



Machine Learning in NOvA and DUNE Neutrino Experiments

Wenjie Wu (吴文杰)

UC Irvine

2024/12/13

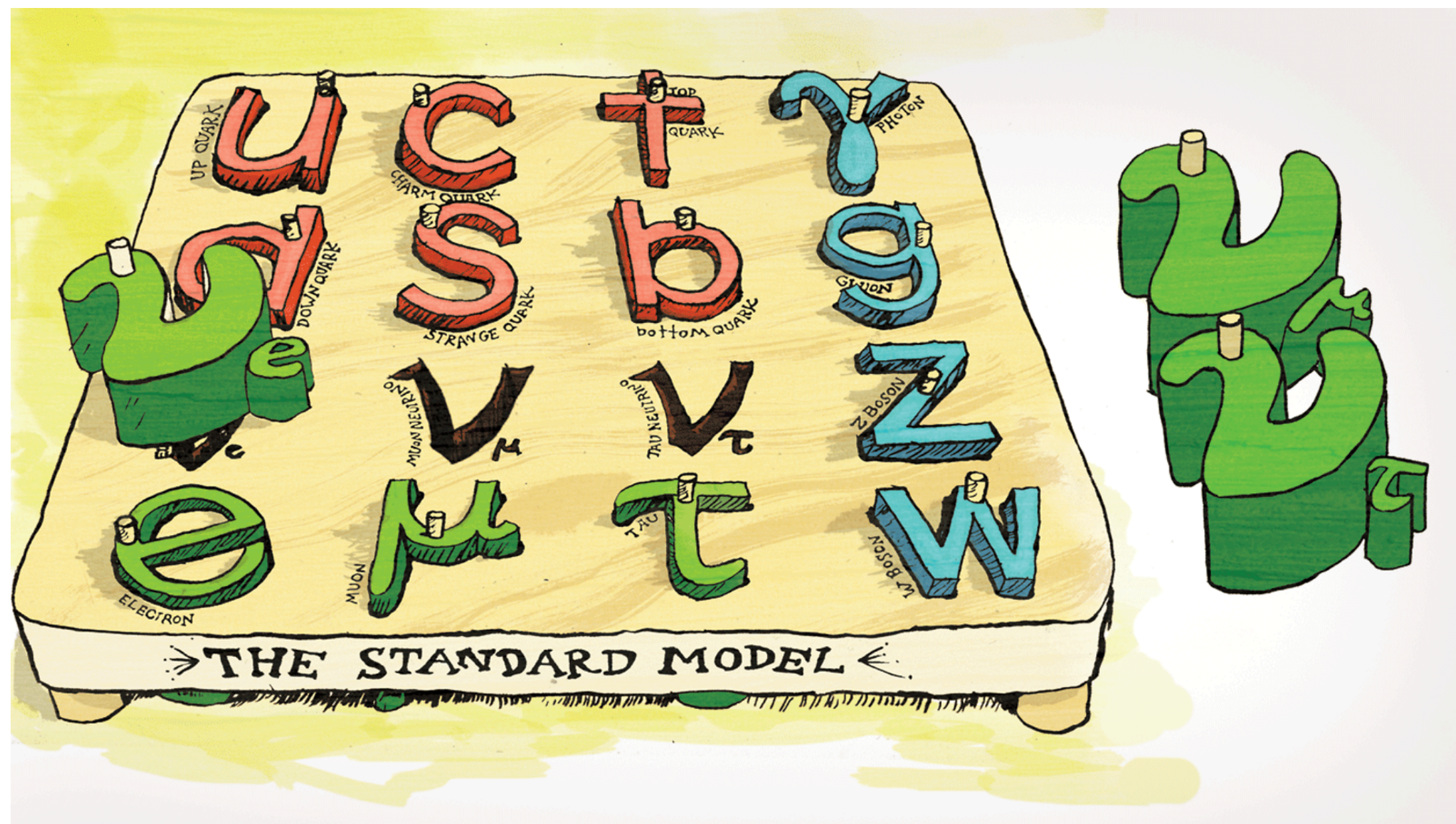
UCI

Neutrinos

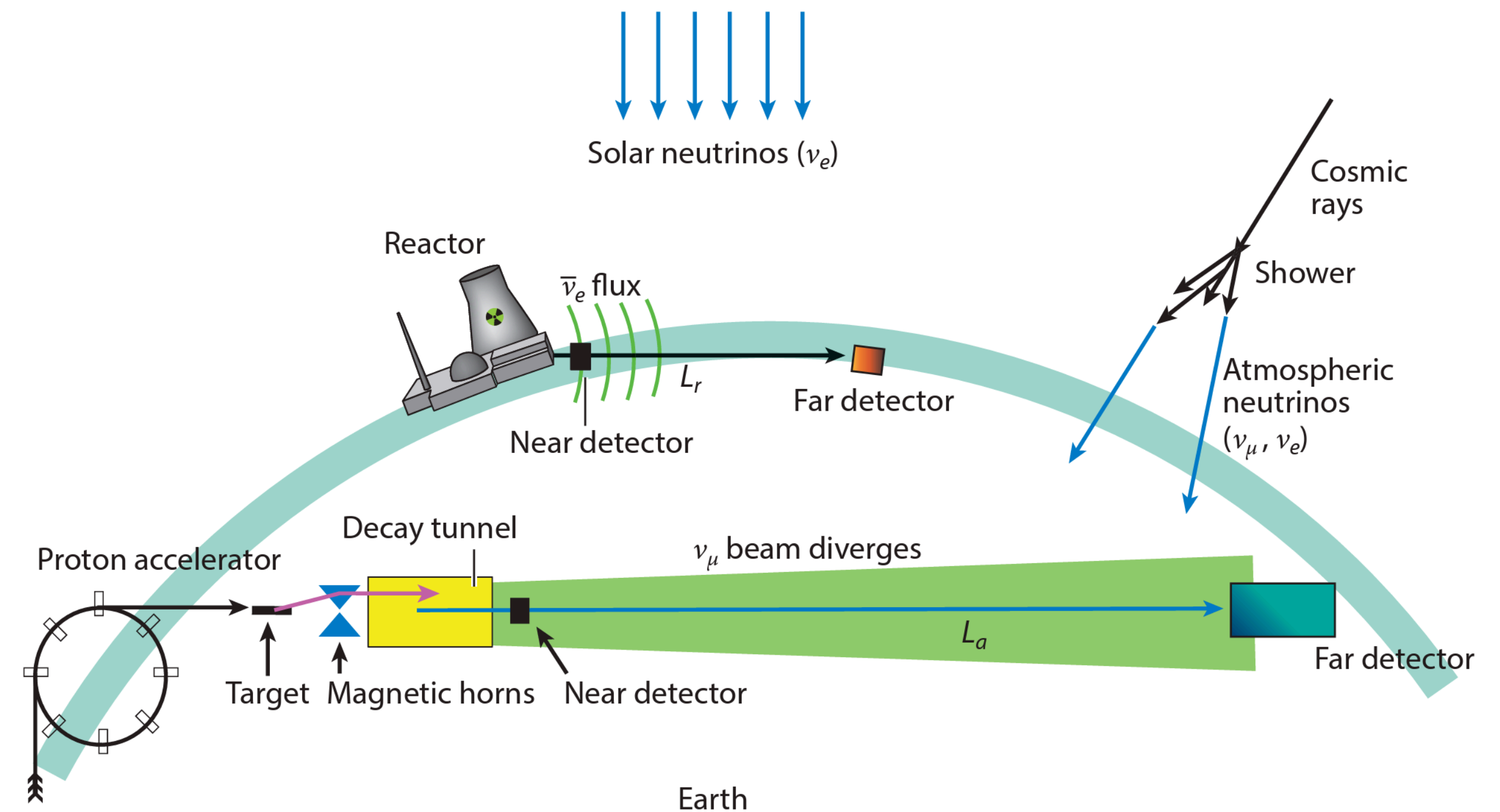
- Neutrinos are elementary particles with little mass and no charge.
- Very abundant, many sources can generate neutrinos.

- Feebly interacting particles, need gigantic detectors to find them.

Reactor, Solar, Atmospheric, **Accelerator**



Credit: [Symmetry](#)



[Annu.Rev.Nucl.Part.Sci.2016.66:47-71](#)

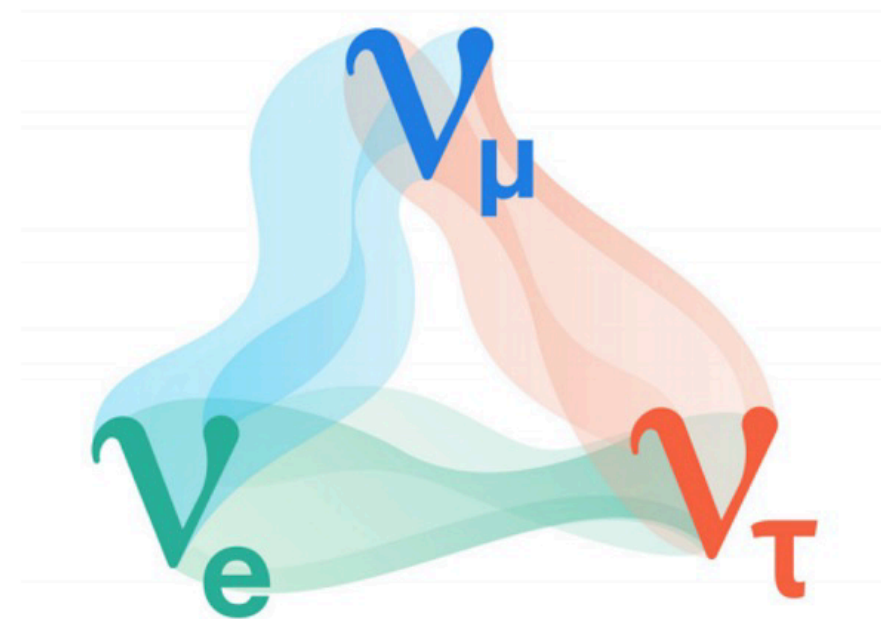
Neutrino oscillation

One neutrino flavor can oscillate into another as a result of quantum mechanical processes

Neutrino oscillation



Neutrinos have non-zero mass

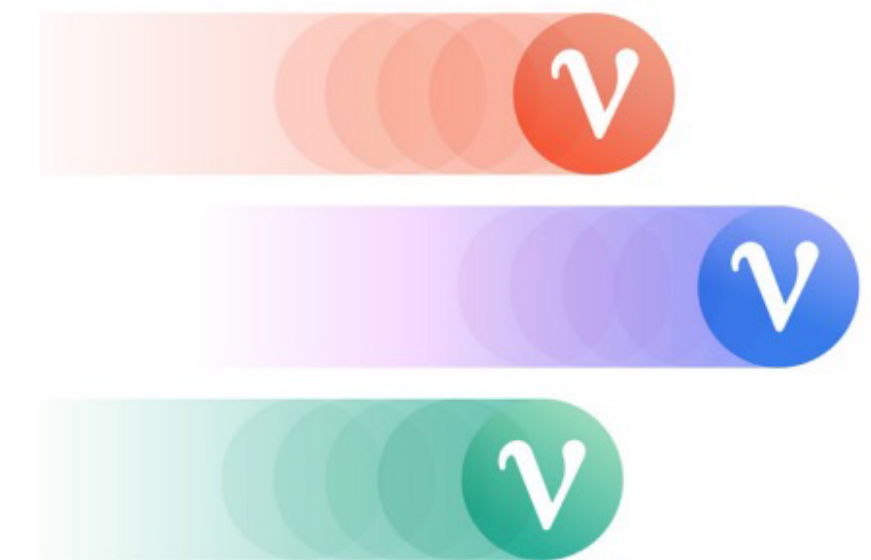


$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{\text{PMNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

Mixing Matrix

Flavor eigenstates

Mass eigenstates



$$U_{\text{PMNS}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

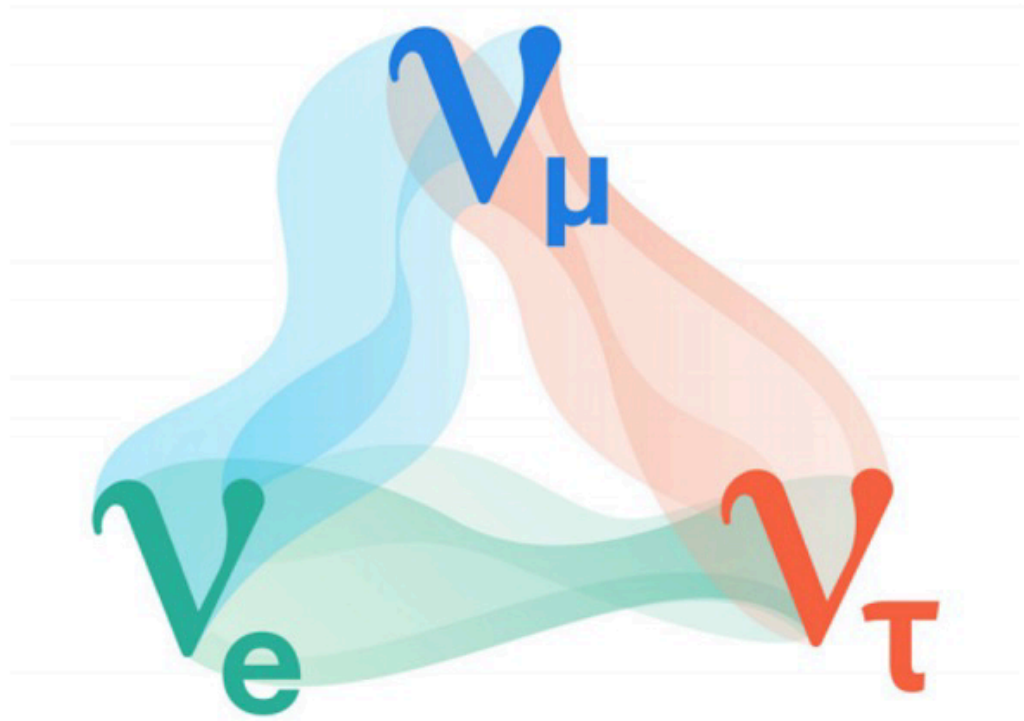
$c_{ij} = \cos \theta_{ij}$
 $s_{ij} = \sin \theta_{ij}$
 $\Delta m_{ij}^2 = m_i^2 - m_j^2$

$$P(\nu_\alpha \rightarrow \nu_\beta) \sim \sin^2(2\theta) \sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right)$$

The probability depends on
L: travel distance (baseline)
E: neutrino energy

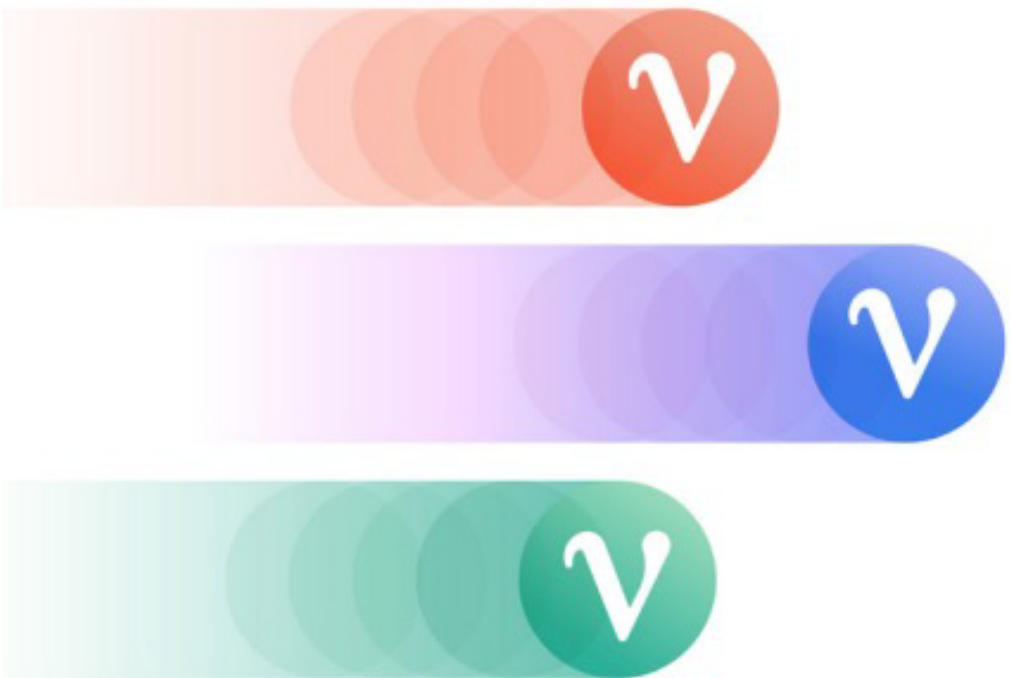
Oscillation parameters:
2 mass square differences
3 mixing angles
1 CP violation phase

Neutrino oscillation



$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{\text{PMNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

Mixing Matrix



Flavor Eigenstates

Mass Eigenstates

$$U_{\text{PMNS}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Atmospheric Neutrino



Accelerator Neutrino

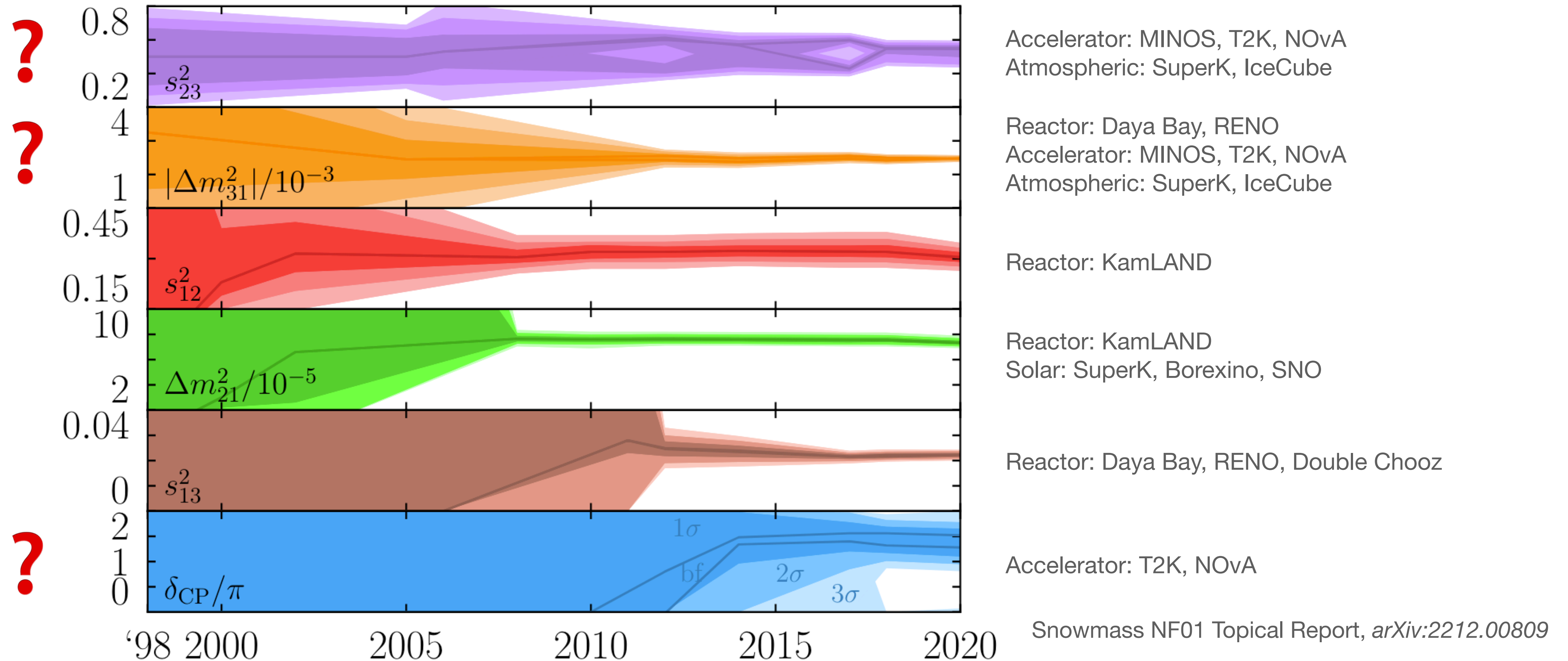


Reactor Neutrino



Solar Neutrino

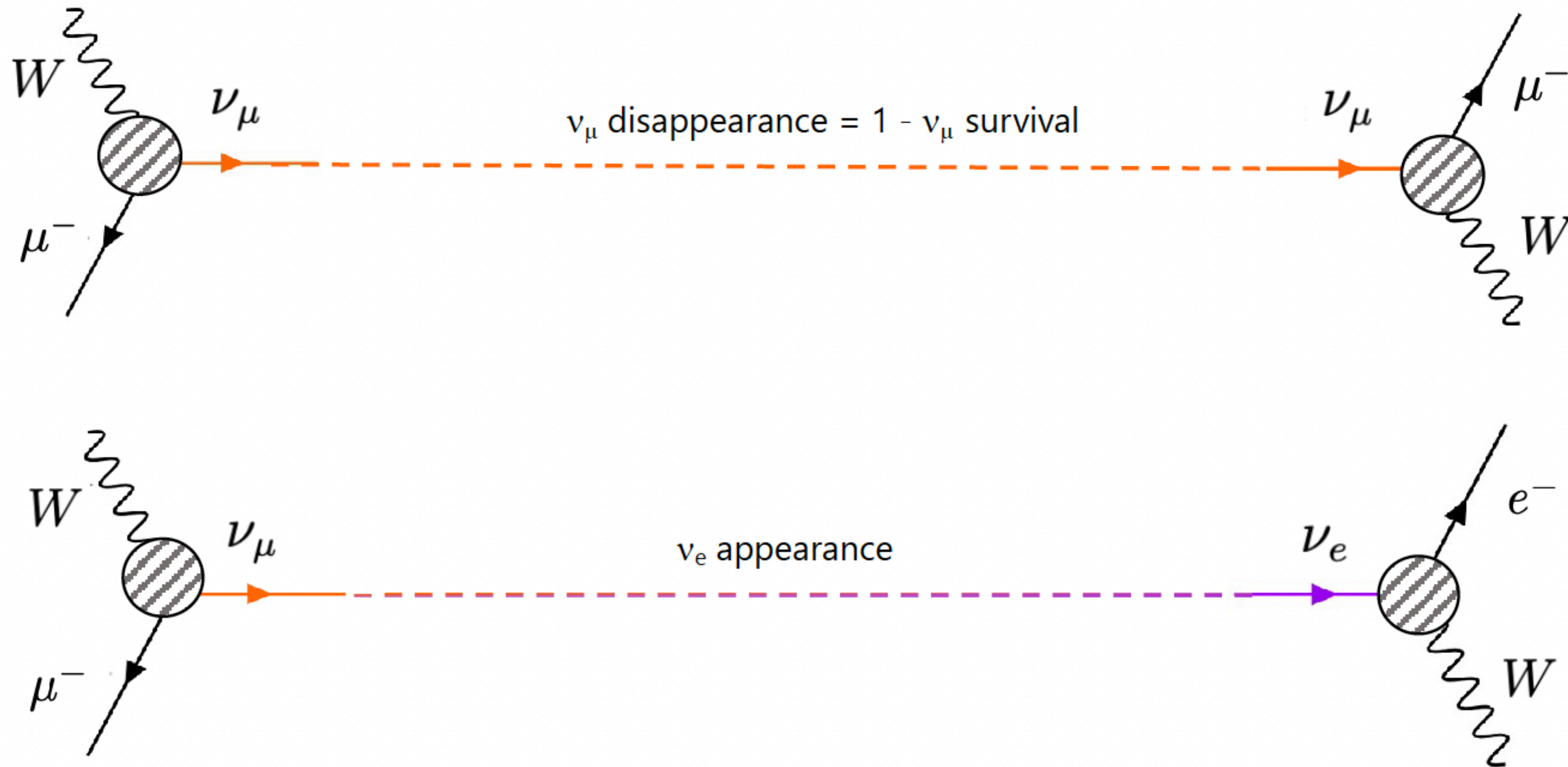
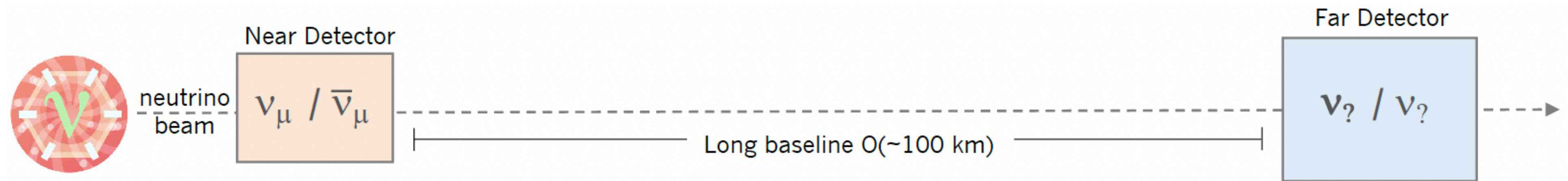
Neutrino oscillation



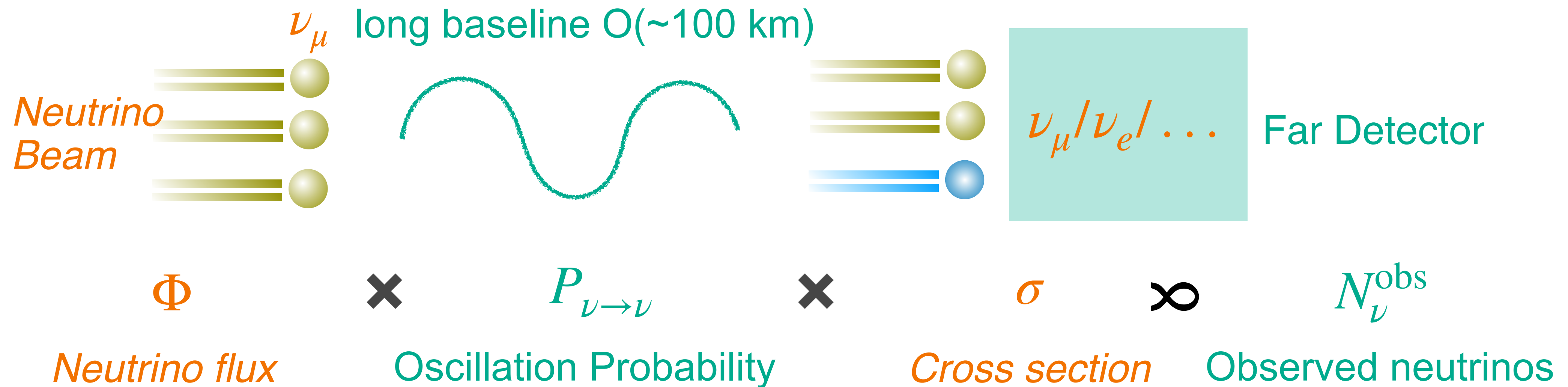
Future neutrino oscillation experiments will measure the **octant of θ_{23}** , the **CP violation phase angle δ** , and the **neutrino mass ordering**

Accelerator-based long-baseline oscillation experiments provide unique handle for CP measurement

Long-baseline experiments



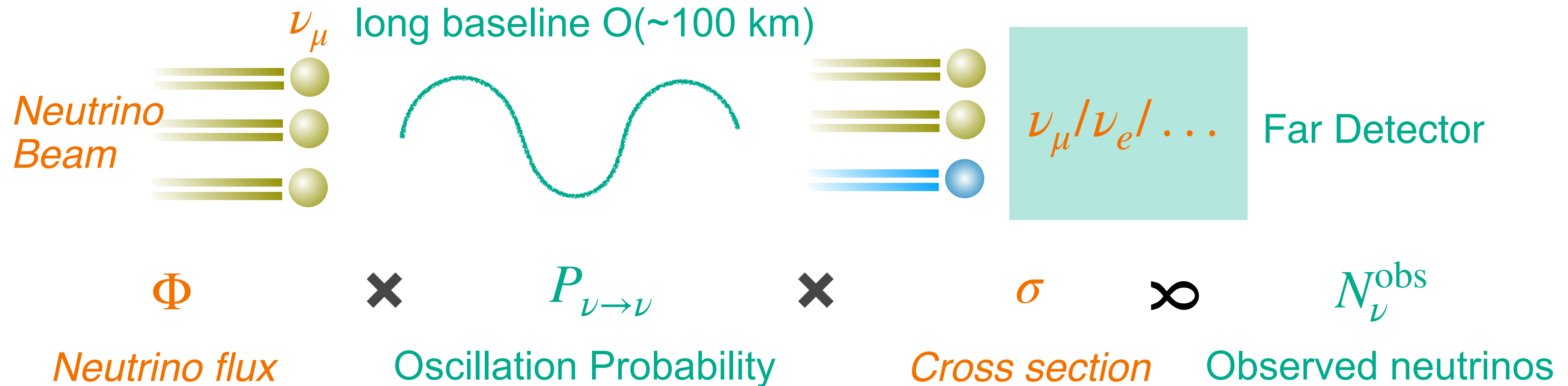
Long-baseline experiments



Oscillation parameters are inferred from observed event (N_ν^{obs})

- ➡ Powerful neutrino beam gives high statistics (Φ)
- ➡ Highly-capable near detector to constrain neutrino flux and cross-section uncertainties (σ)
- ➡ High-resolution detectors reduce ambiguities due to detector response effects

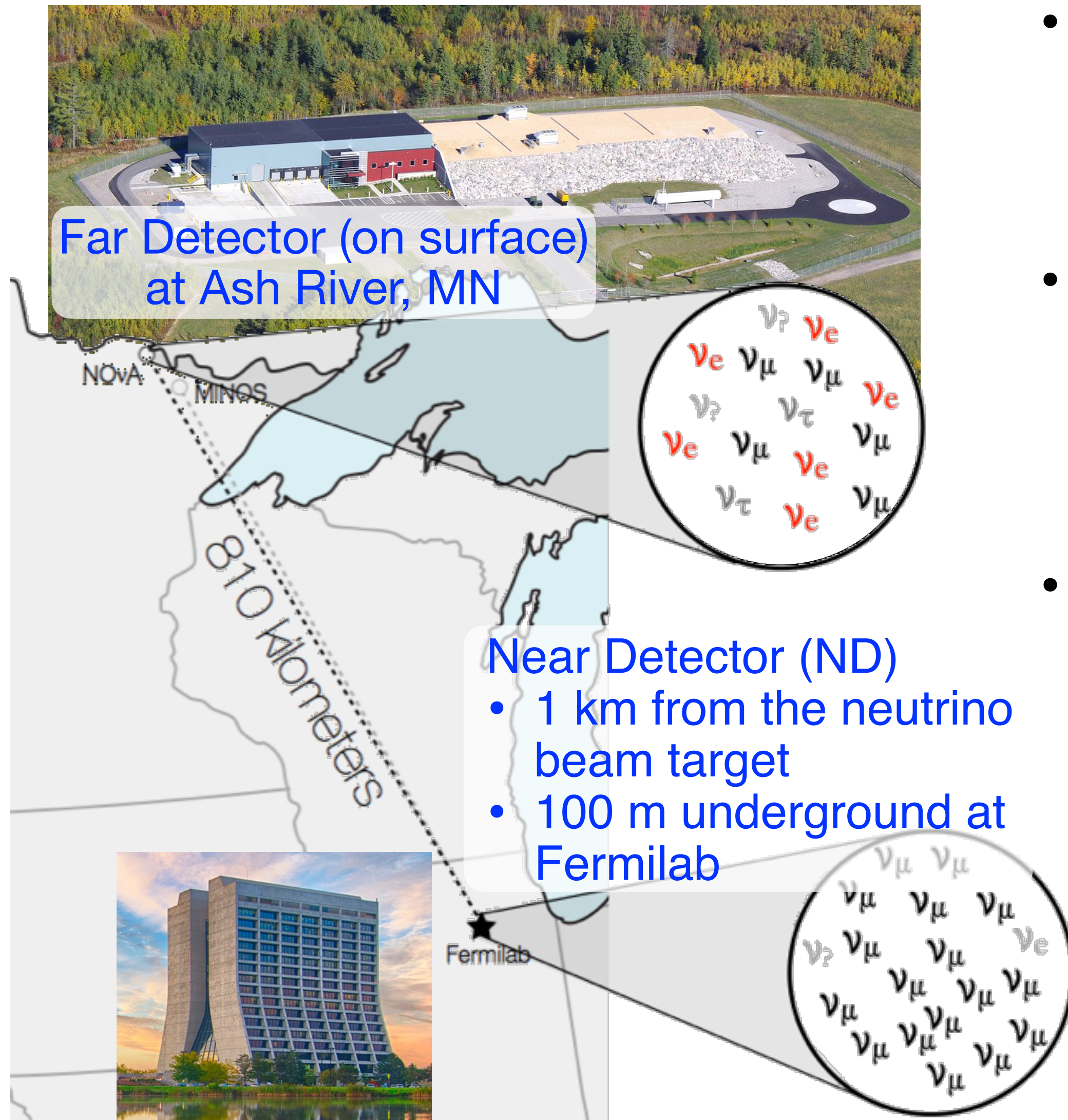
Long-baseline experiments



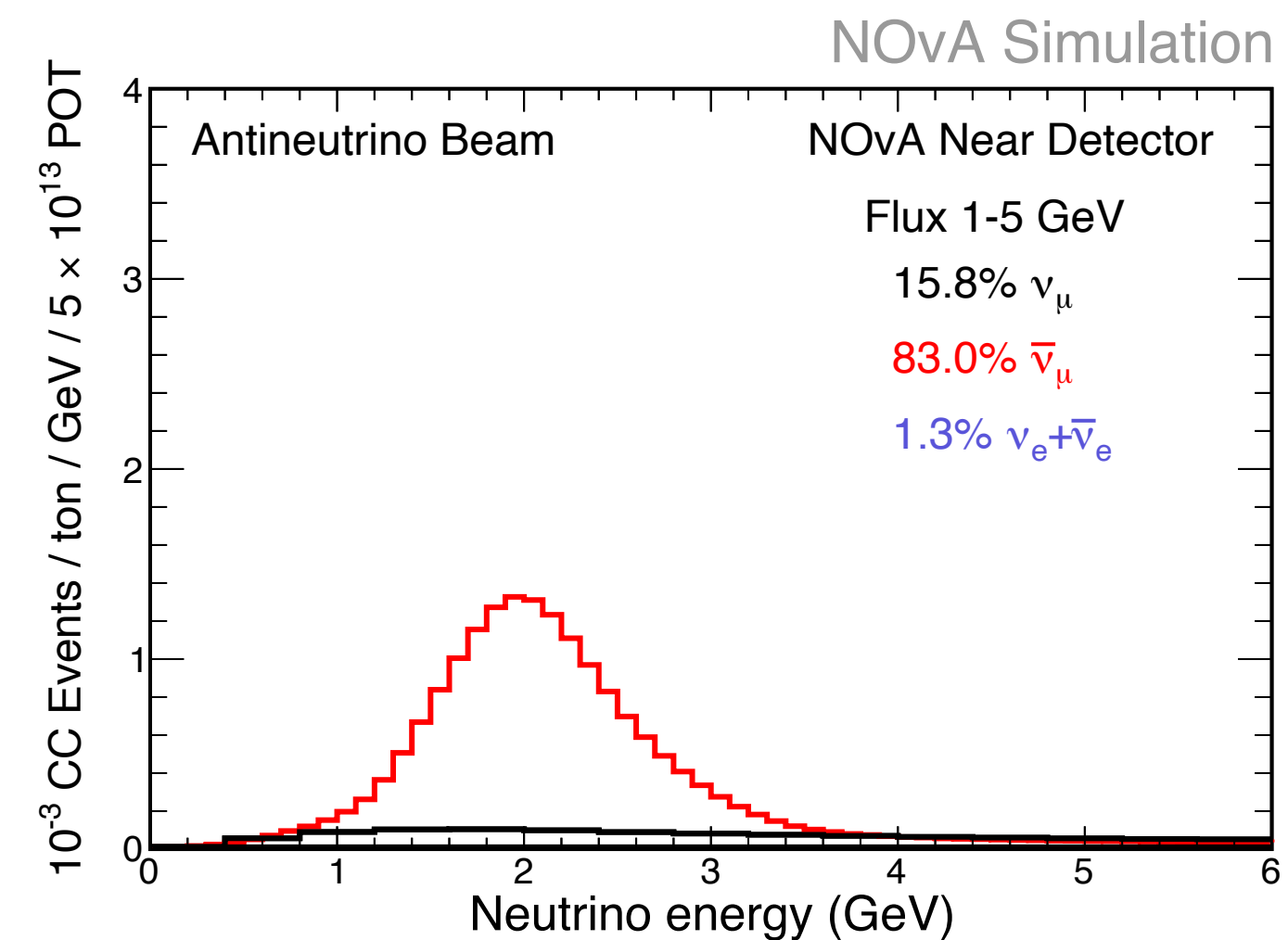
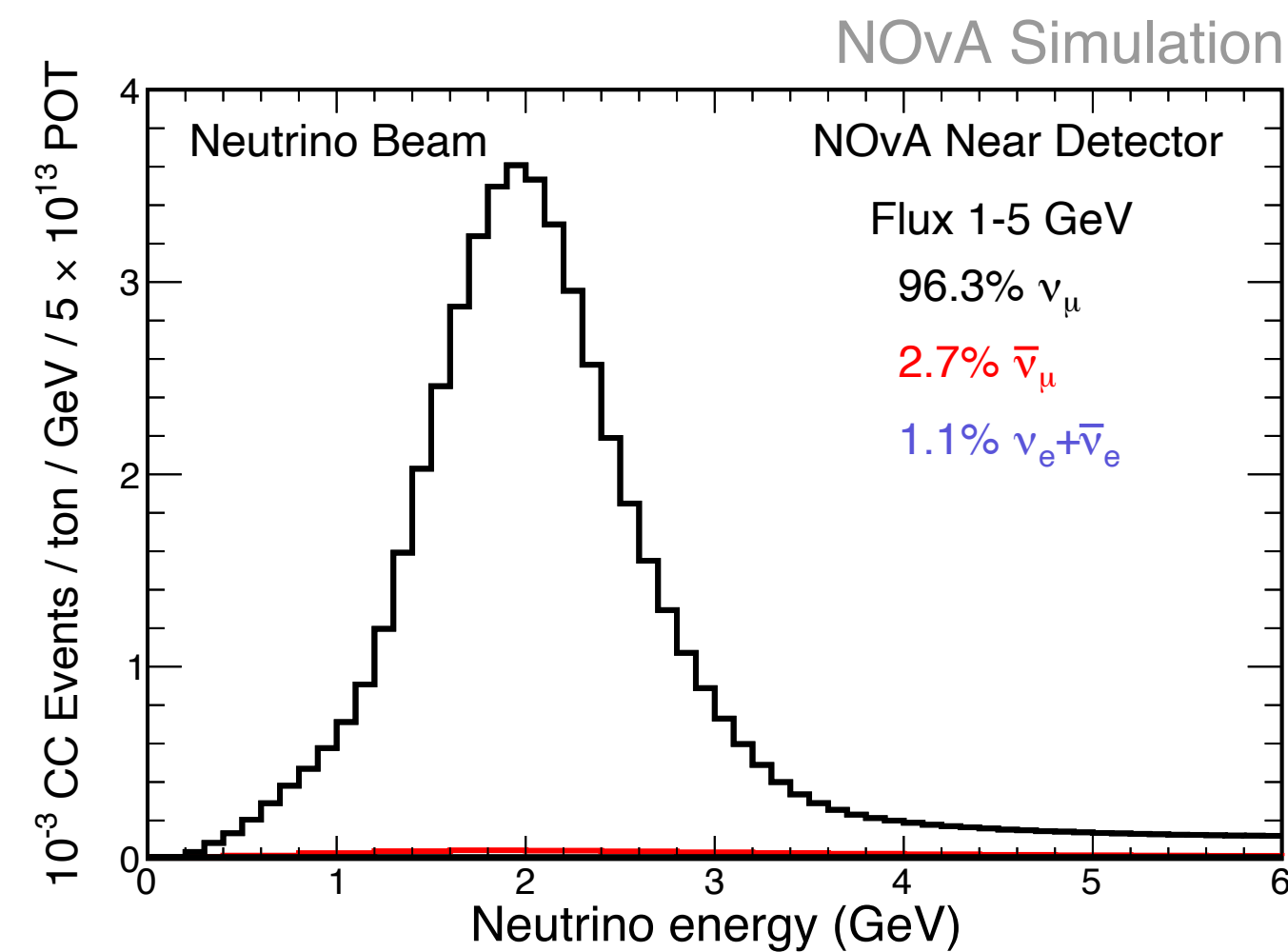
Oscillation parameters are inferred from observed event (N_ν^{obs})

- ➡ Powerful neutrino beam gives high statistics (Φ)
- ➡ Highly-capable near detector to constrain neutrino flux and cross-section uncertainties (σ)
- ➡ High-resolution detectors reduce ambiguities due to detector response effects
- ➡ **Use advanced ML techniques to fully explore detector's potential and improve sensitivity**

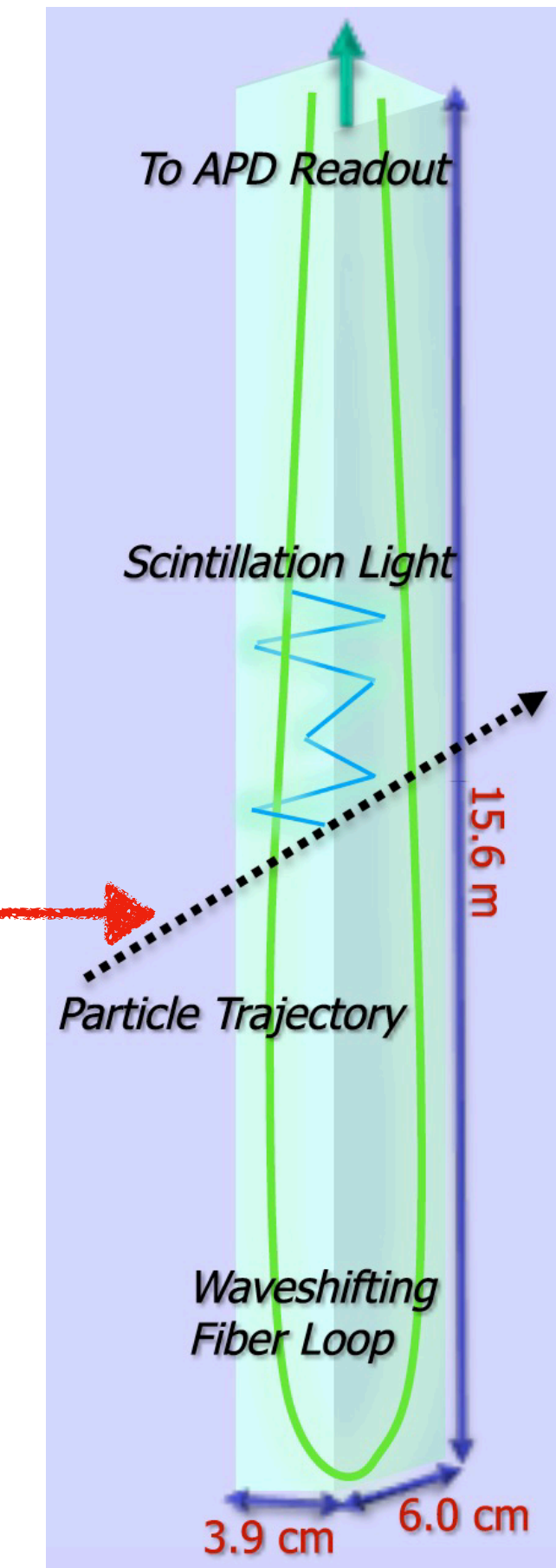
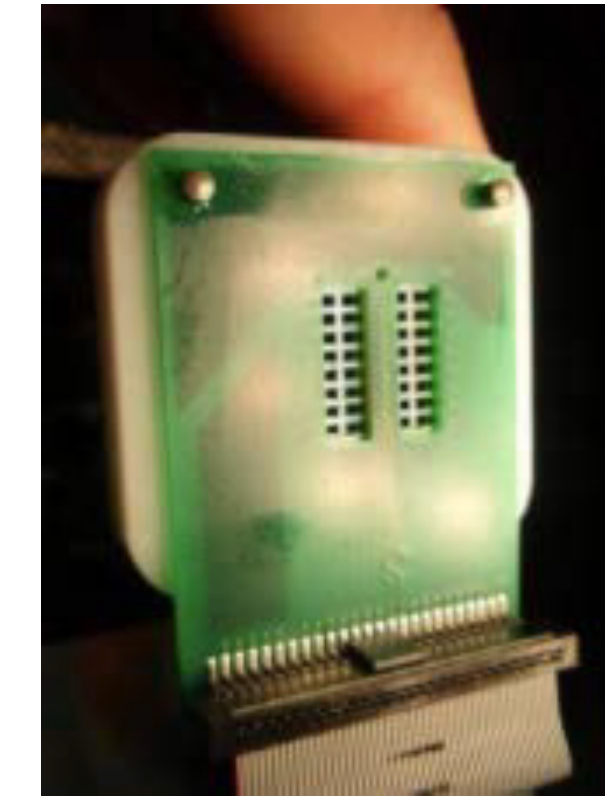
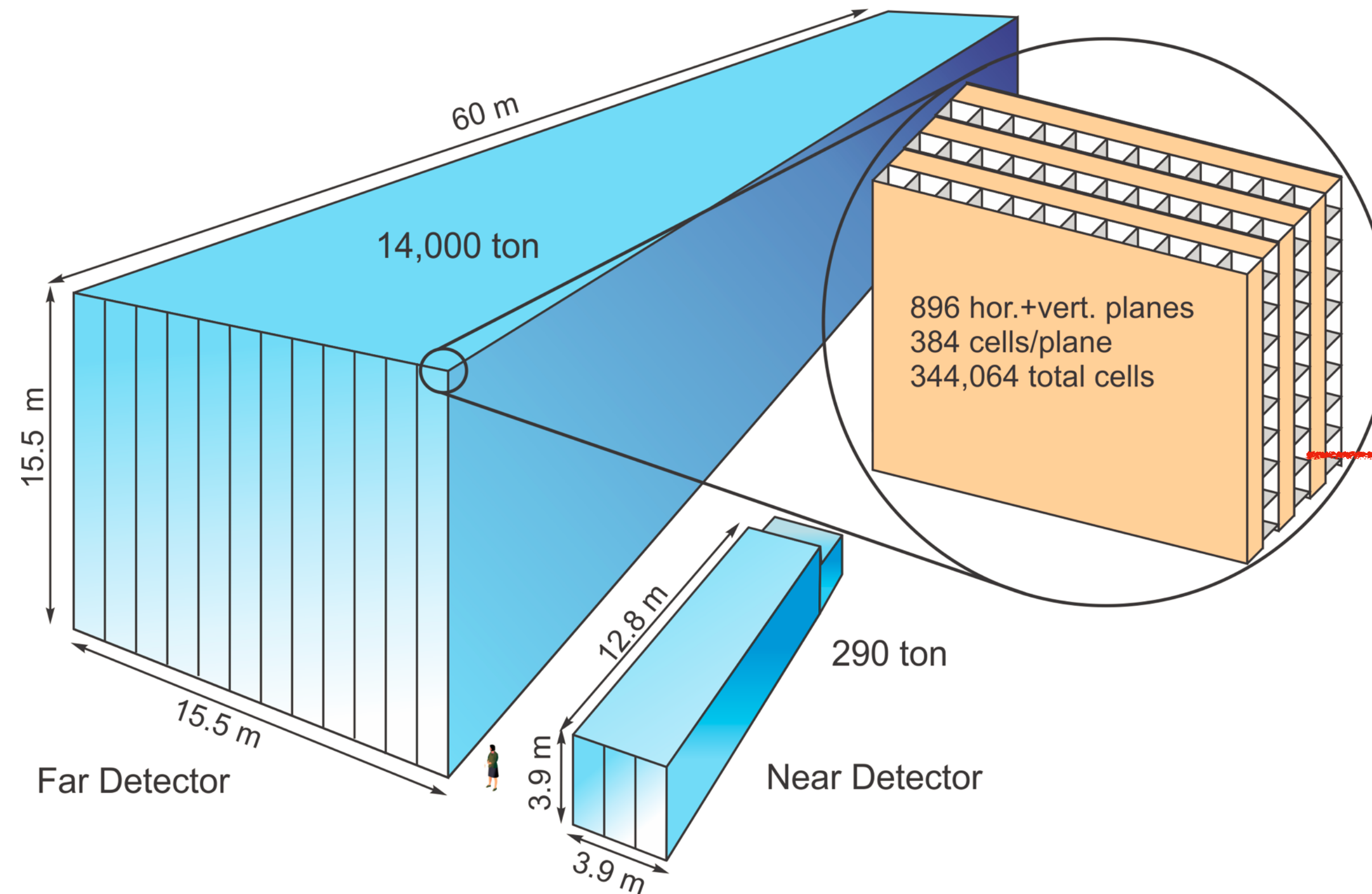
NOvA: NuMI Off-Axis ν_e Appearance Experiment



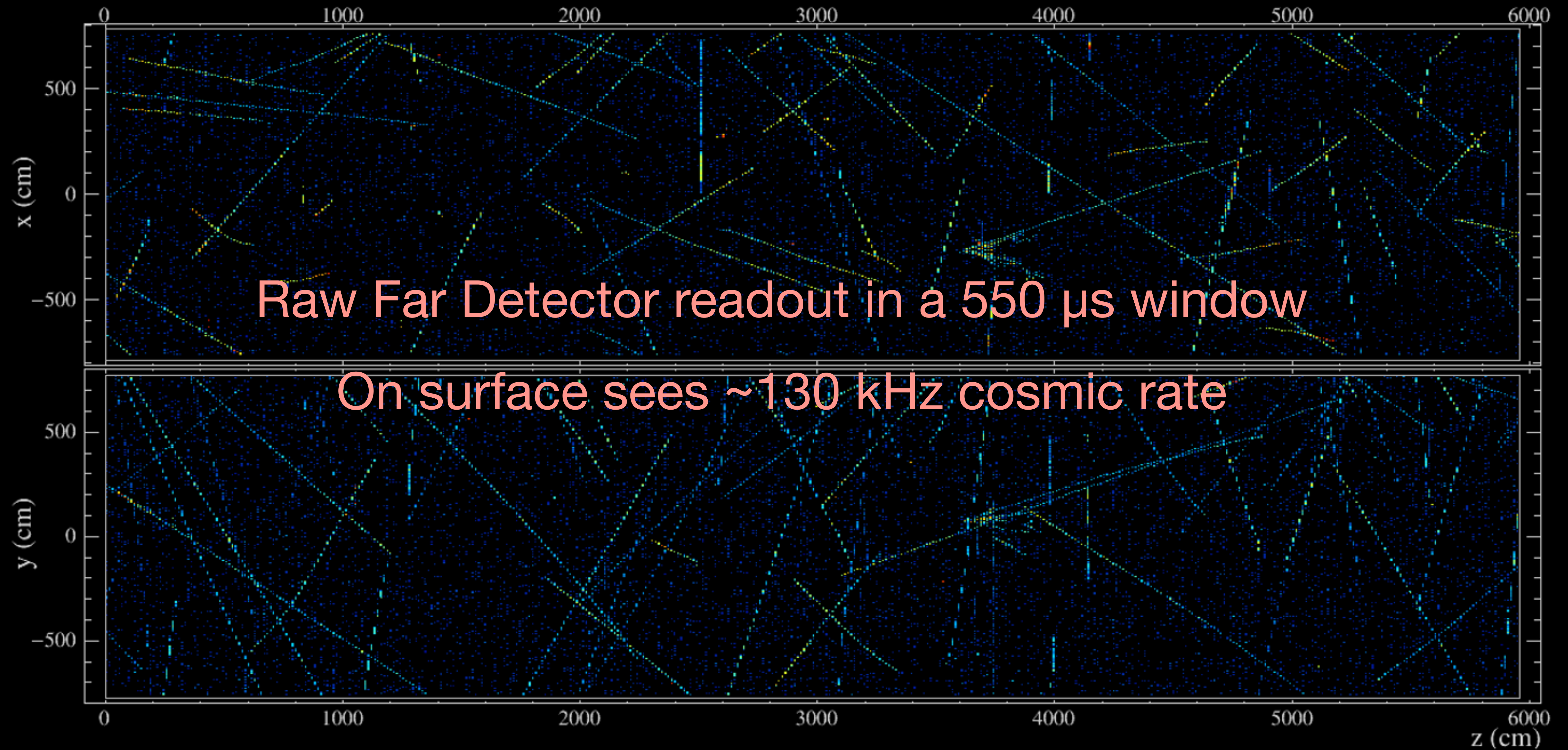
- NOvA is an accelerator-based neutrino experiment
 - Longest baseline in operation (810 km), large matter effect, sensitive to mass ordering
- Muon neutrino beam (NuMI) at Fermilab
 - Two configurations: neutrino mode and antineutrino mode
 - Power record 1.018 MW in 2024
- ~14 mrad off-axis, narrow-band 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



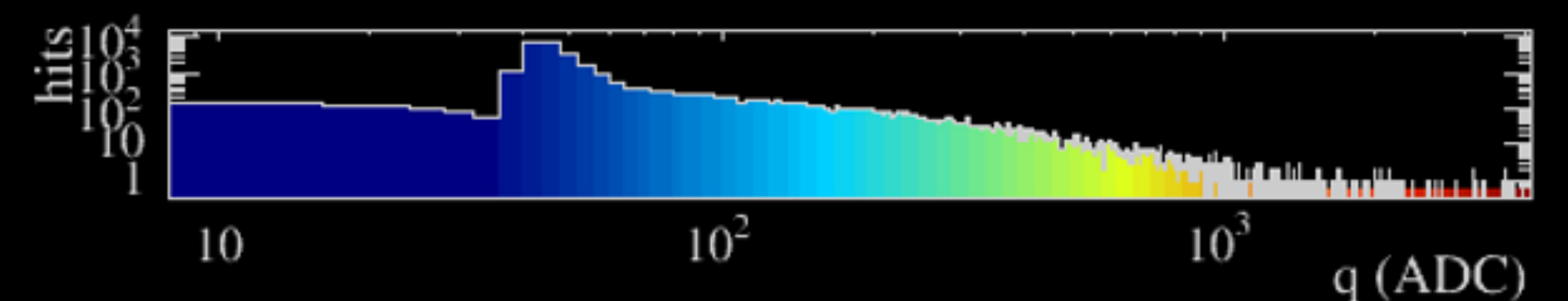
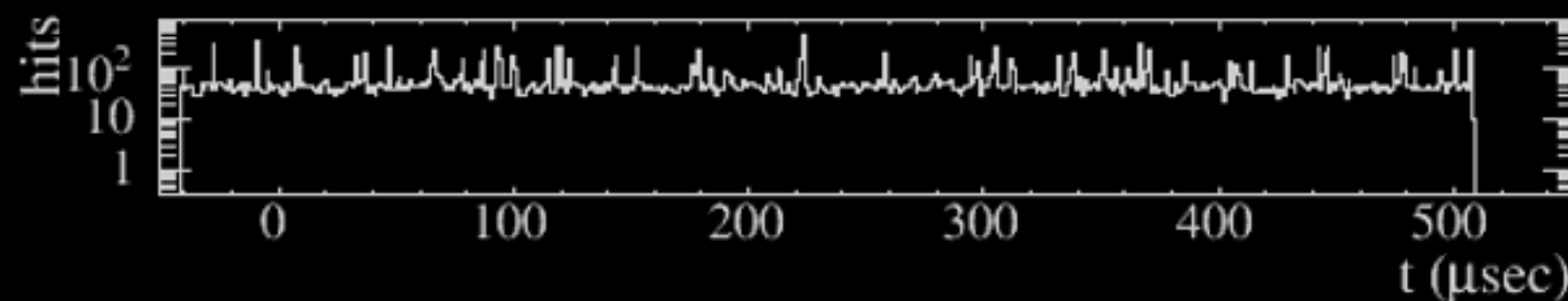
NOvA - FNAL E929

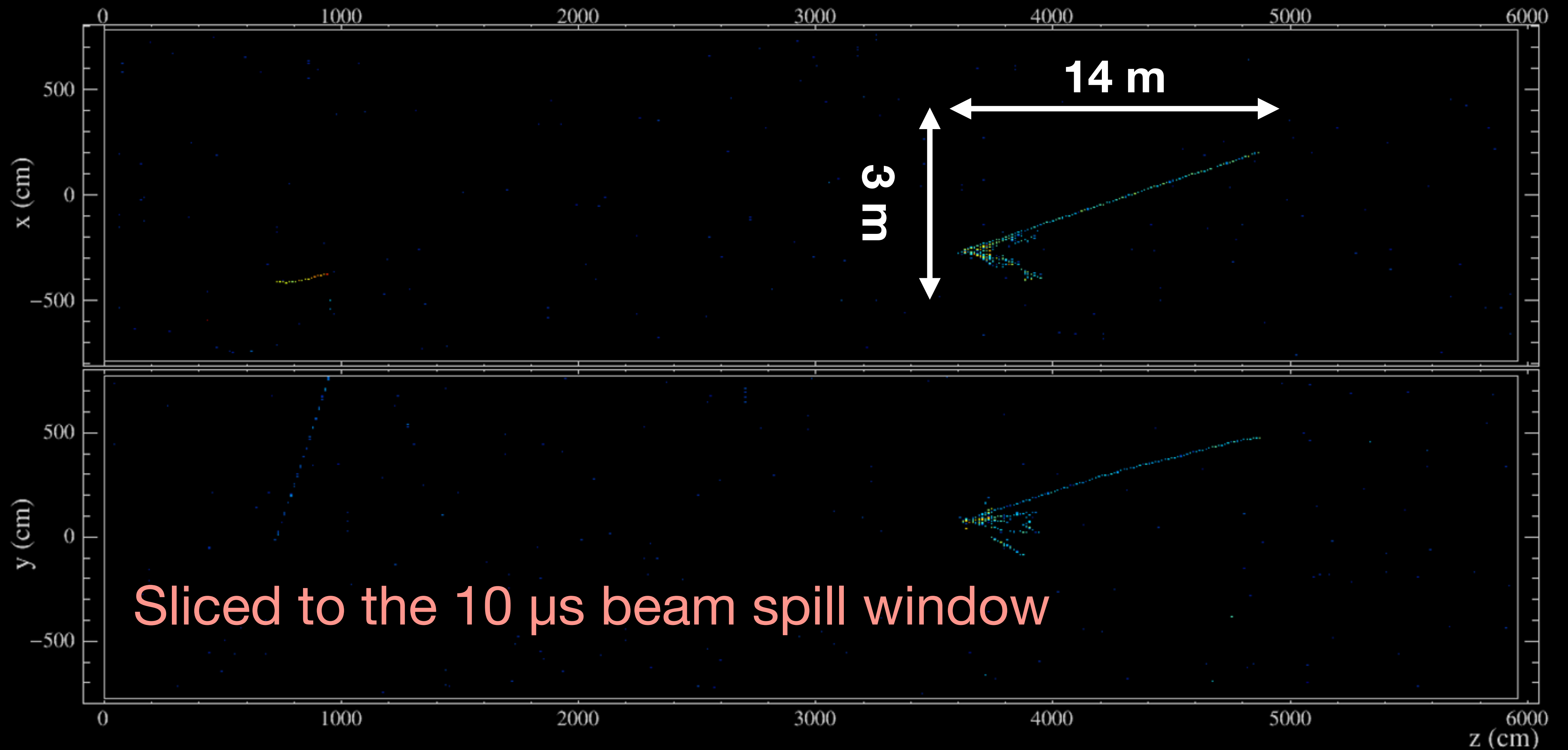
Run: 18620 / 13

Event: 178402 / --

UTC Fri Jan 9, 2015

00:13:53.087341608





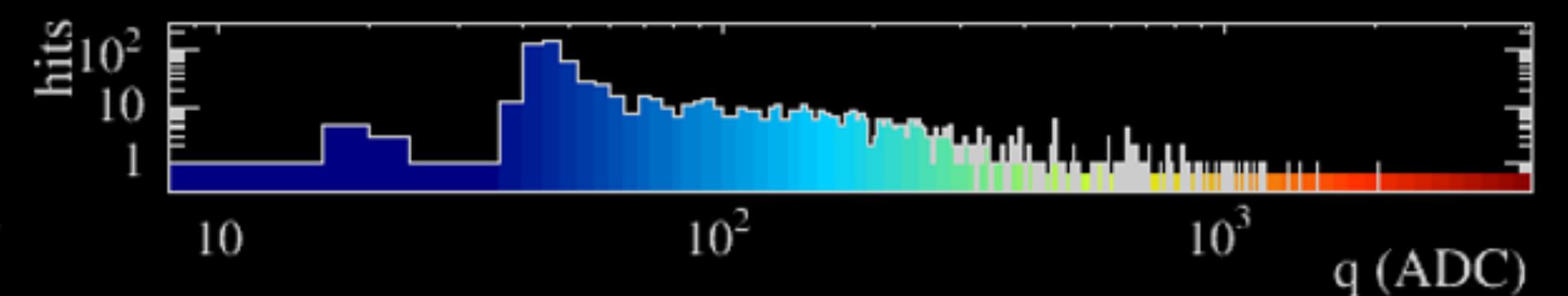
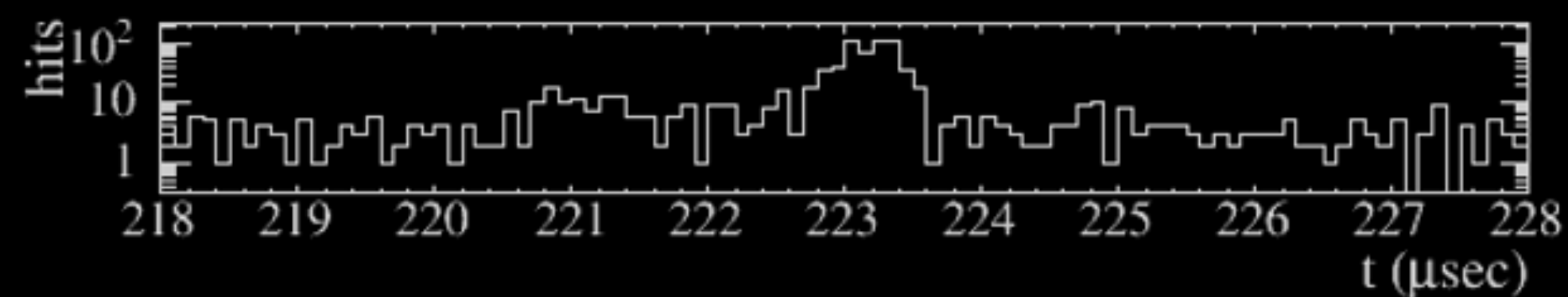
NOvA - FNAL E929

Run: 18620 / 13

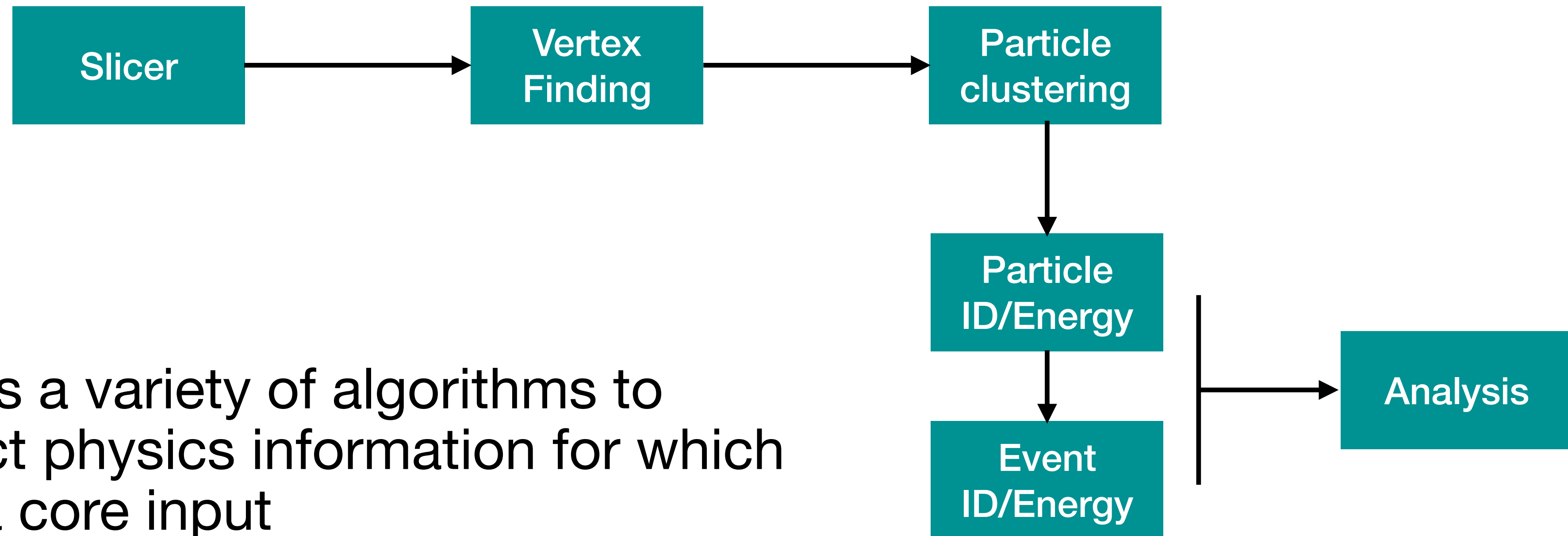
Event: 178402 / --

UTC Fri Jan 9, 2015

00:13:53.087341608



Event Reconstruction in NOvA

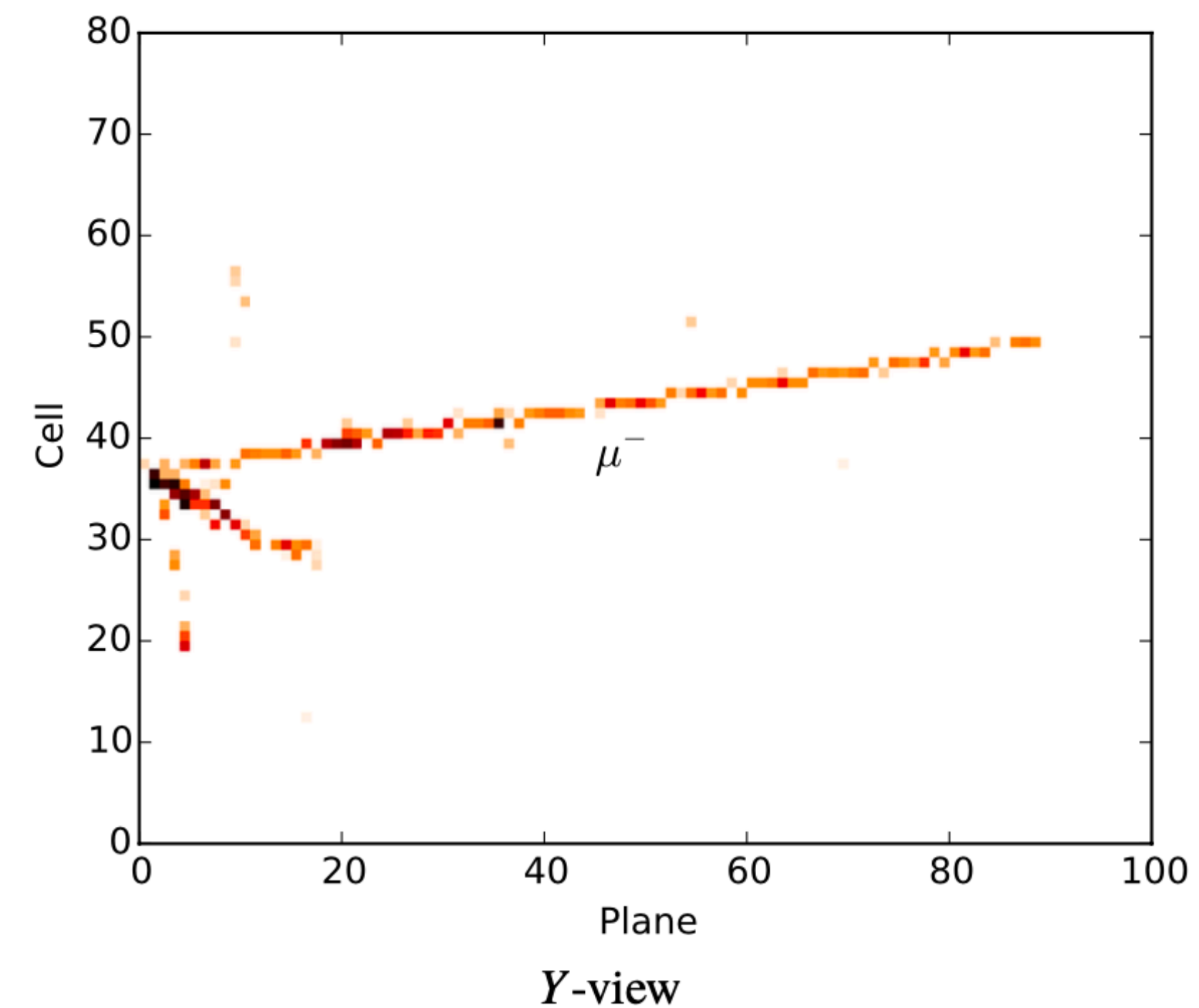
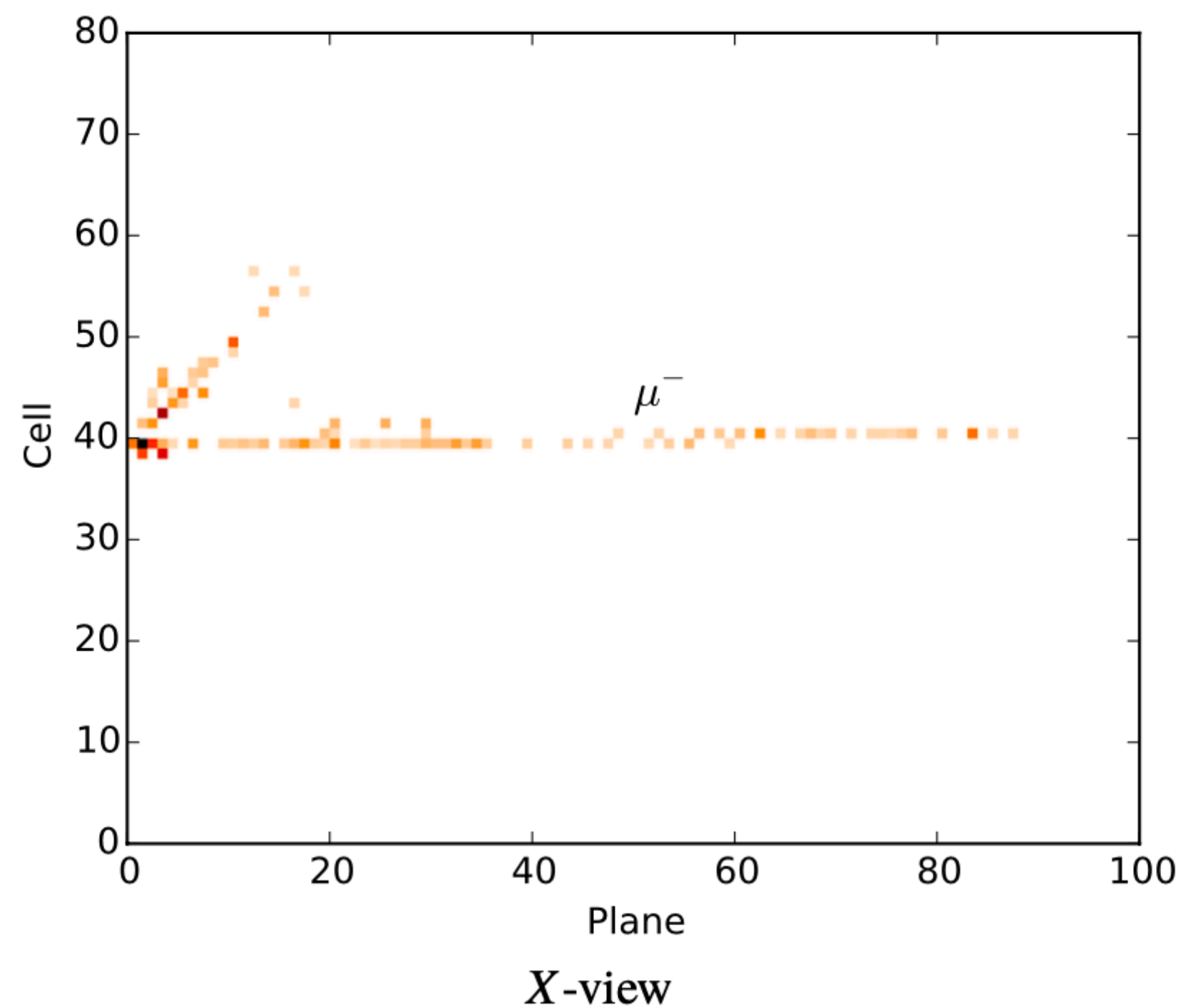
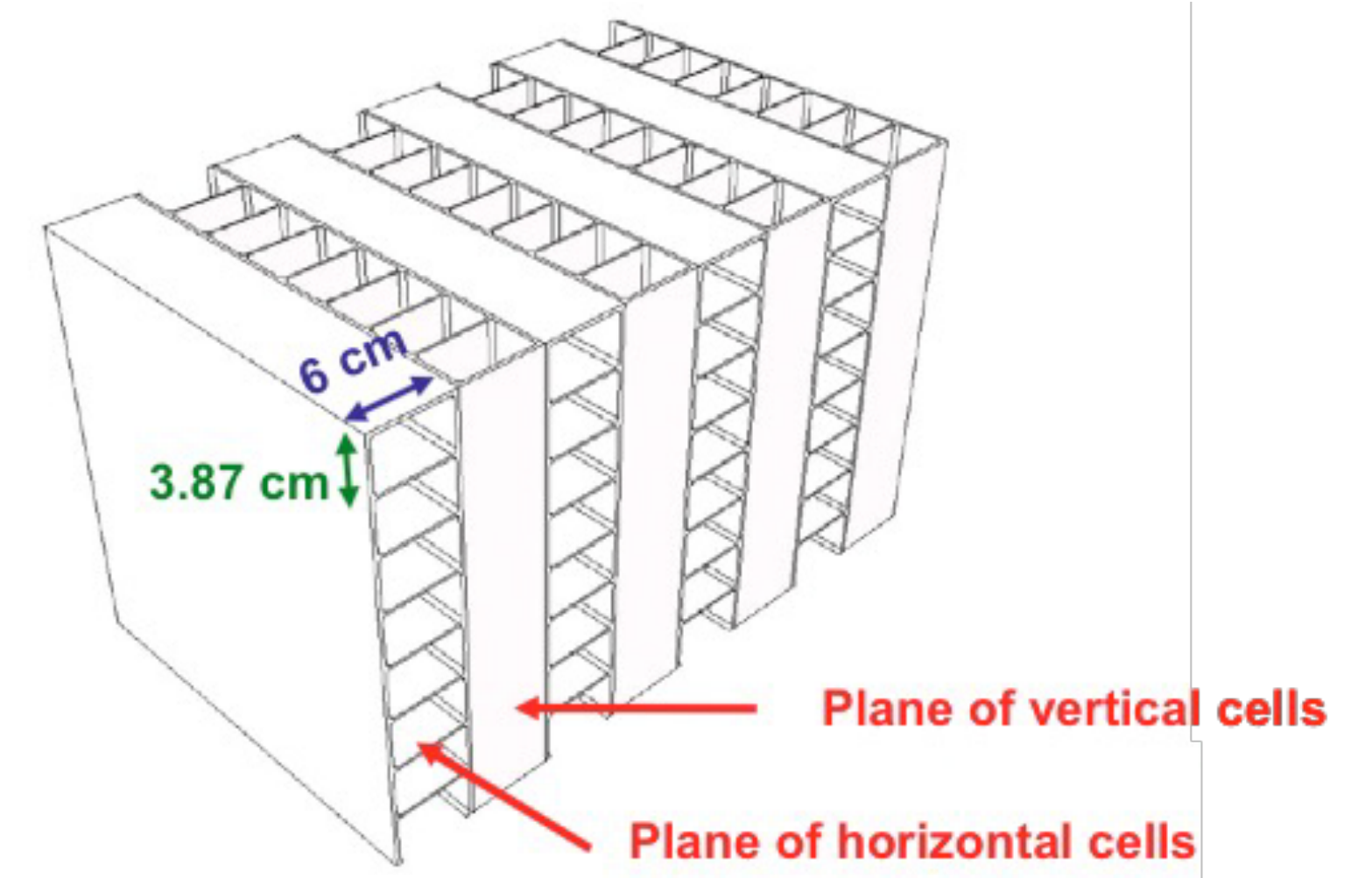


- NOvA uses a variety of algorithms to reconstruct physics information for which slicing is a core input
- Machine learning is making significant contribution in the reconstruction chain and can replace “traditional” kinematic based algorithms in some cases

- **Classification**
- **Regression**

Detector Views

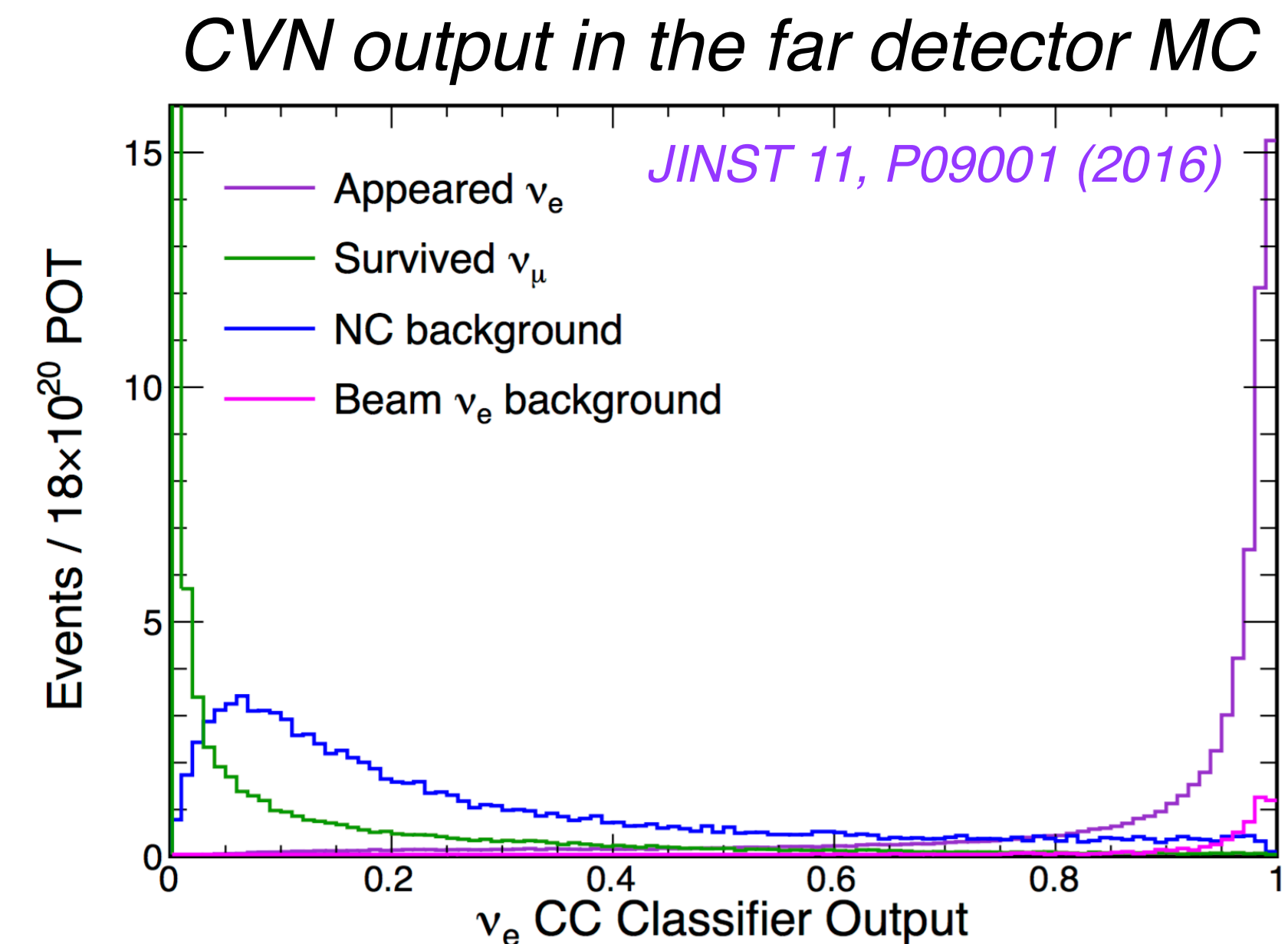
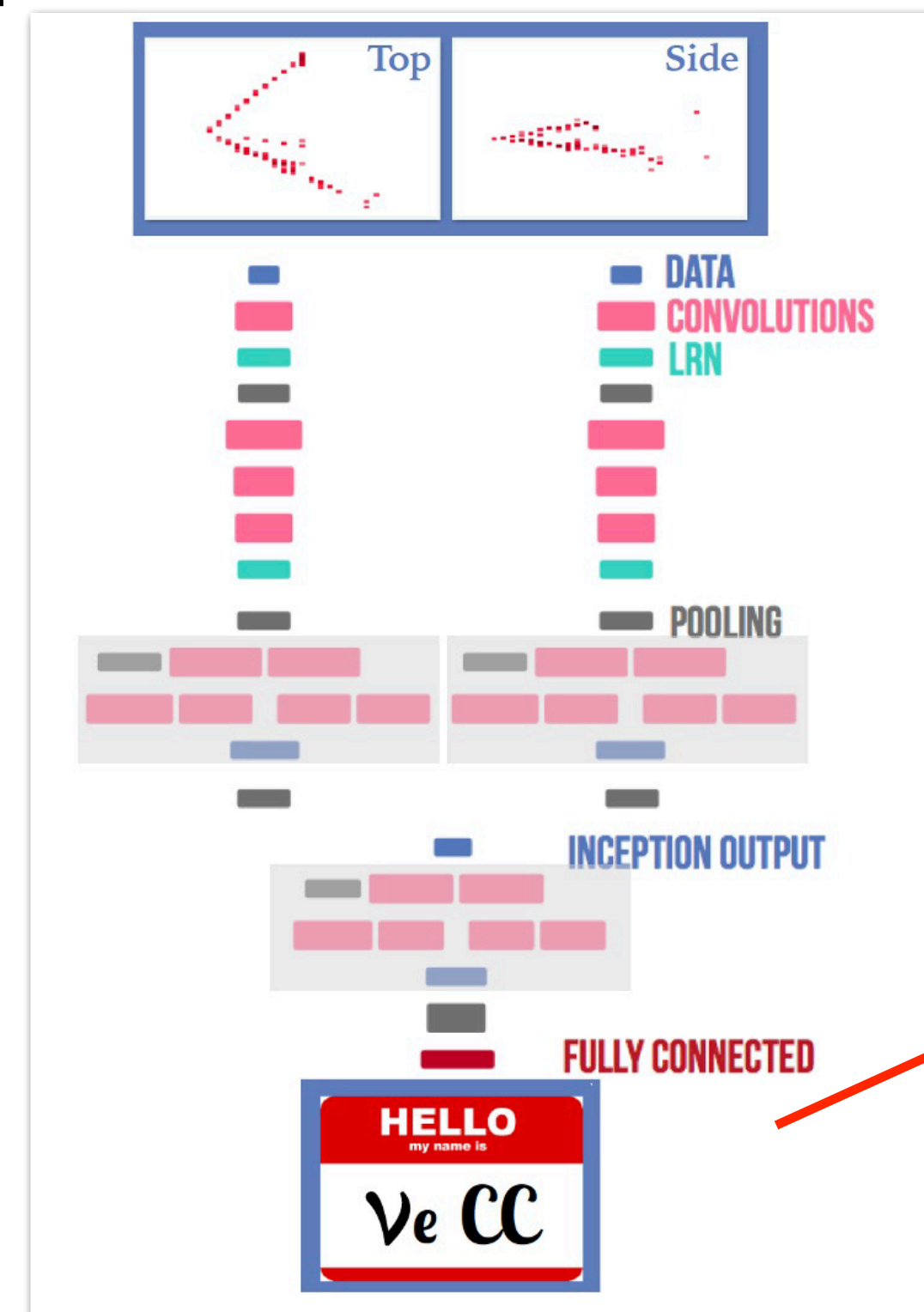
- NOvA detectors are naturally segmented
- Producing a pair of pixel maps (Cell number v.s. Plane number, size: 100x80) for the Top and Side view of each interaction



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



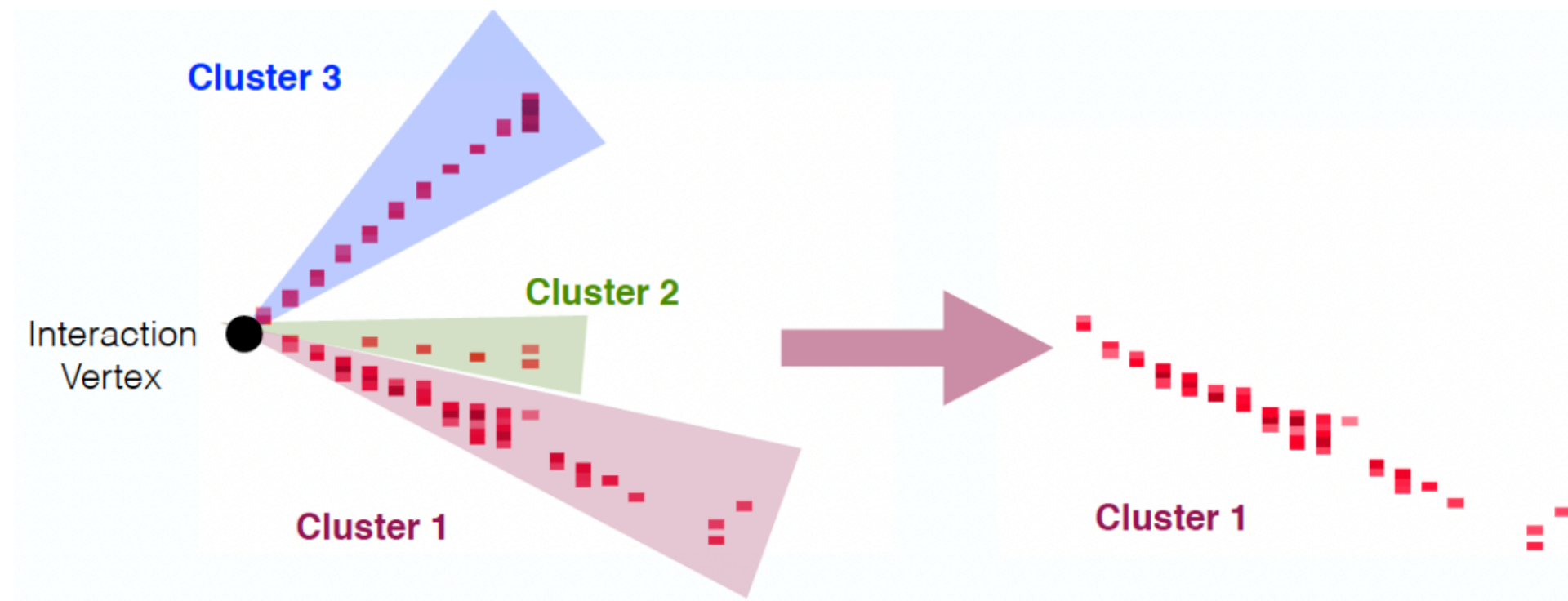
Select ν_μ ($\bar{\nu}_\mu$) CC and ν_e ($\bar{\nu}_e$) CC candidates from neutrino (anti-neutrino) beam with CVN

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

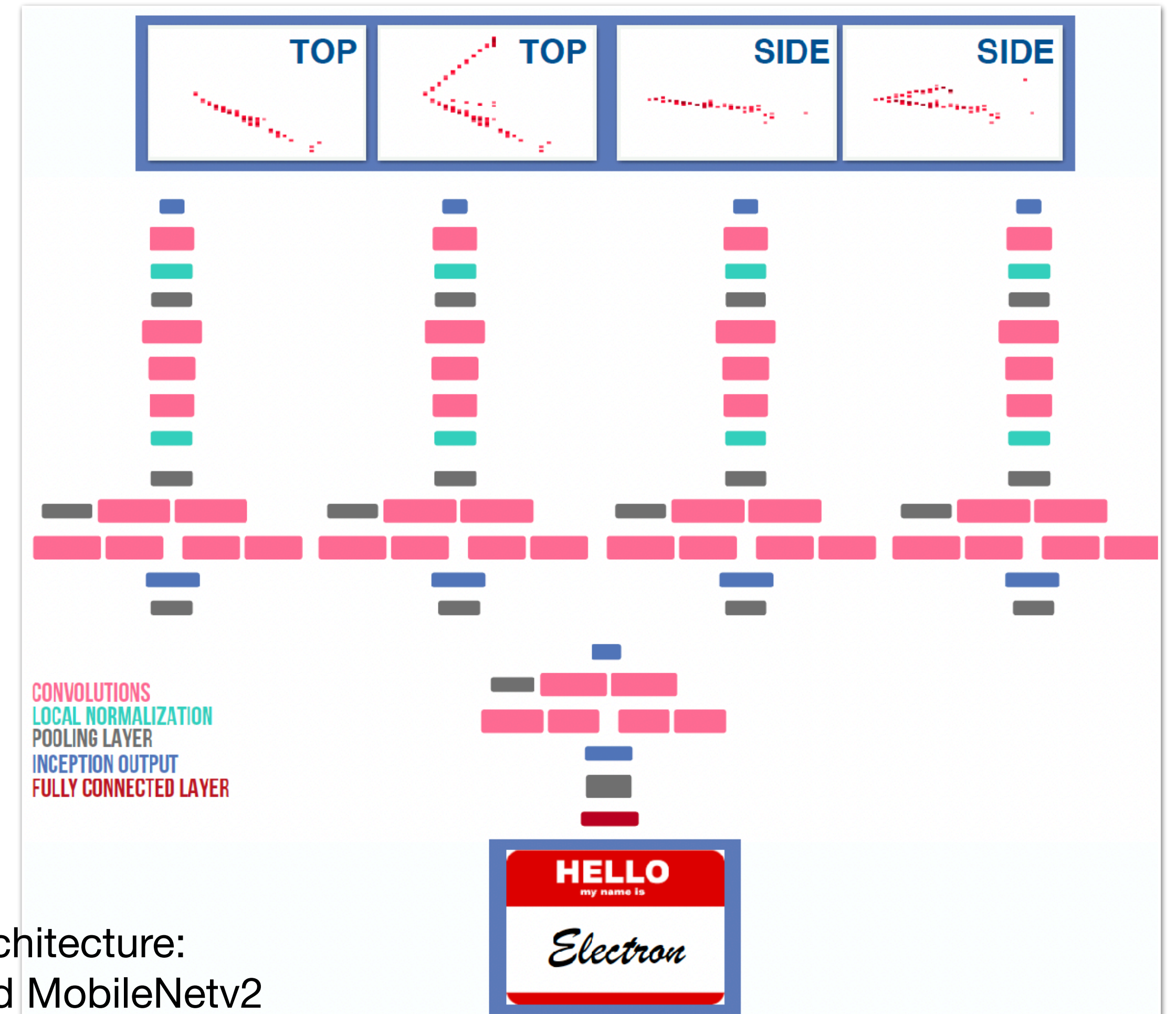


CNN-based Event Classifier (ProngCVN)



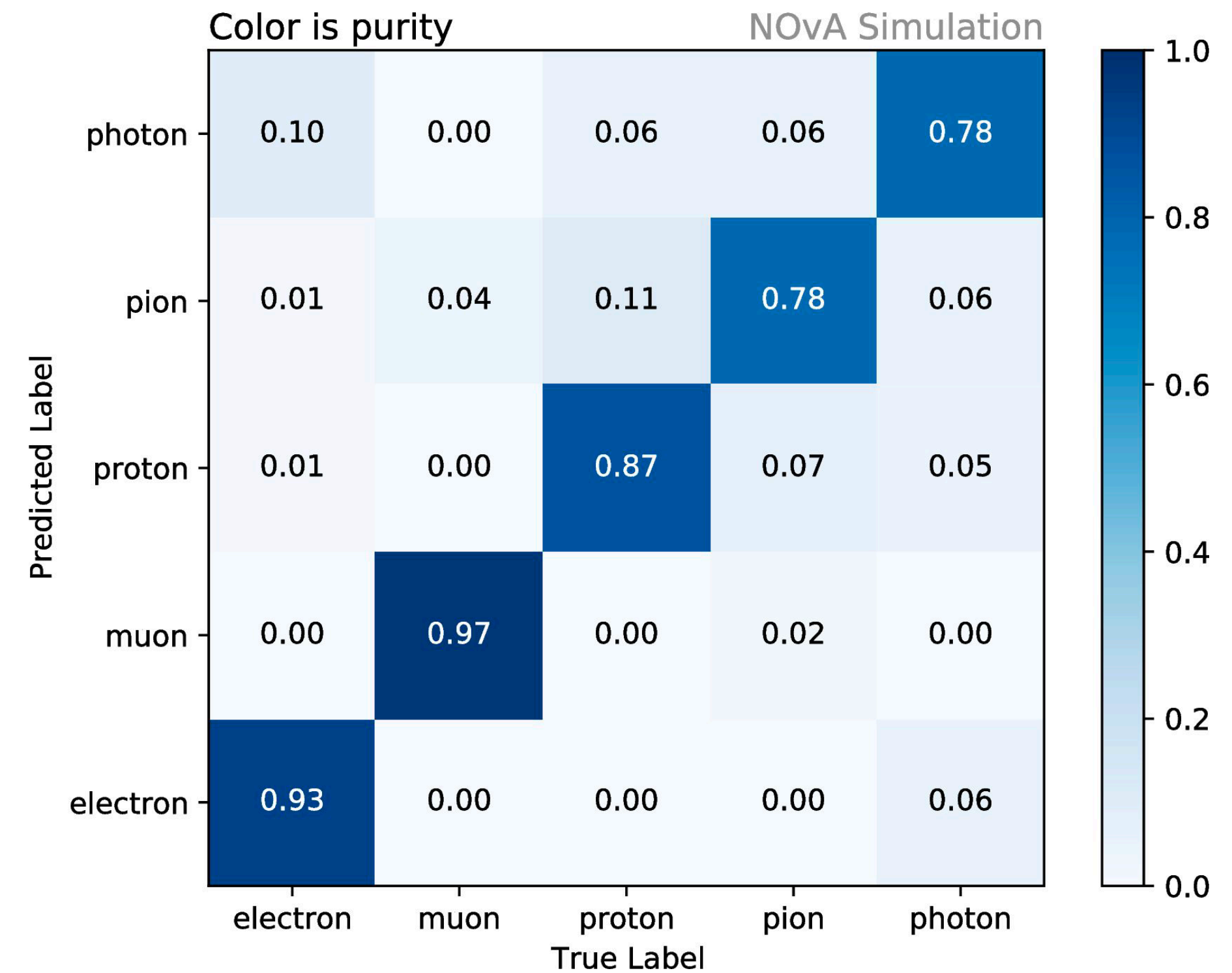
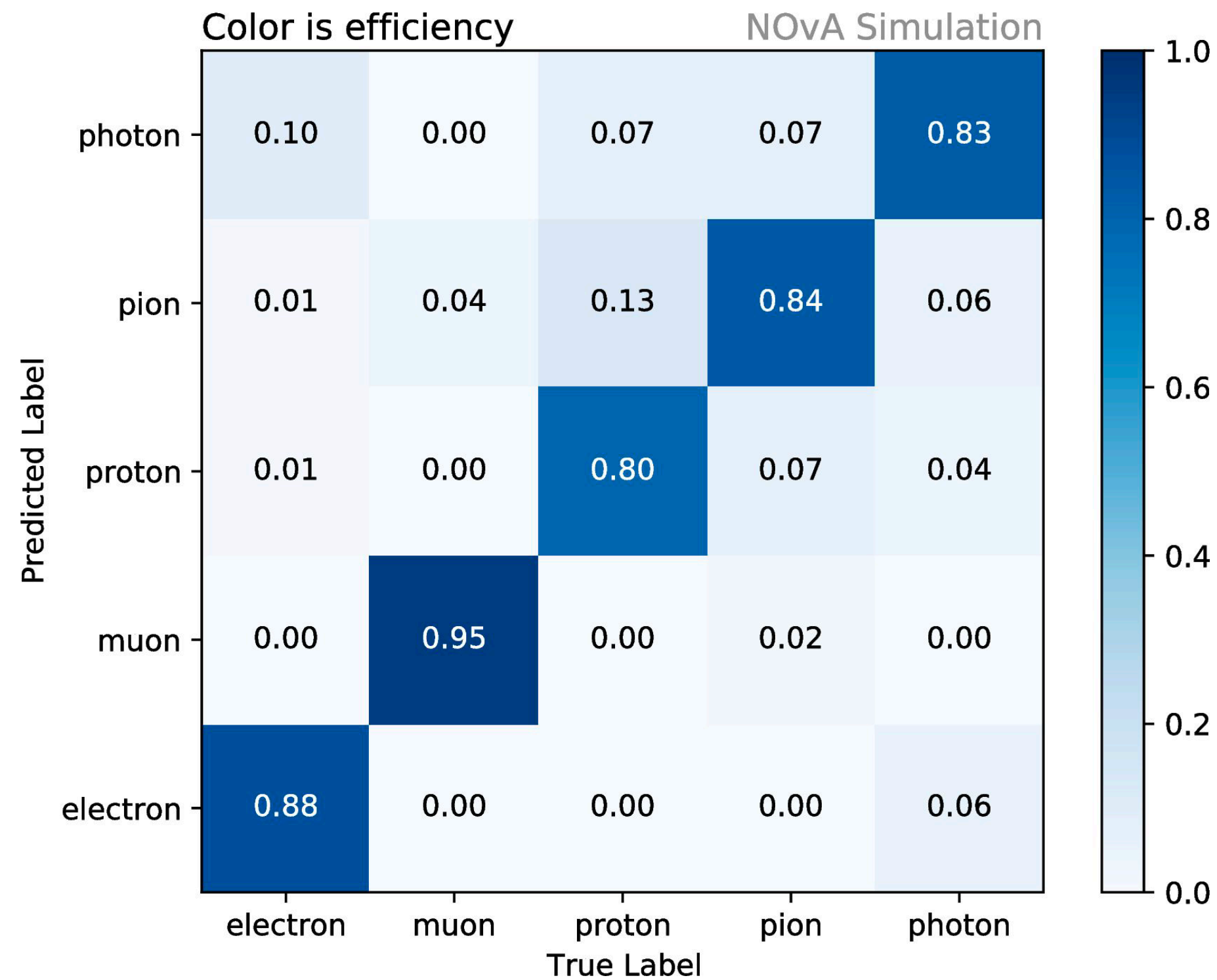
- Single particles are separated using geometric reconstruction methods
- Classify particles using both views of the **particle** and both views of the entire **event**
- This shows the network contextual information about single particles

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



CNN architecture:
Modified MobileNetv2
Four-tower Siamese structure

CNN-based Event Classifier (ProngCVN)



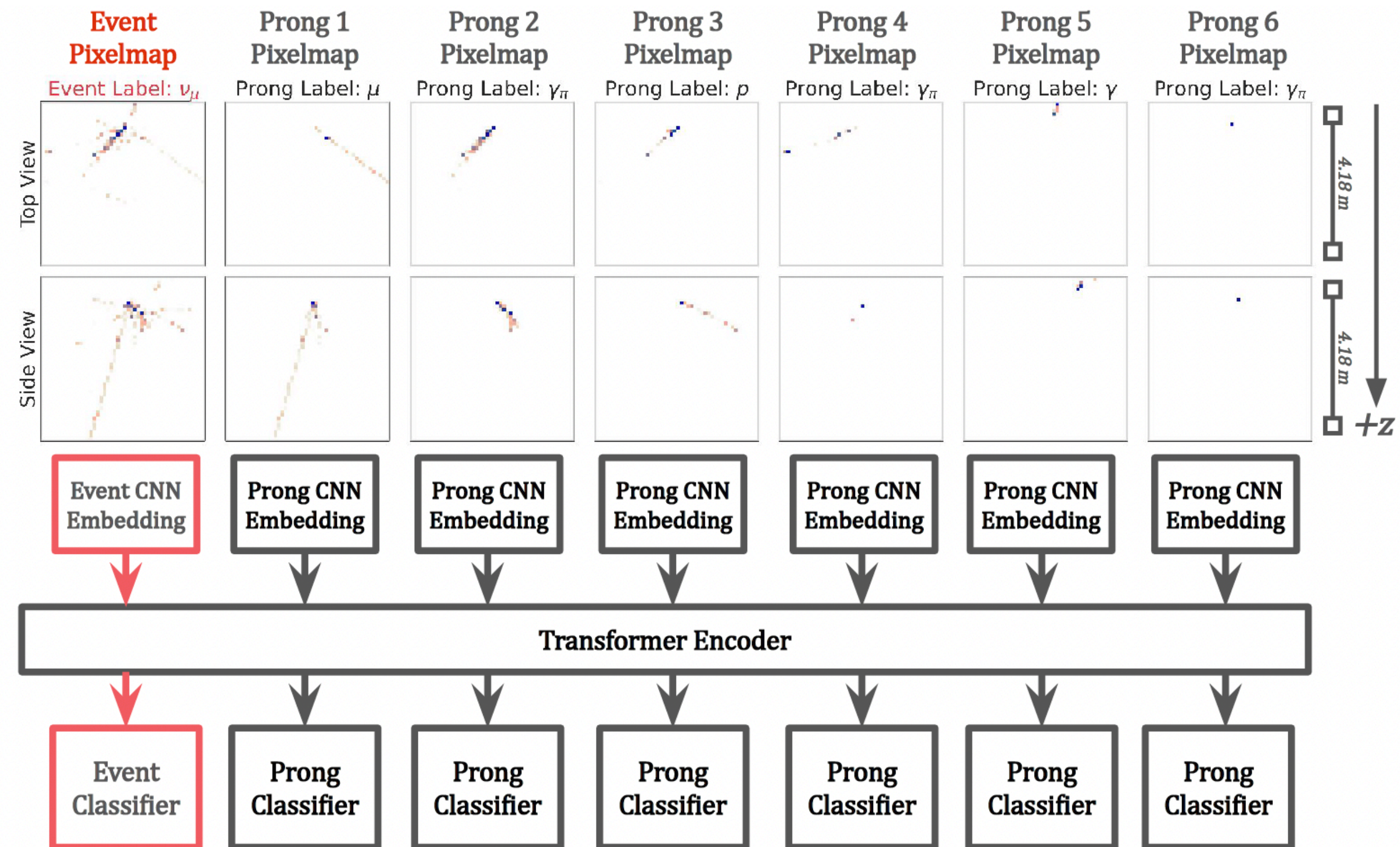
- Improvements were found in both efficiency and purity for all particle types, compared to the particle-only network
- In particular ~10% increase in the efficiency of selecting photons and pions

Transformer for both Event and Particle Classification

- Transformer: attention based network, ideal for training on variable-length collection of objects such as prongs

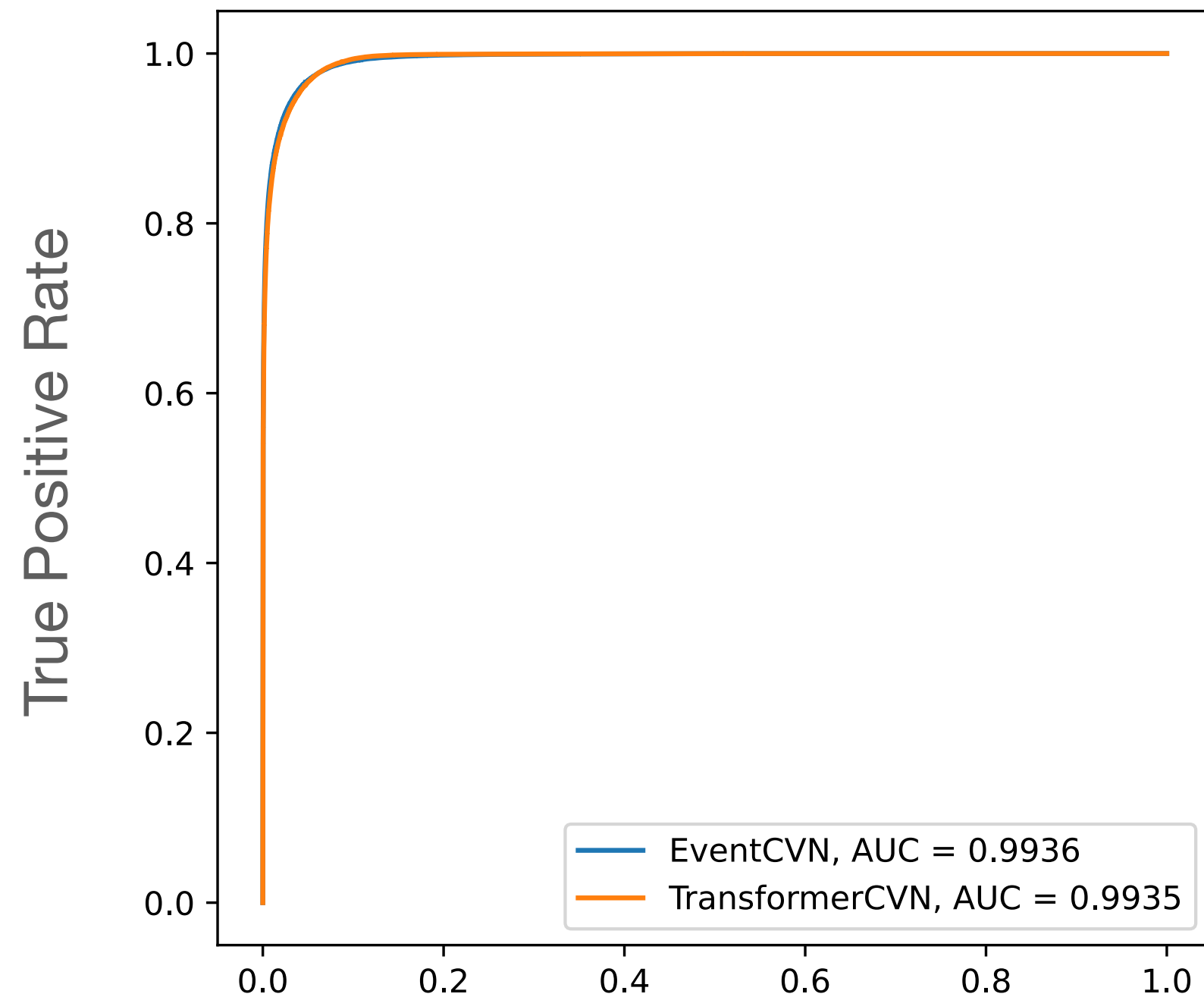
“TransformerCVN” = Transformer+CNN

- Combines the **spatial learning enabled by convolutions** with the **contextual learning enabled by attention**
- Classifies each event and reconstructs every individual particle's identity
- Attention mechanisms
 - focus on regions with high importance, reduce the computing burden and enhance performance
 - enable performing interpretability studies

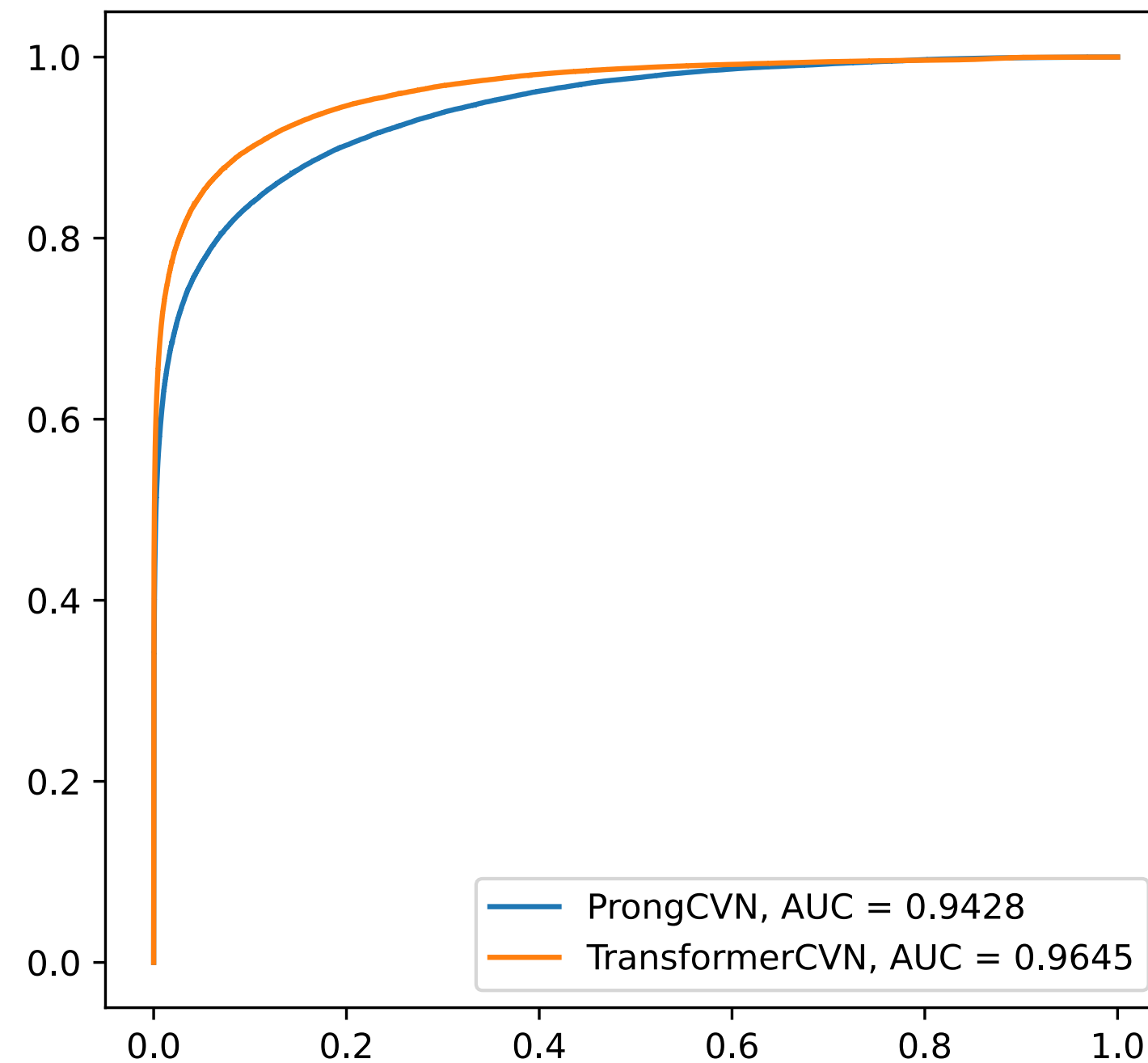


Transformer for both Event and Particle Classification

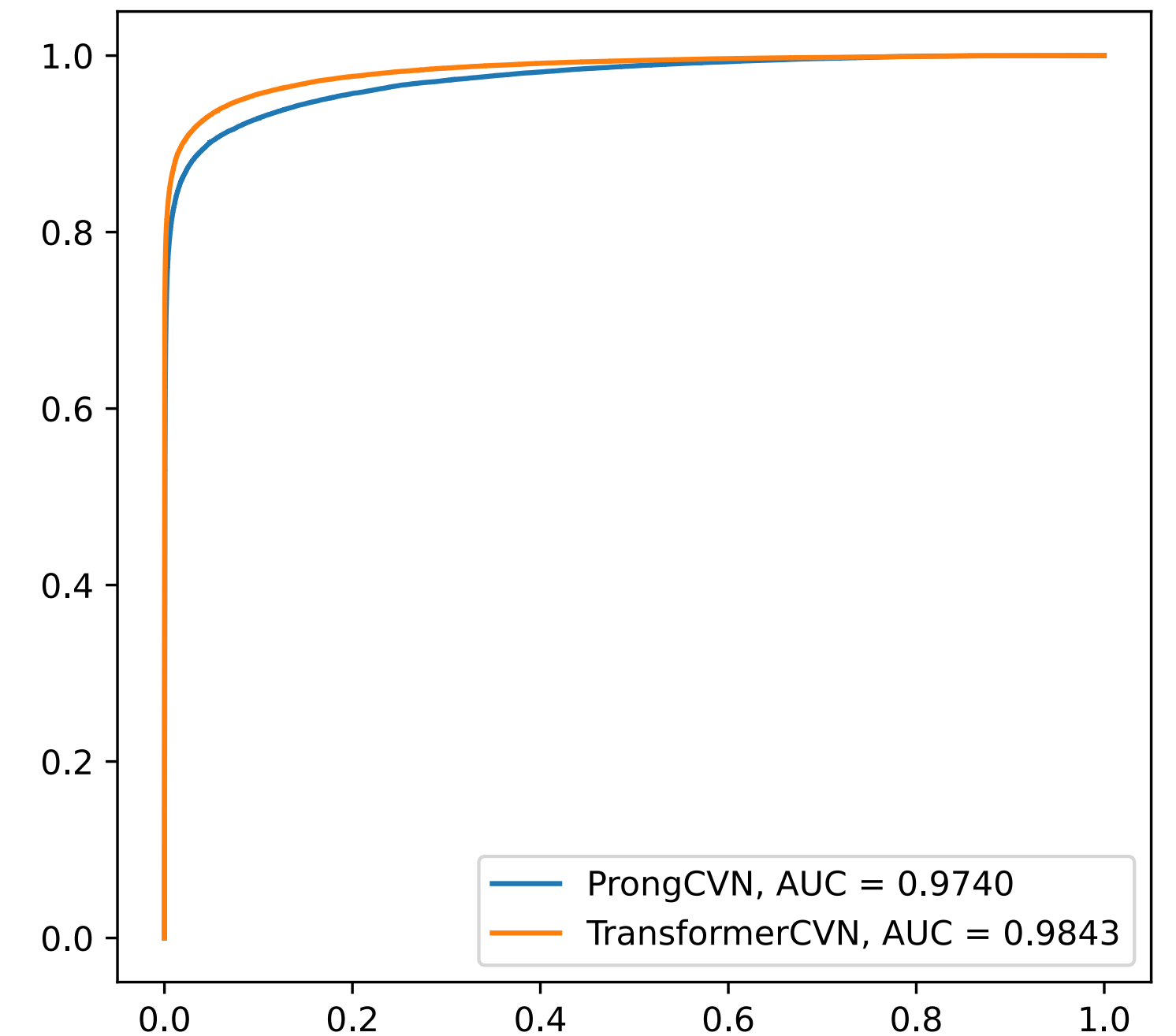
ν_e CC



Electron



Muon

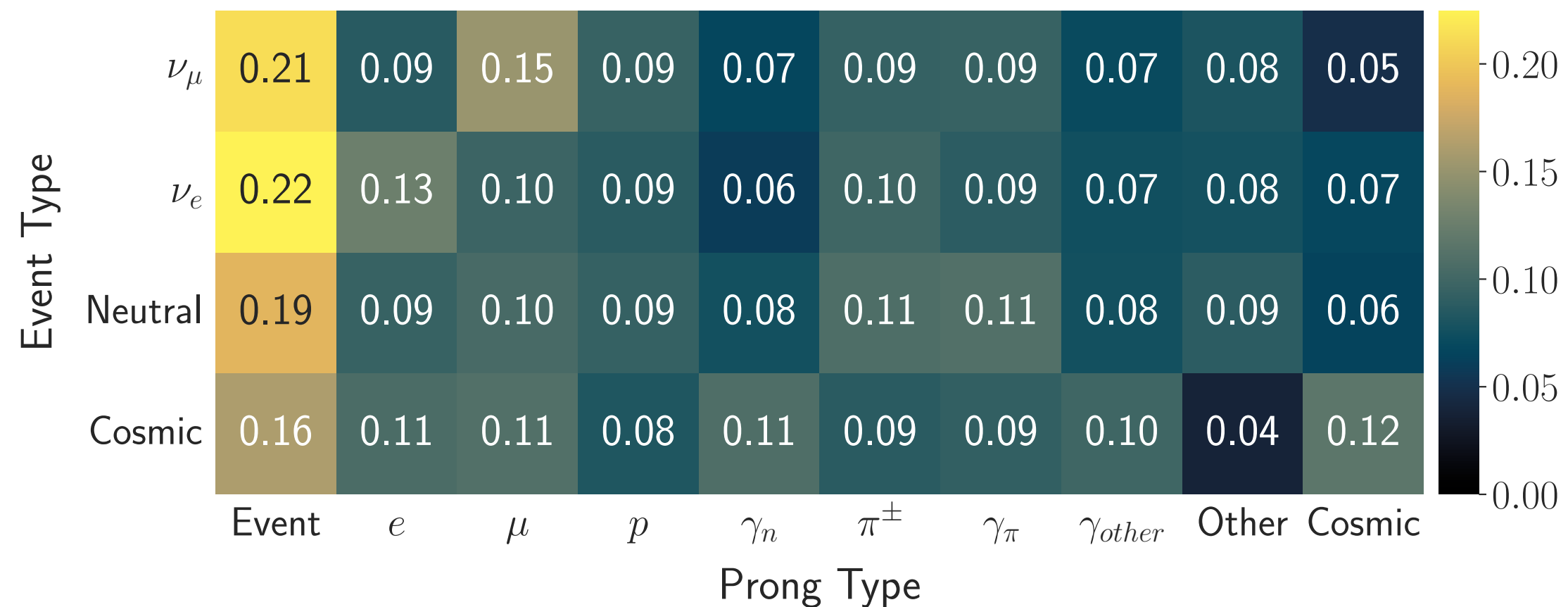


False Positive Rate

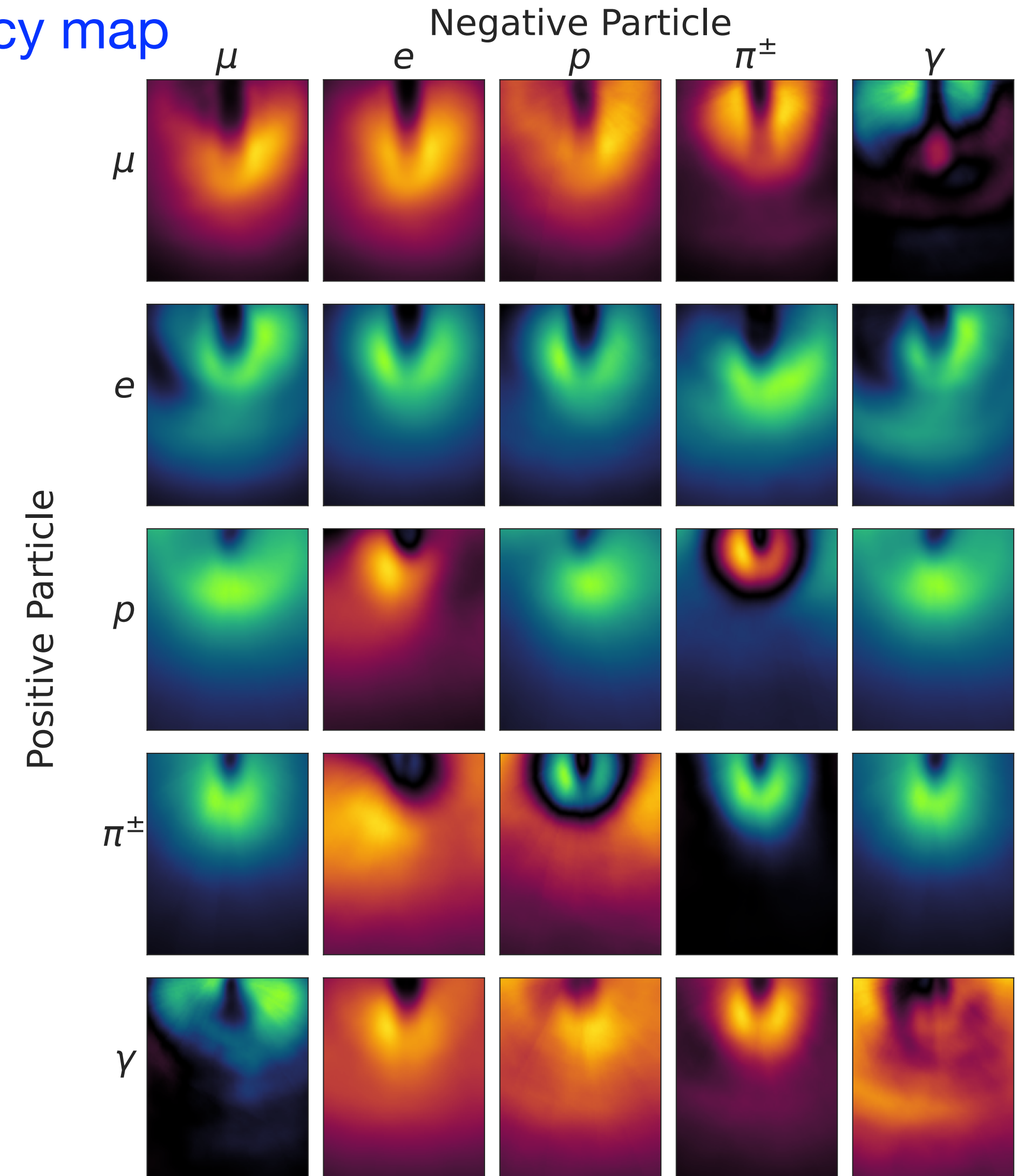
- Comparable performance of identifying neutrino flavors compared to our benchmark network - EventCVN
- Great improvement in particle identification, benefits from the additional context provided by all prongs and the transformer's attention mechanism, compared to our benchmark network - ProngCVN

Transformer for both Event and Particle Classification

Attention map

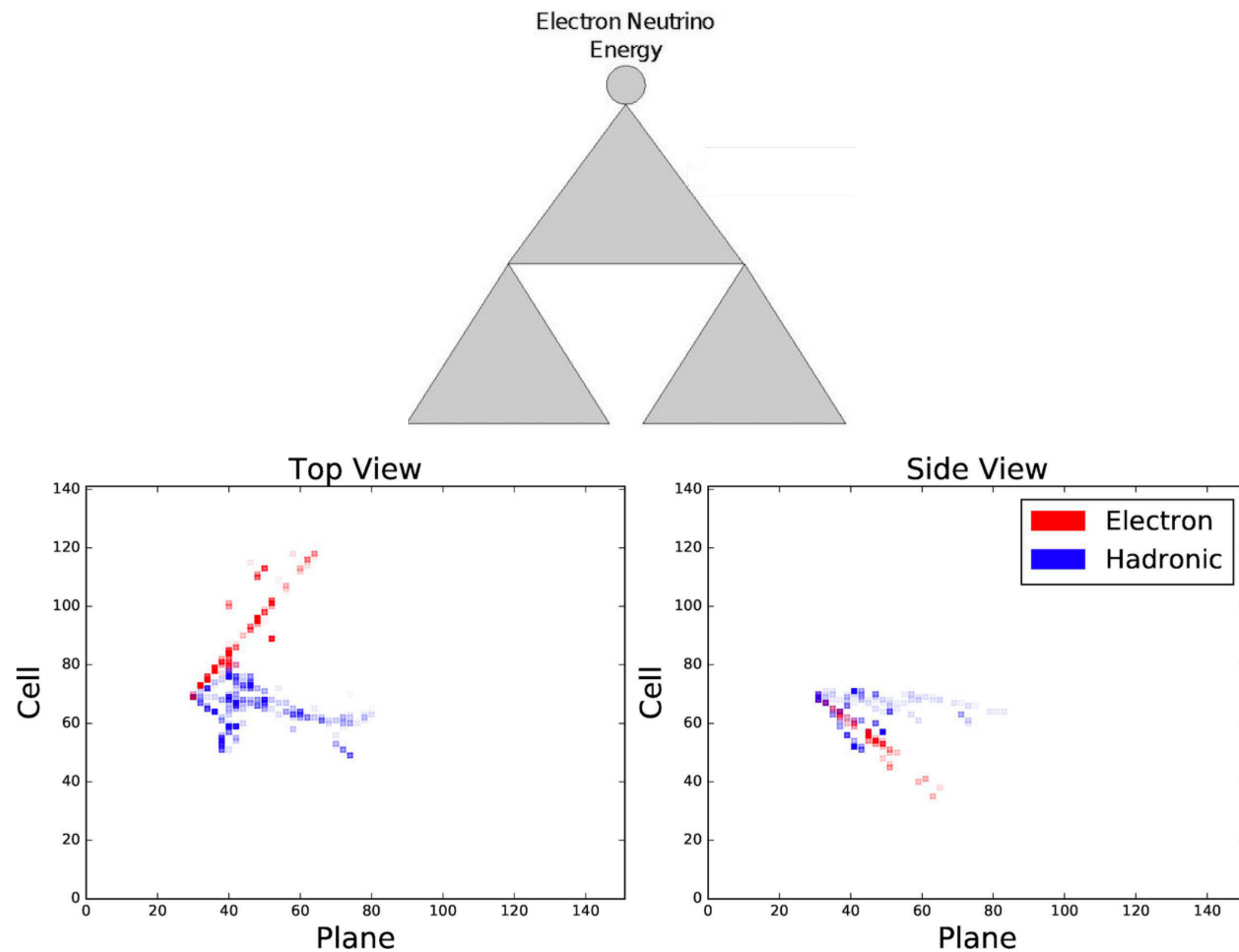


Saliency map



- Interpretability of the network
 - **Attention map**: importance of each input to each output
 - Diagnose neural network and explain decision
 - **Saliency map**: derivative of a network output w.r.t the input pixel
 - Study salience to understand which regions the Transformer focuses on to identify a particle

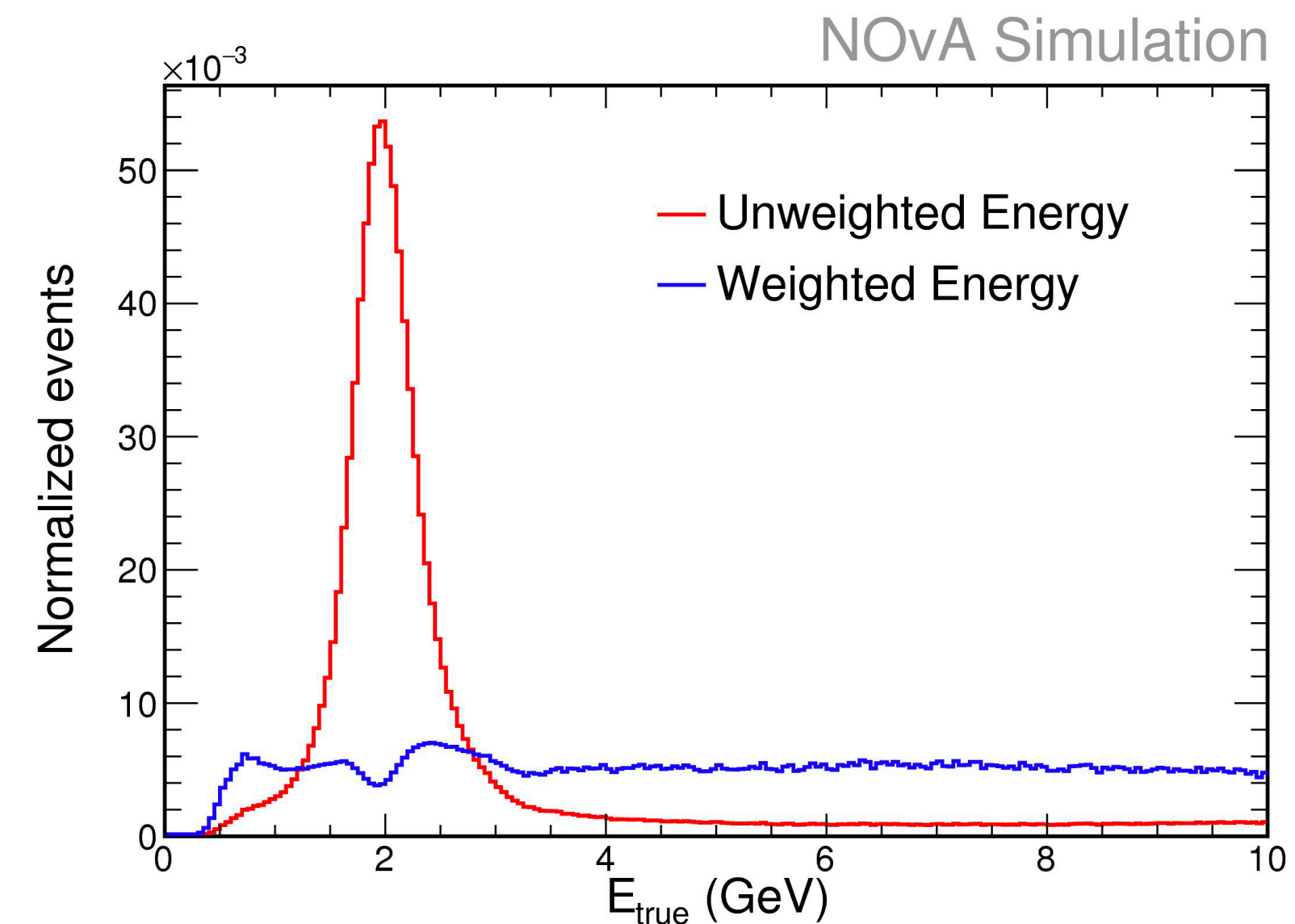
Regression CNNs for Energy Estimation



[*PhysRevD.99.012011*](#)

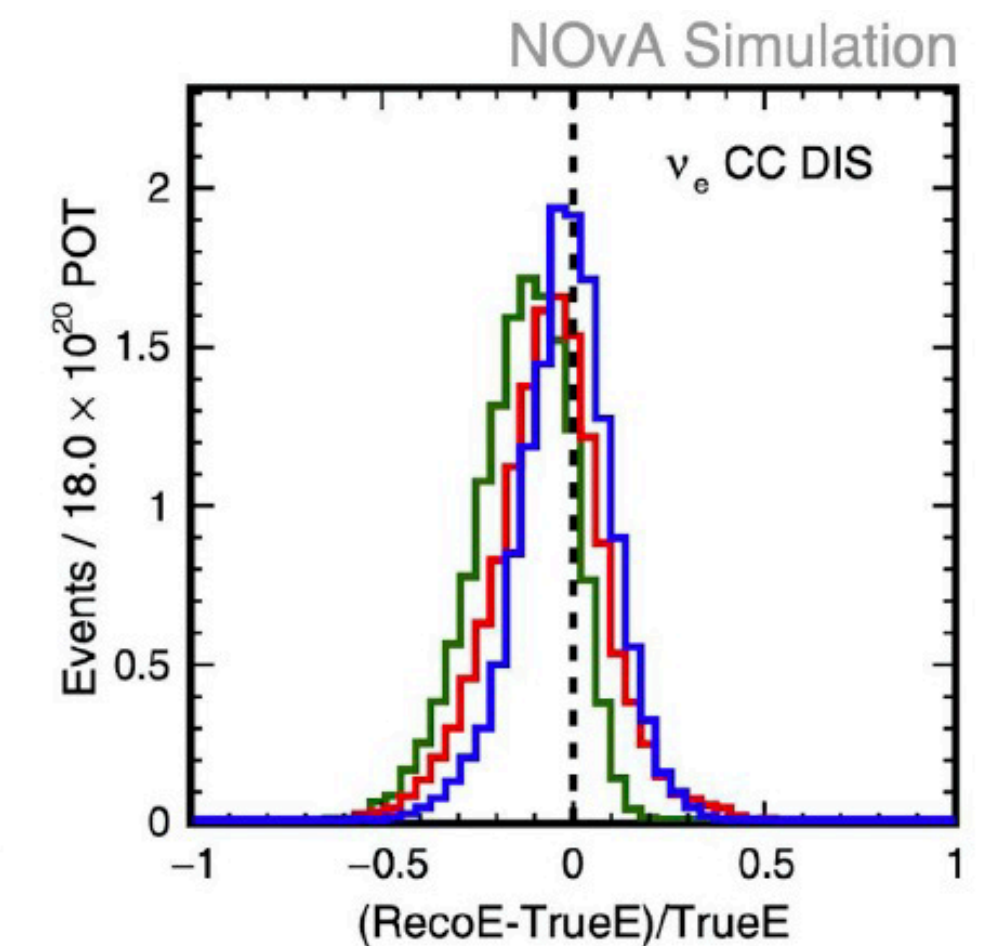
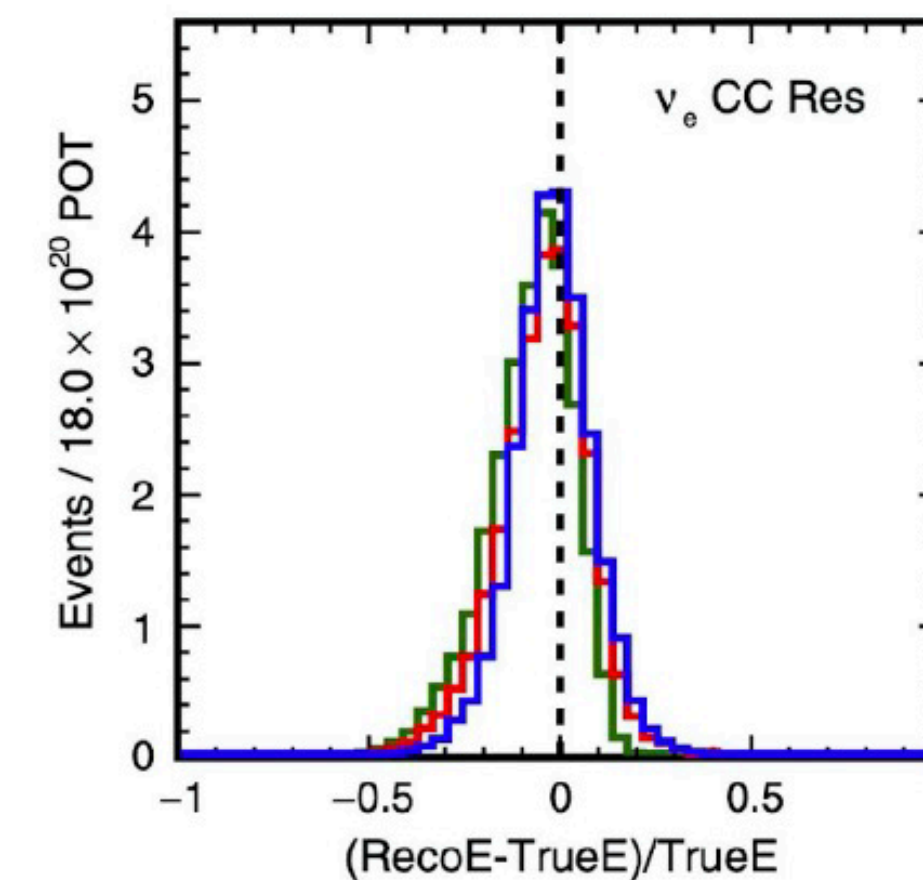
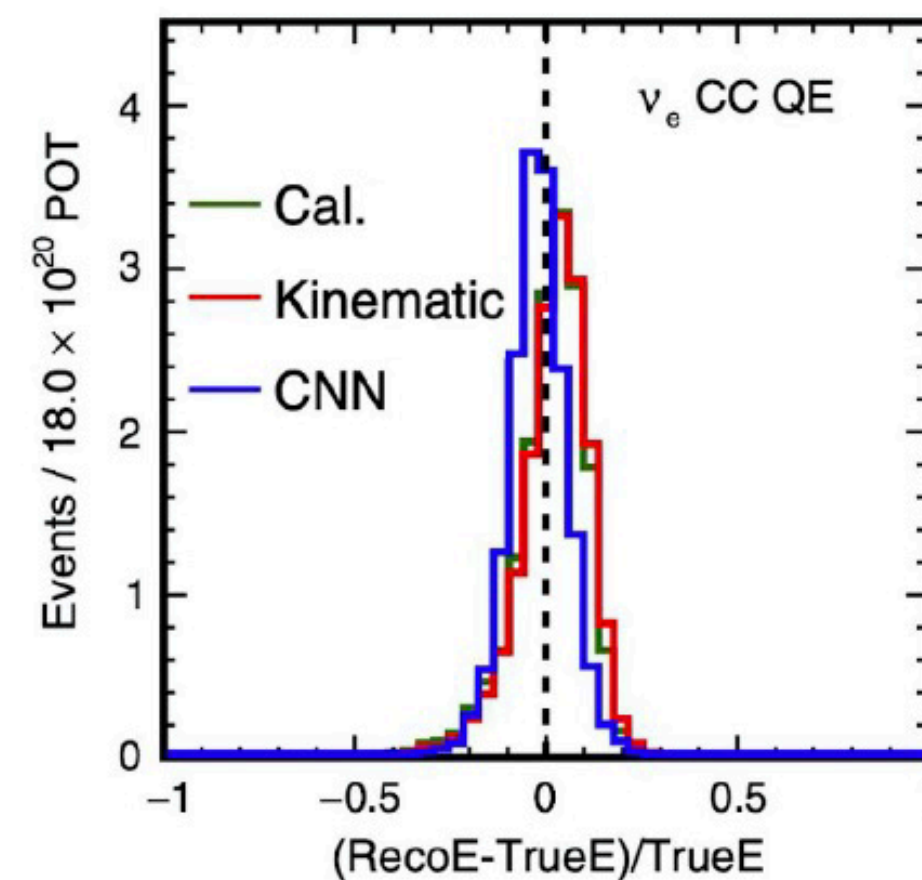
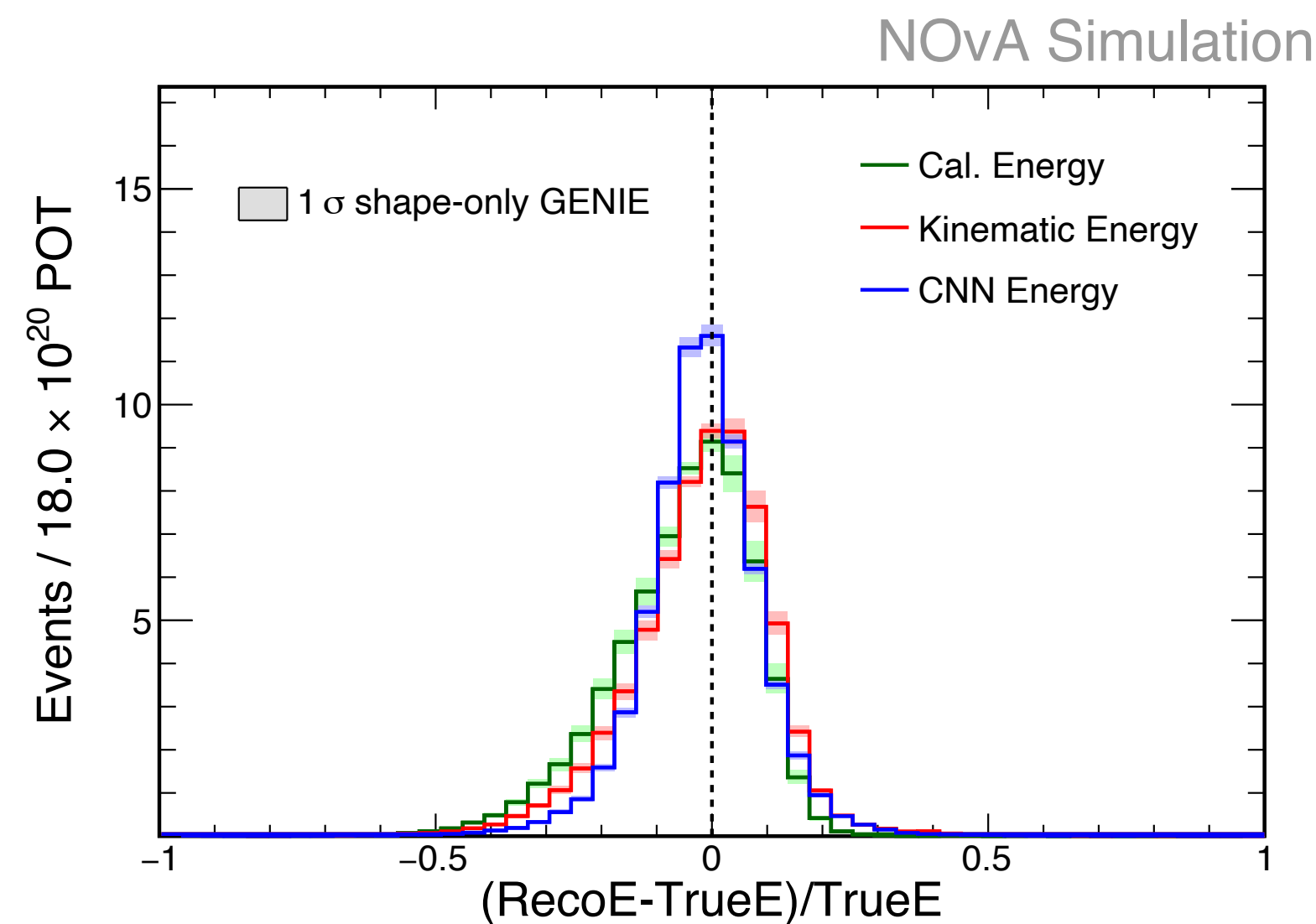
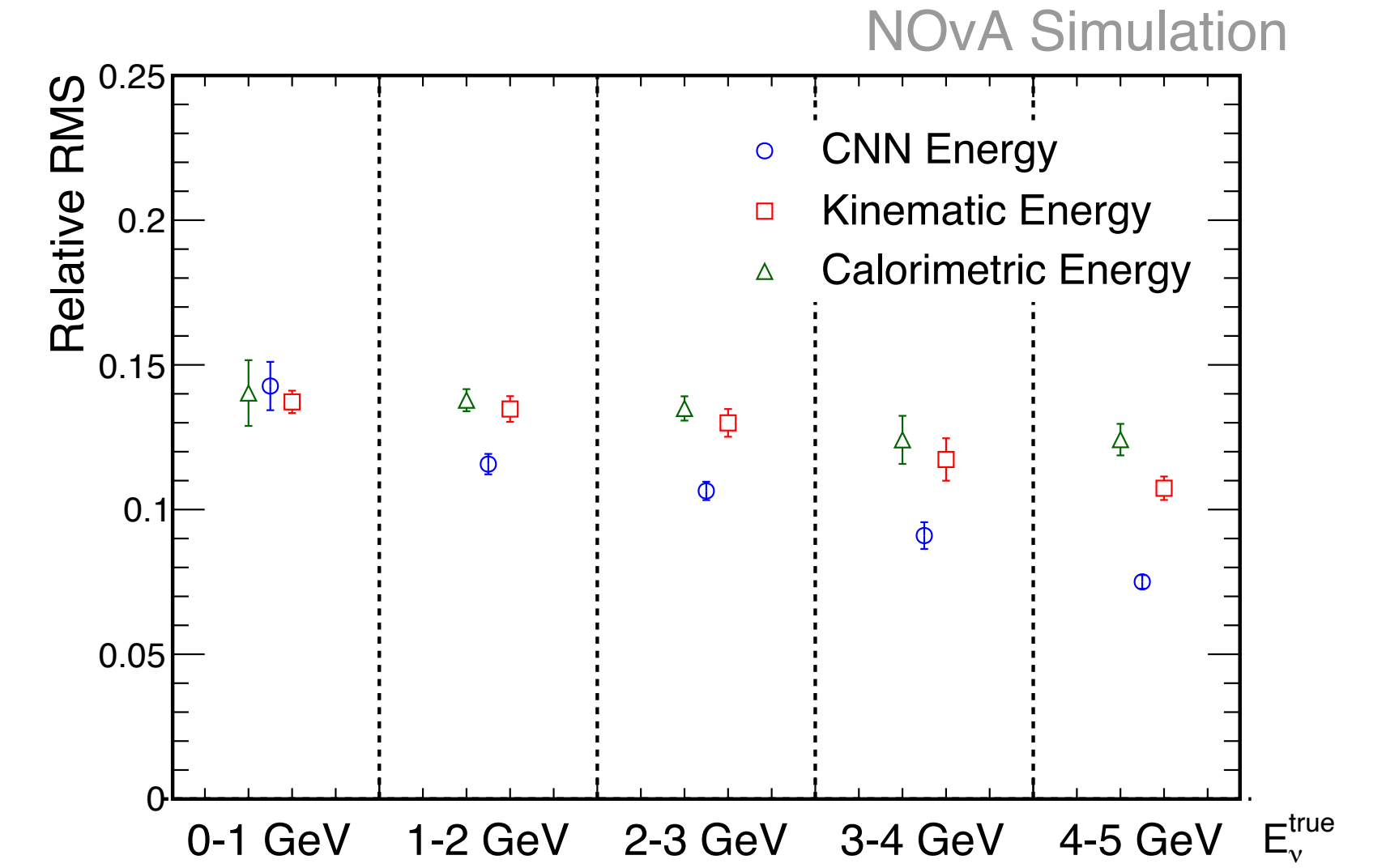
- The CNN architecture used is an adapted ResNet
- Weighting scheme so the loss function sees a flat energy distribution, to control energy dependence
- Use mean absolute percentage error instead of square of errors to decrease the effects of outliers

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



Regression CNNs for Energy Estimation

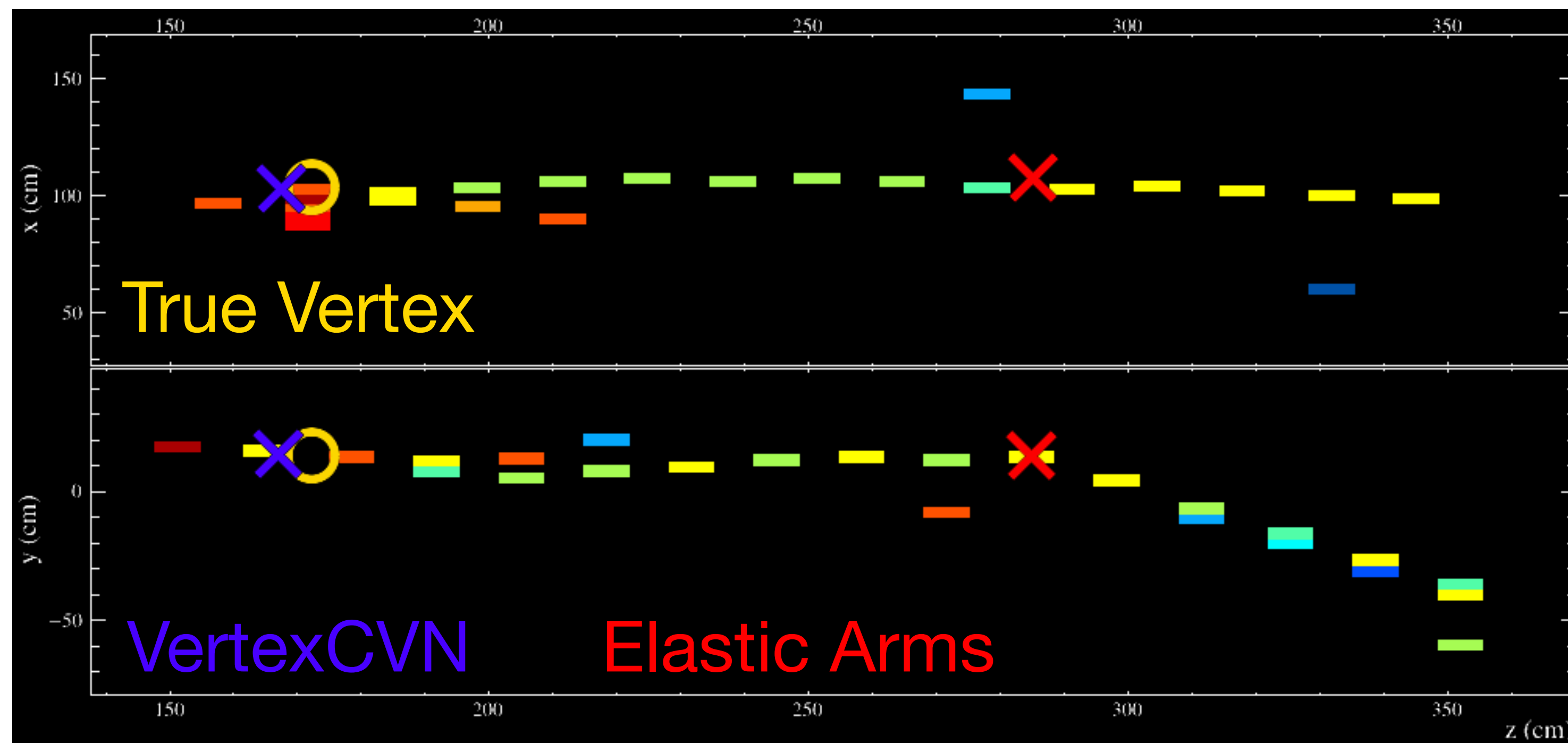
- Regression CNN shows a better resolution compared with kinematics-based energy reconstruction
- Shows smaller systematic uncertainties due to neutrino interaction simulation
- Good stability over interaction types



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

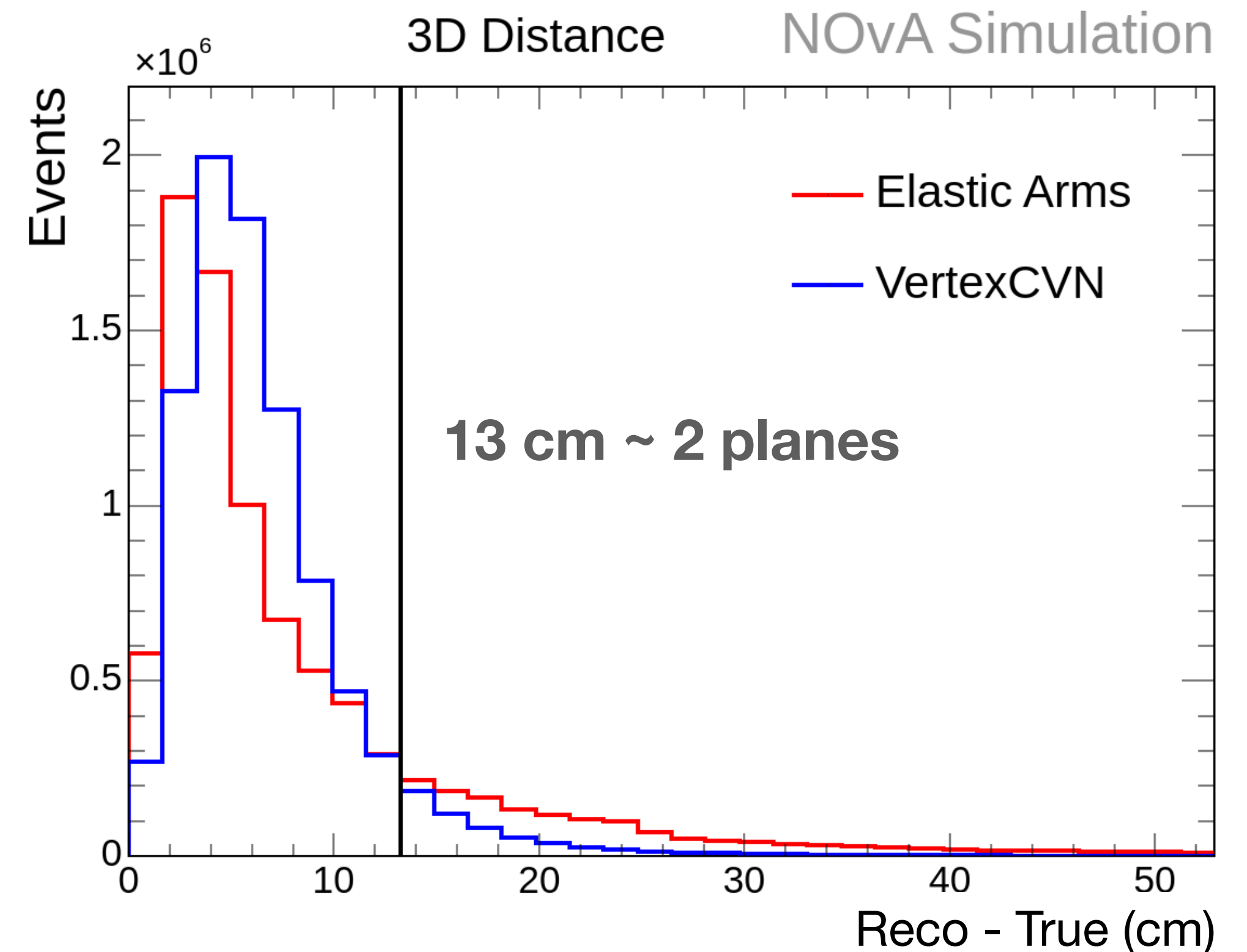
Vertex Reconstruction

- Apply machine learning to estimate the position where the neutrino interactions happen (VertexCVN)
- Developed to address several known failure modes of NOvA's existing algorithm "Elastic Arms"
 - Forward failure - tendency for main prong to be split into two
 - Backward failure - tendency for multiple, small prongs to be combined into one



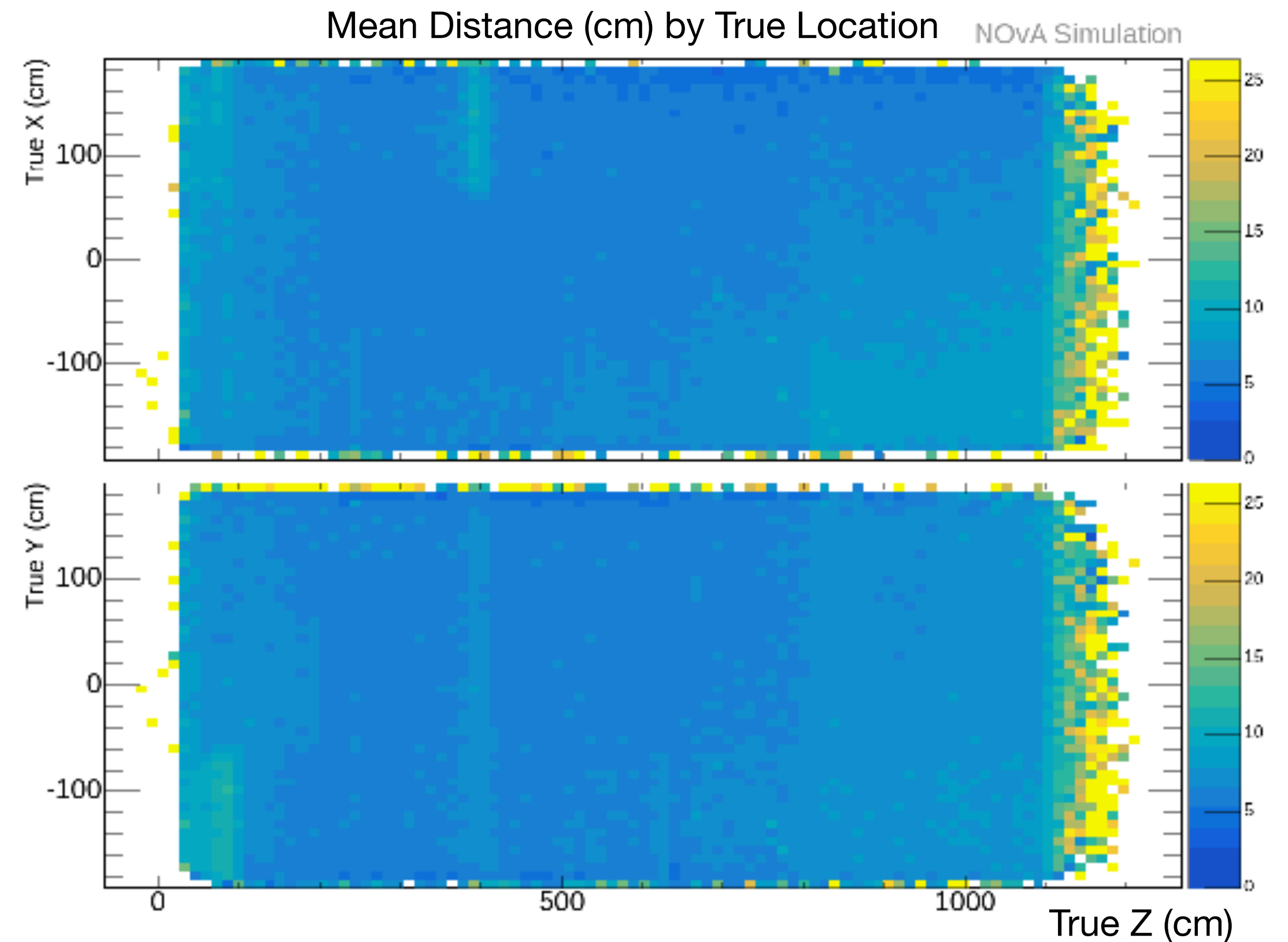
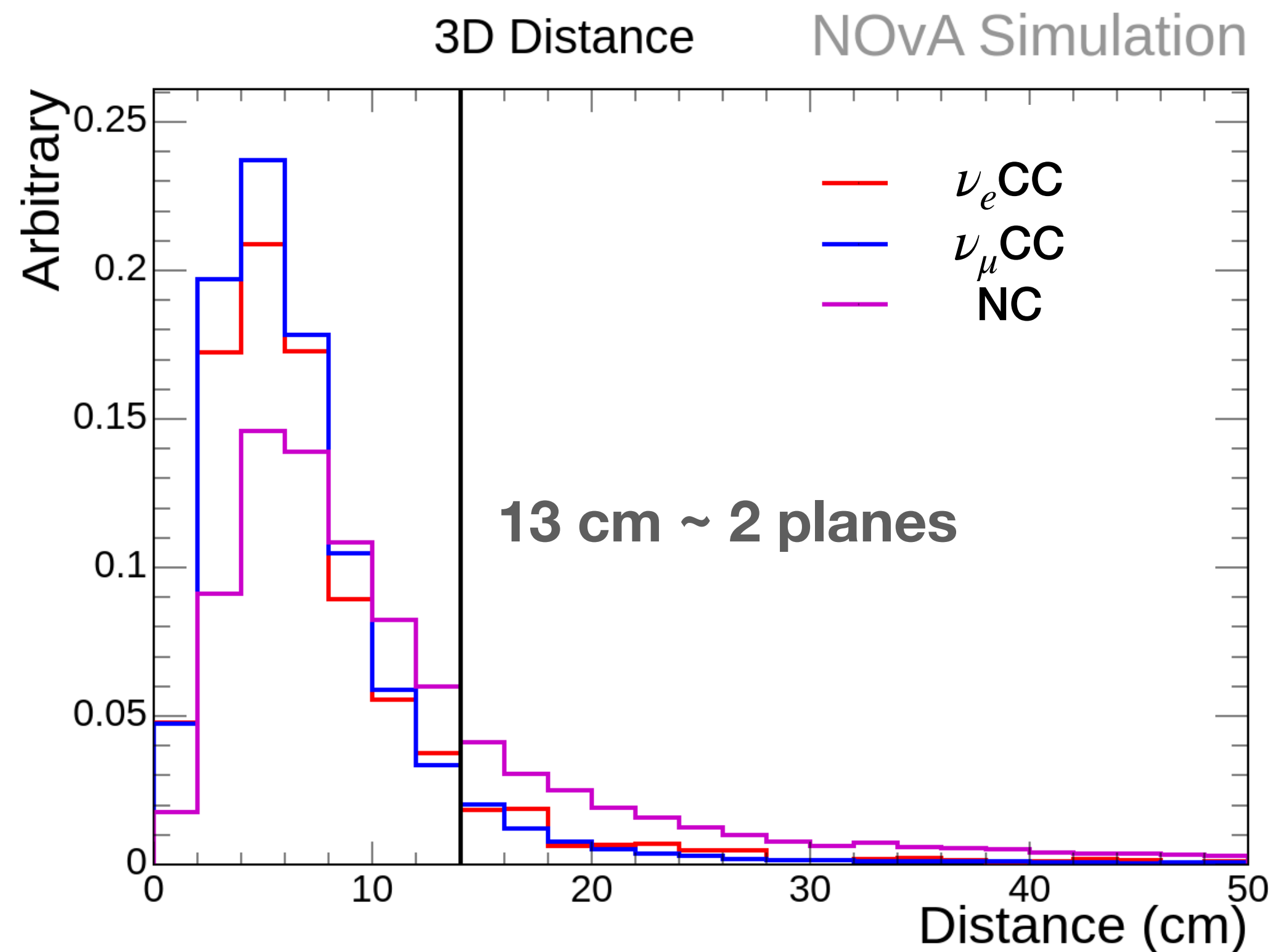
VertexCVN

- Better vertex finding, means better on
 - Clustering hits to form individual particle tracks/showers
 - Identifying particle types
 - Energy estimation
- Same network architecture as EventCVN (modified MobileNetv2, arXiv:1801.04381) was explored to predict one 3D vertex
 - Trained with both beam modes combined, but separately for Near (~18 million events) and Far (~25 million events) detectors
- VertexCVN is more precise than Elastic Arms, but slightly less accurate

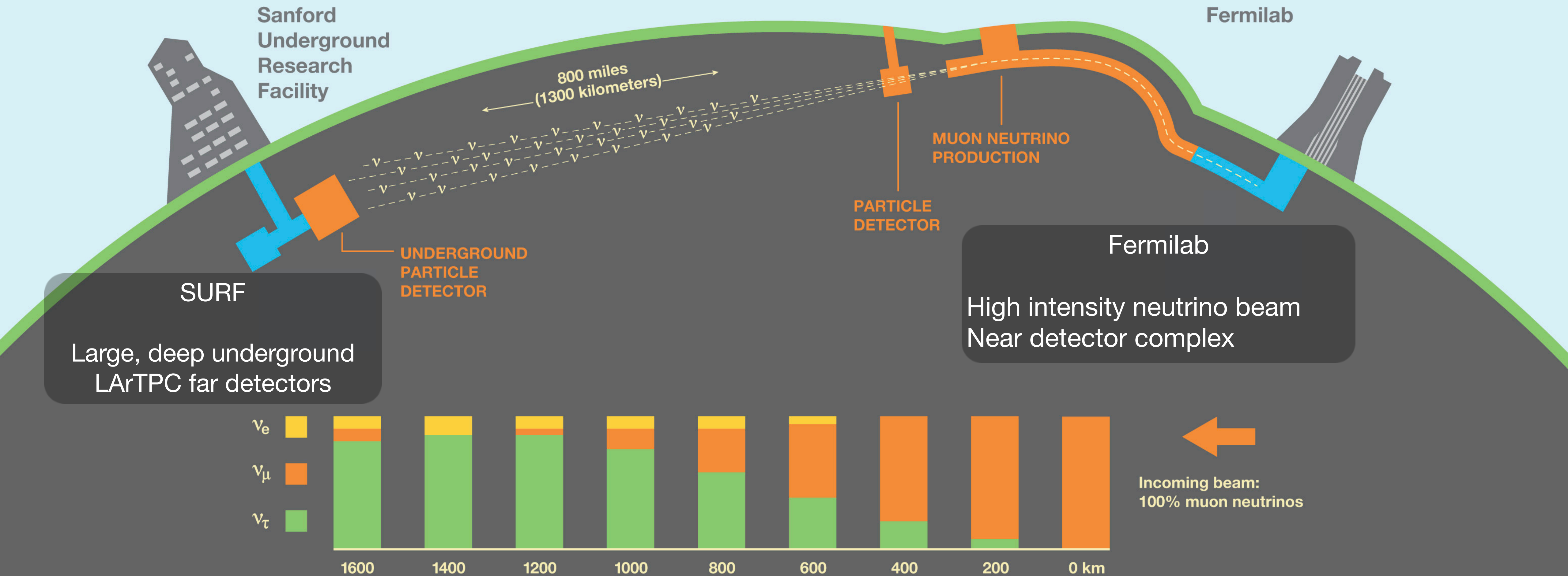


VertexCVN

- Relative performance has been studied for different types of neutrino interaction:
 - Better performance on ν_e CC, less on NC
- VertexCVN performance is largely insensitive to the true position of the vertex

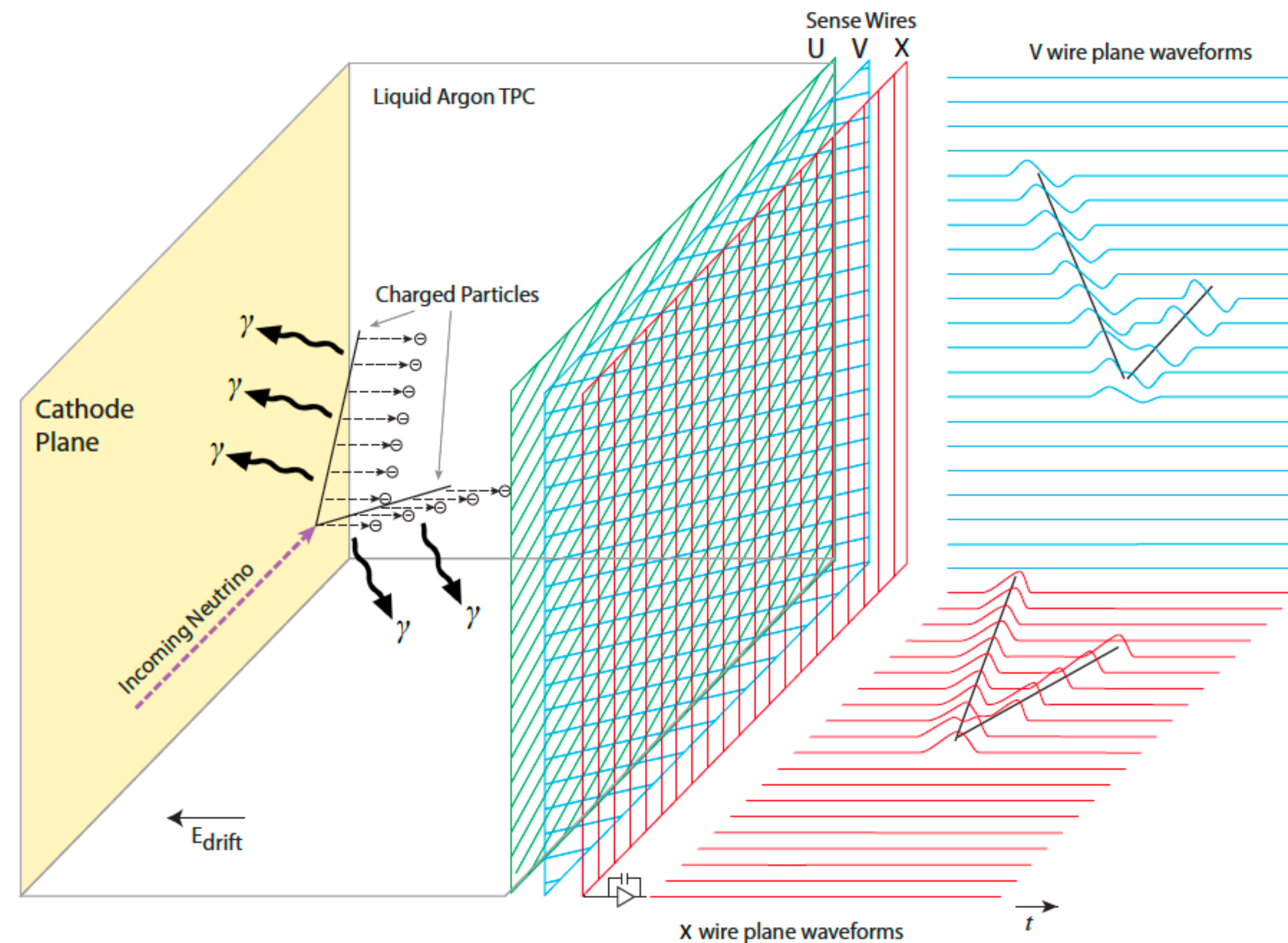


Deep Underground Neutrino Experiment

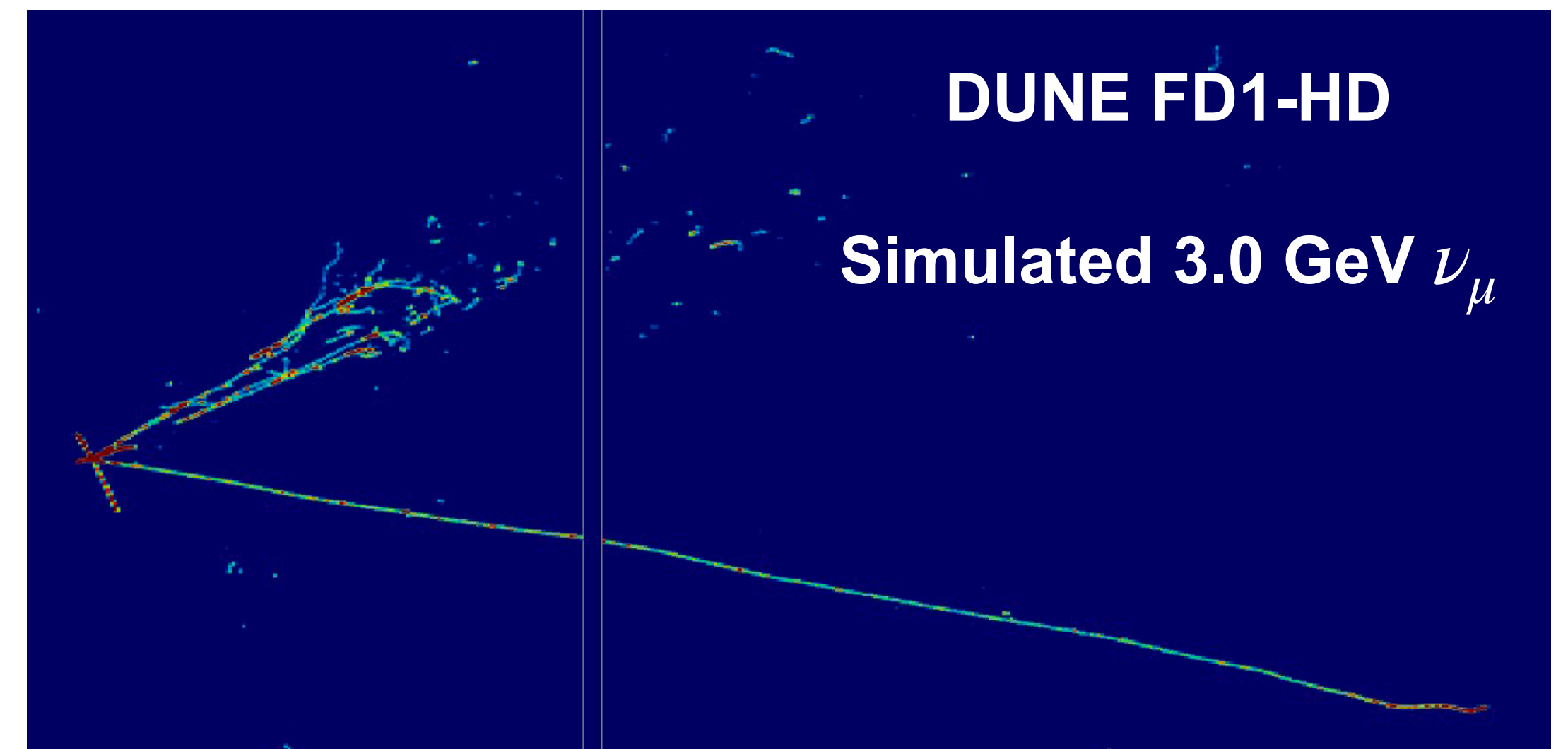


Precision neutrino oscillation measurements,
MeV-scale neutrino physics, BSM searches,

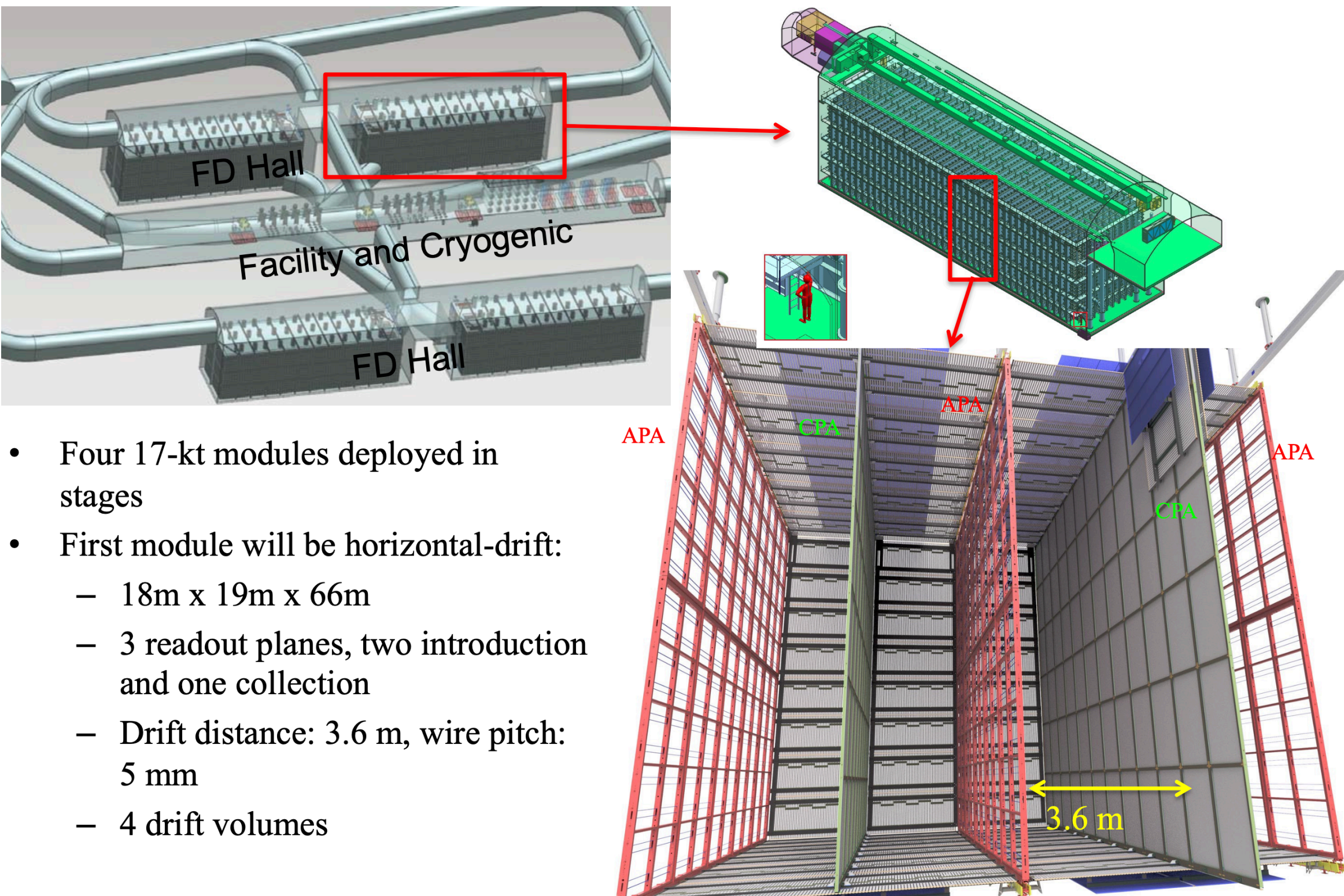
Liquid Argon Time Projection Chamber (LArTPC)



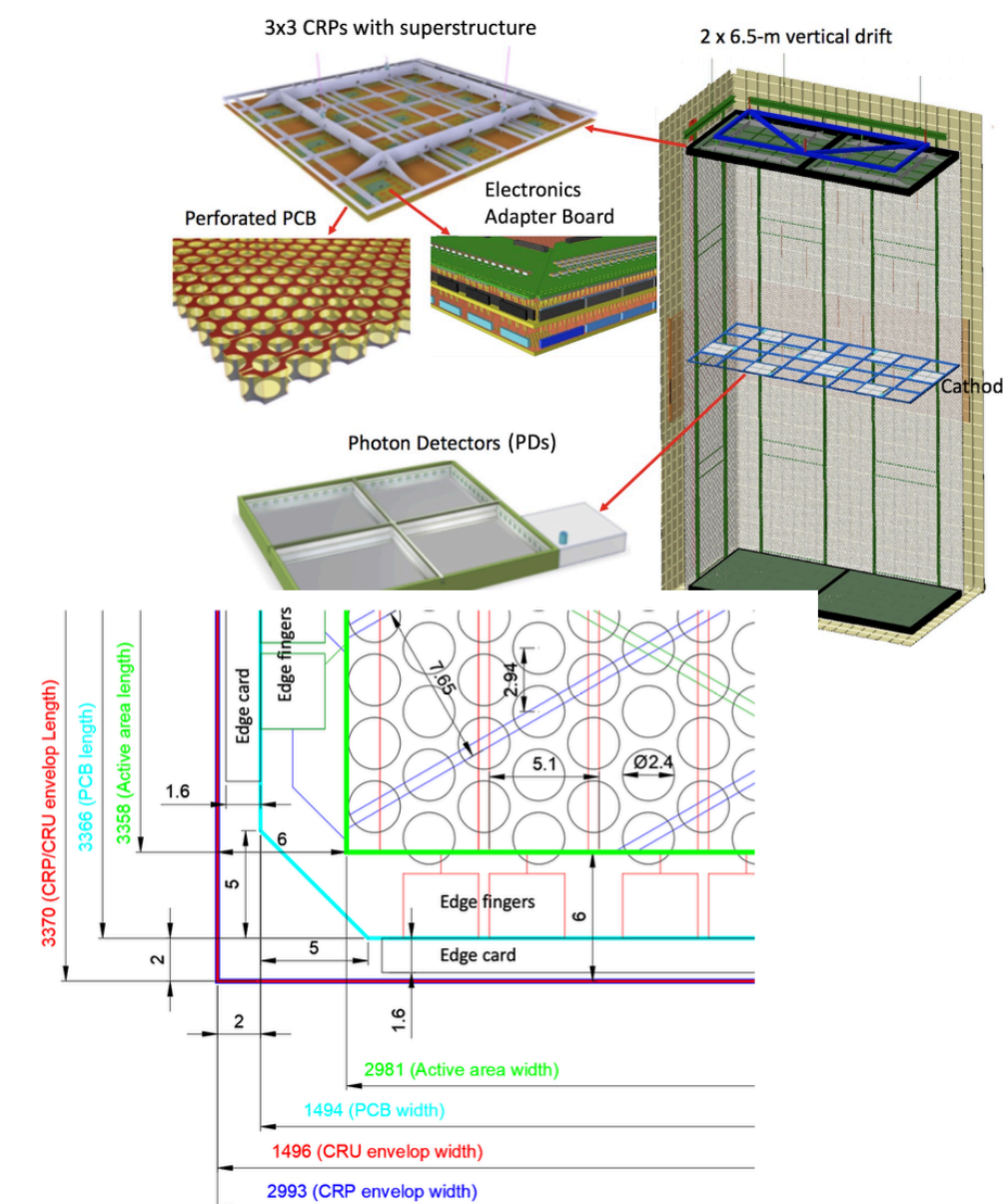
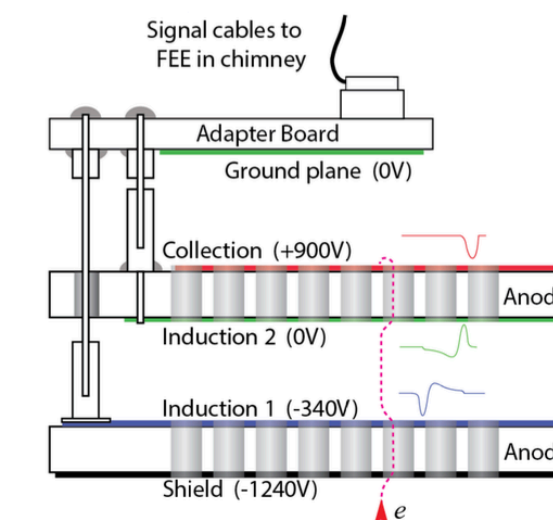
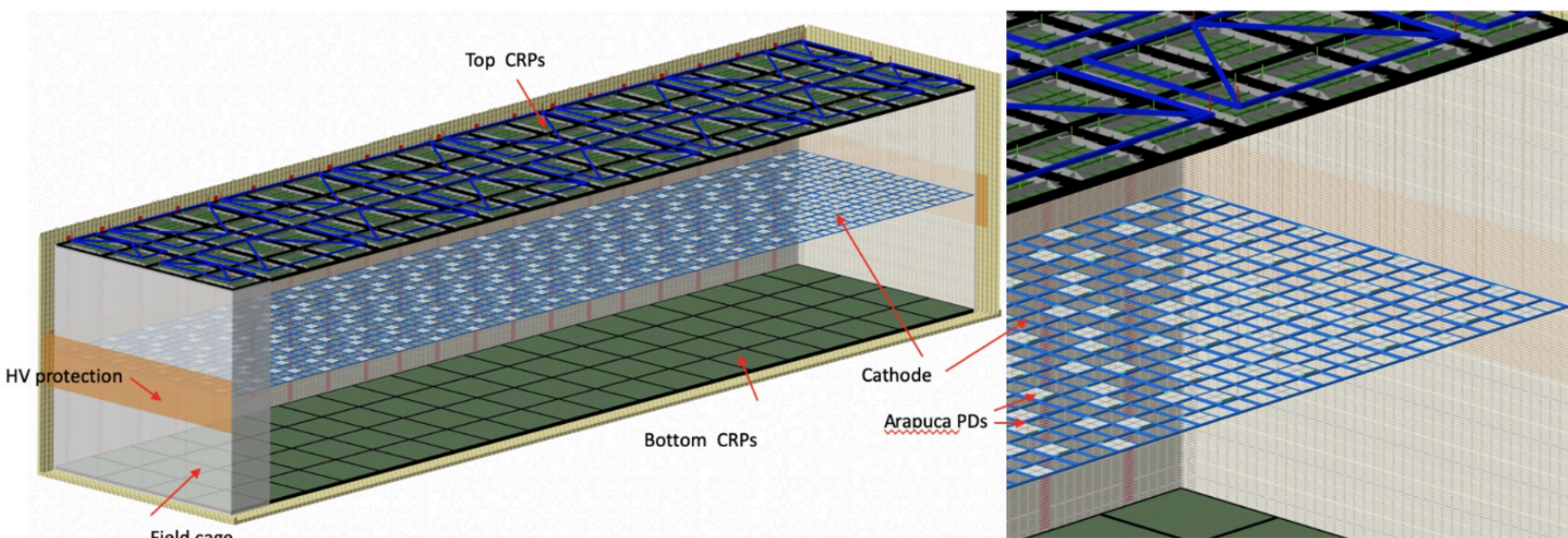
- In the DUNE FD modules, three readout planes collect the ionization charge that is generated when charged particles traverse the liquid argon volume
- Raw detector waveforms are de-convolved to obtain the charge information
- The position of the charge observed in each of the three planes is combined with the drift time to create three views of each interaction



Far Detectors: HD and VD

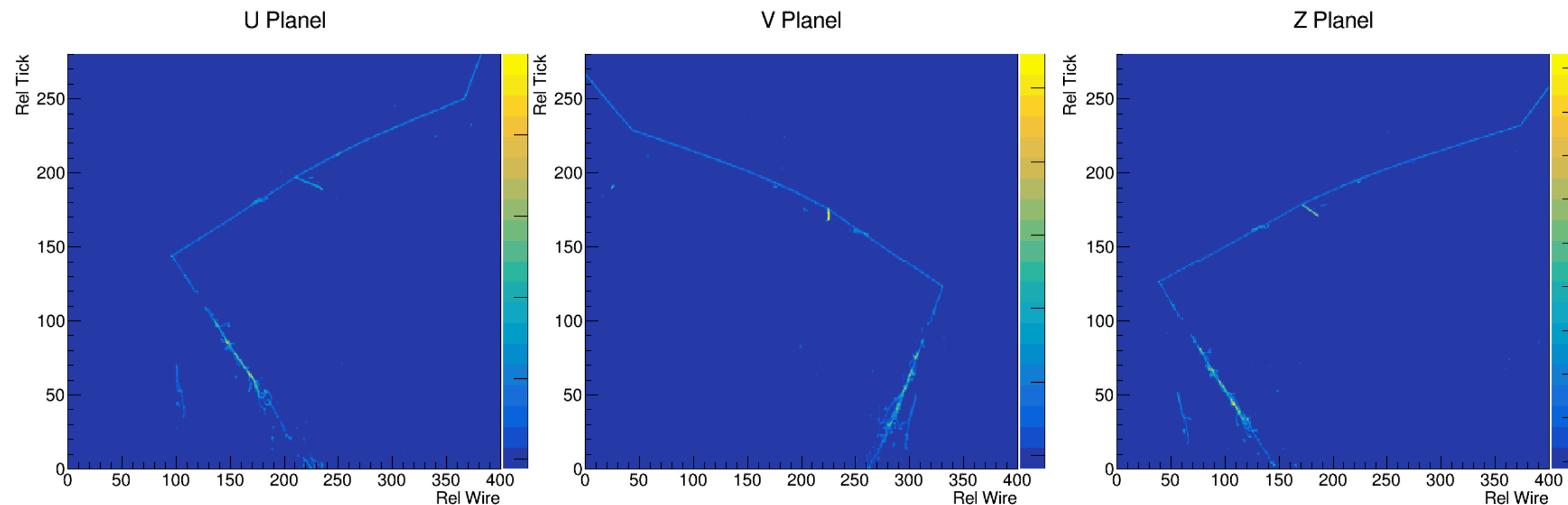


- Four 17-kt modules deployed in stages
- First module will be horizontal-drift:
 - 18m x 19m x 66m
 - 3 readout planes, two introduction and one collection
 - Drift distance: 3.6 m, wire pitch: 5 mm
 - 4 drift volumes

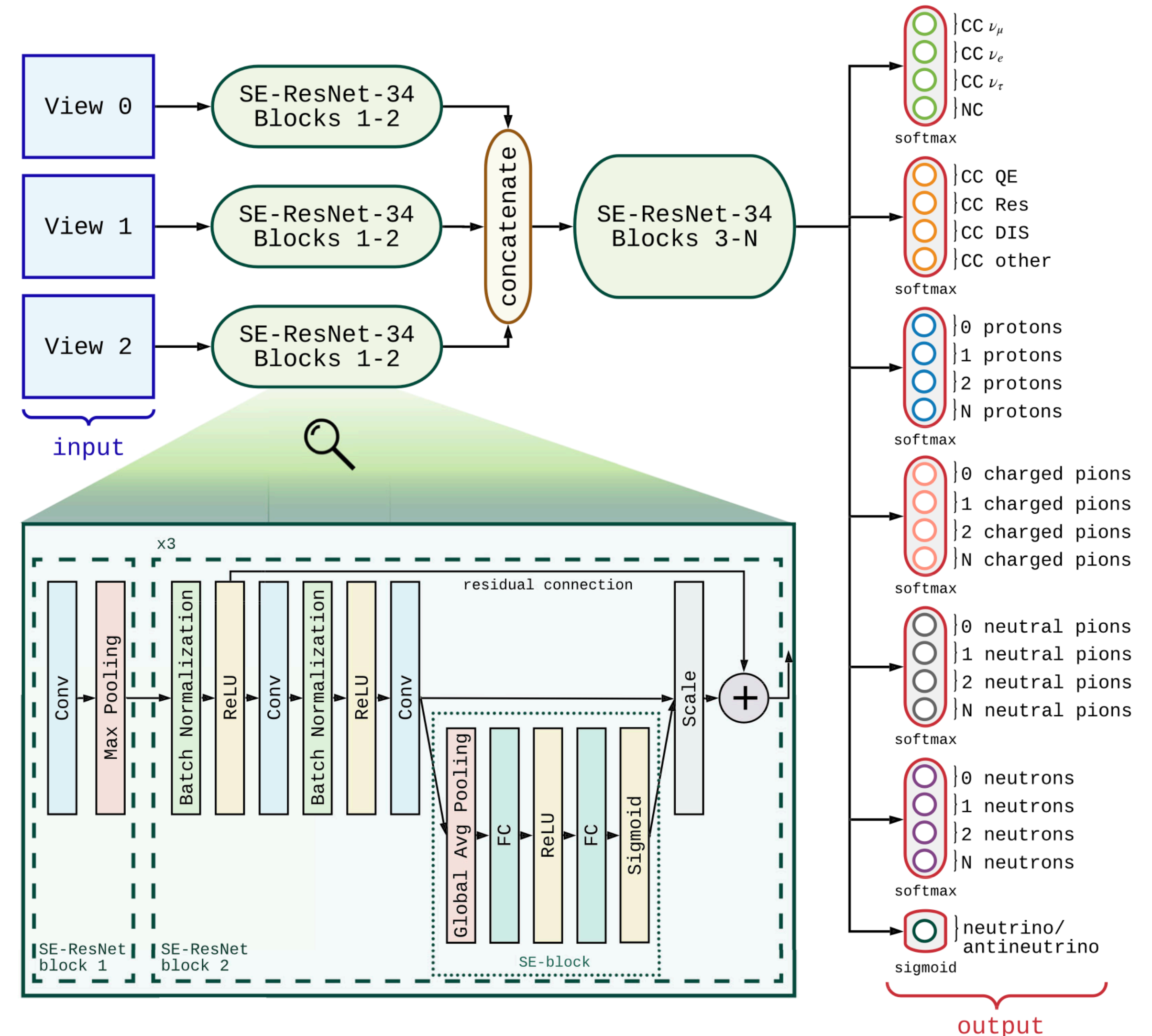


- For the 2nd DUNE FD module, has a vertical drift (VD) path in contrast to HD
- 2 drift volumes, cathode plane on the middle
- Anode: a stack of perforated PCBs with 3 layers of readout etched electrode strips in different orientations
- Modular design allows easy assembly and production. Wires → Strips improves mechanical robustness

CNN for Event ID

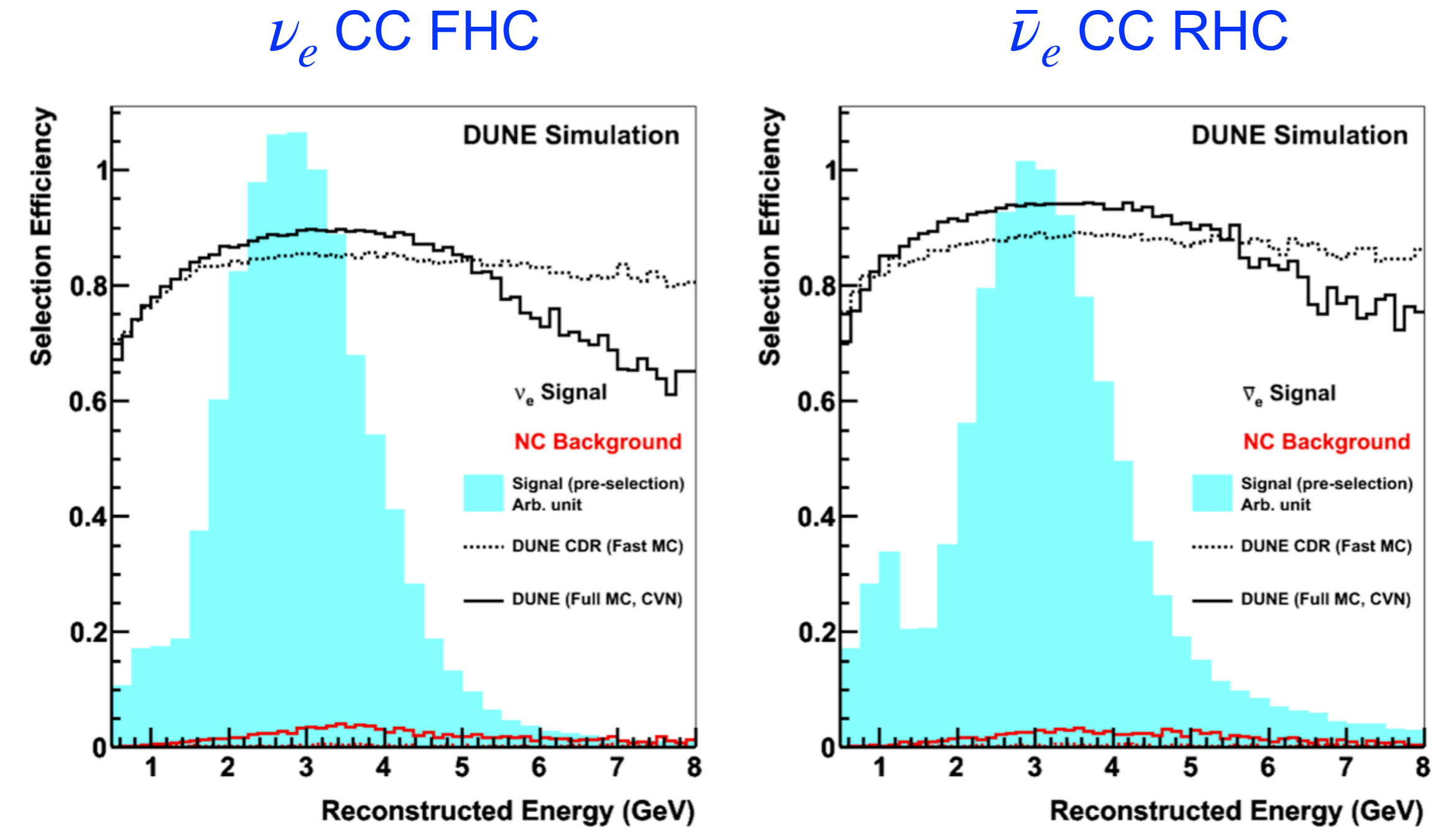
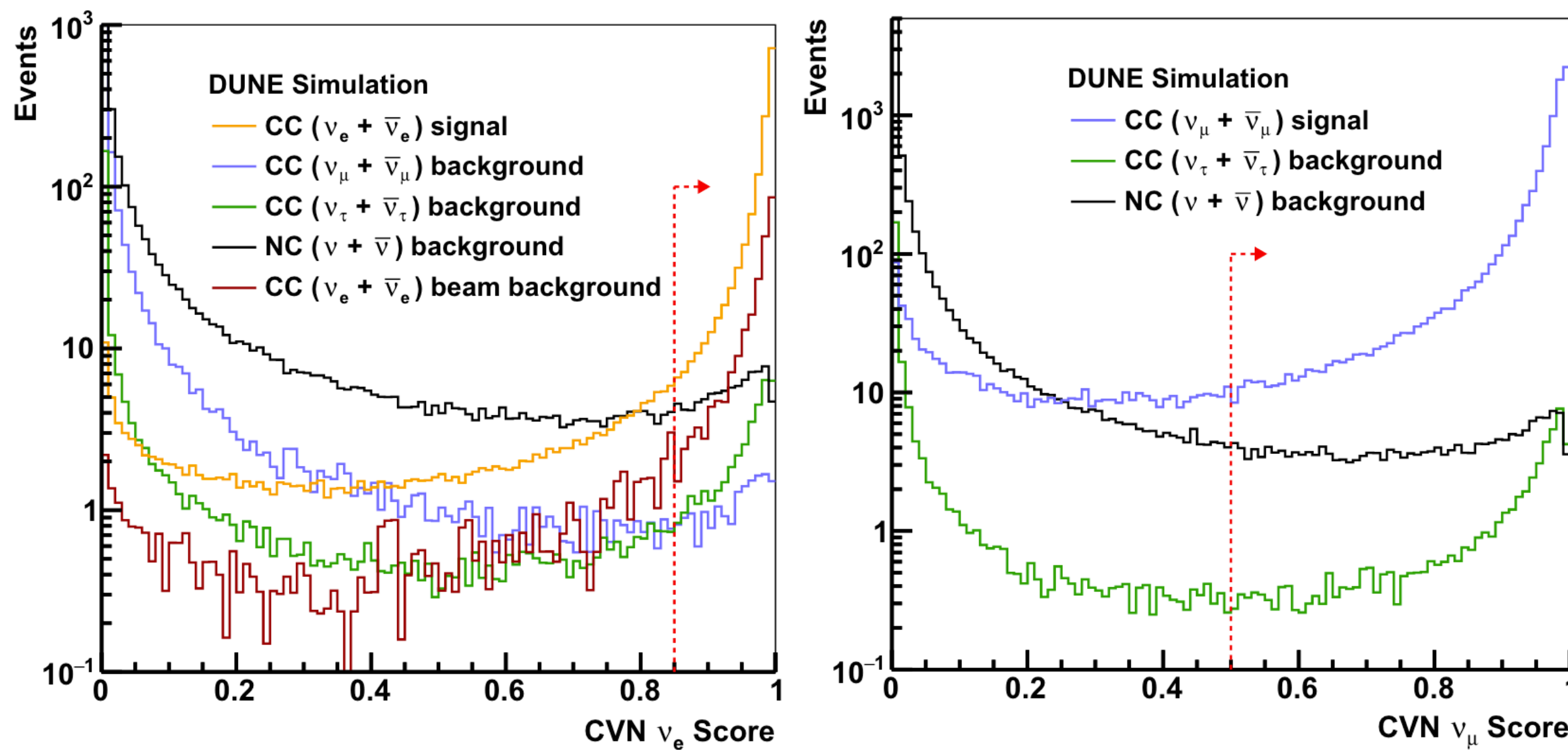


- Uses 3 2D readout images, one for each wire/strip plane, directly as input to a ResNet CNN architecture
- CNN merges information across the 3 planes and uses fully connected layer at the end for neutrino flavor classification or energy regression
- ResNet (residual neural network) is an important feature for DUNE as it allows details of LArTPC images to be propagated further deep



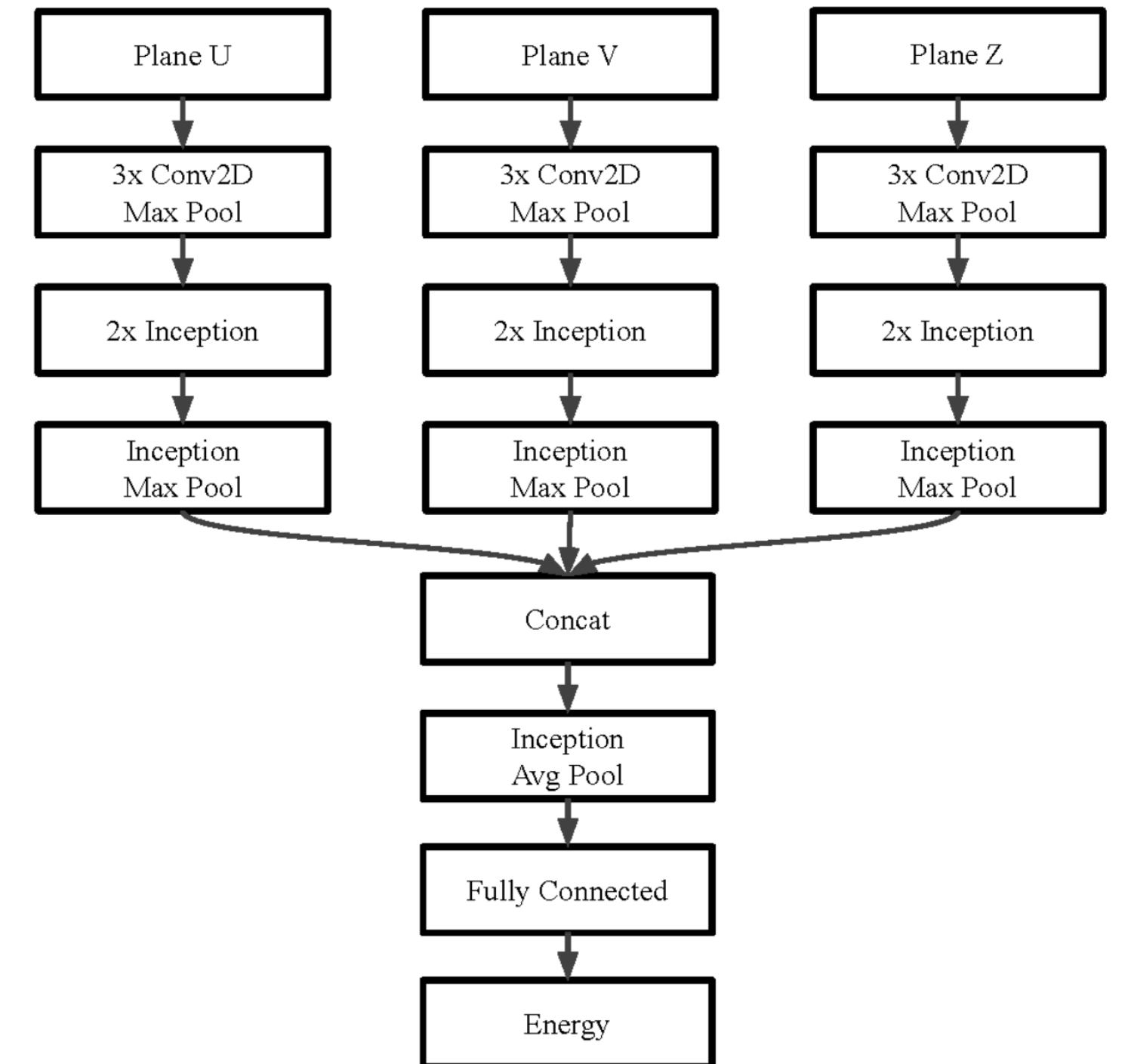
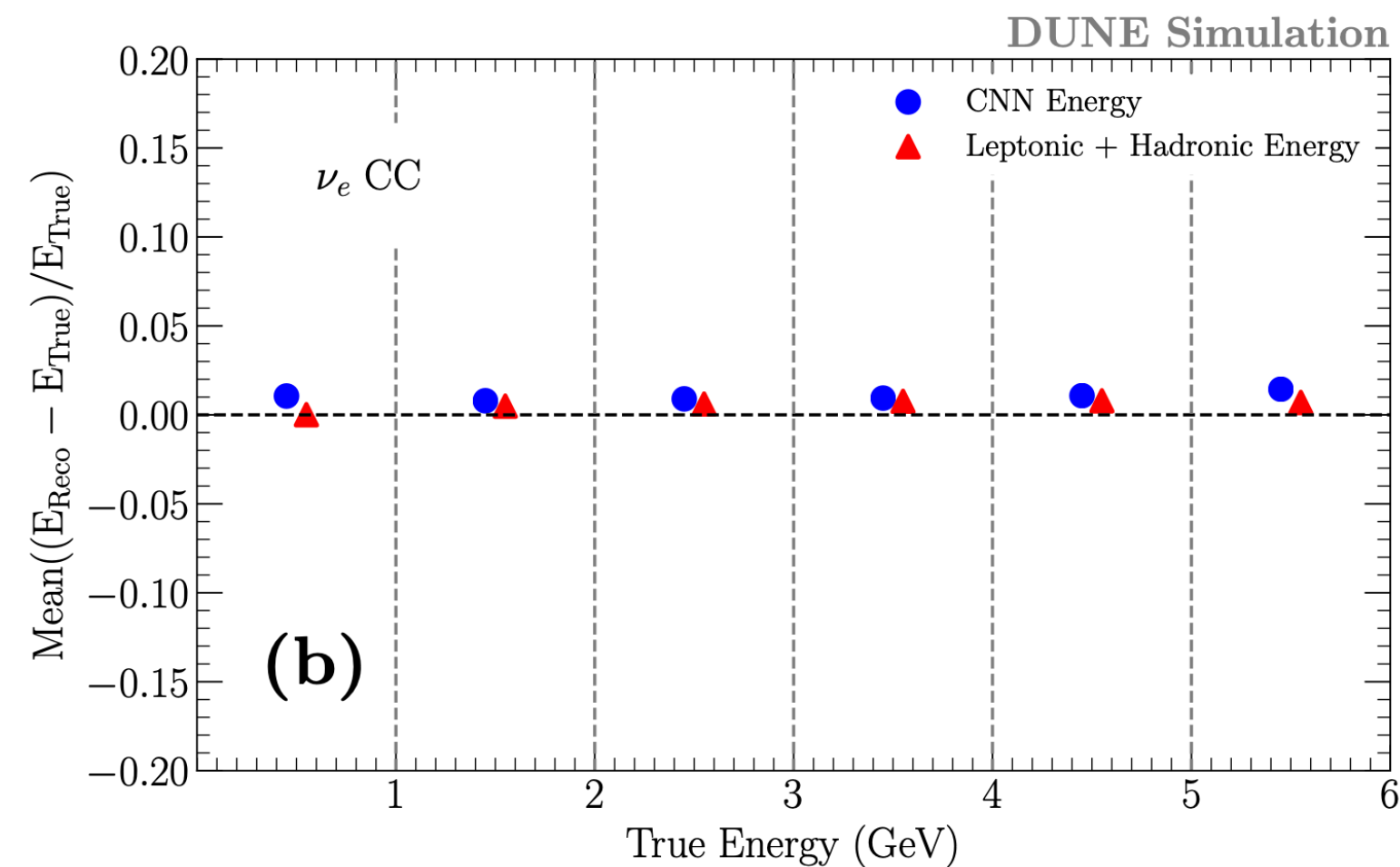
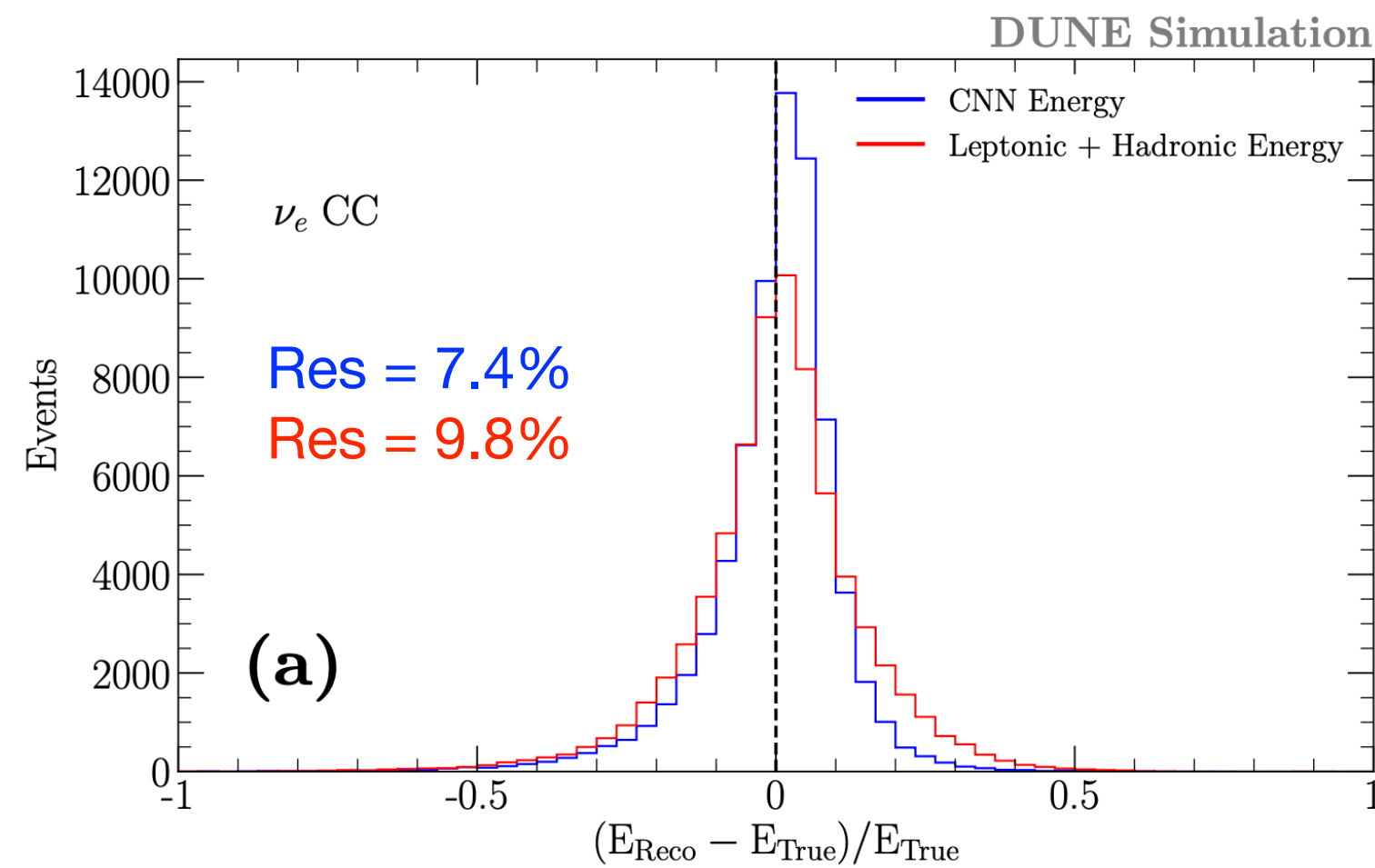
CNN for Event ID

- The primary goal of the CVN is to efficiently and accurately produce event selection of ν_e CC and ν_μ CC in FHC, $\bar{\nu}_e$ CC and $\bar{\nu}_\mu$ CC in RHC.
- The ν_e and ν_μ efficiencies in both FHC and RHC beam modes all exceed 90% in the neutrino flux peak.



CNN Energy and Direction Reconstruction

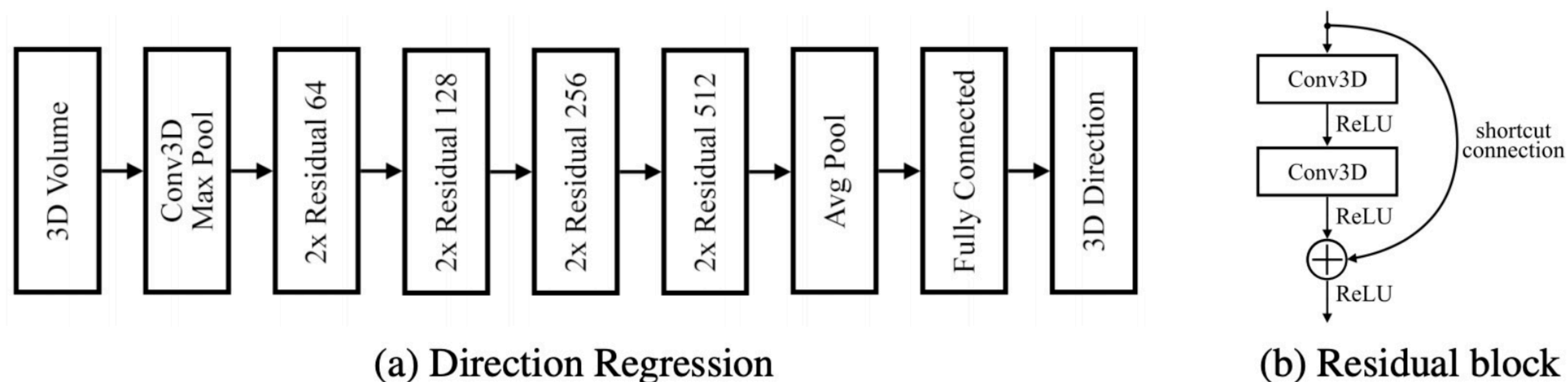
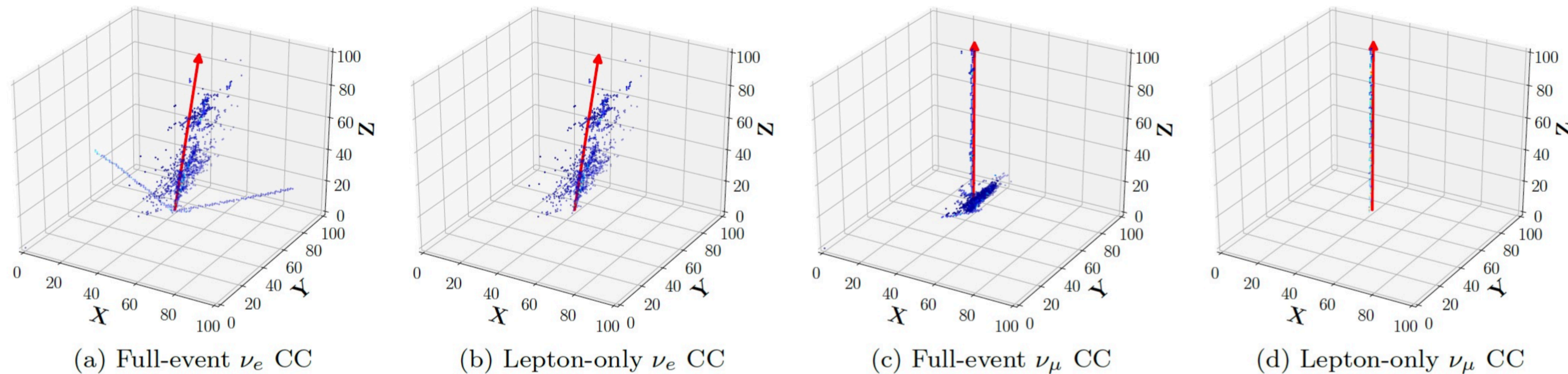
- CNN with linear output regression for event energy, optimizing resolution $(E_{\text{true}} - E_{\text{reco}})/E_{\text{true}}$
- Weighted events by energy to reduce energy dependent bias in the training
- Better resolutions than lepton+hadronic energy method
- Also trained on clustered lepton shower/track pixel maps to predict lepton energy



CNN Energy and Direction Reconstruction

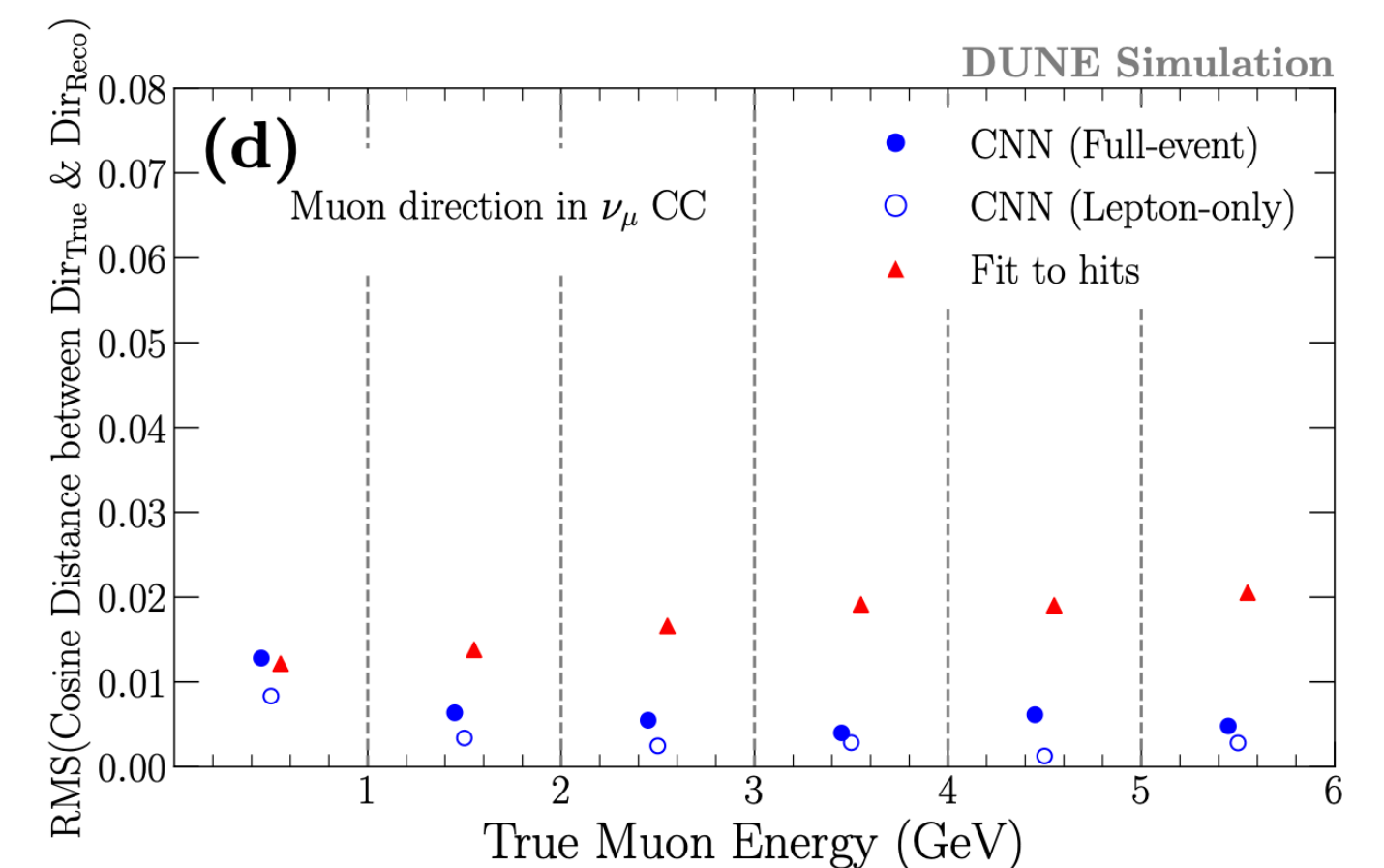
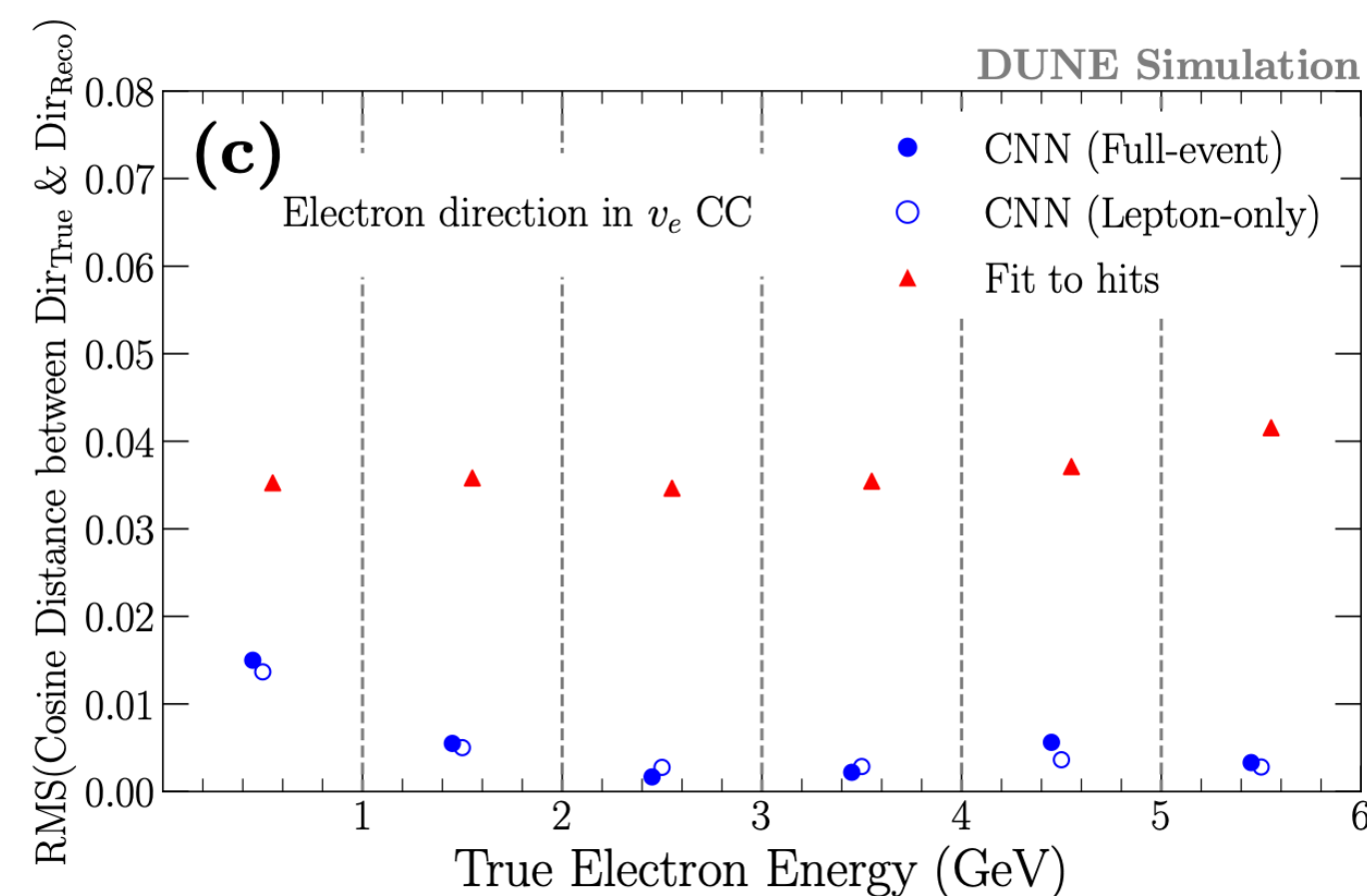
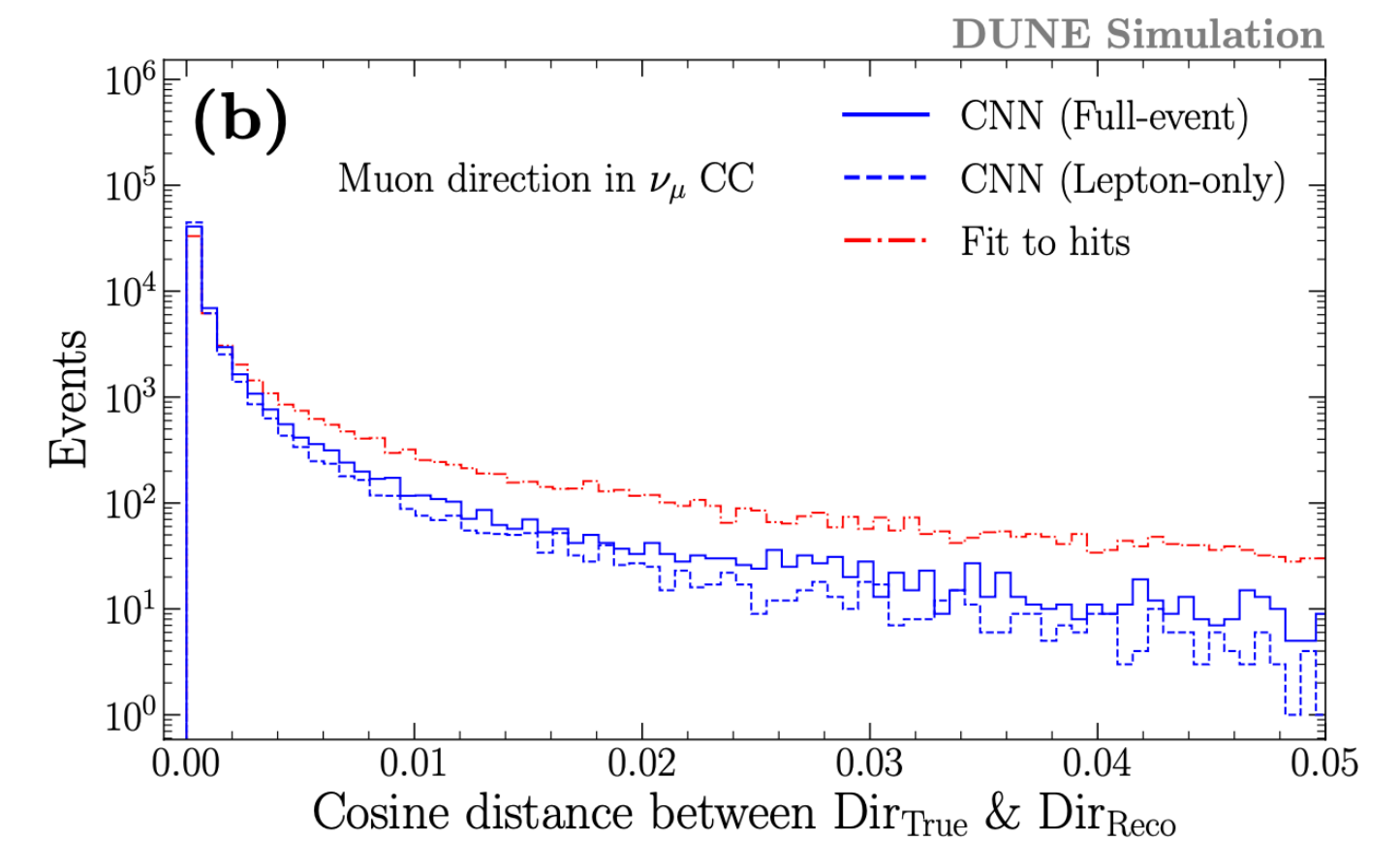
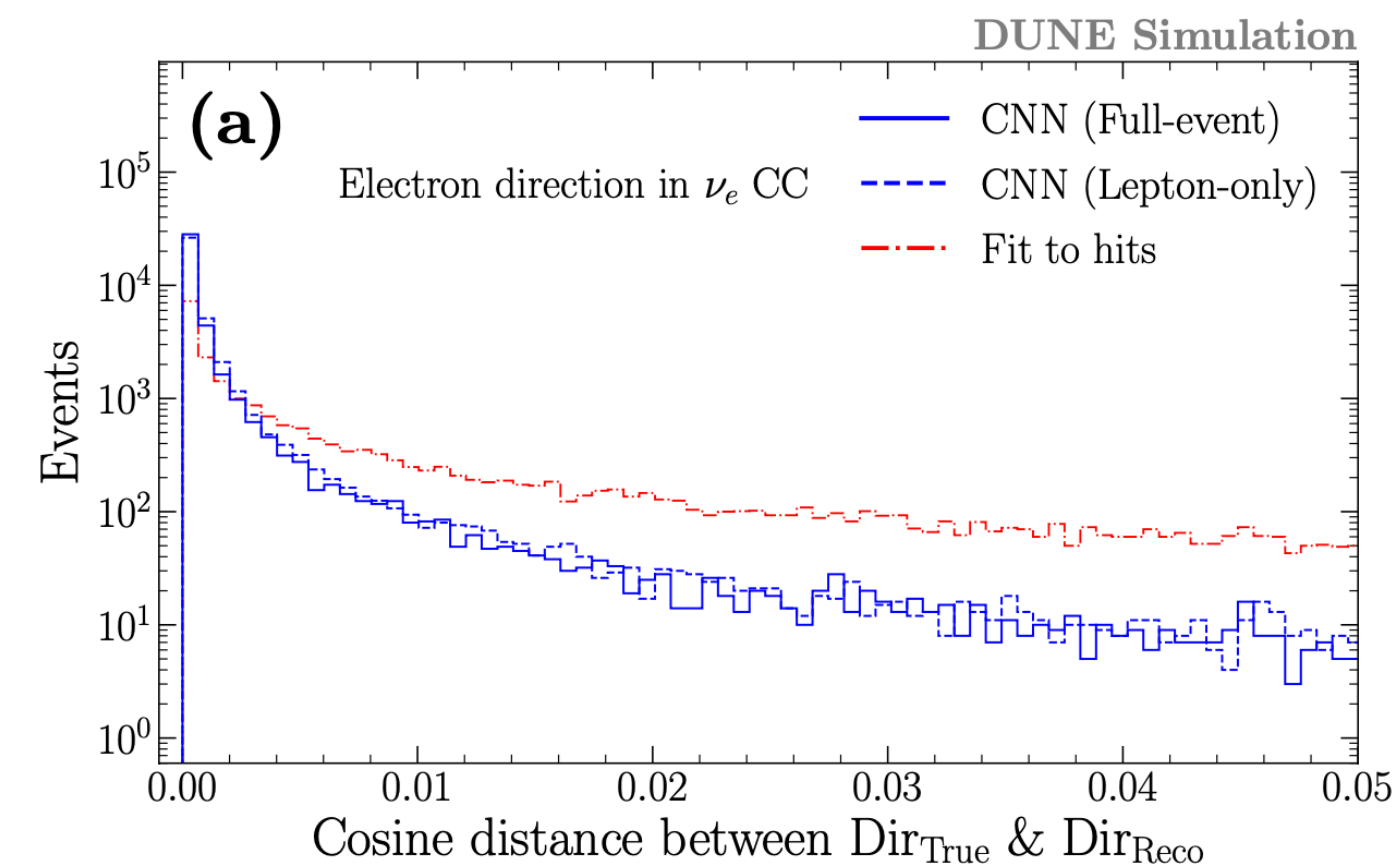
- Direction regression strongly dependent on 3-D information
- A 3D CNN to reconstruct lepton directions
- 3D images constructed from 3 2D detector images

arXiv:2012.06181



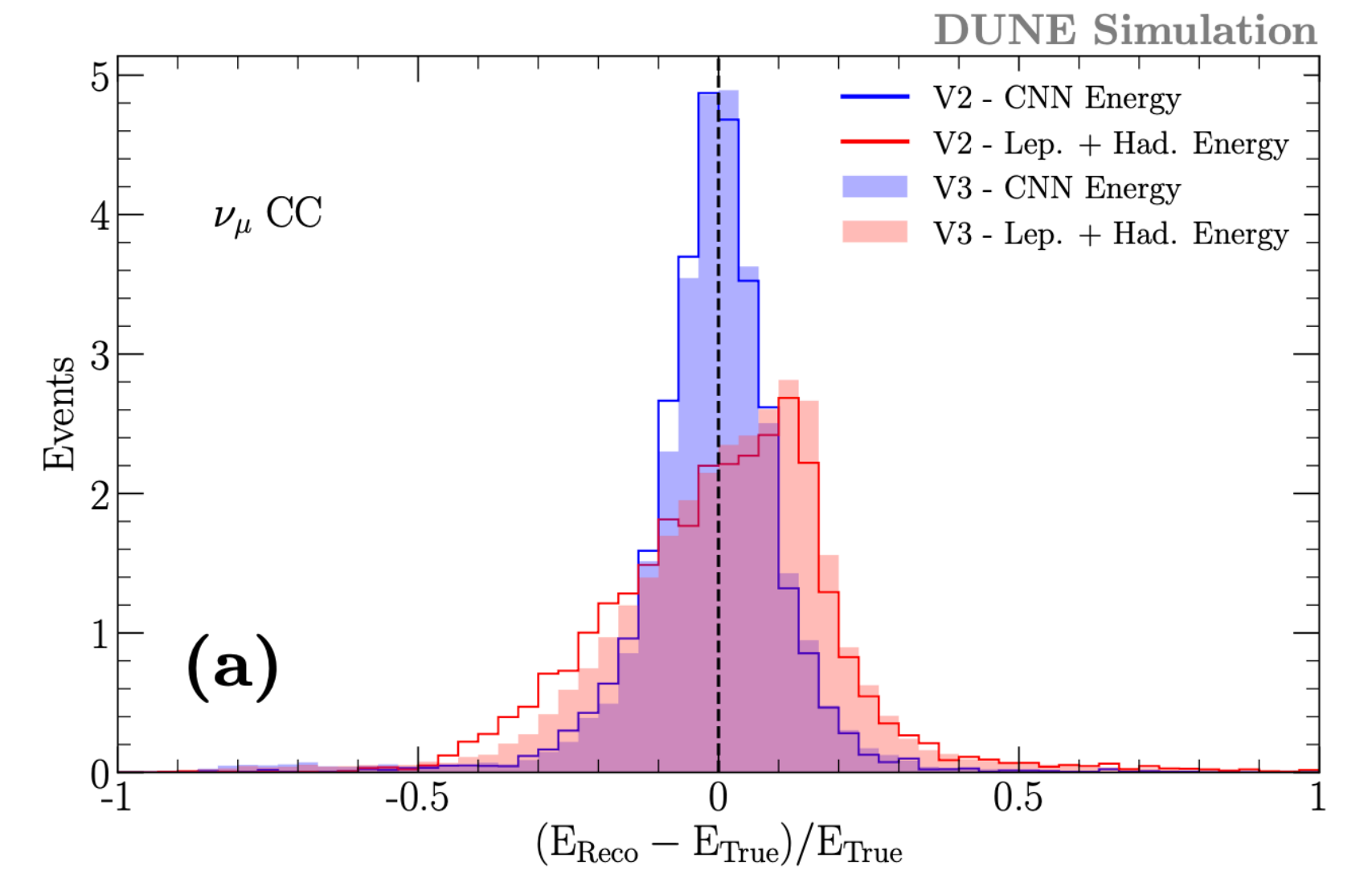
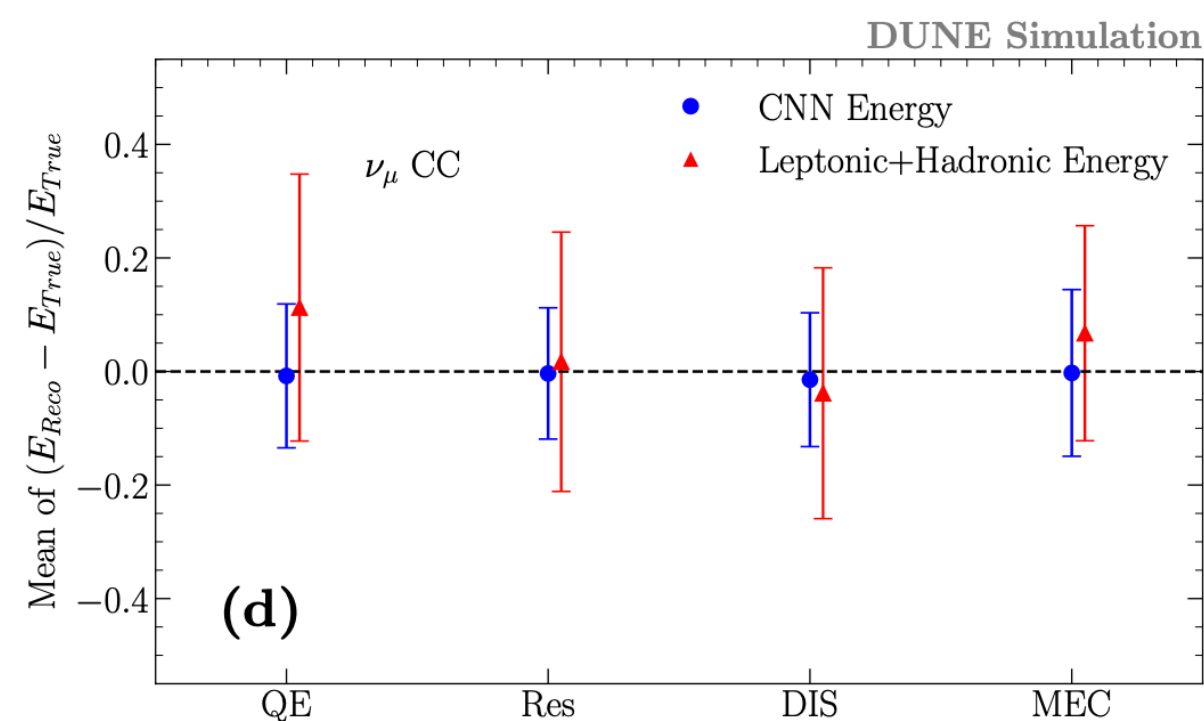
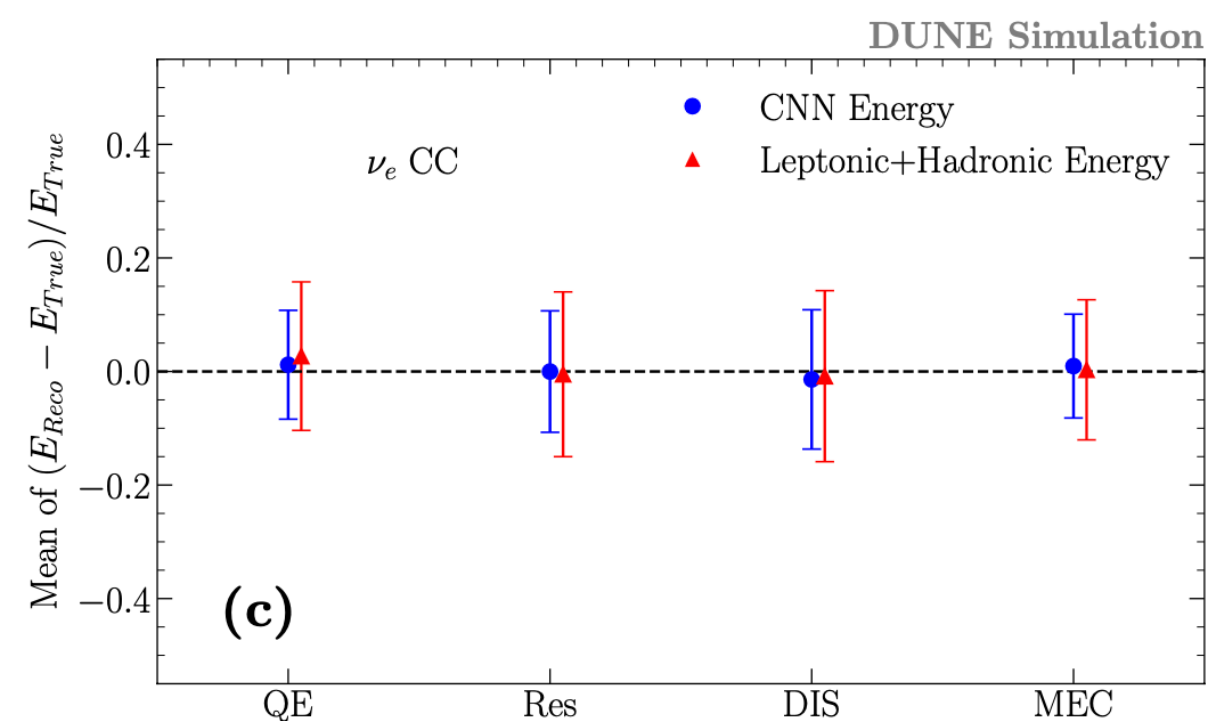
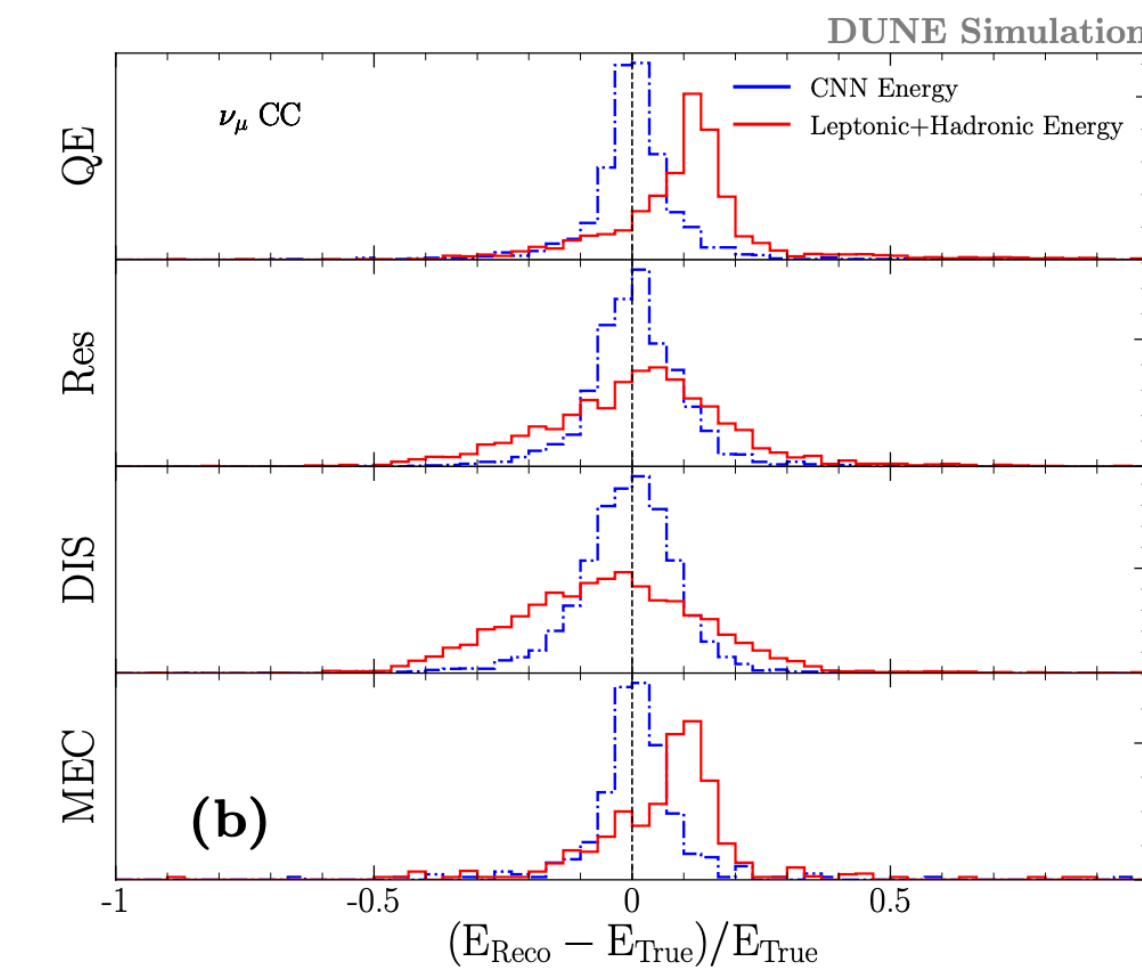
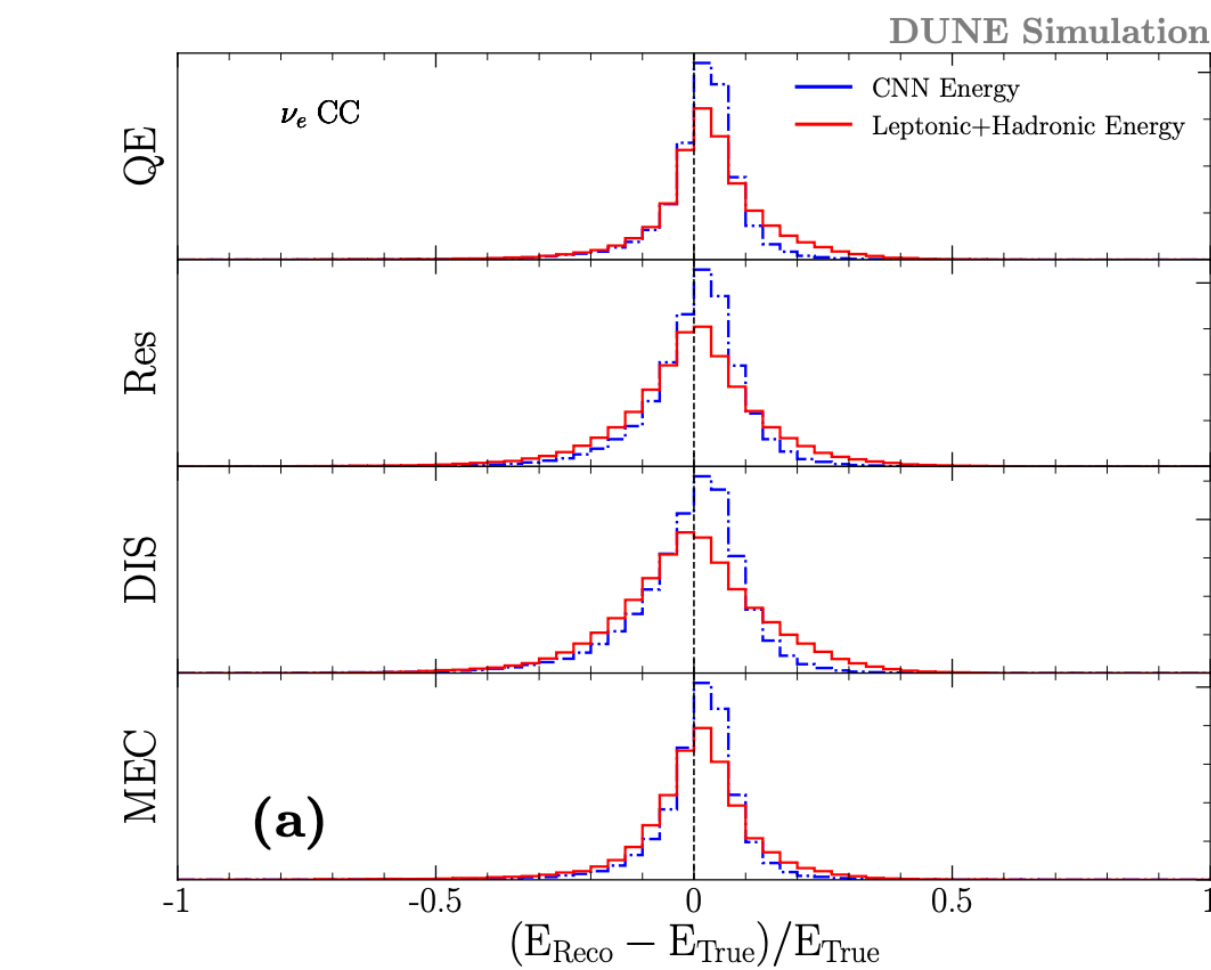
CNN Energy and Direction Reconstruction

- 3D CNN outperforms the traditional method (Fit to hits) with better electron and muon resolutions in all energy regions
- 3D CNN trained with full-event pixel maps shows comparable performance to that trained with clustered lepton track/shower pixel maps
→ extract particle kinematics without clustering/tracking

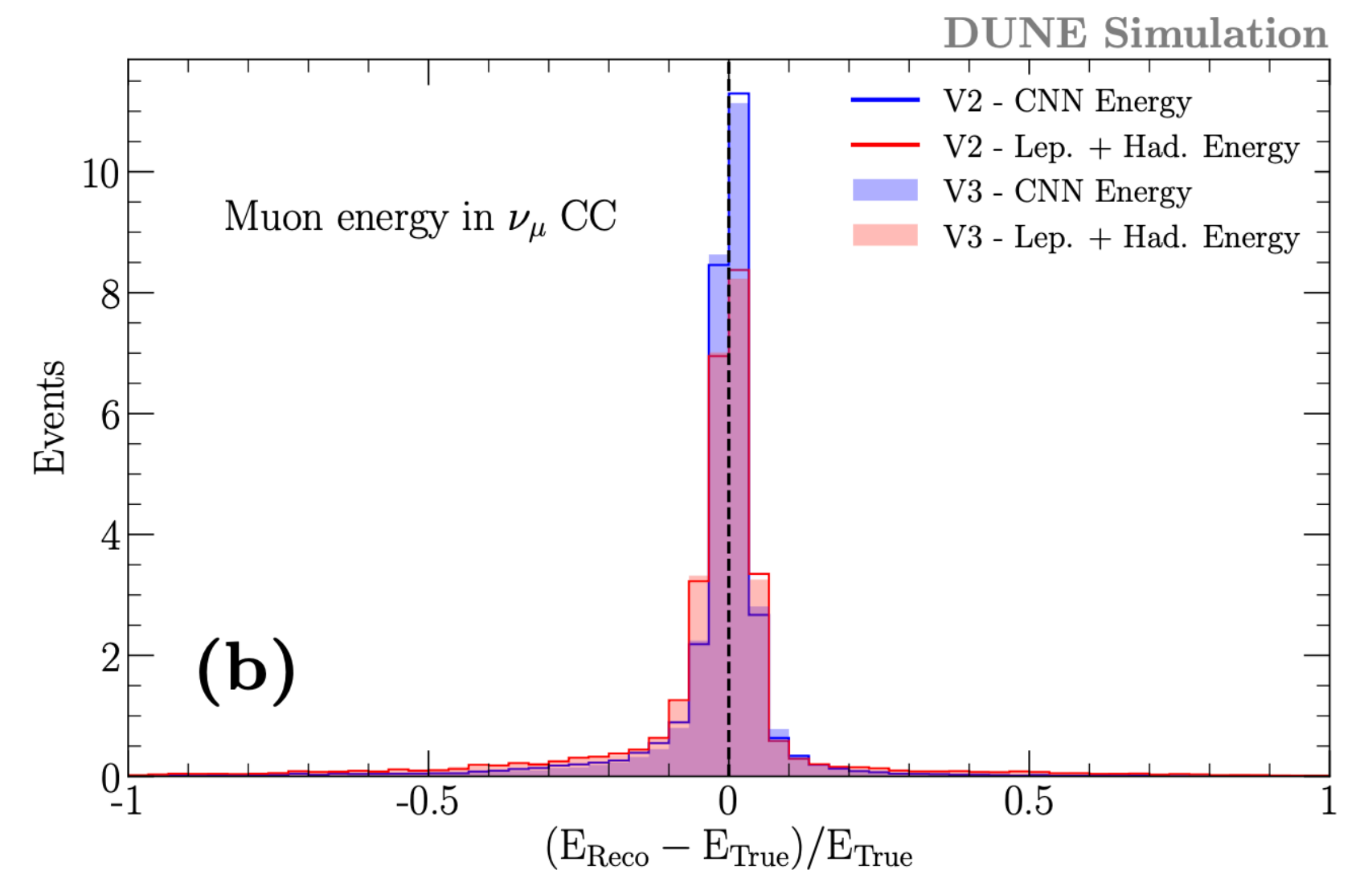


CNN Robustness Tests

- CNN show robustness against neutrino interaction modes
- GENIE versions have small effects on CNNs

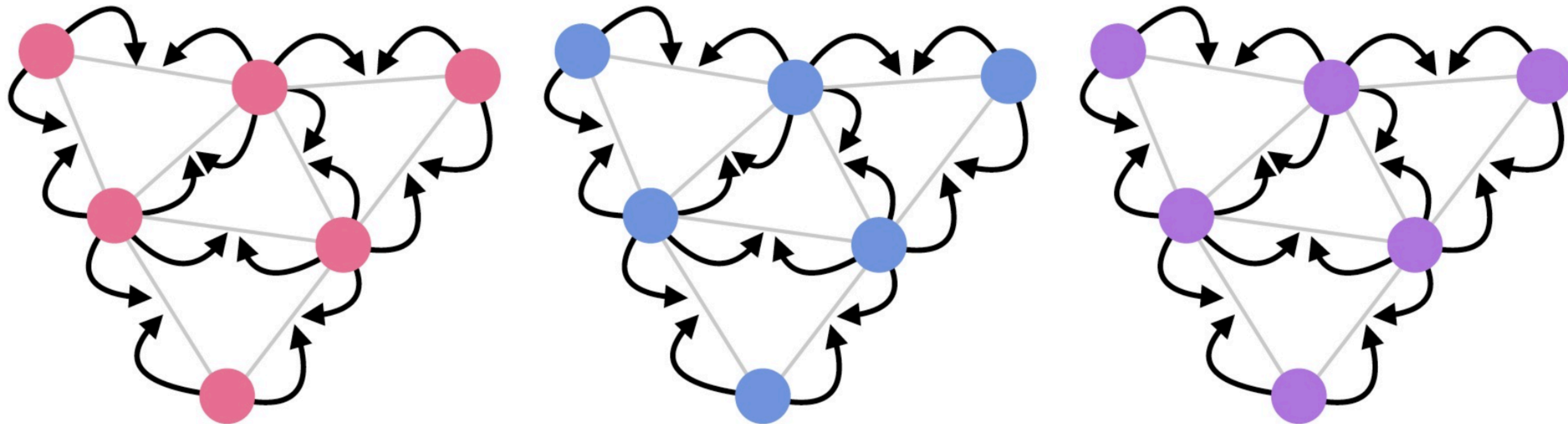


Trained on GENIE 2, tested with GENIE 3



Graph Neural Networks

- Graph Neural Network (GNN)
 - Define input data as a graph represented by nodes and edges
 - Nodes are generalized as quantized objects with some arbitrary set of features
 - Edges describe the relationships between nodes
 - Perform convolutions on nodes and edges rather than the entire pixel map in CNN → speed up NN training
 - Output can be either regression or classification



Graph Neural Networks

- Successfully reconstruct LArTPC showers/tracks with GNN in ExtExa.TrkX project (a collaboration developing GNN reconstruction for HEP)
- Implementing under DUNE context

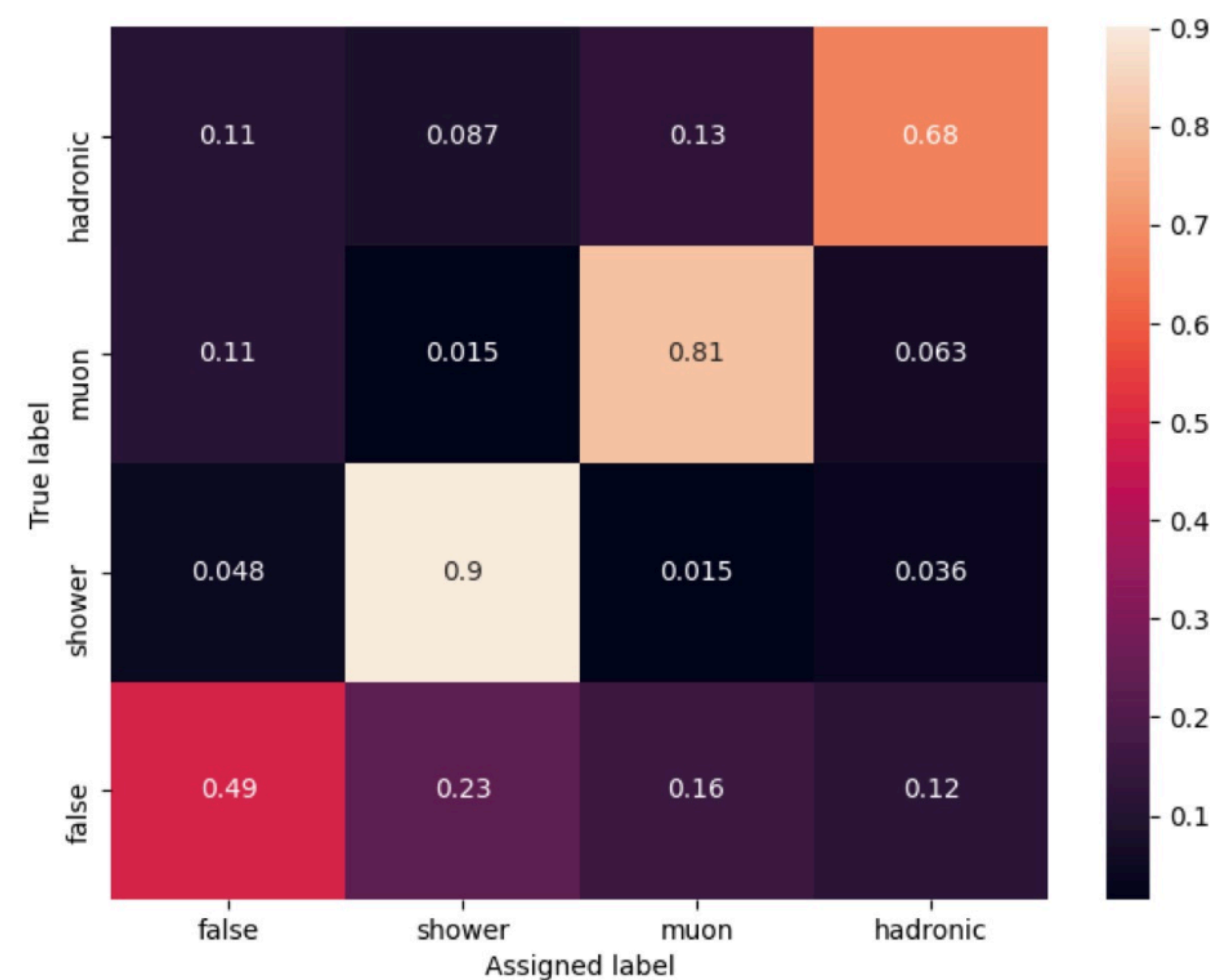


Figure 4. Confusion matrix showing the overlap of true and reconstructed edge labels.

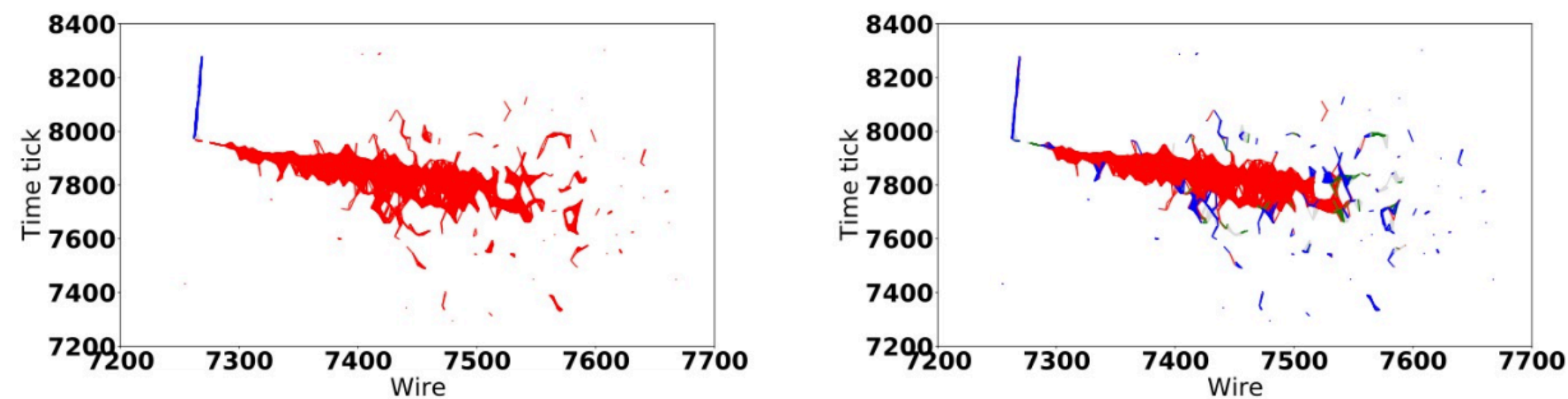
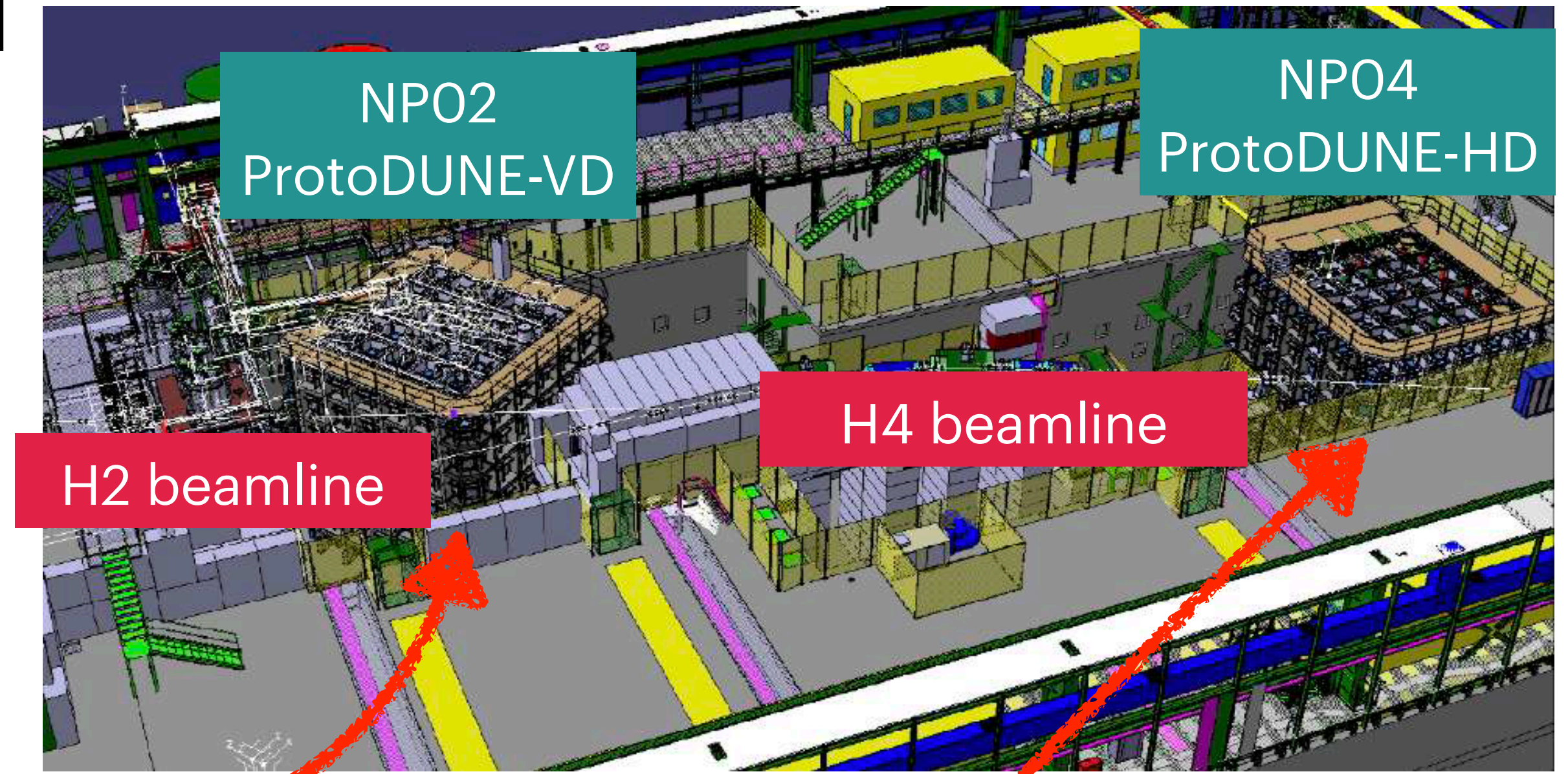


Figure 2. Example graph of a ν_e interaction (left: ground truth, right: model output). Shower-like edges are drawn in red, hadronic edges are drawn in blue, muonic edges are drawn in green and false edges are drawn in grey.

Jeremy Hewes, Adam Aurisano etc., EPJ Web of Conferences 251, 03054 (2021)

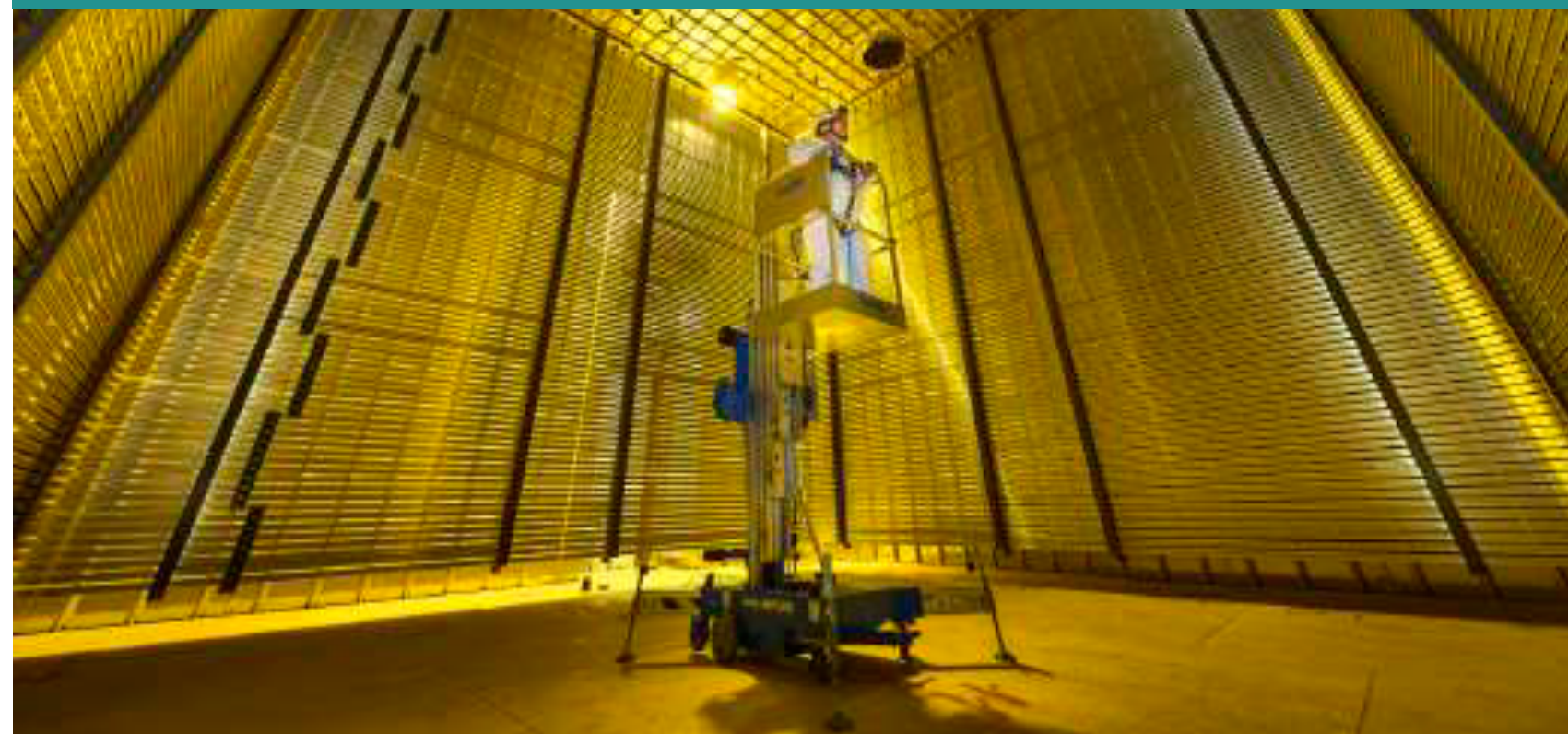
ProtoDUNE at CERN

- ProtoDUNE-HD and VD are DUNE's large scale prototypes (~1 kton-scale) of its far detector modules at CERN Neutrino Platform.
 - Exposed to test beams at CERN, momentum-dependent beam composition contains $e, K^\pm, \mu, p, \pi^\pm$.



Dual Phase (DP) → Vertical Drift (VD)

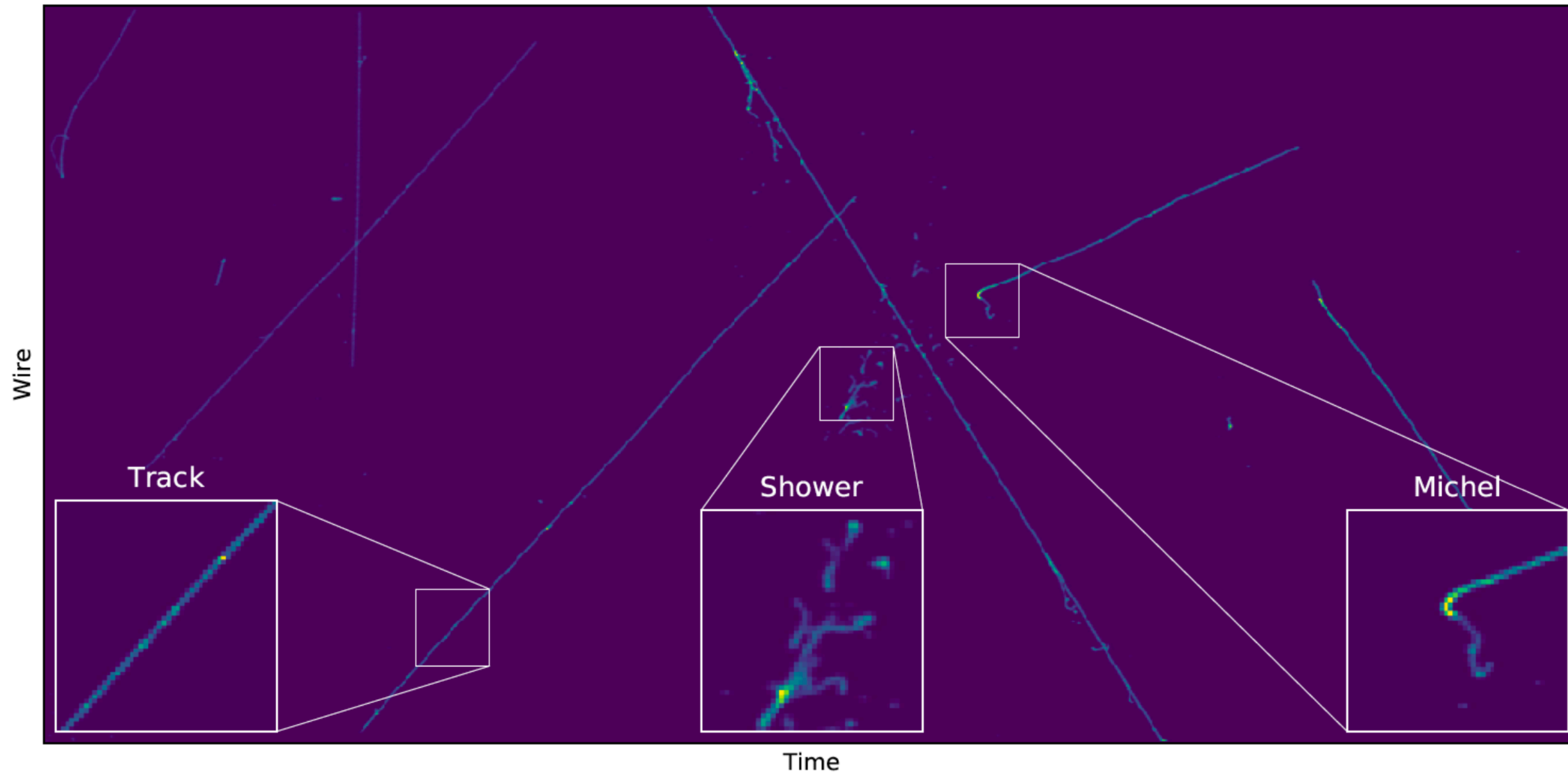
- LArTPC 6 m vertical drift + charge amplification in gas Ar + photon detection
- Cosmic-muon data in 2019 - 2020
- VD is now under commissioning for next year's running (2025)



ProtoDUNE Single Phase (SP) → Horizontal Drift (HD)

- LArTPC 3.6 m horizontal drift + photon detection
- Beam data taken in 2018 & cosmic data in 2018 - 2020
- HD has taken beam and cosmic data this year (2024)

CNN for shower/track separation in PDSP

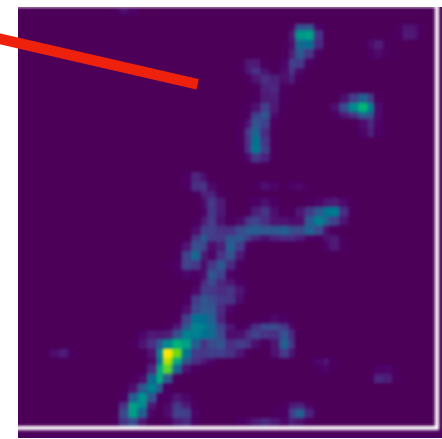
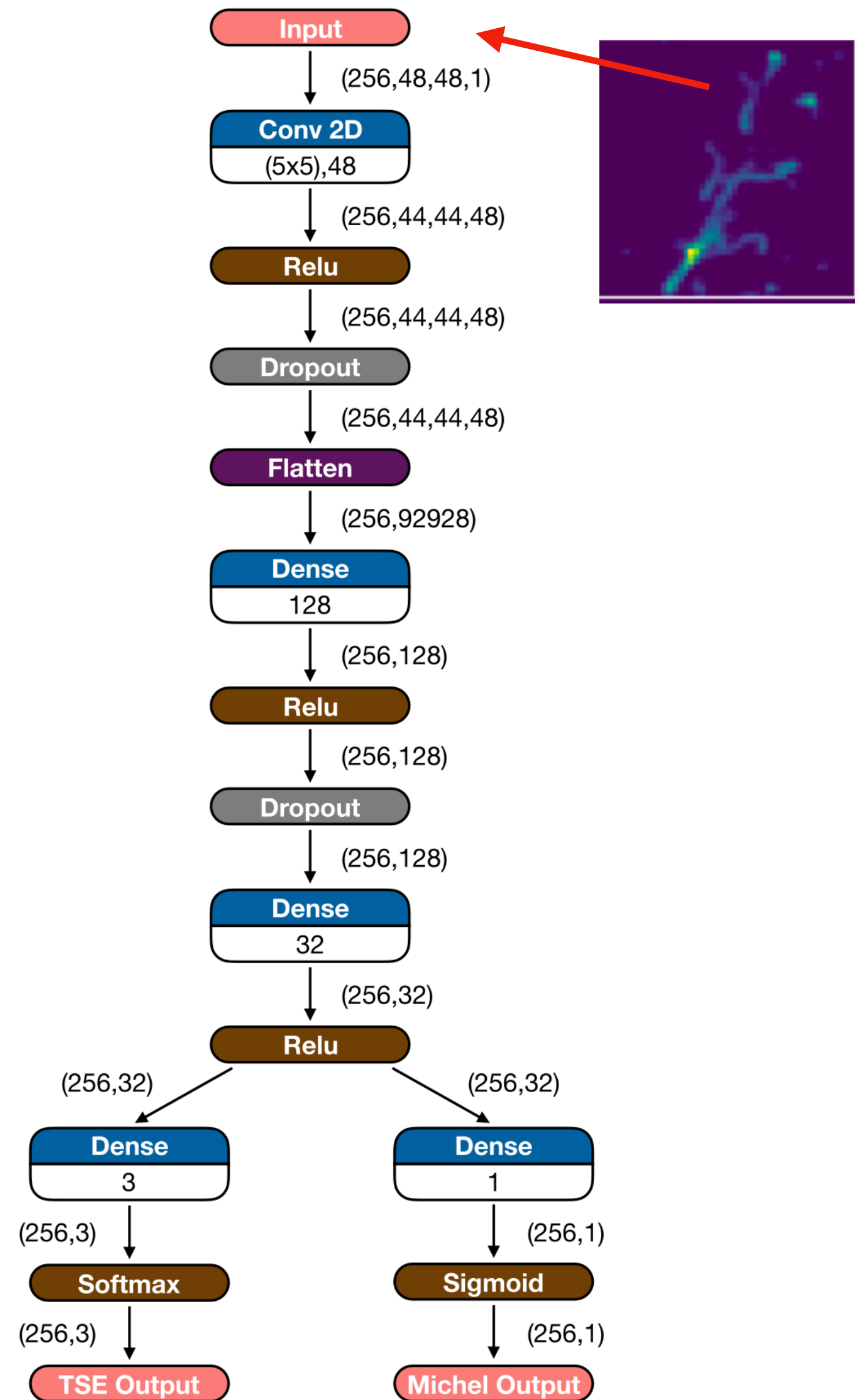


Use CNN to classify energy deposits (hits) from Shower, Track and Michel electrons

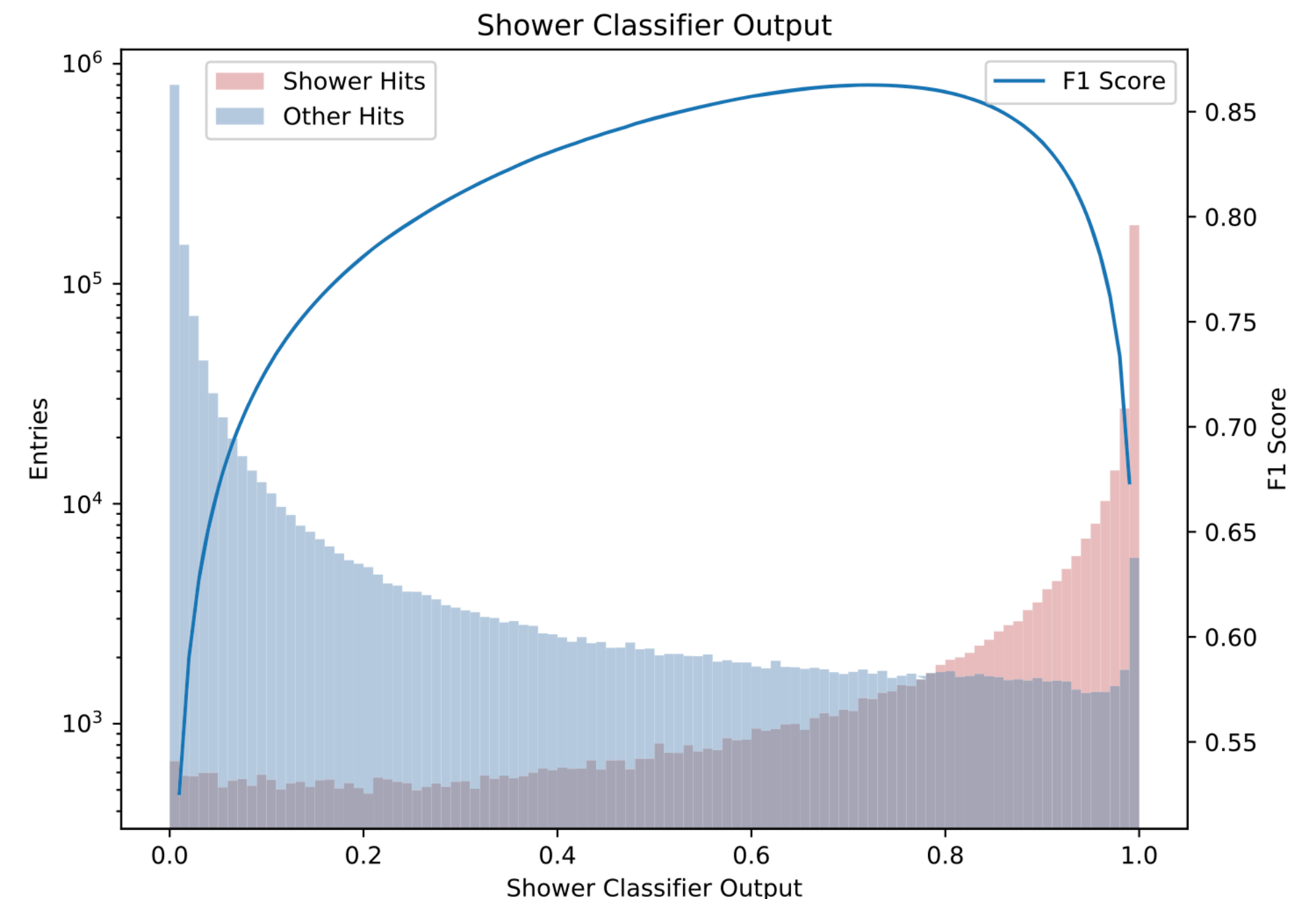
- Showers: Energy deposit pattern caused by electron, gamma, etc
- Tracks: Energy deposit pattern caused by muon, pion, etc
- Michel electrons: Low energy electron from muon decays

Can be used in clustering, PID, etc

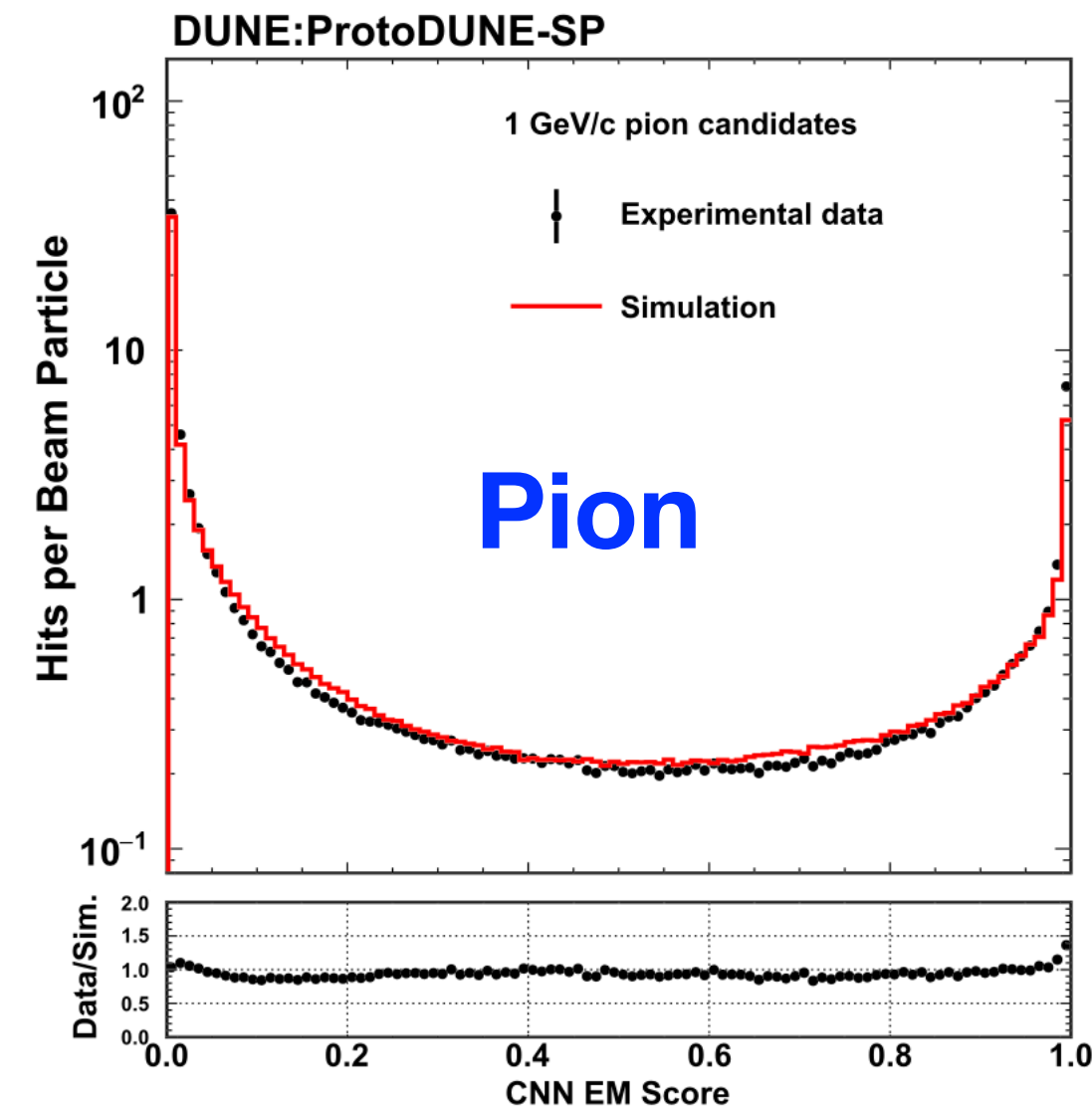
CNN for shower/track separation in PDSP



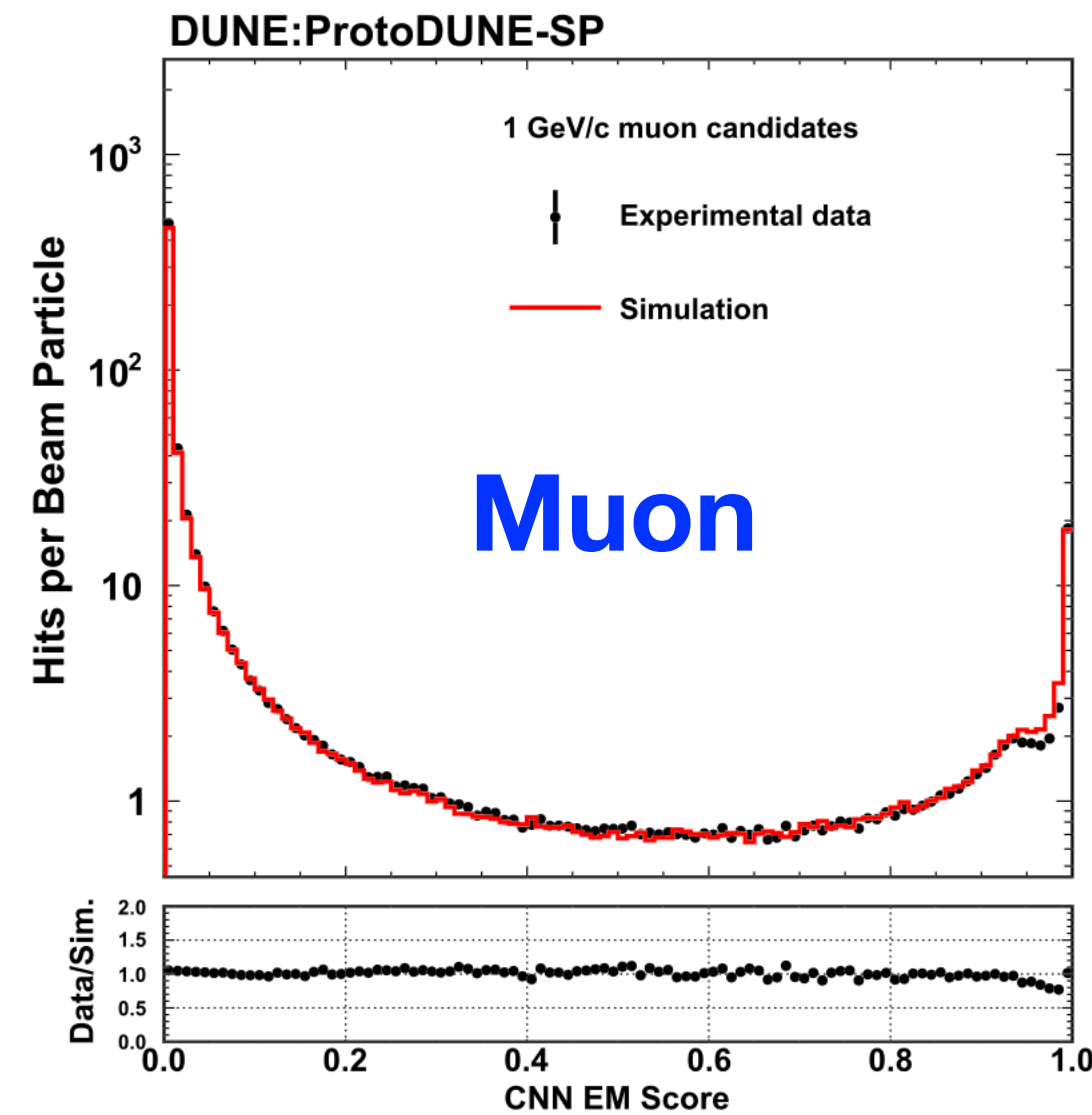
- The inputs are 48x48 pixel maps centered on the reconstructed hit object to be classified
- A single convolutional layer is used to extract feature maps from the images
- These are processed by two dense (fully connected) layers before being split into two branches which classify the images
 - Output: Is the type of hits: from shower? Track? Michel electron?



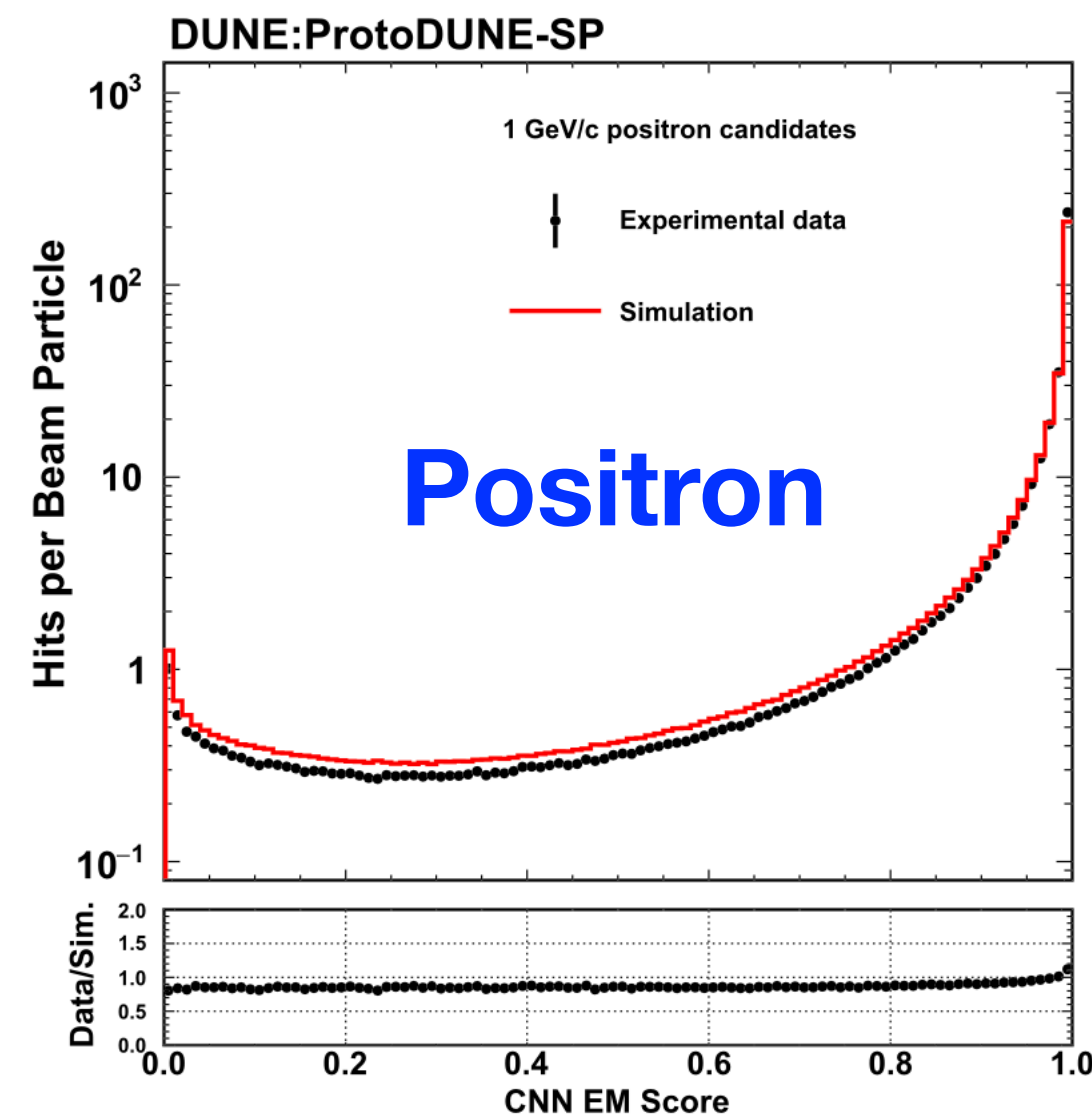
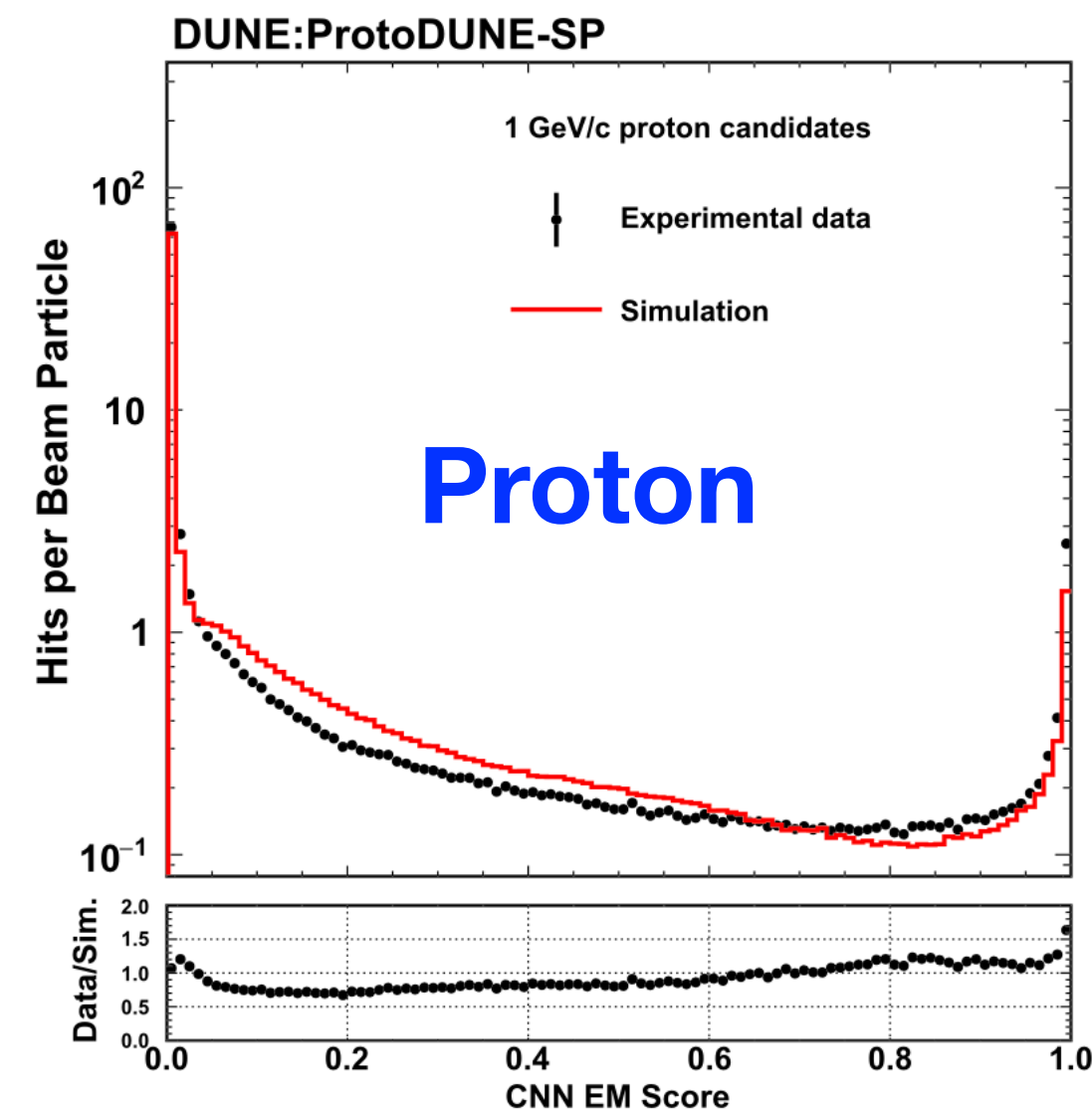
CNN for shower/track separation in PDSP



(a) pion



(b) muon



- Test shower classifier scores for different particle species in the ProtoDUNE-SP
- Hit classification used in clustering and PID
- Reasonable Data/MC agreement

Eur.Phys.J.C 82 (2022) 10, 903

Summary

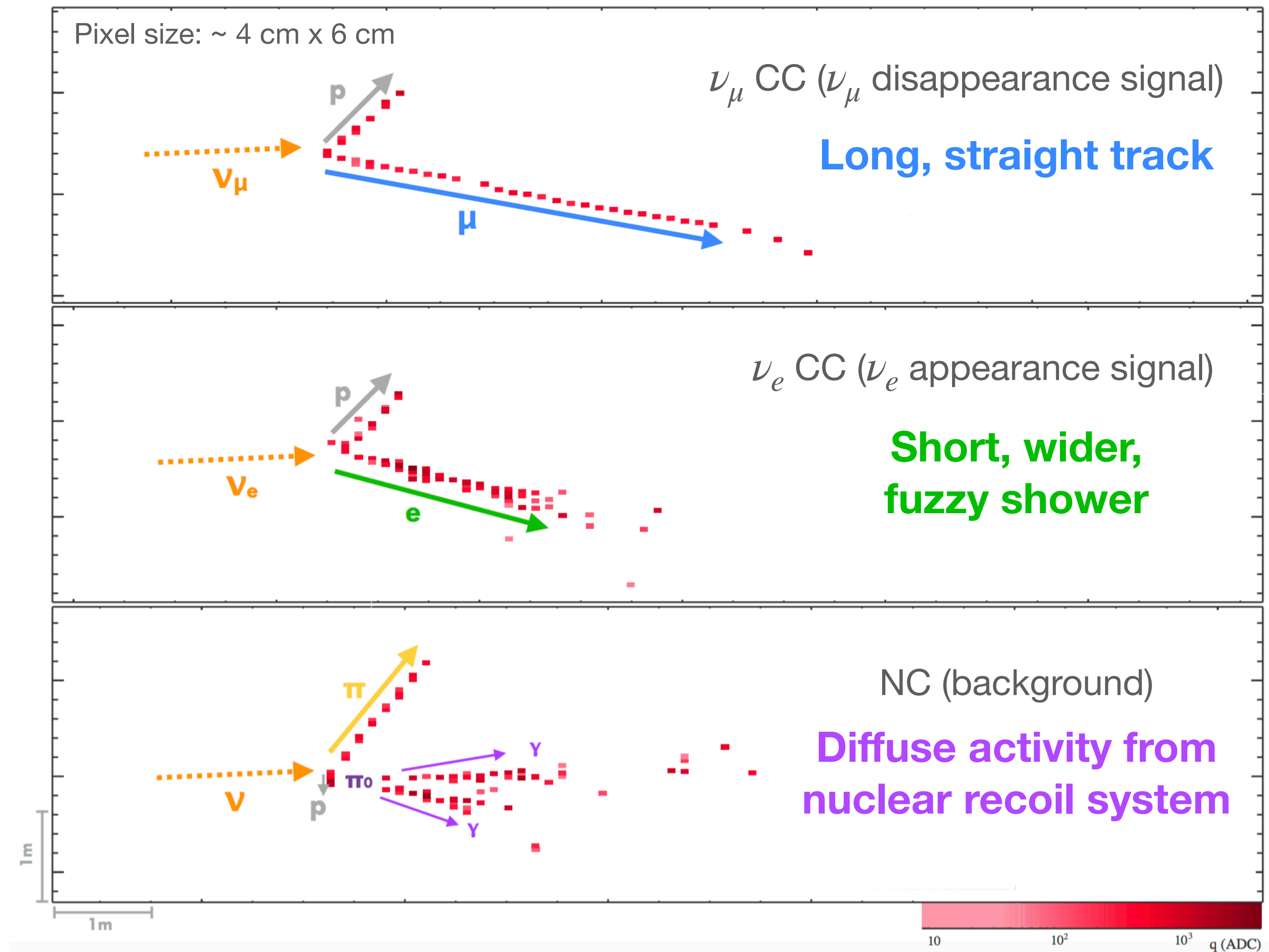
- There are more:
 - In NOvA: Improve ProngCVN with both neutrino and antineutrino sample, Unsupervised training, Panoptic segmentation with point transformer
 - In DUNE: a whole suite of algorithms for the ND reconstruction (semantic segmentation, shower/track separation with GNN, etc)
- ML is covering more aspects in a neutrino experiment: beam monitoring, data taking, simulation, reconstruction, analysis.

Thanks

Backup

NOvA Event Topology

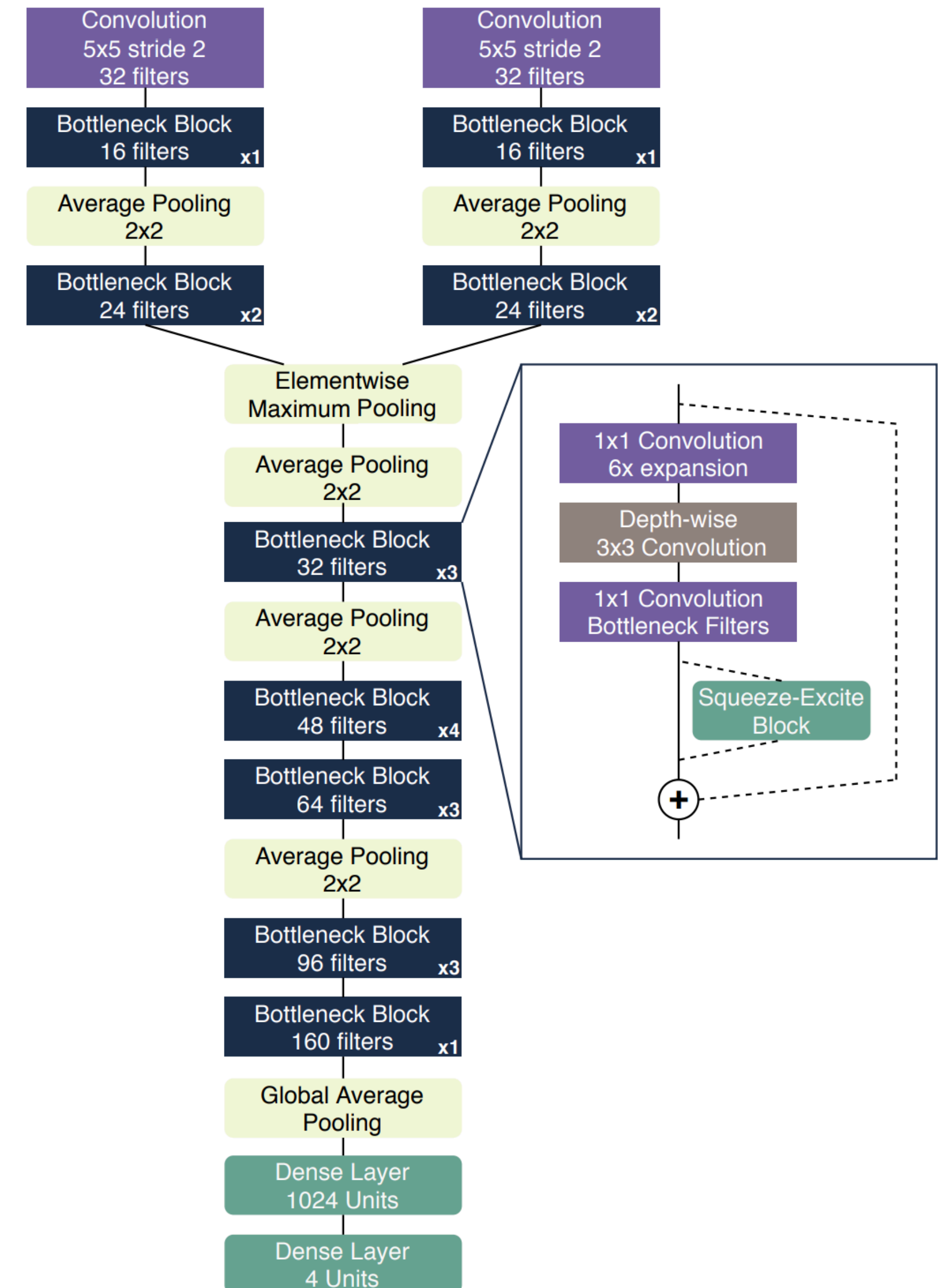
- Charged-current interactions are the “signal” events that allow the neutrino flavor to be identified, via identification of the charged lepton flavor.
- Due to the small ν_e appearance probability (5%), background processes can greatly confuse the measurement.
- For example, NC π^0 events where a π^0 is produced can fake ν_e CC.



J. Inst. 11, P09001 (2016)

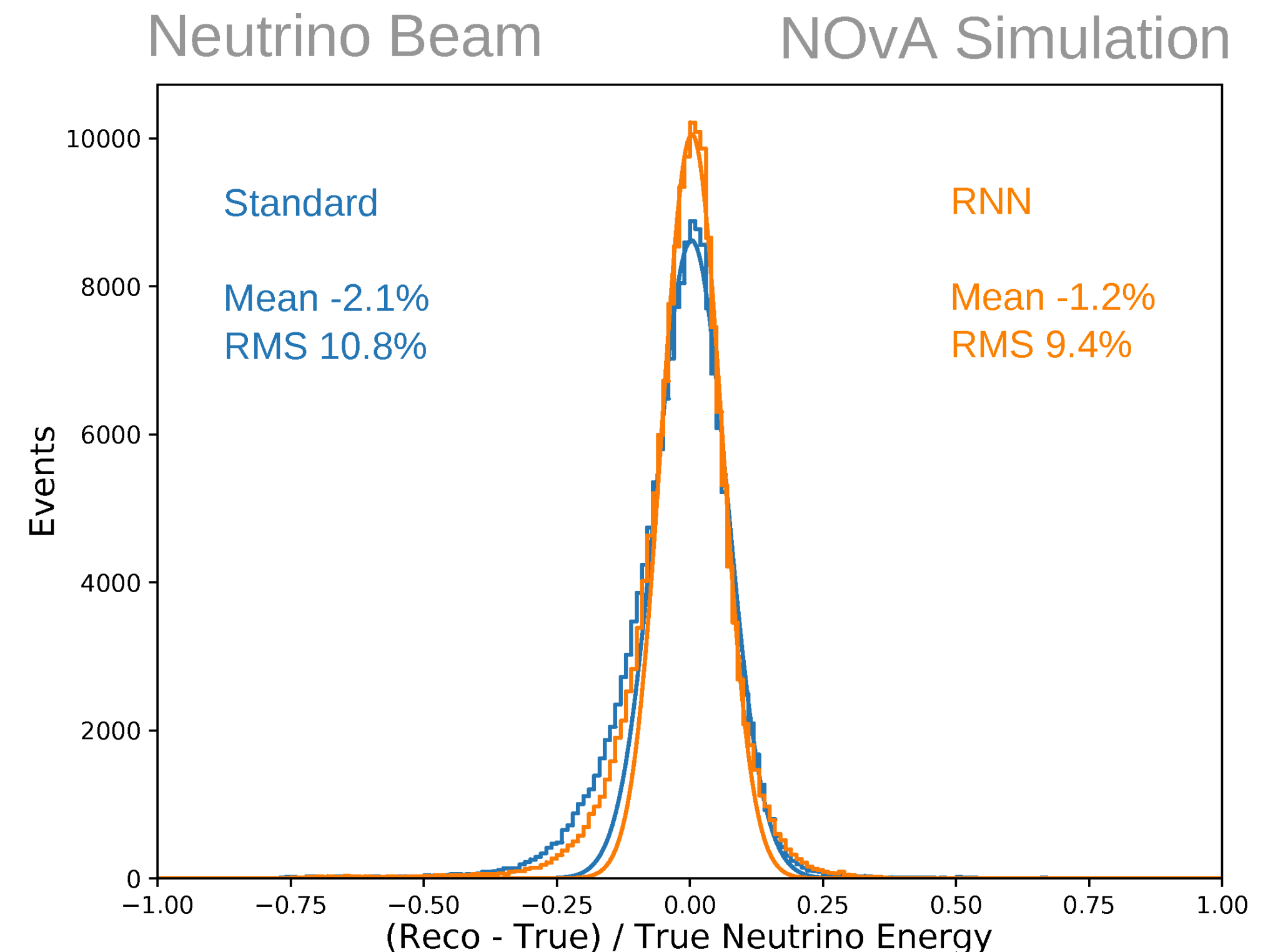
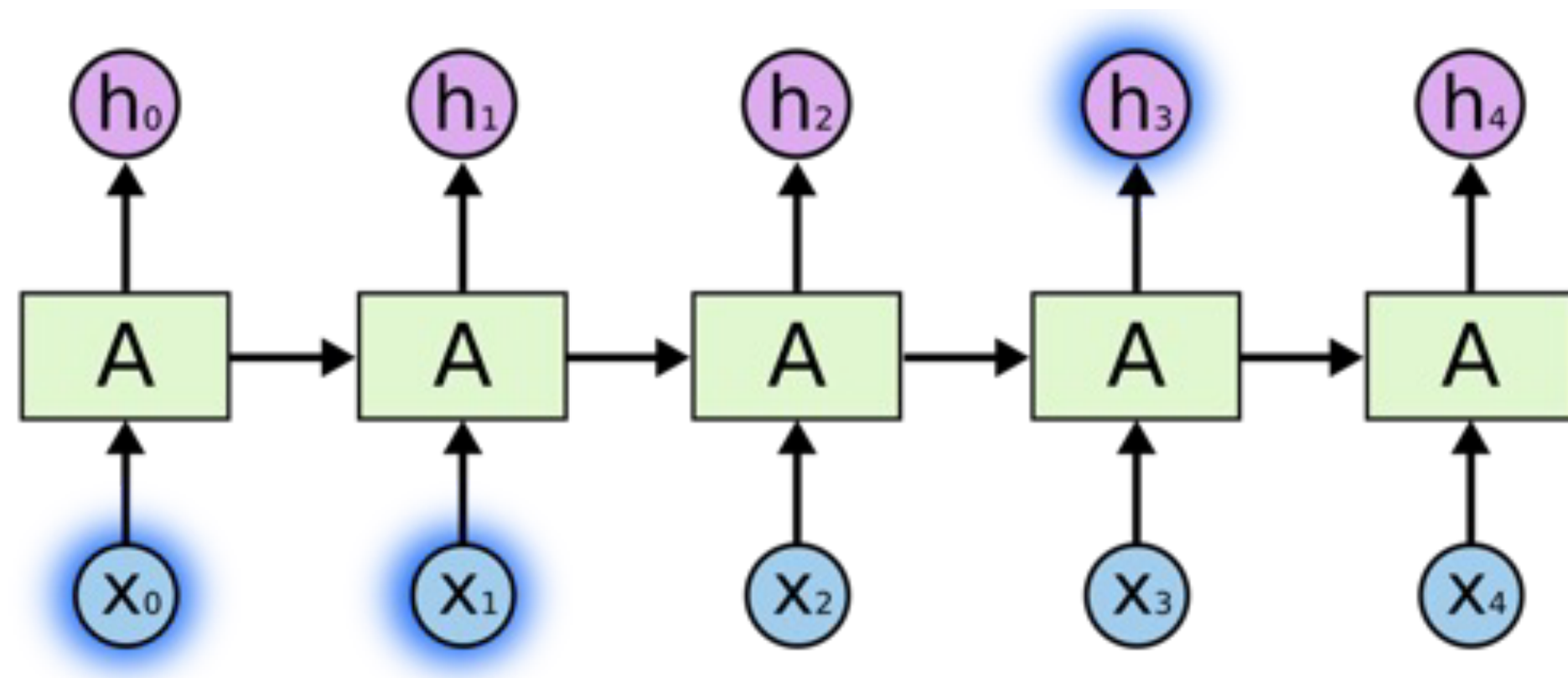
Deep-Learning based neutrino classifier

- NOvA is the **first neutrino experiment using convolutional neural networks (CNN)** for event reconstruction.
- CVN (convolutional visual network) is an event classifier which employs a Deep Convolutional Network in the “image recognition” style.
- Two-tower CNN architecture **learns from the top and side views** independently.
- The statistical power of CVN is equivalent to **30% more exposure** than previous methods.



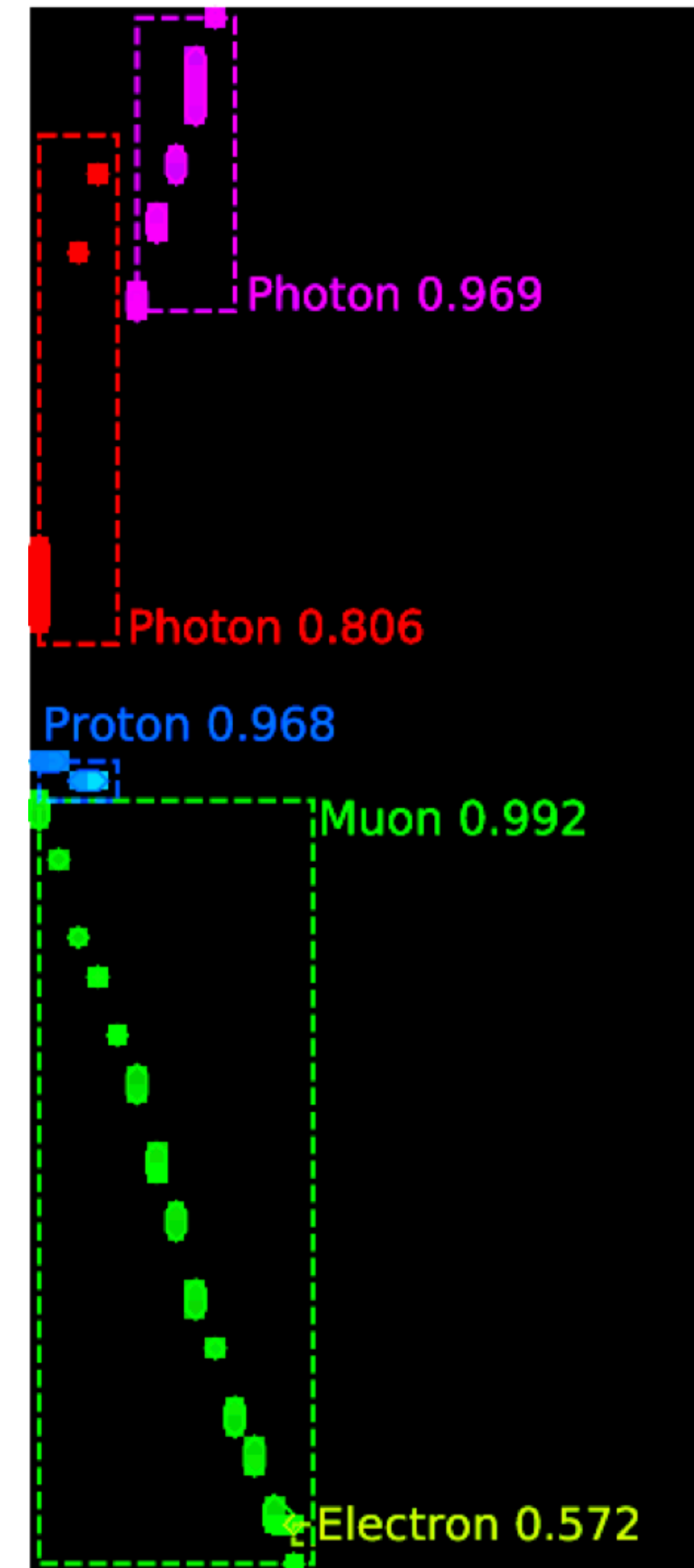
LSTM for Energy Estimation

- 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



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



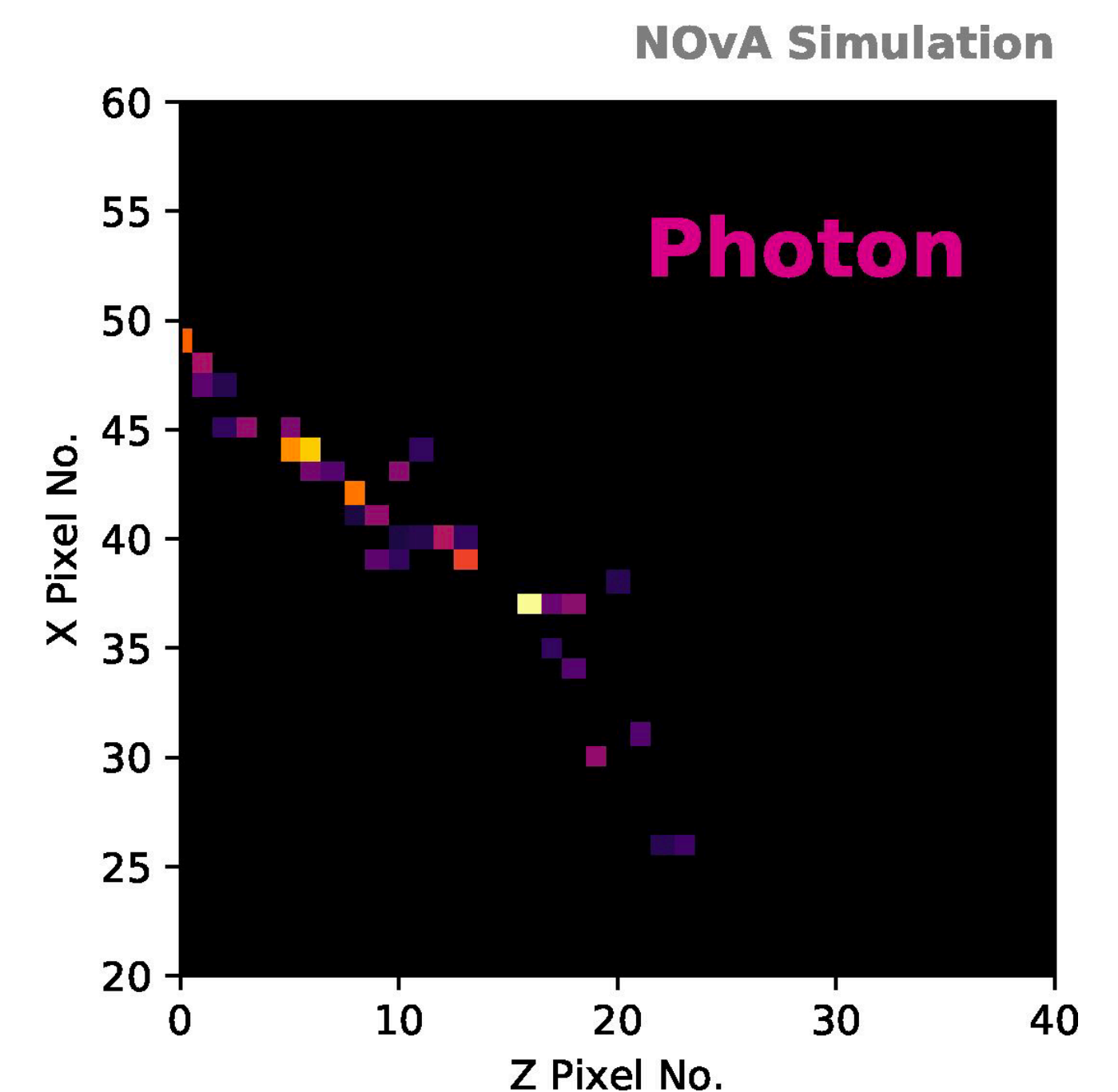
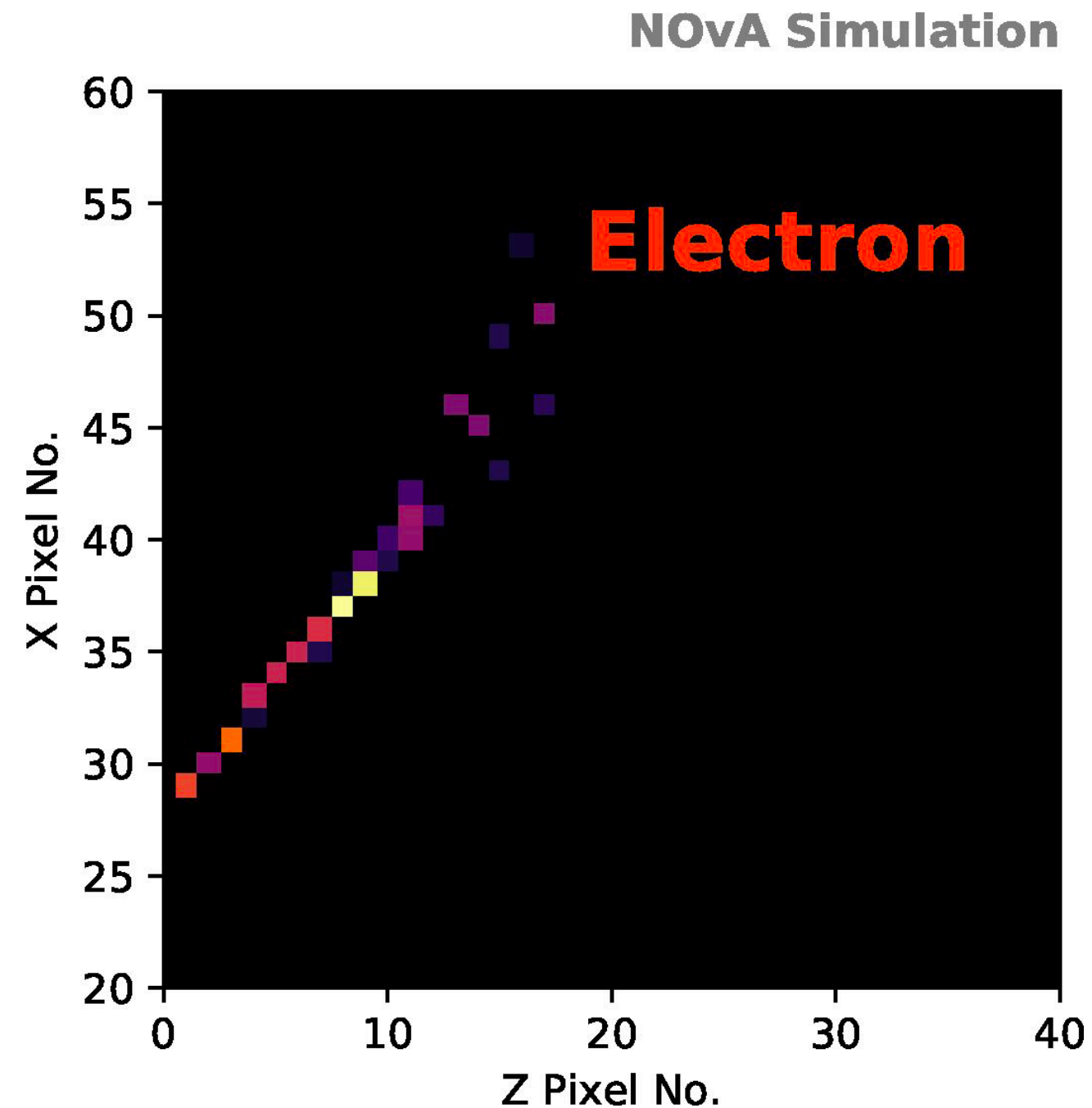
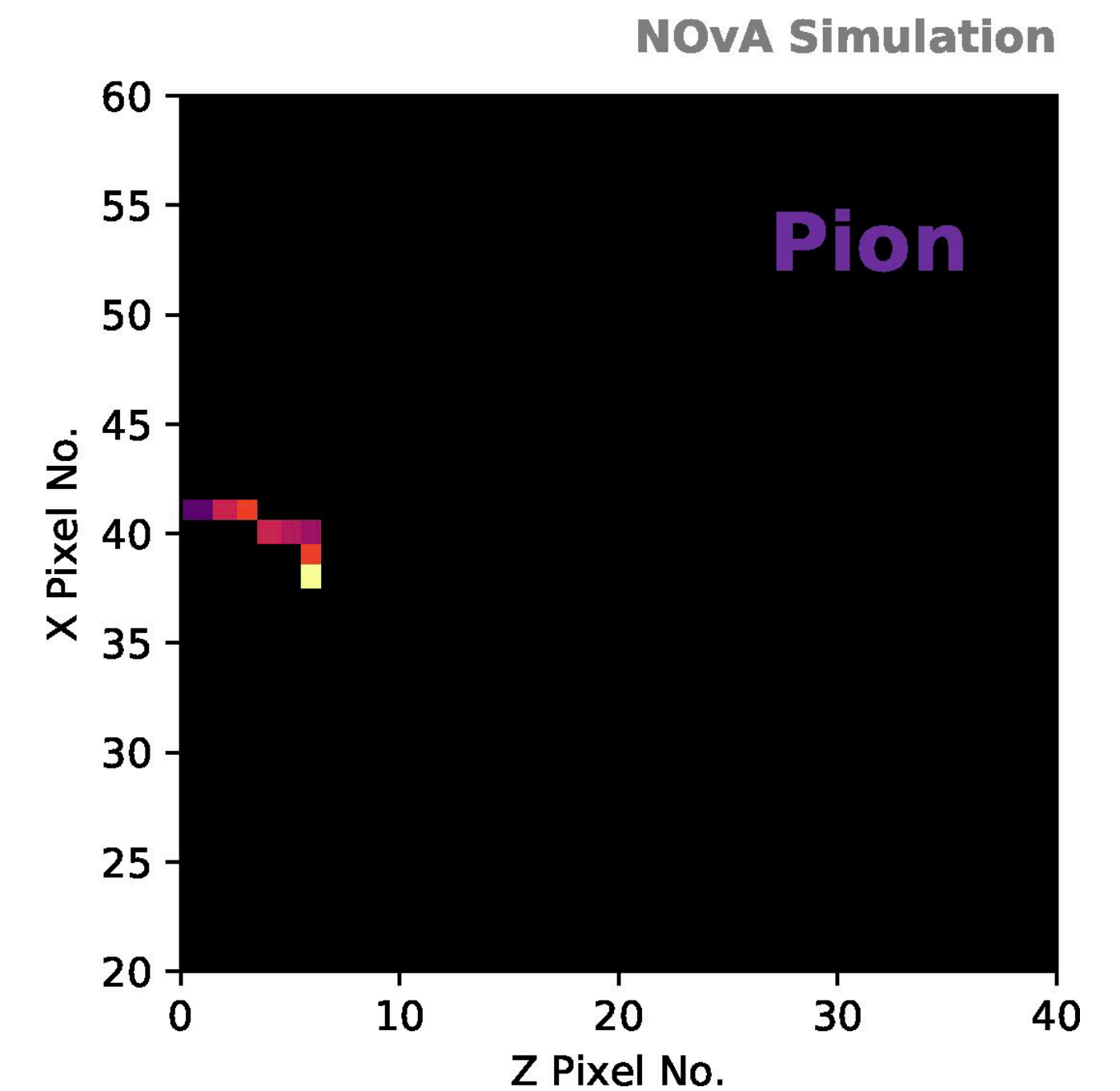
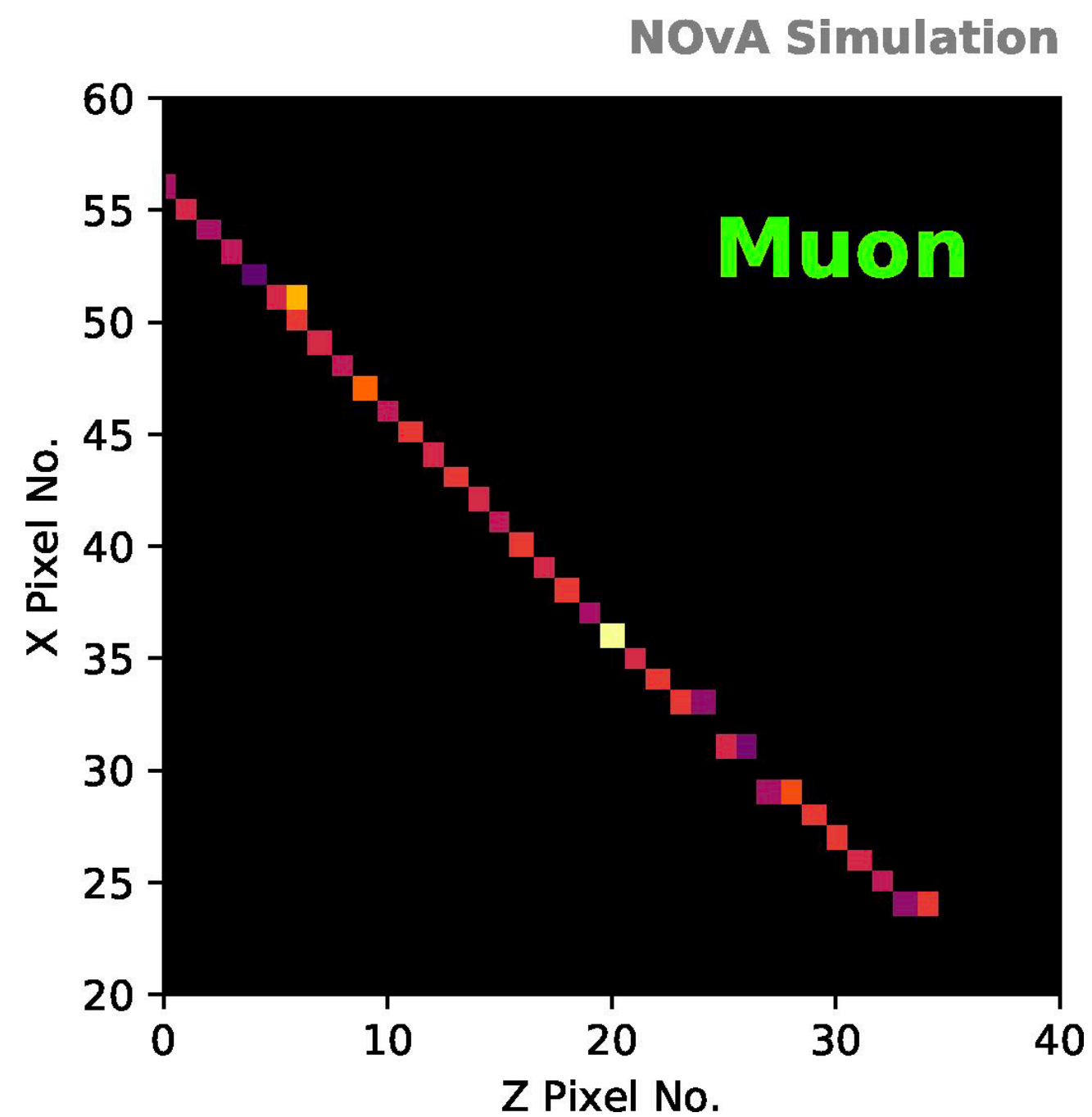
Cosmic filtering with a NN

- 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
$\bar{\nu}_\mu$	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

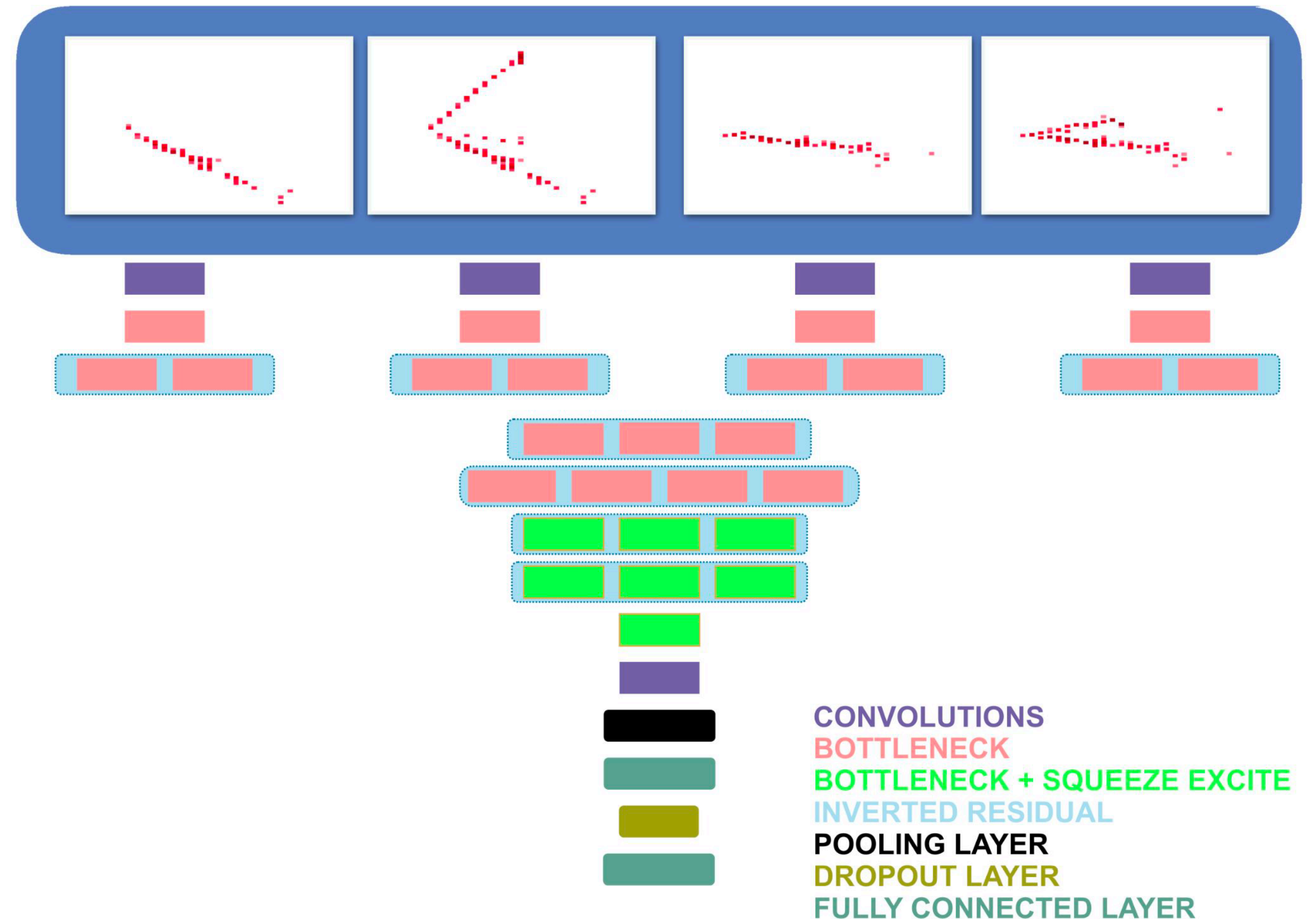
- NOvA also has trained a network using singularly simulated particles for ND analyses → no contextual information
- Also developing a network designed for neutron identification using these samples



Improved ProngCVN

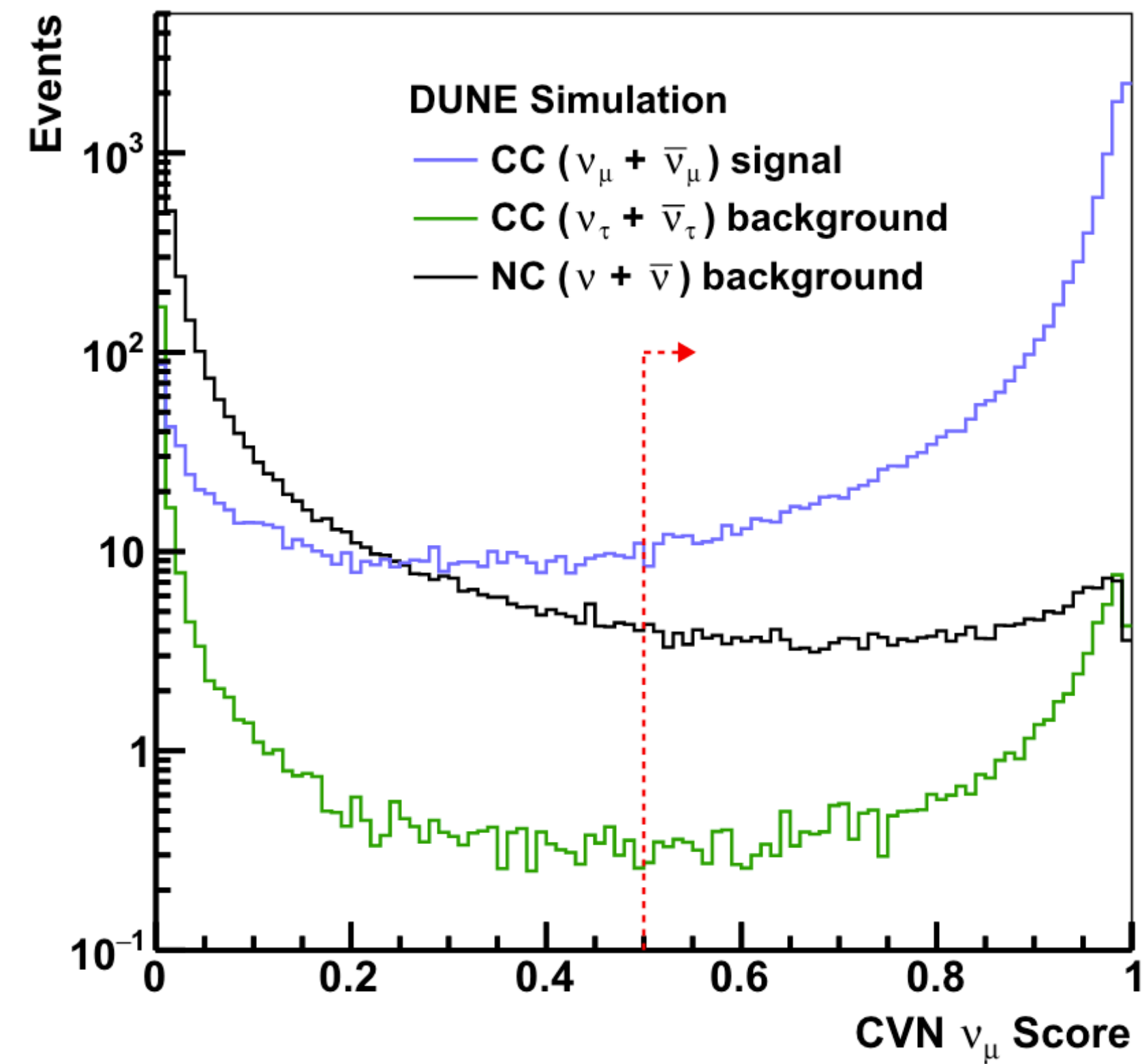
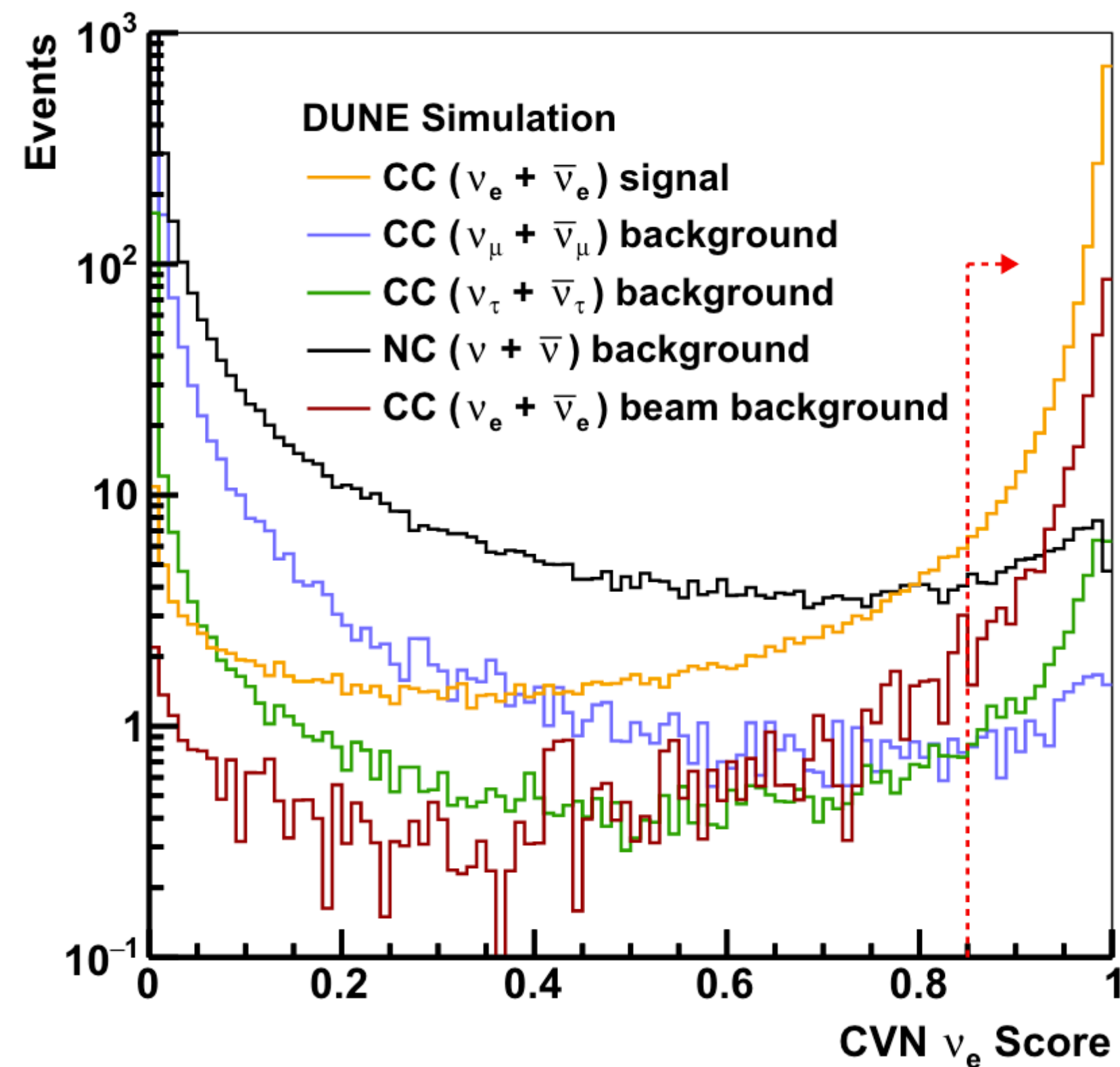
Akshay Chatla, DAE 2022

- 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 selection

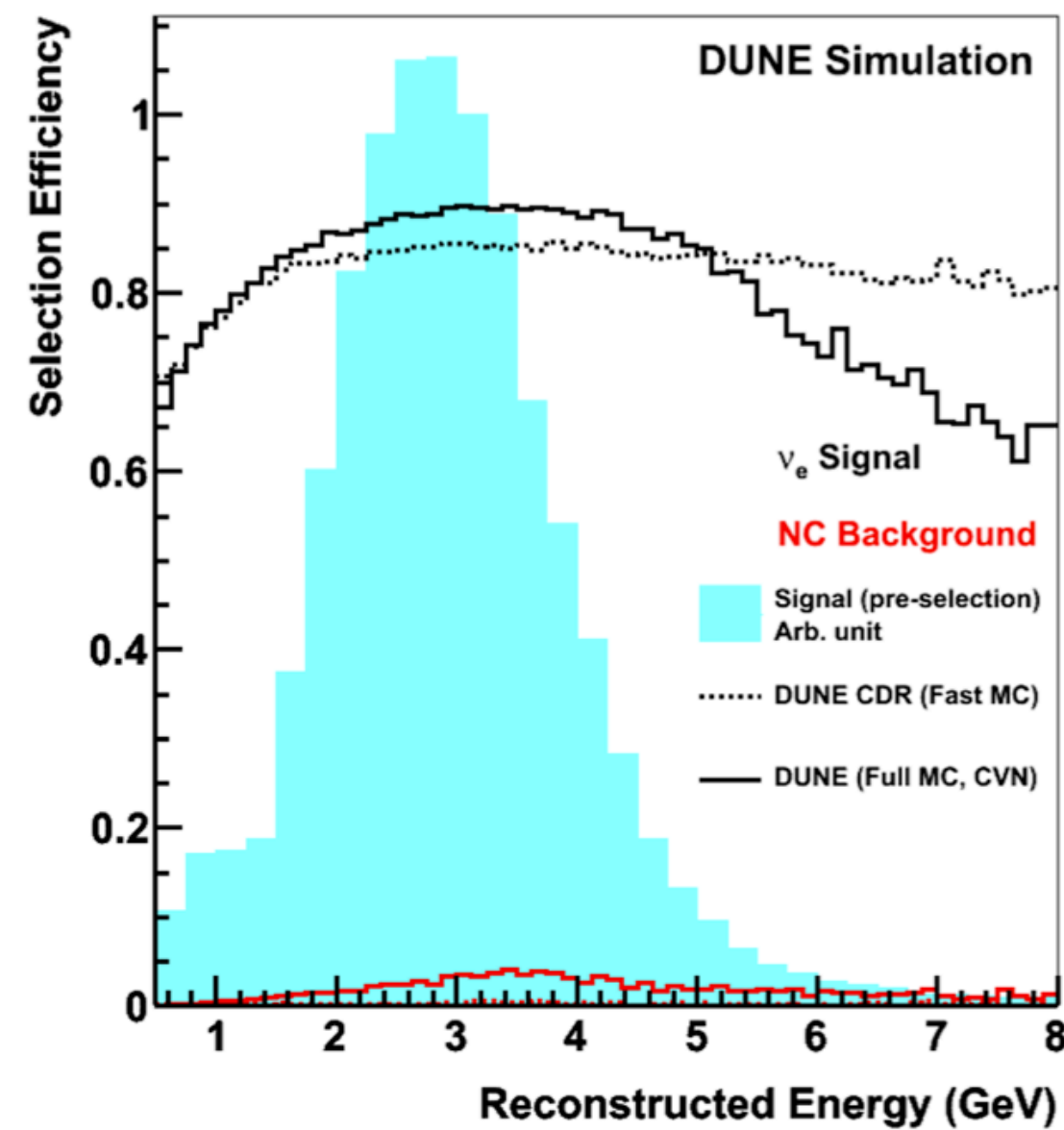
- Event classification is carried out through image recognition techniques using a convolutional neural network, named convolutional visual network (CVN).
 - Build upon the three readout views of the LArTPC, using the reconstructed hits on individual planes.



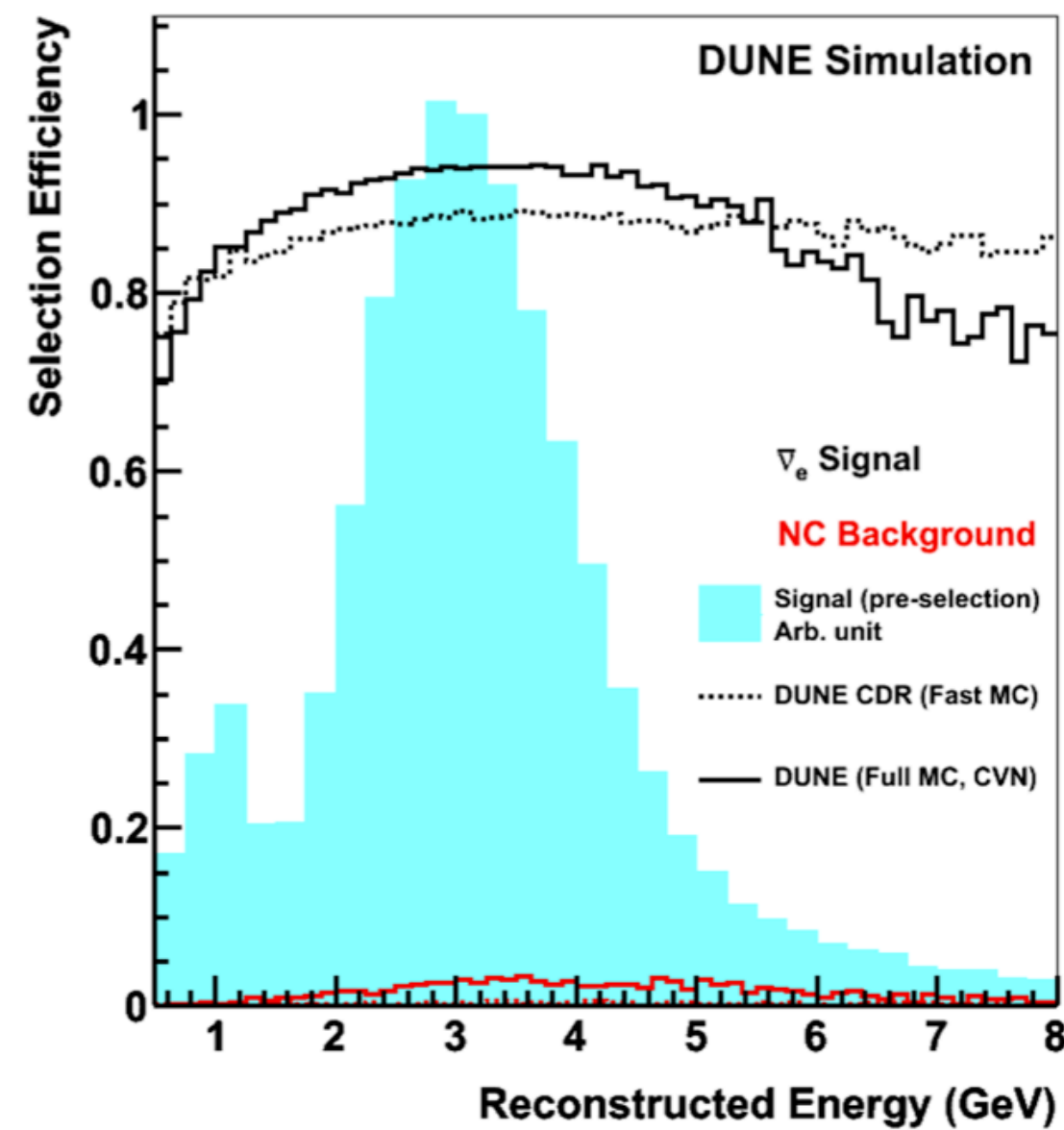
Event selection

- The primary goal of the CVN is to efficiently and accurately produce event selection of ν_e CC and ν_μ CC in FHC, $\bar{\nu}_e$ CC and $\bar{\nu}_\mu$ CC in RHC.
- The ν_e and ν_μ efficiencies in both FHC and RHC beam modes all exceed 90% in the neutrino flux peak.

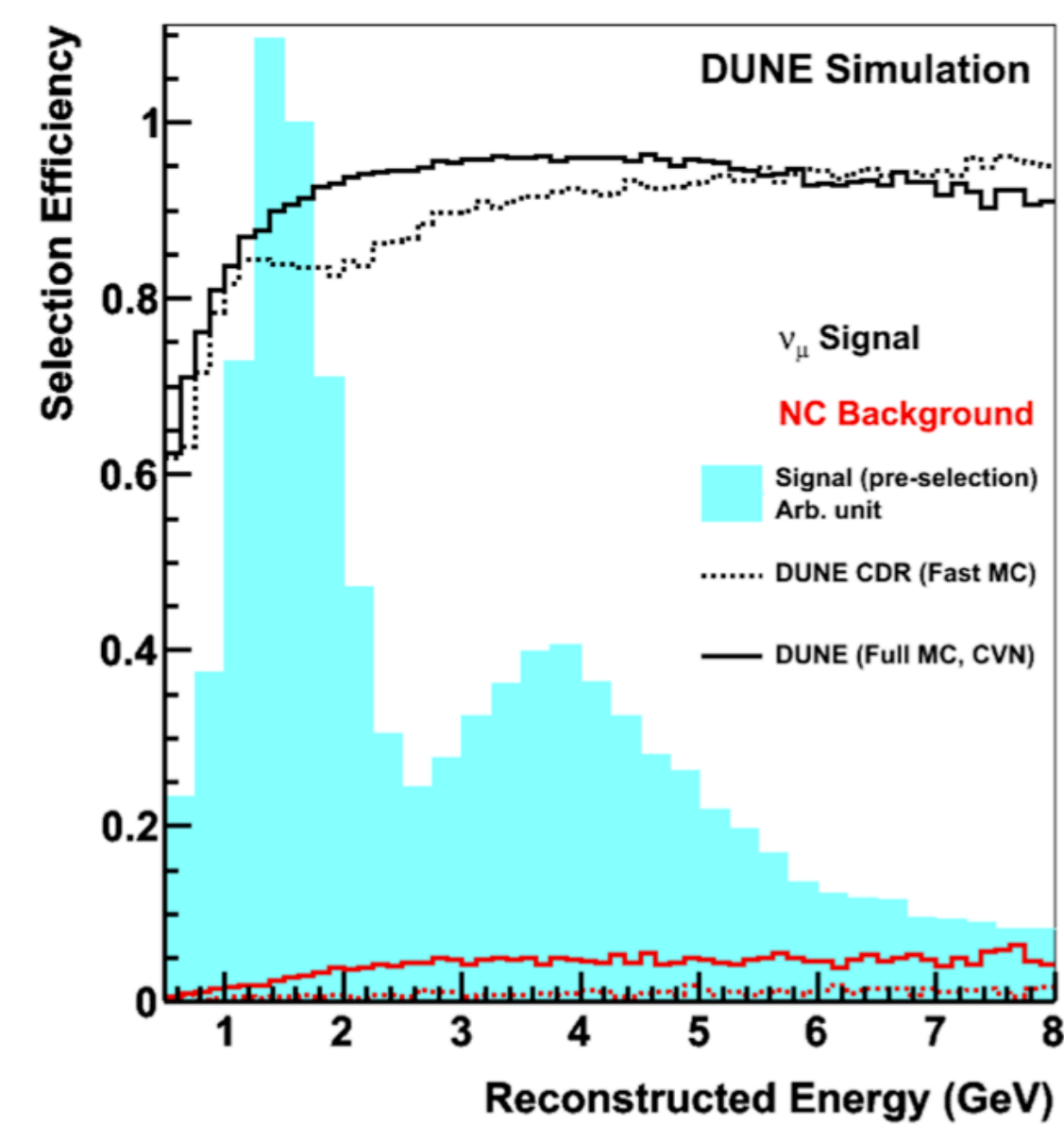
ν_e CC FHC



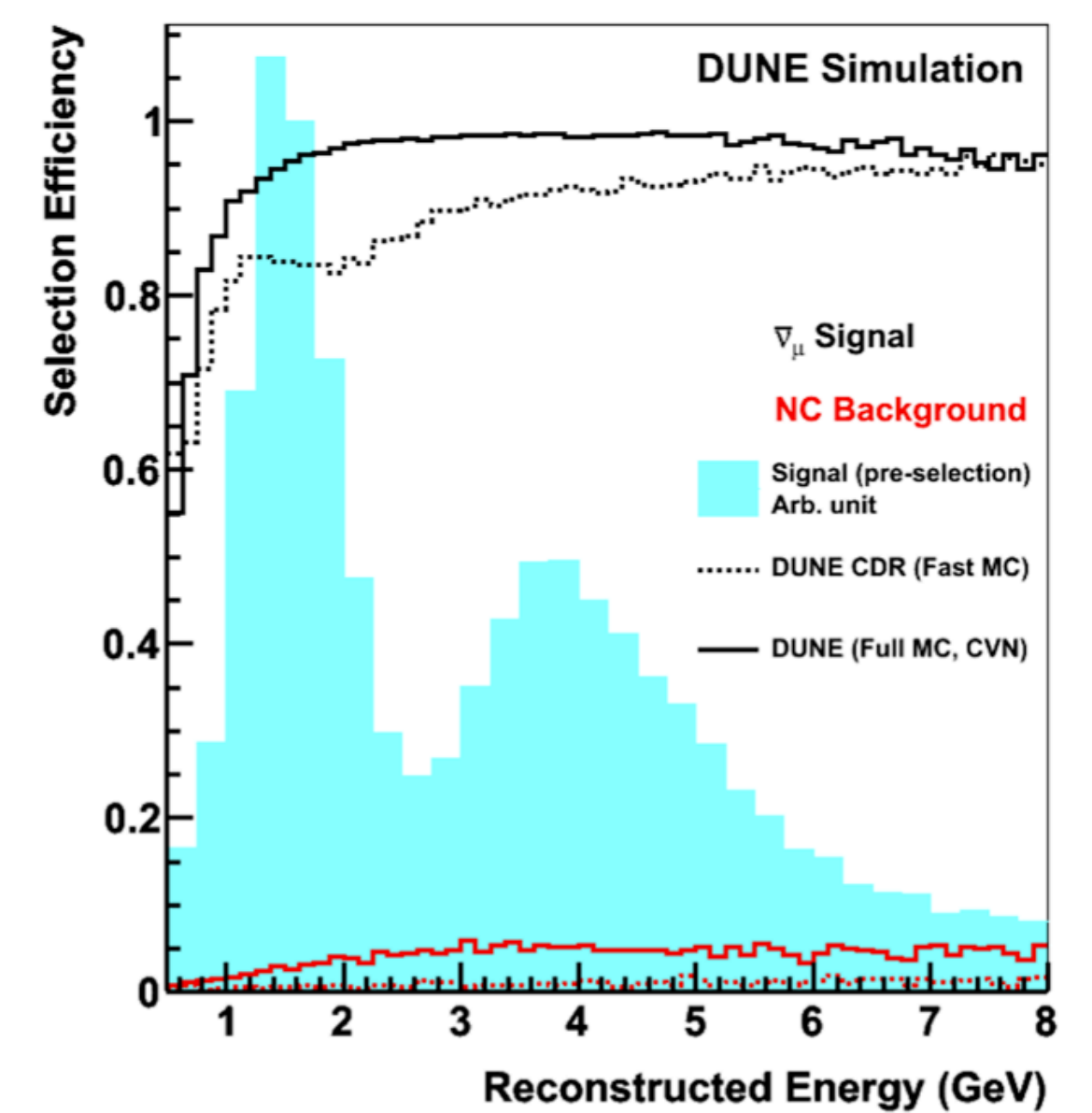
$\bar{\nu}_e$ CC RHC



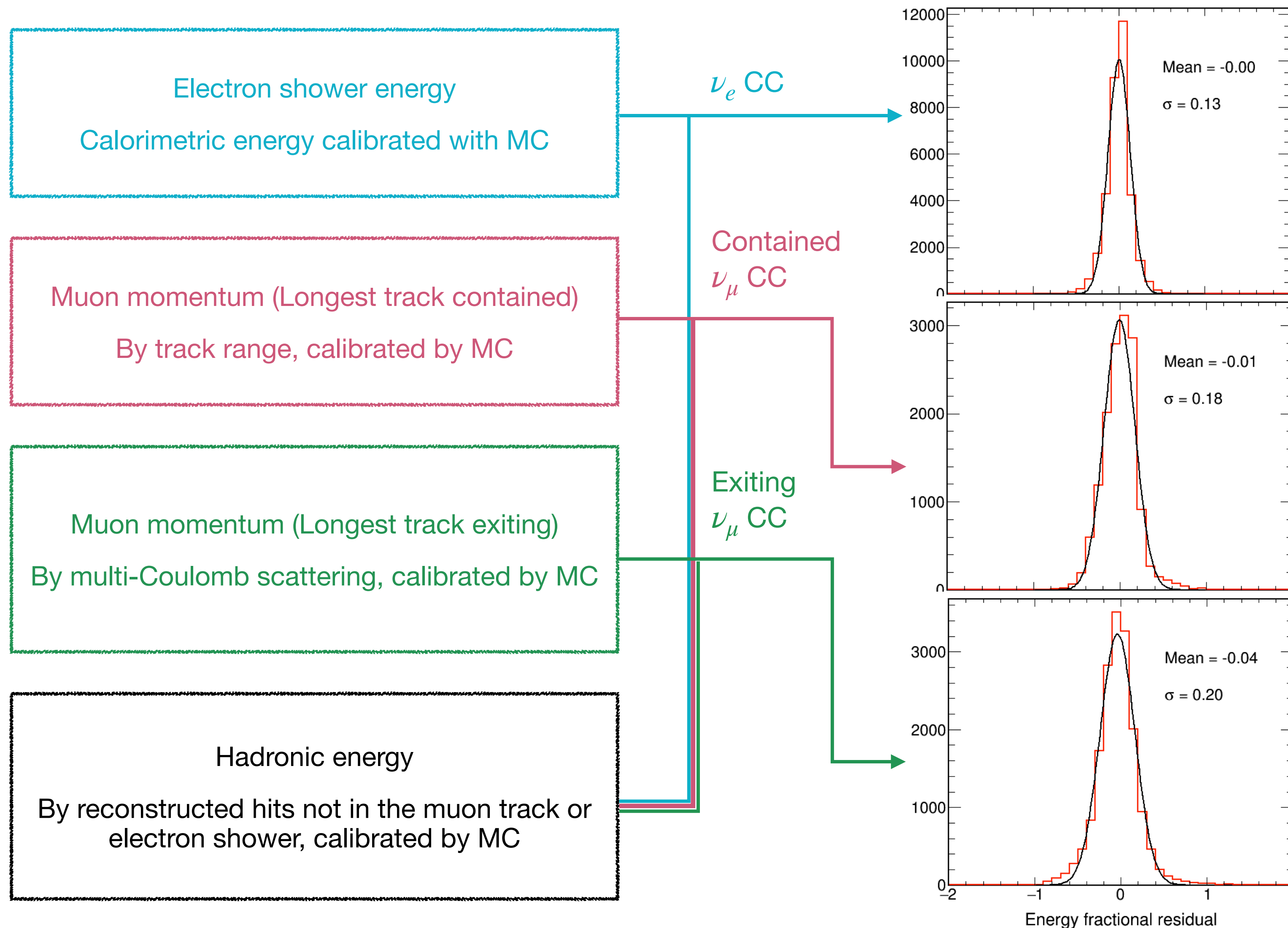
ν_μ CC FHC



$\bar{\nu}_\mu$ CC RHC



Energy resolution



$$E_\nu = E_{\text{lep}}^{\text{cor}} + E_{\text{had}}^{\text{cor}}$$

- ν_e CC energy: sum of the energy of the reconstructed EM shower with the highest energy and the hadronic energy.
- ν_μ CC energy: sum of the energy of the longest reconstructed track and the hadronic energy.
- **15-20% energy resolution in the range of 0.5-4 GeV** that is the most relevant for oscillation measurements.