Investigating data quality metrics for stochastic gravitational-wave detection

Caltech LIGO SURF 2021

Makenzi Fischbach
Wellesley College

Mentors: Derek Davis & Arianna Renzini
California Institute of Technology

July 2021

Investigating data quality metrics for stochastic gravitational-wave detection

Makenzi Fischbach[1]
Wellesley College[1]

Mentors: Derek Davis[2] and Arianna Renzini[2]
California Institute of Technology[2]
(Dated: July 2, 2021)

Abstract. The detection of gravitational waves has created the opportunity for many new discoveries. One such potential discovery is the stochastic gravitational wave background. In order to detect it, stochastic data must be properly monitored and analysed. Stochmon, a low latency stochastic data monitoring pipeline, works to monitor the quality of stochastic data. Stochmon has not been recently updated and is not well integrated with current gravitational wave data analysis tools. The goal of this project is to identify potential improvements to make to Stochmon's analysis functions, implement said changes, and integrate the system with existing analysis tools so that it can be used during the next observing run.

I. INTRODUCTION

Ever since their initial detection in 2015, gravitational waves (GWs) have been at the forefront of scientific research. GWs are notably ripples caused by disruptions to the fabric of space-time typically traced back to high-energy events, such as binary black hole mergers, compact binary coalescence (CBC), and bursts. GWs have the potential to provide unprecedented insight into astrophysical phenomena and the primordial universe [1].

The Laser Interferometer Gravitational-Wave Observatory (LIGO) has the ability to directly detect the GWs permeating from high-energy events and has been doing so since the first successful GW detection on September 14th 2015 [2]. LIGO is a large interferometer consisting of two, four kilometer arms oriented in an L-shape. A laser beam is split using a beam splitter and the two resulting beams are sent down the arms of the detector. If the light beams go undisturbed by GWs, the light from both arms will arrive back at the detector at the same time and cancel each other out, resulting in no GW detection. If a GW is present, it will create a slight disturbance and the two beams will return to the detector at different times. In this instance, the two beams of light will not cancel due to the varying phases, providing evidence of the presence of a GW. There are LIGO detectors in Livingston, Louisiana and Hanford, Washington.

While the sources of GWs are isolated astrophysical events, currently detected GWs can be detected from the stochastic gravitational-wave background (SGWB) [3]. The SGWB is a stochastic signal composed of the weak GW signals from a large number of unidentified events [5]. For instance, the superposition of GW signals from a population of binary black holes would appear stochastic. The SGWB can also be credited to stochastic processes that occurred in the primordial stages of the universe. We expect a successful detection of the SGWB to occur in the near future.

Stochmon is a data-quality monitor which specializes in the analysis of LIGO and Virgo low-latency stochastic data [4]. The monitor has a variety of tools that provide us with useful analysis, such as estimates for the sensitivity at which stochastic data is being collected and analyzed as well as coherence estimates for the two LIGO locations and the noise stationarity of the detectors.

A. Problem

The improvement of Stochmon will lead to a direct improvement in the analysis of stochastic data, the quality of data, and the overall detection of the SGWB. With the current instrumentation, the detection of the SGWB, and GWs in general, is imperfect. This is especially evident in frequency bands where the data is corrupted by noise.

An improvement in Stochmon would improve future stochastic data analysis. This would provide the opportunity for more research to be conducted on the SGWB and its corresponding sources. It would also aid researchers in detecting any problematic data more efficiently and in turn they would have more accurate and meaningful data.

This improvement in stochastic data analysis could potentially lead to a deeper understanding of the primordial universe and the stochastic events which may have occurred around the time of the Big Bang [6]. Additionally, the stochastic data analysis could provide the ability to achieve a deeper understanding of what the universe is composed of and allow for a method of detection free of scientific models.

As of right now, Stochmon exists and is operational. However, it has not been integrated and is not actively being improved or monitored. We want to ensure that Stochmon is updated and that the necessary improvements are implemented to ensure the most efficient detection of gravitational waves.

II. OBJECTIVES

One of the main objectives of this project is to improve Stochmon and its tools to more effectively monitor potential data-quality issues. Improved monitoring allows for the identification of imperfect data and may lead to the development of better data collection and analysis. Additionally, a goal of this project is to characterize the features in the LIGO data that have the potential to impact the stochastic sensitivity. A deeper understanding of the data leads to the identification of new ways to improve and analyze stochastic data.

We aim for an improvement of the Stochmon system and its ability to investigate the performance of LIGO's detectors in detecting the SGWB. An improvement must first be identified through an analysis of the current efficacy of the Stochmon system. Prior to beginning the project, the assumption is that all of the elements of Stochmon can be updated in some way to achieve a higher quality of stochastic data analysis. Another objective is to ensure that Stochmon and its tools are well integrated with other existing online data monitoring tools. These updates must then also be integrated so that they can be utilized during the next LIGO detection run. The project will be considered a success if we are on the path to improving Stochmon and the data it monitors.

III. APPROACH

The approach for this project is entirely contingent on what elements of Stochmon ultimately need improving and how those updates can be accomplished. The first step is to identify the components of Stochmon which are the most beneficial and need the most revision. After the initial identification, the next step would be to identify the ways in which the component could be improved and how those improvements could be implemented. The length of each step in the process is dependent on what approaches are taken and how intensive those approaches may be. One area that has already been identified for improvement is the systematic integration of the output of Stochmon with the LIGO detector Summary Pages. While working on improvements, we will be working closely with the original developers and maintainers of Stochmon. They will provide us with guidance and support throughout the process.

Stochmon consists of many tools which aid in the stochastic data analysis and which can be improved. Stochmon provides a detailed analysis of cross-correlated data between the Hanford and Livingston LIGO locations. This H1-L1 coherence, shown in Figure 1, is determined by dividing the cross power of the two detectors by the product of the auto powers [4]:

$$coh(f) = \frac{|\overline{S_{12}(f)}|^2}{\overline{S_1(f)}\ \overline{S_2(f)}} \tag{1}$$

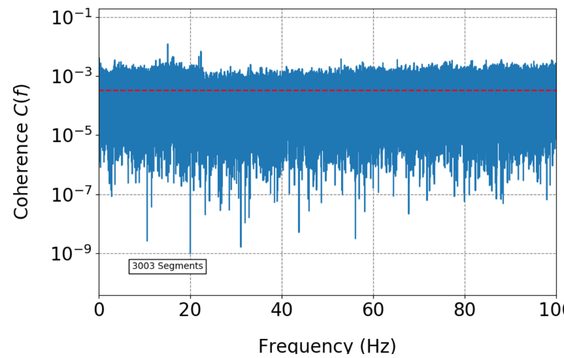

FIG. 1. The coherence between Livingston and Hanford with 1 mHz frequency resolution. The dashed red line signifies the expected level of coherence. This plot shows the coherence between the detectors is strongest from 0 Hz to about 22 Hz. Figure reproduced from the Stochmon summary page: `https://ldas-jobs.ligo.caltech.edu/~thomas.callister/stochmon-O3/stochmon.html`.

Knowing the coherence aids in the cross-analysis of data and therefore in the process of separating the stochastic data from any disruptive external artifacts or noise from instrumentation.

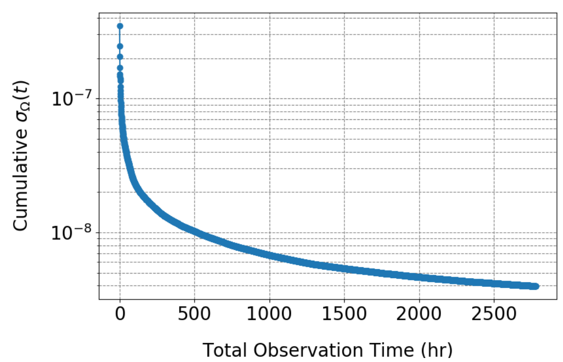

FIG. 2. Energy sensitivity vs. observation time. The search sensitivity decreases as observation time increases. The cumulative sensitivity is at its highest at the start of the observing period. The variance on Omega decreases as a function of time as $1/\sqrt{t}$, implying that the sensitivity to Omega increases through integrating over the whole observation time. Figure reproduced from the Stochmon summary page: `https://ldas-jobs.ligo.caltech.edu/~thomas.callister/stochmon-O3/stochmon.html`.

Stochmon also provides an analysis of the cross amplitude density plots for both detectors [4]:

$$a(f) = |\widetilde{s}_I \widetilde{s}_I(f)|^{1/2} \tag{2}$$

One of Stochmon's main features is the analysis of detector sensitivity to stochastic signals (Figure 2). The strain sensitivity (σ_h) is the sensitivity of what is measured with the detector. The energy sensitivity (σ_Ω), which is the cosmological quantity used in publications, is determined and is then compared to the aforementioned strain sensitivity [4]:

$$\sigma_\Omega(f) = \frac{10\pi^2}{3H_{100}^2}\frac{f^3}{\gamma(f)}\sigma_h(f)^2 \tag{3}$$

An analysis can then be performed by taking a weighted average of both the sensitivity of time and frequency [4]:

$$\sigma = \sum_{t=1}^{n}\sum_{f=1}^{m}(\sigma(f,t)^{-2})^{-1/2} \tag{4}$$

The analysis of sensitivity provides a deeper understanding of the detectors' strengths and weaknesses, as well as how they can be improved.

IV. STOCHASTIC DATA AND STOCHMON

An analysis of the detectors' sensitivities is one of Stochmon's central features which aids in the data analysis process. In order to best implement constructive change to the way in which the sensitivity is monitored, a deeper understanding of the sensitivity of the detectors as a whole had to be developed.

Each detector is most sensitive to different locations in the sky and has different polarisation responses dependent on their location and orientation on Earth. To best visualize these sensitivities, we can determine the detector polarisation response functions of each detector in both the cross and plus polarisation using built in Bilby functions and plot them using healpy (Figure 3) [7].

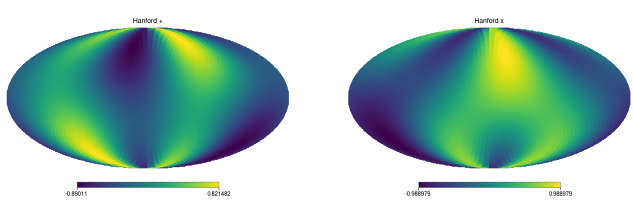

FIG. 3. The H1 cross (right) and plus (left) polarisation response functions. The dark blue and yellow represent locations in the sky in which the detector is most sensitive. The cross and plus polarisations allow for a wider range of high sensitivity.

After finding the detectors' individual polarisation response functions, an overlap function for the H1 and L1 detector pair as a function of time can be determined and plotted (Figure 4).

When the overlap function is plotted as a function of time, the areas of high sensitivity appear to rotate around the map once per sidereal day.

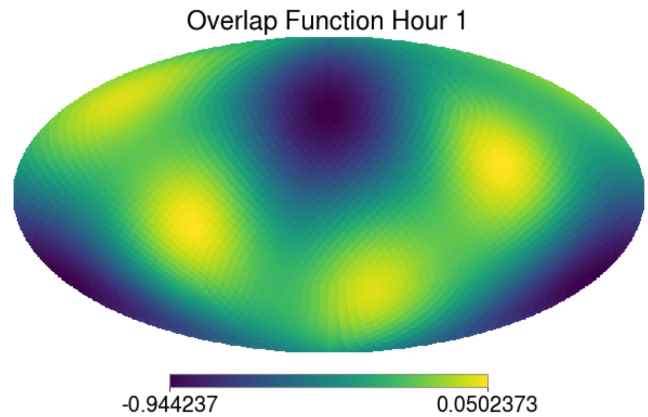

FIG. 4. The overlap function of H1 an L1 as a function of time at hour 1 of a sidereal day. The dark blue and yellow represent locations in the sky in which the detectors are most sensitive. As the sidereal day continues, the areas of high sensitivity rotate along with the detectors relative to a fixed point in the sky.

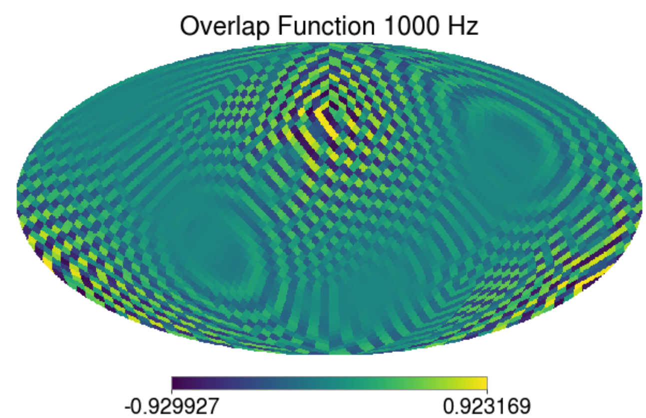

FIG. 5. The overlap function of H1 an L1 as a function of frequency at 1000 Hz. The dark blue and yellow represent locations in the sky in which the detectors are most sensitive. As the frequency increases, the number of waves that fit between the two detectors increases.

Next, the initial overlap function is multiplied with the plane wave term, where Δt is the delay term between both detectors, to get a visualisation (Figure 5) for the full stochastic sky response dependent on frequency:

$$\gamma(f,n;t) = overlap(n;t) * e^{i2\pi f\Delta t} \tag{5}$$

The overlap functions can also be visualized in three dimensions, where the radius of the plot correlates to the value of the overlap function at a given point in the sky. Figure 6 shows the 3D overlap function as a function of time and Figure 7 is the 3D representation as a function of frequency.

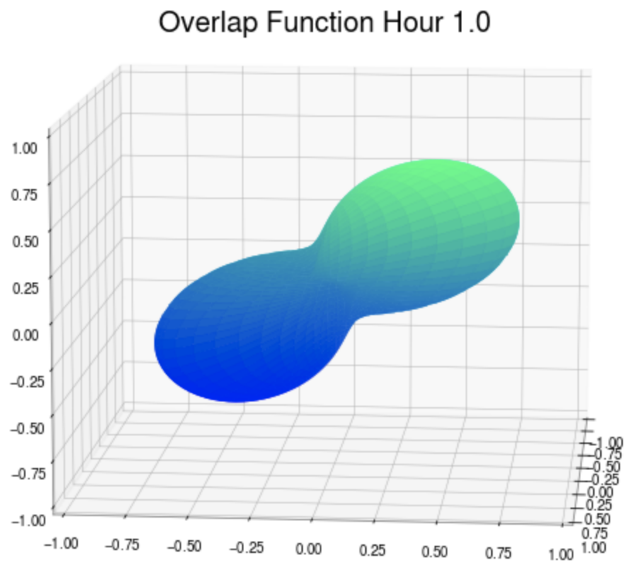

FIG. 6. 3D overlap function dependent on time at hour 1 of a sidereal day. In this plot, the coloring aids in the 3D visualization and has no further significance. The axes represent the value of the overlap function at a given location.

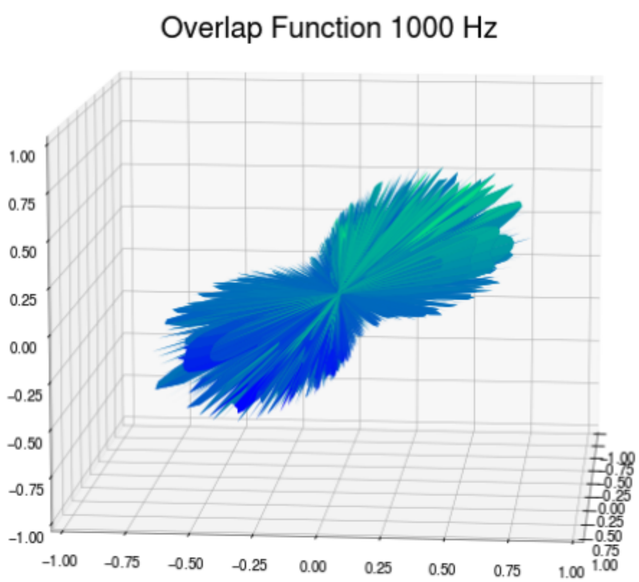

FIG. 7. 3D overlap function dependent on frequency at 1000 Hz. In this plot, the coloring aids in the 3D visualization and has no further significance. The axes represent the value of the overlap function at a given location.

Overlap functions are a helpful tool for both potential stochastic gravitational wave detection and stochastic data analysis.

V. DETCHAR AND DATA QUALITY ANALYSIS

To implement stronger tools for stochastic data quality (DQ) analysis, it is beneficial to turn to general DQ analysis for guidance. At LIGO, one of the ways DQ is evaluated is through DQ shifts. DQ shifts are standard procedure where the designated 'shifters' for the week at both H1 and L1 are tasked with reviewing a week's worth of data and plots and writing a summary of any significant or notable changes. These can be either positive or negative changes.

While doing DQ shifts, shifters refer to the detector characterization (Detchar) summary pages, which contain all of the data and plots for each day of an observing run. On these pages, there are four main plots that shifters analyse. These plots are meant to detect and highlight any noise, glitches, or events in the data that may affect analysis done with said data.

The first plot is the spectrogram, which is a plot of time versus frequency for a day of observing. Glitches and noise in the data present themselves as red or blue lines. Red lines represent high noise relative to the median while blue lines represent a decrease from the standard noise. The binary neutron star (BNS) inspiral range plot shows the detector's range at which it is most sensitive in detecting BNS related events. Shifters look for significant increases or decreases in the range which may be indicative of a change in the detector's functionality. 'Glitchgrams' and glitch rate plots both show potential glitches in the data. Shifters look for loud glitches that may be in clusters. These clusters of noise could have negative impacts on the data and impact its usability. The detchar summary pages also have Hveto and LASSO, tools that identify the potential channels which may be responsible for the noise and glitches in the data[9][10].

VI. NEXT STEPS

The next steps for this project involve more learning, as well as the beginning stages of implementing improvements to Stochmon. My next task is to compare and contrast CBC data quality metrics with stochastic data quality metrics to better identify what changes to make to the current stochastic monitoring pipeline, and how to effectively implement and integrate those changes.

[1] T. L. S. Collaboration, the Virgo Collaboration, the KAGRA Collaboration, *et al.*, Upper limits on the isotropic gravitational-wave background from advanced ligo's and advanced virgo's third observing run (2021),

arXiv:2101.12130 [gr-qc].

[2] J. Aasi and et al., Advanced ligo, Classical and Quantum Gravity **32**, 074001 (2015).

[3] M. Maggiore, Gravitational wave experiments and early universe cosmology, Physics Reports **331**, 283–367 (2000).

[4] G.Hernandez, Stochmon: A ligo data analysis tool (2013).

[5] J. D. Romano and N. J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: a unified treatment, Living Reviews in Relativity **20**, 10.1007/s41114-017-0004-1 (2017).

[6] T. Regimbau, The astrophysical gravitational wave stochastic background, Research in Astronomy and Astrophysics **11**, 369–390 (2011).

[7] G. Ashton and et al., Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy, The Astrophysical Journal Supplement Series **241**, 27 (2019).

[8] D. Davis and et al., Ligo detector characterization in the second and third observing runs (2021), arXiv:2101.11673 [astro-ph.IM].

[9] J. R. Smith and et al., A hierarchical method for vetoing noise transients in gravitational-wave detectors, Classical and Quantum Gravity **28**, 235005 (2011).

[10] M. Walker and et al., Identifying correlations between ligo's astronomical range and auxiliary sensors using lasso regression, Classical and Quantum Gravity **35**, 225002 (2018).