



The Enhanced Population of Extreme Mass-ratio Inspirals in the LISA Band from Supermassive Black Hole Binaries

Smadar Naoz^{1,2} and Zoltán Haiman^{3,4}

¹ Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, CA 90095, USA; snaoz@astro.ucla.edu

² Mani L. Bhaumik Institute for Theoretical Physics, Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095, USA

³ Department of Astronomy, Columbia University, 550 W. 120th St., New York, NY 10027, USA

⁴ Department of Physics, Columbia University, 550 W. 120th St., New York, NY 10027, USA

Received 2023 July 13; revised 2023 September 5; accepted 2023 September 11; published 2023 September 26

Abstract

Extreme mass-ratio inspirals (EMRIs) take place when a stellar-mass black hole (BH) merges with a supermassive BH (SMBH). The gravitational-wave emission from such an event is expected to be detectable by the future Laser Interferometer Space Antenna (LISA) and other millihertz detectors. It was recently suggested that the EMRI rate in SMBH binary systems is orders of magnitude higher than the EMRI rate around a single SMBH with the same total mass. Here we show that this high rate can produce thousands of SMBH–BH sources at a redshift of unity. We predict that LISA may detect a few hundred of these EMRIs with signal-to-noise ratio above $S/N \geq 8$ within a 4 yr mission lifetime. The remaining subthreshold sources will contribute to a large confusion noise, which is approximately an order of magnitude above LISA’s sensitivity level. Finally, we suggest that the individually detectable systems, as well as the background noise from the subthreshold EMRIs, can be used to constrain the SMBH binary fraction in the low-redshift Universe.

Unified Astronomy Thesaurus concepts: Supermassive black holes (1663); Gravitational wave sources (677); Gravitational wave detectors (676); Gravitational waves (678); Gravitational wave astronomy (675); Dwarf galaxies (416); E+A galaxies (424); Post-starburst galaxies (2176); Astrodynamics (76)

1. Introduction

The gravitational-wave (GW) emission from a merger of a stellar-mass black hole (BH) and a supermassive BH (SMBH) is one of the science drivers of the Laser Interferometer Space Antenna (LISA), as well as other planned millihertz detectors, such as TianQin (e.g., Amaro-Seoane et al. 2017; Baker et al. 2019; Robson et al. 2019; Mei et al. 2021). Thus, the formation channel and rate of these extreme mass-ratio inspirals (EMRIs) are of prime importance for these future missions.

One of the main EMRI formation channels is based on the weak gravitational interactions between neighboring stellar-mass BHs surrounding SMBHs at the centers of galaxies. These weak two-body interactions are referred to as “two-body relaxation” and describe the so-called “loss-cone” dynamics in dense nuclear star clusters (e.g., Binney & Tremaine 2008; Merritt 2010). In this channel, the encounters in the dense cluster surrounding the SMBH can lead to extreme eccentricity for a stellar-mass BH, driving it toward the SMBH at the center of the cluster (e.g., Hopman & Alexander 2005; Aharon & Perets 2016; Amaro-Seoane 2018; Sari & Fragione 2019). This scenario yields an EMRI rate of a few tens of EMRIs per gigayear per galaxy (e.g., Hopman & Alexander 2005; Aharon & Perets 2016).⁵

Recently, Naoz et al. (2022; hereafter, N22) suggested that EMRI formation in an SMBH binary configuration is an

extremely efficient process. SMBH binaries are a natural consequence of the hierarchical formation of galaxies (e.g., Di Matteo et al. 2005; Hopkins et al. 2006; Robertson et al. 2006; Callegari et al. 2009; Li et al. 2020). This concept is supported by AGN pairs and observations of SMBHs on wide orbits (e.g., Bianchi et al. 2008; Komossa et al. 2008; Comerford et al. 2009; Green et al. 2010; Smith et al. 2010; Caproni et al. 2017; Stemo et al. 2021). These observations imply that those systems may evolve to become tighter (reaching separations on the order of a parsec) binaries. A handful of promising individual candidates on subparsec orbits have been reported in the literature (e.g., Sillanpää et al. 1988; Rodriguez et al. 2006; Bansal et al. 2017; Ren et al. 2021; O’Neill et al. 2022), and samples of several dozen quasars have been identified as subparsec binary candidates based on the periodicities of their light curves in large time-domain surveys (Graham et al. 2015; Charisi et al. 2016; Liu et al. 2019; Chen et al. 2020). Our own Galactic center environment, too, is consistent with hosting a possible less massive intermediate-mass BH (IMBH) companion to Sgr A* (e.g., Hansen & Milosavljević 2003; Maillard et al. 2004; Gürkan & Rasio 2005; Gualandris & Merritt 2009; Fragione et al. 2020; Generozov & Madigan 2020; GRAVITY Collaboration et al. 2020; Naoz et al. 2020; Zheng et al. 2020; Rose et al. 2022; Gravity Collaboration et al. 2023; Zhang et al. 2023).

In the N22 scenario, the SMBH undergoing a high EMRI rate is the less massive of the SMBH pair. As was shown in their study, the eccentric Kozai–Lidov (EKL; Kozai 1962; Lidov 1962; Naoz 2016) mechanism, combined with two-body relaxation, is more efficient around the lighter, secondary SMBH than around the more massive primary one. This behavior is a result of general relativistic precession, which strongly suppresses eccentricity excitations around the primary SMBH, but less so around the secondary (e.g., Ford et al. 2000;

⁵ Note that collisions between stellar-mass BHs and the surrounding stars can lead to a build-up of the stellar BH’s mass, to reach the IMBH regime (Rose et al. 2022).



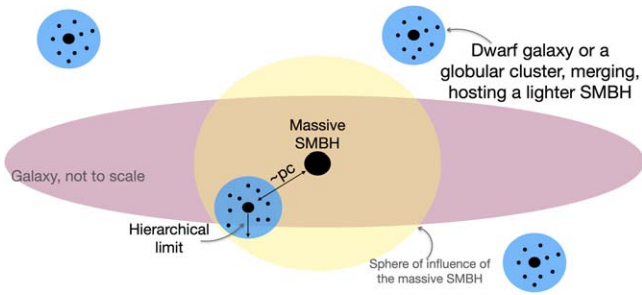


Figure 1. An illustration of the system. The hierarchical nature of galaxy formation suggests that a galaxy will undergo multiple mergers with smaller structures, such as dwarf galaxies or globular clusters. If these smaller objects host SMBHs (or IMBHs) at their centers, the central galaxy’s SMBH can have several lower-mass SMBH (or IMBH) companions.

Naoz et al. 2013; Naoz & Fabrycky 2014; Li et al. 2015). Eccentricity excitations still take place around the secondary, but less efficiently and typically via scattering events (e.g., Chen et al. 2009, 2011; Ivanova et al. 2010; Bode & Wegg 2014). Specifically, N22 demonstrated that the combined channels of EKL and two-body relaxation are vital to producing an efficient rate. They showed that in this mechanism, all available stellar-mass BHs around the lighter SMBH are driven to high eccentricity, yielding a GW source in this process.

This channel has similar implications for the rate of tidal disruption events (TDEs). In particular, it has previously been suggested that SMBH binaries can result in a fast, temporary (a few megayears) increase of the TDE rate (e.g., Chen et al. 2009, 2011; Ivanova et al. 2010; Wegg & Bode 2011; Bode & Wegg 2014; Li et al. 2015; Fragione & Sari 2018; Li et al. 2019). However, most of the investigations in the literature have focused on disrupting stars on the primary (the central, more massive SMBH) due to perturbations from the secondary (the less massive one). In this scenario, EKL eccentricity excitations are suppressed due to general relativity precession (e.g., Naoz et al. 2013) and the majority of TDEs take place due to chaotic orbital crossings (e.g., Chen et al. 2009, 2011; Wegg & Bode 2011). Focusing on TDEs on the secondary (the less massive one) allows for efficient EKL excitations to produce TDEs (Li et al. 2015), which can be used to detect the lighter SMBH in the binary pair (Mockler et al. 2023). Including two-body relaxation to the EKL yields an efficient process similar to the EMRI case. Specifically, it has been shown by Melchor et al. (2023) that it increases the number of stars that undergo TDEs, over a longer period of time. As a result, the TDE rate is consistent with the observed rate in poststarburst (PSB) (E+A) galaxies (see Figure 9, left panel, in Melchor et al. 2023). This scenario also relies on having an unequal-mass SMBH binary configuration.

Interestingly, the hierarchical nature of galaxy formation suggests that a massive SMBH may have a high probability of hosting lighter companions. Since mergers with dwarfs or smaller galaxies are more probable for a given galaxy (e.g., Volonteri et al. 2003a; Rodriguez-Gomez et al. 2015; Pillepich et al. 2018; Paudel et al. 2018; Mićić et al. 2023), an SMBH in the center of a galaxy may have multiple smaller SMBH companions throughout the lifetime of the galaxy (either one at a time or a few at a time; e.g., Madau & Rees 2001; Volonteri et al. 2003b; Rashkov & Madau 2014; Ricarte et al. 2016). See Figure 1 for an illustration of this arrangement. This configuration was adopted in the aforementioned study by

N22, who estimated the EMRI rate to be on the order of $\sim 10^3$ to $\sim 10^5$ EMRIs per primary SMBH (for SMBH masses from $\sim 10^6$ to $\sim 10^8 M_\odot$). It had been suggested that lower-mass dwarf galaxies take longer to merge with a large galaxy and suffer tidal stripping, which reduces their mass and increases the dynamical friction infall timescale. If so, these dwarfs’ low-mass SMBHs will wander away from the large galaxies’ centers (e.g., Madau & Rees 2001; Volonteri et al. 2003a). However, high-resolution simulations, including gaseous dissipation, show that if the low-mass companion galaxy is gas-rich, its SMBHs may sink all the way to the central parts of the galaxy, also holding on to a relic of stars within their own sphere of influence (e.g., Kazantzidis et al. 2005; Callegari et al. 2009; Rashkov & Madau 2014; Lacroix & Silk 2018; Boldrini et al. 2020; Greene et al. 2020; Chu et al. 2023). In these cases, the EMRI rate should be significantly enhanced.

Given this high rate, in this Letter, we ask: what are the consequences of a high event rate for future GW detectors such as LISA? In particular, we obtain a boost in the EMRI rate accounting for binaries with different masses, and estimate the possible confusion noise for LISA at the predicted high event rate. While LISA can resolve individual EMRIs, the aforementioned high rate suggests that the unresolved systems can accumulate incoherently to an unresolved confusion noise or a stochastic background. The possible stochastic background from EMRIs has been estimated, for the single-SMBH case, to produce a low stochastic background, and in some optimistic cases, perhaps a more significant effect is expected (e.g., Barack & Cutler 2004a; Bonetti & Sesana 2020; Pozzoli et al. 2023). We show and quantify below that the unresolved stochastic background expected from EMRIs in SMBH binaries may be significantly above LISA’s sensitivity.

This paper is organized as follows: we begin by using the SMBH mass function in order to estimate the cosmological EMRI rate (Section 2). The noise level within LISA is then estimated in Section 3. Last, we discuss the astrophysical implications in Section 4 and offer our conclusions in Section 5.

2. Cosmological Rate Estimation

N22 estimated the EMRI rate per SMBH as a function of the (secondary) SMBH mass M_* . Here, we present a fit to their result, $\gamma(M_*)$, as

$$\gamma(M_*) \sim 925.4 \text{ Gyr}^{-1} \left(\frac{M_*}{10^6 M_\odot} \right)^{0.76} \times f_{\text{LISA}} \times f_{\text{EKL}} \times \eta_{\text{bin}}, \quad (1)$$

where the value is based on the M - σ relation and a fraction of stellar BHs, of 3.2×10^{-3} (e.g., Kroupa 2001), which is equal to the total mass of stellar-mass BHs divided by the total mass in stars, within the sphere of influence of an SMBH. The total mass enclosed within the sphere of influence of an SMBH is twice the SMBH mass (e.g., Tremaine et al. 2002). In the scenario presented here, each falling SMBH carries its nuclear star cluster with it. The radius of the secondary’s nuclear star cluster shrinks, as the primary’s gravitational potential begins to overtake the secondary’s. This radius, which is often referred to as the “hierarchical limit,” roughly describes the region where the EKL equation of motions can be applied (e.g., Lithwick & Naoz 2011; Naoz 2016; Zhang et al. 2023). In our

case, this is on the order of a parsec, which is of the order of the sphere of influence. The number of enclosed stellar-mass BHs can be approximated by adopting the M – σ relation for these smaller SMBHs’ nuclear star clusters. Adopting the value 3.2×10^{-3} for a fraction of stellar-mass BHs means that we neglect the effects of mass segregation close to the SMBH. Mass segregation may increase the number of stellar-mass BHs, (e.g., Hopman & Alexander 2006), thus yielding an even larger boost to the EMRI rate than the one described below.

We also introduce the following parameters: f_{LISA} is the fraction of stellar-mass BHs seen to merge with SMBHs in the LISA band as EMRIs, rather than as plunges. N22 estimated this fraction to be between 40% and 100% (see their Appendix A). f_{EKL} is the fraction of stellar-mass BHs in the SMBHs’ spheres of influence that are driven to merge with the SMBH when both EKL and two-body relaxation operate. N22 estimated $f_{\text{EKL}} \sim 0.5$ –1. Finally, η_{bin} is the fraction of SMBHs that acquire a more massive SMBH companion. While this fraction is at most unity, below, we allow η_{bin} to be larger than unity due to uncertainties in the SMBH mass function, especially at the low-mass end, where it is extrapolated through population modeling, as well as additional uncertainties that are degenerate with this number and may therefore be captured by this single parameter (see Section 4).

Equation (1) is derived from Figure 5 in N22, which represents a scaling from their fiducial system. It adopts a fixed EKL timescale and rescales the binary separation accordingly. Mockler et al. (2023) and Melchor et al. (2023) have recently showed that keeping the SMBH binary separation constant yields consistent results. In fact, Melchor et al. (2023), considering TDEs via the combined EKL and two-body relaxation mechanisms, showed that the number of stars (proportional to the primary mass) is the main factor determining the TDE rate. The number of available stars in the hierarchical limit is directly related to the SMBH binary eccentricity, with the lowest eccentricity yielding a higher number of particles and, thus, a higher rate. Since Equation (1) is derived from N22, who adopted $e = 0.7$ for their fiducial system, it represents an EMRI rate that must be lower compared to yet smaller SMBH binary eccentricity values. In other words, this implies that the lower limit of f_{EKL} may be higher than estimated in N22. Furthermore, Melchor et al. (2023) considered two mass ratios, $M_2/M_1 = 10$ and 100, demonstrating that they have little effect on the TDE rate, as long as the disruption is taking place on the less massive SMBH.

Furthermore, N22 and Melchor et al. (2023) demonstrated that the SMBH companion drives the stellar-mass BHs (or stars) toward the primary via this channel on timescales that can be as short as 10^5 yr, up to 10^8 yr. The hardening of an SMBH binary may take between millions and billions of years (e.g., Begelman et al. 1980; Milosavljević & Merritt 2001; Yu 2002; Sesana et al. 2006; Haiman et al. 2009; Kelley et al. 2017). Note that the hardening process may assist EMRI formation and still achieve a high rate (Mazzolari et al. 2022).

As highlighted above, there are many uncertainties involved in this process. For example, the SMBH binary mass ratio, the separation and eccentricity, the hardening rate, and the probability of the SMBH having a companion(s) in the first place. For simplicity, here we fold all of these uncertainties into the parameter η_{bin} . We offer possible constraints on this parameter in the discussion below.

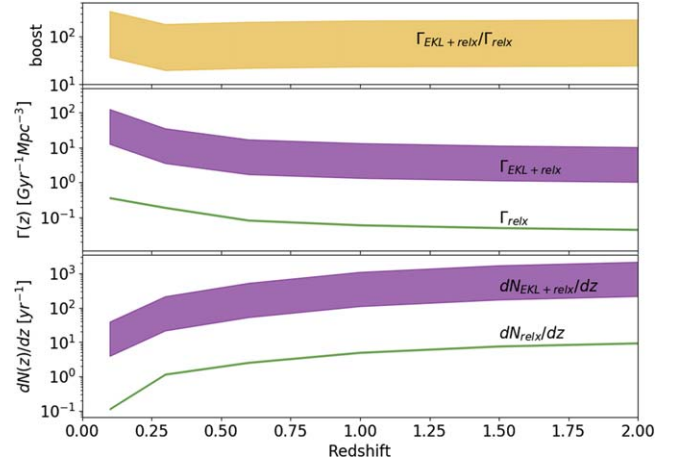


Figure 2. Bottom panel: the expected number of EMRIs per unit redshift per year. We consider the SMBH binary scenario (wide purple band), where the upper limit assumes $\eta_{\text{bin}} = 1$, $f_{\text{LISA}} = f_{\text{EKL}} = 1$ and the lower limit assumes $\eta_{\text{bin}} = 0.5$, $f_{\text{LISA}} = 0.4$ and $f_{\text{EKL}} = 0.5$ (the latter two choices being motivated by N22). The thin green line shows the corresponding rate for single SMBHs via the two-body relaxation channel. Middle panel: the volumetric rate of EMRIs as a function of redshift (see Equation (2)). The bottom panel is obtained by integrating this rate over volume and allowing one year of LISA operation (Equation (3)). Top panel: the enhancement of the EMRI rate in the binary scenario relative to the single-SMBH case. The numbers in this figure do not take into account LISA’s sensitivity (see Section 3).

Merloni & Heinz (2008) estimated the SMBH mass function by solving the continuity equation for its evolution and using the locally determined mass function as a boundary condition (see also Sicilia et al. 2022). Similar estimates were found by Aversa et al. (2015), who calibrated their results to so-called abundance matching. In the Appendix, we show the fit we used for the SMBH mass function. Note that this mass function assumes that the mass of each nucleus is dominated by the more massive (central) SMBH; however, in our scenario, each nucleus may hold more than one SMBH, thus this mass function undercounts small SMBHs. Given the overall uncertainty in the mass function (e.g., Merloni & Heinz 2008; Aversa et al. 2015; Sicilia et al. 2022), we do not expect this multiplicity to alter our conclusions significantly. Given the number density $d^2N(z)/d \log(M_*) dV$ (i.e., the number of SMBHs per comoving volume and mass, as a function of redshift z and mass M_*), we can estimate the EMRI rate density at each redshift as

$$\Gamma(z) = \int \frac{d^2N}{dV d \ln M} \gamma(M) d \ln M, \quad (2)$$

Thus, the number of EMRIs per unit z expected over LISA’s lifetime T_{obs} is

$$\frac{dN}{dz}(z) = T_{\text{obs}} \times \int \frac{\Gamma(z)}{1+z} \frac{dV}{dz}, \quad (3)$$

where dV/dz is the comoving volume per unit redshift, and the factor $(1+z)^{-1}$ accounts for cosmological time dilation.

In the bottom panel of Figure 2, we show an example of the number of EMRIs per unit redshift, expected after $T_{\text{obs}} = 1$ yr of LISA operation. The wide band shows the result of the combined effect of EKL with two-body relaxation around SMBH binaries, following N22, where the upper edge of each band assumes $\eta_{\text{bin}} = 1$ and $f_{\text{LISA}} = f_{\text{EKL}} = 1$. The lower edge of the band instead adopts $\eta_{\text{bin}} = 0.5$, $f_{\text{LISA}} = 0.4$ and $f_{\text{EKL}} = 0.5$, the latter two consistent with the N22 findings. The thin band in the figure

shows the estimated number of EMRIs from single SMBHs based on two-body relaxation, calculated from Equation (3). The EMRI rate per single SMBH (i.e., γ_{relx}) is estimated following Hopman & Alexander (2005); see Equation (15) in N22. The top panel of Figure 2 shows the boost in the EMRIs rate in the binary scenario relative to the single-SMBH case.

As depicted in Figure 2, even at the lower edge of the uncertain ranges shown, the combined EKL + two-body relaxation boosts the EMRI rate by more than an order of magnitude. Further, as expected, higher redshifts allow for a larger volume to be sampled and yield more EMRIs per unit z .

3. LISA Observability

3.1. Characteristic GW Strain

Eccentric EMRIs emit GWs over a wide range of orbital frequency harmonics. The dimensionless GW strain $h(a, e, t)$ of an EMRI with a semimajor axis a and eccentricity e is the sum of the strains at each harmonic $f_n = n f_{\text{orb}}$, where $f_{\text{orb}} = \sqrt{GM_*/a^3}$ (Peters & Mathews 1963). In other words, the strain is

$$h(a, e, t) = \sum_{n=1}^{\infty} h_n(a, e, f_n) e^{i2\pi f_n t}, \quad (4)$$

where

$$h_n(a, e, f_n) = h_0(a) \frac{2}{n} \sqrt{g(n, e)}. \quad (5)$$

Here, $h_0(a)$ is the instantaneous strain amplitude for a circular EMRI with SMBH mass M_* and stellar-mass BH mass m , at a luminosity distance D_l , whose value, averaged over sky location and GW polarization, is given by

$$h_0(a) = \sqrt{\frac{32}{5}} \frac{G^2}{c^4} \frac{M_* m}{D_l a}. \quad (6)$$

Here, G is the gravitational constant, c is the speed of light, and $g(n, e)$ is defined as

$$g(n, e) = \frac{n^4}{32} \left[\left(J_{n-2} - 2eJ_{n-1} + \frac{2}{n}J_n + 2eJ_{n+1} - J_{n+1} \right)^2 + (1 - e^2)(J_{n-1} - 2J_n + J_{n+2})^2 + \frac{4}{3n^2}J_n^2 \right], \quad (7)$$

where J_i is the i th Bessel function evaluated at ne . The characteristic strain of a stationary EMRI $h_c(a, e, f)$ can be approximated using the Fourier transform of $h(a, e, t)$ measured for some time T_{obs} (Kocsis et al. 2012):

$$h_c(a, e, f) = 2\tilde{f}h(a, e, f) \sqrt{\min(T_{\text{obs}}, f_n/\dot{f}_n)}. \quad (8)$$

The minimum term accounts for the (square root of) the number of cycles that can be observed at each frequency, determined by the smaller of the finite observation time and the time for the EMRI to chirp across the frequency band. Last, $\tilde{h}(a, e, f)$ is the Fourier transform of Equation (4) over the observation time T_{obs} ,

$$\tilde{h}(a, e, f) = \int_0^{T_{\text{obs}}} h(a, e, t) e^{-2\pi i f t} dt, \quad (9)$$

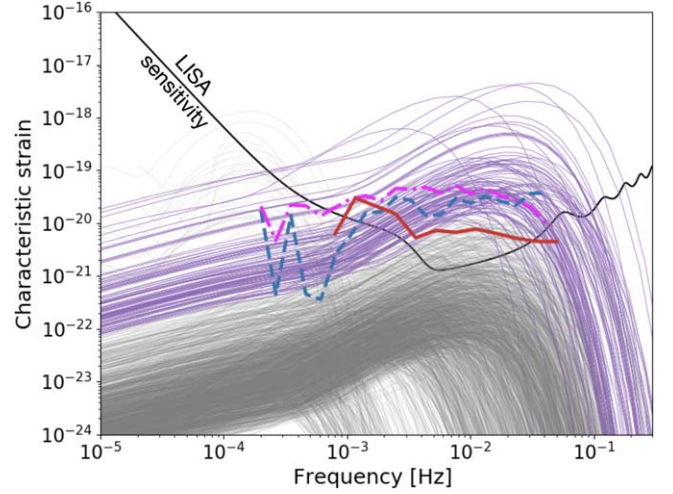


Figure 3. Characteristic GW strain and noise level for LISA. We show characteristic strains in the LISA band as a function of frequency. We show 2000 randomly chosen EMRI systems and adopt $T_{\text{obs}} = 4$ yr. The thin gray lines show the unresolved systems, while the purple lines above the LISA sensitivity curve show the resolved systems with $S/N > 8$. The thick red solid line shows $\sqrt{S(f)} \times f$, as described in the text. The dashed blue line shows $\sqrt{\sum \max(h_c)^2}$ for those 2000 randomly chosen systems, where $\max(h_c)$ is the maximum characteristic strain of each system. The dashed-dotted magenta line shows $\sqrt{\sum (h_c)^2}$, including all the relevant frequencies of the strain for each realization. The black line shows LISA’s sensitivity curve, following Robson et al. (2019).

and has units of Hz^{-1} . Following O’Leary et al. (2009), Kocsis et al. (2012), and Gondán et al. (2018), we assume that there is negligible overlap between the harmonics, and thus the sum of squares of each element dominates the product of the sums in the Fourier transform. Note that a small fraction of systems ($< 1\%$, for a 1 yr observation; see below) are expected to evolve significantly in the LISA band and thus will require a more detailed analysis of their strain amplitude (e.g., Barack & Cutler 2004a), which is beyond the scope of this work. Note that evolving EMRIs tend to have a higher signal-to-noise ratio (S/N); thus, the analysis below represents a conservative estimate for the resolvable sources and the stochastic background.

Figure 3 depicts the envelope of the characteristic strain of 2000 EMRIs for $z \leq 1$. The value of 2000 for the total number of systems is a result of the integral presented in the bottom panel of Figure 2; see also Equation (3). We randomly draw 2000 EMRI systems using the above mass function (see the Appendix). Furthermore, we sample configurations in a and e from the fiducial $10^7 M_\odot$ simulations in N22,⁶ and rescale the results for different primary masses, keeping the periods similar.⁷ The EMRIs shown in Figure 3 are highly eccentric, with an average eccentricity of 0.996 and an average period of ~ 5 yr. The realizations are chosen from the SMBH mass

⁶ The N22 simulations were initialized with a cusp-like Bahcall & Wolf (1976) distribution and a thermal distribution of eccentricities within the hierarchical limit.

⁷ It was noted previously that EKL systems are scalable (e.g., Naoz & Silk 2014; Naoz et al. 2019). In these studies, the EKL timescale is kept constant. However, even when we relax the constant EKL timescale, the main contributor to the rate is the number of particles within the hierarchical limit (Melchor et al. 2023; Mockler et al. 2023). The number of BHs within the hierarchical limit is proportional to the mass of the primary via the M – σ relation (e.g., Tremaine et al. 2002).

function described in the [Appendix](#) and yield an average SMBH mass of $1.6 \times 10^6 M_\odot$. We note that we kept the stellar-mass BH fixed at $10 M_\odot$ throughout this Letter. These wide and eccentric systems result in a long GW merger timescale, with an average value of 2.3×10^4 yr. Thus, only 1% of the systems evolve within 1 yr of observation time.

As is clearly shown, many EMRIs lie below LISA’s sensitivity curve (the thick black line in the figure calculated following Robson et al. 2019). The purple lines (above the LISA sensitivity curve) represent systems with S/N larger than 8, and they represent about ~ 100 (~ 40) detectable sources, after $T_{\text{obs}} = 4$ yr ($T_{\text{obs}} = 1$ yr) from the first year of observations (recall that increasing T_{obs} for a stationary source increases its S/N).⁸ After 4 yr of observations, we expect a few hundred sources to become detectable.

The S/N is defined (e.g., Kocsis et al. 2012; Robson et al. 2019) as

$$\begin{aligned} \langle (S/N)^2(a, e) \rangle &= \frac{16}{5} \int \frac{|\tilde{h}(a, e, f)|^2}{S_n(f)} df \\ &= \frac{16}{5} \int \frac{|\tilde{h}(a, e, f)|^2}{f S_n(f)} d \ln f. \end{aligned} \quad (10)$$

We can thus approximate the S/N for the eccentric source at the peak frequency as (e.g., Kocsis et al. 2012, Z. Xuan et al. 2023, in preparation)

$$S/N \sim \frac{\max(h_c)}{\sqrt{S_n(f_{\text{peak}}) f_{\text{peak}}}} \quad \text{for } P_{\text{orb}} < T_{\text{obs}}, \quad (11)$$

where $S_n(f_{\text{peak}})$ is LISA’s sensitivity threshold. In other words, among the expected 2000 EMRIs from $z \sim 1$ (see Figure 2, bottom panel), we expect 2% to be visible within LISA with $S/N > 8$, per year, and the rest to contribute to confusion noise (depicted as thin gray lines).

Note that in some cases, we may resolve specific harmonics (with a similar treatment as described in Barack & Cutler 2004b; Seto 2012; Katz et al. 2021; Xuan et al. 2023). This detection of higher individual harmonics effectively can increase the S/N and potentially break the degeneracies between inferred parameters. Furthermore, the actual trajectory of the stellar-mass BH and its characteristic strain should be modified by Kerr geometry and special relativity in the case of spinning SMBHs and wide systems, respectively, (e.g., Yunes et al. 2008; Berry & Gair 2013; Johannsen 2013; Schnittman 2015, 2018; Chua et al. 2017, 2021).

3.2. Stochastic Background from Unresolved EMRIs

The population of EMRIs that do not individually rise to detectable levels constitutes an unresolved GW background, which, in turn, can represent confusion noise for LISA. We estimate the level of this confusion background “noise,” $\sqrt{S(f)}f$, using two different methods. The first of these is to generate a single Monte Carlo realization of the sky with 2000 sources, using the results in the previous section and following the approach of computing the corresponding GW background from a catalog of discrete sources by, e.g., Kocsis & Sesana (2011), Nissanke et al. (2012), and Dvorkin & Barausse (2017). The second approach is to directly compute the expectation

value of the local GW background using smooth functions describing the population of sources, following the approach by Phinney (2001) generalized to eccentric binaries, as in Enoki & Nagashima (2007) and Toonen et al. (2009).

In the first method, we use the characteristic strain described above and estimate the noise as

$$\sqrt{S(f)}f \sim \sqrt{\sum_i \max[h_c(f)]^2}, \quad (12)$$

where $\max[h_c(f)]$ is the maximum characteristic strain (from Equation (8)) for each EMRI. Equation (12) is motivated by the S/N definition, Equations (10) and (11). The summation is done in each observed frequency bin. This result is depicted in Figure 3 as the thick (blue) dashed line. For reference, the magenta dotted–dashed line curve in this figure shows the same calculation, except including the full frequency spectrum of each source, rather than just the strain at the peak frequency. Both cases include only sources with $S/N < 8$ and periods longer than 1 yr.

In the second method, we follow Barack & Cutler (2004a), where the one-sided spectral density $S(f)$ in the LISA band can be estimated as

$$S(f) = \frac{4G}{\pi c^2 f^2} \frac{d\rho}{df}, \quad (13)$$

where f is the observed frequency and $(d\rho/df)\Delta f$ is the total GW energy density observed today with frequencies of $(f, f + df)$, which can be estimated as

$$\begin{aligned} \frac{d\rho}{df} \Delta f &= \int_0^{z_{\text{max}}} \Delta f_{\text{em}} \frac{dt}{dz} \frac{\dot{\mathcal{E}}(f_{\text{em}})}{1+z} dz \\ &= \int_0^{z_{\text{max}}} \Delta f \frac{dt}{dz} \dot{\mathcal{E}}(f(1+z)) dz. \end{aligned} \quad (14)$$

Here, $\dot{\mathcal{E}}(f_{\text{em}})$ is the emitted spectrum of the GW sources (i.e., energy per unit comoving volume, unit proper time, and unit emitted frequency), and in the last step we used $f_{\text{em}} = f(1+z)$. The emitted spectrum can be approximated as

$$\dot{\mathcal{E}}(f_{\text{em}}) \sim \int \frac{d^2 N}{dV d \ln M} \gamma(M) \Delta E_{\text{GW}}(M, f_{\text{em}}) d \ln M, \quad (15)$$

where $\Delta E_{\text{GW}}(M, f_{\text{em}})$ is the emitted GW energy as a function of mass in a frequency bin $(f_{\text{em}}, f_{\text{em}} + df_{\text{em}})$. We note that the $z > 1$ events contribute negligibly to the total LISA signal. For each SMBH primary and a stellar-mass BH of $m = 10 M_\odot$, with a given semimajor axis a and eccentricity e , the energy loss is (Peters & Mathews 1963; Peters 1964)

$$\frac{dE}{dt} = \frac{32 G^4 M_*^2 m^2 (M_* + m)}{5 a^5 c^5} f(e), \quad (16)$$

where $f(e)$ is the summation of all the harmonics, which gives

$$f(e) = \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right) (1 - e^2)^{-7/2}. \quad (17)$$

We use the fiducial system of a 10^7 – $10^9 M_\odot$ SMBH binary in N22. We rescale the EMRI’s semimajor axis for each mass based on this fiducial system. For simplicity, we assume that most of the energy is emitted at the peak frequency

⁸ We note that an S/N of 8 is lower by a factor of a few than some of the commonly adopted S/Ns. However, these numbers are uncertain, since LISA is yet to be launched.

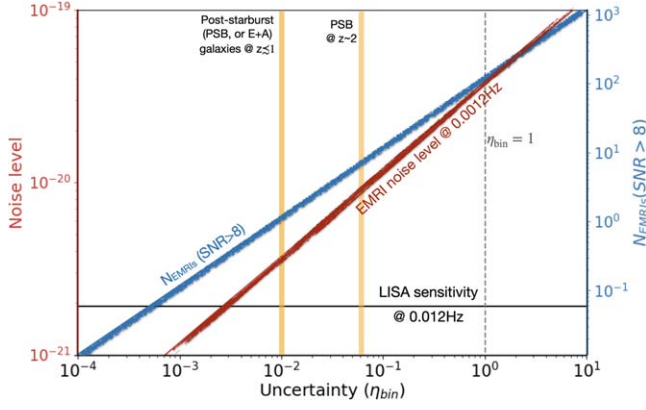


Figure 4. The astrophysical consequences of binary EMRIs. The x-axis shows the SMBH binary fraction, η_{bin} , which incorporates all of the uncertainties mentioned in this work. This term can be larger than unity if the SMBHs in galactic nuclei typically have multiple companion SMBHs/IMBHs. The left y-axis (red) shows the EMRI noise level at ~ 0.012 Hz estimated from the noise calculation described in Section 3.2 and shown in Figure 3, as a function of η_{bin} . LISA’s sensitivity at ~ 0.012 Hz (black horizontal line) is overplotted. This frequency corresponds to the largest difference between the EMRI noise level and the LISA sensitivity curve. The right y-axis (blue) shows the estimated number of EMRIs with $S/N > 8$, and $T_{\text{obs}} = 4$ yr, as a function of η_{bin} . The thick vertical yellow lines mark the approximate fractions of PSB galaxies in the local Universe ($z \lesssim 1$) and at $z \sim 2$.

$f_{\text{peak}} = f_{\text{orb}} \sqrt{(1+e)/(1-e)^3}$, where $f_{\text{orb}} = \sqrt{GM/a^3}$. Note that each $dE(a, e)/dt$ has its characteristic f_{peak} .

In order to find $\Delta E_{\text{GW}}(M, f_{\text{em}})$, we draw 4000 random EMRI configurations (i.e., a set of a and e) from the fiducial simulation and rescale for each M . Then, we find the fraction of the cases in which a dE/dt value appeared in the frequency bin $f_{\text{em}} = f_{\text{peak}}(a, e)$. This process, thus, roughly estimates the energy loss probability of an EMRI for a given primary mass in a frequency bin. We then estimate $\Delta E_{\text{GW}}(M, f_{\text{em}})$ for a typical EMRI, over its lifetime, at mass M , as

$$\Delta E_{\text{GW}}(M, f_{\text{em}}) \sim \frac{1}{N_{\text{real}} \sum_{\text{real}}} \left(\frac{df_{\text{em, peak}}}{dt} \right)^{-1} \frac{dE(a, e, f_{\text{em, peak}})}{dt}, \quad (18)$$

where the summation represents averaging over the (a, e) distribution at fixed mass and frequency, and we assume each EMRI, drawn from the N22 distribution in (a, e) , subsequently evolves purely due to GW emission, i.e., $df_{\text{em, peak}}/dt$ is given by (shown here for completion)

$$\frac{df_{\text{em, peak}}}{dt} = \frac{\partial f_{\text{em, peak}}}{\partial a} \frac{da}{dt} + \frac{\partial f_{\text{em, peak}}}{\partial e} \frac{de}{dt}, \quad (19)$$

where $f_{\text{em, peak}}$ is defined above and da/dt and de/dt are taken from Peters & Mathews (1963) and Peters (1964).

Figure 3 depicts the confusion noise estimated following the above prescriptions. Specifically, the thick red line shows $\sqrt{S(f)f}$. This figure suggests that in an idealized case, in which all galaxies host one SMBH binary, the noise level due to unresolved EMRIs covers most of the LISA band between (3 and 30) mHz at the few $\times 10^{-20}$ level. We note the consistency between the two methods used here to calculate the noise level. The first method (i.e., $\sqrt{\sum \max(h_c)^2}$) is cruder, but appears to be consistent with the analytical calculation (thick red line).

We suggest that the spectrum of this stochastic EMRI background noise may be “jagged” due to the finite number of

EMRIs within the observable volume. As shown in Figure 2, $\mathcal{O}(1000)$ EMRI systems may exist in nature within $z \sim 1$, each contributing primarily to different frequency bins. Thus, while in principle we could sample from a larger simulated sample and create a smoother spectrum, this may be misleading. The precise shape of the jagged noise spectrum is uncertain, as it varies from one set of realizations to another.

4. Discussion and Astrophysical Implications

The hierarchical nature of galaxy formation implies that SMBH binaries or even higher multiples should be the common outcome of galaxy mergers (see Figure 1 for an illustration). Furthermore, since galaxies’ minor mergers are common, a large SMBH mass ratio in SMBH binaries may be a natural consequence (e.g., Rashkov & Madau 2014; Ricarte et al. 2016). It has recently been shown that these systems can yield a high EMRI rate per galaxy (N22). Here we have demonstrated that this high rate implies that thousands of SMBH–stellar-mass BH GW sources may exist within the local Universe (redshift of unity); see Figure 2. Among these EMRIs, a few hundred may be detectable with $S/N > 8$ by LISA during its lifetime, and the rest, unresolved ones, accumulate to a high stochastic background noise level. As depicted in Figure 3, the level of this background can be an order of magnitude above LISA’s sensitivity limit.

We note that the stochastic background from EMRIs described above is analogous to the background from Galactic dwarf binaries. This white-dwarf background has been estimated (assuming circular binary inspirals) and is known to reduce LISA’s sensitivity at ~ 2 mHz (e.g., Nissanke et al. 2012; Robson et al. 2019). The reduced sensitivity manifests as a “bump” in the LISA sensitivity curve, increasing it by a factor of a few. Here we find that the EMRI background may result in a similar, but potentially even higher, “bump” at ~ 3 –30 mHz. However, unlike the nearly 10 billion Galactic white dwarfs that produce a smooth feature, here we find that the EMRI background may result in a noisy, jagged curve. This jaggedness is a result of the combination of the discrete nature of EMRIs (of which there are many fewer than white dwarfs) and their nonzero eccentricities, which spread them over a wider range of frequencies compared to circular inspirals, causing even fewer sources to contribute per frequency bin.

The background noise presented here is much higher than previous, single-SMBH estimates. Specifically, previous analyses of the EMRI background have found that its expected level is very close, or slightly above, LISA’s sensitivity curve, which includes the unresolved Galactic white-dwarf binaries (Barack & Cutler 2004a; Pozzoli et al. 2023). The optimistic model of Bonetti & Sesana (2020) is slightly below the conservative background estimates presented here. Further, their optimistic background noise estimate drops below the LISA sensitivity curve at around 0.01 Hz, while the unresolved background presented here is roughly constant.

We capture the uncertainties in the above estimation with the parameter η_{bin} . This parameter describes the average number of light companions around a central SMBH with overlapping spheres of influence, thus, in principle, η_{bin} can be larger than unity. Additionally, η_{bin} may be degenerate with other uncertainties, including (but not limited to) the fraction of stellar-mass BHs in nuclear star clusters, the fraction of plunges versus slowly inspiraling EMRIs (denoted as f_{LISA} , above), the fraction of stellar-mass BHs that will be driven toward the

primary in the aforementioned channel (denoted as f_{EKL} , above), and the hardening time of SMBH binaries (although an SMBH undergoing a merger has also been suggested to yield a high EMRI rate; e.g., Mazzolari et al. 2022). However, unlike the SMBH binary fraction, f_{LISA} and f_{EKL} may be constrained using the N22 simulations. In particular, the latter is an estimate of the fraction of particles that crossed the hierarchical limit, and the former is derived from estimating the fraction with periods shorter than ~ 10 yr; see Appendix B in N22. Notably, even at the low end of their expected range, these parameters still significantly enhance the EMRI rate (by over an order of magnitude; see the top panel in Figure 2).

We suggest that LISA observations can be used to constrain the astrophysical nature of SMBH binaries—particularly the average number of companions of SMBHs. This is illustrated in Figure 4, where the x -axis is η_{bin} . The left y -axis (red) depicts the noise level from the unresolved EMRIs at ~ 0.012 Hz, shown as the thick red line (labeled). This frequency corresponds to the maximum noise compared to LISA’s sensitivity threshold (depicted in the figure as a black horizontal line at ~ 0.012 Hz). Equation (13) combined with Equation (1) suggests that the noise level $\sqrt{S(f)} \times f \sim \sqrt{\eta_{\text{bin}}}$. Thus, measuring this stochastic background noise level with LISA can be used to constrain η_{bin} . Similarly, the number of LISA-detected EMRIs can be used to constrain the incidence rate of SMBH binaries. This is illustrated by the right y -axis of Figure 4, where we show the number of systems with $S/N > 8$ detectable by LISA per year. This is estimated from the calculations presented in Figure 3, where $\sim 6\%$ of the systems from $z \lesssim 1$ have $S/N > 8$. As illustrated in Figure 4, the combination of the number of individually detected EMRIs and the level of the unresolved background should yield strong constraints on the SMBH binary population.

We highlight that the EMRI rate predicted from SMBH binaries is much higher than the EMRI rate based on single SMBHs (see Figure 2). Thus, for non-negligible values of η_{bin} , the single-SMBH EMRI population has a minor contribution to the background noise. Disentangling between the binary and single channels for a resolved EMRI may be possible if the EMRI can be localized and the mass of the EMRI SMBH is found to deviate from (i.e., fall well below) the observed or inferred central SMBH mass. Such a possibility may indicate that the stellar-mass BH merged with a smaller, hidden SMBH. Another possibility to differentiate the single and binary channels may be through their respective electromagnetic signatures. For example, a concurrent detection of an EMRI and a TDE (e.g., Melchor et al. 2023) may help constrain the mass of the disrupting SMBH, adding an independent measure of the mass.

Notably, taken at face value, if every SMBH at the center of a galaxy hosts at least one smaller SMBH companion, the same channel that produces a high EMRI rate will result in a high TDE rate, perhaps even higher than the observed TDE rate per galaxy (estimated as 10^{-5} – 10^{-4} , per year, per galaxy; e.g., French et al. 2020). It has been shown recently that the combined EKL + two-body relaxation indeed results in a high TDE rate (Melchor et al. 2023). Furthermore, it can be used to detect the lighter SMBH (Mockler et al. 2023). Interestingly, the TDE rate predicted from the combined EKL + two-body relaxation channel is consistent with the TDE rate observed in PSB galaxies (Melchor et al. 2023). Remarkably, the observed TDE rate in PSB galaxies is higher than the averaged observed

TDE rate (e.g., van Velzen et al. 2020). Thus, these results suggest that SMBH binaries near the sphere of influence of each other may be a typical outcome of PSB galaxies. This inference appears to be consistent with the recent NANOGrav result, suggesting that high-mass SMBH binary mergers take place within the first 1 Gyr of galaxy mergers (e.g., Antoniadis et al. 2023).

Thus, in Figure 4, we overplot the estimated PSB galaxy fraction at $z \lesssim 1$ and $z \sim 2$ (Bergvall et al. 2016; Wild et al. 2016; Belli et al. 2019; D’Eugenio et al. 2020; French 2021). Thus, if SMBH binaries exist only in PSB galaxies,⁹ then the PSB galaxy fraction may indicate the fraction of SMBHs with companions, i.e., η_{bin} . The consistency of the TDE rate with the EKL + two-body relaxation estimate (Melchor et al. 2023) may hint at such a scenario. As shown by this figure, limiting the stochastic background noise from binaries to originate only from PSB galaxies will still create a significant level of noise confusion in LISA in its most sensitive frequency band. In other words, a noise level $2\text{--}8 \times 10^{-20}$ (for $z \sim 1\text{--}2$) may imply that the origin of EMRIs is consistent with PSB galaxies.

5. Conclusions

In this paper, we have demonstrated that the high EMRI rate predicted in SMBH binaries (N22) can result in thousands of GW sources. A considerable fraction of these sources may be detected individually by LISA (estimated as a few hundred sources for a 4 yr LISA mission lifetime). The accumulated, unresolved sources, in turn, may accumulate to a significant unresolved stochastic background, reaching an amplitude that is an order of magnitude above LISA’s sensitivity threshold between $\sim (3 \text{ and } 30)$ mHz. Finally, we suggest that in a simple interpretation, LISA observations of individually detectable EMRIs and EMRI noise levels can be used in combination to constrain the prevalence of SMBH binaries in galactic nuclei.

Acknowledgments

We thank Zeyua Xuan, Brenna Mockler, and Suvi Gezari for useful discussions. We also thank the anonymous referee for their useful comments. S.N. acknowledges the partial support from NASA ATP 80NSSC20K0505 and from NSF-AST 2206428 and also thanks Howard and Astrid Preston for their generous support.

Appendix The Mass Function

The SMBH mass function is adopted from Merloni & Heinz (2008). We develop a simple fitting formula for their mass function as a function of redshift and BH mass. We find that the SMBH mass function can be estimated by

$$\log_{10} \left(\frac{dN}{dV d \log M} (z) \right) = a_1(z) (\log_{10} M)^4 + a_2(z) (\log_{10} M)^3 + a_3(z) (\log_{10} M)^2 + a_4(z) \log_{10} M + a_5(z), \quad (\text{A1})$$

⁹ In this case, the rest of the population may have quickly merged or stayed in the galaxy’s bulge. Alternatively, the light SMBH did not hold to its nucleus as it sank in.

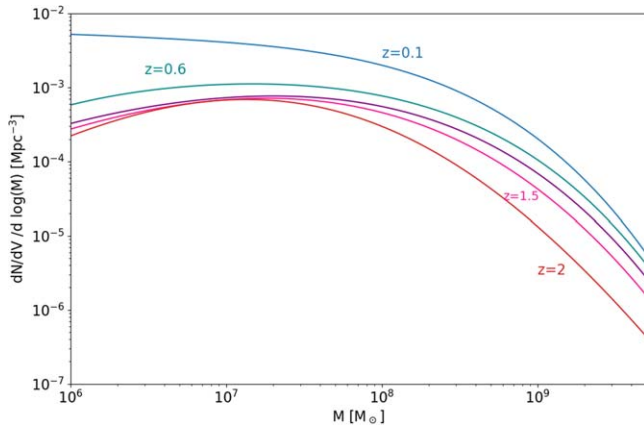


Figure 5. Examples of the mass function fits used in these calculations. These are fits for Merloni & Heinz (2008).

where $a_i(z)$ is fitted to be

$$a_i(z) = p_{i1}z^4 + p_{i2}z^3 + p_{i3}z^2 + p_{i4}z + p_{i5}, \quad (\text{A2})$$

and

$$p_{ij} = \begin{pmatrix} 0.023365 & -0.1031 & 0.158098 & -0.080575 & -0.009329 \\ -0.747931 & 3.335781 & -5.154087 & 2.683332 & 0.176503 \\ 8.8832 & -40.0616 & 62.3869 & -33.2041 & -0.9996 \\ -46.4664 & 211.9867 & 333.0359 & 181.6464 & 0.622461 \\ 90.3990 & -417.6483 & 663.5951 & -373.2478 & 4.26139 \end{pmatrix}. \quad (\text{A3})$$

The results from this fit are depicted in Figure 5.

ORCID iDs

Smadar Naoz <https://orcid.org/0000-0002-9802-9279>
Zoltán Haiman <https://orcid.org/0000-0003-3633-5403>

References

- Aharon, D., & Perets, H. B. 2016, *ApJL*, **830**, L1
Amaro-Seoane, P. 2018, *LRR*, **21**, 4
Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv:1702.00786
Antoniadis, J., Arumugam, P., Arumugam, S., et al. 2023, arXiv:2306.16227
Aversa, R., Lapi, A., de Zotti, G., Shankar, F., & Danese, L. 2015, *ApJ*, **810**, 74
Bahcall, J. N., & Wolf, R. A. 1976, *ApJ*, **209**, 214
Baker, J., Bellovary, J., Bender, P. L., et al. 2019, arXiv:1907.06482
Bansal, K., Taylor, G. B., Peck, A. B., Zavala, R. T., & Romani, R. W. 2017, *ApJ*, **843**, 14
Barack, L., & Cutler, C. 2004a, *PhRvD*, **70**, 122002
Barack, L., & Cutler, C. 2004b, *PhRvD*, **69**, 082005
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, **287**, 307
Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, **874**, 17
Bergvall, Nils, Marquart, Thomas, & Michael, Way 2016, *A&A*, **587**, A72
Berry, C. P. L., & Gair, J. R. 2013, *MNRAS*, **433**, 3572
Bianchi, S., Chiaberge, M., Piconcelli, E., Guainazzi, M., & Matt, G. 2008, *MNRAS*, **386**, 105
Binney, J., & Tremaine, S. 2008, *Galactic Dynamics* (Second Edition; Princeton, NJ: Princeton Univ. Press)
Bode, J. N., & Wegg, C. 2014, *MNRAS*, **438**, 573
Boldrini, P., Mohayaee, R., & Silk, J. 2020, *MNRAS*, **495**, L12
Bonetti, M., & Sesana, A. 2020, *PhRvD*, **102**, 103023
Callegari, S., Mayer, L., Kazantzidis, S., et al. 2009, *ApJL*, **696**, L89
Caproni, A., Abraham, Z., Motter, J. C., & Monteiro, H. 2017, *ApJL*, **851**, L39
Charisi, M., Bartos, I., Haiman, Z., et al. 2016, *MNRAS*, **463**, 2145
Chen, X., Madau, P., Sesana, A., & Liu, F. K. 2009, *ApJL*, **697**, L149
Chen, X., Sesana, A., Madau, P., & Liu, F. K. 2011, *ApJ*, **729**, 13
Chen, Y.-C., Liu, X., Liao, W.-T., et al. 2020, *MNRAS*, **499**, 2245
Chu, A., Boldrini, P., & Silk, J. 2023, *MNRAS*, **522**, 948
Chua, A. J. K., Katz, M. L., Warburton, N., & Hughes, S. A. 2021, *PhRvL*, **126**, 051102
Chua, A. J. K., Moore, C. J., & Gair, J. R. 2017, *PhRvD*, **96**, 044005
Comerford, J. M., Griffith, R. L., Gerke, B. F., et al. 2009, *ApJL*, **702**, L82
D'Eugenio, C., Daddi, E., Gobat, R., et al. 2020, *ApJL*, **892**, L2
Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Natur*, **433**, 604
Dvorkin, I., & Barausse, E. 2017, *MNRAS*, **470**, 4547
Enoki, M., & Nagashima, M. 2007, *PThPh*, **117**, 241
Ford, E. B., Kozinsky, B., & Rasio, F. A. 2000, *ApJ*, **535**, 385
Fragione, G., Loeb, A., Kremer, K., & Rasio, F. A. 2020, *ApJ*, **897**, 46
Fragione, G., & Sari, R. 2018, *ApJ*, **852**, 51
French, K. D. 2021, *PASP*, **133**, 072001
French, K. D., Wevers, T., Law-Smith, J., Graur, O., & Zabludoff, A. I. 2020, *SSRv*, **216**, 32
Generozov, A., & Madigan, A.-M. 2020, *ApJ*, **896**, 137
Gondán, L., Kocsis, B., Raffai, P., & Frei, Z. 2018, *ApJ*, **855**, 34
Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, *MNRAS*, **453**, 1562
GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2020, *A&A*, **636**, L5
Gravity Collaboration, Straub, O., Bauböck, M., et al. 2023, *A&A*, **672**, A63
Green, P. J., Myers, A. D., Barkhouse, W. A., et al. 2010, *ApJ*, **710**, 1578
Greene, J. E., Strader, J., & Ho, L. C. 2020, *ARA&A*, **58**, 257
Gualandris, A., & Merritt, D. 2009, *ApJ*, **705**, 361
Gürkan, M. A., & Rasio, F. A. 2005, *ApJ*, **628**, 236
Haiman, Z., Kocsis, B., & Menou, K. 2009, *ApJ*, **700**, 1952
Hansen, B. M. S., & Milosavljević, M. 2003, *ApJL*, **593**, L77
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, **163**, 1
Hopman, C., & Alexander, T. 2005, *ApJ*, **629**, 362
Hopman, C., & Alexander, T. 2006, *ApJL*, **645**, L133
Ivanova, N., Chaichenets, S., Fregeau, J., et al. 2010, *ApJ*, **717**, 948
Johannsen, T. 2013, *PhRvD*, **87**, 124017
Katz, M. L., Chua, A. J. K., Speri, L., Warburton, N., & Hughes, S. A. 2021, *PhRvD*, **104**, 064047
Kazantzidis, S., Mayer, L., Colpi, M., et al. 2005, *ApJL*, **623**, L67
Kelley, L. Z., Blecha, L., & Hernquist, L. 2017, *MNRAS*, **464**, 3131
Kocsis, B., Ray, A., & Portegies Zwart, S. 2012, *ApJ*, **752**, 67
Kocsis, B., & Sesana, A. 2011, *MNRAS*, **411**, 1467
Komossa, S., Zhou, H., & Lu, H. 2008, *ApJL*, **678**, L81
Kozai, Y. 1962, *AJ*, **67**, 591
Kroupa, P. 2001, *MNRAS*, **322**, 231
Lacroix, T., & Silk, J. 2018, *ApJL*, **853**, L16
Li, G., Naoz, S., Kocsis, B., & Loeb, A. 2015, *MNRAS*, **451**, 1341
Li, K., Bogdanović, T., & Ballantyne, D. R. 2020, *ApJ*, **896**, 113
Li, S., Berczik, P., & Chen, X. 2019, *ApJ*, **883**, 132
Lidov, M. L. 1962, *P&SS*, **9**, 719
Lithwick, Y., & Naoz, S. 2011, *ApJ*, **742**, 94
Liu, T., Gezari, S., Ayers, M., et al. 2019, *ApJ*, **884**, 36
Madau, P., & Rees, M. J. 2001, *ApJL*, **551**, L27
Maillard, J. P., Paumard, T., Stolovy, S. R., & Rigaut, F. 2004, *A&A*, **423**, 155
Mazzolari, G., Bonetti, M., Sesana, A., et al. 2022, *MNRAS*, **516**, 1959
Mei, Jianwei, Bai, Y.-Z., Bao, J., et al. 2021, *PTEP*, **2021**, 05A107
Melchor, D., Mockler, B., Naoz, S., Rose, S., & Ramirez-Ruiz, E. 2023, arXiv:2306.05472
Merloni, A., & Heinz, S. 2008, *MNRAS*, **388**, 1011
Merritt, D. 2010, arXiv:1001.3706
Mičić, M., Holmes, O. J., Wells, B. N., & Irwin, J. A. 2023, *ApJ*, **944**, 160
Milosavljević, M., & Merritt, D. 2001, *ApJ*, **563**, 34
Mockler, B., Melchor, D., Naoz, S., & Ramirez-Ruiz, E. 2023, arXiv:2306.05510
Naoz, S. 2016, *ARA&A*, **54**, 441
Naoz, S., & Fabrycky, D. C. 2014, *ApJ*, **793**, 137
Naoz, S., Kocsis, B., Loeb, A., & Yunes, N. 2013, *ApJ*, **773**, 187
Naoz, S., Rose, S. C., Michaely, E., et al. 2022, *ApJL*, **927**, L18
Naoz, S., & Silk, J. 2014, *ApJ*, **795**, 102
Naoz, S., Silk, J., & Schnittman, J. D. 2019, *ApJL*, **885**, L35
Naoz, S., Will, C. M., Ramirez-Ruiz, E., et al. 2020, *ApJL*, **888**, L8
Nissanke, S., Vallisneri, M., Nelemans, G., & Prince, T. A. 2012, *ApJ*, **758**, 131
O'Leary, R. M., Kocsis, B., & Loeb, A. 2009, *MNRAS*, **395**, 2127
O'Neill, S., Kiehlmann, S., Readhead, A. C. S., et al. 2022, *ApJL*, **926**, L35
Paudel, S., Smith, R., Yoon, S. J., Calderón-Castillo, P., & Duc, P.-A. 2018, *ApJS*, **237**, 36
Peters, P. C. 1964, *PhRv*, **136**, 1224
Peters, P. C., & Mathews, J. 1963, *PhRv*, **131**, 435
Phinney, E. S. 2001, arXiv:astro-ph/0108028
Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, **473**, 4077
Pozzoli, F., Babak, S., Sesana, A., Bonetti, M., & Karnesis, N. 2023, arXiv:2302.07043
Rashkov, V., & Madau, P. 2014, *ApJ*, **780**, 187
Ren, G.-W., Ding, N., Zhang, X., et al. 2021, *MNRAS*, **506**, 3791

- Ricarte, A., Natarajan, P., Dai, L., & Coppi, P. 2016, [MNRAS](#), **458**, 1712
- Robertson, B., Bullock, J. S., Cox, T. J., et al. 2006, [ApJ](#), **645**, 986
- Robson, T., Cornish, N. J., & Liu, C. 2019, [CQGra](#), **36**, 105011
- Rodriguez, C., Taylor, G. B., Zavala, R. T., et al. 2006, [ApJ](#), **646**, 49
- Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, [MNRAS](#), **449**, 49
- Rose, S. C., Naoz, S., Sari, R., & Linial, I. 2022, [ApJL](#), **929**, L22
- Sari, R., & Fragione, G. 2019, [ApJ](#), **885**, 24
- Schnittman, J. D. 2015, [ApJ](#), **806**, 264
- Schnittman, J. D., Dal Canton, T., Camp, J., Tsang, D., & Kelly, B. J. 2018, [ApJ](#), **853**, 123
- Sesana, A., Haardt, F., & Madau, P. 2006, [ApJ](#), **651**, 392
- Seto, N. 2012, [PhRvD](#), **85**, 064037
- Sicilia, A., Lapi, A., Boco, L., et al. 2022, [ApJ](#), **934**, 66
- Sillanpaa, A., Haarala, S., Valtonen, M. J., Sundelius, B., & Byrd, G. G. 1988, [ApJ](#), **325**, 628
- Smith, K. L., Shields, G. A., Bonning, E. W., et al. 2010, [ApJ](#), **716**, 866
- Stemo, A., Comerford, J. M., Barrows, R. S., et al. 2021, [ApJ](#), **923**, 36
- Toonen, S., Hopman, C., & Freitag, M. 2009, [MNRAS](#), **398**, 1228
- Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, [ApJ](#), **574**, 740
- van Velzen, S., Holoien, T. W. S., Onori, F., Hung, T., & Arcavi, I. 2020, [SSRv](#), **216**, 124
- Volonteri, M., Haardt, F., & Madau, P. 2003a, [ApJ](#), **582**, 559
- Volonteri, M., Madau, P., & Haardt, F. 2003b, [ApJ](#), **593**, 661
- Wegg, C., & Bode, J. N. 2011, [ApJL](#), **738**, L8
- Wild, V., Almaini, O., Dunlop, J., et al. 2016, [MNRAS](#), **463**, 832
- Xuan, Z., Naoz, S., & Chen, X. 2023, [PhRvD](#), **107**, 043009
- Yu, Q. 2002, [MNRAS](#), **331**, 935
- Yunes, N., Sopuerta, C. F., Rubbo, L. J., & Holley-Bockelmann, K. 2008, [ApJ](#), **675**, 604
- Zhang, E., Naoz, S., & Will, C. M. 2023, [ApJ](#), **952**, 103
- Zheng, X., Lin, D. N. C., & Mao, S. 2020, [ApJ](#), **905**, 169