Pattern Recognition in KM3NeT:
A multi-dimensional challenge

João Coelho

for the KM3NeT Collaboration

21 February 2017

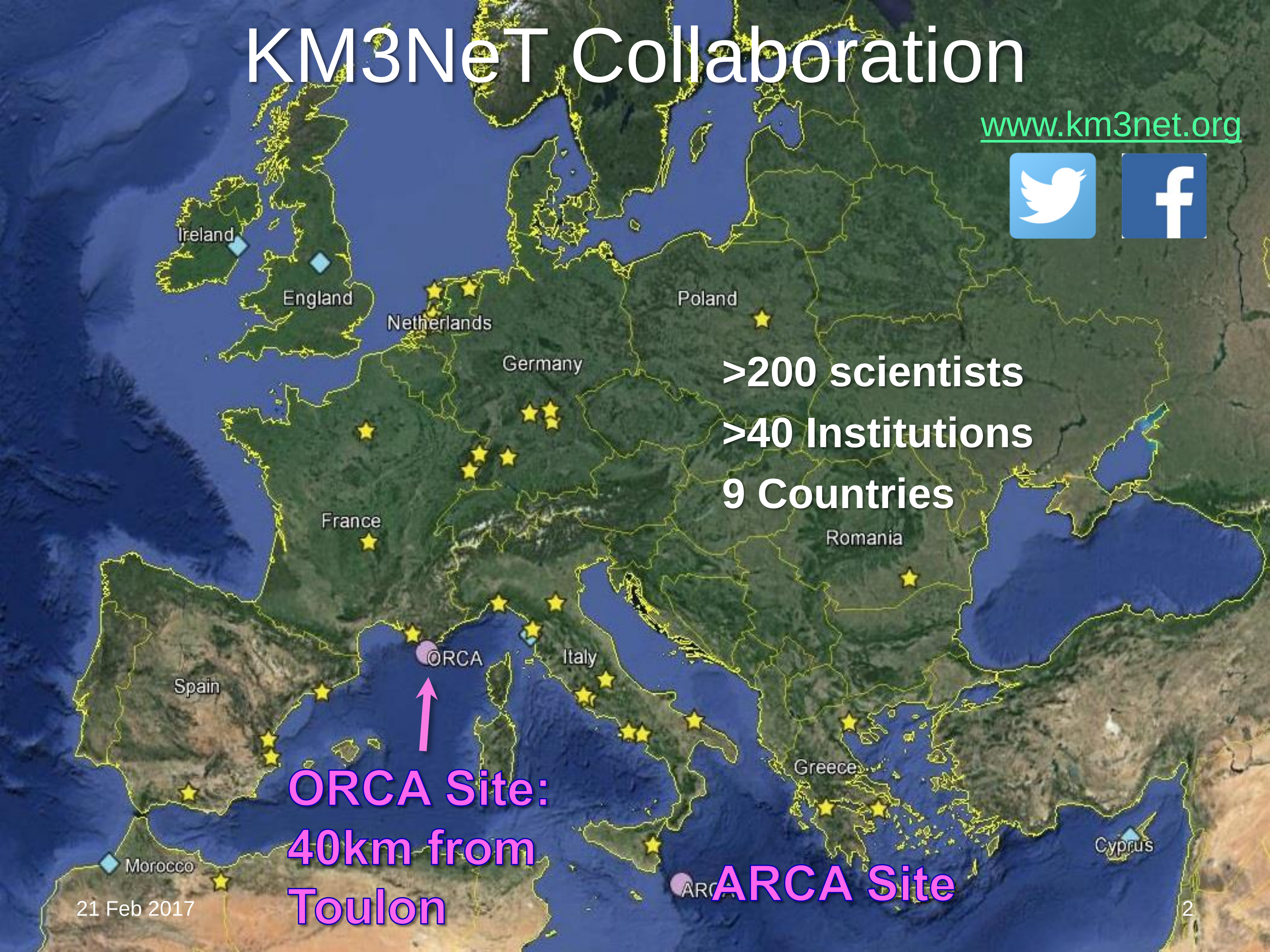
KM3NeT Collaboration

www.km3net.org

>200 scientists

>40 Institutions

9 Countries

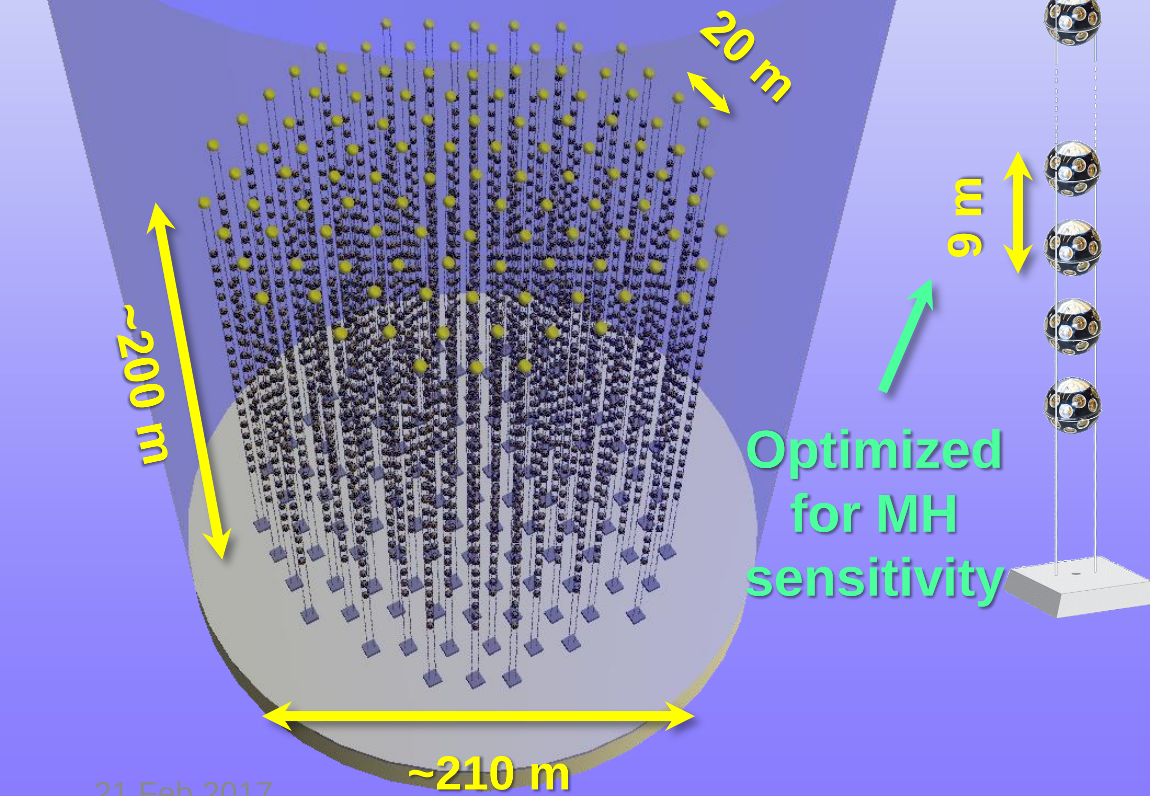


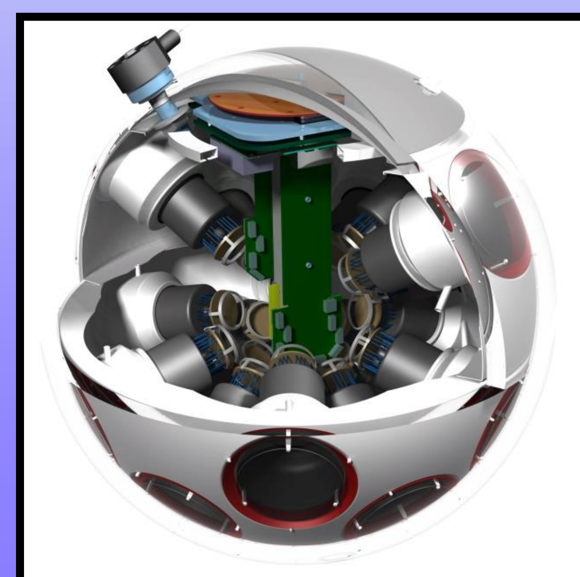
ORCA Site: 40km from Toulon

ARCA Site

The ORCA Detector

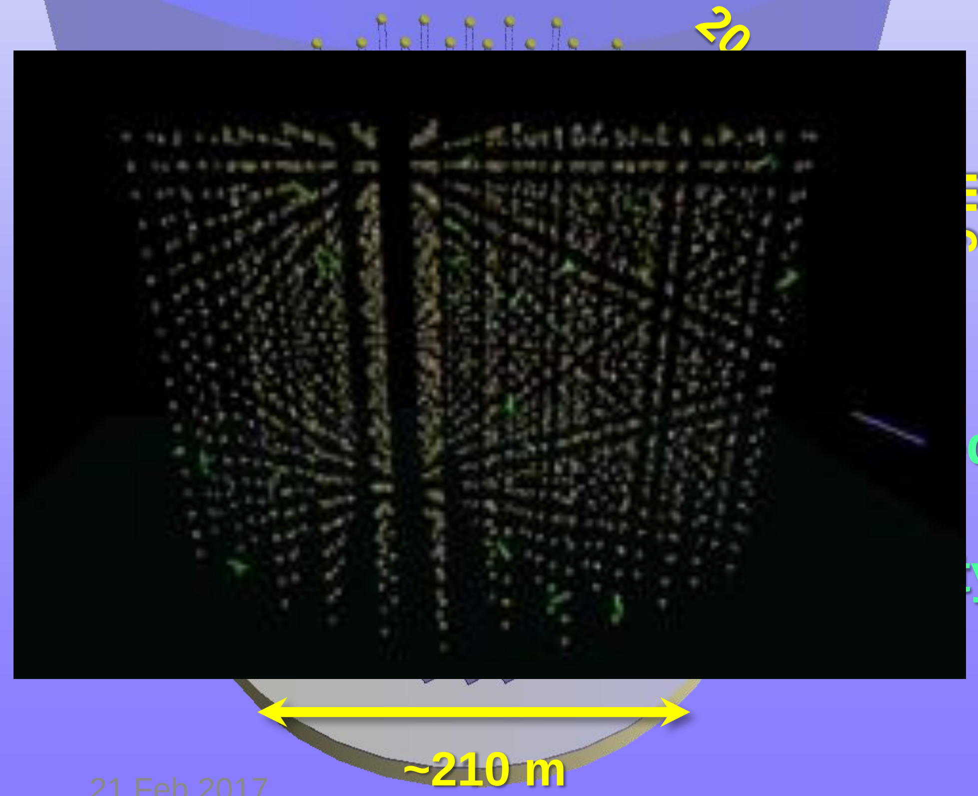
- **~6 Mt** instrumented
- **115** strings
- **18** DOMs / str
- **31** PMTs / DOM
- Total: **64k PMTs**

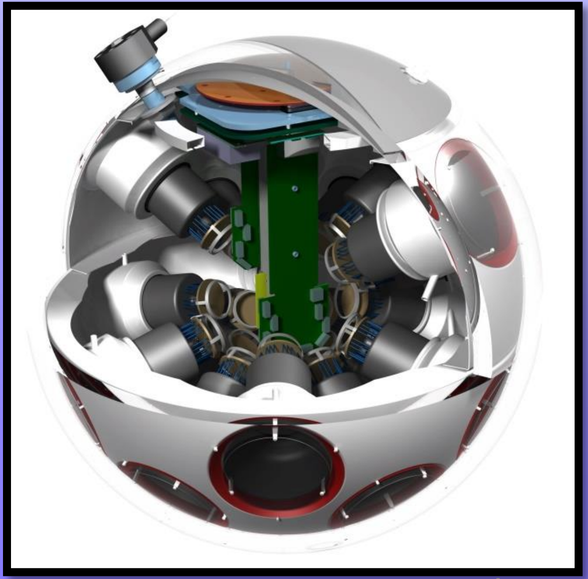




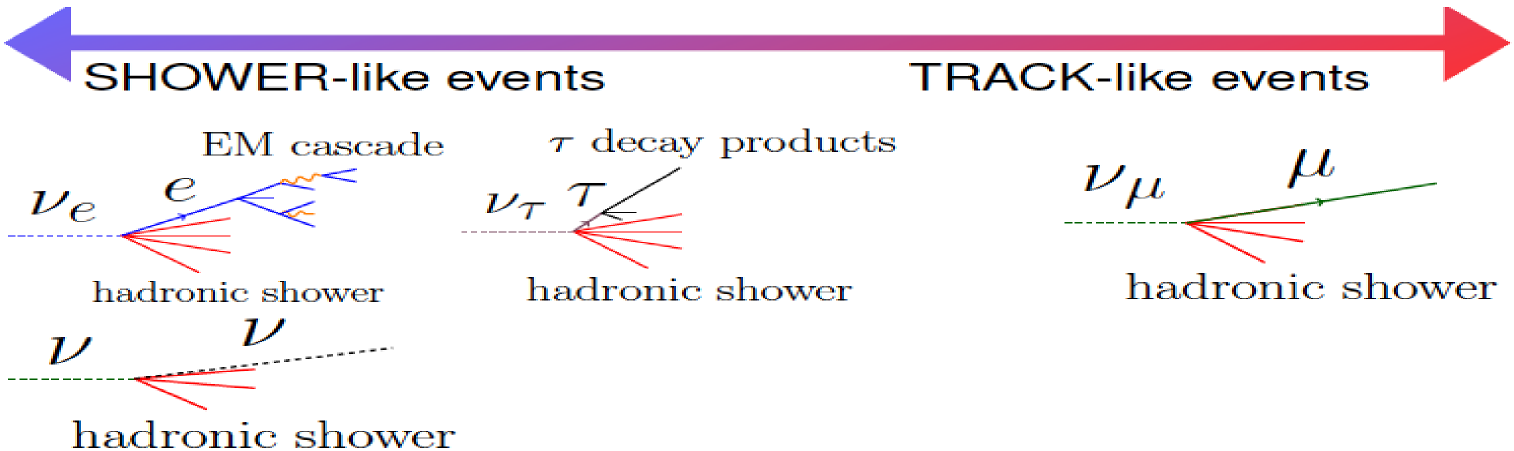
The ORCA Detector

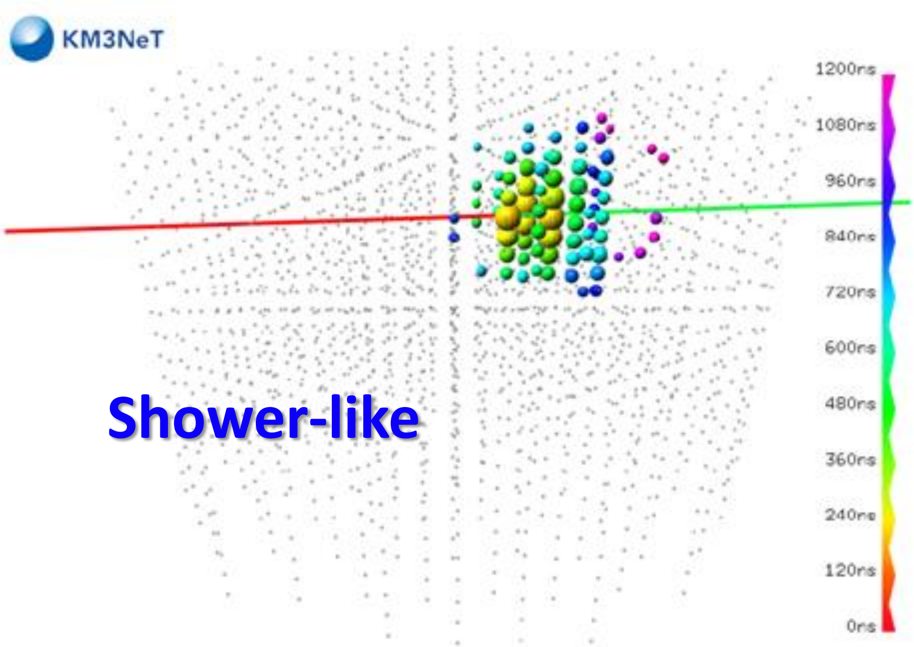
- **~6 Mt** instrumented
- **115** strings
- **18** DOMs / str
- **31** PMTs / DOM
- Total: **64k PMTs**

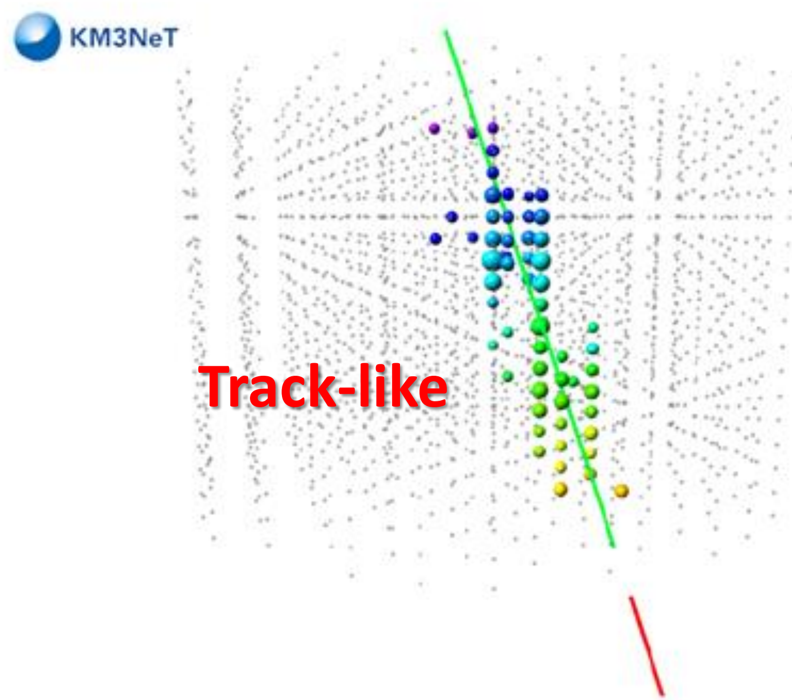




Measuring Neutrinos

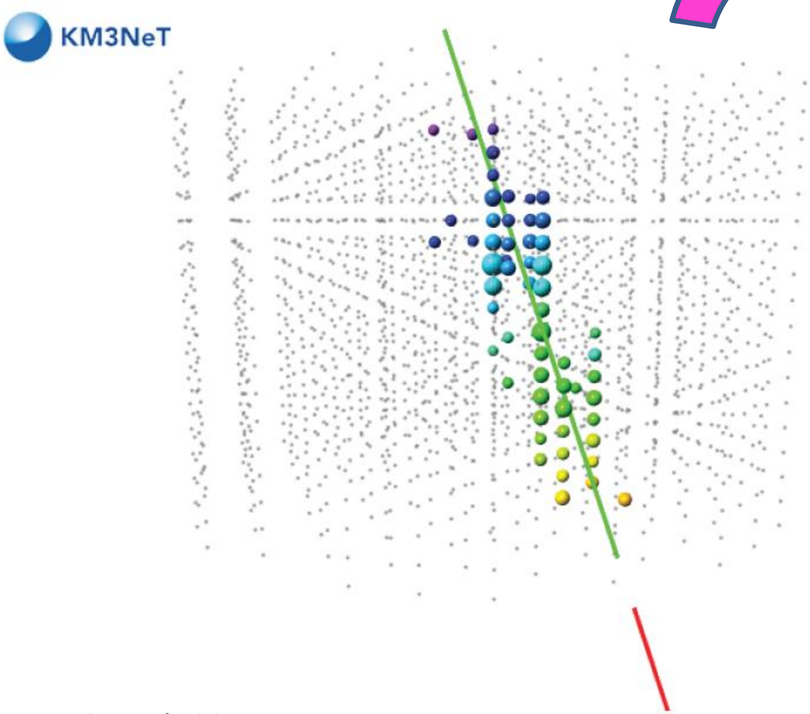


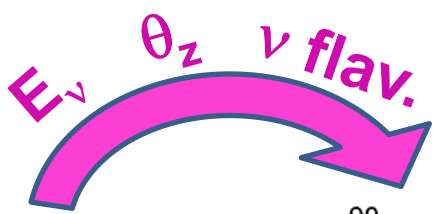


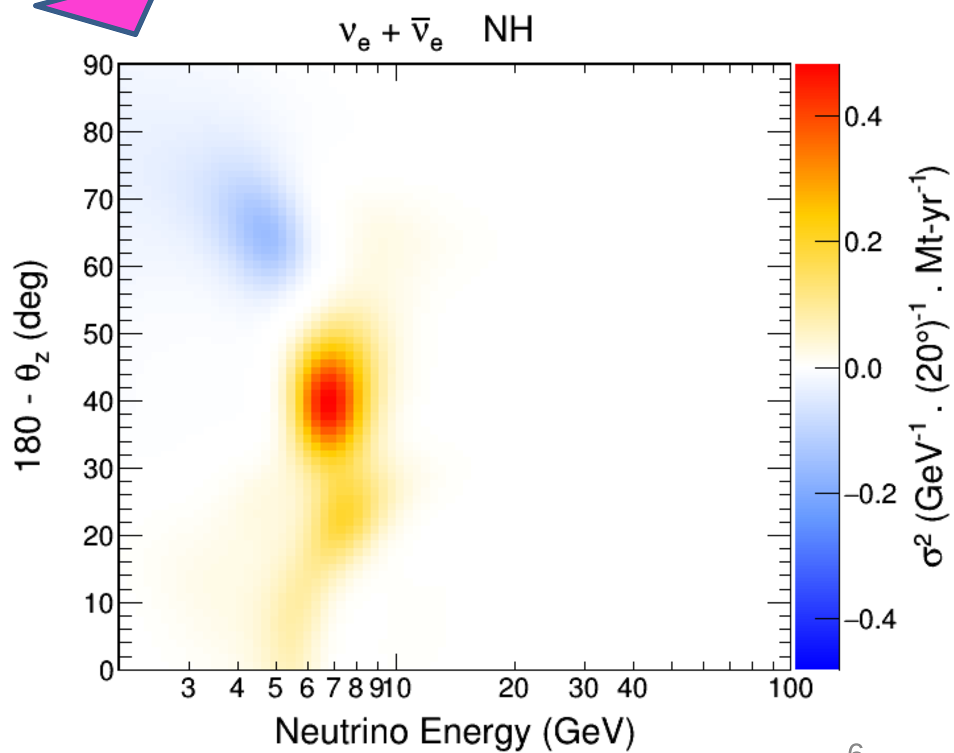


Motivation

- KM3NeT detectors form 3D images (6D maybe?)
- Our goal is to extract physics from these images
- Can we train a neural network to do the job for us?

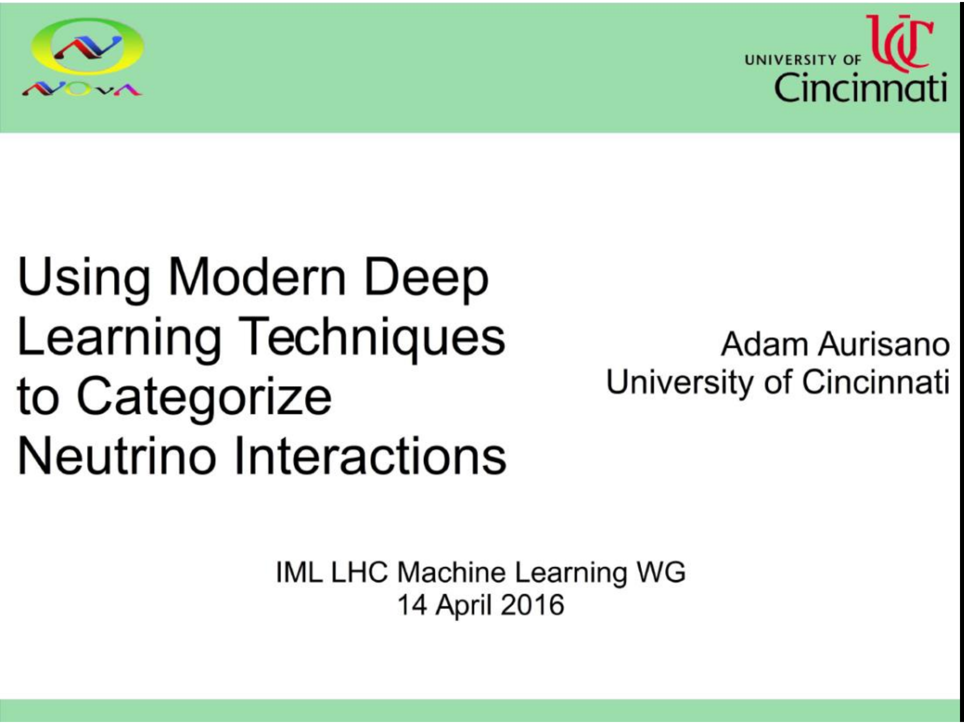


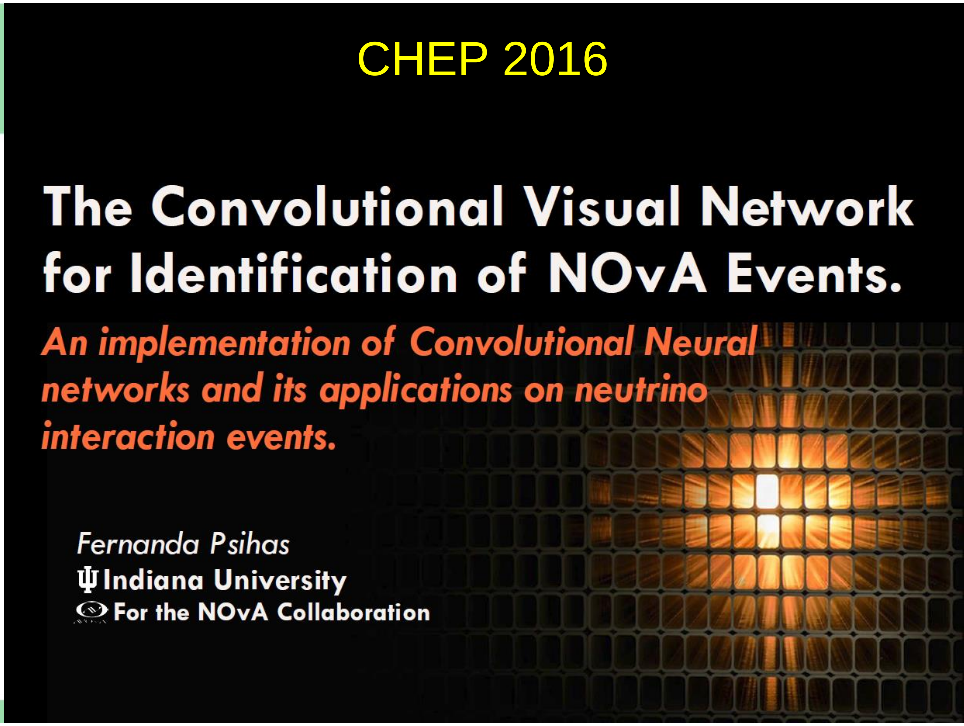




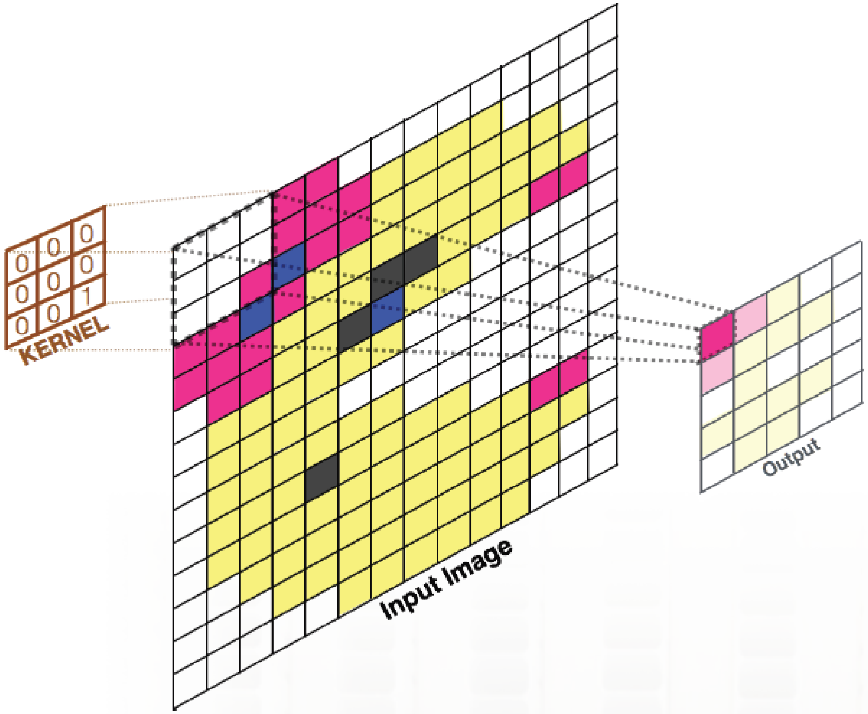
Inspired by:

- Convolutional Neural Network work developed by NOvA
- Paper: [arXiv:1604.01444]
- The following slides were taken from talks below



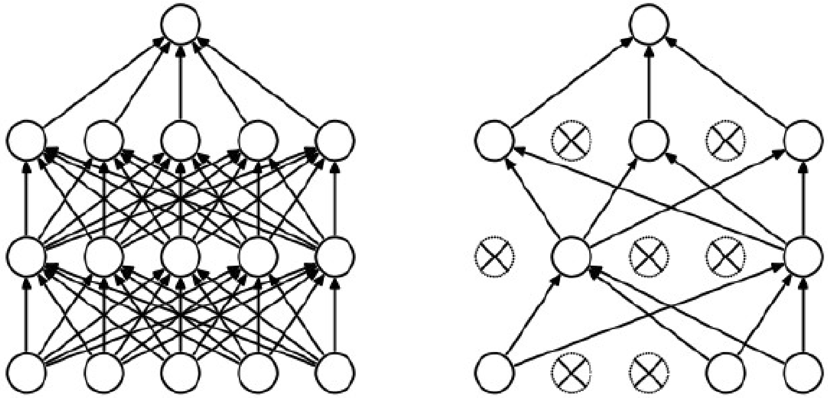


Network Layers



Kernel Renormalization:

Kernels evolve as the training progresses through renormalization. This process uses non saturating functions.



(a) Standard Neural Net

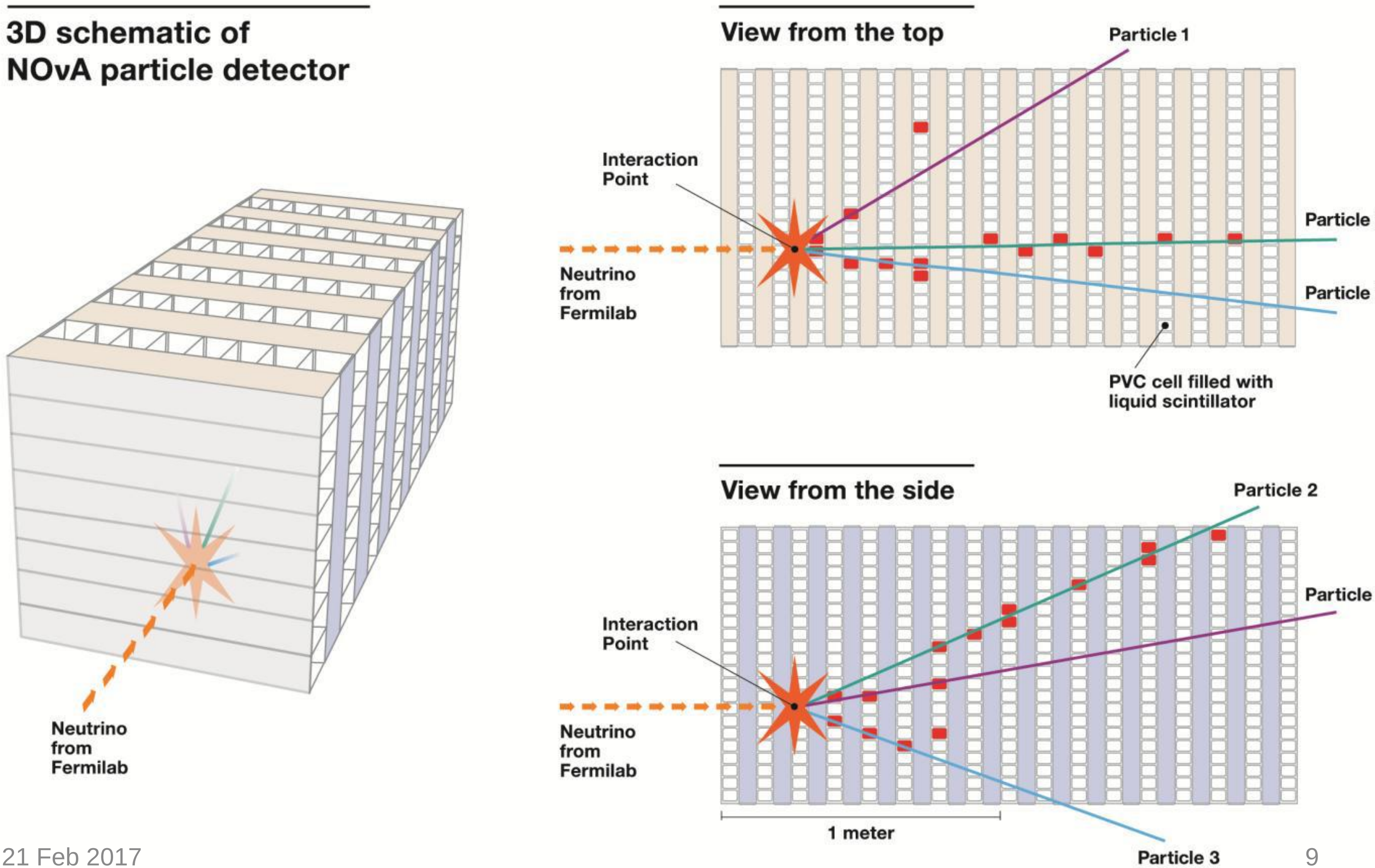
(b) After applying dropout.

Dropout:

Randomly reset weights, effectively removing whole nodes at each step.

Encourages complex dependence and discourages overtraining

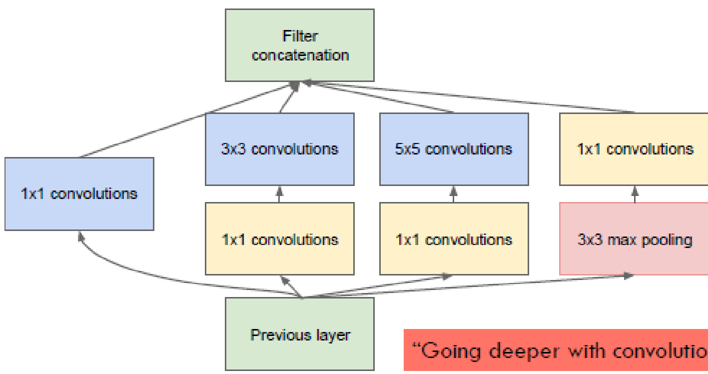
NOvA



3D schematic of
NOvA particle detector
Neutrino
from
Fermilab
View from the top
Particle 1
Interaction
Point
Particle 2
Neutrino
from
Fermilab
Particle 3
PVC cell filled with
liquid scintillator
View from the side
Particle 2
Particle 1
Interaction
Point
Neutrino
from
Fermilab
1 meter
Particle 3

CVN Convolutional Visual Network

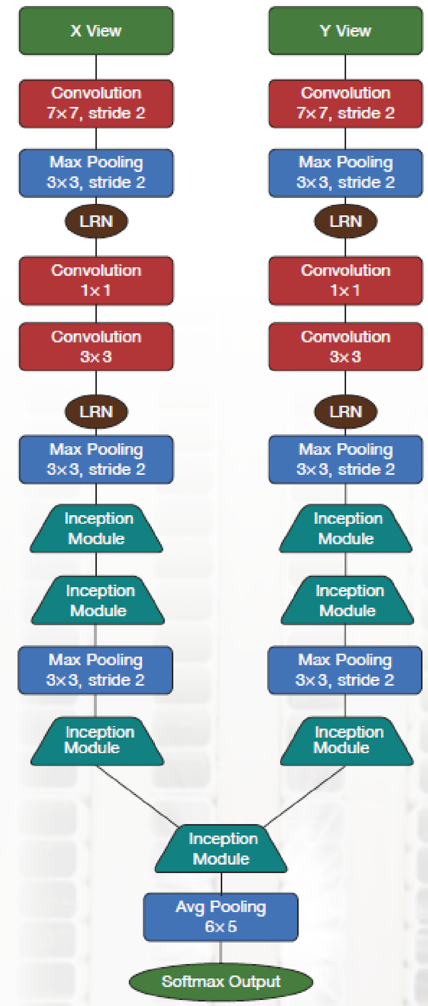
Neutrino Event CVN: Siamese network architecture based on GoogLeNet.



Inception modules:

Network in network model with kernels of multiple dimensions

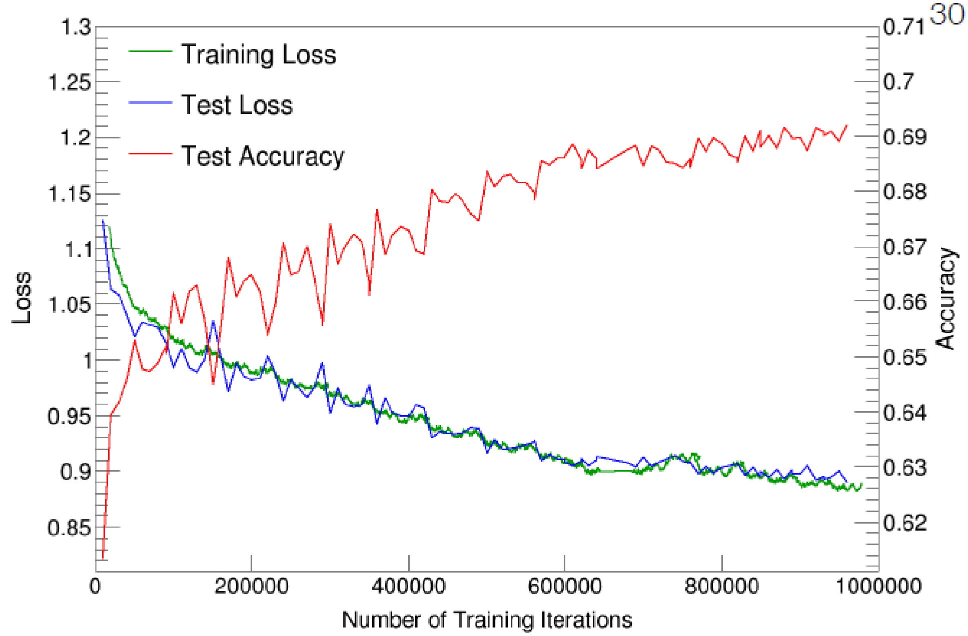
"Going deeper with convolutions" **arXiv:1409.4842**

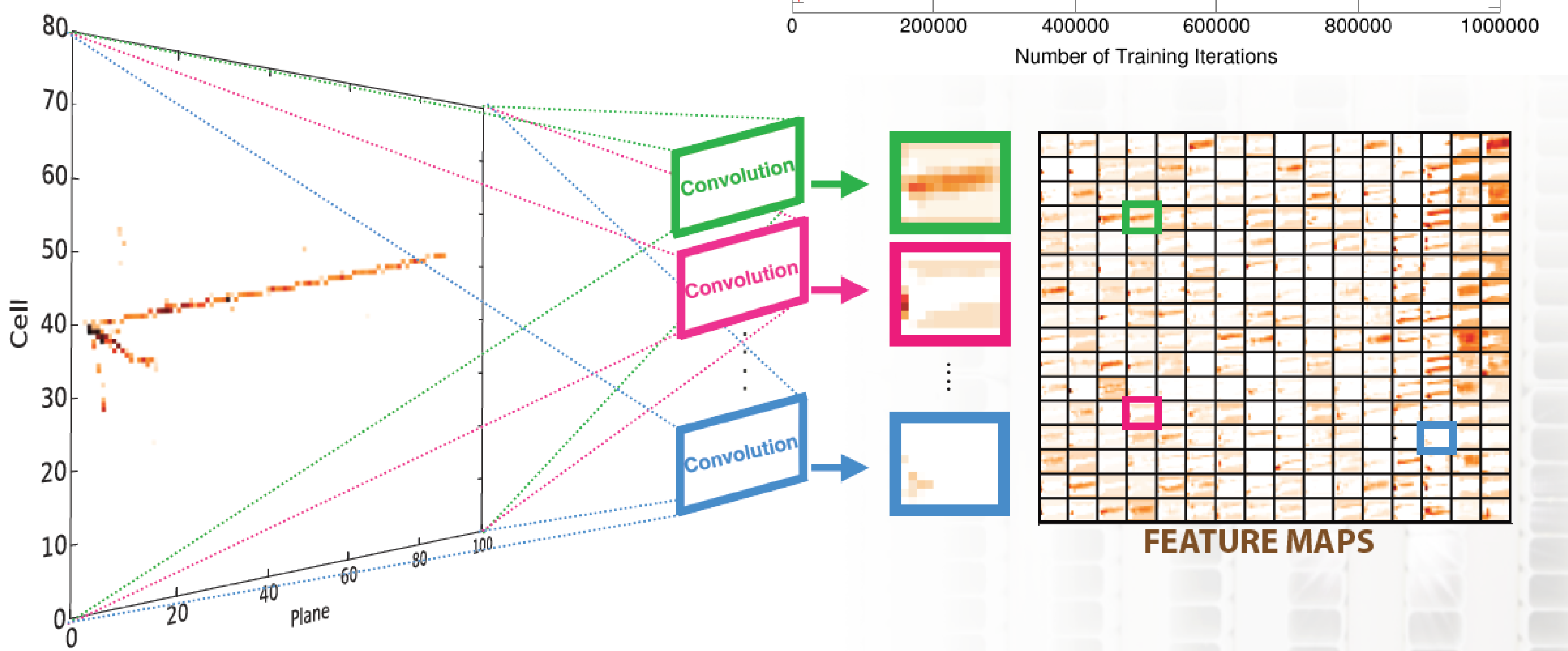


- Inspired by siamese architectures to allow the network to learn from features on each 2D view of the event.
- Using the caffe framework **http://caffe.berkeleyvision.org/**
- We train on Fermilab's Wilson cluster GPUs (**2 K40s**)
- Trained on 4.7 million simulated events of all neutrino interaction types plus cosmic rays

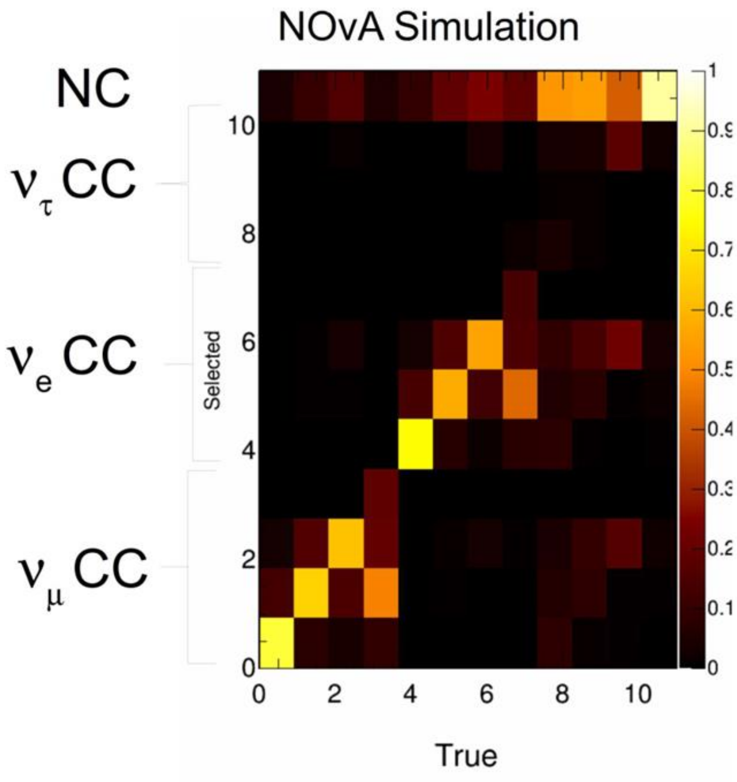
CVN Performance

- Trained on 4.7 million simulated events of all neutrino interaction types plus cosmic rays.
- training sample has minimal preselection



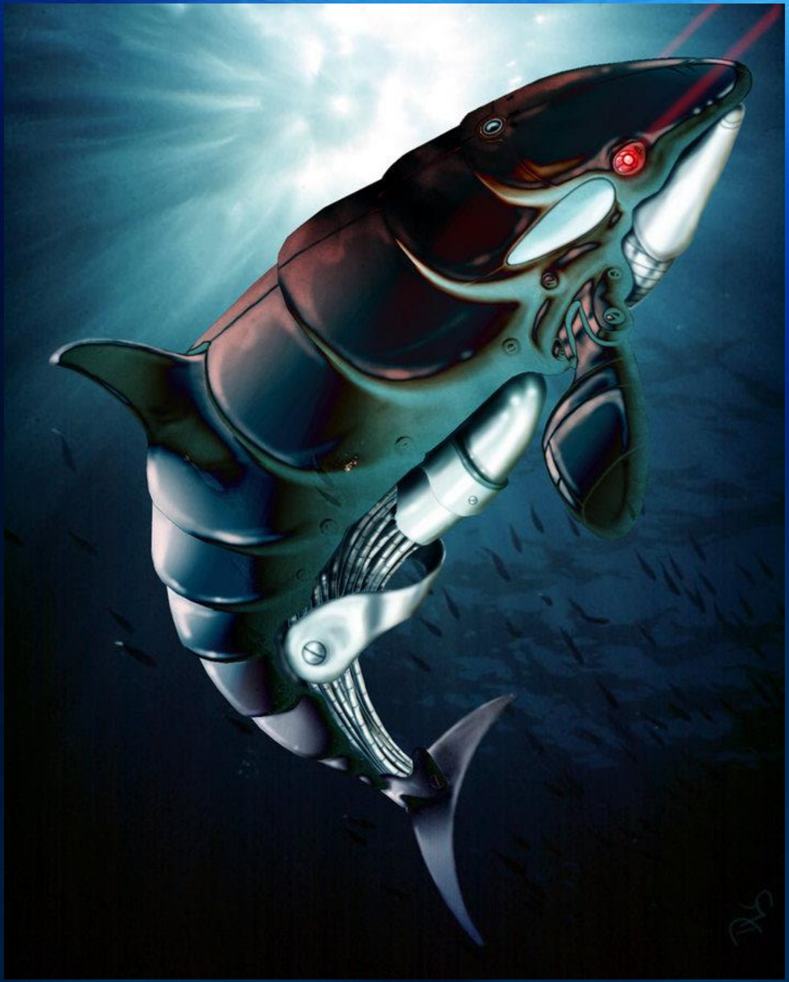


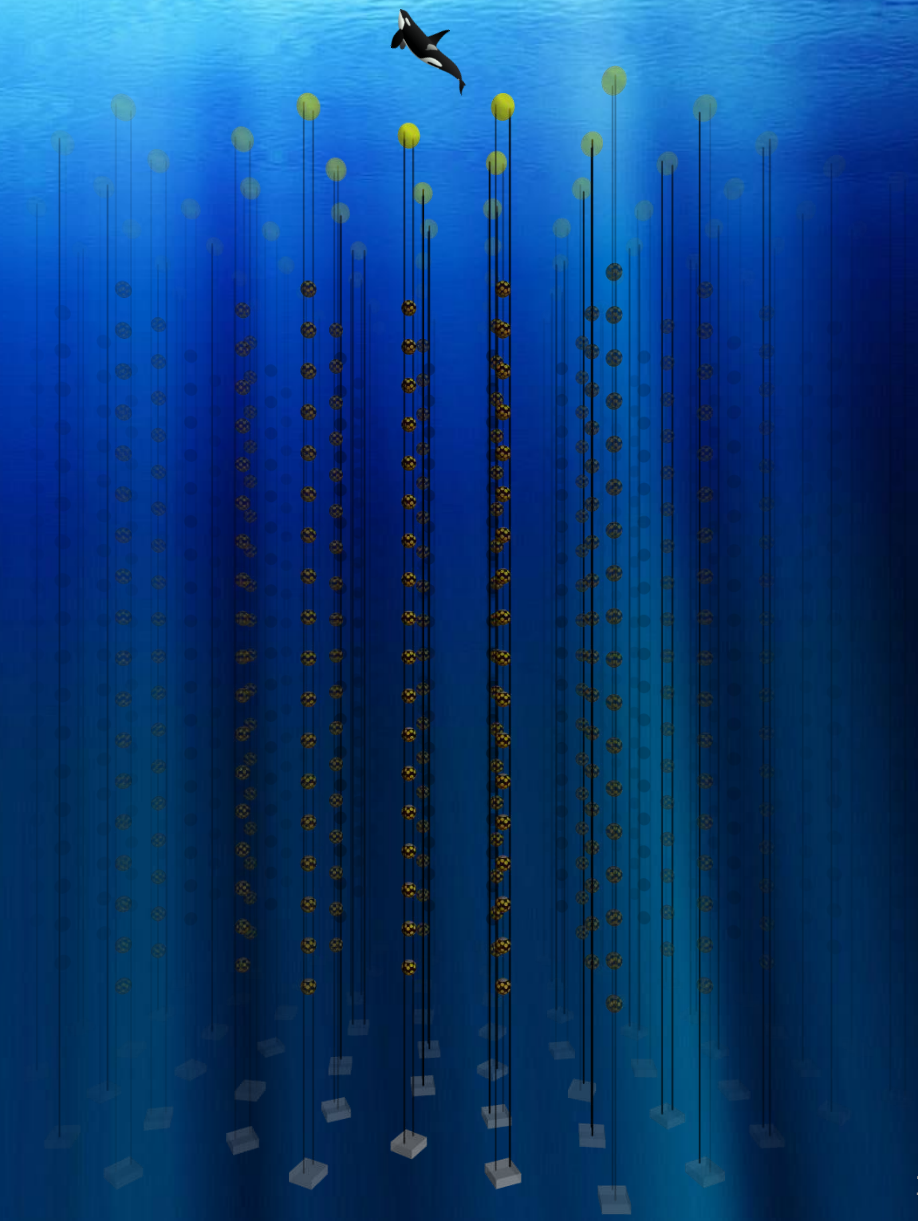
Initial Training



- The initial training consisted of ~10 passes over 3.4 million training images.
  - With two k40 GPUs, this took ~1 week.
- We can judge how well our model works by producing a confusion matrix.
  - This shows the relationship between the true event category and what the PID thought was the most likely event category.
- The matrix is mostly diagonal – events are mostly correctly identified.
- Mis-identified events mostly fall within blocks – while the interaction type is wrong, the selected neutrino flavor is still correct.

What about ORCA?

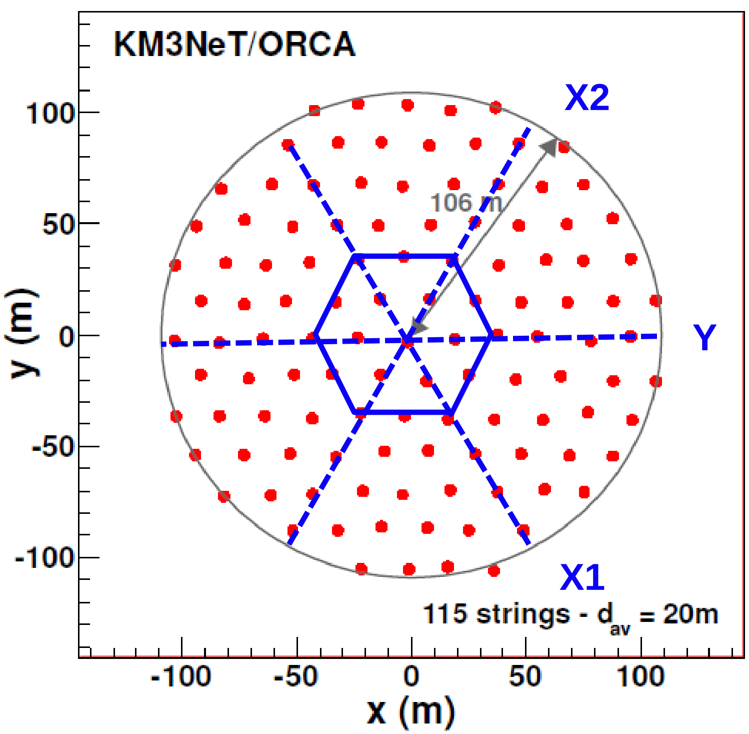


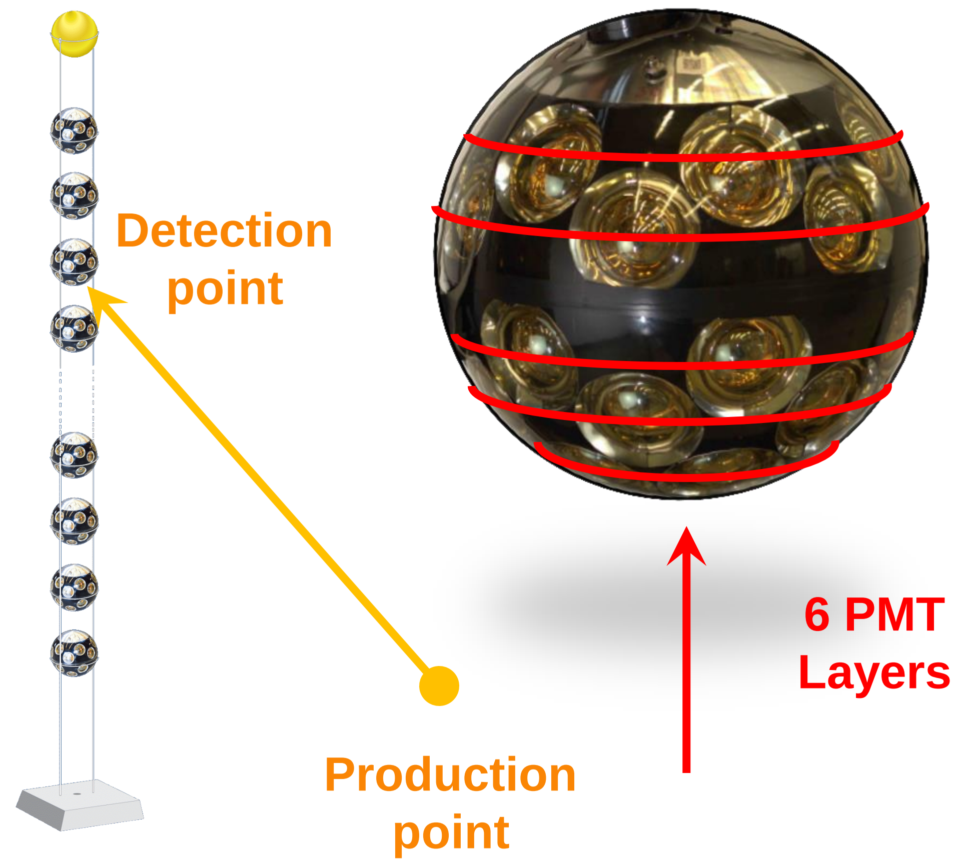


Our Detector

- Cylindrical volume with hexagonal string symmetry
- Extra dimensions from PMT orientation and hit time
- A 2-point model: Light production and detection points

3 well-defined axes





Detection Point

- No need for true detection, just encode time and direction information
- Extrapolate light back in the direction of the PMT for a time t_{hit} - t_0 (event start time)
- Ignores particle that produced the light

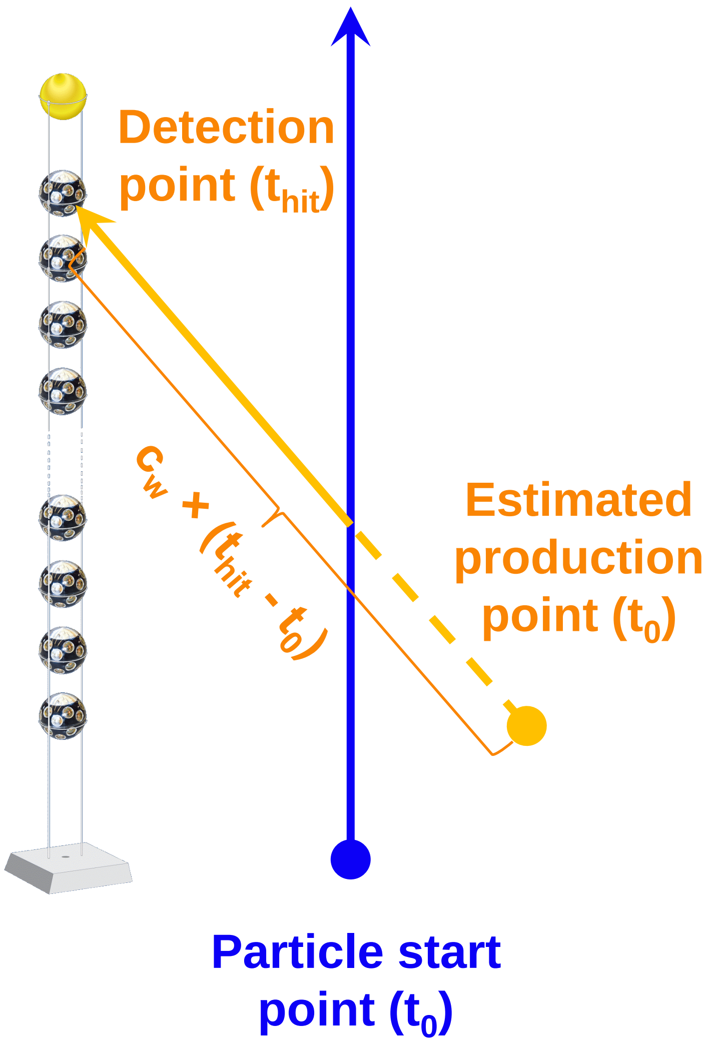
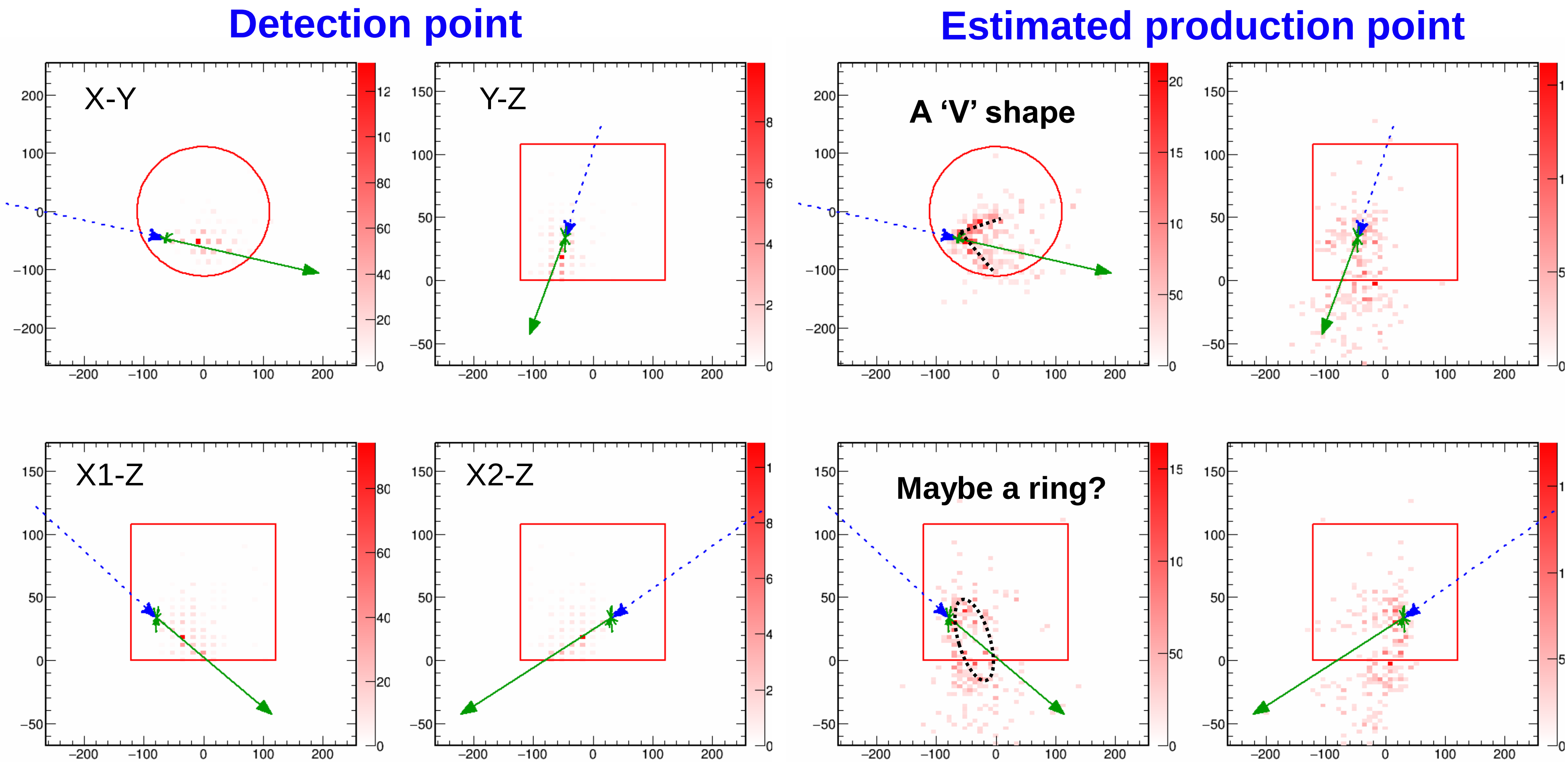


Image Inputs

- Define four (x2) views (2D projections of the detector)
- Plot charge in detection and “production” positions



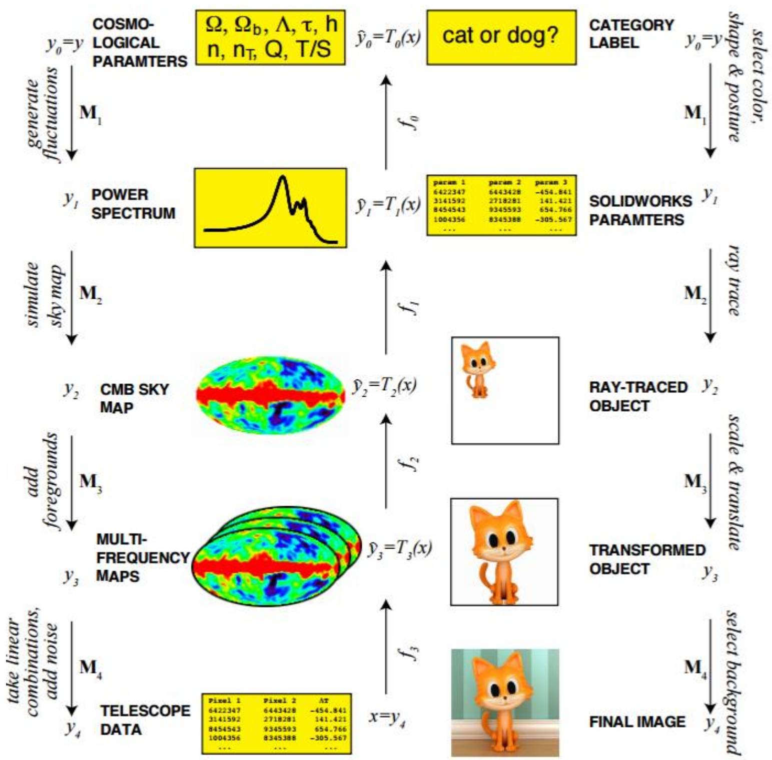
Why Projections?

Why does deep and cheap learning work so well?

Henry W. Lin (Harvard), Max Tegmark (MIT)

(Submitted on 29 Aug 2016 (v1), last revised 28 Sep 2016 (this version, v2))

- **Symmetry, locality**, ... -> neural nets are good with **images and drawings**
- **Hierarchical data** -> deep neural network more efficient



Multi-view Convolutional Neural Networks for 3D Shape Recognition

Hang Su, Subhransu Maji, Evangelos Kalogerakis, Erik Learned-Miller

(Submitted on 5 May 2015 (v1), last revised 27 Sep 2015 (this version, v3))

- **Multi-view 2D** images can **outperform** networks with **truly 3D** shape inputs

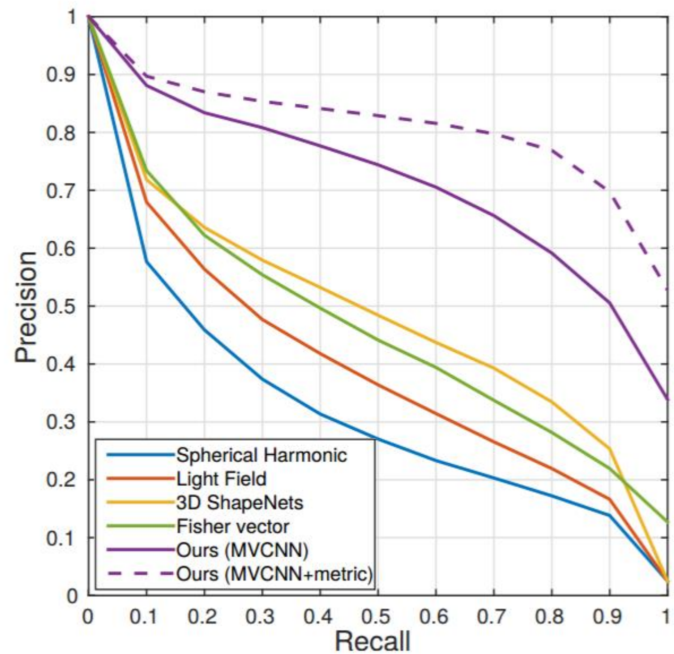


Figure 2. Precision-recall curves for various methods for 3D shape retrieval on the ModelNet40 dataset. Our method significantly outperforms the state-of-the-art on this task achieving 80.2% mAP.

What’s next?

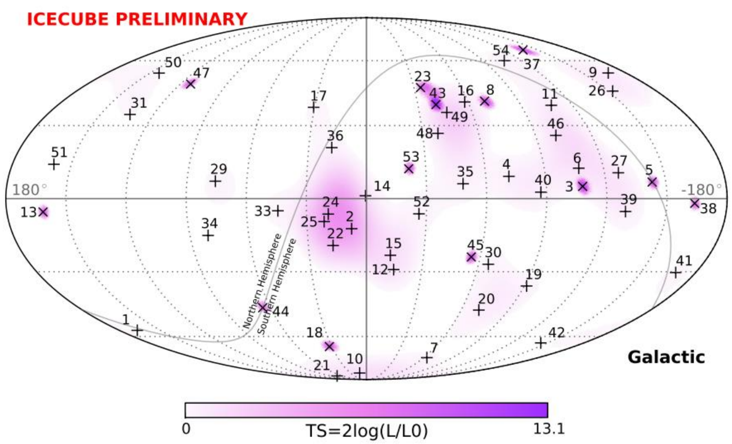
- **Deep learning** is being successfully implemented in HEP
- Can we use it to **reconstruct and/or classify** KM3NeT events?
- We have an initial concept, based on the NOvA experience
- NOvA used the **Caffe framework**, but should we?
- Challenging 6D input space. Is the best solution **projections?**
- Are **Convolutional NNs** a good approach?
- Other sugestions?

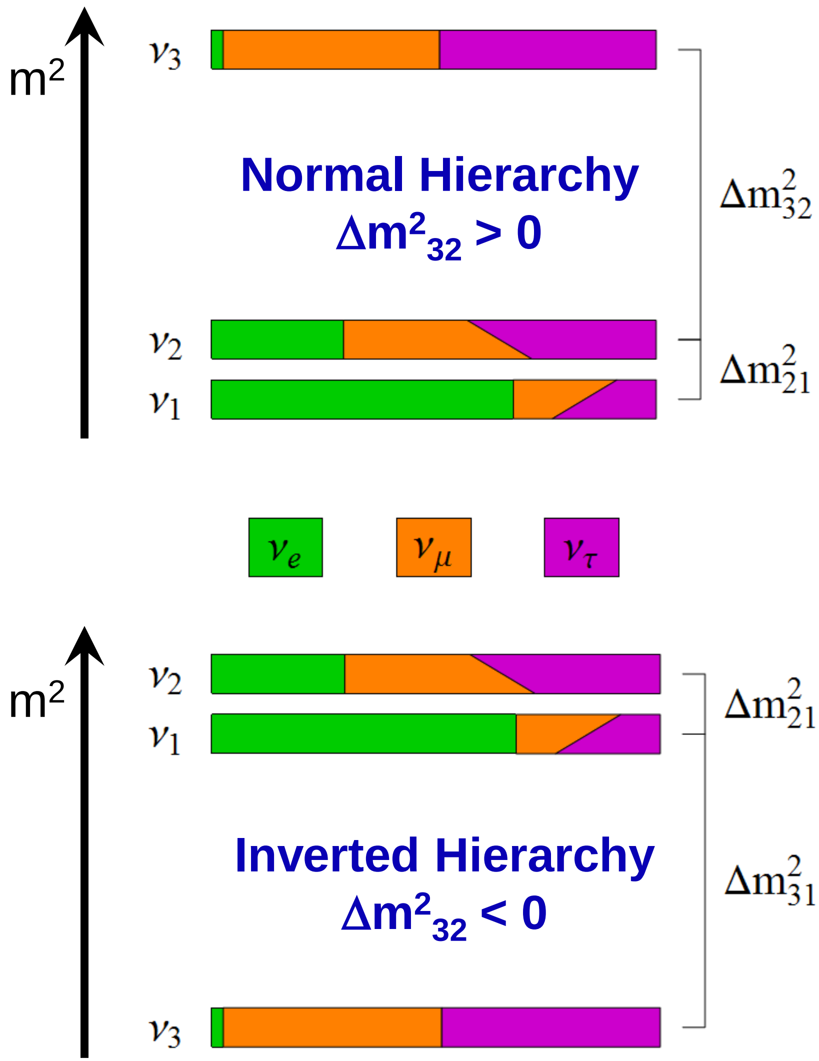
Thank you!

Backup Slides

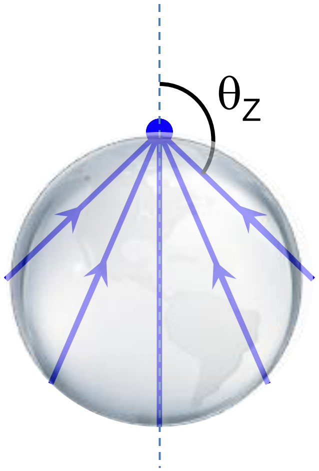
Objectives

- **ORCA**: Determine the **Neutrino Mass Hierarchy (NMH)**
- **ARCA**: Discover/Observe high-energy neutrino sources in the universe

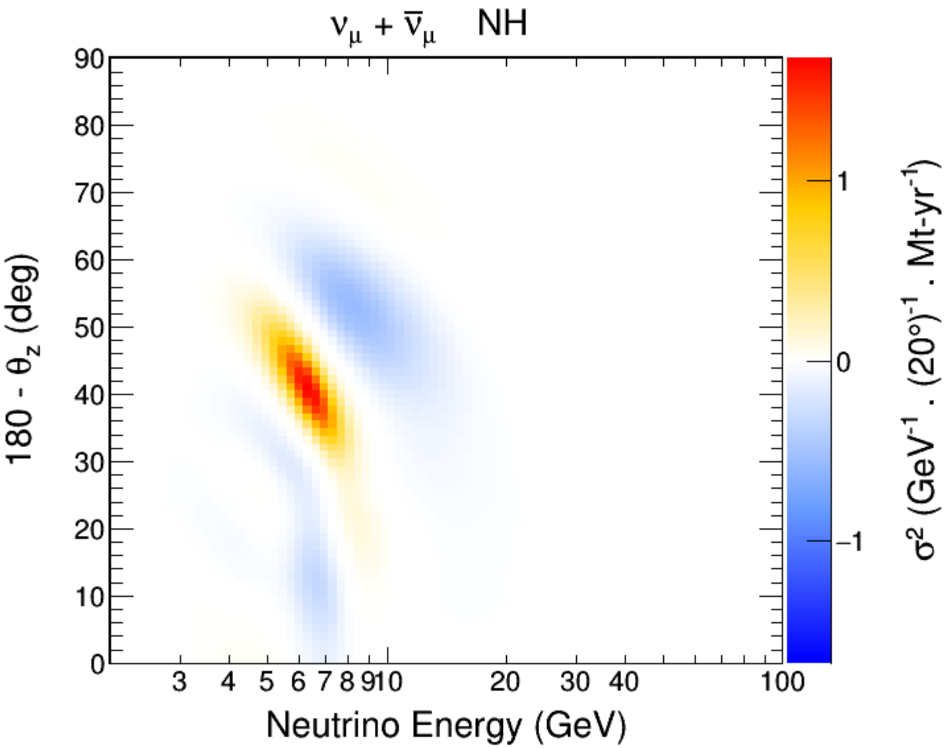


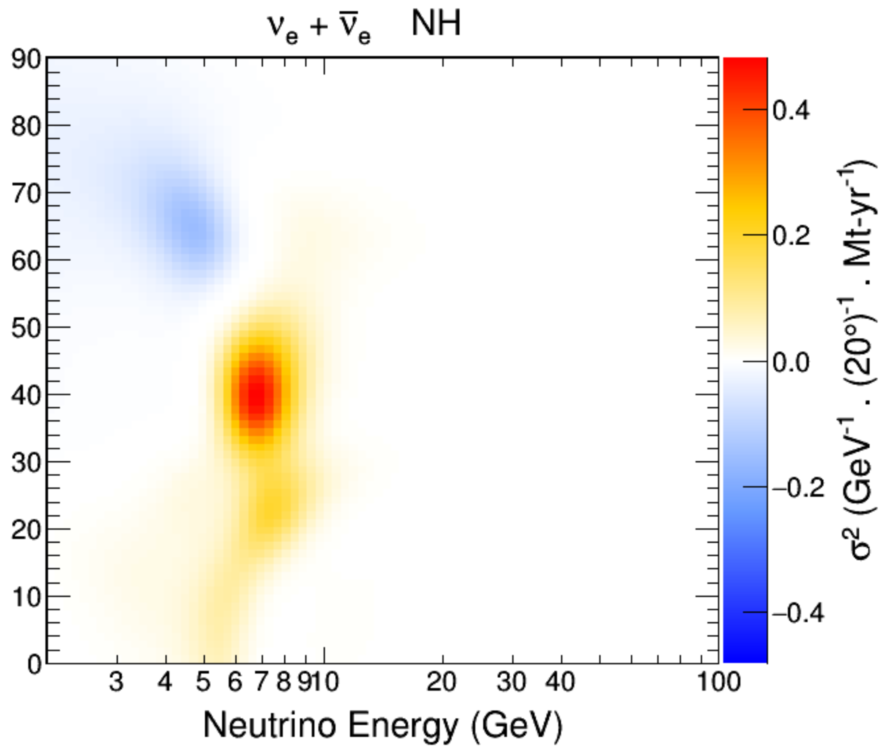


The Challenge



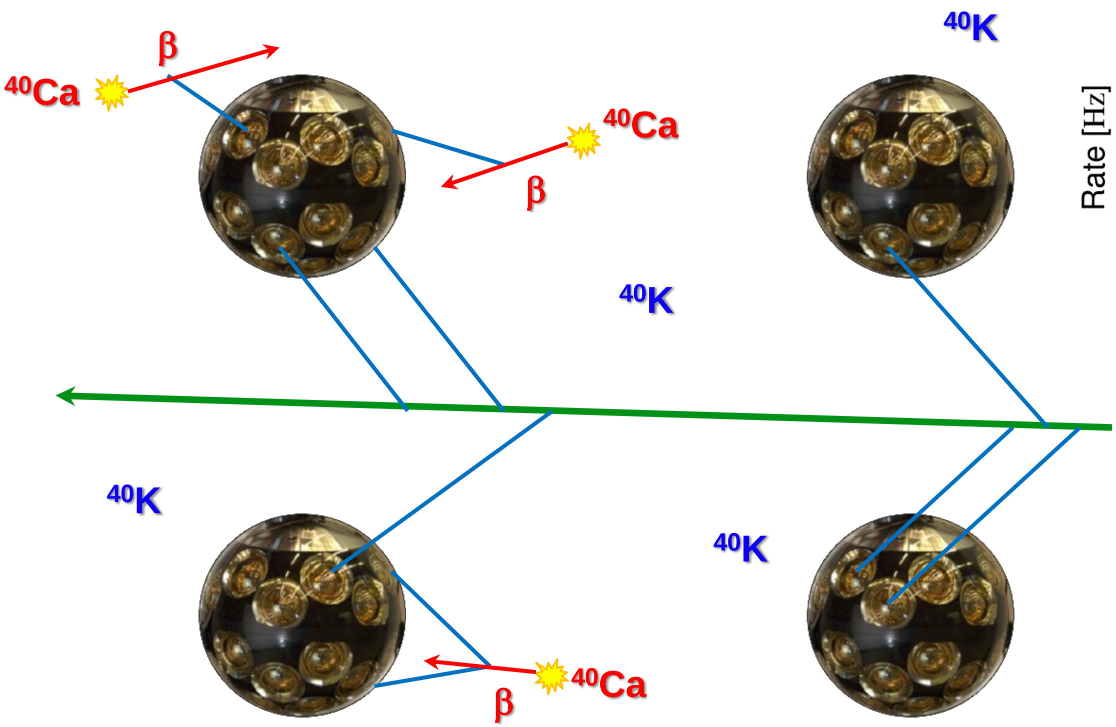
- Measure neutrino direction and energy
- Search for **oscillation patterns** from **matter effects**
- Sensitive to difference in patterns between NH and IH
- Requires **large statistics** and good **energy and direction res.**

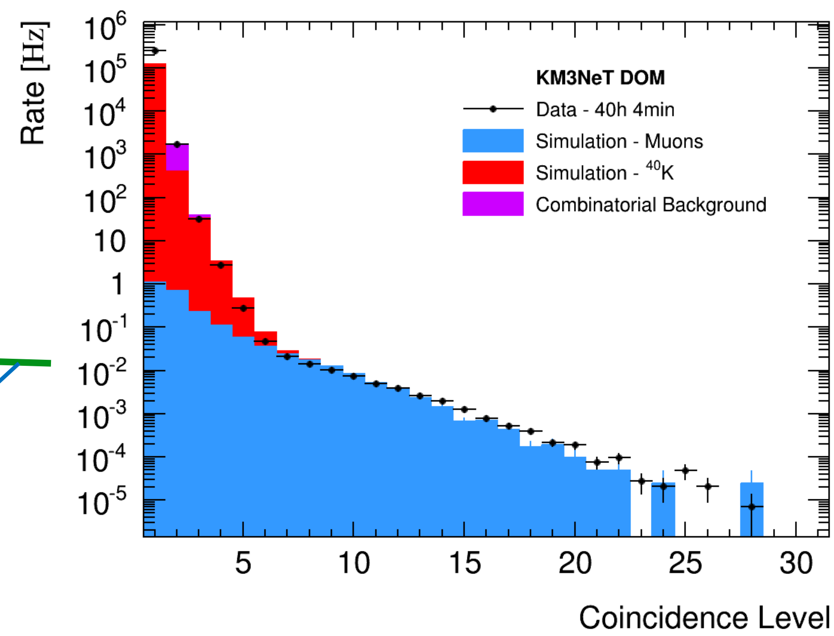




Trigger

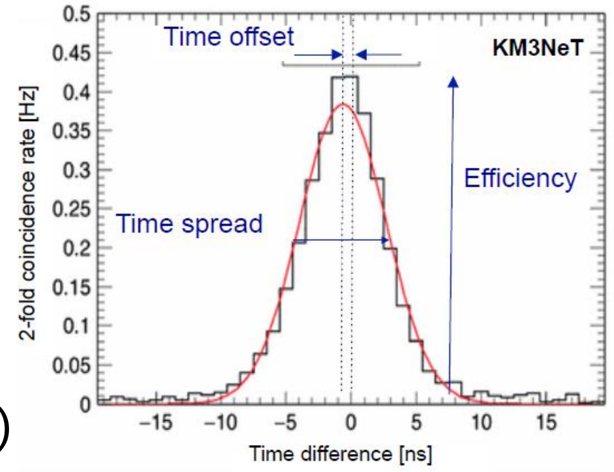
- Optical background mostly from ^{40}K decays in the water
- Measured: **8 kHz** uncorr., **340 Hz** level-two coinc. / PMT **[Eur. Phys. J. C 74, 3056 (2014)]**
- Look for coincidences in time and PMT direction to reduce trigger rate.
- Causality further restricts space and time correlations for extra power.
- Final trigger rate **~59 Hz**, with **70%** of events containing a cosmic ray **muon**.

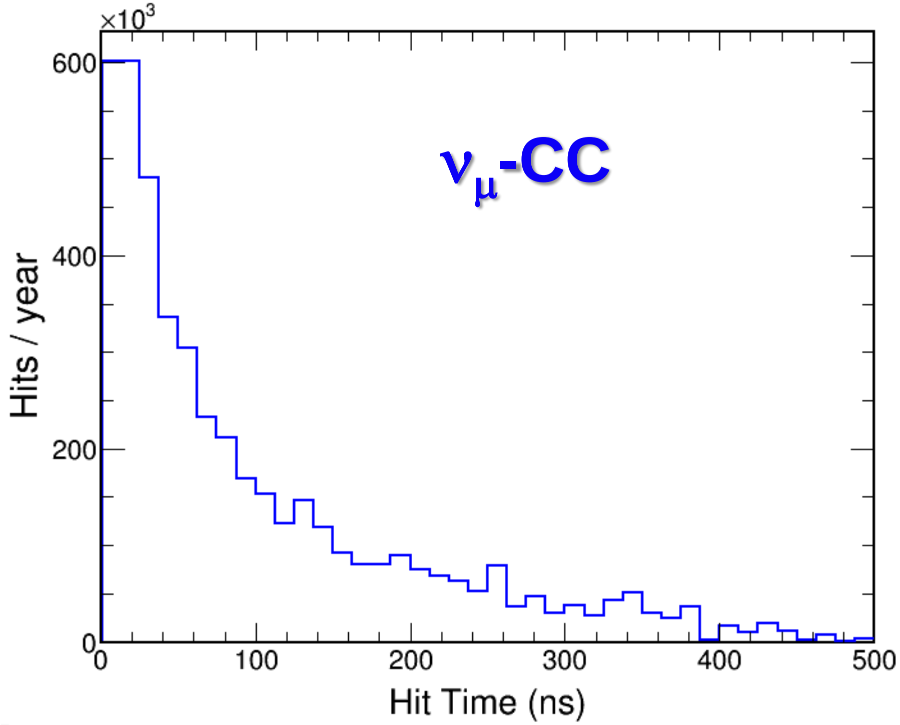


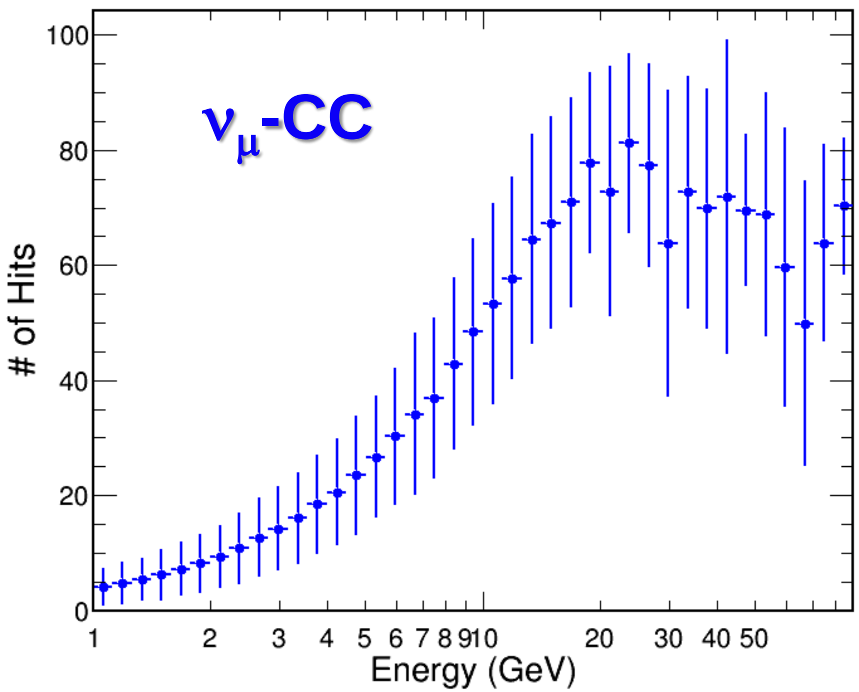


Optical Noise

- Optical background in full detector **~500 MHz**.
- Neutrino events ~40 hits in a **~500 ns** window
- Expect **250 noise hits** (~14% purity)
- Trigger approach **~5 ns time residuals**
- Calibrated using 2-fold coincidences
- Can achieve **~3 noise hits** per trigger (>90% purity)

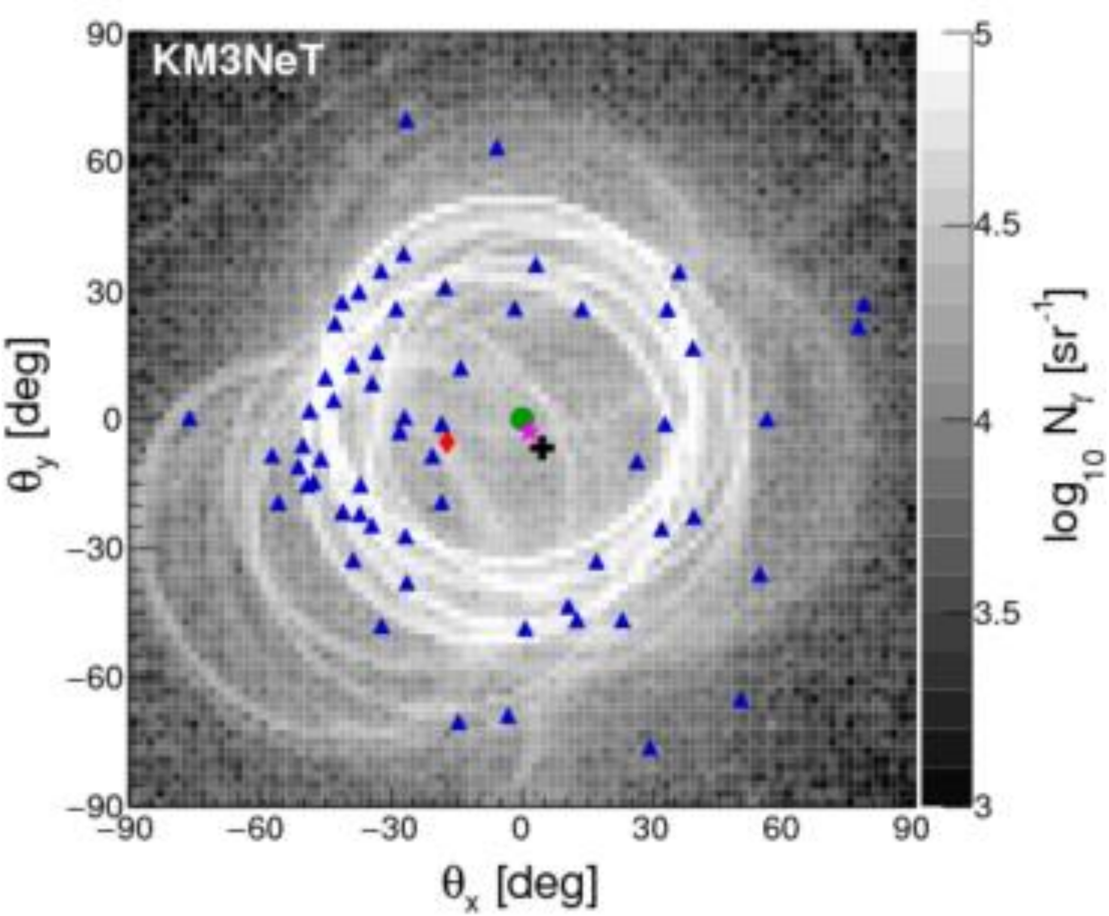






Limitations

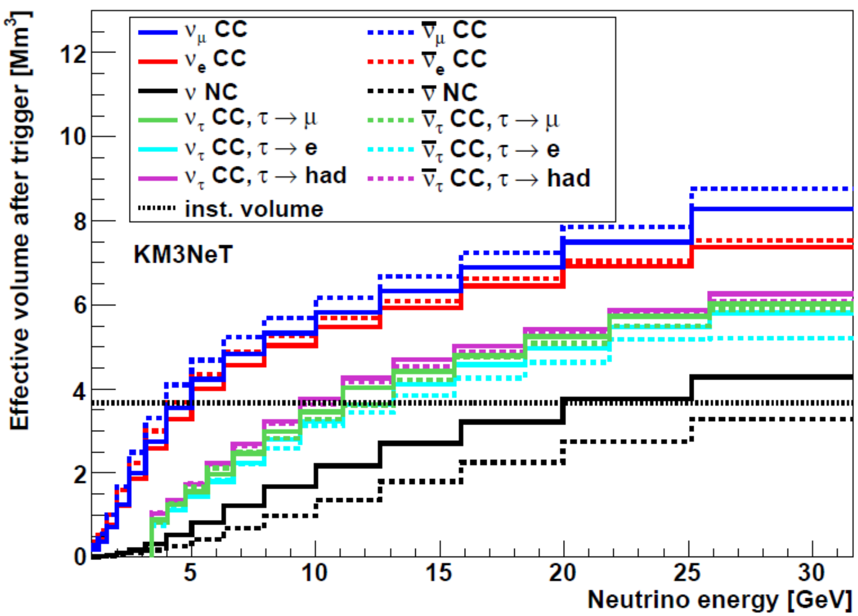
- Hadronic showers can generate multiple cherenkov rings, but reconstructing them is limited by photon statistics
- Without hadronic info, reconstruction is limited by intrinsic kinematics
- Extracting some information on particle multiplicity would be a significant achievement for a deep learning technique

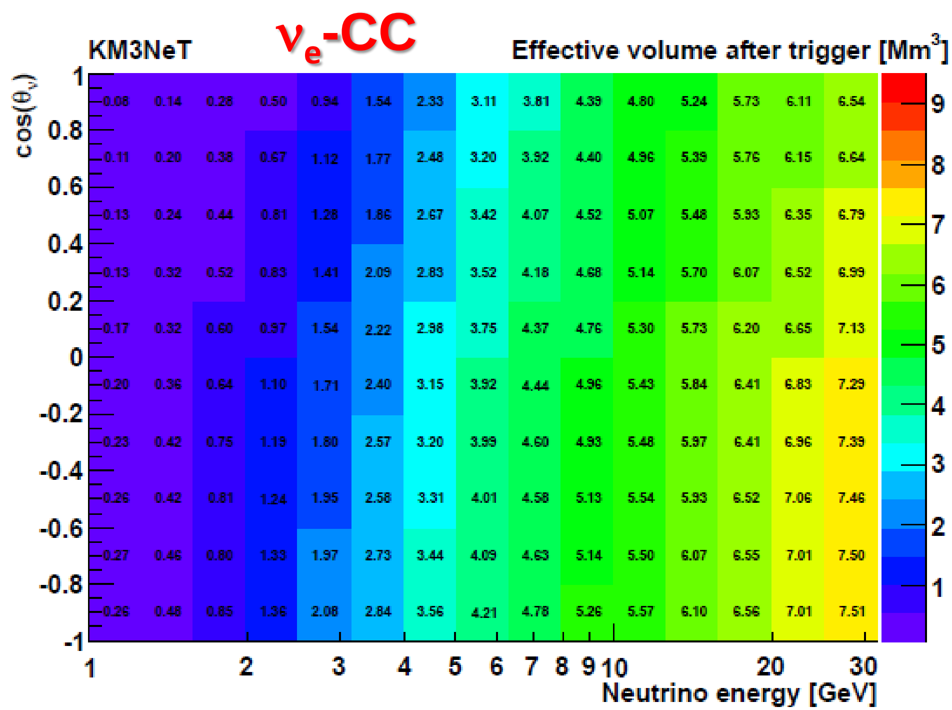


Trigger Performance

- Input a **conservative** noise rate of **10 kHz** uncorr. (**500Hz** level-two coinc.)
- Achieve a total triggered rate of **59 Hz**
- About **70%** of events contain a **muon** (41 Hz)
- High efficiency for ν_μ and ν_e above 4 GeV
- Slightly more efficient for up-going neutrinos (Larger PMT coverage)

Neutrino Rate:
~ 1 ν / 10 min

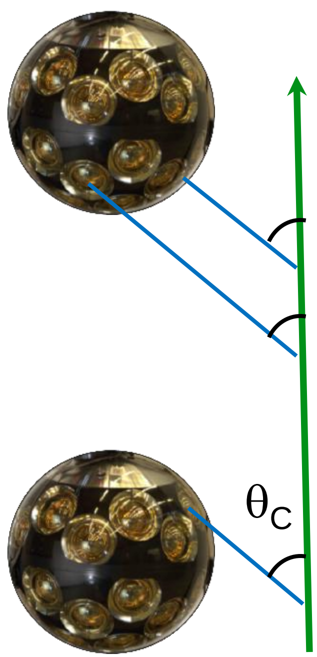




Reconstruction

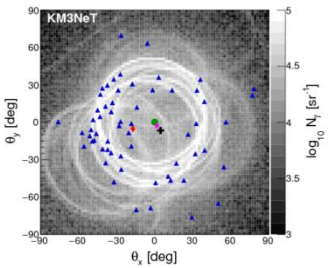
1) Start with a track or shower hypothesis
2) Use **causality** to perform a robust **hit selection**
3) Find **vertex** and **direction** that best match hit pattern
4) Estimate track range for computing **track energy** (0.24 GeV / m)
5) Estimate **Shower energy** and direction from hit distribution after initial fit to the vertex position and time

Track Hypothesis



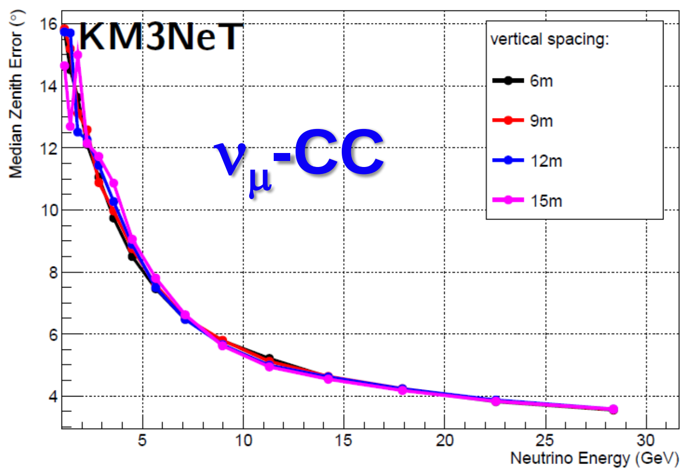
Shower Hypothesis

Reco Performance

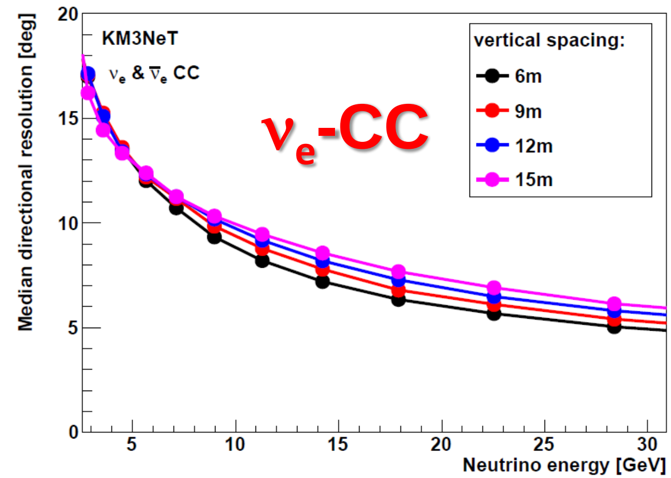


- Energy resolution: ~25% (Close to limit arXiv:1612.05621)
- Angular resolution: Better than 10 degrees at relevant energies

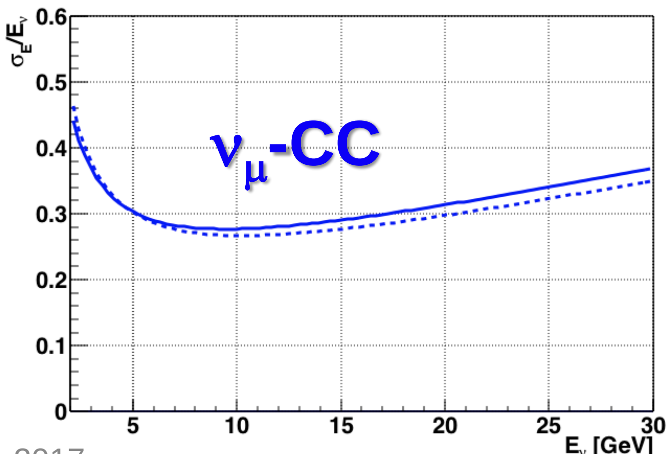
θ Res.



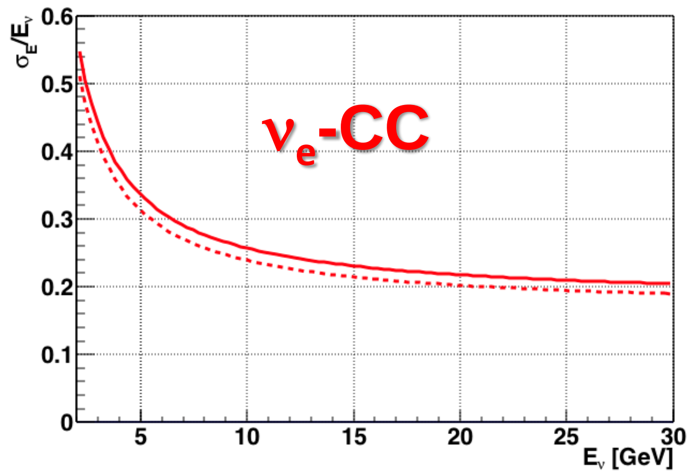
θ Res.



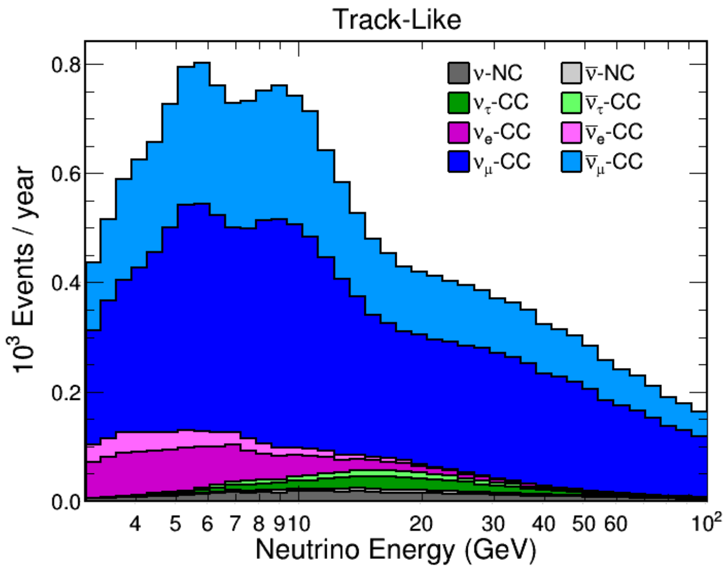
E Res.



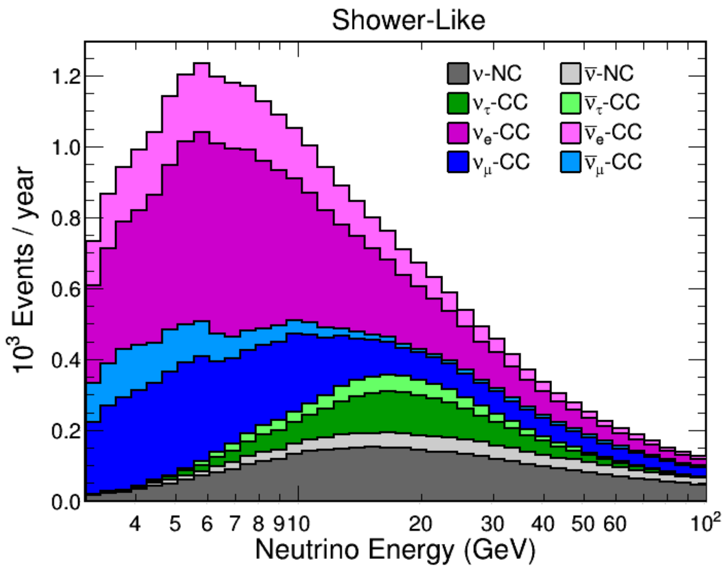
E Res.

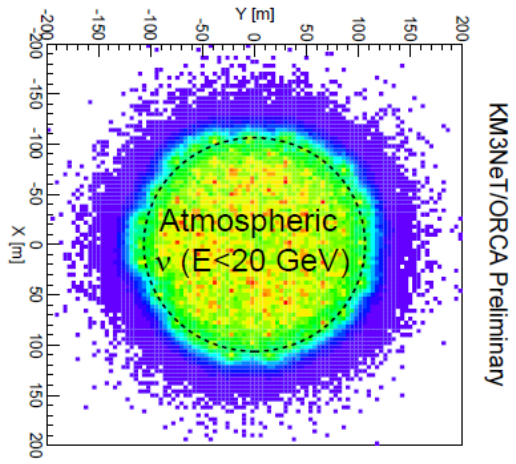


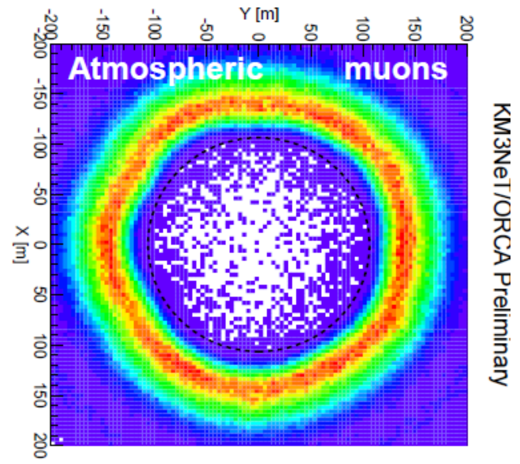
Event Selection



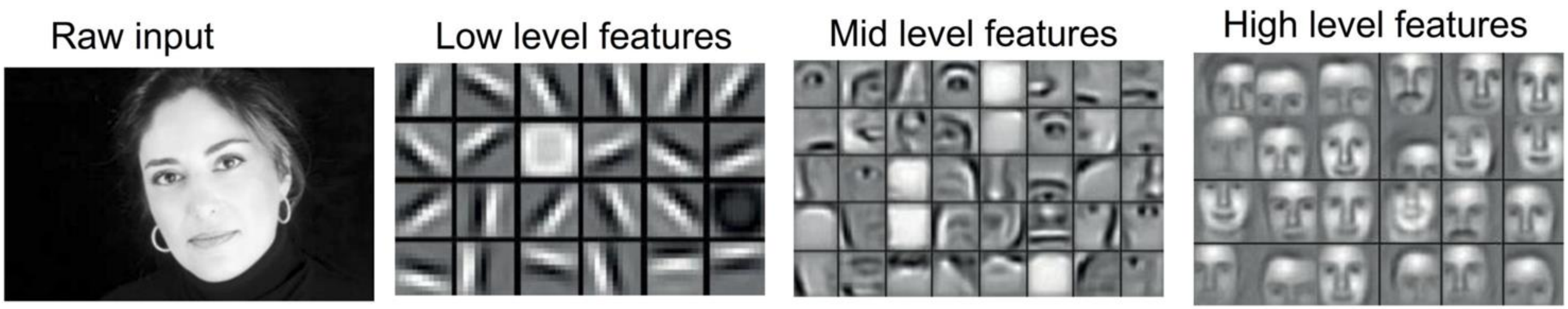
- Events are classified through a Random Decision Forest (RDF)
- At 10 GeV:
  - 90% of ν_e-CC are shower-like
  - 70% of ν_μ-CC are track-like
- Most atmospheric muons are removed by containment cuts







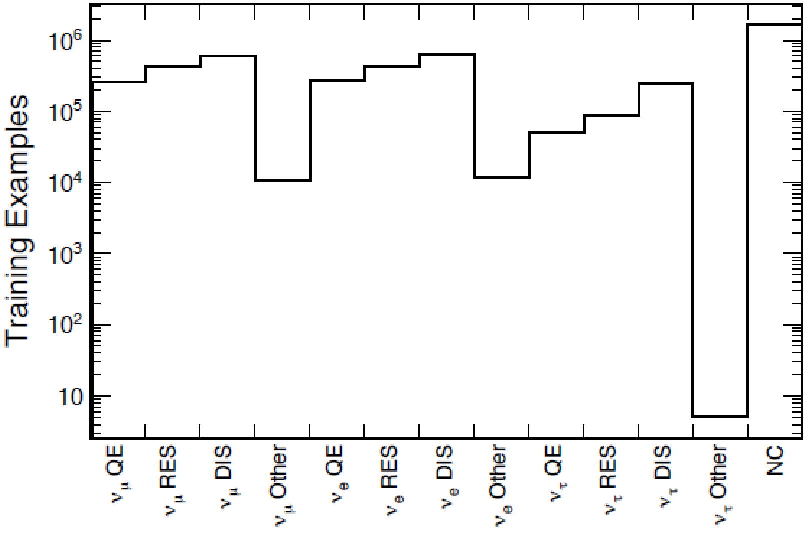
Deep Learning



developer.nvidia.com/deep-learning-courses

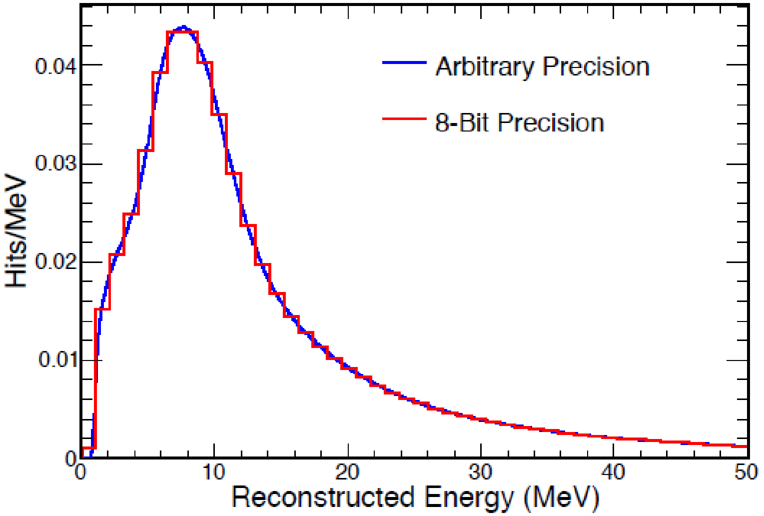
- Deep learning is a new paradigm that has caused a renaissance in the machine learning community.
- Use sparsely connected neurons to allow for many hidden layers.
- Deep structure extracts increasingly complex features from the input data instead of needing engineered features.

CVN Classifier



4.7 million, minimally preselected simulated events, pushed into LevelDB databases: 80% for training and 20% for testing.

Rescale calibrated energy depositions to go from 0 to 255 and truncate to chars for dramatically reduced file size at no loss of information



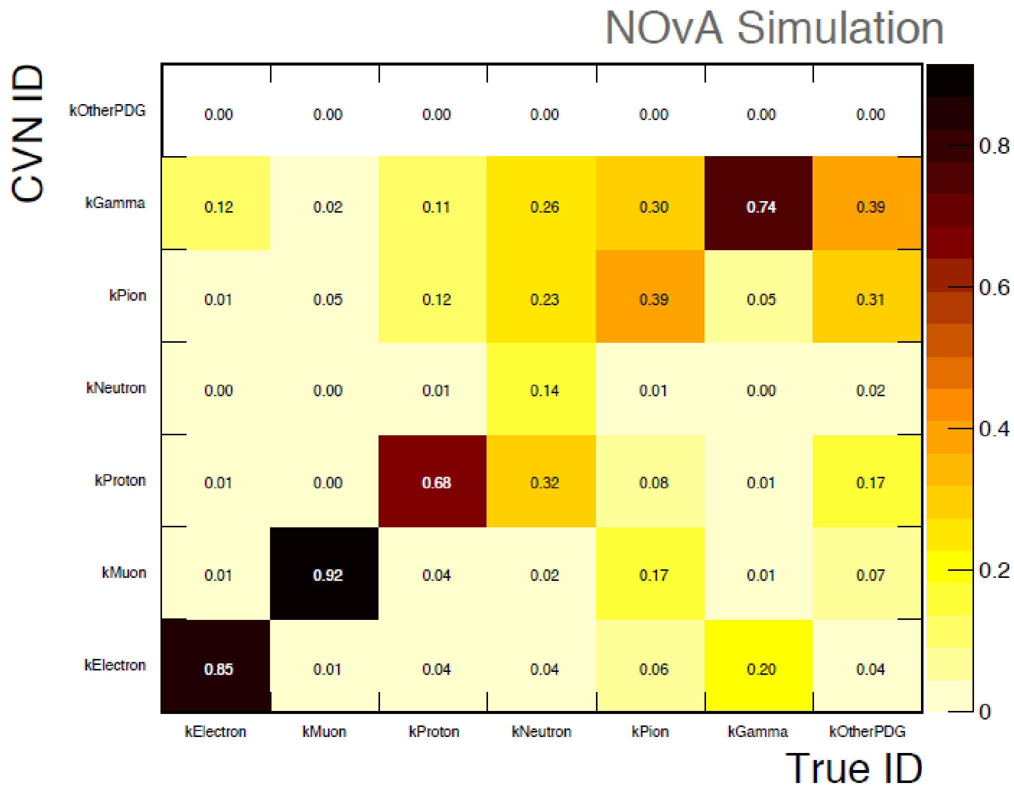
Fine tuned with 5 million cosmic data events taken from an out of beam time minimal bias trigger.

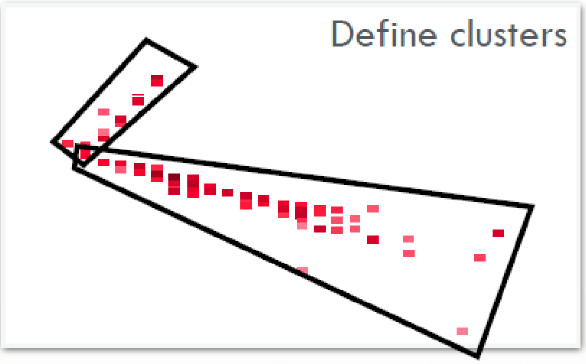
The architecture attempts to categorize events as {ν_μ, ν_e, ν_τ } × {QE,RES,DIS}, NC, or Cosmogenic.

Ongoing Work CVN and Reconstruction

Using the existing reconstruction.

Classify clusters by particle ID





Original CVN network modified to take 4 views (event + prong)

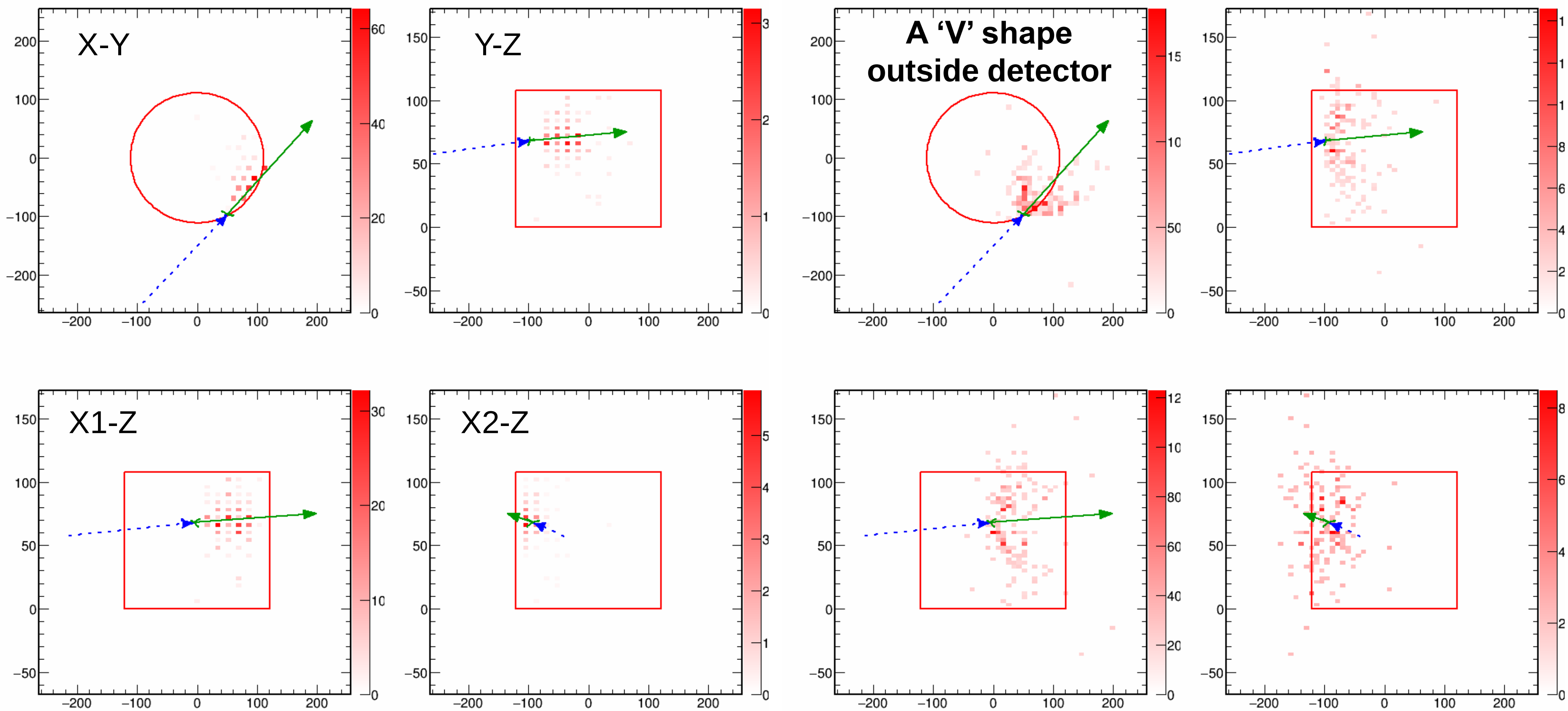
Trained on 50% purity prongs from all events no preselection

Room for improvement in classification and network optimization

OrcaNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and “production” positions

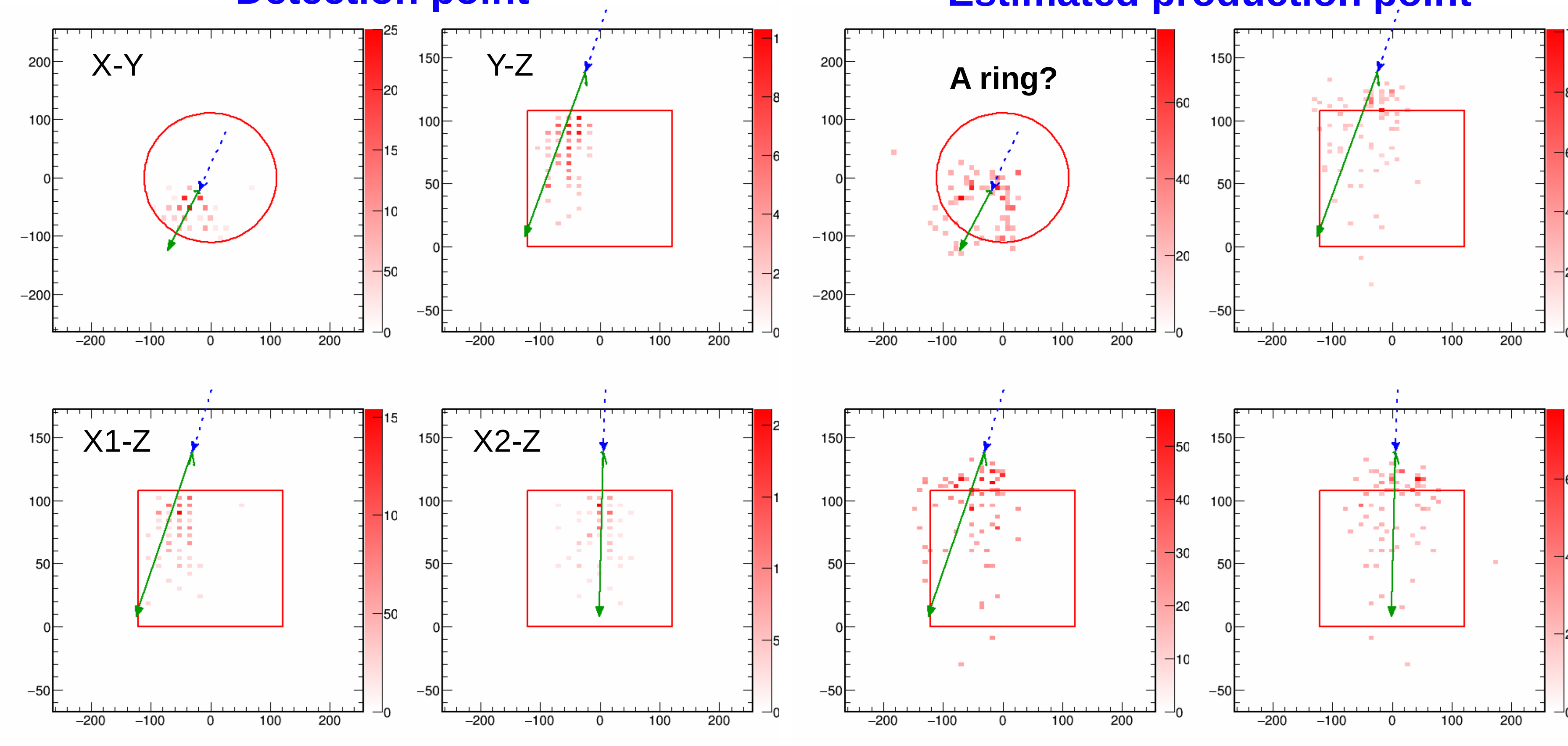
Detection point **Estimated production point**



OrcaNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and "production" positions

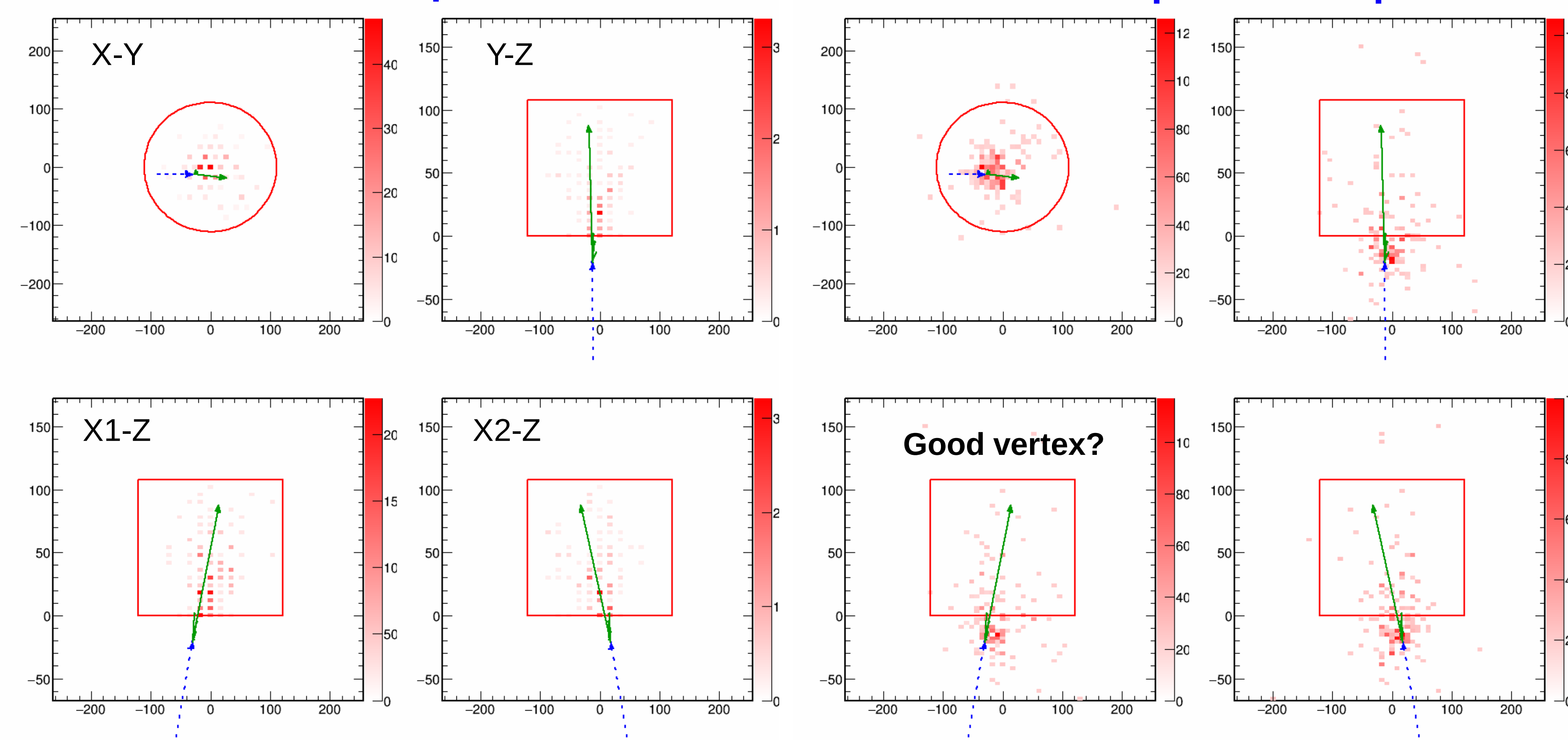
Detection point | **Estimated production point**



OrcaNet Inputs

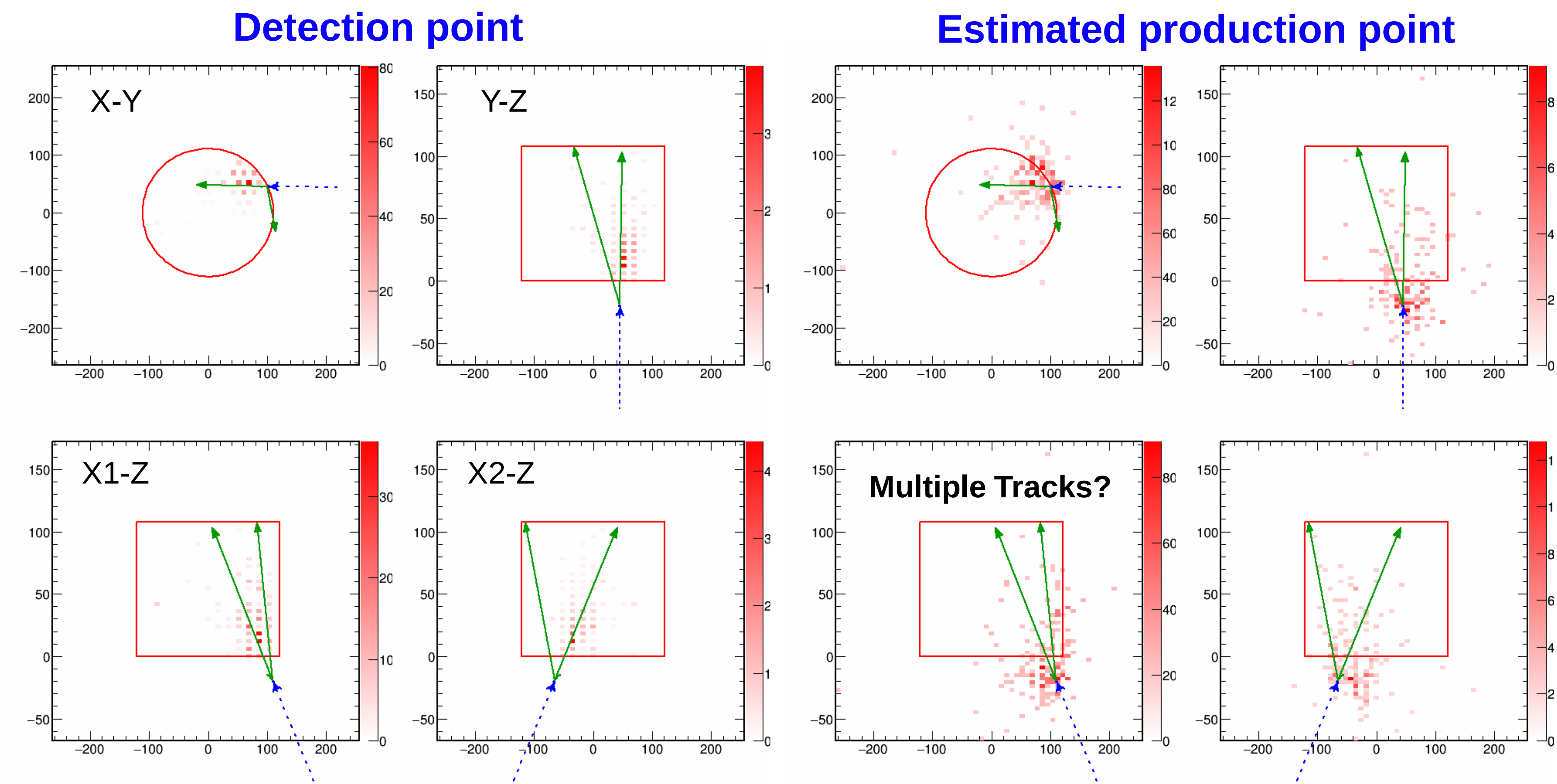
- Define four views (2D projections of the detector)
- Plot charge in detection and “production” positions

Detection point **Estimated production point**



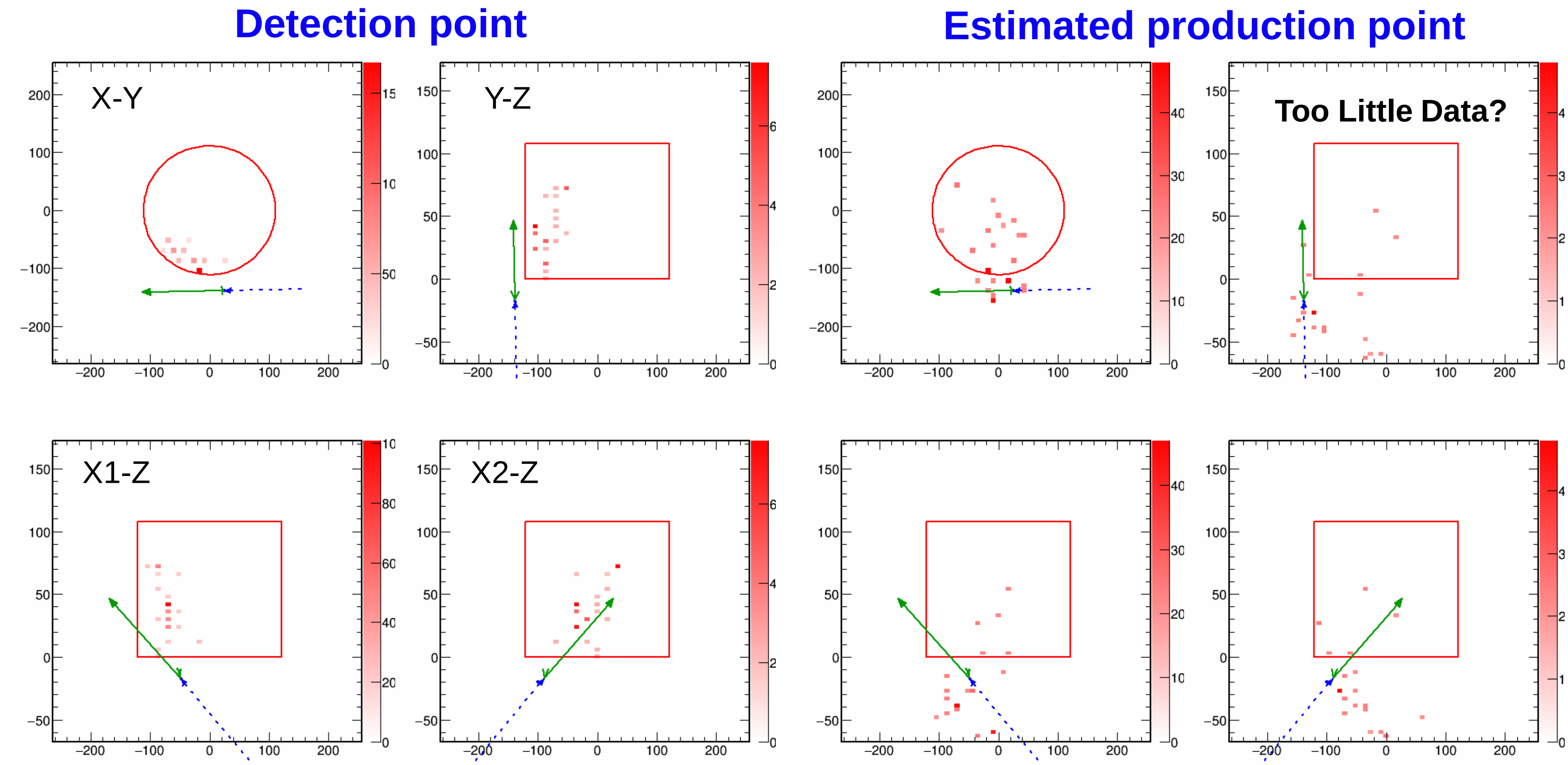
OrcaNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and “production” positions



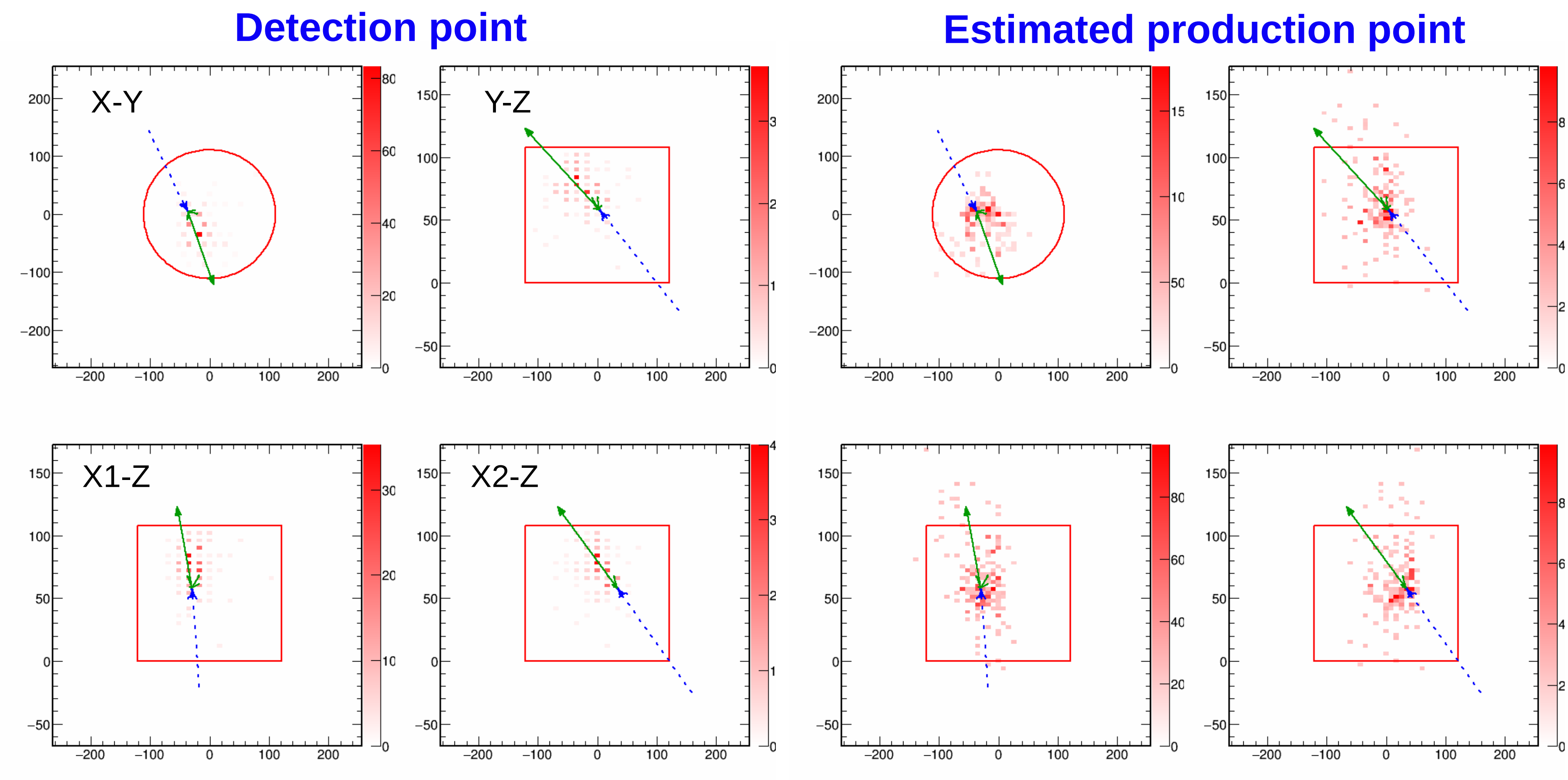
OcraNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and “production” positions



OrcaNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and "production" positions



OrcaNet Inputs

- Define four views (2D projections of the detector)
- Plot charge in detection and "production" positions