Overview of Machine Learning on NOvA

Micah Groh
Indiana University

On behalf of the
NOvA Collaboration

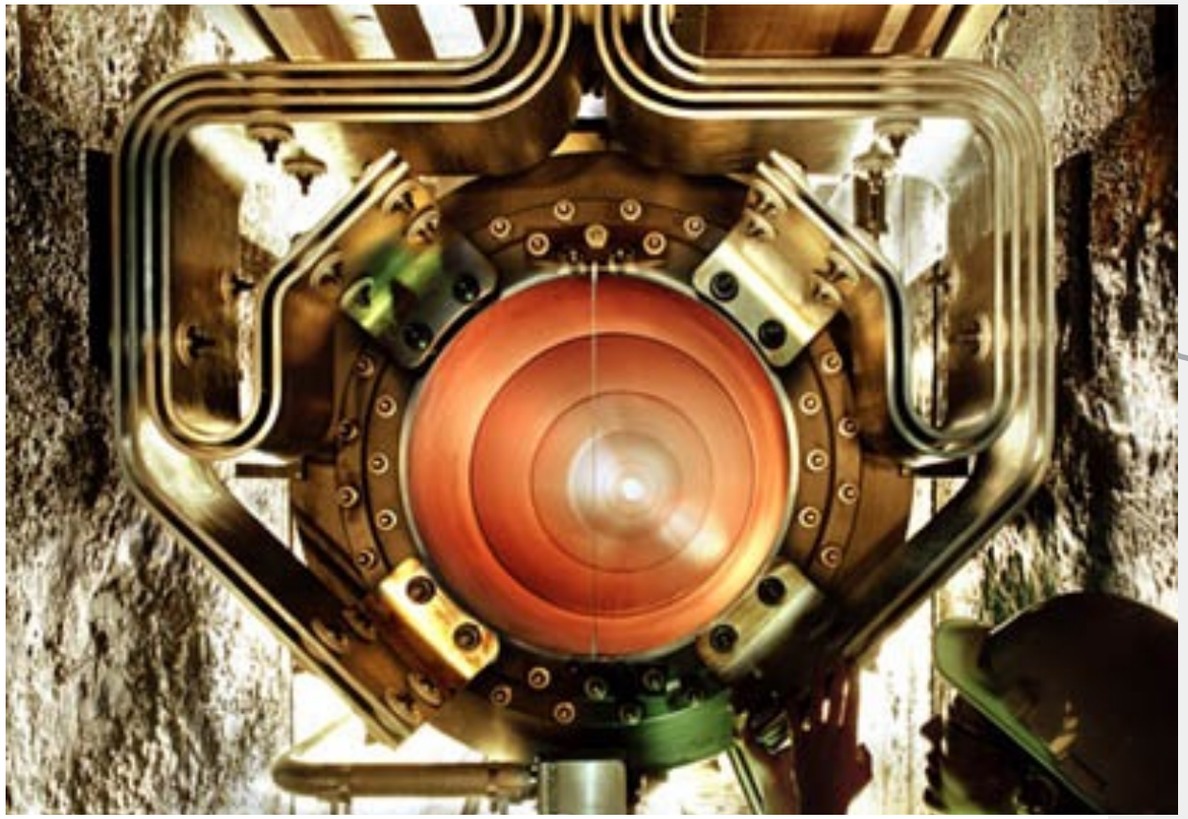
NOvA

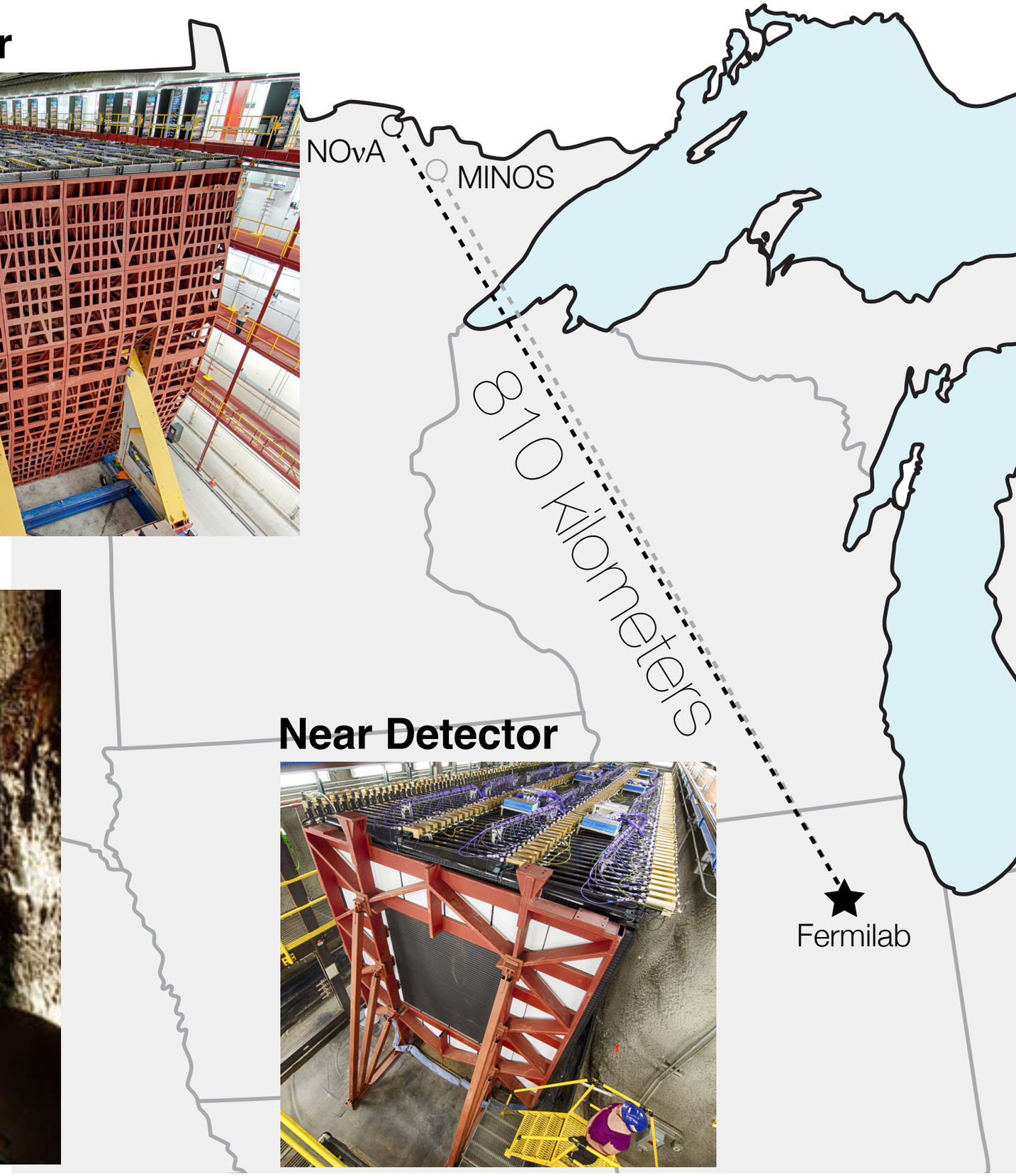
NuMI Off-Axis ν_e Appearance Experiment

A long baseline neutrino oscillation experiment using the NuMI beam at Fermilab.

Far Detector

NuMI Horn

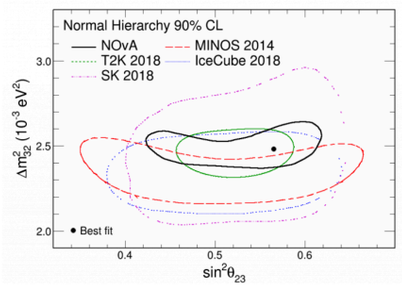




Near Detector

NOvA

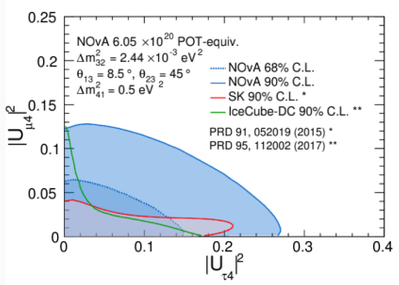
Physics Program



First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOvA

FERMILAB-PUB-19-272-ND

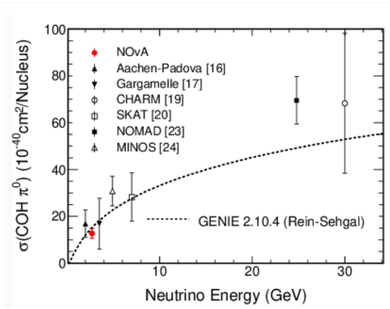
June 2019



Search for active-sterile neutrino mixing using neutral-current interactions in NOvA

FERMILAB-PUB-17-198-ND

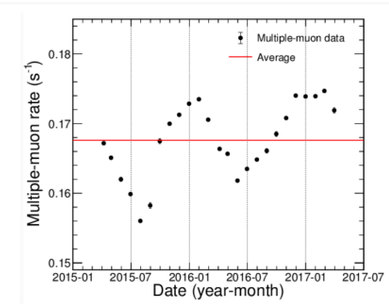
June 2017



Measurement of Neutrino-Induced Neutral-Current Coherent π^0 Production in the NOvA Near Detector

FERMILAB-PUB-19-047-ND

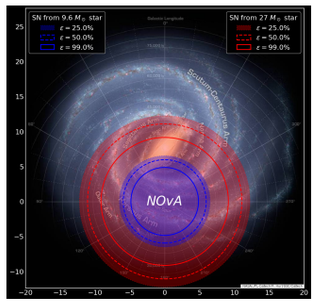
February 2019



Observation of seasonal variation of atmospheric multiple-muon events in the NOvA Near Detector

Phys. Rev. D 99, 122004

April 2019



Supernova Trigger Paper in Preparation

NOvA has a broad physics program using the two detectors.

- Neutrino Oscillations
- Sterile Neutrino Searches
- Neutrino Cross Sections
- Cosmic Ray Physics
- Supernova Neutrinos

NOvA Publications Page

NOvA

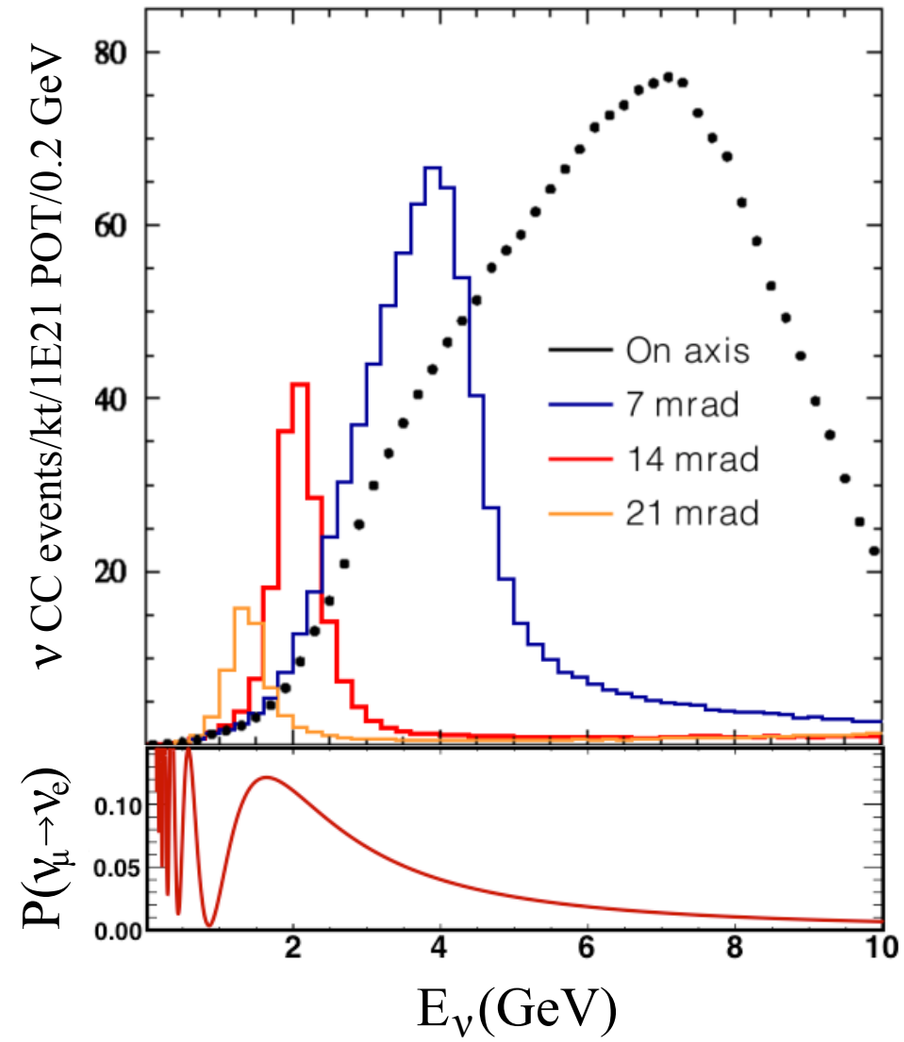
Neutrinos from the Main Injector

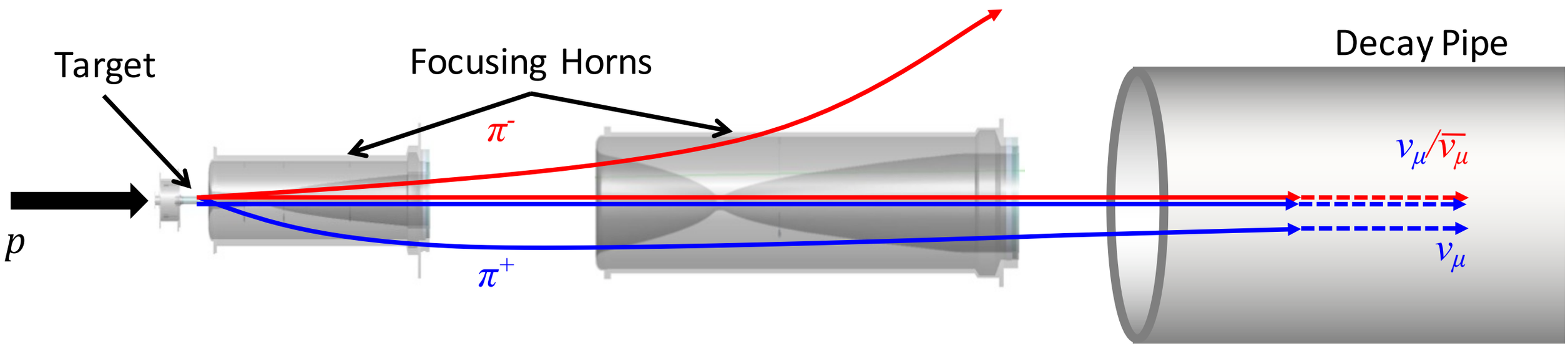
NuMI is currently the **most intense neutrino beam in the world**.

Running at >700 kW since 2017.

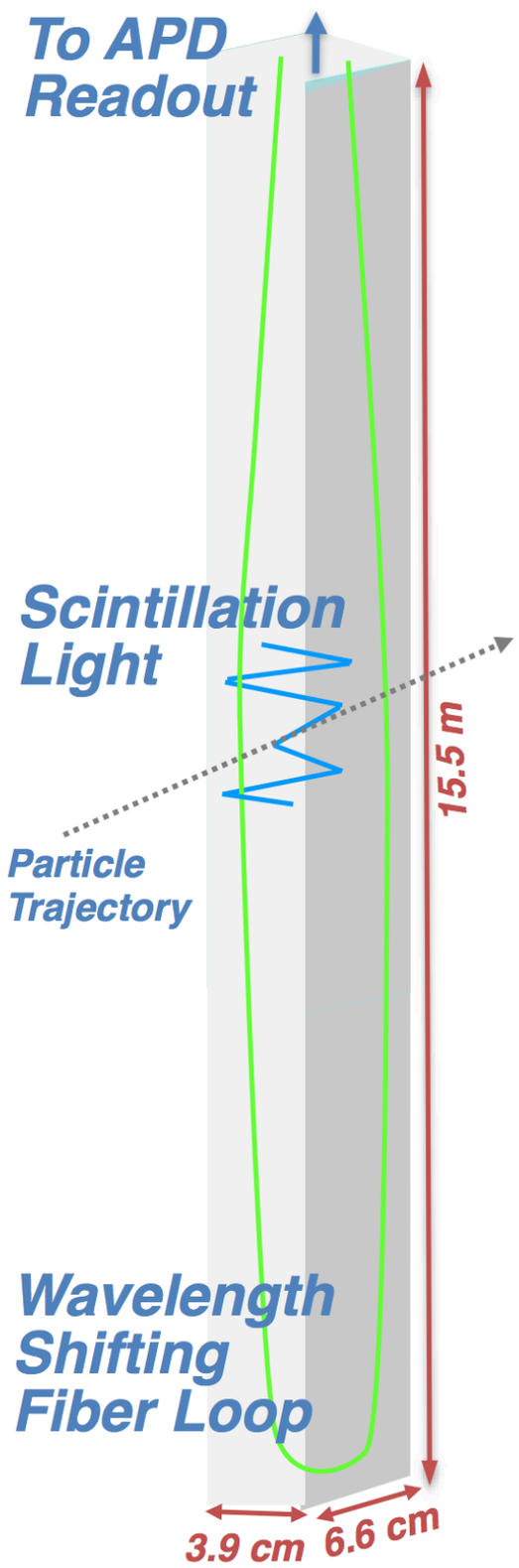
Achieved 758 kW hourly peak average this year.

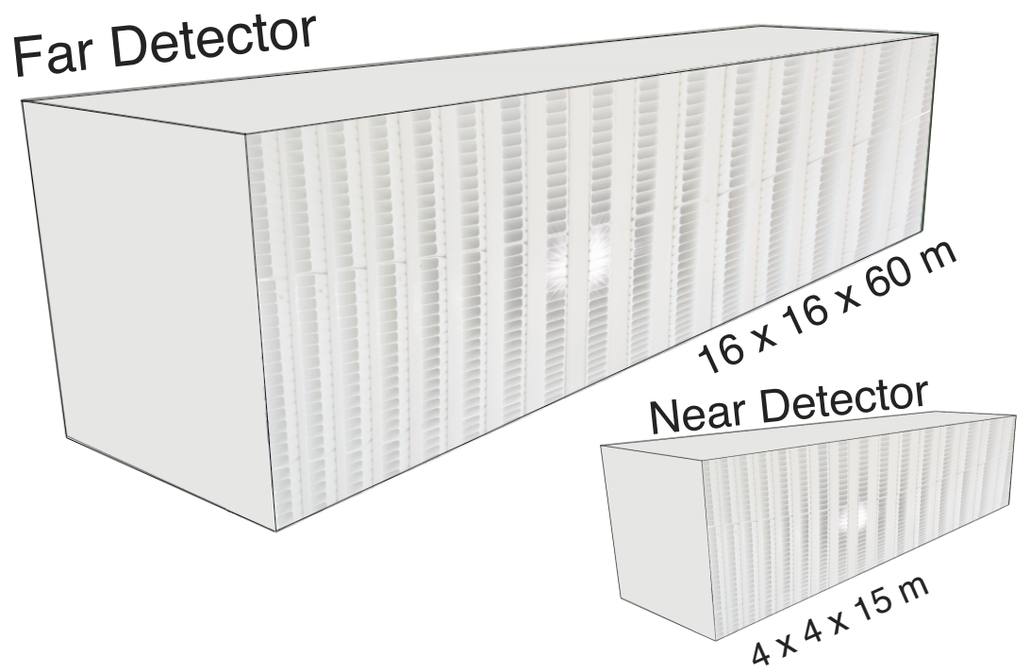
NOvA is positioned 14.6 mrad off-axis to constrain the neutrino energy and reduce backgrounds.





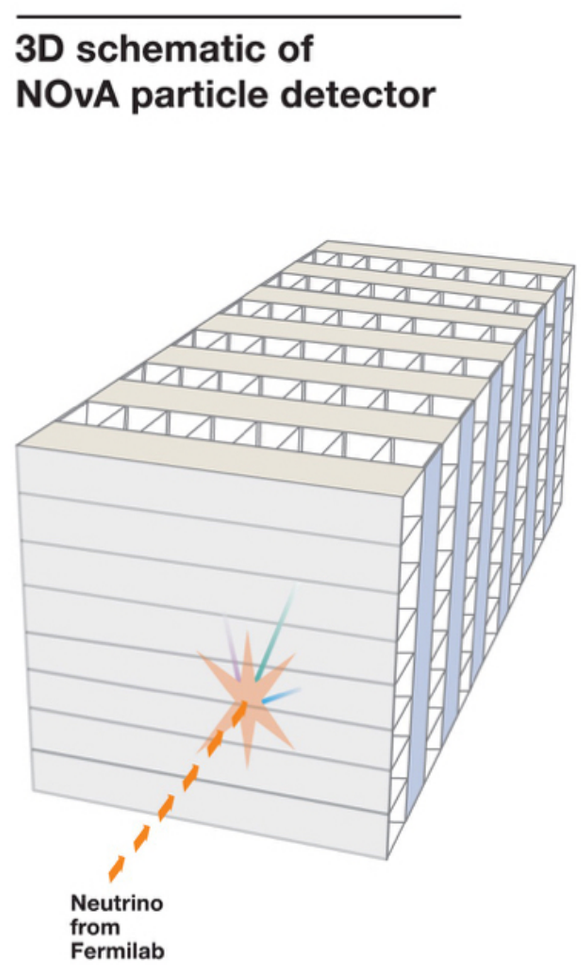
NOvA

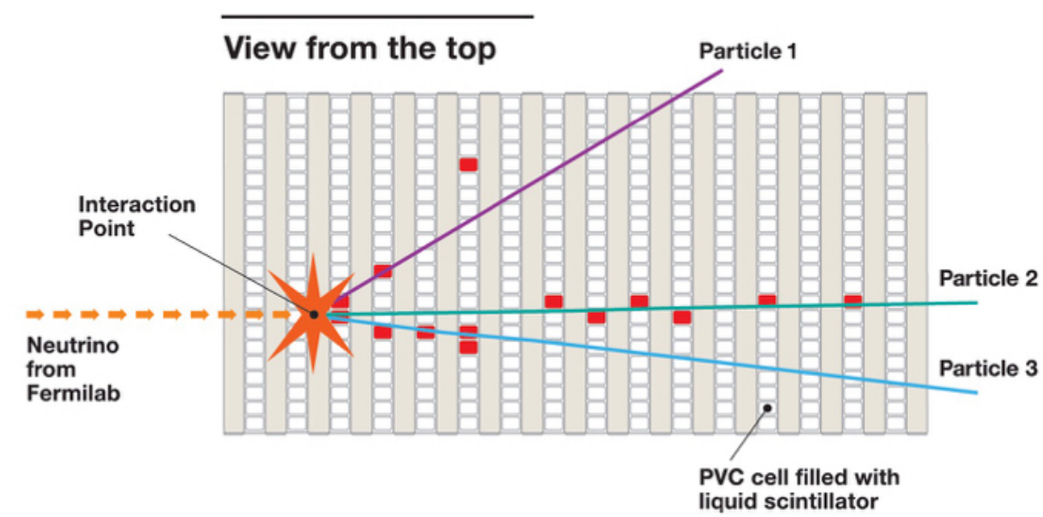


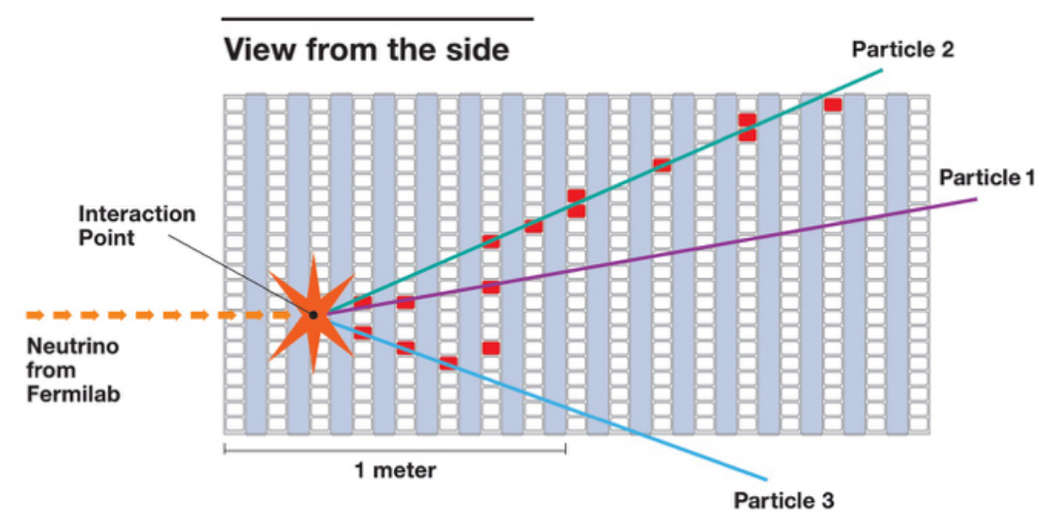


The detectors are functionally equivalent sampling calorimeters.

Detect scintillation light from charged particles.

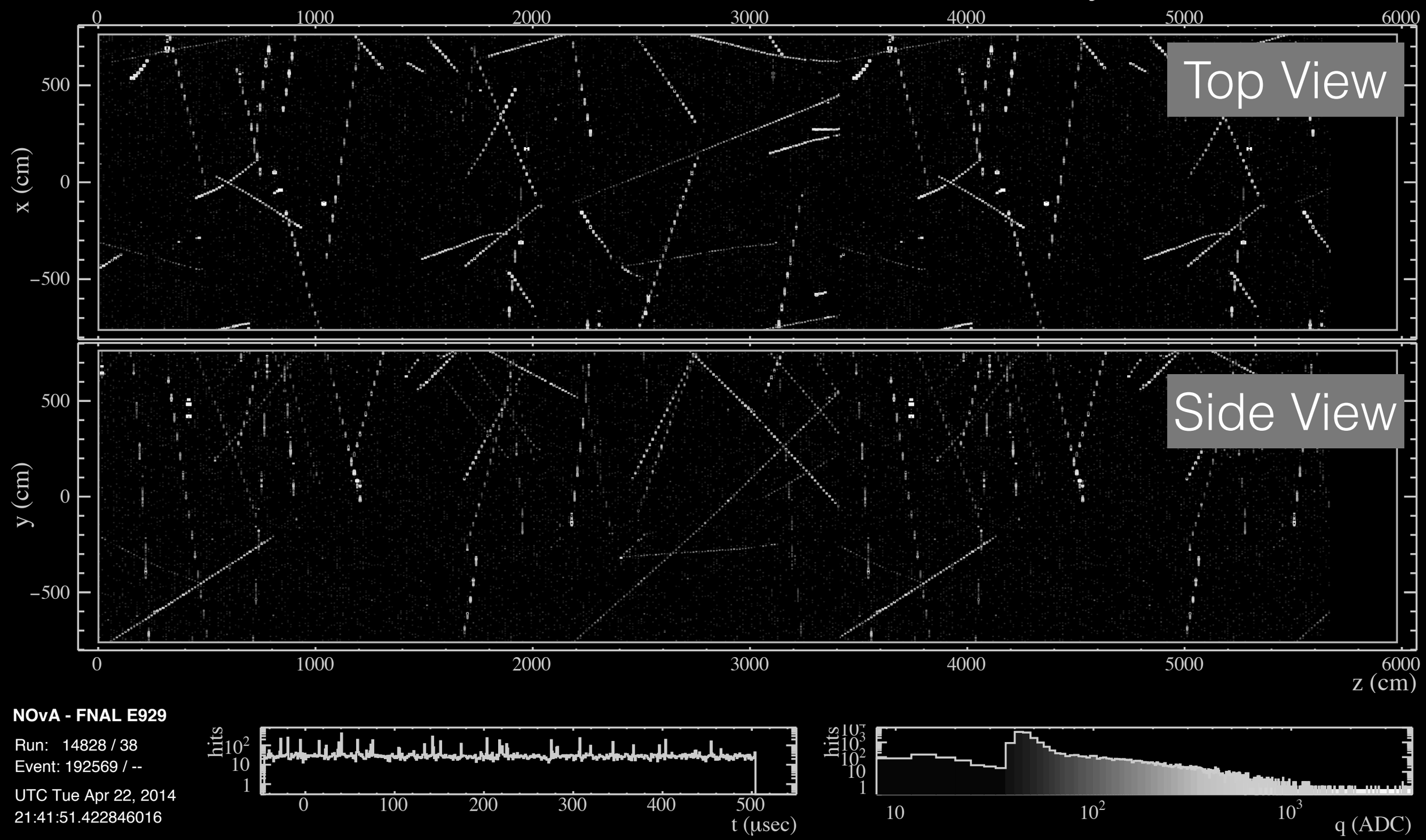




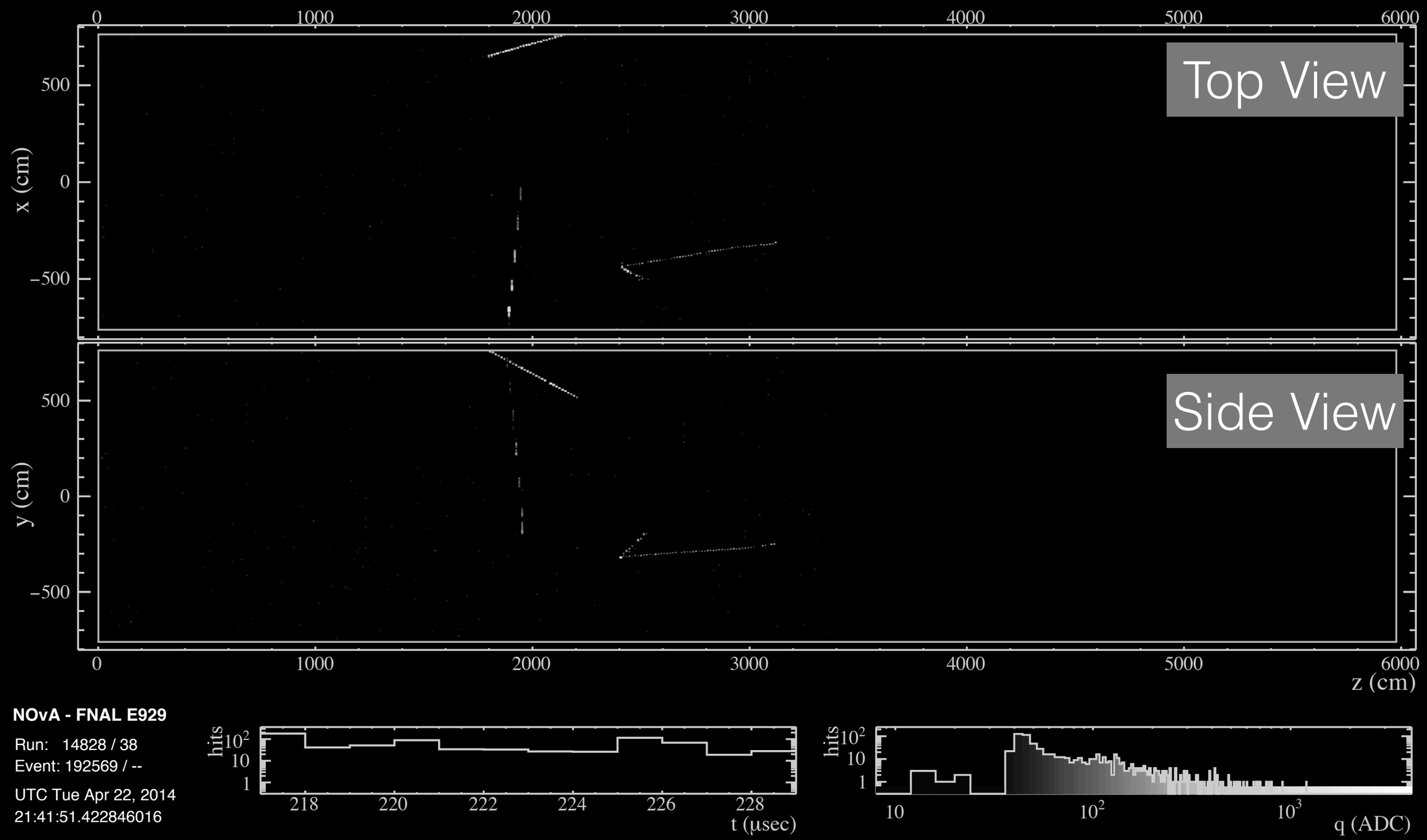


NOvA Events

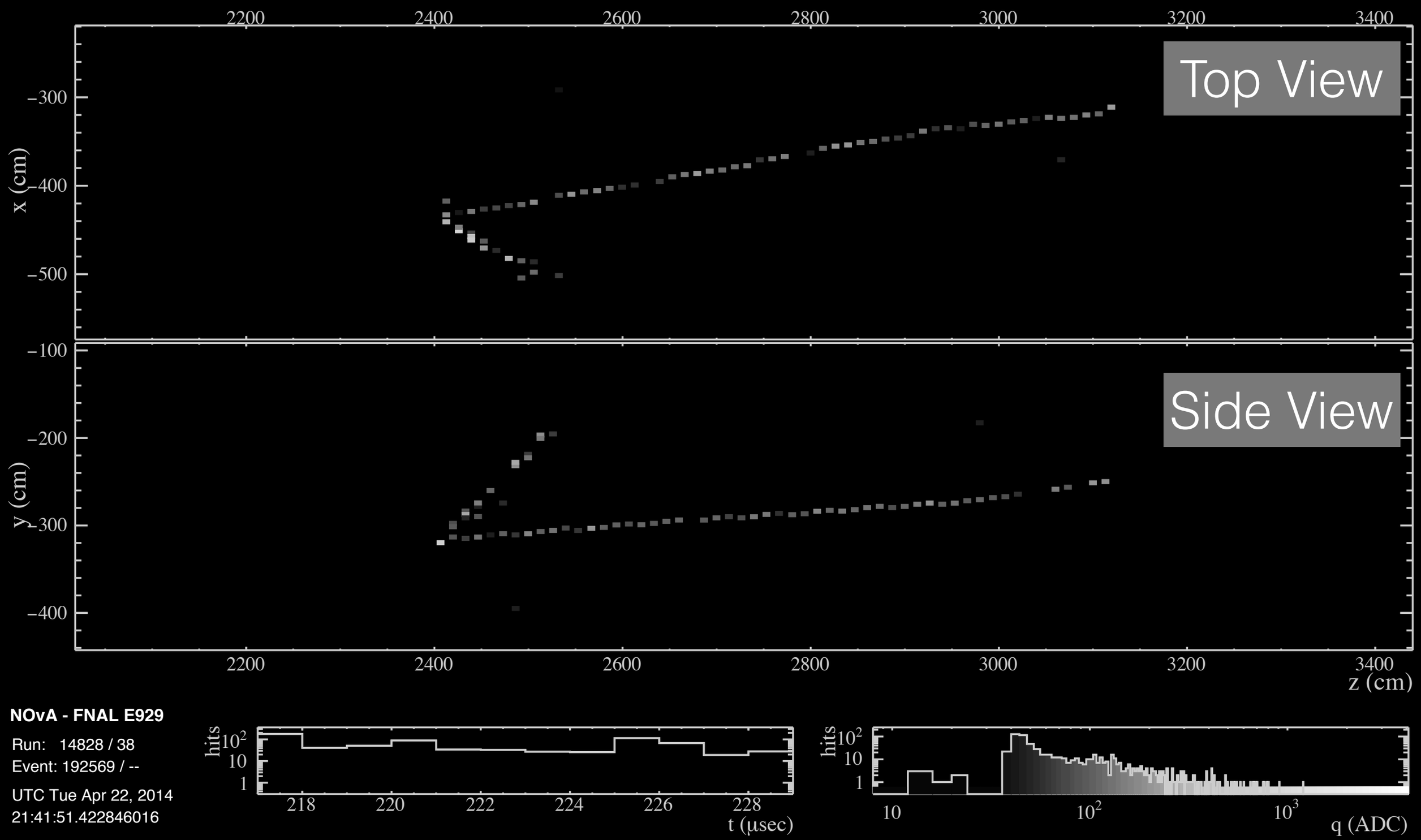
Far detector has cosmic ray flux of ~130 kHz.



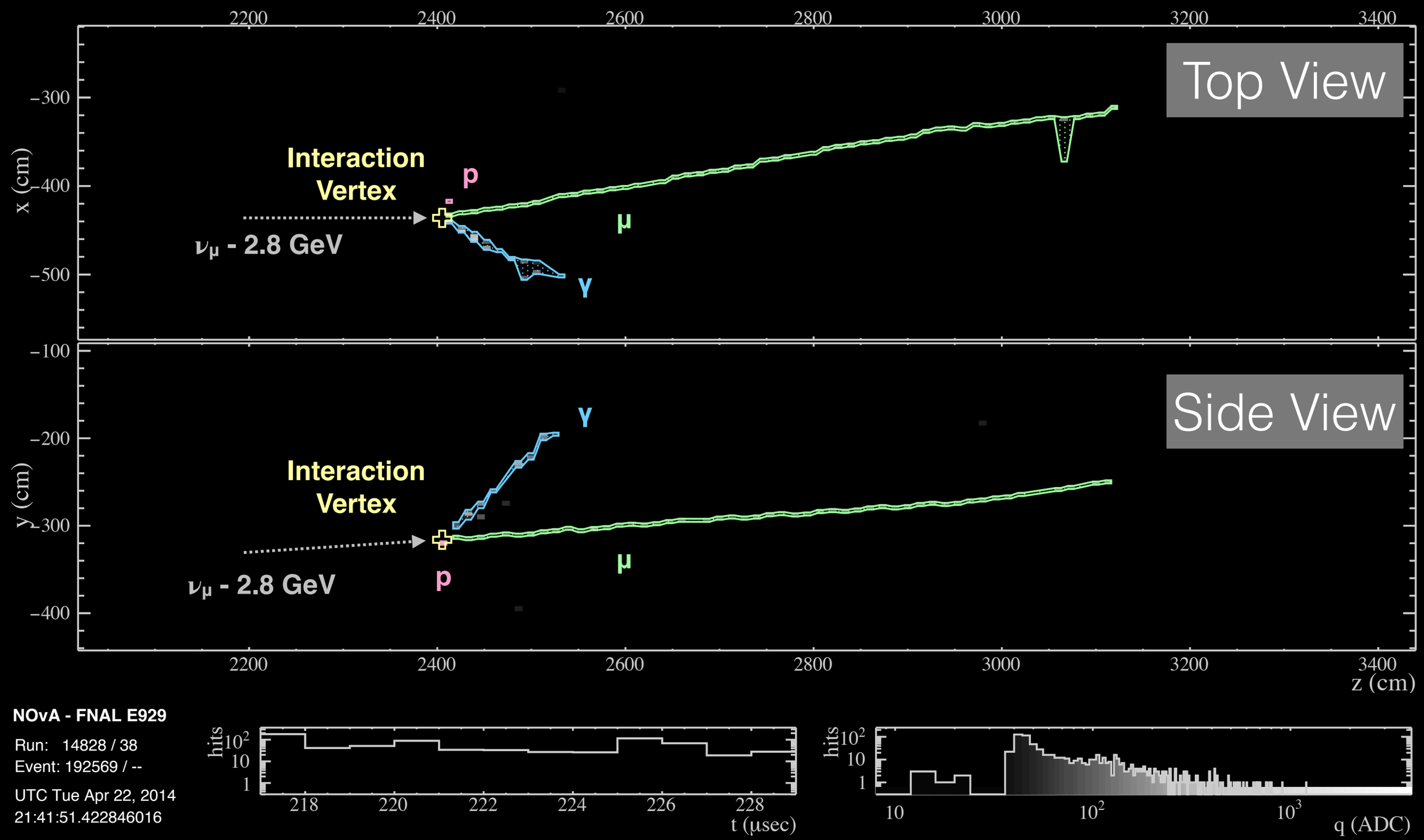
NOvA Events


Top View
Side View
x (cm)
y (cm)
z (cm)
0
1000
2000
3000
4000
5000
6000
500
0
−500
NOvA - FNAL E929
Run: 14828 / 38
Event: 192569 / --
UTC Tue Apr 22, 2014
21:41:51.422846016
hits
t (μsec)
q (ADC)

NOvA Events

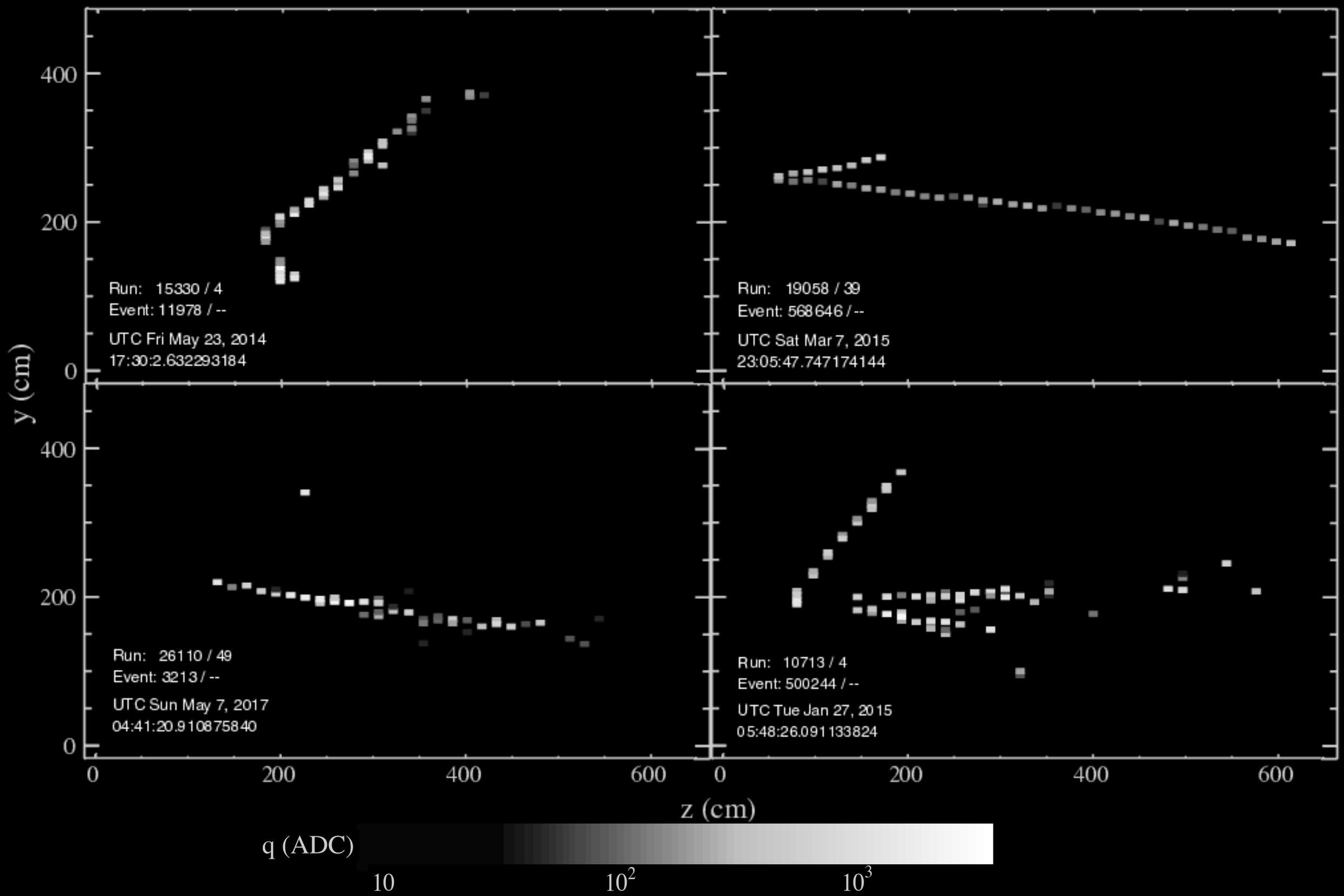

Top View
Side View
x (cm)
y (cm)
z (cm)
NOvA - FNAL E929
Run: 14828 / 38
Event: 192569 / --
UTC Tue Apr 22, 2014
21:41:51.422846016
hits
t (μsec)
q (ADC)

Reconstruction



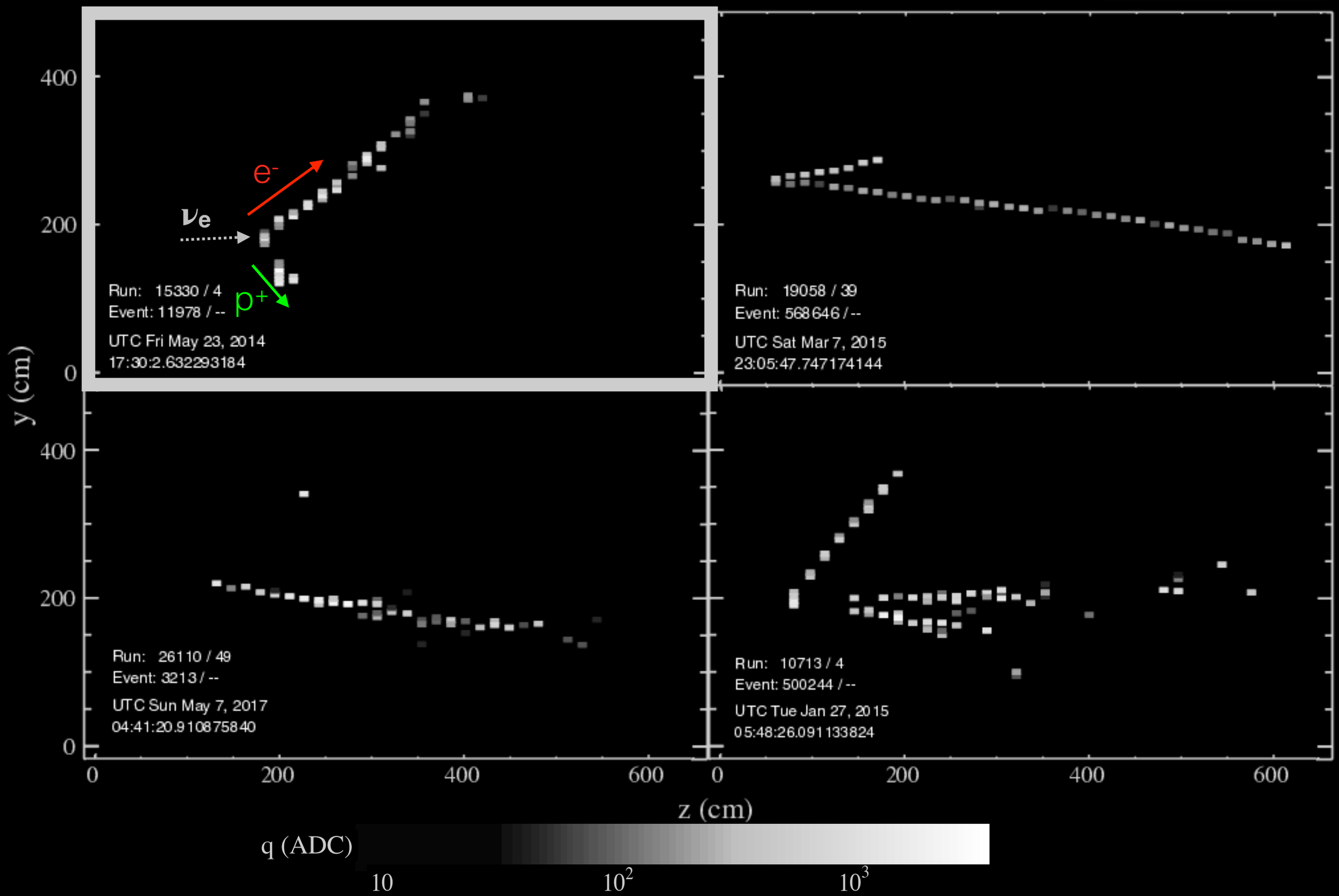
Signature Data Events

Top Views



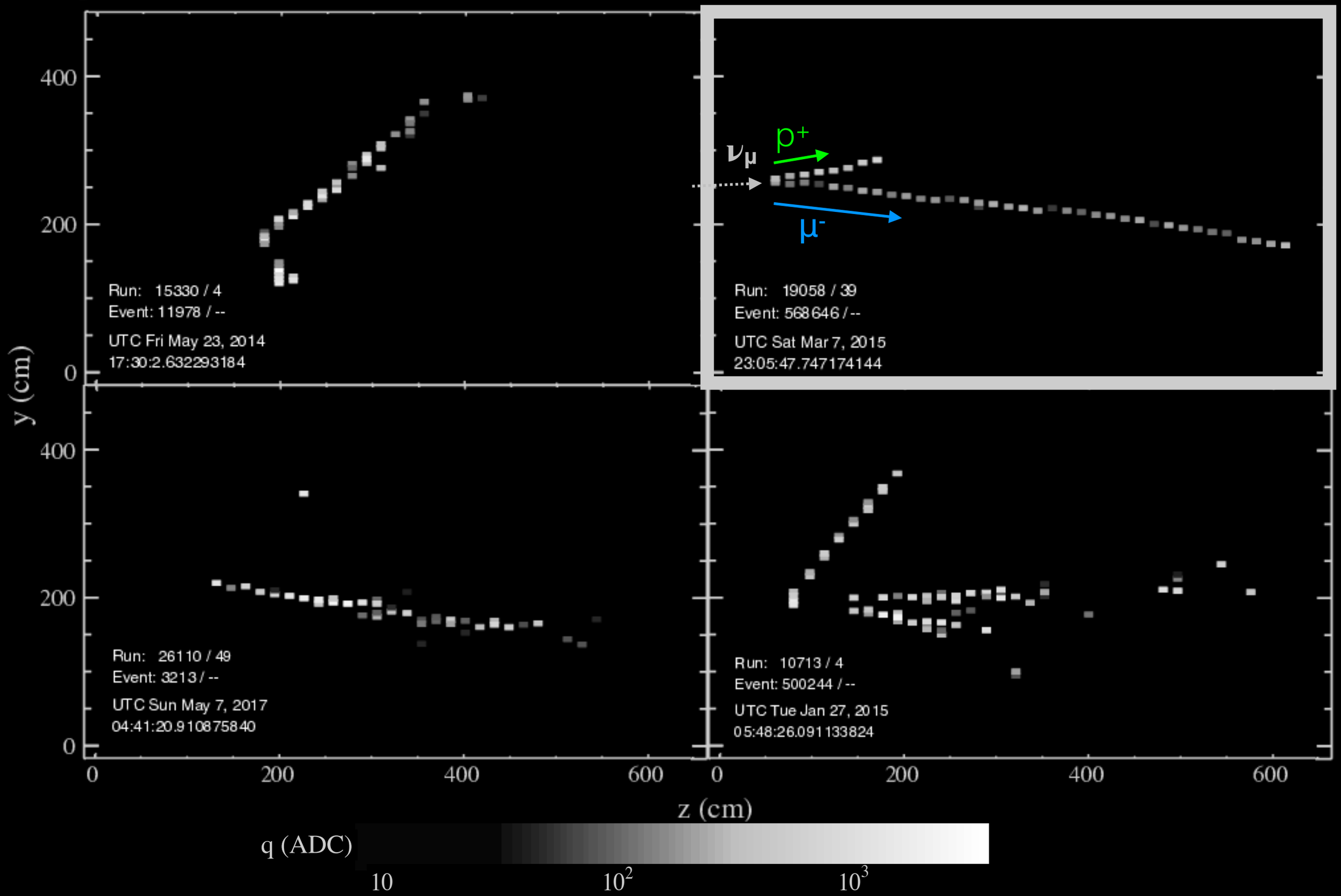
Signature Data Events

Top Views



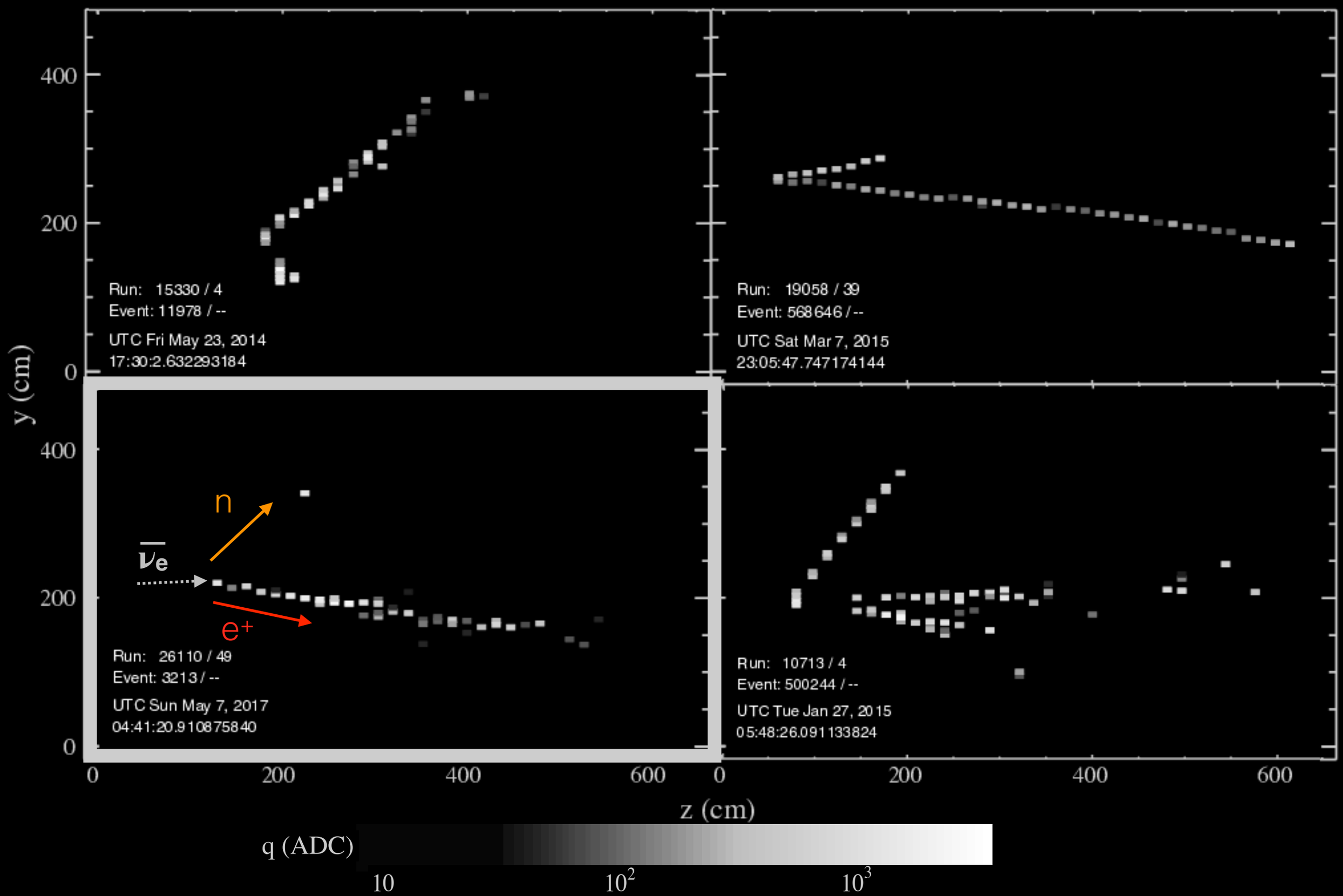
Signature Data Events

Top Views



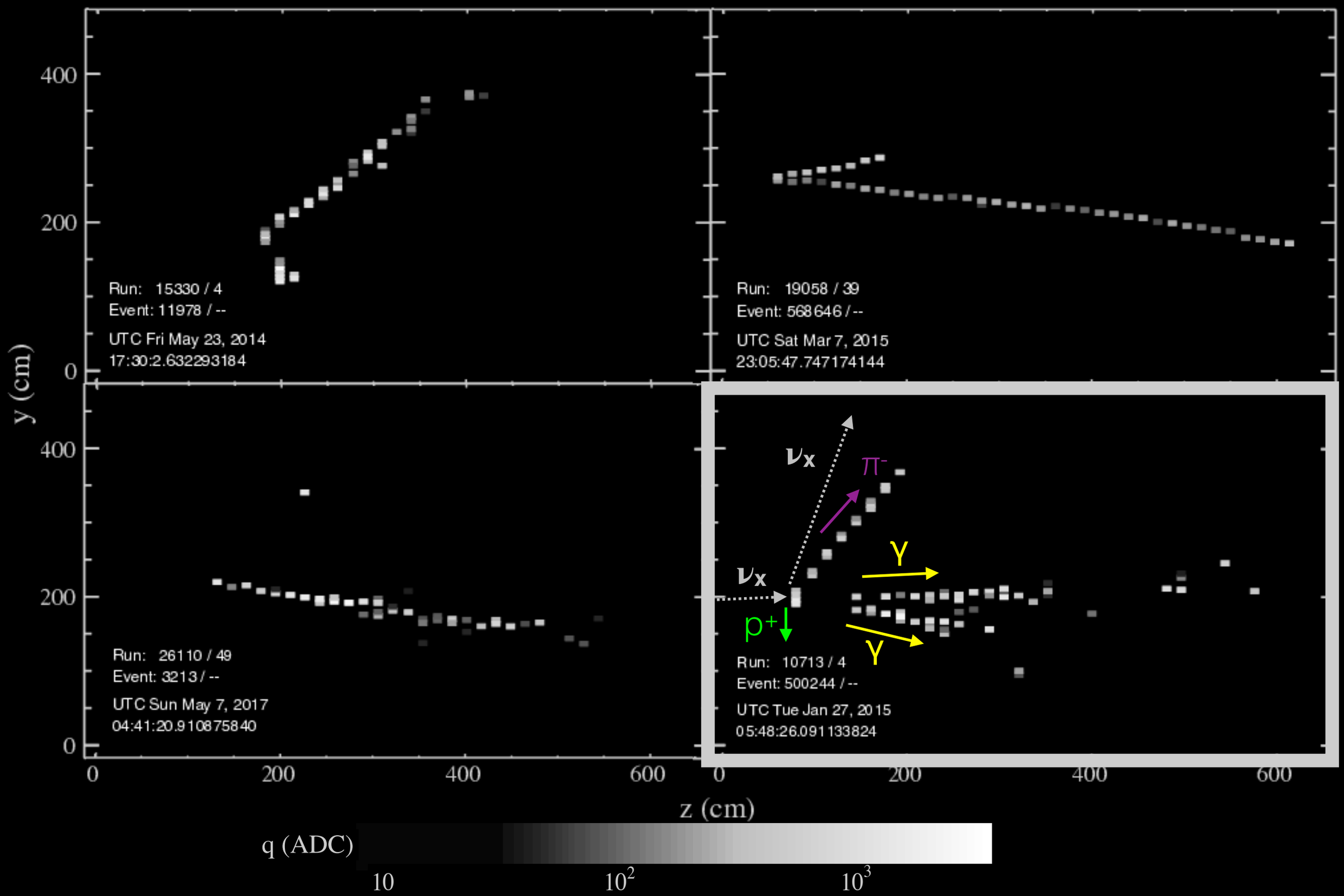
Signature Data Events

Top Views

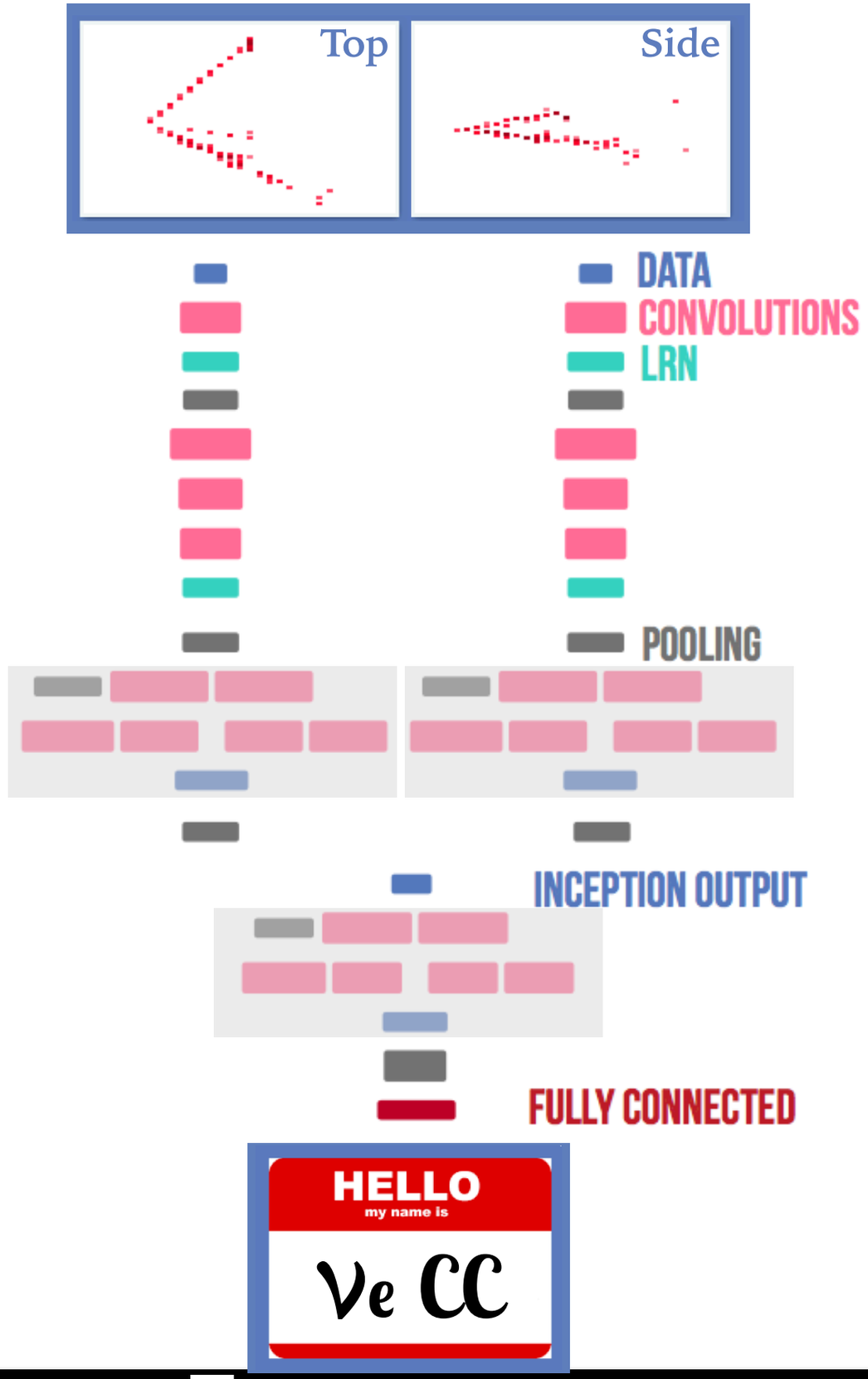


Signature Data Events

Top Views



Event Classification



Classify neutrino events using **two tower convolutional neural network**, called the Convolutional Visual Network or CVN.

Convolutional neural networks are algorithms designed to utilize topological information in images.

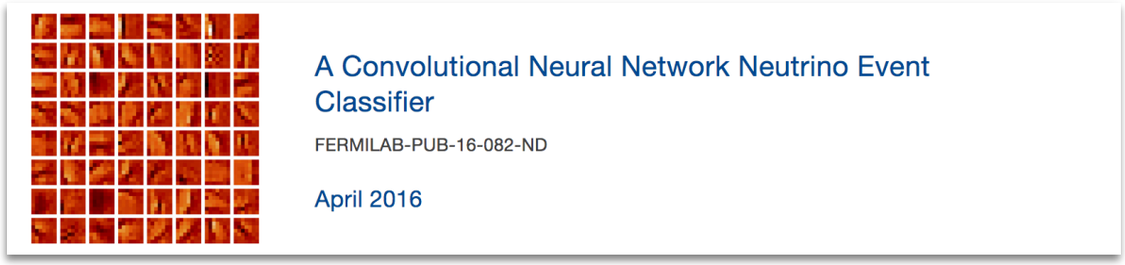
NOvA was the first experiment to apply CNNs to a HEP result in its 2016 oscillation analysis.

Yielded an effective **30% increase in exposure.**

Has now been used in our last four oscillation analyses.

Updates in the latest analyses:

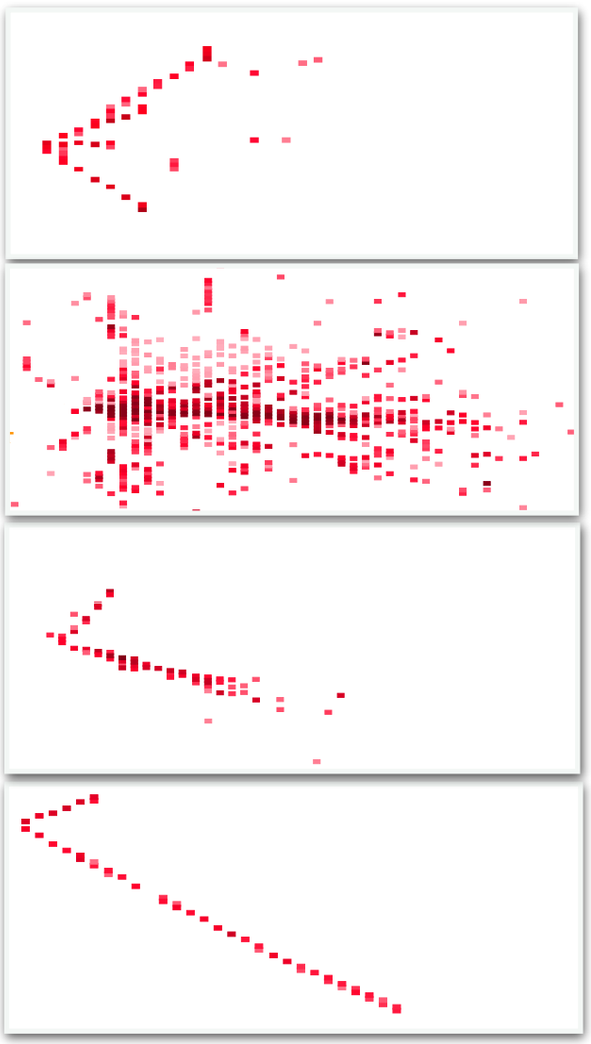
- Updated simulation.
- Classification is done using final states.
- Architecture optimizations.

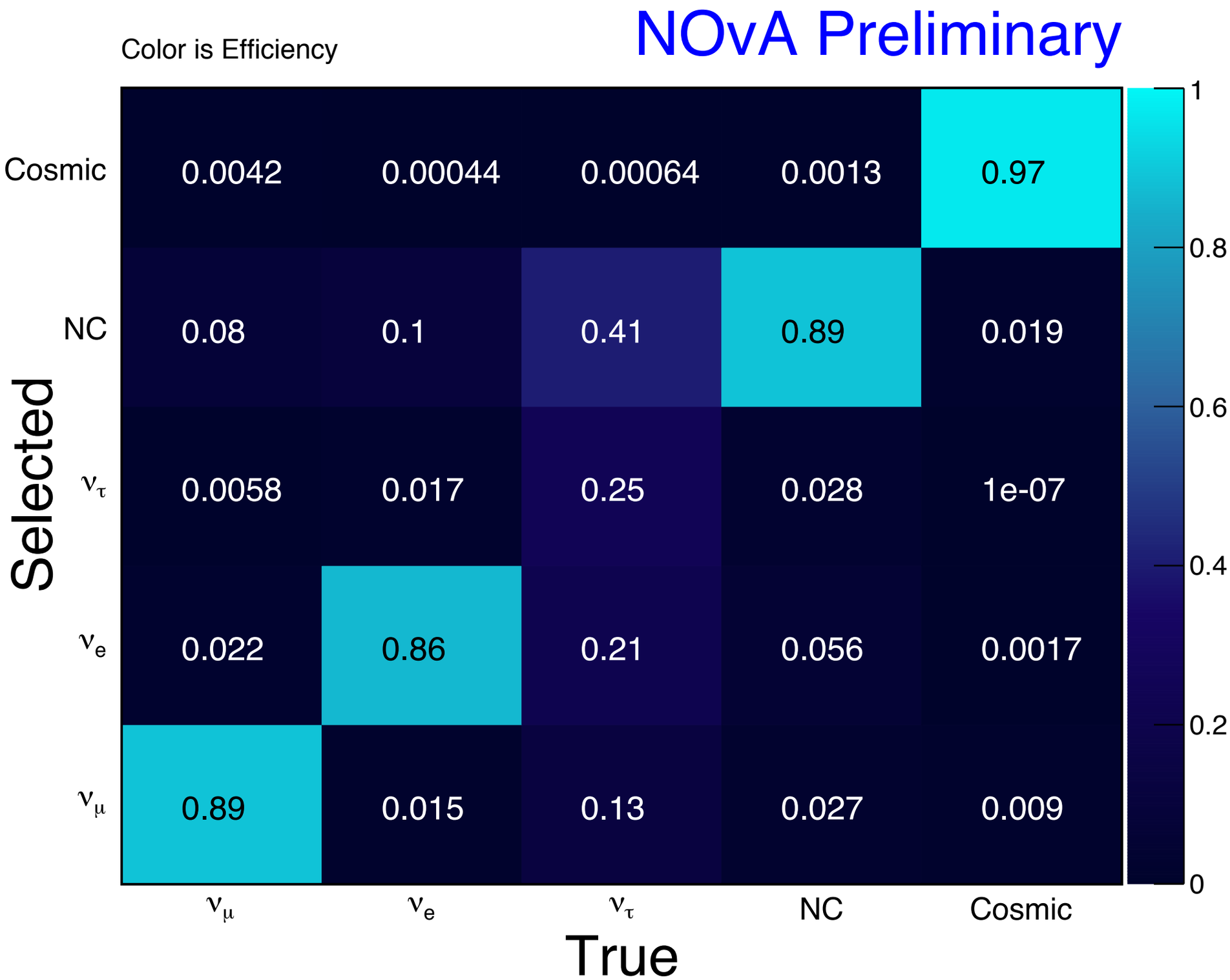

A Convolutional Neural Network Neutrino Event Classifier

FERMILAB-PUB-16-082-ND

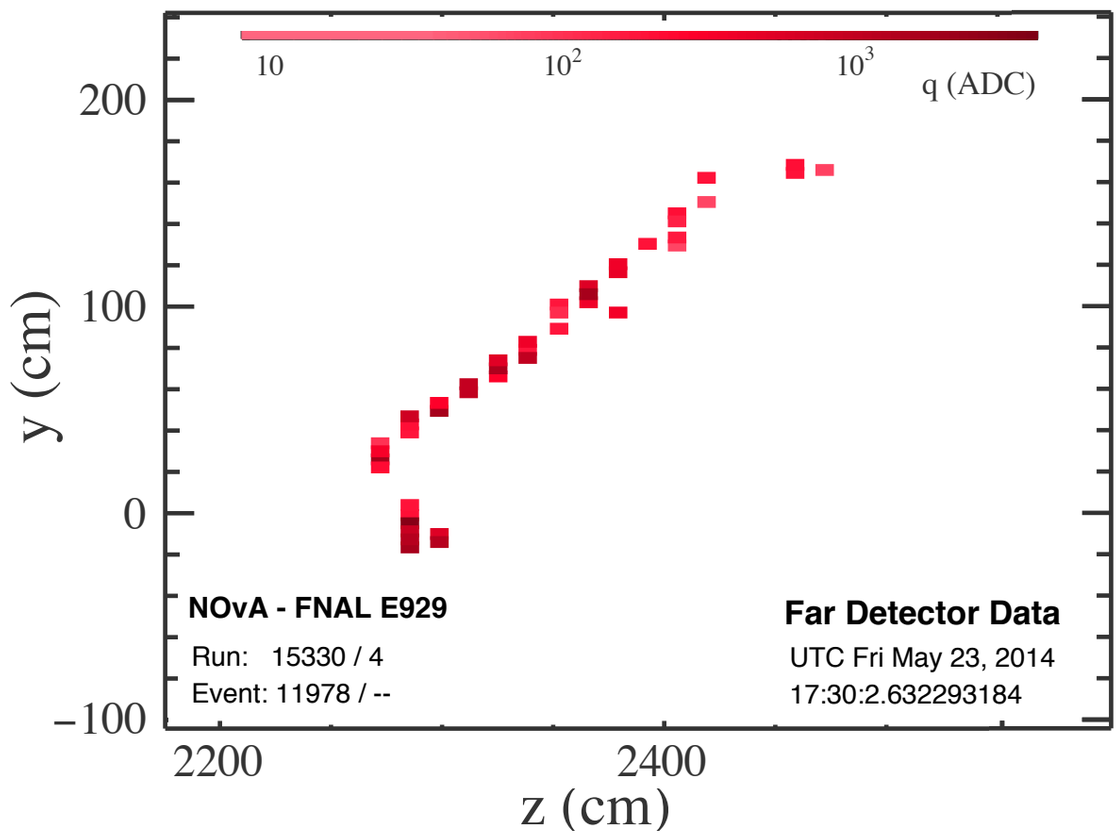
April 2016

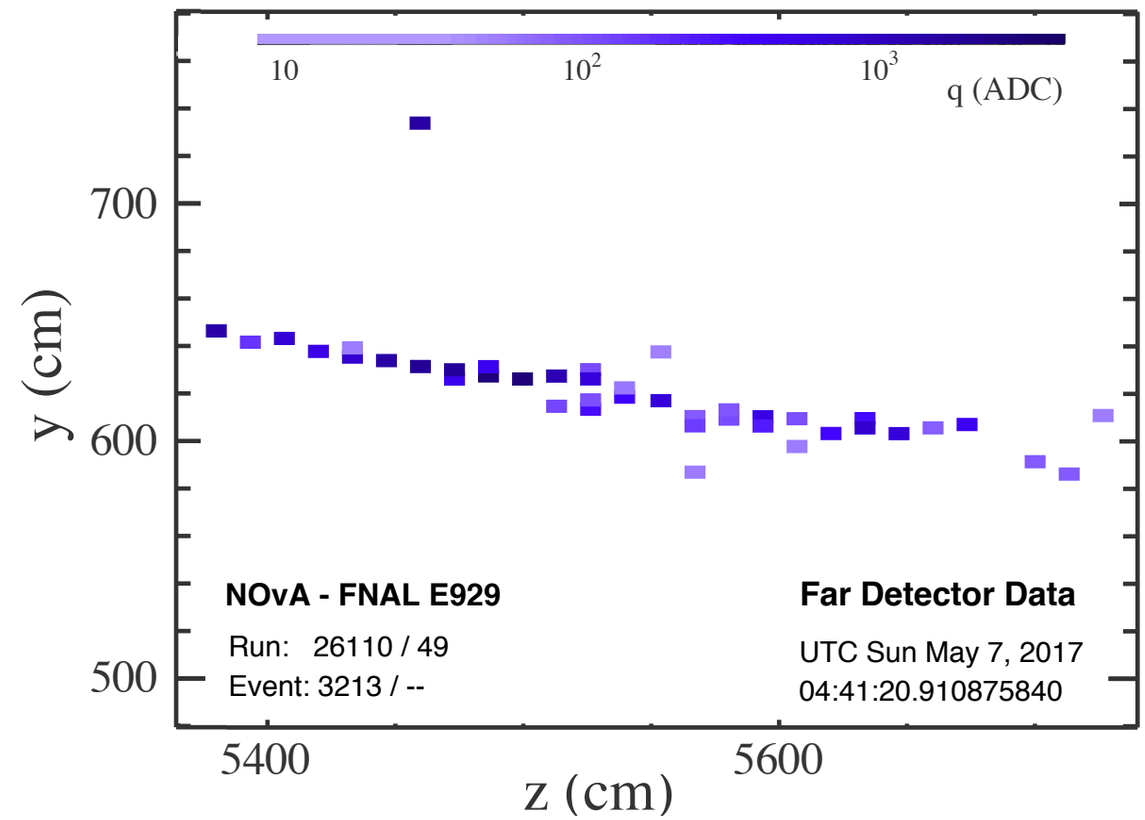
Event Classification





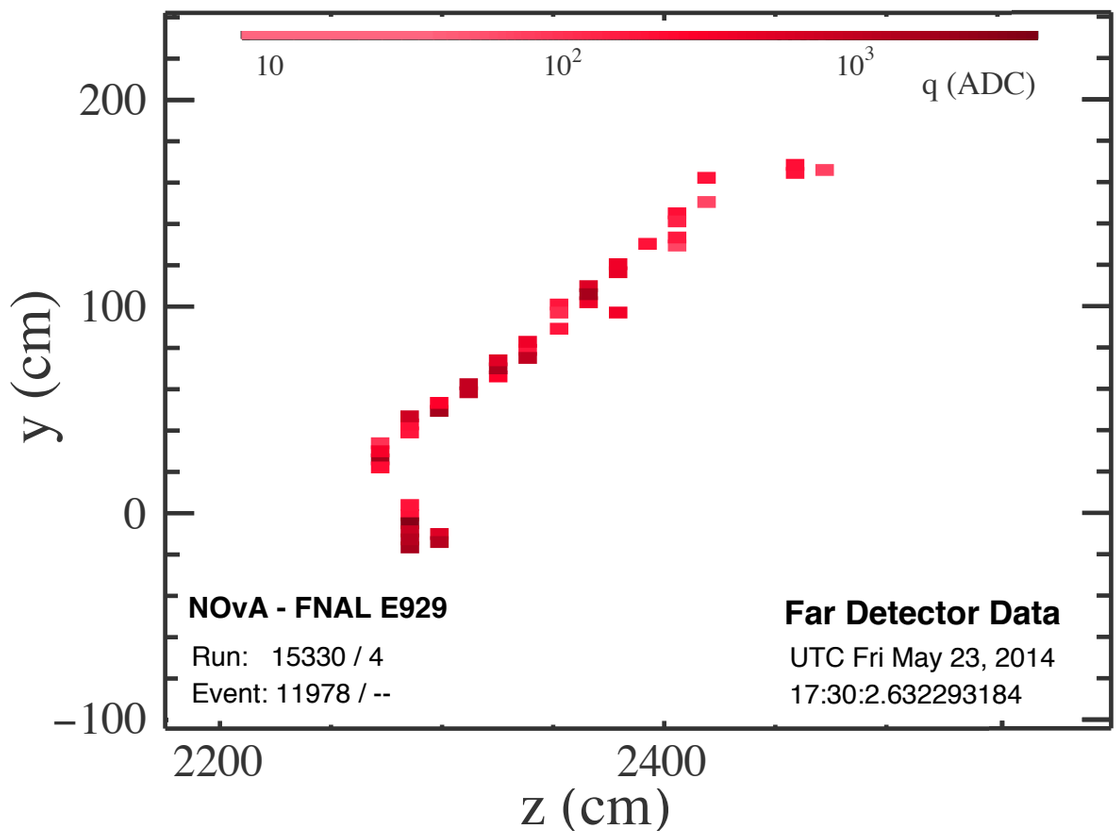
Neutrinos and Anti-Neutrinos

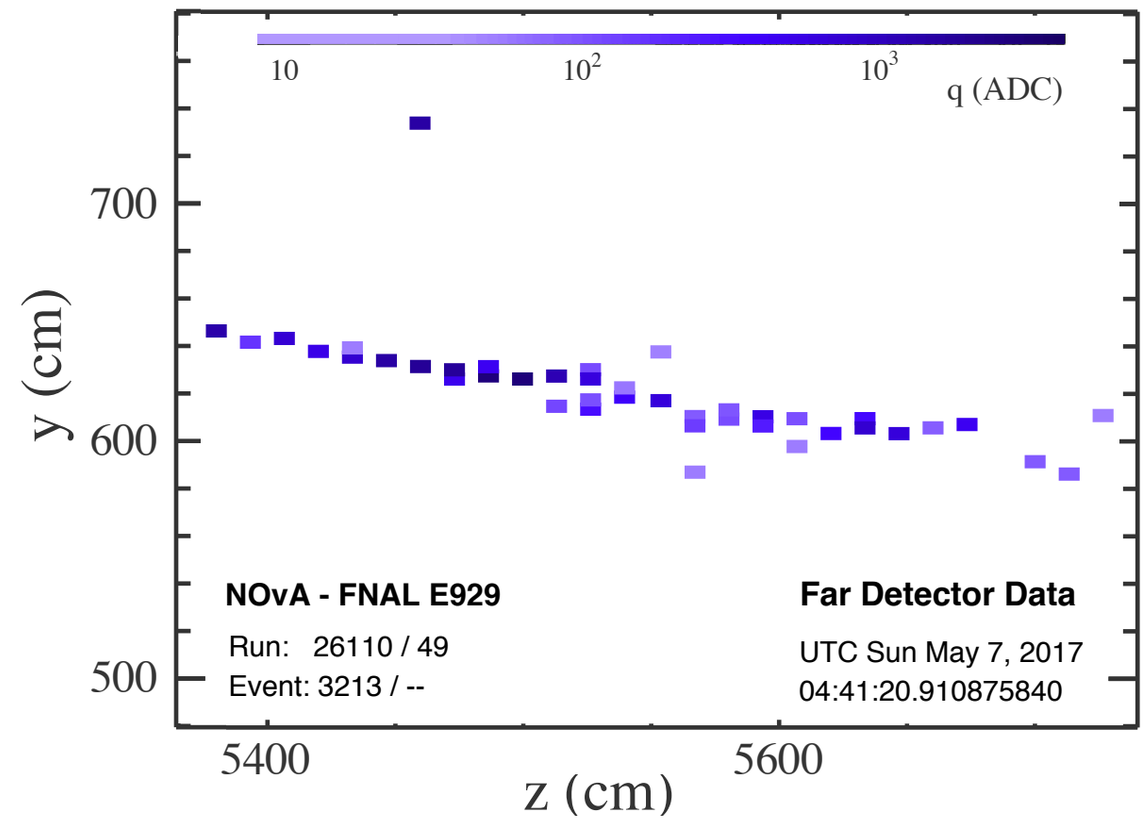




The topology of **neutrino** and **anti-neutrino** interactions are different on average.

Neutrinos and Anti-Neutrinos





Train two networks on **neutrino** beam and **anti-neutrino** beams separately.

The networks can utilize these differences in event topology.

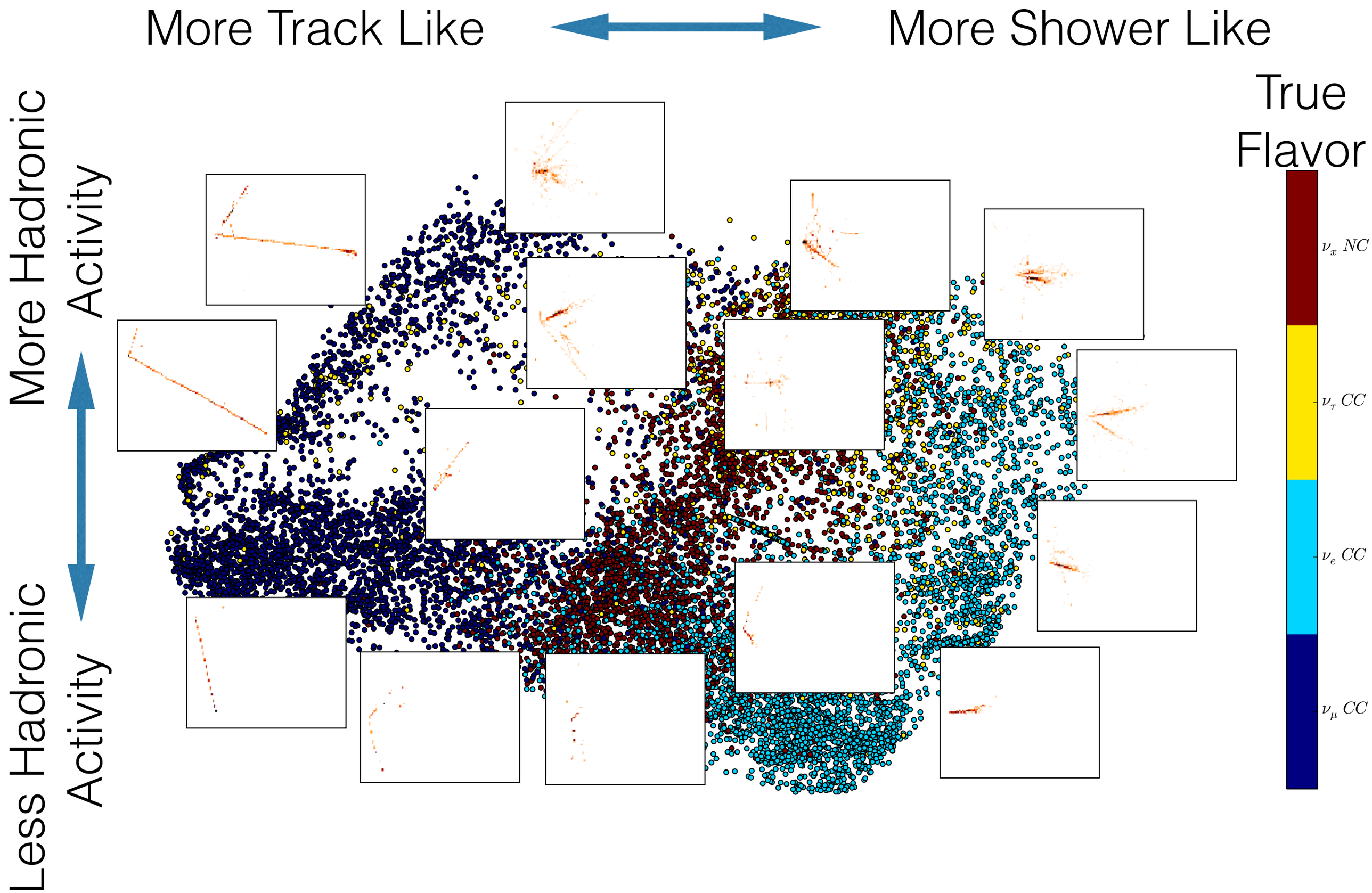
$\bar{\nu}$ **Efficiency Improvement** Training Sample (ID > 0.9)		
$\bar{\nu}_e$ CC Signal	$\bar{\nu}_\mu$ CC Signal	$\bar{\nu}$ NC Signal
14%	6%	10%

t-SNE

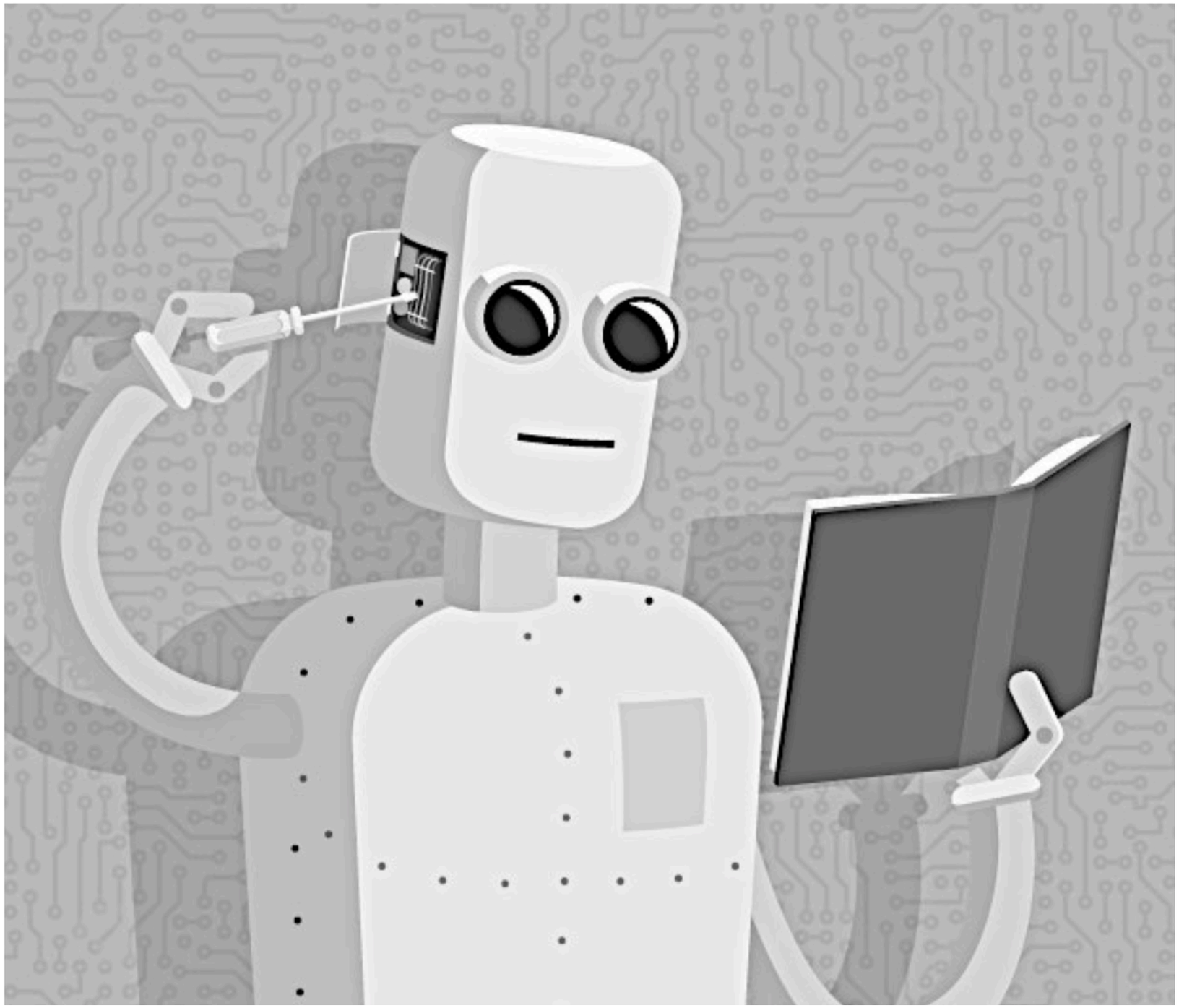
t-Distributed Stochastic Neighbor Embedding

https://indico.io/blog/visualizing-with-t-sne/

https://www.nature.com/articles/s41586-018-0361-2


More Track Like
More Shower Like
More Hadronic Activity
Less Hadronic Activity
True Flavor
ν_x NC
ν_τ CC
ν_e CC
ν_μ CC

For now, let’s discuss some algorithms in development for future analyses.



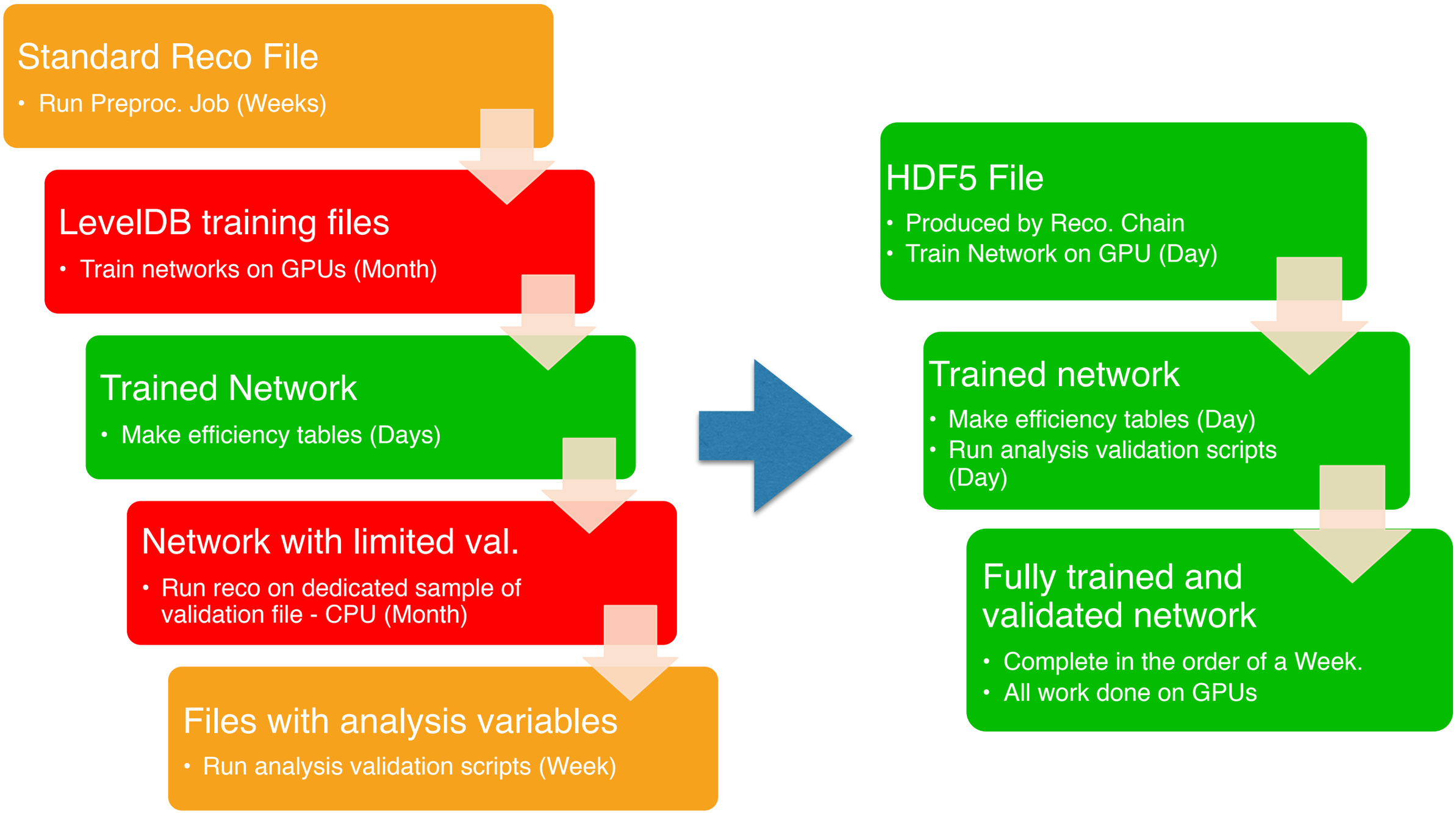
ML Infrastructure

We recently revamped our machine learning infrastructure.

With the help of SciDAC at Fermilab, we have a new file format, hdf5s.

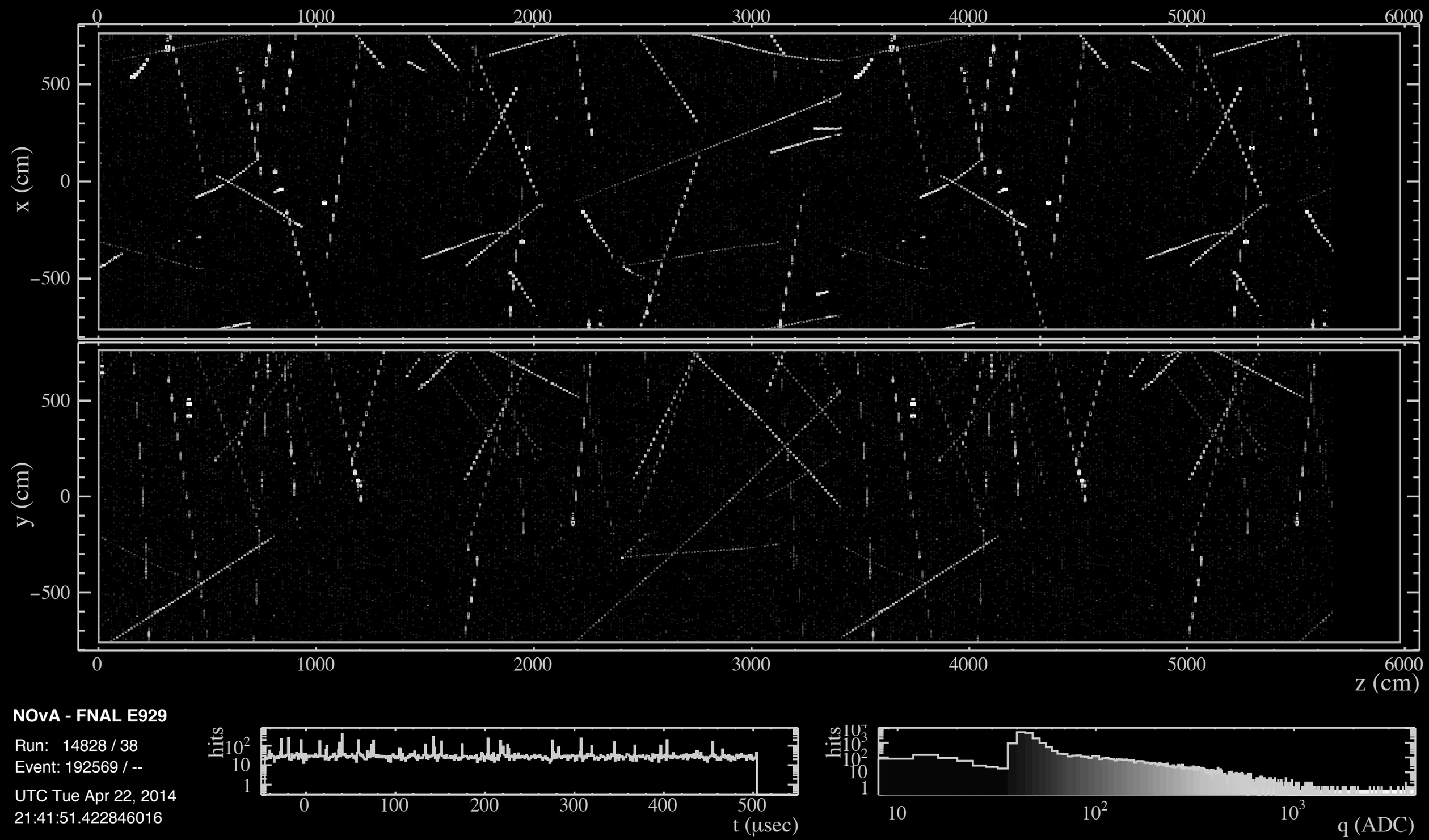
Changed from leveldbs, root, and caffe to now use hdf5, pandas, and tensorflow.

Tensorflow runs ~7x faster in our production environment than caffe.

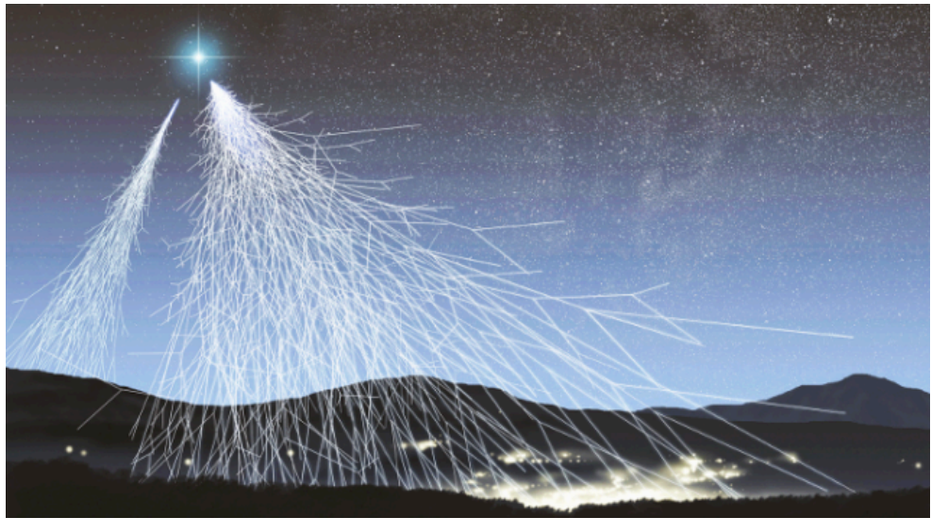


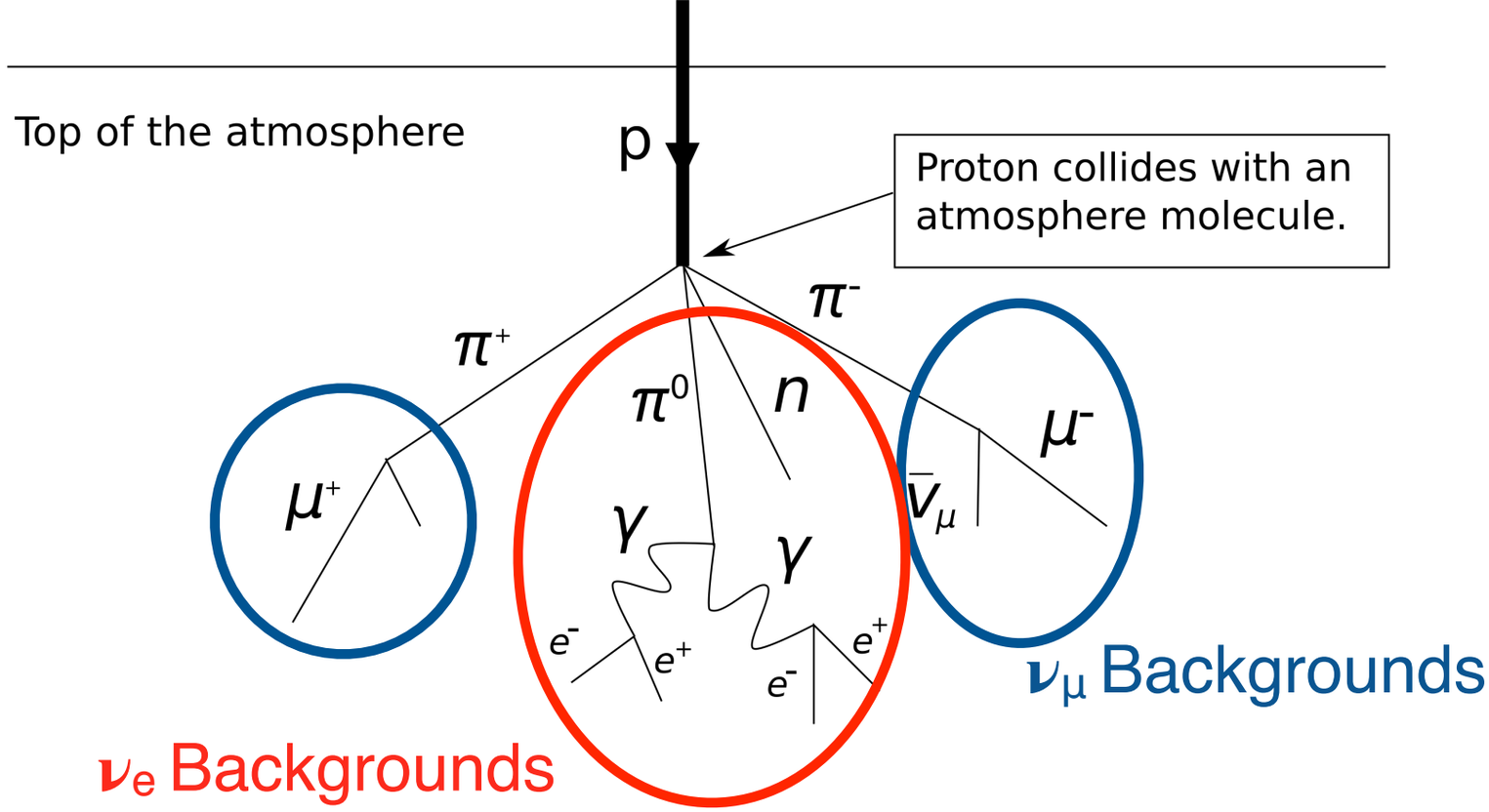
Cosmic Rejection

Far detector has cosmic ray flux of ~130 kHz.



Cosmic Rejection





Cosmic rays make a large background in the oscillation analysis.

Predicting this background requires processing a large number of cosmic events.

1 ν_μ background per ~2.5×10^7 cosmic rays

1 ν_e background per ~1.5×10^7 cosmic rays

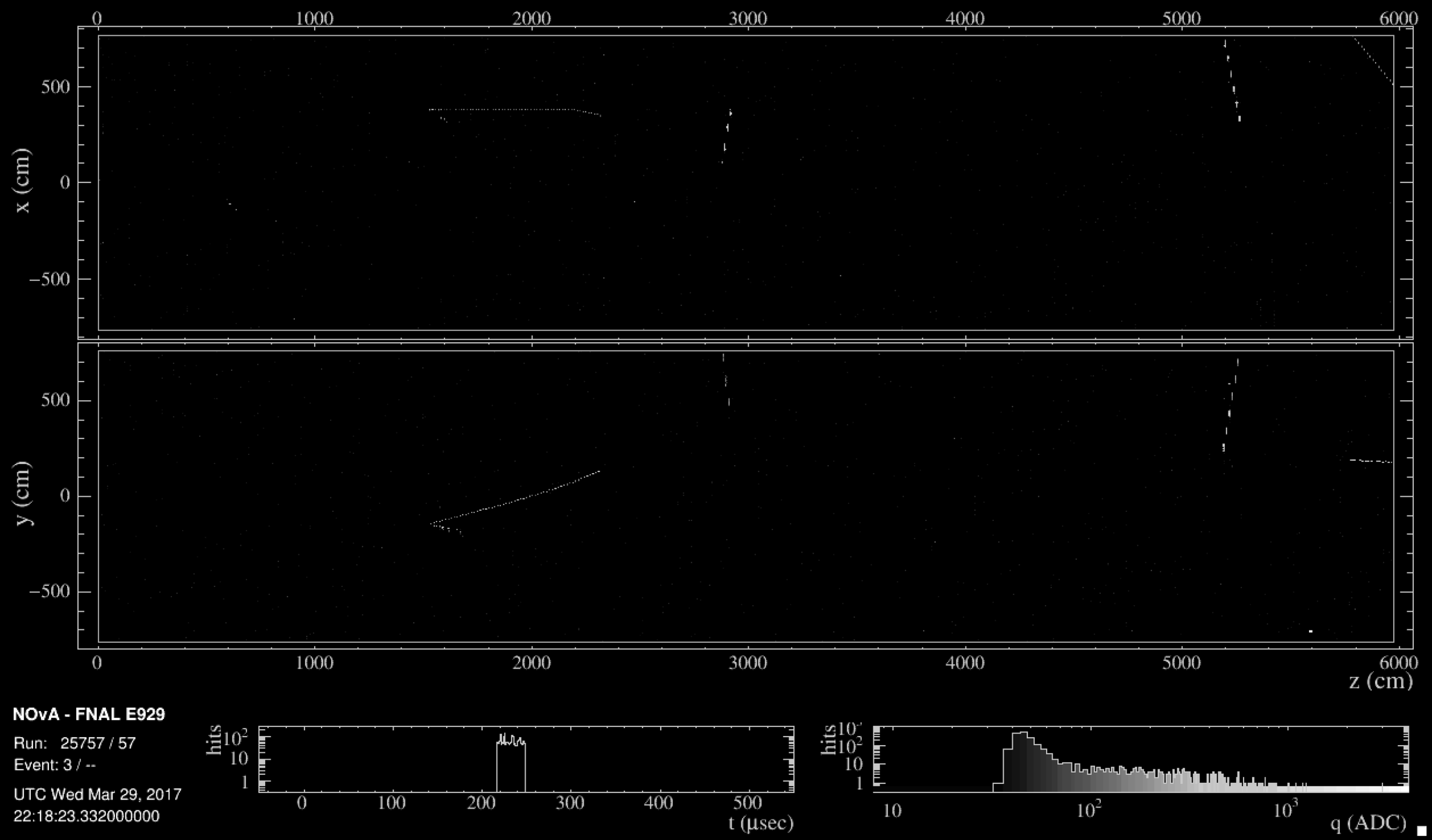
We have ~6 years of cosmic data on disk, almost 10^{11} cosmic rays.

Fully processing all that data with current techniques would take about 2 years.

Instead, use a CNN to evaluate full readout on wether they contain a neutrino or not.

Based on resnet-18, the network processes an entire readout in time chunks very early in processing stages.

Network has no dependencies besides the pattern of hits.



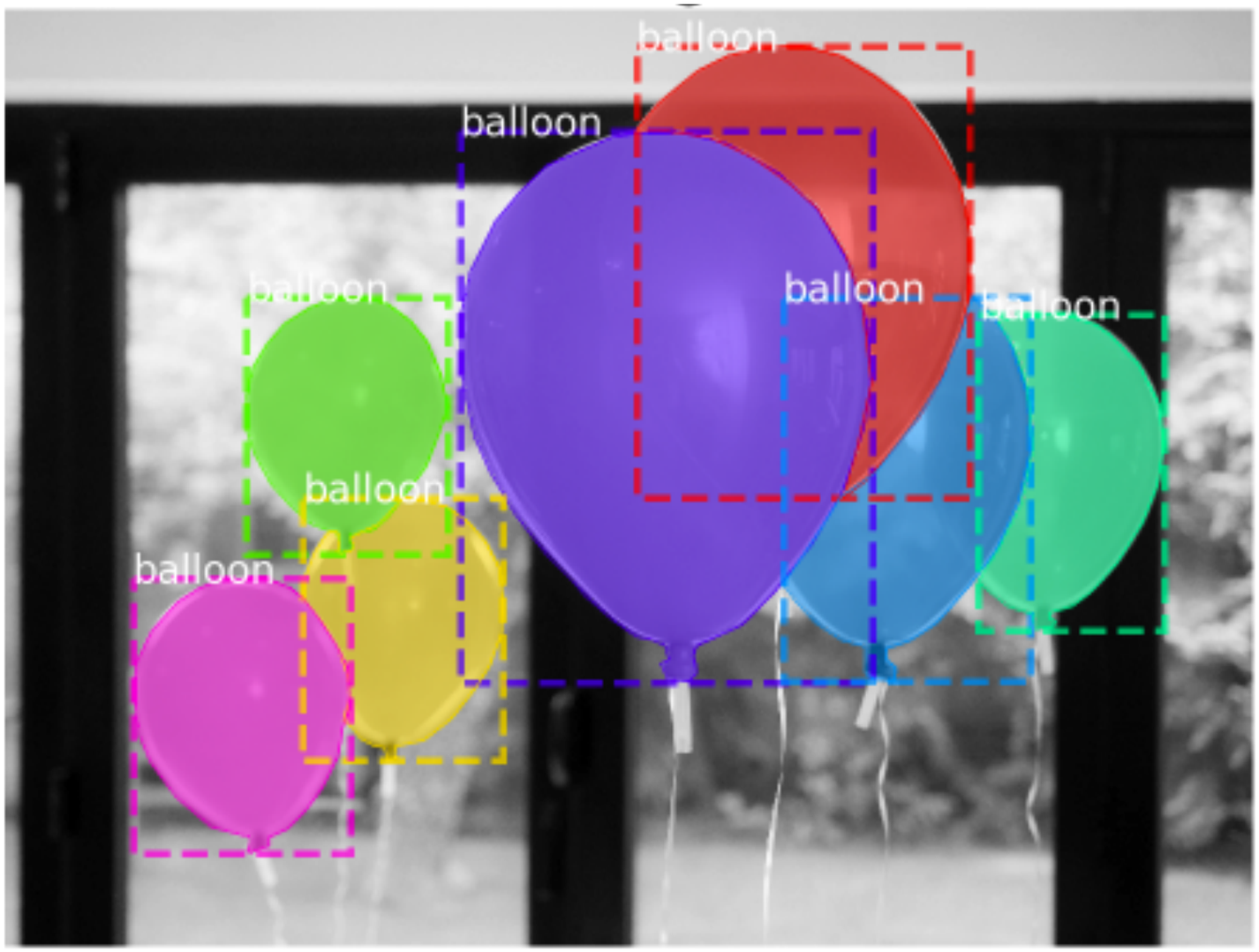
Full Event Reconstruction

Cluster and **classify** objects simultaneously using **instance aware semantic segmentation**.

Use machine learning to reconstruct an event **hit by hit**.

Three outputs:

1. **Bounds**
2. **ID Score**
3. **Clusters**

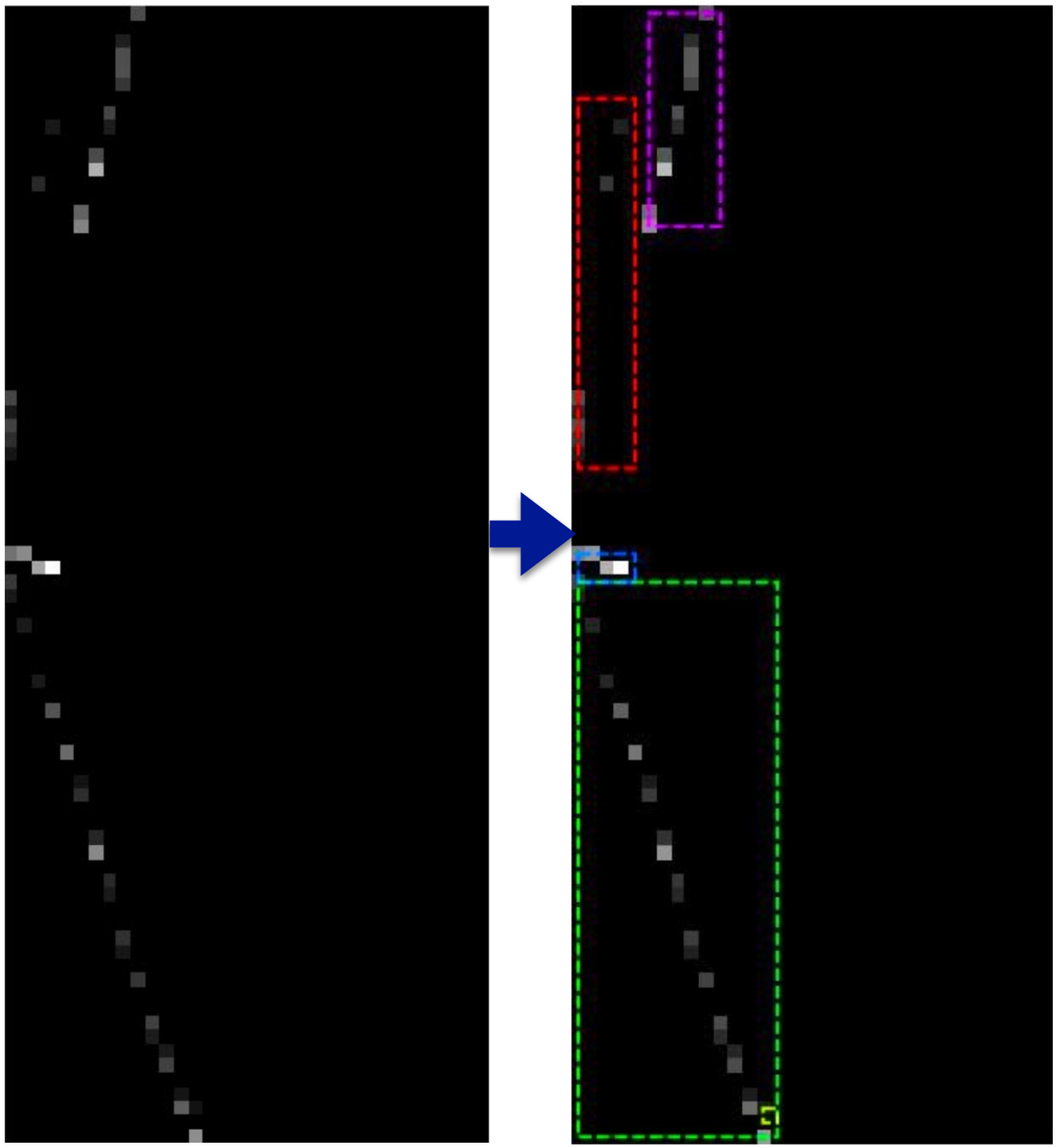


Using an implementation of Mask R-CNN:
K. He, G. Gkioxari, P. Dollar, and R. Girshick. Mask R-CNN. arXiv:1703.06870, 2017.

Full Event Reconstruction

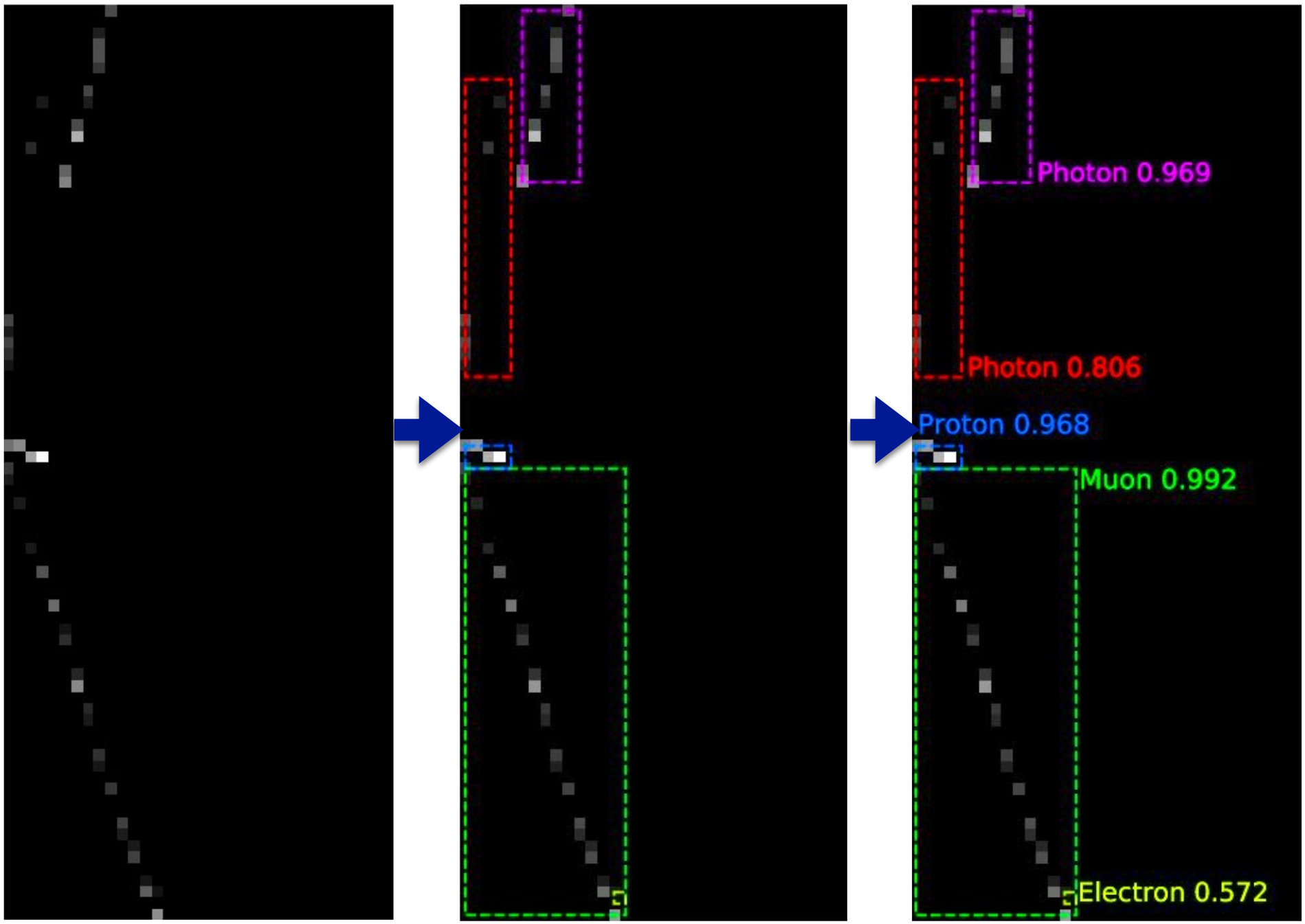
Full Event Reconstruction

Bounds - Look for individual particles within the event and construct bounding boxes containing each.



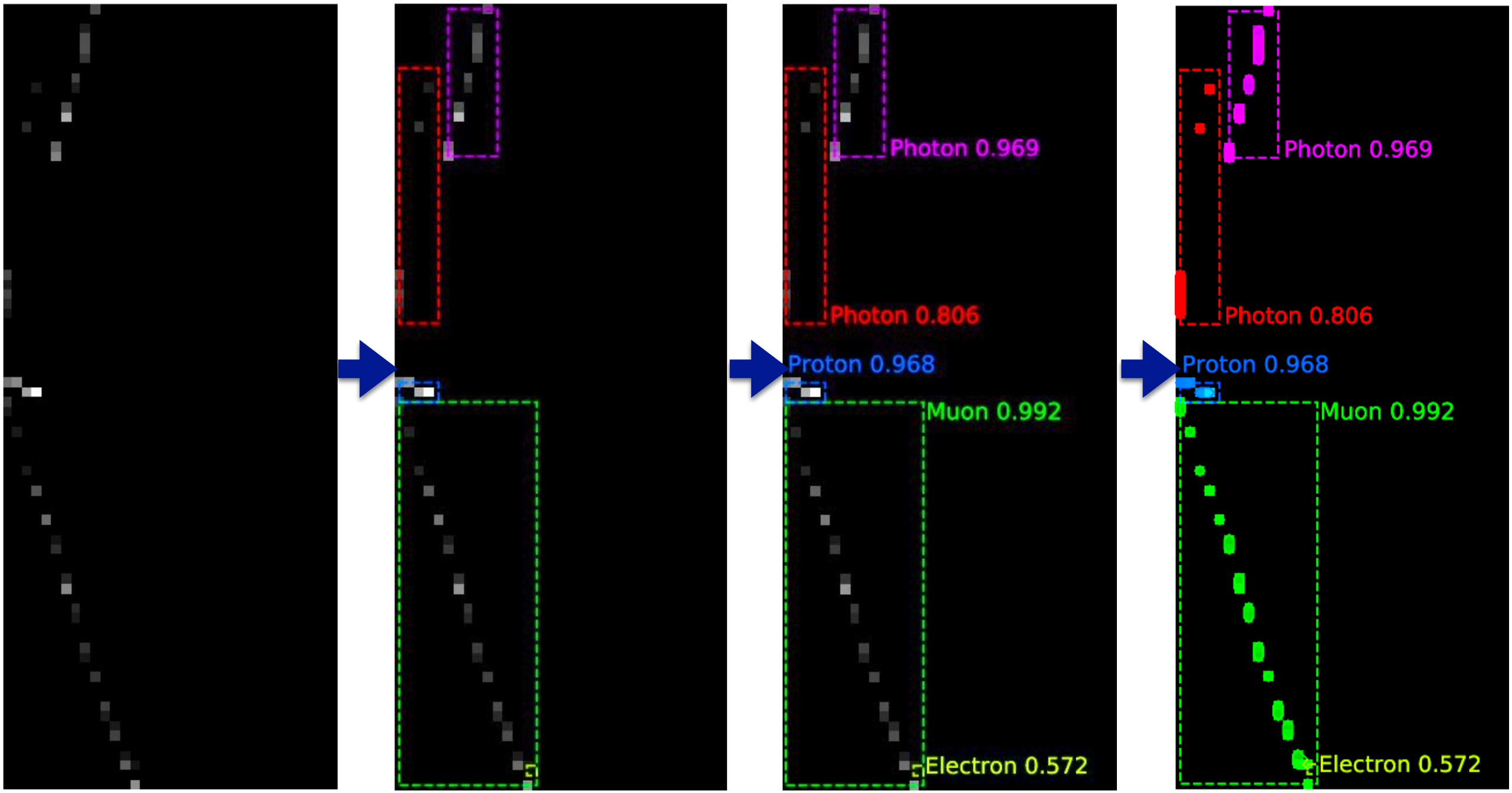
Full Event Reconstruction

ID Score - Use a softmax function to classify the particle contained within each box.

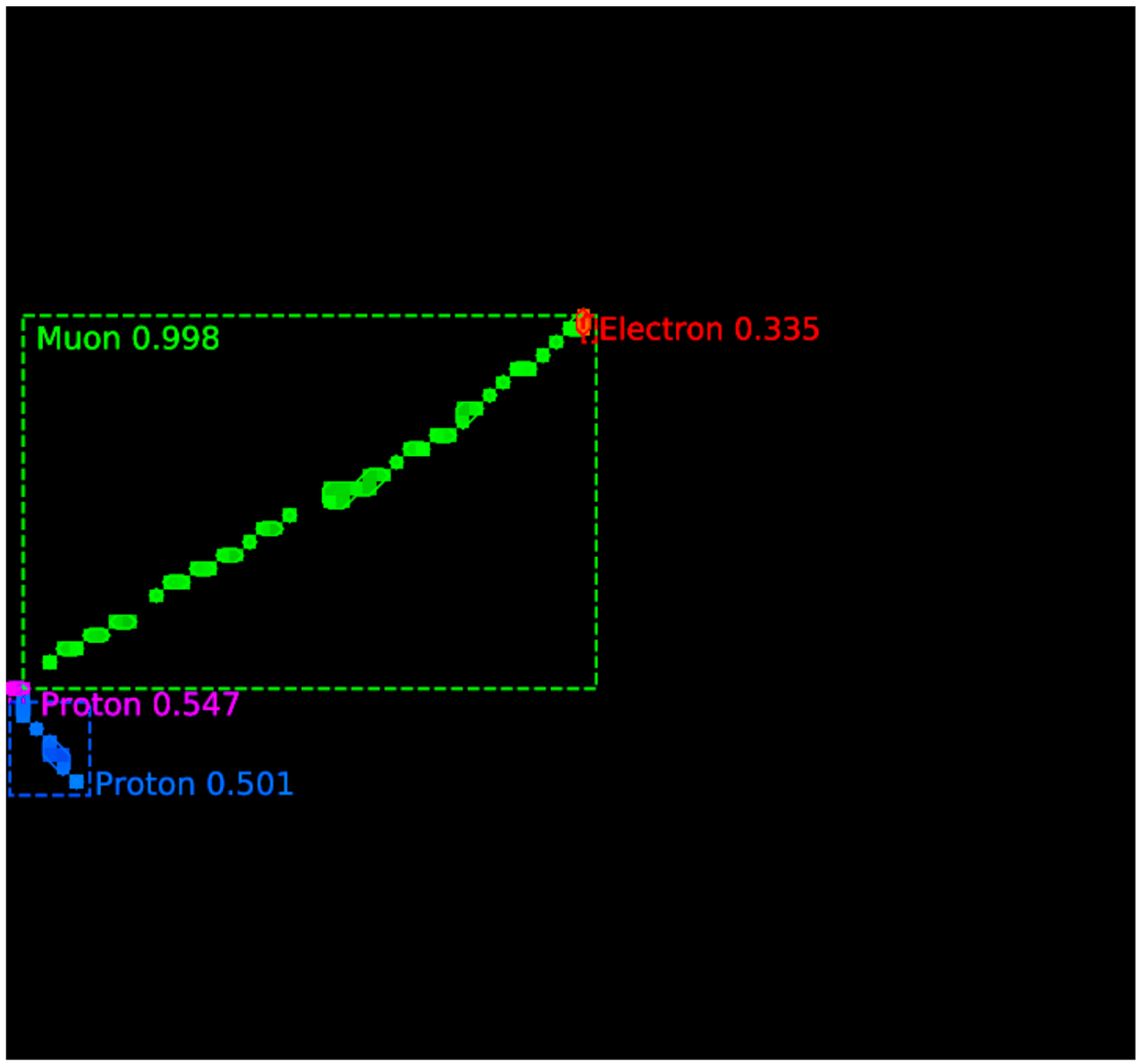


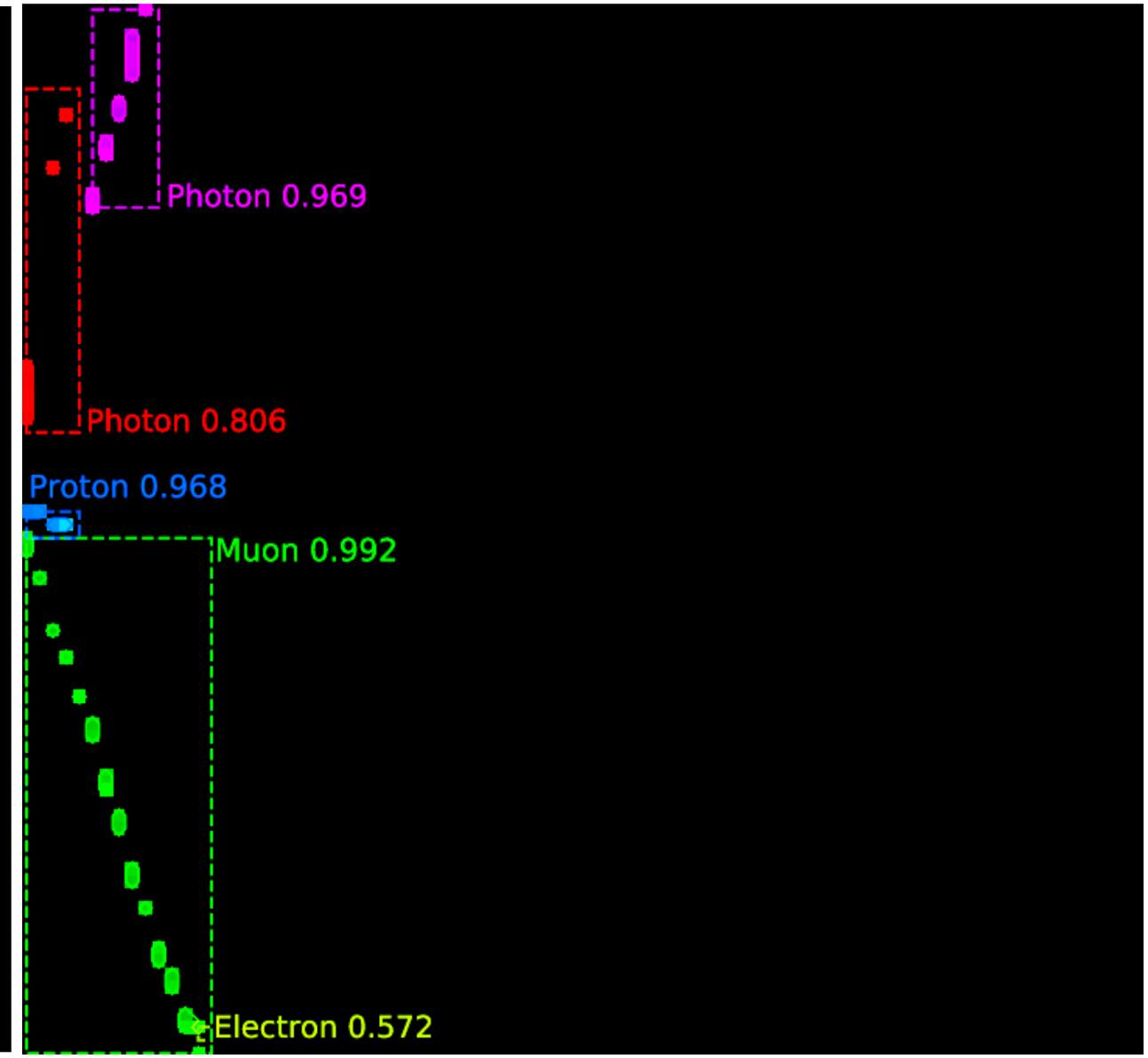
Full Event Reconstruction

Clusters - Group together hits within each box to make clusters for each particle.

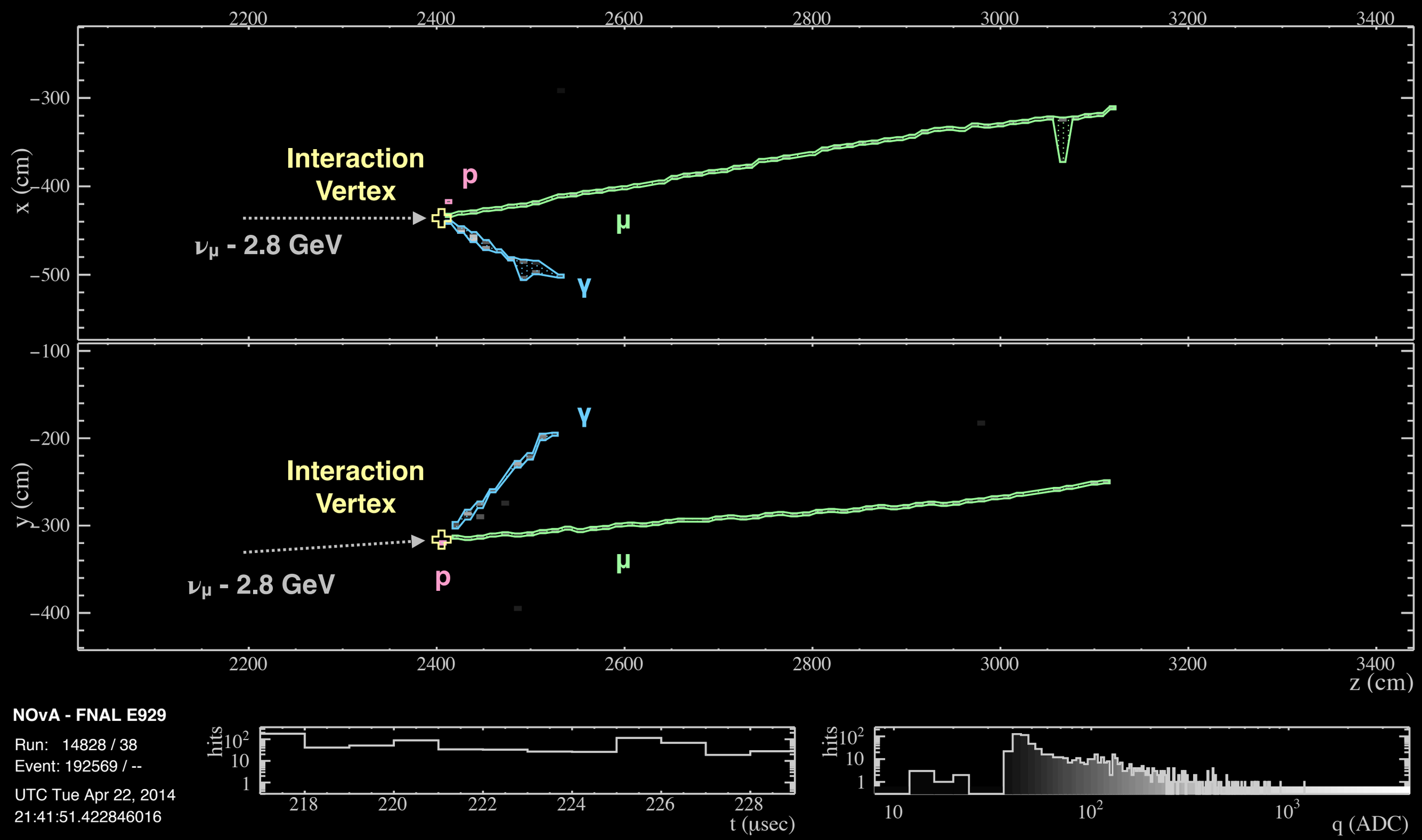


Full Event Reconstruction

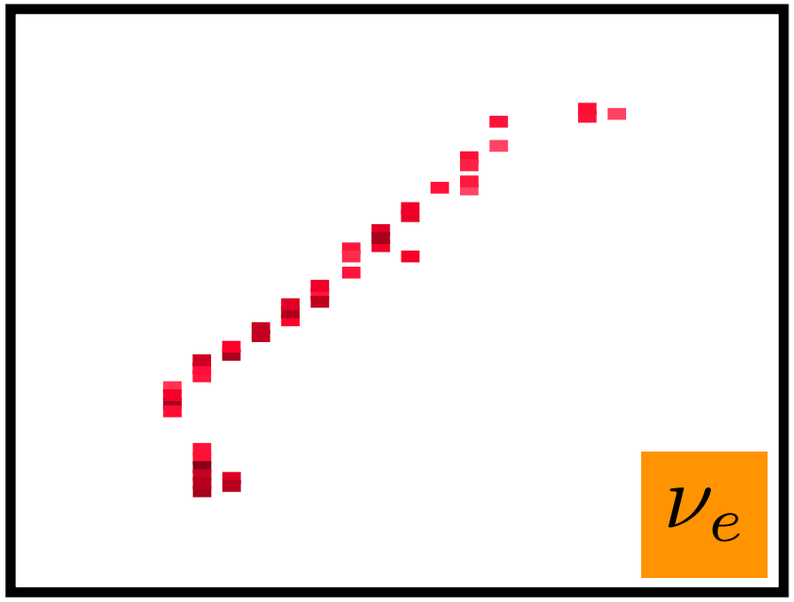


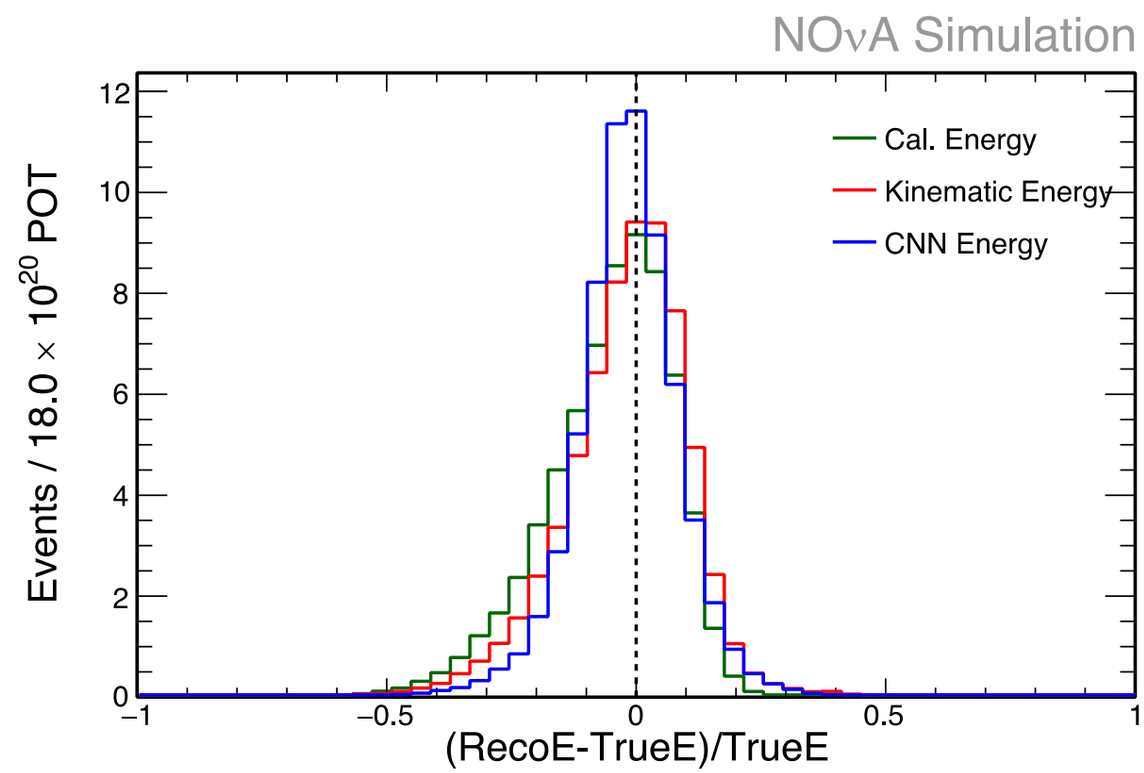


Reconstruction



CNN Energy Estimator



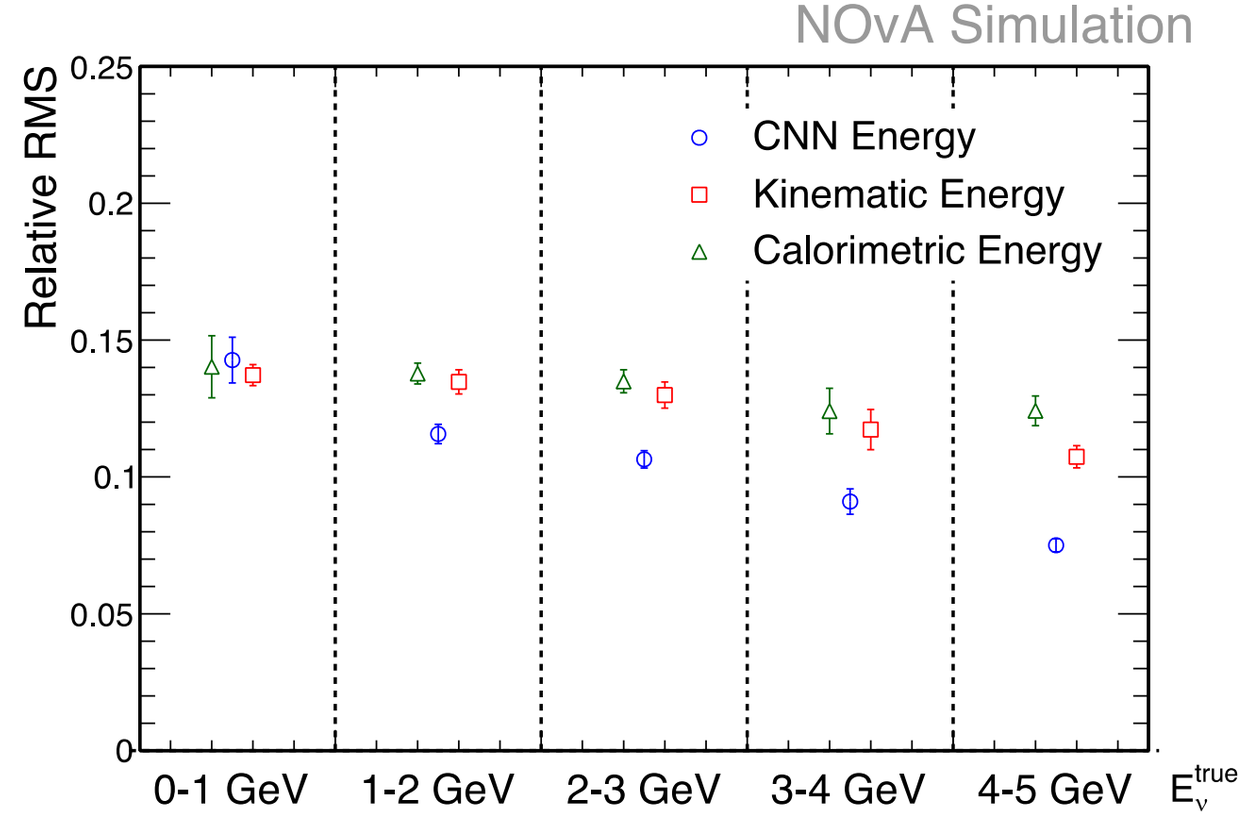


Use a CNN to estimate the energy of electron neutrino interactions.

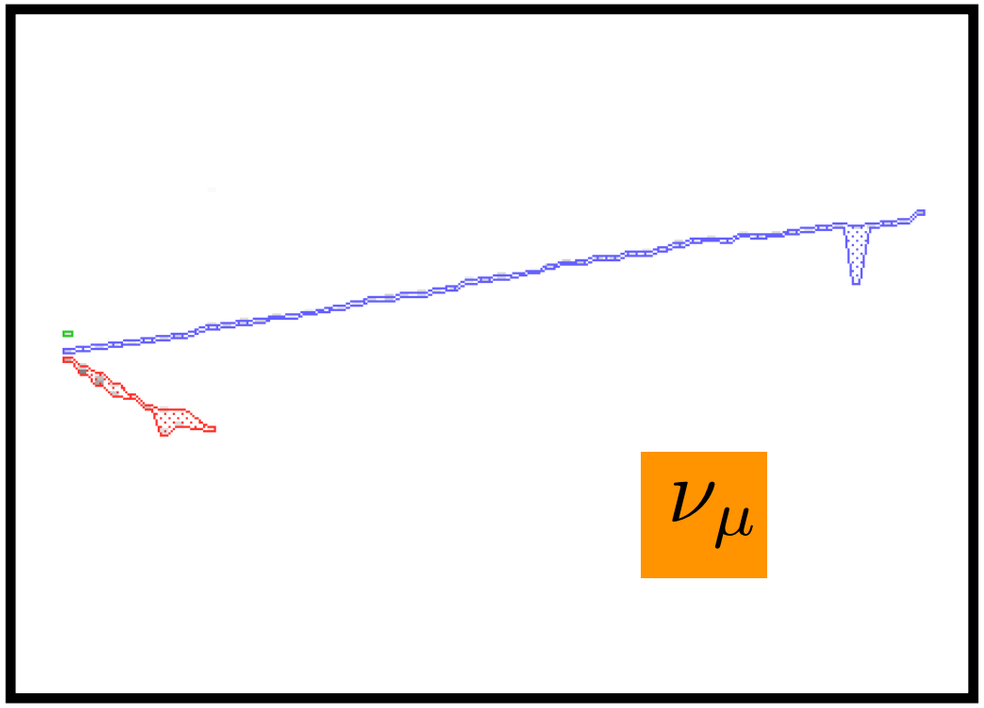
Use linear output rather than classification for continuous variables.

Shows better resolution and smaller dependence on interaction model.

P. Baldi, J. Bian, L. Hertel, and L. Li, Phys. Rev. **D99**, 012011 (2019).



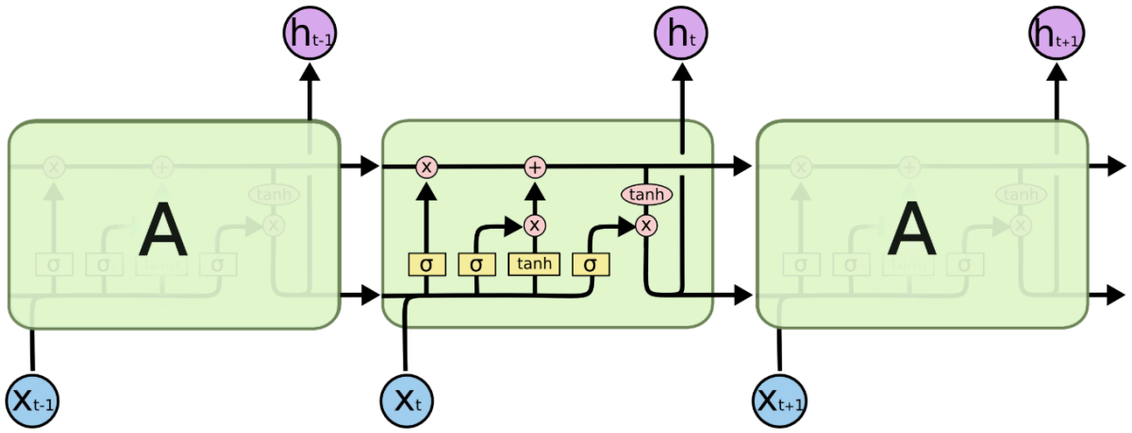
LSTM Energy Estimator



Recurrent Neural Networks:

Takes sequential data where each layer takes the next input and an output from the previous layer.

LSTM: RNN with a Memory Cell



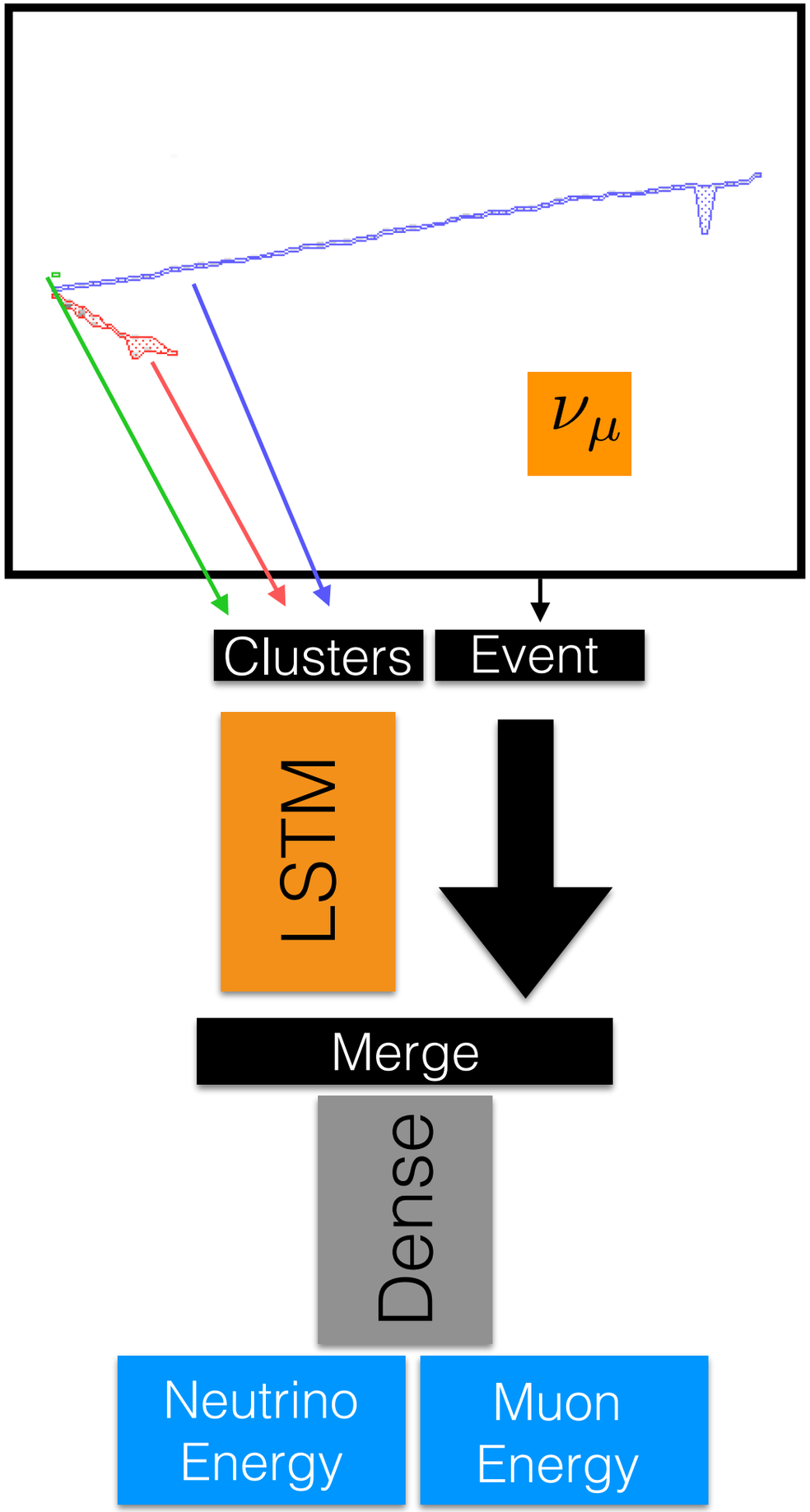
LSTMs are capable of maintaining information across many inputs.

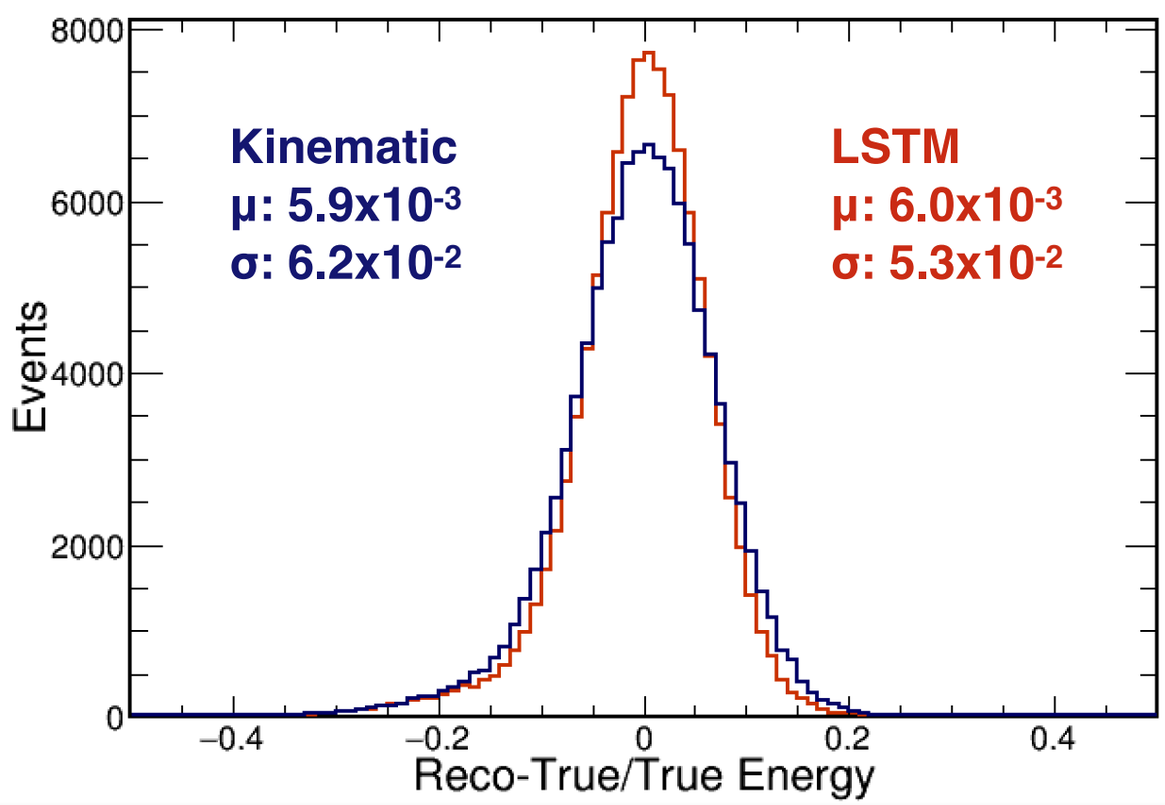
LSTM Energy Estimator

Feed reconstructed quantities from particle clusters into an LSTM:

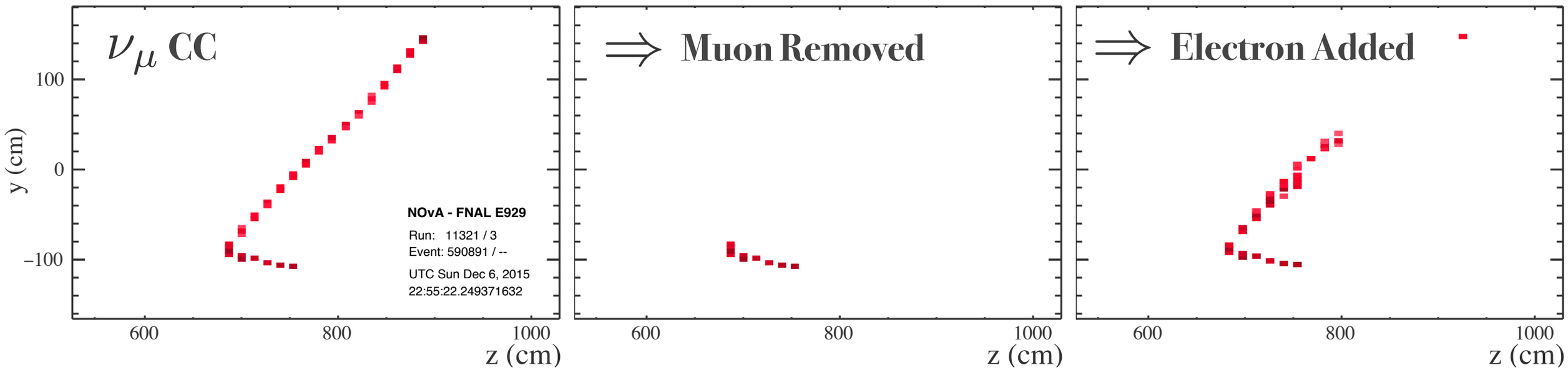
- Number of Hits
- Calorimetric Energy
- Length
- Direction
- Even more…

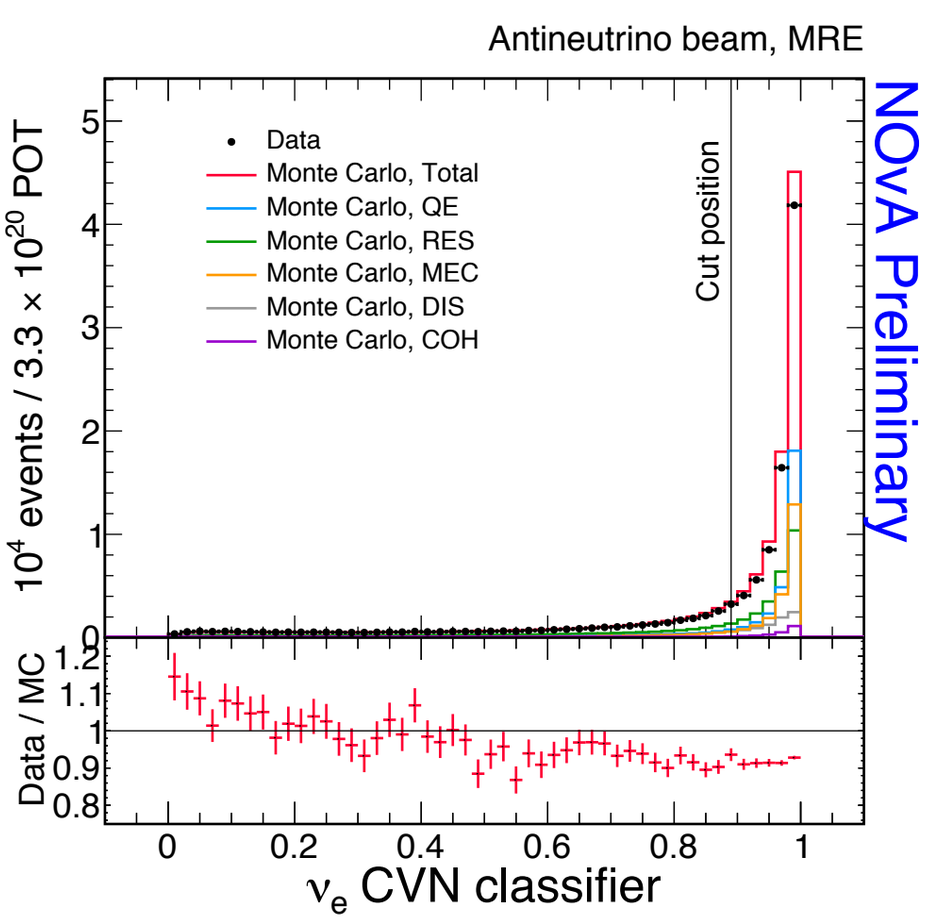
As well as information about the entire event.





Data Check: MRE





Muon Removed - Electron Added:

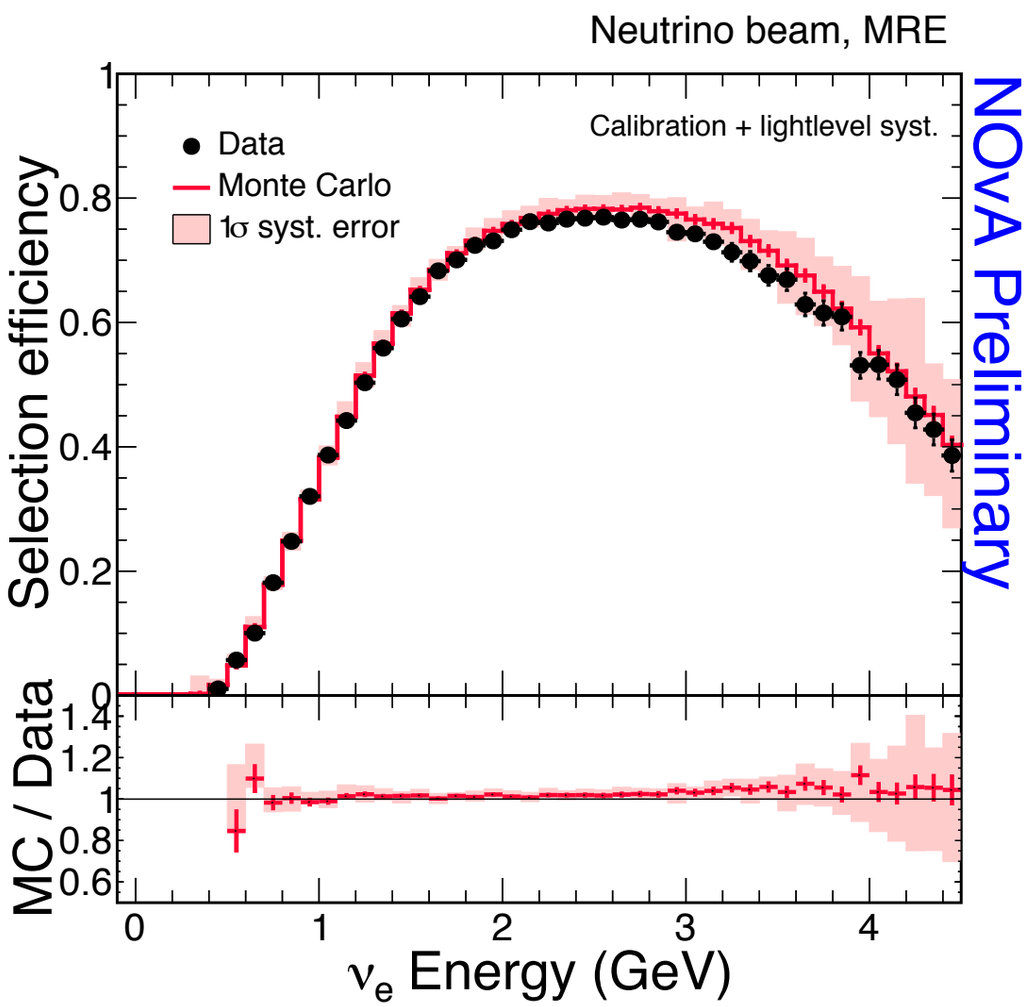
Select a muon neutrino interaction.

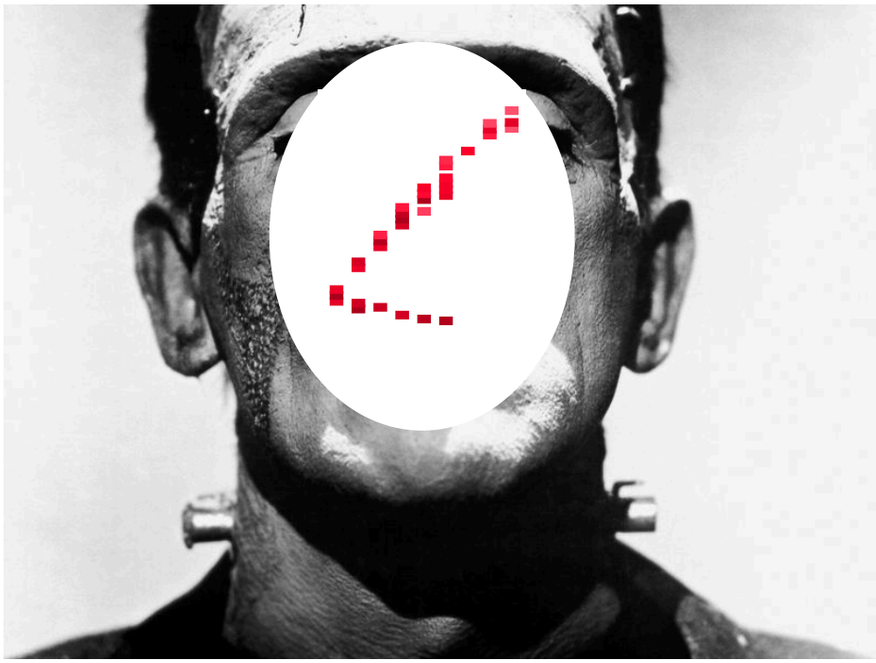
Remove the muon hits and replace with a simulated electron.

Can perform on both data and simulation to compare selection algorithms.

Useful to compare differences between data and simulation in the hadronic system.

Data Check: MRE





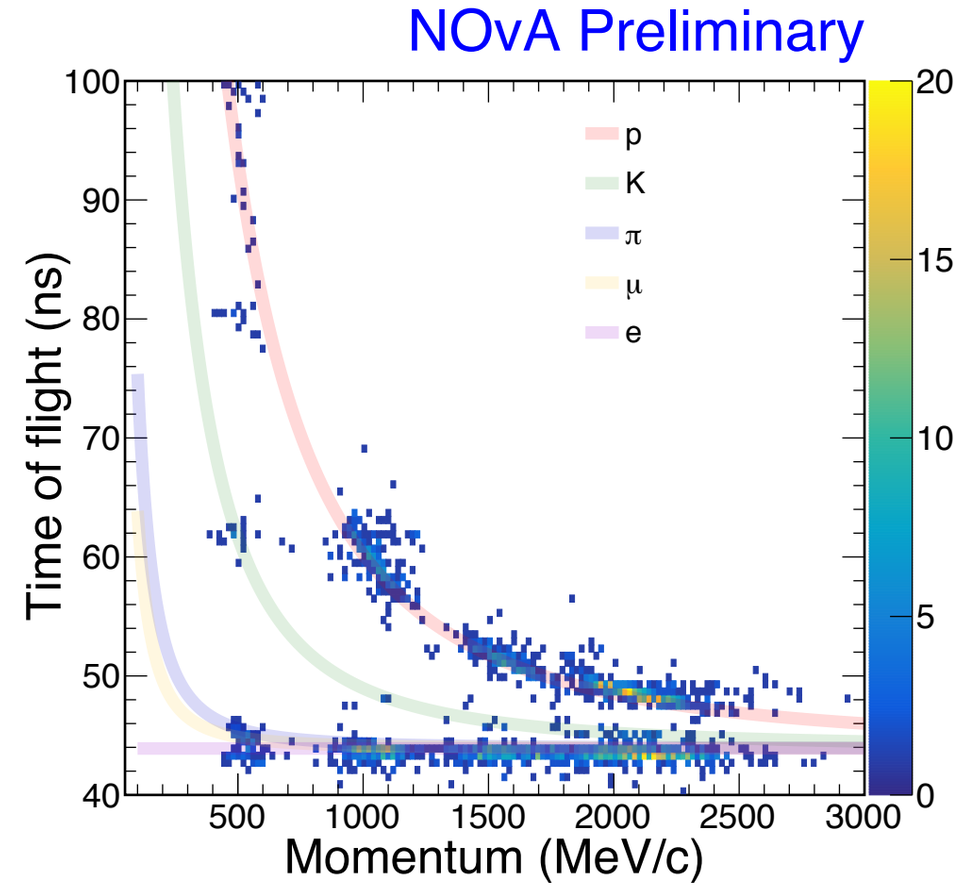
MRE events show good selection agreement between **data** and **simulation**.

		Preselection	Full Selection	Efficiency	Difference
Neutrino beam 6.77×10^{20} POT	Data	486083	316009	0.650	-2.56%
	MC	511287	341119	0.667	
Antineutrino beam 3.29×10^{20} POT	Data	115400	78133	0.677	-1.23%
	MC	123741	84825	0.686	

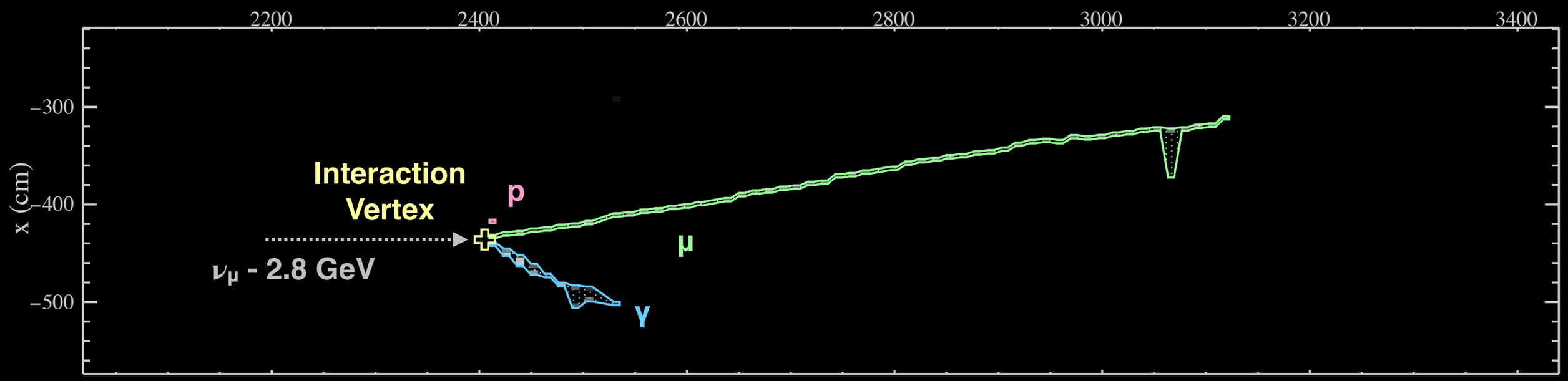
NOvA Test Beam

The NOvA test beam will begin production running this fall.

With a library of labeled data from single particle **interactions of known identity and momentum**, NOvA will expand the data-driven checks of our deep learning algorithms.



Summary



NOvA uses deep learning for neutrino flavor and particle ID in our analyses, Ryan's talk is imminent.

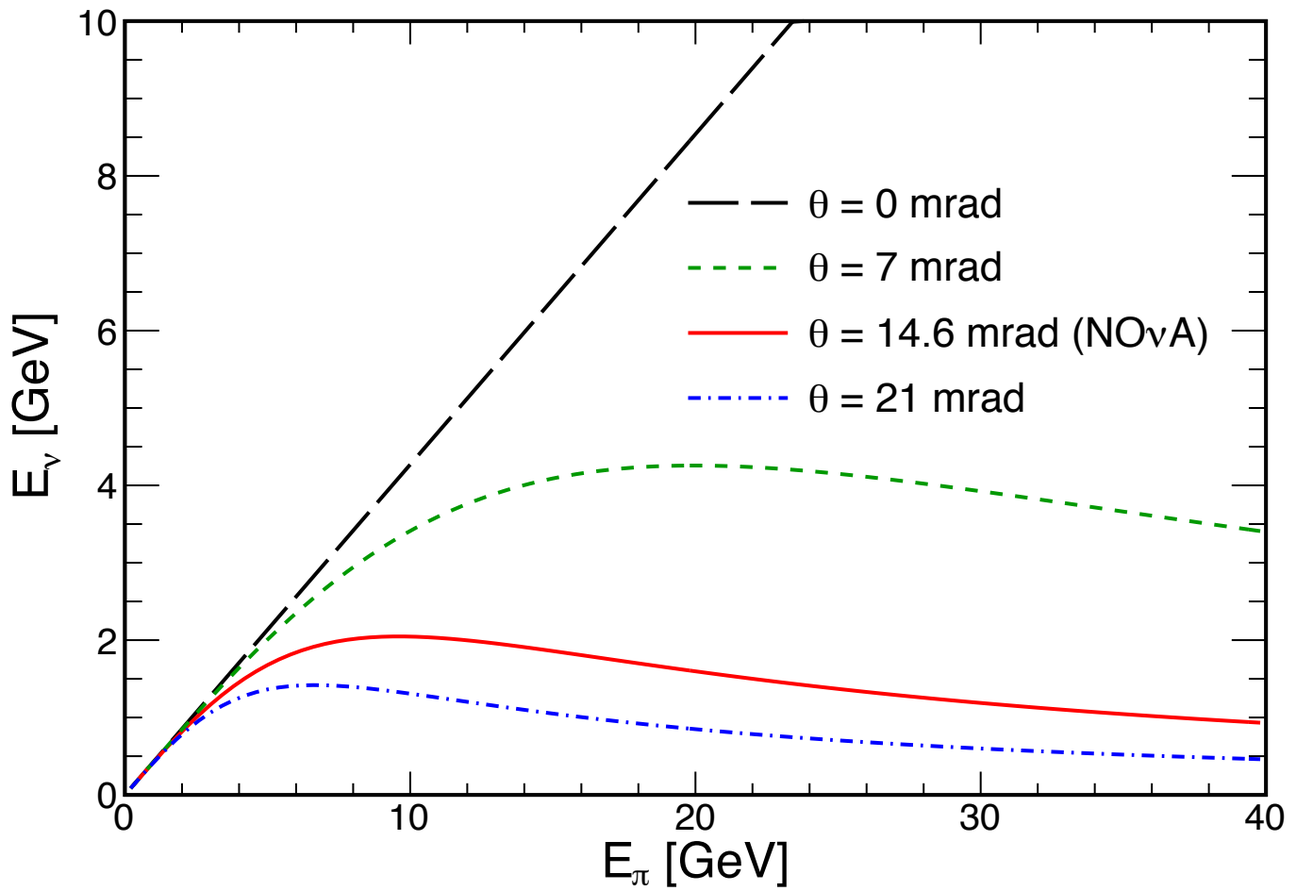
Used to improve oscillation measurements since 2016, and now being incorporated into cross sections analyses.

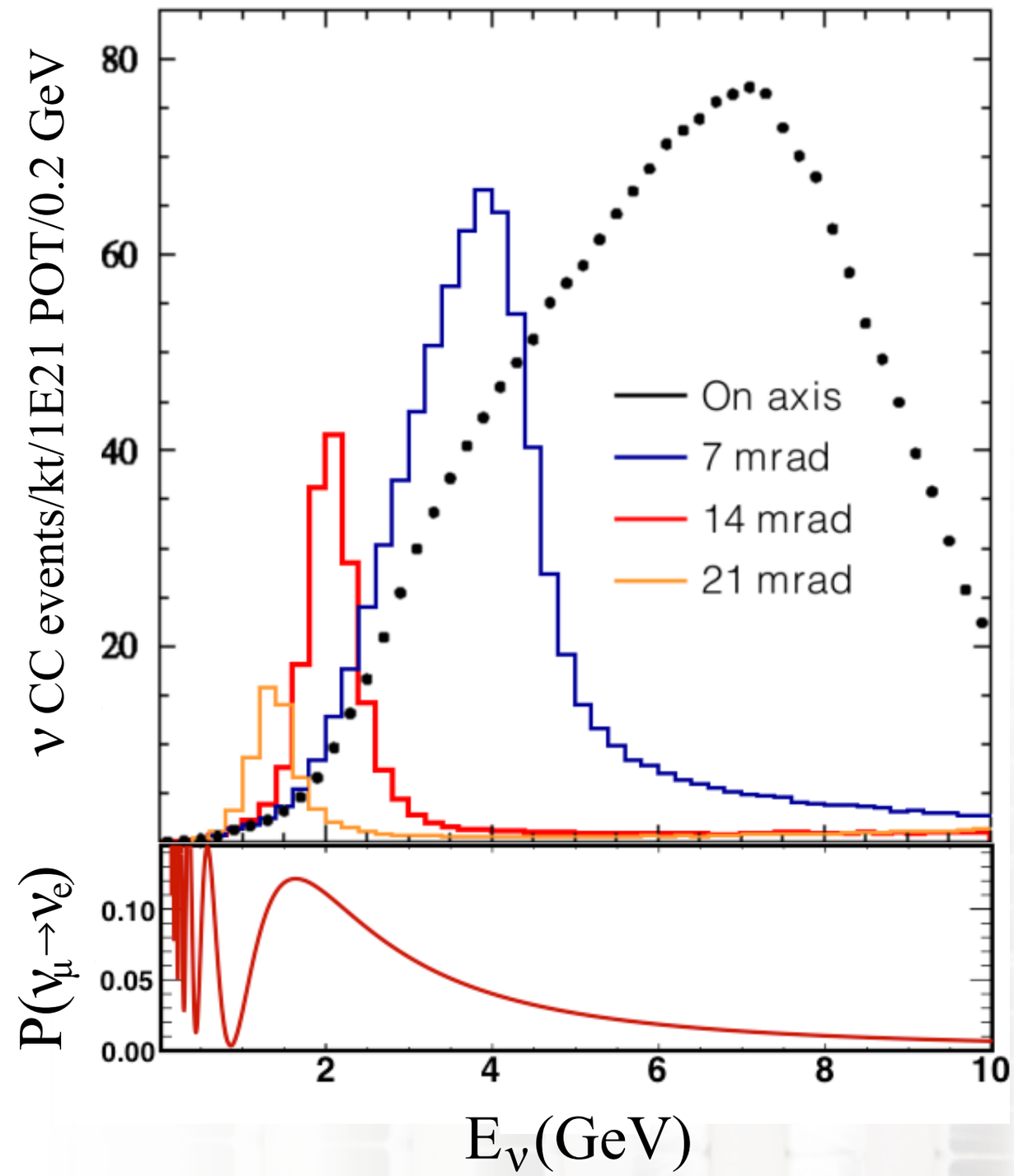
Our deep learning program incorporates data driven cross checks from cosmic, neutrino, and soon test beam data!

NOvA deep learning efforts are broad, including algorithms for identification, clustering, energy reconstruction, and more!

Can you beat our neural network at neutrino classification?

Backup


θ = 0 mrad
θ = 7 mrad
θ = 14.6 mrad (NOνA)
θ = 21 mrad
E_ν [GeV]
E_π [GeV]
0
2
4
6
8
10
0
10
20
30
40

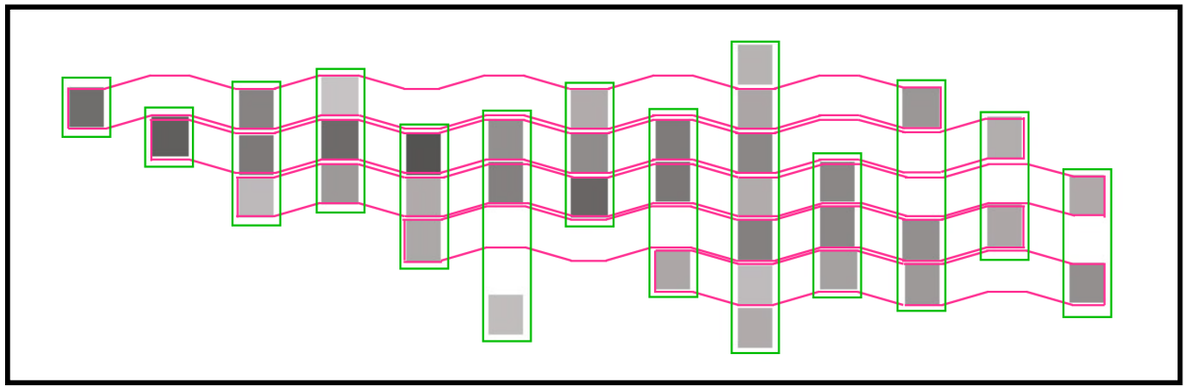

ν CC events/kt/1E21 POT/0.2 GeV
On axis
7 mrad
14 mrad
21 mrad
$P(\nu_\mu \to \nu_e)$
E_ν(GeV)
20
40
60
80
0.00
0.05
0.10
2
4
6
8
10

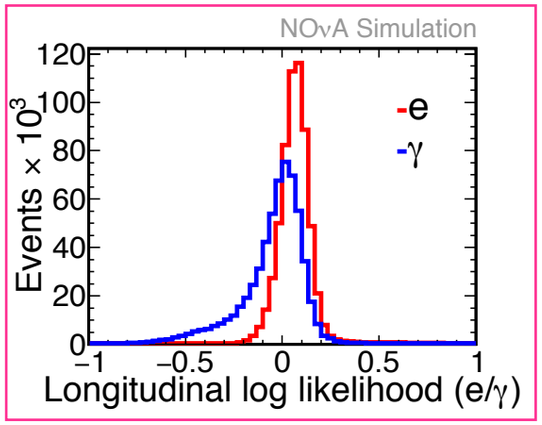
Neutrino Identification (First Analysis)

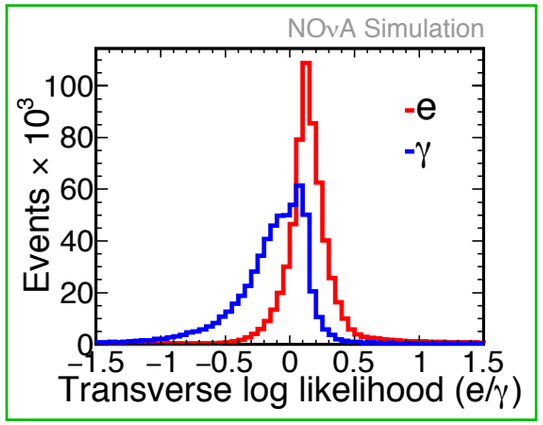
LID

Likelihood Identification

Premise: Electron showers have characteristic transverse and longitudinal energy deposition profiles.







In practice:

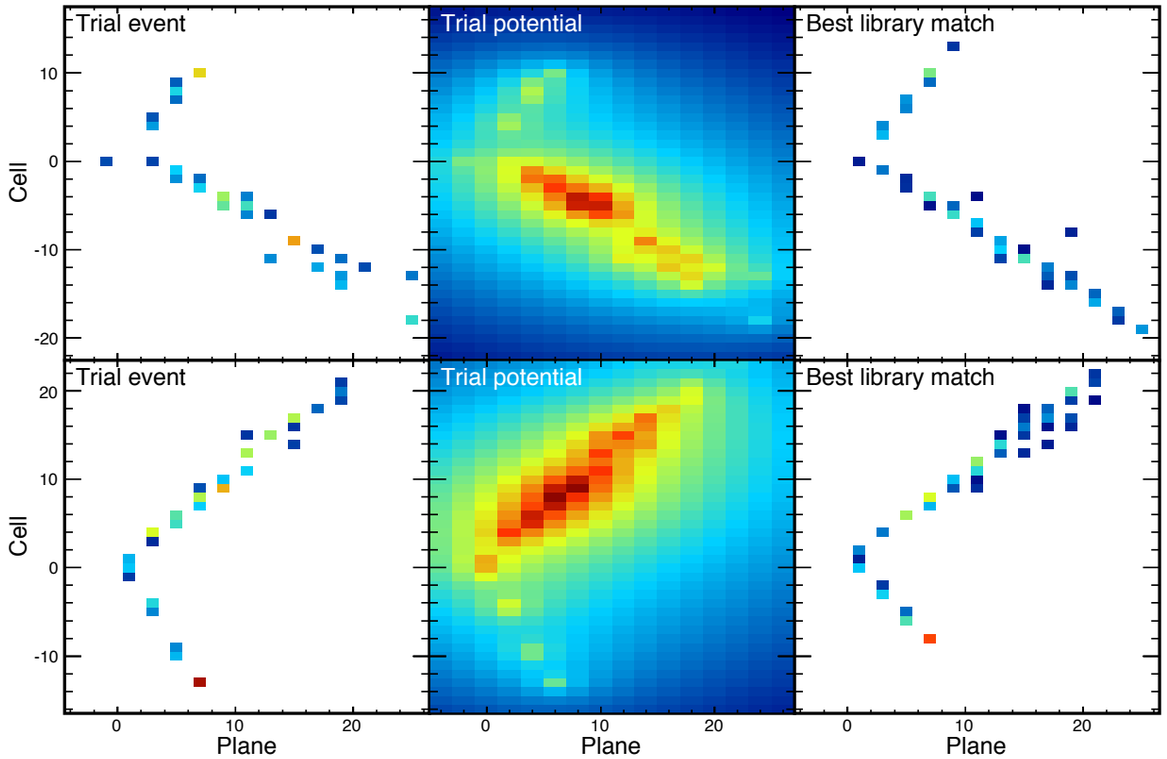
- ★ Reconstruct electron shower.
- ★ Find likelihoods from it's dE/dx profiles compared to particle hypotheses.

Likelihoods → *Neural Network*

LEM

Library Event Matching

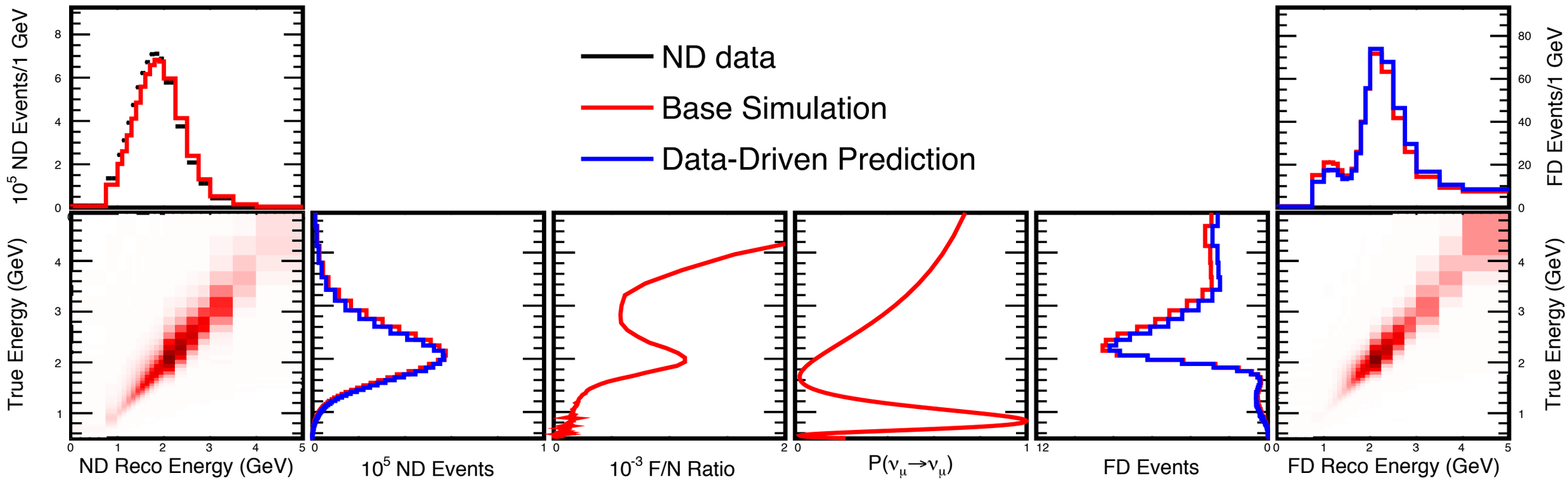
Premise: We have a large library of simulated event templates, large enough that we can use it to compare pixel by pixel.

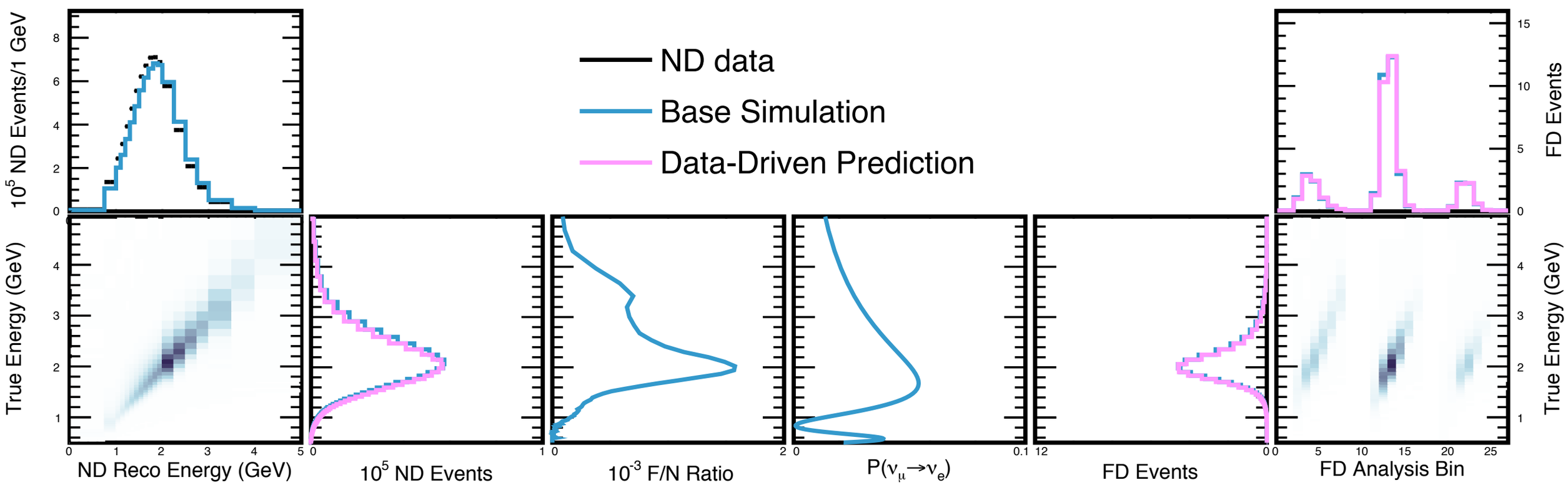


In practice:

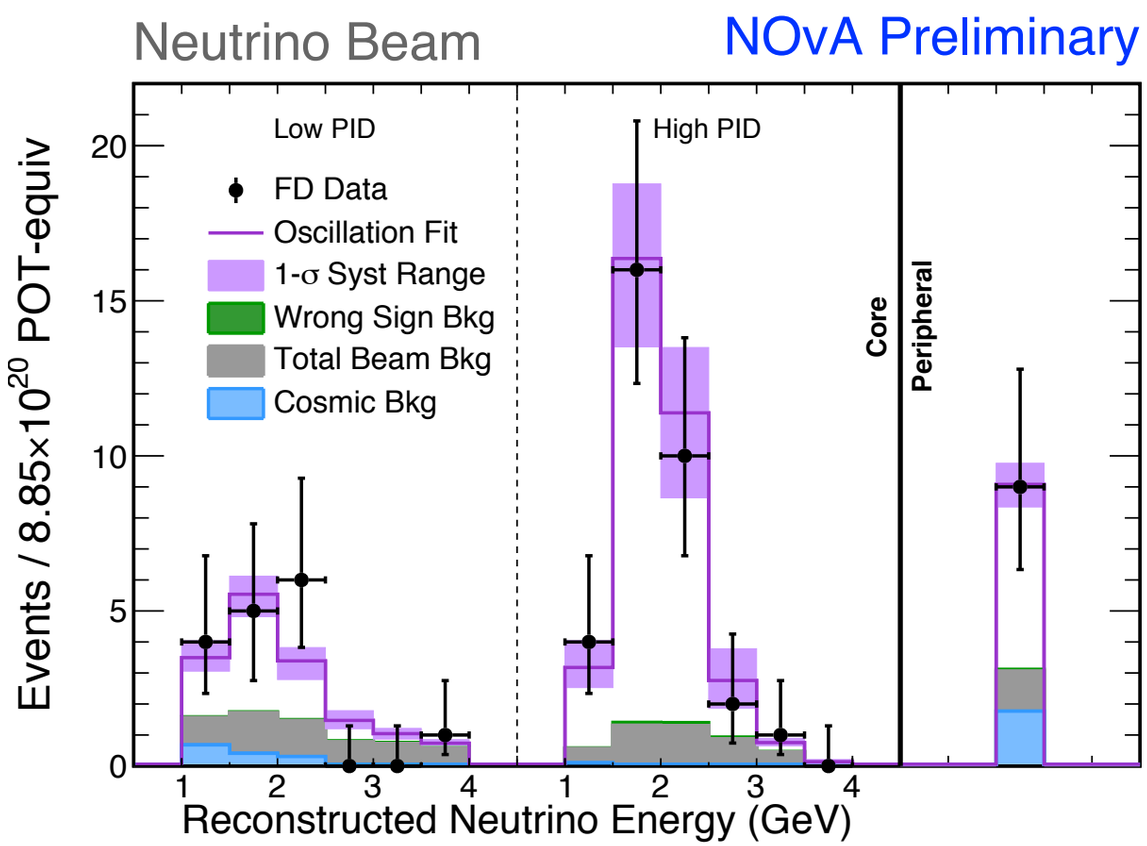
- ★ Find the best matches from the event library.
- ★ Extract features from best matches.

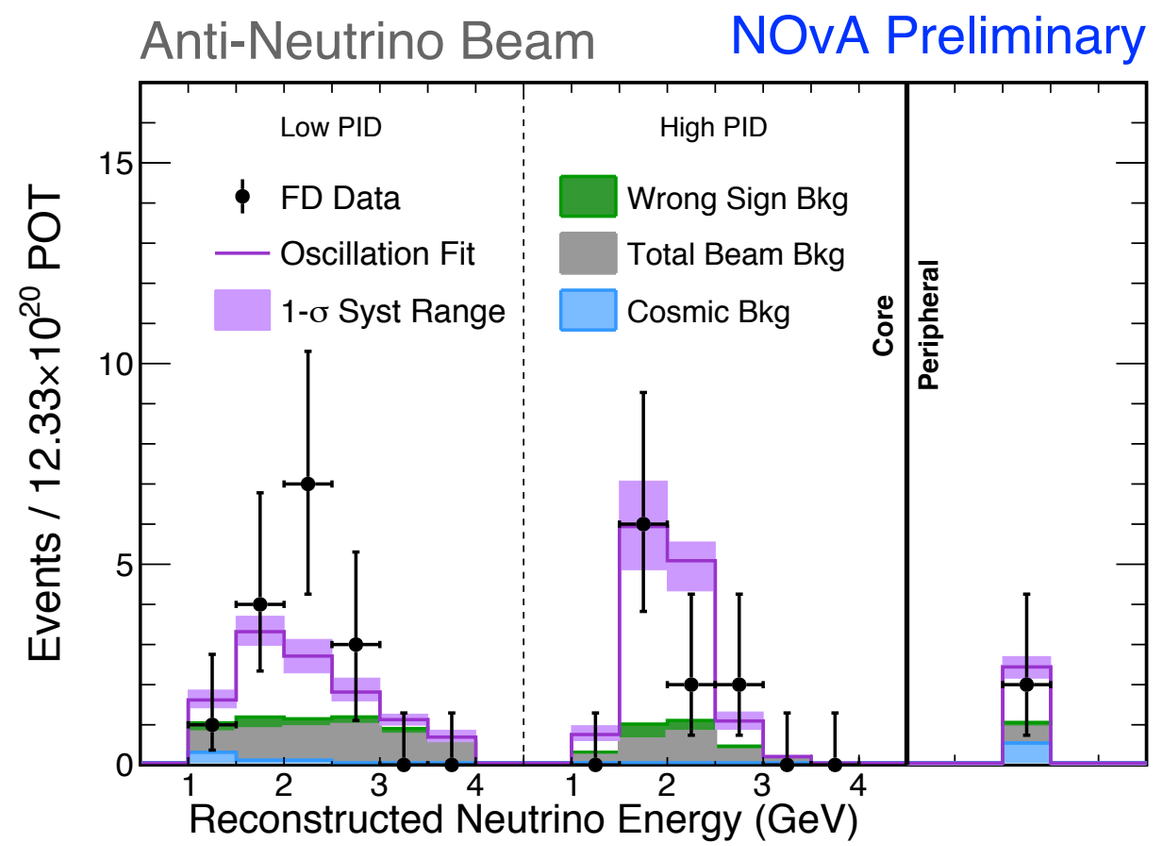
Features → *Decision Tree*


ND data
Base Simulation
Data-Driven Prediction
10^5 ND Events/1 GeV
True Energy (GeV)
ND Reco Energy (GeV)
10^5 ND Events
10^{-3} F/N Ratio
$P(\nu_\mu \rightarrow \nu_\mu)$
FD Events
FD Events/1 GeV
FD Reco Energy (GeV)

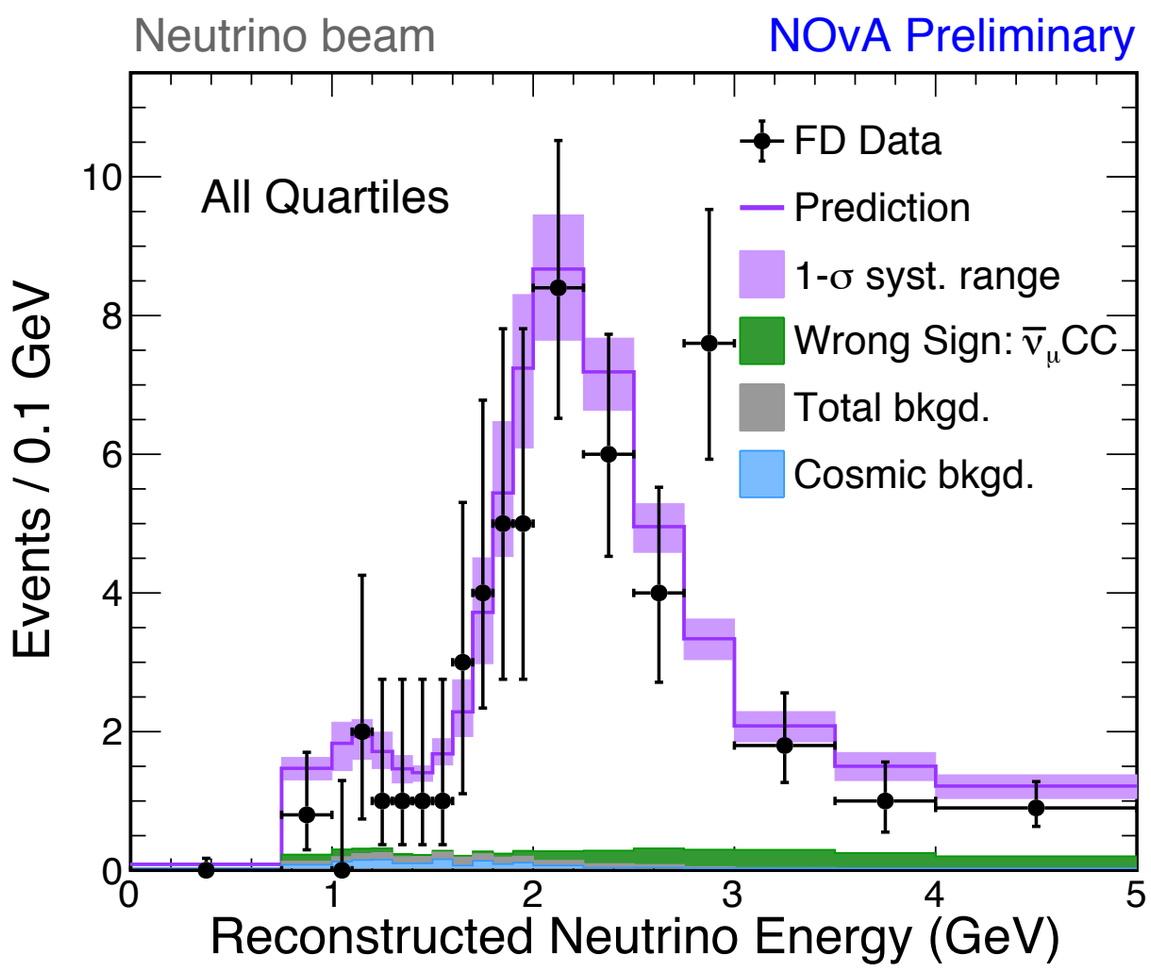

ND data
Base Simulation
Data-Driven Prediction
10^5 ND Events/1 GeV
True Energy (GeV)
ND Reco Energy (GeV)
10^5 ND Events
10^{-3} F/N Ratio
$P(\nu_\mu \rightarrow \nu_e)$
FD Events
FD Analysis Bin

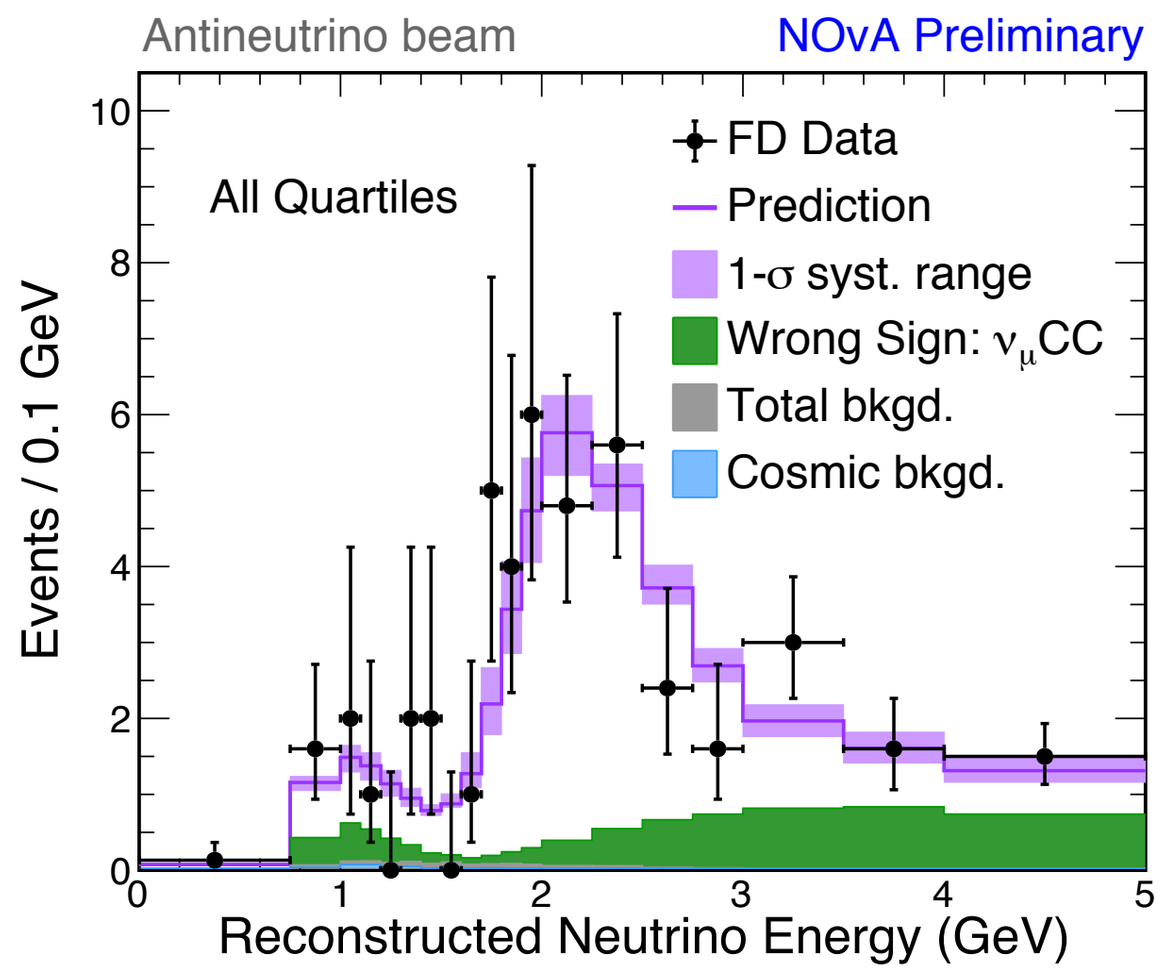
Nue Spectrum





Numu Spectrum





PandAna

The hdf5 files contain the same information as our standard root analysis files, as well as training images.

We created a new python-based analysis framework to compliment the root one, PandAna.

Event	NHit	Max X	Max Y	Max Z
1	75	5.4	5.6	23.9
2	54	3.9	4.8	54.8
3	102	4.5	5.9	12.7
4	63	5.7	2.4	43.6

PANDAna

CAFAna

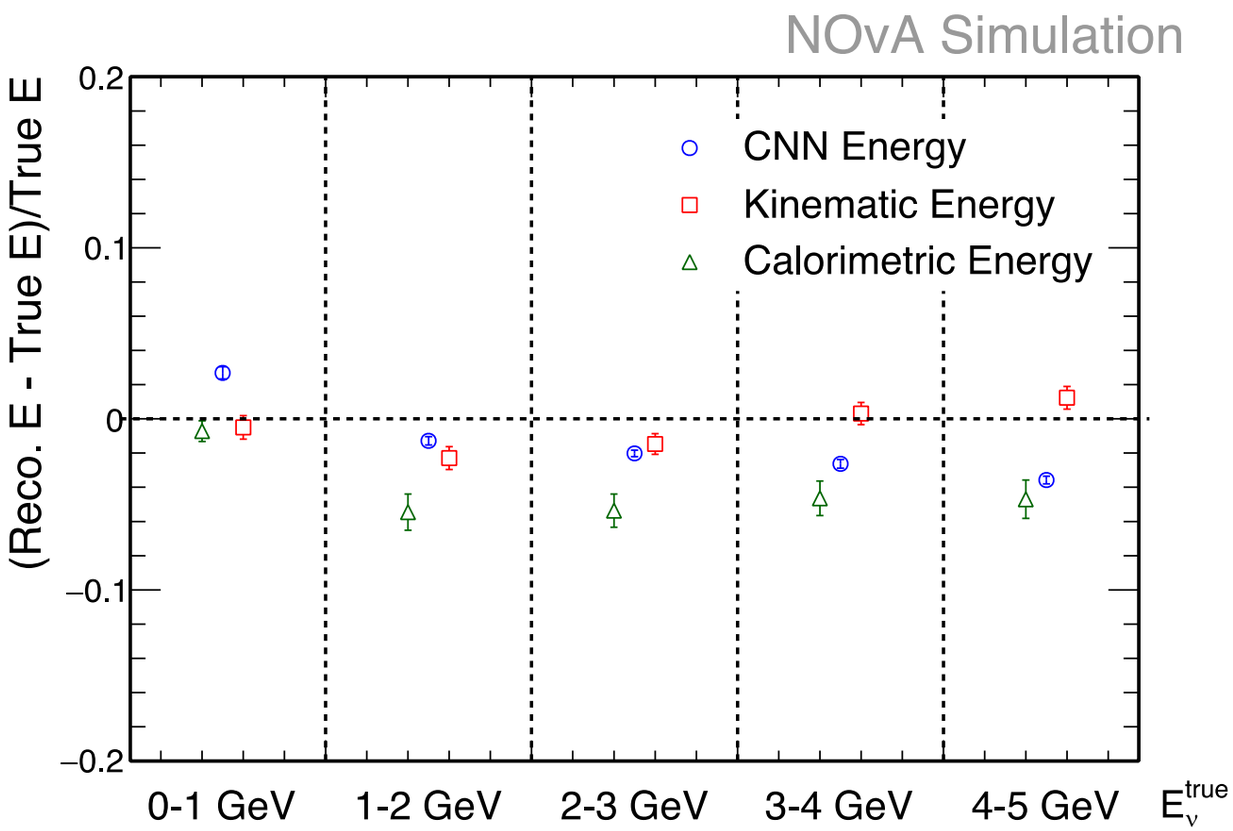
PandAna does a column based analysis with pandas.

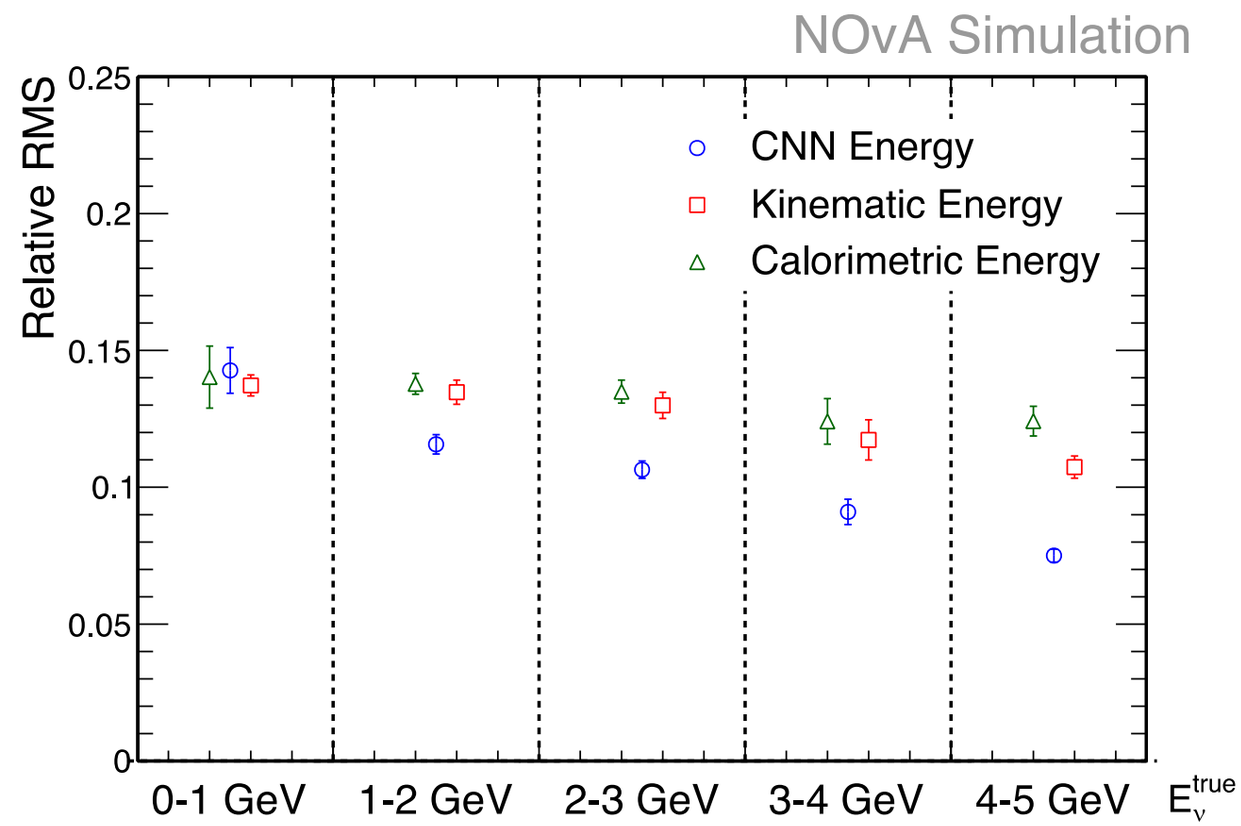
Allows for carefully crafted training sample.

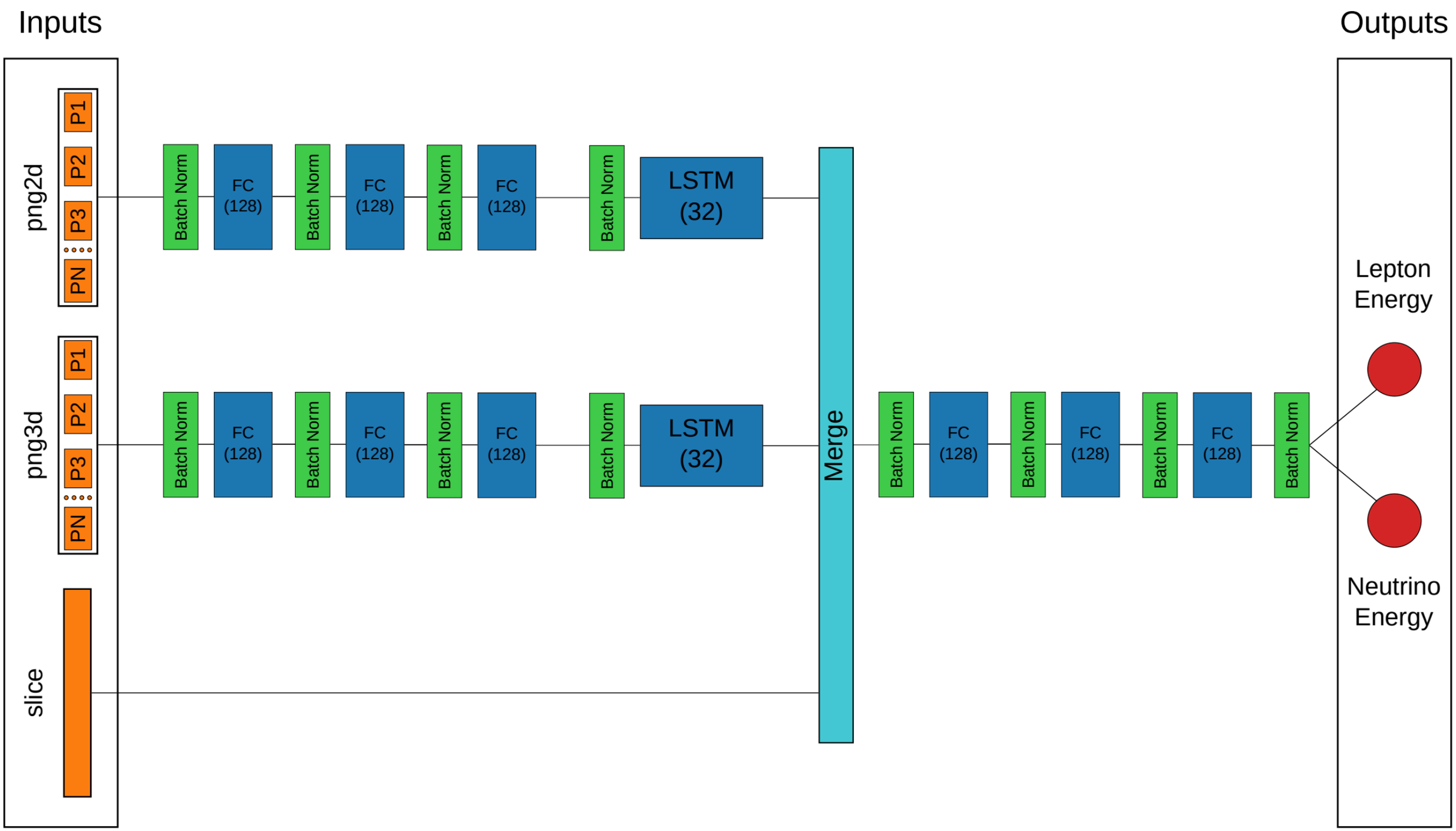
Additionally, we can do a full algorithm validation within the hdf5 files.

Regression CNN

Bias against true energy.






Inputs
Outputs
png2d
P1
P2
P3
PN
png3d
P1
P2
P3
PN
slice
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
LSTM (32)
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
LSTM (32)
Merge
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
FC (128)
Batch Norm
Lepton Energy
Neutrino Energy

LSTM Variables

3D prong vars:
`vtx.elastic[0].fuzzyk.png`:

```
png.bpf[i].energy
png.bpf[i].momentum.{x,y,z}
png.bpf[i].overlapE
png.calE
png.cvnpart.{
  muon,pion,proton,
  electron,photonid
}
png.dir.{x,y,z}
png.len
png.nhit{,.x,.y}
png.nplane
png.start.{x,y,z}
png.weightedCalE
```

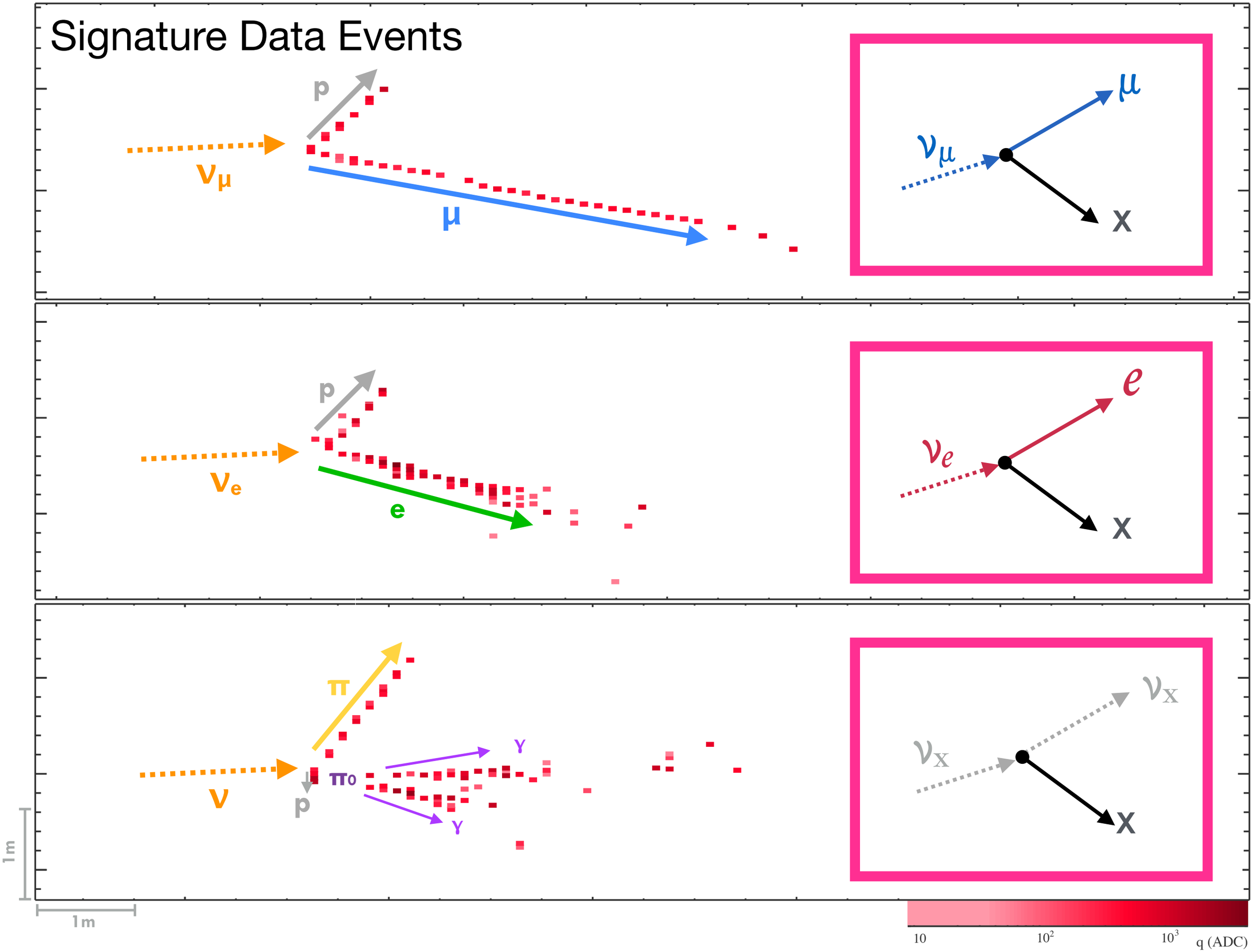
2D prong vars:
`vtx.elastic[0].fuzzyk.png2d`:

```
png2d.calE
png2d.dir.{x,y,z}
png2d.len
png2d.nhit{,.x,.y}
png2d.nplane
png2d.start.{x,y,z}
png2d.weightedCalE
```

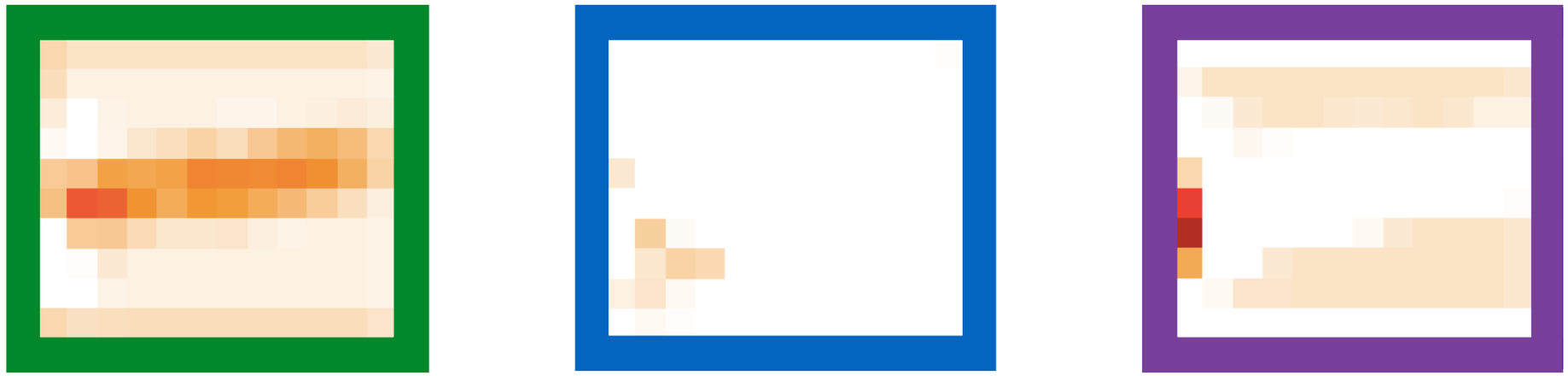
Slice vars:

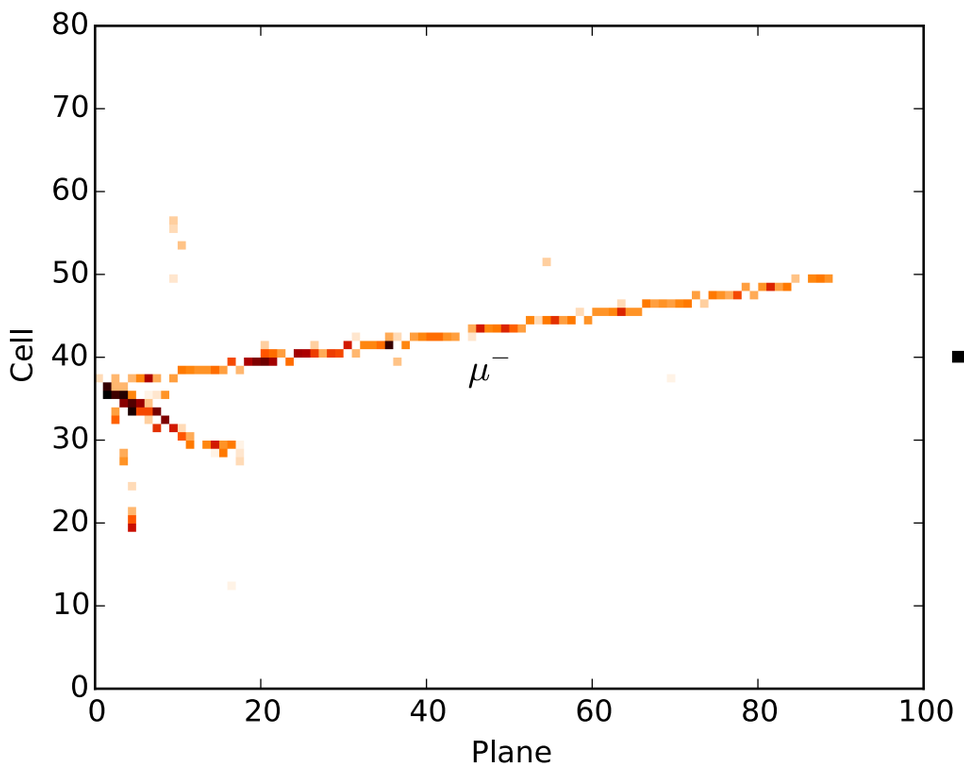
```
calE
IsCoarseTiming
IsLowGain
nHit
orphCalE
remPngCalE
```

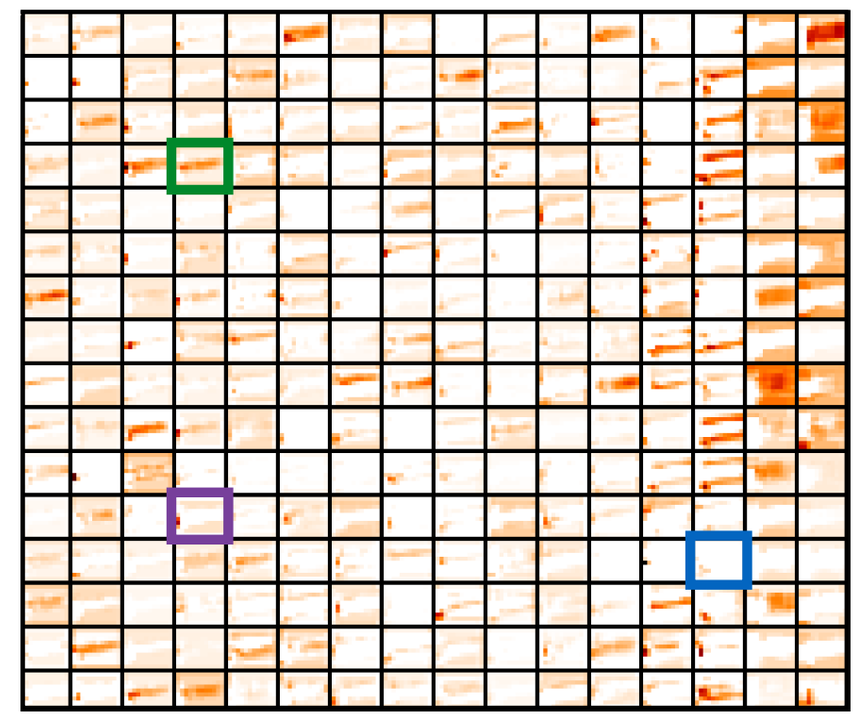
Signature Data Events



NOvA Features







NOvA Features