Development of Machine Learning Based Triggering for ICEBERG Liquid Argon Time Projection Chamber

Josiah Dubovi
Neutrinos & Rare Events Group, 2023 NSF REU at Nevis Laboratories at Columbia University

DUNE NSF

NEVIS LABORATORIES
COLUMBIA UNIVERSITY

PennWest UNIVERSITY

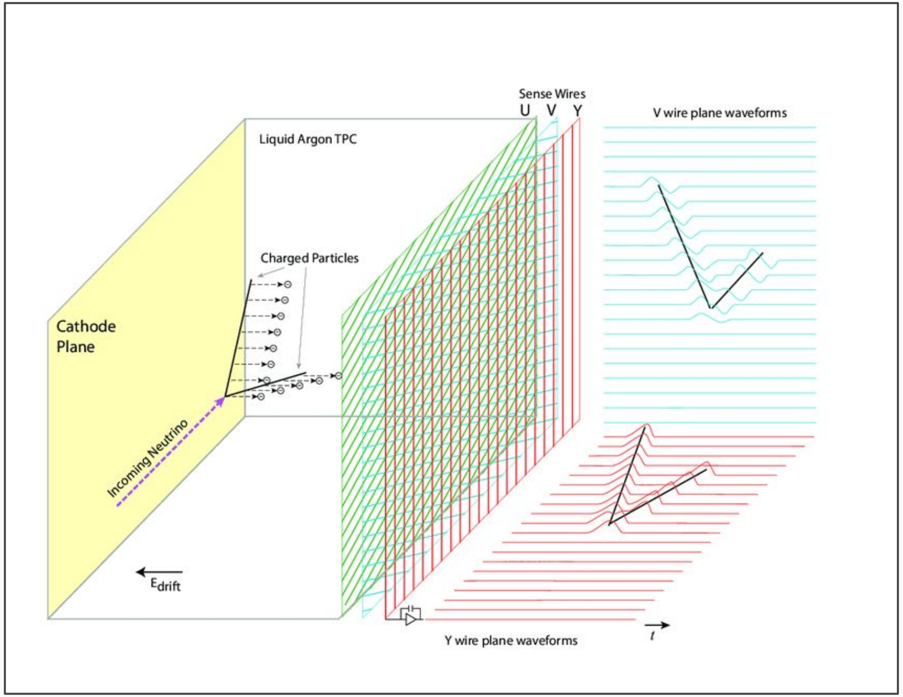
COLUMBIA UNIVERSITY

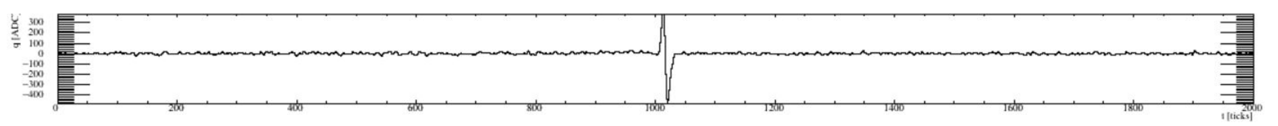
Table of Contents

Background & Motivation

Liquid Argon Time Projection Chambers (LArTPC)

- Detection of charged particles
  - Ionization Electrons
  - Scintillation Light
- Constructs 3D particle path using spatial (channel) information and timing information
- Components:
  - Liquid Argon
  - Electric Field
  - Induction Planes
  - Collection Plane
  - PMTs

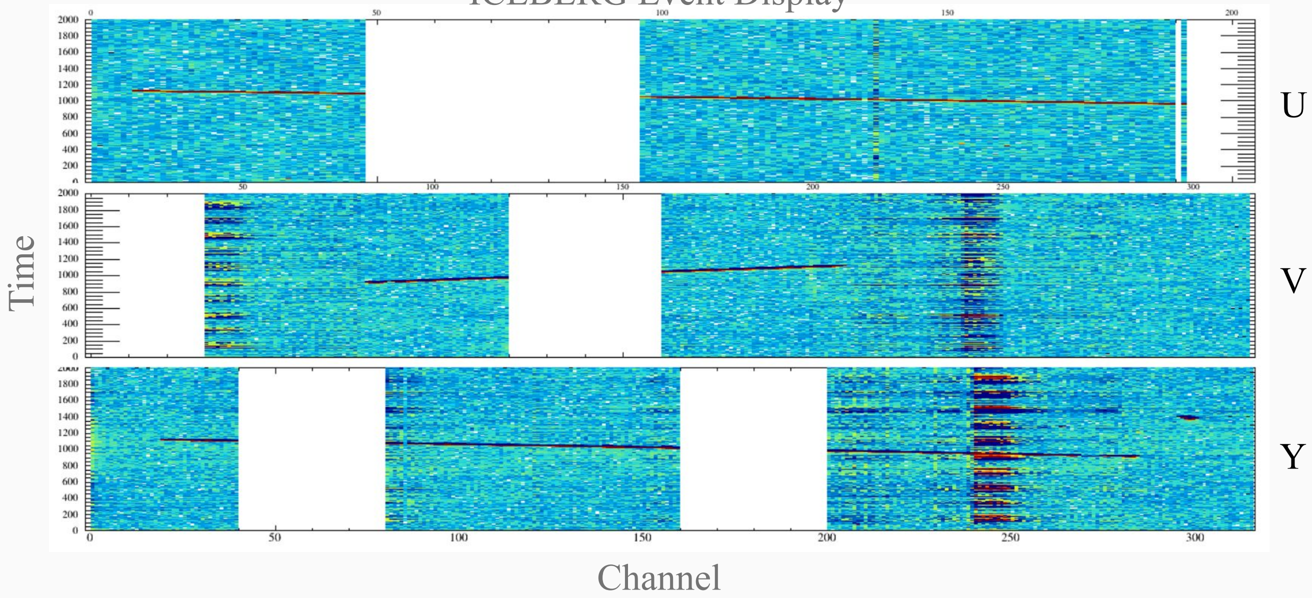




LArTPC Event Displays

- Signals are read as ADC waveforms on the wires
- A 2D event display can be constructed to visualize the particle's path:
    - Channel vs Time vs ADC count

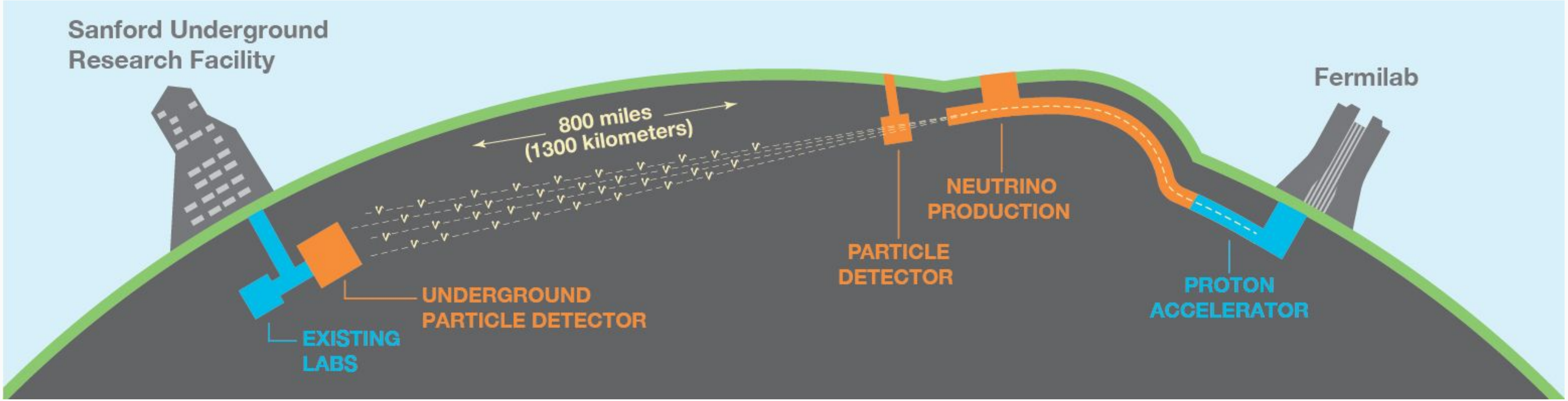
ICEBERG Event Display



COLUMBIA UNIVERSITY

Deep Underground Neutrino Experiment (DUNE)

- LArTPC deep underground to decrease background
- Major Objectives:
  - Neutrino Mass hierarchy
  - CP Violation in neutrinos
  - Observation of proton decay
  - Supernova neutrino detection



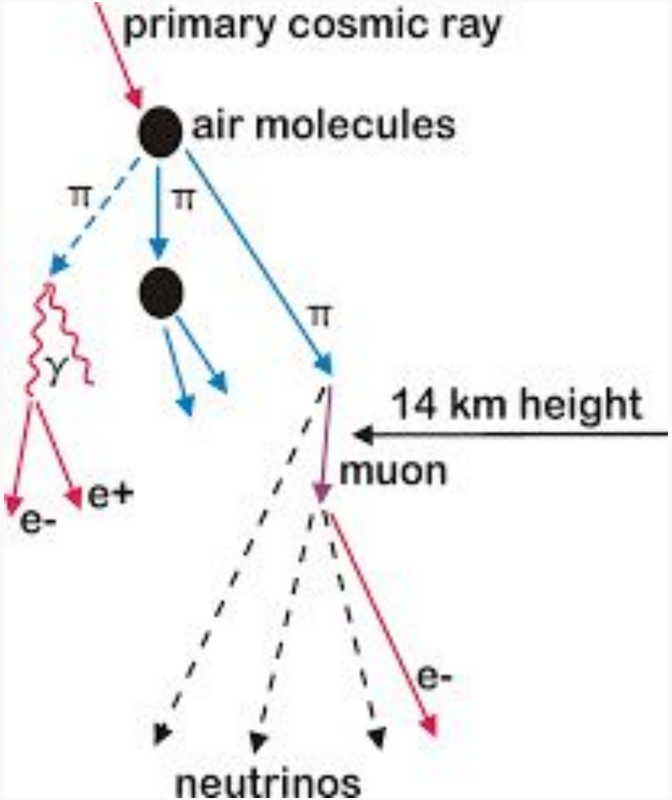
- DUNE is very large and very expensive
- Prototypes must be used to test the hardware/software
- ICEBERG
    - " Integrated Cryostat and Electronics Built for Experimental Research Goals"

- Contains ~4.5 tons of liquid argon
    - Dune will have a total of 70 kilotons
- Not underground!

Michel Decay

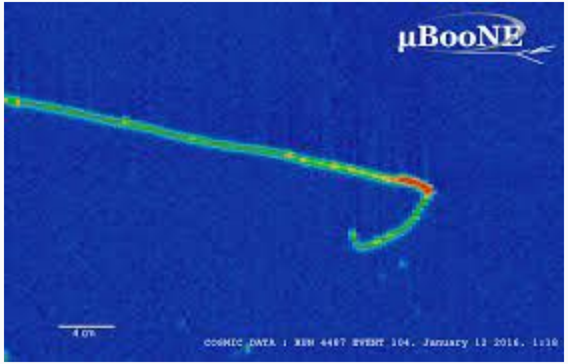
- Cosmic rays
  - Atomic nuclei traversing space at relativistic speeds
- Upon entering Earth's atmosphere, they decay into pions (mesons)
- Pions decay quickly into muons (heavy leptons)
- Muons most common mode of decay (Michel Decay) is:

$$\mu \rightarrow \nu_\mu + \overline{\nu_e} + e$$



Michel Decay *(continued)*

- Muons are unstable and decay within microseconds at rest
    - This is slower from our reference frame because they move at relativistic speeds
- DUNE will be strongly shielded from cosmic ray muons, but ICEBERG isn't
    - Michel decay is well understood and can be used for electronics calibration and testing software



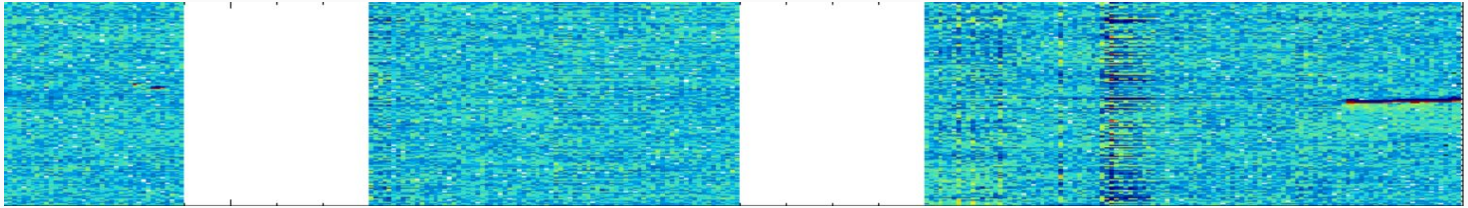
Event Triggering

- I am creating an algorithm for detecting Michel decay against background
- DUNE will generate several terabytes of data every second
  - Most of this cannot be stored
  - Must be reduced by an order of 10,000 to meet hardware limitations

An algorithm that decides what data to store must be developed:

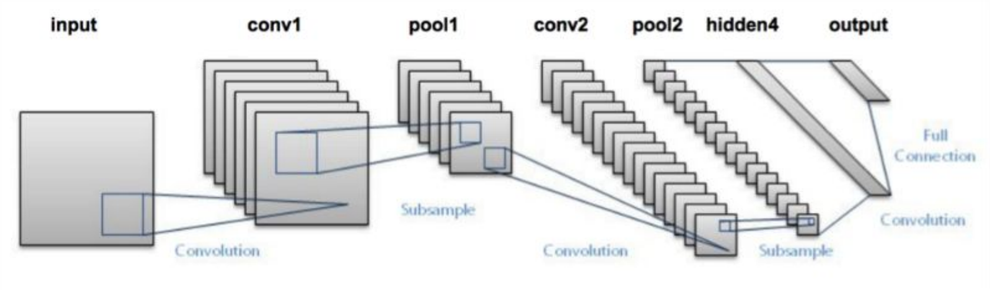
- Must be fast
  - DUNE will collect data in real time
- Must be accurate
  - DUNE aims to observe very rare events; we can't afford to miss them
  - ~99.5% of noise signals must be discarded
- Must not be highly computationally complex
  - Hardware resources are constrained

- Machine Learning has proven to be an effective method for event triggering
- Two-Dimensional Convolutional Neural Networks (2D CNNs) perform excellently in image classification
  - LArTPC event displays are simply images



2D CNN

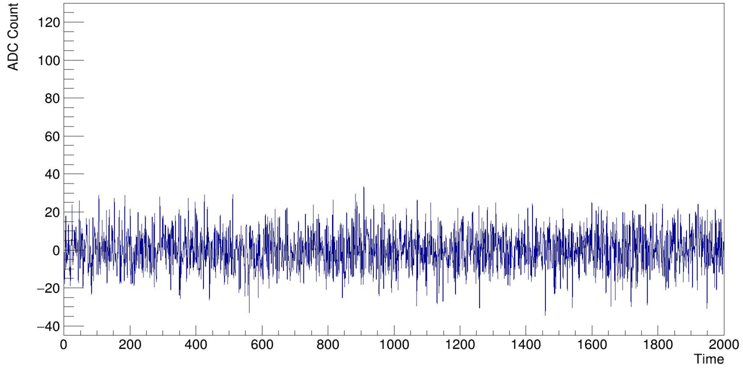
- Input layer
  - Tensor with dimensions: Length × Width × Color Channel
- Zero Padding Layer
  - Changes dimensions of tensor by adding entries with 0 value
  - May be necessary for operations in later layers
- Pooling Layer
  - Groups together nearby entries to reduce parameter count and computational load
- Convolutional Layer
  - Performs mathematical convolution between input tensor and a convolution matrix predetermined from training
- Fully Connected Layer
  - Every neuron applies a preset activation function to the input tensor



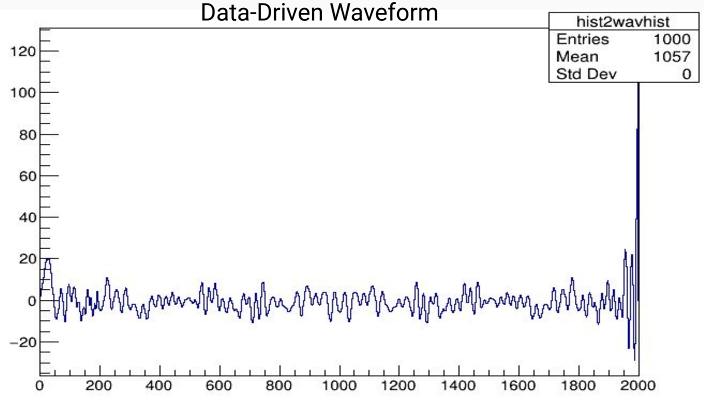
Noise Simulation for Neural Network Training

- A neural network must be trained on a very large, labelled dataset
  - This can effectively be produced via simulation
- Stopping muons can be simulated, but a data-driven noise model for ICEBERG does not yet exist
  - A method for generating noise must be generated using real noise data (because we can't simulate from first principles) to ensure accuracy to real events

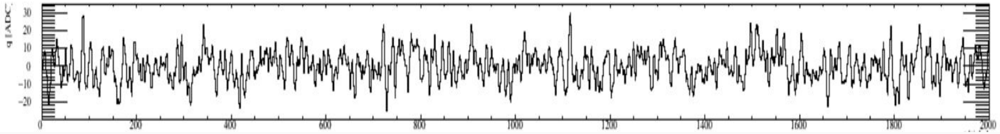
Randomized Waveform



Data-Driven Waveform



Real Waveform



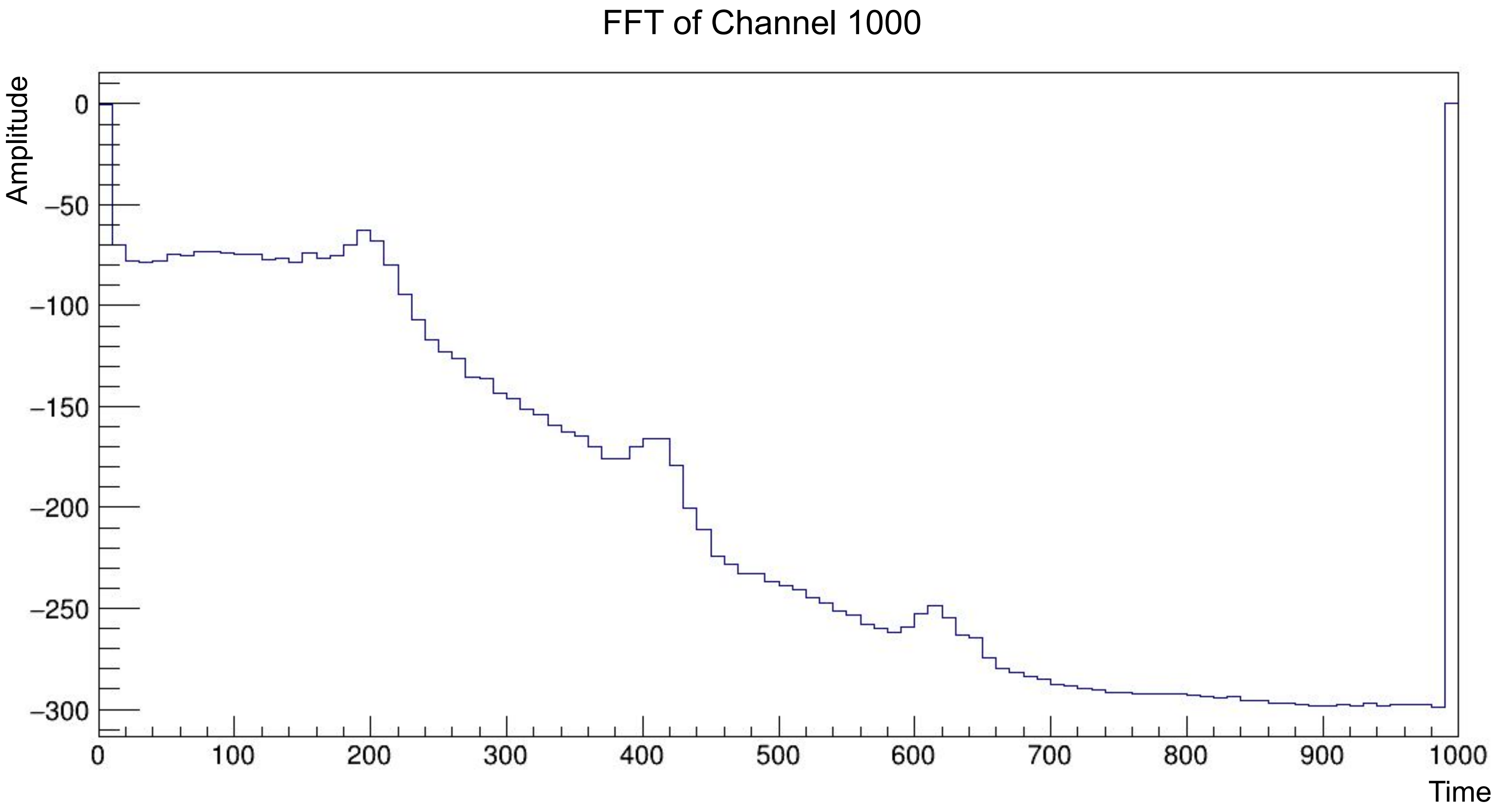
Data-Driven Noise Model

COLUMBIA UNIVERSITY

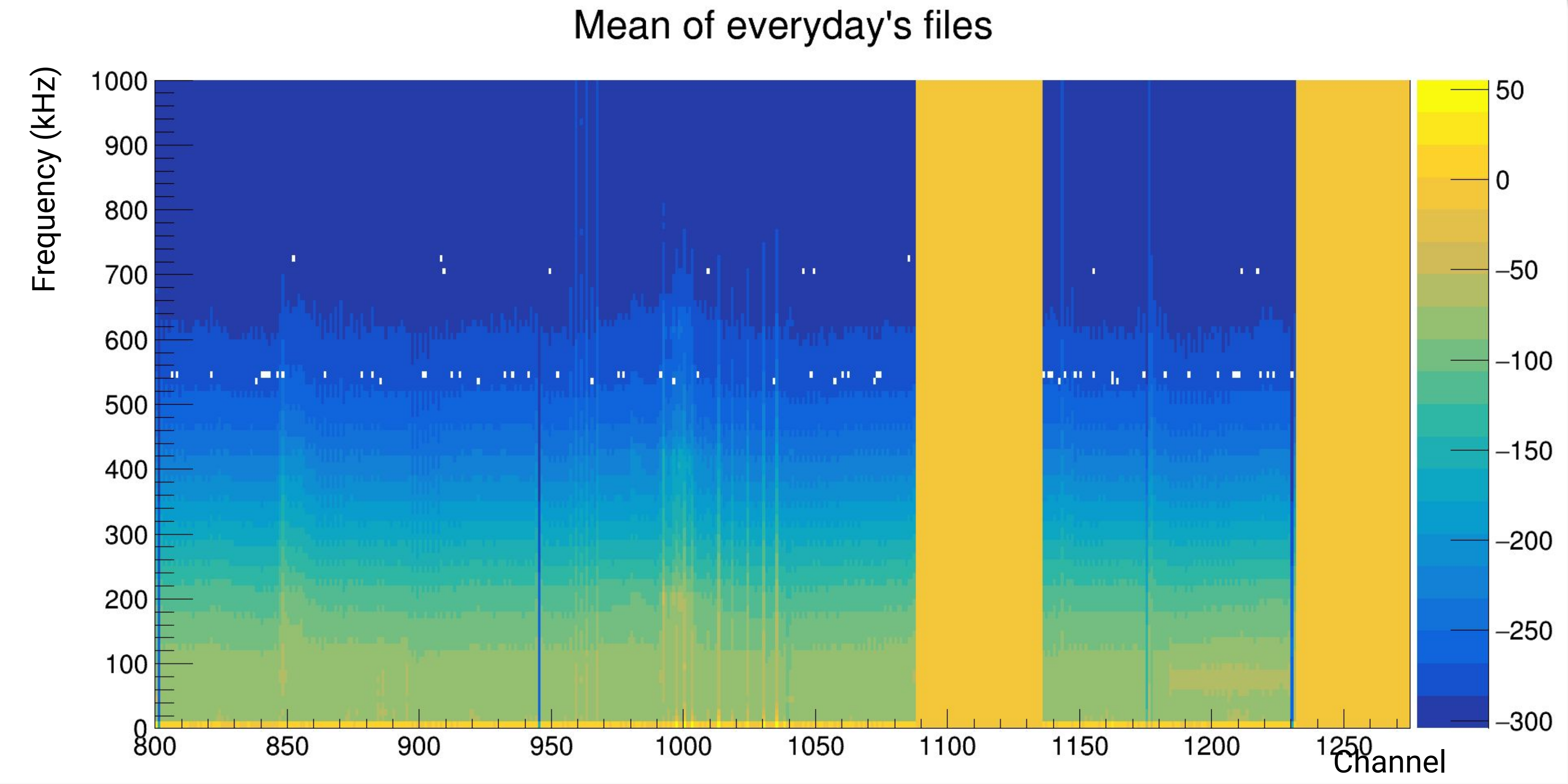
Fast Fourier Transform

- A Fast Fourier Transform (FFT) can be used to calculate the frequencies and associated amplitudes of a given waveform
    - This information is critical for generating a simulated waveform
- Each channel experiences noise slightly different
    - Different lengths and positions within the detector
    - We must therefore analyze the frequencies associated with each channel independent of the others
- Noise varies over time
    - Files containing data collected between 3/13/21 and 4/4/21 were analyzed to avoid overfitting to a single time
    - FFT's for each day were compared to look for trends relating to time

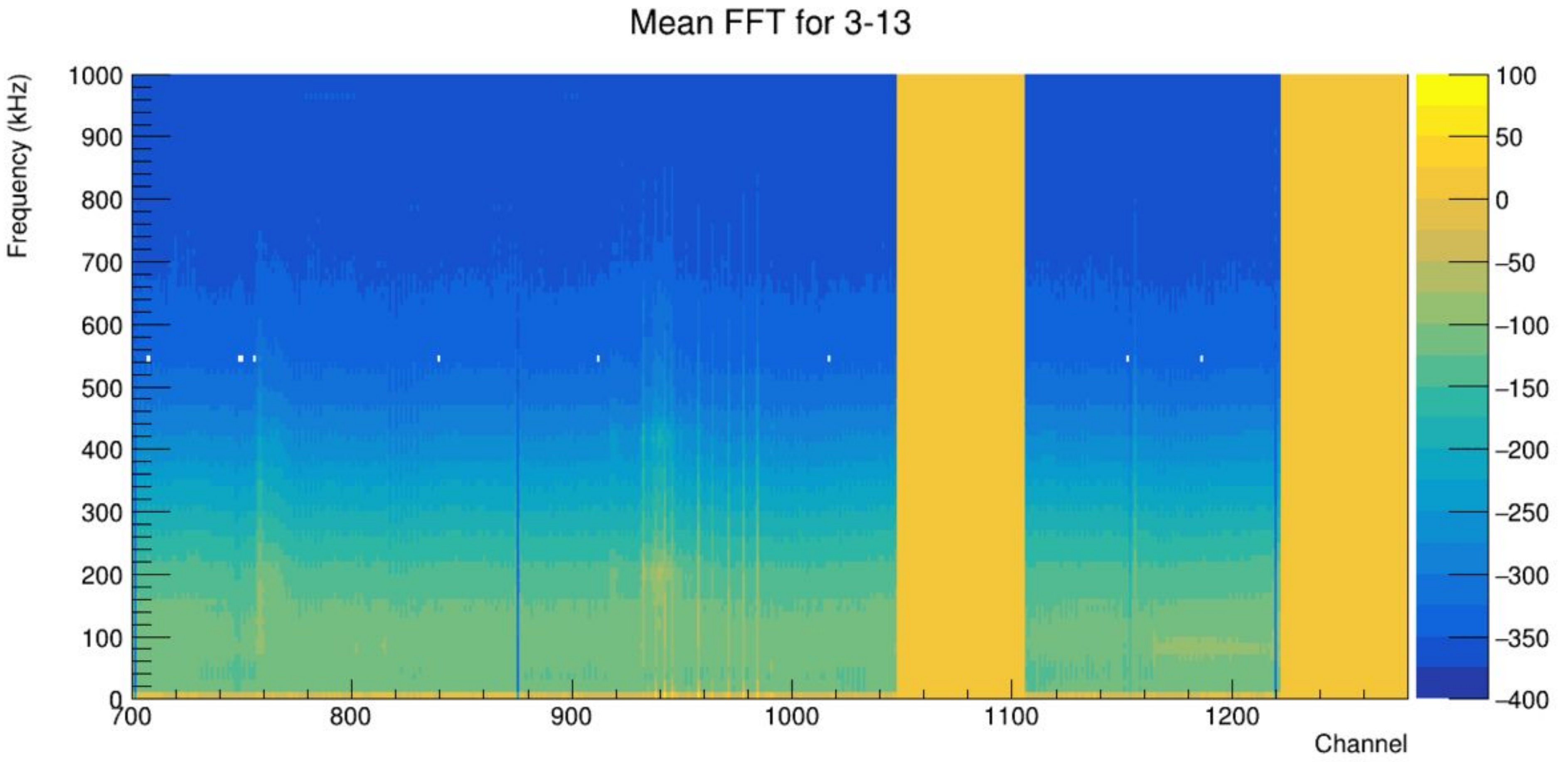
FFT spectrum for one channel, averaged over every analyzed event:



FFT spectra for all channels, averaged over every analyzed event:

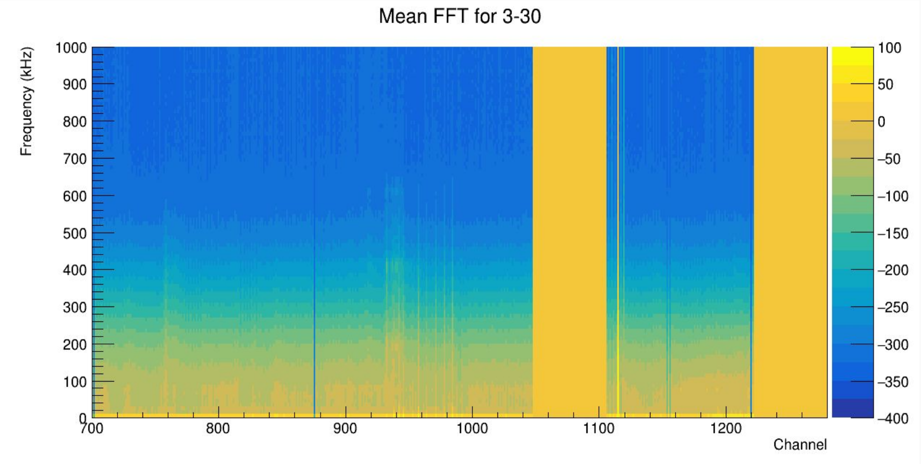


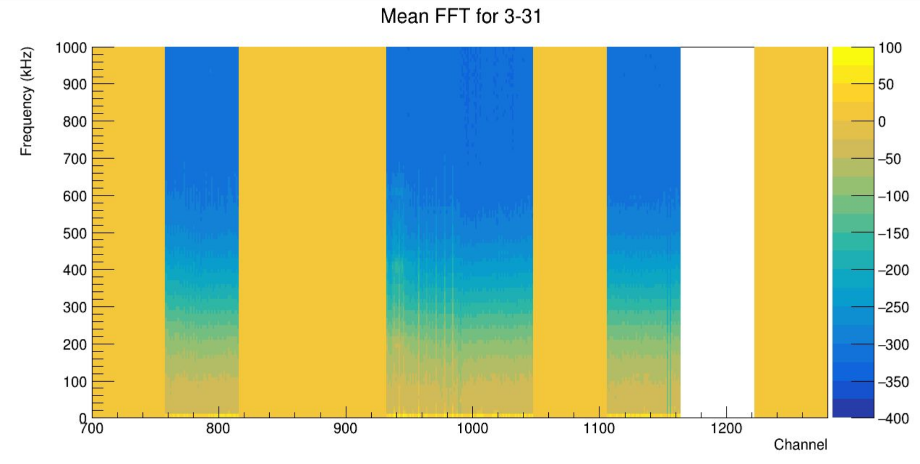
Twenty-Two Days of ICEBERG Data:



FFT Spectra

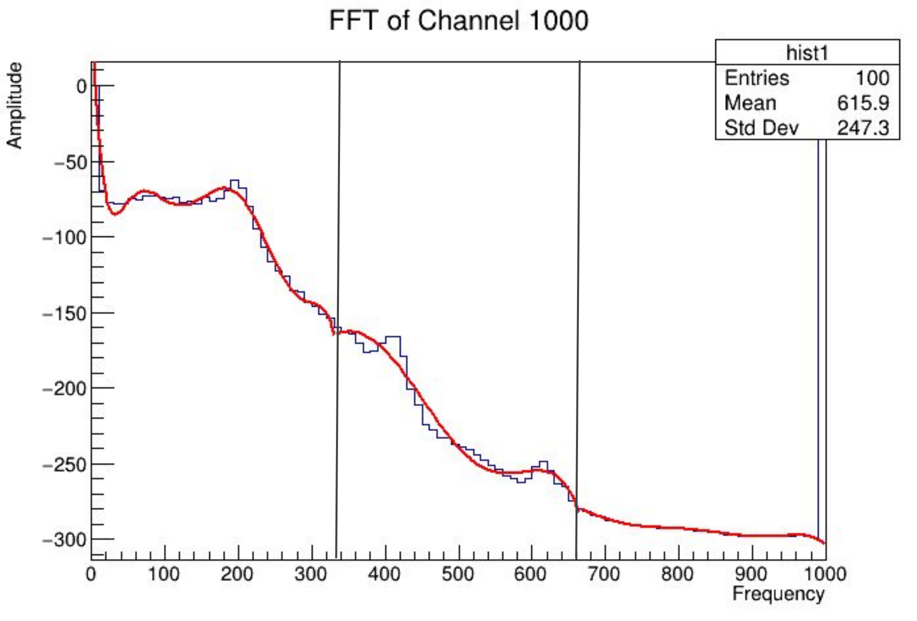
- No trends were noticeable as time passes, but variance is still present
- The same channels tend to consistently be the noisiest
- Some days have different Data Acquisition (DAQ) configurations, causing them to have different dead channels



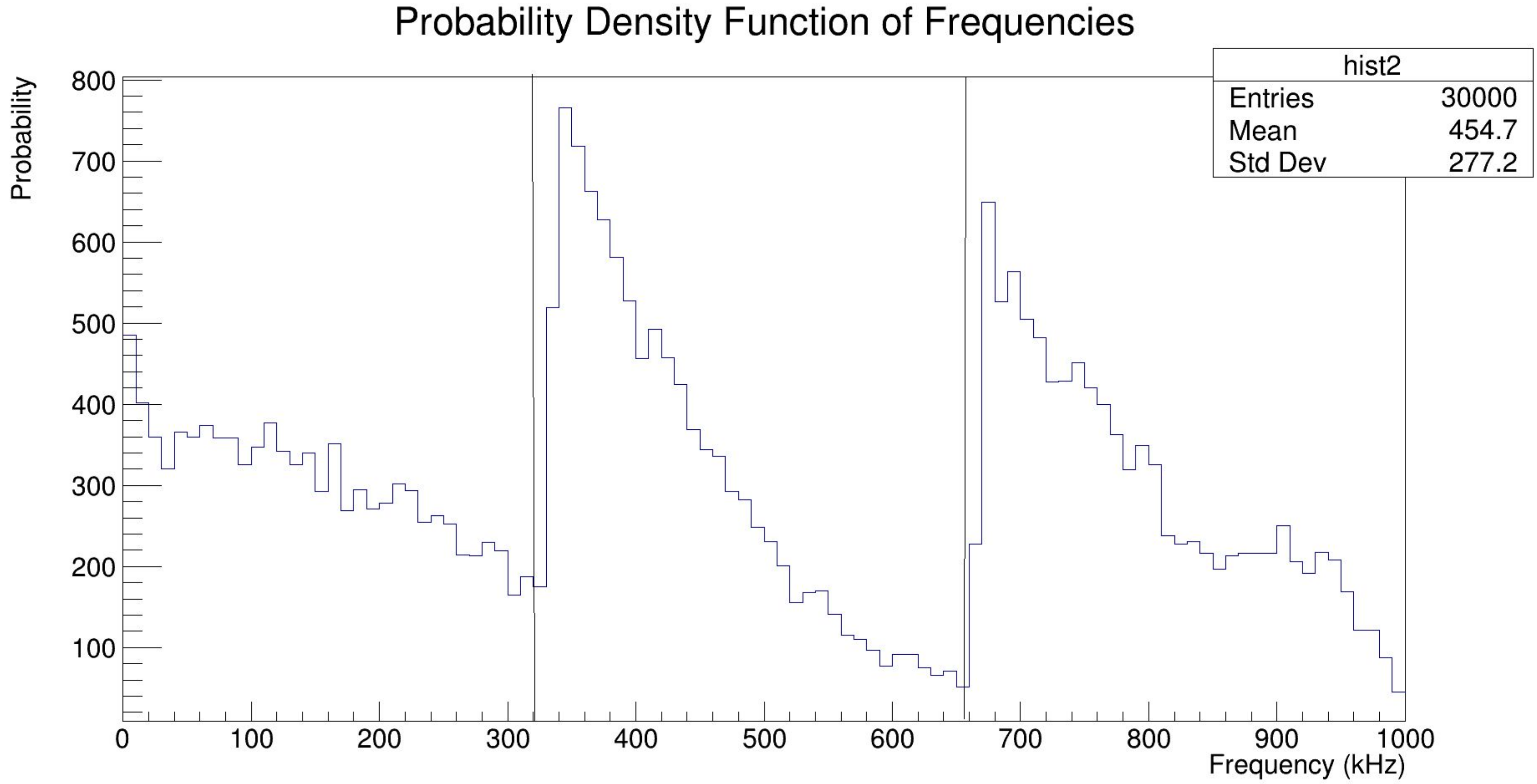


Columbia University

- The average FFT is divided into three regions: low frequency, middle frequency, and high frequency
- A function is fitted to each region using polynomial interpolation
    - This function behaves as a probability density function

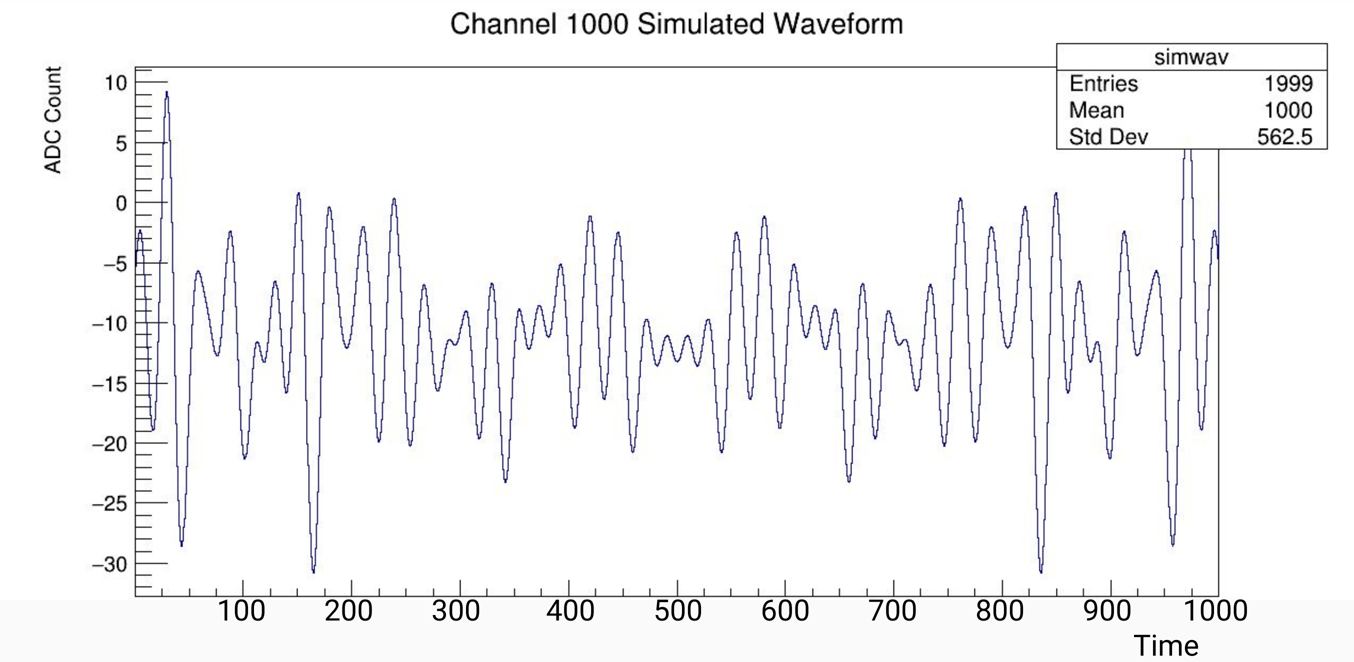


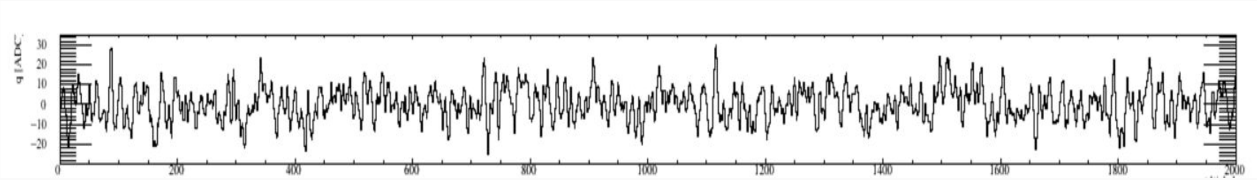
- One frequency is randomly selected from each region based on the probability density function
- An inverse FFT is performed to generate a waveform with these three frequencies


Probability Density Function of Frequencies
hist2
Entries 30000
Mean 454.7
Std Dev 277.2
Probability
800
700
600
500
400
300
200
100
0
100
200
300
400
500
600
700
800
900
1000
Frequency (kHz)

Single Simulated Wave

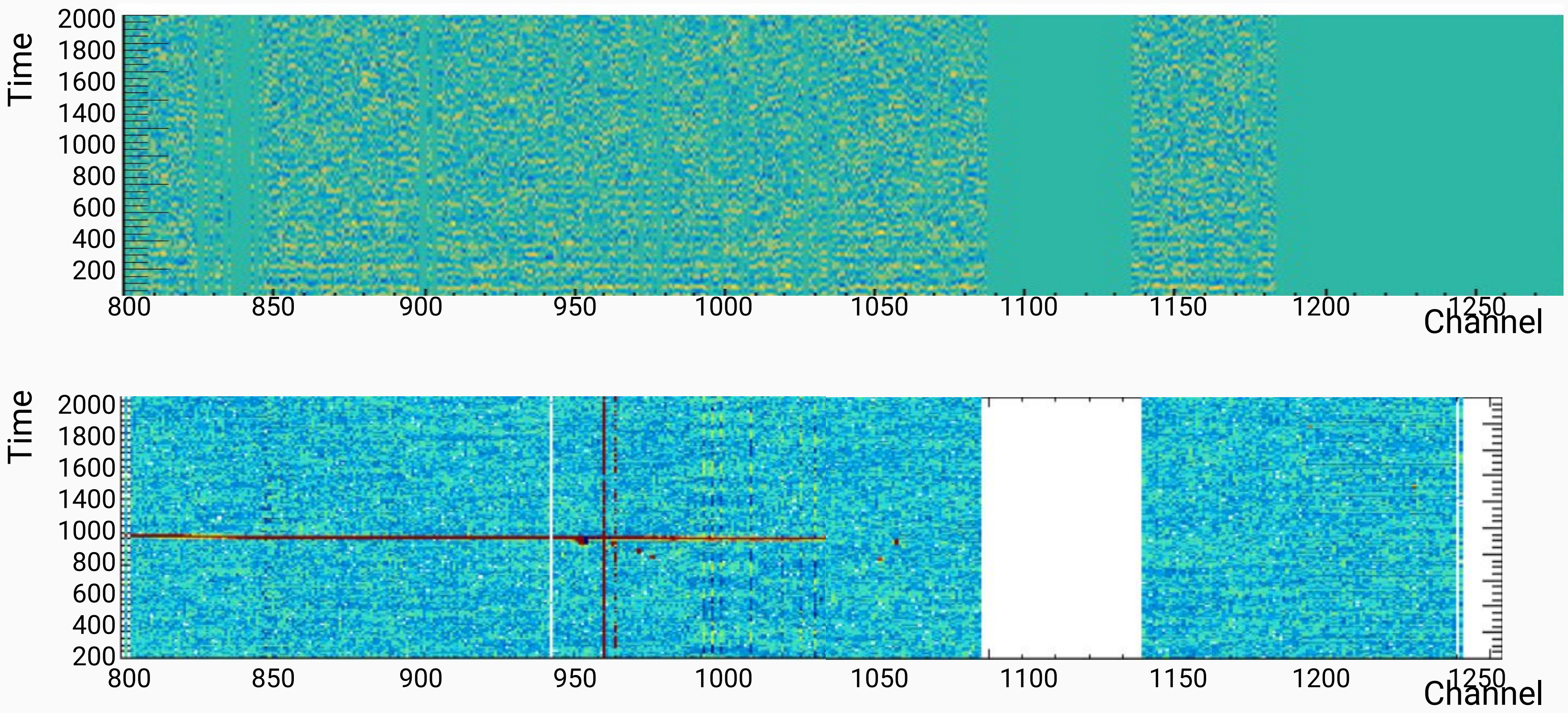
Low: 70 kHz
Mid: 430 kHz
High: 700 kHz





Columbia University

Simulated Event Display (Collection Plane)



Columbia University

Convolutional Neural Network

COLUMBIA UNIVERSITY

Architecture

- Two Dimensional Convolutional Neural Network (2D CNN)
- Based on a network previously developed to detect low energy neutrino events in DUNE
  - Low parameter count, and high accuracy

Layer (type)	Output Shape	Param #
resizing_28 (Resizing)	(None, 64, 480, 3)	0
zero_padding2d_73 (ZeroPadding2D)	(None, 64, 512, 3)	0
max_pooling2d_92 (MaxPooling2D)	(None, 32, 64, 3)	0
zero_padding2d_74 (ZeroPadding2D)	(None, 34, 66, 3)	0
conv2d_70 (Conv2D)	(None, 32, 64, 2)	56
re_lu_60 (ReLU)	(None, 32, 64, 2)	0
max_pooling2d_93 (MaxPooling2D)	(None, 16, 16, 2)	0
zero_padding2d_75 (ZeroPadding2D)	(None, 18, 18, 2)	0
conv2d_71 (Conv2D)	(None, 16, 16, 2)	38
re_lu_61 (ReLU)	(None, 16, 16, 2)	0
max_pooling2d_94 (MaxPooling2D)	(None, 4, 4, 2)	0
reshape_13 (Reshape)	(None, 1, 1, 32)	0
dense_46 (Dense)	(None, 1, 1, 12)	396
re_lu_62 (ReLU)	(None, 1, 1, 12)	0
dense_47 (Dense)	(None, 1, 1, 3)	39

Total params: 529
Trainable params: 529
Non-trainable params: 0

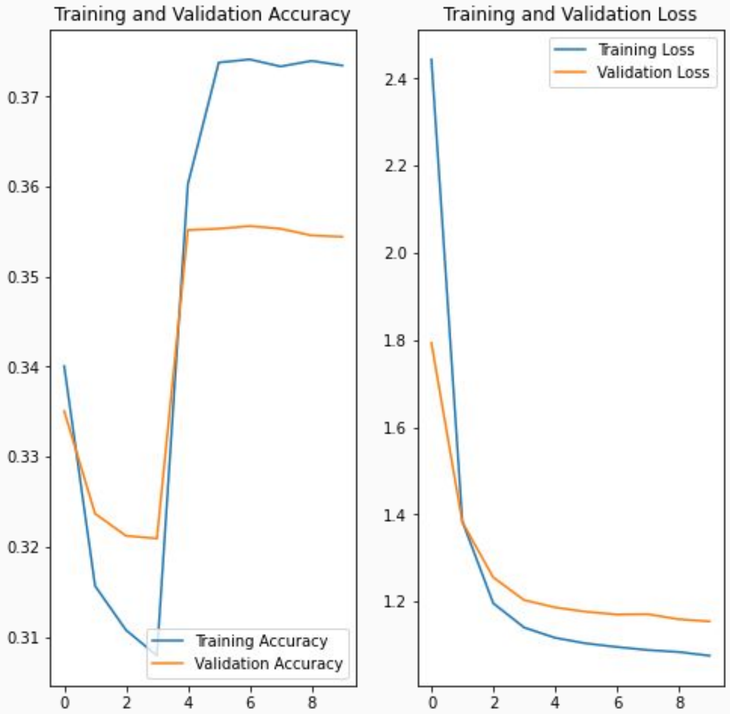
Training and Performance

- Training set contains simulated stopping muons and through-going muons
    - These simulations are only recently accessible, and further formatting is necessary to prepare them for training
- Network was tested on publicly accessible neural network datasets
    - Performance was poor because architecture was not optimized for this data

tulips

sunflowers

roses



Summary & Conclusions

COLUMBIA UNIVERSITY

Summary & Conclusion

- A data-driven noise model has been developed for ICEBERG
  - A similar method will likely be used for DUNE
- A 2D CNN is promising for Michel Decay triggering, but it has not yet been trained and tested on the simulated data

Next Steps

- Improve accuracy of noise simulation
  - Truncate edge effects
  - Use larger dataset
- Improve speed of noise simulation
  - A simulated event display currently takes ~30 minutes to generate
- Train/optimize the Neural Network
  - Modify architecture to find ideal balance between complexity and performance
  - Compare if using all three detection planes improves performance
- Use CNN for detection of other interactions
  - Ar-39

Acknowledgements

This project is funded by NSF Grant No. PHY/1950431

Additional thanks to the Neutrinos and Rare Events group at Nevis Laboratories and the ICEBERG collaboration at Fermilab