Implementation of an inpainting filter to mitigate the effect of glitches on gravitational wave parameter estimation

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

Mentor: Derek Davis
LIGO, California Institute of Technology
(Dated: July 6, 2022)

First Interim Report
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. Glitches in LIGO data may cause a bias in parameter estimation (PE) analyses that can affect the posterior distributions obtained. This project aims to implement inpainting to address this problem in BILBY, one of various PE pipelines used for LIGO analysis. A function has been written and tested to correctly inpaint data. It currently increases the single-likelihood evaluation time by an order of magnitude. Initial PE runs show that inpainting close to merger time may affect the recovery of geocentric time. Future plans include improving the efficiency and testing with different data and configurations.

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. 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) [1]. Much of LSC efforts are dedicated to mitigating the effect of glitches in GW signal searches and parameter estimation (PE) methods.

BILBY is one of various PE pipelines used by the LSC [2]. Like most PE 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 [3].

The BILBY analysis assumes that the noise in the data is stationary and Gaussian. This allows the use of the Gaussian noise likelihood [4]:

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

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 that depends on parameters θ. In practice, Eq. 2 is costly to compute, so the Whittle approximation to the likelihood is used [5]:

$$\mathcal{L}(\mathbf{d}|\theta) \propto \exp\left[-\frac{1}{2}(\mathbf{d}-\mathbf{h}|\mathbf{d}-\mathbf{h})\right] \tag{4}$$

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

$$(\mathbf{a}|\mathbf{b}) \equiv \sum_{j=0}^{\frac{N}{2}-1} 4\,\mathrm{Re}\left(\frac{\tilde{a}_j^* b_j}{S_n(f_j)}\Delta f\right) \tag{5}$$

for a segment with N samples and power spectral density S_n.

A. Inpainting

When glitches are present in the data, they invalidate the fundamental assumption that allows the use of Eq. 2 and Eq. 4. This creates a bias when calculating posterior distributions, especially when glitches are close to a gravitational wave signal [6]. Therefore, segments of strongly non-stationary, non-Gaussian data must be dealt with before these expressions can be used.

A method to address the effect of glitches on data analysis using an inpainting filter was derived in [7] and discussed in the context of PE in [5]. The inpainting filter F is designed to satisfy

$$C^{-1}F\mathbf{d} = 0 \tag{6}$$

in a specified region of data samples, which is known as the *hole*. Unlike gating, which is another popular method for glitch mitigation, inpainting does not introduce artifacts outside the hole once the data is whitened. It therefore allows for minimal data corruption.

The inpainted data can be rewritten as :

$$F\mathbf{d} = \mathbf{d} - \mathbf{d}_{proj} \tag{7}$$

where $\mathbf{d}_{proj}$ is the projection of the data into the overwhitened data space. With this, the Toeplitz system,

$$\mathbf{C}^{-1}\mathbf{d}_{proj} = \mathbf{C}^{-1}\mathbf{d} \tag{8}$$

can be solved for $\mathbf{d}_{proj}$ in the hole region.

Once the inpainted data is obtained, it can be used for Bayesian analysis. Using the inpainted data and waveform in the Whittle approximation (Eq. 4),

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

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

By design of the inpainting filter, none of these terms depend on the data or the waveform inside the hole. Therefore, the glitch should not contribute to the likelihood used in the analysis.

Sec. II discusses the progress that has been made with this project, including the initial training that was completed (Sec. II A), the development of the inpainting function (Sec. II B), initial examinations of its efficiency (Sec. II C), and PE runs that have been completed with the inpainting filter (Sec. II D). Finally, Sec. III explains future plans for this project.

With LIGO's increasing sensitivity, having a glitch mitigation method like inpainting implemented in BILBY for future observing runs will be of great value for astrophysical analyses of gravitational waves.

II. PROGRESS

A. Initial Training

To prepare for the computational tasks necessary for this project, it was important to review some basic concepts related to BILBY, signal filtering, and inpainting. First, short PE runs were completed with BILBY and simple modifications were made to the likelihood calculations. Finite response filters and infinite response filters were also studied to learn about signal processing. Finally, an existing inpainting function from a different software was used to learn about how inpainting has worked in practice outside of parameter estimation.

B. Inpainting function

Following the derivation in [7], a prototype for an inpainting function was written using packages that BILBY already calls to, such as NumPy, SciPy, and GWpy. The function uses the inverse power spectral density of the data to obtain $\mathbf{C}^{-1}$ and the overwhitened data. With this, it solves Eq. 8 for $\mathbf{d}_{proj}$ to obtain the inpainted data.

The performance of this function on a glitch near GW170817 is shown in Fig. 1. As expected, inpainting required less data to be modified to mitigate the glitch than gating.

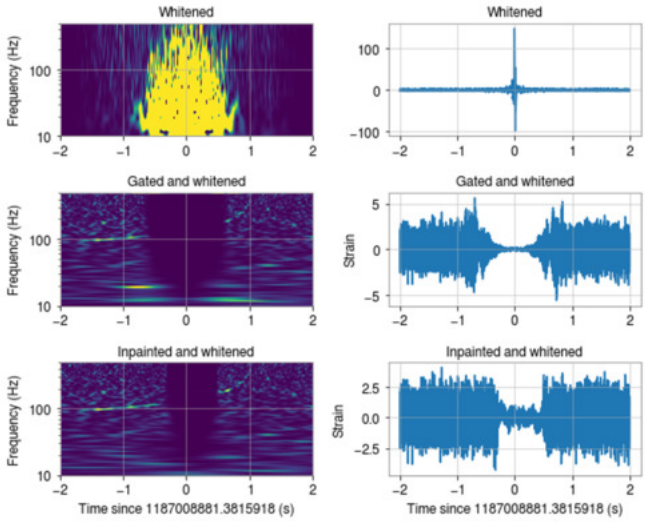

FIG. 1. Q-transforms and time series of data centered on glitch near GW170817. Top panel shows original whitened data with glitch. Middle panel shows data gated with a window of 0.5 s and taper of 0.25 s. Bottom panel shows data inpainted with a window of 0.4 s.

To confirm that the data was being inpainted correctly, the overwhitened data was also plotted. Fig. 2 shows that the overwhitened inpainted data satisfies the requirement of being zeroed in the specified region.

C. Function efficiency

Once the function was verified to work on its own, it was introduced into BILBY's signal-to-noise ratio calculation stage. It was predicted that the inpainting function would increase the run time of a BILBY PE analysis. To study the efficiency of the function, the single-likelihood evaluation time and the calculated likelihood were measured repeatedly for different inpainting windows.

1. *Effect on single likelihood evaluation time*

Fig. 3 shows the single likelihood evaluation times for different inpainting windows (window meaning the half-length in seconds of data to be filtered). This was done by taking 4 seconds of data around GW150914 and applying the filter centered on GPS time 1126259461.4 (one

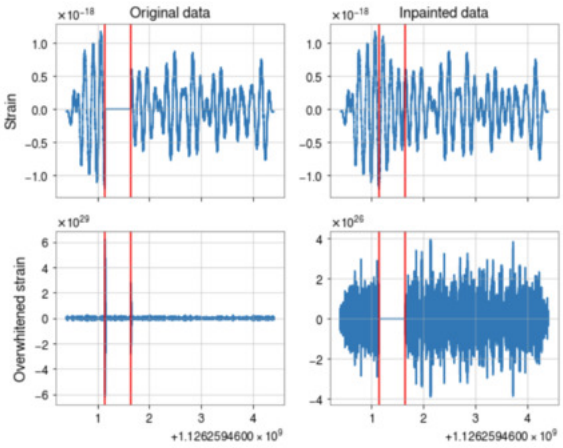

FIG. 2. Original time domain data and inpainted data. Top two panels show raw strain. Bottom two panels show overwhitened strain. Red vertical lines show the region that was set to be inpainted.

second before the merger time). Without the filter, the single likelihood evaluation times were in the order of 10^{-2} seconds. Once the filter was applied for as short as 0.01 seconds, this time increased by a full order of magnitude. As the window was increased, the evaluation time increased until reaching an order of 10^{1} seconds for inpainting the entire 4 seconds of data. Although this increase in run time was to be expected, it is greater than what is aimed for with this project. Future efforts will be dedicated to reducing the run time by studying which lines take the longest to run and considering alternate ways to compute them.

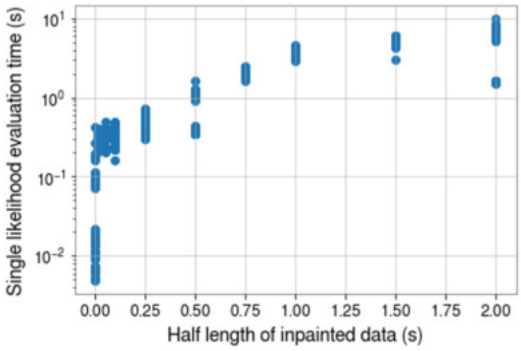

FIG. 3. Single likelihood valuation times for different inpainting windows. Each point represents a single sample in the parameter space.

2. *Effect on likelihood evaluations*

BILBY calculates the first term in Eq. 10, which is referred to as the *noise log-likelihood*, separately from the last two terms, which are collectively referred to as the *log-likelihood ratio*. Currently, the inpainting filter has only been applied to the calculation of the log-likelihood ratio. Fig. 4 shows how the log likelihood ratio was affected by increasing the inpainting window.

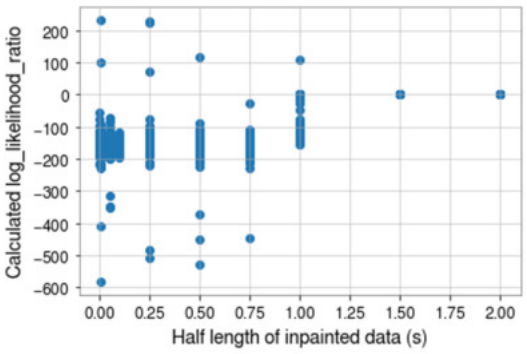

FIG. 4. Calculated log likelihood ratios for different inpainting windows. Each point represents a single sample in the parameter space.

With windows closer to 0, the variance in the likelihoods is greater. This aligns with what is expected of different waveforms being sampled against the data and some being more well-fitting than others. When the entire segment of data is inpainted, there is no signal and the likelihood for all waveforms should equally approach 0, as shown in the plot.

D. PE runs with inpainting

With this preliminary understanding of the efficiency of the initial inpainting function, three PE runs were completed for GW150914 using different inpainting windows and center times. To keep them short, most parameters were fixed. A full list of the values used for each parameter is shown in Table I.

TABLE I. Fixed parameters for GW150914.

Parameter	Value
chirp_mass	31.45
mass_ratio	0.9
a_1	0.0
a_2	0.0
tilt_1	0.0
tilt_2	0.0
phi_12	0
phi_jl	0
dec	-1.2232
ra	2.19432
theta_jn	.89694
psi	0.532268
luminosity_distance	412.066

Posterior distributions were obtained for phase and geocentric time using three different configurations. The

results for the phase are shown in Fig. 5. There does not appear to be significant deviation from the results of the standard analysis for neither of the two runs where the inpainting filter was applied.

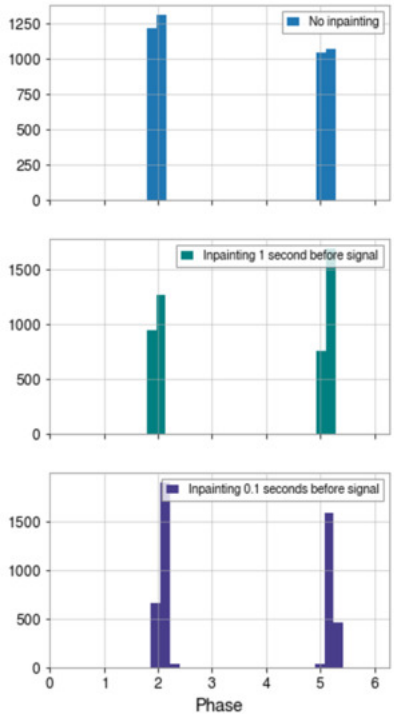

FIG. 5. Posterior distributions obtained for phase. Top panel shows results without the inpainting filter applied. Middle panel shows results for the inpainting filter centered 1 second before merger with window of 0.1 seconds. Bottom panel shows results for the inpainting filter centered 0.1 seconds before merger with window of 0.05 seconds.

Fig. 6 shows the posterior distributions obtained for the geocentric time of the signal. The distribution time for the run with the inpainting filter applied 1 second before the merger time appears to match with that of the standard analysis. However, the distribution for the run with the inpainting filter applied 0.1 seconds before the merger time appears to be slightly shifted right. Geocentric time may be one of the parameters more sensitive to inpainting since bandwidths in the signal that are vital for accurate recovery may be removed with the filter.

III. FUTURE WORK

As mentioned in Section II C 1, one of the present goals is to reduce the run time of the inpainting function. It is predicted that most of the run time is spent performing fast Fourier transforms. This should be verified to ensure that no other operations are unknowingly costing an excess amount of time.

Additionally, various different PE runs will be completed, both on injected and real signals. These will vary in amount of data to be inpainted, distance of inpainting to merger times, presence of glitches, and parameters analyzed. Once the function has been tested for these runs and has improved in efficiency, it will be used to run PE on real signals that have been found near glitches to study how the inpainting filter affects the posteriors obtained.

The long term goal of this project is to have an implementation of inpainting in PE that can be formally added to BILBY.

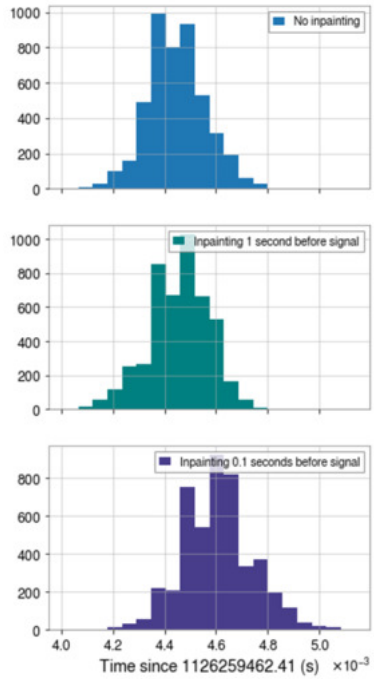

FIG. 6. Posterior distributions obtained for geocentric time. Top panel shows results without the inpainting filter applied. Middle panel shows results for the inpainting filter centered 1 second before merger with window of 0.1 seconds. Bottom panel shows results for the inpainting filter centered 0.1 seconds before merger with window of 0.05 seconds.

IV. ACKNOWLEDGEMENTS

I am extremely thankful for my mentor, Derek Davis, and the guidance they have provided. They have worked hard to make this a rewarding summer for both of us and I look forward to what we will accomplish in the following weeks. I would also like to thank the LIGO team at the California Institute of Technology for being so welcoming and creating a positive work environment. I gratefully acknowledge the National Science Foundation, the Student-Faculty Programs office, and the LIGO Summer Undergraduate Research Fellowship for making this experience possible.

[1] D. Davis and M. Walker, Detector Characterization and Mitigation of Noise in Ground-Based Gravitational-Wave Interferometers, Galaxies **10**, 12 (2022).
[2] 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].
[3] 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].
[4] 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].
[5] 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].
[6] J. Powell, Parameter Estimation and Model Selection of Gravitational Wave Signals Contaminated by Transient Detector Noise Glitches, Class. Quant. Grav. **35**, 155017 (2018), arXiv:1803.11346 [astro-ph.IM].
[7] 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].