Neutrino Physics and Machine Learning 2024
HCI J4-June 25th
ETH Zurich, CH

UNIVERSITY OF Cincinnati

Deep Learning in NOvA

Barbara Yaeggy
On behalf of the NOvA Collaboration

U.S. DEPARTMENT OF ENERGY | Office of Science

The NuMI Off-Axis ν_e Appearance Experiment
NOvA

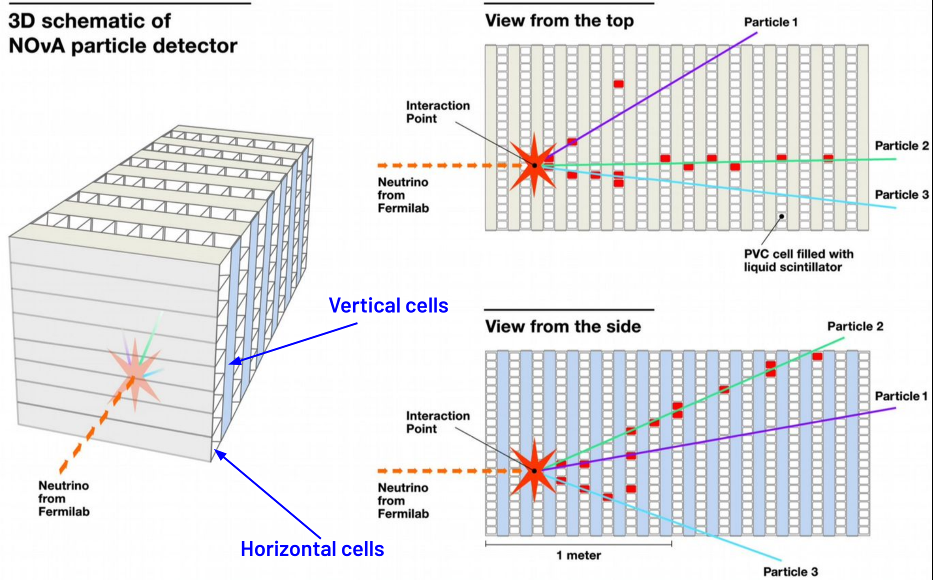
Muon neutrino beam (NuMI) at Fermilab
- Two configurations: neutrino mode (FHC) and antineutrino mode (RHC)

NOvA is an accelerator-based neutrino experiment
- Longest baseline in operation (810 km)
- Large matter effect, sensitive to mass ordering

~14 mrad off-axis, narrow-band beam around oscillation max

NOvA Detectors

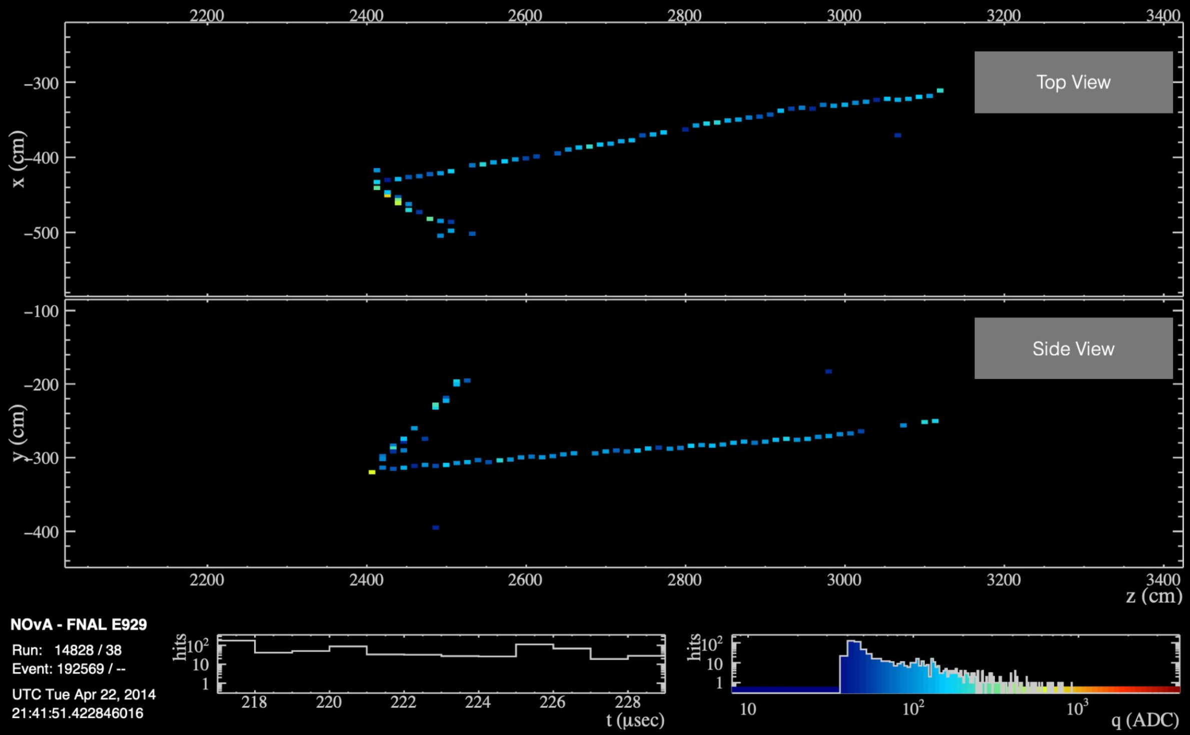
- ND (290-ton) and FD (14-kton) are functionally identical to minimize systematics
  - Time resolution ~ ns, spatial resolution ~ cm, ND beam spill 10μm
- Composed of highly reflective extruded PVC cells filled with liquid scintillator. Scintillation light captured and routed to Avalanche Photodiode (APD) via wavelength shifting fiber (WLS)
- Cells arranged in planes, assembled in alternating horizontal and vertical directions → provide 3D views of the events



Barbara Yaeggy - University of Cincinnati

Event Reconstruction in NOvA

- NOvA detectors are naturally segmented
- Producing a pair of pixel maps: cell number v.s. plane number



Slicer

↓

Vertex Finding

↓

Particle Clustering

↓

Particle ID/Energy

↓

Event ID/Energy

ANALYSIS

NOvA uses a variety of algorithms to reconstruct physics information for which **slicing is the core input**

Barbara Yaeggy - University of Cincinnati

NOvA use DL techniques for reconstruction and analysis

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

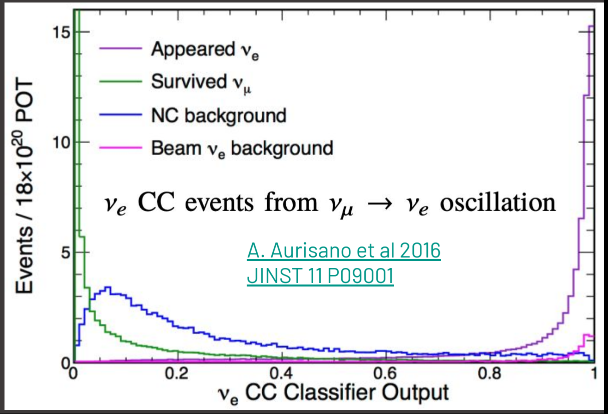
Event Classifier - *Oscillation related*

Event classification (neutrino interactions or cosmic rays) and for particle classification starting from individual clusters

- MobileNetv2 arXiv:1801.04381 - Two-tower Siamese structure

Increase sensitivity to neutrino oscillation parameters over traditional methods (Phys. Rev. Lett. 116, 151806) equivalent to collecting 30% more exposure

These criteria also reject 97.6% of the NC and 99.0% of the ν_μ - CC beam backgrounds

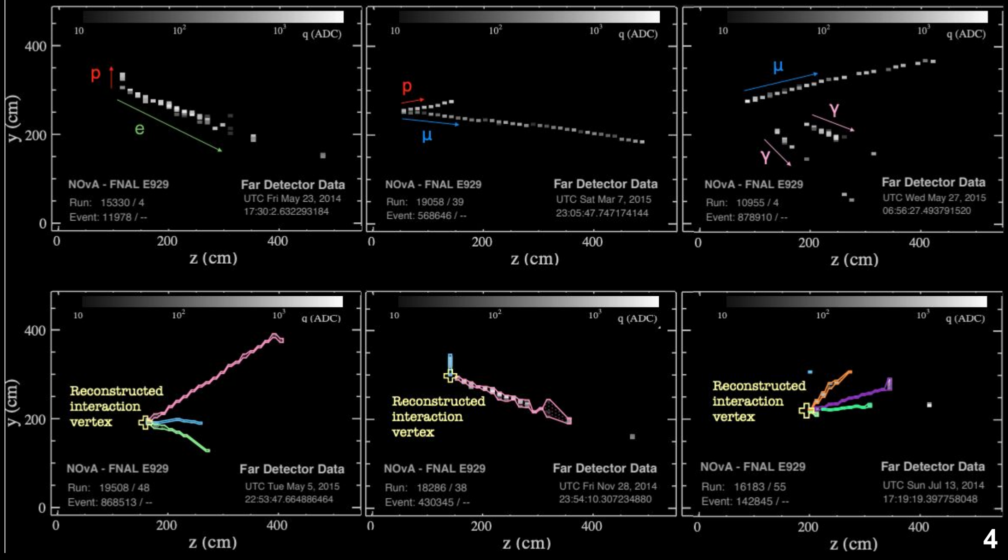


Particle Classifier - *Cross section measurements*

Identify prongs as hadronic or electromagnetic (EM)

It takes the categories assigned by EventCVN and it labels by particle type, corresponding to the largest CVN score for the prong.

- MobileNetv2 - Four-tower Siamese structure



EventCVN Phys. Rev. Lett. 116, 151806

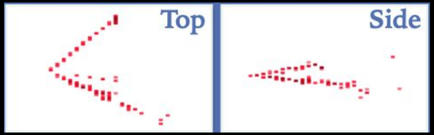
MobileNetv2 arXiv:1801.04381

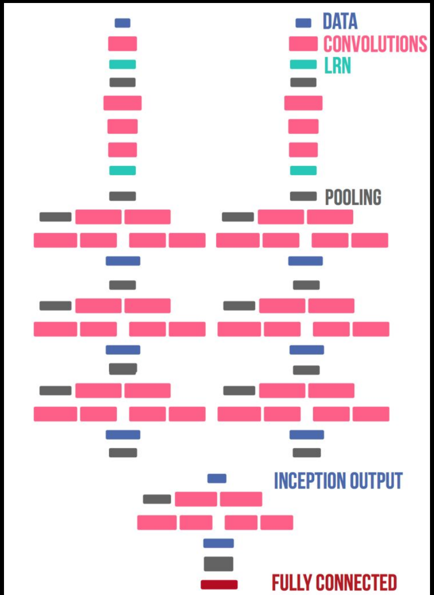
Two-tower Siamese structure

Similar performance for both modes

Anti-nu shows slightly increase in efficiency

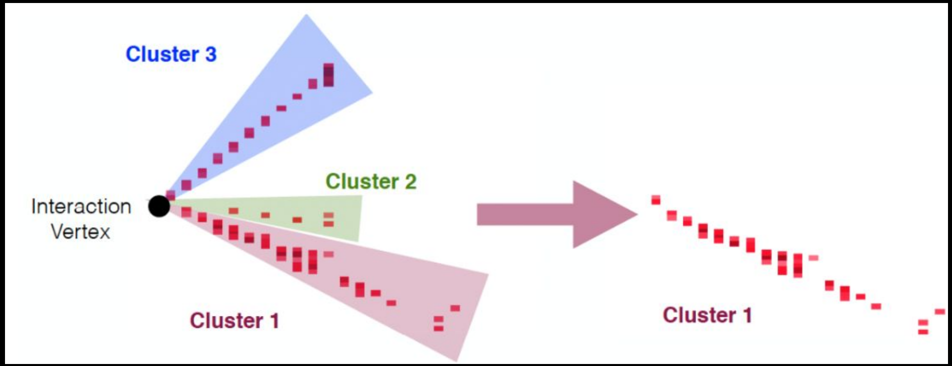
Purity over 90% for all interactions flavors

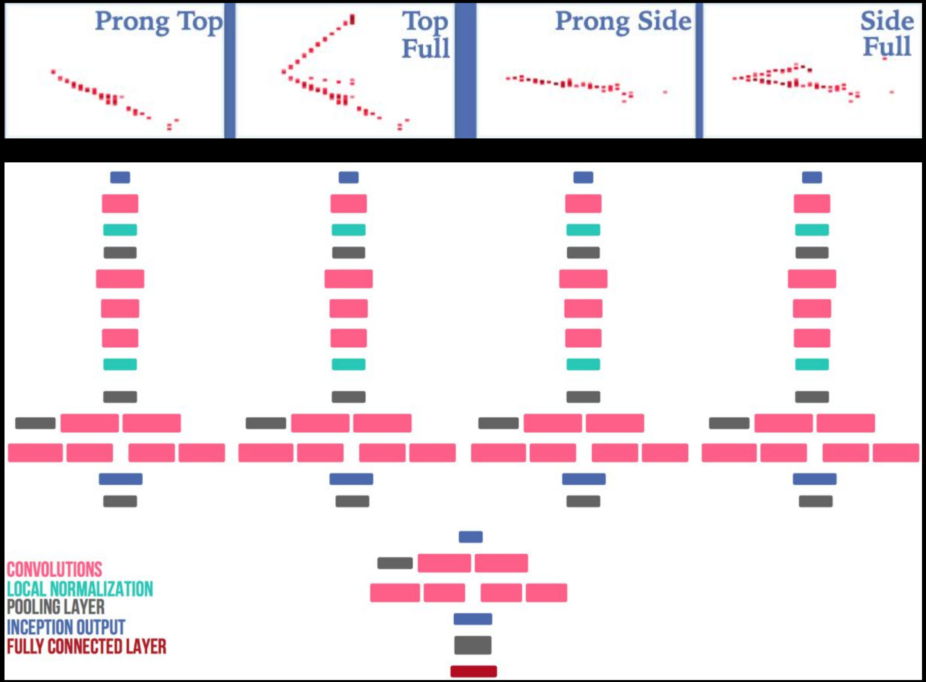




ProngCVN Phys. Rev. Lett. 116, 151806

MobileNetv2 arXiv:1801.04381
Four-tower Siamese structure





- **Single particles are separated using geometric reconstruction methods**
- **Hits with a common direction are grouped together to make clusters.**
  - Individual cluster identification
- **Classifies particles using both views of the particle and both views of the entire event**
- **Currently, trained on separate samples of neutrino and anti-neutrino mode**

TransformerCVN

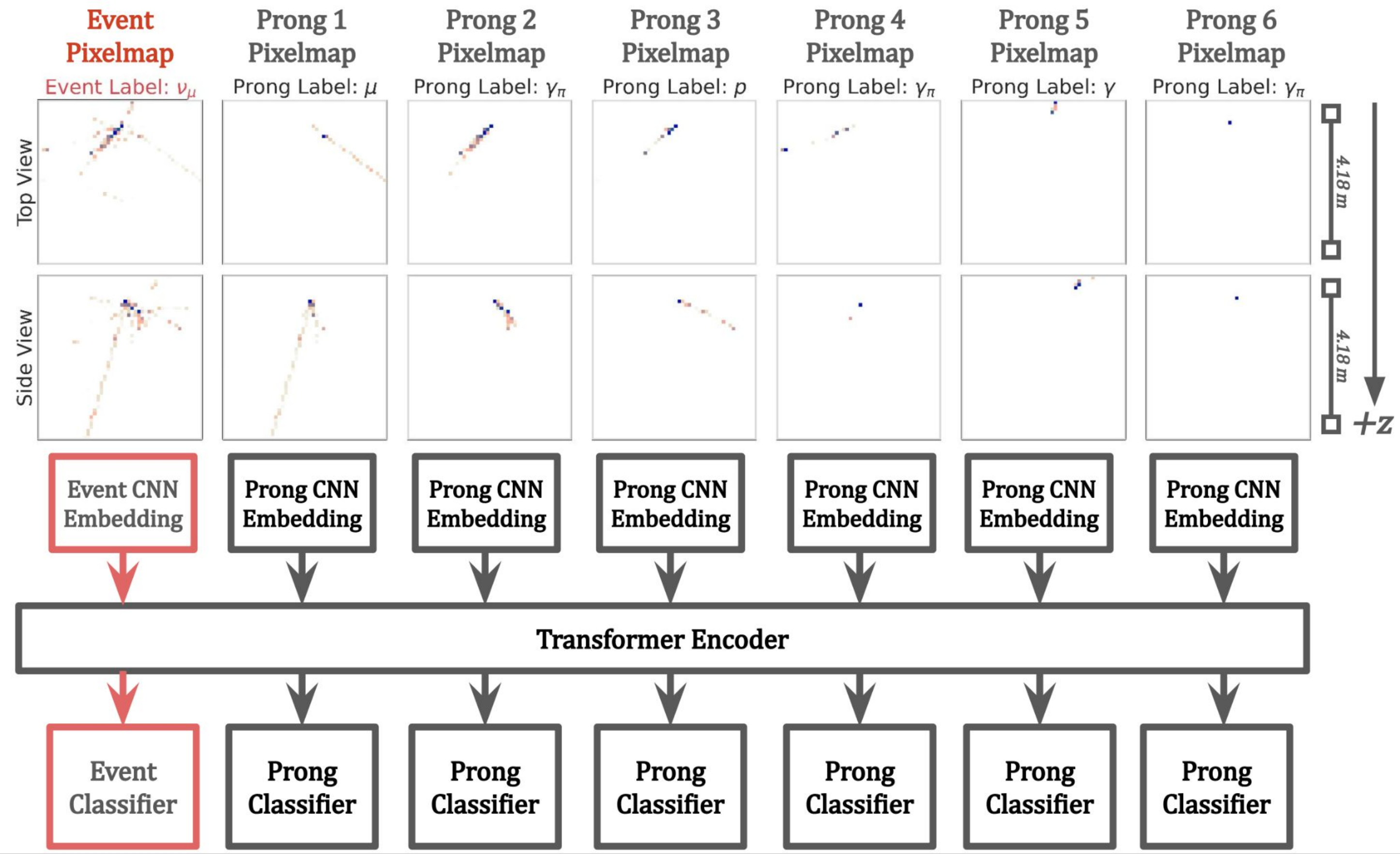
Event & Particle Classification

Coming talk with details about Transformer CVN at NOvA, by Jianming Bian

Spatial learning + attention mechanism: simultaneously classifies each event and reconstructs every individual particle's identity

Classifies events with 90% accuracy & improves the reconstruction of individual particles by 6% over baseline methods

Great improvement in particle identification!
Thanks to the the additional context provided by all prongs and the transformer's attention mechanism



arXiv:2303.06201

Regression CNNs for Energy Estimation

Phys. Rev. D 99, 012011

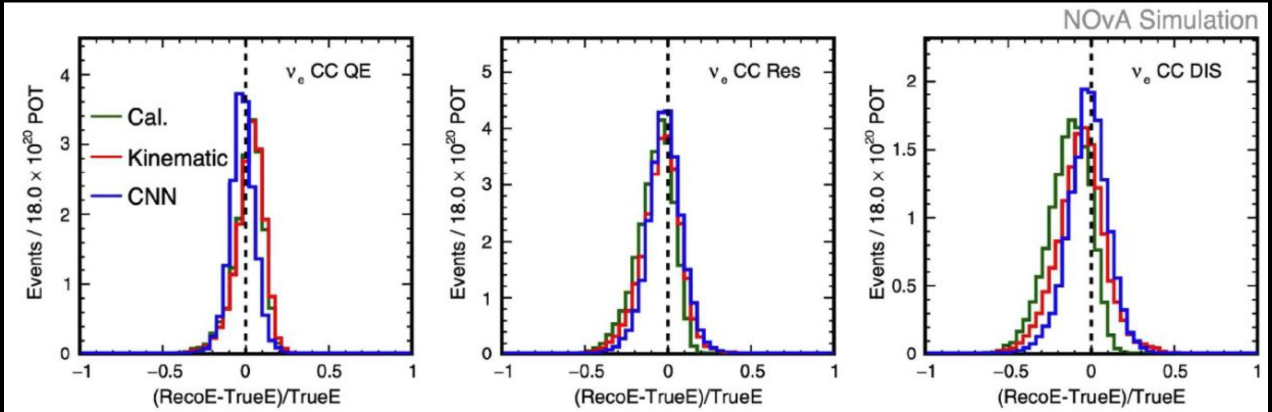
The CNN architecture used is an adapted ResNet

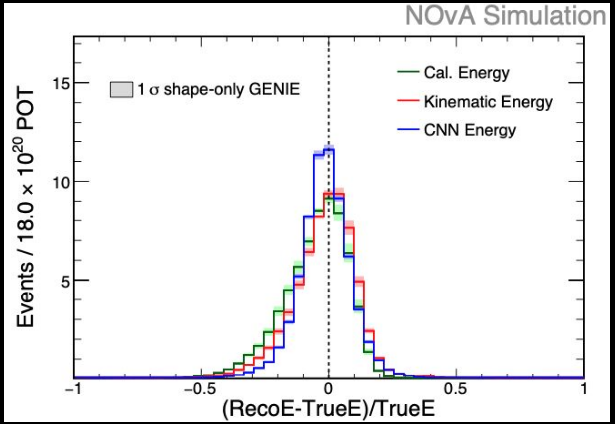
- Weighting scheme, so the loss function sees a flat energy distribution, to control energy dependence
- Use mean absolute percentage error instead of square of errors to decrease the effects of outliers

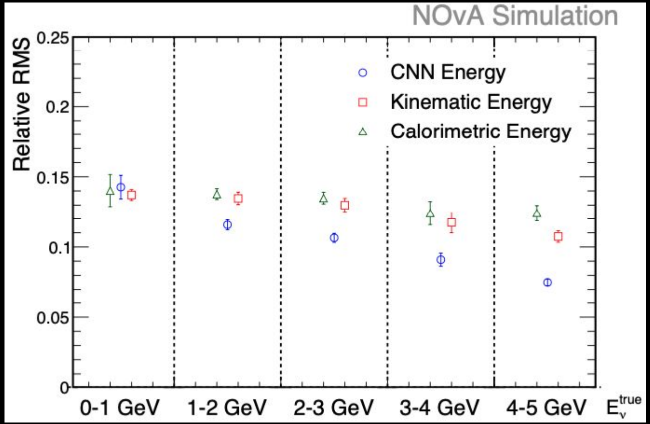
Shows a better resolution compared with kinematics-based energy reconstruction

Provides improvements of 16% and 12% in RMS for ν_e CC and electron, respectively

Good stability over interaction types!

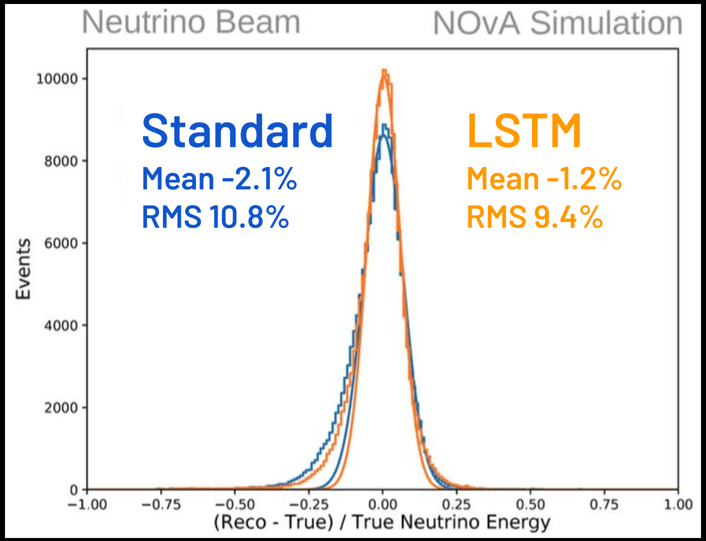






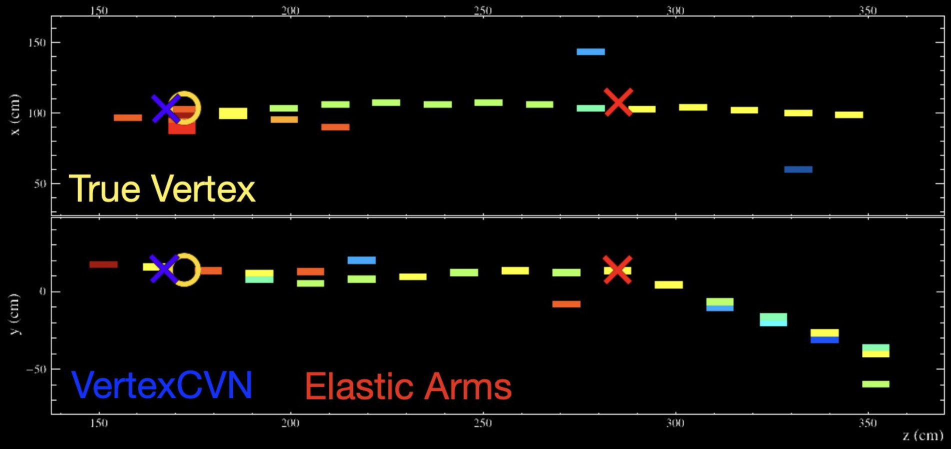
LSTM for Energy Estimation

- Long Short-Term Memory (LSTM) is a type of recurrent neural network
- Takes a number of traditional reconstruction quantities as inputs (prong & slice level) ⟶ energy prediction
- Trained with artificially engineered sample to increase network resilience
- Resolution comparable with regression CNN, (used for quick cross-check)



ML Vertexer, *VertexCVN*

- More accurate vertex finding, means more accurate on:
  - Clustering hits to form individual particle tracks/showers
  - Identifying particle types
  - Energy estimation
- Shows good performance across interaction types!
- 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



NOvA Simulation

We want to do a better reconstruction!

Semantic segmentation : pixel level ID
Instant segmentation: semantic segmentation + object detection

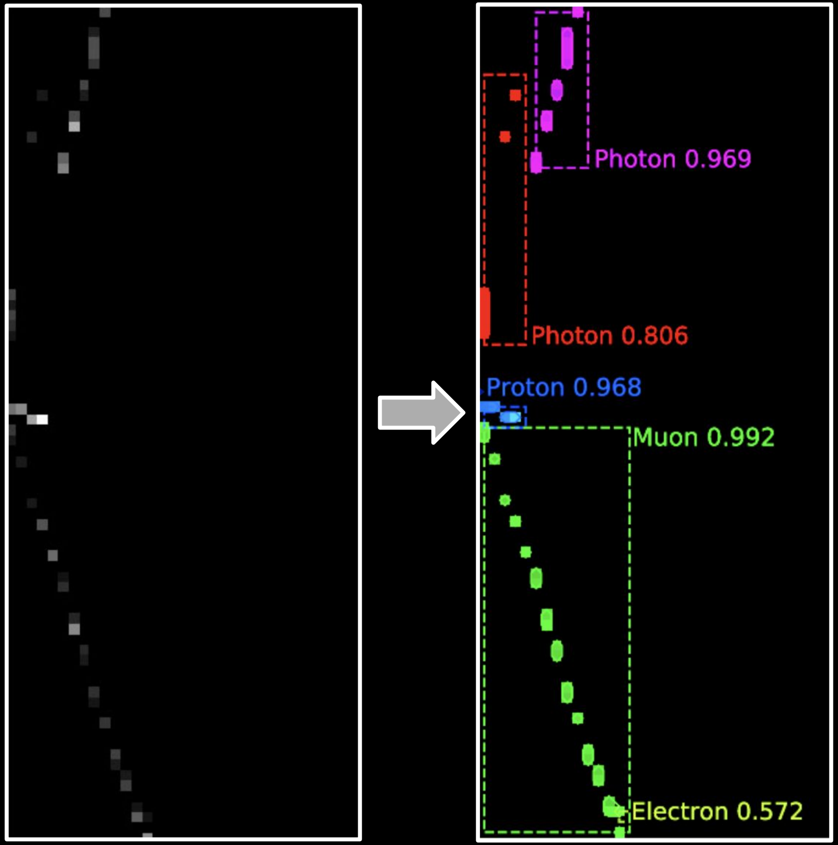
- **Full event reconstruction on a hit-by-hit basis using instance segmentation:**

- 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



NOvA Simulation

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
- Singularly simulated particle for ND analyses (no contextual information)
- Cosmic filtering, based on ResNet18 with a siamese structure
- Reconstruct neutrino energy, final state particle energy, vertex
- Perform full event reconstruction

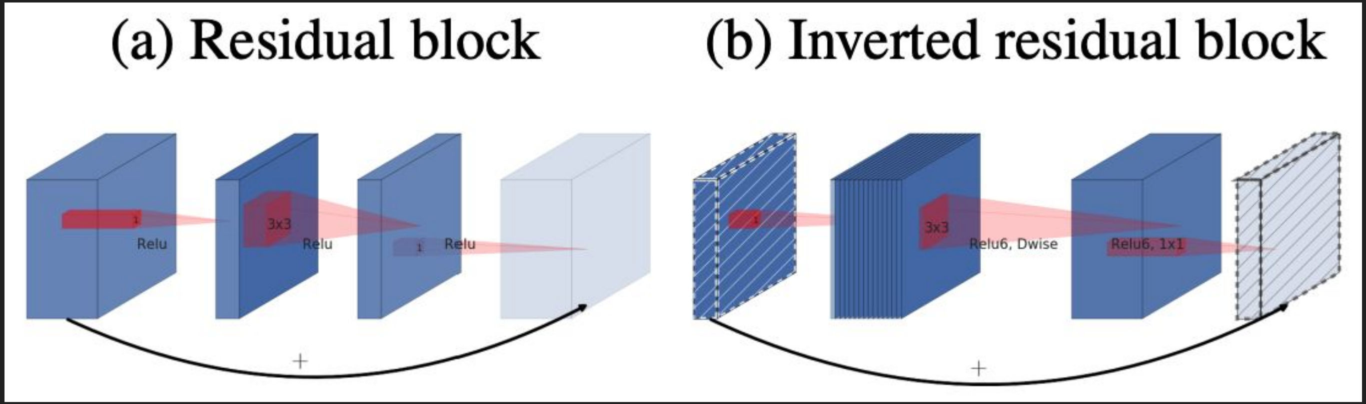
- **Other ongoing ML efforts in NOvA:** improve ProngCVN with both neutrino and antineutrino sample, graph neural networks, unsupervised training.

- NOvA has been performing expansive data comparison, impact analysis, uncertainty studies and cross-checks to improve robustness and interpretability of ML techniques

THANK YOU!

Backup

MobileNetV2



- CNN

- Thickness of each block to indicate its relative number of channels

- The MobileNet v2 architecture is based on an inverted residual structure where the input and output of the residual block are thin bottleneck layers opposite to traditional residual models which use expanded representations in the input. MobileNet v2 uses lightweight depthwise convolutions to filter features in the intermediate expansion layer.

Siamese Neural Networks

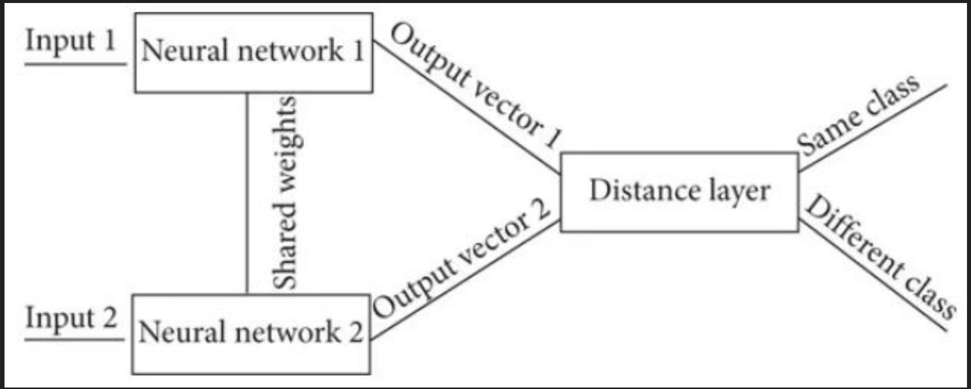
- Innovative technique known as similarity problems which finds if two inputs are similar or not which is known as a siamese neural network

- We require only one training example for each class

- This kind of neural network architecture is scalable and does not require much retraining

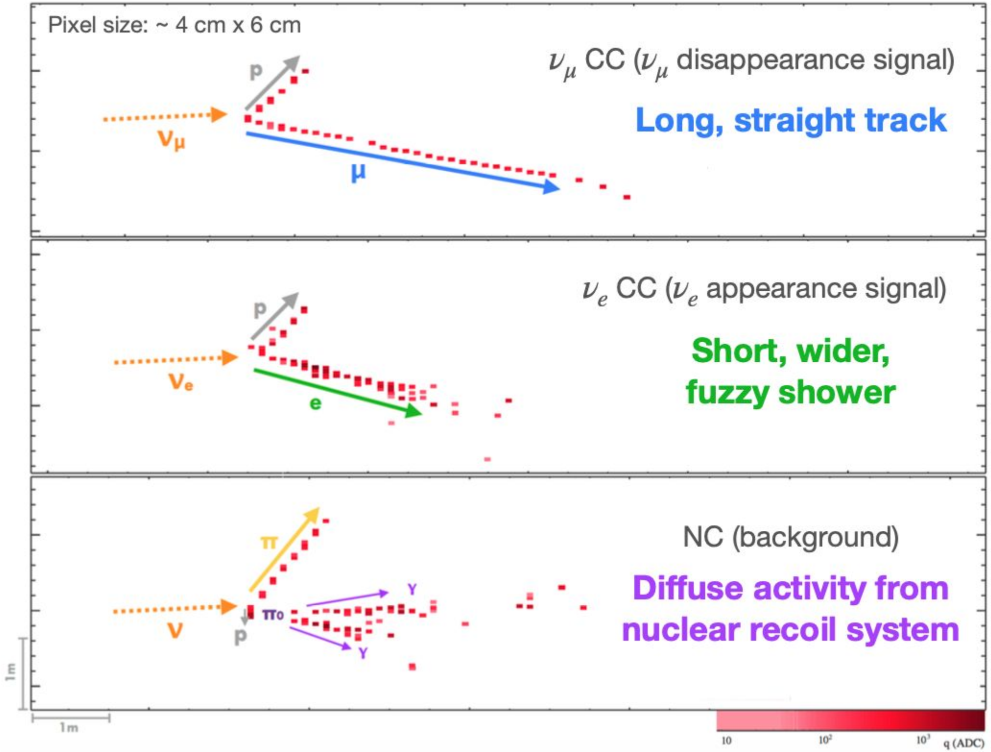
- Siamese networks are not learning to classify an image to any of the output classes.

- But, it is learning with help of a similarity function, which takes two images as input and gives the probability of how similar these images are.



Barbara Yaeggy - University of Cincinnati

Event Topology



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

Cosmic filtering with a NN

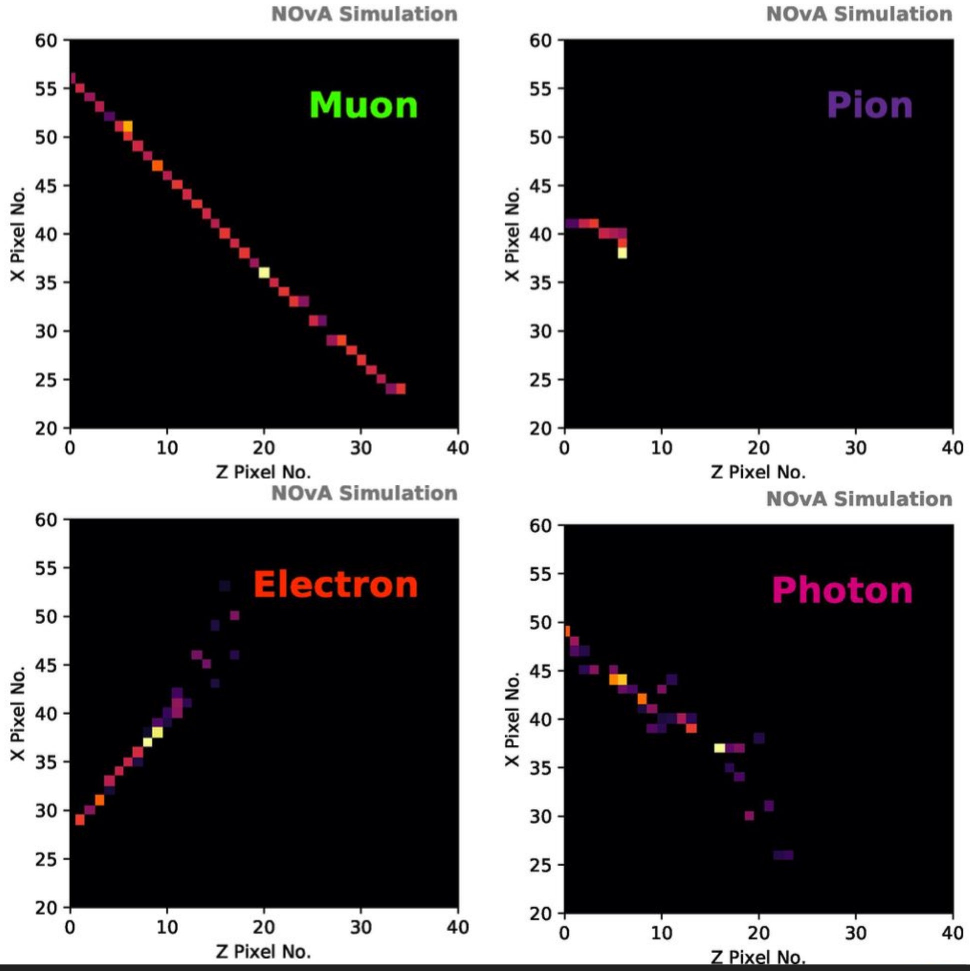
- Network based on ResNet18 backbone with a siamese structure
  - Takes in two event images (top-view and side-view) as input
- Softmax output with five labels: ν_μ, ν_e, ν_τ, NC, and cosmic score
- Training sample contained 1M+ ν_μ, ν_e, and NC events in both beam modes and 5M+ cosmic events
  - Not 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

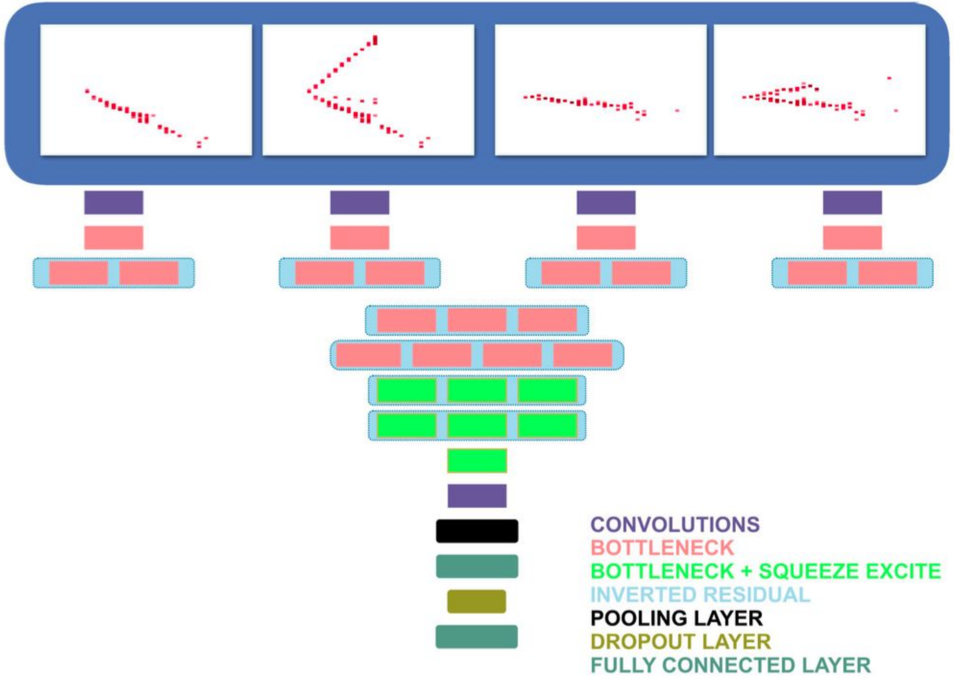
- Also developing a network designed for neutron identification using these samples



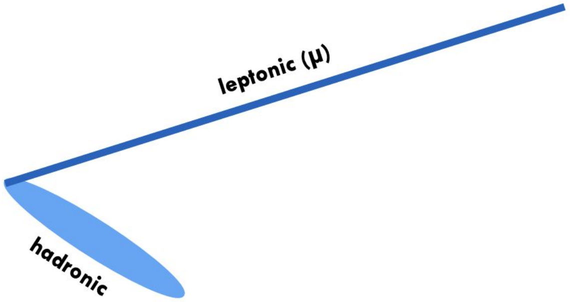
Improved ProngCVN

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



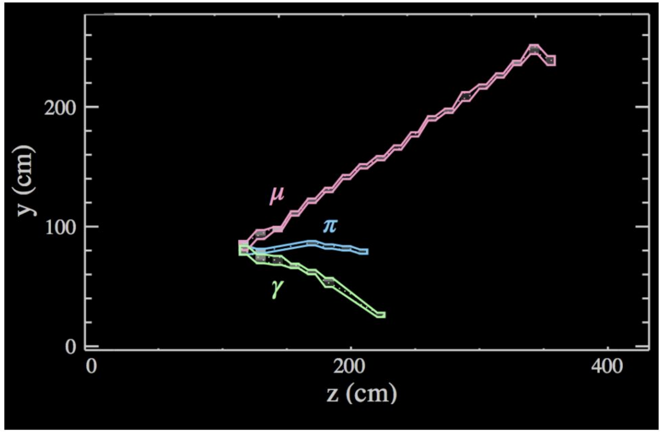
Reconstruction Uses



$$E_\nu = E_{lep} + E_{had}$$

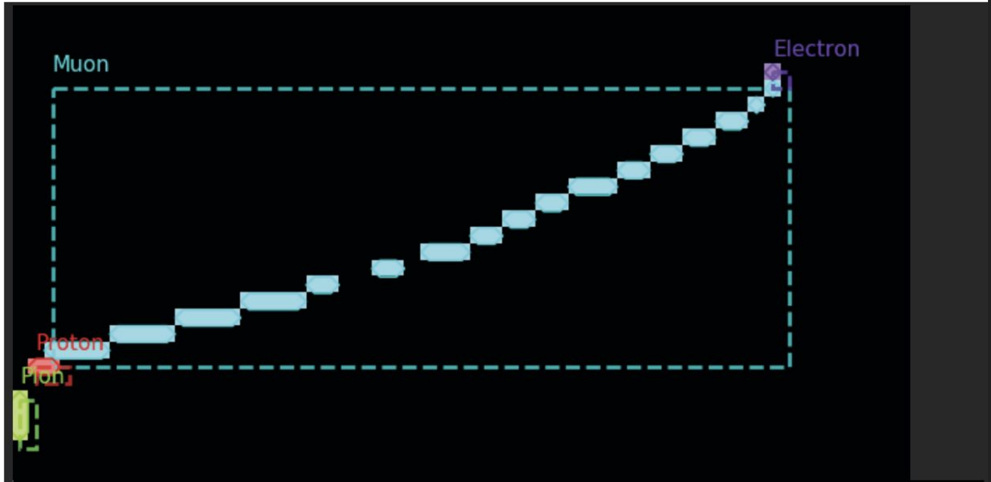
Particle clusters are the base unit for neutrino energy estimation.

Clusters are used to identify specific final states for cross section measurements.

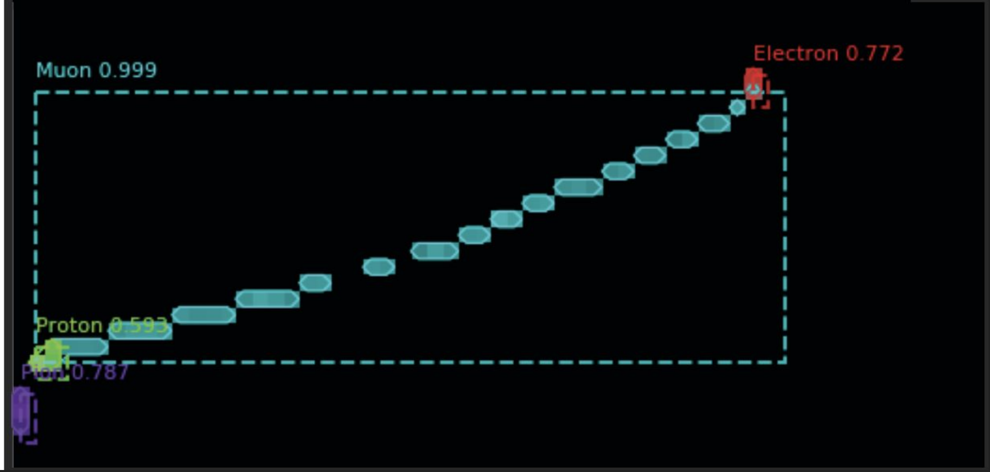


Results Instance Segmentation

True Clusters



Instance Segmentation Clusters



Performance of LSTM EE

To assess the performance of both standard EE and LSTM EE, we define the energy resolution as

$$R = \frac{E_{reco} - E_{true}}{E_{true}}.$$

A smaller Root Mean Square(RMS) of resolution indicates a better performance of the estimator.

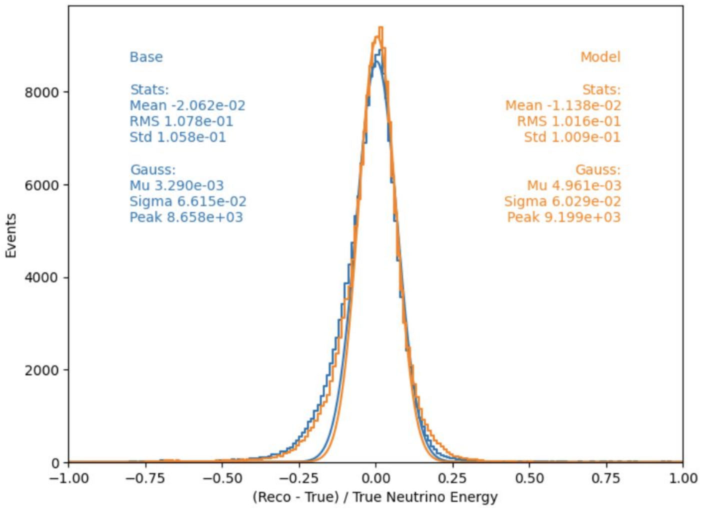


Figure 1: Plot of muon neutrino energy resolutions of standard EE and LSTM EE.

The default neutrino sample we used is an FD FHC sample generated by Genie and GEANT4

LSTM EE has better performance than standard EE in NOvA.

It has
smaller resolution of energy and higher stability with respect to systematics in NOvA.

It help NOvA to determine the neutrino oscillation parameters.