Digging deeper: finding sub-threshold compact binary merger events in LIGO data

LIGO Caltech SURF Program 2019 Project Proposal

Author: Sierra Garza[1,2]

Mentors: Jonah Kanner[2], Alan Weinstein[2], Liting Xiao[2]

[1]*California State Polytechnic University, Pomona, CA 91768, USA*

[2]*LIGO, California Institute of Technology, Pasadena, CA 91125, USA*

(Submitted on 7 July 2019 (v1), last revised July 10 2019 (this version, v3))

INTRODUCTION / BACKGROUND:

The LIGO and Virgo detectors have collected gravitational wave (GW) data from three separate observation runs since 2015, with the third run presently collecting data. There have been 10 signals from binary black hole (BBH) mergers and one binary neutron star (BNS) merger detected from the first two observation runs and several more from the third run. These detections were all confirmed due to high confidence in their signal-to-noise ratio (SNR); however, there are likely many more unconfirmed signals in the data due to lower SNRs. A limitation in the SNR criteria arises when accidental coincidence of "loud" glitches or other rare noise fluctuations in the two LIGO detectors can result in high SNRs but are not the product of real GWs **[1]**.

Initial PyCBC **[2]** and GstLAL**[3]** pipeline searches have flagged potential GW events as triggers based upon matched-filtering and threshold values **[4]**. Since LIGO-Virgo's search sensitivity scales with binary black hole mergers primary component mass, approximately $V \propto M_1{}^{2.1}$, higher mass mergers have been easier to detect than those of lower mass **[5]**. Current compact-binary-coalescence (CBC) searches do not adequately differentiate GWs of lower confidence from inherent noise. Bayesian model comparison of coherence may be a way to effectively discriminate whether a marginal trigger, a flagged event just below the SNR from the pipeline searches, is likely a GW signal or instrumental noise from the detectors.

Coherence requires that the strain signals in multiple instruments share a phase evolution consistent with a single astrophysical source, represent a well-described CBC waveform, and be temporally coincident **[6]**. Instrumental noise transients (glitches) are not expected to fully meet these requirements, whereas GWs are; therefore, allowing the distinction to be made.

The data from these signals contain valuable untapped information to understanding gravitational waves and the characteristics from merging black hole and neutron star systems in the distant universe. Being able to collect and analyze the data from the weaker signals will help fill in the gaps of our understanding of CBC populations. Confirming numerous GW events may allow us to construct a diagram depicting the evolution of these populations, similar to the Hertzsprung-Russell diagram for stellar evolution.

OBJECTIVES / METHODS:

I will be working on analyzing whether the Bayesian coherence ratio (BCR) can be a reliable method to improve confidence in the signal of marginal triggers and/or improve the rejections of inherent noise in the data in conjunction with current pipeline searches.

A BCR compares the odds between the hypotheses that the data comprise coherent CBC signal or incoherent instrumental features. This calculation factors in the evidence of the signal in pure gaussian noise, non-gaussian noise fluctuations that are not coherent between detectors, and the coherent signal in gaussian noise. Priors, such as the BBH merger rate and mass range, will be determined through trial and error and be used to compute the evidence parameters in the BCR calculation. The two weighted terms (α and β) in the BCR calculation are priors that affect the separation between the background and foreground populations and will be estimated to produce useful BCRs. A predetermined surrogate glitch hypothesis will be used and an assumption for the gaussian noise evidence is of a perfect measurement of the detector noise power-spectral-density (PSD) **[6]**.

$$P(\mathbf{w} \mid D, \mathcal{H}_i) = \frac{P(D \mid \mathbf{w}, \mathcal{H}_i) P(\mathbf{w} \mid \mathcal{H}_i)}{P(D \mid \mathcal{H}_i)} \qquad (1)$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}. \qquad (2)$$

Equation 1: Showing Bayes' Theorem. w represents the parameters of H_i (hypothesis). D represents the data.

Equation 2: Shows simplified version of Bayes' Theorem. Analogous to Equation 1.

We want to calculate the BCR for triggers produced in all observing runs to detect weak GW signals and potentially define empirical probability distributions that would allow us to obtain likelihood ratios to use for trigger classification.

The BCRs will be computed with the likelihood values calculated from running Bilby **[7]** jobs on the LIGO cluster. We will run hundreds of Bilby jobs on injected signals, glitches, confirmed signals, and marginal triggers analyzed with an astrophysical and respective noise models for both Advanced LIGO detectors.

We will compute and apply the BCR to Observation 2 and 3 background triggers in effort to reject any glitches. This will allow an updated false alarm rate (FAR) to be used for Observation 3 event analysis. A BCR < 1 would allow for the rejection of that trigger as a GW event because it would favor the odds of the hypothesis that the data is comprised of incoherent instrumental noise.

Various plots will be produced to visualize the results from using the BCR to compare real and simulated signals to inherent noise. I will utilize Thomas Alford's python code and Bilby tutorials to produce plots in Figure 2, following the discussion in **[6]**.

I will now work on determining appropriate priors, such as the weights (α and β) which minimize the overlap between the signal and noise trigger **[6]** to calculate BCRs such that signal and noise are separated effectively. We will also need to determine a signal duration limit that we want to initially evaluate and providing us with a chirp-mass range. We will use standard ranges to restrict parameters on mass ratio, spin magnitude, luminosity distance, etc.. By identifying any likely GW events and/or rejecting any likely triggers due to instrumental noise, will allow us to calculate a new FAR that we will apply to Observation 3 events, improving our confidence in those events. The final step for this project would be to select new parameters, signal time duration limit, and compute BCRs for Observation 3 triggers. After running through the same sequence of steps, we can work on making any improvements/adjustments to method of distinguishing signal from noise.

PROGRESS:

For this project I am developing a thorough understanding of Bayesian inference and parameter estimation and its potential applications to GW research. I have also become familiar with Alford's python code; which has allowed me to understand how to utilize LIGO data and set up to run multiple Bilby jobs on LIGO's computer cluster.

I investigated if there are any overlapping trigger times from several published sub-threshold trigger catalogs, such as GWTC-1 **[4],** 1-OGC **[8]**, two from Princeton **[10-11]**. I extracted the intervals of data that surrounded the trigger times listed in the various catalogs and write corresponding hdf5 files. This was done after first converting all published trigger times into GPS time, the time in seconds from January 06, 1980. For each catalog's list of triggers, I appended them to a list and sorted them from lowest to highest trigger times to discern if there were any overlapping triggers within a catalog. To make this determination, I compared whether the ending time of the first trigger was greater or equal to the starting time of the following trigger. I used a loop to iterate through each list. I then compared multiple catalogs for overlapping trigger times by appending the sub-lists into one list and resorting it so that the triggers were ordered lowest to highest GPS time. By running the new list through the same loop, I was able to determine no overlapping trigger times from the four sub-threshold catalogs, given a specified time range of $\pm$ 0.1 s surrounding the published trigger times. However, I am expanding my specified time range to identify any sub-threshold triggers that are reasonably close in time, given that multiple sub-threshold catalogs utilized different pipeline searches to produce their respective catalogs.

In python, I have produced time-frequency plots (Figure 1) for various trigger times to visualize how the amplitude spectral density (ASD) in the data evolves over specified time intervals.

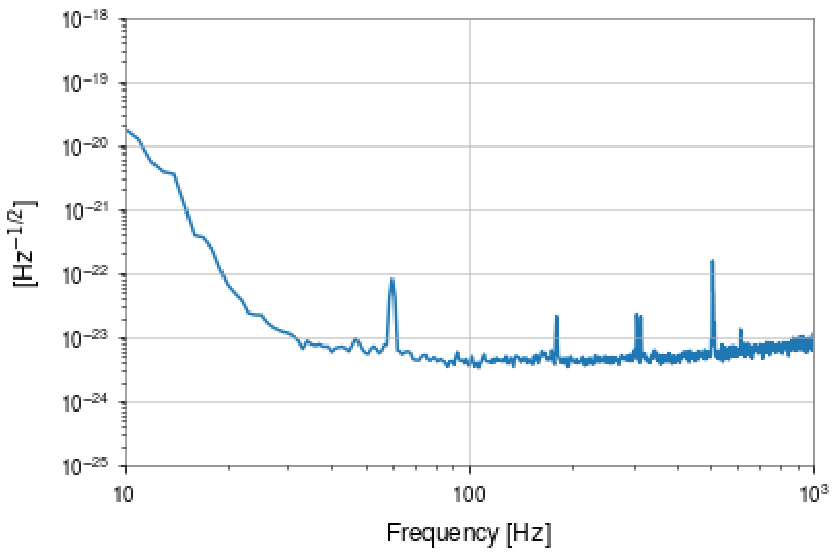

Figure 1: Time-frequency plot of the amplitude spectral density (ASD) over a GPS time frame (1242442965.45, 1242442969.45). This data was taken from the channel L1:GDS-CALIB_STRAIN_CLEAN. Frequency range specified for LIGO-Virgo sensitivity band. Below 10 Hz is not properly calibrated. Above 10^3 Hz is affected by detector quantum noise.

I have used a computed set of BCR and SNR data files to compute the $\log_{10}$BCR to produce density plots against corresponding SNR distributions with varying α and β values, as shown in Figure 2. Contour and scatter plots show how the foreground, background, and selected GW trigger relate $\log_{10}$BCR vs SNR. The relation between these data subsets are also shown in a histogram with counts vs $\log_{10}$BCR, and cumulative distribution function vs $\log_{10}$BCR to represent the survival function from the selection of background triggers and foreground signals **[6]**. These visualizations help indicate whether the α and β values need to be adjusted and whether the BCR is effectively distinguishing GW signals from inherent noise.

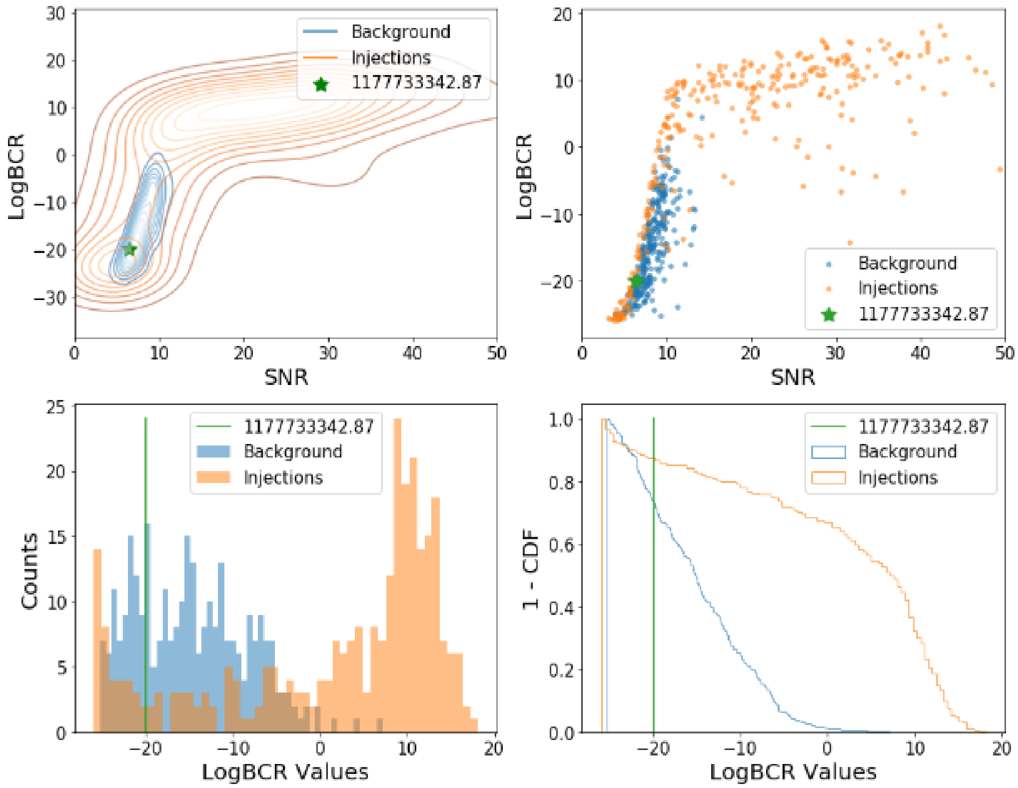

Figure 2: Density Plots with $\alpha = 3.62804 \times 10^{-11}$ and $\beta = 3.26975 \times 10^{-06}$

Recently, I have investigated how Alford's code is fitting and method of estimating the α and β values.

CHALLENGES:

Thus far, I have encountered some difficulty with fully understanding Bayesian inference and parameter estimation methods. I have worked to become familiar with the LIGO data and learning how to set up and run multiple Bilby jobs on a computer cluster. It has taken some time to understand Alford's code, his process, and methods for analyzing LIGO data and determining BCRs. Moving forward, I anticipate difficulties in determining useful α and β values to compute insightful likelihoods and BCRs. I may also find writing and submitting hundreds of Bilby jobs on the LIGO computer cluster since I am unfamiliar with Condor. In addition, I expect a learning curve with being able to analyze these results in a way so that I know which parameters may need to change and by approximately how much.

REFERENCES:

[1] Kanner et al.,
Leveraging waveform complexity for confident detection of gravitational waves, 2016, doi:10.1103/PhysRevD.93.022002

[2] Usman et al.,

The PyCBC search for gravitational waves from compact binary coalescence, 2016, arXiv:1508.02357 [gr-qc]
[3] Messick et al.,
Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data, 2017,
arXiv:1604.04324 [astro-ph.IM]
[4] Abbott et al.,
GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, 2018, arXiv:1811.12907 [astro-ph.HE]
[5] Fishbach M., Holz D. E., Where Are LIGO's Big Black Holes?, 2017, Astrophys. J. Lett., 851, L25, arXiv:1709.08584 [astro-ph.HE]
[6] Isi et al., Enhancing confidence in the detection of gravitational waves from compact binaries using signal coherence, 2018, arXiv:1803.09783 [gr-qc]
[7] Ashton et al.,
Bilby: A user-friendly Bayesian inference library for gravitational-wave astronomy, 2019, arXiv:1811.02042 [astro-ph.IM]
[8] Thrane E., Talbot C., An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models, 2019, arXiv:1809.02293 [astro-ph.IM]
[9] Nitz et al.,
1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data, 2019, arXiv:1811.01921 [gr-qc]
[10] Venumadhav et al.,
New Binary Black Hole Mergers in the Second Observing Run of Advanced LIGO and Advanced Virgo, 2019, arXiv:1904.07214 [astro-ph.HE]

[11] Venumadhav et al.,
A New Search Pipeline for Compact Binary Mergers: Results for Binary Black Holes in the First Observing Run of Advanced LIGO, 2019, arXiv:1902.10341 [astro-ph.IM]