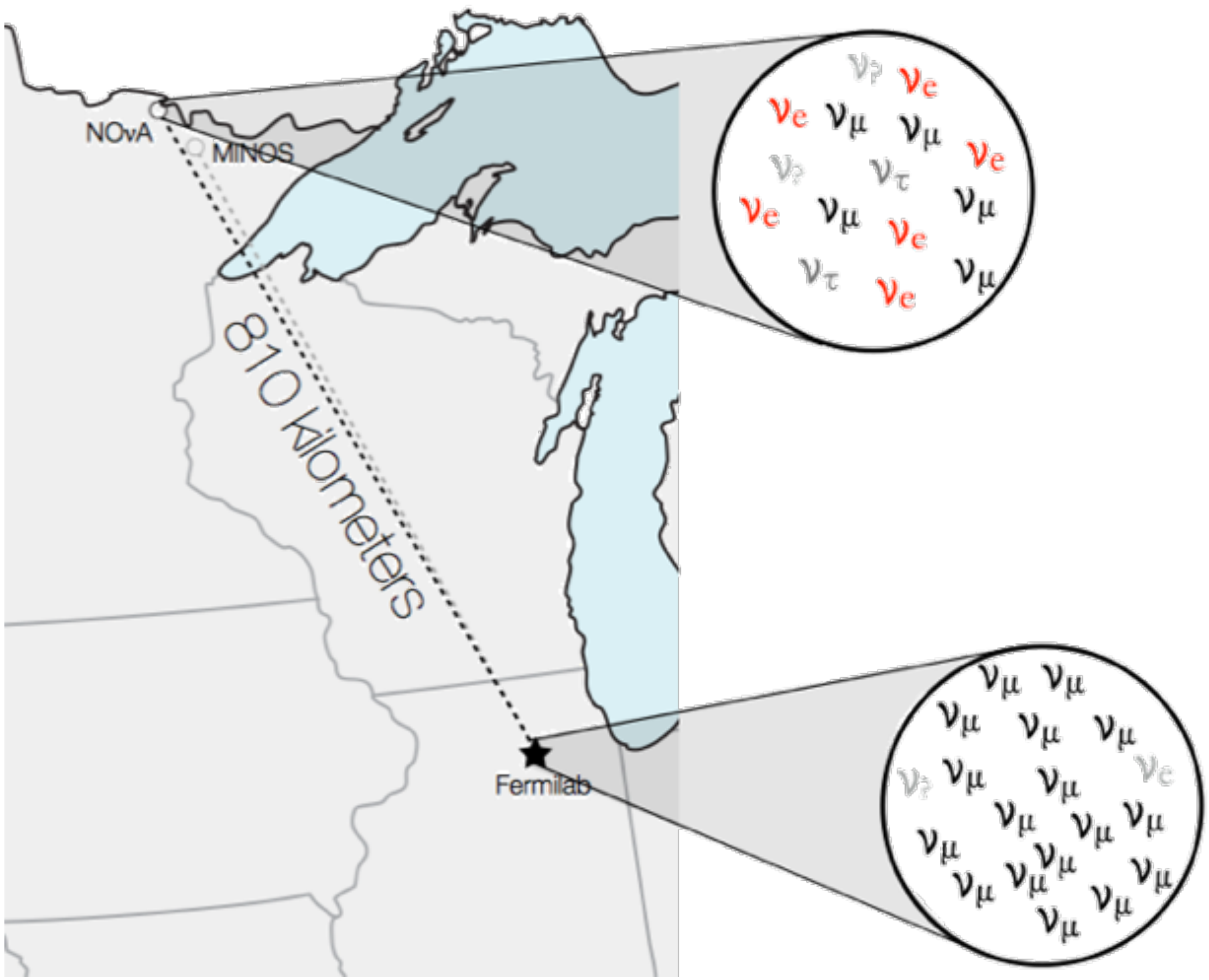
Machine Intelligence Applications for Particle Physics @Fermilab

For the Machine Intelligence and Reconstruction Group
Aristeidis Tsaris - atsaris@fnal.gov
Machine Learning Seminar @ JLab // November 6, 2018

Science @Fermilab

- Fermilab is **America's particle physics and accelerator laboratory**.
- **Diverse program:** neutrinos, collider physics, muons, astronomy and cosmology, theory, dark matter and dark energy searches.
- We are pioneers in detector technology, computing, and quantum initiatives.
- Fermilab is using machine learning across all programs.
- Today: *case studies from NOvA.*

The NOvA Experiment

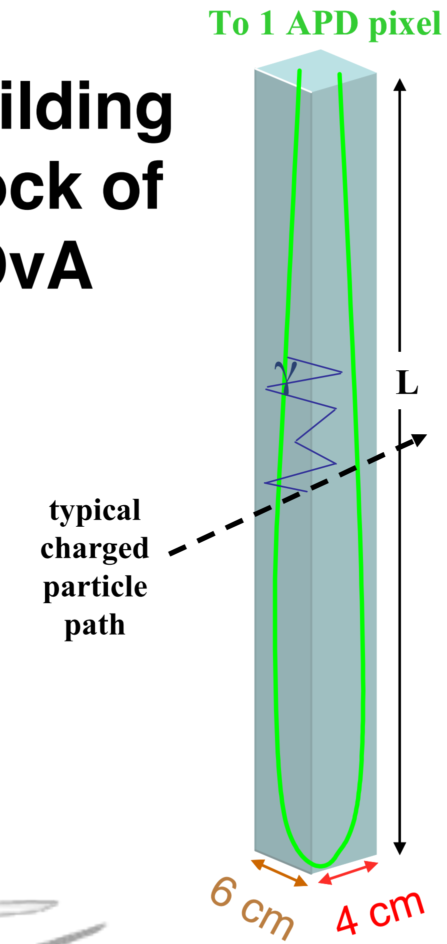


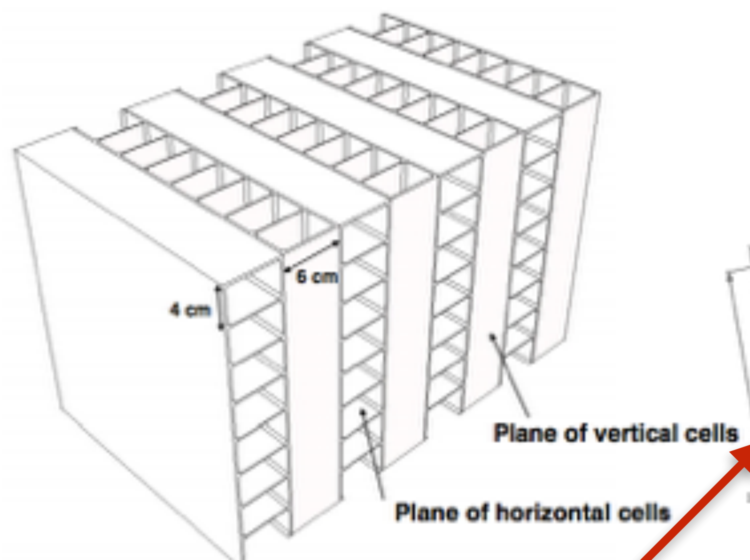
- **NuMI Off-axis ν_e Appearance Experiment**
- NuMI: Neutrinos at the Main Injector
- Long-baseline (anti-)neutrino oscillation experiment
- Two functionally identical detectors, optimized for ν_e identification
- Primary goal: measurement of 3-flavor oscillations via $\nu_\mu \rightarrow \nu_\mu$ and $\nu_\mu \rightarrow \nu_e$
- Other goals include:
  - Searches for sterile neutrinos
  - Neutrino cross sections
  - Supernova neutrinos
  - Cosmic ray physics

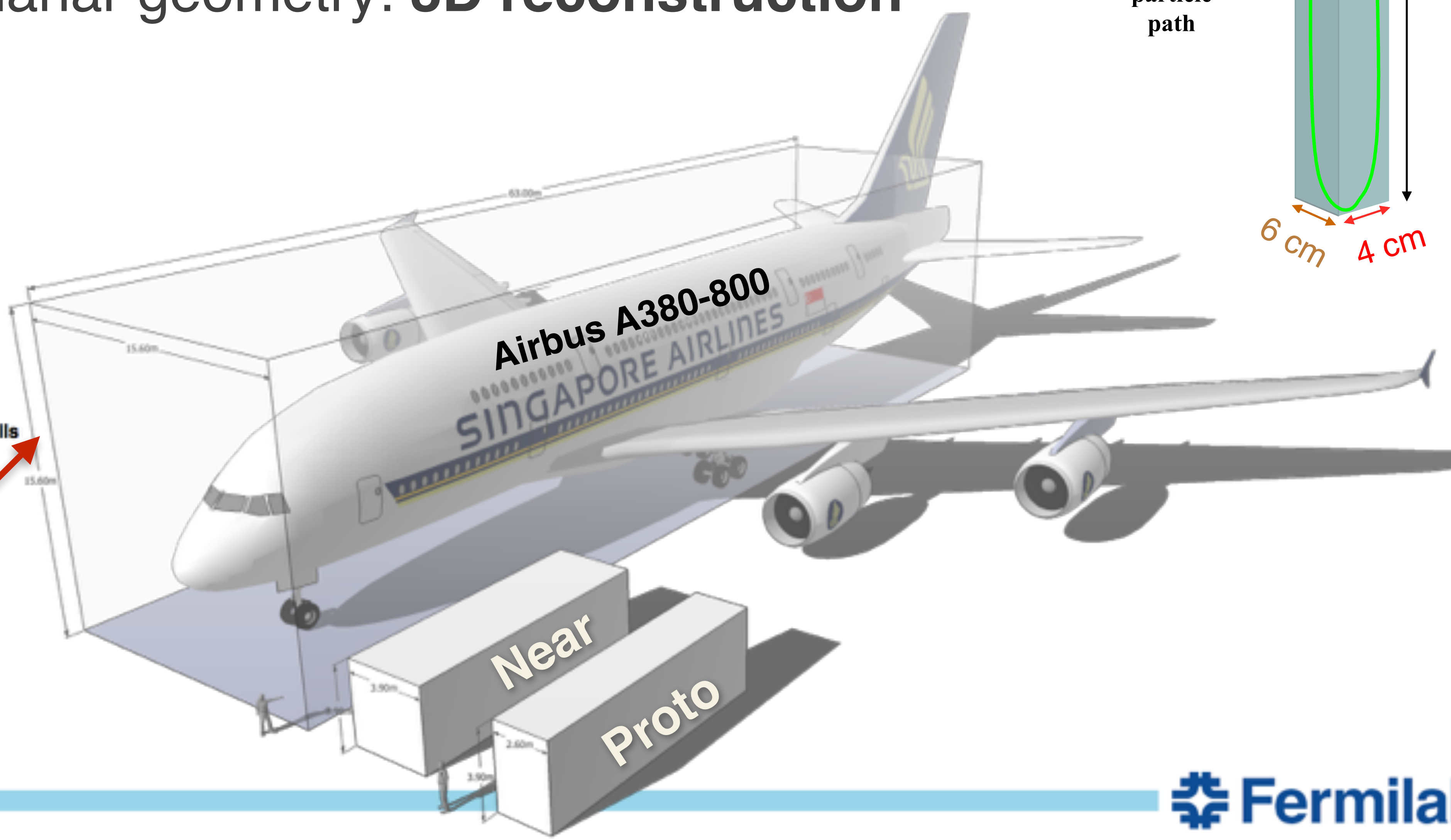
NOvA Detectors

- Highly segmented low Z tracking calorimeter.
- Cells are filled wit liquid scintillator
  - wave shifting fiber readout.
- 65% active by volume
- Detection with avalanche photo diodes.
- Alternating X/Y planar geometry: **3D reconstruction**

Building block of NOvA



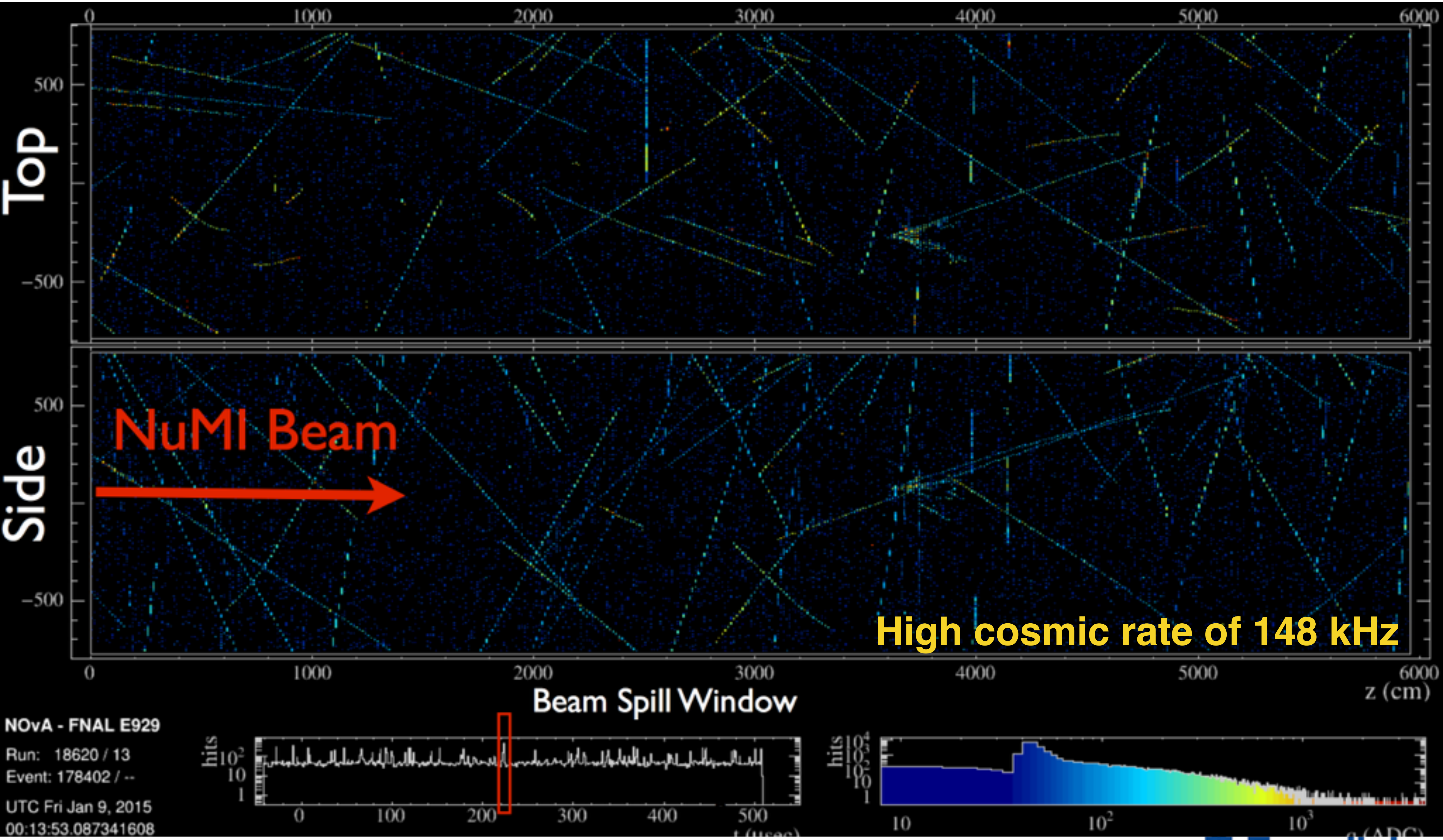




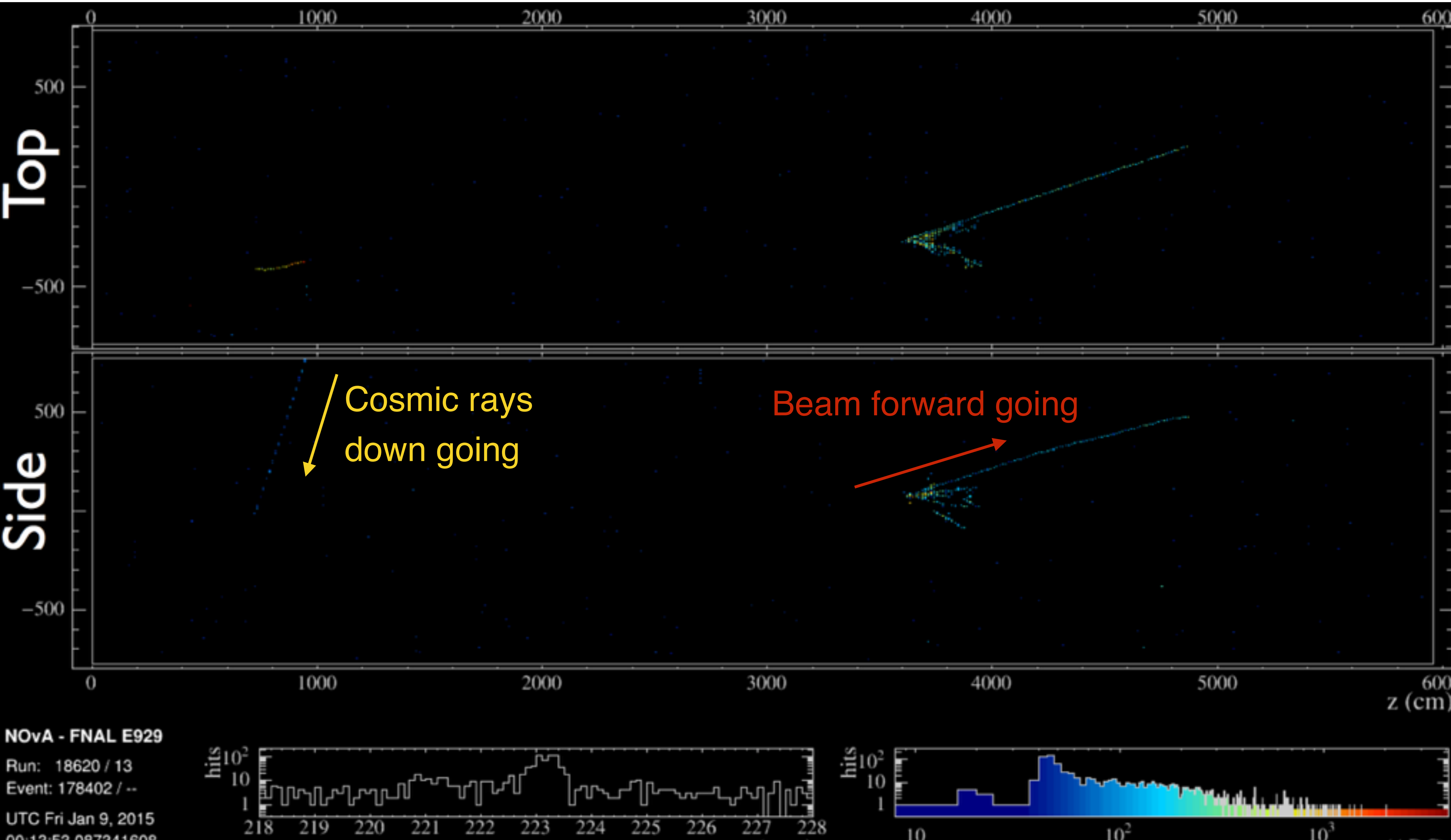
Far Detector, on surface
14 kTons
896 readout planes
344,064 pixels

Part I
Neutrino Flavor Classification

Can You Find the Neutrino?

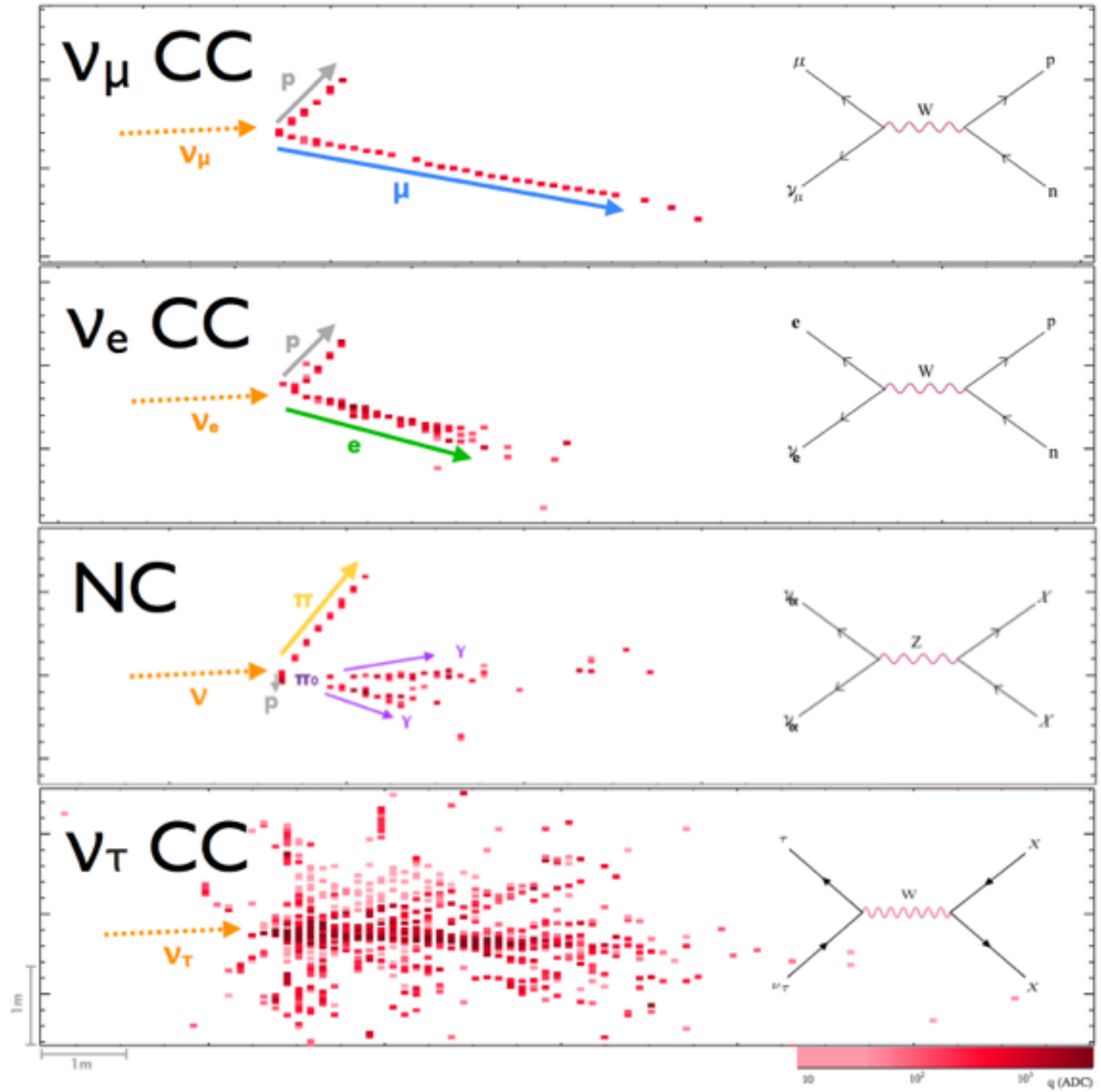


Zooming in …


Top
Side
Cosmic rays
down going
Beam forward going
z (cm)
NOvA - FNAL E929
Run: 18620 / 13
Event: 178402 / --
UTC Fri Jan 9, 2015
00:13:53.087341608
hits
hits

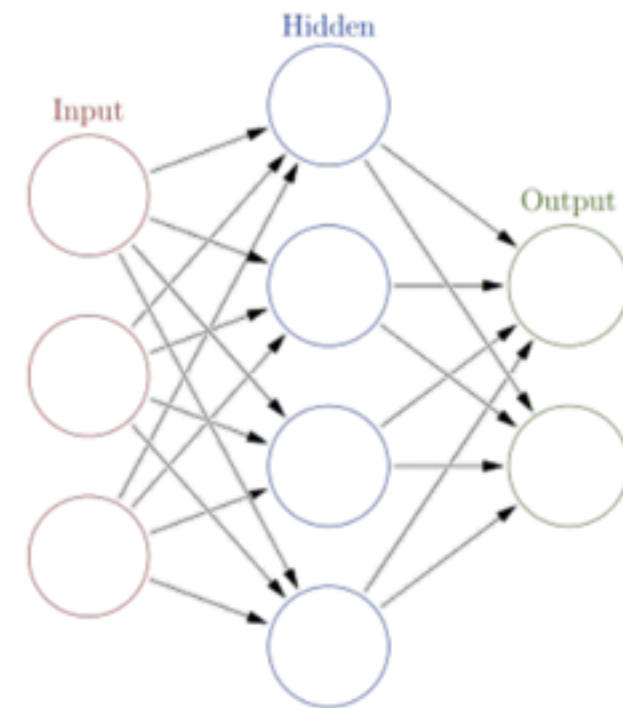
Event Topologies

- Low Z detector materials lead to long tracks and well developed showers
- Key challenges:
  - » Discriminating between muons and charged pions (muons produce longer tracks and less interaction with nuclei)
  - » Discrimination between electron and photons (photons can travel a short distance before showering)



Traditional Electron Neutrino Selectors

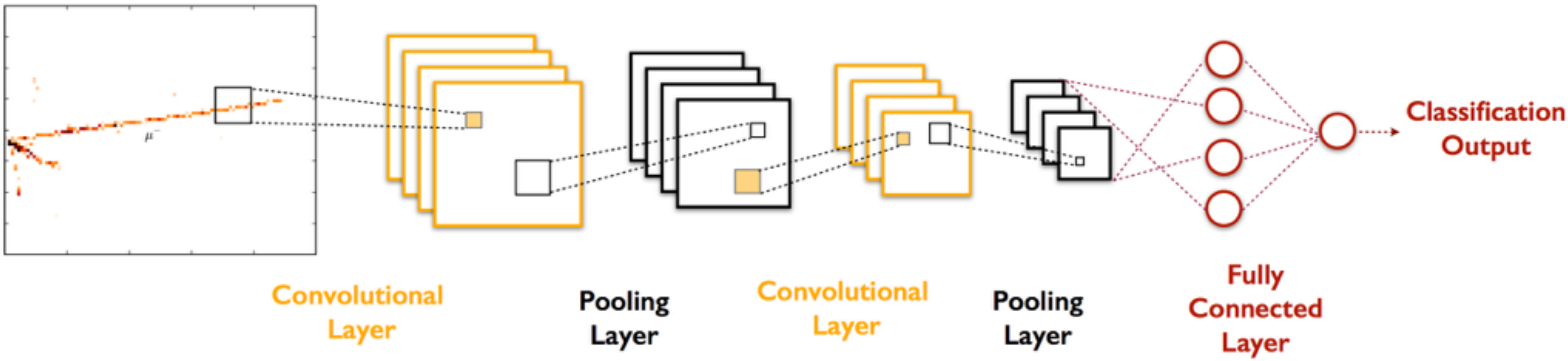
- Likelihood based ID:
  - » Calculates transverse and longitudinal dE/dx likelihoods for various particle hypothesis
  - » There, plus topological features, are fed into a standard neural network
- Library Event Matching ID:
  - » Finds best matches to a library of simulated events
  - » Properties of the best matches are fed into a decision tree



– This techniques are only as good as our ability to think up features with good separation power and our ability to construct them robustly

– Number of hidden layers is limited due to large number of weights produced by fully connected layers

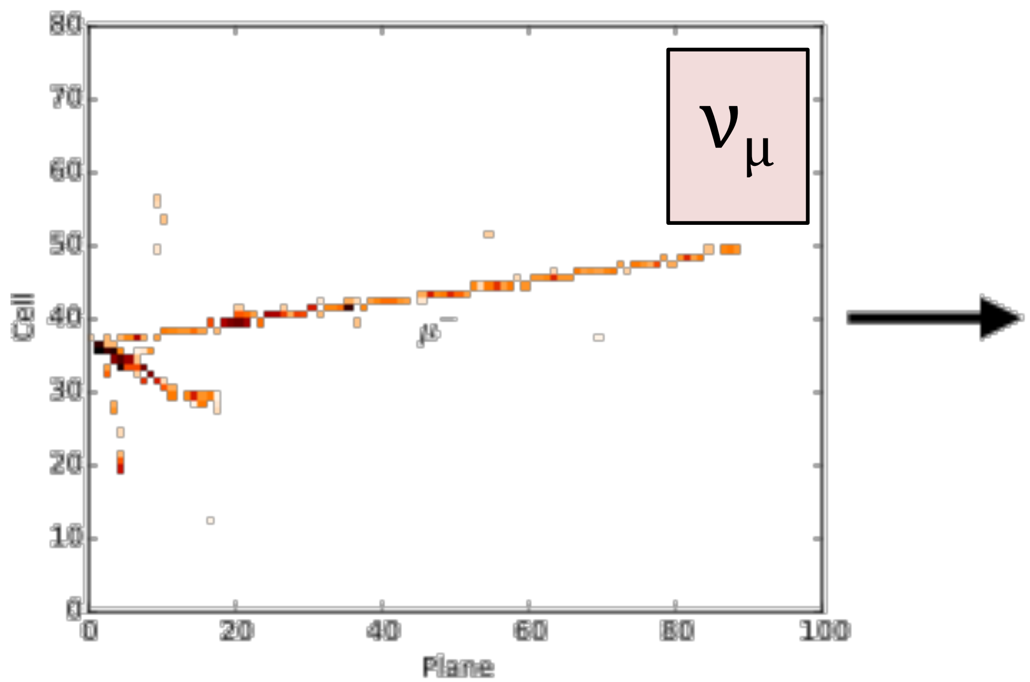
Convolutional Neural Networks



- **Convolutional Layers:** kernels are used to extract features and create feature maps
- **Pooling Layers:** downsample feature maps
- **Fully Connected Layers:** multi-classification output

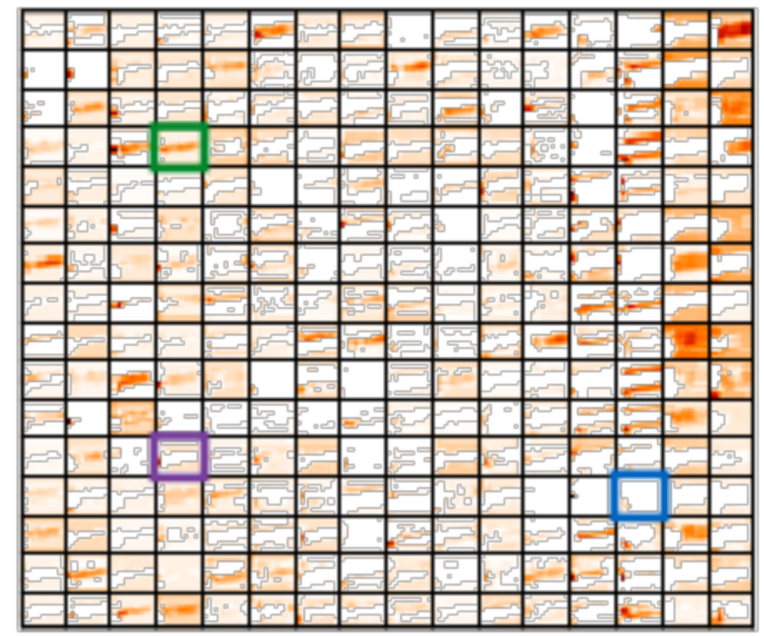
A Muon Neutrino Event

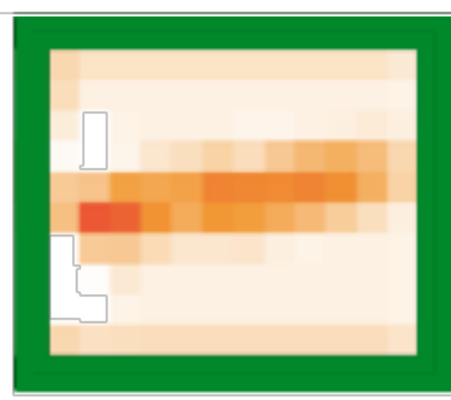
Input Image



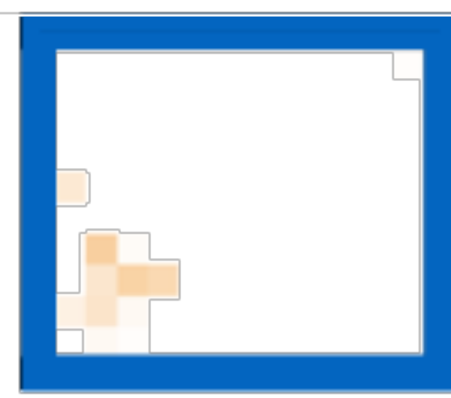
256 Feature Maps

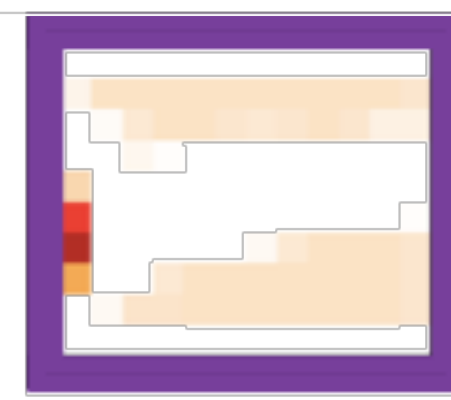
Learned variations on the original image





Responding to *μ tracks*.

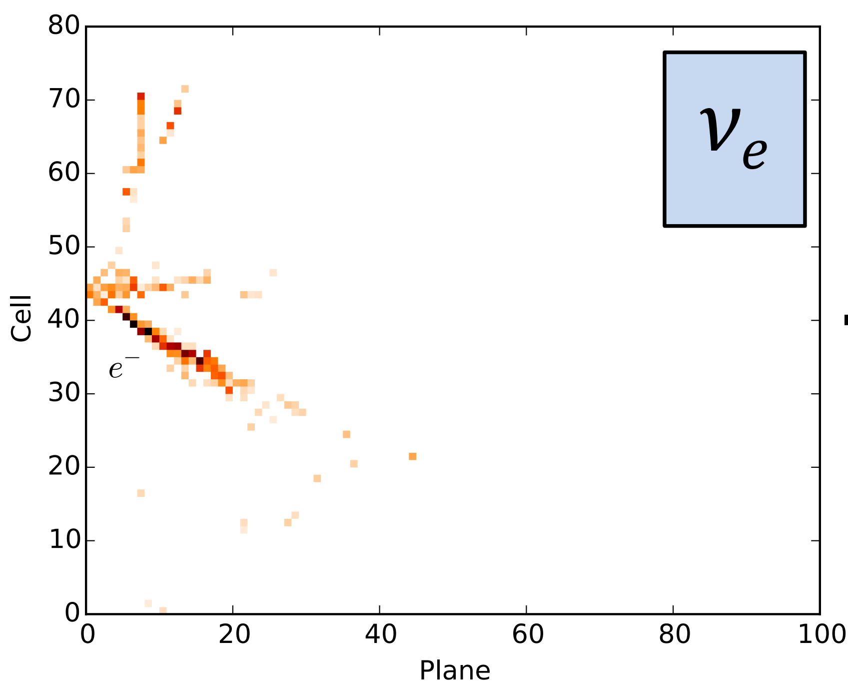




Responding to *hadronic activity*.

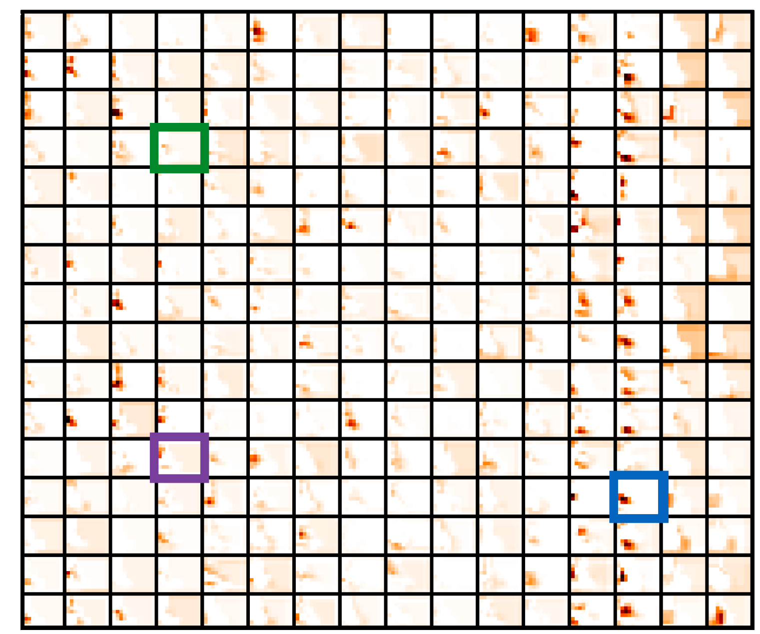
An Electron Neutrino Event

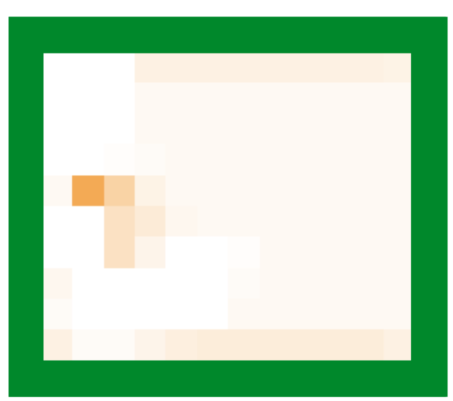
Input Image

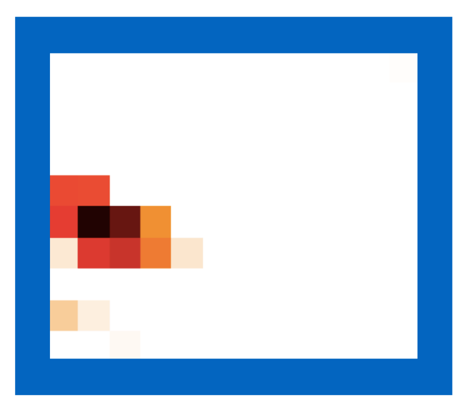


256 Feature Maps

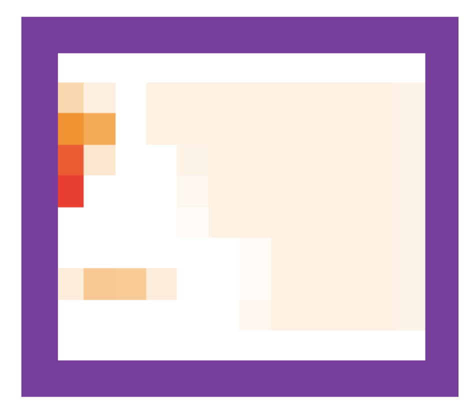
Learned variations on the original image







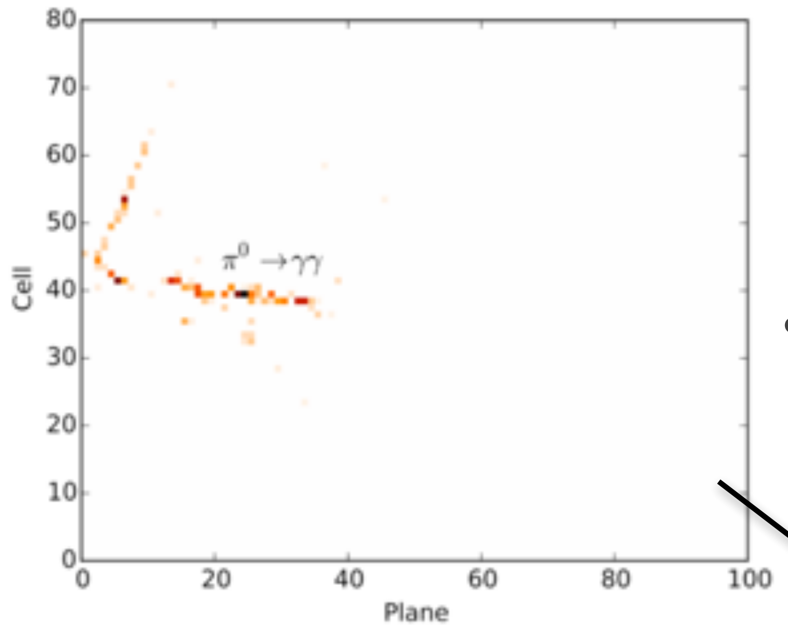
Responding to *e showers*.

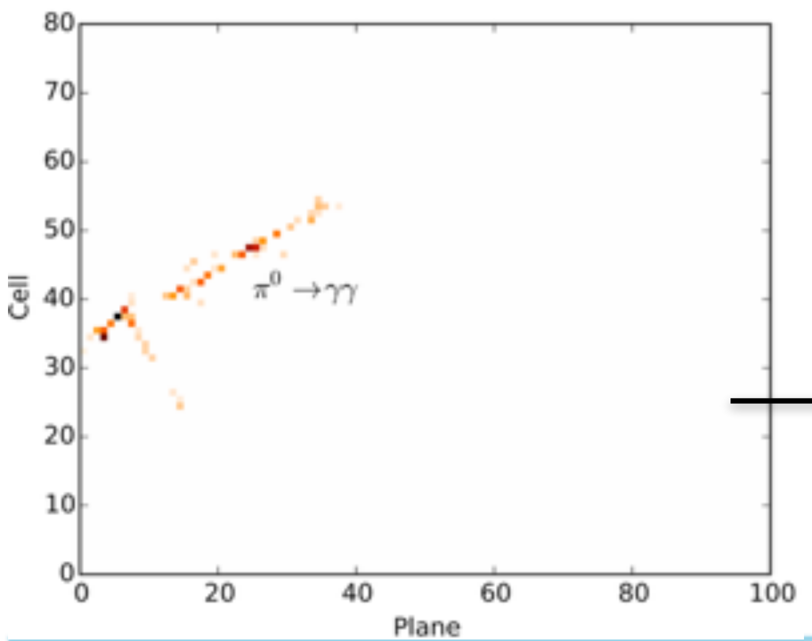


Responding to *hadronic activity*.

Convolutional Visual Network (CVN)

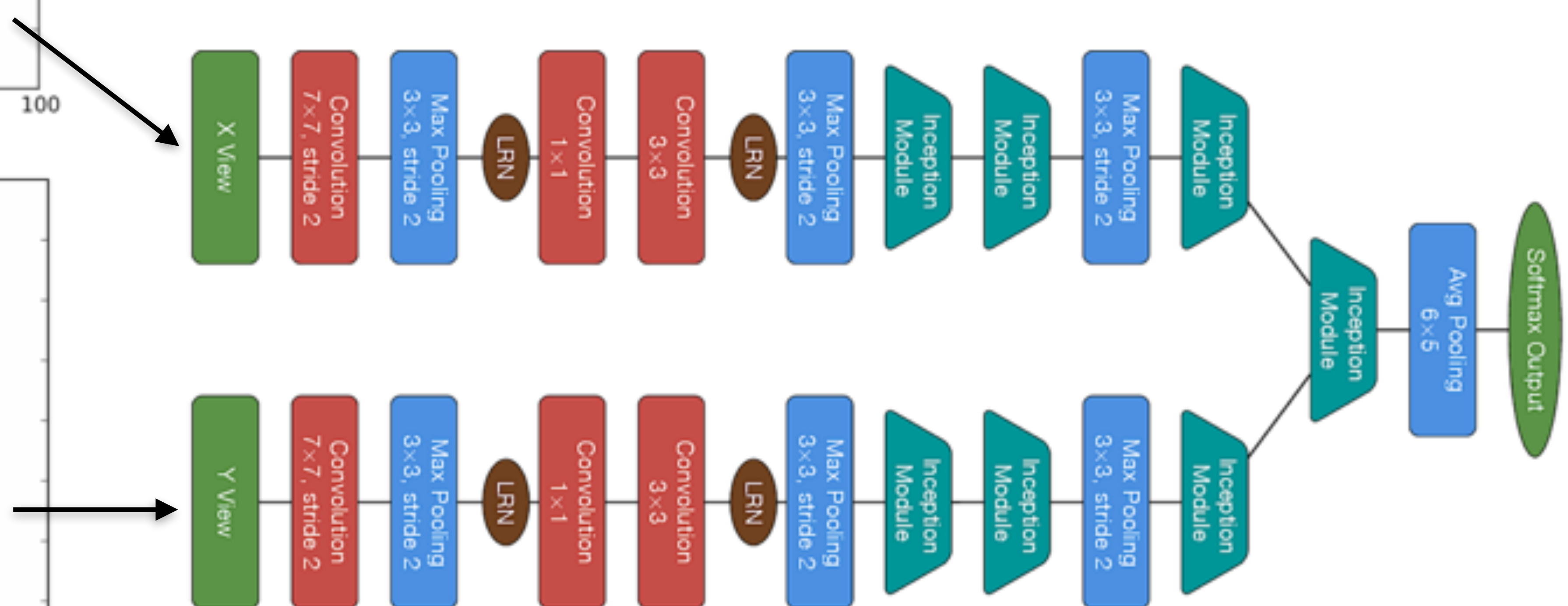
A. Aurisano and A. Radovic and D. Rocco et. al, JINST 11 P09001 (2016)





- » Create a bi-columnar networks with shared weights
- » Split views early to extract parallel features
- » Merge together at the end before going through fully connected layers
- » Ends with a feed forward neural network to create multi-classification

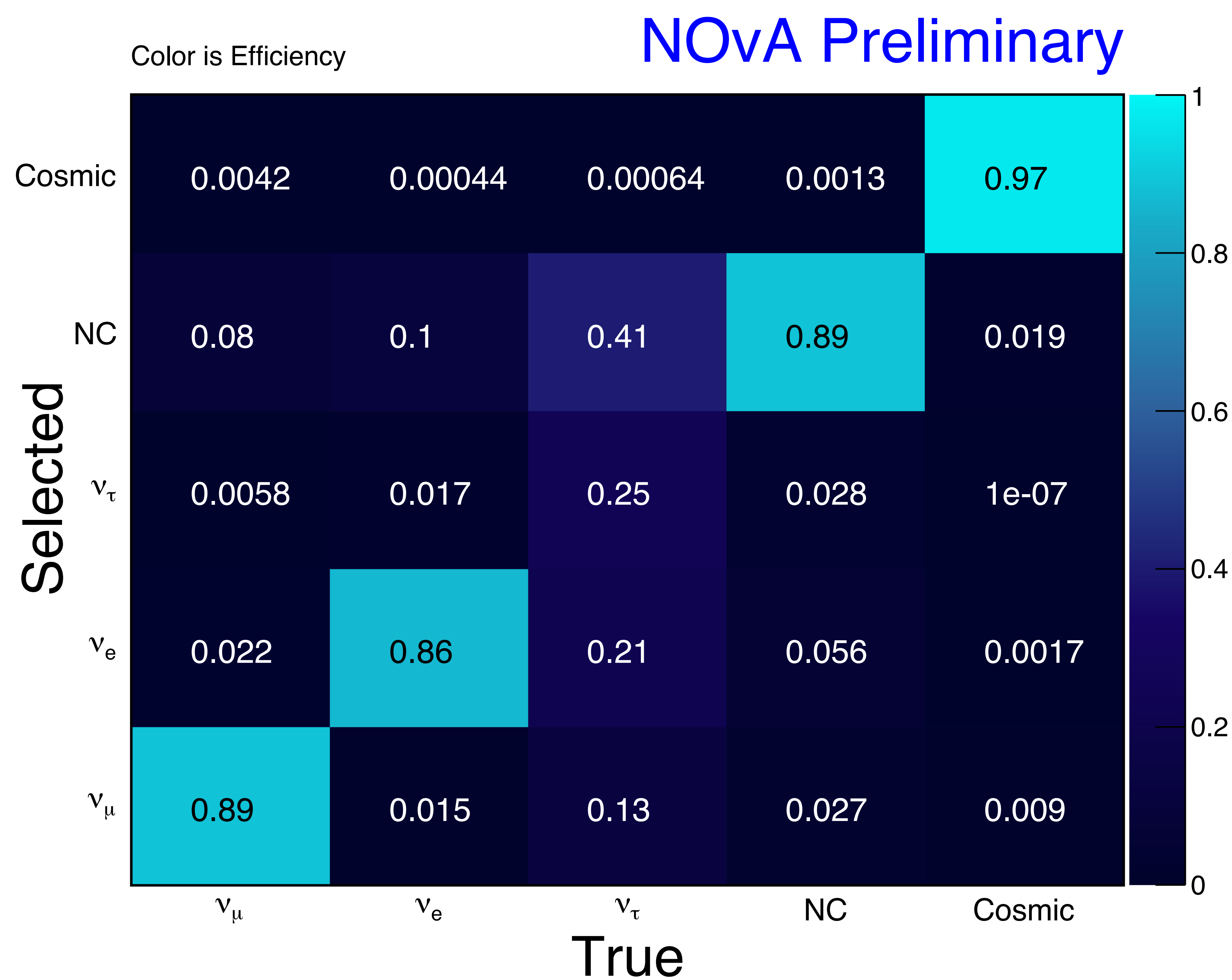
- ***Trained on 4.5+ million Monte Carlo beam events combined with cosmic ray data***



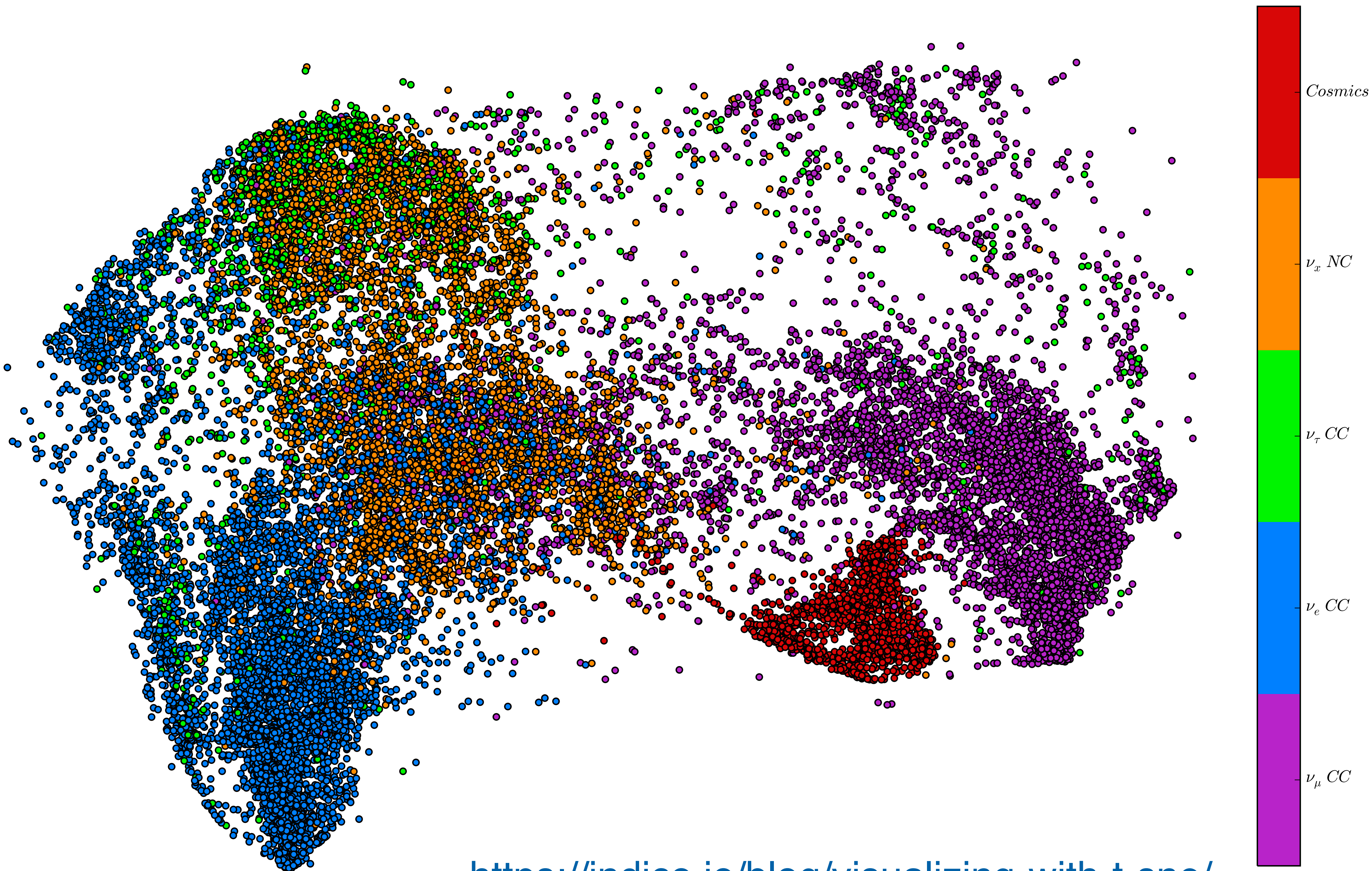
Going Deeper with Convolutions (arXiv:1409.4842)

CVN Event Classification Matrix

- NOvA was the first HEP experiment to use CNN to extract published physics results
- It improved the headline analysis performance by 30%, equivalent to an equipment savings of approximately $72 million

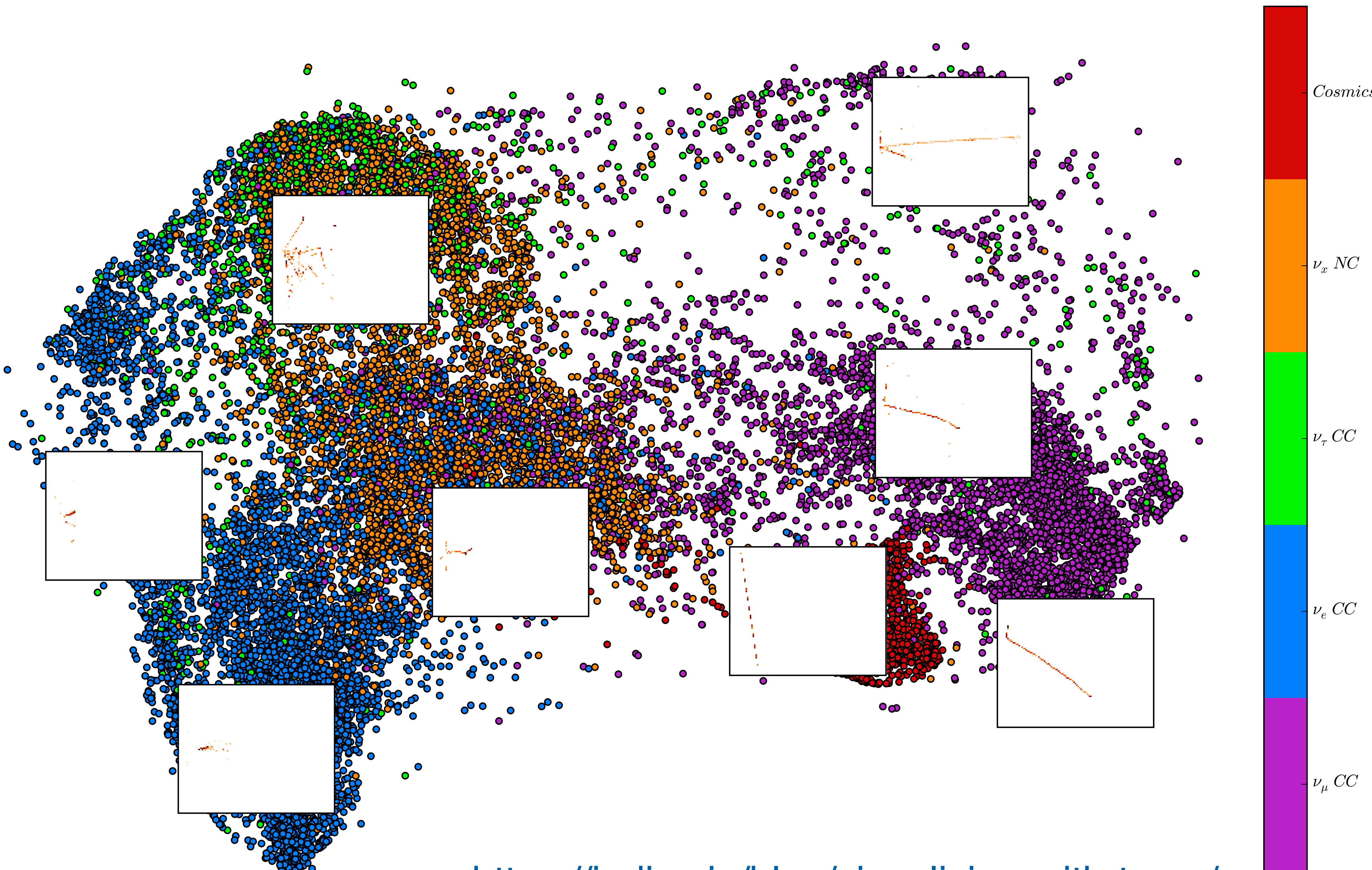


Understanding the Network: Feature Embedding with t-SNE



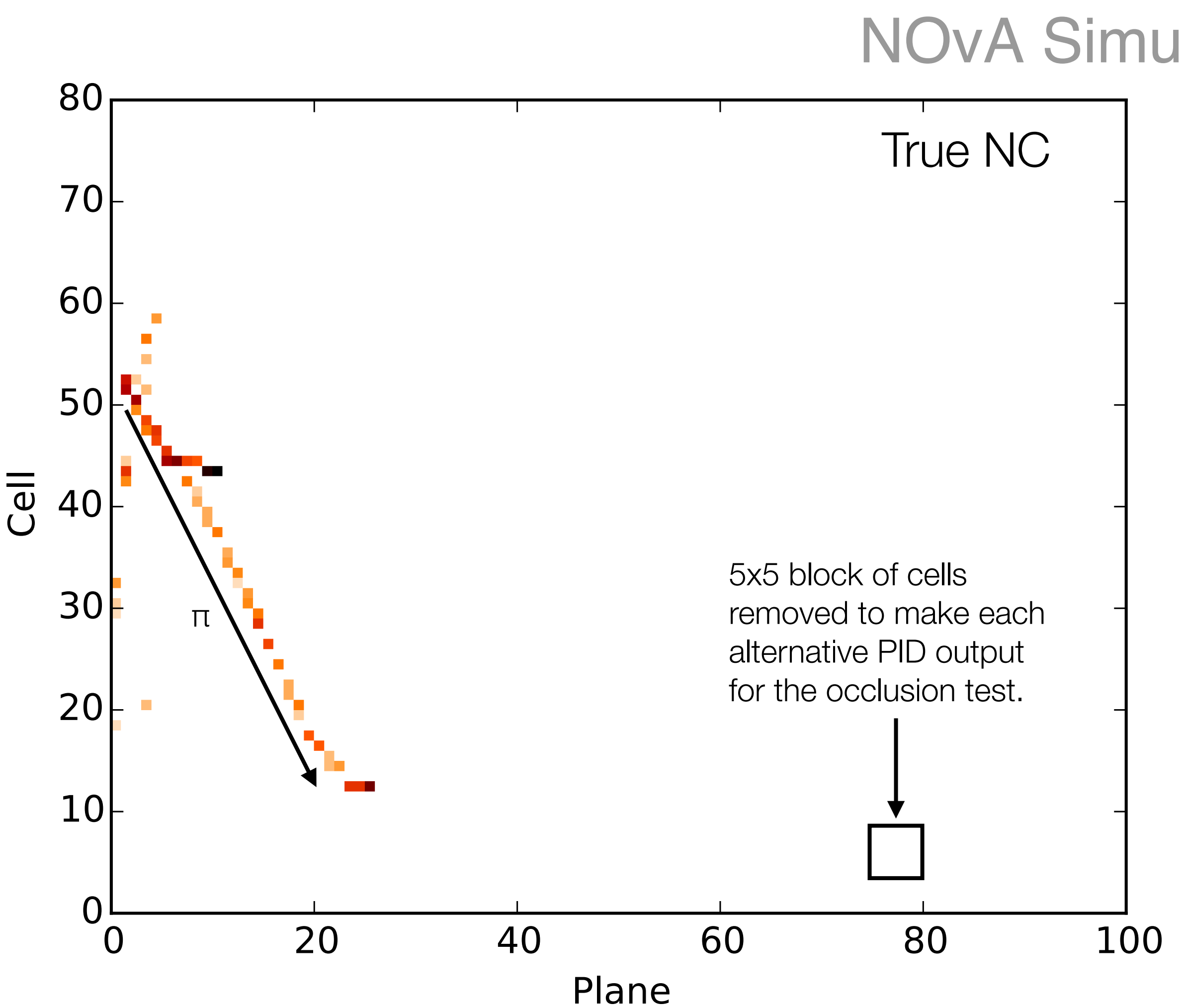
https://indico.io/blog/visualizing-with-t-sne/

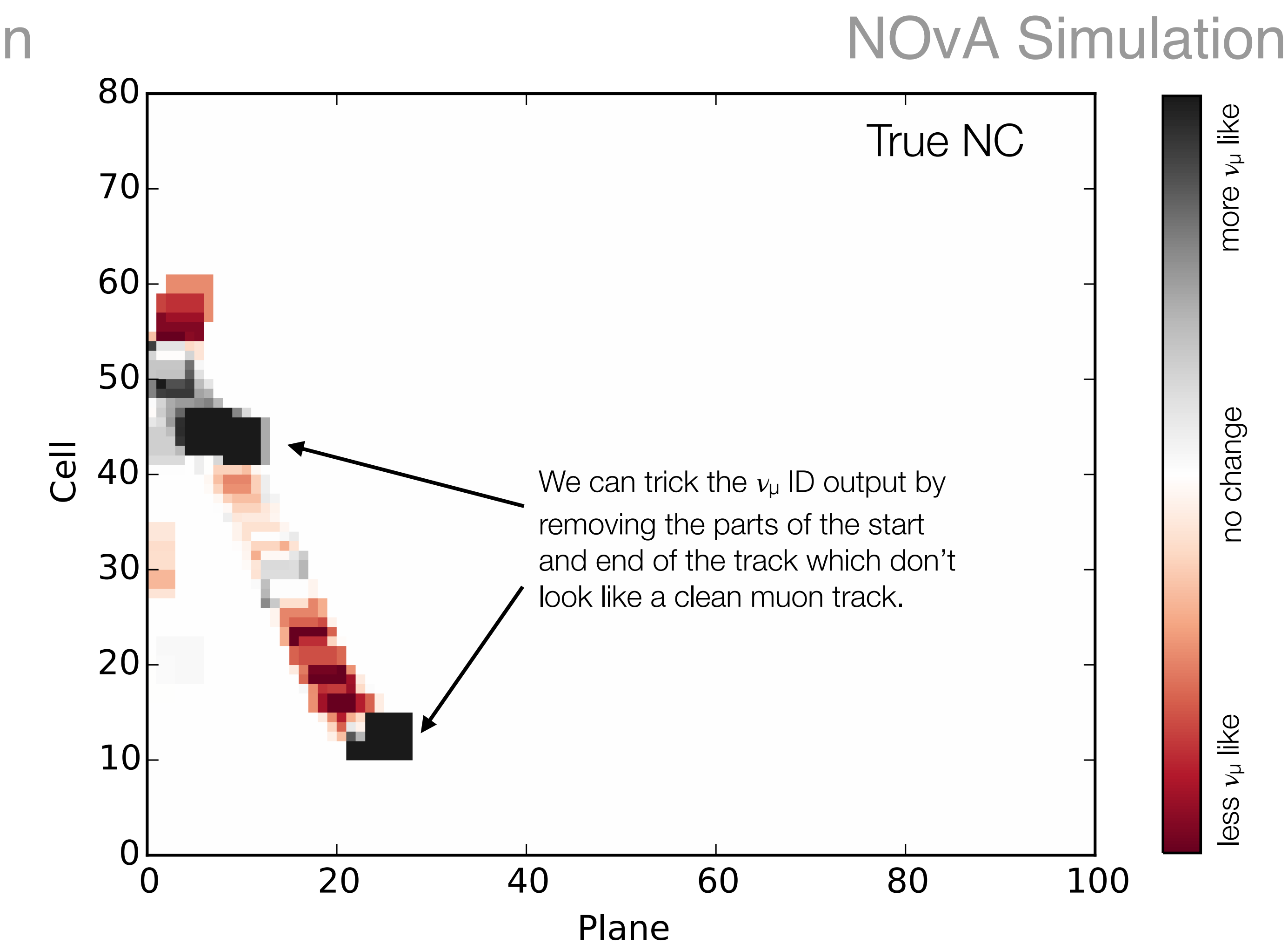
Understanding the Network: Feature Embedding with t-SNE



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

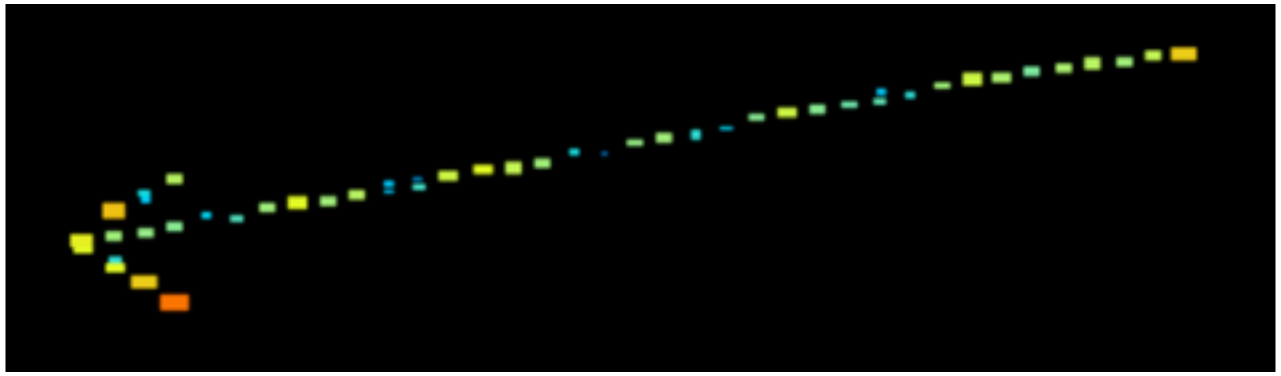
Understanding the Network: Occlusion Tests





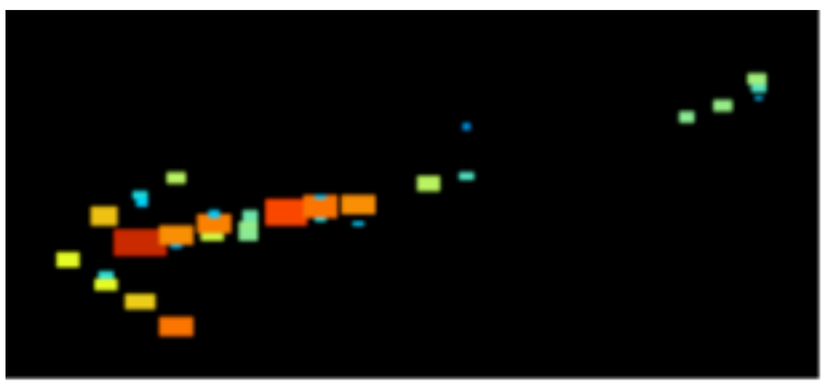
Hybrid Event Testing: CVN on Read Data

Muon Event from Data



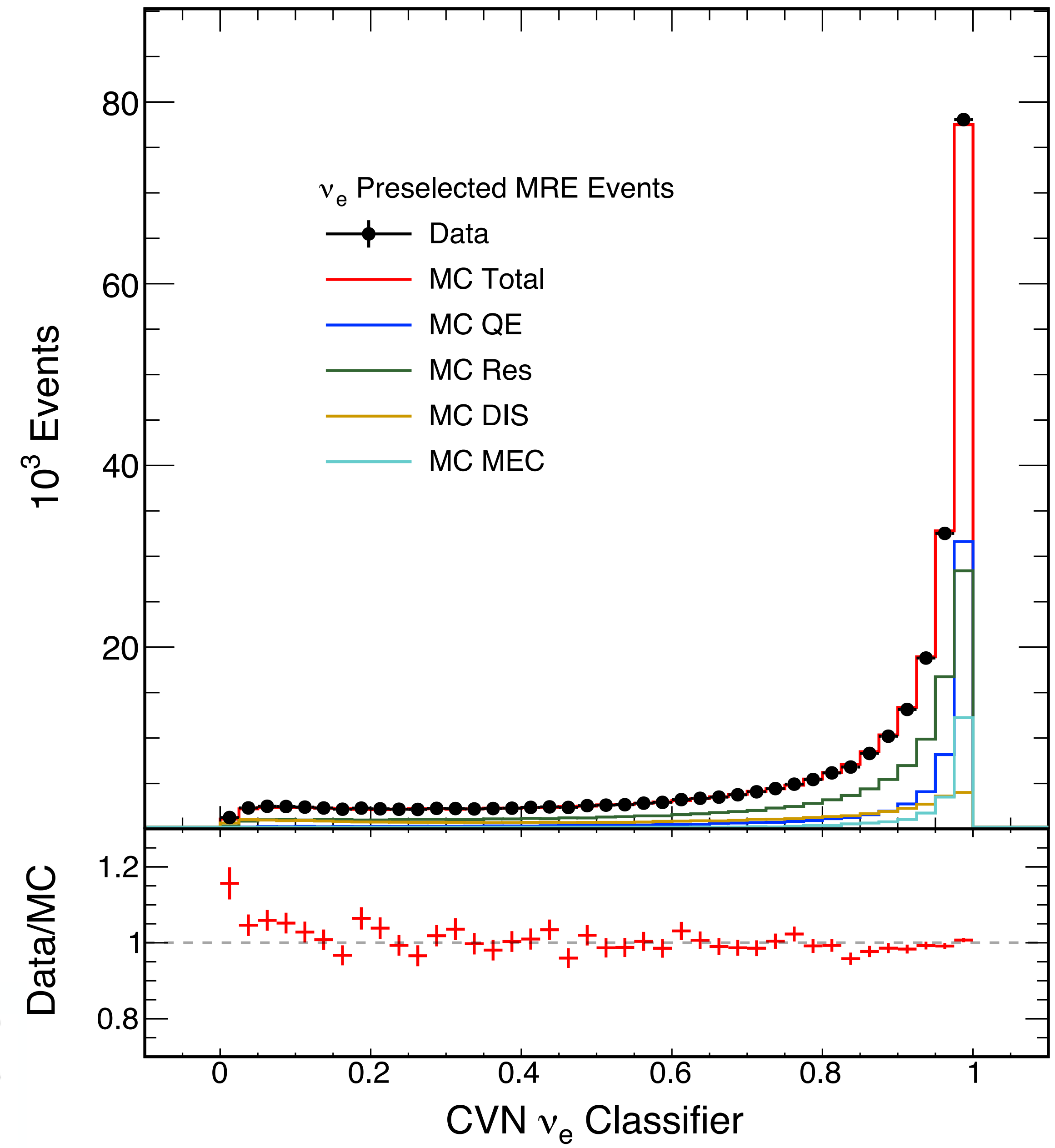
Remove muon hits

and replace with simulated electron



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	

NOvA Preliminary



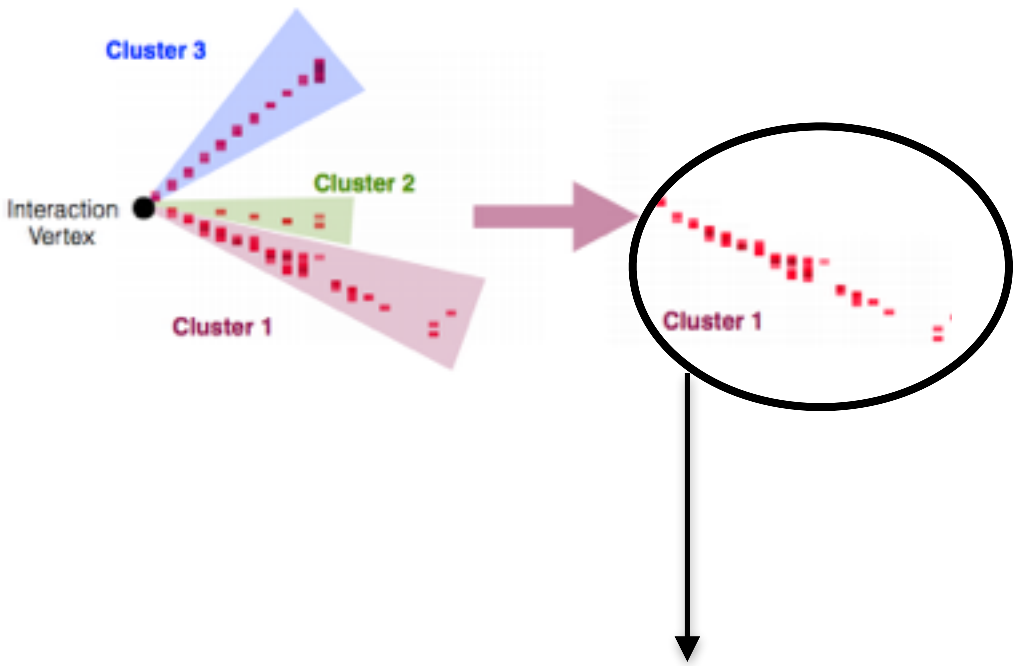
Data and MC efficiency agree to better than 1%

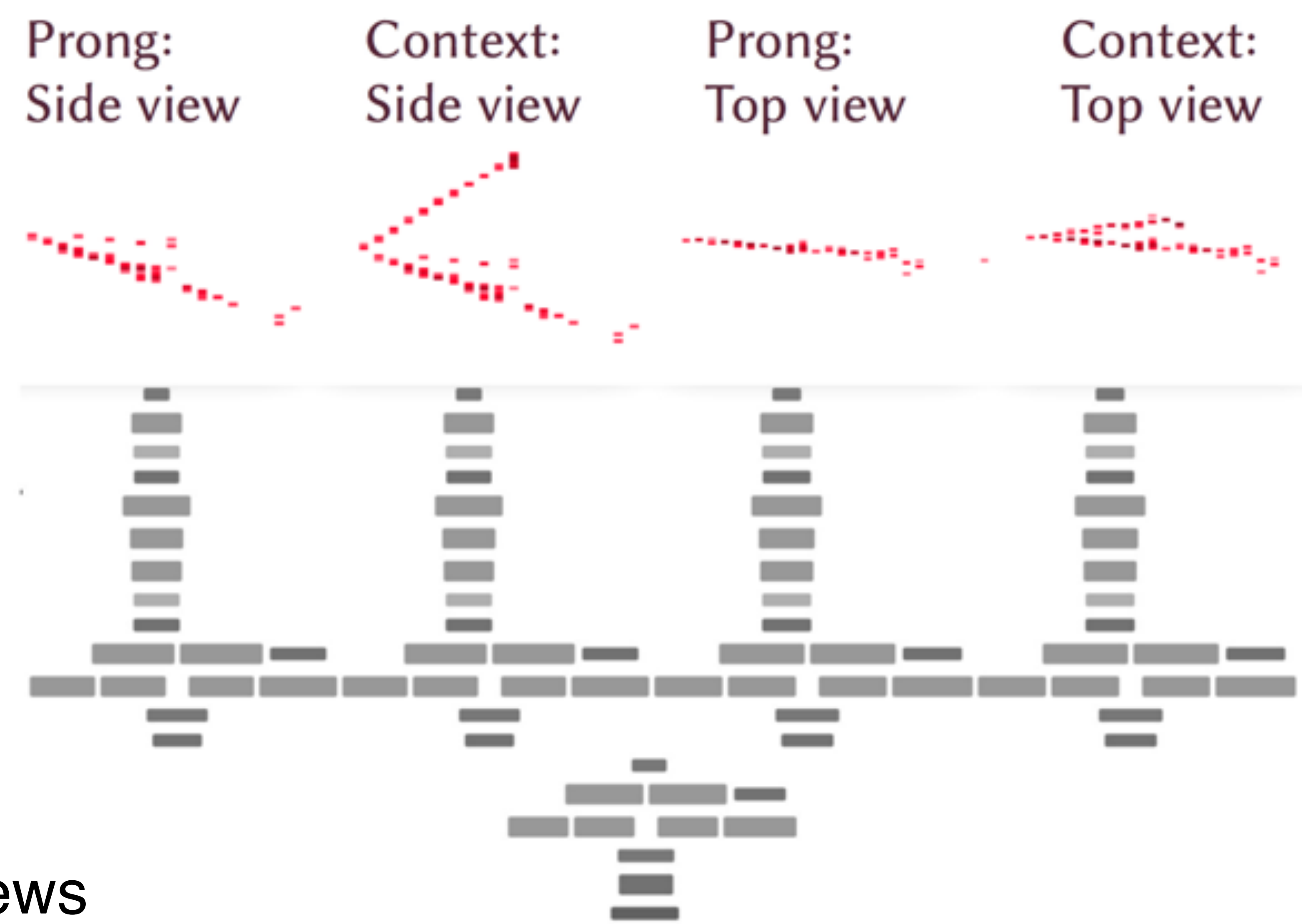
Part II
Beyond Neutrino Flavor Classification

Particle Classification Using CNNs (Prong CVN)

Particle Classification: Prong CVN

Single particles are separated using geometric reconstruction methods.

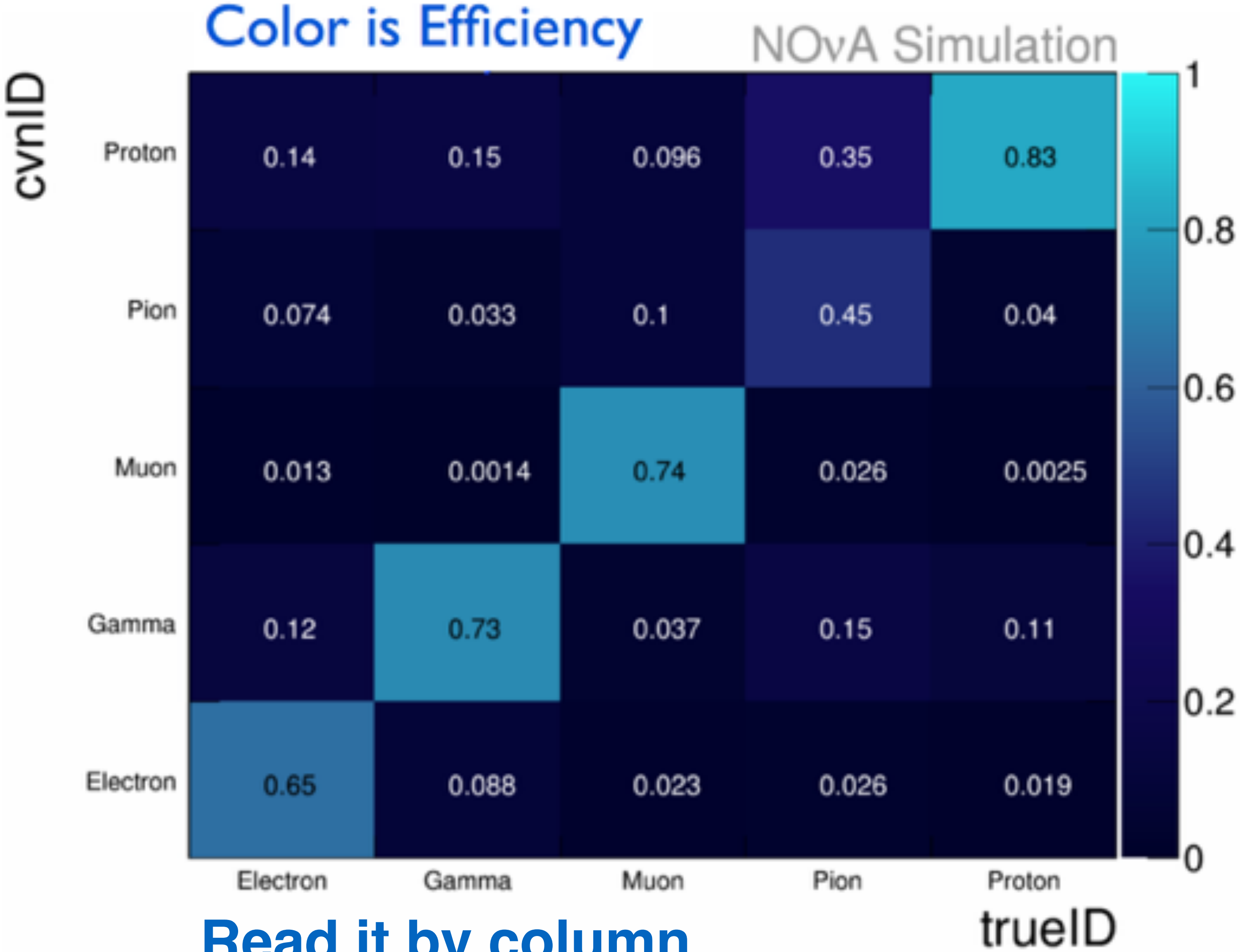




Classify particles using full event topology from both views as well as reconstructed cluster information (***4 views***)

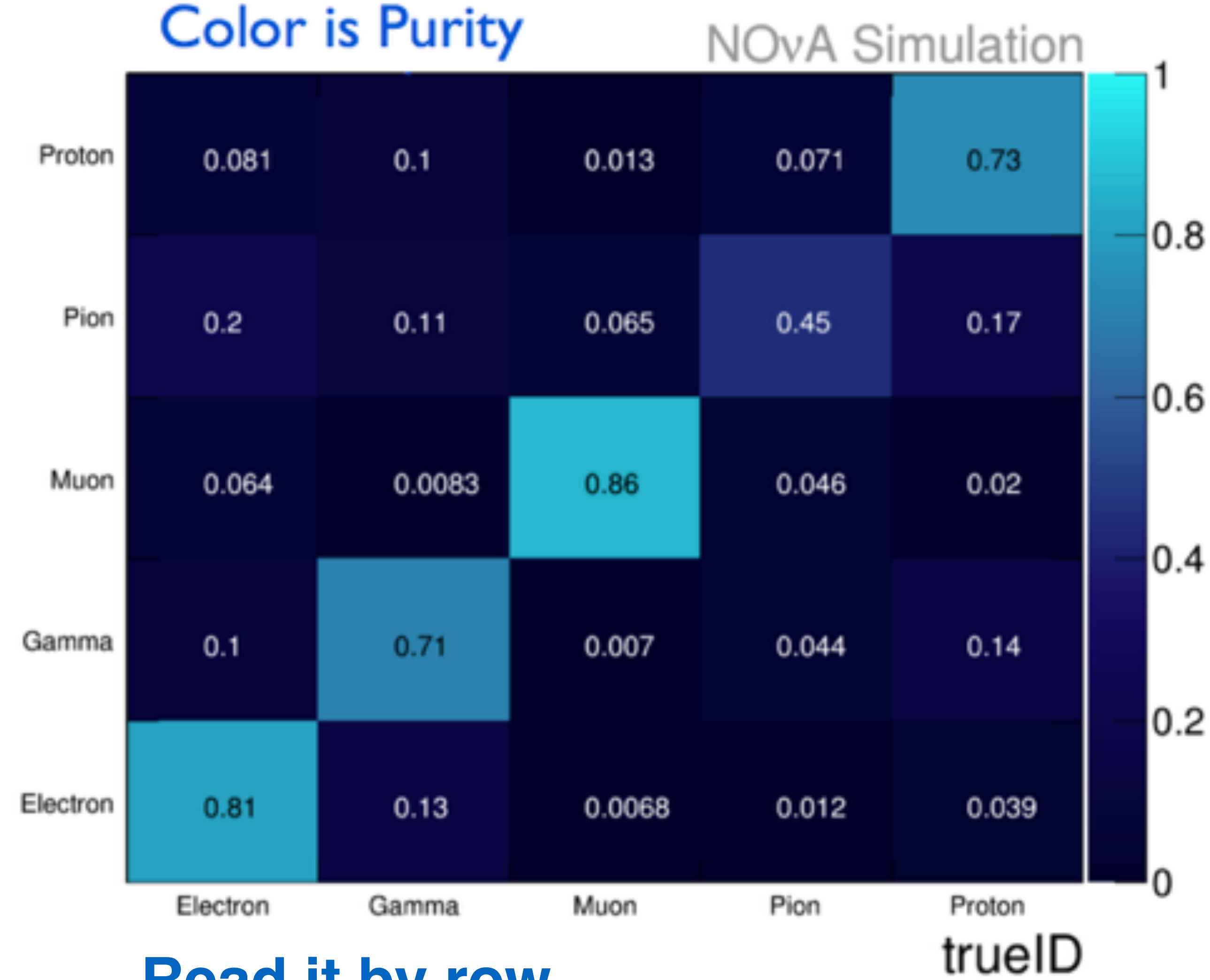
Prong CVN: Particle Classification Matrix

Color is Efficiency



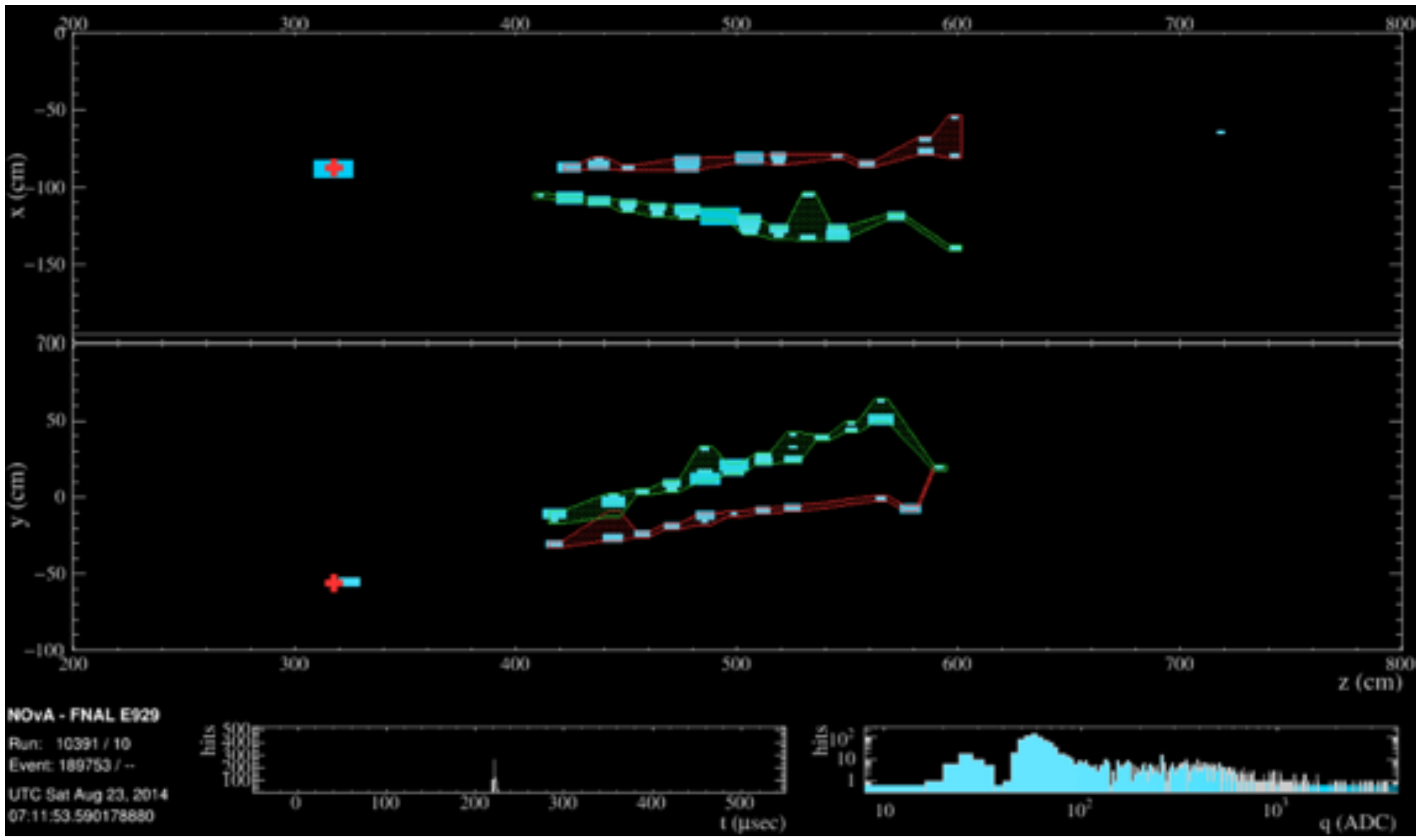
Read it by column

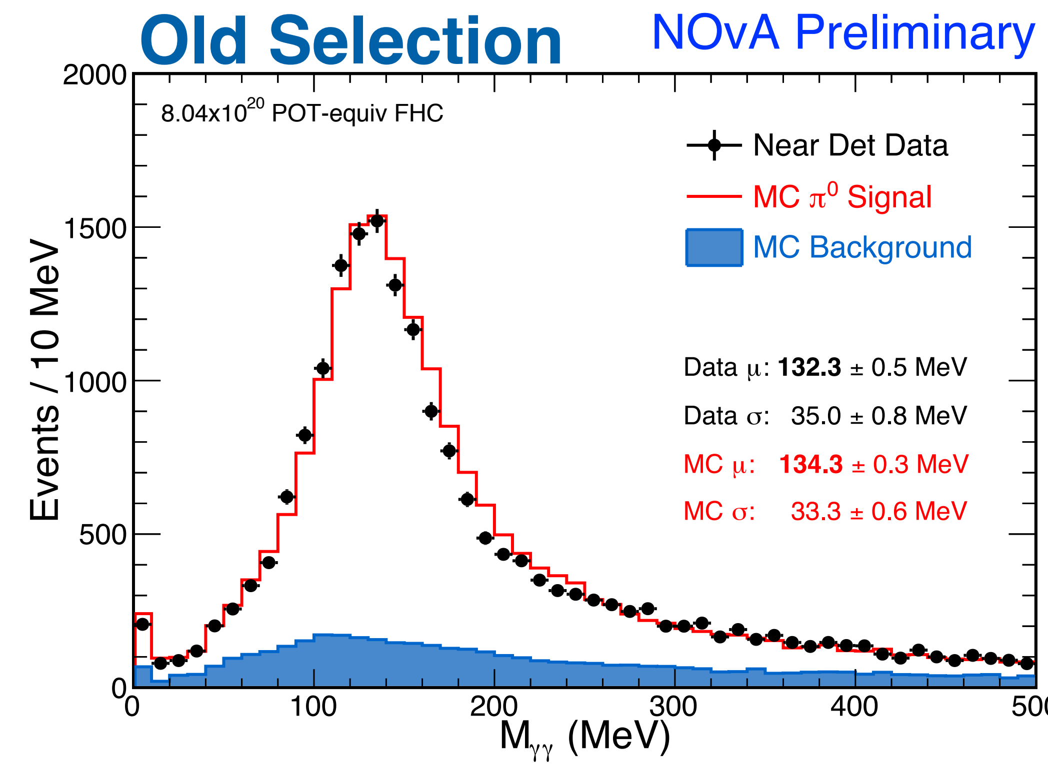
Color is Purity

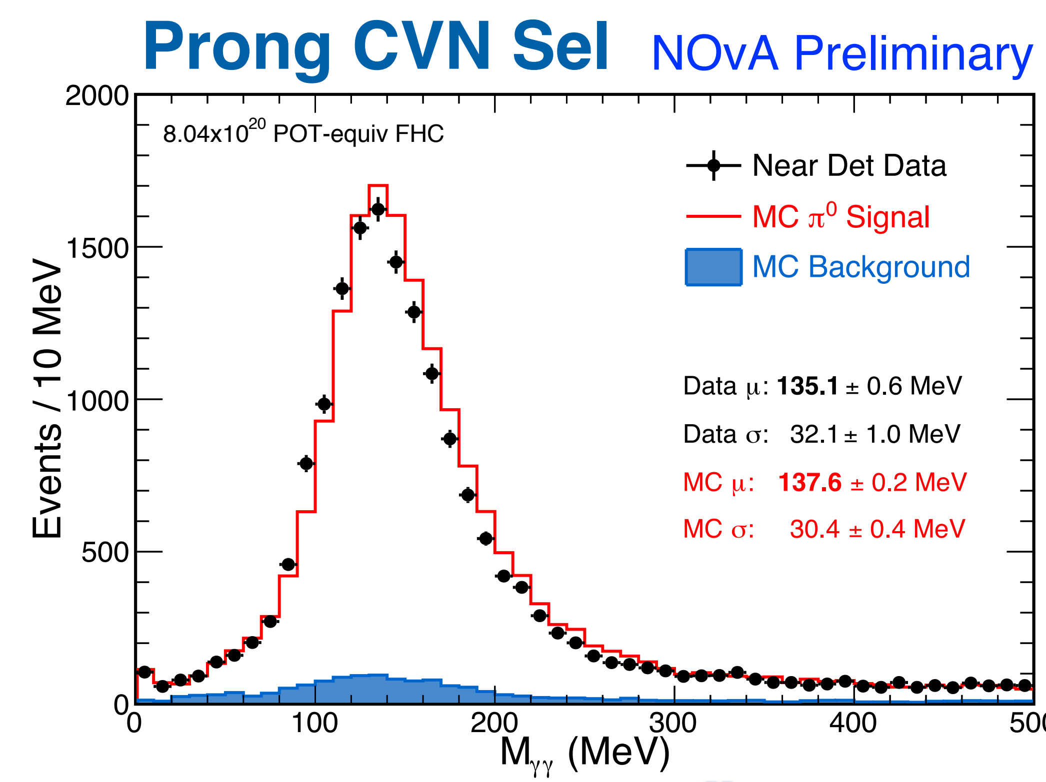


Read it by row

Use case: π^0 Mass Peak





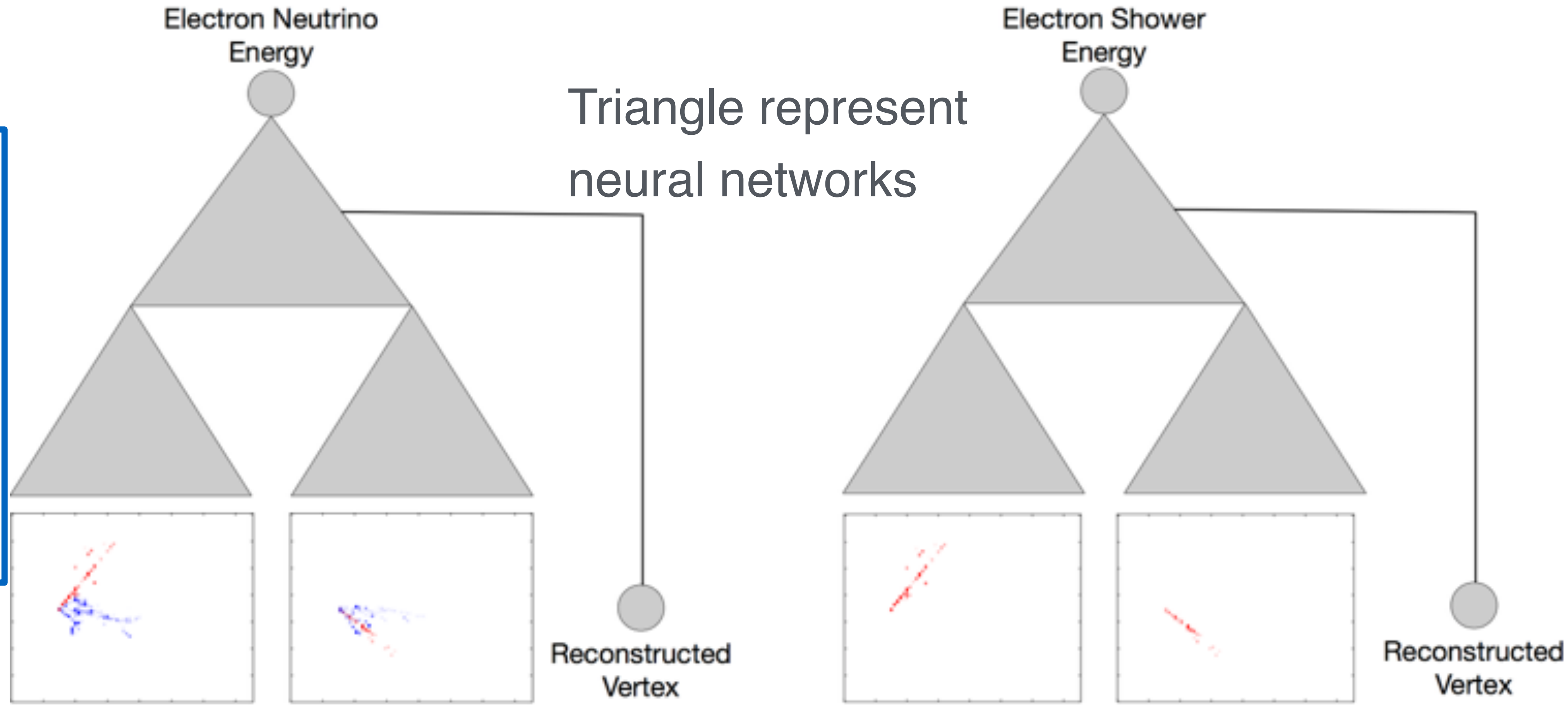


- **Data** driven method to gauge the energy response of our near detector.
- Comparing the old method using traditional reconstruction with deep learning technique lets us gain 12% increase in purity in selection. About the same efficiency.

Energy Reconstruction with Regression CNN

Energy Reconstruction with Regression CNN: Architecture

Estimate energy directly from pixel maps with minimal reconstruction



Triangle represent neural networks

- A variant of the CVN architecture
- To consider position dependence, reconstructed vertex positions in the two views used as neural network inputs
- Linear output for continuous variables
- No regularization

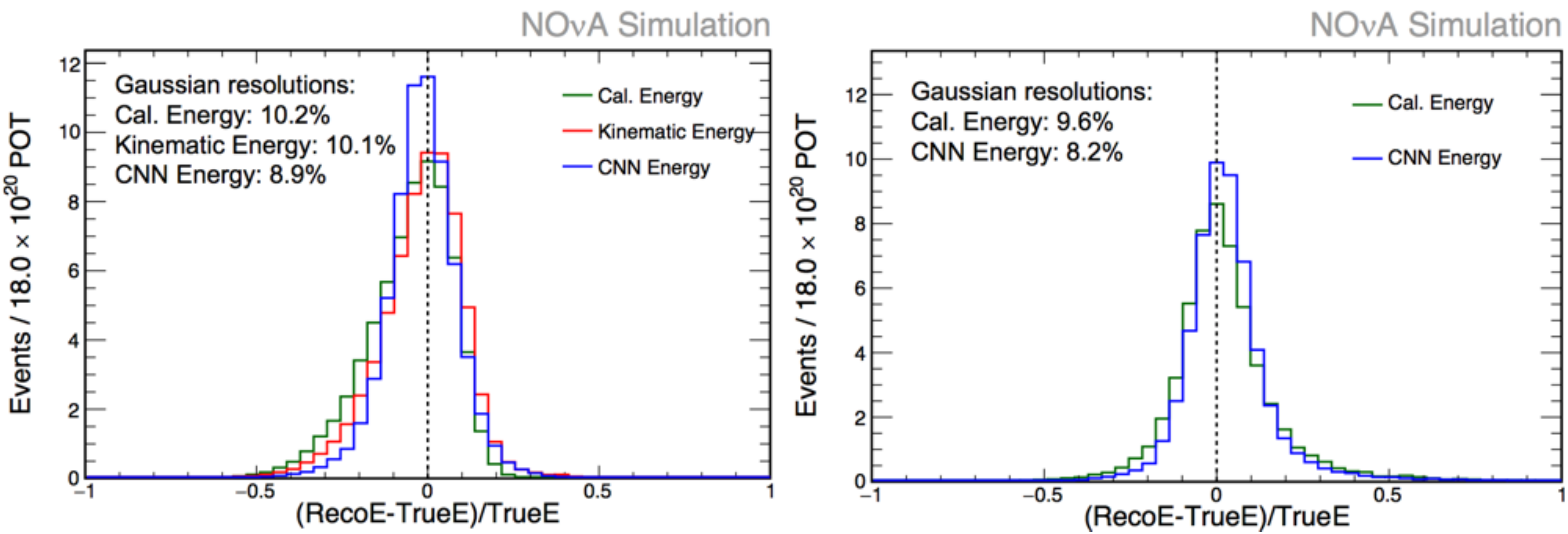
Energy Reconstruction with Regression CNN: Training

- To precisely reconstruct energy, interested in energy resolution E_{reco} - E_{true} / E_{true}, so define loss function as:

$$L(\mathbf{W}, \{\mathbf{x}_i, y_i\}_{i=1}^{n}) = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{f_{\mathbf{W}}(\mathbf{x}_i) - y_i}{y_i} \right|$$

- Use absolute error instead of mean squared error to prevent large impacts from outliers
- New hyperparameter optimization was necessary

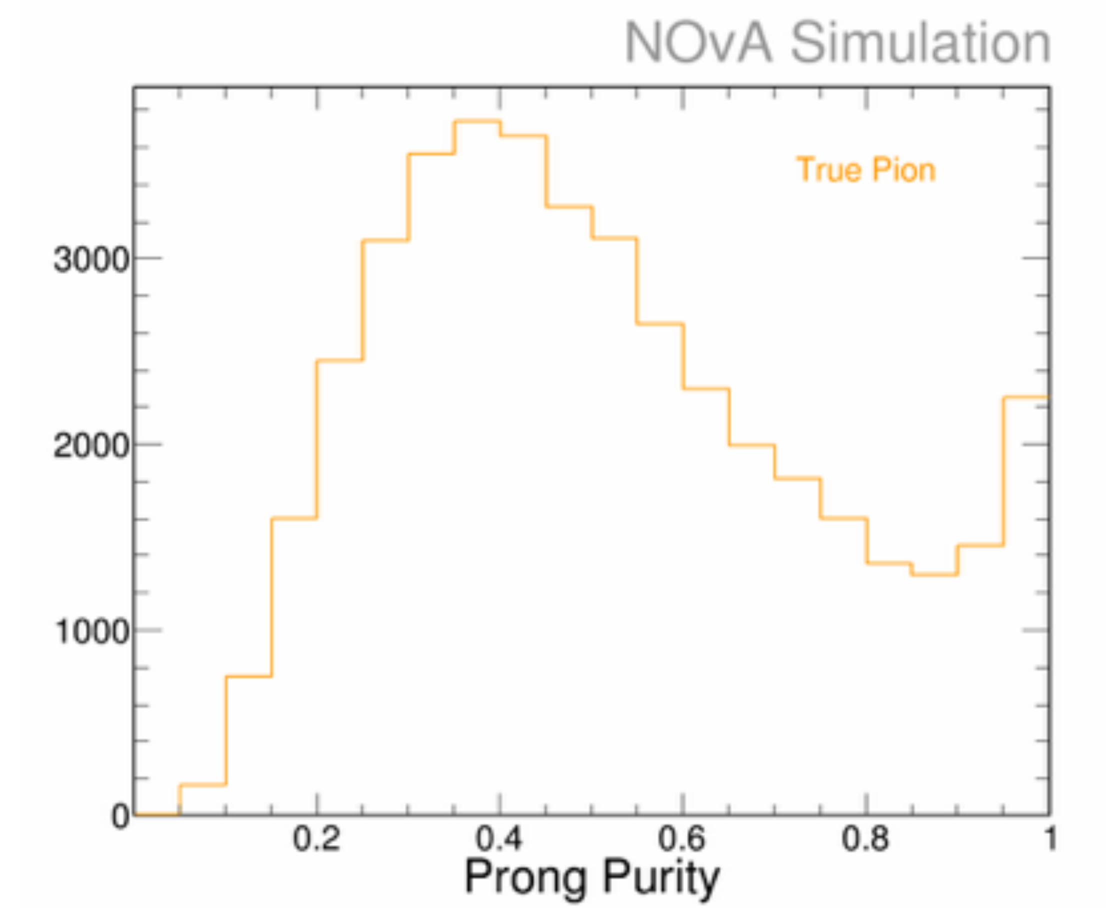
Energy Reconstruction with Regression CNN: Results

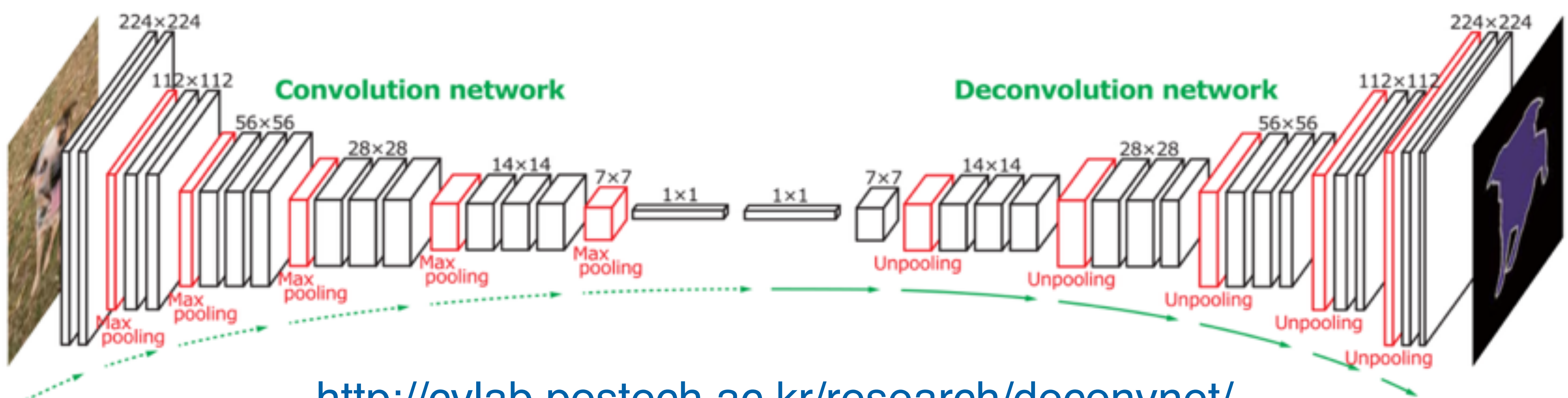


- Calorimetric energy: Sum of calibrated energy with a scale factor
- Kinetic energy: Based on NOvA's ν_e analysis 2017:
  - $E(\nu_e) = A{*}E_{EM} + B{*}E_{HAD} + CE_{EM}^2 + DE_{HAD}^2$
- CNN Energy: Regression CNN energy estimator

Future Endeavors: Full Event Reconstruction

- Single particle classifier depends on the quality of clustering hits
- Hit level identification, were clustering and classifying particles at the same time: **semantic segmentation**





http://cvlab.postech.ac.kr/research/deconvnet/

Fermilab is using machine learning across all programs
Few examples…

Many More Fermilab Machine Intelligence Applications

- ***Minerva:*** Reducing model bias in a deep learning classifier using domain adversarial neural networks in the MINERvA experiment (arXiv:1808.08332)

- ***MicroBooNE:*** A Deep Neural Network for Pixel-Level Electromagnetic Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber (arXiv:1808.07269)

- ***LHC:*** Real Time AI: Fast inference of deep neural networks in FPGAs for particle physics (arXiv:1804.06913v3)

- ***Cosmology/Astronomy:*** DeepCMB: Lensing Reconstruction of the Cosmic Microwave Background with Deep Neural Networks (arXiv:1810.01483v1)

Thanks for Listening!

Special thanks to NOvA collaborators