Deep Learning on NOvA

Applications, Successes and Lessons

Fernanda Psihas

Fermilab

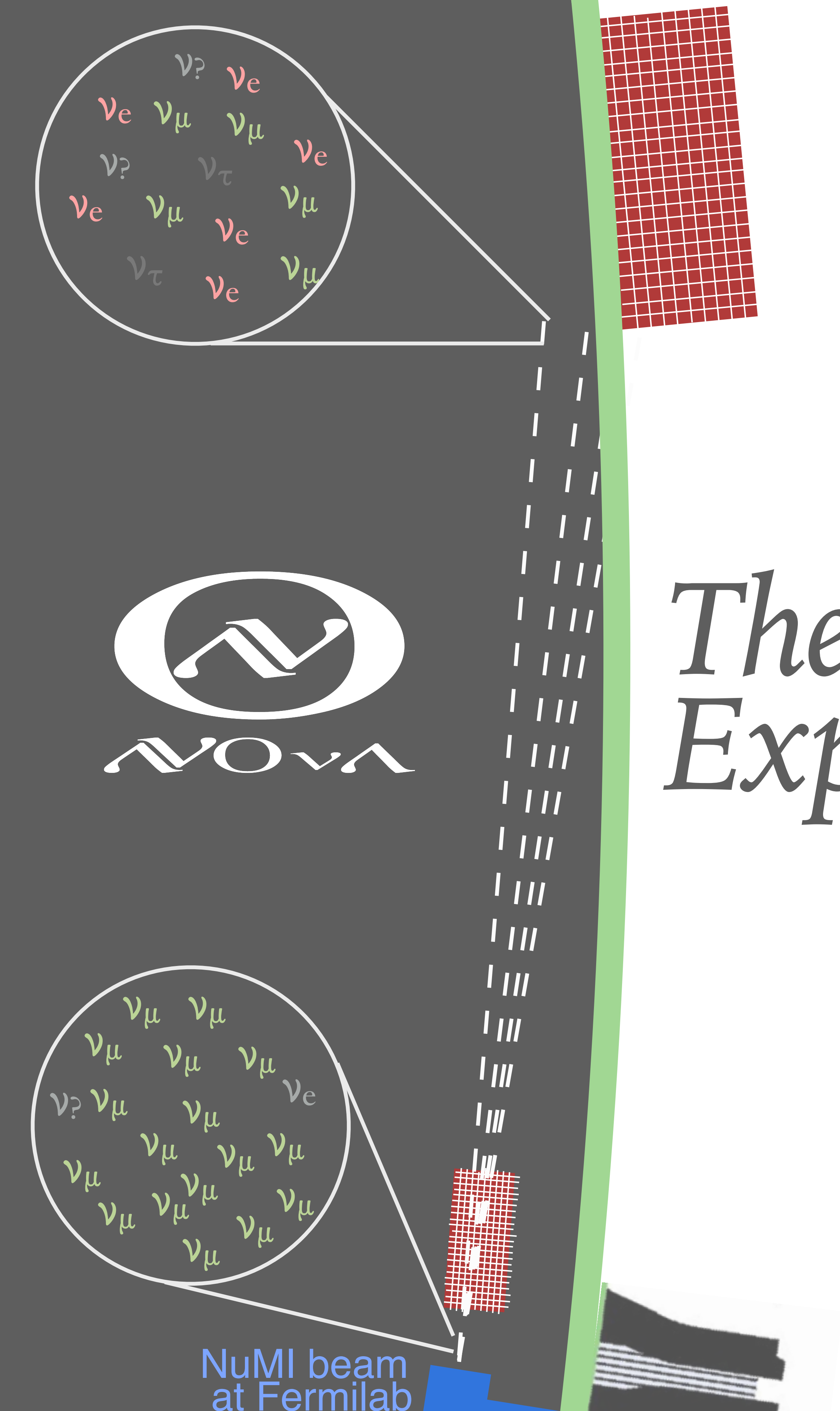
Outline

NOvA & our deep learning program

- Experiment, physics, data
- Summary of DL applications

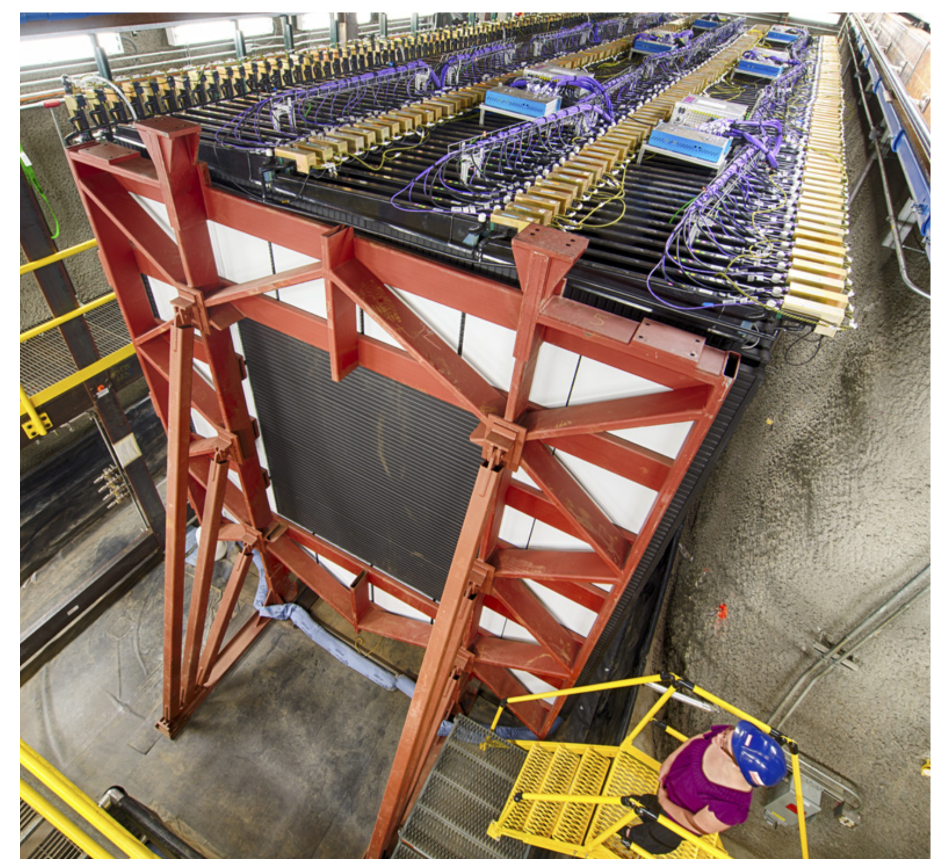
Challenges & Lessons

- Think about the ***physics***: applications & data
- Think about the ***physicists***: framework & training
- Think about the ***problems***: Robustness, uncertainties, etc

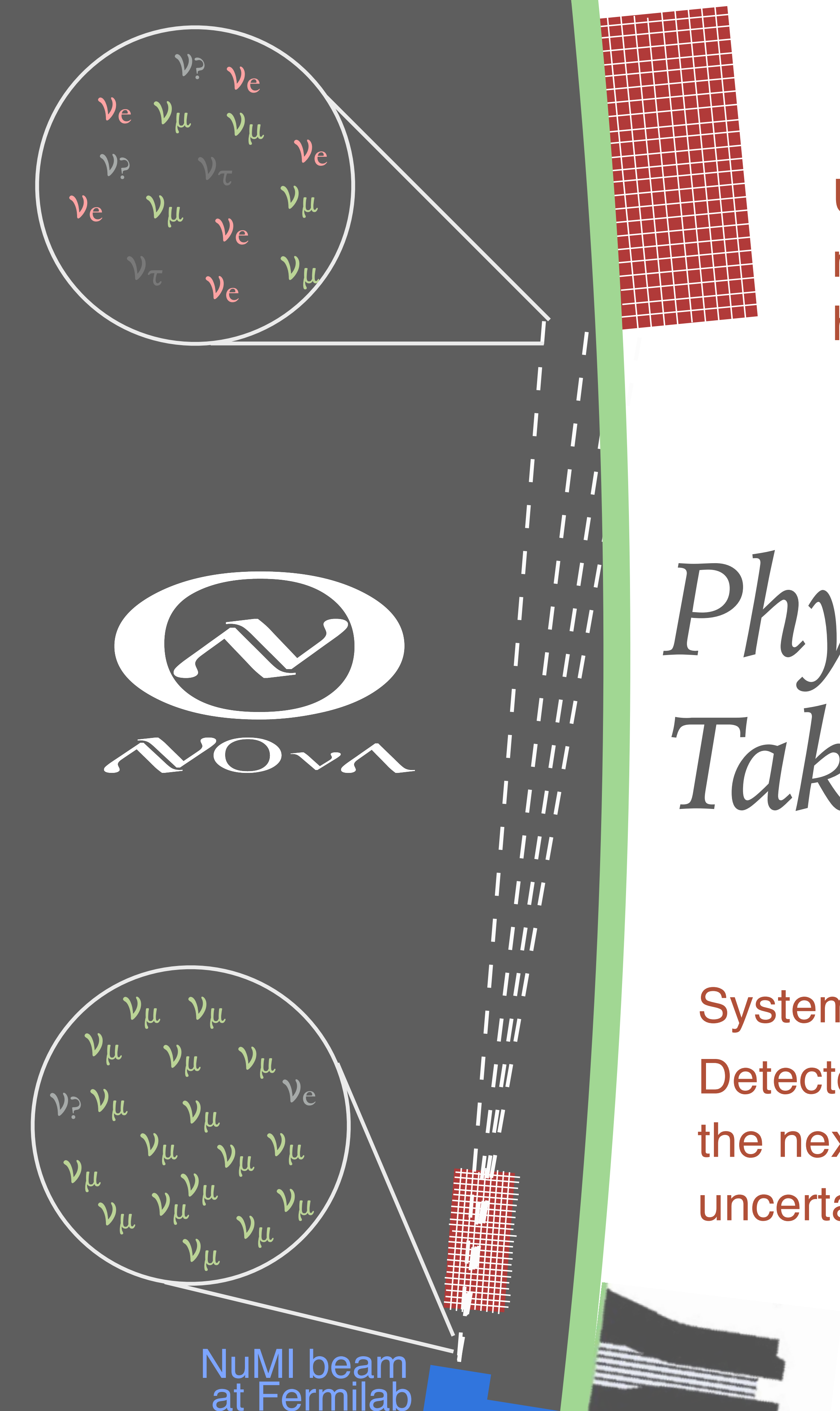


Neutrino oscillation measurements at the Far Detector

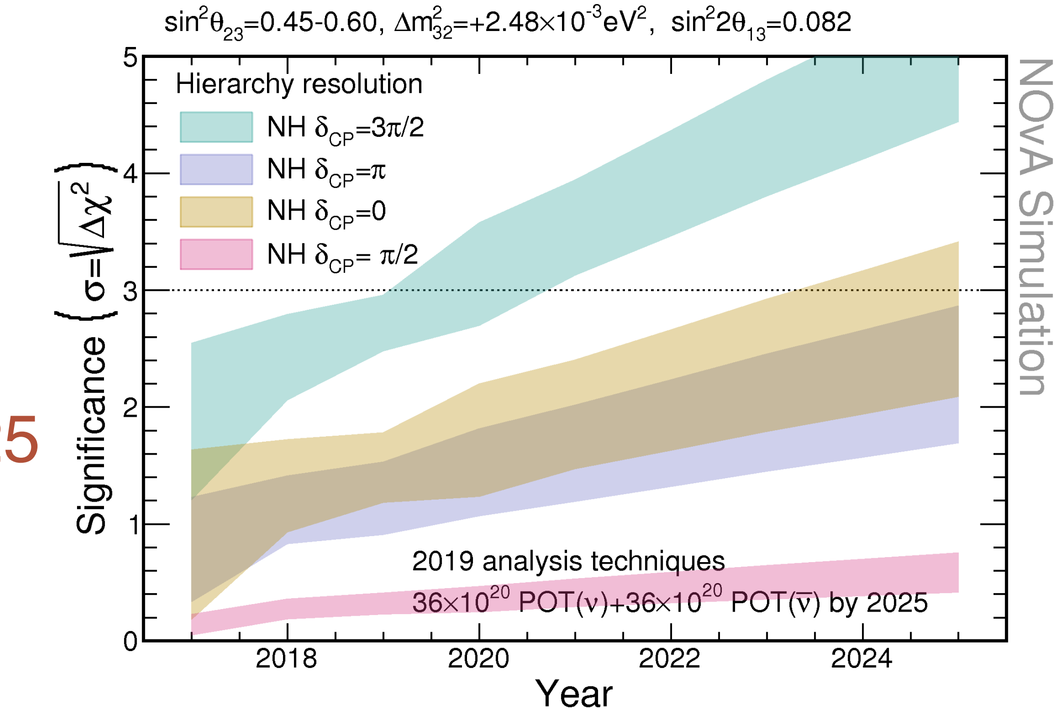
The NOvA Experiment



Neutrino interaction cross-sections measurements at the Near Detector



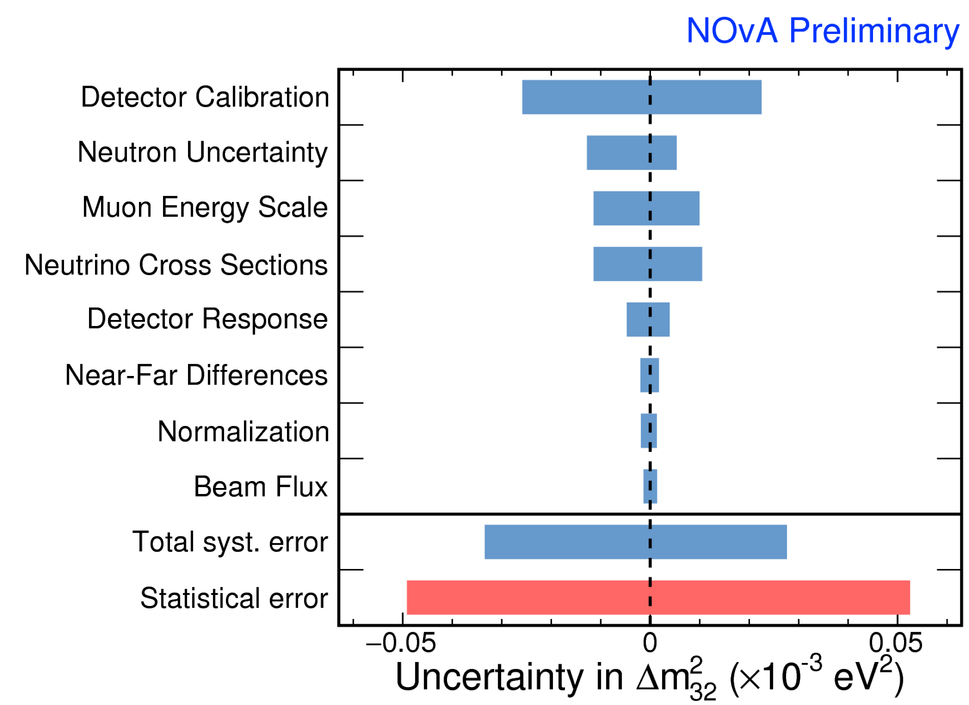
Up to 5σ on the neutrino mass hierarchy by 2025



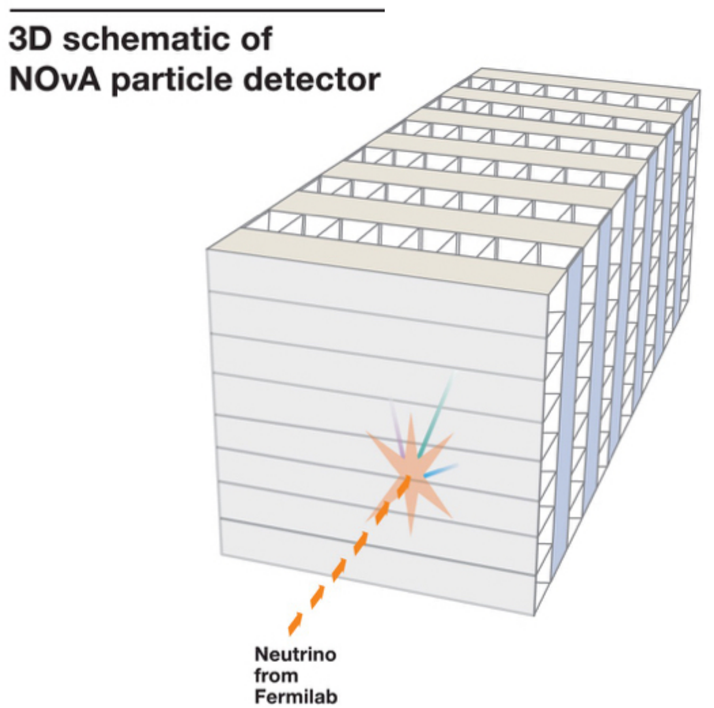
Physics Takeaways

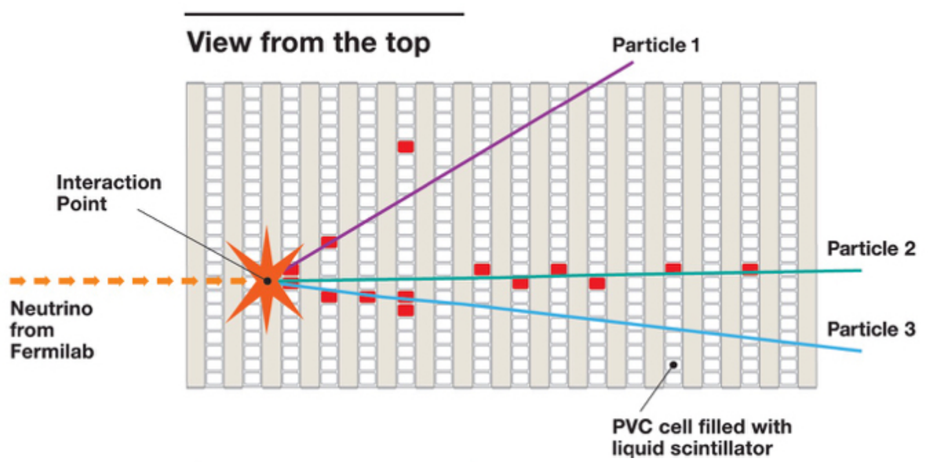
Also… sterile neutrinos, supernova neutrinos, cosmic-ray physics, and more!

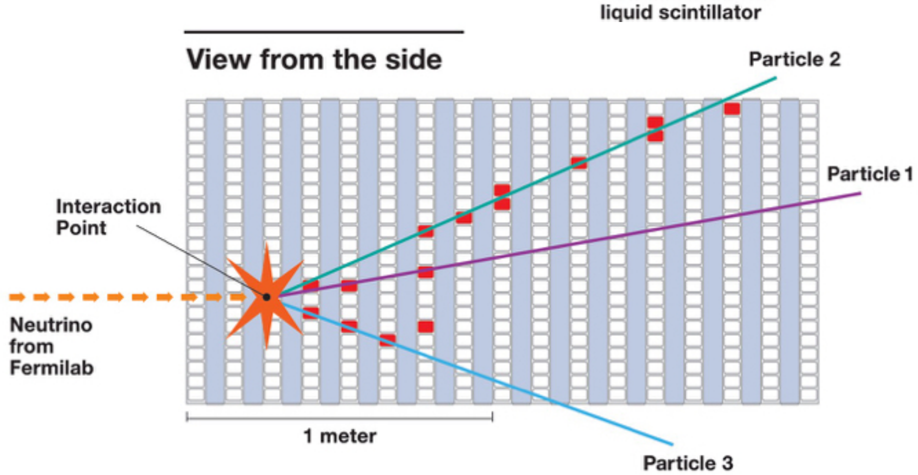
Systematics limited, Detector response is the next largest uncertainty.



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

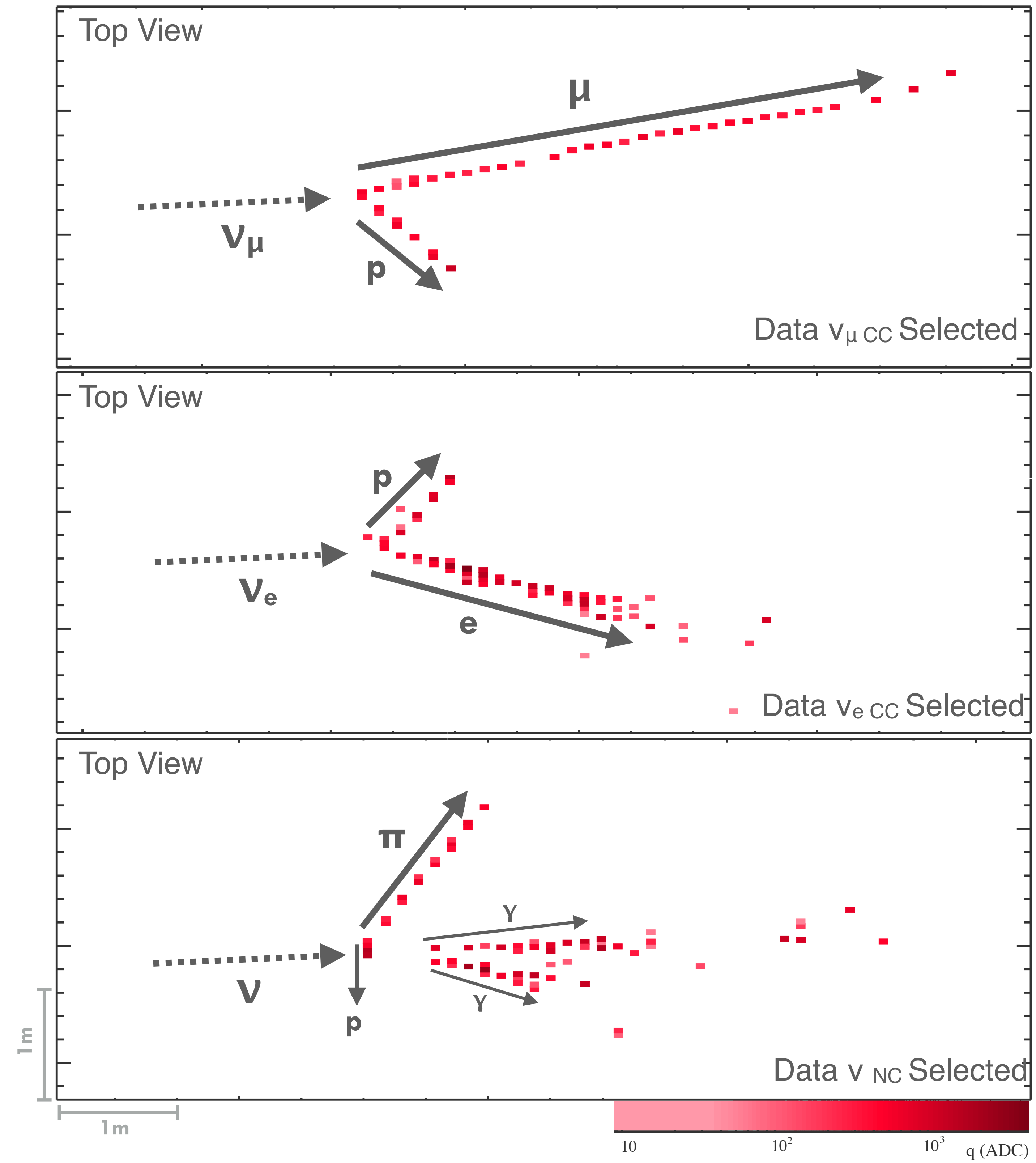
For interaction cross-sections we want to measure the incoming neutrino flavor and energy, as well as the final state.

NOvA event topology

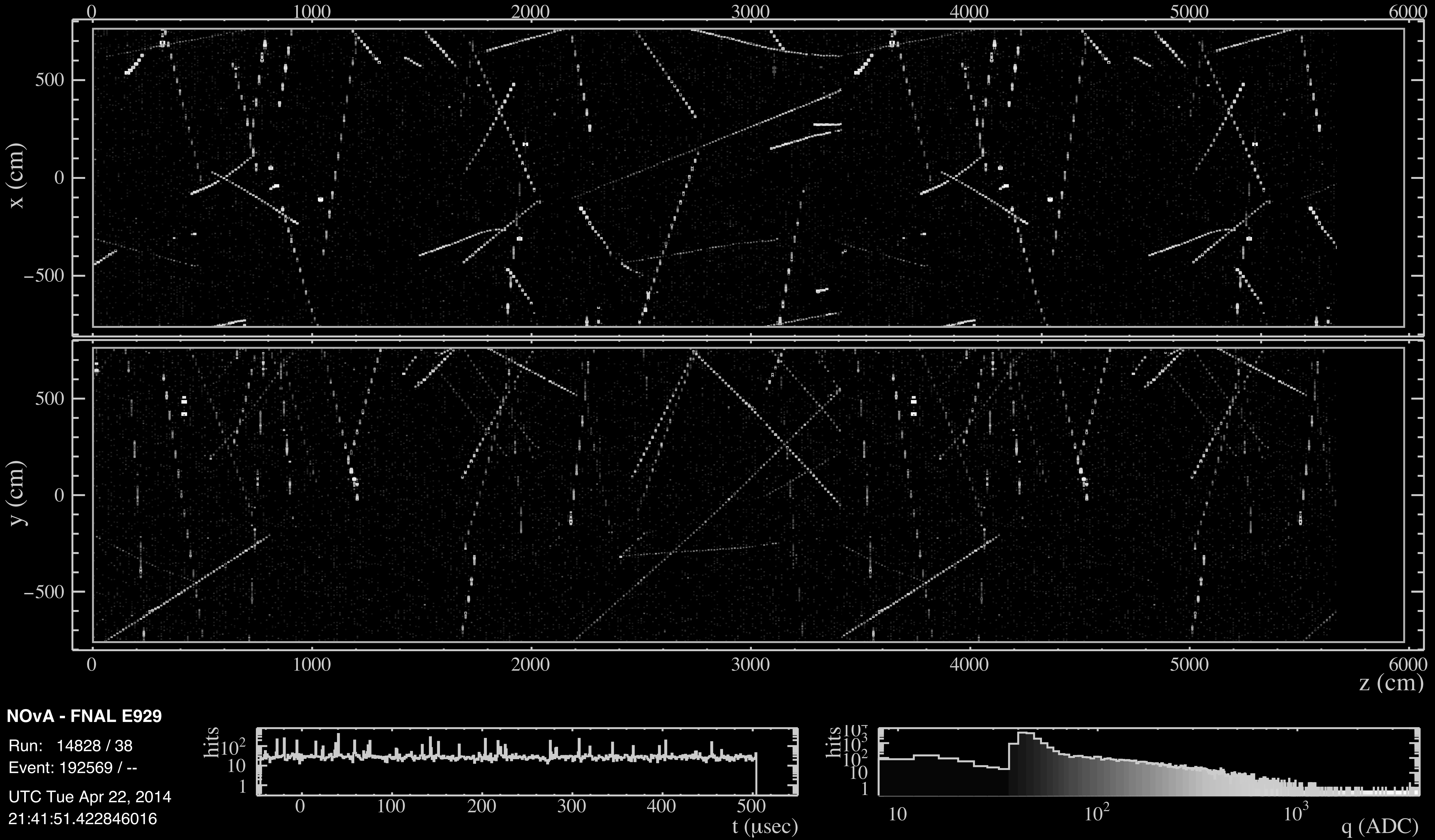
Long tracks are typically muons (minimally ionizing).

e and γ produce electromagnetic showers spanning multiple detector cells.

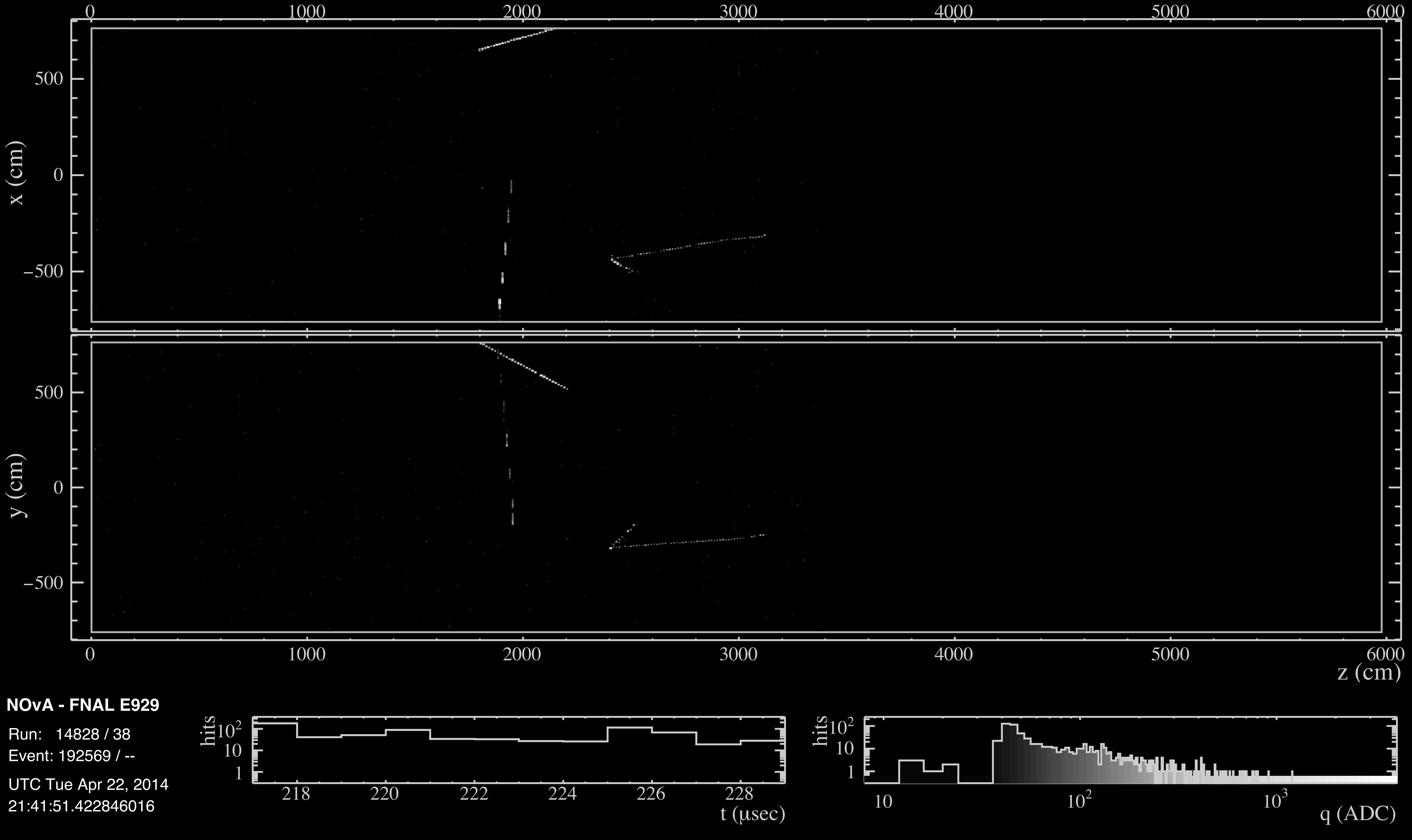
Heavier particles typically shorter, higher dE/dx tracks.



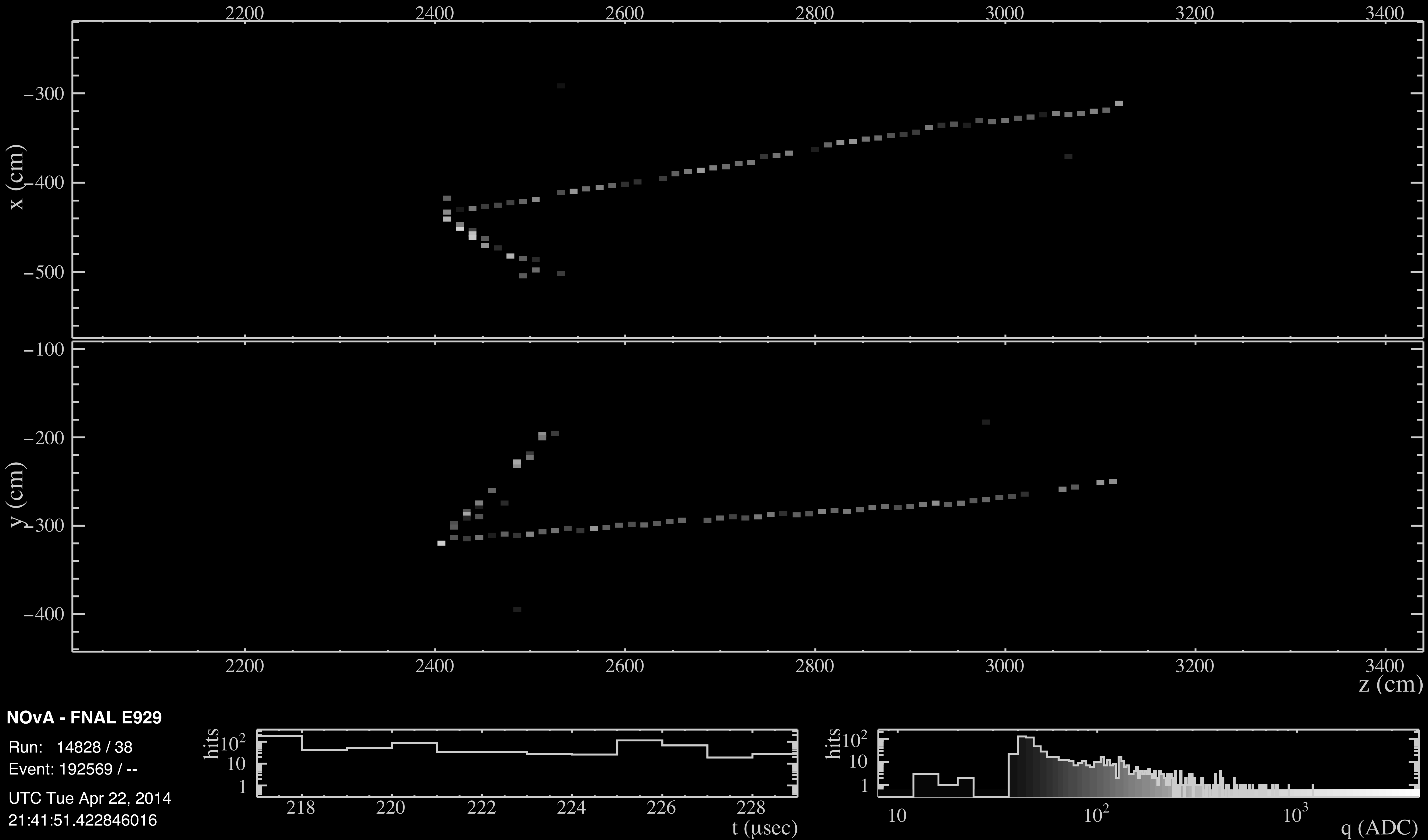
NOvA's Reconstruction in 2 minutes



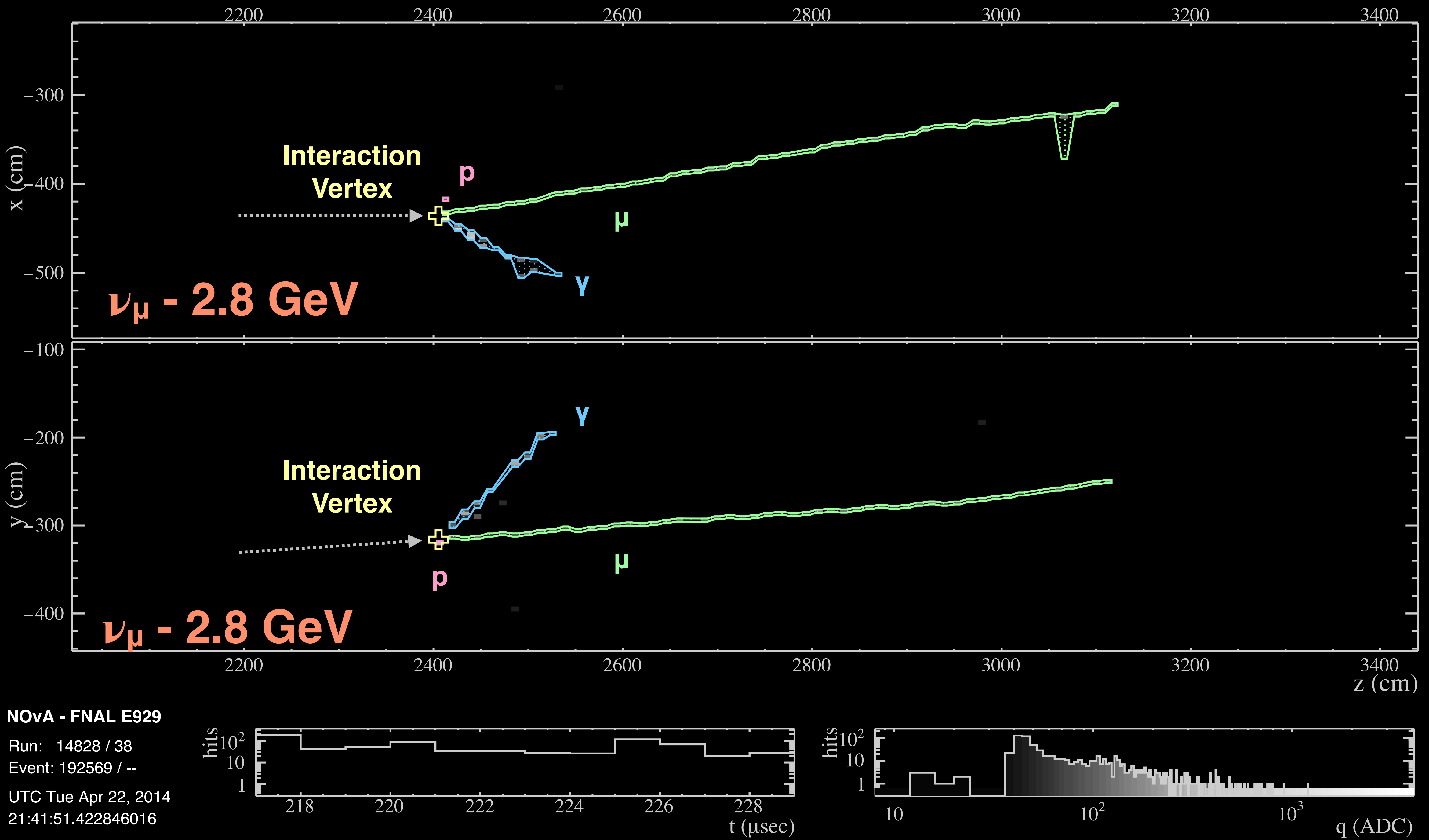
NOvA's Reconstruction in 2 minutes



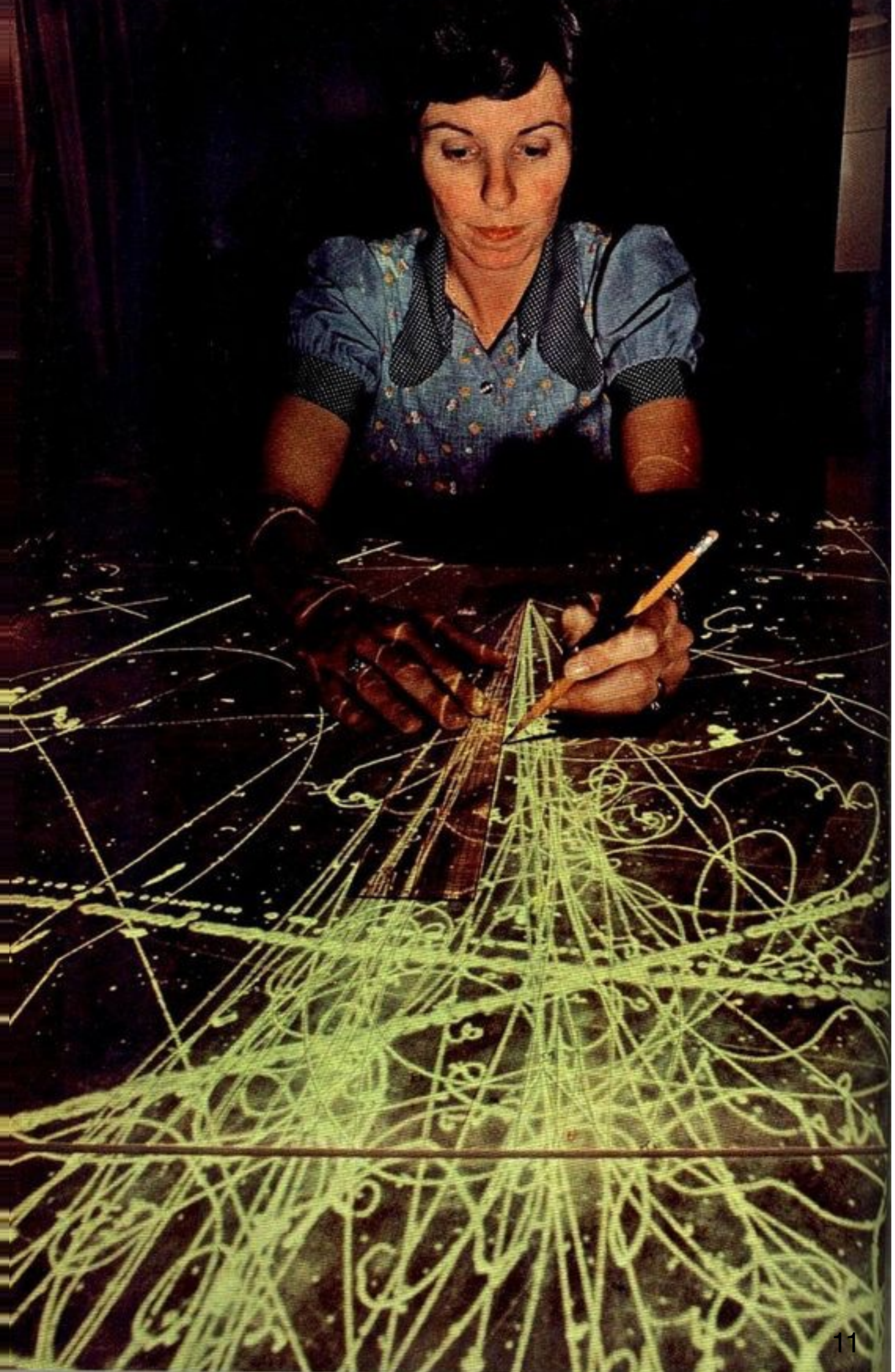
NOvA's Reconstruction in 2 minutes



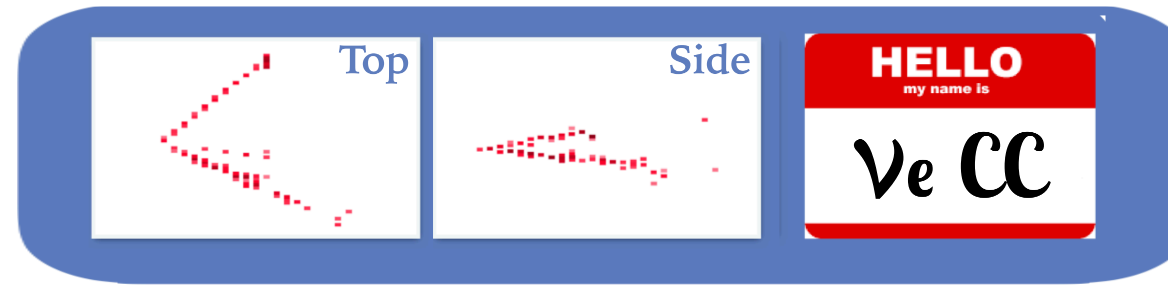
NOvA's Reconstruction in 2 minutes



Reconstruction and Identification



Neutrino Flavor Identification

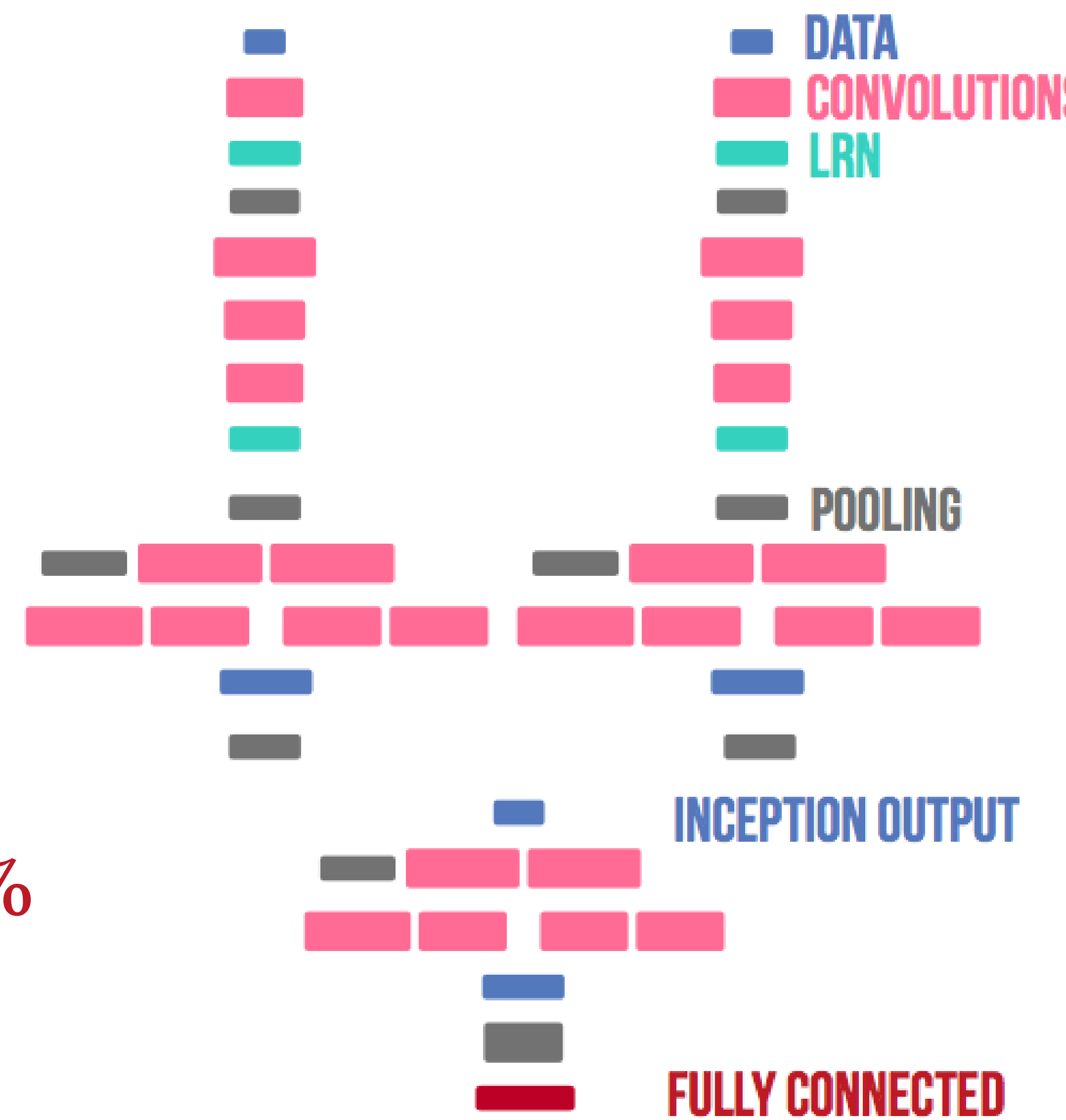


Two-tower CNN architecture **learns from the top and side views** independently first.

We train on **~5M neutrino MC + cosmic data** events.

This was the **first implementation** of Convolutional Neural Networks on a HEP result.

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



Published.

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))

Existing Deep Learning Applications

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))

Improved Energy Reconstruction in NOvA with Regression Convolutional Neural Networks

Pierre Baldi, Jianming Bian, Lars Hertel, Lingge Li

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, and A. Sousa
Phys. Rev. D **100**, 073005 – Published 14 October 2019

In Development

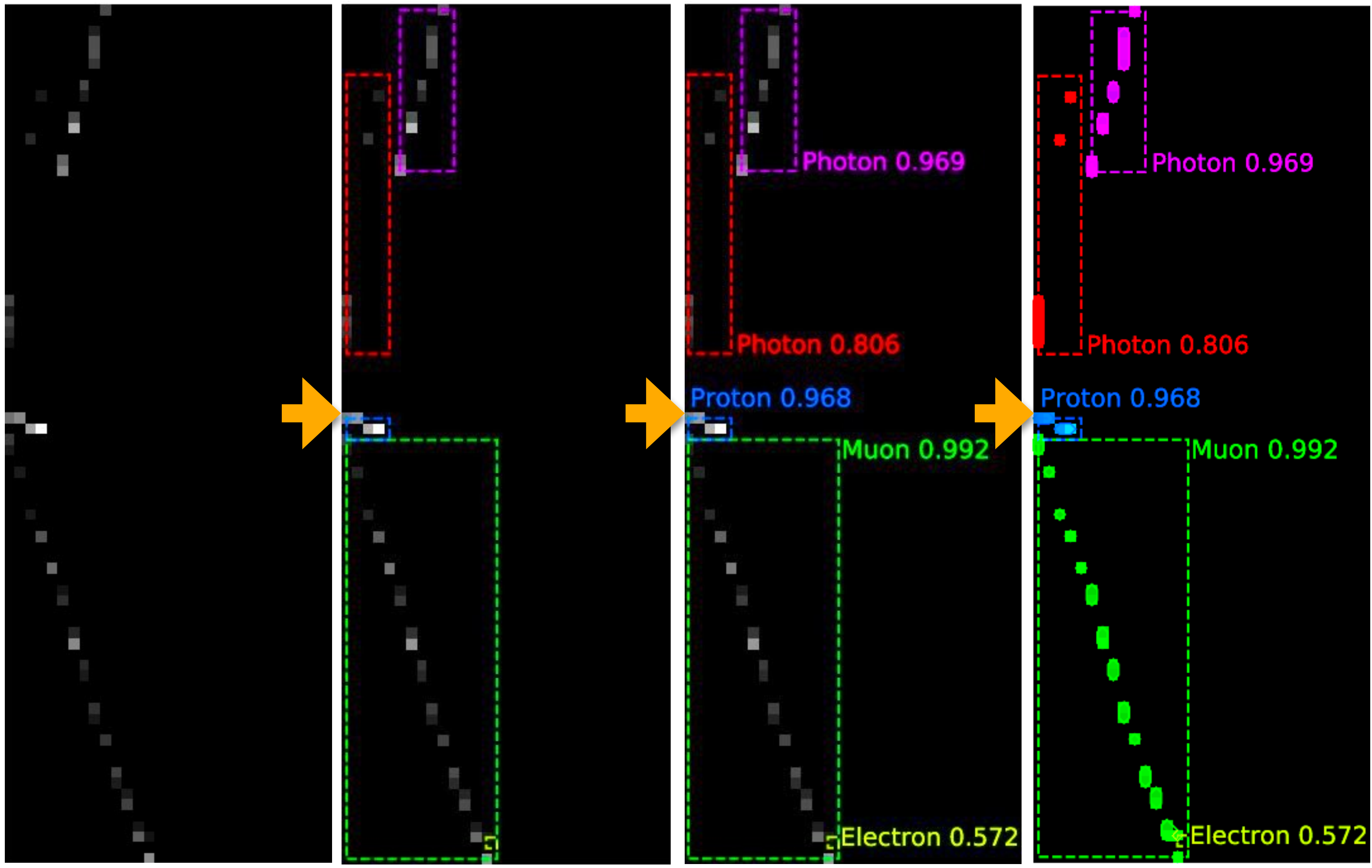
LSTMs for energy reconstruction

CNN cosmic rejection

Systematics-robust trainings

Clustering w/graph networks

Full event reconstruction with semantic segmentation.



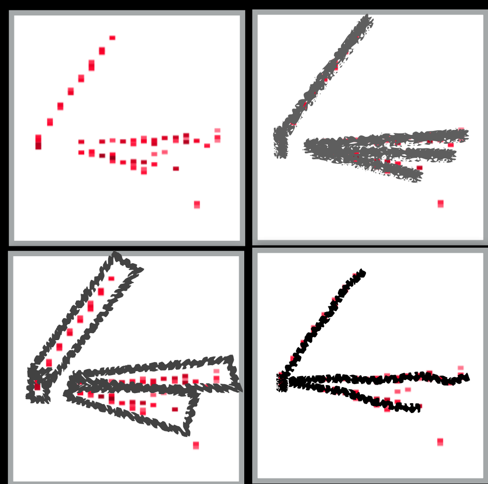
*Think about the **physics***

What is the best tool for your problem-set.

What do you know about your data?

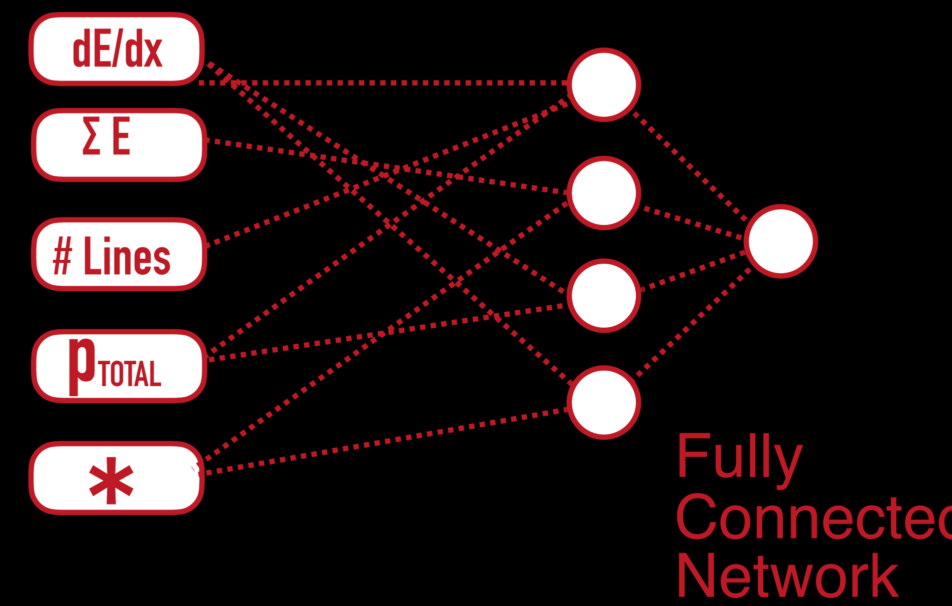
How can you enhance/best use the information contained in your data with

Geometrical Reconstruction

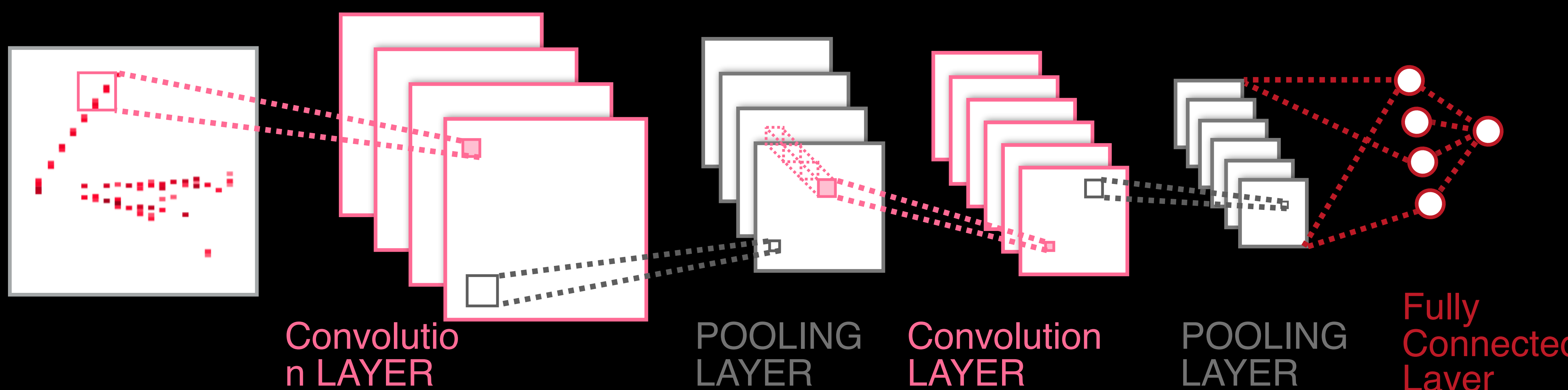


Extract information (features) from the event which can separate signal from background.

Train a neural network for signal vs background discrimination.



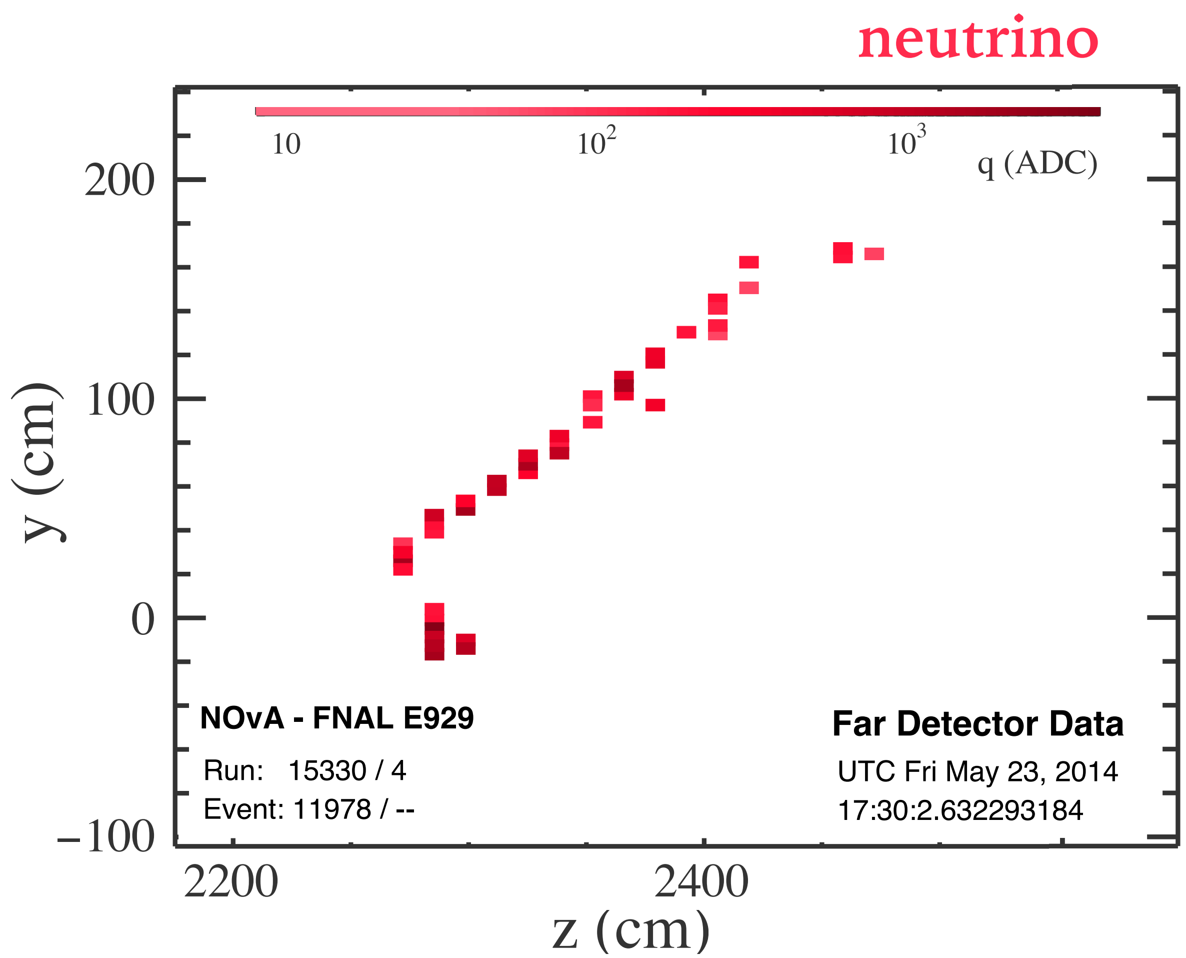
Deep Learning with Convolutional Networks

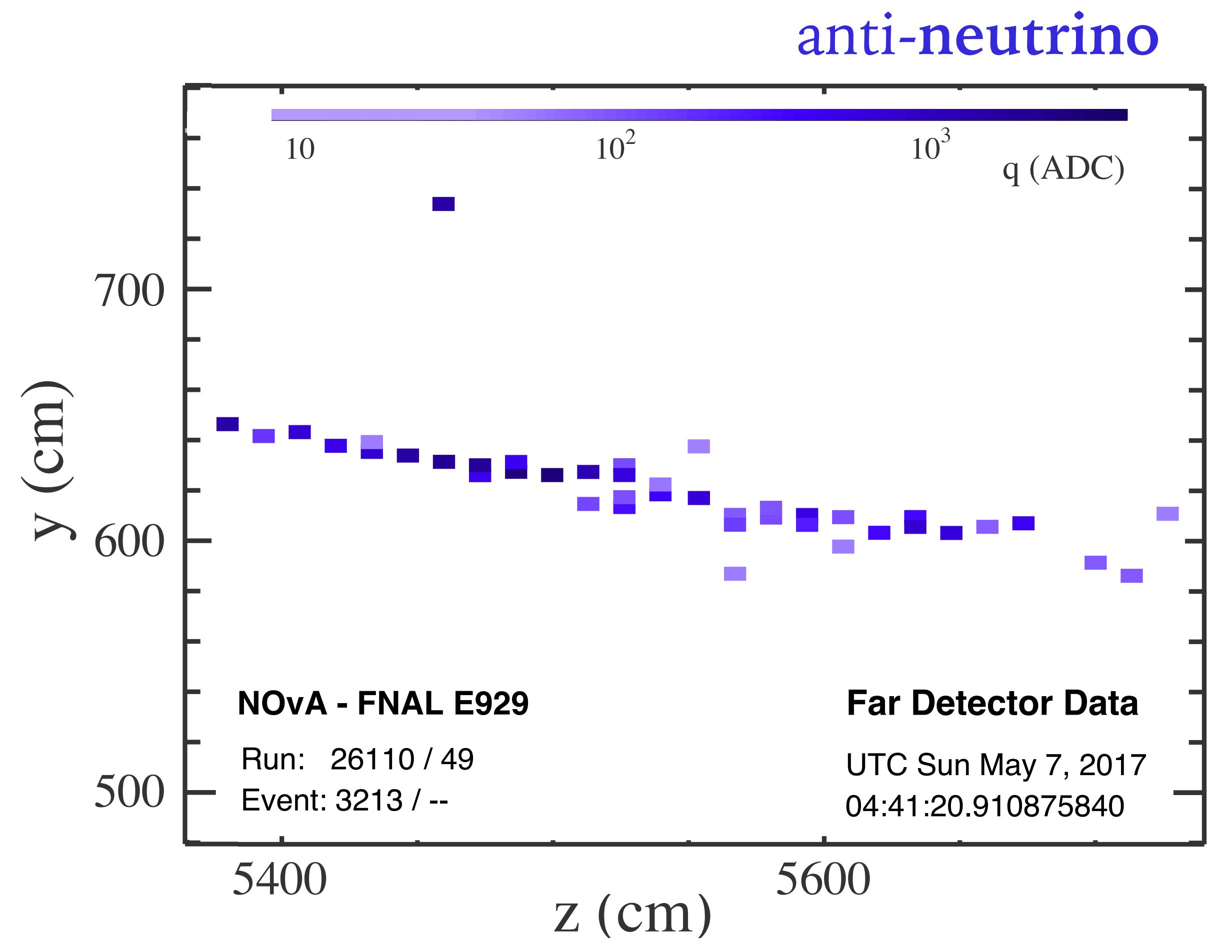


The features are learnt by the network.

Disentangle from traditional reconstruction.

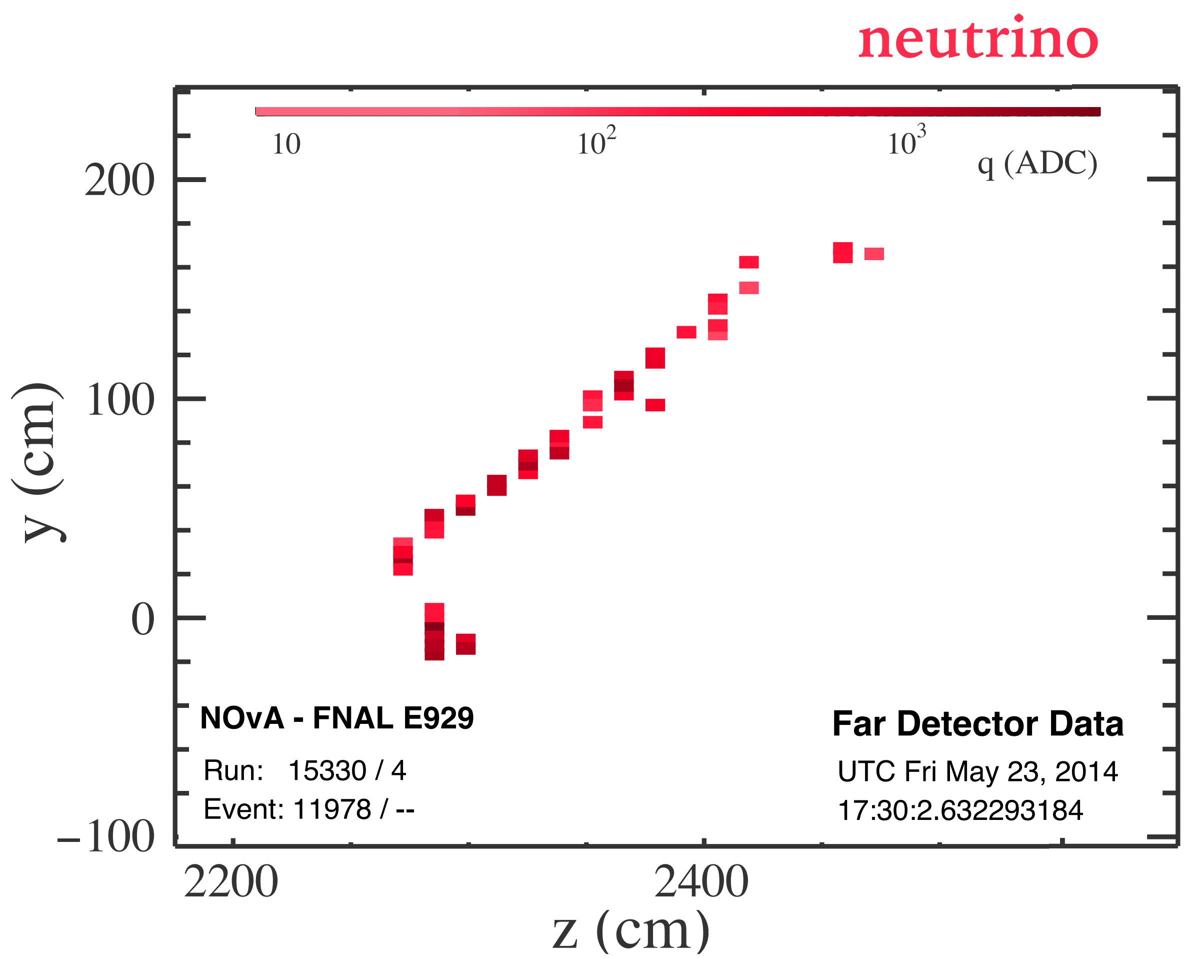
Event Identification

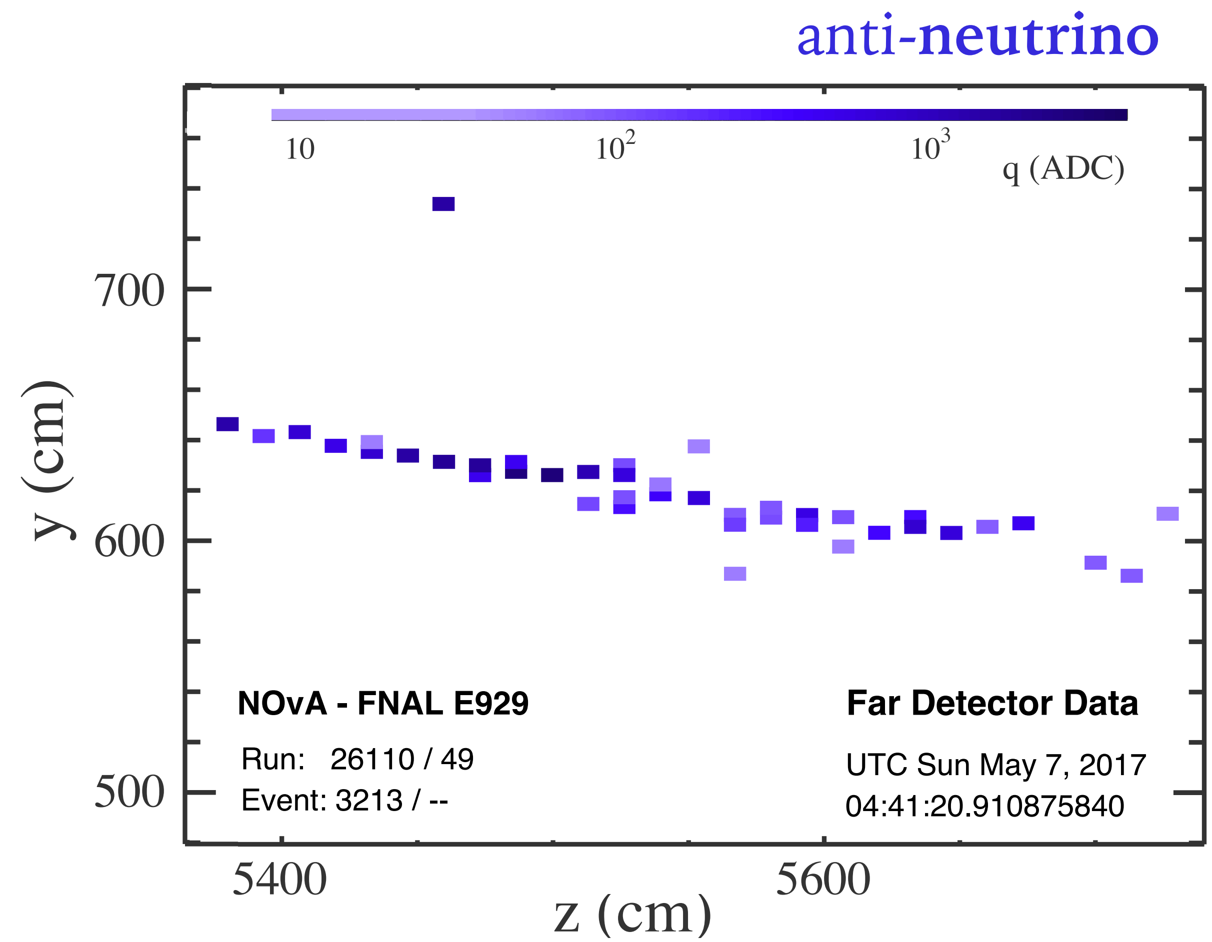




The topology of **neutrino** and **anti-neutrino** interactions is different on average.

Event Identification

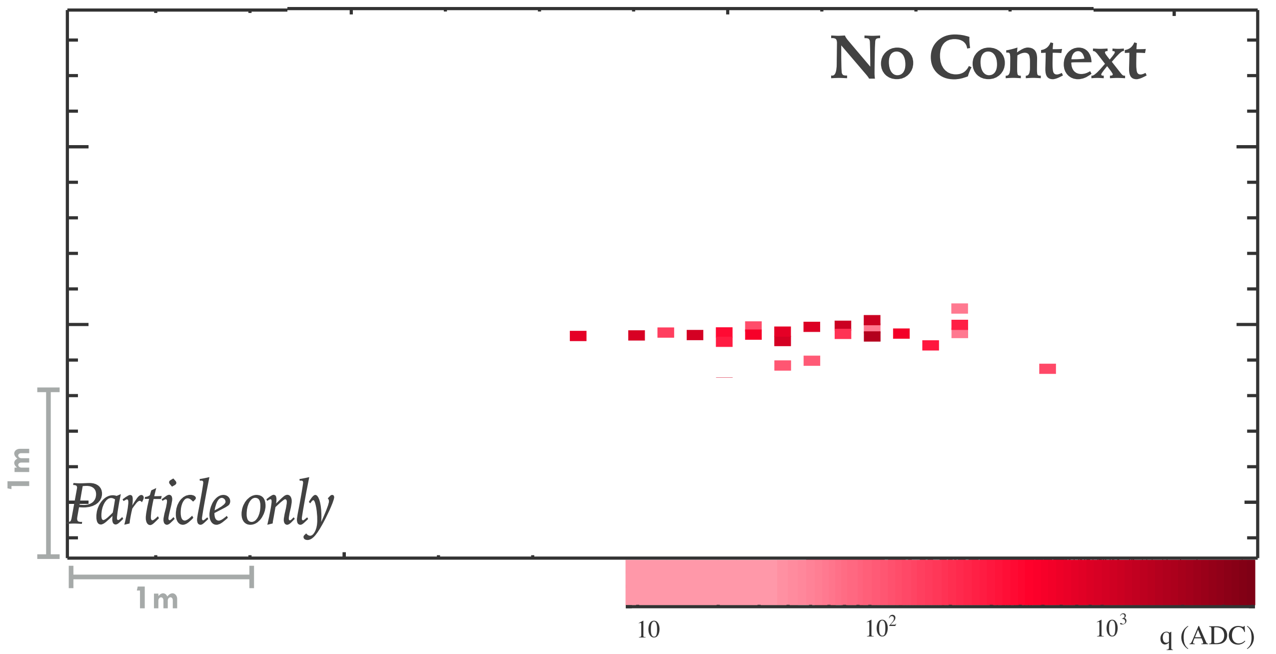


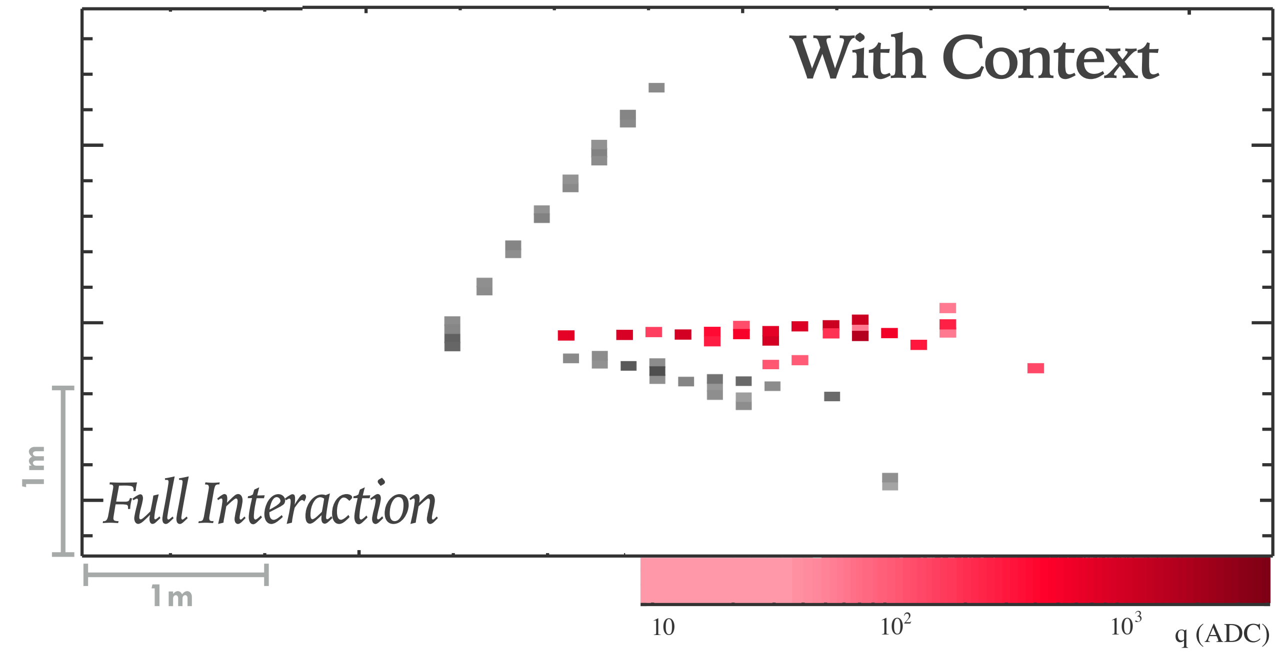


Train on **neutrino** beam and **anti-neutrino** beam simulations separately.

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%





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

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, and A. Sousa
Phys. Rev. D **100**, 073005 – Published 14 October 2019

Think about the ***physicists***

What frameworks are available with:

- reasonable **learning curves**
- **applicability** to other tasks/career paths
- A path to plug-in to your **workflow**.

How do you **train** an army of deep learners?

OLD Framework

Standard art (root) File
- Run pre-processing jobs

Root to LevelDB conversion
- Inputs to caffe. Create DBs and transfer to GPUs

Train Network
- Train plus simple loss-accuracy validation

Physics validation
- Re-run reco on dedicated samples (caffe c++ api) CPU

Analysis validation
- Produce analysis files and full-chain validation

NEW Framework

HDF5 File
- Now produced by the standard Reco. Chain
- Train Network on GPU (Day)

Train network
- Train plus simple loss-accuracy
- Tensorflow infrastructure + physics information

Fully trained and validated network
- Physics and analysis validation
- All work done on GPUs

~month

~weeks

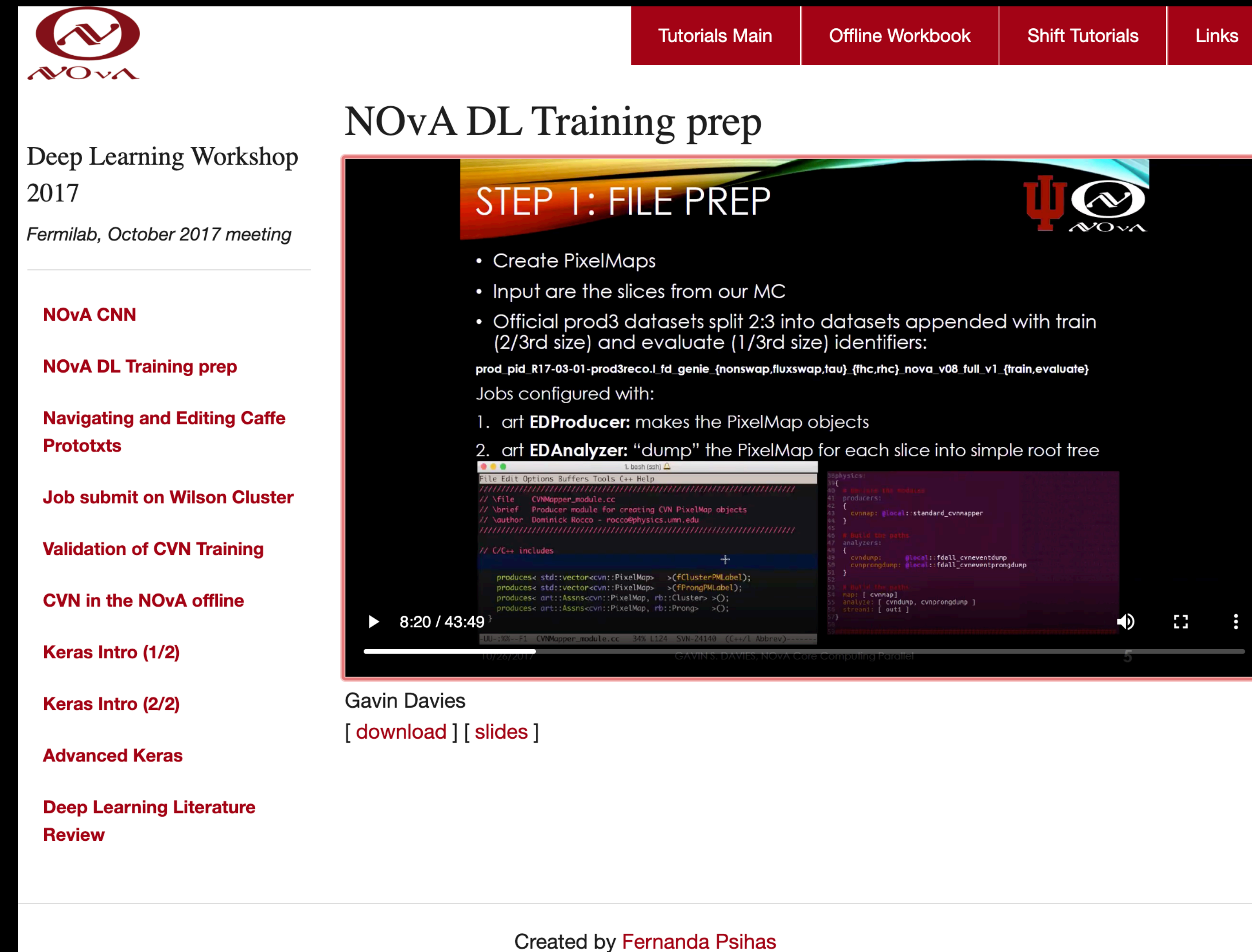
~days

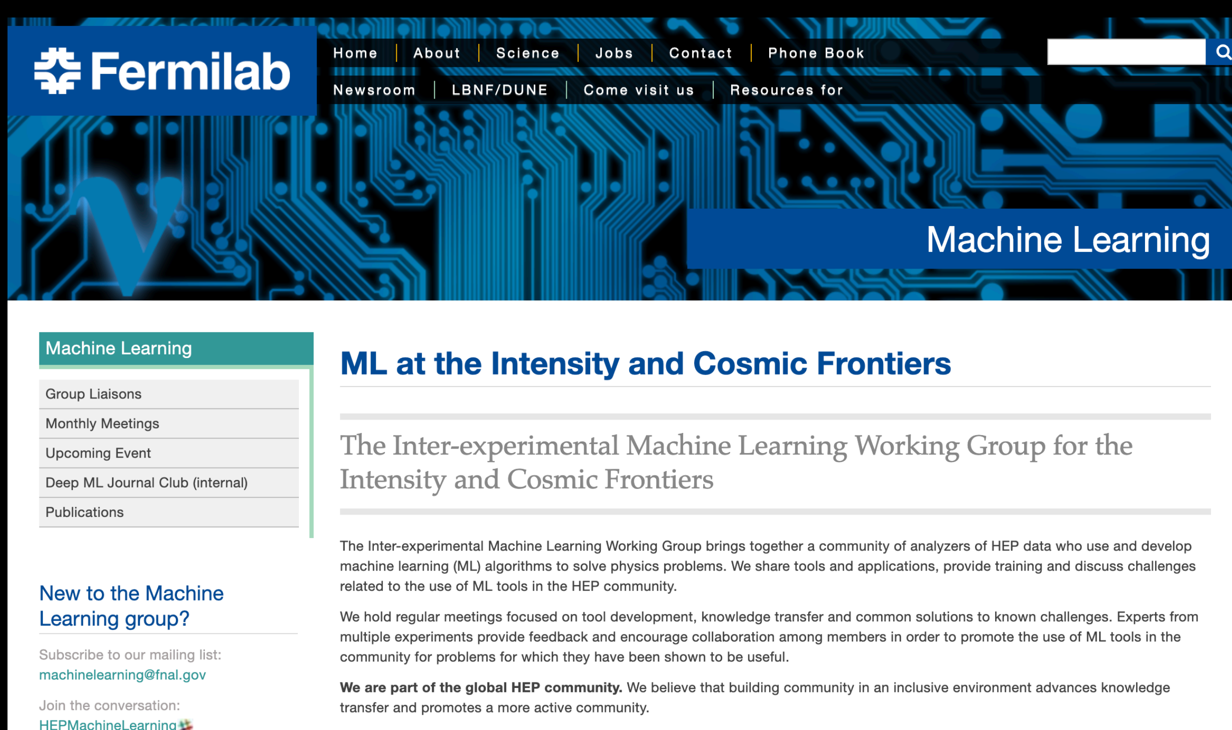
TRAINING In-house

Targeted workshops

Organized by young NOvA

Target specific tasks and experiment specific tasks





CROSS-Experiment

FNAL ML group is a community knowledge-sharing group

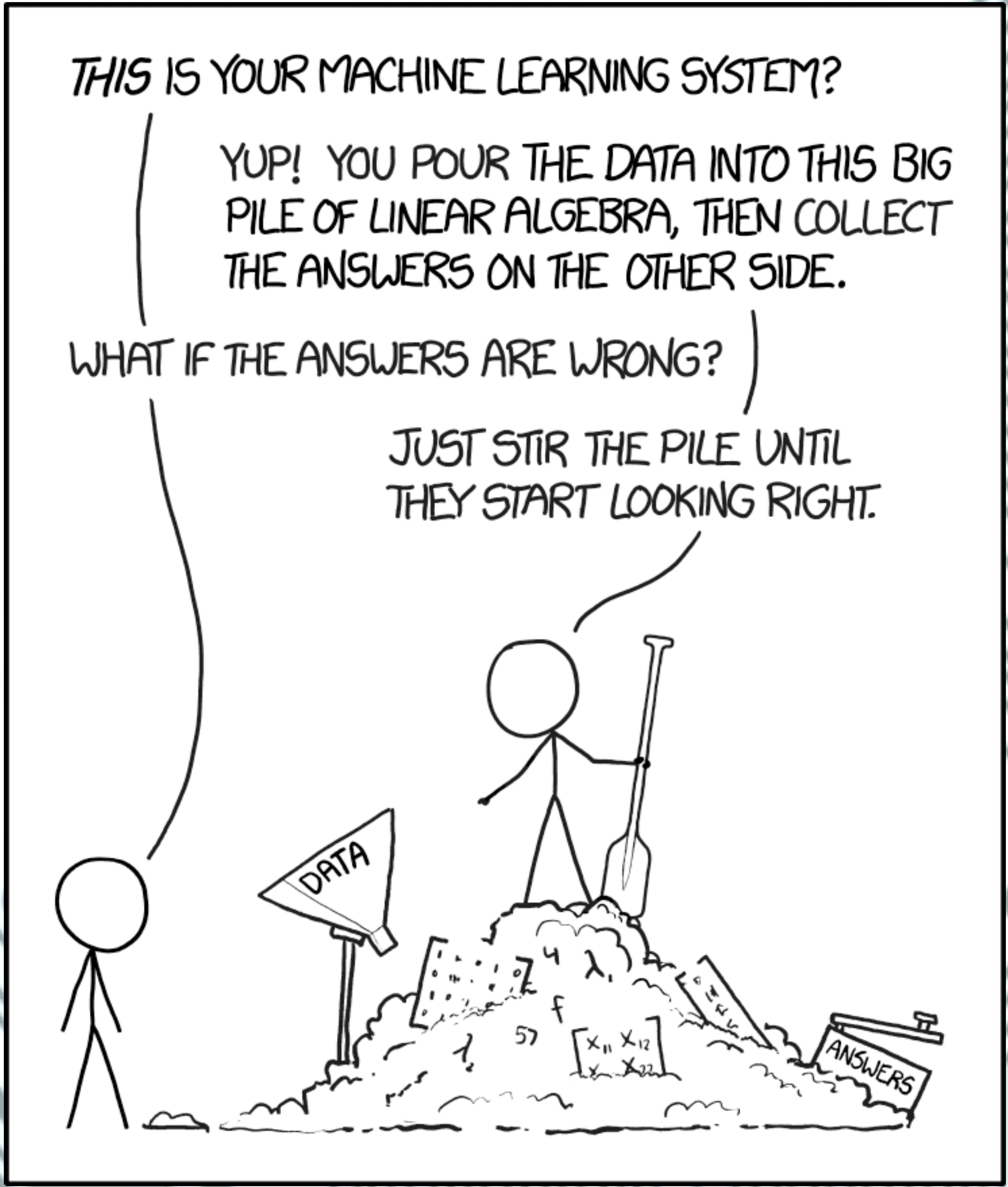
ML experts across experiments discuss technical details of their applications

Think about the problems

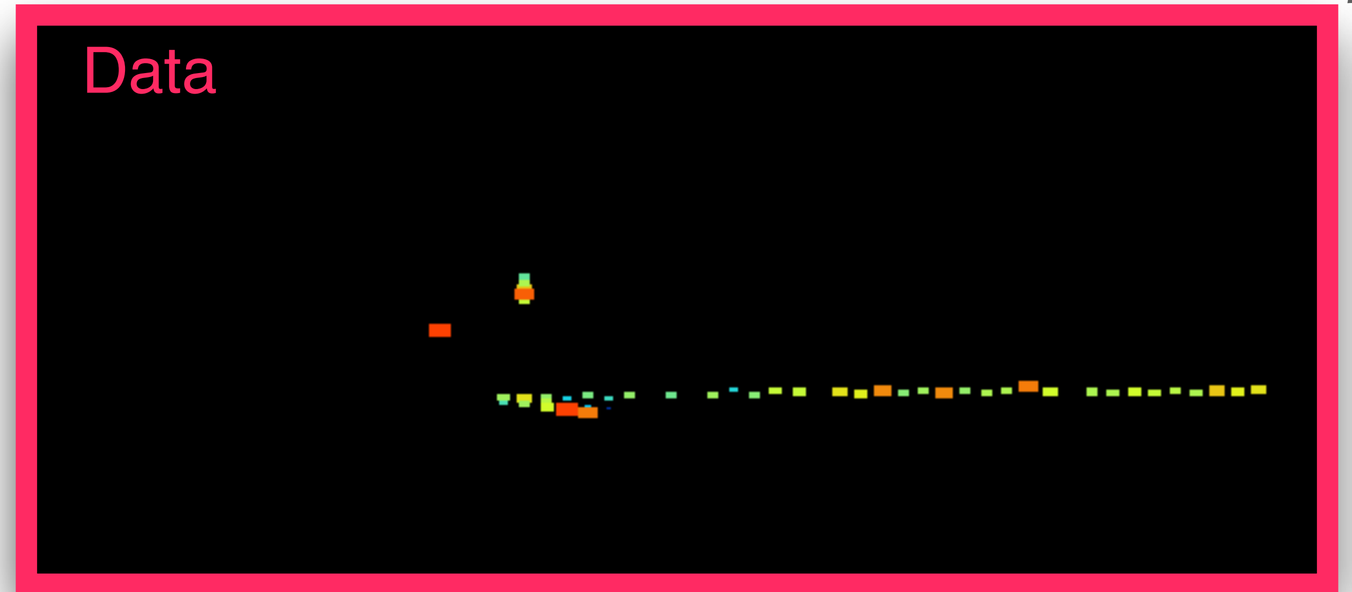
Bias introduced from MC training.

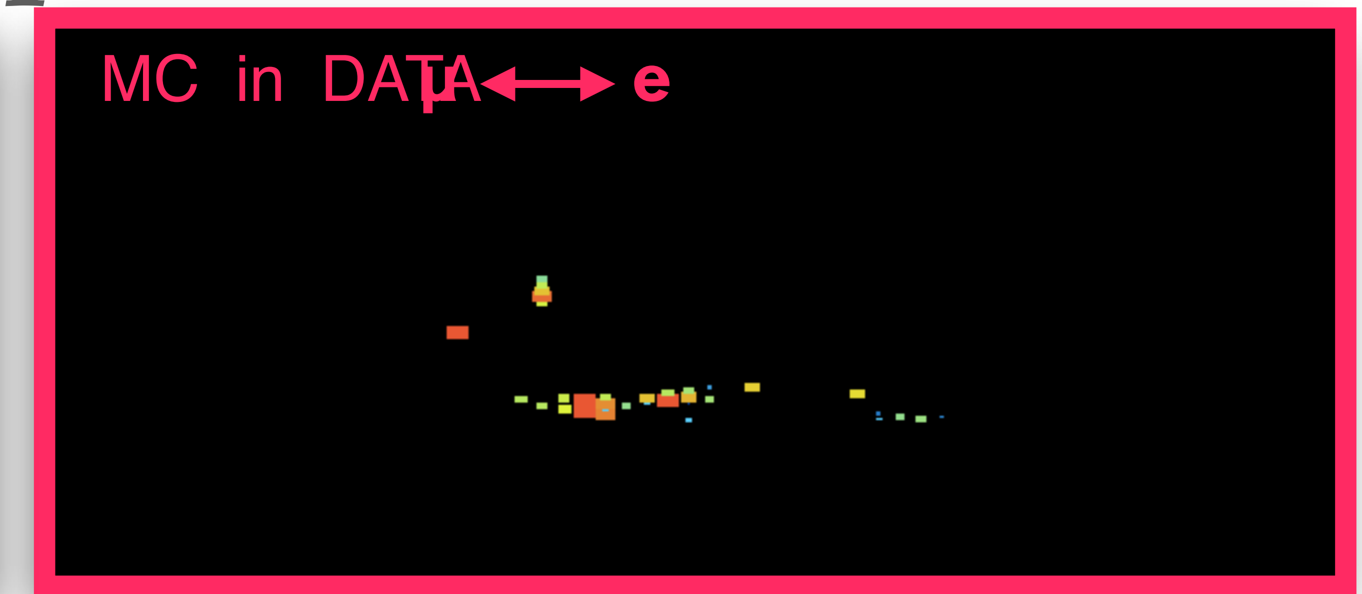
Mis-modeling translating into data/MC differences

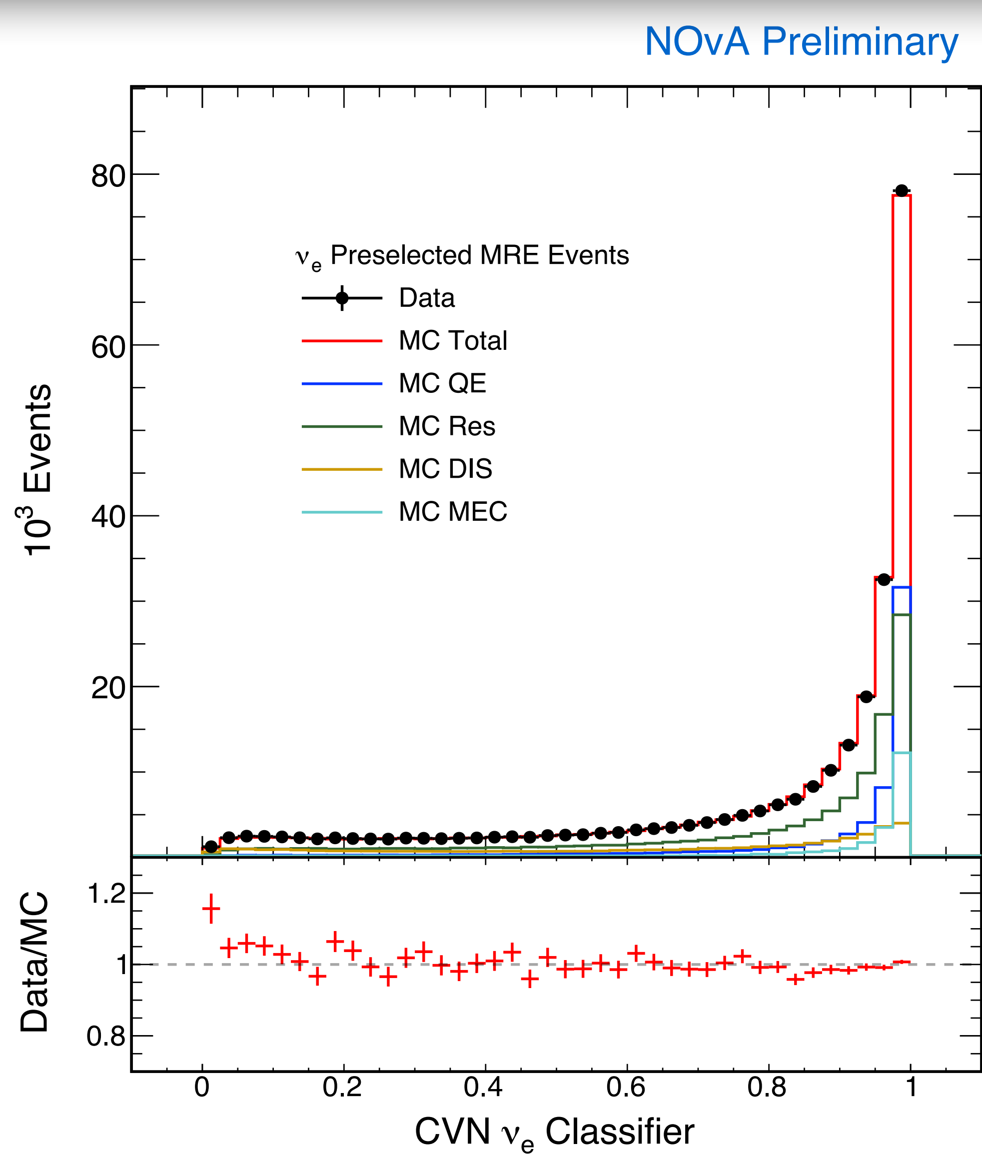
Uncertainties and other practical considerations



Data-driven test example







MRE (Muon Removed - Electron):

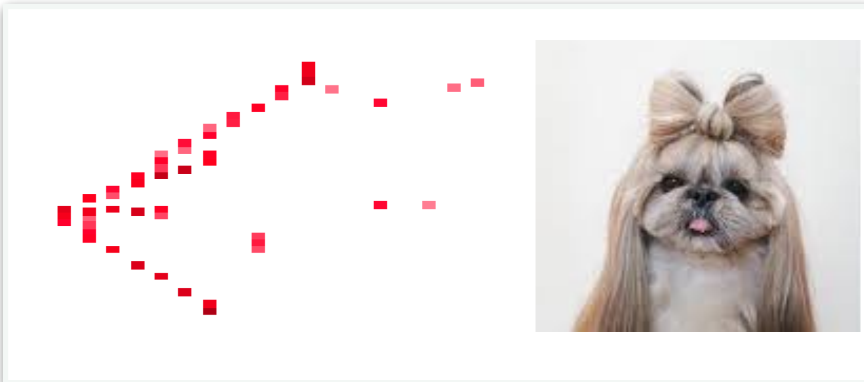
Select a muon neutrino interaction with traditional ID methods.

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

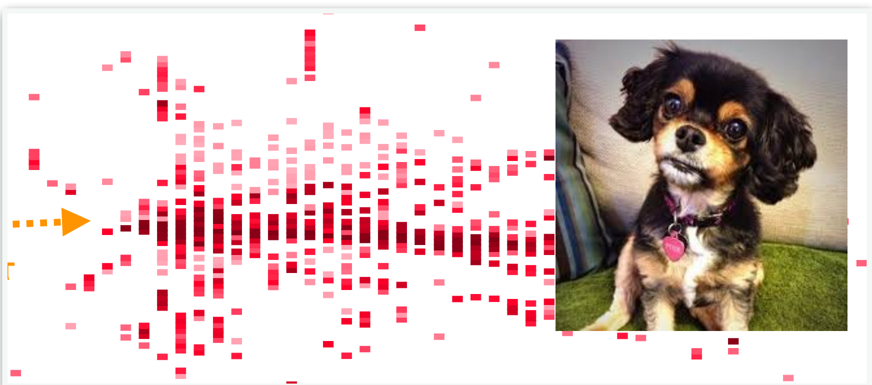
Data/MC comparisons show less than 1% difference in efficiency.

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

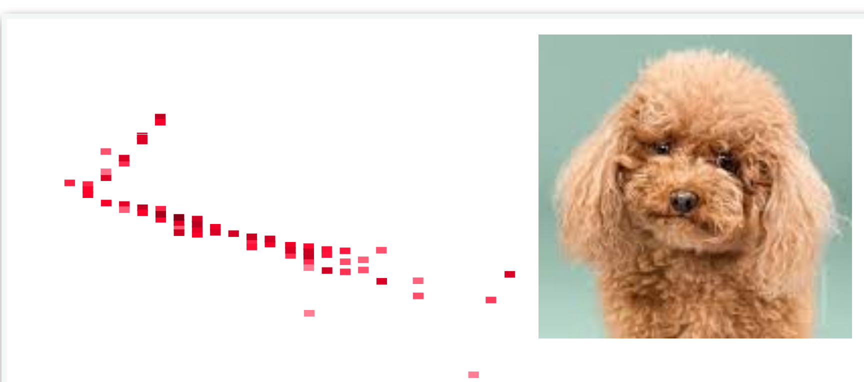
Uncertainty & Bias assessment



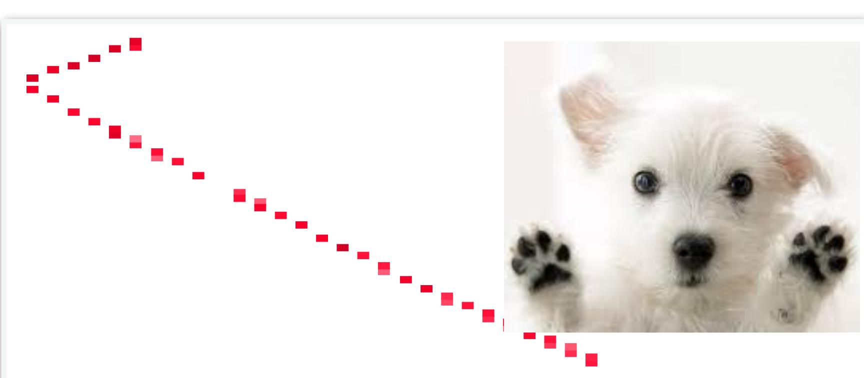
We currently apply the same systematic treatment through the whole analysis, including our DL algorithms.



Data-driven tests are an essential tool to show data/MC robustness & assess some of the potential biases.



NOvA test beam running through 2019-2020 will provide us labeled particle data for training.



We are exploring adversarial training in to reduce the impact of detector uncertainties in the training.

Thank you.

Backup

death star

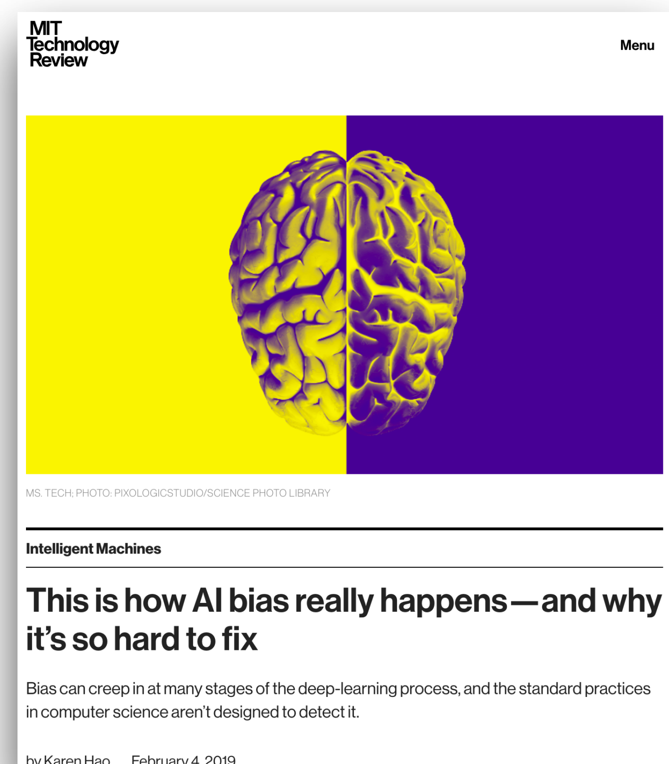
These are not the slides you’re looking for

NOvA Test Beam

The NOvA test beam detector is being commissioned and will be taking data this year.

With library of labeled data for single particles of known **identity** and **momentum**, NOvA will expand th data-driven checks of our deep learning algorithms.

How is your model biased?



MIT Technology Review

Menu

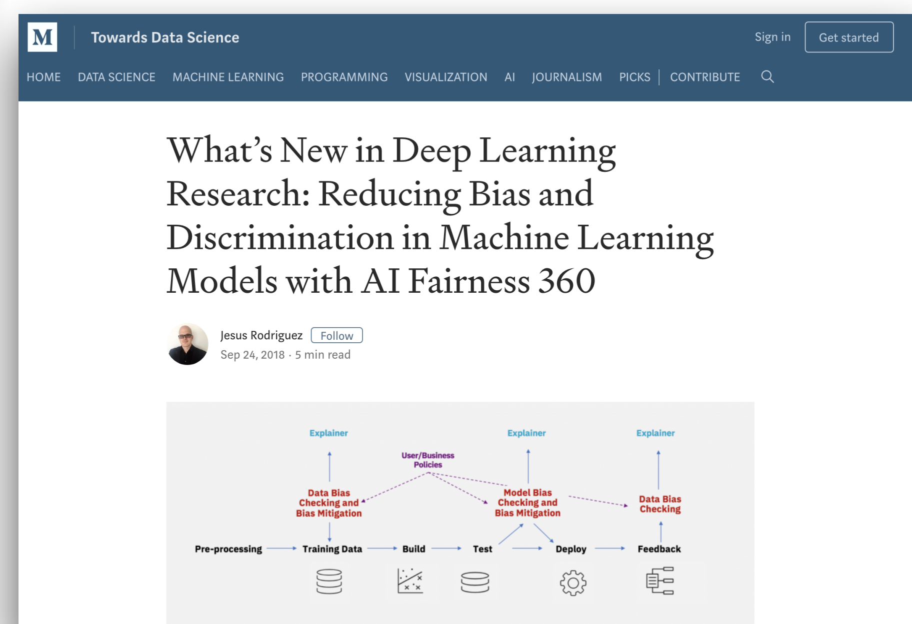
MS. TECH; PHOTO: PIXOLOGICSTUDIO/SCIENCE PHOTO LIBRARY

Intelligent Machines

This is how AI bias really happens—and why it's so hard to fix

Bias can creep in at many stages of the deep-learning process, and the standard practices in computer science aren't designed to detect it.

by Karen Hao February 4, 2019



Towards Data Science

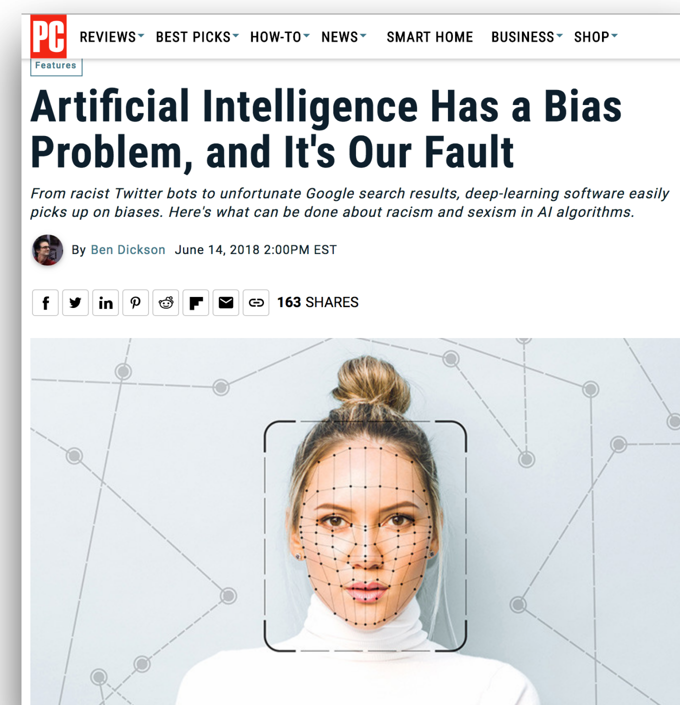
Sign in Get started

HOME DATA SCIENCE MACHINE LEARNING PROGRAMMING VISUALIZATION AI JOURNALISM PICKS CONTRIBUTE

What's New in Deep Learning Research: Reducing Bias and Discrimination in Machine Learning Models with AI Fairness 360

Jesus Rodriguez Follow

Sep 24, 2018 · 5 min read



PC

REVIEWS BEST PICKS HOW-TO NEWS SMART HOME BUSINESS SHOP

Features

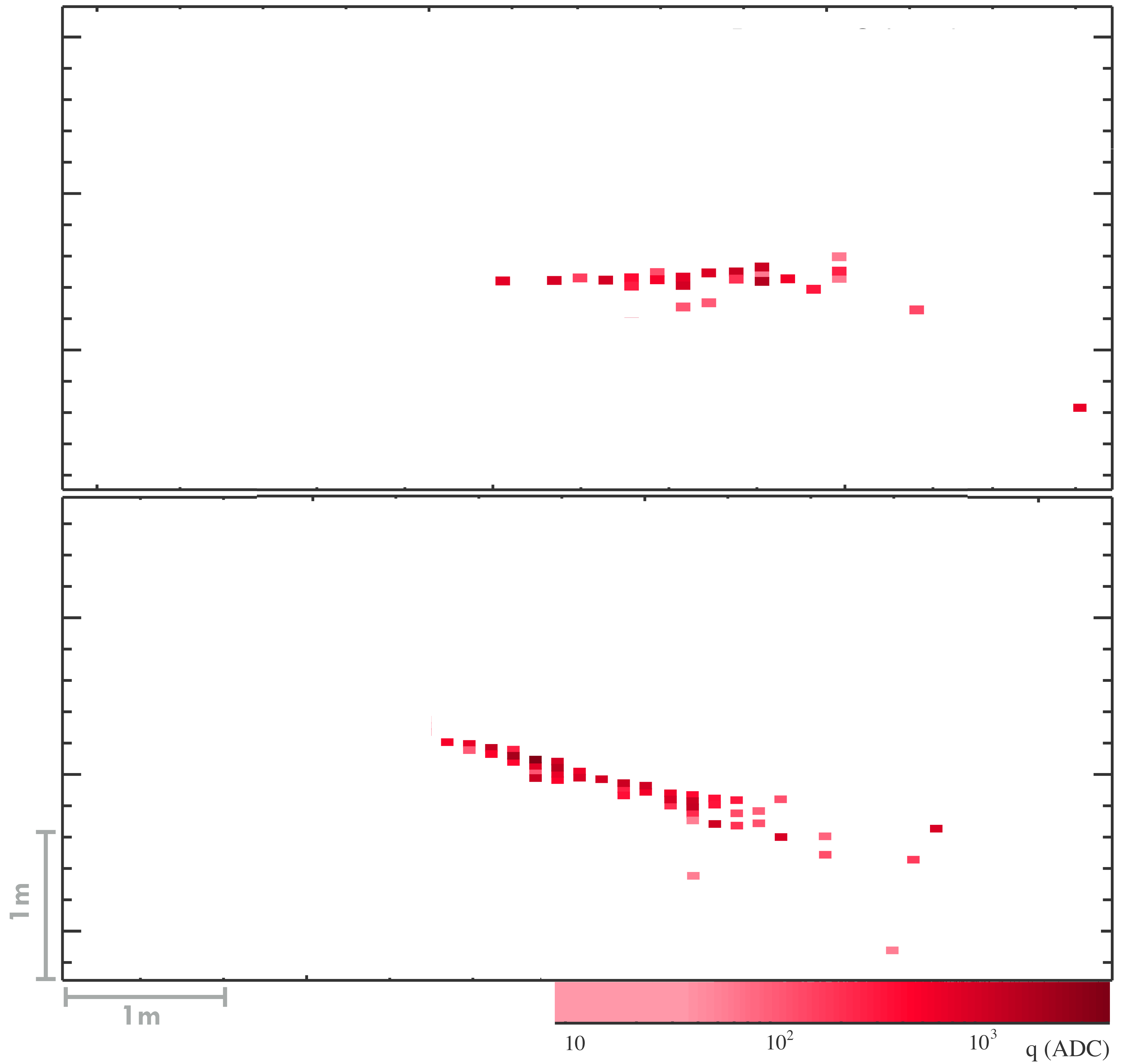
Artificial Intelligence Has a Bias Problem, and It's Our Fault

From racist Twitter bots to unfortunate Google search results, deep-learning software easily picks up on biases. Here's what can be done about racism and sexism in AI algorithms.

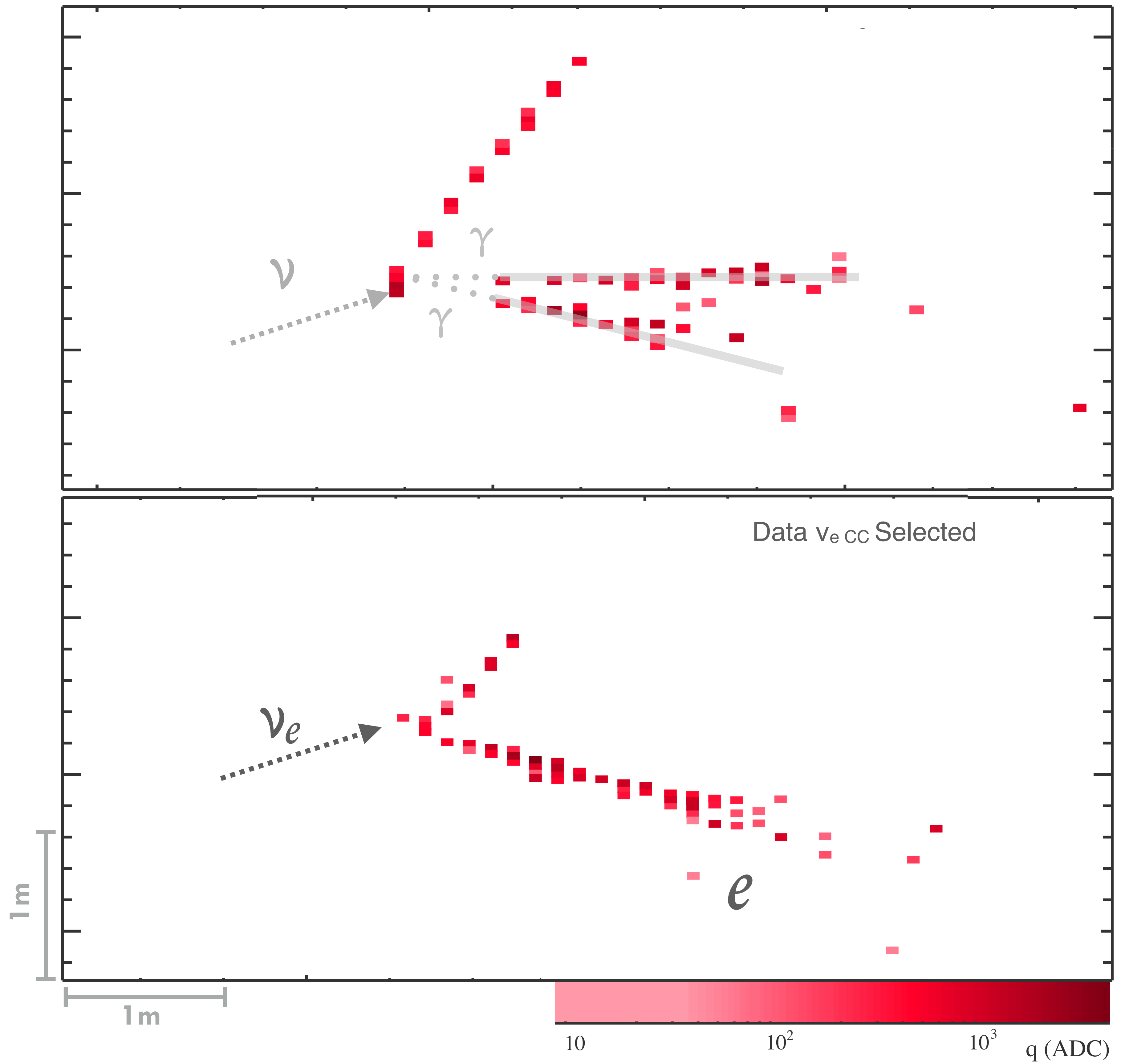
By Ben Dickson June 14, 2018 2:00PM EST

163 SHARES

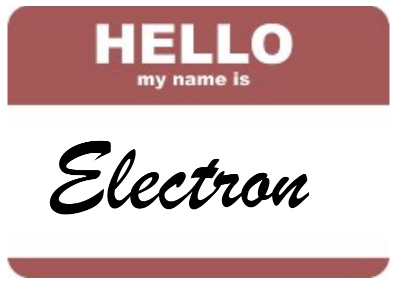
Particle Identification

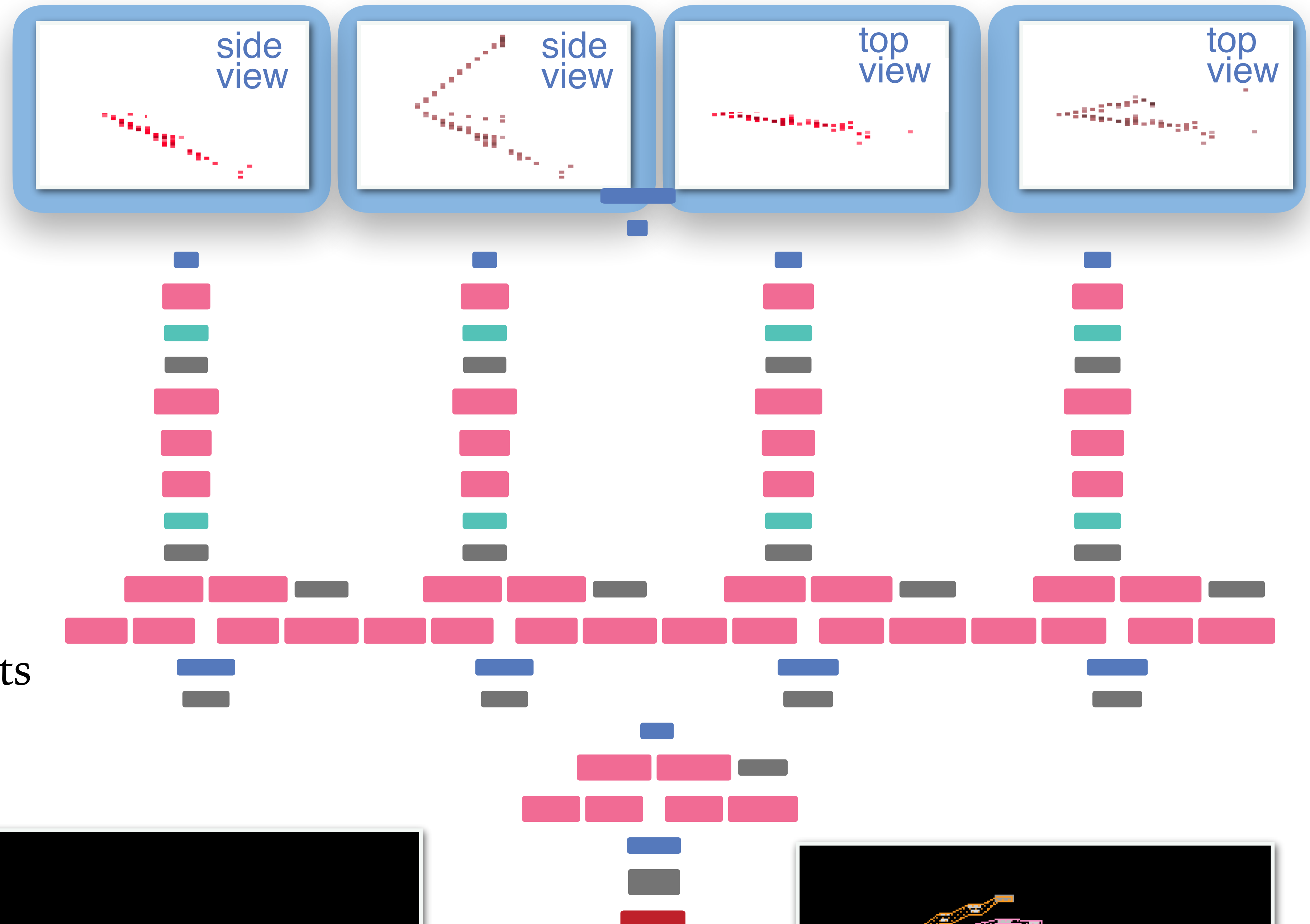


Particle Identification



Context Improves Classification

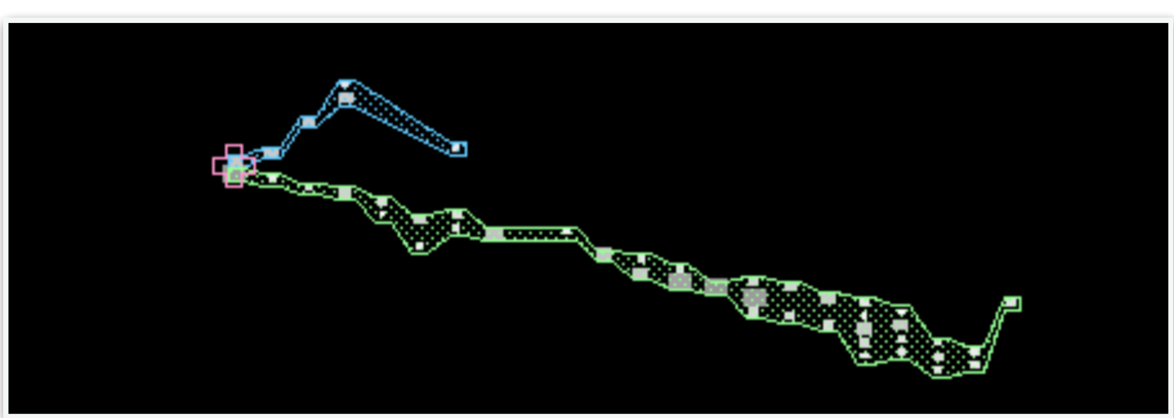


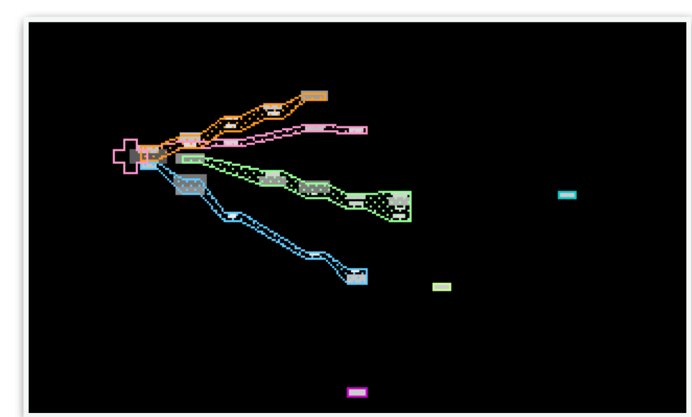


Using the existing reconstruction, classify clusters of hits.

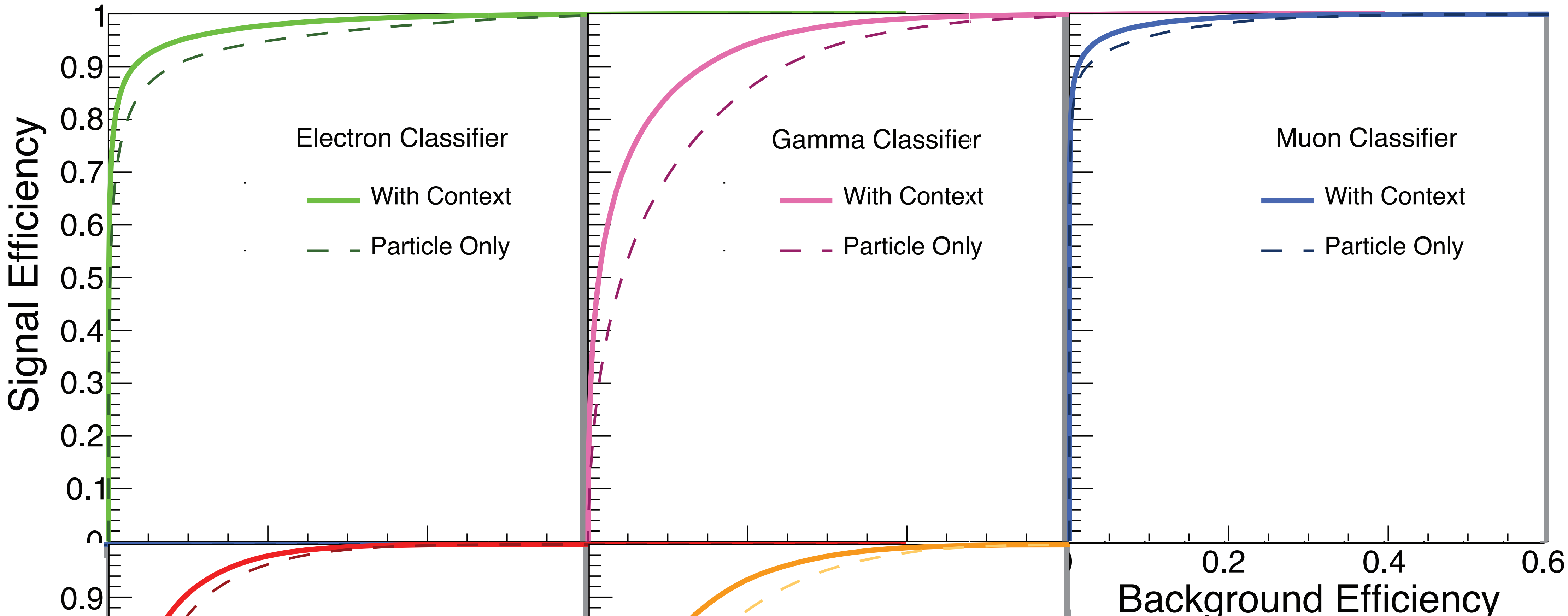
Modified to take 4 views (event + clustered particle hits)

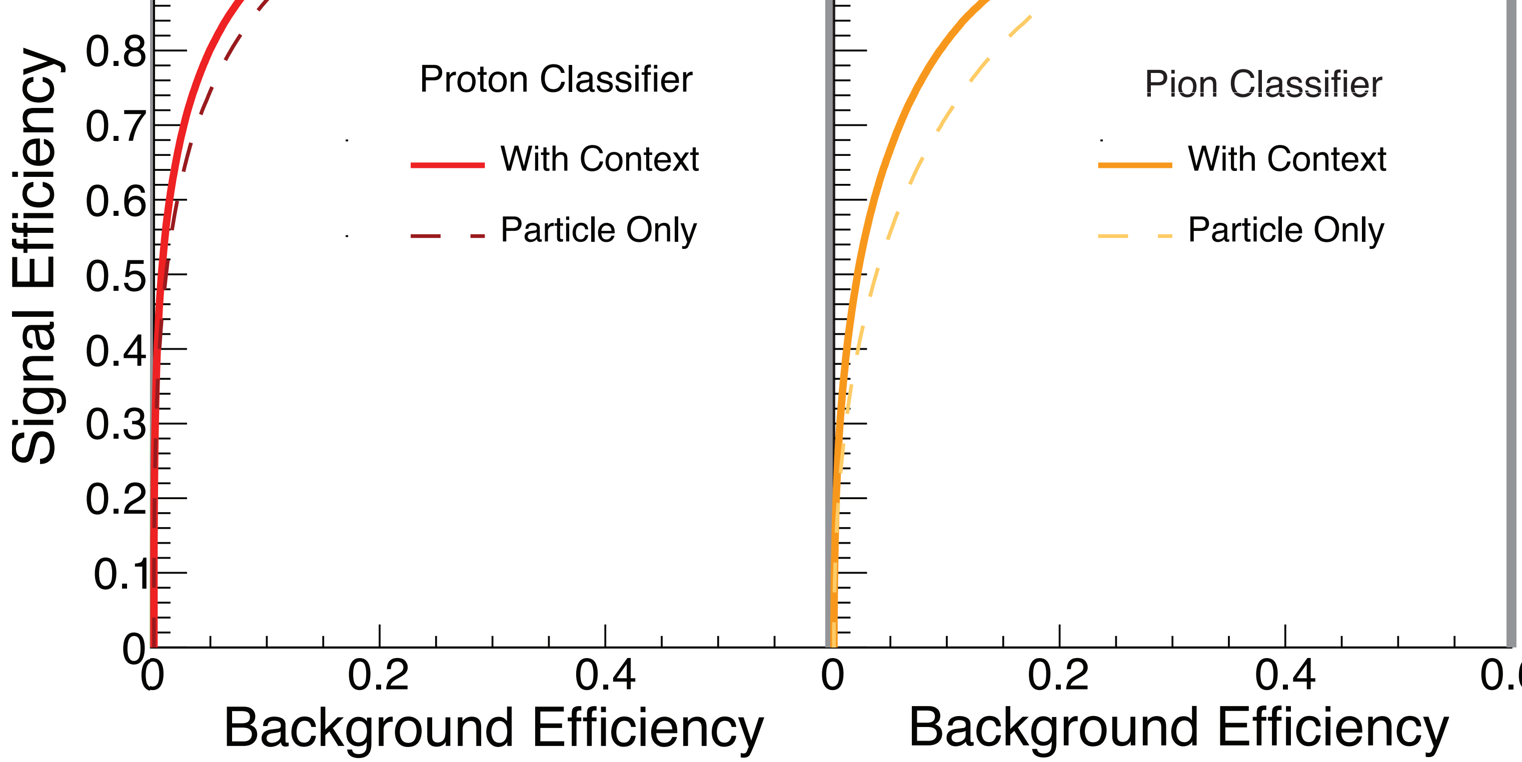
Trained on clusters from all events above some minimum purity.

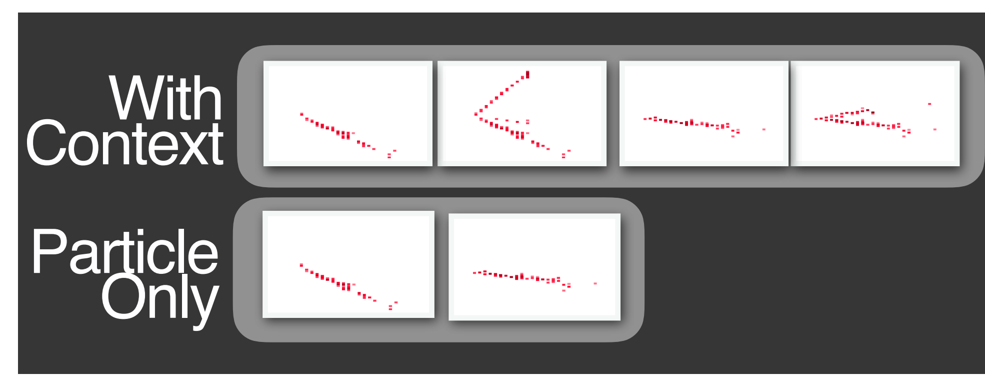




Context Improves Classification

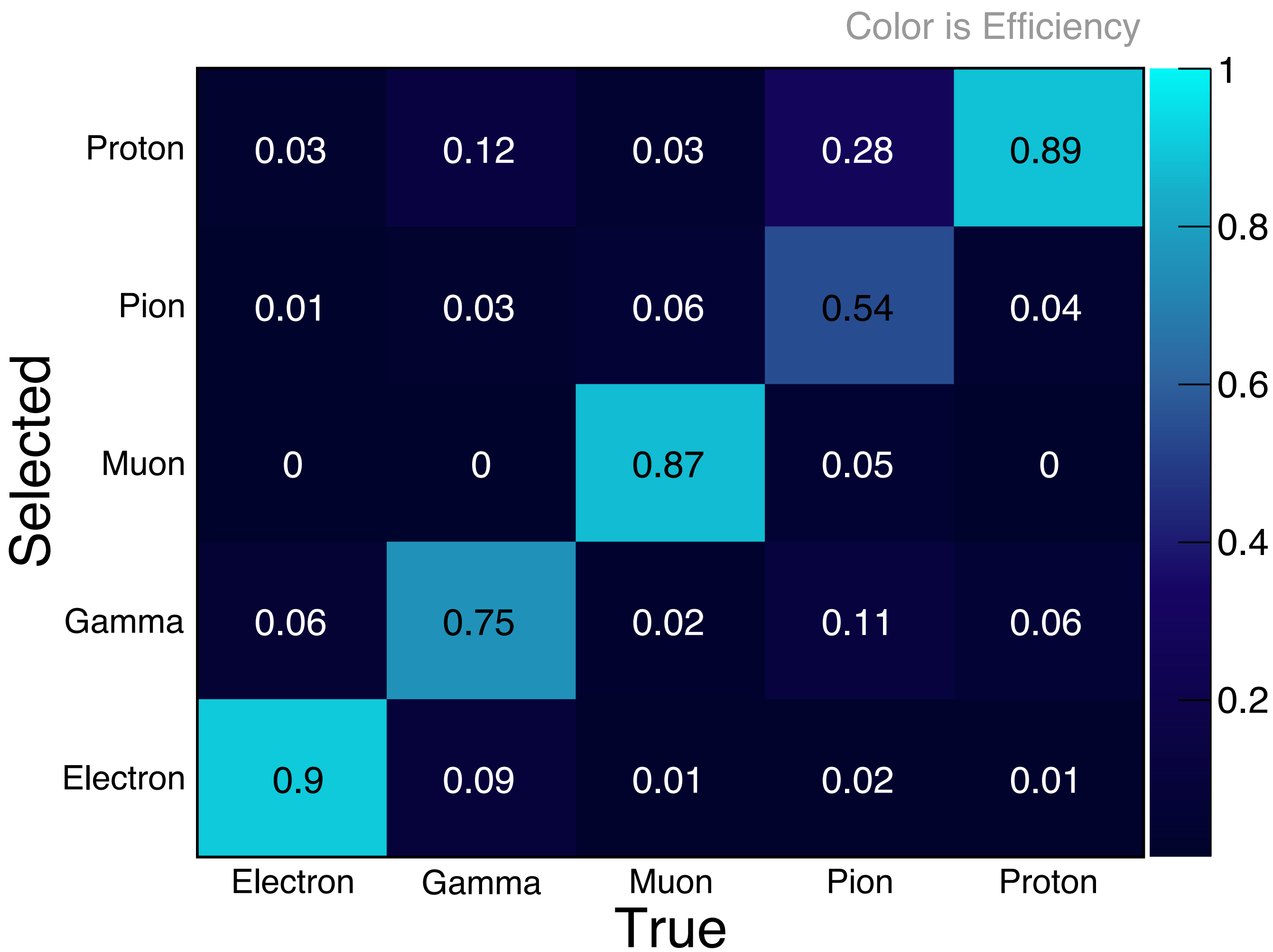


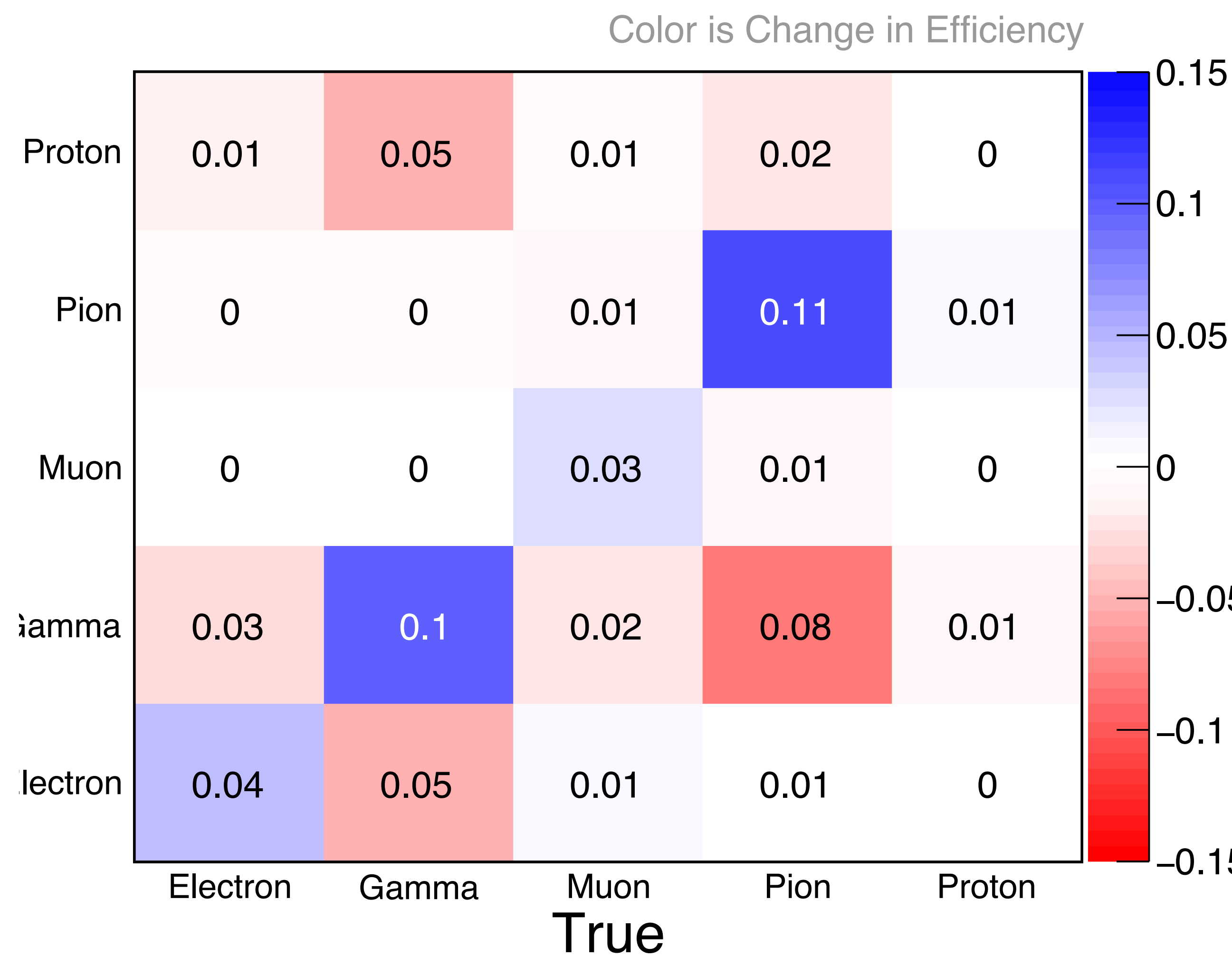




Up to 11% improvement from adding context to the classifier.

Context Improves Classification





Color is efficiency taking the highest softmax output.

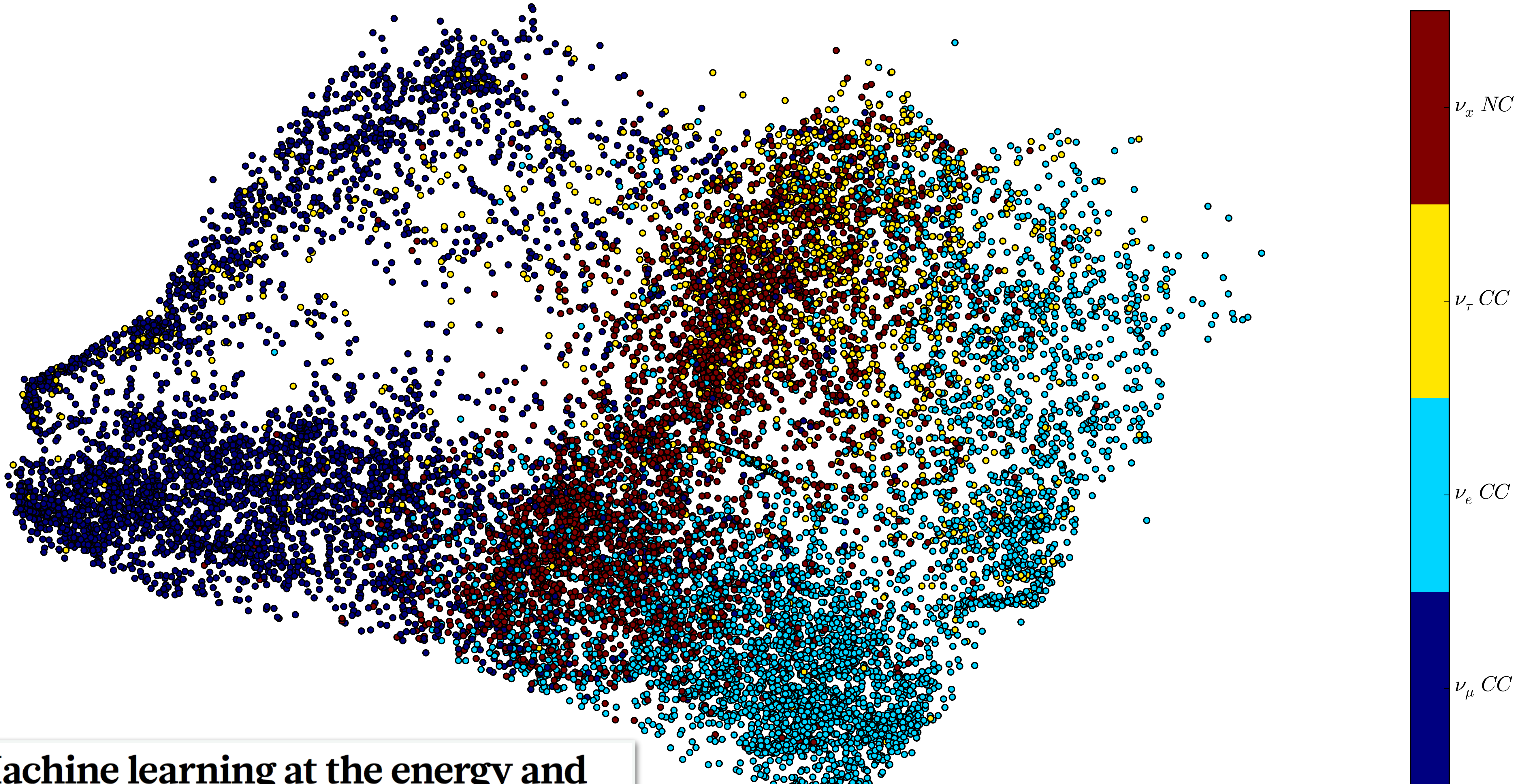
Color is difference in efficiency taking the highest softmax output.

Larger with context

Larger w/o context

t-Distributed Stochastic Neighbor Embedding

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

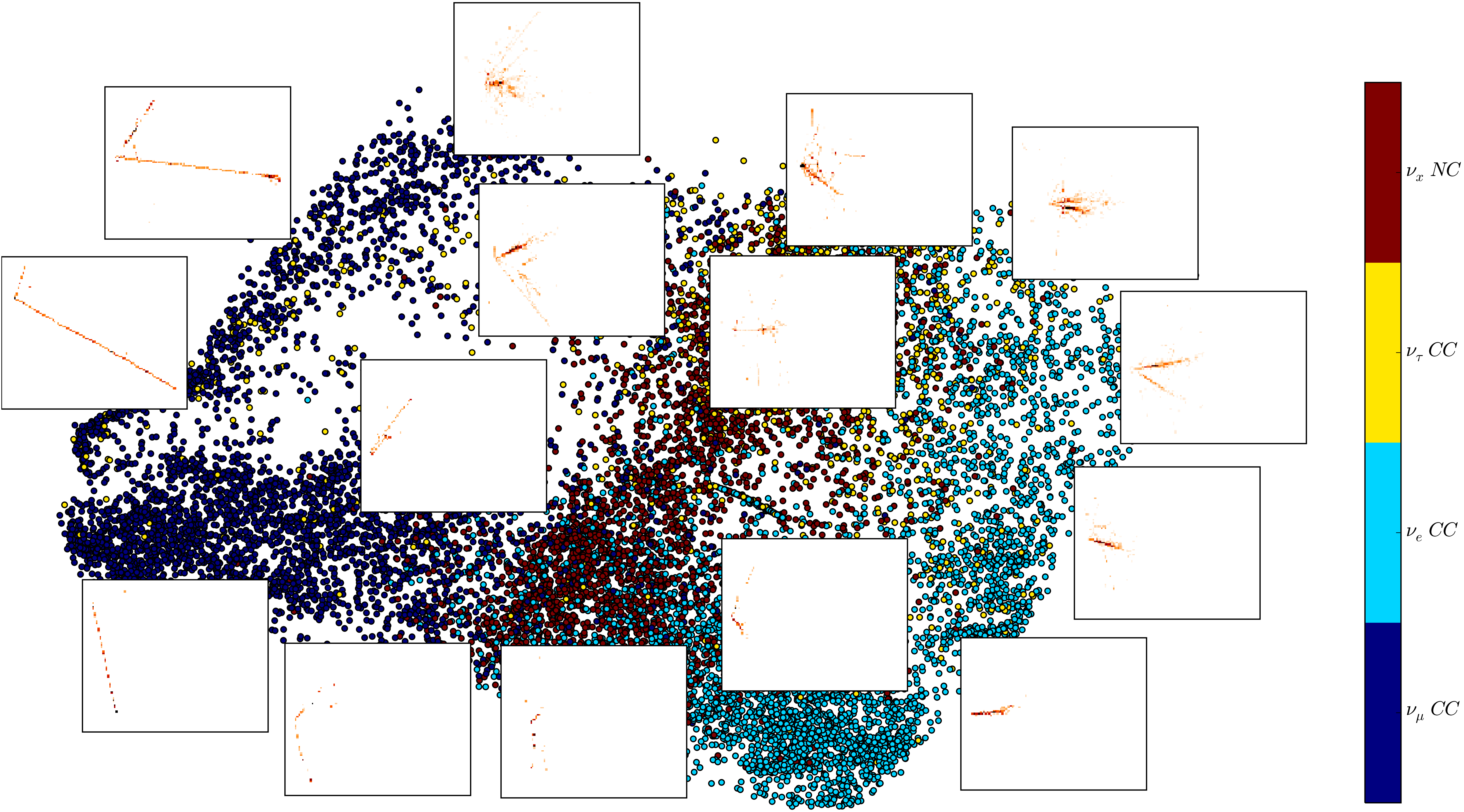


Machine learning at the energy and intensity frontiers of particle physics

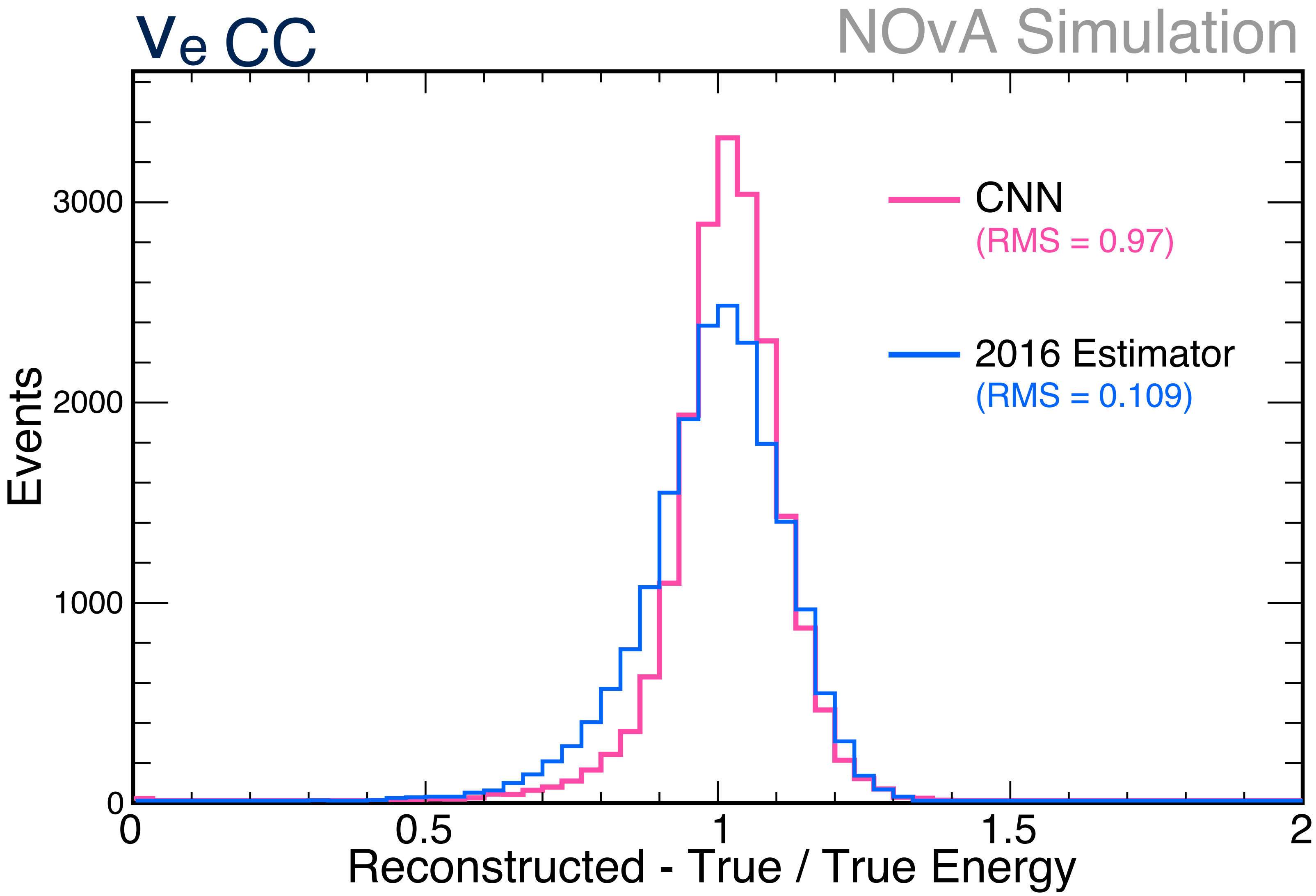
Alexander Radovic[1]*, Mike Williams[2]*, David Rousseau[3], Michael Kagan[4], Daniele Bonacorsi[5,6], Alexander Himmel[7], Adam Aurisano[8], Kazuhiro Terao[4] & Taritree Wongjirad[9]

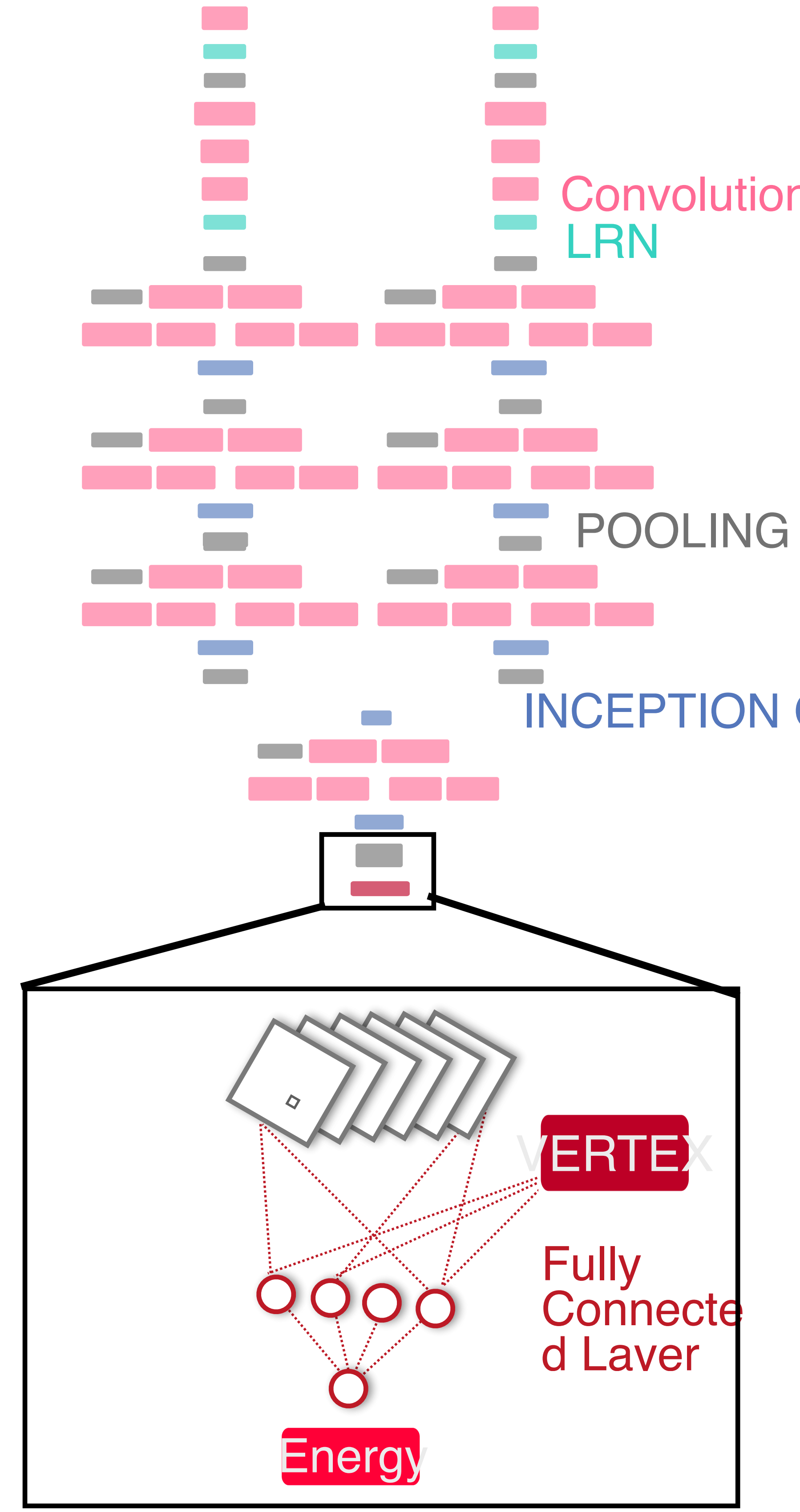
Nature Vol. 560

t-sne



CNNs + Regression for Energy Reconstruction





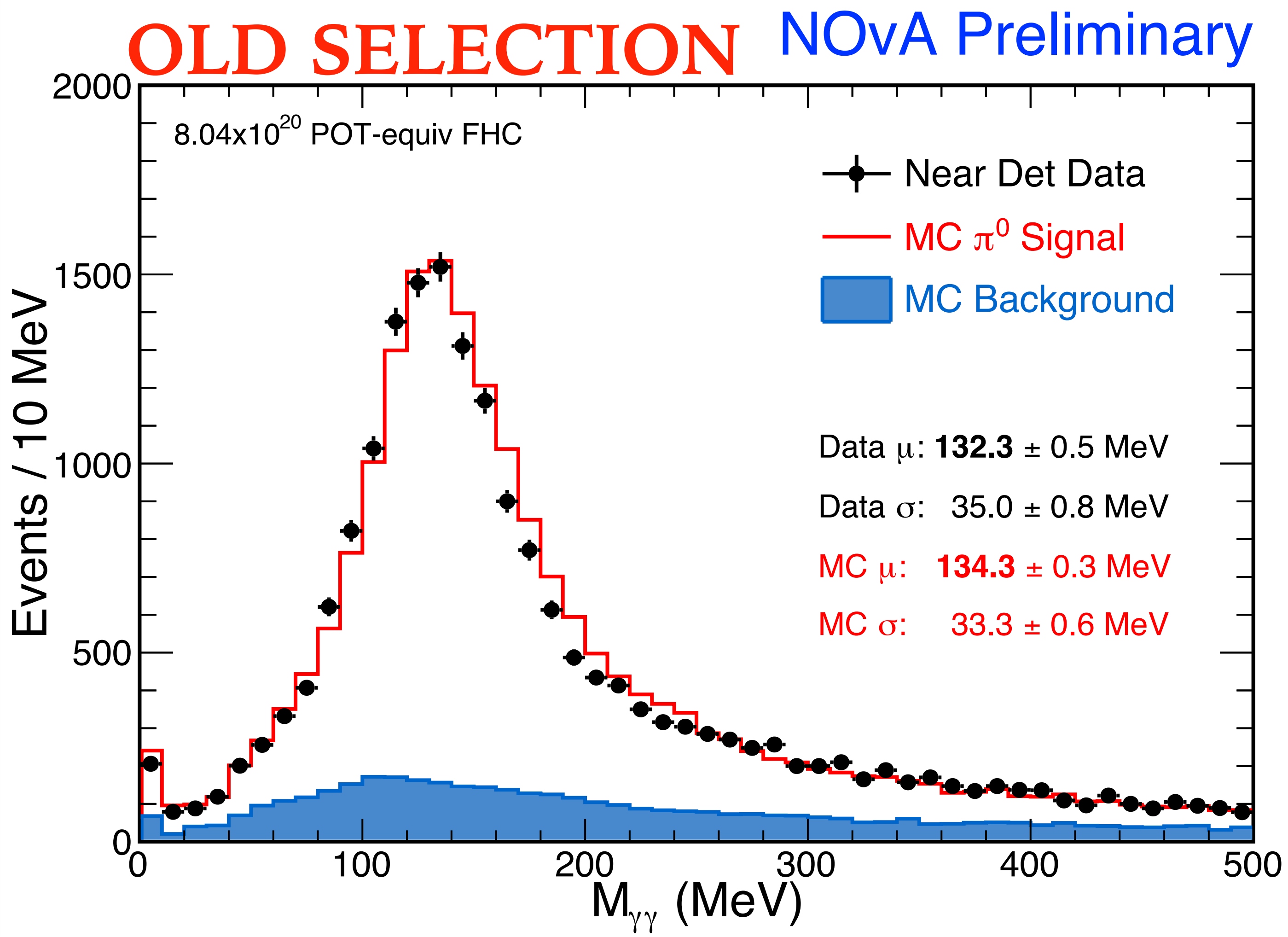
The target is Energy instead of PID values

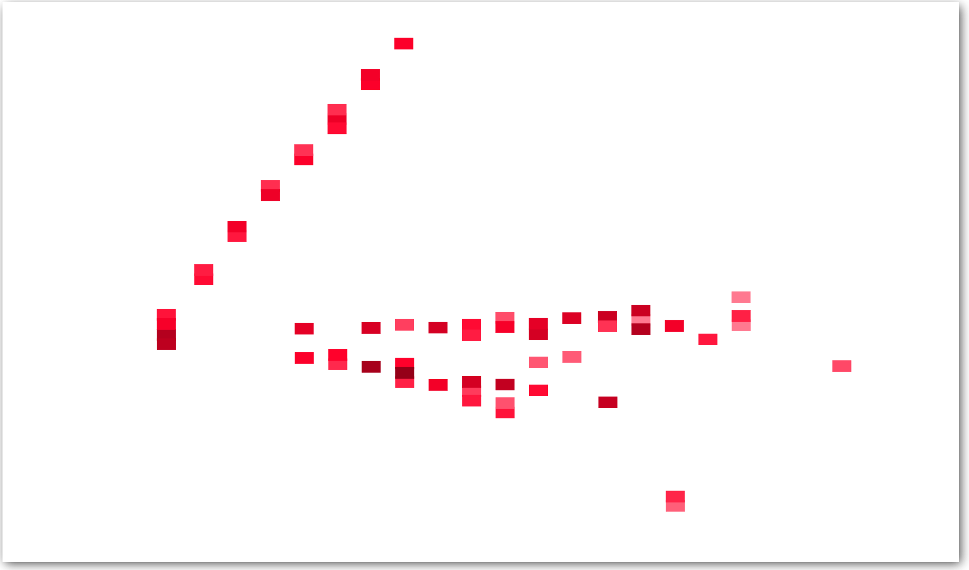
Incorporate information from the reconstruction at the fully connected stage.

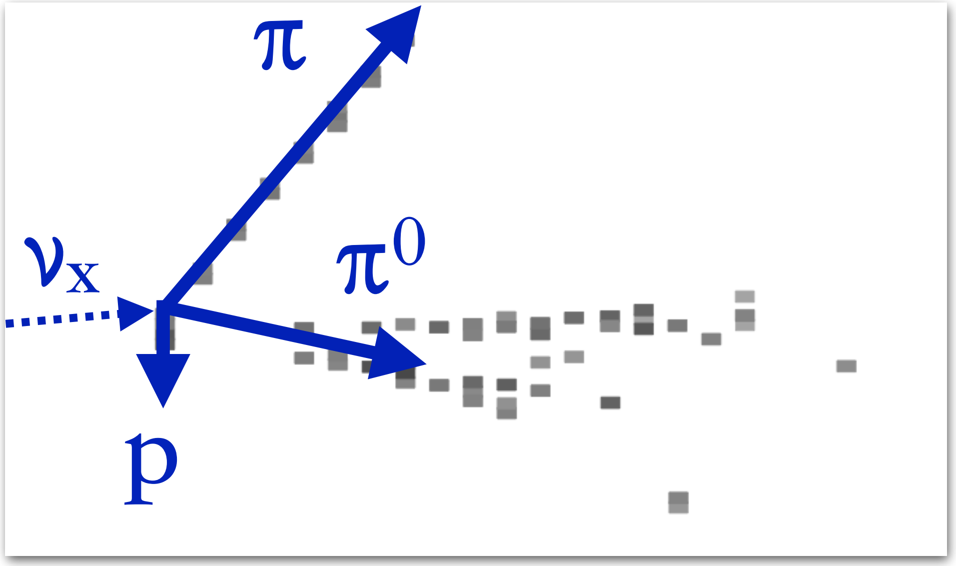
Physics measurements with data

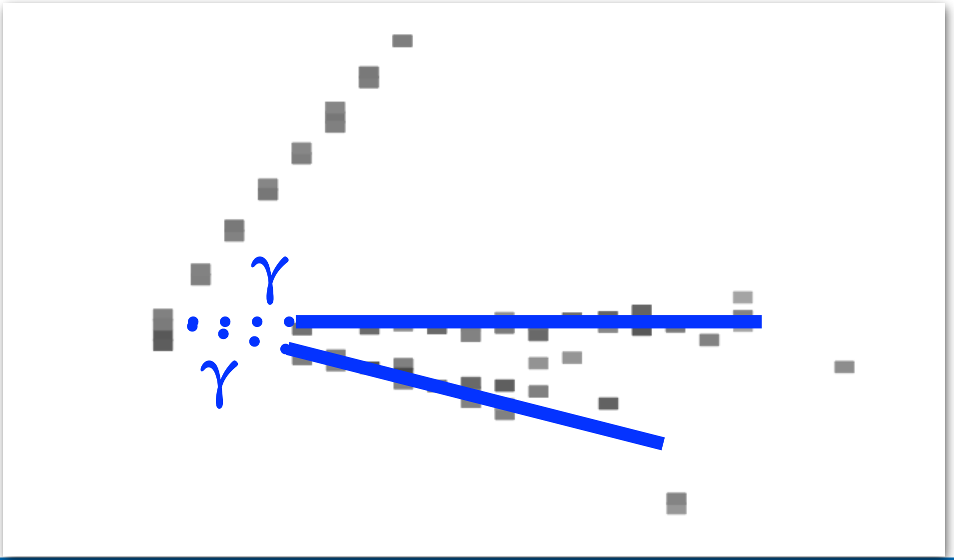
Reconstruct the well known π^0 mass from well identified pairs of photons.

This is a handle on the dead material correction for electromagnetic deposits.





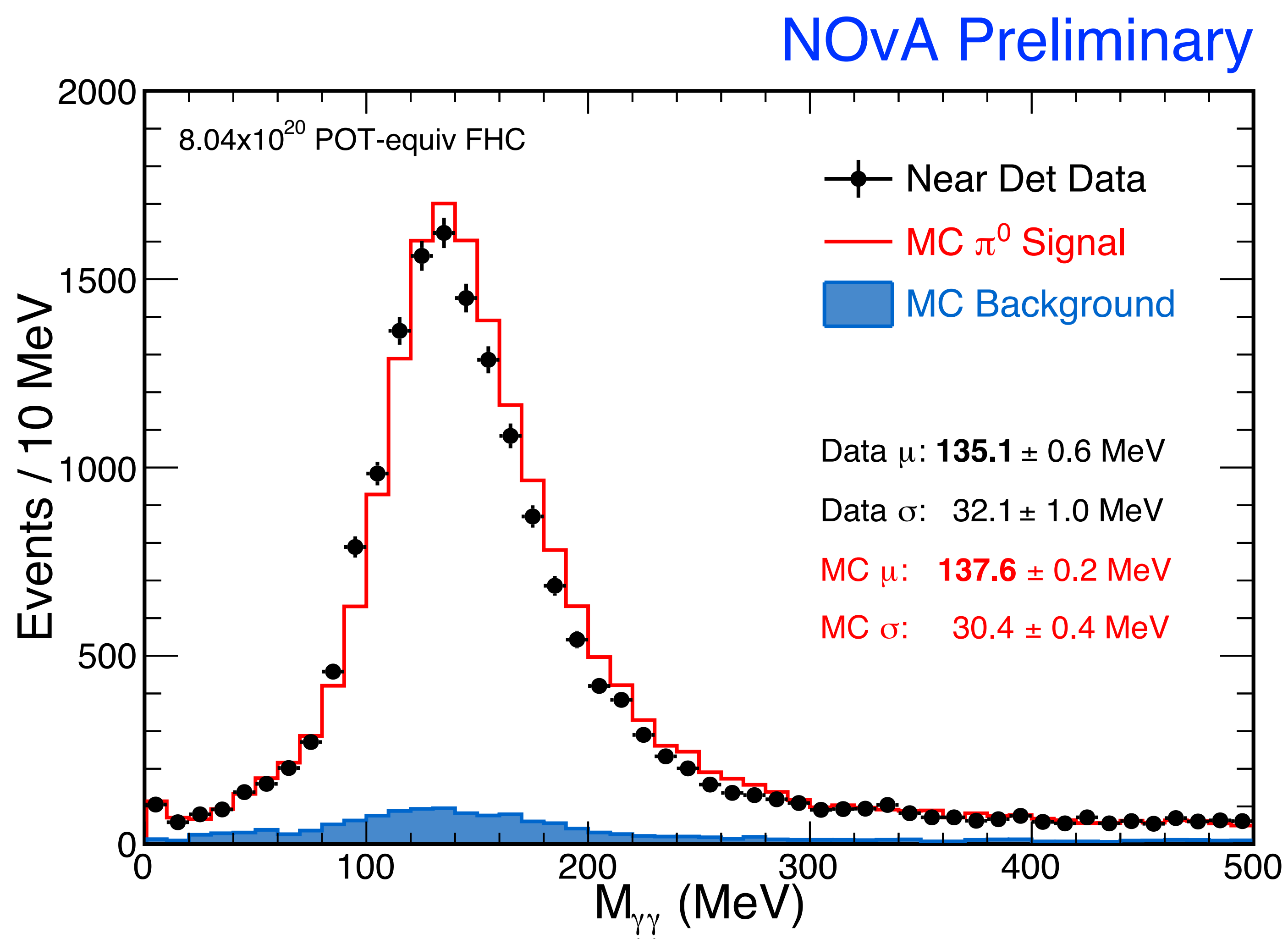


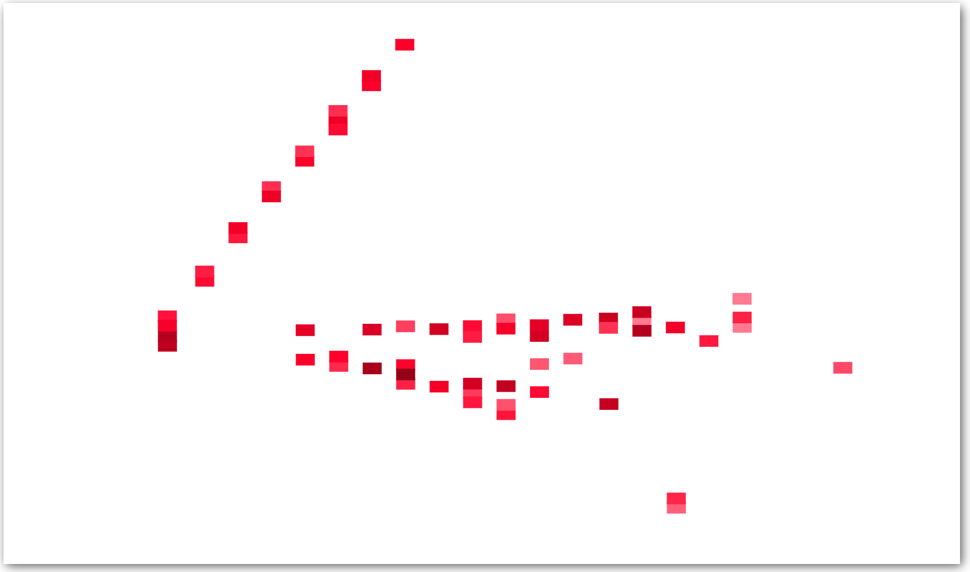


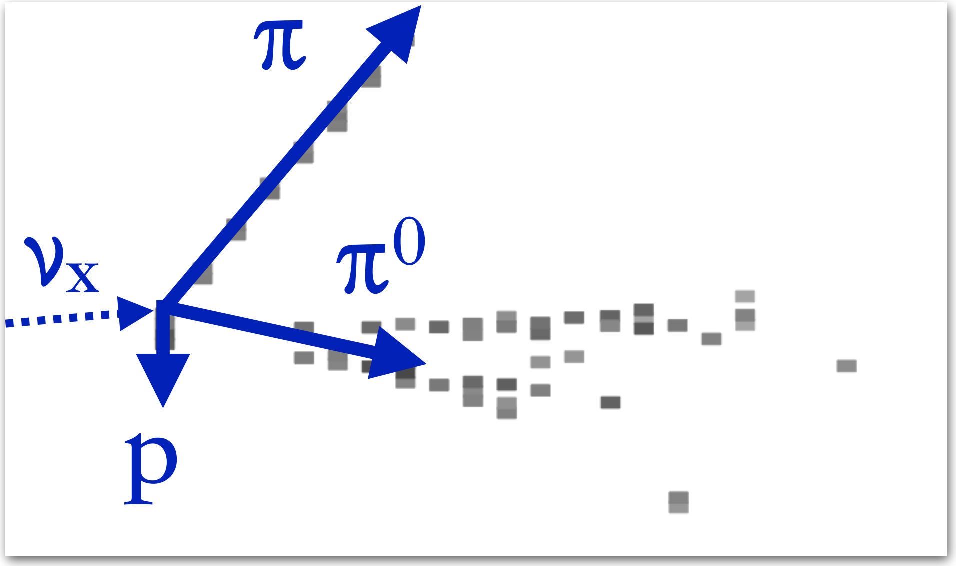
Physics and Calibration with CNNs

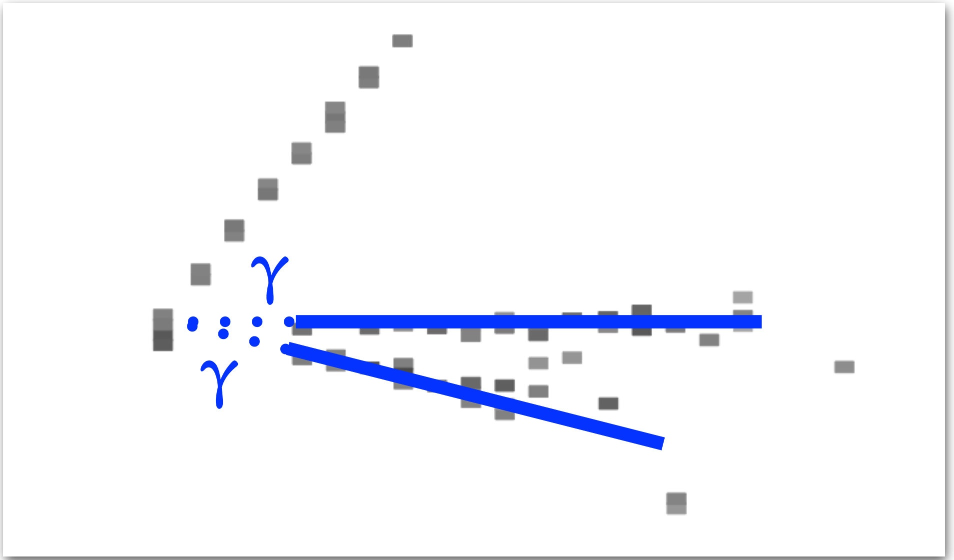
Reconstruct the well known π^0 mass from well identified pairs of photons.

CNN based selection yields **60% background reduction** for the same efficiency







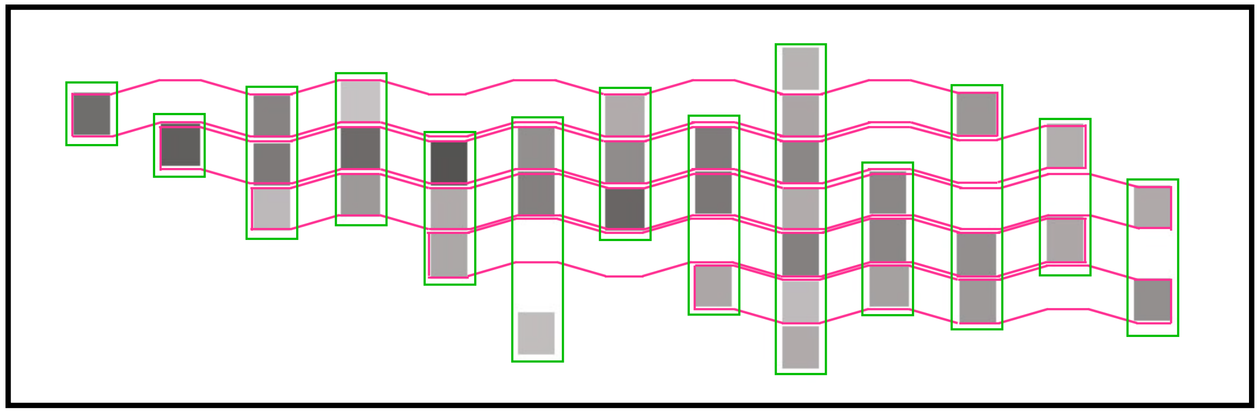


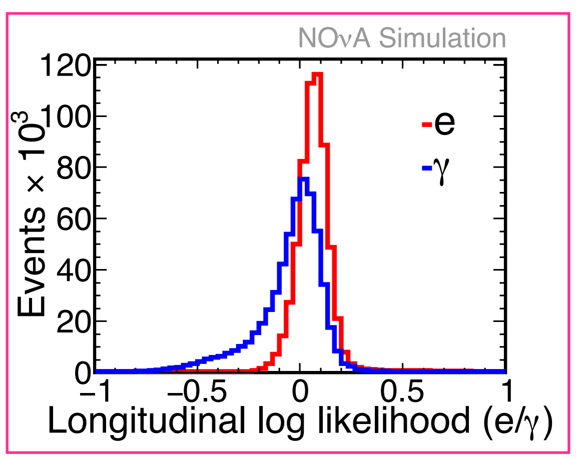
Neutrino Identification (First Analysis)

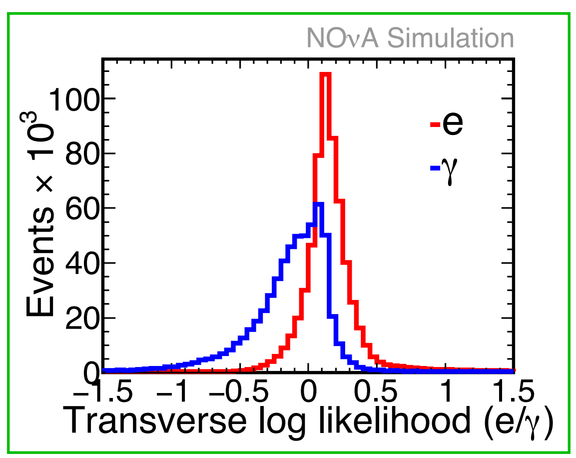
LID

Likelihood Identification

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







In practice:

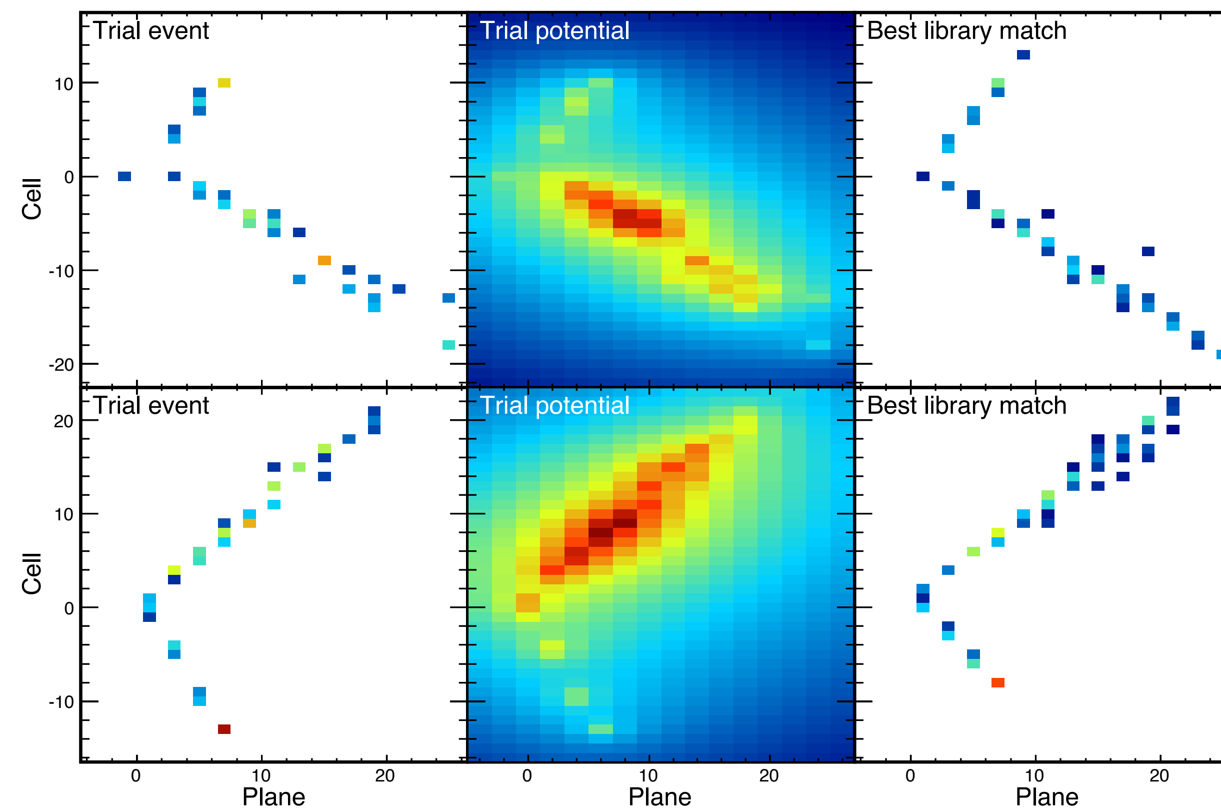
- ★ Reconstruct electron shower.
- ★ Find likelihoods from it's dE/dx profiles compared to particle hypotheses.

Likelihoods → *Neural Network*

LEM

Library Event Matching

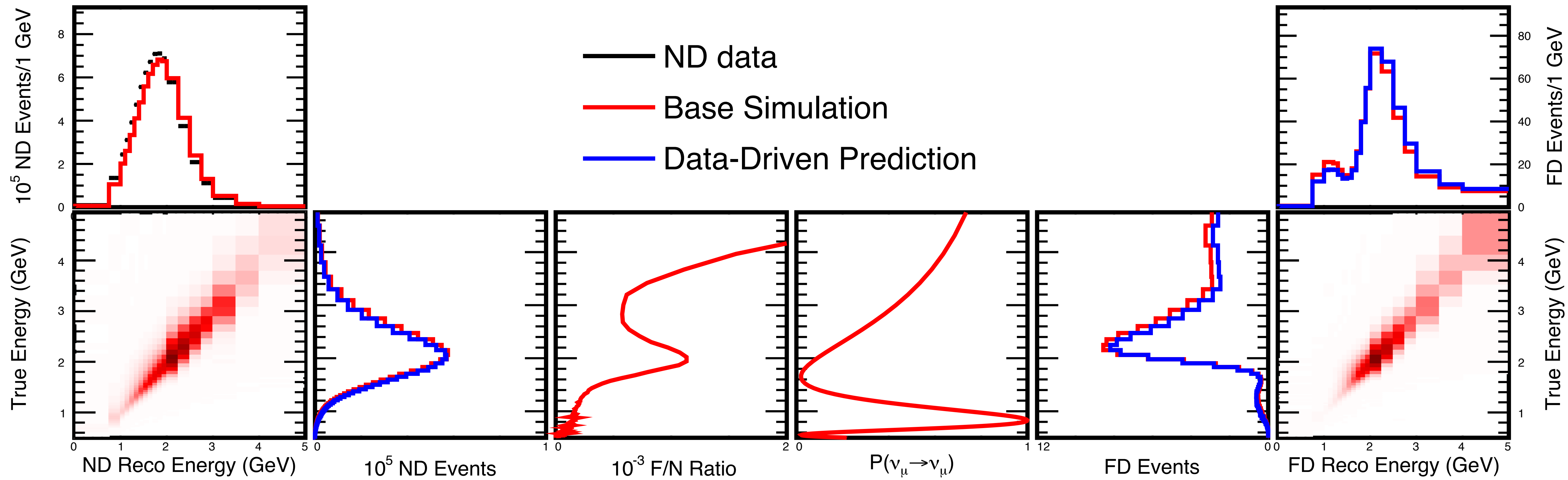
Premise: We have a large library of simulated event templates, large enough that we can use it to compare pixel by pixel.

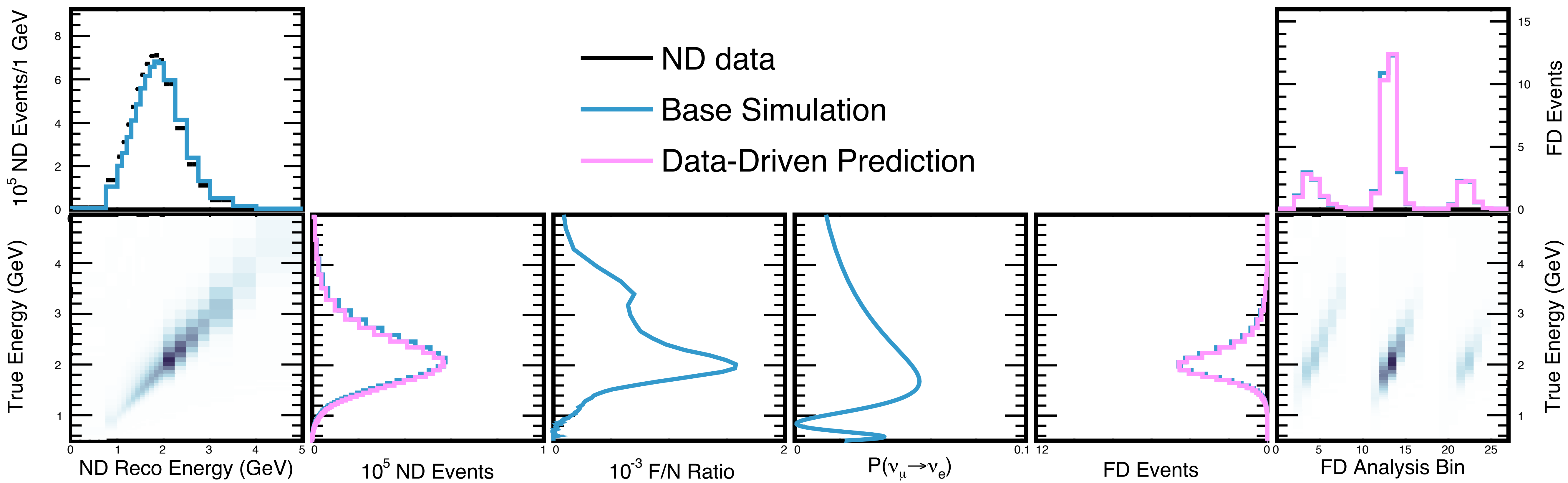


In practice:

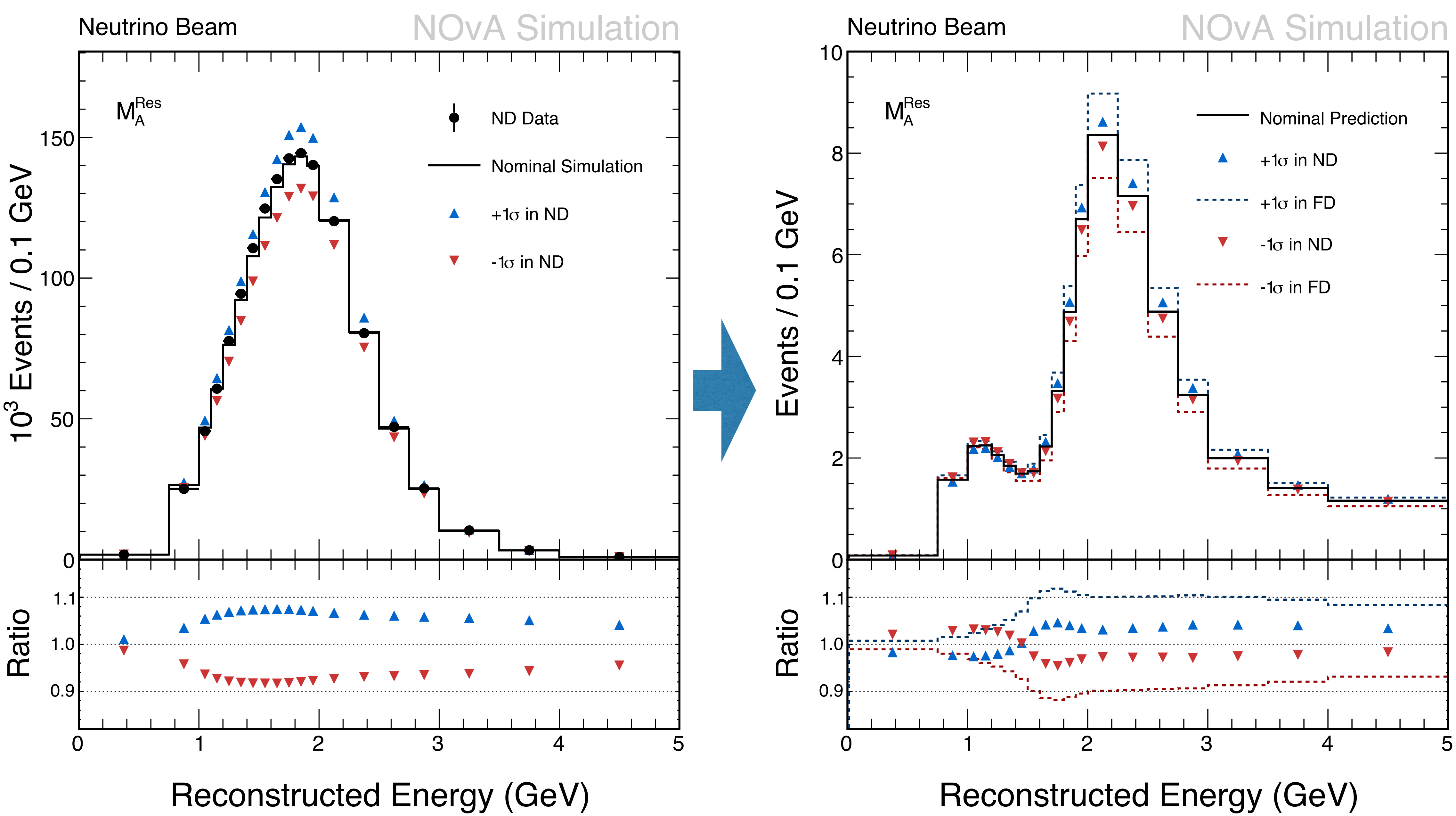
- ★ Find the best matches from the event library.
- ★ Extract features from best matches.

Features → *Decision Tree*


ND data
Base Simulation
Data-Driven Prediction
10^5 ND Events/1 GeV
FD Events/1 GeV
True Energy (GeV)
ND Reco Energy (GeV)
10^5 ND Events
10^{-3} F/N Ratio
$P(\nu_\mu \to \nu_\mu)$
FD Events
FD Reco Energy (GeV)


ND data
Base Simulation
Data-Driven Prediction
10^5 ND Events/1 GeV
FD Events
True Energy (GeV)
ND Reco Energy (GeV)
10^5 ND Events
10^{-3} F/N Ratio
$P(\nu_\mu \to \nu_e)$
FD Events
FD Analysis Bin

Introducing a Systematic Shift



ND **FD**