

DEEP LEARNING ON NOVA

Applications, Successes and Lessons

Fernanda Psihas

 **Fermilab**

Outline

NOvA & our deep learning program

Experiment, physics, data

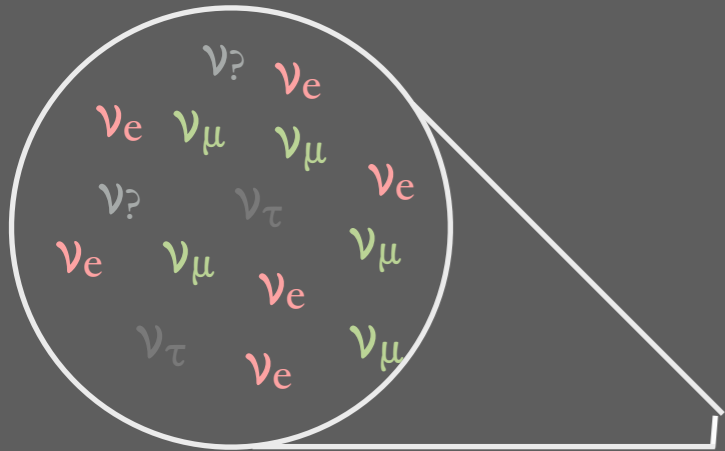
Summary of DL applications

Challenges & Lessons

Think about the **physics**: applications & data

Think about the **physicists**: framework & training

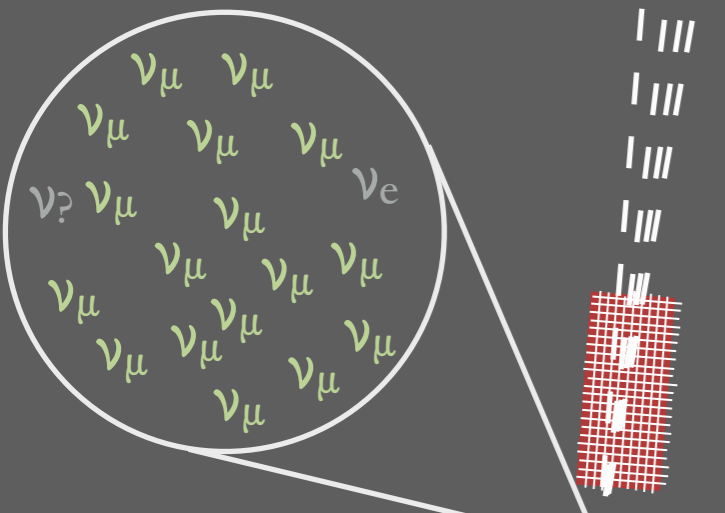
Think about the **problems**: Robustness, uncertainties, etc



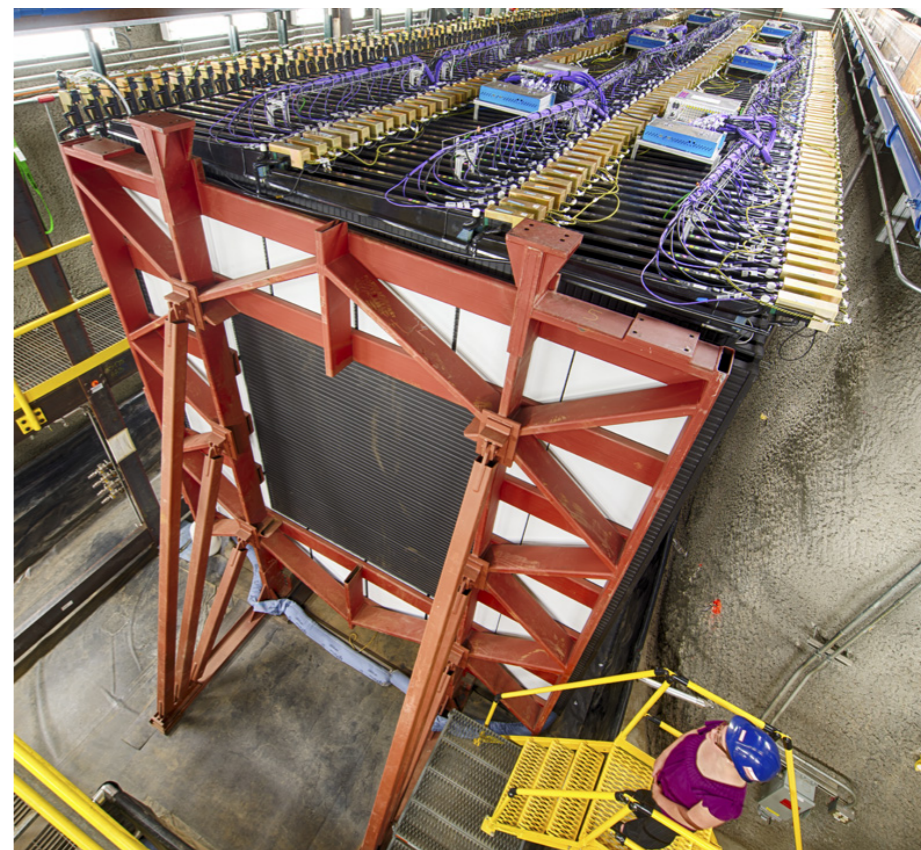
Neutrino oscillation measurements at the Far Detector



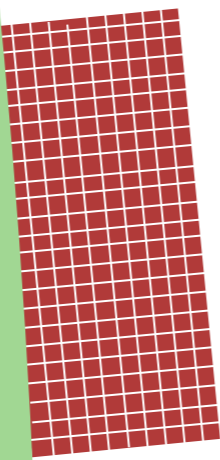
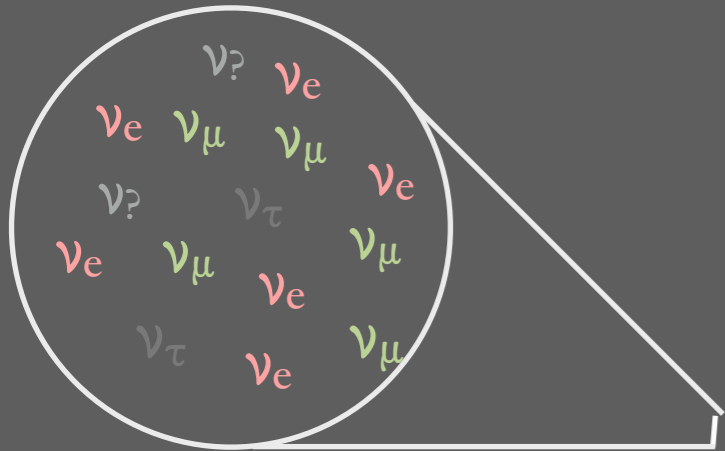
The NOvA Experiment



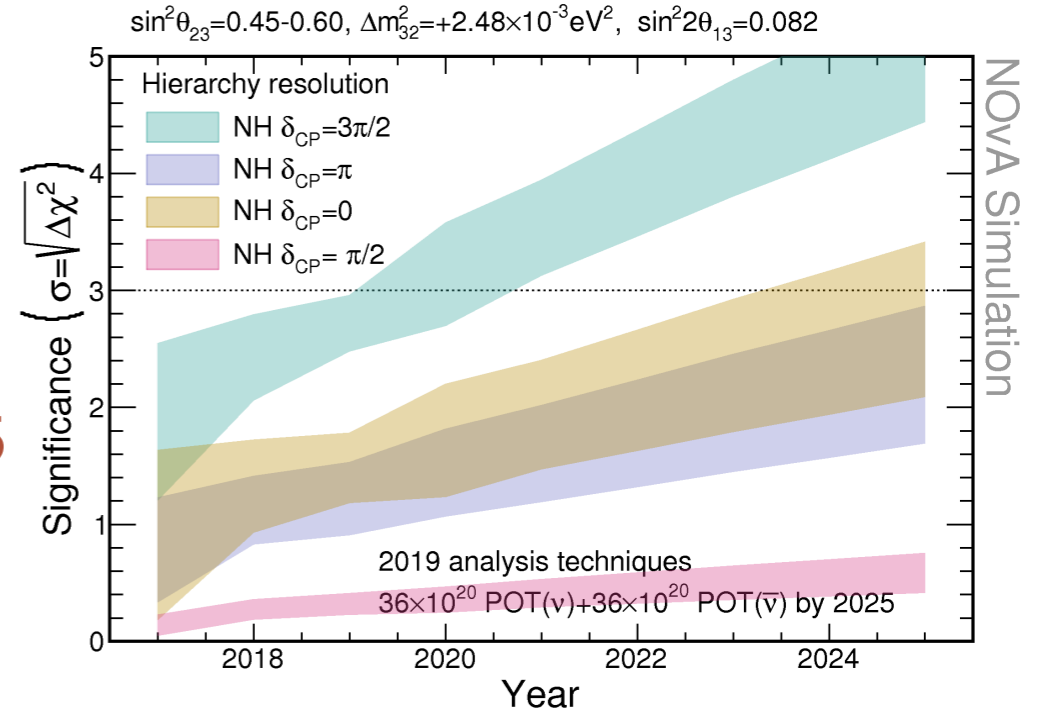
Neutrino interaction cross-sections measurements at the Near Detector



NuMI beam at Fermilab

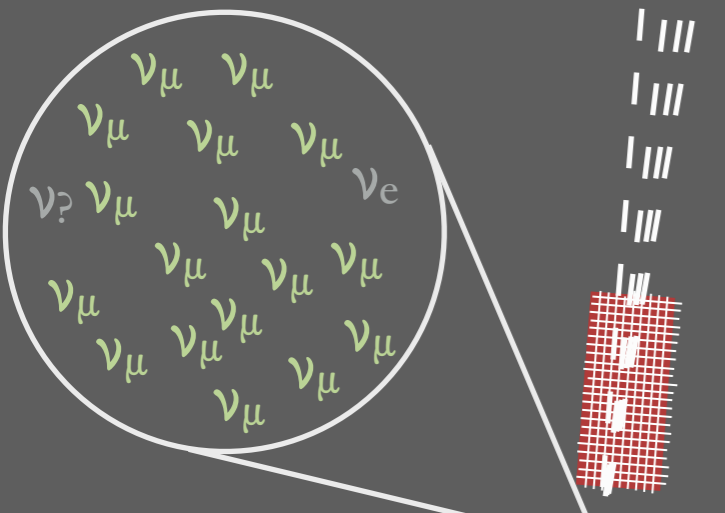


Up to 5σ on the neutrino mass hierarchy by 2025

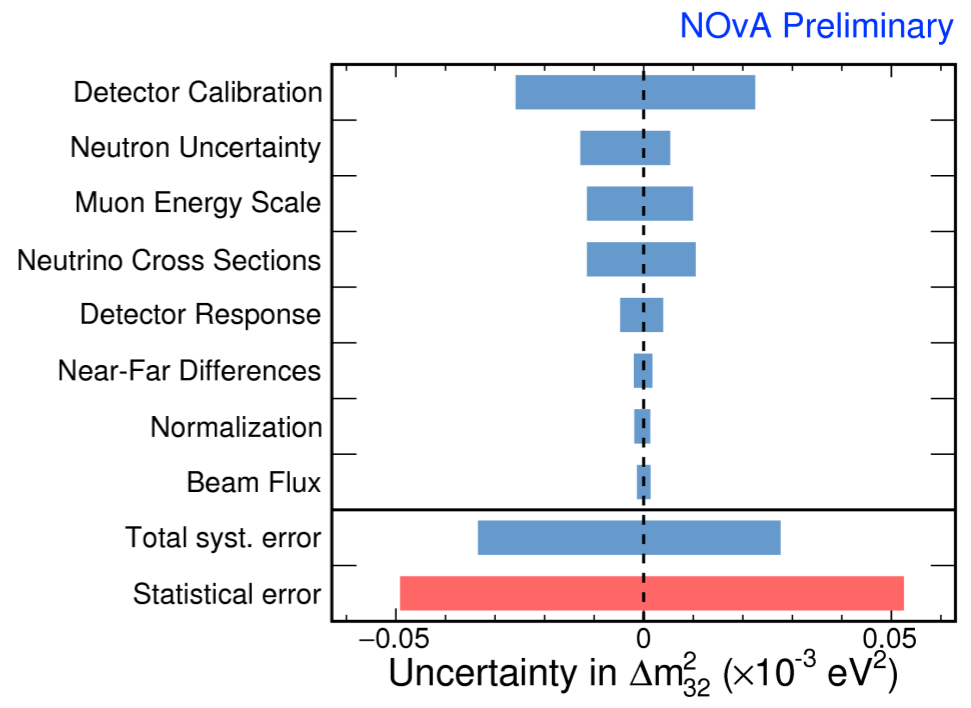


Physics Takeaways

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



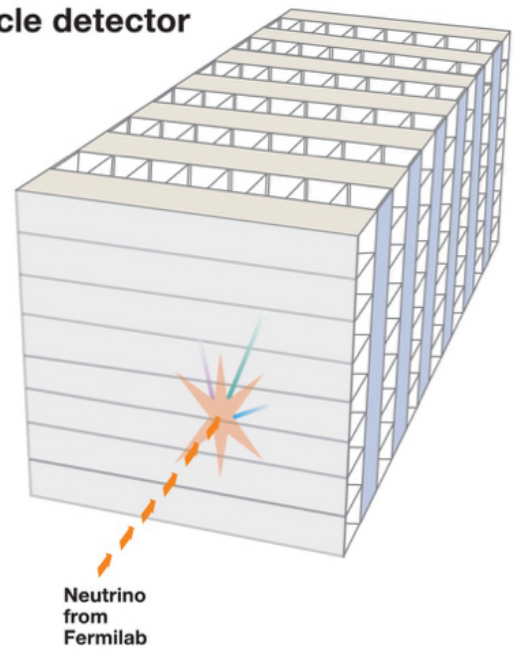
Systematics limited, Detector response is the next largest uncertainty.



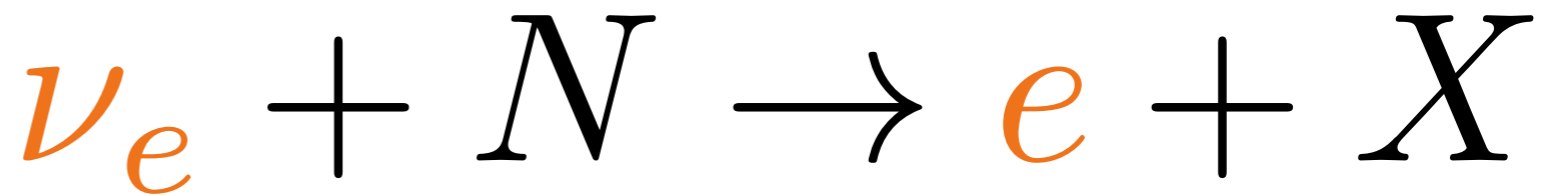
NuMI beam at Fermilab

Measurables are Flavor and Energy

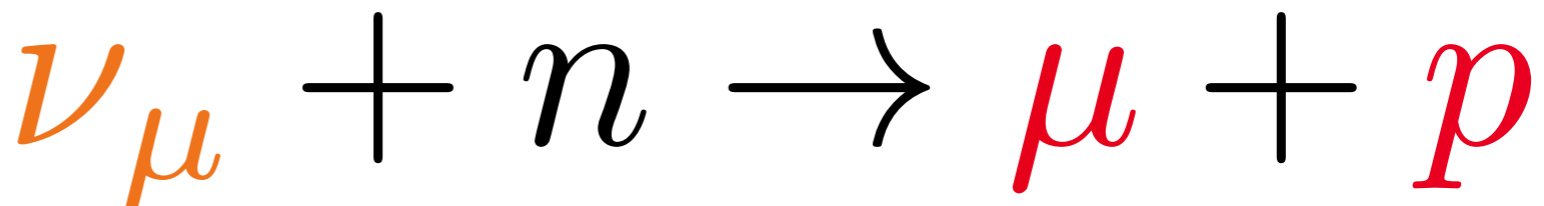
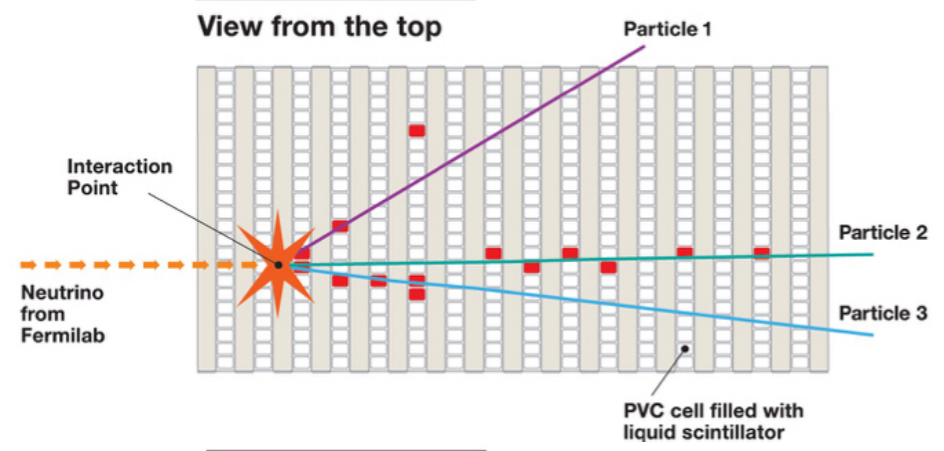
3D schematic of NOvA particle detector



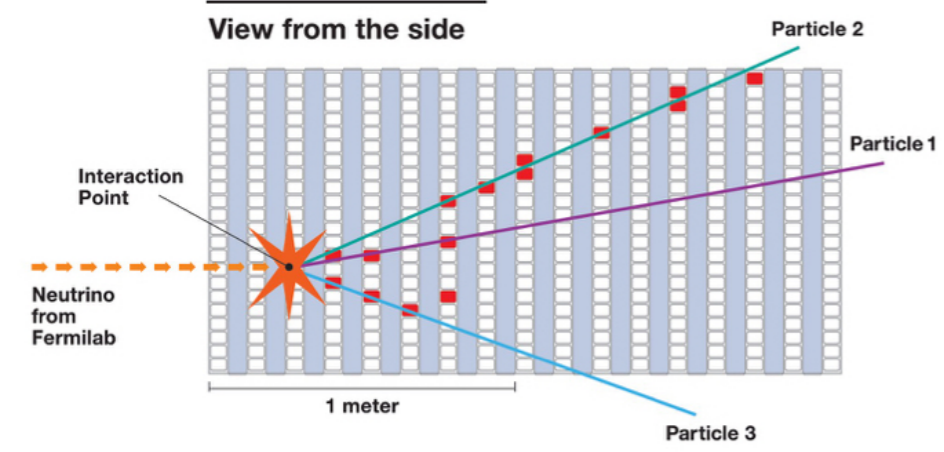
For oscillations we want to measure the incoming neutrino flavor and energy.



Outgoing lepton matches incoming ν flavor



Final state particles



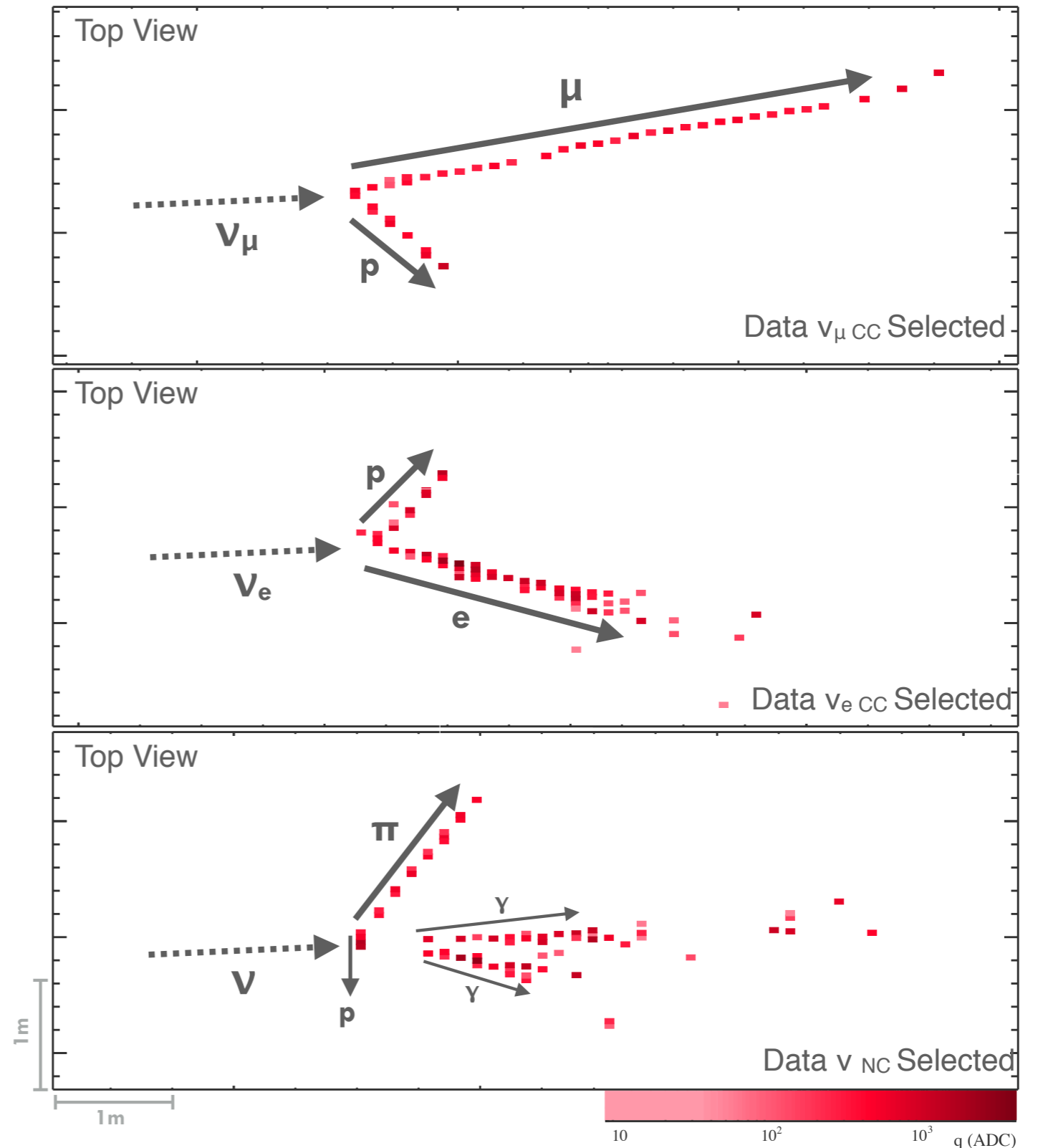
For interaction cross-sections we want to measure the incoming neutrino flavor and energy, as well as the final state.

NO ν A event topology

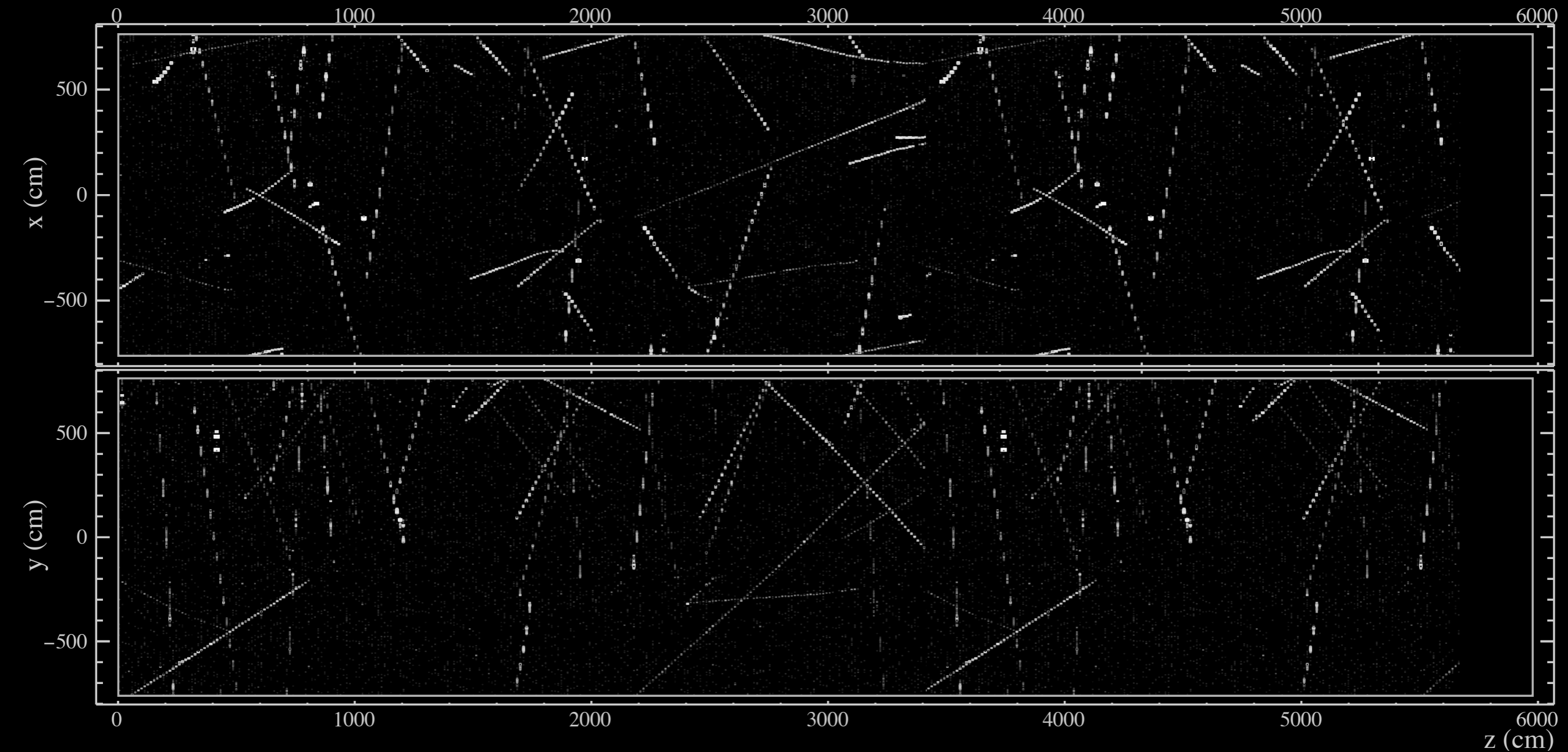
Long tracks are typically muons (minimally ionizing).

e and γ produce electromagnetic showers spanning multiple detector cells.

Heavier particles typically shorter, higher dE/dx tracks.



NOvA's Reconstruction in 2 minutes



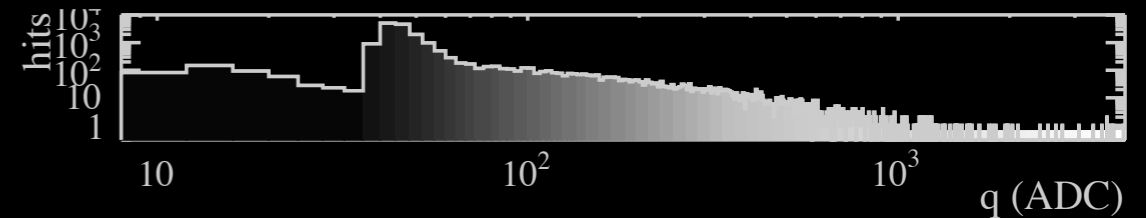
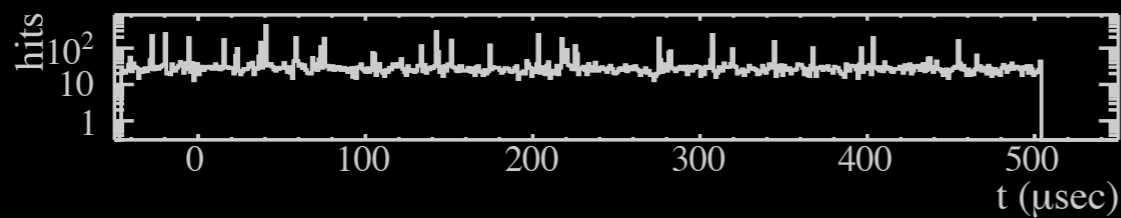
NOvA - FNAL E929

Run: 14828 / 38

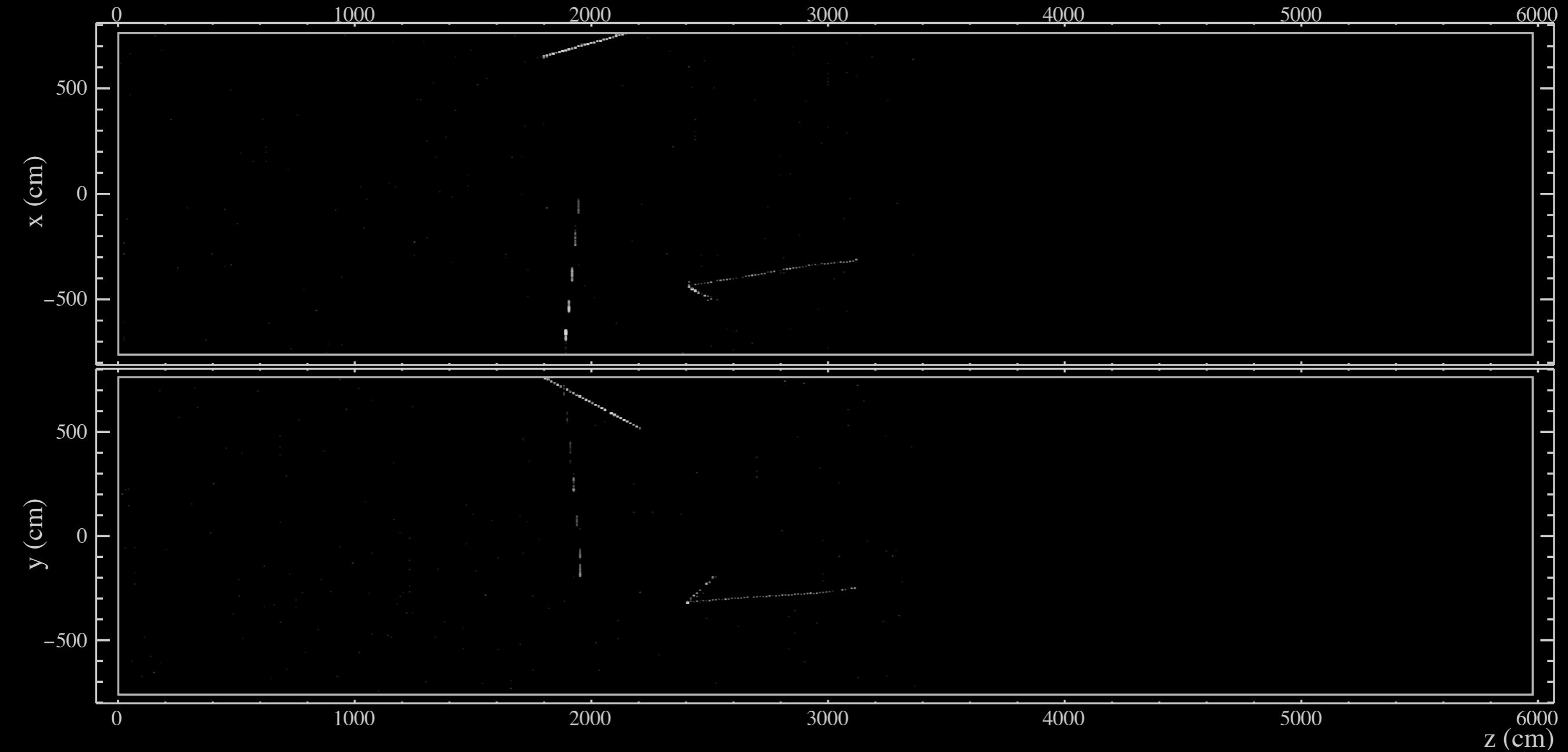
Event: 192569 / --

UTC Tue Apr 22, 2014

21:41:51.422846016



NOvA's Reconstruction in 2 minutes



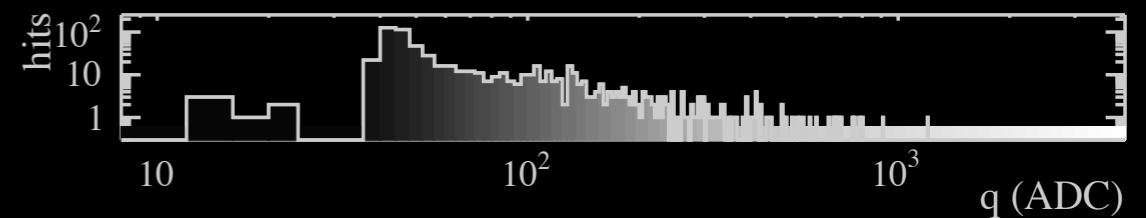
NOvA - FNAL E929

Run: 14828 / 38

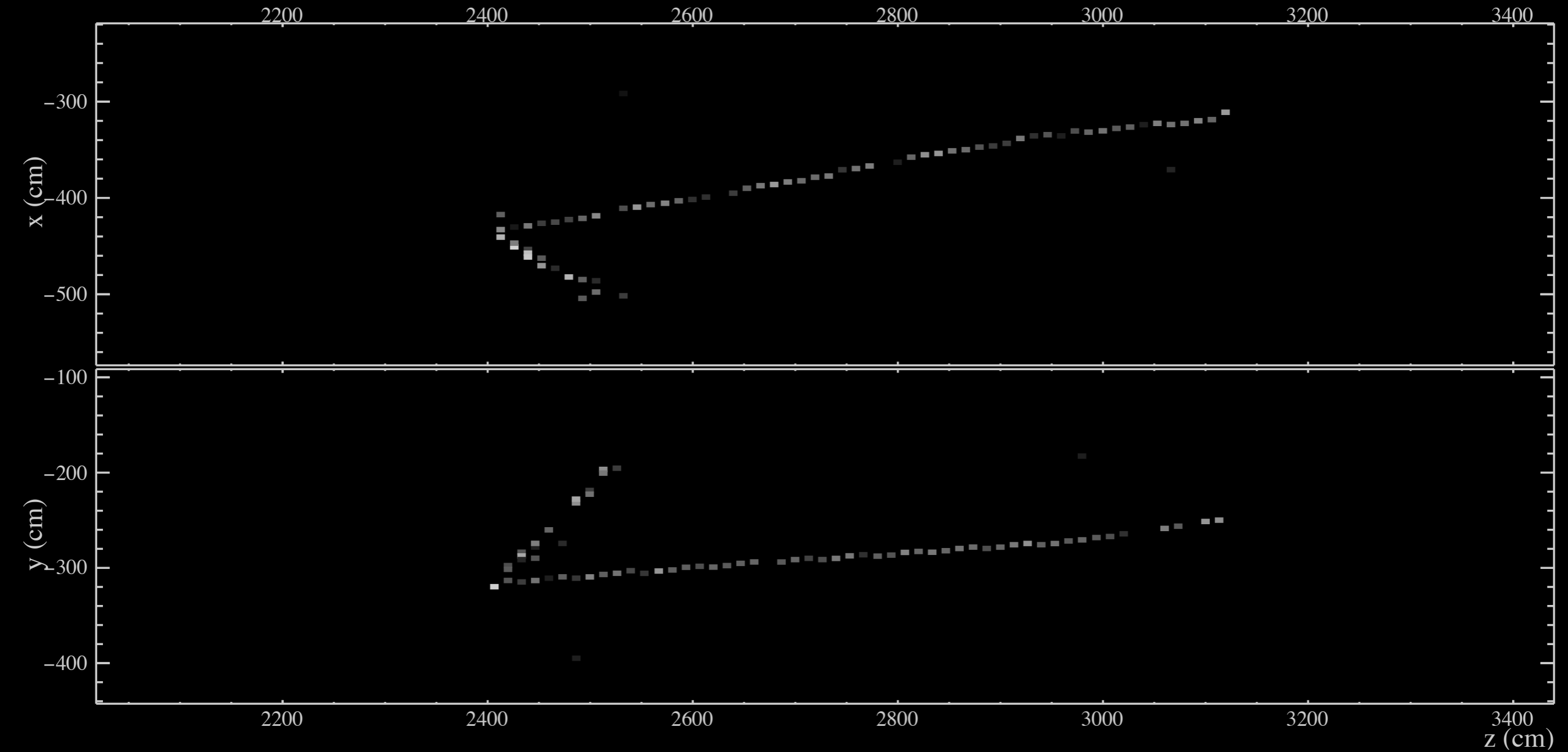
Event: 192569 / --

UTC Tue Apr 22, 2014

21:41:51.422846016



NOvA's Reconstruction in 2 minutes



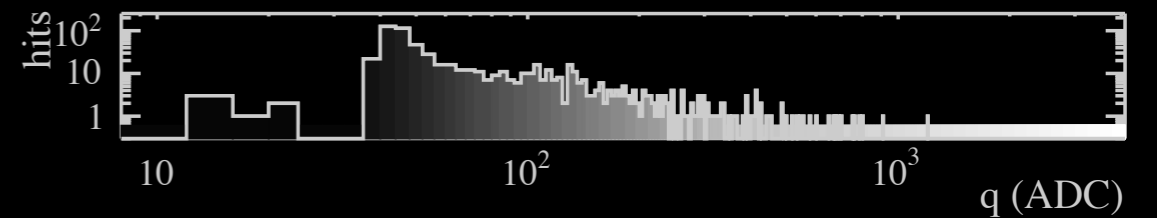
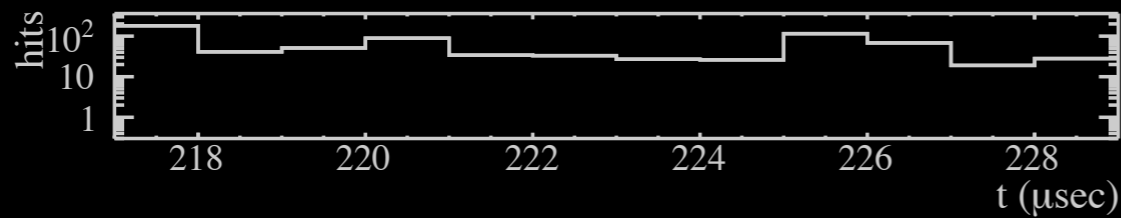
NOvA - FNAL E929

Run: 14828 / 38

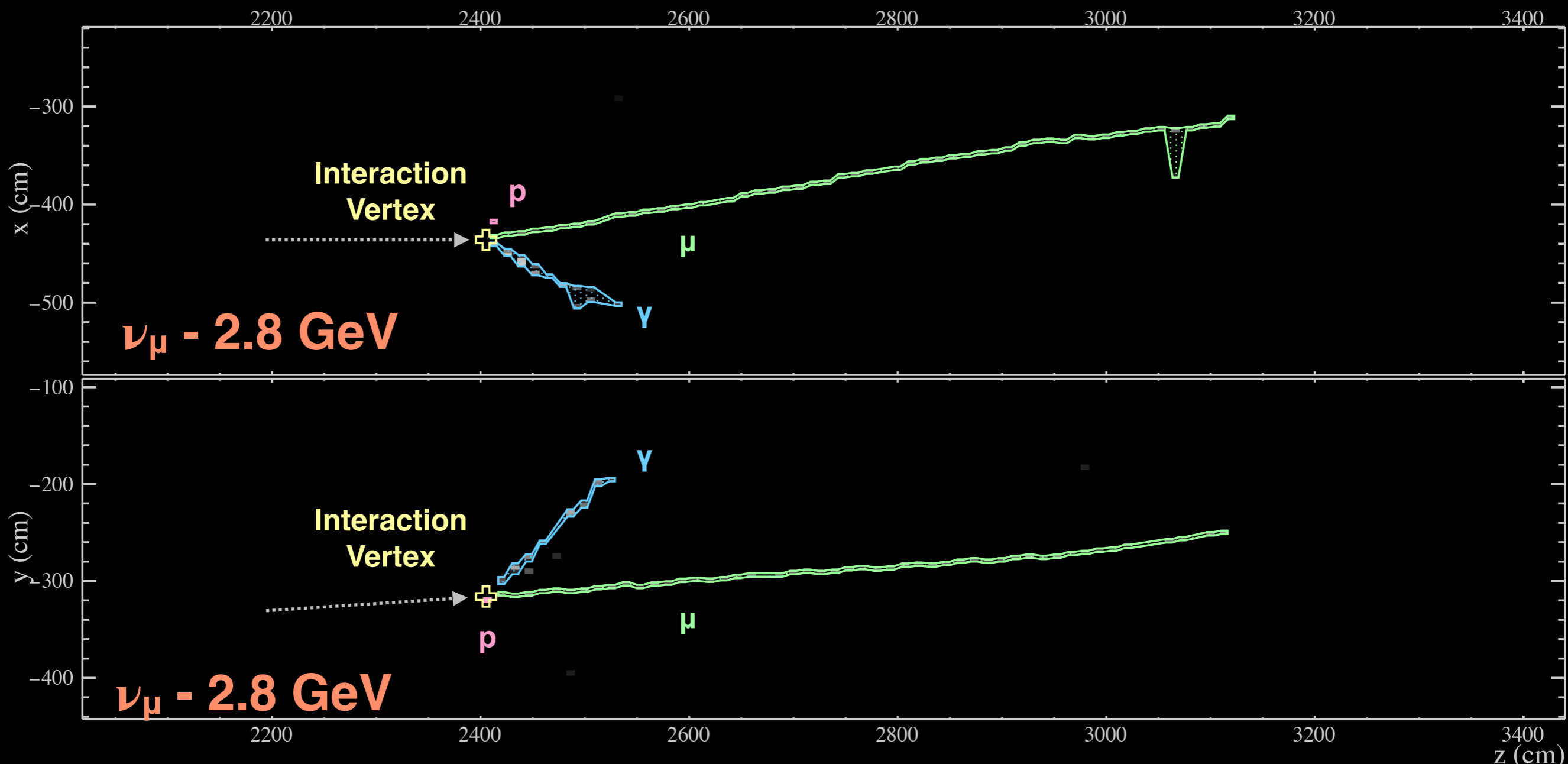
Event: 192569 / --

UTC Tue Apr 22, 2014

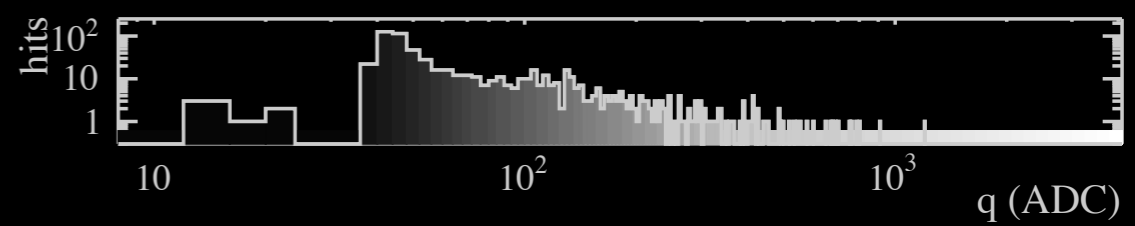
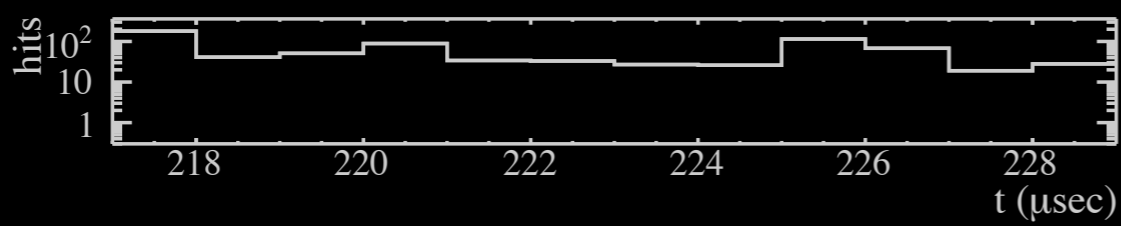
21:41:51.422846016



NOvA's Reconstruction in 2 minutes



NOvA - FNAL E929
Run: 14828 / 38
Event: 192569 / --
UTC Tue Apr 22, 2014
21:41:51.422846016



Reconstruction and Identification



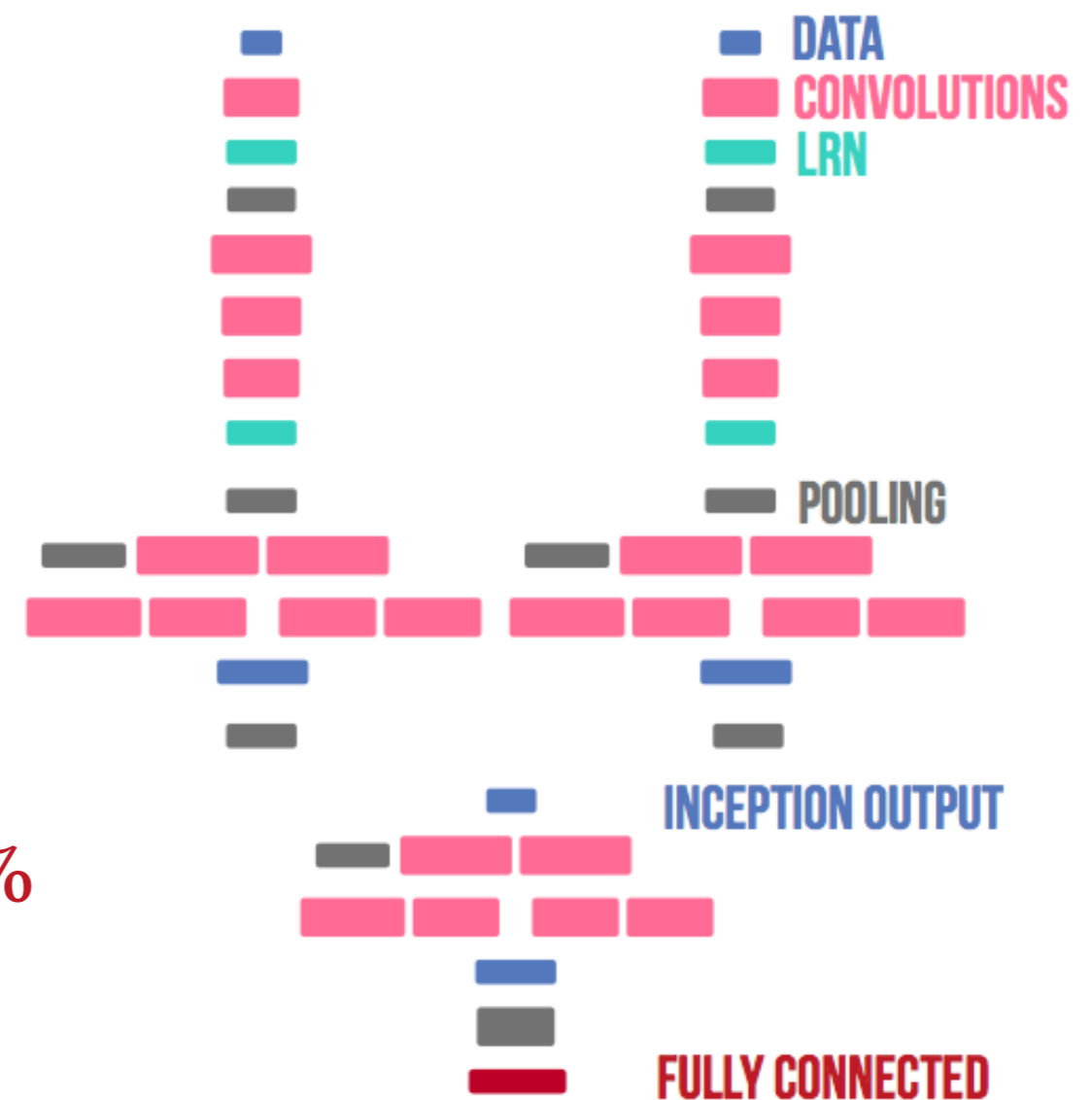
Neutrino Flavor Identification

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

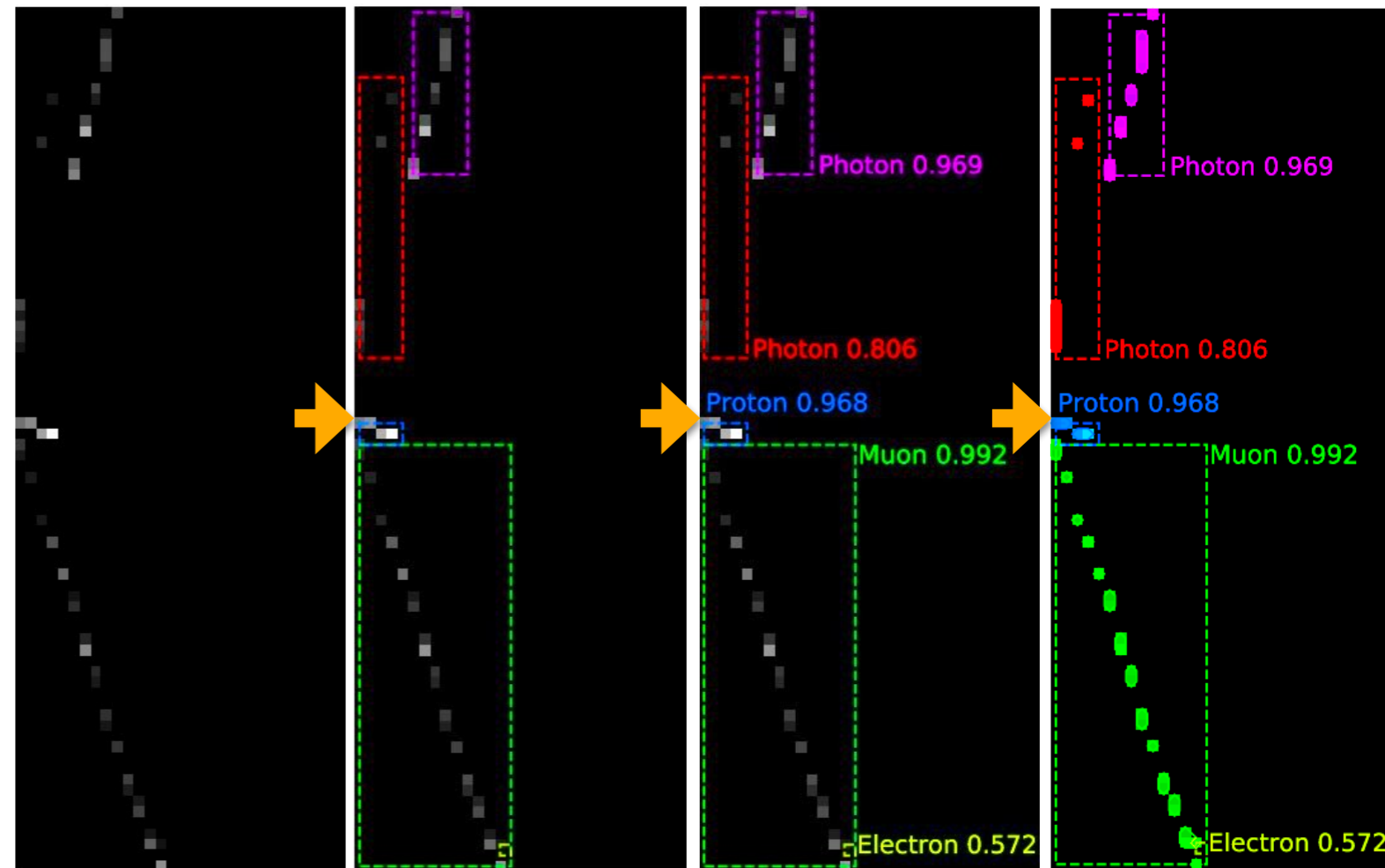
LSTMs for energy reconstruction

CNN cosmic rejection

Systematics-robust trainings

Clustering w/graph networks

Full event reconstruction with semantic segmentation.



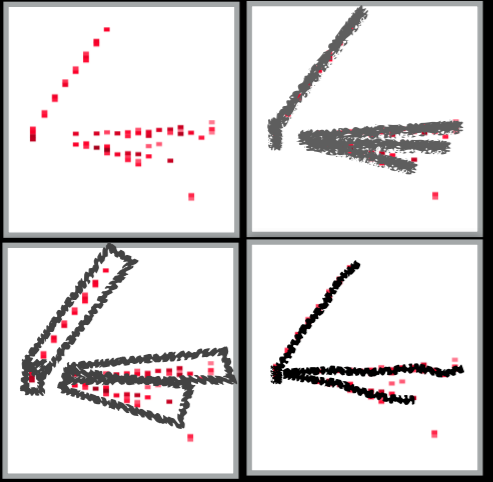
Think about the physics

What is the best tool for your problem-set.

What do you know about your data?

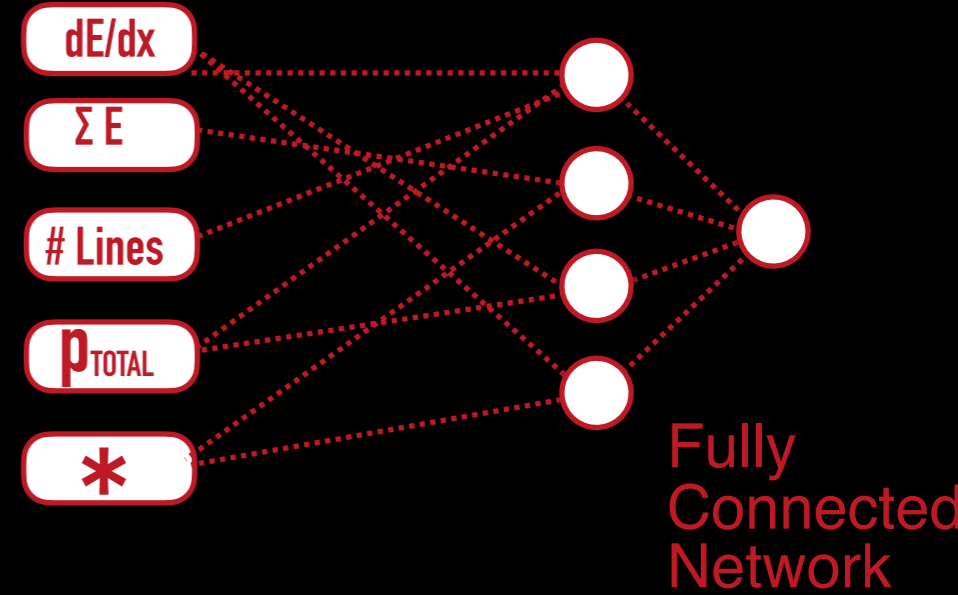
How can you enhance/best use the information contained in your data with

Geometrical Reconstruction

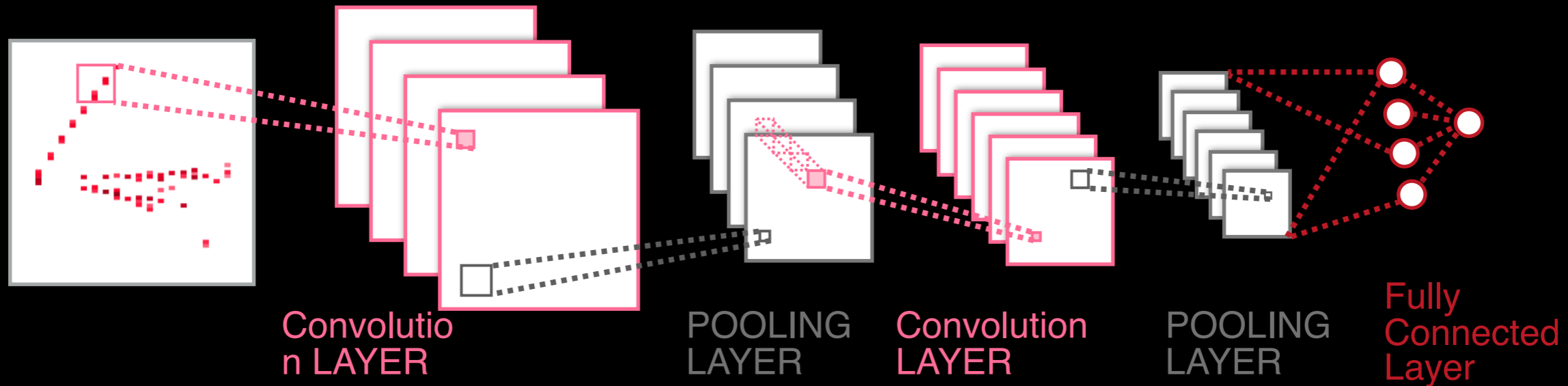


Extract information (features) from the event which can separate signal from background.

Train a neural network for signal vs background discrimination.



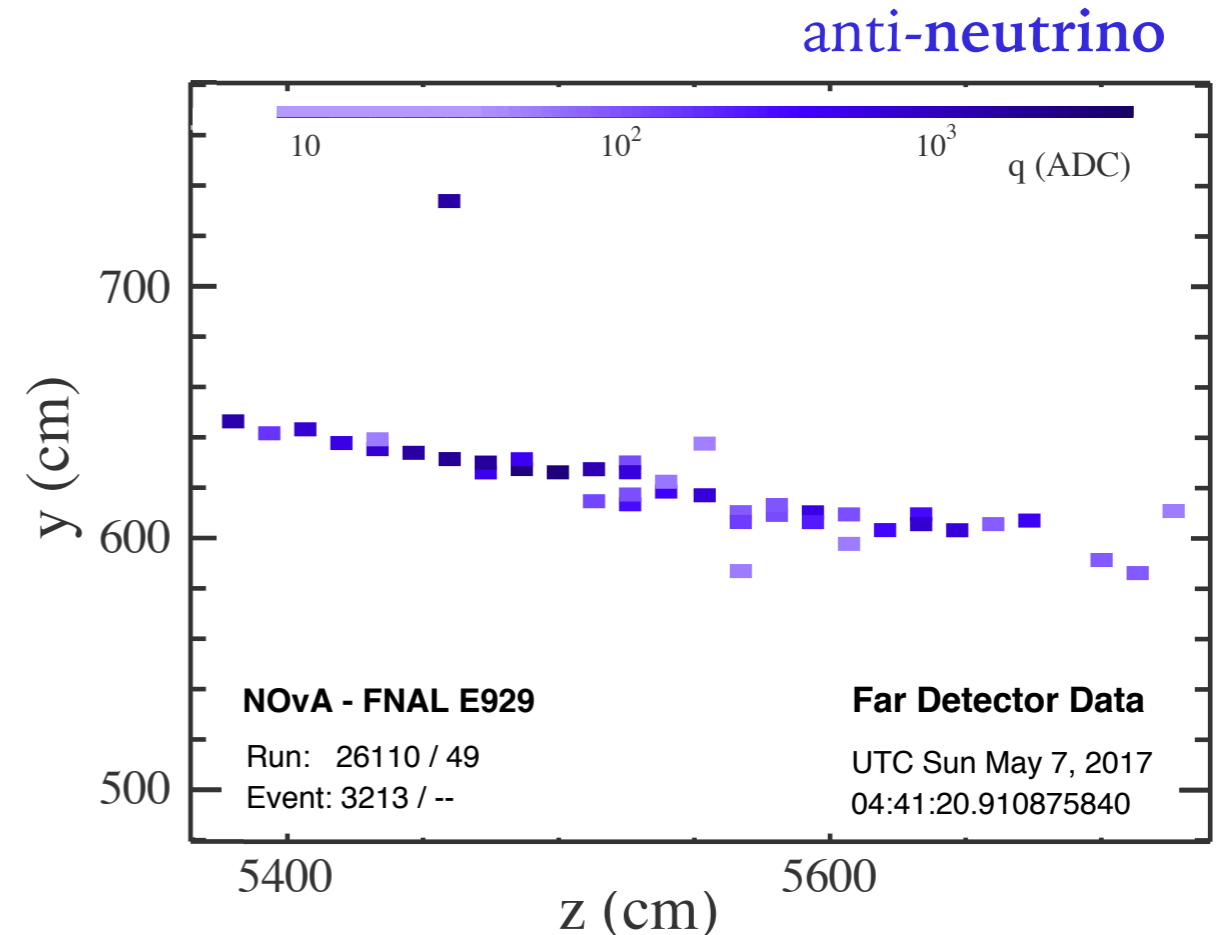
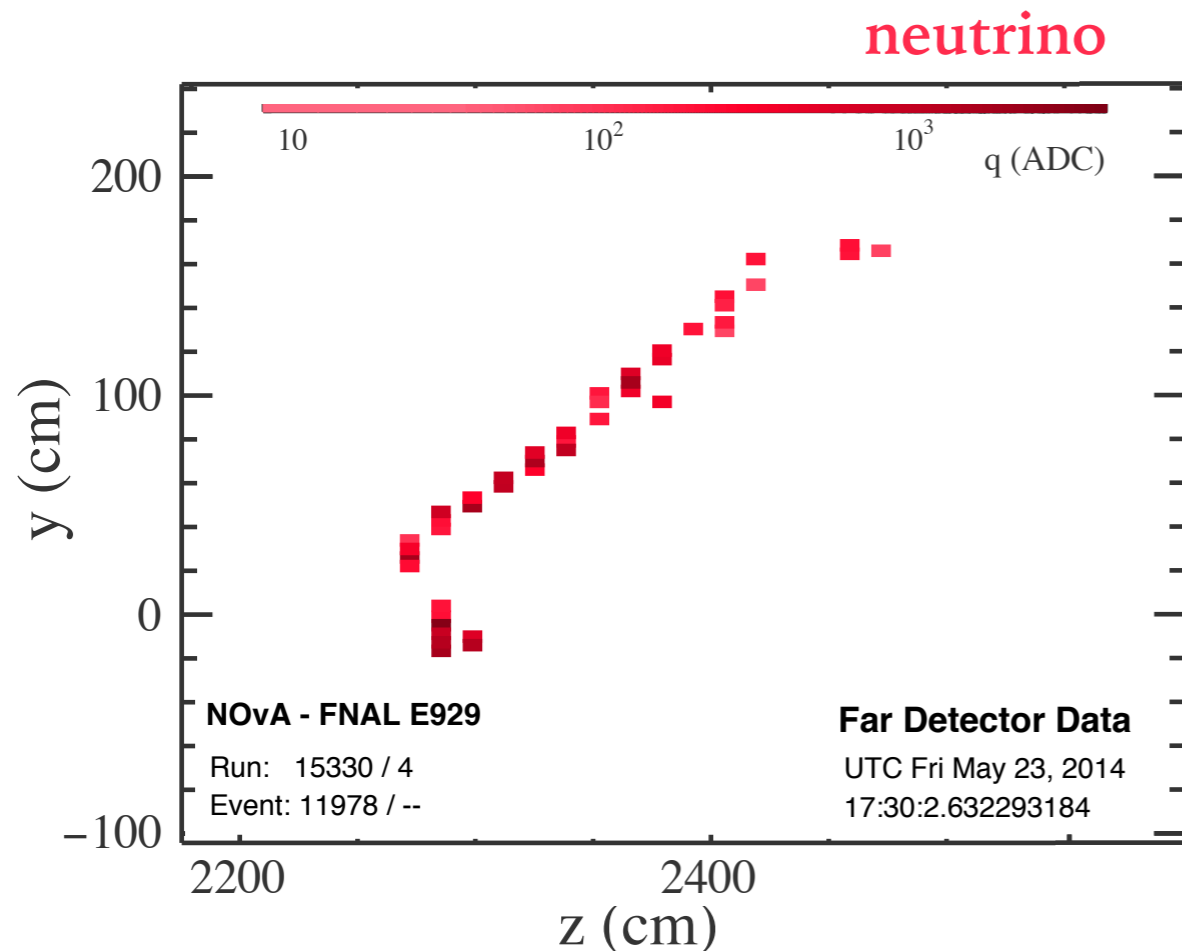
Deep Learning with Convolutional Networks



The features are learnt by the network.

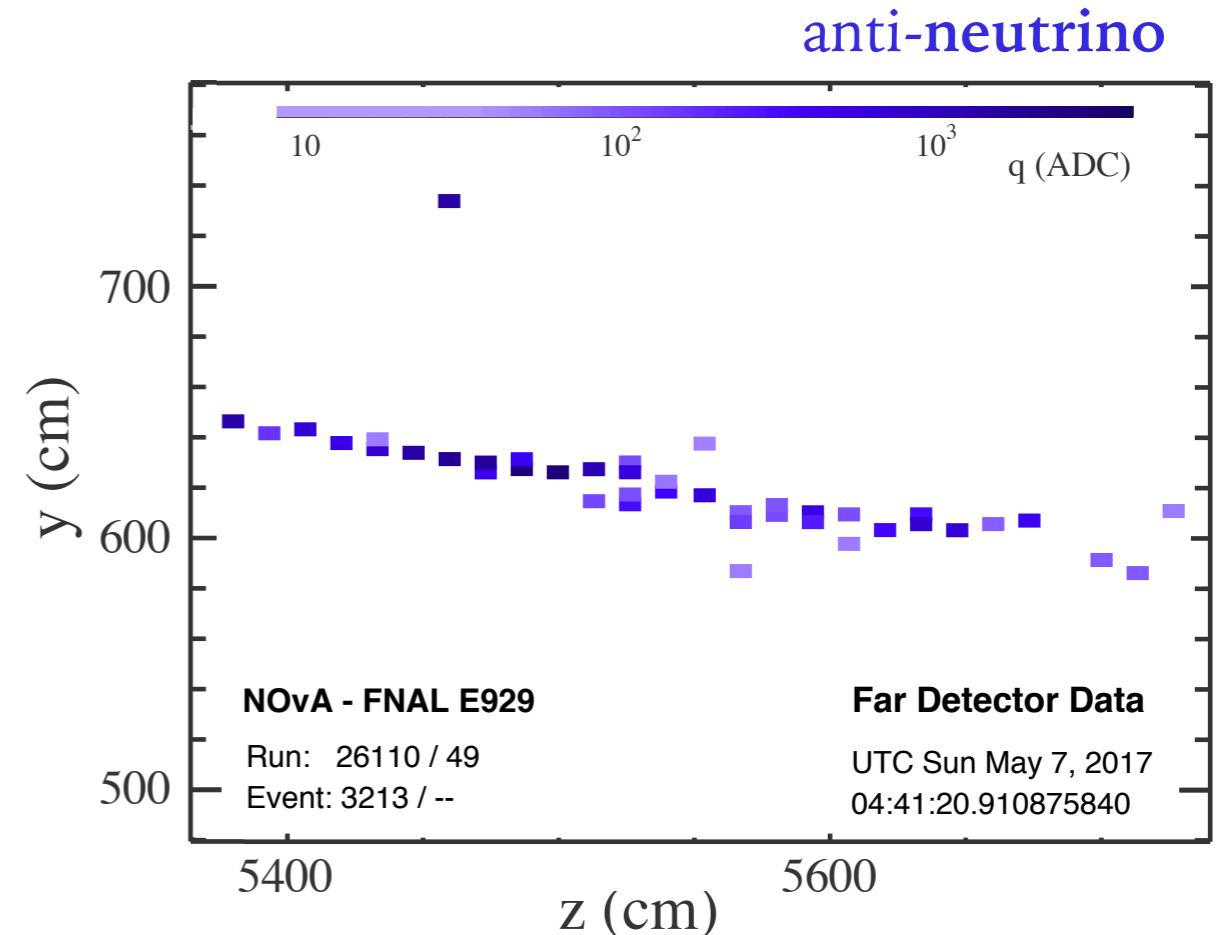
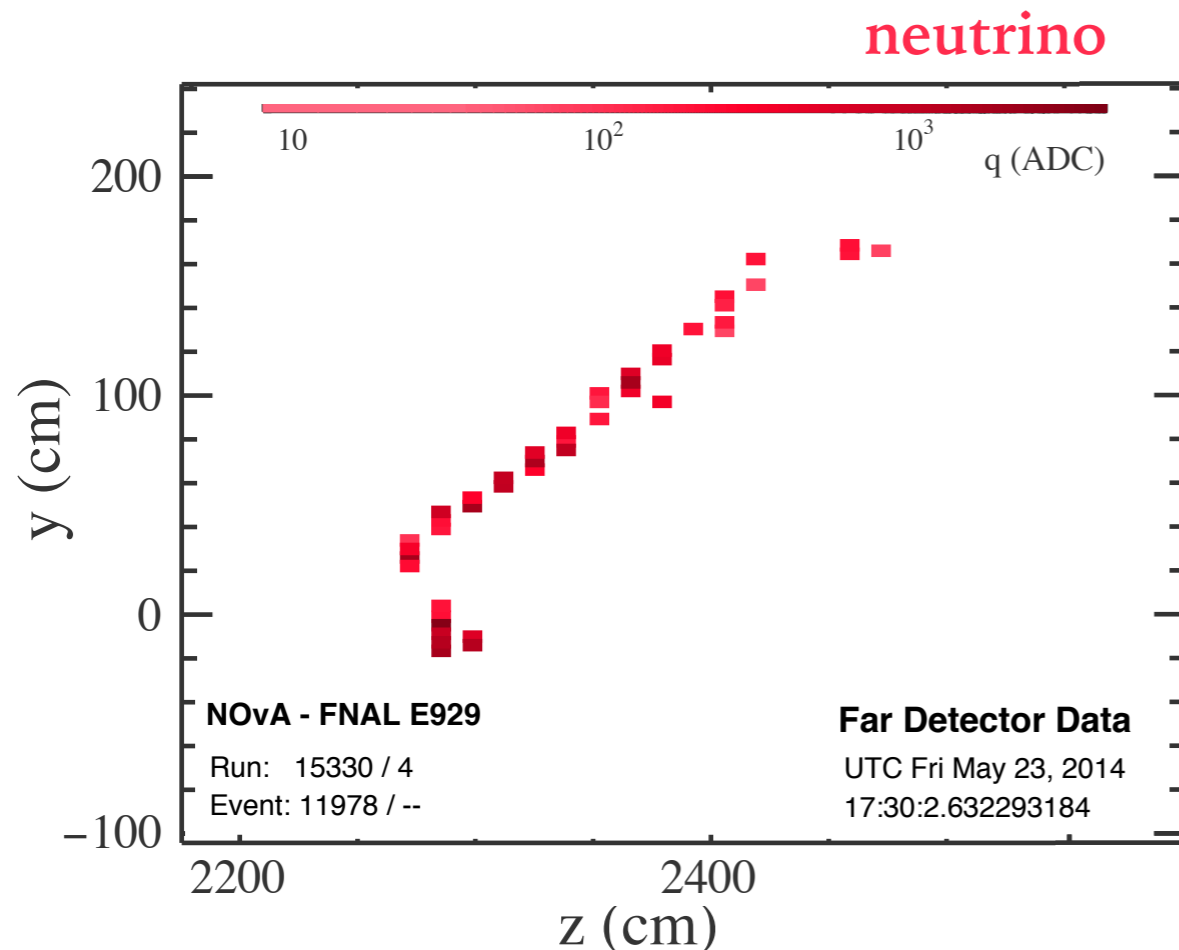
Disentangle from traditional reconstruction.

Event Identification



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

Event Identification



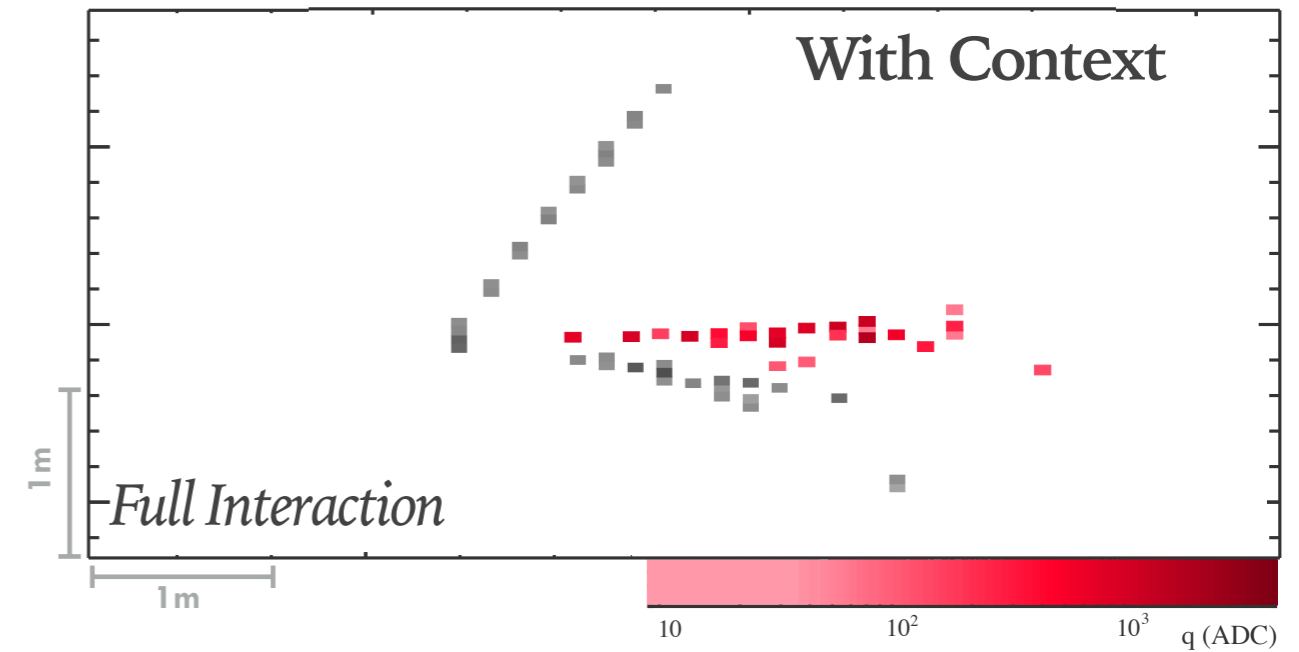
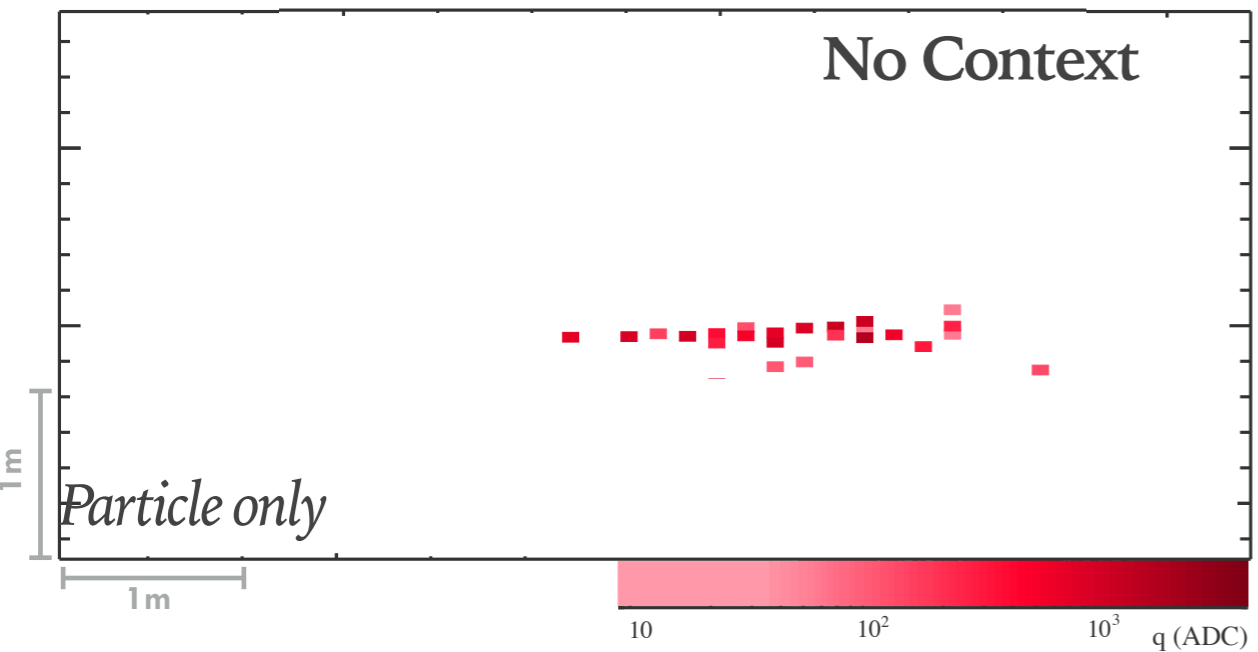
Train on **neutrino** beam and **anti-neutrino** beam simulations separately.

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



Deep Learning Workshop 2017

Fermilab, October 2017 meeting

NOvA CNN

NOvA DL Training prep

Navigating and Editing Caffe Prototxts

Job submit on Wilson Cluster

Validation of CVN Training

CVN in the NOvA offline

Keras Intro (1/2)

Keras Intro (2/2)

Advanced Keras

Deep Learning Literature Review

NOvA DL Training prep

STEP 1: FILE PREP

- Create PixelMaps
- Input are the slices from our MC
- Official prod3 datasets split 2:3 into datasets appended with train (2/3rd size) and evaluate (1/3rd size) identifiers:

```
prod_pid_R17-03-01-prod3reco.lfd_genie_{nonswap,fluxswap,tau}_{thc,rhc}_nova_v08_full_v1_{train,evaluate}
```

Jobs configured with:

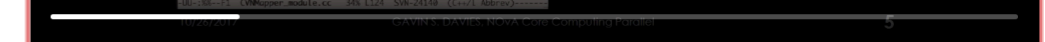
1. art **EDProducer**: makes the PixelMap objects
2. art **EDAnalyzer**: "dump" the PixelMap for each slice into simple root tree

```

//file CVNMapper_module.cc
//brief Producer module for creating CVN PixelMap objects
//author Dominick Bocca - nocap@physics.uiowa.edu

// C/C++ includes
#include <vector>
#include <string>
#include <fstream>
#include <map>
#include <set>
#include <algorithm>
#include <random>
#include <cmath>
#include <string.h>
#include <math.h>
#include <unistd.h>
#include <sys/stat.h>
#include <sys/types.h>
#include <fcntl.h>
#include <sys/mman.h>
#include <sys/time.h>
#include <sys/resource.h>
#include <sys/wait.h>
#include <sys/param.h>
#include <sys/uio.h>
#include <sys/xattr.h>
#include <sys/zfs.h>
#include <sys/fs/extattr.h>
#include <sys/fs/ext2fs.h>
#include <sys/fs/ext3fs.h>
#include <sys/fs/ext4fs.h>
#include <sys/fs/extfs.h>
#include <sys/fs/fuse.h>
#include <sys/fs/nfs.h>
#include <sys/fs/procfs.h>
#include <sys/fs/smbfs.h>
#include <sys/fs/vxfs.h>
#include <sys/fs/ufs.h>
#include <sys/fs/udf.h>
#include <sys/fs/ufs.h>
#include <sys/fs/ufs.h>
#include <sys/fs/ufs.h>
#include <sys/fs/ufs.h>

```



Gavin Davies

[[download](#)] [[slides](#)]

Created by **Fernanda Psihias**

Machine Learning

Machine Learning

Group Liaisons
Monthly Meetings
Upcoming Event
Deep ML Journal Club (internal)
Publications

New to the Machine Learning group?

Subscribe to our mailing list:
machinelearning@fnal.gov

Join the conversation:
[HEPMachineLearning](#)

ML at the Intensity and Cosmic Frontiers

The Inter-experimental Machine Learning Working Group for the Intensity and Cosmic Frontiers

The Inter-experimental Machine Learning Working Group brings together a community of analyzers of HEP data who use and develop machine learning (ML) algorithms to solve physics problems. We share tools and applications, provide training and discuss challenges related to the use of ML tools in the HEP community.

We hold regular meetings focused on tool development, knowledge transfer and common solutions to known challenges. Experts from multiple experiments provide feedback and encourage collaboration among members in order to promote the use of ML tools in the community for problems for which they have been shown to be useful.

We are part of the global HEP community. We believe that building community in an inclusive environment advances knowledge transfer and promotes a more active community.

Our sister group, the Inter-Experimental LHC Machine Learning (IML) working group, is focused on building a community of researchers

CROSS-Experiment

FNAL ML group is a community knowledge-sharing group

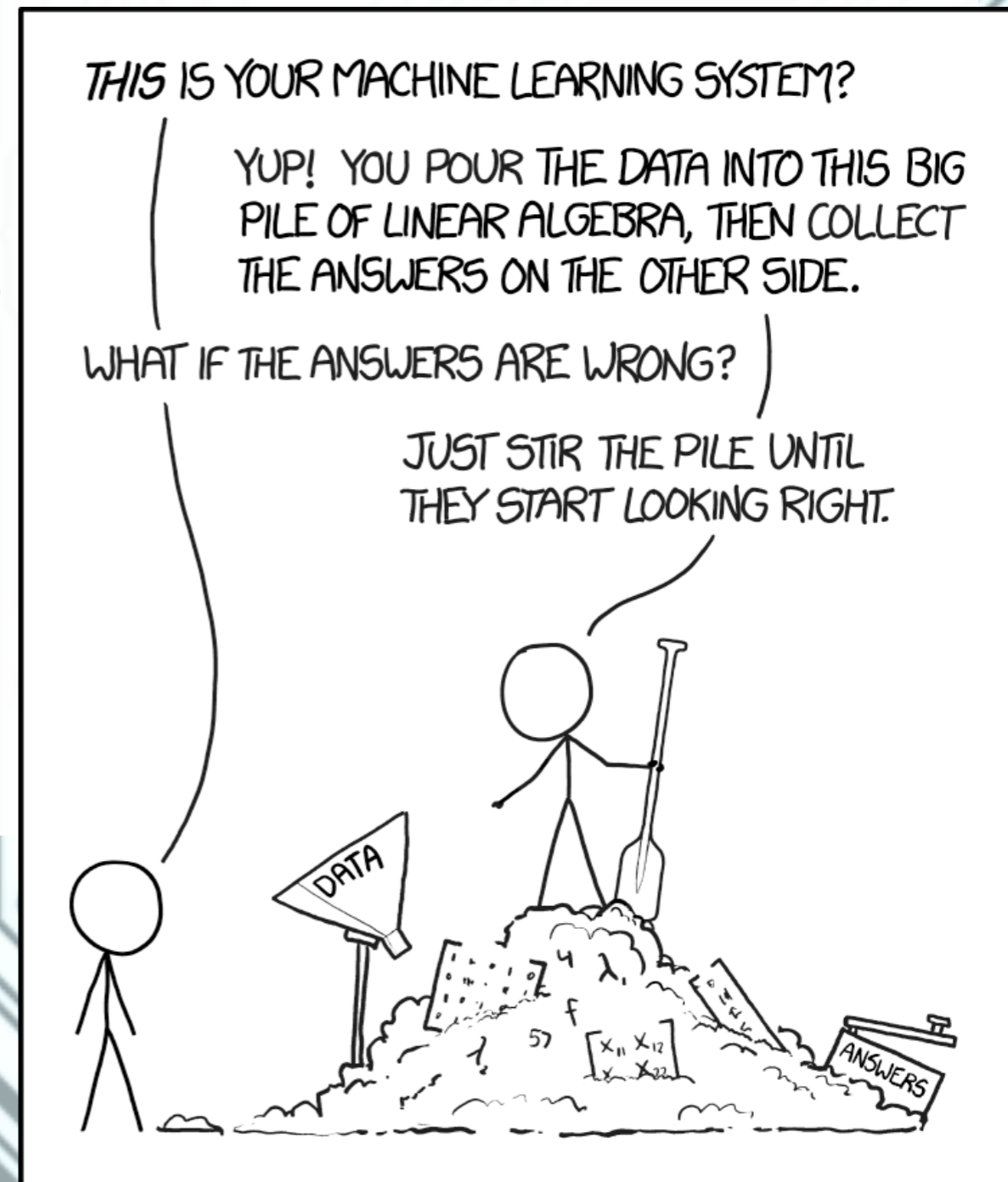
ML experts across experiments discuss technical details of their applications

Think about the problems

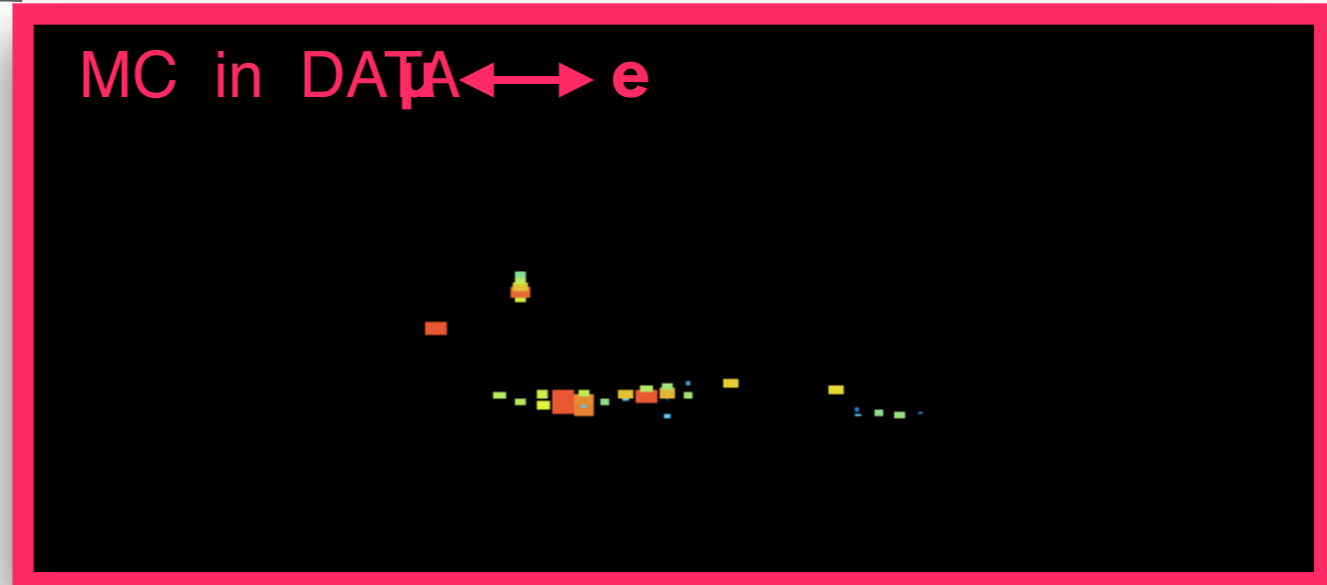
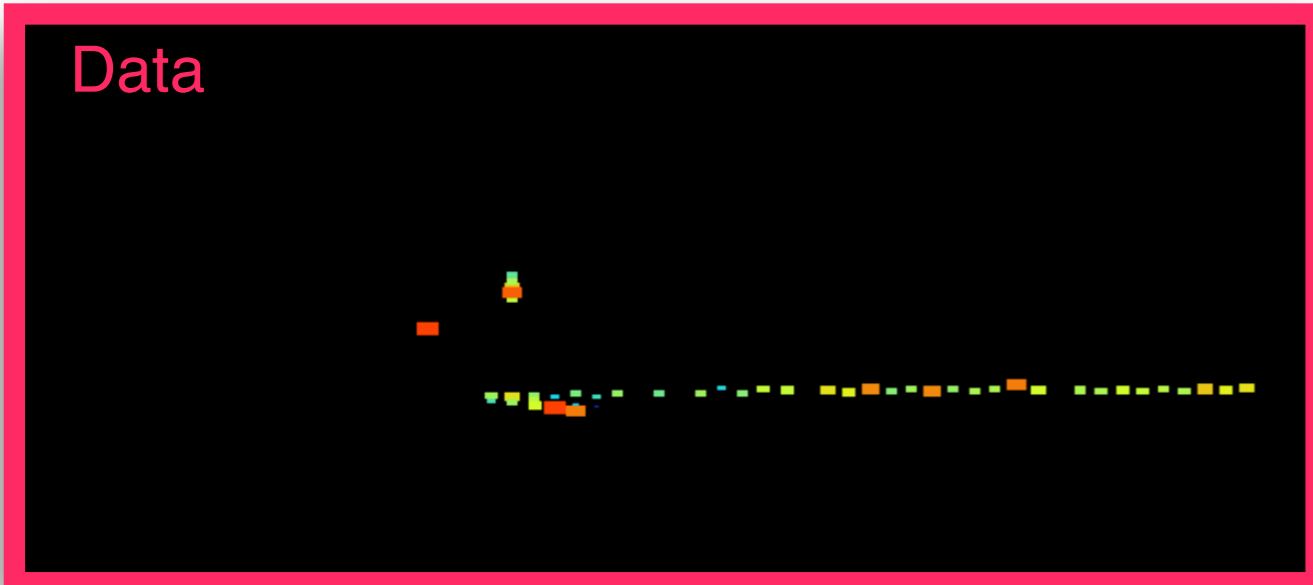
Bias introduced from MC training.

Mis-modeling translating into data/MC differences

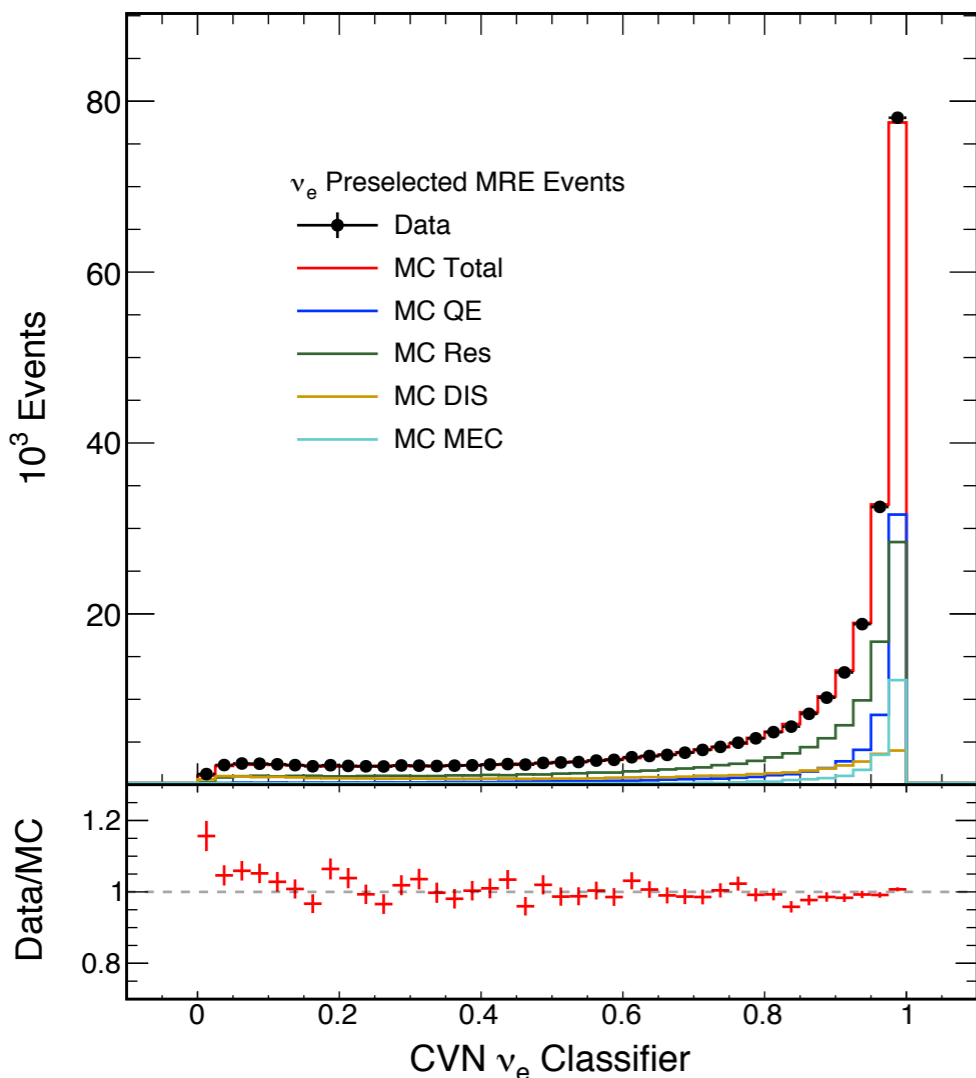
Uncertainties and other practical considerations



Data-driven test example



NOvA Preliminary



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	

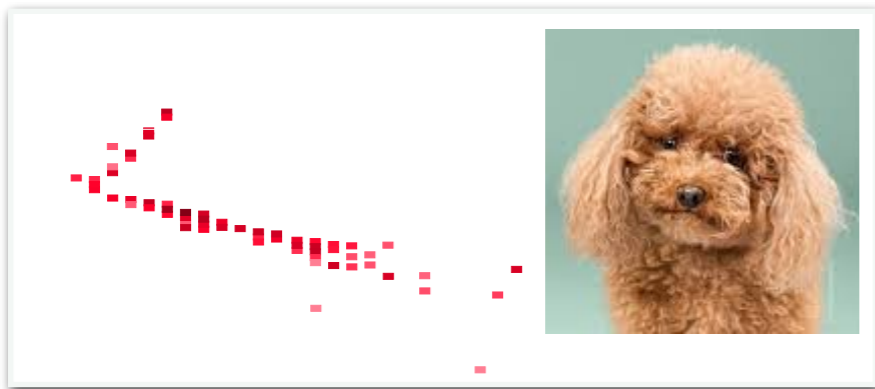
Uncertainty & Bias assessment



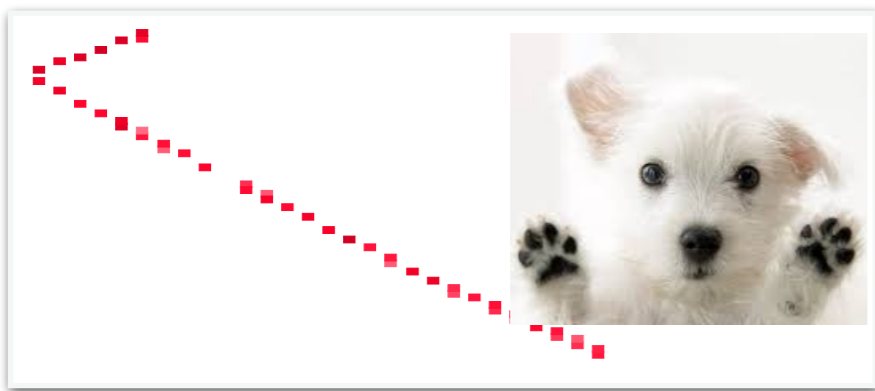
We currently **apply the same systematic** treatment through the whole analysis, including our DL algorithms.



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.

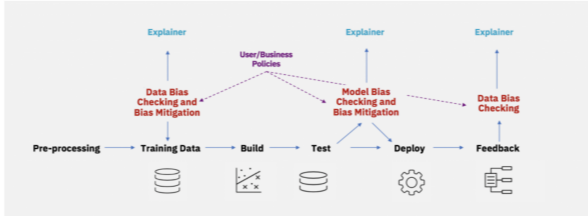
by Karen Hao February 4, 2019

M 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



```
graph LR; subgraph ML_Process; direction LR; P[Pre-processing] --> T[Training Data]; T --> B[Build]; B --> Te[Test]; Te --> D[Deploy]; D --> F[Feedback]; end; subgraph Bias_Mitigation; direction TB; U[User/Business Policies] --> DBCM[Data Bias Checking and Bias Mitigation]; U --> MBCM[Model Bias Checking and Bias Mitigation]; U --> DBC[Data Bias Checking]; end; DBCM -.-> T; MBCM -.-> Te; DBC -.-> F; E1[Explainer] -.-> T; E2[Explainer] -.-> Te; E3[Explainer] -.-> F;
```


PC REVIEWS BEST PICKS HOW-TO NEWS SMART HOME BUSINESS SHOP

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.

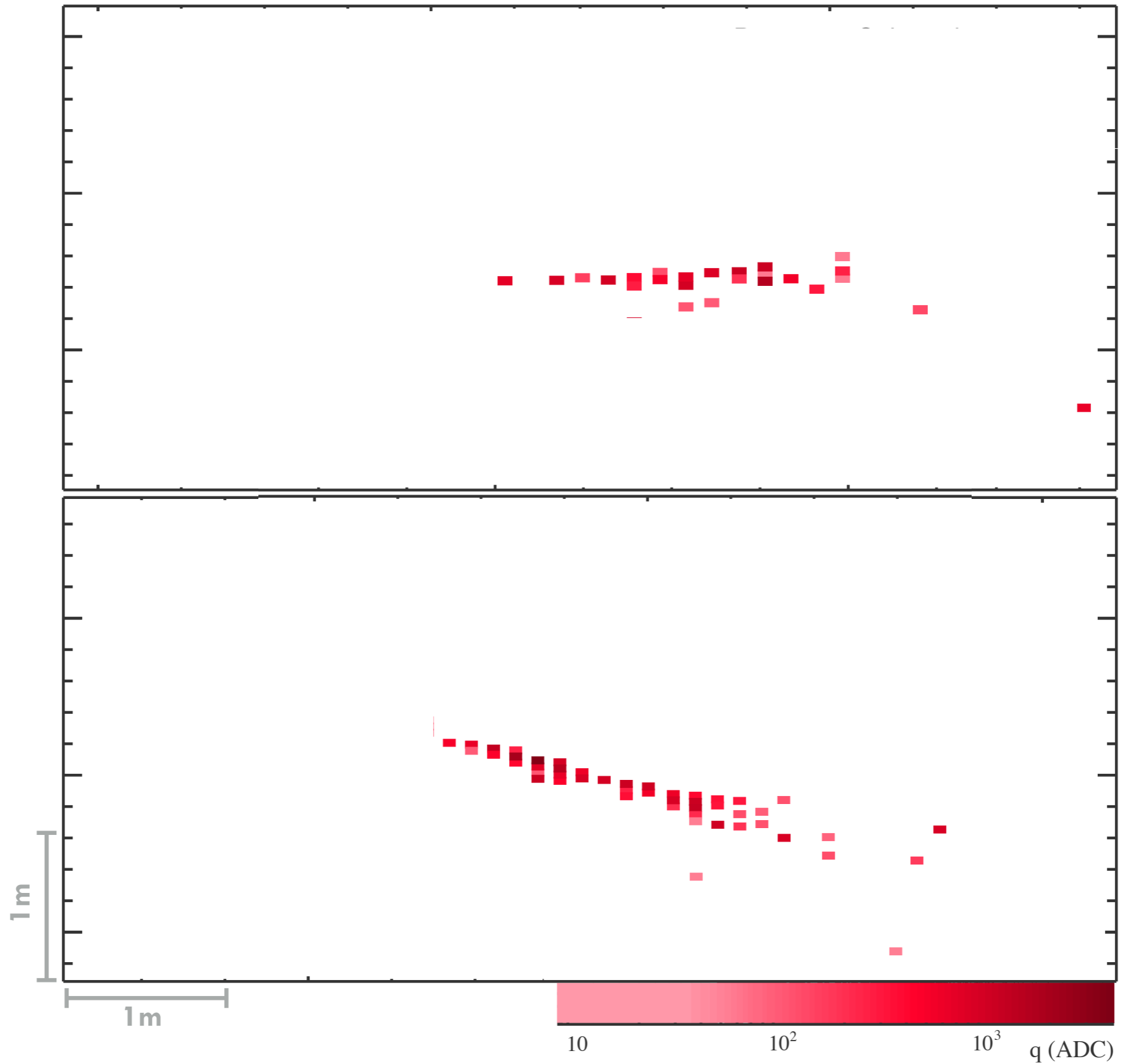
By Ben Dickson June 14, 2018 2:00PM EST

163 SHARES

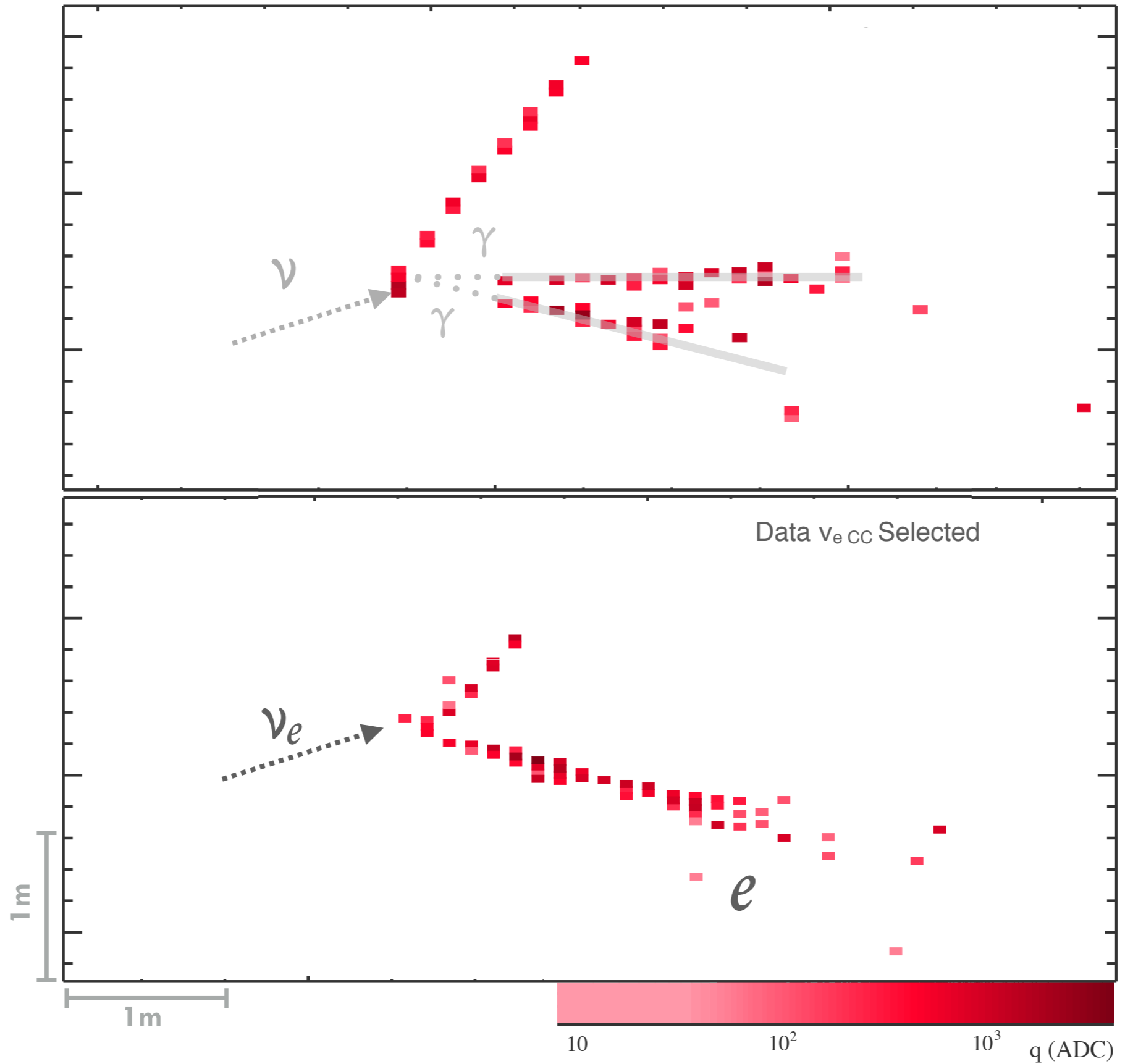




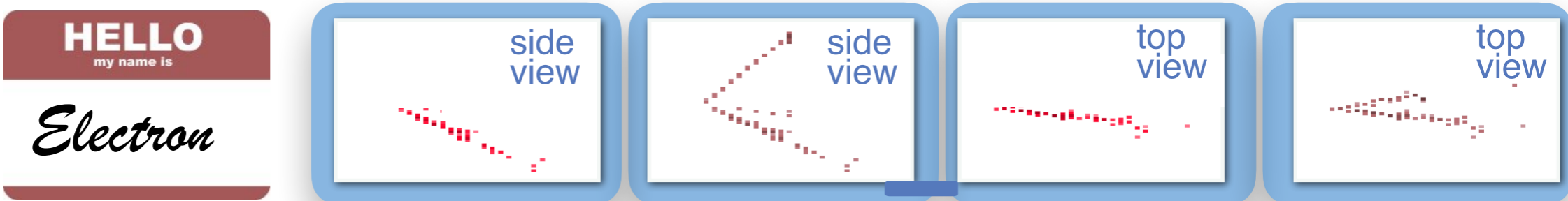
Particle Identification



Particle Identification



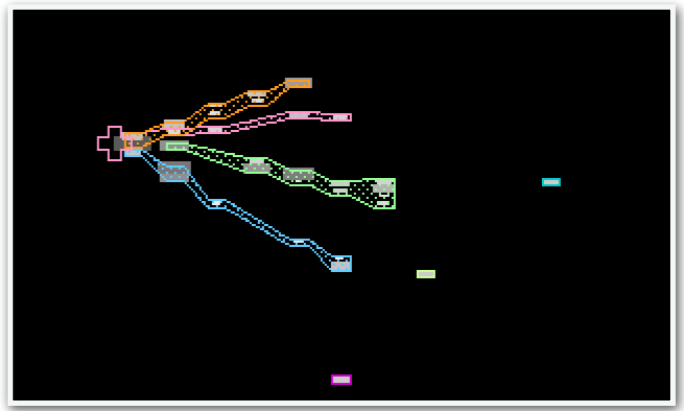
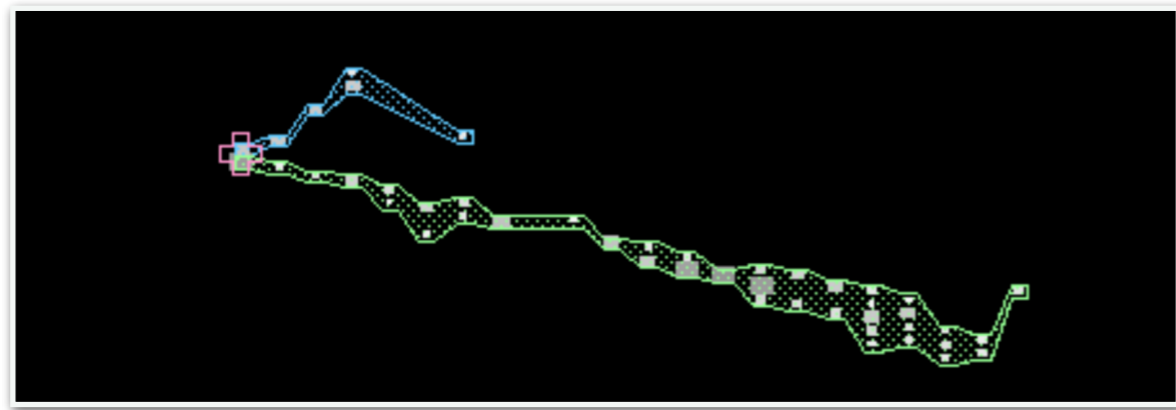
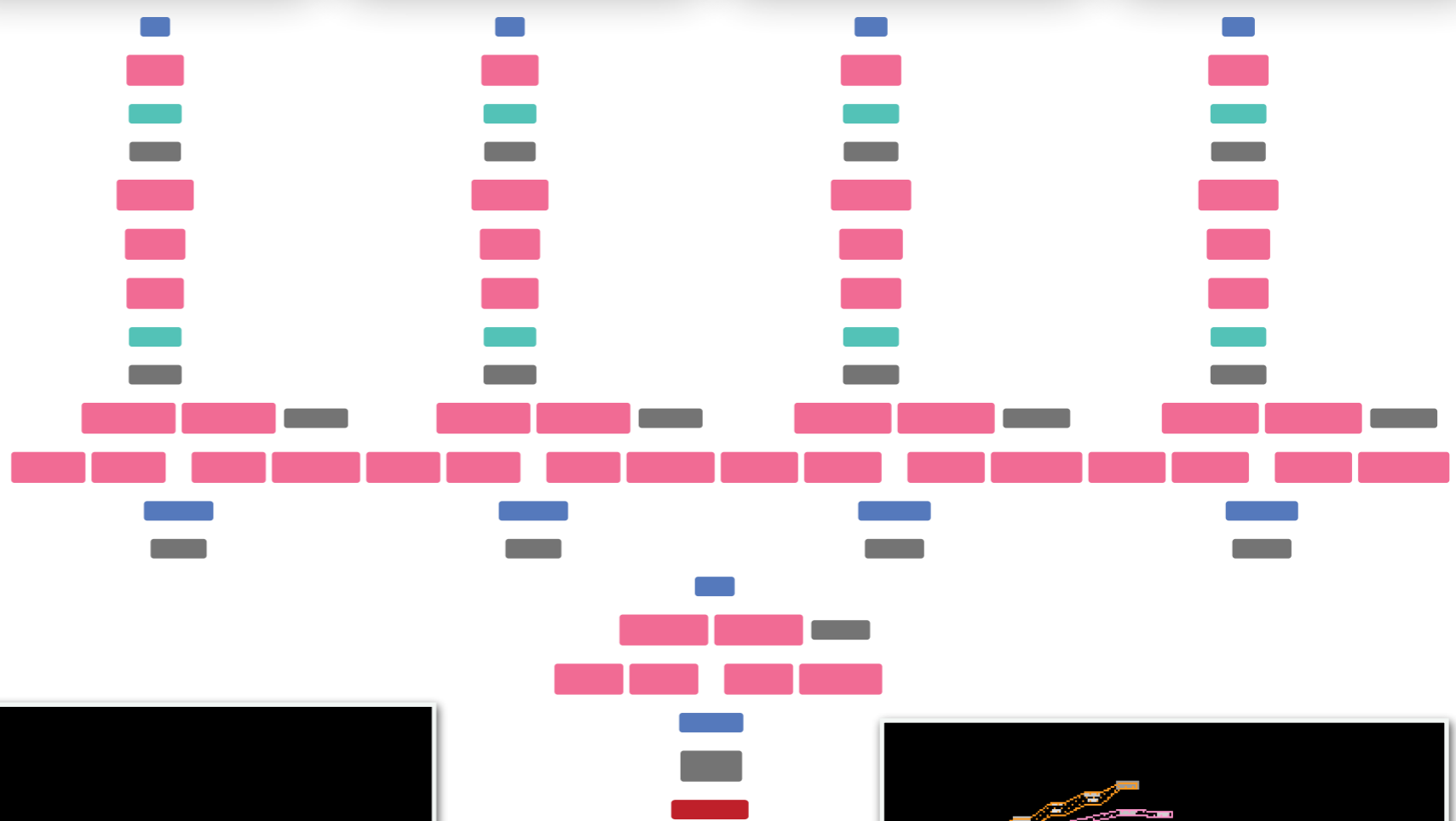
Context Improves Classification



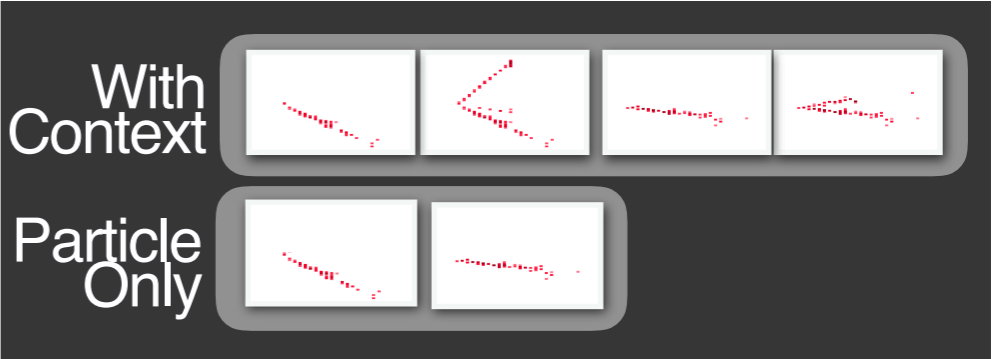
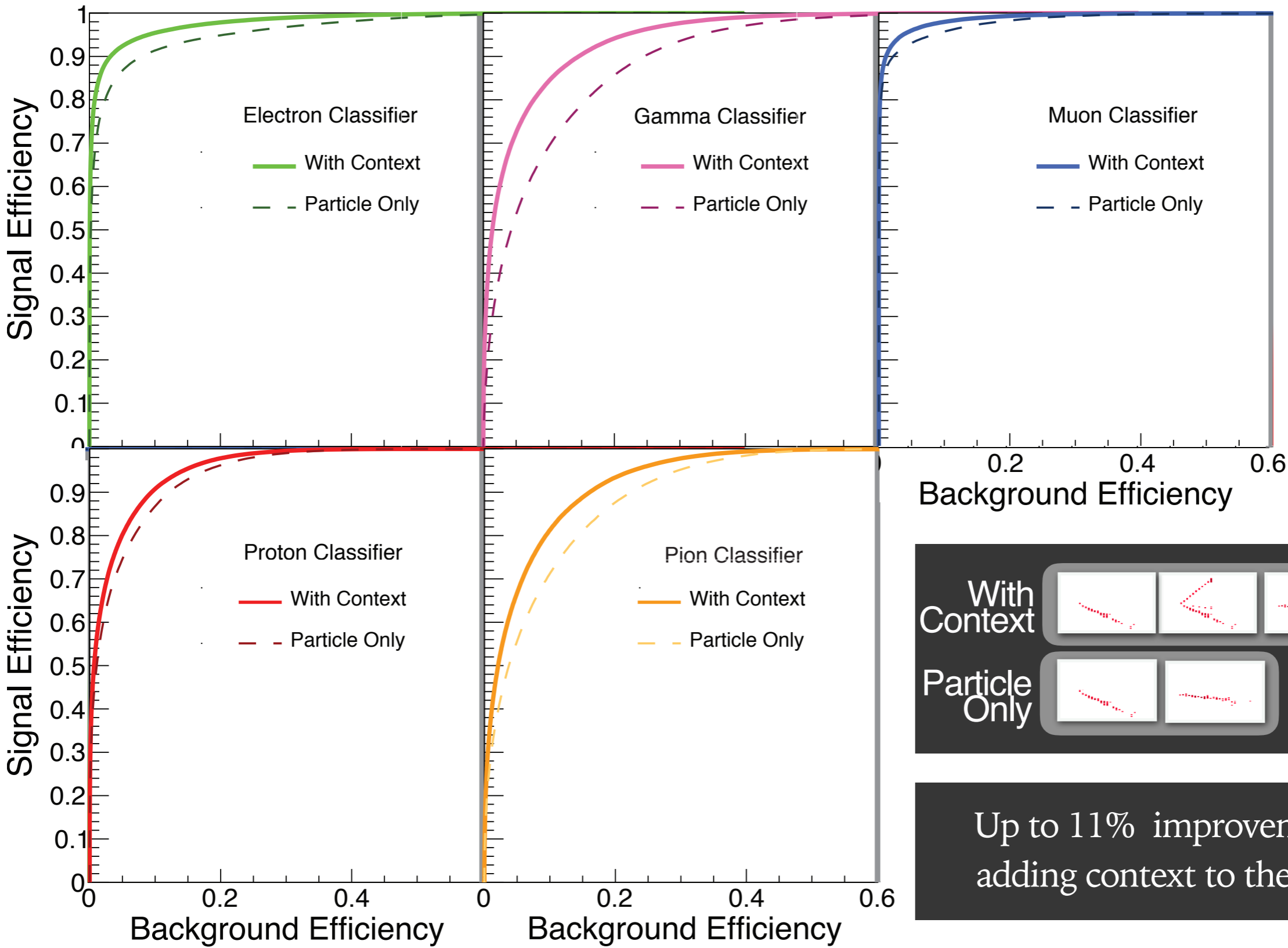
Using the existing reconstruction, classify clusters of hits.

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

Trained on clusters from all events above some minimum purity.

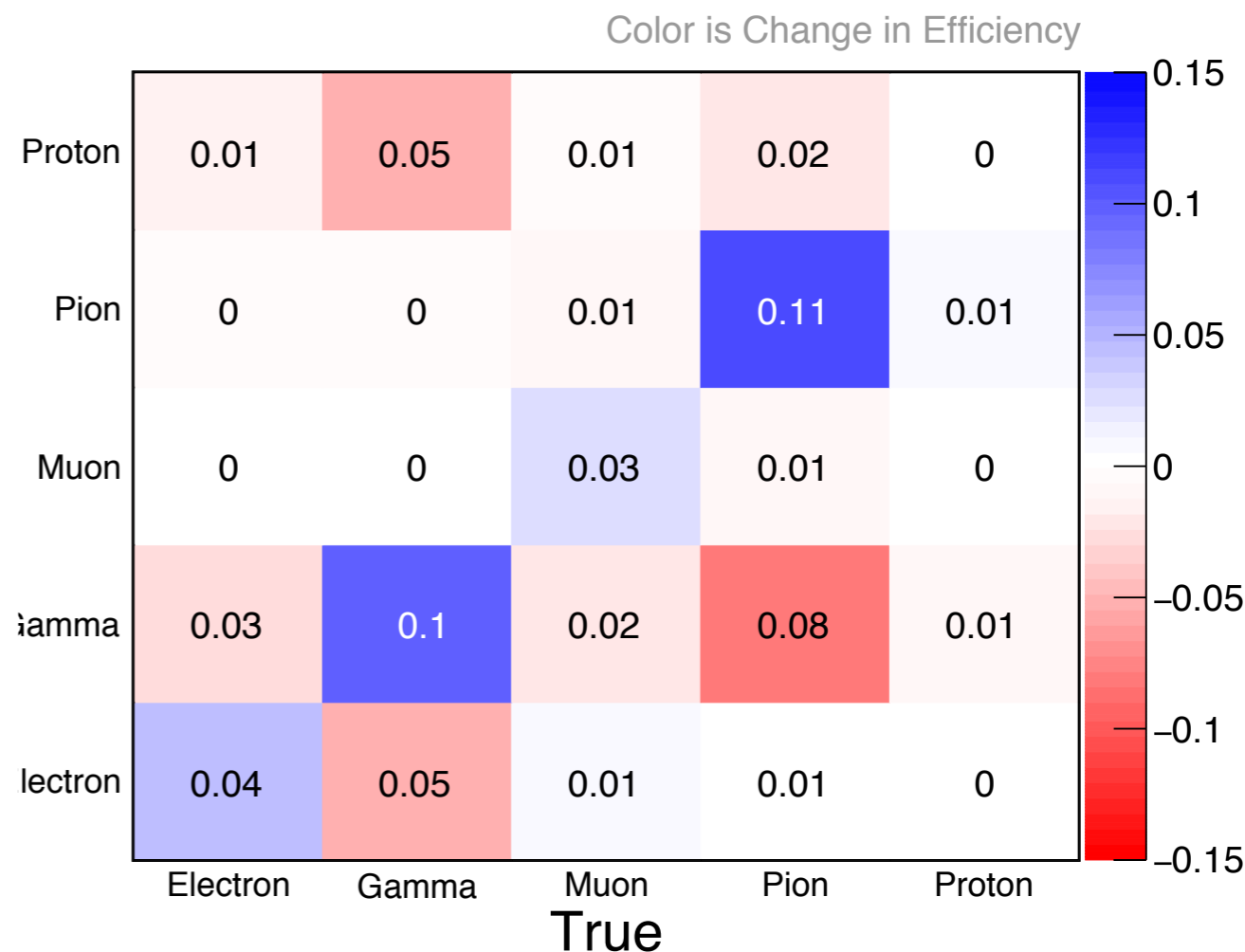
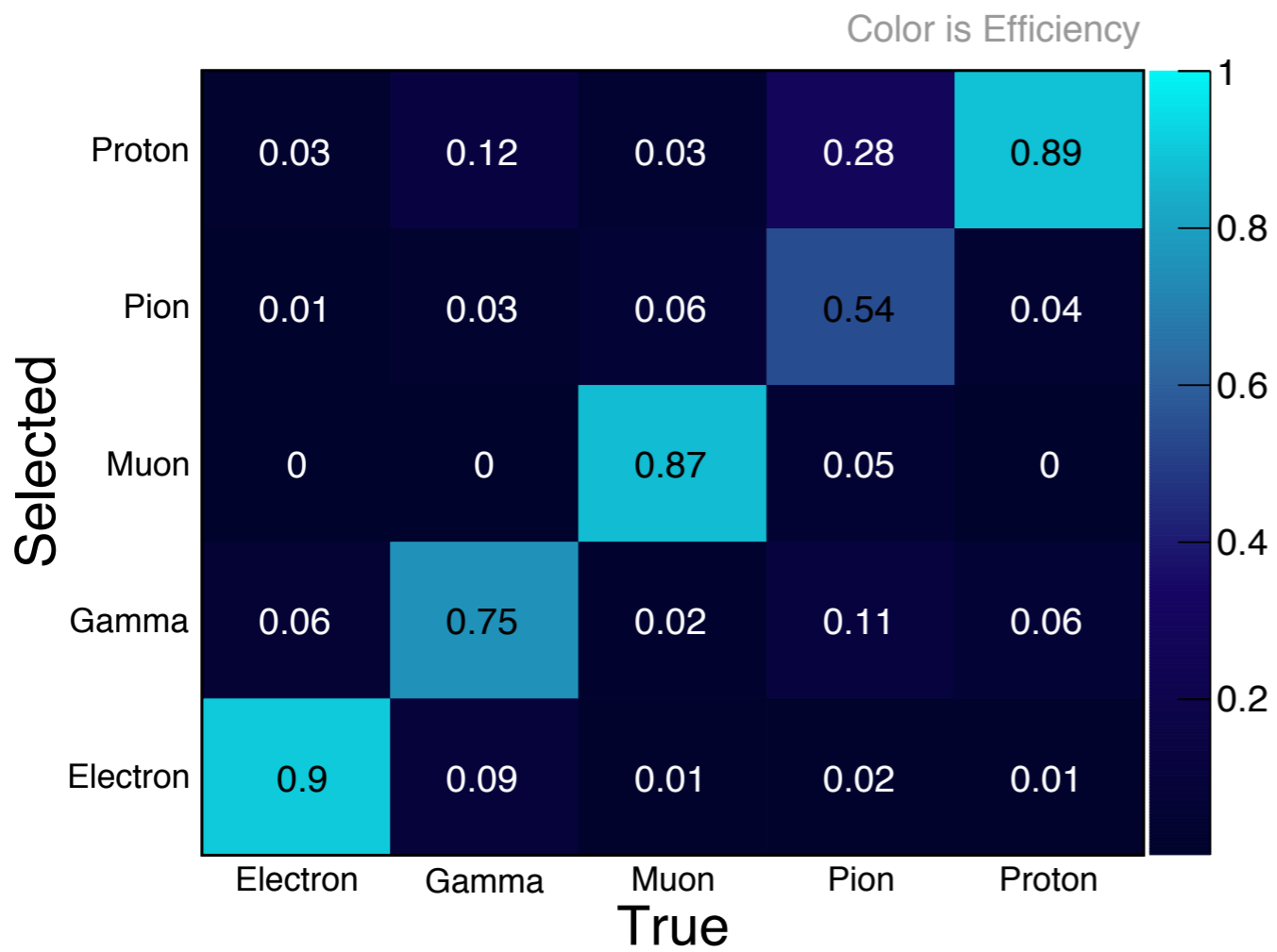


Context Improves Classification



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

Context Improves Classification

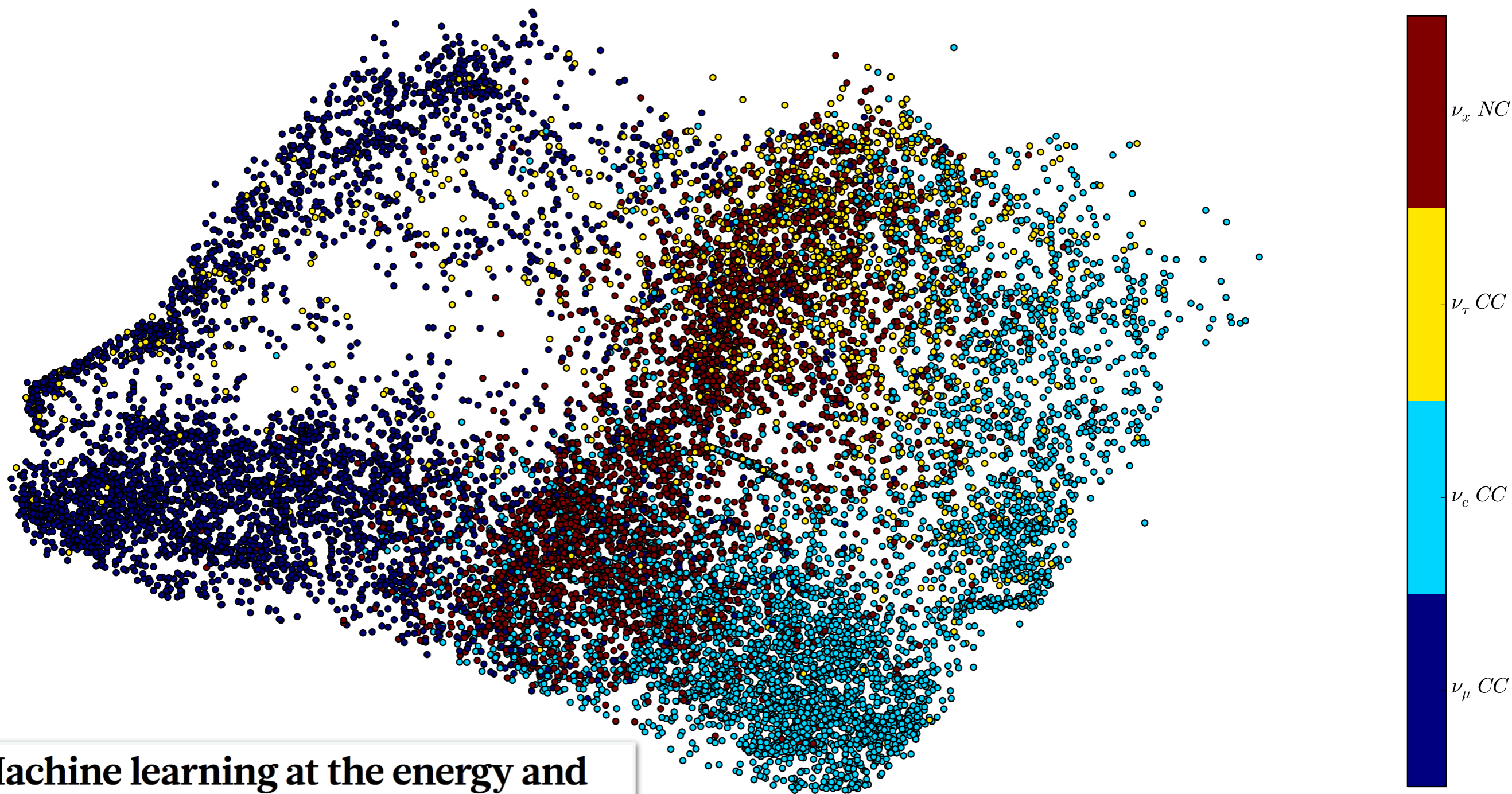


Color is efficiency taking the highest softmax output.

Color is difference in efficiency taking the highest softmax output.

Larger with context

Larger w/o context

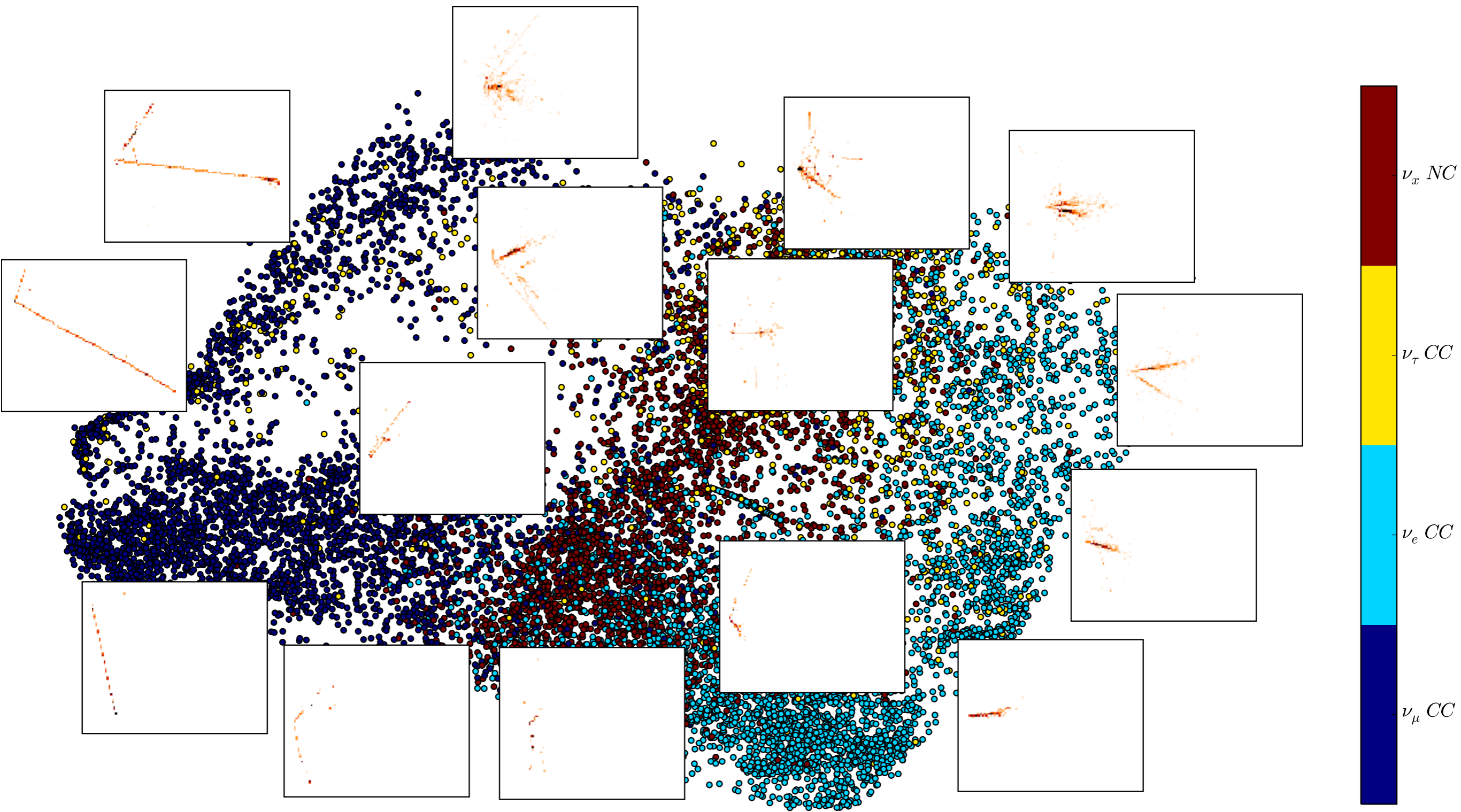


Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic^{1*}, Mike Williams^{2*}, David Rousseau³, Michael Kagan⁴, Daniele Bonacorsi^{5,6}, Alexander Himmel⁷, Adam Aurisano⁸, Kazuhiro Terao⁴ & Taritree Wongjirad⁹

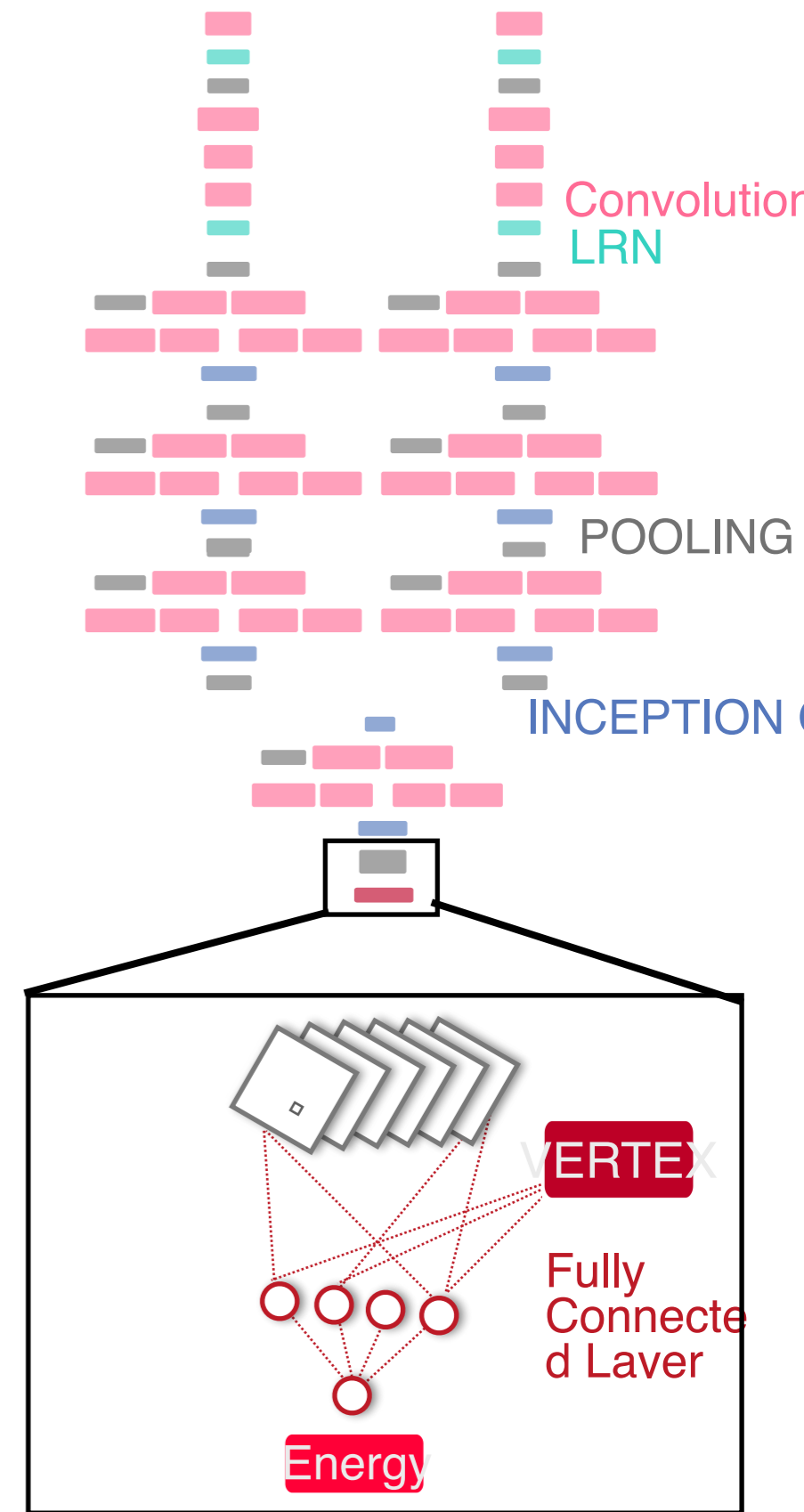
Nature Vol. 560

t-sne



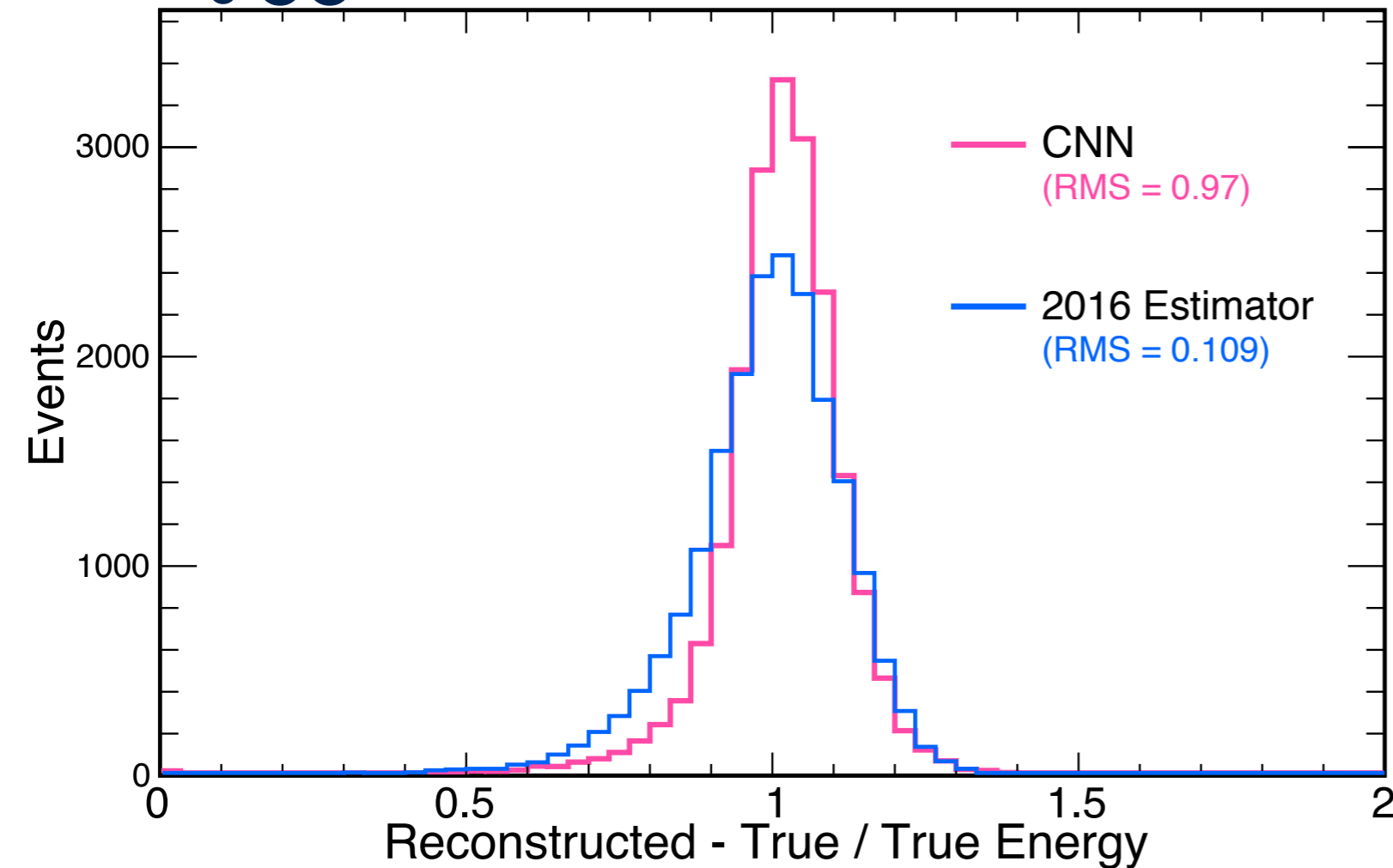


CNNs + Regression for Energy Reconstruction



ν_e CC

NOvA Simulation

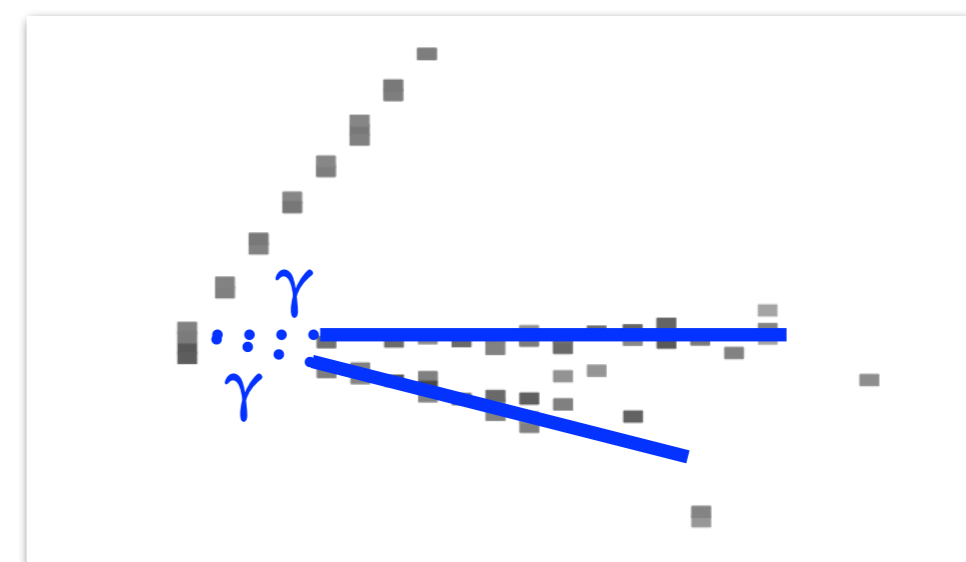
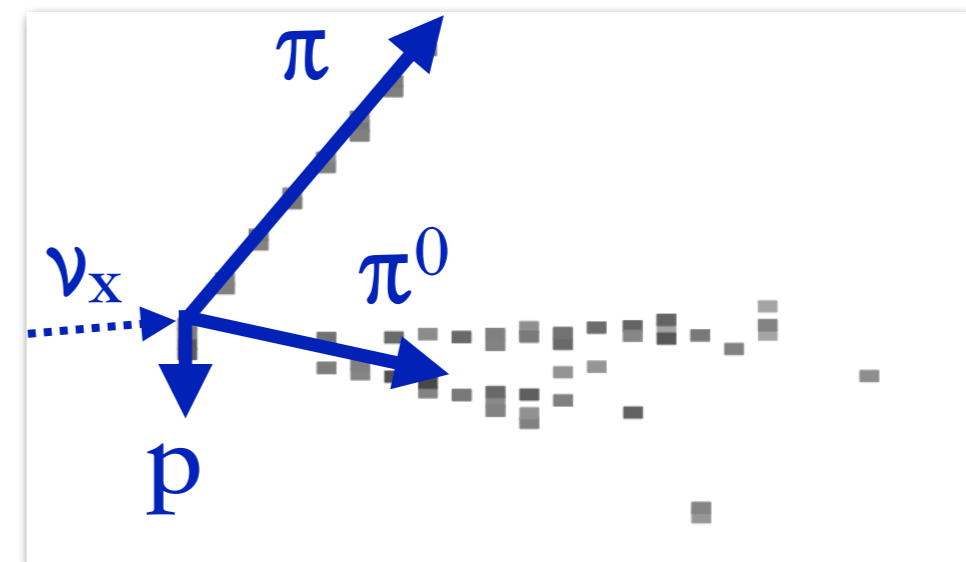
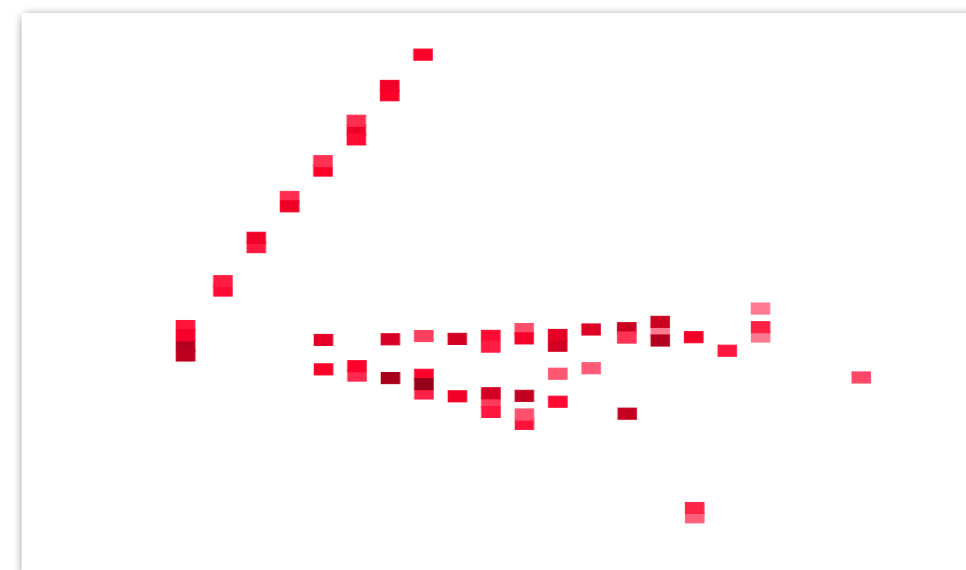
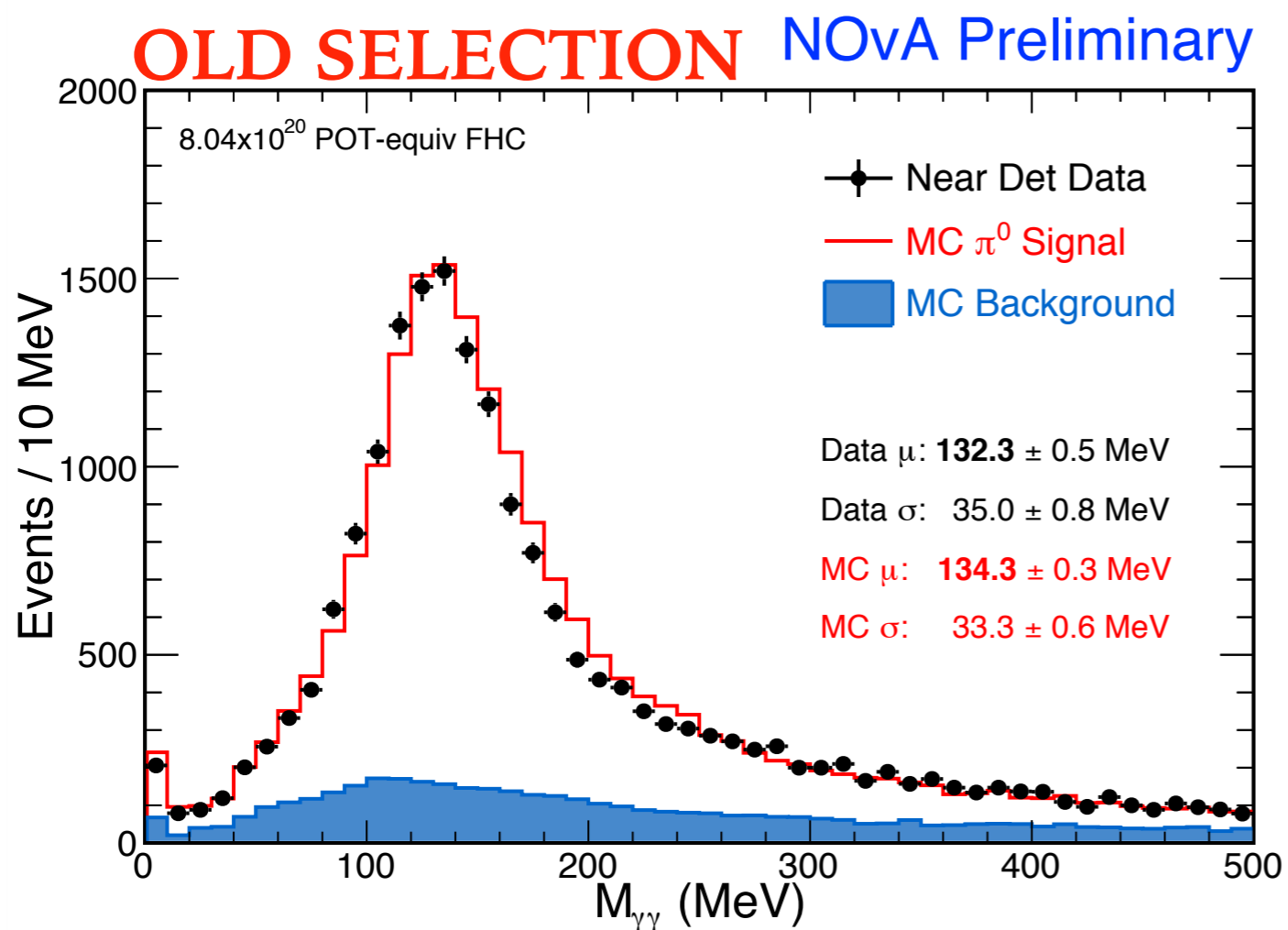


The target is Energy instead of PID values

Incorporate information from the reconstruction at the fully connected stage.

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

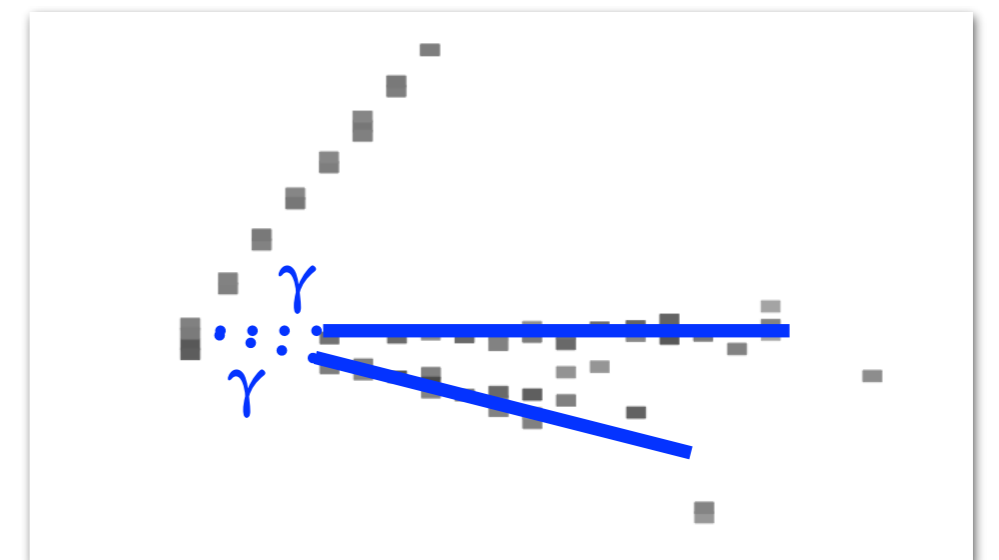
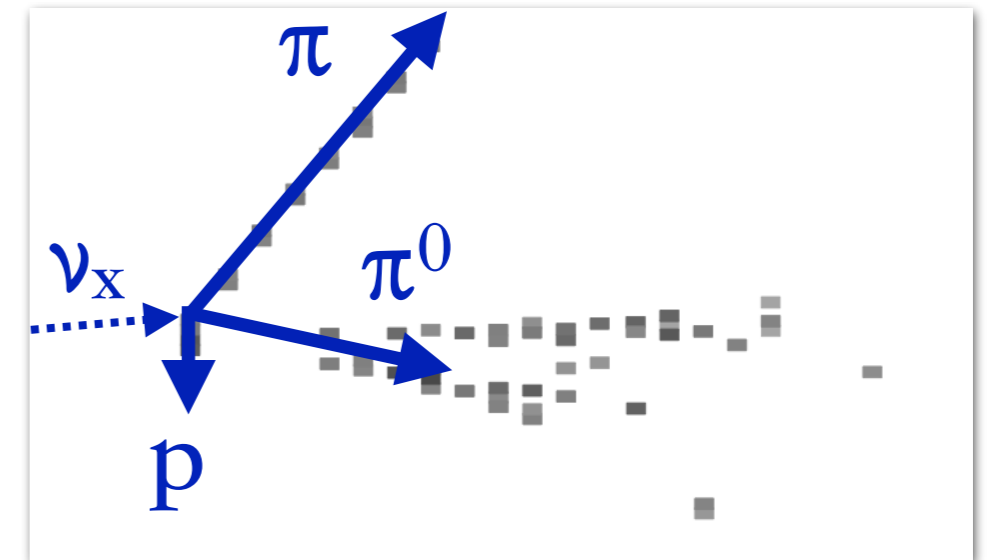
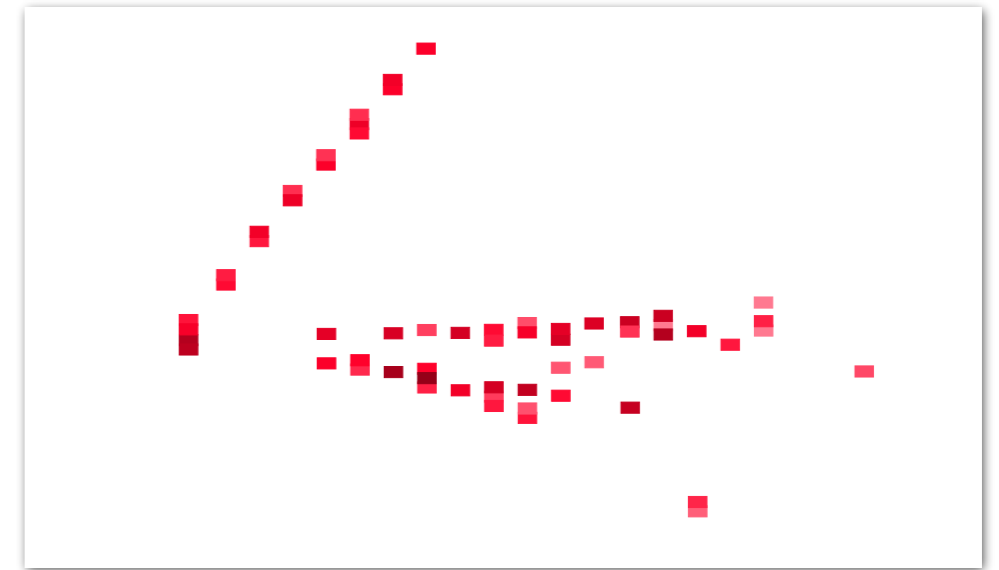
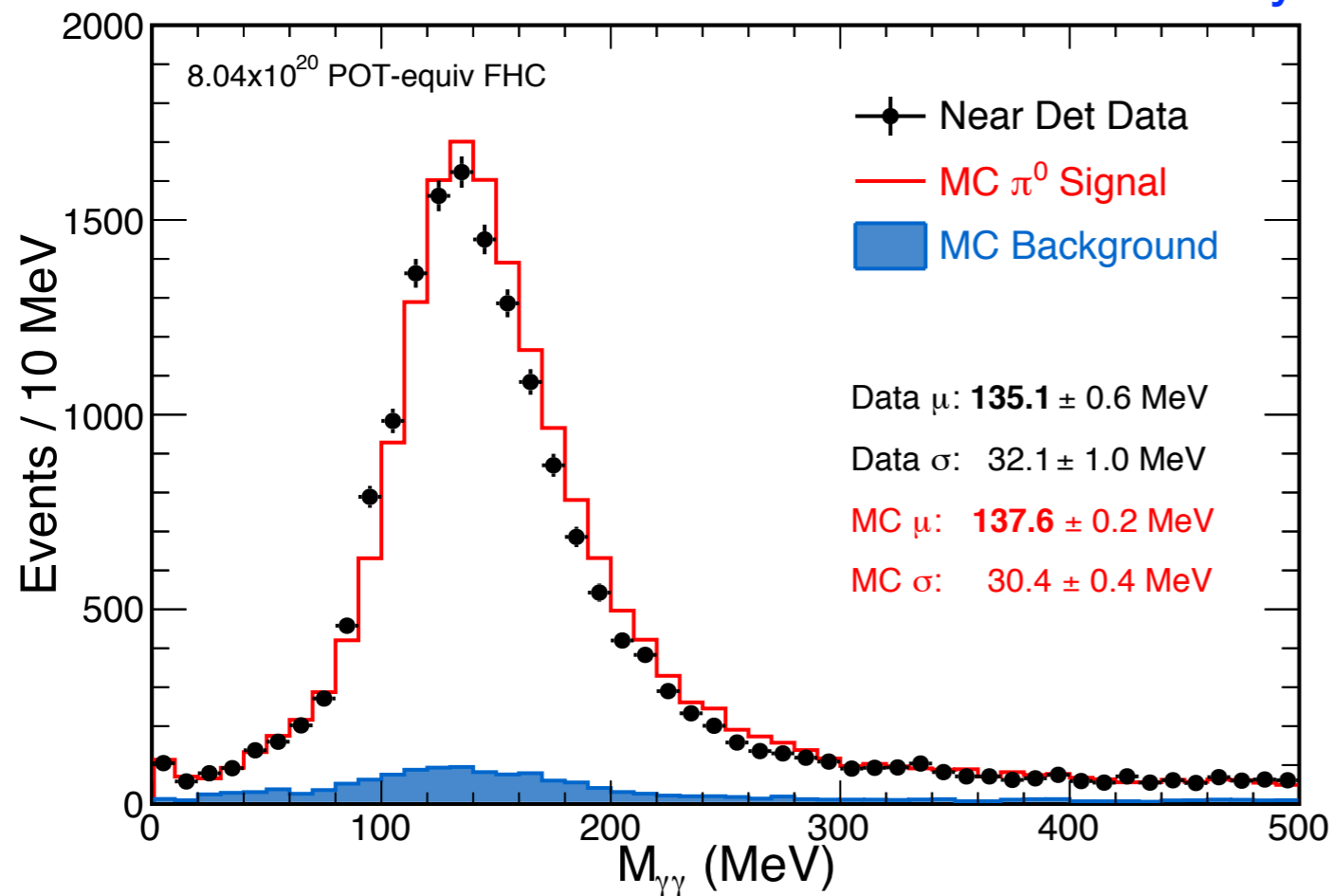
This is a handle on the dead material correction for electromagnetic deposits.



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

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

NOvA Preliminary

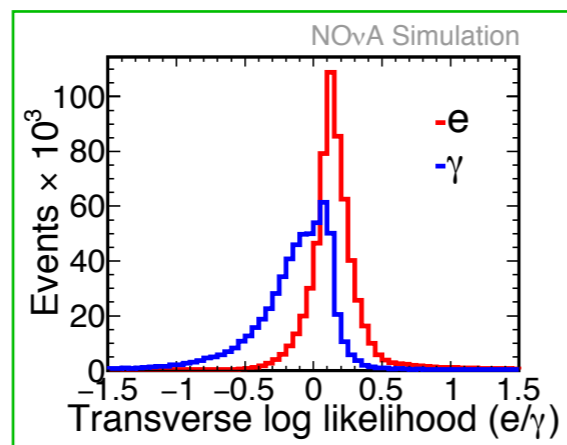
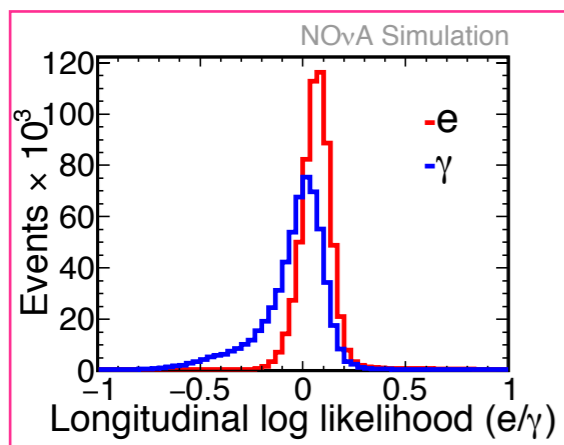
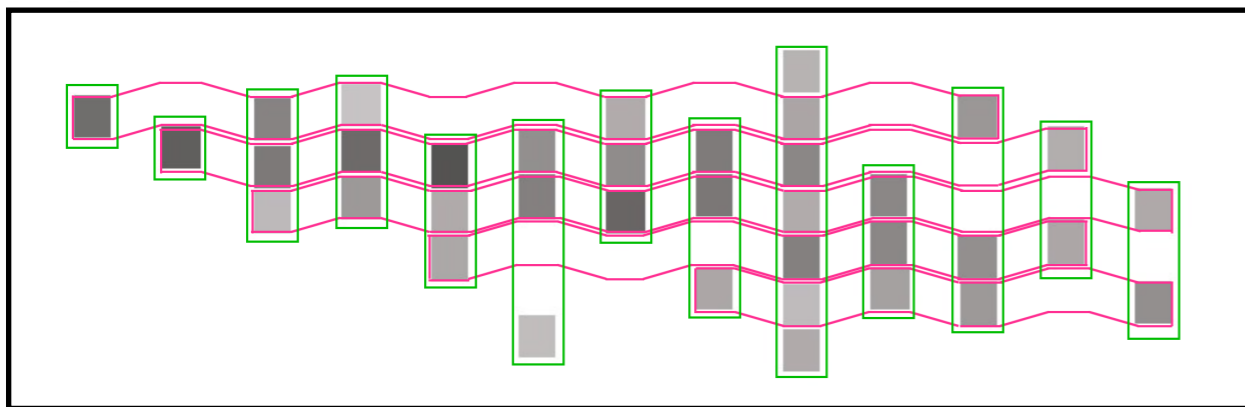


Neutrino Identification (First Analysis)

LID

Likelihood *I*dentification

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



In practice:

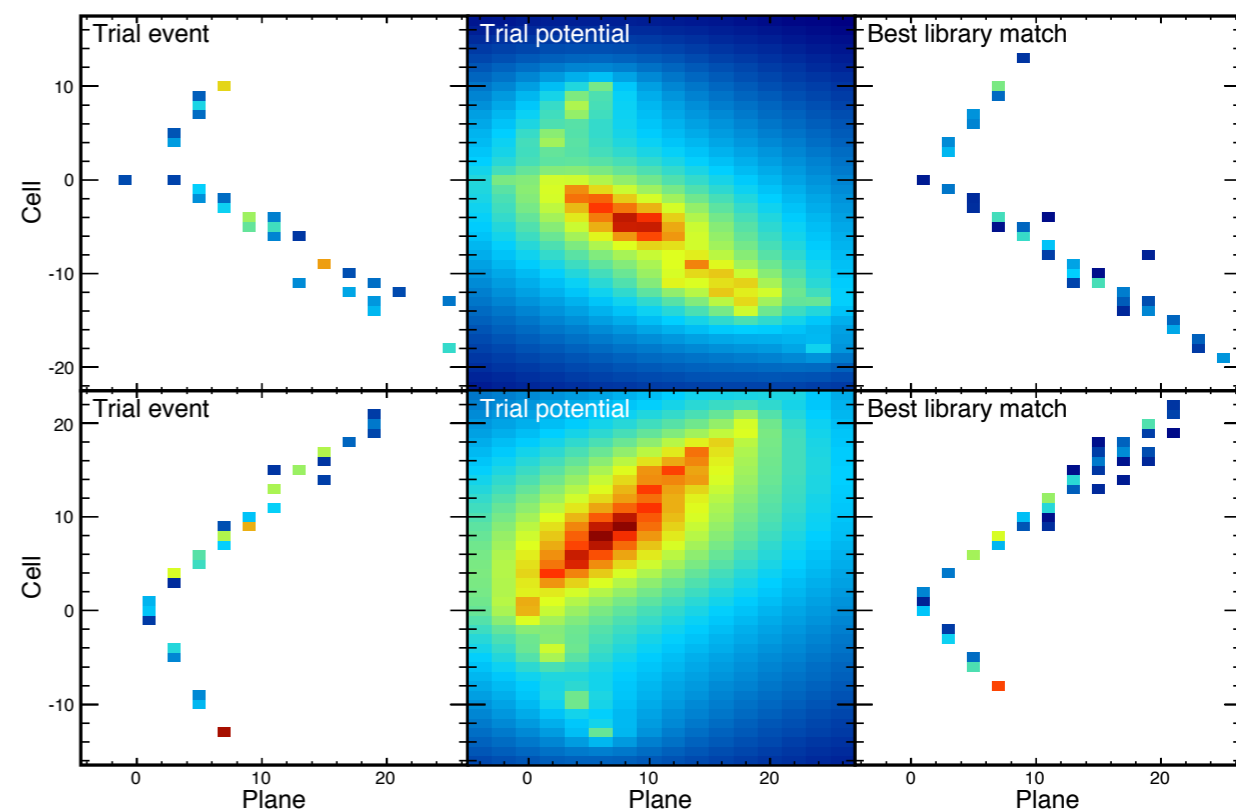
- ★ Reconstruct electron shower.
- ★ Find likelihoods from its dE/dx profiles compared to particle hypotheses.

Likelihoods → Neural Network

LEM

Library *E*vent *M*atching

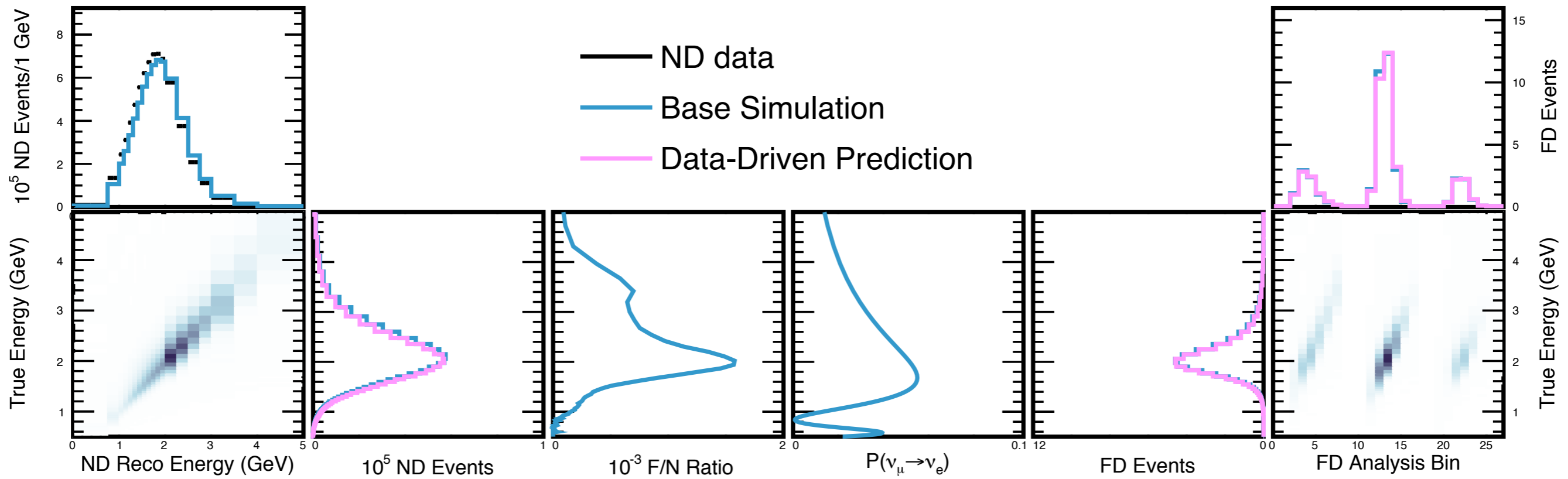
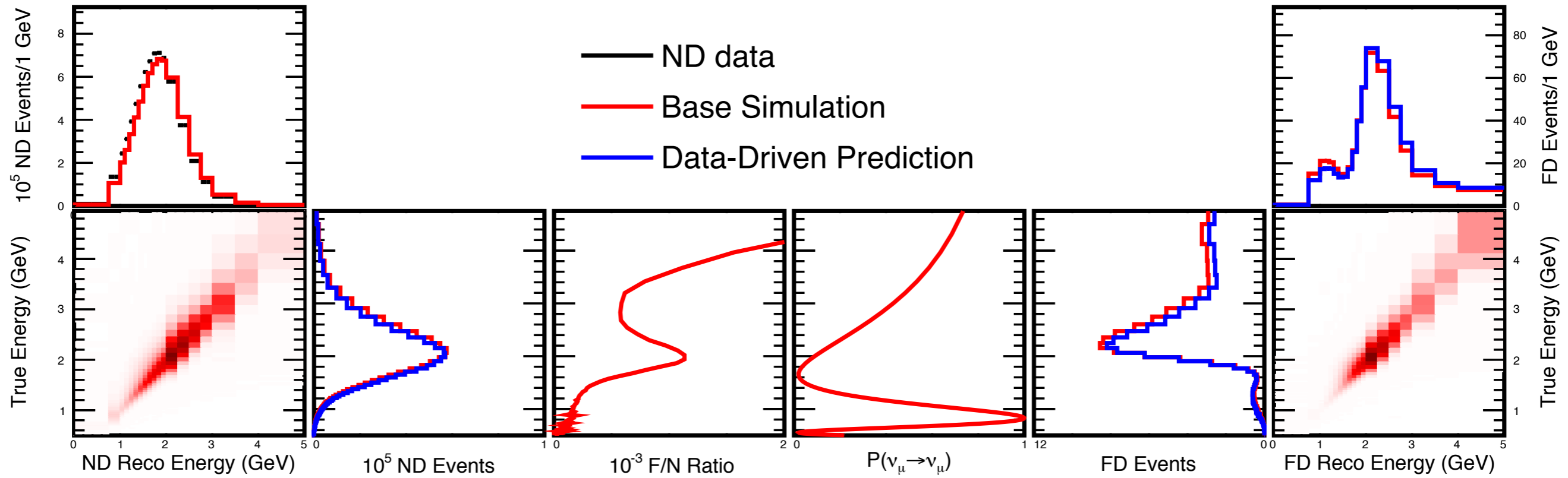
Premise: We have a large library of simulated event templates, large enough that we can use it to compare pixel by pixel.



In practice:

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

Features → Decision Tree



Introducing a Systematic Shift

