

Detectability of higher harmonics with LISA

Chantal Pitte^{1,*}, Quentin Baghi¹, Sylvain Marsat², Marc Besançon¹, and Antoine Petiteau^{1,3}

¹*IRFU, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France*

²*Laboratoire des 2 Infinis—Toulouse (L2IT-IN2P3), Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France*

³*Astroparticule et Cosmologie, Université de Paris, CNRS, Paris, 75013, France*



(Received 7 April 2023; accepted 19 July 2023; published 24 August 2023)

Supermassive black hole binaries (SMBHBs) are expected to be detected by the future space-based gravitational-wave detector LISA with a large signal-to-noise ratio (SNR). This prospect enhances the possibility of differentiating higher harmonics in the inspiral-merger-ringdown (IMR) waveform. In this study, we test the ability of LISA to identify the presence of different modes in the IMR waveform from a SMBHB. We analyze the contribution of each mode to the total SNR for different sources. Higher modes, in particular the mode (3, 3) and (4, 4), can dominate the signal observed through the LISA detector for SMBHB of the order of $10^8 M_\odot$. With Bayesian analysis, we can discriminate models with different harmonics. While spherical harmonics are often considered orthogonal, we observe it is not the case in the merger-ringdown phase observed by LISA. Omitting harmonics not only diminishes the SNR but can also lead to biased parameter estimates. We analyze the bias for each model in a source example and quantify the threshold SNR where we can expect the parameter bias to be comparable to the statistical error. By computing the waveform model error with the Fisher approximation and comparing it with the posterior distribution from our sampler results, we can evaluate the veracity of the analytical bias, which converges with the sampler results as more harmonics are introduced. To conclude, SMBHB events with SNR of a few hundred, as expected in LISA, are required to use templates with at least modes (2, 2), (2, 1), (3, 3), (3, 2), (4, 4), (4, 3) to estimate all intrinsic parameters correctly. Our work highlights the importance of higher modes to describe the gravitational waveform of events detected by LISA.

DOI: [10.1103/PhysRevD.108.044053](https://doi.org/10.1103/PhysRevD.108.044053)

I. INTRODUCTION

In the next decade, the Laser Interferometer Space Antenna (LISA) [1,2] will leave Earth on the quest to find new gravitational-wave (GW) sources. The most intense signals are expected to result from the coalescence of supermassive black hole binaries (SMBHBs). The predicted rate depends on the population and evolution model [3] and vary from 1 to 100 per year with a signal-to-noise ratio (SNR) ranging from a few tens up to thousands [4]. These high-SNR sources allow testing general relativity (GR) with black hole binaries (BHBs). One way to put Einstein's theory to the test with black hole binary coalescences is to probe the no-hair theorem [5,6] through the study of quasinormal modes (QNMs) [7].

After two compact objects have merged into one, a final black hole (BH) in a perturbed state is expected to be obtained. As it stabilizes to quiescence during the ringdown regime, it will emit gravitational radiation that can be described as a superposition of sinusoidal oscillations decaying over, i.e., the QNMs [8–11]. Each QNM has an

associated complex frequency labeled by (l, m, n) , including polar, azimuthal, and overtone indices. In GR, these solutions are entirely determined by the final black hole's mass and spin (M_f, a_f) [12,13]. Nonetheless, the corresponding amplitude and phase of each oscillation depend on the characteristics of the progenitors (the initial BHs) and their relative orientation toward the observer [14,15].

The spin-weighted “spheroidal” harmonics are the eigenfunctions of the ringdown. They can be projected into spin-weighted spherical harmonics, which are the basis of numerical relativity (NR) (see Eq. (3.7) in [8], also see [16–18] for further information and mixing of the modes). This representation is conventionally used to describe the full inspiral-merger-ringdown (IMR) waveform, as it is done, for instance, in phenomenological models, where only the fundamental overtone ($n = 0$) is included in each spherical harmonic in the ringdown regime. From now on, we will refer to spin-weighted spherical harmonics components of the IMR waveform as *modes* and will index them by their mode numbers (l, m) . The response of LISA can be included as a transfer function in the frequency domain separately for each mode, allowing us to produce fast phenomenological

*chantal.pitte@cea.fr

IMR waveforms [19–21] with the appropriate instrumental response [22].

Investigating the role of higher harmonics in SMBHB signals is essential because of the high SNR these events will have in LISA. The analysis of such strong signals will be sensitive to many subdominant features in the waveform and, in particular, to higher harmonics beyond the dominant (2, 2) mode. Modes with different m are often considered orthogonal since their phases scale differently with the orbital phase as $m\phi_{\text{orb}}$, leading to destructive interference. In LISA, this is no longer the case for the merger-ringdown phase, where a large SNR is accumulated over only a few wave cycles. Cross-terms of the harmonics yield an SNR contribution in the likelihood, which can also affect the parameter inference. Thus, the absence of higher harmonics in the template will induce biases in the parameter estimation.

The correct estimation of the parameters of the full waveform is crucial, in particular, to test the no-hair theorem. There are two main approaches to probing the no-hair hypothesis. The first one uses two QNM parameters to find the “hairs” of the final BH, i.e., final mass and spin (M_f, a_f), while a third QNM parameter, or preferably more, are used to check for consistency [7,23]. The second approach compares the estimated final parameters from the full IMR or pre-merger waveform with those obtained from only the ringdown [24]. Their inequality can be understood as a possible deviation from GR. Note that in the latter case, the measure of only one harmonic is necessary to test the hypothesis. A third method called the “merger-ringdown” test has recently been proposed to check for consistency between both regimes. In the same line as the previous method, it is based on the relation of the amplitude and the phase of the QNMs with the properties of the progenitors [25]. To test the hypothesis with any method, one should know the intrinsic parameters with enough accuracy and precision. They include the total mass M , the mass ratio q , and the individual spins S_i of the initial BHs. With the progenitors’ parameters, one can compute the values of final mass and spin (M_f, a_f) (see, e.g., [19,26]) thus allowing one to perform BH spectroscopy, i.e., the study of the BH harmonics. Nonetheless, one should also determine the extrinsic parameters from the IMR or premerger inference to analyze the ringdown regime.

Several analyses of black hole spectroscopy have been proposed over the years [13,27–31]. The work from [23] was the first to use Bayesian analysis to study deviations from GR in the context of the first method; but also see [32,33]. Concerning the second approach, studies of the detectability of overtones ($n \geq 1$) or higher angular modes ($l, m \geq 2$) have been performed by the LIGO-Virgo-Kagra (LVK) collaboration as well as by other authors [24,34–39]. Tension between different authors is observed regarding the presence of higher modes for the detected event GW150914 [40]. We recommend however to read [41]

for a thorough analysis on the detectability of the overtones. The effective-one-body (EOB) formalism [42–44] (full IMR waveform for spinning or nonspinning binaries) was also adapted to study possible deviations from GR, namely pSEOBNR and pEOBNR [45,46], for events in the LVK frequency range. To the best of our knowledge, the full LISA response (including high-frequency effects) has not been taken into account to study ringdown signals.

In this study, we evaluate our ability to identify and differentiate modes of a plausible source detected by LISA and investigate the possible consequences of ignoring modes. To this end, we make use of the software `lisa-beta` [47], which incorporates LISA’s response to the source waveform, as described in Sec. II. We continue explaining the study methodology in Sec. III. In Sec. IV, we analyze the contribution of the modes to the total SNR for general cases. Then we focus on the impact of mode contributions on estimating parameters for a specific event in Sec. V. We also analyze the errors from using an incorrect template relative to the source SNR. Finally, we summarize our conclusions in Sec. VI.

II. SUPERMASSIVE BLACK HOLE WAVEFORMS IN LISA

To study signals observed by LISA, we must incorporate its instrumental response to the GW signature produced by an event. We use the `lisabeta` software developed for this purpose, which accounts for several instrumental effects in LISA. We review the main features implemented in `lisabeta` that are particularly interesting to our study in the following subsections. More detailed information can be found in [22].

A. Waveform in the source and detector frames

We generate the SMBHB waveform with PhenomHM [21], a phenomenological approach based on PhenomD [19,20] for nonprecessing binaries. In addition to the dominant quadrupole ($l = 2, m = 2$), higher modes are introduced, including $(l, m) = (2, 1), (3, 3), (3, 2), (4, 4), (4, 3)$. Incorporating these higher harmonics brings crucial complementary information, not only to the ringdown part of the signal but also to the inspiral and merger, which has been shown to narrow down the posterior of extrinsic parameters in the inference [48–53].

The gravitational-wave signal can be decomposed in spin-weighted spherical harmonics ${}_{-2}Y_{lm}$, which depend on the orientation of the emission parametrized by the inclination ι and the phase φ . The polarizations h_+ and h_\times of a GW are related to their harmonics by

$$h_+ - ih_\times = \sum_{l \geq 2} \sum_{m=-l}^l {}_{-2}Y_{lm}(\iota, \varphi) h_{lm}, \quad (1)$$

where each mode can be described in terms of an amplitude A_{lm} and a phase Φ_{lm} that depends on the intrinsic parameters of the source

$$h_{lm} = A_{lm} e^{-i\Phi_{lm}}. \quad (2)$$

In the frequency domain, nonprecessing binary systems have an advantageous symmetry relation between progrades and retrogrades modes (m and $-m$). It allows us to describe each polarization as

$$\tilde{h}_{+, \times}(f) = \sum_l \sum_{m>0} K_{lm}^{+, \times} \tilde{h}_{lm}(f), \quad (3)$$

where we introduced

$$K_{lm}^+ = \frac{1}{2} (-_2Y_{lm} + (-1)^l {}_{-2}Y_{l-m}^*), \quad (4a)$$

$$K_{lm}^\times = \frac{i}{2} (-_2Y_{lm} - (-1)^l {}_{-2}Y_{l-m}^*). \quad (4b)$$

The GW strain in the traceless-transverse gauge is expressed as

$$h^{TT} = \mathbf{e}_+ h_+ + \mathbf{e}_\times h_\times, \quad (5)$$

where $\mathbf{e}_{+, \times}$ are the polarization tensors,

$$\mathbf{e}_+ = \mathbf{u} \otimes \mathbf{u} - \mathbf{v} \otimes \mathbf{v}, \quad (6a)$$

$$\mathbf{e}_\times = \mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}. \quad (6b)$$

Vectors \mathbf{v} and \mathbf{u} together with the propagation vector \mathbf{k} in spherical coordinates locate the source in the observational frame,

$$\mathbf{u} = \{\sin \lambda, \cos \lambda, 0\}, \quad (7a)$$

$$\mathbf{v} = \{-\sin \beta \cos \lambda, -\sin \beta \sin \lambda, \cos \beta\}, \quad (7b)$$

$$\mathbf{k} = \{-\cos \beta \cos \lambda, -\cos \beta \sin \lambda, -\sin \beta\}. \quad (7c)$$

with (β, λ) as the ecliptic latitude and longitude.

Combining previous equations, we obtain the final expression

$$\tilde{h}^{TT}(f) = \sum_{l,m} P_{lm} \tilde{h}_{lm}(f), \quad (8)$$

where

$$\begin{aligned} P_{lm} &= \mathbf{e}_+ K_{lm}^+ + \mathbf{e}_\times K_{lm}^\times \\ &= \frac{1}{2} [{}_2Y_{lm}(\mathbf{e}_+ + i\mathbf{e}_\times) e^{-2i\psi} \\ &\quad + (-1)^l {}_{-2}Y_{l-m}^*(\mathbf{e}_+ - i\mathbf{e}_\times) e^{2i\psi}]. \end{aligned} \quad (9)$$

Note that P_{lm} depends not only on (l, m) and (i, φ) due to the spherical harmonics, but also on the parameters defining the reference frame, such as the sky localization (β, λ) and the polarization ψ .

B. LISA response

The LISA constellation comprises three spacecraft (S/C) deployed triangularly. As a GW travels across one arm of the constellation, the detectors at both ends of the arm will read the frequency shift between the sent and received signals. This response is known as the link response. One advantage of this triangular setup is that it allows for forming multiple interferometers with different combinations of the links. Each link response is defined as

$$\begin{aligned} y_{rs}(t_r) &\simeq \frac{1}{2(1 - \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}_{rs}(t_r))} [H_{rs}(t_r - L_{rs}(t_r) \\ &\quad - \hat{\mathbf{k}} \cdot \mathbf{x}_r(t_r)) - H_{rs}(t_r - \hat{\mathbf{k}} \cdot \mathbf{x}_s(t_r))], \end{aligned} \quad (10)$$

where we used $c = 1$. The position of the S/C is designated by $\mathbf{x}_{r,s}$ where the subindex s stands for the sender, while r stands for the receiver. Note the ordering of those indices, which follows the last convention adopted in [54]. Their values go from 1 to 3, indexing the three S/C. Therefore six combinations are produced, resulting in six different links. The unit vector $\hat{\mathbf{k}}$ defines the direction of propagation of the GW, while $\hat{\mathbf{n}}_{rs}$ is the direction of the beam. Finally, L_{rs} is the arm's length between the two S/C, and H_{rs} is the source's gravitational strain projected into the arm,

$$\begin{aligned} H_{rs}(t) &= (h_+(t) \cos 2\psi - h_\times(t) \sin 2\psi) \hat{\mathbf{n}}_{rs}(t) \cdot \mathbf{e}_+ \cdot \hat{\mathbf{n}}_{rs}(t) \\ &\quad + (h_+(t) \sin 2\psi + h_\times(t) \cos 2\psi) \hat{\mathbf{n}}_{rs}(t) \cdot \mathbf{e}_\times \cdot \hat{\mathbf{n}}_{rs}(t). \end{aligned} \quad (11)$$

The constellation as a whole will follow Earth in the same orbit. However, each S/C will follow its own orbit around the sun and within the constellation. The orbital motion translates into a time variation in the orientation of the detector relative to the solar system barycenter frame (SSB). Note that this introduces modulations on the signal observed by LISA, as is noticeable from the explicit time-dependent prefactors and delays in the instrumental response given by Eqs. (10) and (11). In LISA's frequency band, the observation of SMBHBs can last from days to months, depending on the total mass and frequency evolution. Therefore, their waveform can be strongly affected by these modulations. Additionally, in the postprocessing, beams are combined in the time-delay

interferometry (TDI, see Refs. [55,56] and Sec. II B 1) to cancel laser frequency noise. As a result, the time delays and their variations leave an imprint in the measured signal.

1. TDI and frequency-domain formulation

Time delay interferometry was first proposed by Armstrong, Estabrook, and Tinto [57] as a solution to cancel the dominant noise produced by fluctuations in the laser frequency. The idea is to combine the links linearly with an adequate time delay to eliminate laser frequency noise. Different combinations have been proposed depending on the characteristics of the constellation. We can find what is known as a first-generation Michelson interferometer (TDI 1.5) for a stationary unequal-arm constellation and a second-generation Michelson interferometry (TDI 2.0) for a rotating unequal- and flexing-arm constellation. Note that any actual constellation will have flexing arms, which prompted the development of TDI 2.0.

It was shown (see e.g., [22]) that it is possible to approximately express the response of the links in terms of the harmonics in the frequency domain as

$$\tilde{y}_{rs}(f) = \sum_{lm} \mathcal{T}_{rs}^{lm}(f) \tilde{h}_{lm}(f), \quad (12)$$

where $\mathcal{T}_{rs}^{lm}(f) = G_{rs}^{lm}(f, t_f^{lm})$ is a kernel carrying information on the modulation and time-delay of the links response. It is defined as

$$G_{rs}^{lm}(f, t) = \frac{i\pi f L_{rs}}{2} \text{sinc}[\pi f L_{rs} (1 - \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}_{rs}(t))] \cdot e^{i\pi f (L_{rs} + \hat{\mathbf{k}} \cdot [\mathbf{x}_r(t) + \mathbf{x}_s(t)])} \hat{\mathbf{n}}_{rs}(t) \cdot P_{lm} \cdot \hat{\mathbf{n}}_{rs}(t). \quad (13)$$

The frequency-domain response can be obtained thanks to a generalized relation giving a time-to-frequency correspondence. This relation extends the one found within the stationary phase approximation (SPA) [58,59], which applies only to the slowly evolving inspiral signal, to the merger-ringdown part of the signal. It reads [22,52]

$$t_f^{lm} \equiv -\frac{1}{2\pi} \frac{d\Phi^{lm}}{df}, \quad (14)$$

where Φ^{lm} is the phase of each mode (l, m) .

To simplify the transfer function in the frequency domain, some assumptions were made:

- (i) the constellation forms an equilateral triangle, then the arms will remain equal and constant;
- (ii) $L_{12} = L_{21}$, i.e., the arm length is similar in both directions since the relative motion of the beam relative to the S/C (also known as the pointing-ahead effect) is not taken into account.

With these assumptions and after factoring out several terms, we can write the response in terms of channels A , E , T ; the optimal linear combinations of Michelson

variables [60] that are approximately orthogonal relative to the noise. They read

$$\tilde{A} = \frac{i\sqrt{2} \sin(2\pi f L)}{e^{-2i\pi f L}} [(1+z)(\tilde{y}_{13} + \tilde{y}_{31}) - \tilde{y}_{32} - z\tilde{y}_{23} - \tilde{y}_{12} - z\tilde{y}_{21}], \quad (15a)$$

$$\tilde{E} = \frac{i\sqrt{2} \sin(2\pi f L)}{\sqrt{3}e^{-2i\pi f L}} [(1-z)(\tilde{y}_{31} + \tilde{y}_{13}) + (2+z)(\tilde{y}_{21} - \tilde{y}_{23}) + (1+2z)(\tilde{y}_{12} - \tilde{y}_{32})], \quad (15b)$$

$$\tilde{T} = \frac{4 \sin(\pi f L) \sin(2\pi f L)}{\sqrt{3}e^{-3i\pi f L}} [\tilde{y}_{12} - \tilde{y}_{21} + \tilde{y}_{23} - \tilde{y}_{32} + \tilde{y}_{31} - \tilde{y}_{13}], \quad (15c)$$

with $z \equiv e^{2i\pi f L}$.

III. METHODOLOGY

This study aims to quantify our ability to identify the presence of different modes in a SMBHB event detected by LISA. We expect the detectability of the modes to be related to the SNR itself and the relative SNR of each mode. In this framework, there are two main parameters to consider in the computation of the SNR: the luminosity distance and the mass. While the distance is just a scaling factor that affects all modes similarly, the mass (and therefore the frequency) moves the signal to lower and higher frequencies affecting the relative weight of the inspiral and merger-ringdown. Since the observed signal is characterized by the redshifted mass, we will distinguish between the source-frame mass and the observed (redshifted) mass.

As an analogy to the luminosity peak in the time domain, we use here an *ad-hoc* frequency-domain definition of *frequency peak* as the frequency at the maximum value of the observed rescaled TDI variable for each mode; see Eq. (29) of [52] for a definition of rescaled TDI variables. With the time-to-frequency correspondence in Eq. (14), this parameter lies about $2.4t_M$ ($t_M = tc^3/MG$, adimensional value) after the time of coalescence. In the waveform observed by LISA, this frequency peak lies in the ringdown regime, even though there is no clear starting point for it (see ongoing discussion on this topic [17,31,41,61,62]). The mass of the source will move the frequency peaks through the spectrum so that the contribution of each mode depends on the sensitivity of LISA at the frequency peaks. For this reason, we study the contribution of each mode in terms of the mass. Results are shown in Sec. IV.

Once we know the contribution of different modes in the general case, we can focus on a specific event as an example. The method we use to assess the detectability of modes is Bayesian model comparison and parameter estimation. Each model corresponds to a different

combination of modes in generating the waveform. This method allows us to compare different models based on Bayes factors, as we shall see.

A. Bayesian analysis

In Bayesian analysis, the posterior distribution of the strain parameters θ given the observed data d , is expressed as

$$p(\theta|d, M) = \frac{p(d|\theta, M)p(\theta|M)}{p(d|M)}, \quad (16)$$

where θ are the physical parameters of the source, M is the model (and any other context) considered, $p(d|M)$ is the evidence, $p(\theta|M)$ the prior of the parameters (in practice independent of M), and $p(d|\theta, M)$ the likelihood. The three latter are also usually denoted by \mathcal{Z} , $\pi(\theta)$, and $\mathcal{L}(\theta)$, as the evidence, the prior, and the likelihood, respectively. In the following, we drop the indication of the model M in the equations.

The likelihood for a Gaussian noise with a covariance matrix \mathbf{C} takes the form

$$\mathcal{L} = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} e^{-\frac{1}{2}(d-h(\theta))^\dagger \mathbf{C}^{-1} (d-h(\theta))}. \quad (17)$$

From now on, we drop the tilde notation for simplicity, even though we work in the frequency domain. One advantage of working in the frequency domain is that each frequency bin is approximately independent when dealing with stationary noise. The noise covariance can therefore be diagonalized and represented by its power spectral density (PSD).

We introduce the definition of the inner product as

$$(a|b) = 4\text{Re} \int_0^\infty \frac{a(f)b^*(f)}{S_n(f)} df, \quad (18)$$

where S_n is the noise's PSD and b^* is the complex conjugate of b . Eq. (17) can thus be rewritten as

$$\ln \mathcal{L} = -\frac{1}{2}(d-h(\theta)|d-h(\theta)) + \text{const.} \quad (19)$$

It can be decomposed as

$$\ln \mathcal{L} = (d|h(\theta)) - \frac{1}{2}(h(\theta)|h(\theta)) - \frac{1}{2}(d|d), \quad (20)$$

where the last term can be neglected since it does not depend on the estimated parameters and represents a multiplicative constant in the likelihood. The full log-likelihood is a sum over the log-likelihoods of the uncorrelated instrumental channels A , E , T . We use an adaptive heterodyned likelihood to speed up the likelihood

computation (see [63,64] and references therein for more information).

Given a chosen prior, the Bayes factor is a means to compute the preference of the data toward one model or another. It is defined as the ratio of the evidence \mathcal{Z} of two models i and j

$$\mathcal{B} = \frac{\mathcal{Z}_i}{\mathcal{Z}_j}, \quad (21)$$

where the evidence is the integral of the likelihood over the whole parameter's hyper-volume,

$$\mathcal{Z} = \int_{\Theta} \mathcal{L}(\theta)\pi(\theta)d\theta. \quad (22)$$

IV. STUDY OF THE MODES CONTRIBUTIONS TO THE SNR

The SNR builds up in time and frequency and is defined by ρ as

$$\rho^2 = \sum_{lm} \sum_{l'm'} \sum_I 4\text{Re} \int \frac{\mathcal{H}_{lm}^l(f)\mathcal{H}_{l'm'}^{l'*}(f)}{S_n(f)} df, \quad (23)$$

where the sum over independent channels (index I) extends over A , E (which have the same noise PSD S_n). The assumed PSD is drawn from the ‘‘science requirement model’’ SciRDv1 [2], including the galactic white dwarf confusion noise, subtracting sources over one year. Note that we are using \mathcal{H}_{lm}^l instead of h_{lm} , because we also include LISA instrumental response and TDI post-processing combining Eqs. (12)–(15). They read

$$\begin{aligned} \mathcal{H}_{lm}^a(f) &= h_{lm}(f) \cdot \frac{i\sqrt{2} \sin(2\pi fL)}{e^{-2i\pi fL}} \\ &\cdot [(1+z(f))(\mathcal{T}_{13}^{lm}(f) + \mathcal{T}_{31}^{lm}(f)) \\ &- \mathcal{T}_{32}^{lm}(f) - z(f)\mathcal{T}_{23}^{lm}(f) \\ &- \mathcal{T}_{12}^{lm}(f) - z(f)\mathcal{T}_{21}^{lm}(f)], \end{aligned} \quad (24a)$$

$$\begin{aligned} \mathcal{H}_{lm}^e(f) &= h_{lm}(f) \cdot \frac{i\sqrt{2} \sin(2\pi fL)}{\sqrt{3}e^{-2i\pi fL}} \\ &\cdot [(1-z(f))(\mathcal{T}_{31}^{lm}(f) + \mathcal{T}_{13}^{lm}(f)) \\ &+ (2+z(f))(\mathcal{T}_{21}^{lm}(f) - \mathcal{T}_{23}^{lm}(f)) \\ &+ (1+2z(f))(\mathcal{T}_{12}^{lm}(f) - \mathcal{T}_{32}^{lm}(f))], \end{aligned} \quad (24b)$$

where $z(f) = e^{2i\pi fL}$. For convenience, we use the notation of inner product for modes,

$$(lm|l'm') = \sum_I 4\text{Re} \int \frac{\mathcal{H}_{lm}^l(f)\mathcal{H}_{l'm'}^{l'*}(f)}{S_n(f)} df. \quad (25)$$

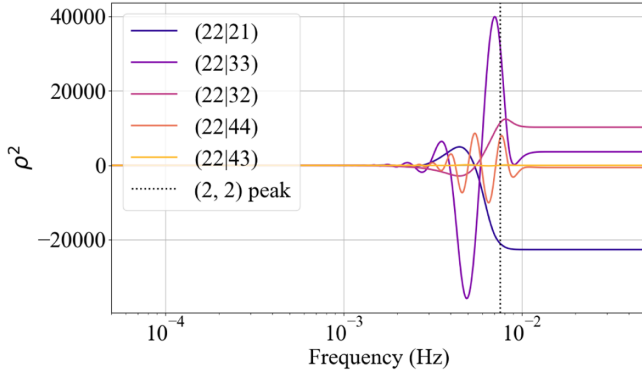


FIG. 1. Contribution of cross-terms $(22|l'm')$ to ρ^2 . We can observe how the cumulative squared SNR changes from positive to negative values and vice-versa until after the ringdown, where it remains constant. The frequency peak of the mode $(2, 2)$ is shown here for guidance. Each mode peaks at a different frequency, so the stabilization period starts at a different point for each pair of modes.

Then the squared SNR can be written as

$$\rho^2 = \sum_{lm} \sum_{l'm'} (lm|l'm'). \quad (26)$$

In Eq. (25), the cross-terms $(lm|l'm')$ have no reason to be positive and can contribute negatively to the total SNR. In other words, the phases can be constructive or destructive, which depends on the values of the ecliptic latitude β and longitude λ , the inclination ι , the phase ϕ , the polarization angle ψ and the mass ratio q . An illustration of this outcome can be seen in Fig. 1. For this example, we use the parameters written in Table I with a redshifted total mass of $\sim 2.44 \times 10^6 M_\odot$. In Fig. 1, we show the $(22|l'm')$ cross-terms of the accumulated squared SNR ρ^2 varying between positive and negative values depending on the frequency. After a certain point, it remains constant since there is no more contribution to the SNR, either positive or negative. This happens at a different frequency for each pair of modes. We plot the frequency peak of the $(2, 2)$ mode with a dotted black line as a guide.

TABLE I. Source parameters in SSB frame with aligned spins and redshifted masses.

Parameter	Value	Parameter	Value
Mass (M_\odot)	$[10^5, 5 \times 10^9]$	β (rad)	$\pi/2$
q (Mass ratio)	2	λ (rad)	π
χ_1	0.5	ϕ (rad)	$\pi/2$
χ_2	0.5	ψ (rad)	$\pi/2$
redshift	3	ι (rad)	$\pi/3$

A. Dependency on mass and frequency

The next natural step to understanding the detectability of modes is to analyze the contribution of each pair of modes to the total SNR for different sources. In general, the SNR depends on all parameters of the source. To simplify our analysis, we fix most parameters to arbitrary values (listed in Table I) and let the mass vary. Note that the results, particularly the details of the hierarchy between modes and cross-terms, might depend on this choice of parameters. For each source's total mass and each frequency bin, we compute the accumulated SNR for each mode and normalize it to the total SNR. Since the ratio of the total mass over the luminosity distance increases for higher masses, the normalization allows us to compare the contribution of each mode regardless of their absolute SNR value.

We gather the results in a set of plots in Figs. 2(a) and 2(b). Each panel of the figure corresponds to a mode and shows the squared SNR dependence on both the source mass and the observed frequency. Then, for a given mass, we can observe how the squared SNR accumulates in frequency for each pair of modes. Each set of plots has its color bar where warmer colors correspond to larger accumulated SNR contributions. The numbers at the right of the bars represent the contribution ratio of each pair of modes to the total squared SNR ρ^2 . Note that we use squared SNR instead of SNR so that the sum of all contributions is equal to 1. This choice also allows us to highlight the negative contributions we mentioned before and their direct impact on the likelihood [see Eqs. (20) and (26)]. That said, the most noticeable feature is the positive contribution of the square terms $(lm|lm)$ over all the frequencies, while in contrast, the cross-terms $(lm|l'm')$ can have negative contributions. We decided to plot the frequency peak of the $(2, 2)$ mode (diagonal dotted line) as a mapping guide since the contribution to the SNR changes considerably in the merger-ringdown regime. Depending on the mode, the SNR variation will start before or after this frequency line.

In the left plot of Fig. 2(a), we note that the contribution of $(22|22)$ to ρ^2 is between 80% to 94% of the total, up to masses $\sim 4 \times 10^6 M_\odot$, while it decreases to 30% around masses of $\sim 10^8 M_\odot$. The rectangular darker area at the bottom right, between frequencies $[10^{-3} - 5 \times 10^{-2}]$ Hz and masses $[10^5 - 2 \times 10^6] M_\odot$, indicates that most of the SNR comes from the inspiral part. This is expected since the waveform peaks outside or at the limit of the LISA frequency band. The pair $(33|33)$ (in the center) has a small contribution for low masses but exhibits a considerable increase up to 21% for high masses around $10^8 M_\odot$. Finally, the case of $(44|44)$ (right plot) shows a similar behavior but is augmented by a factor ~ 2.5 , representing about 54% of the total SNR for large-mass SMBHBs. In other words, if the redshifted total mass of the system is larger than

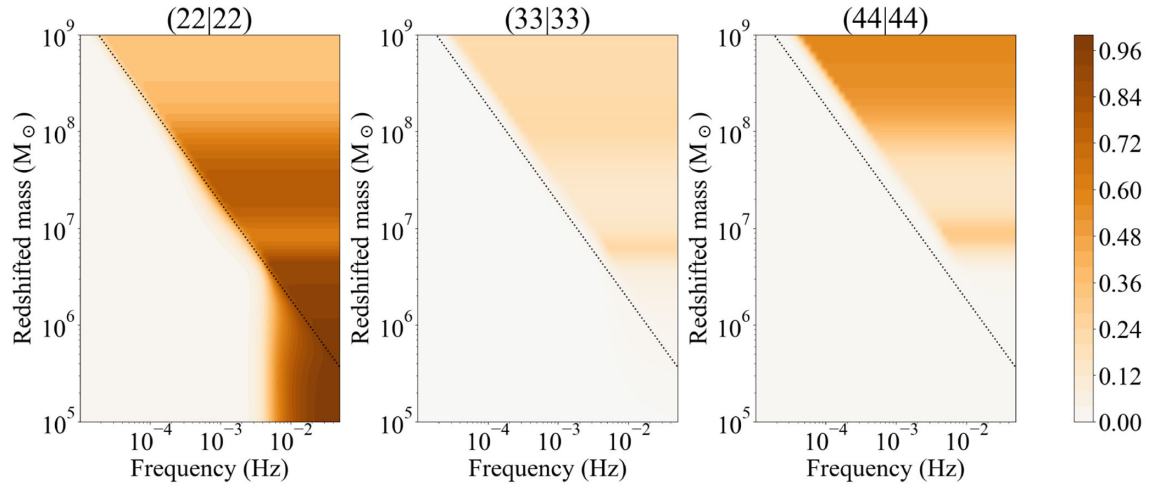
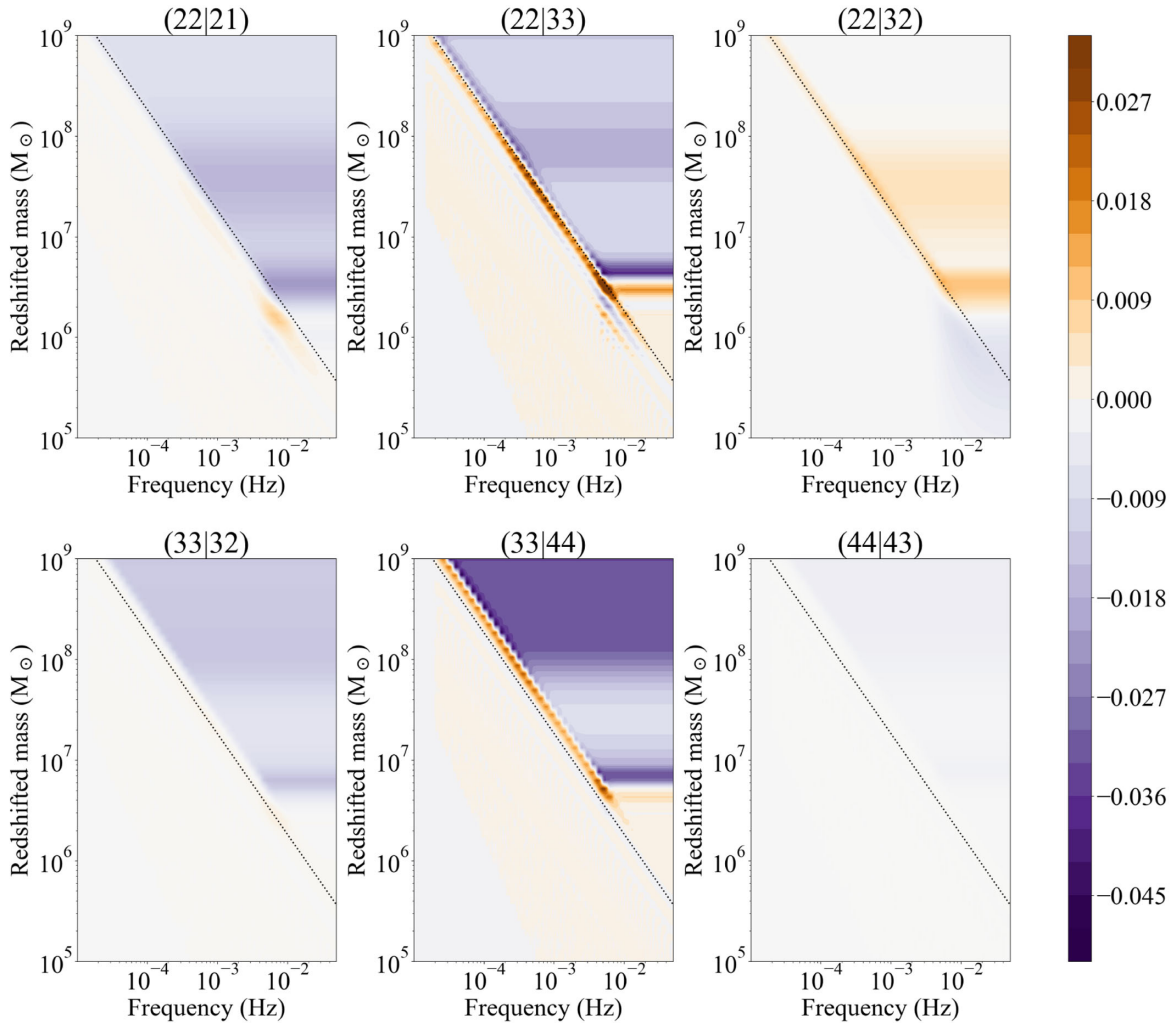
(a) Ratio contribution of the square terms ($lm|lm$) to the total squared SNR.(b) Highest ratio contributions of the cross-terms ($lm|l'm'$) to the total squared SNR.

FIG. 2. Cumulative contribution to the squared SNR of pairs of modes depending on the total redshifted mass. In the top row, we find the square terms with the largest ratios. The quadrupolar square term represents most of the contribution for low-mass sources, while it slowly decreases for more massive ones. The opposite happens for terms (33|33) and (44|44) albeit to a smaller extent. We show the cross-terms in the last two rows. They oscillate between $\sim 1\%$ and 5% , especially near the ringdown, illustrated with a dotted line for the mode (2, 2) as guidance. This percentage can represent an SNR of tens for a total SNR of thousands.

$10^8 M_\odot$, the contribution of the quadrupolar mode will no longer predominate. It can represent half of the mode (4, 4) and only a factor of 1.5 bigger than the mode (3, 3). This highlights the importance of including higher harmonics to describe SMBHB signals.

The second set of plots in Fig. 2(b) shows the pairs of modes that make the highest contributions to the squared SNR, representing approximately from 1% to 5% of the total. This percentage might seem small, but for SMBHBs with a total SNR of 1000, the cross-term contribution can be from 10 to 50. Those contributions are even higher than other square terms not shown here. Even though the contributions can be negative, we use their absolute value for comparison. The reason is that the sign depends on the extrinsic parameters and the mass ratio, which vary from source to source, but are fixed here to arbitrary values for the sake of an illustration.

The position of the mode's frequency peak relative to the LISA sensitivity curve is driven by the event mass. Therefore, the SNR of higher modes can become more relevant than the (2, 2) mode for large mass events.

B. Dependency on source mass and redshift

To represent the impact of the mass and the distance on the SNR, as mentioned in Sec. III, we show in Figs. 3(a) and 3(b) the contour plots for different modes depending on the total mass in the source frame. We choose a source with parameters listed in Table I for the sake of an illustration. However, the contributions would change with different parameters. Note that the lower bound of the total SNR is 10, which is the adopted threshold for SMBHB detection in LISA [1,2].

In these figures, we can see the high contribution of the square terms $|(l, m = l|l, m = l)|^{1/2}$ as well as the highest-contributing cross-terms such as $|(22|21)|^{1/2}$, $|(22|32)|^{1/2}$ and $|(33|32)|^{1/2}$. The complete array of pairs of modes is shown in Fig. 11 of Appendix. The pairs $|(22|33)|^{1/2}$ and $|(22|32)|^{1/2}$ exhibit a drop in SNR for systems with a total source mass around $10^6 M_\odot$, depending on the redshift. This effect results from the specific choice of parameters of the source we study. It is also visible in Fig. 1, where some cross-terms have a null cumulative contribution to the total squared SNR.

We also investigated the effect of the mass ratio on mode SNR by performing the same analysis, observing the SNR cancellation expected for equal-mass systems (see Appendix). Contour plots are shown in Fig. 10 also in Appendix.

V. EXAMPLE OF BAYESIAN ANALYSIS

A. Data and models

To understand the impact of higher modes in the parameter inference, we inject a SMBHB source signal with 6 modes. The parameters used to generate this source

were taken from the LISA Data Challenges (LDC) *Sangria*'s catalog [65], and detailed in Table II, where the subscript L in the extrinsic parameters means that it is expressed in the LISA frame at the time of coalescence t_c . We also use the redshifted chirp mass M_c and mass ratio $q \geq 1$ instead of the individual masses. The parameters χ_i represent the aligned (nonprecessing binaries) BH's dimensionless spin and D_L is the luminosity distance.

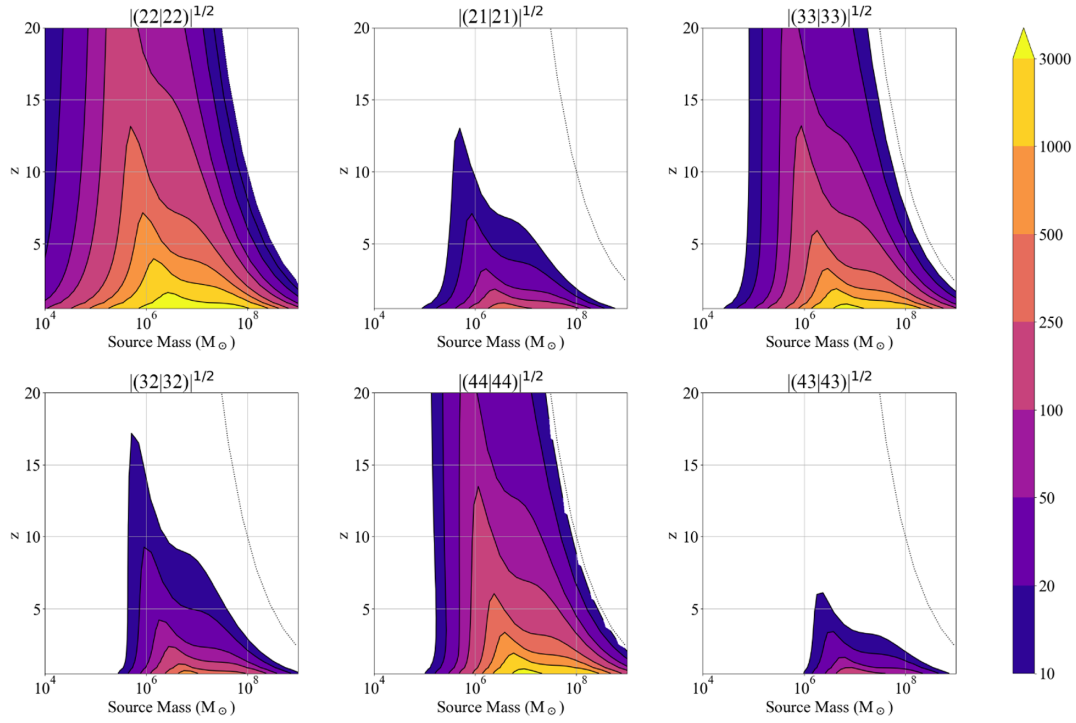
Two datasets were considered, one without noise and another with noise. The noise was generated with the PSD assumed in Sec. IV. Each dataset includes two TDI channels, A and E . In both cases, we restricted the frequency band to the interval $[10^{-5} - 5 \times 10^{-2}]$ Hz and ran a nested sampling algorithm to estimate the source parameters for various models. We chose the sampler *dynesty* [66], as it allows us to obtain approximate evidence estimates [see Sec. III A and particularly Eq. (22)]. As a consistency check, we also ran *ptemcee*, a parallel-tempered Markov Chain Monte Carlo ensemble sampler [67,68]. We obtained consistent results with the two samplers, with a slightly better convergence for *ptemcee* (which does not allow for direct evidence computation, however). We report *dynesty* results in the following.

We consider six models for the parameter estimation, where each one describes the waveform with a certain number of modes. The first model generates the gravitational signature with only the quadrupolar mode (2, 2). The other models (see Table III for models' definition) include an increasing number of higher harmonics, ranked by their SNR contribution ($lm|lm$), as observed in Fig. 3(a). This amounts to first selecting successive $(l, m = l)$ modes with increasing l and then the $(l, m = l - 1)$ modes.

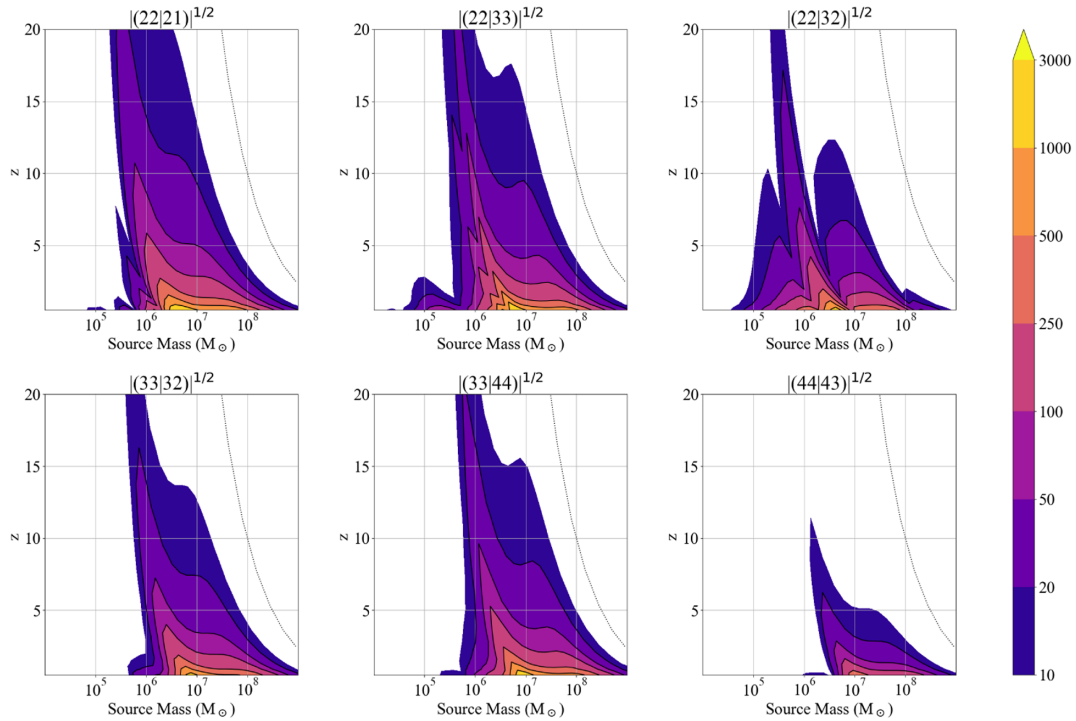
The priors we use are flat for all parameters in the intervals written in Table II, except for the chirp mass, which has a uniform prior in logarithmic scale. We use the whole physically allowed interval for the extrinsic parameters, while we use a raw estimation of the expected values for the intrinsic ones. We use a narrow prior for the coalescence time as it can easily be spotted in the detection process but with a difference of up to 600 seconds between the LISA and the SSB reference frame. Note that the polarization ψ is allowed to go from 0 to π (and not 2π) to prevent parameter degeneracy, given that in the antenna pattern, the polarization is always preceded by a factor 2 as shown in Eq. (9).

B. SNR and Bayes factor

Before presenting the Bayes factor and parameter estimation results, we discuss our expectations regarding the contribution of modes. Our example considers a source event randomly chosen from the LDC *Sangria*'s catalog with a SNR ~ 744 (see Table II). Converting the redshifted total mass to the source-frame total mass with a redshift of 4.3, we obtain the value of $2.28 \times 10^5 M_\odot$. Using Fig. 3,



(a) Contour plot with the contributions of square pairs of modes $|(\ell m|\ell m)|^{1/2}$.



(b) Contour plot with the highest SNR contributions of cross-terms $|(\ell m|\ell' m')|^{1/2}$.

FIG. 3. Contour plot for the root squared absolute cumulative value of contributions $|(\ell m|\ell' m')|^{1/2}$. The dotted line corresponds to the total SNR with a ≥ 10 threshold, plotted here for comparison. We remark the subdominant contribution of $|(\ell m|\ell m)|^{1/2}$ after the quadrupolar square term. Note the 0-contributions in some cross-terms due to representing a single system with fixed parameters. The complete array for all contributions is plotted in Fig. 11 of the Appendix.

TABLE II. Parameters of the SMBHB source with SNR 744 chosen from *Sangria*'s LDC catalog for redshifted mass without precession, along with the flat priors intervals used in the inference.

Parameter	Value	Prior
$M_c (M_\odot)$	857080.8396	$[10^4, 5 \times 10^7]$
q	2.7589	$[1, 10]$
χ_1	-0.5488	$[-1, 1]$
χ_2	0.2317	$[-1, 1]$
D_L (Mpc)	40084.6792	$[10^4, 5 \times 10^6]$
t_c (s)	0.0	$[-600, 600]$
β_L (rad)	-0.6186	$[-\pi/2, \pi/2]$
λ_L (rad)	2.2782	$[0, 2\pi]$
ϕ (rad)	0.2492	$[-\pi, \pi]$
Ψ_L (rad)	1.5158	$[0, \pi]$
ι (rad)	2.5969	$[0, \pi]$

TABLE III. Each model is indexed according to the number of modes included in the waveform generation.

Model	Modes (l, m)
M_1	(2, 2)
M_2	(2, 2), (3, 3)
M_3	(2, 2), (3, 3), (4, 4)
M_4	(2, 2), (3, 3), (4, 4), (2, 1)
M_5	(2, 2), (3, 3), (4, 4), (2, 1), (3, 2)
M_6	(2, 2), (3, 3), (4, 4), (2, 1), (3, 2), (4, 3)

with this source-frame mass, we can expect the term (22|22) to be the dominant contributor, whereas the rest of the SNR will come from (33|33) and (44|44) and the cross-terms (22|21) and (22|32). This is a simple estimation since the source parameters are not precisely the same as the ones listed in Table I.

The actual contribution from each pair of modes is plotted in Fig. 4, where we show the squared SNR contribution for each pair of modes for the full IMR signal of our example source. Both axes correspond to the modes, so the intersection represents the pair of modes $(lm|l'm')$. In the diagonal of the matrix, we find the square terms, while in the upper and lower triangle, we encounter the symmetric cross-terms. The value in each box indicates the squared SNR of each pair, whose absolute value is shown by the color bar. We can observe that the pair (22|22) indeed accounts for the largest contribution as expected, followed by the pairs (22|21), (33|33), and (22|33). The color gradient we observe when descending in the diagonal line is a consequence of the hierarchic ordering of modes. One striking difference with Fig. 3 is the high contribution of (22|21) when compared to (33|33), showing that the details of the mode contributions will vary for different sources with different intrinsic and extrinsic parameters.

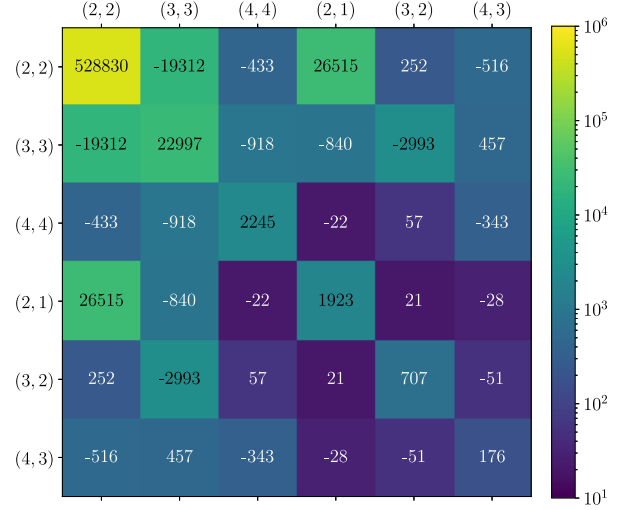


FIG. 4. Final value of squared SNR of each pair of modes $(lm|l'm')$ for our example. Note the positive values for the square terms and the decreasing values in the diagonal. In contrast, note the negative values for pairs with different “m” except for cross-terms (22|21) and (44|32). An interesting result is the relatively high value of (22|21), the second highest value.

Even if the final value of a mode’s SNR is small for the complete IMR signal, it does not mean that its impact is negligible in relative terms everywhere in frequency, particularly in the pre-merger phase. This feature can be observed in Fig. 5, where we represent the squared SNR absolute value as a function of the frequency. The pairs’ contributions are separated into groups to make the figures readable. We can see how some terms dominate in their group below a frequency that approximately corresponds to the merger, after which they later decrease. This happens for the pair (22|32) in the top-right figure or (21|32) in the bottom-left figure. Thus, statements about modes’ relative importance generally depend on the total accumulated SNR.

To quantify our ability to identify the presence of modes, we compute the Bayes factor using the *dynesty* sampler. We compare all the models M_k with $k = 1, \dots, 5$ with M_6 . The results gathered in Table IV show clear negative values for all of them. This means the model with 6 modes is preferred and describes the data better than all other models, as expected. Even the value of -84 (-100 with noise) shows a significant preference for the model M_6 over M_5 , where only the mode (4, 3) is absent. Thus, even the weakest modes in our setting should be identified as present in the data, which indicates that LISA observations will be capable of identifying waveform modes beyond the ones available in current waveform models. This result advocates using waveforms with all available higher harmonics to capture all the physics in LISA signals and further developing waveforms with higher mode contents. In the following section, we investigate whether ignoring these

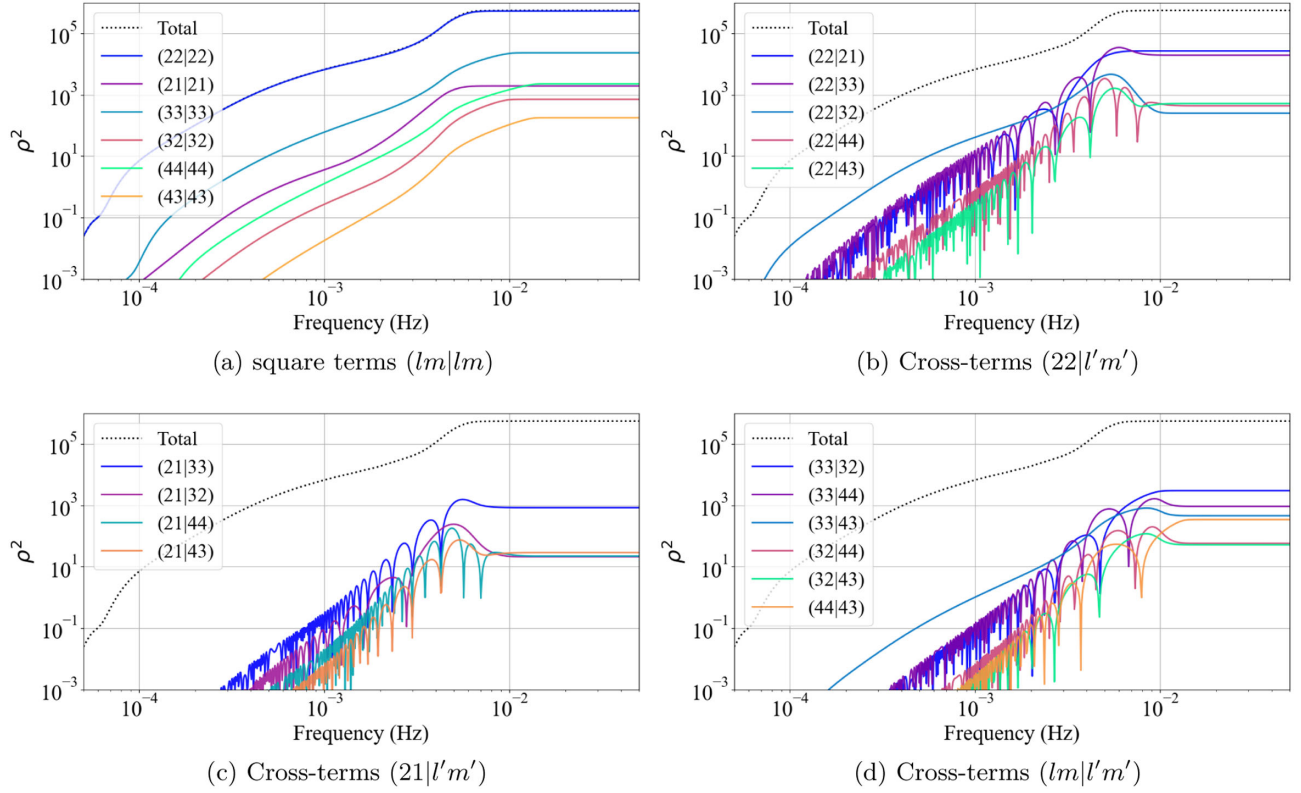


FIG. 5. Cumulative contribution to the squared SNR (in absolute value) of different pairs of modes depending on the frequency. The upper left figure shows the contribution of the square pairs, where we can see the quadrupolar making the higher contribution. All the other figures show cross-terms, from where we can highlight the contribution of the quadrupolar mode with higher modes ($22|l'm'$), which are predominant over other cross-terms.

weaker higher modes would produce biased parameter estimation results.

C. Posterior and parameter bias

To assess the impact of the mode's contribution on the estimation of the parameters, we show in Fig. 6 the posterior distribution of the logarithm of the chirp mass $\log M_c$, the mass ratio q and individual adimensional spins χ_1, χ_2 . Note that these parameters are relevant to the description of the remnant BH (see Sec. I and [19,26]). The complete array of the parameter's posterior distribution can be found in Fig. 12 in Appendix.

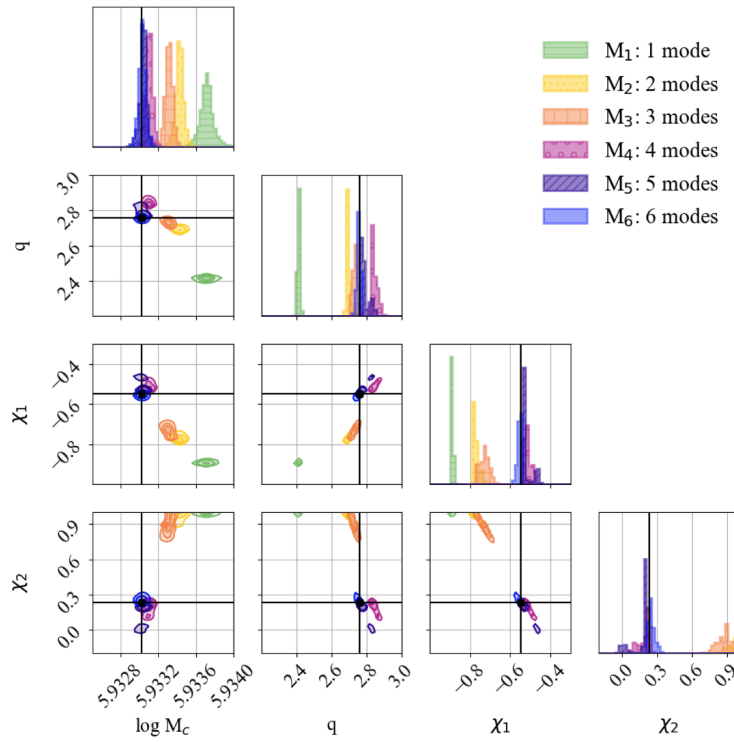
We show here the six models for comparison. Intersected black lines represent the true values, and each model's

TABLE IV. Bayes factor for all models compared to the injected model.

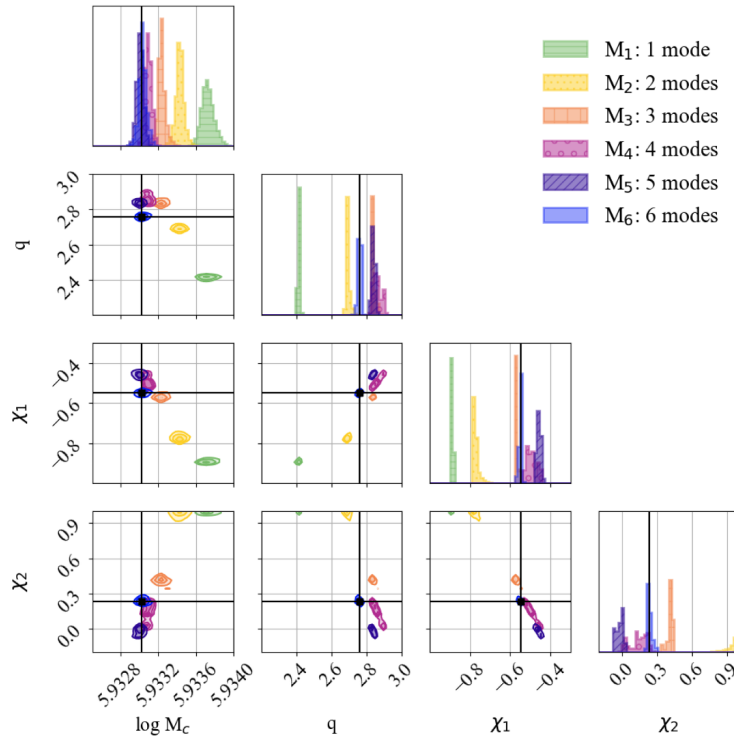
Bayes factor	Noiseless dataset	Noisy dataset
$\log(\mathcal{Z}_1/\mathcal{Z}_6)$	-6845	-6873
$\log(\mathcal{Z}_2/\mathcal{Z}_6)$	-976	-1015
$\log(\mathcal{Z}_3/\mathcal{Z}_6)$	-237	-259
$\log(\mathcal{Z}_4/\mathcal{Z}_6)$	-109	-134
$\log(\mathcal{Z}_5/\mathcal{Z}_6)$	-84	-100

posterior distribution is distinguished by the color code indicated in the legend. The parameter estimation for the model with only the quadrupolar mode (green color) leads to biased estimates. As we increase the number of harmonics in the models, the parameter posterior means get closer to the injected values. We observe that the posterior of model M_6 (in blue) is centered on the true value for all parameters, which is expected since the signal is injected and recovered with the same model. By comparing models with 3 and 4 modes (M_3 and M_4 , orange and pink, respectively), we observe a better estimation of spin parameters when the mode (2, 1) is included in the waveform (model M_4). This observation is consistent with the large relative contribution of $(22|21)$ indicated by Fig. 4. The explanation of the importance of the (2, 1) mode and whether this is generic or specific to our example source are left for future investigations. Note that PhenomHM generates the inspiral phase of the waveform with post-Newtonian approximation, however, the SNR is dominated by the late-inspiral-merger regime.

In Table V, we list the parameter's injected values and the estimated values, with models M_1 and M_6 in the presence and absence of noise. The super- and subscripts indicate the 68% confidence interval. In both cases, i.e., with and without noise, the model featuring only the dominant



(a) Marginalized posterior distribution on mass and spin parameters without noise.



(b) Marginalized posterior distribution on mass and spin parameters for noisy data.

FIG. 6. The six models with different numbers of modes are represented here. The posterior of the model M_6 (blue) finds the true values with high accuracy, while other models tend to induce biases, especially for spin parameters. In the bottom figure, we see how the introduction of noise does not decrease our ability to find the true values with 6 modes for this particular source.

TABLE V. Estimated value for models M_1 and M_6 , for the two datasets, without and with noise.

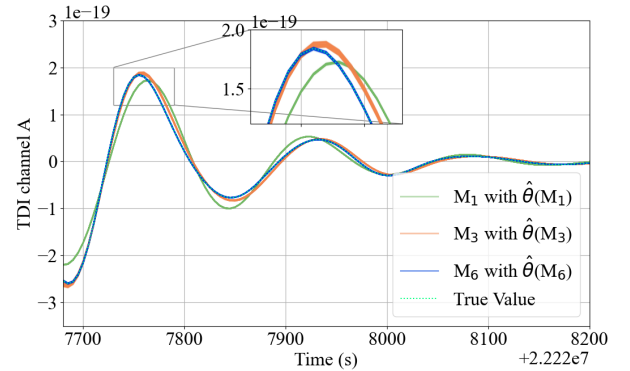
Parameter	True value	Estimated value with M_1 (noiseless)	Estimated value with M_6 (noiseless)	Estimated value with M_1 (with noise)	Estimated value with M_6 (with noise)
$\log M_c (M_\odot)$	5.93302	$5.93371^{+0.00019}_{-0.00016}$	$5.93303^{+0.00010}_{-0.00010}$	$5.93374^{+0.00019}_{-0.00016}$	$5.93304^{+0.00009}_{-0.00010}$
q	2.759	$2.411^{+0.012}_{-0.011}$	$2.751^{+0.019}_{-0.021}$	$2.414^{+0.012}_{-0.012}$	$2.759^{+0.013}_{-0.023}$
χ_1	-0.549	$-0.890^{+0.009}_{-0.007}$	$-0.559^{+0.016}_{-0.024}$	$-0.888^{+0.009}_{-0.008}$	$-0.549^{+0.011}_{-0.021}$
χ_2	0.232	$0.996^{+0.004}_{-0.019}$	$0.261^{+0.064}_{-0.043}$	$0.996^{+0.004}_{-0.018}$	$0.231^{+0.057}_{-0.030}$

quadrupole mode is inaccurate in finding the true values. In contrast, the estimated mean value with all modes is consistent with the injection. Surprisingly, the values obtained with the noisy data and the M_6 template appear closer to the injection than the ones obtained without noise. Note that the posterior distributions are not perfectly Gaussian and the mean value can be shifted due to tails. We performed another run with a different realization of the noisy data, obtaining similar results as the data without noise. The evidences we obtained with both noisy realization encompass the evidence obtained the data without noise. Therefore, we interpret this particular result as a fortuitous outcome of statistical fluctuations. In conclusion, for the medium-SNR and medium-mass case that we study (see Fig. 4), the absence of higher modes would already result in biased estimated values. These biases would only get worse for higher-SNR and higher-mass systems.

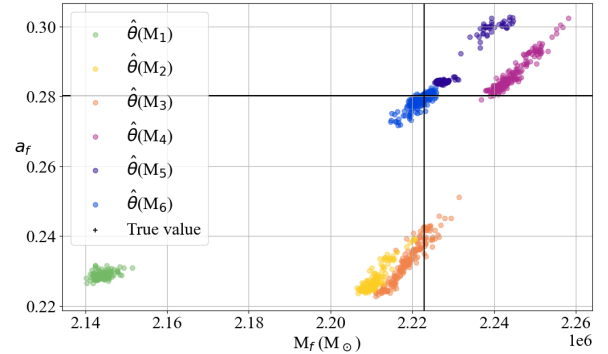
In Fig. 7, we illustrate the impact of the parameter biases on the reconstruction of the post-merger waveform. We randomly select 200 samples from posterior distributions obtained with models M_1 , M_3 , and M_6 and generate the waveform in the time domain with the same model. Fig. 7(a) serves as a visual representation of the amplitude and phase from the results obtained by each model. We cannot distinguish individual lines due to the small statistical error: the posteriors are centered around biased parameters but with a small dispersion. The waveform reconstruction would therefore be “confidently wrong”. Trying to infer a ringdown analysis with IMR information from biased analyses would presumably raise issues. If we compare M_3 and M_6 (in orange and blue, respectively), enlarging the image, we see visible differences in the post-merger phase. This feature, consistently with the significant Bayes factor for model M_6 over model M_3 , highlights the contribution of less dominant modes such as (2, 1), (3, 2), and (4, 3).

One of the tests looking for deviations from GR in ringdown signals consists in comparing the final mass and spin inferred from the ringdown signal with the values derived from the IMR posteriors using formulas fitted on numerical relativity; the consistency between the two estimates is the focus of the test. We do not perform a ringdown analysis here. Still, we illustrate in Fig. 7(b) how the parameter biases found in our IMR parameter estimation would translate into erroneous mass and spin. Using

the same fitted formulas as in PhenomHM [21] (see Eqs. (3.6)–(3.8) in [19]), we derived the final mass and spin for 2000 randomly distributed points within the posterior distribution for each model. The IMR parameter



(a) Waveform obtained with models M_1 , M_3 , and M_6 with parameters sampled from posterior distributions of the corresponding model.



(b) Derivation of final BH mass and spin from posteriors using models M_1 to M_6 compared with the true values.

FIG. 7. Illustrative effect of biased parameters in the ringdown. In the top figure, we see waveforms generated with the same model as the one used in the inference, obtained from 200 posterior samples for three models $\hat{\theta}(M_{1,3,6})$. We can see the consistency of the models with the injection, albeit its parameter bias. The bottom figure shows the mass and spin of the remnant BH derived from each set of parameters for all models (colored dots) and the true value (crossing black lines). This is a visual representation of the impact of biased parameters on the remnant BH.

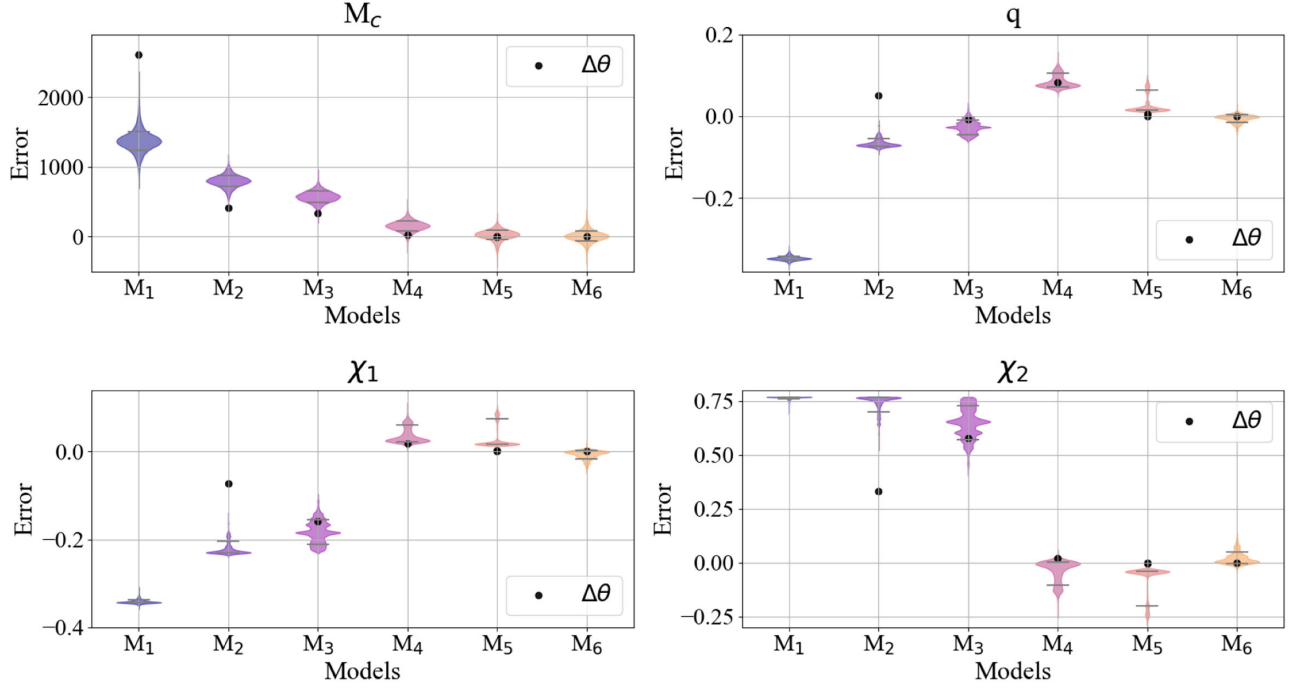


FIG. 8. Comparison of the error in the intrinsic parameters posterior distribution for all models (M_k , $k = 1, \dots, 6$ in color) with the error derived from the Fisher approximation (in black). Due to its large value, the Fisher bias for M_1 in the spins and mass ratio lies outside the plot. We observe consistency between the Fisher computation and posteriors for models M_k with $k = 3, 4, 5, 6$ as the bias decreases.

biases would translate into significantly biased M_f and a_f . In this figure, the addition of higher modes shows no clear trend for the final mass, but we can appreciate how the introduction of weaker higher modes helps to obtain a more accurate final spin for the remnant BH. Those models are M_4 , M_5 , and M_6 , in pink, purple, and blue, respectively. It is worth mentioning the scattered distribution of M_5 , which is the result of the bimodal distribution of individual spins (see Fig. 6).

D. Modeling error and SNR dependency

The magnitude of the bias on each parameter depends on the SNR and on the template waveform used for the inference, hence on the model M_k (with $k = 1, \dots, 6$). To properly analyze this issue, we introduce some definitions. Within the Fisher approximation, valid in principle in the high-SNR limit, the statistical error in each parameter σ_θ produced by the noise for a given waveform is determined as

$$\sigma_{\theta_i} = \sqrt{\Gamma_{ii}^{-1}}, \quad (27)$$

where $\Gamma_{ij} = \left(\frac{\partial h}{\partial \theta_i} \middle| \frac{\partial h}{\partial \theta_j}\right)$ is the Fisher information matrix. In this framework, the statistical error scales directly as SNR^{-1} .

The bias $\Delta\theta(M_k)$ or “modeling error” due to the use of an incorrect template is defined as [69]

$$\Delta\theta_i = \hat{\theta}_i^0 - \hat{\theta}_i^{\text{temp}} = \sum_j \Gamma_{ij}^{-1}(\theta_k) \left(\frac{\partial h_k}{\partial \theta_j} \middle| \delta h_k \right), \quad (28)$$

where k refers to model M_k , with $k = 1, \dots, 5$ and $\delta h_k = h_0 - h_k$ is the difference between the true waveform and the template waveform with model M_k . The template model corresponds to waveforms generated with the modes defined in Table III, so δh_k is simply the sum of the ignored modes.

If the statistical error σ_θ is larger than the error produced using an incorrect template $\Delta\theta(M_k)$, one can consider the bias irrelevant. On the contrary, if the statistical error is smaller than the modeling error, the waveform model is not sufficiently accurate to describe the data, and the bias becomes relevant.

We first check in Fig. 8 whether the bias observed in the posterior distributions is consistent with the value obtained from the approximate Fisher bias formula in Eq. (28). We show the bias from each model (black dot) and the error distribution obtained from the sampler (in colors) for intrinsic parameters. In the case of M_1 , we observe that the Fisher bias is much larger than the one found by sampling for all parameters, sometimes lying outside the plot. For adimensional spins, this may be because their values are limited to the interval $[-1, 1]$ in the sampler, whereas they are unconstrained in the Fisher matrix computation. In the case of M_2 , the opposite happens, and the posterior distribution exhibits a slightly larger error

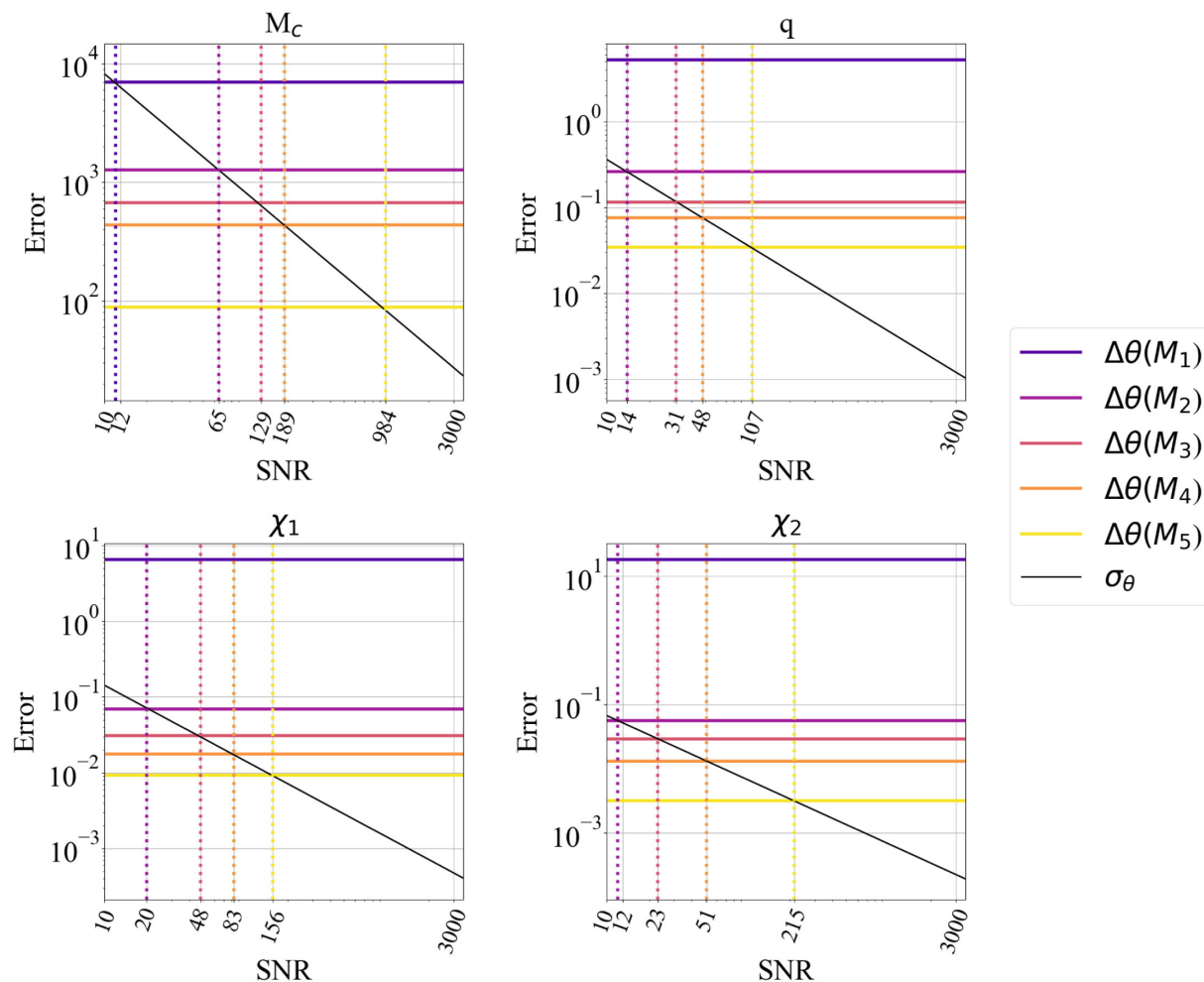


FIG. 9. Comparison of the statistical error (in black) and modeling errors (in colors) for relevant parameters in function of the SNR. When the statistical error becomes smaller than the modeling error, that model no longer fits the data well and biases the estimated parameters. The SNR for this transition is marked with dotted lines in the correspondent color for each incorrect template (models M_k , $k = 1, \dots, 5$).

than the bias predicted by the Fisher analysis. Overall, from model M_3 to M_6 , both errors become more and more consistent as parameter biases shrink. Thus, we can rely on the analytical bias obtained with the Fisher approximation for the intrinsic parameters. The same analysis for extrinsic parameters gave disagreeing results, with typically an overestimation of the bias with the Fisher approximation compared to the sampled posteriors; see Appendix for a short discussion on this matter.

A natural question arises about the minimum SNR at which higher modes become important in parameter estimation. In other words, given a certain SNR, how many modes do we need to describe the waveform adequately? One way to answer this question is by comparing the approximate statistical error for each parameter σ_θ with the systematic bias induced by an incorrect model $\Delta\theta(M_k)$. We perform this comparison in Fig. 9, where we show the errors for the intrinsic parameters as a

function of SNR. Varying the SNR amounts to changing the value of the luminosity distance D_L , leaving all other parameters unchanged. The Fisher bias in models M_1 and M_2 are inconsistent with posteriors, as observed in Fig. 8. For this reason, we will not discuss them, although they are plotted in the figure. The black diagonal line corresponds to the statistical error in the model parameters (σ_θ), and the color lines represent the modeling error produced by the wrong waveform template (M_k). With the same color code as the modeling error, we mark the value of SNR in dotted lines at which the modeling error becomes higher than the statistical error.

Figure 9's top left panel shows that the model M_3 (in pink), which includes modes (2, 2), (3, 3), and (4, 4), does not describe accurately enough the signal for sources with $\text{SNR} \geq 129$ since the chirp mass bias becomes more significant than the statistical error. Similarly, the model M_4 is sufficient until the SNR reaches 189 and M_5 until 984.

Since the SNR of the source we consider in this work is around 744, the value of the chirp mass inferred with the model M_5 (in yellow) should be within the estimated error. However, to accurately infer the other parameters, the model M_5 works until an SNR of 107 for the mass ratio, 156 for χ_1 , and 215 for χ_2 . Then, with an SNR of 744, this model will correctly estimate the chirp mass but induce biases for all other parameters. For consistency, we can look at the inferred values from the Bayesian analysis with M_5 in Fig. 6 and confirm this statement within the 68% confidence level.

This analysis does not derive a limit on the number of modes needed to describe an event observed by LISA. Still, it provides maximum SNR values to correctly estimate the parameters with a given model if less than these six modes are present in the waveform. Extrapolating from our example, we find that generally, sources detected by LISA with SNR of hundreds will require using waveforms with at least six modes to estimate all intrinsic parameters correctly. Note that the Fisher error in Eq. (27) and its scaling with SNR^{-1} are only approximate and are best valid for high-SNR and nondegenerate posteriors, so this estimate does not replace a complete parameter estimation study.

VI. CONCLUSIONS

We studied the contribution of IMR higher modes of a SMBHB source to the total SNR. We observed how this contribution depends on the event's redshifted mass through the observed frequency, directly related to LISA's response and sensitivity. We also showed that the cross-terms could contribute constructively or destructively to the total SNR, depending on the signal frequency and observational parameters. We presented a map guide of the relevance of each mode given the mass of an event. We highlighted the role of higher modes for SMBHBs with masses of the order of $10^8 M_\odot$. In LISA, large mass sources enhance the contribution of modes with higher frequencies so that the quadrupolar mode will no longer dominate.

To compare sensitivity performances, we defined different models, each including different harmonics. We could distinguish higher modes by comparing the Bayesian evidence for different models. In our example of a noisy signal with six modes, the model M_6 , which includes the same higher modes as in the injection, was the preferred one, as expected. The model M_6 showed a very significant Bayes factor compared to models with fewer modes. Furthermore, we found that the absence of modes in the waveform template can bias the parameter estimation for high SNR sources due to the nonorthogonality of the modes in the merger-ringdown phase. Biased binary parameters can lead to a biased inference of the remnant BH's mass and spin. This effect can corrupt the no-hair theorem test and lead to misinterpretations.

We could quantify the SNR needed to distinguish models by comparing statistical errors of the injected waveform parameters with the modeling errors produced using an incorrect template (M_k , $k = 1, \dots, 5$). In other words, given a certain SNR, we can specify the modes needed to infer the parameters without significant bias. This quantification depends on the actual waveform, which includes six modes in our analysis. In reality, such a situation is unlikely, as we expect more modes in the dynamics. Hence, this study does not derive a limit on the number of observable modes, which is still an open question that can be answered once more harmonics are implemented. However, our work demonstrates the need for higher modes in the waveform templates to perform accurate GW source characterization with LISA. Besides, featuring precession and eccentricity in the inspiral stage will also be necessary, while mode-mixing and nonlinearity [16,17,70,71] will become essential features in the ringdown.

The ability of LISA to identify different modes allows us to consider GR tests on more solid grounds, including the test of the no-hair theorem, which will be the subject of a forthcoming study.

ACKNOWLEDGMENTS

This work was supported by CNES, focused on LISA Mission. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon—France) for providing computing and data-processing resources needed for this work. *CP* acknowledges financial support from CNES Grant No. 51/20082 and CEA Grant No. 2022-039.

APPENDIX: COMPLEMENTARY RESULTS

To assess how the mass ratio affects the SNR, we show in Fig. 10 the contour plots representing mass ratio versus total source mass for a source with parameters given in Table II and redshift $z = 2$. The absence of contribution at leading PN order of modes with odd m for equal mass ratio (see, e.g., [72]) is illustrated here. We can also note the importance of higher modes for sources with masses around $10^6 M_\odot$, and mass ratios between 2 and 15.

For completeness, we also include here the square and cross-terms of the SNR contributions from Sec. IV B. As well as the marginalized posterior on all parameters for M_1 , M_3 , and M_6 with noisy data from Sec. V C.

1. Modeling error

If we compare the modeling error on the extrinsic parameters obtained from the Fisher information with the ones obtained from posterior distributions, see Fig. 13, we observe some disagreeing results. The model M_1 is the least accurate, as some Fisher bias values lie outside the range in the plots. As we introduce more modes, the difference between analytical and experimental results tends toward zero. Thus, we expect the analytical bias to

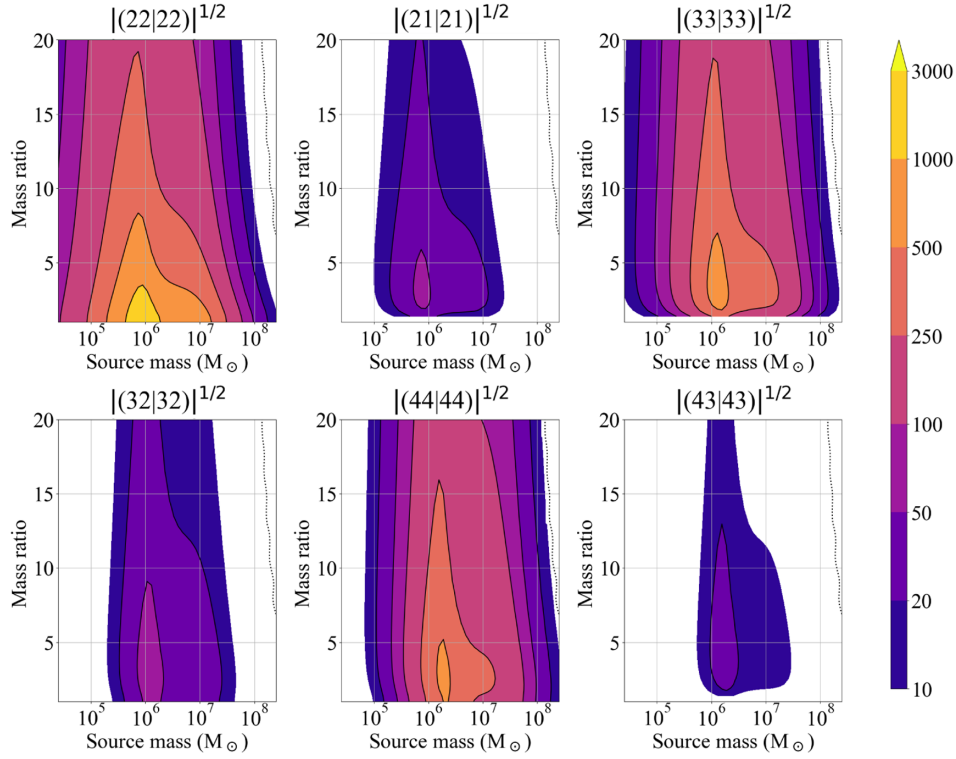


FIG. 10. Contour plots for mass ratio with respect to the total mass. The plot highlights the contribution of higher modes for sources of $\sim 10^6 M_\odot$ and mass ratios up to 15. SNR decreases slowly when increasing the mass ratio by a few units for a fixed mass.

be trustworthy for M_k with $k \geq 4$. As previously mentioned, the discrepancy could come from the constraint set on the parameter space explored by the sampler, which is absent from the Fisher derivation. This can be seen for example in the polarization ψ_L and phase ϕ , where some analytical points are outside the allowed range. Another possible explanation is the multimodality of some extrinsic parameters, such as the ecliptic latitude β_L in LISA's frame

and the source inclination ι , especially for M_1 . We also observe that most of the errors obtained with the Fisher information can change if we use for instance adimensional spin parameters such as χ_+ and χ_- instead of individual spins χ_1 and χ_2 . Furthermore, the numerics of Fisher matrices are notoriously delicate, and we leave the observed discrepancy between extrinsic errors for future investigation.

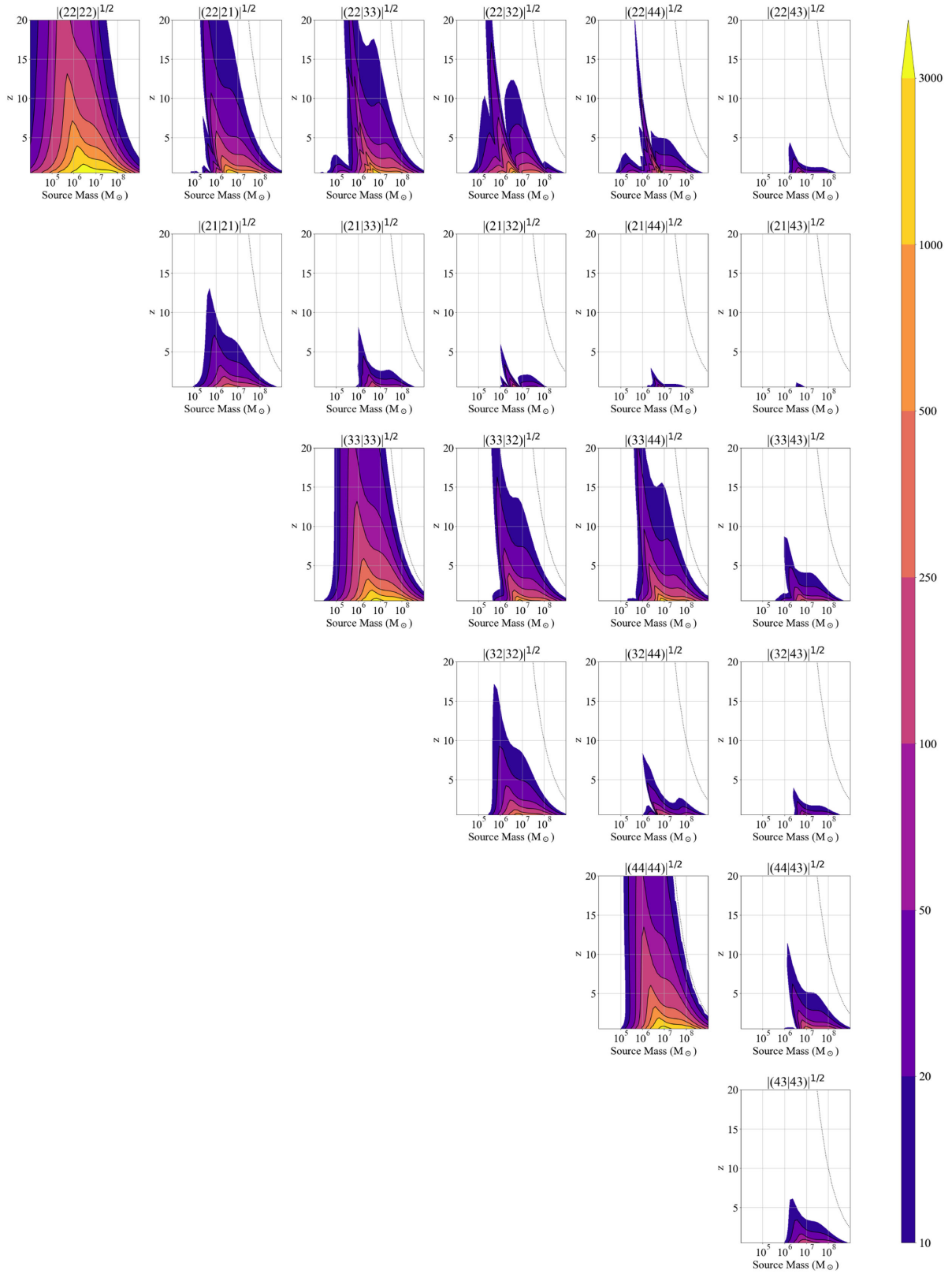


FIG. 11. SNR contribution for each pair of modes.

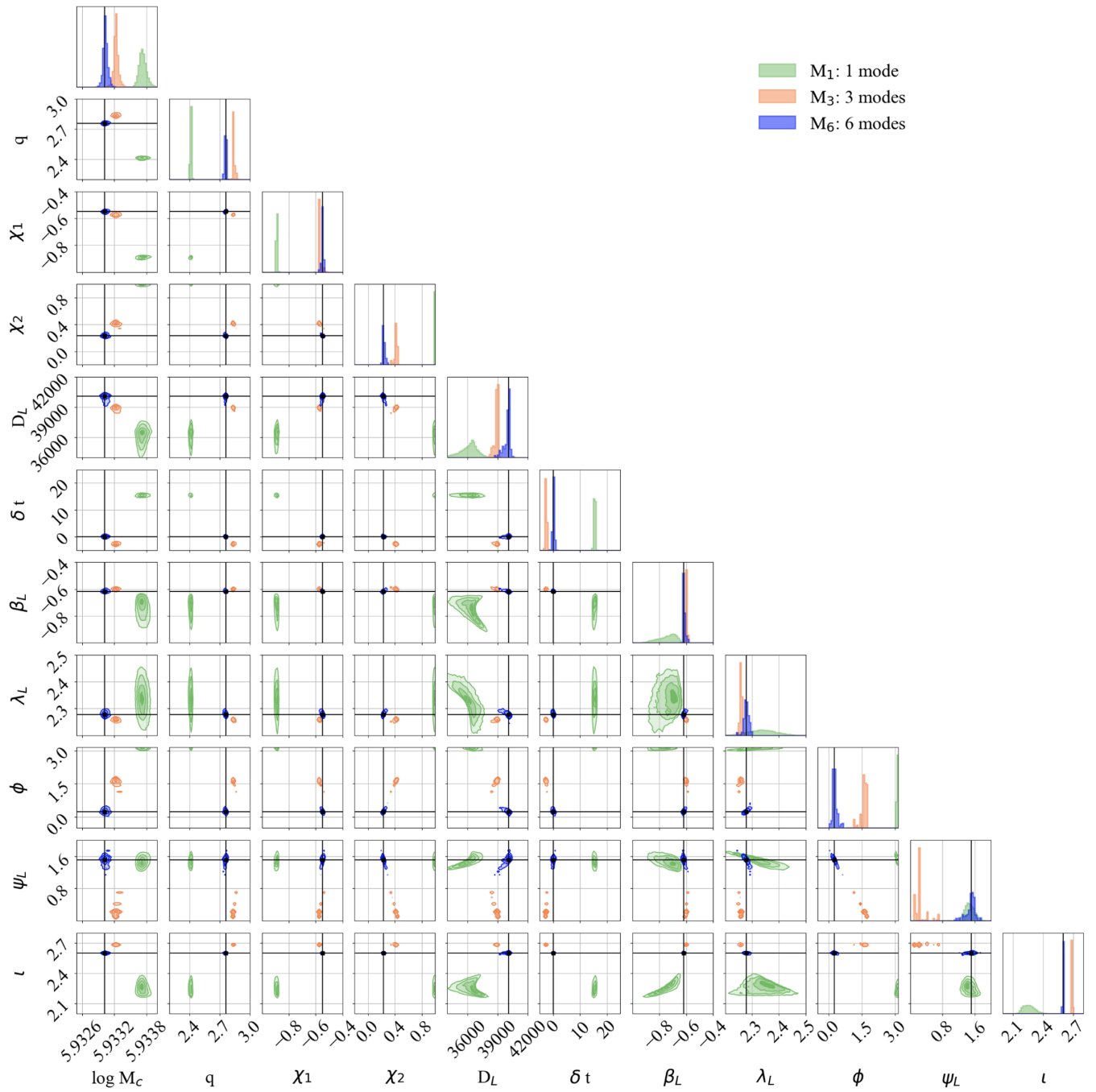


FIG. 12. Marginalized posterior distribution for all parameters. Here M_1 in green, M_3 in orange and M_6 in blue with true values represented with black lines. Note the high accuracy of the estimated mean value of all parameters found with the model M_6 .

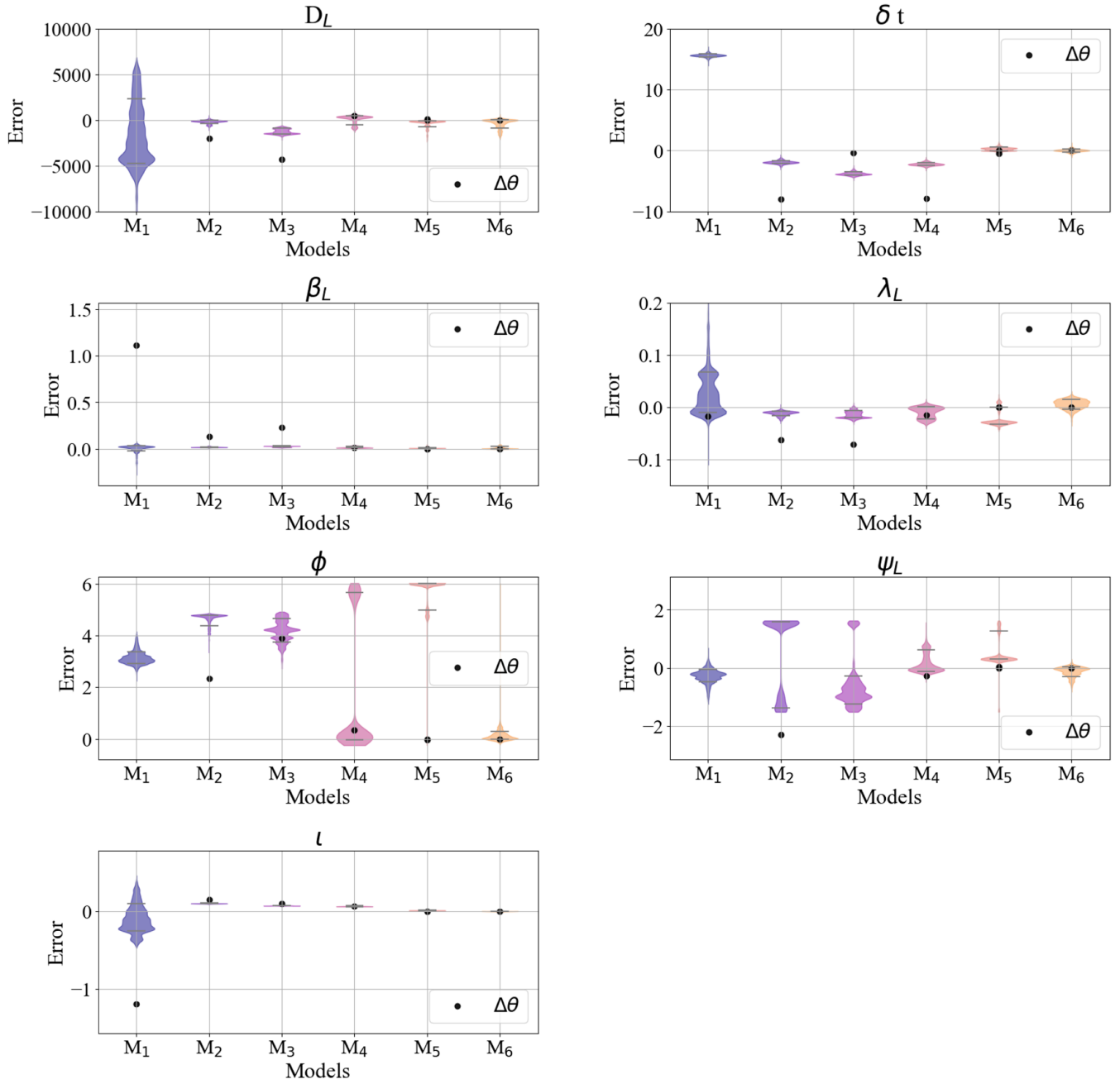


FIG. 13. Comparison of the error in the posterior distribution of extrinsic parameters for all models (M_k , $k = 1, \dots, 6$ in color) and modeling error (in black).

[1] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).

[2] LISA Science Study Team, LISA Science Requirements Document, Requirement Document Report No. ESA-L3-EST-SCI-RS-001-i1.0, ESA, 2018, <https://www.cosmos.esa.int/web/lisa/lisa-documents/>.

[3] P. Amaro-Seoane *et al.*, Astrophysics with the laser interferometer space antenna, *Living Rev. Relativity* **26**, 2 (2023).

[4] A. Sesana, Black hole science with the laser interferometer space antenna, *Front. Astron. Space Sci.* **8**, 601 (2021).

[5] B. Carter, Axisymmetric Black Hole Has Only Two Degrees of Freedom, *Phys. Rev. Lett.* **26**, 331 (1971).

- [6] S. L. Detweiler, Resonant oscillations of a rapidly rotating black hole, *Proc. R. Soc. A* **352**, 381 (1977).
- [7] O. Dreyer, B. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Black-hole spectroscopy: Testing general relativity through gravitational-wave observations, *Classical Quantum Gravity* **21**, 787 (2004).
- [8] W. H. Press and S. A. Teukolsky, Perturbations of a rotating black hole. II. Dynamical stability of the Kerr metric, *Astrophys. J.* **185**, 649 (1973).
- [9] S. Chandrasekhar and S. Detweiler, The quasi-normal modes of the Schwarzschild black hole, *Proc. R. Soc. A* **344**, 441 (1975).
- [10] H.-P. Nollert, Quasinormal modes: The characteristic ‘sound’ of black holes and neutron stars, *Classical Quantum Gravity* **16**, R159 (1999).
- [11] K. D. Kokkotas and B. G. Schmidt, Quasi-normal modes of stars and black holes, *Living Rev. Relativity* **2**, 2 (1999).
- [12] F. Echeverria, Gravitational-wave measurements of the mass and angular momentum of a black hole, *Phys. Rev. D* **40**, 3194 (1989).
- [13] E. Berti, V. Cardoso, and C. M. Will, Gravitational-wave spectroscopy of massive black holes with the space interferometer LISA, *Phys. Rev. D* **73**, 064030 (2006).
- [14] I. Kamaretsos, M. Hannam, S. Husa, and B. S. Sathyaprakash, Black-hole hair loss: Learning about binary progenitors from ringdown signals, *Phys. Rev. D* **85**, 024018 (2012).
- [15] I. Kamaretsos, M. Hannam, and B. S. Sathyaprakash, Is Black-Hole Ringdown a Memory of its Progenitor?, *Phys. Rev. Lett.* **109**, 141102 (2012).
- [16] E. Berti and A. Klein, Mixing of spherical and spheroidal modes in perturbed Kerr black holes, *Phys. Rev. D* **90**, 064012 (2014).
- [17] L. London, D. Shoemaker, and J. Healy, Modeling ringdown: Beyond the fundamental quasinormal modes, *Phys. Rev. D* **90**, 124032 (2014).
- [18] B. J. Kelly and J. G. Baker, Decoding mode mixing in black-hole merger ringdown, *Phys. Rev. D* **87**, 084004 (2013).
- [19] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal, *Phys. Rev. D* **93**, 044006 (2016).
- [20] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, *Phys. Rev. D* **93**, 044007 (2016).
- [21] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries, *Phys. Rev. Lett.* **120**, 161102 (2018).
- [22] S. Marsat and J. G. Baker, Fourier-domain modulations and delays of gravitational-wave signals, [arXiv:1806.10734](https://arxiv.org/abs/1806.10734).
- [23] S. Gossan, J. Veitch, and B. S. Sathyaprakash, Bayesian model selection for testing the no-hair theorem with black hole ringdowns, *Phys. Rev. D* **85**, 124056 (2012).
- [24] B. P. Abbott and others. (LIGO Scientific and Virgo Collaborations), Tests of General Relativity with GW150914, *Phys. Rev. Lett.* **116**, 221101 (2016).
- [25] S. Bhagwat and C. Pacilio, Merger-ringdown consistency: A new test of strong gravity using deep learning, *Phys. Rev. D* **104**, 024030 (2021).
- [26] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism, *Phys. Rev. D* **84**, 124052 (2011).
- [27] E. Berti, V. Cardoso, and A. O. Starinets, Quasinormal modes of black holes and black branes, *Classical Quantum Gravity* **26**, 163001 (2009).
- [28] V. Baibhav and E. Berti, Multimode black hole spectroscopy, *Phys. Rev. D* **99**, 024005 (2019).
- [29] V. Baibhav, E. Berti, and V. Cardoso, LISA parameter estimation and source localization with higher harmonics of the ringdown, *Phys. Rev. D* **101**, 084053 (2020).
- [30] M. Giesler, M. Isi, M. A. Scheel, and S. A. Teukolsky, Black Hole Ringdown: The Importance of Overtones, *Phys. Rev. X* **9**, 041060 (2019).
- [31] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Ringdown overtones, black hole spectroscopy, and no-hair theorem tests, *Phys. Rev. D* **101**, 044033 (2020).
- [32] E. Thrane, P. D. Lasky, and Y. Levin, Challenges for testing the no-hair theorem with current and planned gravitational-wave detectors, *Phys. Rev. D* **96**, 102004 (2017).
- [33] J. Meidam, M. Agathos, C. V. D. Broeck, J. Veitch, and B. Sathyaprakash, Testing the no-hair theorem with black hole ringdowns using TIGER, *Phys. Rev. D* **90**, 064009 (2014).
- [34] R. Abbott and others (LIGO Scientific Collaboration and Virgo Collaboration), Tests of general relativity with binary black holes from the second LIGO-VIRGO gravitational-wave transient catalog, *Phys. Rev. D* **103**, 122002 (2021).
- [35] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Testing the No-Hair Theorem with GW150914, *Phys. Rev. Lett.* **123**, 111102 (2019).
- [36] J. C. Bustillo, P. D. Lasky, and E. Thrane, Black-hole spectroscopy, the no-hair theorem, and GW150914: Kerr versus occam, *Phys. Rev. D* **103**, 024041 (2021).
- [37] G. Carullo, W. D. Pozzo, and J. Veitch, Observational black hole spectroscopy: A time-domain multimode analysis of GW150914, *Phys. Rev. D* **99**, 123029 (2019).
- [38] C. D. Capano and A. H. Nitz, Binary black hole spectroscopy: A no-hair test of GW190814 and GW190412, *Phys. Rev. D* **102**, 124070 (2020).
- [39] R. Cotesta, G. Carullo, E. Berti, and V. Cardoso, Analysis of ringdown overtones in GW150914, *Phys. Rev. Lett.* **129**, 111102 (2022).
- [40] B. P. Abbott and others (LIGO Scientific Collaboration and Virgo Collaboration), Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [41] V. Baibhav, M. H.-Y. Cheung, E. Berti, V. Cardoso, G. Carullo, R. Cotesta, W. D. Pozzo, and F. Duque, Agnostic black hole spectroscopy: Quasinormal mode content of numerical relativity waveforms and limits of validity of linear perturbation theory, [arXiv:2302.03050](https://arxiv.org/abs/2302.03050).
- [42] A. Buonanno and T. Damour, Effective one-body approach to general relativistic two-body dynamics, *Phys. Rev. D* **59**, 084006 (1999).

- [43] A. Buonanno and T. Damour, Transition from inspiral to plunge in binary black hole coalescences, *Phys. Rev. D* **62**, 064015 (2000).
- [44] A. Bohé *et al.*, Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors, *Phys. Rev. D* **95**, 044028 (2017).
- [45] E. Maggio, H. O. Silva, A. Buonanno, and A. Ghosh, Tests of general relativity in the nonlinear regime: A parametrized plunge-merger-ringdown gravitational waveform model, [arXiv:2212.09655](https://arxiv.org/abs/2212.09655).
- [46] R. Brito, A. Buonanno, and V. Raymond, Black-hole spectroscopy by making full use of gravitational-wave modeling, *Phys. Rev. D* **98**, 084038 (2018).
- [47] Lisabeta, for the moment only accessible for LISA Consortium members.
- [48] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, and S. Sinha, Higher harmonics increase LISA's mass reach for supermassive black holes, *Phys. Rev. D* **75**, 124002 (2007).
- [49] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, S. Sinha, and C. Van Den Broeck, Higher signal harmonics, LISA's angular resolution, and dark energy, *Phys. Rev. D* **76**, 104016 (2007).
- [50] M. Trias and A. M. Sintes, LISA parameter estimation of supermassive black holes, *Classical Quantum Gravity* **25**, 184032 (2008).
- [51] K. G. Arun, S. Babak, E. Berti, N. Cornish, C. Cutler, J. Gair, S. A. Hughes, B. R. Iyer, R. N. Lang, I. Mandel, E. K. Porter, B. S. Sathyaprakash, S. Sinha, A. M. Sintes, M. Trias, C. V. D. Broeck, and M. Volonteri, Massive black-hole binary inspirals: Results from the LISA parameter estimation taskforce, *Classical Quantum Gravity* **26**, 094027 (2009).
- [52] S. Marsat, J. G. Baker, and T. D. Canton, Exploring the bayesian parameter estimation of binary black holes with LISA, *Phys. Rev. D* **103**, 083011 (2021).
- [53] G. Pratten, A. Klein, C. J. Moore, H. Middleton, N. Steinle, P. Schmidt, and A. Vecchio, On the LISA science performance in observations of short-lived signals from massive black hole binary coalescences, *Phys. Rev. D* **107**, 123026 (2023).
- [54] J.-B. Bayle, O. Hartwig, and M. Staab, Adapting time-delay interferometry for LISA data in frequency, *Phys. Rev. D* **104**, 023006 (2021).
- [55] M. Tinto and J. W. Armstrong, Cancellation of laser noise in an unequal-arm interferometer detector of gravitational radiation, *Phys. Rev. D* **59**, 102003 (1999).
- [56] M. Vallisneri, Geometric time delay interferometry, *Phys. Rev. D* **72**, 042003 (2005); **76**, 109903(E) (2007).
- [57] J. W. Armstrong, F. B. Estabrook, and M. Tinto, Time-delay interferometry for space-based gravitational wave searches, *Astrophys. J.* **527**, 814 (1999).
- [58] C. Cutler and É. E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?, *Phys. Rev. D* **49**, 2658 (1994).
- [59] A. Klein, N. Cornish, and N. Yunes, Gravitational waveforms for precessing, quasicircular binaries via multiple scale analysis and uniform asymptotics: The near spin alignment case, *Phys. Rev. D* **88**, 124015 (2013).
- [60] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, LISA optimal sensitivity, *Phys. Rev. D* **66**, 122002 (2002).
- [61] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Brügmann, Inspiral, merger, and ringdown of unequal mass black hole binaries: A multipolar analysis, *Phys. Rev. D* **76**, 064034 (2007).
- [62] S. Bhagwat, M. Okounkova, S. W. Ballmer, D. A. Brown, M. Giesler, M. A. Scheel, and S. A. Teukolsky, On choosing the start time of binary black hole ringdowns, *Phys. Rev. D* **97**, 104065 (2018).
- [63] N. J. Cornish, Fast fisher matrices and lazy likelihoods, [arXiv:1007.4820](https://arxiv.org/abs/1007.4820).
- [64] N. J. Cornish, Heterodyned likelihood for rapid gravitational wave parameter inference, *Phys. Rev. D* **104**, 104054 (2021).
- [65] M. Le Jeune and S. Babak, LISA Data Challenge Sangria (LDC2a), Zenodo, [10.5281/zenodo.7132178](https://zenodo.org/record/7132178) (2022).
- [66] J. S. Speagle, DYNESTY: A dynamic nested sampling package for estimating Bayesian posteriors and evidences, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
- [67] W. D. Vousden, W. M. Farr, and I. Mandel, Dynamic temperature selection for parallel tempering in Markov Chain Monte Carlo simulations, *Mon. Not. R. Astron. Soc.* **455**, 1919 (2015).
- [68] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [69] C. Cutler and M. Vallisneri, LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms, *Phys. Rev. D* **76**, 104018 (2007).
- [70] M. H.-Y. Cheung, V. Baibhav, E. Berti, V. Cardoso, G. Carullo, R. Cotesta, W. Del Pozzo, F. Duque, T. Helfer, E. Shukla, and K. W. K. Wong, Nonlinear Effects in Black Hole Ringdown, *Phys. Rev. Lett.* **130**, 081401 (2023).
- [71] K. Mitman, M. Lagos, L. C. Stein, S. Ma, L. Hui, Y. Chen, N. Deppe, F. Hébert, L. E. Kidder, J. Moxon, M. A. Scheel, S. A. Teukolsky, W. Throwe, and N. L. Vu, Nonlinearities in Black Hole Ringdowns, *Phys. Rev. Lett.* **130**, 081402 (2023).
- [72] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, Erratum: Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiraling compact binaries: Ready-to-use gravitational waveforms, *Phys. Rev. D* **79**, 104023 (2009); **84**, 049901(E) (2011).