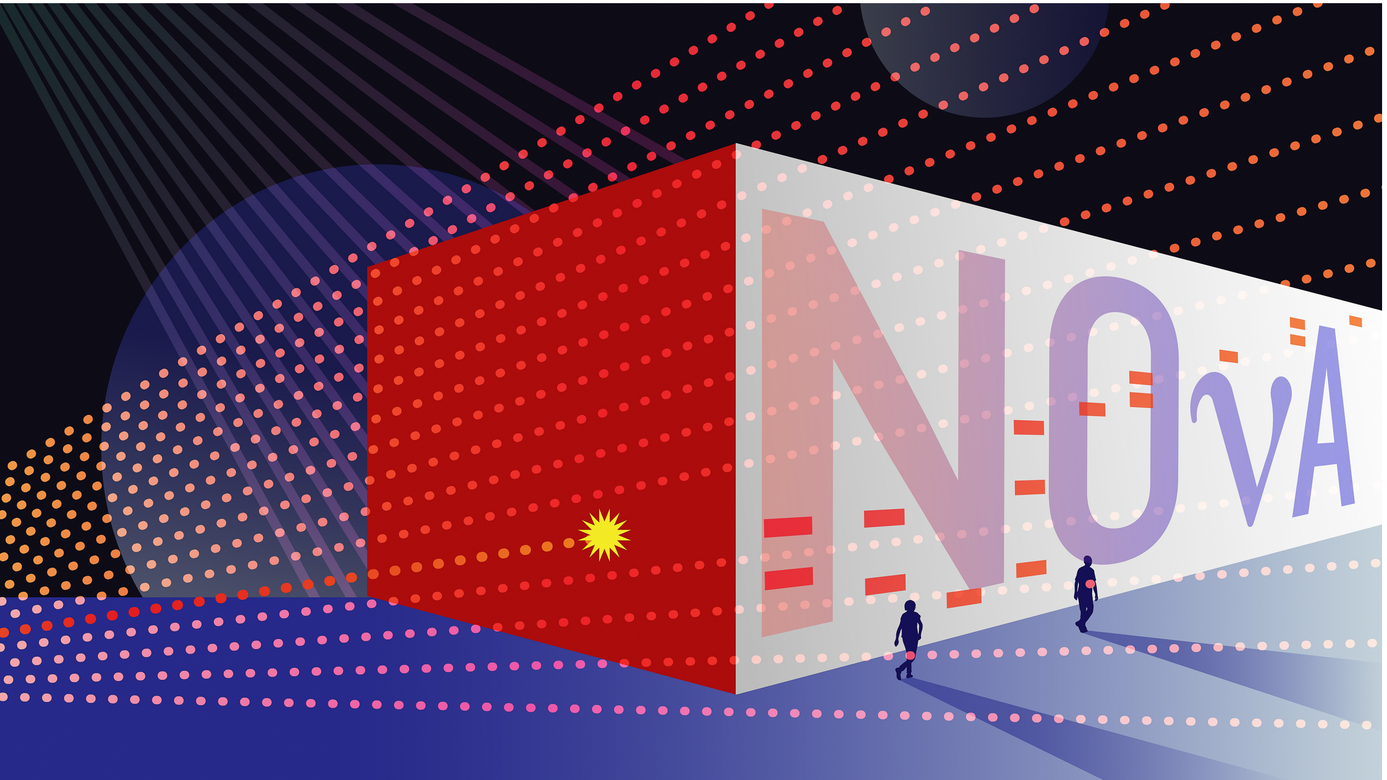
Exploring Computing Methods for Improved Cosmic Background Rejection in NOvA's Sterile Neutrino Searches



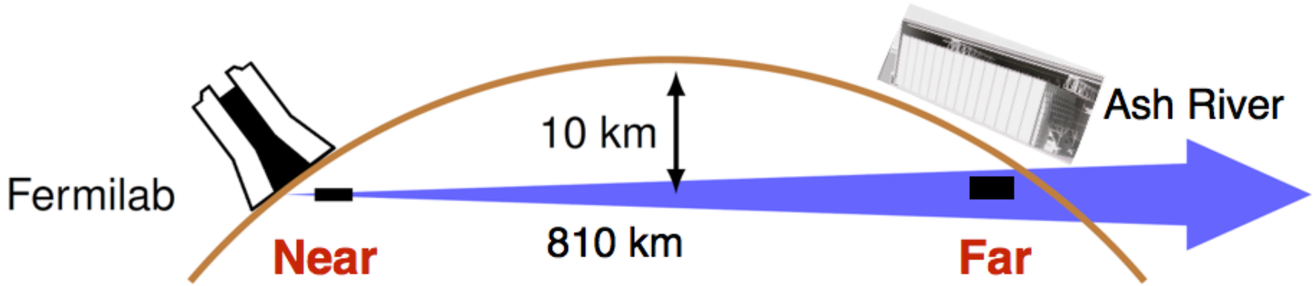
Shaokai Yang
University of Cincinnati
On behalf of the NOvA Collaboration

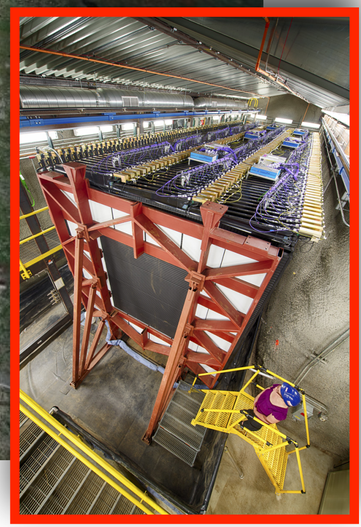
DPF 2017 Meeting
Fermilab, Batavia, IL
08/02/2017

NOvA Experiment

NOvA(NuMI Off-Axis ν_e Appearance) is :

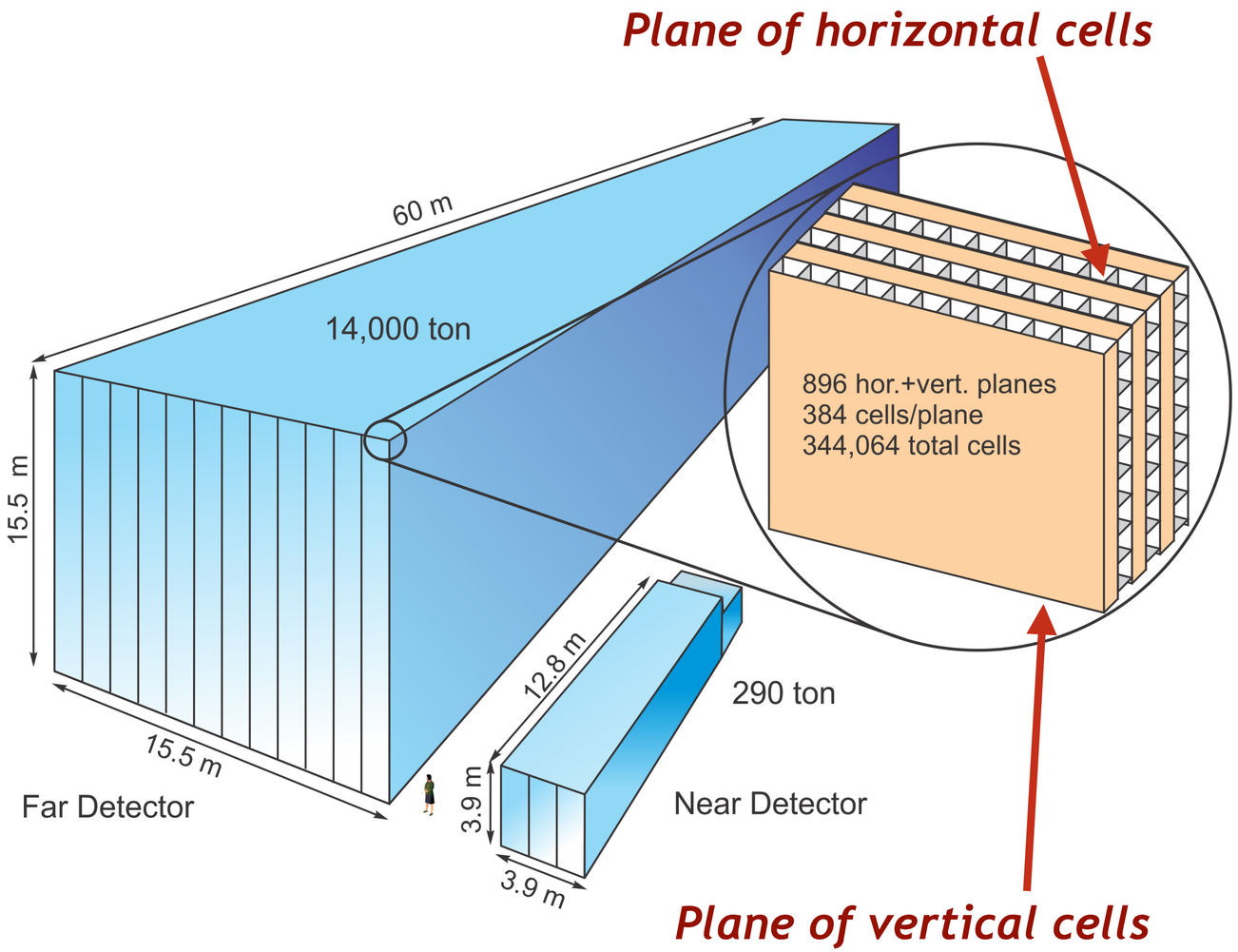
- A second generation long-baseline experiment designed to detect neutrinos in the Fermilab's NuMI (Neutrino at the Main Injector) beam;
- A *14 kt* , Far Detector (FD), 62% active tracking liquid scintillator calorimeter sited *14.6 mrad* off the NuMI beam axis at a distance of *810 km* in Ash River, Minnesota
- A *300* ton Near Detector(ND) functionally identical to the far detector sited underground at a distance of *1 km* from the NuMI beam origin at Fermilab.





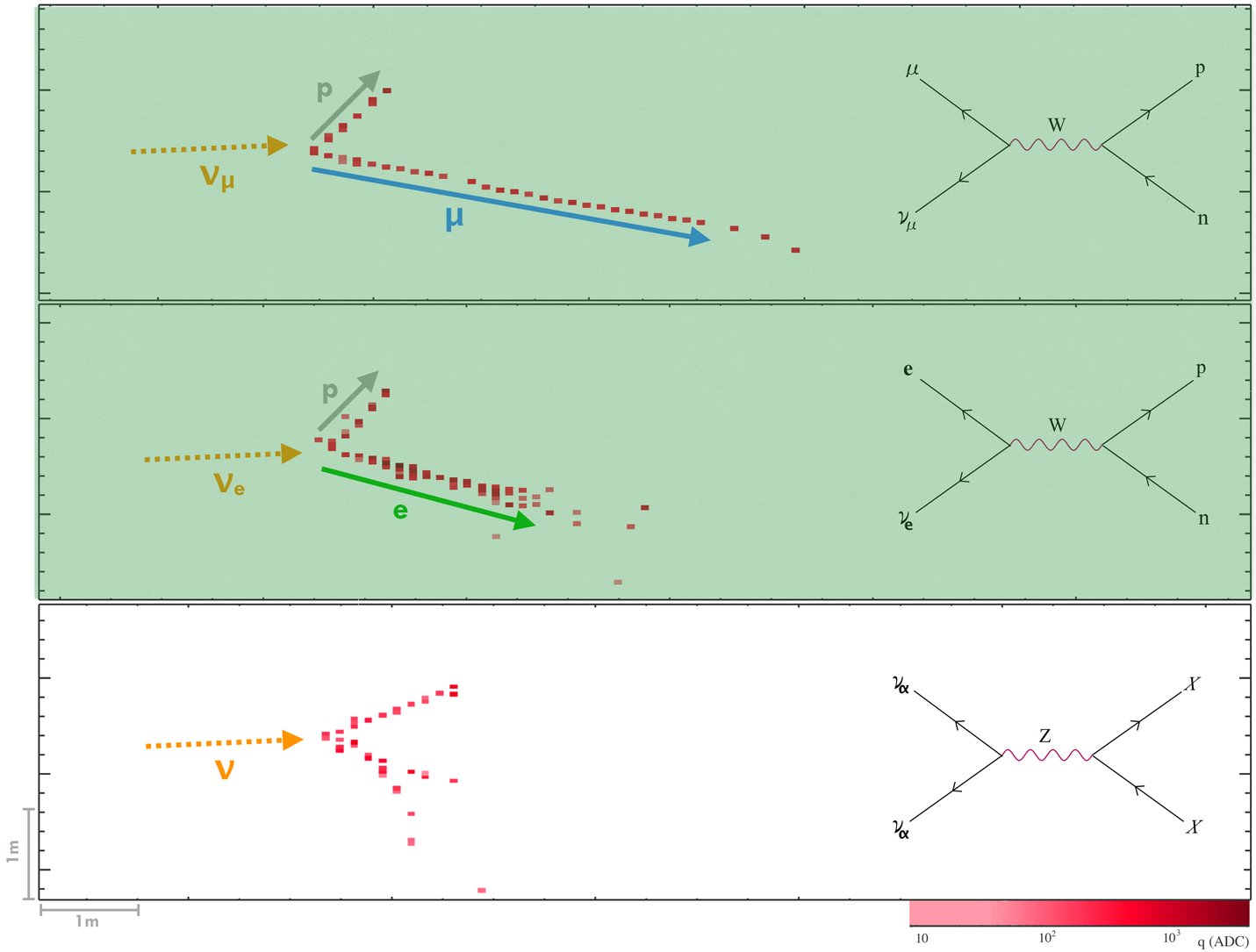
NOvA Detectors

- Composed of extruded PVC cells filled with liquid scintillator;



- NOvA Far Detector is on the surface, hence exposed to abundant rate of cosmic rays. The detector was covered by overburden (4 feet concrete and 6 inches or barite) which designed to shield about 14 radiation lengths.

NOvA Beam Events Topology



ν_μ CC:

a long Muon Track

ν_e CC:

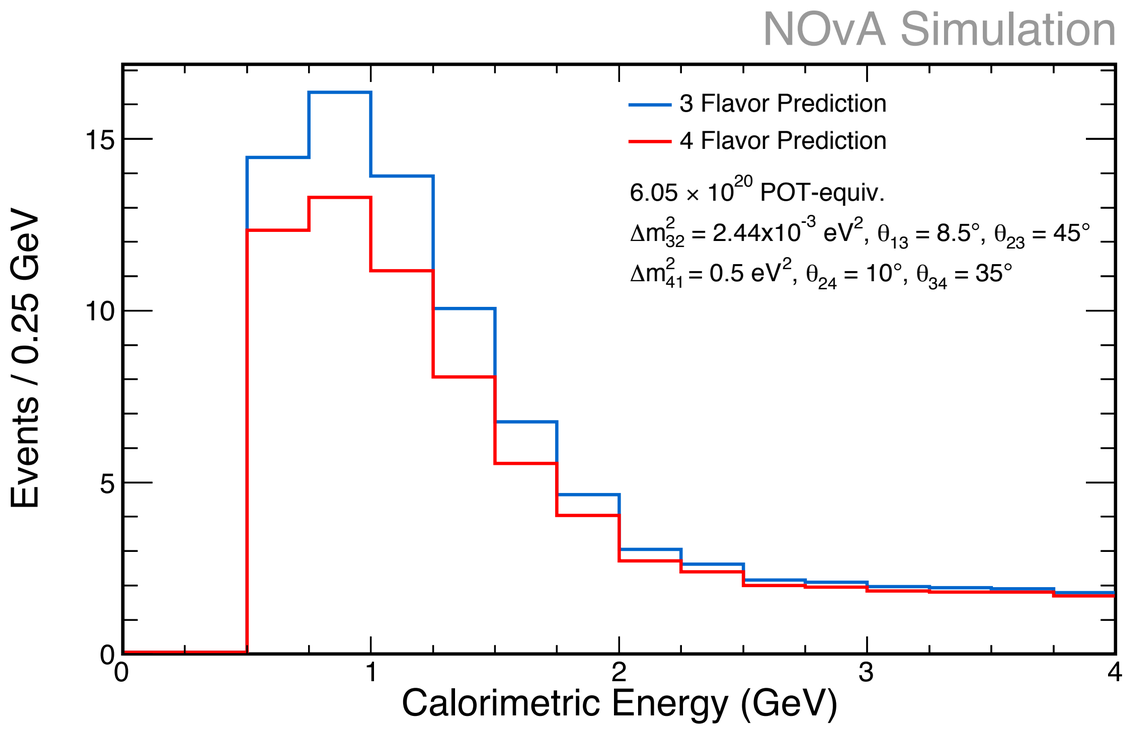
EM Shower with shorter, wider, fuzzy prongs

NC Signal:

only have recoil cascade and missing outgoing neutrino

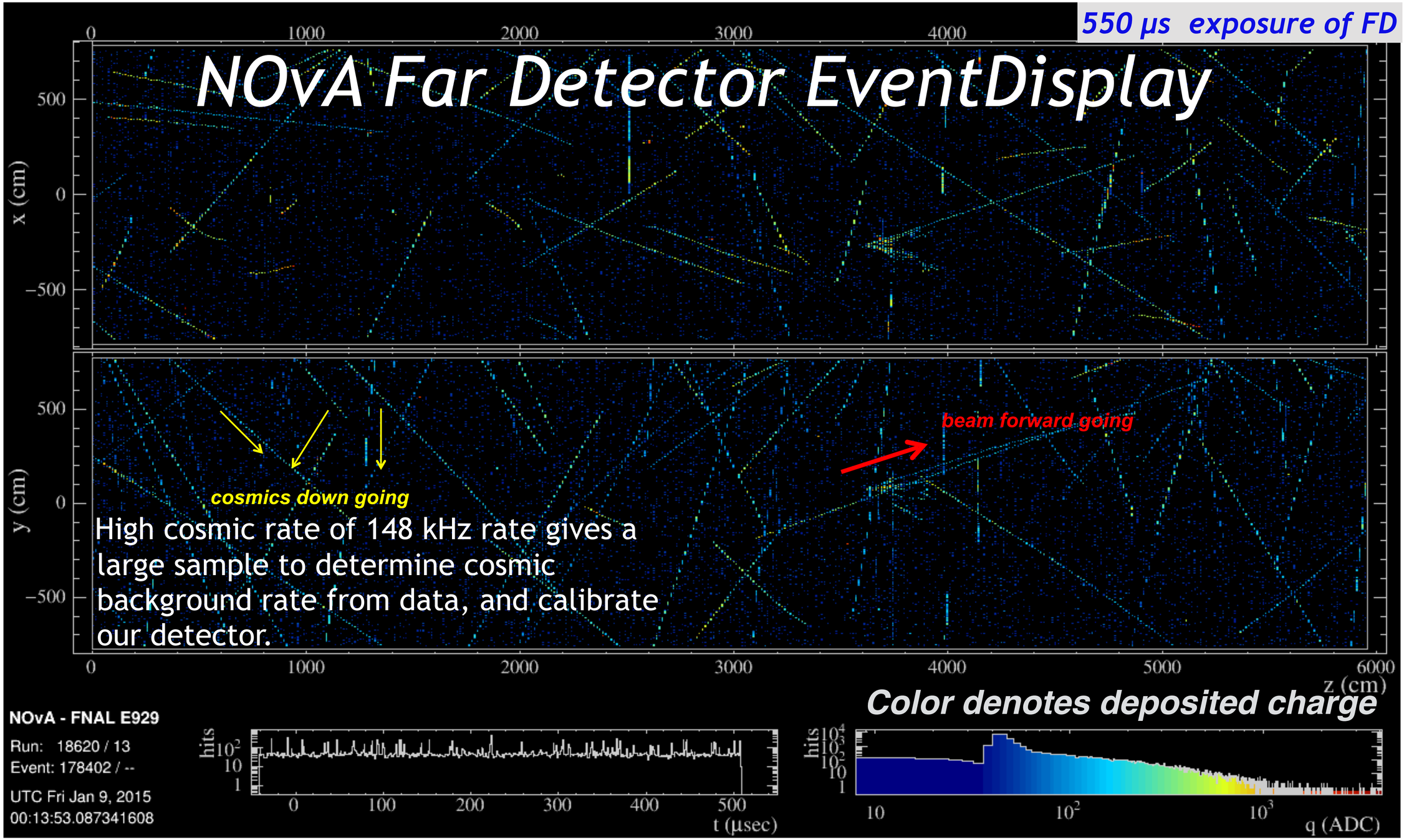
Looking for Sterile Neutrinos in NOvA

- Long-baseline experiments, such as NOvA, provide us an excellent opportunity to test for sterile neutrinos by comparing the observed interaction rates in the Near and Far detectors;
- Unlike charged-current (CC) interactions, neutral-current (NC) interaction rates are not affected by three-flavor oscillations;
- Active-Sterile neutrino mixing would reduce the NC interaction rate in FD as sterile neutrinos would not interact in the detector;
- An energy dependent reduction of the number of NC interactions in the FD with respect to the ND would be a clear evidence of oscillations into sterile neutrinos.



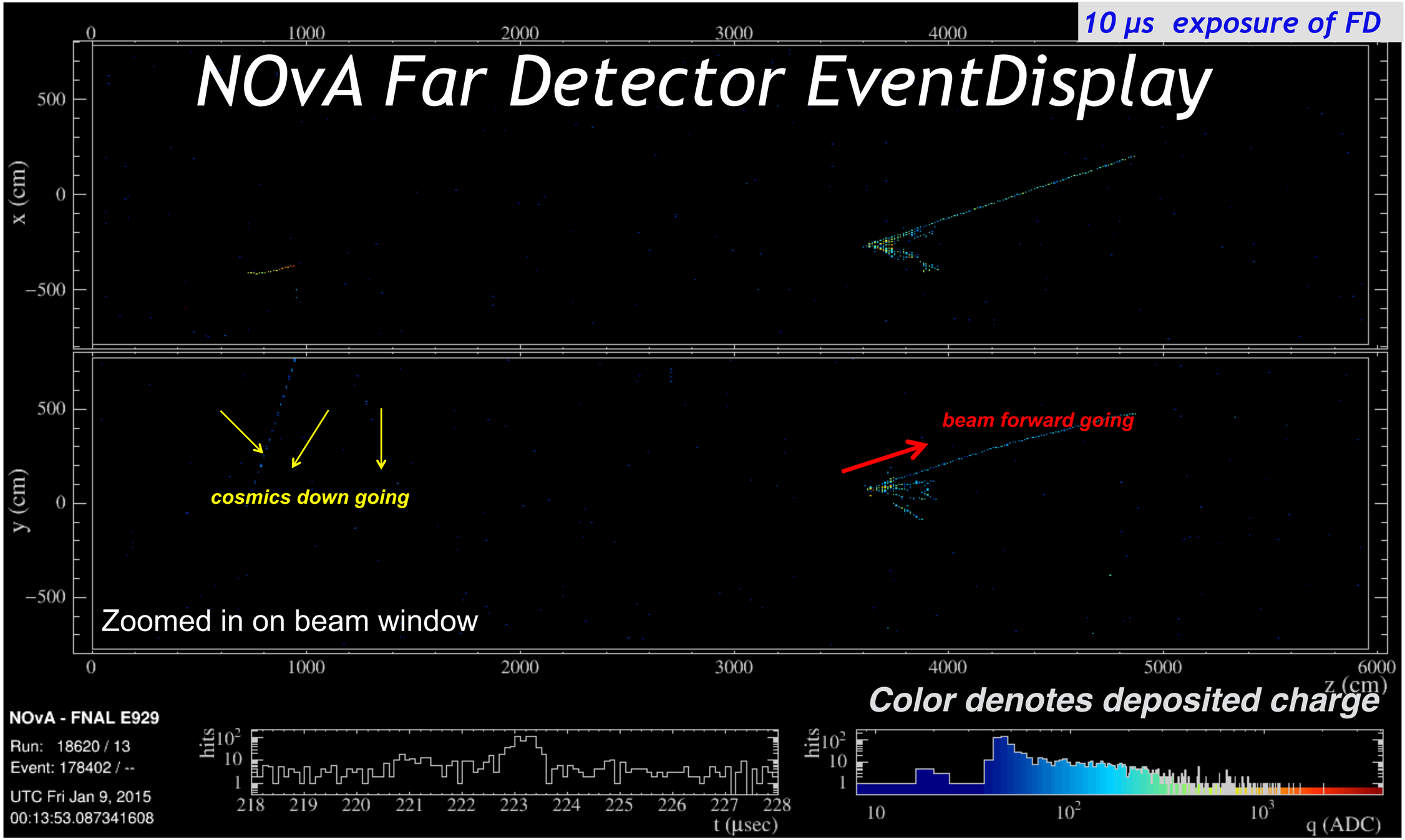
NC disappearance relative to 3-flavour predictions is model independent

NOvA Far Detector EventDisplay



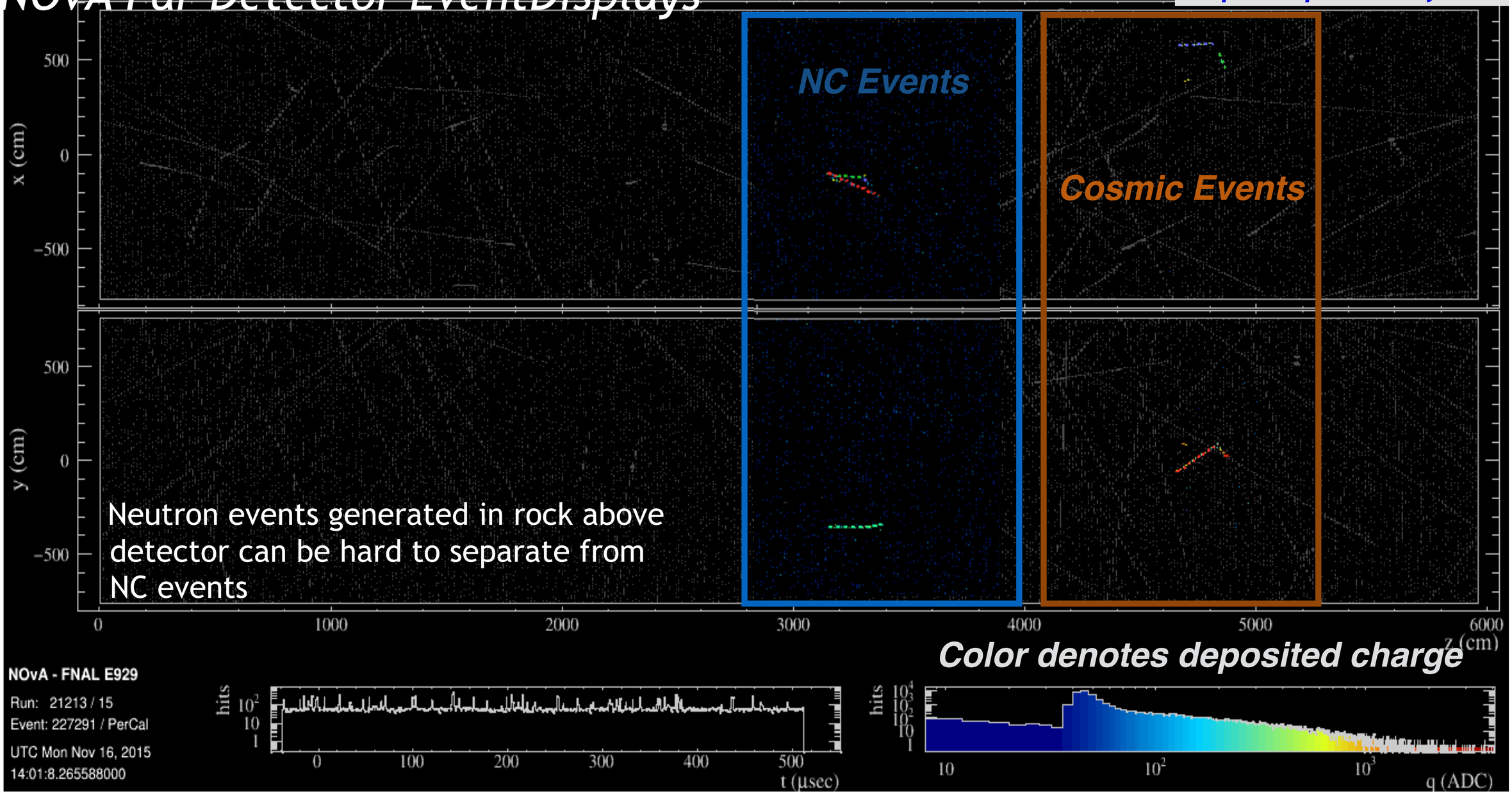
High cosmic rate of 148 kHz rate gives a large sample to determine cosmic background rate from data, and calibrate our detector.

NOvA Far Detector EventDisplay

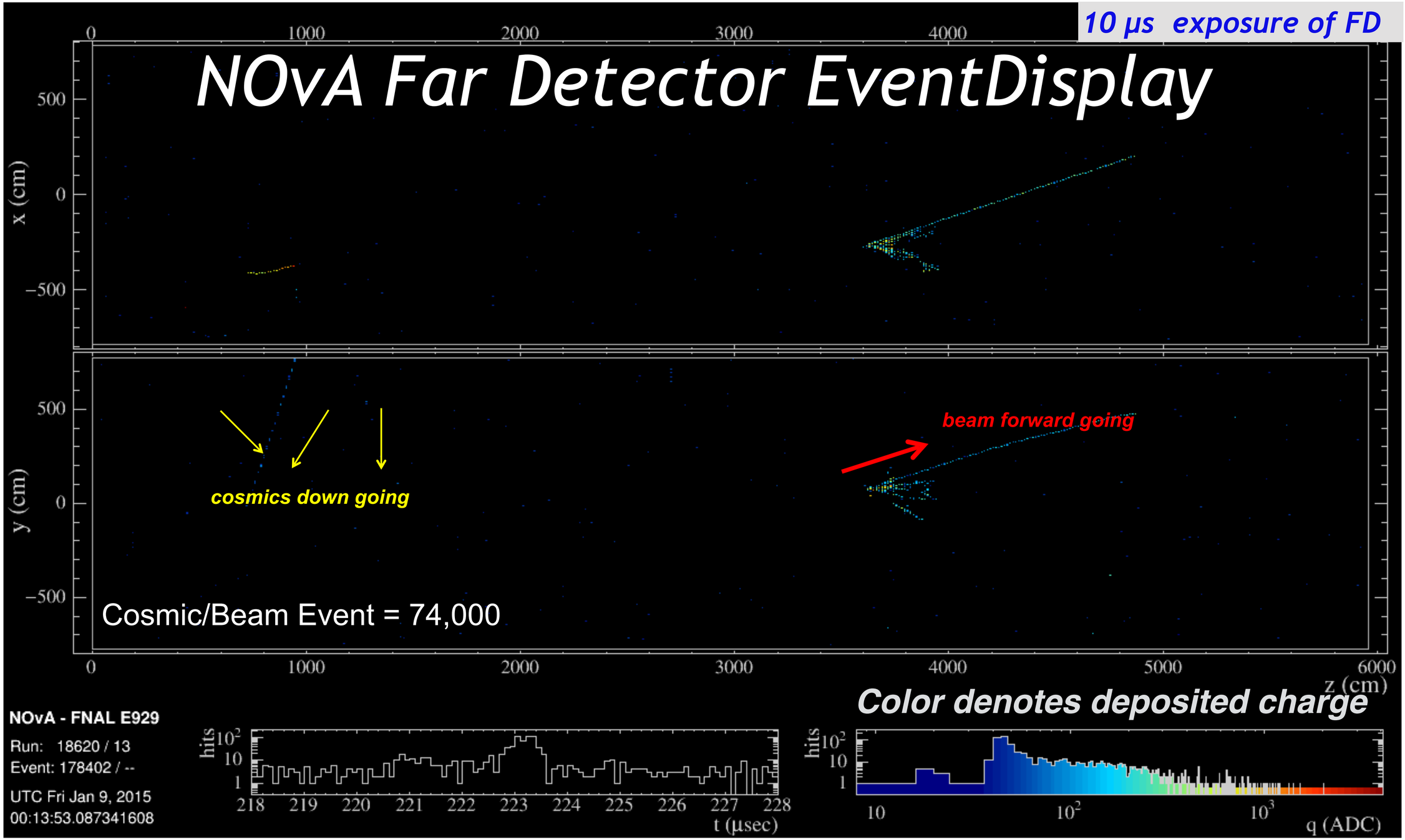


NOvA Far Detector EventDisplays

10 μs exposure of FD

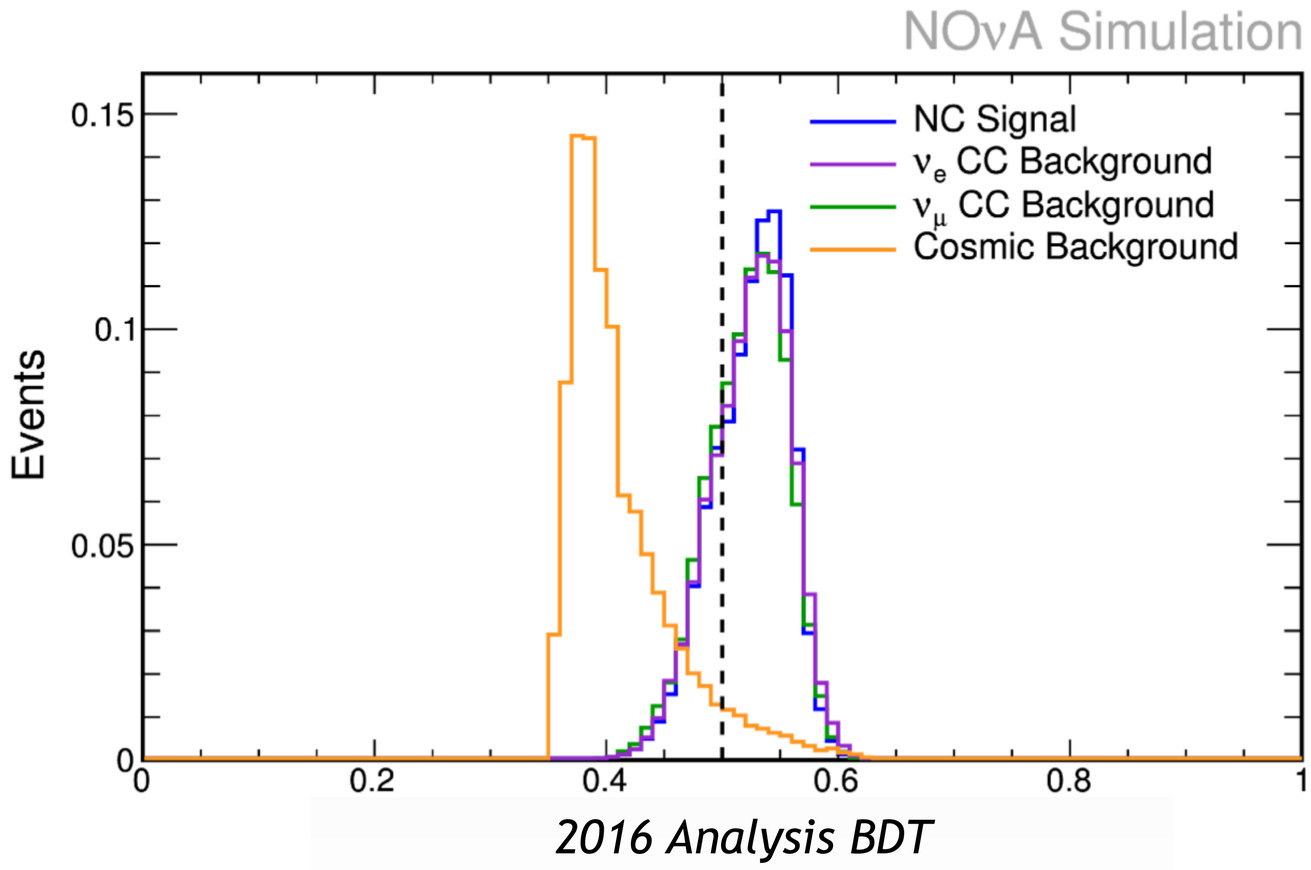


NOvA Far Detector EventDisplay



NOvA Cosmic Rejection Methods for 2016 Analysis

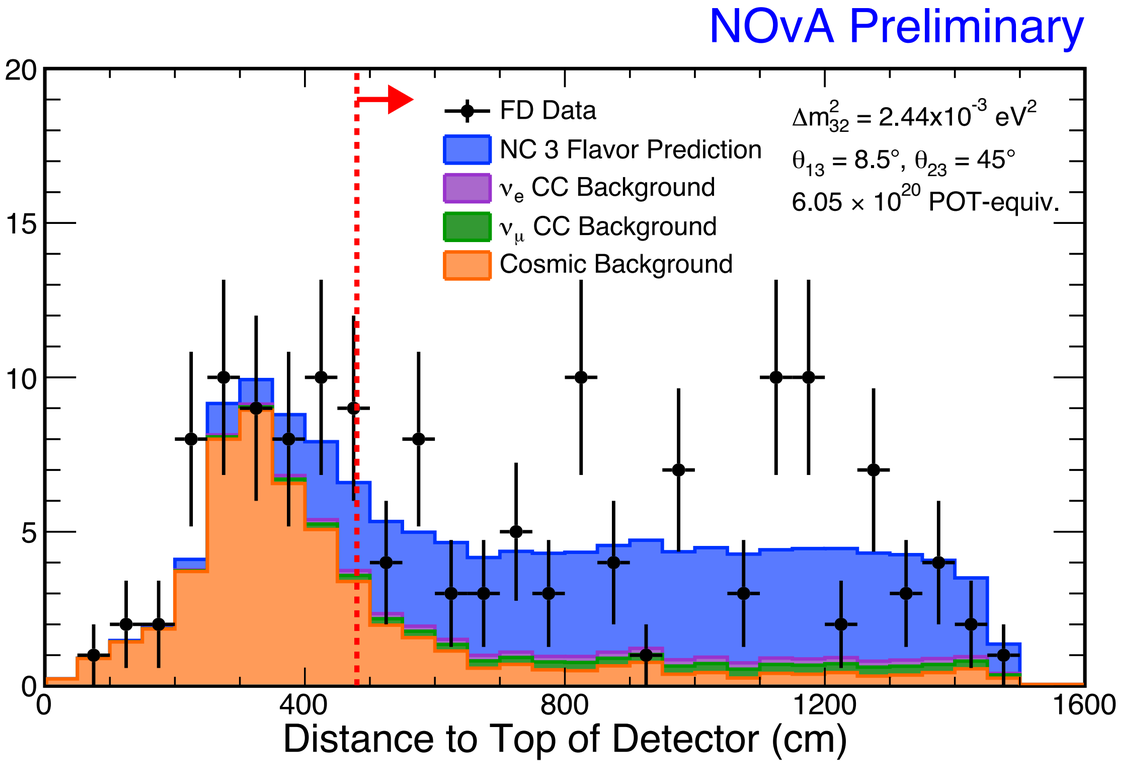
- Used cosmic rejection based on TMVA framework trained to reject cosmic muons from beam events both NC and CC events;
- Training Algorithm: AdaBoosting+ Decision Trees;
- Trained after preselection with containment cuts and event quality cuts;
- Fourteen reconstructed track based variables are taken as the training Inputs:
  1. Leading track direction;
  2. Leading track length;
  3. Number of hits in leading track;
  4. Angle between leading and sub-leading tracks and *so on*.



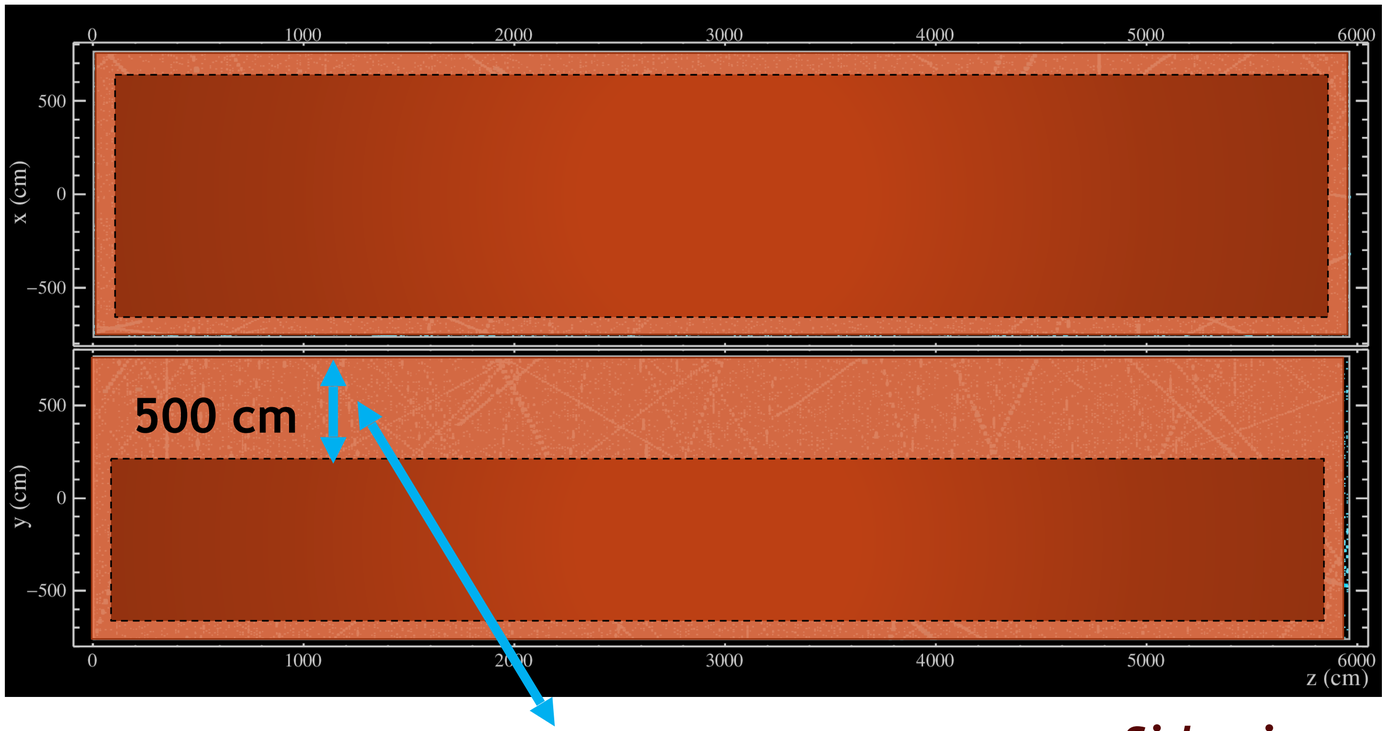
2016 Cosmic-Neutron Rejection

FD overburden creates large number of neutrons
Not removed by standard 2016 Analysis BDT

Remove this background by removing events near top of detector



Top view

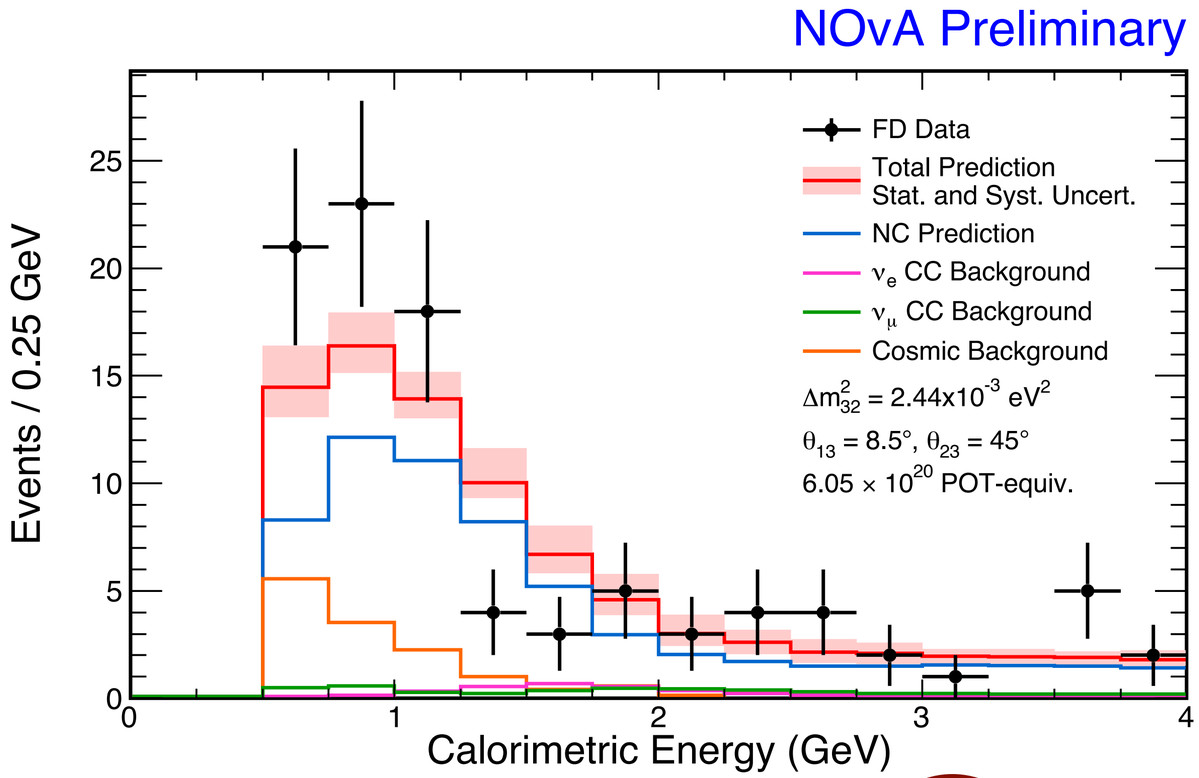


cut all events in top 500 cm of detector

Side view

2016 Analysis Selection Results

- After selection 1 in every 1.7 million cosmic is misidentified as a NC signal event or 1 cosmic per 6 signal.



Observed	Predicted Total	NC	CC	Cosmic	POT
95	83.5	60.6	8.6	14.3	6.05E+20

See poster:
"Sterile neutrino search in the NOvA Far Detector"
Sijith Edayath

See talk:
"Sterile neutrino search in the NOvA Far Detector"
Gavin Davies

See paper:
"Search for active-*sterile neutrino mixing using neutral-current interactions in NOvA"*
(arXiv:1706.04592)

Cosmogenic events are still the dominant background after cutting off 5 meters from the top of the far detector

Improving NOvA NC-Cosmic Rejection

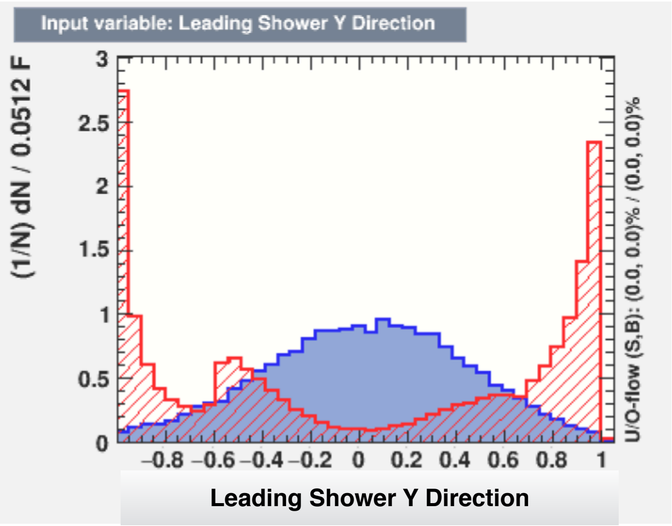
- 2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers
  - AdaBoosting DT
  - Gradient Boosting DT

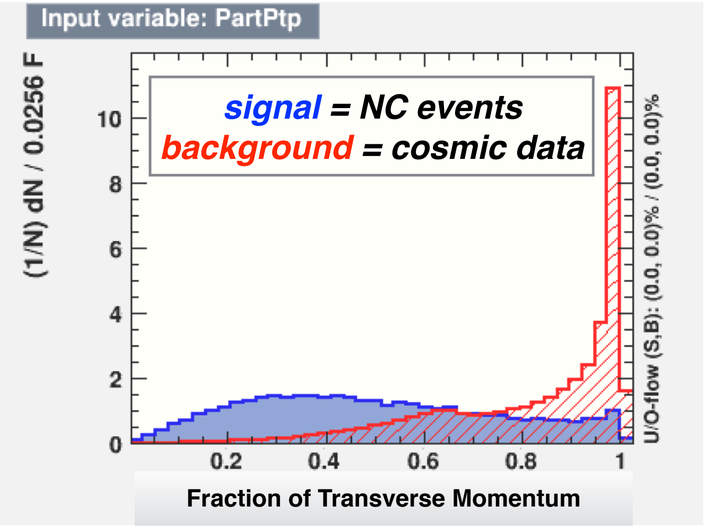
- Future Analysis NC-Cosmic Rejection: TensorFlow-based Particle Identifiers
  - Neural Network Classification
  - Convolutional Neutral Networks Classification (*Image-based Classification)*
    - VGG Net
    - Residual Net

2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

- Training Aim: Rejection cosmic events from Neutral-Current Neutrino Interactions;
- Trained after preselection with containment cuts, event quality cuts and cuts to remove muons;
- Thirteen reconstructed variables are selected as the training inputs:

1. CVN Cosmic ID
2. Leading Shower Y Direction
3. Fraction of Transverse Momentum
4. Number of Shower
5. Leading Shower Length
6. X View Number of Hits - Y View Number of Hits
7. X View Number of Hits + Y View Number of Hits
8 .(X View - Y View)/(X View + Y View)
9. Leading Shower Width
10. Leading Shower Gap
11. Leading Prong Number of Hits
12. Number of MIP Hits in the slice
13. Leading Prong Calibrated Energy





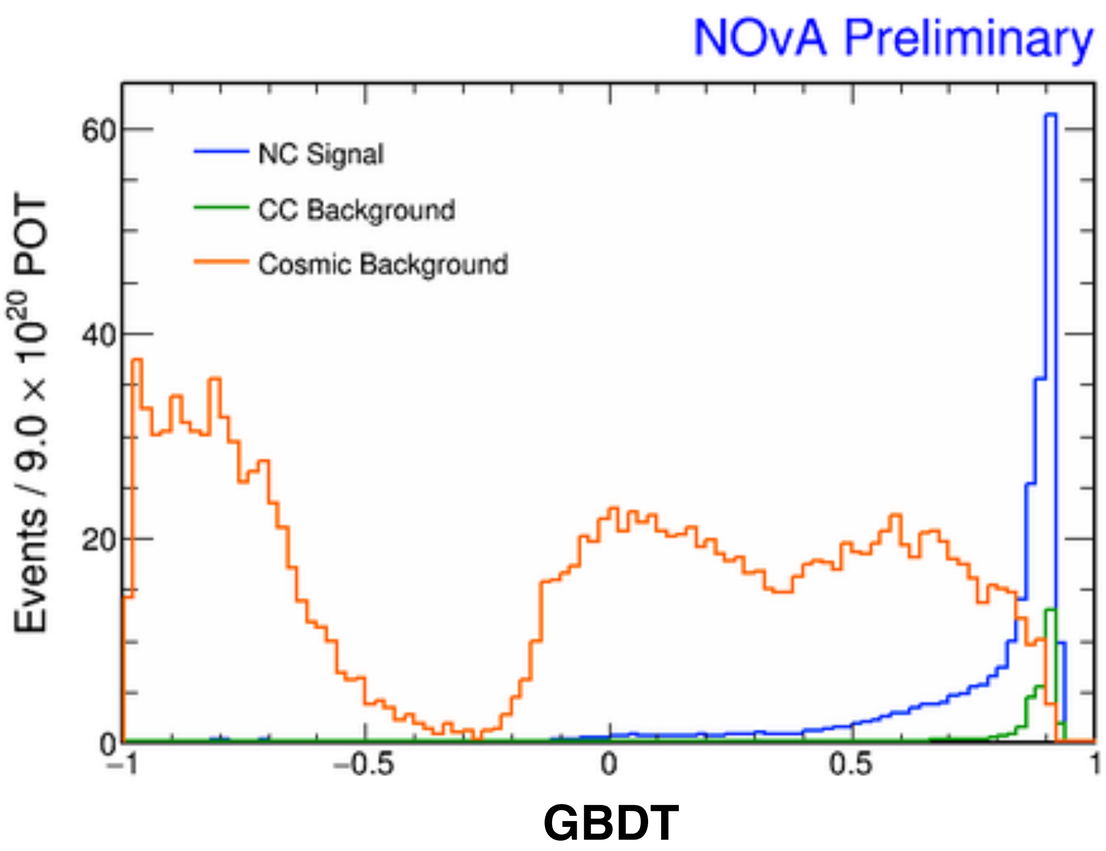
2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

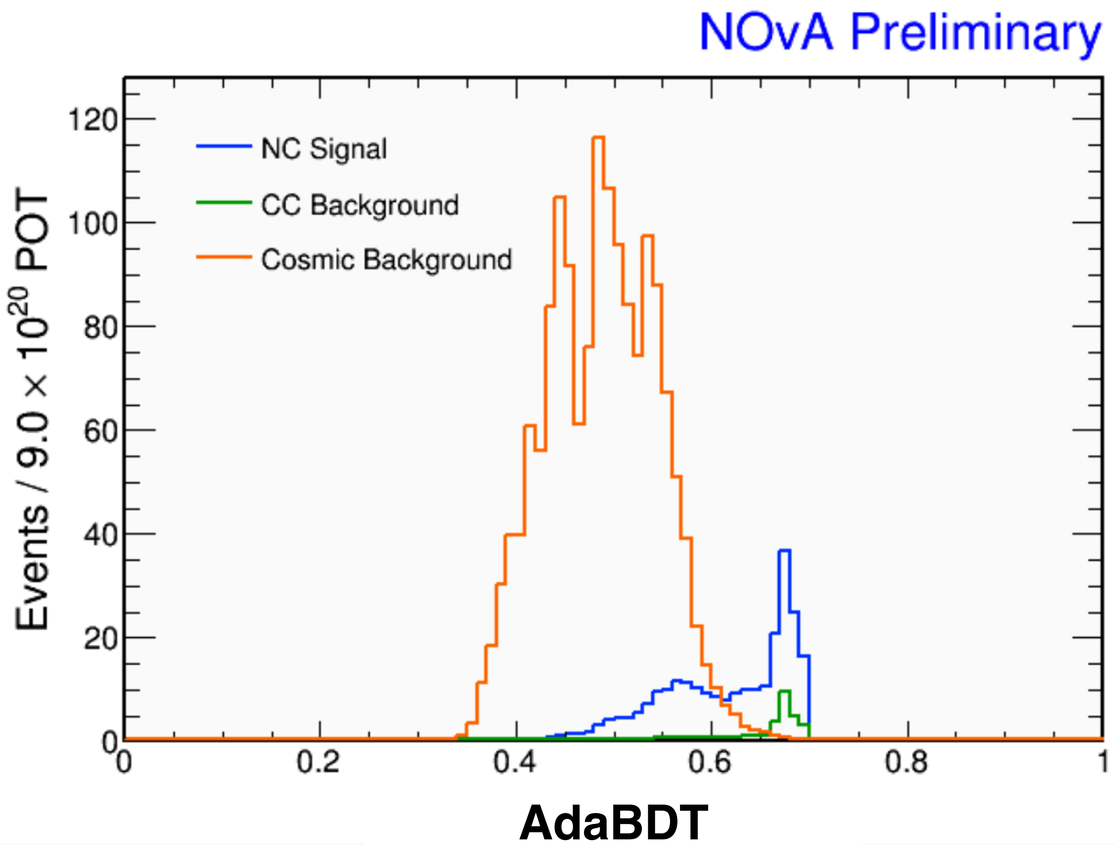
1. Adaptive Boosted Decision Trees:
   a. 500 Trees;
   b. Based on exponential loss function : $L(F, y) = e^{-F(\mathbf{x})y}$
   c. Tree depth = 5;
   d. Node Purity Limit = 0.9;

2. Gradient Boosted Decision Trees:
   a. 1000 Trees;
   b. Based on binomial log-likelihood loss function : $L(F, y) = \ln\left(1 + e^{-2F(\mathbf{x})y}\right)$
   c. Tree depth = 5;
   d. Node Purity Limit = 0.9;

2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

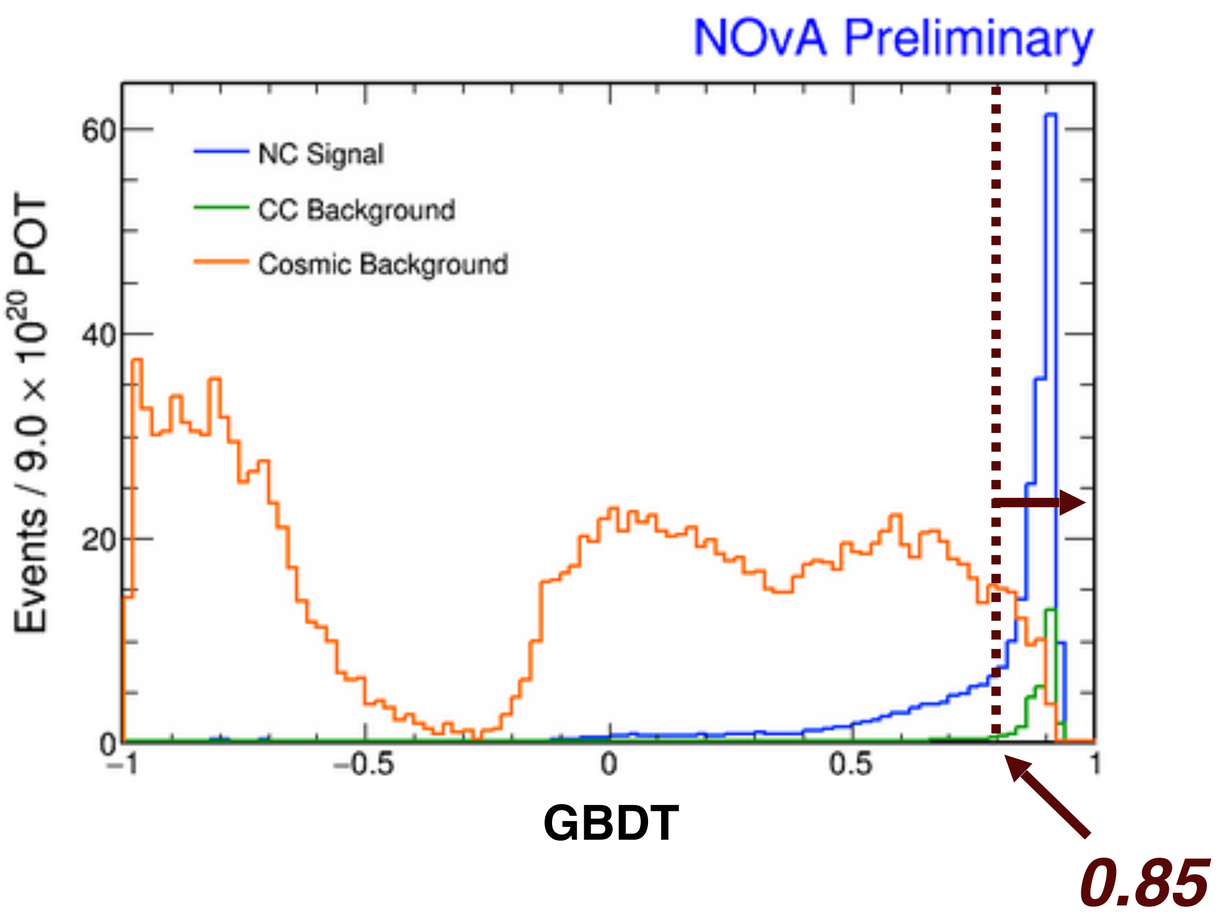
- Output of training on NOvA FD simulation and with cosmic data for GBDT and AdaBDT

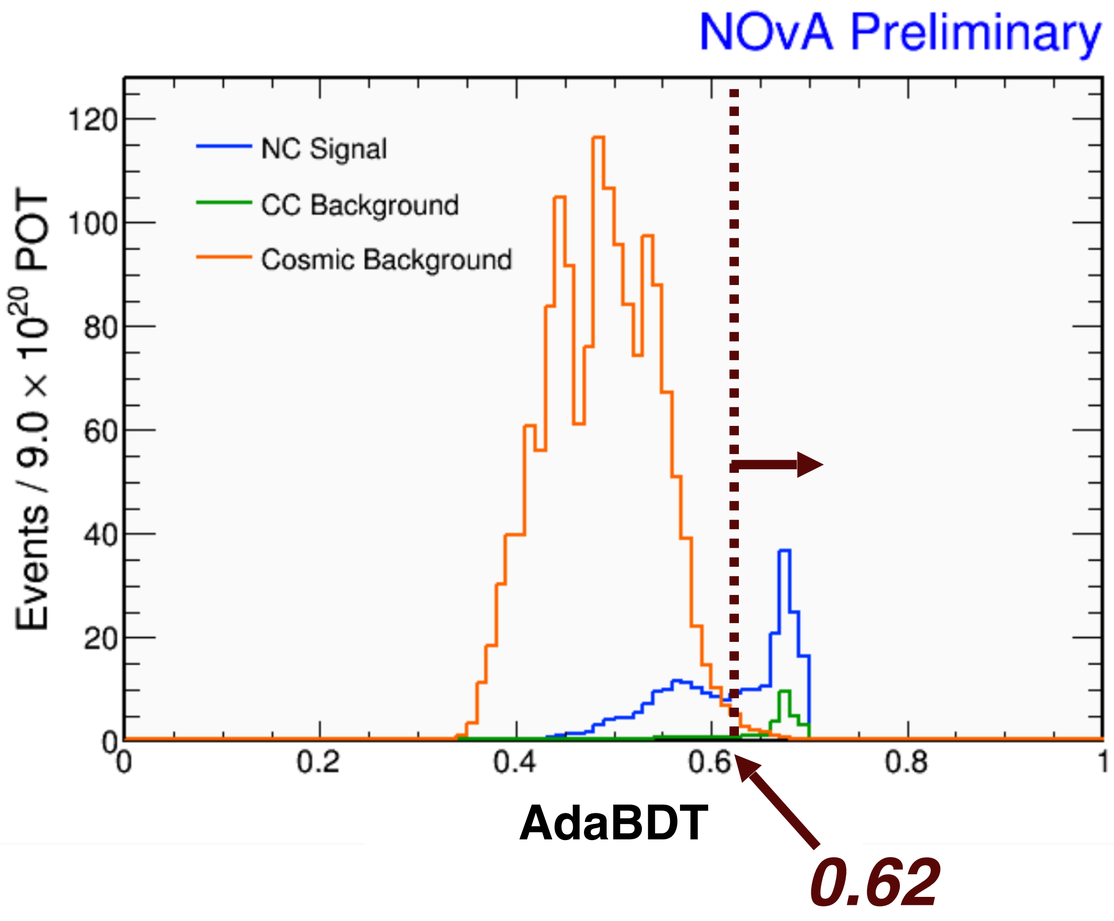




2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

- Output of training on NOvA FD simulation and with cosmic data for GBDT and AdaBDT





2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

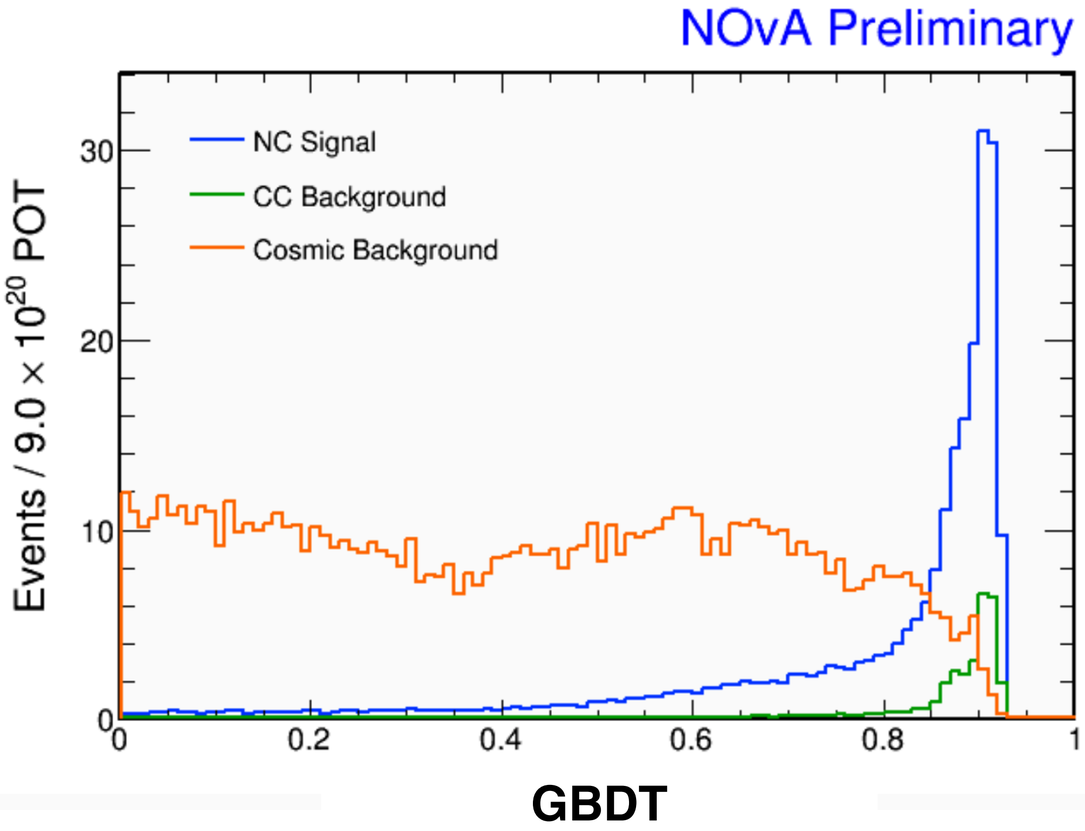
- Output of training on NOvA FD simulation and with cosmic data for GBDT and AdaBDT

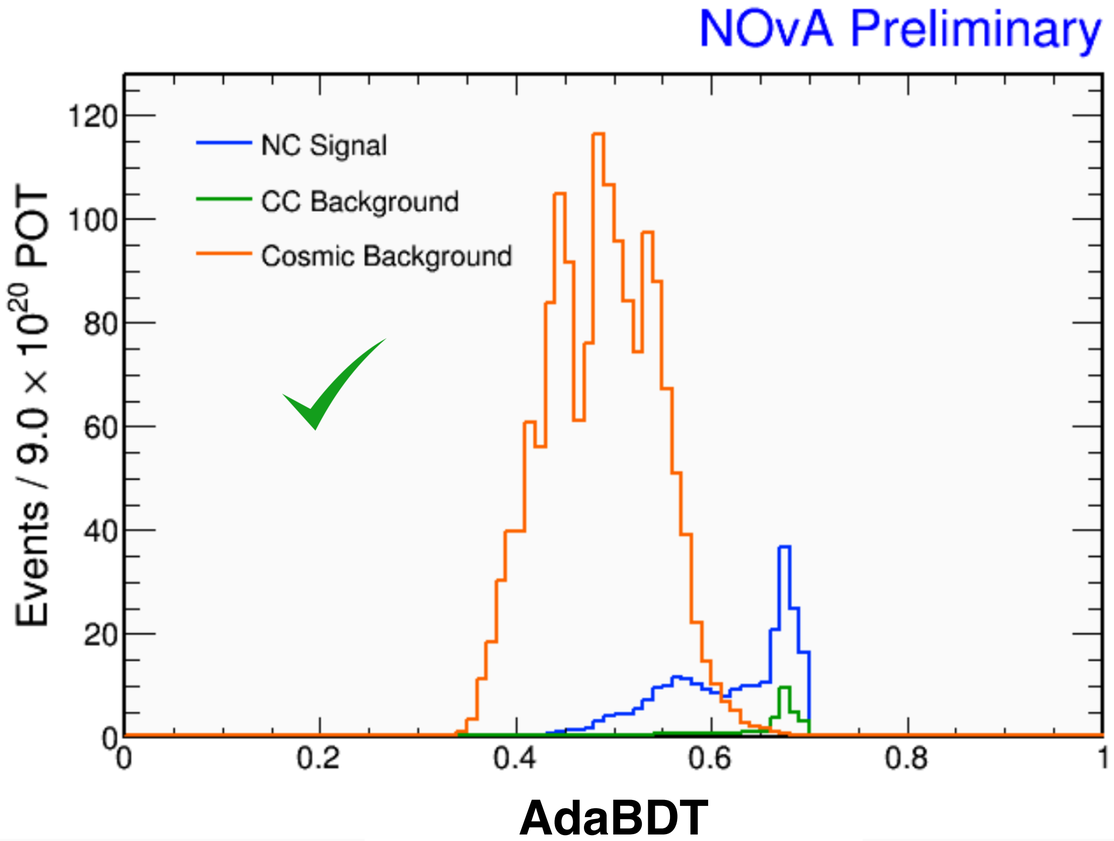
	Predicted Total	NC	CC	Cosmic	POT	FoM
2016 Analysis BDT	124.21	90.1	12.8	21.2	9E+20	8.08
GBDT (cut on 0.85)	201.91	140.7	25.9	35.25	9E+20	9.90
AdaBDT (cut on 0.62)	178.05	137.4	25.4	17.18	9E+20	10.24

FoM =Signal/Sqrt(Signal+Background)

2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

- Output of training on NOvA FD simulation and with cosmic data for GDBT and AdaBDT





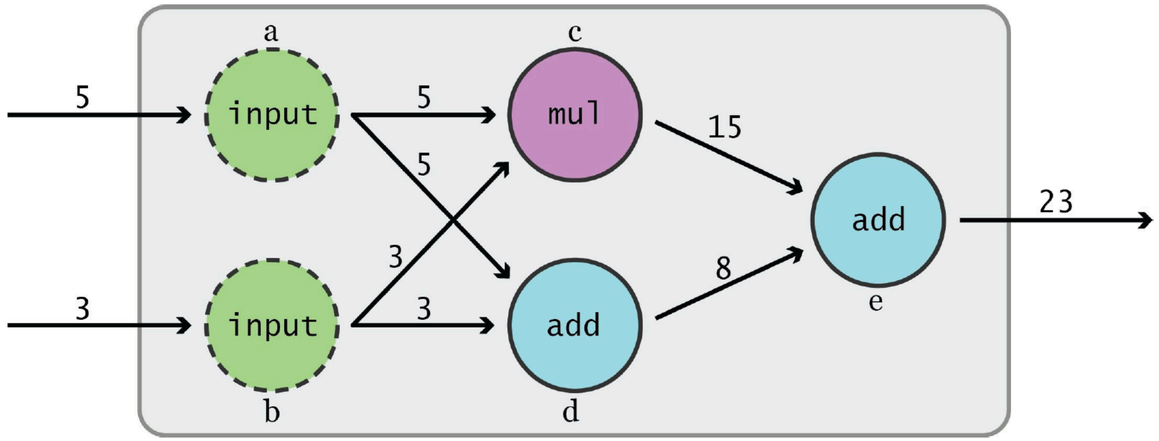
Future Analysis NC-cosmic rejection: TensorFlow-based Particle Identifiers

So, what is TensorFlow?

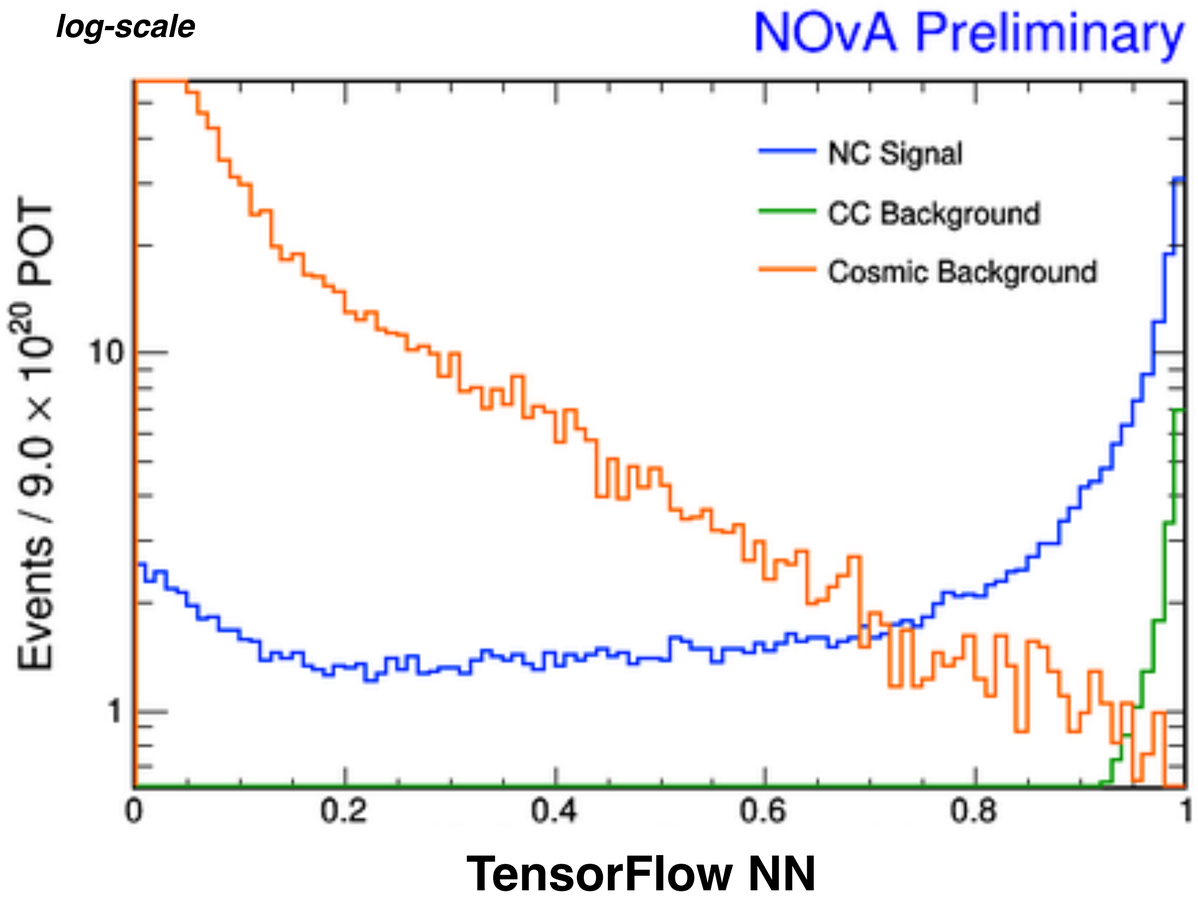
TensorFlow is the interface for defining machine learning models, training them with data, and exporting them for further use. At a high level, TensorFlow is a Python library that allows users to express arbitrary computation as a graph of *data flows*.

Then, what is Tensor?

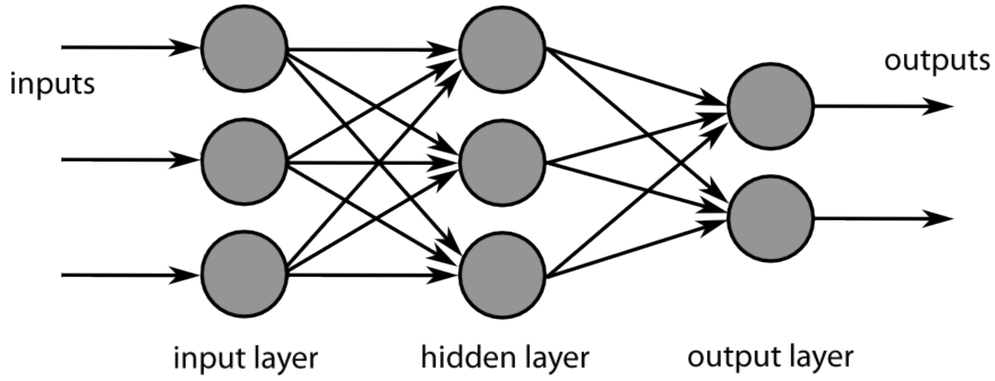
Data in TensorFlow are represented as tensors, which are multidimensional arrays.



Neural Network Classification

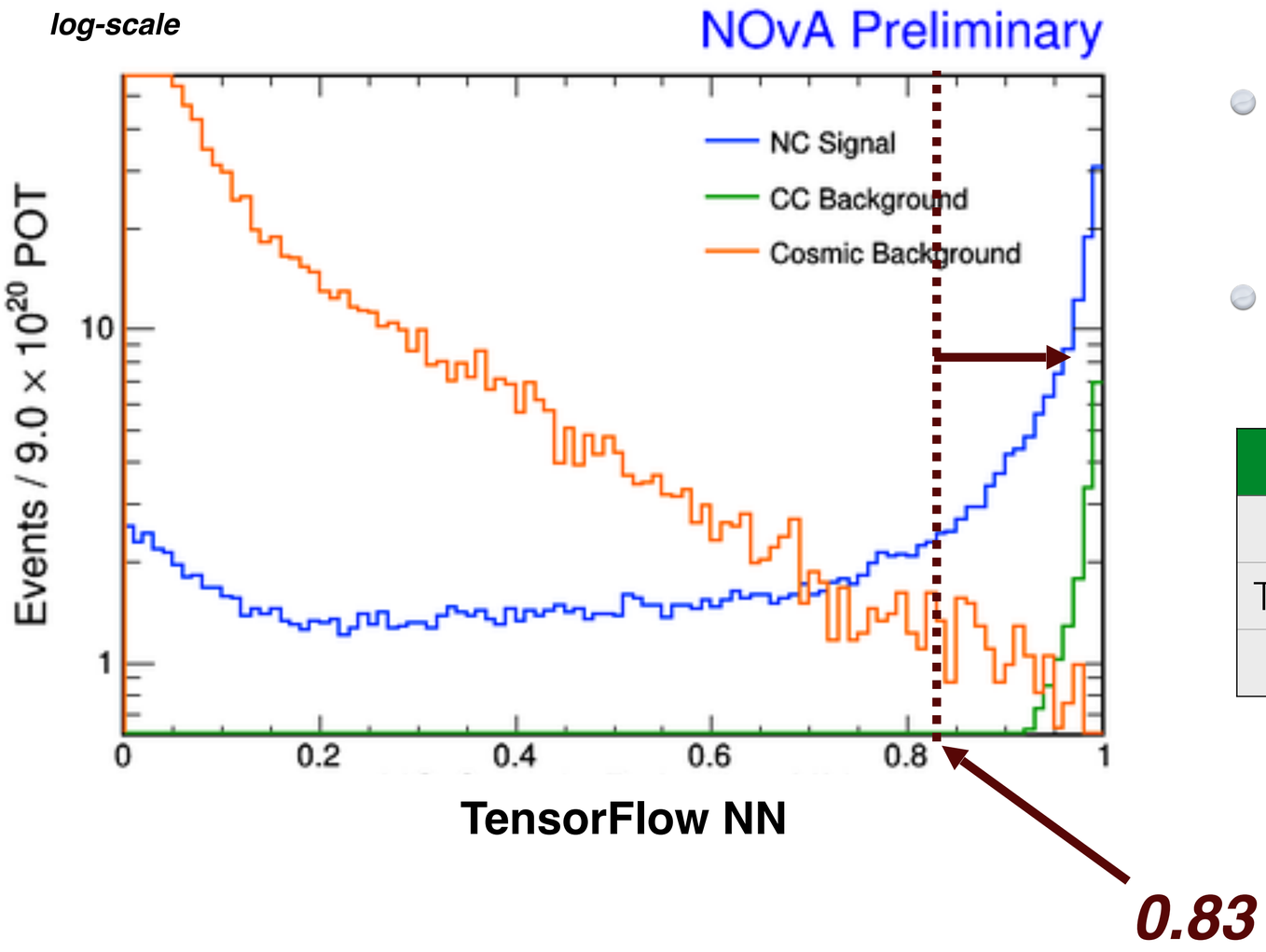


Same Inputs as BDTs



Five Hidden Layers

Neural Network Classification

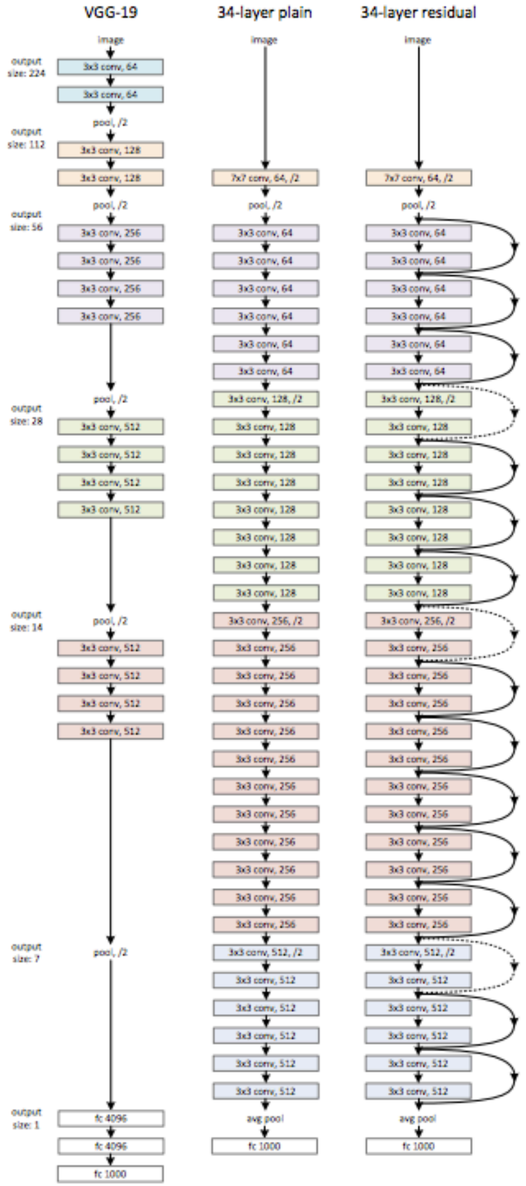


- Using same training file and selection and variables we trained NN
- Five hidden layers

	Total	NC	CC	Cosmic	FoM
2016 BDT	124.21	90.1	12.8	21.2	8.08
TensorFlow NN	169.83	123.14	20.36	16.43	9.7
AdaBDT	178.05	137.4	25.4	17.18	10.24

Convolutional Neural Network Architectures

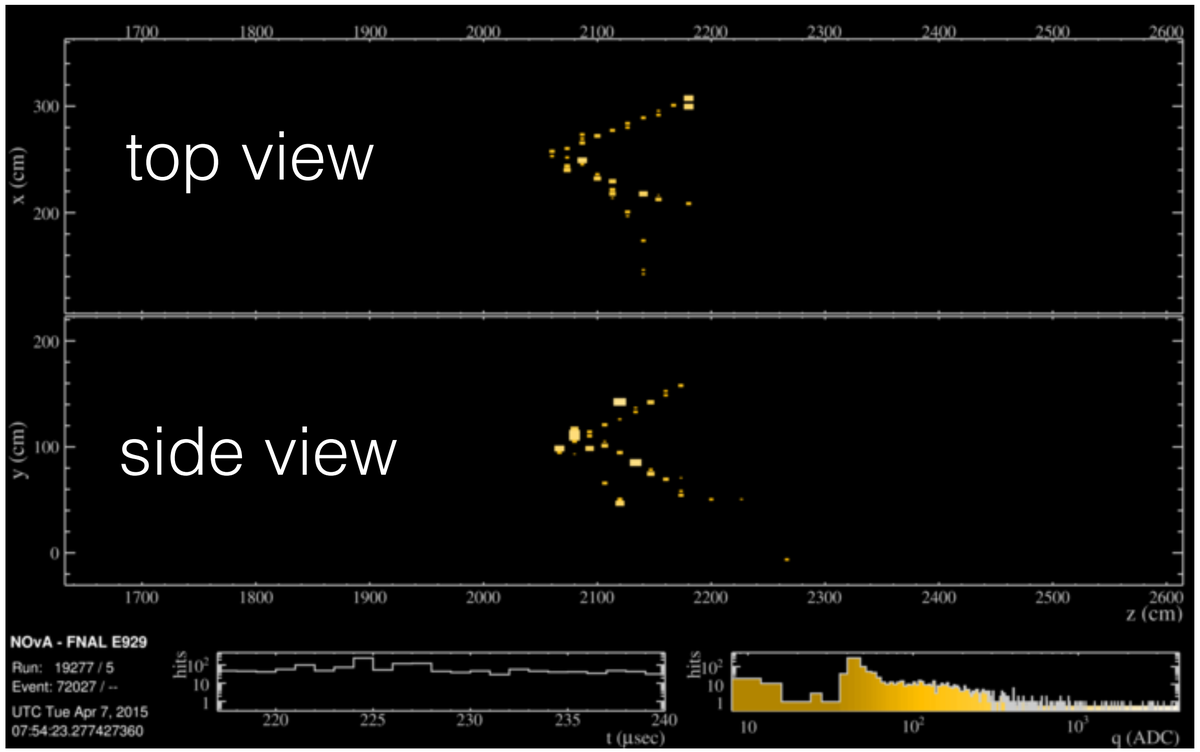
- Visual Geometry Group (VGG)
  - An extremely homogeneous architecture that only performs 3x3 convolutions and 2x2 pooling from the beginning to the end;
  - Final best network contains 16 CONV/FC layers;
  - Therefore, using a lot more memory.



- Residual Network (ResNet)
  - Make it much easier to train really deep models;
  - Really easy-to-implement

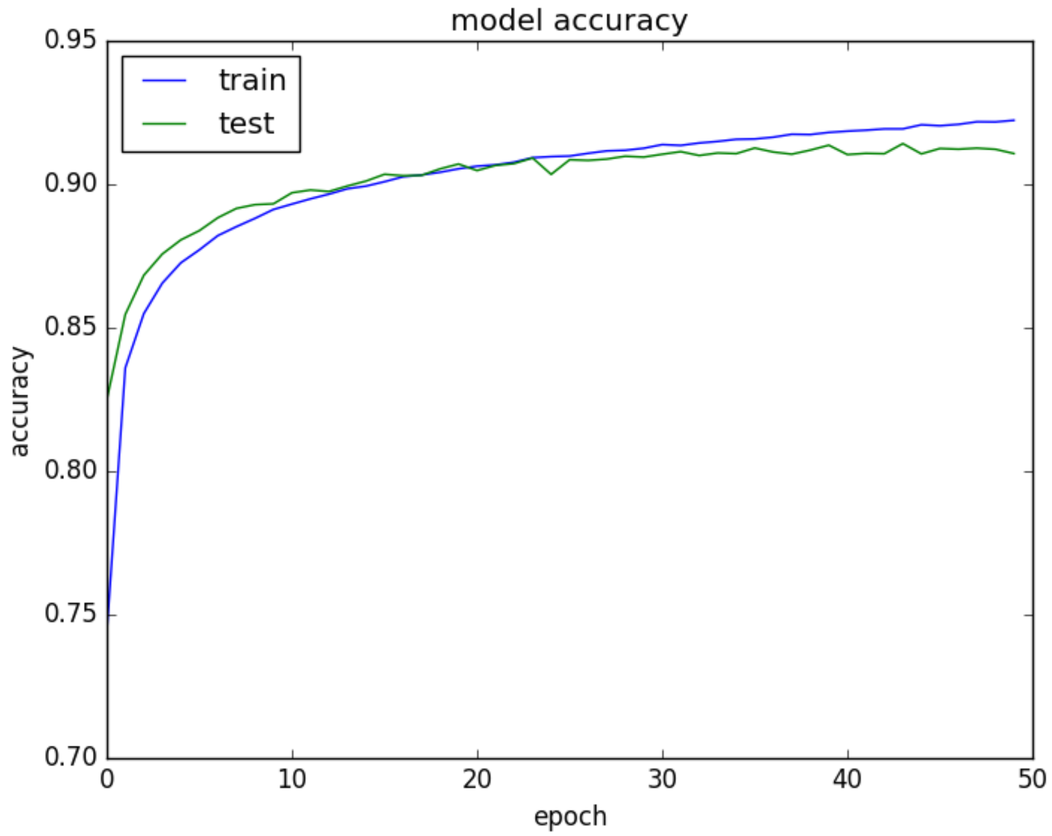
Future Analysis NC-cosmic rejection: TensorFlow-based Particle Identifiers

- We trained VGG-Like network with 3 Blocks and 2 Dropout Layers;
- We also trained a 50 layers ResNet;
- We used the eventdisplay as the input with the top and side views trained separately;
- We applied containment and event quality cuts to select the input events.



Future Analysis NC-cosmic rejection: TensorFlow-based Particle Identifiers

VGG-like architecture Results

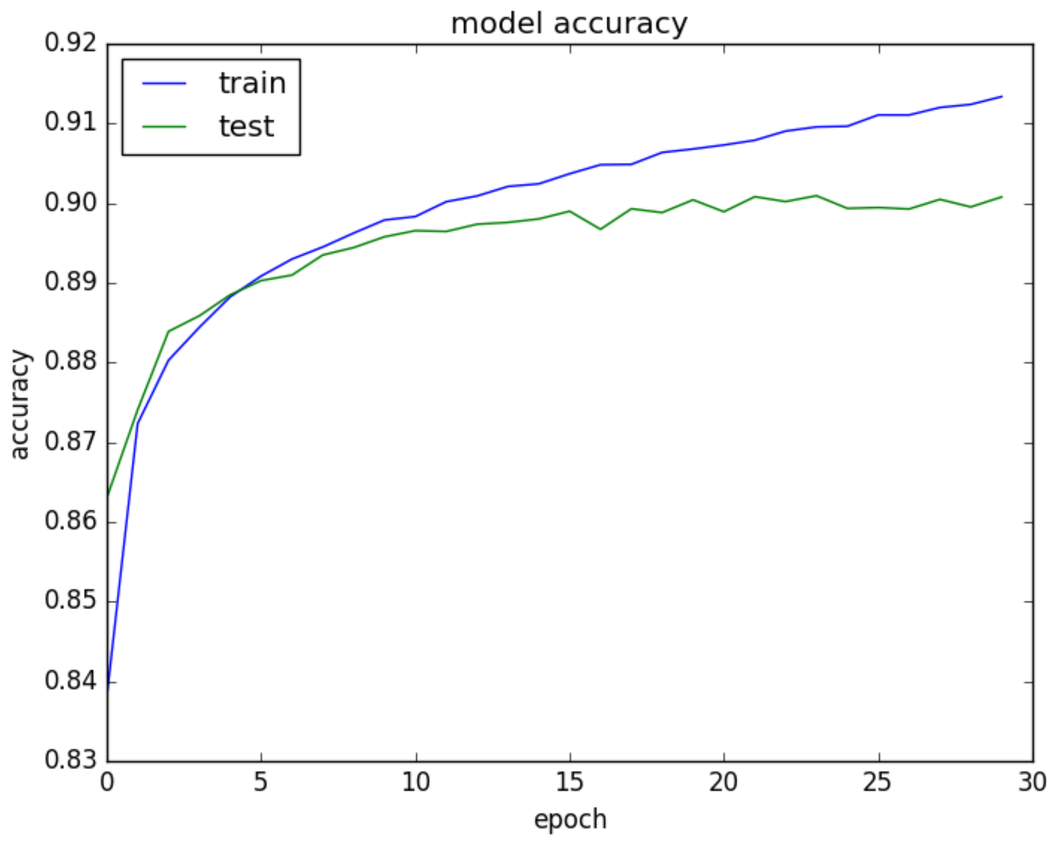


- Selected 4.5 million events.
- 2 million cosmic, 0.5 million NC, 1million NuE, and 1 million NuMu events.
- Only consider the calibrated hits' energy deposition;
- 80% train 20% test.

We produced preliminary results looking at accuracy
Accuracy = *91.26%*

Future Analysis NC-cosmic rejection: TensorFlow-based Particle Identifiers

ResNet-50 Architecture Results



- Selected 4.5 million events.
- 2 million cosmic, 0.5 million NC, 1million NuE, and 1 million NuMu events.
- Only consider the calibrated hits' energy deposition;
- 80% train 20% test.

We also looked at ResNet and see very similar results
Accuracy = *90.09%*

Conclusions

- NOvA Far Detector is on the surface, which make us face the challenge to remove high rate of cosmic events;
- Dominant background of NOvA 2016 NC analysis is still of cosmic events even all selection cuts has been applied;
- For 2017 analysis, the new TMVA-based cosmic rejection classifier, which show big improvements over the 2016 analysis, will be implemented for the NOvA NC analysis;
- We continue to investigate new methods based on TensorFlow to improve the selection purity and efficiency.

NOvA Collaboration

Backup

2017 Analysis NC-Cosmic Rejection: TMVA-based Particle Identifiers

- Both methods use a set of weak learners. They try to boost these weak learners into a strong learner. I assume that the strong learner is additive by the weak learners;

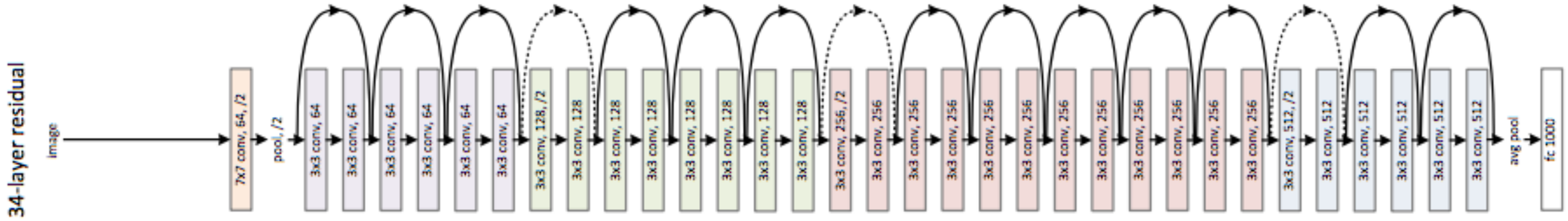
- Gradient boosting generates learners during the learning process. It build first learner to predict the values/labels of samples, and calculate the loss (the difference between the outcome of the first learner and the real value). It will build a second learner to predict the loss after the first step. The step continues to learn the third, forth… until certain threshold;

- Adaboost requires users specify a set of weak learners (alternatively, it will randomly generate a set of weak learner before the real learning process). It will learn the weights of how to add these learners to be a strong learner. The weight of each learner is learned by whether it predicts a sample correctly or not. If a learner is mispredict a sample, the weight of the learner is reduced a bit. It will repeat such process until converge;

- Formulate Adaboost as gradient descent with a special loss function. Generalize Adaboost to Gradient Boosting in order to handle a variety of loss functions.

Convolutional Neural Network Classification

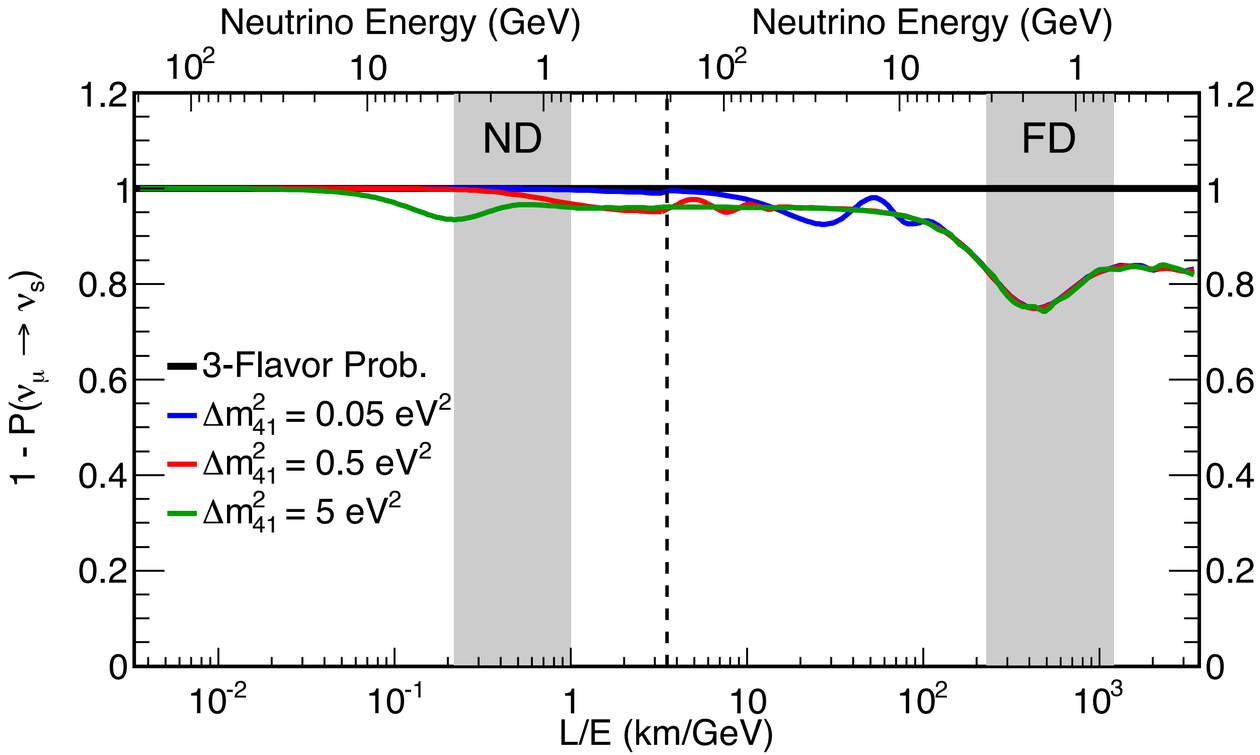
- Residual Network (ResNet):



It features special skip connections and a heavy use of batch normalization. The arch
is also missing fully connected layers at the end of the network.

arXiv:1512.03385 K. He. et.al

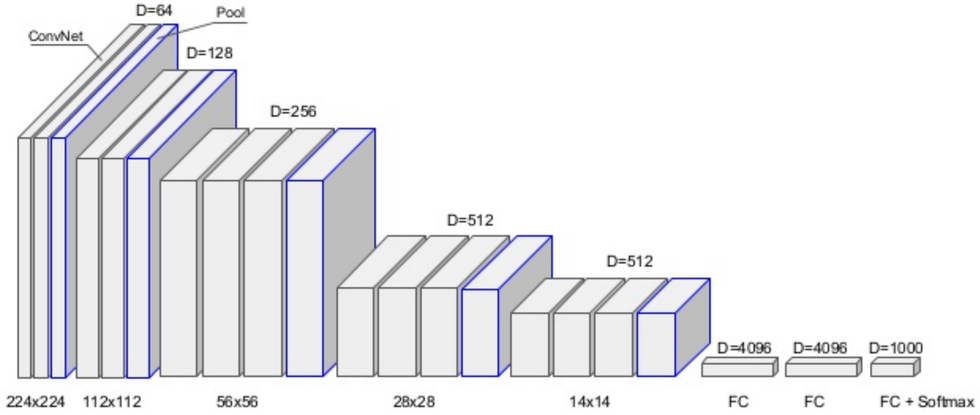
Looking for Sterile Neutrinos in NOvA



- An energy dependent reduction of the number of NC interactions in the FD with respect to the ND would be a clear evidence of oscillations into sterile neutrinos;
- Assuming Δm^2_{41} = 0.5 eV^2, so that we have no oscillations at the ND, but we have rapid oscillations at the FD.

Convolutional Neural Network Classification

- Visual Geometry Group (VGG):



Convolutional Neural Network Classification

- Visual Geometry Group (VGG):