

## Exploring the Bayesian parameter estimation of binary black holes with LISA

Sylvain Marsat,<sup>1,2</sup> John G. Baker,<sup>3</sup> and Tito Dal Canton<sup>3,4,\*</sup>

<sup>1</sup>*APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*

<sup>2</sup>*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam-Golm 14476, Germany*

<sup>3</sup>*Gravitational Astrophysics Laboratory, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA*

<sup>4</sup>*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France*



(Received 1 March 2020; accepted 15 March 2021; published 15 April 2021)

The space-based gravitational wave detector LISA will observe mergers of massive black hole binary systems (MBHBs) to cosmological distances, as well as inspiralling stellar-origin (or stellar-mass) binaries (SBHBs) years before they enter the LIGO/Virgo band. Much remains to be explored for the parameter recovery of both classes of systems. Previous MBHB analyses relied on inspiral-only signals and/or a simplified Fisher matrix analysis, while SBHBs have not yet been extensively analyzed with Bayesian methods. We accelerate likelihood computations by (i) using a Fourier-domain response of the LISA instrument, (ii) using a reduced order model for nonspinning waveforms that include a merger-ringdown and higher harmonics, and (iii) setting the noise realization to zero and computing overlaps in the amplitude/phase representation. We present the first simulations of Bayesian inference for the parameters of massive black hole systems including consistently the merger and ringdown of the signal, as well as higher harmonics. We clarify the roles of LISA response time and frequency dependencies in breaking degeneracies and illustrate how degeneracy breaking unfolds over time. We also find that restricting the merger-dominated signal to its dominant harmonic can make the extrinsic likelihood very degenerate. Including higher harmonics proves to be crucial to breaking degeneracies and considerably improves the localization of the source, with a surviving bimodality in the sky position. We also present simulations of Bayesian inference for the extrinsic parameters of SBHBs, and show that although unimodal, their posterior distributions can have non-Gaussian features.

DOI: [10.1103/PhysRevD.103.083011](https://doi.org/10.1103/PhysRevD.103.083011)

### I. INTRODUCTION

Gravitational waves from the coalescence of black hole binaries are now regularly observed [1] by the ground-based interferometers Advanced LIGO [2], Advanced Virgo [3], and, soon, KAGRA [4].

Ground-based interferometers are fundamentally limited at low frequency by terrestrial noise, and cannot be used to study mergers of compact objects much heavier than a few  $10^2 M_{\odot}$ . The spaceborne detector LISA [5] will overcome this limitation, enabling the observation and precise characterization of gravitational waves from binary black holes, from the coalescence of systems with millions of  $M_{\odot}$  [massive black hole binaries (MBHB) [6]] to the inspiral of systems with tens of  $M_{\odot}$  [stellar-mass black hole binaries (SBHB) [7]]. The observation of black hole mergers and

inspirals in the LISA frequency band will yield major scientific rewards, and they form the main target of LISA among other classes of sources.

In order to realize the full scientific objectives of LISA with respect to black hole mergers, adequate data analysis tools must be prepared in advance. Similar to what happens in LIGO and Virgo, one has to first identify the presence of a merger waveform in the LISA data, possibly in the presence of other superposed signals (notably galactic white dwarf binaries [8]), and then infer the distribution of the physical parameters of its source from the data, enabling the construction of a catalog of black hole binaries. Inferring the parameters of binaries, particularly their distance and sky location, and refining the analysis as the signal accumulates with observation time will also be necessary to organize multimessenger observations of their late inspirals and mergers [9,10]. Cosmological applications using LISA observations as standard sirens [11,12] also depend on the ability to localize individual sources.

\*NASA Postdoctoral Program fellow

Understanding parameter estimation of SBHBs will also be important to understanding the outcomes and challenges of possible multiband gravitational wave observations [7]. In this paper we focus on the latter part of the analysis problem, i.e., the inference of the black hole parameters from the LISA data, once we have reasons to believe the presence of the signal in the data, leaving aside the identification problem. We also limit our analysis to extrinsic parameters and ignore the effect of superposed gravitational wave signals from various sources.

The inference problem amounts to producing samples from the posterior distribution for source parameters related by Bayes theorem to the likelihood function, i.e., the probability of observing the measured data given the source parameters and a model of the source and detector. In LISA, this problem is more complex than in kilometer-scale ground-based interferometers, as additional challenges arise from the much larger expected signal-to-noise ratios (SNRs) for high-mass systems, or the much longer duration of the waveform for low-mass systems. In particular, for MBHBs, the large SNR will require a very accurate modeling of the waveform, including the merger and ringdown regimes, the effects of spin, and corrections of higher harmonics in the signal beyond the quadrupole.

The response of the LISA instrument to a given waveform shows a time- and frequency-dependence [13–16]. Contrary to LIGO and Virgo, which are approximately fixed in an inertial frame during the observation of any merger, the LISA constellation will move around the Sun and change its orientation appreciably if the signal is observable for a significant fraction of a year, leading to additional modulations of the signals. Furthermore, due to the LISA arm length, for many sources the long-wavelength approximation will not hold, which introduces a frequency-dependence in the response. As we will see in this study, proper treatment of these effects is necessary in order to understand degeneracies between the source parameters.

Previous parameter recovery studies with LISA used mainly simplified signal and response models, and often relied on a Fisher matrix approximation to parameter recovery instead of full Bayesian simulations. Numerous works used the combination of inspiral-only post-Newtonian waveforms (sometimes including precession), the low-frequency approximate response, and Fisher matrix estimates [13,17–20]. The Mock LISA Data Challenges [21] used such a setup for the signals and the response, albeit moving towards Markov chain Monte Carlo (MCMC) tools with a focus on detection. Bayesian methods going beyond Fisher, albeit still restricted to inspiral signals, were developed in [22–31]. The importance of the higher harmonics for LISA was stressed already in several studies [32–35]. To explore the importance of the merger-ringdown using numerical relativity waveforms, the studies [35–38] used a Fisher approach, while Ref. [39] used Bayesian analyses limited to extrinsic

parameters. More recently, in the context of the redesign of the LISA mission [40], the study [6] explored the performance of various LISA instrumental designs using Fisher matrix estimates and inspiral signals, but used a reweighting procedure to represent the role of the merger-ringdown signal. We also note a recent work [41] investigating parameter recovery for ringdown-dominated signals. To this date, no full Bayesian parameter estimation studies with inspiral-merger-ringdown (IMR) signals have been performed.

On the other hand, important advances have been registered in recent years in providing fast and accurate waveforms including the merger and ringdown for the LIGO/Virgo data analysis through phenomenological waveforms [42,43], reduced order models (ROM) of effective-one-body (EOB) waveforms [44,45], and numerical relativity surrogates [46,47]. This progress has yet to be transposed to LISA applications.

Here we demonstrate that standard Bayesian inference can be performed for MBHBs using a self-consistent waveform model that includes inspiral, merger, ringdown, and higher-order modes (albeit no spins) and a full model of the LISA response. We investigate and explain degeneracies in the posterior distribution of the source’s parameters, and highlight the crucial role played by the higher harmonics in the signal and the frequency-dependency in the instrument response.

In the present study, we will consider a somewhat reduced parameter space, not including the effects of the spins of the two black holes. These could lead to more complicated waveforms, notably through precession effects, which we leave for future work. It should be noted, however, that aligned spin components, though quantitatively important, are not expected to change many of our conclusions due to the relative disconnect between the inference of intrinsic and extrinsic parameters.

We start by introducing a fast method to calculate the likelihood function for black hole merger signals. This method includes: a fast computation of a Fourier-domain inspiral-merger-ringdown waveform with higher-order modes based on a reduced order model; a fast Fourier-domain computation of the LISA response based on [16]; and a fast method for computing the noise-weighted product between waveforms when setting the noise realization to zero. We couple our likelihood with two codes for Bayesian inference based on different sampling techniques, constructing an inference engine which enables us to perform a variety of investigations with simulated black hole mergers.

Focusing on two examples of moderately massive black holes binaries, we perform Bayesian parameter estimation simulation for full inspiral-merger signals, and we show the crucial role of higher-order modes. We highlight the challenges we encounter in the sampling, and compare our two samplers with each other and with the Fisher matrix approximation. We explore degeneracies appearing

between some of the parameters when ignoring these higher harmonics, and derive an analytic explanation of their origin via simplified approximations to the instrument response. We show for our examples how the posterior distributions evolve over time as more and more data becomes available for the inference, and explain how degeneracies between different positions of the source in the sky are broken by features in the instrument response.

Finally, we repeat some of our investigations for stellar-origin black hole mergers. Parameter estimation studies for SBHB systems have relied so far on the Fisher matrix approach [7,48–50], with Bayesian inference tools yet to be developed. Focusing on the posterior distribution of extrinsic parameters (fixing the masses and time to coalescence), we demonstrate that Bayesian analyses can be run with a speed for the likelihood computation that is comparable to the MBHB case, and discuss the features of the posteriors.

Section II introduces comprehensive notations for the frequency-domain signal and response, and Sec. III explains our methods for likelihood computations and Bayesian sampling. Section IV presents the MBHB example signals that we analyze and their accumulation with time, together with their instrumental response in various approximations. In Sec. V, we present our main results for the parameter estimation of MBHBs; we contrast results obtained with and without higher harmonics, contrast the Fisher matrix with Bayesian results, and describe and explain degeneracies in the parameter space. In Sec. VI, we turn to the refinement of parameter estimation as the signal accumulates with ongoing observations, describing how multiple inferred sky positions can coexist before features of the full instrument response break the degeneracies. In Sec. VII, we turn to the application of our fast likelihood computation to investigate the extrinsic parameter inference of SBHBs. We summarize and discuss our findings in Sec. VIII.

## II. LISA INSTRUMENT RESPONSE

In this section, we introduce a complete set of notations for the LISA instrument response that will be useful in the later discussion of degeneracies in parameter space. We will use units with  $G = c = 1$  and assume the proposed arm length for LISA,  $L = 2.5$  Gm [5].

### A. The GW signal

To describe the gravitational wave signal, we first need to define a conventional source frame associated to the binary system, as detailed in the Appendix A. The gravitational wave propagation vector  $k$ , pointing from the source to the observer, has polar angles  $(\iota, \varphi)$  in this source frame. Next, one introduces polarization vectors  $p, q$  so that  $(p, q, k)$  is a direct triad. The precise choice of  $(p, q)$  is a matter of convention, for which we refer to Appendix A.

The gravitational waveform in the transverse-traceless gauge is described by the two polarizations  $h_+, h_\times$ . If  $H_{ij} = h_{ij}^{\text{TT}}$  represents the gravitational wave signal in matrix form,

$$H = h_+ P_+ + h_\times P_\times, \quad (1)$$

with the polarization tensors

$$P_+ = p \otimes p - q \otimes q, \quad (2a)$$

$$P_\times = p \otimes q + q \otimes p. \quad (2b)$$

Conversely, the polarizations are

$$h_+ = \frac{1}{2}(p \otimes p - q \otimes q) : H, \quad (3a)$$

$$h_\times = \frac{1}{2}(p \otimes q + q \otimes p) : H, \quad (3b)$$

with the notation  $A : B = A_{ij} B_{ij}$ .

One can further decompose the gravitational wave signal, seen as a function of the direction of emission  $(\iota, \varphi)$  in the source frame, in spin-weighted spherical harmonics [51] as

$$h_+ - ih_\times = \sum_{\ell \geq 2} \sum_{m=-\ell}^{\ell} {}_{-2}Y_{\ell m}(\iota, \varphi) h_{\ell m}. \quad (4)$$

Explicit expression of the  ${}_{-2}Y_{\ell m}$  can be found in, e.g., Sec. 3 of [52]. In the following, we will use exclusively this parametrization of the waveform as a set of modes  $h_{\ell m}$ . The dominant harmonic is  $h_{22}$ , while the others are called higher modes (HM) or higher harmonics.

We will both generate the waveforms and apply the response directly in the Fourier domain. Our convention for the Fourier transform of a function  $F$  is<sup>1</sup>

$$\tilde{F}(f) = \int dt e^{2i\pi f t} F(t). \quad (5)$$

### B. Mode decomposition and polarization angle

We now translate the mode decomposition (4) in the Fourier domain. We have

$$h_+ = \frac{1}{2} \sum_{\ell, m} ({}_{-2}Y_{\ell m} h_{\ell m} + {}_{-2}Y_{\ell m}^* h_{\ell m}^*), \quad (6a)$$

$$h_\times = \frac{i}{2} \sum_{\ell, m} ({}_{-2}Y_{\ell m} h_{\ell m} - {}_{-2}Y_{\ell m}^* h_{\ell m}^*), \quad (6b)$$

<sup>1</sup>Note that this differs by a change  $f \rightarrow -f$  from the more usual convention used, e.g., in LAL [53].

which is valid in general [we dropped the  $(t, \varphi)$  arguments of the  ${}_{-2}Y_{\ell m}$ ]. Now, for nonprecessing binary systems, an exact symmetry relation between modes reads

$$h_{\ell, -m} = (-1)^\ell h_{\ell, m}^*. \quad (7)$$

Using this symmetry, we can write

$$h_{+, \times} = \sum_{\ell, m} K_{\ell m}^{+, \times} h_{\ell m}, \quad (8)$$

with

$$K_{\ell m}^+ = \frac{1}{2} ({}_{-2}Y_{\ell m} + (-1)^\ell {}_{-2}Y_{\ell, -m}^*), \quad (9a)$$

$$K_{\ell m}^\times = \frac{i}{2} ({}_{-2}Y_{\ell m} - (-1)^\ell {}_{-2}Y_{\ell, -m}^*). \quad (9b)$$

Going to the Fourier domain, an approximation often used is to neglect support for negative/positive frequencies according to

$$\tilde{h}_{\ell m}(f) \simeq 0 \quad \text{for } m < 0, \quad f > 0 \quad (m > 0, f < 0), \quad (10)$$

and neglecting modes  $h_{\ell 0}$ . We will use this approximation throughout this paper. Note that we picked our Fourier convention (5) to ensure that modes  $\tilde{h}_{\ell m}(f)$  with  $m > 0$  have support for  $f > 0$ . Using (10), for  $f > 0$  we have

$$\tilde{h}_{+, \times} = \sum_{\ell} \sum_{m>0} K_{\ell m}^{+, \times} \tilde{h}_{\ell m}. \quad (11)$$

Next, it is convenient to introduce mode-by-mode polarization matrices,

$$P_{\ell m} = P_+ K_{\ell m}^+ + P_\times K_{\ell m}^\times, \quad (12)$$

so that

$$H = \sum_{\ell, m>0} P_{\ell m} h_{\ell m}. \quad (13)$$

The polarization angle  $\psi$  can be seen (see Appendix A) as a degree of freedom in the relation between the source frame and the detector frame, parametrizing a rotation around the wave vector  $k$ . To define this angle, one introduces reference vectors  $(u, v)$  orthogonal to  $k$ , that define the zero of the polarization angle as  $(p, q)(\psi = 0) = (u, v)$ . Our convention for  $(u, v)$  is detailed in Appendix A.

To make explicit the dependence in polarization, we can define polarization tensors for a zero polarization angle as

$$P_+^0 = P_+(\psi = 0) = u \otimes u - v \otimes v, \quad (14a)$$

$$P_\times^0 = P_\times(\psi = 0) = u \otimes v + v \otimes u. \quad (14b)$$

This allows us to write the dependence in polarization as

$$P_+ + iP_\times = e^{-2i\psi} (P_+^0 + iP_\times^0), \quad (15)$$

or explicitly in  $P_{\ell m}$  as

$$P_{\ell m}(t, \varphi, \psi) = \frac{1}{2} {}_{-2}Y_{\ell m}(t, \varphi) e^{-2i\psi} (P_+^0 + iP_\times^0) + \frac{1}{2} (-1)^\ell {}_{-2}Y_{\ell, -m}^*(t, \varphi) e^{+2i\psi} (P_+^0 - iP_\times^0). \quad (16)$$

Writing the  $P_{\ell m}$  matrices in this way allows us to factor out explicitly all dependencies in the extrinsic parameters  $(t, \phi, \psi)$ . Together with the luminosity distance  $D$  scaling the overall amplitude of the signal, these parameters enter as constant (time- and frequency-independent) prefactors in the response for each mode  $h_{\ell m}$ .

Since  $\psi$  always appears with a factor of 2, it has an exact  $\pi$ -degeneracy, and we choose the convention of restricting  $\psi \in [0, \pi]$ .

### C. Frequency-domain LISA response

The LISA response [13–15, 54, 55] can be built from single-link observables  $y_{slr} = (\nu_r - \nu_s)/\nu$ , representing a laser frequency shift between the transmitting spacecraft  $s$  and the receiving spacecraft  $r$  along the link  $l$ . We use the expression [56, 57]

$$y_{slr} = \frac{1}{2} \frac{n_l \otimes n_l}{1 - k \cdot n_l} : [H(t - L - k \cdot p_s) - H(t - k \cdot p_r)], \quad (17)$$

where  $H$  is the transverse-traceless matrix representing the gravitational wave,  $k$  is the wave propagation vector,  $L$  is the delay along one arm, taken to be fixed,  $n_l$  are the link unit vectors (from  $s$  to  $r$ ), and  $p_s, p_r$  are the positions of the spacecrafts. Here  $n_l, p_s,$  and  $p_r$  are evaluated at the same time,  $t$ . In the following, we will use interchangeably the notation  $y_{sr}$  instead of  $y_{slr}$ , since  $n_l$  can be deduced from the sending and receiving indices  $s, r$ .

Our response formalism applies to waveforms which can be represented as a combination of harmonics with slowly varying amplitude and phase as<sup>2</sup>

$$\tilde{h}_{\ell m}(f) = A_{\ell m}(f) e^{-i\Psi_{\ell m}(f)}. \quad (18)$$

We will use the analysis of [16] and write the response in individual observables  $y_{slr}$ , with a transfer function for each spherical harmonic mode as

<sup>2</sup>Note our Fourier convention (5).



$$\tilde{y}_{slr} = \sum_{\ell,m} \mathcal{T}_{slr}^{\ell m}(f) \tilde{h}_{\ell m}. \quad (19)$$

How to take the Fourier transform of (17) was investigated in detail in the perturbative formalism of [16]. The main timescales to compare are the radiation-reaction timescale  $\propto 1/\sqrt{\dot{\omega}}$  of the binary and the LISA orbital timescale 1 yr. At leading order in the separation of these timescales, for a signal that is chirping fast enough, we have simply

$$\mathcal{T}_{slr}^{\ell m}(f) = G_{slr}^{\ell m}(f, t_f^{\ell m}), \quad (20)$$

where

$$G_{slr}^{\ell m}(f, t) = \frac{i\pi f L}{2} \text{sinc}[\pi f L(1 - k \cdot n_l)] \cdot \exp[i\pi f(L + k \cdot (p_r + p_s))] n_l \cdot P_{\ell m} \cdot n_l, \quad (21)$$

is the kernel from [16] with  $n_l$ ,  $p_s$ ,  $p_r$  evaluated at  $t$ , with  $P_{\ell m}$  defined by the decomposition (13). In (20),

$$t_f^{\ell m} = -\frac{1}{2\pi} \frac{d\Psi_{\ell m}}{df}, \quad (22)$$

is the effective time-frequency correspondence, defined across the whole frequency band and including the merger and ringdown. This definition generalizes the stationary phase approximation (SPA).

The analysis of [16] has shown that higher-order corrections in the separation of timescales in the LISA response are small in general for MBHB systems, and are also small for SBHB systems provided that they are not too far from the coalescence. The fact that we use the same Fourier-domain treatment of the transfer functions for both the signal and the templates should also mitigate the importance of modeling errors. We will therefore limit ourselves to the leading order in the treatment of [16] in the rest of this paper.

In (21) it is convenient to decompose the spacecraft positions as

$$p_{r,s} = p_0 + p_{r,s}^L, \quad (23)$$

where  $p_0$  is the position of the center of the LISA constellation. This allows us to note an important feature of (21): apart from a global phase delay factor  $\exp[i\pi f k \cdot p_0]$ , frequency-dependent terms only feature the projections

$$k_{\parallel} \cdot n_l, \quad k_{\parallel} \cdot p_{s,r}^L, \quad (24)$$

where  $k_{\parallel}$  is the projection of the wave vector  $k$  in the (instantaneous) LISA plane, and thus the frequency-dependent factors are invariant when we reflect  $k$  across

this plane. The delay  $k \cdot p_0$  is the only one with a baseline outside the LISA plane [see (36) below]. This will remain true when constructing time delay interferometry (TDI) variables below.<sup>3</sup> We will investigate sky degeneracies in detail in Sec. VB and Sec. VIB.

In Eqs. (17) and (21), one can distinguish two types of delays: on one hand the delay associated to the position  $p_0$  of the center of the constellation on its orbit around the Sun, with baseline  $R = 1$  au, and on the other hand the delays associated to the individual spacecraft positions in the constellation, with the baseline of the arm length  $L$ . This defines the transfer frequencies

$$f_R = 1/R = 2.0 \times 10^{-3} \text{ Hz}, \quad (25a)$$

$$f_L = 1/L = 0.12 \text{ Hz}. \quad (25b)$$

The transfer frequency  $f_L$  corresponds to fitting a full wavelength in the arm length; as we will see below in Sec. VI, departures from the long-wavelength approximation starts to be important at significantly lower frequencies.

#### D. Time delay interferometry

The basic one-arm observables of Eq. (17) are affected by laser noise whose amplitude is orders of magnitude larger than the astrophysical signal. However, TDI allows one to construct a new set of observables from delayed combinations of  $y_{slr}$ , where laser noise is suppressed by orders of magnitude [58–62]. Various generations of TDI schemes have been proposed in order to deal with a nonrigid and rotating LISA constellation [63–66]. However, these refinements only marginally affect the response to the gravitational waves. Hence, in this work we only consider first-generation TDI and adopt a rigid approximation for the constellation, where delays are all constant and equal to  $L$ . Using the notation  $y_{slr,nL} = y_{slr}(t - nL)$ , the first-generation TDI Michelson observable  $X$  reads [56]:

$$X = y_{31} + y_{13,L} + (y_{21} + y_{12,L})_{,2L} - (y_{21} + y_{12,L}) - (y_{31} + y_{13,L})_{,2L}, \quad (26)$$

with the other Michelson observables  $Y$ ,  $Z$  being obtained by cyclic permutation. Uncorrelated combinations  $A$ ,  $E$ , and  $T$  [67] are then expressed as

$$A = \frac{1}{\sqrt{2}}(Z - X), \quad (27a)$$

<sup>3</sup>Strictly speaking, it will not be true in the most general response formalism of [16], where corrections are either slightly nonlocal in time or involve velocities that are out of the LISA plane.

$$E = \frac{1}{\sqrt{6}}(X - 2Y + Z), \quad (27b)$$

$$T = \frac{1}{\sqrt{3}}(X + Y + Z). \quad (27c)$$

These channels are independent under the assumption of an identical and uncorrelated noise in the detector arms. Note that various conventions coexist in the literature. With constant delays in the rigid approximation, and using the notation  $z \equiv \exp[2i\pi fL]$ , the TDI combinations in the frequency domain take the form

$$\tilde{a} = (1+z)(\tilde{y}_{31} + \tilde{y}_{13}) - \tilde{y}_{23} - z\tilde{y}_{32} - \tilde{y}_{21} - z\tilde{y}_{12}, \quad (28a)$$

$$\begin{aligned} \tilde{e} = \frac{1}{\sqrt{3}} & [(1-z)(\tilde{y}_{13} - \tilde{y}_{31}) + (2+z)(\tilde{y}_{12} - \tilde{y}_{32}) \\ & + (1+2z)(\tilde{y}_{21} - \tilde{y}_{23})], \end{aligned} \quad (28b)$$

$$\tilde{t} = \frac{\sqrt{2}}{\sqrt{3}}[\tilde{y}_{21} - \tilde{y}_{12} + \tilde{y}_{32} - \tilde{y}_{23} + \tilde{y}_{13} - \tilde{y}_{31}], \quad (28c)$$

where we have eliminated frequency-dependent prefactors that are common to the signal and to the noise by introducing the rescalings

$$\tilde{a}, \tilde{e} = \frac{e^{-2i\pi fL}}{i\sqrt{2}\sin(2\pi fL)} \times \tilde{A}, \tilde{E}, \quad (29a)$$

$$\tilde{t} = \frac{e^{-3i\pi fL}}{2\sqrt{2}\sin(\pi fL)\sin(2\pi fL)} \times \tilde{T}. \quad (29b)$$

We will use mode-by-mode transfer functions for these reduced channels as

$$\tilde{a}, \tilde{e}, \tilde{t} = \sum_{\ell, m} \mathcal{T}_{a,e,t}^{\ell m} \tilde{h}_{\ell m}. \quad (30)$$

To make the connection between these reduced TDI observables and the gravitational strain, more familiar in the context of ground-based instruments, we also introduce the notations

$$\tilde{h}_{a,e,t} \equiv \frac{1}{(-6i\pi fL)} \times \tilde{a}, \tilde{e}, \tilde{t}, \quad (31a)$$

$$\mathcal{T}_{h_a, h_e, h_t}^{\ell m} = \frac{1}{(-6i\pi fL)} \mathcal{T}_{a,e,t}^{\ell m}. \quad (31b)$$

Scaling out the same square factors from the noise power spectral density (PSD) as

$$S_n^A, S_n^E = 2\sin^2(2\pi fL) \times S_n^a, S_n^e, \quad (32a)$$

$$S_n^T = 8\sin^2(\pi fL)\sin^2(2\pi fL)S_n^t, \quad (32b)$$

the reduced PSD for the three channels takes the form

$$\begin{aligned} S_n^a = S_n^e = 2(3 + 2\cos(2\pi fL) + \cos(4\pi fL))S^{\text{pm}}(f) \\ + (2 + \cos(2\pi fL))S^{\text{op}}(f), \end{aligned} \quad (33a)$$

$$S_n^t = 4\sin^2(2\pi fL)S^{\text{pm}}(f) + S^{\text{op}}(f), \quad (33b)$$

where  $S^{\text{pm}}$  is the test-mass noise PSD and  $S^{\text{op}}$  is the optical noise PSD. We also include a confusion noise coming from the background of galactic binaries in the LISA band that is added to the instrumental noise. For the instrument performance defining these noise levels, we take values from [68] (see Appendix A 4). We can define a strainlike noise PSD associated with the strainlike TDI observables (31) as

$$S_n^{a,e,t}(f) = \frac{S_n^{a,e,t}(f)}{(6\pi fL)^2}. \quad (34)$$

The prefactors (29) are oscillatory and have zero-crossings at high frequencies, with the first one occurring at  $f_L/2 = 0.06$  Hz, which is why it is convenient to factor them out to avoid 0/0 numerical instabilities. This treatment would not apply directly to a more realistic model for LISA with varying arm lengths and residual laser noise. In that case, imperfect cancellations in the vicinity of the zero-crossings would likely result in a localized loss of sensitivity.

### E. The low-frequency limit

Though our parameter estimation calculation primarily applies the full LISA response, it will be useful to consider some simplifying asymptotic limits to understand parameter degeneracies.

As is well-known [13], in the low-frequency limit (also called the long-wavelength approximation), the finite arm length effects vanish, and the response of LISA is analogous to the response of two LIGO-type detectors rotated from each other by  $\pi/4$  and set in motion.

For  $f \ll f_L$ , we have  $2\pi fL \ll 1$  and the kernel (21) reduces to

$$G_{slr}^{\ell m} \simeq \frac{i\pi fL}{2} \exp[2i\pi f k \cdot p_0] n_l \otimes n_l : P_{\ell m}. \quad (35)$$

The exponential factor is a delay phase, for which we introduce the notation

$$\Phi_R \equiv 2\pi f k \cdot p_0, \quad \Delta\Phi_R = \Phi_R - \Phi_R(t = t_f^{\text{peak}}). \quad (36)$$

The quantity  $\Phi_R$  is often called the Doppler phase in the literature; it may appear to be large for  $f \gg f_R$ , but we should remember that it corresponds in part simply to the

fixed delay between the time of arrival at the SSB and the time of arrival at the LISA constellation. In the limit of short-lived coalescence signals, only the Doppler phase variation  $\Delta\Phi_R$  carries useful information about the sky position.

For  $2\pi fL \ll 1$ , we have  $z \simeq 1$ , and the link reversal symmetry  $\tilde{y}_{slr} \simeq \tilde{y}_{r-ls}$ , so that (28) becomes (dropping  $l$  indices and symmetrizing  $r, s$ )

$$\tilde{a} \simeq 4\tilde{y}_{31} - 2\tilde{y}_{23} - 2\tilde{y}_{12}, \quad (37a)$$

$$\tilde{e} \simeq 2\sqrt{3}[\tilde{y}_{12} - \tilde{y}_{23}], \quad (37b)$$

$$\tilde{i} \simeq 0, \quad (37c)$$

with the  $T$ -channel becoming negligible in this limit. Using (35), we can write

$$\tilde{a}, \tilde{e} = (-2i\pi f) \exp[2i\pi f k \cdot p_0] \sum_{\ell, m > 0} \tilde{h}_{\ell m} D_{a,e} : P_{\ell m}, \quad (38)$$

where we introduced the detector tensors

$$D_a = \frac{L}{2} (n_1 \otimes n_1 + n_3 \otimes n_3 - 2n_2 \otimes n_2), \quad (39a)$$

$$D_e = \frac{L\sqrt{3}}{2} (n_1 \otimes n_1 - n_3 \otimes n_3). \quad (39b)$$

Here, we have made apparent factors  $(-2i\pi f)$ , which correspond to a time derivative in the Fourier domain. The observables  $\tilde{a}, \tilde{e}$  are therefore analogous to time derivatives of the strain commonly used for ground-based detectors.

We can now map these two channels to two fictitious LIGO-type detectors as follows. For an orthogonal detector of the same setup as the ground-based LIGO and Virgo with arm length  $L_D$ , rotated by an angle  $\epsilon_D$  from the basis vectors  $x, y$ , the detector tensor is

$$D = \frac{L_D}{2} [\cos 2\epsilon_D (x \otimes x - y \otimes y) + \sin 2\epsilon_D (x \otimes y + y \otimes x)]. \quad (40)$$

Compared with (39), using the LISA frame as a detector frame (see Appendix A) we obtain

$$L_a = L_e = 3L, \quad \epsilon_a = \frac{2\pi}{3}, \quad \epsilon_e = \frac{5\pi}{12}. \quad (41)$$

Note a degeneracy  $\pm\pi$  in  $\epsilon_{a,e}$ , reflecting a freedom in the choice of the orientation of these effective detectors. The effective arm length  $3L$  is a matter of convention, since it depends on an arbitrary overall scaling in (27).

Next, we factor out the effective length, defining

$$F_{a,e}^{+,\times} \equiv \frac{1}{3L} D_{a,e} : P_{+,\times}^0, \quad (42)$$

which gives, in terms of the LISA-frame sky position angles  $\lambda_L, \beta_L$  (see Appendix A 3),

$$F_a^+ = \frac{1}{2} (1 + \sin^2 \beta_L) \cos \left( 2\lambda_L - \frac{\pi}{3} \right), \quad (43a)$$

$$F_a^\times = \sin \beta_L \sin \left( 2\lambda_L - \frac{\pi}{3} \right), \quad (43b)$$

$$F_e^+ = \frac{1}{2} (1 + \sin^2 \beta_L) \cos \left( 2\lambda_L + \frac{\pi}{6} \right), \quad (43c)$$

$$F_e^\times = \sin \beta_L \sin \left( 2\lambda_L + \frac{\pi}{6} \right). \quad (43d)$$

These functions are the familiar pattern functions of ground-based detectors (for  $\psi_L = 0$ ) (see, e.g., [69]). One can check that the expressions for the channel  $e$  are obtained from the expressions for  $a$  with the replacement  $\lambda_L \rightarrow \lambda_L + \pi/4$ .

If we introduce

$$F_{a,e}^{\ell m} = \frac{1}{2} {}_{-2}Y_{\ell m} e^{-2i\psi_L} (F_{a,e}^+ + iF_{a,e}^\times) + \frac{1}{2} (-1)^\ell {}_{-2}Y_{\ell, -m}^* e^{+2i\psi_L} (F_{a,e}^+ - iF_{a,e}^\times), \quad (44)$$

the mode transfer functions (30) become

$$\tilde{a}, \tilde{e} = \sum_{\ell, m} \mathcal{T}_{a,e}^{\ell m} \tilde{h}_{\ell m} = (-6i\pi f L) \exp[2i\pi f k \cdot p_0] \sum_{\ell, m} F_{a,e}^{\ell m} \tilde{h}_{\ell m}. \quad (45)$$

We see that scaling out  $(-6i\pi f L)$  as in (31) brings us back to strainlike observables. The only difference with the response of a ground-based observatory is the time-dependency entering  $p_0(t)$  and the LISA frame angles  $\lambda_L(t), \beta_L(t), \psi_L(t)$  [see (A15)], with the time evaluated at  $t_f^{\ell m}$ .

Finally, we note that we could include the polarization angle in the pattern functions as follows. Above we have chosen to write  $\psi$  as an outside prefactor, but we could write  $\psi$ -dependent pattern functions as

$$F_{a,e}^+(\lambda_L, \beta_L, \psi_L) = \cos 2\psi_L F_{a,e}^+ + \sin 2\psi_L F_{a,e}^\times, \quad (46a)$$

$$F_{a,e}^\times(\lambda_L, \beta_L, \psi_L) = -\sin 2\psi_L F_{a,e}^+ + \cos 2\psi_L F_{a,e}^\times, \quad (46b)$$

so that in the low-frequency approximation the response for the strainlike observables (31) is

$$\tilde{h}_{a,e} = F_{a,e}^+(\lambda_L, \beta_L, \psi_L) \tilde{h}_+ + F_{a,e}^\times(\lambda_L, \beta_L, \psi_L) \tilde{h}_\times, \quad (47)$$

if we ignore the delay  $k \cdot p_0$  and treat the angles as constants. This is the more familiar form of the instrument response used for ground-based detectors.

### III. METHODOLOGY

#### A. Bayesian setting

Introducing the standard matched-filter inner product [69] (also called overlap) as

$$(\tilde{a}|\tilde{b}) = 4\text{Re} \int_0^{+\infty} df \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}, \quad (48)$$

for stationary Gaussian noise with PSD  $S_n$ , the likelihood  $\mathcal{L} = p(d|\boldsymbol{\theta})$  for a given gravitational wave channel takes the form

$$\ln \mathcal{L} = -\frac{1}{2}(h(\boldsymbol{\theta}) - d|h(\boldsymbol{\theta}) - d), \quad (49)$$

with the data stream being a superposition of the gravitational wave signal for the true parameters  $\boldsymbol{\theta}_0$  and the noise realization in the experiment,  $d = h(\boldsymbol{\theta}_0) + n$ . The posterior distribution for the physical parameters  $\boldsymbol{\theta}$  given the observed data  $d$  is then

$$p(\boldsymbol{\theta}|d) = \frac{p(d|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(d)}, \quad (50)$$

where  $p(\boldsymbol{\theta})$  are the priors on the parameters, and  $p(d)$  the evidence. The focus of this work is not selecting between different models, so we will treat  $p(d)$  as a normalization constant and will not consider it further.

In general, the templates  $h(\boldsymbol{\theta})$  used for the analysis will only be an approximation of the physical signals present in the data stream  $d$ . In this work we will entirely ignore this distinction and assume that our template waveforms exactly match the astrophysical waveforms. In other words, we will not explore the issue of systematic errors arising from incomplete modeling of compact binary merger waveforms.

In our analysis, we will simulate signals and perform Bayesian analyses of them with Eqs. (49) and (50) to obtain the parameter posteriors. In doing so, we will use the so-called zero-noise approximation, i.e., setting  $n = 0$ . This allows us to accelerate the likelihood computation as explained in Sec. III C. The approximation is sufficient to explore the structure of the likelihood and the parameter degeneracies, greatly improving them with respect to the Fisher matrix approach. The extension of our analysis to include nonzero-noise realizations will be discussed in future work.

The log-likelihood function that we use, with a zero-noise realization, is a sum over the three independent channels  $A$ ,  $E$ ,  $T$  (27) rescaled according to (29):

$$\begin{aligned} \ln \mathcal{L} = & -\frac{1}{2}[(\tilde{a} - \tilde{a}_{\text{inj}}|\tilde{a} - \tilde{a}_{\text{inj}}) + (\tilde{e} - \tilde{e}_{\text{inj}}|\tilde{e} - \tilde{e}_{\text{inj}}) \\ & + (\tilde{t} - \tilde{t}_{\text{inj}}|\tilde{t} - \tilde{t}_{\text{inj}})], \end{aligned} \quad (51)$$

with the ‘‘inj’’ subscript indicating the simulated signals, and where the inner products  $(\cdot|\cdot)$  are given by (48) with the noise PSDs (33).

#### B. Reduced order model for EOBNRv2HM waveforms

As opposed to previous studies of Bayesian parameter recovery for LISA, we wish to use full inspiral-merger-ringdown signals. Since higher harmonics will play a crucial role in our analysis, in this study we use the EOB waveform model EOBNRv2HM [70]. These waveforms are based on the EOB formalism [71,72], and are calibrated to numerical relativity waveforms. The model is limited to nonspinning systems on quasicircular orbits, but includes a set of higher harmonics in the signal which are among the most important quantitatively:

$$(\ell, m) = (2, 2), (2, 1), (3, 3), (4, 4), (5, 5). \quad (52)$$

The waveforms are generated in the time domain by integrating a system of ordinary differential equations, an operation requiring up to seconds of computing time for long signals. For this reason, we developed a ROM for these waveforms, EOBNRv2HMRM, following the methods of [44] and enabling a much faster generation of the Fourier-domain amplitudes and phases of the modes in Eq. (18). As the total mass  $M$  scales out of the problem, the parameter space reduces to the mass ratio  $q = m_1/m_2$  only. The parameter space interpolation in this ROM is therefore significantly simpler than the models [44,45,73], which include aligned spins. The inclusion of higher harmonics, however, mandates modeling the relative dephasing of different modes. The output of the code consists of a Fourier-domain amplitude and phase for each mode, sparsely sampled on  $\sim 300$  frequencies. The unfaithfulness with the original waveforms, assuming advanced LIGO noise curves [74], is  $\lesssim 10^{-4}$ . The computational cost is submillisecond and smaller than other stages of our likelihood computation.

We note that the assumption of nonspinning black holes is an important limitation to a realistic assessment of intrinsic parameter uncertainties. In particular, it means that our results ignore the well-known degeneracy between mass ratio and spin [75–77], thus underestimating the uncertainty in the recovered mass ratio. By ignoring misaligned spins, we also neglect the possible effects of orbital precession on parameter degeneracies. Nevertheless, consistently including merger, ringdown, and higher harmonics already represents a significant improvement over previous studies, as these features in the signal carry a large amount of SNR [6,35–39]. In the absence of precession, we expect intrinsic and extrinsic parameters to be weakly



correlated, so that our investigations of the extrinsic part of the likelihood should be valid for aligned spins.

### C. Likelihood computation with zero noise

Since typical Bayesian analyses will need to evaluate millions of likelihood values, the computational performance of the likelihood function is crucial. In general, both the template and the injection in (49) are given as frequency series sampled with the Nyquist criterion  $\Delta f = 1/(2T)$ , where  $T$  is the duration of the signal and the overlaps (48) are simply computed as discrete sums over those frequency samples. Although transforming (48) in a discrete sum is a straightforward operation, the size of the frequency series can lead to a significant cost. In the context of LIGO/Virgo observations, this has prompted the development of reduced order quadratures (ROQs) [78–80] to accelerate the likelihood computation. In the LISA case, the computational cost of the standard likelihood implementation depends primarily on the mass of the system and the duration of the signal. This cost can range from tens of milliseconds for MBHBs and short-lived signals (a few days), to impracticable values for SBHB signals lasting for years and sampled at high frequency. By setting the noise realization to zero, however, we can represent the amplitude and phase of the signals over a sparse frequency grid and obtain a large speedup.

Applying the response as in Sec. II C, mode-by-mode on the sparsely sampled amplitude and phase generated as in Sec. III B, we obtain a complete representation of both the waveform and of the instrument transfer functions (30) on a few hundred points in the Fourier domain, with the full signal being implicitly reconstructed with a standard cubic spline interpolation over frequencies. Decomposing (49) as

$$\ln \mathcal{L} = (h(\boldsymbol{\theta})|h(\boldsymbol{\theta}_0)) - \frac{1}{2}(h(\boldsymbol{\theta})|h(\boldsymbol{\theta})) - \frac{1}{2}(h(\boldsymbol{\theta}_0)|h(\boldsymbol{\theta}_0)), \quad (53)$$

we have to compute inner products of the form  $(h(\boldsymbol{\theta})|h(\boldsymbol{\theta}'))$ . Note that individual terms in (53) can be individually large (of the order of  $\text{SNR}^2/2$ ), and computing  $\ln \mathcal{L}$  relies on accurate cancellations between terms.

Taking the likelihood (51) decomposed in TDI channels, decomposing further in harmonics as in (30), we have a sum of terms with the structure (symbolically)

$$\sum_{\text{chan.}} \sum_{\ell m, \ell' m'} \int \frac{df}{S_n} \mathcal{T}_{1,\text{chan}}^{\ell m} \mathcal{T}_{2,\text{chan}}^{\ell' m'} A_1^{\ell m} A_2^{\ell' m'} e^{-i(\Psi_1^{\ell m} - \Psi_2^{\ell' m'})}, \quad (54)$$

with transfer functions  $\mathcal{T}$  for each TDI channel, and Fourier-domain amplitudes  $A$  and phases  $\Psi$  [see (18)].

The difficulty of numerically computing such an overlap depends on the phase difference. A large phase difference

causes the integrand to be very oscillatory, requiring more frequency resolution than nonoscillatory integrands. On the other hand, oscillatory integrands result in a small integral due to cancellation effects, while nonoscillatory integrands contribute the most. For a single mode, SNR terms of the type  $(h|h)$  have a zero phase difference by construction, but including different modes generates cross-terms,  $(\ell, m) \neq (\ell', m')$ , with a large phase difference.

Note that in a Bayesian analysis, a large fraction of the time (after burn-in) will be spent exploring signals that are rather close to the injection with single mode phase differences that will be mostly small. In this work, we wish to include oscillatory cross-terms between modes, and we use a generic numerical treatment for the integrals (53) and keep all terms with no further approximation. Other methods to accelerate likelihoods, applicable in the presence of noise, include heterodyning with a reference signal [81], using a variable frequency resolution (multibanding) [82,83], and, as discussed before, ROQs [78–80]. We leave for future work the generalization of our likelihood computations to accommodate for noise.

First, we build a joint sparse frequency grid suitable to represent both signals and the transfer functions. We then resample the integrand on this grid, separating prefactors [noise PSD, transfer function without the Doppler phase (36), and signal amplitude] from phases (Doppler phases and signal phase difference). We then build a cubic spline for the prefactor, and a quadratic spline for the phase. We obtain the structure

$$\sum_{j \in \text{grid}} \int_{f_j}^{f_{j+1}} df P_j(f) e^{i\phi_j(f)}, \quad (55)$$

where for each grid interval  $P_j$  is a cubic polynomial and  $\phi_j$  is a quadratic polynomial. We can then compute the elementary integrals (55) using a combination of numerical methods. Integrations by parts can reduce the polynomial order in (55), leaving only Fresnel integral functions to be evaluated. Numerical instabilities can occur in these integrations by parts when the coefficients of the phase polynomial are very small or very large, in which case we resort to asymptotic expansions instead.

We combine this implementation of overlap computations with our fast waveform generation in Sec. III B and fast treatment of the Fourier-domain response in Sec. II C. Since the number of overlaps to compute in (30) is quadratic in the number of modes, likelihoods with higher harmonics are significantly more expensive. The final likelihood cost depends also on the number of grid intervals that is chosen to limit spline interpolation errors (Table I). The accuracy of our likelihood computations is typically  $|\Delta \ln \mathcal{L}| \lesssim 0.2$ , due to the numerical errors being magnified by the required cancellations between the different terms in (53).

TABLE I. Approximate frequency grid resolution and likelihood cost for the different choices of source type and the number of modes in the waveform.

Source type	Modes	Grid	Cost
SBHB	(2, 2)	500	2 ms
MBHB	(2, 2)	300	1.2 ms
MBHB	(2, 2), (2, 1), (3, 3), (4, 4), (5, 5)	300	15 ms

#### D. Bayesian sampling

For this exploratory work we take a “brute force” approach to inferring the posterior probability distribution of the parameters, without applying specific knowledge or expectations about distributions and degeneracies. Such assumptions might, in fact, lead to an incomplete exploration of the posterior probability in ways that would be difficult to recognize and diagnose. In order to demonstrate that the problem is tractable without critically depending on the details of the methodology, we also use two independent approaches to sample the posterior distributions: parallel tempering MCMC methods and nested sampling.

Our MCMC code `ptmcmc` [84] has been developed through several astrophysics projects, tested on several sampling problems, and modularly designed with the aim of minimizing opportunities for errors with new applications. The code performs parallel tempered MCMC [85,86] with temperatures slowly adjusted to achieve nearly equal exchange rates among all pairs of adjacent temperatures. Most runs here performed comparably with either 80 or 240 parallel temperature chains. We applied a general-purpose proposal distribution composed of a weighted set of subproposals, including differential-evolution steps [87,88] (although the proposal in this case is based on each single chain, not an ensemble), and a collection of several Gaussian step draws of varying sizes scaled off the prior domain at about 0.01–1% of its scale and weighted to favor the smaller scale proposal draws. Effective sample sizes (ESS) were estimated from the number of post-burn-in samples and reduced by the autocorrelation length, with the burn-in size chosen to maximize the ESS. In most cases, the runs were continued until achieving  $ESS > 2000$ .

For comparison and verification we also computed posterior samples using the nested sampling code `bambi` [89], a variant<sup>4</sup> of `Multinest` [27]. The nested sampling algorithm [90] evolves a set of “live points” toward regions of high probability by iteratively resampling the lowest probability point from within an ever-narrowing region covering the set of live-point samples; in `Multinest`, this region is built as a collection of overlapping ellipsoids.

<sup>4</sup>`bambi` comes with the additional option to train a neural network to learn the likelihood, but this was not used in the present study.

The computation proceeds until the estimated Bayesian evidence, or marginal probability, within the region covered by the live points is smaller than some threshold. Posterior-distributed samples are then drawn based on the set of sub-regions and the sample posterior values thereby generated. In our runs we used 4000 live points.

#### E. Fisher matrix parameter estimation

For comparison with our Bayesian inference results we also compute estimates of parameter uncertainties using the Fisher information matrix, which, despite its limitations, has been the most common workhorse of LISA science studies to date [91]. For measurements with additive Gaussian noise, as assumed here, the Fisher information matrix can be computed by

$$F_{ij} = (\partial_i h | \partial_j h), \quad (56)$$

where  $\partial_i$  is the derivative with respect to component  $i$  of the parameter vector  $\theta$ . In the Fisher matrix approach, the likelihood is approximated as

$$\ln \mathcal{L} \simeq -\frac{1}{2} F_{ij} \Delta \theta^i \Delta \theta^j, \quad (57)$$

where  $\Delta \theta^i$  is the parameter deviation from the true signal. The inverse of the Fisher matrix,  $\Sigma = F^{-1}$ , is the Gaussian covariance matrix that can be used as an estimate of the true parameter uncertainties.

We compute the derivatives by second-order finite differences, with the signal inner-products expanded and computed before differencing. For these calculations, where numerical smoothness is more of a concern than speed, we compute the inner products (48) explicitly on a fine grid. Generally, finite-difference Fisher matrices can be problematic, with small numerical defects potentially having outsized impact, so the finite-difference step-size  $\epsilon_i$  used in each parameter derivative  $\partial_i h$  must be chosen carefully. We target the step-size to be scaled off the diagonal Fisher matrix elements by  $\epsilon_i \approx \delta / \sqrt{F_{ii}}$  using  $\delta = 0.001$  for the results shown here. To achieve this, we begin with an initial choice of  $\epsilon_i$  scaled off the parameter values or prior widths, then we estimate the Fisher diagonals and iterate until convergence. In this process we also impose the constraint that the step is never larger than  $10^{-10}$  times the initial scaling.

## IV. MASSIVE BLACK HOLES SIGNALS

### A. Signals and transfer functions

We focus our analysis on a single choice of mass parameters, representative of canonical MBHB sources expected for LISA [5,6,92]: we pick a total redshifted mass,  $M = m_1 + m_2 = 2 \times 10^6 M_\odot$ , and a mass ratio,  $q = m_1/m_2 = 3$ , and place the source at the redshift

TABLE II. Parameters of the simulated MBHB mergers. Angles are given in the SSB frame. The SNR labelled  $h_{\ell m}$  is obtained including all harmonics listed in (52).

Identifier	I	II
Mass 1 ( $M_{\odot}$ )	$1.5 \times 10^6$	
Mass 2 ( $M_{\odot}$ )	$0.5 \times 10^6$	
Source-frame Mass 1 ( $M_{\odot}$ )	$3 \times 10^5$	
Source-frame Mass 2 ( $M_{\odot}$ )	$1 \times 10^5$	
Redshift	4.	
Lum. Distance (Mpc)	36594.3	
Inclination (rad)		$\pi/3$
Phase (rad)	2.140	-1.249
Ecliptic longitude (rad)	3.335	2.275
Ecliptic latitude (rad)	1.468	-1.376
Polarization (rad)	2.237	1.635
LISA SNR $h_{22}$	857.4	645.7
LISA SNR $h_{\ell m}$	944.8	666.0

$z = 4$ .<sup>5</sup> The corresponding source-frame total mass is  $M/(1+z) = 4 \times 10^5 M_{\odot}$ . The inclination plays an important role in deciding the importance of higher harmonics; we set it to the value  $\iota = \pi/3$ . We will study two such systems in detail, named system I and system II, stemming from a series of twelve runs with randomized orientation parameters, and chosen to exemplify qualitative differences in the parameter estimation. The parameters are summarized in Table II, along with the SNR of these signals with and without higher harmonics. The randomized orientations and the inclusion of higher harmonics both have a sizeable impact on the SNR.

We note that, contrasted with the current ground-based observations that are horizon-limited and that have a strong selection bias towards systems with favorable orientations (face-on or face-off), the LISA horizon for MBHB systems extends to essentially the entirety of the observable universe [5]. It is therefore natural to consider randomized orientations. Here, although we focus on just two systems, we take  $\iota = \pi/3$ , the median value when inclination is randomized.

We illustrate the Fourier-domain signal in Fig. 1, with mode-by-mode contributions to the characteristic strain, as well as the characteristic noise PSD, defined to be

$$\tilde{h}_{a,e,t}^c(f) = 2f\tilde{h}_{a,e,t}(f), \quad (58a)$$

$$S_c^{a,e,t}(f) = fS_h^{a,e,t}(f), \quad (58b)$$

with the strainlike observables and noise PSD defined in (31) and (34). We only show the  $A$  TDI channel, as the  $E$  channel is qualitatively similar and the  $T$  channel is negligible at low frequencies. Although the individual

<sup>5</sup>We convert between luminosity distances and redshifts using the cosmological parameters of [93].

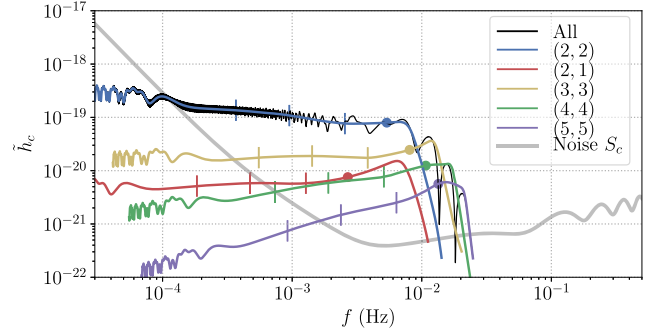


FIG. 1. Characteristic strain signal (58), shown here for the TDI channel  $A$  of system II, decomposed as a set of  $(\ell, m)$  contributions from the modes available in the waveform model. The gray curve shows the characteristic noise PSD (58). Vertical bars indicate for each mode the frequencies corresponding to the time-to-merger cuts marking the points where  $(1/64, 1/16, 1/4)$  of the final SNR has accumulated as in our premerger analysis as described in Sec. IV B, while the dots mark the merger frequency.

harmonics have fairly smooth amplitudes as a function of frequency, the full signal shows strong oscillations caused by the beating between the harmonics. The effect of the LISA response can be seen in the amplitude oscillations at low frequency, which are caused by the modulation resulting from the LISA motion.

Fig. 2 focuses on the details of the transfer functions. The Doppler phase variation  $\Delta\Phi_R$  (36) is a small effect at low frequencies, being suppressed by a factor of  $\sim 2\pi fR$ ; it is also small at higher frequencies, because the chirp happens over a short time interval during which  $\Phi_R$  does not have time to vary. We also show transfer functions for the mode  $h_{22}$  in the three TDI channels  $A, E, T$  ( $T$  is suppressed at low frequencies), after factoring out the Doppler phase. The low-frequency features show the time-dependency of the response created by the LISA motion, while frequency-dependency in the response appears at high frequencies. Section IV C will further investigate these different physical effects in the response.

## B. Accumulation of signal with time

We now illustrate the accumulation of the gravitational wave signal with time, which is crucial to understanding the ability of the instrument to identify and localize MBHB signals in advance of their coalescence, as well as the requirements put on the instrumental configuration (data gaps and downlink cadence). As will be shown in more detail in a separate publication [94] (see also [95]), within the uncertainties of astrophysical models we can expect the bulk of MBHB signals to be detectable only a few days prior to merger, while a tail of more favorable events could be observable for a significantly longer time, up to months.

We will focus on system II of Table II, and we will highlight four different epochs, corresponding to the time before the merger where a certain fraction of the total

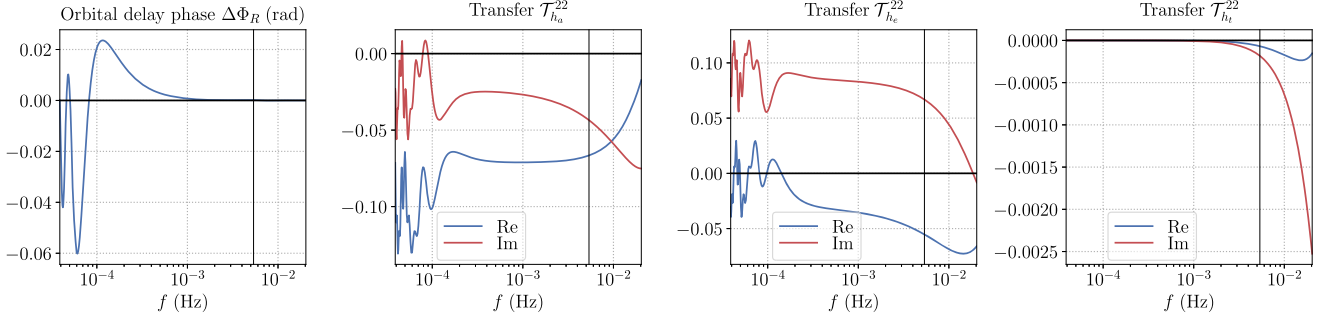


FIG. 2. Example of instrument response for system II in Table II. The first panel isolates the Doppler phase  $\Delta\Phi_R$  as defined in (36), while the three other panels show the TDI transfer functions for  $A$ ,  $E$ ,  $T$ . We show here the rescaled transfer functions for the 22-mode as defined below (30). The vertical lines correspond to the frequency at merger.

SNR = 666 has been accumulated. We include higher harmonics here. These epochs are:

- (i) SNR/64:  $\sim 42$  hr before the merger, which corresponds roughly with SNR  $\simeq 10$ , the first time we could confidently claim a detection;
- (ii) SNR/16:  $\sim 2.5$  hr before the merger;
- (iii) SNR/4:  $\sim 7$  min before the merger.

We approximate the abrupt interruption of the observation as an upper frequency cutoff. The cutoff frequency is derived from the end time of the observation via the time-to-frequency correspondence (22). In reality, an abrupt interruption of the signal would produce nonlocal features in the Fourier domain (unless tapering is applied), which we do not consider here.

The time-to-frequency correspondence (22) gives time as a function of frequency for each  $(\ell, m)$  mode. Due to the scaling  $h_{\ell m} \propto e^{-im\phi_{\text{orb}}}$  with  $\phi_{\text{orb}}$  as the orbital phase, a given time corresponds to different frequencies according to

$$\omega_{\ell m}(t) \simeq \frac{m}{2} \times \omega_{22}(t) \quad (\text{inspiral}), \quad (59a)$$

$$t_f^{\ell m} \left( \frac{m}{2} f \right) \simeq t_f^{22}(f) \quad (\text{inspiral}), \quad (59b)$$

where  $\omega_{\ell m} = \dot{\phi}_{\ell m}$  is the instantaneous mode frequency in the time domain. Figure 3 illustrates these relations and we use them to mark in Fig. 1 the frequencies corresponding to our time cuts that differ for each mode, together with the merger frequency (also called the peak frequency).

The relations (59) are only accurate for the inspiral regime, as such a correspondence is at the heart of the SPA. Although  $t_f^{\ell m}$  (22) can be formally extended even past the merger time, close to the merger it starts to lose accuracy and physical interpretation, eventually becoming nonmonotonic with the frequency [16]. The departure from the scaling (59) between modes can be seen in Fig. 3: the scaling holds up to a few minutes before the merger, and the mode  $h_{21}$  shows the earliest signs of a deviation.

Figure 4 shows the cumulative contributions to the total SNR of individual mode combinations. For a generic signal  $s = \sum_{\ell m} s_{\ell m}$  with mode contributions  $s_{\ell m}$  (where  $s$  stands for any of the channels  $\tilde{a}$ ,  $\tilde{e}$ ,  $\tilde{i}$ ),

$$\text{SNR}^2 = (s|s) = \sum_{\ell m} \sum_{\ell' m'} (s_{\ell m} | s_{\ell' m'}). \quad (60)$$

We choose to show contributions to  $\text{SNR}^2$  because this is the scale on which the contributions of different modes and channels are additive, and because it is the relevant scale for the log-likelihood (49), since  $\ln \mathcal{L} \sim \text{SNR}^2$ . We also sum over the three channels. Together with the total contribution to  $\text{SNR}^2$ , the three panels of Fig. 4 show diagonal terms  $(s_{\ell m} | s_{\ell m})$ , cross-terms involving the dominant quadrupolar mode  $(s_{22} | s_{\ell m})$ , and finally other cross-terms between subdominant modes. Results are displayed as a function of the time-to-merger, and we show separately the result obtained for the full postmerger signal.

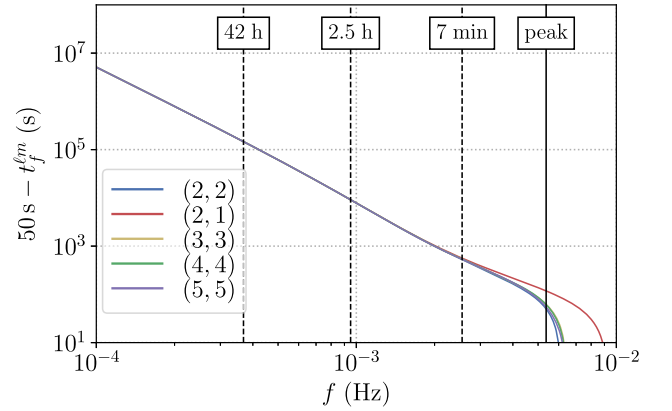


FIG. 3. Relation between time and frequency for MBHB systems I and II with  $M = 2 \times 10^6 M_{\odot}$ ,  $q = 3$ , and  $z = 4$ . The different lines correspond to the  $(\ell, m)$  harmonics available in the waveform model. The time-to-frequency correspondence is computed according to  $t_f^{\ell m}$  defined in (22), evaluated at  $mf/2$  to map to the same time according to (59). The vertical lines represent the time cuts explained in Sec. IV B and the merger frequency.



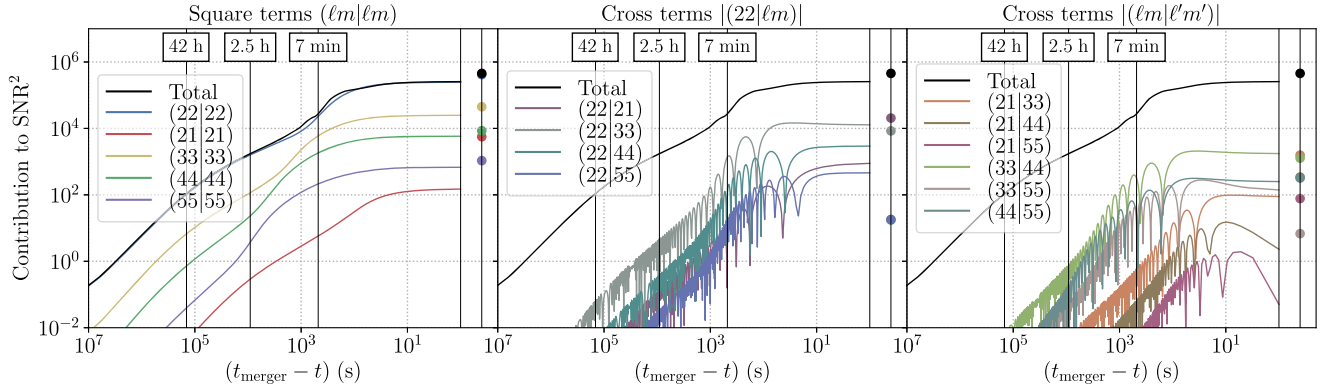


FIG. 4. Cumulative contribution to  $\text{SNR}^2$  as defined in (60) for system II. The left panel shows inner products of modes with themselves, the middle panel shows cross-terms involving  $h_{22}$ , while the right panel shows other cross-terms. Note that oscillatory cross-terms change sign, we show their absolute value. In all panels the black line represents the total. The horizontal axis gives the time-to-merger, with vertical lines for the cuts of Sec. IV B, and the dots on the right of each panel show the result for the full signal including the ringdown.

We can draw several conclusions from Fig. 4. The signal reaches a roughly detectable level of  $\text{SNR} = 10$  about two days before the merger. We see that the SNR accumulates rapidly in the last instant before the merger, as shown in particular by reaching  $\text{SNR}/4$  only seven minutes before coalescence. For a period following the first detection most or all of the higher modes are not significant. Stopping the signal at a given time in the inspiral somewhat alters the hierarchy between subdominant modes that was seen in their power spectra. Contrasting Fig. 4 with Fig. 1, we see, e.g., that  $h_{21}$  is now subdominant before the merger (and increases significantly in strength when including the postmerger signal). This is because the modes  $h_{\ell m}$  with a higher  $m$  will reach a higher frequency at any given time, while the weighting by the noise favors higher frequencies. This is also visible in Fig. 1, where ticks translate a cut in time into a different cut in frequency for different modes.

We also see that since even the most subdominant modes contribute significantly to  $\text{SNR}^2$ , our set  $(22, 21, 33, 44, 55)$  seems to be incomplete; we will need waveforms with a richer set of higher harmonics to analyze real LISA data. Diagonal-terms  $(s_{\ell m}|s_{\ell m})$  and cross-terms with  $(\ell m) \neq (\ell' m')$  are qualitatively different. Diagonal-terms accumulate coherently, while cross-terms are oscillatory, as they feature two modes with different phasings. This oscillatory character tends to suppress the contribution of the cross-terms, however, we see they are not negligible, particularly the ones involving the dominant harmonic and a subdominant harmonic.

### C. Decomposing the instrument response

The LISA instrument response, as recalled in Sec. II, is both time- and frequency-dependent. As we will see, these features play an important role in breaking degeneracies in the parameter estimation of the source, notably allowing us to localize the source in the sky. It is therefore important to

understand these features, also in the light of the premerger accumulation of signal with time.

The time-dependency in the response follows the motion of the LISA constellation. In the low-frequency picture of Sec. II E, the time-dependency enters both in the Doppler phase term (36) (LISA moves in the wavefront) and in the time-dependent LISA-frame angles (A15) (the LISA arms change orientation). However, we have seen in the previous section that the MBHB signals we consider here are quite short (less than two days). This limits the effect of the LISA motion; between the time where the signal becomes detectable and the merger, LISA has barely changed in orientation and position. Moreover, the period when the signal accumulates most of its SNR is even shorter, as shown in Fig. 4.

To disentangle these different physical effects, we distinguish four different approximations of the response:

- (i) *Full*: in this case, we keep the full response of Secs. II C–II D, with its complete time- and frequency-dependency;
- (ii) *Frozen*: we neglect the LISA motion by evaluating all time-dependent vectors at  $t_f = t_f^{\text{peak}}$ , effectively freezing LISA in its orbit while keeping the frequency-dependency in the response;
- (iii) *Low- $f$* : we implement the low-frequency (long-wavelength) approximation as in Sec. II E, while still keeping the time-dependency due to the motion;
- (iv) *Frozen low- $f$* : we neglect both the time- and frequency-dependency in the response.

Case (iii), *Low- $f$* , has been extensively used in the past for the analysis of MBHB signals. Since it is based on the  $f \ll f_L$  approximation, it is more appropriate for inspiral-only signals (to which most of the previous studies were limited) than for a full IMR signal.

In case (iv), *Frozen low- $f$* , the response is equivalent to two LIGO-type detectors, lying motionless in the same location and rotated from each other by  $\pi/4$ , and there is

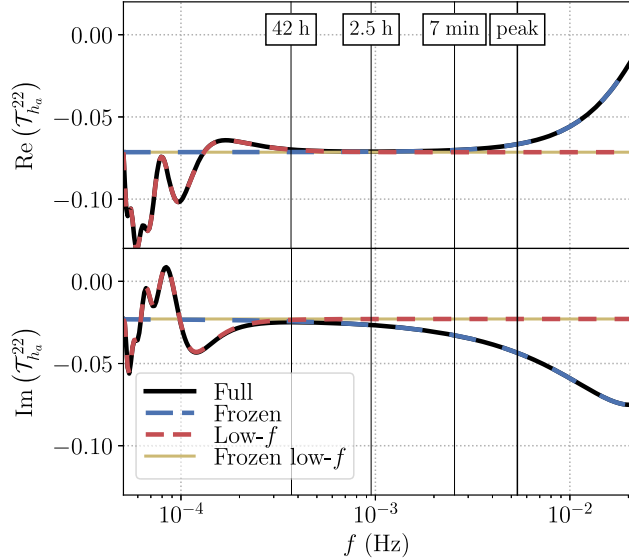


FIG. 5. Real and imaginary parts of the transfer function  $\mathcal{T}_{h_a}^{22}$  (31) in different response approximations. The four lines stand for the four approximation levels (i)–(iv) defined in Sec. IV C. Vertical lines indicate the frequencies of the time cuts of Sec. IV B and the peak frequency.

no other information on the source’s sky position other than the frequency-independent pattern functions of the two effective detectors. Contrary to networks of ground-based observatories, we have no triangulation information from times of arrival at different detectors. This case is the most degenerate, and it will be useful to get an analytical understanding of approximate degeneracies that occur when using the more complicated full response [case (i)]. This limit can also be a representative approximation for some short-duration premerger LISA MBHB observations.

The effect on the response of adopting these approximations is shown in Fig. 5. We display the strain transfer function  $\mathcal{T}_{h_a}^{22}(f)$  as defined in (31) for the mode  $h_{22}$ , with the four lines showing the four approximations [(i)–(iv)]. Vertical lines also indicate the times to merger highlighted in Sec. IV B, as well as the peak frequency. Relating this figure to Figs. 1 and 4, we stress that most of the SNR is accumulated in the very last instant before the merger and at the merger itself, due to the noise normalization not visible at the level of the transfer function.

The *Frozen low- $f$*  transfer function is just a constant factor, similar to the LIGO response where  $h_{+, \times}$  are simply multiplied by pattern functions. The *Low- $f$*  transfer function goes to the same constant at high frequencies, where the signal chirps so rapidly that the LISA motion is negligible. At low frequencies, however, modulations due to the LISA motion appear. The *Frozen* transfer function asymptotes to the constant of the *Frozen low- $f$*  case at low frequencies, for  $f \ll f_L$ . At higher frequencies, we see a growing departure from this approximation. As noted in Sec. II E, when reaching the arm length transfer

frequency  $f_L = 0.12$  Hz, the long-wavelength approximation has completely broken down; departures start to be significant at much lower frequencies. Finally, the *Full* transfer function displays all the features we discussed.

We finally note a coincidence in Fig. 5: the frequency below which we see the imprint of the LISA motion and the frequency above which we see the breakdown of the long-wavelength approximation appear to be the same. This will not be true in general—the former is essentially a measure of the time-to-frequency correspondence  $t_f$ , with lower-mass signals having support and showing these features at higher frequencies, while the latter only marks the magnitude of  $2\pi fL$  factors and is source-independent.<sup>6</sup>

## V. MASSIVE BLACK HOLES PARAMETER ESTIMATION

### A. Analysis of IMR signals

Using the methodology summarized in Sec. III, here we present the results of Bayesian parameter estimation for the two MBHB sources listed in Table II. The priors used are logarithmic in mass, flat in luminosity distance, and uniform for the angles: flat on the sphere for the pairs of angles  $(\iota, \varphi)$  and  $(\lambda, \beta)$ , and flat for the polarization  $\psi$ . We expect priors to be unimportant for the masses, since they are well-determined in this very high SNR limit. The luminosity distance, as we will see, can be less well-determined in the absence of higher harmonics, and one should keep in mind that the prior choice does affect the posterior there.

We find that both our samplers, `ptmcmc` and `Multinest`, require between  $\sim 10^7$  and  $\sim 10^8$  likelihood evaluations to produce a final set of posterior samples. Thanks to our fast likelihood implementation, this already represents a manageable computing cost. We stress again that the settings of both samplers were not optimized by taking into account the characteristics of the specific problem or the expected correlations between parameters. Therefore, we expect that the cost can be reduced with future optimizations.

To illustrate the role of higher harmonics in the analysis, we will present two classes of results: “22”, where we inject and recover signals including only the dominant harmonic  $h_{22}$ , and “HM”, where we both inject and recover with the higher harmonics  $(h_{22}, h_{21}, h_{33}, h_{44}, h_{55})$ . We do not attempt to recover an injection with higher modes using 22-mode waveforms; one would expect parameter biases in that case, induced by the inadequacy of the simplified waveforms. A related question would be to assess the mode content necessary for waveform models to mitigate systematic biases when analyzing full physical signals. We leave such waveform systematic studies for future work; we only note that, as shown in Fig. 4, from SNR only the set of

<sup>6</sup>The effect in the transfer functions is source-independent; however, the total SNR will determine the impact of these features on the parameter estimation.

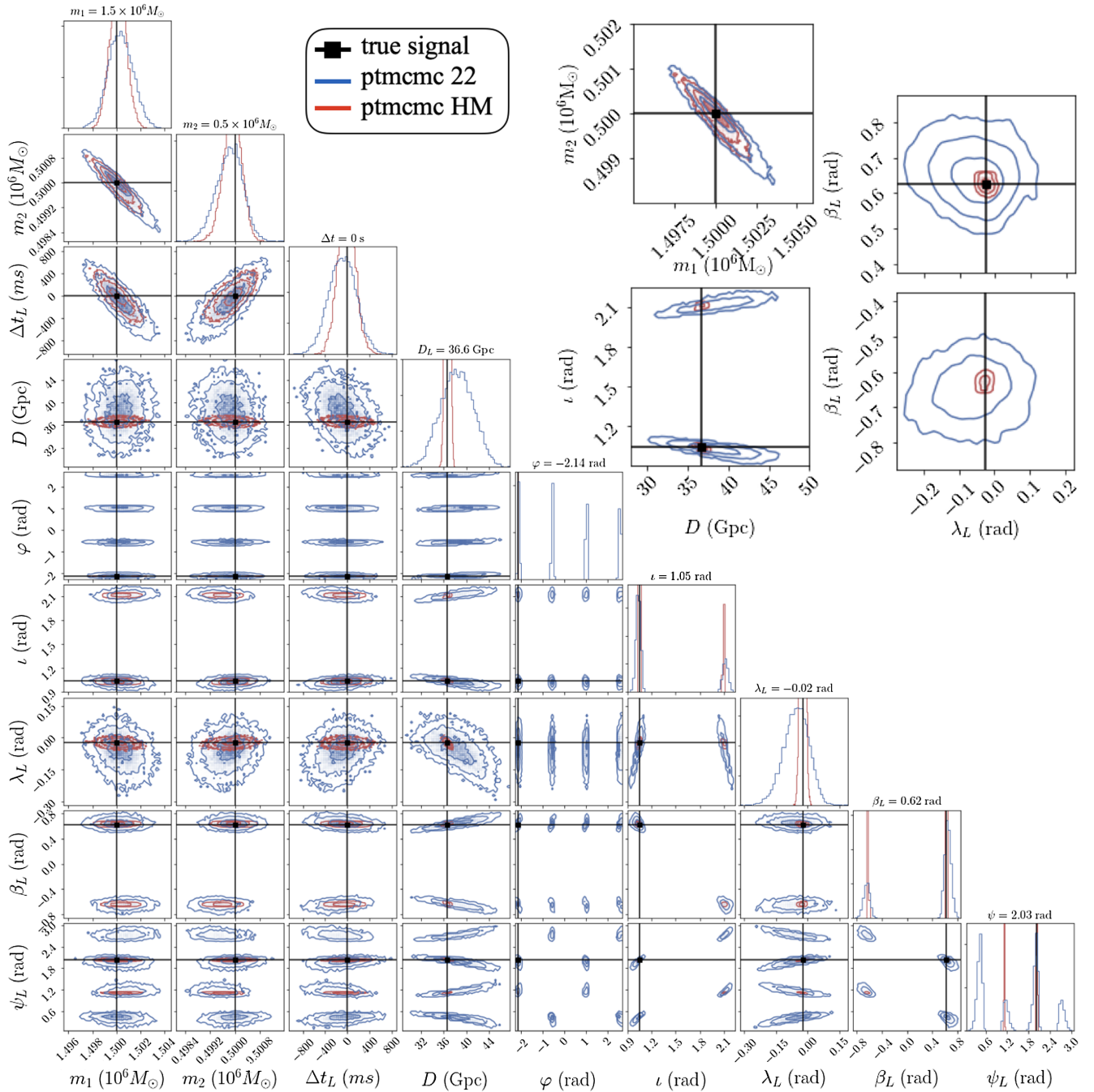


FIG. 6. Inferred parameter posterior distribution for the MBHB system I of Table II, with 1-2-3- $\sigma$  contours and the black cross indicating the true parameters. The result injecting and recovering with only the 22-mode is shown in blue, and the result injecting and recovering with higher harmonics is in red. The top right panels zoom on the masses, distance-inclination, and sky position marginalized posteriors. The time is centered so that  $\Delta t = 0$  corresponds to the injection. All extrinsic parameters are given in the LISA-frame as defined in Appendix A 3.

five modes we include in the present study is quite obviously incomplete.

While the 22-mode-only signals here are nonphysical, studying them is instructive not only because 22-only waveforms have been a common approximation in previous work but also because (as seen in Fig. 4) upon first

detection LISA MBHB signals will often be fairly approximated by 22-mode-only signals.

Figures 6 and 7 show posterior distributions for all parameters for cases I and II respectively. We overlay the “22” and “HM” posteriors. All results for MBHB systems are presented using LISA-frame parameters, as defined in



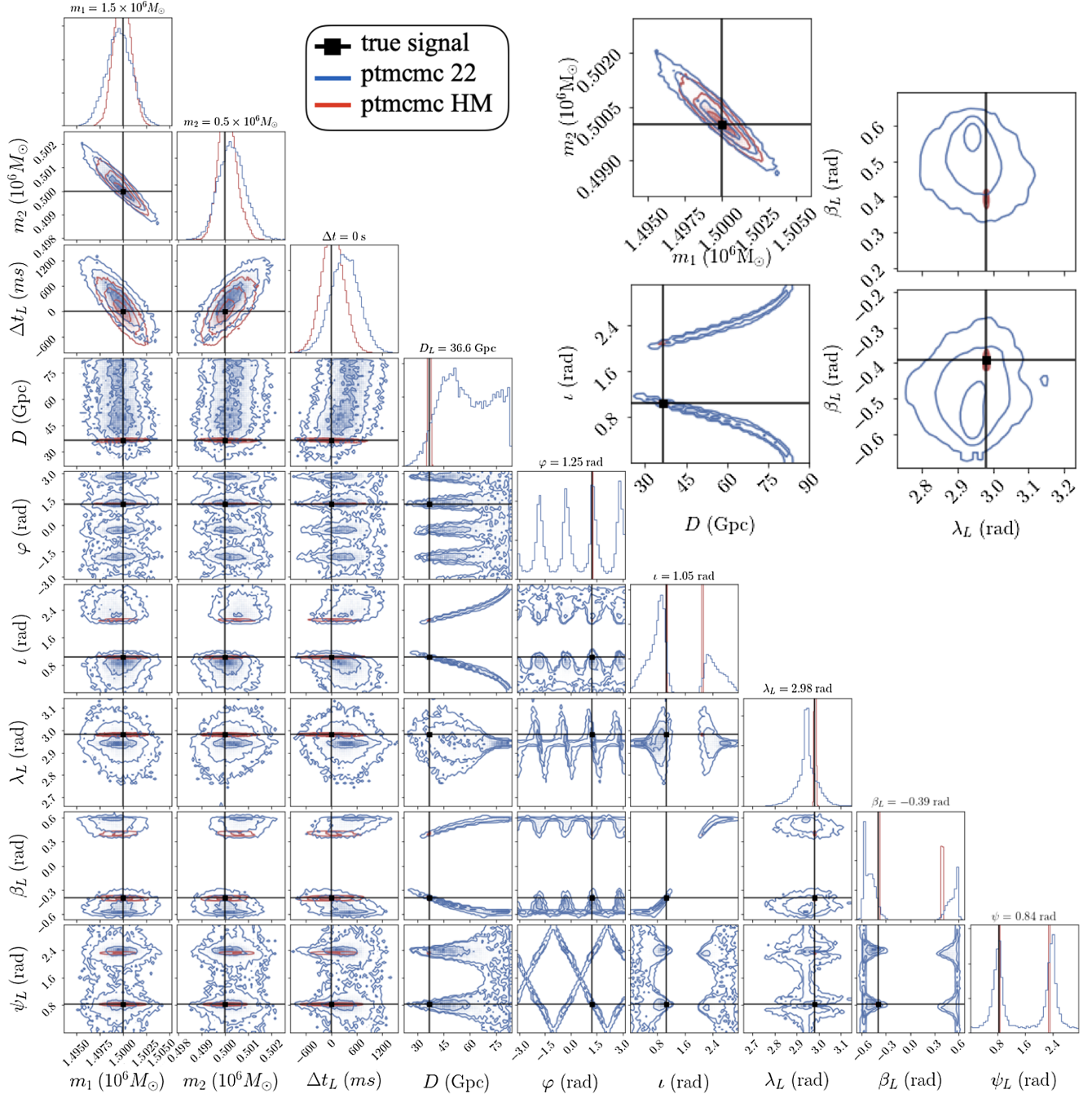


FIG. 7. Same as Fig. 6, for the MBHB system II of Table II. In the extrinsic parameters, the red 2D contours are barely visible on this scale.

Appendix A. We should keep in mind the difference in SNR of cases I and II (SNR  $\approx 945$  versus SNR  $\approx 666$  with higher harmonics, see Table II) resulting from their different orientation. However, their posteriors have qualitative differences that go beyond a simple scaling of the errors as  $\sim 1/\text{SNR}$ .

In case I, the “22” posterior appears to be well represented by a multimodal Gaussian, with multimodality for the angular parameters. The sky position, most notably,

admits a degenerate mode at  $\beta_L \rightarrow -\beta_L$ , although it has less weight than the main mode around the injection. By contrast, in case II, the “22” posterior is much more degenerate. The distance and inclination are very degenerate with each other, with a support extending all the way to  $\iota = 0$  (or  $\iota = \pi$ ). The phase  $\varphi$  and polarization  $\psi_L$  have a distinct extended degeneracy along lines of constant  $\varphi + \psi_L$  and  $\varphi - \psi_L$ . Most notably, the sky position, if retaining the same overall bimodality  $\beta_L \rightarrow -\beta_L$  as in case I, shows



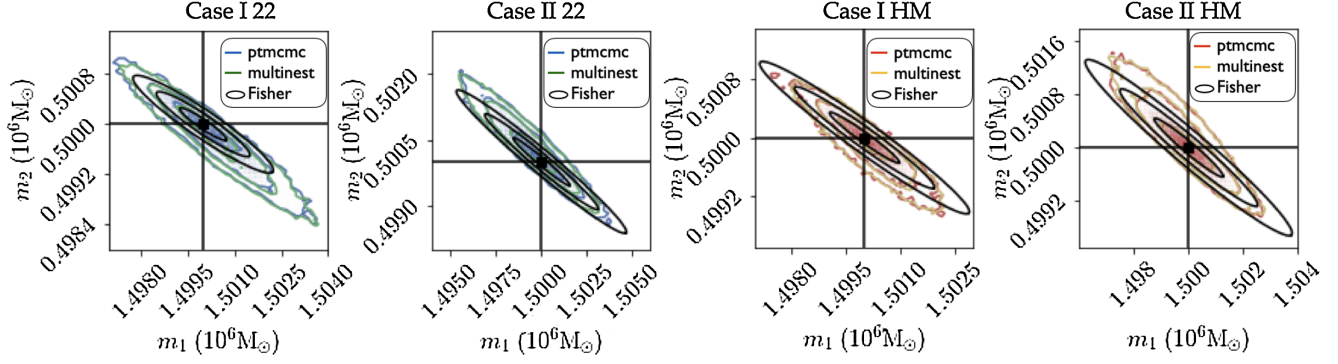


FIG. 8. Comparisons between inferred posteriors obtained with our two different Bayesian samplers described in Sec. III D and with the Fisher matrix approximation described in Sec. III E. Results are shown for systems I and II of Table II, with the 22-mode-only for the two left panels and with higher harmonics in the two right panels. Blue (22) and red (HM) show the ptmcmc result, and green (22) and yellow (HM) the multinest result. In all panels, the black ellipses show the Fisher matrix estimate for the posterior.

here a curious feature: the marginalized sky posterior peaks away from the injected value. Looking at correlations, we see that the shifted peak corresponds to the region of high distance (and extremal inclination). This is a genuine feature of the multidimensional posterior, even without noise, and will be explained in Sec. V C.

In both cases I and II, including the higher harmonics has a major effect on the parameter recovery (as was already stressed in [32–35]), much beyond what we would expect solely from the modest gain in total SNR shown in Table II (and a scaling of statistical errors  $\sim 1/\text{SNR}$ ). The marginalized posterior of the masses is narrower, although not qualitatively different; the major change is in the extrinsic parameters.

We can understand the dramatic effect of the higher harmonics on distance and inclination by noting that, in a signal of the form  $\sum_{\ell m} {}_{-2}Y_{\ell m}(t, \varphi) h_{\ell m}$  where the  ${}_{-2}Y_{\ell m}$

all have a different dependency with the inclination  $\iota$ , measuring independently two separate contributions,  $(\ell m) \neq (\ell' m')$ , gives us an independent measurement of  $\iota$  by the relative amplitude of the two mode contributions. When only the  $h_{22}$  mode is included, both the luminosity distance and inclination determine the overall signal amplitude and they are therefore degenerate.

Higher harmonics also lead to an independent determination of the phase  $\varphi$ , which in turn breaks degeneracies in the other angular parameters. In both cases I and II, the sky localization is vastly improved, with a remaining multimodality that we will discuss in the next section.

We compare the Bayesian inference results obtained with our two samplers, ptmcmc and multinest, with the error estimates, given by the Fisher matrix approximation from Sec. III E, in Fig. 8 for the masses, Fig. 9 for distance/inclination, and Fig. 10 for the sky position. We find a good

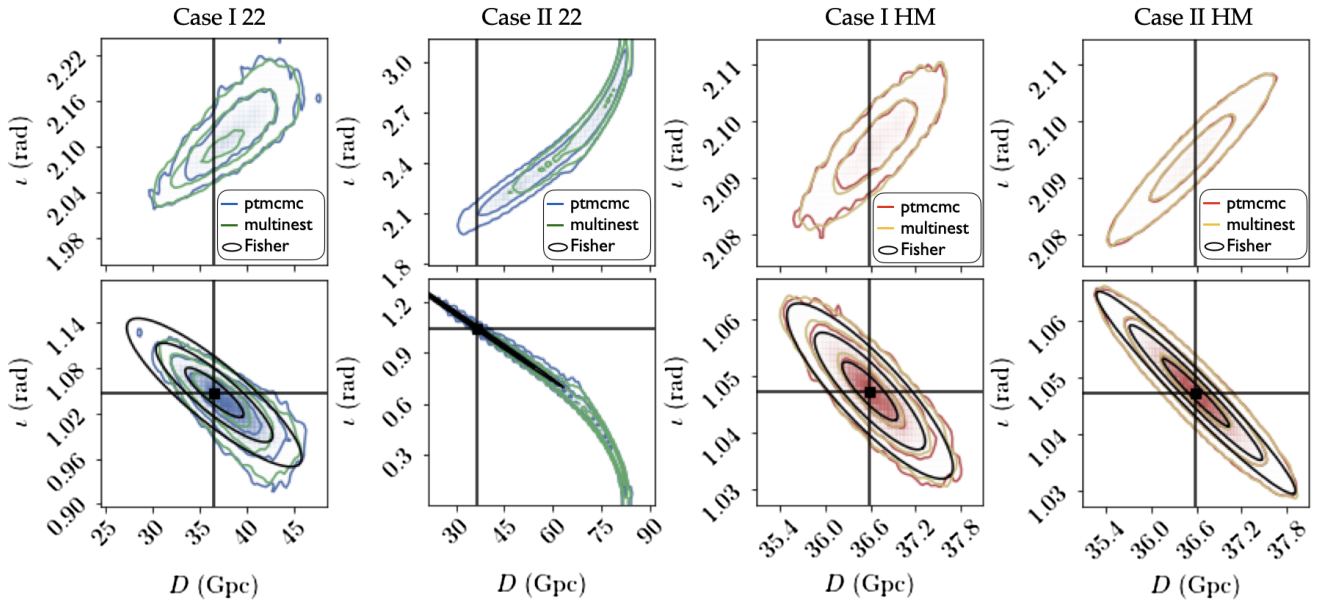


FIG. 9. Same comparisons as in Fig. 8, for the distance and inclination parameters, zooming on the relevant regions.

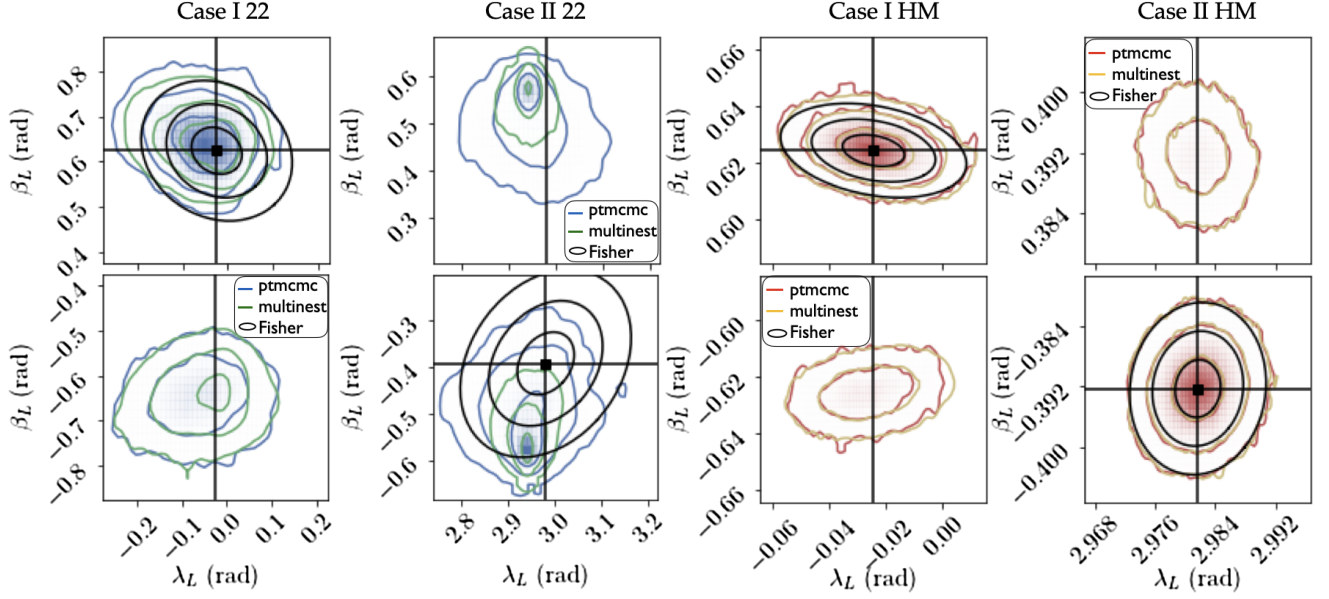


FIG. 10. Same comparisons as in Fig. 8, for the sky position parameters, zooming on the relevant regions for inclination. The parameters shown are in the LISA-frame as defined in Appendix A 3.

agreement between the two Bayesian samplers, except for in case II when using only 22-mode signals. There, MultiNest seems to fail to explore the full posterior, remaining stuck in the large-distance region of the very degenerate parameter space. We will further explore the nature of this degeneracy in Sec. V C and in Appendix B. The Fisher matrix approach focuses on the vicinity of the true signal’s parameters and is by construction unable to handle multimodality. It is also insufficient in capturing the degeneracy features of case II with  $h_{22}$  signals. It gives good results, however, for the mass parameters and for the main mode of the posterior in nondegenerate cases, particularly when including higher harmonics.

### B. Degeneracies in the sky

A remarkable feature in Figs. 6 and 7 is the existence of a degenerate mode (or secondary maximum) in the posterior distribution for the sky position  $(\lambda_L, \beta_L)$ , located at  $\beta_L^* = -\beta_L^{\text{inj}}$ , i.e., by reflecting the wave vector  $k$  across the plane of the LISA constellation. The secondary mode contains less probability than the “correct” mode close to the true parameters, but it is present in both simulations and it survives the inclusion of higher harmonics.

We can gain insight about degeneracies in the parameter space by considering the simple expressions representing the instrument response in the *Frozen low- $f$*  approximation described in Sec. IV C, appropriate for the low-frequency limit and signals short enough that the LISA motion can be neglected. We will explore the respective influence of these sometimes neglected effects in Sec. VI.

The *Frozen low- $f$*  response is given by Eqs. (43)–(45), ignoring the  $k \cdot p_0$  delay as a mere constant delay, and

neglecting time-dependency in the LISA frame angles  $(\lambda_L, \beta_L, \psi_L)$ . The response is then summarized by the pattern functions for harmonics  $F_{a,e}^{\ell m}(t, \varphi, \lambda_L, \beta_L, \psi_L)$  given in (44), which we reproduce here:

$$F_{a,e}^{\ell m} = \frac{1}{2} {}_{-2}Y_{\ell m}(t, \varphi) e^{-2i\psi_L} (F_{a,e}^+ + iF_{a,e}^{\times})(\lambda_L, \beta_L) + \frac{1}{2} (-1)^\ell {}_{-2}Y_{\ell, -m}^*(t, \varphi) e^{+2i\psi_L} (F_{a,e}^+ - iF_{a,e}^{\times})(\lambda_L, \beta_L). \quad (61)$$

Changing the sign of  $\beta_L$  in the pattern functions  $F_{a,e}^{+,\times}$  given in (43) has no effect on  $F_{a,e}^+$  and it changes the sign of  $F_{a,e}^{\times}$  for both channels. Thus, the  $(F_{a,e}^+ \pm iF_{a,e}^{\times})$  factors in the two terms of Eq. (61) are exchanged. Moreover, since spin-weighted spherical harmonics [51] obey the relation

$${}_{-2}Y_{\ell m}(\pi - \iota, \varphi) = (-1)^\ell {}_{-2}Y_{\ell, -m}^*(\iota, \varphi), \quad (62)$$

we see that simultaneously changing  $\beta_L \rightarrow -\beta_L$ ,  $\iota \rightarrow \pi - \iota$  and  $\psi_L \rightarrow \pi - \psi_L$  leaves  $F_{a,e}^{\ell m}$ ,  $F_e^{\ell m}$  unchanged.

This defines a transformation of extrinsic parameters yielding an exact degeneracy in the *Frozen low- $f$*  approximation, which we call the *reflected sky position* (for a reflection with respect to the LISA plane):

$$\begin{aligned} \lambda_L^* &= \lambda_L, \\ \beta_L^* &= -\beta_L, \\ \psi_L^* &= \pi - \psi_L, \\ \iota^* &= \pi - \iota, \\ \varphi^* &= \varphi, \end{aligned} \quad (63)$$

TABLE III. Sketch of the degeneracy structure of the eight modes in the sky, with the cases where the degeneracy is exact (degen.), and with the qualitative effects that break the degeneracy with the injected signal:  $t$ -dependency,  $f$ -dependency, both, or only the Doppler phase,  $\Delta\Phi_R$  (36). This qualitative structure is the same with and without higher harmonics.

Sky mode	<i>Full</i>	<i>Frozen</i>	<i>Low-f</i>	<i>Frozen low-f</i>
<i>reflected</i> : $-\beta_L, \lambda_L$	$t$ -dep.	degen.	$t$ -dep.	degen.
<i>antipodal</i> : $-\beta_L, \lambda_L + \pi$	$f$ -dep. + $\Delta\Phi_R$	$f$ -dep.	$\Delta\Phi_R$	degen.
$\beta_L, \lambda_L + \pi/2$	$t$ - $f$ -dep.	$f$ -dep.	$t$ -dep.	degen.
$\beta_L, \lambda_L + \pi$	$t$ - $f$ -dep.	$f$ -dep.	$t$ -dep.	degen.
$\beta_L, \lambda_L - \pi/2$	$t$ - $f$ -dep.	$f$ -dep.	$t$ -dep.	degen.
$-\beta_L, \lambda_L + \pi/2$	$t$ - $f$ -dep.	$f$ -dep.	$t$ -dep.	degen.
$-\beta_L, \lambda_L - \pi/2$	$t$ - $f$ -dep.	$f$ -dep.	$t$ -dep.	degen.

where we chose the transformation for  $\psi_L$  to keep this parameter in the range  $[0, \pi]$ .

From the structure of (61), other points in parameter space are degenerate in the low-frequency limit. First,  $\lambda_L \rightarrow \lambda_L + \pi$  leaves the pattern functions  $F_{a,e}^{+, \times}$  invariant. For  $\lambda_L \rightarrow \lambda_L \pm \pi/2$ , these pattern functions acquire an overall minus sign. Such an overall minus sign is readily compensated by a shift,  $\psi_L \rightarrow \psi_L \pm \pi/2$ , so that we have the other transformation,

$$\begin{aligned}\lambda_L^{(k)} &= \lambda_L + \frac{k\pi}{2} \pmod{2\pi}, \quad k = 0, \dots, 3, \\ \psi_L^{(k)} &= \psi_L + \frac{k\pi}{2} \pmod{\pi}, \quad k = 0, \dots, 3.\end{aligned}\quad (64)$$

Combining Eqs. (63) and (64), we arrive at eight different degenerate positions in the sky, equally spaced in  $\lambda_L$  and symmetric above and below the LISA plane, with an inclination  $\pi - \iota$  for the reflected positions and various values for the polarization  $\psi_L$ . Among these eight secondary modes in the sky, two play special roles: first, the *reflected* mode (63) already mentioned; and second, the *antipodal* mode with

$$\begin{aligned}\lambda_L^{(a)} &= \lambda_L + \pi, \\ \beta_L^{(a)} &= -\beta_L, \\ \psi_L^{(a)} &= \pi - \psi_L, \\ \iota^{(a)} &= \pi - \iota, \\ \varphi^{(a)} &= \varphi.\end{aligned}\quad (65)$$

How does this situation change when considering a more complete instrument response, moving away from the *Frozen low-f* approximation? As in Sec IVC we separately consider relaxing each of the qualifiers *Frozen* and *Low-f*.

When adding back the frequency-dependence while keeping LISA motionless, in the *Frozen* response, the *reflected* mode is the only mode remaining degenerate. Indeed, as noted in Sec. II C, in (21) the frequency-dependent terms feature  $k \cdot n_l$ ,  $k \cdot p_{s,r}$  projections that

depend only on the projection of the wave vector  $k$  in the plane of LISA, and are invariant for  $\beta_L \rightarrow -\beta_L$ . The LISA motion, however, will break this degeneracy in general.

When adding back the LISA motion while still ignoring the frequency-dependence in the response, in the *Low-f* response, the *antipodal* mode is the only one that keeps pattern functions that are exactly degenerate with the injection: all other modes will evolve with the time-dependent LISA frame. The antipodal mode is not moving, because the antipode of the true direction of the arriving signal is defined as  $k \rightarrow -k$  independently of the orientation of LISA; the only degeneracy-breaking term is then the orbital delay  $k \cdot p_0$  in the Doppler phase (36). When considering the *Full* response, the frequency-dependence terms in (21) featuring  $k \cdot n_l$ ,  $k \cdot p_{s,r}$  will break this antipodal degeneracy.

Thus, we have found that in the *Frozen low-f* approximation for the response we expect a pattern of eight degenerate positions in the sky (with certain rules for inclination and polarization). Table III summarizes the qualitative picture of degeneracy breaking by the features of the response. We note that the recent work [41], focusing on low-frequency ringdown-dominated signals for which the LISA motion can be neglected, remarked the same eight-mode sky degeneracy that we described in this section. We will see in Sec. VI that the eight-mode degeneracy pattern indeed appears when doing a premerger analysis. On the other hand, in the results of Figs. 6 and 7 for full IMR signals, out of the eight possible sky modes only the *reflected* mode (63) survives. We will investigate in detail in Sec. VI how time-dependence and frequency-dependence in the response break part, but not all, of the degeneracies.

However, this analysis does not explain why one obtains, with a zero-noise realization, marginalized posteriors for the sky positions that appear biased from the injected signal, when ignoring higher harmonics. This is the question that we will address in the next section.

### C. Apparent sky position bias for 22-mode signals

In this section, we investigate the cause of the apparent bias in sky position in the posterior distribution of system II

when injecting and recovering with 22-mode-only waveforms, as shown in Fig. 7 (and in the middle left panel of Fig. 10). The posterior for the sky forms a peak that appears shifted from the injection. This feature is surprising since we set the noise realization to zero and the maximum likelihood is, by construction, reached at the injection. It occurs only when 22-mode-only signals are used, and is present for both Bayesian samplers, although `ptmcmc` and `multineest` differ noticeably for this case with `ptmcmc` recovering more parameter volume. We will show that we can understand this feature as a projection effect in the multidimensional degenerate posterior. Understanding the structure of these features could prove useful to inform Bayesian samplers, for instance by adapting jump proposals to speed up mixing of MCMC chains.

In the following we will aim at explaining these features analytically by building a simplified model for the response. In this *simplified likelihood* approximation, we:

- (i) pin the masses and the LISA-frame coalescence time to their injected values;
- (ii) use the *Frozen low- $f$*  response, ignoring the LISA motion and the frequency-dependence in the response.

The first point allows us to focus only on the extrinsic parameters, as we find weak correlations between intrinsic and extrinsic parameters. Such a decoupling is also at play in LIGO/Virgo, and makes low-latency sky localization possible [96]. The second point is partly justified by the fact that our signal is short, as shown in Fig. 4. For the MBHB system that we picked as an example, the SNR accumulates in a matter of days, with SNR = 10 reached  $\sim 40$  h before the merger. Neglecting the high-frequency features is in fact a stronger approximation, as will be explored in Sec. VI B. In this limit we can apply the simple analytic expressions for the response given in Sec. II E.

Under all these simplifying assumptions, the likelihood becomes pure function of the extrinsic parameters ( $D, \iota, \varphi, \lambda_L, \beta_L, \psi_L$ ), and takes the form of trivial geometric factors multiplying constant mode overlaps that can be precomputed. Equation (37) shows that the  $T$ -channel is negligible in this approximation, so that the likelihood (51) is

$$\ln \mathcal{L} = -\frac{1}{2}(\tilde{a} - \tilde{a}_{\text{inj}}|\tilde{a} - \tilde{a}_{\text{inj}} - \frac{1}{2}(\tilde{e} - \tilde{e}_{\text{inj}}|\tilde{e} - \tilde{e}_{\text{inj}}). \quad (66)$$

Having fixed the intrinsic parameters and the time, in (45) the modes  $h_{\ell m}$  are fixed as well, and it is convenient to introduce the notation (note that these quantities are defined as complex)

$$\langle \ell m | \ell' m' \rangle = 4 \int \frac{df}{S_n^{a,e}} (6\pi f L)^2 \tilde{h}_{\ell m} \tilde{h}_{\ell' m'}^*. \quad (67)$$

for mode overlaps that are constant factors and can be precomputed for a given system. Here, the noise PSD  $S_n^{a,e}$  is given by (33) and is identical between the channels  $a$  and  $e$ .

Using the results of Sec. II E, and ignoring in (45) the factor  $\exp[2i\pi f k \cdot p_0]$  as a pure constant corresponding to a redefinition of time, we obtain

$$\ln \mathcal{L} = -\frac{1}{2} \text{Re} \sum_{\ell m} \sum_{\ell' m'} [(s_a^{\ell m} - s_{a,\text{inj}}^{\ell m})(s_a^{\ell' m'} - s_{a,\text{inj}}^{\ell' m'})^* + (s_e^{\ell m} - s_{e,\text{inj}}^{\ell m})(s_e^{\ell' m'} - s_{e,\text{inj}}^{\ell' m'})^*] \langle \ell m | \ell' m' \rangle, \quad (68)$$

where we introduced

$$s_{a,e}^{\ell m} = \frac{1}{d} F_{a,e}^{\ell m}, \quad (69)$$

where  $d = D/D_{\text{inj}}$  is the dimensionless ratio of luminosity distances, and with the mode transfer functions  $F_{a,e}^{\ell m}$  given in (44). In each term the intrinsic/extrinsic parameter dependence is thus separated with the intrinsic parameters (that we keep fixed), intervening only in  $\langle \ell m | \ell' m' \rangle$ .

In the case where we only include the dominant harmonic  $h_{22}$ , the likelihood (68) simplifies to

$$\ln \mathcal{L} = -\frac{1}{2} \langle 22 | 22 \rangle [ |s_a^{22} - s_{a,\text{inj}}^{22}|^2 + |s_e^{22} - s_{e,\text{inj}}^{22}|^2 ]. \quad (70)$$

For  $(\ell m) = (22)$ , the functions (69) take the explicit form

$$s_{a,e}^{22} = \frac{1}{4d} \sqrt{\frac{5}{\pi}} \cos^4 \frac{\iota}{2} e^{2i(\varphi - \psi_L)} (F_{a,e}^+ + iF_{a,e}^\times) + \frac{1}{4d} \sqrt{\frac{5}{\pi}} \sin^4 \frac{\iota}{2} e^{2i(\varphi + \psi_L)} (F_{a,e}^+ - iF_{a,e}^\times), \quad (71)$$

with the pattern functions for the channels  $a$  and  $e$  defined in Eqs. (42) and (43).

In order to elucidate the degeneracies in the problem, it will be useful to introduce the following notation. First, since  $\iota \in [0, \pi]$  while  $\beta_L \in [-\pi/2, \pi/2]$ , it will be more convenient to work with the colatitude  $\theta_L = \pi/2 - \beta_L$ . We will further abbreviate notation by using the variables

$$t_\iota \equiv \tan \frac{\iota}{2}, \quad t_\theta \equiv \tan \frac{\theta_L}{2}. \quad (72)$$

We also introduce the azimuthal angles  $\lambda_a = \lambda_L - \pi/6$ ,  $\lambda_e = \lambda_L + \pi/12 = \lambda_a + \pi/4$ , and form the combinations

$$\sigma_\pm \equiv \frac{1}{2} [s_a^{22} \pm i s_e^{22}]. \quad (73)$$

In this notation, we obtain

$$\sigma_+ = \rho e^{2i\varphi} [t_\theta^A e^{-2i\psi_L} + t_\iota^A e^{2i\psi_L}] e^{-2i\lambda_a}, \quad (74a)$$

$$\sigma_- = \rho e^{2i\varphi} [e^{-2i\psi_L} + t_\theta^A t_\iota^A e^{2i\psi_L}] e^{2i\lambda_a}, \quad (74b)$$

with a common prefactor



$$\rho(d, \iota, \theta_L) = \frac{1}{4d} \sqrt{\frac{5}{\pi}} \frac{1}{(1+t_i^2)^2(1+t_\theta^2)^2}. \quad (75)$$

Finding points in the parameter space that are degenerate with the injection, i.e., with  $\ln \mathcal{L} \simeq 0$ , amounts to finding choices for the parameters which obtain the same values as the injection for both quantities  $\sigma_+$  and  $\sigma_-$ .

We can now use the simplified response written in the form shown in Eqs. (74) and (75) to look for symmetries and degeneracies. First, it is easy to check in this new notation that the transformations (63) and (64) indeed leave the likelihood exactly invariant. There is also a symmetry based on exchanging  $\iota \leftrightarrow \theta_L$ . While leaving  $\sigma_-$  unchanged, this conjugates the factor inside brackets for  $\sigma_+$ , which can be compensated using the phase term  $\varphi - \lambda_a$ . If  $\Phi = \text{Arg}[t_\theta^4 e^{-2i\psi_L} + t_i^4 e^{2i\psi_L}]$ , we obtain the symmetry

$$\begin{aligned} \iota' &= \theta_L, & \theta_L' &= \iota, \\ \varphi' &= \varphi + \frac{1}{2}\Phi \pmod{\pi}, \\ \lambda_a' &= \lambda_a - \frac{1}{2}\Phi \pmod{\pi}. \end{aligned} \quad (76)$$

Beyond these discrete symmetries, since likelihood function dependence on the six extrinsic parameters is funneled through just two complex functions, we should expect a two-dimensional degenerate subspace. Indeed we can explicitly find a general solution for parameter values that solve  $\sigma_\pm(D, t_i, t_\theta, \varphi, \lambda_a, \psi_L) = \sigma_{\pm, \text{inj}}$ , which we just sketch here as needed to explain features of the degeneracies.

Defining the ratio

$$r(\lambda_a, t_i, t_\theta, \psi_L) = \frac{\sigma_+}{\sigma_-} = \frac{t_\theta^4 + t_i^4 e^{4i\psi_L}}{1 + t_\theta^4 t_i^4 e^{4i\psi_L}} e^{-4i\lambda_a}, \quad (77)$$

makes clear that  $r = r_{\text{inj}}$  provides one complex condition on four real unknown parameters, eliminating parameters  $D$  and  $\varphi$  while retaining the features of the full degenerate subspace. Going further,  $|r|^2 = |r_{\text{inj}}|^2$  yields one condition on three parameters, eliminating  $\lambda_a$ . In fact, with a little rearrangement, this can be written as a quadratic expression for either  $t_i^4$  or  $t_\theta^4$  given the other variable ( $t_\theta$  or  $t_i$ ) and  $\psi_L$ . Given a solution for  $(t_i, t_\theta, \psi_L)$ , the rest of the solution then proceeds backwards. An explicit expression for commensurate  $\lambda$  comes from solving  $r = r_{\text{inj}}$ , and then  $D$  and  $\varphi$  are obtained from solving, e.g.,  $\sigma_- = \sigma_{\text{inj}}$ .

We can now understand the degeneracies we saw by considering limits of the ratio  $r$ . For instance, when  $t_i^4 \ll 1$ , we have simply  $|r| \simeq t_\theta^4$ . This means that in this limit all values of  $\psi_L$  are allowed, with  $\theta_L$  fixed to a specific value. Similarly, for  $t_i^4 \gg 1$ , we have  $|r| \simeq t_\theta^{-4}$ . A large part of the degenerate subspace volume then tends to be found with a parameter  $\beta_L$  near these special values fixed by the

modulus constraint, with a special value of  $\lambda_L$  coming from the complex argument constraint. By contrast, intermediate values of  $t_i^4$  will not give as much parameter space allowed for the degeneracy, as illustrated by the case  $t_i = 1$  (i.e.,  $\iota = \pi/2$ ). In that case,  $|r| = 1$  irrespective of  $\theta_L$  and  $\psi_L$ , and no solution exists if  $|r_{\text{inj}}| \neq 1$ .

In terms of the original parameters, these special sky positions are

$$\beta_L \approx \beta_L^\dagger = \pm \left( \frac{\pi}{2} - 2\text{Arctan}|r_{\text{inj}}|^{1/4} \right), \quad (78a)$$

$$\lambda_L \approx \lambda_L^\dagger = \frac{\pi}{6} - \frac{1}{4} \text{Arg} r_{\text{inj}} \pmod{\frac{\pi}{2}}, \quad (78b)$$

$$\rho e^{2i(\varphi - \psi_L)} \approx e^{-2i(\lambda_L^\dagger - \pi/6)} \sigma_{\text{inj}}^-, \quad (78c)$$

with the  $\lambda_L \rightarrow \lambda_L + k\pi/2$  and  $\beta_L \rightarrow -\beta_L$  symmetries corresponding to the eight-mode sky symmetry discussed in Sec. V B. The degeneracy for the pair of parameters  $(\varphi, \psi_L)$  is exact, and the constraint  $\rho = |\sigma_{\text{inj}}^-|$  gives an approximate degeneracy for the pair  $(d, \iota)$  as long as we remain in the regime  $t_i^4 \ll 1$  or  $t_i^4 \gg 1$  (which is quite extended, thanks to the quartic power). Thus, for each of these sky positions built from (78), many different values of  $(\varphi, \psi_L)$  and  $(d, \iota)$  produce a waveform very close to the injection, resulting in apparent peaks in the marginalized posterior distribution for the sky position, located at the special sky positions  $(\lambda_L^\dagger, \theta_L^\dagger)$ , which are offset from the injected value.

Similarly, considering the limits  $t_\theta^4 \ll 1$ ,  $t_\theta^4 \gg 1$ , we can expect a significant part of the degenerate parameter space near special values for inclination and correspondingly for  $\lambda_L + \varphi$ ,

$$\iota \approx \iota^\dagger = \frac{\pi}{2} \mp \left( \frac{\pi}{2} - 2\text{Arctan}|r_{\text{inj}}|^{1/4} \right), \quad (79a)$$

$$\lambda_L - \psi_L \approx (\lambda_L - \psi_L)^\dagger = \frac{\pi}{6} - \frac{1}{4} \text{Arg} r_{\text{inj}} \pmod{\frac{\pi}{2}} \quad (79b)$$

$$\rho e^{2i\varphi} \approx e^{-2i((\lambda_L - \psi_L)^\dagger - \pi/6)} \sigma_{\text{inj}}^-. \quad (79c)$$

When considering more harmonics beyond the dominant mode  $h_{22}$ , such degeneracies will be broken easily. As in the case of the full likelihood, the simplified likelihood (68) will have several terms with different inclination and phase dependencies. If the signal is loud enough for at least two modes to be detected, then the inclination  $\iota$  and the phase  $\varphi$  are fixed by the relative amplitude and phase of these modes. This will break the degeneracy  $(d, \iota)$  and the degeneracy  $(\varphi, \psi_L)$ .

We illustrate our findings in by running a Bayesian parameter estimation for the simplified likelihood (70) with our two samplers, MultiNest and ptmcmc. The results display the degeneracy structure that we discussed in this

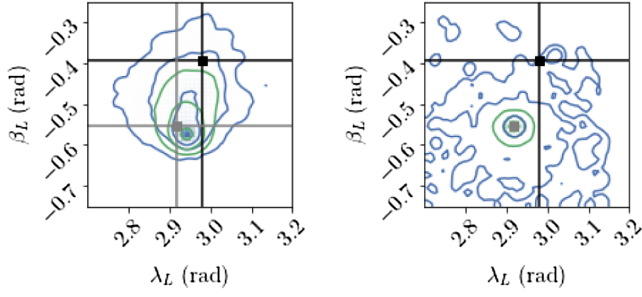


FIG. 11. Sky position posterior for system II in Table II in the presence of strong degeneracies, zooming close to the injected value and discarding the secondary mode in the sky. The left panel shows the complete likelihood, while the right panel shows the simplified likelihood (68) in the *Frozen, low- $f$*  approximation and with fixed masses and time. The results of `ptmcmc` (blue) and `MultiNest` (green) are superimposed. The black cross is the injected value, and the grey cross is the analytic prediction (78) for the offset peak created by degeneracies.

section. Focusing on the sky position in the vicinity of the true source parameters, Fig. 11 contrasts the posterior distribution obtained with the *full* and the *simplified* likelihoods. Although differences are visible, we see that both show a similar peak shifted from the true signal’s sky position that agrees well with our prediction (78). More details on the full posterior distribution with the simplified likelihood are given in Appendix B.

## VI. MASSIVE BLACK HOLES: ACCUMULATION OF INFORMATION WITH TIME

### A. Premerger analysis

The rate at which parameter information accumulates during the observation of an inspiral is crucial for establishing the LISA downlink and data processing requirements, as well as for planning multimessenger observations [9,10]. In this section, we explore how parameter information accumulates on approach to the merger by performing parameter estimation studies with temporally truncated signals, as explained in Sec. IV B: The signals are cut at points where the accumulated SNR is about  $\{10, 42, 166\}$  corresponding to  $\{1/64, 1/16, 1/4\}$  of the total SNR = 666.0, and times  $\{41 \text{ hr}, 2.5 \text{ hr}, 7 \text{ min}\}$  before the merger. As shown in Fig. 4, most of the SNR accumulates over the last few minutes of this signal.

We implement these temporal cuts in the Fourier domain, using the time-frequency correspondence (22) adapted for higher harmonics following (59). Our cuts are sufficiently early before the merger for this relation to be a good approximation for temporal cuts, and consistent among the various harmonics, as shown in Fig. 3.

In Fig. 12 we show how parameter information accumulates at these time points. The results are shown for case II, including the higher harmonics in the signal. Individual masses are loosely constrained, within a factor of 2, at first

detection; the chirp mass combination is better determined. They start to be individually constrained to better than 10% by 2.5 hours before merger. The luminosity distance is poorly determined, within a factor of 4, when reaching the detection threshold 41 hours before merger; the constraint improves to roughly 10% at 2.5 hours prior to merger. Merger time is estimated with an uncertainty of roughly 2 hours at first detection, which improves to a few minutes by 2.5 hours before the merger.

The corresponding sky position posteriors are shown in Fig. 13, using the LISA frame angles in (A15). The lines show the 1-2-3- $\sigma$  uncertainty contours on the sky, for each of our time cuts as well as for the full, postmerger signal, with the two panels differing by the inclusion of higher harmonics. We see clearly, in the premerger analysis, the eight-mode sky degeneracy introduced in Eqs. (63) and (64). The different cuts give us an idea of the continuous evolution from a badly determined sky position at first, when reaching detection, to an eight-mode degeneracy structure, to finally only two modes surviving when reaching the merger, the true injection and the *reflected* sky position.

### B. Degeneracy breaking by the time- and frequency-dependence in the response

The premerger analysis of the previous section made apparent the eight-mode degeneracy pattern that can be predicted analytically from the structure of the *Frozen, low- $f$*  response approximation [see (63) and (64)]. However, the analysis of IMR signals in Sec. VA shows that only the *reflected* mode (63) survives postmerger. In this section we explain how and why this transition occurs.

We have already presented a discussion of different qualitative effects in the instrument response in Sec. IV C and Fig. 5: on one hand, the motion of LISA leaves an imprint (the time-dependency) on the low frequencies, where the time-frequency map is steep enough that a short interval in frequency maps to a large interval of time; on the other hand, the breakdown of the long-wavelength approximation leaves an imprint (the frequency-dependency) at high frequencies. Here we look at the quantitative importance of these features on the inference as a function of frequency (or as a function of time). We can readily select one or the other feature by using the *Frozen* response, ignoring time-dependency, or the *Low- $f$*  response, ignoring frequency-dependency.

In Fig. 14 we show log-likelihood values obtained with either the *Full*, *Frozen*, or *Low- $f$*  response, with and without higher harmonics, for each of the eight modes in the sky described by Eqs. (63) and (64). The likelihood (51) is computed with the same response approximation for the injection and for the template. The results are shown as a function of frequency in a cumulative sense: We compute the likelihood by accumulating signal up to the frequency shown in the  $x$ -axis. When including higher harmonics, this

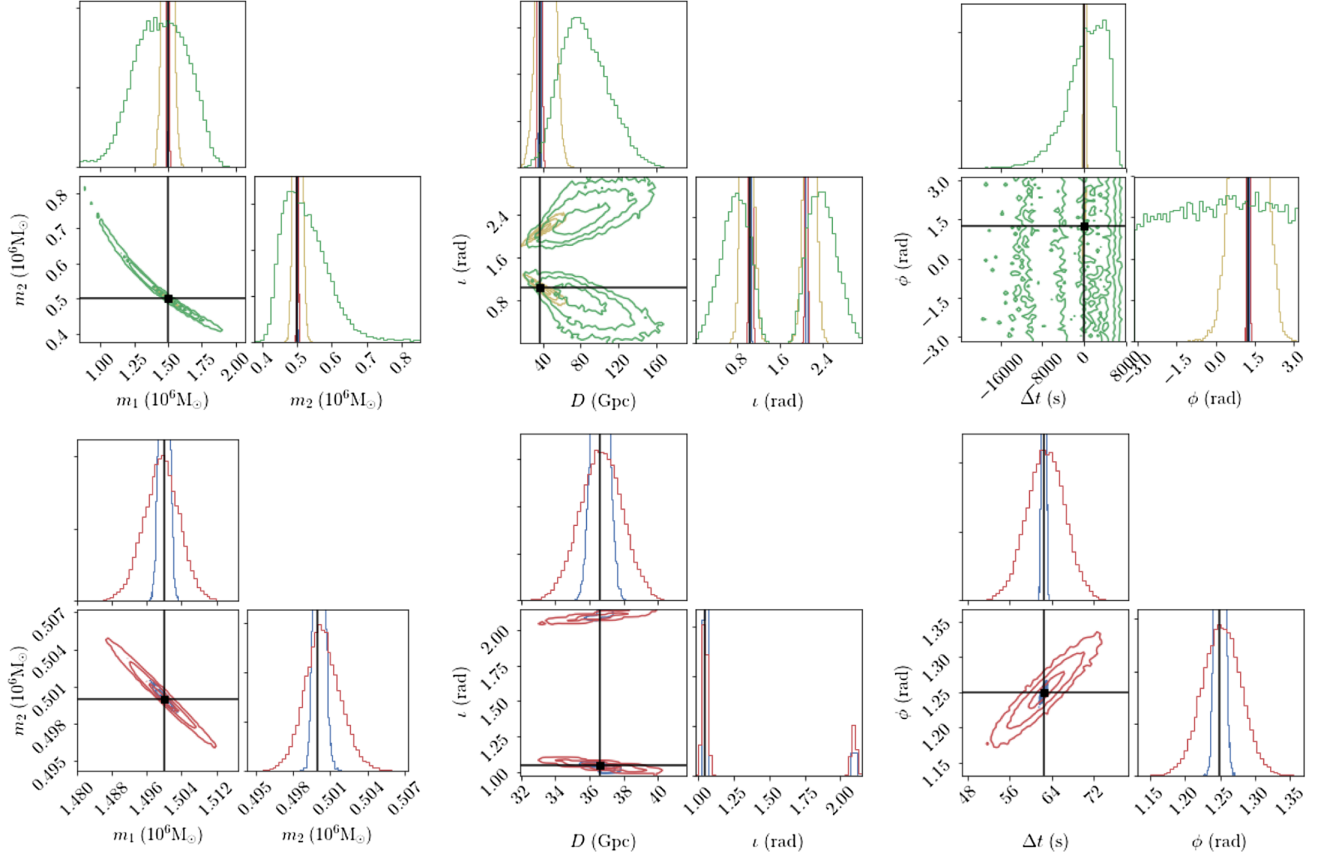


FIG. 12. Refinement of parameter inference over time for system II. The posteriors shown correspond to the time cuts defined in Sec. IV B at SNR/64 (green), SNR/16 (yellow), SNR/4 (red), and postmerger (blue) for the masses (left), distance/inclination (center), and time/phase (right). On the first row, all four posteriors are shown, while in the second row we display only SNR/4 and postmerger, zooming closer to the true parameters. We inject and recover with higher harmonics. The black cross shows the true parameters.

cut in frequency is interpreted as a cut in time and propagated to other harmonics, according to (59).

A value of  $\ln \mathcal{L} = 0$  means that the template signal with shifted parameters is identical to the injection, while a very negative  $\ln \mathcal{L}$  means that these parameter values are ruled out. The injection (full black line) always has  $\ln \mathcal{L} = 0$  by construction. The *reflected* sky mode (63) is the dashed black line, and the *antipodal* sky mode (65) is the dashed red line.

We choose here to show a likelihood measure instead of the perhaps more familiar mismatch between signals. Mismatches are by definition normalized and therefore SNR-insensitive, while the log-likelihood scales are with  $\text{SNR}^2$ . Since LISA could observe SNR values in the hundreds or thousands and intuition based on previous familiarity with mismatch measures used in the LIGO/Virgo context might be lost.

With the *Low- $f$*  response, the *antipodal* mode remains almost exactly degenerate with the injection, differing only by the Doppler phase (36), which has a small effect on our short signals as shown in Fig. 2. As explained below (65), other modes fail to reproduce the injected signal because of the time-dependence of the pattern

functions. They acquire a moderate penalty at low frequency, but  $\ln \mathcal{L}$  then goes to a constant since the motion becomes negligible at high frequencies, as shown by the transfer functions in Fig. 5.

With the *Frozen* response, the *reflected* mode remains exactly degenerate as anticipated in Sec. V B. All other modes see their  $\ln \mathcal{L}$  fall rapidly at around 2 mHz due to the onset of frequency-dependency in the response. Although the frequency-dependency vanishes in the limit  $f \ll f_L \approx 0.12$  Hz, note that the effect is crucial even for  $f/f_L \approx 0.02$ . Compared with Fig. 5, we see that, around 2 mHz, the breaking of the long-wavelength approximation still appears mild; however, the SNR starts to accumulate a lot from SNR/16  $\sim$  40 at 2.5 h to SNR/4  $\sim$  160 at 7 min, magnifying the effect of these features. Comparing the two rows of the figure, we see that the higher harmonics have an effect here, in making the elimination of secondary modes happen slightly earlier in frequency (or time). This is expected as they start to contribute significantly to the SNR (see Fig. 4) and reach higher in frequency than the  $h_{22}$  harmonic according to (59).

The *Full* response case is essentially a superposition of the previous two response approximations, with a



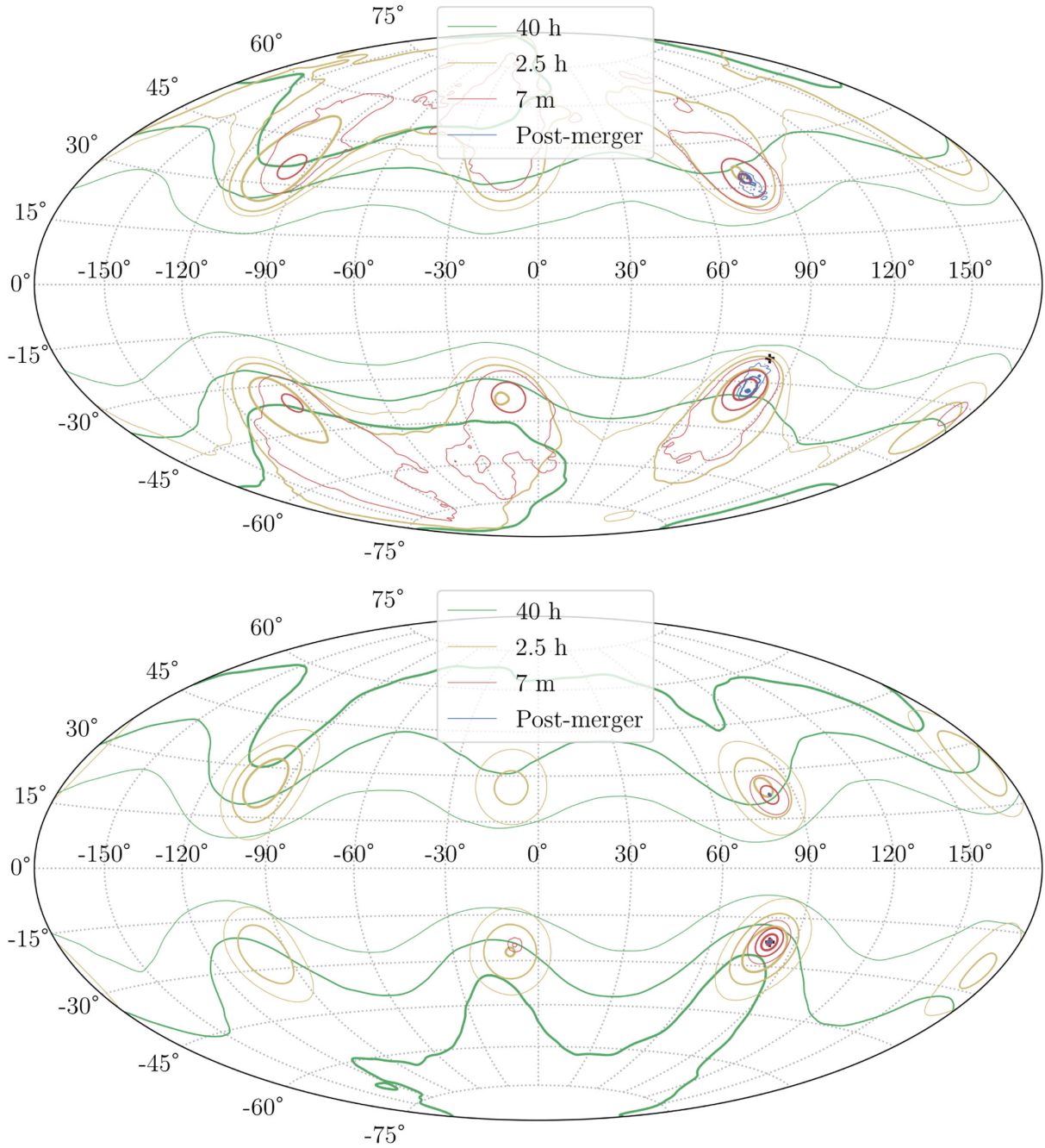


FIG. 13. Refinement of the LISA-frame sky position inference over time for system II. The posteriors shown correspond to the time cuts defined in Sec. IV B at SNR/64 (green), SNR/16, SNR/4, and postmerger (blue). For the top figure, we inject and recover with a 22-mode-only waveform, while for the bottom figure we inject and recover with higher harmonics. The black cross is the injected sky position; the apparent offset in the 22-mode analysis corresponds to the phenomenon highlighted in Sec. V C. The blue contour for the analysis with the full signal corresponds to the posterior of Fig. 7 (where blue and red were used for “22” and “HM”); in the bottom figure with higher harmonics, it is almost reduced to two dots on this scale.

successive onset of the time-dependency and frequency-dependency to break degeneracies. We note that overall, including the higher harmonics does not change this qualitative picture, despite an earlier onset of frequency-dependent degeneracy-breaking features.

Limitations of such explorations should be made clear. By using a pointwise estimate, we cannot make statistical statements and are missing volume effects discussed in Sec. V C; by transforming the angular parameters with an analytical prescription while keeping the other fixed, we



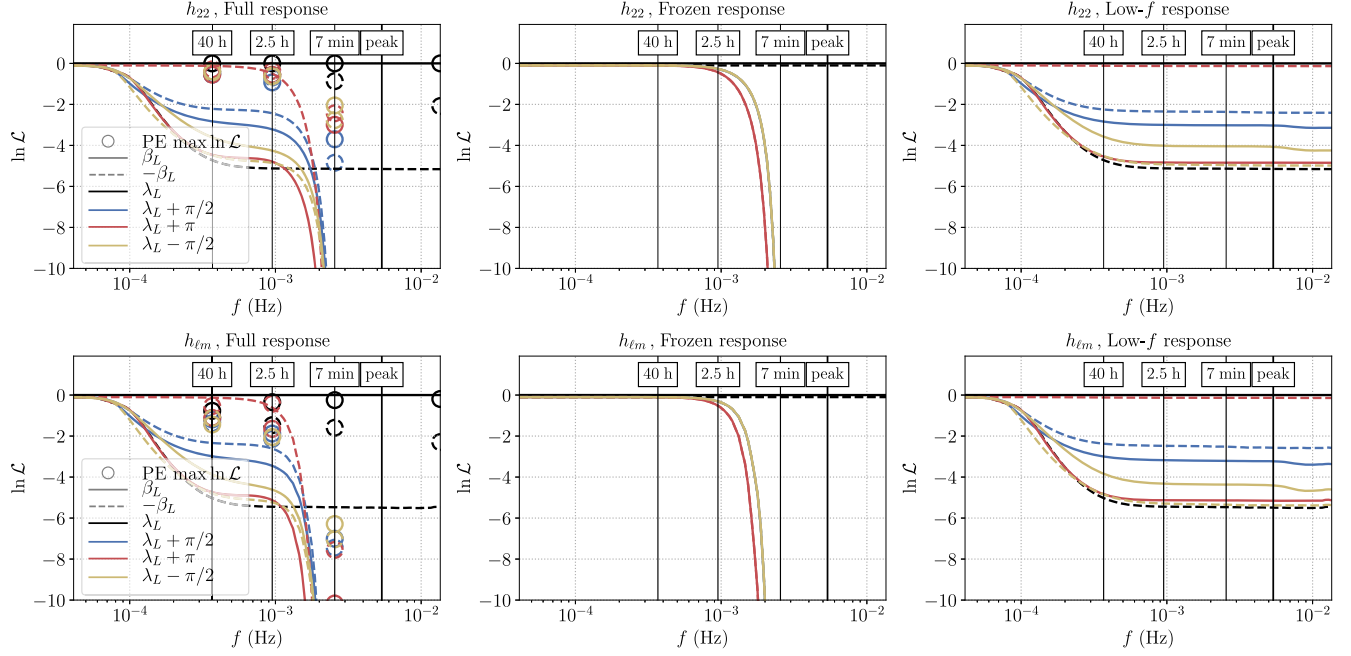


FIG. 14. Log-likelihood (51) at degenerate modes in the sky summarized in Table III, for different response approximations (from left to right: *Full*, *Frozen*, *Low-f*), with 22-mode only (top) and with higher harmonics (bottom). The log-likelihood is shown as a function of the maximal frequency where we cut the 22-mode signal [with the rescaling (59) for the other modes], which can be thought of as accumulating signal with time (the vertical lines represent the time cuts of Sec. IV B and the merger). The lines represent the sky positions listed in Table III: In terms of LISA-frame parameters full or dashed stands for the reflection in latitude, and the color stands for the  $\pi/2$ -shifted longitudes. The dashed black curve corresponds to the *reflected* position, and the dashed red curve to the *antipodal* position. At the true position (solid black),  $\ln \mathcal{L}$  is zero by construction while it is decreasing as a function of  $f$  for other sky positions. Reaching a very negative  $\ln \mathcal{L}$  means that this sky position is excluded by the data accumulated thus far. In the left panel, the circles indicate the maximum log-likelihood found among the posterior samples, when the parameters are allowed to deviate slightly from the theoretical degenerate point.

are, in a sense, overconstraining the degeneracy. In particular, it is conceivable that one could find a better match to the injected signal in the vicinity of a degenerate sky mode, by adjusting all parameters slightly.

This is indeed what happens in multidimensional Bayesian parameter estimation, and is illustrated in Fig. 14. In the first panel, for the three premerger analyses as well as for the postmerger analysis, we overlay circles indicating the best  $\ln \mathcal{L}$  found among the posterior samples, in each eighth of the sky corresponding to the eight sky modes (when no samples are present in this region of the sky, nothing is displayed). A sampler is not an optimizer, so the precise value achieved is not optimal and would vary when repeating runs; nevertheless, this measure gives a good proxy for how much closer to the injected signal we could get by slightly biasing the parameters.

We see that the best  $\ln \mathcal{L}$  among samples can be higher than the point estimate would tell us, especially for the analysis at  $\sim 7$  min before merger, and higher harmonics make a visible difference. This has consequences for

approximations like the Fisher analysis, if secondary peaks are shifted from the analytical predictions. We leave for future work the investigation of approximate representations of sky degeneracies.

## VII. STELLAR-MASS BLACK HOLES

### A. Signals and transfer functions

Stellar-mass black hole binary inspirals have been recently recognized as a potentially important source of LISA detections [7,48]. Full Bayesian parameter estimation studies for these LISA signals have not yet been developed, but the Fisher approach has been used in [7,48–50]. In this section we present two case studies of parameter inference for such sources using the formalism and methods described earlier, while limiting the analysis to the extrinsic parameters (excluding the masses, and fixing the time to coalescence). The parameters of the simulated mergers were drawn from the Radler LISA Data Challenge, which itself followed the population of [7], and are listed in

TABLE IV. Parameters of the simulated SBHB inspiral signals.

Identifier	I	II
Mass 1 ( $M_\odot$ )	21.44	63.51
Mass 2 ( $M_\odot$ )	20.09	45.15
Redshift	0.01	0.07
Lum. Distance (Mpc)	49.1	318.2
Inclination (rad)	0.65	1.92
Phase (rad)	0	0
Ecliptic longitude (rad)	3.44	5.66
Ecliptic latitude (rad)	-0.074	-0.055
Polarization (rad)	1.74	2.20
Time-to-merger (yr)	2	2
LISA SNR $h_{22}$	26.6	11.9

Table IV. System I is moderately massive, placed at a favorable close distance; with an SNR of 26.6 it would be at the tail of the most favorable detections that we can hope for [7,97,98]. System II is a higher-mass, moderate-SNR system. Both systems happen to be close to the ecliptic plane.

Figure 15 illustrates, for the system I in this table, the instrumental transfer functions (30) as well as the Doppler phase (36). Contrasting with the MBHB case, we see that the transfer functions have much more structure in this case. Both the time-dependency and the frequency-dependency in the response lead to oscillatory features in the transfer functions. The Doppler phase (36), that was negligible for our short MBHB systems, now varies by tens of radians already at the lowest frequency.

### B. Bayesian inference of extrinsic parameters

We caution the reader that our purpose here is to show the feasibility of Bayesian inference for SBHB sources in

LISA using the methodology presented in this paper, and to illustrate the qualitative features of the posterior distribution for extrinsic parameters, in relation to our previous investigations for MBHB systems. To this end, the problem we pose is simplified in several ways: we assume that we know the signal to be present, bypassing the challenges of detection itself [98], we fix the masses and the time-to-merger to the injected value, and we neglect spins. The parameter recovery of the intrinsic parameters, masses, and spins is important to address as one can expect strong degeneracies in mass ratio and spin for these inspiral signals; it will be tackled in a forthcoming publication [99].

The waveform model used to work with SBHB systems is a standard post-Newtonian frequency-domain waveform (we use a version of the `TaylorF2` approximant as implemented in `LAL` [53]). In practice, because LISA only observes the early inspiral of SBHB systems, the signal is well within the domain of applicability of the post-Newtonian approximation [100]. In the early inspiral, harmonics beyond the dominant quadrupole give a negligible contribution, so we only include the (2,2) mode.

As done in the previous section, we apply our parameter inference procedure to these systems using both the `MultiNest` and `ptmcmc` samplers, and we numerically evaluate the Fisher matrix at the parameters of the injection. In order to generate these results, our `ptmcmc` sampler requires between  $5 \times 10^5$  and  $3 \times 10^6$  steps with 100 temperatures, taking on the order of days on a single CPU core without careful optimization. The `MultiNest` sampler, on the other hand, completes in a few hours only.

The resulting posterior distributions are shown in Figs. 16 and 17. We find close agreement between the `MultiNest` and `ptmcmc` samplers. The joint distributions for

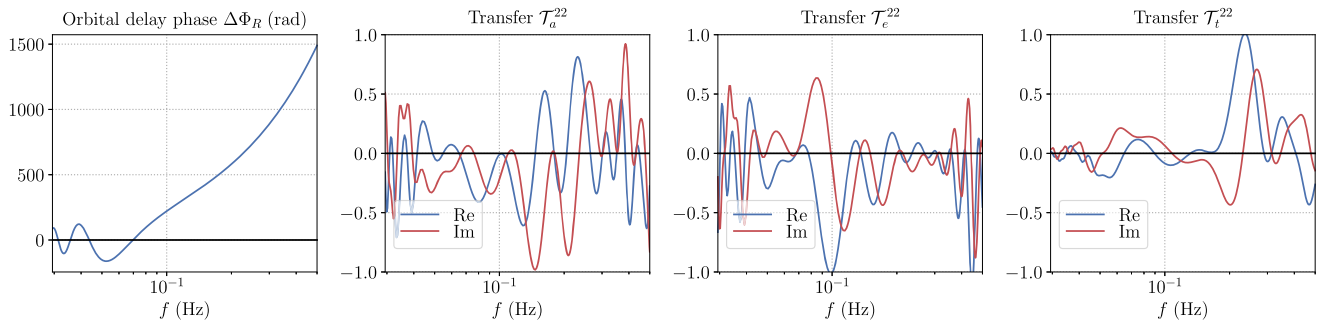


FIG. 15. Example of instrument response for the SBHB system I in Table IV. The first panel isolates the Doppler phase  $\Delta\Phi_R$  as defined in (36), while the three other panels show the TDI transfer functions for  $A$ ,  $E$ ,  $T$ . We show here the transfer functions for the 22-mode contribution to  $a$ ,  $e$ ,  $t$  as defined in (30).

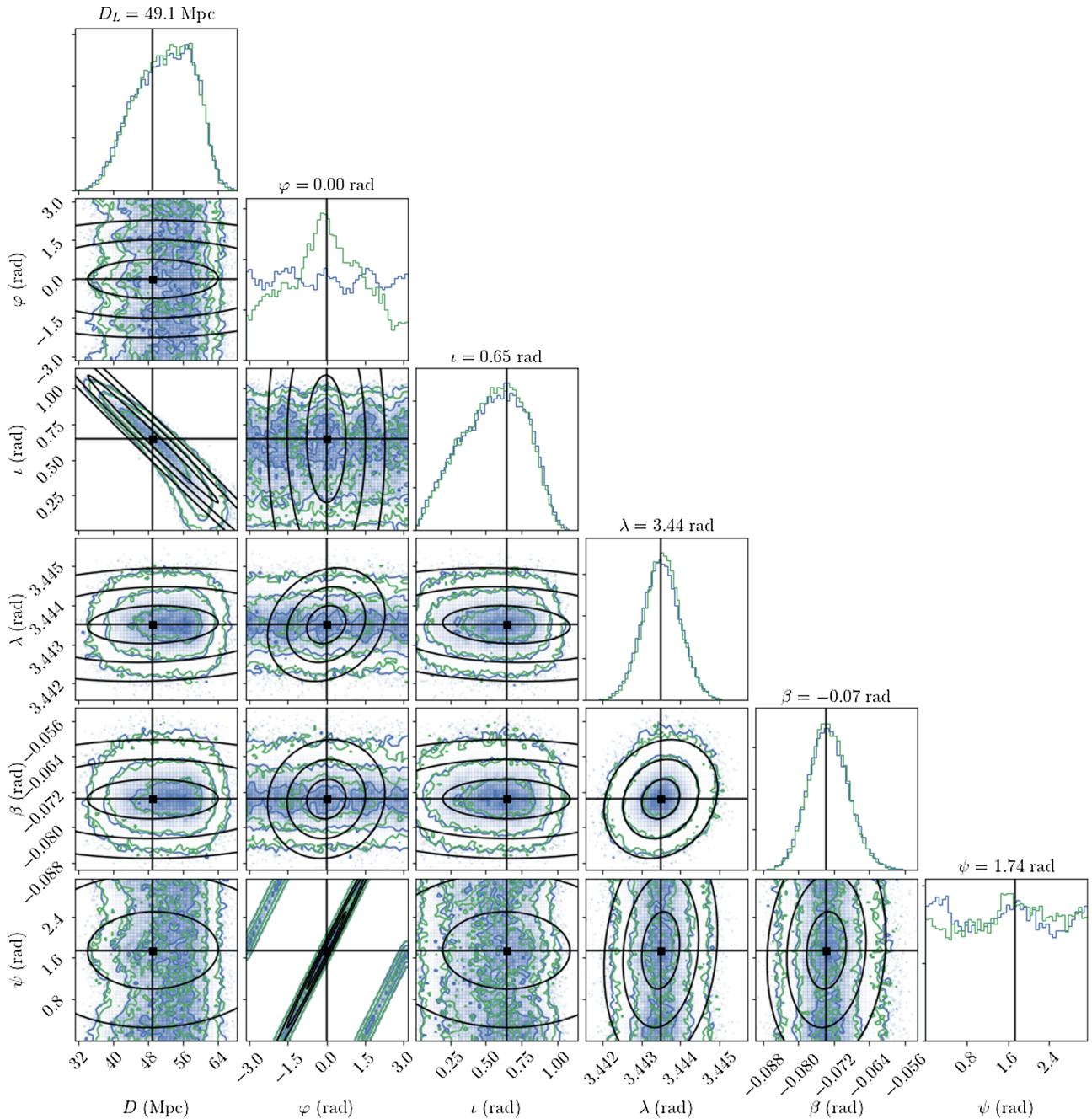


FIG. 16. Inferred parameter posterior distribution for the SBHB system I of Table IV, with 1-2-3- $\sigma$  contours and the black cross indicating the true parameters. The black ellipses show the same contours for the Fisher matrix approximation. Waveforms include only the 22-mode. The results obtained with our two different samplers are shown, `ptmcmc` in blue and `multimc` in green. The initial frequency is fixed to obtain a time-to-merger of two years. All extrinsic parameters are given in the SSB frame.

extrinsic parameters are fairly simple, with extended degeneracies occurring only in the angular parameters ( $\varphi, \psi$ ). Thanks to the complexity of the instrument transfer functions shown in Fig. 15 carrying a lot of information, the sky position determination is very good. The typical degeneracy between inclination and distance is evident, as there is no degeneracy breaking from higher harmonics. Even though the SNR values are much lower than for our

MBHB systems, the Fisher matrix estimates represent quite well the marginal distributions of most parameters. We note, however, some notable deviation from Gaussianity for the system II: the sky position is almost bimodal, with an approximate symmetry across the ecliptic plane (we recall here that the ecliptic plane is not the plane of the LISA instrument that was used to report MBHB results), owing probably to the localization being close to this plane; and



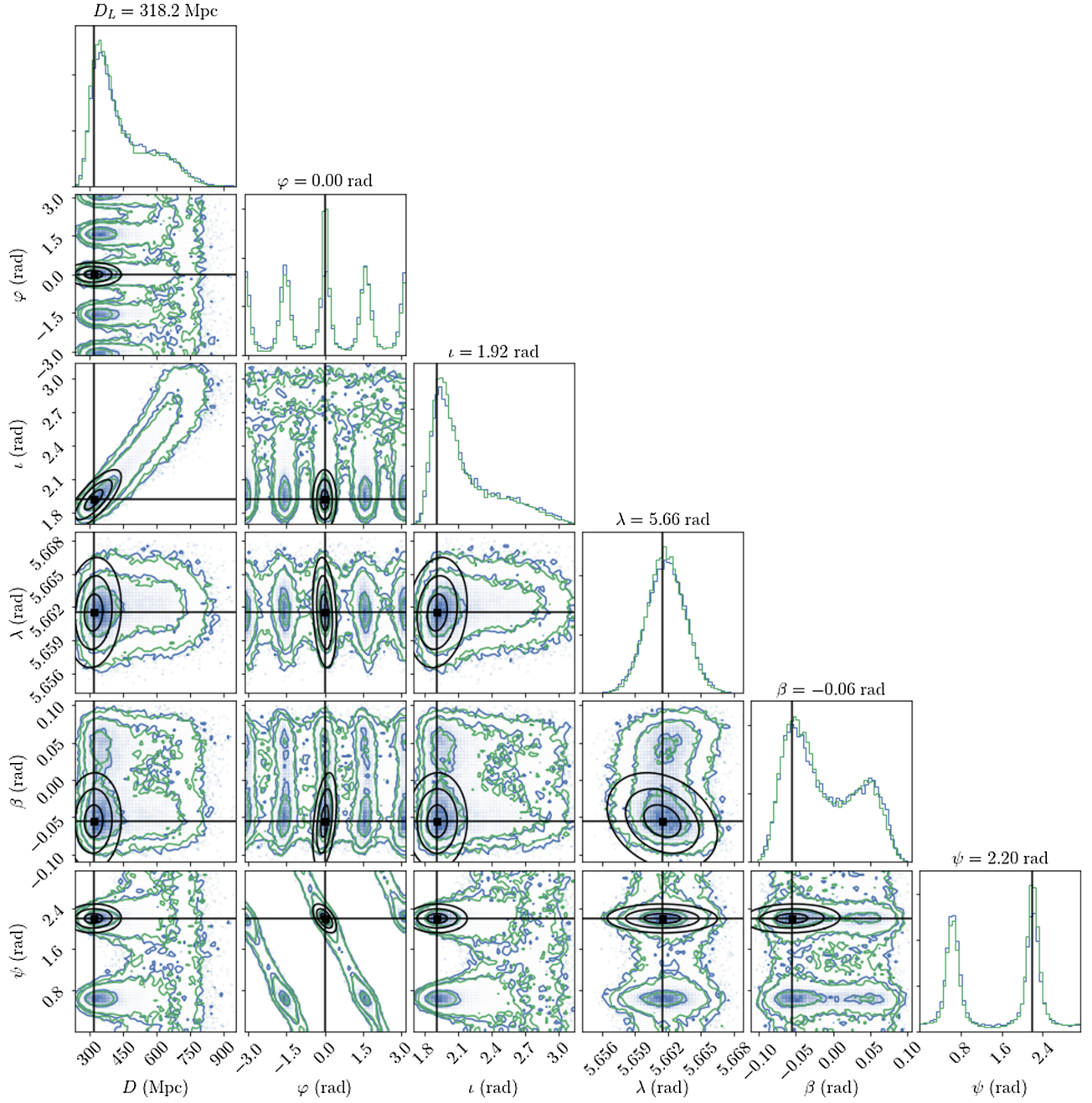


FIG. 17. Same as Fig. 16, for the SBHB system II of Table IV.

the distance and inclination posteriors show a long tail extending way beyond the Fisher estimate.

### VIII. SUMMARY AND DISCUSSION

In this work, we have explored the Bayesian parameter estimation for full, inspiral-merger-ringdown MBHB signals including higher harmonics, albeit neglecting the spin degrees of freedom, and we have also produced posterior distributions for the extrinsic parameters of SBHB signals.

We improved the speed of likelihood computations by combining several ingredients: fast IMR Fourier-domain waveforms, a fast Fourier-domain treatment of the instrument response, and accelerated overlap computations between amplitude/phase signals in the absence of noise. The resulting likelihood costs, of about 15 ms for waveforms with higher harmonics and 1 ms for waveforms including only the dominant harmonic, allowed us to easily perform Bayesian parameter estimation on high-SNR signals requiring up to  $10^8$  likelihood evaluations with standard Bayesian samplers.



Using these tools to simulate the parameter recovery of two examples of MBHB signals, we produced multidimensional posterior distributions. Because our study did not consider spins and their well-known degeneracy with masses, our posterior distributions for masses cannot yet be taken as representative of the actual LISA capabilities. Nevertheless, we encountered and investigated two different kinds of degeneracies in the extrinsic parameter space.

Considering first IMR signals, we found a *two-fold* degeneracy for the sky position corresponding to the *reflected* sky localization with respect to the plane of the LISA constellation (opposite sign of the LISA-frame latitude). We also found that, when limiting the signals to their dominant harmonic  $h_{22}$ , strong degeneracies in distance-inclination can occur, causing the posterior distribution in the sky to appear shifted from the injected value. We gave an analytic explanation of this degeneracy by eliminating the time and frequency dependencies in the response, reducing it to an explicitly degenerate form. However, including higher harmonics in the analysis strongly breaks those degeneracies, and leads to a much better determination of both distance-inclination and sky position. Although the reflected sky localization survives with higher harmonics, we also note that the posteriors appear as much more Gaussian in all extrinsic parameters; using a simplified signal with only the dominant harmonic significantly complicates the posterior sampling.

Second, considering premerger analyses to investigate how information accumulates with time, we found an *eight-fold* degeneracy for the sky position, obtained by rotating the original and reflected sky positions by integer multiples of  $\pi/2$  around the axis of the LISA constellation. For our example systems, the posterior only collapses on the two-fold degeneracy, noticed previously for IMR signals, minutes before the merger. By decomposing the LISA instrument response into different effects, separating its time-dependency from its frequency-dependency, we could pinpoint the role of both effects. The time-dependency of the response, consequence of the motion of LISA, leaves a moderate imprint at low frequencies, penalizing moderately other modes in the sky but leaving the *antipodal* sky position degenerate. By contrast, the frequency-dependency of the response, not included in previous parameter studies, is responsible for the degeneracy breaking and the collapse from eight to two modes in the sky. Note that the relative importance of the time- and frequency-dependency of the response will change for heavier or lighter systems.

We also explored the recovery of extrinsic parameters of SBHB signals (fixing the masses and spin parameters as well as the merger time), using the same methodology as for MBHB signals to arrive at a likelihood cost of  $\sim 2$  ms. Investigating two example systems, we found that the distribution of extrinsic parameters is mostly Gaussian, with a bimodality in the sky position found for one source

lying close to the ecliptic plane, and an extended tail in the distance posterior.

Comparing the parameter estimation results obtained with our two samplers, `ptmcmc` and `multinest`, we found an overall good agreement. We noted however that `multinest` can fall short of exploring the full degeneracies in parameter space, in the degenerate case of a signal without higher harmonics; this limitation is further illustrated by the example of a fully degenerate response. We also found that using Fisher matrix estimates does not reproduce Bayesian results in general: The complicated degeneracies in the sky cannot be covered by a Fisher computation that is intrinsically unimodal, and neither can the specific features of the 22-mode degenerate posteriors. That degeneracies even occur with the high SNRs of our MBHB signals is remarkable, and reminds us that the usual statement of the validity of the Fisher matrix approximation in the high-SNR limit is very dependent on the morphology of the signal. In many respects, the moderate-SNR posteriors for our SBHB systems are more Gaussian than our high-SNR MBHB 22-mode posteriors.

Our exploratory work calls for a number of natural extensions that will help us to better understand the future capabilities of the LISA instrument.

First, we only investigated a handful of sources. In particular, we chose our MBHB systems to be representative of the bulk of the expected population; they are not “golden sources”, i.e., they are not part of the tail of lower-redshift systems that will deliver the most interesting science outputs with LISA, notably in terms of electromagnetic counterparts. It is likely that the parameter recovery here will be different, with a longer detectable signal and the motion of LISA playing a role in the premerger localization. We also limited ourselves to a single redshifted mass, while the morphology of the signals will change strongly between the two ends of the LISA spectrum, from intermediate mass black holes ( $M \sim 10^3 M_\odot$ ) to massive systems ( $M > 10^7 M_\odot$ ).

Second, it will be necessary to incorporate more physics in our waveform models. In this study our signals were limited to nonspinning, quasicircular systems. The spin components along the orbital angular momentum will be degenerate with the masses, but we do not expect this to qualitatively change the posterior distributions for the extrinsic parameters. This will be investigated in a forthcoming publication [101]. Orbital precession and eccentricity, however, will change the harmonic structure of the signals and we expect they will play a much more important role.

Third, we used zero-noise realizations in order to accelerate the likelihood computation. Although we do not expect that the noise realization will cause drastic changes, particularly with the structure of degeneracies of signals in parameter space remaining relevant, fast likelihoods for presampled noisy data will eventually be

necessary. Analyzing signals with a noise injection will also become more important when considering more realistic instrumental noise models and in prototyping algorithms for analysis of actual LISA data.

We also note that our sampling tools themselves (MultiNest and `ptmcmc`) were not tailored for the problem at hand; it is very possible that better sampling algorithm would require much less evaluations of the likelihood. The understanding we gained of the parameter space degeneracies will allow us to inform the sampler with, e.g., tailored jump proposals, effectively telling MCMC chains where to look for degenerate signals in parameter space. For instance, parameter estimation tools used in ground-based gravitational wave astronomy make use of an analytical knowledge of degeneracies [102].

Finally, we recall that many scientific questions about LISA rely on parameter estimation tools, and would benefit from proper Bayesian analyses with realistic signals: premerger localization of sources and advance warnings for electromagnetic instruments, cosmography using LISA sources as standard sirens, astrophysical inference from the source population, accuracy requirements on waveform models, and trade-offs between instrumental design and scientific outputs. We will keep extending our tools to address all these applications.

## ACKNOWLEDGMENTS

We would like to thank Stanislav Babak, Michael Katz, Antoine Klein, Vivien Raymond, and Alexandre Toubiana for useful discussions. We are indebted to Philip Graff for his help with the `bambi` sampler. This study benefited strongly from the use of the HPC clusters Hypatia and Minerva at the Max Planck Institute for Gravitational Physics, T. D. C. was supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Universities Space Research Association under contract with NASA, during part of this work.

## APPENDIX A: CONVENTIONS FOR ORBITS AND FRAMES

For completeness, we give in this Appendix the definitions and convention choices that were sketched in Sec. II but were not essential for the discussion.

### 1. Source frame, wave frame, and SSB frame

We first introduce a source frame  $(\hat{x}_S, \hat{y}_S, \hat{z}_S)$  attached to the binary system emitting gravitational waves. Relating this frame to the physical configuration of the binary is understood as being part of the waveform model. For comparable-mass systems without spin or with aligned spins,<sup>7</sup> the natural choice is to take the normal to the orbital

<sup>7</sup>For precessing systems different choices are possible, like  $\hat{z}_S = \hat{J}$ , the direction of the total angular momentum.

plane as  $\hat{z}_S$ , and we will assume that the remaining rotation around  $\hat{z}_S$  is fixed by the phase convention of the waveform model.

Introducing  $k$ , the wave propagation unit vector going from the source towards the observer, we define the inclination  $\iota$  and the observer phase  $\varphi$  as its spherical angular coordinates in the source frame, so that in that frame

$$k_S = (\sin \iota \cos \varphi, \sin \iota \sin \varphi, \cos \iota). \quad (\text{A1})$$

By convention, for our polarization vectors  $p$  and  $q$  we will use the spherical coordinate vectors  $p = e_\theta^S, q = e_\phi^S$ . In terms of only  $\hat{z}_S$ ,

$$p = q \times k, \quad (\text{A2a})$$

$$q = \hat{z}_S \times k / |\hat{z}_S \times k|. \quad (\text{A2b})$$

The vectors  $(p, q, k)$  form the wave frame, such that the gravitational wave takes in this frame the familiar form

$$H_W = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{A3})$$

We can now relate the wave frame to a detector frame  $(\hat{x}, \hat{y}, \hat{z})$ , which will fix our convention for the sky position and polarization angle. We choose here this detector frame to be based on the plane of ecliptic, and centered on the Solar System Barycenter; we will call this frame the SSB frame. Below, we will introduce another detector frame more suitable for short-lived signals, the LISA frame.

The source position in the sky is given by  $(\lambda, \beta)$ , the ecliptic longitude and latitude in this SSB frame. Like with the source frame, we can introduce spherical vectors  $(e_r^{\text{SSB}}, e_\theta^{\text{SSB}}, e_\phi^{\text{SSB}})$ . Since the propagation vector is  $k = -e_r^{\text{SSB}}$ ,

$$k = (-\cos \beta \cos \lambda, -\cos \beta \sin \lambda, -\sin \beta). \quad (\text{A4})$$

The last degree of freedom between the frames represents a rotation along the line-of-sight, parametrized by the polarization angle  $\psi$ . We introduce reference polarization vectors in the SSB frame as

$$u = \hat{z} \times k / |\hat{z} \times k|, \quad (\text{A5a})$$

$$v = k \times u. \quad (\text{A5b})$$

In terms of spherical vectors,  $(k, u, v) = (-e_r^{\text{SSB}}, -e_\phi^{\text{SSB}}, -e_\theta^{\text{SSB}})$ .

In the following, we will use the notation  $R(v, \varpi)$  to denote the matrix of an active rotation around the vector  $v$

by an angle  $\varpi$ . The polarization angle  $\psi$  is then defined such that  $(p, q)$  are obtained by rotating  $(u, v)$  by the angle  $\psi$  around  $k$ , i.e.,

$$(p, q) = R(k, \psi) \cdot (u, v). \quad (\text{A6})$$

These relations can be summarized by the active rotation matrix  $R_W$  from the SSB frame to the wave frame:

$$R_W = R(z, \lambda - \pi/2) \cdot R(x, \beta + \pi/2) \cdot R(z, \psi). \quad (\text{A7})$$

## 2. LISA trajectories

The orbits of the three spacecrafts around the Sun can be chosen so that, at leading order in the eccentricity of the orbits, the constellation retains the shape of an equilateral triangle in a cartwheeling motion following the Earth's orbit. We denote by  $a$  the semi-major axis,  $e$  the eccentricity, and pose  $\alpha = \Omega_0(t - t_0)$ , with  $t_0$  a reference time for the initial position ( $t_0 = 0$  in our case) and  $\Omega_0 = 2\pi/(1 \text{ yr})$ . The reference SSB frame is  $(\hat{x}, \hat{y}, \hat{z})$ .

Using the notation  $c, s = \cos, \sin \alpha$ , the trajectory of the center of the constellation is simply

$$p_0 = ac\hat{x} + as\hat{y}, \quad (\text{A8})$$

and we take  $a \equiv R = 1 \text{ au}$ . We denote the positions of the spacecrafts as  $p_A$  for  $A = 1, 2, 3$ , and the positions relative to the constellation center as  $p_A^L = p_A - p_0$ . Setting  $\beta_A = 2(A - 1)\pi/3 + \beta_0$  (with  $\beta_0$  and initial condition set to zero in our case), the Cartesian coordinates in the SSB frame of the position of the spacecrafts read

$$\begin{aligned} p_A^L &= ae[\sin\beta_A cs - \cos\beta_A(1 + s^2)]\hat{x} \\ &+ ae[\cos\beta_A cs - \sin\beta_A(1 + c^2)]\hat{y} \\ &- ae\sqrt{3}\cos(\alpha - \beta_A)\hat{z}, \end{aligned} \quad (\text{A9a})$$

The arm length, constant in this approximation, is related to the eccentricity and semi-major axis by

$$L = 2\sqrt{3}ae. \quad (\text{A10})$$

The rigid approximation for the constellation, at first order in  $e$ , can be seen as a first-order approximation in the small parameter  $L/R \simeq 0.017$  for a 2.5 Gm arm length.

## 3. The LISA frame

It will be very useful to introduce a time-dependent frame  $(\hat{x}_L, \hat{y}_L, \hat{z}_L)(t)$  following the detector in its motion. Specifically, we choose this LISA frame (L-frame for short) such that at any time,

$$p_1^L = -\frac{L}{\sqrt{3}}\hat{x}_L, \quad (\text{A11a})$$

$$p_2^L = \frac{L}{2\sqrt{3}}\hat{x}_L - \frac{L}{2}\hat{y}_L, \quad (\text{A11b})$$

$$p_3^L = \frac{L}{2\sqrt{3}}\hat{x}_L + \frac{L}{2}\hat{y}_L. \quad (\text{A11c})$$

This frame also provides us with an equivalent representation of the trajectories making use of rotation matrices. Since in our rigid approximation the constellation remains an equilateral triangle in its cartwheeling motion around the Sun, the configuration of the constellation at a later time is given by a rotation around the Sun composed with a rotation of the constellation around its symmetry axis.

If we denote by  $R_L$  the active rotation from the SSB frame to the L-frame such that for each of the three basis vectors  $(\hat{x}_L, \hat{y}_L, \hat{z}_L) = R_L(\hat{x}, \hat{y}, \hat{z})$ , we have

$$R_L = R(z, \alpha) \cdot R(y, -\pi/3) \cdot R(z, -\alpha), \quad (\text{A12})$$

where we recall that  $\alpha = \Omega_0(t - t_0)$ . For all vectors  $X$  among  $p_A^L, n_A$  we have  $X(t) = R_L \cdot X(t = t_0)$ . If a vector  $X$  is given by its components in the SSB frame, in the L-frame the components are  $X_L = R_L^{-1} \cdot X$ .

While long-lasting signals like SBHBs will see a strong imprint of the LISA orbital motion over the course of observation, MBHB signals are dominated in SNR by a short-lived burst of emission at the merger. For such signals, it will be useful to use a parametrization based on the LISA frame at the time of merger instead of the SSB frame.

The new parameters are defined as playing the same role as the SSB parameters, but relative to the L-frame. Namely, we define  $R_{LW}$  the active rotation matrix from the L-frame to the wave frame, expressed in the L-frame basis, so that for each basis vector expressed in the L-frame basis  $(\hat{x}_W, \hat{y}_W, \hat{z}_W)_L = R_{LW} \cdot (\hat{x}_L, \hat{y}_L, \hat{z}_L)_L$ . The defining condition on  $(\lambda_L, \beta_L, \psi_L)$  is

$$R_{LW} = R(z, \lambda_L - \pi/2) \cdot R(x, \beta_L + \pi/2) \cdot R(z, \psi_L). \quad (\text{A13})$$

Coming back to vectors expressed in the SSB-basis,  $R_L^{-1} \cdot (\hat{x}_W, \hat{y}_W, \hat{z}_W) = R_{LW} \cdot R_L^{-1} \cdot (\hat{x}_L, \hat{y}_L, \hat{z}_L)$ . Since  $(\hat{x}_W, \hat{y}_W, \hat{z}_W) = R_W \cdot (\hat{x}, \hat{y}, \hat{z})$  and  $(\hat{x}_L, \hat{y}_L, \hat{z}_L) = R_L \cdot (\hat{x}, \hat{y}, \hat{z})$ , we arrive at

$$R_{LW} = R_L^{-1} \cdot R_W. \quad (\text{A14})$$

Combining (A14), (A12), (A7), and (A13) yields the following expressions for the L-frame angular parameters<sup>8</sup>:

<sup>8</sup>The convention here is that the point of coordinates  $(x, y)$  in the plane has for argument  $\arctan[x, y]$ .

$$\beta_L = \arcsin \left[ \cos \frac{\pi}{3} \sin \beta - \sin \frac{\pi}{3} \cos \beta \cos (\lambda - \alpha) \right], \quad (\text{A15a})$$

$$\begin{aligned} \lambda_L = \arctan & \left[ \cos \beta \cos \lambda \left( \cos \frac{\pi}{3} \cos^2 \alpha + \sin^2 \alpha \right) \right. \\ & + \cos \beta \sin \lambda \cos \alpha \sin \alpha \left( \cos \frac{\pi}{3} - 1 \right) \\ & + \sin \frac{\pi}{3} \sin \beta \cos \alpha, \\ \cos \beta \sin \lambda & \left( \cos \frac{\pi}{3} \sin^2 \alpha + \cos^2 \alpha \right) \\ & + \cos \beta \cos \lambda \cos \alpha \sin \alpha \left( \cos \frac{\pi}{3} - 1 \right) \\ & \left. + \sin \frac{\pi}{3} \sin \beta \sin \alpha \right], \quad (\text{A15b}) \end{aligned}$$

$$\begin{aligned} \psi_L = \psi + \arctan & \left[ \cos \frac{\pi}{3} \cos \beta + \sin \frac{\pi}{3} \sin \beta \cos (\lambda - \alpha), \right. \\ & \left. - \sin \frac{\pi}{3} \sin (\lambda - \alpha) \right]. \quad (\text{A15c}) \end{aligned}$$

Through  $\alpha = \Omega_0(t - t_0)$ , these expressions are time-dependent, reflecting the rotation of this frame to follow LISA on its orbit.

Finally, it is useful to introduce a “time-at-LISA” parameter  $t_L$ , taken into account the propagation delay from the SSB to the center of the LISA constellation, as

$$t_L = t + k \cdot p_0. \quad (\text{A16})$$

Labeling short-lived MBHB signals by their time of arrival at the SSB is a bad parametrization. For a given measured arrival time at the LISA instrument  $t_L$ , different sky positions will give separate peaks in the “time-at-SSB” variable  $t$ . The use of  $t_L$  eliminates this issue.

#### 4. LISA instrumental noise

The instrumental noise we consider in this study is constructed from [68], which specifies the LISA noise budget for different subsystems. The numerical expressions we use are (with all quantities in SI units):

$$S^{\text{pm}}(f) = \left( \frac{1}{2\pi f c} \right)^2 (3 \times 10^{-15})^2 \left[ 1 + 36 \left( \left( \frac{10^{-4}}{f} \right)^2 + \left( \frac{3 \times 10^{-5}}{f} \right)^{10} \right) \right] + \frac{1}{4} (1.7 \times 10^{-12})^2 \left( \frac{2\pi f}{c} \right)^2 + S^{\text{WD}}(f), \quad (\text{A17a})$$

$$S^{\text{op}}(f) = \left( \frac{2\pi f}{c} \right)^2 [(8.9 \times 10^{-12})^2 + (1.7 \times 10^{-12})^2 + (2 \times 10^{-12})^2]. \quad (\text{A17b})$$

#### APPENDIX B: SAMPLING THE FULLY DEGENERATE EXTRINSIC POSTERIOR

In this Appendix, we further illustrate our findings of Sec. V C by running our Bayesian parameter estimation codes with the *simplified likelihood* (70) consisting in pinning intrinsic parameters (masses and time) and using the *Frozen, low- $f$*  response approximation.

Since the inner products (67) are constants computed only once, the external factors become trivially inexpensive, and the likelihood is very fast to compute. We take advantage of this to run `ptmcmc` for a very large number of steps, ensuring that we do not miss degenerate regions of the parameter space. The resulting posterior distribution is extremely degenerate, as shown in Fig. 18. The sampler `MultiNest` fails to resolve this structure, stopping after having only explored the degenerate region  $\iota \rightarrow 0$  and  $\iota \rightarrow \pi$  and the eight sky positions (78). In accordance with (78), this corresponds to lines in  $\varphi \pm \psi_L = \text{const}$ . The sampler `ptmcmc` finds more degenerate regions, particularly it finds similar structures with the roles of  $\iota$  and  $\beta_L$  exchanged following (79) [as shown notably by the  $(d, \iota)$  and  $(d, \beta_L)$  panels], as well as intermediate regions connecting all these features.

Figure 18 also shows the result obtained when including all available harmonics (52) in (68). The degenerate structures collapse to leave the eight modes corresponding to  $\beta_L \rightarrow -\beta_L$  and  $\lambda_L \rightarrow \lambda_L + k\pi/2$ , this time centered around the injected value and not the shifted value  $(\lambda_L^\dagger, \beta_L^\dagger)$ . We found that this degeneracy-breaking occurs already when including a single subdominant harmonic (for instance  $h_{21}$ ).

These results also give us a hint to explain the results of Sec. V A, and particularly the difference between Bayesian samplers shown in the center left panels of Figs. 9 and 10. Since making the likelihood more degenerate in the simplified likelihood limit (70) magnifies the difference between `ptmcmc` and `MultiNest`, with `ptmcmc` exploring a much larger volume in parameter space, their disagreement with the full likelihood can likely be attributed to an apparent shortcoming of `MultiNest`<sup>9</sup> for resolving extended degenerate regions.

<sup>9</sup>We only used standard values for the metaparameters for this sampler, and we do not exclude that it would be possible to improve results with minor changes.



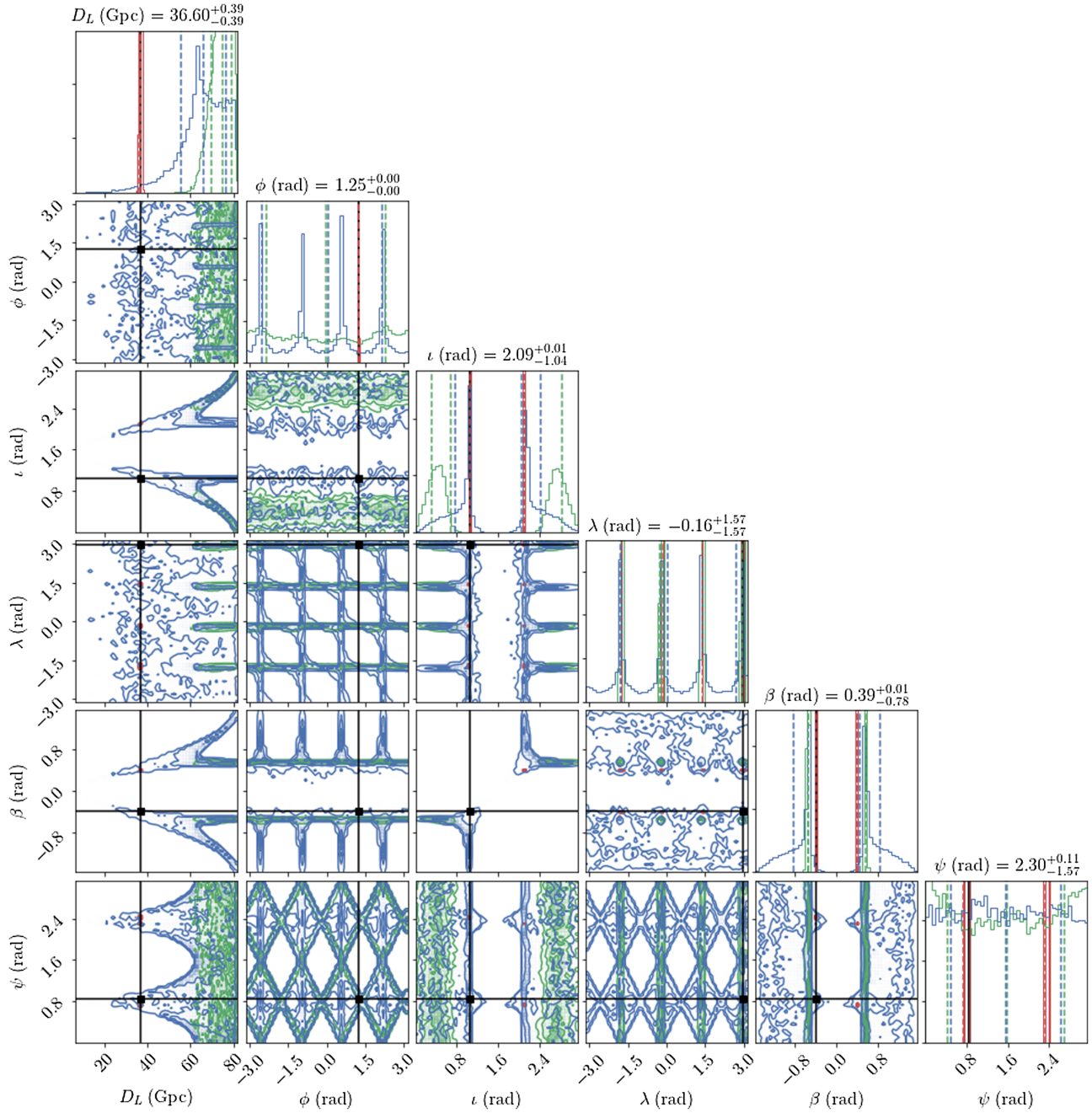


FIG. 18. Posterior distributions obtained with the two samplers `ptmcmc` (blue) and `multiNest` (green) for the simplified likelihood (68), using only the dominant mode,  $(\ell, m) = (2, 2)$ . Masses and time are pinned to the injected value, and the motion of the detector is neglected. We also superimpose (red) the contours obtained with `MultiNest` by including all harmonics  $(2, 2)$ ,  $(2, 1)$ ,  $(3, 3)$ ,  $(4, 4)$ ,  $(5, 5)$ .

- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).  
 [2] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).

- [3] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).  
 [4] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto

- (KAGRA Collaboration), *Phys. Rev. D* **88**, 043007 (2013).
- [5] P. Amaro-Seoane *et al.*, [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).
- [6] A. Klein, E. Barausse, A. Sesana, A. Petiteau, E. Berti, S. Babak, J. Gair, S. Aoudia, I. Hinder, F. Ohme, and B. Wardell, *Phys. Rev. D* **93**, 024003 (2016).
- [7] A. Sesana, *Phys. Rev. Lett.* **116**, 231102 (2016).
- [8] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart, *Astron. Astrophys.* **375**, 890 (2001).
- [9] P. J. Armitage and P. Natarajan, *Astrophys. J. Lett.* **567**, L9 (2002).
- [10] T. Dal Canton, A. Mangiagli, S. C. Noble, J. Schnittman, A. Ptak, A. Klein, A. Sesana, and J. Camp, *Astrophys. J.* **886**, 146 (2019).
- [11] B. F. Schutz, *Nature (London)* **323**, 310 (1986).
- [12] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, *J. Cosmol. Astropart. Phys.* **04**(2016) 002.
- [13] C. Cutler, *Phys. Rev. D* **57**, 7089 (1998).
- [14] S. L. Larson, W. A. Hiscock, and R. W. Hellings, *Phys. Rev. D* **62**, 062001 (2000).
- [15] N. J. Cornish and L. J. Rubbo, *Phys. Rev. D* **67**, 022001 (2003); **67**, 029905(E) (2003).
- [16] S. Marsat and J. G. Baker, [arXiv:1806.10734](https://arxiv.org/abs/1806.10734).
- [17] A. Vecchio, *Phys. Rev. D* **70**, 042001 (2004).
- [18] K. G. Arun, *Phys. Rev. D* **74**, 024025 (2006).
- [19] E. Berti, A. Buonanno, and C. M. Will, *Phys. Rev. D* **71**, 084025 (2005).
- [20] R. N. Lang and S. A. Hughes, *Phys. Rev. D* **74**, 122001 (2006).
- [21] S. Babak *et al.*, *Classical Quantum Gravity* **27**, 084009 (2010).
- [22] D. A. Brown, J. Crowder, C. Cutler, I. Mandel, and M. Vallisneri, *Classical Quantum Gravity* **24**, S595 (2007).
- [23] N. J. Cornish and E. K. Porter, *Classical Quantum Gravity* **23**, S761 (2006).
- [24] J. Crowder, N. J. Cornish, and J. L. Reddinger, *Phys. Rev. D* **73**, 063011 (2006).
- [25] E. D. L. Wickham, A. Stroeer, and A. Vecchio, *Classical Quantum Gravity* **23**, S819 (2006).
- [26] C. Röver, A. Stroeer, E. Bloomer, N. Christensen, J. Clark, M. Hendry, C. Messenger, R. Meyer, M. Pitkin, J. Toher, R. Umstätter, A. Vecchio, J. Veitch, and G. Woan, *Classical Quantum Gravity* **24**, S521 (2007).
- [27] F. Feroz, J. R. Gair, M. P. Hobson, and E. K. Porter, *Classical Quantum Gravity* **26**, 215003 (2009).
- [28] J. R. Gair and E. K. Porter, *Classical Quantum Gravity* **26**, 225004 (2009).
- [29] A. Petiteau, Y. Shang, and S. Babak, *Classical Quantum Gravity* **26**, 204011 (2009).
- [30] E. K. Porter and J. Carré, *Classical Quantum Gravity* **31**, 145004 (2014).
- [31] E. K. Porter and N. J. Cornish, *Phys. Rev. D* **91**, 104001 (2015).
- [32] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, S. Sinha, and C. van den Broeck, *Phys. Rev. D* **76**, 104016 (2007).
- [33] M. Trias and A. M. Sintes, *Phys. Rev. D* **77**, 024030 (2008).
- [34] E. K. Porter and N. J. Cornish, *Phys. Rev. D* **78**, 064005 (2008).
- [35] S. T. McWilliams, J. I. Thorpe, J. G. Baker, and B. J. Kelly, *Phys. Rev. D* **81**, 064014 (2010).
- [36] J. I. Thorpe, S. T. McWilliams, B. J. Kelly, R. P. Fahey, K. Arnaud, and J. G. Baker, *Classical Quantum Gravity* **26**, 094026 (2009).
- [37] S. T. McWilliams, B. J. Kelly, and J. G. Baker, *Phys. Rev. D* **82**, 024014 (2010).
- [38] S. T. McWilliams, R. N. Lang, J. G. Baker, and J. I. Thorpe, *Phys. Rev. D* **84**, 064003 (2011).
- [39] S. Babak, M. Hannam, S. Husa, and B. Schutz, [arXiv:0806.1591](https://arxiv.org/abs/0806.1591).
- [40] eLISA Consortium, P. Amaro Seoane, S. Aoudia, H. Audley, G. Auger, S. Babak, J. Baker, E. Barausse, S. Barke, M. Bassan, V. Beckmann *et al.*, [arXiv:1305.5720](https://arxiv.org/abs/1305.5720).
- [41] V. Baibhav, E. Berti, and V. Cardoso, *Phys. Rev. D* **101**, 084053 (2020).
- [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] M. Pürrer, *Classical Quantum Gravity* **31**, 195010 (2014).
- [45] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi, *Phys. Rev. D* **95**, 044028 (2017).
- [46] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, *Phys. Rev. D* **96**, 024058 (2017).
- [47] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. D* **99**, 064045 (2019).
- [48] S. Vitale, *Phys. Rev. Lett.* **117**, 051102 (2016).
- [49] A. Nishizawa, E. Berti, A. Klein, and A. Sesana, *Phys. Rev. D* **94**, 064020 (2016).
- [50] A. Nishizawa, A. Sesana, E. Berti, and A. Klein, *Mon. Not. R. Astron. Soc.* **465**, 4375 (2017).
- [51] J. N. Goldberg, A. J. Macfarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, *J. Math. Phys. (N.Y.)* **8**, 2155 (1967).
- [52] L. Blanchet, *Living Rev. Relativity* **17**, 2 (2014).
- [53] LSC algorithm library, <http://www.lsc-group.phys.uwm.edu/lal>.
- [54] F. B. Estabrook and H. D. Wahlquist, *Gen. Relativ. Gravit.* **6**, 439 (1975).
- [55] L. J. Rubbo, N. J. Cornish, and O. Poujade, *Phys. Rev. D* **69**, 082003 (2004).
- [56] M. Vallisneri, *Phys. Rev. D* **71**, 022001 (2005).
- [57] A. Krolak, M. Tinto, and M. Vallisneri, *Phys. Rev. D* **70**, 022003 (2004); **76**, 069901(E) (2007).
- [58] M. Tinto and J. W. Armstrong, *Phys. Rev. D* **59**, 102003 (1999).
- [59] J. W. Armstrong, F. B. Estabrook, and M. Tinto, *Astrophys. J.* **527**, 814 (1999).
- [60] F. B. Estabrook, M. Tinto, and J. W. Armstrong, *Phys. Rev. D* **62**, 042002 (2000).
- [61] S. V. Dhurandhar, K. R. Nayak, and J. Y. Vinet, *Phys. Rev. D* **65**, 102002 (2002).

- [62] M. Tinto and S. V. Dhurandhar, *Living Rev. Relativity* **8**, 4 (2005).
- [63] D. A. Shaddock, *Phys. Rev. D* **69**, 022001 (2004).
- [64] N. J. Cornish and R. W. Hellings, *Classical Quantum Gravity* **20**, 4851 (2003).
- [65] D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong, *Phys. Rev. D* **68**, 061303 (2003).
- [66] M. Tinto, F. B. Estabrook, and J. W. Armstrong, *Phys. Rev. D* **69**, 082001 (2004).
- [67] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, *Phys. Rev. D* **66**, 122002 (2002).
- [68] A. Petiteau, M. Hewitson, G. Heinzel, E. Fitzsimons, and H. Halloin, Lisa noise budget, LISA Consortium LISA Tech. Report No. IISA-CST-TN-0001, 2016.
- [69] É. É. Flanagan and S. A. Hughes, *New J. Phys.* **7**, 204 (2005).
- [70] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, *Phys. Rev. D* **84**, 124052 (2011).
- [71] A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999).
- [72] A. Buonanno and T. Damour, *Phys. Rev. D* **62**, 064015 (2000).
- [73] M. Pürrer, *Phys. Rev. D* **93**, 064041 (2016).
- [74] M. Evans, R. Sturani, S. Vitale, and E. Hall, Unofficial sensitivity curves (ASD) for aLIGO, Kagra, Virgo, Voyager, cosmic explorer and ET, Tech. Rep. (2018).
- [75] C. Cutler and E. Flanagan, *Phys. Rev. D* **49**, 2658 (1994).
- [76] E. Poisson and C. M. Will, *Phys. Rev. D* **52**, 848 (1995).
- [77] E. Baird, S. Fairhurst, M. Hannam, and P. Murphy, *Phys. Rev. D* **87**, 024035 (2013).
- [78] R. J. E. Smith, C. Hanna, I. Mandel, and A. Vecchio, *Phys. Rev. D* **90**, 044074 (2014).
- [79] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, *Phys. Rev. Lett.* **114**, 071104 (2015).
- [80] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016).
- [81] N. J. Cornish, arXiv:1007.4820.
- [82] E. K. Porter, arXiv:1411.0598.
- [83] S. Vinciguerra, J. Veitch, and I. Mandel, *Classical Quantum Gravity* **34**, 115006 (2017).
- [84] <https://github.com/JohnGBaker/ptmcmc>
- [85] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **57**, 2607 (1986).
- [86] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **80**, 063007 (2009).
- [87] C. J. F. T. Braak, *Stat. Comput.* **16**, 239 (2006).
- [88] C. J. F. ter Braak and J. A. Vrugt, *Stat. Comput.* **18**, 435 (2008).
- [89] P. Graff, F. Feroz, M. P. Hobson, and A. Lasenby, *Mon. Not. R. Astron. Soc.* **421**, 169 (2012).
- [90] J. Skilling, *Bayesian Anal.* **1**, 833 (2006).
- [91] M. Vallisneri, *Phys. Rev. D* **77**, 042001 (2008).
- [92] E. Barausse, *Mon. Not. R. Astron. Soc.* **423**, 2533 (2012).
- [93] Planck Collaboration, *Astron. Astrophys.* **594**, A13 (2016).
- [94] S. Marsat and S. Babak, arXiv:2003.00357.
- [95] M. L. Katz and S. L. Larson, *Mon. Not. R. Astron. Soc.* **483**, 3108 (2019).
- [96] L. P. Singer and L. R. Price, *Phys. Rev. D* **93**, 024013 (2016).
- [97] D. Gerosa, S. Ma, K. W. K. Wong, E. Berti, R. O’Shaughnessy, Y. Chen, and K. Belczynski, *Phys. Rev. D* **99**, 103004 (2019).
- [98] C. J. Moore, D. Gerosa, and A. Klein, *Mon. Not. R. Astron. Soc.* **488**, L94 (2019).
- [99] A. Toubiana, S. Marsat, S. Babak, J. G. Baker, and T. Dal Canton, *Phys. Rev. D* **102**, 124037 (2020).
- [100] A. Mangiagli, A. Klein, A. Sesana, E. Barausse, and M. Colpi, *Phys. Rev. D* **99**, 064056 (2019).
- [101] M. L. Katz, S. L. Larson, S. Marsat, S. Babak, and C. A. J. K., *Phys. Rev. D* **102**, 023033 (2020).
- [102] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).