Concurrent estimation of noise and compact-binary signal parameters in gravitational-wave data

Cailin Plunkett[,1,*] Sophie Hourihane[,2] and Katerina Chatziioannou[2,3]

[1]*Department of Physics and Astronomy, Amherst College, Amherst, Massachusetts 01002, USA*
[2]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*
[3]*LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA*

Gravitational-wave parameter estimation for compact binary signals typically relies on sequential estimation of the properties of the detector Gaussian noise and of the binary parameters. This procedure assumes that the noise variance, expressed through its power spectral density, is perfectly known in advance. We assess the impact of this approximation on the estimated parameters by means of an analysis that simultaneously estimates the noise and compact binary parameters, thus allowing us to marginalize over uncertainty in the noise properties. We compare the traditional sequential estimation method and the new full marginalization method using events from the GWTC-3 catalog. We find that the recovered signals and inferred parameters agree to within their statistical measurement uncertainty. At current detector sensitivities, uncertainty about the noise power spectral density is a subdominant effect compared to other sources of uncertainty.

I. INTRODUCTION

Estimating the source properties of compact binary coalescences (CBC) using data from gravitational-wave (GW) observatories [1,2] relies on accurate models of both the astrophysical signal and the noise in the detector. While the signal model employs waveform templates that are based on analytic approximations to Einstein's field equations, the noise model typically takes the form of the power spectral density (PSD) of the detector Gaussian noise, $S_n(f)$ [3]. Given GW data $d = h + n$ that are a sum of a coherent signal h between detectors and uncorrelated random noise n, and an accurate signal model h', the residual $r = d - h'$ should follow the same distribution as the noise. Assuming Gaussian and stationary noise, the likelihood function $\mathcal{L}(d|h')$ in the frequency domain is [4,5]

$$\mathcal{L}(d|h') = \prod_i \frac{2}{\pi T S_n(f_i)} \exp\left[-2\frac{|\tilde{r}_i|^2}{T S_n(f_i)}\right], \quad (1)$$

which explicitly depends on $S_n(f)$; T is the duration of the data analyzed and i counts frequency bins.

The likelihood of Eq. (1), and thus analyses of GW signals, depend on multiple formal assumptions about the detector noise and its PSD. Parameter estimation [6–8] assumes that the noise is Gaussian and stationary in time, though mitigation techniques and methods to relax these assumptions have been explored [9–13]. Furthermore, the

PSD is typically a required input for CBC parameter estimation. Using a point estimate for $S_n(f)$ assumes the noise PSD is perfectly known or at least that is it estimated to precision far greater than other statistical uncertainties. The noise PSD must be estimated from the data [3], typically in one of two ways. The first, termed the "off-source" method, relies on the median PSD of many data segments around, but not including, the data segment containing the signal [6]. This calculation requires the stationarity assumption to hold over the entire stretch of data used to estimate the PSD, typically $\mathcal{O}(10^3\ \mathrm{s})$ [6,14]. The second, termed the "on-source" method, uses only the data segment containing the signal.

The "on-source" method relies on BayesLine [15], integrated into the broader BayesWave algorithm [12,16], and models the PSD with a two-component phenomenological fit. A cubic spline describes the broadband noise behavior, while narrow-band spectral features are fit using Lorentzians. The number and parameters of the spline points and Lorentzians are marginalized over with a trans-dimensional Markov chain Monte Carlo (MCMC) algorithm [17], also known as a reversible-jump MCMC (RJMCMC). Any non-Gaussian components of the data (namely, instrumental glitches or astrophysical signals) are modeled with sine-Gaussian wavelets, thus relaxing the assumption that the noise is Gaussian during PSD estimation. By using the same data segment for both PSD and signal parameter estimation, the on-source method additionally restricts the stationarity assumption to a shorter data segment, typically $\mathcal{O}(4\ \mathrm{s})$. A point estimate from the

---

*cplunkett23@amherst.edu

104021-1

PSD posterior, typically a fair draw or the median, is then selected and passed on for downstream analyses. The on-source and off-source methods were compared in [18] where it was demonstrated that the former results in PSDs that are more consistent with the Gaussian likelihood assumptions. Both methods, however, still provide a single point-estimate for the PSD.

Traditional parameter estimation [19–21] involves sequential estimation of the noise PSD and the binary parameters. Once a point-estimate for the noise PSD has been provided with either the on-source or the off-source methods, LALInference [6], BILBY [7,8], or other similar pipelines marginalize over the binary parameters using waveform templates as the signal model. During this two-step sequential process, care must be taken to ensure that the spectral resolution and maximum and minimum frequencies of the estimated PSD match the desired parameter estimation analysis settings. Besides these technical complications, the sequential process assumes that the PSD is perfectly estimated with negligible uncertainty, or at least that any uncertainty is not correlated with the binary parameters.

Assessing the impact of the assumption that the noise PSD is perfectly known requires marginalizing over noise uncertainty during parameter estimation. The choice of PSD estimation method and PSD prior results in different functional forms for the marginalized CBC likelihood [15,22–29]. Röver *et al.* [22] showed that assuming a PSD prior that is an inverse χ^2-distribution allows one to analytically marginalize over PSD uncertainty and converts the standard Whittle likelihood to a Student-t distribution. Talbot and Thrane [29] extended [22] and computed the marginalized likelihood for both the mean and median off-source PSD. Finally, Biscoveanu *et al.* [28] explored a method of incorporating PSD uncertainty in CBC parameter estimation by considering the full PSD posterior from BayesLine rather than a point-estimate. They still analyzed the data sequentially, but used 200 fair draws from the PSD posterior to perform 200 parameter estimation analyses and combined the resulting posteriors. Besides the computational cost, one drawback of this method is that it ignores potential correlations between the PSD and the CBC parameter posterior.

Both Biscoveanu *et al.* [28] and Talbot and Thrane [29] concluded that PSD uncertainty is likely a subdominant effect compared to typical current statistical uncertainties given their respective PSD and binary parameter estimation methods. Biscoveanu *et al.* [28] found posteriors under the median and marginalized on-source PSD that agree in position and width to within a few percent, which they determine is an order of magnitude higher than expected due to statistical fluctuations. Among the GWTC-1 events [19], GW151012 [30] displays the largest effect due to uncertainty on the LIGO Hanford PSD. However, the choice of PSD estimation method, specifically the off-source mean or median studied in [29], had a greater impact on the inferred

parameter posterior than PSD marginalization. Similar conclusions were reached in [31] in the context of the impact of PSD misestimation on inferring the tidal deformability of neutron star binaries. The impact of marginalization could be more significant when precise calculations of Bayesian evidence are needed [29].

In this study we revisit the issue of uncertainty in the PSD estimation and its effect on the inferred CBC parameters on events from GWTC-3 [21]. We take advantage of recent developments on the BayesWave algorithm[1] that augment its list of models with a CBC model that describes the signal with waveform templates [10,32]. This allows us to *simultaneously* model the noise PSD and the GW signal from the same on-source data, thus obtaining CBC parameter posteriors that are marginalized over noise uncertainty. Compared to the approach of [28], our method allows us to obtain a joint posterior for the noise and the signal rather than assume the signal and noise posteriors are uncorrelated and obtained by sequential analysis of the same data.

We apply our analysis to binary black hole (BBH) signals from GWTC-3 [21]. To reduce computational cost, we restrict to confident detections with signal-to-noise ratio (SNR) > 10. We further discard signals with analysis segments > 16 seconds and events with technical complications of different lower frequency cutoffs in different detectors, for a total sample of 50 events. We analyze each event with two methods. The *sequential* method follows the procedure employed for parameter estimation in [21] by first estimating the median for the on-source noise PSD and then using this point estimate to sample from the posterior of the CBC source parameters. The *marginalized* method samples from the joint posterior of the noise PSD and CBC source parameters. We compare the resulting source parameter posteriors and find broad agreement between the two methods. In particular, over all events the marginalized method results in an average change of $0.0^{+1.4}_{-1.6}\%$ in the marginalized chirp mass posterior median and a decrease of $2^{+16}_{-13}\%$ in the marginalized posterior 90% width. Neither are statistically distinguishable from zero. In addition, we find no significant trend between the change in posterior width and the posterior median or signal SNR. Finally, we devise a simple toy model to demonstrate how a broadband increase or reduction in the noise PSD results in a higher-order, and thus subdominant, effect in the CBC parameter posteriors.

The rest of the paper is organized as follows. In Sec. II, we describe the BayesWave algorithm and models for the GW signal available within it. In Sec. III, we discuss our analyses on a set of events from GWTC-3 [21], considering the waveform reconstructions as well as the parameter posteriors. We use a toy model of a shifted PSD to help explain the minimal impact of PSD marginalization on parameter posteriors in Sec. IV. Finally, in Sec. V we conclude.

---

[1]From now on we use the term BayesWave to describe the combined BayesLine and BayesWave algorithms.

II. METHODS AND ANALYSIS

In this section, we describe the features of BayesWave that are relevant to our study as well as details about our analysis of GWTC-3, for which we use data from the Gravitational Wave Open Science Center (GWOSC) [33,34].

A. Algorithm description

The BayesWave algorithm is described in detail in [5,10,16,32] and here we describe the aspects relevant to this work. BayesWave simultaneously models GW signals, detector noise, and glitches using different models. We provide a glossary of terms in Table I, a schematic algorithm workflow in Fig. 1, and further describe each model below.

(i) The *noise model* uses the BayesLine algorithm, fully integrated into BayesWave, to model the PSD $S_n(f)$ as a sum of a broadband spline and Lorentzians for spectral lines [15]. The number and parameters of the spline control points and Lorentzians are marginalized over with an RJMCMC. In the workflow of Fig. 1, the noise RJMCMC is depicted with a blue box.

(ii) The *CBC model* uses waveform templates derived within general relativity to model a CBC signal. The CBC sampler and its integration within BayesWave are described in detail in [32]. For this work, we use the IMRPhenomD [35,36] waveform model with identical settings and priors as [32]. The CBC likelihood is computed with the heterodyning method to reduce the computational cost [37,38]. In the workflow of Fig. 1, the CBC model parameters are split between the extrinsic MCMC that handles parameters θ_{ext} (purple box) and the CBC MCMC that handles parameters θ^{cbc} (pink box); see [32] for details.

(iii) The *glitch model* targets any incoherent non-Gaussian noise in the detectors by modeling such noise with a sum of sine-Gaussian wavelets. The number and morphology of the wavelets are not fixed *a priori* but instead marginalized over with an RJMCMC. This model is typically used to target instrumental glitches, but can capture any excess power that is unaccounted for, including signals. In the workflow of Fig. 1, the glitch RJMCMC is depicted with a yellow box and its parameters are collectively denoted as θ_{w}.[2]

BayesWave samples from the multidimensional posterior of all its models using a blocked Gibbs sampler. Each model is sampled iteratively using independent MCMC or RJMCMC samplers that update their corresponding block

---

[2]For completeness we also mention the *signal model* though it is not used in this study and thus not depicted in Fig. 1. The *signal model* targets astrophysical signals by modeling coherent power between the detectors through a sum of sine-Gaussian wavelets. The number and morphology of the wavelets are again marginalized over with an RJMCMC. The *signal model* is very similar to the *glitch model*, only the wavelets are coherent rather than incoherent across the detector network.
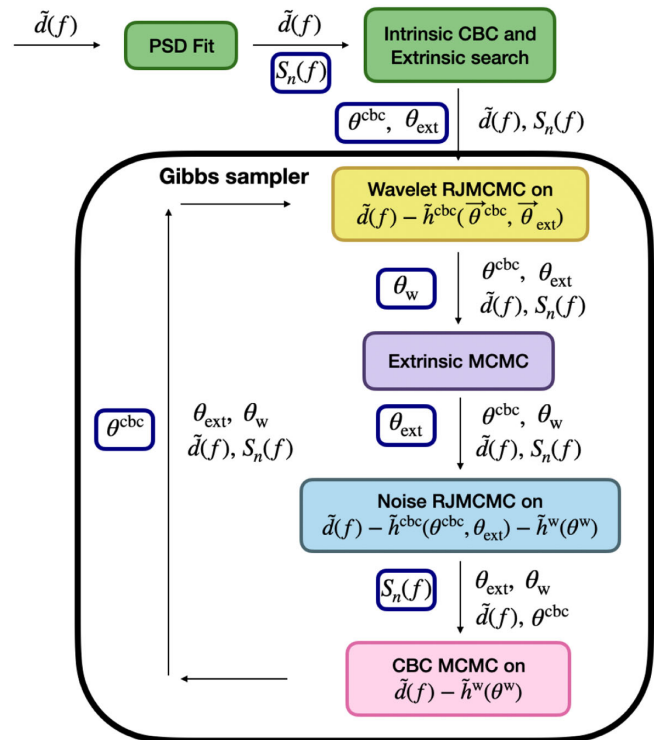

FIG. 1. BayesWave workflow diagram. Preconditioned frequency-domain data $\tilde{d}(f)$ undergo quick initial parameter fits to use as starting points for the blocked Gibbs sampler (black box). The Gibbs sampler cycles through the component samplers, which are given in filled colored boxes with their names and input parameters, using independent MCMCs. The yellow (purple, blue, pink) box represents the glitch (extrinsic parameter, noise PSD, intrinsic CBC parameter) sampler. Below each component sampler, the parameters that are held fixed (sampled) are unboxed (boxed in navy).

of parameters while keeping others parameters fixed. Each independent sampler proceeds for a given number of iterations, before returning the last set of parameters and switching to the next sampler. This procedure provides an efficient way of approximating the joint distribution of all parameters, provided that the parameters across blocks are not too correlated. The breakdown of parameters within blocks is described in [32], while [11] shows that this procedure converges even when the parameters across blocks are correlated.

BayesWave can be used with any combination of these models, resulting in different analysis modes as shown in Table I. The "cleaning phase" is the standard BayesWave mode used to compute median PSDs for analyses that use a point-estimate for the PSD. Data in each detector are analyzed separately with the *noise model* and the *glitch model*. The former targets the noise PSD, while the latter targets any excess non-Gaussian power (glitch or GW signal) to ensure that its presence does not affect PSD estimation. A posterior for the noise (and the excess power) is returned from which a median can be computed for further parameter estimation analyses such as [19–21].

TABLE I.   BAYESWAVE term glossary. The first column gives the model name, while the second column provides a brief description of its configuration. The third, fourth, and fifth columns correspond to the three analyses modes used in this study and denote which models are active during each.

		Analysis mode		
Model name	Model configuration	Cleaning Phase	CBC + noise	CBC + fixed
CBC model	GW signal with CBC templates	✗	✓	✓
glitch model	Incoherent power with wavelets	✓	✗	✗
noise model	Noise PSD with splines and Lorentzians	✓	✓	✗
signal model	GW signal with wavelets	✗	✗	✗

During the "CBC + noise" analysis mode, data from all detectors are analyzed simultaneously. The GW signal is modeled with the *CBC model* while the noise PSD in each detector is modeled with the *noise model* and thus fully marginalized over. For comparison to the standard analysis of [19–21], we also use the "CBC + fixed" analysis mode where again the GW signal is modeled with the *CBC model*, but now the noise PSD is fixed, typically acquired from a previous "Cleaning Phase" analysis.

B. Analysis description and data

We analyze the BBH events from GWTC-3 with SNR above 10, excluding events that contain neutron stars to limit computational cost. A list of events and analysis settings are provided in Table II. In order to examine the effect of uncertainty on the noise PSD on inference, we perform two analyses on each event, which we label *sequential* and *marginalized*.

TABLE II.   Table of events and analysis settings. The columns in order provide the event name, its geocenter GPS time, the analysis segment length, the analysis sampling rate, the detector network (H: LIGO Hanford, L: LIGO Livingston, V: Virgo), and network SNR as reported in the GW Open Science Center. For all events, we use a lower frequency cutoff of $f_{\text{low}} = 20$ Hz.

Event	GPS time (s)	Segment length (s)	Sampling rate (Hz)	Detector network	Network SNR
GW150914	1126259462.420	4	2048	H, L	24.4
GW151012	1128678900.400	4	2048	H, L	10.0
GW151226	1135136350.647	8	2048	H, L	13.1
GW170104	1167559936.599	8	2048	H, L	13.0
GW170729	1185389807.328	4	1024	H, L, V	10.2
GW170809	1186302519.747	4	2048	H, L, V	12.4
GW170814	1186741861.527	4	2048	H, L, V	15.9
GW170818	1187058327.082	4	2048	H, L, V	11.3
GW170823	1187529256.518	4	2048	H, L	11.5
GW190408_181802	1238782700.286	8	2048	H, L, V	14.7
GW190412	1239082262.170	8	4096	H, L, V	18.9
GW190421_213856	1239917954.260	4	1024	H, L	10.6
GW190503_185404	1240944862.298	4	2048	H, L, V	12.1
GW190512_180714	1241719652.419	8	2048	H, L, V	12.3
GW190513_205428	1241816086.747	4	2048	H, L, V	12.3
GW190517_055101	1242107479.838	4	2048	H, L, V	10.2
GW190519_153544	1242315362.398	4	1024	H, L, V	12.1
GW190521_030229	1242442967.460	4	512	H, L, V	14.3
GW190521_074359	1242459857.466	4	1024	H, L	24.4
GW190602_175927	1243533585.089	4	1024	H, L, V	12.1
GW190620_030421	1245035079.311	4	1024	L, V	10.9
GW190630_185205	1245955943.180	4	2048	L, V	15.6
GW190701_203306	1246048404.580	4	1024	H, L, V	11.6
GW190706_222641	1246487219.346	4	2048	H, L	12.4
GW190707_093326	1246527224.169	16	4096	H, L	13.0
GW190708_232457	1246663515.384	8	4096	L, V	13.1
GW190720_000836	1247616534.707	16	4096	H, L, V	11.7
GW190728_064510	1248331528.534	16	8192	H, L, V	13.6
GW190828_063405	1251009263.800	4	2048	H, L, V	16.0

(Table continued)

TABLE II. *(Continued)*

Event	GPS time (s)	Segment length (s)	Sampling rate (Hz)	Detector network	Network SNR
GW190828_065509	1251010527.890	8	2048	H, L, V	11.1
GW190910_112807	1252150105.324	4	1024	L, V	13.4
GW190915_235702	1252627040.692	8	2048	H, L, V	13.1
GW191109_010717	1257296855.217	4	1024	H, L	17.3
GW191129_134029	1259070047.197	16	8192	H, L	13.1
GW191204_171526	1259514944.092	8	4096	H, L	17.5
GW191215_223052	1260484270.334	8	2048	H, L, V	11.2
GW191216_213338	1260567236.472	16	8192	H, V	18.6
GW191222_033537	1261020955.124	8	1024	H, L	12.5
GW191230_180458	1261764316.407	4	2048	H, L, V	10.4
GW200112_155838	1262879936.091	4	2048	L, V	19.8
GW200128_022011	1264213229.901	4	1024	H, L	10.6
GW200129_065458	1264316116.433	8	2048	H, L, V	26.8
GW200202_154313	1264693411.556	16	8192	H, L, V	10.8
GW200208_130117	1265202095.950	4	1024	H, L, V	10.8
GW200219_094415	1266140673.197	4	2048	H, L, V	10.7
GW200224_222234	1266618172.402	4	1024	H, L, V	20.0
GW200225_060421	1266645879.396	8	2048	H, L	12.5
GW200302_015811	1267149509.516	8	2048	H, V	10.8
GW200311_115853	1267963151.300	4	2048	H, L, V	17.8
GW200316_215756	1268431094.158	16	4096	H, L, V	10.3

(1) The *sequential* method is constructed to closely follow the traditional fixed PSD analysis. First, we use the "cleaning phase" mode (see Table I) to obtain a posterior for the PSD. From that posterior, we compute the per-frequency-bin median PSD. Second, we use the "CBC + fixed" mode with the previously computed median PSD as input.

(2) The *marginalized* method includes a single analysis of the data with the "CBC + noise" mode (see Table I) where the CBC and PSD are inferred simultaneously.

III. RESULTS ON GWTC-3 EVENTS

In this section we present the results of the *sequential* and *marginalized* noise methods for each event. We begin with an in-depth discussion of GW150914 [39] and then present results for all events.

A. GW150914

We show the time- and frequency-domain reconstructions of GW150914 in Fig. 2 with both methods. We find that the recovered waveforms agree in median and uncertainty; despite the marginalized method adding free parameters to the analysis, the resulting reconstructions have comparable widths. The marginalized and median PSDs agree quite well overall, with minor deviation around the low-frequency spectral lines in LIGO Hanford.

We then consider the posteriors for select CBC parameters. In Fig. 3 we show one- and two-dimensional marginalized posteriors for the detector-frame chirp mass,

$$\mathcal{M} \equiv \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}; \tag{2}$$

the effective spin,

$$\chi_{\text{eff}} \equiv \frac{m_1 \chi_1 \cos\theta_1 + m_2 \chi_2 \cos\theta_2}{m_1 + m_2}, \tag{3}$$

which is a relatively well-measured spin parameter [40]; the mass ratio $q \equiv m_2/m_1 < 1$; and the luminosity distance D_{L}. In the above equations, m_i, χ_i, and $\cos\theta_i$ with $i \in \{1, 2\}$ are the component masses, spin magnitudes, and angles between the spin and the Newtonian orbital angular momentum, respectively. As with the waveforms, the parameter posteriors are consistent within the statistical measurement uncertainty. For all parameters, the median appears shifted to slightly larger values but the effect is small, with only a 0.4% shift in the chirp mass. The marginalized method additionally recovers ~1% higher network SNR, which is a minimal difference possibly attributable to the slightly closer distance posterior.

B. Remaining events

We present the one-dimensional marginalized posterior for $\mathcal{M}$, χ_{eff}, D_{L}, and q for the entire set of GWTC-3 events in Fig. 4. While we find small variation in posterior position and width between the marginalized and the sequential methods, overall there is broad agreement.

To quantify the difference between the methods, we compute the Jensen-Shannon divergence (JSD) [41]
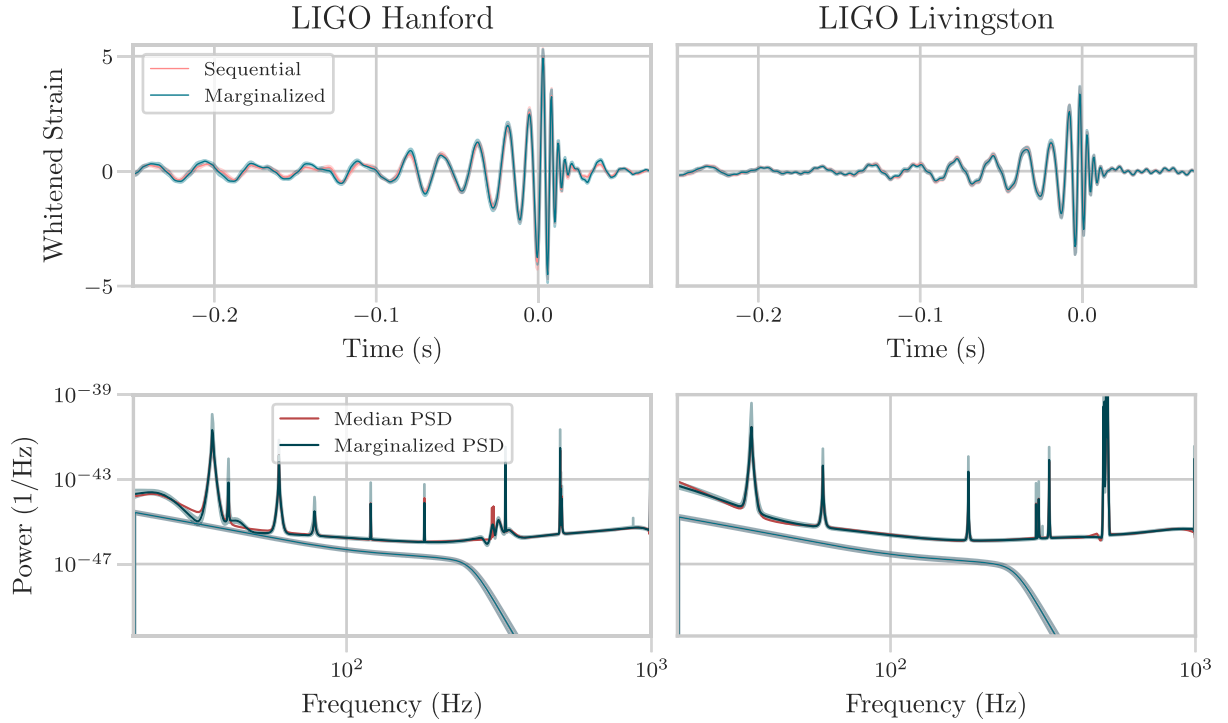
FIG. 2. Whitened time-domain reconstruction (top) and spectrum (bottom) of GW150914 in LIGO Hanford (left) and LIGO Livingston (right). The median and 90% credible interval for the signal are shown in pink and teal for the sequential and the marginalized analyses, respectively. The red line gives the median PSD used for the sequential analysis, while the median and 90% credible interval for the PSD from the marginalized analysis are in dark blue. Time is with respect to GPS 1126259462.42.
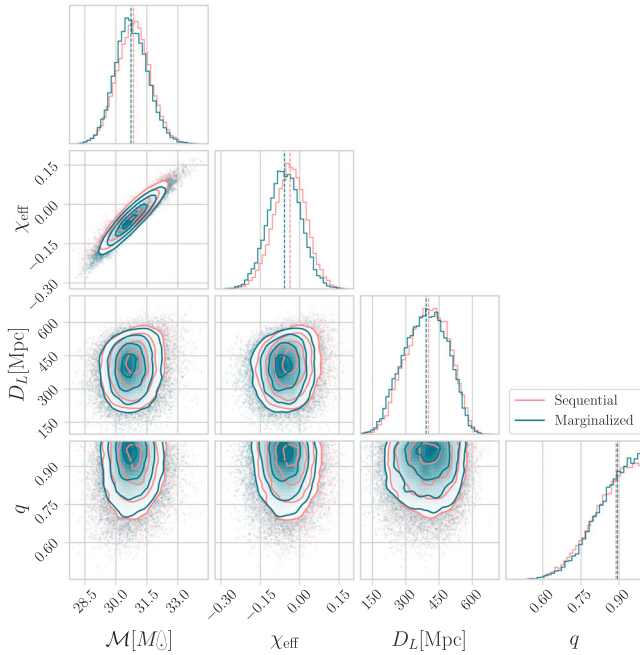

FIG. 3. One- and two-dimensional marginalized posterior for the detector-frame chirp mass $\mathcal{M}$, the effective spin χ_{eff}, the luminosity distance D_{L}, and the mass ratio q of GW150914. We show results obtained with the sequential and the marginalized noise estimation methods in pink and teal, respectively. The methods give consistent results to within the statistical uncertainty.

between the sequential and marginalized posteriors. This summary statistic describes the similarity of two distributions and ranges from 0 for identical distributions to 1 for dissimilar distributions. In Fig. 5, we show the JSD for the marginalized $\mathcal{M}, \chi_{\text{eff}}, D_L$, and q posteriors as a function of the network matched-filter SNR as reported in GWOSC [33,34]. The average JSD is 0.006 across all events and parameters and the JSD is smaller than 0.05 for all except one, indicating overall agreement between the methods [19]. The largest JSD is 0.052, for the chirp mass of GW190512_180714. The magnitude arises from a sizeable posterior width discrepancy, which is further discussed below and is shown in Appendix to depend on the analysis low frequency cutoff. Overall, the difference between the marginalized and sequential posteriors is below the difference in posteriors induced by choice of waveform model [19]. In addition, the JSD shows minimal trend with SNR with the exception of a slight trend of increasing chirp mass JSD with SNR. However, the underdensity of events at high SNR makes it premature to confirm a relation.

Further, we examine the chirp mass posterior width in Fig. 6, where we plot the relative difference between the two methods as a function of the posterior median (i.e., the event chirp mass) and the SNR. We find no significant trends between these parameters, concluding that the effect of marginalizing over PSD uncertainty does not depend on the chirp mass or signal strength. Indeed, a linear fit of the
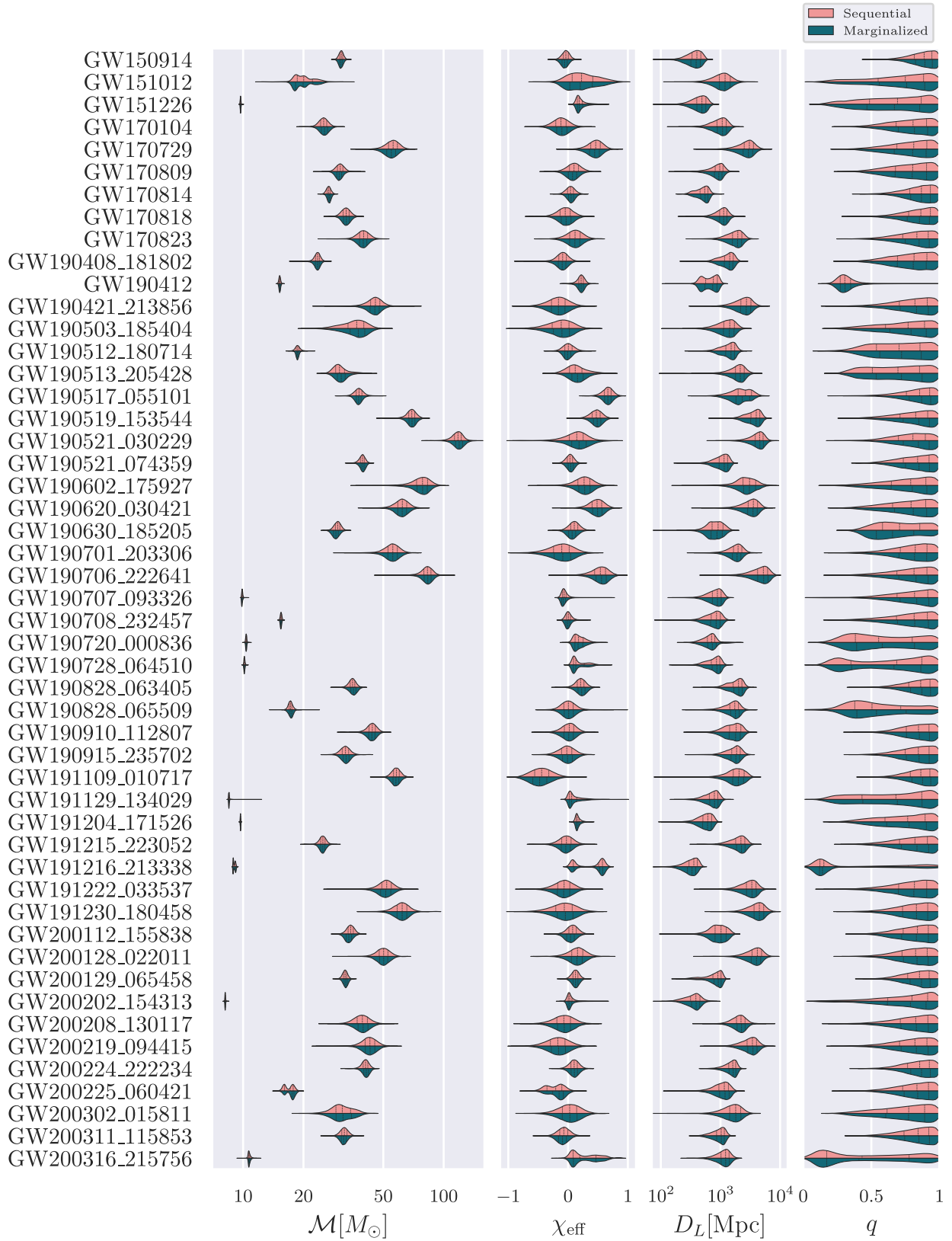
FIG. 4.   Violin plots of the posterior probability density of the sequential (pink) compared to the marginalized (teal) noise estimation method for the detector-frame chirp mass $\mathcal{M}$, effective spin χ_{eff}, luminosity distance D_L, and mass ratio q of each event. The dashed and dotted lines depict the quartiles for each posterior.
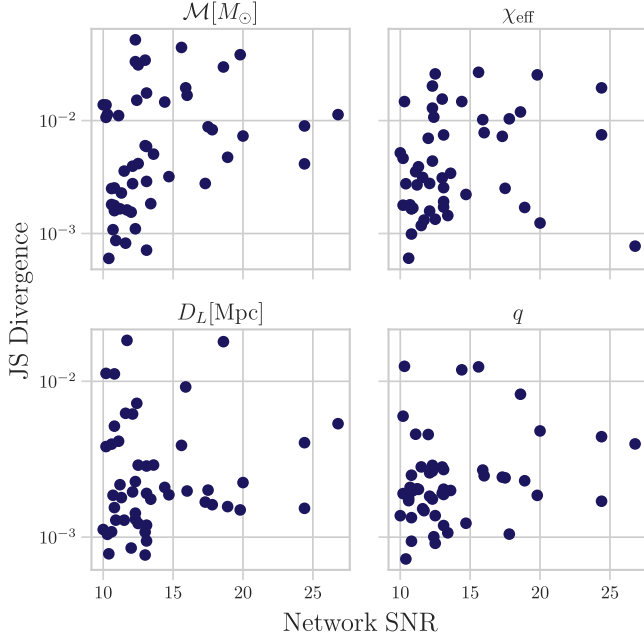
FIG. 5. Jensen-Shannon divergence (JSD) between the sequential and marginalized parameter posteriors as a function of the network matched-filter SNR for $\mathcal{M}$, χ_{eff}, D_L, and q. Each dot corresponds to one event from Table II. Across all events we find JSDs ≤ 0.05.

data of Fig. 6 returns a slope consistent with zero. The biggest outlier is GW190512_180714, for which we obtain a $\sim$40% *decrease* in posterior width using the marginalized PSD. Upon further testing, reducing the lower frequency limit to $f_{\mathrm{low}} = 16$ Hz from $f_{\mathrm{low}} = 20$ Hz reduces these differences. We apply the same test to another event with a large posterior width discrepancy, GW190513_205428 ($\sim$25% decrease) and reach the same conclusion. We continue to use the standard settings for the remaining analyses, but present the $f_{\mathrm{low}} = 16$ Hz results in Appendix.
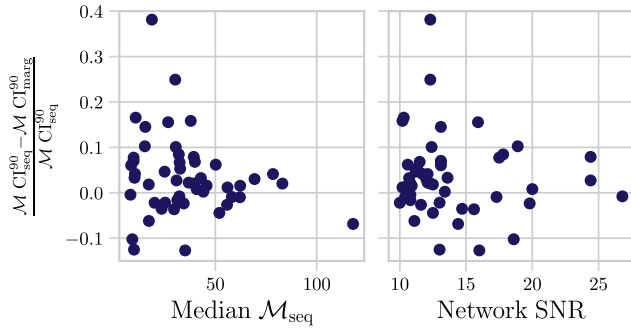
FIG. 6. Fractional difference between the width of the 90% credible interval of the sequential and marginalized methods as a function of the network matched-filter SNR and the median posterior $\mathcal{M}$. We find no overall trend.

IV. TOY MODEL OF PSD VARIATION

The results of Sec. II B suggest that the sequential noise estimation method results in CBC parameter posteriors that are qualitatively similar to the marginalized method. The main difference between the two is that the former uses the median of the PSD posterior, while the latter uses the full posterior. To further explore the difference between the two, we consider what happens if we instead had selected the upper or lower 90% estimate of the PSD with the sequential method. A larger noise PSD should lead to underestimation of the signal SNR, and as a consequence, overestimation of the posterior uncertainty. Similarly, a lower PSD should lead in more narrow posteriors. In effect, we expect a correlation between the noise posterior height and the CBC parameter posterior width.

To explore this behavior in the data, we revisit the "CBC + fixed" analysis for GW159014 using not only the median PSD, but also the lower and upper 90% credible interval PSDs. The chirp mass posterior for each analysis is plotted in the top panel of Fig. 7; they show the expected posterior broadening and narrowing but also a small shift in the posterior median. The bottom panel of Fig. 7 shows the width of the 90% credible interval for $\mathcal{M}$ as a function of
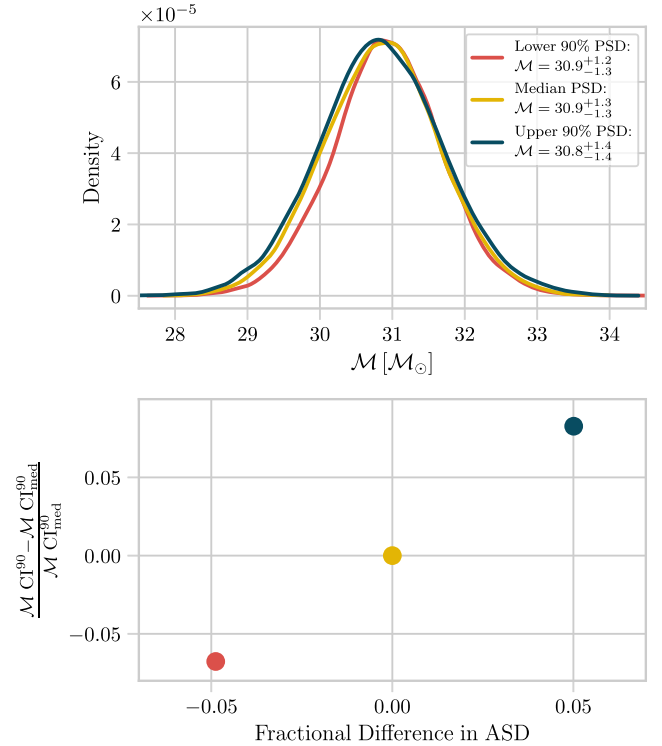

FIG. 7. Top: marginalized detector-frame chirp mass posterior for GW150914 with the "CBC + fixed" analysis using the median, lower 90%, and upper 90% estimate for the PSDs. Bottom: fractional difference between the 90% credible interval widths of the chirp mass posteriors from the top panel against the fractional difference between amplitude spectral density and the median at 100 Hz.

the height of the amplitude spectral density (i.e., the square root of the PSD that is proportional to the SNR and thus the posterior width), given as fractional difference to the median at 100 Hz. Though with three points it is premature to conclude that the trend is linear, it is still evident that the uncertainty increases as the PSD increases (and the SNR decreases). In the case of GW159014, we also find an increase in the posterior median with decreased PSD height. However, tests on other events confirm the trend of the bottom panel of Fig. 7 regarding the posterior width, but no trend regarding the median. This suggests that the median shift is not a generic feature and further confirms the expected correlation between the posteriors.

We further explore a toy model to illuminate this behavior. When marginalizing over PSD uncertainty, each step in the CBC MCMC uses a different PSD that is drawn from the PSD posterior. Overall, the sampler therefore uses a range of PSDs that can be either higher or lower than the median, and thus the CBC posterior width of each sample will be either increased or decreased compared to the median. As a result, the CBC posterior that results from marginalizing over PSD parameters will be a collection of samples from posteriors with different standard deviations. Simplifying to the case of a Gaussian posterior, this effect is comparable to a normal distribution integrated over a range of standard deviations. The difference between the posteriors from using a fixed median PSD (the sequential method) versus the full PSD posterior (the marginalized method) can then be approximated by the difference between a normal distribution and one integrated over a range of standard deviations,

$$\delta(\theta) \equiv \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathcal{N}(\theta|\mu, (1+x)\sigma)dx - \mathcal{N}(\theta|\mu, \sigma), \quad (4)$$

where θ is the parameter of interest, μ is the mean of the posterior, σ is the standard deviation when using the median PSD, and ε encodes the relative PSD uncertainty. Here we choose ε symmetric about σ, which arises from the expectation (and observation, as shown in Fig. 7) that if the ASD uncertainty is Gaussian and thus symmetric about the median, then the resulting range of posterior widths will also be roughly symmetric. Figure 8 shows $\delta(\theta)$ for $\mu = 0$, $\sigma = 1$, and several values of ε. A Taylor expansion around $\varepsilon = 0$ shows that the difference between the two distributions peaks at their mean and goes as $\delta(\theta) \sim \varepsilon^2$. This scaling suggests that a small PSD uncertainty (and thus a small ε) results in a suppressed effect in the parameter posteriors as it is a higher order effect. Indeed as shown in Fig. 7, the typical PSD uncertainty at current sensitivities is $\varepsilon \sim 0.05$, which is much smaller than the typical width of the posteriors.

V. CONCLUSION

In this study, we assessed the impact of assuming a known and fixed PSD during parameter estimation for CBCs. We compared two methods, a *sequential* method that first computes the PSD posterior and then uses the median to estimate the posterior for the CBC parameters, and a *marginalized* method that simultaneously infers the posterior for the PSD and the CBC parameters. We recovered similar parameter posteriors, in agreement with previous results [28,29]. In particular, the difference in the chirp mass posterior median (width) between the two methods is $0^{+1.4}_{-1.6}\%$ ($2^{+16}_{-13}\%$) across all events. We found no trends between these differences and the event chirp masses and network SNRs. The marginalized method involves simultaneous modeling of more parameters and thus comes with a greater computational cost than the sequential method. This increase is smaller for high-mass events, but can reach an increase of up to ~1.5 in computing time for low-mass events.

The simultaneous modeling of the CBC signal and the noise PSD allows us to marginalize over uncertainty in the latter. Compared to the method explore in [28], our method involves a single parameter estimation analysis and can account for potential correlations between the CBC signal and the PSD model parameters. Additionally, our method enables estimation of the Bayesian evidence for the marginalized PSD analysis and thus computation of the Bayes factor between different models. However, correctly interpreting the Bayes factor within the context of the chosen priors of the PSD model is a nontrivial task and lies outside the scope of this paper.

We explained our results within the context of a toy model that considers the impact of a systematic increase or decrease of the PSD. In the case of simple Gaussian distributions, the difference between the posterior from the median PSD and the one from the full PSD posterior is of higher order in the PSD uncertainty. Thus the *sequential* method that is based on the *median* PSD already results in posteriors that are very similar to the *marginalized* method.
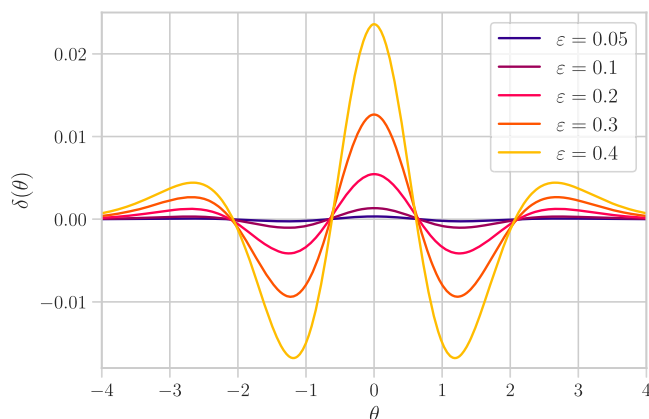

FIG. 8. Plot of $\delta(\theta)$, Eq. (4), the difference between $\mathcal{N}(\theta|0, 1)$ and $\int_{-\varepsilon}^{\varepsilon} \mathcal{N}(\theta|0, 1 + x)dx$ for various values of ε. A Taylor expansion shows that $\delta(\theta) \sim \varepsilon^2$ at $\theta = 0$.

We conclude that at current detector sensitivities, PSD uncertainty is subdominant compared to other sources of uncertainty, especially when the median on-source PSD is used.

ACKNOWLEDGMENTS

APPENDIX: OUTLIER FOLLOW-UP

We revisit the two events with the largest differences in parameter posteriors between the marginalized and sequential method, GW190512_180714 and GW190513_205428. In the analyses with the standard settings presented in Sec. III B, we found fractional differences in the chirp mass 90% posterior width of 38.1% and 24.9%, respectively.
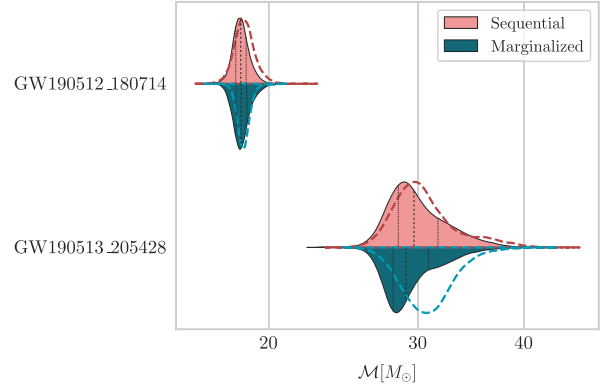

FIG. 9. Detector-frame chirp mass posterior comparison between the sequential (pink) and marginalized (teal) analyses for GW190512_180714 and GW190513_205428 using $f_{low} = 16$ Hz. These events have larger discrepancies between the posteriors when using $f_{low} = 20$ Hz, presented in dashed for comparison.

Among the various parameter settings, reducing the lower frequency cutoff from $f_{low} = 20$ Hz to $f_{low} = 16$ Hz significantly mitigates the observed differences. We compare the different $\mathcal{M}$ posteriors for GW190512_180714 and GW190513_205428 in Fig. 9. Using $f_{low} = 16$ Hz reduces the posterior width differences to -7.1% and 11.3%, respectively. Similarly, the JSDs are reduced from 0.052 and 0.033 to 0.0025 and 0.015 for each event, respectively. This follow-up analysis confirms that the differences between the posteriors of these events goes down with smaller lower frequency cutoffs, and suggests that analysis settings can sometimes significantly influence parameter estimation results.

[1] J. Aasi *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **32,** 074001 (2015).

[2] F. Acernese *et al.* (VIRGO Collaboration), Classical Quantum Gravity **32,** 024001 (2015).

[3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Classical Quantum Gravity **37,** 055002 (2020).

[4] J. D. Romano and N. J. Cornish, Living Rev. Relativity **20,** 2 (2017).

[5] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, Phys. Rev. D **103,** 044006 (2021).

[6] J. Veitch *et al.*, Phys. Rev. D **91,** 042003 (2015).

[7] G. Ashton *et al.*, Astrophys. J. Suppl. Ser. **241,** 27 (2019).

[8] I. M. Romero-Shaw *et al.*, Mon. Not. R. Astron. Soc. **499,** 3295 (2020).

[9] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D **104,** 063034 (2021).

[10] K. Chatziioannou, N. Cornish, M. Wijngaarden, and T. B. Littenberg, Phys. Rev. D **103,** 044013 (2021).

[11] S. Hourihane, K. Chatziioannou, M. Wijngaarden, D. Davis, T. Littenberg, and N. Cornish, Phys. Rev. D **106,** 042006 (2022).

[12] N. J. Cornish, Phys. Rev. D **102,** 124038 (2020).

[13] O. Edy, A. Lundgren, and L. K. Nuttall, Phys. Rev. D **103,** 124061 (2021).

[14] J. Aasi *et al.* (LIGO Scientific and VIRGO Collaborations), Phys. Rev. D **88,** 062001 (2013).

[15] T. B. Littenberg and N. J. Cornish, Phys. Rev. D **91,** 084034 (2015).

[16] N. J. Cornish and T. B. Littenberg, Classical Quantum Gravity **32,** 135012 (2015).

[17] P. J. Green, Biometrika **82,** 711 (1995).

[18] 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).

[19] B. Abbott, R. Abbott, and T. e. a. Abbott, Phys. Rev. X **9,** 031040 (2019).

[20] R. Abbott, T. D. Abbott, and S. e. a. Abraham, Phys. Rev. X **11,** 021053 (2021).

[21] R. Abbott *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), arXiv:2111.03606.

[22] C. Röver, R. Meyer, and N. Christensen, Classical Quantum Gravity **28,** 015010 (2011).

[23] C. Rover, Phys. Rev. D **84,** 122004 (2011).

[24] T. B. Littenberg, M. Coughlin, B. Farr, and W. M. Farr, Phys. Rev. D **88,** 084044 (2013).

[25] M. C. Edwards, R. Meyer, and N. Christensen, Phys. Rev. D **92,** 064011 (2015).

[26] S. Banagiri, M. W. Coughlin, J. Clark, P. D. Lasky, M. A. Bizouard, C. Talbot, E. Thrane, and V. Mandic, Mon. Not. R. Astron. Soc. **492,** 4945 (2020).

[27] R. Smith and E. Thrane, Phys. Rev. X **8,** 021019 (2018).

[28] S. Biscoveanu, C.-J. Haster, S. Vitale, and J. Davies, Phys. Rev. D **102,** 023008 (2020).

[29] C. Talbot and E. Thrane, Phys. Rev. Res. **2,** 043298 (2020).

[30] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. X **6,** 041015 (2016); **8,** 039903(E) (2018).

[31] K. Chatziioannou, Phys. Rev. D **105,** 084021 (2022).

[32] M. Wijngaarden, K. Chatziioannou, A. Bauswein, J. A. Clark, and N. J. Cornish, Phys. Rev. D **105,** 104019 (2022).

[33] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, J. Phys. Conf. Ser. **610,** 012021 (2015).

[34] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), SoftwareX **13,** 100658 (2021).

[35] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, Phys. Rev. D **93,** 044006 (2016).

[36] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Phys. Rev. D **93,** 044007 (2016).

[37] N. J. Cornish, arXiv:1007.4820.

[38] N. J. Cornish, Phys. Rev. D **104,** 104054 (2021).

[39] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **116,** 061102 (2016).

[40] K. K. Y. Ng, S. Vitale, A. Zimmerman, K. Chatziioannou, D. Gerosa, and C.-J. Haster, Phys. Rev. D **98,** 083007 (2018).

[41] J. Lin, IEEE Trans. Inf. Theory **37,** 145 (1991).

[42] https://www.gw-openscience.org.

[43] 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).

[44] J. D. Hunter, Comput. Sci. Eng. **9,** 90 (2007).

[45] D. Foreman-Mackey, J. Open Source Software **1,** 24 (2016).