Implications of the pulsar timing array detections for massive black hole mergers in the LISA band

Enrico Barausse[1,2] Kallol Dey[3] Marco Crisostomi[1,2] Akshay Panayada[3]
Sylvain Marsat,[4] and Soumen Basak[3]

[1]*SISSA, Via Bonomea 265, 34136 Trieste, Italy and INFN Sezione di Trieste*
[2]*IFPU—Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy*
[3]*School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram,*
Maruthamala PO, Vithura, Thiruvananthapuram 695551, Kerala, India
[4]*Laboratoire des 2 Infinis—Toulouse (L2IT-IN2P3), Université de Toulouse,*
CNRS, UPS, F-31062 Toulouse Cedex 9, France

The recent evidence of a stochastic background of gravitational waves in the nHz band by pulsar timing array (PTA) experiments has shed new light on the formation and evolution of massive black hole binaries with masses $\sim 10^8$–$10^9 M_\odot$. The PTA data are consistent with a population of such binaries merging efficiently after the coalescence of their galactic hosts, and presenting masses slightly larger than previously expected. This momentous discovery calls for investigating the prospects of detecting the smaller ($\sim 10^5$–$10^7 M_\odot$) massive black hole binaries targeted by the Laser Interferometer Space Antenna (LISA). By using semianalytic models for the formation and evolution of massive black hole binaries calibrated against the PTA results, we find that LISA will observe at least a dozen and up to thousands of black hole binaries during its mission duration. The minimum number of detections rises to ~ 70 if one excludes models that only marginally reproduce the quasar luminosity function at $z = 6$. We also assess LISA's parameter estimation capabilities with state-of-the-art waveforms including higher modes and realistic instrumental response, and find that the masses, sky position, and distance will typically be estimated to within, respectively, 1%, 10 sq deg, and 10% for the detected systems (assuming a four-year mission).

I. INTRODUCTION

The possible detection of a stochastic background of gravitational waves (GWs) by the European Pulsar Timing Array (EPTA) [1], the Indian Pulsar Timing Array (InPTA) [2], the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [3], the Parkes Pulsar Timing Array (PPTA) [4], and the Chinese Pulsar Timing Array (CPTA) [5] in June 2023 opened a way for a deeper understanding of GW sources in the nHz frequency band. Various exotic explanations have been proposed to account for the origin of the observed stochastic background. These explanations encompass concepts like ultralight dark matter [6], cosmic strings [7], or cosmological background [8–10]. However, the most plausible possibility, by far, is that the pulsar timing array (PTA) signal is produced by an astrophysical population of merging massive black hole (MBH) binaries [8,11].

MBHs, with masses ranging from $\sim 10^5 M_\odot$ (or even lower) to $\sim 10^9 M_\odot$ are ubiquitous in massive galaxies [12,13] and in a fraction of low-mass dwarf galaxies [14–16] in the local Universe. The evolution of these black holes is intricately connected with that of their galactic hosts, from which they accrete matter, and it is believed that they exert feedback, either through radiation or jets, when they are active galactic nuclei (AGNs) and shine brightly [17]. This synergistic evolution is reflected in a wealth of scaling relations [18–20] between the MBH properties and those of their galactic hosts, and is fundamental for galaxy formation models. Indeed, AGN feedback is probably responsible, at least in part, for the occurrence of these scaling relations (see, however, Refs. [21–24] for some complications in the interpretation of these relations). Moreover, it is also crucial to reconcile the bottom-up hierarchical formation predicted by the ΛCDM model for dark matter halos with the "antihierarchical" evolution (or "downsizing") of galaxies [25,26], i.e., the fact that massive galaxies mostly contain old stellar populations and show little star formation, while low-mass galaxies are typically dominated by young stellar populations and enjoy more vigorous star formation.

Despite their importance for galaxy formation, MBHs are still poorly understood, especially at the low end of their mass spectrum and at high redshift, where electromagnetic observations are difficult. GW observatories are crucial in this respect, as gravitational signals decay slowly with redshift (scaling with the inverse of the luminosity

distance D_L) and interact weakly with matter. In this context, it is clear that the PTA detection of a stochastic GW background is potentially ripe for consequences for our understanding of MBHs. Indeed, as stressed in Refs. [8,9,11], the PTA signal is in broad agreement with expectations from our current understanding of the formation and evolution of MBH binaries, provided that the latter merge (and possibly accrete) rather efficiently. However, as the MBHs responsible for the PTA background have masses $\gtrsim 10^8 M_\odot$, different experimental facilities are needed to explore the low end of their mass function, down to $10^5 M_\odot$ or smaller. Among the latter, the Laser Interferometer Space Antenna (LISA) [27], a joint Euro-American space mission, will play a major role.

LISA will target MBH mergers with masses between 10^4 and $10^7 M_\odot$, up to very high redshifts ($z \sim 15$ or even larger), and will therefore greatly enhance our understanding of the assembly of MBHs in their galactic hosts [27]. Numerous attempts have been made at predicting the number of MBH mergers that will be observed by LISA, based on semi-analytic galaxy formation models and hydrodynamic simulations [28–37], but these estimates are plagued by large uncertainties due to the lack of resolution of the models/ simulations and most of all to our incomplete understanding of the "subgrid" physics that regulates the evolution of MBHs and galaxies on small scales. The PTA detection, therefore, offers a unique opportunity to test the models used thus far to predict LISA event rates and decrease the error bars of the latter [38]. This will be the subject of this paper.

In more detail, the paper is structured as follows: In Sec. II, we delve into the physics of semianalytic models for the formation and evolution of galaxies and MBHs. In Sec. III, we compare the predictions of these models with the data obtained from the PTA observations. Moving on to Sec. IV, we provide a comprehensive description of GW emission and detection with LISA. Section V is dedicated to presenting our findings concerning LISA's detection rates and parameter estimation. Finally, in Sec. VI, we summarize our study and draw our conclusions based on the results obtained.

II. THE SEMIANALYTIC MODEL

We adopt the semianalytic model of [39] for the formation and evolution of galaxies and MBHs. Additions and improvements to specific aspects of the model were subsequently introduced in [35,37,40–42]. Here, we review concisely the model's framework, referring the interested reader to the aforementioned works for more details.

A. The dark matter merger trees

The model is constructed upon a merger tree of dark matter, created using an extended Press-Schechter algorithm [43,44], modified to reproduce N-body simulation results [44]. The resolution of the merger tree (i.e., the minimum halo mass below which dark matter growth via mergers is not followed in detail but collectively modeled as accretion) is redshift dependent. In more detail, it is set to a fixed fraction of the mass of the most massive halo at the previous (later) redshift step of the tree.[1] The redshift step is chosen adaptively to be sufficiently small to ensure that multiple halo fragmentation is unlikely [28]. A procedure to account for the finite merger tree resolution and extrapolate MBH merger rates to infinite resolution was put forward in Fig. 4 of [33], and will be adopted below whenever explicitly mentioned.

B. The extrapolation to infinite resolution

More precisely, the extrapolation procedure for the (intrinsic) MBH merger rates works as follows. In our simulations for low-mass halos, we notice a linear correlation between a halo's mass at $z = 0$ (M_0) and the number of MBH mergers taking place in the past history of that halo. This trend is easily understood as follows. If one assumes that seed black holes form in halos of mass M_s, the halo at $z = 0$ will have formed from $N \approx M_0/M_s$ seed halos. By defining $N \approx 2^n$ and considering a perfect hierarchy where mergers proceed in subsequent rounds, the total number of mergers by $z = 0$ is $\sum_{i=0}^{n-1} 2^i = 2^n - 1 \approx M_0/M_s$, which explains the linear trend at low masses. At high masses $\gtrsim 10^{13} M_\odot$, this trend is lost, presumably due to lack of resolution. We therefore correct for it by appropriately reweighing the merger rates in those halos.

C. The baryonic structures

The evolution of baryonic structures along the merger tree is followed using semianalytic prescriptions. Among these baryonic structures is a chemically unprocessed intergalactic medium, which accretes onto dark matter halos and undergoes shock heating to the virial temperature in low-redshift, high-mass systems, or which flows into halos along cold filaments on a timescale comparable to the dynamical time in higher-redshift and/or lower-mass systems [45–48].

The intergalactic medium, cooling or streaming along cold filaments into halos, gives rise to a cold gas medium that eventually forms stars (and which is therefore known as "interstellar medium"). The model follows the evolution of the interstellar medium and the stellar population in disks and bulges (i.e., spheroids) accounting for the disruption of disks as a result of major galaxy mergers and bar instabilities, as well as for supernova (SN) feedback and the interstellar medium's chemical evolution.

In galactic nuclei, the model forms nuclear star clusters via *in situ* star formation and/or migration of globular

---

[1]This fixed fraction was chosen to be 10^{-2} in Ref. [39] and 3×10^{-3} in the later works using our model, e.g., Refs. [33,35,37,40–42].

clusters [41,49]. MBHs are also included in the model, forming from seeds at high redshifts and then growing by accretion of nuclear gas and mergers with other black holes when galaxies coalesce. MBH accretion and nuclear star cluster *in situ* formation is assumed to occur from a nuclear gas reservoir, whose growth is modeled as linearly correlated with star formation in the bulge [39,50–52]. The gas in this nuclear reservoir is then assumed to accrete onto the central black hole on a viscous timescale evaluated (at the black hole influence radius) [40], but the accretion rate is capped at a maximum rate comparable to the Eddington rate (or slightly larger for the light-seed scenario; see below). Starting from Ref. [37], the model started accounting also for a possible effect of SN feedback on the formation of the nuclear gas reservoir. Indeed, the simulations of [53] show that in SN feedback models where gas cooling is delayed (as a result of shocks), accretion is quenched in low-mass systems. The model of Ref. [37] therefore suppresses the reservoir growth in bulges with escape velocities $\lesssim 270$ km/s (i.e., the typical speed of SN winds). The model also accounts for AGN feedback from MBHs onto the bulge interstellar medium and the diffuse chemically pristine gas, both from radio jets [39] and radiation [37].

D. The black hole seeds

A crucial aspect of the model for predicting the event rate for LISA is the initial mass function of the black hole seeds at high redshift. Several physical models for the latter have been proposed; see, e.g., [54] for a review. Here, we adopt two representative scenarios: a light-seed (LS) mass function, assuming that seeds form from the remnants of population III stars in high-redshift, low-metallicity systems [55], and a heavy-seed (HS) mass function, assuming that the seeds arise from the collapse of protogalactic disks induced by bar instabilities [56]. In more detail, in the LS scenario we draw the mass of the initial population III stars from a log-normal distribution centered on $300 M_\odot$ and with a standard deviation of 0.2 dex, but with a pair-instability gap between 140 and $260 M_\odot$ [57]. We then assume that the black hole seed has mass $\sim 2/3$ of the mass of the initial population III stars, to account for mass losses during the stellar collapse. Following Ref. [28], we assume that LSs form only in the most massive halos, corresponding to the 3.5σ peaks of the matter density field, at $z \gtrsim 15$. Since it is well known that LS models struggle to reproduce the high-redshift luminosity function [58], we allow for mildly super-Eddington accretion (up to twice the Eddington rate) in the LS scenario. In the HS scenario, black hole seeds have masses depending on the properties of the host halo, typically of the order of $\sim 10^5 M_\odot$. Several implementations of this scenario have been put forward, with Ref. [39] adopting, e.g., the model of Ref. [59], and later versions of the model (from 2015 onward) adopting instead the model of Ref. [56]. In the latter, in particular,

HSs form from bar instabilities of protogalactic disks at $z \gtrsim 15$ in halos of masses $\lesssim$ a few times $10^7 M_\odot$. The instability is assumed to occur below a critical Toomre parameter Q_c, which is believed to be in the range $2 \lesssim Q_c \lesssim 3$. Different choices of Q_c affect the seed occupation fraction at high redshifts, with lower (larger) Q_c producing fewer (more) seeds. For instance, Ref. [42] used $Q_c = 2.5$, while Refs. [33,37,41] used $Q_c = 3$. It should be noticed that because the HSs of Ref. [56] form in relatively small halos ($\lesssim$ a few times $10^7 M_\odot$) at high redshift, merger tree branches that fall below the tree's resolution at low redshifts (and which are therefore not followed to high z) would be artificially devoid of MBH seeds. To alleviate this problem, in Ref. [37] our model started following these subresolution branches and their dark matter merger history on the fly up to $z \gtrsim 15$, without evolving baryons (for computational efficiency).[2] If any one of these $z \gtrsim 15$ progenitor halos contains a seed, we place a seed black hole in the subresolution branch. Reference [37] uses this mitigation strategy for both LSs and HSs, although it is more important for the latter, since our population III seed forms in the largest halos at high z. We stress that this procedure, like the other details for the seeding mechanism, has little effect on low z observables, including the PTA stochastic background. This is because, by the time the black holes have evolved to low redshift or have grown enough to emit in the PTA band, they have lost memory of the initial seeding conditions. However, the seed model, and also our prescription for subresolution seeding, has a big impact on LISA merger rates, which are sensitive also to high redshift. In particular, we will see below that the HS models of Ref. [37] predict more LISA events than similar earlier models (e.g., those of Ref. [33]), and this difference can be ascribed to the updated prescription for subresolution seeding.

E. The delays

Also crucial for the prediction of LISA merger rates is the modeling of the "delays" between halo/galaxy mergers and MBH mergers. Unlike the seeding mechanism at high redshift, these delays also have an important effect on predicting the signal for PTAs [42]. In more detail, the delays are implemented in our model as follows. When two halos merge according to our dark matter tree, we assume that the smaller one survives inside the newly formed system as a satellite subhalo, until it has sunk to the center as a result of dynamical friction [60]. In this phase, which typically lasts a few Gyr, the subhalo and its baryon content also undergo tidal stripping and evaporation as a result of the tidal field of the primary, which in turn changes the

---

[2]While the correction for subresolution branches was already present in Ref. [37], unfortunately it was not reported explicitly there.

evolution of the system (decreasing the efficiency of dynamical friction [61]).

When the subhalo has sunk to the bottom of the host, the baryonic components (i.e., the galaxies) do not merge right away. The satellite galaxy keeps falling toward the center of the host galaxy, again as a result of dynamical friction and tidal stripping/evaporation [37,62,63]. This phase can last for several Gyr, especially for galaxies with unequal stellar masses [63], and is crucial to driving the MBHs from ∼kpc to ∼pc separation. However, as a result of tidal stripping and evaporation, the satellite galaxy may be disrupted and its MBH left as a naked black hole (surrounded at most by a core of stars [62]). This can lead to a potentially numerous population of "stalled" MBHs wandering at separations of hundreds of pc [62,63].

At separations, ∼pc, MBHs eventually form bound binaries, for which dynamical friction becomes inefficient compared to other processes. These include, for instance, three-body interactions between the black hole binary and individual stars ("stellar hardening" [64,65]). In more detail, stars with low angular momentum (i.e., in the "loss cone") can remove energy from the binary via the slingshot effect. As the process repeats for many stars, the black hole binary shrinks progressively, as stars in the loss cone are ejected from the nucleus (e.g., as hypervelocity stars [66]). This causes hardening eventually to become inefficient unless the loss cone gets replenished by stellar diffusion. If this diffusion-induced replenishment is efficient, e.g., as a result of triaxiality of the galaxy potential (resulting, for instance, from a recent galaxy merger) [67–71] or galaxy rotation [72], stellar hardening can drive the binary separation down to $\sim 10^{-2}$–10^{-3} pc on a timescale of few Gyr. Once those small separations have been reached, GW emission alone leads the binary to merger in less than a Hubble time.

Other processes may also help the binary reach sub-pc separations (see, e.g., Ref. [73] for a review). For instance, in gas-rich galactic nuclei, the MBH binary may shrink on timescales of $\sim 10^7$–10^8 yr thanks to interactions with the nuclear gas (planetarylike migration). Furthermore, even in gas-poor environments with inefficient loss-cone replenishment (and thus, inefficient stellar hardening), triple MBH systems will eventually form as a result of later galaxy mergers. In such triple systems, Kozai-Lidov oscillations [74,75] and/or chaotic three-body interactions [76,77] can trigger the merger of at least two of the black holes in a sizeable fraction of systems [35,37,42,77]. Interestingly, these triplet-induced mergers are expected to present a significant eccentricity (≳0.99 when they enter the LISA band, and ∼0.1 at the merger) [35].

III. COMPARISON WITH PTA DATA

Preliminary implications of the PTA detections for the formation and evolution of MBHs have been presented in Refs. [8,9,11]. Here, we follow and extend Sec. 3.3.1 of

Ref. [8] and perform a comparison between the EPTA measurement of the stochastic background and the predictions of our semianalytic model. Like in Ref. [8], we consider in particular the semianalytic model of [39] in its original version (B12) and in its subsequent updates, which were introduced and used in Refs. [33,41] (K + 16), Refs. [35,42] (B + 18), and [37] (B + 20).

The specific models that we consider (which include as a subset those considered in Ref. [8]) can be roughly divided into three classes: (i) "LS-nod (B12)," "HS-nod (B12)," "Q3-nod (K + 16)," "LS-nod-noSN (B + 20)," "HS-nod-noSN (B + 20)," "LS-nod-SN (B + 20),","HS-nod-SN (B + 20)," and "HS-nod-SN-high-accr (B + 20)" assume no delays between galaxy and MBH mergers (although they account for the delays between halo and galaxy mergers; cf. Sec. II)[3]; (ii) "popIII-d (K + 16)," "Q3-d (K + 16)," "LS-d (B + 18)," "HS-d (B + 18)," "LS-noSN-sd (B + 20)," "LS-SN-sd (B + 20)," "HS-noSN-sd (B + 20)," and "HS-SN-sd (B + 20)" account not only for the delays between halo and galaxy mergers but also for stellar hardening, MBH triplets, and planetarylike migration; (iii) "LS-noSN-d (B + 20)," "LS-SN-d (B + 20)," "HS-noSN-d (B + 20)," "HS-SN-d (B + 20)," "LS-inf-d (B + 18)," and "HS-inf-d (B + 18)" account for the full delays; i.e., they also include the evolution of black hole pairs at separations of hundreds of pc (cf. Sec. II). The labels "SN" and "noSN," respectively, indicate models that do and do not account for the effect of SN feedback on the growth of the nuclear gas reservoir, while "LS"/"popIII" and "HS"/"Q3" denote, respectively, LS and HS scenarios for the high-redshift initial mass function.

Each model's predictions are compared to the EPTA measurement in Fig. 1, which is adapted from Fig. 8 of Ref. [8] and which we report here for completeness. The model predictions assume quasicircular orbits for the MBH binaries, as a result of which the spectrum has a slope $\gamma = 13/3$, and are obtained by summing the GW energy spectra of the whole theoretical binary population. For this reason, the predictions do not have any cosmic variance; i.e., at this stage we do not consider the scatter among different realizations of the binaries in our past light cone. The shaded area is the EPTA measurement of the amplitude at $f = 1/10$ yr assuming $\gamma = 13/3$.

Note that each model's prediction [except model HS-nod-SN-high-accr (B + 20)] is shown as a range. The lower end represents the prediction of finite merger tree resolution, while the upper end is the extrapolation to infinite resolution. This extrapolation is performed following

---

[3]Models LS-nod-noSN (B + 20), HS-nod-noSN (B + 20), LS-nod-SN (B + 20), HS-nod-SN (B + 20), and HS-nod-SN-high-accr (B + 20) are produced with the semianalytic model of B + 20, but with no delays between galaxy and MBH mergers (except for the dynamical friction timescale—including tidal effects—between dark matter halos). These models were not presented in B + 20, but only in Ref. [8].
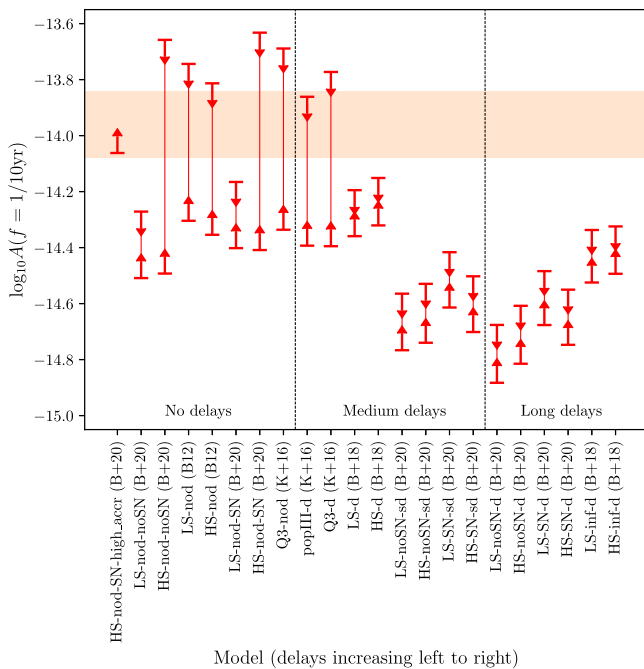
FIG. 1. Characteristic strain amplitude at $f = 1/10$ yr predicted by various published semianalytic galaxy formation models, assuming quasicircular inspiral, no interactions with matter, and without accounting for cosmic variance (i.e., for the scatter among different realizations of the black hole population). The models feature different physical assumptions for the delays between the galaxy and black hole mergers (increasing from left to right). The "no delays," "medium delays," and "long delays" models correspond to the model classes (i), (ii), and (iii) defined in the text, respectively. For each model, the lower bound is the result at finite resolution, while the upper one is the extrapolation to infinite resolution (see text for details). The shaded area is the 95% confidence region from the EPTA measurement assuming $\gamma = 13/3$ [8]. This figure is adapted from Fig. 8 of Ref. [8].
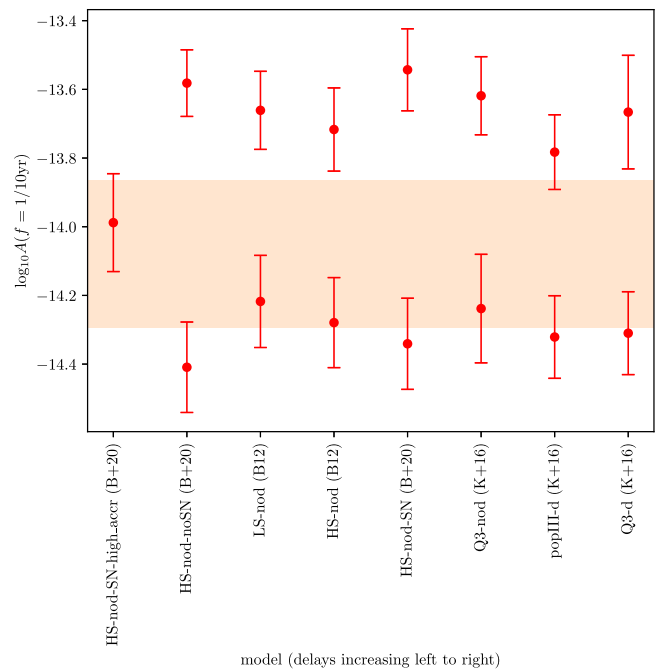

FIG. 2. Predicted characteristic strain amplitude at $f = 1/10$ yr from various semianalytic galaxy formation models. The amplitude is obtained by producing multiple realizations of the population of MBH binaries, and then by fitting the resulting spectrum in the first nine frequency bins (assuming an observation time of 10.3 yr) with a power law. The error bars account for the scatter among the different realizations (cosmic variance) and represent 95% confidence intervals. For each model, we present both the result at finite resolution (lower) and the extrapolation to infinite resolution (higher), except for model HS-nod-SN-high-accr (B + 20) (finite resolution only). The shaded area is the 95% confidence region from the EPTA measurement (marginalizing over the spectrum's slope) [8].

Fig. 4 of [33] (i.e., as described in Sec. II B), but because of the uncertainties involved, it should be considered as an upper limit. For model HS-nod-SN-high-accr (B + 20), we report only the (more robust) finite resolution prediction, which agrees already with the measurement. This agreement is the result of stronger black hole accretion, which is obtained in this model by boosting the influx of gas into the nuclear reservoir during star formation events in the spheroid by a factor ~4.

In Fig. 2, we consider only the models in better agreement with the data.[4] Here, we do not sum the spectra of the whole theoretical population, but we consider multiple realizations of the binaries in our past light cone, following Ref. [78]. We then fit each realization's predicted spectrum

---

[4]Given the uncertainties in the PTA measurements and the fact that we assume quasicircular binaries in our predictions, we consider all models in Fig. 2 to be in broad agreement with the PTA results.

with a power law in the first nine frequency bins and report the average signal amplitude $A(f = 1/10 \text{ yr})$ and its 95% confidence region. Note that this region (i.e., the error bars) should therefore be interpreted as representing cosmic variance. These predictions are compared to the corresponding range of the measurement, i.e., the 95% confidence region of the amplitude $f = 1/10$ yr, but this time we marginalize over the slope γ (cf. Fig. 1 of Ref. [8]), since the model predictions do not have $\gamma = 13/3$ as a result of the finite number of realizations [78].

As stressed in Ref. [8], this comparison allows one to draw the following qualitative conclusions: (i) Large delays at separations of hundreds of pc are disfavored, and MBHs merge efficiently after galaxy mergers, and (ii) accretion onto MBHs may be more efficient than previously thought, resulting potentially in a larger local MBH mass function at the high-mass end.

To check that the boost in the reservoir growth rate in model HS-nod-SN-high-accr (B + 20) did not result in an overprediction of the local bolometric quasar luminosity
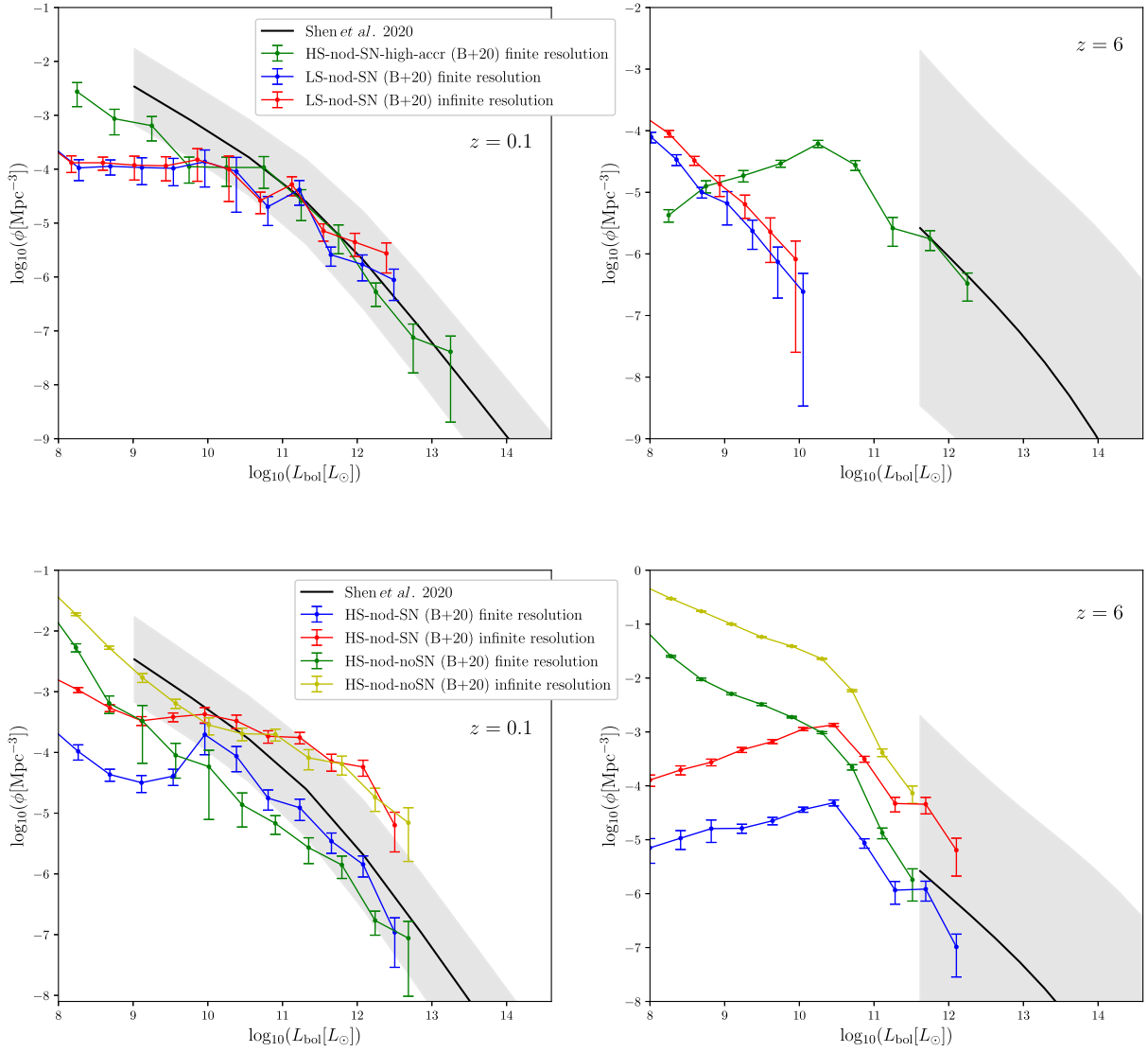
FIG. 3.   Bolometric quasar luminosity function at $z = 0.1$ (left panels) and $z = 6$ (right panels) for different models described in the text. The black line with shaded 95% confidence interval is the observed bolometric luminosity function from Ref. [79].

function, in the upper panels of Fig. 3 we compare to the data reported in Ref. [79], which show good agreement.[5] Similarly, because it is expected [37,58] that LS models can only reproduce the luminosity function of quasars at $z = 6$ in the presence of a sustained phase of super-Eddington accretion, in the same panels we also show results from LS-nod-SN (B + 20). As can be seen, model LS-nod-SN (B + 20) struggles to reproduce the observed luminosity function as a result of SN feedback, as also found in Ref. [37]. We will include model LS-nod-SN (B + 20)

when making predictions for LISA, but one should keep in mind that it is marginally disfavored by the $z = 6$ quasar luminosity function. In the lower panels of Fig. 3, we show results from models HS-nod-noSN (B + 20) and HS-nod-SN (B + 20) for comparison.

IV. GW EMISSION, DETECTION, AND PARAMETER ESTIMATION

We produce synthetic catalogs of MBH binaries to characterize their emitted GW signals. Our main goal is to assess their detection rate and the parameter estimation prospects for LISA. We describe the GW signal from an MBH binary by a set of 11 parameters: the masses M_1 and M_2, the (aligned) spins χ_1 and χ_2, the time of coalescence t_c, the luminosity distance D_L, the inclination of the angular momentum of the source relative to the line

---

[5]Note that it is not obvious that our boosted accretion model will agree with the local luminosity function. In our case, it does agree because of our treatment of the radiative efficiency at high accretion rates [cf. Eq. (40) of B12, which approximately represents the transition from thin to slim disk accretion as the mass accretion rate increases].

of sight ι, the sky longitude and latitude λ and β, the polarization angle ψ, and the phase at coalescence ϕ. If one were to include precession in the analysis, there would be extra parameters describing the orientation of the spins. In this study, we employ IMRPhenomHM [80], which is a nonprecessing waveform model that describes binaries with aligned spins,[6] but which includes contributions from higher order modes. The inclusion of higher harmonics helps measure the source parameters, and particularly the luminosity distance.

The signal for LISA has two main components: the GW waveform from the source and the LISA response function. In the frequency domain, the waveform takes the form $h_{\ell m} = A_{\ell m}(f)e^{-i\Psi_{\ell m}(f)}$. The single-arm interferometric measurement, which describes the shift in frequency of the laser linkage l between the spacecraft's s and r, is

$$y_{slr} = \sum_{\ell,m} \mathcal{T}_{slr}^{\ell m}(f)h_{\ell m}, \qquad (1)$$

where $\mathcal{T}_{slr}^{\ell m}$ is the transfer function for the ℓm mode. A computation of $\mathcal{T}_{slr}$ without any approximations, although possible, is unfeasible for parameter estimation due to its high computational cost. A fast method of calculating it using a perturbative approach was extended to merger ringdown in Ref. [82], making the computation suitable for use in parameter estimation.

Laser noise dominates the single link observable y_{slr}. However, time-delay interferometry (TDI) [83–86], which relies on three independent time-delayed linear combinations (referred to as the A, E, and T channels) of the single link outputs, removes the laser noise. The T channel has very low signal content, and therefore is not included in our analysis; i.e., we will only work with the A and E channels.

We simulate gravitational waveforms for each binary in our synthetic catalogs, neglecting the fact that the true LISA data will contain overlapping signals (i.e., we consider each source "in isolation"). We assume noiseless simulated data when computing signal-to-noise ratios (SNRs) and parameter estimation. We also neglect possible data gaps and glitches [87–89], but include the LISA response function. The inclusion of a Gaussian noise realization in the simulated data is expected to shift the posterior distribution up to the statistical error but is unlikely to change the variance of the posterior distribution [90–92].

To assess the detectability of a GW source with LISA, we compute its SNR ρ defined as

$$\rho^2 = (h|h), \qquad (2)$$

where h is the GW signal, and the inner product $(h|h)$ is defined as

$$(a|b) = 4\text{Re}\int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f)}{S_n(f)}\,df, \qquad (3)$$

where $S_n(f)$ is the noise power spectral density. For the latter, we use the SciRDv1 noise model [93] with the addition of an unresolved white dwarf background from Galactic binaries [94,95]. As we have higher modes in the waveform, the calculation of ρ^2 results in non-negligible cross terms between different modes [Eq. (26) of Ref. [96]]. Moreover, the total SNR is the sum (in quadrature) of the contributions of two independent TDI observables A and E.

Each detection (with $\rho \geq 8$) is followed by a full parameter estimation. This provides information on the number of sources for which we can estimate the posterior distribution with errors lower than a specified threshold. According to Bayes's theorem, the posterior distribution of the parameters θ given the data d is provided by

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)p(\theta)}{p(d)}, \qquad (4)$$

where $\mathcal{L}(d|\theta)$ is the likelihood, $p(\theta)$ is the prior, and $p(d)$ is the evidence. The evidence acts as a normalization factor and is useful when comparing different models. While sampling, we choose the following parameters: total mass M, mass ratio q, χ_1, χ_2, t_c, D_L, ι, λ, β, ψ, and ϕ. We choose uniform priors on M, q, χ_1, χ_2, t_c, D_L, $\cos\iota$, ψ, and ϕ. For the sky position, we choose uniform priors on λ and $\sin\beta$.

The likelihood is computed using the LISAbeta code described in [97]. To sample the posterior distribution, we use ptemcee [98] which implements a parallel-tempered ensemble sampler. Parallel tempering [99,100] is a variation of the commonly used Markov-chain-Monte-Carlo-based techniques, and it is efficient at sampling complex posterior surfaces with multiple sharp peaks. As the source parameters of the injected data are known, and our aim is to estimate the parameter errors from the posterior distribution, we can further help the sampler by initializing it at the injection's ground truth, with a preliminary covariance of the parameters from the Fisher matrix. The Fisher matrix elements are calculated as

$$F_{ij}(\theta_{\text{inj}}) = (\partial_i h|\partial_j h)|_{\theta_{\text{inj}}}, \qquad (5)$$

where ∂_i is the derivative with respect to the ith element in the parameter set θ, and θ_{inj} is the injection's ground truth. Initializing the sampler around the true value of the parameters, and using the Fisher matrix in the preliminary stage, we can focus directly on the regions with high

---

[6]While a recent work [81] included both precession and higher modes in the analysis of MBH binaries observable by LISA, we focus here on the nonprecessing case for simplicity. It should be noted, however, that precession can play a crucial role in binary evolution. This will be addressed in a future work.

posterior probability, without having to sample the full parameter space. The chains, therefore, converge faster, leading to lower computational costs.

This approach, unfortunately, will ignore secondary peaks in the posterior distribution, if proposal distributions are not tuned accordingly. MBH signals are known to present degeneracies among sky position, inclination, and polarization [97]. We, therefore, choose the proposal distributions to allow the sampler to jump between these degenerate modes.

V. RESULTS

A. LISA detection rates

For detectability, we consider an SNR threshold of 8. Predictions for the number of LISA detections for a four-year mission are listed in Table I for the different models. As can be seen, LS models predict fewer detectable sources than HS ones, because lower BH masses lead to lower SNRs. In more detail, lower-mass sources merge at higher frequencies. Although LISA is nominally sensitive in the band ~0.01 mHz to ~0.1 Hz, at frequencies close to these boundary values the noise dominates, as can be seen from the noise power spectral density [93]. As the merger is the loudest part of an MBH binary signal, missing it results in a lower SNR. In fact, for very low masses $\sim 10^3 M_\odot$, the merger may even be completely outside the LISA sensitivity band. This effect is exacerbated by the effect of SN feedback, which may hamper accretion in low-mass/high-redshift galaxies (where SN winds can exceed the typical escape velocities) [37,53]. While this lack of accretion growth does not prevent binaries from radiating in the LISA band in HS models (where the seed masses already correspond to frequencies LISA is sensitive at), in LS models it leads to very low binary masses at high z. As a result, the LS-nod-SN (B + 20) model is the one with the lowest number of events. However, as mentioned above, this model is in marginal tension with the observed quasar luminosity function at $z = 6$. Also noticeable in Table I is the difference in the number of detections between the models of B12 and K + 16[7] and the later models of B + 20. This is due to the updated procedure to resolve early mergers in subresolution branches of the merger tree, which we discussed in Sec. II and which was introduced in the models of B + 20. An exception is model HS-nod (B12), which predicts a large number of detections because it adopts a different seeding scenario, namely, that of

---

[7]We stress that the detection rates computed for the models K + 16 are consistent with those reported in Ref. [33] for the N2A2M5L6 configuration, once one accounts for (i) the different mission duration (five years in Ref. [33] vs four years here), (ii) the slightly different detector configuration (2 Gm arms in Ref. [33] vs 2.5 Gm here), and (iii) the inclusion of the full LISA response and higher modes (both absent in Ref. [33]), which are particularly important for LS models.

TABLE I. Number of detections (N_{det}) for the different models described in the text, in a four-year LISA mission.

Model	$N_{\text{det}}(4 \text{ yr})$ finite res.	$N_{\text{det}}(4 \text{ yr})$ inf. res.
HS-nod-SN-high-accr (B + 20)	8901	…
LS-nod-noSN (B + 20)	203	250
HS-nod-noSN (B + 20)	15821	38712
LS-nod (B12)	432	570
HS-nod (B12)	6154	7184
LS-nod-SN (B + 20)	11	12
HS-nod-SN (B + 20)	16133	36090
Q3-nod (K + 16)	468	656
popIII-d (K + 16)	183	339
Q3-d (K + 16)	33	74

Ref. [59]. The results of B12 and K + 16 can be therefore viewed as conservative.

Note also that the impact of the extrapolation to infinite resolution depends on the seeding model and the distribution of the MBH mergers. For instance, our HS models are more affected by finite resolution effects, because the seeds can form in relatively small halos [56], while LSs are assumed to form in the largest halos at high redshift [39,55]. Similarly, no-delay models tend to predict more mergers at high z, where finite resolution may be an issue. These subtle effects are reflected in the different extrapolated rates in Table I.[8]

To further visualize the models' physical content and understand their implications for LISA, in Fig. 4 we show the number of mergers (with no threshold in the SNR) in four years for models HS-nod-SN-high-accr (B + 20) (left panel) and LS-nod-noSN (B + 20) (right panel), as a function of the redshift and source-frame chirp mass. Also shown are SNR contours (computed assuming $q = 1$) highlighting the parameter space region in which LISA is sensitive. The side panels report marginalized distributions over chirp mass and redshift, for all systems (intrinsic) and the detected ones. Similar figures are shown in Appendix for the other models (at finite resolution; figures at infinite resolution only show a slightly higher normalization but are otherwise very similar). Again, one can notice that the LS models (and especially those with SN feedback) lead to very few systems in the region where LISA is sensitive; i.e., only a fraction of the whole population is detectable. In the HS models, instead, the detection fraction is close to 100%. Moreover, in the HS models, the detected event has typically larger SNR (up to thousands) than in the LS models, again as a result of the higher masses. These larger SNRs will result in correspondingly low errors on the parameters in the next

---

[8]On top of this, the old results of Ref. [39] were produced with slightly lower resolution than the more recent results of Refs. [35,37,40–42], as mentioned above.

HS-nod-SN-high_accr (B+20) finite resolution

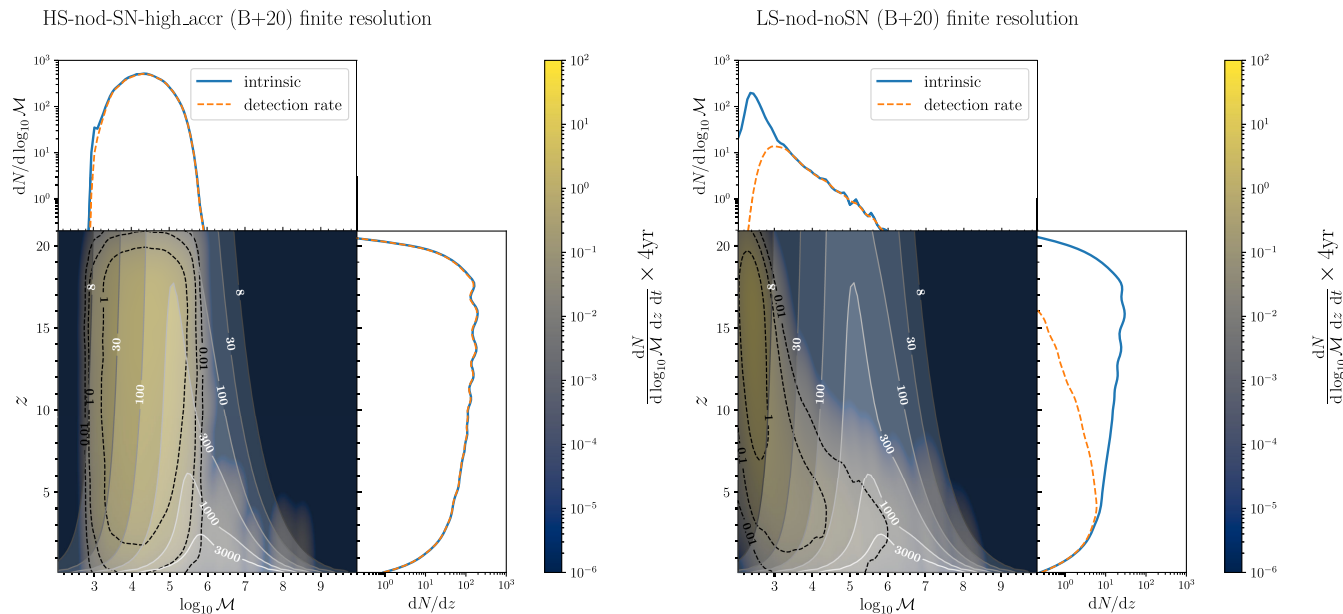LS-nod-noSN (B+20) finite resolution

FIG. 4.   Number of mergers (color gradient and black dashed lines) in four years for models HS-nod-SN-high-accr (B + 20) (left panel) and LS-nod-noSN (B + 20) (right panel), as a function of the source-frame chirp mass and redshift. Superimposed (labeled by white numbers), we show the LISA sensitivity in terms of contours of constant SNR (assuming $q = 1$). In the upper and right panels, we show the distribution of the total (intrinsic) number of mergers (blue line) and that of the detected number (orange dashed line), marginalizing, respectively, over the redshift and chirp mass.

section, where we will explore the parameter estimation capabilities of LISA.

B. Parameter estimation

To assess the LISA parameter estimation capabilities, which are of course crucial to extract science from the data, we consider the detected sources (with $\rho > 8$) and apply to them our Bayesian pipeline described in Sec. IV. We run the sampler with 64 walkers and 10 temperatures to produce 8000 samples per detected source. For the calculation of the likelihood integral, we set $f_{\min} = 10^{-5}$ Hz and $f_{\max} = 0.5$ Hz, which corresponds to the region of the LISA sensitivity curve most favorable to MBH binary detections.[9] As can be observed in Table I, the number of detections is very large for models HS-nod-SN-high-accr (B + 20), HS-nod-noSN (B + 20), HS-nod (B12), and HS-nod-Sn (B + 20); therefore, for these models, we run

---

[9]It is worth noting that these frequency limits indicate the mission goals, while the mission requirement for the lower bound of the LISA band is $f_{\min} = 10^{-4}$ Hz [93]. This difference in the lower bounds is unlikely to have a significant impact on SNR calculations, as the sensitivity is quite low in the band 10^{-5}–10^{-4} Hz. As for parameter estimation, inclusion of higher modes (as we do here) can mitigate potential degradation in the estimated posterior distributions for short-lived signals [101]. Moreover, for the vast majority of signals, which are in the LISA band for longer periods (months or years), the exact choice of the lower frequency bound is unlikely to have an appreciable impact.

the parameter estimation pipeline on a subset of sources and extrapolate to the whole detected population. Model LS-nod-SN (B + 20) has the opposite problem; i.e., because it has a very low detection rate, to obtain statistically relevant conclusions, we draw our Bayesian inference for more sources than those that would be
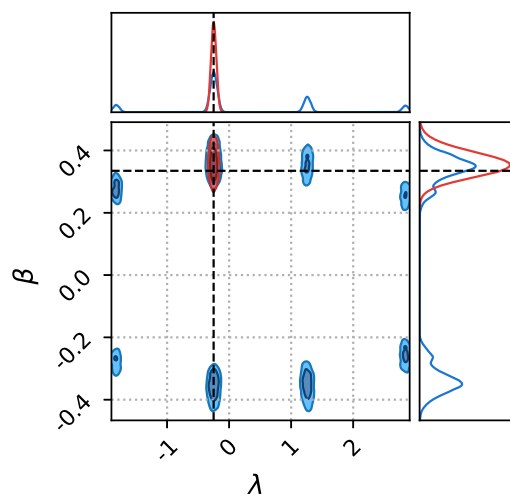

FIG. 5.   Multimodality in the sky position measurement. The "raw" posterior distribution without any postprocessing is shown in blue. The primary maximum, to which we restrict for error computation, is shown in red. The parameters of this source are $(M_1 = 2.3 \times 10^6 M_\odot, M_2 = 1.8 \times 10^5 M_\odot, \chi_1 = 0.03, \chi_2 = -0.02, t_c = 2522880$ s, $D_L = 152$ Gpc, $\iota = 2.4, \lambda = -0.25, \beta = 0.33, \psi = 0.92, \phi = -0.64)$.

detected in four years. More precisely, we consider 100 years of data from the synthetic catalogs and rescale the rate to the LISA mission duration.

From each of the estimated posteriors, we derive the covariance matrix of the source parameters. The error in the measurement of a parameter θ can be described by $\Delta\theta = \sqrt{\mathrm{var}(\theta)}$ where $\mathrm{var}(\theta)$ is the variance of the parameter θ computed from the posterior samples. We are mainly interested in the masses, the spins, the luminosity distance, and the sky position of the sources. The error in the sky position ($\Delta\Omega$), which is very important for performing multimessenger astronomy [102], is calculated as $\Delta\Omega = -2\pi \log(1-p)\sqrt{(\mathrm{var}(\lambda)\mathrm{var}(\sin\beta) - \mathrm{cov}(\lambda,\sin\beta)^2)}$ [103], where λ and β are the longitude and latitude of the source in the sky, and p is the credibility level (which we fix to 90%). For the luminosity distance, in addition to the statistical error from the posteriors, the weak-lensing error

is expected to be non-negligible. As in Ref. [102], we estimate it following Ref. [104], which gives

$$\sigma_{\mathrm{WL}}(z) = D_L \times 0.066 \left(\frac{1 - (1+z)^{-0.25}}{0.25} \right)^{1.8}, \quad (6)$$

and we add this error in quadrature to the statistical error. Errors on individual masses, spins, and sky position are presented without the effect of weak lensing. The effect of the latter, however, is expected to be small, as these quantities have little correlation with luminosity distance.

Our choice to summarize the posteriors by their covariance matrix is potentially problematic due to multimodalities. The posterior distribution of the sky position of the MBH binaries detected by LISA may indeed be multimodal [97,105]. Depending on the parameters of the system, one may observe up to eight modes in the sky. This is due to the
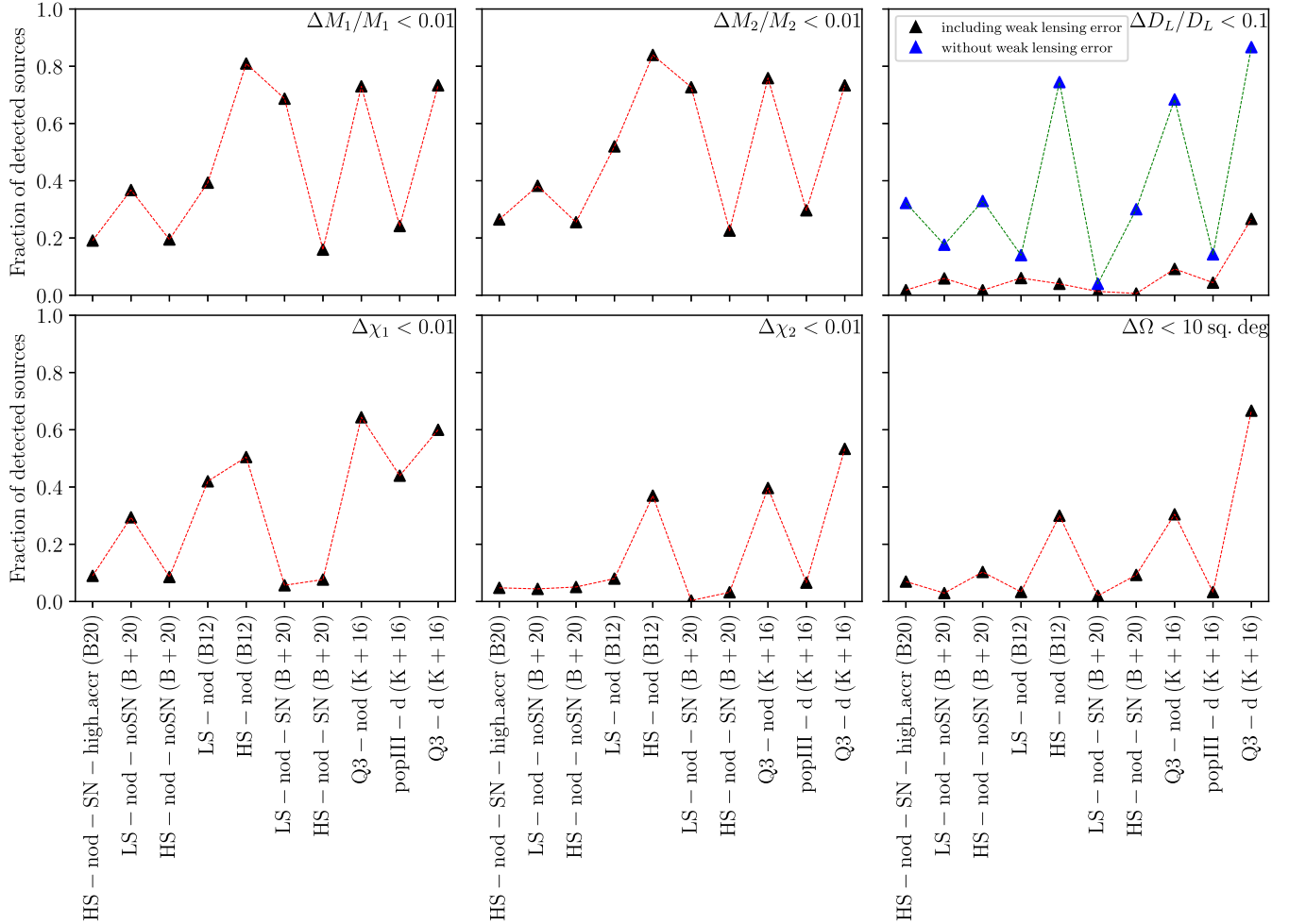

FIG. 6. Fraction of detected sources satisfying given thresholds on the parameter errors. In more detail, the top panels, from left to right, show sources with $\Delta M_1/M_1 < 0.01$, $\Delta M_2/M_2 < 0.01$, and $\Delta D_L/D_L < 0.1$; the bottom panels show sources with $\Delta\chi_1 < 0.01$, $\Delta\chi_2 < 0.01$, and $\Delta\Omega < 10$ sq deg. In the panel for D_L, we also show the fraction of detected sources within the threshold when the weak-lensing error is excluded.

LISA antenna pattern functions, but the motion of the LISA instrument during observations and the frequency-dependent features of the instrument response at high frequencies can help to break this degeneracy. Indeed, the mode containing the true source parameter (the ground truth) typically has the highest posterior probability. Figure 5 shows an example of the multimodality commonly seen in sky location. In other cases, one may observe only bimodality in β, or the degeneracy may even be entirely absent with only a single peak in the posterior. In our analysis, we restrict ourselves to the primary peak alone (shown in red in Fig. 5): We compute the location of the degenerate modes [97], divide the sky into quadrants so that each quadrant contains only one mode, and then we restrict the posterior to the quadrant with the true sky position and compute the associated sky position error. As a proxy for the effect of multimodalities, in cases where the true sky location does not lie in the quadrant with the maximum probability, we discard the source from further analysis and assume that it is simply not localized. This treatment is approximate, but multimodal posteriors are expected to represent a minority of the loud, nearby sources that are the best candidates for multimessenger observations [105].

Next, to characterize the LISA performance, we compute the fraction of the sources for which the parameter estimation can be performed better than the given thresholds. More precisely, we compute the fraction of detected sources for which the true sky position lies in the same quadrant as the peak of the posteriors, and the masses can be estimated to 1% level or better, the spins can be estimated to within an absolute error of 0.01, the luminosity distance can be estimated to 10% or better, or the sky position can be estimated with an error of less than 10 sq deg. These thresholds are meant to be at 1σ, except for the sky position (for which we set $p = 0.9$). The results of this analysis are shown in Fig. 6. As can be seen, the masses can be estimated well (i.e., to the 1% threshold set above or better) for $\sim$20%–80% of the detections. This fraction is lower for the spins (with the primary spin χ_1 better determined than the secondary one, χ_2). Similarly, the luminosity distance can only be estimated to 10% error in $\sim$1%–30% of the detected events, partly as a result of weak lensing (Fig. 6 also shows results without the lensing error). The sky position is well estimated for $\sim$2%–70% of the detected sources, depending on the model.

The difference between different models depends on the details of the populations that they predict. For instance, the Q3-d (K + 16) model predicts more sources at low z, which explains the larger fraction of events with well-estimated distance and sky location. HS models predict larger SNRs, which favor precise parameter estimation, but LS models can also lead to good determination of the component masses and spins, thanks to the inspiral of these sources being in the LISA band for a long time. We also stress that although the fractions are shown in Fig. 6 may appear low,

they need to be multiplied by the event rates shown in Table I, and can therefore result in a significant number of well-characterized sources, especially for the HS models.

VI. CONCLUSIONS

In this paper, we have reassessed the prospects for detection and parameter estimation of MBH binaries with LISA, accounting for the constraints on semianalytic models for the formation and evolution of MBHs from the recent detection of a GW background from PTA observatories. This extraordinary discovery seems to suggest that large delays between galaxy and MBH mergers are disfavored and that MBHs merge efficiently after galaxy mergers. Moreover, accretion onto MBHs seems to be more efficient than previously thought, leading to a larger local MBH mass function at high masses. We have therefore analyzed models that are in good agreement with PTA data and produced synthetic catalogs of MBH binaries to characterize their emitted GW signals for LISA.

We have found that LISA will detect at least a dozen MBH mergers in four years if we include models that struggle to reproduce the quasar luminosity function at $z = 6$. However, this number rises to $\gtrsim 100$ if we exclude such models. In fact, models that are in better agreement with the quasar luminosity function at $z = 6$ can predict from several hundred to a few tens of thousands of detections (in some cases; cf. Table I), potentially raising the question of whether one will be able to resolve these signals singularly. It should be noticed that similarly high detection numbers for LISA have been recently found by Ref. [38] on the basis of agnostic and astrophysically informed merger rate models calibrated to the PTA data. In fact, the event rates found by Ref. [38] are even higher than ours, which can probably be ascribed to the differences in the population modeling (ours is based on a full-fledged semianalytic galaxy formation model, while theirs is based on a phenomenological parametrization of the merger rate).

We have found that the models with the highest detection rate (namely, HS models) also have a larger SNR, which translates into better accuracy when performing parameter estimation. The fraction of detected sources for each model, satisfying given thresholds on the parameter errors, is summarized in Fig. 6. An important caveat is that we have performed Bayesian parameter estimation on the single events without accounting for the presence of the other (overlapping) MBH binaries. The superposition of hundreds or thousands of MBH signals would pose a serious challenge to LISA data analysis, and the recent detection of a stochastic GW background from PTA experiments is warning us that this may be a concrete possibility.

We stress that the fact that the PTA data seem to favor models with short or no delays is surprising, but not to the point of being (overly) concerning. First, as explained in Sec. II, even the no-delay models include at least the timescale related to the dynamical friction (including tidal

effects) between dark matter halos. That may be slightly overestimated (due to the uncertainties in the underlying models [60,61]), and at least partially compensate for additional "baryonic" delays. More in general, the latter are computed in our semianalytic model based on simple formulas [35,37,41,42] calibrated to simulations (which are themselves affected by uncertainties due to resolution and subgrid physics), which could cause significant errors. Finally, the no-delay and short-delay models being preferred may point to gas being more important than previously thought in MBH mergers, at least at high masses and low redshift. This possibility may be checked with future PTA measurements, as gas interactions, besides providing very short delays, also tend to flatten the PTA spectrum [42]. Eventually, since the dynamics leading to the delays between galaxy/halo mergers and MBH mergers may (and probably will) be different at high redshift and low masses than it is in the local Universe for MBHs in the PTA band, the ultimate and most complete answer to this puzzle will be provided by the LISA observations themselves.

APPENDIX: MERGER AND DETECTION RATES FOR DIFFERENT MODELS

In this appendix, we show the mass and redshift distribution of the total and the detected number of binaries, in Fig. 7, following the same format as in Fig. 4, but for the other models used in this work.
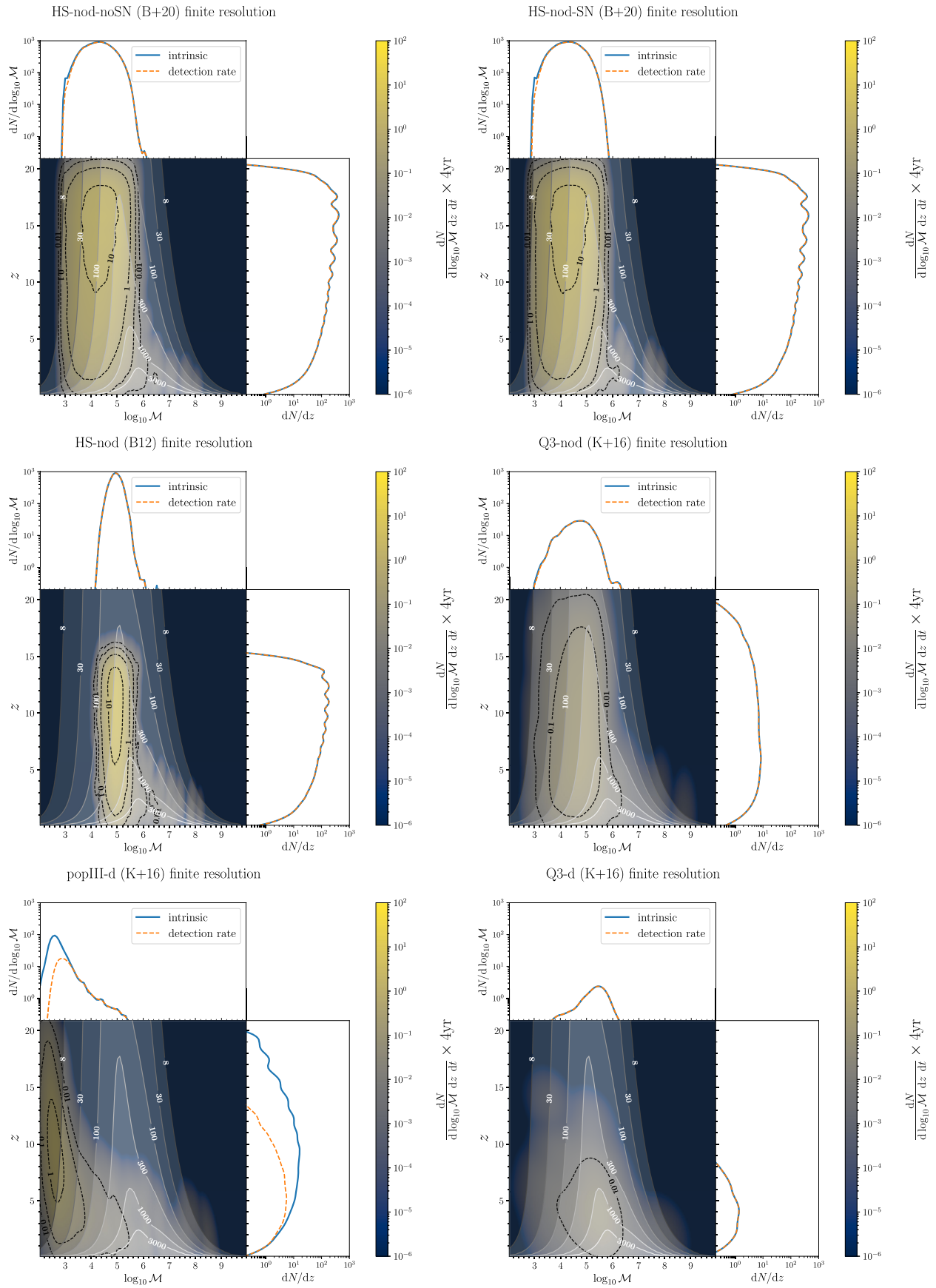
FIG. 7.   The same as Fig. 4, for the other models used in this work.

[1] J. Antoniadis *et al.*, The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals, Astron. Astrophys. **678,** A50 (2023).

[2] P. Tarafdar *et al.*, The Indian Pulsar Timing Array: First data release, Pub. Astron. Soc. Aust. **39,** e053 (2022).

[3] G. Agazie *et al.* (NANOGrav Collaboration), The NANO-Grav 15-year data set: Evidence for a gravitational-wave background, Astrophys. J. Lett. **951,** L8 (2023).

[4] D. J. Reardon *et al.*, Search for an isotropic gravitational-wave background with the Parkes Pulsar Timing Array, Astrophys. J. Lett. **951,** L6 (2023).

[5] H. Xu *et al.*, Searching for the nano-hertz stochastic gravitational wave background with the Chinese Pulsar Timing Array Data Release I, Res. Astron. Astrophys. **23,** 075024 (2023).

[6] C. Smarra *et al.* (European Pulsar Timing Array Collaboration), The second data release from the European Pulsar Timing Array: VI. Challenging the ultralight dark matter paradigm, Phys. Rev. Lett. **131,** 171001 (2023).

[7] H. Quelquejay Leclere *et al.* (EPTA Collaboration), Practical approaches to analyzing PTA data: Cosmic strings with six pulsars, arXiv:2306.12234.

[8] J. Antoniadis *et al.*, The second data release from the European Pulsar Timing Array: V. Implications for massive black holes, dark matter and the early Universe, arXiv:2306.16227.

[9] A. Afzal *et al.* (NANOGrav Collaboration), The NANO-Grav 15-year data set: Search for signals from new physics, Astrophys. J. Lett. **951,** L11 (2023).

[10] D. G. Figueroa, M. Pieroni, A. Ricciardone, and P. Simakachorn, Cosmological background interpretation of Pulsar Timing Array Data, arXiv:2307.02399.

[11] G. Agazie *et al.* (NANOGrav Collaboration), The NANO-Grav 15-year data set: Constraints on supermassive black hole binaries from the gravitational wave background, Astrophys. J. Lett. **952,** L37 (2023).

[12] T. Gehren, J. Fried, P. A. Wehinger, and S. Wyckoff, Host galaxies of quasars and their association with galaxy clusters, Astrophys. J. **278,** 11 (1984).

[13] J. Kormendy and D. Richstone, Inward bound—the search for supermassive black holes in galactic nuclei, Annu. Rev. Astron. Astrophys. **33,** 581 (1995).

[14] A. E. Reines, G. R. Sivakoff, K. E. Johnson, and C. L. Brogan, An actively accreting massive black hole in the dwarf starburst galaxy Henize2-10, Nature (London) **470,** 66 (2011).

[15] A. E. Reines, J. E. Greene, and M. Geha, Dwarf galaxies with optical signatures of active massive black holes, Astrophys. J. **775,** 116 (2013).

[16] V. F. Baldassare, M. Geha, and J. Greene, A search for optical AGN variability in 35,000 low-mass galaxies with the Palomar Transient Factory, Astrophys. J. **896,** 10 (2020).

[17] D. J. Croton, V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida, The many lives of active galactic nuclei: Cooling flows, black holes and the luminosities and colours of galaxies, Mon. Not. R. Astron. Soc. **365,** 11 (2006).

[18] J. Kormendy and L. C. Ho, Coevolution (or not) of supermassive black holes and host galaxies, Annu. Rev. Astron. Astrophys. **51,** 511 (2013).

[19] N. J. McConnell and C.-P. Ma, Revisiting the scaling relations of black hole masses and host galaxy properties, Astrophys. J. **764,** 184 (2013).

[20] M. Schramm and J. D. Silverman, The black hole-bulge mass relation of active galactic nuclei in the extended Chandra Deep Field-South survey, Astrophys. J. **767,** 13 (2013).

[21] M. Volonteri and A. E. Reines, Inferences on the relations between central black hole mass and total galaxy stellar mass in the high-redshift universe, Astrophys. J. Lett. **820,** L6 (2016).

[22] F. Shankar, M. Bernardi, R. K. Sheth, L. Ferrarese, A. W. Graham, G. Savorgnan, V. Allevato, A. Marconi, R. Läsker, and A. Lapi, Selection bias in dynamically measured supermassive black hole samples: Its consequences and the quest for the most fundamental relation, Mon. Not. R. Astron. Soc. **460,** 3119 (2016).

[23] E. Barausse, F. Shankar, M. Bernardi, Y. Dubois, and R. K. Sheth, Selection bias in dynamically measured supermassive black hole samples: Scaling relations and correlations between residuals in semi-analytic galaxy formation models, Mon. Not. R. Astron. Soc. **468,** 4782 (2017).

[24] J. E. Greene, J. Strader, and L. C. Ho, Intermediate-mass black holes, Annu. Rev. Astron. Astrophys. **58,** 257 (2020).

[25] L. L. Cowie, A. Songaila, E. M. Hu, and J. G. Cohen, New insight on galaxy formation and evolution from Keck spectroscopy of the Hawaii deep fields, Astron. J. **112,** 839 (1996).

[26] F. Fontanot, G. De Lucia, P. Monaco, R. S. Somerville, and P. Santini, The many manifestations of downsizing: Hierarchical galaxy formation models confront observations, Mon. Not. R. Astron. Soc. **397,** 1776 (2009).

[27] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[28] M. Volonteri, F. Haardt, and P. Madau, The assembly and merging history of supermassive black holes in hierarchical models of galaxy formation, Astrophys. J. **582,** 559 (2003).

[29] A. Sesana, M. Volonteri, and F. Haardt, The imprint of massive black hole formation models on the LISA data stream, Mon. Not. R. Astron. Soc. **377,** 1711 (2007).

[30] A. Sesana, J. Gair, E. Berti, and M. Volonteri, Reconstructing the massive black hole cosmic history through gravitational waves, Phys. Rev. D **83,** 044036 (2011).

[31] J. E. Plowman, R. W. Hellings, and S. Tsuruta, Constraining the black hole mass spectrum with gravitational wave observations—II. Direct comparison of detailed models, Mon. Not. R. Astron. Soc. **415,** 333 (2011).

[32] V. Rodriguez-Gomez, S. Genel, M. Vogelsberger, D. Sijacki, A. Pillepich, L. V. Sales, P. Torrey, G. Snyder, D. Nelson, V. Springel, C.-P. Ma, and L. Hernquist, The merger rate of galaxies in the Illustris simulation: A comparison with observations and semi-empirical models, Mon. Not. R. Astron. Soc. **449,** 49 (2015).

[33] A. Klein, E. Barausse, A. Sesana, A. Petiteau, E. Berti, S. Babak, J. Gair, S. Aoudia, I. Hinder, F. Ohme, and

B. Wardell, Science with the space-based interferometer eLISA: Supermassive black hole binaries, Phys. Rev. D **93,** 024003 (2016).

[34] A. Ricarte and P. Natarajan, The observational signatures of supermassive black hole seeds, Mon. Not. R. Astron. Soc. **481,** 3278 (2018).

[35] M. Bonetti, A. Sesana, F. Haardt, E. Barausse, and M. Colpi, Post-Newtonian evolution of massive black hole triplets in galactic nuclei—IV. Implications for LISA, Mon. Not. R. Astron. Soc. **486,** 4044 (2019).

[36] P. Dayal, E. M. Rossi, B. Shiralilou, O. Piana, T. R. Choudhury, and M. Volonteri, The hierarchical assembly of galaxies and black holes in the first billion years: Predictions for the era of gravitational wave astronomy, Mon. Not. R. Astron. Soc. **486,** 2336 (2019).

[37] E. Barausse, I. Dvorkin, M. Tremmel, M. Volonteri, and M. Bonetti, Massive black hole merger rates: The effect of kiloparsec separation wandering and supernova feedback, Astrophys. J. **904,** 16 (2020).

[38] N. Steinle, H. Middleton, C. J. Moore, S. Chen, A. Klein, G. Pratten, R. Buscicchio, E. Finch, and A. Vecchio, Implications of pulsar timing array observations for LISA detections of massive black hole binaries, Mon. Not. R. Astron. Soc. **525,** 2851 (2023).

[39] E. Barausse, The evolution of massive black holes and their spins in their galactic hosts, Mon. Not. R. Astron. Soc. **423,** 2533 (2012).

[40] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, Linking the spin evolution of massive black holes to galaxy kinematics, Astrophys. J. **794,** 104 (2014).

[41] F. Antonini, E. Barausse, and J. Silk, The coevolution of nuclear star clusters, massive black holes, and their host galaxies, Astrophys. J. **812,** 72 (2015).

[42] M. Bonetti, A. Sesana, E. Barausse, and F. Haardt, Post-Newtonian evolution of massive black hole triplets in galactic nuclei—III. A robust lower limit to the nHz stochastic background of gravitational waves, Mon. Not. R. Astron. Soc. **477,** 2599 (2018).

[43] W. H. Press and P. Schechter, Formation of galaxies and clusters of galaxies by self-similar gravitational condensation, Astrophys. J. **187,** 425 (1974).

[44] H. Parkinson, S. Cole, and J. Helly, Generating dark matter halo merger trees, Mon. Not. R. Astron. Soc. **383,** 557 (2008).

[45] A. Dekel and Y. Birnboim, Galaxy bimodality due to cold flows and shock heating, Mon. Not. R. Astron. Soc. **368,** 2 (2006).

[46] A. Cattaneo, A. Dekel, J. Devriendt, B. Guiderdoni, and J. Blaizot, Modelling the galaxy bimodality: Shutdown above a critical halo mass, Mon. Not. R. Astron. Soc. **370,** 1651 (2006).

[47] A. Dekel, Y. Birnboim, G. Engel, J. Freundlich, T. Goerdt, M. Mumcuoglu, E. Neistein, C. Pichon, R. Teyssier, and E. Zinger, Cold streams in early massive hot haloes as the main mode of galaxy formation, Nature (London) **457,** 451 (2009).

[48] C. A. Correa, J. Schaye, J. S. B. Wyithe, A. R. Duffy, T. Theuns, R. A. Crain, and R. G. Bower, The formation of hot gaseous haloes around galaxies, Mon. Not. R. Astron. Soc. **473,** 538 (2018).

[49] F. Antonini, E. Barausse, and J. Silk, The imprint of massive black hole mergers on the correlation between nuclear star clusters and their host galaxies, Astrophys. J. Lett. **806,** L8 (2015).

[50] G. L. Granato, G. De Zotti, L. Silva, and L. Bressan, and A. Danese, A physical model for the coevolution of QSOs and their spheroidal hosts, Astrophys. J. **600,** 580 (2004).

[51] A. Lapi, S. Raimundo, R. Aversa, Z.-Y. Cai, M. Negrello, A. Celotti, G. De Zotti, and L. Danese, The coevolution of supermassive black holes and massive galaxies at high redshift, Astrophys. J. **782,** 69 (2014).

[52] A. Ricarte, M. Tremmel, P. Natarajan, and T. Quinn, Tracing black hole and galaxy co-evolution in the ROMULUS simulations, Mon. Not. R. Astron. Soc. **489,** 802 (2019).

[53] M. Habouzit, M. Volonteri, and Y. Dubois, Blossoms from black hole seeds: Properties and early growth regulated by supernova feedback, Mon. Not. R. Astron. Soc. **468,** 3935 (2017).

[54] M. A. Latif and A. Ferrara, Formation of supermassive black hole seeds, Pub. Astron. Soc. Aust. **33,** e051 (2016).

[55] P. Madau and M. J. Rees, Massive black holes as population III remnants, Astrophys. J. Lett. **551,** L27 (2001).

[56] M. Volonteri, G. Lodato, and P. Natarajan, The evolution of massive black hole seeds, Mon. Not. R. Astron. Soc. **383,** 1079 (2008).

[57] A. Heger and S. E. Woosley, The nucleosynthetic signature of population III, Astrophys. J. **567,** 532 (2002).

[58] P. Madau, F. Haardt, and M. Dotti, Super-critical growth of massive black holes from stellar-mass seeds, Astrophys. J. Lett. **784,** L38 (2014).

[59] S. M. Koushiappas, J. S. Bullock, and A. Dekel, Massive black hole seeds from low angular momentum material, Mon. Not. R. Astron. Soc. **354,** 292 (2004).

[60] M. Boylan-Kolchin, C.-P. Ma, and E. Quataert, Dynamical friction and galaxy merging time-scales, Mon. Not. R. Astron. Soc. **383,** 93 (2008).

[61] G. Taffoni, L. Mayer, M. Colpi, and F. Governato, On the life and death of satellite haloes, Mon. Not. R. Astron. Soc. **341,** 434 (2003).

[62] F. Dosopoulou and F. Antonini, Dynamical friction and the evolution of supermassive black hole binaries: The final hundred-parsec problem, Astrophys. J. **840,** 31 (2017).

[63] M. Tremmel, F. Governato, M. Volonteri, T. R. Quinn, and A. Pontzen, Dancing to CHANGA: A self-consistent prediction for close SMBH pair formation time-scales following galaxy mergers, Mon. Not. R. Astron. Soc. **475,** 4967 (2018).

[64] G. D. Quinlan, The dynamical evolution of massive black hole binaries I. Hardening in a fixed stellar background, New Astron. **1,** 35 (1996).

[65] A. Sesana and F. M. Khan, Scattering experiments meet N-body—I. A practical recipe for the evolution of massive black hole binaries in stellar environments, Mon. Not. R. Astron. Soc. **454,** L66 (2015).

[66] A. Sesana, F. Haardt, and P. Madau, Interaction of massive black hole binaries with their stellar environment. I. Ejection of hypervelocity stars, Astrophys. J. **651,** 392 (2006).

[67] Q. Yu, Evolution of massive binary black holes, Mon. Not. R. Astron. Soc. **331,** 935 (2002).

[68] F. M. Khan, A. Just, and D. Merritt, Efficient merger of binary supermassive black holes in merging galaxies, Astrophys. J. **732,** 89 (2011).

[69] E. Vasiliev, Rates of capture of stars by supermassive black holes in non-spherical galactic nuclei, Classical Quantum Gravity **31,** 244002 (2014).

[70] E. Vasiliev, F. Antonini, and D. Merritt, The final-parsec problem in nonspherical galaxies revisited, Astrophys. J. **785,** 163 (2014).

[71] E. Vasiliev, F. Antonini, and D. Merritt, The final-parsec problem in the collisionless limit, Astrophys. J. **810,** 49 (2015).

[72] K. Holley-Bockelmann and F. M. Khan, Galaxy rotation and rapid supermassive binary coalescence, Astrophys. J. **810,** 139 (2015).

[73] M. Colpi, Massive binary black holes in galactic nuclei and their path to coalescence, Space Sci. Rev. **183,** 189 (2014).

[74] Y. Kozai, Secular perturbations of asteroids with high inclination and eccentricity, Astron. J. **67,** 591 (1962).

[75] M. L. Lidov, The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies, Planet. Space Sci. **9,** 719 (1962).

[76] M. Bonetti, F. Haardt, A. Sesana, and E. Barausse, Post-Newtonian evolution of massive black hole triplets in galactic nuclei—I. Numerical implementation and tests, Mon. Not. R. Astron. Soc. **461,** 4419 (2016).

[77] M. Bonetti, F. Haardt, A. Sesana, and E. Barausse, Post-Newtonian evolution of massive black hole triplets in galactic nuclei—II. Survey of the parameter space, Mon. Not. R. Astron. Soc. **477,** 3910 (2018).

[78] A. Sesana, A. Vecchio, and C. N. Colacino, The stochastic gravitational-wave background from massive black hole binary systems: Implications for observations with Pulsar Timing Arrays, Mon. Not. R. Astron. Soc. **390,** 192 (2008).

[79] X. Shen, P. F. Hopkins, C.-A. Faucher-Giguère, D. M. Alexander, G. T. Richards, N. P. Ross, and R. C. Hickox, The bolometric quasar luminosity function at $z = 0$–7, Mon. Not. R. Astron. Soc. **495,** 3252 (2020).

[80] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, First higher-multipole model of gravitational waves from spinning and coalescing black-hole binaries, Phys. Rev. Lett. **120,** 161102 (2018).

[81] G. Pratten, P. Schmidt, H. Middleton, and A. Vecchio, Precision tracking of massive black hole spin evolution with LISA, arXiv:2307.13026.

[82] S. Marsat and J. G. Baker, Fourier-domain modulations and delays of gravitational-wave signals, arXiv:1806 .10734.

[83] F. B. Estabrook, M. Tinto, and J. W. Armstrong, Time delay analysis of LISA gravitational wave data: Elimination of spacecraft motion effects, Phys. Rev. D **62,** 042002 (2000).

[84] J. W. Armstrong, F. B. Estabrook, and M. Tinto, Time-delay interferometry for space-based gravitational wave searches, Astrophys. J. **527,** 814 (1999).

[85] S. V. Dhurandhar, K. Rajesh Nayak, and J. Y. Vinet, Algebraic approach to time-delay data analysis for LISA, Phys. Rev. D **65,** 102002 (2002).

[86] M. Tinto and S. V. Dhurandhar, Time delay, Living Rev. Relativity **8,** 4 (2005).

[87] K. Dey, N. Karnesis, A. Toubiana, E. Barausse, N. Korsakova, Q. Baghi, and S. Basak, Effect of data gaps on the detectability and parameter estimation of massive black hole binaries with LISA, Phys. Rev. D **104,** 044035 (2021).

[88] Q. Baghi, I. Thorpe, J. Slutsky, J. Baker, T. Dal Canton, N. Korsakova, and N. Karnesis, Gravitational-wave parameter estimation with gaps in LISA: A Bayesian data augmentation method, Phys. Rev. D **100,** 022003 (2019).

[89] A. Spadaro, R. Buscicchio, D. Vetrugno, A. Klein, D. Gerosa, S. Vitale, R. Dolesi, W. J. Weber, and M. Colpi, Glitch systematics on the observation of massive black-hole binaries with LISA, arXiv:2306.03923.

[90] C. L. Rodriguez, B. Farr, V. Raymond, W. M. Farr, T. B. Littenberg, D. Fazi, and V. Kalogera, Basic parameter estimation of binary neutron star systems by the Advanced LIGO/Virgo Network, Astrophys. J. **784,** 119 (2014).

[91] L. Sampson, N. Cornish, and N. Yunes, Gravitational wave tests of strong field general relativity with binary inspirals: Realistic injections and optimal model selection, Phys. Rev. D **87,** 102001 (2013).

[92] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, Exploring short gamma-ray bursts as gravitational-wave standard sirens, Astrophys. J. **725,** 496 (2010).

[93] LISA Science Study Team, LISA science requirements document, European Space Agency, 2018, https://www.cosmos .esa.int/documents/678316/1700384/SciRD.pdf/25831f6b-3c01-e215-5916-4ac6e4b306fb?t=1526479841000.

[94] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, Science with the space-based interferometer LISA. V. Extreme mass-ratio inspirals, Phys. Rev. D **95,** 103012 (2017).

[95] T. Robson, N. J. Cornish, and C. Liu, The construction and use of LISA sensitivity curves, Classical Quantum Gravity **36,** 105011 (2019).

[96] C. Pitte, Q. Baghi, S. Marsat, M. Besançon, and A. Petiteau, On the detectability of higher harmonics with LISA, Phys. Rev. D **108,** 044053 (2023).

[97] S. Marsat, J. G. Baker, and T. Dal Canton, Exploring the Bayesian parameter estimation of binary black holes with LISA, Phys. Rev. D **103,** 083011 (2021).

[98] W. D. Vousden, W. M. Farr, and I. Mandel, Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations, Mon. Not. R. Astron. Soc. **455,** 1919 (2016).

[99] R. H. Swendsen and J.-S. Wang, Replica Monte Carlo simulation of spin-glasses, Phys. Rev. Lett. **57,** 2607 (1986).

[100] D. J. Earl and M. W. Deem, Parallel tempering: Theory, applications, and new perspectives, Phys. Chem. Chem. Phys. **7,** 3910 (2005).

[101] G. Pratten, A. Klein, C. J. Moore, H. Middleton, N. Steinle, P. Schmidt, and A. Vecchio, LISA science performance in observations of short-lived signals from massive black hole binary coalescences, Phys. Rev. D **107,** 123026 (2023).

[102] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, Science with the space-based

interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens, J. Cosmol. Astropart. Phys. 04 (2016) 002.

[103] C. Cutler, Angular resolution of the LISA gravitational wave detector, Phys. Rev. D **57,** 7089 (1998).

[104] C. M. Hirata, D. E. Holz, and C. Cutler, Reducing the weak lensing noise for the gravitational wave Hubble diagram using the non-Gaussianity of the magnification distribution, Phys. Rev. D **81,** 124046 (2010).

[105] A. Mangiagli, C. Caprini, M. Volonteri, S. Marsat, S. Vergani, N. Tamanini, and H. Inchauspé, Massive black hole binaries in LISA: Multimessenger prospects and electromagnetic counterparts, Phys. Rev. D **106,** 103017 (2022).