COMPUTATIONAL IMAGING

ACTIVITY 03

COMPRESSIVE SENSING FOR AUDIO SIGNALS

SUBMITTED BY:
RENE L.PRINCIPE JR.

SUBMITTED TO:
DR. MARICOR N. SORIANO

Background

An original signal, say an audio waveform $\boldsymbol{x}$ can be represented by a sparse vector **s** and a basis transformation $\boldsymbol{\Psi}$. Essentially,

$$\boldsymbol{x} = \boldsymbol{\Psi}\,\boldsymbol{s}.$$

In this activity, we exploit the sparsity of audio signals at low frequencies to recover the the original signal in a technique called compressive sensing. Instead of uniformly sampling the audio signal, we generate a measurement matrix $\boldsymbol{C}$ through random sampling measurements $\boldsymbol{y}$, hence

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} = \boldsymbol{C}\boldsymbol{\Psi}\boldsymbol{s}.$$

The sparse vector can be solved through minimization given by

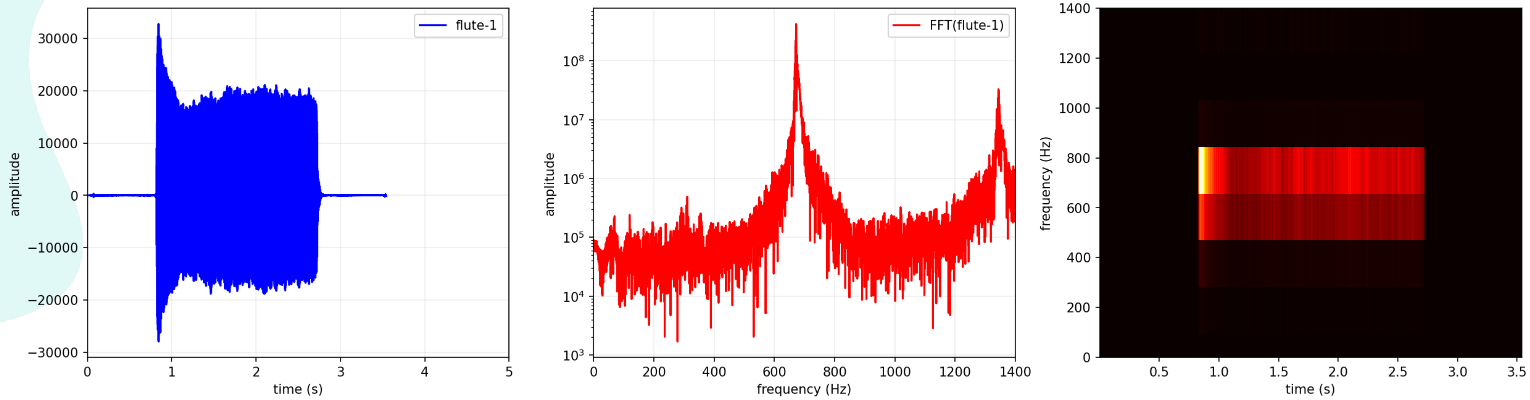
$$\hat{\boldsymbol{s}} = argmin\|\boldsymbol{s}\|_1,$$

where $\|\cdot\|_1$ is the l_1 norm given by $\|\boldsymbol{s}\|_1 = \sum_{k=1}^{n}|s_k|$.

Solving for $\boldsymbol{s}$ and since $\boldsymbol{\Psi}$ is a known basis, the original signal $\boldsymbol{x}$ can be recovered using the first equation.

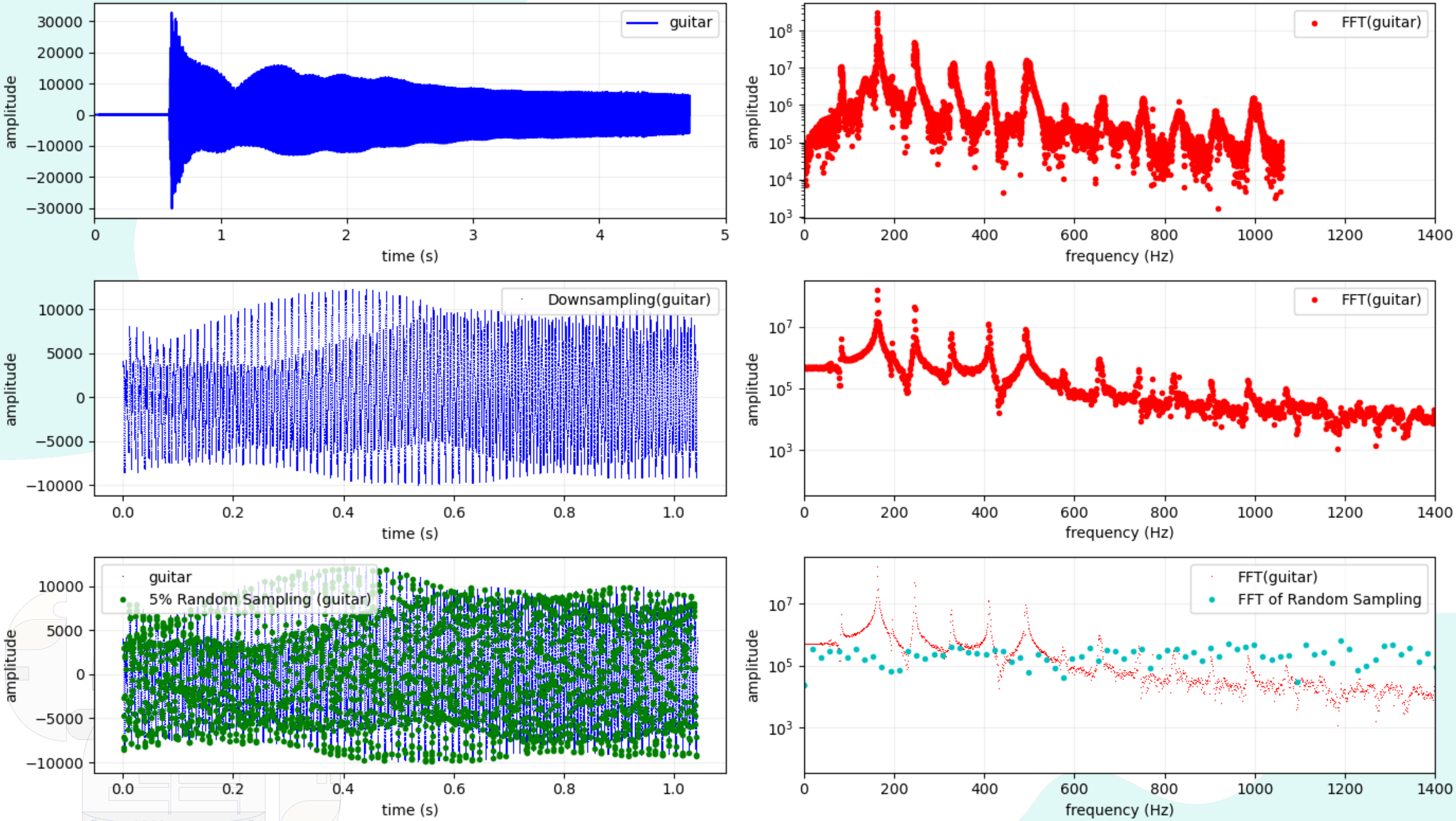
Overview

Shown below is an audio signal, its Fourier transform, and the spectrogram. Notice how there's only two distinct peaks in its frequency space. Some, if not most, of these other frequencies can be discarded and it still preserves the characteristics of the original signal.

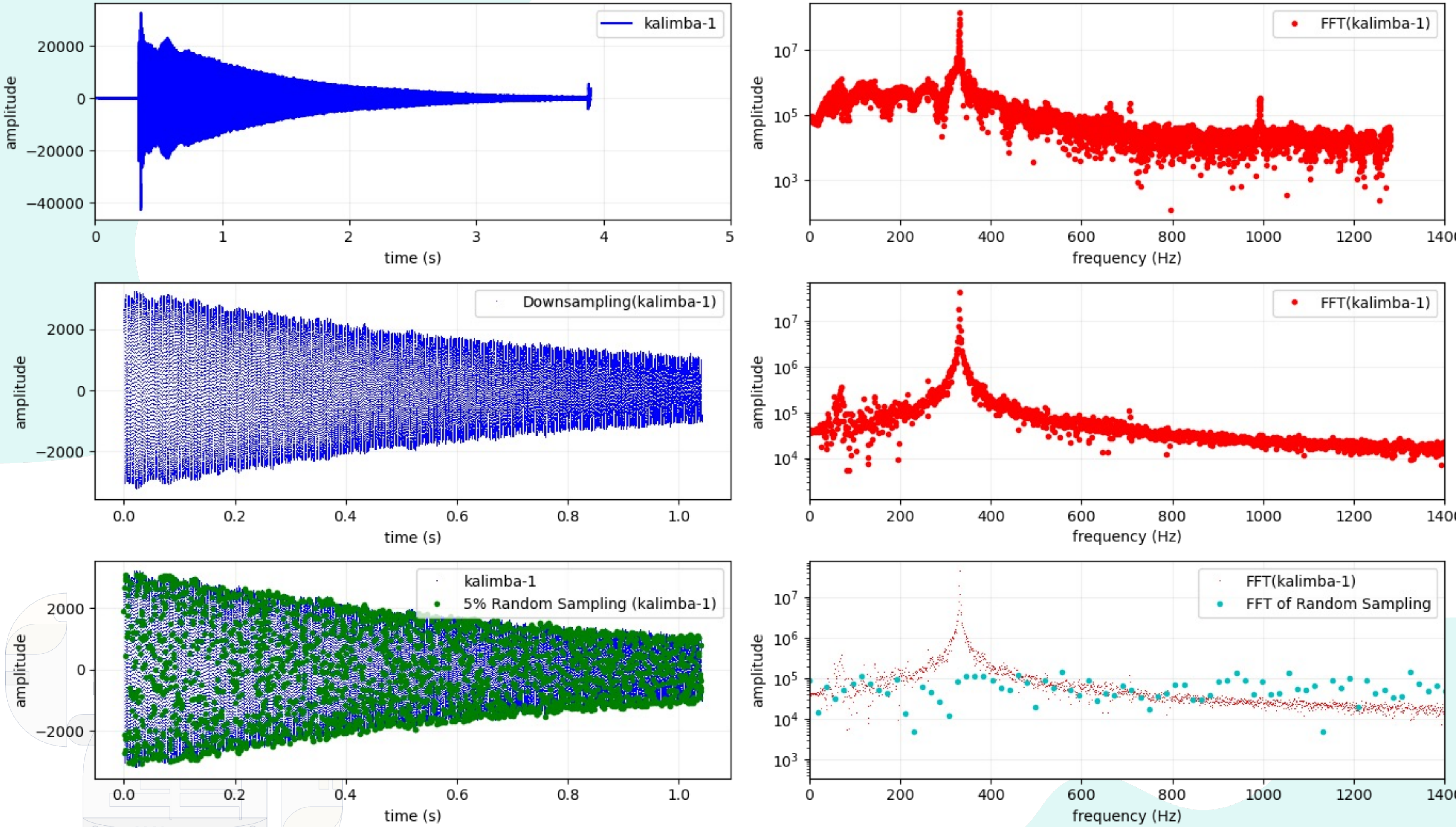


In this activity, Guitar, Kalimba, and Flute audio signals (files from Reinier Ramos) were randomly sampled at 5%, 10%, 20% and 50%. A Discrete Cosine Transform (DCT) was used as the basis transformation. The lasso optimization algorithm then estimates the missing samples for reconstruction. The original and reconstructed signals were compared in real and frequency space.

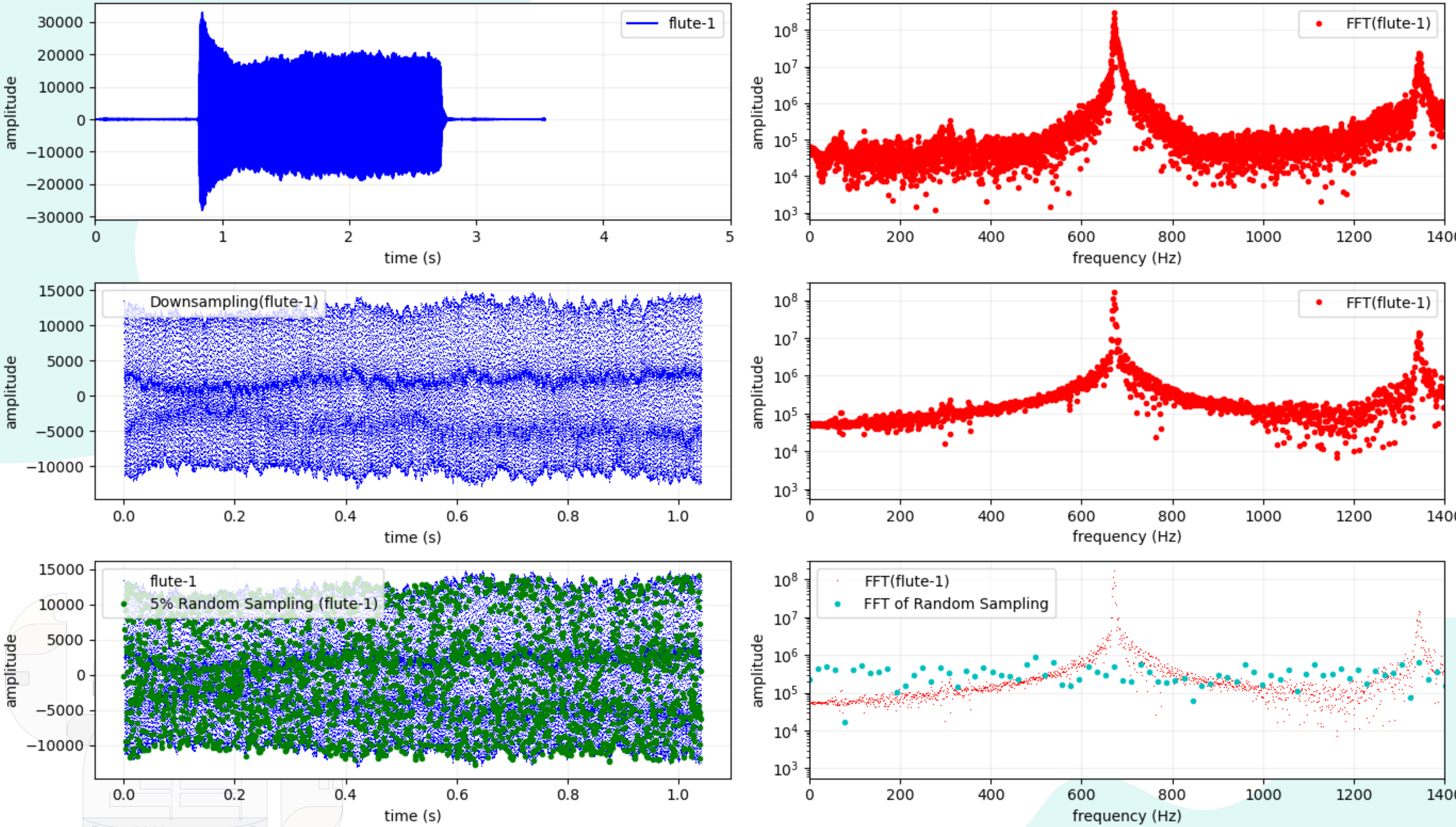
Guitar


guitar
amplitude
time (s)
FFT(guitar)
frequency (Hz)
Downsampling(guitar)
5% Random Sampling (guitar)
FFT of Random Sampling

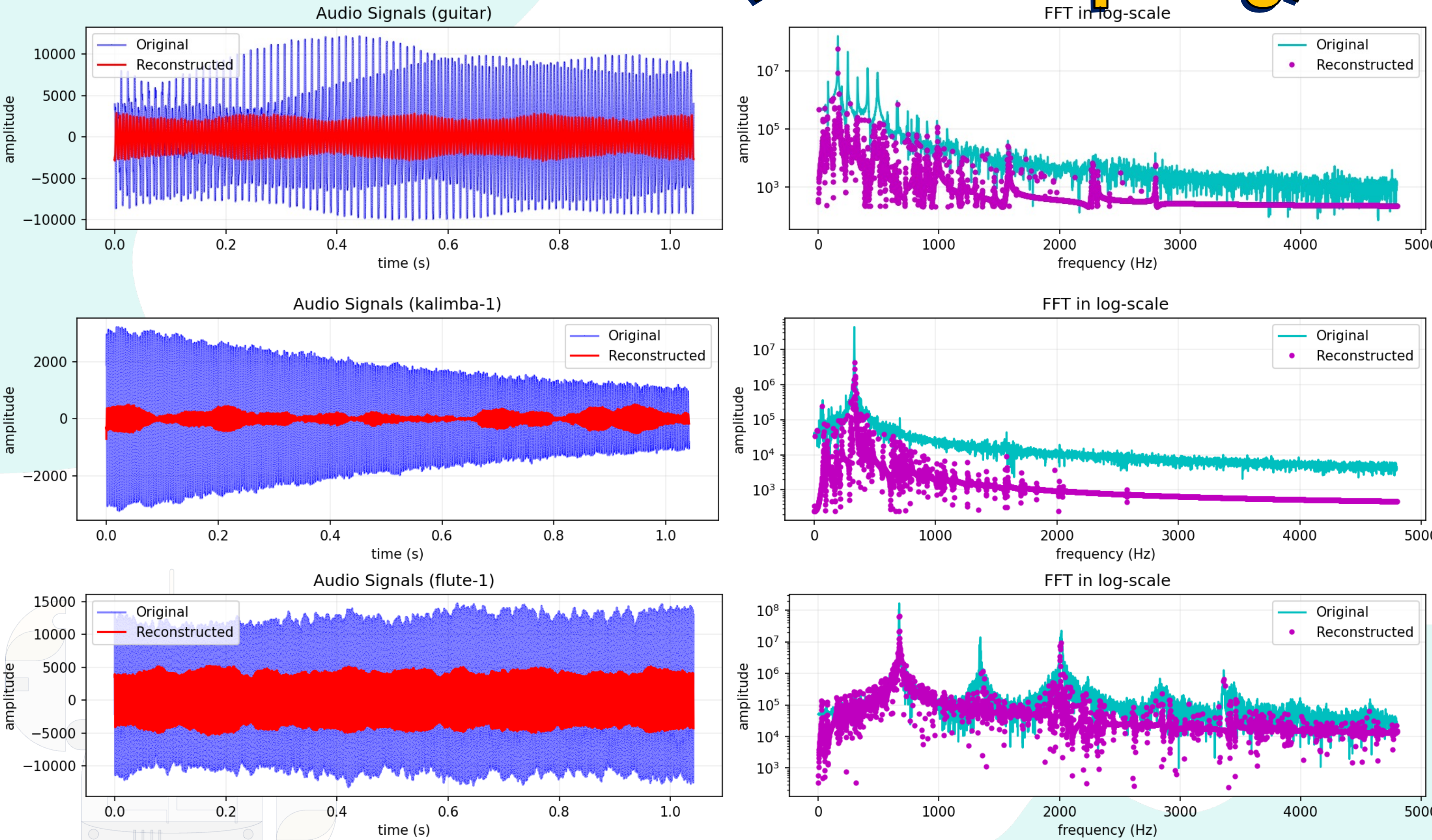
Kalimba


kalimba-1
amplitude
time (s)
FFT(kalimba-1)
frequency (Hz)
Downsampling(kalimba-1)
5% Random Sampling (kalimba-1)
FFT of Random Sampling

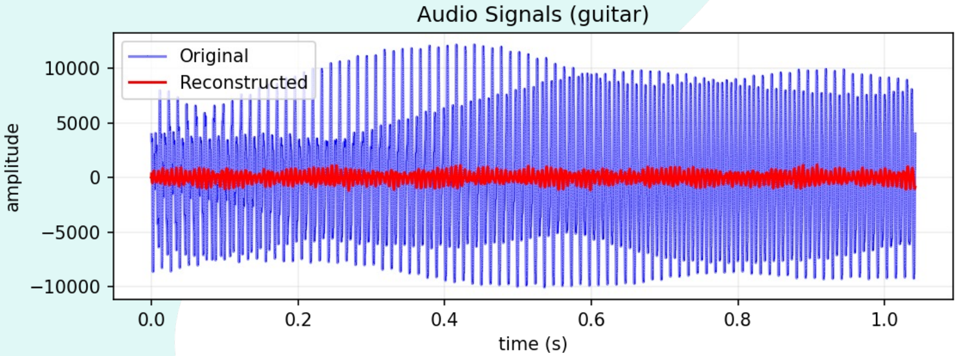
Flute

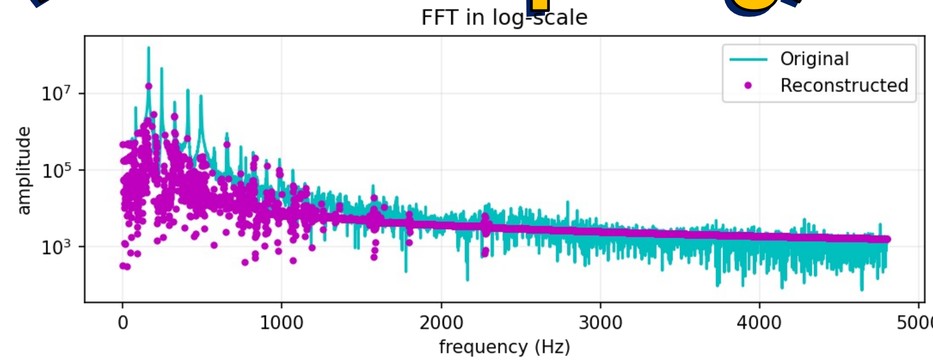

flute-1
amplitude
time (s)
FFT(flute-1)
frequency (Hz)
Downsampling(flute-1)
5% Random Sampling (flute-1)
FFT of Random Sampling

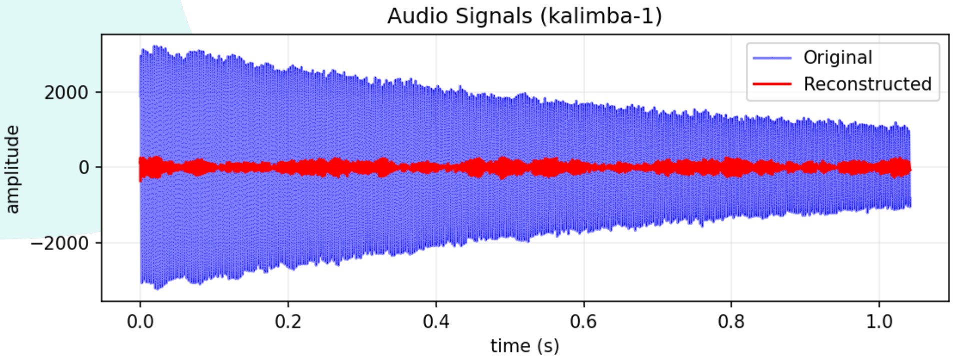
Reconstruction (5% Sampling)

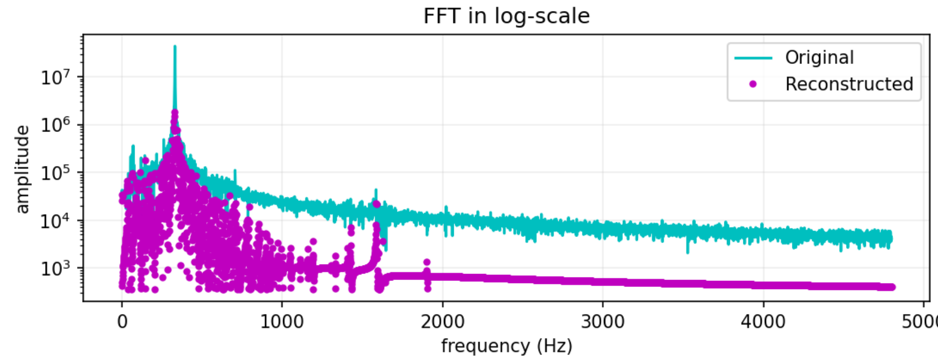


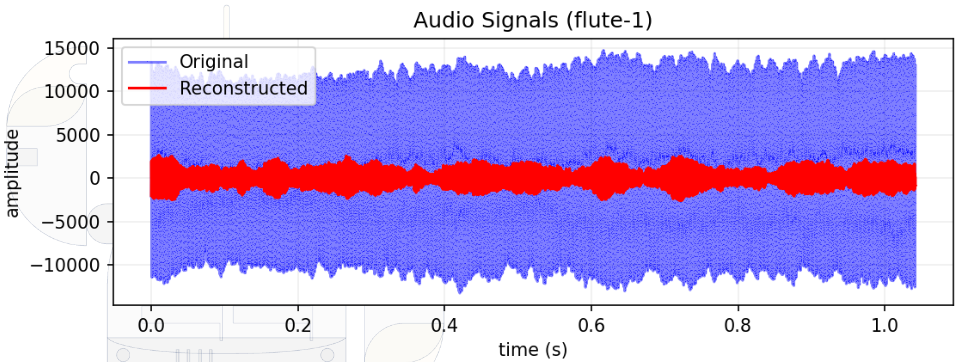
Reconstruction (10% Sampling)

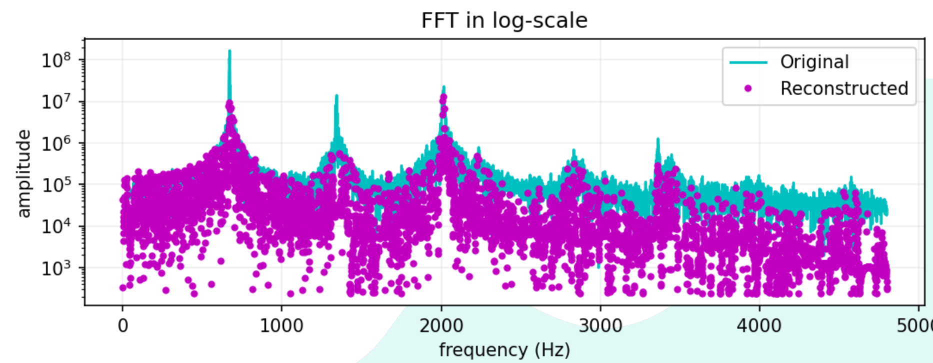




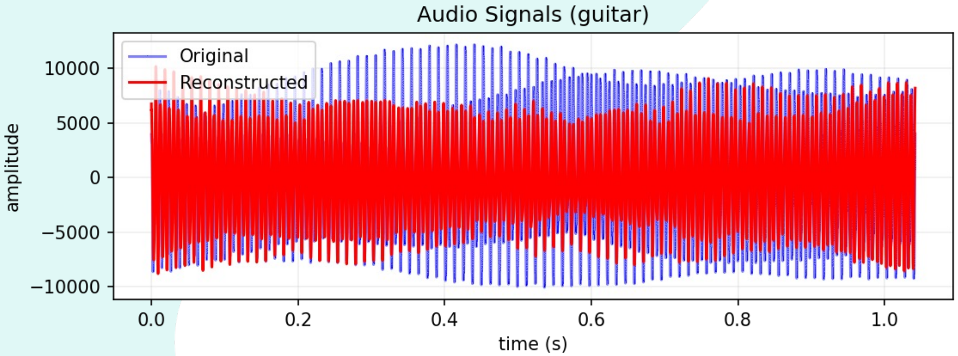


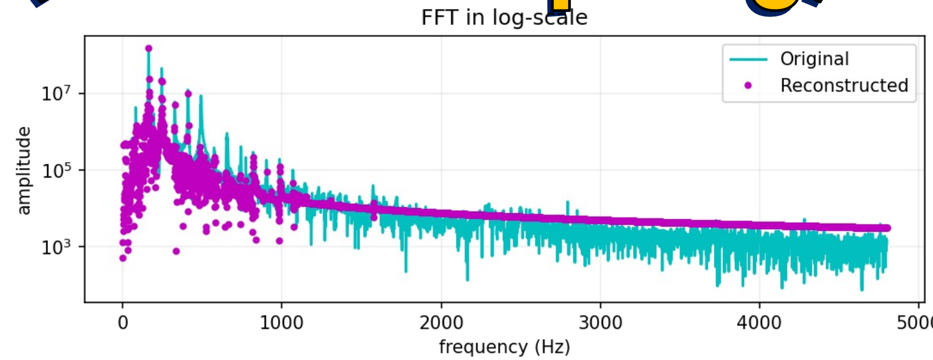


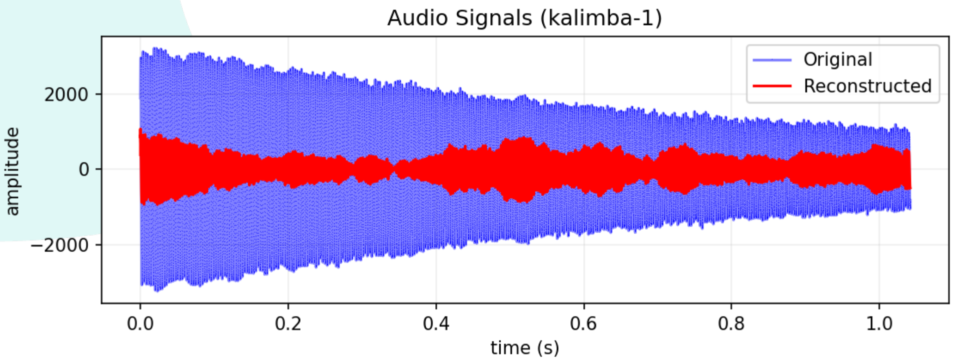


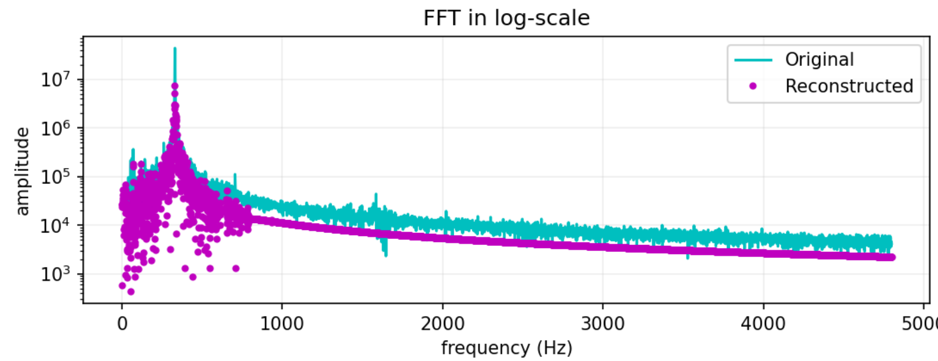


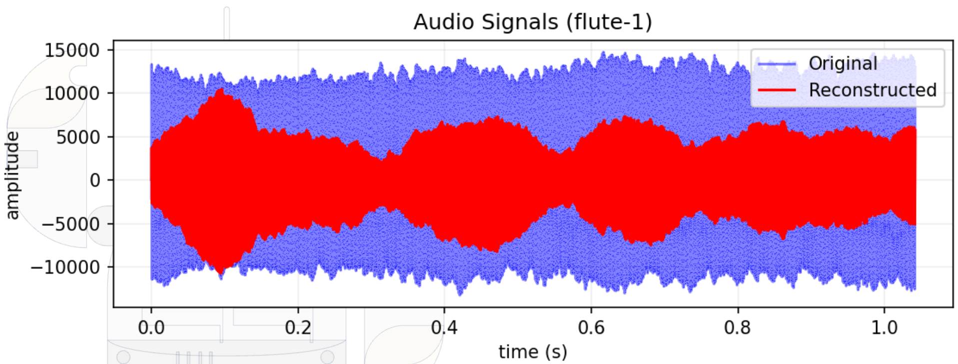
Reconstruction (20% Sampling)

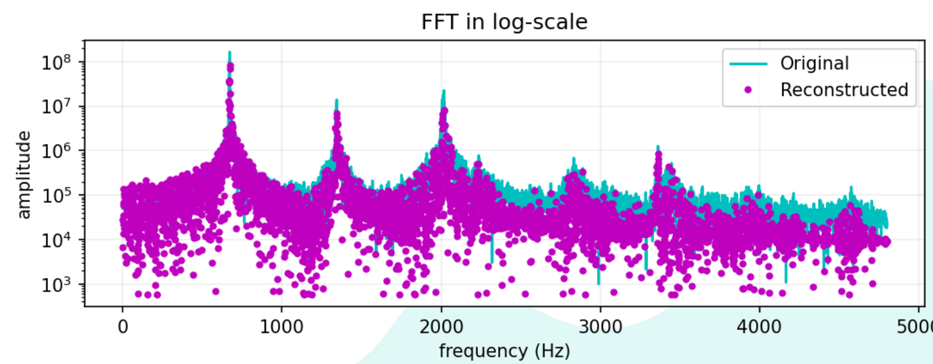




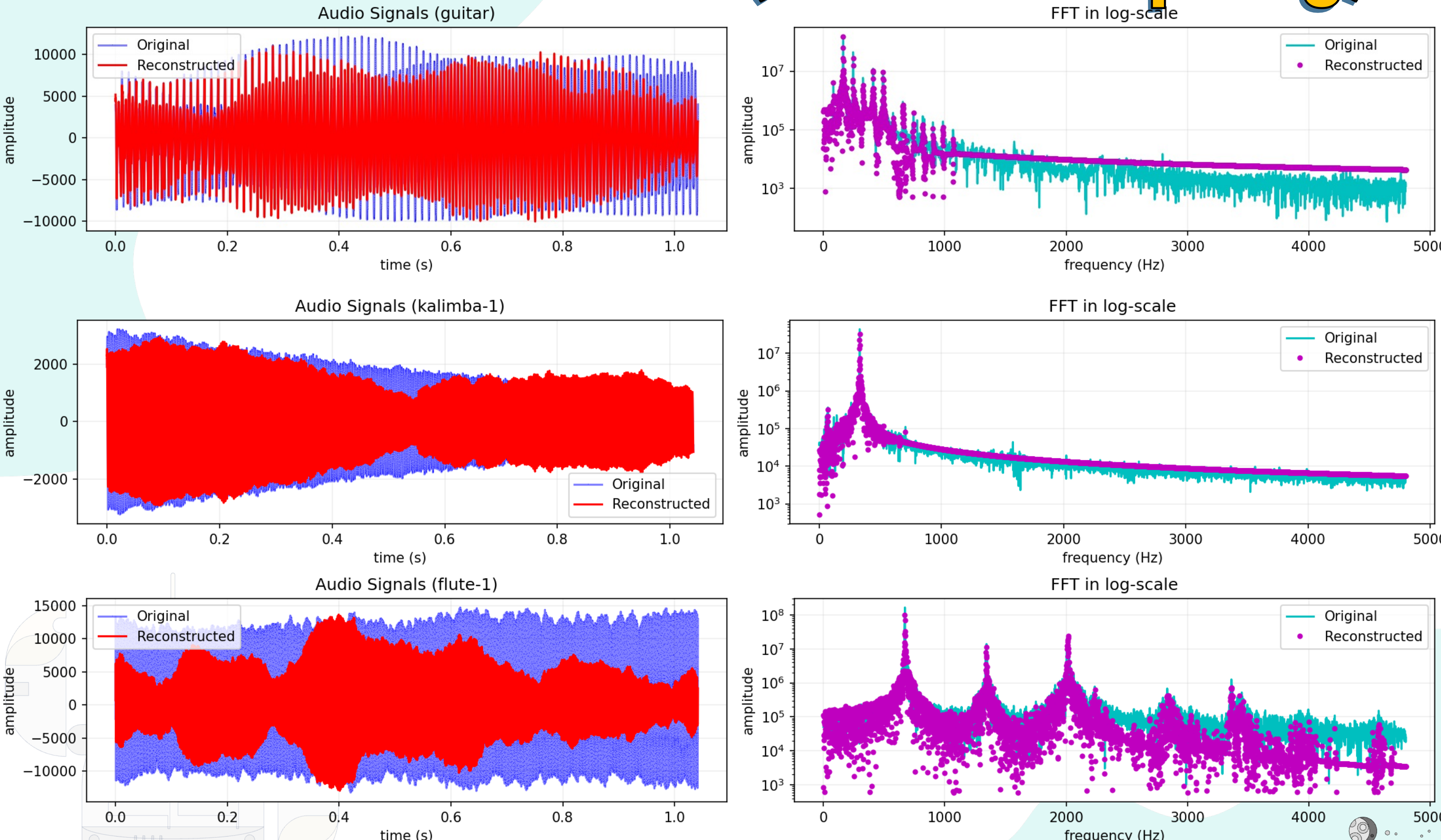








Reconstruction (50% Sampling)



Discussion

Compressive sensing as a concept is very interesting as I see it as counter-analogous to the super-resolution techniques explored in previous activities. In this activity, a raw audio signal from an instrument was resampled and randomly sampled at varying rates. Using a known and established DCT as basis, the signal was reconstructed, and the lasso optimization algorithm was used.

By viewing the reconstructed signal in frequency space, the reconstruction using a very small number of measurements (even at 5%) proved to be sufficient as it still retains the most prominent frequencies.

reflection

The concept is very straightforward as I can see analogies between principal components analysis (PCA) which I have extensively studied and compressive sensing. The difference is that DCT is a widely used an universal basis as opposed to PCA's custom basis given a training data. The possibility that a 5% sampling can restore or mimic the original signal seems like magic and seeing how the prominent frequencies were retained really amazed me. I would like to acknowledge Reinier Ramos for giving me a copy of the audio signals, and (together with Janelle Manuel, Lemuel Saret, Kenneth Domingo, and ChatGPT) for allowing me to use their codes as a reference. A sufficient set of code, results, and discussion was presented in this activity; however, I genuinely would want to explore and elaborate on this in the future… if time permits. #ThesisIt :'(

With that said, I'd give myself a score of **100/100.**

references

- [1] M. Soriano, Physics 305 – Compressive Sensing, (2023).

SOURCE CODE

https://github.com/reneprincipejr/Physics-305/tree/main/Activity%203%20-%20%20Compressive%20Sensing%20for%20Audio