Glitch mitigation methods for parameter estimation of compact binary coalescences

Viviana A. Cáceres Barbosa
Physics Department, University of Puerto Rico at Mayagüez

Mentor: Derek Davis
LIGO, California Institute of Technology
(Dated: May 9, 2022)

Project Proposal
LIGO Caltech SURF Program 2022

Recovering accurate distributions for the source parameters of gravitational wave signals is essential to confirm current models of general relativity. Bayesian inference used in most parameter estimation pipelines assumes that the noise distribution of the data is stationary and Gaussian. However, if there are glitches in the data, making this assumption can reduce the validity of the posterior distributions obtained. We propose an implementation of inpainting to address this issue in BILBY, which is one of various parameter estimation pipelines routinely used for LIGO analysis. Inpainting a segment of glitch data involves solving a Toeplitz system, which can be done computationally effectively. We aim to maintain BILBY's efficiency and improve its accuracy in preparation for LIGO's future observing runs.

I. INTRODUCTION

Raw strain data recorded by gravitational wave (GW) detectors such as the Laser Interferometer Gravitational Wave Observatory (LIGO) is typically dominated by noise coming from a variety of different sources [1]. Although many of these sources are known and well-modeled, there are often short noise bursts, known as glitches, of unknown origin that impact the sensitivity of the data analysis softwares used by the LIGO Scientific Collaboration (LSC) [2]. Much of LSC efforts are dedicated to mitigating the effect of glitches in GW signal searches and parameter estimation methods.

BILBY is one of various parameter estimation pipelines used by the LSC [3]. Like most parameter estimation pipelines, it uses Bayesian inference to produce posteriors, which are probability distributions of the GW source parameters. These are computed using Bayes' theorem:

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

where $\mathcal{L}(d|\theta)$ is the likelihood of measuring the data d given some source parameters θ, $\pi(\theta)$ is the prior distribution of these source parameters, and $z(d)$ is the evidence [4]. An example of how these posterior distributions may look is shown in Fig. 1

The BILBY analysis assumes that the noise in the data is stationary and Gaussian. This allows the use of the Gaussian noise likelihood:

$$\mathcal{L}(d|\theta) = \frac{1}{|2\pi\mathbf{C}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\chi^2(\mathbf{d},\mathbf{h})\right\} \tag{2}$$

with

$$\chi^2(\mathbf{d},\mathbf{h}) = [\mathbf{d}-\mathbf{h}(\theta)]\mathbf{C}^{-1}[\mathbf{d}-\mathbf{h}(\theta)] \tag{3}$$

where $\mathbf{d}$ is a vector representation of the data, $\mathbf{C}$ is the noise covariance matrix and $\mathbf{h}(\theta)$ is the waveform with parameters θ [1]. When glitches are present in the data, they invalidate this fundamental assumption. Therefore, segments of strongly non-stationary, non-Gaussian data must be dealt with before this likelihood can be used.

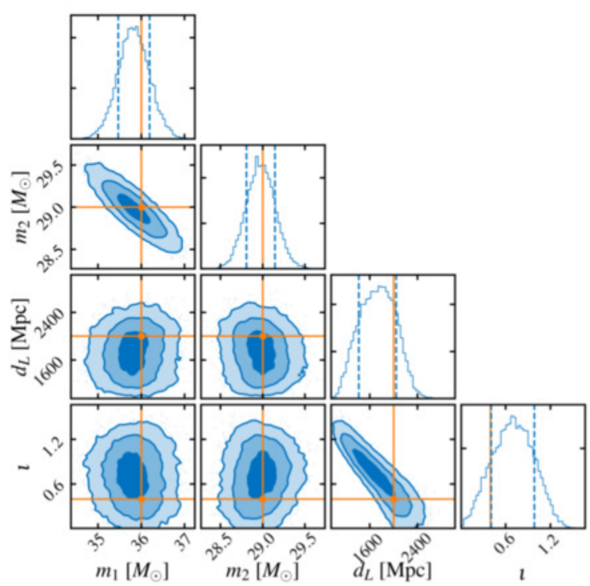

FIG. 1. Posterior distributions obtained using BILBY for the luminosity distance d_L, source masses m_1, m_2, and inclination angle i of an injected binary black hole signal. The injected values are shown in orange. Image reproduced from [3].

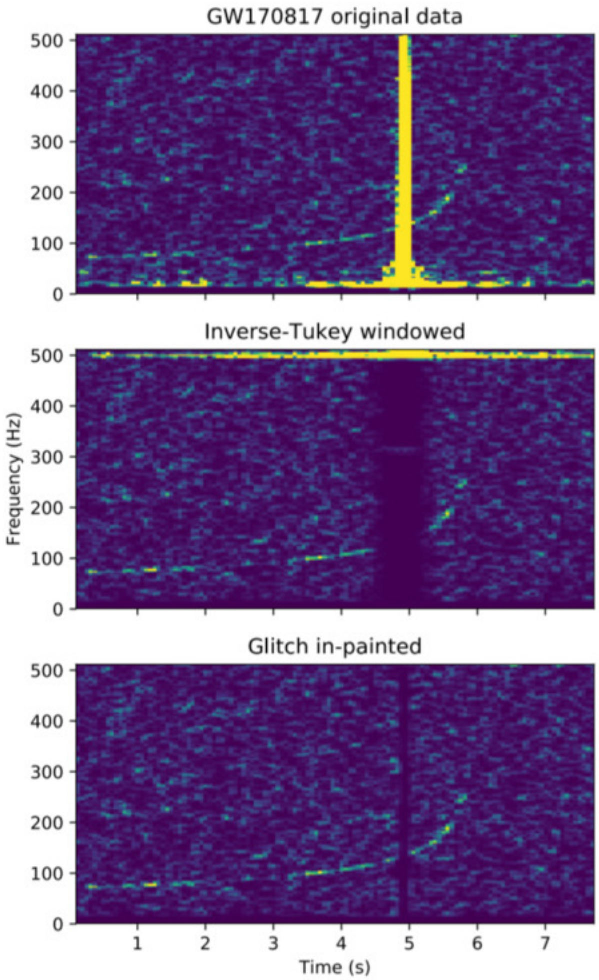

FIG. 2. Spectrograms showing the effect of gating and inpainting for a glitch near the merger time of GW170817. *Top panel*: Spectrogram of original data. *Middle panel*: Spectrogram of gated data. *Bottom panel*: Spectrogram of inpainted data. Image reproduced from [5].

A. Inpainting

Two methods often used to address glitches are gating and inpainting. Gating directly zeroes out and tapers segments with glitches, but can create artifacts if done before whitening, or corrupt the data around the glitch if done after [2]. Applying an inpainting filter avoids these problems by modifying the region of interest such that when the data is whitened, this region will be zeroed out. Fig. 2 shows a comparison between gating and inpainting for a glitch that overlapped with the signal for GW170817.

Some GW data analysis pipelines such as PyCBC [6] have incorporated inpainting into their workflow. In this proposal, we outline a plan to implement inpainting in BILBY, which will aid in the recovery of accurate parameters for past and future GW observing runs. Section II lists the objectives of this project, Section III explains the methods that will be used to carry out these objectives, and Section IV organizes the timeline of the project.

II. OBJECTIVES

The goal of this project is to develop a method to prevent glitches from biasing parameter estimation with BILBY. We propose to do this using an inpainting filter, with the intention that segments of data with glitches do not contribute to the outcome of our Bayesian analysis. Assuming that we have already identified segments that contain glitches and that the noise excluding the glitch can be modeled as stationary and Gaussian, then applying an inpainting filter would improve the precision and accuracy of the calculated posterior distributions. We do not expect the filter to significantly reduce the run time for an analysis.

We can analyze both the validity of our results and the effect on BILBY's speed by running tests with and without the inpainting filter and comparing the resulting posterior distributions and run times. This can be done by injecting a GW signal with known parameters into data that has segments with glitches and performing different BILBY analyses on it. If the inpainting filter improves our results, we can also run tests on actual GW signals from past observing runs.

III. APPROACH

A. Inpainting the data

Following the derivation in [5], if the data contains N_h samples to be masked, we denote $u^{(\alpha)}$ a list of N_h vectors, each of which are 0 everywhere except at one of the samples to be masked, where they are 1. Functionally, these vectors pick out the samples of the data that contain the glitch. Because we want the whitened data to be 0 for these samples, the inpainting filter F applied to data d satisfies

$$u^{(\alpha)T}\mathbf{C}^{-1}F\mathbf{d} = 0 \tag{4}$$

where $\mathbf{C}$ is the noise covariance matrix. The inpainted data Fd equals the original data outside of the glitch region but can take any value inside since these samples will be nulled after whitening.

To obtain the inpainted data, Eq. 4 must be solved for Fd. The inpainted data can be written as

$$F\mathbf{d} = \mathbf{d} - \mathbf{d}_{proj} \tag{5}$$

where $\mathbf{d}_{proj}$ is the projection of the zeroed data into the whitened data space. Therefore, Eq. 4 is equivalent to

$$u^{(\alpha)T}\mathbf{C}^{-1}\mathbf{d}_{proj} = u^{(\alpha)T}\mathbf{C}^{-1}\mathbf{d} \tag{6}$$

Since C^{-1} is a diagonal-constant matrix, this is a matter of solving a Toeplitz system, which can be done computationally efficiently. Once the inpainted data is obtained, we can proceed to use it for our Bayesian analysis.

B. Likelihood approximation

Because the Gaussian likelihood in Eq. 2 is applied to a finite-duration discretely sampled time series (the GW strain data), it can be approximated using the Whittle likelihood, as done in [7]:

$$\mathcal{L}(\mathbf{d}|\theta) \propto \exp\left[-\frac{1}{2}(\mathbf{d}-\mathbf{h}|\mathbf{d}-\mathbf{h})\right] \tag{7}$$

where $(\mathbf{d}-\mathbf{h}|\mathbf{d}-\mathbf{h})$ is the noise-weighted inner product defined as

$$(\mathbf{a}|\mathbf{b}) \equiv \sum_{j=0}^{\frac{N}{2}-1} 4\,\mathrm{Re}\left(\frac{\tilde{a}_j^* b_j}{S_n(f_j)}\Delta f\right) \tag{8}$$

for a segment with N samples and power spectral density S_n. If we use inpainted data instead of the original data in the Whittle approximation, we obtain

$$\mathcal{L}(\mathbf{d}_{inp}|\theta) \propto \exp\left[-\frac{1}{2}(\mathbf{d}_{inp}-\mathbf{h}|\mathbf{d}_{inp}-\mathbf{h})\right] \tag{9}$$

$$= \exp\left[-\frac{1}{2}(\mathbf{d}_{inp}|\mathbf{d}_{inp}) + (\mathbf{d}_{inp}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h})\right] \tag{10}$$

In this expression, $(\mathbf{d}_{inp}|\mathbf{d}_{inp})$ does not depend on the waveform, and, because we have defined the inpainting filter such that it can take any value inside the hole, $(\mathbf{d}_{inp}|\mathbf{h})$ is independent of the waveform inside the hole. Since $(\mathbf{h}|\mathbf{h})$ does not depend on the data, inpainting only the data will not be enough to ensure that the likelihood is not biased. The generated waveforms themselves must be inpainted as well.

Once this method has been implemented, its performance can best be studied by inserting a gravitational wave signal whose parameters are known, called an *injection*, to data with known glitches, and then comparing the posterior distributions obtained with and without the inpainting filter.

IV. TIMELINE

This project will take place over ten weeks. Tentatively, the first half will be dedicated to training in the following areas:

- Using BILBY: completing tutorials for obtaining posterior distributions, studying the syntax of BILBY code
- Inpainting data: learning how to apply the inpainting filter described in Sec. III to test data

After the appropriate training is completed, the second half of the project will be dedicated to implementation, which will require:

- Writing the code in BILBY that will inpaint the data
- Creating an injection to insert into strain data with glitches
- Running the new code with the injected data and studying the posterior distributions
- Running the code with real gravitational wave signals and comparing with the posteriors obtained in previous LIGO analyses

Finally, the last weeks will be spent organizing the results for the final presentation and written report.

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals, Class. Quant. Grav. **37**, 055002 (2020), arXiv:1908.11170 [gr-qc].

[2] D. Davis and M. Walker, Detector Characterization and Mitigation of Noise in Ground-Based Gravitational-Wave Interferometers, Galaxies **10**, 12 (2022).

[3] G. Ashton *et al.*, BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy, Astrophys. J. Suppl. **241**, 27 (2019), arXiv:1811.02042 [astro-ph.IM].

[4] I. M. Romero-Shaw *et al.*, Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue, Mon. Not. Roy. Astron. Soc. **499**, 3295 (2020), arXiv:2006.00714 [astro-ph.IM].

[5] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, Detecting gravitational waves in data with non-stationary and non-Gaussian noise, Phys. Rev. D **104**, 063034 (2021), arXiv:1908.05644 [astro-ph.IM].

[6] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, A. B. Nielsen, and B. Krishnan, Observation of a multimode quasi-normal spectrum from a perturbed black hole, (2021), arXiv:2105.05238 [gr-qc].

[7] J. Y. L. Kwok, R. K. L. Lo, A. J. Weinstein, and T. G. F. Li, Investigation of the effects of non-Gaussian noise transients and their mitigation in parameterized gravitational-wave tests of general relativity, Phys. Rev. D **105**, 024066 (2022), arXiv:2109.07642 [gr-qc].