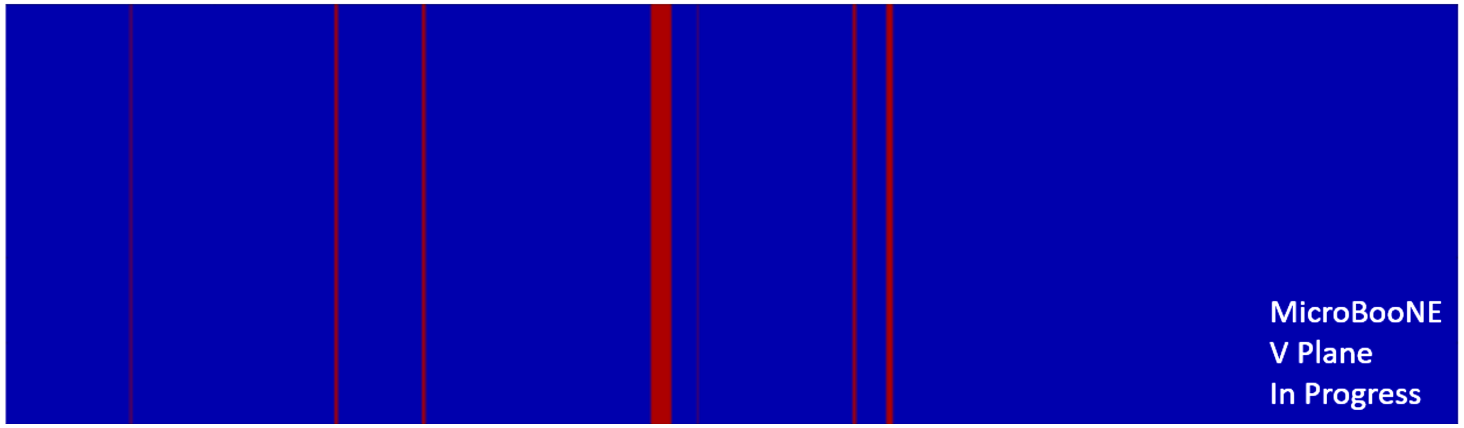
PID using CNN

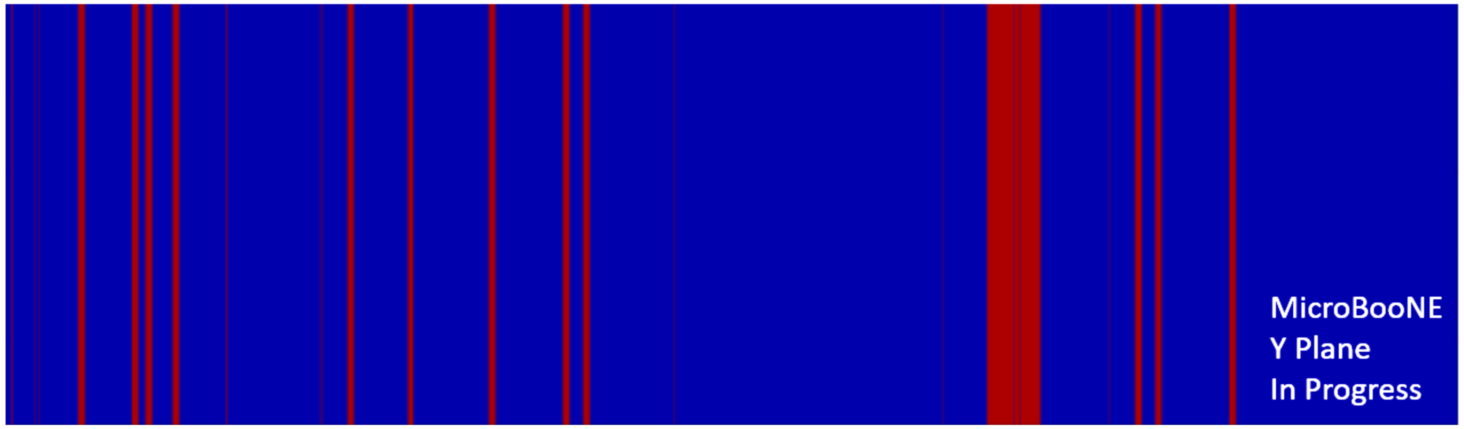
$1\mu 1$p

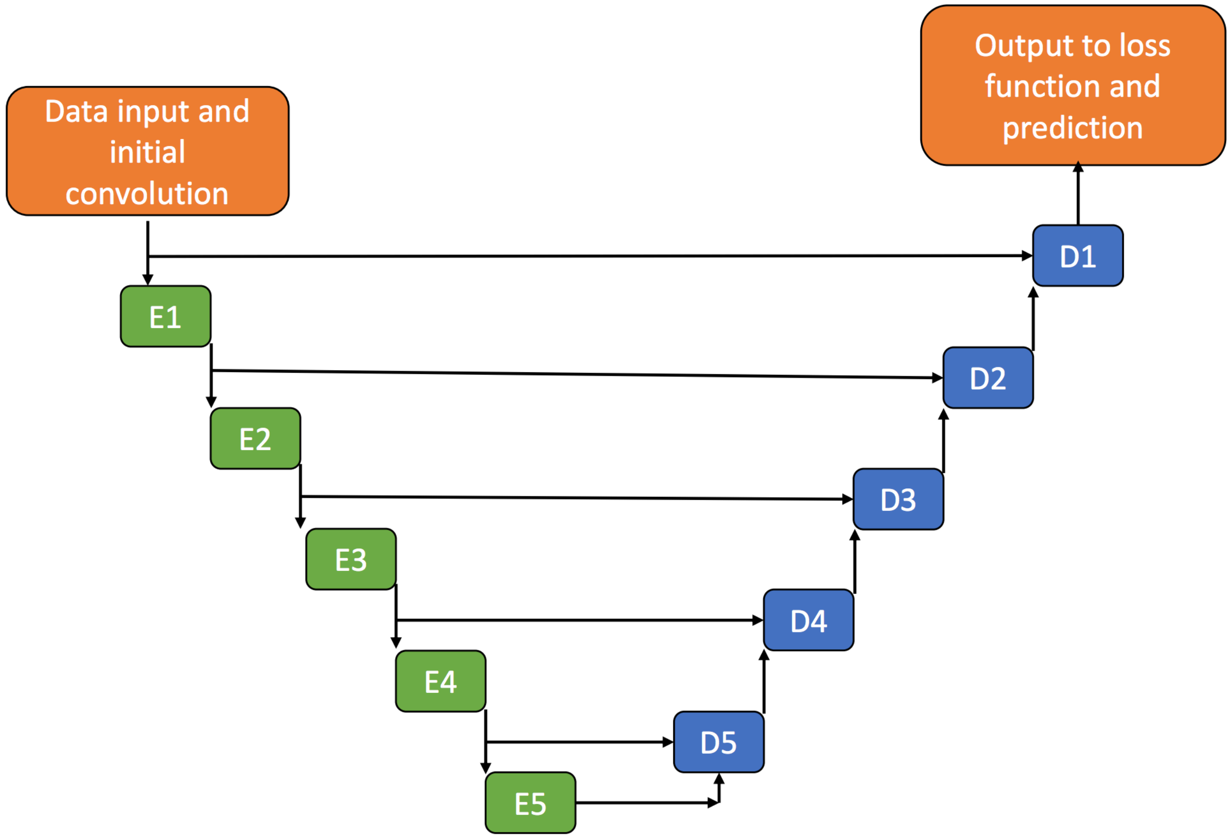
↓

$1e1$p

Infill Network

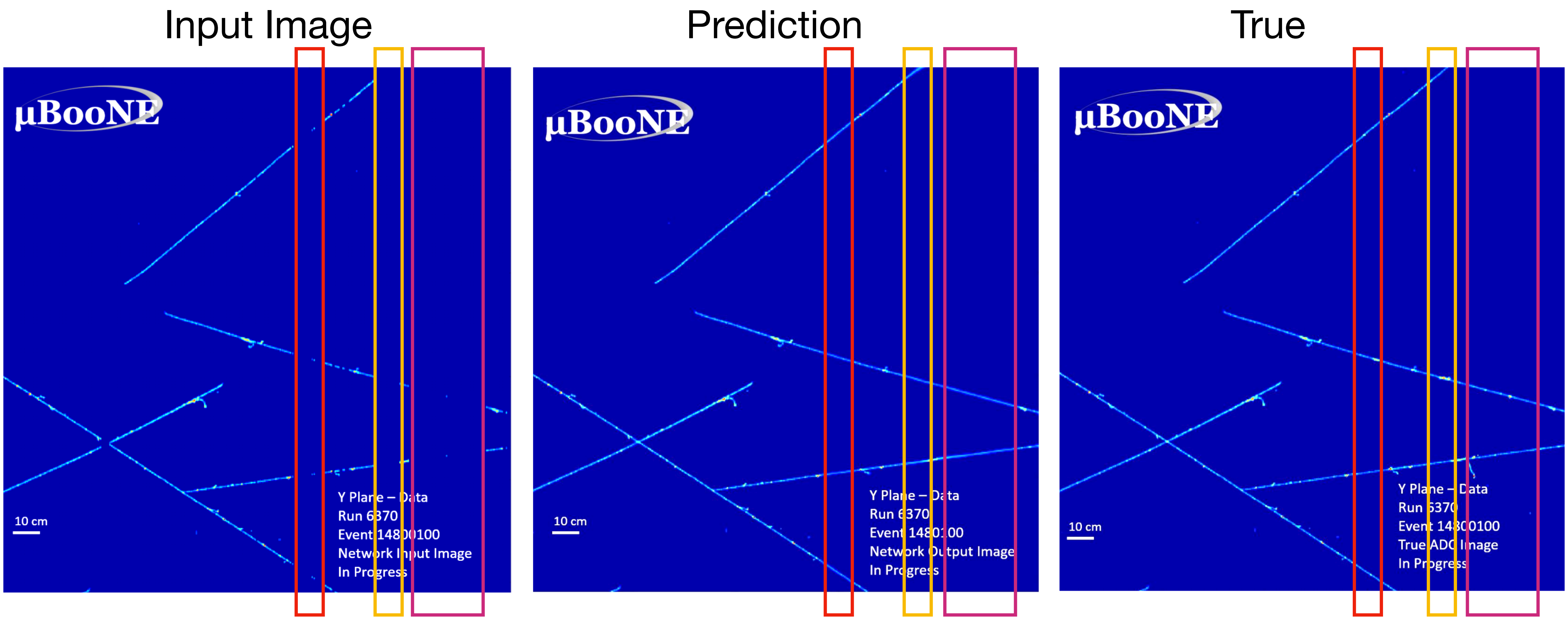






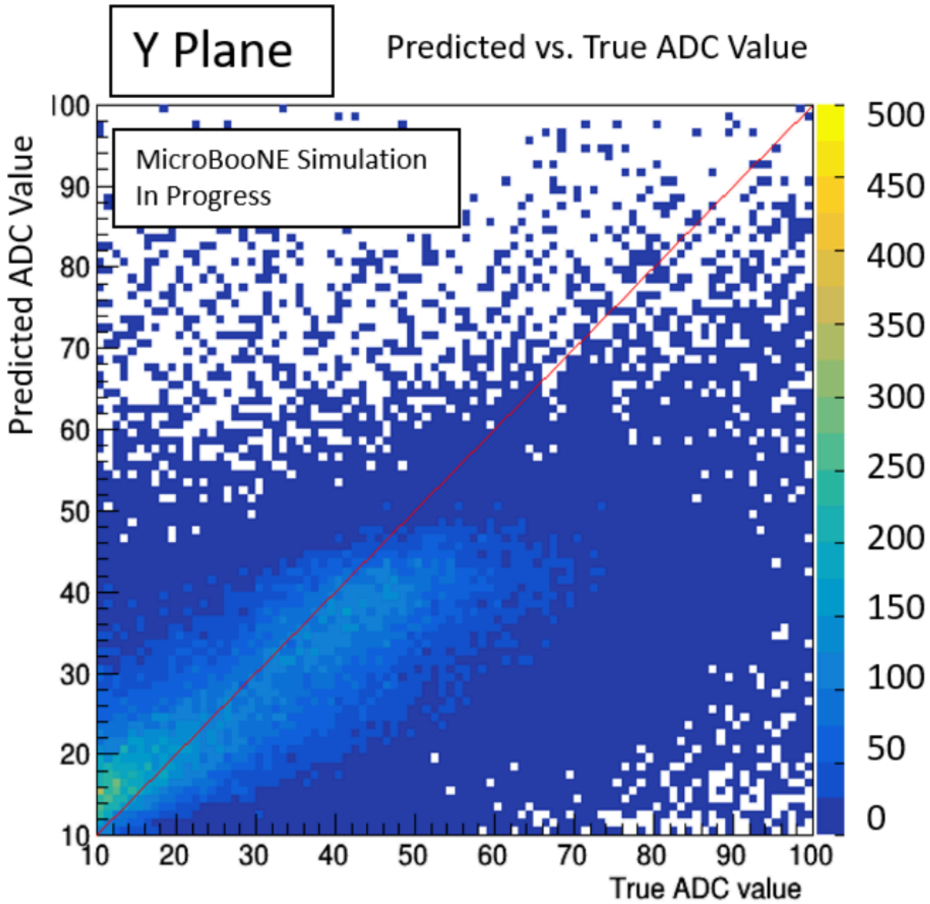
- MicroBooNE has large number of dead wires.
- Infill network tries to fill these region with realistic values and improve tracking performance for cosmic muon.
- Training on off-beam data plus dead wire pattern crop.
- U-net + encoder and decoder layers.

Infill Network



- MicroBooNE simulation example. Colored boxes are dead wire regions.
- Pixels are predicted in reasonable region.

Infill Network



	U Plane (3 epochs)	V Plane (3 epochs)	Y Plane (4 epochs)
Within 2 ADC	27.24%	20.22%	25.20%
Within 5 ADC	45.33%	37.36%	43.83%
Within 10 ADC	66.82%	56.37%	66.73%
Within 20 ADC	84.47%	75.14%	84.88%
Binary Accuracy	98.40%	99.13%	99.16%

- Predicated pixel ADCs are in a reasonable region and have good efficiencies.

Mask RCNN

- Mask Regional-CNN, a model well known for its good performance on instance segmentation, can provide:
  - ★ ROIs for objects,
  - ★ Label for the object,
  - ★ Pixels for the object.
- Researched with MicroBooNE data in two ways:
  - ★ Mask RCNN-Ancestor: Neutrino vs. Cosmic.
  - ★ Mask RCNN-Particle:
    Four particle types of e^--like, μ^-, $\pi^\pm$ and proton.

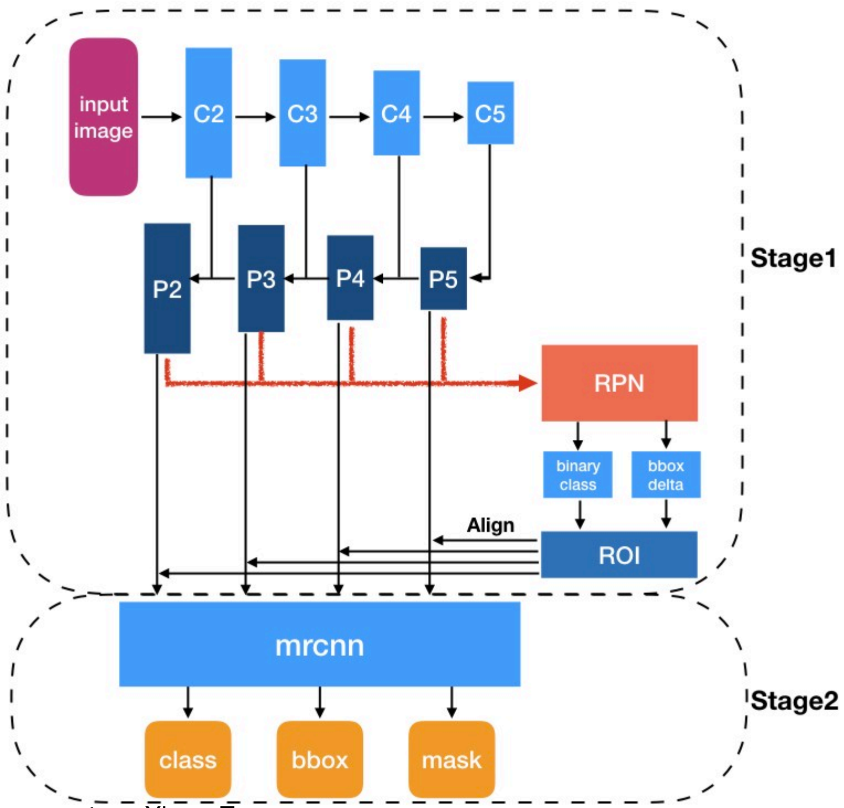
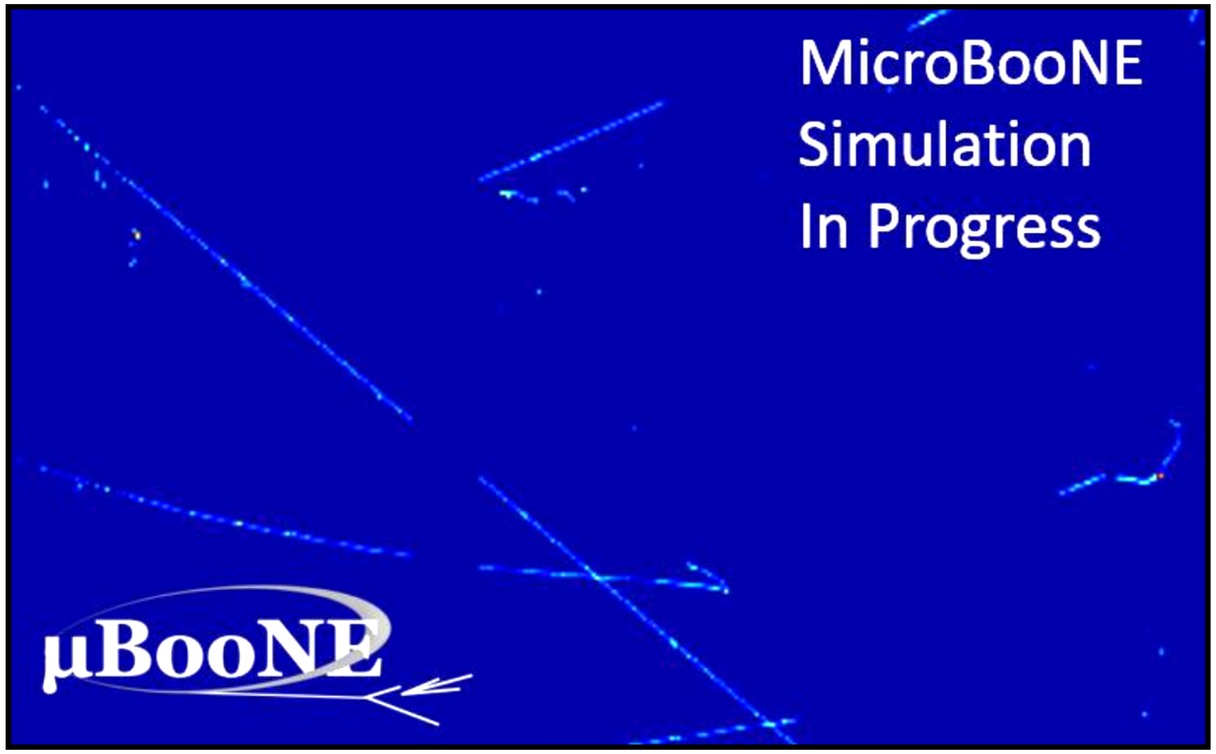


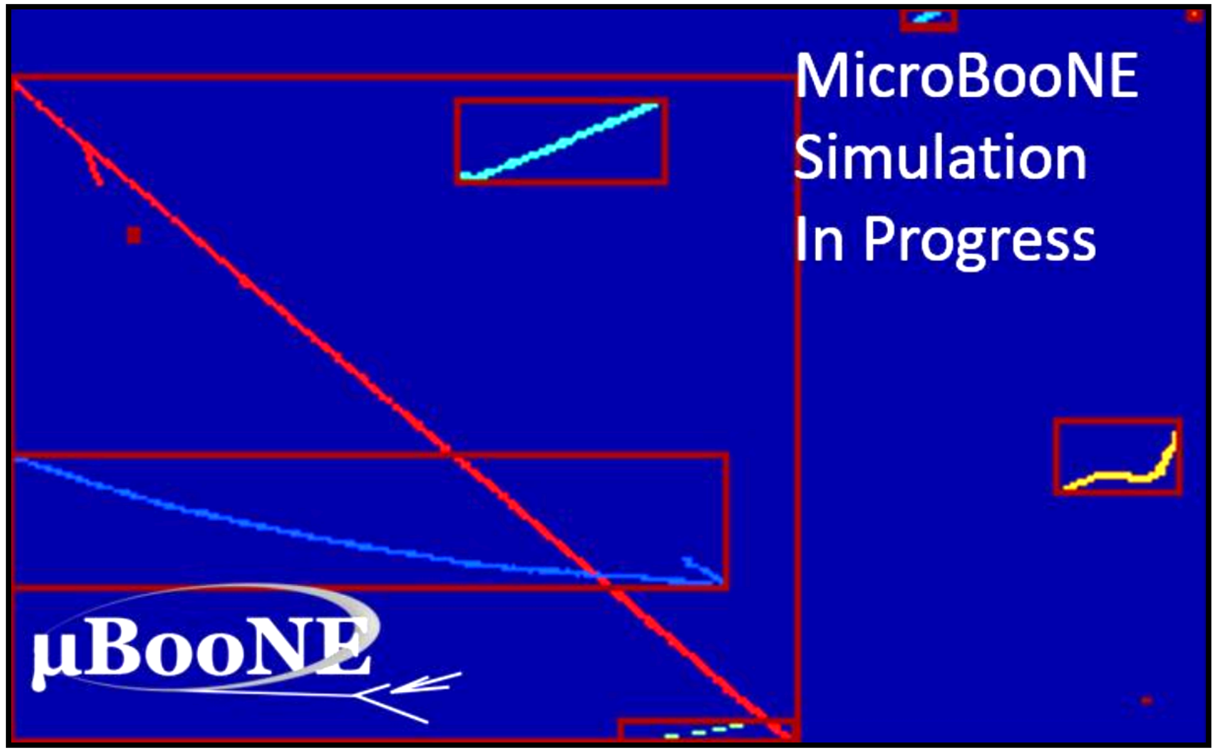
Image courtesy Xiang Z.

Mask RCNN-Ancestor

ADC Image

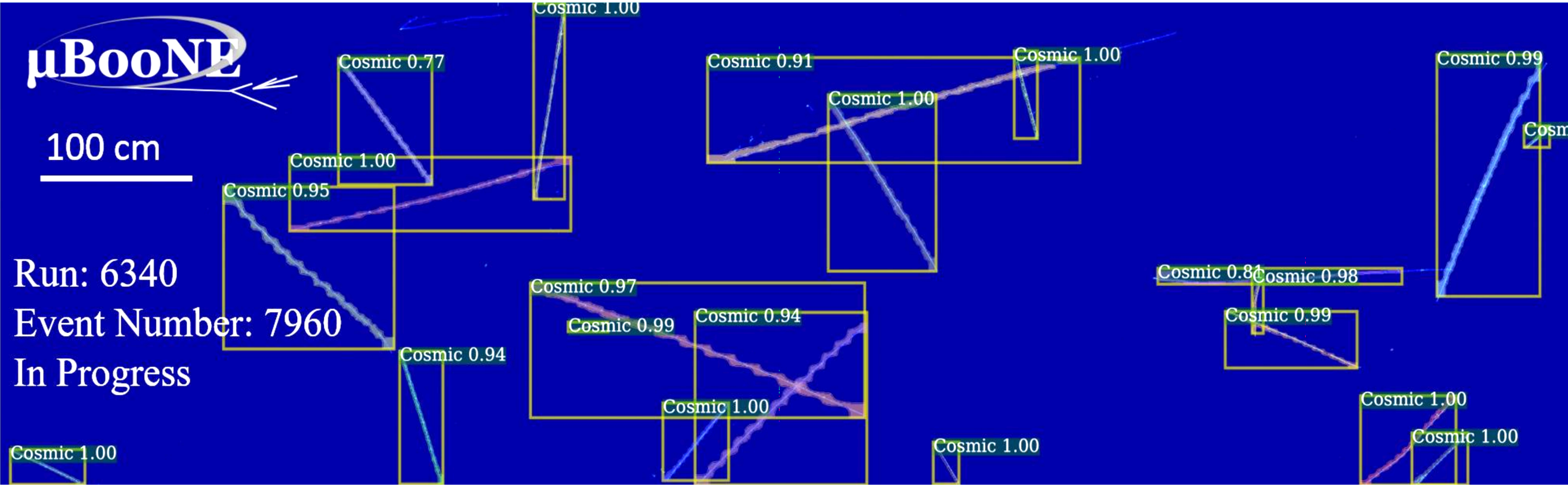


Training Image



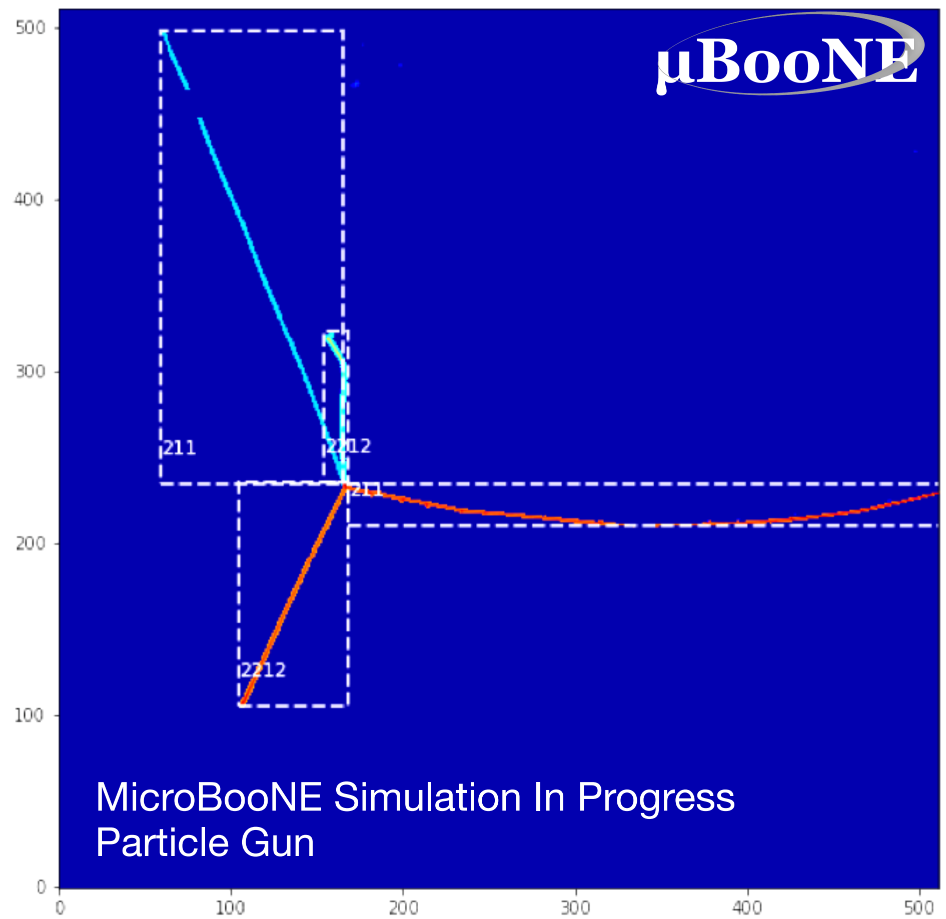
- Training image size of 512x832 cropped from full event display.
- Training sample has 93% cosmic, 3% neutrino and 4% other.
- Training with ResNet-50 pre-trained weights.

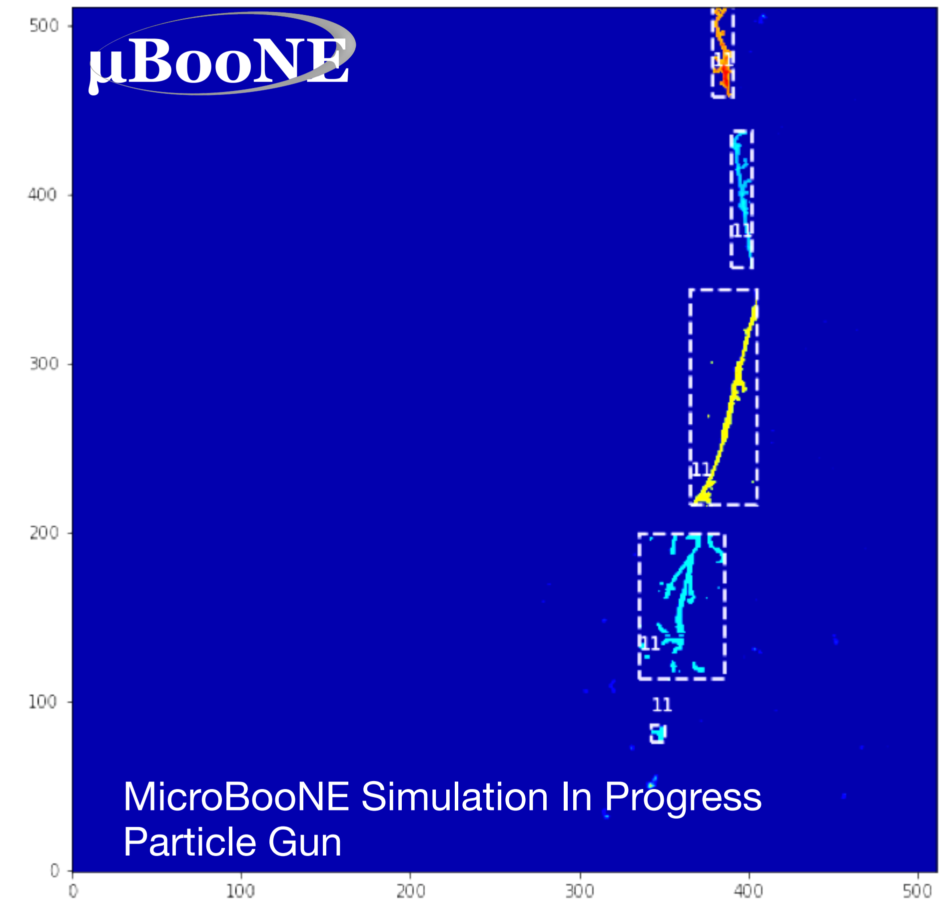
Mask RCNN-Ancestor



- Applied Mask RCNN-Ancestor to MicroBooNE data.
- ~90% particle covered by the proposed clusters of network output.
- Proposed cluster maps strongly to one particle
- Great tool for cosmic pixel tagging.

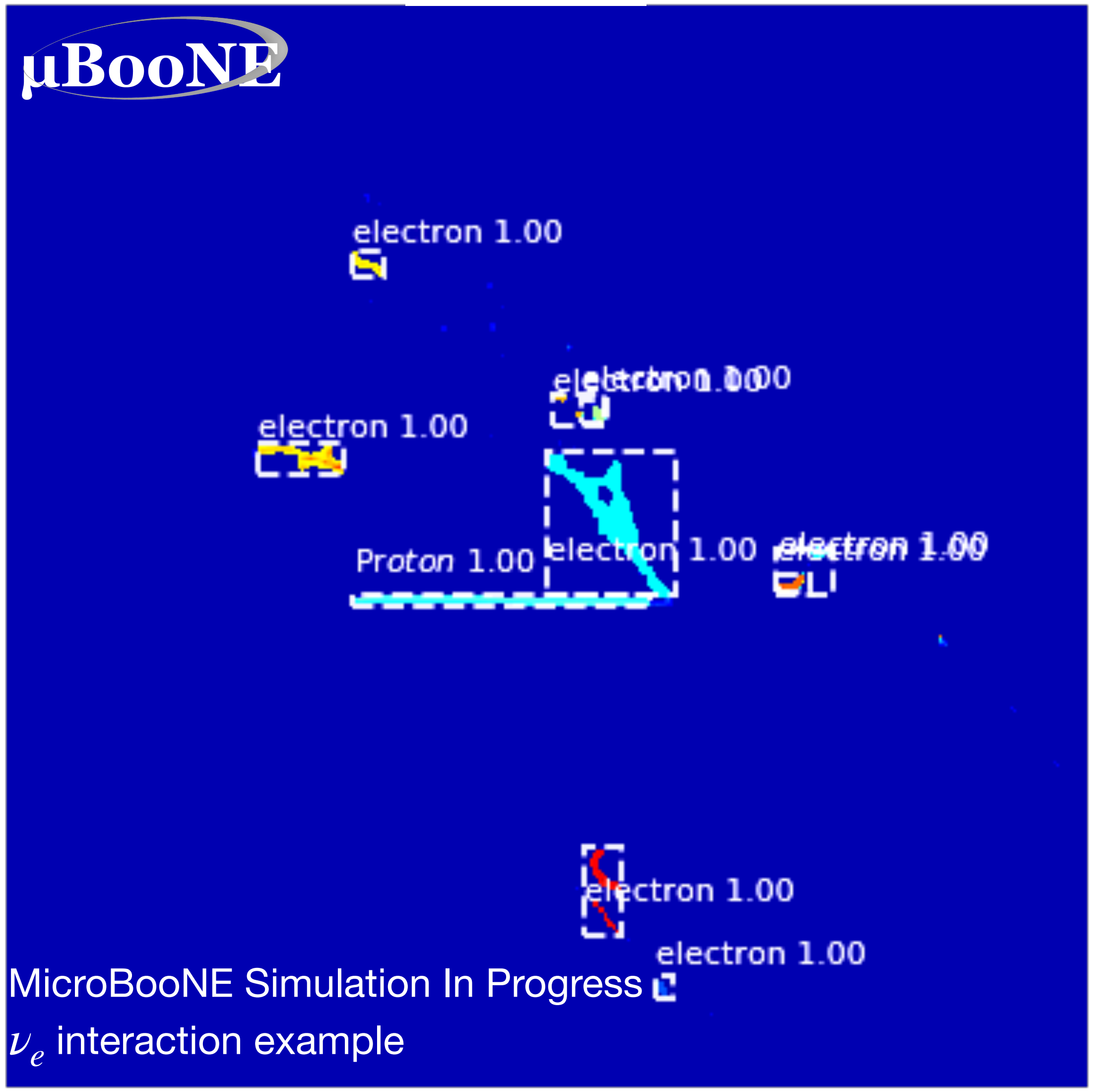
Mask RCNN-Particle

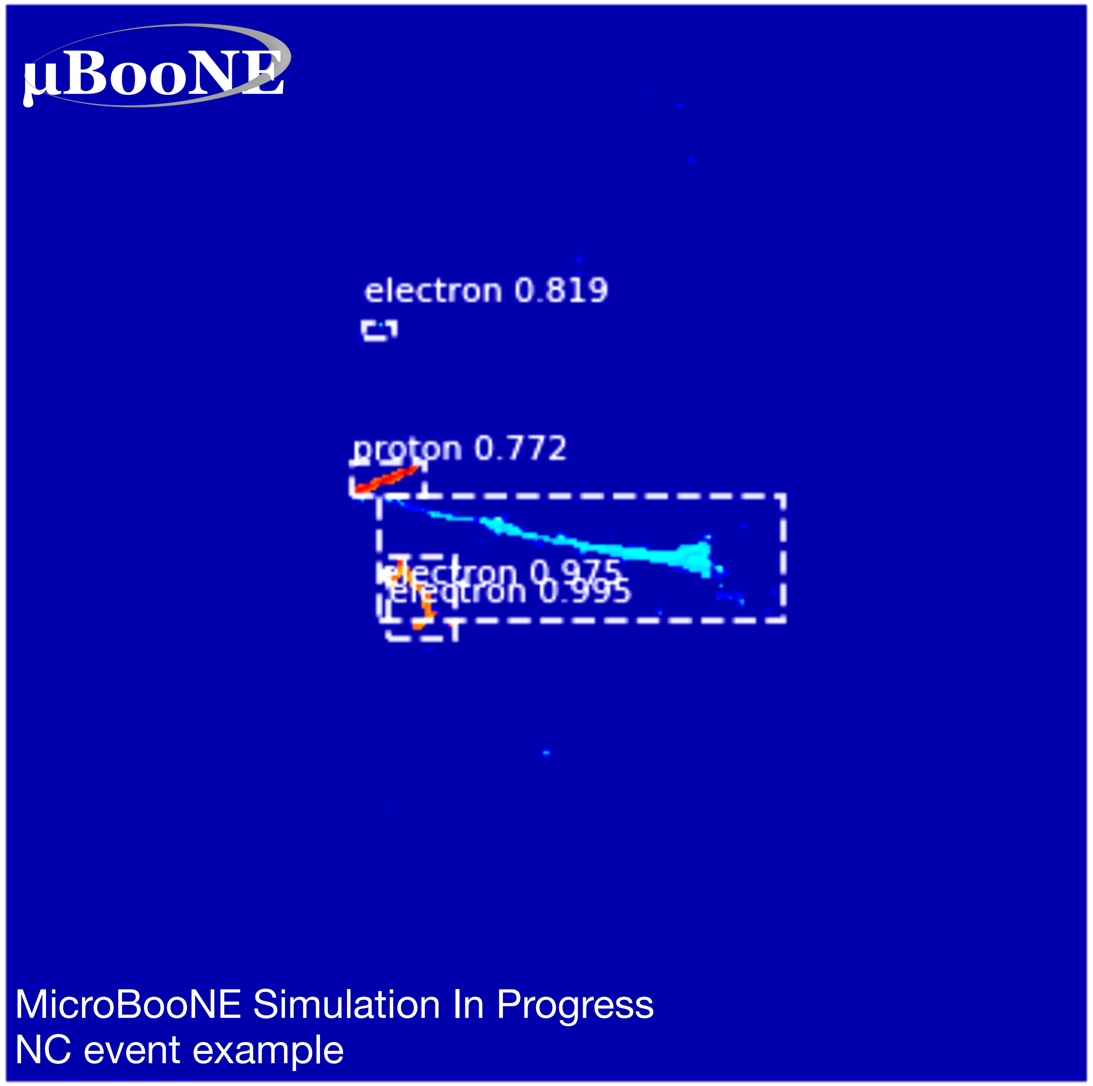




- DBScan applied to EM shower particles to break large sparse ROIs introduced by EM showers,
- Four particle types in training sample, P, μ^-, $\pi^\pm$ and electron-like shower piece,
- Applied similar pixel weighted as SSNet and a class weighted loss (to unbias the dominating electron labels from DBScan),
- Trained with ResNet-101 weights from COCO, improving the ROI proposal on tracks.

Mask RCNN-Particle





- Mask RCNN-particle shows potential in improving particle clustering, especially for event with detached shower.
- Challenge: the center-based ROI proposal step does not seem ideal for particles coming from one vertex (e.g. compared to corner-based ROI proposal).

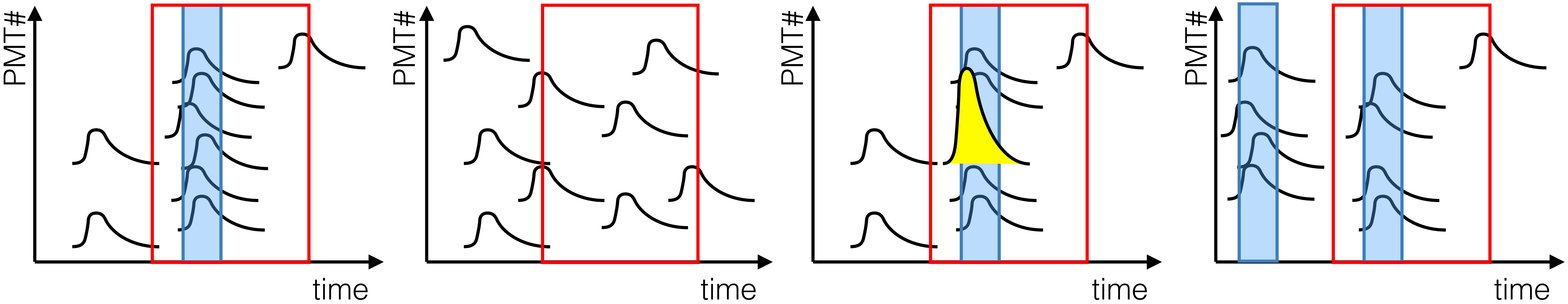
Conclusion

- MicroBooNE has developed an imaged-based reconstruction tool for 1μ1Proton and $1e$1Proton topologies,
  - ★ on the way to a LEE result,
  - ★ being applied to higher multiplicity topologies.

- A various of deep learning networks and how to apply the output have been researched by MicroBooNE.
- Good demonstrations of applying deep learning to LArTPC.

Back UP

PMT PreCuts

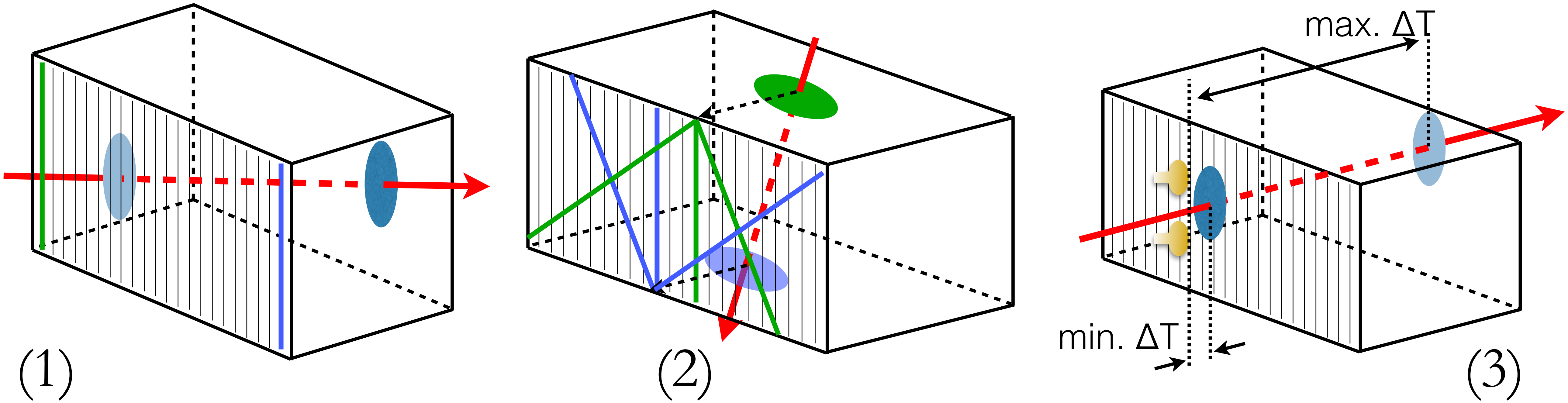


Precuts applied to reconstructed flash in a short beam spill window(93.75ns) to reject:

1. Random single PE noise.
2. Single PMT noise.
3. Flash from Michel electron prior to beam spill window.

Achieved:
Neutrino efficiency: > 96%
Backgrounds rejection: >75%

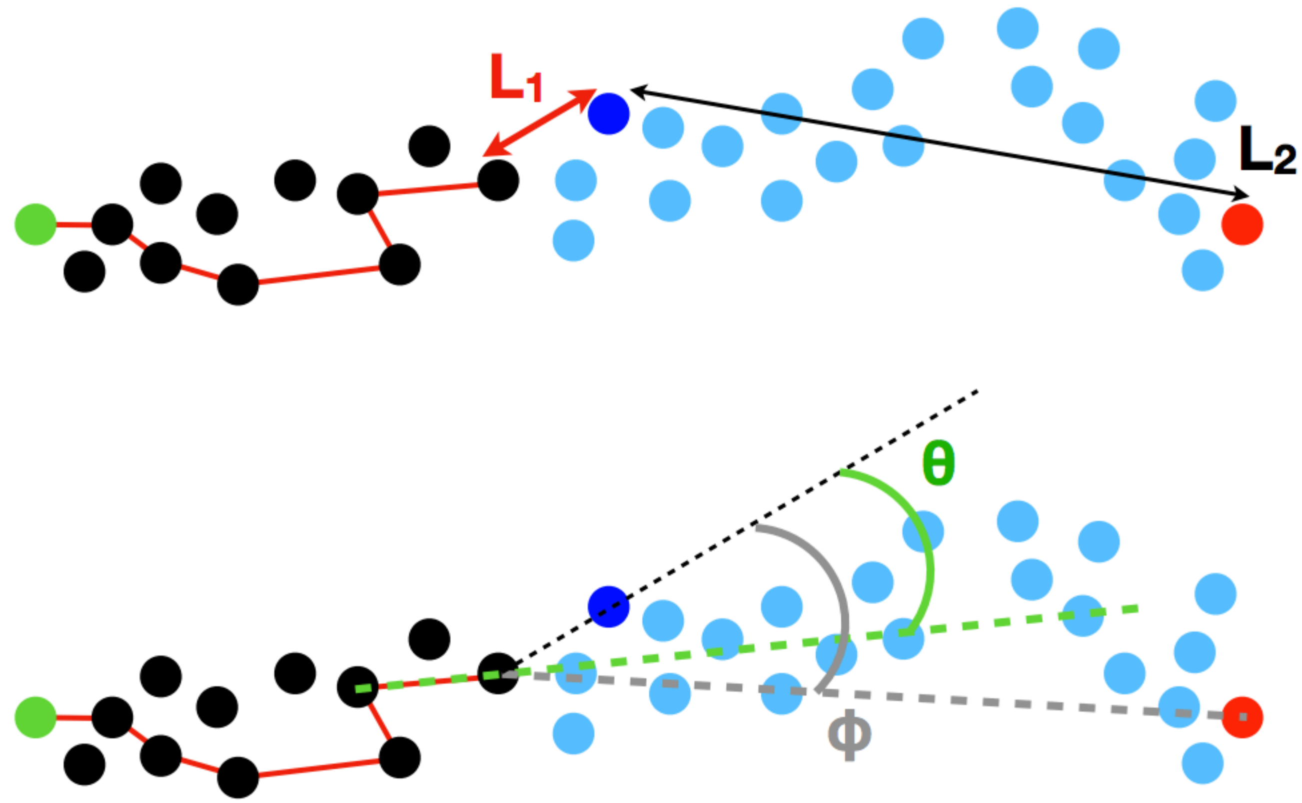
Cosmic Tagging



Dedicated cosmic tagging tool developed to tag cosmic objects and find ROI.

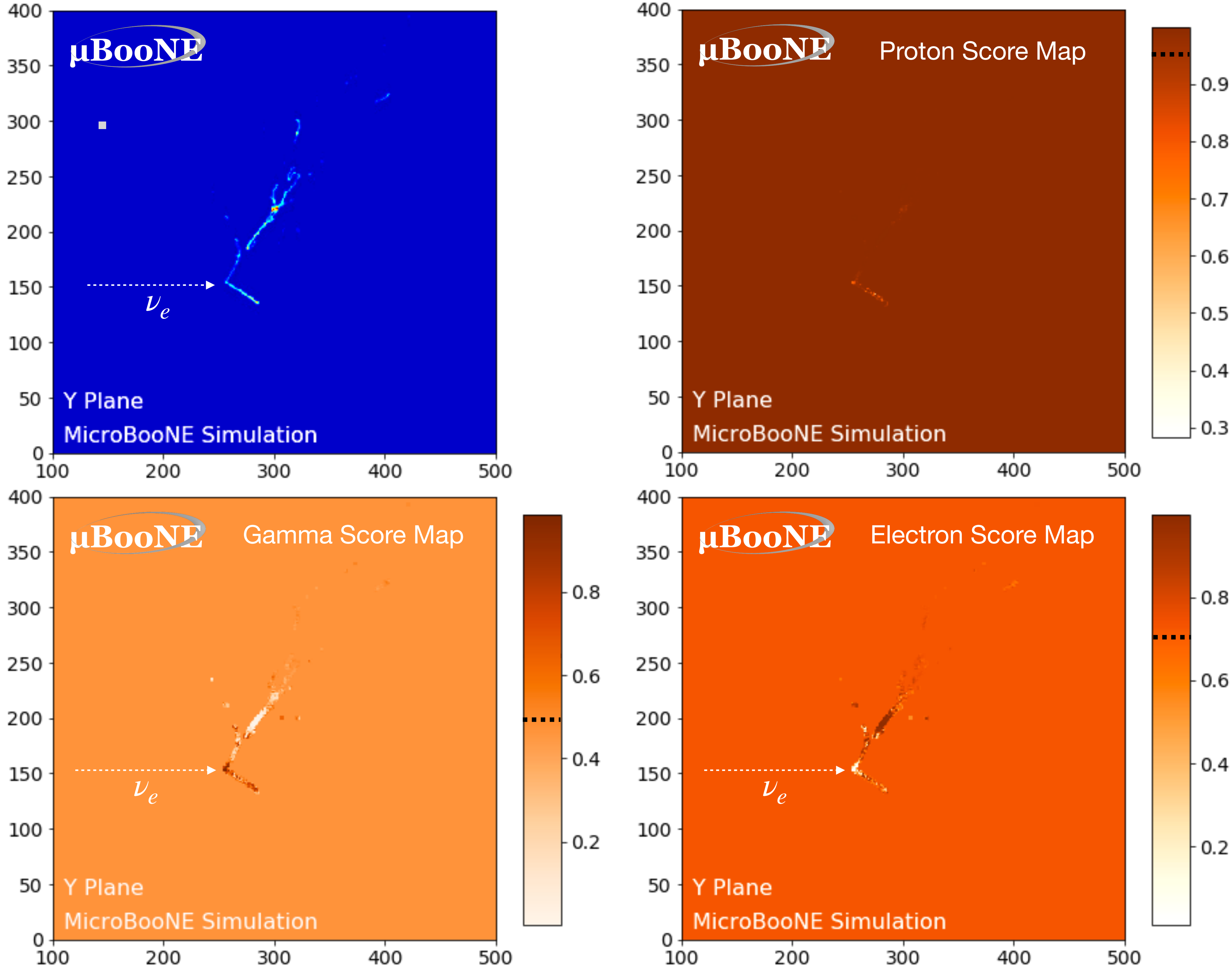
1. Cosmics down beam direction cross the front and end wires
2. Cosmics across top-bottom have unique triplet of wires of three planes.
3. Cosmics traversing TPC have the max Δt between the first and last charges on the track.

3D Track Reconstruction

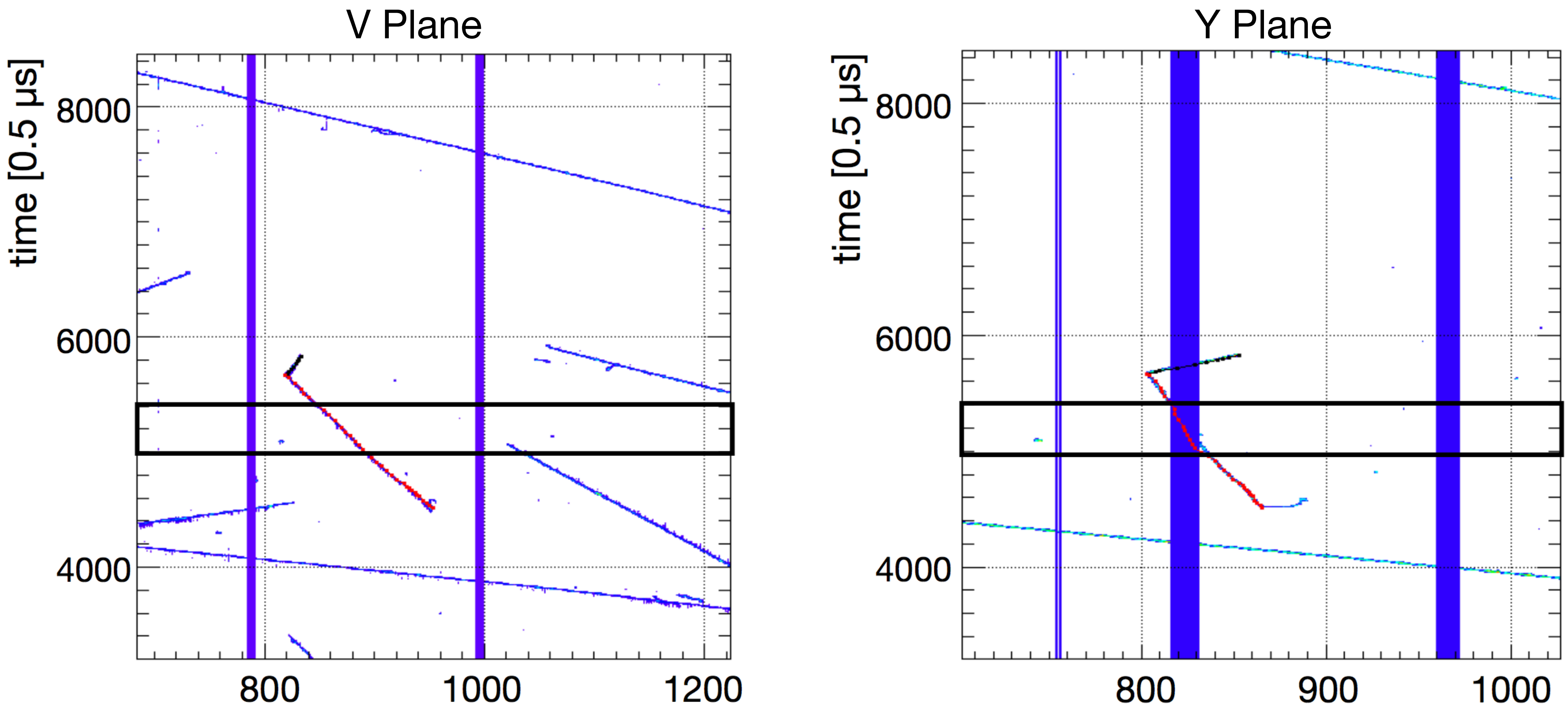


- This reconstruction uses computer vision and clustering tools to find 3D-consistent vertices, and a 3D stochastic best neighbor search.

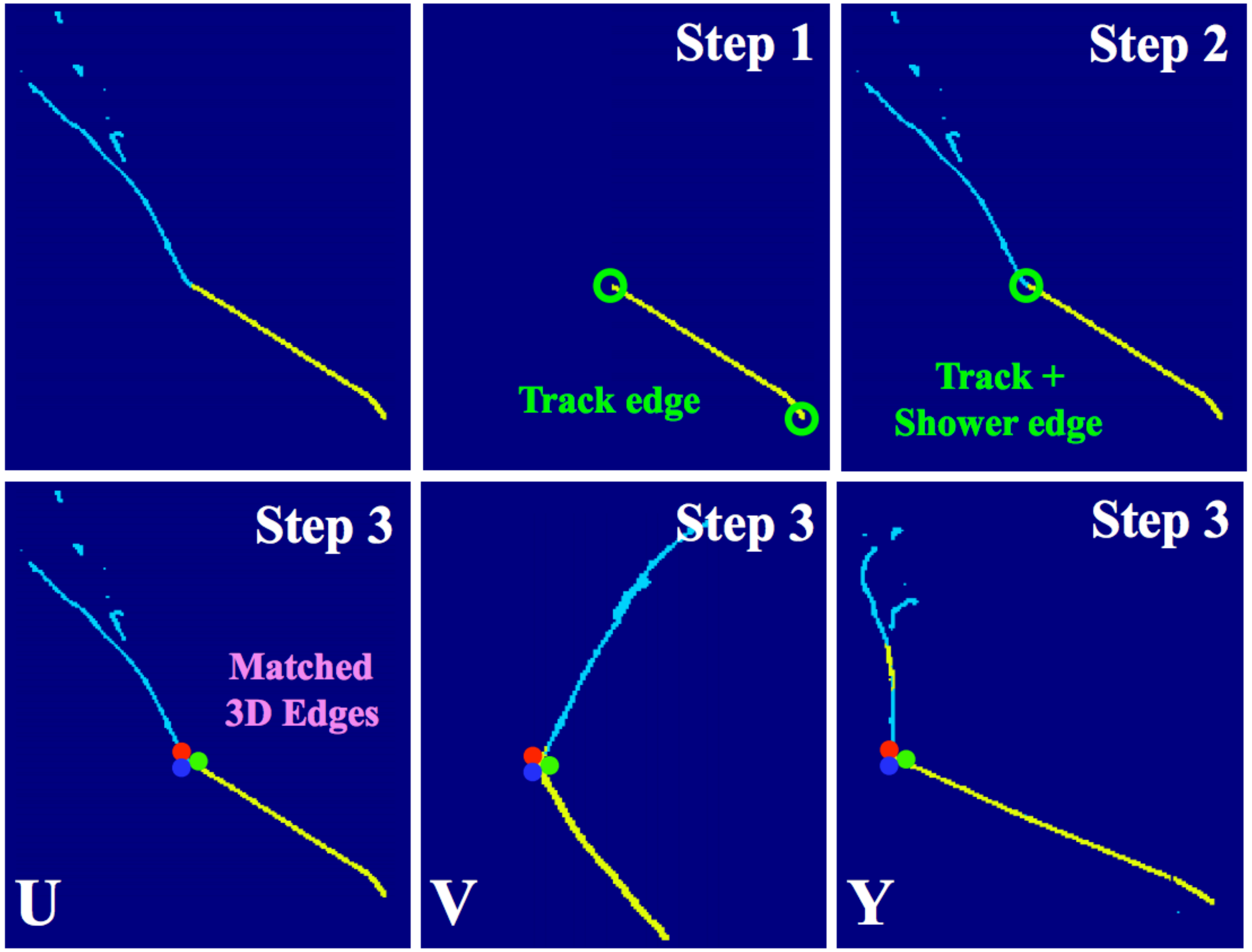
Occlusion Analysis on MPID



3D Track Reconstruction

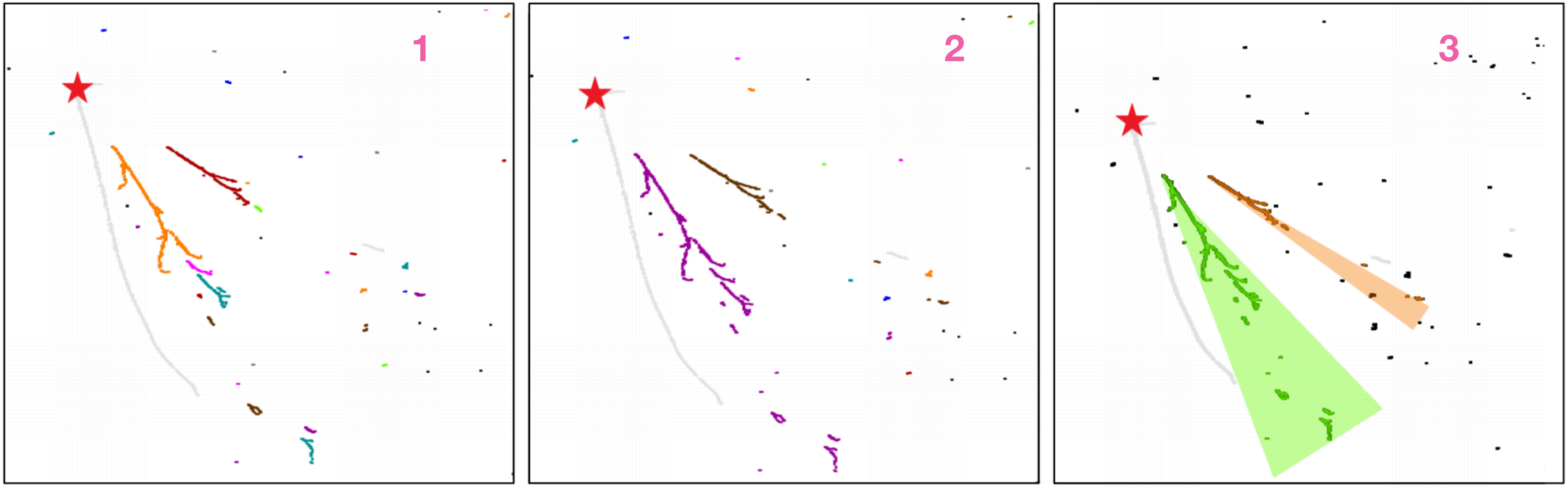


1e1P Vertex Finding



- Step 1: Find edge points on track,
- Step 2: Keep edge point crossing shower pixels,
- Step 3: Match edge points across planes

Shower Reconstruction

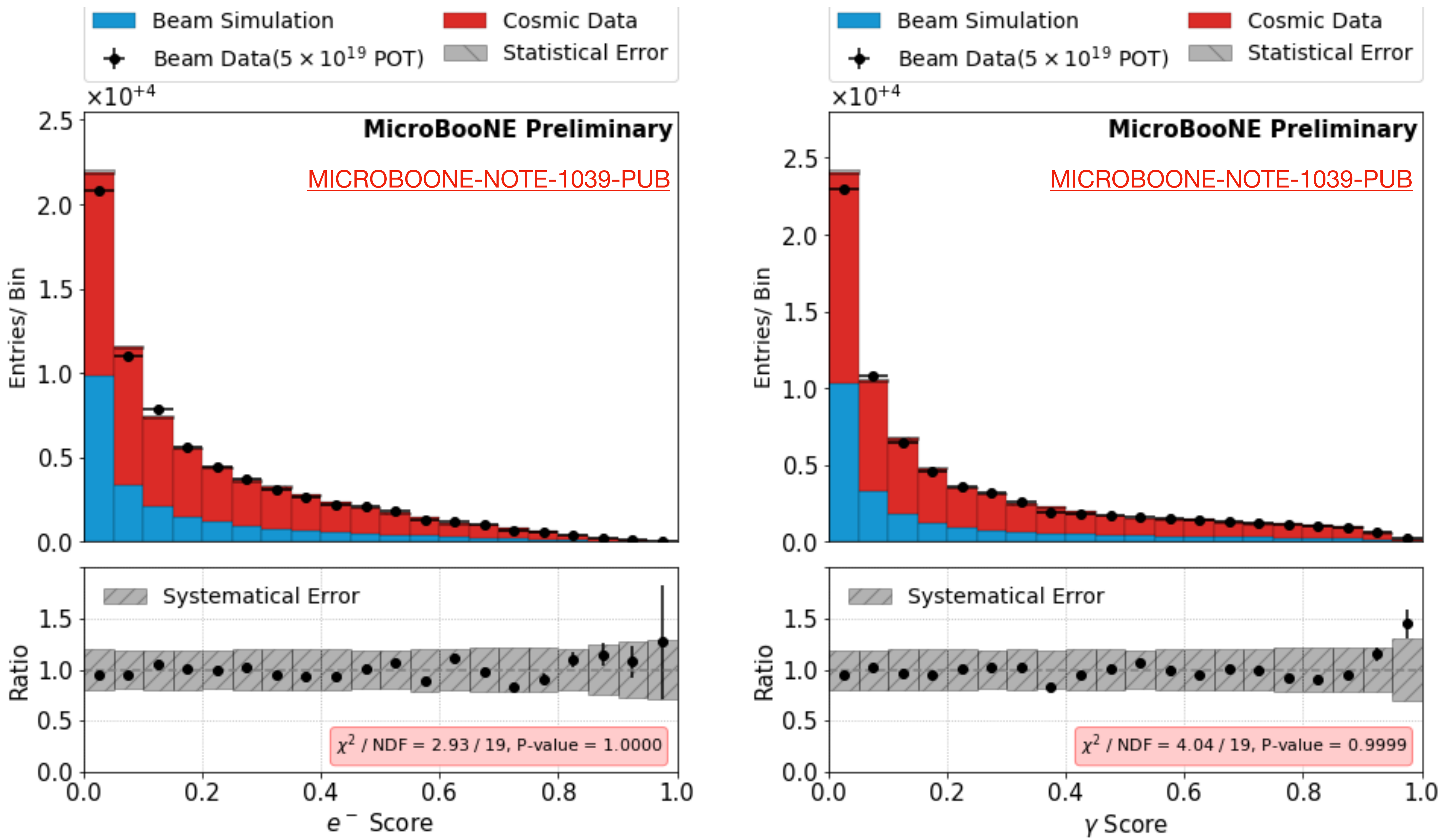


1. Mask out track pixels(in gray). Cluster shower pixels by proximity.
2. Merge showers based on pixel location(length, angle) in a polar coordinate.
3. Reconstructed showers in fit cones.

MPID Training Sample

- Training sample generated with customized event generator,
  - ⋆ Create “3D interaction vertex”,
  - ⋆ One vertex per event, random, uniformly distributed in TPC,
  - ⋆ Random particle multiplicity [1, 4] particles per event,
  - ⋆ Random, isotropic particle momentum directions,
  - ⋆ Random particle types from, P, e^-, γ, μ^- and $\pi^{\pm}$,
  - ⋆ Two particle types mixtures,
    - ✳ 80% events with kinetic energy in [100,1000]MeV and proton in [100,400]MeV,
    - ✳ 20% events with kinetic energy in [30,100]MeV and proton in [40,100]MeV.
- ~45,000 events for training and ~40,000 for validation.
- Why not train with overlay image for training, (1) Cosmic muon would make the network fail for detecting neutrino induced muon (2) Michel & deltas would make the MPID fail for detecting neutrino-induced electron.

MPID Data MC Comparison



MPID Data MC Comparison