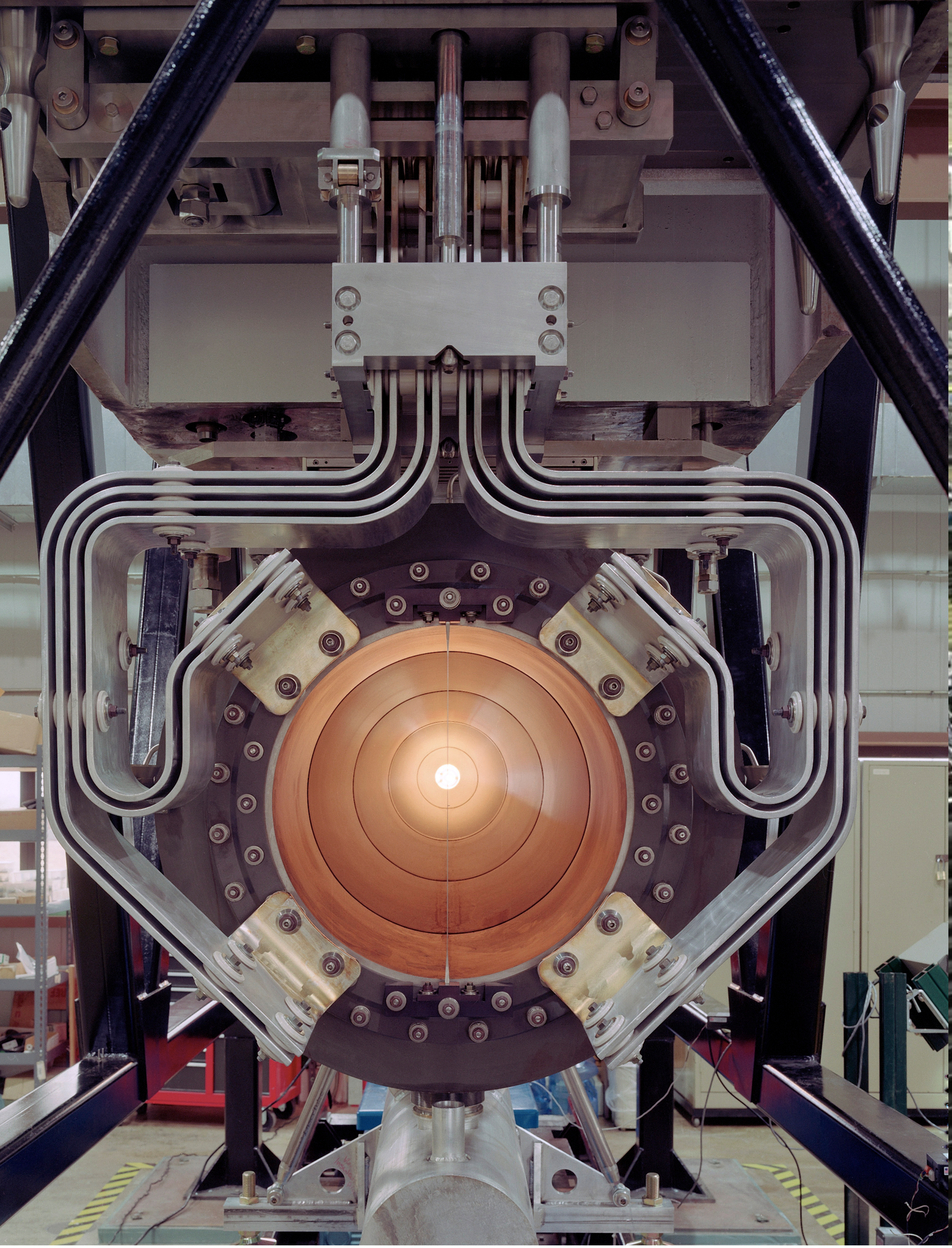
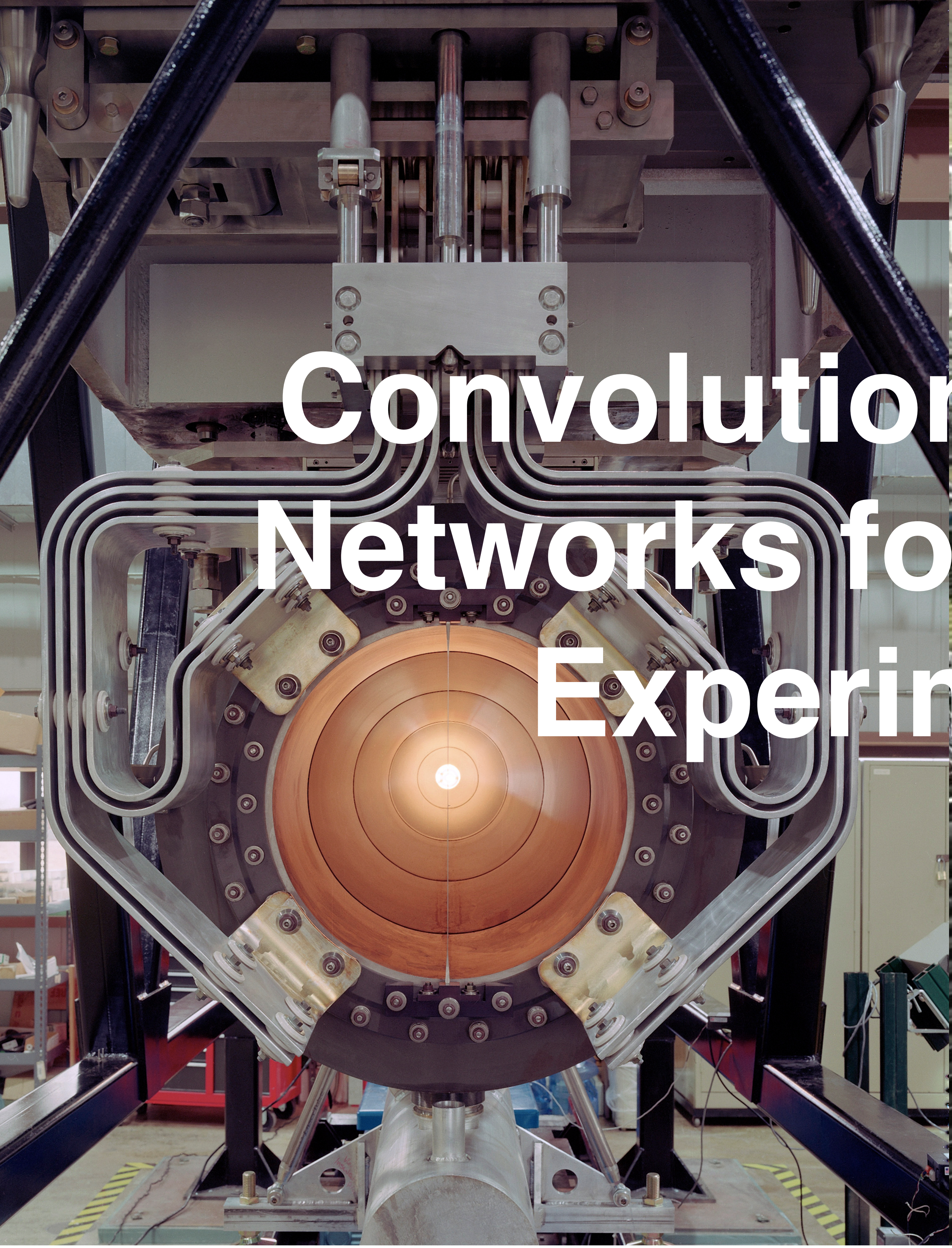


Alexander Radovic

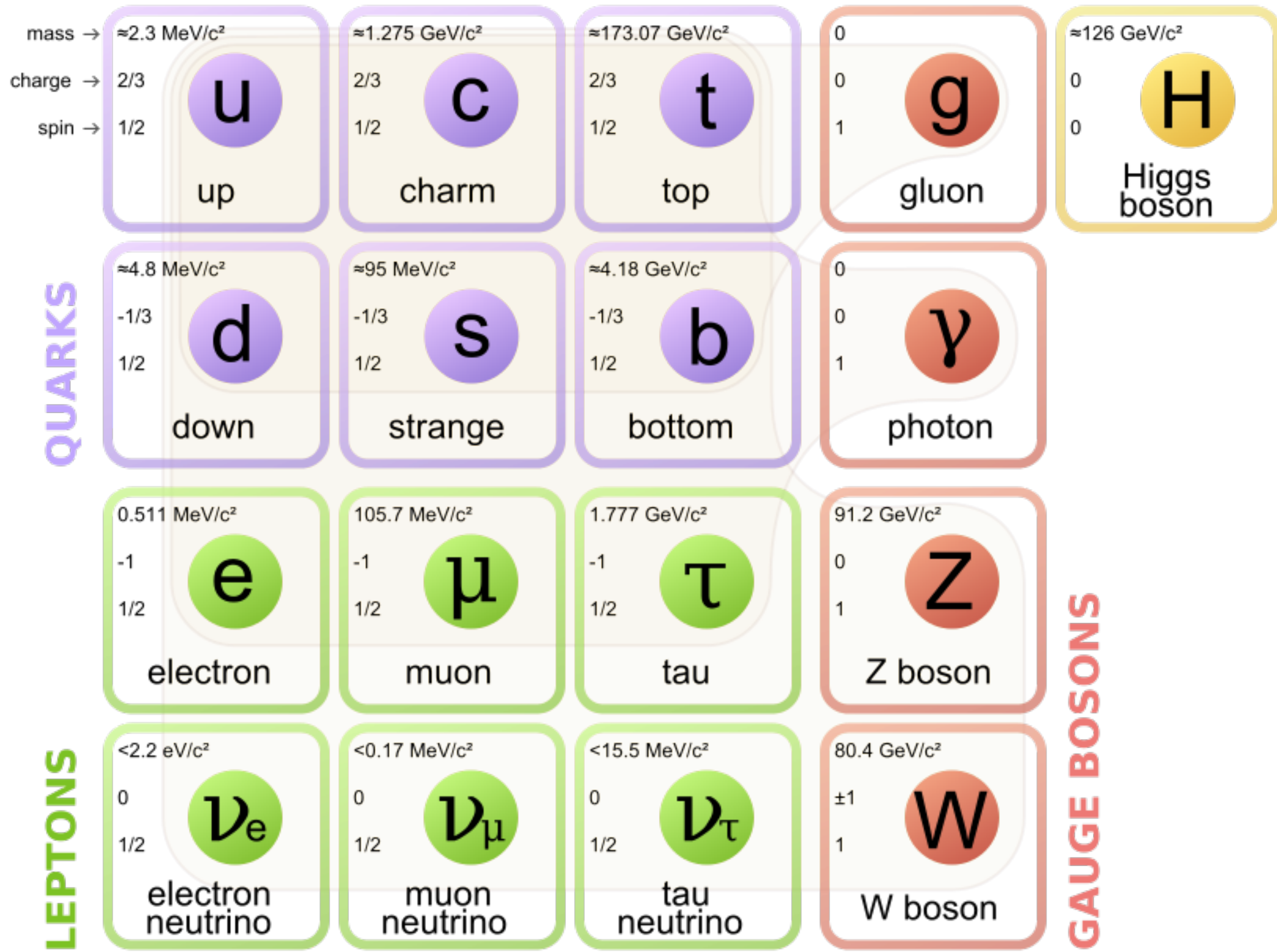
College of William and Mary



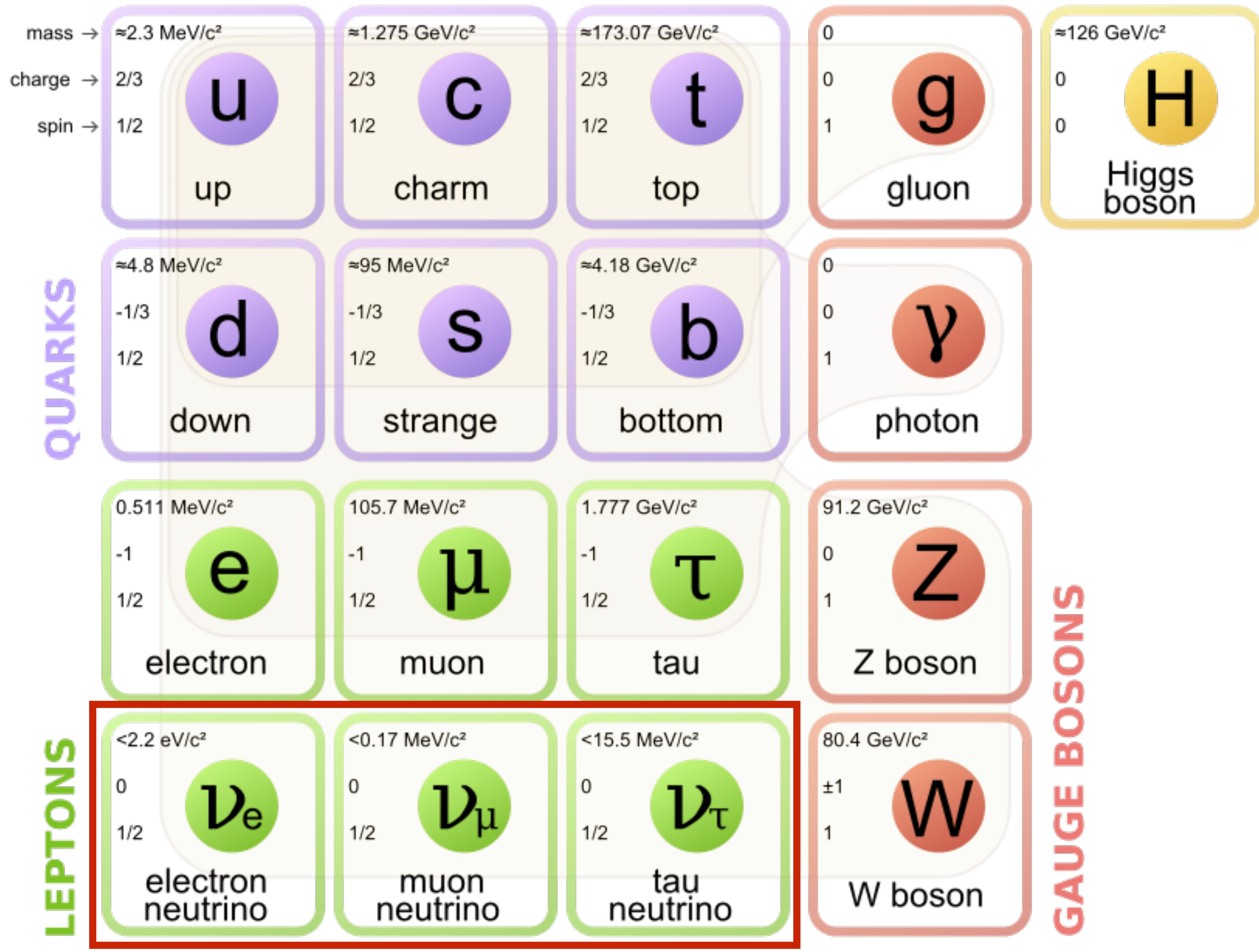
Convolutional Neural Networks for Neutrino Experiments

Alexander Radovic
College of William and Mary

Neutrino Oscillations 101



Neutrino Oscillations 101

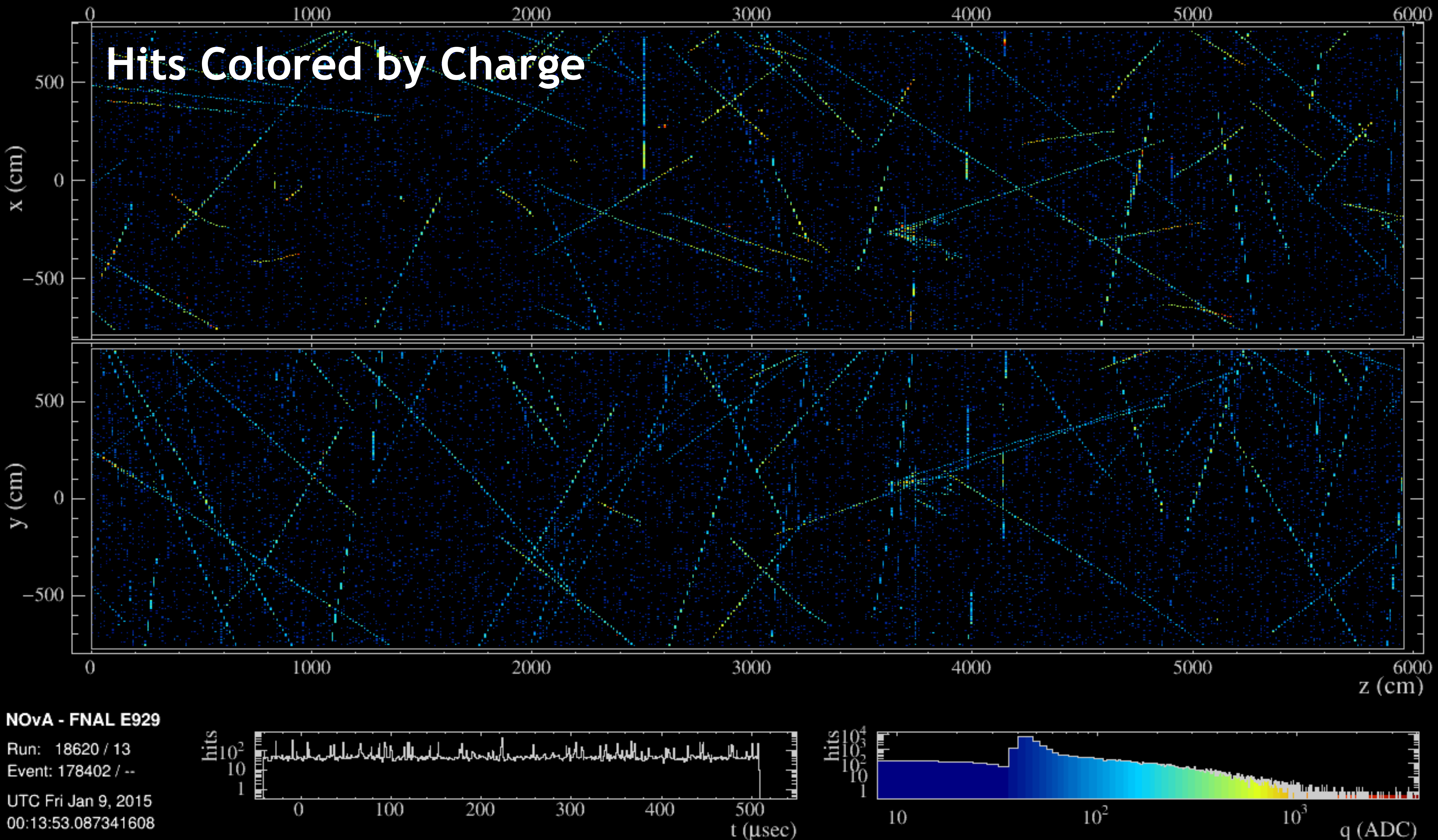


“I have done a terrible thing, I have postulated a particle that cannot be detected.”
-Wolfgang Pauli

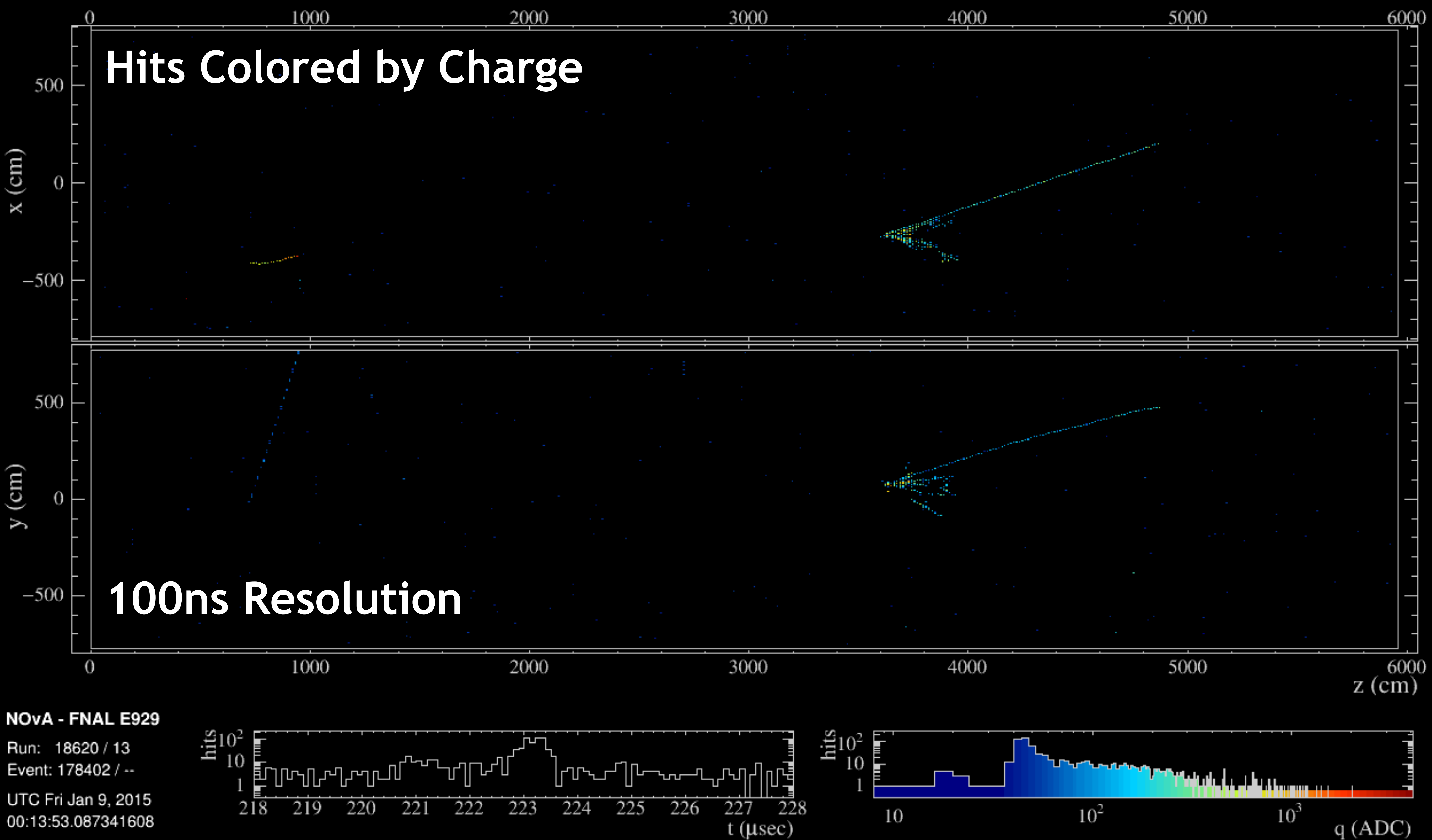
“I have done a terrible thing, I have postulated a particle that ~~cannot~~ can be detected.”

-Wolfgang Pauli

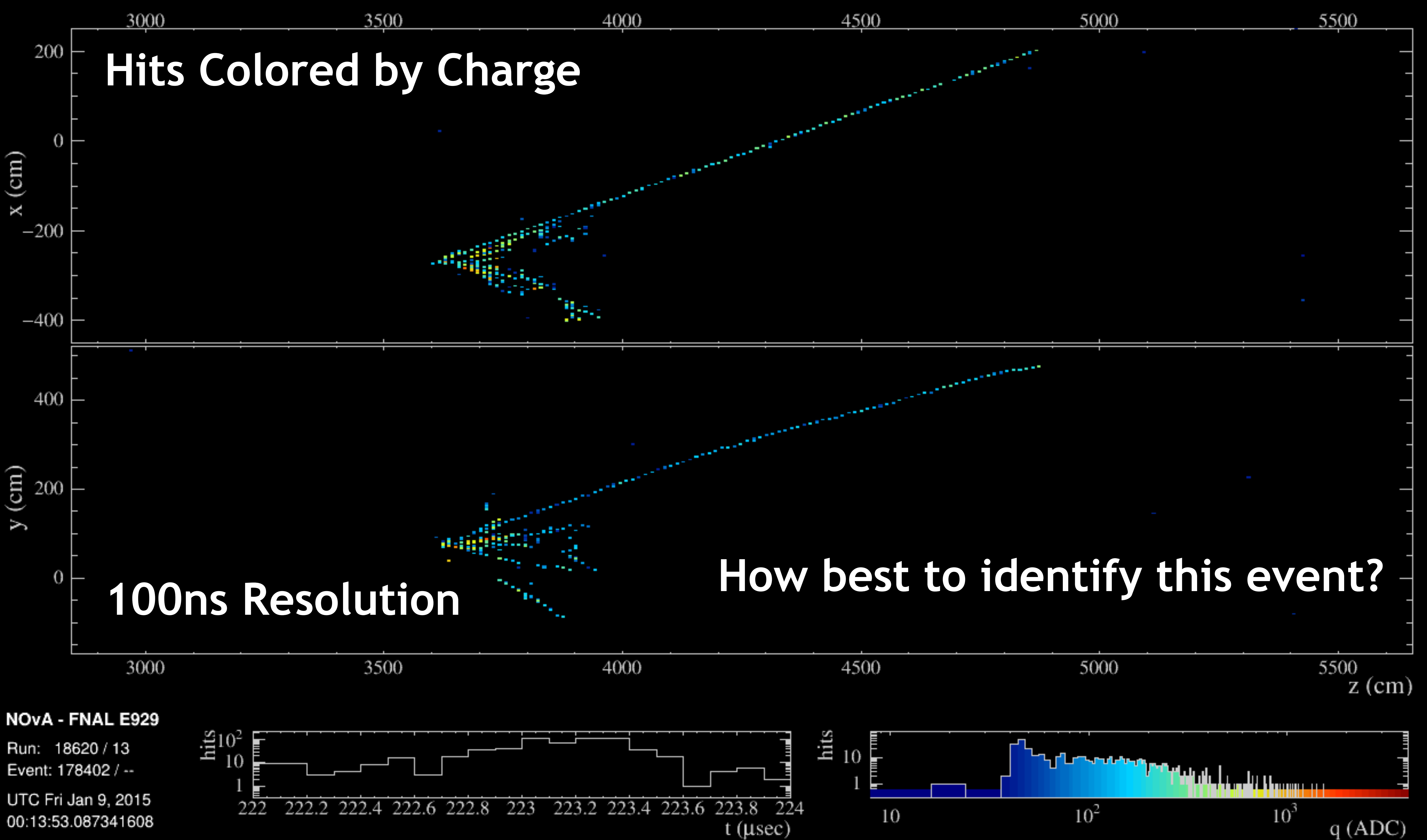
550 μs exposure of the NOvA Far Detector



Time-zoom on 10 μs interval during NuMI beam pulse



Close-up of neutrino interaction in the NOvA Far Detector



The Motivation

Why Study Neutrinos?

Neutrino oscillations raises as many questions as it answers:

- Why is lepton sector mixing much larger than quark sector mixing?
- What is the hierarchy of neutrino masses and how does this effect searches for a majorana neutrino?
- Is there CP violation in the lepton sector? Could it be large enough to explain observed matter antimatter asymmetry of our universe?

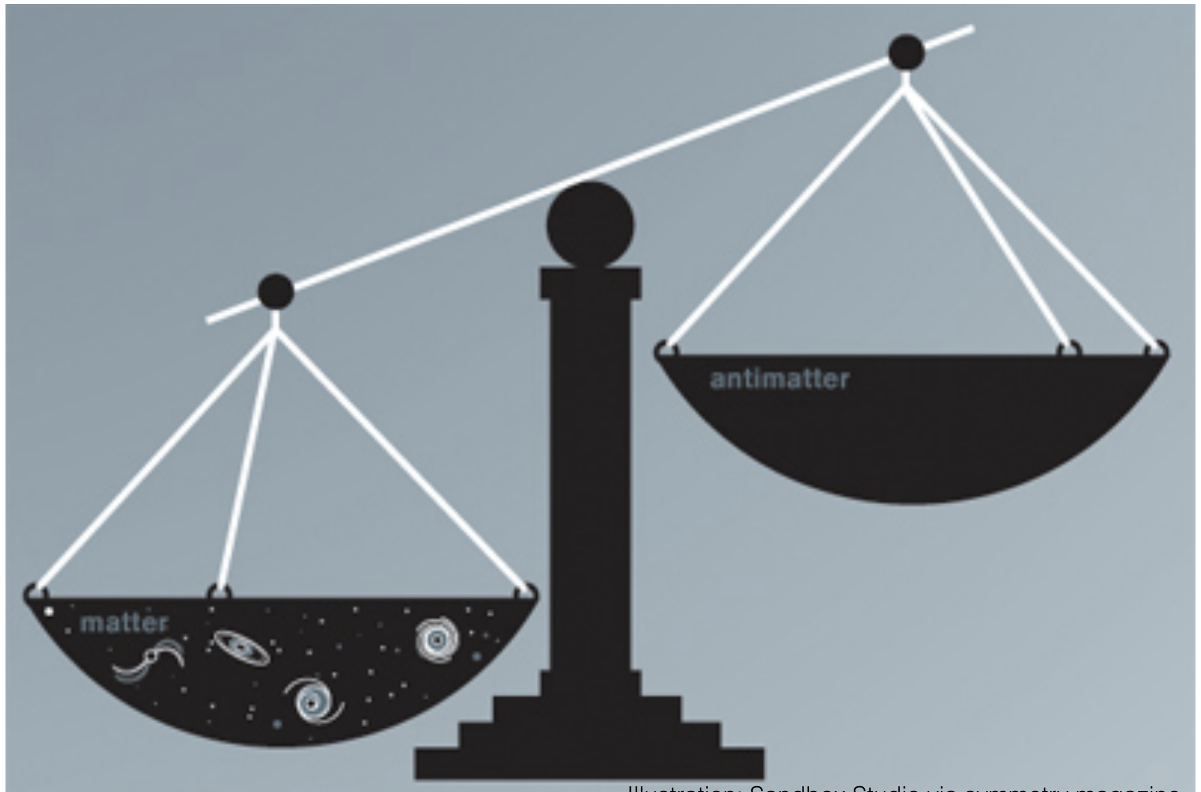
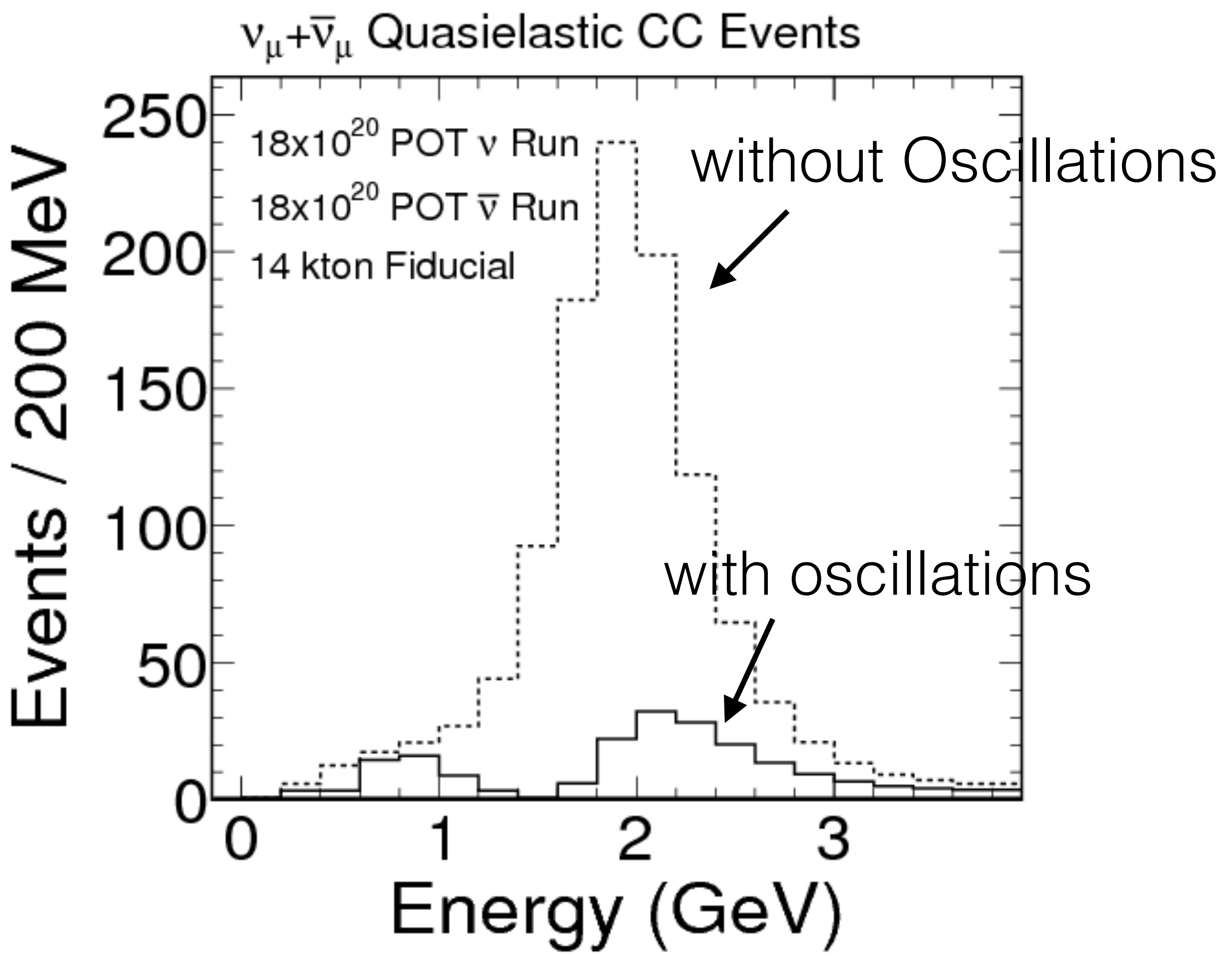


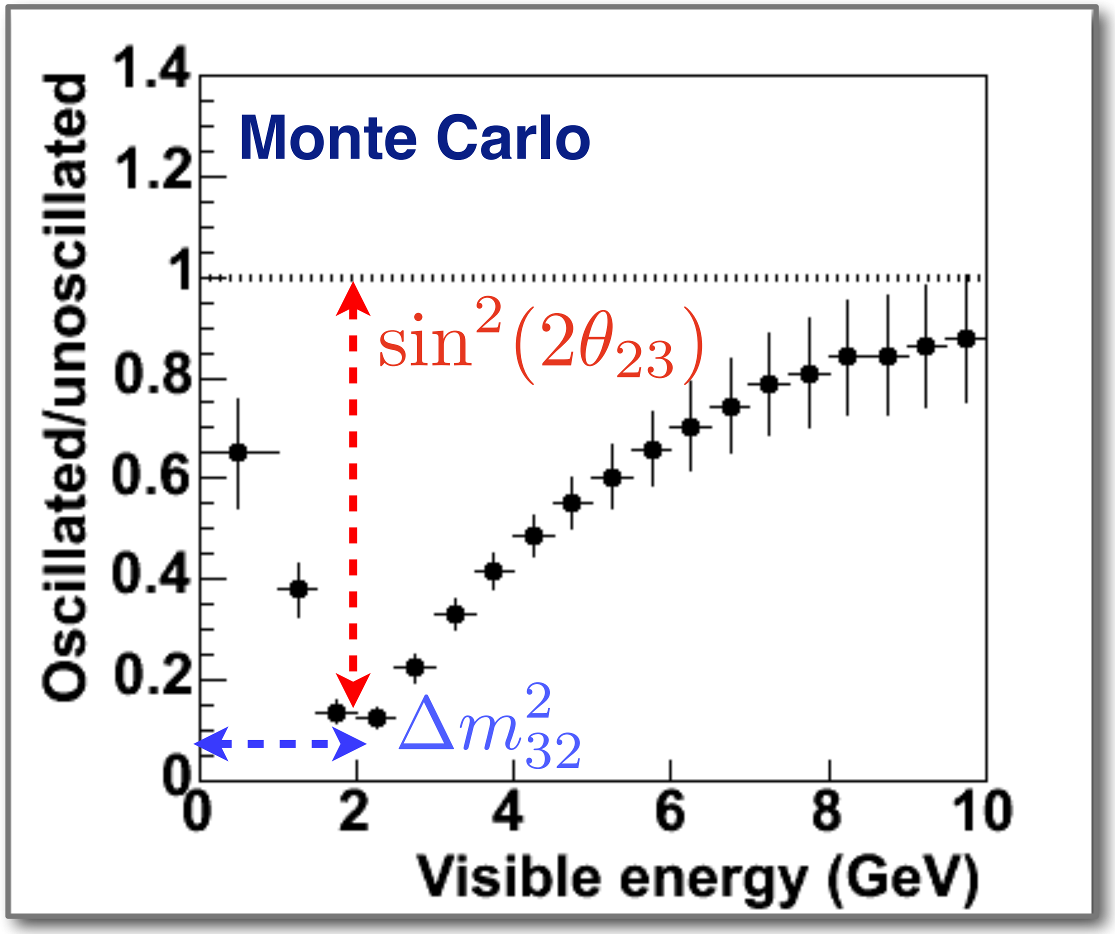
Illustration: Sandbox Studio via symmetry magazine

Why Convolutional Neural Networks?

- Measuring neutrino oscillations is all about measuring how neutrinos change between different lepton flavor states as a function of distance traveled and neutrino energy.

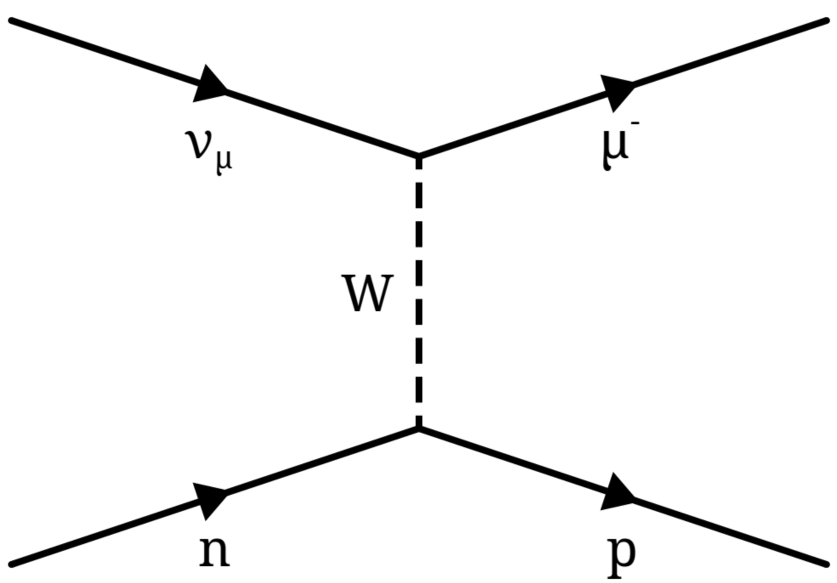
$$P(\nu_\mu \rightarrow \nu_\mu) \approx 1 - \sin^2(2\theta_{23}) \sin^2\left(\frac{1.27\Delta m^2_{atm} L}{E}\right)$$



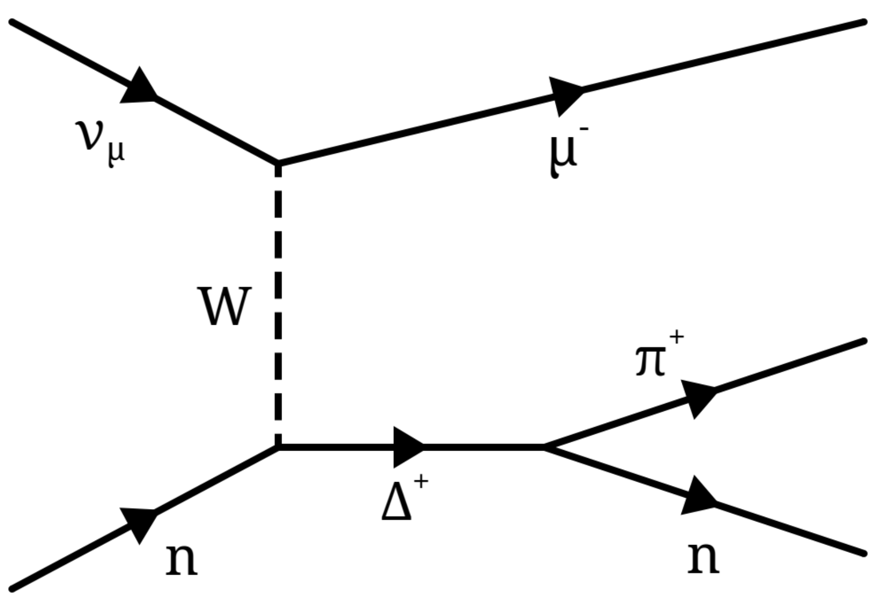


Why Convolutional Neural Networks?

- That means that any oscillation analysis can benefit from precise identification of the interaction in two ways:
  - Estimating the lepton flavor of the incoming neutrino.
  - Correctly identifying the type of neutrino interaction, to better estimate the neutrino energy, aka is it a quasi elastic event or a resonance event?



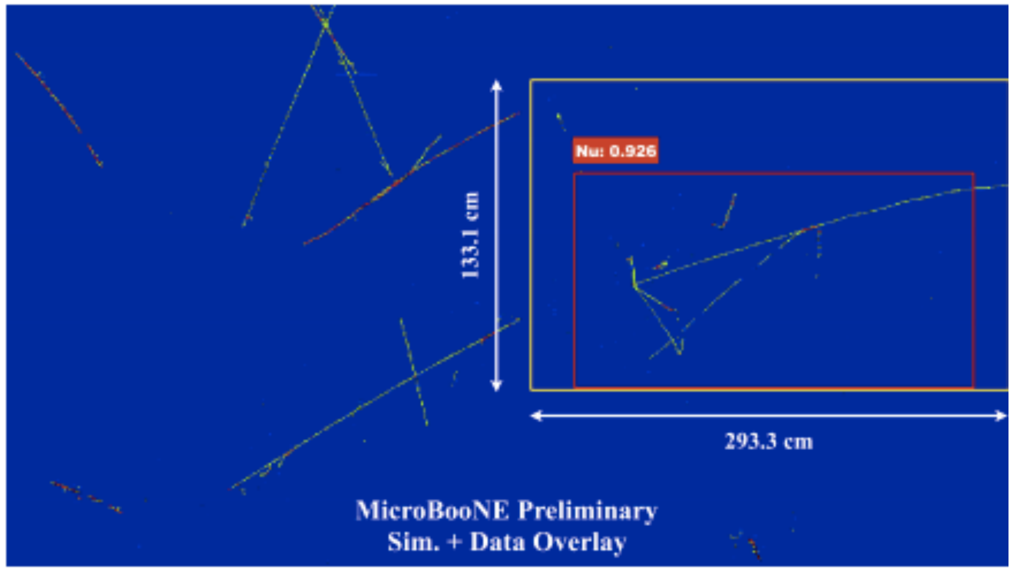
Quasi-Elastic

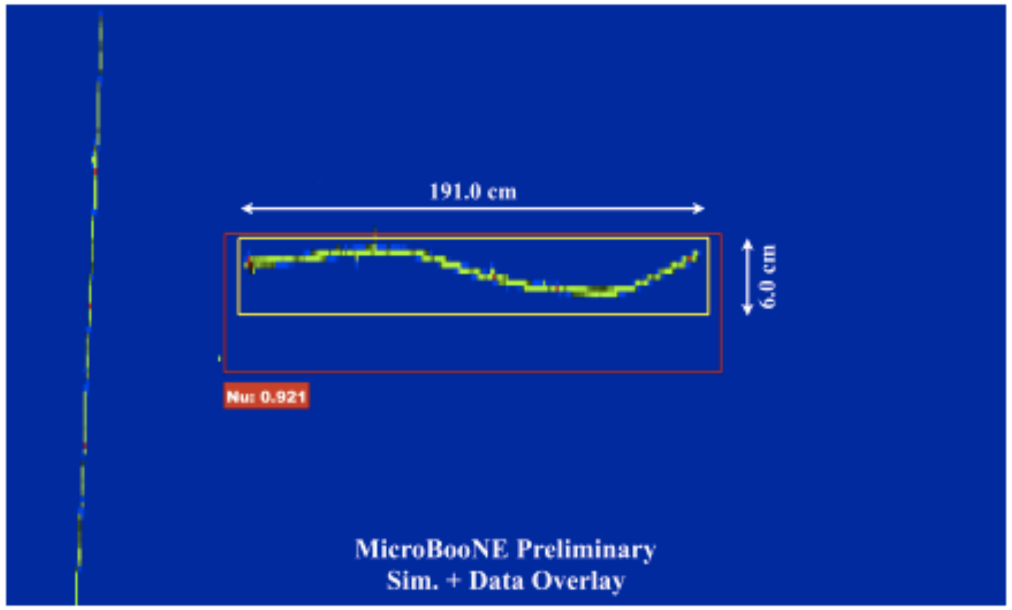


Resonance

Why Convolutional Neural Networks?

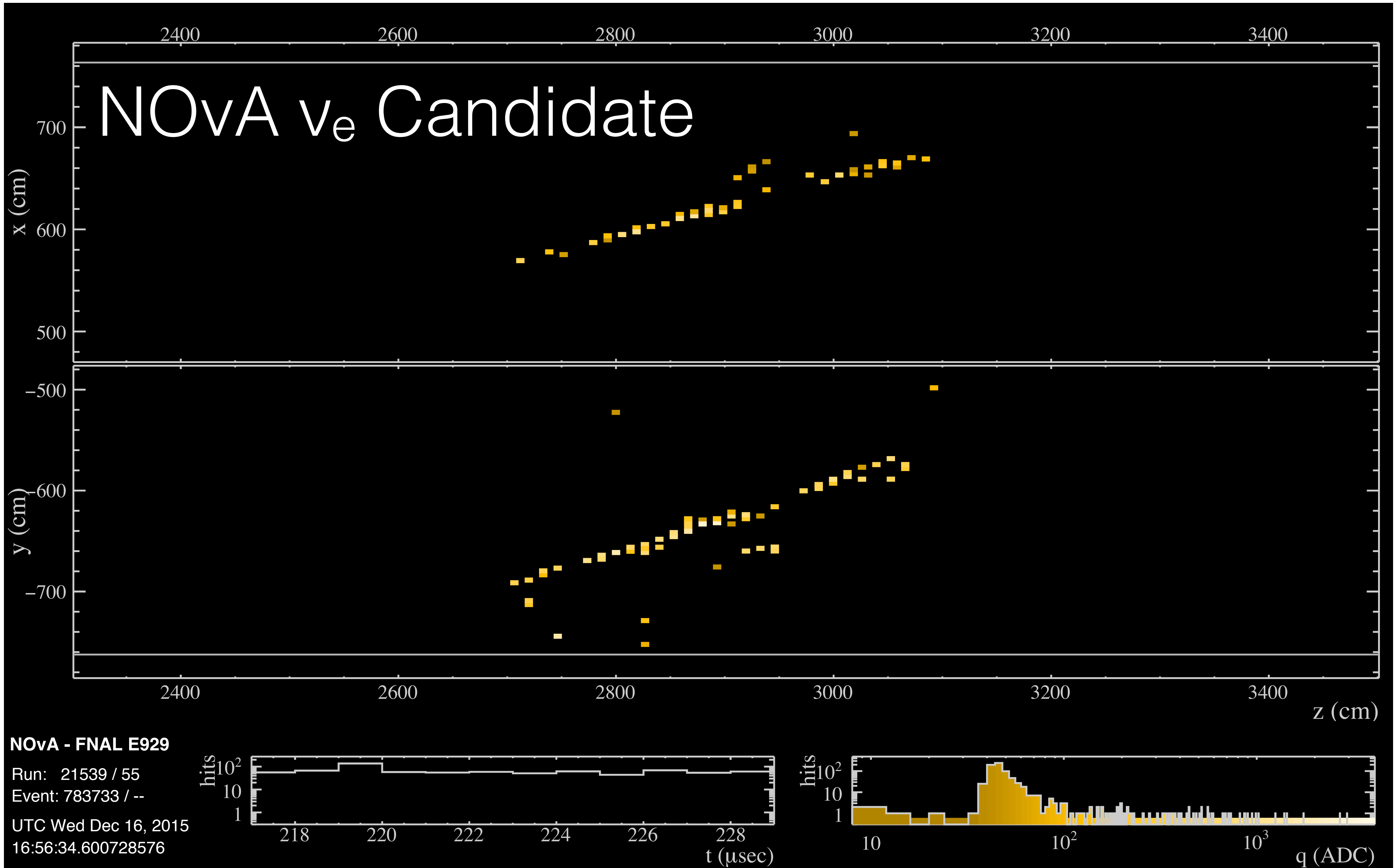
- Furthermore we as we want to accurately estimate the neutrino energy we need precise & robust reconstruction:
  - Identifying reconstructed objects in the event.
  - R-CNNs to isolate activity related to the interaction of interest.
  - Ultimately semantic segmentation to combine local and global information smoothly for hit-by-hit identification.





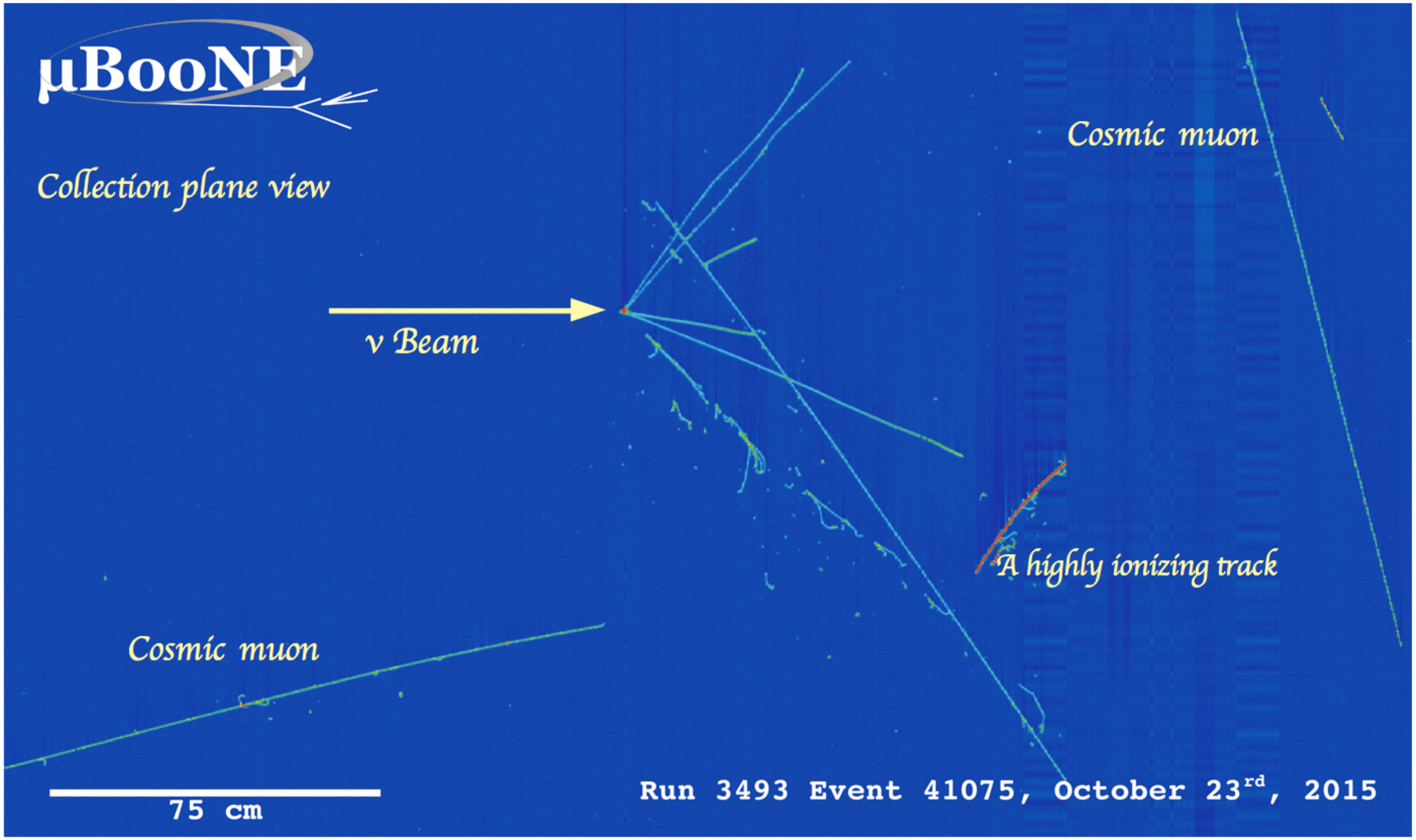
Why Convolutional Neural Networks?

- Our detectors are also often the perfect domain:
  - Large ~uniform volumes where spatially invariant response is a benefit.
  - Usually only one or two detector systems.



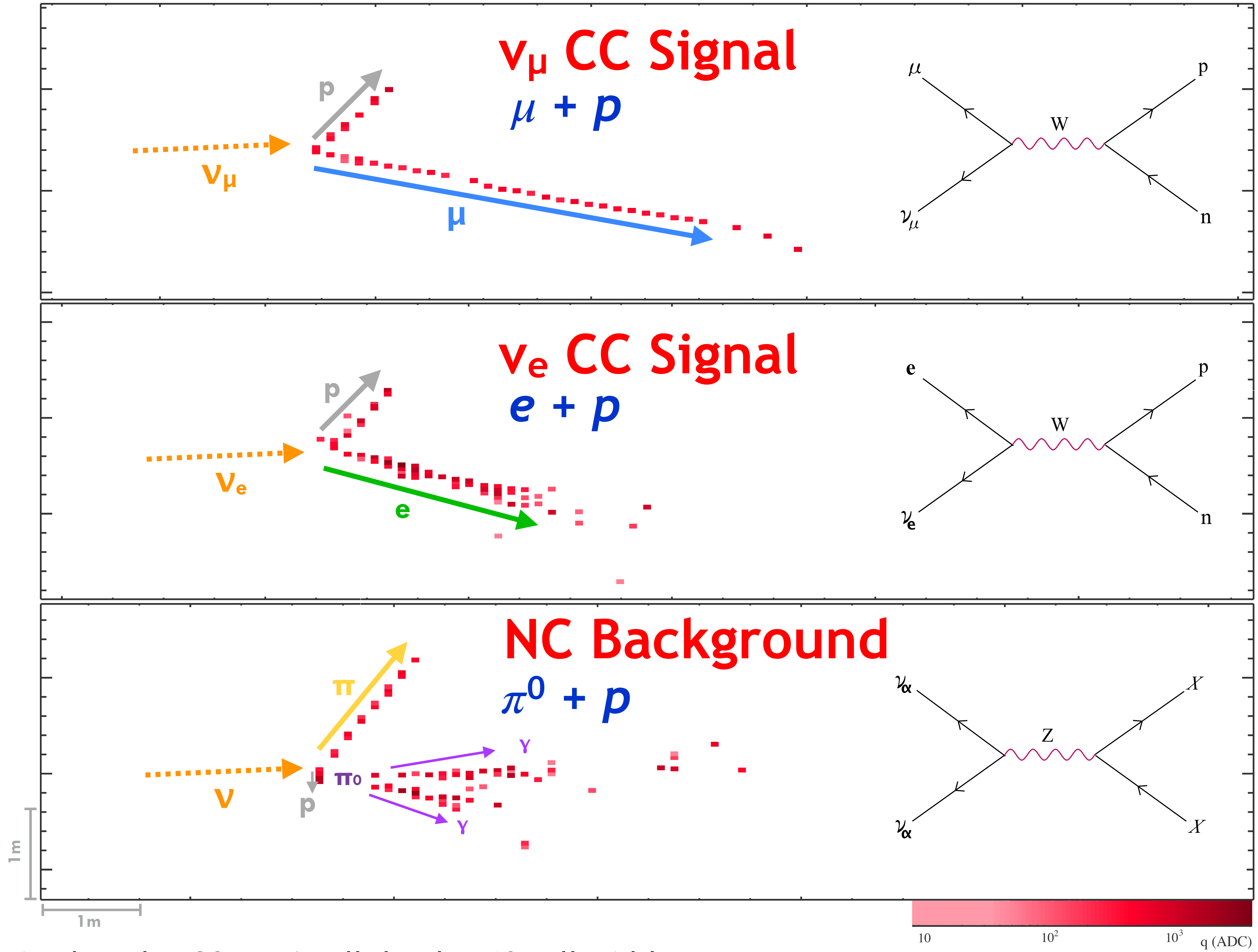
Why Convolutional Neural Networks?

- Our detectors are also often the perfect domain:
  - Large ~uniform volumes where spatially invariant response is a benefit.
  - Usually only one or two detector systems.



NOvA

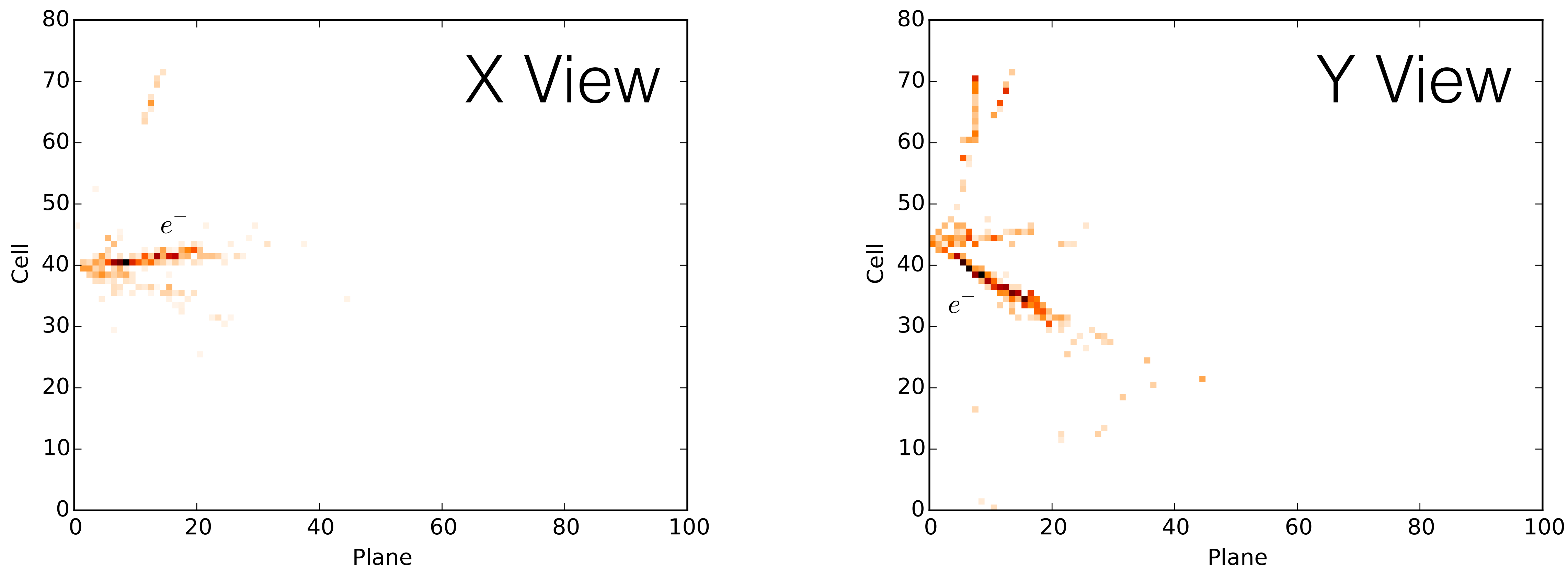
NOvA Event Topologies



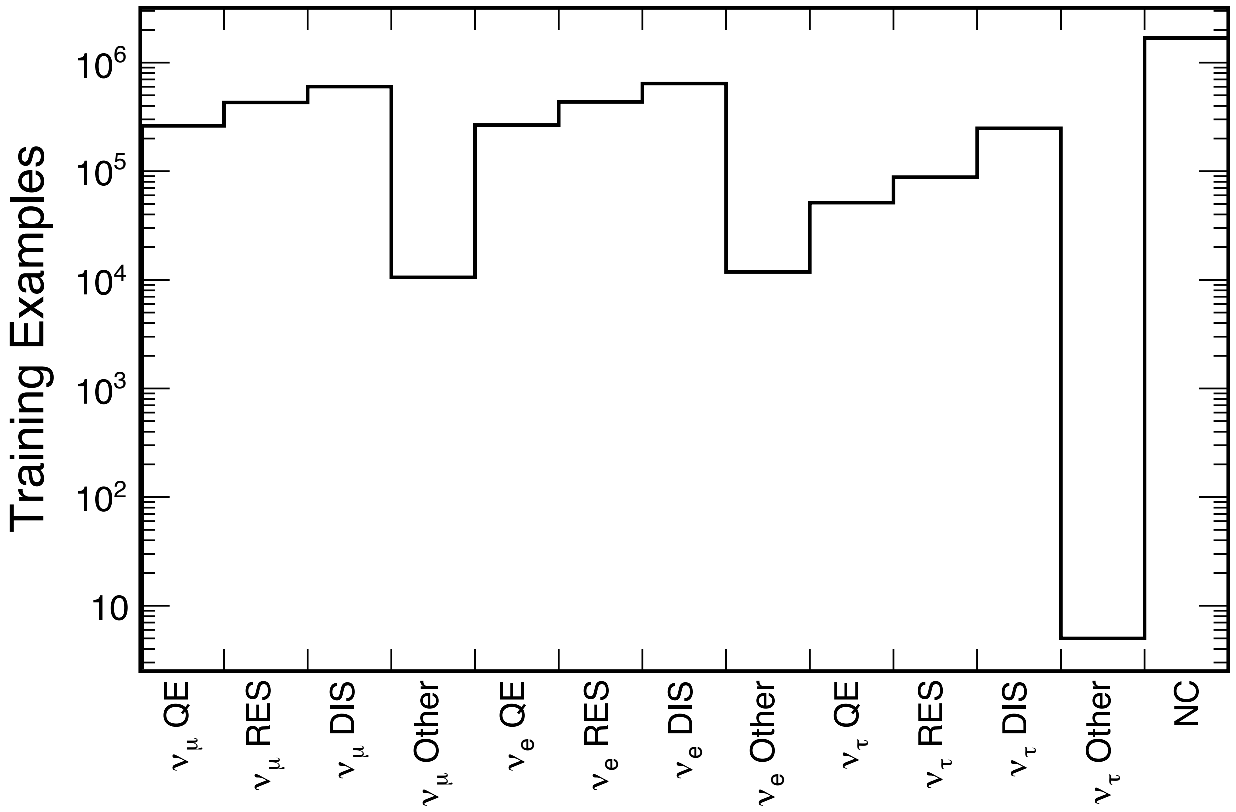
1 radiation length = 38cm (6 cell depths, 10 cell widths)

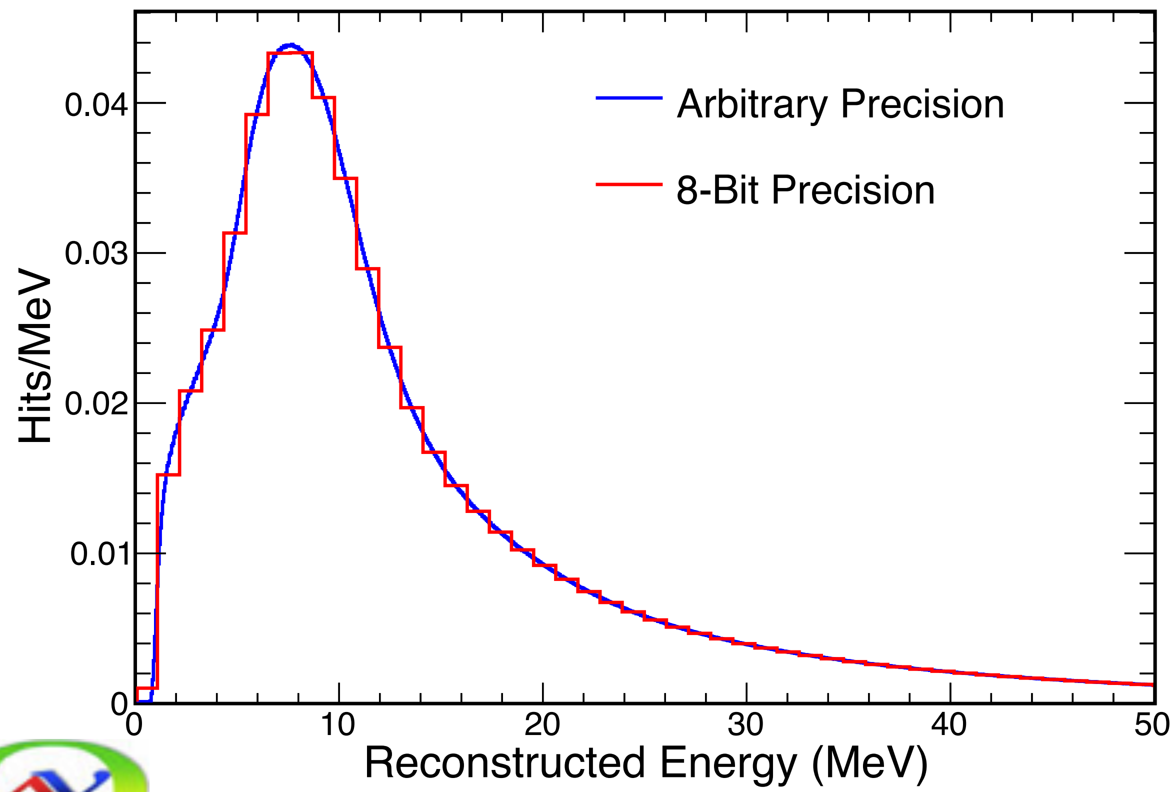
Our Input

Our input “image”, is a pair of maps of the hits in a tight space/time window. One for the X view and another for the Y view. Each “pixel” is the calibrated energy response in that cell. All “images” have the same dimensions- 100 planes by 80 cells.



The Training Sample





- 4.7 million, minimally preselected simulated events, pushed into LevelDB databases: 80% for training and 20% for testing.
- Rescale calibrated energy depositions to go from 0 to 255 and truncate to chars for dramatically reduced file size at no loss of information
- Fine tuned with 5 million cosmic data events taken from an out of beam time minimal bias trigger.

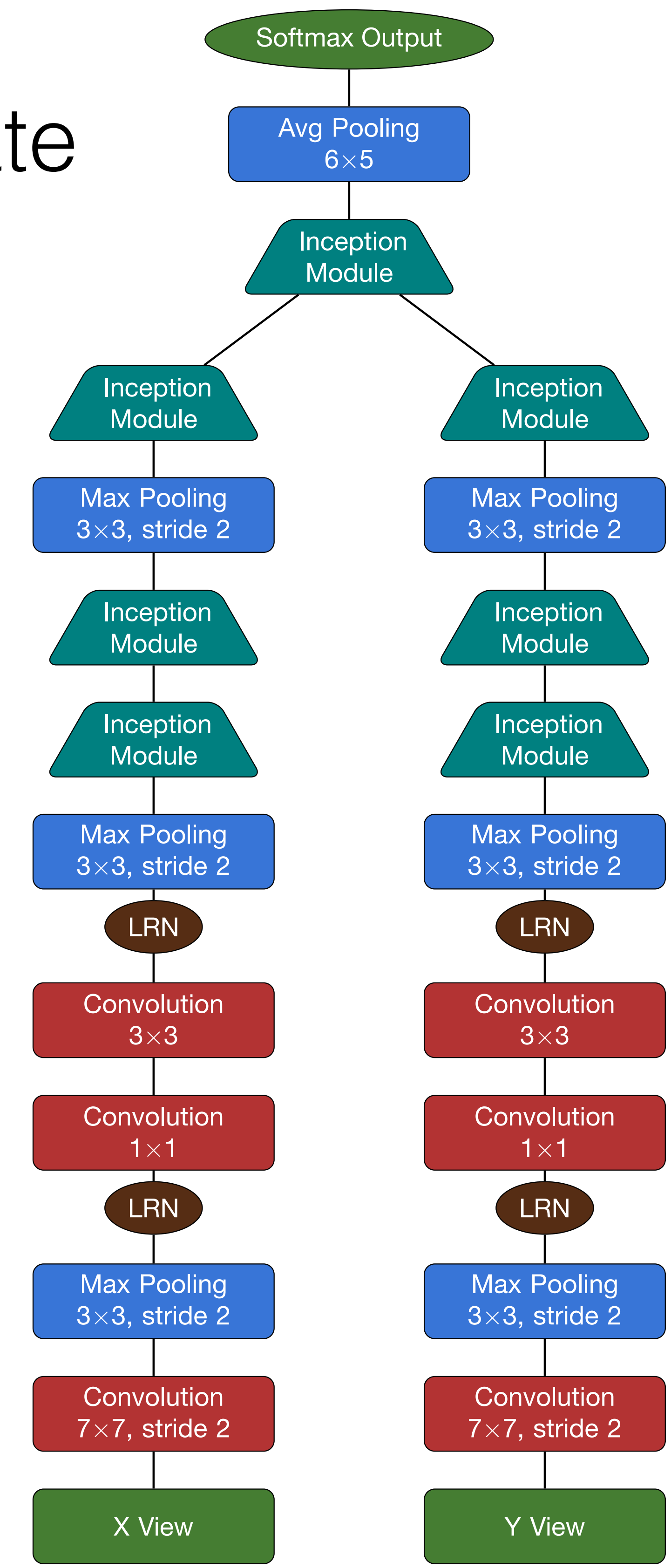
Our Architecture

Based on the first googlenet. Largest innovation is splitting each view into a separate sequence of layers and concatenating the outputs near the end of the network. Named “Convolutional Visual Network”, or **CVN**.

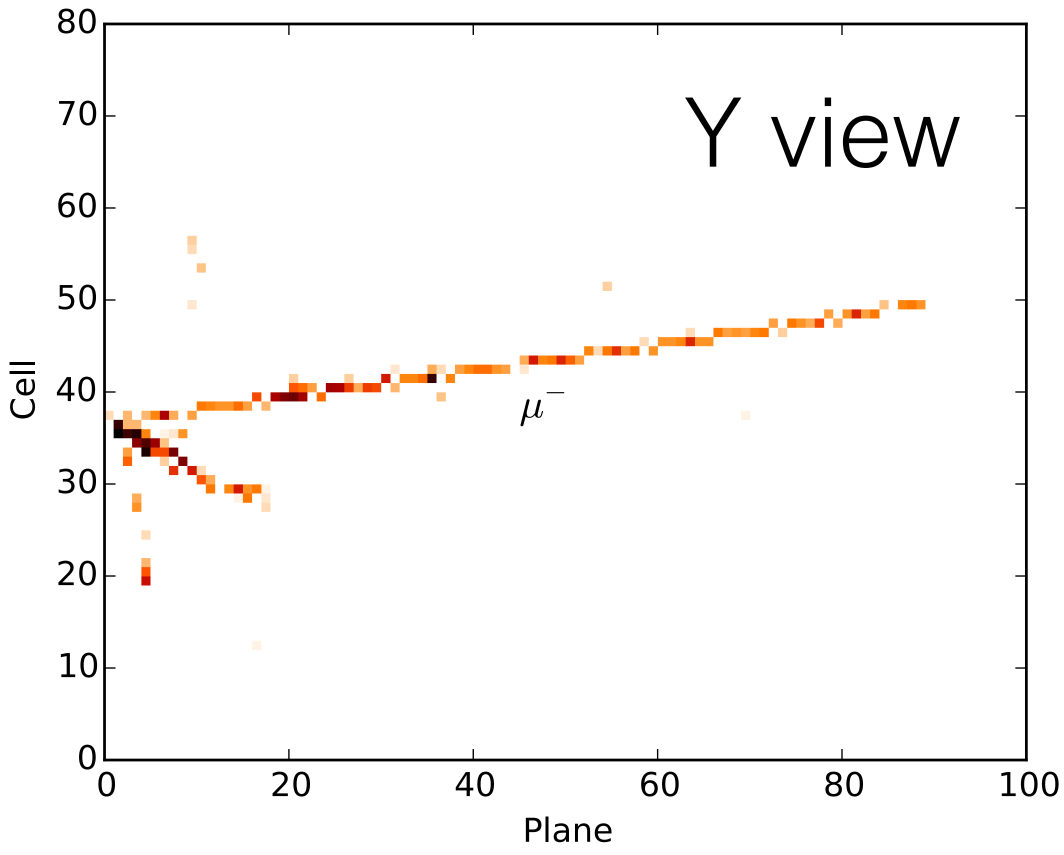
The architecture attempts to categorize events as $\{\nu_\mu, \nu_e, \nu_\tau\} \times \{QE, RES, DIS\}$, NC, or Cosmogenic.

Utilizes a “softmax” output.

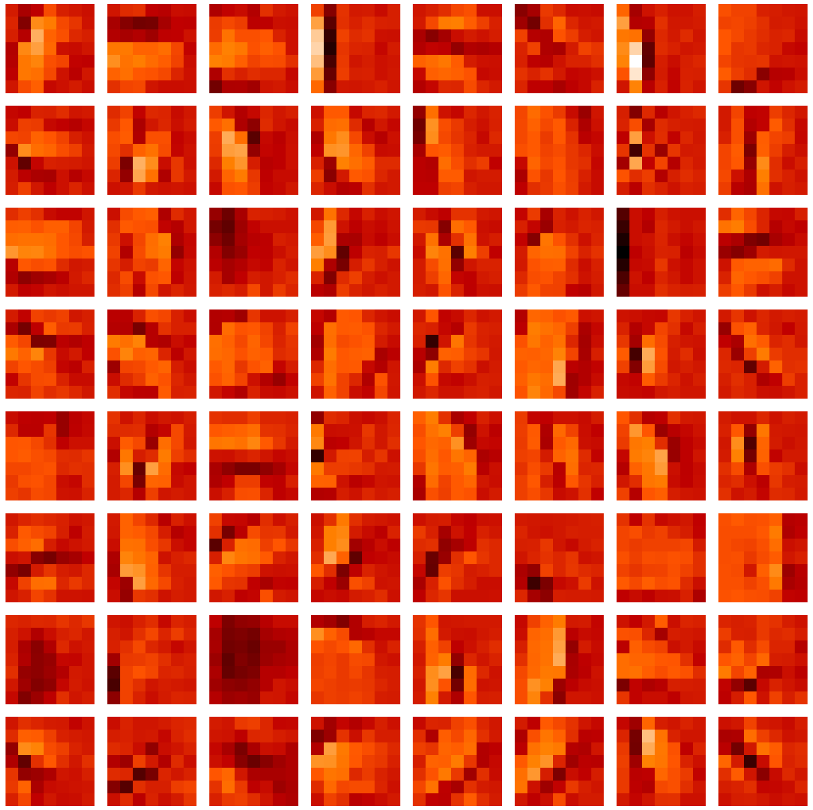
Built in the excellent CAFFE framework: http://caffe.berkeleyvision.org/



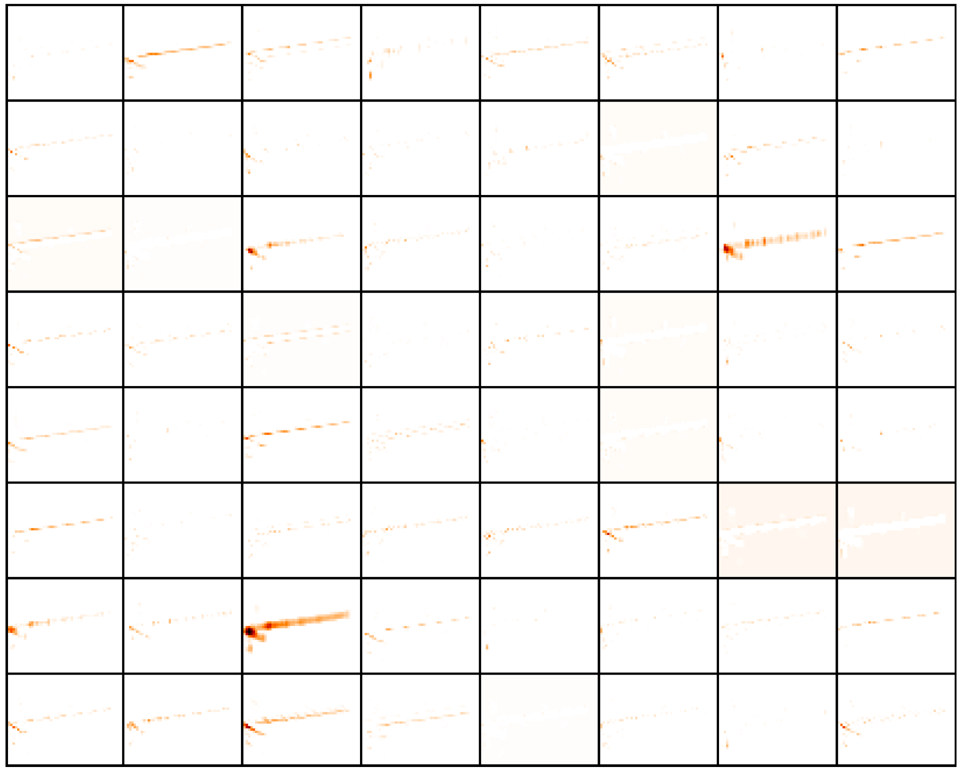
Example CVN Kernels In Action: First Convolution



X



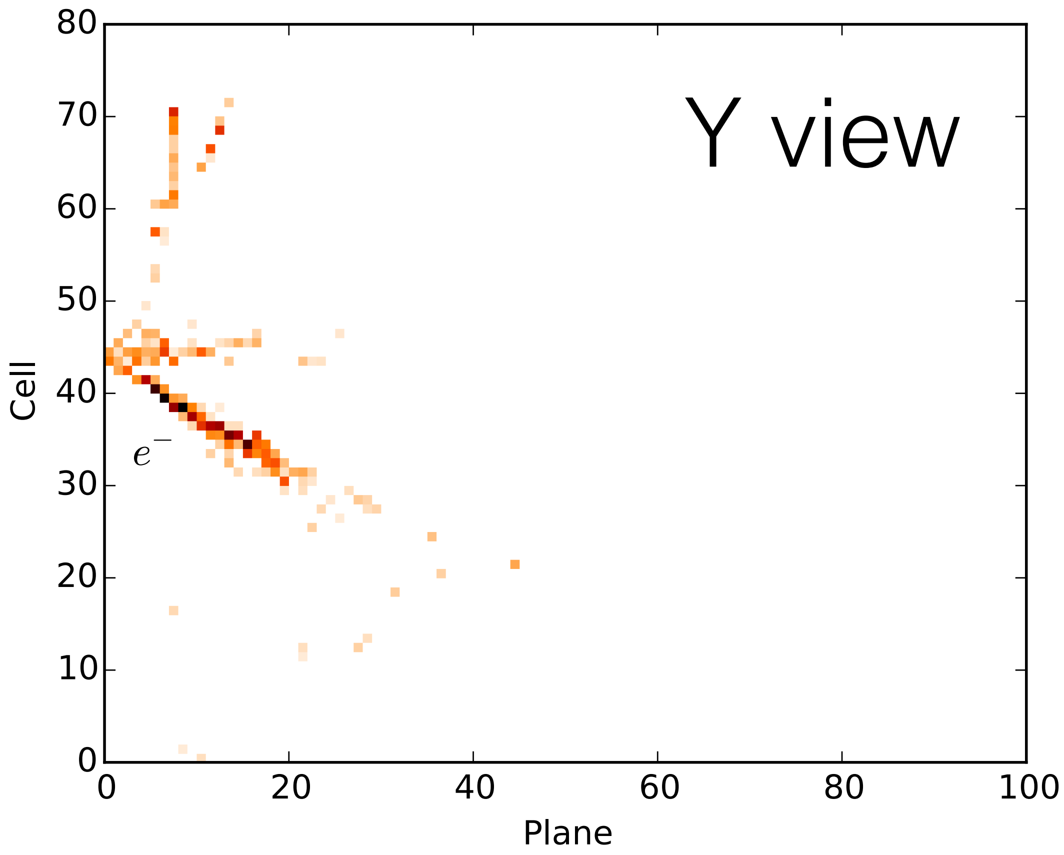
=



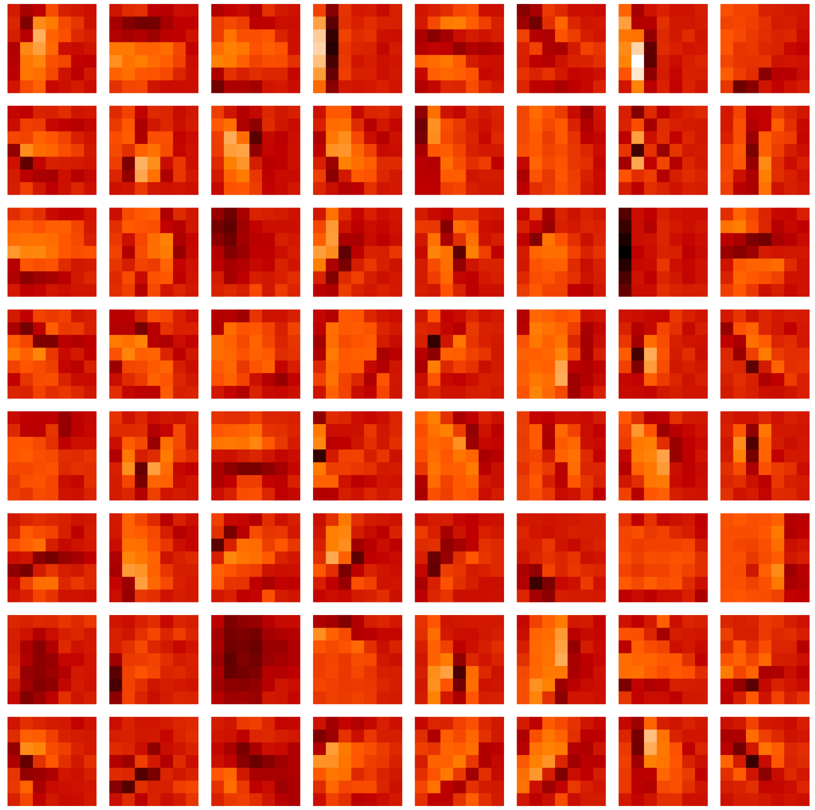
Here the earliest convolutional layer in the network starts by pulling out primitive shapes and lines.

Already "showers" and "tracks" are starting to form.

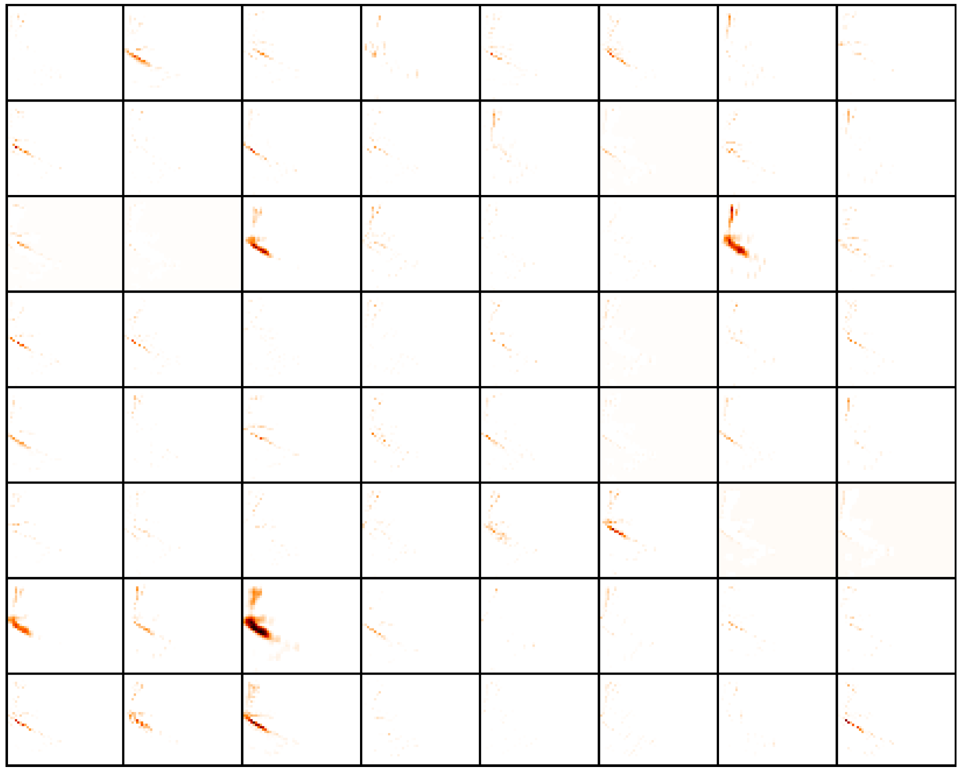
Example CVN Kernels In Action: First Convolution



X



=



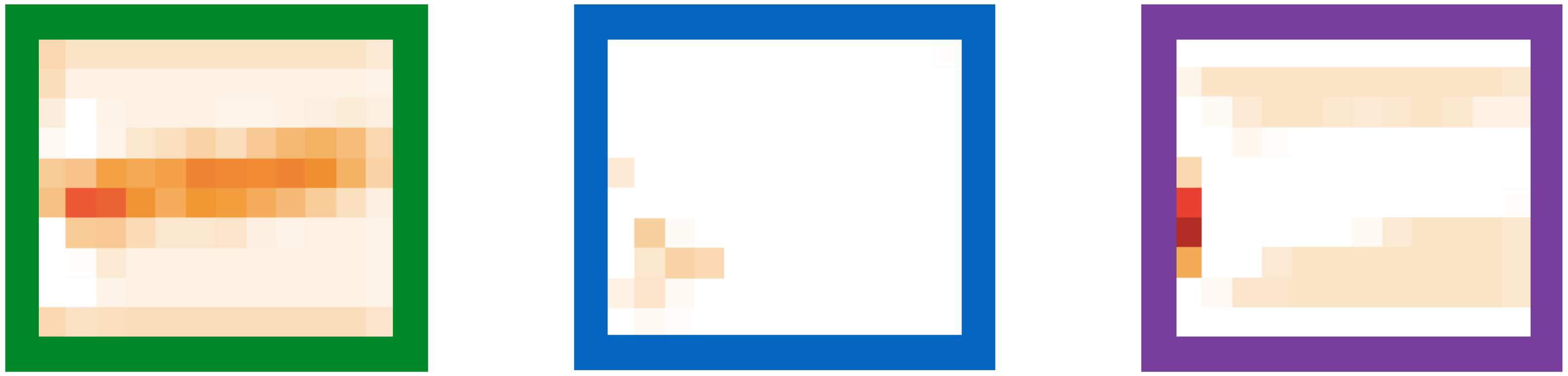
Here the earliest convolutional layer in the network starts by pulling out primitive shapes and lines.

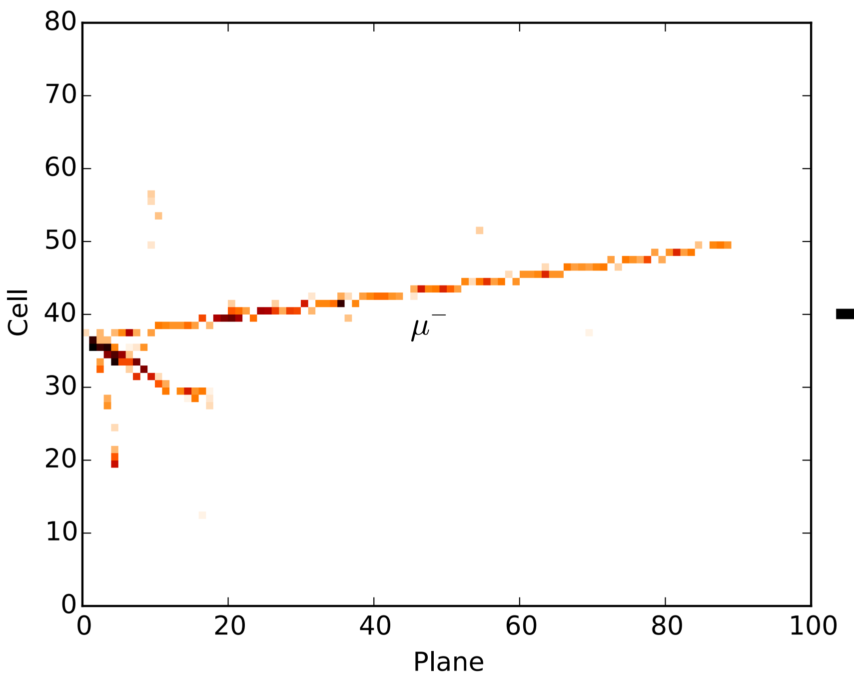
Already "showers" and "tracks" are starting to form.

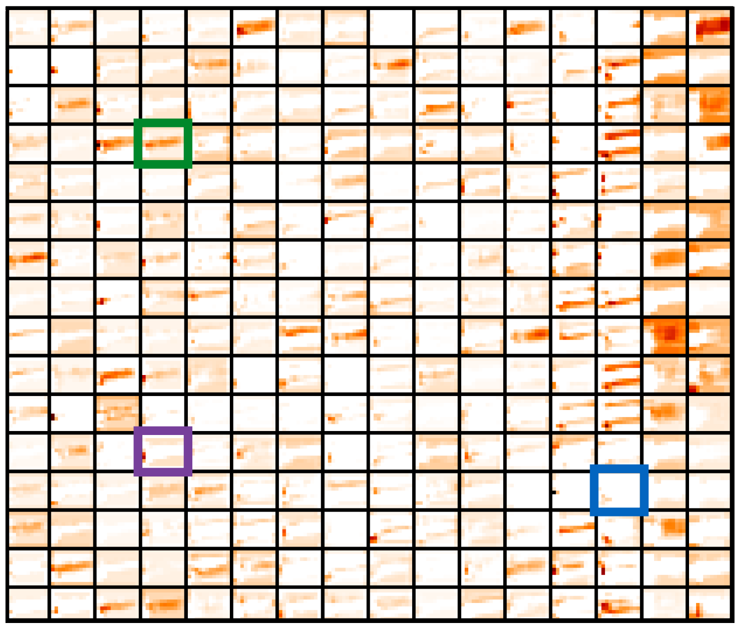
Example CVN Kernels In Action: First Inception Module Output

Deeper in the network, now after the first inception module we can see more complex features have started to be extracted.

Some seem particularly sensitive to muon tracks, EM showers, or hadronic activity.

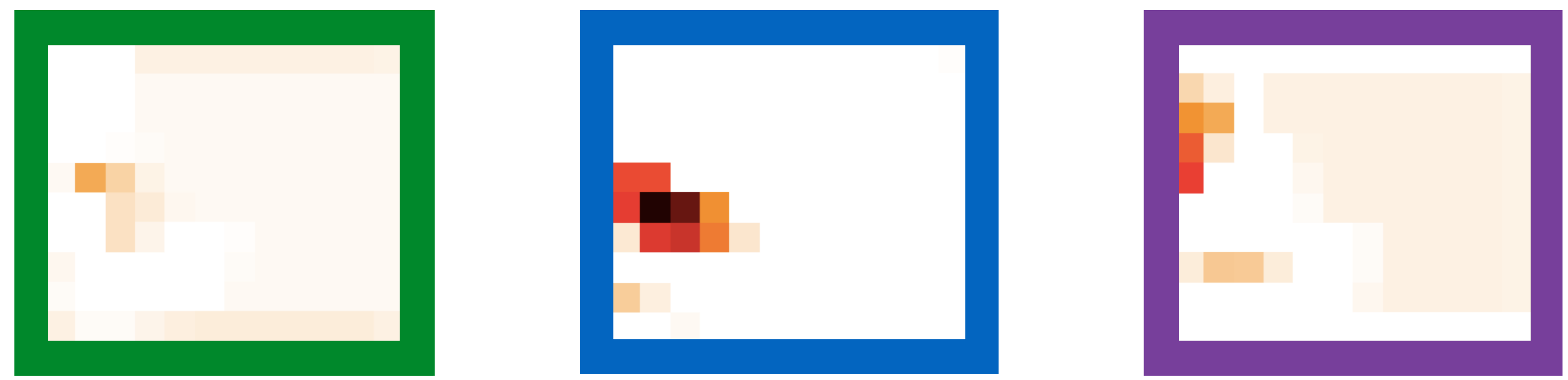


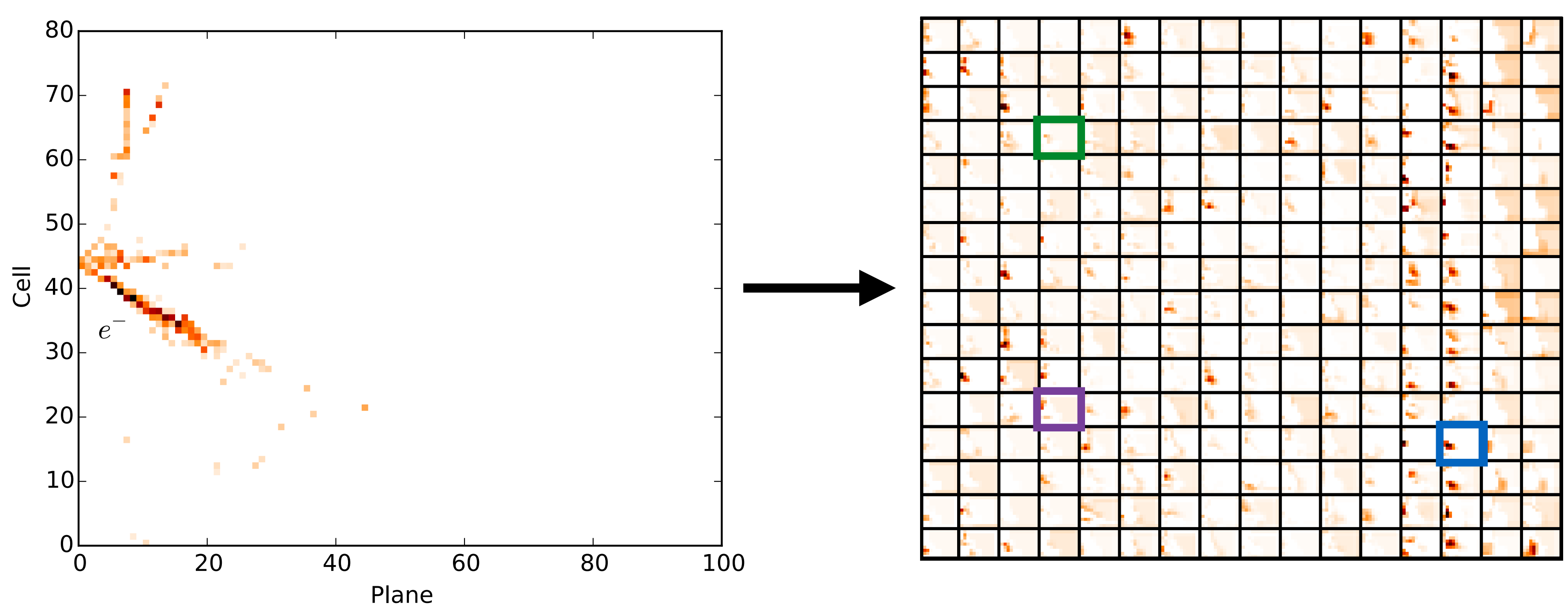




Example CVN Kernels In Action: First Inception Module Output

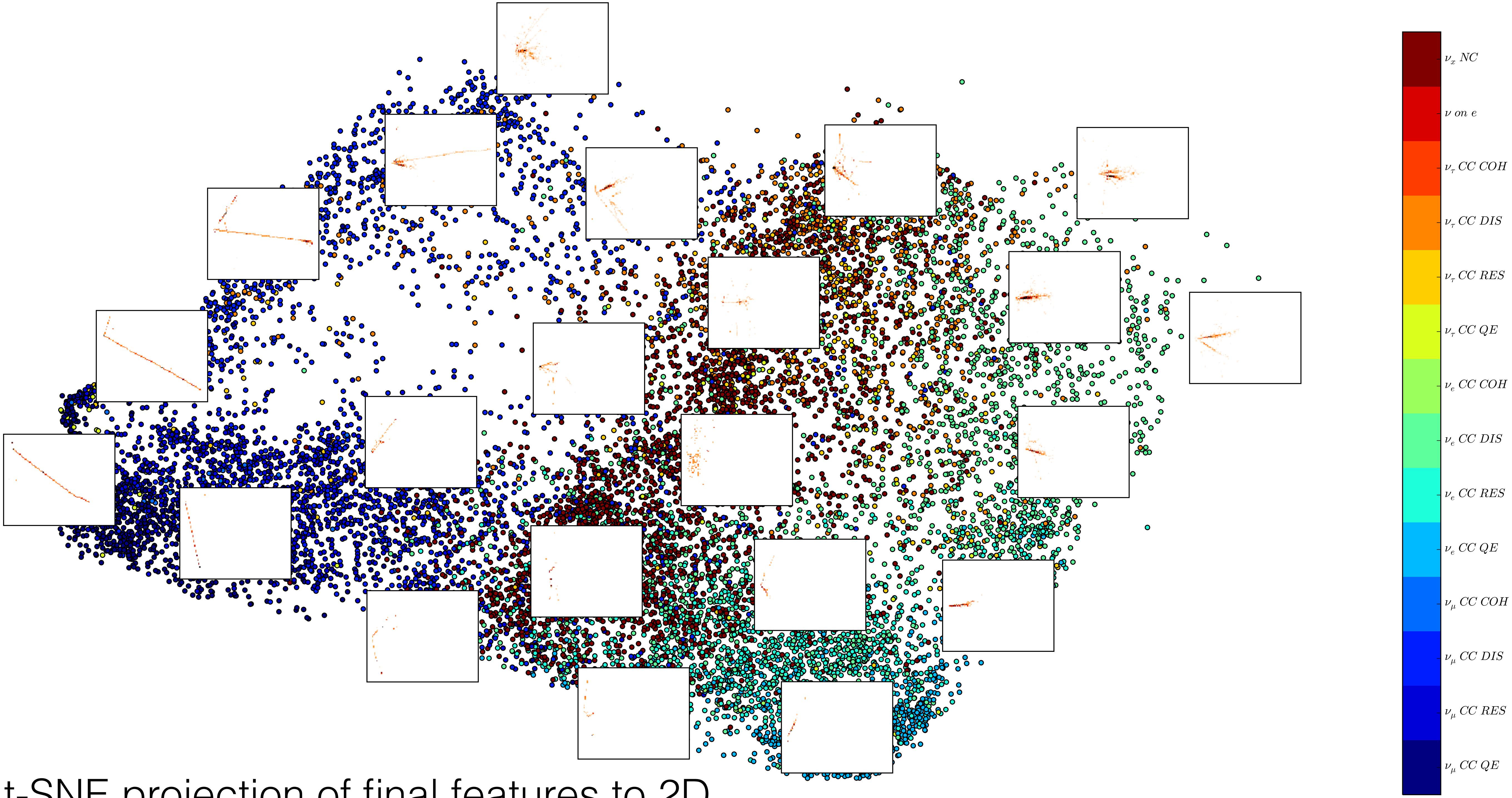
Deeper in the network, now after the first inception module we can see more complex features have started to be extracted.





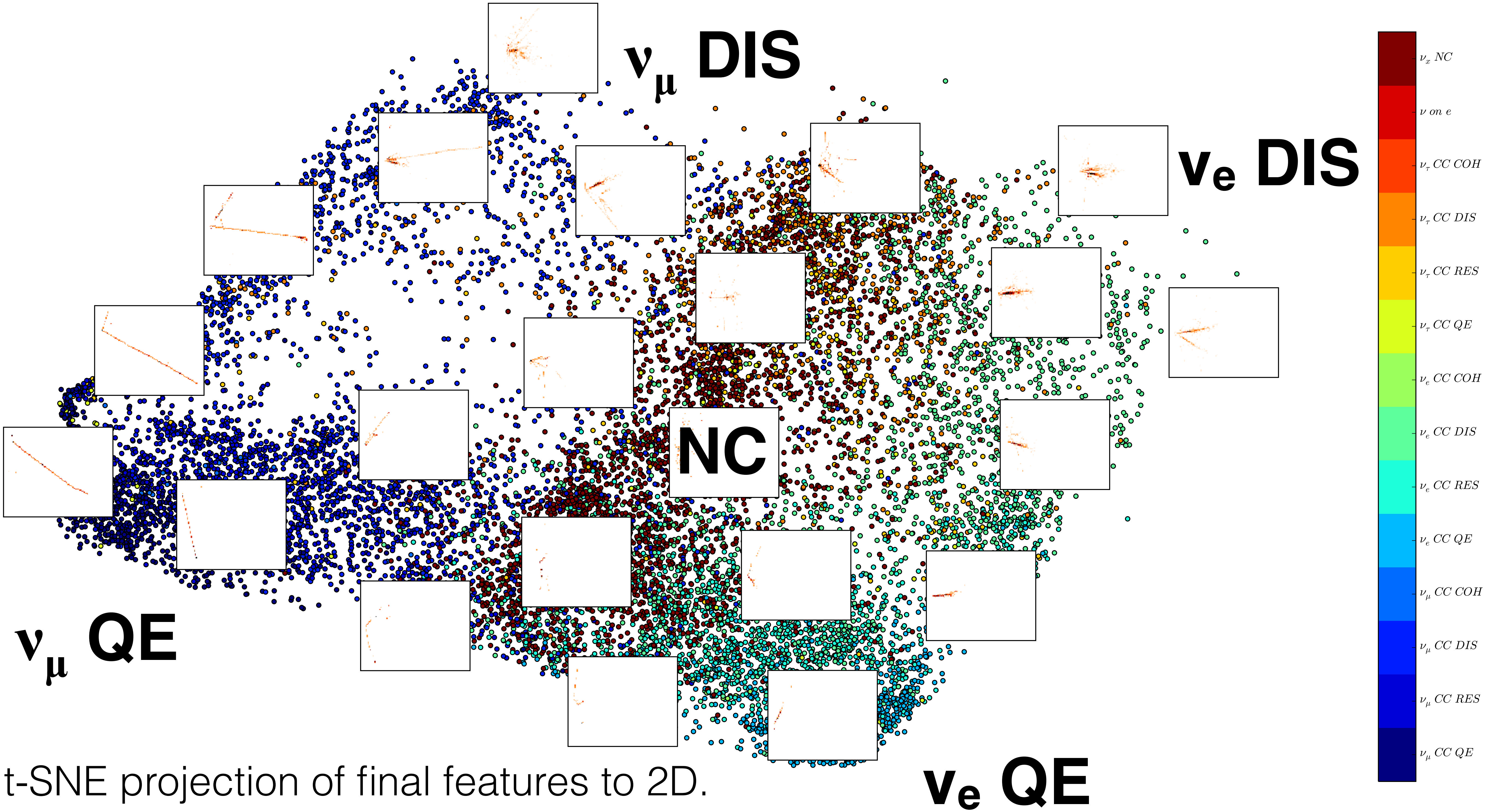
Some seem particularly sensitive to muon tracks, EM showers, or hadronic activity.

t-SNE Representation of Test Sample



t-SNE projection of final features to 2D.
Truth labels, training sample subset.

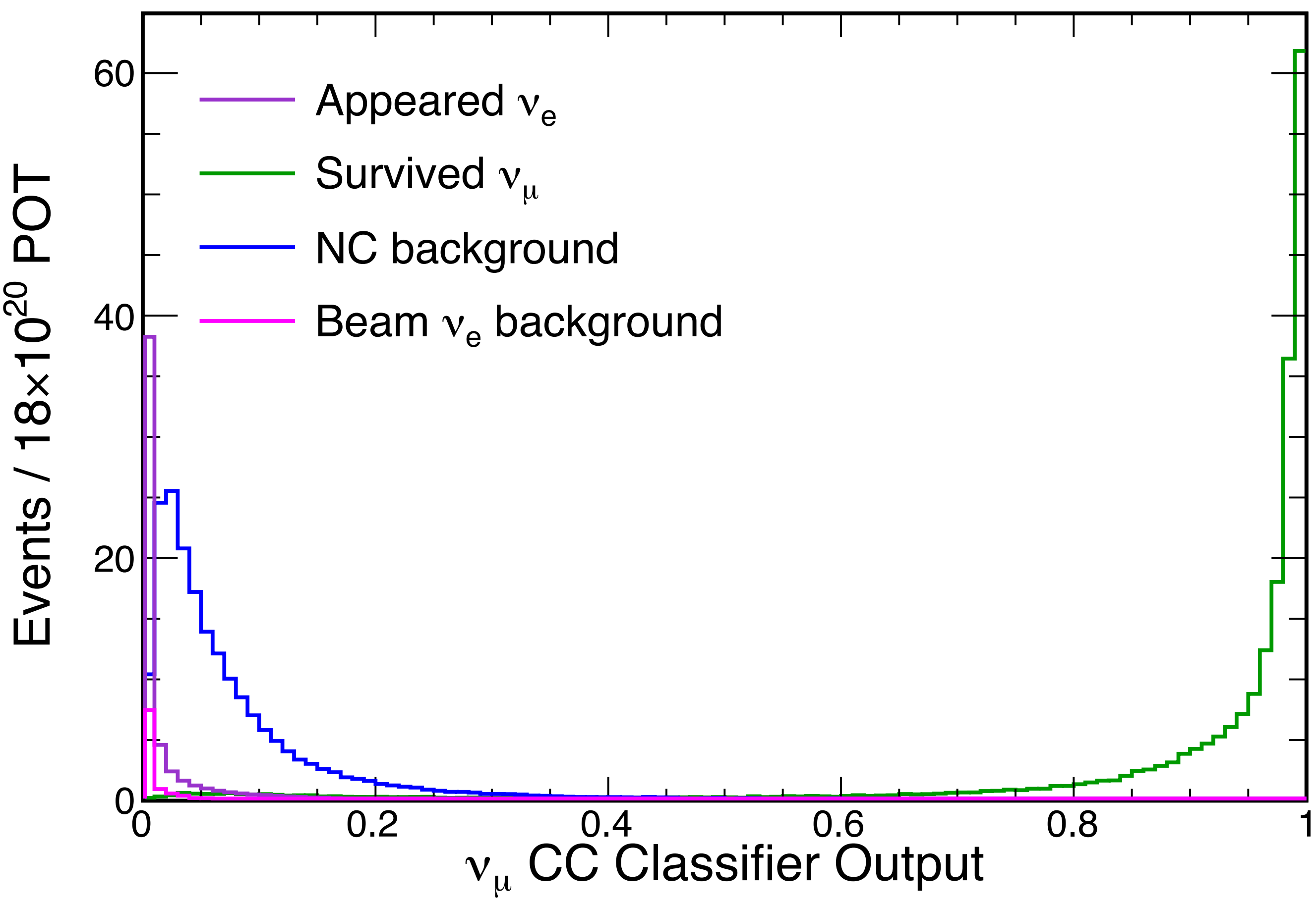
t-SNE Representation of Test Sample

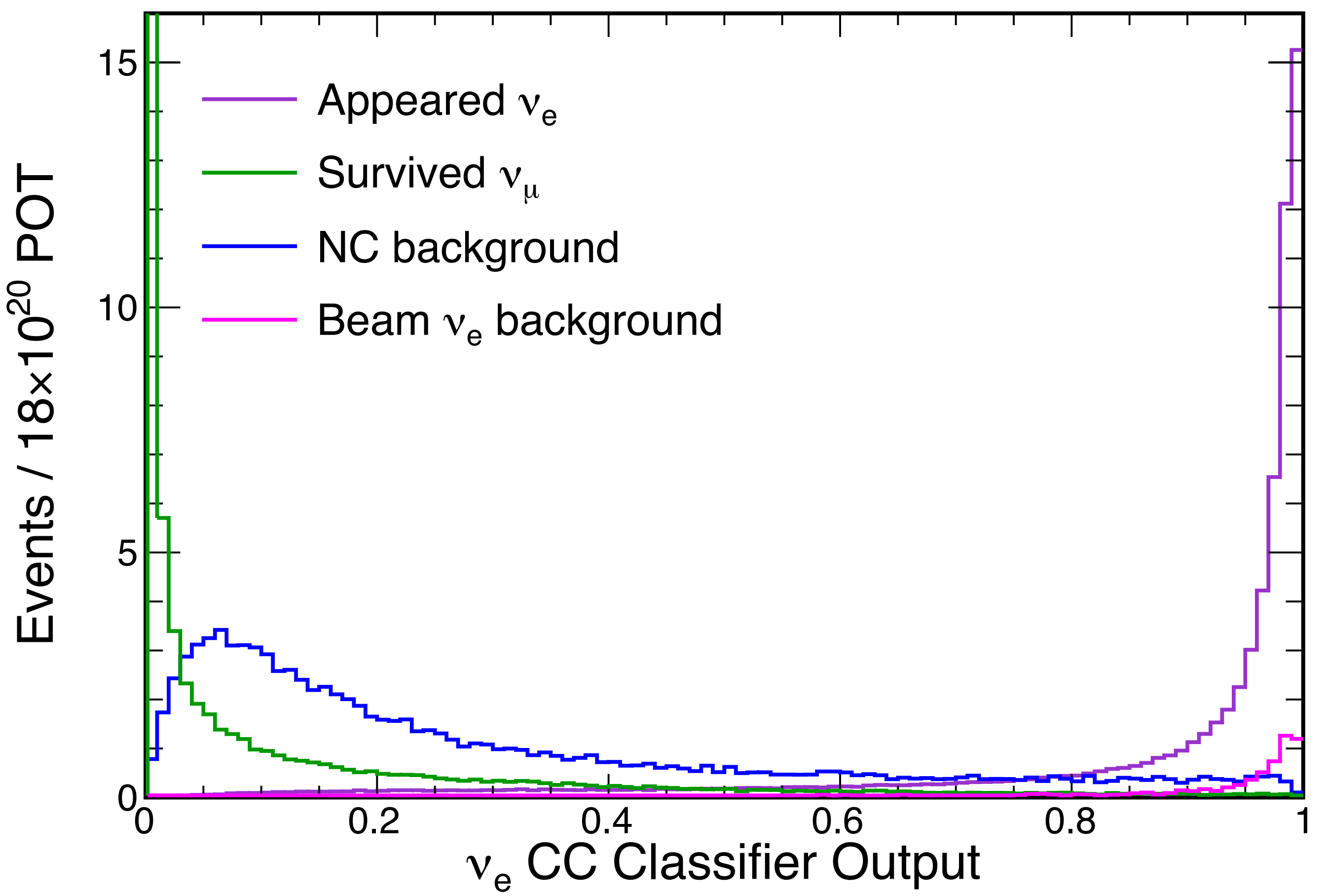


t-SNE projection of final features to 2D.
Truth labels, training sample subset.

The Bottom Line

After oscillations, cosmic rejection cuts, data quality cuts:

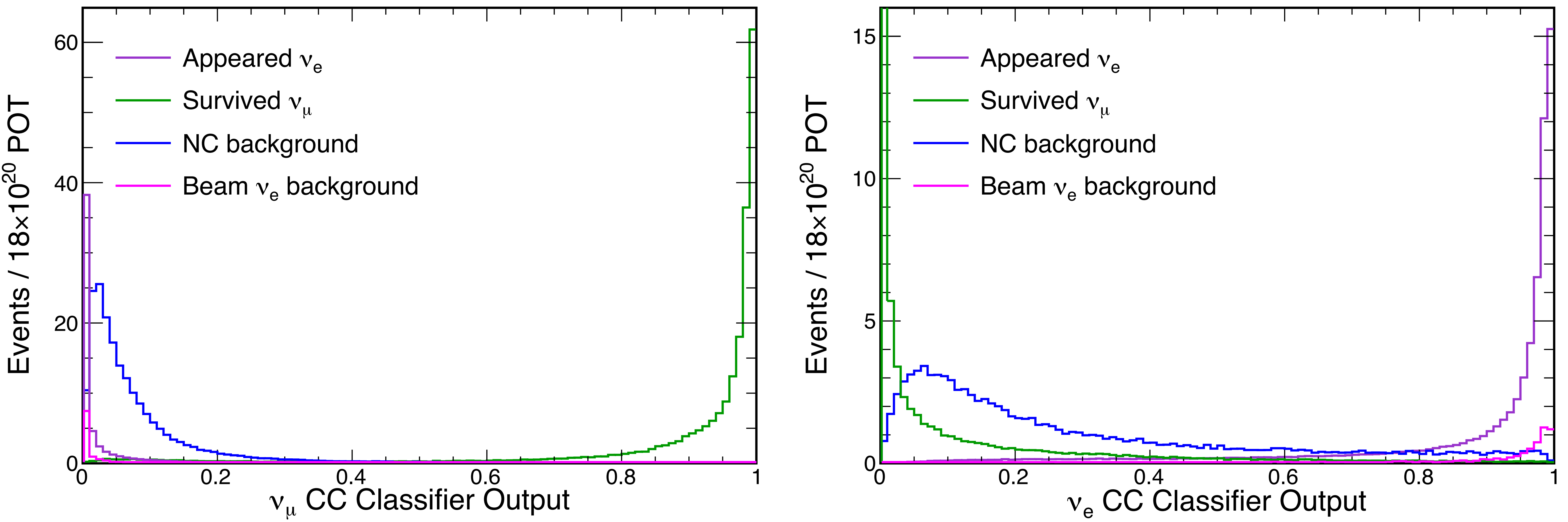




Excellent separation of the $\boldsymbol{\nu_\mu}$ sample, but ~identical to existing, much simpler, KNN selector. Matches expectation- hard to miss a muon track. Space to improve in cosmic rejection.

The Bottom Line

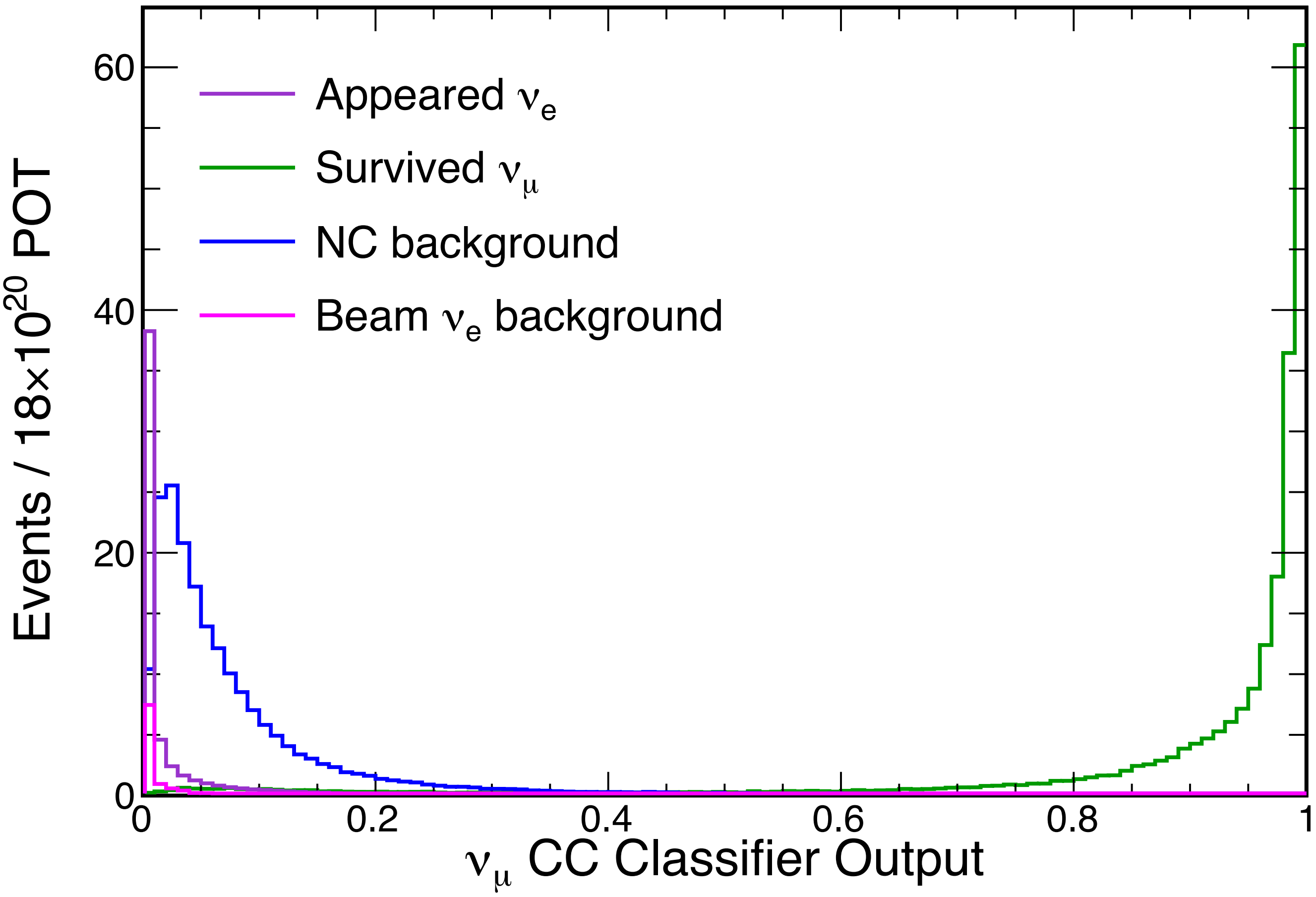
After oscillations, cosmic rejection cuts, data quality cuts:

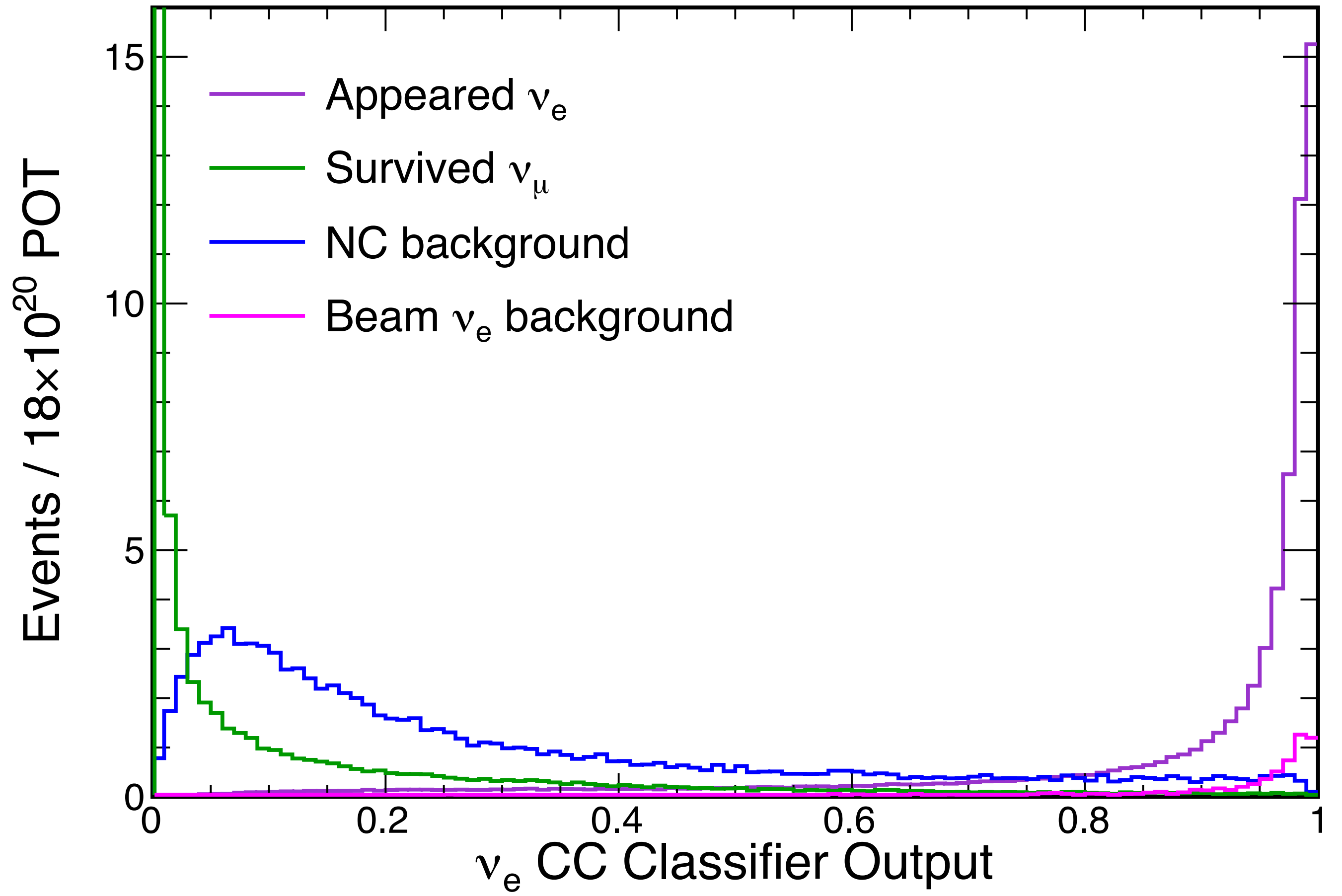


However our CNN achieves **73%** efficiency and **76%** purity on **ν_e** selection at the $s/\sqrt{s+b}$ optimized cut.
Equivalent to **30%** more exposure with the old PIDs.

The Bottom Line

After oscillations, cosmic rejection cuts, data quality cuts:

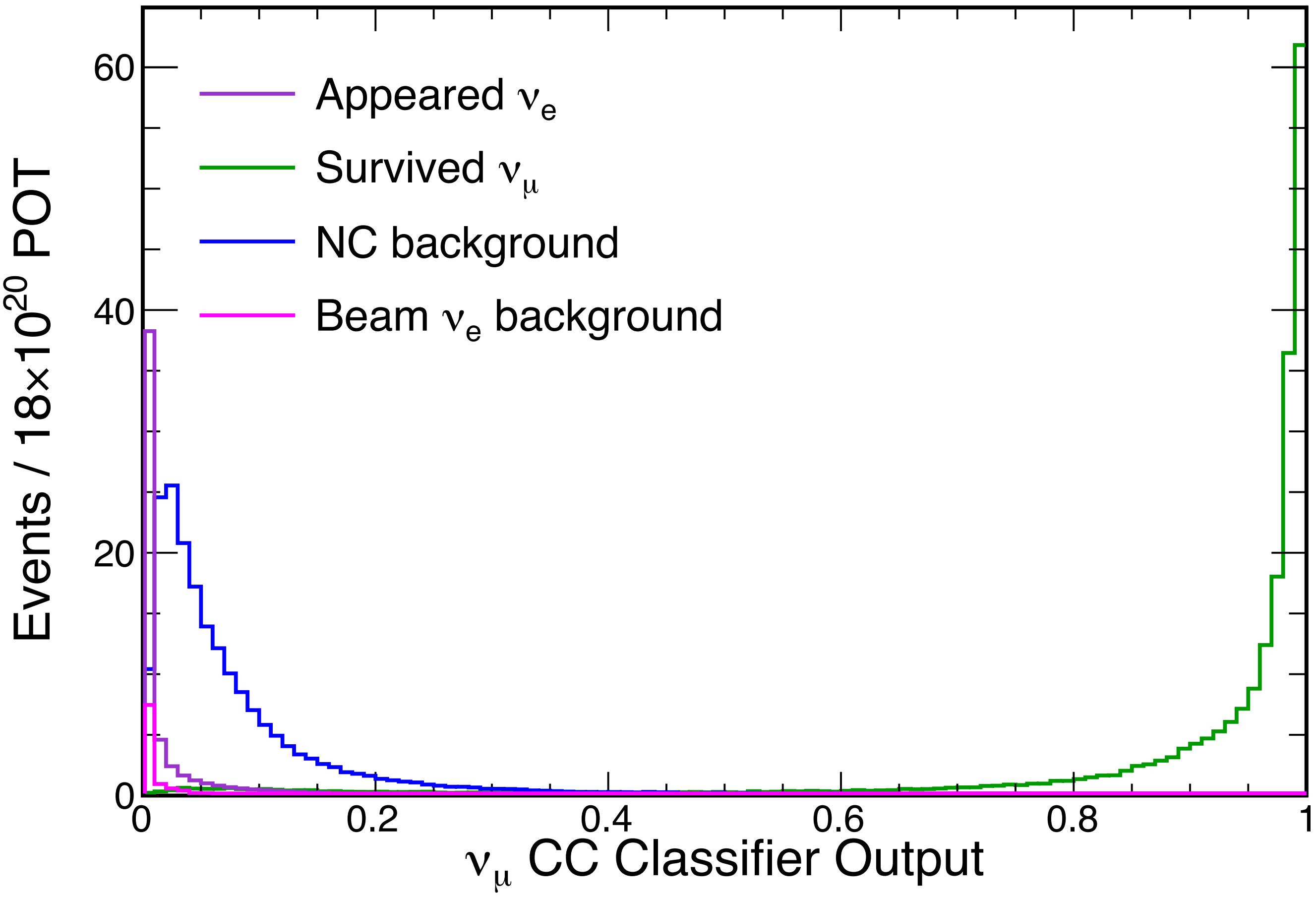


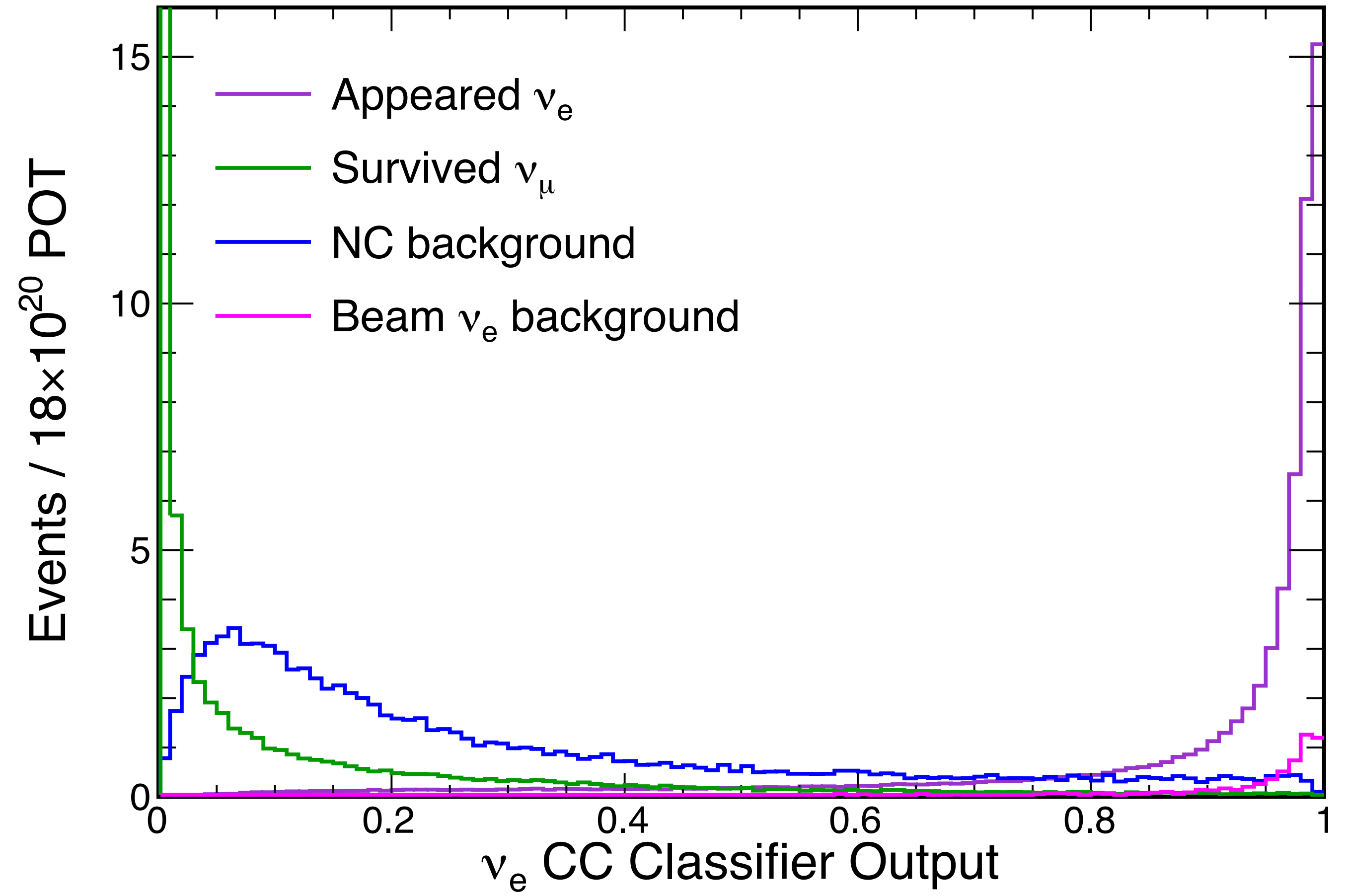


"A Convolutional Neural Network Neutrino Event Classifier"
A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, P. Vahle
https://arxiv.org/abs/1604.01444
Journal of Instrumentation, Volume 11, September 2016

The Bottom Line

After oscillations, cosmic rejection cuts, data quality cuts:



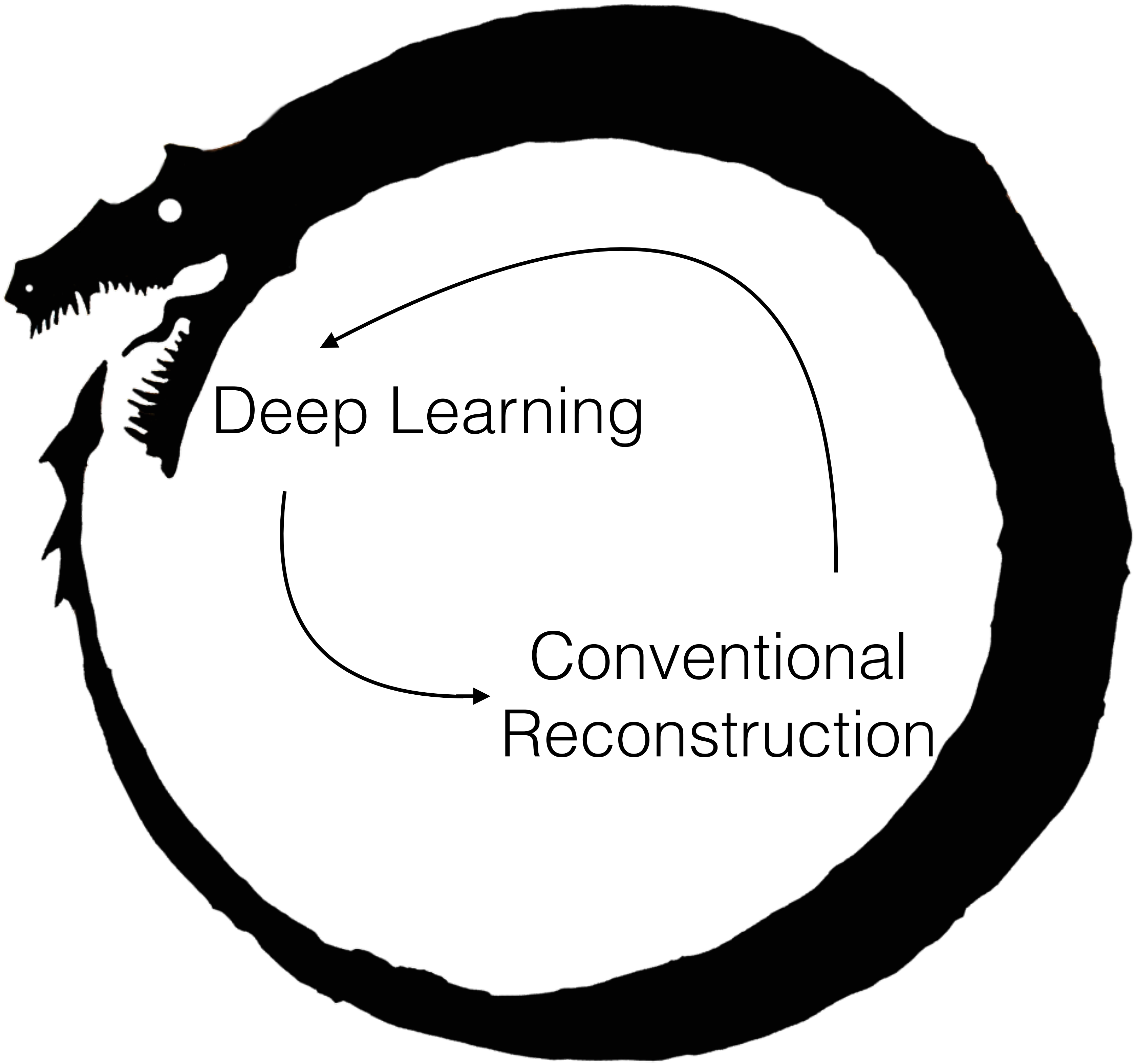


"Constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in NOvA"
NOvA Collaboration
https://arxiv.org/pdf/1703.03328.pdf
Submitted to PRL

The Future

The original dream

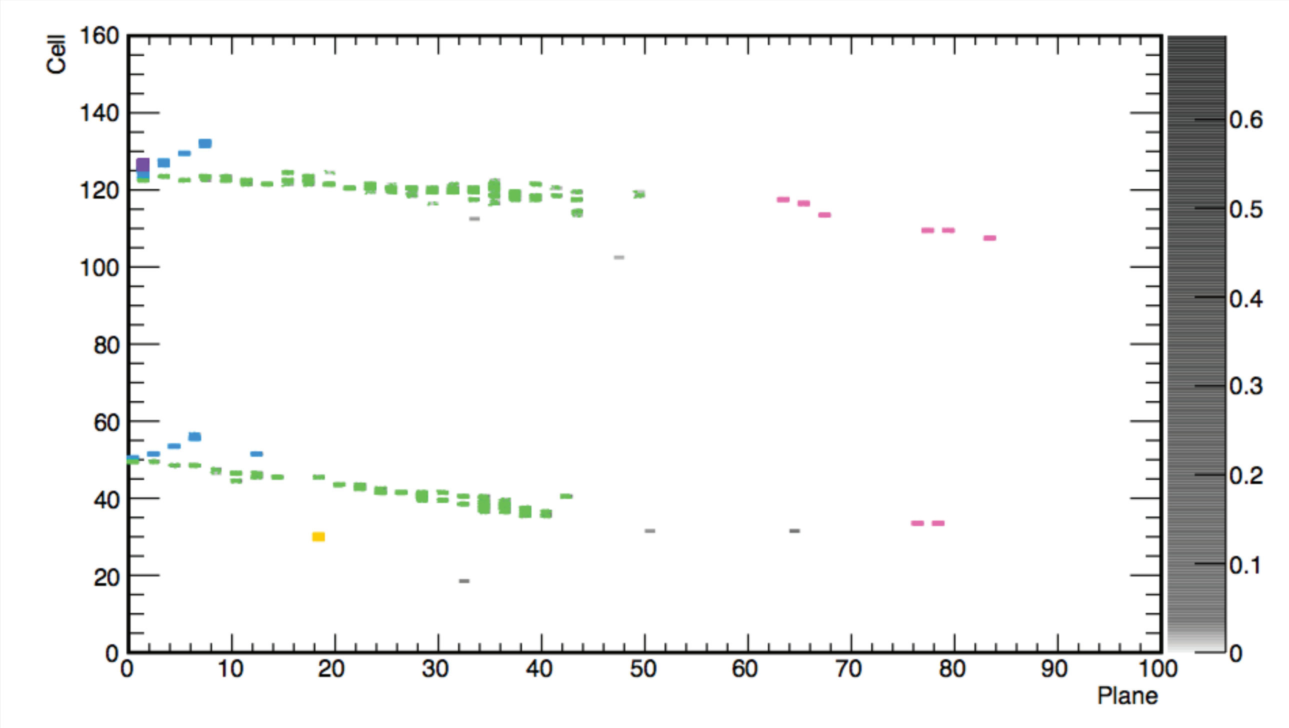
Where we're really going



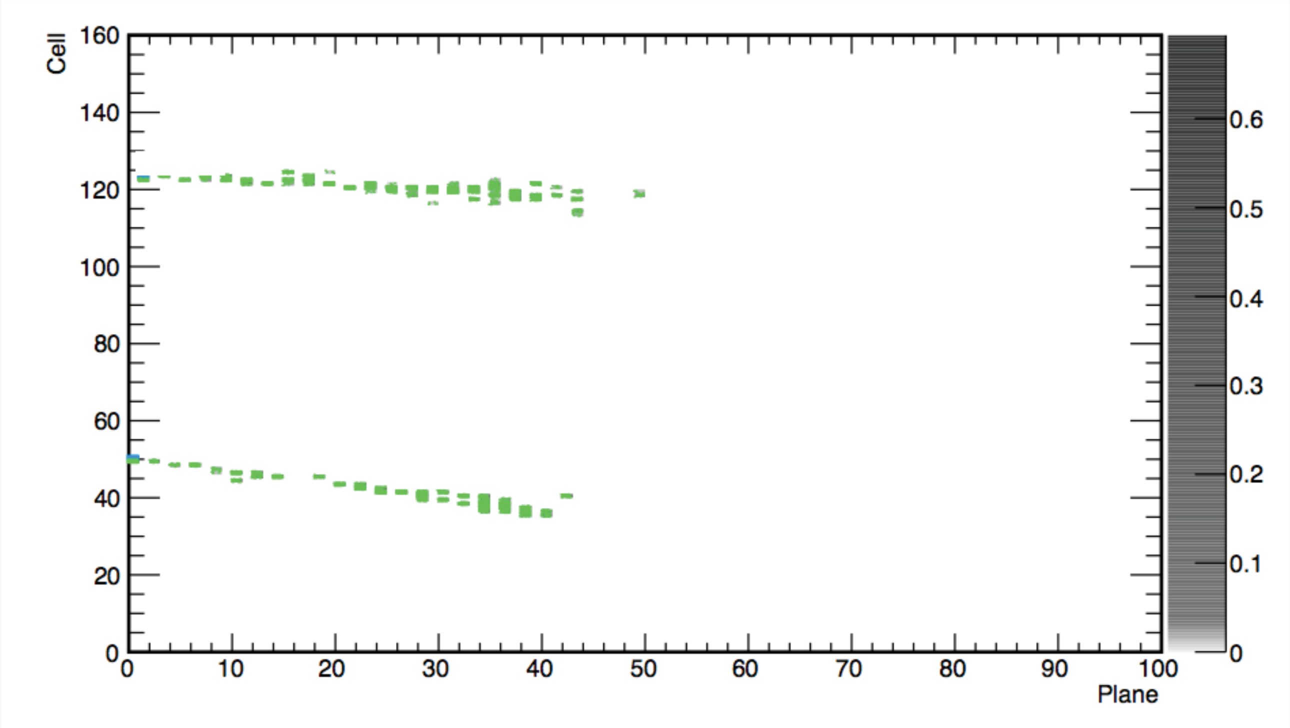
Prong ID

Can we use CVN to ID our reconstructed objects, like showers?

CVN Input: 4 views



2 complete views

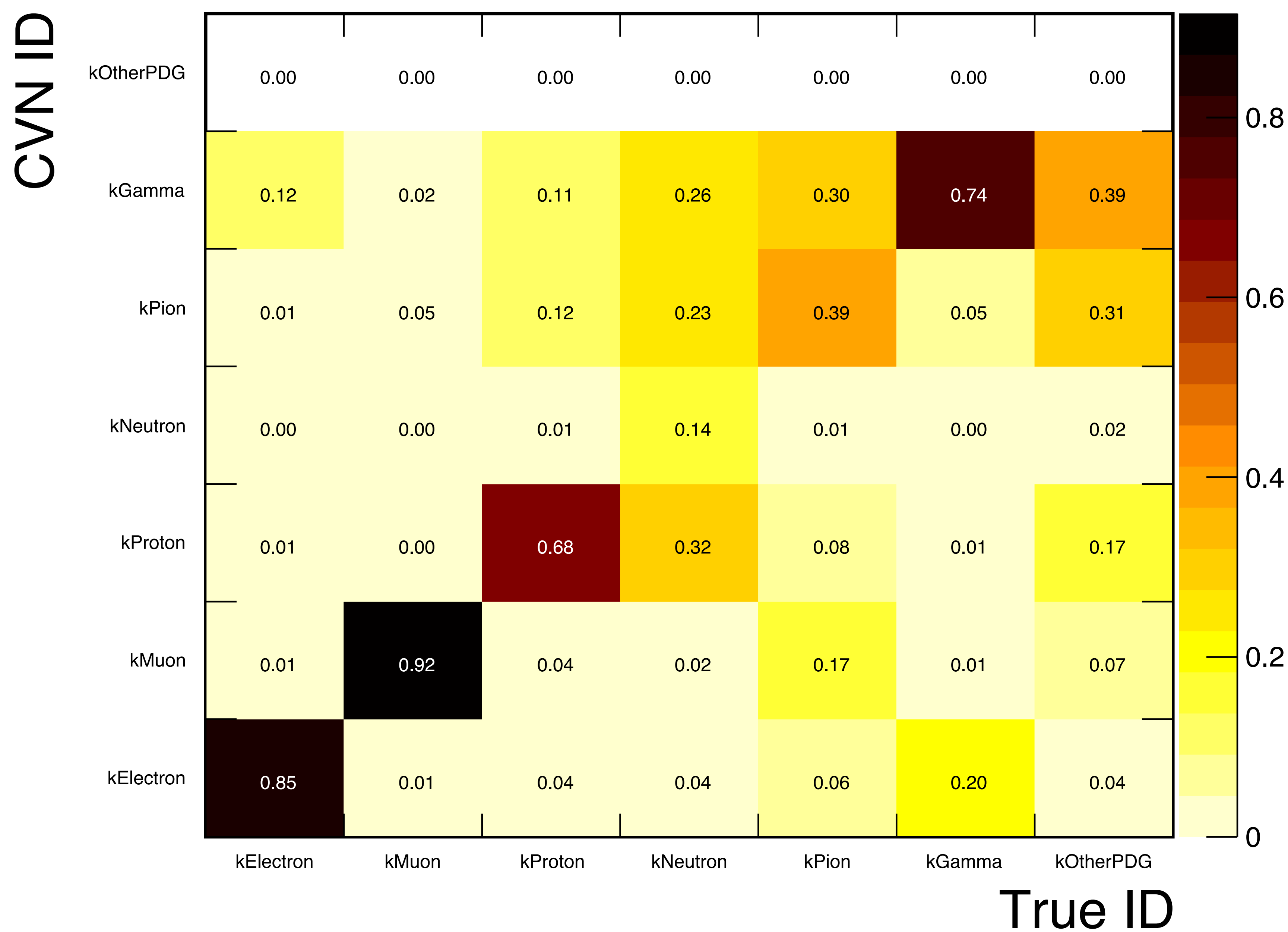


2 prong views

Prong ID

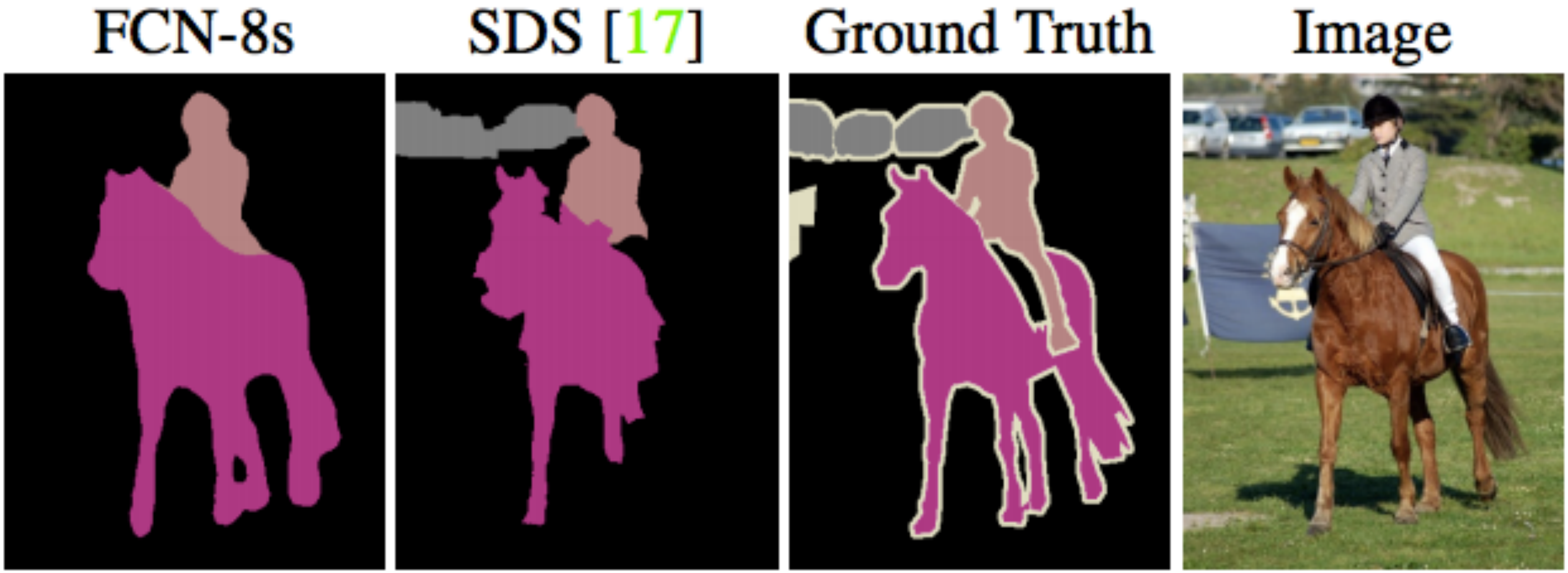
Very promising! Why stop at reconstructed showers though?

Confusion Matrix (prong purity > 0.50)



Semantic Segmentation

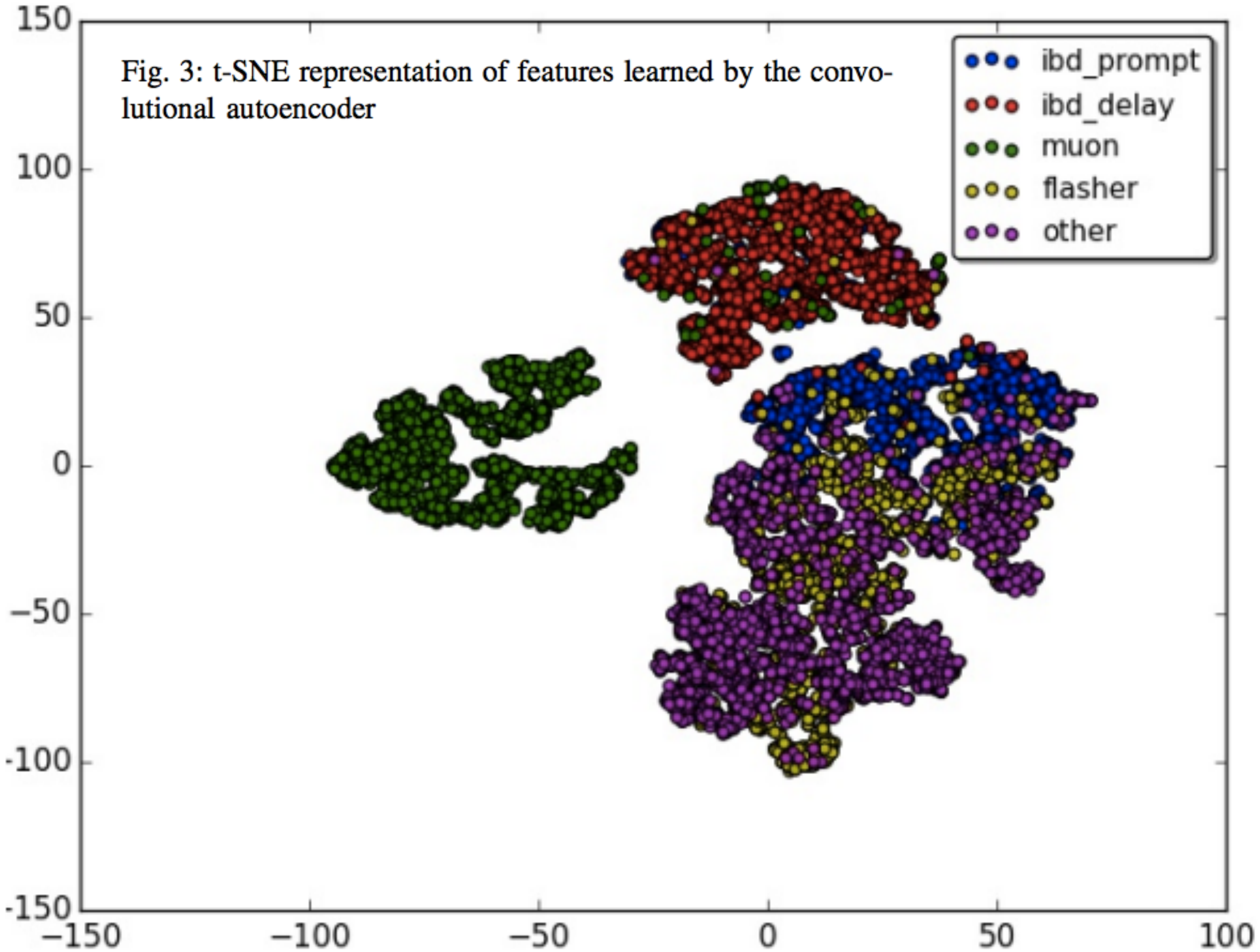
Semantic segmentation takes advantage of information at every lay of a CNN to perform a identification at the pixel level.



http://www.cs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf

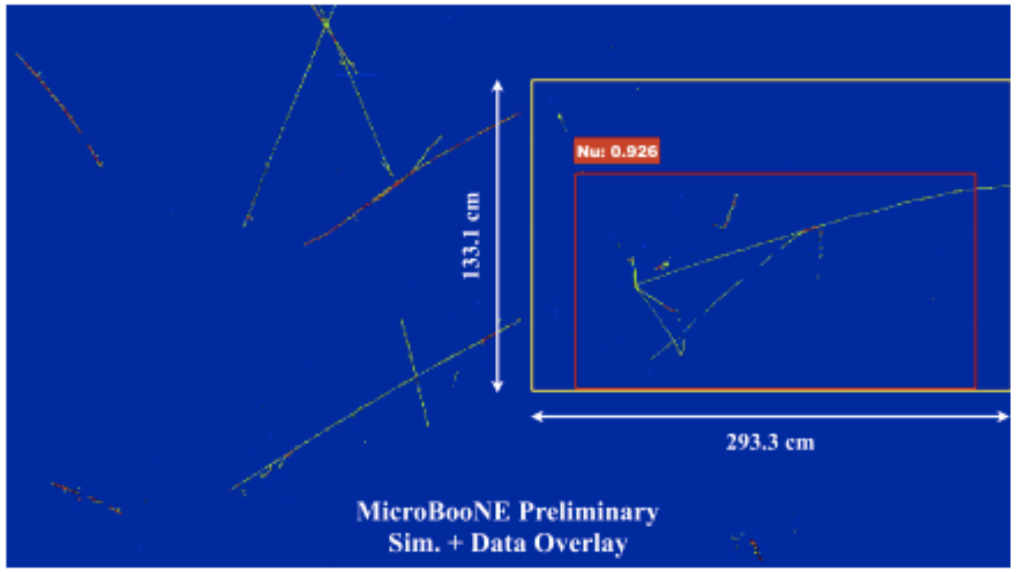
An Active Field!

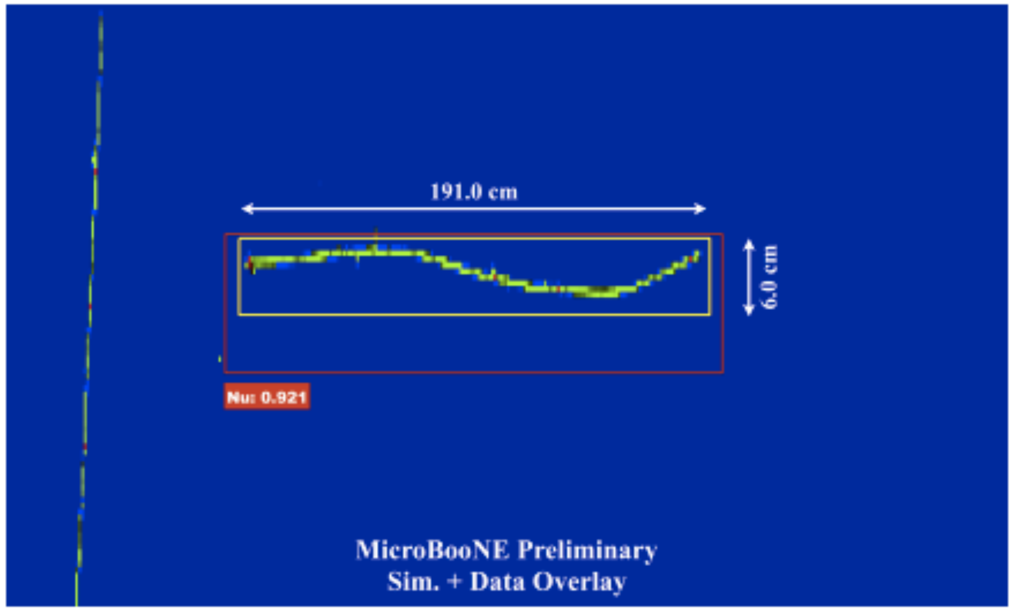
Daya Bay



arXiv:1601.07621

MicroBooNE

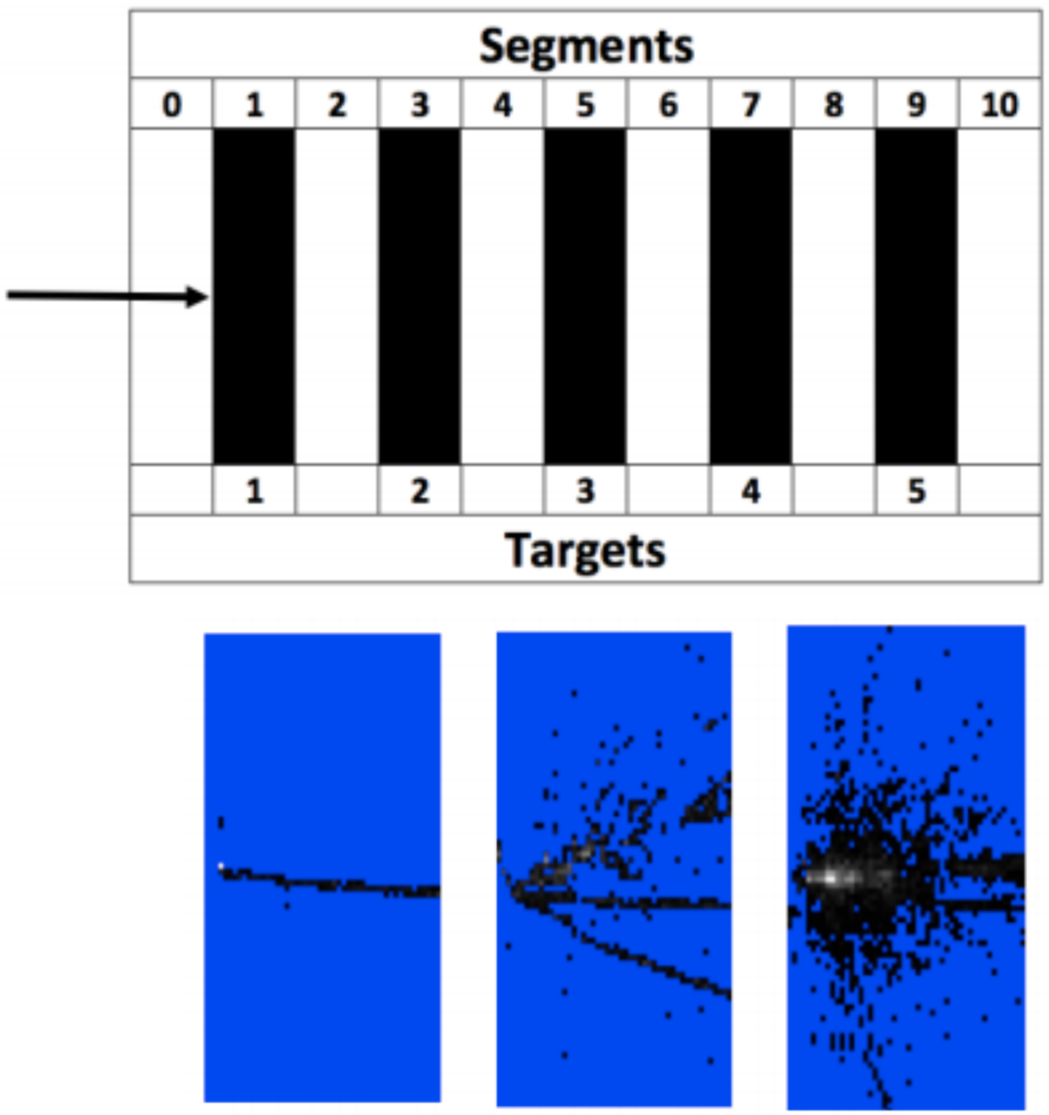




JINST 12 (2017) no.03, P03011

An Active Field!

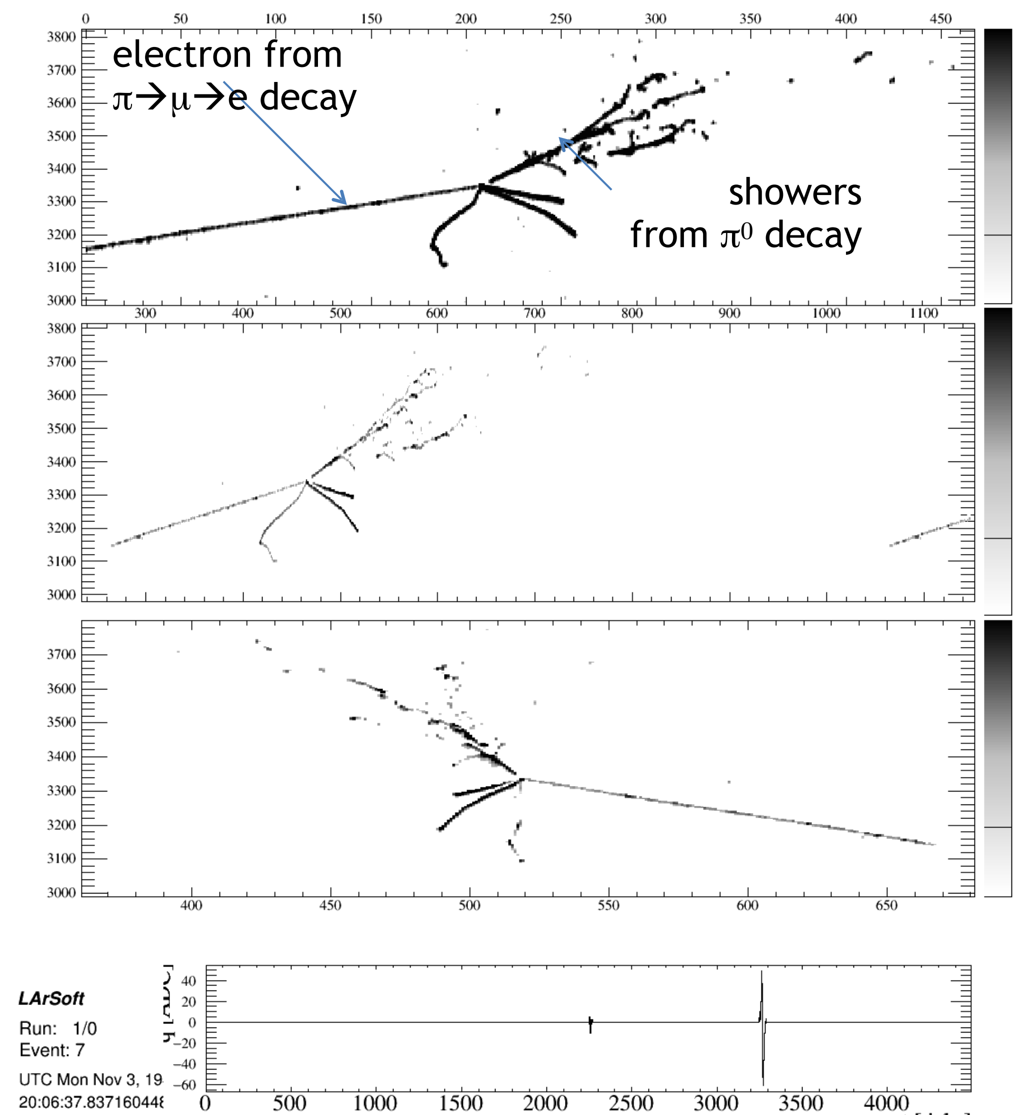
MINERvA and MENNDL



PhyStat-nu Fermilab 2016 (19-September 21, 2016)

Lariat and ProtoDUNE

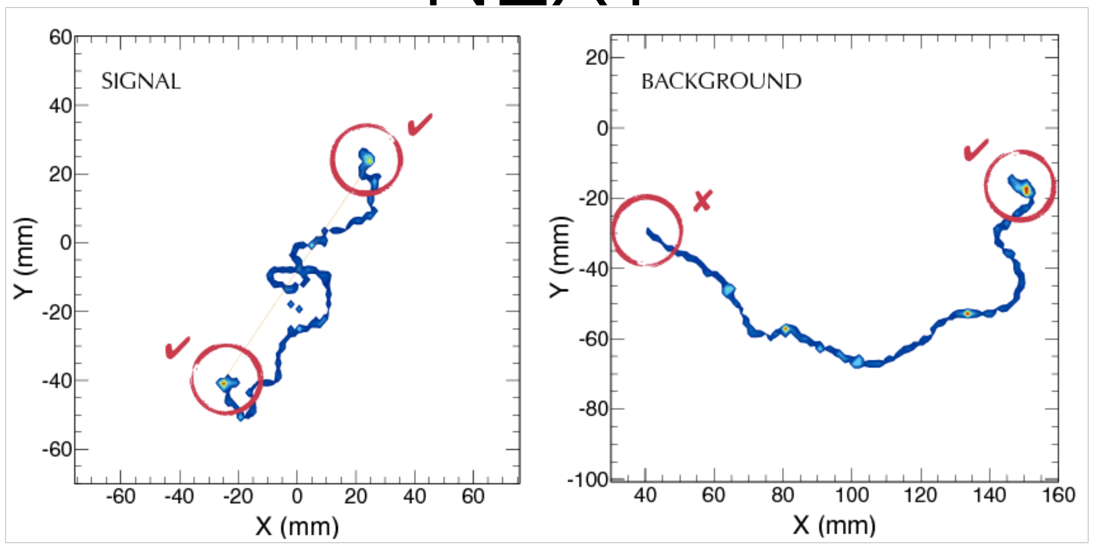
EM / hadronic component discrimination

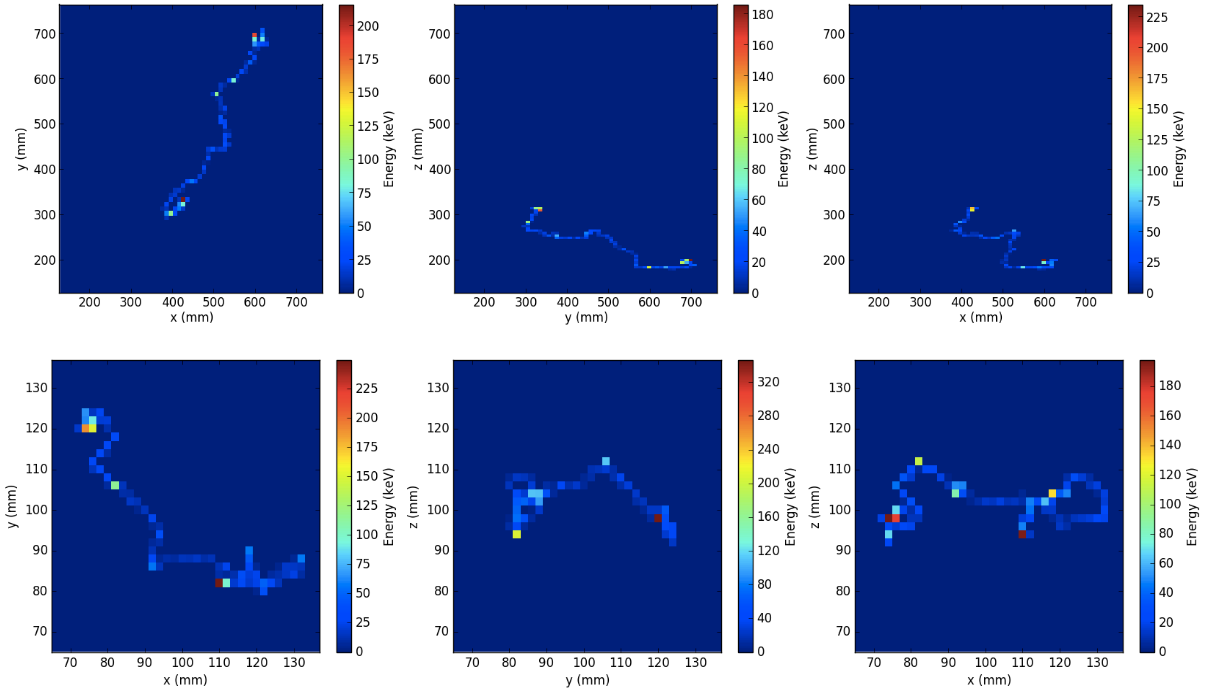


Private communication, Robert Sutlej

An Active Field!

NEXT





JINST 12 (2017) no.01, T01004

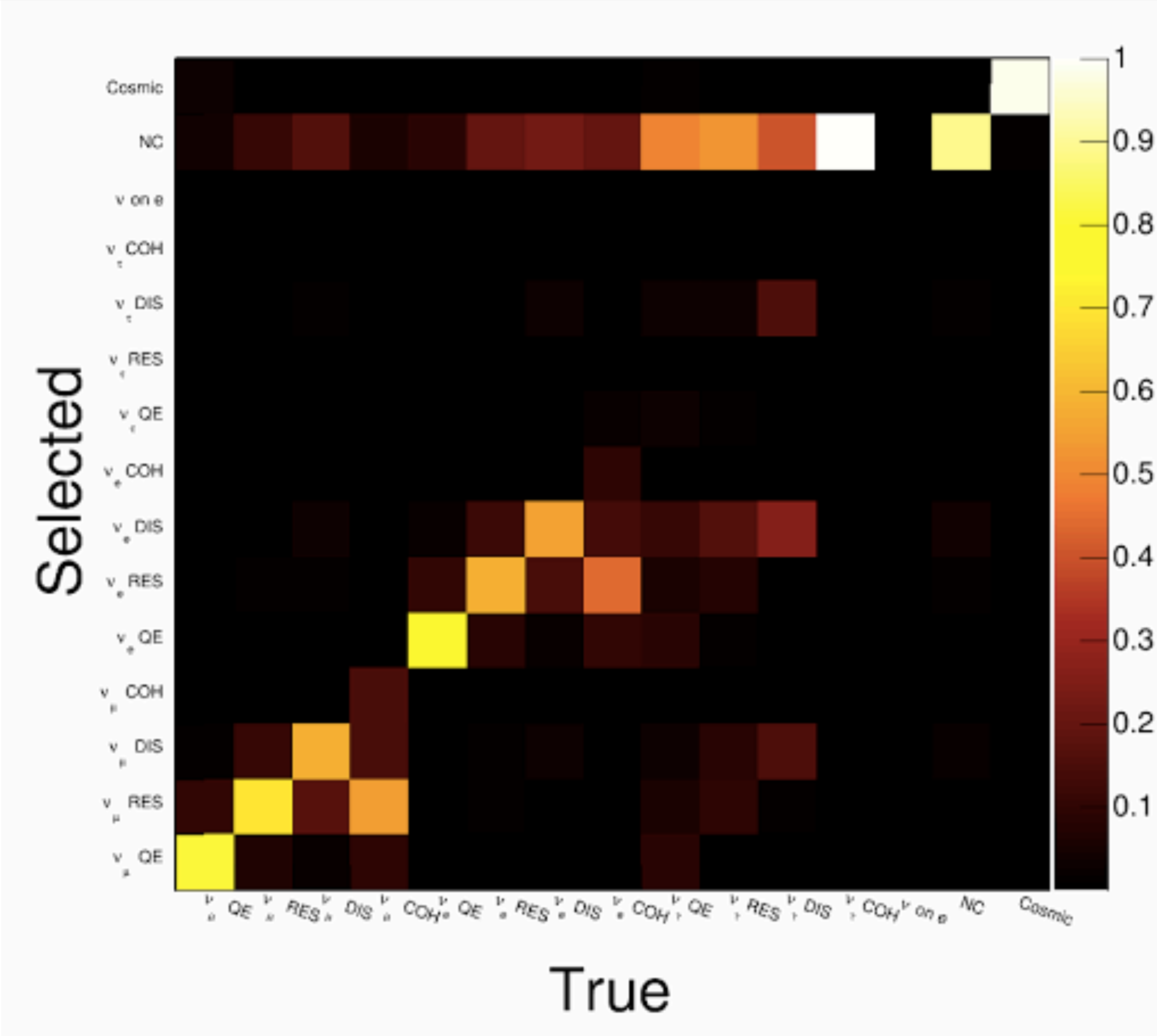
Conclusions

- The first high energy particle physics measurement to use a convolutional neural network learning was a neutrino oscillation measurement from the NOvA experiment!
- Just the tip of the iceberg at NOvA! Huge amounts of room to optimize our classification network, and to explore other applications of convolutional neural networks.
- Almost every neutrino experiment seems to be investigating the use of CNNs in their analysis and reconstruction of neutrino interactions, set to be particularly vital for liquid argon detectors which are a core part of the planned DUNE LBL oscillation experiment.

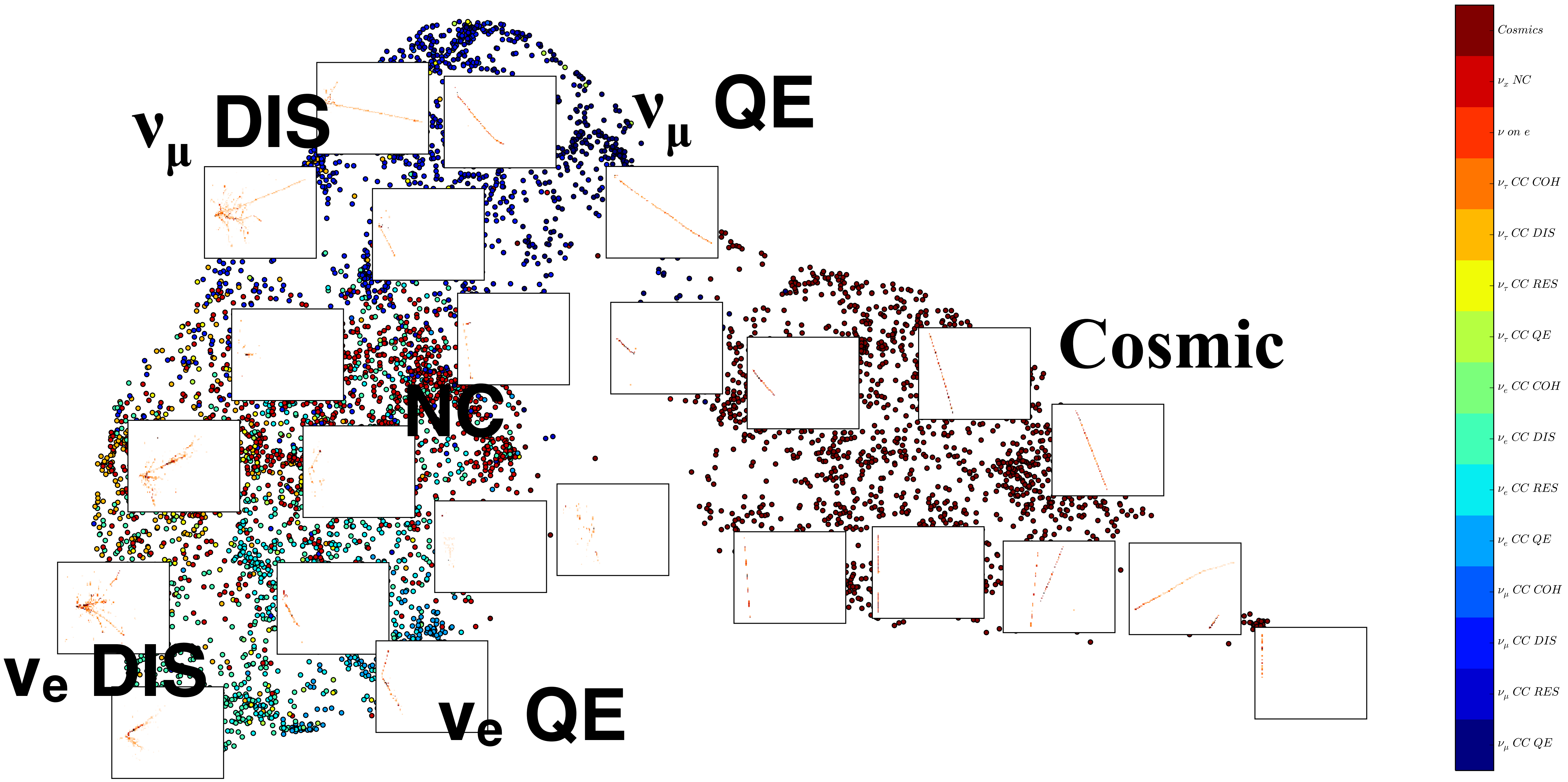
Q&A

Many thanks to the NOvA collaboration, Fermilab National Accelerator laboratory, and to the National Science Foundation.

Confusion Matrix

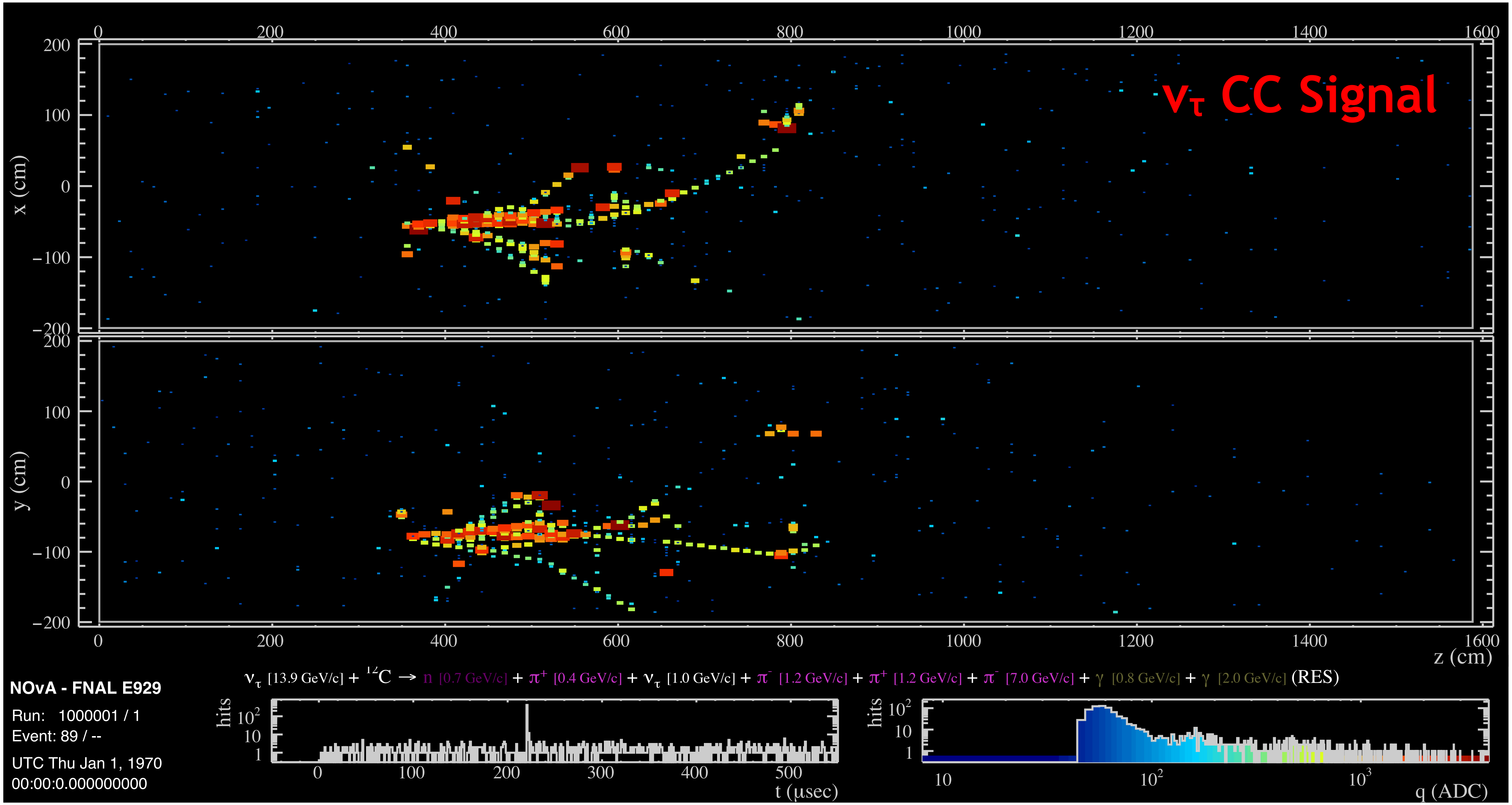


t-SNE Representation of Test Sample



Truth labels, training sample subset.

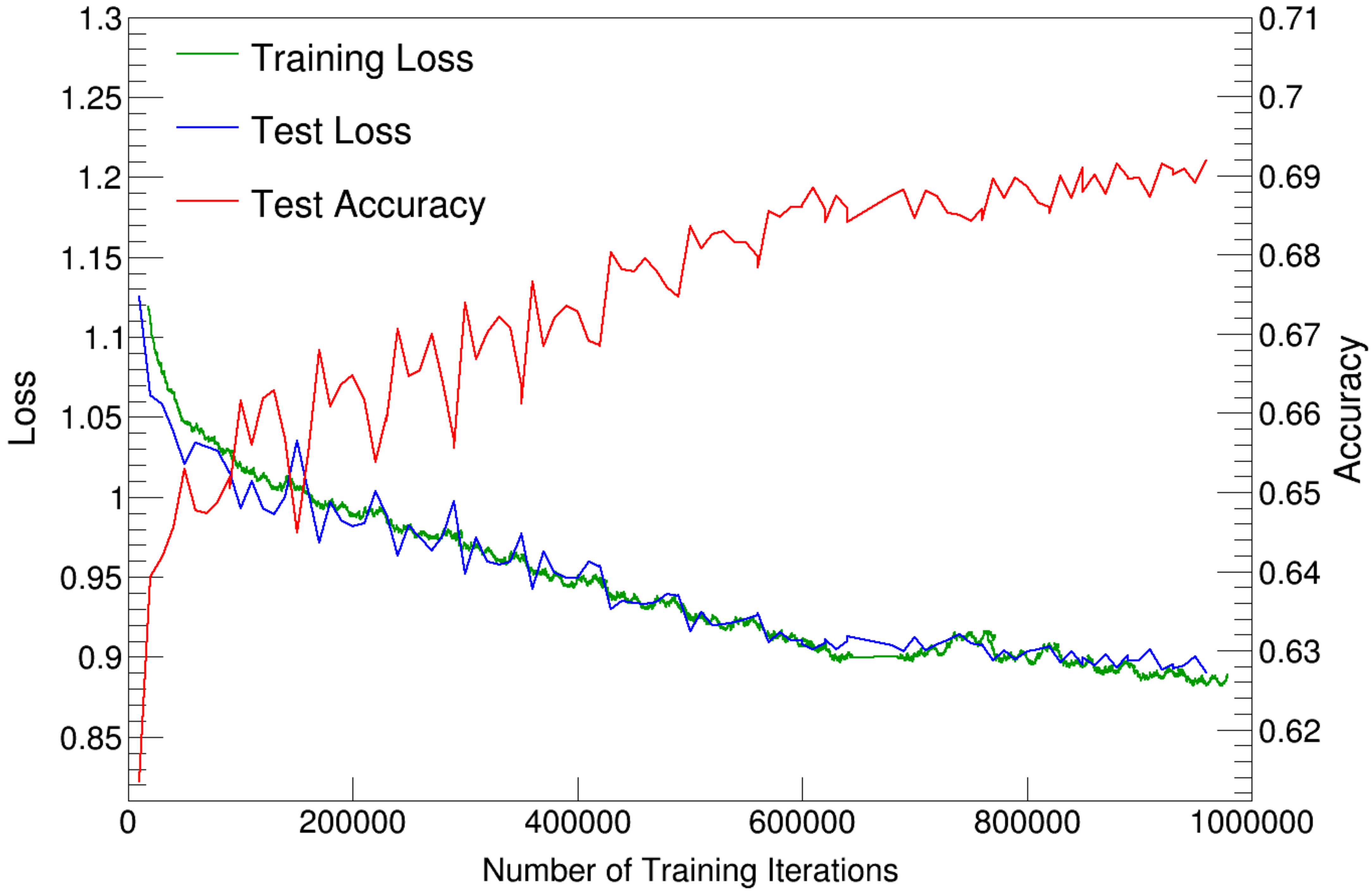
NOvA Event Topologies



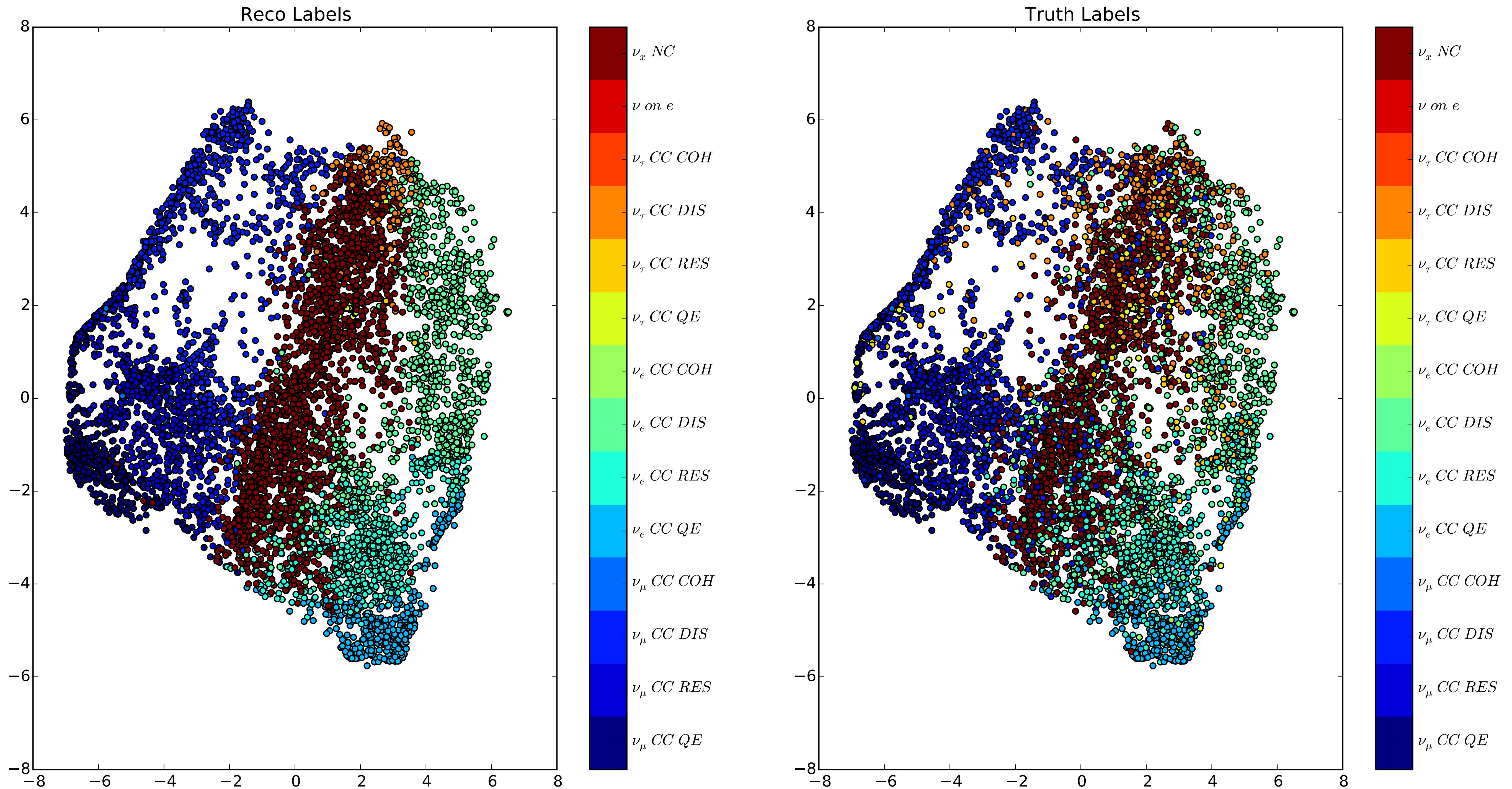
1 radiation length = 38cm (6 cell depths, 10 cell widths)

Training Performance

No sign of overtraining- exceptional training test set performance agreement!

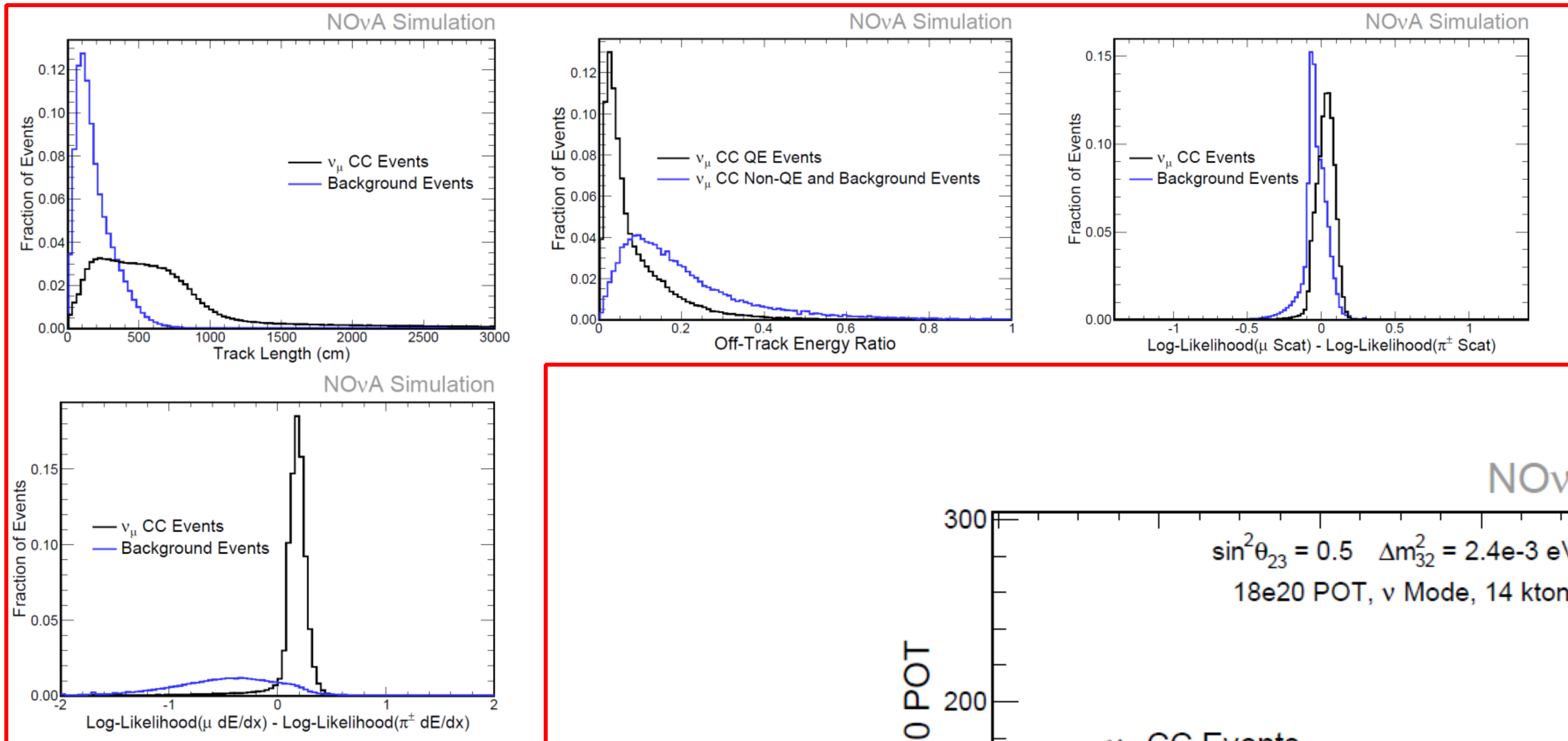


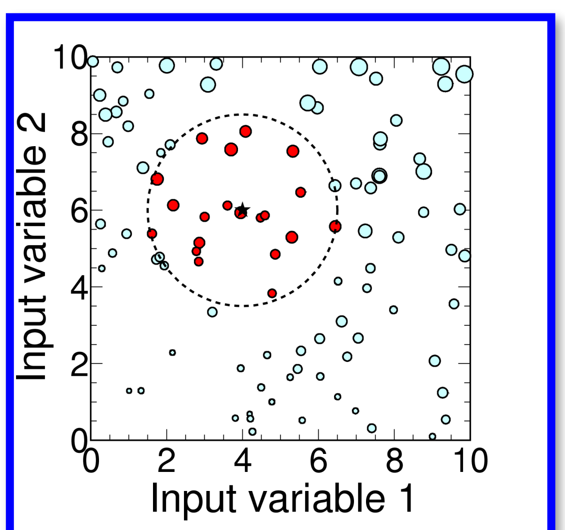
t-SNE Representation of Test Sample

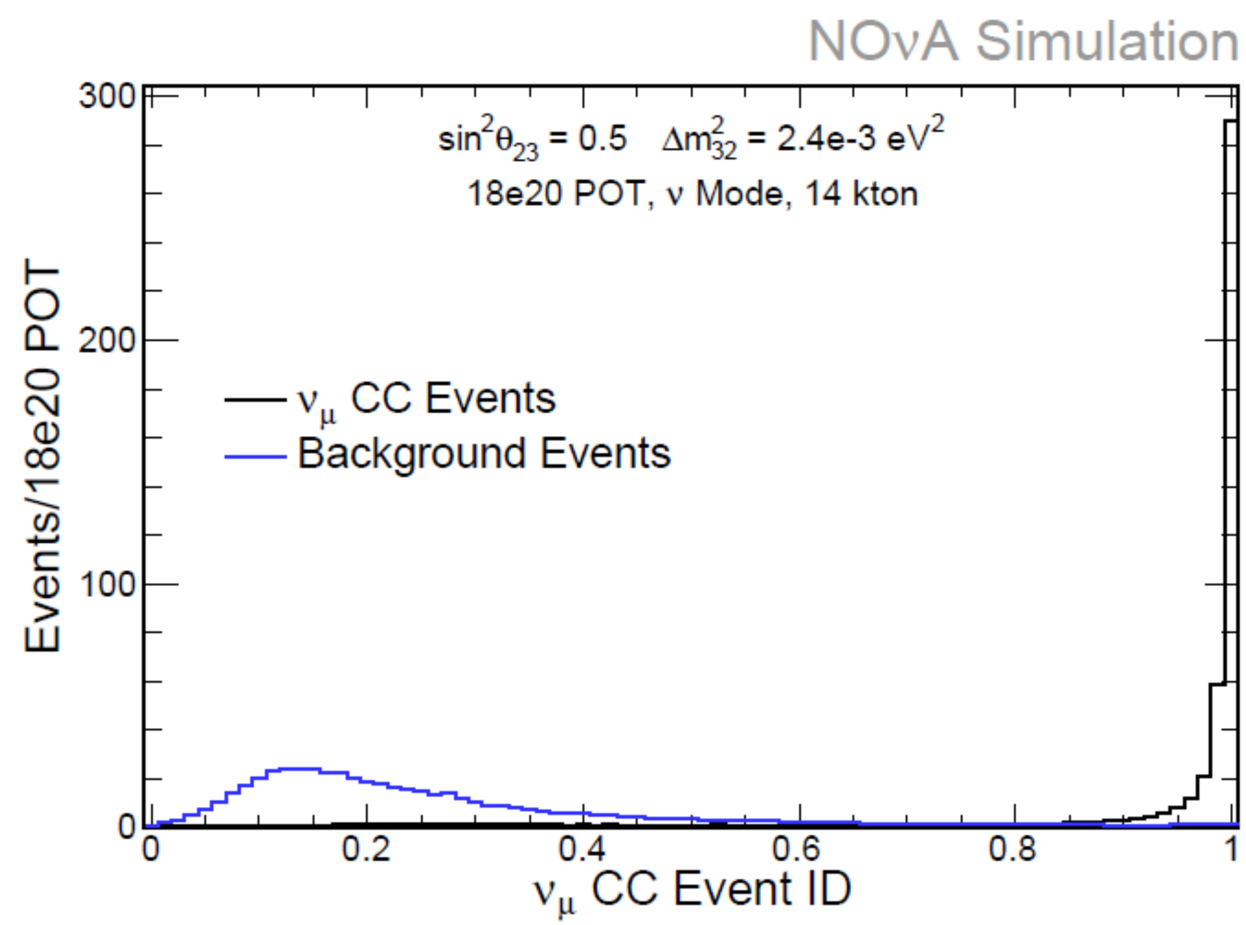


t-SNE projection of final features to 2D. Reco (left) or truth (right) labels, training sample subset.

Conventional PID: ν_μ Selection



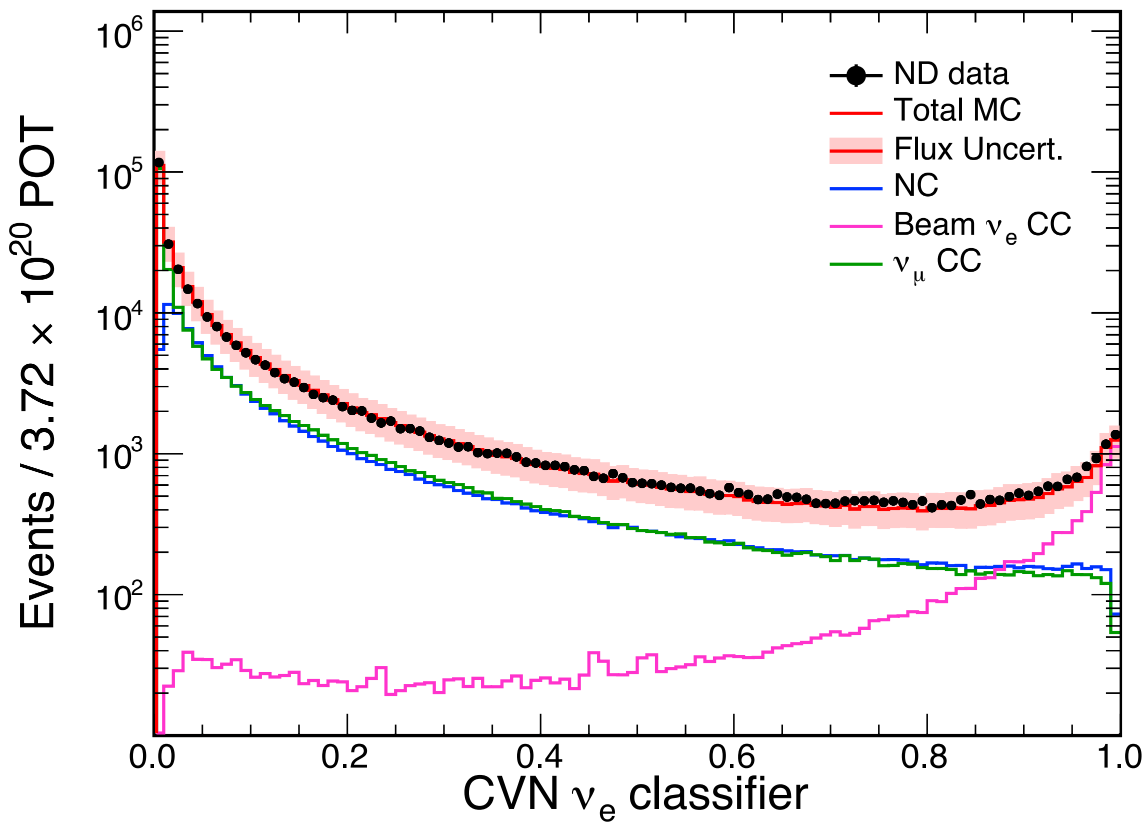


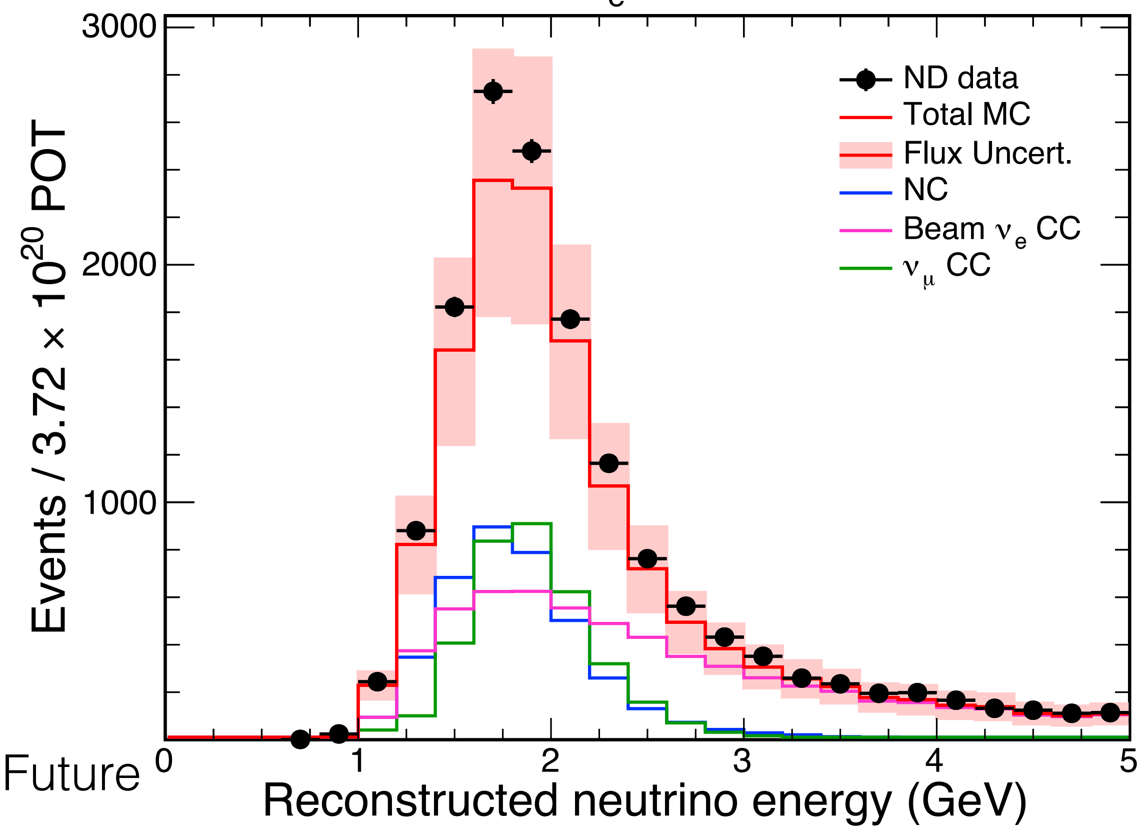


ν_e ND Selected Sample

- Selection optimized to favor parameter measurement- increased signal efficiency by including lower purity bins
- Use ND data to predict background in FD
  - NC, CC, beam ν_e each propagate differently
  - constrain beam ν_e using selected ν_μ CC spectrum
  - constrain ν_μ CC using Michel Electron distribution
- Final correction: beam ν_e up by 4%, NC up by 10%, ν_μ CC up by 17%.

NOvA Preliminary

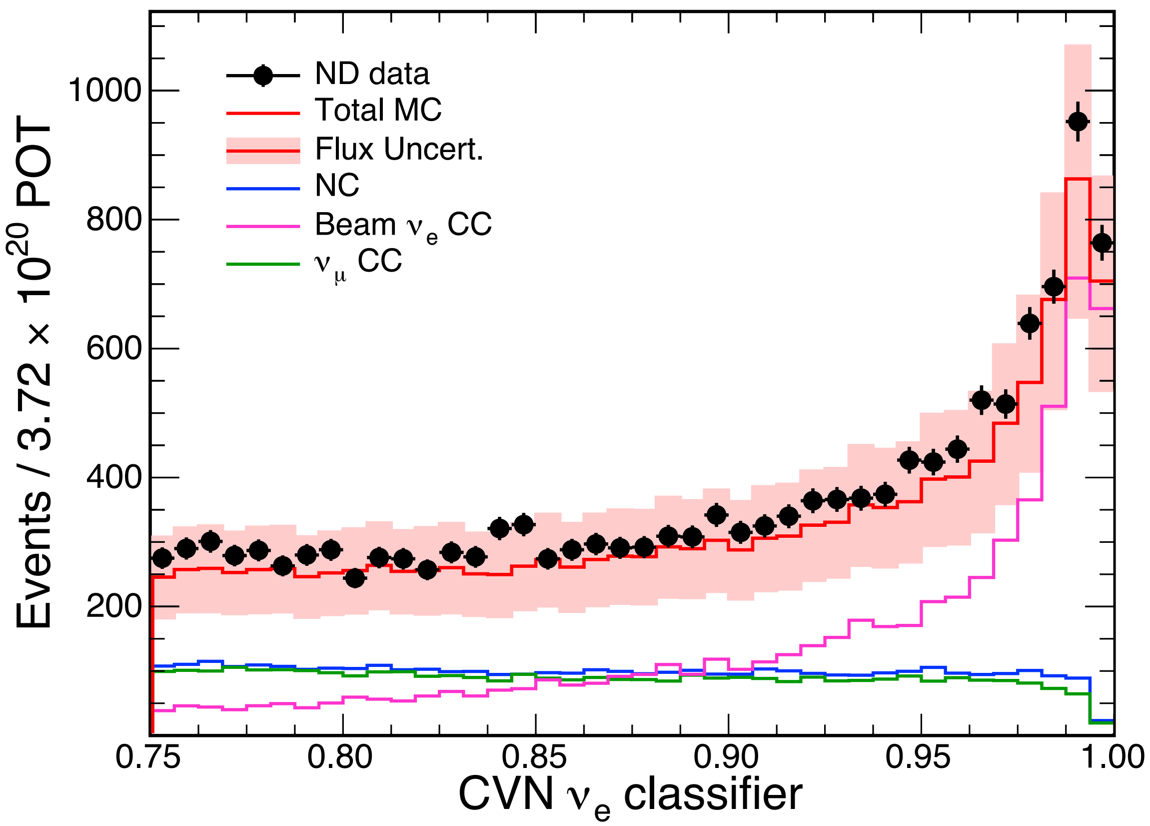


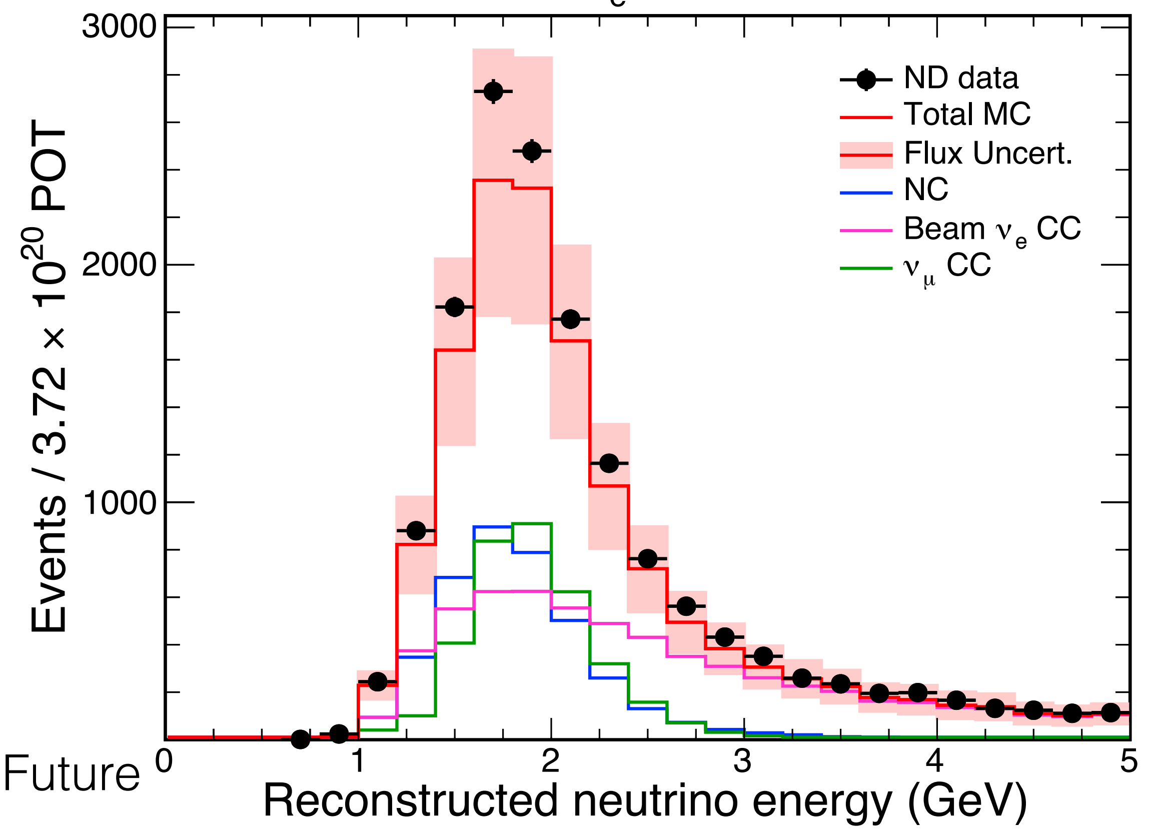


ν_e ND Selected Sample

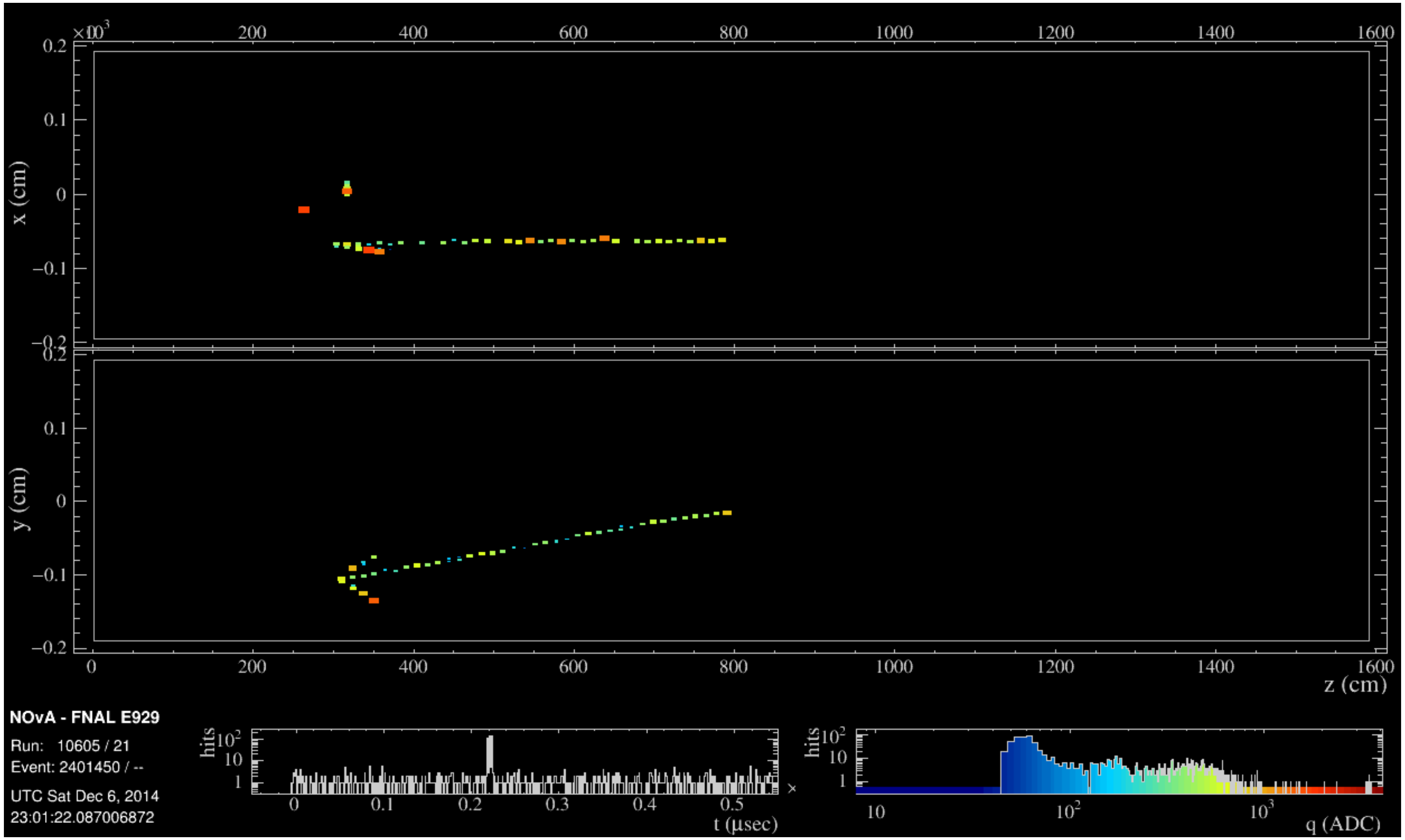
- Selection optimized to favor parameter measurement- increased signal efficiency by including lower purity bins
- Use ND data to predict background in FD
  - NC, CC, beam ν_e each propagate differently
  - constrain beam ν_e using selected ν_μ CC spectrum
  - constrain ν_μ CC using Michel Electron distribution
- Final correction: beam ν_e up by 4%, NC up by 10%, ν_μ CC up by 17%.

NOvA Preliminary



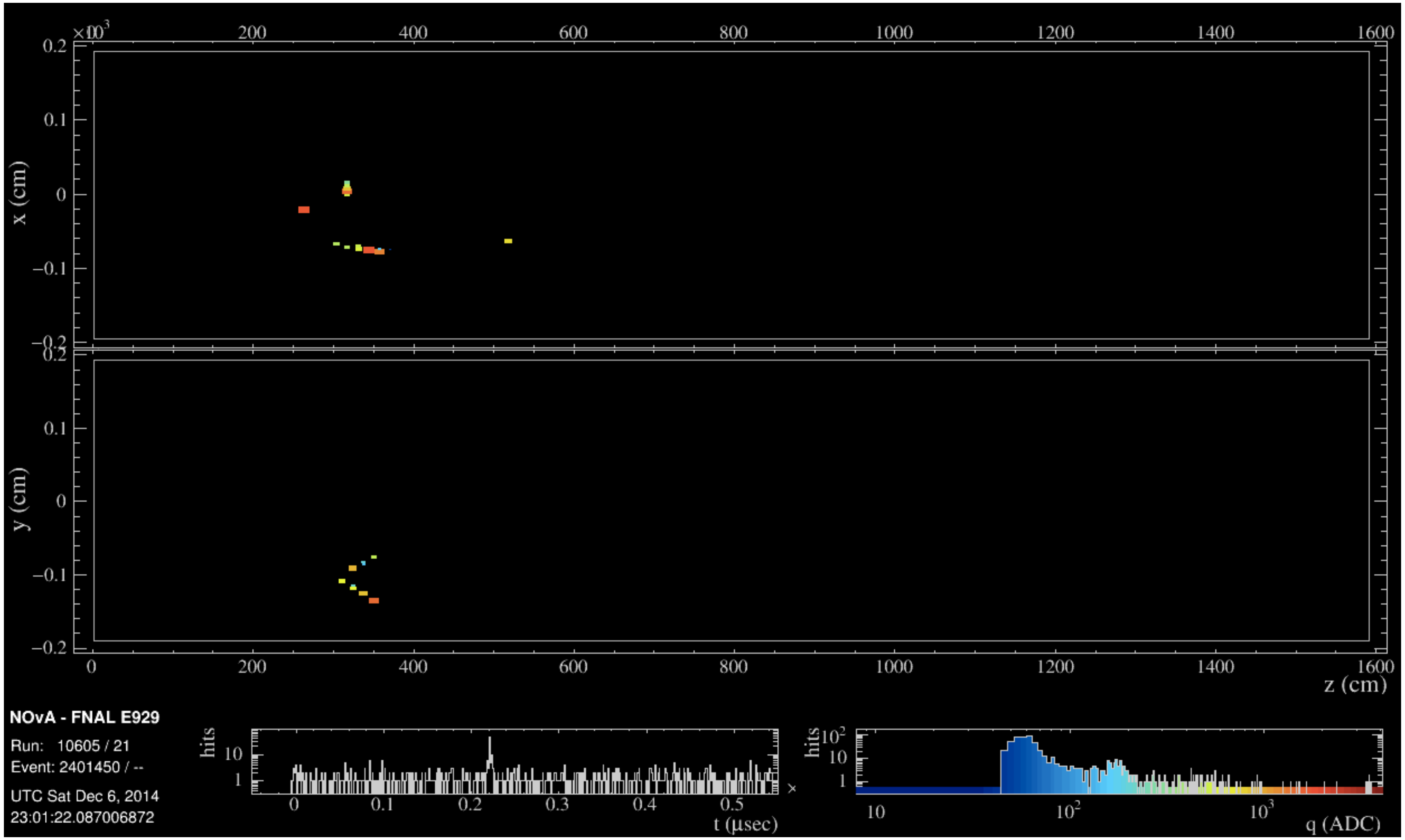


Signal Cross Checks: Muon Removed Electron Added



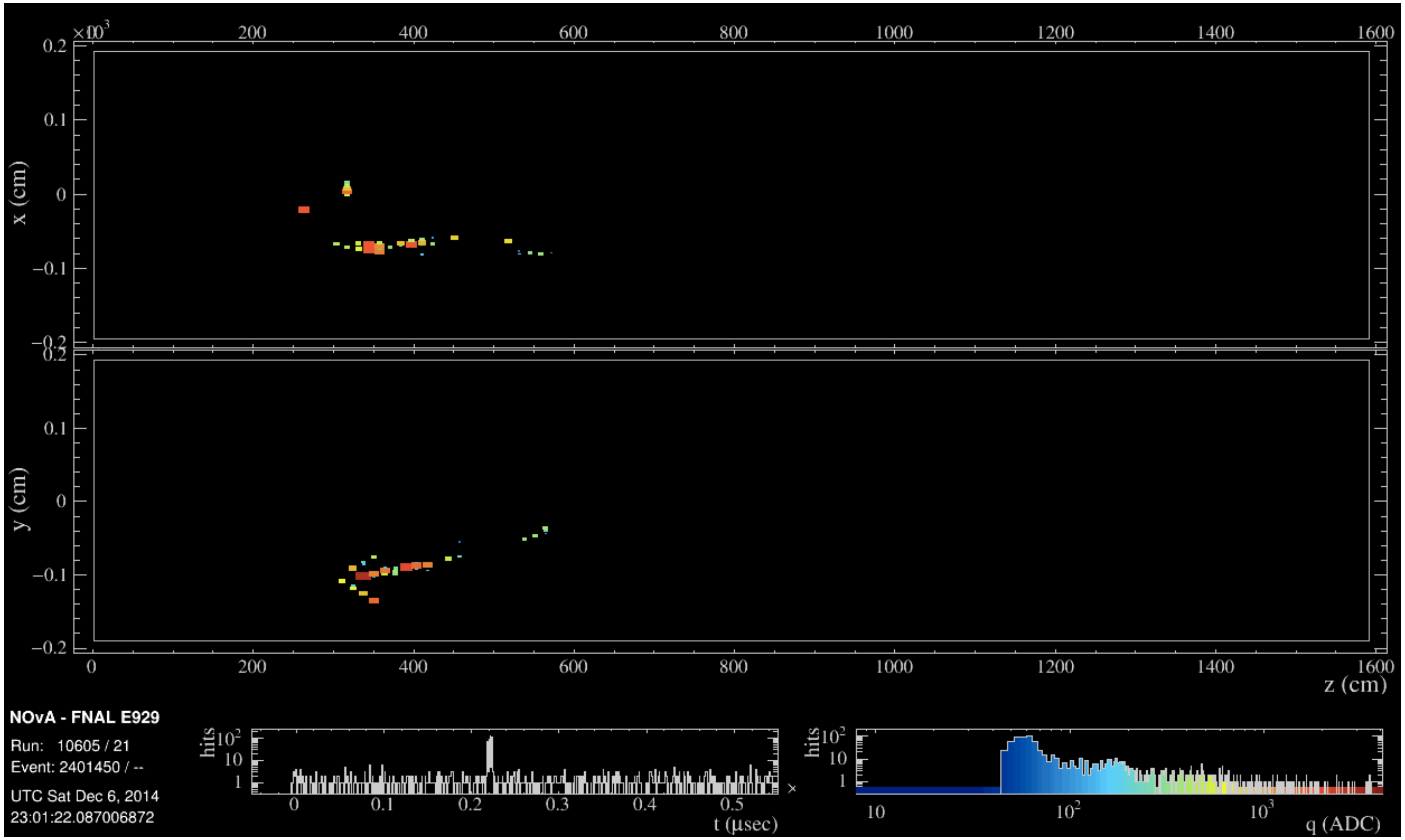
How to check our performance on our signal sample using the Near Detector? Try faking appeared electron neutrinos by creating hybrid data/simulation events.

Signal Cross Checks: Muon Removed Electron Added



How to check our performance on our signal sample using the Near Detector? Try faking appeared electron neutrinos by creating hybrid data/simulation events.

Signal Cross Checks: Muon Removed Electron Added



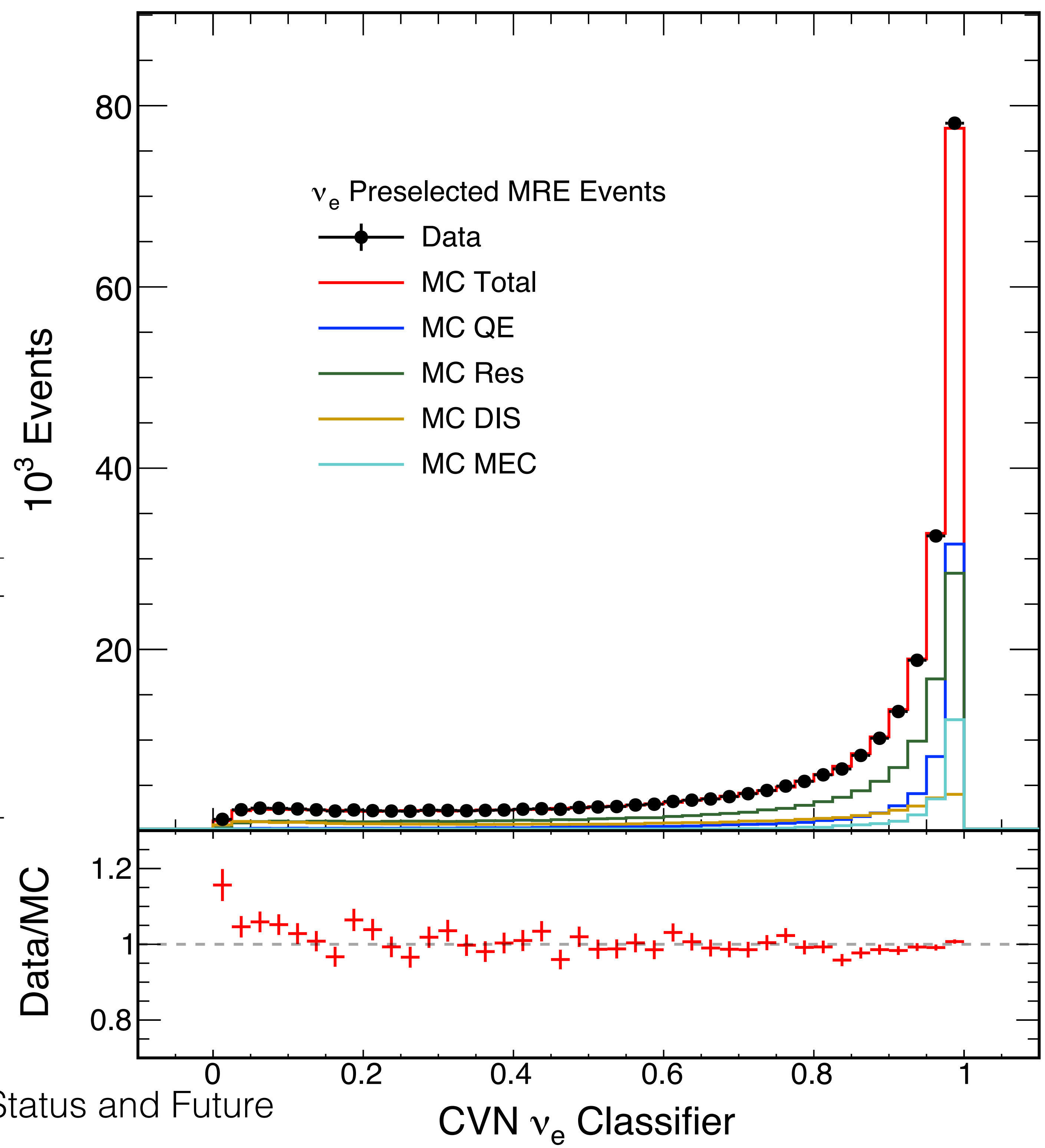
How to check our performance on our signal sample using the Near Detector? Try faking appeared electron neutrinos by creating hybrid data/simulation events.

Signal Cross Checks: Muon Removed Electron Added

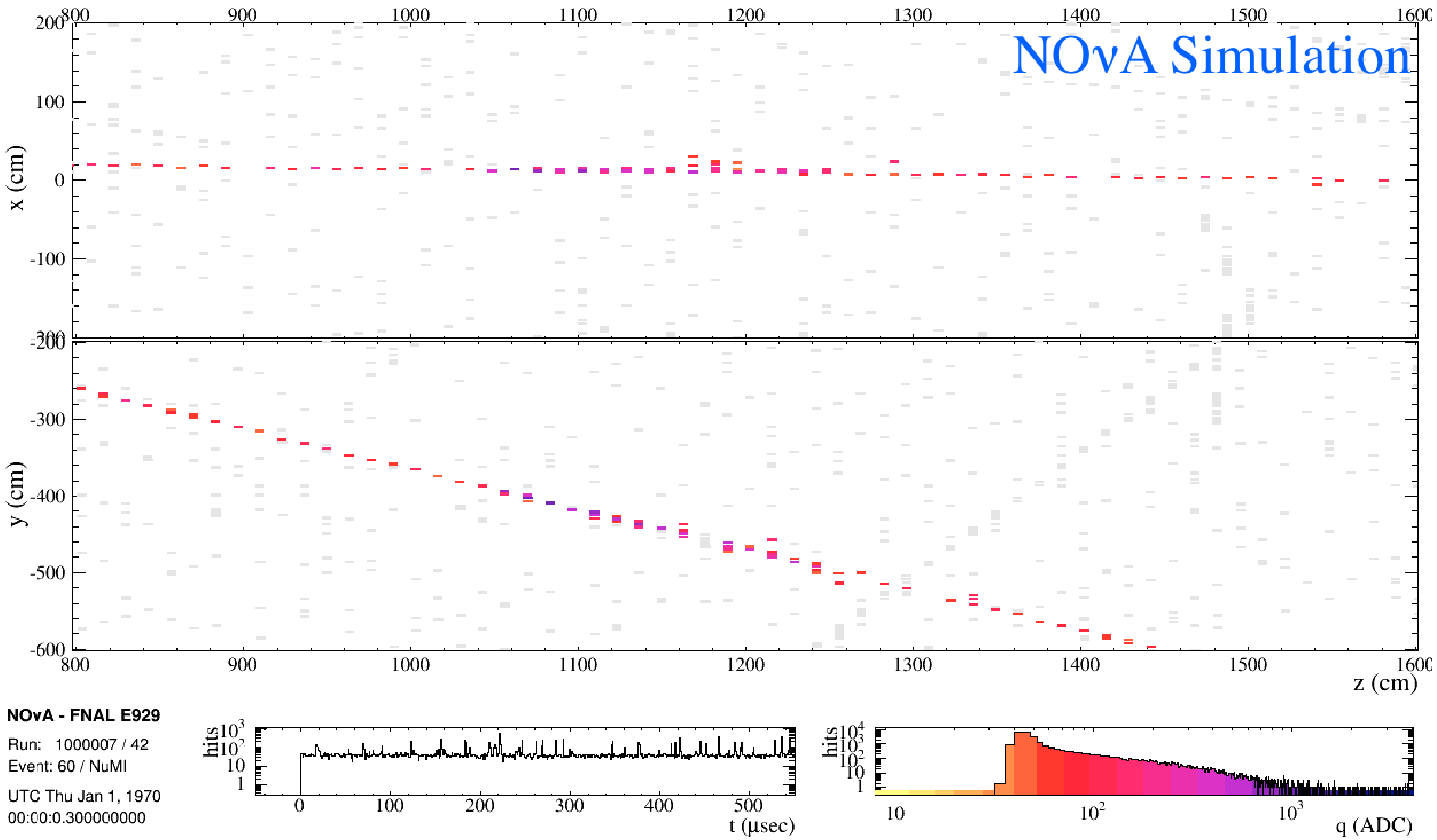
NOvA Preliminary

Excellent data/MC agreement in MRE sample. Efficiency difference <1%. Smaller than for previous PIDs:

PID	Sample	Preselection	PID	Efficiency	Efficiency diff %
CVN	Data	262884	188809	0.718222	-0.36%
	MC	277320	199895	0.720809	
LEM	Data	262884	153599	0.584284	-0.73%
	MC	277320	163218	0.588555	
LID	Data	262884	175492	0.667564	2.09%
	MC	277320	181267	0.653638	

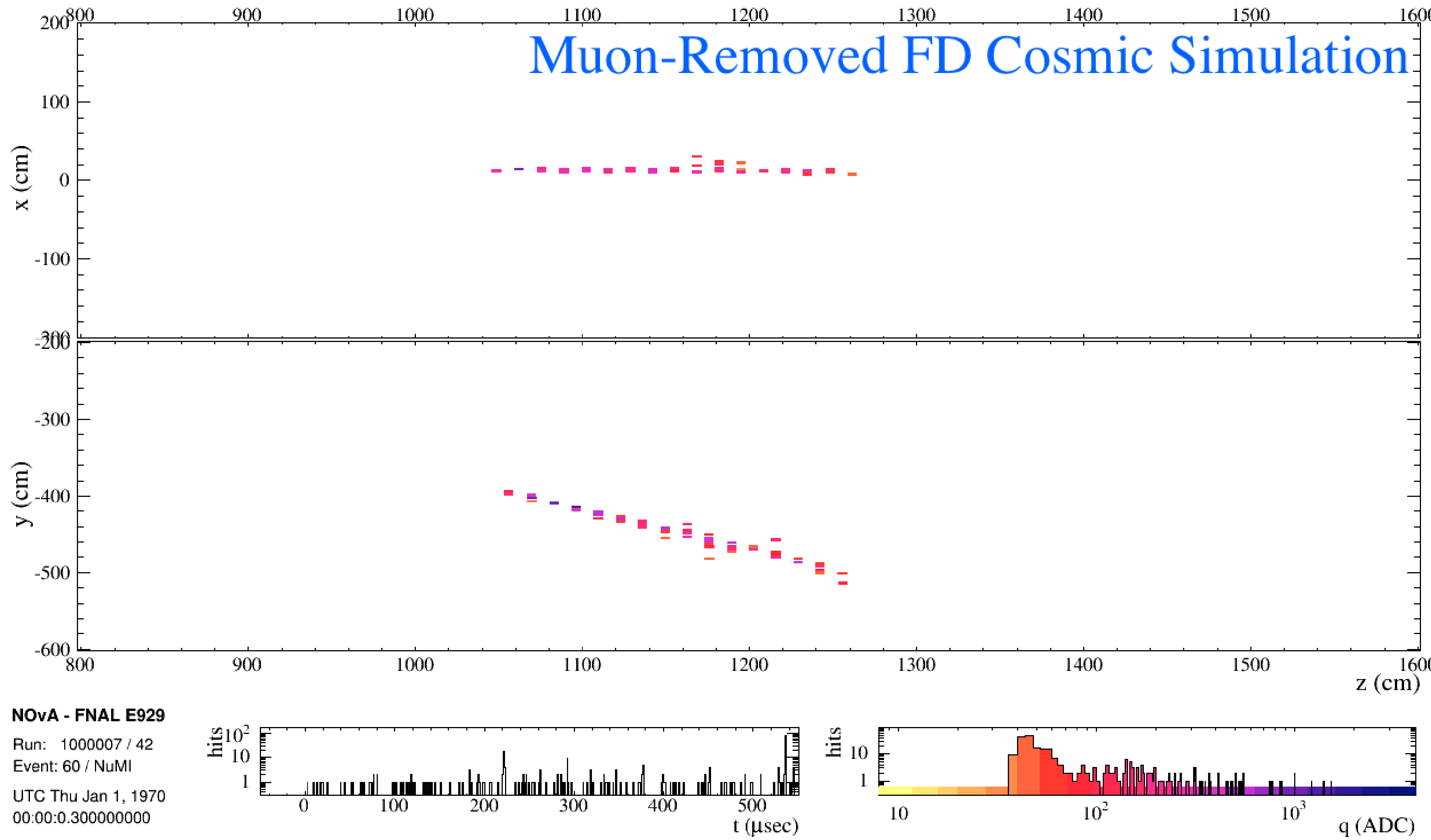


Signal Cross Checks: Muon Removed Bremsstrahlung



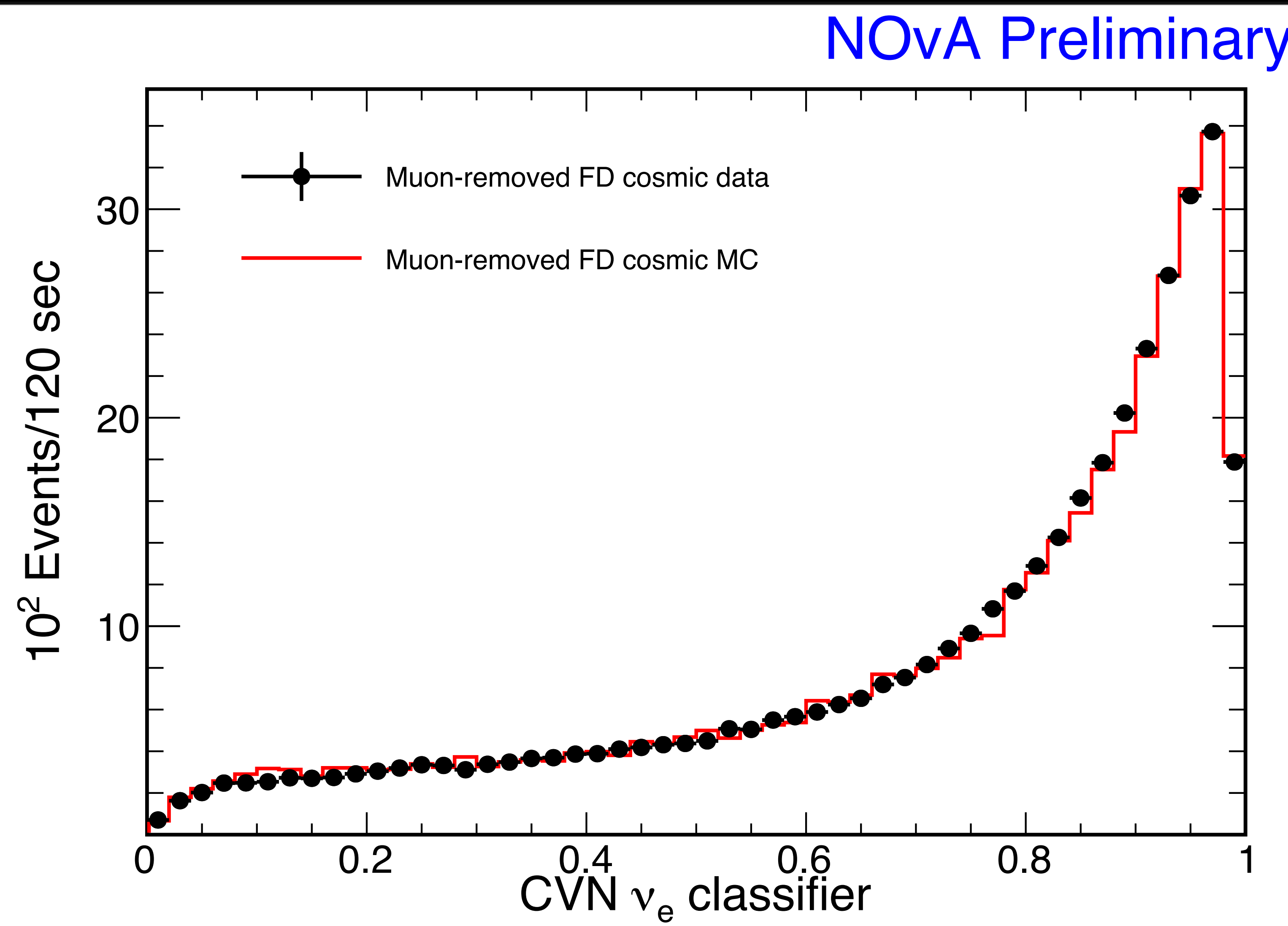
But what about the Far Detector? Try using cosmogenic activity. We find Bremsstrahlung, remove the associated muon, and see what CVN does in data vs. simulation.

Data Driven Cross Checks: Muon Removed Bremsstrahlung



But what about the Far Detector? Try using cosmogenic activity. We find Bremsstrahlung, remove the associated muon, and see what CVN does in data vs. simulation.

Signal Cross Checks: Muon Removed Bremsstrahlung



But what about the Far Detector? Try using cosmogenic activity. We find Bremsstrahlung, remove the associated muon, and see what CVN does in data vs. simulation.

ν_e FD Predicted Sample

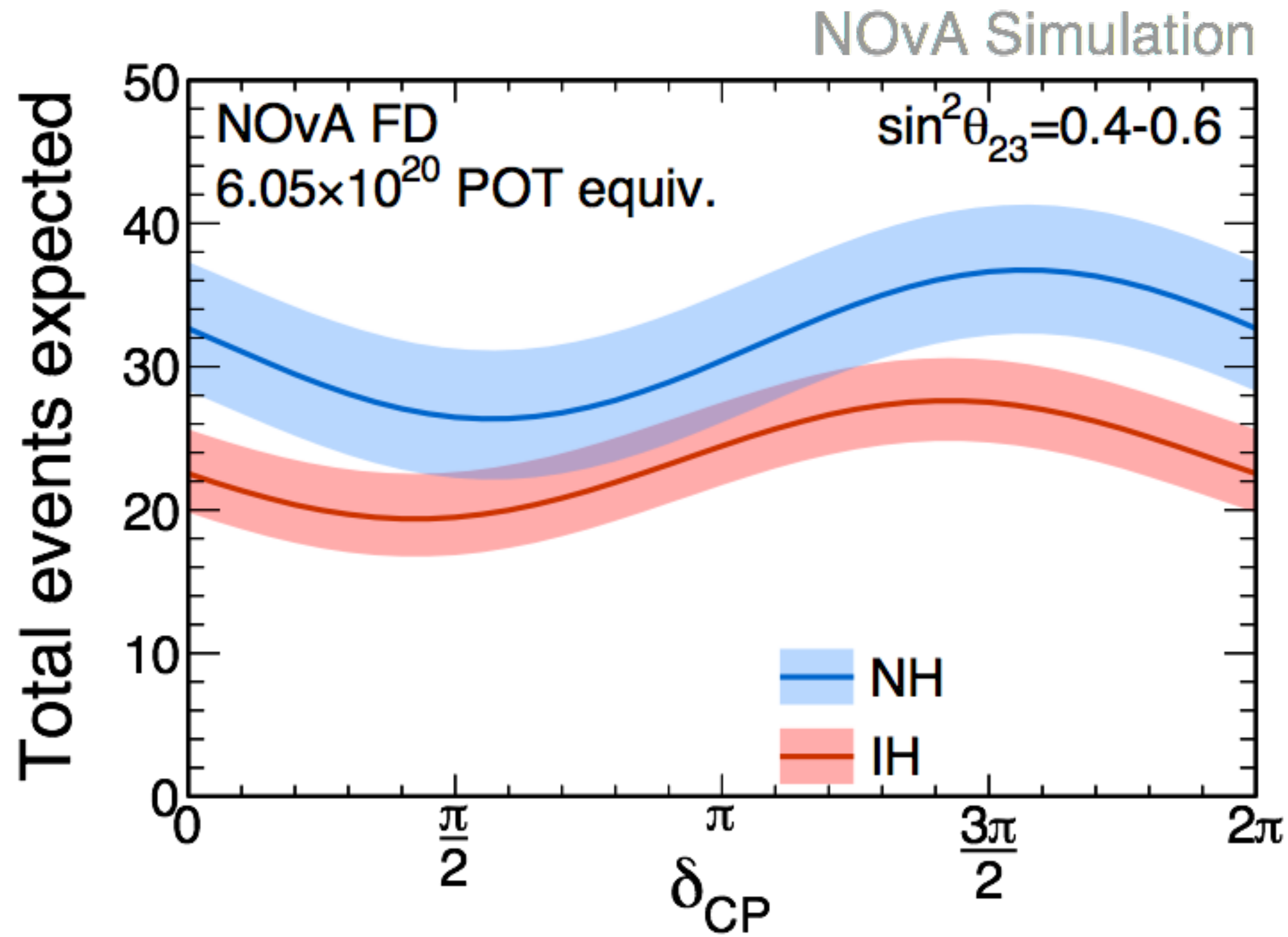
- Extrapolate each component in bins of energy and CVN output.
- Expected event counts depend on oscillation parameters.

Signal events
(±5% systematic uncertainty):

NH, $3\pi/2$,	IH, $\pi/2$,
28.2	11.2

Background by component
(±10% systematic uncertainty):

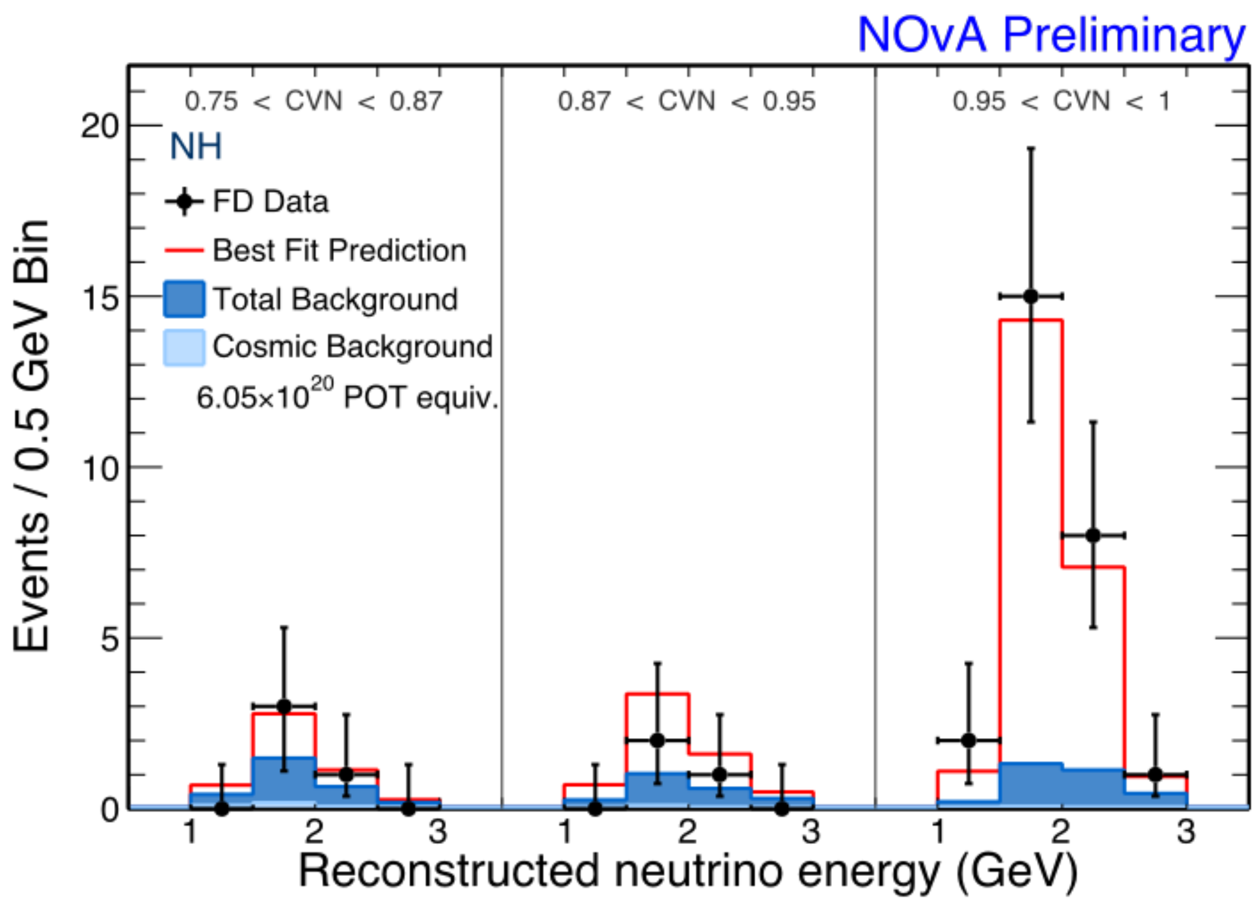
Total BG	NC	Beam ν_e	ν_μ CC	ν_τ CC	Cosmics
8.2	3.7	3.1	0.7	0.1	0.5

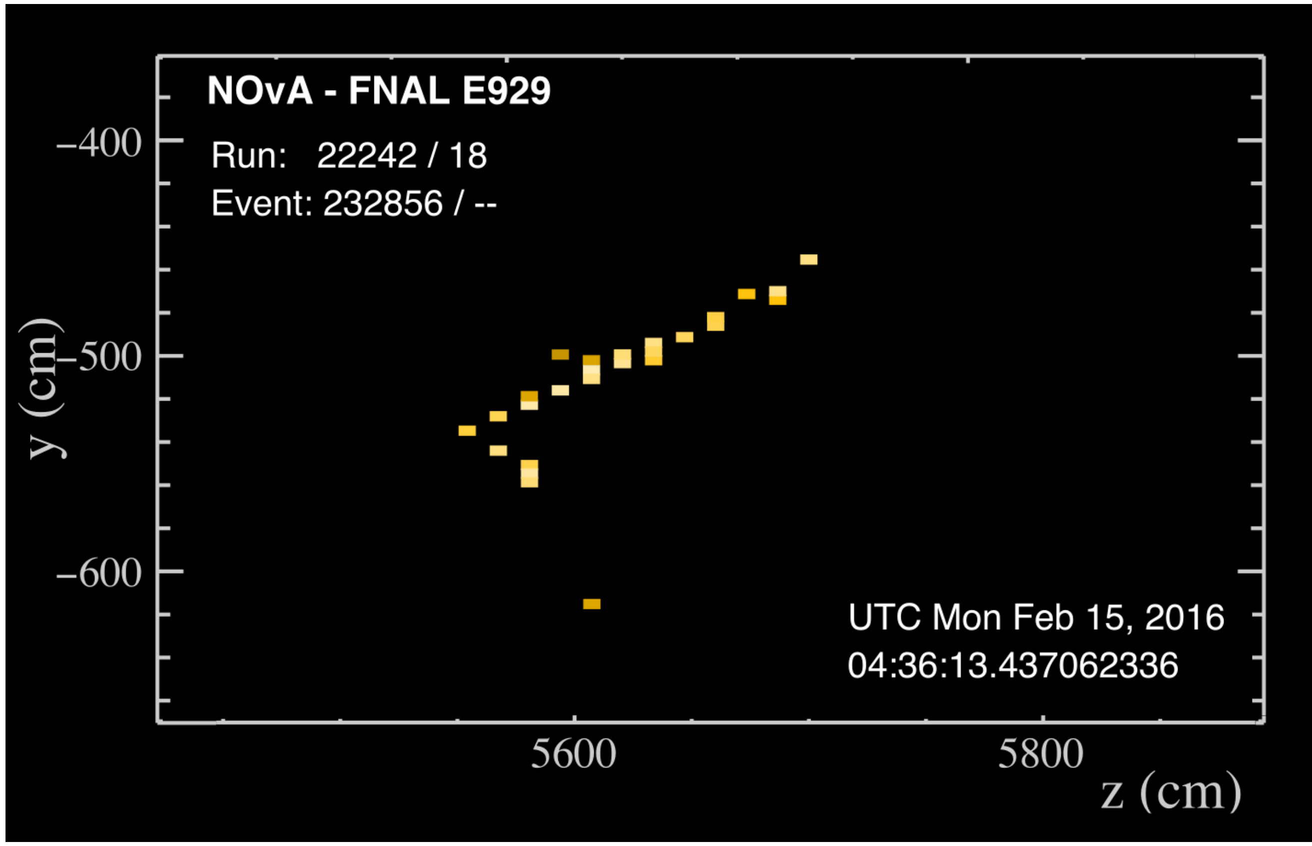


ν_e FD Selected Sample

>8σ electron neutrino appearance signal

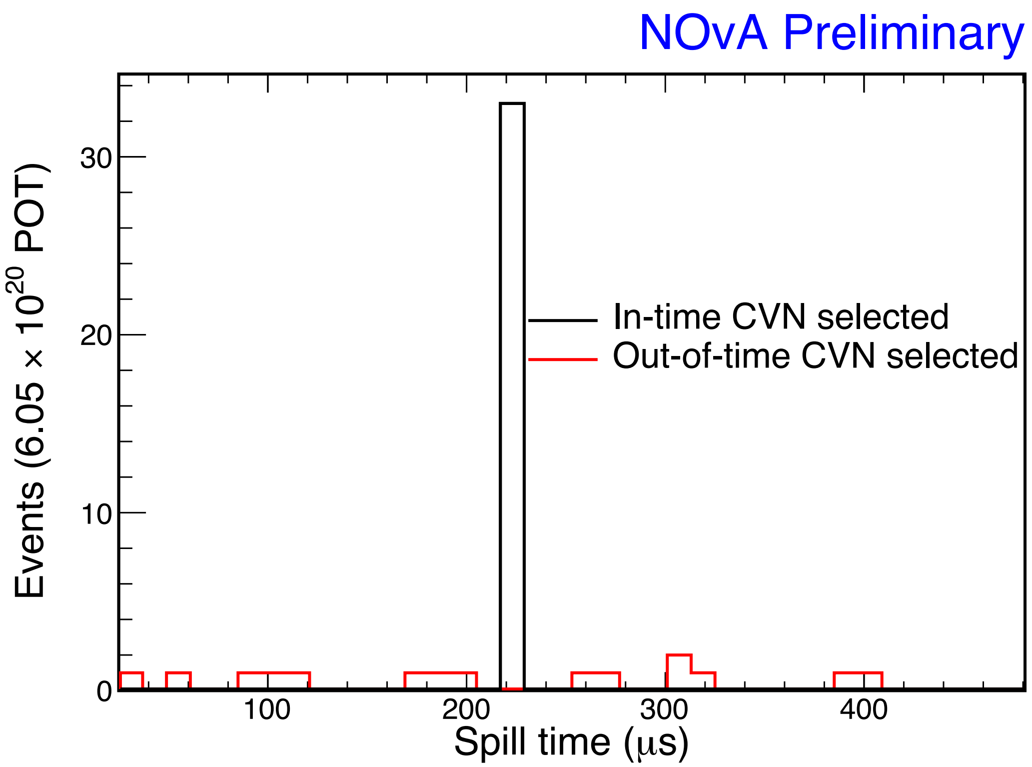
Observe **33 events in FD**. Background Expectation 8.2±0.8.

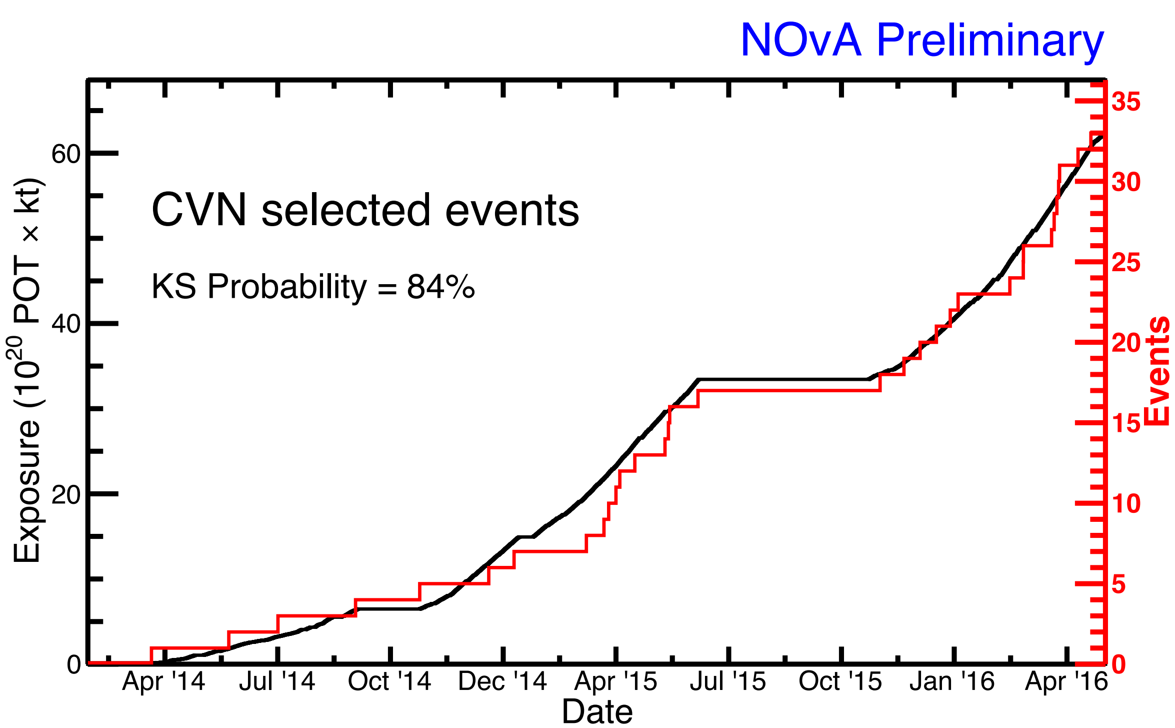


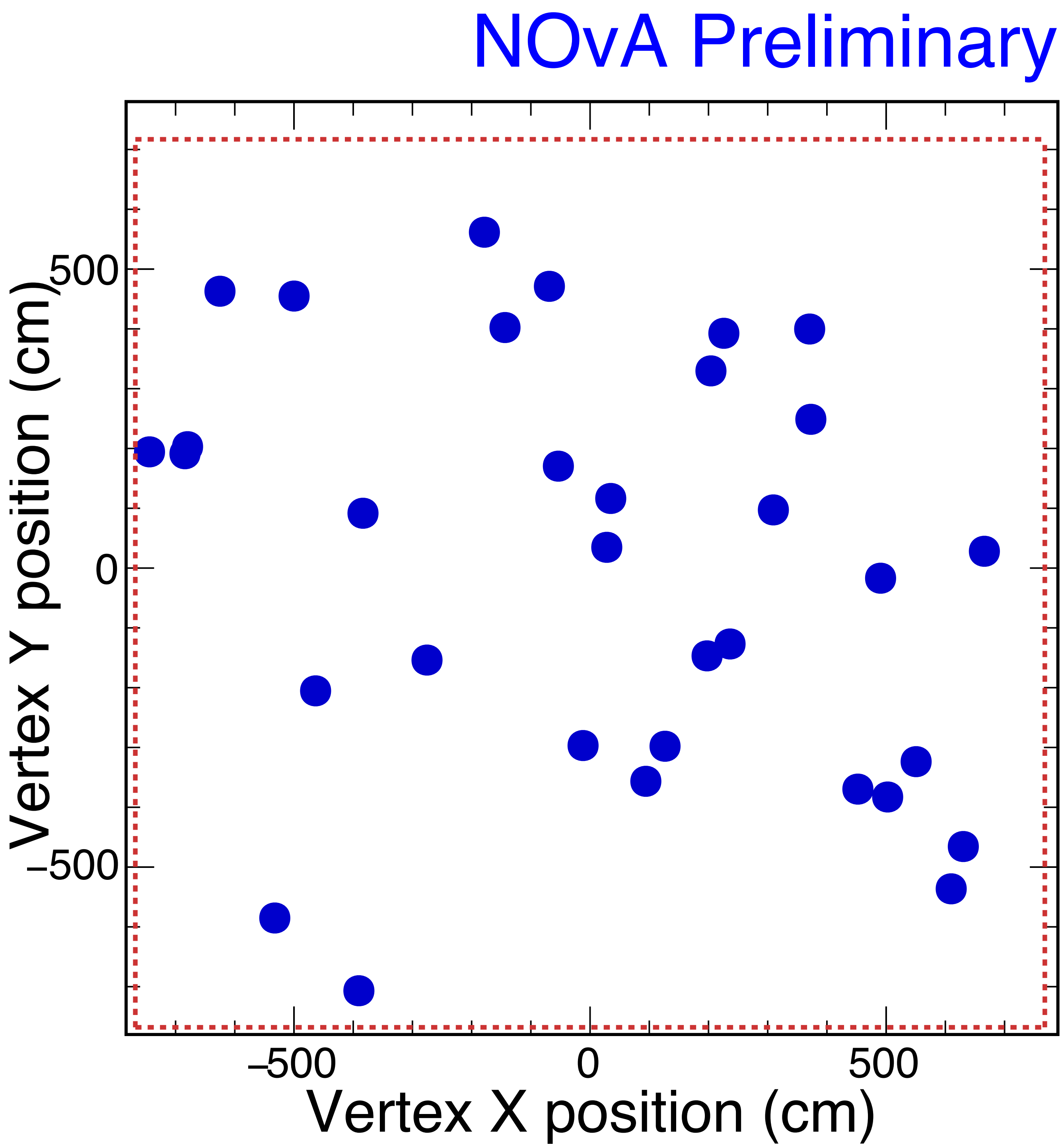


Alternate selectors from 2015 analysis show consistent results
LID: 34 events, 12.2±1.2 BG expected
LEM: 33 events, 10.3±1.0 BG expected

ν_e candidates, when & where?





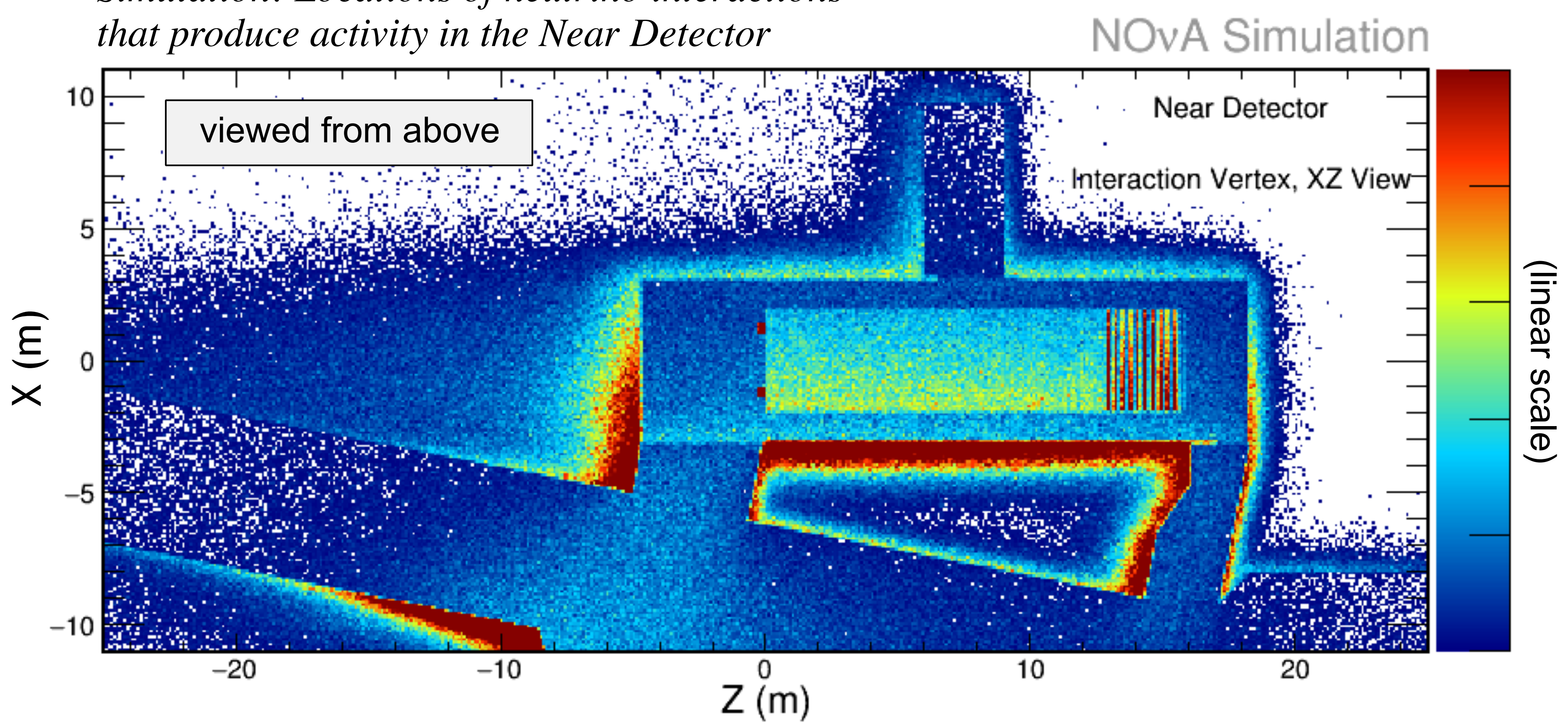


Simulation

Highly detailed end-to-end simulation chains:

- Beam hadron production, propagatio, neutrino flux: **FLUKA/FLUGG**
- Cosmic ray flux: **CRY**
- Neutrino Interactions and FSI modeling: **GENIE**
- Detector Simulation: **GEANT4**
- Readout electronics and DAQ: **Custom simulation routines**

Simulation: Locations of neutrino interactions that produce activity in the Near Detector



Reconstruction

Three key pieces:

- **Vertexing:** use lines of energy deposition formed with hough transforms to find intersections
- **Clustering:** find clusters in angular space around the vertex and merge views via topology and prong dE/dx
- **Tracking:** Trace particle trajectories using a kalman filter, example below