CNN-based neutrino event classification in DUNE

Caveat: a number of plots will be updated to use the latest results from MCC11

Leigh Whitehead

(for the DUNE Collaboration)

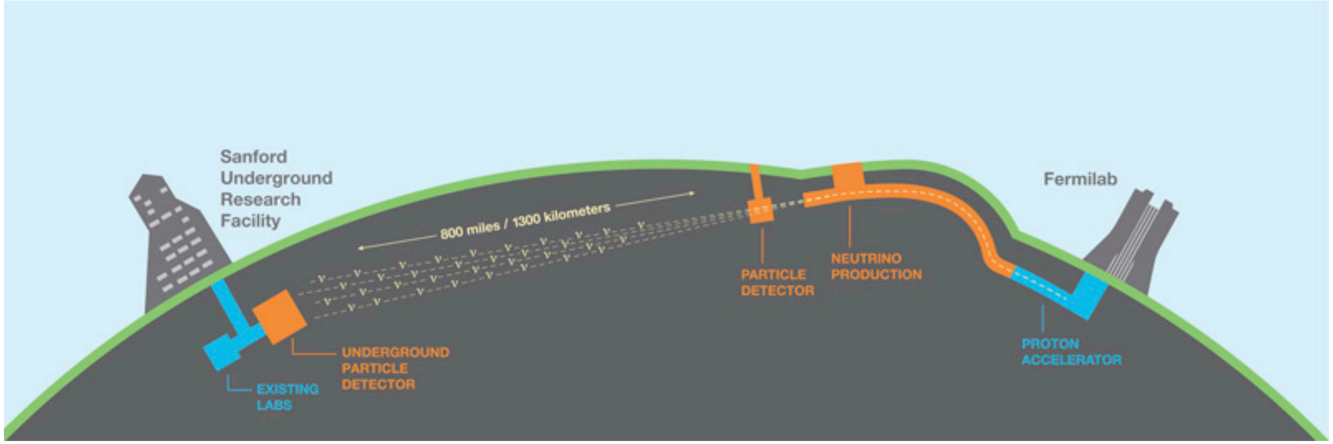
LArTPC Calibration and Reconstruction Workshop

12/12/18

CERN

The DUNE Experiment

- DUNE is a next-generation neutrino oscillation experiment



- Far Detectors (FD) are 800 miles from the neutrino beam source
  - Four modules, each with 10,000 ton of liquid argon
- High power muon neutrino beam produced at Fermilab
  - Can switch polarity to produce a muon antineutrino beam
- Look for the appearance of electron (anti)neutrinos at the FD
  - Measure *CP*-violation

NB: I will only write neutrino from now on, but the same is applicable for antineutrinos

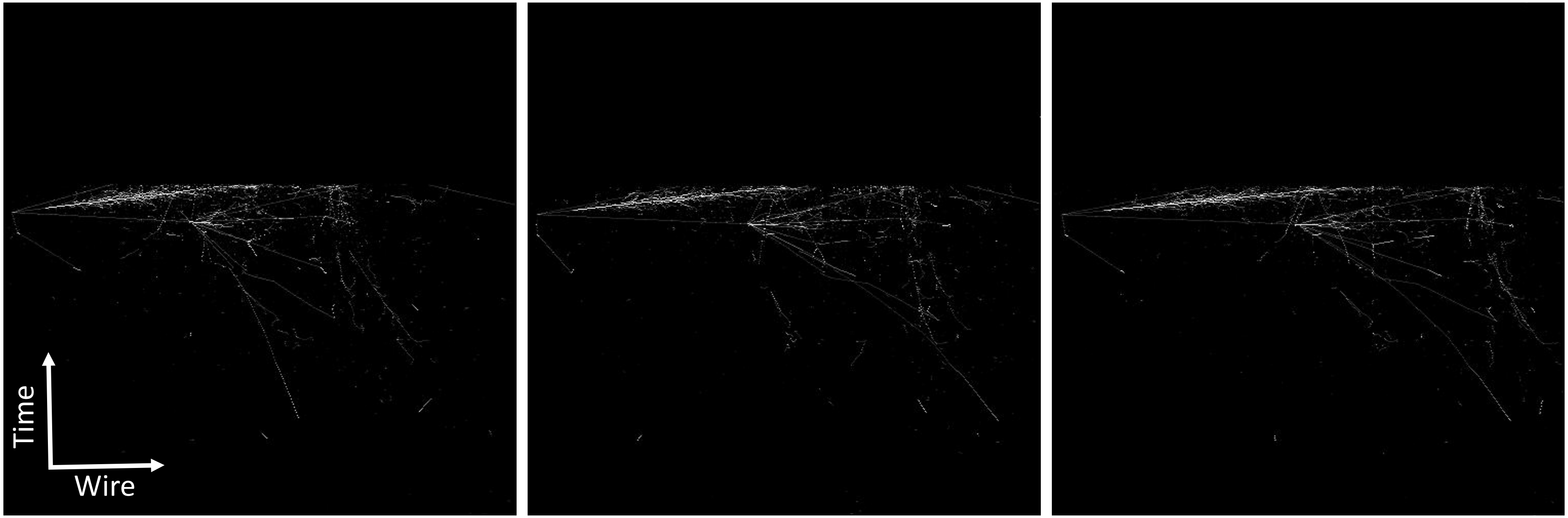
Ingredients for the CP-violation analysis

- We need to consider two signal channels and their backgrounds
  - Charged current $\overset{(-)}{\nu}_\mu$ disappearance – main background is NC $1\pi^{\pm}$
  - Charged current $\overset{(-)}{\nu}_e$ appearance – main background is NC $1\pi^0$
- Primary goal:
  - Classify the neutrino flavour as ν_e ,ν_μ , ν_τ or NC
- Secondary goal:
  - Can we go beyond flavor classification to individual interaction mode classification?
    - Different event classes will have different energy resolutions and systematic uncertainties, so separation can provide increased sensitivity

Far Detector Data

- The Far Detectors contain three wire readout planes
  - This provides three "images" of each neutrino interaction

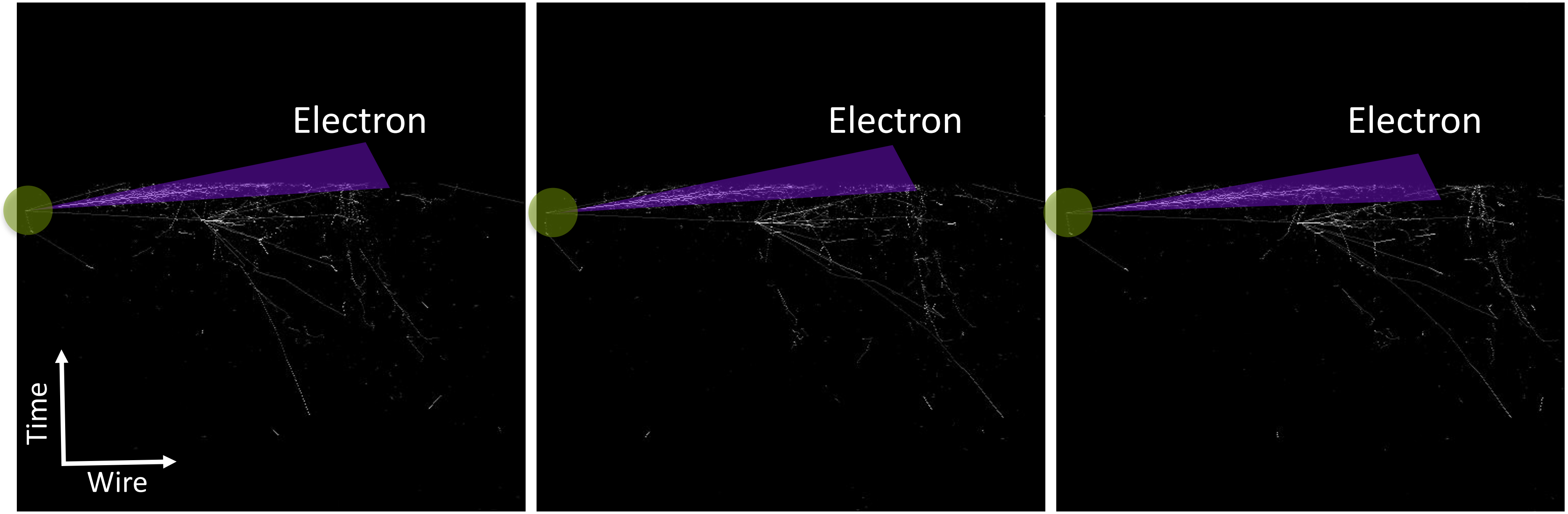
- Simulated electron neutrino interaction (signal)



Far Detector Data

- The Far Detectors contain three wire readout planes
  - This provides three "images" of each neutrino interaction

- Simulated electron neutrino interaction (signal)

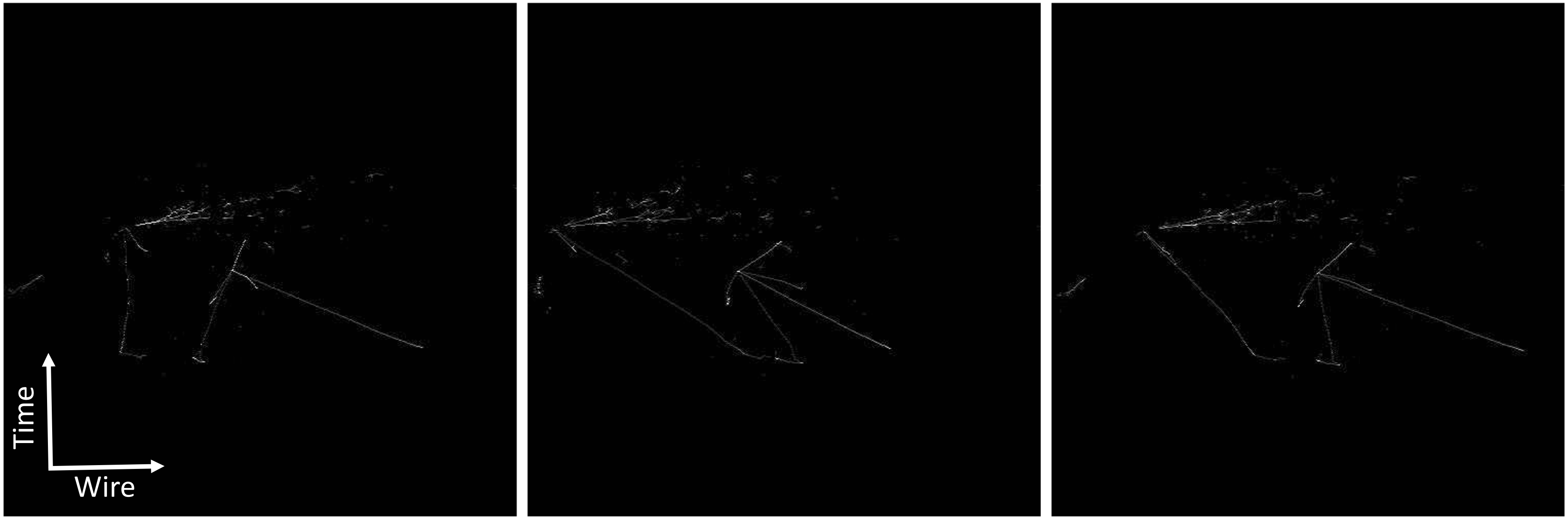


- Electron produces the highlighted shower, beginning at the vertex

Far Detector Data

- The Far Detectors contain three wire readout planes
  - This provides three "images" of each neutrino interaction

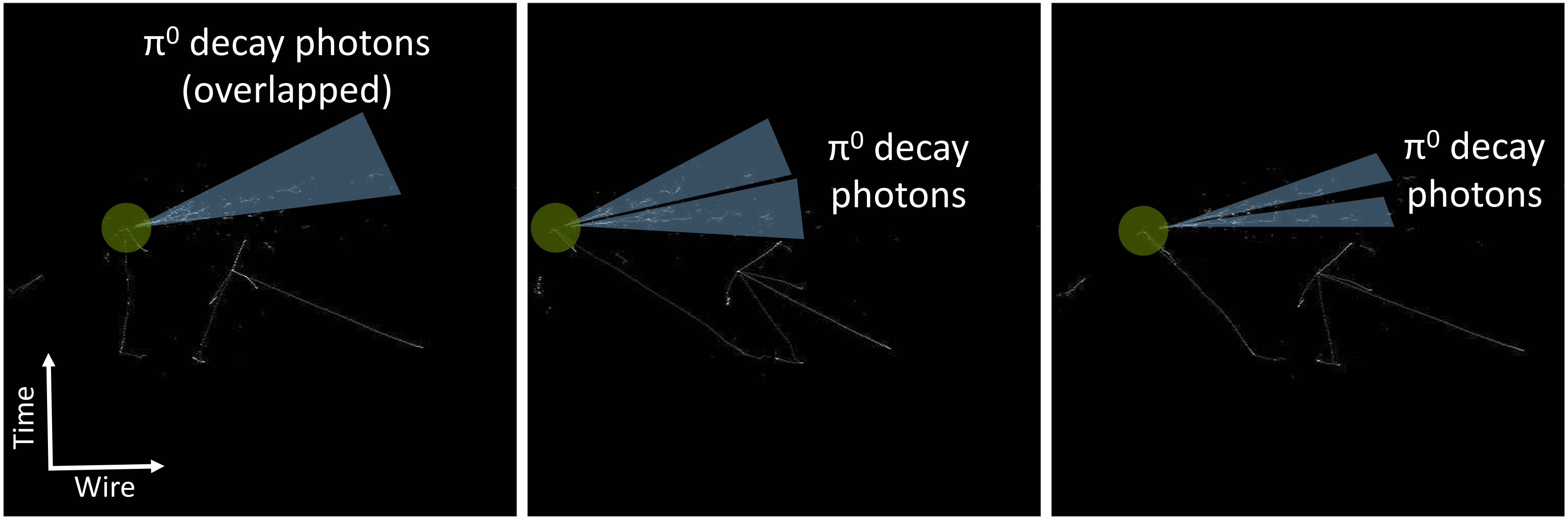
- Simulated neutral current π^0 interaction (background)



Far Detector Data

- The Far Detectors contain three wire readout planes
  - This provides three “images” of each neutrino interaction

- Simulated neutral current π^0 interaction (background)



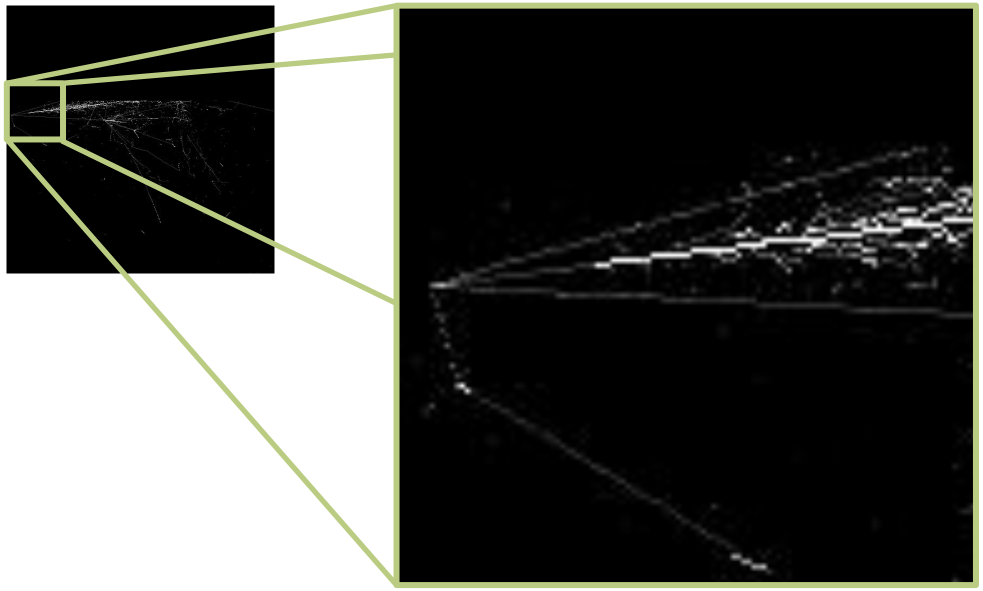
- π^0 decay photon showers are displaced from vertex

Image Recognition

- Fine-grained detail of LArTPCs lends itself to image recognition
- The human eye is a remarkably good image recognition tool
  - Once you know what to look for, it is fairly easy to find distinguishing features of different types of interactions
- Realistically, the experiment will produce too much data for scanning the interactions by eye
- We need to be able to train a computer to do this task
  - Recent years have shown rapid development of automated image recognition. One of the most promising approaches is the Convolutional Neural Network (CNN)

Convolutional Neural Networks

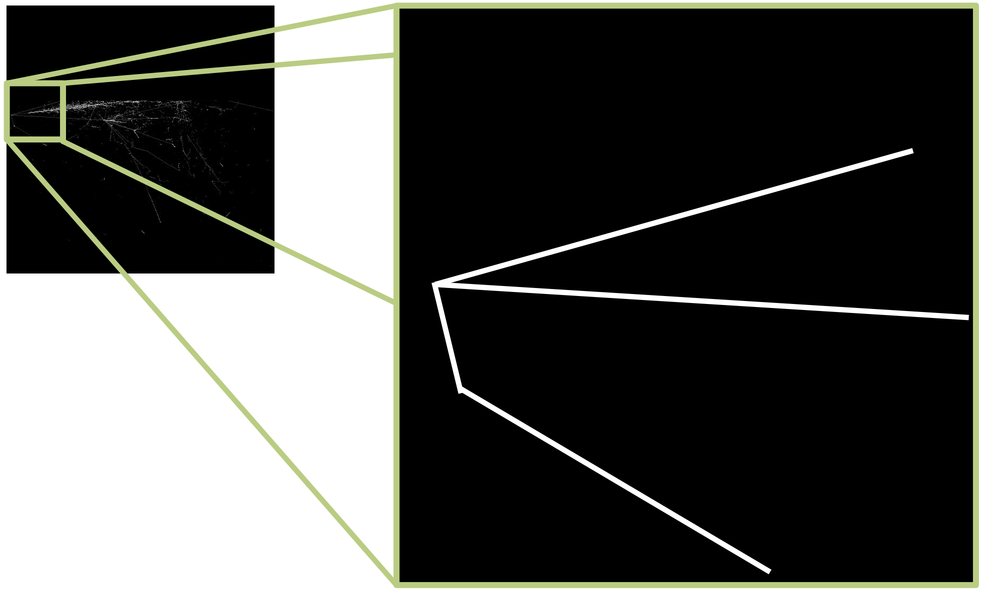
- CNNs are used to classify images by applying filters to small patches of the image (using a convolution)
- Scans over the image with a number of N x N pixel filters



- Each filter extracts some feature from the image

Convolutional Neural Networks

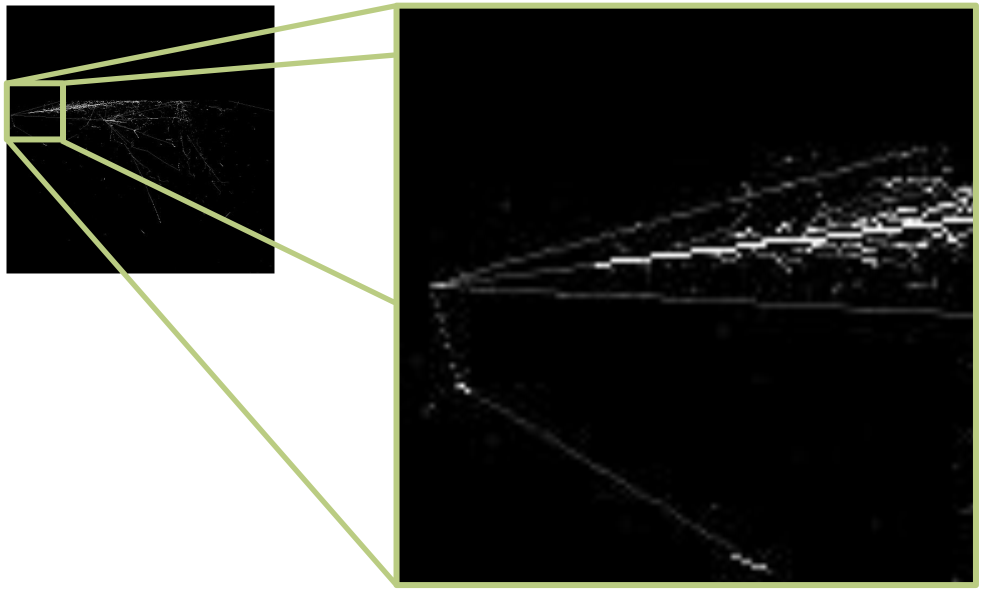
- CNNs are used to classify images by applying filters to small patches of the image (using a convolution)
- Scans over the image with a number of N x N pixel filters



- Each filter extracts some feature from the image
- For example, filter 1 may find tracks

Convolutional Neural Networks

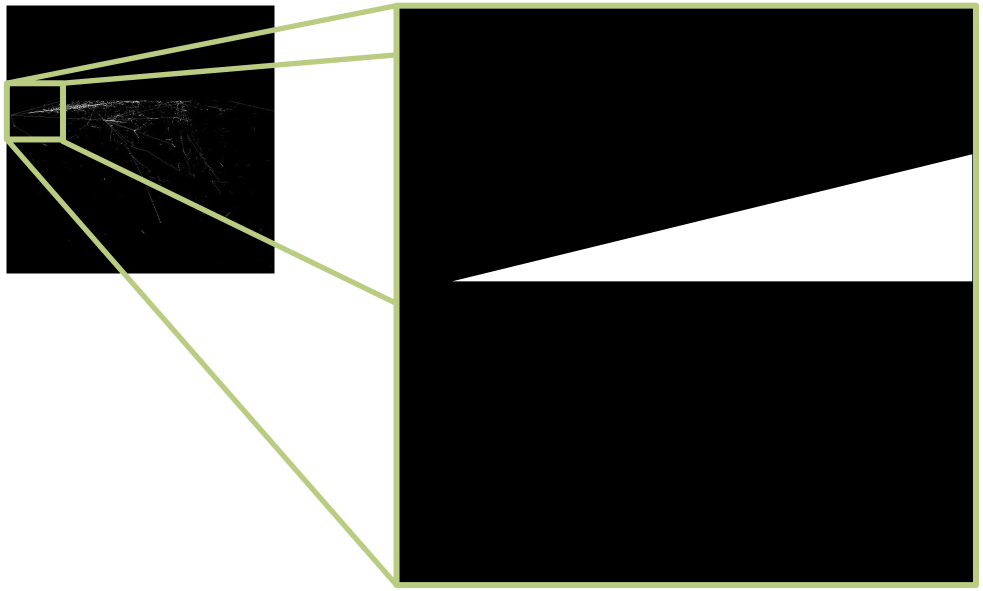
- CNNs are used to classify images by applying filters to small patches of the image (using a convolution)
- Scans over the image with a number of N x N pixel filters



- Each filter extracts some feature from the image
- For example, filter 1 may find tracks

Convolutional Neural Networks

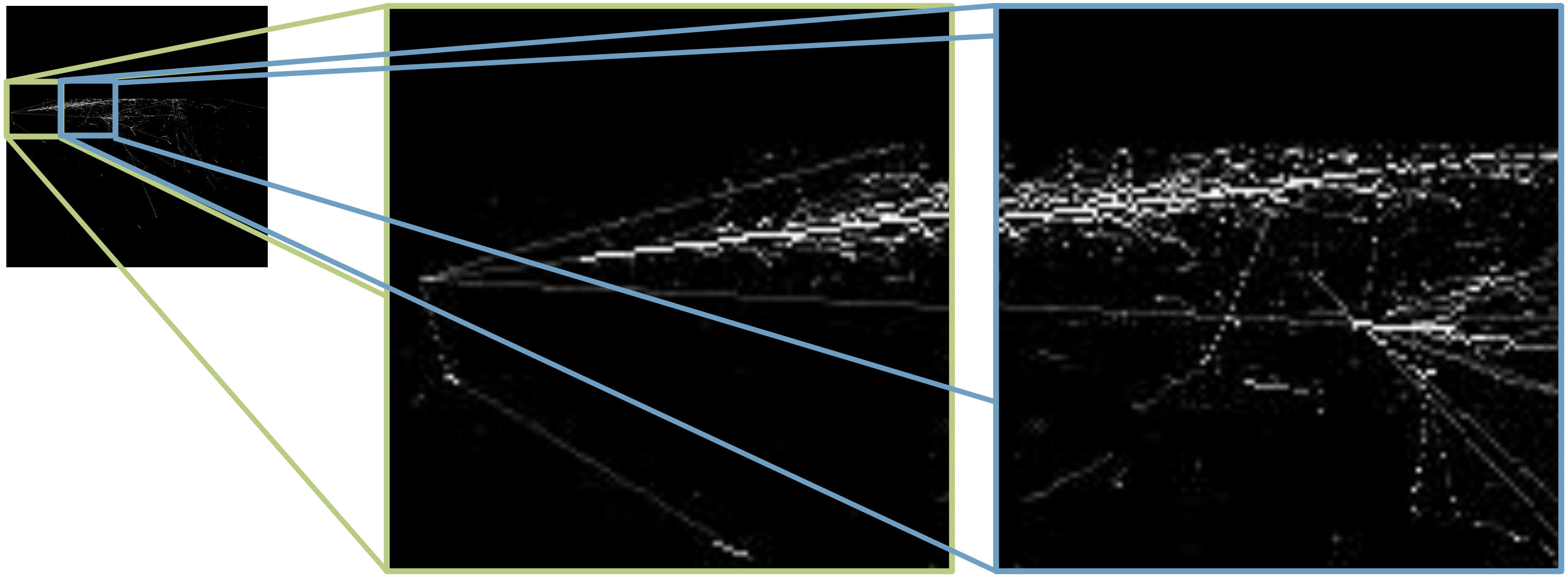
- CNNs are used to classify images by applying filters to small patches of the image (using a convolution)
- Scans over the image with a number of N x N pixel filters



- Each filter extracts some feature from the image
- For example, filter 1 may look for tracks
- Filter 2 might look for showers

Convolutional Neural Networks

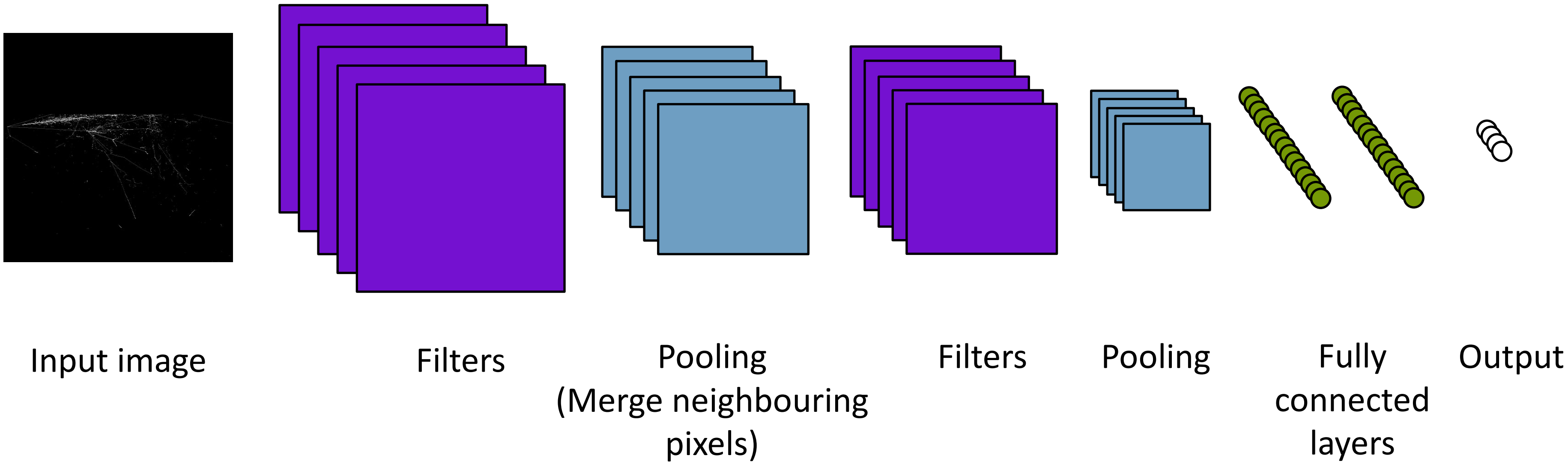
- CNNs are used to classify images by applying filters to small patches of the image (using a convolution)
- Scans over the image with N x N pixel filters



- Then move onto the next patch of the image and repeat the process

Convolutional Neural Networks

- The output from each filter then forms the basis of the next layer which can include further filters

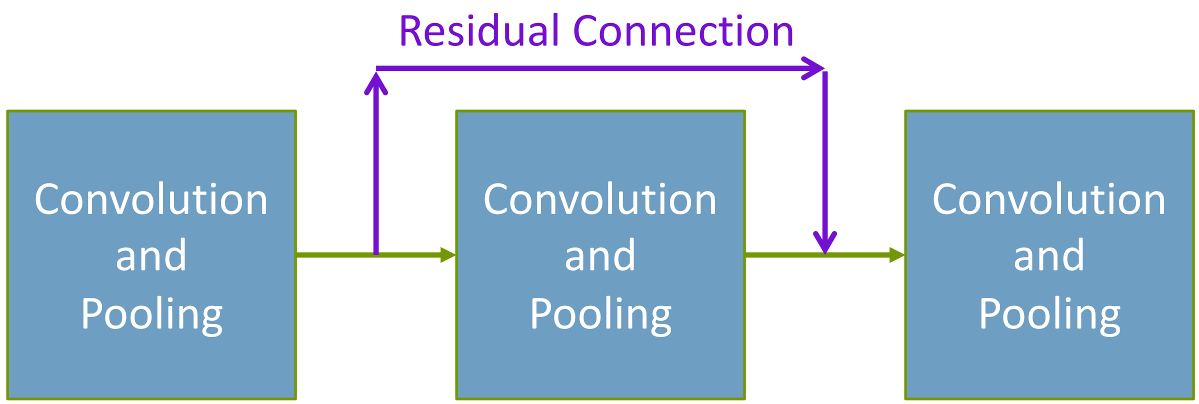


- Different architectures can be considerably more complex than the above toy example

DUNE CNN introduction

- The initial DUNE CVN was based on the NOvA implementation

- In the last year we have moved to a completely new architecture and framework
  - We now use a SE-ResNet[1,2] based architecture

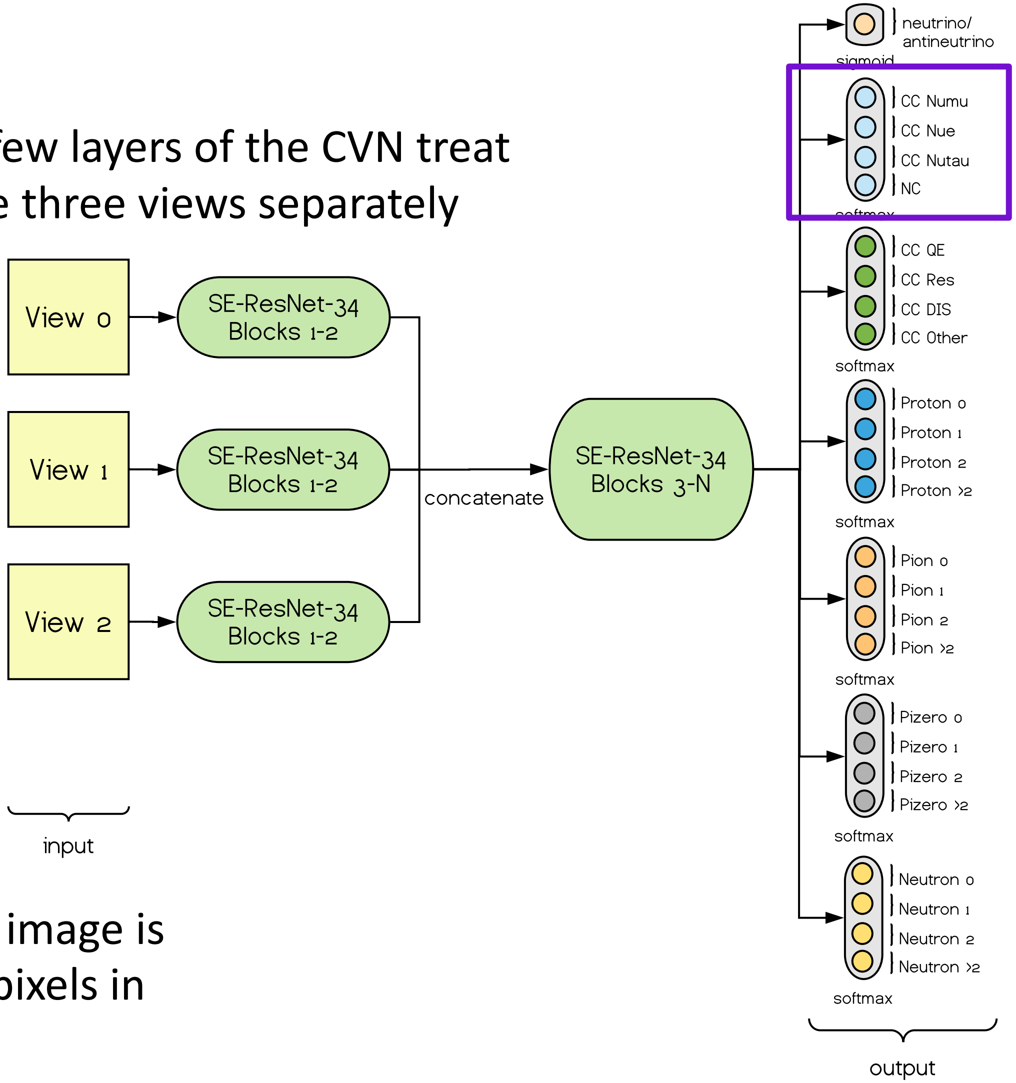


  - Helps preserve the fine-grained detail deeper into the network

[1] H. Kaiming et al., Deep residual learning for image recognition, CoRR, arXiv 1512.03385, 2015
[2] J. Hu et al., Squeeze-and-Excitation Networks, arXiv 1709.01507, 2017

Architecture Overview

First few layers of the CVN treat the three views separately

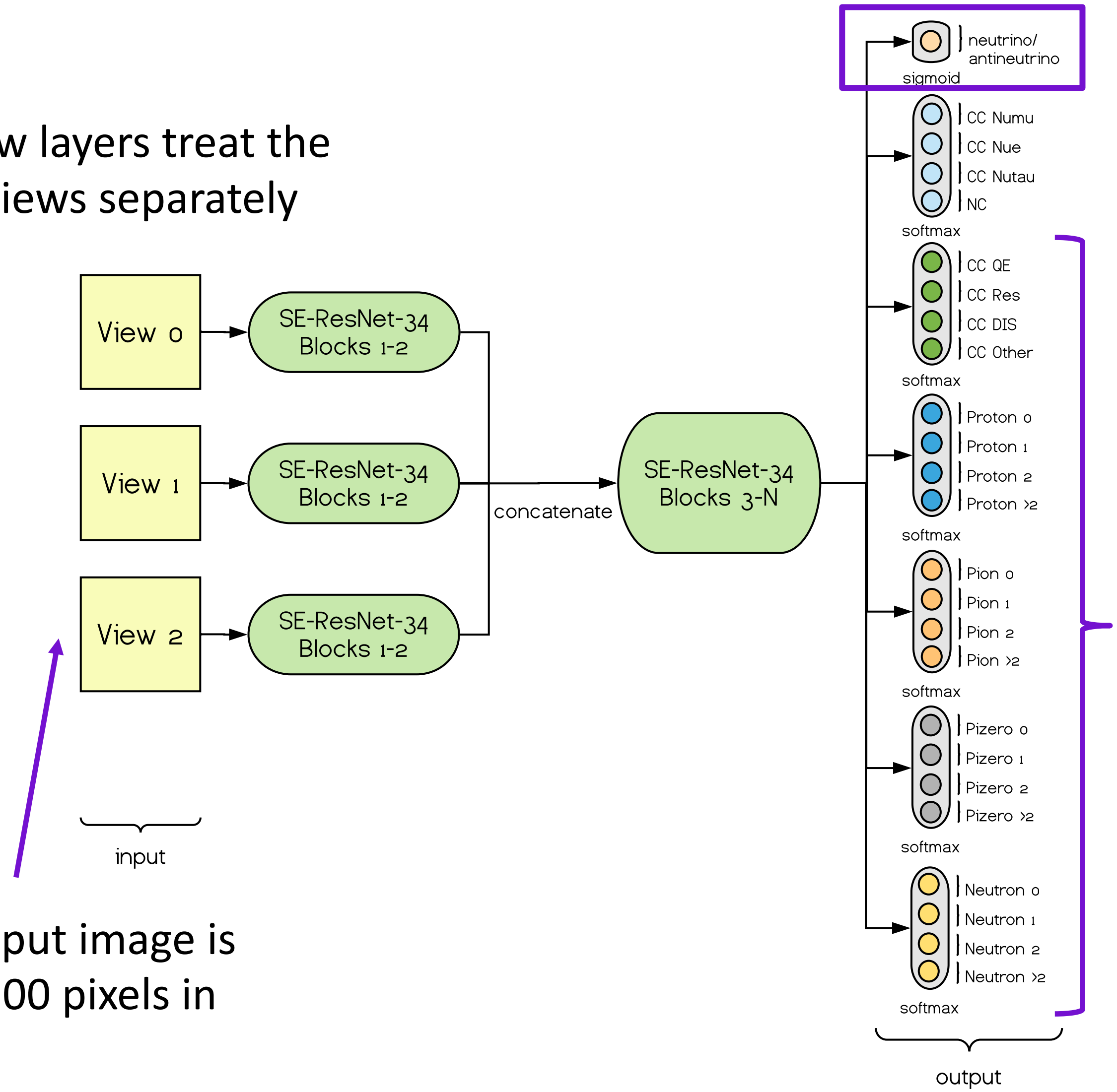


Primary output: Flavour identification

Each input image is 500 x 500 pixels in size

Architecture Overview

First few layers treat the three views separately



Neutrino / antineutrino

Secondary outputs: Exclusive final-states and particle counting

Each input image is 500 x 500 pixels in size

Training and using the CNN

- Use millions of images of simulated neutrino interactions with the true neutrino flavour known
  - Allows the CNN to learn the features of each type of neutrino interaction
  - The CNN filters are not predefined – it needs to learn which filters to use to extract the information required to classify events

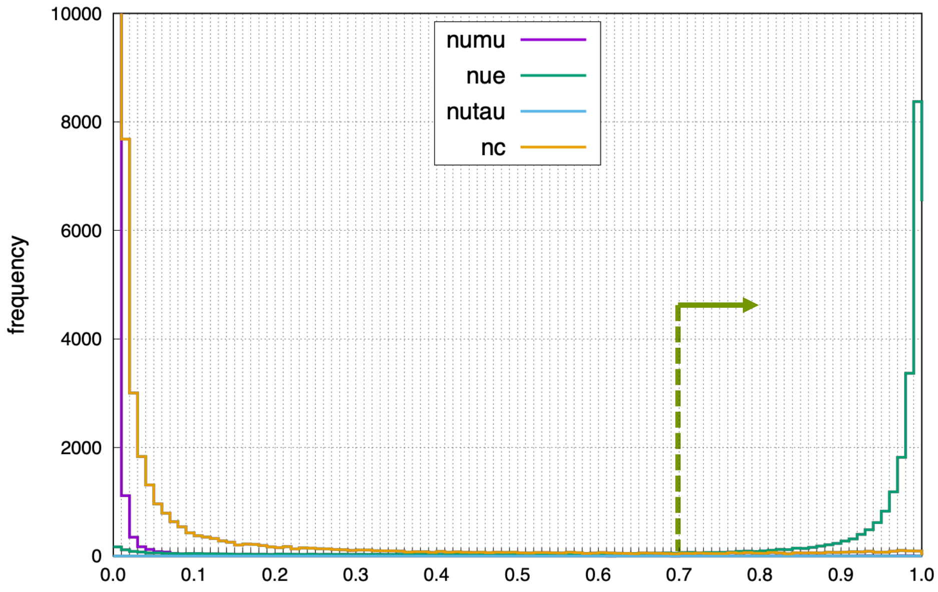
- Once the CNN is trained it is applied to images with no truth information attached – eventually the experimental data

- The CNN gives probabilities for each event to be the following:
  - Charged-current ν_e, ν_μ, ν_τ and neutral-current (all flavours)
  - Outputs sum to one
  - Use these probabilities for the event selection

Selecting Electron Neutrinos

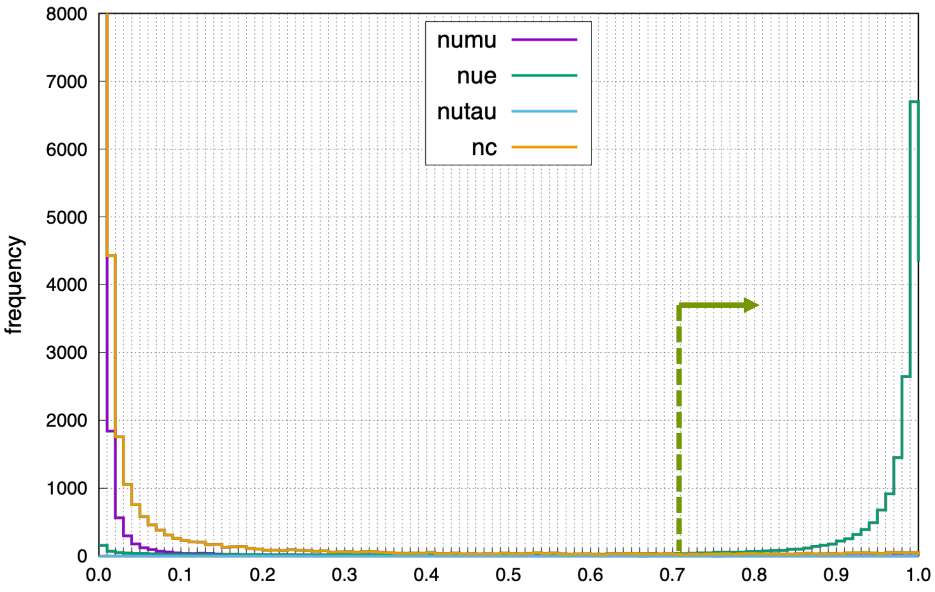
- Electron neutrino probability spectra from the DUNE CVN

Neutrino beam



ν_e probability

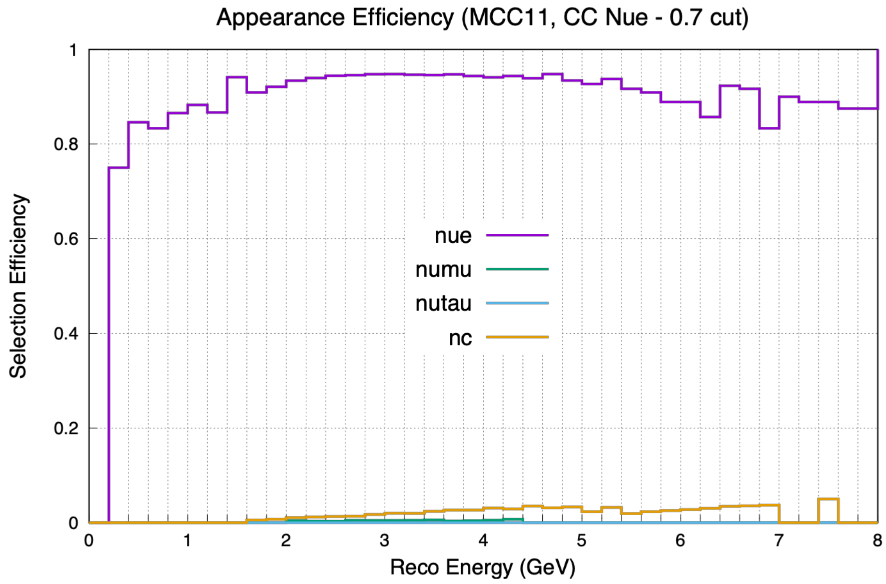
Antineutrino beam

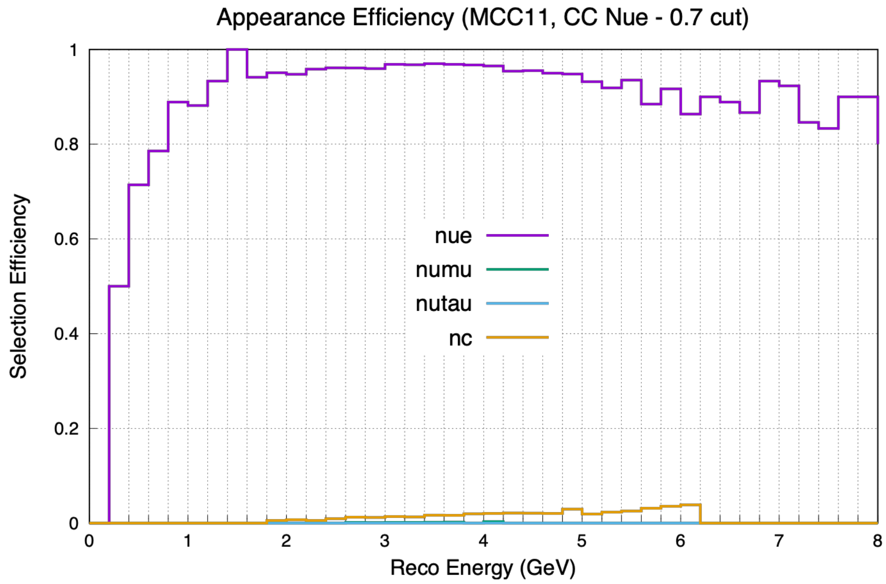


ν_e probability

Electron Neutrino Efficiency

- Select all events that are more than 70% likely to be electron neutrinos
- Over 90% selection efficiency in the flux peak
- Efficiency better for antineutrinos due to typically cleaner final state (neutron instead of proton)

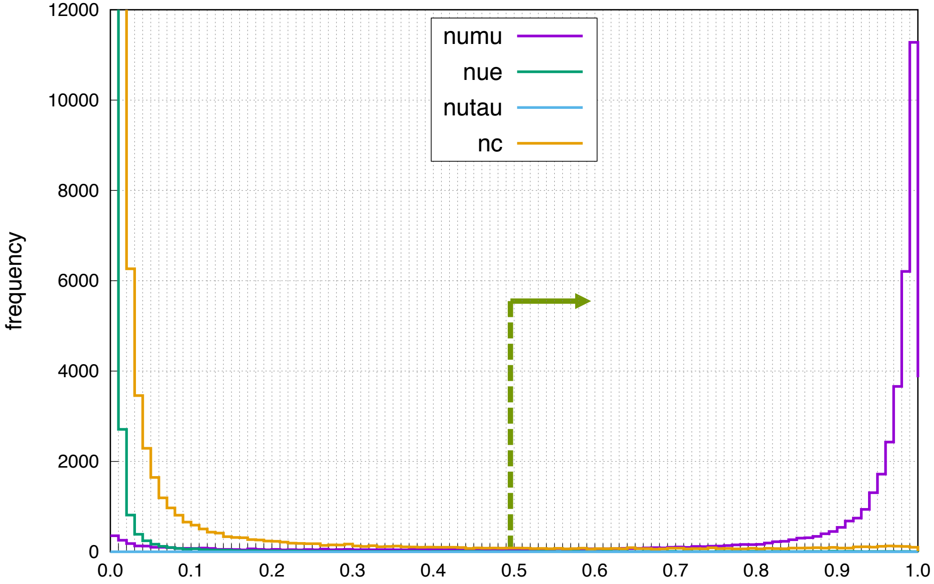




Selecting Muon Neutrinos

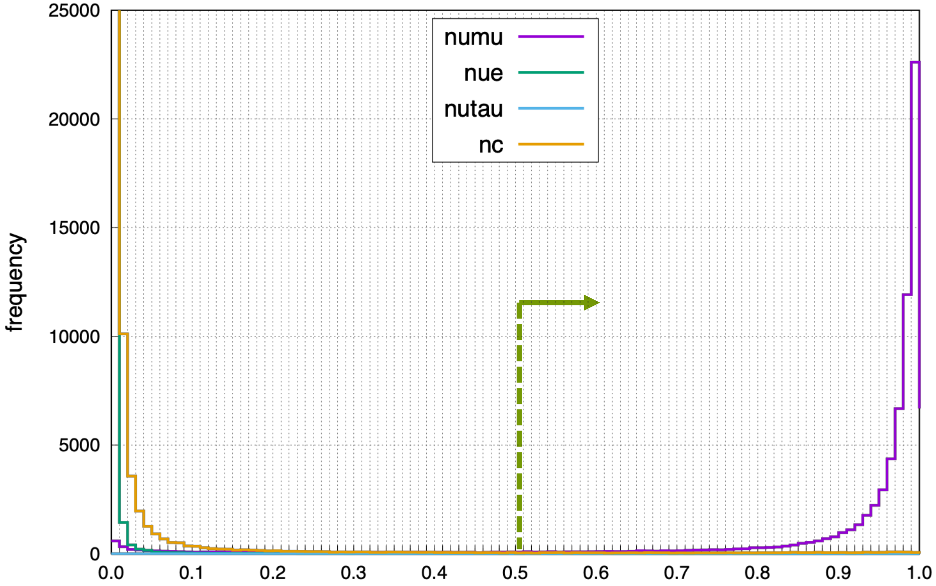
- Muon neutrino probability spectra from the new CVN

Neutrino beam



ν_μ probability

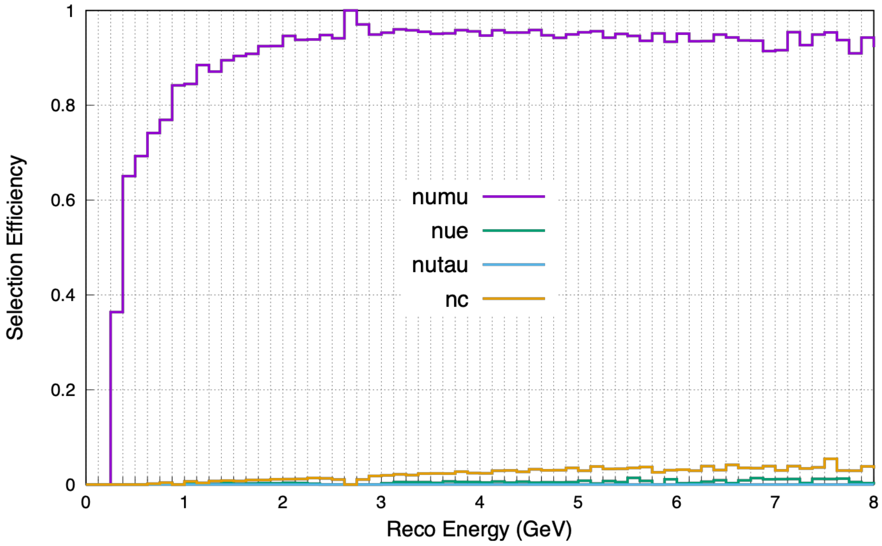
Antineutrino beam

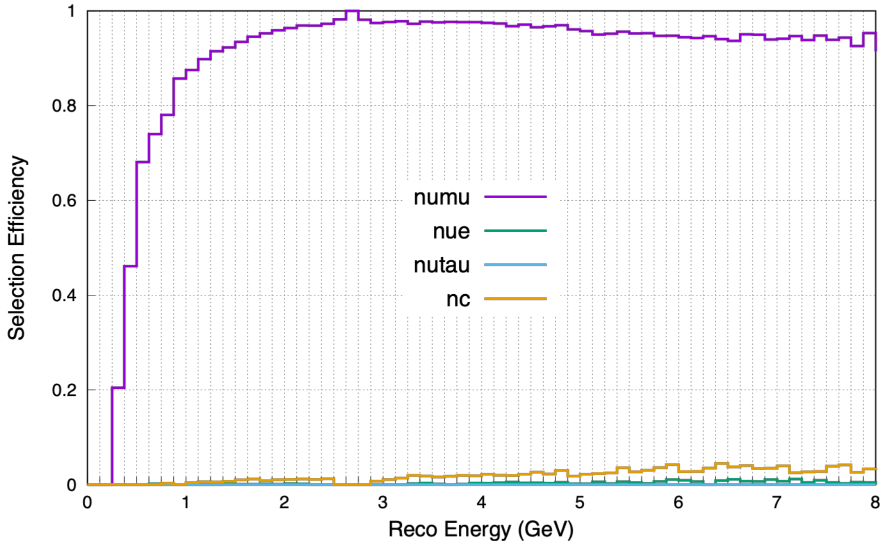


ν_μ probability

Muon Neutrino Efficiency

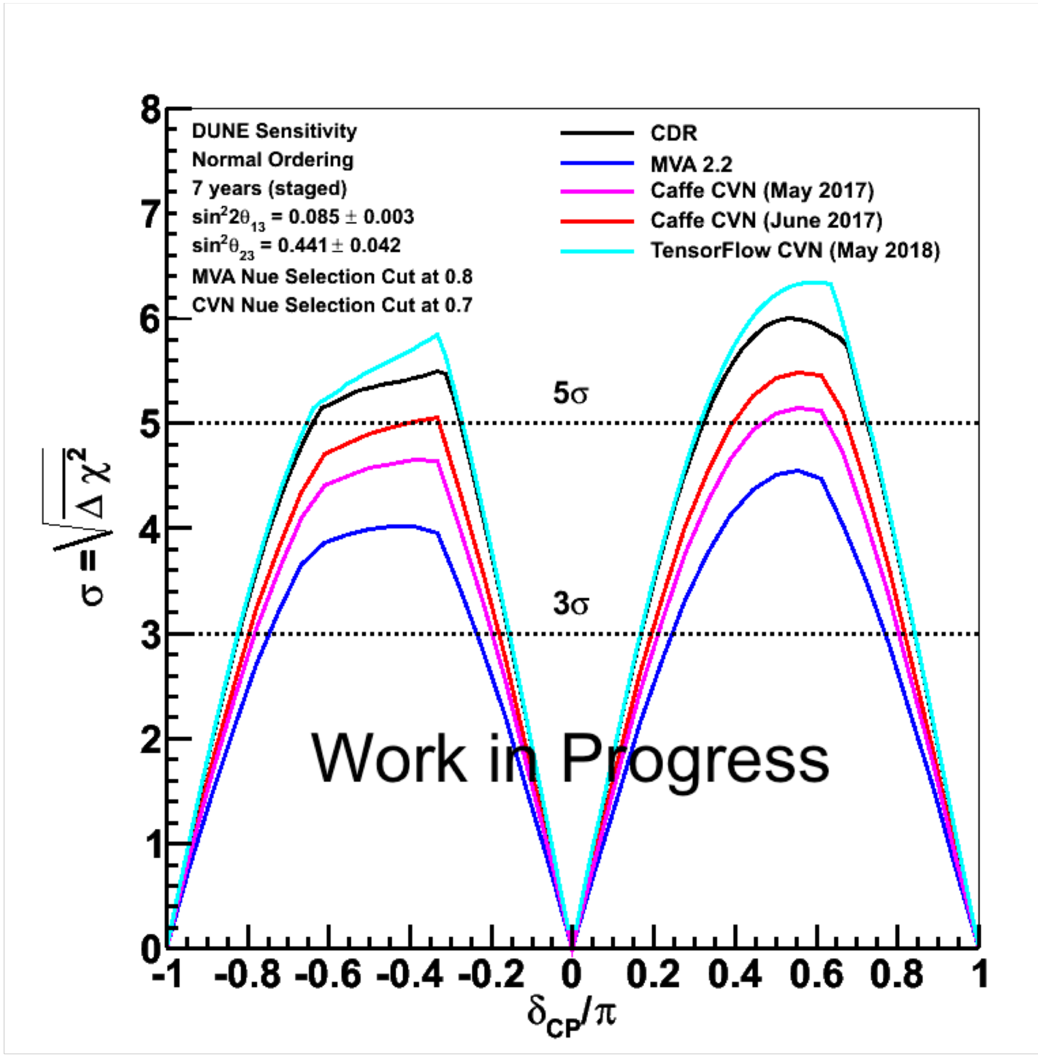
- Select all events that are more than 50% likely to be muon neutrinos
- Over 90% selection efficiency in the flux peak
- Efficiency better for antineutrinos due to typically cleaner final state (neutron instead of proton)





CP-Violation Sensitivity – May 2018

- Same selection criteria
  - P(ν_e) > 70%
  - P(ν_μ) > 50%
- Very large improvement over the previous CVN
- Exceeded the DUNE conceptual design report sensitivity
  - Very big milestone for DUNE!



Robustness

- We will use protoDUNE to test the CVN on real data
  - There are no neutrinos at protoDUNE, but we can use single particles extracted from events to approximate simple neutrino interactions
- Select individual reconstructed objects and pass into the CVN
  - Cosmic muon tracks mimic CCQE ν_μ interactions
  - Beam electron showers mimic CCQE ν_e events
  - The CVN should return classifications of CC ν_μ and CC ν_e with no hadronic system, respectively
- We will also use fake data studies to ensure robustness against systematic effects, including those from alternative event generators

Summary

- The DUNE CVN provides powerful neutrino interaction flavour classification
  - Hope to demonstrate good performance of exclusive final-states in the coming months
- ProtoDUNE provides an excellent opportunity to test the CVN on data using single particles to mimic simple neutrino interactions
- Further improvements will provide diminishing returns on the experimental sensitivity
  - The focus now shifts to ensuring robustness and equal performance when applied to data and simulation