Parameter estimation of stellar-mass black hole binaries with LISA

Alexandre Toubiana,1,2 Sylvain Marsat,1 Stanislav Babak,1,3 John Baker,4 and Tito Dal Canton5

1APC, AstroParticule et Cosmologie, Université de Paris, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France
2Institut d’Astrophysique de Paris, CNRS & Sorbonne Universités, UMR 7095, 98 bis bd Arago, 75014 Paris, France
3Moscow Institute of Physics and Technology, Dolgoprudny, Moscow region, Russia
4Gravitational Astrophysics Laboratory, NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD 20771, USA
5Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Stellar-mass black hole binaries (SBHBs), like those currently being detected with the ground-based gravitational-wave (GW) observatories LIGO and Virgo, are also an anticipated GW source in the LISA band. LISA will observe them during the early inspiral stage of evolution; some of them will chirp through the LISA band and reappear some time later in the band of 3rd generation ground-based GW detectors. SBHBs could serve as laboratories for testing the theory of General Relativity and inferring the astrophysical properties of the underlying population. In this study, we assess LISA’s ability to infer the parameters of those systems, a crucial first step in understanding and interpreting the observation of those binaries and their use in fundamental physics and astrophysics. We simulate LISA observations for several fiducial sources, setting the noise realization to zero, and perform a full Bayesian analysis. We demonstrate and explain degeneracies in the parameters of some systems. We show that the redshifted chirp mass and the sky location are always very well determined, with typical errors below $10^{-4}$ (fractional) and 0.4 deg$^2$. The luminosity distance to the source is typically measured within $40-60\%$, resulting in a measurement of the chirp mass in the source frame of $O(1\%)$. The error on the time to coalescence improves from $O(1$ day) to $O(30$ s) as we observe the systems closer to their merger. We introduce an augmented Fisher-matrix analysis which gives reliable predictions for the intrinsic parameters compared to the full Bayesian analysis. Finally, we show that combining the use of the long-wavelength approximation for the LISA instrumental response together with the introduction of a degradation function at high frequencies yields reliable results for the posterior distribution when used self-consistently, but not in the analysis of real LISA data.

I. INTRODUCTION

Debuting with the first detection of a stellar-mass black hole binary (SBHB) in 2015 [1], the LIGO/Virgo collaboration has issued the first catalog of the gravitational-wave (GW) sources identified during the first and second observing runs (O1 and O2) [2–7] and three noteworthy detections from the third observing run (O3) [8–10]. These observations of GW in the 10–1000 Hz band, which include eleven SBHBs mergers and two binary neutron star (BNS) mergers, have inaugurated the era of GW astronomy and opened a new window to the universe, allowing to infer for the first time the properties of the population of compact binaries [11, 12] and providing new tests of general relativity (GR) [13–15].

The Laser Interferometer Space Antenna (LISA) [16], scheduled for launch in 2034, will observe GWs in a different frequency band (the mHz band) and, therefore, complement ground-based detectors. The strongest anticipated GW sources in the LISA data will be massive black hole binaries (MBHBs), with total mass in the range $10^4 - 10^7 M_\odot$ [17], and galactic white dwarf binaries (GBs). The latter are so numerous that they will form a stochastic foreground signal dominating over instrumental noise in addition to a smaller number ($\sim 10^4$) of individually resolvable binaries [18]. SBHBs with a total mass as large as those observed by LIGO and Virgo could also be detected by LISA during their early inspiral phase, long before entering the frequency band of ground-based detectors and merging [19]. SBHBs in the mHz band can be at very different stages of their evolution, ranging from almost monochromatic sources to chirping sources which leave the LISA band during the mission [19]. We focus on resolvable SBHBs, one of the best candidates for multiband observations [19, 20].

Although SBHBs are not LISA’s main target, the scientific potential of multiband observations with LISA and ground based detectors is considerable. These could be used to probe low-frequency modifications due to deviations from GR [19, 21–26] or to environmental effects [27–31], to facilitate electromagnetic follow up observations [27], or simply to improve parameter estimation (PE) over what is possible with ground-based interferometers alone [19, 24]. More precise measurements would improve the testing of competing astrophysical formation models. Several scenarios have been suggested for the formation of SBHBs, such as stellar evolution of field binaries versus dynamical formation channels [32, 33]. Moreover, the possibility that these black holes are of primordial origin [34] cannot be completely discarded. The various possible formation channels typically predict different distributions for the parameters of SBHBs, especially eccentricity and spin orientation/magnitude [35–37] providing discriminating power in astrophysical model selection. A recent study has suggested that already a few “special” events, binaries with high primary mass and/or spin, would have a huge discriminating power [38]. The specific impact of observations of SBHBs with LISA on astrophysical inference was considered in [39–43].

LISA will observe MBHBs somewhere between a few days and few months before their merger, which corresponds to the end of inspiral when the binary is evolving rapidly [17]. On the contrary, GBs are slowly evolving, almost monochromatic sources and they will remain in band during the whole
LISA mission [44]. Resolvable SBHB signals will fall in between these two behaviours: they are long lived sources but are not monochromatic and some SBHBs can chirp and leave the LISA band. In addition, all resolvable SBHBs are clustered at the high end of LISA’s sensitive band. Thus, SBHBs will produce very peculiar signals of great diversity. In this work we do not address the question of how to detect these sources, although it has been argued to be challenging [45]. Instead, we assume that we have at our disposal an efficient detection method, and concentrate on inferring the parameters of the detected signals. We also consider only one signal at the time, while in reality the SBHB signals will be superposed in the LISA data stream. The PE study presented here will also be valuable in building search tools as we discuss at the end of the paper.

Most previous studies of PE for SBHBs with LISA relied on the Fisher approach, and used simple approximations to LISA’s output response to GW signals. While a quick and efficient method for forecasting studies, the Fisher approach might not be suited for the systems with low signal to noise ratio (SNR) and non-Gaussian parameter distributions [46]. SBHBs signals are long lived and emit at wavelengths comparable to LISA’s size, this implies that we need to properly take into account the complete LISA response. As a result, the use of the long-wavelength approximation (also called low-frequency approximation) [47] might not hold and could seriously bias the PE [48, 49]. In this work we consider the full LISA response as described in [50, 51] and perform a full Bayesian analysis in 0-noise of all the systems we consider. We also provide a comparison to Fisher-matrix-based PE and briefly comment on the impact of using the long-wavelength approximation.

Despite the on-going effort to infer the astrophysical formation channel of the SBHBs observed by LIGO/Virgo, a huge uncertainty still remains. The situation will improve as we detect more signals. We expect that 3rd generation detectors (Einstein Telescope [52–54], Cosmic Explorer [55]) will be operational in parallel with LISA, with SNR figures reaching hundreds or thousands, thus significantly improving the PE over current observations. Given the current uncertainty in the population of SBHBs, we cannot reliably specify the properties of most detectable sources [19, 39, 45, 56–59]. In our study we focus on a fiducial system consistent with the population of currently observed SBHBs instead of working with a randomized catalog of sources. From there, we perform a systematic scan of the parameter space by varying few parameters at a time, investigating their qualitative impacts on the PE. We consider quasi-circular binaries consisting of spinning black holes, with the spins aligned or anti-aligned with the orbital angular momentum, merging no later than 20 years from the beginning of observations. We start with a GW150914-like system [1] and explore the parameter space by varying at most three parameters at a time. For each of these systems we infer the posterior distribution assessing correlations between parameters and the accuracy in measuring each parameter across the parameter space.

The paper is organised as follows. In section II we describe how we generate GW signals and our tools to perform PE. Then we give details on how we choose all the systems on which we perform PE in section III. Detailed description and analysis of PE results are given in section IV. There we also provide a comparison to a slightly modified version of Fisher matrix analysis and assess the validity of long-wavelength approximation for the PE. Finally, in section V we discuss the scientific opportunities offered by LISA observations of SBHBs in the light of our results.

II. ANALYSIS METHOD

A. Bayesian framework

Data measured by LISA ($d$) will consist in a superposition of GW signals ($s$) and a noise realisation ($n$): $d = s + n$. The instantaneous amplitude of a GW signal is much lower than noise making their detection very challenging. We use matched filtering as a main technique for detection, the main idea is to search for a specific pattern (GW template) in the data [60]. It is done by correlating the data with a set of GW templates in frequency domain ($h(f, \theta)$ which are functions of source parameters $\theta$). This correlation is given by the matched-filter overlap

$$ (d|h) = 4\Re \left( \int_0^{\infty} \frac{\tilde{d}(f) \tilde{h}^*(f)}{S_n(f)} df \right), \quad (2.1) $$

where $S_n(f)$ is the power spectral density (PSD) of the detector noise, assumed to be stationary. In this work, we use the LISA “Proposal” noise level as given in [16]. We do not discuss detectability of SBHBs in this paper, we assume that all sources discussed here could be detected and we concentrate on the parameter extraction/estimation. We should note that the detection itself could be a challenge, at least for the traditional method of template banks [45].

We work in a Bayesian framework for the parameter estimation, treating the set of parameters of the source, $\theta$, as random variables. The Bayes theorem tells us that given the observed data $d$, the posterior distribution $p(\theta|d)$ is given by:

$$ p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)} \quad . \quad (2.2) $$

In the right hand side of this equation $p(d|\theta)$ is the likelihood, $p(\theta)$ is the prior distribution and $p(d)$ is the evidence. As the noise and GW signal models will be fixed in this study, the evidence can be seen as a normalisation constant which does not need an explicit calculation. Assuming noise to be stationary and Gaussian, the likelihood is given by:

$$ \mathcal{L} = p(d|\theta) = \exp \left[ -\frac{1}{2} (d - h(\theta)) (d - h(\theta))^T \right] \quad . \quad (2.3) $$

In order to speed up the computation, we set the noise realization to zero as $n = 0$, so that $d = s$. The addition of noise to the GW signal is not expected to drastically affect the PE, leading at most to a displacement of the centroid of the posterior distribution within the confidence intervals (CI) (with
the probability defined by CI). Thus, the analysis of the posterior distribution itself should remain representative in presence of noise (still assuming Gaussianity). We consider only one source at a time and we neglect all possible systematic errors due to signal mismodelling: \( s = h(\theta_0) \) with \( \theta_0 \) the parameters of the GW source. Under these simplifications, the log-likelihood is given by (up to a normalization constant in \( \mathcal{L} \)):
\[
\log \mathcal{L} = -\frac{1}{2} (h(\theta_0) - h(\theta)) h(\theta_0) - h(\theta)
\] (2.4)

Since LISA will only observe the inspiral phase of these binaries, we expect that the dominant 22-mode is sufficient and we neglect the contribution of all other subdominant harmonics. We use the model called PhenomD [61, 62] to generate \( h_{2222} \) and compute the LISA response to generate the time delay interferometry (TDI) observables \( \mathcal{A}, \mathcal{E}, \mathcal{T} \) (see e.g. [63]) as described in [50, 51]. The three TDI observables constitute independent data sets, therefore the log-likelihood is actually a sum of three terms like Eq. (2.4), one per TDI observable.

Similarly to the treatment of galactic binaries, we parametrise the sources by their initial frequency and phase at the start of observations, instead of the time to coalescence and phase at coalescence (more suitable in LIGO/Virgo data analysis or for MBHBs with LISA). We define the initial time as the moment LISA starts observing the system. A system is characterised by (i) 5 intrinsic parameters: the masses \( (m_1 \text{ and } m_2) \), the GW frequency at which LISA starts observing the system \( (f_0) \) and spins \( (\chi_1 \text{ and } \chi_2) \); and (ii) 6 extrinsic parameters: the position in the sky defined in the solar system barycenter frame (SSB) \( (\lambda \text{ and } \beta) \), the polarisation angle \( (\psi) \), the azimuthal angle of the observer in the source frame \( (\phi) \), the inclination of the orbital angular momentum with respect to the line of sight \( (i) \) and the luminosity distance to the source \( (D_L) \).

We introduce a set of sampling parameters for which we believe the posterior distribution should be a simple function, i.e. close to either a uniform or Gaussian distribution. The choice of sampling parameters is made based on the post-netwonian (PN) expressions for GW [64–66]. We use the set of parameters \( \theta = (M_c, \eta, f_0, \chi_+, \chi_-, \lambda, \sin(\beta), \psi, \varphi, \cos i, \log_{10}(D_L)) \), where \( M_c = \frac{m_1 m_2}{m_1 + m_2} \) is the chirp mass, \( \eta = \frac{q}{(1+q)^{3/5}} \) is the symmetric mass ratio with \( q = \frac{m_1}{m_2} \geq 1 \) being the mass ratio, \( \chi_+ = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2} \) is the effective spin (often denoted \( \chi_{\text{eff}} \) in the literature) and \( \chi_- = \frac{m_1 \chi_1 - m_2 \chi_2}{m_1 - m_2} \) is an antisymmetric spin combination.

For the simulated data we assume a sampling rate of 1 Hz so the Nyquist frequency is \( f_{\text{Nyq}} = 0.5 \text{ Hz} \). When computing inner products given by Eq. (2.1), we generate templates from \( f_0 \) up to \( f_{\text{max}} = \min(f_{\text{Nyq}}, f_{\text{obs}}) \) where \( f_{\text{obs}} \) is the frequency reached by the system after the observation time \( T_{\text{obs}} \). We consider two mission durations: \( T_{\text{obs}} = 4 \text{yr} \text{ and } T_{\text{obs}} = 10 \text{yr} \). Details on fast LISA response generation and likelihood computation are given in [51].

Due to the high dimensionality of the problem, we need an efficient way to explore the parameter space. We do this by means of a Markov Chain Monte Carlo (MCMC) algorithm [67]. More specifically we designed a Metropolis-Hastings MCMC (MHMCMC) [68] for this purpose that we present next. A less costly alternative would be to use the parameter estimation based on the Fisher information matrix. We will show how one can modify the Fisher matrix to make it a robust PE tool in the following subsections. We exploit the metric interpretation of Fisher matrix in the MHMCMC, thus, we start by reviewing some basics on the Fisher matrix approach and delegate comparison with Bayesian PE to subsection IV C.

### B. Fisher matrix

In the Fisher matrix approach, the likelihood is approximated by a multivariate Gaussian distribution [46]:
\[
p(\theta | d) \propto e^{-\frac{1}{2} F_{ij}(\theta_0) \Delta \theta^i \Delta \theta^j}
\]
(2.5)

where \( \Delta \theta = \theta - \theta_0 \) and \( F \) is the Fisher matrix given by:
\[
F_{ij}(\theta) = \left( \partial_i h(\partial_j h) \right)_\theta
\]
(2.6)

where the partial derivative \( \partial_i \) denotes the derivative with respect to \( \theta_i \). Here again we actually have a sum of three terms, one for each TDI observable. We assume this to be implicit in the following. The inverse of \( F \) is the Gaussian covariance matrix of the parameters, it gives an estimate of the error on each parameter. This approach has been extensively used in the studies of LISA’s scientific capability thanks to its simplicity. However for systems with low SNR as the ones we consider, the Fisher approximation might not be valid [46] and we need to perform a full Bayesian analysis.

The Fisher matrix has an alternative interpretation: it can be seen as a metric on the parameter space associated with the distance defined by the inner product (2.1):
\[
||h(\theta + \delta \theta) - h(\theta)||^2 = \langle h(\theta + \delta \theta) - h(\theta), h(\theta + \delta \theta) - h(\theta) \rangle
\]
\[
\approx \langle \partial_i h(\partial_j h) \delta \theta^i \delta \theta^j \rangle
\]
\[
= F_{ij}(\theta) \delta \theta^i \delta \theta^j.
\]
(2.7)

We exploit this property in our MHMCMC sampler.

### C. Metropolis-Hastings MCMC

We sample the target distribution \( p(\theta | d) \) by means of a Markov chain, generated from a transition function \( P(\theta, \theta') \) satisfying the detailed balance condition:
\[
p(\theta | d) P(\theta, \theta') = p(\theta' | d) P(\theta', \theta).
\]
(2.8)

We build the transition function from a proposal function \( \pi \) such that \( P(\theta, \theta') = \pi(\theta', \theta) \alpha(\theta, \theta') \) where \( \alpha(\theta, \theta') \) is the acceptance ratio defined as:
\[
\alpha(\theta, \theta') = \begin{cases} 
\min(1, \frac{p(\theta' | d) \pi(\theta, \theta')}{p(\theta | d) \pi(\theta', \theta)}) & \text{if } \pi(\theta, \theta') \neq 0 \\
0 & \text{otherwise}.
\end{cases}
\]
(2.9)
It is easy to verify that $P$ satisfies the detailed balance condition for any choice of $\pi$. In practice, a jump from a point $\theta$ to $\theta'$ is proposed using the function $\pi(\theta, \theta')$ and the new point is accepted with probability $a(\theta, \theta')$. If the point is not accepted the chain remains at $\theta$. In both cases, we update the current state of the chain. By repeating this procedure, we obtain a sequence of samples of the target distribution. We see from the expression of the acceptance ratio that points with higher posterior are more likely to be accepted, thus the chain will tend to move towards regions of higher posterior exploring part of the parameter space compatible with the observed data. In theory, the chain should converge independently of the chosen proposal function and of where the chain is started but it could take an infinite time. Since for PE we are interested in high posterior regions we start the chain from the injection point, i.e. the maximum likelihood point. The maximum likelihood point coincides with the maximum posterior point if all priors are flat, but it is not true in general and the maximum posterior point can depend on the adopted prior distribution. Even though the posterior does not depend on the proposal, the convergence, efficiency and resolution of tails of distribution very strongly depends on a particular choice, the best proposal should closely resemble the target (posterior) distribution. Thus, most of the work goes into building an efficient proposal function. Note that for a symmetric proposal $(\pi(\theta, \theta') = \pi(\theta', \theta))$, the acceptance ratio is simply given by the ratio of the posterior distributions. This specific case is called Metropolis MCMC [69] and is the one we consider.

Our runs are done in two steps: we first run a short MCMC chain ($\approx 10^5$ points) to explore the parameter space and then use the covariance matrix of the points obtained from this chain to build a multivariate Gaussian proposal that we use in a longer chain. During the first stage, called burn-in, we use a block diagonal covariance matrix. We split the set of parameters in 3 groups: intrinsic parameters ($M_o, \eta, f_o, \chi_+, \chi_-)$, angles except the inclination ($\lambda, \sin(\beta), \psi, \varphi$) and inclination-distance ($\cos(\iota), \log(D_L)$). Each block is computed inverting the Fisher matrix of that group of parameters. This separation was based on the intuition, well verified in practice, that the stronger correlations are within these groups of parameters and is intended to avoid numerical instabilities that may arise when dealing with full Fisher matrices. Notice that by making this choice we do not discard possible correlations between parameters of different groups, we are simply not taking them into account when proposing points based on the Fisher matrix. If those correlations exist, they should appear in the resulting covariance matrix that we use to build a proposal for the main chain. Failing to include existing correlations could reduce the efficiency of our sampler in its exploratory, or burn-in phase: however the splitting can easily be adapted if needed.

We rotate the current state vector $\theta$ to the basis of the covariance matrix’s eigenvectors. In this basis the covariance matrix is diagonal, formed by the eigenvalues of the covariance matrix in the original basis. Because for some parameters the distribution is very flat, the eigenvalues of the covariance matrix predicted by Fisher can be very large, reducing the efficiency of the sampler. This is usually the case for poorly constrained but bounded parameters like $\cos \iota$ and spins. To avoid this issue we truncate the eigenvalues of the $(\cos \iota, D_L)$ matrices and define an effective Fisher matrix accounting for the finite extent of the prior on spins: $F_{\text{eff}} = F + F_p$. We take $F_{\text{eff}} = F_{\text{eff}}^p = \frac{1}{\sigma^2}$ with $\sigma = 0.5$. We motivate this choice in Sec. IV C.

In order to improve efficiency of the sampler in the event of complicated correlations between intrinsic parameters, we exploit the metric interpretation of Fisher matrix and occasionally recompute the covariance matrix for the first group of parameters with a given probability. By doing so we might violate the balance equation, but it is done only during the burn-in stage (exploration of the parameter space) and those points are discarded later from the analysis.

We test the convergence of the chains by running multiple chains with different random number generator seeds, checking that they all give similar distributions and computing the Gelman–Rubin diagnosis [70] for all the parameters. Potential scale reduction factors below 1.2, as the ones we get, indicate that the chains converged [70]. For each chain we accumulate $10^3 - 10^4$ independent samples (by thinning the full chain by the autocorrelation length) which takes 4 – 7 hours on a single CPU thanks to the fast likelihood computation and LISA response generation presented in [51].

III. SETUPS

A. Systems

We start by considering a system with masses and spins compatible with the first detected GW signal (GW150914) and label it Fiducial system. Its parameters are given in Table I along with its SNR assuming LISA mission lifetimes $T_{\text{obs}} = 4$yr and $T_{\text{obs}} = 10$yr. We give both the detector and source frame masses, related by $m = (1 + z)m_s$, where $z$ is the cosmological redshift and subscripts $s$ denote parameters in the source frame. We adopt the cosmology reported by the Planck mission (2018) [72]. Note that $T_{\text{obs}}$ is the mission duration, not the time spent by the system in the LISA band and we assume an ideal 100% duty cycle. The initial frequency is derived from the time to coalescence from the beginning of LISA observation that we fix to 8 years for the Fiducial system, thus for $T_{\text{obs}} = 4$yr the Fiducial system is observed for a fraction of its inspirals, while for $T_{\text{obs}} = 10$yr the system is observed for 8 years before exiting the LISA band when nearing coalescence. The sky location is given in the SSB frame. In the following subscripts $f$ refers to the Fiducial system.

We explore the parameter space of SBHBS by changing a few parameters of the Fiducial system at a time. We list all the systems we consider in the following subsections, specifying what are the changes with respect to the Fiducial system and the corresponding labels. For all systems we consider two possible mission duration quoted above (unless something else is specified).

In Table II we show the considered systems and their respective SNR. Note that we chose to use the LISA proposal noise level [16], which does not include a 50% margin in-
| $m_1$ (M$_\odot$) | 40 |
|-----------------|----|
| $m_2$ (M$_\odot$) | 30 |
| $m_{1,f}$ (M$_\odot$) | 36.2 |
| $m_{2,f}$ (M$_\odot$) | 27.2 |
| $t_c$ (yrs) | 8 |
| $f_0$ (mHz) | 12.7215835397 |
| $c$ | 0.6 |
| $x_1$ | 0.4 |
| $\lambda$ (rad) | 1.9 |
| $\beta$ (rad) | $\pi/3$ |
| $\psi$ (rad) | 1.2 |
| $\varphi$ (rad) | 0.7 |
| $\iota$ (rad) | $\pi/6$ |
| $D_L$ (Mpc) | 250 |
| $z$ | 0.054 |
| $T_{\text{obs}}$ (yrs) | 4 10 |
| SNR | 13.5 21.5 |

**TABLE I.** Parameters of a representative SBHB system labeled *Fiducial.* The masses and spins of this system are compatible with GW150914 [71]. The initial frequency is computed such that the system is merging in 8 years from the start of LISA observations. We consider two possible durations of the LISA mission: 4 and 10 years (in the latter case, the signal stops after 8 years at coalescence). Subscripts denote quantities in the source frame, bare quantities are in the detector frame. The sky location is given in the SSB frame.

To introduce to form the “Science requirements” SciRDv1 [73]. The SNRs would thus be significantly more pessimistic with SciRDv1. From the point of view of the PE, using one or the other noise model amounts to a constant rescaling of the noise PSD $S_n$, with the same effect as rescaling the distance to the source.

### 1. Intrinsic Parameters

Unless specified we take $t_c$ = 8 years and we compute the initial frequency corresponding to the chosen $t_c$. Changing $t_c$ (or equivalently $f_0$) amounts in shifting the GW signal in frequency and also defines its frequency bandwidth (within the chosen observation time).

- **Time left to coalescence at the beginning of LISA observations:**
  - Earlier: $t_c = 20$ years,
  - Later: $t_c = 2$ years

- **Chirp mass keeping the mass ratio unchanged:**
  - Heavy: $M_c = 1.5M_{c,f}$, $q=q_f$, $D_L = 445$ Mpc

  - Mass ratio, keeping the chirp mass unchanged:
    - $q3$: $q = \frac{m_1}{m_2} = 3$, $M_c = M_{c,f}$
    - $q8$: $q = \frac{m_1}{m_2} = 8$, $M_c = M_{c,f}$

  - **Spins:**
    - $Spin_{Up}$: $\chi_1 = 0.95$, $\chi_2 = 0.95$
    - $Spin_{Down}$: $\chi_1 = -0.95$, $\chi_2 = -0.95$
    - $Spin_{Op12}$: $\chi_1 = 0.95$, $\chi_2 = -0.95$
    - $Spin_{Op21}$: $\chi_1 = -0.95$, $\chi_2 = 0.95$

For the *Heavy* and *Light* systems we scaled the distance so that the SNR remains the same as for the *Fiducial* system in the case $T_{\text{obs}} = 10$ yr. Changing spins or mass ratio barely affects the SNR so we do not change the distance for those systems. Since the Earlier system merges in 2 years, increasing the observation time from 4 to 10 years has no impact.

### 2. Extrinsic parameters

Changes in extrinsic parameters do not affect the time to coalescence so all systems below have the same initial frequency as the *Fiducial* system.

- **Position on the sky (in the SSB frame):**
  - Polar: $\beta = \frac{\pi}{2} - \frac{\pi}{36}$, $\lambda = \lambda_f$
  - Equatorial: $\beta = \frac{\pi}{36}$, $\lambda = \lambda_f$

- **Inclination:**
  - Edge-on: $\iota = \frac{\pi}{2} - \frac{\pi}{36}$, $D_L = 150$ Mpc, $T_{\text{obs}} = 10$ yr

- **Distance:**
  - Close: $D_L = 190$ Mpc, $T_{\text{obs}} = 4$ yr
  - Far: $D_L = 350$ Mpc, $T_{\text{obs}} = 10$ yr
  - Very Far: $D_L = 500$ Mpc, $T_{\text{obs}} = 10$ yr

The drop in SNR being very large for an almost edge-on system, we decrease the distance of the *Edge-on* system to sustain a reasonably high SNR. For the same reason, we use only $T_{\text{obs}} = 10$ yr in this case. The purpose of variation in distance is to assess the impact of the SNR on the PE, all other things being equal. This also mimicks the effect of varying the noise level and the duty cycle. For the *Close* system we only consider the $T_{\text{obs}} = 4$ yr case and for the *Far* and *Very Far* systems we only consider the $T_{\text{obs}} = 10$ yr case.

### B. Prior

Regarding the Bayesian analysis, we take our fiducial prior to be flat in $m_1$ and $m_2$ with $m_1 \geq m_2$, flat in spin magnitude between $-1$ and $1$, flat in initial frequency, volume uniform for the source location and flat in the source orientation, its polarisation and its initial phase. For phase and polarization,
In the Flatmag system, varying a few parameters at once. We use the Flatsampl proposal noise level given in [16]. Different systems are derived from the Fiducial system, varying a few parameters at once. We use the Full response.

| System      | $T_{\text{obs}} = 4\text{yr}$ | $T_{\text{obs}} = 10\text{yr}$ |
|-------------|-------------------------------|-------------------------------|
| Fiducial    | 13.5                          | 21.1                          |
| Earlier     | 10.3                          | 17.2                          |
| Later       | 11.8                          | /                             |
| Heavy       | 12.8                          | 20.9                          |
| Light       | 14.1                          | 21.1                          |
| SpinUp      | 13.5                          | 21.1                          |
| SpinDown    | 13.5                          | 21.1                          |
| SpinOp12    | 13.5                          | 21.1                          |
| SpinOp21    | 13.5                          | 21.1                          |
| Polar       | 12.8                          | 20.1                          |
| Equatorial  | 14.9                          | 23.1                          |
| Edgeon      | /                             | 14.7                          |
| Close       | 17.8                          | /                             |
| Far         | /                             | 15.1                          |
| Very Far    | /                             | 10.6                          |

TABLE II. SNR of all systems considered, computed with the LISA proposal noise level given in [16]. Different systems are derived from the Fiducial system, varying a few parameters at once. We use the Full response.

since only $2\varphi$ (for a 22-mode waveform) and $2\psi$ intervene, we restrict to an interval of $\pi$. We obtain the prior probability density function (PDF) in terms of the sampling parameters by computing the Jacobian of the transformation from $(m_1, m_2, \chi_+, \chi_-, D_L)$ to $(M_q, \eta, \chi_+, \chi_-, \log_{10}(D_L))$ which gives:

$$p_f(\theta) = \begin{cases} N \frac{M_q g^{\frac{1}{15}} D_L}{\sqrt{1 - \eta^2}} & \text{if } 0.05 \leq \eta \leq 0.25, \\ 0 & \text{otherwise.} \end{cases}$$ (3.1)

Just like the evidence in Eq. (2.2), $N$ acts only as a normalisation constant and thus it is of no importance for us. The lower limit for $\eta$ was set according to the maximum mass ratio up to which PhenomD is calibrated ($q = 16$) [51, 62]. The range of chirp mass, initial frequency and distance are orders of magnitude larger than the posterior support so they do not affect the posterior. We label this prior as Flatphys and use it by default unless we specify some other choice, for example, we will consider two additional priors:

- **Flatmag**: uniform prior for the spins orientation and magnitude
- **Flatsampl**: flat prior in $M_q$, $\eta$ and $\log_{10}(D_L)$.

In the Flatmag case we start from a full 3D spin prior, uniform in [0,1] for the spins amplitude and uniform on the sphere for the spins orientation. We then consider only the spin projections on the orbital momentum, thus ignoring the in-plane spin components. The resulting prior is $p(\chi) = -\frac{1}{2 \ln |\chi|}$.

The Flatmag PDF is: $p(\theta) = p_f(\theta) p(\chi_1) p(\chi_2)$ where $p_f(\theta)$ is given in Eq. (3.1). This is the prior generally used by the LIGO/Virgo collaboration [2]. The Flatsampl PDF is given by:

$$p(\theta) = \begin{cases} N \frac{1}{\eta} & \text{if } 0.05 \leq \eta \leq 0.25, \\ 0 & \text{otherwise.} \end{cases}$$ (3.2)

This prior has no astrophysical motivation, we will use it to compare Fisher-based-PE to our full Bayesian inference in Sec. IV C.

We find it instructive to illustrate how the non-trivial priors look like. As we will show later in section IV, the chirp mass can be constrained by Bayesian analysis to a fractional error of $10^{-4}$, so we can impose a narrow constrain on the prior. The chirp mass is non-trivially coupled to other parameters (as we will show in great details in the following sections), and constraining it to the narrow interval introduces non-linear slicing in other parameters. Note that the imposed interval ($10^{-3}$ in relative terms) is still much broader than the typical measurement error. In Fig. 1 we display the Flatphys, Flatmag and Flatsampl prior distributions for $\eta$, $\chi_+$ and $\chi_-$ obtained by restraining the chirp mass to the specified interval. The remarkable features of our fiducial prior, the Flatphys prior, are the double peak at $\eta = 0.25$ and $\eta = 0$ and the bell-like shape for the $\chi_+$ and $\chi_-$ priors with almost zero support at extreme values. The Flatmag is singled out by the strong peak at $\chi_+ = 0$. As we will discuss in section IV, these non-trivial shapes of the priors can strongly affect the resulting posterior distributions in some cases.

C. LISA response

We briefly review how the LISA response is computed and refer to [51] for a more extensive discussion. We recall that LISA is composed of three spacecraft linked by lasers across arms of length $L = 2.5\text{ Gm}$. The TDI observables $A$, $E$ and $T$ are time-delayed linear combinations of the single-link observables $y_{slr}$ that measure the laser frequency shift due to an incoming GW across link $l$ between spacecrafts $s$ and $r$. We consider only the dominant 22-mode of the waveform, and following [51] we exploit a mode symmetry (for non-precessing systems) between $h_{22}$ and $h_{22,-2}$ to write the signal in terms of $h_{22}$ only. The single-link observables can then be written with a transfer function:

$$\tilde{y}_{slr} = F_{slr} h_{22}.$$ (3.3)

Denoting the amplitude and phase of the mode 22 as $\tilde{h}_{22} = A_{22}(f) e^{i \varphi_{22}(f)}$, working at leading order in the separation of timescales in the formalism of [50] the transfer functions are
given by (we set $c = 1$):

$$ T_{sli}^{22} = G_{sli}^{22}(f, t_f^{22}) \quad \text{(3.4)} $$

$$ G_{sli}^{22}(f, t) = \frac{i\pi f L}{2} \text{sinc} \left[ \pi f L (1 - k \cdot n_0) \right] $$

$$ \cdot \exp \left[ i\pi f \left( L + k \cdot (p_l^0 + p_r^0) \right) \right] $$

$$ \cdot \exp(2i\pi f k \cdot p_0) n_1 \cdot P_{22} \cdot n_l \quad \text{(3.5)} $$

$$ t_f^{22} = -\frac{1}{2\pi} \frac{d\Psi_{22}}{df}, \quad \text{(3.6)} $$

where $k$ is the unit gravitational wave propagation vector, $n_s(t)$ is the link unit vector pointing from the spacecraft $s$ to $r$, $p_0(t)$ is the position vector of the center of the LISA constellation in the SSB frame, $p_l^0(t)$ is the position of spacecraft $r$ measured from the center of the LISA constellation and $P_{22}$ is the polarisation tensor defined in [51]. We dropped the $t$ dependence in Eq. (3.5) for more clarity. The global factor $\exp(2i\pi f k \cdot p_0)$ is the Doppler modulation in GW phase and the $n_1 \cdot P_{22} \cdot n_l$ term is the projection of the GW tensor on the interferometer axes which is associated with the antenna pattern function. Note that both the Doppler modulation in phase and the antenna pattern are time dependent due to LISA’s motion, moreover, they depend on the sky position of the source, so that the annual variation in the phase and amplitude allows us to localize the GW source.

All our results are obtained using the full LISA response but we also assess the impact of using the long-wavelength approximation, a simplified version of the LISA response [47]. In this approximation, LISA is somewhat similar to two LIGO/Virgo-type detectors moving around the sun with the angle between arms being $\pi/4$. It is obtained by taking the $2\pi f L \ll 1$ limit in the LISA response so that:

$$ G_{sli}^{22}(f, t) = \frac{i\pi f L}{2} \exp(2i\pi f k \cdot p_0) n_1 \cdot P_{22} \cdot n_l \quad \text{(3.7)} $$

The sinc function appearing in Eq. (3.5) leads to a damping of the signal amplitude at high frequencies but in the long-wavelength approximation it is replaced by 1, leading to unrealistically high SNRs. To compensate for this, inspired by the computation of the sky averaged sensitivity [74] we introduce a degradation function that multiplies the GW amplitude:

$$ R(f) = \frac{1}{1 + 0.6(2\pi f L)^2} \quad \text{(3.8)} $$

To explore the validity of this approximation for SBHBs, we will compare the PE for the *Fiducial, Polar* and *Equatorial* systems using the full response and the long-wavelength approximation labeled *Full and LW* respectively. We will only use the leading order in the separation of timescales in the framework of [50], keeping in mind that corrections could be needed in general, in particular for almost-monochromatic signals.

### IV. PARAMETER ESTIMATION OF SBHBs

In order to test the performance of our MHMCMC sampler we compared it to our well tested parallel tempering MCMC code *ptmcmc* [1]. The quality of agreement between two distributions $p_1$ and $p_2$ can be quantified by computing the Kullback-Leibler (KL) divergence [75]:

$$ D_{KL} = \sum_\theta p_1(\theta) \log \left( \frac{p_1(\theta)}{p_2(\theta)} \right) \quad \text{(4.1)} $$

The KL divergence is zero if two distributions are identical. We computed $D_{KL}$ for the marginalised distributions of each parameter obtained with two samplers using the *Flatsampl* prior, and assuming 4 and 10 years of observation. Apart from $\psi$ and $\varphi$, all divergences were below 0.1 for 4 years of observation and below 0.01 for 10 years of observation showing a very good agreement between samplers. For $\psi$ and $\varphi$, less well determined in general, we get slightly higher values (up to $\simeq 0.6$) but still showing a good agreement. The results presented in this paper were obtained with our MHMCMC code and, unless otherwise specified, we use the *Flatphys* prior and the *Full* response. In our discussion, we use redshifted masses (rather than source frame) because they are directly inferred from the observed data. We give the full “corner plot” [76] for our fiducial system, comparing results for the two observation times in Fig. 2, this plot shows pair-wise correlation between parameters and the fully marginalised posterior for each parameter. The inset on the right top of the figure shows posterior distributions for $(m_1, m_2, \chi_1, \chi_2)$.

---

[1] https://github.com/JohnGBaker/ptmcmc
It would be difficult to represent the posterior distributions for all possible variations (deviations from the Fiducial) discussed above. Instead, we will summarize our results by underlining qualitative differences whenever we observe them and show comparative corner plots only when necessary. We start by discussing the structure of correlation between intrinsic parameters, move to extrinsic parameters, then compare the full Bayesian analysis with predictions from the Fisher matrix and, finally show the effect on the PE of the LW approximation to the response.

A. Intrinsic parameters

One of the main features appearing in Fig. 2 is the strong correlation between intrinsic parameters, in particular the one between $M_c$ and $\eta$ which is especially pronounced for 4 years of observation. The main reason for this degeneracy is the limited evolution of the GW frequency: in 4 years of observation the Fiducial system spans a very narrow range from $f_0 = 12.7$ mHz to $f_{4yr} = 16.5$ mHz.

To understand this issue better we consider a simplified problem by reducing the dimensionality: we fix $f_0, \chi+, \chi-$ and all extrinsic parameters to the “true” values and investigate the correlation between $M_c$ and $\eta$ for the Fiducial system. Keep-
FIG. 3. Analysis of the degeneracy between $M_c$ and $\eta$. The blue (orange) dots were obtained running a PE on the Fiducial system in the $T_{\text{obs}} = 4\,\text{yr}$ ($T_{\text{obs}} = 10\,\text{yr}$) case allowing only $M_c$ and $\eta$ to vary. The injection point is indicated by the black dashed lines. The orange dotted and the green dashed curves are given by (4.5) using the full PhenomD phase and the 1.5 PN truncation of the phase respectively. The red solid line was obtained by minimising the phase difference between the injected signal and templates over the whole frequency range spanned over 4 years of observation.

FIG. 4. Individual PN phase contributions $\Delta \Phi_n$ for the Fiducial system. The linestyle indicates the nature of the term, non-spinning (NS), spin-orbit (SO) or spin-spin (SS), while the color indicates the PN order. Note that these contributions are individually aligned at $f_0$, as explained in the text, and that interpreting the magnitude of these terms is not easy due to the alignment freedom. The vertical line shows $f_{4\text{yr}}$, and the greyed area shows the frequency range contributing less than 1 in SNR$^2$.

FIG. 5. Value of $\delta I \Psi$ along the curve in the $(M_c, \eta)$ plane that minimises it for $I = [f_0, f_{4\text{years}}]$ (blue) and $I = [f_0, f_{\max, LISA}]$ (orange). When LISA observes the system at low frequencies, the phase difference can be kept small over an extended region far from the injection. When LISA observes the chirp of the system, the phase difference becomes very large immediately at the vicinity of the injection point, reducing the extent of the degeneracy between $M_c$ and $\eta$.

where $v = (\pi M f)^{1/3}$ and where the $a_i$ are PN coefficients (we scaled out the leading term, so that $a_0 = 1$) that depend on the mass ratio and on the spins and can be separated between non-spinning terms (NS), spin-orbit terms (SO) and spin-square terms (SS). It was argued in [77] that most SBHBs would require terms up to the 2PN order. In Fig. 4, we show the magnitude of the know PN terms in the phasing for our Fiducial system. In general, the magnitude of phase contributions is delicate to interpret because of the alignment freedom, as some of the phasing error can typically be absorbed in a time and phase shift. In Fig. 4 we align contributions individually at $f_0$ with a zero phase and zero time according to (3.6). We see that, for $T_{\text{obs}} = 4\,\text{yr}$, PN orders beyond 1.5PN appear negligible due to

\[
\Psi(f) = \frac{3}{128\eta v^2} \sum_i a_i v^i, \quad (4.2)
\]
the limited chirping in frequency, while more terms become relevant for \(T_{\text{obs}} = 10\text{yr}\) where much higher frequencies are reached. We also gray out the area (\(f > 123\text{mHz}\)) beyond which the signal contributes less than 1 in SNR\(^2\), which we take as a somewhat conventional limit to indicate that ignoring the signal beyond this point would not affect the likelihood (2.4) and therefore the PE.

Since for \(T_{\text{obs}} = 4\text{yr}\) the GW frequency changes little from \(f_0\) to \(f_{3\text{yr}}\), we can Taylor expand the phase around \(f_0\):

\[
\Psi(f) \approx \Psi(f_0) + \frac{d\Psi}{df}\bigg|_{f_0} (f - f_0) + \frac{1}{2} \frac{d^2\Psi}{df^2}\bigg|_{f_0} (f - f_0)^2. \tag{4.3}
\]

We consider the inner product between the data \(d = A_d(f)e^{-i\Psi_d(f)}\) and the template \(h = A_h(f)e^{-i\Psi_h(f)}\). From our convention, the initial phase at \(f_0\) is the same, \(\Psi_d(f_0) = \Psi_h(f_0)\). The initial time is zero at \(f_0\), so the stationary phase approximation (3.6) gives: \(\frac{d\Psi}{df}\bigg|_{f_0} = 0\). The inner product becomes:

\[
(d|h) = 4\text{Re} \int_{f_0}^{f_{3\text{yr}}} df \frac{A_d(f)A_h(f)e^{i\Psi_d(f) - \Psi_h(f)}}{S_n(f)} \approx A_d(f_0)A_h(f_0) 4\text{Re} \left[ \int_{f_0}^{f_{3\text{yr}}} df e^{i\Psi_d(f) - \Psi_h(f)} \right]. \tag{4.4}
\]

where we used the fact that the amplitude is a slowly varying function of the frequency. The overlap is maximized when the template is in phase with the data, making the integral non-oscillating. In our quadratic approximation to the dephasing, this defines a curve in the \((M_c, \eta)\) plane according to

\[
\frac{d^2\Psi}{df^2}\bigg|_{f_0} = \frac{d^3\Psi(M_c, \eta_0)}{df^3}\big|_{f_0}. \tag{4.5}
\]

In Fig. 3 we display in blue (orange) dots points from the sampling in the \((M_c, \eta)\) plane in the \(T_{\text{obs}} = 4\text{yr} (T_{\text{obs}} = 10\text{yr})\) case and over-plot (in orange dotted line) the curve obtained by solving (4.5). The true (injection) value is indicated by black dashed lines. The curve closely follows the shape obtained from PE in the \(T_{\text{obs}} = 4\text{yr}\) case. The green dashed line is obtained by solving (4.5) truncating the phase to 1.5 PN (post-newtonian) order. We verified that adding higher PN terms does not produce any noticeable changes, which is in a good agreement with [77] and Fig. 4. We can even better reproduce the degeneracy by minimising the phase difference between injection and template over the whole frequency range spanned by the injected signal. More specifically, defining:

\[
\delta I(M_c, \eta) = \max_f |\Psi(M_c, \eta_0)(f) - \Psi(M_c, \eta)(f)|, \tag{4.6}
\]
for each value of $M_\odot$, we find $\eta$ such that $\delta_\eta \Psi$ is minimised. Note that all parameters are kept fixed in the dephasing measure we use here, in particular there is no optimization over a constant phase or time shift. The subscript $l$ stands for the frequency interval and we plot this curve for $I = [f_0, f_{\text{ears}}]$ in Fig. 3. One can see that we almost perfectly reproduce the shape of the correlation between $M_\odot$ and $\eta$ in the $T_{\text{obs}} = 4$yr case. In the $T_{\text{obs}} = 10$yr case, the system evolves until it leaves the band so it spans a broader frequency range. In Fig. 5 we show the value of the minimised $\delta_\eta \Psi$ for $I = [f_0, f_{\text{ears}}]$ and $I = [f_0, f_{\text{LISA}}]$ with $f_{\text{LISA}} = 0.5$Hz taken at the conventional end of the LISA frequency band. In practice, in the latter case the maximal dephasing occurs typically around $\sim 0.1$Hz. For the observation span of 4 years, we can find $\delta_\eta \Psi$ to be quite small ($<0.5$ rad) over a large range of $\chi$. As the bandwidth of the signal becomes broader, we cannot efficiently compensate for a change in the chirp mass by varying $\eta$ which results in the significant reduction of the degeneracy and great improvement in measuring those two parameters as seen by the narrower region covered by the orange dots on Fig. 3.

We now come back to the full Bayesian analysis and consider the estimation of the black holes spins. Following [62, 78] we introduce:

$$\chi_{\text{PN}} = \frac{1}{113} \left(94\chi_+ + 19q - 1 \frac{q - 1}{q + 1}\chi_+\right)$$

(4.7)

$$= \frac{\eta}{113} \left(113q + 75\chi_1 + \frac{113}{q} + 75\chi_2\right).$$

(4.8)

This term defines how the spins enter the GW phase at the leading (1.5 PN) order [64] and, therefore, should be determined the best from observation of SBHBs. We found this to be indeed well verified. As an illustration, we plot samples obtained for the $q_2$, $q_8$, SpinUp, SpinDown, SpinOp12 and SpinOp21 systems in the $T_{\text{obs}} = 10$yr case in Fig. 6. The points are the samples obtained in PE analysis and the lines show $\chi_{\text{PN}} = \chi_{\text{PN,0}}$ (fixing the mass ratio to its “true” (injection) value) for all those systems in the $(\chi_1, \chi_2)$ plane. In all these cases $\chi_{\text{PN}}$ is extremely well measured, within $10^{-2}$, but the combination of spins orthogonal to $\chi_{\text{PN}}$ is constrained only by the prior boundaries.

For slowly evolving binaries, only terms up to 1.5 PN in the GW phase are found to be relevant. At this order we expect a strong correlation between $\chi_{\text{PN}}$ and $\eta$: any change in $\chi_{\text{PN}}$ can be efficiently compensated by a change in $\eta$ such that the 1.5 PN term $(-16r + 1133\chi_{\text{PN}})\eta^{-3/5}$ is kept (almost) constant. We have verified this by plotting the curve $(-16r + 1133\chi_{\text{PN}})\eta^{-3/5} = \text{const}$ on top of the samples obtained for the Fiducial system and reproducing the shape formed by the posterior samples. Thus, we obtain and explained the three-way correlation between chirp mass, mass ratio and spins for the mildly relativistic systems spanning a narrow frequency band during the observation time. The increase in the observation time allows further chirping of the system, making the contribution of the 1PN and 1.5PN corrections in the phasing significant, thus breaking strong correlations between intrinsic parameters; however, the effect of higher-order PN terms is weak, as consistent with [77] and 4, which leads to only the spin combination $\chi_{\text{PN}}$ to be measured. This study also suggests that $\chi_{\text{PN}}$, being the most relevant mass-weighted spin combination, should be used as sampling parameter. The component of $\chi_{\text{PN}}$ along $\chi_+$ is always much larger than the one along $\chi_-$ (at least by a factor $94/79 \approx 5$), so we find that $\chi_+$ is also measured reasonably well and it is more relevant for the astrophysics of SBHBs. We will alternate between $\chi_{\text{PN}}$ and $\chi_+$ in our next discussions.

In order to further quantify the dependence of PE on the frequency bandwidth spanned by the signal during the observation time, we consider the Earlier, Fiducial and Later systems which differ in the initial frequency chosen so that the SBHBs merge in 20, 8 and 2 years respectively. We compute the KL divergence between the marginalised posterior and the marginalised prior for each intrinsic parameter, and report our findings in Table III. The larger values of $D_{\text{KL}}$ indicate that knowledge has been gained from the GW observations as compared to the prior. The results show a strong dependence on the observation time (therefore on the frequency bandwidth), especially for spins, for which the $D_{\text{KL}}$ varies by an order of magnitude. For the Earlier system we find that only the chirp mass measurement is truly informative.
systems and choices of prior. When observing the system at low frequencies, only $M_{c}$ shows a sensible deviation from the prior. The likelihood is informative on $\eta$ and $\chi_{+}$ (and $\chi_{PN}$) only for chirping systems. Different choices of prior give similar results.

|          | $Fiducial$ | $Earlier$ | $Later$ |
|----------|------------|-----------|---------|
|          | $M_{c}$ | $\eta$ | $\chi_{+}$ | $\chi_{-}$ | $\chi_{PN}$ | $M_{c}$ | $\eta$ | $\chi_{+}$ | $\chi_{-}$ | $\chi_{PN}$ | $M_{c}$ | $\eta$ | $\chi_{+}$ | $\chi_{-}$ | $\chi_{PN}$ |
| $Flatphys$ | $T_{obs} = 4yr$ | 3.6 | 0.4 | 0.2 | 0.1 | 0.3 | 2.7 | 0.04 | 0.04 | 0.03 | 0.04 | 6.1 | 1.7 | 3.1 | 0.4 | 3.6 |
|          | $T_{obs} = 10yr$ | 7.6 | 2.5 | 3.7 | 0.5 | 4.3 | 4.5 | 0.7 | 0.5 | 0.2 | 0.6 | / | / | / | / | / |
| $Flatmag$ | $T_{obs} = 4yr$ | 3.4 | 0.6 | 0.07 | 0.04 | 0.08 | / | / | / | / | / | / | / | / | / | / |
|          | $T_{obs} = 10yr$ | 7.5 | 2.5 | 4.4 | 0.4 | 4.8 | / | / | / | / | / | / | / | / | / | / |
| $Flatsamp$ | $T_{obs} = 4yr$ | 3.7 | 0.4 | 0.3 | 0.2 | 0.3 | / | / | / | / | / | / | / | / | / | / |
|          | $T_{obs} = 10yr$ | 7.3 | 3.2 | 3.7 | 0.5 | 4.4 | / | / | / | / | / | / | / | / | / | / |

TABLE III. Kullback-Leibler divergences between the marginalised posterior and prior distribution of the intrinsic parameters for different systems and choices of prior. When observing the system at low frequencies, only $M_{c}$ shows a sensible deviation from the prior. The likelihood is informative on $\eta$ and $\chi_{+}$ (and $\chi_{PN}$) only for chirping systems. Different choices of prior give similar results.

Note that the longer frequency evolution plays a bigger role than the SNR. For instance, $Later$ which leaves LISA after 2 years with $SNR = 11.8$ is more informative than $Earlier$ with $T_{obs} = 10yr$ which has an $SNR = 17.2$. We repeated this analysis using the $Flatmag$ and $Flatsamp$ priors for the $Fiducial$ system. For all choices of prior, the KL divergences are similar, proving that the $\eta, \chi_{+}, \chi_{-}$ distributions are prior dominated when observing slowly evolving systems. Notice that KL divergences for spins are slightly smaller when using the $Flatmag$ prior, meaning that the posterior is even more dominated by the prior. This is because the $Flatmag$ prior peaks strongly at $\chi_{+} = \chi_{-} = 0$ as discussed in section III B. Note that the values of $D_{KL}$ are always larger for $\chi_{PN}$ than for the other spin combinations, reflecting the fact that it is the best measured spin combination. Still, for systems evolving through a narrow frequency interval, the $\chi_{PN}$ distribution is also prior dominated. The effect of the prior is especially well seen for the $Fiducial$ system and $T_{obs} = 4yr$ in Fig. 7: the strong peak of $\eta$ at 0.25 is what we expect due to prior (see section III B). The same peak is also observed for the $Later$ system (predominantly due to low $SNR$) but $\eta$ is much better constrained for this system, the likelihood is informative enough to reduce the width of the distribution, but not large enough to suppress the prior. Let us reiterate this important finding: for intrinsic parameters beyond the chirp mass, the chirping (extend of the frequency evolution) of the observed SBHB has stronger influence on $PE$ than the $SNR$ or observation time per se.

We note that, although the frequency is slowly evolving, the signal is far from monochromatic unlike most of galactic binaries (e.g. double withe dwarf binaries). As an element of comparison, using the quadrupole formula to compute the frequency derivative at $f_{0}$, for the $Earlier$ system we find $f_{0} = 1.9 \times 10^{-11}$ Hz$^{2}$ which is 4 orders of magnitude higher than the fastest evolving galactic binaries [18]. Thus, despite the strong correlation between intrinsic parameters, the chirp mass is always well measured, with a relative error of order $10^{-4}$ for the $Earlier$ system when observing for 4 years and below $10^{-6}$ for the chirping systems. The tight constraint on $M_{c}$ leads to the banana-like shape correlation between $m_{1}$ and $m_{2}$ seen on the top right part of Fig. 2. As a result, we can determine individual masses (within $20-30\%$) only for chirping systems.

We give the $90\%$ CI for parameters of the $Fiducial$ system in Table IV. Whenever the marginalised distribution of a given parameter is leaning against the upper (lower) boundary of the prior as for $m_{1}$ ($m_{2}$) we define the $90\%$ CI as the value between the 0.1 and 1 quantile ($0$ and $0.9$). Otherwise, in all other situations we define the $90\%$ CI as the values between the 0.05 and 0.95 quantiles. In all cases we report the median as a point estimate.

Systems with a higher mass ratio ($q_{3}$ and $q_{8}$, keeping the chirp mass the same as for $Fiducial$) give an error on the chirp mass similar to the $Fiducial$ system, but the mass ratio is better determined. This is because, when keeping the chirp mass fixed, the PN expansion of the GW phase features negative powers of $\eta$, notably in the 1PN term. Moreover, what should matter is the derivative of the phase with respect to $\eta$ which contains only negative powers ($\eta^{-1/5}, \eta^{-2/5}$) which makes it more sensitive to the small mass ratio as compared to the equal-mass systems. For an observation time of 4 years, the uncertainty on individual masses is still of order of $100\%$, but for an observation time of 10 years, it reaches below $10\%$ and $1\%$ for the $q_{3}$ and $q_{8}$ system respectively.

We now discuss the effect of priors on PE for high-spin systems. Consider $SpinUp$ system in the $T_{obs} = 4yr$ case shown in Fig. 8. As discussed above, in this case we have the correlation between spin ($\chi_{PN}$) mass ratio $\eta$ and the chirp mass. In the posterior we observe the interplay between the $\eta$ and $\chi_{+}$ priors which push samples towards $\eta = 0.25$ and $\chi_{+} = 0$ and the likelihood which peaks at the true value of $\chi_{+}$ (0.95). This, together with the correlation between parameters, leads to the resulting posterior distribution which has double peak in $\eta$ and broad distribution for $\chi_{+}$ (the 2-D histogram is more informative). The distribution (overall) is shifted away from the true values (well evident in the right panel of Fig. 8), though they are still contained within $90\%$ CI. In the case of $T_{obs} = 10yr$, the system chirps, so the information provided by the likelihood dominates over the prior, therefore, this bias is corrected and most of the degeneracies (at least partially) broken. In general, the posterior for the spins for weakly chirping sys-
FIG. 8. The left panel shows the inferred distribution on $\eta$ and $\chi_+$ for the SpinUp system. Because of a “competition” between the prior and the likelihood the distributions of $\eta$ and $\chi_-$ peak away from the true value indicated by the black lines and the square. The $M_c$ distribution, not shown, is marginally affected. Because of the bias in $\eta$, the inferred distribution of masses is significantly biased. However with our definitions, the true value is within the 90% CI.

FIG. 9. The left (right) panel shows the inferred distribution on $\eta$ and $\chi_+$ ($m_2, \chi_1$) for the Fiducial system using the Flatphys, Flatmag and Flatsampl priors. Under the effect of the prior, the posterior distribution can be significantly shifted away from the true value indicated by black lines and squares.

tems are badly constrained and closely resemble the priors. For chirping systems, the determination of spins can be understood from Fig. 6. Because of the orientation of lines $\chi_{PN} = \text{const}$, $\chi_1$ is better constrained than $\chi_2$. As the mass ratio increases the slope of these lines changes, accentuating this difference. Spins of same (opposite) sign, are better (worse) determined as their magnitude increase because of the narrowing (broadening) of the allowed region. For the Fiducial system, the error on $\chi_1$ is quite large but we can infer that the spin is positive with 0 (and negative values) being outside the 90% CI given in Table IV. The effective spin ($\chi_+$) is measured within 0.1 for chirping systems.

All results (for masses) so far were for redshifted masses. Since $M_{c,s} = M_c/(1+z)$, we get:

$$\frac{\Delta M_{c,s}}{M_{c,s}} = \frac{\Delta z}{1+z} + \frac{\Delta M_c}{M_c}. \quad (4.9)$$

As we discuss in Sec. IV B 2, $D_L$ is typically measured within $40-60\%$ which implies a measurement of the redshift $z$ within $\sim 40-60\%$ (at the low redshifts we are considering,
TABLE IV. 90% CI on the parameters of the Fiducial system whose parameters are given in Table I using the Flatphys prior. For masses and distance, we give the relative errors. The redshifted chirp mass is extremely well determined for both mission durations but individual masses can be measured only if the mission is long enough and we can observe the system chirping. The measurement of the source frame chirp mass is worse, being dominated by the error on the distance measurement and therefore the redshift in (4.9). The error on individual masses is dominated by their intrinsic degeneracy. For chirping systems, we can also measure $\chi_{\text{PN}}$ which translates into a good constraint on the effective spin $\chi'$. The error on individual spins remains large for the chirping system, but we can start to constrain the spin of the primary black hole (in our example, excluding negative values). As a consequence of the overall improvement in the determination of the intrinsic parameters, the inference of the time to coalescence improves drastically. The sky location (given by Eq. 4.11) is very well determined for both mission durations, within the field of view of next generation electromagnetic instruments like Athena and SKA [79, 80].

$D_L$ and $z$ are linearly related). Thus, the second term on the right hand side of Eq. (4.9) is clearly subdominant and the error on the source frame chirp mass is dominated by the error on redshift, as the result we get

$$\frac{\Delta \mathcal{M}_c}{\mathcal{M}_c} \approx \frac{0.5 z}{1 + z}.$$ (4.10)

This error is typically of the order of a few percent for systems detectable by LISA (up to $z \sim O(10^{-1})$), which is better than current LIGO/Virgo measurements [2]. This estimate is in good agreement with the results presented in Table IV. The error for individual masses in the source frame are dominated by the error on the masses (like the second term on the left hand side of Eq. (4.9)) due to poorly constrained mass ratio.

The initial frequency is always extremely well determined, with relative errors below $10^{-5}$. Its determination improves for chirping systems due to reduction of correlation with other intrinsic parameters. The frequency of the system at the beginning of the LISA observations is coincidental, as it is directly linked to the time that is left for the system to coalesce. We apply the stationary phase approximation (3.6) to the full GW phase to infer $\iota$. This transformation involves all intrinsic parameters, so the error on $\iota$ is typically smaller for chirping systems. We find an error of the order of 1 day for systems far from merger, while for more strongly chirping systems $\iota$ can be determined to within 30 s.

We find that increasing or decreasing the total mass of the system (while preserving the SNR) as in the Heavy and Light systems has little consequence on the estimation of intrinsic parameters. The error on spins and symmetric mass ratio are the same as in the Fiducial case. The relative error on chirp mass and initial frequency is slightly smaller for lighter systems (factor $\approx 1.4$ between the Heavy and Light systems) because of the larger number of cycles. However, we do not find a simple scaling with the chirp mass of the system for a fixed level of SNR. In particular, we do not find the error on the chirp mass to scale with $\mathcal{M}_c^{1/3}$ as computed in [81, 82]. This was to be expected since as discussed in this section the error on intrinsic parameters depends crucially on the frequency interval through which we observe the binary.

Finally, the choice of prior only marginally affects the posterior distribution for chirping systems. On the other hand, it can have a significant impact for non-chirping systems as can be seen in Fig. 9. For example, the Flatmag prior completely dominates the posterior distribution of spins as the KL divergences suggested and shown in Fig. 9. Because of the noted correlations, the prior on spins propagates into determination of mass ratio and individual masses.

B. Extrinsic parameters

1. Sky location

The sky location of the source is very well determined and, except for systems close to the equator, its posterior distribution is very similar to a Gaussian unimodal distribution. We define the solid angle as in [47]:

$$\Delta \Omega = 2 \pi \sqrt{|\Sigma_{\lambda,\lambda}|(\Sigma_{\lambda,\lambda} \sin(\beta) \sin(\beta) - (\Sigma_{\lambda,\lambda} \sin(\beta))^2),}$$ (4.11)

where $\Sigma$ is the covariance matrix. This defines a 63% confidence region around the true location. The error for the Fiducial system, reported in Table IV, is below 0.4 deg$^2$ which is within the field of view of most planned electromagnetic instruments such as Athena and SKA. [79, 80]. With the exception of the Equatorial system, the sky position is constrained with a similar precision for all systems considered in this work. The good localization comes from the complicated modulations imprinted on the signal by the orbital motion of
FIG. 10. We plot $f|\tilde{h}^2|/S_n$ (normalised to its maximum value) as a function of the time to coalescence. This quantity is the integrand of the diagonal Fisher element (2.6) and indicates where in frequency the information on the parameter $\theta_i$ comes from. We show this quantity for $\lambda$ (the behaviour for $\sin(\beta)$ is very similar) and $\mathcal{M}_c$ (the behaviour for the other intrinsic parameters is very similar). We indicate the initial and end frequencies of the $\text{Later}$ (red), $\text{Fiducial}$ (red) and $\text{Early}$ (black) systems for $T_{\text{obs}} = 10\text{yr}$. As discussed in Sec. IV A, most of the information on intrinsic parameters comes from the high end of the frequency band, whereas the contribution to sky-parameters mainly comes from the low frequencies.

LISA, according to (3.5). To understand how the sky localization evolves as a function of the frequency band we observe a system, in Fig. 10 we plot $f|\tilde{h}^2|/S_n$ (normalised with respect to its maximum value) as a function of the frequency. The quantity $f|\tilde{h}^2|/S_n$ is the integrand entering the computation of the diagonal elements of the Fisher matrix (2.6) and indicates (for each parameter) the most informative frequency range. Using a logarithmic scale for frequencies, the factor $f$ ensures that we can visualize the contributions to the integral as the area under the curve (up to a normalisation factor). We also indicate the corresponding values of the time to coalescence $t_c$ on the upper x-axis. We indicate the initial (dashed line) and end (solid line) frequencies of the $\text{Later}$ (red), $\text{Fiducial}$ (red) and $\text{Early}$ (black) systems for $T_{\text{obs}} = 10\text{yr}$. The behaviour for $\sin(\beta)$ is similar to $\lambda$. For comparison we show the same quantity for $\mathcal{M}_c$, the behaviour for other intrinsic parameters is similar. As discussed in the previous section, most of the information on intrinsic parameters comes from high frequencies. On the other hand, there is more information on the sky location at low frequencies, where a given range of frequencies corresponds to more orbital cycles of the LISA constellation. However, this is to be balanced with the narrower frequency range spanned by systems evolving at lower frequencies, for a fixed observation time. For this reason, the $\text{Later}$ system gives a better localization than the $\text{Earlier}$ system even in the $T_{\text{obs}} = 10\text{yr}$ case as reported in Table V (0.05 against 0.2 deg$^2$).

We can distinguish two main effects in (3.5) informing us about the sky localization: the time-dependency (through $t_f$, see (3.6)) of the response reflects the orbital cycles of LISA, and the Doppler modulation $\exp(2\text{im}(f/k \cdot \mathbf{p}_o)$ of the phase. The Doppler modulation shows this time dependency, but also scales with $f$, so this term is larger for chirping signals reaching high frequencies. We find a better sky localization for lighter systems: $\Delta \Omega_{\text{Light}} < \Delta \Omega_{\text{Fiducial}} < \Delta \Omega_{\text{Heavy}}$ (ranging from 0.1 to 0.3 deg$^2$ in the $T_{\text{obs}} = 4\text{yr}$ case and from 0.02 to 0.05 deg$^2$ in the $T_{\text{obs}} = 10\text{yr}$ case). This is a result of keeping fixed the time to coalescence $t_c$ and the SNR (by adjusting the dis-

| $T_{\text{obs}}$ | \(\Delta \Omega\) (deg$^2$) |
|---|---|
| 4yr | Fiducial | 0.18 |
| 10yr | Fiducial | 0.3 |
| 4yr | Earlier | 0.7 |
| 10yr | Earlier | 0.2 |
| 4yr | Later | 0.05 |
| 10yr | Later | 0.02 |
| 4yr | Polar | 0.14 |
| 10yr | Polar | 0.24 |

TABLE V. Solid angle around the injection point corresponding to a 63% confidence region, computed with (4.11). The sky localization is slightly better for the Polar sky position $\beta = \pm \pi/2$, but much worse for the $\text{Equatorial}$ sky position $\beta = 0$. The sky localization is better for the $\text{Later}$ system than for the $\text{Earlier}$ system (despite a lower SNR in the $T_{\text{obs}} = 10\text{yr}$ case) due to the broader frequency range spanned during its observation.

FIG. 11. Inferred distribution on the angles parametrising the position of the source for the Polar, Fiducial and Equatorial systems, with $T_{\text{obs}} = 4\text{yr}$. As explained in the main text, to avoid coordinate effects near the pole we do not compare the angles in the SSB frame ($\lambda, \beta$) but transformed angles ($\mu, \gamma$) defined by placing the injection point at the equator in each case (note that the scale of the two axis is not the same). The injection corresponds to $\mu = \gamma = 0$ as indicated by the black solid lines and squares. $\mu$ is equally well recovered in the three cases. For the $\text{Equatorial}$ system we find a tail extending to the position $\beta \rightarrow -\beta$.

| $\mu$ (rad) | $\gamma$ (rad) |
|---|---|
| $\text{Fiducial}$ | $\text{Polar}$ |
| $\text{Equatorial}$ | $\text{Equatorial}$ |

| $\mu$ (rad) | $\gamma$ (rad) |
|---|---|
| $\text{Fiducial}$ | $\text{Polar}$ |
| $\text{Equatorial}$ | $\text{Equatorial}$ |
where \( \tilde{\mu}, \tilde{\gamma} \) the latitude, with a singularity at the pole. To evade this issue quite misleading because the metric on a sphere depends on the chirp mass. Namely, \( f_0 = 9.9 \text{mHz}, 2.7 \text{mHz}, 16.4 \text{mHz} \) for Heavy, Fiducial, and Light, respectively. Since we keep the SNR fixed in this comparison, this means that the lighter system has a stronger sky-dependent Doppler modulation of the phase, helping with the localization.

When comparing the Polar, Fiducial and Equatorial systems, a direct comparison of the sky localization could be quite misleading because the metric on a sphere depends on the latitude, with a singularity at the pole. To evade this issue we define a system of coordinates on the sphere \((\mu, \gamma)\) such that the injection point is always on the equator. The transformation from the ecliptic coordinates to this frame is source dependent. The spherical coordinates at the equator are locally Cartesian and simplify the comparison of the results. We show the results of the sky localization in Fig. 11 for the Polar, Equatorial and Fiducial systems in the \((\mu, \gamma)\) frame and for \(T_{\text{obs}} = 4\text{yr}\). All three systems recover \(\mu\) (azimuthal angle) similarly well but the determination of \(\gamma\) worsens as \(\beta \to 0\). Furthermore, for the Equatorial system we find a tail extending all the way to a secondary sky position corresponding to \(\beta \to -\beta\). This behaviour is due to the dominant Doppler phase in the frequency response which goes as \(\cos \beta\): although the amplitude of the effect itself is maximized, its variation with \(\beta\) is minimal as \(\cos \beta\) is flat for \(\beta = 0\). For \(T_{\text{obs}} = 10\text{yr}\) this partial degeneracy is broken thanks to a combination of effects: there are more cycles of LISA’s orbit contributing, the signal reaches high frequencies where the \(f\)-dependent in the response \((3.5)\) are larger, and the total SNR itself is larger. The solid angle for the Equatorial system is larger as compared to other systems (as reported in Table V) but remains well below the current sky localization with ground-based observatories [2].

2. Other extrinsic parameters

We find strong correlations between inclination and distance, and between the polarisation and the initial phase. These degeneracies are commonly seen in the analysis of LIGO/Virgo sources when using only the dominant 2, \(\pm 2\) mode in the analysis. With only the dominant 2, \(\pm 2\) mode, the gravitational wave in the radiation frame is given as:

\[
\tilde{h}_\ell(f) = \tilde{A}(f) \left( \frac{1 + \cos^2(\ell)}{2} \right) e^{i\phi} e^{-2i\Psi(f)},
\]

\[
\tilde{h}_x(f) = i \tilde{A}(f) \cos(2\psi) e^{-2i\Psi(f)},
\]

(4.12a)

(4.12b)

where \(\tilde{h}_\ell(f) = A(f) \exp(-i\Psi(f))\) is the frequency domain amplitude and phase decomposition of the mode \(h_\ell\), with \(\tilde{A} = \sqrt{S/I_{\text{obs}}}A(f)\) absorbing conventional factors. We refer to [51] for notations; in particular we exploit the symmetry between \(h_{22}\) and \(h_{-2}\) for non-precessing systems to write the waveform in terms of \(h_{22}\) only. Going to the SSB frame we rotate by the polarisations angle \(\psi\):

\[
\tilde{h}^{\text{SSB}}_\ell = \tilde{h}_\ell \cos(2\psi) - \tilde{h}_x \sin(2\psi),
\]

\[
\tilde{h}^{\text{SSB}}_x = \tilde{h}_x \sin(2\psi) + \tilde{h}_\ell \cos(2\psi).
\]

(4.13a)

(4.13b)

For a face-on system, \(\ell = 0\) leading to:

\[
\tilde{h}^{\text{SSB}}_\ell(f) = \tilde{A}(f) e^{2i\psi} e^{-2i\Psi(f)},
\]

\[
\tilde{h}^{\text{SSB}}_x(f) = i \tilde{A}(f) e^{2i\psi} e^{-2i\Psi(f)},
\]

(4.14a)

(4.14b)

Thus we see that \(\psi\) and \(\varphi\) appear only through the combination \(\psi - \varphi\) yielding a true degeneracy corresponding to \(\psi - \varphi = \text{const}\). For systems close to face-on face-off, like the Fiducial system, this gives the strong correlation between \(\psi\) and \(\varphi\) observed in Fig. 2. For edge-on systems \((\ell = \pi/2)\) we have instead:

\[
\tilde{h}^{\text{SSB}}_\ell(f) = \tilde{A}(f) \cos(2\varphi) e^{-2i\Psi(f)},
\]

\[
\tilde{h}^{\text{SSB}}_x(f) = \tilde{A}(f) \sin(2\varphi) e^{-2i\Psi(f)},
\]

(4.15a)

(4.15b)

and the degeneracy between \(\psi\) and \(\varphi\) is then broken as also shown in Fig. 12. There we compare the distributions of \(\psi\), \(\varphi\), \(\ell\) and \(D_L\) for the Edge-on system to the Far system (which is almost face-on). Those systems have similar SNR. When
the degeneracy between $\psi$ and $\phi$ is broken, we observe a correlation between $\varphi$ and $f_0$. This is an artificial correlation due to relating $\varphi$ to the value of the phase at $f_0$ for each template. Using a fixed reference frequency, such as the the initial frequency of the injected signal for example, eliminates this correlation.

In Fig. 12 we also plot distance and inclination, which show a significant correlation for the Far system, that is absent for the Edge-on system. Distance and inclination are purely extrinsic parameters, and the degeneracy features when subdominant (higher order) modes are negligible appear in the same way for LIGO/Virgo and LISA. For LISA, see e.g. a discussion in the context of galactic binaries in [83]. In short, in the limit of face-on/off systems the inclination acts as a scaling factor over a rather broad range of inclination values, thus changes in $\cos \iota$ can be compensated by changes in $D_L$. For close to edge-on systems, the $x$-polarisation of the wave is suppressed (in the wave-frame, before transforming to the SSB frame as in (4.13)). The important point is that this suppression of $h_x$ depends itself quite rapidly on the inclination, so that reproducing the injected signal leads to a rather tight constraint on $\iota$, and as a consequence on $D_L$. For MBHBs observations with LISA, higher modes play an important role and help breaking these degeneracies [51]; but SBHBs are observed by LISA far from coalescence and higher modes are negligible for these signals.

In Fig. 13 we show the effect of the distance prior on the posterior distribution for $\cos \iota$ and $D_L$ using the Flatphys and Flatsampl priors for $T_{\text{obs}} = 4\text{yr}$. The former favors larger distances and, to keep the correct overall signal amplitude, compensates by preferring the face-on configuration. In the case of the Flatsampl prior, the posterior distribution of $\cos \iota$ is flat because the likelihood itself is very flat around $\iota = 0, \pi$ ($\cos \iota$ is a slowly varying function around its extrema). Thus, the choice of prior shifts the peak of the posterior, but the 90% CI still contains the true value and its width is largely unaffected.

Among all the cases we have considered, $D_L$ can be at best determined within 40%, with the exception of the Edge-on system for which we can determine distance to within 20%. However, the edge-on systems will have lower SNR for a fixed distance to the source, and, therefore there is an observational selection effect where the the face-on/off systems are preferred (that is what we observe with LIGO/Virgo). If we fix all other parameters of the Fiducial system and set $\iota = \pi/2 - \pi/36$, the SNR drops from 21 to 9 for $T_{\text{obs}} = 10\text{yr}$. For the fixed inclination, time to coalescence and source position, the error on intrinsic parameters, distance and sky position scale, in first approximation, as $1/\text{SNR}$.

### C. Fisher matrix analysis

In this subsection we consider PE using a slightly improved version of Fisher information matrix analysis, inspired by [46]. We have introduced the Fisher matrix in subsection II B and discussed its augmented version, the effective Fisher, in subsection II C for computing the covariance matrix. As we mentioned in subsection II C and showed in subsection IV B 2, the likelihood is very flat around $\iota = 0, \pi$ leading Fisher-based-PE to overestimate the errors on $\cos \iota$ and $D_L$. To correct for this, we add an additional term ($F_\iota$) to the effective Fisher matrix: $F_{\text{eff}} = F + F_\iota + F_t$ where $F$ is the “original” Fisher matrix given by eq (2.6) and $F_\iota$ is introduced to account for the prior on spins. Empirically, we found the choice $F_{\cos \iota, \cos \psi} = \frac{1}{0.2(20/\text{SNR})^2}$ and 0 elsewhere to give good results for $\cos \iota$ and $D_L$. The prior matrix $F_\iota$ does more than truncating the error on spins: it mimicks the non-trivial prior on $\chi_+$ and $\chi_-$. Indeed, requiring the spins to be in the physically allowed range ($-1 \leq \chi_{1,2} \leq 1$) leads to a parabola-shaped prior on $\chi_{+,-}$ as seen on 1. We approach this non-trivial prior by a Gaussian distribution centered at $\chi_{+,,-} = 0$ with standard deviation $\sigma = 0.5$. We invert the effective Fisher matrix to obtain the covariance matrix and use it to draw points from a multivariate Gaussian distribution. To fully account for the effect of the prior on spins, the point at which the Gaussian distribution is centred is shifted to: $\theta_{\text{eff}} = F_\iota^{-1} F_\theta$. We only keep points within the boundaries given in Eq. (3.2). For $\psi$ and $\phi$ we draw points in an interval of width $\pi$ around the central value. In Figs. 14 and 15 we compare our Fisher analysis to the inferred distribution for the Fiducial system using the Flatsampl prior.

We find a very good agreement despite the rather low SNR of this system, especially in the $T_{\text{obs}} = 4\text{yr}$ case. In particular, the sky localization is the same for the full PE and the effective Fisher analysis. Naturally, this method cannot reproduce the secondary maximum we found for the Equatorial system.
FIG. 14. Comparison between the inferred distribution for the *Fiducial* system using the *Flatsamp1* prior and our Fisher analysis with $T_{\text{obs}} = 4\text{yr}$. Black lines and squares indicate the true values.

but it does predict a higher error as the system approaches the equatorial plane. The good agreement for $\chi_+$ and $\chi_-$ and for $T_{\text{obs}} = 4\text{yr}$ is because the effective Fisher and posterior distribution are both prior dominated. For $M_c$ and $\eta$, Fisher agrees with the full PE on a 2-sigma level but cannot reproduce the banana-like correlation. In case of $T_{\text{obs}} = 10\text{yr}$, the likelihood becomes more informative for $\chi_+$, reducing the error predicted by the “original” Fisher while the $\chi_-$ distribution is still prior dominated. Without adding $F_t$ to the effective Fisher, the direction for the correlation between $\cos\iota$ and $D_L$ is predicted well but the Fisher matrix severely overestimates the error for nearly face-on/face-off systems. For the *Edgeon* system, the

FIG. 15. Similar to Fig. 14 but with $T_{\text{obs}} = 10\text{yr}$. 
FIG. 16. Evolution of the error as function of the time before merger we start observing the system in the $T_{\text{obs}} = 4\text{yr}$ and $T_{\text{obs}} = 10\text{yr}$ cases. The SNR is given in the upper panel. The errors on $M_c$, $\eta$ and $\chi_+$ correspond to the width of the 90% CIs, and $\Delta \Omega$ is defined in (4.11).

Based on the rather good agreement we found with Bayesian PE, we can exploit the simplicity of Fisher analysis to further explore how does the PE evolves with the time (left) to coalescence. In Fig. 16, assuming $T_{\text{obs}} = 4\text{yr}$ and $T_{\text{obs}} = 10\text{yr}$, we plot the errors on $M_c$, $\eta$, $\chi_+$ and $\Delta \Omega$ as a function of the time to coalescence $t_c$, keeping all the parameters of the Fiducial system fixed but varying the initial frequency in accordance with the chosen $t_c$. We plot the corresponding evolution of the SNR in the top panel, with the lowest SNR of 8 being reached for $t_c \approx 1\text{yr}$. Dashed lines mark $t_c = T_{\text{obs}}$ in each case which corresponds to the maximum achievable SNR given the observation time and it also corresponds to the best estimation of parameters. Note two different regimes on the two sides of the dashed line: to the left, the PE is governed by the decrease in the signal duration in LISA band and reduction in SNR, while to the right the PE is determined mainly by the bandwidth of the signal spanned over the observation time. As discussed in Sec. IV B 1 the sky localization comes mainly from modulations caused by the motion of LISA, therefore it worsens rapidly if the system spends too little time in band (below 1 year).

D. Long-wavelength approximation

In Table VI we compare the SNR for the Fiducial, Polar and Equatorial systems using the Full and LW responses for two observation times. We find that accounting for the degradation at high frequencies (Eq. (3.8)), the LW approximation seems to barely affects the PE as can be seen in Fig. 17. We find similar behaviour for the Polar and Equatorial systems. Some care is needed in interpreting this result: this comparison shows that the high frequency terms neglected in the LW approximation have little impact on the posterior of the sky position if the likelihood is computed self-consistently (signal and template are produced using same response, either LW or full). However, when analysing real data these high frequency terms cannot be neglected. In other words, these effects in the full response can indeed be subdominant in the parameter recovery, if more information comes from other effects like the LISA motion and the main Doppler modulation, while not being negligible in the signal itself. To illustrate this, we simulate data for the Fiducial, Polar and Equatorial systems using the full response and perform a Bayesian analysis using the LW approximation to compute templates. In Table VII, we

| System     | $T_{\text{obs}} = 4\text{yr}$ | $T_{\text{obs}} = 10\text{yr}$ |
|------------|-------------------------------|-------------------------------|
| Fiducial   | 13.5                          | 12.9                          |
| Polar      | 12.8                          | 12.2                          |
| Equatorial | 14.9                          | 14.2                          |

TABLE VI. Comparison between the SNRs for the Fiducial, Polar and Equatorial systems using the Full and the LW response.
FIG. 17. Comparison of inferred distributions of intrinsic parameters and sky location using the Full and Lowf responses for the Fiducial system in the $T_{\text{obs}} = 10\text{yr}$ case. Black lines and squares indicate the true values.

|                   | $T_{\text{obs}} = 4\text{yr}$ | $T_{\text{obs}} = 10\text{yr}$ |
|-------------------|-------------------------------|-------------------------------|
|                   | log $\mathcal{L}(\theta_0)$  | max (log $\mathcal{L}$)      | $\tilde{\rho}$ | log $\mathcal{L}(\theta_0)$  | max (log $\mathcal{L}$) | $\tilde{\rho}$ |
| Fiducial          | -50                           | -2                            | 0.99            | -268                          | -38                       | 0.91            |
| Polar             | -45                           | -3                            | 0.99            | -234                          | -30                       | 0.92            |
| Equatorial        | -55                           | -2                            | 0.99            | -288                          | -34                       | 0.94            |

TABLE VII. Loglikelihood at the true point, maximum likelihood and relative SNR (defined in Eq. (4.16)) when using the LW approximation in the Bayesian analysis for data generated with the Full response.

give the log-likelihood evaluated at the true point, the maximum likelihood and the maximum overlap:

$$\tilde{\rho} = \max_h \left( \frac{(d|h)}{\sqrt{(d|d)(h|h)}} \right).$$

(4.16)

In practice, we compute max by optimizing over our samples. The quantity $1 - \tilde{\rho}$ indicates how much SNR would be lost if wrong templates were used for the detection of signal. We find that up to $\sim 10\%$ of the SNR could be lost, given the already low SNR of SBHBs in LISA this would severely compromise our chances of detecting such sources. The very small value of the likelihood at the true point by itself shows that using the LW approximation will have an impact on the PE. In fig. 18, we compare posterior distributions obtained by using template generated with Full or LW response while analysing the Fiducial system, with $T_{\text{obs}} = 10\text{yr}$ and generated with the Full response. This system has a significant bandwidth and the LW template cannot fit simultaneously the low and high frequency content of the signal, causing severe biases in PE and loss of SNR. The same system but with $T_{\text{obs}} = 4\text{yr}$, shows different result, the LW template is effectual enough to fit the signal rather well with the largest bias appearing only in $\psi - \varphi$ distribution as a compensation for terms neglected in the response and with a mild drop in the SNR. However, those signals are quite weak and we do not have the luxury to loose even a small portion of SNR.

Thus, our findings seem to validate the LW approximation for prospective parameter estimation studies, if it is used consistently for injecting and recovering the signal, while it would be inappropriate to analyze real data. However, we should remember that we did not explore the full parameter space, while (3.8) is valid as an average over orientations, so a different choice of parameters could yield worse results. We also note that the full response (3.5) is actually quite simple and not more expensive computationally, while being unambiguous and eliminating the need for the averaging entering (3.8).
V. DISCUSSION

Merging binary stellar mass black holes are detected almost weekly during the third LIGO/Virgo observational run (O3). In this work we explored what LISA will be able to tell us about those binaries. While ground-based detectors observe the last seconds before the merger, LISA will see the early inspiral evolution of those systems. The results of O3 run are not publicly available, so we used a GW150914-like system as a fiducial system in our study. We varied parameters of the system in turn, investigating the corresponding changes in the PE. We worked on the simulated (noiseless) data and we used the full LISA response. We employed a Bayesian analysis for the parameter estimation and cross-checked our results using two independent samplers. We have found that the PE results are most sensitive to the frequency span of the GW and its position within LISA sensitivity given the observation duration, or in other words, how much the signal chirps during the observation time.

For weakly chirping systems that do not reach high frequencies during LISA’s observations, the GW phase dominated by the leading PN order, with smaller contributions from higher PN terms. As a result, the best measured parameter is the chirp mass (entering at the leading order) with typical relative error below $10^{-4}$. The weak contributions of sub-leading terms up to 1.5PN lead to a three-way correlation between spins, sym-

FIG. 18. Comparison of the inferred distributions for the Fiducial system in the $T_{\text{obs}} = 10\text{yr}$ case using the Full and the LW approximation in the Bayesian analysis. In both cases, data was generated with the Full response. Black lines and squares indicate the true values.
metric mass ratio and the chirp mass. The mass ratio is very poorly constrained and the posterior for the spins is dominated by the priors. We still can recover very well the sky position of the source (typically $< 0.4 \, \text{deg}^2$) which comes from the modulation of the amplitude and the phase of the GW signal due to LISA's motion. Such an area in the sky is within the field of view of electromagnetic instruments, such as Athena and/or SKA.

For chirping systems, that coalesce during the observations and therefore reach the high end of the LISA frequency band, higher order PN terms become more important and help breaking the correlations in intrinsic parameters, leading to a significant improvement in parameter estimation. The individual masses for chirping systems are measured within $20 - 30\%$ and even better for systems with higher mass ratio. The constraints on individual spins result from the combination of the measurement of $\chi_{\text{PN}}$ and the physical boundaries of the prior on spins: $-1 \leq \chi_{1,2} \leq 1$. This suggests suggest using $\chi_{\text{PN}}$ as a sampling parameter. We find that the measurement of the time to coalescence improves as we observe the systems closer to merger, from $O(1 \, \text{day})$ (for mildly chirping binaries) to $O(30 \, \text{s})$. We note that the only way to increase our chances to observe SBHBs chirping is by increasing LISA’s mission duration.

The measurement of the luminosity distance is less impacted by whether the systems are chirping or not, and it is essentially a function of their SNR. Much like LIGO/Virgo observations when higher modes can be neglected, the degeneracy between distance and inclination is important. In our example, the distance is typically measured within $40 - 60\%$ if the system is close to face-on and within $20\%$ if the system is edge-on (when adjusting the distance to keep the same SNR). The distance and therefore redshift uncertainty dominates the measurement of the source-frame chirp mass at a percent level. The precision on individual masses in the source frame is dominated by intrinsic degeneracies.

We have suggested an augmentation of the usual Fisher matrix approach, that we called the effective Fisher matrix, and we have shown that it gives rather reliable results for the sky position and intrinsic parameters of the system when compared to Bayesian PE. We also showed that combining the use of the long-wavelength approximation for LISA with the introduction of a degradation factor at high frequencies yields very similar results as compared to using the full response for computing likelihoods self-consistently (using the same response for injected data and templates). However, using the long-wavelength approximation to analyse real data could decrease the SNR by $10\%$, drastically reducing our chances of detecting the signals and has a significant impact on the PE, in particular on the measurement of intrinsic parameters. The computational cost of the full response being essentially the same than the long-wavelength approximation, we recommend its use in future works.

We can utilise the knowledge and understanding obtained in PE study for development of the search: (i) the PE for those systems is mainly monomodal, the secondary modes appear either in special cases (like Equatorial system considered in the main body of the paper) or under the effect of prior for parameters on which the likelihood is weakly informative (like $\eta$ for non-chirping systems) (ii) the chirp mass and the sky coordinates are the best measured parameters, so we can make hierarchical search starting with those parameters taking into account the correlations which are explored and understood (see section IV) (iii) the effective Fisher is a good proposal for Bayes-based search once we started to see sign of the GW signal in the data. In addition we can perform data volume-increasing analysis starting with half year long data segment progressively increasing it. This works as a natural annealing scheme and should help in recovering (especially) chirping systems.

Detection of even few SBHBs by LISA which merge somewhat later with a very high SNR in the band of ground based detectors [37] will constitute “golden events”. Beyond all the benefits of multiband detections per se, the information provided by LISA itself will be very valuable. For example, [84] suggested that the measurement of the time to coalescence could be used to inform ground based detectors and improve black hole spectroscopy. The good estimate on the time to coalescence and the sky location could be used for electromagnetic follow-up of the source as suggested in [27]. Finally, these measurements could be used to tighten the constraints on the Hubble constant ($H_0$) even if no electromagnetic counterparts are detected, using galaxy catalogues [85–87].

Modifications to the GW phase induced by either modified theories of gravity or environmental effects will generically involve additional coefficients parametrising the underlying mechanisms and their correlation with the parameters of the system ($M_c, \eta, \ldots$). As a consequence, even for low frequency modifications the best constraints/measurements will come from chirping systems as we found in [21] in the context of testing modified theories of gravity with LISA observations.

A major improvement to this work would be the inclusion of eccentricity. Astrophysical formation models predict that binaries formed dynamically should have large eccentricities [35, 36]. However, by the time these binaries reach the frequency band of ground based detectors they will have circularised. Thus LISA could play an important role in the discrimination between different formation channels [40–43]. Furthermore, neglecting eccentricity could affect PE and detection efficiency. We are currently limited by the lack of fast eccentric waveforms but there is a good progress in this direction [88–92]. Concerning spins, binaries formed dynamically are expected to have misaligned spins [37] but precession effects should be weak in the early inspiral. Thus, although interesting, the use of precessing waveforms should not appreciably change conclusions drawn in this work.

We conclude with the claim that this work, together with [21, 27], has confirmed the scientific potential of the observation of SBHBs with LISA and should be seen as a first step towards an extensive study of the PE for the multiband observations.
ACKNOWLEDGEMENTS

A.T. is grateful to Nikolaos Karnesis for his help at the start of the project and his support all along it. This work has been supported by the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 690904. A.T. acknowledges financial support provided by Paris-Diderot University (now part of Université de Paris). The authors would also like to acknowledge networking support from the COST Action CA16104.

[1] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.
[2] B.P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev. X*, 9(3):031040, 2019.
[3] B. P. Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119(14):141101, 2017.
[4] B. P. Abbott et al. GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence. *Astrophys. J.*, 851(2):L35, 2017.
[5] B. P. Abbott et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev. X*, 6(4):041015, 2016. [erratum: Phys. Rev.X8,no.3,039903(2018)].
[6] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspirial. *Phys. Rev. Lett.*, 119(16):161101, 2017.
[7] Benjamin P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017. [Erratum: Phys.Rev.Lett. 121, 129901 (2018)].
[8] R. Abbott et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. 4 2020.
[9] B.P. Abbott et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4M_\odot$. *Astrophys. J. Lett.*, 892:L3, 2020.
[10] R. Abbott et al. GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object. *The Astrophysical Journal*, 896(2):L44, jun 2020.
[11] B. P. Abbott et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *Astrophys. J.*, 882(2):L24, 2019.
[12] Benjamin P. Abbott et al. Upper Limits on the Rates of Binary Neutron Star and Neutron Star–black Hole Mergers From Advanced Ligo’s First Observing run. *Astrophys. J.*, 832(2):L21, 2016.
[13] B. P. Abbott et al. Tests of general relativity with the binary black hole signals from the ligo-virgo catalog gwrtc-1. 2019.
[14] B. P. Abbott et al. Tests of general relativity with GW150914. *Phys. Rev. Lett.*, 116(22):221101, 2016. [Erratum: Phys. Rev. Lett.121.no.12,129902(2018)].
[15] B. P. Abbott et al. Tests of General Relativity with GW170817. *Phys. Rev. Lett.*, 123(1):011102, 2019.
[16] Heather Audley et al. Laser Interferometer Space Antenna. 2017.
[17] Antoine Klein et al. Science with the space-based interferometer eLISA: Supermassive black hole binaries. *Phys. Rev.*, D93(2):024003, 2016.
[18] Valeriya Korol, Elena M. Rossi, Paul J. Groot, Gijs Nelemans, Silvia Toonen, and Anthony G.A. Brown. Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA. *Mon. Not. Roy. Astron. Soc.*, 470(2):1894–1910, 2017.
[19] Alberto Sesana. Prospects for Multiband Gravitational-Wave Astronomy after GW150914. *Phys. Rev. Lett.*, 116(23):231102, 2016.
[20] Pau Amaro-Seoane and Lucia Santamaria. Detection of IMBHs with ground-based gravitational wave observatories: A biography of a binary of black holes, from birth to death. *Astrophys. J.*, 722:1197–1206, 2010.
[21] Alexandre Toubiana, Sylvain Marsat, Enrico Barausse, Stanislav Babak, and John Baker. Tests of general relativity with stellar-mass black hole binaries observed by LISA. 2020.
[22] Giuseppe Gnocchi, Andrea Maselli, Tiziano Abdelsaﬁn, Nicola Giaocchino, and Michela Mapelli. Bounding Alternative Theories of Gravity with Multi-Band GW Observations. 2019.
[23] Zack Carson and Kent Yagi. Multi-band gravitational wave tests of general relativity. 2019.
[24] Salvatore Vitale. Multiband Gravitational-Wave Astronomy: Parameter Estimation and Tests of General Relativity with Space- and Ground-Based Detectors. *Phys. Rev. Lett.*, 117(5):051102, 2016.
[25] Enrico Barausse, Nicolás Yunes, and Katie Chamberlain. Theory-Agnostic Constraints on Black-Hole Dipole Radiation with Multiband Gravitational-Wave Astrophysics. *Phys. Rev. Lett.*, 116(24):241104, 2016.
[26] Katie Chamberlain and Nicolas Yunes. Theoretical Physics Implications of Gravitational Wave Observation with Future Detectors. *Phys. Rev.*, D96(8):084039, 2017.
[27] Andrea Caputo, Laura Sherma, Alexandre Toubiana, Stanislav Babak, Enrico Barausse, Sylvain Marsat, and Paolo Pani. Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart. 2020.
[28] Enrico Barausse, Vitor Cardoso, and Paolo Pani. Can environmental effects spoil precision gravitational-wave astrophysics? *Phys. Rev.*, D89(10):104059, 2014.
[29] Enrico Barausse, Vitor Cardoso, and Paolo Pani. Environmental Effects for Gravitational-wave Astrophysics. *J. Phys. Conf. Ser.*, 610(1):012044, 2015.
[30] Nicola Tamanini, Antoine Klein, Camille Bonvin, Enrico Barausse, and Chiara Caprini. The peculiar acceleration of stellar-origin black hole binaries: measurement and biases with LISA. 2019.
[31] Vitor Cardoso and Andrea Maselli. Constraints on the astrophysical environment of binaries with gravitational-wave observations. 2019.
[32] Konstantin A. Postnov and Lev R. Yungelson. The Evolution of Compact Binary Star Systems. *Living Rev. Rel.*, 17.3, 2014.
[33] Matthew J. Benacquista and Jonathan M. B. Downing. Relativistic Binaries in Globular Clusters. *Living Rev. Rel.*, 16:4, 2013.
[34] Simeon Bird, Ilias Cholis, Julian B. M. Muñoz, Yacine Ali-Haïmoud, Marc Kamionkowski, Ely D. Kovetz, Alvise Racanelli, and Adam G. Riess. Did LIGO detect dark matter?
[74] Travis Robson, Neil J. Cornish, and Chang Liu. The construction and use of LISA sensitivity curves. *Class. Quant. Grav.*, 36(10):105011, 2019.

[75] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[76] Daniel Foreman-Mackey. corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 24, 2016.

[77] Alberto Mangiagli, Antoine Klein, Alberto Sesana, Enrico Barausse, and Monica Colpi. Post-Newtonian phase accuracy requirements for stellar black hole binaries with LISA. *Phys. Rev.*, D99(6):064056, 2019.

[78] Eric Poisson and Clifford M. Will. Gravitational waves from inspiraling compact binaries: Parameter estimation using second postNewtonian wave forms. *Phys. Rev.*, D52:848–855, 1995.

[79] N. Meidinger. The Wide Field Imager instrument for Athena. *Contributions of the Astronomical Observatory Skalnate Pleso*, 48(3):498–505, Jul 2018.

[80] Lee Samuel Finn and David F. Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. *Phys. Rev. D*, 47:2198–2219, 1993.

[81] Curt Cutler and Eanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral wave form? *Phys. Rev.*, D49:2658–2697, 1994.

[82] Sweta Shah, Marc van der Sluys, and Gijs Nelemans. Using electromagnetic observations to aid gravitational-wave parameter estimation of compact binaries observed with LISA. *Astron. Astrophys.*, 544:A153, 2012.

[83] Rhondale Tso, Davide Gerosa, and Yanbei Chen. Optimizing LIGO with LISA forewarnings to improve black-hole spectroscopy. *Phys. Rev.*, D99(12):124043, 2019.

[84] Bernard F. Schutz. Determining the Hubble Constant from Gravitational Wave Observations. *Nature*, 323:310–311, 1986.

[85] Walter Del Pozzo, Alberto Sesana, and Antoine Klein. Stellar binary black holes in the LISA band: a new class of standard sirens. *Mon. Not. Roy. Astron. Soc.*, 475(3):3485–3492, 2018.

[86] Koutarou Kyutoku and Naoki Seto. Gravitational-wave cosmography with LISA and the Hubble tension. *Phys. Rev.*, D95(8):083525, 2017.

[87] Zhoujian Cao and Wen-Biao Han. Waveform model for an eccentric binary black hole based on the effective-one-body-numerical-relativity formalism. *Phys. Rev. D*, 96(4):044028, 2017.

[88] Brennan Ireland, Ofek Birnholtz, Hiroyuki Nakano, Eric West, and Manuela Campanelli. Eccentric Binary Black Holes with Spin via the Direct Integration of the Post-Newtonian Equations of Motion. *Phys. Rev. D*, 100(2):024015, 2019.

[89] Tanja Hinderer and Stanislav Babak. Foundations of an effective-one-body model for coalescing binaries on eccentric orbits. *Phys. Rev. D*, 96(10):104048, 2017.

[90] E.A. Huerta et al. Eccentric, nonspinning, inspiral, Gaussian-process merger approximant for the detection and characterization of eccentric binary black hole mergers. *Phys. Rev. D*, 97(2):024031, 2018.