The NOvA Deep Learning Program

A Summary of Applications, Successes and Lessons.

Fernanda Psihas

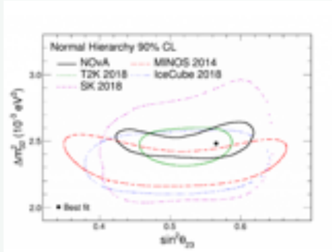
TEXAS
The University of Texas at Austin

ConTex

for the NOvA collaboration

The NOvA Experiment

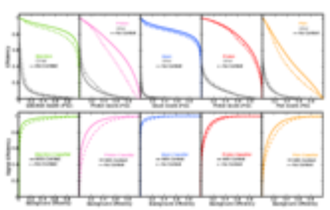
Technology: A neutrino beam, two functionally equivalent detectors.



First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOvA

FERMILAB-PUB-19-272-ND

June 2019

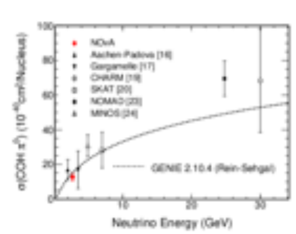


Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events

FERMILAB-PUB-19-258-PPD

June 2019

Particle detectors record the interactions of subatomic particles and their passage through matter. The identification of these particles is necessary for in-depth physics

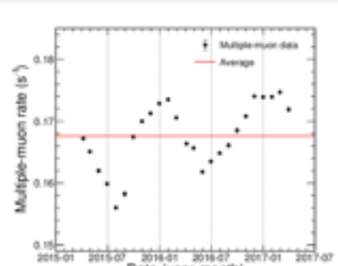


Measurement of Neutrino-Induced Neutral-Current Coherent π^0 Production in the NOvA Near Detector

FERMILAB-PUB-19-047-ND

February 2019

The cross section of neutrino-induced neutral-current coherent π^0 production on a carbon-dominated target is measured in the NOvA near detector. This measurement uses a narrow-band neutrino beam with the average neutrino energy of 2.7 GeV, which is of

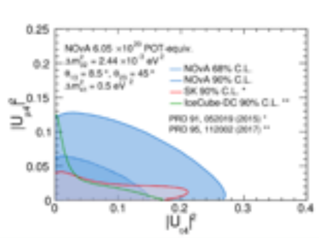


Observation of seasonal variation of atmospheric multiple-muon events in the NOvA Near Detector

Phys. Rev. D 99, 122004

April 2019

Using two years of data from the NOvA Near Detector at Fermilab, we report a seasonal variation of cosmic ray induced multiple-muon ($N_\mu \geq 2$) event rates which has an opposite phase to the seasonal variation in the atmospheric temperature. The strength of



Search for active-sterile neutrino mixing using neutral-current interactions in NOvA

FERMILAB-PUB-17-198-ND

June 2017

We report results from the first search for sterile neutrinos mixing with active neutrinos through a reduction in the rate of neutral-current interactions over a baseline of 810 km

Physics:

Neutrino oscillation,

Sterile neutrinos,

Cross Sections,

Cosmic Ray Physics,

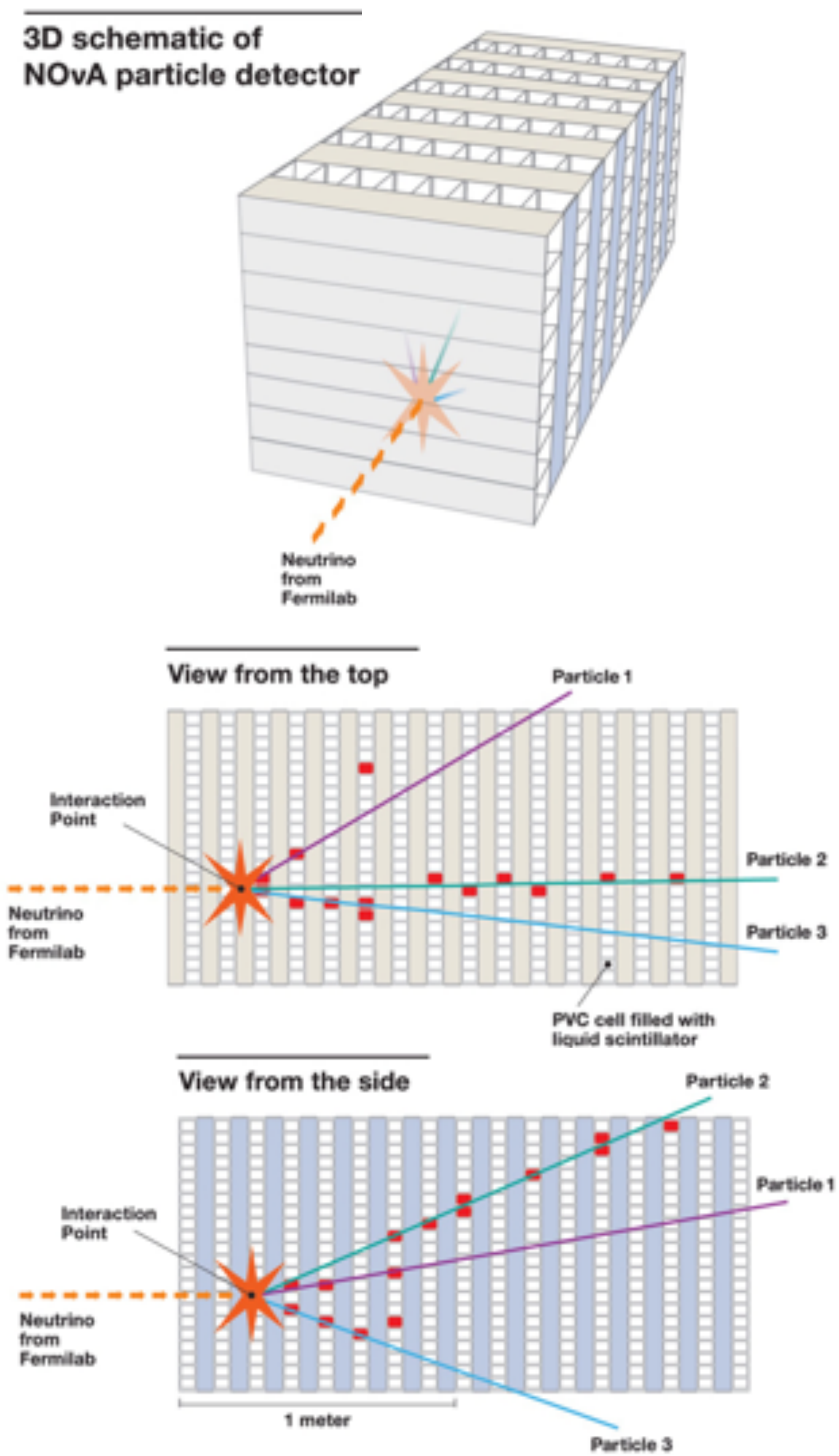
Supernova Physics,

NSI Searches,

Exotics searches

https://novaexperiment.fnal.gov

Measurables are Flavor and Energy



For oscillations we want to measure the incoming neutrino flavor and energy.

$$\nu_e + N \rightarrow e + X$$

Outgoing lepton matches incoming ν flavor

$$\nu_\mu + n \rightarrow \mu + p$$

Final state particles

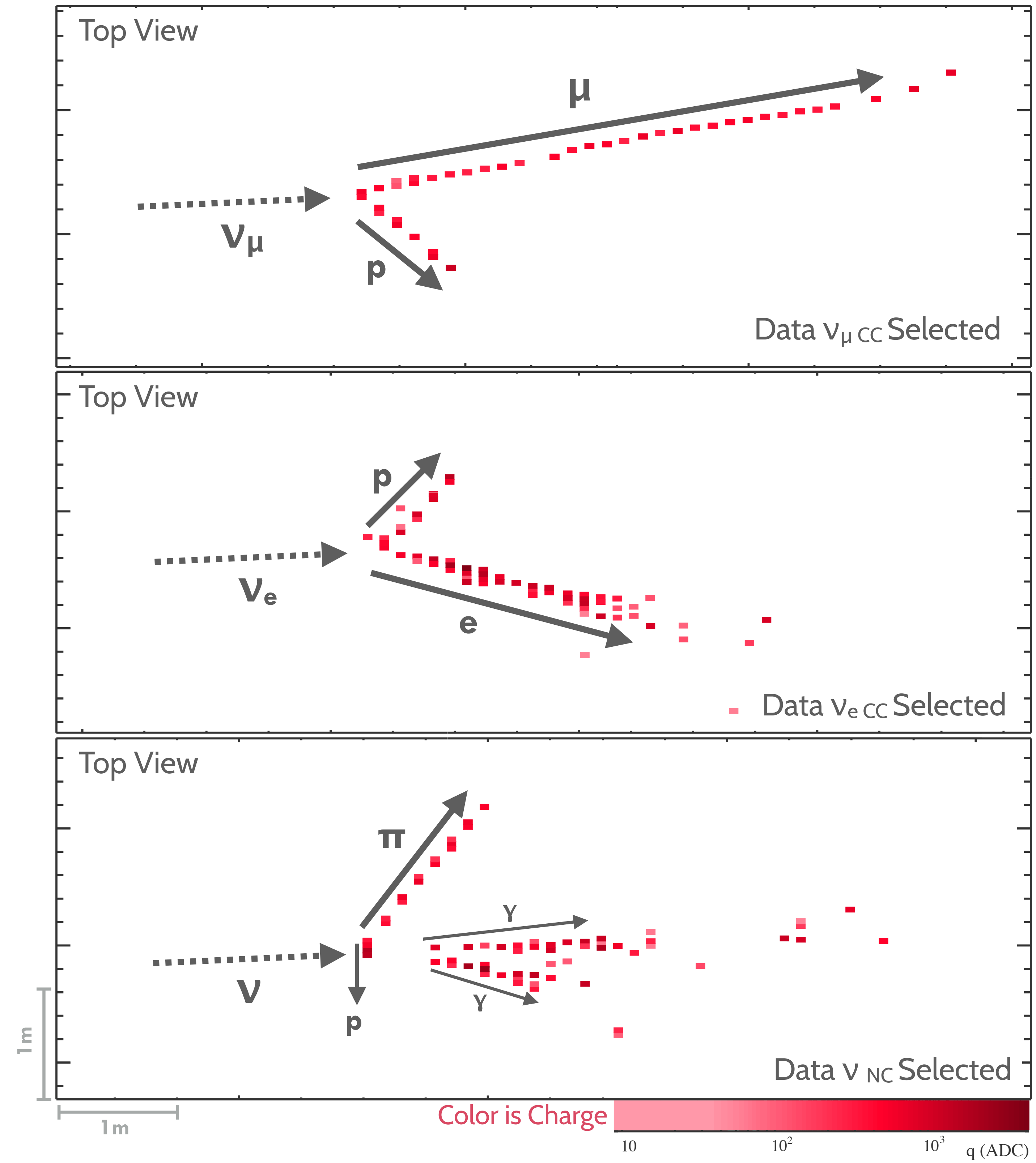
For interaction cross-sections we want to measure the incoming neutrino flavor and energy, as well as the final state.

NOvA event topology

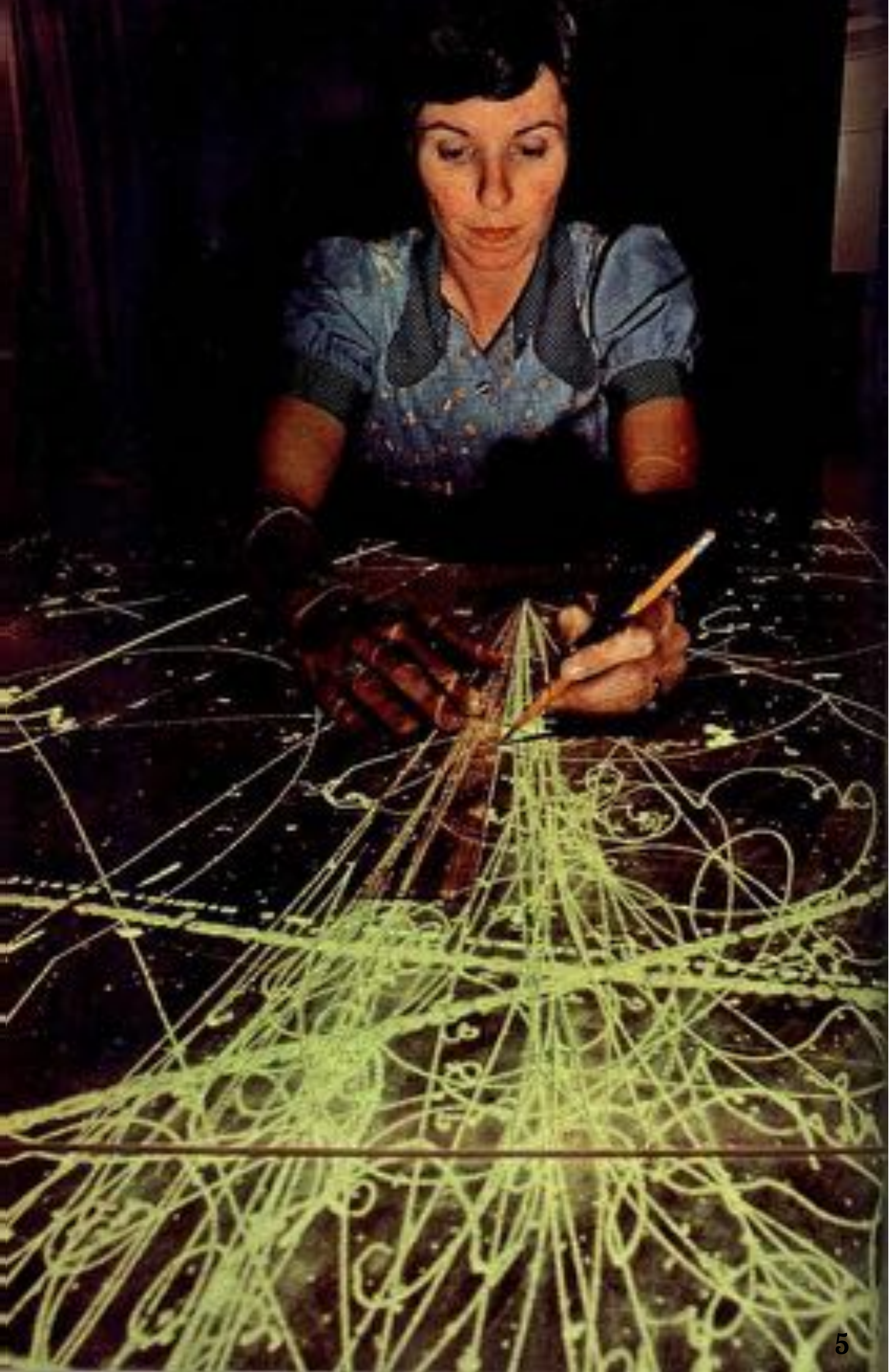
Long tracks are typically muons (minimally ionizing).

e and γ produce electromagnetic showers spanning multiple detector cells.

Heavier particles typically shorter, higher dE/dx tracks.

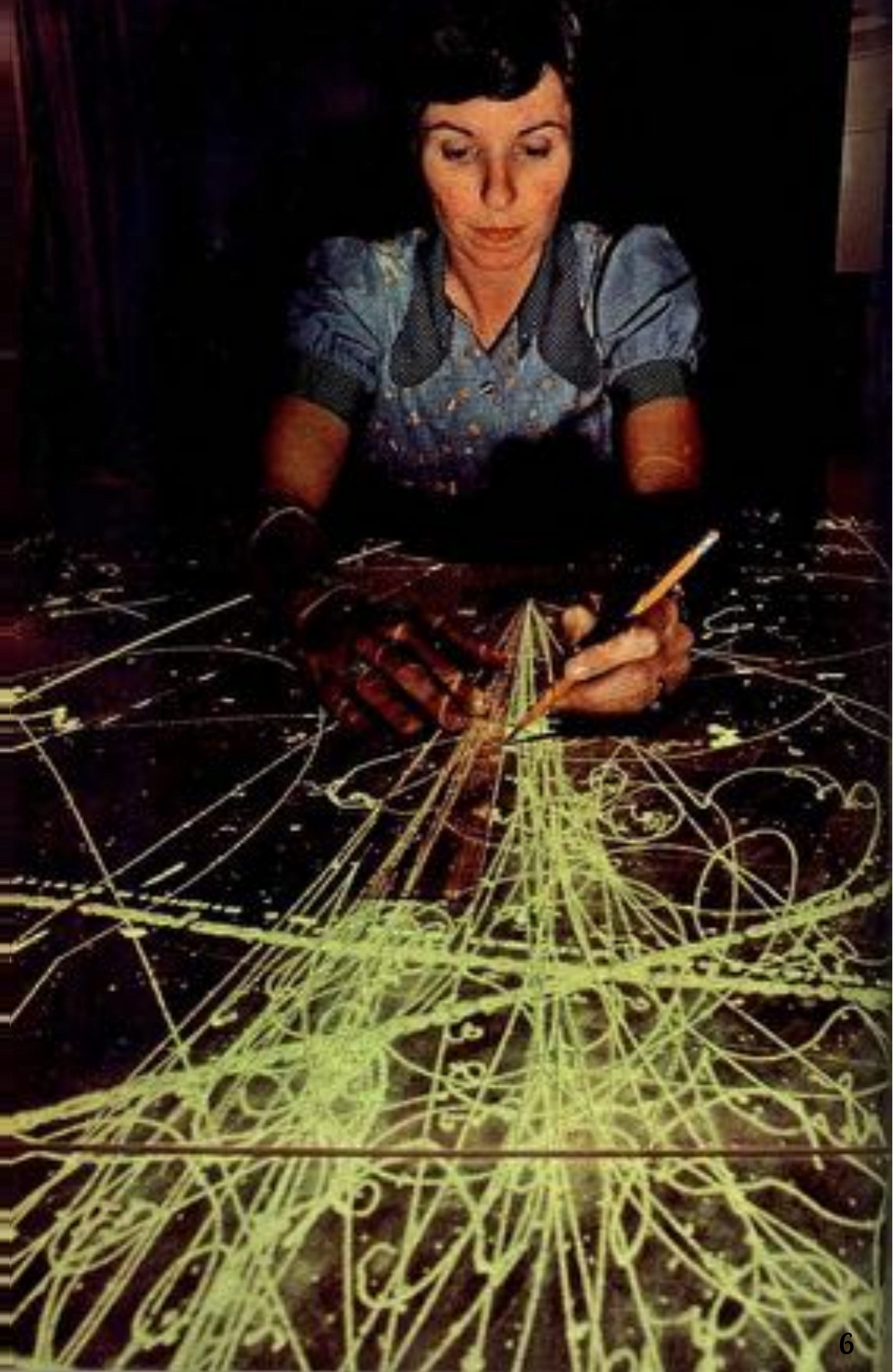


Why deep learning?

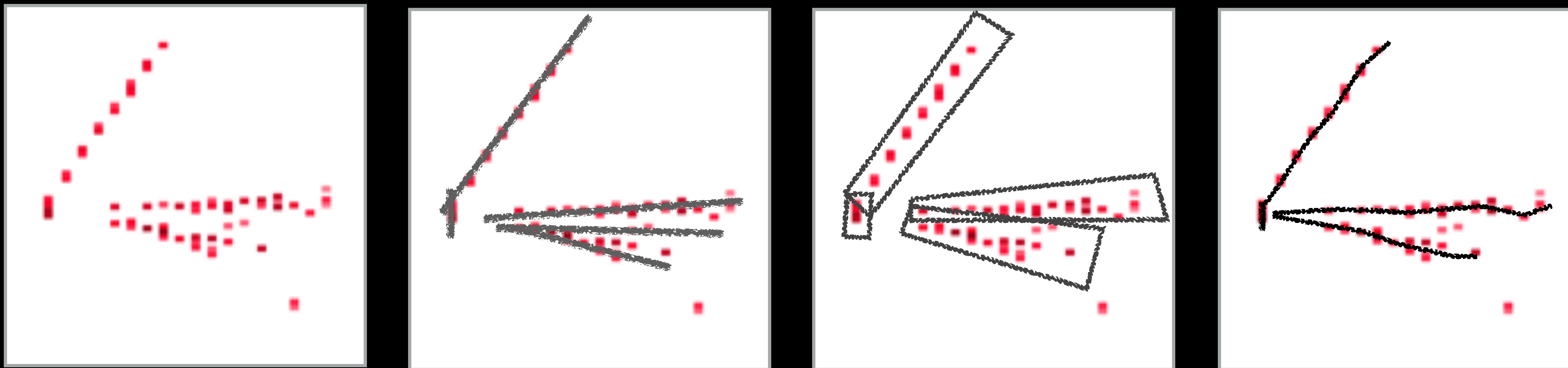


Why deep learning?

1. Data rates
2. Functionality

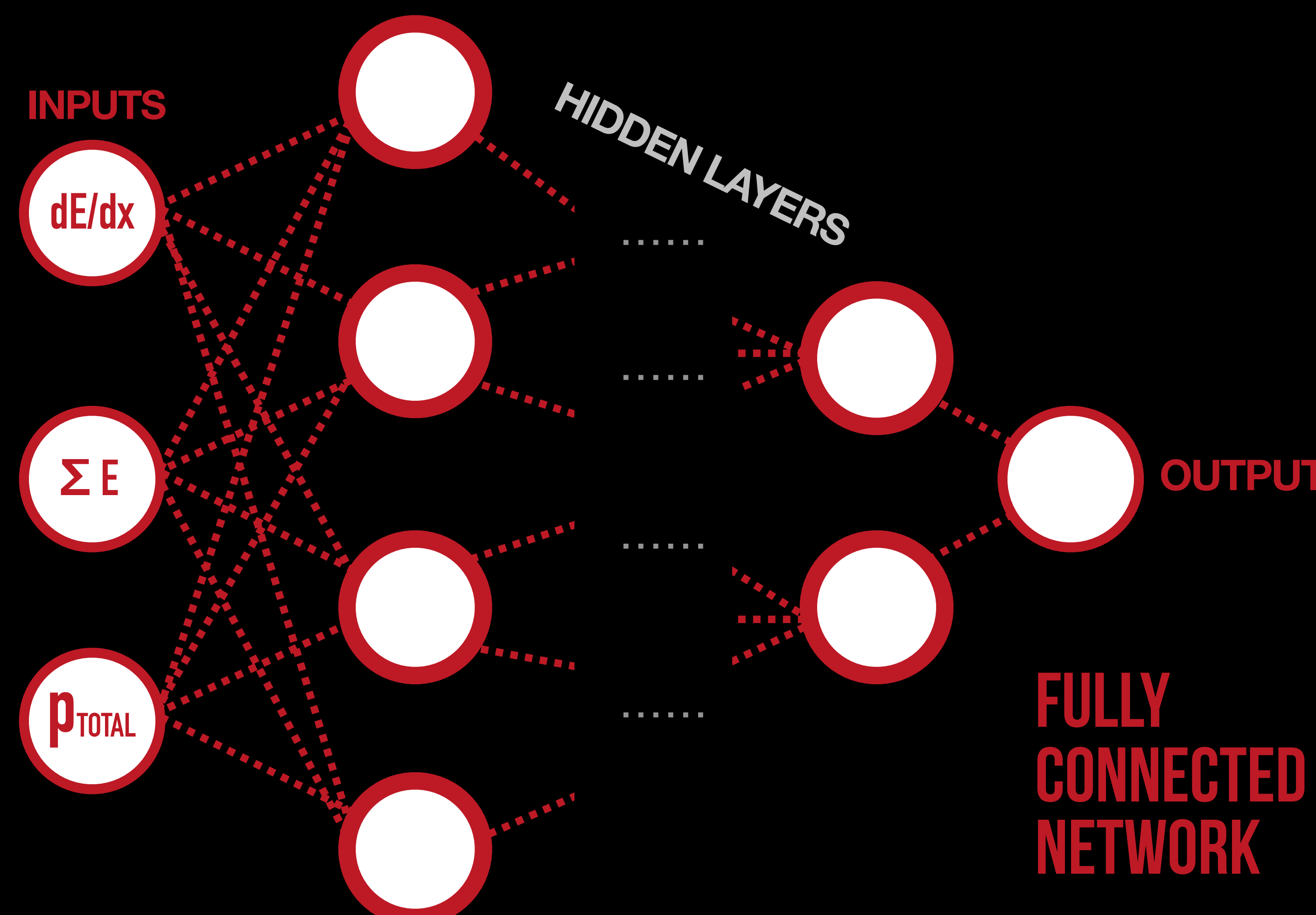


Geometrical Reconstruction

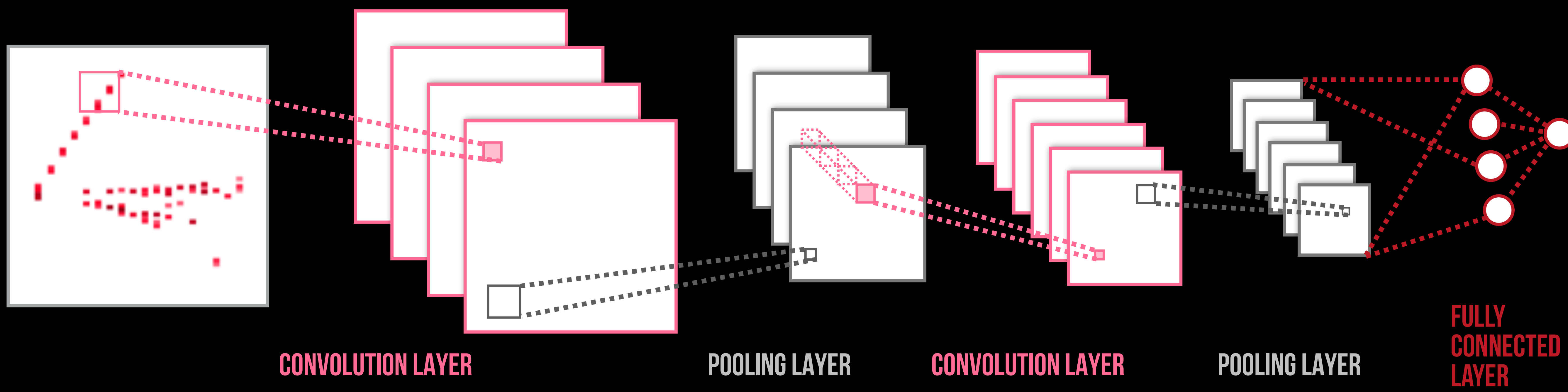


Extract information (features) from the event which can separate signal vs background.

Train NNs with these features from MC libraries



Convolutional Neural Networks

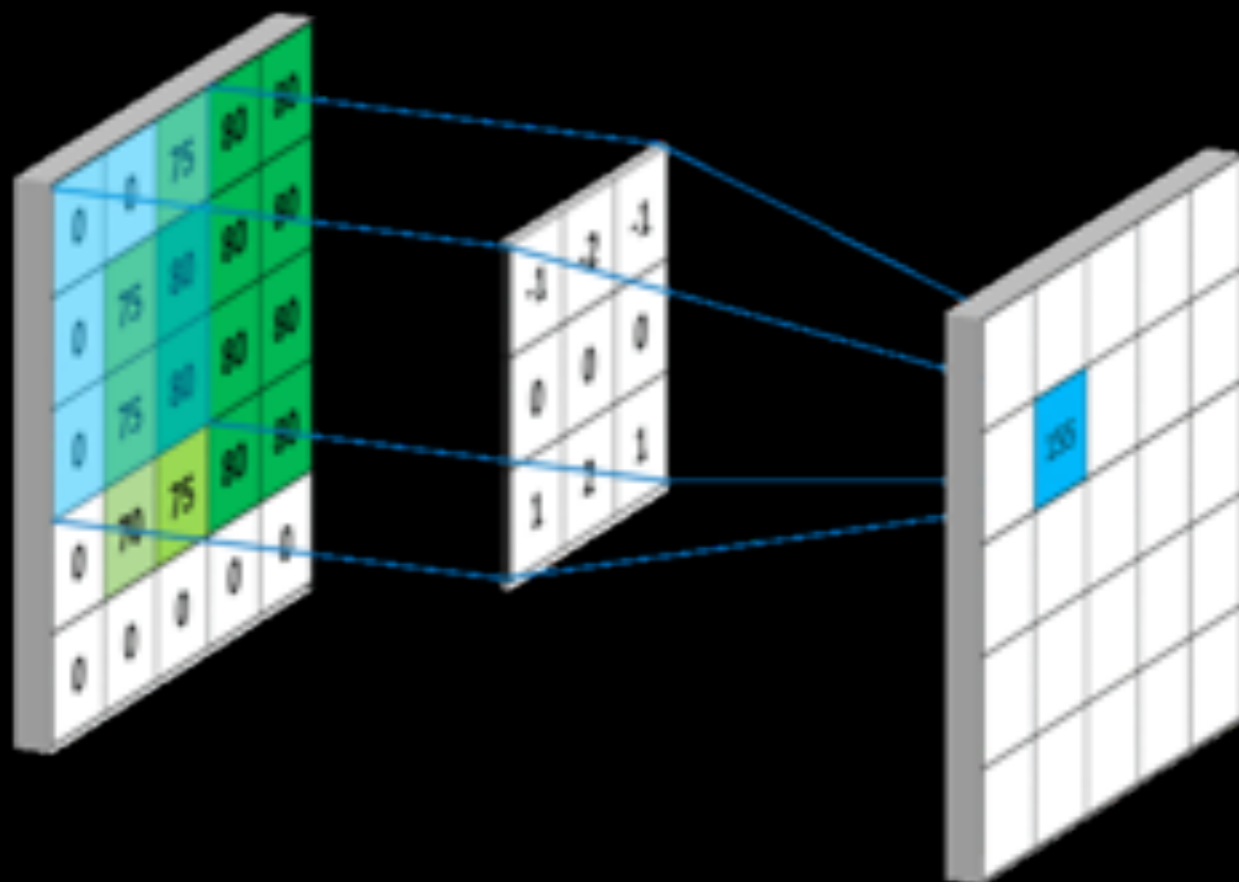


Convolution and **pooling** layers extract features and reduce the map dimensions.

Architectures vary by application, but the **basic structure** has convolutions, pooling and ends in a fully connected layer

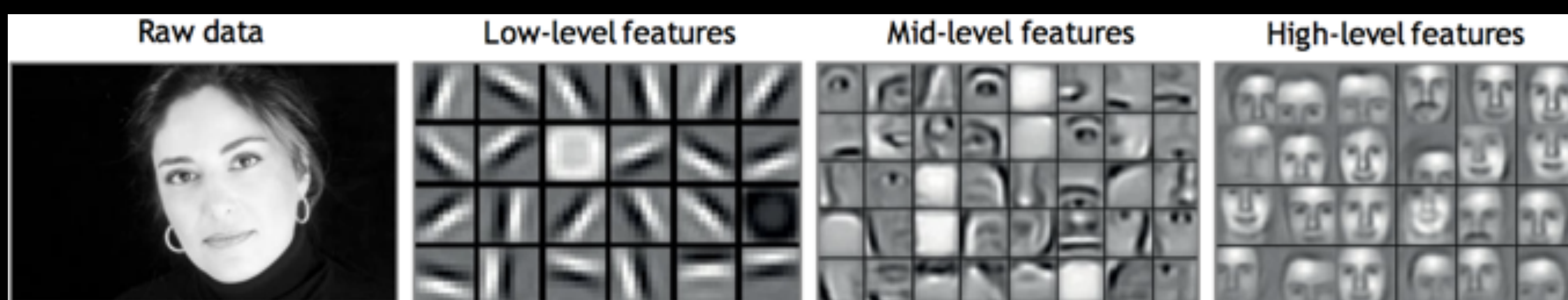
Feature extraction + Classification

Convolutions for feature extraction



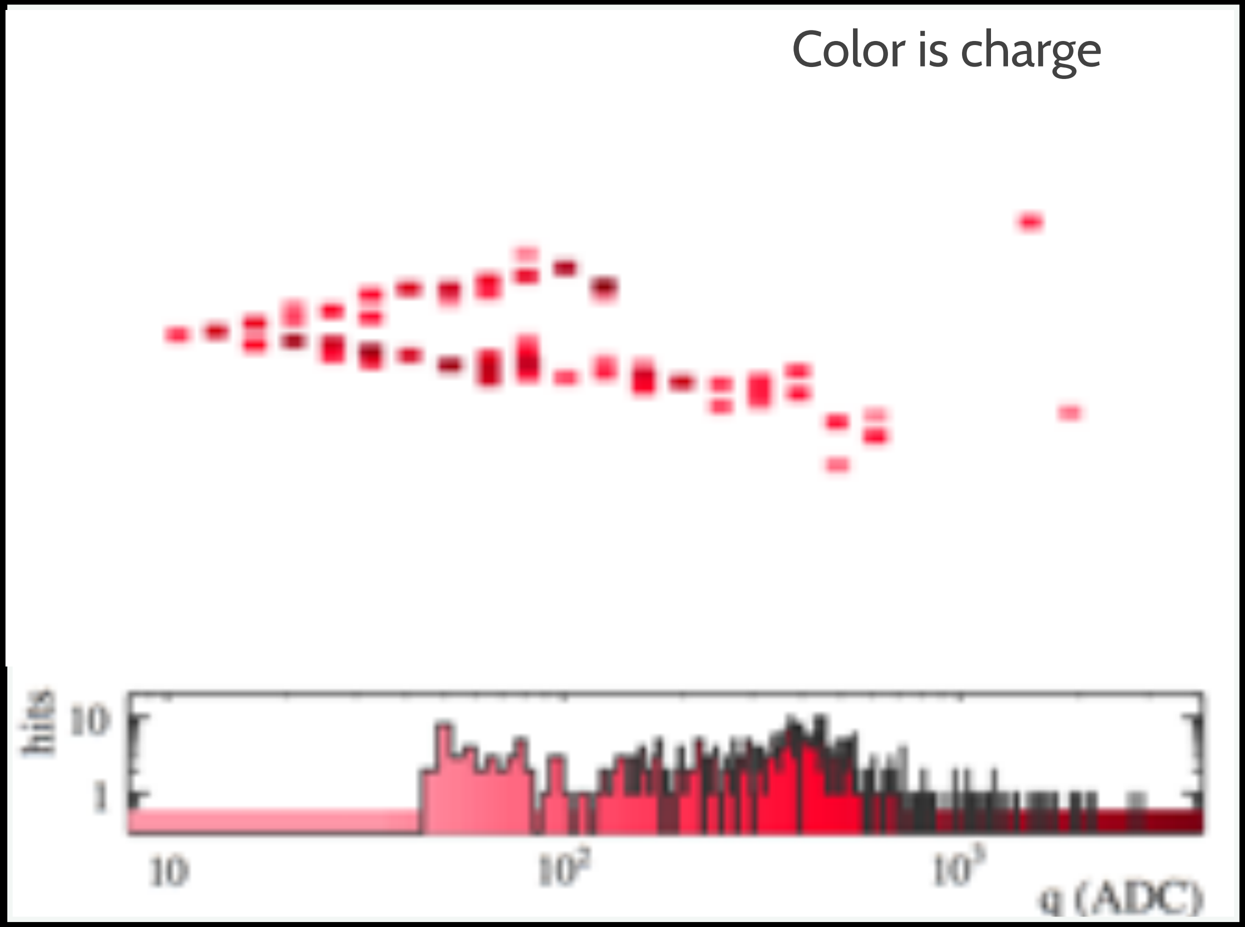
Convolutions operate kernels over a matrix/map to extract features.

The size of the kernel determines the level of the features which can be extracted.



Deep Learning for Identification

Premise: Rather than select a set of features a priori, let a deep learning network extract **learned features**.

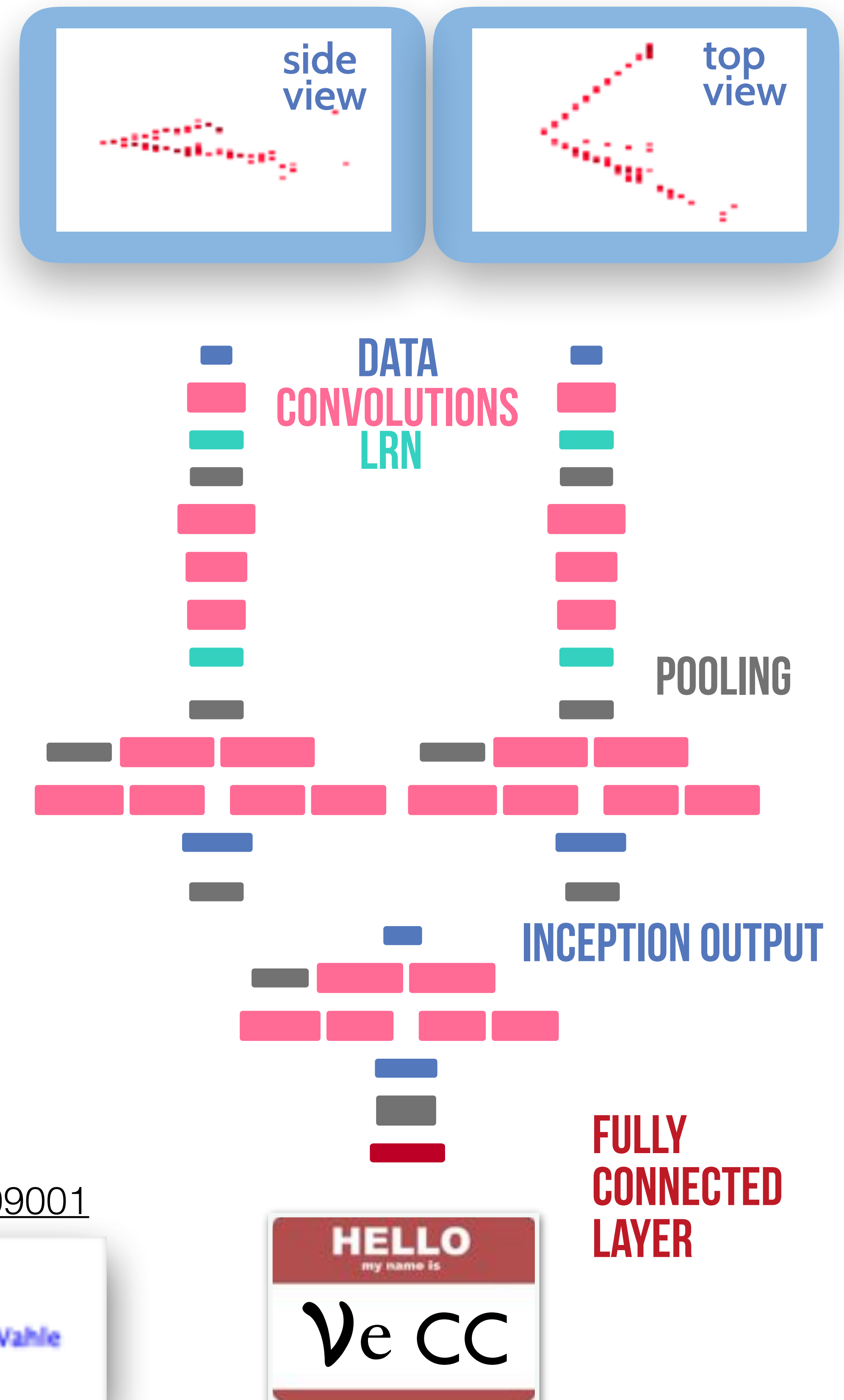


In practice, using convolutional neural networks CNNs also **decouples from traditional reconstruction**, reducing the inefficiencies and bias upstream.

NOvA event input are fixed-size **2D-view** pixel maps of the **calibrated hits**.

Event Identification

*NOvA's convolutional visual network CVN architecture is a "siamese" CNN



Neutrino ID is done mostly with our **CVN or convolutional visual network**

CVN is a two-tower convolutional neural network which **learns from the top and side detector views** independently first.

Trained on **cosmic ray data events** and **neutrino MC with overlaid cosmic data.**

Published. 2016

JINST 11 (2016) no.09, P09001

A Convolutional Neural Network Neutrino Event Classifier

A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, P. Vahle

(Submitted on 5 Apr 2016 (v1), last revised 12 Aug 2016 (this version, v3))

Event Identification

NOvA had the **first implementation** of Convolutional Neural Networks on a particle physics result.

It **increased our effective exposure by 30%** compared to traditional ID methods.

Can you beat CVN?

http://nusoft.fnal.gov/nova/public/nova-events/

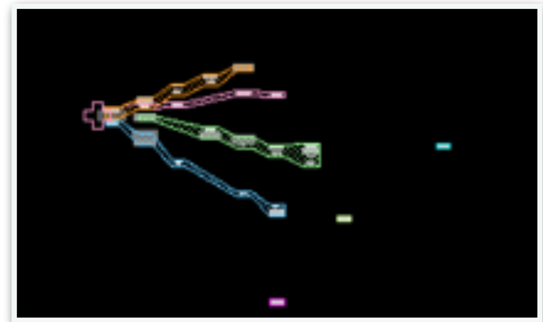
NEW for the anti-neutrino analysis:
Improved simulation.
Training is final-state based.
Network optimizations*
Beam-mode training

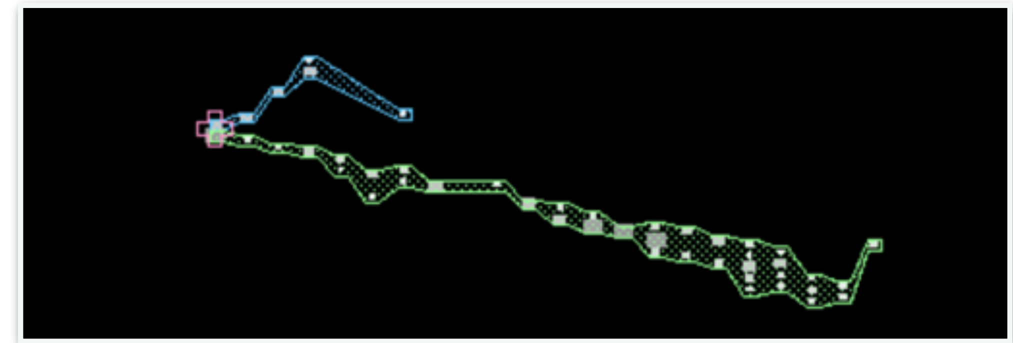
New beyond 2019:
Moving from caffe to TensorFlow
NOvA is now using HDF5 files + Python based analysis infrastructure

*F. Psihas, Ph.D. thesis, Indiana University, 2018,doi:10.2172/1437288.

Particle Identification with CNNs

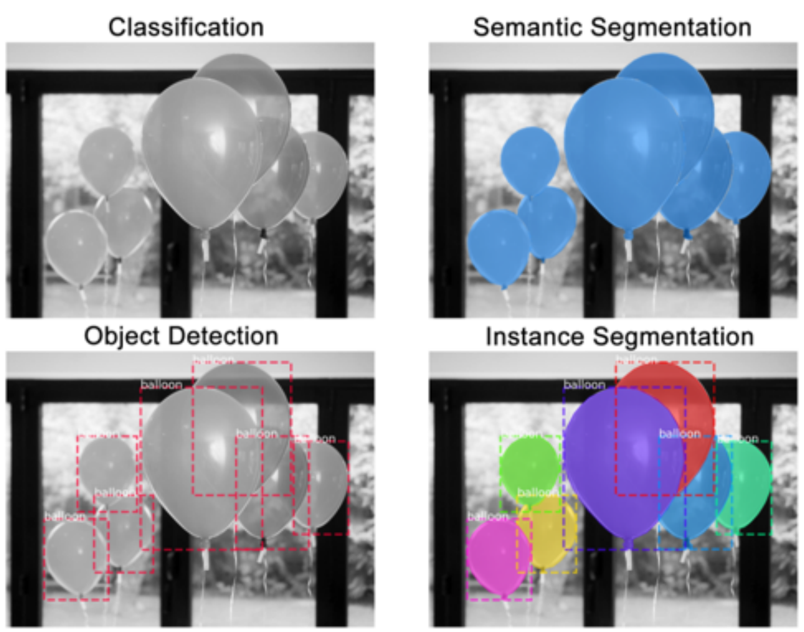
We use our Deep CNN classifier to identify each cluster.





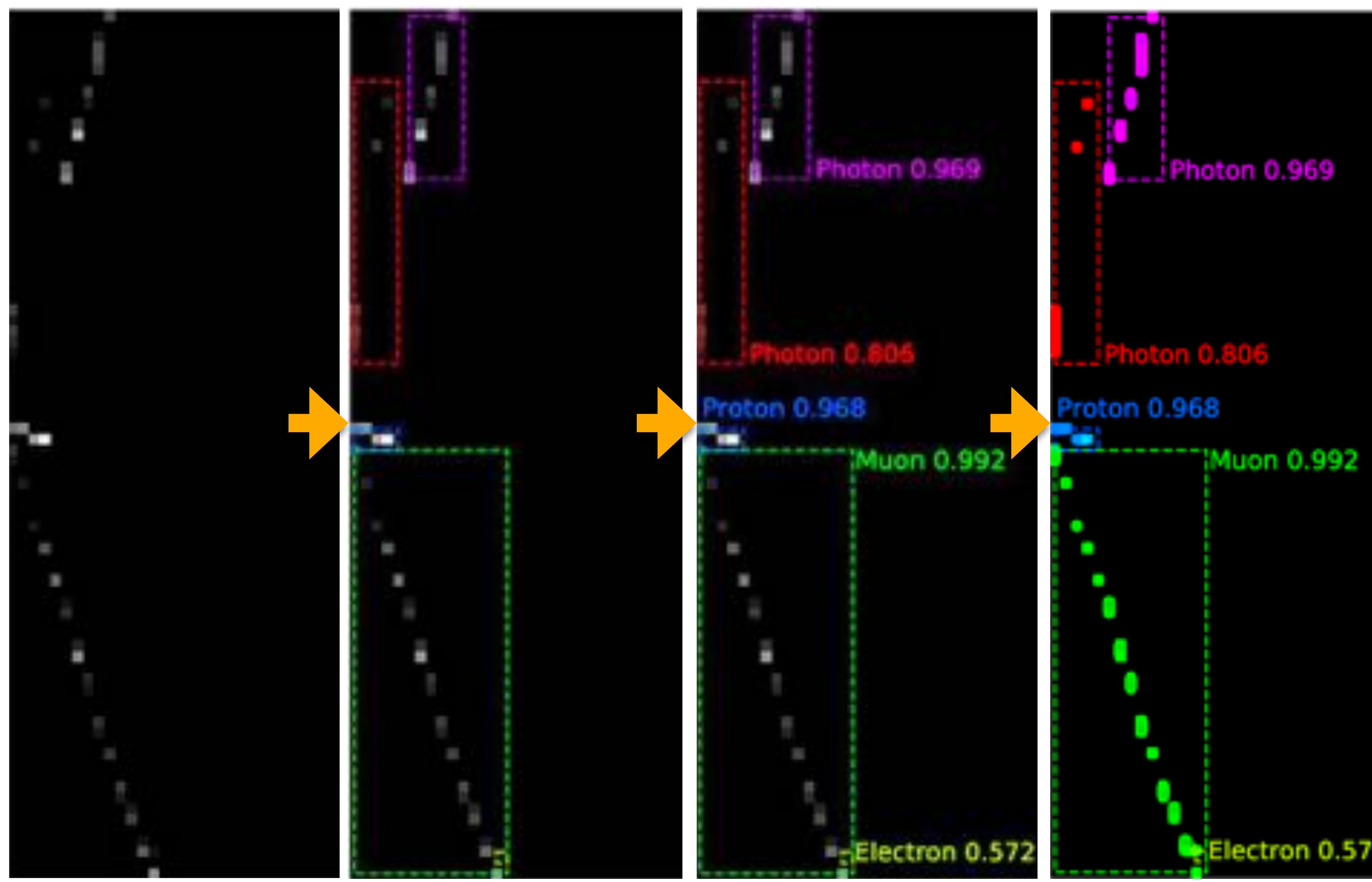
Localizing + Clustering + Identification

Under development

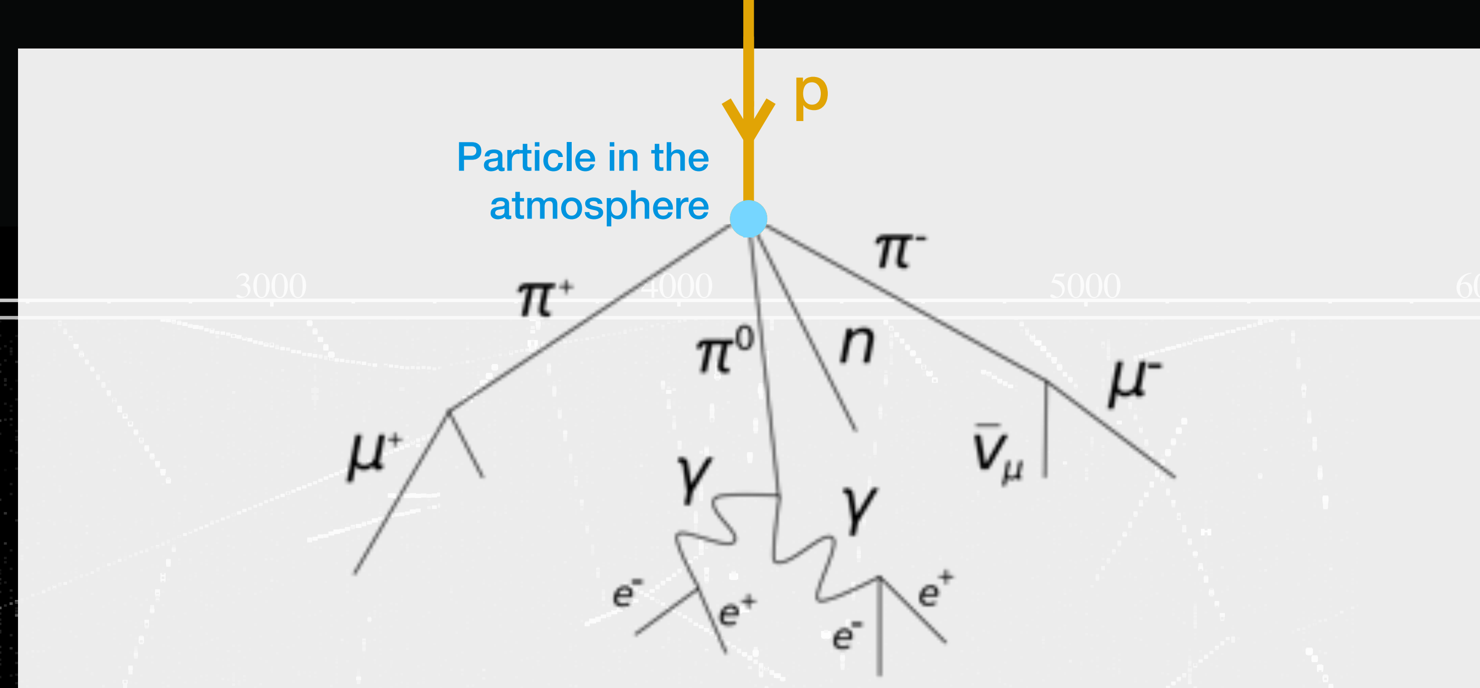


MASK-RCNN

We use an adaptation of MASK-RCNN to cluster and identify all activity.



Reducing Cosmic rays



Cosmic rays make up an important background to our oscillation signal.

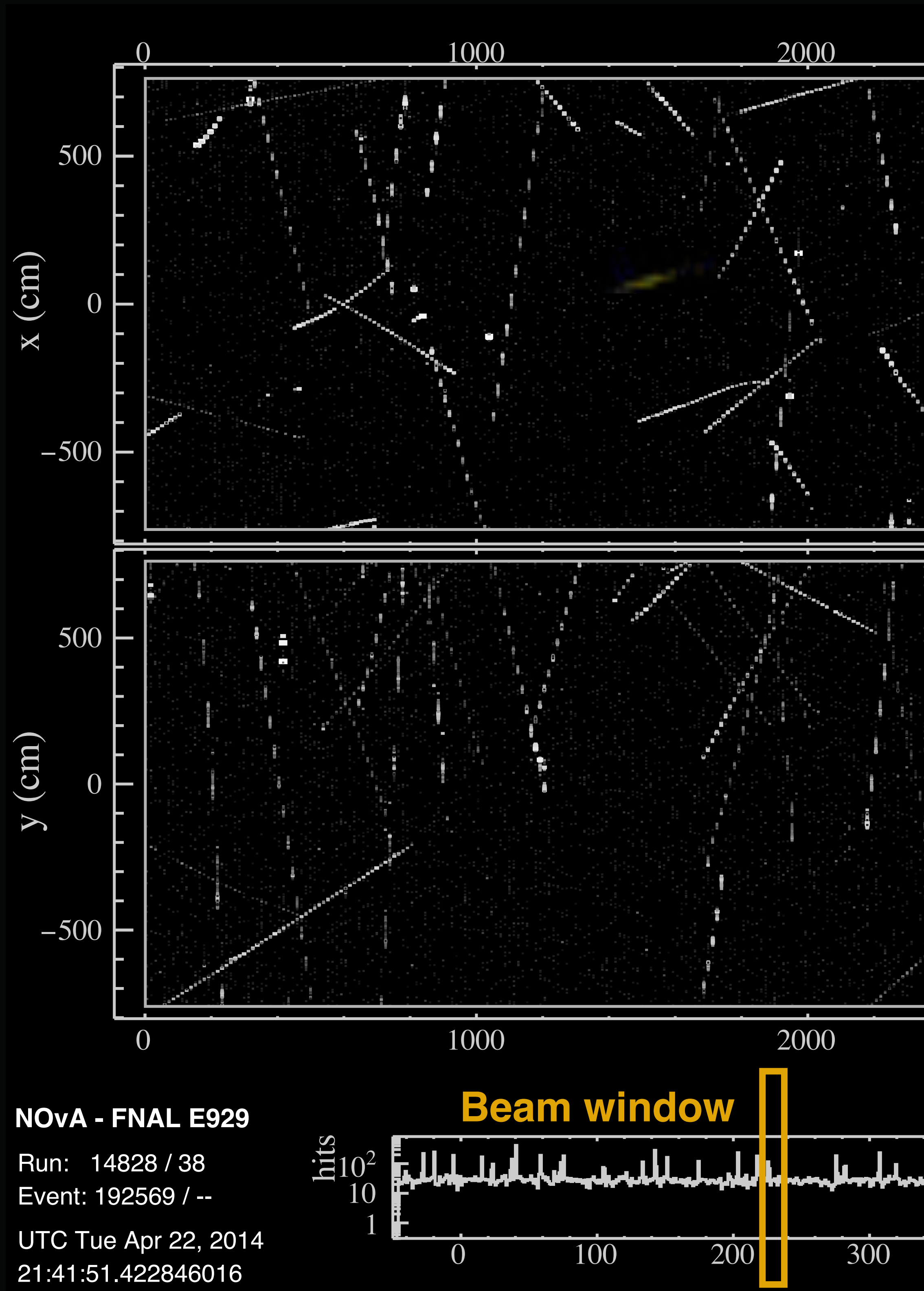
We use cosmic ray data to predict our backgrounds.
1 **ν** background per ~10^7 cosmic rays

NOvA is exploring a CNN software trigger to reduce **cosmic ray background reduction**

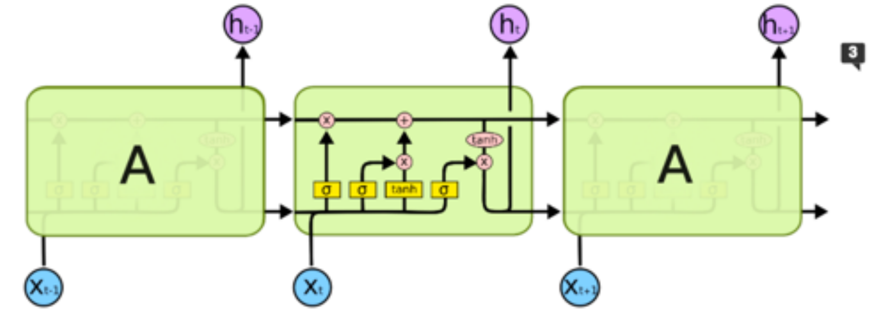
INPUTS: X and Y views of the whole detector

Using TensorFlow and HDF5s.
Evaluated on GPUs.

The target is to reject full detector readouts.



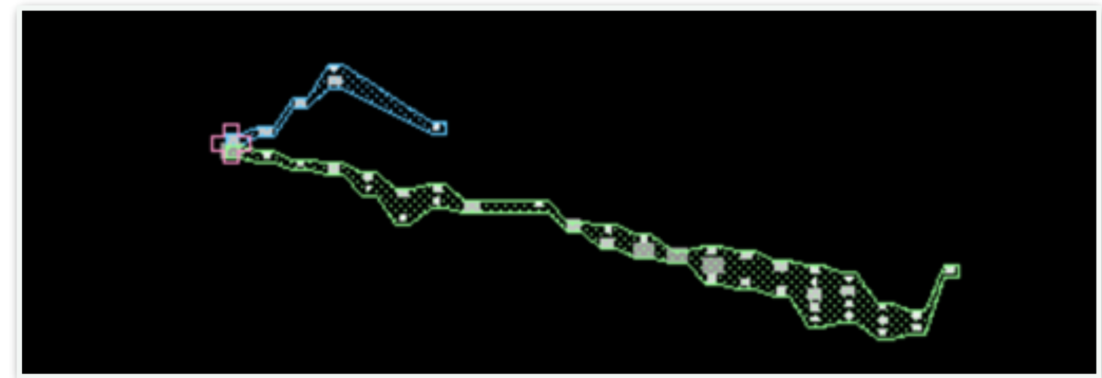
Energy Reconstruction with LSTMs



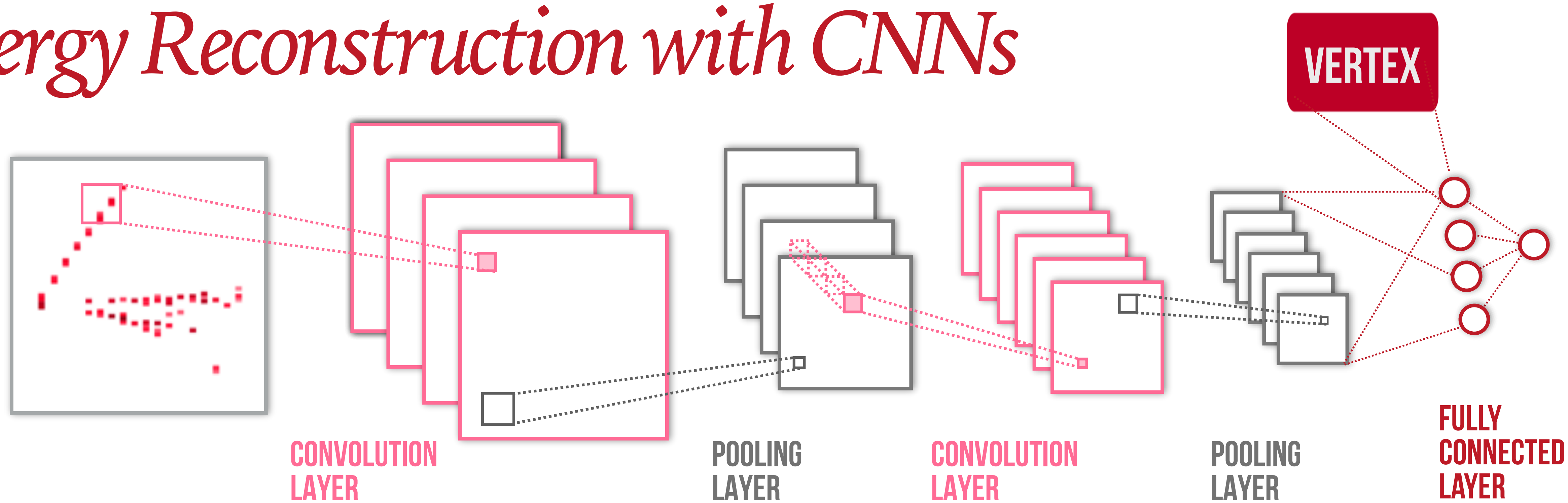
Recurrent Neural Networks:

Sequential network using the current state of the system + the output from last iteration.

Long + **S**hort **T**erm **M**emory Nets add a long term memory cell to RNNs



Energy Reconstruction with CNNs



The target is Energy instead of PID values.

Published. 2019

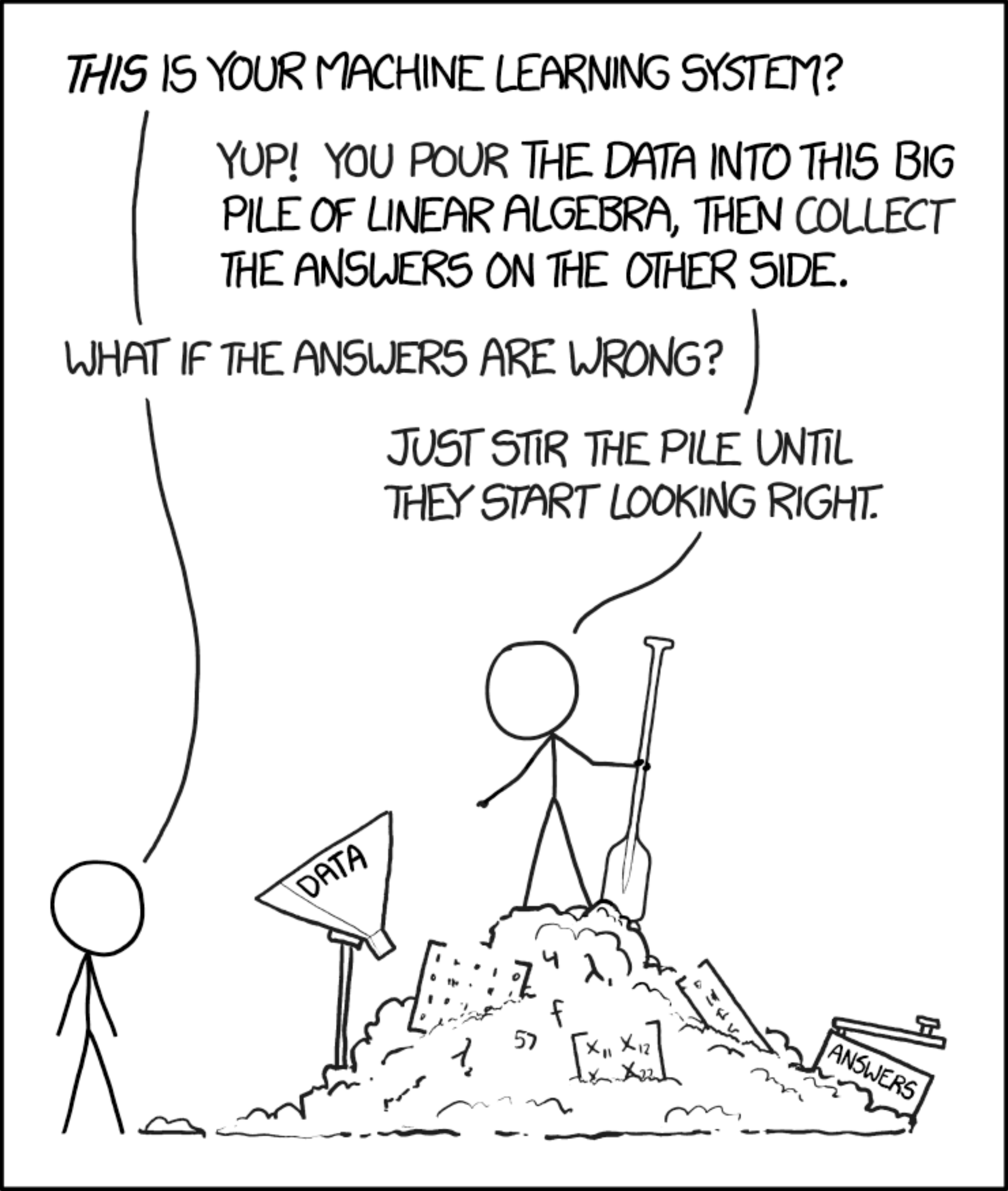
Phys.Rev. D99 (2019) no.1, 012011

Improved Energy Reconstruction in NOvA with Regression Convolutional Neural Networks

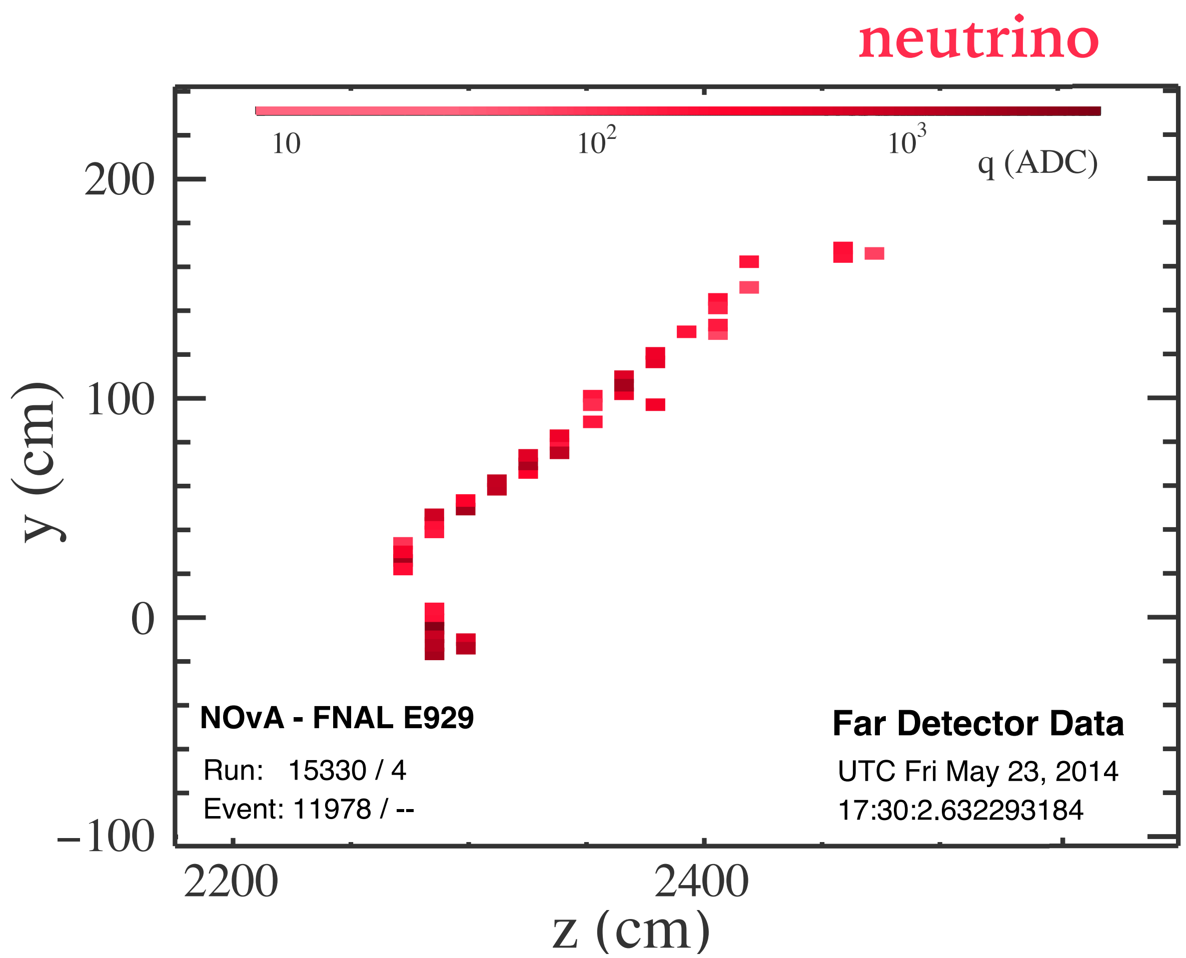
Pierre Baldi, Jianming Bian, Lars Hertel, Lingge Li

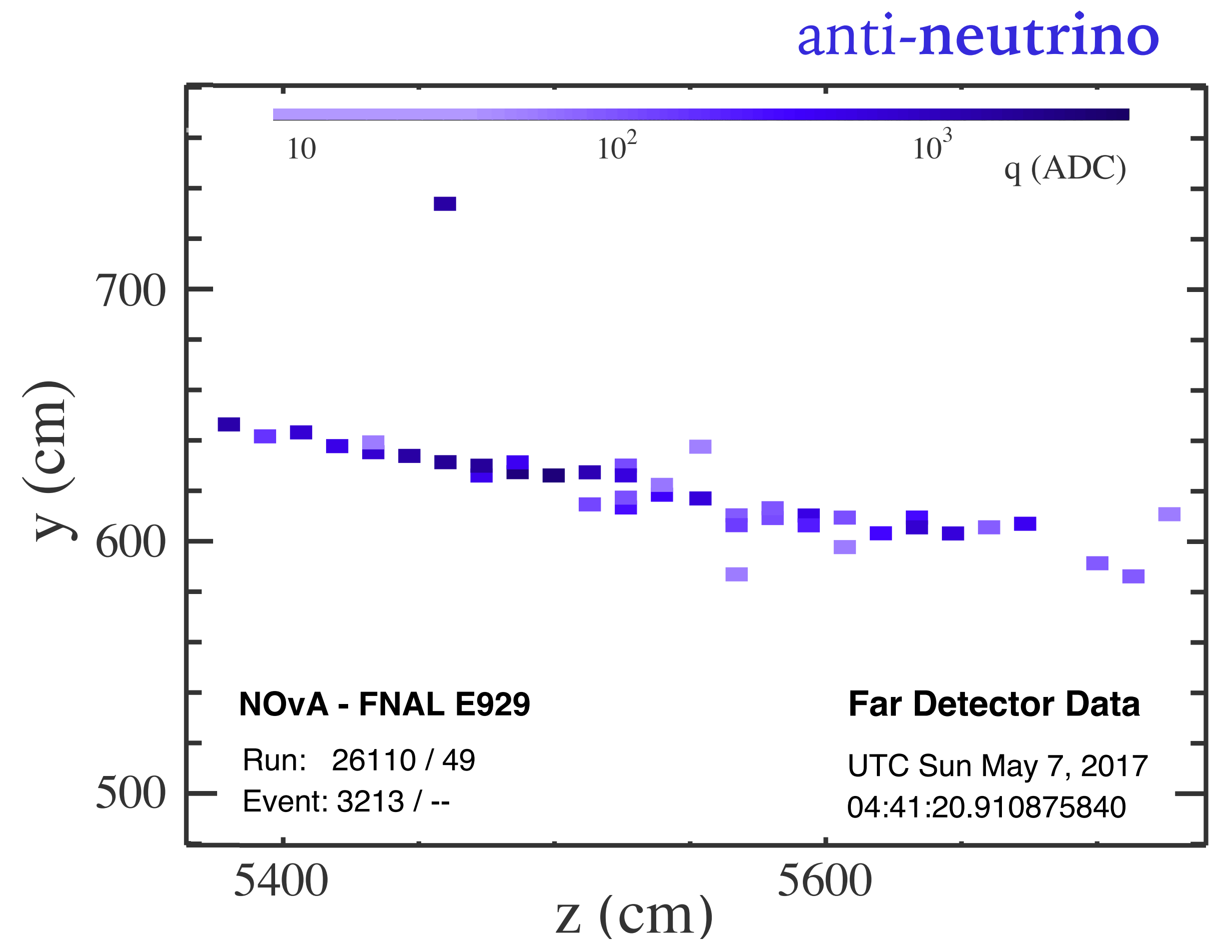
Lessons:

Apply an understanding of your data…



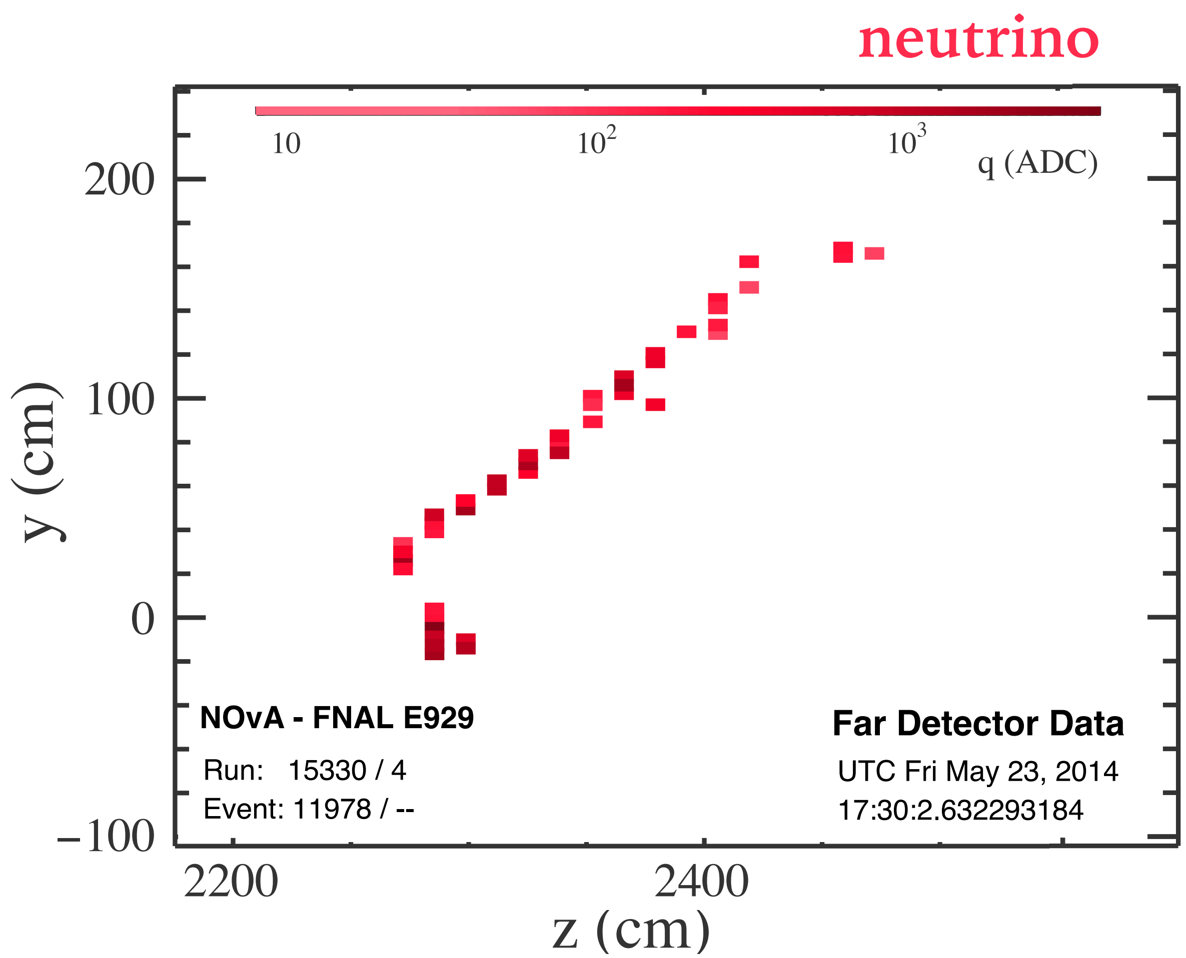
Event Identification

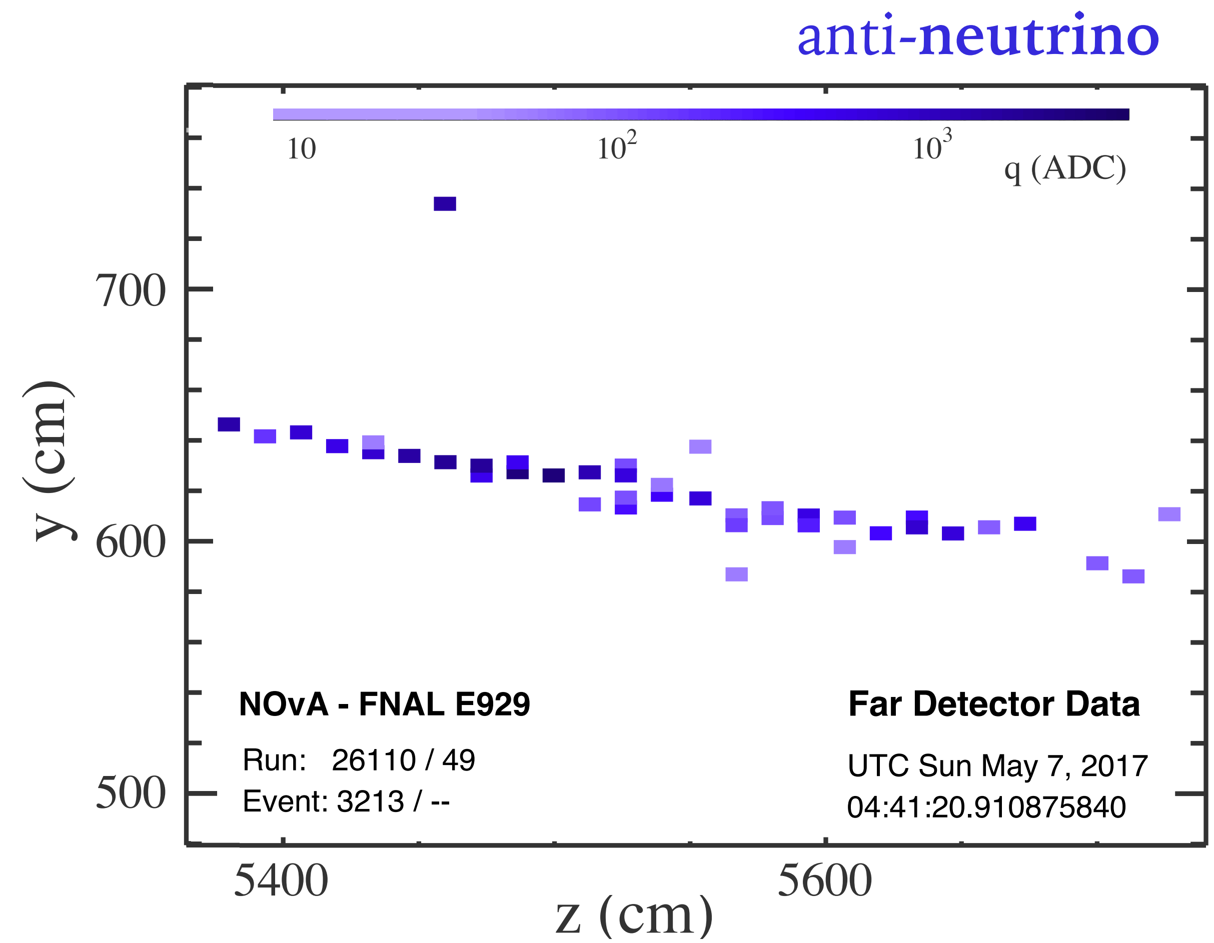




The topology of **neutrino** and **anti-neutrino** interactions is different on average.

Event Identification



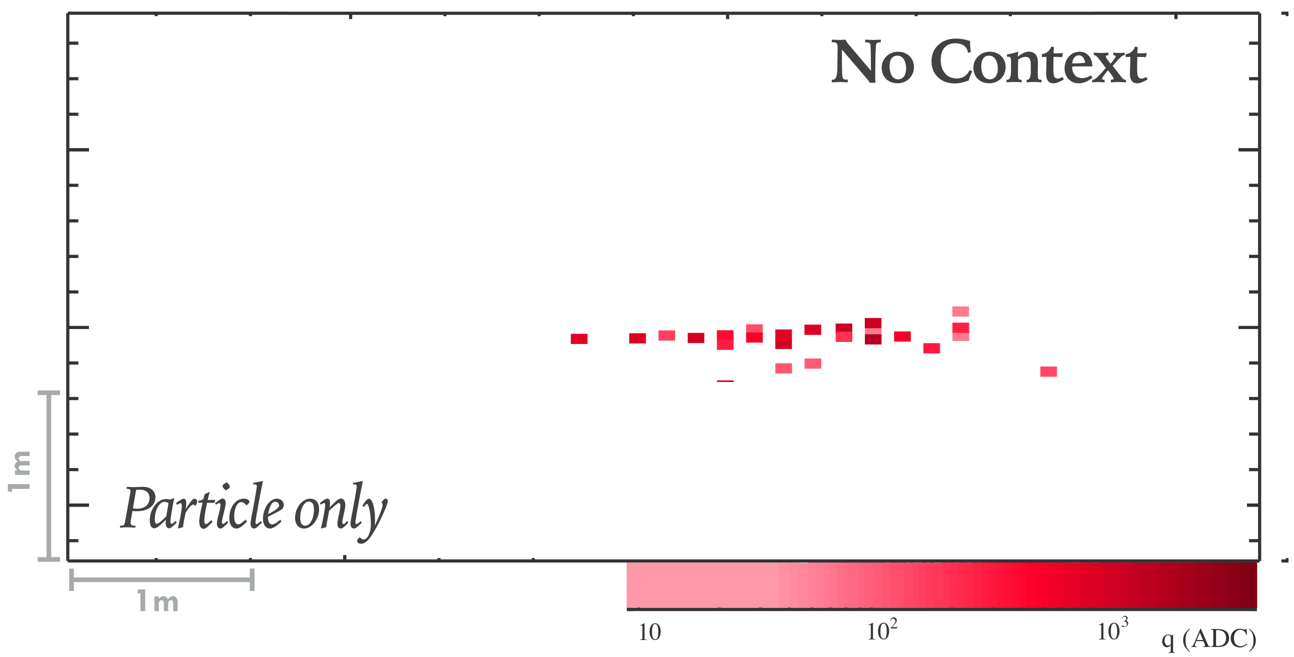


Train on **neutrino** beam and **anti-neutrino** beam simulations separately.

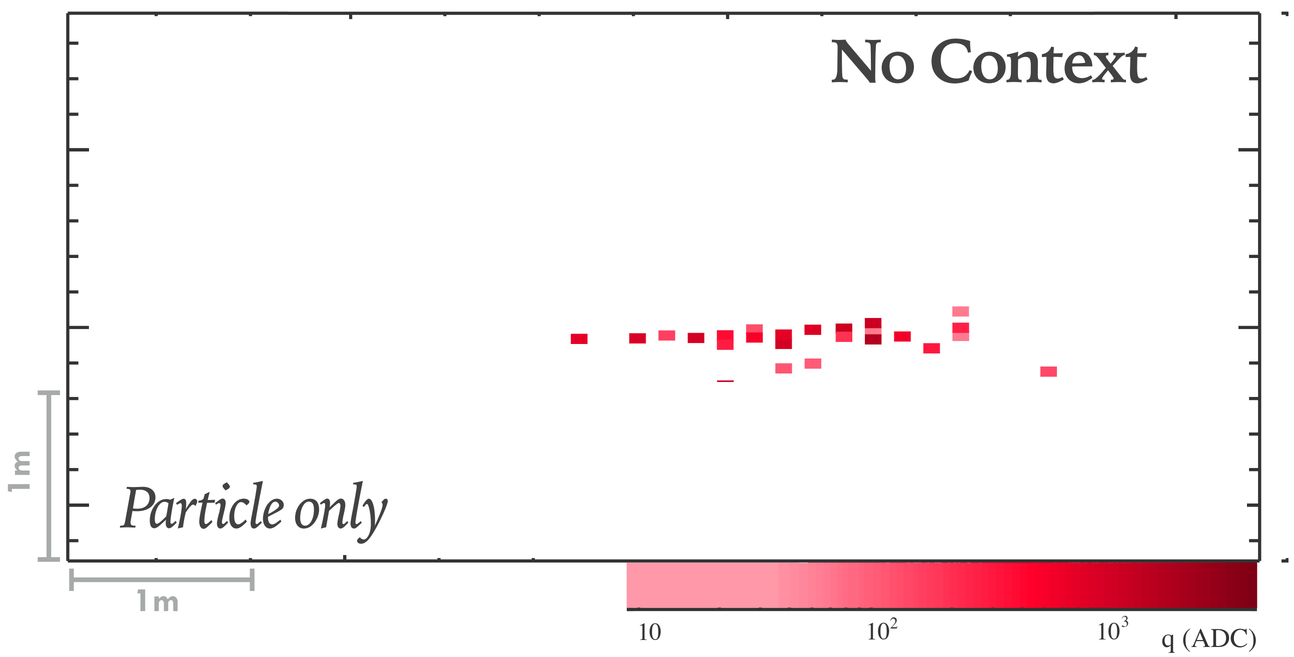
Utilize differences in event topology.

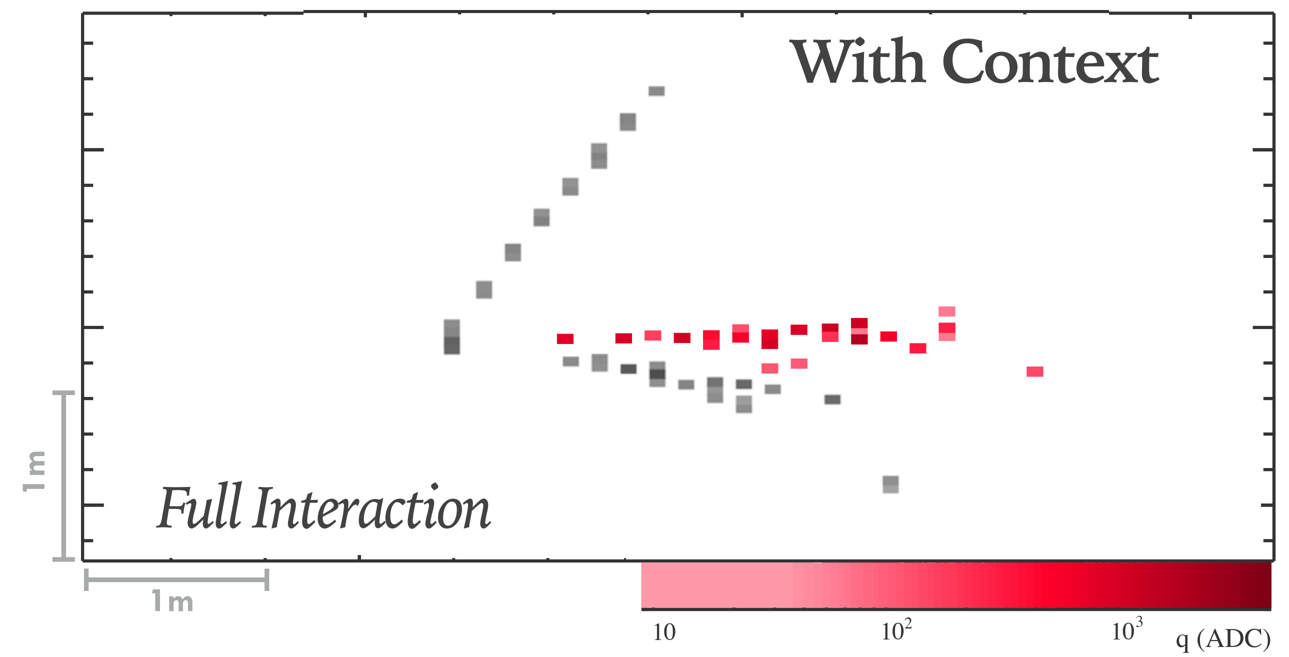
$\bar{\nu}$ **Efficiency Improvement** Training Sample (ID > 0.9)		
$\bar{\nu}_e$ CC Signal	$\bar{\nu}_\mu$ CC Signal	$\bar{\nu}$ NC Signal
14%	6%	10%

Context-aided classification



Context-aided classification





The task of classification can be aided by **providing context.**

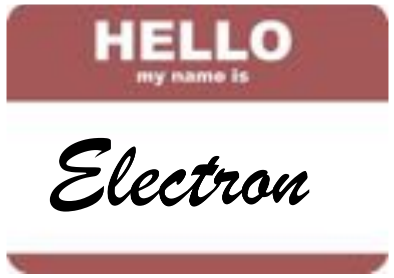
Single Particle Tagging

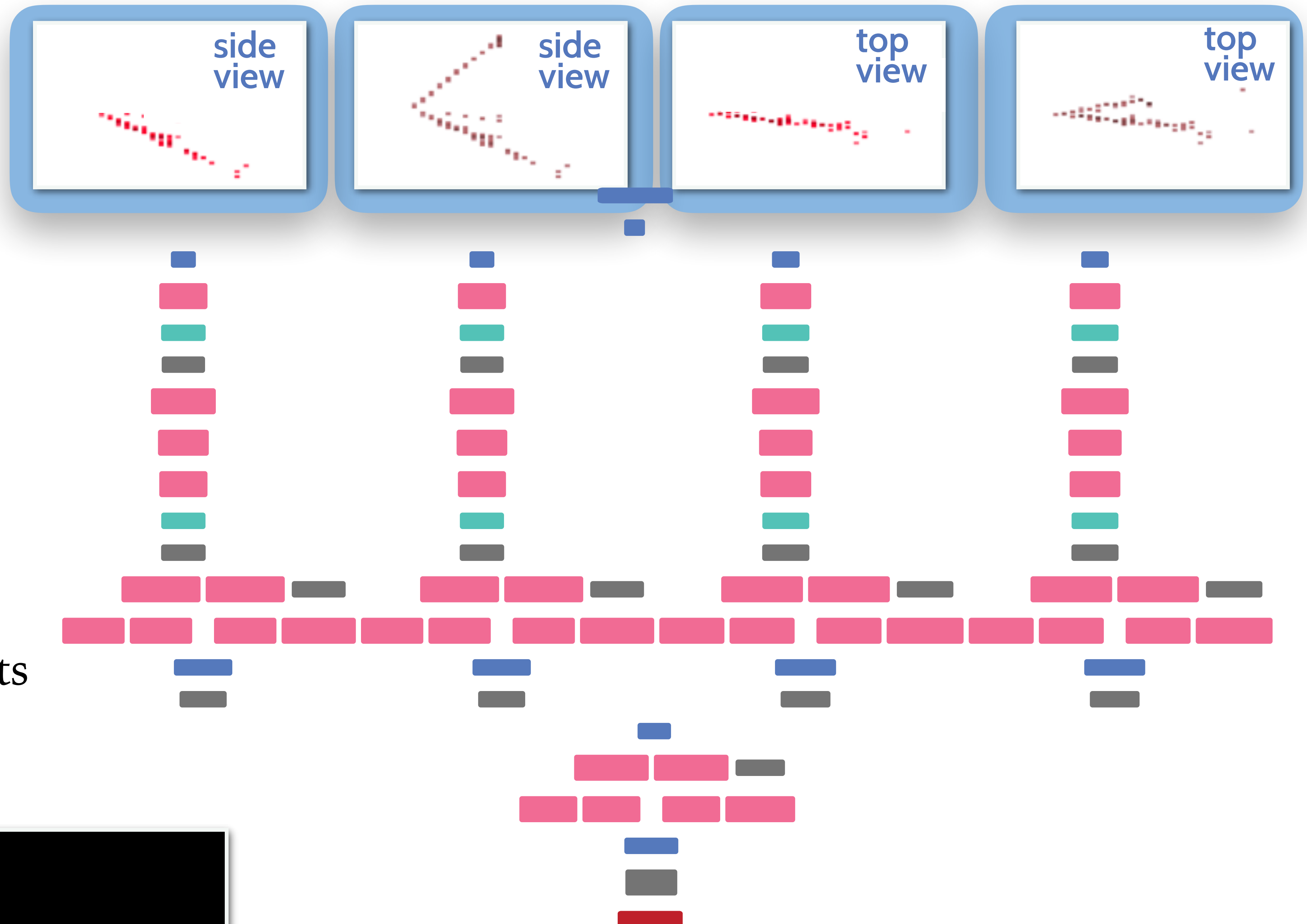
Soon to be on PRD

Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events

F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic, A. Sousa

(Submitted on 3 Jun 2019)

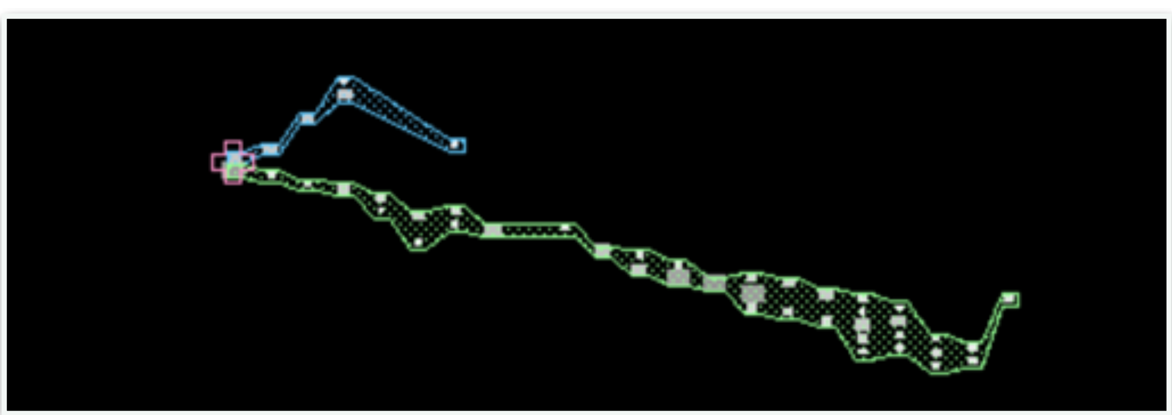




Using the existing reconstruction, classify clusters of hits.

Modified to take 4 views (event + clustered particle hits)

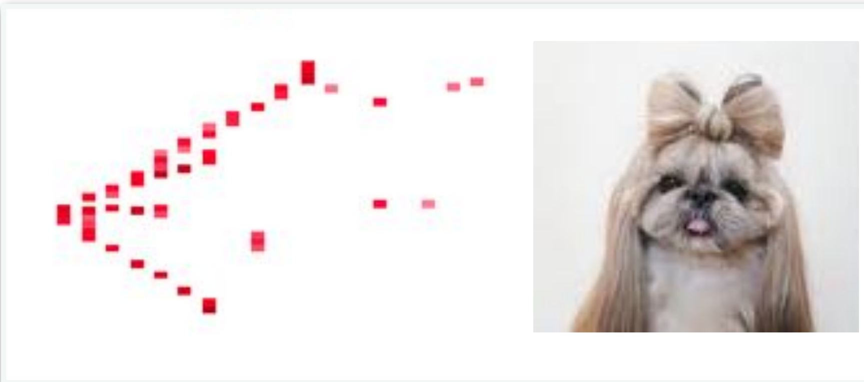
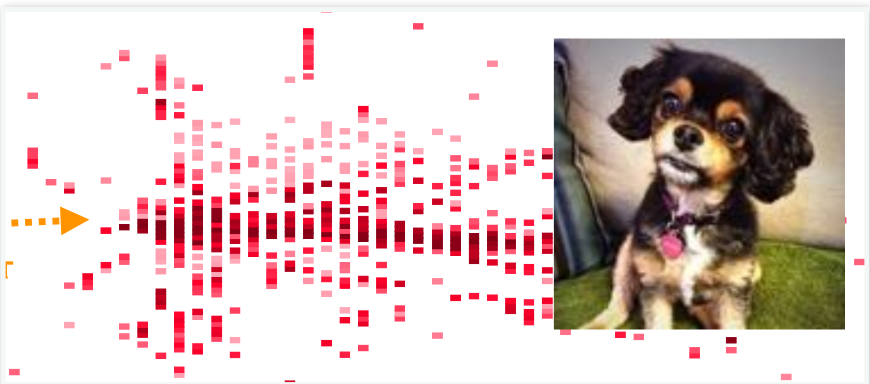
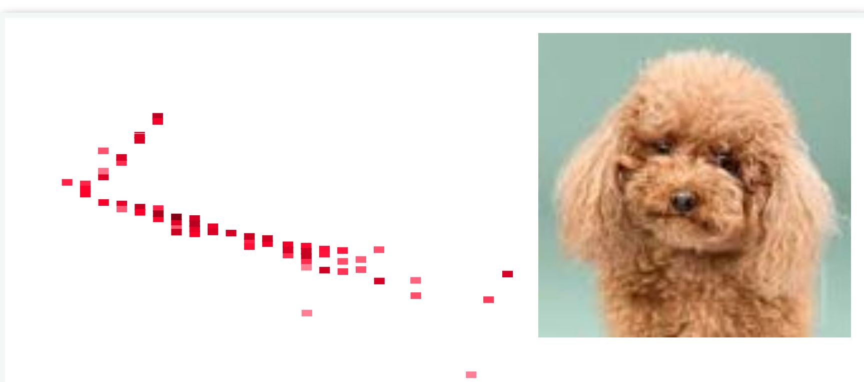
Trained on clusters from all events above some minimum purity.



Up to **11% improvement from adding context** to the classifier.

Challenges of applying ML

Performance and Robustness


Performance…

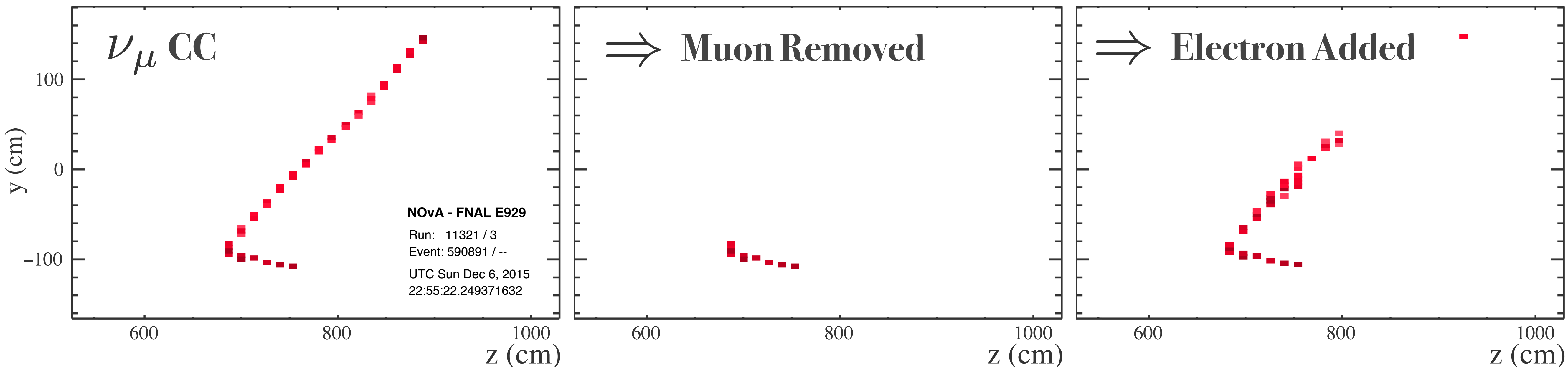
- ⭑ Overall accuracy
- ⭑ Behavior of loss functions, etc

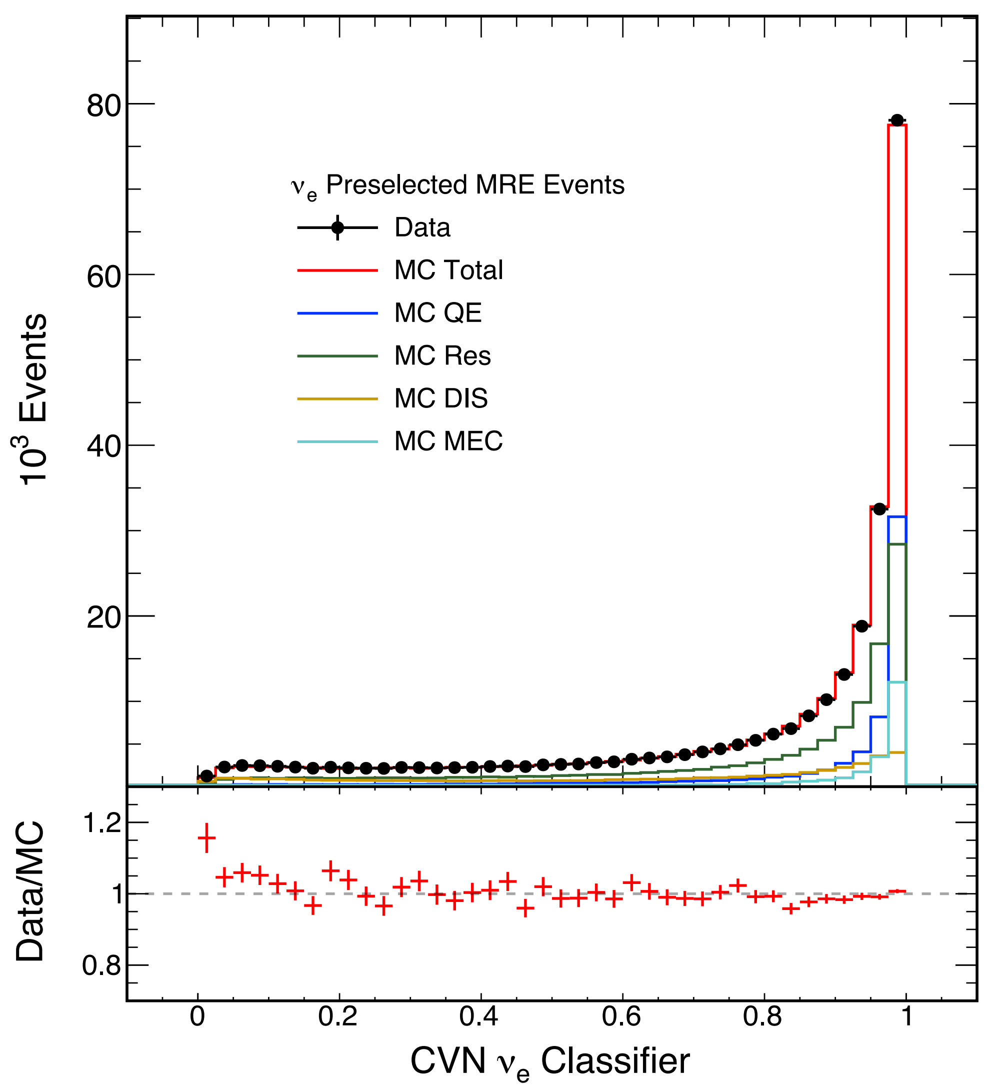
Robustness of the algorithms…

- ⭑ Effects of training sample composition
- ⭑ Studies of systematic uncertainties
- ⭑ Data driven performance tests

How do we find the biases we have introduced in our training?

Data-driven test example





MRE (Muon Removed - Electron added):

Select a muon neutrino interaction with traditional ID methods.

Remove the muon hits and replace them with a single simulated electron of matching momentum.

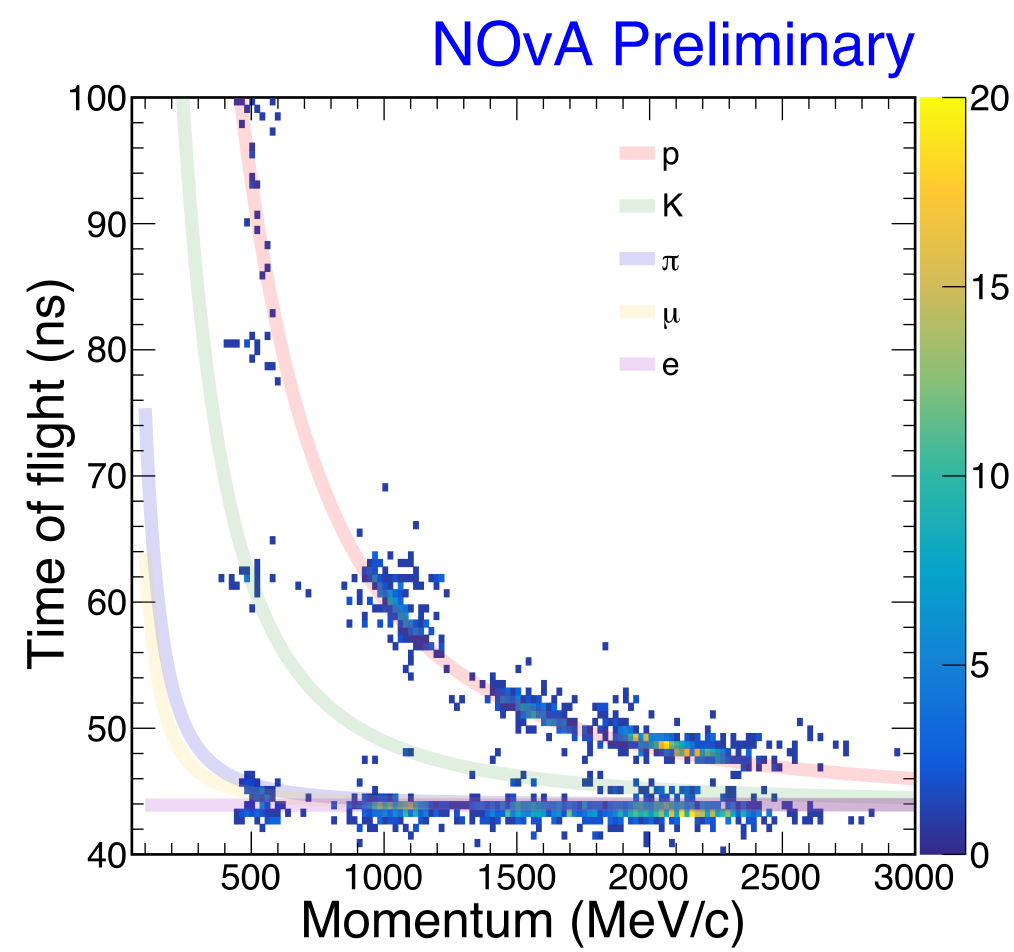
Data/MC comparisons show less than 1% difference in efficiency.

PID	Sample	Preselection	PID	Efficiency	Efficiency diff %
CVN	Data	262884	188809	0.718222	-0.36%
	MC	277320	199895	0.720809	

NOvA Test Beam

The NOvA test beam detector is currently taking data and will continue throughout 2019.

With a **library of labeled data from single particle interactions** of known identity and momentum, NOvA will expand the data-driven checks of our deep learning algorithms.



Summary

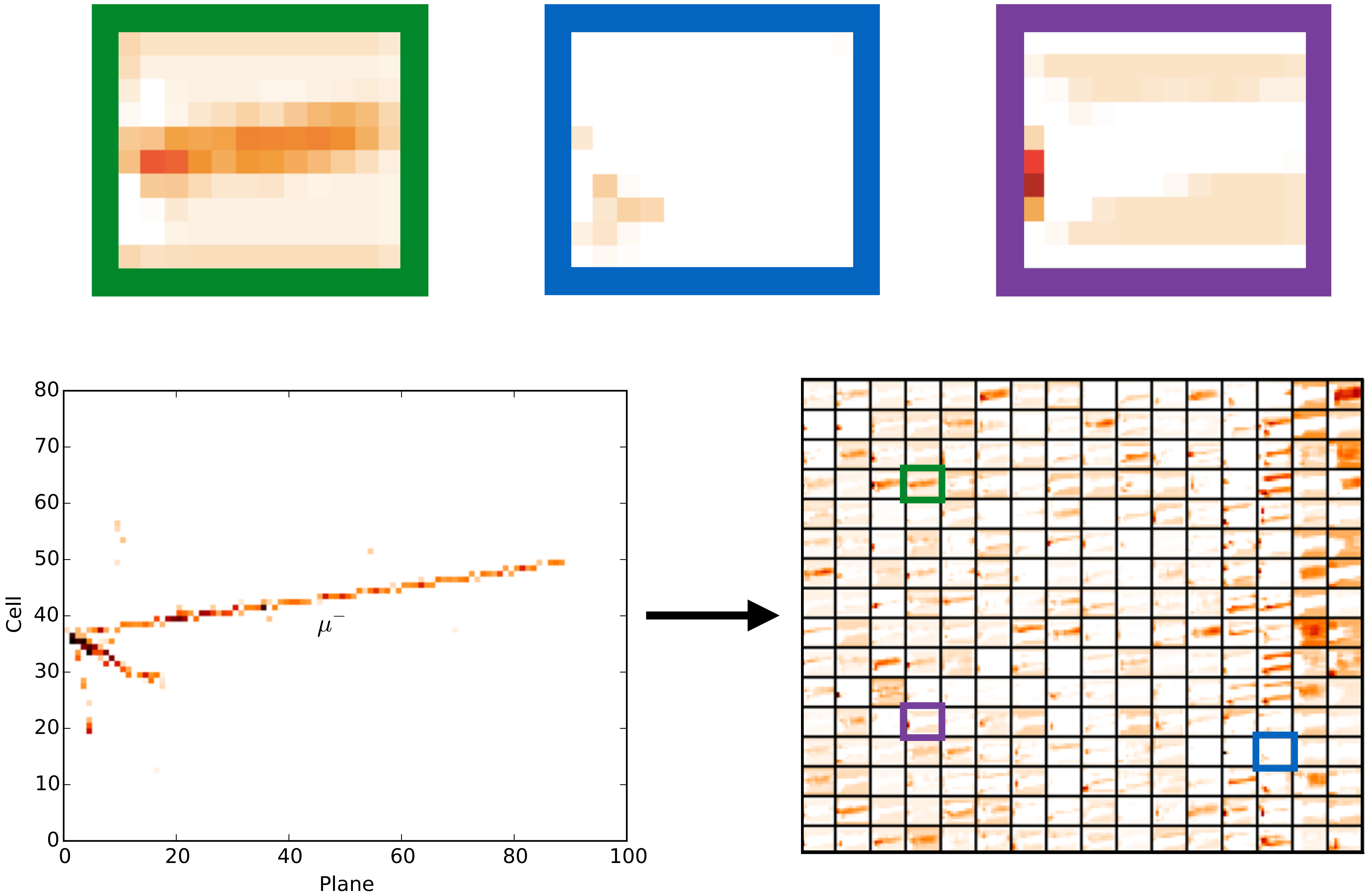
NOvA has adapted deep learning algorithms for classification, energy reconstruction, full event reconstruction, and other applications are being explored.

NOvA has demonstrated **substantial improvements** in performance from the use of deep learning.

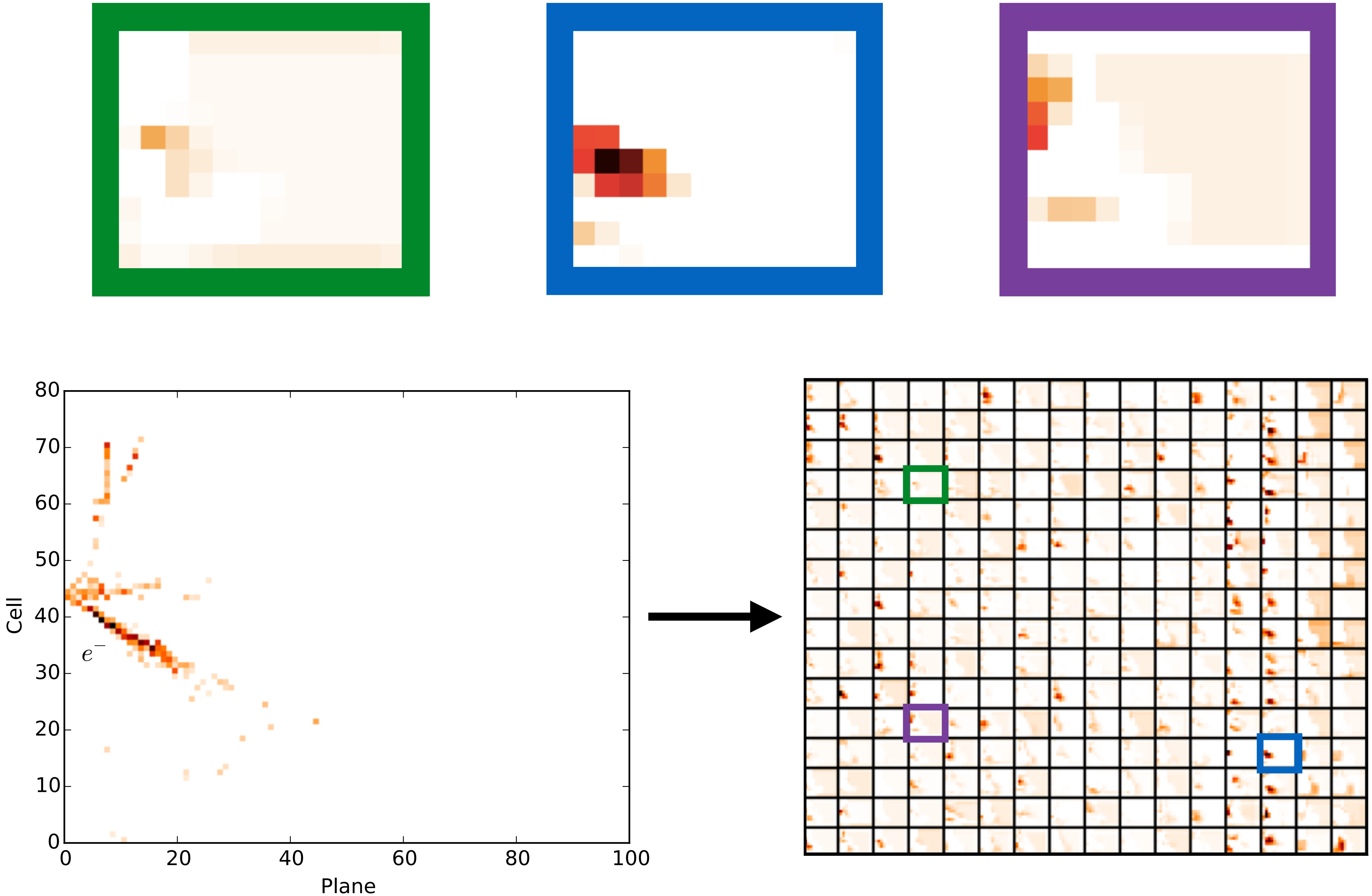
Applying our understanding of physics, our problem set, and our data can yield comparably **large improvements** in performance.

Data-driven tests are essential to show robustness of algorithms and assess some of the potential biases.

Thank you.



80
70
60
50
40
30
20
10
0
Cell
0
20
40
60
80
100
Plane
μ^-


80
70
60
50
40
30
20
10
0
Cell
e^-
0
20
40
60
80
100
Plane

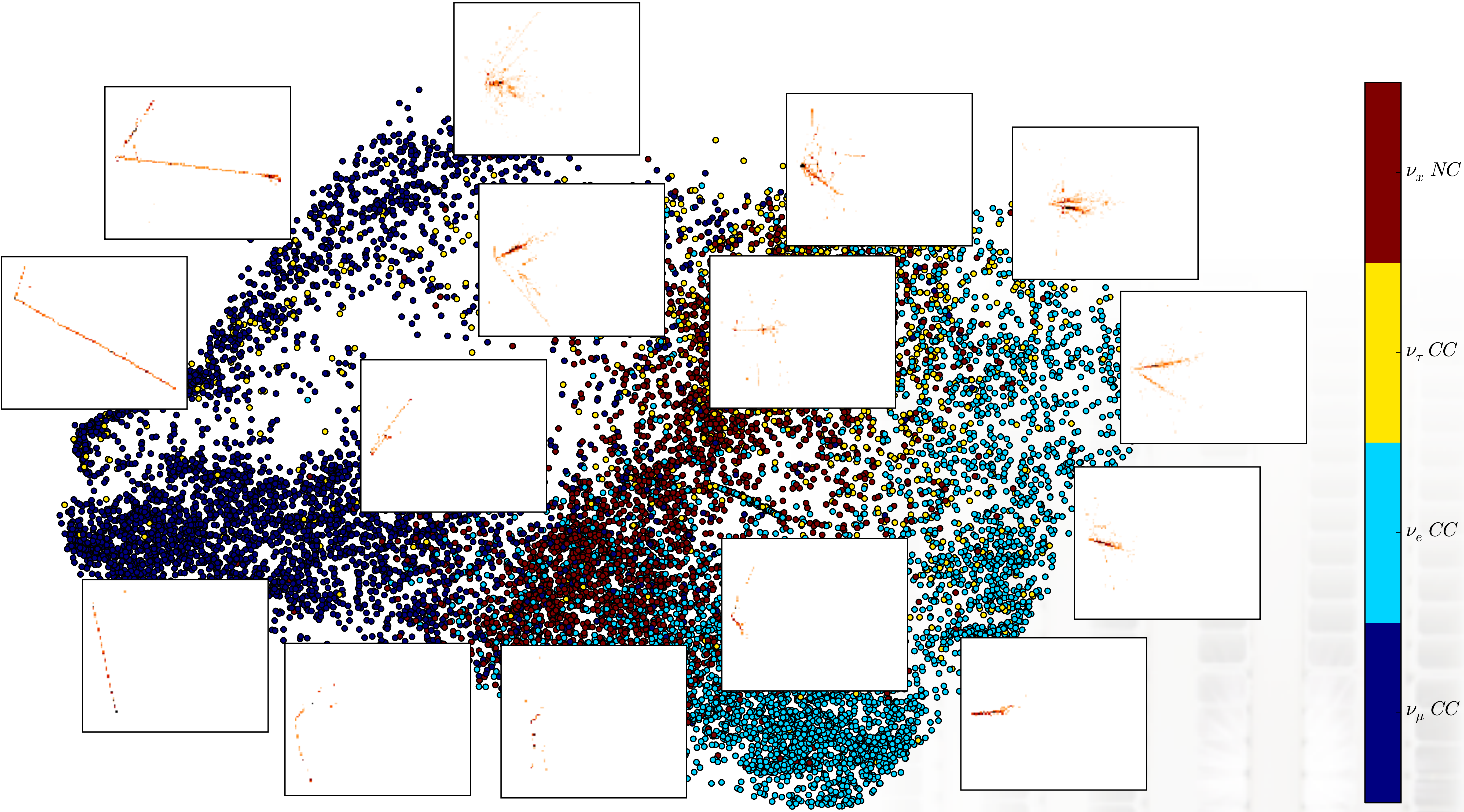
t-SNE t-Distributed Stochastic Neighbor Embedding

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

https://www.nature.com/articles/s41586-018-0361-2

t-SNE

t-Distributed Stochastic Neighbor Embedding



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

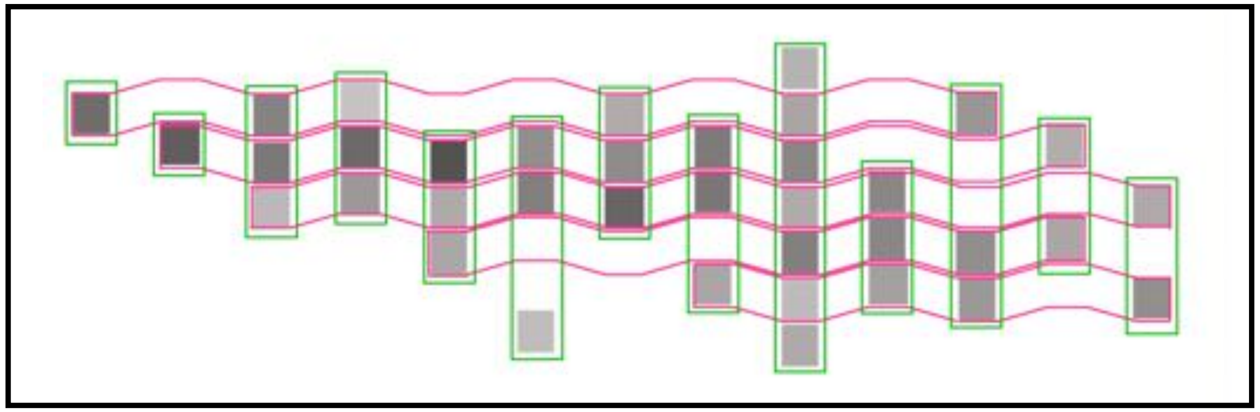
https://www.nature.com/articles/s41586-018-0361-2

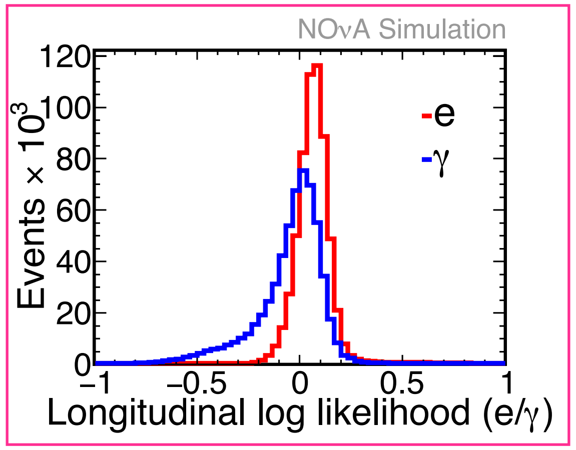
Neutrino Identification (First Analysis)

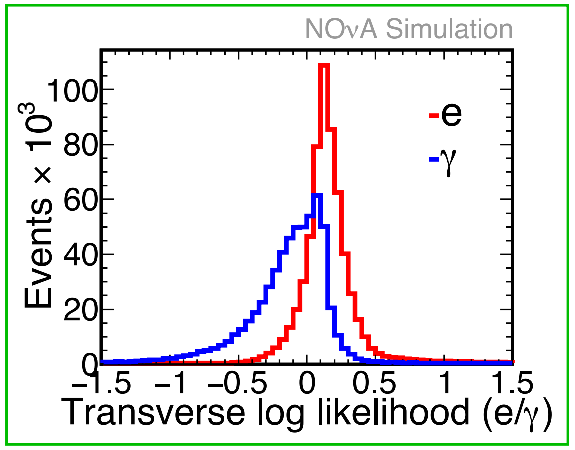
LID

Likelihood Identification

Premise: Electron showers have characteristic transverse and longitudinal energy deposition profiles.







In practice:

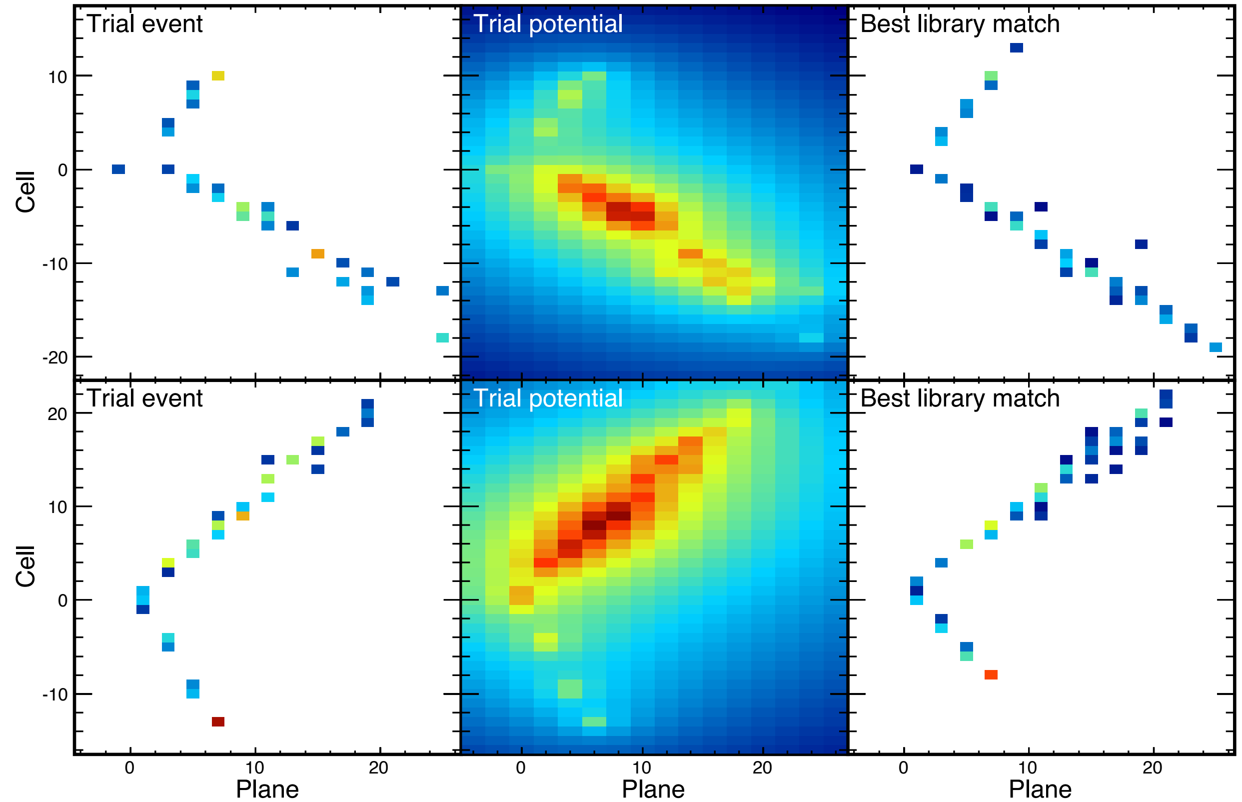
- ★ Reconstruct electron shower.
- ★ Find likelihoods from it's dE/dx profiles compared to particle hypotheses.

Likelihoods ➡ *Neural Network*

LEM

Library Event Matching

Premise: We have a large library of simulated event templates, large enough that we can use it to compare pixel by pixel.



In practice:

- ★ Find the best matches from the event library.
- ★ Extract features from best matches.

Features ➡ *Decision Tree*

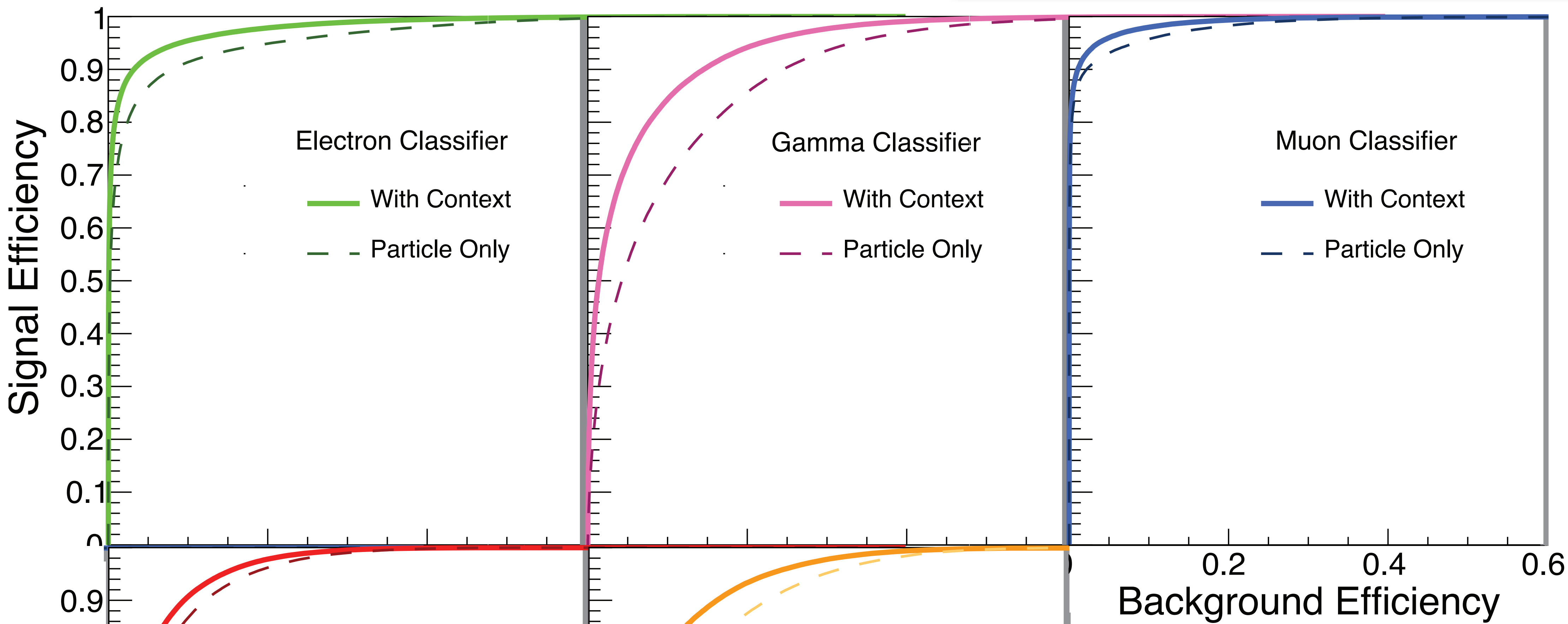
Context Improves ID

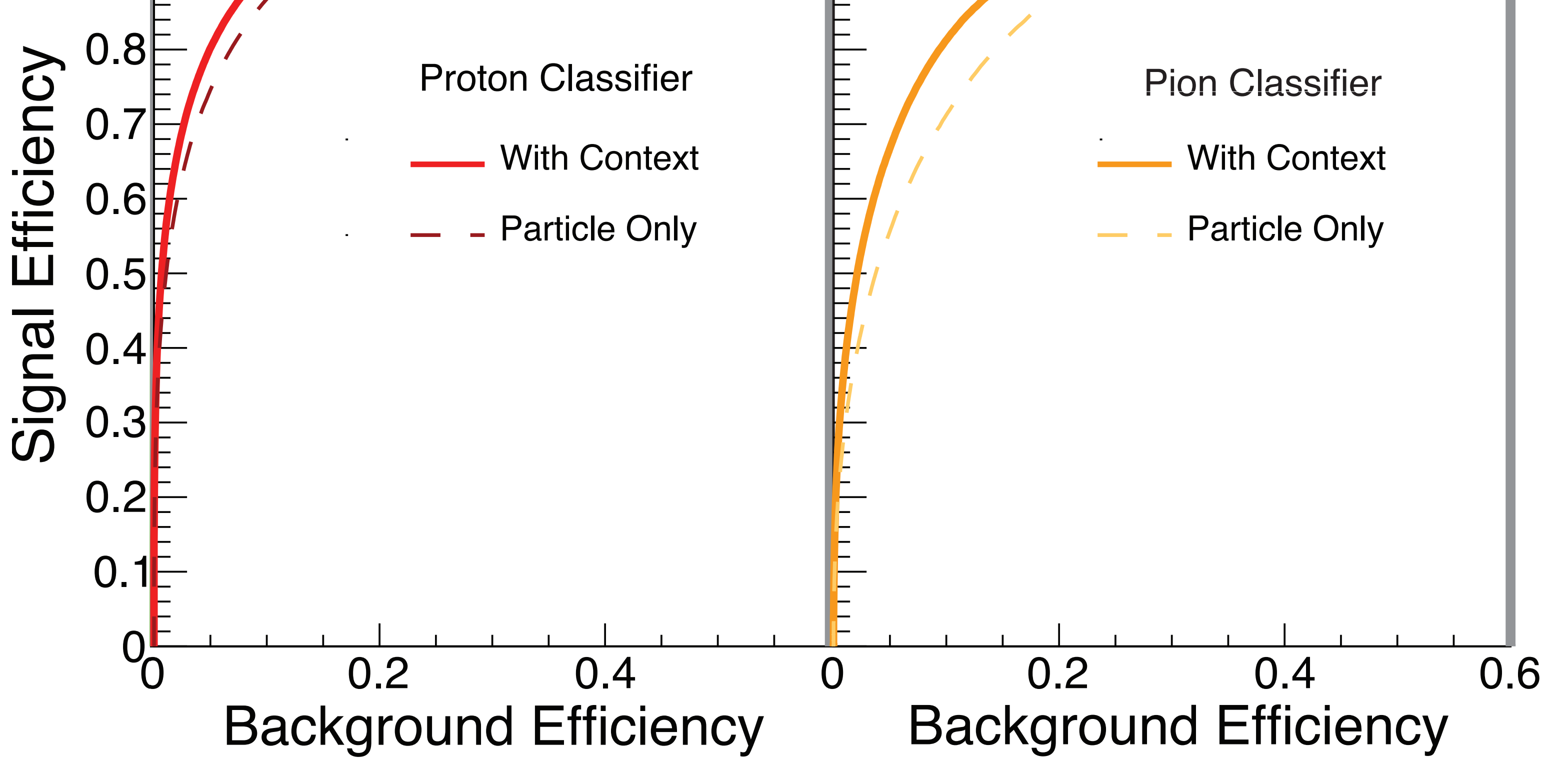
Soon to be on PRD

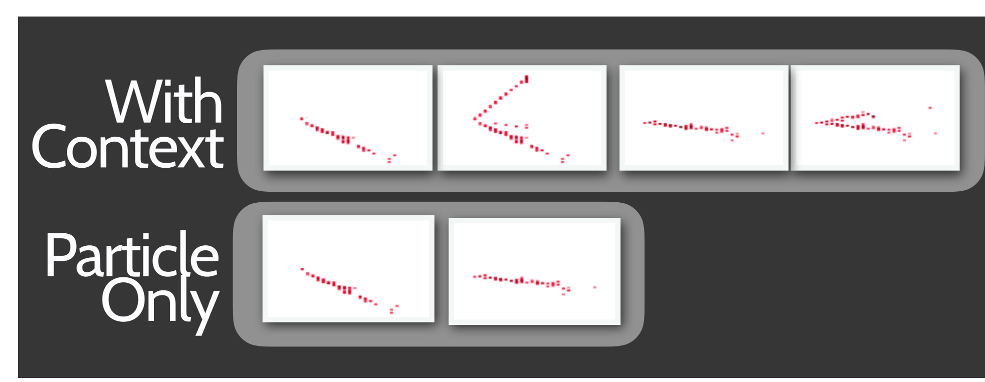
Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events

F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic, A. Sousa

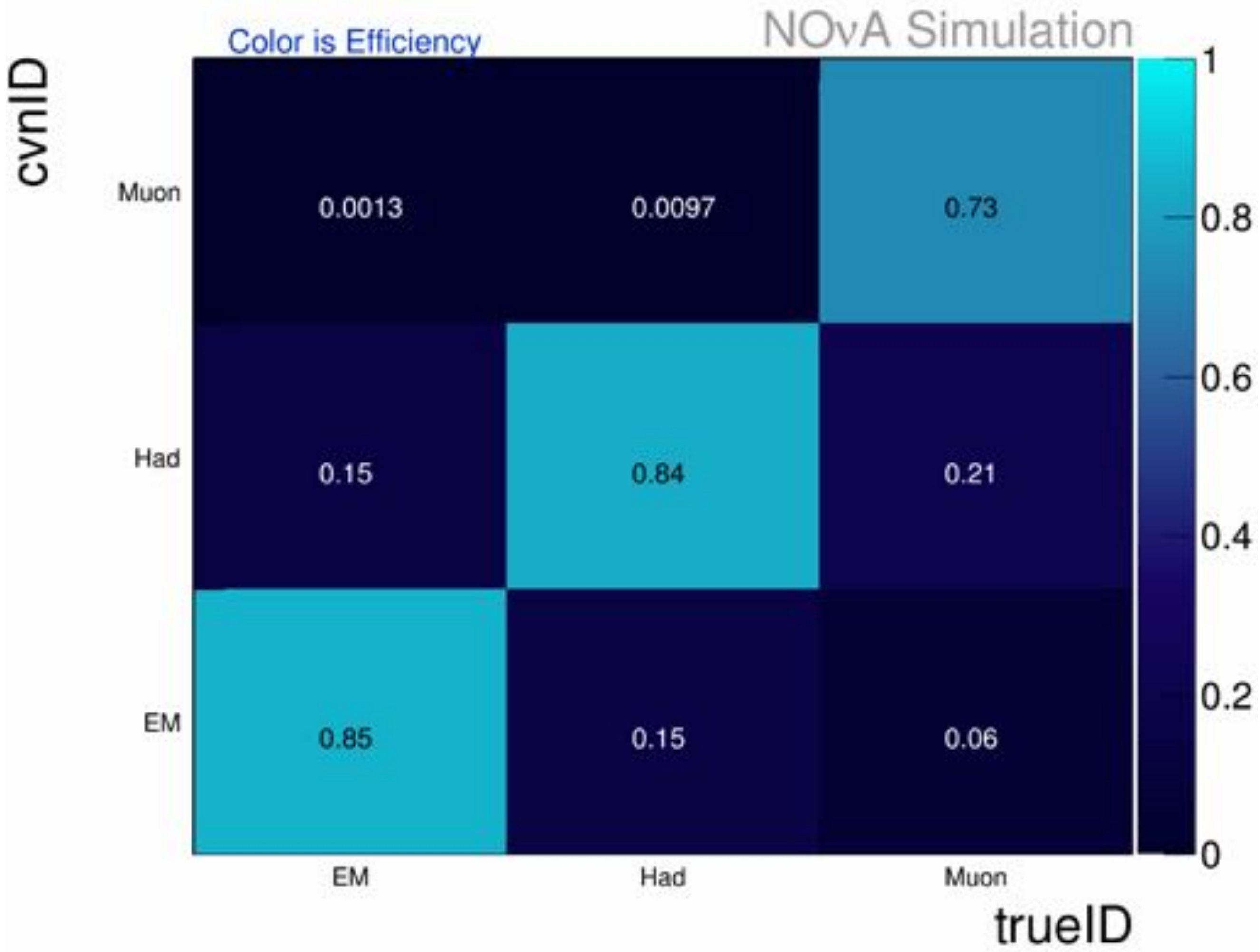
(Submitted on 3 Jun 2019)

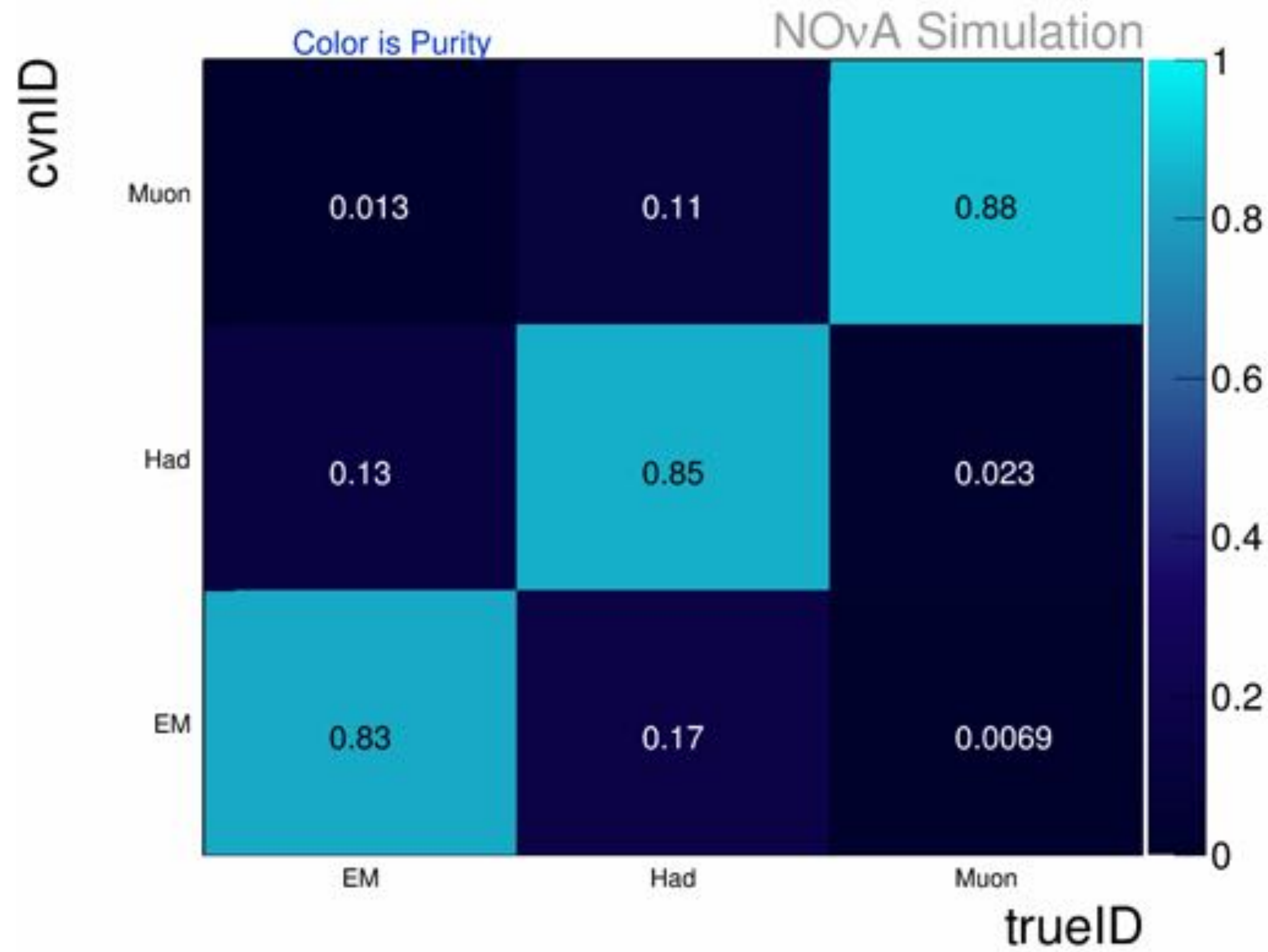


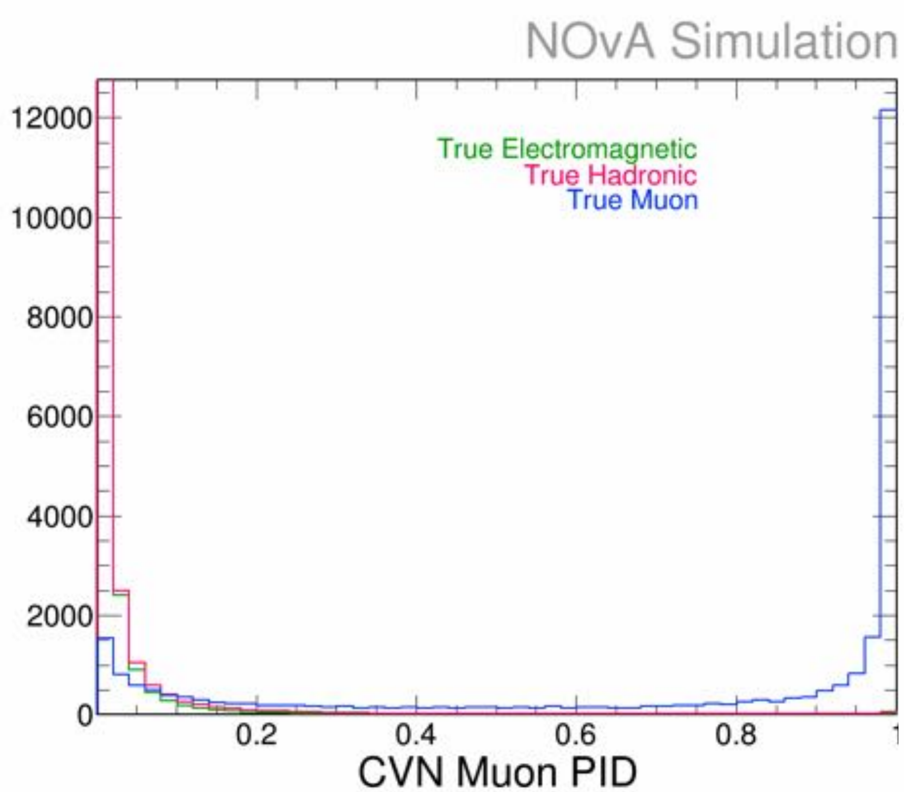


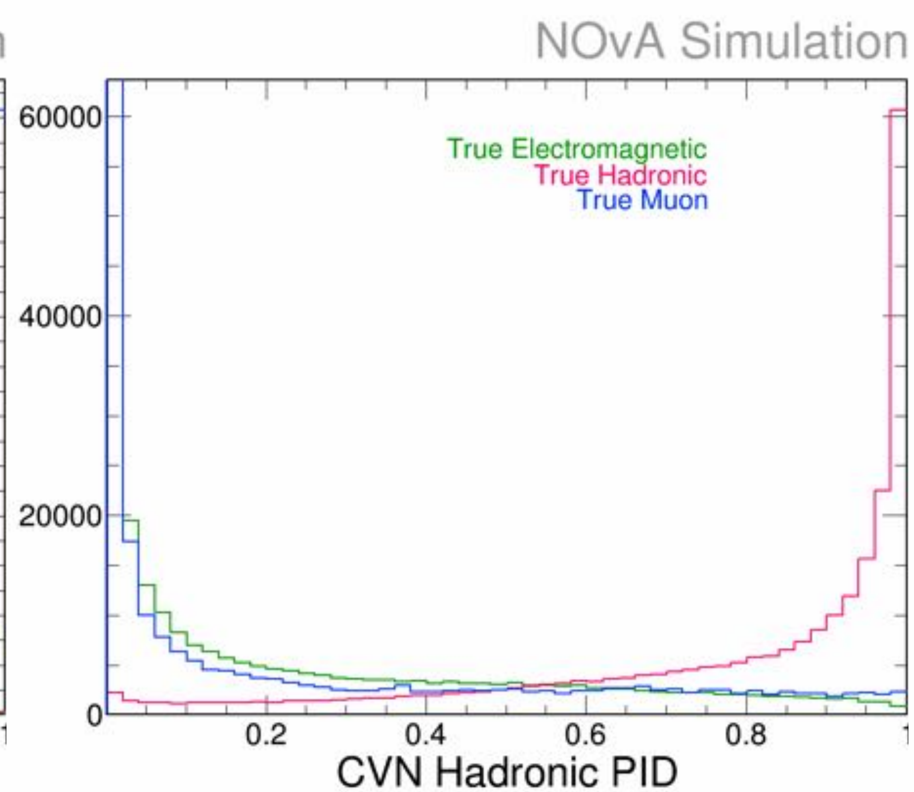


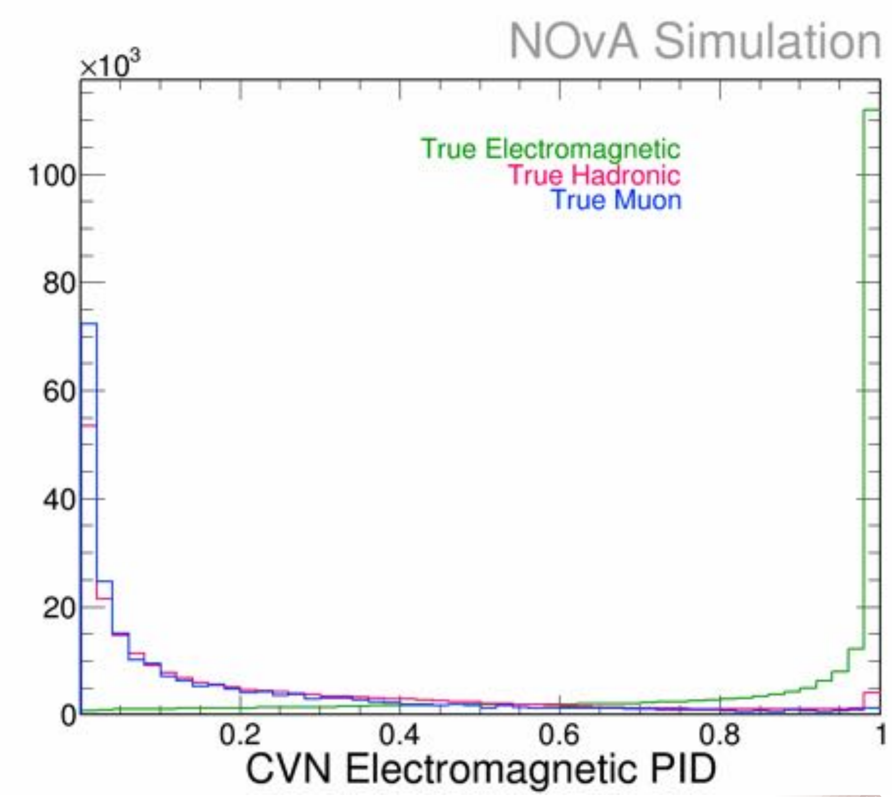
Up to 11% improvement from adding context to the classifier.


Color is Efficiency
NOvA Simulation
cvnID
Muon
Had
EM
0.0013
0.0097
0.73
0.15
0.84
0.21
0.85
0.15
0.06
EM
Had
Muon
trueID


Color is Purity
NOvA Simulation
cvnID
Muon
Had
EM
0.013
0.11
0.88
0.13
0.85
0.023
0.83
0.17
0.0069
EM
Had
Muon
trueID


NOvA Simulation
True Electromagnetic
True Hadronic
True Muon
CVN Muon PID


NOvA Simulation
True Electromagnetic
True Hadronic
True Muon
CVN Hadronic PID


NOvA Simulation
True Electromagnetic
True Hadronic
True Muon
CVN Electromagnetic PID


Fermilab