Accelerating parameter estimation of gravitational waves from compact binary coalescence using adaptive frequency resolutions

Soichiro Morisaki◉

Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA

Bayesian parameter estimation of gravitational waves from compact binary coalescence (CBC) typically requires the generation of millions of computationally expensive template waveforms. We propose a technique to reduce the cost of waveform generation by exploiting the chirping behavior of CBC signal. Our technique does not require waveforms at all frequencies in the frequency range used in the analysis, and does not suffer from the fixed cost due to the upsampling of waveforms. Our technique speeds up the parameter estimation of typical binary neutron star signal by a factor of $\mathcal{O}(10)$ for the low-frequency cutoff of 20 Hz, and $\mathcal{O}(10^2)$ for 5 Hz. It does not require any offline preparations or accurate estimates of source parameters provided by detection pipelines.

I. INTRODUCTION

The discovery of gravitational waves from binary black hole coalescence opened a new window to the Universe [1]. In 2017, gravitational waves from binary neutron star coalescence were also detected [2]. The multimessenger observations of this event enabled us to measure the Hubble constant in a way independent from the cosmic ladder [3], learn the origin of heavy elements [4–9] and the structure of ultrarelativistic jet from the merger [10–14]. So far, several tens of compact binary coalescence (CBC) events have been detected [15,16] by the LIGO-Virgo collaboration [17,18], enabling us to learn the population properties of binary black holes [19,20].

In this era of gravitational-wave astronomy, an accurate inference of the source properties from gravitational-wave data is important. The LIGO-Virgo collaboration employs Bayesian inference with stochastic sampling [21,22], which generates thousands of random samples following the probability distribution of the source parameters conditioned on data. It typically requires generation of millions of computationally expensive template waveforms. The waveform generation becomes very costly for light binaries, whose signals have long durations and extend to high frequencies. For a typical binary neutron star signal, the parameter estimation can take a few weeks, or even years, without any approximate methods. To solve this issue, various techniques to speed up the inference have been proposed [23–36].

One of the techniques widely used in the detection and parameter estimation of CBC signal is multibanding [29,37–39]. It exploits the chirping behavior of CBC signal, whose frequency simply increases with time. In the time domain, it means the sampling frequency can be lowered at the early stage of inspiral, which significantly reduces the

number of time samples at which waveforms are evaluated. This idea has been utilized to speed up matched filtering of data in the detection of CBC signals [39].

On the other hand, the standard parameter estimation is performed in the frequency domain. The Fourier transform of a CBC waveform is an oscillating function of frequency, and the frequency scale of the oscillations is the inverse of time-to-merger. Since the time-to-merger decreases as the frequency increases, we can use coarser frequency resolutions at high frequencies for resolving that oscillatory behavior. The previous study [29] has proposed an efficient technique, which computes waveforms with coarser frequency resolutions at high frequencies and upsamples them to the original frequency resolution. This technique was named MB-Interpolation, and it significantly reduces the number of waveform evaluations at high frequencies.

However, the overall speed-up gain of MB-Interpolation is more modest than the reduction of the number of waveform evaluations. For example, it reduces the number of waveform evaluations by a factor of ∼60 for typical BNS signal and the low-frequency cutoff of 20 Hz, but the speed-up gain of parameter estimation with the TaylorF2 [40] waveform model is ∼3 (See Table 1 of [29]). It arises because MB-Interpolation requires the upsampling of waveforms and the computation of inner products of upsampled waveforms and data. Their costs are proportional to the original frequency samples, and are not reduced by multibanding.

In this paper, we propose another technique, which exploits the chirping behavior of CBC signal but does not require the upsampling. If the length of data is T, the inner products of waveforms and data require waveforms at frequency points with the frequency interval of $1/T$.

044062-1

On the other hand, if $f_{1/2}$ is the frequency from which the time-to-merger is $T/2$, the inner products in $f > f_{1/2}$ can be approximately computed with the latter half of data and waveforms at frequency points with the frequency interval of $2/T$. Generalizing this idea, we divide the total frequency range into multiple bands in a way that the time-to-merger of each band is much smaller than that of the previous band, and use coarser frequency resolutions in high-frequency bands. It significantly reduces the number of waveform evaluations at high frequencies, and does not require upsampled waveforms.

This paper is organized as follows. First, we formulate our technique and evaluate its speed-up gains in Sec. II. Next, we investigate the accuracy of our technique in Sec. III. Finally, we summarize the results and conclude this paper in Sec. IV.

II. PARAMETER ESTIMATION WITH ADAPTIVE FREQUENCY RESOLUTIONS

In this section, we formulate our technique and evaluate its speed-up gains. First, we review the basics of Bayesian parameter estimation of CBC signal in Sec. II A. Next, we formulate our technique in Sec. II B–II F. Finally, we evaluate the speed-up gains in Sec. II G.

A. Bayesian inference

In the Bayesian inference, the probability distribution of model parameters conditioned on data is calculated via the Bayes' theorem,

$$p(\boldsymbol{\theta}|\boldsymbol{d}) \propto \pi(\boldsymbol{\theta})\mathcal{L}(\boldsymbol{d}|\boldsymbol{\theta}), \qquad (1)$$

where $\boldsymbol{d} = (d_0, d_1, ..., d_{N-1})^T$ represents time-domain data at N time samples, whose sampling rate is $1/\Delta t$, and $\boldsymbol{\theta}$ represents the model parameters. $p(\boldsymbol{\theta}|\boldsymbol{d})$, $\pi(\boldsymbol{\theta})$ and $\mathcal{L}(\boldsymbol{d}|\boldsymbol{\theta})$ are referred to as posterior, prior and likelihood respectively. The prior is determined based on our prior knowledge or belief on $\boldsymbol{\theta}$.

In the standard parameter estimation of CBC signal, the noise is modeled as stationary Gaussian random process. In this model, the logarithm of likelihood for a single detector is given by [21]

$$\ln \mathcal{L}(\boldsymbol{d}|\boldsymbol{\theta}) = (\boldsymbol{d}, \boldsymbol{h}(\boldsymbol{\theta})) - \frac{1}{2}(\boldsymbol{h}(\boldsymbol{\theta}), \boldsymbol{h}(\boldsymbol{\theta})) + \text{const}, \qquad (2)$$

where $\boldsymbol{h}(\boldsymbol{\theta})$ is the CBC waveform for $\boldsymbol{\theta}$. The inner products are given by

$$(\boldsymbol{d}, \boldsymbol{h}) \equiv \frac{4}{T}\Re\left[\sum_{k=1}^{\lceil(N-1)/2\rceil} \frac{\tilde{d}_k^* \tilde{h}(f_k)}{S_k}\right], \qquad (3)$$

$$(\boldsymbol{h}, \boldsymbol{h}) \equiv \frac{4}{T}\sum_{k=1}^{\lfloor(N-1)/2\rfloor} \frac{|\tilde{h}(f_k)|^2}{S_k}, \qquad (4)$$

where T is the duration of data $T \equiv N\Delta t$, f_k is the frequency of the kth bin $f_k \equiv k/T$, $\lfloor x \rfloor$ is the greatest integer less than or equal to x, $\tilde{d}_k$ is the Fourier component of data defined by

$$\tilde{d}_k \equiv \Delta t \sum_{m=0}^{N-1} d_m e^{-2\pi ikm/N}, \qquad (5)$$

S_k is the one-sided power spectral density (PSD) of the detector's noise, and $\tilde{h}(f)$ is the template waveform in the frequency domain. The DC and Nyquist frequencies have been excluded, and $\boldsymbol{\theta}$ has been omitted for ease of notation. Typically, the low and high frequency cutoffs, f_{low} and f_{high}, have been determined before the analysis. Here we assume

$$S_k = +\infty, \quad (f_k < f_{\text{low}} \quad \text{or} \quad f_k > f_{\text{high}}) \qquad (6)$$

so that the components outside the frequency range are automatically vanishing. The likelihood for multiple detectors is the product of likelihood for each detector.

In the inference with stochastic sampling, the non-constant part of log-likelihood, which is often referred to as log-likelihood-ratio,

$$\ln \Lambda(\boldsymbol{d}|\boldsymbol{\theta}) \equiv (\boldsymbol{d}, \boldsymbol{h}(\boldsymbol{\theta})) - \frac{1}{2}(\boldsymbol{h}(\boldsymbol{\theta}), \boldsymbol{h}(\boldsymbol{\theta})), \qquad (7)$$

is computed tens to hundreds of millions of times [21,34]. It requires the evaluations of $\tilde{h}(f)$ at K_{orig} frequency points, where

$$K_{\text{orig}} \equiv \lfloor f_{\text{high}}T \rfloor - \lceil f_{\text{low}}T \rceil + 1 \sim (f_{\text{high}} - f_{\text{low}})T. \qquad (8)$$

Our technique is an approximate method to compute log-likelihood-ratio with fewer waveform evaluations.

B. Window functions

For dividing the total frequency range into multiple frequency bands, we introduce the following overlapping window functions,

$$w^{(b)}(f) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\pi \frac{f - f^{(b)}}{\Delta f^{(b)}}\right)\right), & (f^{(b)} - \Delta f^{(b)} < f < f^{(b)}) \\ 1, & (f^{(b)} \le f \le f^{(b+1)} - \Delta f^{(b+1)}) \\ \frac{1}{2}\left(1 - \cos\left(\pi \frac{f - f^{(b+1)}}{\Delta f^{(b+1)}}\right)\right), & (f^{(b+1)} - \Delta f^{(b+1)} < f < f^{(b+1)}) \\ 0, & (\text{otherwise}) \end{cases} \tag{9}$$

where

$$f_{\text{low}} = f^{(0)} < f^{(1)} < \ldots < f^{(B)} = f_{\text{high}} + \Delta f^{(B)},$$
$$f^{(b)} < f^{(b+1)} - \Delta f^{(b+1)}, \qquad \Delta f^{(0)} = 0, \tag{10}$$

and B represents the number of frequency bands. $\Delta f^{(B)}$ needs to be positive to smooth the high-frequency end of waveform. We explain its necessity and our choice of $\Delta f^{(B)}$ in Sec. II F. The window functions are constructed so that their sum becomes unity,

$$\sum_{b=0}^{B-1} w^{(b)}(f) = 1. \quad (f_{\text{low}} \le f \le f_{\text{high}}) \tag{11}$$

The reason for using smooth window functions rather than rectangular window functions is explained in the Appendix A.

C. How to compute (d, h)

First, we introduce an approximate method to compute (d, h). It can be rewritten as follows thanks to (11),

$$(d, h) = \sum_{b=0}^{B-1} \frac{4}{T} \Re\left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k) \frac{\tilde{d}_k^* \tilde{h}(f_k)}{S_k}\right], \tag{12}$$

where $N^{(b)}$ is an integer satisfying

$$\left\lfloor \frac{N^{(b)} - 1}{2} \right\rfloor \ge f^{(b+1)} T. \tag{13}$$

We use the minimum power of 2 satisfying this condition as $N^{(b)}$ for efficient Fourier transforms, which specifically speeds up the IFFT-FFT computations for (h, h) introduced in Sec. II D. If $N^{(b)} > N$, we pad zeros to data so that (12) is satisfied,

$$\frac{\tilde{d}_k}{S_k} = 0. \quad \left(k > \left\lfloor \frac{N-1}{2} \right\rfloor\right). \tag{14}$$

The inner product can be transformed into a sum over times as follows,

$$\frac{4}{T} \Re\left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k) \frac{\tilde{d}_k^* \tilde{h}(f_k)}{S_k}\right] = 2\Delta t^{(b)} \sum_{m=0}^{N^{(b)}-1} D_m^{(b)} h_m^{(b)}, \tag{15}$$

where $\Delta t^{(b)} \equiv N\Delta t / N^{(b)}$, and

$$D_m^{(b)} \equiv \frac{2}{T} \Re\left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} \frac{\tilde{d}_k}{S_k} e^{2\pi i k m / N^{(b)}}\right], \tag{16}$$

$$h_m^{(b)} \equiv \frac{2}{T} \Re\left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k) \tilde{h}(f_k) e^{2\pi i k m / N^{(b)}}\right]. \tag{17}$$

Since the frequency of the CBC waveform simply increases with time, $h_m^{(b)}$ is almost vanishing for $m\Delta t^{(b)} \lesssim T - \tau(f^{(b)} - \Delta f^{(b)})$, where $\tau(f)$ is the time to merger from a frequency f. If the waveform contains multiple gravitational-wave moments, $\tau(f)$ is defined on the moment with the maximum magnetic number, whose time to merger is the longest. Figure 1 shows $h_m^{(b)}$ for 1.4 $M_\odot$ − 1.4 $M_\odot$ BNS
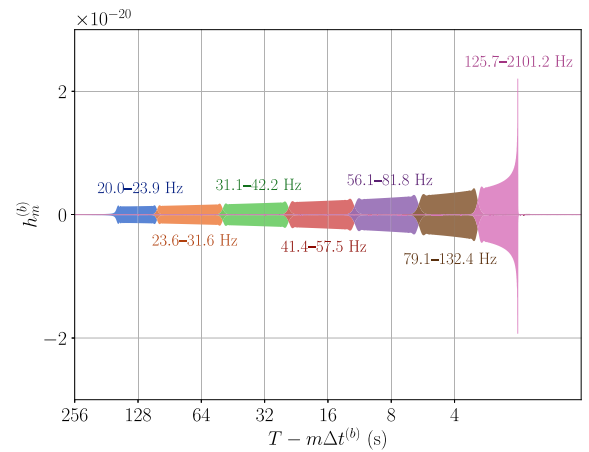

FIG. 1. The inverse Fourier transform of the waveform in each frequency band. The waveforms are for nonspinning 1.4 $M_\odot$ − 1.4 $M_\odot$ BNS signal, whose coalescence time is at $T - 2$ s. The total frequency range starting from 20 Hz is divided into 7 frequency bands, and each band is constructed so that the waveform in the bth band is vanishing except for in the last 2^{8-b} s. Each label presents the start and end frequencies of the band, $f^{(b)} - \Delta f^{(b)}$ and $f^{(b+1)}$.

with zero spins, which ends at $T - 2$ s. Each frequency band is constructed so that $\tau(f^{(b)} - \Delta f^{(b)}) < (2^{8-b} - 2)$ s. The figure shows that $h_m^{(b)}$ is almost vanishing at $m\Delta t^{(b)} < T - 2^{8-b}$ s, which validates ignoring $h_m^{(b)}$ there. It implies we can make the approximation that

$$h_m^{(b)} \simeq 0, \qquad (m = 0, 1, ..., N^{(b)} - M^{(b)} - 1) \qquad (18)$$

where $T^{(b)} \equiv M^{(b)}\Delta t^{(b)}$ is long enough compared to $\tau(f^{(b)} - \Delta f^{(b)})$, and

$$T^{(B-1)} < T^{(B-2)} < ... < T^{(0)} \leq T. \qquad (19)$$

This approximation leads to

$$\sum_{m=0}^{N^{(b)}-1} D_m^{(b)} h_m^{(b)} \simeq \sum_{m=N^{(b)}-M^{(b)}}^{N^{(b)}-1} D_m^{(b)} h_m^{(b)}. \qquad (20)$$

Finally, the inner product can be transformed into a sum over frequencies as follows,

$$2\Delta t^{(b)} \sum_{m=N^{(b)}-M^{(b)}}^{N^{(b)}-1} D_m^{(b)} h_m^{(b)}$$

$$\simeq \frac{4}{T^{(b)}} \Re \left[\sum_{k=K_s^{(b)}}^{K_e^{(b)}} w^{(b)}\left(f_k^{(b)}\right) \tilde{D}_k^{(b)*} \tilde{h}\left(f_k^{(b)}\right) \right], \qquad (21)$$

where $K_s^{(b)} = \lceil (f^{(b)} - \Delta f^{(b)})T^{(b)} \rceil$, $K_e^{(b)} = \lfloor f^{(b+1)}T^{(b)} \rfloor$, $f_k^{(b)} = k/T^{(b)}$ and

$$\tilde{D}_k^{(b)} = \Delta t^{(b)} \sum_{m=N^{(b)}-M^{(b)}}^{N^{(b)}-1} D_m e^{-2\pi ikm/M^{(b)}}. \qquad (22)$$

Here, we have made the following approximation,

$$\Delta t^{(b)} \sum_{m=N^{(b)}-M^{(b)}}^{N^{(b)}-1} h_m^{(b)} e^{-2\pi ikm/M^{(b)}} \simeq w^{(b)}\left(f_k^{(b)}\right) \tilde{h}\left(f_k^{(b)}\right). $$

$$(23)$$

Finally, the inner product is reduced to

$$(d, h) \simeq \sum_{b=0}^{B-1} \frac{4}{T^{(b)}} \Re \left[\sum_{k=K_s^{(b)}}^{K_e^{(b)}} w^{(b)}\left(f_k^{(b)}\right) \tilde{D}_k^{(b)*} \tilde{h}\left(f_k^{(b)}\right) \right]. \qquad (24)$$

$\tilde{D}_k^{(b)}$ can be computed from (16) and (22), and stored before the sampling. The frequency interval of the bth band is $1/T^{(b)}$, which is larger than the original frequency

interval of $1/T$ for $b \geq 1$. It means (24) requires fewer waveform evaluations for $b \geq 1$. The number of waveform evaluations is

$$K_{\mathrm{MB}} = \sum_{b=0}^{B-1} (K_e^{(b)} - K_s^{(b)} + 1), \qquad (25)$$

and its cost is reduced by a factor of $K_{\mathrm{orig}}/K_{\mathrm{MB}}$. Since the computation of (24) does not require the upsampling of $\tilde{h}(f_k^{(b)})$, the computation of (d, h) is sped up by the same factor.

D. How to compute (h, h)

Next, we introduce approximate methods to compute (h, h). It can be rewritten as follows thanks to (11),

$$(h, h) = \sum_{b=0}^{B-1} \frac{4}{T} \sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k) \frac{|\tilde{h}(f_k)|^2}{S_k}. \qquad (26)$$

As shown in the following, it can be approximately computed with waveforms at $f_k^{(b)}(b = 0, 1, ..., B - 1; k = K_s^{(b)}, K_s^{(b)} + 1, ..., K_e^{(b)})$, which have been computed for (d, h).

The waveform of CBC signal can be expressed as the linear combination of the -2 spin-weighted spherical harmonics $_{-2}Y_{lm}(\theta, \phi)$ [41]. The dominant moments are the quadrupole moments $(l, m) = (2, \pm 2)$, and moments with $|m| \geq 3$ are referred to as higher-order moments. Figure 2 shows $|\tilde{h}(f)|^2$ for waveform models containing
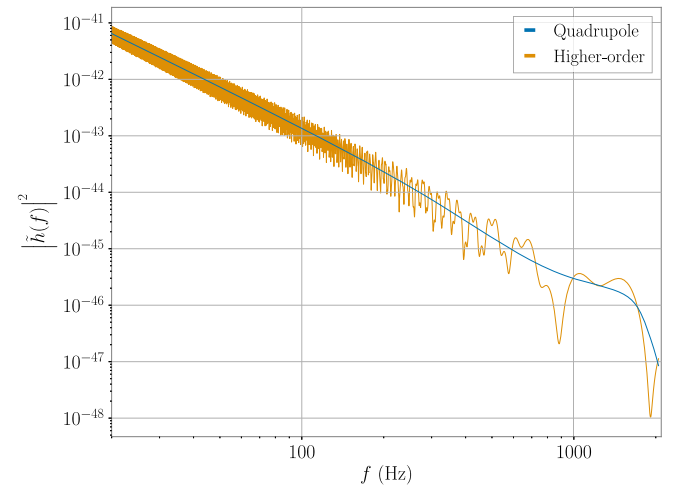

FIG. 2. $|\tilde{h}(f)|^2$ for waveform models containing only dominant quadrupole moments (blue) and containing higher-order moments (orange). They are for nonspinning $7\,M_\odot - 1.4\,M_\odot$ binary, whose inclination angle between the line of sight and the orbital angular momentum is $\pi/2$. The waveform models are IMRPhenomD [42] for the quadrupole case and IMRPhenomHM [43] for the higher-order case.

only dominant quadrupole moments and containing higher-order moments. In the former case, $|\tilde{h}(f)|^2$ is a smooth function as the phase is canceled out. In the latter case, the cross terms between different moments give rise to an oscillatory behavior. To compute $(\boldsymbol{h}, \boldsymbol{h})$ efficiently and accurately in each case, we propose two methods: *Linear interpolation* and *IFFT-FFT*. The former method is more efficient but may not be accurate for waveform models containing higher-order moments. The latter method is accurate for such waveform models but more costly.

1. Linear interpolation

In this method, $|\tilde{h}(f)|^2$ is approximated by the linear interpolation of $|\tilde{h}(f_k^{(b)})|^2$,

$$|\tilde{h}(f)|^2 \simeq T^{(b)}(f_{k+1}^{(b)} - f)|\tilde{h}(f_k^{(b)})|^2 + T^{(b)}(f - f_k^{(b)})|\tilde{h}(f_{k+1}^{(b)})|^2, \quad (27)$$

for $f_k^{(b)} \leq f < f_{k+1}^{(b)}$ when the inner product in the bth band is computed. The approximate forms in $f_{K_s^{(b)}}^{(b)} \leq f < f_{K_s^{(b)}+1}^{(b)}$ and $f_{K_e^{(b)}-1}^{(b)} \leq f < f_{K_e^{(b)}}^{(b)}$ are extrapolated to $f < f_{K_s^{(b)}}^{(b)}$ and $f \geq f_{K_e^{(b)}}^{(b)}$ respectively.

Substituting the linear interpolation into (26), we obtain

$$(\boldsymbol{h}, \boldsymbol{h}) = \sum_{b=0}^{B-1} \sum_{k=K_s^{(b)}}^{K_e^{(b)}} c_k^{(b)} \left|\tilde{h}\left(f_k^{(b)}\right)\right|^2. \quad (28)$$

The coefficients are given by

$$c_k^{(b)} = \frac{4T^{(b)}}{T} \sum_{\bar{f}_k^{(b)} \leq f_l < \bar{f}_{k+1}^{(b)}} \left(f_{k+1}^{(b)} - f_l\right) \frac{w^{(b)}(f_l)}{S(f_l)}$$
$$+ \frac{4T^{(b)}}{T} \sum_{\bar{f}_{k-1}^{(b)} \leq f_l < \bar{f}_k^{(b)}} \left(f_l - f_{k-1}^{(b)}\right) \frac{w^{(b)}(f_l)}{S(f_l)}, \quad (29)$$

where

$$\bar{f}_{K_s^{(b)}-1}^{(b)} = \bar{f}_{K_s^{(b)}}^{(b)} = 0, \quad \bar{f}_{K_e^{(b)}}^{(b)} = \bar{f}_{K_e^{(b)}+1}^{(b)} = \infty,$$
$$\bar{f}_k^{(b)} = f_k^{(b)}. (k = K_s^{(b)}+1, K_s^{(b)}+2, ..., K_e^{(b)}-1). \quad (30)$$

The computation of (28) only requires $\mathcal{O}(K_{\mathrm{MB}})$ floating-point operations, and its cost is negligible compared to that of the waveform evaluations. The interpolation may not be accurate if the waveform model takes into account higher-order moments due to the oscillatory behavior of $|\tilde{h}(f)|^2$.

2. IFFT-FFT

The approximation used in this method is similar to that used for computing $(\boldsymbol{d}, \boldsymbol{h})$. First, the inner product can be transformed into a sum over times as follows,

$$\frac{4}{T} \sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k) \frac{|\tilde{h}(f_k)|^2}{S_k} = 2\Delta t^{(b)} \sum_{m=0}^{N^{(b)}-1} I_m^{(b)} H_m^{(b)}, \quad (31)$$

where

$$I_m^{(b)} \equiv \frac{2}{T} \Re \left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} \frac{1}{S_k} e^{2\pi i k m/N^{(b)}} \right], \quad (32)$$

$$H_m^{(b)} \equiv \frac{2}{T} \Re \left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} w^{(b)}(f_k)|\tilde{h}(f_k)|^2 e^{2\pi i k m/N^{(b)}} \right]. \quad (33)$$

$H_m^{(b)}$ is the convolution of the windowed waveform,

$$H_m^{(b)} = \Delta t^{(b)} \sum_{m'=0}^{N^{(b)}-1} \hat{h}_{\mathrm{mod}(m+m',N^{(b)})}^{(b)} \hat{h}_{m'}^{(b)}, \quad (34)$$

where $\mathrm{mod}(a, b)$ is the remainder of a by b, and

$$\hat{h}_m^{(b)} \equiv \frac{2}{T} \times \Re \left[\sum_{k=1}^{\lfloor (N^{(b)}-1)/2 \rfloor} \sqrt{w^{(b)}(f_k)} \tilde{h}(f_k) e^{2\pi i k m/N^{(b)}} \right]. (35)$$

The same argument in Sec. II C leads to,

$$\hat{h}_m^{(b)} \simeq 0. \quad (m = 0, 1, ..., N^{(b)} - M^{(b)} - 1) \quad (36)$$

If $2M^{(b)} - 1 < N^{(b)}$, it means $H_m^{(b)}$ is vanishing for $M^{(b)} \leq m \leq N^{(b)} - M^{(b)}$. Thus, we can make the following approximation,

$$\sum_{m=0}^{N^{(b)}-1} I_m^{(b)} H_m^{(b)} \simeq \sum_{m=0}^{\hat{N}^{(b)}-1} I_{\mathrm{c},m}^{(b)} H_{\mathrm{c},m}^{(b)}, \quad (37)$$

where $\hat{N}^{(b)} \equiv \min[2M^{(b)}, N^{(b)}]$, and $I_{\mathrm{c},m}^{(b)}$ and $H_{\mathrm{c},m}^{(b)}$ are the cropped sequences with the sizes of $\hat{N}^{(b)}$,

$$I_{\mathrm{c},m}^{(b)} \equiv \begin{cases} I_m^{(b)}, & (m \leq \lfloor \hat{N}^{(b)}/2 \rfloor) \\ I_{m+N^{(b)}-\hat{N}^{(b)}}^{(b)}, & (m \geq \lfloor \hat{N}^{(b)}/2 \rfloor + 1) \end{cases} \quad (38)$$

$$H_{\mathrm{c},m}^{(b)} \equiv \begin{cases} H_m^{(b)}, & (m \leq \lfloor \hat{N}^{(b)}/2 \rfloor) \\ H_{m+N^{(b)}-\hat{N}^{(b)}}^{(b)}. & (m \geq \lfloor \hat{N}^{(b)}/2 \rfloor + 1) \end{cases} \quad (39)$$

$H_{c,m}^{(b)}$ can be expressed as the convolution of the cropped waveform,

$$H_{c,m}^{(b)} = \Delta t^{(b)} \sum_{m'=0}^{\hat{N}^{(b)}-1} \hat{h}_{c,\mathrm{mod}(m+m',\hat{N}^{(b)})}^{(b)} \hat{h}_{c,m'}^{(b)}, \qquad (40)$$

where

$$\hat{h}_{c,m}^{(b)} \equiv \hat{h}_{m+N^{(b)}-\hat{N}^{(b)}}^{(b)}. \quad (m = 0, 1, ..., \hat{N}^{(b)}-1) \qquad (41)$$

Using the properties of the Fourier transformation, we obtain

$$2\Delta t^{(b)} \sum_{m=0}^{\hat{N}^{(b)}-1} I_{c,m}^{(b)} H_{c,m}^{(b)} = \frac{4}{\hat{T}^{(b)}} \sum_{k=1}^{\lfloor (\hat{N}^{(b)}-1)/2 \rfloor} \tilde{I}_{c,k}^{(b)} |\tilde{h}_{c,k}^{(b)}|^2, \quad (42)$$

where $\hat{T}^{(b)} \equiv \hat{N}^{(b)} \Delta t^{(b)}$ and

$$\tilde{I}_{c,k}^{(b)} = \Delta t^{(b)} \sum_{m=0}^{\hat{N}^{(b)}-1} I_{c,m}^{(b)} e^{-2\pi i k m / \hat{N}^{(b)}}, \qquad (43)$$

$$\tilde{h}_{c,k}^{(b)} = \Delta t^{(b)} \sum_{m=0}^{\hat{N}^{(b)}-1} h_{c,m}^{(b)} e^{-2\pi i k m / \hat{N}^{(b)}}. \qquad (44)$$

Substituting (31), (37) and (42) into (26) leads to

$$(\boldsymbol{h}, \boldsymbol{h}) \simeq \sum_{b=0}^{B-1} \frac{4}{\hat{T}^{(b)}} \sum_{k=1}^{\lfloor (\hat{N}^{(b)}-1)/2 \rfloor} \tilde{I}_{c,k}^{(b)} |\tilde{h}_{c,k}^{(b)}|^2. \qquad (45)$$

$\tilde{I}_{c,k}^{(b)}$ can be computed from (32), (38) and (43), and stored before the sampling. $\tilde{h}_{c,k}^{(b)}$ can be approximately computed as follows. First, the last $M^{(b)}$ components of $\hat{h}_{c,m}^{(b)}$ are computed as the inverse Fourier transform of the windowed waveform,

$$\hat{h}_{c,m}^{(b)} \simeq \frac{2}{T^{(b)}} \Re \left[\sum_{k=1}^{\lfloor (M^{(b)}-1)/2 \rfloor} \sqrt{w^{(b)}(f_k^{(b)})} \tilde{h}(f_k^{(b)}) e^{2\pi i k m / M^{(b)}} \right],$$
$$(m = \hat{N}^{(b)} - M^{(b)}, \hat{N}^{(b)} - M^{(b)} + 1, ..., \hat{N}^{(b)} - 1)$$
$$(46)$$

and the first $\hat{N}^{(b)} - M^{(b)}$ components of $\hat{h}_{c,m}^{(b)}$ are set to be zeros. Then, $\tilde{h}_{c,k}^{(b)}$ can be computed as the Fourier transform of $\hat{h}_{c,m}^{(b)}$. Thus, the computation of $\tilde{h}_{c,k}^{(b)}$ requires an inverse fast Fourier transform (IFFT) and a fast Fourier transform (FFT), and this method is more costly than the linear interpolation method. This IFFT-FFT operation requires $\mathcal{O}(\hat{N}^{(b)} \log_2 \hat{N}^{(b)})$ floating-point operations for each band.

Since $\hat{N}^{(b)} \leq 2M^{(b)} \sim 4f^{(b+1)}T^{(b)}$, $2f^{(b+1)} \ll 1/\Delta t$ for small b, and $T^{(b)} \ll T$ for large b, we have $\hat{N}^{(b)} \ll N$. Thus, unless we have a lot of redundant bands with similar values of $T^{(b)}$, the IFFT-FFT operations do not cause a fixed cost of $\mathcal{O}(N)$.

E. How to determine the frequency bands $\{f^{(b)}\}_{b=0}^{B}$

We assume $\{T^{(b)}\}_{b=0}^{B-1}$ are specified by the user. Then, the frequency bands $\{f^{(b)}\}_{b=0}^{B}$ should be determined so that the windowed waveform in the bth band is vanishing at $t < T - T^{(b)}$, where the start time of data is $t = 0$.

The asymptotic behavior of the windowed waveform is studied in Appendix A. If we choose the following value as $\Delta f^{(b)}$,

$$\Delta f^{(b)} = \frac{1}{\sqrt{-\tau'(f^{(b)})}}, \qquad (47)$$

the windowed waveforms behave as follows,

$$h_m^{(b)} \propto \left(\frac{\sqrt{-\tau'(f^{(b)})}}{t_c - \tau(f^{(b)}) - m\Delta t^{(b)}} \right)^3, \qquad (48)$$

$$\hat{h}_m^{(b)} \propto \left(\frac{\sqrt{-\tau'(f^{(b)})}}{t_c - \tau(f^{(b)}) - m\Delta t^{(b)}} \right)^2, \qquad (49)$$

where t_c is the time at which the coalescence part of signal arrives at the detector, according to (A16) and (A20). Thus, $f^{(b)}$ should satisfy

$$\frac{\sqrt{-\tau'(f^{(b)})}}{t_c - \tau(f^{(b)}) - T + T^{(b)}} \ll 1. \qquad (50)$$

For this condition to be satisfied, $f^{(b)}$ is determined by the following equation,

$$\tau(f^{(b)}) + L\sqrt{-\tau'(f^{(b)})} = T^{(b)} + t_{c,\mathrm{min}} - T, \qquad (51)$$

where L is a user-specified constant satisfying $L \gg 1$, and $t_{c,\mathrm{min}}$ is the minimum of t_c. L refers to the duration of the tail part of the windowed waveform taken in the segment $T^{(b)}$ per $\sqrt{-\tau'(f^{(b)})}$. Larger L increases the accuracy of the approximation, as the power of neglected part of waveform becomes less. As shown in Sec. III, $L = 5$ is large enough for signals with signal-to-noise ratios of ~ 25. Given $f^{(b)}$ satisfying (51), $\Delta f^{(b)}$ is determined by (47).

$t_{c,\mathrm{min}}$ is determined by the prior range of $t_\oplus$, which is the time at which the coalescence part of signal arrives at the

geocenter, and the light-traveling time from the geocenter to the detector. We use the following conservative estimate,

$$t_{c,min} = t_{\oplus,min} - \frac{R_\oplus}{c}, \tag{52}$$

where $t_{\oplus,min}$ is the minimum of $t_\oplus$ in the prior range, $R_\oplus$ is the radius of the Earth, and c is the light speed. In the standard parameter estimation performed by the LIGO-Virgo collaboration, $t_{\oplus,min} = T - 2.1$ s [21], and hence $t_{c,min} - T = -2.12$ s. We use that standard prior range of $t_{\oplus,min}$ and that value of $t_{c,min}$ throughout this paper unless specified otherwise.

For $\tau(f)$, we use the following leading-order expression in the post-Newtonian (PN) expansion [44],

$$\tau_{0PN}(f) = \frac{5}{256} \frac{G\mathcal{M}}{c^3} \left(\frac{\pi G \mathcal{M} f}{c^3} \right)^{-\frac{8}{3}}, \tag{53}$$

if the waveform contains only dominant quadrupole moments. G is the gravitational constant, and $\mathcal{M}$ is so-called chirp mass defined by

$$\mathcal{M} = \frac{(m_1 m_2)^{\frac{3}{5}}}{(m_1 + m_2)^{\frac{1}{5}}}, \tag{54}$$

where m_1 and m_2 are the masses of colliding objects. We compute $\tau(f)$ with the minimum of chirp mass in the prior range, which gives the most conservative estimates. If the waveform model takes into account higher-order moments, we use
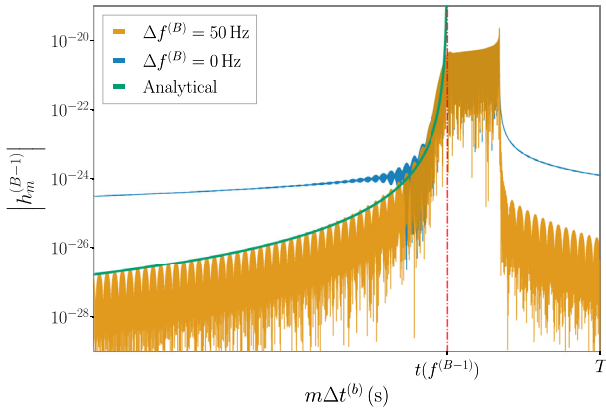

FIG. 3. The inverse Fourier transforms of the waveforms in the $(B-1)$th frequency band for $\Delta f^{(B)} = 0$ Hz and $\Delta f^{(B)} = 50$ Hz. The waveforms are for nonspinning $1.4\ M_\odot - 1.4\ M_\odot$ BNS signal, whose coalescence time is at $T - 2$ s. The high-frequency cutoff is $f^{(B)} - \Delta f^{(B)} = 2048$ Hz. The red dashed-dotted line represents the time at which the frequency of the waveform is $f^{(B-1)}$, where $f^{(B-1)} = 125.7$ Hz for this plot. The green line represents the analytical prediction given by the right-hand side of (A16).

$$\tau(f) = \tau_{0PN}\left(\frac{2}{m_{max}} f \right), \tag{55}$$

where m_{max} is the maximum of the magnetic numbers of moments for conservative estimates.

F. Choice of $\Delta f^{(B)}$

$\Delta f^{(B)}$ needs to be positive to smooth the high-frequency end of waveform. Figure 3 shows the absolute value of $h_m^{(B-1)}$ for $\Delta f^{(B)} = 0$ Hz and $\Delta f^{(B)} = 50$ Hz, in comparison with the analytical prediction given by the right-hand side of (A16). For $\Delta f^{(B)} = 0$ Hz, $h_m^{(B-1)}$ has a long tail, which does not decay following the analytical prediction, due to the abrupt cutoff at the high-frequency end. It significantly degrades the accuracy of our approximation. For $\Delta f^{(B)} = 50$ Hz, $h_m^{(B-1)}$ quickly decays following the analytical prediction.

Assuming that the smoothed waveform decays with the timescale of $1/\Delta f^{(B)}$, we use the following value,

$$\Delta f^{(B)} = \frac{100}{T - t_{c,max}}, \tag{56}$$

where $t_{c,max}$ is the maximum of t_c, so that $h_m^{(b)}$ rapidly decays in the last $T - t_{c,max}$ of data. Following the same argument for deriving (52), we use the following conservative value of $t_{c,max}$,

$$t_{c,max} = t_{\oplus,max} + \frac{R_\oplus}{c}, \tag{57}$$

where $t_{\oplus,max}$ is the maximum of $t_\oplus$ in the prior range. In the standard parameter estimation performed by the LIGO-Virgo collaboration, $t_{\oplus,max} = T - 1.9$ s [21], which leads to $\Delta f^{(B)} \simeq 53$ Hz. We use that standard value throughout this paper unless specified otherwise.

G. Speed-up gains

Finally, we evaluate the speed-up gains of our technique. We implemented our technique based on the likelihood class of BILBY [22,45], and measured the speed-up gains in evaluations of log-likelihood-ratio. Since the run time of parameter estimation is approximately the product of the evaluation time of log-likelihood-ratio and the number of their evaluations, it approximates the overall speed-up gain in parameter estimation.

Table I shows K_{orig}/K_{MB} and speed-up gains for nonspinning $1.4\ M_\odot - 1.4\ M_\odot$ BNS. The table lists their values for various values of f_{low}, and T is chosen as the minimum power of 2 larger than the time-to-merger of dominant quadrupole moments from f_{low}. It effectively sets higher low-frequency cutoffs on higher-order moments. The high-frequency cutoff is 2048 Hz. The total frequency range is

TABLE I. $K_{\mathrm{orig}}/K_{\mathrm{MB}}$ and speed-up gains in evaluations of log-likelihood-ratio for nonspinning 1.4 $M_\odot$–1.4 $M_\odot$ BNS. The table lists their values for various values of low-frequency cutoffs f_{low} and durations T, and the IMRPhenomD and IMRPhenomHM waveform models. The high-frequency cutoff is 2048 Hz. The speed-up gains were measured on 8-core Intel Core i9 with the clock rate of 2.4 GHz.

f_{low}(Hz)	T(s)	K_{orig}	IMRPhenomD		IMRPhenomHM	
			$K_{\mathrm{orig}}/K_{\mathrm{MB}}$	Speed up	$K_{\mathrm{orig}}/K_{\mathrm{MB}}$	Speed up
20	256	5.2×10^5	4.5×10	5.1×10	2.7×10	2.1×10
10	1024	2.1×10^6	1.2×10^2	1.5×10^2	5.7×10	4.6×10
5	8192	1.7×10^7	4.4×10^2	4.9×10^2	1.6×10^2	1.2×10^2

divided into frequency bands determined by (51) with $\{T^{(b)}\}_{b=0}^{B-1} = \{T, T/2, T/4, \ldots, 4\text{ s}\}$ and $L = 5$. IMRPhenomD [42] and IMRPhenomHM [43] are chosen as representative waveform models. IMRPhenomHM includes the effects of higher-order multiple moments, and $m_{\mathrm{max}} = 4$ is used to compute $\tau(f)$ when the frequency bands are calculated. For the computation of $(\boldsymbol{h}, \boldsymbol{h})$, the linear-interpolation method and the IFFT-FFT method were used for IMRPhenomD and IMRPhenomHM respectively.

The speed-up gains are roughly equal to $K_{\mathrm{orig}}/K_{\mathrm{MB}}$, which means our method does not suffer from the fixed cost present in MB-Interpolation [29]. The speed-up gains are smaller for IMRPhenomHM, as higher-order multiple moments have longer time-to-merger from a given frequency, and (51) gives more severe constraints. For $f_{\mathrm{low}} = 20$ Hz, which is used in most of the analyses by the LIGO-Virgo collaboration [15,16], the speed-up gains are $\mathcal{O}(10)$ in both cases. For $f_{\mathrm{low}} = 5$ Hz, which can be used for the third-generation detectors with improved sensitivities at low frequencies [46,47], the speed-up gains are $\mathcal{O}(10^2)$.

III. VALIDATION

In the previous section, we have formulated our technique, and shown that it significantly speeds up the parameter estimation. In this section, we investigate the accuracy of our technique.

A. Likelihood errors for GW190814

First, we investigate the errors of log-likelihood-ratio $\ln \Lambda$ from our approximation for GW190814 [48], gravitational-wave signal detected by the LIGO-Virgo collaboration. We computed $\ln \Lambda$ with and without our approximation on the posterior samples from the parameter estimation of this signal, and took their differences $\Delta \ln \Lambda$ as the errors of our approximation. This signal has a relatively large signal-to-noise ratio (SNR) of ~ 25, and it is appropriate for our study as systematic errors become prominent for a large SNR. This signal contains higher-order multipole moments at high confidence, which enables us to study the accuracy of our technique in their presence. The data, PSD, and posterior samples were obtained from the Gravitational Wave Open Science Center [49].

We computed the errors for two different waveform models, IMRPhenomD and IMRPhenomPv3HM [50,51]. For each waveform model, we used the posterior samples from the analysis using the same waveform model. The results are shown in Fig. 4. We used 16 s of data around the time of detection. Following [48], we analyzed the frequency range of 20–1024 Hz for LIGO-Hanford and Virgo, and 30–1024 Hz for LIGO-Livingston. The total frequency range is divided into 3 bands determined by (51) with $\{T^{(b)}\}_{b=0}^2 = \{16\text{ s}, 8\text{ s}, 4\text{ s}\}$, and $L = 5$ or $L = 50$. For computing $\tau(f)$, we used the reference chirp mass of 6.4 $M_\odot$, which is the median of the inferred detector-frame chirp mass, and $m_{\mathrm{max}} = 4$ for IMRPhenomPv3HM. We ignored the calibration errors of detectors as their effects are expected to be negligible for the SNR of this signal [52].

For IMRPhenomD, we used the linear-interpolation method to compute $(\boldsymbol{h}, \boldsymbol{h})$. The median error is 4×10^{-3} for $L = 5$ and 2×10^{-4} for $L = 50$, which shows increasing L improves the accuracy. The number of waveform evaluations is reduced by a factor of 3.6 for $L = 5$ and 3.2 for $L = 50$. For IMRPhenomPv3HM, we primarily used the IFFT-FFT method to compute $(\boldsymbol{h}, \boldsymbol{h})$. The median error is 2×10^{-4} for $L = 5$ and 5×10^{-5} for $L = 50$, which again shows increasing L improves the accuracy. The number of waveform evaluations is reduced by a factor of 3.3 for $L = 5$ and 2.8 for $L = 50$. For IMRPhenomPv3HM, the figure also shows the errors from the linear-interpolation method and $L = 5$. These errors are larger than those with the IFFT-FFT method, but they are still well below unity. In any case, the systematic errors due to our approximation are well below unity and the statistical errors.

B. Consistency of parameter estimation

To investigate the consistency of our technique, we performed parameter estimation of hundreds of simulated CBC signals using our technique. For each signal, we constructed the credible interval of each parameter centered on its median, and computed the credible level at which its true value is found. For the inference to be consistent, the credible levels should be uniformly distributed from 0 to 1 [53,54].
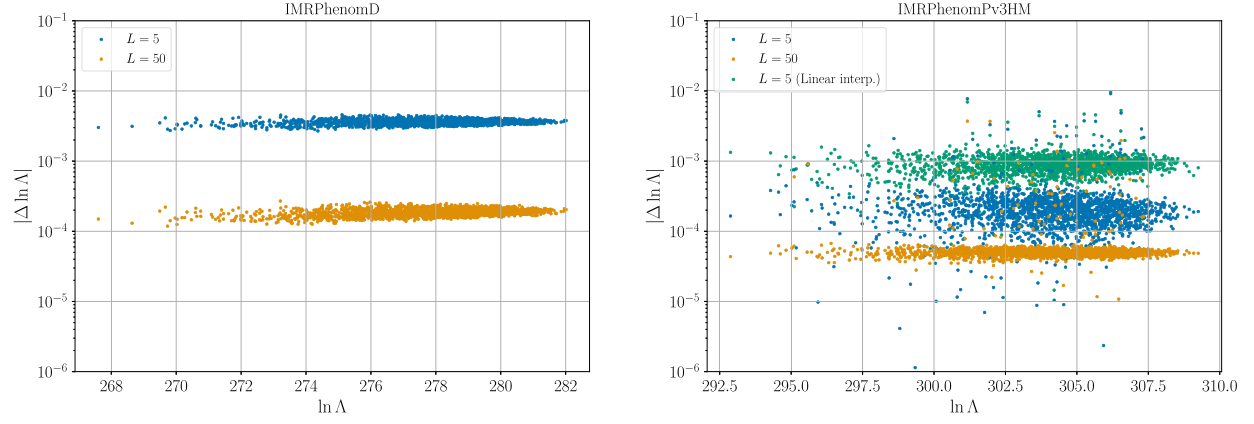
FIG. 4.   The errors of log-likelihood-ratio, $\ln \Lambda$, from our approximation for GW190814. The left figure shows the errors for the IMRPhenomD waveform model, and the right figure shows those for IMRPhenomHM. In each figure, the errors for $L = 5$ and $L = 50$ are shown in blue and orange respectively. For the computation of $(\boldsymbol{h}, \boldsymbol{h})$, the linear-interpolation method and the IFFT-FFT method are used for IMRPhenomD and IMRPhenomHM respectively by default. For IMRPhenomHM, the errors from the linear-interpolation method and $L = 5$ are also shown in green. For visibility, the data points are downsampled to 2000.

We considered the network of the two LIGO detectors and the Virgo detector, and signals were injected into Gaussian noise colored by their design sensitivities. We simulated 256 nonspinning BNS signals, whose chirp masses and mass ratios $q \equiv m_2/m_1$ are distributed uniformly within

$$1.15 \ M_\odot \leq \mathcal{M} \leq 1.25 \ M_\odot, \qquad 0.2 \leq q \leq 1. \qquad (58)$$

The luminosity distance D_{L} ranges from 10 Mpc to 100 Mpc, and its distribution is proportional to D_{L}^2. The
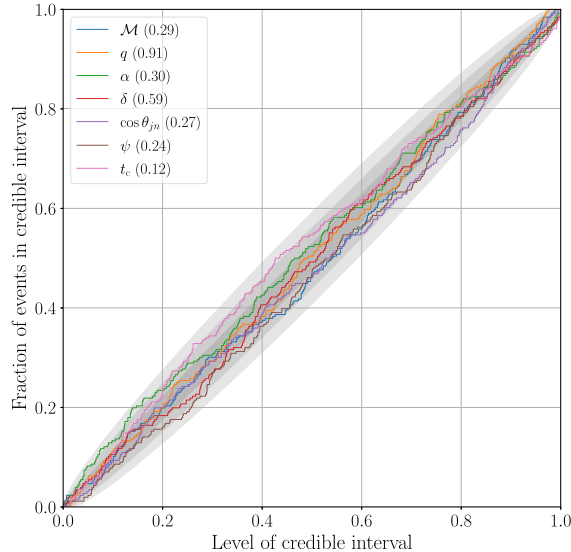
FIG. 5.   The cumulative distribution of credible levels for each source parameter, obtained from 256 simulated CBC signals. The gray regions represent the $1 - \sigma$, $2 - \sigma$ and $3 - \sigma$ confidence intervals of statistical errors due to a finite number of samples. Each label shows the p-value of the Kolmogorov-Smirnov test between the credible levels and a uniform distribution.

locations of the sources and the directions of the orbital angular momenta are isotropically distributed. The waveform model of simulated signals is IMRPhenomD, and the same waveform model was used for parameter estimation. The median network SNR of simulated signals is 24.3.

For parameter estimation, we used BILBY as an interface between likelihood and sampler, and DYNESTY [55] as sampler. The prior is the same as the distribution of simulated signals. The coalescence phase was analytically marginalized over and the luminosity distance was marginalized over with the look-up table method [56,57]. The total frequency range is 20–2048 Hz, and it is divided into 7 bands determined by (51) with $\{T^{(b)}\}_{b=0}^6 = \{256 \ \mathrm{s},$ $128 \ \mathrm{s}, \ldots, 4 \ \mathrm{s}\}$, $L = 5$ and the reference chirp mass of $1.15 \ M_\odot$. The number of waveform evaluations is reduced by a factor of 44.

Figure 5 shows the cumulative distribution of credible levels for each parameter. If credible levels are uniformly distributed, they should be diagonal lines for an infinite number of samples. The gray regions represent the $1 - \sigma$, $2 - \sigma$, and $3 - \sigma$ confidence intervals of statistical errors due to a finite number of samples, and the distributions are inside the $3 - \sigma$ interval for most of the range. The figure also presents the p-values of Kolmogorov-Smirnov tests between the credible levels and a uniform distribution in the legend. The moderate p-values indicate that the credible levels are consistent with uniformly distributed random numbers.

IV. CONCLUSION

In this paper, we have presented a technique to speed up the parameter estimation of gravitational waves from compact binary coalescence (CBC), which exploits the chirping behavior of CBC signal. It does not require the upsampling of waveforms, which is required by the

MB-Interpolation technique, another implementation of this idea proposed by [29]. Thus, our technique does not suffer from the fixed cost due to it, and the speed-up gains are larger. In Sec. II G, we have found that our technique speeds up the parameter estimation of 1.4 $M_\odot$–1.4 $M_\odot$ binary neutron star (BNS) signal by a factor of $\mathcal{O}(10)$ for the low-frequency cutoff of 20 Hz, which is the standard low-frequency cutoff used by the LIGO-Virgo analyses. The speed-up gain is increased to $\mathcal{O}(10^2)$ for the low-frequency cutoff of 5 Hz, which can be used for the third-generation detectors. We have investigated the errors of log-likelihood-ratio from our approximation and the consistency of the inference using our technique in Sec. III. The results indicate our technique is accurate enough to be used for signals, which have relatively large signal-to-noise ratios (SNRs) of ∼25. The errors of log-likelihood-ratio imply our technique is applicable to signals with higher SNRs. Investigating the limitation on the accuracy of our technique is the future work.

We note that there are various other techniques proposed to reduce the cost of waveform evaluations in parameter estimation. The reduced order quadrature (ROQ) technique [23–25] approximates waveforms by the linear combinations of basis vectors, and significantly reduces the number of frequency samples where waveforms are evaluated. The speed-up gain for BNS signal with the low-frequency cutoff of 20 Hz is $\mathcal{O}(10^2)$ [23,24], and it is increased to $\mathcal{O}(10^4)$ if basis vectors are constructed in narrow parameter space [25]. The heterodyned likelihood [32,33] and relative binning [30] methods assume waveforms sampled over in parameter estimation are very similar to the template waveform triggering the detection. The speed-up gain of relative binning is $\mathcal{O}(10^4)$ for GW170817 [30], BNS signal detected by the LIGO-Virgo collaboration. Compared to those speed-up gains, the speed-up gain of our technique is more modest. On the other hand, ROQ requires offline basis construction, which needs to be done for each waveform model we are interested in. The heterodyned likelihood and relative binning methods require a reference waveform, which is very similar to the true waveform. Since our technique does not require any offline preparations or reference waveforms, it is more easy-to-use than the other techniques. We also note that our technique can be used to reduce the file size of ROQ basis vectors and speed up the precomputations of ROQ, which is explained in Appendix B.

APPENDIX A: CHOICE OF WINDOW FUNCTION

Our technique is based on the approximation that the windowed waveform in the bth band is vanishing at $t \lesssim T - \tau(f^{(b)} - \Delta f^{(b)})$. We investigate the asymptotic behavior of the windowed waveform at $t < T - \tau(f^{(b)} - \Delta f^{(b)})$ for various window functions.

1. Stationary phase approximation

Each moment of a CBC waveform can be modeled as follows,

$$h(t) = A(t) \cos \Phi(t), \tag{A1}$$

where $\Phi(t)$ is defined so that $\Phi' > 0$, $\Phi'' > 0$. In the inspiral regime, the amplitude changes more slowly than the phase,

$$\left| \frac{A'}{A} \right| \ll \Phi', \qquad \Phi'' \ll (\Phi')^2. \tag{A2}$$

In this regime, we can apply the stationary phase approximation to calculate the Fourier transform of $h(t)$ [58],

$$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} dt \tag{A3}$$

$$\simeq \begin{cases} B(f) e^{-i\Psi(f)}, & (f > 0) \\ B(f) e^{i\Psi(f)}, & (f < 0) \end{cases} \tag{A4}$$

where

$$B(f) = \frac{\sqrt{t'(f)}}{2} A(t(f)), \tag{A5}$$

$$\Psi(f) = -\Phi(t(f)) + 2\pi f t(f) - \frac{\pi}{4}, \qquad \text{(A6)}$$

and $t(f)$ is the time at which $\Phi'(t) = 2\pi f$.

2. Rectangular window

First, we consider the waveform windowed by a rectangular window,

$$h^{(b)}(t) = 2\Re\left[\int_{f^{(b)}}^{f^{(b+1)}} df B(f) e^{2\pi i f t - i\Psi(f)}\right]. \qquad \text{(A7)}$$

For $t < t(f^{(b)})$, most of the contributions to the integral come from around $f = f^{(b)}$. Thus, we expand the integrand around $f = f^{(b)}$,

$$B(f) \simeq B(f^{(b)}), \qquad \text{(A8)}$$

$$\Psi(f) \simeq \Psi(f^{(b)}) + 2\pi t(f^{(b)})(f - f^{(b)}) \\ + \pi t'(f^{(b)})(f - f^{(b)})^2, \qquad \text{(A9)}$$

and approximately evaluate the integral as follows,

$$h^{(b)}(t) \simeq 2\Re\left[B(f^{(b)}) e^{2\pi i f^{(b)} t - i\Psi(f^{(b)})}\right. \\ \left. \times \int_{f^{(b)}}^{\infty} df e^{-2\pi i (f - f^{(b)})(t(f^{(b)}) - t) - i\pi t'(f^{(b)})(f - f^{(b)})^2}\right] \\ \simeq A(t(f^{(b)}))\Re\left[e^{2\pi i f^{(b)} t - i\Psi(f^{(b)})}\right. \\ \left. \times \left(-\frac{i}{2\pi}\frac{\sqrt{t'(f^{(b)})}}{t(f^{(b)}) - t} + \frac{1}{4\pi^2}\left(\frac{\sqrt{t'(f^{(b)})}}{t(f^{(b)}) - t}\right)^3\right)\right],$$

$$\text{(A10)}$$

where we have used

$$\int_0^{\infty} e^{-ipx - iqx^2} dx = -\frac{i}{p} + \frac{2q}{p^3} + \mathcal{O}\left(\frac{1}{p^5}\right), \qquad \text{(A11)}$$

for $p > 0$ and $q > 0$. (A10) means $h^{(b)}(t)$ has a long tail inversely proportional to $t(f^{(b)}) - t$, which degrades the accuracy of our technique.

3. Smooth window

Next, we consider the smooth window given by (9),

$$h^{(b)}(t) = 2\Re\left[\int_{f^{(b)} - \Delta f^{(b)}}^{f^{(b+1)}} df w^{(b)}(f) B(f) e^{2\pi i f t - i\Psi(f)}\right].$$

$$\text{(A12)}$$

Using the approximations, (A8) and (A9), we can approximately evaluate the integral from $f^{(b)} - \Delta f^{(b)}$ to $f^{(b)}$,

$$\int_{f^{(b)} - \Delta f^{(b)}}^{f^{(b)}} df w^{(b)}(f) B(f) e^{2\pi i f t - i\Psi(f)} \\ \simeq A(t(f^{(b)})) e^{2\pi i f^{(b)} t - i\Psi(f^{(b)})} \\ \times \left[\frac{i}{4\pi}\frac{\sqrt{t'(f^{(b)})}}{t(f^{(b)}) - t} - \frac{1}{8\pi^2}\left(\frac{\sqrt{t'(f^{(b)})}}{t(f^{(b)}) - t}\right)^3\right. \\ + \frac{i}{32\pi}\frac{\sqrt{t'(f^{(b)})}}{(\Delta f^{(b)})^2(t(f^{(b)}) - t)^3} \\ \left. \times \left(1 + e^{2\pi i \Delta f^{(b)}(t(f^{(b)}) - t) - \pi i t'(f^{(b)})(\Delta f^{(b)})^2}\right)\right], \qquad \text{(A13)}$$

where we have used

$$\int_{-1}^0 (1 + \cos(\pi x)) e^{-ipx - iqx^2} \\ = \frac{2i}{p} - \frac{4q - i\pi^2(1 + e^{i(p-q)})}{p^3} + \mathcal{O}\left(\frac{1}{p^4}\right). \qquad \text{(A14)}$$

The integral from $f^{(b)}$ to $f^{(b+1)}$ is the same as (A10), and the windowed waveform is given by

$$h^{(b)}(t) \simeq A(t(f^{(b)})) \\ \times \Re\left[\frac{i}{16\pi}\frac{\sqrt{t'(f^{(b)})}}{(\Delta f^{(b)})^2(t(f^{(b)}) - t)^3} e^{2\pi i f^{(b)} t - i\Psi(f^{(b)})}\right. \\ \left. \times \left(1 + e^{2\pi i \Delta f^{(b)}(t(f^{(b)}) - t) - \pi i t'(f^{(b)})(\Delta f^{(b)})^2}\right)\right]. \qquad \text{(A15)}$$

Its amplitude is quickly attenuated in proportion to $(t(f^{(b)}) - t)^{-3}$,

$$|h^{(b)}(t)| \le \frac{A(t(f^{(b)}))}{8\pi}\frac{\sqrt{t'(f^{(b)})}}{(\Delta f^{(b)})^2(t(f^{(b)}) - t)^3}, \qquad \text{(A16)}$$

and this smooth window is more appropriate than the rectangular window to be used for our technique.

4. Square-root of smooth window

Finally, we consider the square-root of the smooth window,

$$h^{(b)}(t) = 2\Re \left[\int_{f^{(b)} - \Delta f^{(b)}}^{f^{(b+1)}} df \sqrt{w^{(b)}(f)} B(f) e^{2\pi i f t - i\Psi(f)} \right],$$

$$(A17)$$

which is used in the IFFT-FFT method for the computation of $(\boldsymbol{h}, \boldsymbol{h})$. The integral from $f^{(b)} - \Delta f^{(b)}$ to $f^{(b)}$ can be approximately evaluated as follows,

$$\int_{f^{(b)} - \Delta f^{(b)}}^{f^{(b)}} df \sqrt{w^{(b)}(f)} B(f) e^{2\pi i f t - i\Psi(f)}$$

$$\simeq A(t(f^{(b)})) e^{2\pi i f^{(b)} t - i\Psi(f^{(b)})}$$

$$\times \left[\frac{i}{4\pi} \frac{\sqrt{t'(f^{(b)})}}{t(f^{(b)}) - t} - \frac{1}{16\pi} \frac{\sqrt{t'(f^{(b)})}}{\Delta f^{(b)}(t(f^{(b)}) - t)^2} \right.$$

$$\left. \times e^{2\pi i \Delta f^{(b)}(t(f^{(b)}) - t) - \pi i t'(f^{(b)})(\Delta f^{(b)})^2} \right],$$

$$(A18)$$

where the following formula has been used,

$$\int_{-1}^{0} \cos\left(\frac{\pi}{2} x\right) e^{-ipx - iqx^2} dx = \frac{i}{p} - \frac{\pi e^{i(p-q)}}{2p^2} + \mathcal{O}\left(\frac{1}{p^3}\right).$$

$$(A19)$$

Thus, the windowed waveform is approximately given by

$$h^{(b)}(t) \simeq -\frac{A(t(f^{(b)}))}{8\pi} \frac{\sqrt{t'(f^{(b)})}}{\Delta f^{(b)}(t(f^{(b)}) - t)^2}$$

$$\times \cos(2\pi f^{(b)} t - \Psi(f^{(b)}) + 2\pi \Delta f^{(b)}(t(f^{(b)}) - t)$$

$$- \pi t'(f^{(b)})(\Delta f^{(b)})^2),$$

$$(A20)$$

and it is quickly attenuated in proportion to $(t(f^{(b)}) - t)^{-2}$.

APPENDIX B: APPLICATION TO ROQ

Our technique can be used to reduce the file size of ROQ basis vectors and speed up the precomputations of ROQ. ROQ approximates template waveforms by the linear combinations of reduced basis vectors $\{B_j\}_{j=1}^{N_L}$ [23–25],

$$\tilde{h}(f_k) \simeq \sum_{j=1}^{N_L} B_j(f_k) \tilde{h}(F_j; t_c = 0) e^{-2\pi i f_k t_c}, \quad \text{(B1)}$$

where $\{F_j\}_{j=1}^{N_L}$ is the subset of frequency samples determined by the empirical interpolation algorithm (See the algorithm 2 of [59]). Substituting it into the original form of $(\boldsymbol{d}, \boldsymbol{h})$, (3), we obtain

$$(\boldsymbol{d}, \boldsymbol{h}) \simeq \Re \left[\sum_{j=1}^{N_L} \omega_j(t_c) \tilde{h}(F_j; t_c = 0) \right], \quad \text{(B2)}$$

$$\omega_j(t_c) \equiv \frac{4}{T} \sum_{k=1}^{\lfloor (N-1)/2 \rfloor} \frac{\tilde{d}_k^* B_j(f_k)}{S_k} e^{-2\pi i f_k t_c}. \quad \text{(B3)}$$

$\{\omega_j(t_c)\}_{j=1}^{N_L}$ are referred to as ROQ weights, and need to be precomputed before sampling. On the other hand, substituting (B1) into our approximate form of $(\boldsymbol{d}, \boldsymbol{h})$, (24), we obtain

$$(\boldsymbol{d}, \boldsymbol{h}) \simeq \Re \left[\sum_{j=1}^{N_L} \omega_j^{\text{MB}}(t_c) \tilde{h}(F_j; t_c = 0) \right], \quad \text{(B4)}$$

$$\omega_j^{\text{MB}}(t_c)$$

$$\equiv \sum_{b=0}^{B-1} \frac{4}{T^{(b)}} \times \Re \left[\sum_{k=K_s^{(b)}}^{K_e^{(b)}} w^{(b)}(f_k^{(b)}) \tilde{D}_k^{(b)*} B_j(f_k^{(b)}) e^{-2\pi i f_k^{(b)} t_c} \right].$$

$$(B5)$$

The computation of (B5) requires basis vectors only at K_{MB} frequency samples while (B3) requires them at all the K_{orig} frequency samples. This means we do not need to store basis vectors at all the frequency samples under our approximation, and their file size can be reduced by a factor of $K_{\text{orig}}/K_{\text{MB}}$. Since the size of basis vectors for BNS waveforms can be $\mathcal{O}(10)$ GB or even larger, this is practically useful. Comparing (B3) and (B5), we also find that our technique reduces the floating-point operations required to calculate ROQ weights by a factor of $K_{\text{orig}}/K_{\text{MB}}$.

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

[2] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **119,** 161101 (2017).

[3] B. Abbott *et al.* (LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, and MASTER Collaborations), Nature (London) **551,** 85 (2017).

[4] M. Drout *et al.*, Science **358**, 1570 (2017).
[5] M. Kasliwal *et al.*, Science **358**, 1559 (2017).
[6] P. Cowperthwaite *et al.*, Astrophys. J. Lett. **848**, L17 (2017).
[7] N. Tanvir *et al.*, Astrophys. J. Lett. **848**, L27 (2017).
[8] Y. Utsumi *et al.* (J-GEM Collaboration), Publ. Astron. Soc. Jpn. **69**, 101 (2017).
[9] M. Tanaka *et al.*, Publ. Astron. Soc. Jpn. **69**, psx12 (2017).
[10] E. Troja *et al.*, Nature (London) **551**, 71 (2017).
[11] G. Hallinan *et al.*, Science **358**, 1579 (2017).
[12] K. Alexander *et al.*, Astrophys. J. Lett. **848**, L21 (2017).
[13] R. Margutti *et al.*, Astrophys. J. Lett. **848**, L20 (2017).
[14] K. Mooley *et al.*, Nature (London) **554**, 207 (2018).
[15] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. X **9**, 031040 (2019).
[16] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. X **11**, 021053 (2021).
[17] G. M. Harry (LIGO Scientific Collaboration), Classical Quantum Gravity **27**, 084006 (2010).
[18] F. Acernese *et al.* (VIRGO Collaboration), Classical Quantum Gravity **32**, 024001 (2015).
[19] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Astrophys. J. Lett. **882**, L24 (2019).
[20] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Astrophys. J. Lett. **913**, L7 (2021).
[21] J. Veitch *et al.*, Phys. Rev. D **91**, 042003 (2015).
[22] G. Ashton *et al.*, Astrophys. J. Suppl. Ser. **241**, 27 (2019).
[23] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, Phys. Rev. Lett. **114**, 071104 (2015).
[24] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Prrer, V. Raymond, and P. Schmidt, Phys. Rev. D **94**, 044031 (2016).
[25] S. Morisaki and V. Raymond, Phys. Rev. D **102**, 104020 (2020).
[26] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy, Phys. Rev. D **92**, 023002 (2015).
[27] J. Lange, R. O'Shaughnessy, and M. Rizzo, arXiv:1805.10457.
[28] D. Wysocki, R. O'Shaughnessy, J. Lange, and Y.-L. L. Fang, Phys. Rev. D **99**, 084026 (2019).
[29] S. Vinciguerra, J. Veitch, and I. Mandel, Classical Quantum Gravity **34**, 115006 (2017).
[30] B. Zackay, L. Dai, and T. Venumadhav, arXiv:1806.08792.
[31] C. Talbot, R. Smith, E. Thrane, and G. B. Poole, Phys. Rev. D **100**, 043030 (2019).
[32] N. J. Cornish, arXiv:1007.4820.
[33] N. J. Cornish, Phys. Rev. D **103**, 104057 (2021).
[34] R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, Mon. Not. R. Astron. Soc. **498**, 4492 (2020).
[35] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini, and R. Murray-Smith, arXiv:1909.06296.

[36] S. R. Green, C. Simpson, and J. Gair, Phys. Rev. D **102**, 104057 (2020).
[37] F. Marion, in *Rencontres de Moriond Gravitational Waves and Experimental Gravity, Moriond Workshops*, edited by T. T. V. J. Dumarchez (The GIOI Publishers, Les Arcs, France, 2003), pp. 145–150.
[38] D. Buskulic (LIGO Scientific and Virgo Collaborations), Classical Quantum Gravity **27**, 194013 (2010).
[39] K. Cannon *et al.*, Astrophys. J. **748**, 136 (2012).
[40] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. Sathyaprakash, Phys. Rev. D **80**, 084043 (2009).
[41] K. Thorne, Rev. Mod. Phys. **52**, 299 (1980).
[42] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Phys. Rev. D **93**, 044007 (2016).
[43] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, Phys. Rev. Lett. **120**, 161102 (2018).
[44] L. Blanchet, Living Rev. Relativity **9**, 4 (2006).
[45] I. M. Romero-Shaw *et al.*, Mon. Not. R. Astron. Soc. **499**, 3295 (2020).
[46] S. Hild *et al.*, Classical Quantum Gravity **28**, 094013 (2011).
[47] B. P. Abbott *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **34**, 044001 (2017).
[48] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Astrophys. J. Lett. **896**, L44 (2020).
[49] Gravitational Wave Open Science Center, GW190814, https://www.gw-openscience.org/eventapi/html/O3_Discovery_Papers/GW190814/v1/.
[50] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, Phys. Rev. D **100**, 024059 (2019).
[51] S. Khan, F. Ohme, K. Chatziioannou, and M. Hannam, Phys. Rev. D **101**, 024056 (2020).
[52] E. Payne, C. Talbot, P. D. Lasky, E. Thrane, and J. S. Kissel, Phys. Rev. D **102**, 122004 (2020).
[53] S. R. Cook, A. Gelman, and D. B. Rubin, J. Comput. Graph. Stat. **15**, 675 (2006).
[54] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, arXiv:1804.06788.
[55] J. S. Speagle, Mon. Not. R. Astron. Soc. **493**, 3132 (2020).
[56] L. P. Singer and L. R. Price, Phys. Rev. D **93**, 024013 (2016).
[57] E. Thrane and C. Talbot, Pub. Astron. Soc. Aust. **36**, e010 (2019); **37**, E036 (2020).
[58] J. D. Creighton and W. G. Anderson, *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis* (Wiley-VCH, Weinheim, Germany, 2011).
[59] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, Phys. Rev. X **4**, 031006 (2014).