# Deep Learning for Classification and Denoising of Cosmic-Ray Radio Signals

Abdul Rehman, Dana Kullgren, Alan Coleman, Frank Schroeder

Workshop on Machine Learning for Cosmic-Ray Air Showers, Delaware USA.

UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

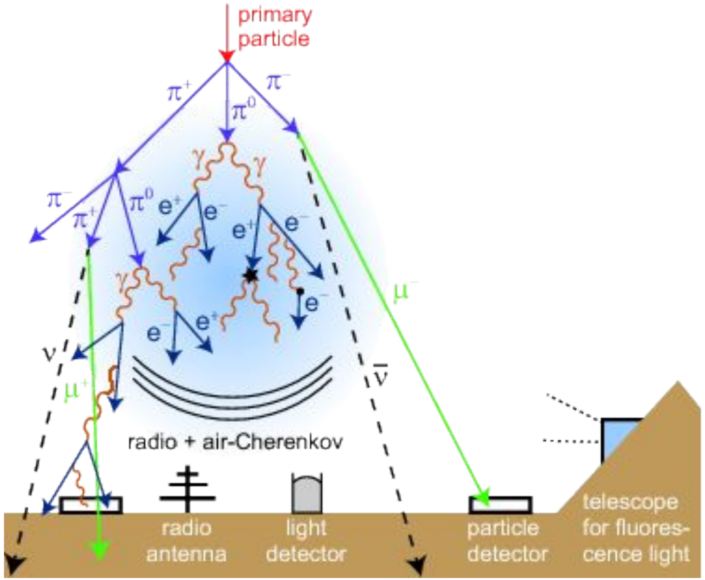
IceCube

# Outline

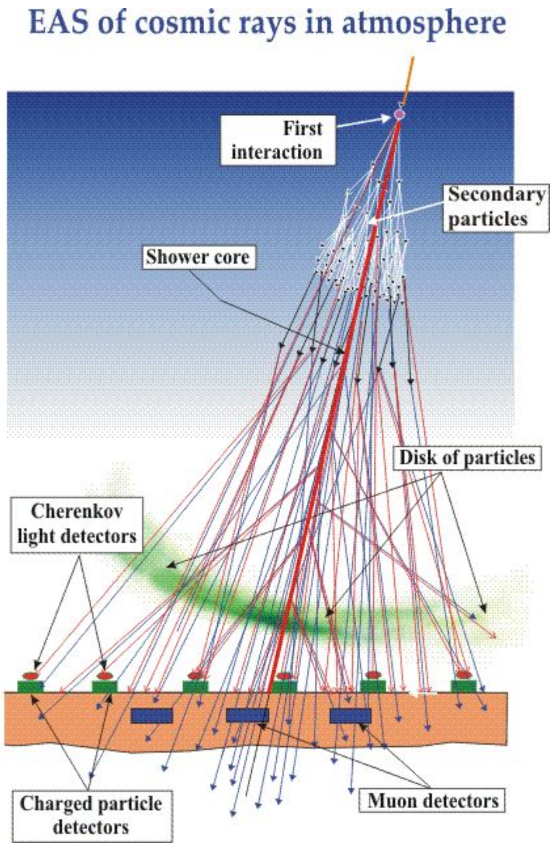
- ❑ Introduction:
  - ❑ CR Air showers.
  - ❑ Radio emission from air showers.
  - ❑ Use of Deep Learning for Classification and Denoising.
- ❑ Preparation of DataSet for model training.
- ❑ CNN Network architecture
- ❑ Results of trained models.
- ❑ Other network that we have tried.
- ❑ Work in progress and Outlook.

# Cosmic-Ray Air Showers

- Primary CRs produce Air Showers when they enter the atmosphere.
- Number of secondary particles first grows, then reach a maximum at X__max and then declines as the shower dies out.

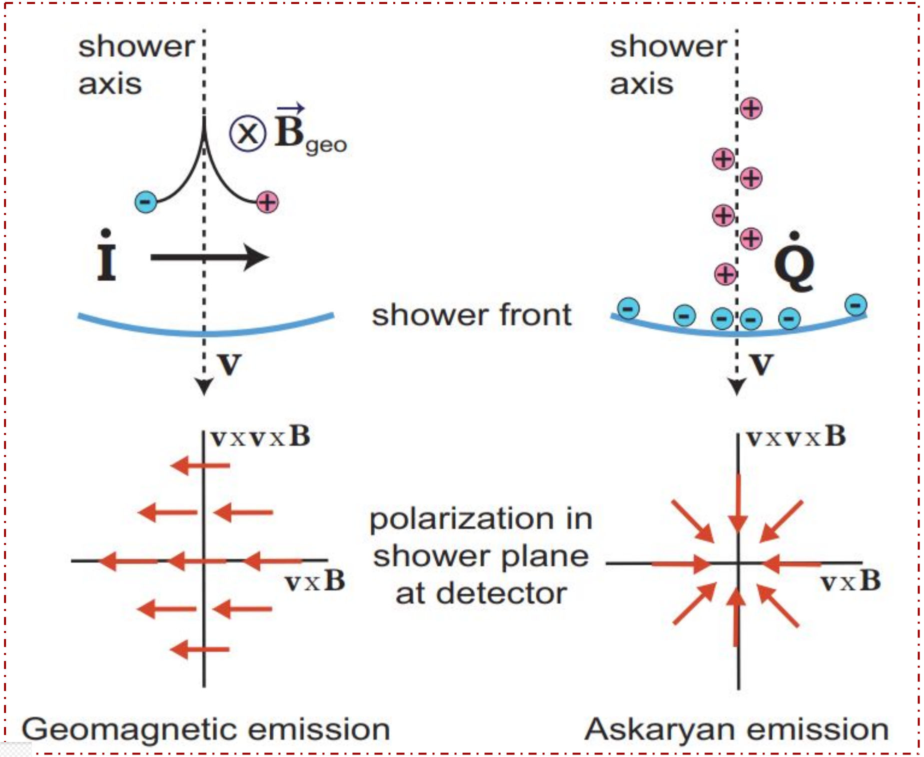



F.G. Schroder, Prog. Part. Nucl. Phys. 93 (2017), 1,: arXiv: 1607.08781




ikfia.ysn.ru/

# Radio Emission from Air Showers




- **Geomagnetic emission**: Due to time-variation of the transverse current

- **Askaryan emission**: Due to time-varying charge excess in the shower front

Radio competes in precision of X_max, arrival direction and energy, with other established techniques like fluorescence detector.

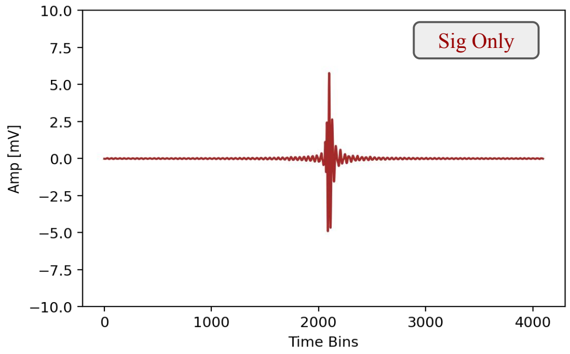
100% duty cycle.

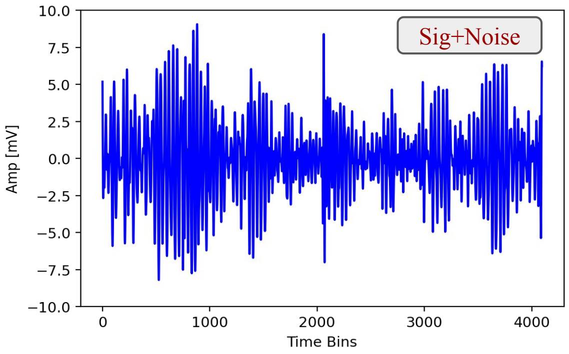
Combining radio with particle detection will significantly increase the total accuracy of Air shower measurements.

F.G. Schroder, Prog. Part. Nucl. Phys. 93 (2017), 1,: arXiv: 1607.08781

# Introduction

- Radio detection, like other techniques, has to deal with the continuous background.
- Because of the noise contaminating the radio signals, radio detection threshold is very high.
- We are using ML (Convolutional Neural Networks) to try to mitigate the effects of background.
- To train networks we used simulated radio signals and background.
- Example Traces are shown in the bottom plots.



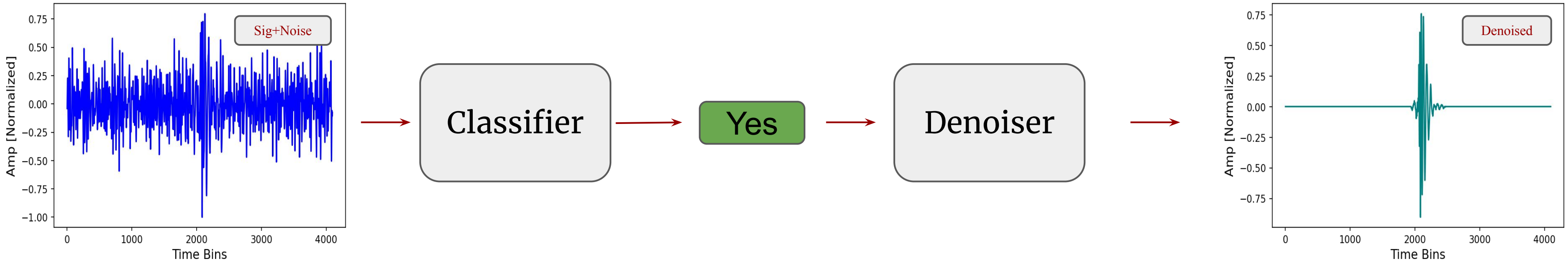




UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

* Previous work by Tunka-Rex Collaboration: arXiv:2101.02943, arXiv:1812.03347.

# Classifier and Denoiser

- Using **Keras/Tensorflow** to construct 1D Convolutional Neural Networks.
- **Classifier:** Identify radio signals and backgrounds.
- **Denoiser:** Recover the underlying signals from the Noisy traces.




UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Data Preparation

➢ **Dataset:**

- CORSIKA → Simulation of air-showers.

  ( IceTop atmosphere )
- CoREAS → Radio emission from air-showers.

  ( Sibyll 2.3d as a high-energy hadronic interaction model. )
- Zenith angles → [0, 65]deg in steps of 5 deg, random azimuth angles.
- For background: Modeled (Cane) Noise → Average Galactic + Thermal background

  ( Dana kullgren will show results using measured background from prototype station at the South Pole)
- After adding antenna and electronic responses the signal and background are combined to produces Noisy traces.
- Filtered band [50-350] MHz.
- Traces are also normalized before inputting them to the network.

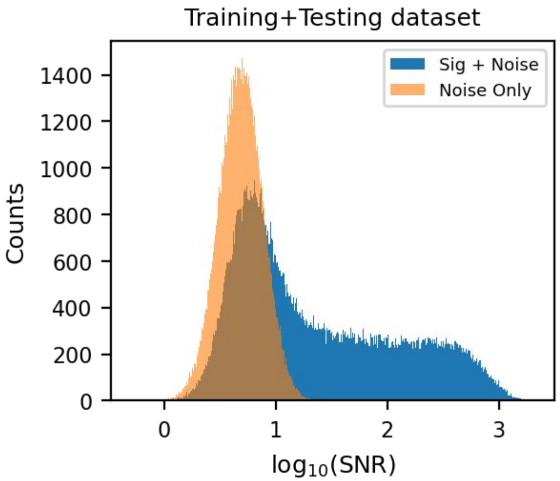
UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Training and Testing Sets

- Signal to Noise ratio (SNR) is used to quantify the signal strength.
- Signals are scaled before adding noise.
- 103k Signal + 135k noise traces.
- 80% for training, 20% for testing.

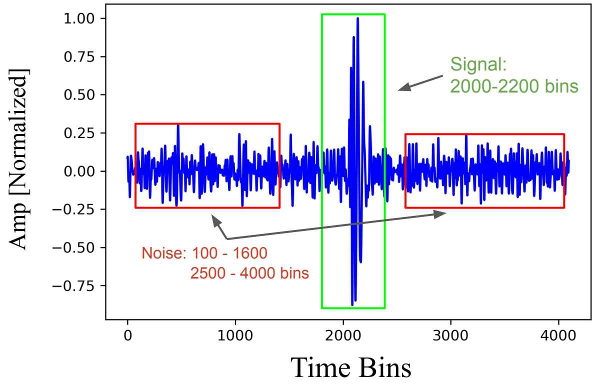
$$\mathrm{SNR} = \left(\frac{\mathrm{Signal}_{\mathrm{Peak}}}{\mathrm{Noise}_{\mathrm{RMS}}}\right)^2$$

❑ Mean SNR of Noise ~ 7




SNR distribution of data set

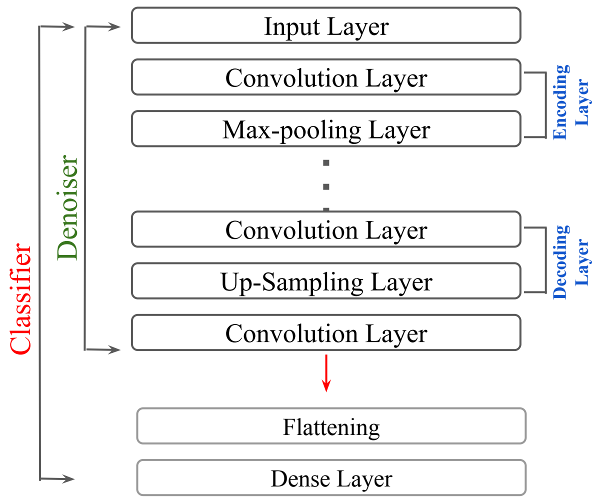
Trace length = 1000 ns




UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Network Architecture

- Networks are based on Autoencoder technique.
- 1D convolutional layers are used with Max-pooling and Up-sampling layers to create encoding and decoding layers respectively.
- Classifier also includes Flattening and Dense layers.
- ReLu activation function is used in all except the last layer which uses Sigmoid.

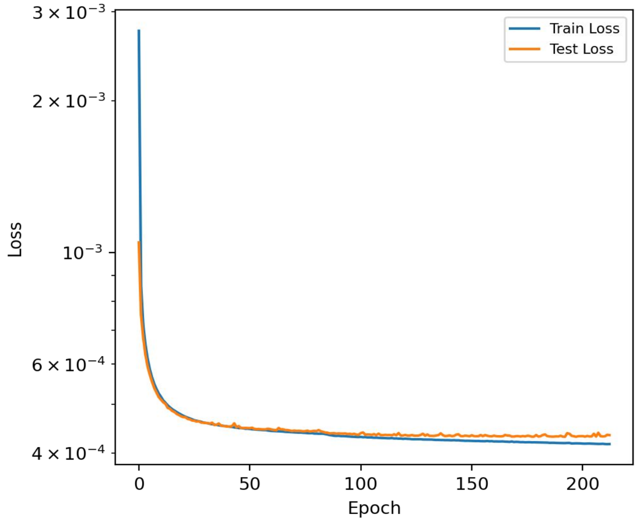



Model Architecture

# Learning Curve

- Mean squared error (MSE) is used as a loss function.
- Early stopping is used:
    - Stop training If test loss is not decreasing after 20 epochs.

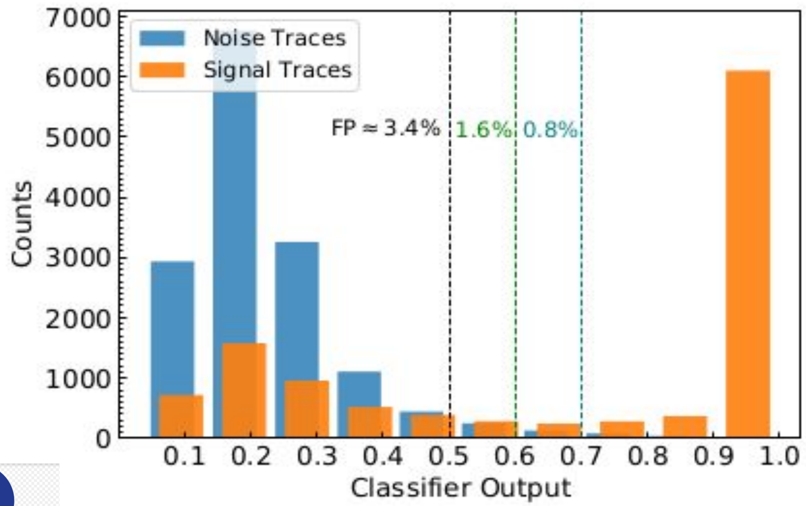
- Not much difference in the training and testing loss, which means that the network is not overfitting.

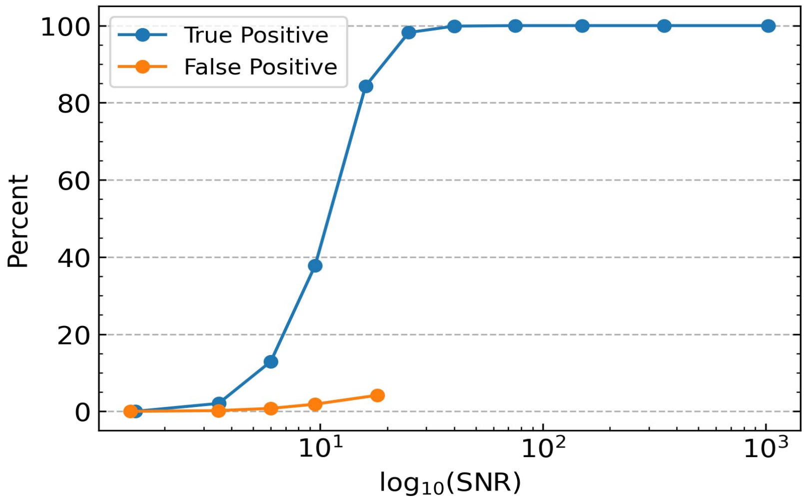



UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Classifier Results

- **Validation set:** 11k signal + 15K background traces. Similar SNR distribution as training and testing set.
- Threshold for signal trace: output value ≥ **0.6** .
- TP and FP rates (in percent), shown in the right plot.

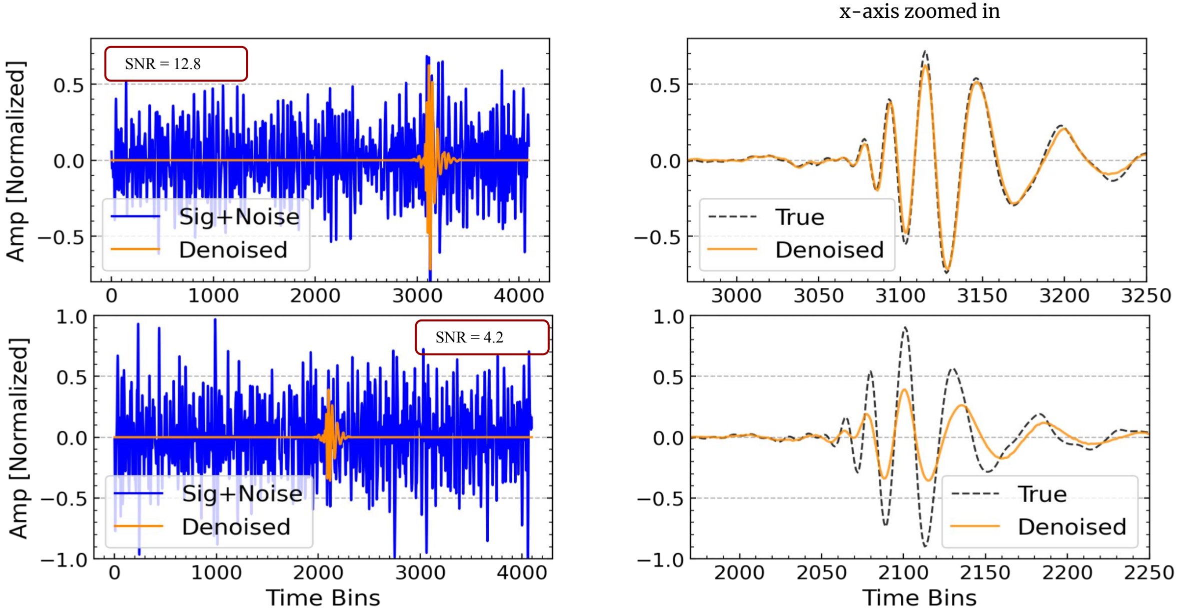



UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Denoiser Results

- Classified Signal traces are passed to the Denoiser for cleaning.
- Two examples shown in bottom left.
- 1st row → best case scenario.
- 2nd row → worse case.

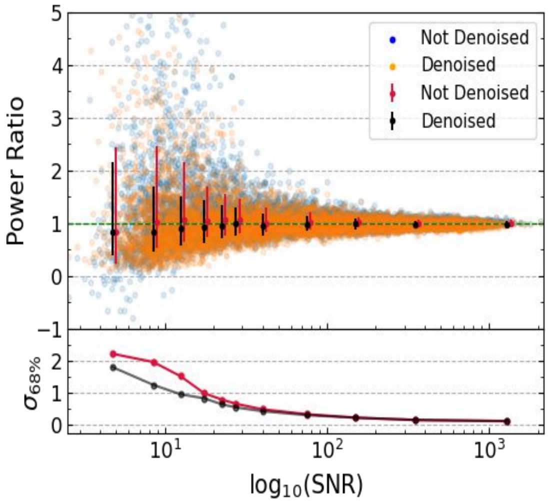


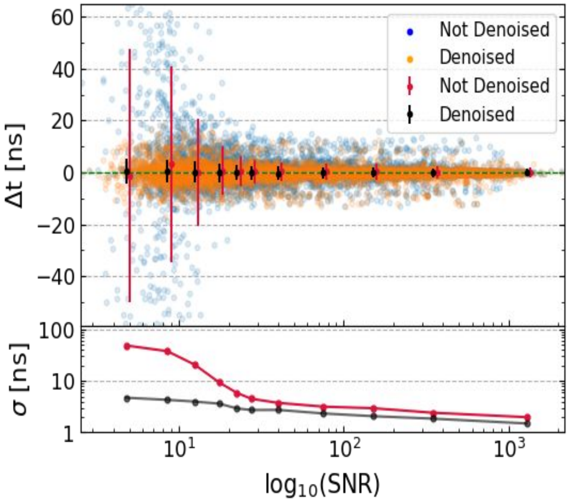
# Accuracy Metrics

➢ Left Plot ⟶ $\text{Power Ratio} = \dfrac{[P_S - P_N]_{\text{Measured}}}{[P_S - P_N]_{\text{True}}}$ .

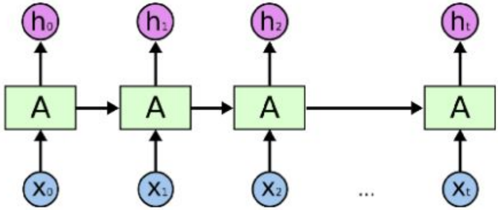
$P_S$ , $P_N$ = power in the signal and noise window, respectively.

➢ Right Plot ⟶ $\Delta t = T_{\text{measured}} - T_{\text{true}}$ .



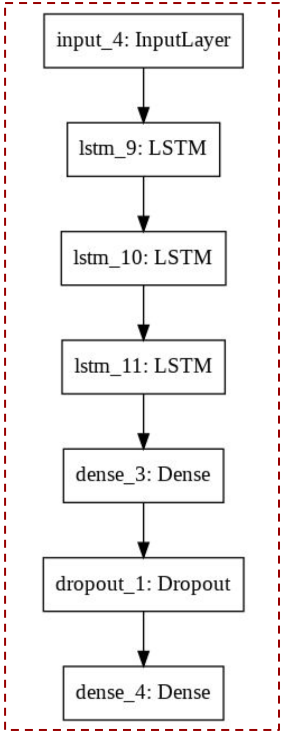



# Other Networks tried




- RNNs process inputs in a **sequential manner**, where the information from the previous input is considered when computing the output of the current step.
- We have used the type of RNN called LSTM.
- LSTM's have the ability to preserve the long-term memory.

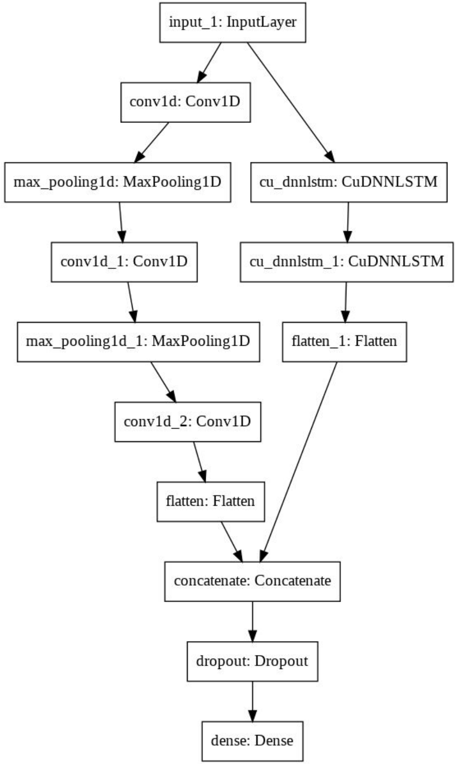
.

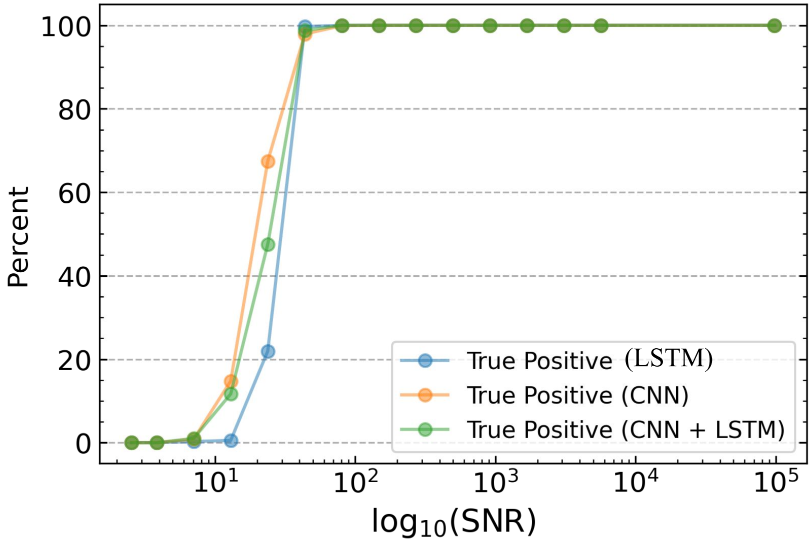



UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# CNN + LSTM

- We have also tried combining the two networks.





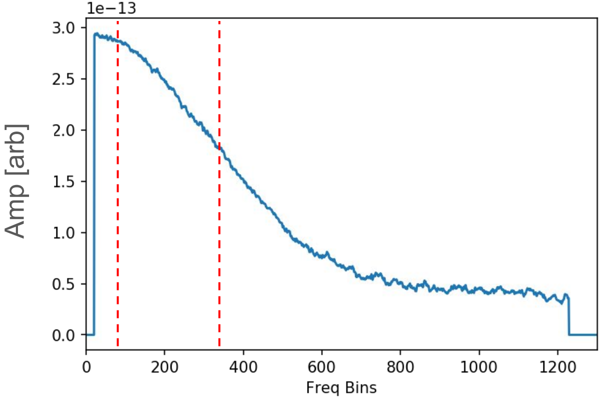

UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

# Using the Frequency spectrum

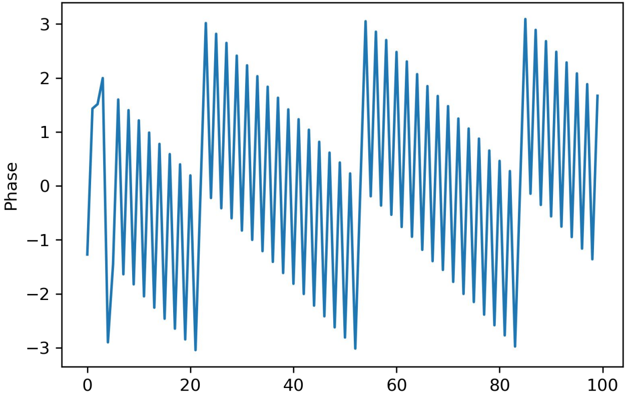
- FFT of time series gives us complex frequency spectrum.
- We can either use the Real and Imaginary part of the spectrum.
- Or we can use Frequency Amplitude and Phases.
- So far this does not improve the results.

Question: How to deal with the Phases?

Amplitudes
[Full Band]




Phases



# Summary and Outlook

- ❑ We are using Convolution networks to study radio signal from cosmic-ray air showers.
- ❑ CoREAS simulations are used to produce radio signals.
- ❑ For background we used modeled noise.
- ❑ Other networks like LSTM are also tried but they did not perform better than the CNNs.
- ❑ We have used time series information for networks training, Currently working on using the frequency spectrum as well.

UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE