Paula Gálvez Molina, Abdul Rehman, Frank Schröder, for the IceCube collaboration
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

Wednesday January 29th, 2025

UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

ICECUBE
NEUTRINO OBSERVATORY

Detection of Radio Signals from Cosmic Rays Using Convolutional Neural Networks with Data from SKALA antennas at IceTop

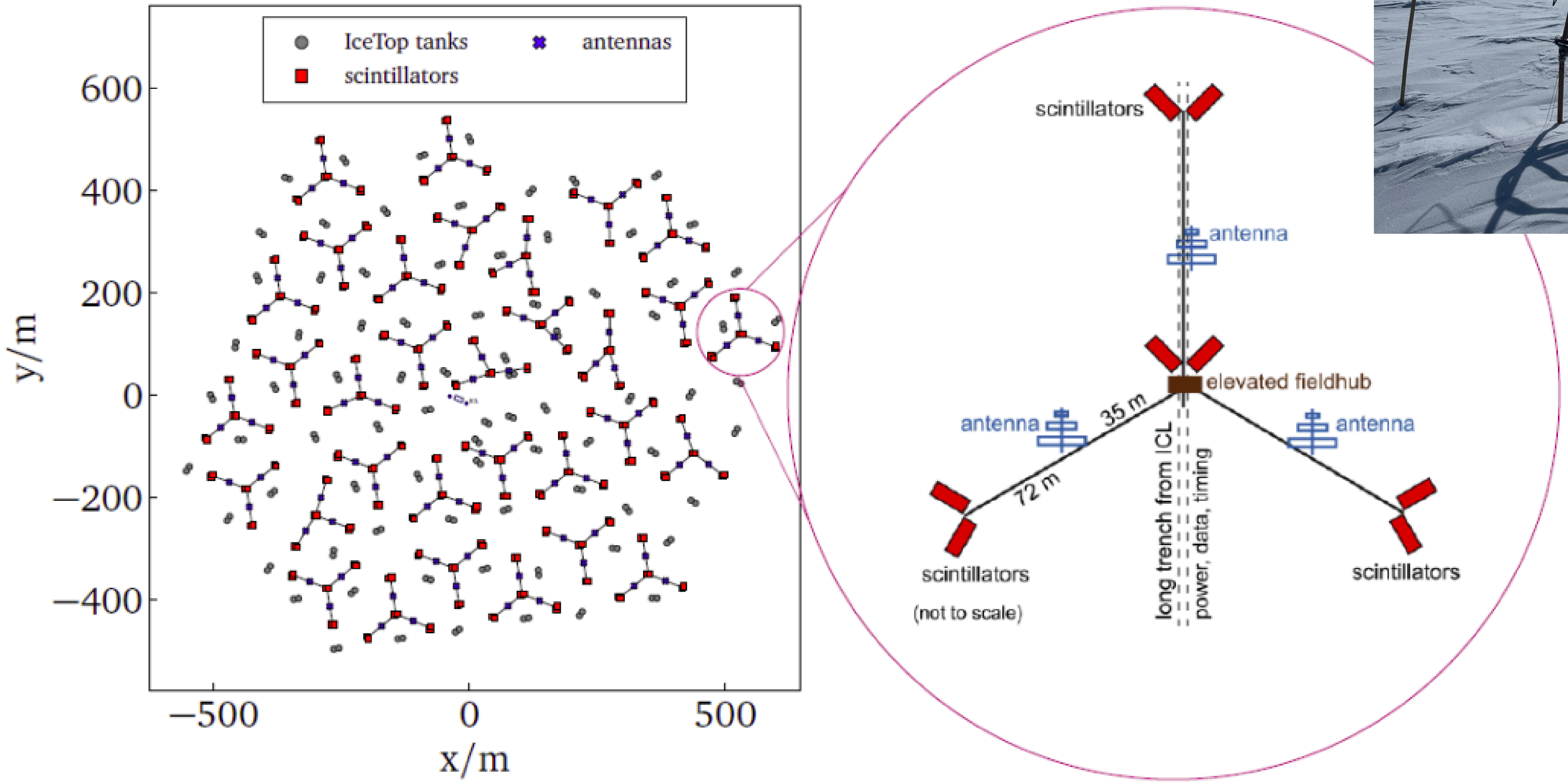
Workshop on Machine Learning for Analysis of High-Energy Cosmic Particles

01

MOTIVATION

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating station of the Surface Enhancement for IceCube measures radio signals from air showers



Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise!

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)

SNR: Signal to Noise Ratio

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)

SNR: Signal to Noise Ratio

Peak value of absolute value of the waveform.

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)

SNR: Signal to Noise Ratio

Bin the waveform and compute the Root Mean Squared (RMS). Choose the **Median RMS.**[2]

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)

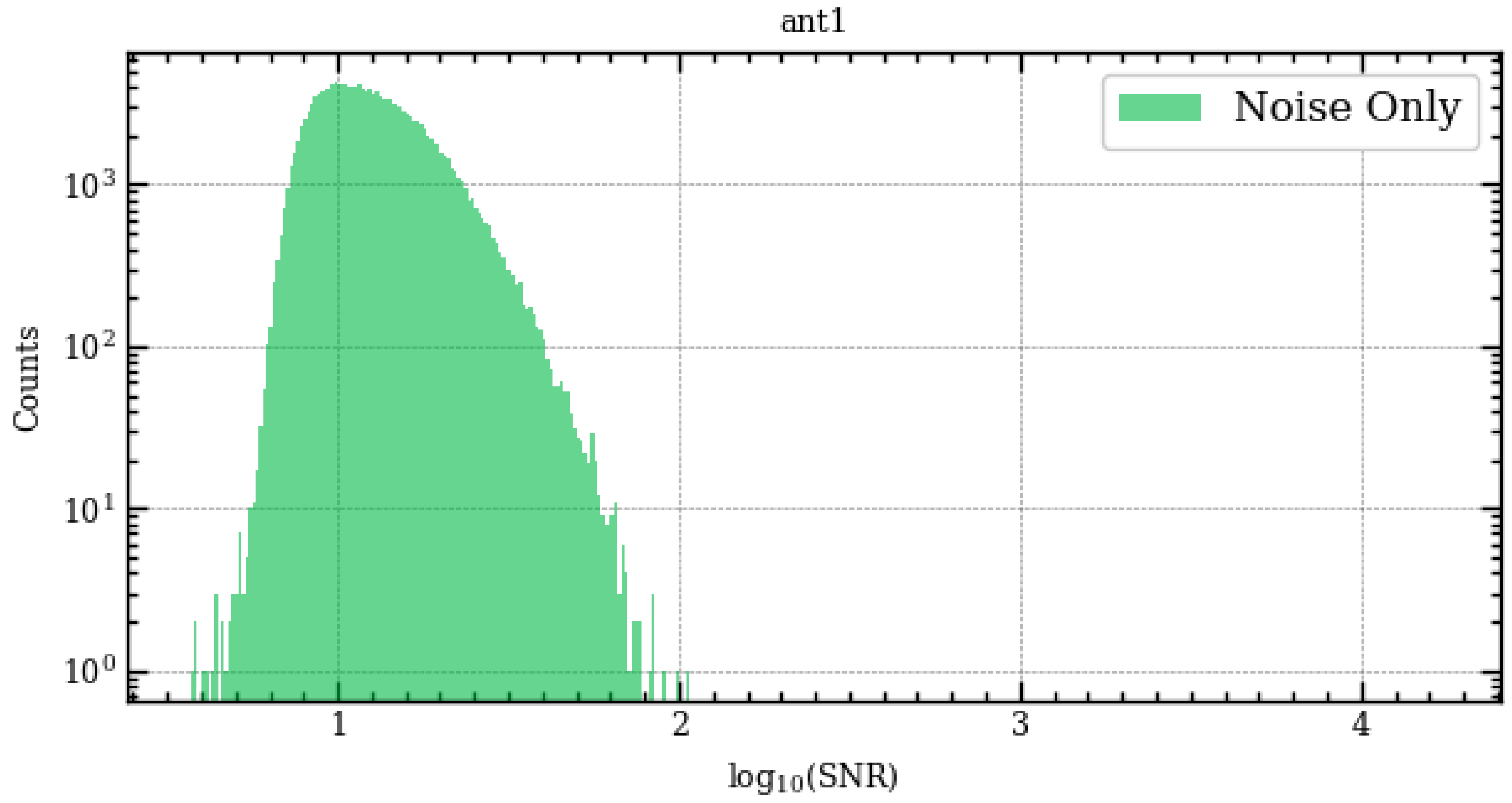
SNR: Signal to Noise Ratio

$$\text{SNR} = \frac{\text{Peak value}^2}{\text{Median RMS}^2}$$

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)

SNR: Signal to Noise Ratio

$$\mathrm{SNR} = \frac{\text{Peak value}^2}{\text{Median RMS}^2}$$

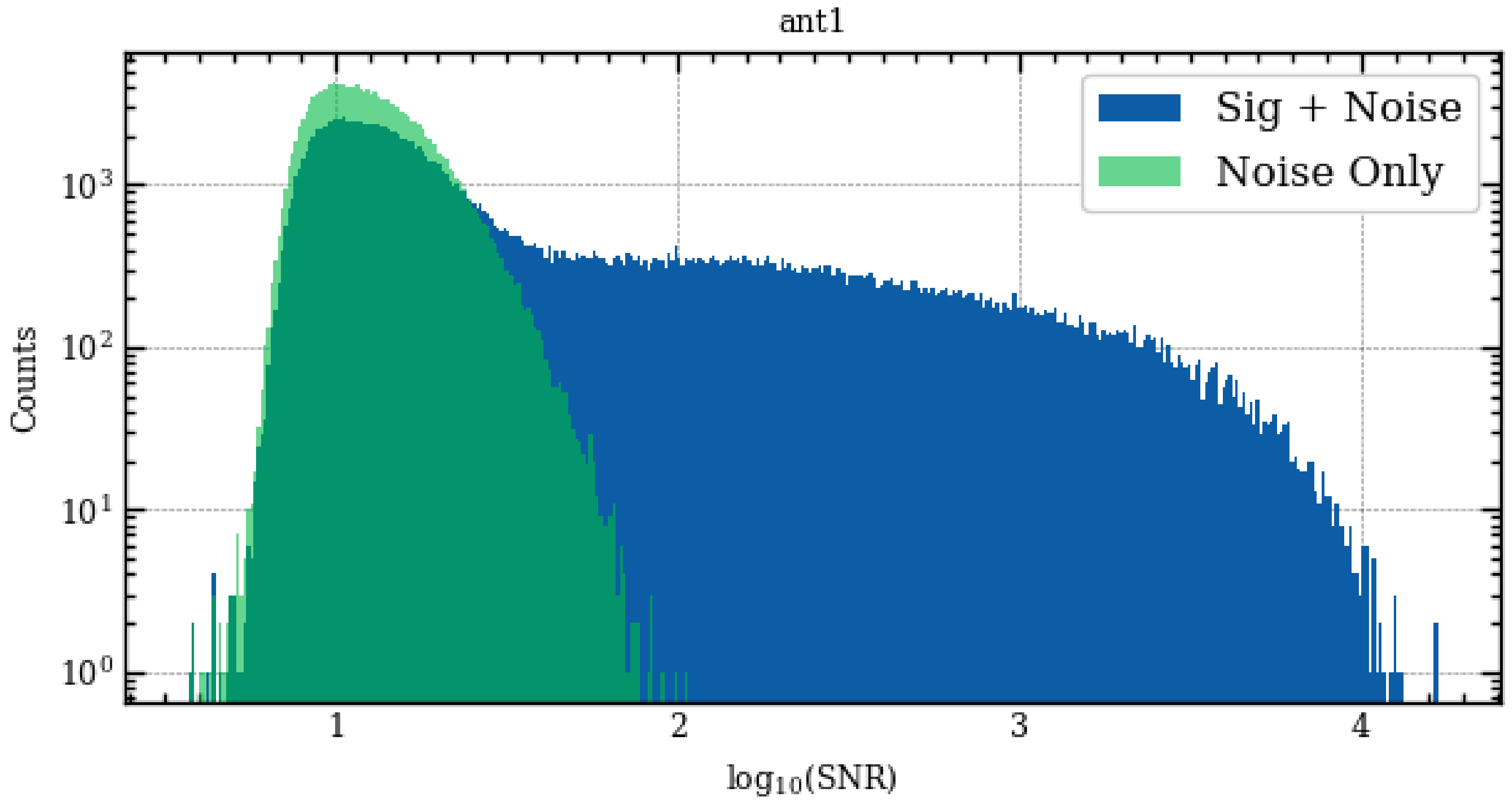


Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)
- Previous detection techniques (such as SNR cuts) lose information at low SNR

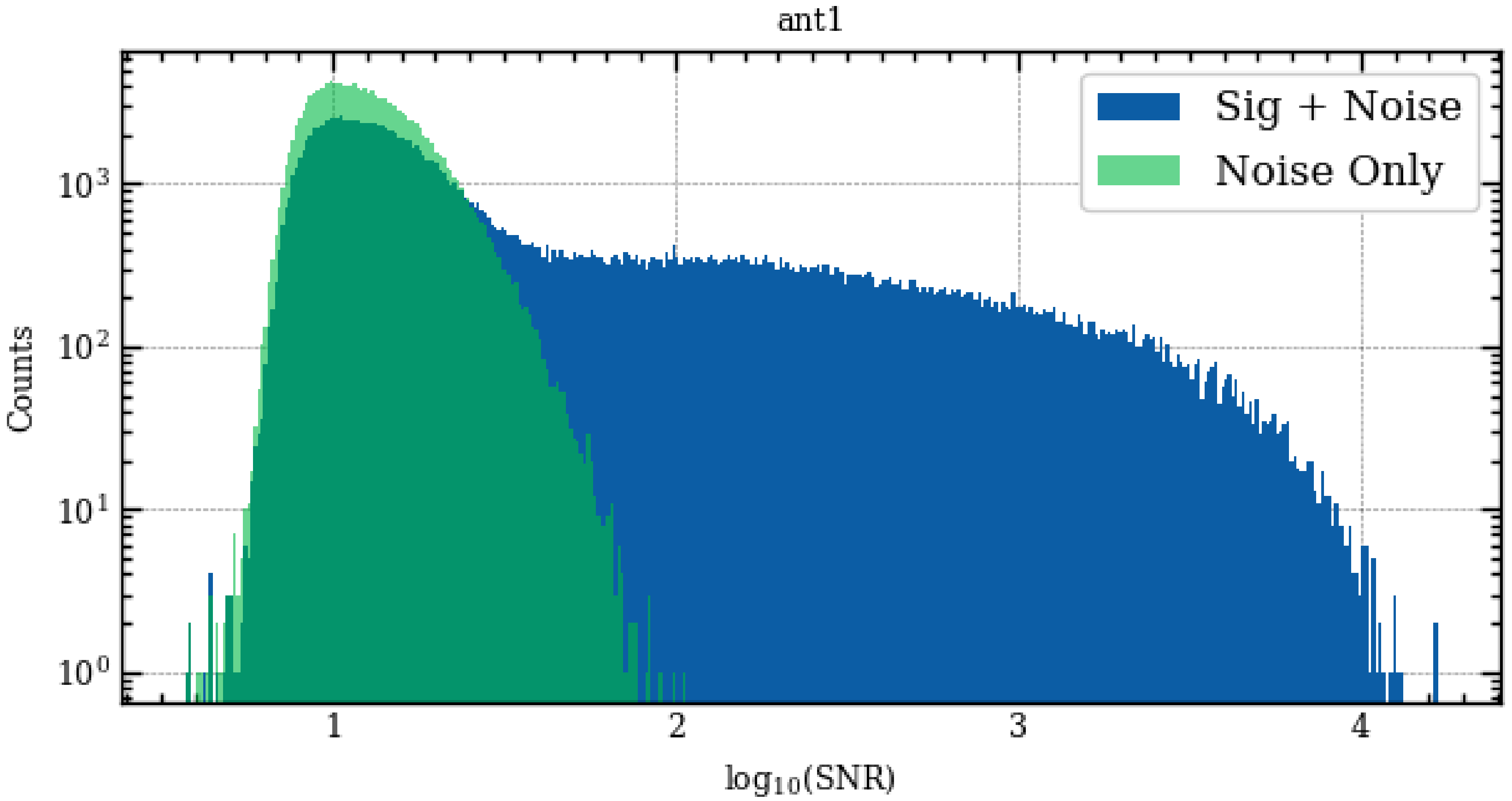
SNR: Signal to Noise Ratio

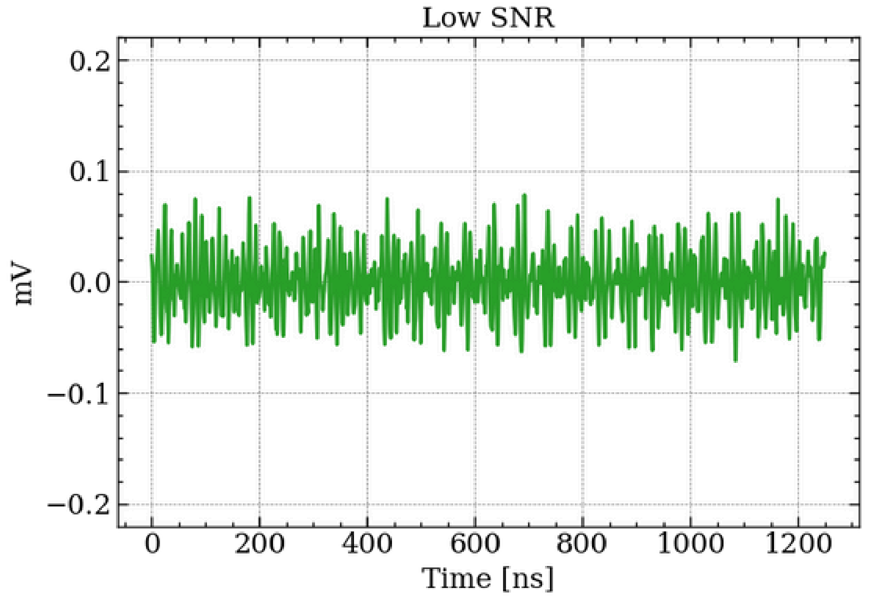
$$\mathrm{SNR} = \frac{\text{Peak value}^2}{\text{Median RMS}^2}$$

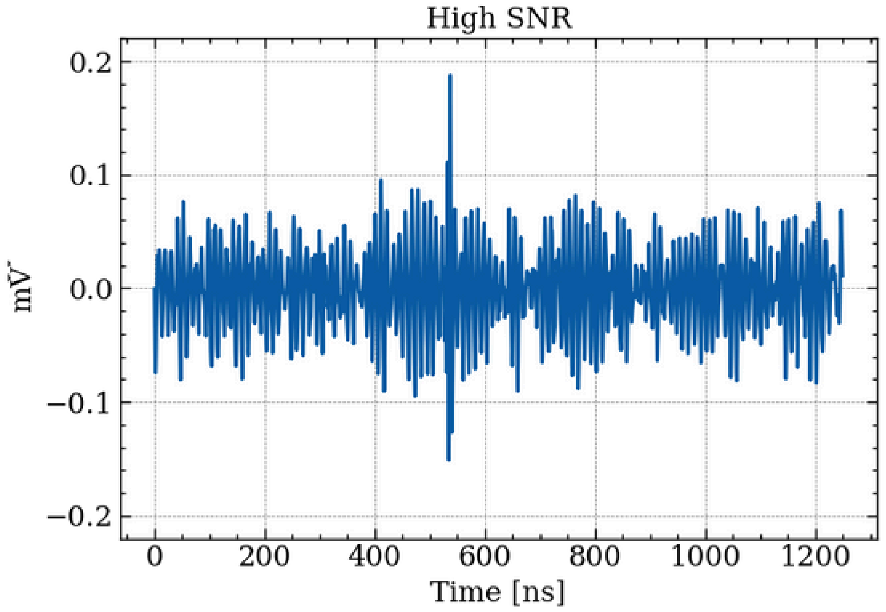


Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

- Currently an operating prototype station of the Surface Enhancement for IceCube measures radio signals from air showers
- There is noise! (SNR: proxy for noisiness)
- Previous detection techniques (such as SNR cuts) lose information at low SNR







Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

GOAL

Apply Convolutional Neural Networks to Classify radio waveforms and Denoise them.

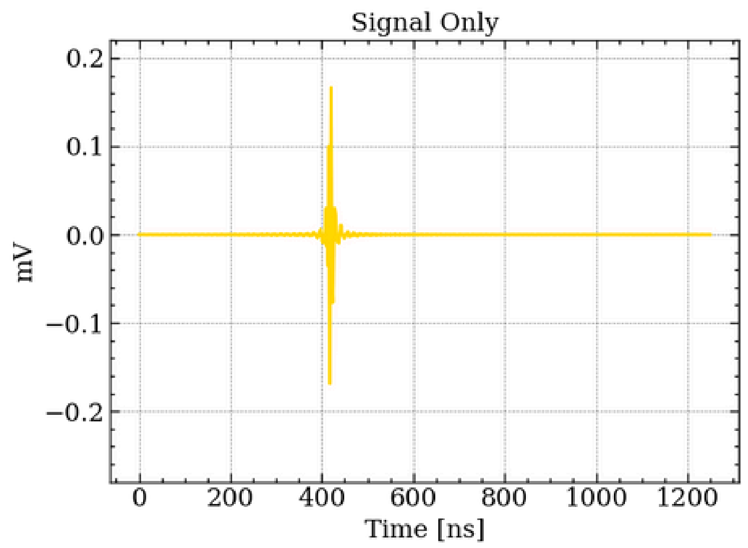
02

METHODOLOGY

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

DATA PREPARATION

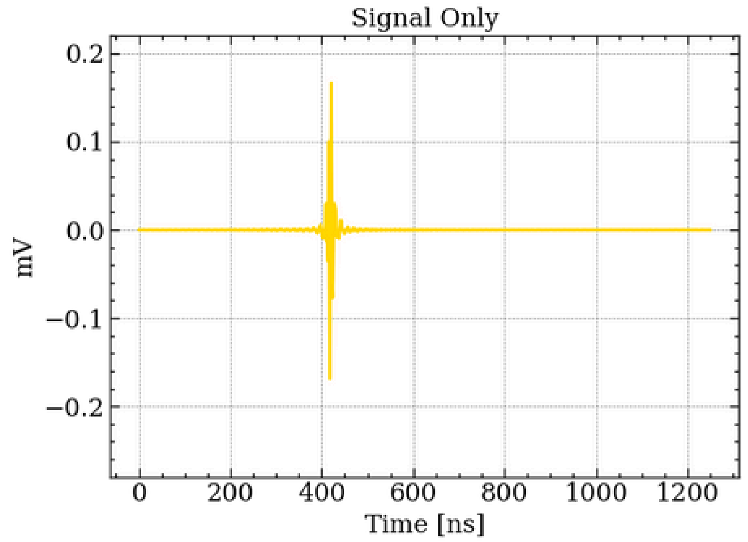
- Generate waveform data set:
  - (CoREAS) Simulated Signals

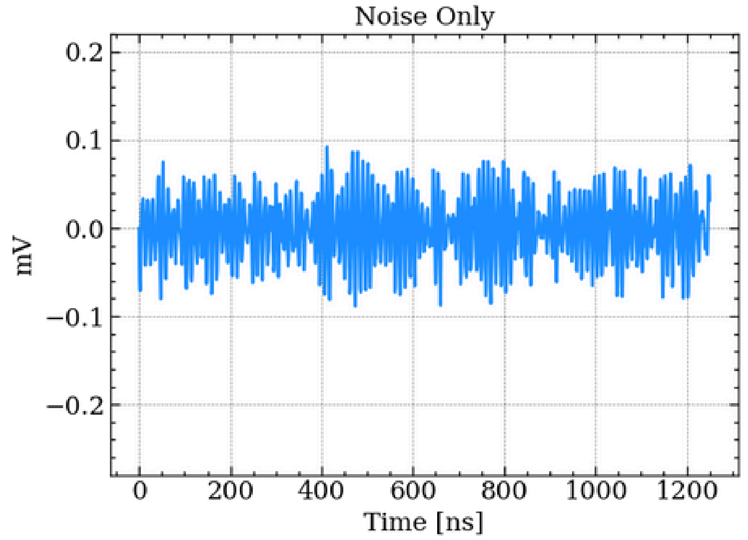


Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

DATA PREPARATION

- Generate waveform data set:
  - (CoREAS) Simulated Signals
  - Measured Noise

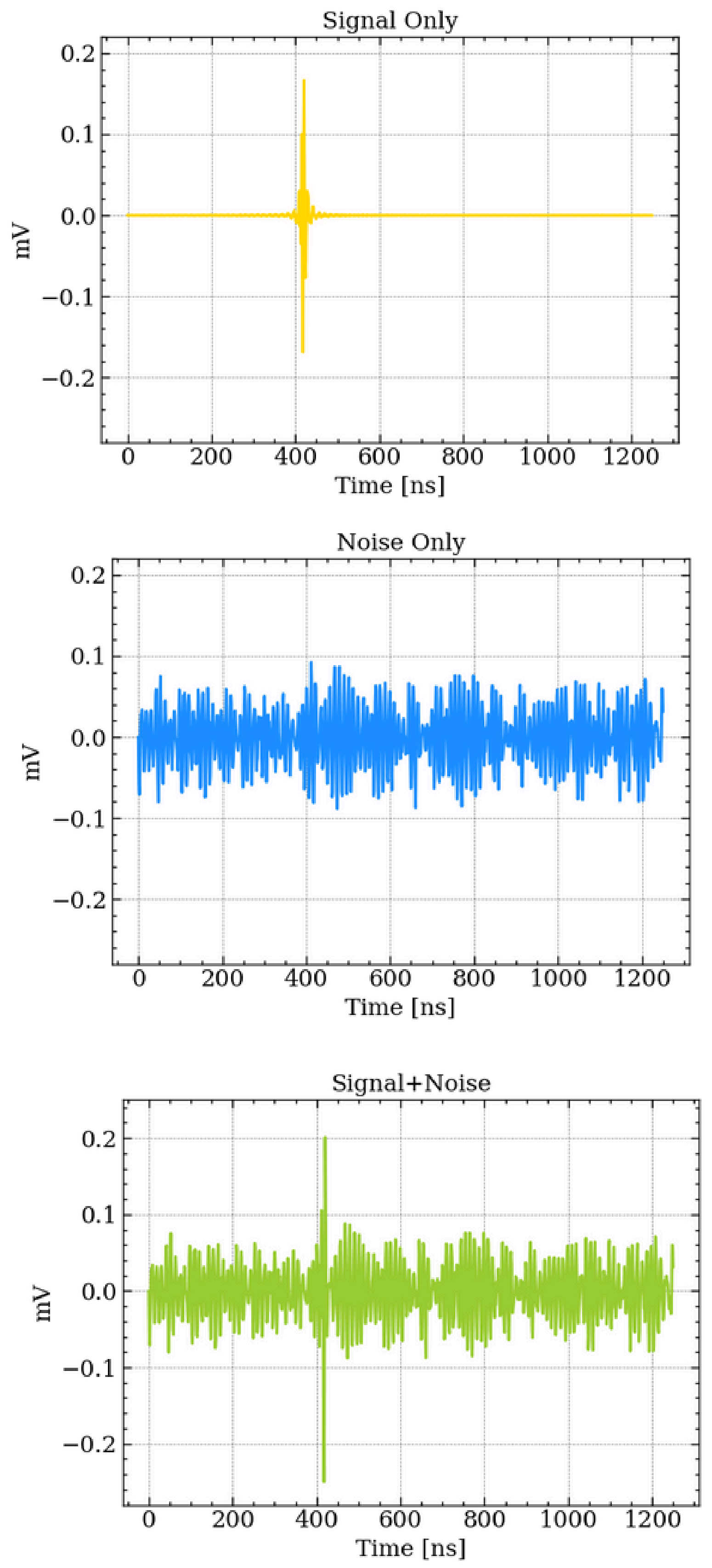




Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

DATA PREPARATION

- Generate waveform data set:
  - (CoREAS) Simulated Signals
  - Measured Noise
  - Signal + Noise waveforms



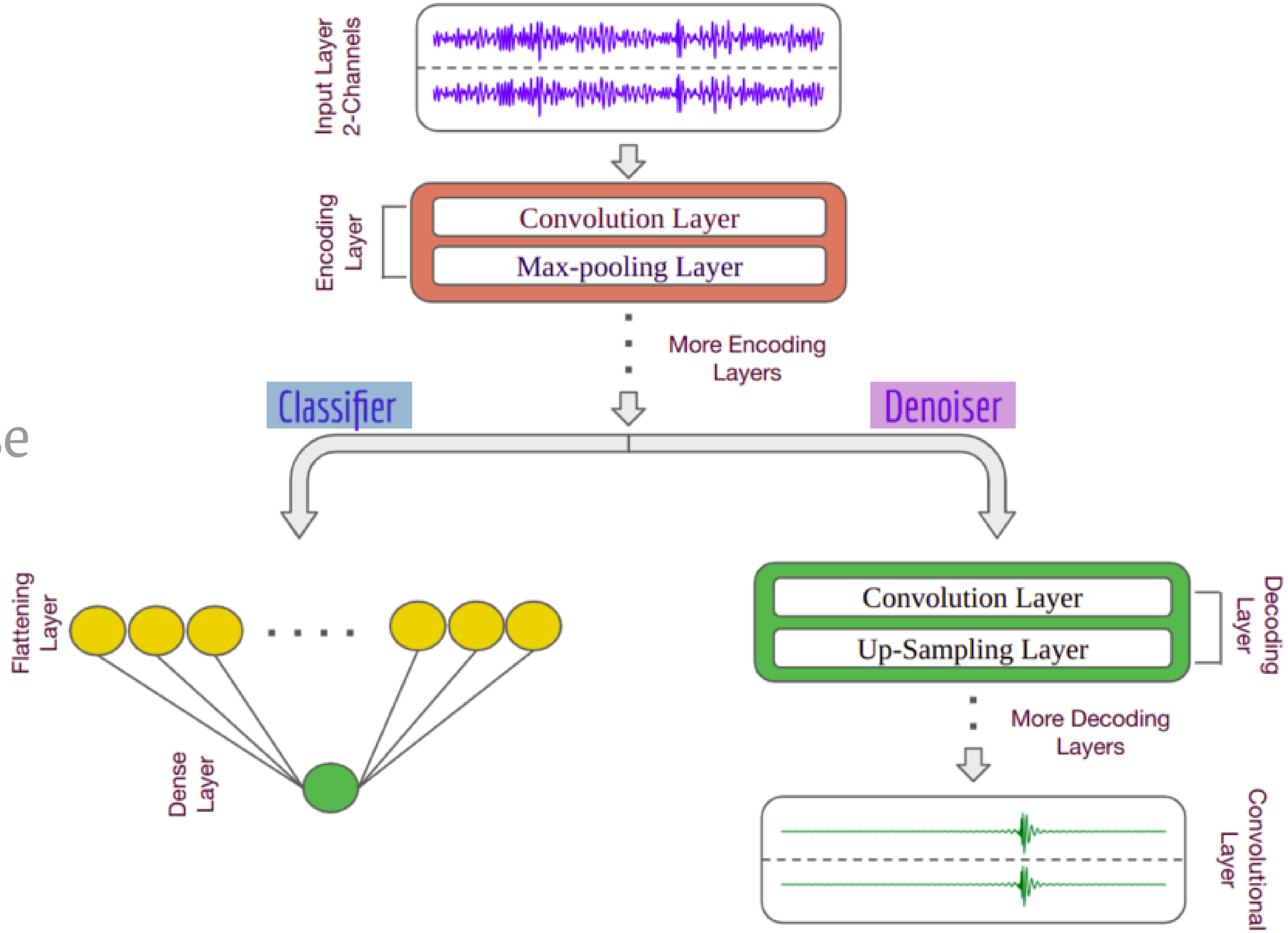
DATA PREPARATION

- Generate waveform data set:
  - (CoREAS) Simulated Signals, Measured Noise, Signal + Noise waveforms
    - Signal:
      - 10^16 to 10^18 eV, and zenith angles from 0 to 0.9 sin^2 (theta)
    - Noise:
      - January - July 2022 Measured Noise
      - Frequency band of 70-350 MHz (Two polarization channels)
      - Traces of 1000 samples with 1 ns binning

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

TRAINING AND TESTING

- Generate waveform data set:
  - (CoREAS) Simulated Signals, Measured Noise, Signal + Noise waveforms
- Create 2 CNNs for Classifying and Denoising



PoS(ARENA2024)034

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

TRAINING AND TESTING

- Generate waveform data set:
    - (CoREAS) Simulated Signals, Measured Noise, Signal + Noise waveforms
- Create 2 CNNs for Classifying and Denoising
- Split the dataset in 80%-20% for training and testing and train the CNNs

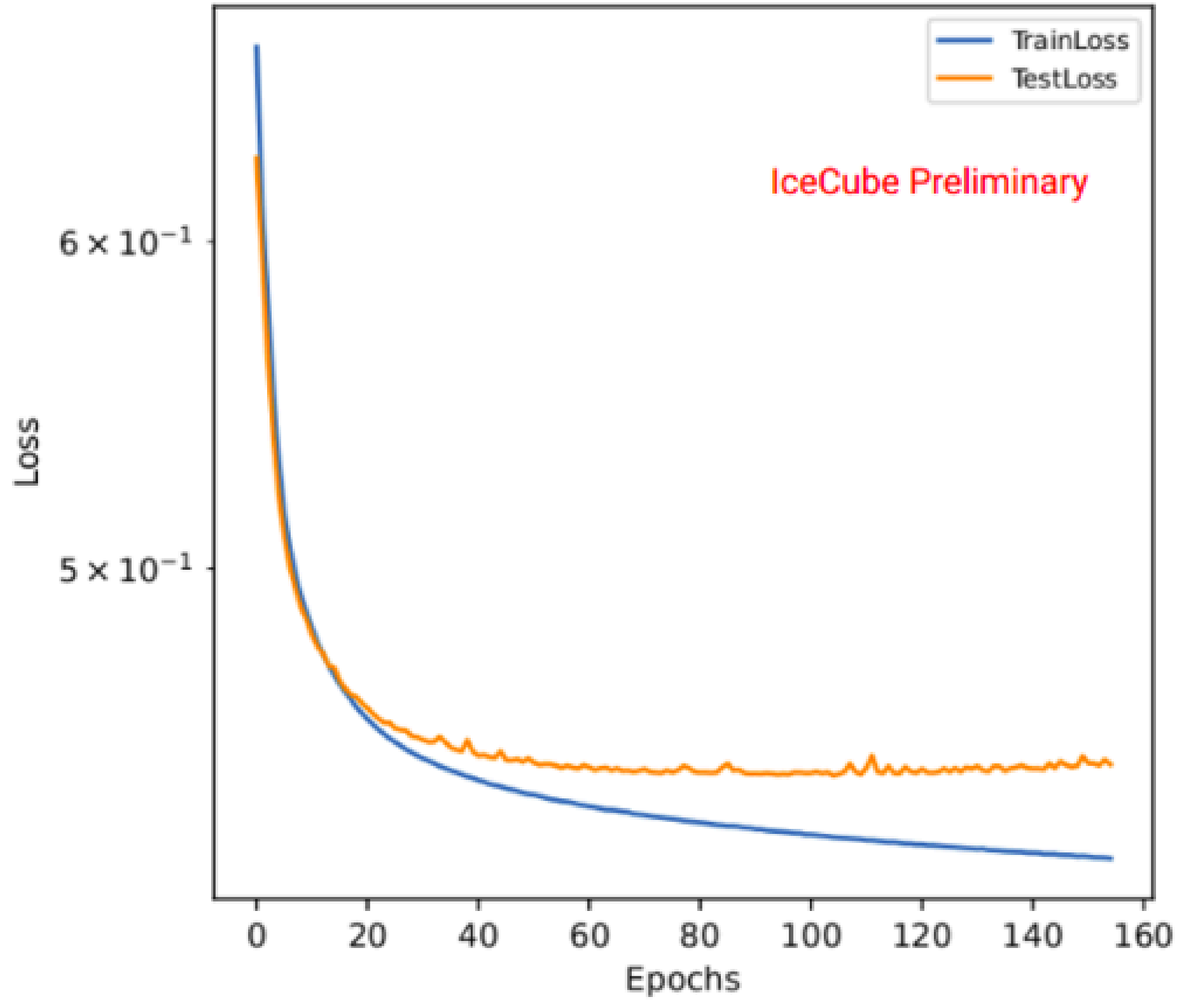
03

FIRST RESULTS

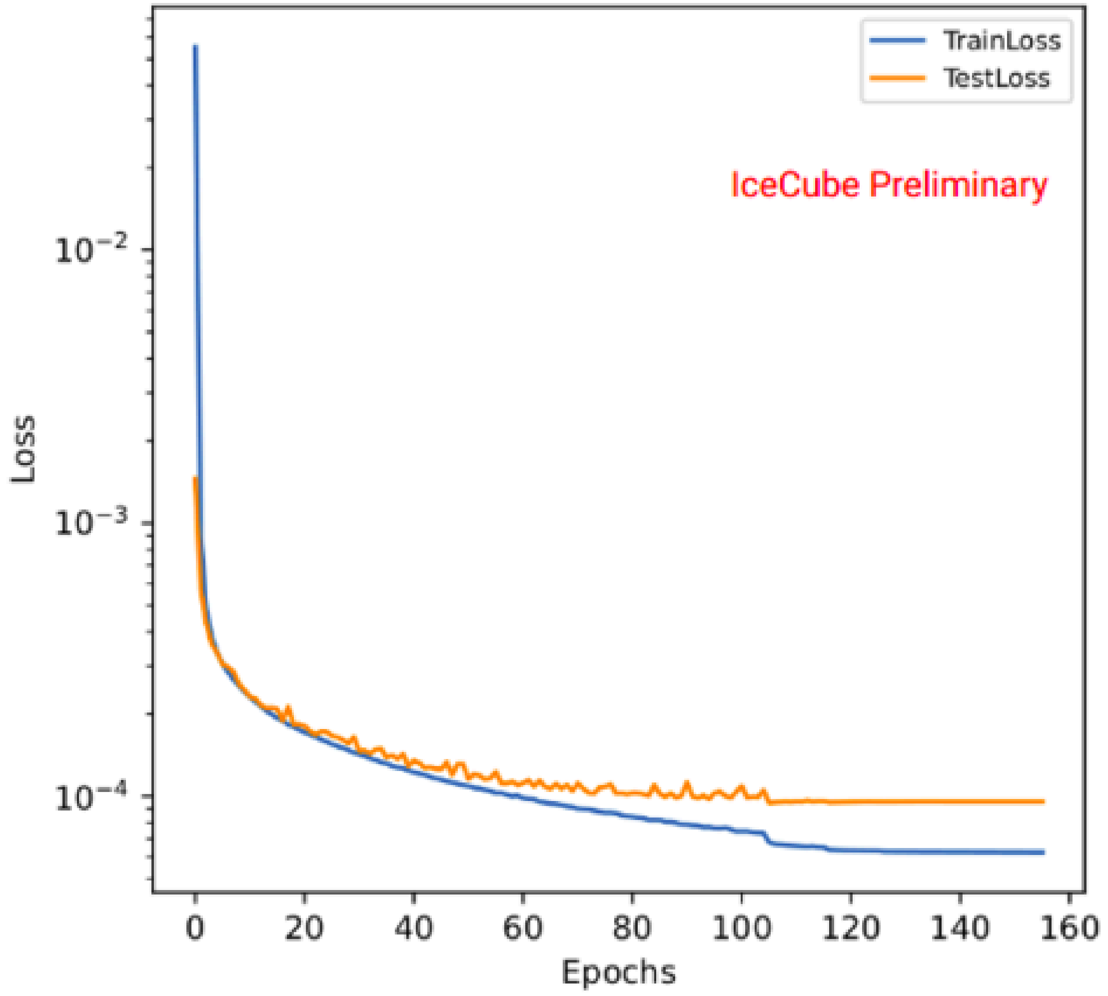
Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

LOSS CURVES

Classifier



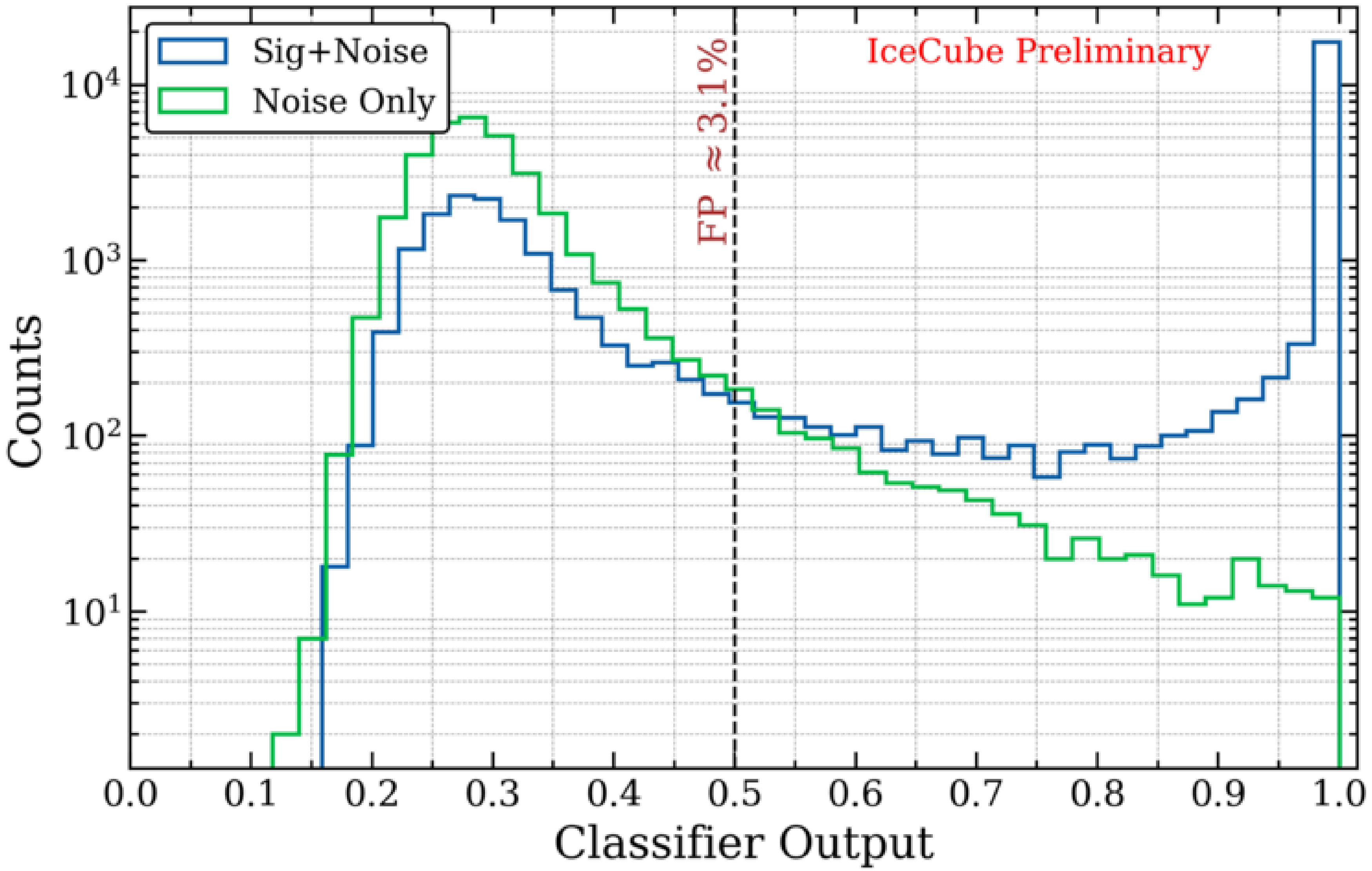
Denoiser

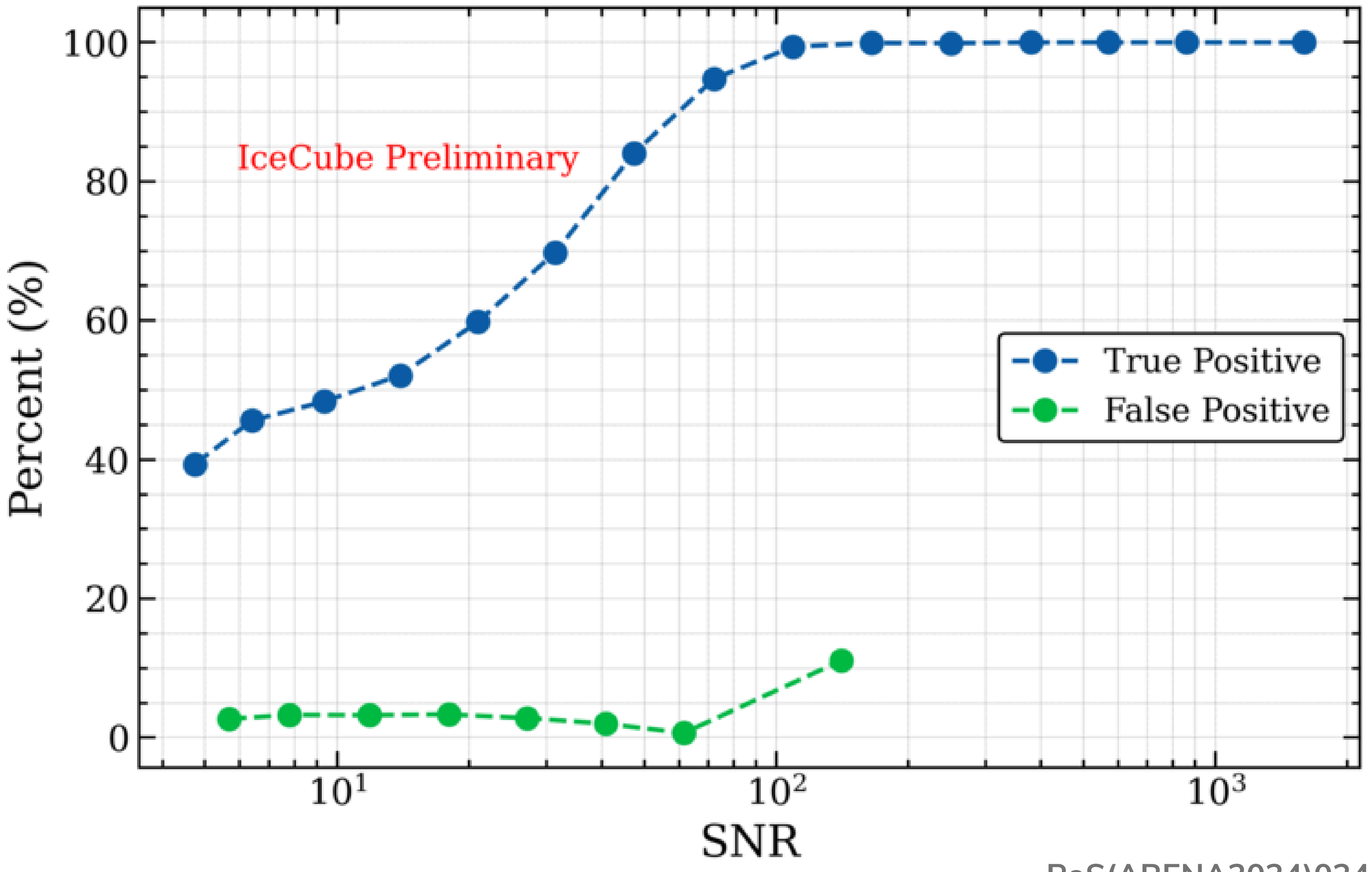


PoS(ARENA2024)034

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

RESULTS CLASSIFIER



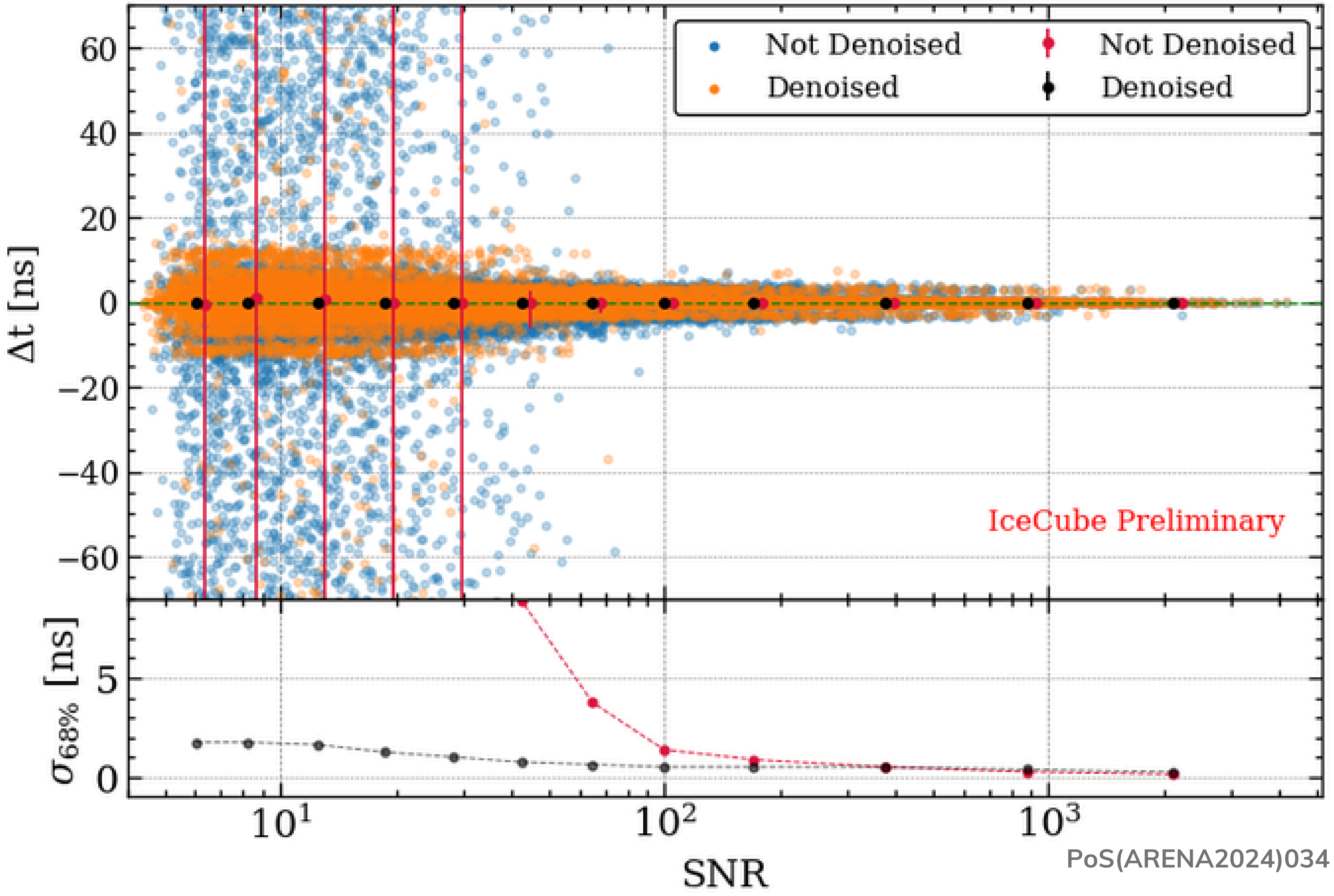


PoS(ARENA2024)034

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

RESULTS

DENOISER



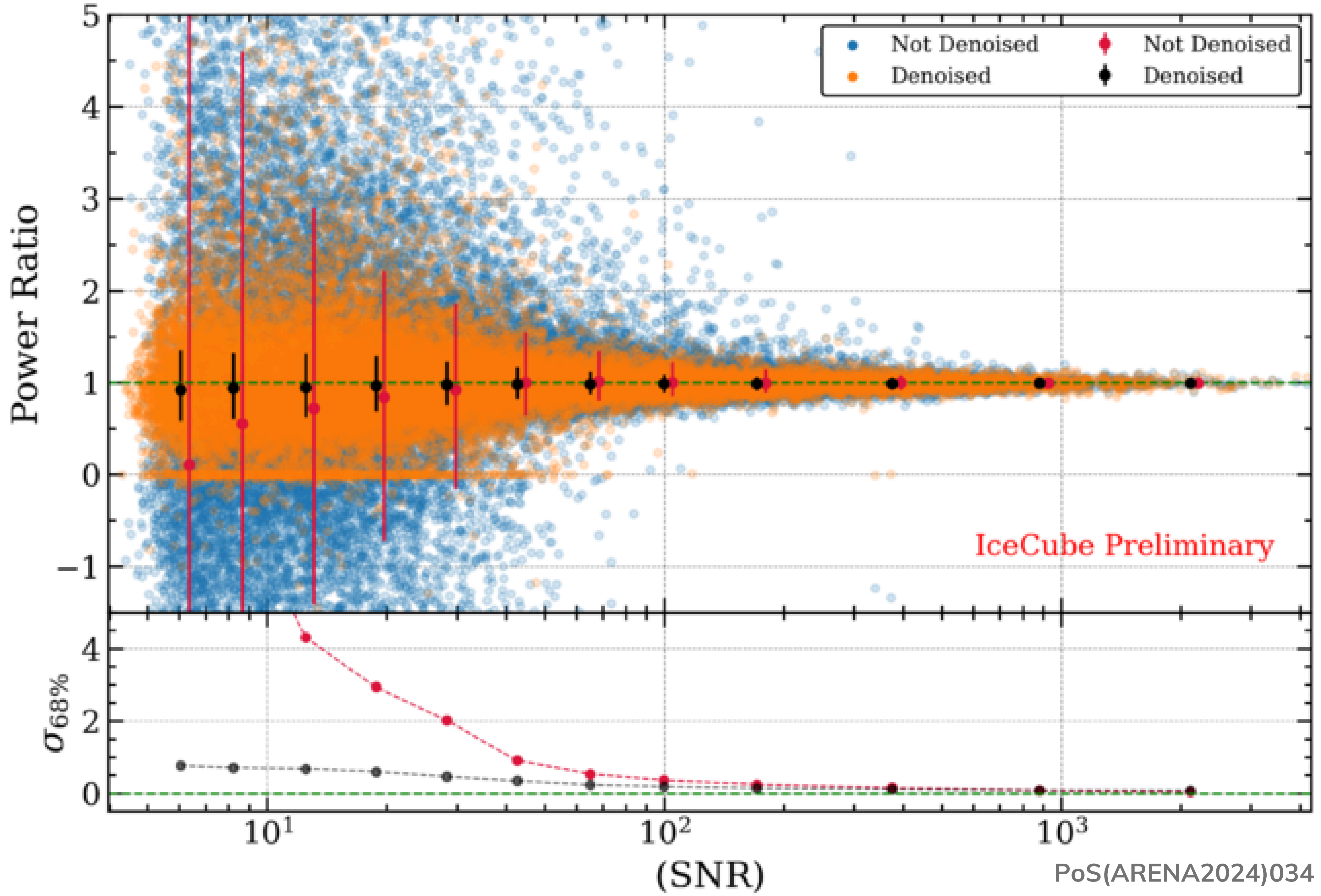
PoS(ARENA2024)034

Peak Time Difference (Δt)

Peak time of denoised signal - peak time of simulated signal

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

RESULTS
DENOISER



Power Ratio

(E_signal,den - E_noise,den) / (E_signal,sim - E_noise,sim)

04

ONGOING WORK

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

MOTIVATION

Can we improve the performance of the CNNs by upsampling the waveforms?

DATA PREPARATION

- Generate waveform data set 1: (CoREAS) Simulated Signals, Measured Noise, Signal + Noise waveforms:
  - Signal:
    - 10^17.5 to 10^18 eV, and zenith angles from 0 to 0.9 sin^2 (theta)
  - Noise:
    - January - May 2024 Measured Noise
    - Frequency band of 70-350 MHz (Two polarization channels)
    - Traces of 1000 samples with 1.25 ns binning

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

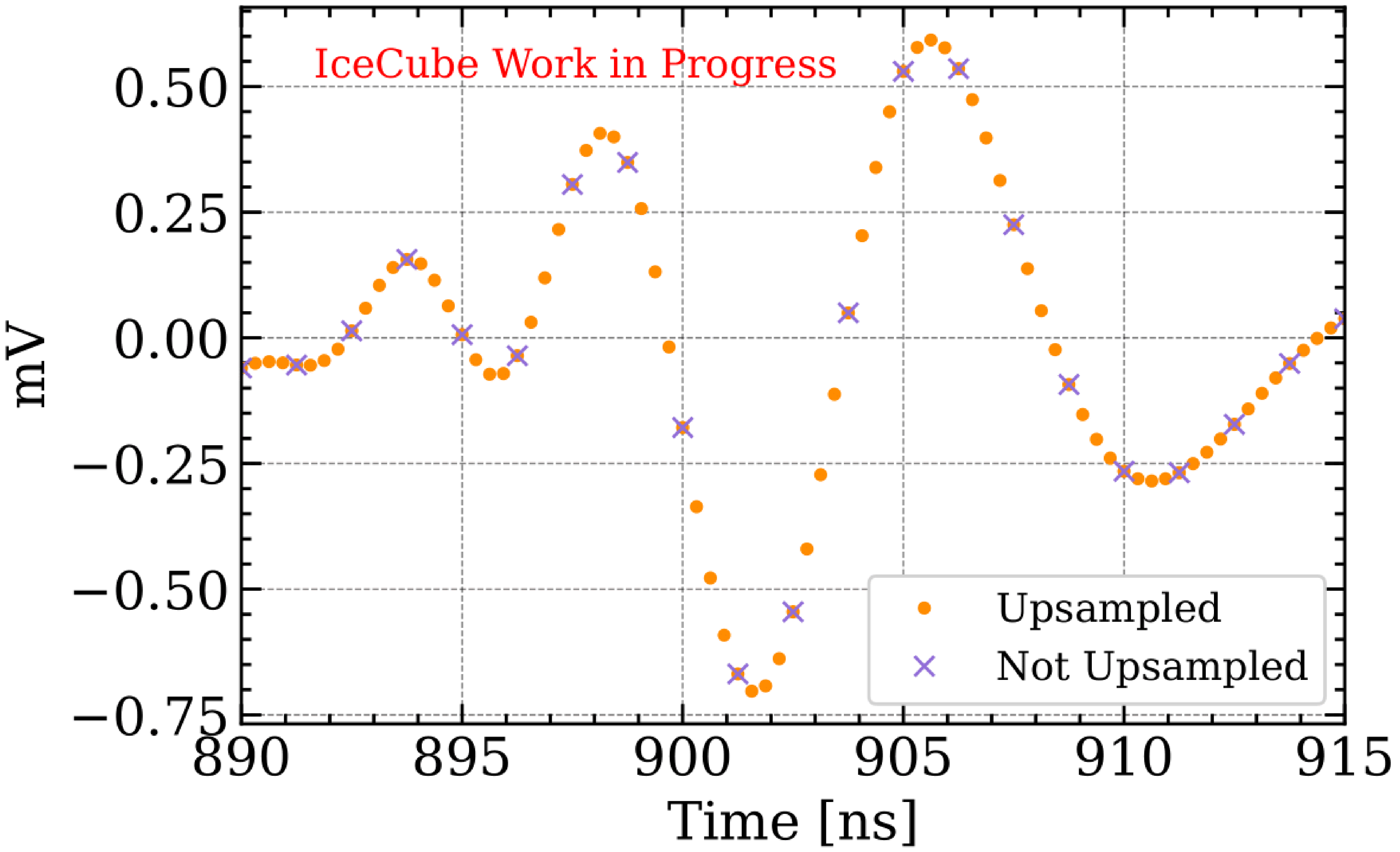
DATA PREPARATION

- Generate waveform data set 1:
  (CoREAS) Simulated Signals,
  Measured Noise, Signal + Noise
  waveforms:
  - Signal:
    - 10^17.5 to 10^18 eV, and zenith angles from 0 to 0.9 sin^2 (theta)
  - Noise:
    - January - May 2024 Measured Noise
    - Frequency band of 70-350 MHz (Two polarization channels)
    - Traces of 1000 samples with 1.25 ns binning

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

DATA PREPARATION

- Generate waveform data set 1
- Generate data set 2:
  - Upsample (x4) data set 1 using zero-padding in frequency domain



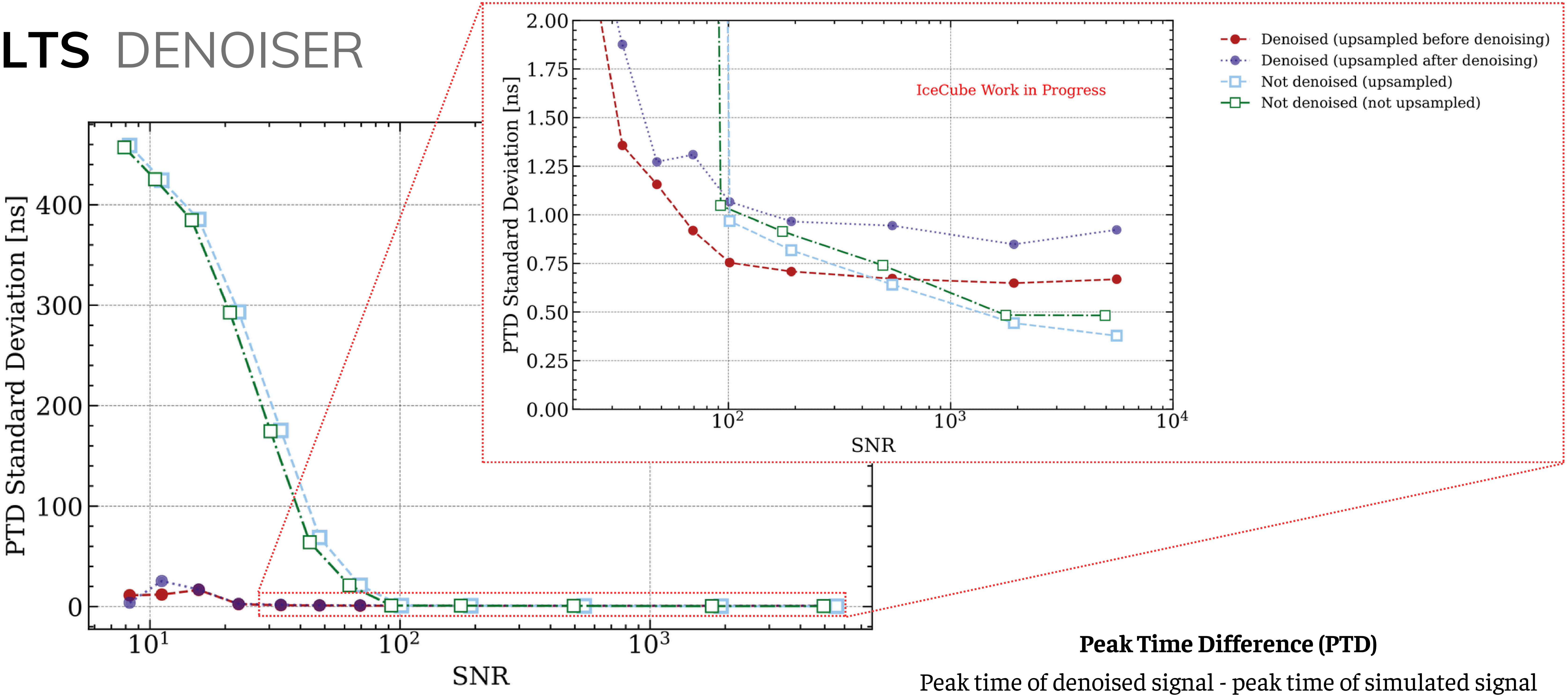
DATA PREPARATION

- Generate waveform data set 1
- Generate data set 2
- Train (using same network architecture)*

* Training was performed 11 times to ensure robustness of results.

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

RESULTS DENOISER

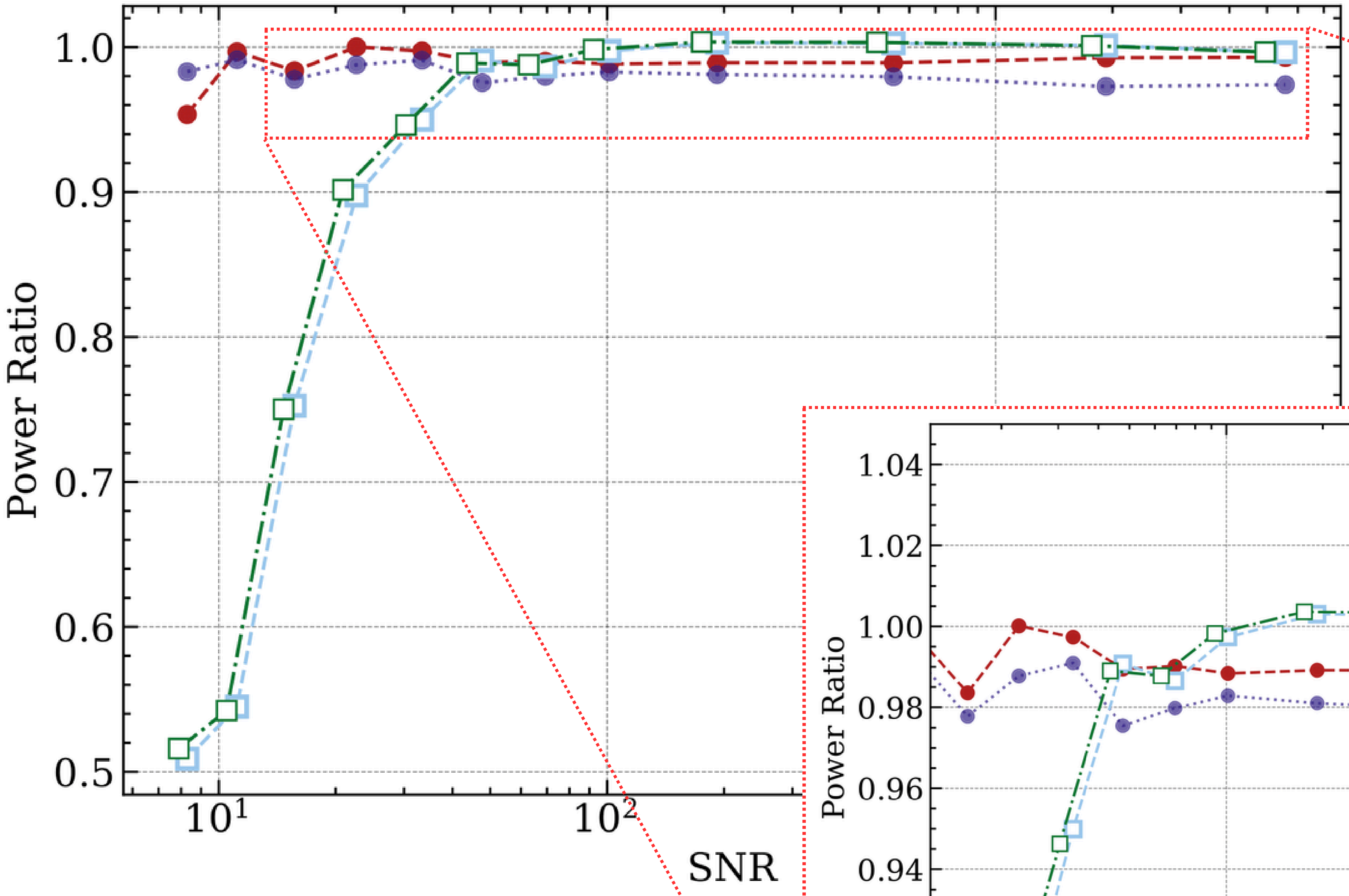


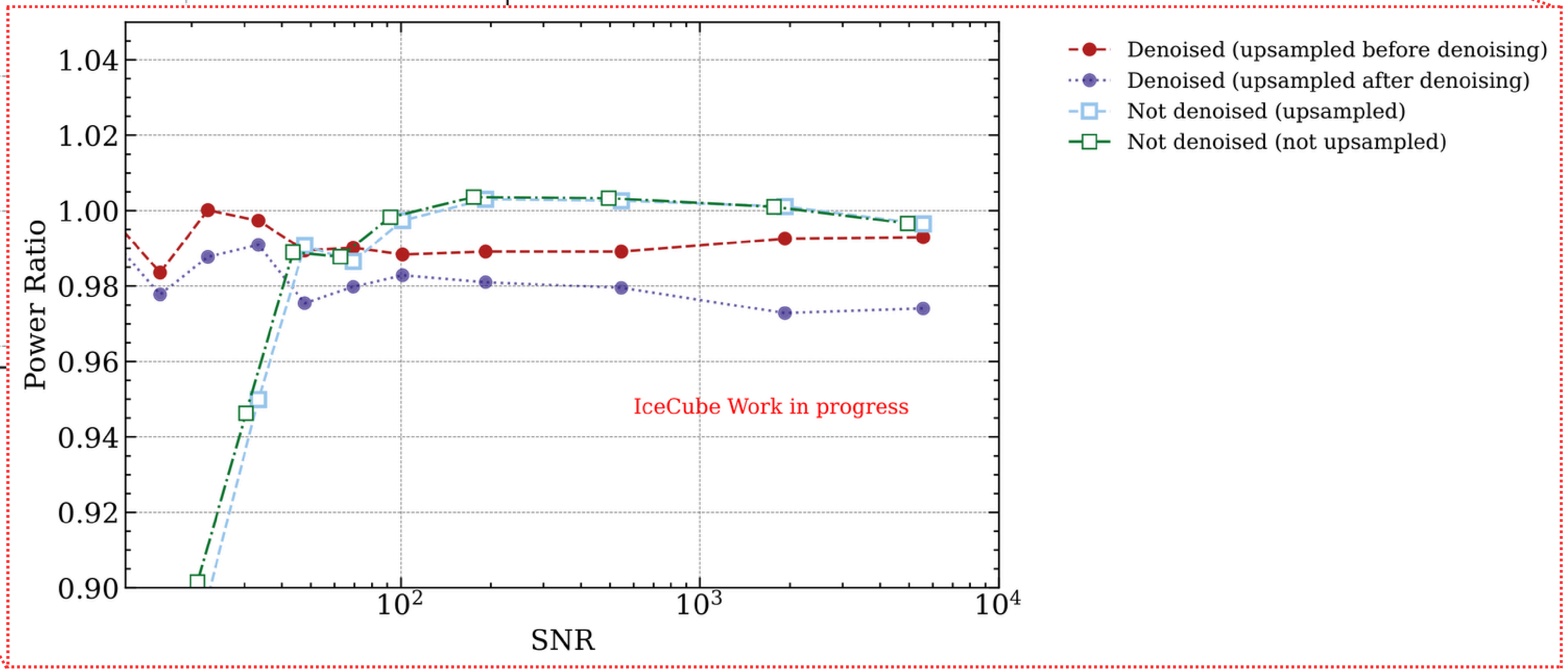
Peak Time Difference (PTD)

Peak time of denoised signal - peak time of simulated signal

Paula Gálvez Molina, Abdul Rehman, Frank Schröder
paulagm@udel.edu; arehman@udel.edu; fgs@udel.edu

RESULTS DENOISER





Power Ratio

(E_signal,den - E_noise,den) / (E_signal,sim - E_noise,sim)

- Training on upsampled traces improves accuracy on location (direction) and amplitude (energy) of the pulse compared to upsampling after denoising or no upsampling at all.

CONCLUSIONS

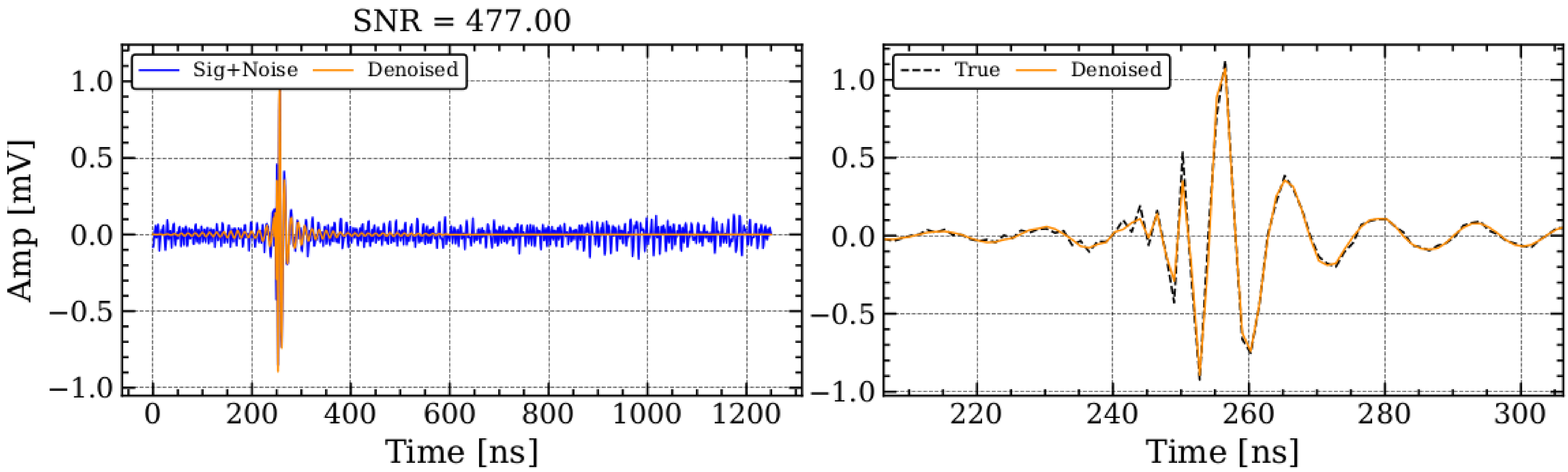
- Training on upsampled traces improves accuracy on location (direction) and amplitude (energy) of the pulse compared to upsampling after denoising or no upsampling at all.

- Upsampling improves pulse denoising of traces at low and intermediate SNR.

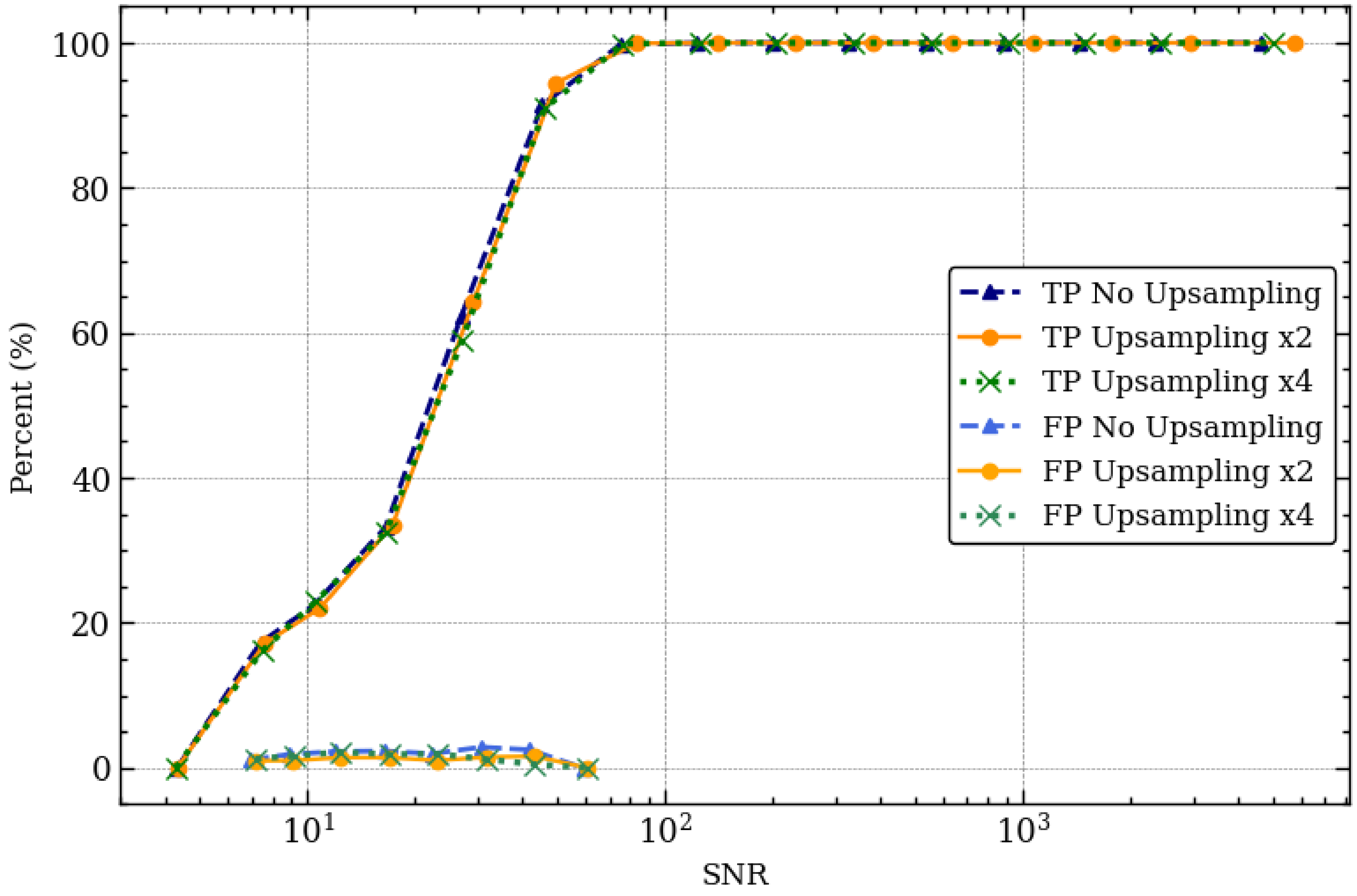
CONCLUSIONS

THANKS

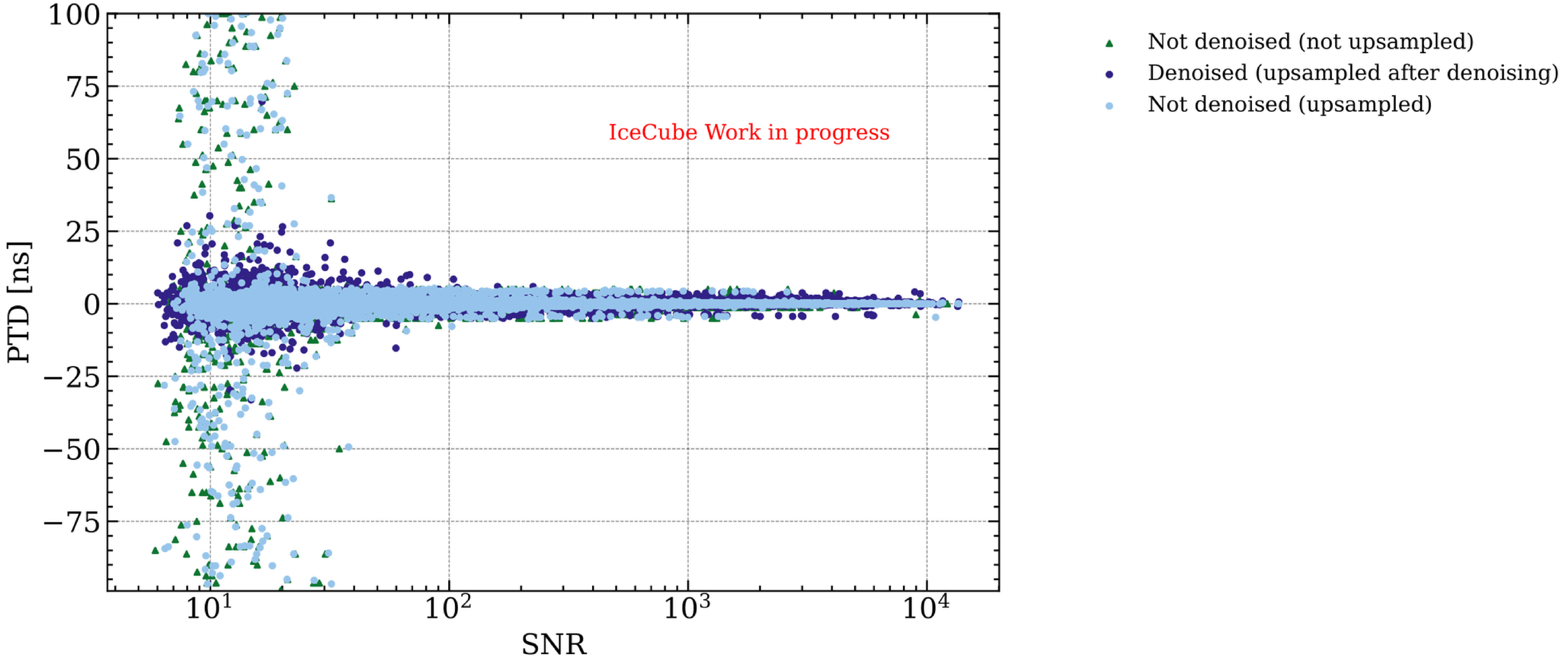
SUPPORTING SLIDES



SAMPLE DENOISED WAVEFORM

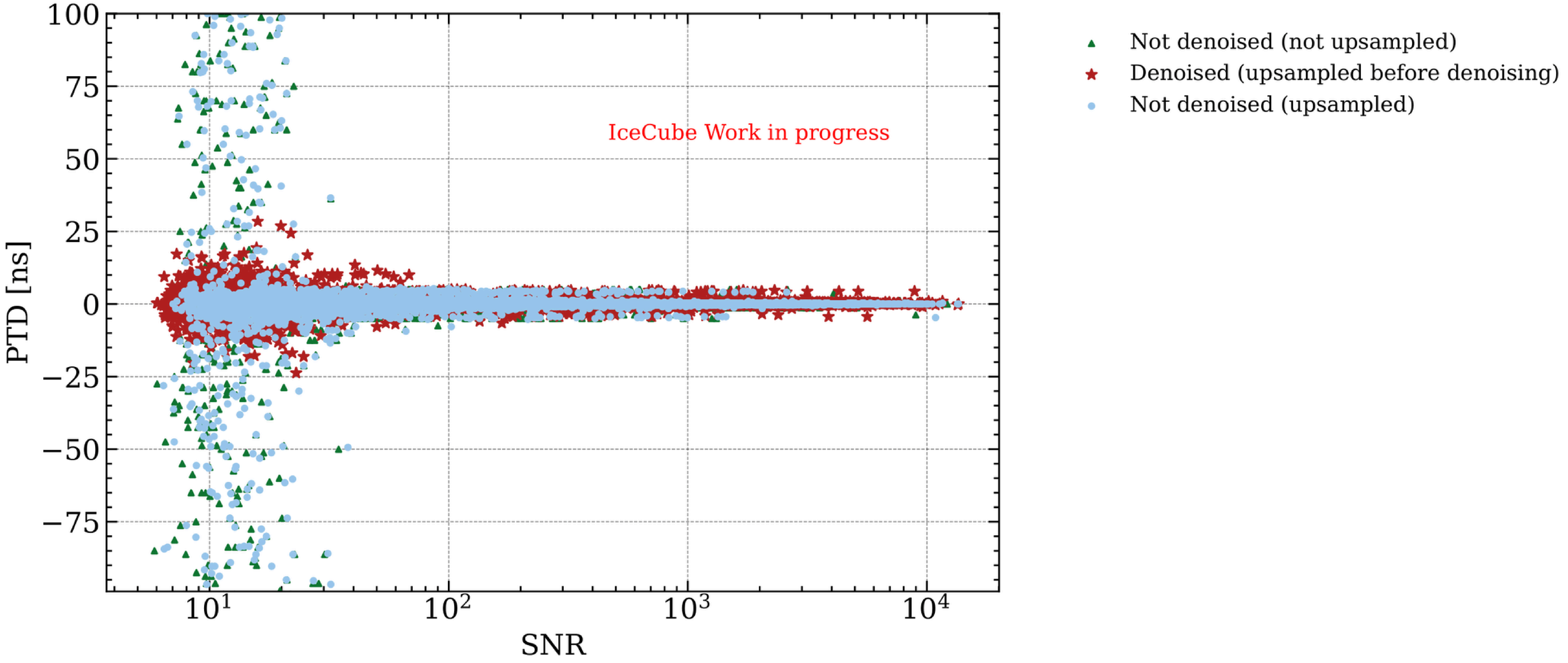


RESULTS CLASSIFIER



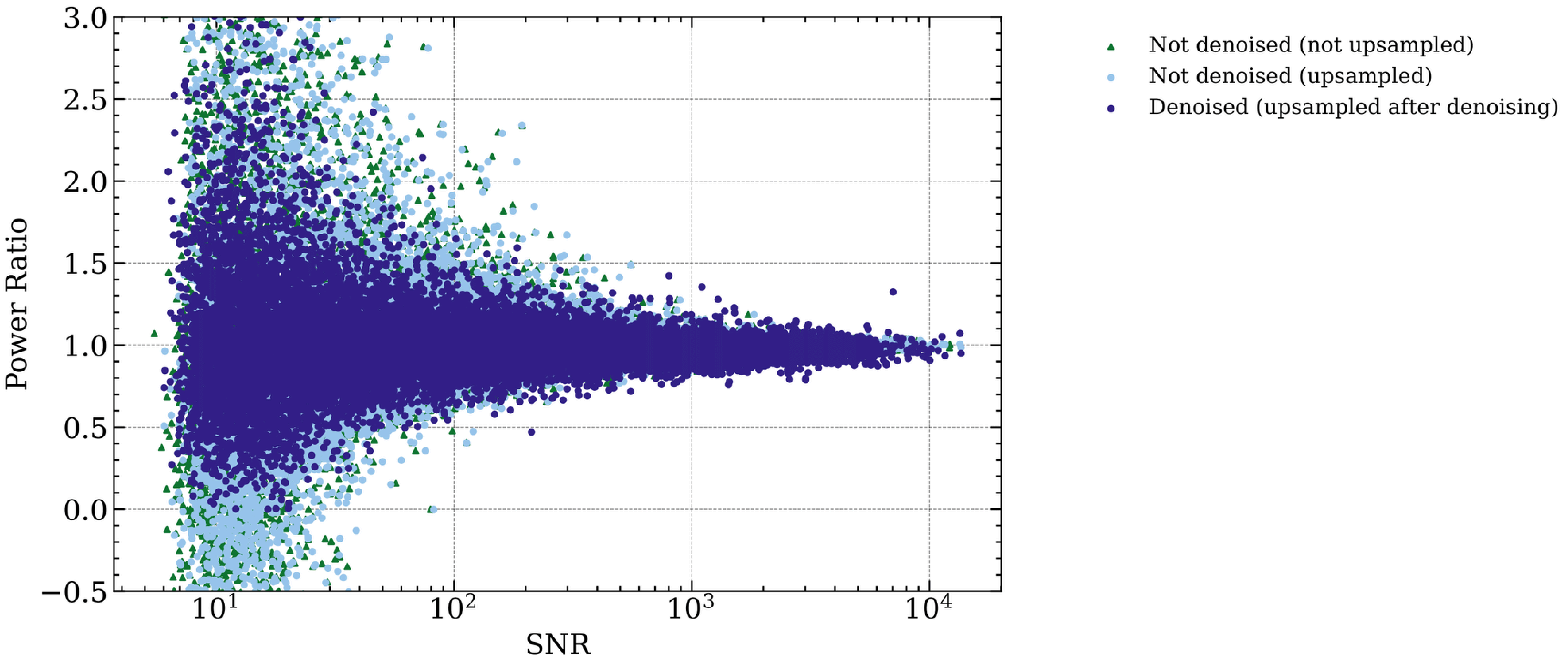
100
75
50
25
0
−25
−50
−75
PTD [ns]
10^1
10^2
10^3
10^4
SNR
IceCube Work in progress
Not denoised (not upsampled)
Denoised (upsampled after denoising)
Not denoised (upsampled)

PTD SCATTER

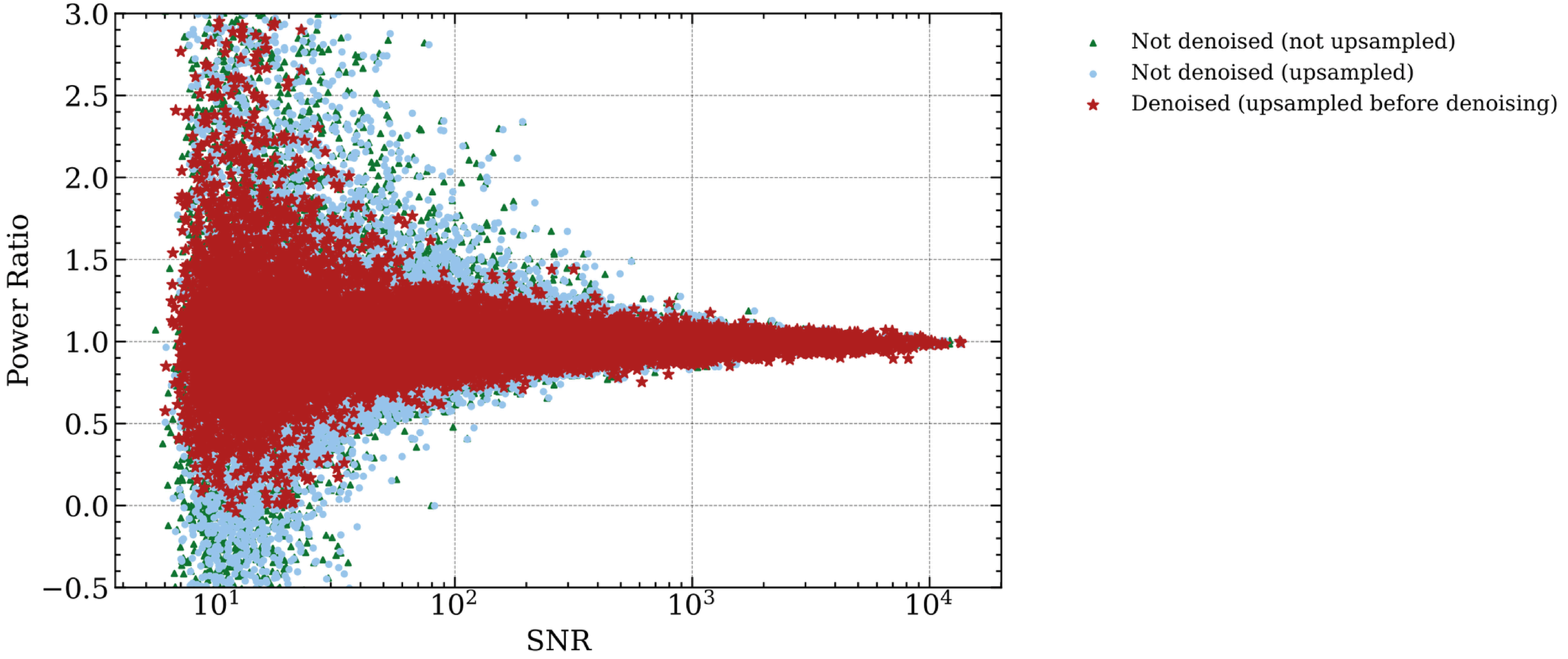


IceCube Work in progress
Not denoised (not upsampled)
Denoised (upsampled before denoising)
Not denoised (upsampled)
PTD [ns]
SNR

PTD SCATTER

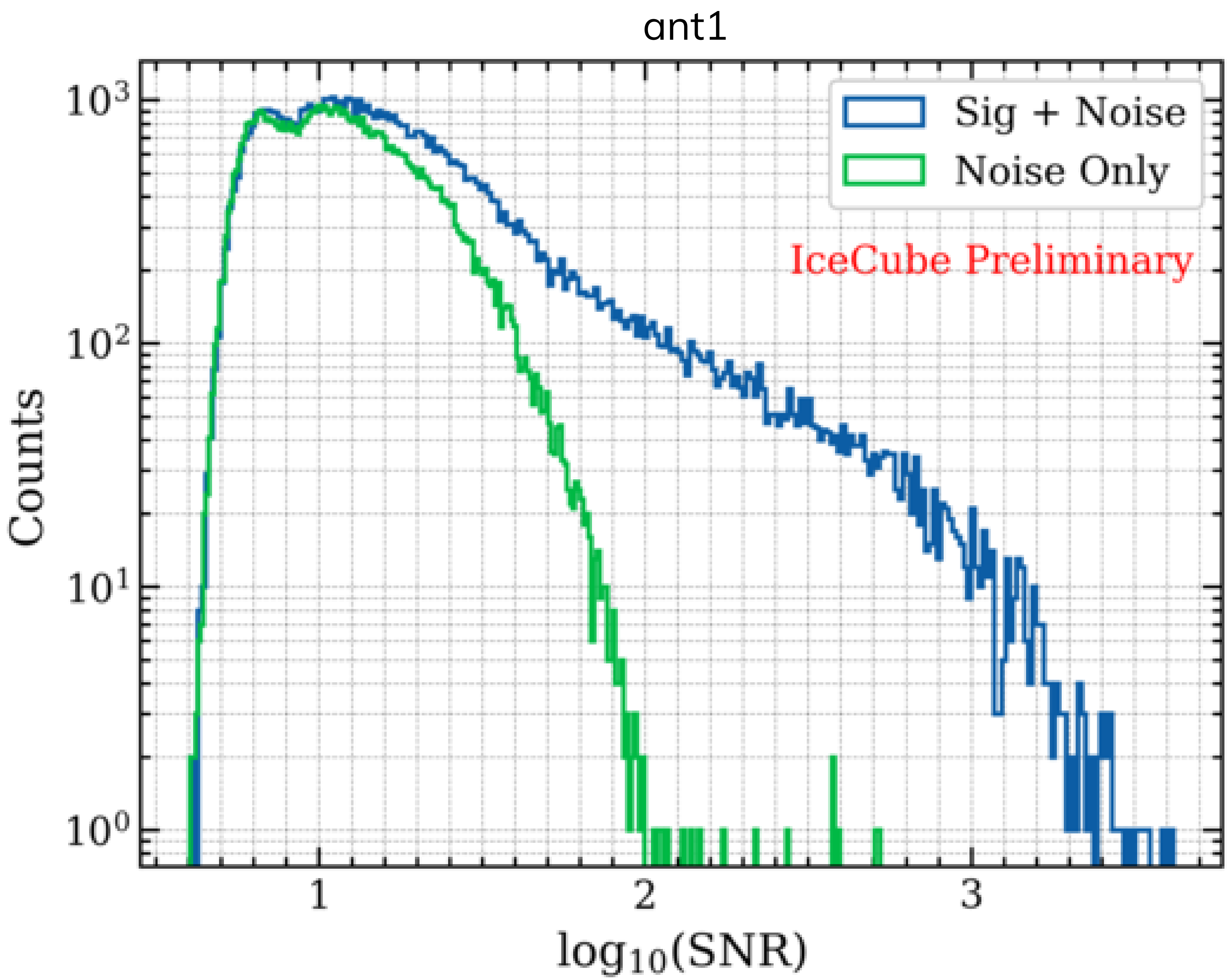

Not denoised (not upsampled)
Not denoised (upsampled)
Denoised (upsampled after denoising)
Power Ratio
SNR
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
10^1
10^2
10^3
10^4

PR SCATTER

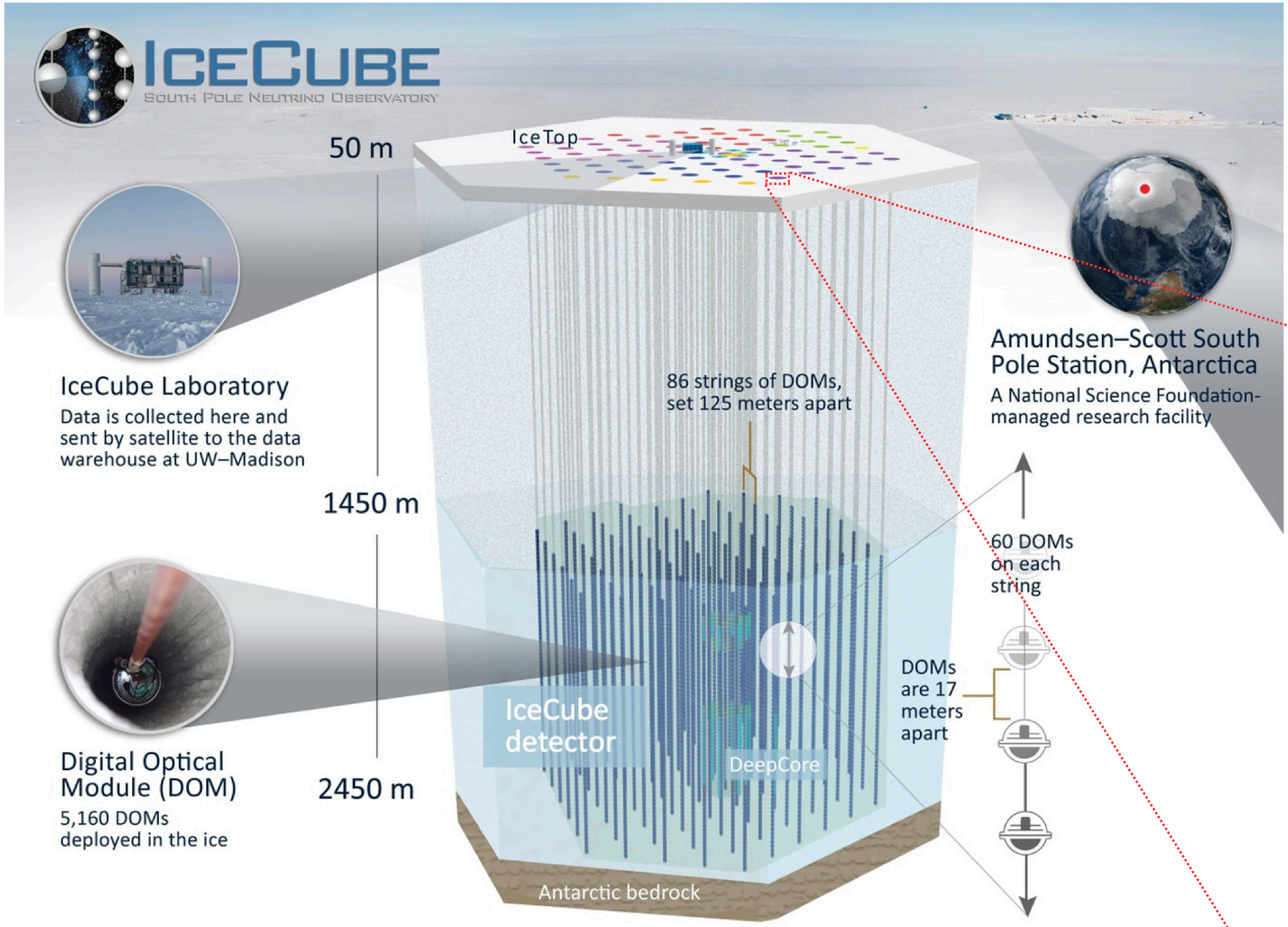


Power Ratio
SNR
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
10^1
10^2
10^3
10^4
Not denoised (not upsampled)
Not denoised (upsampled)
Denoised (upsampled before denoising)

PR SCATTER


ant1
Sig + Noise
Noise Only
IceCube Preliminary
Counts
log10(SNR)

PR SCATTER


ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY
IceTop
50 m
IceCube Laboratory
Data is collected here and
sent by satellite to the data
warehouse at UW–Madison
86 strings of DOMs,
set 125 meters apart
Amundsen–Scott South
Pole Station, Antarctica
A National Science Foundation-
managed research facility
1450 m
60 DOMs
on each
string
DOMs
are 17
meters
apart
IceCube
detector
DeepCore
Digital Optical
Module (DOM)
5,160 DOMs
deployed in the ice
2450 m
Antarctic bedrock

ICECUBE