Optimal Settings for Fast Low-Latency Skymaps of Neutron Star Binaries
Caltech LIGO SURF 2021, Final Report

Celine Wang and Mentors: Katerina Chatziioannou and Isaac Legred
(Dated: October 29, 2021)

The detection of gravitational waves is instrumental to our understanding of astrophysical processes and the fate and evolution of the sources of such waves. One source of gravitational waves is compact binary coalescences (CBC's)–binary systems which consist of black holes, neutron stars, or both. Here, we examine neutron star binary systems and their intrinsic and extrinsic parameters and determine how to optimize data intake to improve the localization of such systems. We confirm that longer signal durations maximize relevant information we can extract from the signals but increase typical computing efforts. For analyses where prompt results are essential, such as for multi-messenger followup, we present the ideal conditions for fast and accurate analysis.

I. INTRODUCTION

Since commencing observing operations in 2015, the Advanced LIGO and Advanced Virgo detectors have discovered a plethora of gravitational waves. Their first gravitational wave transient catalog includes ten binary black hole coalescences and one binary neutron star coalescence. Additionally, an updated second gravitational wave catalog from the third observing period has yielded dozens more events, with signals encompassing even more compact binary coalescences [1].

The binary neutron star merger was detected through GW170817 [2]; the first signal from this low-mass, compact binary inspiral was detected August 17, 2017, and was inferred by astronomers to be located in the NGC 4993 galaxy. Gravitational waves from binary neutron stars exhibit a chirplike, or frequency-increasing, time evolution which to leading order depends both on the system's chirp mass with additional contribution from its mass ratio and spins. Also unique to neutron star systems is the influence of their internal structure on its waveform; such properties can be inferred from tidal interactions.

The localization of GW170817 to the NGC 4993 galaxy was aided by the skymap that was computed from gravitational wave data. Once the skymap was constructed, astronomers were able to follow up on this data and were able to localize the source within a few hours after initial detection. This initial skymap took around 4-5 hours to be sent out, and an updated version using methods relying on Bayesian inference took about 10 hours. Figure 1 demonstrates the improved localization of GW170817 from gravitational wave data alone due to re-calibration of Virgo data. Assuming the previously deduced location in NGC 4993, the 90% localization region was reduced from 28 $\deg^2$ to 16 $\deg^2$.

Aiding in this endeavor of gravitational wave detection is BILBY [3], a Bayesian inference library which infers source properties from individual signals of compact binary coalescences. For any analysis, BILBY is able to vary certain source parameters and produces samples from the posterior for the source parameters. Some parameters may remain fixed while others are sampled; for example, sky parameters are frequently fixed for GW170817 analyses since its exact location is known. To date, BILBY has produced reliable results for both simulated and real gravitational wave data from compact binary mergers and coalescences.The recovered simulated parameters are consistent with the injected parameters, and the inferred parameters from real signals are consistent with other pipelines. Using these source properties, in upcoming observing runs BILBY will be used to compute skymaps which will then be utilized by astronomers to localize the gravitational wave signal sources, such as in the case of GW170817.

II. MOTIVATIONS

Gravitational wave data allows for localization of the source signal itself, upon which astronomers may then search to identify the exact location of the CBC event. Ideally, this continuous process of intaking gravitational wave data and computing the subsequent skymap should occur in as minimal a timeframe as possible, since an electromagnetic signal from the CBC event could fade rapidly within the span of a few hours or even minutes. The runtime will vary depending on which parameters are sampled but ideally should last from a few minutes up to half an hour.

For every signal analysis, there exists a trade-off between accuracy and computational efficiency. A typical BNS signal may last for a span of up to two minutes with a frequency of interest for the analysis ranging from 20 Hz to 2000 Hz; Figure 2 shows such an example of a BNS signal. Conversely, the amount

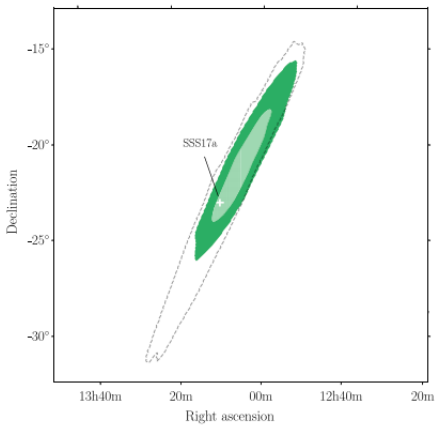

Figure 1. The improved localization of GW170817, with the dotted gray line indicating previous localization of 90% credibility from 2017 and lighter and darker green regions corresponding to increased region credibility of 50% and 90%, respectively, in 2019. The electromagnetic counterpart of GW170817 was found based on the 2017 skymap.

of time it takes to compute the associated skymaps may range from a few hours to even months, depending on various factors including models and tools for analyses. This proves problematic when we consider that the electromagnetic signal from the CBC event may last only up to a few hours at most and a skymap must be computed and distributed for astronomical follow-up as soon as possible. As such, we explore how settings such as segment length, sampling rate, and the sampler dictionary may be optimized in order to achieve reliable sky localization in a minimal timeframe. The segment length is the duration of the signal we are analyzing; the sampling rate is the number of samples taken per second, and the sampler dictionary allows us to vary the sampler we use and its functions.

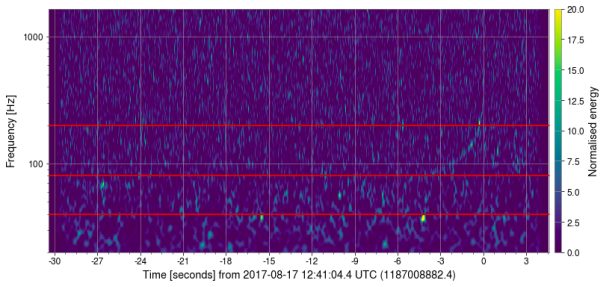

Figure 2. Spectrogram of GW170817 from the LIGO Hanford detector with the merger placed at the 0 second time mark. Frequencies of 40, 80, and 200 Hz, respectively, are overplotted in red; these correspond to various times that different parameters may be inferred.

Ultimately, many inferred source parameters through gravitational wave data of merging binary neutron stars will be further improved by and electromagnetic detection and identification of the host galaxy. These include specific properties of the binary system itself, such as its mass, spin, and tidal parameters, which may also better equip our general understanding of binary. Improved localization may also better our understanding of short gamma-ray burst properties and, on a grander scale, the equation of state of neutron-star matter, the nature of gravity, the value of the cosmological constant, and even allow us to test the theory of general relativity.

III. METHODS

A. Bayesian Inference

We rely on Bayesian inference for this project. We begin with a posterior probability distribution which is calculated using Bayes' Theorem:

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}} \tag{1}$$

Here θ represents the source parameters, $\mathcal{L}(d|\theta)$ is the likelihood function, or probability of the detectors measuring data d assuming a model hypothesis, $\pi(\theta)$ is the prior distribution, which incorporates any prior knowledge about our parameters, and $\mathcal{Z}$ is the normalization factor, or evidence, which is defined as:

$$\mathcal{Z} = \int \mathcal{L}(d|\theta)\pi(\theta)d\theta \tag{2}$$

This evidence indicates how well the data is modeled by the hypothesis, which is vital for model selection. In this case, the posterior probability distributions are calculated by BILBY, and we apply restricted analysis by focusing on restricted parameters and/or timeframes in order to narrow down the posterior distribution area. Ultimately the parameters that we are primarily interested in are the sky location and the distance to the source, but we also consider some intrinsic parameters such as masses, spins, and tidal deformability.

B. Stochastic Sampling

To generate its posterior distributions, BILBY relies on two types of stochastic sampling, where independent samples are drawn randomly from the parameter space; these two methods are the Monte

Carlo Markov Chain (MCMC) and nested sampling. For MCMC, particles within the given parameter space will take a "random walk" along the posterior distribution. The probability of a particle going to a certain position is determined by the Markov chain transition probability. On the other hand, nested sampling initially calculates the evidence instead of drawing samples from the posterior, which is generated as a byproduct. For nested sampling, the parameter space contains a certain number of live points drawn from the prior distribution. For every iteration, the lowest likelihood live point is removed; this process continues until the entire prior distribution has been sampled [4].

The default sampler that BILBY uses is Dynesty [5], which utilizes both MCMC and nested sampling. First, Dynesty extracts N number of likelihood live points from the prior, which is controlled by the setting `nlive`. The uniform sampling is iteratively improved by selecting new live points through random walkers using the MCMC algorithm [3]. The setting `maxmcmc` determines the maximum number of random walkers used. Both `nlive` and `maxmcmc` impact runtime and efficiency of CBC analyses, and we experiment with these settings for this project.

IV. PROCEDURE AND RESULTS

For all our injections, we use BILBY_PIPE, a Python package which automates the process of running BILBY on a computing cluster [3] due to the large volume of data. We have analyzed several of these run results using PESummary, a parameter estimation summary page builder which compiles relevant data into a user-friendly and comprehensible webpage [6].

A. Priors

To begin our analysis, certain settings for the prior needed to be implemented. Instead of sampling over the individual mass components, we use the mass ratio:

$$q = \frac{m_2}{m_1} \leqslant 1 \tag{3}$$

and chirp mass:

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{4}$$

because it is easier and faster to sample over these parameters.

Next, we substituted the six separate spin parameters a_i, `tilt_i`, and `phi_jl` and `phi_12`, with the single spin parameter X_i, one for each neutron star; the former constitute the dimensionless spin magnitude and associated angle parameters, while the latter is the ith object aligned spin, or the projection of the ith object spin onto the orbital angular momentum:

$$\mathcal{X}_i = a_i cos(\theta_i) \tag{5}$$

Since we use nonprecessing injections for this project where all spin magnitudes and angles are set to zero, using `chi_i` is computationally more efficient because it gives us less parameters to sample over.

B. Optimal Settings for Runtime

Two of the primary settings that we experiment with are the segment length, or time duration of the signal we are analyzing, and sampling rate, or number of samples taken per second which corresponds to half the highest frequency of the signal that we analyze. We use decreasing segment lengths from 128 seconds down to 16 seconds and sampling rates from 2048 Hz down to 128 Hz. Both the sampling rate and segment length have a direct relationship with the runtime; the larger the segment length and sampling rate, the longer the runtime, as Figure 3 demonstrates. The sampling rates of 2048 Hz and 1024 Hz took such a large amount of time to run that for practical purposes, those rates should not be used for successful multi-messenger followup efforts. However, sampling rates of 512 Hz and below proved more realistic with a maximum runtime of 69.5 hours and a minimum runtime of 40 minutes.

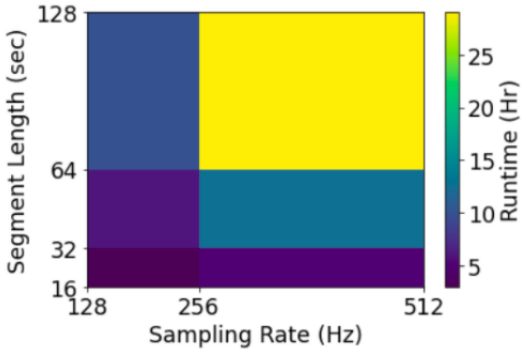

Figure 3. Heat map showing the effects of sampling rate and segment length settings on runtime. Smaller sampling rates and segment lengths result in a shorter runtime.

C. Varying the Segment Length and Sampling Rate

After comparing the runtimes, we then proceed to compare how the different settings impact the posterior distributions that are generated. For these injections, we use three detectors–LIGO Livingston, LIGO Hanford, and VIRGO–and inject noise as well. We fix either the segment length or the sampling rate and then vary the other setting to see how removing available data reduces the accuracy of the marginalized posteriors that are returned. Two parameters that are particularly affected by data reductions include the chirp mass and effective spin. Chirp mass and effective spin are primarily inferred along lower frequencies in the signal, and so decreasing the segment length, which removes lower frequency data, would cause the posterior distribution to be wider, as shown in Figures 4 and 5. This effect in chirp mass is even more prominent when we use a lower sampling rate, as shown in Figure 6; nonetheless, all of the posteriors still agree with the true value that is denoted by the black vertical line.

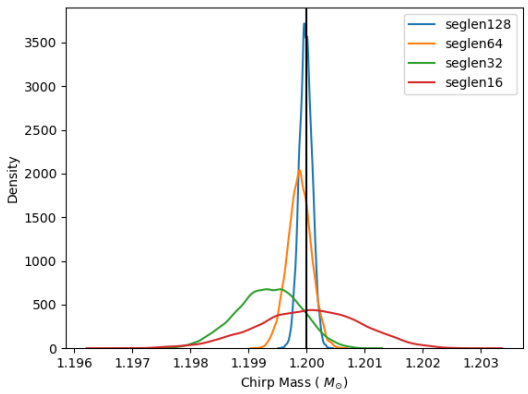

Figure 4. Posterior distribution for chirp mass at a sampling rate of 256 Hz with decreasing segment lengths.

Conversely, tidal parameters in binary neutron star systems are mostly inferred among higher frequencies. Decreasing the sampling rate removes higher frequency data, so fixing the segment length while varying the sampling rate should exhibit the same pattern in the tidal parameter distributions as in the previous examples with chirp mass and effective spin. Our recent runs so far have produced the expected tidal parameter posterior distributions only for higher sampling rates of 2048 Hz and 1024 Hz as shown in Figure 7. However, at lower sampling rates there is not as much information to be extracted about the tides compared to the higher sampling rates, and so the posteriors look fairly similar.

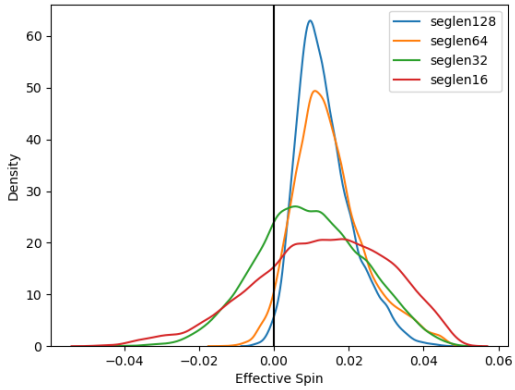

Figure 5. Posterior distribution for effective spin at a sampling rate of 512 Hz with decreasing segment lengths. The peak offset from the true value is due to the well known correlation between effective spin and mass ratio and is seen in all signals, both real or injected.

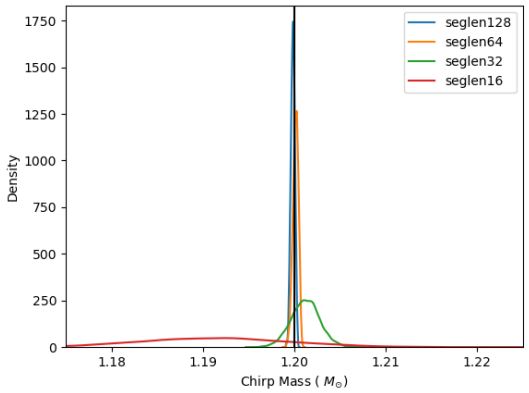

Figure 6. Posterior distribution for chirp mass at a sampling rate of 128 Hz with decreasing segment lengths.

Although not as drastic as the chirp mass and effective spin examples, the luminosity distance also exhibits a similar, albeit more subtle, pattern as the posterior distributions for the aforementioned parameters when one decreases the sampling rate for a constant segment length, as shown in Figure 8. This shows that a decreased sampling rate of 256-512 Hz does not affect the distance posterior generation as much, since all relevant information for the distance has already been obtained from lower frequencies. However, continuing to decrease the sampling rate will cause the loss of valuable information from low frequencies.

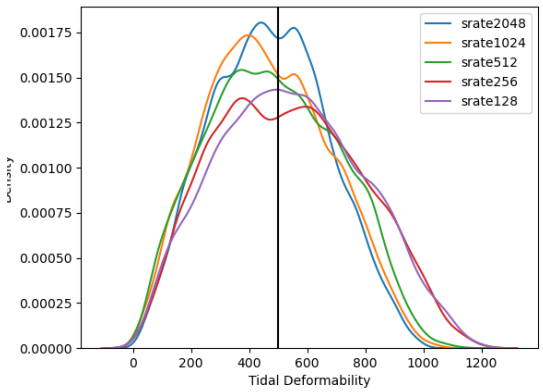

Figure 7. Posterior distribution for tidal deformability at a segment length of 16 seconds with decreasing sampling rate.

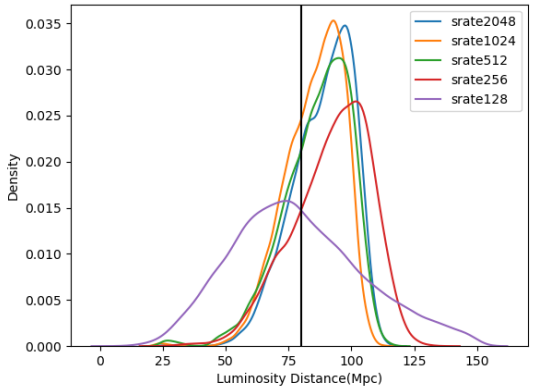

Figure 8. Posterior distribution for luminosity distance at a segment length of 16 seconds with decreasing sampling rate.

D. Skymaps

Removal of certain frequencies of data by decreasing the sampling rate or segment length also affects the accuracy of skymaps generated from our BNS injections. In general, decreases in both the sampling rate and the segment length exhibit a similar pattern for the reduced accuracies. The following skymaps are all generated from an injected sky location of a right ascension of 4.2 and a declination of -0.7. Figure 9 shows how a decreasing sampling rate for a constant segment length of 32 seconds produces a worse constrained sky location posterior distribution; however, the highest shown sampling rate of 512 Hz generates an extremely well constrained skymap, which Figure 10 zooms in on. A similar pattern in the

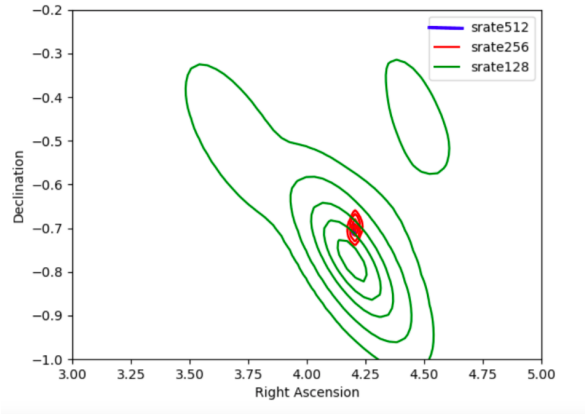

Figure 9. Skymaps generated at a segment length of 32 seconds with decreasing sampling rate.

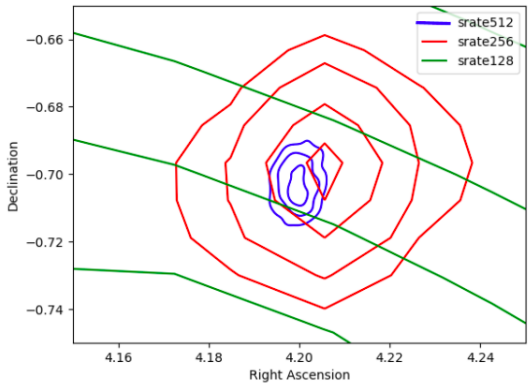

Figure 10. Zoomed in skymaps generated at a segment length of 32 seconds with decreasing sampling rate.

accuracy and constraint of skymaps occurs for other fixed segment lengths with varying sampling rates, such as in Figure 11, as well as for fixed sampling rates with decreasing segment lengths, as in Figure 13.

From Figure 12, we see that as the segment length decreases and less data is retrieved, we obtain a worse localization in return. A sampling rate of 512 Hz produces a very well constrained sky localization with a 90% credible region of 2 deg^2 and a 50% credible region of 1 deg^2 according to the PESummary generated skymap; this run was completed in around 12 hours. Although not as well constrained as the 512 Hz run, a sampling rate of 256 Hz still generates a fairly well constrained skymap with a 90% credible region of 15 deg^2 and a 50% credible region of 3 deg^2. The runtime for these settings was around 2/3 of an hour, which is pragmatic and within a reasonable margin of time in which a counterpart electromagnetic signal from a binary neutron star system

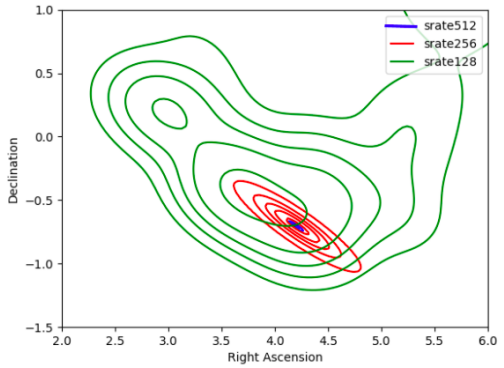

Figure 11. Skymaps generated at a segment length of 16 seconds with decreasing sampling rate.

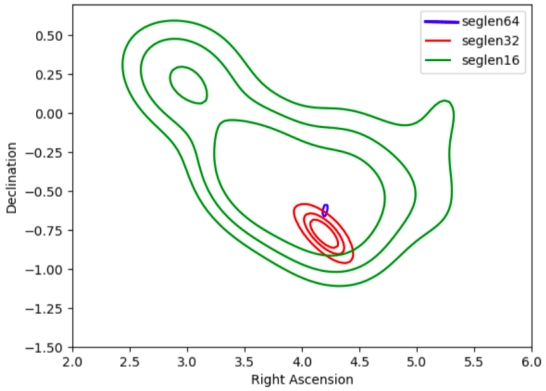

Figure 12. Skymaps generated at a sampling rate of 128 Hz with decreasing segment lengths.

might be visible and detectable for multimessenger followup.

V. CONCLUSIONS AND FUTURE WORK

From these results we have identified the ideal settings for generating skymaps of neutron star binaries via Bayesian inference in a minimal time frame. We may confirm that both a smaller sampling rate and segment length lead to a shorter runtime. We also conclude that the removal of certain data frequencies from our analysis results in less accurate posterior distributions for certain parameters such as the chirp mass and effective spin. Assuming O3 design sensitivity and setup, we have successfully localized an injected binary neutron star source up to 80 Mpc–twice the distance of the signal GW170817–away very well with a reasonable runtime.

One area to experiment upon in the future would be the generation of skymaps for different locations in the sky. We may also experiment with the number of detectors used in runs–including versus excluding the VIRGO detector–and see how those results compare. Another area of exploration lies in the sampler dictionary. So far we have merely reduced `nlive` to 500 from a default of 1000 and `maxmcmc` to 3000 from a default of 5000 in an attempt to reduce the runtime; however, in the future we may reduce these even more and compare how various values of these sampler settings impact the runtime.

VI. ACKNOWLEDGMENTS

I would like to thank my mentors Katerina Chatziioannou and Isaac Legred for their guidance and support along this project; Alan Weinstein for his wisdom and management of the LIGO SURF program; Colm Talbot for his additional advising on BILBY; the National Science Foundation, California Institute of Technology, and LIGO Scientific Collaboration for providing me the opportunity to conduct this research; and Noah Wolfe, for no particular reason at all.

[1] B. 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, (2019), arXiv:1811.12907.
[2] B. Abbott *et al.*, Properties of the binary neutron star merger GW170817, (2019), arXiv:1805.11579.
[3] I. Romero-Shaw *et al.*, Bayesian inference for compact binary coalescences with BILBY: Validation and application to the first LIGO–Virgo gravitational-wave transient catalogue, (2020), arXiv:2006.00714.
[4] E. Thrane and C. Talbot, An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models, Publications of the Astronomical Society of Australia **36**, 10.1017/pasa.2019.2 (2019).
[5] J. S. Speagle, Dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences, Monthly Notices of the Royal Astronomical Society **493** (2020).
[6] C. Hoy and V. Raymond, Pesummary: the code agnostic parameter estimation summary page builder (2021), arXiv:2006.06639 [astro-ph.IM].