

Glitches far from transient gravitational-wave events do not bias inference

Sophie Hourihane^{*} and Katerina Chatziioannou[†]

*TAPIR, California Institute of Technology, Pasadena, California 91125, USA
and LIGOLaboratory, California Institute of Technology, Pasadena, California 91125, USA*



(Received 30 June 2025; accepted 19 September 2025; published 6 October 2025)

Non-Gaussian noise in gravitational-wave detectors, known as “glitches,” can bias the inferred parameters of transient signals when they occur nearby in time and frequency. These biases are addressed with a variety of methods that remove or otherwise mitigate the impact of the glitch. Given the computational cost and human effort required for glitch mitigation, we study the conditions under which it is strictly necessary. We consider simulated glitches and gravitational-wave signals in various configurations that probe their proximity both in time and in frequency. We determine that glitches located outside the time-frequency space spanned by the gravitational-wave model prior and with a signal-to-noise ratio, conservatively, below 50 do not impact estimation of the signal parameters.

DOI: [10.1103/kxbb-dpwp](https://doi.org/10.1103/kxbb-dpwp)

I. INTRODUCTION

The properties of compact binary coalescences (CBCs) observed via gravitational-waves (GWs) [1,2], offer insights on the astrophysical properties of black holes and neutron stars [3] and test General Relativity [4]. Since the first direct detection of GWs [5] the LIGO-Virgo-Kagra (LVK) detector network [6,7] has identified 89 more CBCs [1]. The current fourth observing run has so far tallied a further ~ 200 event candidates [8], for an approximate event rate of once every other day.

Also present in the data are transient, short-duration bursts of power of terrestrial origin, “glitches,” with a rate of 0.5–1.28 per minute *in each detector* [9]. When glitches overlap with a signal, i.e., they are coincident in both time and frequency, they impact the inferred source parameters [10–19]. This is because glitches violate two fundamental assumptions that underlie GW inference, namely that the noise is stationary and Gaussian. In practice, the impact of glitches ranges from biasing inference of subdominant effects like spin-precession [20] or spin-alignment [14], to mimicking an entirely different signal [12,21]. Simply put, glitches make the noise non-Gaussian such that the common Whittle likelihood no longer describes it.

GW inference is based on a time-frequency “analysis window” that is determined by the detector and signal properties; see Appendix E of Abbott *et al.* [1] for details. The window frequency extent is defined by a lower bound, $f_{\text{low}} = 20$ Hz, set by the detector low-frequency sensitivity, and an upper bound, f_{high} , chosen to contain the merger frequency of the $(\ell, |m|) = (3, 3)$ signal mode. The time

extent is designed to enclose the signal from f_{low} to merger (and ringdown) with an additional 2 s of data post-merger [22]. For reference, a typical analysis window for a $30M_{\odot} + 30M_{\odot}$ (detector-frame masses) binary is 4–8 s whereas for a $1.4M_{\odot} + 1.4M_{\odot}$ binary the analysis window is 128–256 s in length.

Every event candidate is vetted for glitches within the analysis window [9,23], and if one is identified, it is subtracted [23]. Notably, this procedure flags glitches anywhere in the analysis window, regardless of whether they overlap, with the signal, i.e., whether they intercept the actual signal time-frequency track [9,23]. For instance, of the 16 events flagged for glitch subtraction in the third LVK observing period [1,2],¹ in only 12 cases did the glitch overlap with the dominant, $(\ell, |m|) = (2, 2)$, signal mode. Since each event requires considerable computer and person time [including vetting, review, run time, etc. [23]] it is desirable to restrict to only cases where mitigation is necessary to avoid biases.

In this study, we explore the conditions under which glitch mitigation can be avoided. Specifically, we address *whether glitches that do not overlap with signals in time-frequency bias inference*. We focus on high-detector-frame mass events, $m \in (20, 100)M_{\odot}$, as glitches disproportionately impact shorter, higher-mass events compared to longer ones containing neutron stars [10]. This is likely due to the fact that inference for short signals hinges on less data, sometimes a single cycle [20,25], and can thus be affected by short glitches. We schematically lay out the potential glitch-signal configurations in Fig. 1 that outlines four (nonexclusive) regions.

^{*}Contact author: sohour@caltech.edu

[†]Contact author: kchatziioannou@caltech.edu

¹An additional 10 were initially flagged by visual inspection, but were ultimately deemed consistent with Gaussian noise [23,24].

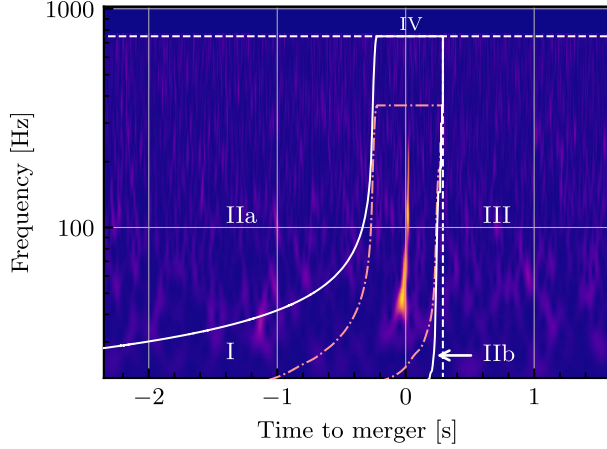


FIG. 1. Time-frequency breakdown of the CBC analysis window, schematically describing the ways in which glitches can be positioned with respect to signals. Region I (solid lines) encloses all time-frequency tracks (including those of higher order modes) within the analysis prior. For reference, in orange we plot the time-frequency content of the (2, 2) mode prior. Region II contains glitches coincident in time and frequency, but never concurrently. We split between Region IIa, those above the (4,4) mode in frequency, and Region IIb, those below the (2,1) mode. Region III contains glitches coincident in frequency, but not in time. Region IV contains glitches not coincident in frequency.

- (i) Region I contains glitches that overlap with *at least one signal within the CBC model prior*. Previous studies [10–19] have shown that such overlaps can lead to biases when not mitigated properly. We confirm these results, primarily using this region for comparison to others.
- (ii) Region II contains glitches that are coincident in time *and* frequency, but not simultaneously. We further distinguish between Region IIa, glitches above the signal, and Region IIb, glitches below the signal (in frequency). Since we are focusing on high-mass events that evolve faster and merge at comparatively lower frequencies than their lower-mass counterparts, Region IIb is very small. We therefore focus on the comparatively much larger Region IIa. We evaluate differences in CBC parameter posteriors from data with and without glitches within this region. Glitches can induce a bias only when they have signal-to-noise ratio (SNR) (conservatively) above 50 *and* they are close in frequency to the CBC merger frequency.
- (iii) Region III includes glitches that share frequency content with the signal, but are not coincident in time. We again simulate glitches within that region and vary the distance (in time) from the signal, as well as the glitch SNR.² Via a standard P-P test,

²We only consider glitches *after* the signal, as those before it are typically cut from the analysis window.

we find that glitches in this region leave no statistical imprint on a population of signals. When considering individual glitches and a GW150914-like signal [5], glitches with $\text{SNR} < 50$ never induce a bias, neither do glitches with $\text{SNR} < 100$ if they occur more than 0.5 s after the signal.

- (iv) Region IV contains glitches that do not share any frequency content with the event. We analytically show that such glitches do not impact inference.

The rest of the paper is organized as follows. In Sec. II we recap current glitch mitigation techniques and lay out the noise assumptions that are the foundations for CBC parameter estimation. In Sec. III we lay out the methodology of our study. In Sec. IV we go through each Region in Fig. 1 and present our results. In Sec. V we conclude.

II. BACKGROUND

In this section we recap glitch mitigation in Sec. II A and standard inference under Gaussian noise in Sec. II B.

A. Glitches and glitch mitigation

During the third observing run, glitches occurred more than once every minute in each detector [1,26]. Of the total of 79 events with an astrophysical probability greater than 0.5, 16 contained glitches within their analysis window. In 4 of 16 glitch-mitigated events, the glitch did *not* overlap the signal time-frequency track [1,2,22,23]. Assuming a fixed glitch rate between 1 and 1.5 glitches per minute, the number of signals requiring glitch mitigation during the fourth LVK observing run is expected to increase simply due to the increased event rate. Adopting a 4 s analysis window and two detectors, we expect 26–40 (13–20%) of the 203 current candidates to contain glitches within their analysis window [8]. The increased demand for glitch mitigation motivates our detailed look into the conditions that require it.

The amount of glitch-signal overlap will determine the degree to which inference is biased. A glitch and a GW signal overlap in frequency when, under a Fourier decomposition, there are frequencies, f_i , for which they are both nonzero. A glitch and a signal overlap in time when there are times, t_i , for which they are both nonzero. A glitch and a signal overlap in time-frequency if, when decomposed into time-frequency space, there are bins with nonzero content from both. Therefore a glitch and signal can overlap in both time and frequency, but *not* overlap in time-frequency. These definitions apply even for GWs with higher-order modes, for which there is no 1:1 relationship between time and frequency. In this case, we consider all CBC modes $(\ell, |m|)$ to determine overlaps, e.g., Fig. 1.

Currently, if a glitch cannot be excluded from the analysis by redefinition of the analysis window [subject to the established criteria [1]], it is subject to mitigation. Common mitigation methods include the following:

(i) subtracting a single estimate of the glitch from the data, a process known as “glitch subtraction” [23,27], (ii) modeling both the signal and the glitch and thus marginalizing over the uncertainty of both models [10,11,14,28], and (ii) removing all affected data by zeroing [29] or by replacing with Gaussian noise [30]. Glitch subtraction requires some (time or frequency) reconstruction of the glitch. Some glitch classes can be described with physical models such as slow and fast scattering [31–33]. Another option is to utilize auxiliary data channels that “witness” noise. By determining the transfer function between these witness channel(s) and the measured strain data, it is possible to subtract the associated glitch [27], but only if such channels exist. In the absence of physically-motivated models or witness channels, most glitches are targeted phenomenologically with `BayesWave` [34,35], further described in Sec. III A.

B. Gaussian noise likelihood

The noise assumptions underlying GW data analysis determine the form of the likelihood [36]. The data, $\mathbf{d}$, are a combination of a GW signal $\mathbf{h}$ and noise $\mathbf{n}$, which is a sum of Gaussian, $\mathbf{n}_G$, and transient, non-Gaussian noise (glitches), $\mathbf{g}$. All quantities are considered in the frequency domain. For stationary $\mathbf{n}_G$, the noise is uncorrelated between frequency bins, meaning that the frequency domain noise-covariance matrix is diagonal. Gaussianity means that the noise is described by a Gaussian distribution at each frequency. The per-frequency Gaussian distribution is then entirely described by the variance, leading to a frequency-domain noise-covariance matrix that is proportional to the noise power spectral density (PSD), $\mathbf{S}_n$.

The likelihood for the GW model with parameters θ , $\mathbf{h}_\theta$ in a single detector is then [36]

$$\mathcal{L}(\mathbf{d}|\theta) = \exp\left(-\frac{1}{2}|\mathbf{d} - \mathbf{h}_\theta|^2 - \sum_{f_i} \ln(2\pi S_{ni})\right), \quad (1)$$

where $|\mathbf{a}|$ is the noise-weighted magnitude

$$|\mathbf{a}| = \sqrt{(\mathbf{a}|\mathbf{a})}, \quad (2)$$

$(\mathbf{a}|\mathbf{b})$ is the noise-weighted inner product

$$(\mathbf{a}|\mathbf{b}) = 4\Delta f \sum_{f_i=f_{\text{low}}}^{f_{\text{high}}} \frac{a_i b_i^*}{S_{ni}}, \quad (3)$$

and Δf is the frequency resolution. The likelihood across all detectors is the product of each individual-detector likelihood.

III. METHODS

In this section, we describe our methodology. We simulate data containing CBC signals and glitches and analyze them ignoring the presence of the glitch. In Sec. III A we introduce the `BayesWaveCpp` analysis package that is used to sample from the posterior for the CBC parameters, Sec. III A 1, and to simulate glitches, Sec. III A 2. We then introduce “glitch reweighting” in Sec. III B as a method to (i) quickly generate posteriors and (ii) to quantify the degree of similarity between two probability distributions, specifically those obtained from data with and without glitches. We introduce the Jensen Shannon divergence in Sec. III C as another quantity with which to compare two posteriors.

A. BayesWaveCpp

`BayesWaveCpp` [37] is a rewrite of and upgrade to `BayesWave` [34,35,38], a software package used to stochastically sample the posteriors of signals, glitches, and noise PSDs in GW data. GWs can be modeled through coherent sums of sine-Gaussian wavelets, physical CBC waveform models, and combinations thereof. Glitches are modeled as sums of sine-Gaussian wavelets. The noise PSD is modeled via broadband splines and Lorentzians for the spectral lines. For the purposes of this study, `BayesWaveCpp` is only used to sample the CBC posterior under a CBC waveform model. Though it has the capacity to do so, we do not sample the glitch or PSD posterior, assuming no glitch for the former and a known PSD for the latter.

1. CBC model

In this study, we model all CBC signals with `IMRPhenomXPHM` [39], an inspiral-merger-ringdown model that includes precession and higher-order modes, but does not include eccentricity.³ The CBC parameters and their priors are listed in Table I; here we briefly describe some parameters of interest that feature in the following figures. The total mass, M , is the sum of the component masses of the compact binary in the detector frame. The spin angular momentum of the compact binary system can be summarized with “effective” parameters: the effective precessing parameter χ_p [Eq. (3.4) in [40]] and the effective spin-aligned parameter χ_{eff} [Eq. (2) in [26]]. Both parameters have been shown to be susceptible to biases due to the presence of glitches in real data [14,20].

2. Simulated glitches

We simulate glitches using `BayesWaveCpp`’s glitch model for convenience, which consists of sums of sine-Gaussian Morlet-Gabor wavelets. Such wavelets constitute an over-complete basis over a smooth function space and are thus

³The ability to model precession is novel to `BayesWaveCpp`, whereas `BayesWave` is restricted to spin-aligned waveforms.

TABLE I. Parameter definition, notation, and prior distribution for all CBC parameters. Here $i \in \{1, 2\}$ indexes the compact objects; $i = 1$ (2) is the larger (smaller) mass object. The time in the Hanford detector is centered 2 s before the end of the 4 s analysis window.

Parameter	Symbol	Prior
Mass	m_i	$U[20, 100] M_\odot$
Spin amplitude	χ_i	$U[0, 1]$
Spin in-plane angle	ϕ_i	$U[0, 2\pi]$
Spin polar angle	θ_i	$\cos \theta_i \sim U[-1, 1]$
Hanford time	t_{LHO}	$U[0.05, 0.05] \text{ s}$
Luminosity distance	D_L	$D_L^3 \sim U[1, 10000^3] \text{ Mpc}^3$
Inclination	ι	$\cos \iota \sim U[0, 1]$
Right ascension	α	$U[0, 2\pi]$
Declination	δ	$\sin \delta \sim U[-1, 1]$
Polarization	ψ	$U[0, 2\pi]$
Coalescence phase	ϕ	$U[0, 2\pi]$

flexible enough to mimic most noise transients in GW data. In addition to each of the five parameters describing each wavelet (time, frequency, quality factor, amplitude, phase), the number of wavelets itself can be varied. We adopt the `BayesWaveCpp` glitch prior as the distribution from which we draw glitches to simulate, listed in Table II.

The use of Morlet-Gabor wavelets allows us to quantify the glitch support in time and frequency space, e.g., Fig. 1. Each wavelet is characterized by a central time t_0 and a central frequency f_0 , around which its power decreases, forming a Gaussian envelope in both time and frequency. In the time domain, the decay has an e-folding time scale of

$$\tau = \frac{Q}{2\pi f_0}, \quad (4)$$

where Q is the wavelet quality factor. Similarly, in the frequency domain, the decay e-folding frequency scale is $1/\tau$. So, a wavelet that is well-localized in time will be poorly localized in frequency and vice-versa. We use these

TABLE II. Parameter definition, notation, and prior distribution for the glitch model, used to simulate glitches. Here ρ_i is the approximate SNR of the wavelet in Eq. (12) in Ref. [34]. This distribution is only used to simulate glitches. The wavelet central time distribution is centered at $\{0, 0.25, 0.5, 1\}$ s after the center of the CBC time prior.

Parameter	Symbol	Prior
Dimension	D_g	$U[1, 10]$
Wavelet central time	t_0	$U[-0.1, 0.1] \text{ s}$
Wavelet central frequency	f_0	$U[16, 1024] \text{ Hz}$
Wavelet quality factor	Q	$U[0.1, 40]$
Wavelet amplitude	A	$\rho_i \in [1, 100]$
Wavelet phase	ϕ	$U[0, 2\pi]$

estimates to approximately place the glitch with respect to the signal in subsequent sections.

In reality, the majority of LIGO glitches cannot be described by a *single* wavelet. Regardless, we expect our main conclusions to apply to more complex morphologies. First, within `BayesWave` glitches are described as a sum of wavelets. Studying the impact of a single wavelet therefore serves as a baseline. Second, the choice to use a single wavelet is driven by practical considerations: we can estimate its time-frequency footprint analytically with f_0 and Eq. (4). More generically, regardless of the exact glitch and signal morphologies, the amount of bias is determined by how much the glitch overlaps with the signal in time-frequency, and wavelets allow for a simple quantification of this.

B. Glitch reweighting

We quantify the impact of glitches on the CBC posterior by reweighting the posterior from data *without* to data *with* a glitch. Equal-weighted samples from one (reference) probability distribution can be “reweighted” to approximate another (target) probability distribution; this is a form of importance sampling called “reweighting”. The exact formalism for reweighting is laid out in Sec. II of Hourihane *et al.* [41], which we adopt here. Reweighting provides a metric to compare how similar or disparate the two distributions are via the efficiency,

$$\mathcal{E} = \frac{n_{\text{eff}}}{N_s}, \quad (5)$$

where n_{eff} is the effective number of samples [Eq. (11) in Ref. [41]] after reweighting and N_s is the original number of samples. Distributions with an efficiency of 1 are identical, whereas disparate distributions have an efficiency of 0.

In this work we use reweighting in a somewhat novel way. While reweighting has been employed to change the likelihood function [41], the model [42–46], the prior, or to evaluate posteriors from a neural network approximate [47], here we change the *data*. That is, we use reweighting to transform a CBC posterior distribution in Gaussian noise, $p(\theta|\mathbf{d}_{\text{NG}})$, to a posterior distribution of identical data plus a glitch, $p(\theta|\mathbf{d}_{\text{G}})$. The (unnormalized) log weights are

$$\begin{aligned} \ln w(\theta) &= \ln \mathcal{L}(\mathbf{d}_{\text{G}}|\theta) - \ln \mathcal{L}(\mathbf{d}_{\text{NG}}|\theta) \\ &= -\frac{1}{2}|\mathbf{d}_{\text{NG}} + \mathbf{g} - \mathbf{h}_\theta|^2 + \frac{1}{2}|\mathbf{d}_{\text{NG}} - \mathbf{h}_\theta|^2 \\ &= (\mathbf{g}|\mathbf{h}_\theta) - C_g, \end{aligned} \quad (6)$$

where $C_g = (\mathbf{g}|\mathbf{g})/2 + (\mathbf{d}|\mathbf{g})$ is constant.

Reweighting has a number of advantages. First, it provides a sensitive, direct estimate of the difference

between two probability distributions via the efficiency, Eq. (5).⁴ Second, reweighting bypasses expensive stochastic sampling and allows us to consider a wide range of glitches from a single glitch-free set of samples. Third, reweighting is not subject to sampling uncertainty when stochastically sampling from a posterior. That is, $\mathcal{E} \neq 1$ if and only if the standard deviation of the likelihood over the posterior changes between the approximate and the target distribution. Since glitches far away from the signal have a small impact on the posterior, we expect reweighting to be a particularly effective method for our purposes.

Such glitch-reweighting is possible on simulated data because we have access to glitch-free data. It is still possible to use glitch-reweighting on real data if there is a model for the glitch; for instance, the model used in glitch subtraction. The efficiency of such post-facto glitch-reweighting would serve as an estimate of the impact of that glitch on inference. In cases where the efficiency is low, it might be important to consider the uncertainty of the glitch model, e.g. [10,14].

C. Jensen Shannon divergence

We also quantify the difference between two posteriors using the Jensen Shannon (JS) divergence [48]; see Appendix of Abbott *et al.* [26] for more details. The JS divergence between probability distributions p and q , $D_{\text{JS}}(p, q)$, is smoothed, normalized, and symmetrized. We adopt the threshold of 0.007 bit [26]. For a Gaussian, this corresponds to a 20% shift in the mean, which is larger than the 0.002 bit expected between two identical posteriors from sampling uncertainty alone [49]. We compute the JS divergence of the 15 marginalized, 1-dimensional CBC posteriors and report the maximum:

$$\max_{\theta} D_{\text{JS}}^{\theta} = \max_{i \in 1 \dots 15} D_{\text{JS}}[p(\theta_i | \mathbf{d}_G), p(\theta_i | \mathbf{d}_{\text{NG}})]. \quad (7)$$

IV. RESULTS

In this section we study how glitches impact CBC inference in each region of Fig. 1.

A. Region IV, glitches above the signal

When the glitch and signal do not overlap in frequency, the CBC posterior is unaffected by the presence of the glitch, even when the glitch frequency is part of the analysis bandwidth. The proof presented below is based on the fact that the noise-covariance is diagonal in the frequency domain, and thus the likelihood is a 1-dimensional integral in frequency. Consider a GW signal only containing frequency content up to frequency F , coincident in time

with a glitch that contains only frequency content above F . Then $(\mathbf{g} | \mathbf{h}) = 0$. If additionally the frequency content of $h_{\theta}(f > F) = 0$ for *all* θ in the CBC prior, π , it follows that, for all $\theta \in \pi$, the likelihood is independent of any glitch-signal cross-terms. That is,

$$\begin{aligned} \mathcal{L}(\mathbf{d} | \theta) &\propto \exp \left[-\frac{1}{2} (|\mathbf{d}_{\text{NG}} + \mathbf{g}|^2 + |\mathbf{d}_{\text{NG}} - \mathbf{h}_{\theta}|^2 - 2(\mathbf{g} | \mathbf{h}_{\theta})) \right] \\ &= \exp \left[-\frac{1}{2} (|\mathbf{d}_{\text{NG}} + \mathbf{g}|^2 + |\mathbf{d}_{\text{NG}} - \mathbf{h}_{\theta}|^2) \right]. \end{aligned} \quad (8)$$

The $|\mathbf{d}_{\text{NG}} + \mathbf{g}|^2$ term is constant and does not depend on θ .⁵ Since the likelihood is not normalized in θ , *the extra glitch term does not change the shape of the posterior*. It will, however, scale the evidence, the expected value of the likelihood over the prior.

B. Region III, glitches after the signal

We now turn to glitches occurring after a GW event that do not overlap with the GW in time. The proof of Sec. IV A—glitches that do not overlap in frequency do not bias inference—no longer applies. This is because the noise-covariance is no longer diagonal in the time domain, and distinct times are correlated; see Sec. II B. Even if a glitch and signal do not overlap in time, the correlations between their times *could* induce a bias. We study this bias for a population of signals and glitches in Sec. IV B 2, and for individual glitches on a GW150914-like event in Secs. IV B 3 and IV B 4.

1. Percentile-percentile test

We first introduce the percentile-percentile (P-P) test, a standard method to determine if an ensemble of posteriors are statistically robust [50,51]. In a P-P test, a set of simulations is generated by drawing parameters from the prior distribution of a model. The posterior is then computed with the same model and prior. If the prior reflects the underlying population (including the noise model) and the posteriors have the correct statistical coverage, the percentiles of the true parameter values within their respective posteriors should follow a uniform distribution. Consequently, the cumulative distribution function (CDF) of these percentiles should form a diagonal line with a slope of 1.

We perform this test with 400 simulated CBC signals (without glitches) in Gaussian noise with parameters drawn from the CBC prior of Table I. The leftmost panel of Fig. 2 shows the CDF for each of the 15 CBC model parameters. All parameters lie within the 3-sigma confidence region (indicated by the gray outline) as expected. This demonstrates that the BayesWaveCpp CBC sampler is unbiased under the conditions of the P-P test.

⁴Another option is to use the Kullback-Liebler divergence which can likewise be computed from the weights. However, efficiency is normalized so we prefer it.

⁵In analyses where the glitch is also sampled, this term would no longer be constant.

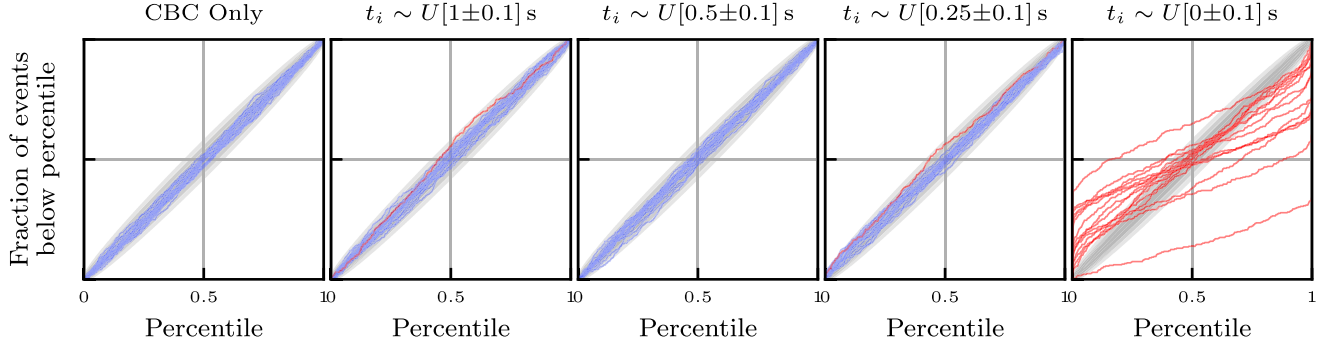


FIG. 2. Percentile-percentile (P-P) plots for various simulations, each drawn from the same CBC prior (Table I) but varying in glitch content (Table II). The titles of each plot specify the relative time between the glitch distribution and the CBC time distribution, with the leftmost plot representing data without glitches and subsequent plots showing glitches progressively closer to the CBC. Each plot comprises 400 simulations, with recovery performed using only a CBC model. Each plot includes 15 lines, one for each CBC parameter, displaying the cumulative distribution function of the percentiles of the true values within their marginal posteriors. Lines are color-coded in red (blue) to indicate whether the parameter failed, $p \leq 0.05$ (passed, $p > 0.05$) the P-P test. A failure rejects the null hypothesis: the percentiles of the true values are uniformly distributed across their posteriors. Three-sigma confidence intervals are plotted in gray. Left: P-P plot when only the CBC model is simulated and recovered. All parameters pass the P-P test which serves as a baseline for the test and its implementation. Center (left to right): P-P plots for glitches in Region III. Right: P-P plot for glitches in Region I. When glitches overlap with the signals, all parameters fail the P-P test.

2. Varying distance between the glitch and signal

To assess the impact of glitches, we perform a variation of the standard P-P test described in Sec. IV B 1. Instead of pure Gaussian noise, the simulated data now also contain glitches drawn from the distribution of Table II; their SNR follows their prior which peaks at SNR 5 and is bounded between 1 and 100. The glitch time distribution is progressively moved closer to the signal. Still, just as in Sec. IV B 1, the posterior is computed only over the CBC model, leaving the glitch unaccounted for. Five scenarios are compared in Fig. 2: a control case with no glitches, and four cases with glitches whose time distributions have a width of 0.2 and are centered at 1, 0.5, 0.25, and 0 s after the center of the CBC time prior. All simulations are performed with a single, LIGO Hanford (LHO) detector (to maximize the impact of glitches), and the Advanced LIGO sensitivity [6]. By leaving the glitch unaccounted for during recovery and varying the timing of the glitch relative to the CBC, we explore how the timing of glitches impacts inference.

Results are shown in Fig. 2. The leftmost panel displays the control case with no glitches; all parameters pass the P-P test, indicating unbiased recovery. The center three panels (from left to right) depict results on data containing glitches after, but progressively closer to, the CBC signals. Recovery remains unaffected, with most of the 15 CBC parameters passing the test for each distance. A few cases with $p < 0.05$, are not unexpected given the large number of tests performed. In contrast, the rightmost panel, corresponding to glitches coincident in time with the CBC signals, reveals that all parameters fail the P-P test, highlighting significant biases. These results suggest that glitches occurring after a GW signal (Region III) have minimal impact on recovery for an ensemble of signals.

However, glitches coincident with the signal (Region I) severely bias the recovered parameters, confirming previous results [10–19].

3. Can very loud glitches bias inference?

We now consider a single instance of a CBC and a glitch and study at which SNR a glitch can create a bias. We compare two datasets, again in a single LHO detector:

- (1) Control Dataset: A simulated GW150914-like signal with a network SNR of 25, injected into Gaussian noise. The CBC parameters are detailed in Table III.

TABLE III. Simulated CBC parameters for Secs. IV B 3, IV B 4, and IV C 2.

Parameter	Value
Masses	$m_1 = 38.2M_\odot$, $m_2 = 32.9M_\odot$
Spin amplitudes	$\chi_1 = 0.998$, $\chi_2 = 0.126$
Spin in-plane angle	$\phi_1 = 0.55\pi$ rad $\phi_2 = 0.364\pi$ rad
Spin polar angle	$\theta_1 = 0.14\pi$ rad $\theta_2 = 0.49\pi$ rad
Hanford time	1126259462.424 s 2.42455 s (in segment)
Luminosity distance	$D_L = 415.66$ Mpc
Inclination	$\iota = 0.88\pi$ rad
Right ascension	$\alpha = 0.35\pi$ rad
Declination	$\delta = -0.36\pi$ rad
Polarization	$\psi = 0.42\pi$ rad
Coalescence phase	$\phi = 1.30\pi$ rad

TABLE IV. Simulated glitch parameters used in Sec. IV B 3.

Parameter	Value
Glitch dimension	$D_g = 1$
Wavelet central time	$t_1 = 2.38$ s
Wavelet central frequency	$f_1 = 100$ Hz
Wavelet quality factor	$Q_1 = 28.73$
Wavelet amplitude	A_1 : Varied such that $\rho_1 \in \{5, 10, 50, 100, 500, 1000, 5000\}$
Wavelet phase	$\phi_1 = 1.61\pi$ rad

- (2) Glitch-impacted Dataset: The glitch-impacted dataset is identical to the control one in terms of the CBC and Gaussian noise realization but includes seven distinct glitch configurations. The glitches share all parameters, Table IV, but SNR which we incrementally increase from 5 and 5000.

We analyze both datasets with BayesWaveCpp, modeling *only* the CBC signal, *leaving again the glitch power entirely unmodeled*. Results for the control dataset serve as a baseline for comparison; if the glitches have no impact, the posteriors from the glitch-impacted dataset will match those from the glitch-free dataset. We compute glitch-impacted posteriors both with direct sampling (BayesWaveCpp) and via reweighting from the control to the glitch-impacted dataset; see Sec. III B. If the glitches have no impact, then the reweighting efficiency will be 1.

We present the results in Fig. 3. The top panel shows the whitened time-domain waveforms of the CBC (magenta) and the glitches (various colors) for reference. Below that, we plot the reweighting efficiency and posteriors for select CBC parameters (total mass, χ_{eff} , and χ_p) as a function of glitch SNR. The direct-sampled, glitch-impacted posteriors are colored, and the corresponding reweighted posteriors are plotted with black dashed lines. The fiducial, glitch-free posterior is displayed in gray. For glitch SNR ≤ 500 ,⁶ we find that not only are the glitch-impacted posteriors *visually* indistinguishable to the glitch-free one, but also they are *actually* identical; the reweighting efficiency between the two distributions is 100%. As the glitch SNR is increased, the efficiency does drop. At glitch SNR of 1000 the posteriors are still visually identical, but the efficiency is 80%, suggesting it is a more sensitive test than visual comparison (it also depends on the full 15-dimensional posterior and not select marginal ones). By a glitch SNR of 5000, the efficiency is zero, signaling severe biases.

⁶For reference, the glitch on the binary neutron star merger, GW170817, was SNR ~ 800 [12].

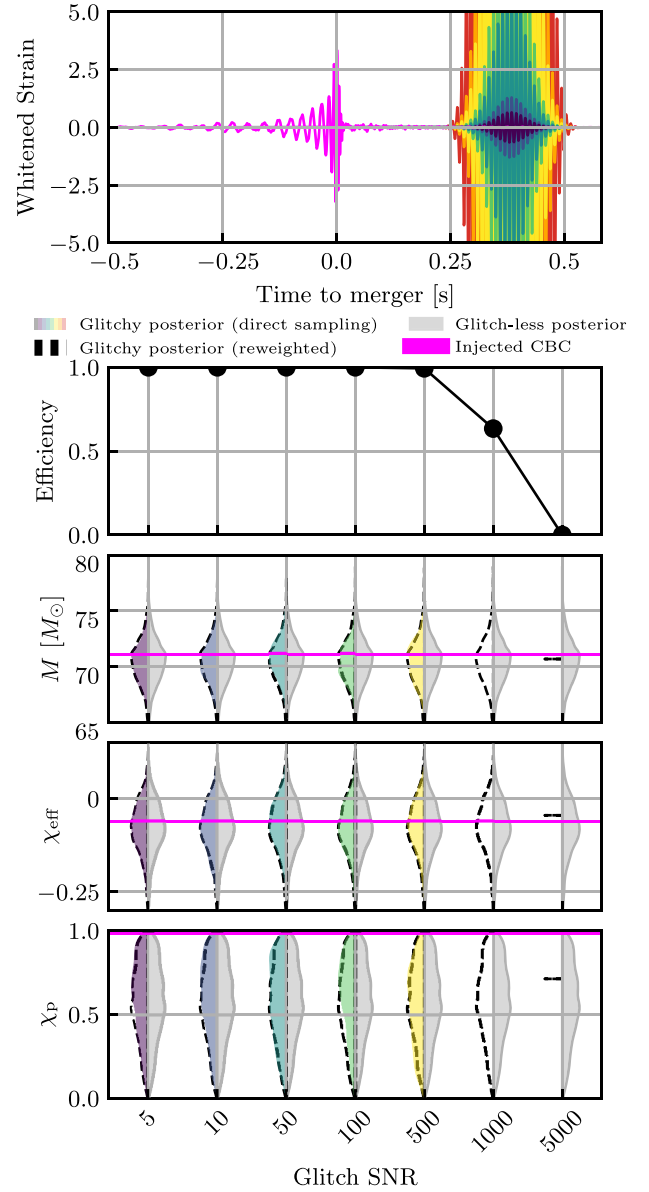


FIG. 3. Top: Time domain whitened waveforms for the CBC (magenta) and a glitch with increasing SNR from 5 to 5000 (various colors) 0.38 s after the signal, see Table III and Table IV for details. Second down: Efficiency when reweighting from a posterior on data with no glitch to a posterior on data with a glitch as a function of the glitch SNR. Bottom 3: 1-dimensional posteriors for select CBC parameters as a function of the glitch SNR. True values are marked in magenta. The direct-sampled posterior on glitch-impacted data is colored and the corresponding reweighted posterior is marked with black dashed lines. To the right we show the control dataset in gray, a posterior recovered from data with identical Gaussian noise but without any glitch; all gray posteriors are identical. For glitch SNR ≥ 1000 we omit the direct-sampled posteriors due to nonphysical waveform behavior, further discussed in the Appendix. The efficiency indicates that glitches louder than SNR 500 start impacting inference.

TABLE V. Parameter distribution of simulated glitches in Sec. IV B 4, additionally restricted to Region III, and in Sec. IV C 2, likewise restricted to Region IIa.

Parameter	Value
Glitch dimension	$D_g = 1$
Wavelet central time	$t_1 \sim U[0, 4] \text{ s}$
Wavelet central frequency	$f_1 \sim U[20, 700] \text{ Hz}$
Wavelet quality factor	$Q_1 \sim U[0.1, 40]$
Wavelet amplitude	A_1 : Varied
Wavelet phase	$\phi_1 \sim U[0, 2\pi]$

4. At what glitch SNR can we expect a bias?

Finally, we consider a random collection of 200 glitches per Table V in Region III after the same CBC signal from Table III. Since we expect glitch-induced biases to be low, we omit direct sampling, and instead only compute posteriors via reweighting. Since the weights are proportional to the glitch SNR, per Eq. (6), the reweighted posterior for a single glitch can be trivially scaled in SNR. We leverage this for each of the 200 glitches to compute the maximum JS divergence over all CBC parameters, and determine the glitch SNR at which it exceeds a JS threshold.

Results are displayed in Fig. 4. The time-frequency locations of the 200 simulated glitches are displayed in the leftmost panel, corresponding to three e-folds of exponential decay in each direction. Boxes are colored by the SNR

at which that glitch induces a posterior with a JS divergence above 0.007 compared to glitch-free data. Glitches are colored gray if the requisite SNR is above 500. On the bottom right we plot the maximum (among all 15 CBC parameters) JS divergence between the glitch-impacted and glitch-less datasets, plotted as a function of glitch SNR. Each curve corresponds to a glitch in the left figure, colored in the same manner. The top right shows the cumulative distribution of the number of glitches that induce an above-threshold divergence, plotted as a function of glitch SNR.

The majority of glitches do not incur a bias unless their SNR is greater than 500. However, glitches closer to the signal (per the left plot, within 0.5 s after merger) can incur a bias at SNR as low as ~ 56 . Those are rare (2% at SNR 100 and 17% at SNR 500) and will likely need mitigation. Even at SNR 1000, fewer than 40% of the glitches induced a bias greater than one would expect from stochastic sampling. Overall, *no glitches below SNR 50 in Region III induce a measurable bias, no matter how close to (but confidently after) the signal.*

C. Region II, glitches sharing time and frequency content but not concurrently

Now we move to glitches that overlap with the GW signal (or its prior) in either time or frequency, but not concurrently. These are glitches “just above,” Region IIa, or “just below,” Region IIb, the signal. No pure time- or frequency-domain analysis can exclude those with analysis window redefinition.

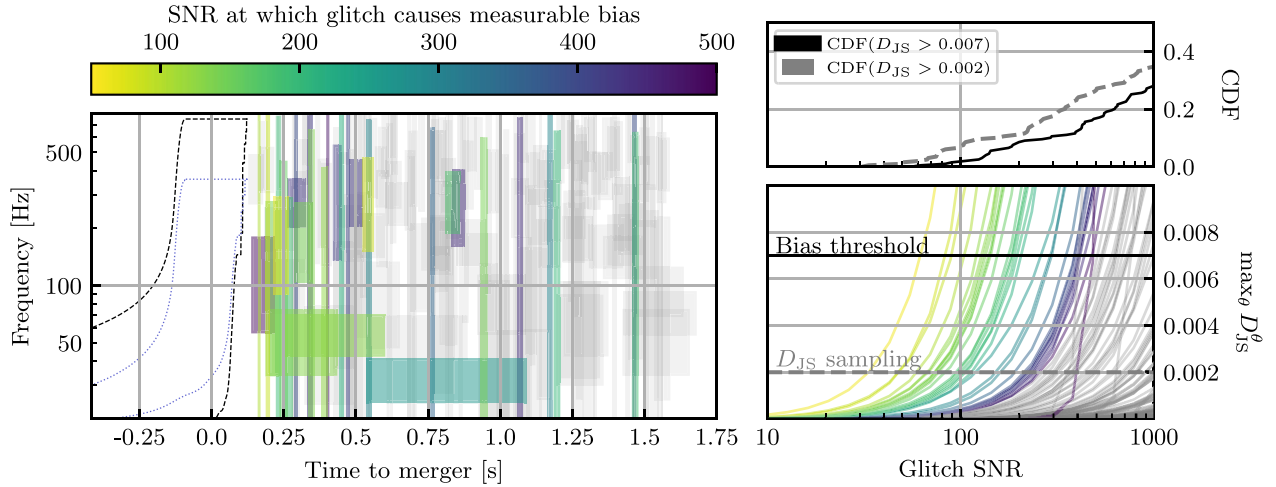


FIG. 4. Left: time-frequency locations of 200 glitches simulated in Region III; after the signal. The glitch time-frequency locations are colored by the SNR at which that glitch induces a measurable change in the CBC posterior and are gray if the requisite SNR is above 500. For reference, Region I (black-dashed) and the time-frequency content of the (2,2) mode across the CBC prior (blue dots) are displayed. Bottom right: maximum Jensen-Shannon divergence across CBC parameters between a CBC posterior on glitch-impacted data and the posterior on glitch-free data, plotted as a function of glitch SNR. Each curve corresponds to a glitch in the left figure. The black horizontal line is the threshold for posteriors considered distinct [26]. The gray dashed line is the JS divergence due only to stochastic sampling uncertainty [49], plotted for reference. As the glitch SNR increases, so does D_{JS} . Top right: cumulative distribution function of the number of glitches that induce a requisite divergence as a function of glitch SNR. Glitches below SNR 50 never induce a measurable bias. Higher-SNR glitches might induce a bias if within 0.5s from the signal merger.

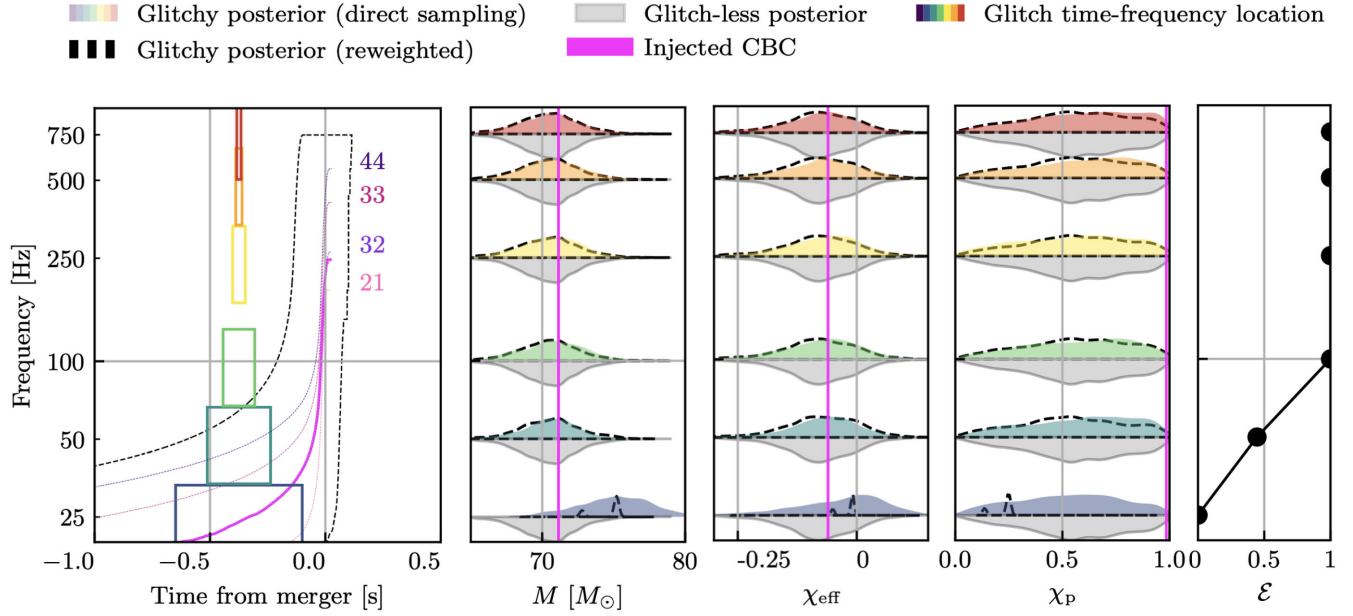


FIG. 5. Inference results on a GW150914-like signal with SNR 25 glitches of increasing frequency. Left: time-frequency track of the GW150914-like signal with glitches overlaid. Region I is outlined in black-dashed lines; it includes the time-frequency content of the entire CBC model prior. The solid magenta line shows the (2,2) frequency track of the simulated CBC and the dotted lines display the higher-order modes, each labeled with its corresponding color. The location in time-frequency space of the glitches is shown in a colored box, each displaying three e-folds of the exponential time and frequency glitch decay. The only glitches that share time-frequency content with the GW model are the ones at 25 and 50 Hz. Center: 1-dimensional posteriors for select CBC parameters from data with a glitch at the corresponding frequency (y-axis and color). In the lower half of each violin plot we show the posterior from identical data but without a glitch (gray); all such posteriors are identical. On the top of each violin plot are the posteriors recovered in glitch-impacted data; the colored posterior are those recovered with BayesWaveCpp (direct sampling), and the black-dashed lines are the posterior obtained via reweighting. Right: reweighting efficiency as a function of glitch frequency. Where the posteriors differ the most, at glitch frequency 25 Hz, the efficiency drops to 0. At 50 Hz, even though the posteriors are visually similar, the efficiency also dips to 50%, meaning that the glitch is nonetheless impacting the posterior.

1. Region IIa: Glitches just above the signal

We follow the methodology of Sec. IV B 3 and consider two simulated datasets:

- (1) Control Dataset: identical to that in Sec IV B 3, with a GW150914-like signal given in Table III in Gaussian noise.
- (2) Glitch-Impacted Dataset: identical to the control dataset in terms of CBC and Gaussian noise, but it includes six distinct glitch configurations. All glitches have an SNR of 25 and are centered ~ 13 cycles before merger; further parameters are given in Table VI. The frequency of the glitch is

TABLE VI. Simulated glitch parameters in Sec. IV C 1.

Parameter	Value
Glitch dimension	$D_g = 1$
Wavelet central time	$t_1 = 2.04$ s
Wavelet central frequency	$f_1 \in \{25, 50, 100, 250, 500, 750\}$ Hz
Wavelet quality factor	$Q_1 = 28.73$
Wavelet amplitude	A_1 : Varied such that $\rho_1 = 25$
Wavelet phase	$\phi_1 = 1.61\pi$ rad

incrementally increased, starting with one that overlaps with the CBC's (2,2) mode and continuing until the glitch is outside the entire CBC model prior.

The left panel of Fig. 5 displays the time-frequency tracks for the modes of the injected CBC signal (colored lines), overlaid with the time-frequency support of the simulated glitches (colored boxes). The solid magenta line represents the dominant (2,2) frequency track, while the dashed lines indicate higher-order modes. The dashed black line encloses the CBC model prior, including all higher-order modes. The glitches at 25 and 50 Hz overlap with the time-frequency GW prior, but only the glitch at 25 Hz overlaps with the (2,2) mode.

As always, we analyze the data leaving the glitch unmodeled. In the central panels of Fig. 5 we plot select marginalized CBC posteriors corresponding to each glitch (same color as the boxes) as well as the fiducial, glitch-less posterior (gray). The y-axis corresponds to the glitch central frequency. The black, dashed lines are again the glitch-reweighted distributions which should be identical to the colored distributions (for sufficiently high efficiency). Above 25 Hz, all distributions agree on visual inspection for all parameters. Glitches that do not overlap with the

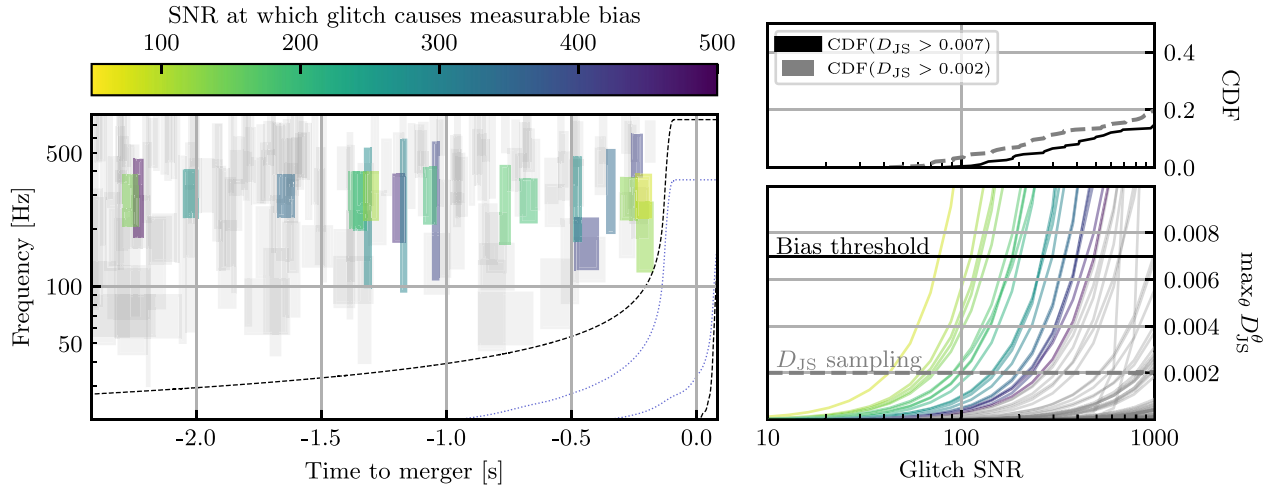


FIG. 6. Left: time-frequency locations of 200 glitches simulated in Region IIa. The glitch time-frequency locations are colored by the SNR at which that glitch induces a measurable change in the CBC posterior and are gray if the requisite SNR is above 500. For reference, Region I (black-dashed) and the time-frequency content of the (2,2) mode across the CBC prior (blue dots) are displayed. Bottom right: maximum Jensen-Shannon divergence across CBC parameters between a CBC posterior on glitch-free data and the posterior on glitch-impacted data, plotted as a function of glitch SNR. Each curve corresponds to a glitch in the left figure, and are colored in the same manner. The black horizontal line is the threshold for posteriors considered distinct [26]. The gray dashed line is the JS divergence expected to arise due to stochastic sampling uncertainty [49], displayed for reference. As the glitch SNR increases, so does D_{JS} . Top right: cumulative distribution function of the number of glitches that induced a requisite divergence as a function of glitch SNR. Only a single glitch with SNR less than 100 induces a measurable bias. Even at SNR 1000 fewer than 20% of the glitches induce a difference greater than one would expect from stochastic sampling.

signal (2,2) mode do not have a noticeable visual effect on the posteriors.

The 25 Hz glitch overlaps the CBC’s (2,2) mode in time-frequency space, and thus falls squarely into Region I. It is clearly biasing inference, increasing the total mass and χ_{eff} . The reweighing efficiency (right panel) is close to 0%, and thus reweighing is unable to reconstruct the direct-sampled posteriors.

Glitches with central frequencies 100 Hz and above have efficiencies close to 100%, meaning that posteriors from data with and without the glitch are indistinguishable. This is visually evident for M , χ_{eff} , and χ_p (middle panels), but the efficiency is a stronger test that considers the full 15-dimensional posteriors. Thus, glitches that do not overlap in time-frequency with the CBC model prior do not impact inference in this example.

The 50 Hz glitch presents a middle case: identical-looking posteriors but an efficiency of 50%, indicating a measurable bias. This is because that glitch shares time-frequency content with the signal’s high-order modes as well as the general CBC model prior (left panel). Glitches that overlap with a signal’s high-order modes can therefore still lead to small biases, even if the dominant (2,2) mode is not impacted.

2. At what glitch SNR can we expect a bias?

Finally, we repeat the analysis of Sec. IV B 4 with 200 random glitches in Region IIa. The analysis setup is

identical, and we refer to that discussion for details. We present our results in Fig. 6 in similar format to Fig. 4. All glitches that result in a bias for SNR less than 500 have central frequencies close to the CBC merger frequency, 150 Hz to 400 Hz. Even then, biases are rare. *Only a single glitch* (0.005%) causes a bias by SNR 100. Somewhat surprisingly, glitches in this region have, on average, a smaller impact on inference than those in Region III.

3. Region IIb: Glitches just below the signal

Since we are focusing on high-mass events [where inference is most heavily impacted by glitches [10]], Region IIb is very small. Figure 1 shows that there is very little time-frequency space below the waveform and outside the prior. We therefore do not perform any analyses here, instead expecting similar results to Region IIa. Region IIb will be bigger (and could contain glitches) for lower-mass signals.

D. Region I, glitches on top of the signal

It has been shown both in real events [12–14] and simulations [10] that glitches overlapping GWs can be detrimental to inference. We do not present further dedicated studies of this Region here, but nonetheless confirm the previous results in the right panel on Figs. 2 and 5. The former shows that a population of glitches that overlap the time-frequency prior of a population of CBCs will result in strong biases. The “s-curves” in a P-P plot are indicative of

the standard deviation of the posterior being incorrect. The latter shows that glitches that overlap with dominant signal mode can lead to large (visible) biases, but a smaller impact is also expected when the glitch overlaps the higher-order modes.

V. CONCLUSION

The purpose of this study was to determine if there are time-frequency locations where one can ignore glitches nearby a transient GW, even when it is in the analysis window. We split our findings and recommendations into the same Regions as in Fig. 1.

- (i) Region I includes glitches that are coincident in time-frequency with the GW event and have been shown to impact inference on real signals [12,14,20]. We confirm these results even at the population level; such glitches need to be carefully mitigated.
- (ii) Region II includes glitches just above and just below the GW model prior in time-frequency. That is, these glitches overlap in time and in frequency with the signal, but not concurrently. Glitches in this region do not impact inference and thus do not require mitigation, unless they have SNR above 50 *and* are close in frequency to the CBC merger frequency.
- (iii) Region III includes glitches that share frequency content with the CBC and its prior, but occur *after* the GW, and thus do not overlap in time. For a fiducial population of glitches, biases are not expected for a glitch SNR below 50, or an SNR below 100 if the glitch is more than 0.5 s after the merger. Mitigation is again not strictly required, though in that case it is significantly simpler to perform though inpainting [30], gating [29], or modeling solely the glitch with BayesWaveCpp with an appropriate analysis window that excludes the CBC [23].
- (iv) Region IV includes glitches that do not share frequency content with the GW signal. Even if these frequencies are included in the analysis window, they cannot impact inference.

Throughout this study we have assumed that the noise PSD is perfectly known and can be estimated regardless of the glitch. When estimating the PSD “on source” [34,52,53] the glitch model is present to account for non-Gaussian noise that could affect the PSD if the glitch is loud. The uncertainty associated with the power spectrum estimate is also not accounted for here, although such uncertainty is a subdominant effect in CBC analyses on *glitch-subtracted* data [54]. Other noise-mitigation techniques such as inpainting do require knowledge of the PSD which could be estimated off-source.

In summary, glitches outside Region I in Fig. 1 are unlikely to affect parameter estimation unless they are sufficiently loud and close to the merger in time and/or frequency. Though the results quantifying the glitch SNR required for biases were based on a single, representative CBC signal, we probed a large number of potential glitch parameters and time-frequency configurations. Moreover, the population analysis of Sec. IV B 1 confirms these expectations for CBC signals drawn from the prior. Quieter glitches (conservatively, below SNR 50) can thus be left in the data. Loud glitches are fortunately much rarer than their low SNR counterparts.

This work made use of BayesWaveCpp [37], GWPY [55], SciPy [56], Matplotlib [57], PANDAS [58], and NumPy [59].

ACKNOWLEDGMENTS

Thank you to Colm Talbot, Reed Essick, Jacob Golomb, Lucy Thomas, and Ryan Magee for useful discussions about the IMRPhenomXPHM waveforms systematics. Thank you to Eliot Finch for help with determining the time-frequency location of the waveform ringdown. Thank you to Derek Davis for their expertise on glitches and glitch mitigation. Thank you to Jane Glanzer and Siddharth Soni for help with glitch rate estimates as a function of SNR. S. H. was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1745301. K. C. was supported by National Science Foundation Grant No. PHY-2409001. This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation, and operates under cooperative Agreement No. PHY-2309200. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459.

DATA AVAILABILITY

The data that support the findings of this article are openly available [38].

APPENDIX: WAVEFORM CONDITIONING AND SYSTEMATICS

In Sec. IV B 3, while we showed that the reweighting efficiency decreases for glitch SNRs of 1000 and 5000, we omitted posteriors created by direct-sampling on the glitch-impacted data. This was because in these cases the CBC

model ended up fitting part of the glitch due to some waveform unphysical behavior. In Fig. 7, we show the maximum likelihood waveform recovered on that data, which has a clear, nonphysical spike caused by a known failure in the multibanding [39] of the waveform.

Regardless of waveform systematics, we recover a drop in efficiency due to the presence of these high SNR glitches. This efficiency drop (computed via reweighting) *cannot* be caused by the aforementioned waveform systematic. On glitch-free data, such a pathological waveform is so disfavored by the likelihood and occupies a minute region of parameter space, that it has a vanishing probability of appearing in a direct-sampled posterior. We interpret the fact that the BayesWaveCpp sampler located this pathological behavior as a testament to its efficacy.

If such inference was obtained from a real signal, the issue would be immediately obvious. Nonetheless, we recommend that glitches with an SNR above 500 be mitigated.

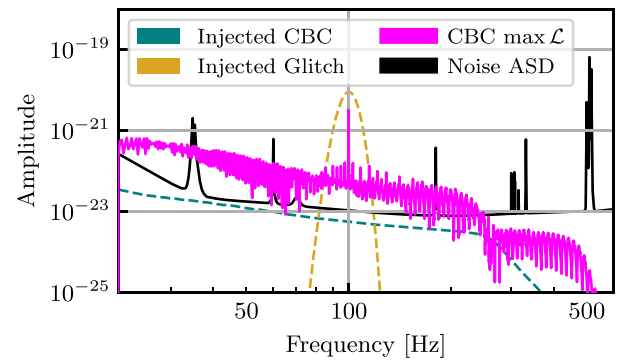


FIG. 7. Spectral amplitude for the SNR 5000 glitch from Sec. IV B 3. We show the noise ASD (square root of the PSD) in black as well as the simulated CBC (gold) and glitch (green). The pink line shows the maximum likelihood posterior sample recovered when directly sampling the posterior with BayesWaveCpp and IMRPhenomXPHM. The sampler stumbled upon a rare waveform pathology that resulted in a spike in frequency that was able to partially fit the glitch.

-
- [1] R. Abbott *et al.* (KAGRA Collaboration, VIRGO Collaboration, and LIGO Scientific Collaboration), *Phys. Rev. X* **13**, 041039 (2023).
 - [2] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. X* **11**, 021053 (2021).
 - [3] R. Abbott *et al.* (KAGRA Collaboration, VIRGO Collaboration, and LIGO Scientific Collaboration), *Phys. Rev. X* **13**, 011048 (2023).
 - [4] R. Abbott *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), *arXiv:2112.06861*.
 - [5] B. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
 - [6] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
 - [7] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
 - [8] LIGO Scientific Collaboration and Virgo Collaboration, LIGO/Virgo public alerts (o4), <https://gracedb.ligo.org/superevents/public/O4/> (2024), accessed: 2024-10-09.
 - [9] S. Soni, B. K. Berger, D. Davis *et al.*, *Classical Quantum Gravity* **42**, 085016 (2025).
 - [10] S. Hourihane, K. Chatzioannou, M. Wijngaarden, D. Davis, T. Littenberg, and N. Cornish, *Phys. Rev. D* **106**, 042006 (2022).
 - [11] S. Ghonge, J. Brandt, J. M. Sullivan, M. Millhouse, K. Chatzioannou, J. A. Clark, T. Littenberg, N. Cornish, S. Hourihane, and L. Cadonati, *Phys. Rev. D* **110**, 122002 (2024).
 - [12] C. Pankow *et al.*, *Phys. Rev. D* **98**, 084016 (2018).
 - [13] E. Payne, S. Hourihane, J. Golomb, R. Udall, R. Udall, D. Davis, and K. Chatzioannou, *Phys. Rev. D* **106**, 104017 (2022).
 - [14] R. Udall, S. Hourihane, S. Miller, D. Davis, K. Chatzioannou, M. Isi, and H. Deshong, *Phys. Rev. D* **111**, 024046 (2025).
 - [15] R. Macas, J. Pooley, L. K. Nuttall, D. Davis, M. J. Dyer, Y. Lecoche, J. D. Lyman, J. McIver, and K. Rink, *Phys. Rev. D* **105**, 103021 (2022).
 - [16] J. Powell, *Classical Quantum Gravity* **35**, 155017 (2018).
 - [17] J. Y. L. Kwok, R. K. L. Lo, A. J. Weinstein, and T. G. F. Li, *Phys. Rev. D* **105**, 024066 (2022).
 - [18] S. Mozzon, G. Ashton, L. K. Nuttall, and A. R. Williamson, *Phys. Rev. D* **106**, 043504 (2022).
 - [19] K. Chatzioannou, N. Cornish, M. Wijngaarden, and T. B. Littenberg, *Phys. Rev. D* **103**, 044013 (2021).
 - [20] E. Payne, S. Hourihane, J. Golomb, R. Udall, D. Davis, and K. Chatzioannou, *Phys. Rev. D* **106**, 104017 (2022).
 - [21] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [22] R. Abbott *et al.* (LIGO Scientific Collaboration and VIRGO Collaboration), *Phys. Rev. D* **109**, 022001 (2024).
 - [23] D. Davis, T. B. Littenberg, I. M. Romero-Shaw, M. Millhouse, J. McIver, F. Di Renzo, and G. Ashton, *Classical Quantum Gravity* **39**, 245013 (2022).
 - [24] L. Vazsonyi and D. Davis, *Classical Quantum Gravity* **40**, 035008 (2023).
 - [25] S. J. Miller, M. Isi, K. Chatzioannou, V. Varma, and I. Mandel, *Phys. Rev. D* **109**, 024024 (2024).

- [26] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. X* **11**, 021053 (2021).
- [27] D. Davis, T. J. Massinger, A. P. Lundgren, J. C. Driggers, A. L. Urban, and L. K. Nuttall, *Classical Quantum Gravity* **36**, 055011 (2019).
- [28] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, *Phys. Rev. D* **100**, 104004 (2019).
- [29] J. Zweizig and K. Riles, Information on self-gating of $h(t)$ used in O3 continuous-wave and stochastic searches, Tech. Rep. (<https://dcc.ligo.org/LIGO-T2000384>, 2021).
- [30] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **104**, 063034 (2021).
- [31] S. Soni, J. Glanzer, A. Effler, V. Frolov, G. González, A. Pele, and R. Schofield, *Classical Quantum Gravity* **41**, 135015 (2024).
- [32] S. Soni *et al.* (LIGO Collaboration), *Classical Quantum Gravity* **38**, 025016 (2020).
- [33] R. Udall and D. Davis, *Appl. Phys. Lett.* **122**, 094103 (2023).
- [34] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, *Phys. Rev. D* **103**, 044006 (2021).
- [35] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).
- [36] J. D. Romano and N. J. Cornish, *Living Rev. Relativity* **20**, 2 (2017).
- [37] H. Deshong, S. Hourihane, K. Chatziioannou, A. Johnson, M. Holst, M. Millhouse, T. Littenberg, and N. Cornish, Bayeswave c++ repository (2025), gitLab repository, accessed May 21, 2025.
- [38] LIGO Scientific Collaboration and Virgo Collaboration, BayesWave, <https://git.ligo.org/lscsoft/bayeswave> (2018).
- [39] G. Pratten *et al.*, *Phys. Rev. D* **103**, 104056 (2021).
- [40] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).
- [41] S. Hourihane, P. Meyers, A. Johnson, K. Chatziioannou, and M. Vallisneri, *Phys. Rev. D* **107**, 084045 (2023).
- [42] E. Payne, C. Talbot, and E. Thrane, *Phys. Rev. D* **100**, 123017 (2019).
- [43] I. M. Romero-Shaw, P. D. Lasky, E. Thrane, and J. C. Bustillo, *Astrophys. J. Lett.* **903**, L5 (2020).
- [44] I. M. Romero-Shaw, P. D. Lasky, and E. Thrane, *Astrophys. J.* **940**, 171 (2022).
- [45] I. M. Romero-Shaw, N. Farrow, S. Stevenson, E. Thrane, and X.-J. Zhu, *Mon. Not. R. Astron. Soc.* **496**, L64 (2020).
- [46] I. M. Romero-Shaw, P. D. Lasky, and E. Thrane, *Astrophys. J. Lett.* **921**, L31 (2021).
- [47] M. Dax, S. R. Green, J. Gair, N. Gupte, M. Pürrer, V. Raymond, J. Wildberger, J. H. Macke, A. Buonanno, and B. Schölkopf, *Nature (London)* **639**, 49 (2025).
- [48] J. Lin, *IEEE Trans. Inf. Theory* **37**, 145 (1991).
- [49] I. M. Romero-Shaw *et al.*, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [50] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference, Fourth Edition* (Marcel Dekker, Inc, New York, 2003), p. 145.
- [51] S. R. Cook, A. Gelman, and D. B. Rubin, *J. Comput. Graph. Stat.* **15**, 675 (2006).
- [52] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **91**, 084034 (2015).
- [53] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, *Phys. Rev. D* **100**, 104004 (2019).
- [54] C. Plunkett, S. Hourihane, and K. Chatziioannou, *Phys. Rev. D* **106**, 104021 (2022).
- [55] D. Macleod, A. L. Urban, S. Coughlin, T. Massinger, M. Pitkin, paulaltin, J. Areeda, E. Quintero, T. G. Badger, L. Singer, and K. Leinweber, gwpy/gwpy: 1.0.1 (2020).
- [56] P. Virtanen *et al.* (SciPy 1.0 Contributors), *Nat. Methods* **17**, 261 (2020).
- [57] J. D. Hunter, *Comput. Sci. Eng.* **9**, 90 (2007).
- [58] T. pandas development team, pandas-dev/pandas: Pandas (2020).
- [59] C. R. Harris *et al.*, *Nature (London)* **585**, 357 (2020).