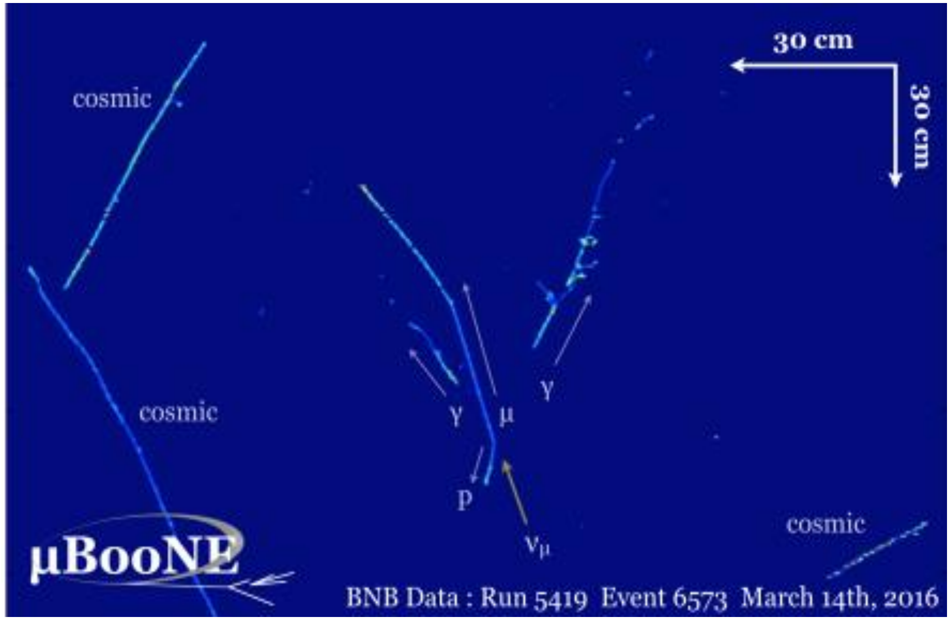
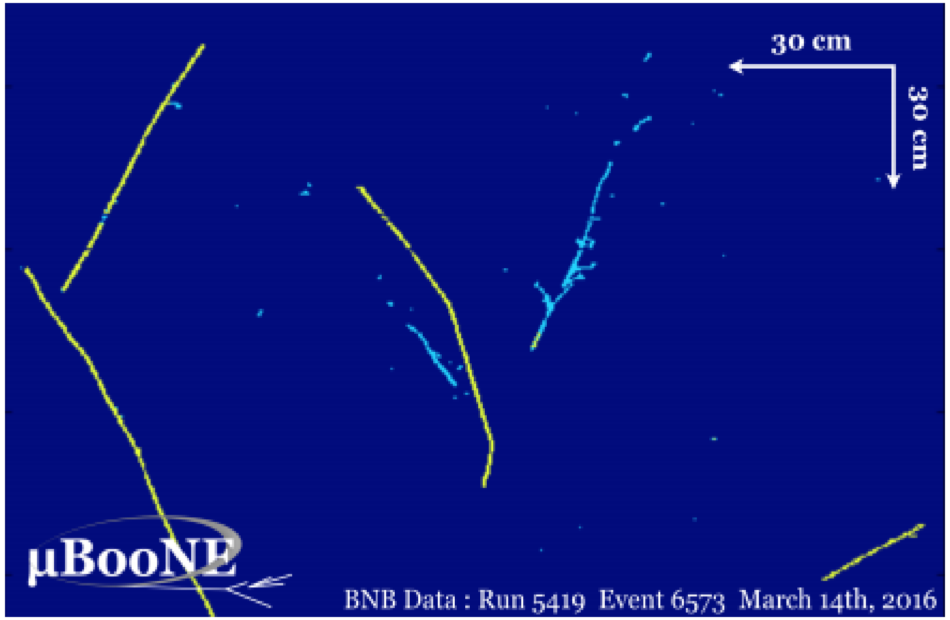
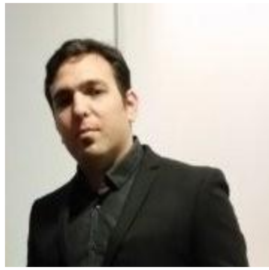
APPLYING SPARSE CNN IN MICROBOONE





Ran Itay

On behalf of MicroBooNE



μBooNE

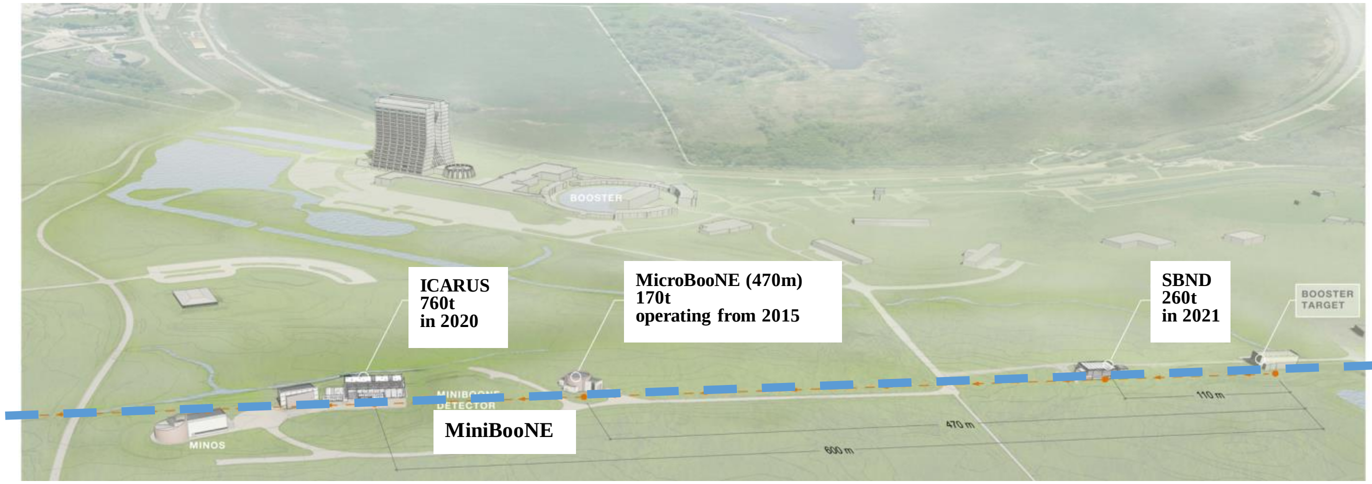
SLAC NATIONAL ACCELERATOR LABORATORY

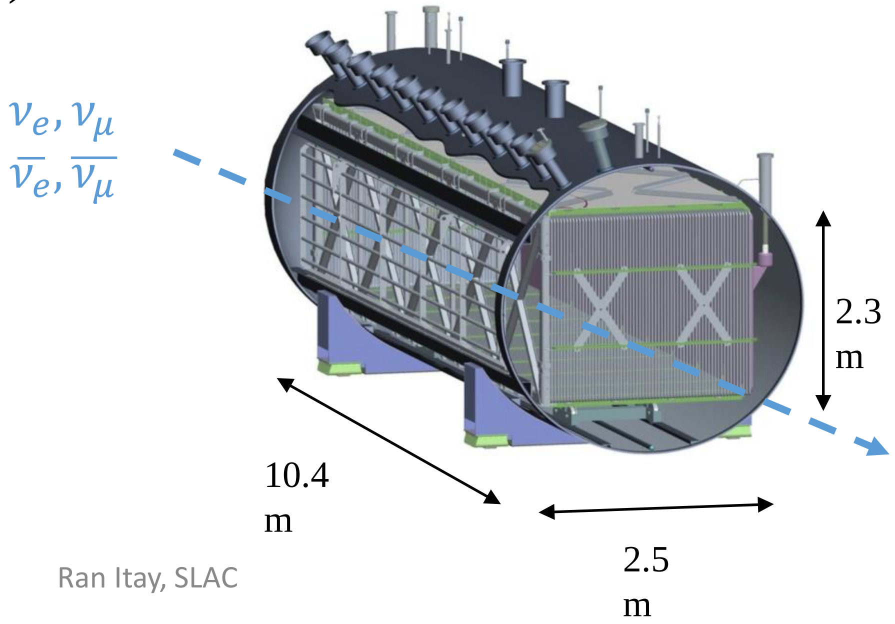
CTD

April, 2020

MicroBooNE

- **Micro Boo**ster **N**eutrino **E**xperiment
  - 85 ton Liquid Argon (LAr) TPC (active mass)
  - Located at Boosted Neutrino Beam (BNB), Fermilab
  - O(1) GeV ν_μ beam
  - Appearance experiment ($\nu_\mu \rightarrow \nu_e$)
    - Lepton flavor distinguishing (track/shower)
  - Operating since 2015
  - Surface detector (~5 kHz cosmics)

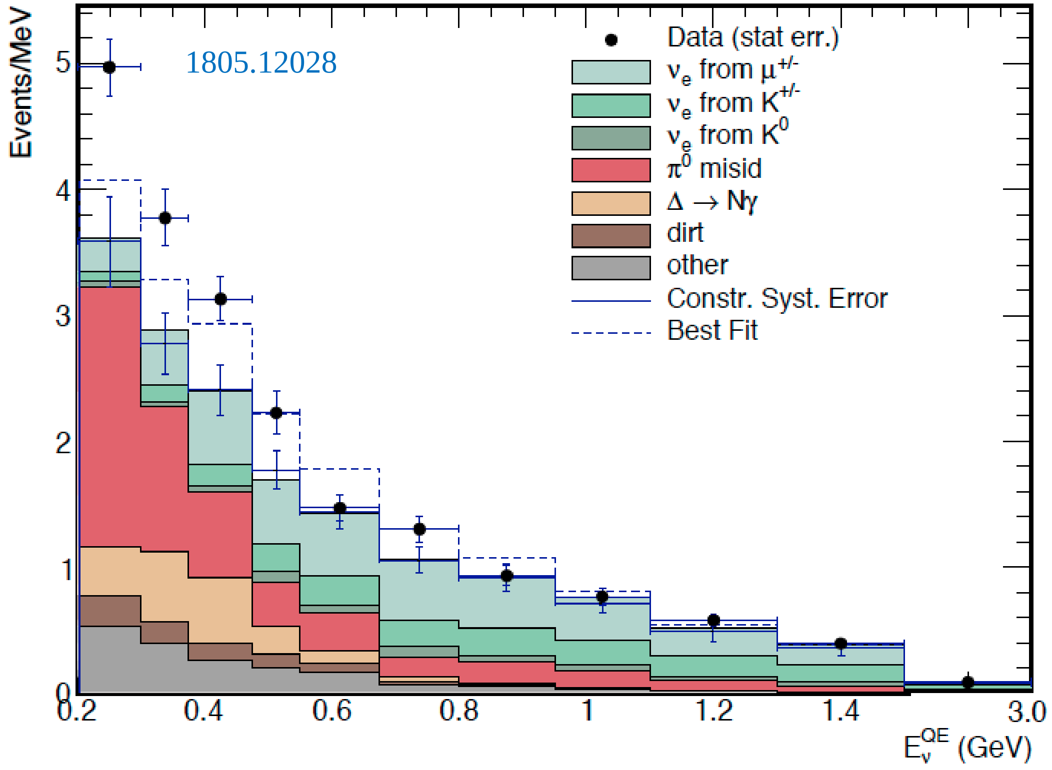




Introduction

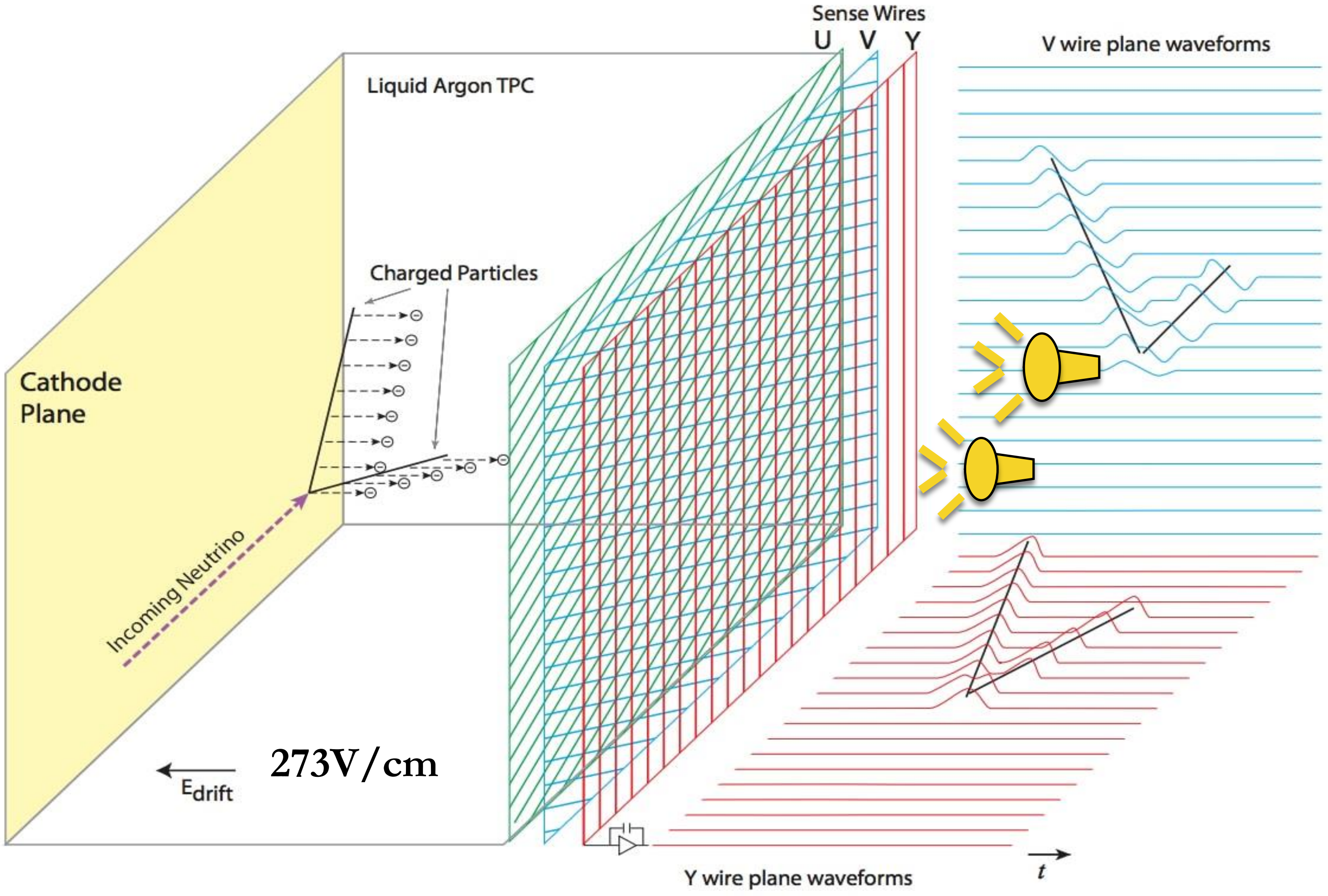
Motivation

- MiniBooNE sees 4.5σ ν_e-like Low Energy Excess (LEE)
- New physics to explain this.
- limited capability to distinguish e vs γ
- MicroBooNE
  - Same beam
  - Similar baseline
  - LArTPC
- Goals
  - LEE search
  - Cross-section measurements
  - Detector physics (DUNE)



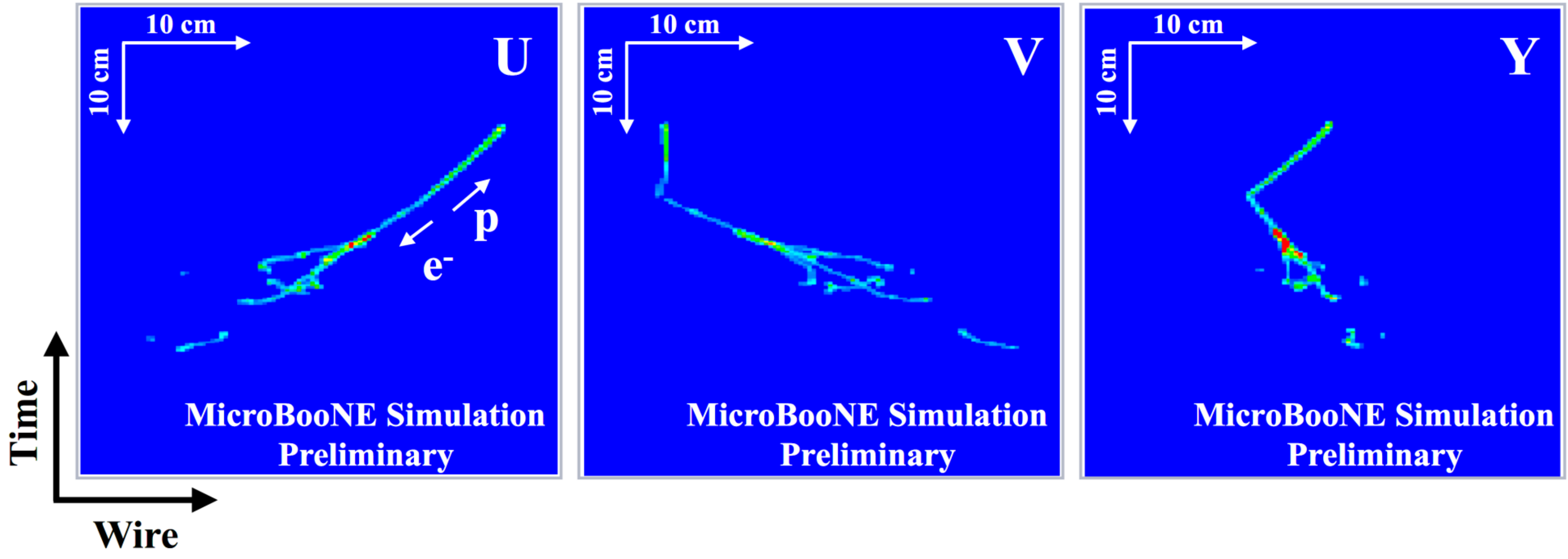
MicroBooNE

- Three wire planes
  - 2 induction planes(2,400 wires each).
  - 1 collection plane(3,456 wires).
  - 3mm wire pitch.

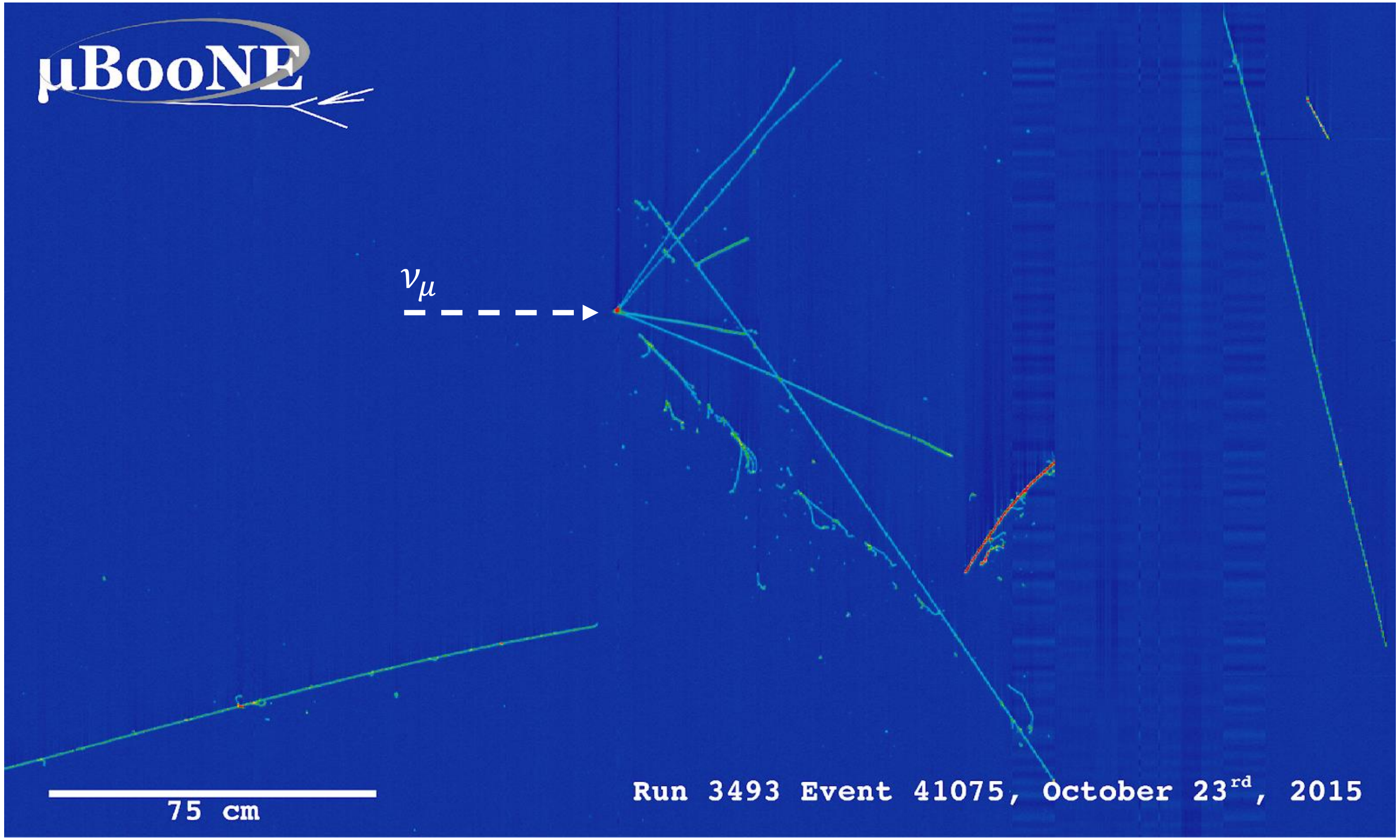


MicroBooNE

- Three wire planes
  - 2 induction planes(2,400 wires each).
  - 1 collection plane(3,456 wires).
  - 3mm wire pitch.

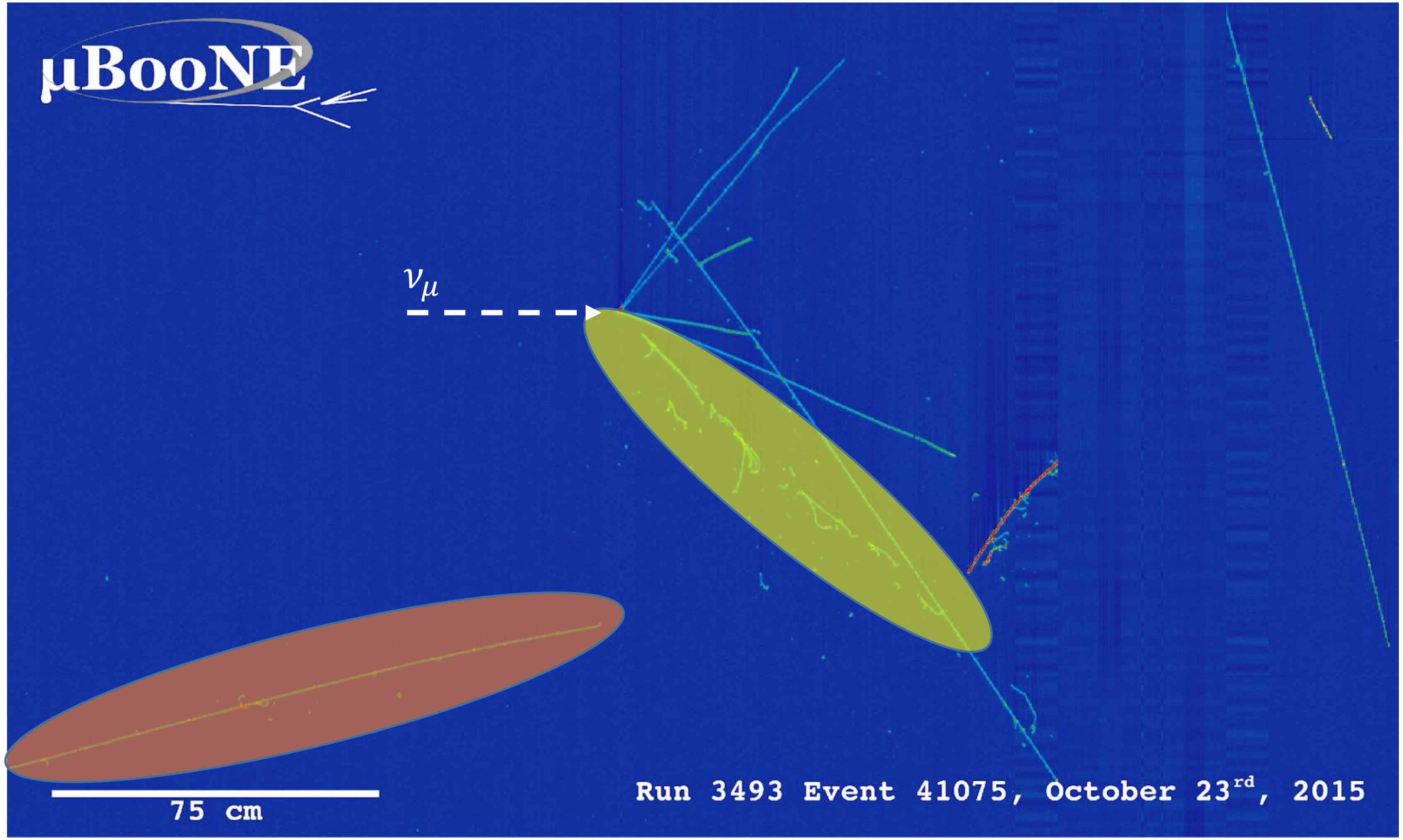


MicroBooNE



MicroBooNE

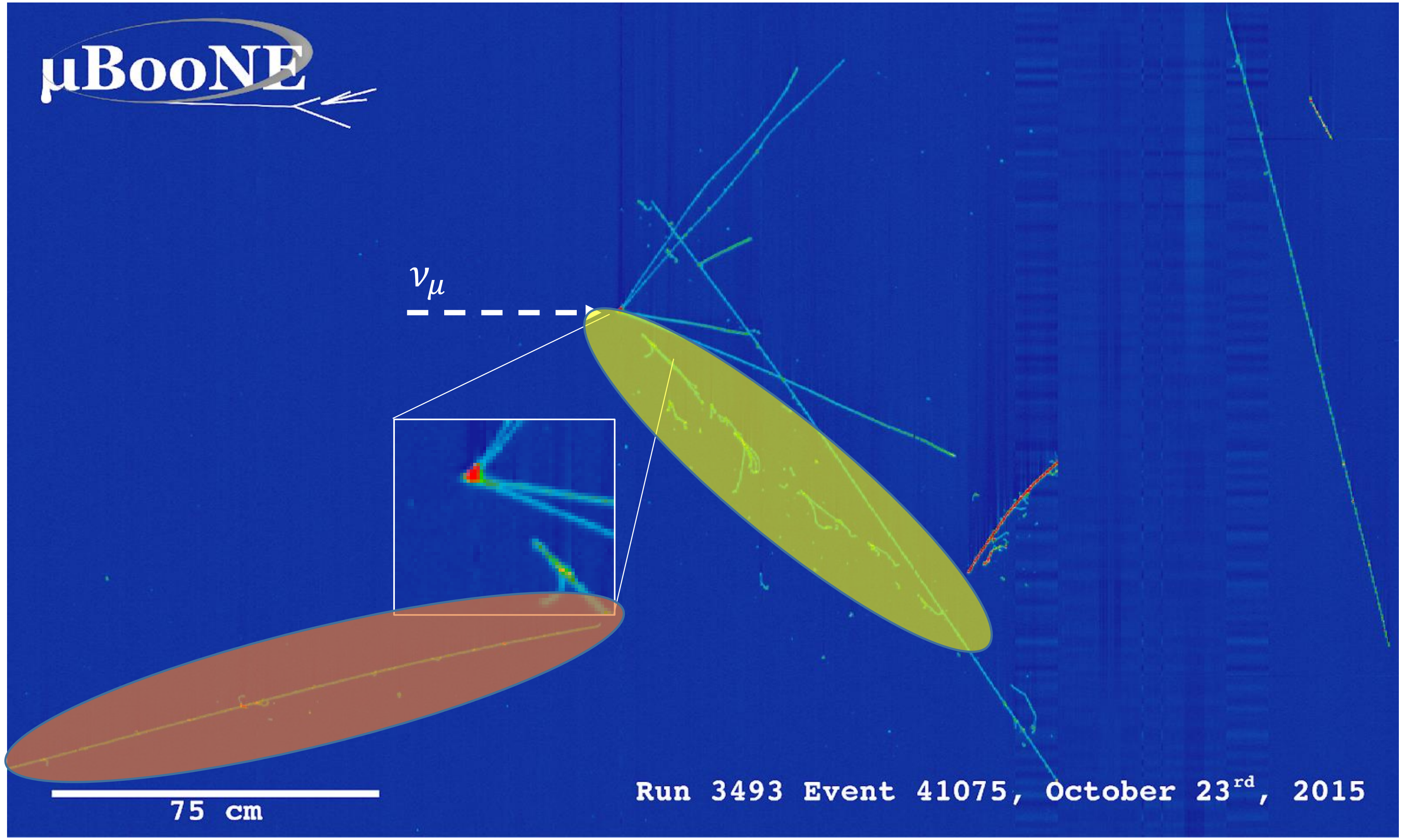
- **Shower Vs Track**
  distinct topologies



MicroBooNE

- **Shower Vs Track**
  distinct topologies

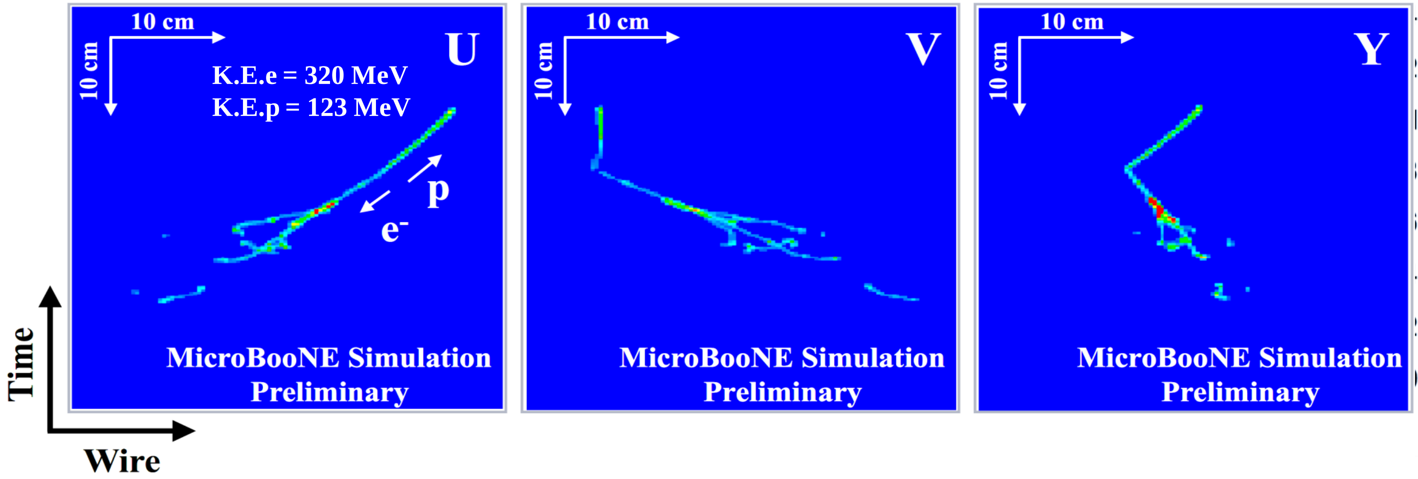
- γ, π^0 Vs e
  Gap from vertex

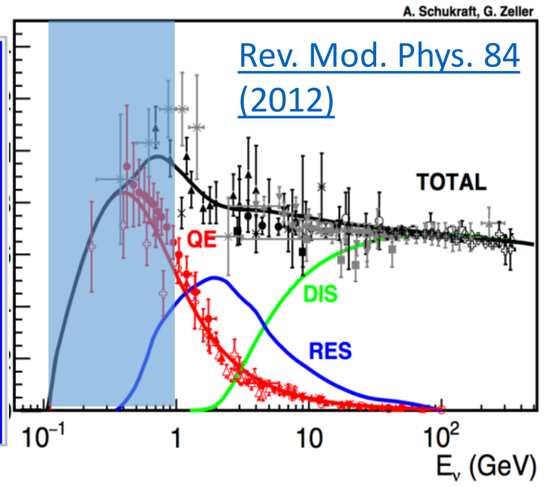


MicroBooNE

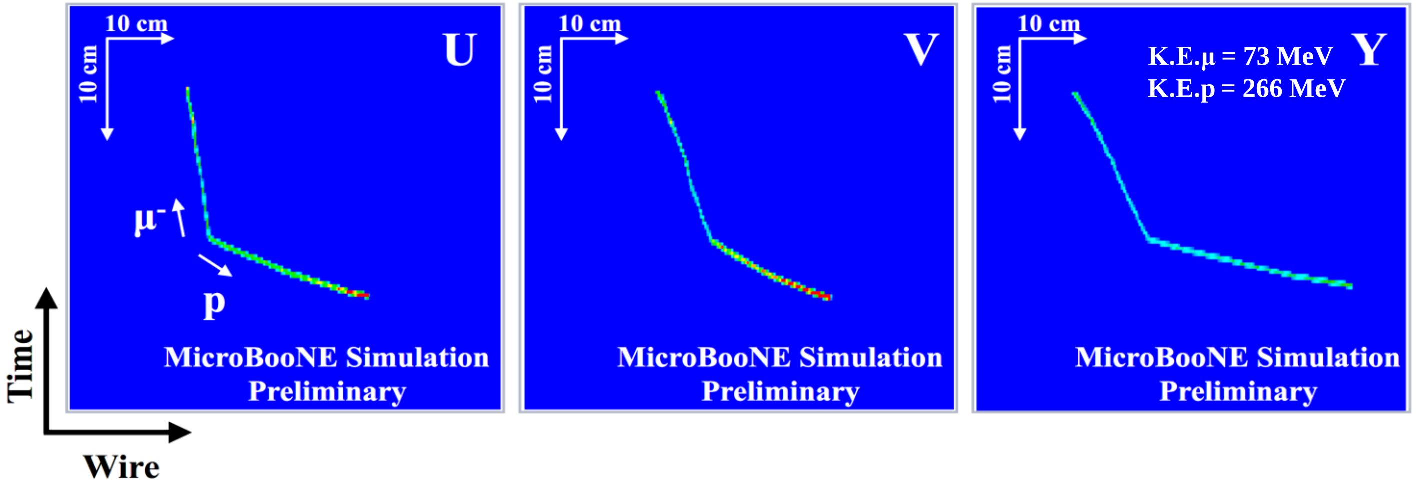
- CCQE is dominant in ROI
- Looking for 1l-1p-0π topologies:

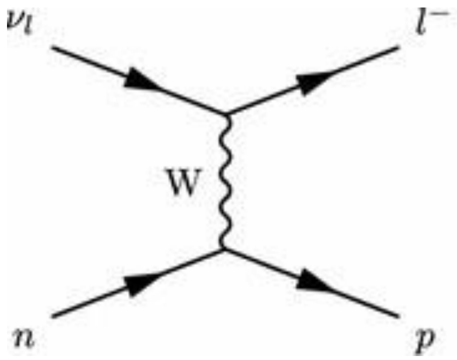
1e1p
Signal





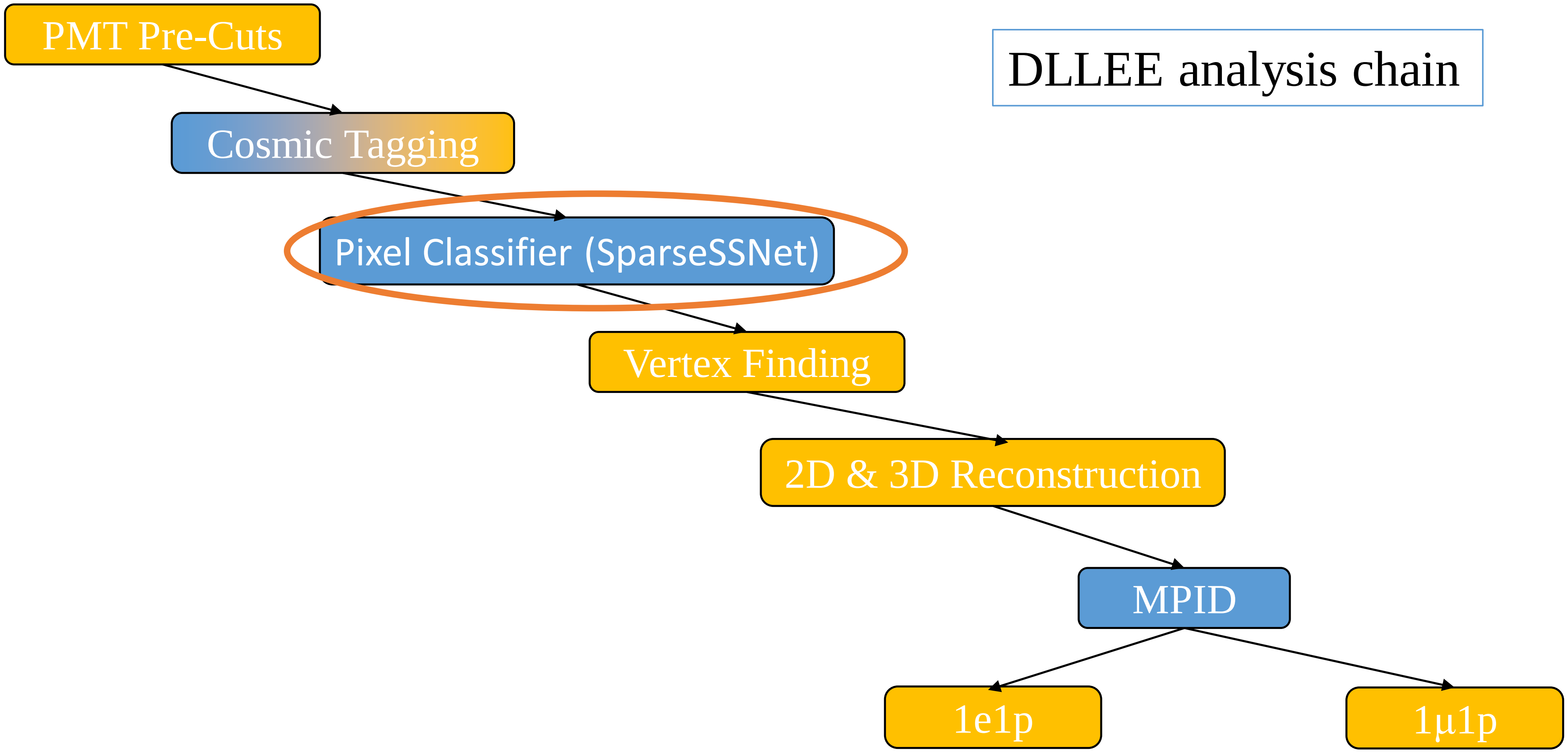
1µ1p
For constraining:
Flux
Systematics





MicroBooNE

DLLEE analysis chain



Sparse Semantic Segmentation

- Semantic Segmentation (pixel level)

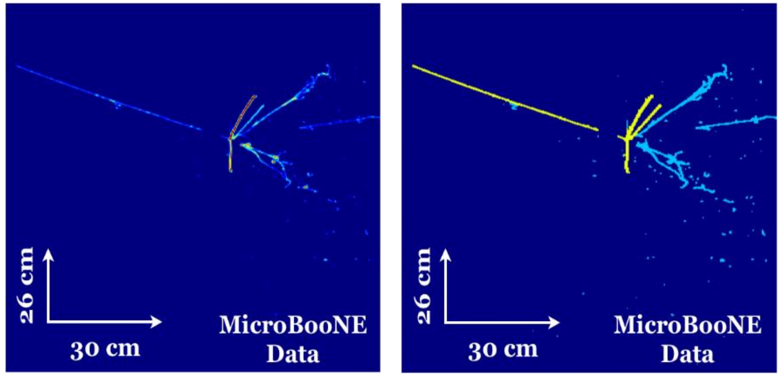
- Works in MicroBooNE, using CNN for track/shower classification using **standard dense algebra** (cropped images)

- In the close vicinity of an event dense data is important.

- Occupancy, < 0.5% pixels within threshold.

Input Image Truth Label Network Prediction

Input Image Network Prediction

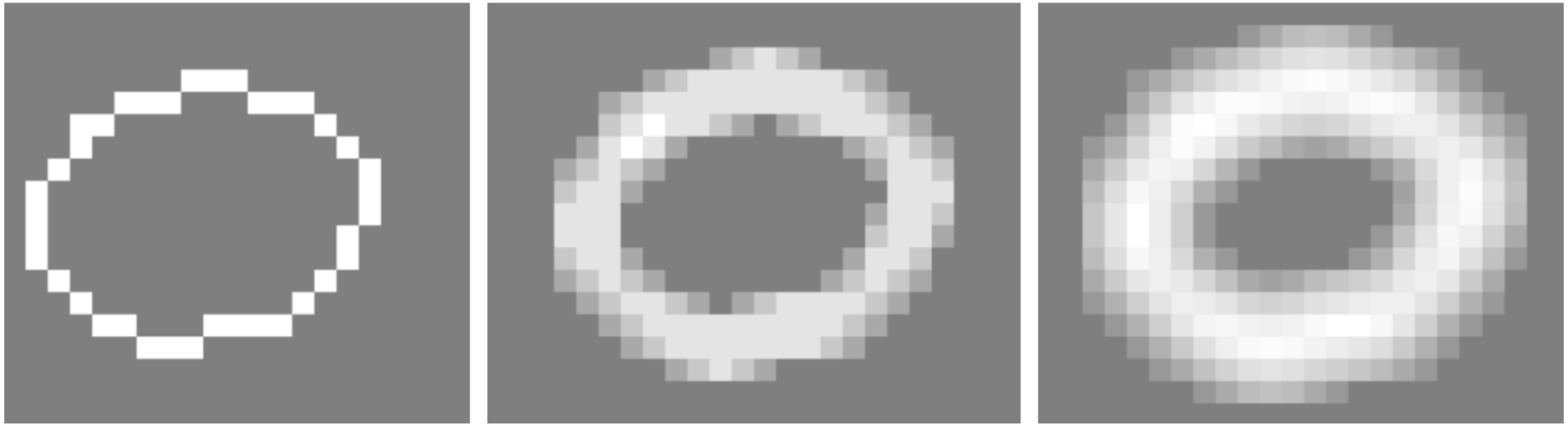


1808.07269

Sparse Semantic Segmentation

- Performing dense matrix multiplication -> **Not optimal**
  - Many multiplication of non important pixels -> **Takes more time**
  - Saving non important pixels -> **Takes more memory**

- Convolutional layers create “blurring” of sparse data, which decreases their spatial precision

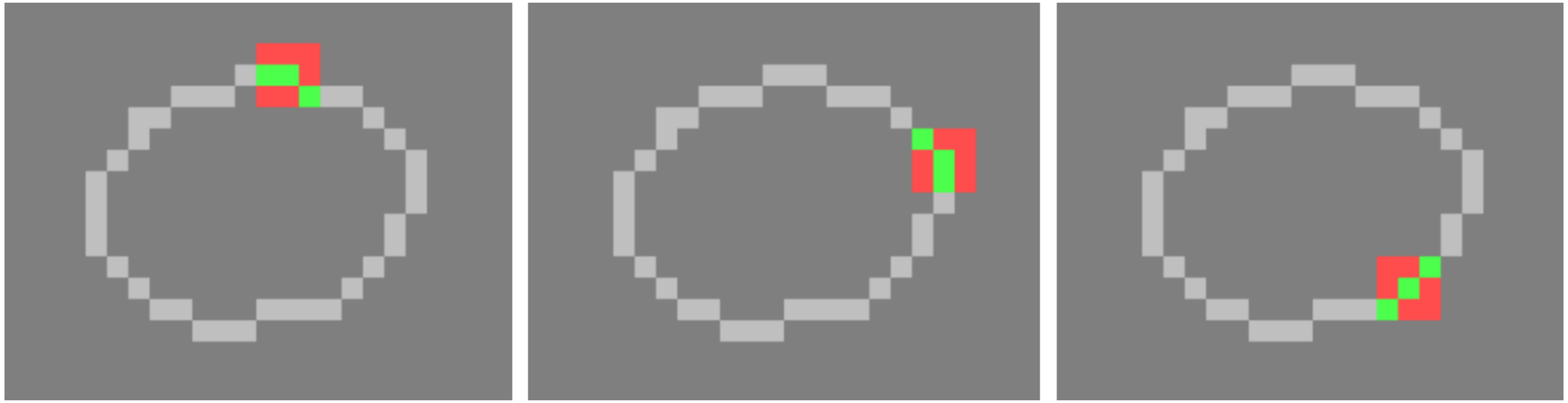


Sparse Semantic Segmentation

- Developed by Facebook AI Research / Oxford - CVPR2018, best 3D semantic segmentation record for ShapeNet (open 3D point-cloud dataset)
- **LAr TPC data sparsity < 1% (0.5% within threshold)**
- Optimized for 3D (see Francois Drielsma's talk, "Particle Clustering and Flow Reconstruction for Particle Imaging Neutrino Detectors Using Graph Neural Networks").
- Submanifold – Important data has an effective lower dimensionality (e.g., track – 1D)

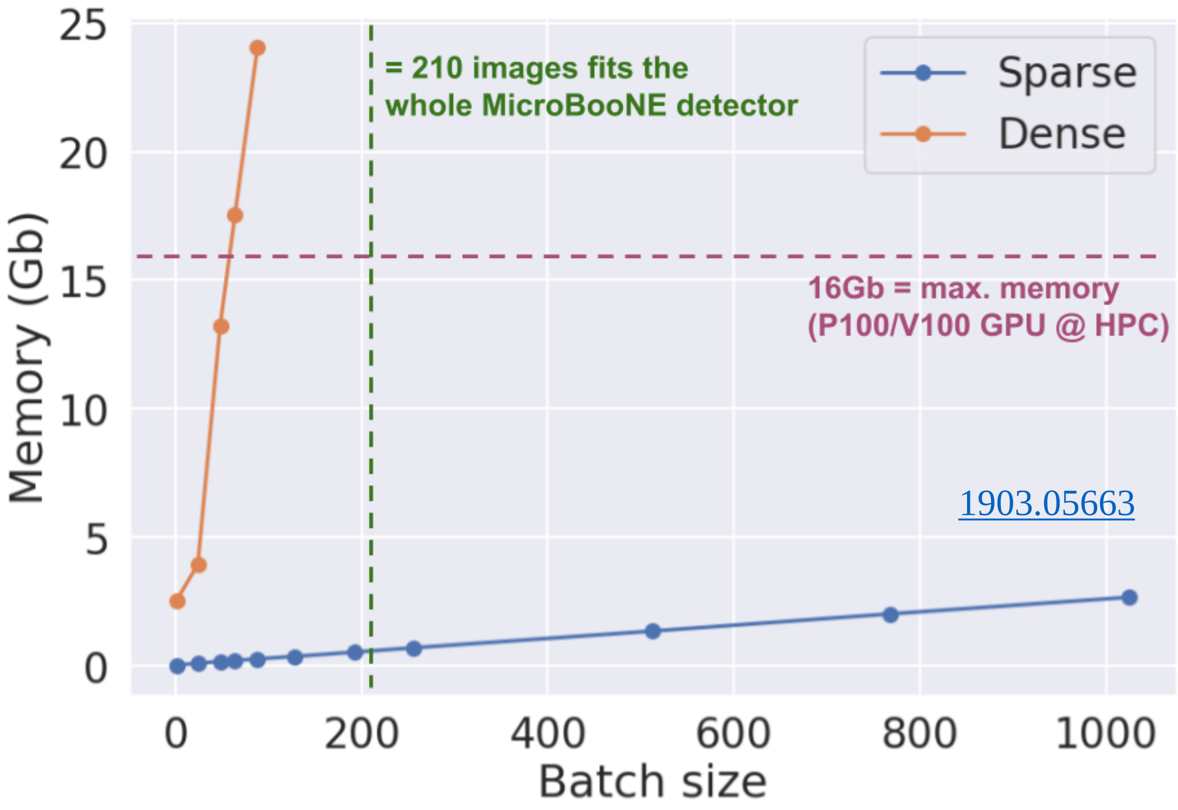
Sparse Semantic Segmentation

- Prior training, define a “mask” to select only important pixels (e.g., pixel value > 0).
  (val) X H X W -> (R , C , Val) X N_mask
- Consider any non-important pixel as 0 and do not perform its multiplications -> **Time and space saving**
- In hidden layers only pixels which were activated in input can be activated -> **prevent “blurring”**



Sparse Semantic Segmentation

- Work done on public available LAr TPC data set shows a big improvement in:
  - Accuracy
  - Computational time
  - Memory usage

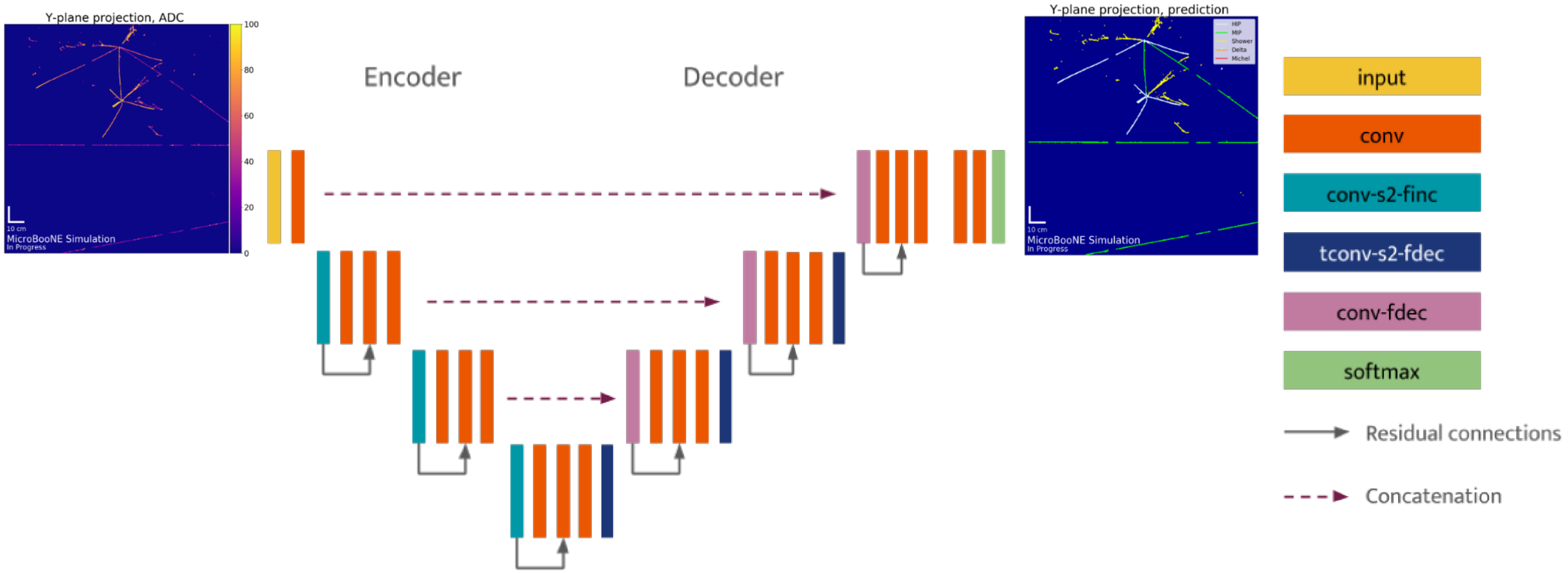


Data Sample

- Data sample – particles generated and propagated including detector response (e.g., dead wires, recombination, diffusion, etc.)
  - "Particle bomb" (e, γ , μ , $\pi^{\pm}$, p) - interaction
  - Muons – cosmic rays
- Single hot labels – HIP (protons), MIP (μ , $\pi^{\pm}$), shower (e, γ), delta (knock-on electron), Michel electrons (decay of muons)
- Label priority – important for downstream analysis (e.g., vertexing) HIP > MIP > Shower > Delta > Michel
- Training sample 120k events; Test sample 23.5k events
- Mask to remove noise while keeping important info – $10 < \text{ADC} < 300$ (~0.5% of pixels in image)
- Distinct weights per plane

Network

- UResNet – a hybrid of Unet & ResNet (demonstrated with standard CNN dense algebra 1611.05531)
- Depth - 5 layers
- 32 initial filters
- Loss - cross entropy sum

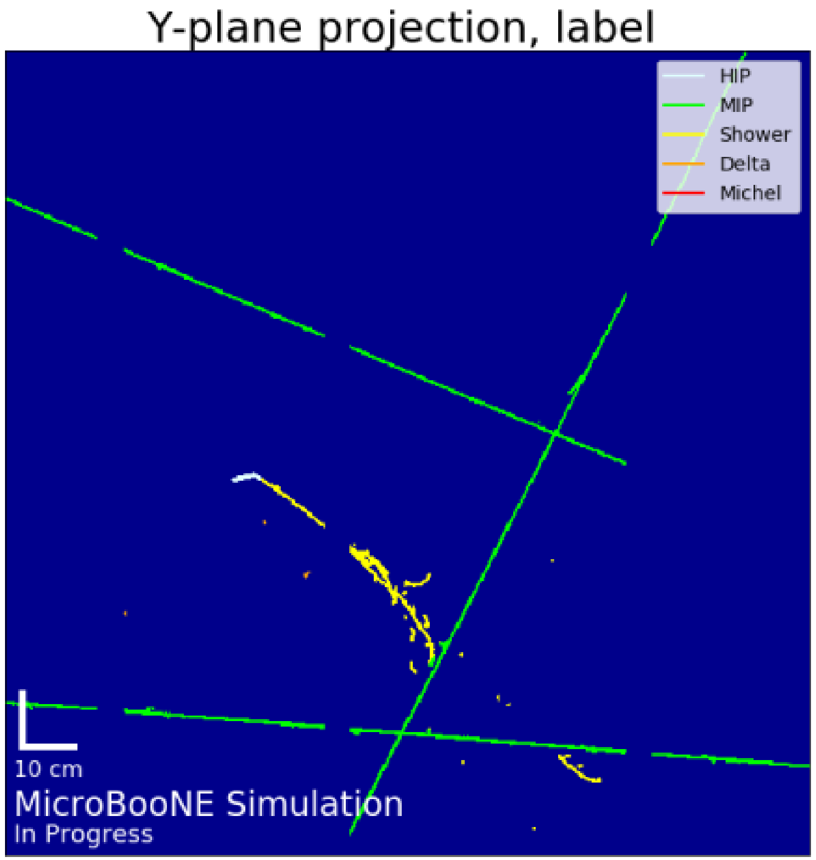


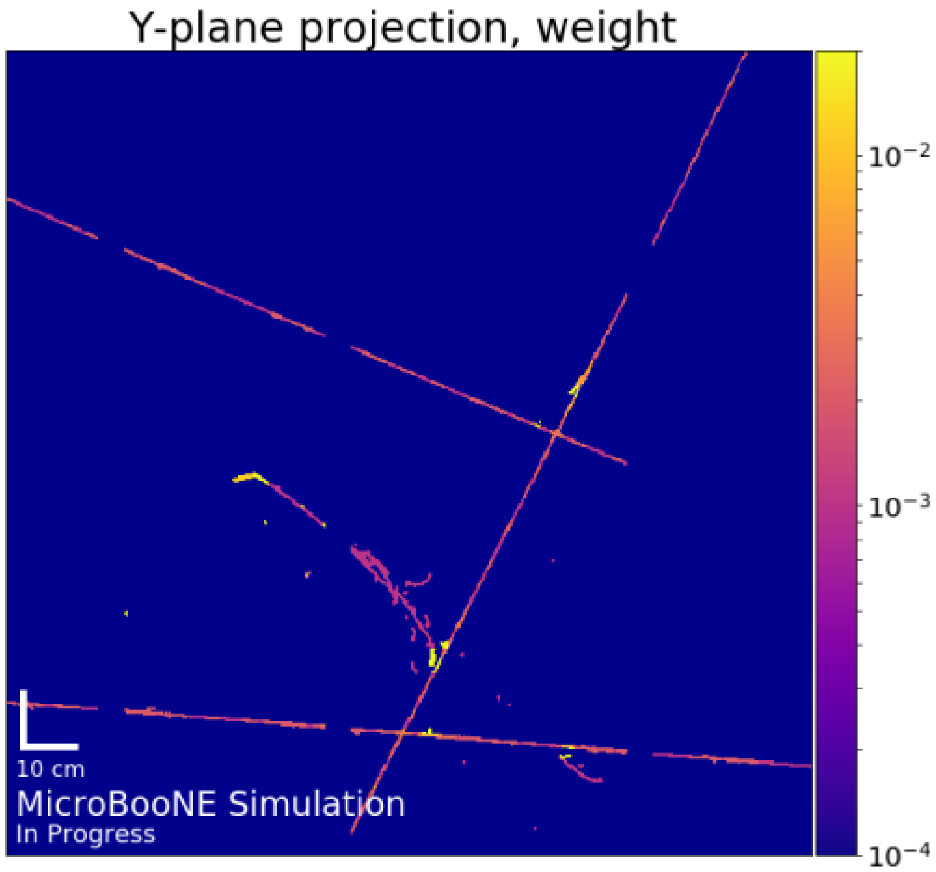
Network

Class Stat imbalance

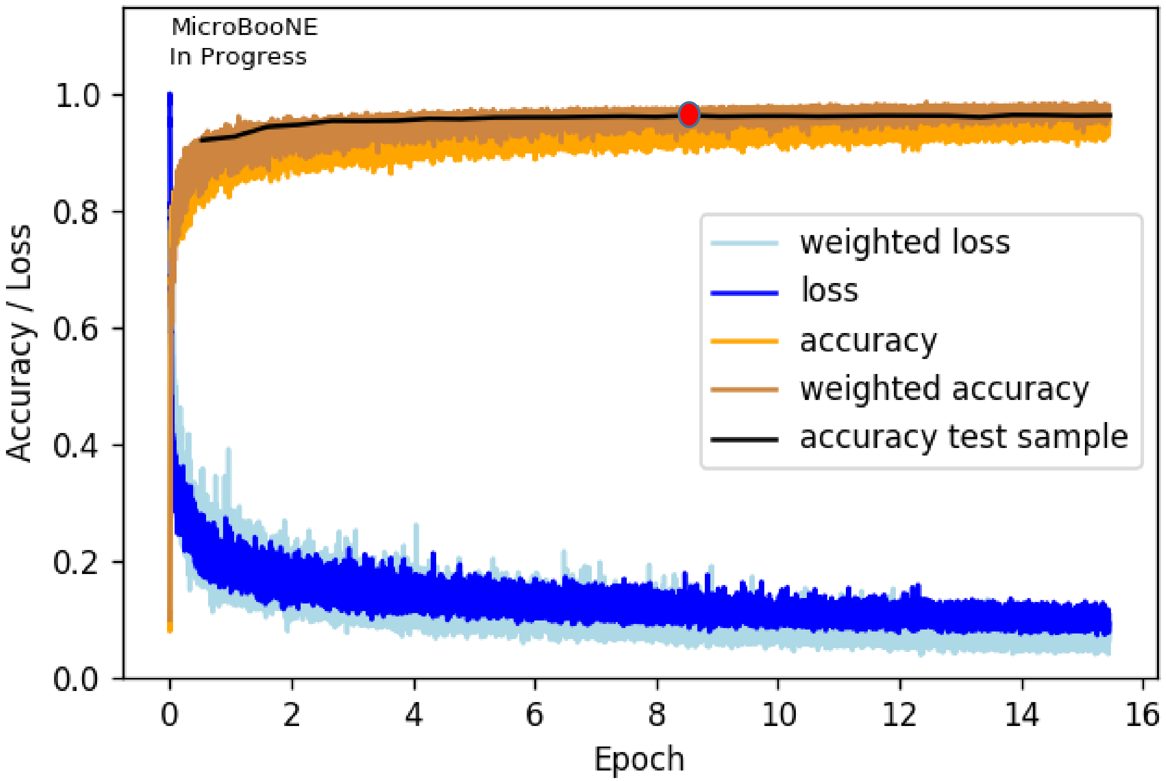
Pixel weighting:

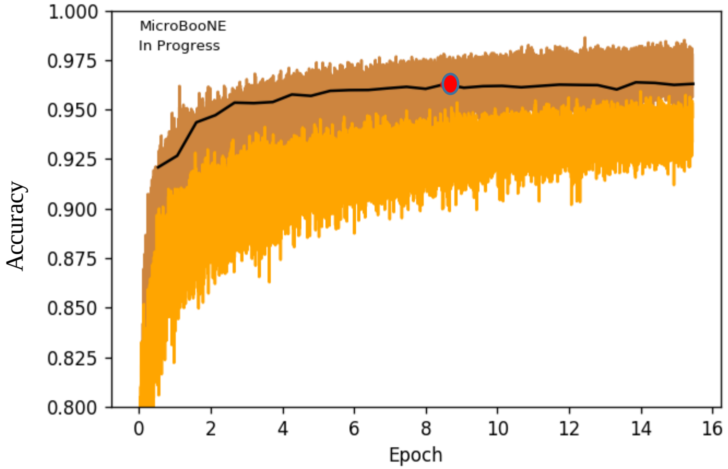
- Cluster weighting - $w_i \propto 1/N$ $(0.02 - 2) \times 10^{-2}$
- Vertex weighting (3 pixel distance) $w_i = 2 \times 10^{-2}$

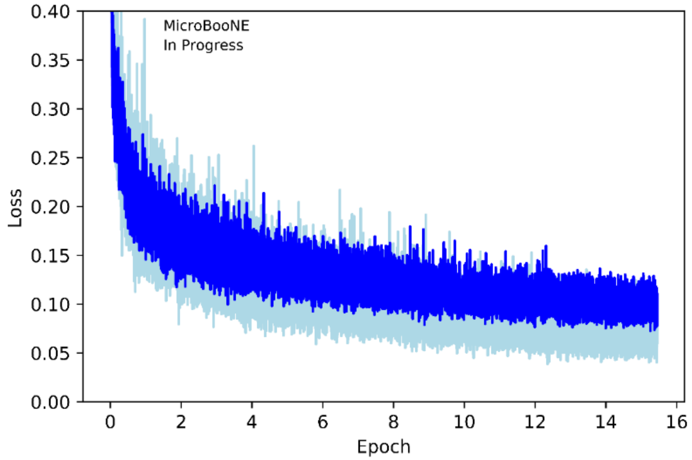




Results





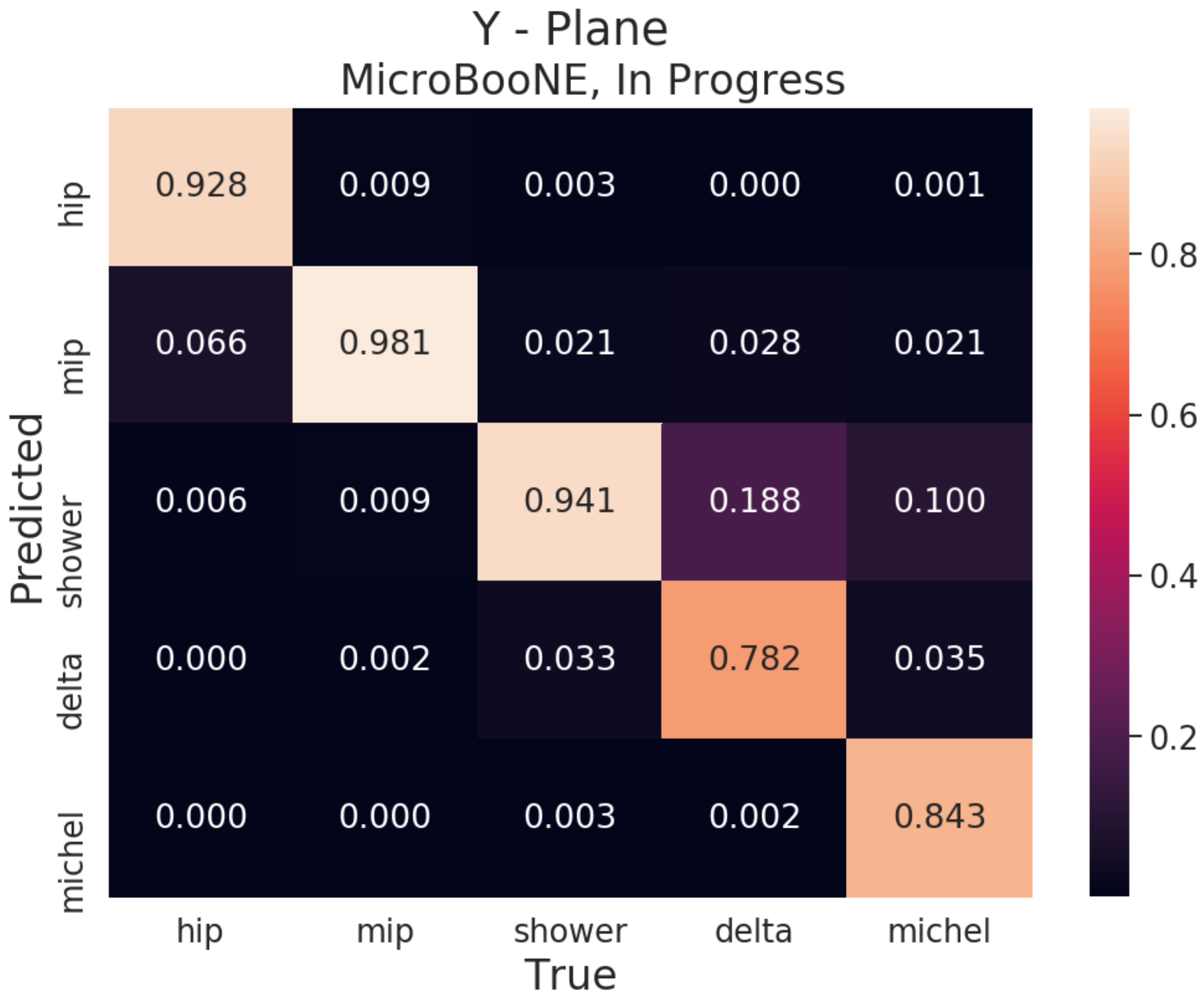


- Total accuracy on training ~ 0.97
- Total accuracy on inference ~ 0.95
- Iter 15,999 = ~epoch 8.5

- **Total training time ~ 8 hours**
- **Average Inference time (CPU) ~ 2s/image**
- **Memory Consumption O(GB/image)**

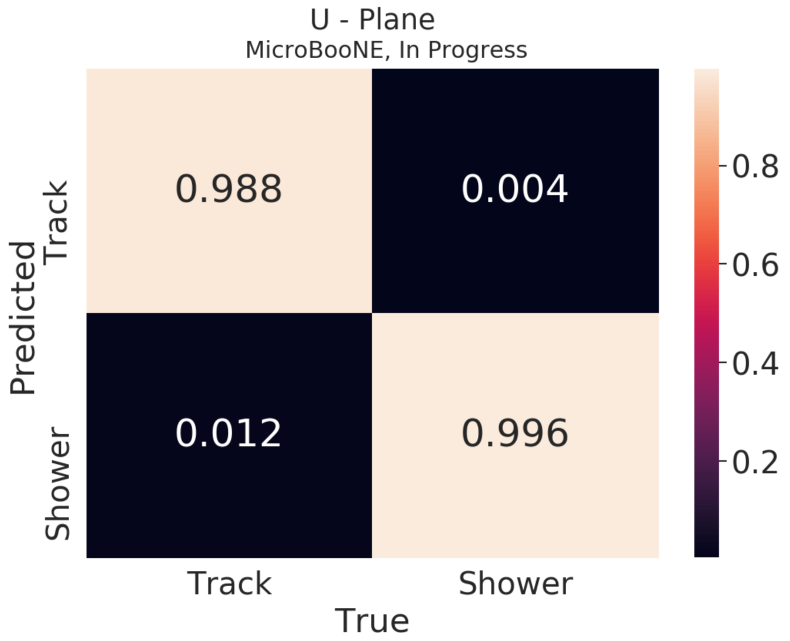
Results

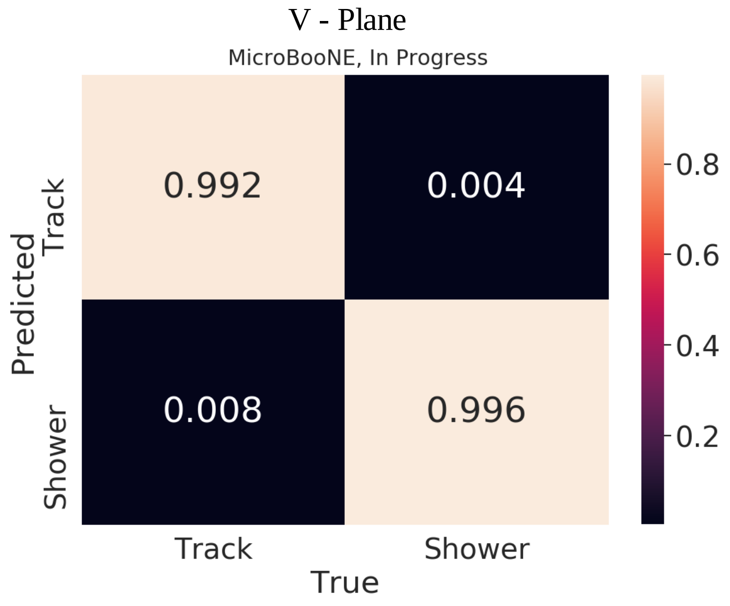
- Normalized to True (accuracy)
- Highest probabilities on diagonal (as expected)
- Most misidentification Delta & Michel
- Other Planes similar (see backup slides)

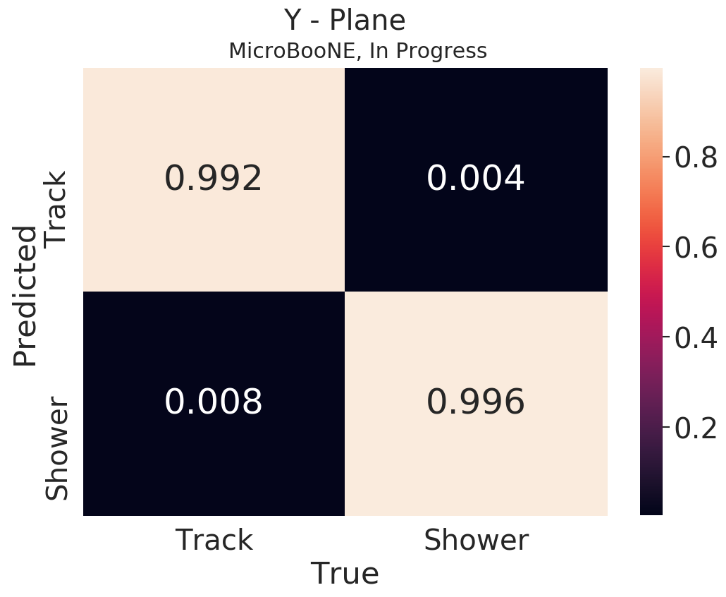


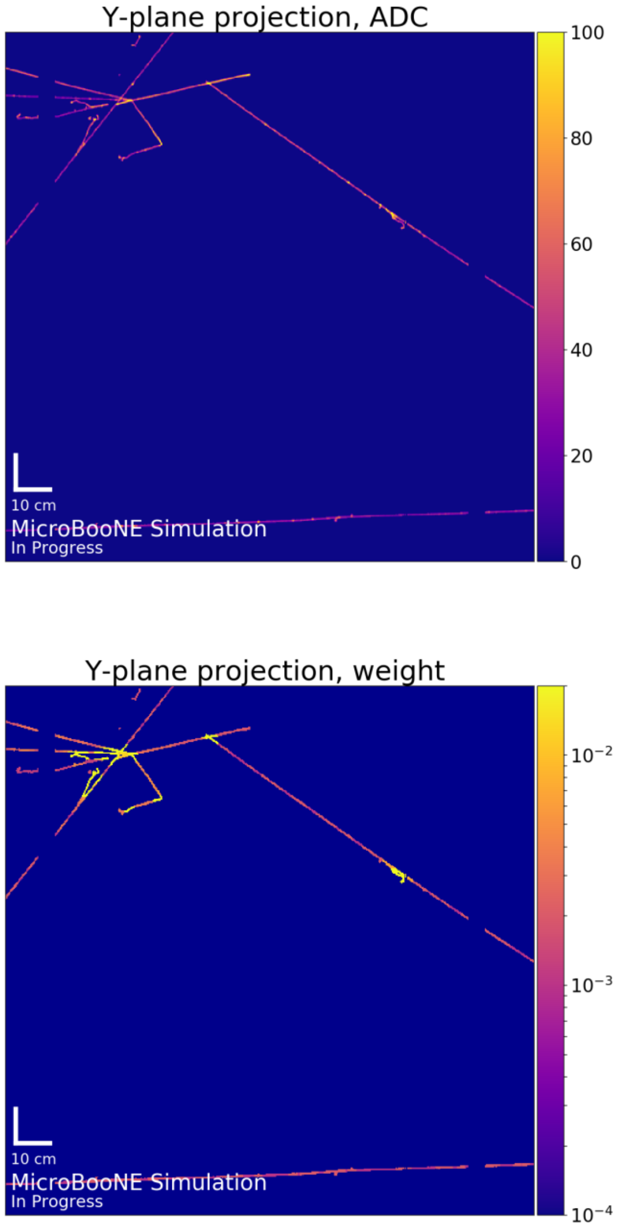
Results

- Comparing only track/shower (current DL LEE analysis)
- Total accuracy = **0.99**
- Total Purity = 0.99

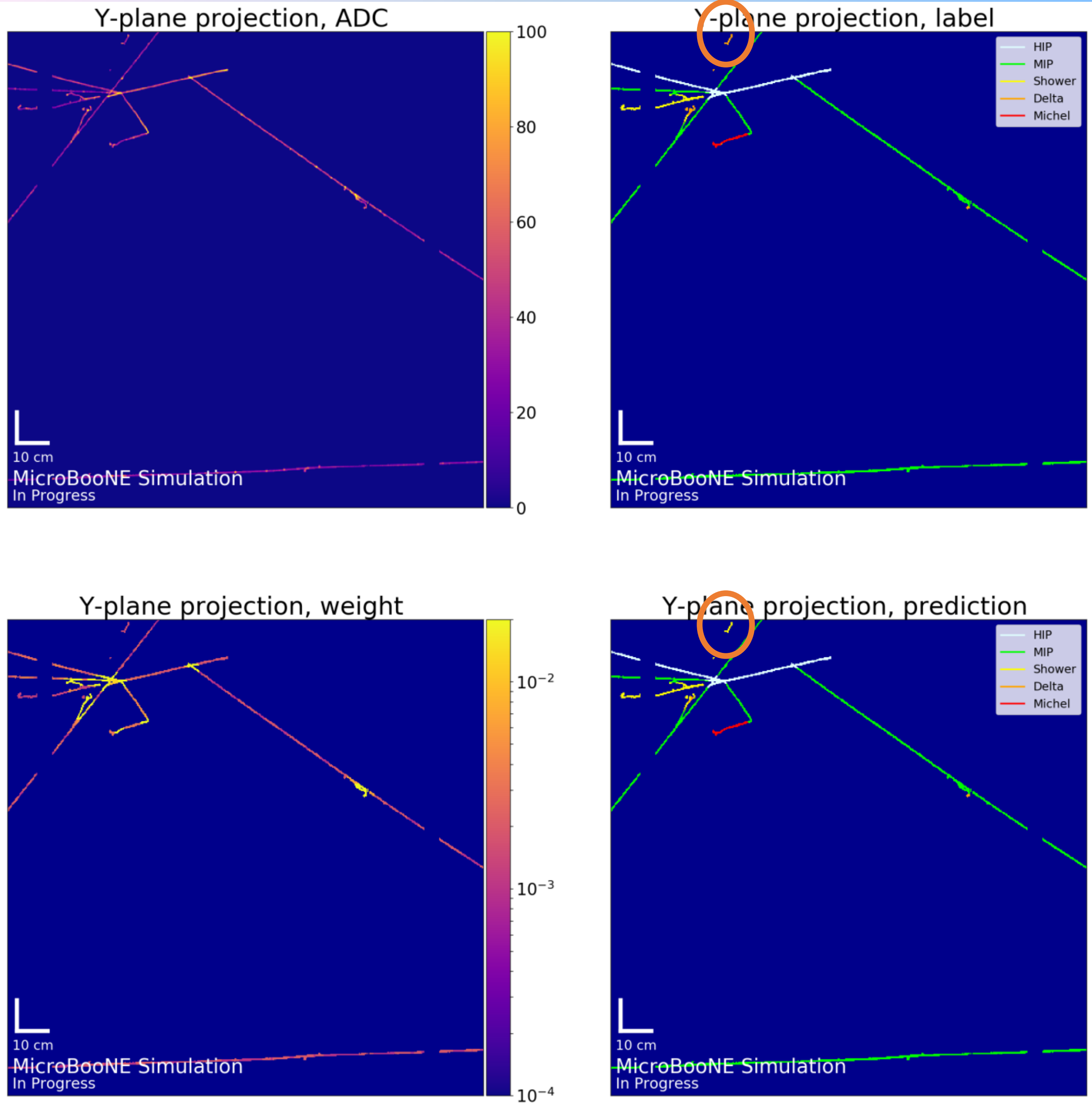






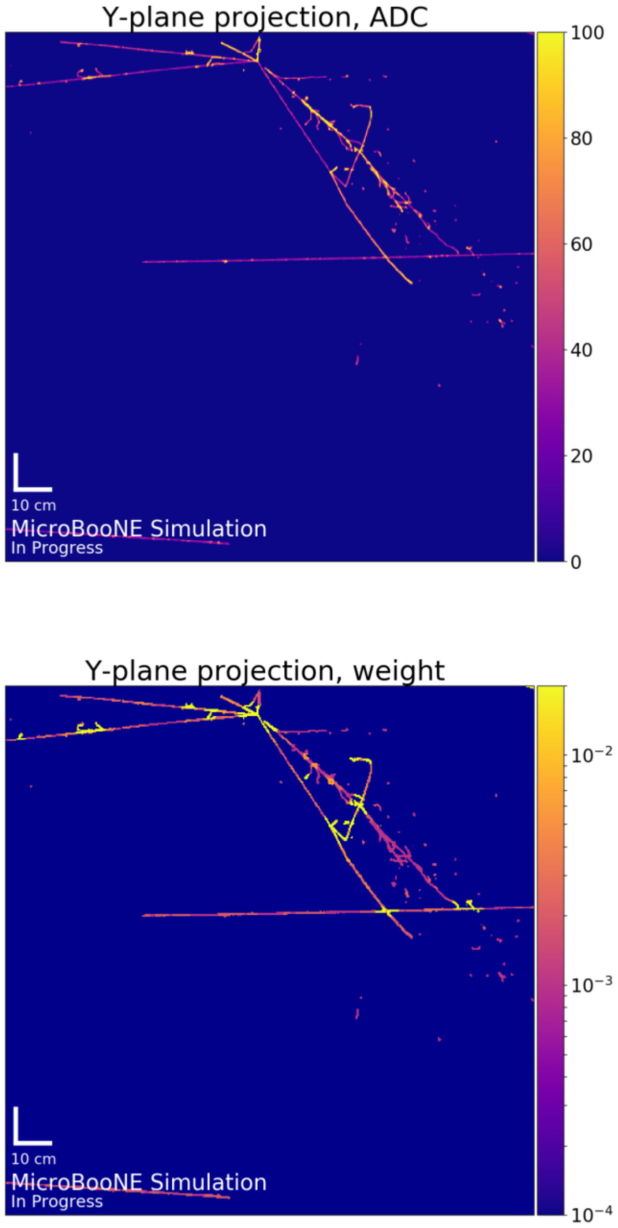


Predictions From Test Sample

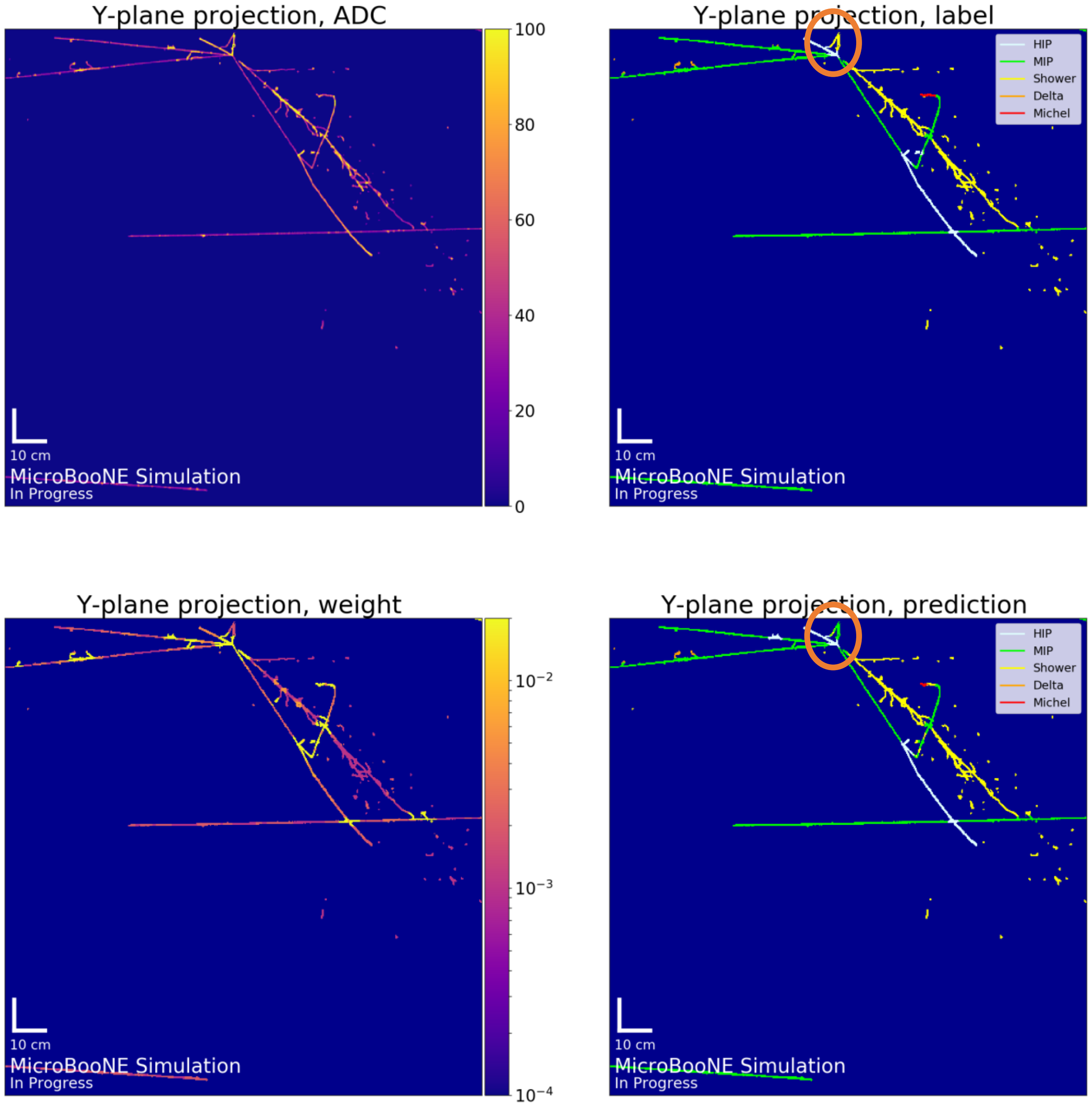


Predictions From Test Sample

- Predicted class is chosen by argmax
- Post processing can improve accuracy (e.g., delta score > 0.5)



Predictions From Test Sample

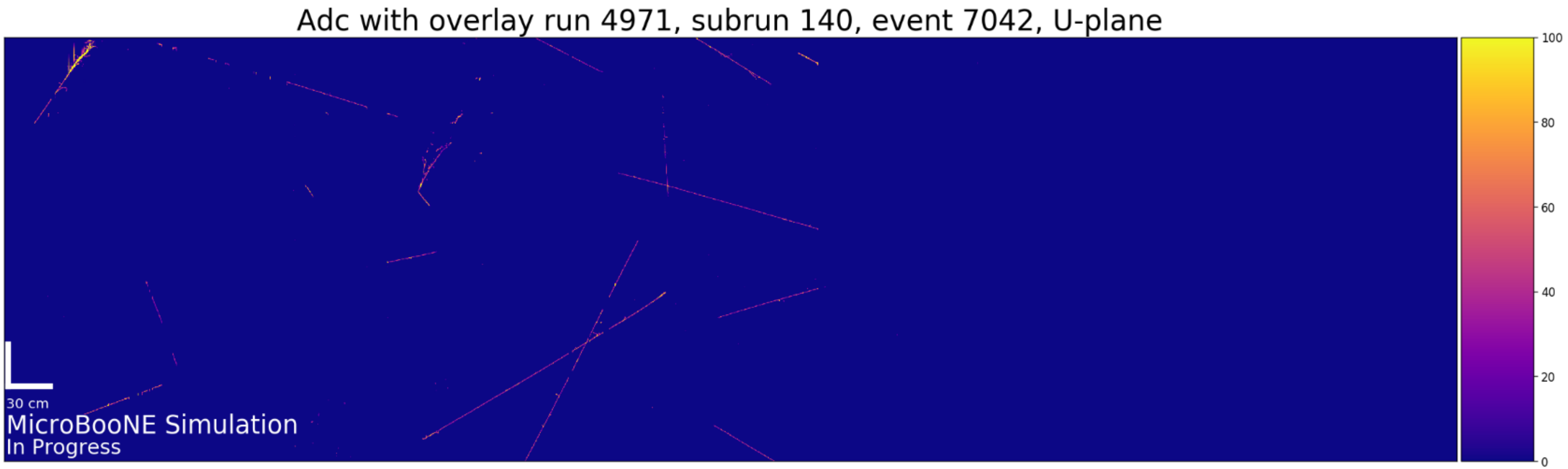


Predictions From Test Sample

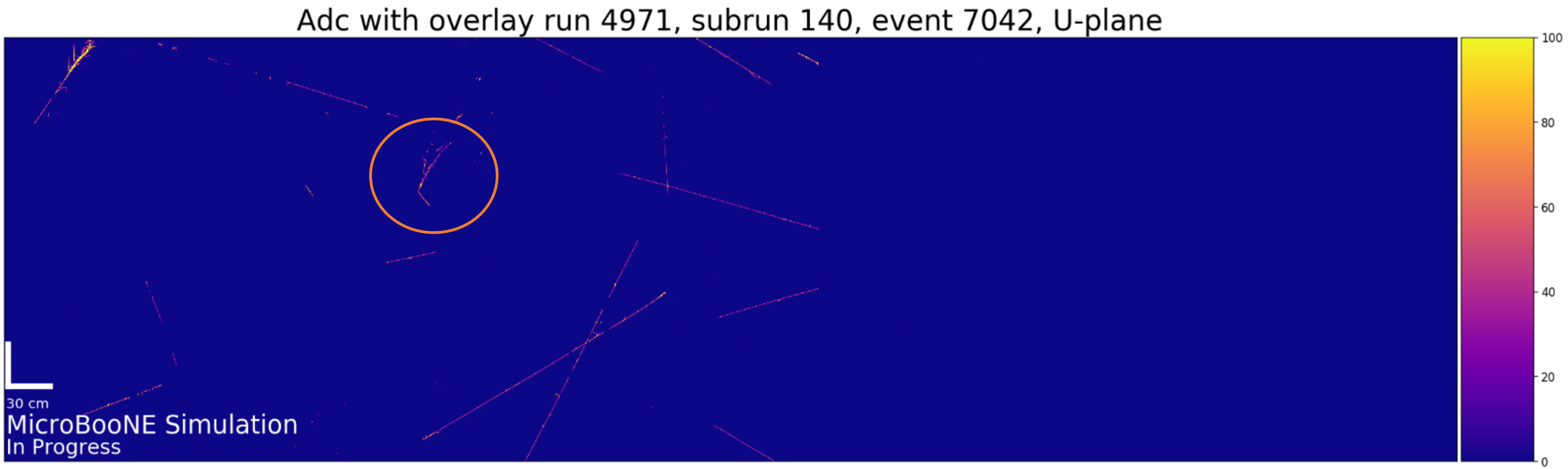
- Predicted class is chosen by argmax
- Post processing can improve accuracy (e.g., MIP score > 0.5)

- Previous images from test sample MC (512X512)
- Using same weights to predict full BNB Simulation (1008 X 3456)
- BNB Simulations:
    - Simulate interaction
    - Overlay cosmics (from off beam data)
    - Network prediction

Predictions from BNB simulations



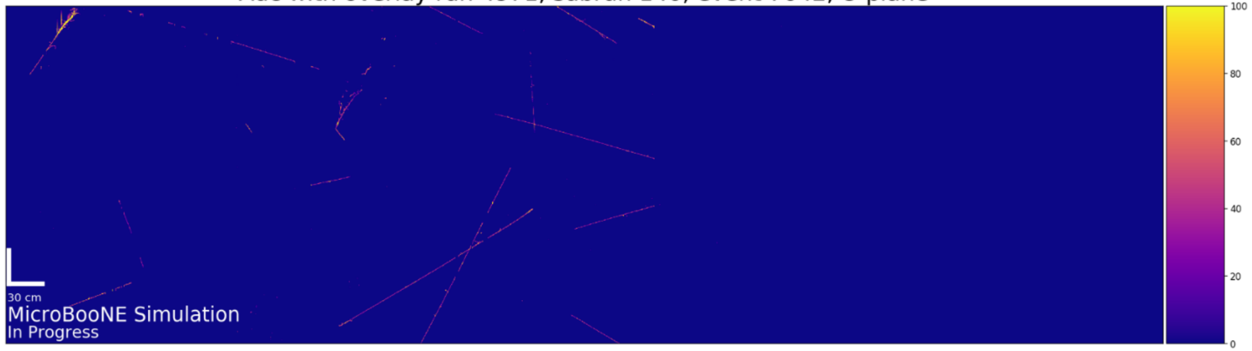
Predictions from BNB simulations



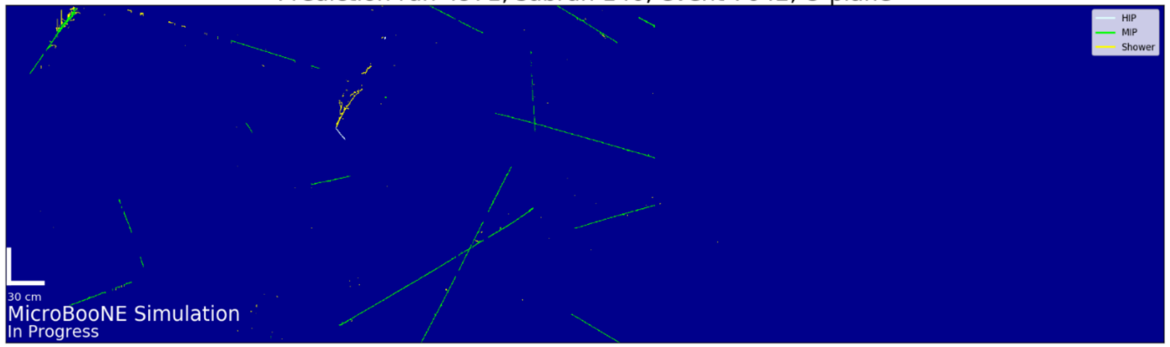
Predictions from BNB simulations

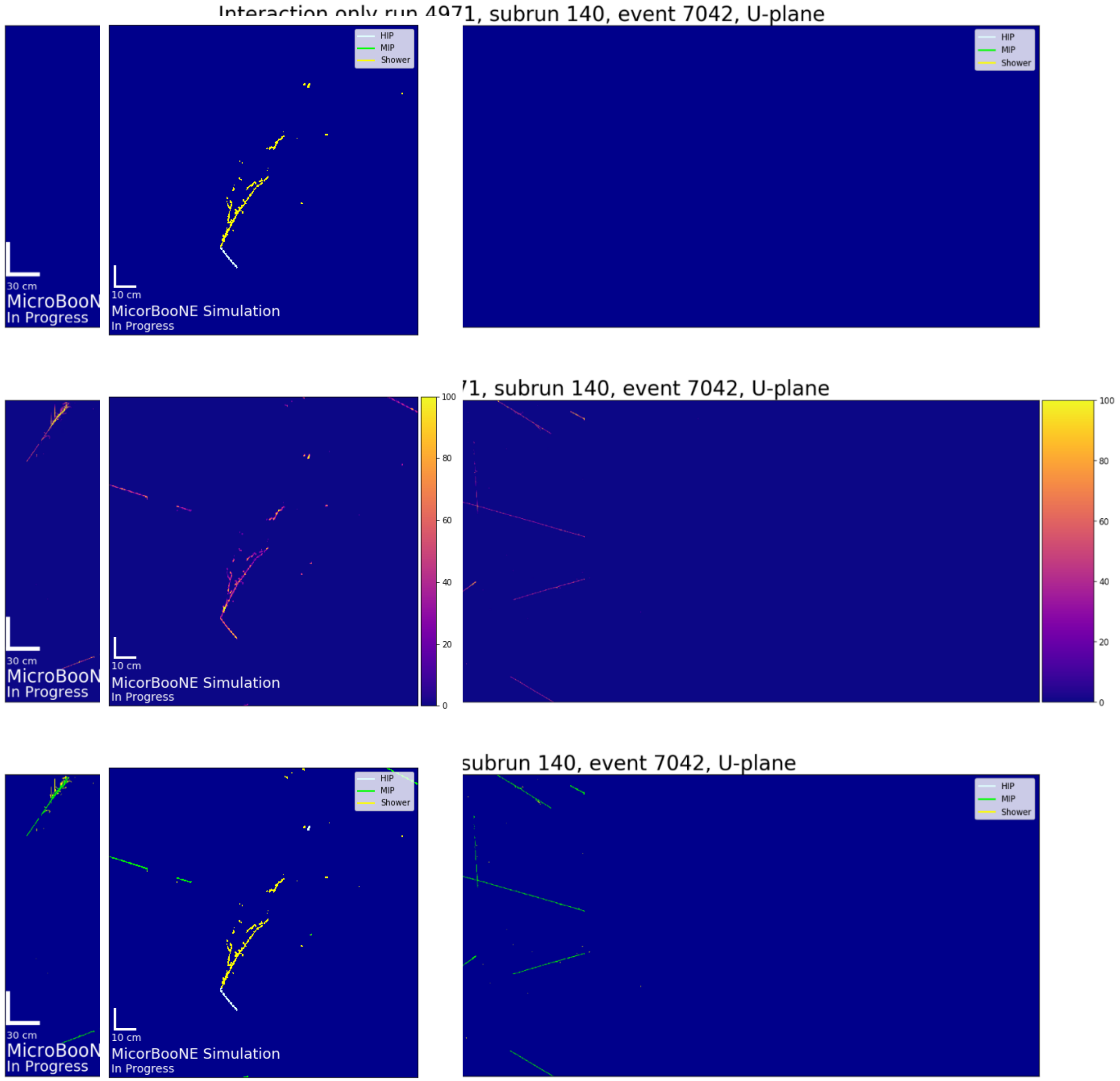
Interaction only run 4971, subrun 140, event 7042, U-plane

Adc with overlay run 4971, subrun 140, event 7042, U-plane



Prediction run 4971, subrun 140, event 7042, U-plane

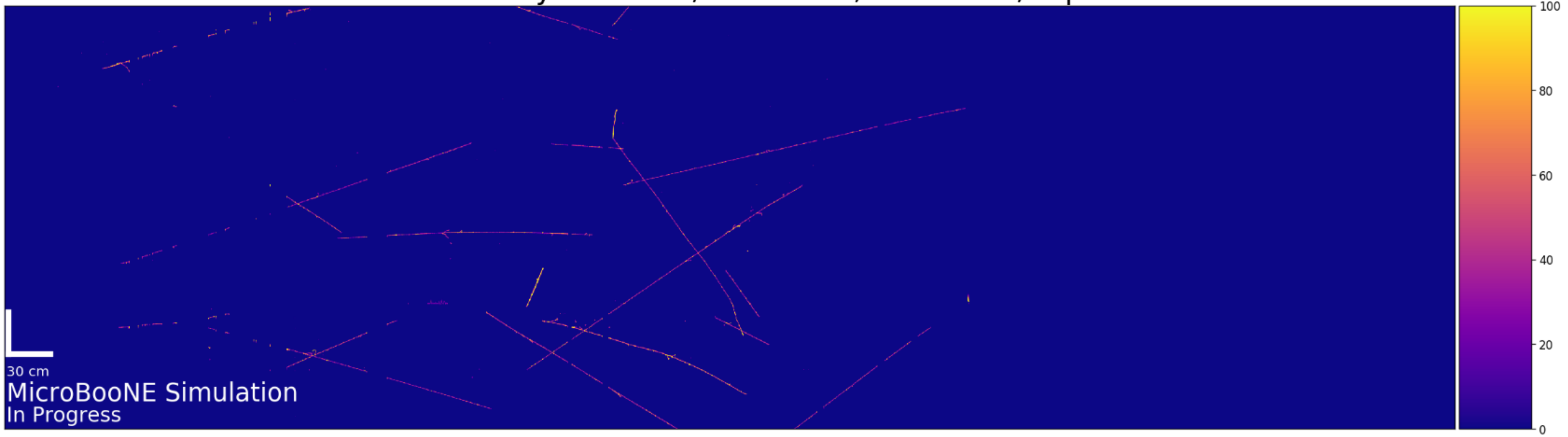




Predictions from BNB simulations

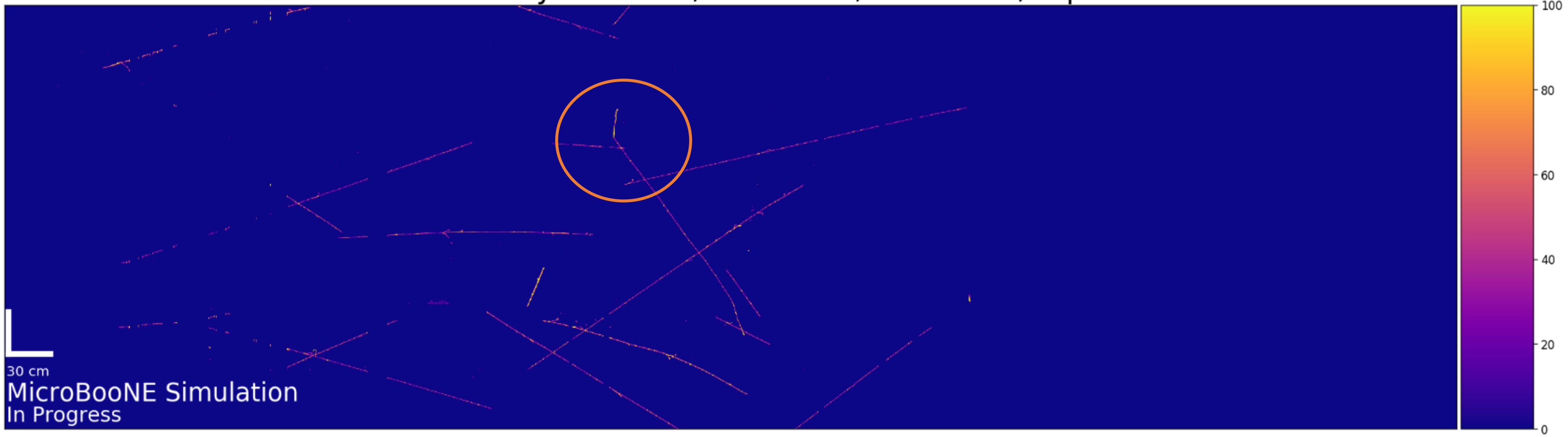
Predictions
from BNB simulations

Adc with overlay run 6837, subrun 11, event 599, U-plane

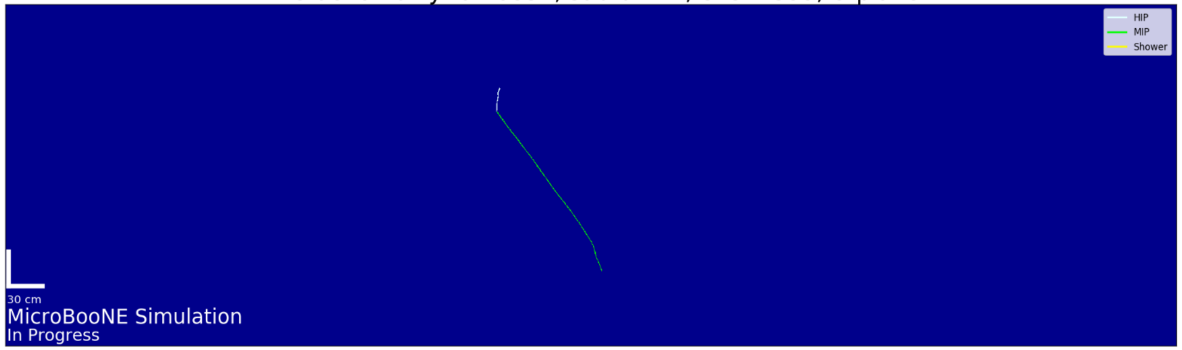


Predictions from BNB simulations

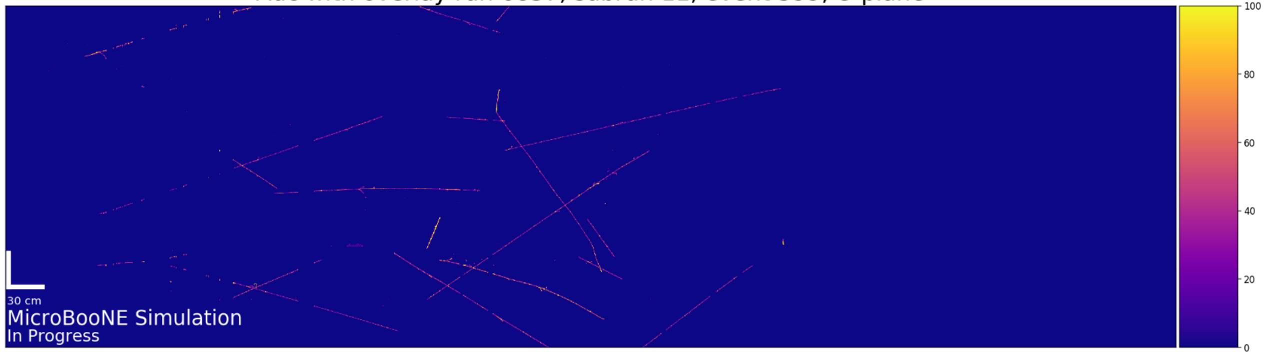
Adc with overlay run 6837, subrun 11, event 599, U-plane



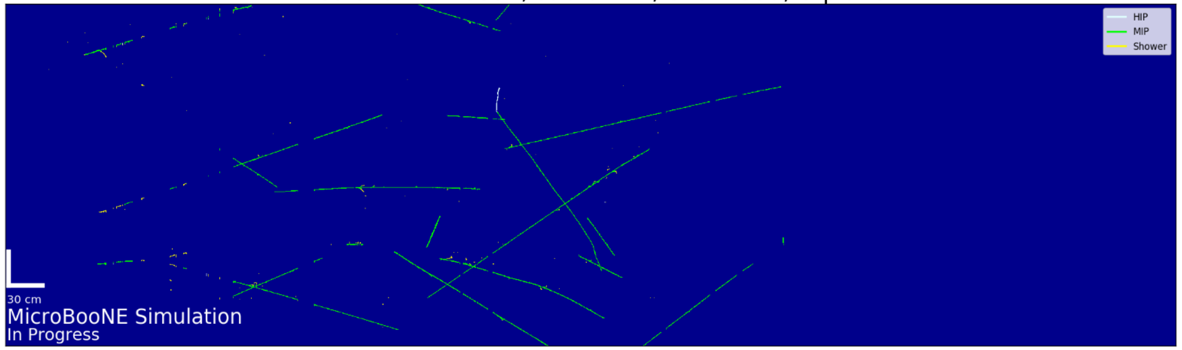
Interaction only run 6837, subrun 11, event 599, U-plane



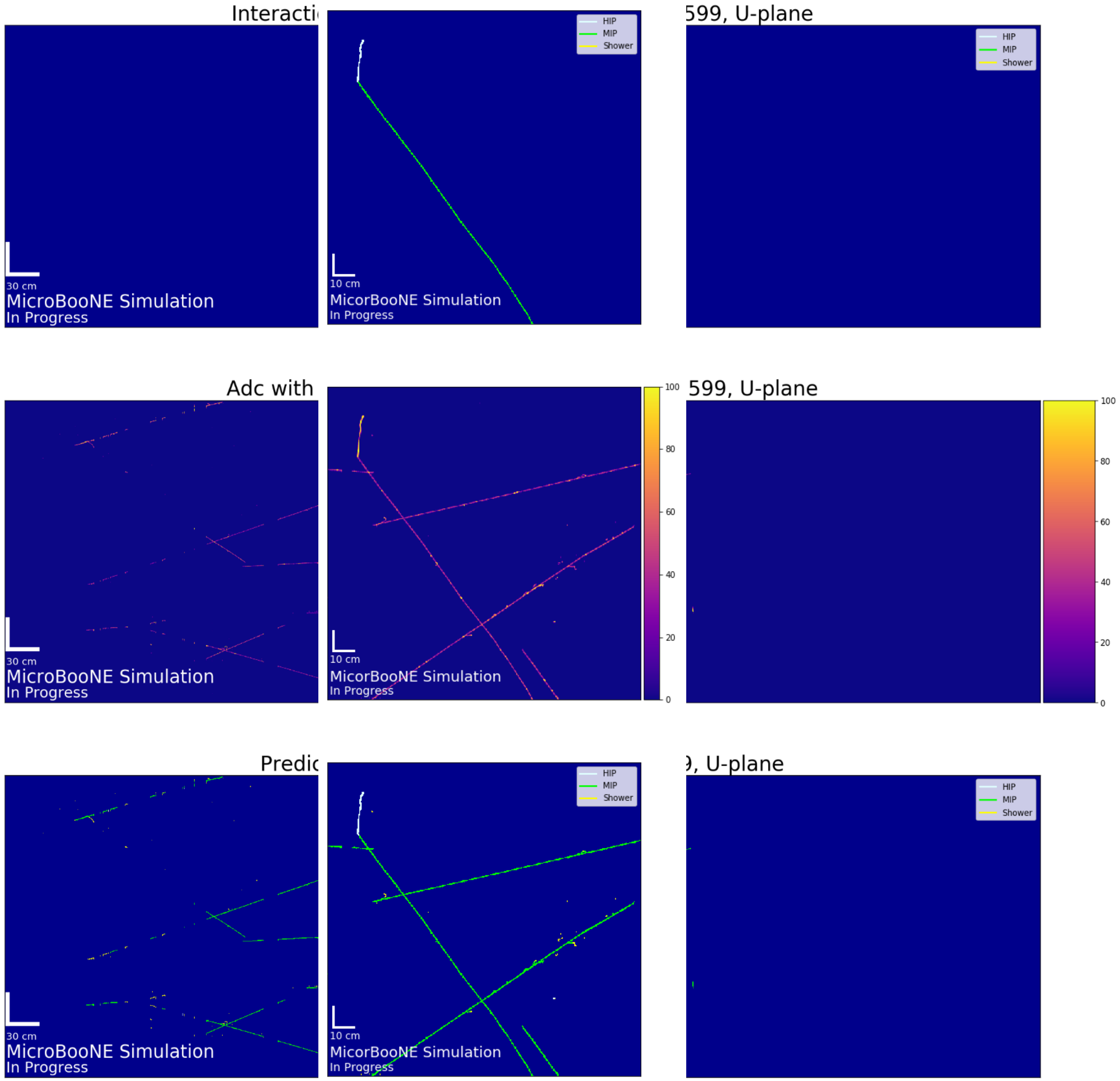
Adc with overlay run 6837, subrun 11, event 599, U-plane



Prediction run 6837, subrun 11, event 599, U-plane



Predictions from BNB simulations



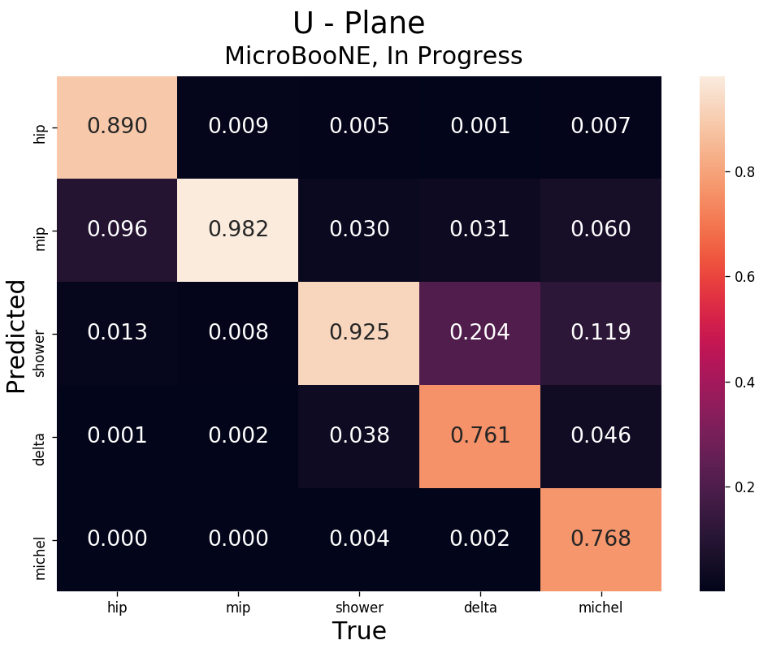
Predictions
from BNB simulations

Summary

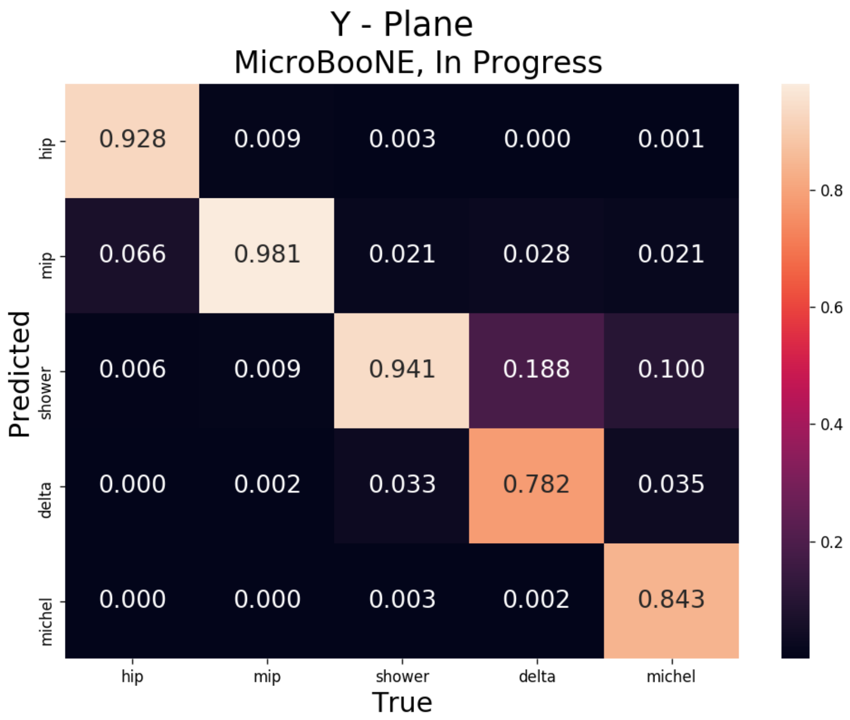
- MicroBooNE data can be represented by 2D images
- One of the first task in DL LEE analysis chain is semantic segmentation
- Data is very sparse **<1%** non zero pixels (in MicroBooNE <0.5% within threshold)
- SparseSSNet
    - Crop image to 60 images (GPU memory limitation) -> no cropping of the image.
    - Improvement in memory usage ~GB/full_image
    - Improvement in time ~2s/full_image (CPU)
    - Allows to utilize existing infrastructure with computing resource improvement, no need to "optimize" network.
    - Higher accuracy -> No "blurring" of data
- Scalability to 3D and bigger detectors (see Francois Drielsma's talk)

Questions

Backup Slides


U - Plane
MicroBooNE, In Progress
Predicted
True
hip mip shower delta michel
hip: 0.890 0.009 0.005 0.001 0.007
mip: 0.096 0.982 0.030 0.031 0.060
shower: 0.013 0.008 0.925 0.204 0.119
delta: 0.001 0.002 0.038 0.761 0.046
michel: 0.000 0.000 0.004 0.002 0.768
0.8 0.6 0.4 0.2


V - Plane
MicroBooNE, In Progress
Predicted
True
hip mip shower delta michel
hip: 0.898 0.008 0.005 0.002 0.004
mip: 0.092 0.982 0.024 0.035 0.038
shower: 0.009 0.008 0.938 0.213 0.133
delta: 0.000 0.002 0.030 0.747 0.051
michel: 0.000 0.000 0.003 0.003 0.773
0.8 0.6 0.4 0.2


Y - Plane
MicroBooNE, In Progress
Predicted
True
hip mip shower delta michel
hip: 0.928 0.009 0.003 0.000 0.001
mip: 0.066 0.981 0.021 0.028 0.021
shower: 0.006 0.009 0.941 0.188 0.100
delta: 0.000 0.002 0.033 0.782 0.035
michel: 0.000 0.000 0.003 0.002 0.843
0.8 0.6 0.4 0.2