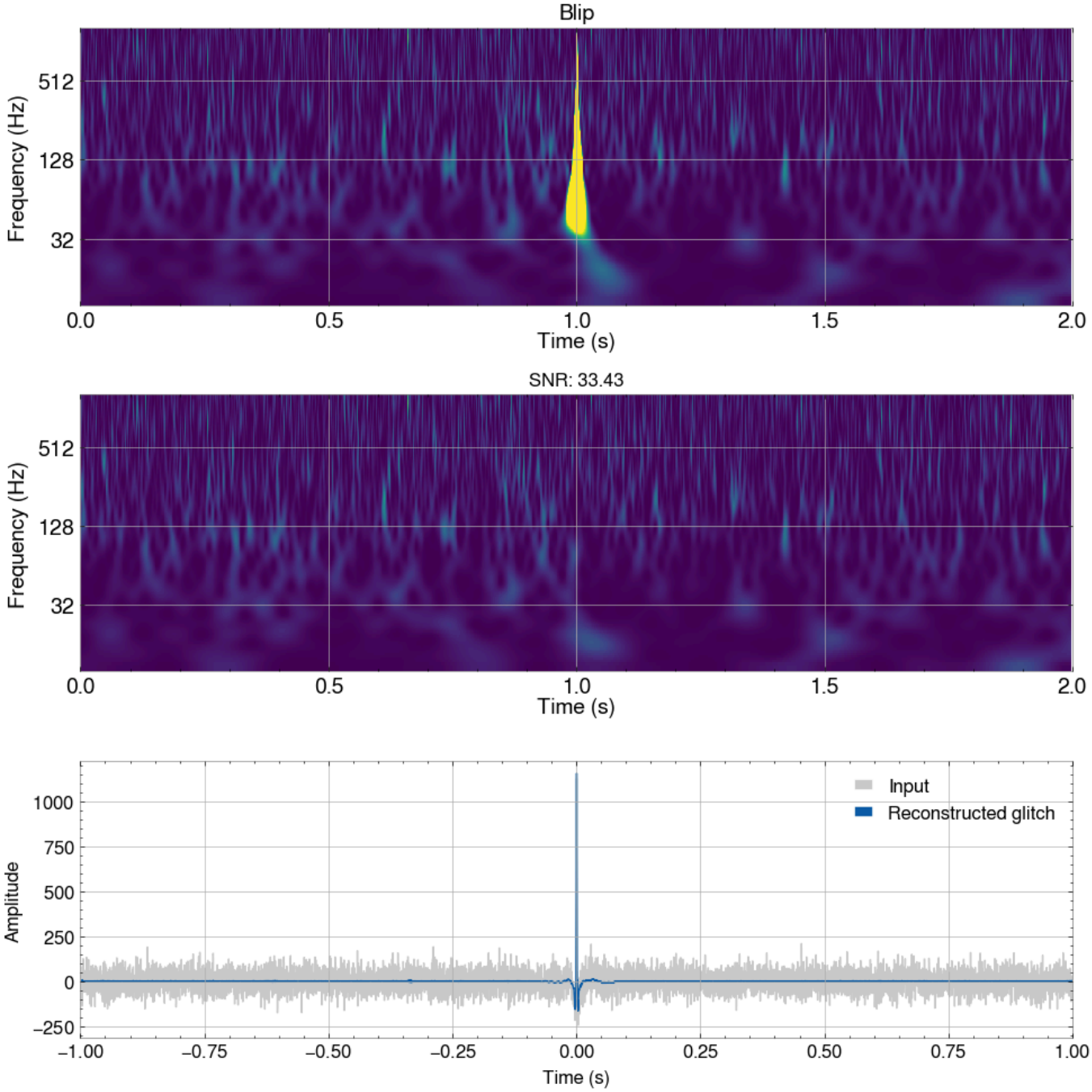
Unsupervised clustering of GW glitches

Glitches are short transient noise bursts that appear in gravitational wave detectors. While many efforts have been conducted to characterize and mitigate them, many questions remain unanswered. The project Gravity Spy [1] is a machine-learning and citizen-science-based pipeline that has identified ~20 different types of glitches, classifying them based on the appearance of their spectrograms. While these classes are helpful, the glitch population can often change over time: new classes can appear while others disappear. Spectrograms also do not retain phase information, and it is still unclear whether this information could improve glitch characterization.

Efforts have been made to develop an unsupervised classifier of glitches that can capture small differences beyond the 20 classes defined by Gravity Spy [2]. However, these approaches also rely on spectrogram-like inputs, which again discard phase information. The DeepExtractor [3] framework, on the other hand, allows for cleaning gravitational-wave detector data by modeling

glitches and subtracting them from the data stream in the time domain. This makes it possible to access the phase of the noise in unprecedented detail.



The goal

The goal of this project is to obtain glitch time-domain waveforms through DeepExtractor and apply clustering algorithms to identify potentially new classes of glitches. Such an approach could provide a fast way to diagnose detector status and discover glitch classes even during ongoing observations, thereby helping instrument teams to mitigate their causes.

[1] inspirehep.net/literature/2142630 (Gravity Spy Project)
[2] arxiv.org/pdf/2412.16796 (t-SNE for glitches)
[3] inspirehep.net/literature/2874224 (Deepextractor)

Data

Deepextractor github : git.ligo.org/tom.dooney/deepextractor/
Gravity spy datasets : zenodo.org/records/5649212