Modeling compact binary signals and instrumental glitches in gravitational wave data

Katerina Chatziioannou, Neil Cornish, Marcella Wijngaarden, and Tyson B. Littenberg

Modeling compact binary signals and instrumental glitches in gravitational wave data

Katerina Chatziioannou,[1, 2, 3] Neil Cornish,[4] Marcella Wijngaarden,[3, 5] and Tyson B. Littenberg[6]

[1]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*
[2]*LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA*
[3]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA*
[4]*eXtreme Gravity Institute, Department of Physics, Montana State University, Bozeman, Montana 59717, USA*
[5]*Mathematical Sciences and STAG Research Centre, University of Southampton, SO17 1BJ, Southampton, UK*
[6]*NASA Marshall Space Flight Center, Huntsville, AL 35812, USA*
(Dated: January 12, 2021)

Transient non-gaussian noise in gravitational wave detectors, commonly referred to as glitches, pose challenges for detection and inference of the astrophysical properties of detected signals when the two are coincident in time. Current analyses aim towards modeling and subtracting the glitches from the data using a flexible, morphology-independent model in terms of sine-gaussian wavelets before the signal source properties are inferred using templates for the compact binary signal. We present a new analysis of gravitational wave data that contain both a signal and glitches by simultaneously modeling the compact binary signal in terms of templates and the instrumental glitches using sine-gaussian wavelets. The model for the glitches is generic and can thus be applied to a wide range of glitch morphologies without any special tuning. The simultaneous modeling of the astrophysical signal with templates allows us to efficiently separate the signal from the glitches, as we demonstrate using simulated signals injected around real O2 glitches in the two LIGO detectors. We show that our new proposed analysis can separate overlapping glitches and signals, estimate the compact binary parameters, and provide ready-to-use glitch-subtracted data for downstream inference analyses.

I. INTRODUCTION

During the first half of their third observing run (O3a), the advanced ground based gravitational wave (GW) detectors LIGO [1] and Virgo [2] observed an astrophysical transient signal about every 5 days of data [3]. The large detection rate increases the chance of observing an event while one of the detectors experiences transient non gaussian noise, also known as instrumental glitches. Indeed, this scenario has come to pass for one event from the second observing run (O2) [4] and 8 events from the first half of the third observing run [3].

Such coincidences are expected to become even more frequent in the coming years. Planned improvements in the detectors' sensitivity will be directly reflected by an even larger rate of astrophysical discoveries [5]. Moreover, O3a was characterized by an increase in the rate of glitch occurrence in the two LIGO detectors, a trend that might persist during the fourth observing run (O4) as the decreased average detector noise might help reveal weaker sources of transient noise. For example the rate of glitches in the LIGO Livingston detector increased from 0.2 per minute in O2 to 0.8 per minute in O3a [3].

The presence of a non gaussian noise feature in the data, a glitch, poses challenges for nearly all inference analyses. GW inference is based on a model for the detector noise, expressed through the likelihood function. In the absence of glitches, detector noise is colored and gaussian to a very good approximation [6], with a spectrum that is described through the noise power spectral density (PSD). The above considerations give rise to a gaussian likelihood function whose variance is the noise PSD, a choice that is almost ubiquitous [7, 8]. Different choices for estimating the PSD or treating its uncertainty can result in different functional forms for the likelihood, but they are all based on the assumption of colored gaussian noise [9, 10].

Since instrumental glitches violate the basic assumptions of GW inference, they need to be effectively mitigated before the data are analyzed. One option is to remove the offending data all together [4, 11–13], which can be done quickly, allowing for low latency estimation of source parameters that enable followup observations [4]. The downside of this approach is that part of the astrophysical signal is lost making it prohibitive for binary black hole (BBH) signals whose duration is comparable to the glitch duration. In order to avoid signal, and thus information loss, another option is to model the glitch and regress it from the data, leaving behind not only the astrophysical signal but also the gaussian noise. This approach is the topic of the current study[1].

The wide variety of glitch morphologies, and even variations within a certain glitch type, make constructing exact models for glitches challenging [21]. A more flexible approach is based on `BayesWave` [22, 23] which models various components of the GW data in a morphology-independent way. Nongaussian features in the data are modeled in terms of sums of sine-gaussian wavelets whose number and parameters are marginalized over with a suite of Markov Chain Monte Carlo (MCMC) and Reversible Jump MCMC (RJMCMC) [24] samplers. Coherent features (i.e. features that appear in all detectors in a

---

[1] An independent effort to mitigate the effect of broadband and/or nonstationary detector noise is based on information from auxiliary sensors [14–20]. This approach does not remove entire data segments either and thus is not expected to lead to loss of information.

manner consistent with an astrophysical signal originating from a specific sky location) are modeled by a single sum of wavelets that is projected onto the detector network; these features are interpreted as having an astrophysical origin. Incoherent features are instead modeled by independent sums of wavelets in each GW detector and are interpreted as instrumental glitches. The PSD of the gaussian noise is also modeled in terms of splines and Lorentzians using an algorithm sometimes known as `BayesLine` [6, 25]. `BayesWave` and `BayesLine` are fully integrated and we will refer to the combined analysis with the name `BayesWave` in this paper.

Modeling instrumental glitches with `BayesWave` and subtracting them from the data in order to make ready-to-use data for downstream inference has been a standard step of LIGO/Virgo analyses since O2 [3, 4]. The GW signal from the first binary neutron star (BNS) coalescence detection, GW170817, overlapped with a glitch in the LIGO Livingston detector approximately 1.1s before coalescence [4]. The glitch was modeled with `BayesWave`'s *glitch model* in terms of a sum of wavelets and removed from the data, a procedure documented and released in [26]. Despite the glitch overlapping with the actual astrophysical signal, the subtraction process was robust against inadvertently removing the signal together with the glitch. The reason is that the specific glitch was short in duration (less than a second) and extended in frequency, unlike the signal that lasted for about 2 minutes in the detector sensitive band. As such, the sine-gaussian wavelets that would fit the glitch and the signal are distinct in terms of their time-frequency features; the wavelets that model the glitch are short and hence do not model the long-lasting BNS signal. This procedure was further shown to not introduce biases in the astrophysical parameter inference of the underlying signal by analyzing simulated signals injected on instances of the same glitch type in LIGO Livingston data [27].

Motivated by the success of this first attempt at glitch mitigation and in preparation for the increased detection rate of O3, `BayesWave` was extended to be able to simultaneously model both the signal and the glitch [23]. Both signals and glitches are modeled with a sum of sine-gaussian wavelets, the only difference being that the signal is coherent across the detectors in the network, while the glitch is not. The analysis effectively uses data from all detectors available to determine which part of the non gaussian data are coherent (and would thus correspond to an astrophysical signal), and which part is incoherent (and would thus correspond to an instrumental glitch). The combined signal+glitch analysis was applied to one O3a detection [3], enabling glitch mitigation even for data that contained short-duration BBH signals.

The signal+glitch analysis models compact binary coalescence (CBC) signals in terms of wavelets, and is thus agnostic to the signal morphology. However, accurate models exist for CBCs in terms of solutions to the Einstein field equations that are routinely used both for detection and parameter estimation. In this paper we take

another step towards efficient separation of CBCs and glitches by constructing an analysis that simultaneously models the CBC signal in terms of CBC templates and the glitch in terms of sine-gaussian wavelets. Similar to the initial glitch-only analysis and the subsequent signal+glitch analysis, we also model and marginalize over the detector noise PSD. We test our analysis using public O2 data that contain common glitch types and simulated CBC signals. We demonstrate that we can efficiently separate the glitch from the CBC, estimate the CBC parameters, and provide ready-to-use glitch-subtracted data for downstream inference analyses.

The rest of the paper is organized as follows. In Sec. II we describe the updates to the standard `BayesWave` algorithm in terms of the CBC analysis. In Sec. III we apply our analysis to simulated signal overlapping with known detector glitches from O2 data. In Sec. IV we analyze a selection of detected signals, namely GW170817 and GW150914. Finally, in Sec. V we conclude and point to future work.

II. GENERAL ALGORITHM DESCRIPTION

The combined `BayesWave` algorithm is presented in detail in [23] and here we describe only the features relevant to our study. `BayesWave` simultaneously models signals, glitches, and gaussian noise in GW data by means of different models. The *signal model* describes astrophysical signals through a sum of Morlet Gabor wavelets that are coherent across the detector network. The number of wavelets and the parameters of each are marginalized over, as are the extrinsic parameters that determine how the signal is projected in each detector. The *glitch model* describes instrumental glitches with an incoherent sum of Morlet Gabor wavelets whose number and parameters are again marginalized over. Glitch power in each detector is described by an independent sum of such wavelets. The *noise model* describes the gaussian noise PSD with a broadband spline model and sharp Lorentzians. As above, the number of spline points and Lorentzians as well as their parameters are marginalized over.

In order to sample the multidimensional posterior density of all models, `BayesWave` uses a blocked Gibbs sampler that takes turns between sampling each model with completely independent MCMC or RJMCMC samplers. This includes (i) an RJMCMC that samples the signal and glitch wavelet parameters, (ii) an MCMC that samples the signal extrinsic parameters, and (iii) an RJMCMC that samples the splines and Lorentzians for the noise PSD. Each sampler in turn updates its parameters for a predetermined number of iterations, typically $\mathcal{O}(10^2)$, while all other parameters are kept fixed. For example, the extrinsic sampler updates the extrinsic signal parameters while the wavelet parameters and noise PSD are kept constant. Once the predetermined number of updates has been reached, the extrinsic sampler returns its current parameters and the noise sampler begins up-

dating the noise model while keeping the wavelet and extrinsic parameters fixed. This process of alternating sampling between different blocks of model parameters is repeated for $\mathcal{O}(10^4)$ iterations.

The construction of the algorithm in terms of a blocked Gibbs sampler makes adding further models and samplers straightforward. In the current version described in [23], the astrophysical signal is modeled with coherent sine-gaussian wavelets that allow us to describe signals with a large level of flexibility. We extend `BayesWave`'s blocked Gibbs sampler by adding one more element, namely a model of the signal in terms of quasicircular CBC waveforms. In fashion with the existing implementation, the MCMC that samples the posterior distribution for the CBC parameters is completely independent from the remaining code samplers. The result is a flexible algorithm that can be used with any combination of CBC, signal[2], glitch, and noise models for the detector data.

The CBC model is integrated with `LALSimulation` [28] and can operate with any non-precessing model available there[3]. The eleven parameters of a spin-aligned quasicircular CBC signal, namely the four intrinsic parameters (the two masses and spin magnitudes) and seven extrinsic parameters (the time of coalescence, the phase of coalescence, two sky location angles, the polarization angle and the inclination angle the distance), are updated in overlapping blocks. Common to both blocks is the phase of coalescence since `BayesWave`'s extrinsic sampler updates the overall phase of the signal as described in [23]. The CBC MCMC sampler updates the four intrinsic parameters, the time of coalescence, the phase of coalescence, and the distance. The existing extrinsic sampler in `BayesWave` updates the two sky angles, the polarization angle, the inclination angle, and the phase of coalescence while holding all other parameters fixed. We use standard priors for all parameters: uniform over the detector-frame masses and spin magnitudes, uniform in time and phase, and uniform in luminosity volume.

The CBC sampler is custom and not based on any existing samplers used in LIGO-Virgo parameter estimation. The CBC sampler is taken from the recently developed `QuickCBC` [29] analysis pipeline. A closely related sampler [30] has been developed for analyzing data from the future Laser Interferometer Space Antenna. The CBC sampler is a replica exchange (parallel tempered)

Markov Chain Monte Carlo (PTMCMC) algorithm that uses a mixture of proposal distributions. The default collection of proposals are: Gaussian jumps along eigenvectors of the Fisher information matrix, scaled by the reciprocal of the square root of the corresponding eigenvalue; differential evolution using a rolling history array at each temperature, updated every 10 iterations and holding 1000 past samples; and small, Gaussian jumps along each parameter direction. Each chain carries its own Fisher information matrix, which is updated periodically. The Fisher and differential evolution proposals are effective at exploring parameter correlations, while the small jumps prevent the chains from getting stuck in regions where the Fisher matrix becomes ill-conditioned.

The CBC sampler is not optimized for blindly finding signals, so it is best to initialize the sampler with a good starting solution for the source parameters such as the output from a CBC search pipeline, or the injected parameters for a simulated signal. Alternatively the sampler can be initialized using a custom built CBC search algorithm from the `QuickCBC` [29] analysis pipeline that has been incorporated into the `BayesWave` pre-processing steps. The search is broken into two stages, a rapid network-coherent search with analytic maximization over extrinsic parameters, followed by a fast MCMC over the extrinsic parameters using a likelihood function that precomputes the waveform inner products [31]. This procedure returns the starting point for all 11 CBC parameters. More details about the initial search step and discussion of its robustness against instrumental glitches are presented in [29].

III. SIMULATED SIGNALS

We test the efficacy of separating CBCs from glitches with our CBC+glitch model by selecting 3 common glitch types from O2 data [32, 33] that are known to have an adverse effect on searches for CBCs [3]. We then add simulated CBC signals consistent with a BBH with detector-frame masses of $36 M_\odot$ and $29 M_\odot$ and vanishing spin at different times with respect to the glitch. All simulated signals have a signal-to-noise (SNR) ratio of 15. We use the `IMRPhenomD` [34, 35] waveform model both for simulation and recovery as implemented in `LALSimulation` [28]. We then analyze the data from the two LIGO detectors with our CBC+glitch+noise model, where the coherent signal is modeled by the CBC template, the glitch is modeled by incoherent wavelets, and the noise PSD is modeled with splines and Lorenzians. Spectrograms for the 3 glitches are shown in Fig. 1: blip glitch (left), scattered light (middle), and blue mountain (right). Further details and run settings for each type of glitch are shown in Table I.

---

[2] We retain the original model names in `BayesWave`, hence the *signal model* refers to the wavelet signal model, while the *CBC model* refers to the model in terms of CBC templates. Both models target astrophysical signals. Since we do not use the *signal model* in the remainder of the paper, we trust that this will not lead to confusion.

[3] Both the sampling and the jump proposals for the CBC parameters are constructed to expect the signal amplitude and phase from the waveform generator. There is therefore no fundamental limitation to non-precessing signals and we plan to extend our analysis to include the effect of spin-precession in the future.
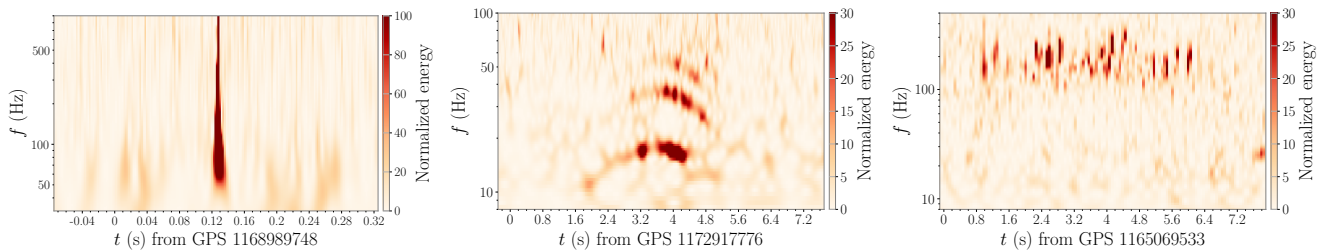
FIG. 1. Spectrograms for the three glitches of different types studied here: blip glitch (left), scattered light (middle), blue mountain (right). The three types of glitches are characterized by very different time-frequency properties.

Glitch	GPS time (s)	Detector	Segment length (s)	Sampling rate (Hz)	f_{low}(Hz)	Q_{max}	CBC SNR
Blip	1168989748	Hanford	4	2048	16	40	15
Scattered light	1172917779	Livingston	8	2048	8	160	15
Blue mountain	1165069536	Hanford	16	2048	16	40	15

TABLE I. Settings for the runs of Sec. III. From left to right, columns correspond to the type of glitch, the GPS time, the affected detector, the segment length, the sampling rate, to low frequency cut off, the maximum quality factor of the glitch wavelets, and the SNR of the injected signals.

A. Glitch type 1: Blip

Blip glitches are one of the most common glitch types for the two LIGO detectors. They are characterized by short duration, and hence pose a challenge for the detection of high mass BBH signals [36]. Their origin is largely unknown. Figures 2-5 show our results for simulated signals injected at different times with respect to a blip glitch in the LIGO Hanford detector during O2. Details about the glitch, including its GPS time, and the run settings are presented in Table I. A spectrogram of the data containing the glitch is given on the left panel of Fig. 1, where the short duration and large frequency extent are shown.

The whitened data and reconstructions for the CBC signal and the glitch are shown in Fig. 2 where we plot the 90% credible intervals for each reconstruction in LIGO Hanford (top) and LIGO Livingston (bottom). The glitch is easily visible in LIGO Hanford as a short duration $\sim 15\sigma$ noise excursion. No glitch power is identified in LIGO Livingston at that time, but the CBC signal is clearly identified. This allows us to separate the corresponding coherent CBC signal in LIGO Hanford from the instrumental glitch, even when the latter overlaps with the merger phase of the signal (left panel). The glitch reconstruction is also consistent across the three simulated signals, suggesting that the glitch model is not fitting any part of the CBC signal.

Source parameters for the simulated CBC are presented in Figs. 3 and 4 both for the CBC+glitch+noise analysis and a CBC+noise analysis for selected recovered parameters for the leftmost simulated CBC signal together with the injected values with black crosses or vertical lines as appropriate. Figure 3 shows the mass ratio q, the effective spin χ_{eff}, and the detector frame chirp mass $\mathcal{M}$, while Fig. 4 shows the luminosity distance and the cosine of the inclination angle. In all cases the posterior distributions recovered under the CBC+glitch+noise model are consistent with the injected parameters, though the marginalized posteriors do not peak at the injected values, as expected from inference of signals in gaussian noise. For reference, we show posteriors under the CBC+noise model in orange that assumes that the data are consisted of just a CBC signal and gaussian noise, without any provision for a glitch. Since this assumption is violated by the presence of the blip glitch, the resulting posteriors are expected to be biased compared to the true parameters and the orange contours in Figs. 3 and 4 quantify this bias. We find that the extrinsic parameters that are primarily determined by the signal amplitude are more biased than the intrinsic ones that are measured through the GW phase, as also discussed in [37].

The separation of the CBC signal from the glitch demonstrated in Fig. 2 can be used to produce ready-to-use deglitched data for downstream inference analyses, as was done in [3]. An estimate of the glitch reconstruction (the median or a fair draw from the glitch model posterior) is subtracted from the data to produce strain data that contain only the CBC signal and gaussian noise. The result of the glitch subtraction is shown in the spectrograms of Fig. 5 that show the LIGO Hanford data before (left) and after (middle) the subtraction of a fair draw glitch reconstruction for the leftmost injection of Fig. 2. The left panel includes both the chirping signal and the blip glitch, while only the former is visible in the middle panel. The right panel shows the data after a fair draw from both the CBC and the glitch models has been subtracted, resulting in residual gaussian noise only.
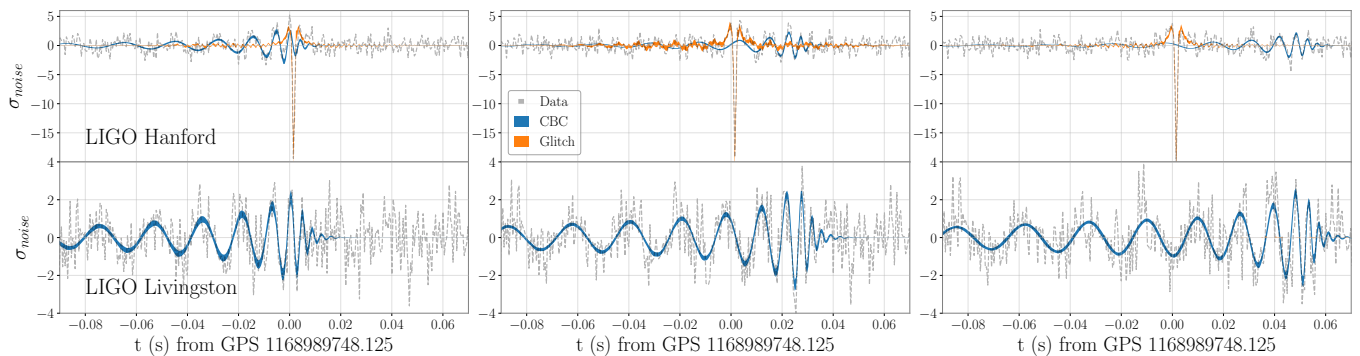
FIG. 2. Credible intervals for the glitch (orange) and the CBC (blue) signal reconstruction for data containing a blip glitch in LIGO Hanford and a simulated CBC signal at 3 different times with respect to the glitch (left to right). Shaded regions correspond to 90% credible intervals for the whitened reconstruction, while in grey dashed lines we plot the data whitened with a fair draw PSD from our noise model posterior. The top row corresponds to LIGO Hanford and the bottom row corresponds to LIGO Livingston.
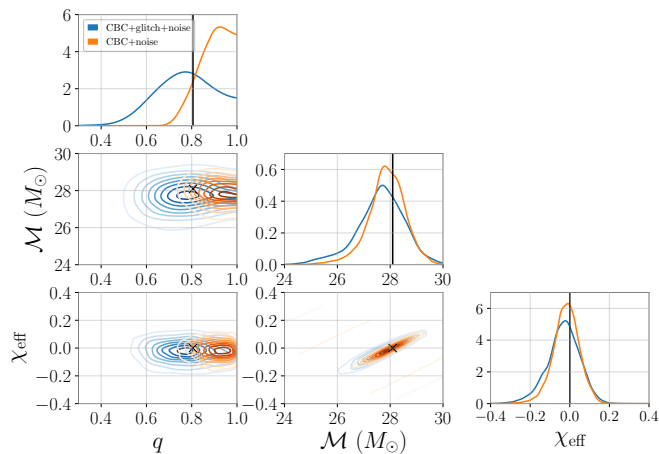

FIG. 3. One- and two-dimensional posterior distributions for selected source parameters of the simulated signal from the left panel of Fig. 2 injected on top of a LIGO Hanford blip glitch. We include the mass ratio q, the effective spin $\chi_{\rm eff}$, and the detector frame chirp mass $\mathcal{M}$ posteriors, while black crosses or black vertical lines denote the true parameters of the injection. Blue (orange) contours and lines correspond to the CBC+glitch+noise (CBC+noise) run.
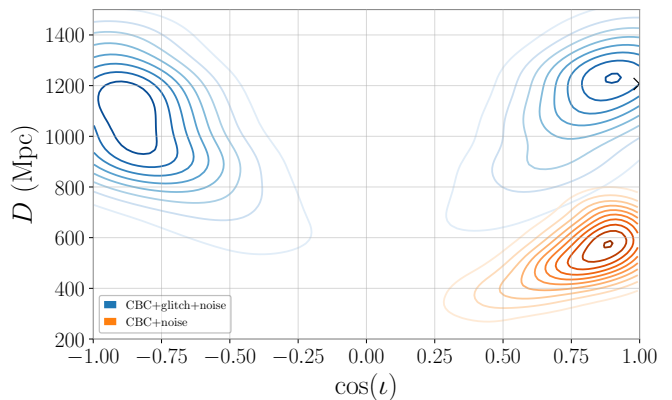

FIG. 4. Two-dimensional posterior distributions for the luminosity distance and the binary inclination of the simulated signal from the left panel of Fig. 2 injected on top of a LIGO Hanford blip glitch. A black cross at (1,1200Mpc) denotes the true parameters of the injection. Blue (orange) contours correspond to the CBC+glitch+noise (CBC+noise) run.

B. Glitch type 2: Scattered light

Glitches caused by scattered light in the interferometer became particularly prominent during O3 [3]. Unlike the blip glitches studied above, scattered light glitches have a longer temporal duration of a few seconds and are characterized by arches in a time-frequency spectrogram [38, 39], as depicted in the middle panel of Fig. 1. We inject simulated signals on an instance of such a glitch in LIGO Livingston and analyze the data from both LIGO detectors with our CBC+glitch+noise model. Details of the glitch and the run settings are given in Table I. Due to the duration of the glitch and its low frequency

power we extend our analysis duration and bandwidth. The longer duration helps the noise model determine the low-frequency gaussian noise PSD and thus separate the low frequency part of the glitch from gaussian noise. We also increase the maximum quality factor $Q_{\rm max}$ of the wavelets due to the glitch's long duration.

Figure 6 shows the data and reconstructed CBC and glitch models. We zoom in around the CBC signals, though the glitch extends beyond the time range plotted. In all cases the CBC signal is separated from the glitch, aided by the presence of a coherent signal in LIGO Hanford. The glitch reconstruction is also consistent for all 3 simulated signals, as expected for runs on the same glitch. The reconstruction exhibits oscillations at around 32Hz and 16Hz, consistent with expectations from the glitch spectrogram. Figure 7 shows posterior distributions for selected source parameters for the left-most injection in
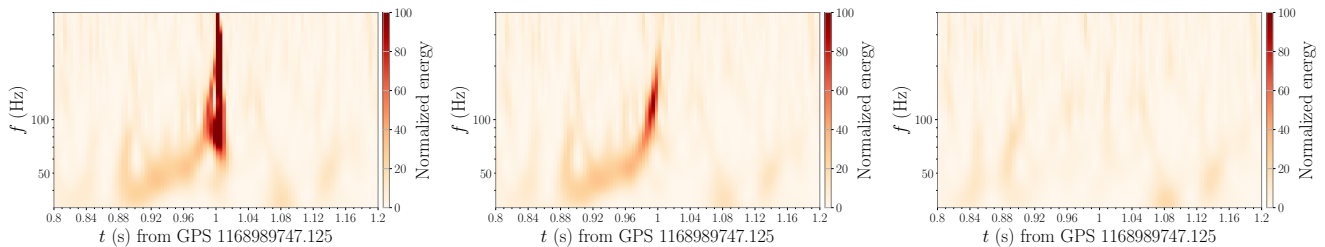
FIG. 5. Spectrogram of the LIGO Hanford data around the time of the blip glitch for the leftmost injection from Fig. 2. Left panel: Data containing the blip glitch and the simulated CBC signal. Middle panel: Data after a fair draw from the glitch model has been subtracted leaving behind only the chirping CBC signal. Right panel: Data after a fair draw from the glitch and CBC models has been subtracted, leaving behind only gaussian detector noise.

blue, as well as the injected parameters. In all cases the recovered parameters are consistent with their injected values. In orange, we plot results from a CBC+noise run and find small biases in the source intrinsic parameters, most notably the mass ratio.

Finally, Fig. 5 shows the spectrogram of the data before and after various components of the model have been subtracted from the data. The left panel corresponds to data that contain both a signal and the glitch and thus both the signal chirp and the characteristic glitch arches are visible. In the middle panel we plot data after a fair draw from the glitch model has been subtracted, resulting in both the high and the low frequency arches of the glitch having been regressed, leaving only the chirping signal behind. The right panel corresponds to data where a fair draw from the CBC model has further been subtracted and is consistent with gaussian noise.

C. Glitch type 3: Blue mountain

The final type of glitch we consider is the blue mountain; the spectrogram of the LIGO Hanford instance of a blue mountain glitch we consider is shown in the right panel of Fig. 1. The glitch has a duration of multiple seconds and is characterized by higher frequencies ~ 200Hz. We inject simulated signals at different times relative to the glitch and again analyze data from the two LIGO detectors with the CBC+glitch+noise model with settings shown in Table I. Due to the large glitch duration we have to increase the length of the analyzed segment even further to 16s. Despite the glitch's overall long duration, we do not find it necessary to increase the wavelet maximum quality factor Q_{max}, as the glitch is composed of short individual bursts of power, each of which is modeled by individual wavelets with a small quality factor.

Figure 9 shows the whitened data and credible intervals for the whitened CBC and glitch reconstruction in each detector for each of the injected signals. Due to the large glitch duration, the signals are injected sufficiently wide apart that the reconstruction plots show non overlapping parts of the data and the glitch. The glitch reconstructions are therefore not expected to match. As

expected from the glitch spectrogram, the glitch is characterized by a series of short high frequency bursts, each of which is modeled by different wavelets within our glitch model. Figure 10 shows posterior distributions for selected source parameters for the left-most injection in blue, as well as the injected parameters. In all cases the recovered parameters are consistent with their injected values, suggesting that the presence of the glitch does not incur biases on the inferred source properties if the two are modeled simultaneously. As before, we also plot results from a CBC+noise run that neglects the glitch in the data in orange and again find small biases by the presence of the glitch in the source intrinsic parameters.

The glitch subtraction process is detailed in Fig. 11 that again shows spectrograms of the original data containing both the glitch and the signal (left), data after a fair draw glitch model has been subtracted (middle), and data after both the glitch and the fair draw CBC model have been removed (right). As before, data from the middle panel could be used for further data processing. The right panel shows data where a model for both the glitch and the CBC have been subtracted. Even though the majority of the glitch power is absent (compare the left and right panels), some small non gaussian power might be left behind. The reason for this is that the blue mountain glitch is manifested as individual short bursts of glitch power, which our flexible analysis attempts to model completely independently. Indeed, the glitch model for this run uses $\mathcal{O}(70)$ wavelets. Each of these wavelets, needs to model sufficient nongaussian power in the data in order to overcome the parsimony penalty incurred by adding more parameters to the model. As such, we expect that some of the weaker "bursts" of the glitch will not be recovered. Possible ways to alleviate this are discussed in Sec. V.

IV. GRAVITATIONAL WAVE EVENTS

As a further demonstration of our CBC+glitch+noise model, we also analyze two astrophysical events, GW170817 [4] and GW150914 [40] whose data are available from GWOSC [32, 33]. Though not the main focus
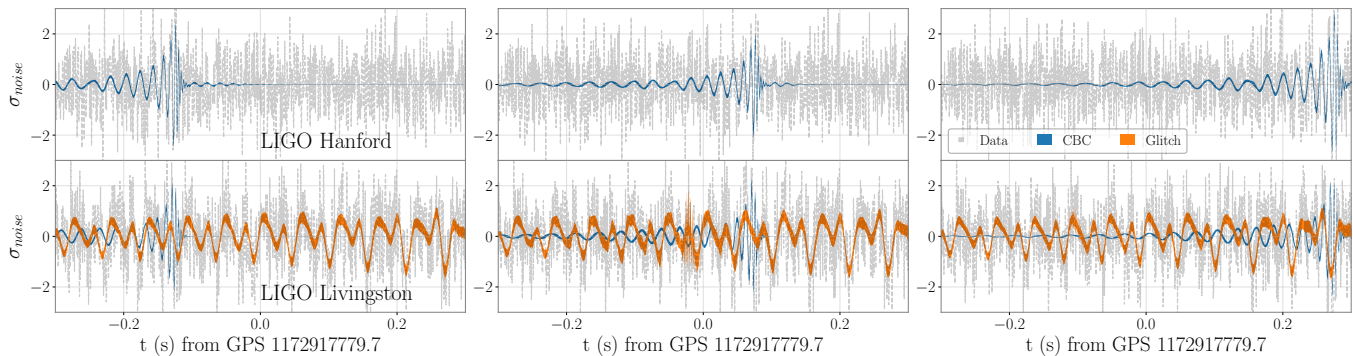
FIG. 6. Credible intervals for the glitch (orange) and the CBC (blue) signal reconstruction for data containing a scatted light glitch in LIGO Livingston and a simulated CBC signal at 3 different times with respect to the glitch (left to right). Shaded regions correspond to 90% credible intervals, while in grey dashed lines we plot the data whitened with a fair draw PSD from our noise model posterior. The top row corresponds to LIGO Hanford while the bottom row corresponds to LIGO Livingston.
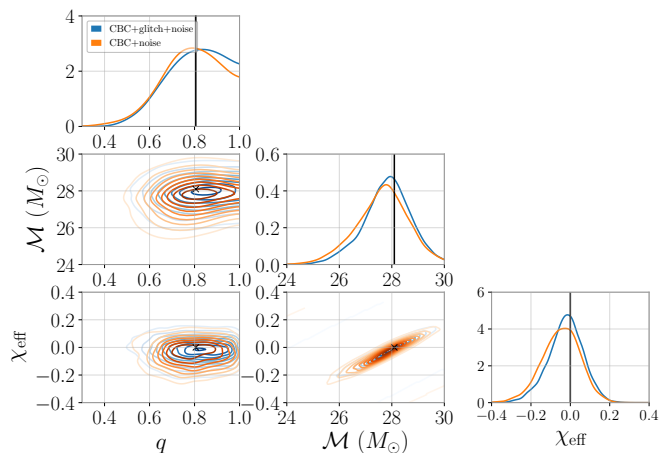

FIG. 7. One- and two-dimensional posterior distribution for selected source parameters of the simulated signal from the left panels of Fig. 6 injected on top of a LIGO Livingston scattered light glitch. We include the mass ratio q, the effective spin $\chi_{\rm eff}$, and the detector frame chirp mass $\mathcal{M}$ posteriors, while black crosses or black vertical lines denote the true parameters of the injection. Blue (orange) contours and lines correspond to the CBC+glitch+noise (CBC+noise) run.

of this paper, the analysis presented below also provides an estimate of the effect of marginalizing over the noise PSD has on the inferred astrophysical parameters. More details about this effect will be presented in a separate study.

A. GW170817

Perhaps the most known instance of a GW signal overlapping with an instrumental glitch is GW170817 [4]. Inference on the GW170817 source properties is performed on data where the glitch in LIGO Livingston has been modeled with `BayesWave`'s glitch-only model

and subtracted. Analysis of simulated signals suggests that this procedure leads to unbiased inference, while any analysis on data that contain the glitch results in highly biased source parameters [27]. Both versions of the data are publicly available, both with and without the glitch [26], so we analyze them both with different models. We use data from the LIGO Hanford and the LIGO Livingston detectors and analyze 64s of data from 16Hz to 2048Hz using the `IMRPhenomD_NRTides` waveform model that includes finite-size effects [41]. We employ our CBC+glitch+noise model on the data with the glitch and the CBC+noise model on data where the glitch has already been subtracted. For the CBC+glitch+noise case we use `GlitchBuster` [23] to provide a quick fit to the glitch and use that as a starting point for our glitch model during sampling.

Credible intervals for the signal and glitch reconstructions are shown in Fig. 12 for each detector for ~ 150ms of data around the glitch. Despite its high SNR, GW170817 had a relatively low amplitude, so the LIGO Hanford plot has been zoomed in to make the signal visible. The LIGO Livingston data are dominated by the glitch, peaking at $\sim 150\sigma$ relative to the background detector gaussian noise. The signal is not visible in the LIGO Livingston data given the plotting scale. Figure 13 shows selected source parameters obtained from data both with and without the glitch. We find consistent results, showing that our combined CBC+glitch+noise analysis can faithfully fit the CBC signal and the glitch simultaneously, without the need for the two step process of first removing the glitch and then reanalyzing the data.

B. GW150914

The first GW signal directly detected by the LIGO detectors, GW150914 [40], did not overlap with an instrumental glitch [42]. However, since it is one of the best studied and loudest signals, we select it as a demonstration of our analysis on data without glitches. Our
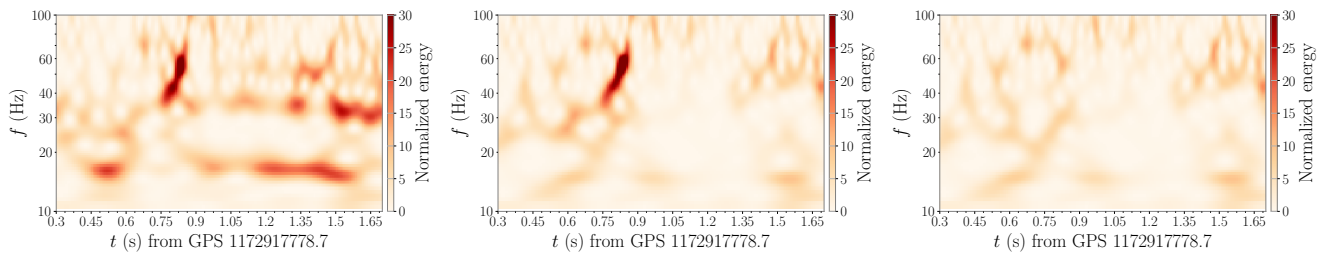
FIG. 8. Spectrogram of the LIGO Livingston data around the time of the scattered light glitch for the leftmost injection from Fig. 6. Left panel: Data containing the scattered light glitch and the simulated CBC signal. Middle panel: Data after a fair draw from the glitch model has been subtracted leaving behind only the chirping CBC signal. Right panel: Data after a fair draw from the glitch and CBC models has been subtracted, leaving behind only gaussian detector noise.
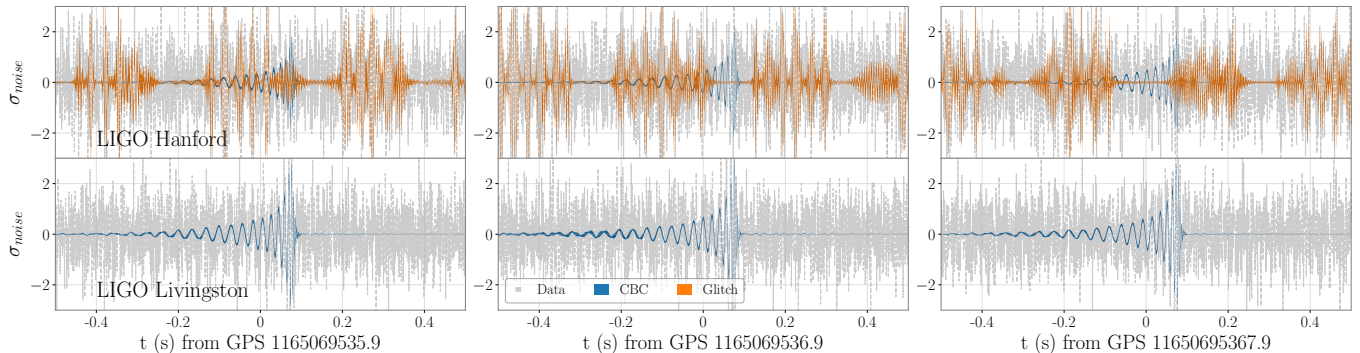

FIG. 9. Credible intervals for the glitch (orange) and the CBC (blue) signal reconstruction for data containing a blue mountain glitch in LIGO Hanford and a simulated CBC signal at 3 different times with respect to the glitch (left to right). Shaded regions correspond to 90% credible intervals, while in grey dashed lines we plot the data whitened with a fair draw PSD from our noise model posterior. The top row corresponds to LIGO Hanford while the bottom row corresponds to LIGO Livingston.

glitch model has the flexibility to use no glitch wavelets, we therefore expect many samples in the glitch model posterior to contain exactly zero glitch power. We analyze 4s of data starting at 16Hz and with a sampling rate of 2048Hz. We perform two runs, one with the CBC+glitch+noise model and one with the CBC+glitch model using otherwise identical settings.

Relevant results are shown in Figs. 14 and 15 where as before we plot the CBC and glitch reconstructions of the CBC+glitch+noise model in the two detectors and the recovered source parameters. The CBC reconstruction of Fig. 14 is consistent with previous results [43]. The glitch reconstruction is too small to identify in the scale of the plot, as we find that 86% and 14% of our posterior samples had exactly zero glitch wavelets in LIGO Hanford and LIGO Livingston respectively. Figure 15 shows the posterior distribution for selected source parameters of GW150914 obtained under the CBC+glitch+noise and the CBC+noise models. The two posteriors yield consistent results, showing that the glitch model does not affect the CBC parameters when no glitch is present in the data.

V. CONCLUSIONS

We construct and validate an analysis of GW data that simultaneously models astrophysical CBC signals and instrumental glitches. We test the analysis against real instances of glitches in the two LIGO detectors from O2 data and simulated CBC signals injected at different times with respect to the glitch. We find that our analysis can separate the two, and provide both estimates for the CBC source parameters and glitch-subtracted data for subsequent analyses. The glitch model we employ is a sum of sine-gaussian wavelets that is not tuned to any specific glitch type and morphology; it can thus handle even novel glitch types that might first appear during O4. Even though this flexibility is desirable given the unpredictable and evolving nature of glitches, the efficacy of glitch subtraction can be improved by employing targeted priors for different glitch types. One such example would be a prior that anticipates arches at frequency multiples in the case of scattered light glitches. We leave such targeted priors to future work.

Our analysis considered only simulated BBH signals, though we also present an analysis of the BNS GW170817. We expect overlapping CBCs and glitches of similar duration to be a worse-case-scenario due to
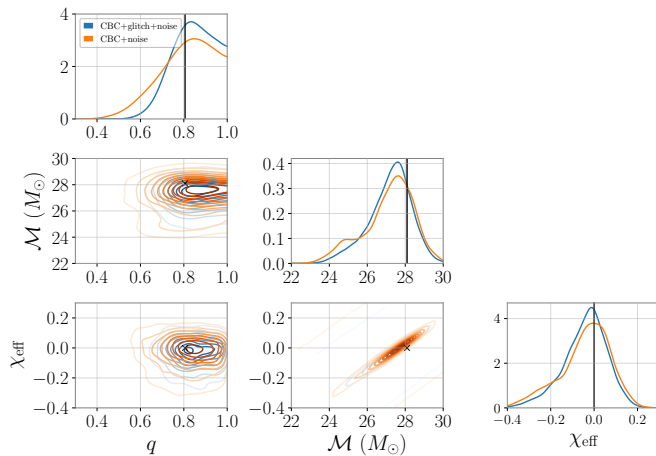
FIG. 10. One- and two-dimensional posterior distribution for selected source parameters of the simulated signal from the left panels of Fig. 9 injected on top of a LIGO Hanford blue mountain glitch. We include the mass ratio q, the effective spin χ_{eff}, and the detector frame chirp mass $\mathcal{M}$ posteriors, while black crosses or black vertical lines denote the true parameters of the injection.

their similar morphology [44, 45]. Given that, we plan to carry out a larger scale study of our CBC+glitch analysis that includes more glitch types and CBC classes, such as BNSs and lower mass BBHs. Additionally, the analysis presented here did not make use of `GlitchBuster` [23] to provide initial fits to the glitch, apart from the GW170817 case. In the future we plan to investigate interfacing `GlitchBuster` and `BayesWave` in more detail, in the hopes that an efficient starting point for

the glitch model during sampling will decrease the sampler's convergence time and result in ready-to-use glitch-subtracted data more quickly. We hope that our analysis will contribute to robust and efficient glitch mitigation against the increased event rate anticipated in O4; our goal is to facilitate analysis of as much data as possible and maximize the science output of the upcoming observations.

[1] J. Aasi et al. (LIGO Scientific), Class. Quant. Grav. 32, 074001 (2015), arXiv:1411.4547 [gr-qc].

[2] F. Acernese et al. (Virgo Collaboration), Class. Quant. Grav. 32, 024001 (2015), arXiv:1408.3978 [gr-qc].

[3] R. Abbott et al. (LIGO Scientific, Virgo), (2020), arXiv:2010.14527 [gr-qc].

[4] B. P. Abbott et al. (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. Lett. 119, 161101 (2017), arXiv:1710.05832 [gr-qc].

[5] B. P. Abbott et al. (LIGO Scientific Collaboration, Virgo Collaboration), Living Rev. Rel. 19, 1 (2013), arXiv:1304.0670 [gr-qc].

[6] 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).

[7] J. Veitch et al., Phys. Rev. D 91, 042003 (2015), arXiv:1409.7215 [gr-qc].

[8] B. P. Abbott et al. (LIGO Scientific, Virgo), Class. Quant. Grav. 37, 055002 (2020), arXiv:1908.11170 [gr-qc].

[9] C. Rover, R. Meyer, and N. Christensen, Class. Quant. Grav. 28, 015010 (2011), arXiv:0804.3853 [stat.ME].

[10] C. Talbot and E. Thrane, (2020), arXiv:2006.05292 [astro-ph.IM].

[11] S. A. Usman et al., Class. Quant. Grav. 33, 215004 (2016), arXiv:1508.02357 [gr-qc].

[12] S. Sachdev et al., (2019), arXiv:1901.08580 [gr-qc].

[13] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, (2019), arXiv:1908.05644 [astro-ph.IM].

[14] R. DeRosa, J. C. Driggers, D. Atkinson, H. Miao, V. Frolov, M. Landry, J. A. Giaime, and R. X. Adhikari, Classical and Quantum Gravity 29, 215008 (2012), arXiv:1204.5504 [physics.ins-det].

[15] V. Tiwari et al., Class. Quant. Grav. 32, 165014 (2015), arXiv:1503.07476 [gr-qc].

[16] G. D. Meadors, K. Kawabe, and K. Riles, Class. Quant. Grav. 31, 105014 (2014), arXiv:1311.6835 [astro-ph.IM].

[17] J. Driggers et al. (LIGO Scientific), Phys. Rev. D 99, 042001 (2019), arXiv:1806.00532 [astro-ph.IM].

[18] D. Davis, T. Massinger, A. Lundgren, J. Driggers, A. Urban, and L. Nuttall, Class. Quant. Grav. 36, 055011 (2019), arXiv:1809.05348 [astro-ph.IM].

[19] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepanczyk, and S. Vitale, Phys. Rev. D 101,
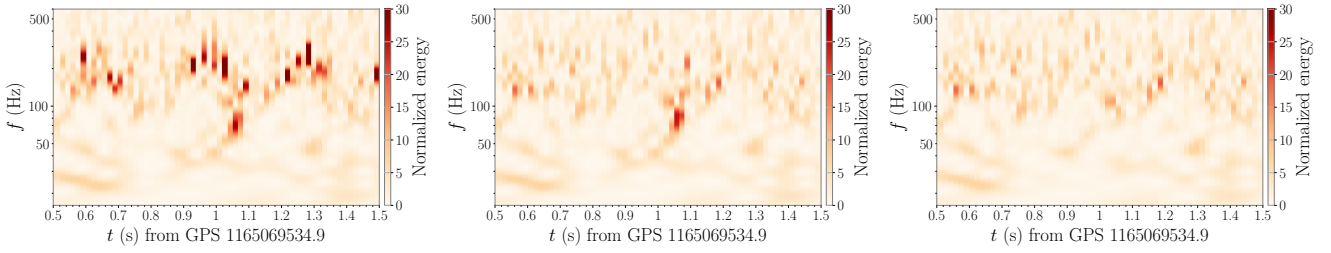
FIG. 11. Spectrogram of the LIGO Livingston data around the time of the blue mountain glitch for the leftmost injection from Fig. 9. Left panel: Data containing the blue mountain glitch and the simulated CBC signal. Middle panel: Data after a fair draw from the glitch model has been subtracted leaving behind only the chirping CBC signal. Right panel: Data after a fair draw from the glitch and CBC models has been subtracted, leaving behind only gaussian detector noise.
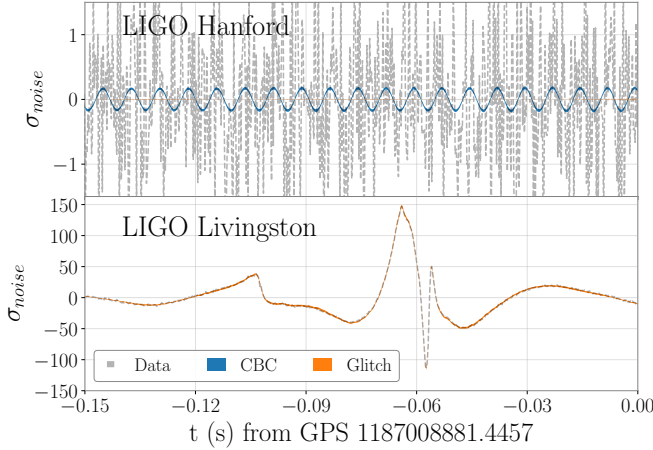

FIG. 12. Credible intervals for the glitch (orange) and the CBC (blue) signal reconstruction for GW170817. Shaded regions correspond to 90% credible intervals, while in grey dashed lines we plot the data whitened with a fair draw PSD from our noise model posterior. The top row corresponds to LIGO Hanford while the bottom row corresponds to LIGO Livingston. The LIGO Hanford plot zooms in to show the signal that is invisible in the LIGO Livingston plot due to the size of the glitch; note the y-scale difference in the two plots
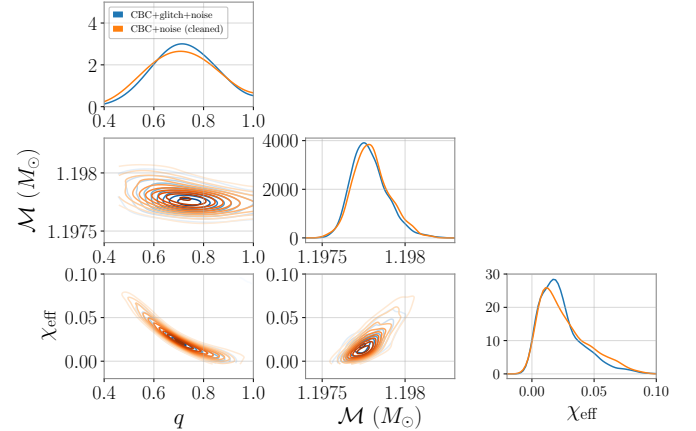

FIG. 13. One- and two-dimensional posterior distribution for selected source parameters for GW170817. We include the mass ratio q, the effective spin χ_{eff}, and the detector frame chirp mass $\mathcal{M}$ posteriors. Blue curves show posteriors under the CBC+glitch+noise model on the full data, while orange curves correspond to the CBC+noise model on data where the glitch has already been subtracted. The two sets of results are consistent with each other.

042003 (2020), arXiv:1911.09083 [gr-qc].

[20] R. Ormiston, T. Nguyen, M. Coughlin, R. X. Adhikari, and E. Katsavounidis, Phys. Rev. Res. **2**, 033066 (2020), arXiv:2005.06534 [astro-ph.IM].

[21] S. Coughlin *et al.*, Phys. Rev. D **99**, 082002 (2019), arXiv:1903.04058 [astro-ph.IM].

[22] N. J. Cornish and T. B. Littenberg, Class. Quant. Grav. **32**, 135012 (2015), arXiv:1410.3835 [gr-qc].

[23] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, (2020), arXiv:2011.09494 [gr-qc].

[24] P. J. Green, Biometrika **82**, 711 (1995), http://oup.prod.sis.lan/biomet/article-pdf/82/4/711/699533/82-4-711.pdf.

[25] T. B. Littenberg and N. J. Cornish, Phys. Rev. **D91**, 084034 (2015), arXiv:1410.3852 [gr-qc].

[26] BayesWave Glitch Subtraction for GW170817, https://dcc.ligo.org/LIGO-T1700406/public.

[27] C. Pankow *et al.*, Phys. Rev. **D98**, 084016 (2018), arXiv:1808.03619 [gr-qc].

[28] LIGO Scientific Collaboration, Virgo Collaboration, "LALSuite," https://git.ligo.org/lscsoft/lalsuite (2018).

[29] N. J. Cornish, (2021), arXiv:2101.01188 [gr-qc].

[30] N. J. Cornish and K. Shuman, Phys. Rev. D **101**, 124008 (2020), arXiv:2005.03610 [gr-qc].

[31] N. J. Cornish, (2016), arXiv:1606.00953 [gr-qc].

[32] Gravitational Wave Open Science Center (GWOSC), https://www.gw-openscience.org/.

[33] R. Abbott *et al.* (LIGO Scientific, Virgo), (2019), arXiv:1912.11716 [gr-qc].

[34] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, Phys. Rev. D **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].

[35] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Phys. Rev. D **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].

[36] M. Cabero *et al.*, Class. Quant. Grav. **36**, 15 (2019), arXiv:1901.05093 [physics.ins-det].
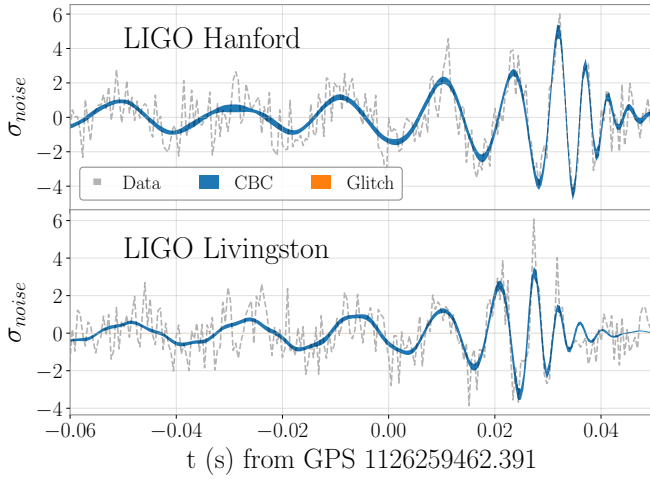
FIG. 14. Credible intervals for the glitch (orange) and the CBC (blue) signal reconstruction for GW150914. Shaded regions correspond to 90% credible intervals, while in grey dashed lines we plot the data whitened with a fair draw PSD from our noise model posterior. The top row corresponds to LIGO Hanford while the bottom row corresponds to LIGO Livingston. Our glitch model recovers essentially no incoherent power coincident with the astrophysical signal and therefore the reconstruction is not visible in the scale of the plot.
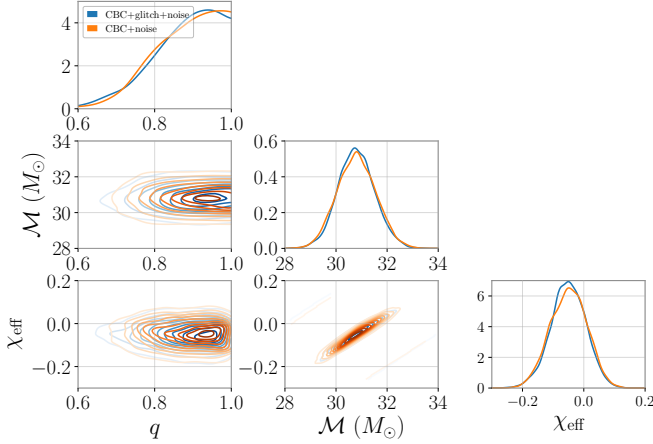

FIG. 15. One- and two-dimensional posterior distribution for selected source parameters for GW150914. We include the mass ratio q, the effective spin χ_{eff}, and the detector frame chirp mass $\mathcal{M}$ posteriors. Blue curves show posteriors under the CBC+glitch+noise model, while orange curves correspond to the CBC+noise model. The two sets of results are consistent with each other.

[37] J. Powell, Class. Quant. Grav. **35**, 155017 (2018), arXiv:1803.11346 [astro-ph.IM].

[38] T. Accadia *et al.*, Classical and Quantum Gravity **27**, 194011 (2010).

[39] S. Soni *et al.*, arXiv e-prints , arXiv:2007.14876 (2020), arXiv:2007.14876 [astro-ph.IM].

[40] B. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[41] T. Dietrich *et al.*, Phys. Rev. **D99**, 024029 (2019), arXiv:1804.02235 [gr-qc].

[42] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Class. Quant. Grav. **33**, 134001 (2016), arXiv:1602.03844 [gr-qc].

[43] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 241102 (2016), arXiv:1602.03840 [gr-qc].

[44] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Class. Quant. Grav. **35**, 065010 (2018), arXiv:1710.02185 [gr-qc].

[45] D. Davis, L. V. White, and P. R. Saulson, Class. Quant. Grav. **37**, 145001 (2020), arXiv:2002.09429 [gr-qc].

[46] 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).

[47] J. D. Hunter, Computing In Science & Engineering **9**, 90 (2007).