n-nbar oscillations in DUNE with CNNs

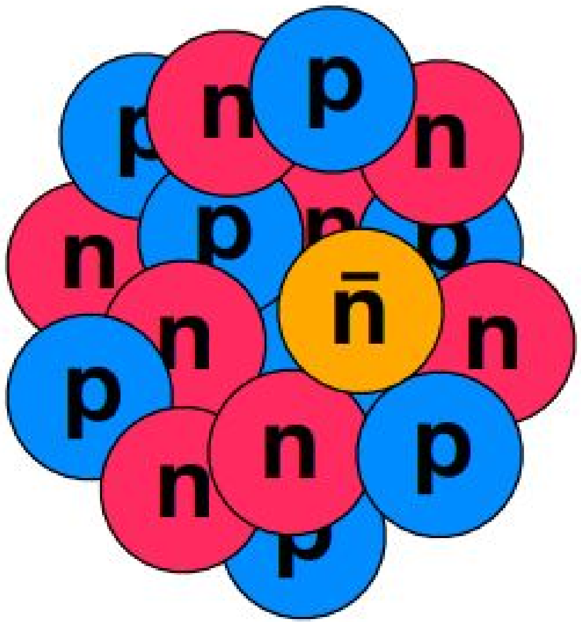
Yuyang Zhou (Columbia University)
Jeremy Hewes (University of Manchester)
Georgia Karagiorgi (Columbia University)

DUNE Physics Week
November 15th, 2017

Overview

1. Introduction to n-nbar
2. Analysis method
3. Results in DUNE
4. Current work

Neutron-antineutron oscillation



1. Inside the nucleus, a neutron spontaneously oscillates into an antineutron
2. Antineutron annihilates with nearby nucleons
3. Isotropic emission of 2-6 pions with ~2 GeV invariant mass and ~300 MeV net momentum
4. Violates baryon number conservation!

Motivation: Baryogenesis, GUT

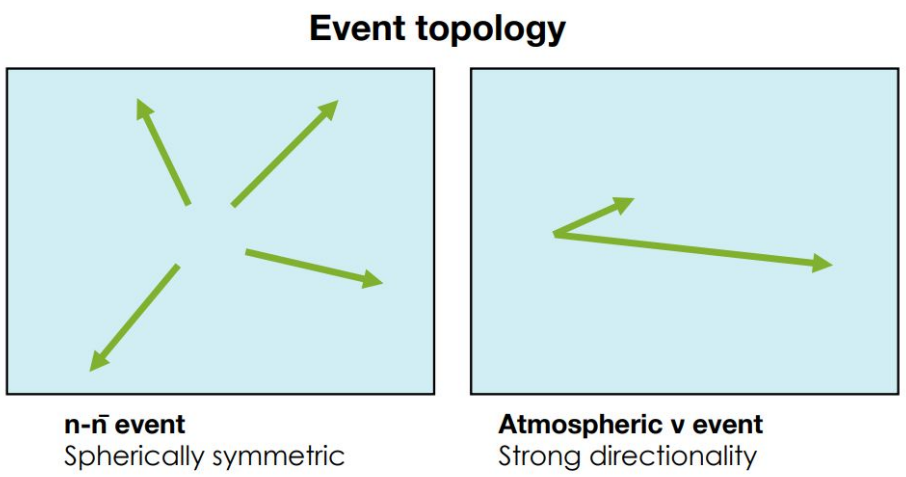
Right: n-nbar/p-nbar branching ratios, Super-K (arXiv 1109.4227)

$\bar{n}+p$		$\bar{n}+n$	
$\pi^+\pi^0$	1%	$\pi^+\pi^-$	2%
$\pi^+2\pi^0$	8%	$2\pi^0$	1.5%
$\pi^+3\pi^0$	10%	$\pi^+\pi^-\pi^0$	6.5%
$2\pi^+\pi^-\pi^0$	22%	$\pi^+\pi^-2\pi^0$	11%
$2\pi^+\pi^-2\pi^0$	36%	$\pi^+\pi^-3\pi^0$	28%
$2\pi^+\pi^-2\omega$	16%	$2\pi^+2\pi^-$	7%
$3\pi^+2\pi^-\pi^0$	7%	$2\pi^+2\pi^-\pi^0$	24%
		$\pi^+\pi^-\omega$	10%
		$2\pi^+2\pi^-2\pi^0$	10%

Current status of experimental n-nbar searches:
- Super-K has set the n-nbar lifetime at 2.7 x 10^8 s
- No searches have been done in LArTPCs!

Analysis method

Analysis overview



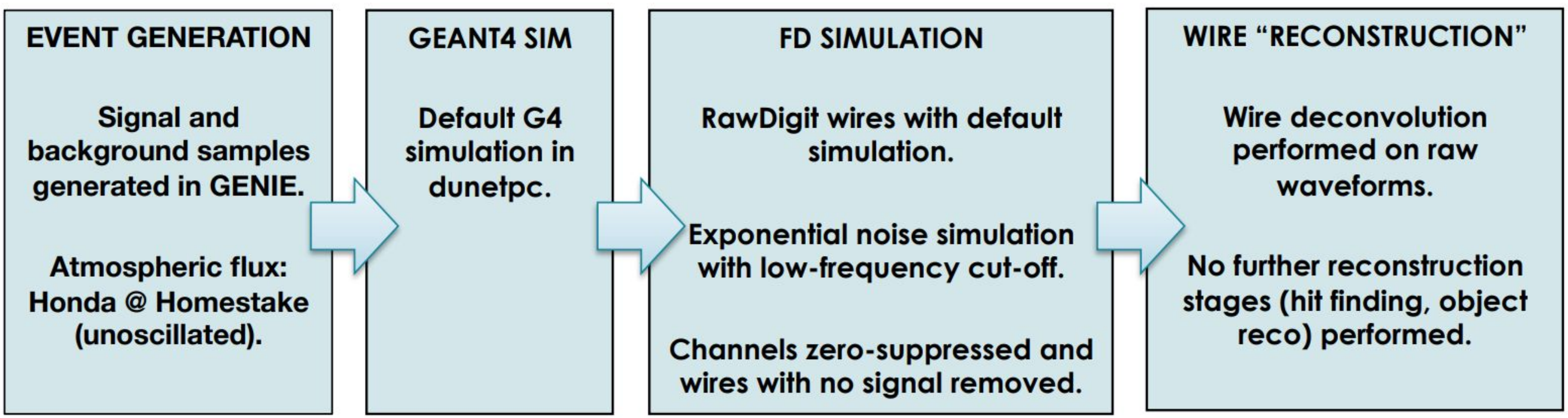
Principle: use convolutional neural networks (CNNs) to classify n-nbar events inside DUNE

- N-nbar events and its predominant background, atmospheric neutrinos events, can be distinguished by their unique "shapes" inside the DUNE detector
  - Use deep learning image classification to distinguish between the two

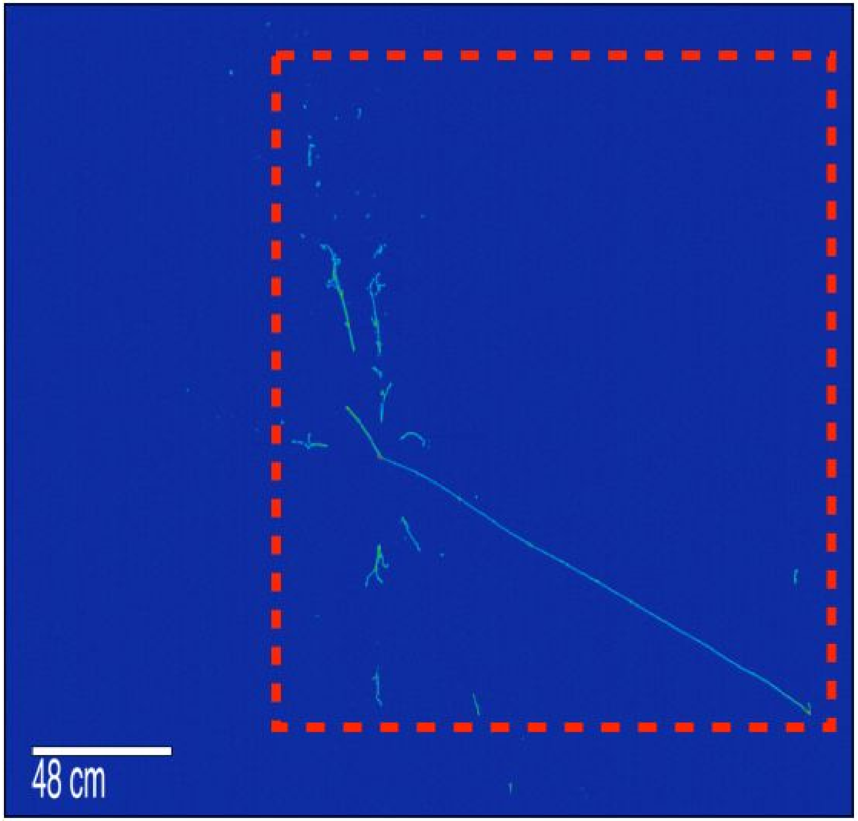
1. Train a network on Monte Carlo (MC) simulations of n-nbar (signal) and atmospheric neutrinos (background)
2. Have the network classify an unknown, separate MC sample of signal and background

Simulation stages

- MC samples generated using **dunetpc v06_24_00**
- APA geometry: 1x2x6
- Images created from **recob::Wire** after deconvolution and without hit finding



Network input image



An n-nbar event in the LArTPC MC, with an ROI selected

1. Cross-APA stitching is nontrivial, so select **single** APA with highest ADC sum

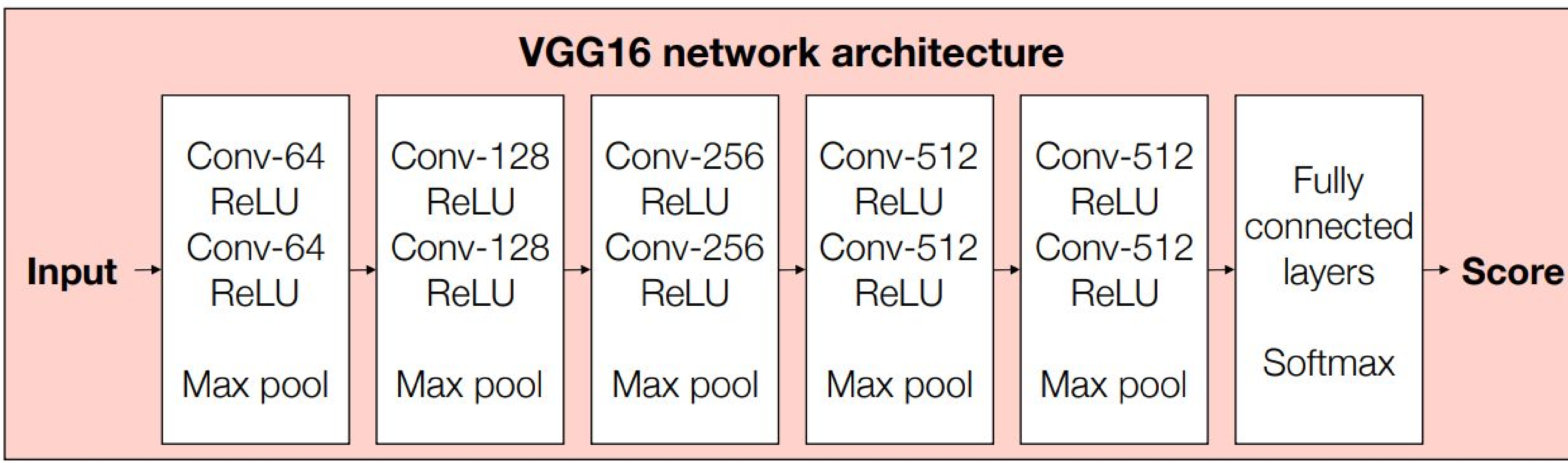
2. Identify region of interest (ROI)

3. ROI embedded onto 600 x 600 empty image and passed to network

Network structure

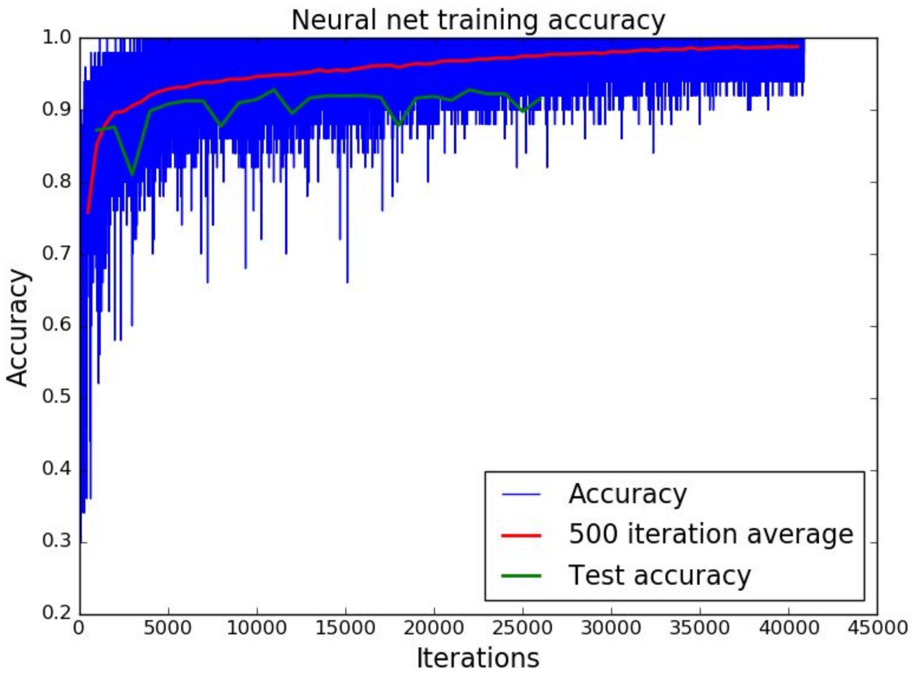
- Software used: **Caffe** CNN framework modified for LArTPC
  - VGG16 architecture

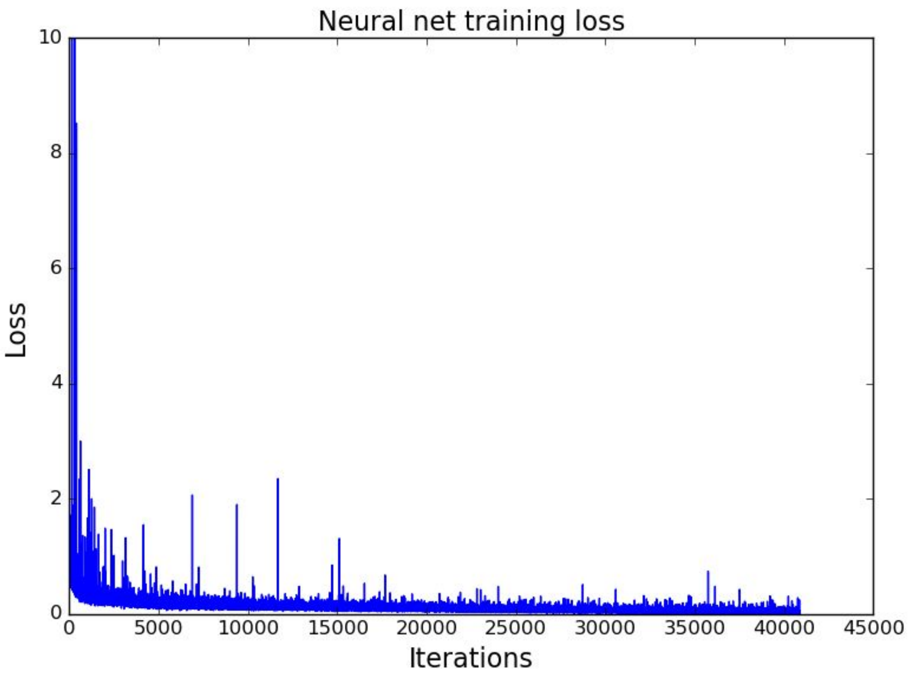
- Network uses multiple layers to apply a series of convolutions to the input image
  - Outputs a score on a range from 0 to 1, with 0 being background-like and 1 being signal-like



Network training

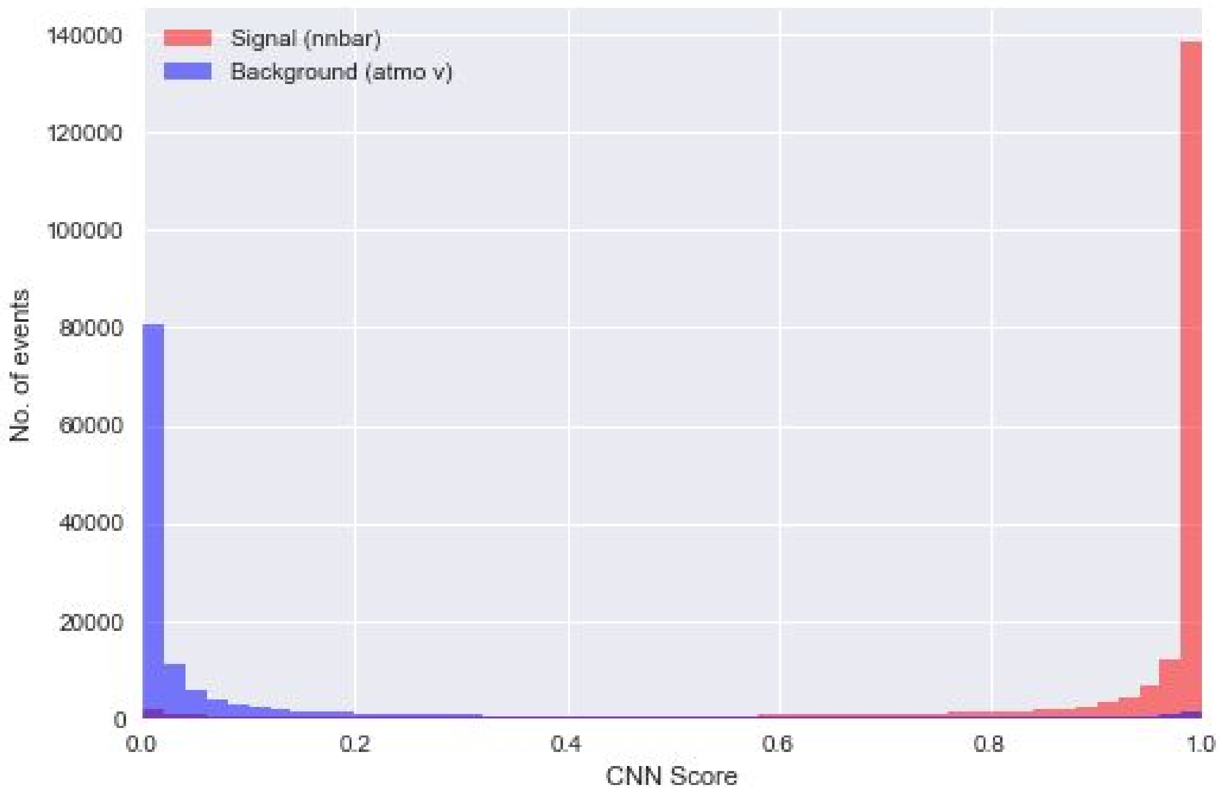
- Successfully trained on 50,000 signal and background images
- During training, network monitors **accuracy** and **loss**
  - Over time, accuracy increases and loss decreases





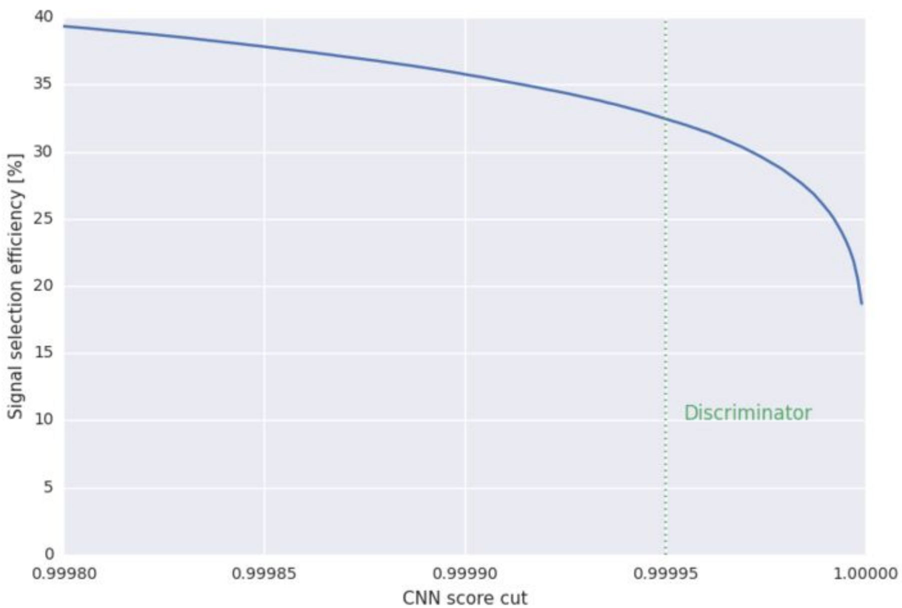
Results in DUNE

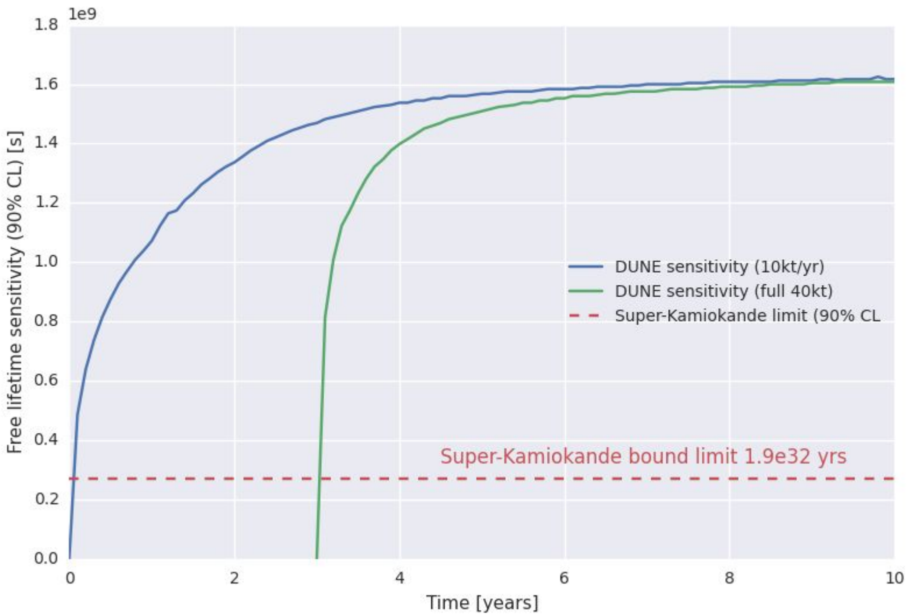
Network inference



- Test the network by having it classify a separate MC sample of 200,000 signal and background images
  - Network scores each image from 0 (background-like) to 1 (signal-like)

Sensitivity calculation





- By choosing an optimal CNN score cut of 0.99995, we obtain the signal selection efficiency of **14%** and background rejection rate of **99.997%**
- Over 10 years, the projected sensitivity is **1.6 x 10^9** s (90% CL)
  - Factor of ~5 improvement over current best limit, 2.7 x 10^8 s (Super-K)

Feature validation

Motivation:

- Is the network actually learning on physical event features?
- Network performance on events near the APA boundary (no fiducialization applied beforehand)

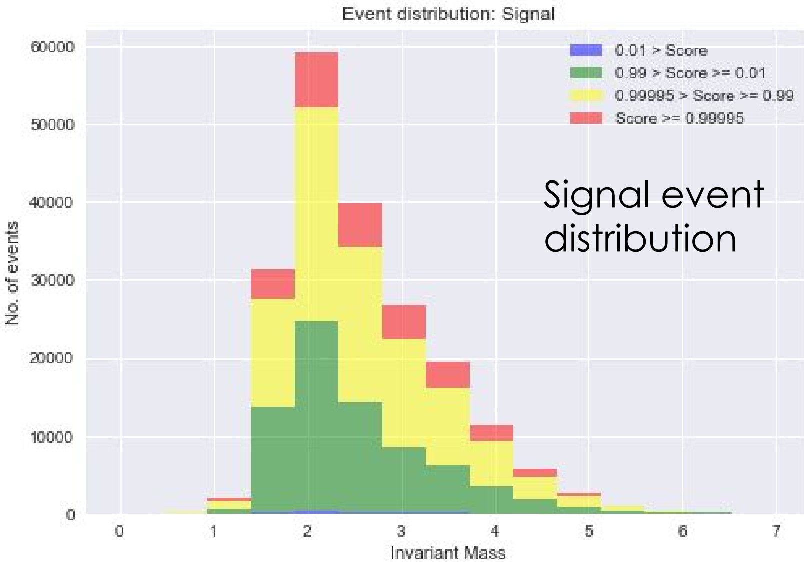
Our study:

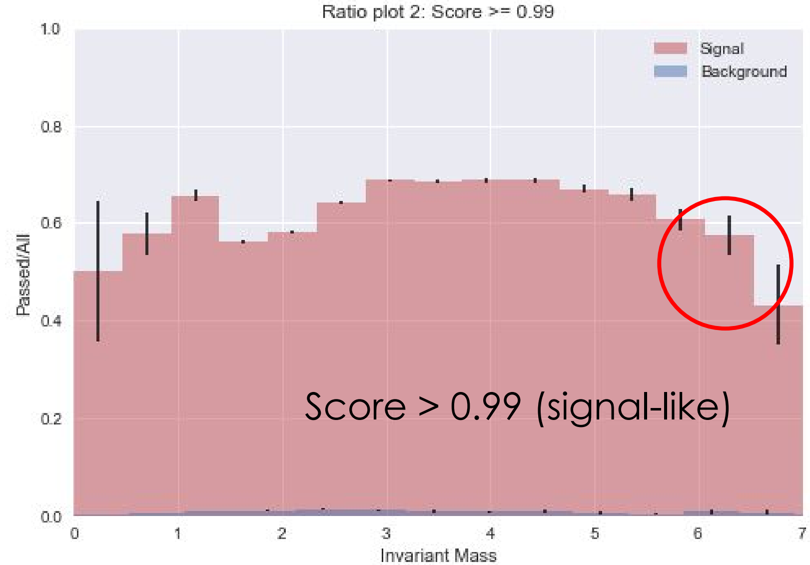
- Look at MC truth variables which we think the network "should" be learning on (e.g. net momentum, invariant mass)

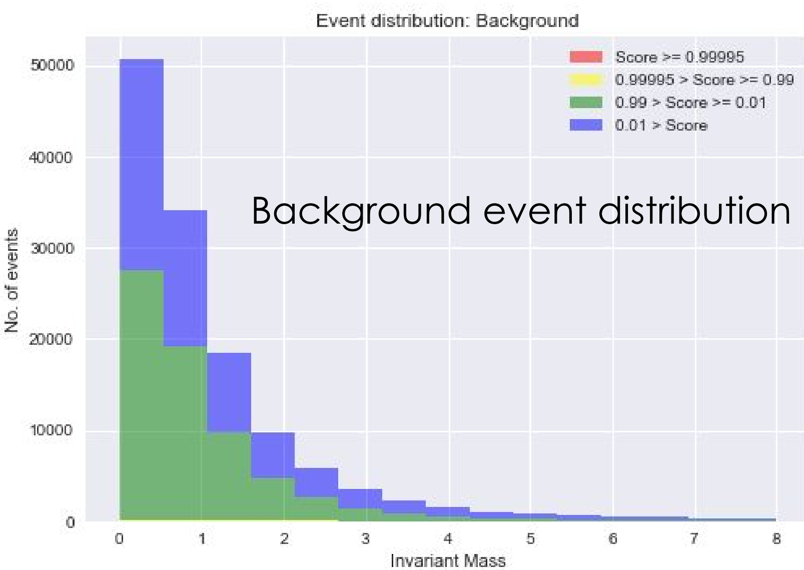
- 4 plots for each variable:

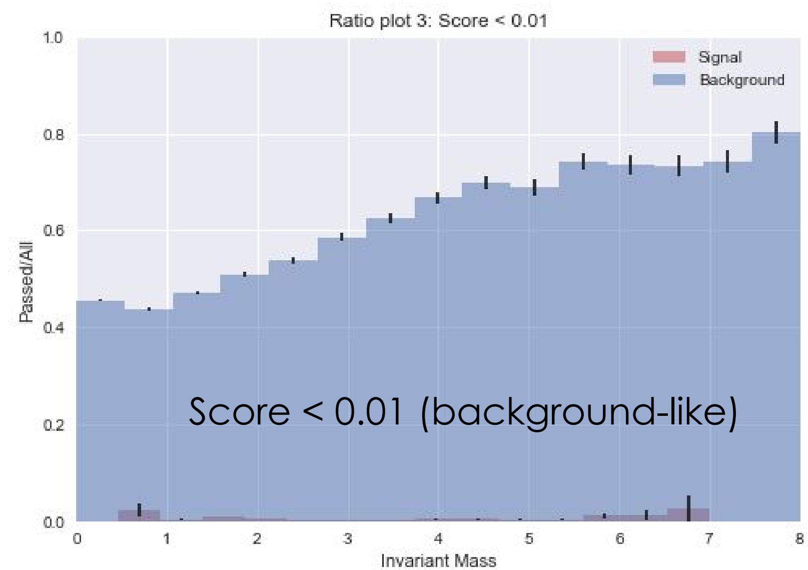
1. Event distribution for signal
2. Event distribution for background
3. Ratio plot: Score > 0.99 (classified signal-like)
4. Ratio plot: Score < 0.01 (classified background-like)

Invariant Mass

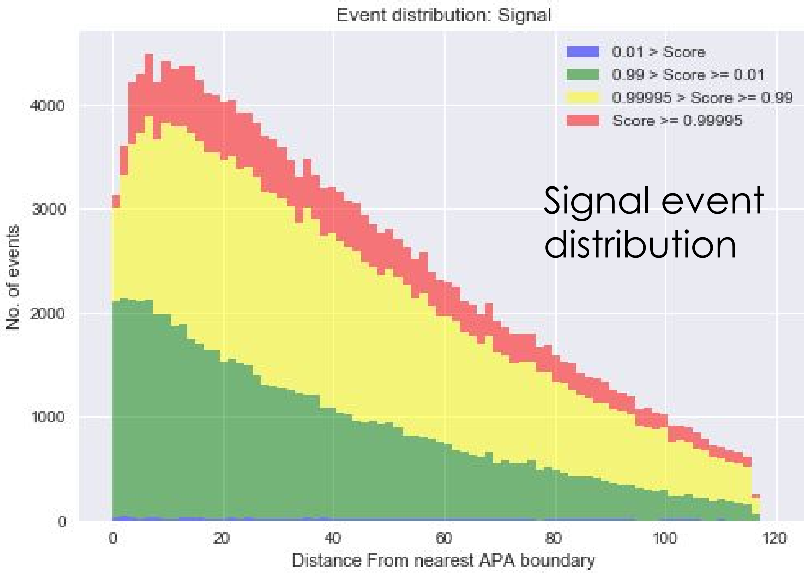


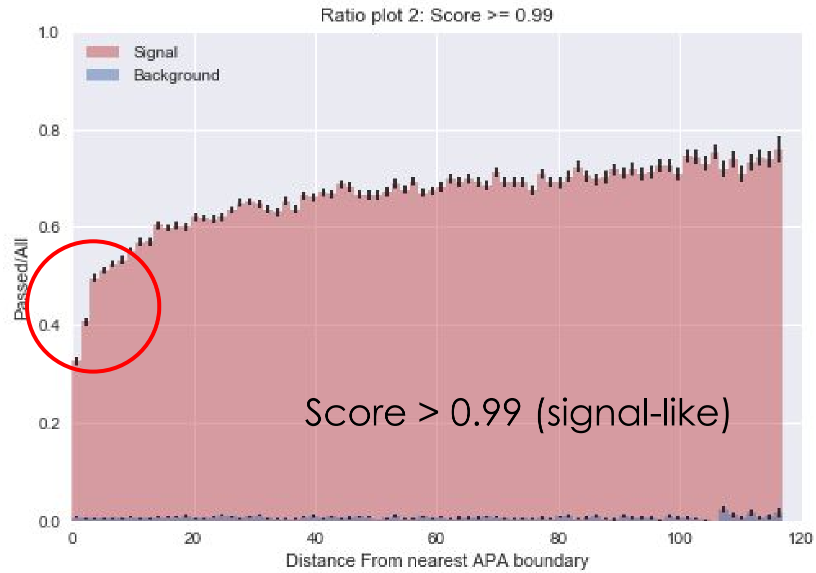


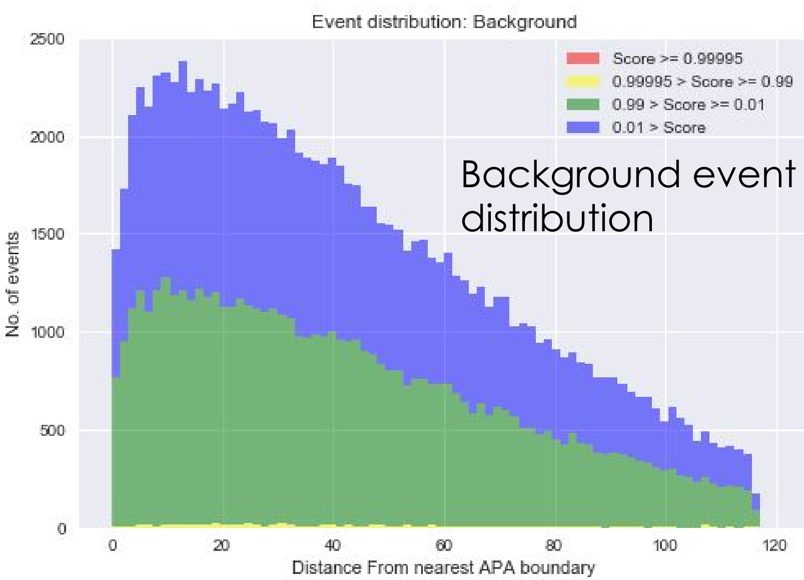


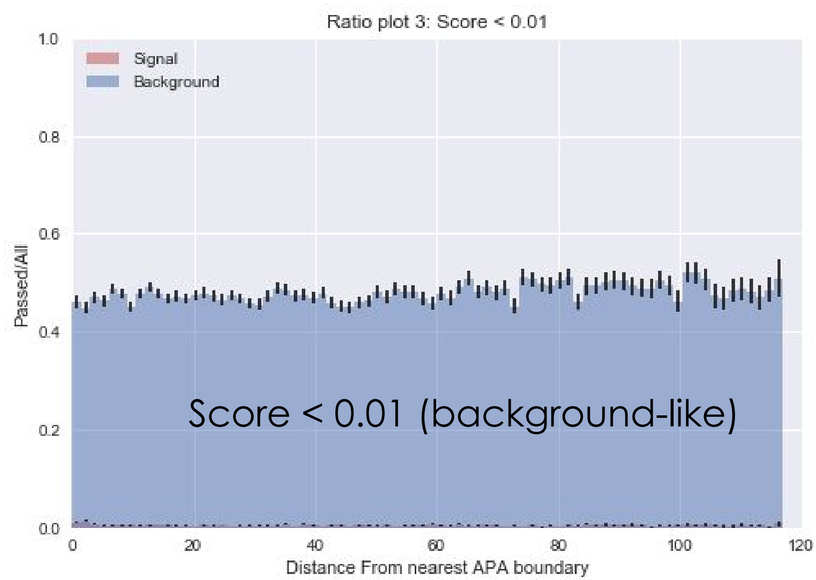


Distance from nearest APA boundary









Current work

Setting the n-nbar lifetime in MicroBooNE

- We have shown the network to be effective in classifying MC simulations in DUNE
  - However, need to show network can also classify real data

- Introducing MicroBooNE
  - DUNE's predecessor, collecting data since late 2015
  - 170t LArTPC compared to DUNE's 40kt
  - Due to size, we expect sensitivity to be worse than DUNE

Goal: benchmark this methodology in MicroBooNE with real detector data

- MC generated with uboonecode **v06_35_00**
- Images produced from **recob::Wire**

Using MicroBooNE's existing data, we will set the world's first n-nbar lifetime in liquid argon!

MicroBooNE CNN training and inference on MC

Repeating the study:

- Using the same CNN architecture, successfully trained a network on 50,000 signal and background events
- Tested the network with 200,000 events of MC signal and background images
- Network performs well overall, although the sensitivity is much worse than DUNE, as expected

The same methodology produces comparable results in MicroBooNE!

Inference on data

However, network runs into issues classifying actual detector data

- Looking at cosmics, the network heavily misclassifies real data
- Network very sensitive to differences between data and MC

To probe this issue, changed the noise simulation method

- Obtained very different score distributions

This means the network highly susceptible to differences in noise simulation

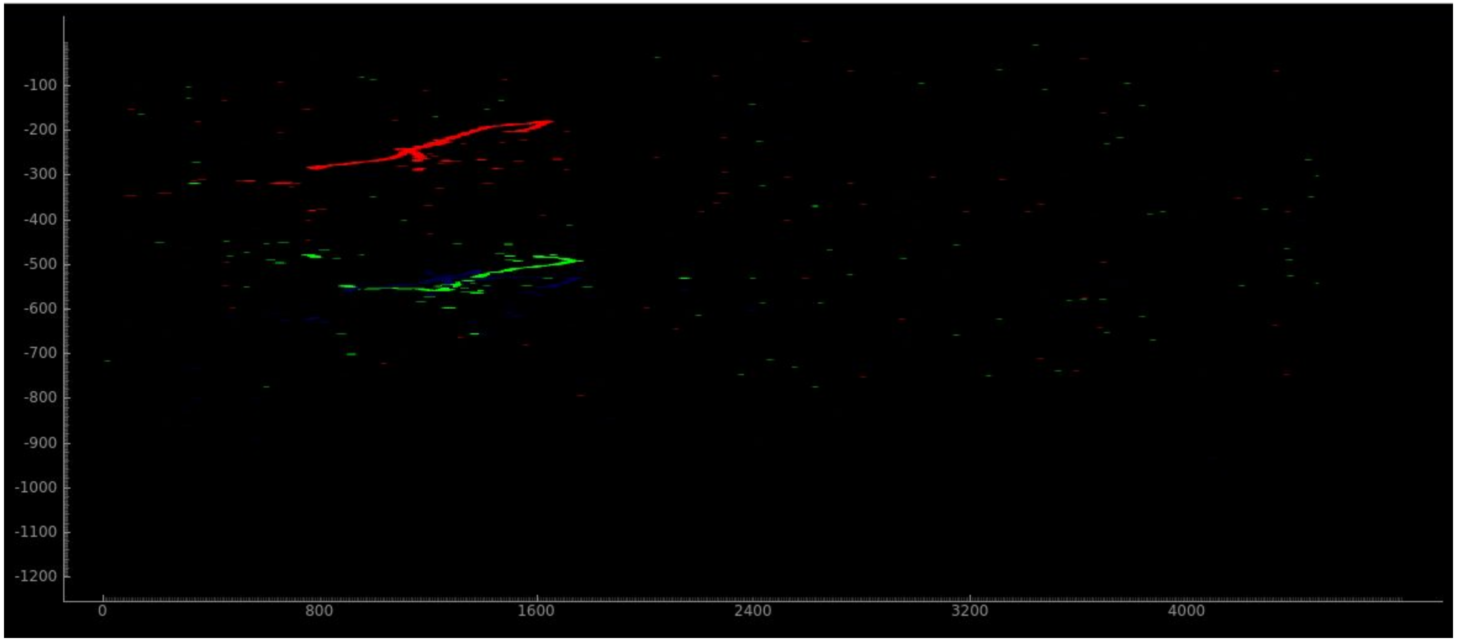
- Very likely candidate for misclassification of data vs MC
- Several ways to combat this
    - Refine noise simulation
    - Produce images from hit finding instead of wires
    - Generative adversarial networks (GANs)

Back to DUNE

- Explore CNN sensitivity to various simulation parameters
  - Noise levels, deconvolution, other detector effects

Raw waveform training:

- So far, all images have been created from **recob::Wire**, which is after deconvolution
- Anticipate minimal reduction in network performance when moving from deconvoluted to raw waveforms (with sufficient training)
  - Potentially interesting from perspective of online classification



An n-nbar event display in raw waveforms

Summary

- We can train a CNN to classify MC n-nbar and atmospheric neutrinos in DUNE and MicroBooNE

- Demonstrated that network learns on physics-motivated features

- Network heavily misclassifies MicroBooNE detector data
    - Looking into possible solutions, such as refining noise simulations, GANs

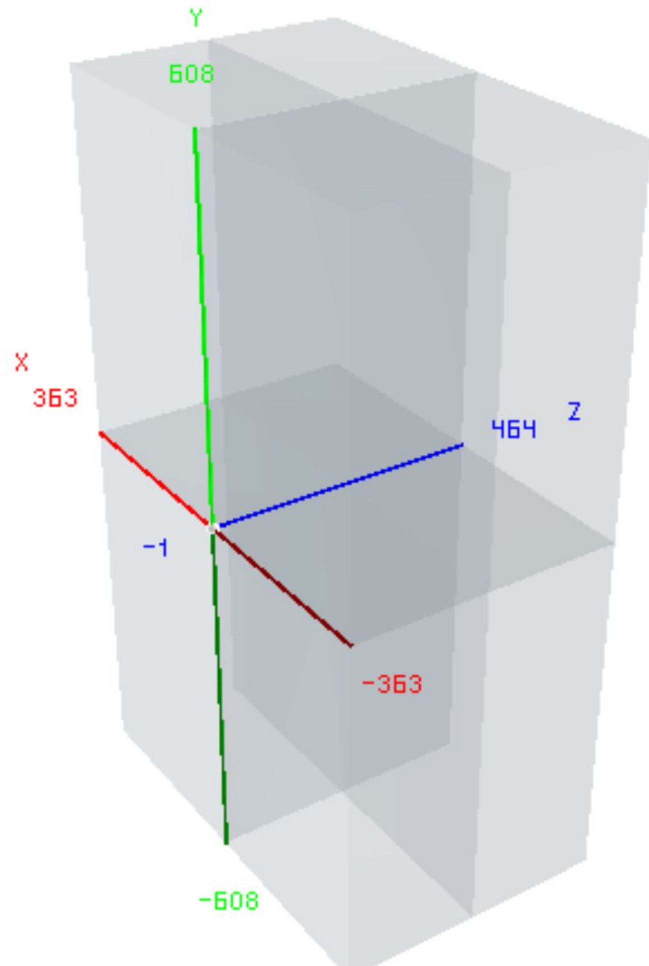
- Technote will be distributed (hopefully) shortly

Thank you!

Backup slides

APA geometry

dune10kt_v1_workspace



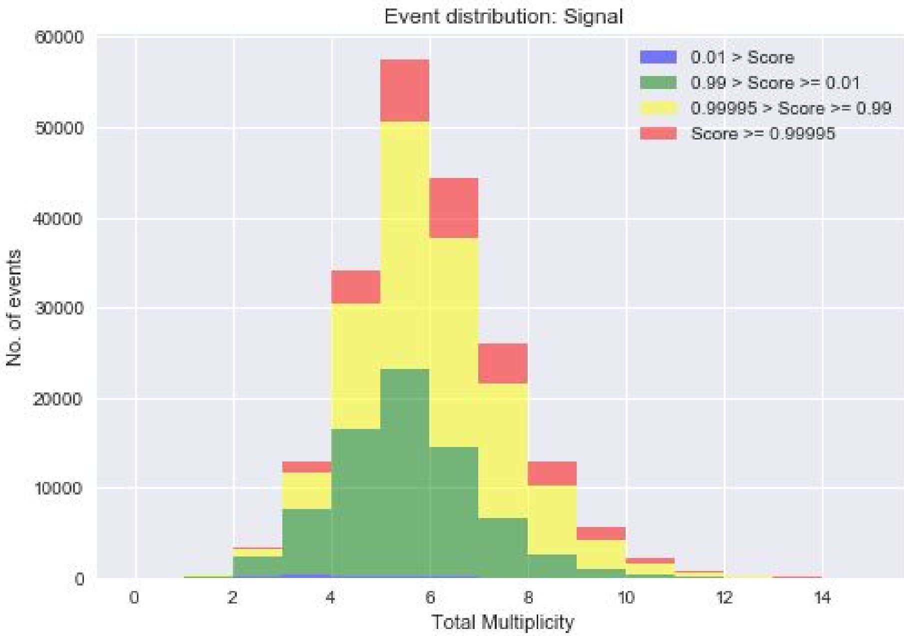
- 1x2x6 APA geometry (1x2x2 shown on the left)
- Two detector modules separated by X-axis
- Each drift region has its own collection plane located at the Y-axis, effectively making it another boundary

Previous results from past collaboration meetings:

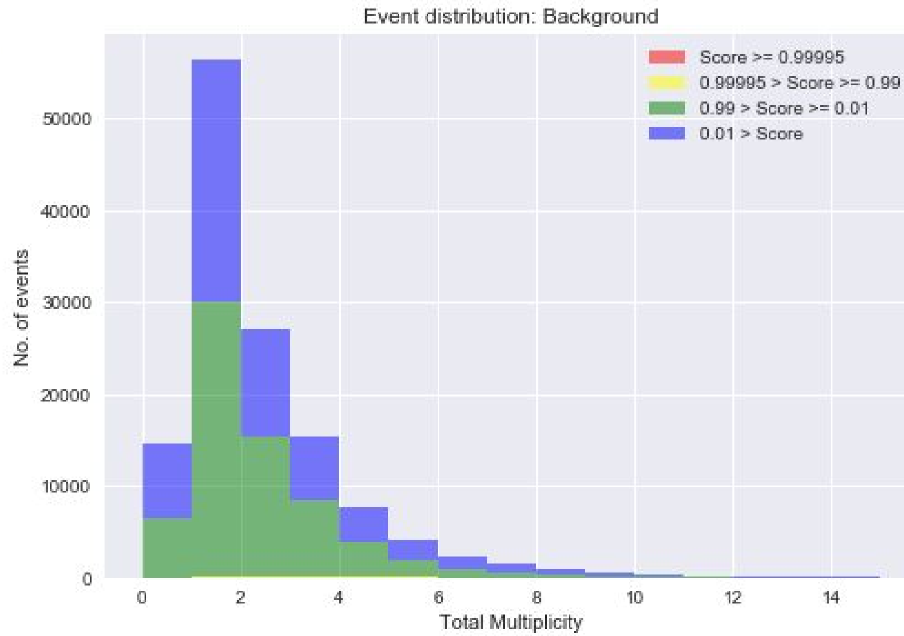
- A first pass on DUNE n-nbar sensitivity utilizing Deep Learning: https://indico.fnal.gov/getFile.py/access?contribId=137&sessionId=16&resId=0&materialId=slides&confId=12345
- CNN training results with oscillated atmospheric neutrino backgrounds: https://indico.fnal.gov/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=14836
    - Network feature studies with un-oscillated backgrounds
- Network inference results with oscillated atmospheric neutrino backgrounds: https://indico.fnal.gov/event/13293/session/20/contribution/102
    - Network feature studies with oscillated backgrounds

Total particle multiplicity (event dist.)

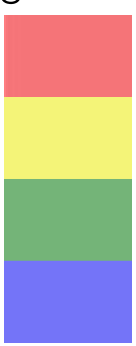
Signal (N-nbar) distribution



Background (atmos. v) distribution



Signal-like



Score >= 0.99995

0.99995 > Score >= 0.99

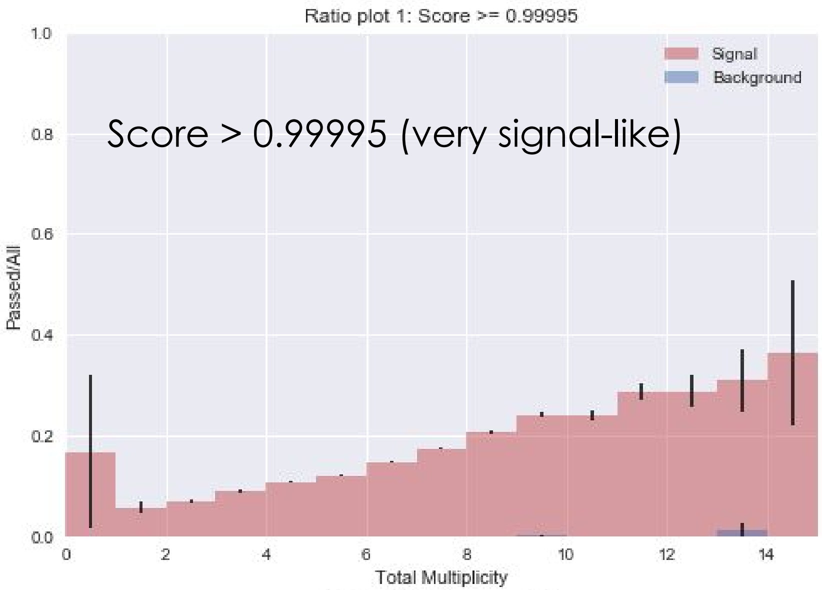
0.99 > Score >= 0.01

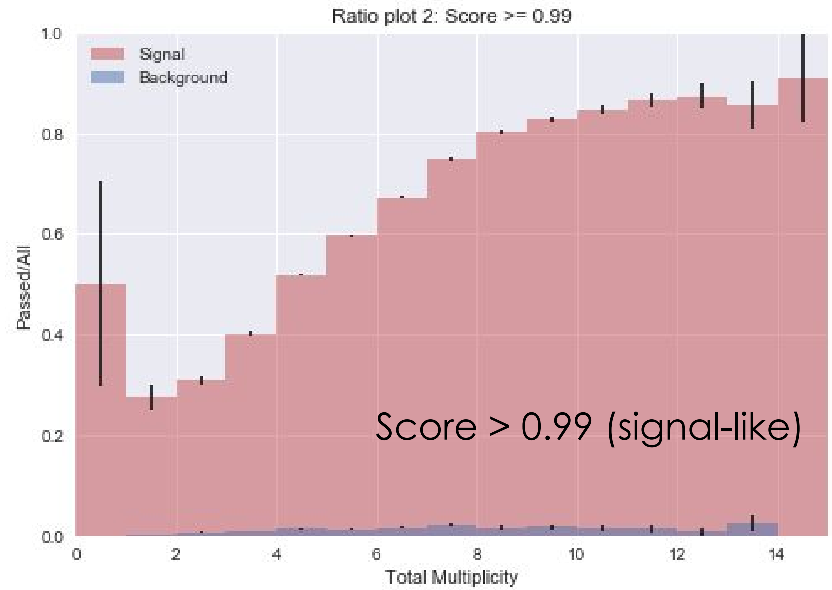
0.01 > Score

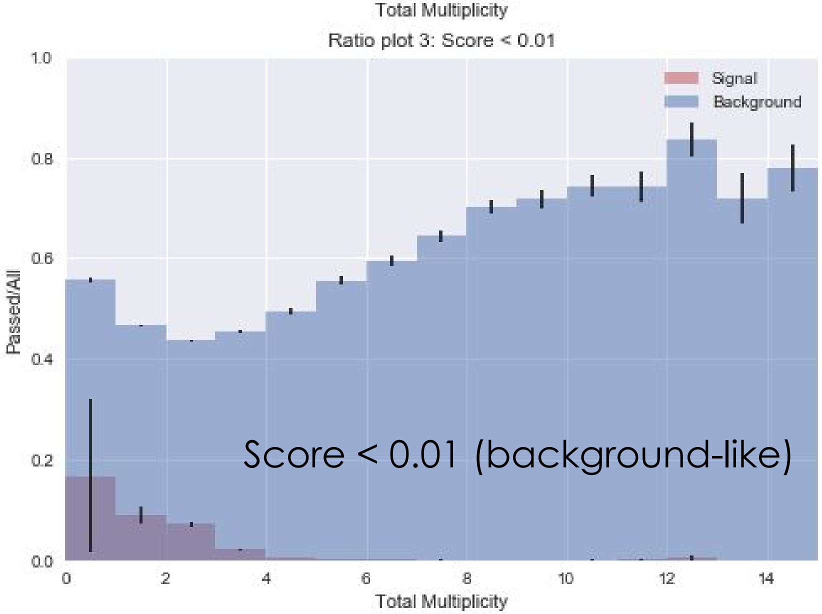
Background-like

- Includes pions, leptons, and protons
- N-nbar events tend to have slightly higher particle multiplicity on average

Total particle multiplicity (ratio)



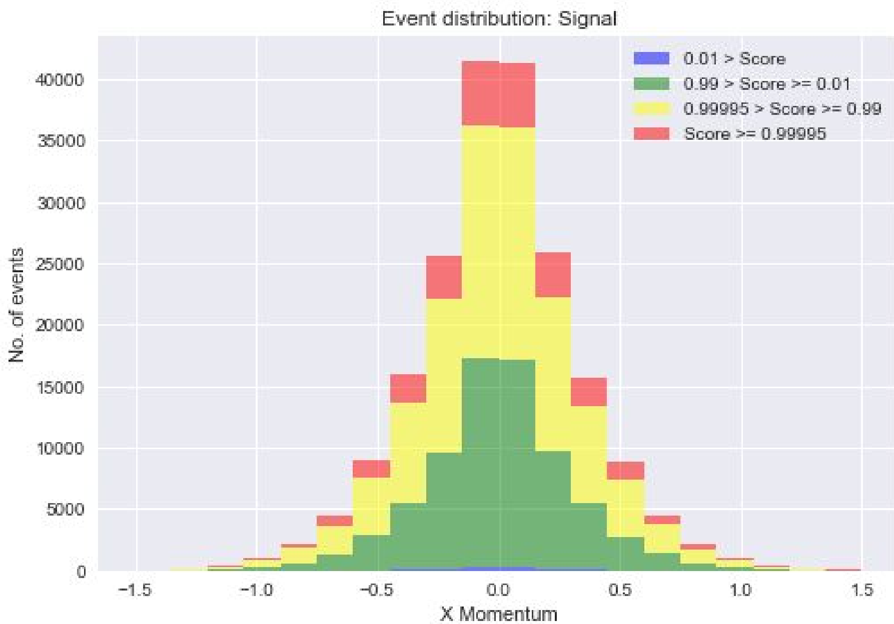




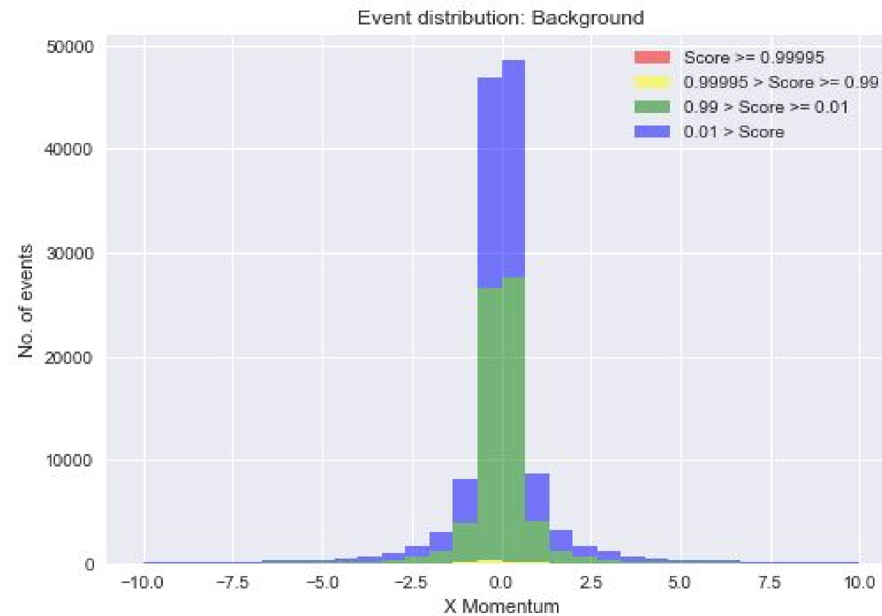
- Network does better at classifying high multiplicity events in general
- Smaller dip in background cut suggests particle multiplicity matters less in classifying background than signal

Momentum in X (event dist.)

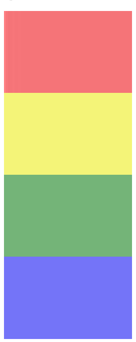
Signal (N-nbar) distribution



Background (atmos. ν) distribution



Signal-like



Score >= 0.99995

0.99995 > Score >= 0.99

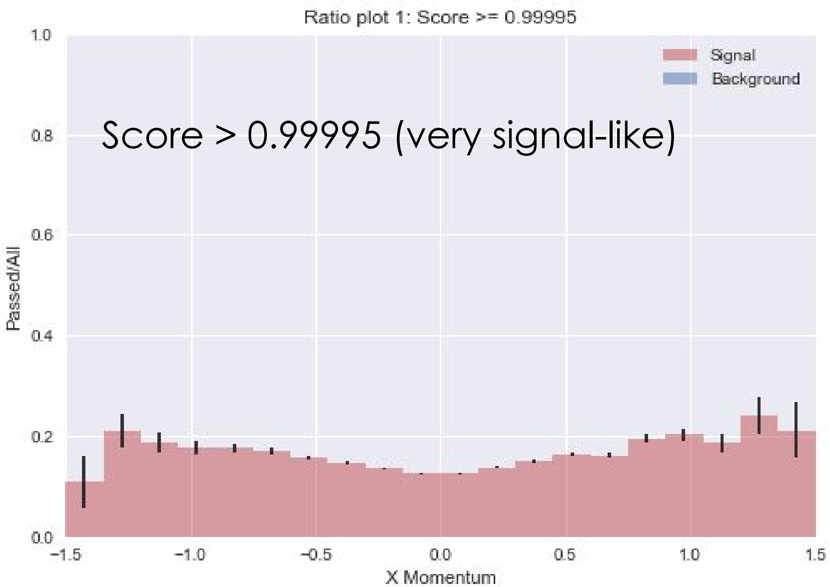
0.99 > Score >= 0.01

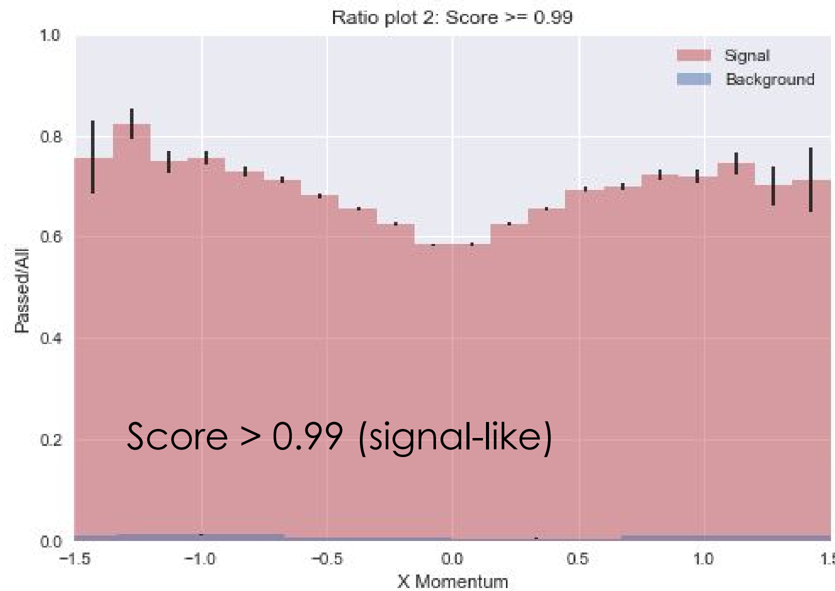
0.01 > Score

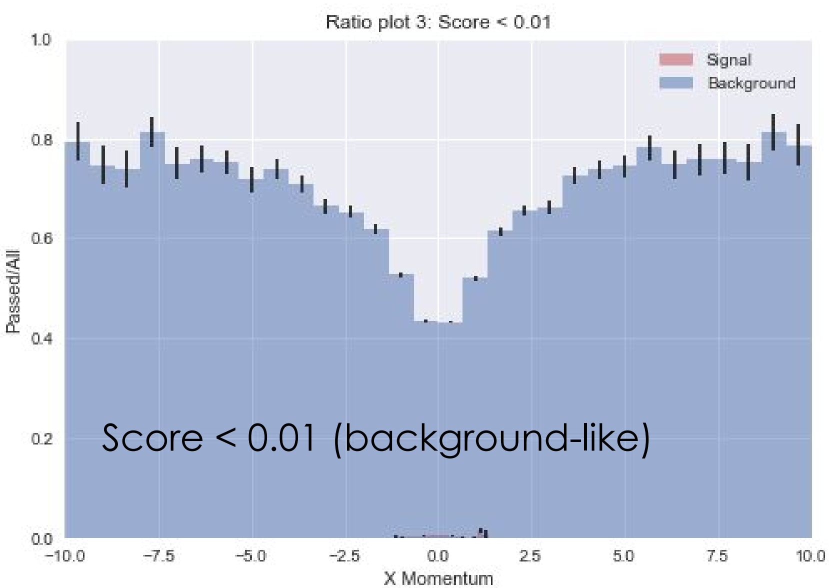
Background-like

- Signal distribution is more tightly huddled around 0 momentum

Momentum in X (ratio)







- Network has trouble classifying events with low component momentum
- Larger dip in background suggests network has more trouble classifying background than signal at low momenta