Using machine learning to gain information about glitches in LIGO strain data

Melissa Kohl

May 3, 2018

1 Introduction

In the Advanced LIGO observation runs, detection of gravitational waves has occurred for multiple black hole mergers and one neutron star collision [1]. Although the sensitivity of the detectors is enough to detect gravitational waves from these huge astronomical events, gravitational waves are also created from other astronomical events. To detect those gravitational waves, the sensitivity of the detectors needs to be better. The strain data from both the science and observation runs of the Advanced LIGO detectors contain "glitches" that can mimic gravitational waves and obscure real gravitational waves [1]. If these glitches were eliminated, the sensitivity of the detectors would improve, making detection of gravitational waves from astronomical events other than black hole and neutron star mergers (including smaller black hole or neutron merger events) possible.

2 Classification of Glitches

To eliminate glitches, they must first be identified. The easiest way to identify a glitch is by looking at the same time frame at another detector that is in a different geographic location. Although the glitches can be eliminated from the data, they then must be classified so the source of the glitch can be identified and eliminated for future observing runs [2].

A group of scientists created a machine learning software package called GravitySpy to aid in classifying the monstrous amounts of glitches [1]. Unlike previous machine learning techniques used on LIGO data that compared the waveforms of the glitches [2], GravitySpy's neural network uses spectrograms from four different time frames to create a multi-layer network that utilizes image classification techniques. Since different types of glitches have different durations, the multiple views not only provide complementary data across time frames, but also allow for identification of a broader group of glitches of different durations [3].

GravitySpy is great at classifying glitches into known classifications, but has trouble identifying new classifications [1]. The output function in GravitySpy's neural network is *softmax* [3], which essentially just classifies the input glitch into the classification with the highest correlation, regardless of how high that correlation is. In typical LIGO Collaboration-style, the task of designating new classifications is therefore largely reliant on volunteers. To date, volunteers have identified two new glitch classifications [1]. Although human sorting is still quite inefficient, even with a large amount of people, GravitySpy is overall largely successful in classifying glitches.

3 Identification of Sources of Glitches

Classification of the glitches themselves is only one half of the problem with glitches in the strain data. Once glitches are classified, the source must be identified so the glitches can be eliminated from future observing runs. A fair amount of glitches have been eliminated by locating the source of the glitch, but many glitches, including the most frequent glitch, referred to as a "blip," still have unidentified sources.

Each LIGO detector has hundreds of channels of information keeping track of environmental and instrumental data that could contain the sources of glitches. LIGO collaborators have used machine learning to match the waveforms of the glitches in the strain data to waveforms in the individual channels [1], but there are too many channels for this technique to efficiently find glitch sources.

4 Project Ideas

One of the main goals related to glitches is identifying the source of blip glitches. Since they are so frequent, there is a strong possibility that there are different types, i.e. subclasses of blip glitches. By creating a new training set to classify existing blip glitches into subclasses,

5 References

1. Zevin, M. *et al* 2017 *Class. Quantum Grav.* **34** 064003

2. Mukherjee, S. *et al* 2010 *J. Phys.: Conf. Ser.* **243** 012006

3. Bahaadini, S. *et al* 2017 *International Conference on Acoustics, Speech and Signal Processing*