Deep Learning Methods in the NOνA Experiment

Chris Kullenberg

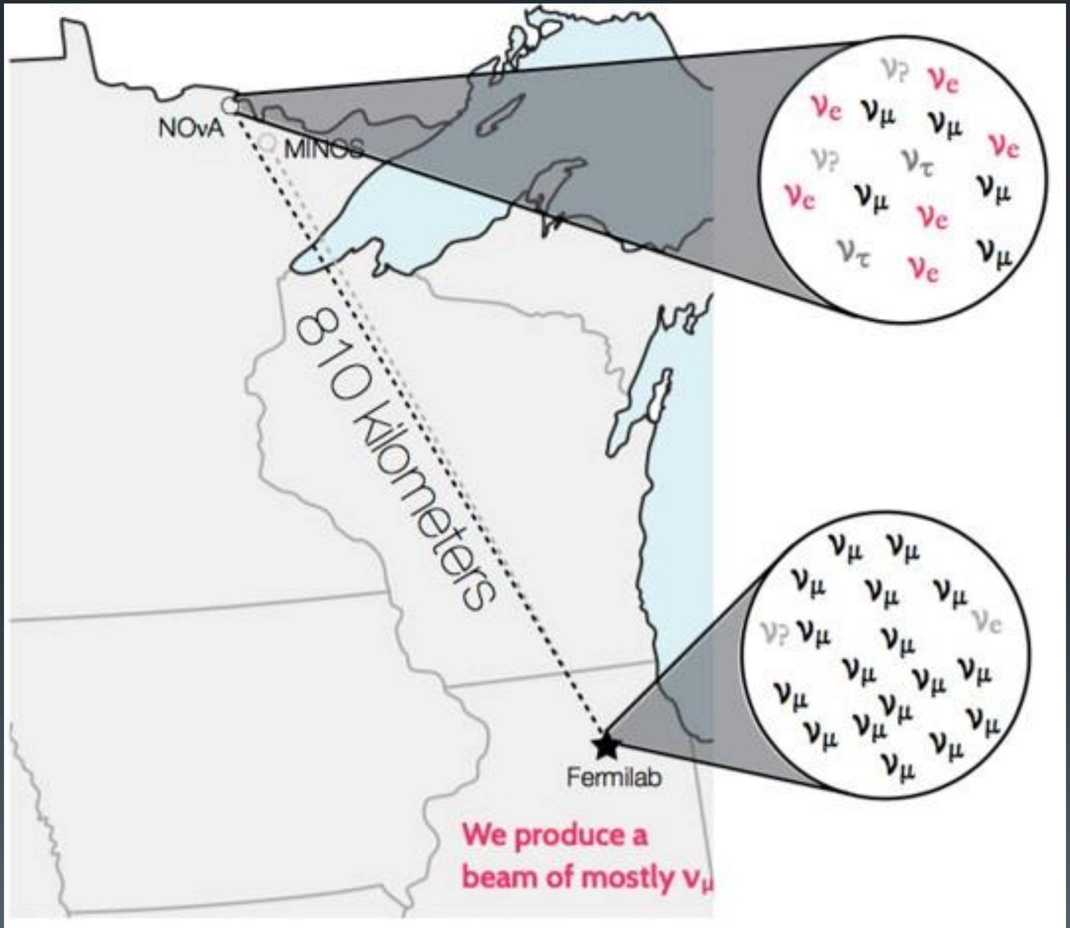
Nuclear Electronics & Computing 2019

Why NOνA?

NuMI **O**ff-axis ν_e **A**ppearance

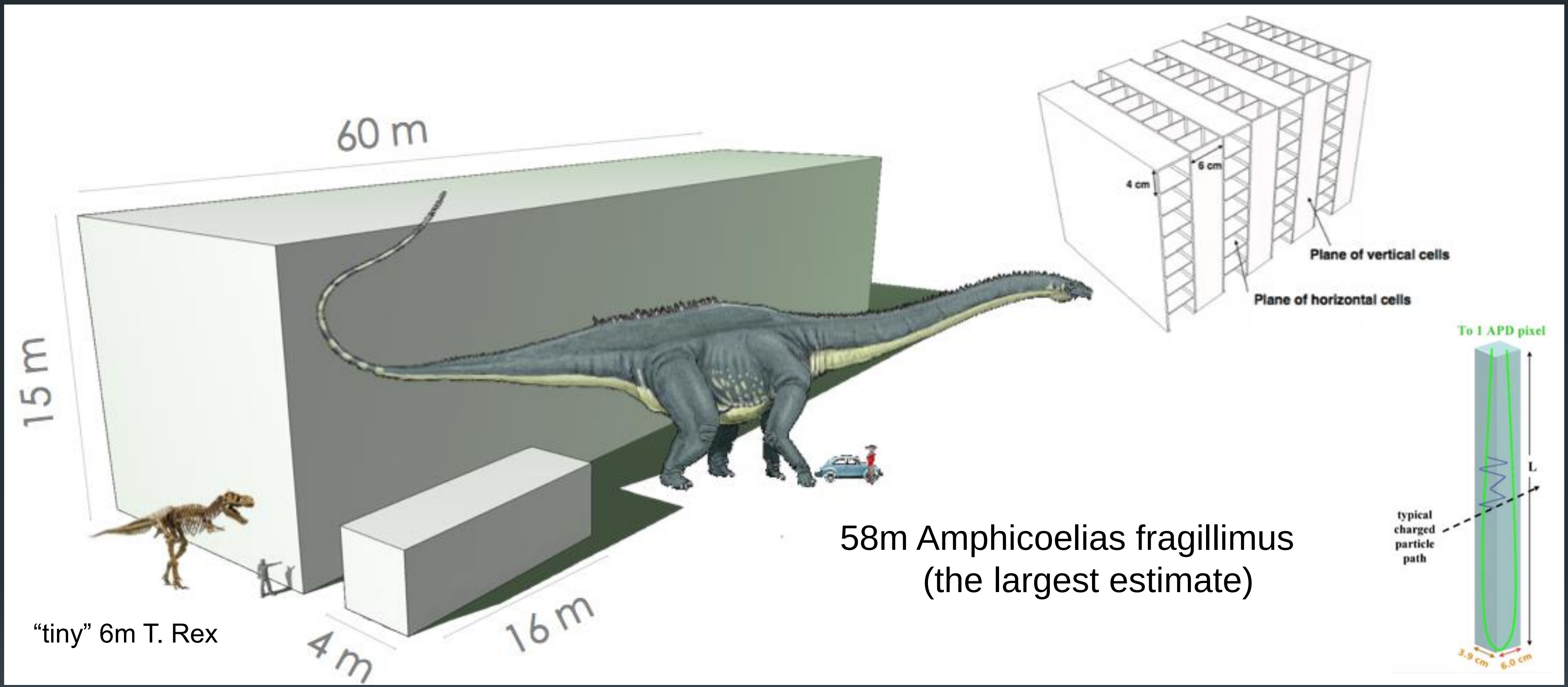
Measure neutrino oscillation parameters!

(among other interesting neutrino topics)



- Neutrino oscillations cause different neutrino flavors to appear during interaction than when created during production
- Understanding these oscillations helps us understand fundamental properties of the neutrino
- The beam composition is measured at the Near Detector (mostly muon neutrinos)
- The number of electron/muon neutrinos appearing at the Far Detector provides the oscillation measurement

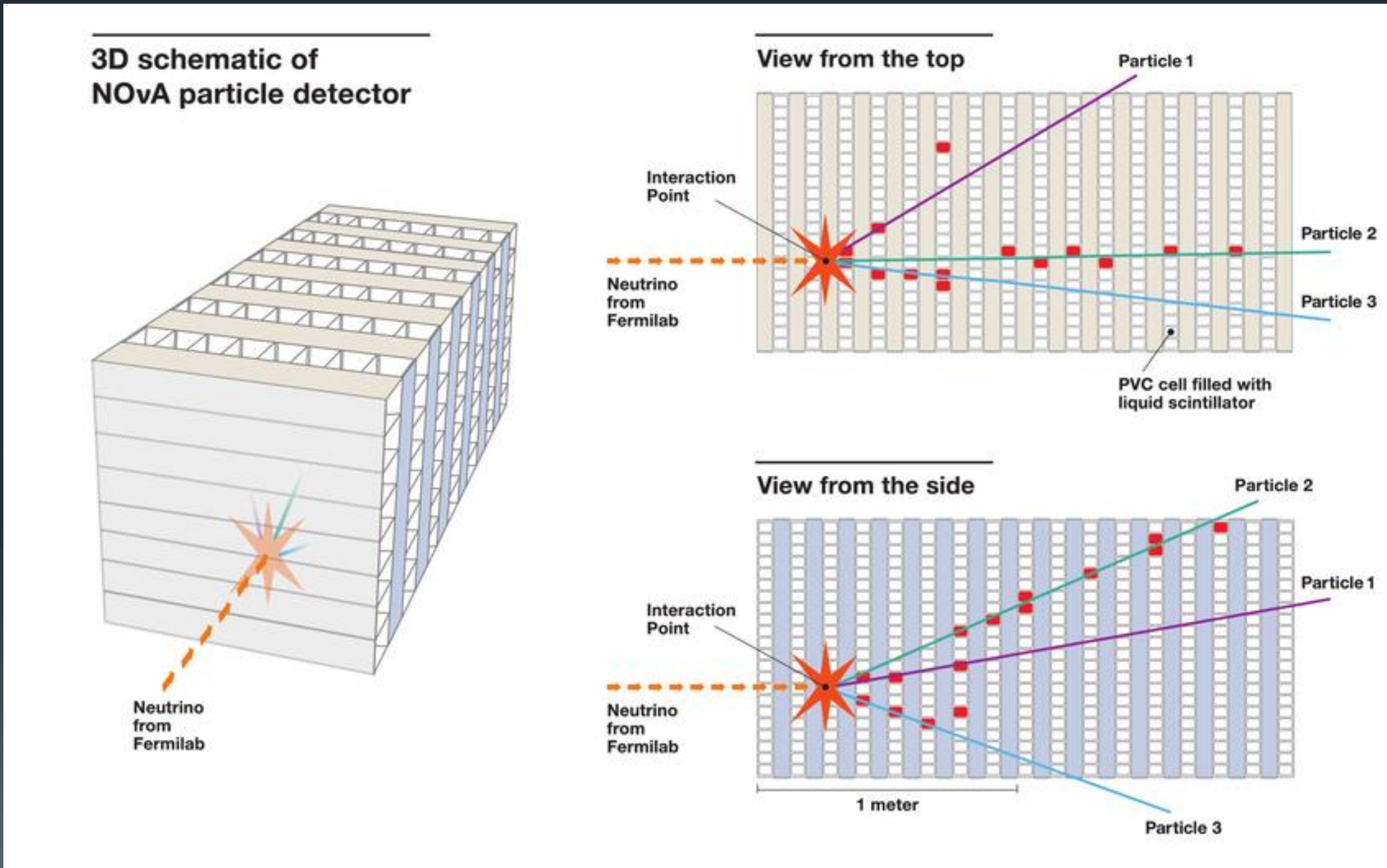
The NOvA Experiment

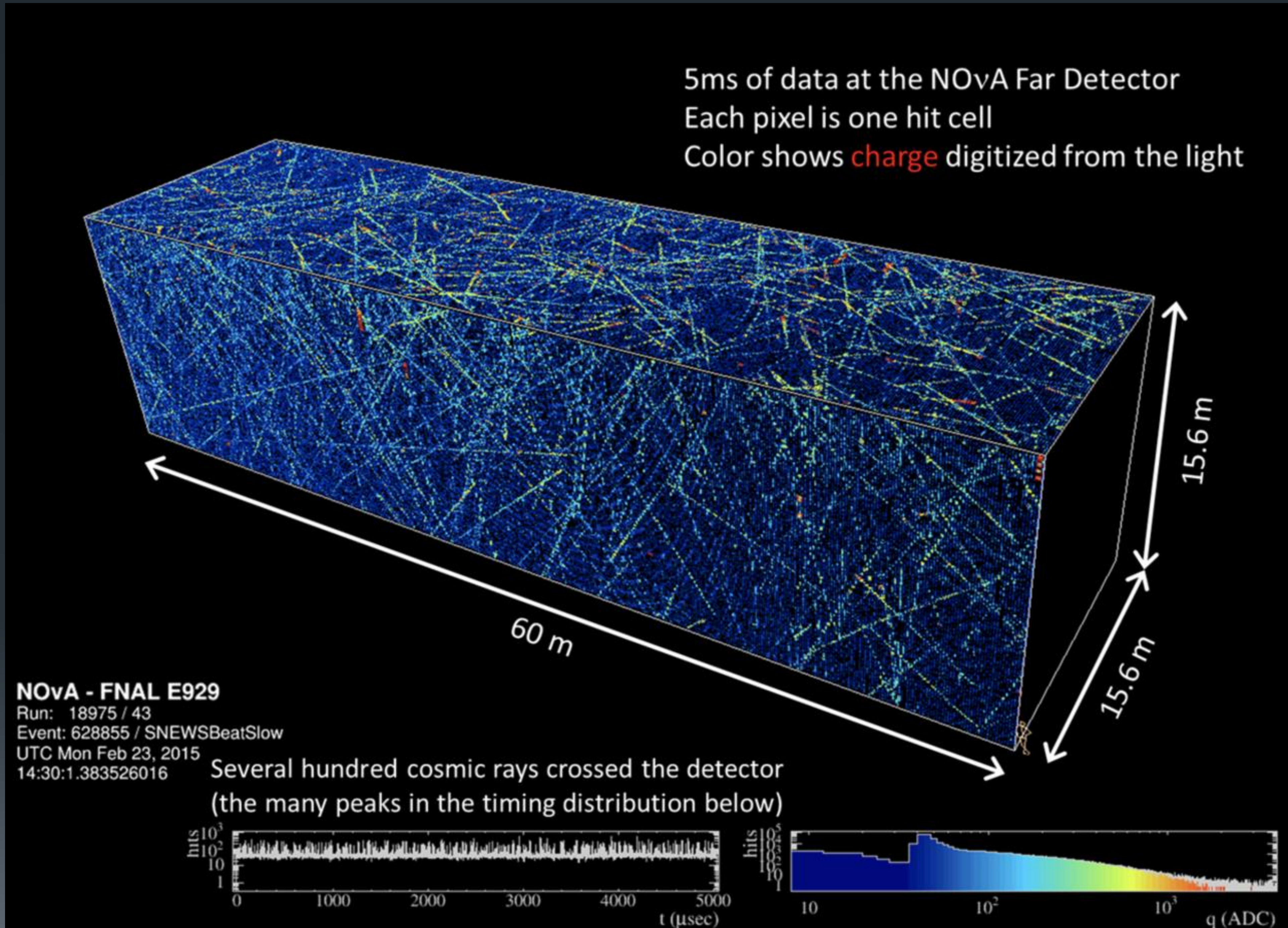


NOvA is basically a pair of neutrino detectors, with identical construction but different sizes. The Far Detector as large as the estimated size of the largest known dinosaur.

Made of alternating vertical/horizontal (X/Y) oriented tubes filled with liquid scintillator.

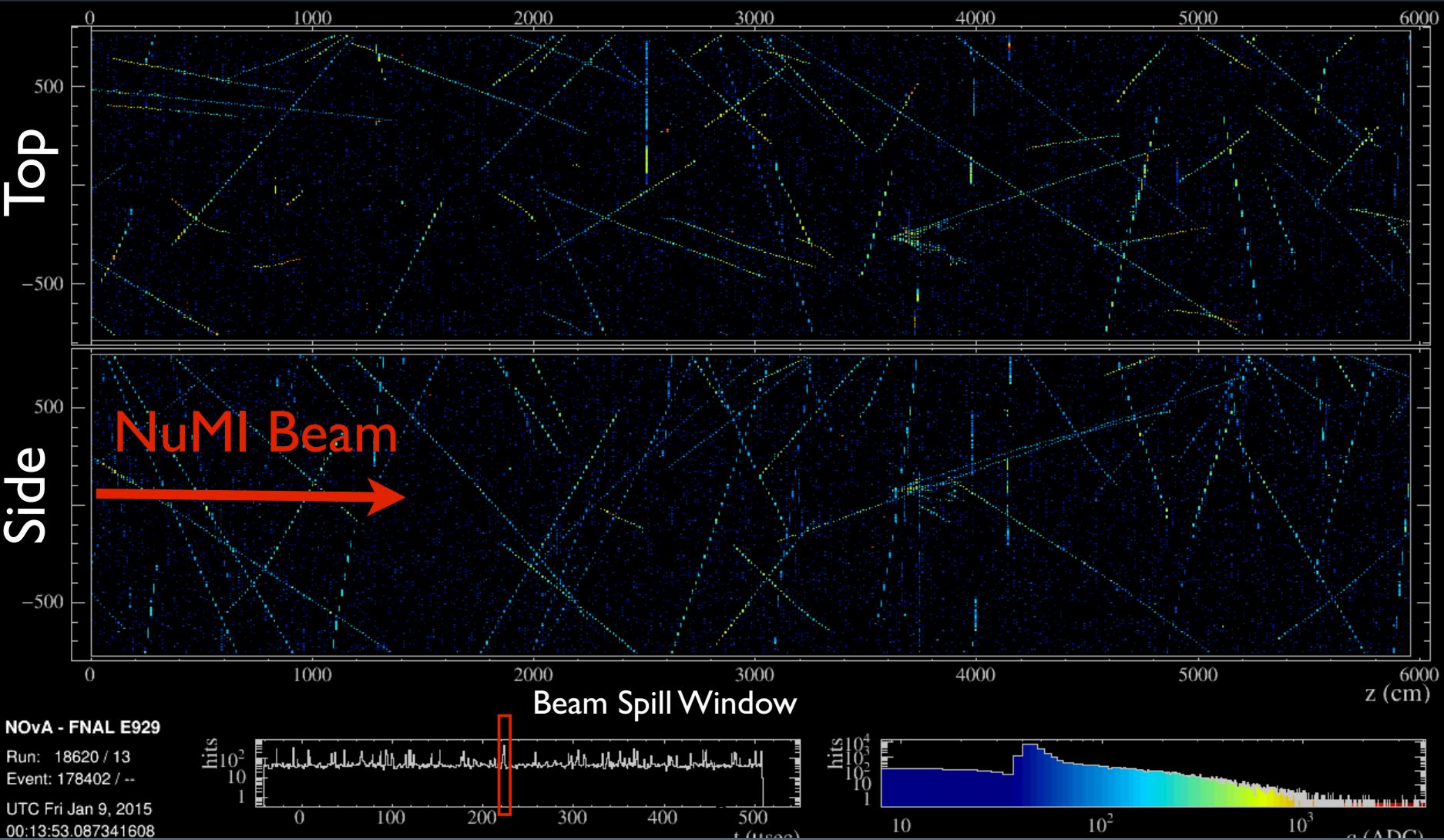
NOvA Data


3D schematic of
NOvA particle detector
Neutrino from Fermilab
View from the top
Particle 1
Interaction Point
Particle 2
Neutrino from Fermilab
Particle 3
PVC cell filled with liquid scintillator
View from the side
Particle 2
Particle 1
Interaction Point
Neutrino from Fermilab
1 meter
Particle 3


5ms of data at the NOνA Far Detector
Each pixel is one hit cell
Color shows charge digitized from the light
15.6 m
60 m
15.6 m
NOνA - FNAL E929
Run: 18975 / 43
Event: 628855 / SNEWSBeatSlow
UTC Mon Feb 23, 2015
14:30:1.383526016
Several hundred cosmic rays crossed the detector
(the many peaks in the timing distribution below)
hits
t (μsec)
hits
q (ADC)

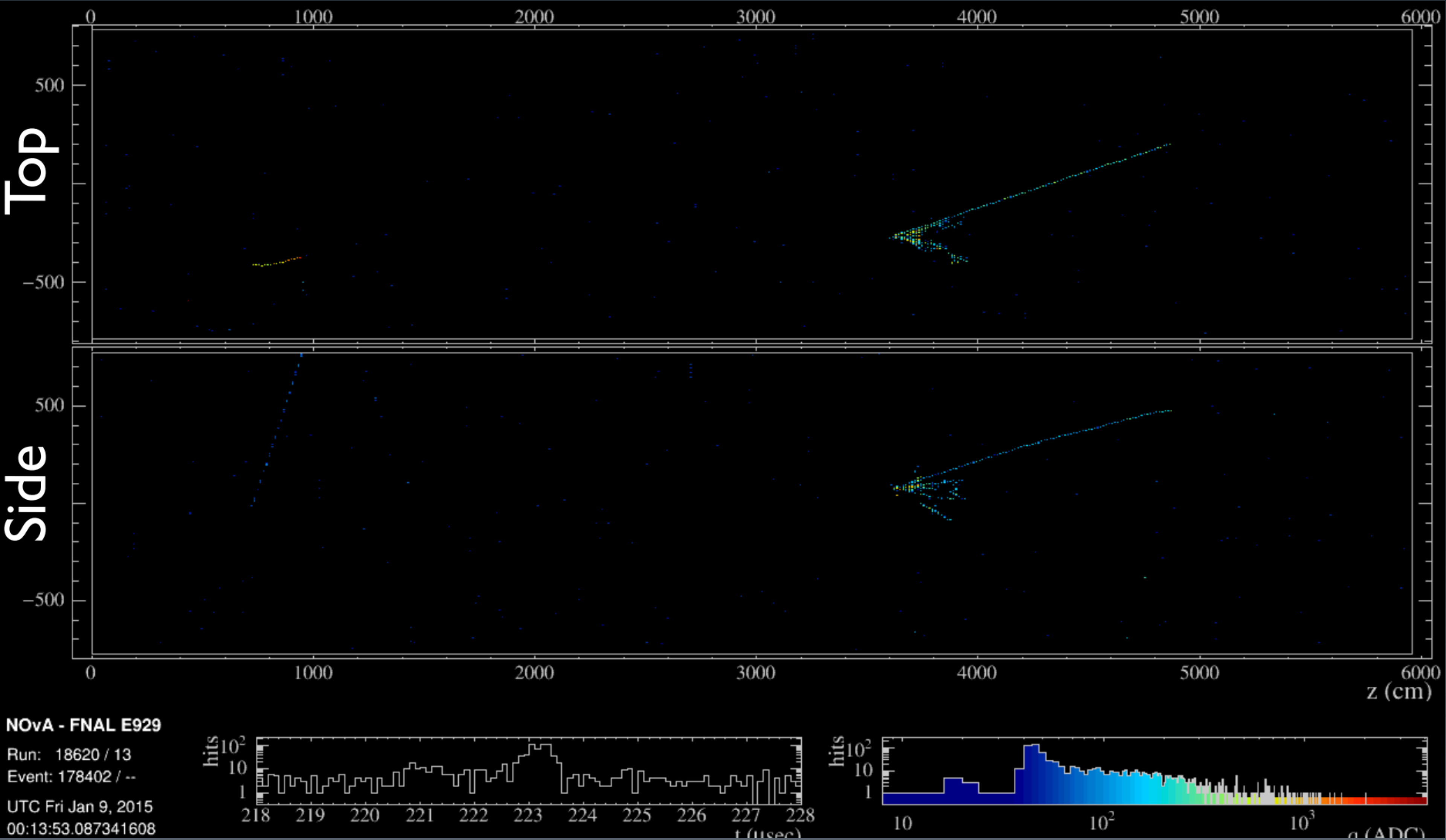
NOvA Data

550 μs window in the Far Detector

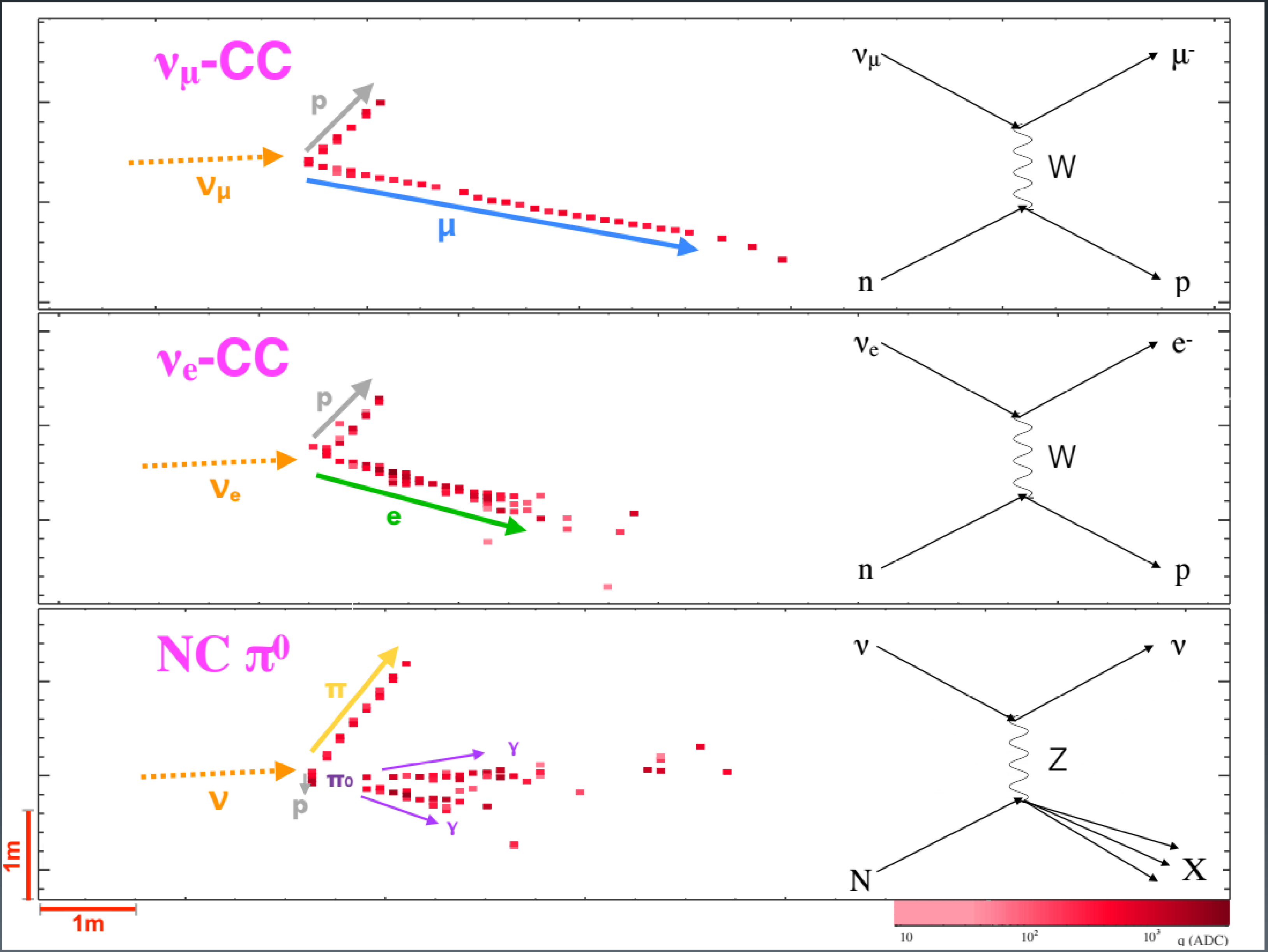


NOvA Data

10 μs window during active beam



How to Distinguish Neutrino Types?



ν_μ

ν_e

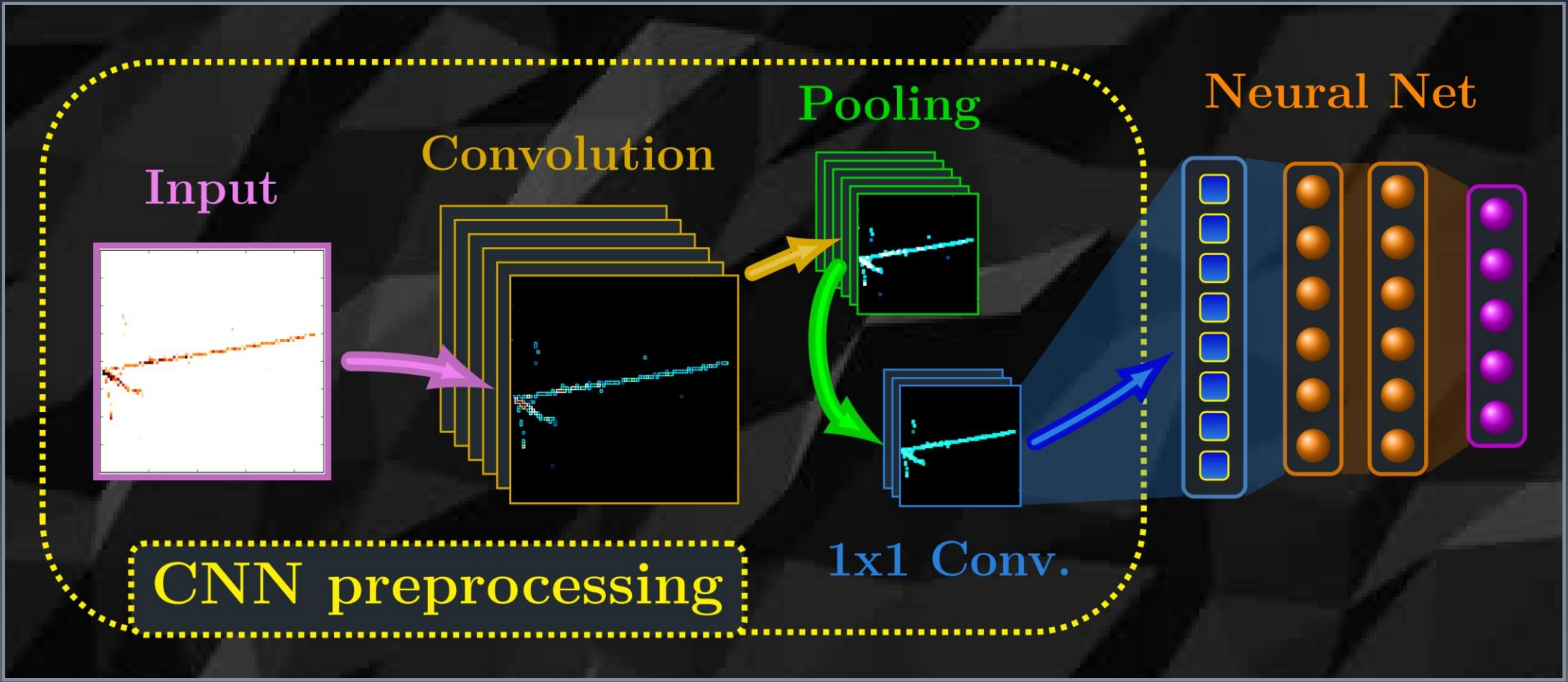
NC

Categorize Events Using CNN

A Convolutional Neural Network uses image preprocessing to reduce data flowing to the Neural Network, while identifying and retaining useful patterns.

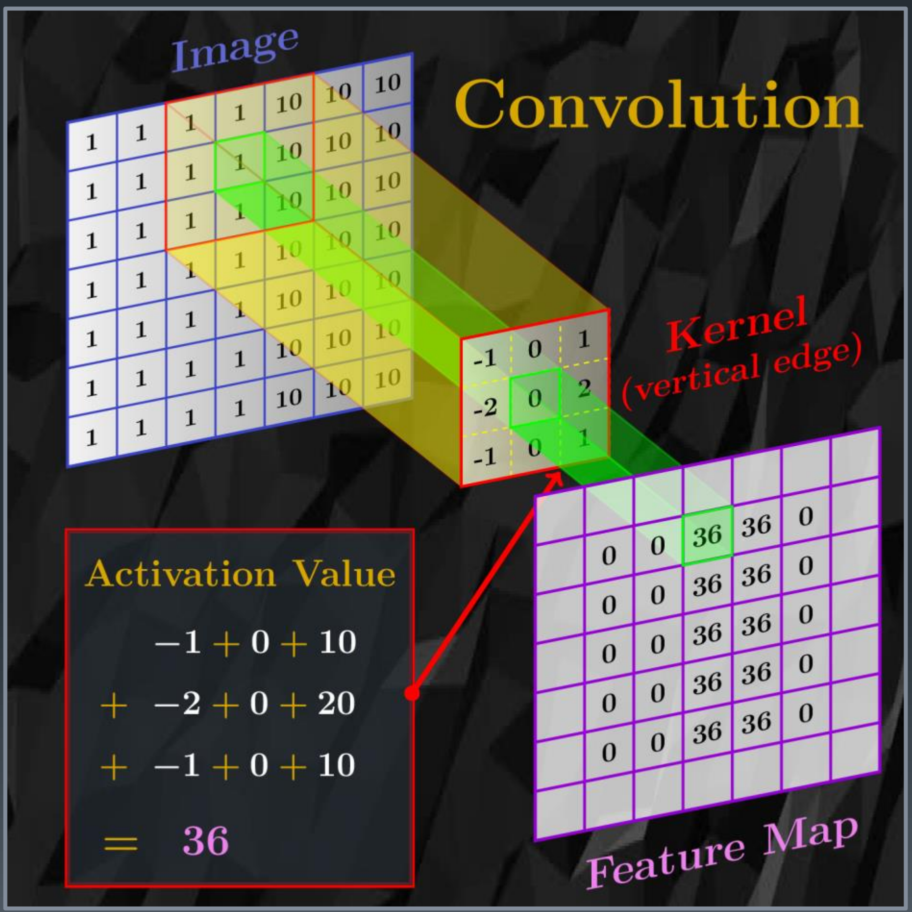
Preprocessing generally consists of:

Multiple layers of Convolution (pattern searching), Pooling (information reduction), and “1x1 Convolution” (information reduction and pattern interweaving)

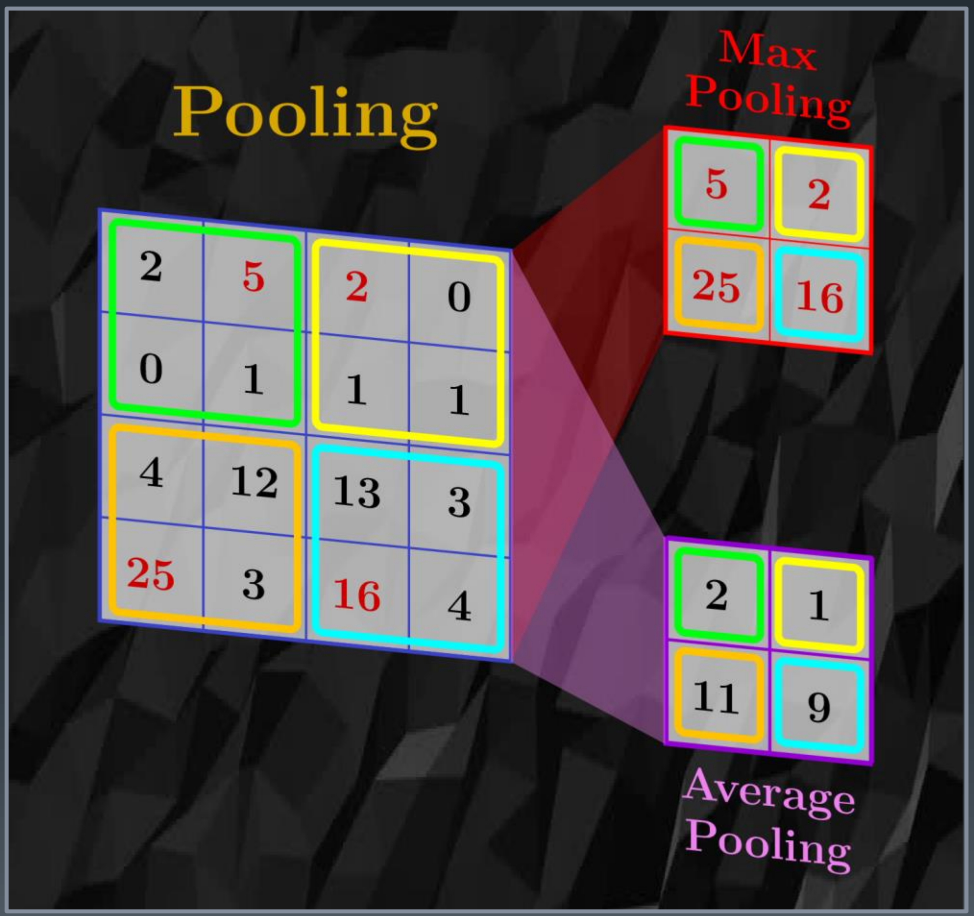


Basic Network Layers

Kernels (patterns) used in convolution are learned by the network!

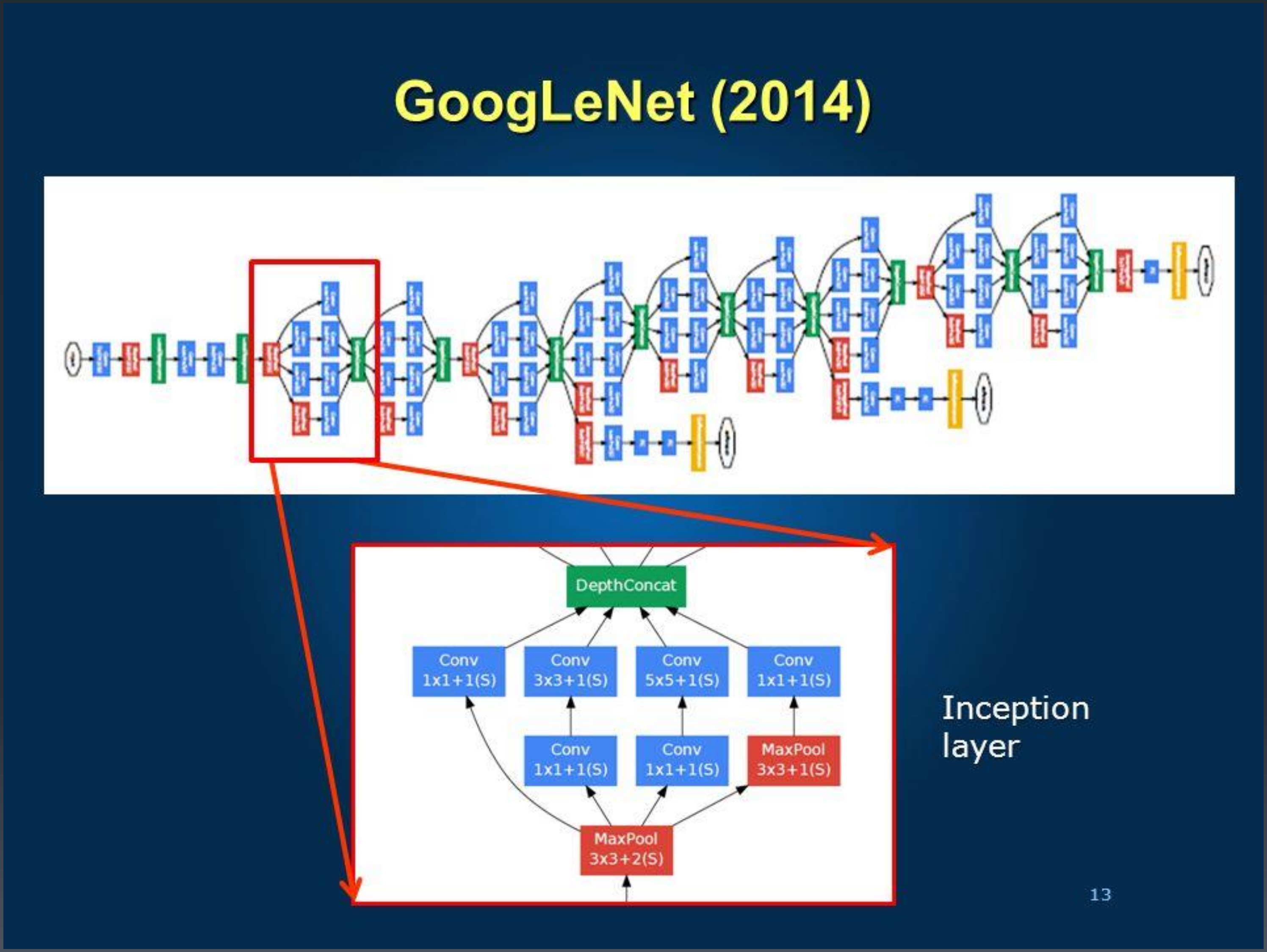


(pattern search)



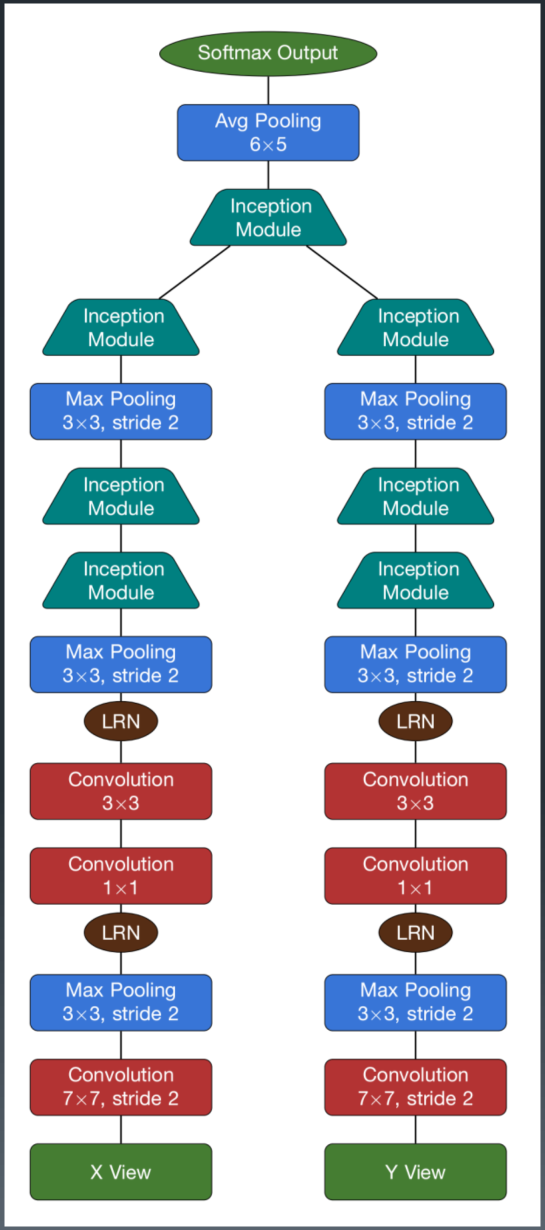
(information reduction)

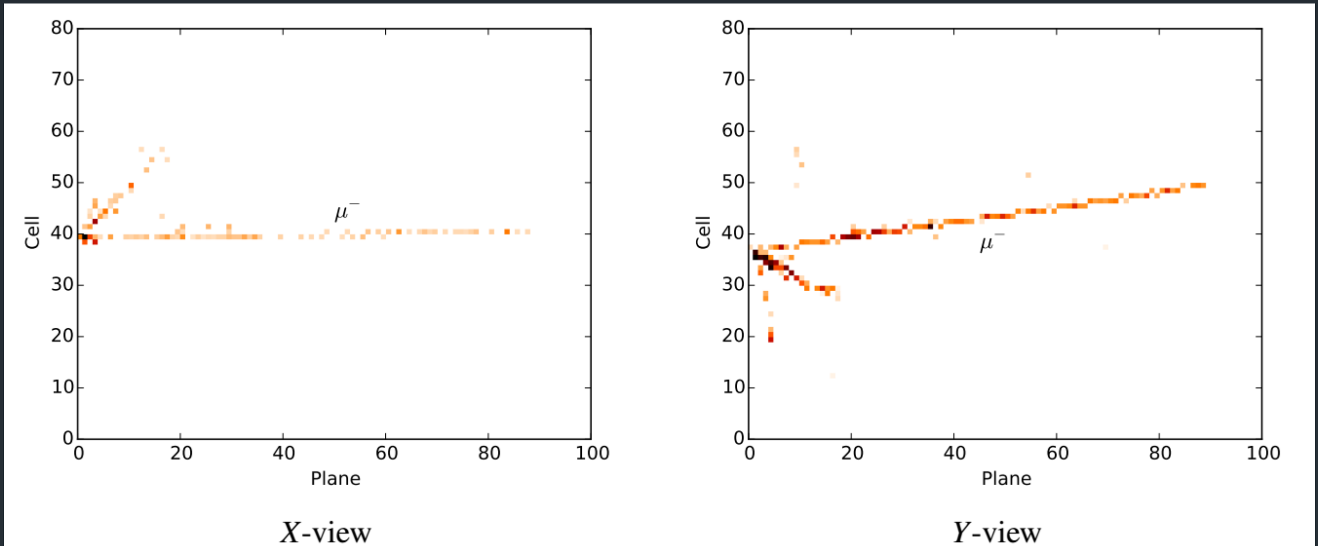
Network Inspiration



Inception modules are groups of Convolution/Pooling layers that include 1x1 Conv. to reduce the parameters in the network, allowing for deeper networks.

NOvA's Event Classification CNN

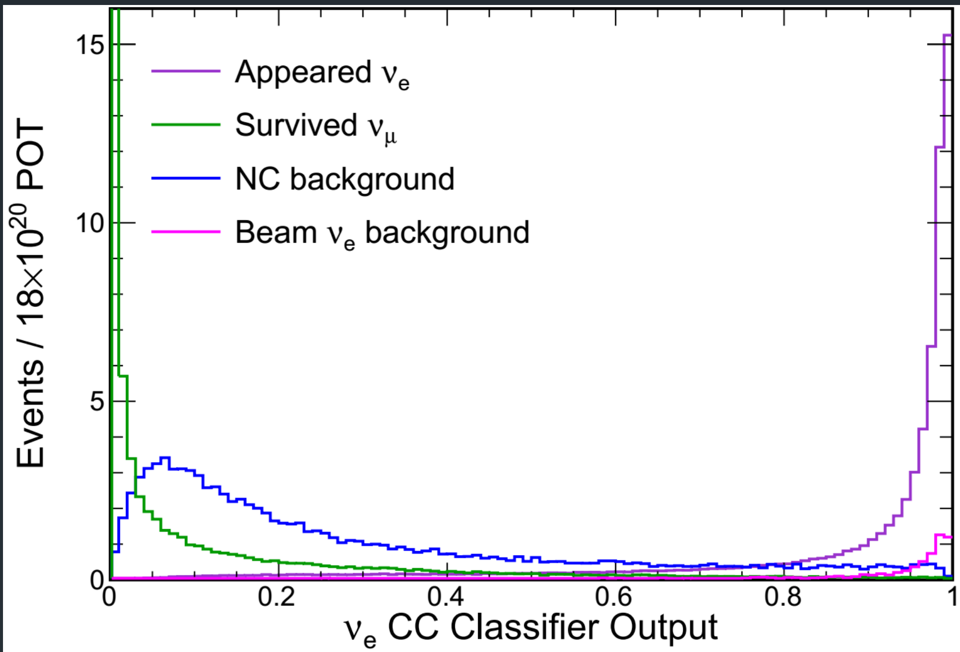


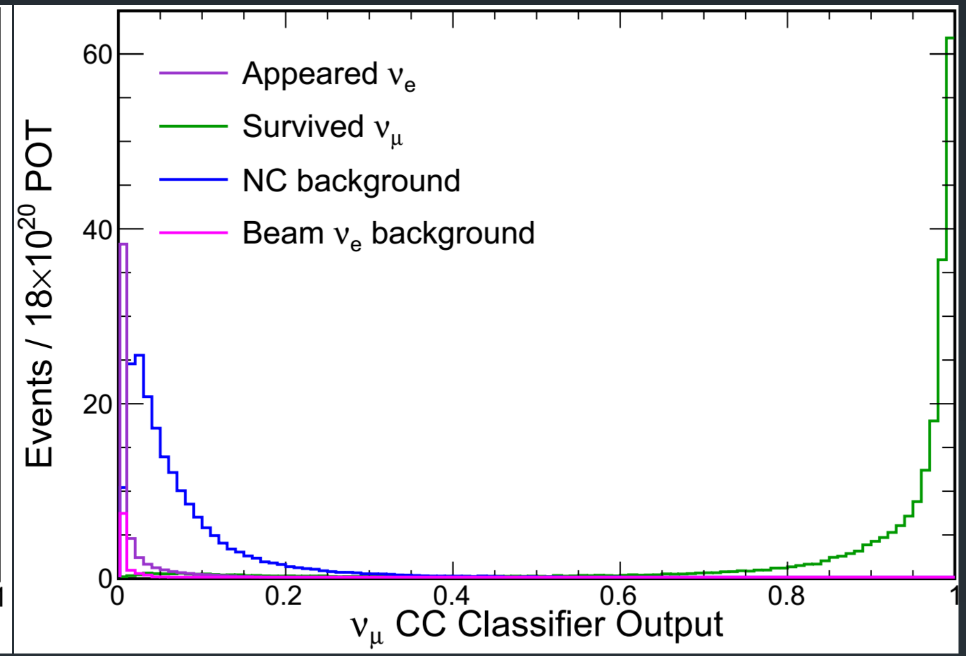


(a) ν_μ CC interaction.

- 80x100 matrix of energy deposits are extracted from the detector data. Greatly reduces data flow through network
- X and Y-views are sent separately through the two towers of NOvA's network
- The towers are merged, and the network outputs a value for each interaction type, which is Softmax normalized

Event Classification Network Results





Published. 2016

JINST 11 (2016) no.09, P09001

A Convolutional Neural Network Neutrino Event Classifier

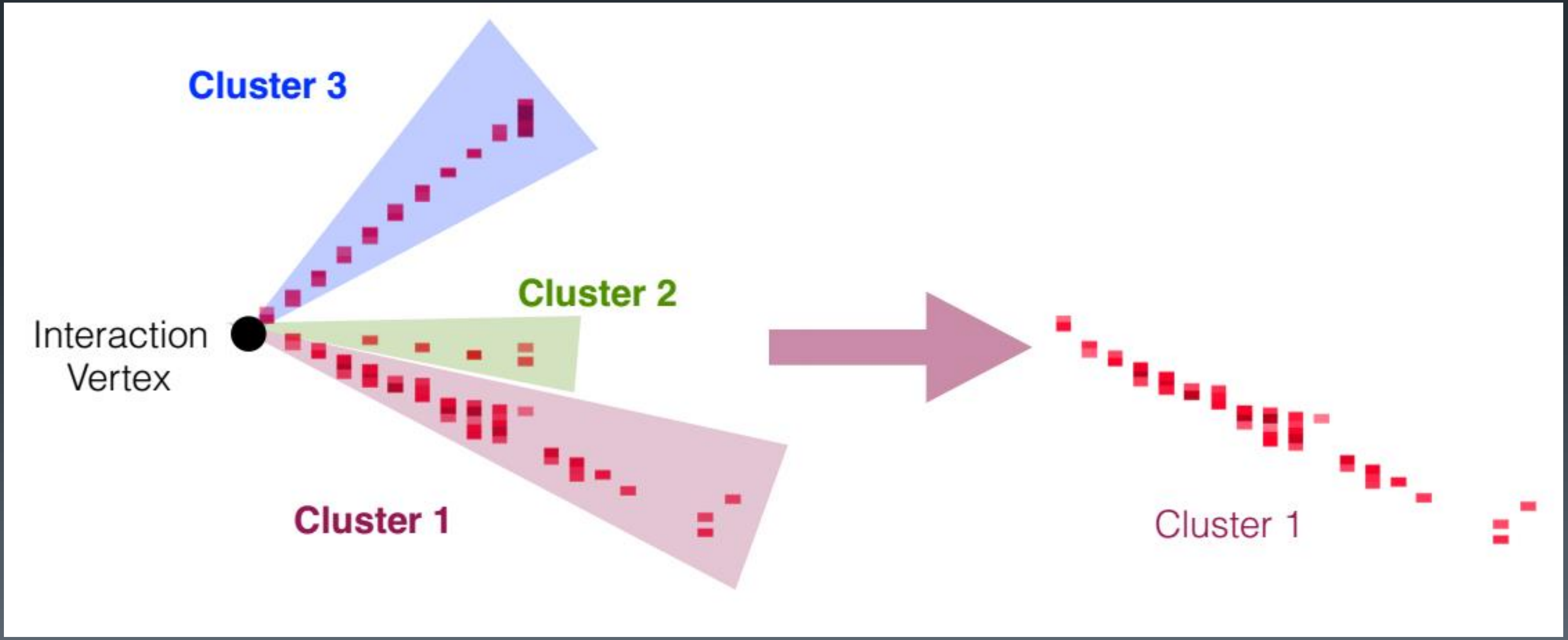
A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, P. Vahle

(Submitted on 5 Apr 2016 (v1), last revised 12 Aug 2016 (this version, v3))

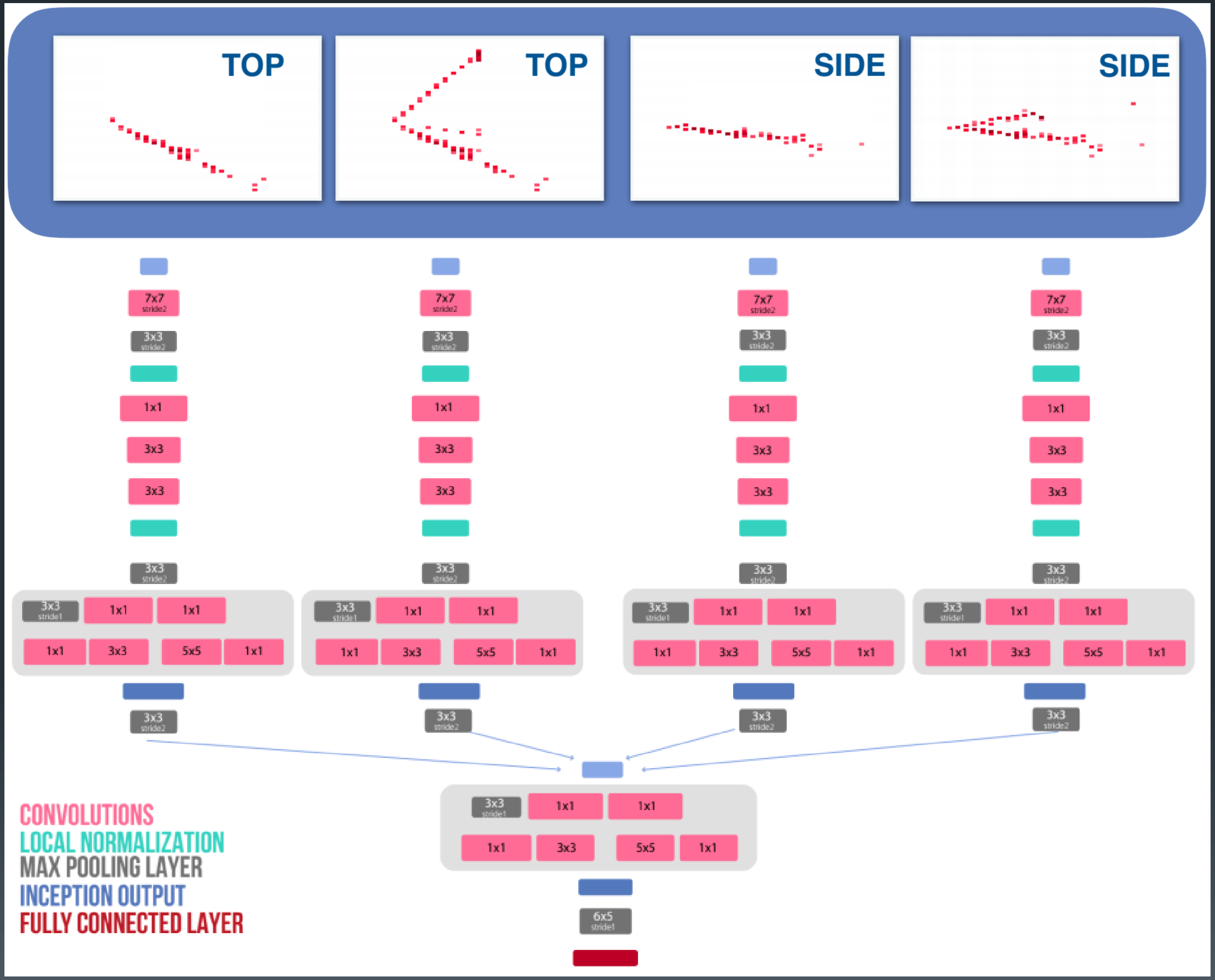
Particle Identification

Particle identification is necessary for in-depth physics analysis.

- Neutrino interactions in the detector are simulated
- Hits are clustered into tracks
- An interaction vertex is determined
- Single particle tracks are separated for training

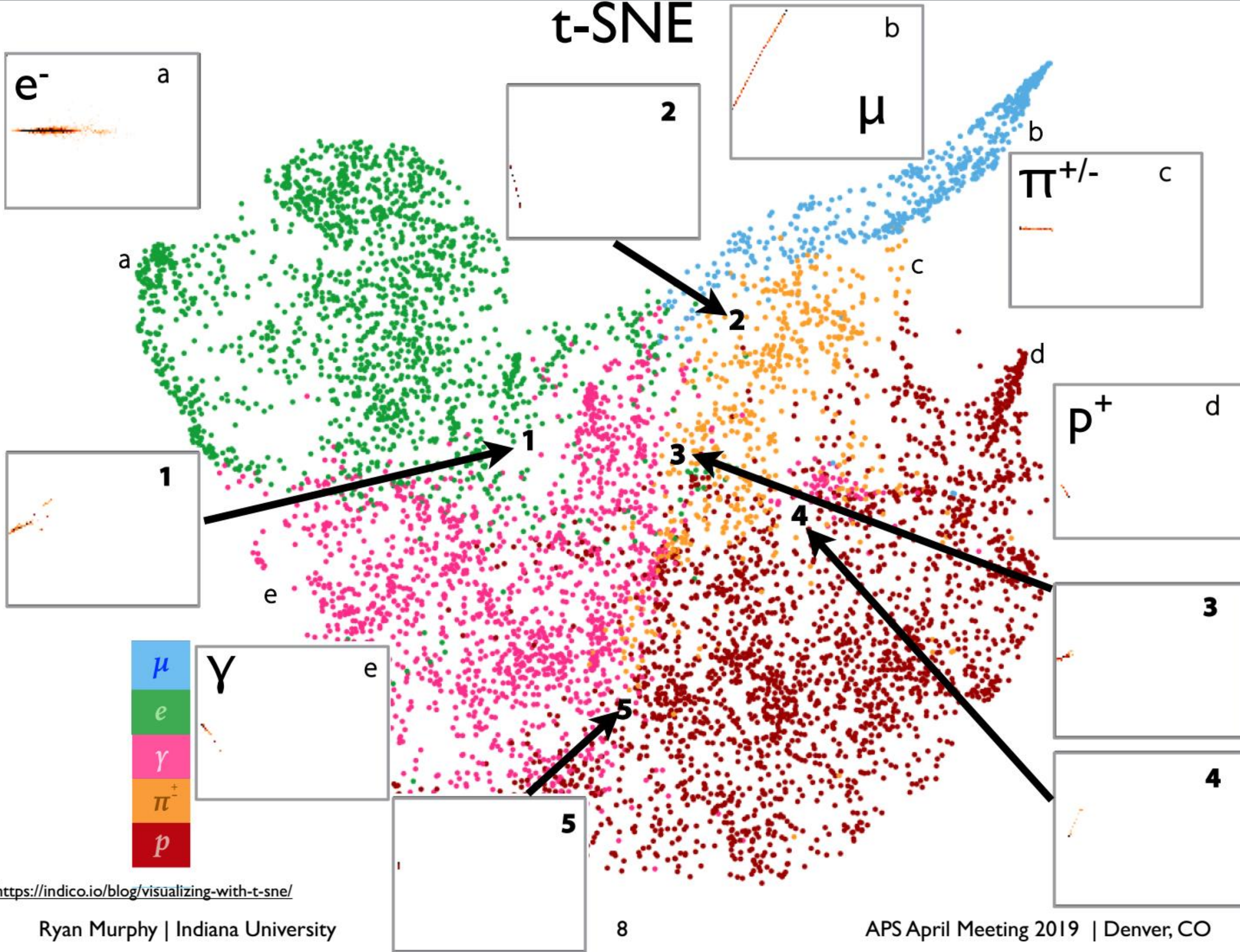


Particle ID Network



Context provides up to an 11% improvement in efficiency and purity!

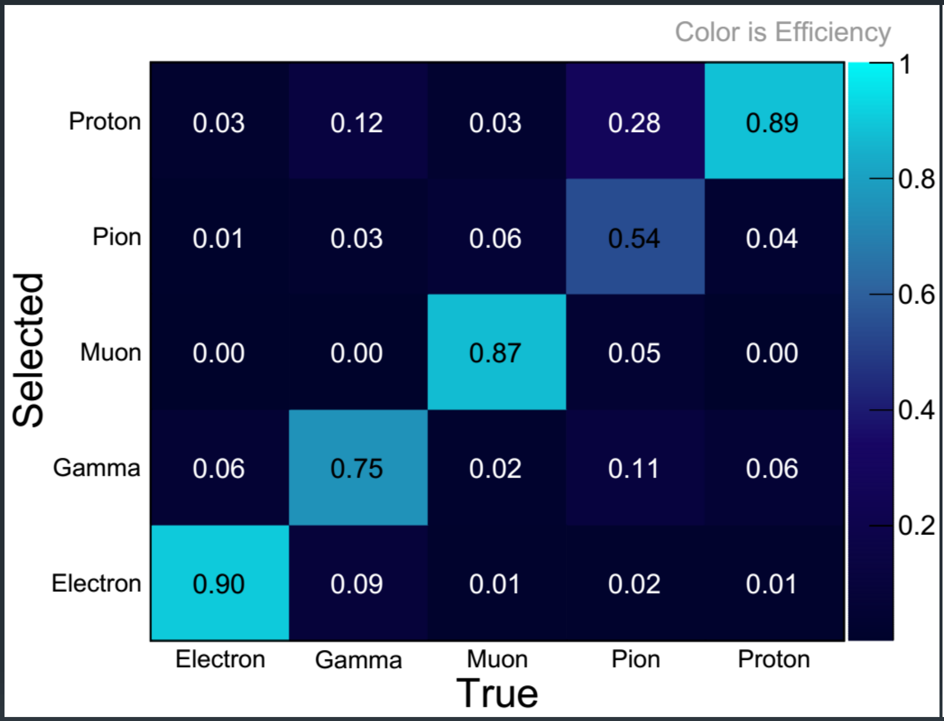
t-SNE



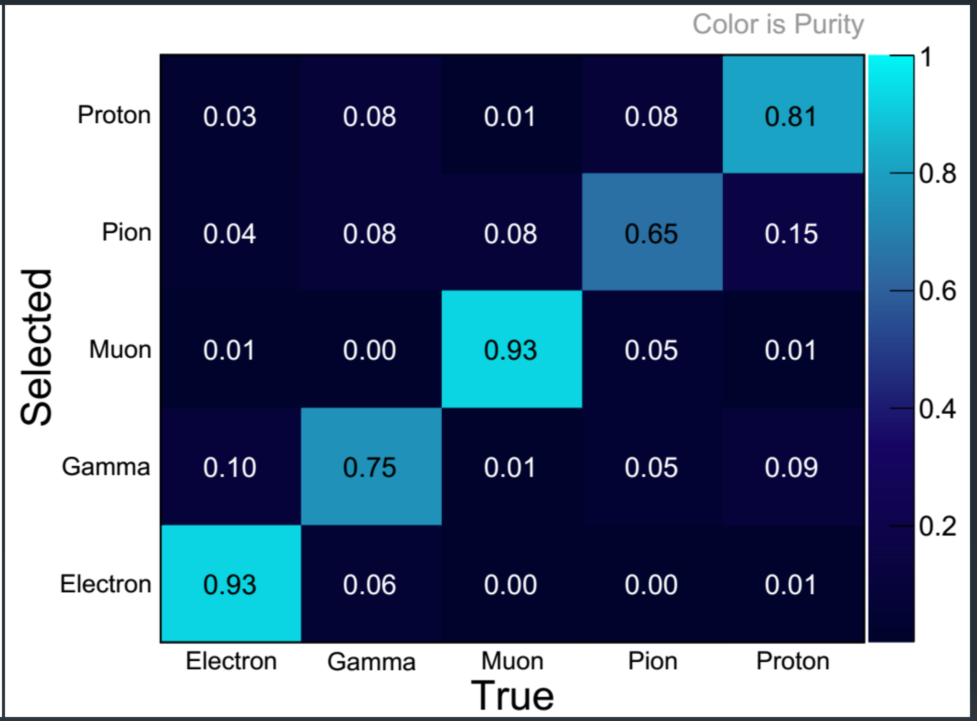
https://indico.io/blog/visualizing-with-t-sne/

Particle ID Network Results

Efficiency



Purity



Using this network as an input to the electron neutrino energy estimator the energy resolution is 11%, an improvement of 20% compared to previous methods.

Submitted to PRD

arxiv:1906.00713

Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events

F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic, A. Sousa

Future Implementations/Improvements

- Continue to improve event classification network and particle ID network

- NOvA has created a CNN to improve ν_e and electron energy estimation:
  PRD: DOI: 10.1103/PhysRevD.99.012011

- Creation of LSTM network to improve ν_μ energy estimation

- CNN to reduce cosmic ray background

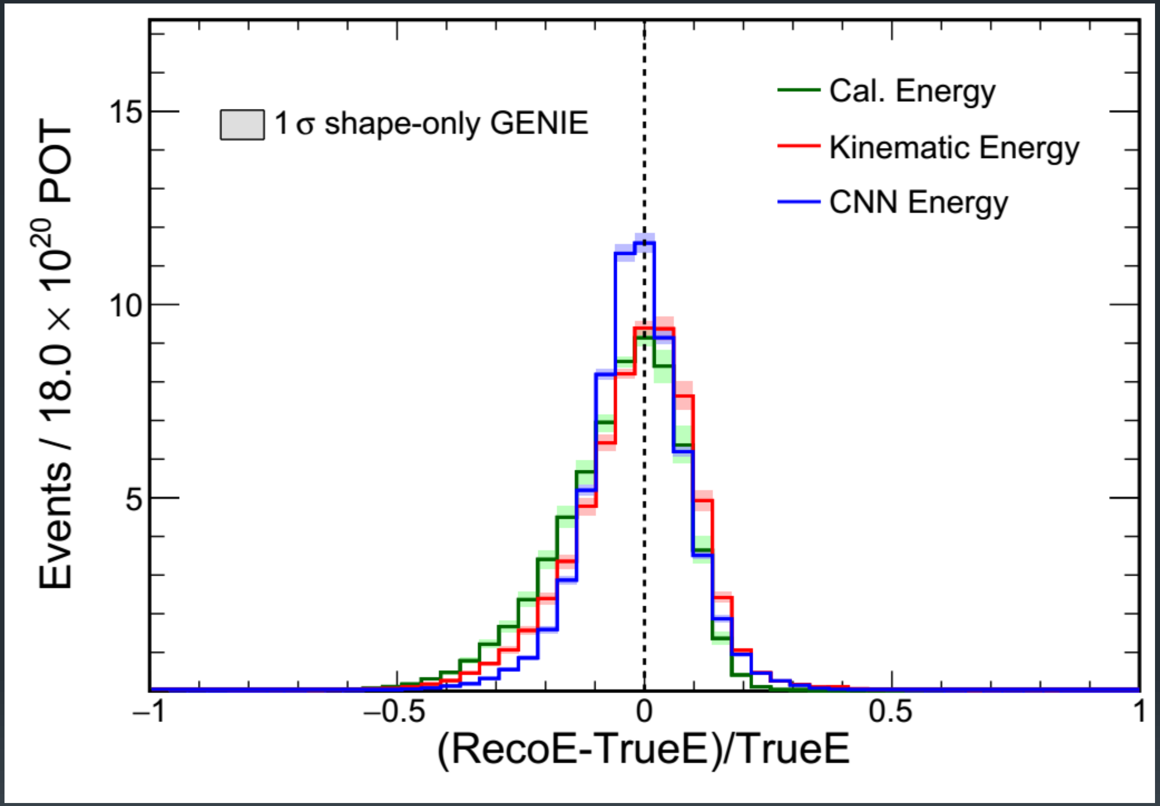
- The NOvA test beam detector will provide labeled data from single-particle interactions allowing for data-driven checks of deep learning methods

- And more….

Backup Slides

ν_e Energy Estimator Network

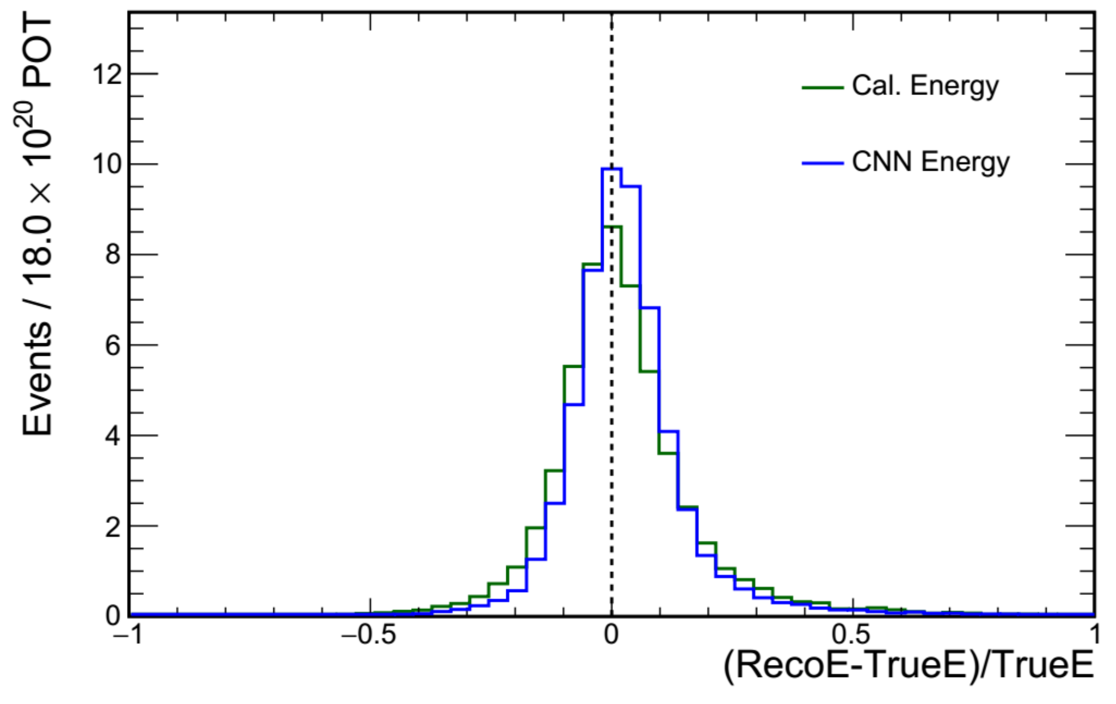
- Uses flat neutrino flux to increase low energy events
- 0.98 Million ν_e interactions
  0.75/0.23 Mil. train/validate
- Keras/Tensorflow
- Used SHERPA to optimize hyperparameters



The network uses a modified version of the classification network, where the event's vertex position is concatenated with the preprocessing output to correct for readout threshold.

Electron Energy Estimator Network

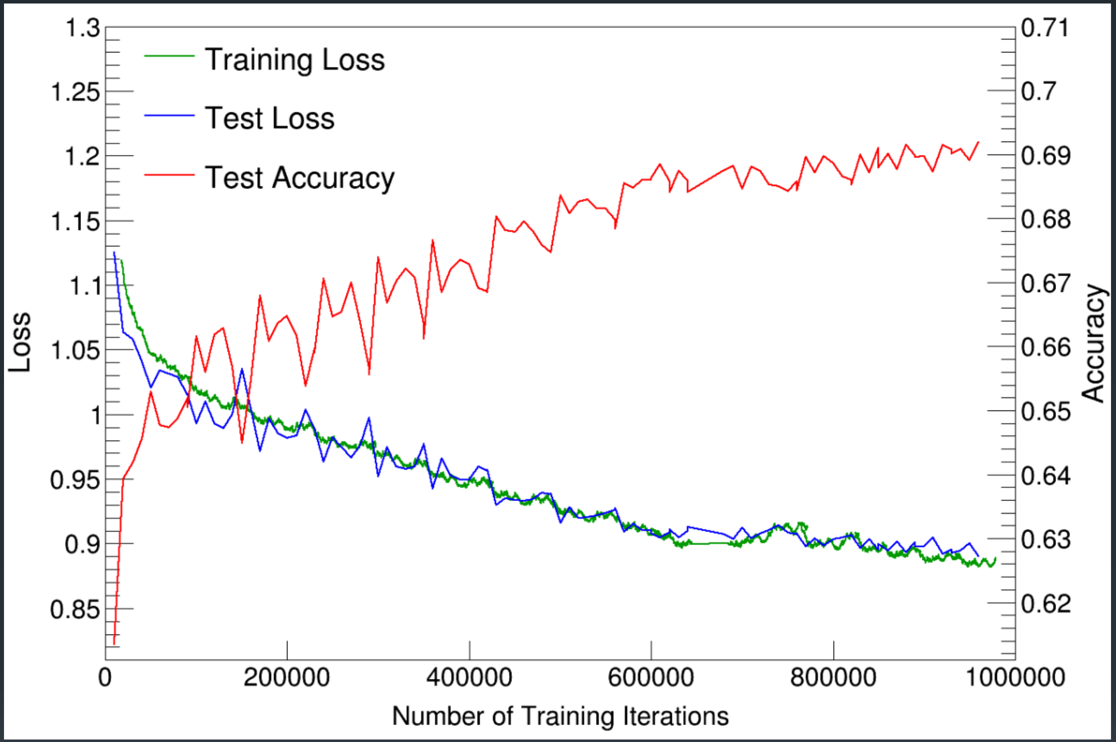
- Uses standard NOvA neutrino flux
- 660 K electron tracks
  610 K / 5 K train/validate
- Keras/Tensorflow
- Used SHERPA to optimize hyperparameters



Uses the same network structure as the ν_e energy estimator, but with different weights. Both use 151x141 pixel detector readouts as input, which include 99.5% of hits.

Classification Network Training

- 4.7 Million events for training (80%/20% train/validate)
- Mini-Batches of 32 events
- Dropout rate of 0.4
- 1 week of GPU-hours on 2 NVIDIA Tesla K40s
- Uses Caffe



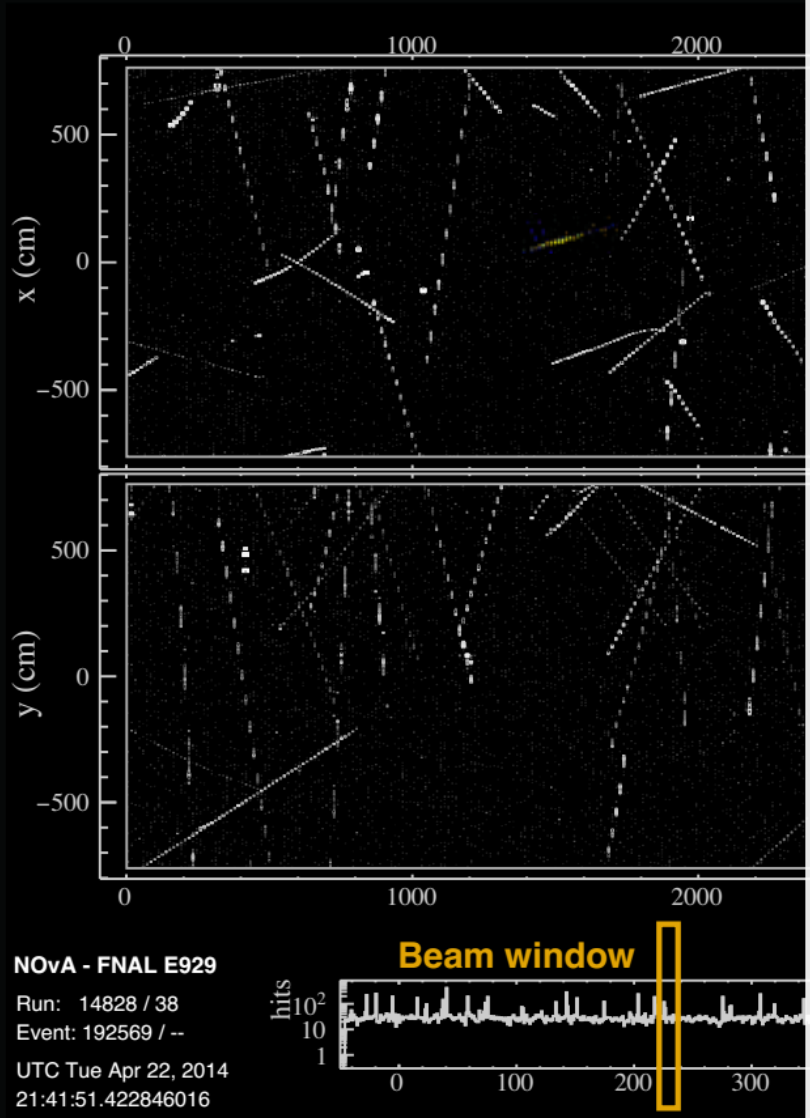
Pixel intensities varied with Gaussian noise, and a portion of events were reflected along the z-axis (nearly parallel to beam) to increase sample size and reduce the importance of individual pixel intensities.

Particle ID Network Training

- 2.95 Million particles for training (80%/20% train/validate)
- Mini-Batches of 64 events
- 700,000 iterations
- 4 NVIDIA Tesla K20 GPUs
- Uses Caffe

Event selection of π^0 decays (photon pairs) has a purity of 92%, which is a great improvement over 60% gotten from previous methods!

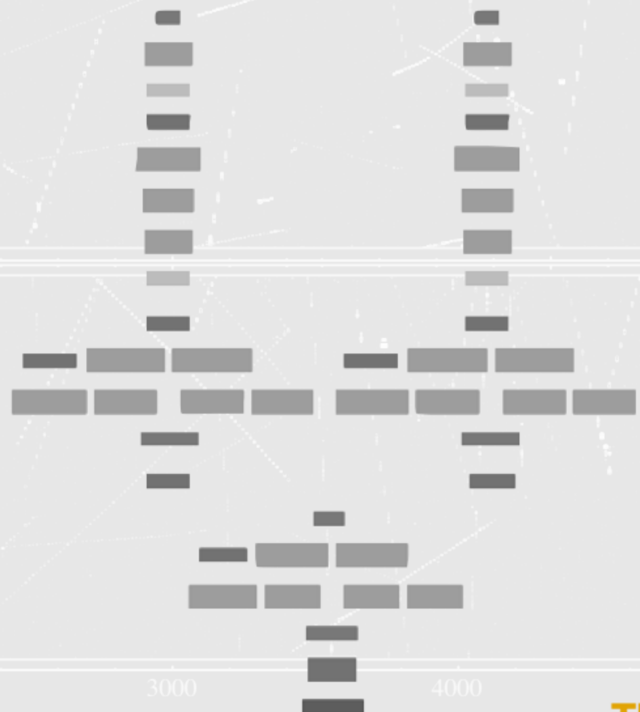
Reducing Cosmic rays



Under development

NOvA is exploring cosmic ray background reduction with CNNs.

INPUTS: X and Y views of the whole detector



CVN-cosmic veto uses a CVN architecture with Resnet-18 backbone

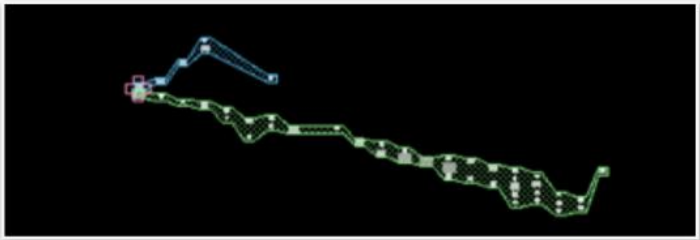
Using TensorFlow and HDF5s.

Evaluated on GPUs.

The target is to reject time slices of full detector readout.

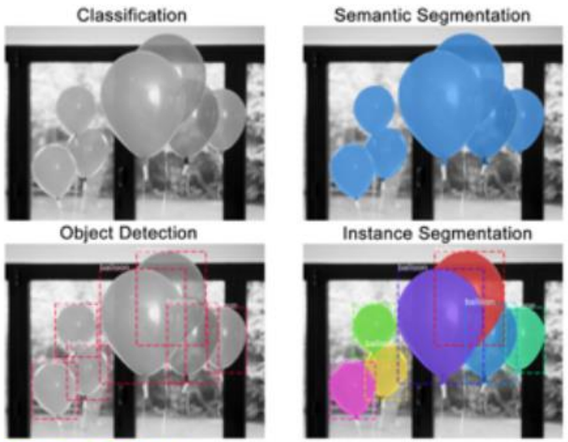
Particle Identification with CNNs

We use our Deep CNN classifier to identify each cluster.



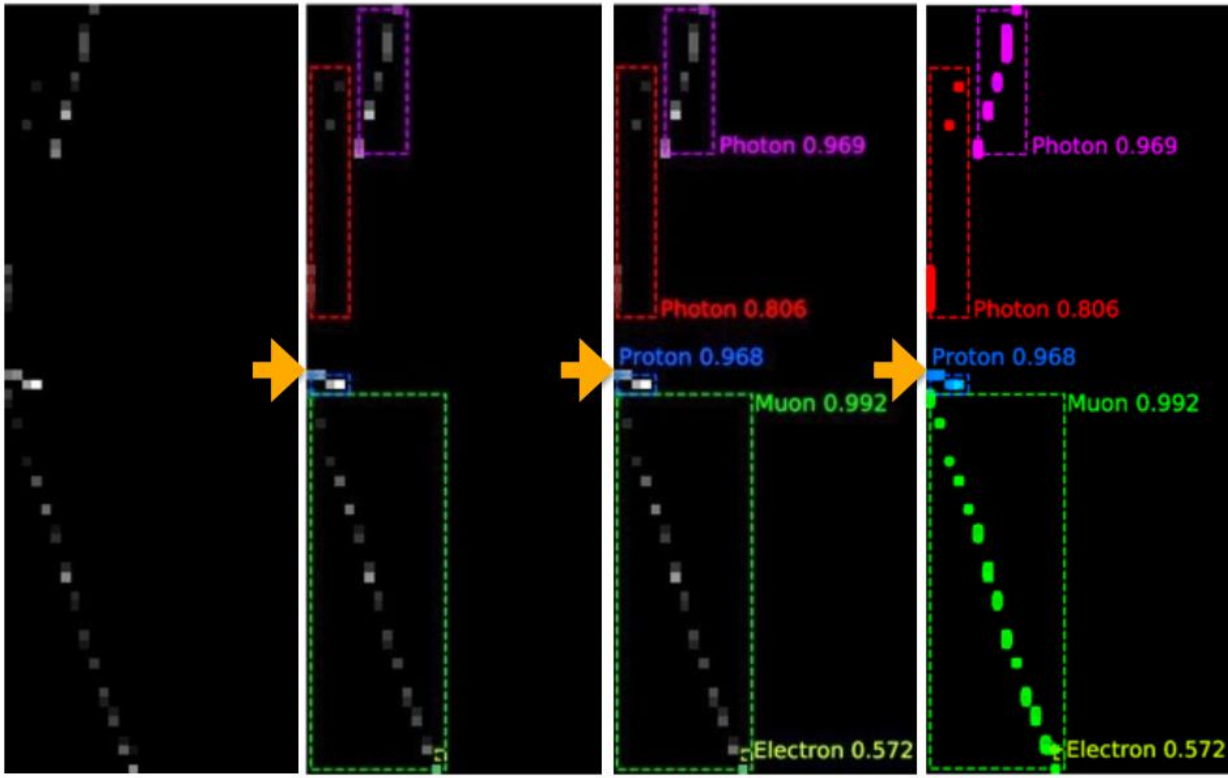
Localizing + Clustering + Identification

In development

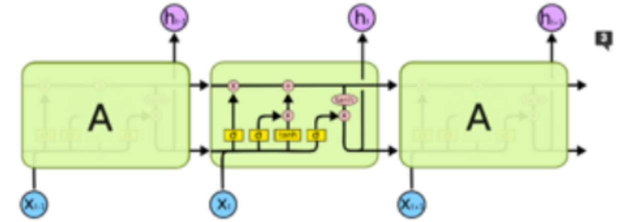


MASK-RCNN

We use an adaptation of MASK-RCNN to cluster and identify all activity.



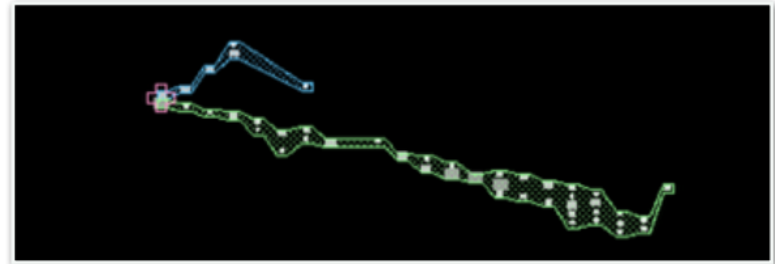
Energy Reconstruction with LSTMs



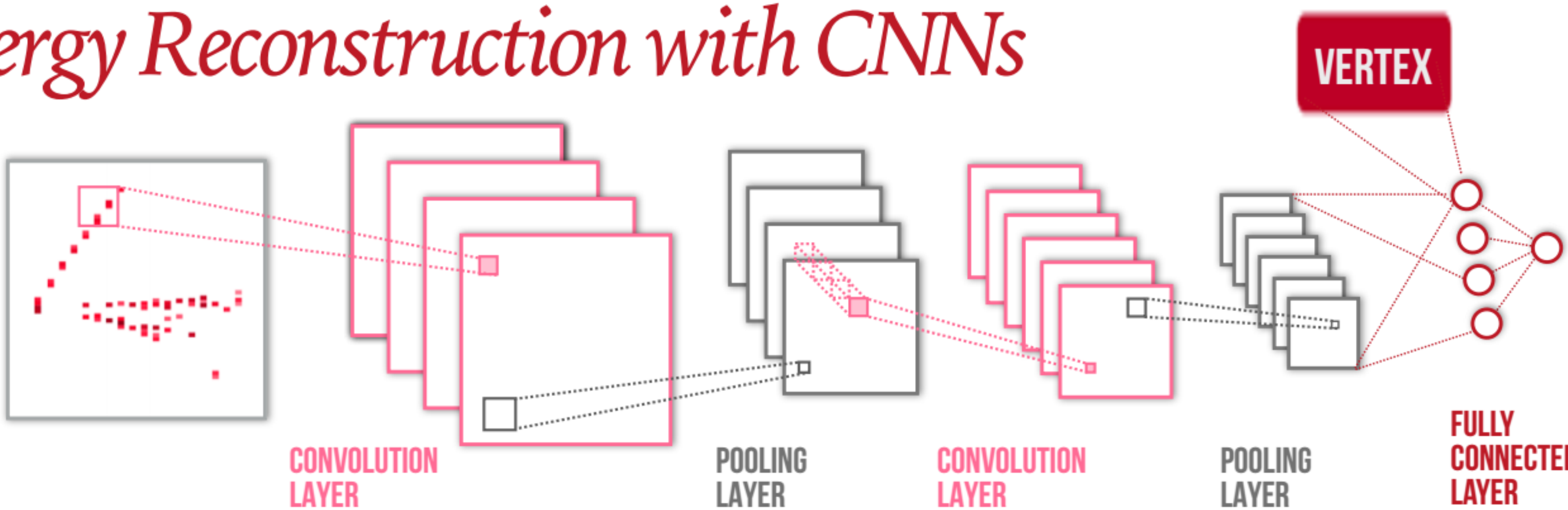
Recurrent Neural Networks:

Sequential network using the current state of the system + the output from last iteration.

LSTM: RNN + Long term memory cell



Energy Reconstruction with CNNs



The target is Energy instead of PID values.

see poster
from yesterday's session K01.00029

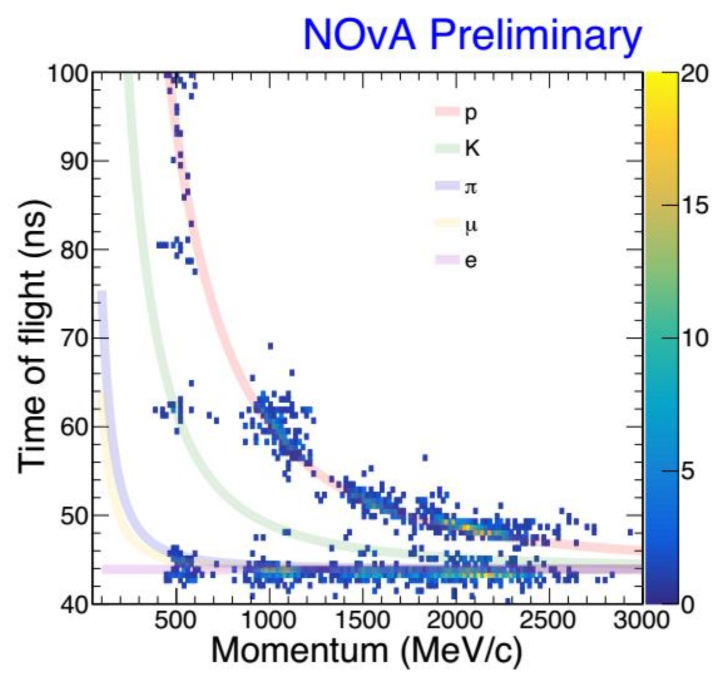
Published. 2019

Phys.Rev. D99 (2019) no.1, 012011

Improved Energy Reconstruction in NOvA with Regression Convolutional Neural Networks

Pierre Baldi, Jianming Bian, Lars Hertel, Lingge Li

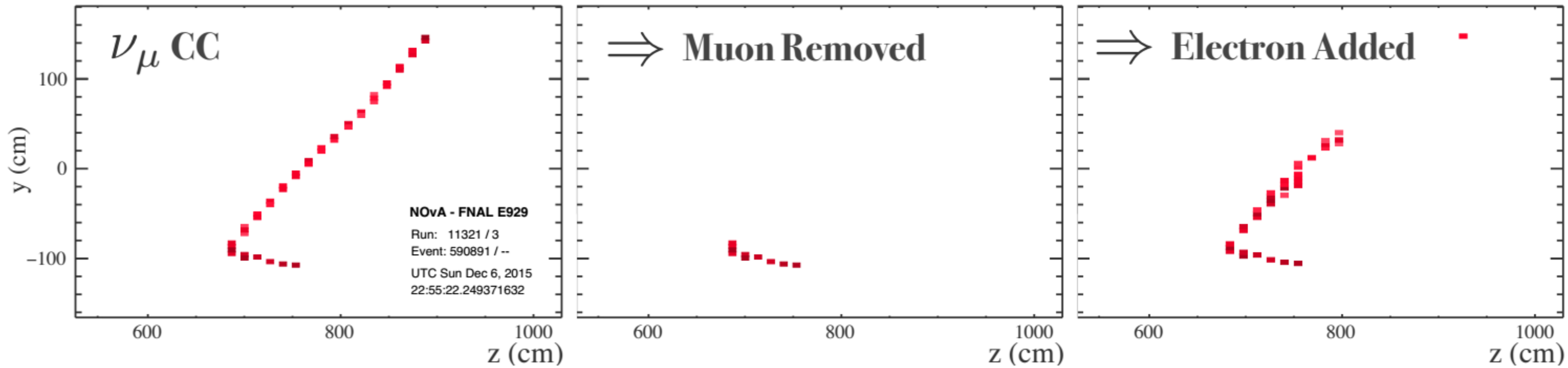
NOvA Test Beam

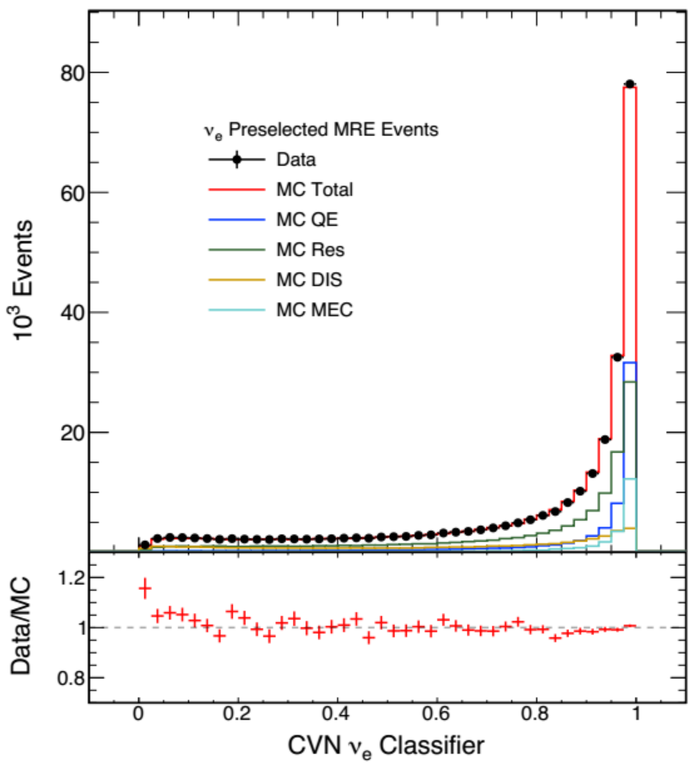


The NOvA test beam detector is being currently taking data and will continue throughout 2019.

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.

Data-driven test example





MRE (Muon Removed - Electron added):

Select a muon neutrino interaction with traditional ID methods.

Remove the muon hits and replace them with a single simulated electron of matching momentum.

Data/MC comparisons show less than 1% difference in efficiency.

PID	Sample	Preselection	PID	Efficiency	Efficiency diff %
CVN	Data	262884	188809	0.718222	-0.36%
	MC	277320	199895	0.720809	