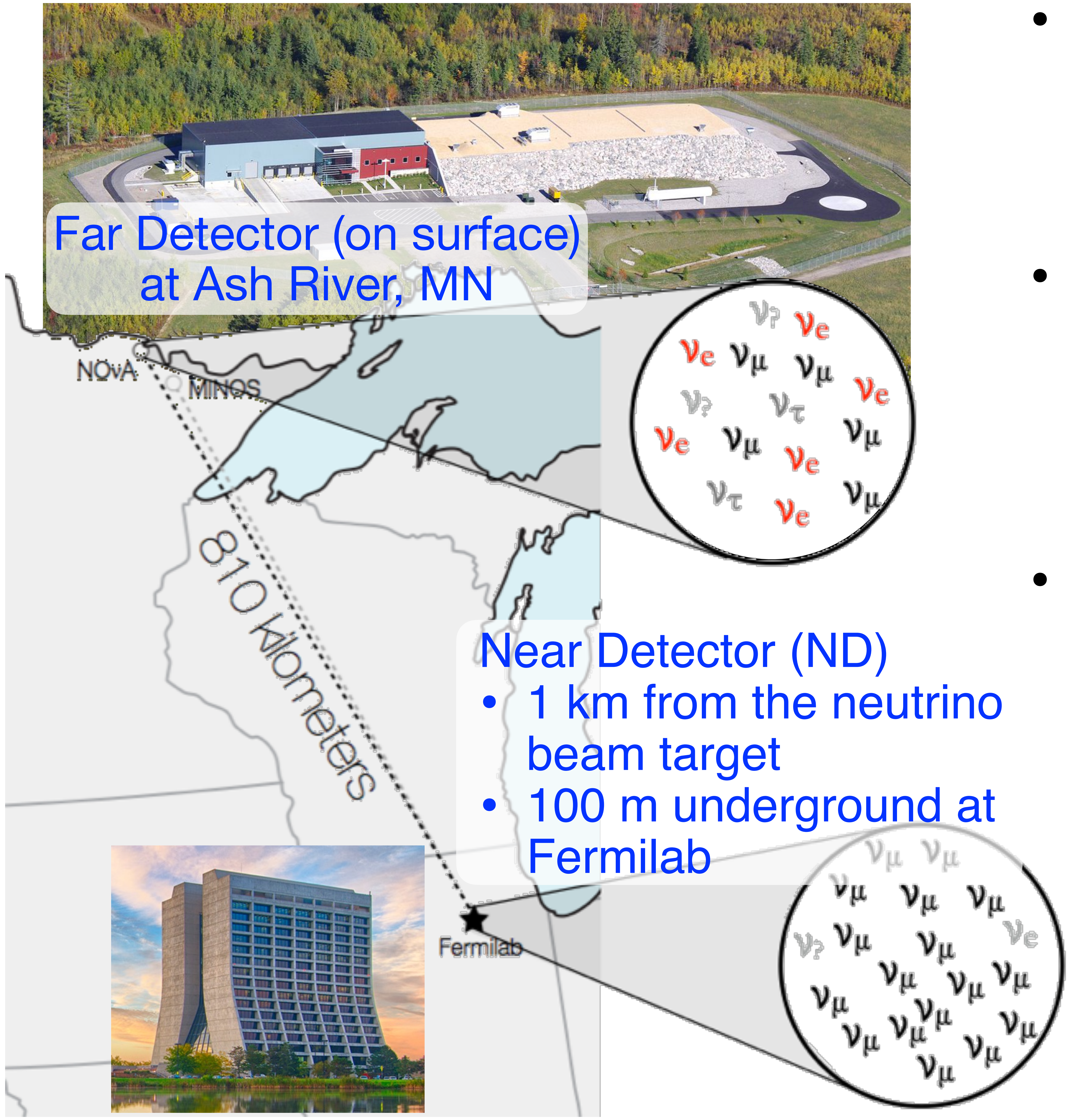
Machine Learning in NOvA

Wenjie Wu (UC Irvine), for the NOvA Collaboration

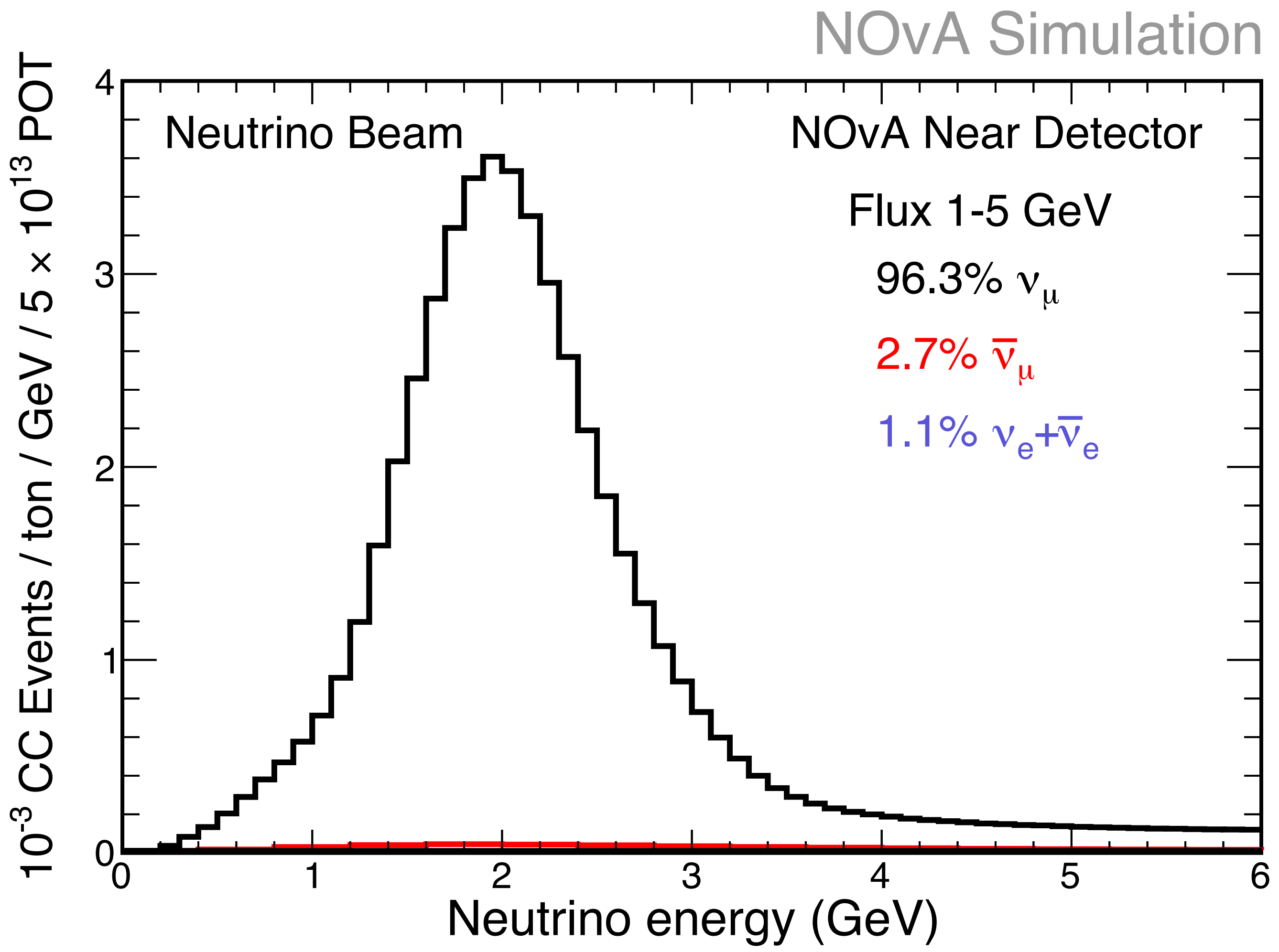
The Second Wire-Cell Reconstruction Summit

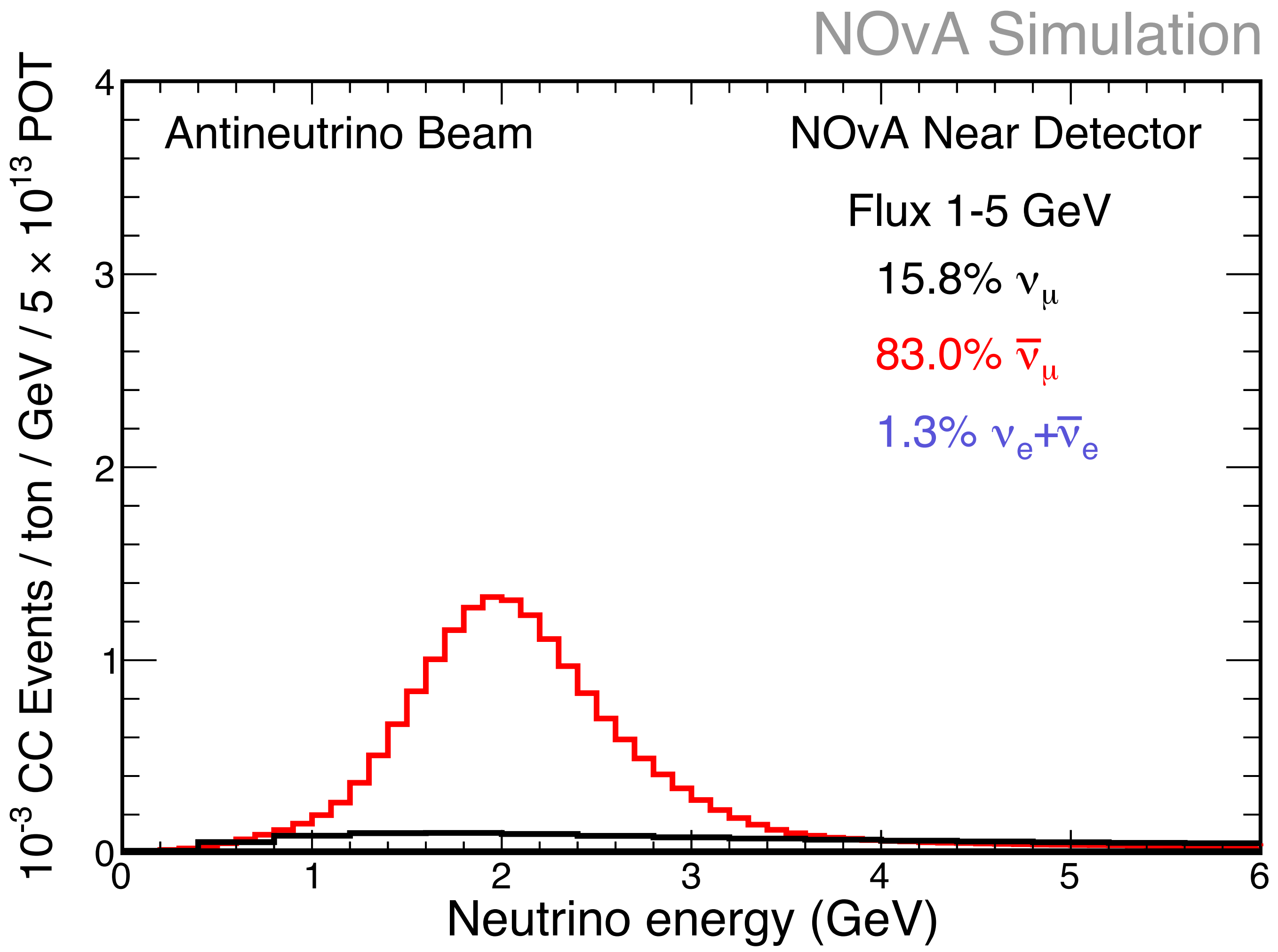
April 12th, 2024

NOvA: NuMI Off-Axis ν_e Appearance Experiment

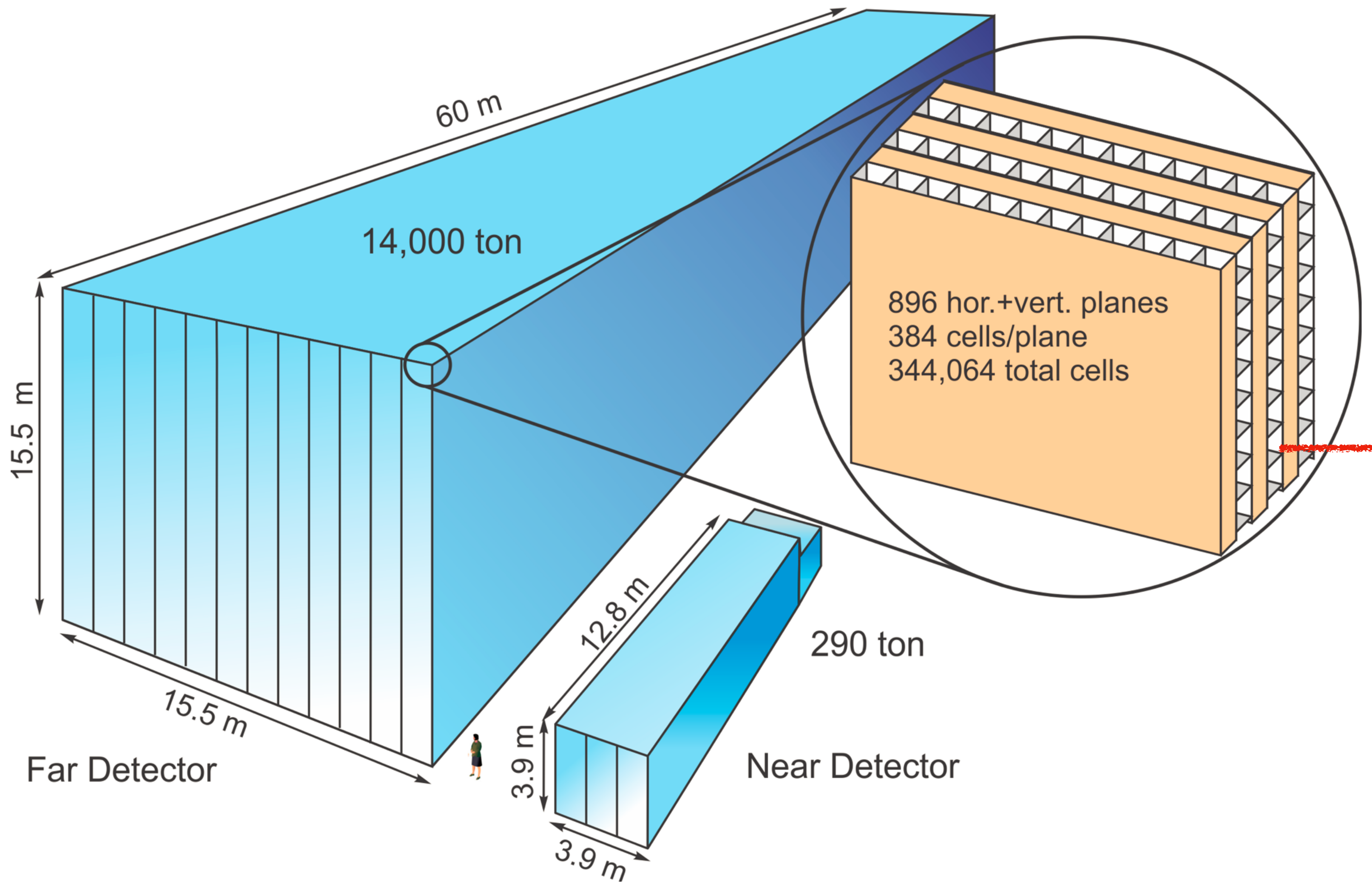


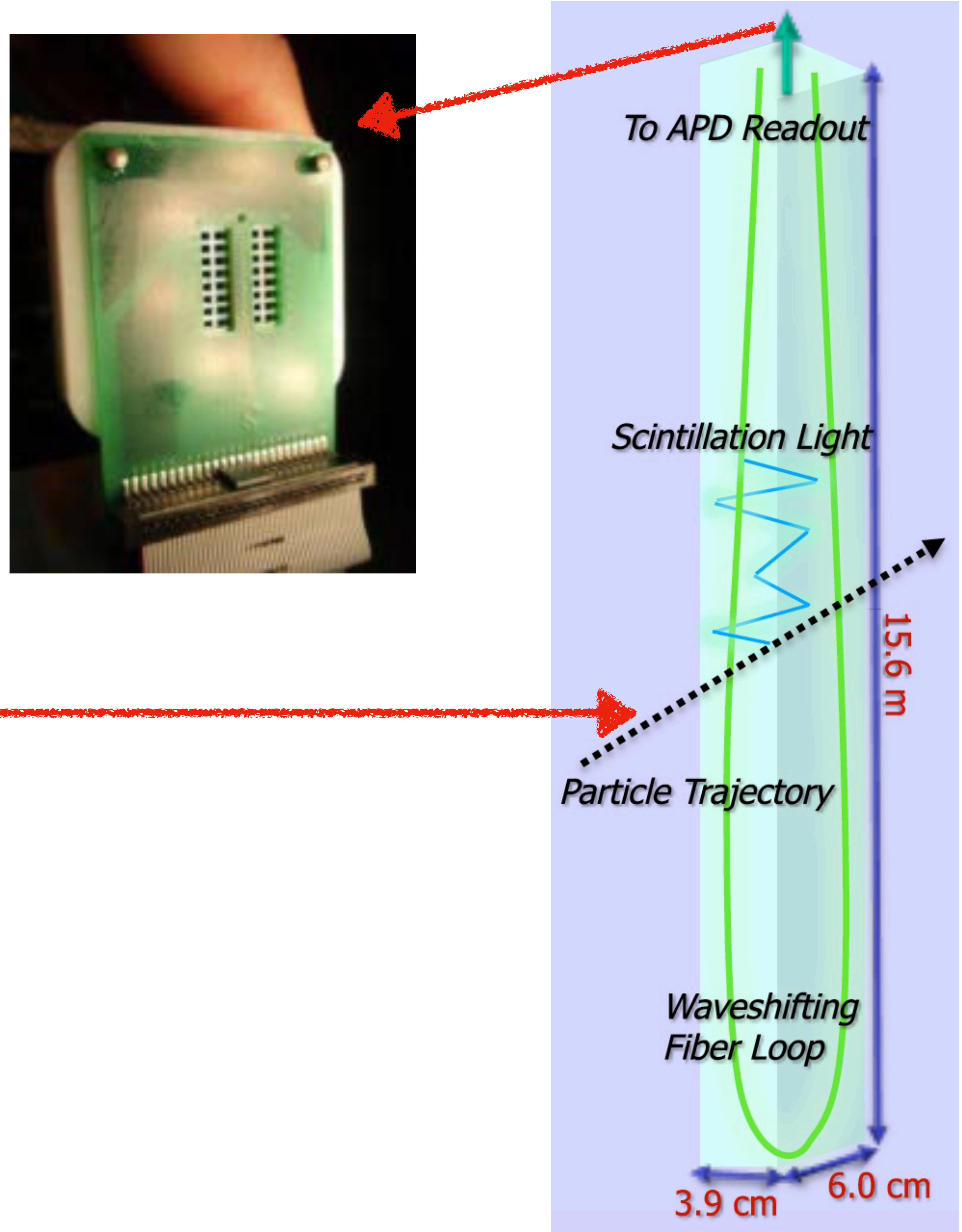
- NOvA is an accelerator-based neutrino experiment
  - Longest baseline in operation (810 km), large matter effect, sensitive to mass ordering
- Muon neutrino beam (NuMI) at Fermilab
  - Two configurations: neutrino mode and antineutrino mode
  - Power record 954 kW in 2023
- ~14 mrad off-axis, narrow-band beam around oscillation max





NOvA Detectors

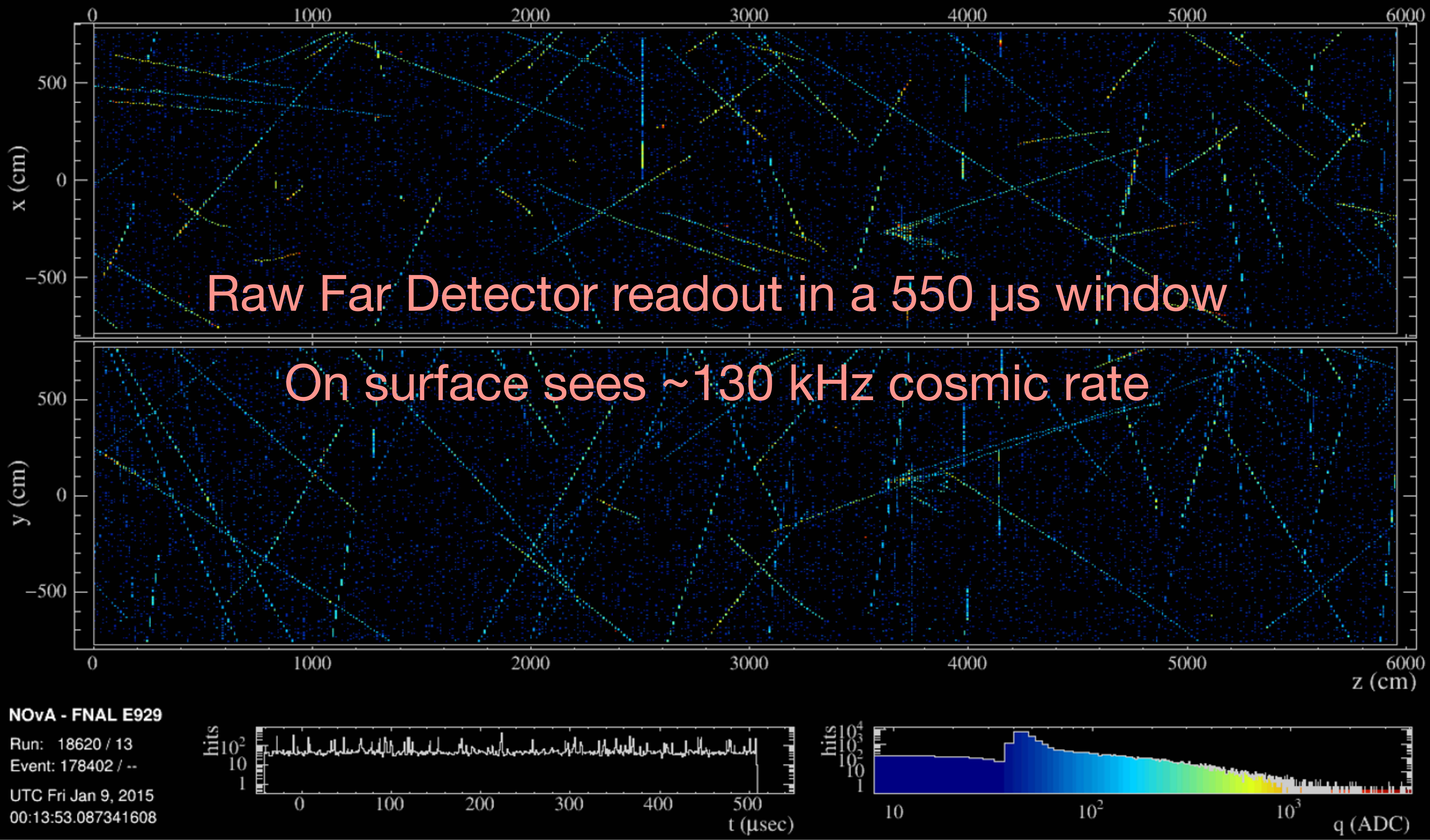


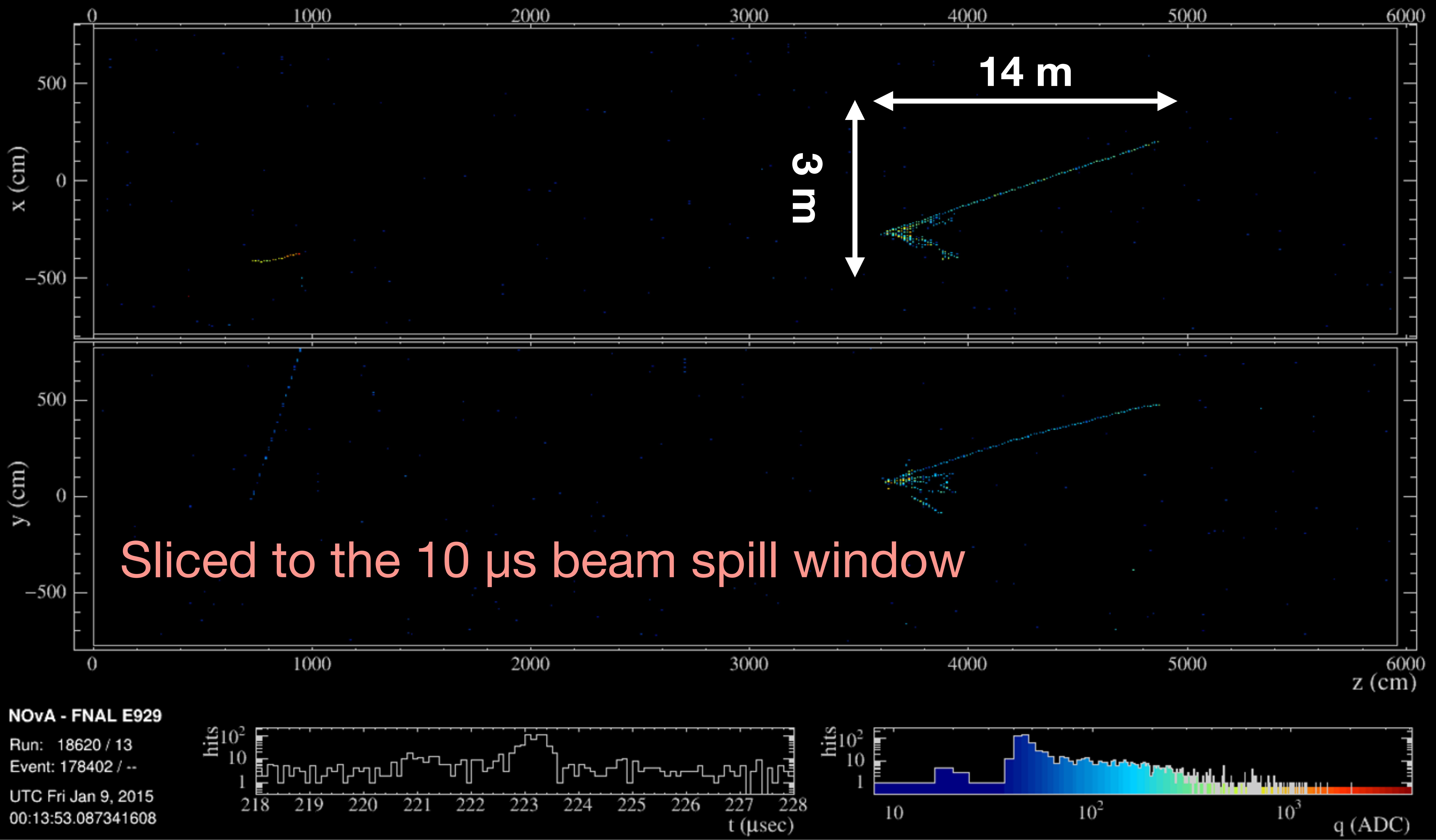


- FD and ND are functionally identical to minimize systematics
- Composed of highly reflective extruded PVC cells filled with liquid scintillator. Scintillation light captured and routed to Avalanche Photodiode (APD) via wavelength shifting fiber (WLS)
- Cells arranged in planes, assembled in alternating horizontal and vertical directions → provide 3D views of the events

Raw Far Detector readout in a 550 μs window

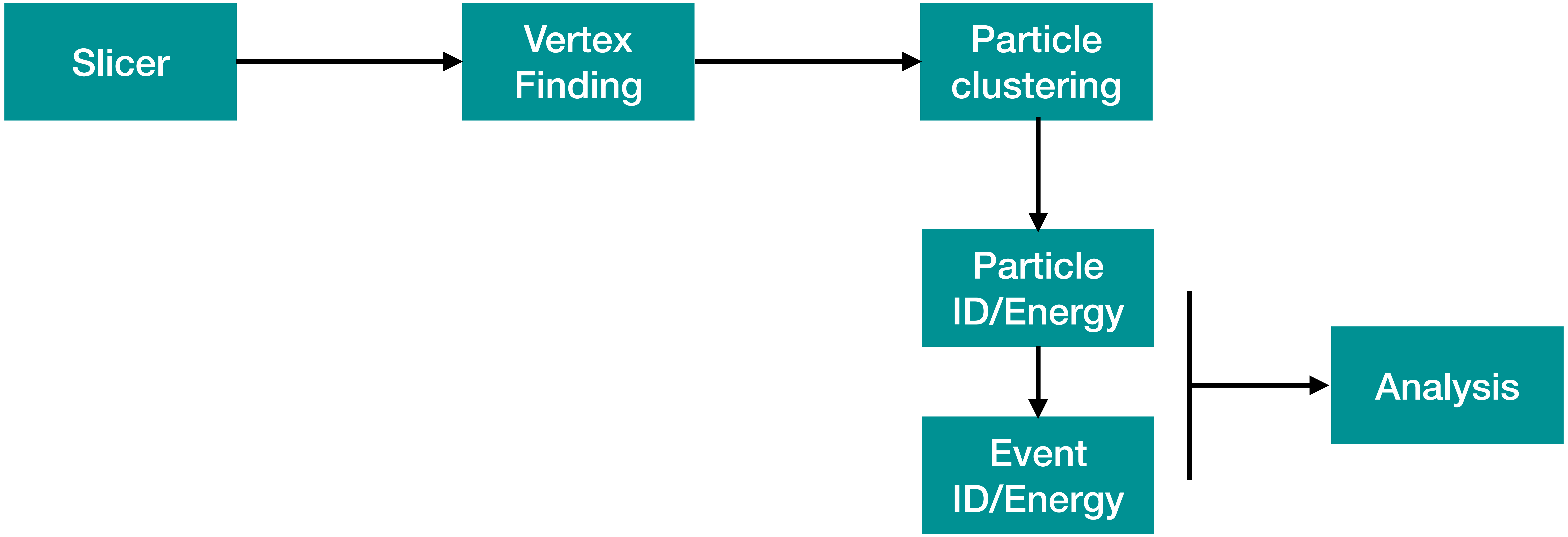
On surface sees ~130 kHz cosmic rate





14 m
3 m
x (cm)
y (cm)
z (cm)
Sliced to the 10 μs beam spill window
NOvA - FNAL E929
Run: 18620 / 13
Event: 178402 / --
UTC Fri Jan 9, 2015
00:13:53.087341608
hits
t (μsec)
q (ADC)

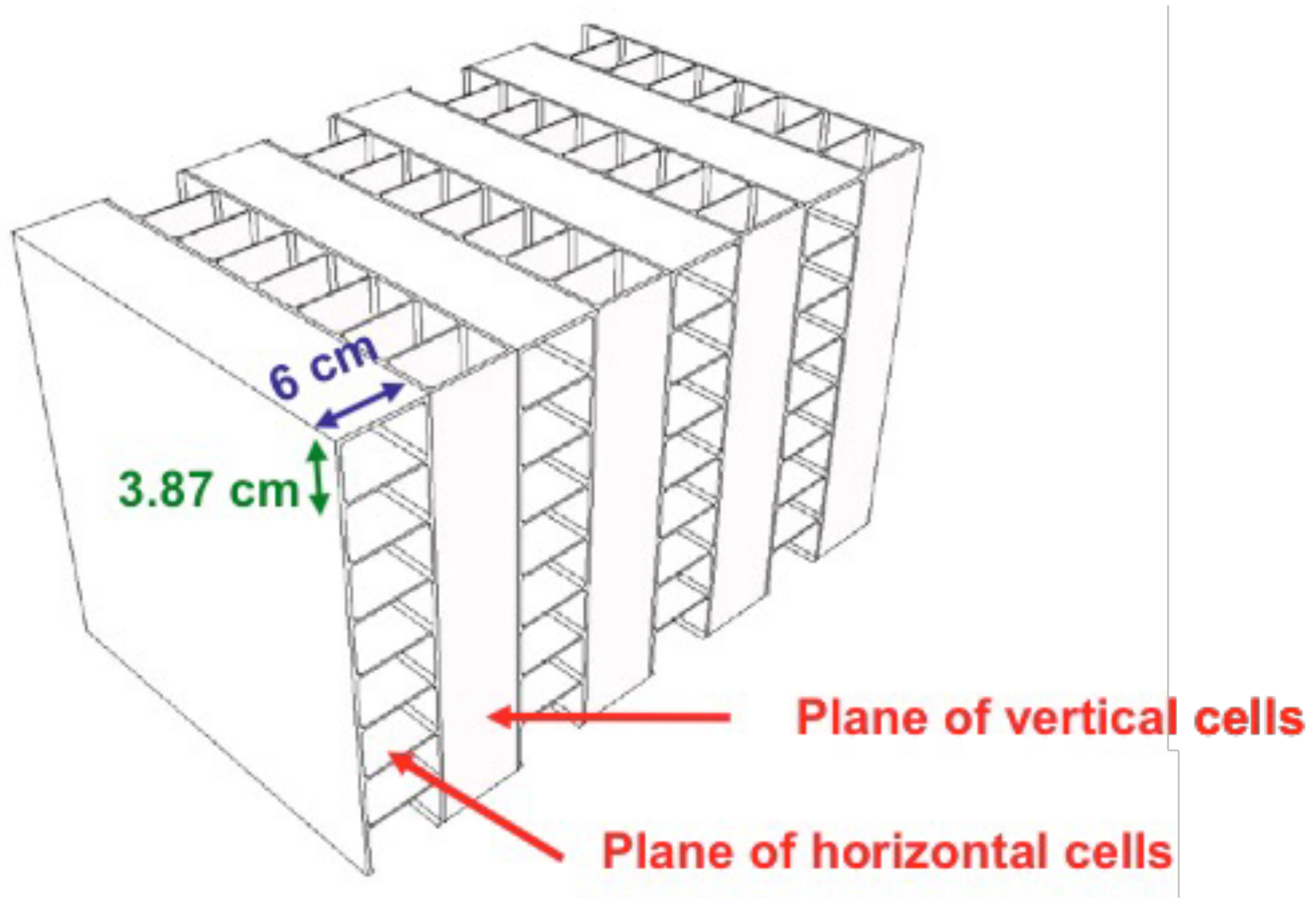
Event Reconstruction in NOvA

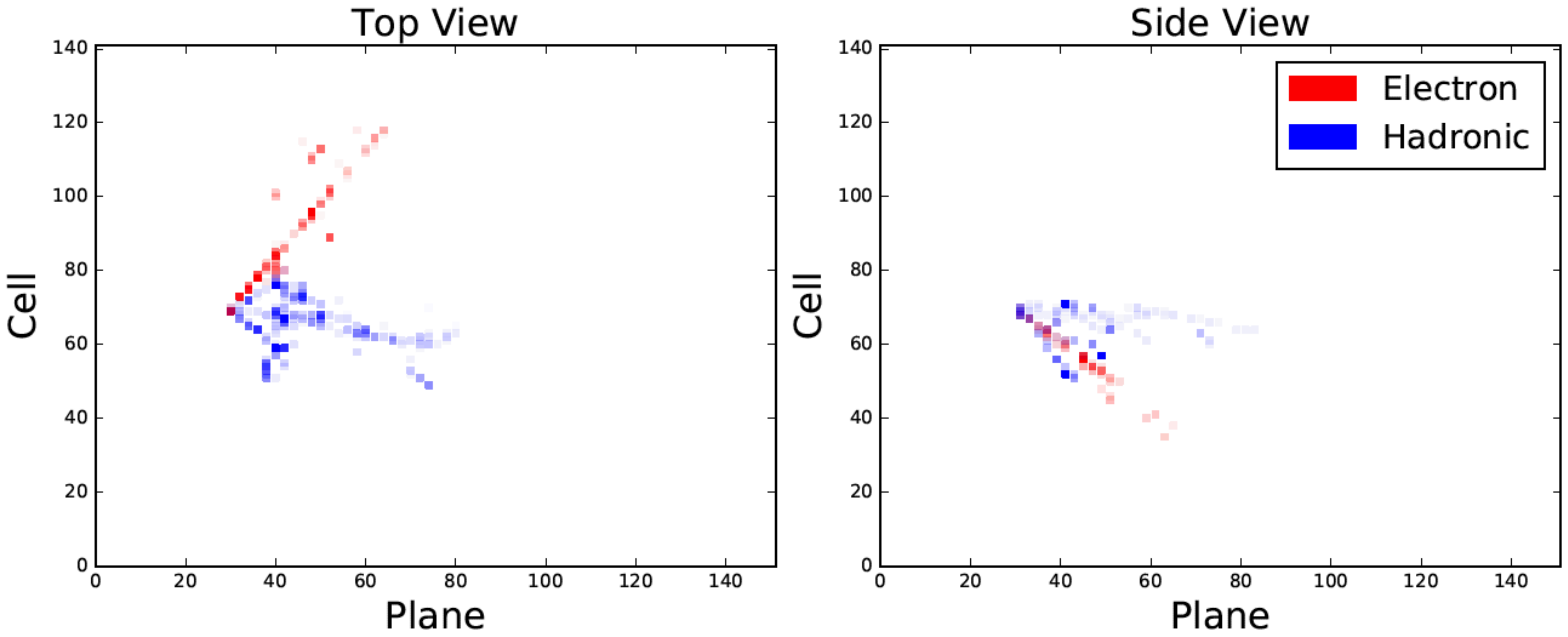


- NOvA uses a variety of algorithms to reconstruct physics information for which slicing is a core input
- Machine learning is making significant contribution in the reconstruction chain and can replace “traditional” kinematic based algorithms in some cases

Detector Views

- NOvA detectors are naturally segmented
- Producing a pair of pixel maps (Cell number v.s. Plane number) for the Top and Side view of each interaction





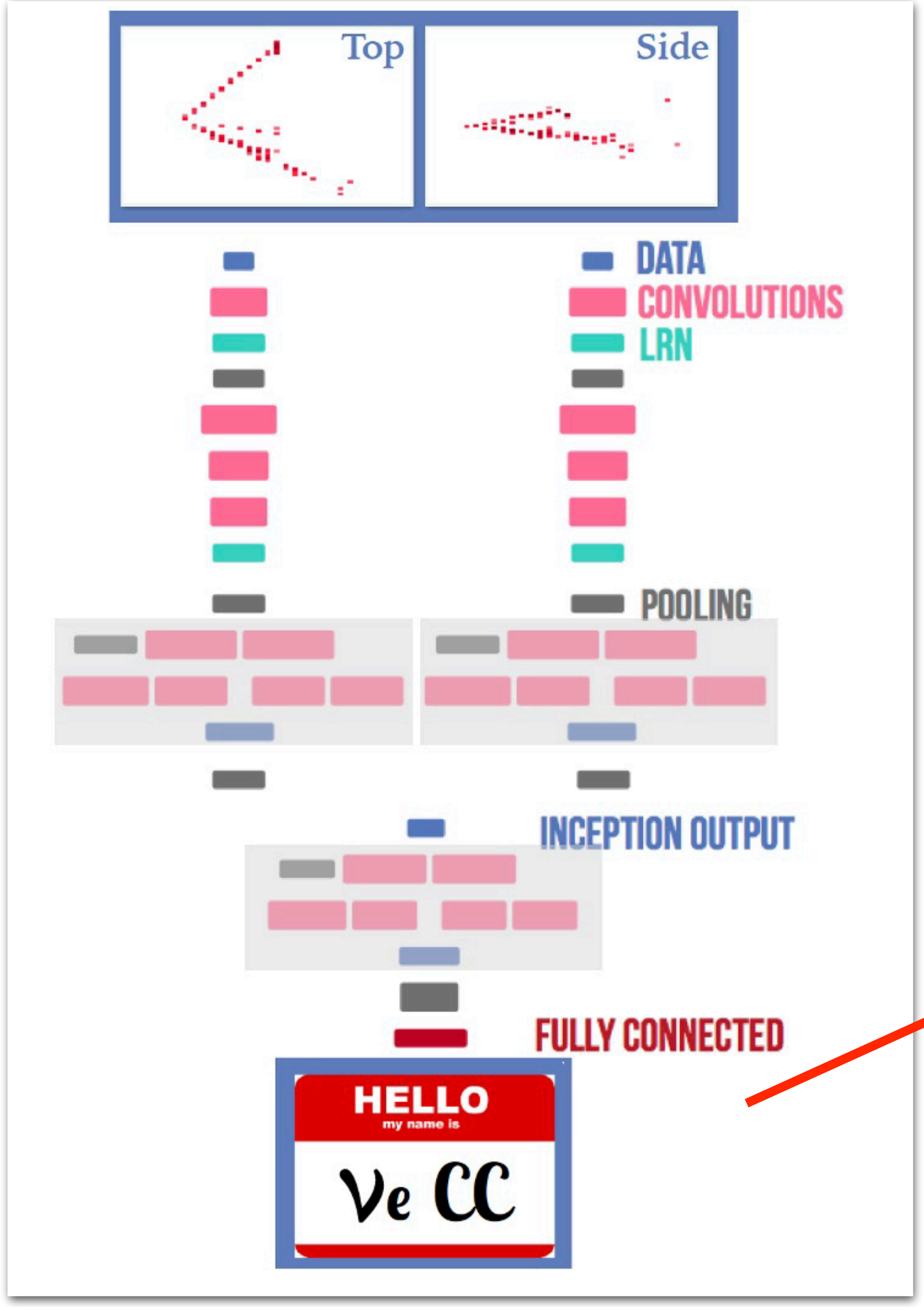
CNN-based Event Classifier (EventCVN)

- CVN: a convolutional neural network, based on modern image recognition technology, identifies neutrino interactions directly from pixel maps

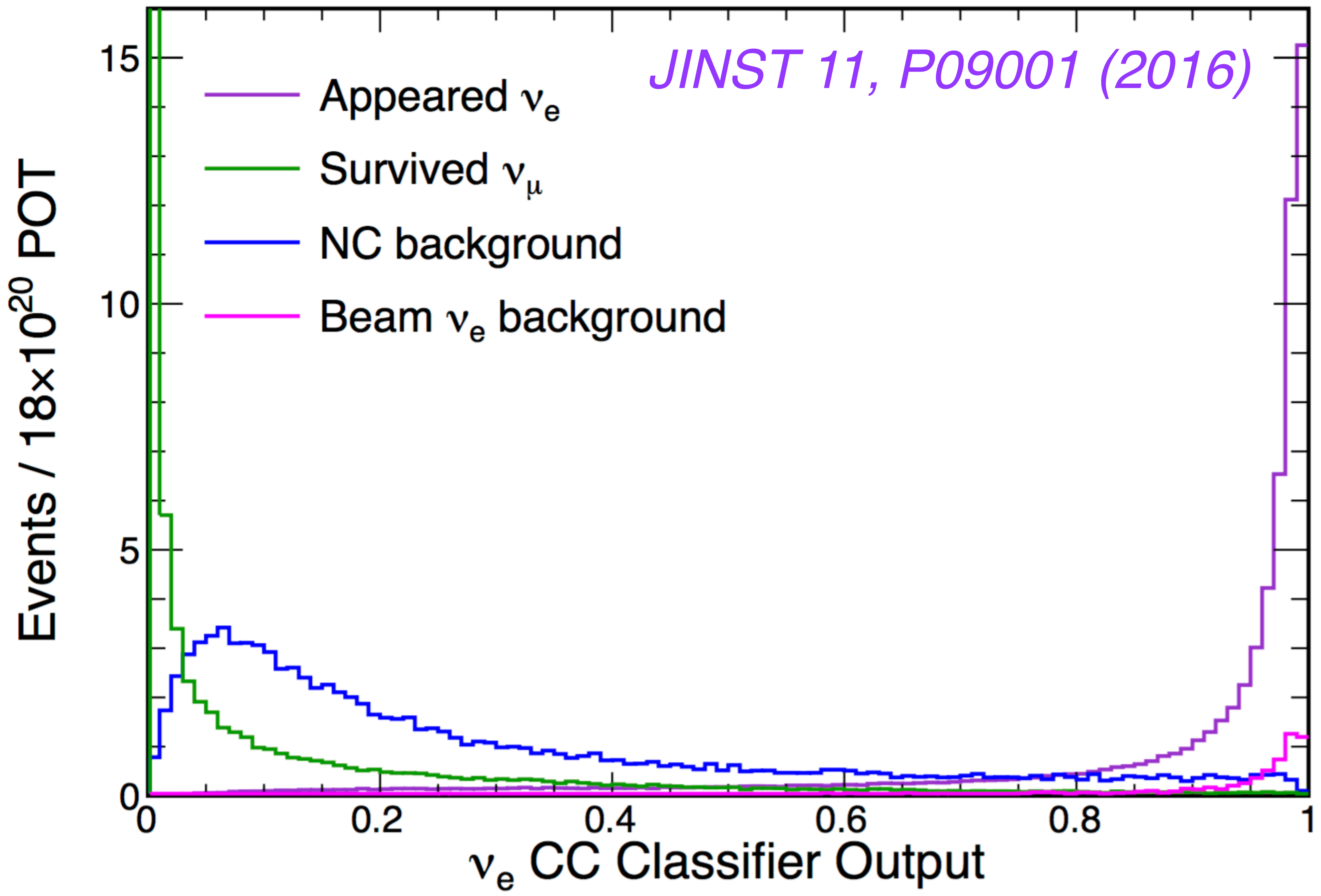
- NOvA is the first HEP experiment to apply CNNs to publish physics results: *Phys.Rev.Lett. 118 (2017)*

- Increased in sensitivity to neutrino oscillation parameters over traditional methods equivalent to collecting 30% more exposure

CNN architecture
2016: GoogleNet
Now: Modified MobileNetv2



CVN output in the far detector MC

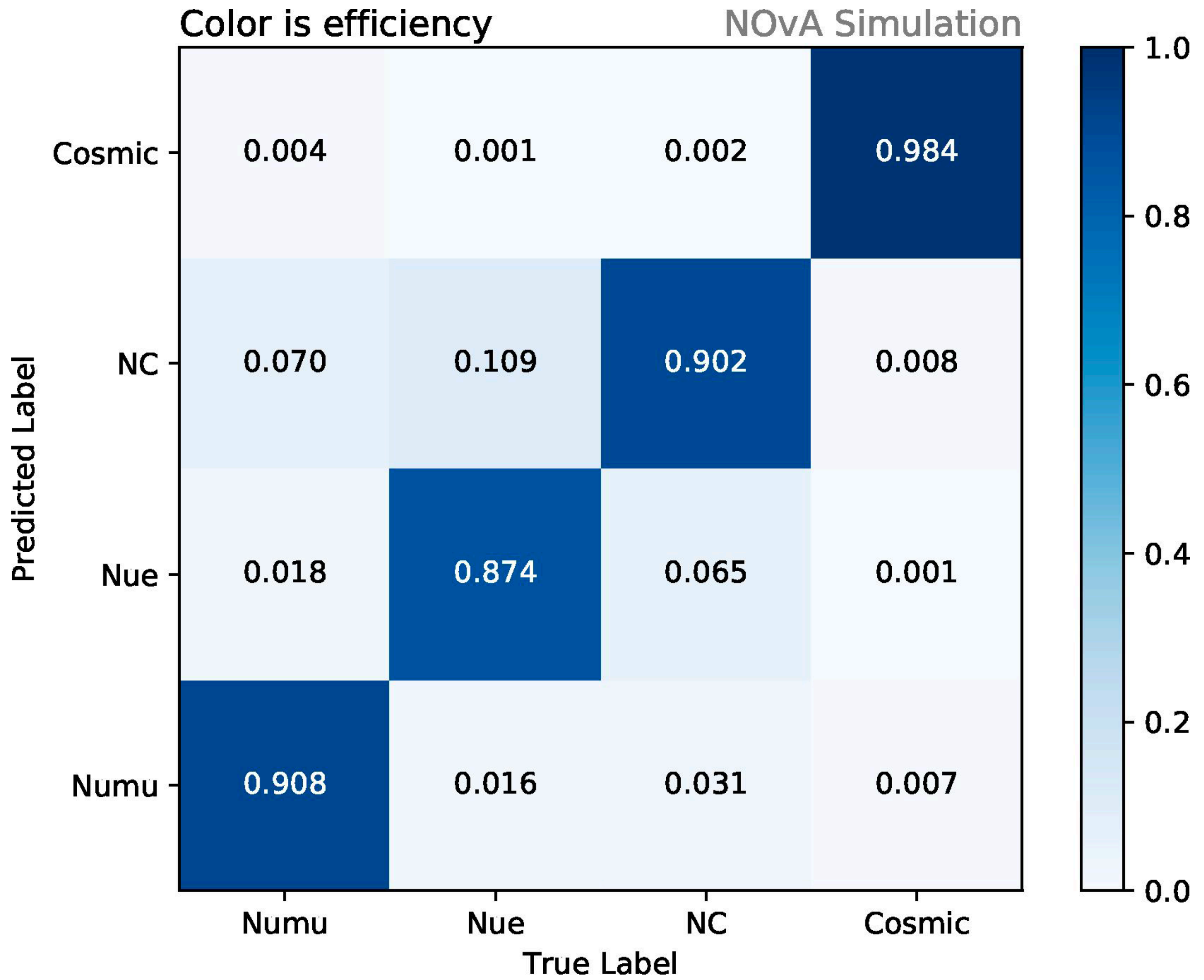


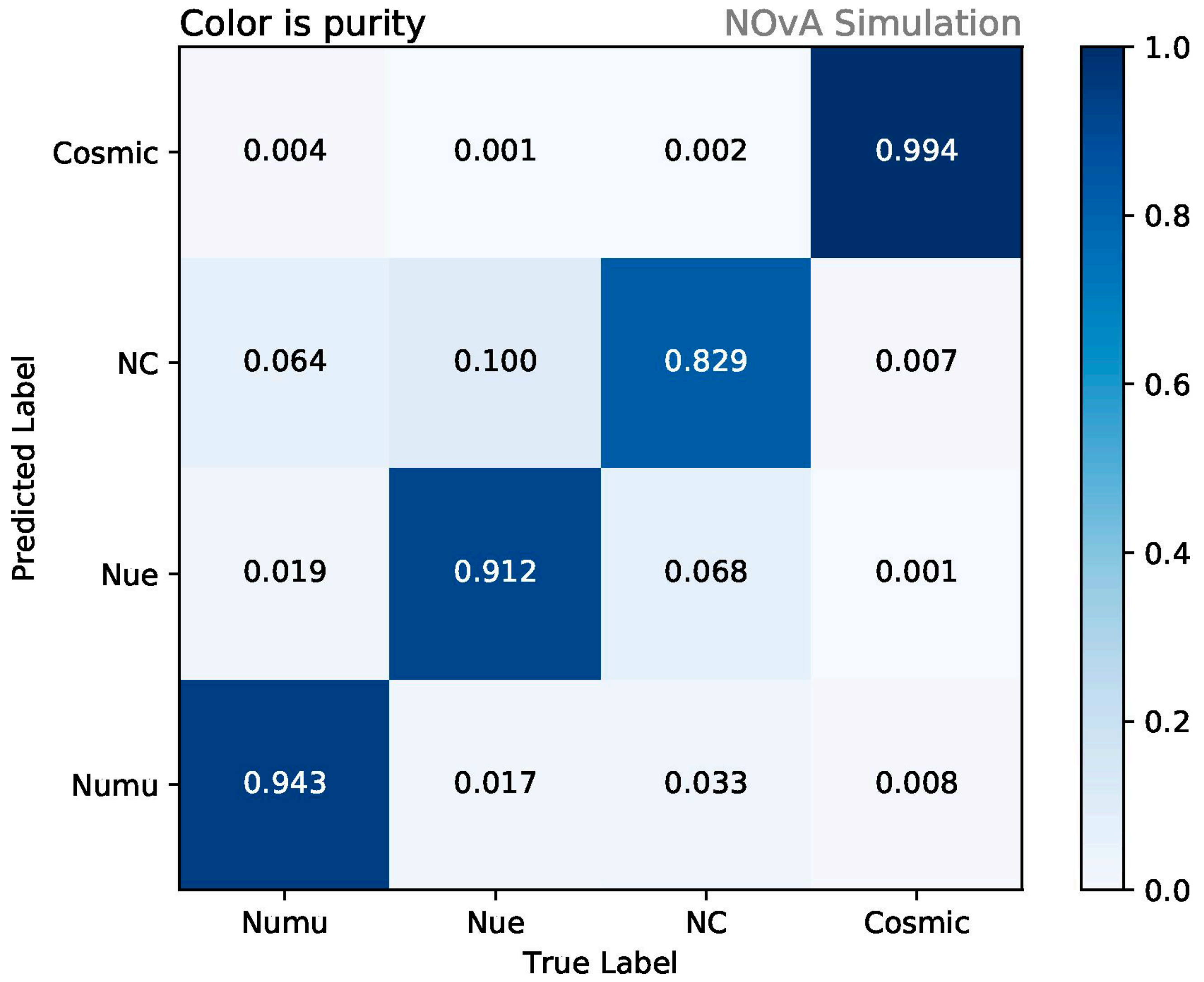
Select ν_μ ($\bar{\nu}_\mu$) CC and ν_e ($\bar{\nu}_e$) CC candidates from neutrino (anti-neutrino) beam with CVN

CNN-based Event Classifier (EventCVN)

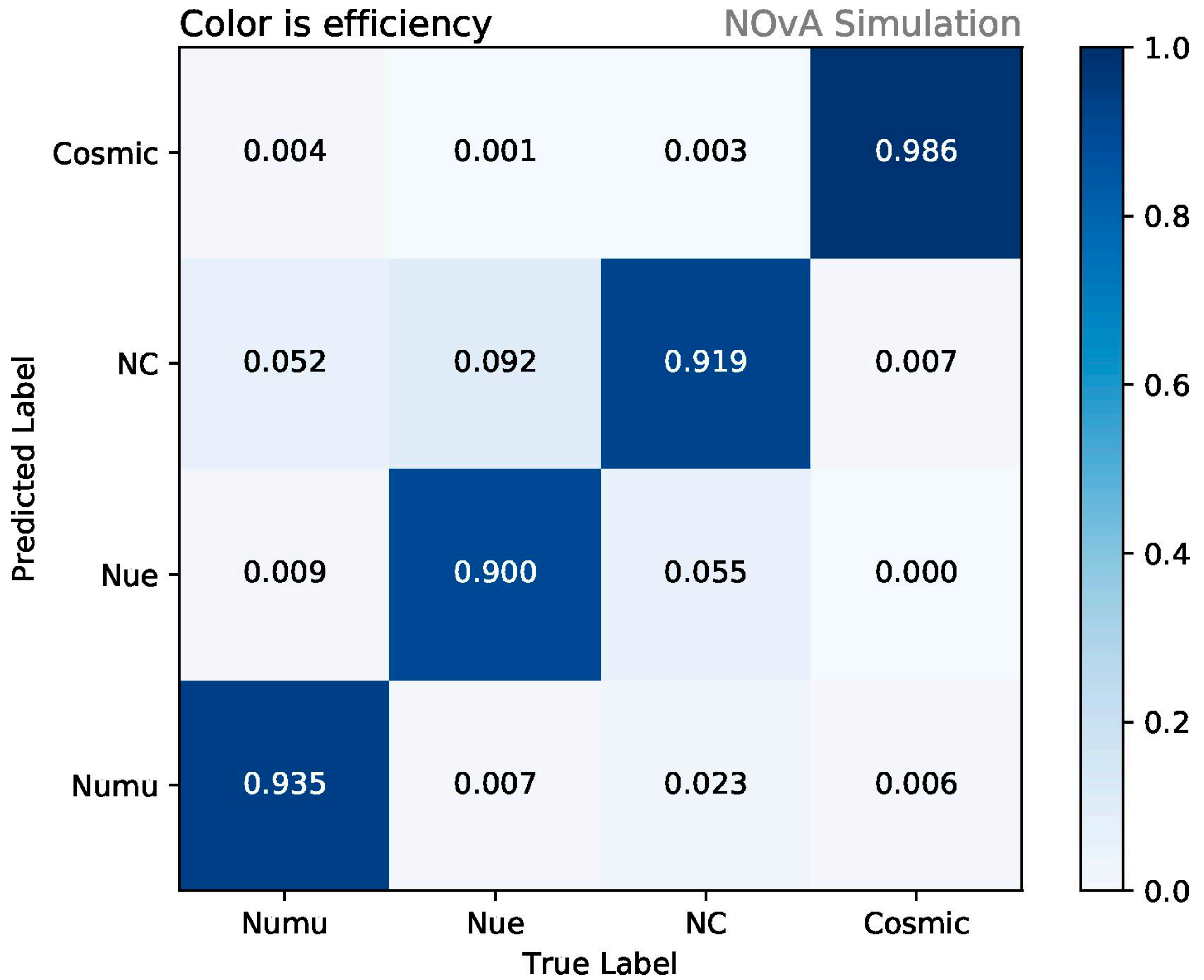
- Similar performance for neutrino and anti-neutrino modes
- Anti-neutrino mode shows slight increase in efficiency
- Purity over 90% for all interaction flavors

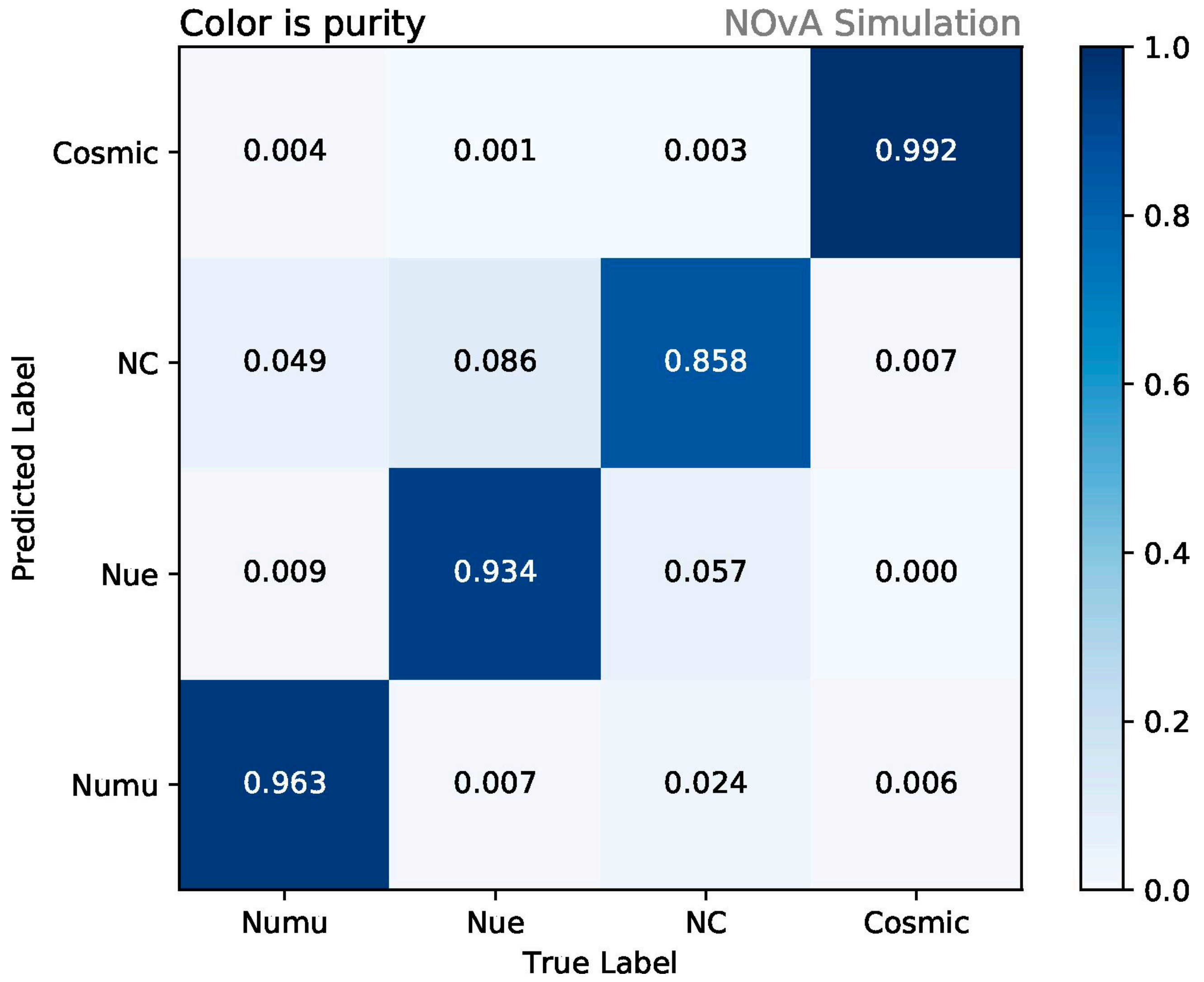
neutrino mode



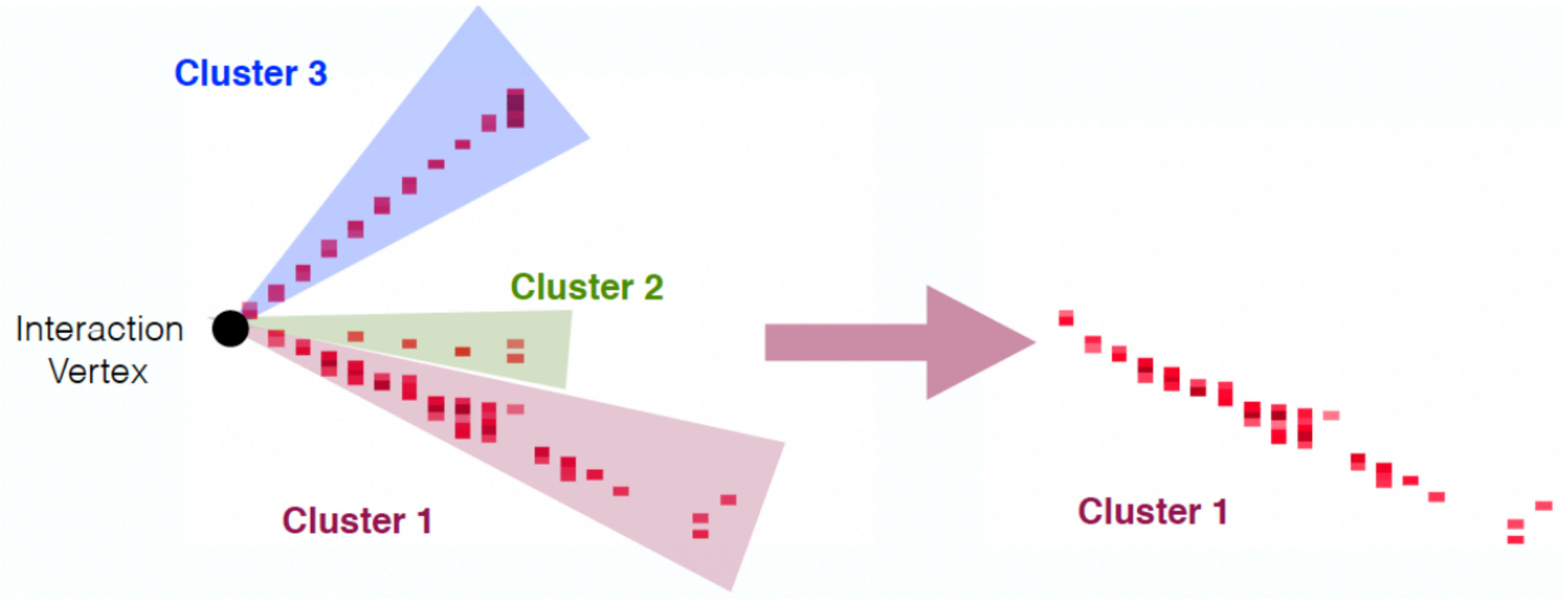


anti-neutrino mode



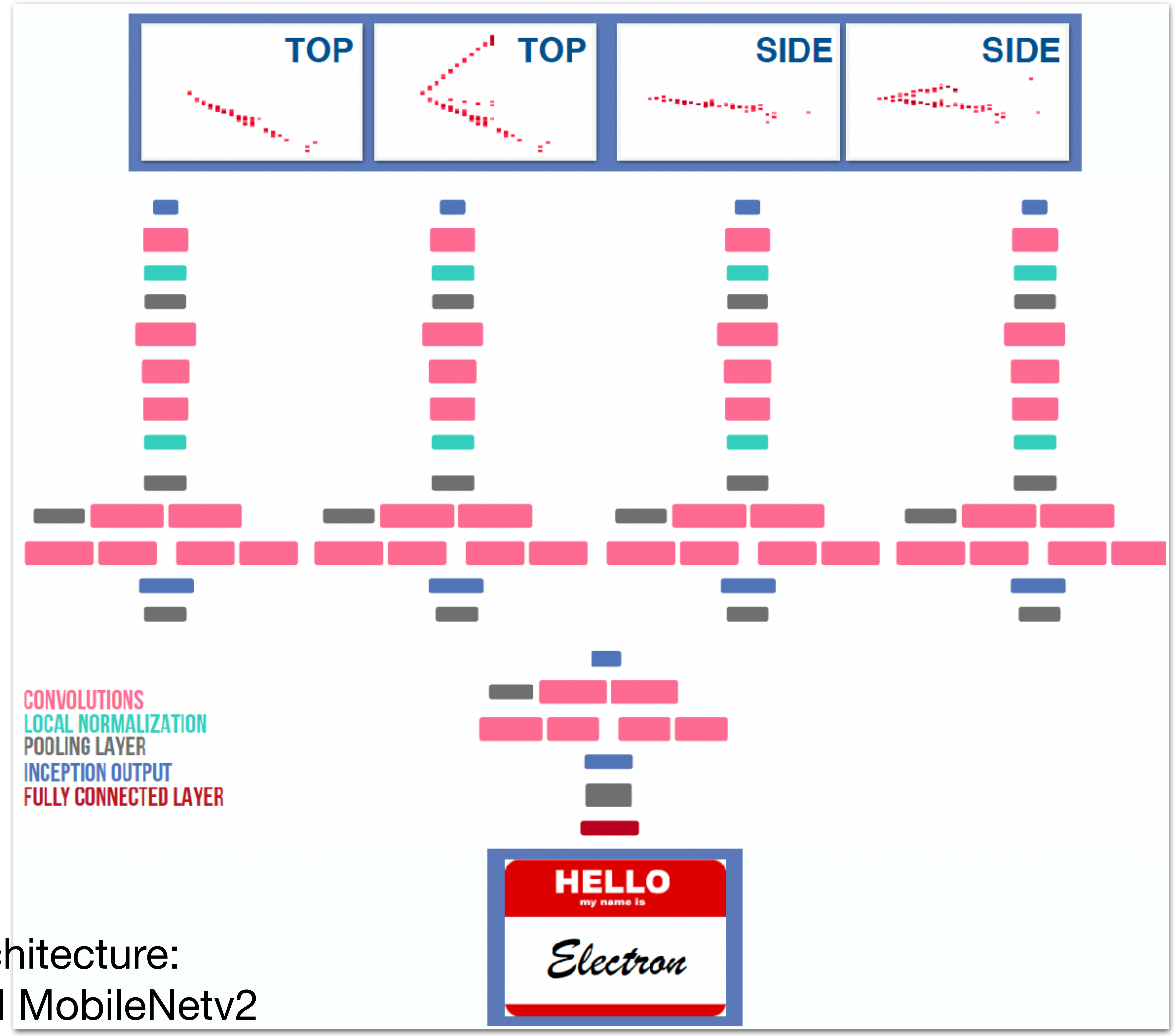


CNN-based Particle Classifier (ProngCVN)



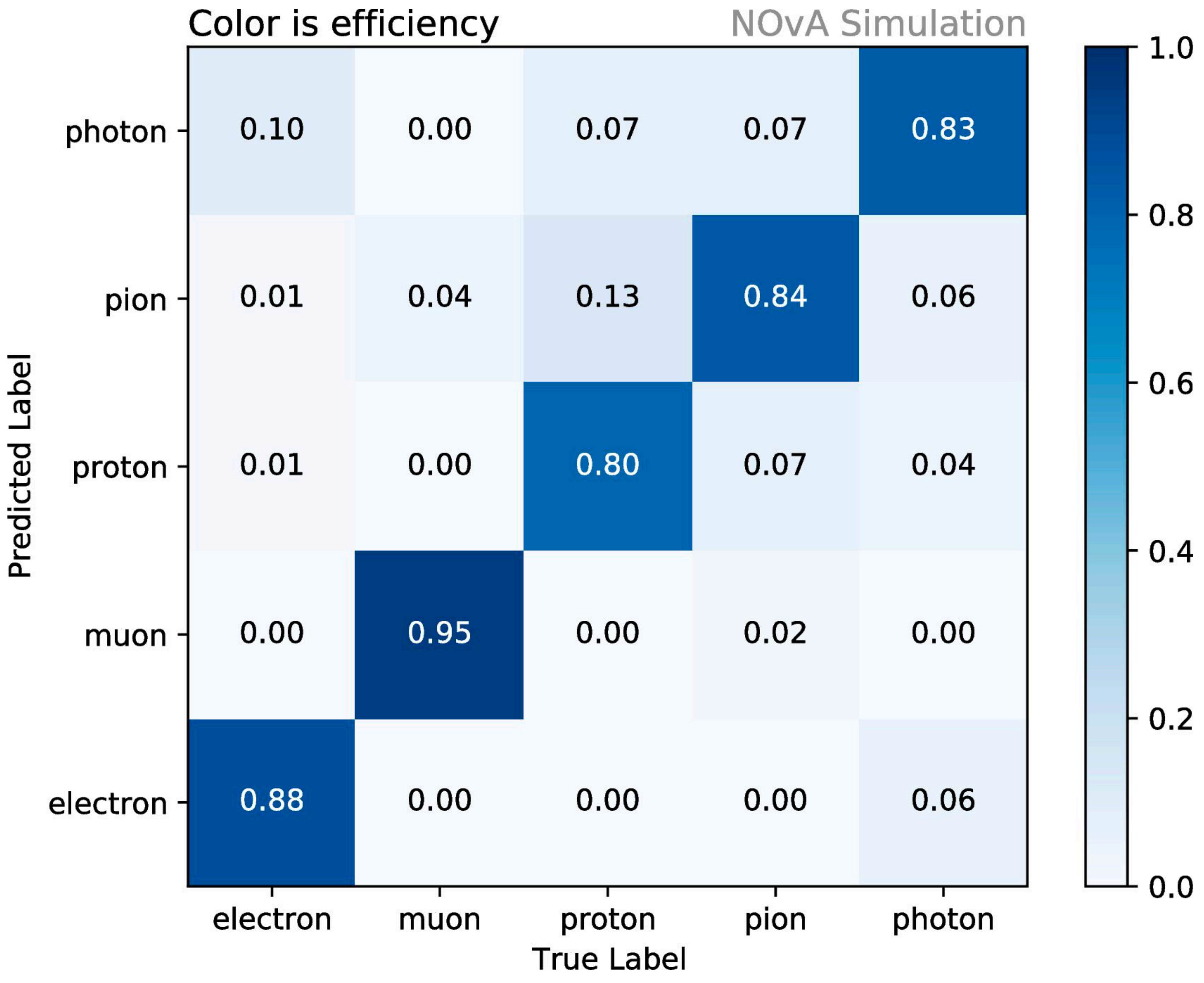
- Single particles are separated using geometric reconstruction methods
- Classify particles using both views of the **particle** and both views of the entire **event**
- This shows the network contextual information about single particles

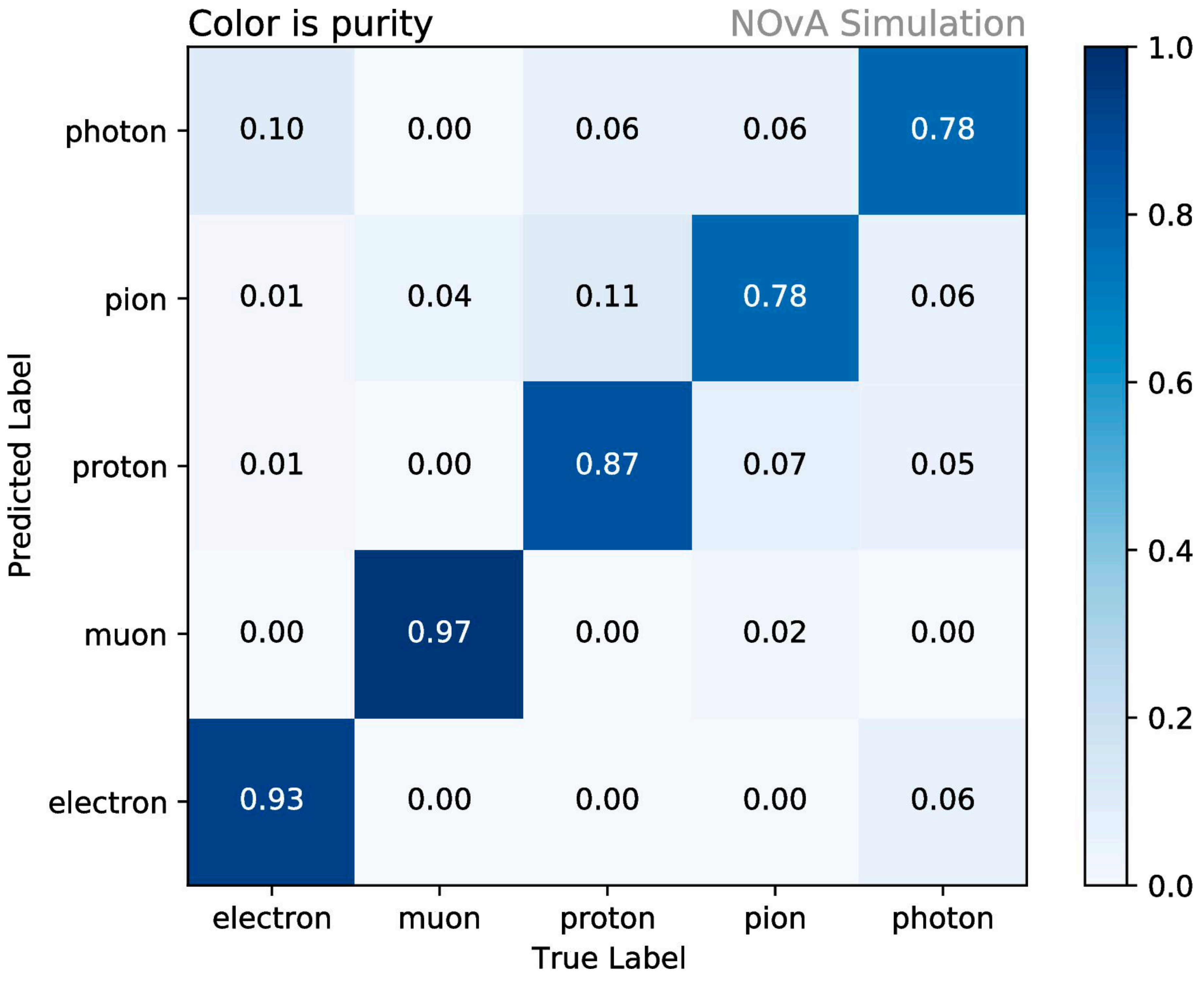
Phys.Rev.D 100 (2019) 7, 073005



CNN architecture:
Modified MobileNetv2
Four-tower Siamese structure

CNN-based Particle Classifier (ProngCVN)



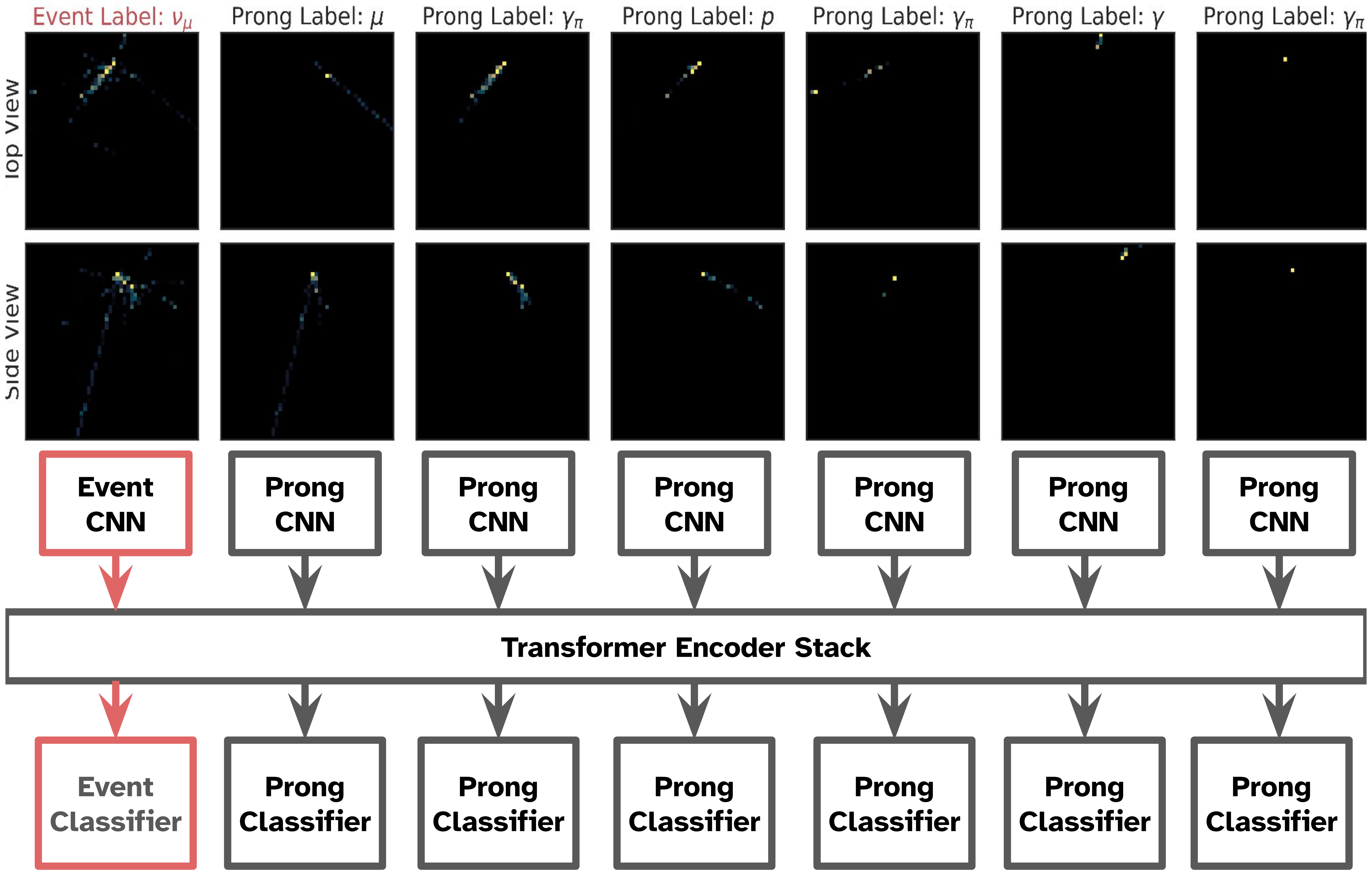


- Improvements were found in both efficiency and purity for all particle types, compared to the particle-only network

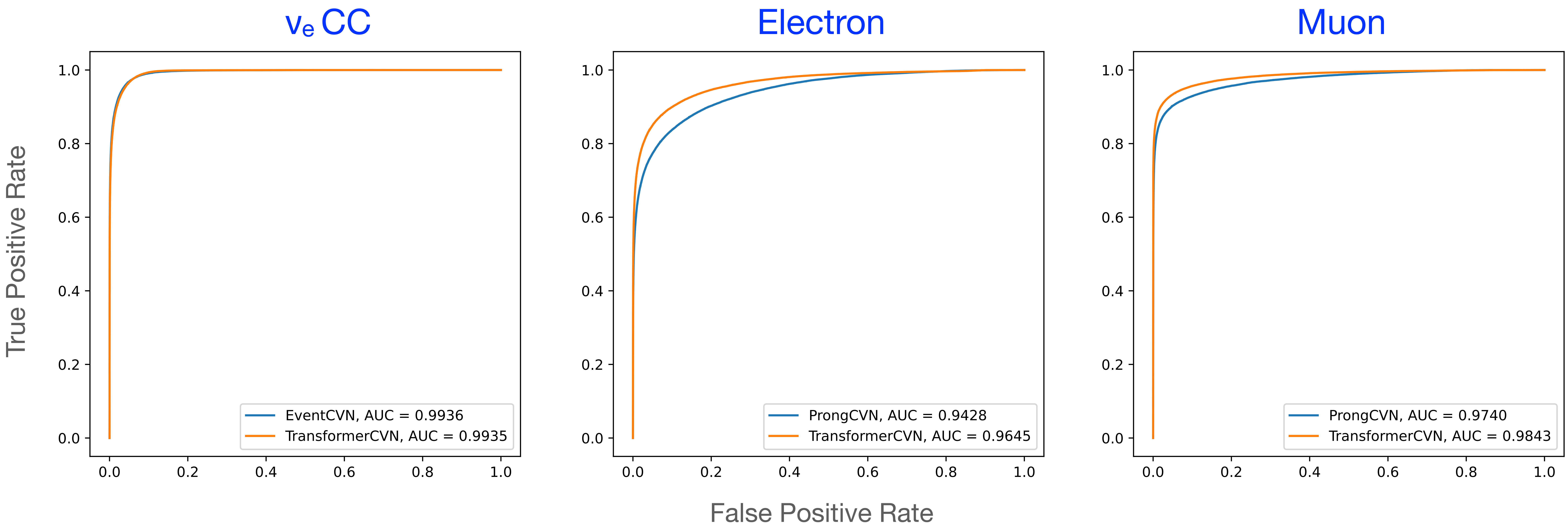
- In particular ~10% increase in the efficiency of selecting photons and pions

Transformer for both Event and Particle Classification

- TransformerCVN, a novel NN that combines the spatial learning enabled by convolutions with the contextual learning enabled by attention, simultaneously classifies each event and reconstructs every individual particle's identity
- Ideal for learning combinatorial relationships of variable-length sets, and has been commonly used in language processing
- It also enables performing interpretability studies



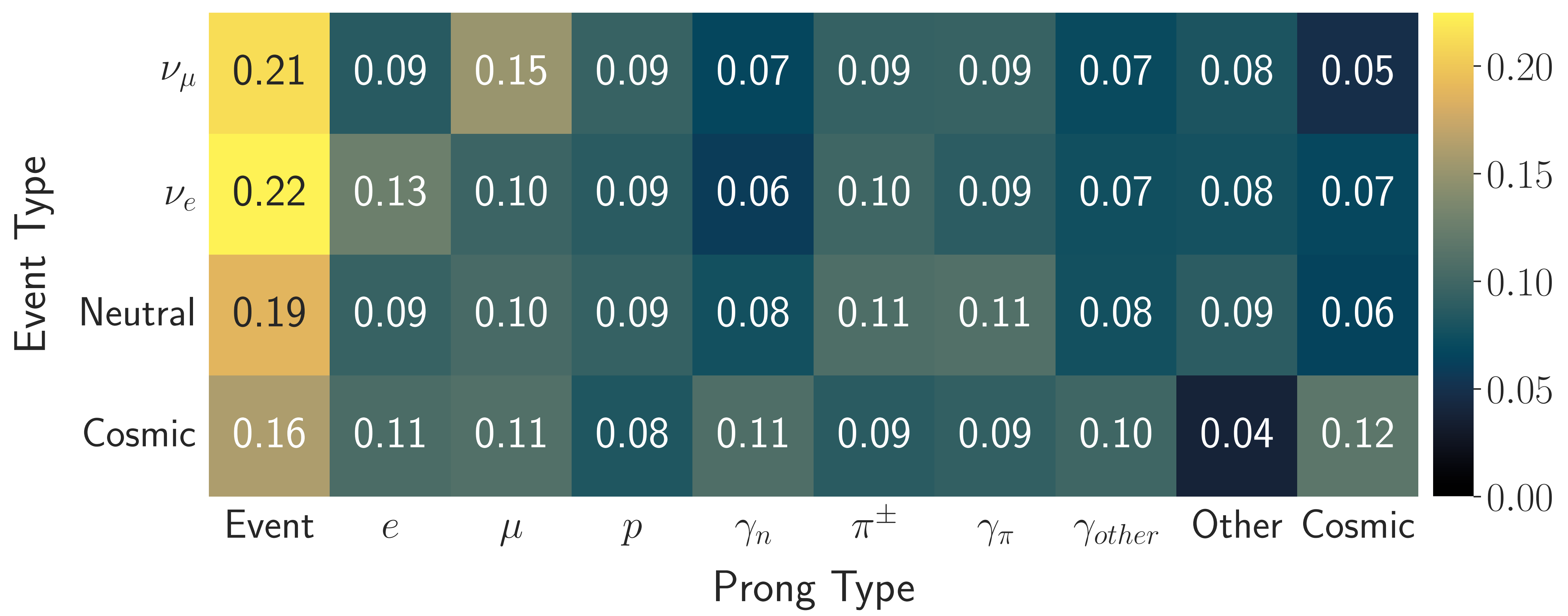
Transformer for both Event and Particle Classification



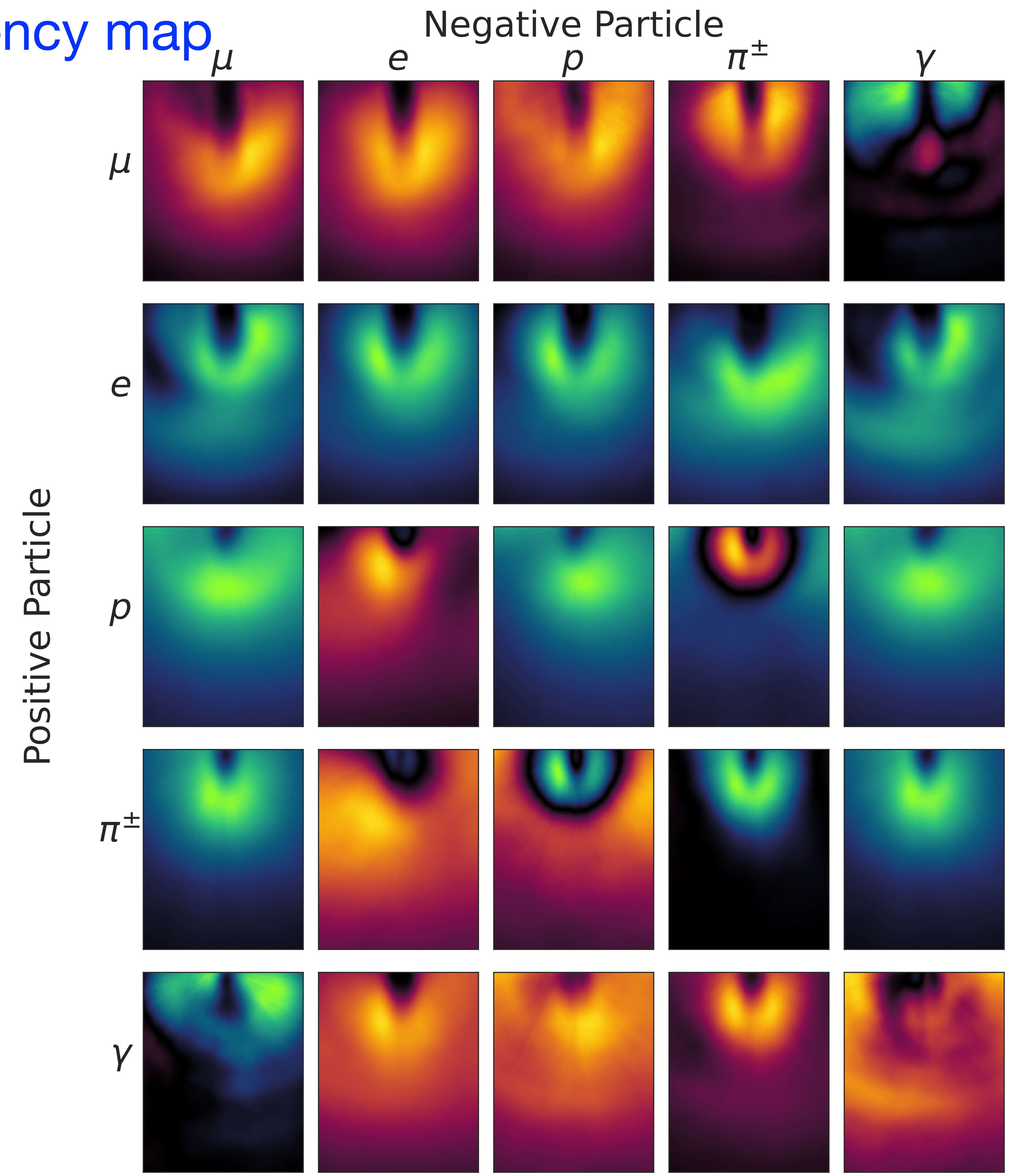
- Comparable performance of identifying neutrino flavors compared to EventCVN
- Great improvement in particle identification, benefits from the additional context provided by all prongs and the transformer's attention mechanism

Transformer for both Event and Particle Classification

Attention map



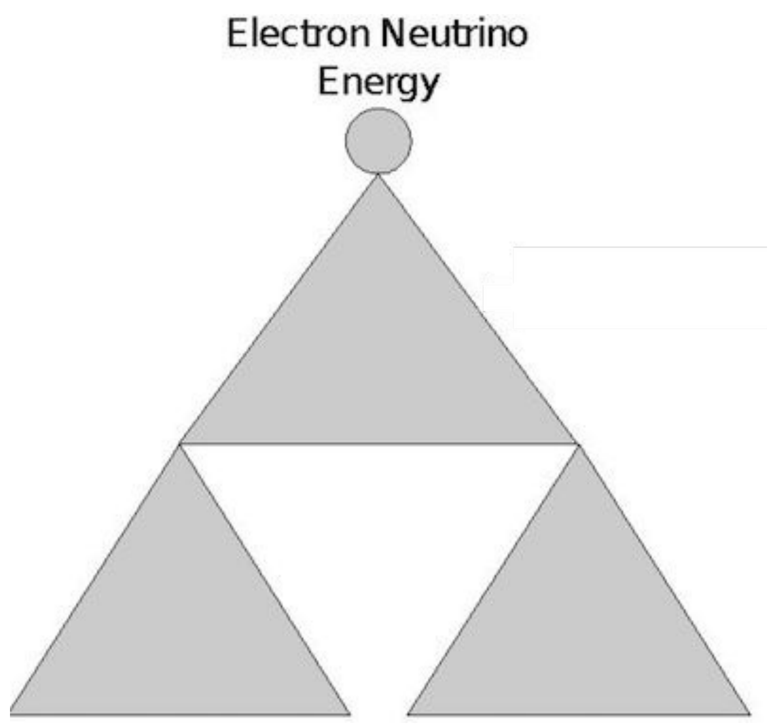
Saliency map

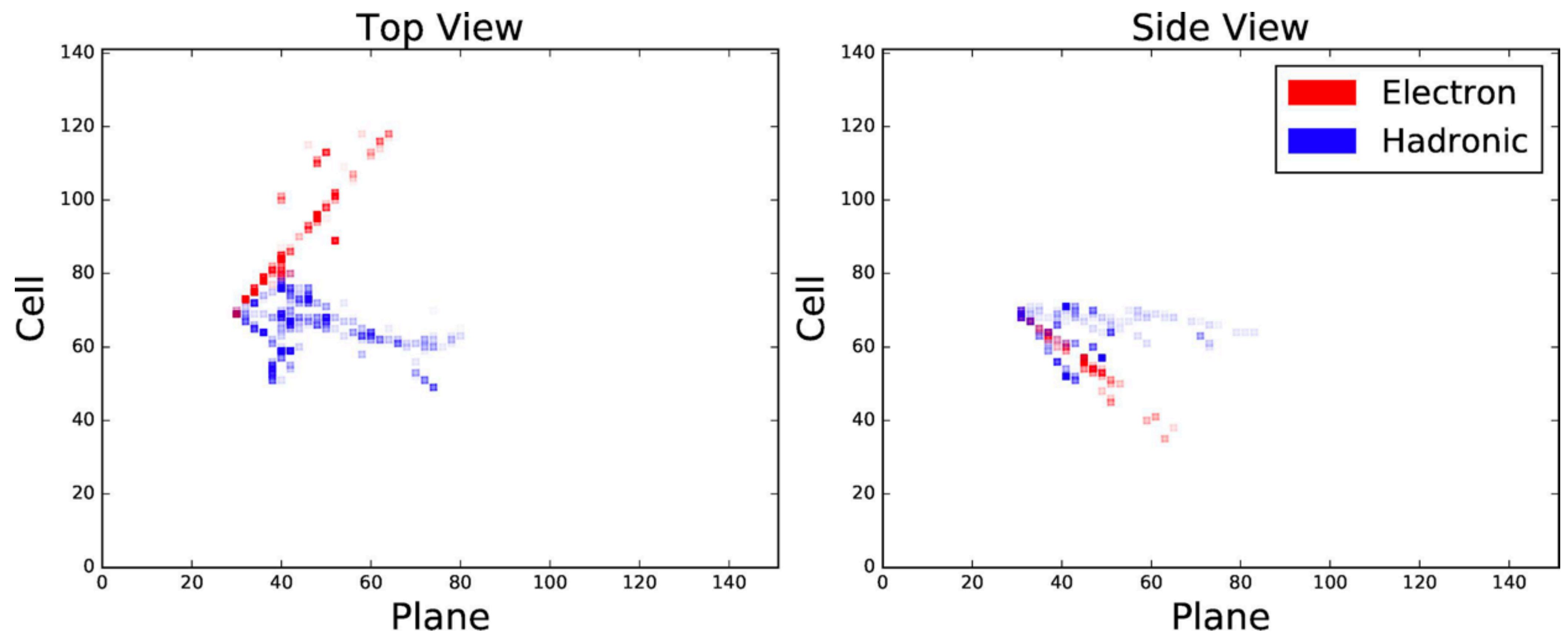


- Interpretability of the network
  - Attention map: importance of each input to each output
  - Saliency map: derivative of a network output w.r.t the input pixel

See Alejandro's talk for the application of transformer to DUNE

Regression CNNs for Energy Estimation

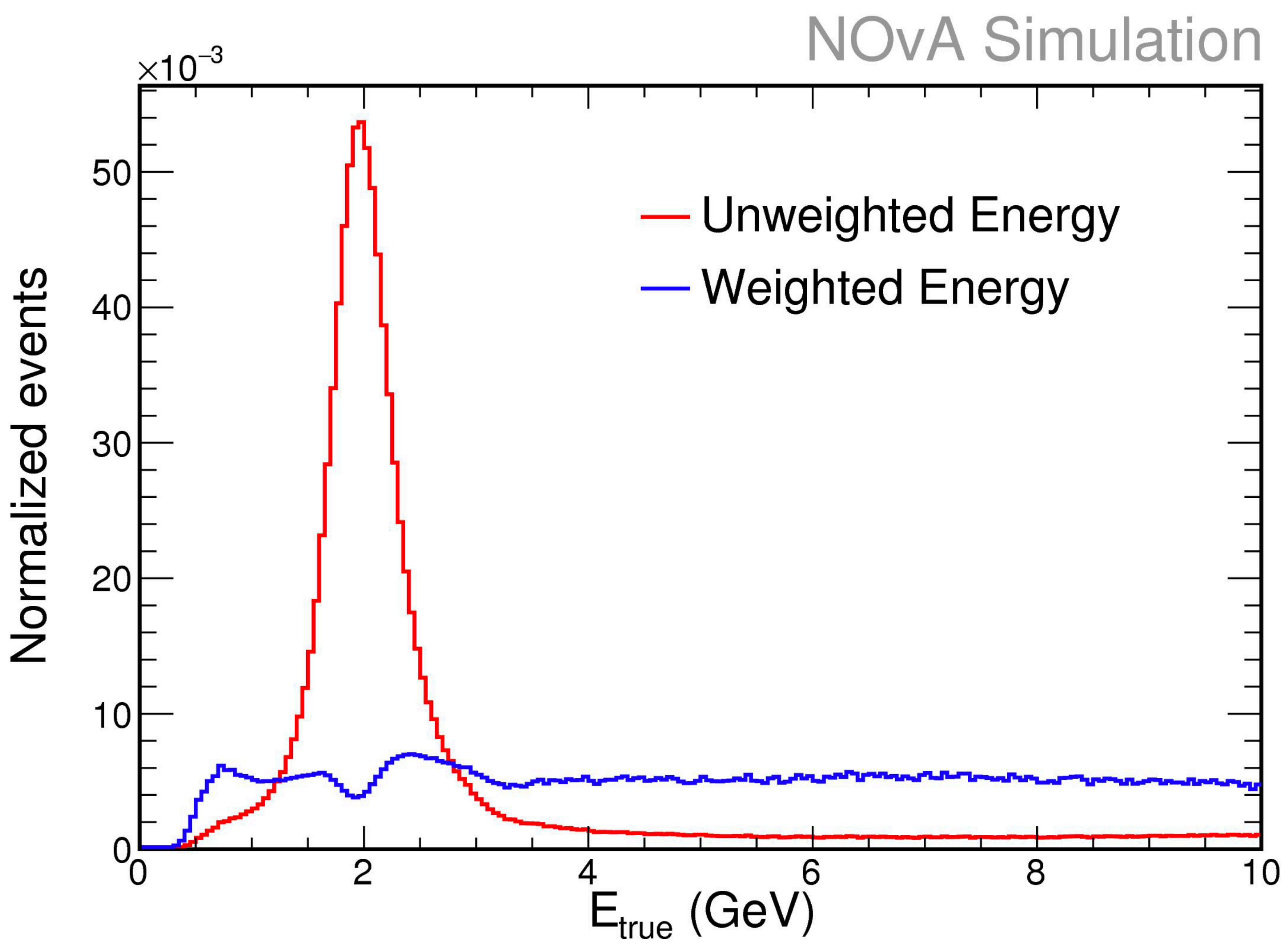




PhysRevD.99.012011

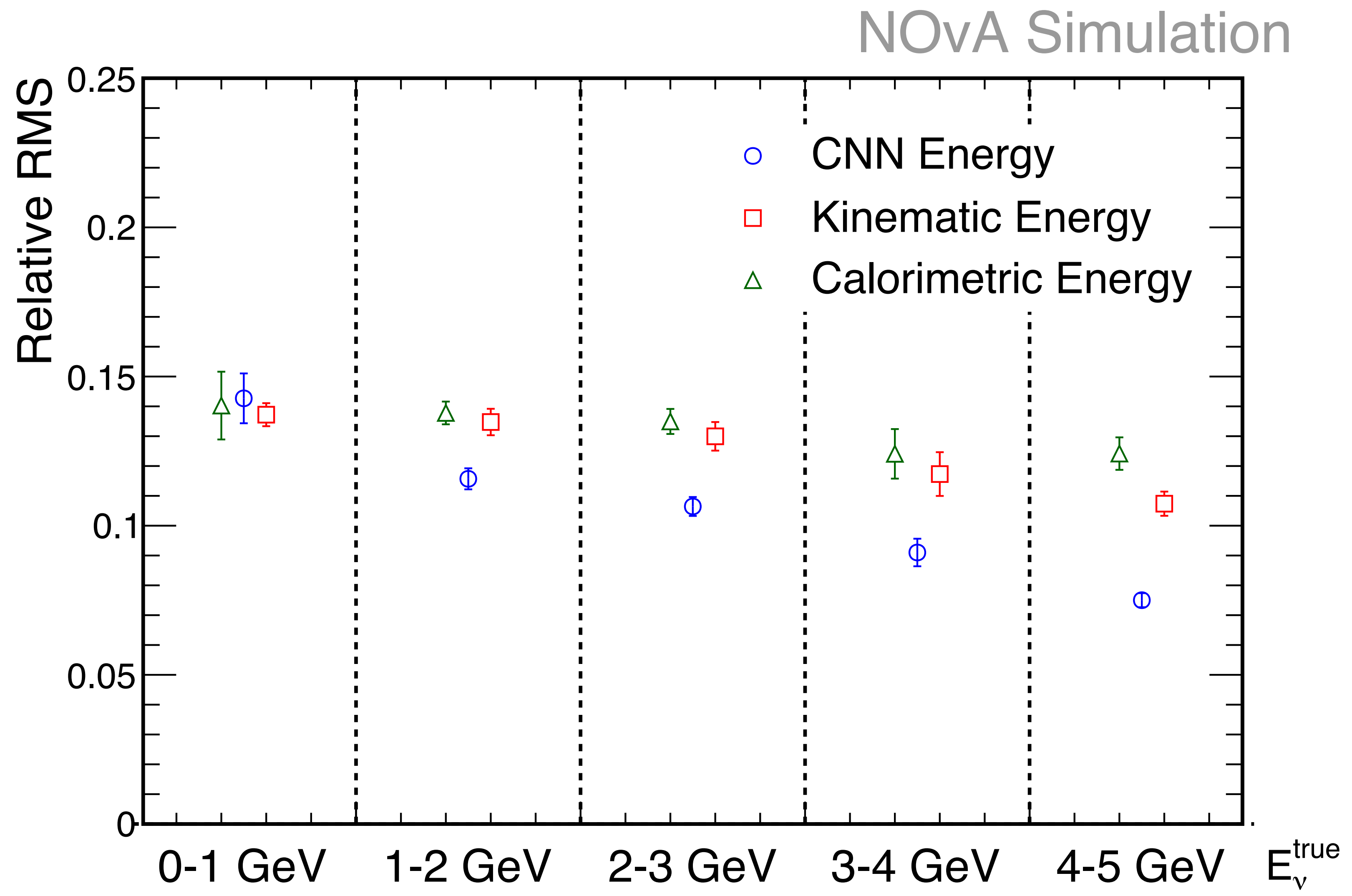
- The CNN architecture used is an adapted ResNet
- Weighting scheme so the loss function sees a flat energy distribution, to control energy dependence
- Use mean absolute percentage error instead of square of errors to decrease the effects of outliers

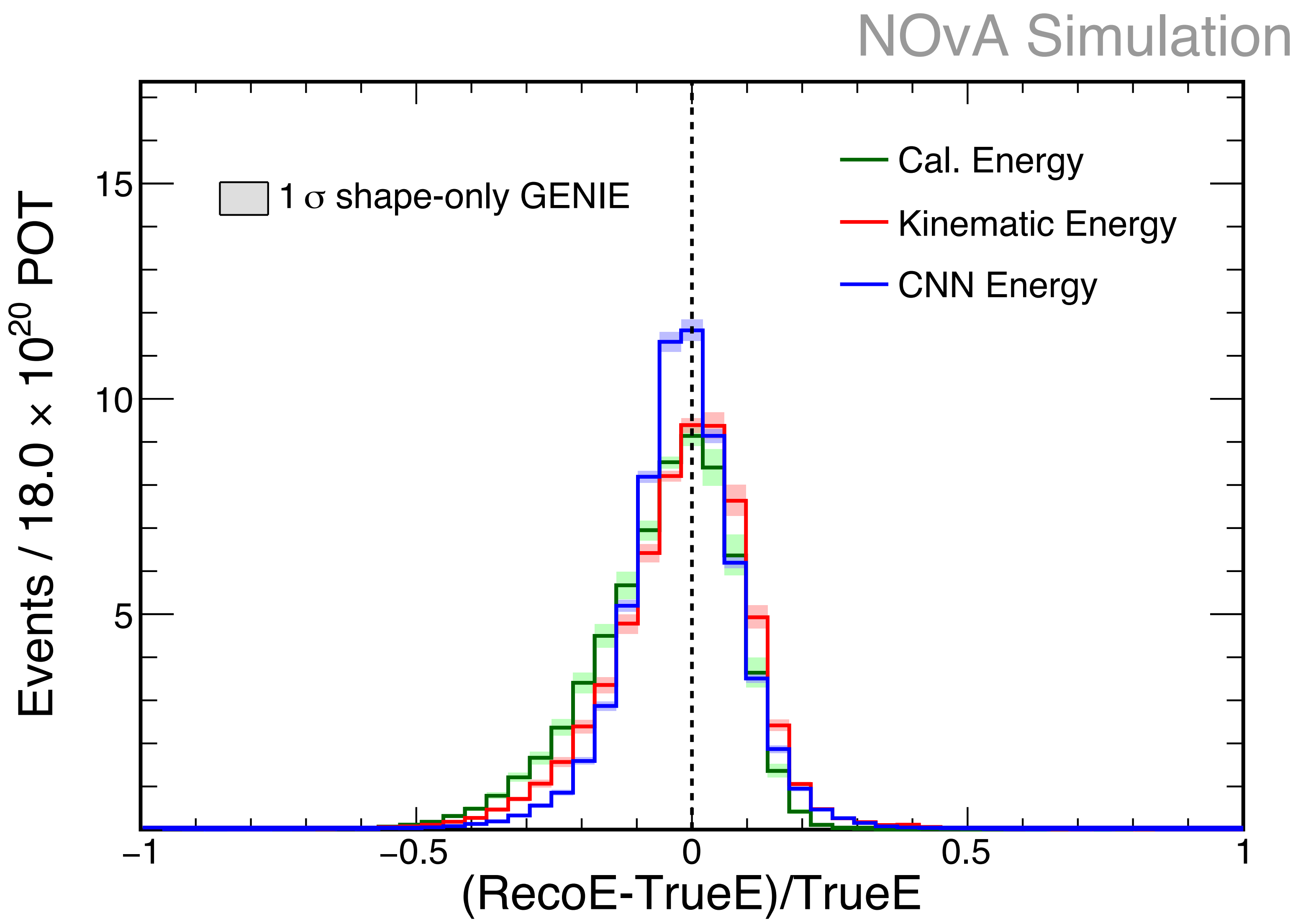
$$L(\mathbf{W}, \{\mathbf{x}_i, y_i\}_{i=1}^{n}) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{f_{\mathbf{W}}(\mathbf{x}_i) - y_i}{y_i}\right|$$

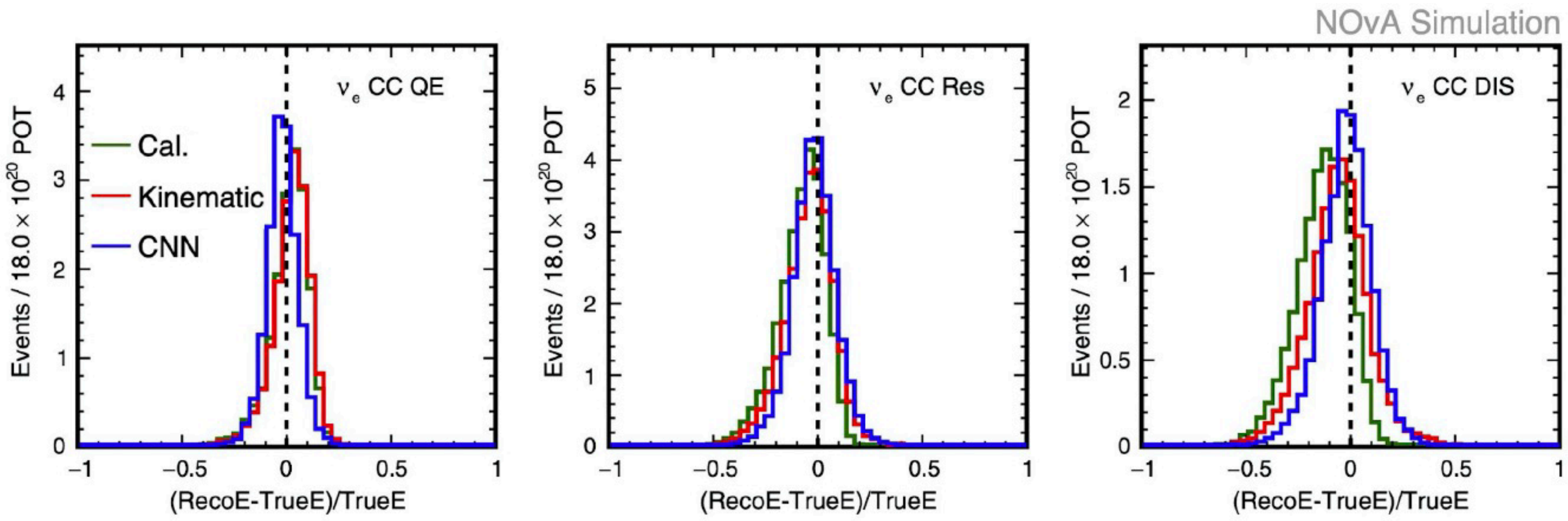


Regression CNNs for Energy Estimation

- Regression CNN shows a better resolution compared with kinematics-based energy reconstruction
- Shows smaller systematic uncertainties due to neutrino interaction simulation
- Good stability over interaction types



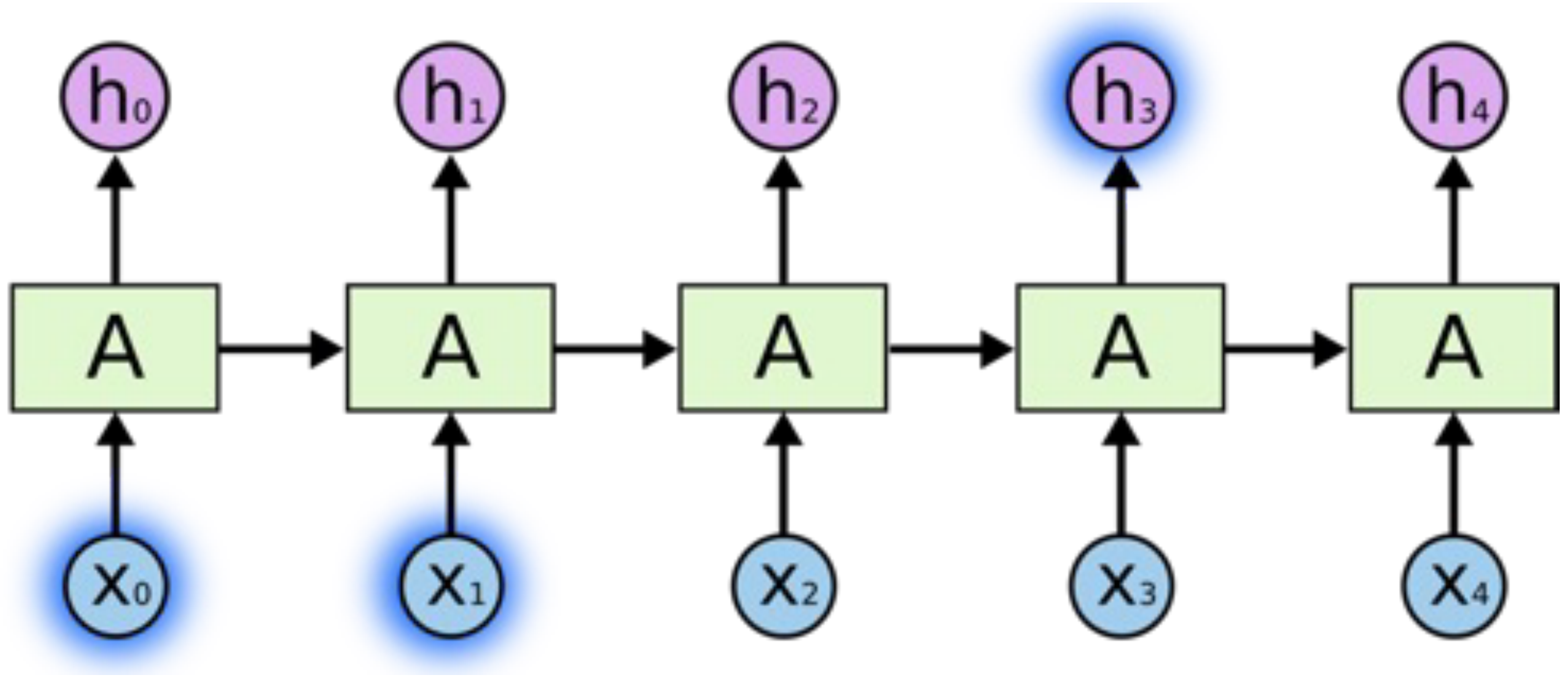


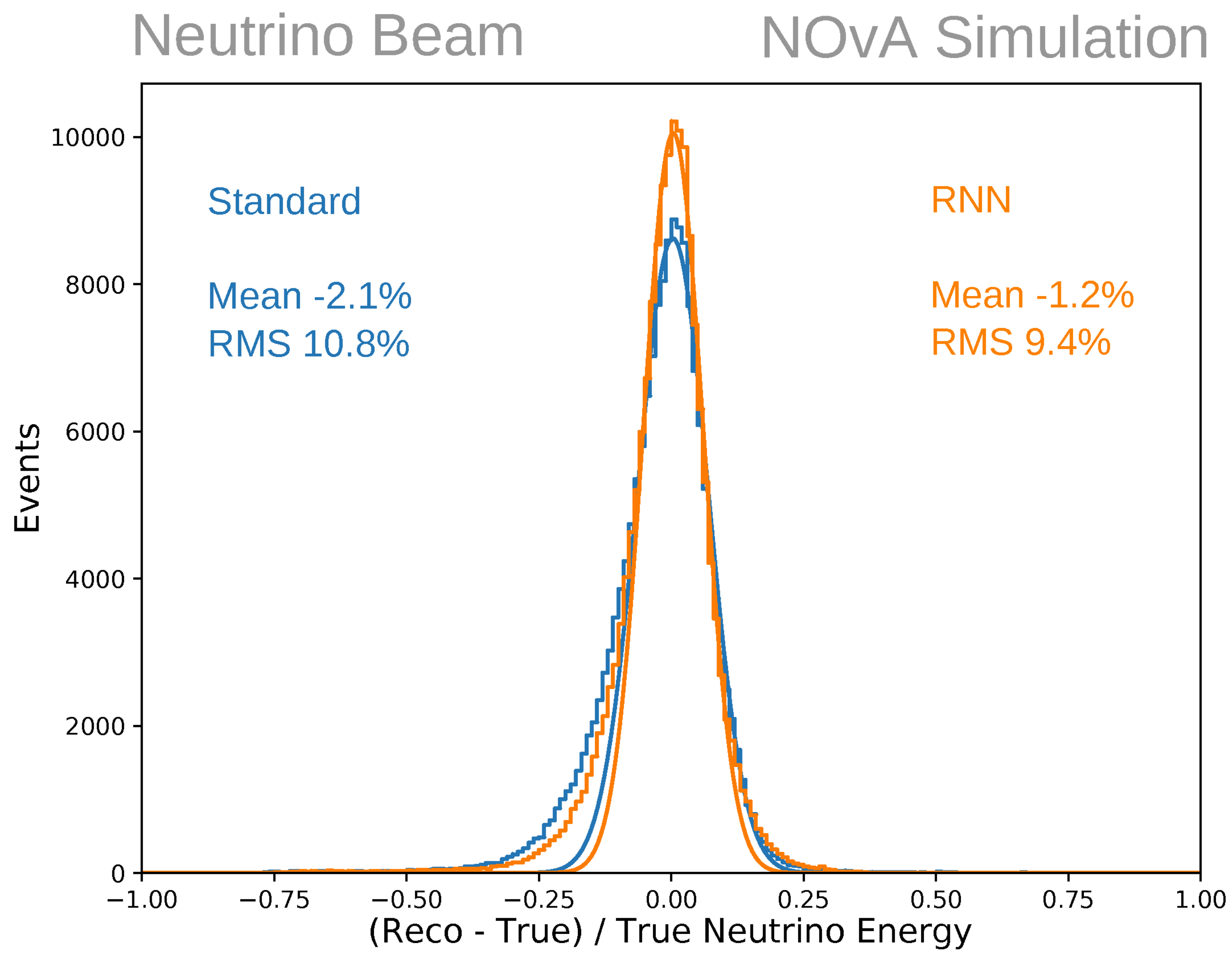


Also trained for electron energy, hadronic energy, ν_μ energy, etc

LSTM for Energy Estimation

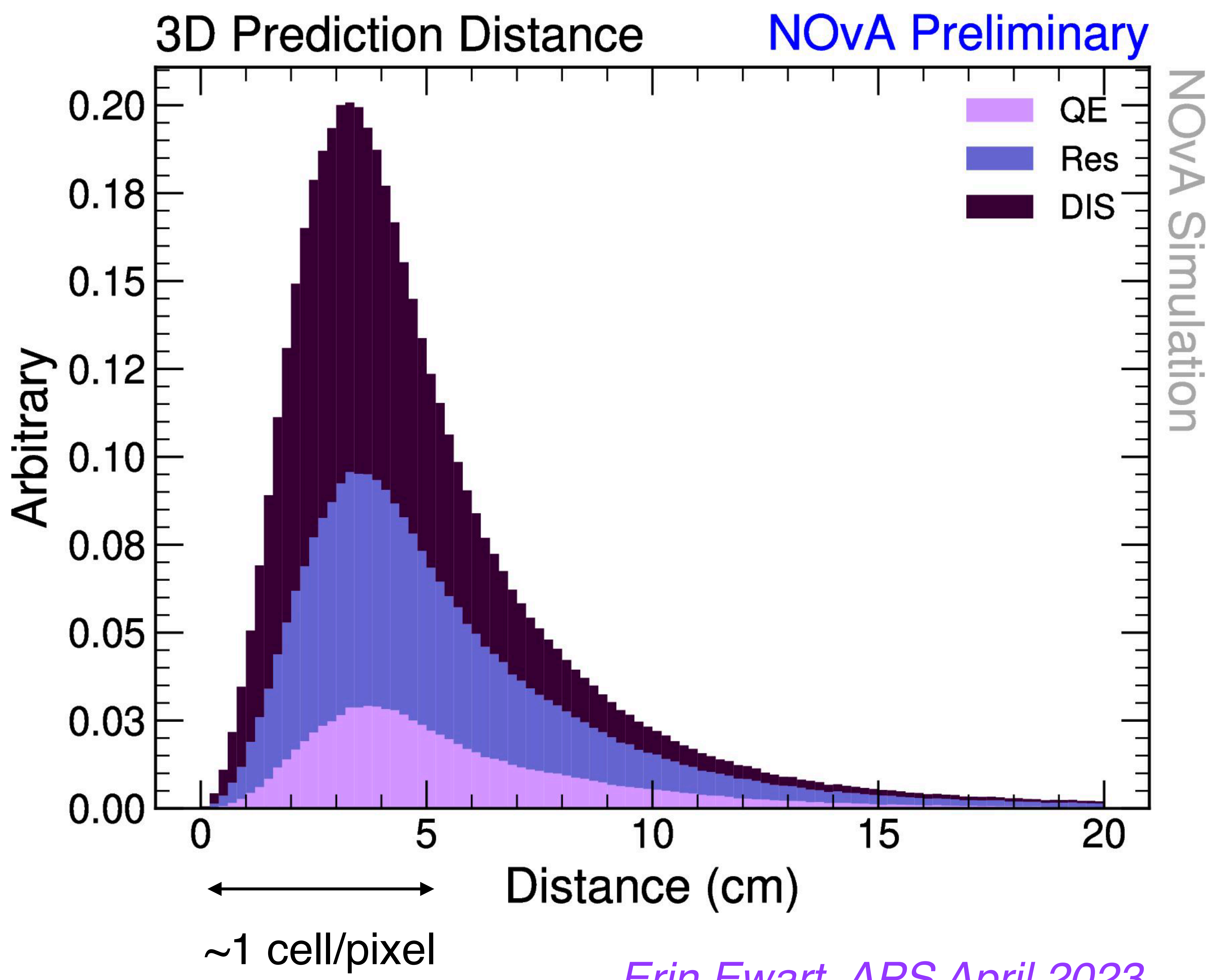
- Long Short-Term Memory (LSTM) is a type of recurrent neural network
- Takes a number of traditional reconstruction quantities as inputs
- Trained with artificially engineered sample to increase network resilience
- Resolution comparable with regression CNN





ML Vertexer (VertexCVN)

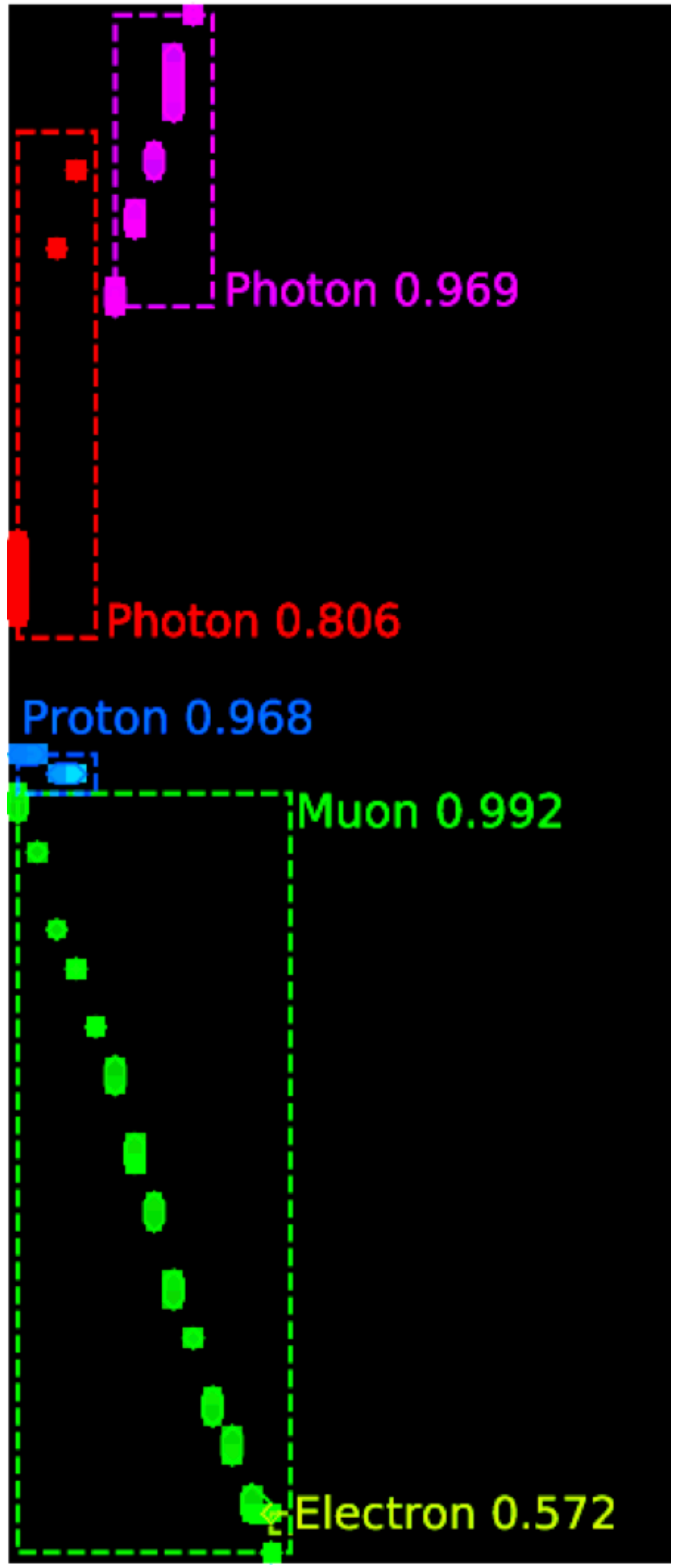
- More accurate vertex finding, means more accurate on
  - Clustering hits to form individual particle tracks/showers
  - Identifying particle types
  - Energy estimation
- Same network architecture as EventCVN (modified MobileNetv2) was explored to predict one 3D vertex
- Shows good performance across interaction types



Erin Ewart, APS April 2023

Full Event Reconstruction with Image Segmentation

- Full event reconstruction on a hit-by-hit basis using instance segmentation:
  - Bounds: Create a bounding box around each particle with a Region-based CNN (RCNN)
  - ID Score: Use a softmax function to classify the particle contained within each box
  - Clusters: Group together hits, identify hits, then individual hits are combined to form clusters
- Very powerful in PID and clustering efficiency
- No dependence on other reconstruction (vertex, etc)
- However, it's quite slow to run on CPUs, and more work needs to be done to run at scale



Summary

- NOvA pioneered the use of CNNs for event classification in HEP and implemented improved networks for recent analyses
- In NOvA, machine learning has been developed to:
  - Identify events and final state particles
  - Reconstruct neutrino energy, final state particle energy, vertex
  - Perform full event reconstruction
- Other ongoing ML efforts in NOvA: Improve ProngCVN with both neutrino and antineutrino sample, Graphical Neural Networks, Unsupervised training
- NOvA has been performing expansive data comparison, impact analysis, uncertainty studies and cross-checks to improve robustness and interpretability of ML techniques

Backup

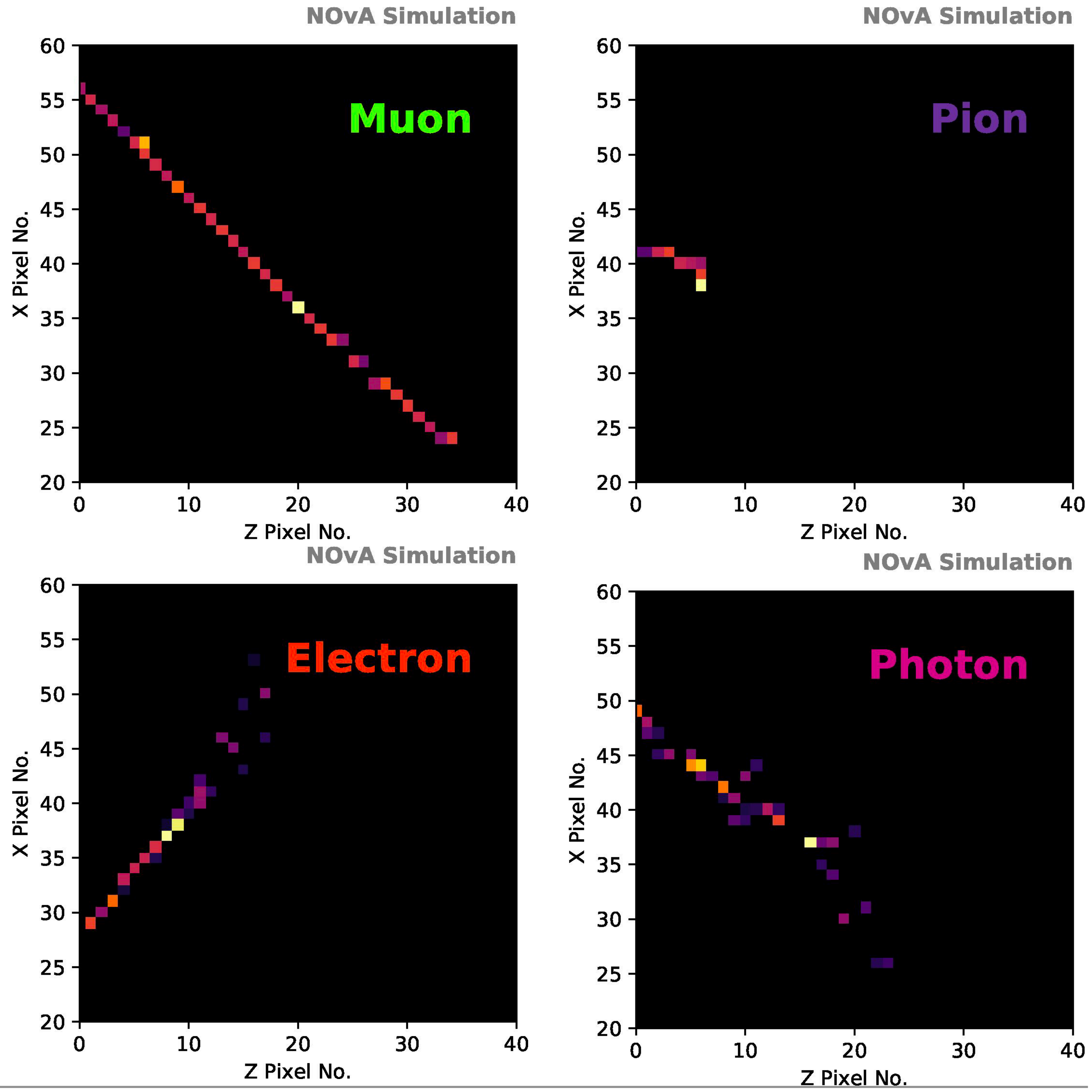
Cosmic filtering with a NN

- Network based on ResNet18 backbone with a siamese structure
  - Takes in two event images (top-view and side-view) as input
- Softmax output with five labels: ν_μ, ν_e, ν_τ, NC, and cosmic score
- Training sample contained 1M+ ν_μ, ν_e, and NC events in both beam modes and 5M+ cosmic events
  - Not trained separately for neutrino/antineutrino mode
- Performs better than traditional cosmic rejection in all samples

Data Sample	Traditional Cosmic Rejection	Cosmic Rejection Neural Network
ν_e	93.21	99.71
$\bar{\nu}_e$	92.81	99.82
ν_μ	93.22	99.20
$\overline{\nu_\mu}$	92.82	99.20
ν NC	93.24	97.08
$\bar{\nu}$ NC	92.79	96.82
Cosmic ν	7.80	5.00

Single particle ID

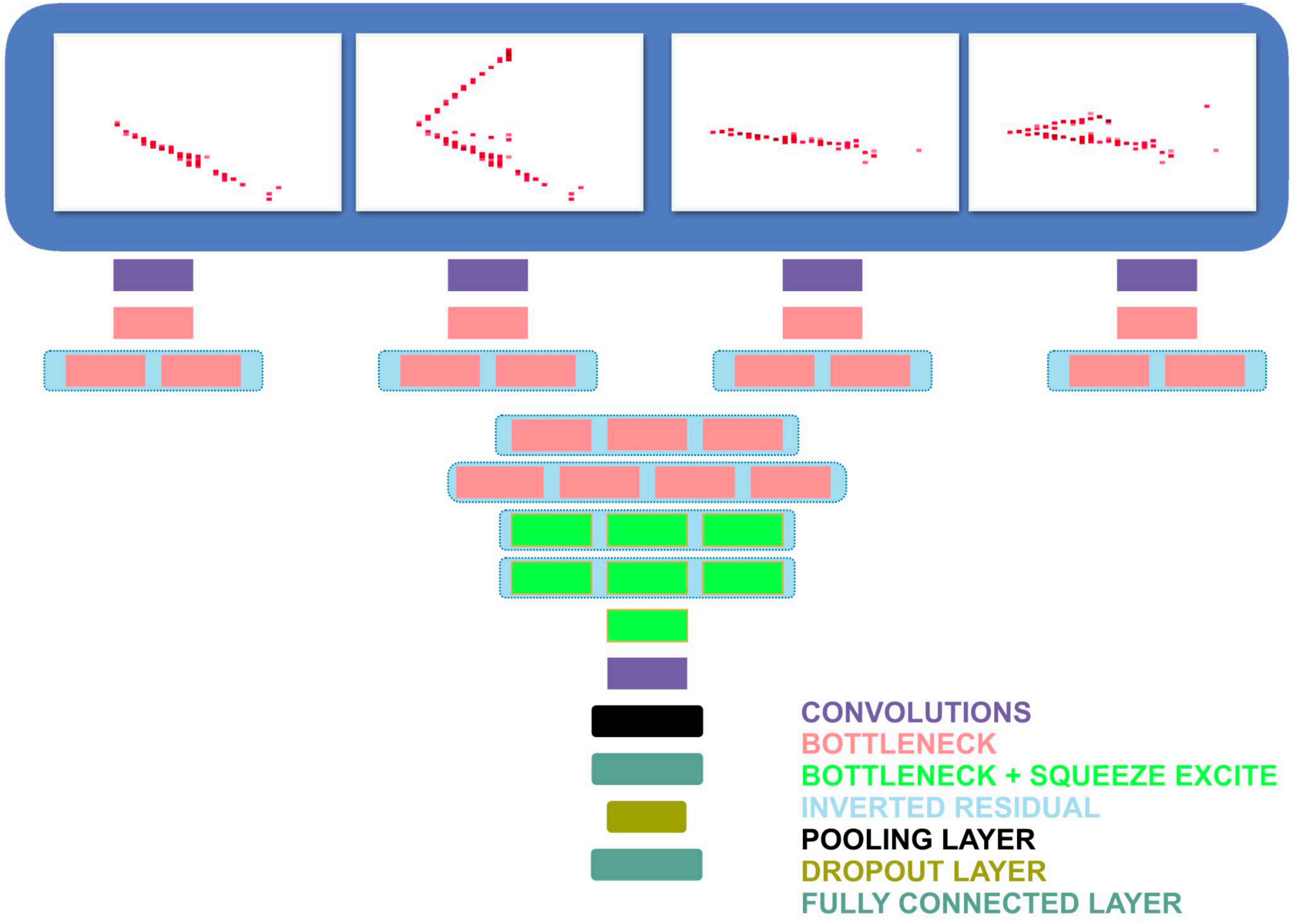
- NOvA also has trained a network using singularly simulated particles for ND analyses → no contextual information
- Also developing a network designed for neutron identification using these samples



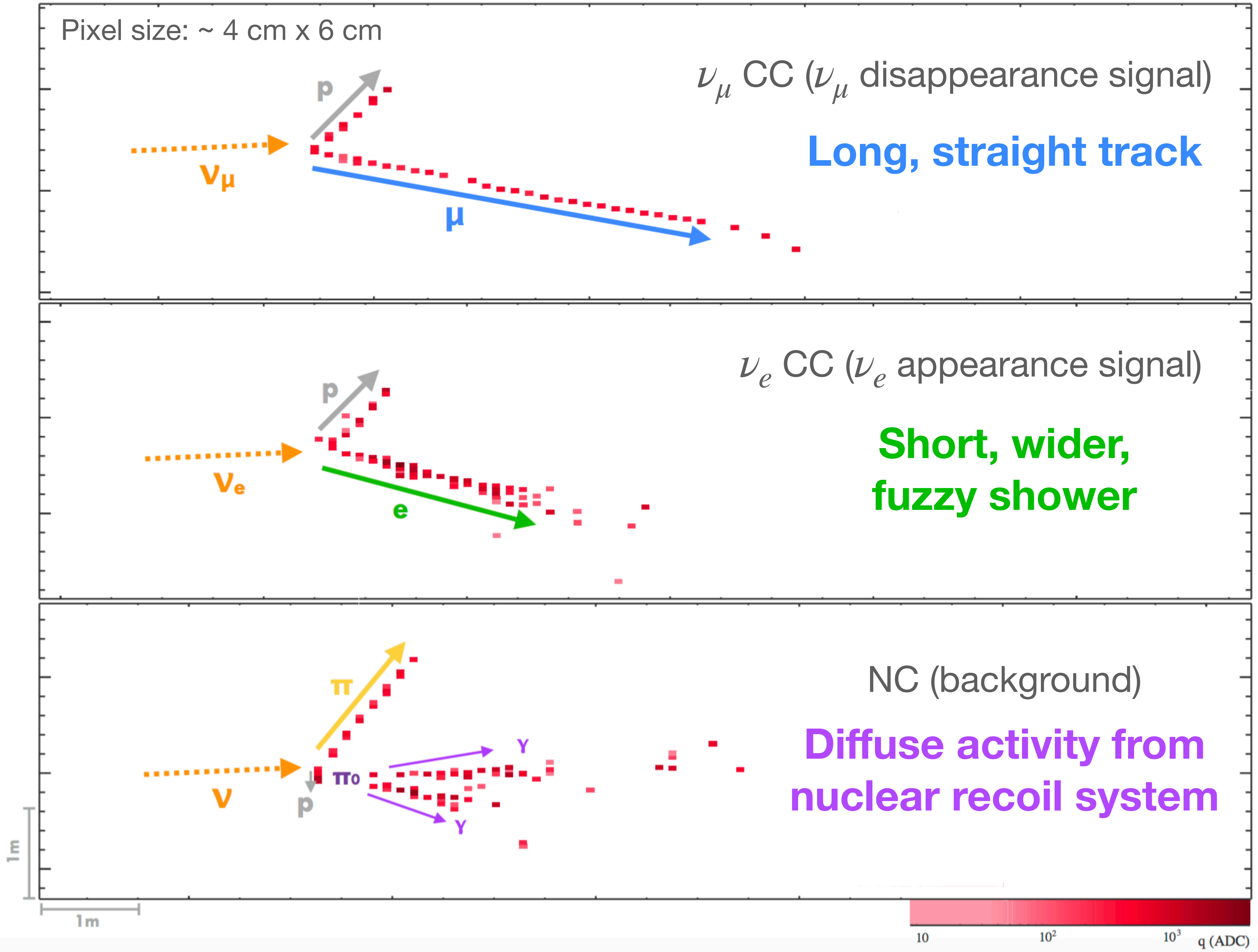
Improved ProngCVN

Akshay Chatla, DAE 2022

- Modifies ProngCVN (modified MobileNetv2) architecture by adding Squeeze-Excite block for channel attention
- Trained on a combined sample of neutrino and antineutrino mode
- Shows good performance for particle classification



Event Topology



J. Inst. **11**, P09001 (2016)